

Understanding and Exploiting Fungal PKS-NRPS
Biosynthesis in *Pyricularia* and
Magnaporthe species

Von der Naturwissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover

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

genehmigte Dissertation

von

Verena Hantke, geb. Belt, M. Sc.

2019

Referent: Prof. Dr Russell Cox

Korreferent: Prof. Dr Andreas Kirschning

Tag der Promotion: 26.11.2019

Abstract

Keywords: natural products, biosynthesis, PKS-NRPS, cytochalasans

The fungus *Pyricularia oryzae* is a major pathogen of rice (*Oryza sativa*) that causes substantial loss of yields every year. Remarkably, certain rice strains carrying the *Pi33* gene are resistant to strains of *P. oryzae* encoding a PKS-NRPS called the avirulence conferring enzyme1 (ACE1).

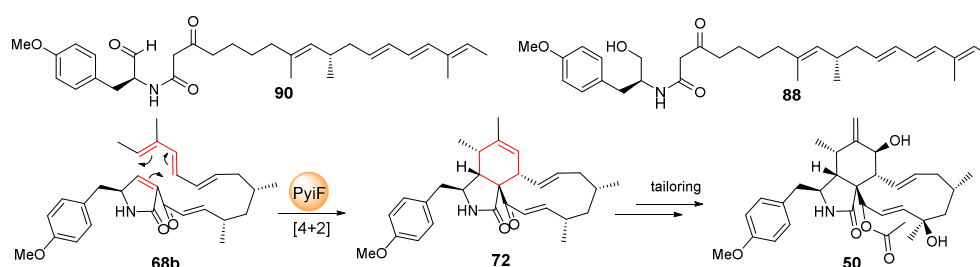
By heterologous expression studies in *Aspergillus oryzae*, more insights into the early biosynthetic steps of this ACE1 metabolite were obtained. *O*-Methyl-L-tyrosine **71** was identified as the preferred substrate for the adenylation domain of the ACE1 NRPS, indicating that the *OME1* gene (*O*-methyltransferase) belongs to the ACE1 cluster. However, even after all biosynthetic genes of the ACE1 cluster were co-expressed; the produced compound was **88** - a shunt intermediate. *In-vitro* enzyme assays confirmed that reduction of the aldehyde **90** to the alcohol **88** is catalysed by a native *A. oryzae* enzyme at an early biosynthetic step; thus stalling the biosynthesis.

Ectopic expression of *BC1* (transcription factor from the ACE1 BGC) in *P. oryzae* did not lead to the production of ACE1-related compounds. However, the compound hinnulin A **105** was formed, which belongs to the class of DHN-melanins. A putative partial BGC potentially involved in the biosynthesis of **105** was validated by RT-PCR and a possible biosynthetic pathway was proposed.

Another cytochalasan BGC was previously demonstrated to be responsible for the production of pyrichalasin H **50** in *Magnaporthe grisea* and now revealed to be homologous to the ACE1 BGC. Thus, the **50** BGC was used as model system to elucidate the function of two potential key enzymes in cytochalasan biosynthesis: an $\alpha\beta$ -hydrolase (HYD) and a putative Diels-Alderase (pDA).

Targeted gene deletion experiments in *M. grisea* revealed that the HYD PylE is involved in early steps during the biosynthesis of **50**, but its exact role remained elusive. The pDA PylF was shown to be involved in the intramolecular [4+2] cycloaddition forming **72**. Complementation studies with *ORF3* from the ACE1 BGC indicated a similar catalytic function of both enzymes.

By recombinant gene expression studies in *E. coli* soluble ORFZ (HYD) and ORF3 (pDA) protein was obtained. *In-vitro* assays are underway to gain further insights into their biosynthetic role.



Zusammenfassung

Schlagwörter: Naturstoffe, Biosynthese, PKS-NRPS, Cytochalasane

Der Pilz *Pyricularia oryzae* ist ein Hauptpathogen der Reispflanze (*Oryza sativa*) und führt zu einer erheblichen jährlichen Ertragssenkung. Bestimmte Reisstämme tragen das Pi33-Gen, wodurch sie resistent gegen jene *P. oryzae* Stämme sind, die das Gen einer bestimmten PKS-NRPS tragen: dem „Avirulenz verleihenden Enzym1 (ACE1)“.

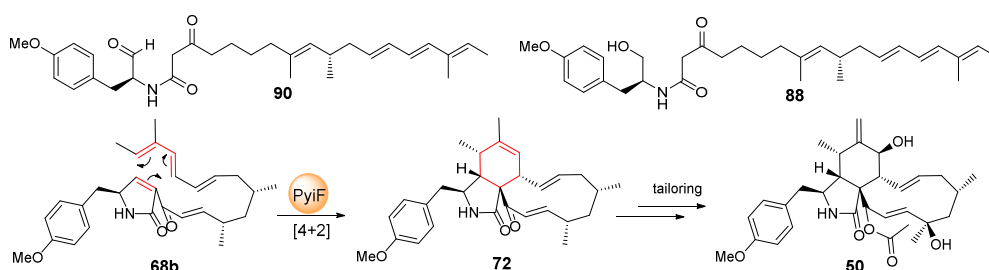
Durch heterologe Expression des ACE1 Clusters in *Aspergillus oryzae* konnten die frühen Biosyntheseschritte charakterisiert werden. *O*-Methyl-L-Tyrosin **71** wurde als bevorzugtes Substrat für die NRPS Adenylierungs-Domäne identifiziert und daher OME1 (*O*-Methyltransferase) dem ACE1 Cluster zugeordnet. Trotz Co-Expression aller ACE1 Biosynthesegene wurde nur **88** als Hauptprodukt nachgewiesen, ein Shunt-Intermediat. *In-vitro* Tests bestätigten, dass die Reduktion des Aldehyds **90** zum Alkohol **88** durch ein natives *A. oryzae* Enzym katalysiert wird.

Die Expression von BC1 (Transkriptionsfaktor aus dem ACE1 Cluster) in *P. oryzae* führte nicht zur Produktion von ACE1-verwandten Verbindungen. Stattdessen wurde jedoch die Verbindung Hinnulin A **105** gebildet. Ein partielles Gencluster, welches womöglich an der Biosynthese von **105** beteiligt ist, wurde durch RT-PCR validiert und ein möglicher Biosyntheseweg vorgeschlagen.

In *Magnaporthe grisea* wurde kürzlich ein weiteres Cytochalasan Gencluster identifiziert, welches für die Synthese von Pyrichalasin H **50** verantwortlich ist. Aufgrund der hohen Homologie mit dem ACE1 Cluster wurde dieses Pyrichalasin Gencluster als Modellsystem verwendet, um die Funktion von zwei Enzymen aufzuklären, die wahrscheinlich eine Schlüsselrolle bei der Biosynthese von Cytochalasanen spielen: eine $\alpha\beta$ -Hydrolase (HYD) und eine mutmaßliche Diels-Alderase (mDA).

Gezielte Deletionsexperimente in *M. grisea* zeigten, dass die HYD PyiE an frühen Schritten der Biosynthese von **50** beteiligt ist. Es wurde gezeigt, dass die mDA PyiF an der [4+2]-Cycloaddition beteiligt ist, wodurch **72** gebildet wird. Komplementationstudien mit ORF3 aus dem ACE1 Gencluster zeigten eine ähnliche katalytische Funktion beider Enzyme.

In *E. coli* konnte lösliches ORFZ (HYD) und ORF3 (mDA) produziert werden, wodurch die Biosynthese des ACE1 Metaboliten in Zukunft durch *in-vitro*-Tests weiter aufgeklärt werden kann.



Acknowledgment

First of all, I would like to thank Prof Russell Cox for always being a great supervisor during the last 3.5 years. I am incredibly thankful for all the support and a very interesting PhD project. Your door was constantly open and somehow you always found the time to discuss my research project, which was very helpful.

I would like to thank Prof Andreas Kirschning for taking the time to be the co-referee of my PhD thesis and the chair of my examination board. No less, I thank Prof Kürşad Turgay for being an examiner. My scientific career started in your lab during my Bachelor thesis.

Many thanks to the BMWZ media kitchen team, especially Katja Körner, and all the analytical departments at the OCI. Without your help this work would not have been possible. I also would like to thank the AK Braun, especially Dr Jennifer Senkler, for all MS analyses.

I would also like to thank Prof Marc-Henri Lebrun (UMR BIOGER INRA AgroParisTech) and Dr Didier Tharreau (CIRAD UMR BGPI) for the *P. oryzae* Guy11 and *M. grisea* NI980 strains and fruitful discussions. In addition, I would like to thank our cooperation partners at the CeBiTEC Bielefeld, especially Dr Daniel Wibberg and Prof Jörn Kalinowski, for all their effort with sequencing and annotation.

Beyond that, I would like to thank all past and current Cox group members, in particular Carsten, Chongqing, Lukas, Karen, Raissa and all my students. Especially, I want to thank Liz for her support during the last years and Haili for providing me with chemically synthesized substrates.

Further, I thank Gesche, Anna, Janna and Janina for their support and a really great time during my PhD.

Finally, I would like to thank my family and friends for their continuous support throughout my studies. Especially, I would like to thank my husband Ingo, who always covered my back during the last years and who always believed in me. Thank you.

Abbreviations

A	adenine	KS	β -ketosynthase
A	adenylation	KO	knockout
Ac	acetate	L	litre
ACE1	avirulence conferring enzyme 1	LC	liquid chromatography
ACP	acyl carrier protein	LUMO	lowest occupied molecular orbital
APS	ammonium persulfate	M	molar
AT	acyltransferase	mg	milligram
AA	amino acid	min	minute
BGC	biosynthetic gene cluster	mL	millilitre
BLAST	basic local alignment search tool	MO	molecular orbital
BLASTp/n	BLASTprotein/nucleotide	mRNA	messenger RNA
bp	base pair	MS	mass spectrometry
BSA	bovine serum albumin	MW	molecular weight
C	condensation	<i>m/z</i>	mass to charge ratio
C	cytosine	NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
cDNA	complementary DNA	NEDDA	Normal electron demand Diels-Alder reaction
Cm	chloramphenicol	NHEJ	non-homologous end-joining
CD	conserved domain	NMR	nuclear magnetic resonance
CFE	cell free extract	NR	non-reducing
C-MeT	C-methyltransferase	NRPS	non-ribosomal peptide synthetase
CoA	Co-enzyme A	nt	nucleotide(s)
COSY	correlation spectroscopy	NTA	nitrilotriacetic acid
Cyc	cyclase	OD ₆₀₀	optical density at 600 nm
Da	Dalton	ORF	open reading frame
DA(ase)	Diels-Alder(ase)	PCP	peptide carrier protein
DAD	diode array detector	PCR	polymerase chain reaction
ddH ₂ O	double distilled H ₂ O	PEG	polyethylene glycol
DH	dehydratase	PKS	polyketide synthase
DHN	1,8-dihydroxynaphthalene	PP	phosphopantetheine
DNA	deoxyribonucleic acid	PR	partially reducing
dNTP	deoxynucleotide triphosphate	PT	product template
E	epimerase	Q-TOF	quadrupole time-of-flight
EDG	electron donating group	R	reductive release domain
EDTA	ethylenediaminetetraacetic acid	R*	reductive-like release domain
eGFP	enhanced green fluorescent protein	rpm	rotations per minute
ER	enoyl reductase	RT	room temperature
ELSD	evaporative light scattering detector	RT-PCR	reverse transcriptase PCR
ESI	electro-spray ionisation	SAM	S-adenosyl methionine
EWG	electron withdrawing group	SAT	starter unit acyl carrier protein
FAS	fatty acid synthase	SDR	short chain dehydratase/reductase
FPLC	fast protein liquid chromatography	SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
FMO	Frontier Molecule Orbital	SEC	size exclusion chromatography
G	guanine	SN	supernatant
gDNA	genomic DNA	SNAC	N-acetylcysteamine thioesters
GOI	gene of interest	SM	secondary metabolite(s)
h	hour(s)	<i>sp.</i>	species
HOMO	highest occupied molecular orbital	T	thiolation
HPLC	high performance liquid chromatography	T	thymine
HR	highly reducing	TE	thiolesterase
HRMS	high resolution mass spectrometry	TEMED	N,N,N',N'-tetramethylethane-1,2-diamine
HSQC	heteronuclear single quantum correlation	TF	transcription factor
<i>hyg</i> ^R	hygromycin B resistance cassette	Tris	tris(hydroxymethyl)aminomethane
Hz	Hertz	TMS	trimethylsilyl
IEDDA	inverse electron demand Diels-Alder	U	uracil
IMAC	immobilized metal ion affinity chromatography	UV	ultra violet
IPTG	isopropyl β -D-1 thiogalactopyranoside	v/v	volume per volume
Kan	kanamycin	WT	wild type
kb	kilobases	w/v	weight per volume
kDa	kilodalton		
KR	β -ketoreductase		

Table of Contents

Abstract	I
Zusammenfassung	II
Acknowledgment	III
Abbreviations	IV
Table of Contents	V
1. Introduction	1
1.1 Fungal Secondary Metabolites	1
1.2 Biosynthetic Gene Clusters in Filamentous Fungi	2
1.3 Biosynthesis of Polyketides and Fatty Acids	4
1.4 Non-ribosomal Peptide (NRP) Biosynthesis and NRP Synthetases	9
1.5 Polyketide Synthase - Non-ribosomal Peptide Synthetase (PKS-NRPS) Hybrids	10
1.6 Cytochalasans	15
1.7 The ACE1 Metabolite from <i>Pyricularia oryzae</i>	19
1.7.1 Introduction	19
1.7.2 ACE1 Gene Cluster and Proposed Biosynthesis	21
1.7.3 Melanization and Appressorium Formation	25
1.8 The Pyrichalasin H Biosynthetic Gene Cluster	26
1.9 Biomolecular Methods Used in this Study	29
1.9.1 Targeted Gene Knockout in Filamentous Fungi	29
1.9.2 Complementation Studies in <i>M. grisea</i> Gene Deletion Strains	31
1.10 Aim of the Thesis	32
2. Bioinformatic Comparison of the ACE1 and the Putative Pyrichalasin H BGC	34
2.1 Aim of the Studies	34
2.2 Results and Discussion	34
3. Heterologous Expression of the <i>P. oryzae</i> ACE1 BGC in <i>A. oryzae</i>	37
3.1 Vector Construction by Homologous Recombination in <i>S. cerevisiae</i>	37
3.2 Aim of the Heterologous Expression Studies	39
3.3 Results	41
3.3.1 Early Steps of ACE1 Biosynthesis: Expression of the ACE1 Core Genes	41
3.3.2 Co-Expression of ACE1, RAP1, ORF3 and ORFZ	43

3.3.3	Feeding Studies with <i>O</i> -Methyl-L-Tyrosine	45
3.3.4	Simultaneous Expression of 11 Genes of the <i>ACE1</i> BGC.....	49
3.4	Discussion.....	51
4.	<i>In-vitro</i> Screening for an Oxidoreductase from <i>A. oryzae</i>	56
4.1	Results	56
4.2	Protein Identification by ESI Q TOF.....	61
4.3	Discussion and Outlook.....	62
5.	Overexpression of the Transcription Factor <i>BC2</i> from the <i>ACE1</i> BGC in <i>P. oryzae</i>	64
5.1	The Transcription Factor <i>BC2</i>	64
5.2	Activation Strategies for Silent Fungal BGC	64
5.3	Aim of the Studies	66
5.4	Results	67
5.4.1	Transformation of <i>P. oryzae</i> Guy11 with pTH-GS- <i>egfp</i>	67
5.4.2	Homologous overexpression of <i>BC2</i> in <i>P. oryzae</i> Guy11	67
5.4.3	Reverse Transcriptase (RT) Analysis of Transformant VBI27-5.....	72
5.5	Discussion.....	76
6.	Investigations of the $\alpha\beta$-Hydrolases ORFZ and <i>PyiE</i>.....	80
6.1	Introduction	80
6.1.1	The Superfamily of $\alpha\beta$ -Hydrolases.....	81
6.1.2	The Catalytic Triad.....	82
6.2	Aim of the Studies.....	85
6.3	Results	86
6.3.1	Targeted Gene Deletion of <i>pyiE</i>	86
6.3.2	Complementation of <i>M. grisea</i> <i>pyiE</i> Deletion Strains.....	89
6.3.3	<i>In-vitro</i> Studies with the $\alpha\beta$ -Hydrolase ORFZ.....	99
6.4	Discussion.....	104
7.	Investigations of the putative Diels-Alderase.....	109
7.1	Introduction	109
7.1.1	Frontier Molecule Orbital Theory	109
7.1.2	How Does a Diels-Alderase Operate?	111
7.2	Aims of This Study.....	115

7.3	Results	116
7.3.1	Targeted Gene Deletion of <i>pyiF</i>	116
7.3.2	Structure Elucidation of Compound 179 - 186 by NMR.....	121
7.3.3	Complementation Studies with Δ <i>pyiF</i> <i>M. grisea</i> Strains	124
7.3.4	Heterologous Complementation with Homologous Genes of <i>pyiF</i>	128
7.3.5	<i>In-vitro</i> Studies with ORF3.....	131
7.4	Discussion.....	135
8.	Overall Conclusion and Outlook	140
9.	Experimental.....	147
9.1	Growth Media, Buffers and Solutions.....	147
9.2	Enzymes and Antibiotics	149
9.3	Microbiology Methods.....	149
9.3.1	<i>Escherichia coli</i> Strains	150
9.3.2	<i>Saccharomyces cerevisiae</i> CEN.PK2.....	151
9.3.3	<i>Aspergillus oryzae</i> NSAR1.....	151
9.3.4	<i>Magnaporthe grisea</i> and <i>Pyricularia oryzae</i>	152
9.4	Molecular Biology Methods	153
9.4.1	Oligonucleotides.....	153
9.4.2	Polymerase Chain Reaction (PCR)	156
9.4.3	Obtained and Constructed Vectors in this Thesis	156
9.4.4	RNA Extraction and cDNA Preparation	157
9.4.5	Agarose Gel Electrophoresis	158
9.4.6	DNA Purification from Gel or PCR	158
9.4.7	Isolation of Plasmid DNA from <i>E. coli</i>	158
9.4.8	Isolation of Plasmid DNA from <i>S. cerevisiae</i>	158
9.4.9	Isolation of Genomic DNA from Fungi.....	158
9.4.10	Cloning Procedure	159
9.4.11	Heterologous Protein Production and Purification.....	159
9.4.12	Protein Analysis by SDS-PAGE, Bradford and ESI Q-TOF	161
9.4.13	Enzyme Activity Assay with ORFZ.....	162
9.4.14	Enzyme Activity Assay with <i>A. oryzae</i> Cell Free Extract	162
9.5	Chemical Analysis.....	163
9.5.1	Extraction of <i>A. oryzae</i> Cultures.....	163

9.5.2	Extraction of <i>M. grisea</i> and <i>P. oryzae</i> Cultures	163
9.5.3	Liquid Chromatography Mass Spectrometry (LCMS) Analysis	163
9.5.4	High Resolution Mass Spectrometry (HRMS).....	164
9.5.5	Nuclear Magnetic Resonance (NMR) Analysis	164
10.	Bibliography	165
11.	Appendix.....	170
11.1	NMR Data Chapter 6	170
11.2	NMR Data Chapter 7	171

1. Introduction

1.1 Fungal Secondary Metabolites

Filamentous fungi produce a wide range of natural products called secondary metabolites (SM), which are not essential for the immediate survival of the organism, but may fulfil important ecological roles. These compounds are usually low-molecular-weight metabolites with diverse chemical structures and biological activities. These activities range from compounds that are rather harmless or even beneficial for humankind up to highly toxic metabolites.¹ Unlike primary metabolites (such as amino acids or sugars) which are essential for the living organism, SM are often exclusively produced by a specific organism or only within its respective phylogenetic group; they do not occur ubiquitously.

Notably, despite immense structural diversity of SM, most SM are constructed of simple building blocks (such as acetyl- and malonyl-CoA, shikimic acid or amino acids) which derive from primary metabolism. Based on their biosynthetic origin, SM are classified in different groups. Polyketides form a large and structurally diverse group. Examples include squalestatin S1 **1** (*Phoma* species), a potent competitive inhibitor of cholesterol biosynthesis,² lovastatin **2** (*Aspergillus terreus*), a cholesterol lowering agent,³ verrucosidin **3** (*Penicillium aurantiogriseum*), a neurotoxin⁴ and the simple tetraketide 6-methylsalicylic acid (6-MSA) **4** (*Penicillium patulum*), which has long been accepted as a model compound for polyketide research (Figure 1.1).⁵

Non-ribosomal peptides like penicillin G **5**, terpenes such as β -bisabolene **6** and alkaloids like ergotamine **7** are also major sub-classes of SM. Furthermore, hybrid compounds exist such as fusarin C **8** (*Fusarium* species)⁶ and tenellin **9** (*Beauveria bassiana*),⁷ where a polyketide is fused to an amino acid (Figure 1.1).

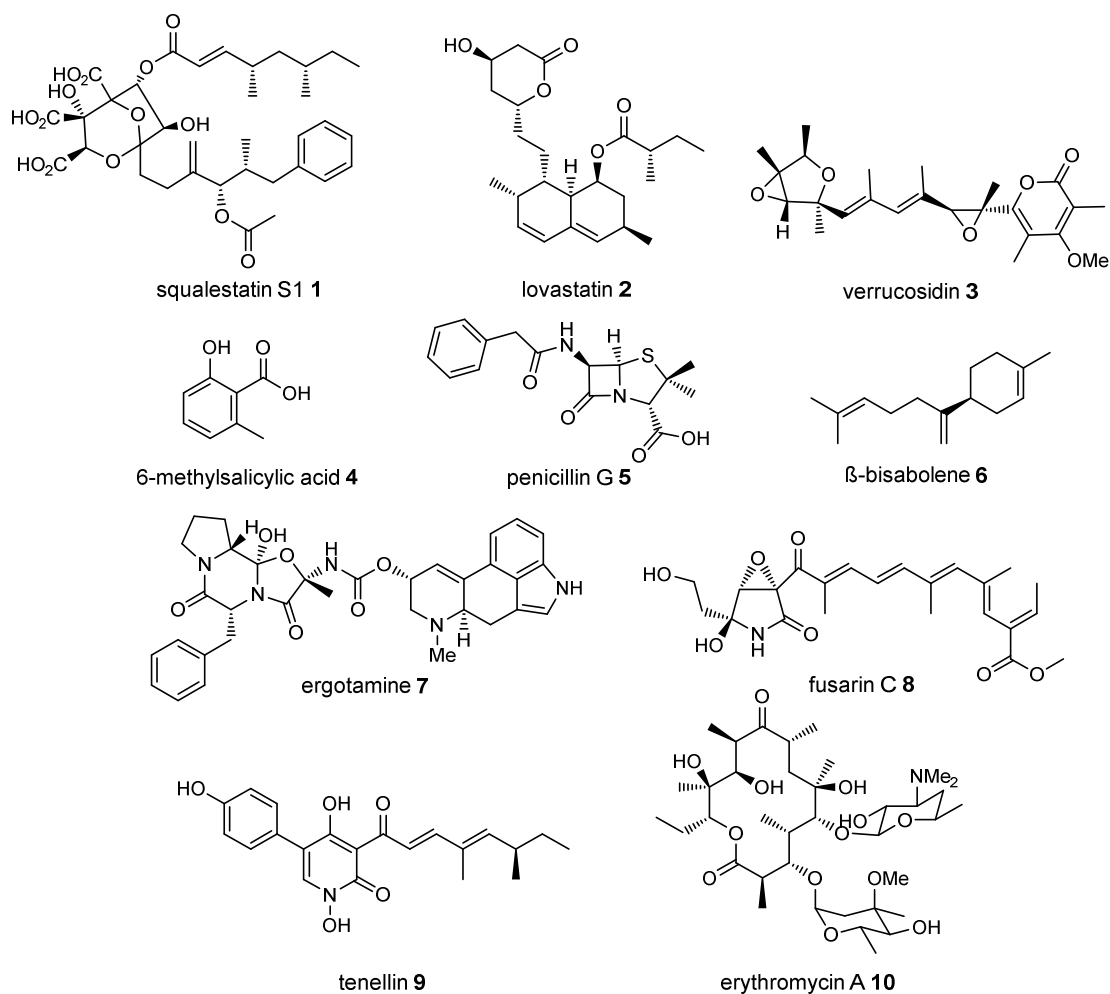


Figure 1.1: Chemical structures of classic fungal secondary metabolites.

1.2 Biosynthetic Gene Clusters in Filamentous Fungi

In fungi, genes involved in the biosynthesis of a given secondary metabolite are usually clustered within the genome. The complete set of biosynthetic genes forms the biosynthetic gene cluster (BGC).⁸ There are rare exceptions to this rule: in *Aspergillus nidulans* the genes required for the biosynthesis of the meroterpenoids austinol and dehydroaustinol are located in two separate clusters (located on different chromosomes).⁹

Genome sequencing, which has increased rapidly in the last years, coupled with computational methods for data processing, allows the rapid identification of BGC. *In silico* analysis of fungal genomes to search for BGC is often referred to as genome mining.^{10,11} For example, sequencing and analysis of the genome of the filamentous fungus *Aspergillus nidulans* revealed the presence of 56 putative BGC that could encode proteins involved in SM biosynthetic pathways.¹² Since the identity of most of the SM produced by the identified BGC remains unknown, the majority of clusters are referred to as cryptic. Various experiments can be performed to link an identified BGC

conformation. Subsequently, the chromatin structure has an impact on the expression pattern by allowing or restricting TF to access the promoter regions. Thus, modifications of histones are able to affect several biological processes, which can be epigenetically inherited. Epigenetic information is not coded by DNA but can still be transmitted from one generation to the next generation (*e.g.* on chromatin).^{17,18}

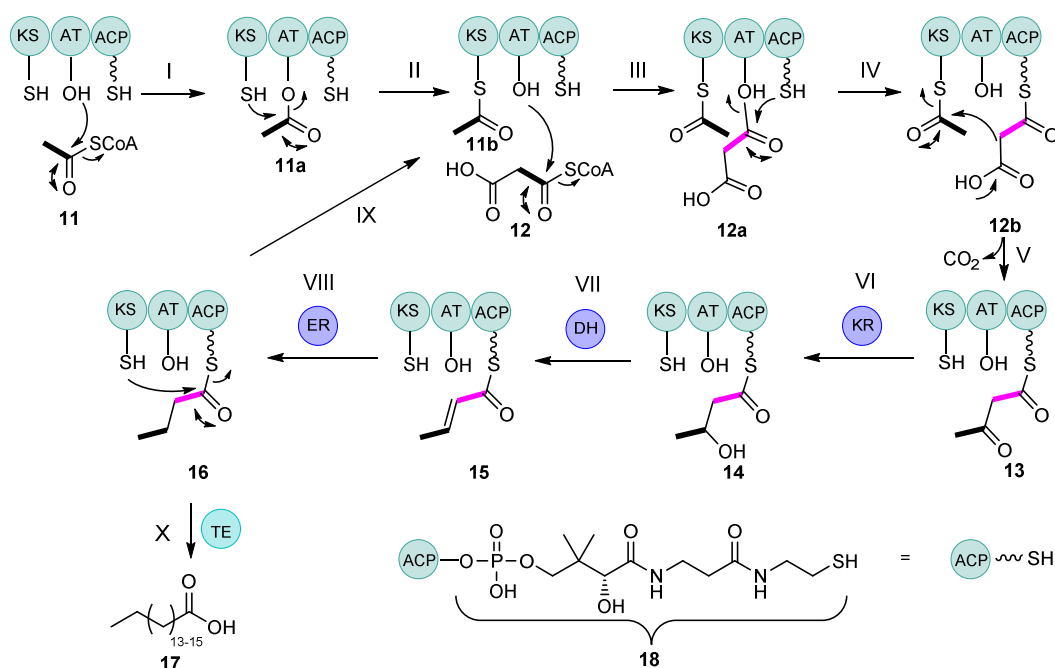
Collectively, the transcriptional activation of a BGC may be dependent on several different parameters, including transcriptional regulators, environmental cues and chromatin structure. In a bigger picture the production of SM is controlled by a complex regulatory network which consists of various interconnecting sub-networks containing multiple proteins that respond to several environmental stimuli. This specific regulatory network can be elucidated as a strategy to characterize the biological role of the compound and to identify the optimal conditions for the production of a metabolite to elucidate its structure.

1.3 Biosynthesis of Polyketides and Fatty Acids

Polyketides are a large and structurally diverse class of secondary metabolites produced by fungi, plants and bacteria. Small aromatic compounds such as 6-methylsalicylic acid **4**, as well as highly complex macromolecules such as erythromycin A **10**, are part of this group (Figure 1.1). Despite the diversity of these compounds, polyketides are defined by their biosynthesis from simple precursors such as acetate and propionate. In the 1950s Birch and co-workers demonstrated by ¹⁴C-feeding experiments that the aromatic ring of **4** is formed by head-to-tail condensation of acetate units. This observation led to the conclusion that polyketide biosynthesis is closely related to the biosynthesis of fatty acids.^{19,20} Since fatty acid biosynthesis is well understood, it is convenient to begin by describing this pathway.

Fatty acids are synthesized from acetyl-CoA **11** and malonyl-CoA **12** precursors through the catalytic activity of enzymes called fatty acid synthases (FAS). The acetyl-CoA starter unit **11** is loaded by an acyltransferase (AT) to a highly conserved cysteine thiol of the ketosynthase (KS) domain (Scheme 1.1, I + II). Next, the malonyl-CoA extender unit **12** is loaded by the AT onto the flexible phosphopantetheine (PP) moiety **18** of the acyl carrier protein (ACP, Scheme 1.1, III). The KS together with the ACP catalyses chain elongation by a decarboxylative Claisen condensation, resulting in a C₂ chain extension to form **13** (Scheme 1.1, IV + V). The produced β-keto thiolester is located on the ACP PP for β-processing (Scheme 1.1, IV-V). The flexible PP moiety functions as a long arm carrying the growing β-keto ester chain, enabling it to reach all catalytic sites of the FAS.

In subsequent steps known as β -processing, the β -keto thioester is reduced to a secondary alcohol **14** by the NADPH-dependent ketoreductase (KR) domain (Scheme 1.1, VI), dehydrated by the dehydratase (DH) domain to form an unsaturated thioester **15** (Scheme 1.1, VII) and finally reduced by the NADPH-dependent enoyl reductase (ER) domain to form a fully saturated thioester **16** (Scheme 1.1, VIII). Intermediate **16** is then transferred back to the KS domain for another cycle of chain extension and β -processing (Scheme 1.1, IX). With each cycle the chain grows in length by two methylene units until a chain length between 16 to 18 carbons is reached and the fatty acid is released from ACP PP by hydrolysis by a thioesterase (TE) domain (Scheme 1.1, X).²¹



Scheme 1.1: Overview of the biosynthesis of fatty acids.²¹

All known fatty acids are synthesized in an iterative process, in which the same catalysts are re-used: after elongation and modification, the chain is reloaded onto the KS. The cycles of extension and β -processing continue until a thioesterase (TE) releases chains of defined length.

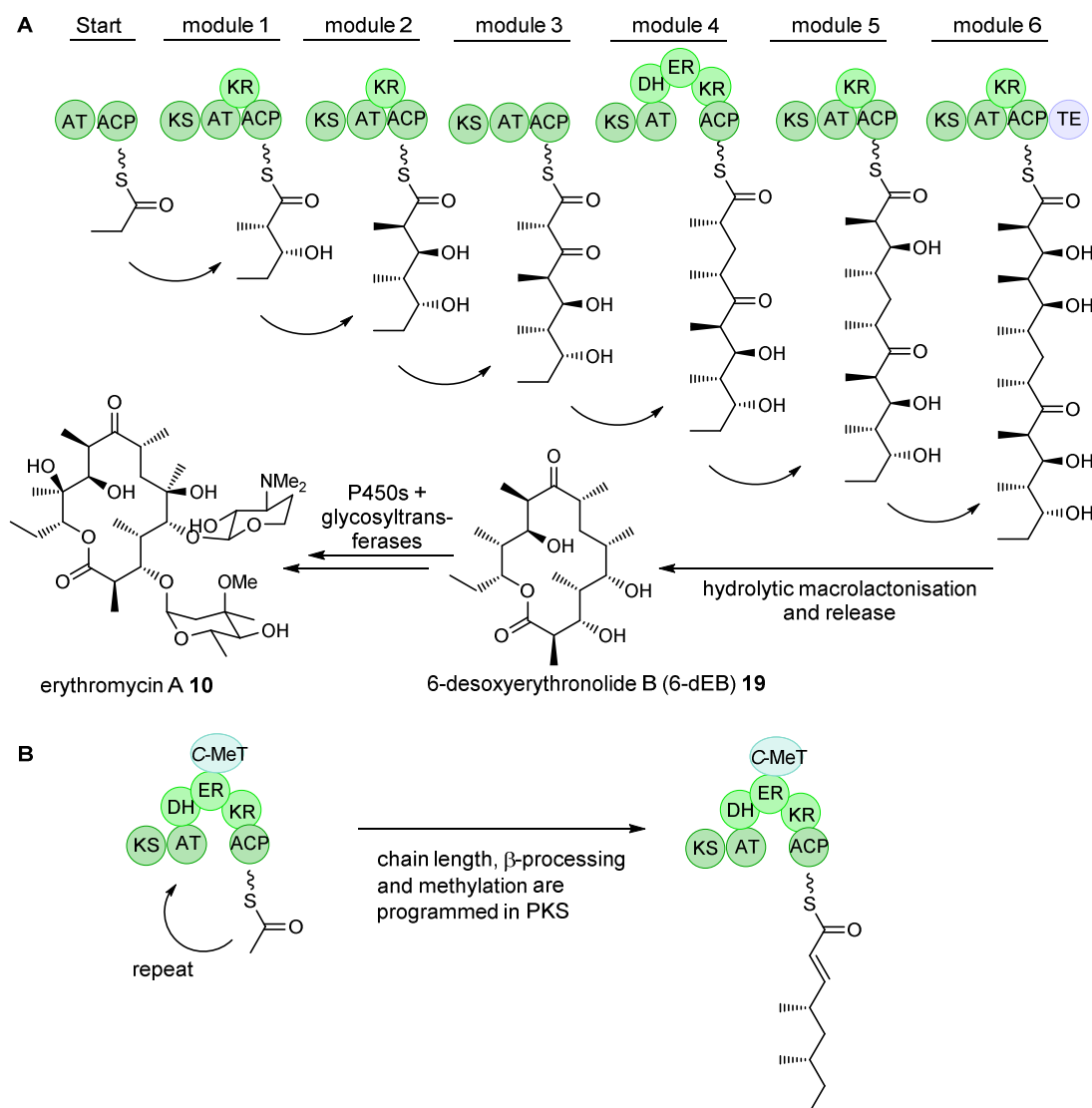
PKS differ from FAS in their chain length and their ability to use a broader range of biosynthetic starter and extender units. The β -processing steps are also optional, resulting in a remarkable level of complex and diverse natural products. In fungi β -processing can also include methylation at the α -carbon.²² Post assembly modifications, like methylations and post-PKS oxidations, further increase the structural diversity and functionality of polyketides.

Based on their domain structures FAS and PKS are classified into three main groups: type I, II and III (III only found in PKS). Type I systems are large multifunctional proteins in which the domains, responsible for chain elongation and β -processing, are covalently linked. In type II systems the catalytic domains are separate enzymes which form multiprotein complexes. Type III PKS only consist of a single KS domain and lack the other domains present in type I and II PKS and FAS.²³

Apart from their structural architecture, PKS can be further categorized as modular (type I) or iterative (type I, II, III), depending on whether each KS catalyses one or more rounds of elongation. In modular type I PKS, found in bacteria, each module carries an individual subset of domains responsible for a single round of chain extension and β -processing. Multiple modules are present, one for each step of the biosynthesis, which allows a good prediction of the structure of the polyketide based on the knowledge of its genetic modular architecture. A minimal module consists of a KS, an AT and an ACP domain, which are needed for chain elongation. Domains for β -processing (KR, DH, ER) are optional. At the end of the last module, a TE domain usually catalyses the release of the product from the PKS. One of the most well-known examples of a modular PKS is encoded by the DEBS BGC responsible for the biosynthesis of erythromycin A **10**, a macrolide antibiotic (Scheme 1.2 A).²⁴

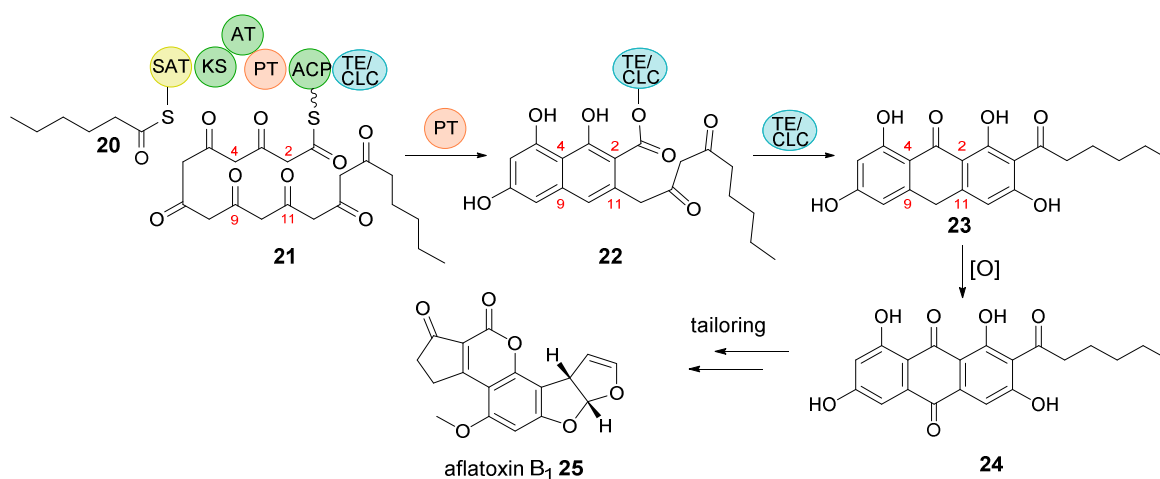
All known PKS found in fungi are iterative type I synthases, consisting of a single module which is used repetitively until the final chain length is reached. After each round the growing chain is reloaded onto the KS. Unlike during the biosynthesis of fatty acids, β -processing is variable after each elongation step, a phenomenon called programming.^{25,26} An example of an iterative PKS is the tetraketide synthase (*mfpks1*) from the squalestatin S1 **1** BGC (Scheme 1.2 B).²⁷

Depending on the variation of the optional processing domains (KR, DH, ER), fungal iterative type I PKS can be further categorised into three classes: non-reducing PKS (NR-PKS), partially reducing PKS (PR-PKS) and highly-reducing PKS (HR-PKS).



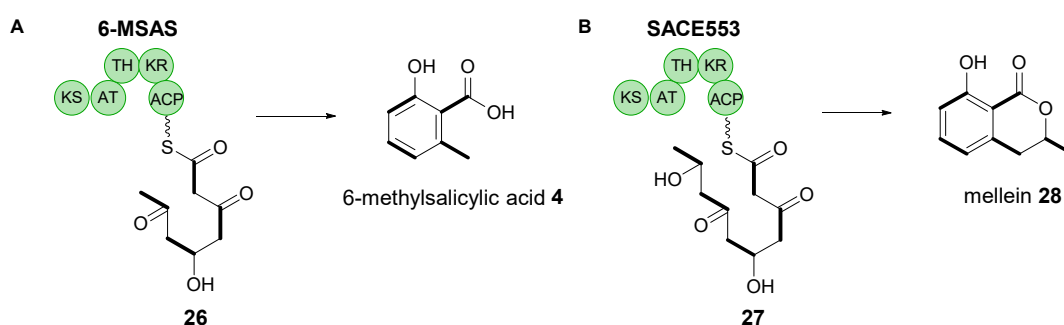
Scheme 1.2: Type I PKS: **A**, Modular PKS; **B**, Iterative PKS.

NR-PKS produce aromatic compounds such as norsolorinic acid **24** which is a precursor of the mycotoxin aflatoxin **25**.²⁸ In addition to the basic domains (AT, KS, ACP) NR-PKS usually consist of a product template (PT) domain and a starter unit acyltransferase (SAT) domain.^{29,30} All of the β -processing domains (KR, DH, ER) are missing, therefore NR-PKS are only able to produce poly- β -keto intermediates. PT domains are involved in regio selective aldol cyclizations of poly- β -keto intermediates during polyketide biosynthesis.^{31,32} SAT domains are able to transfer more complex starter units to the ACP such as hexanoic acid **20** during the biosynthesis of **24** (Scheme 1.3).²⁹



Scheme 1.3: Aflatoxin **25** biosynthesis. The unusual hexanoyl starter unit **20** is selected by the SAT domain; seven rounds of elongation lead to a C₂₀ linear intermediate **21**. Regioselective aldol cyclization between C-4/C-9 and C-2/C-11 is catalysed by the PT domain (intermediate **22**). Via a Claisen-like cyclisation, release of the NR-PKS product is catalysed by the TE/CLC domain. Further biosynthetic tailoring steps yield aflatoxin B₁ **25**.

PR-PKS are known to produce small aromatic compounds such as 6-methylsalicylic acid **4** and mellein **28**. The 6-methylsalicylic acid synthase (6-MSAS) from *Penicillium patulum* is one of the best studied PR-PKS and is responsible for the production of **4**.⁵ Mellein **28** was demonstrated to be produced by the PR-PKS SACE553 from the soil bacterium *Saccharopolyspora erythraea*.³³ Its associated KR domain selectively reduces the β-keto group of the diketide and tetraketide intermediates, whereas in the biosynthesis of **4** only reduction of the triketide intermediate is catalysed (Scheme 1.4). Comparative studies of KR domains indicated that the KR domain is the main determinant of the programming mechanism.³⁴



Scheme 1.4: Partially reduced polyketides: **A**, 6-MSAS reduces the keto group of the triketide intermediate during the biosynthesis of 6-methylsalicylic acid **4**; **B**, SACE5532 reduces the keto group of the diketide and tetraketide intermediates during the biosynthesis of mellein **28**.

HR-PKS produce complex, highly reduced compounds. All domains for β-processing (KR, DH, ER) are present, resulting in highly/ fully saturated carbon chains similar to fatty acids. In addition, C-methyltransferases (C-MeT) can catalyse the methylation of the α-carbon on the extended polyketide, a modification which is not observed during the biosynthesis of fatty acids. S-adenosyl

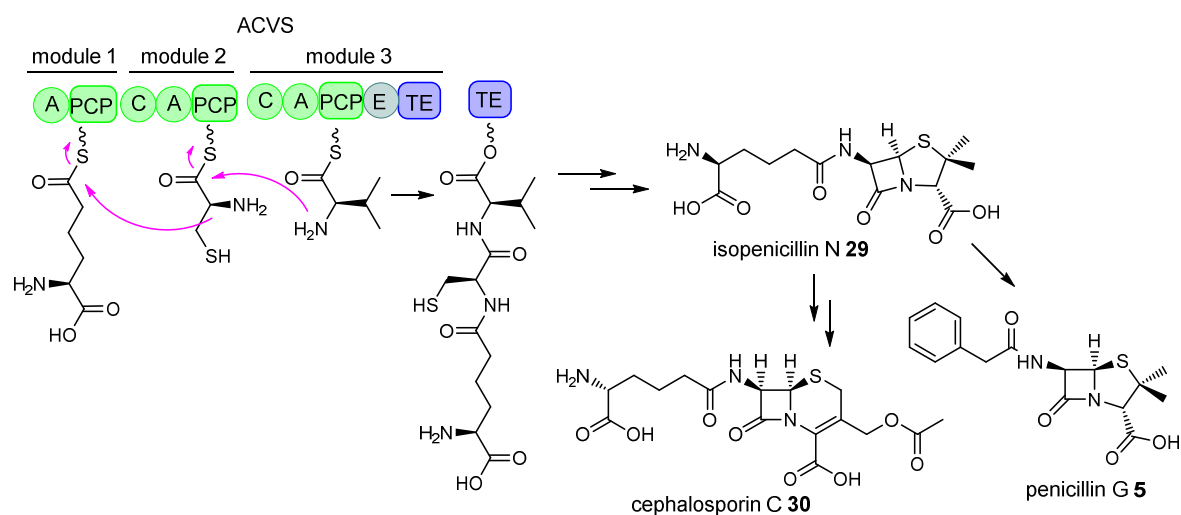
methionine (SAM) is required as cofactor.^{35,36} Prominent examples of HR-polyketides are lovastatin **2** from *Aspergillus terreus* and squalestatin S1 **1** from *Phoma* species (Figure 1.1).^{2,3}

1.4 Non-ribosomal Peptide (NRP) Biosynthesis and NRP Synthetases

While most peptides are regarded as primary metabolites, some of them also belong to the group of secondary metabolites. Notably, a few of them are gene encoded, but have undergone several posttranslational modification events (tailoring) to gain structural diversity.³⁷ For example, lantibiotics are peptides of ribosomal origin but have been post-translationally cyclized. Non-ribosomal peptides, like the precursor of penicillin G **5**, are not gene encoded but are synthesized on large multifunctional enzymes called non-ribosomal peptide synthetases (NRPS). Non-ribosomal peptides are assembled from a large pool of possible precursors including non-proteinogenic and D-amino acids as well as amino acids with other modifications such as *N*-methyl groups.³⁸

NRPS are organized in a modular fashion in a comparable manner to modular PKS; each module is responsible for the incorporation of one amino acid into the final product. All modules consist of an adenylation domain (A), a condensation (C) domain and a thiolation domain (T, or peptidyl carrier protein domain (PCP)). Together, these three core domains comprise a minimal NRPS module. The final module of a NRPS usually contains a thioesterase domain (TE) to catalyse the hydrolytic release of the product. Reductive release (R) domains are also known. In addition, a NRPS may also have additional catalytic domains such as epimerases (E), cyclases (Cyc), methyltransferases (MeT) and formylases (F).^{39,40}

A prominent example of a linear NRPS is the ACV synthetase which is responsible for the production of the tripeptide δ -(L- α -amino adipoyl)-L-cysteinyl-D-valine (ACV), which is a precursor of isopenicillin N **29**. Subsequent processing of **29** leads to the formation of natural penicillins and cephalosporins such as penicillin G **5** and cephalosporin C **30**. β -lactam antibiotics are produced by various prokaryotic and eukaryotic species including the filamentous fungi *Cephalosporium acremonium* and *Penicillium chrysogenum*.⁴¹ Notably, the δ -carboxyl moiety of the non-proteinogenic amino acid α -amino adipate participates in peptide bond formation (Scheme 1.5).^{42,43}



Scheme 1.5: Biosynthesis of penicillin G 5 and cephalosporin 30.

Iterative NRPS systems as observed for many fungal PKS, are also known.^{38,44} They use one module more than once during the biosynthesis of a single peptide.

1.5 Polyketide Synthase - Non-ribosomal Peptide Synthetase (PKS-NRPS) Hybrids

Microorganisms, especially fungi, produce a wide range of bioactive compounds. In addition to PKS and NRPS *per se*, natural products can derive from polyketides fused to amino acids, which are synthesised by hybrids between a polyketide synthase and non-ribosomal peptide synthetase modules (PKS–NRPS).⁴⁵ The construction of each part of the hybrid synthetase (PKS and NRPS part) is similar to the separate systems mentioned previously. Among the variety of polyketide and non-ribosomal hybrid compounds, pyrroline-2,4-dione systems, generally known as tetramic acids, are found in many structures.⁴⁶ Examples of natural products which are derived from fungal tetramate synthases include cyclopiazonic acid **31** (*Aspergillus flavus*),⁴⁷ equisetin **32** (*Fusarium heterosporum*)⁴⁸ and pretenellin A **33** (*Beauveria bassiana*).^{7,49} Their corresponding BGC contain a conserved PKS-NRPS gene encoding an N-terminal iterative HR-PKS and a single C-terminal NRPS module. In each case a polyketide 3-ketoacyl chain is formed by the PKS module. The NRPS module appears to have two functions: a) the selection and installation of the amino acid and b) releasing the NRPS bound acyl chain as tetramic acid.⁴⁶ The PKS-NRPS genes involved in the biosynthesis of fusarin C **8** and cytochalasans also belong to this class of fungal tetramate BGC.^{50,51}

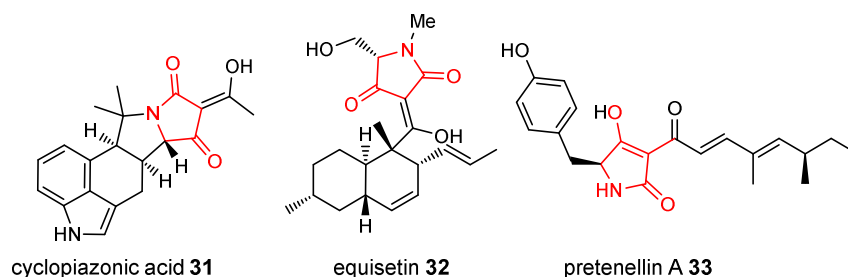


Figure 1.3: Selected natural products containing a tetramic acid moiety.

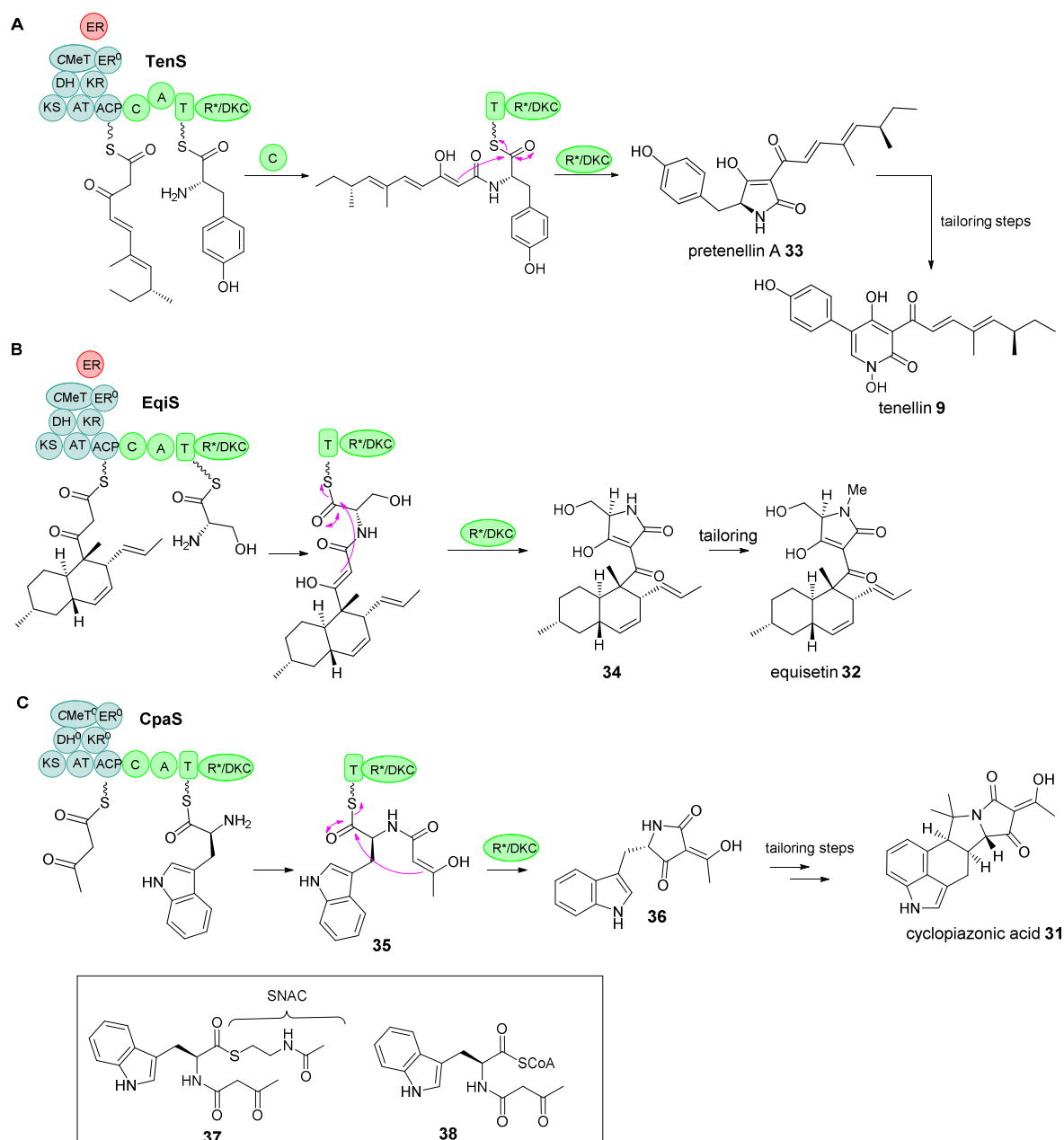
The tetramic acid moiety is recognised by a variety of biological targets due to its conformational constraint and its dicarbonyl functionality which offers H-bond interactions with target proteins. This was shown for example by crystallization experiment of cyclopiazonic acid **31** complexed with its target sarcoplasmic reticulum Ca^{2+} -ATPase.^{52,53} Another example is equisetin **32** which also demonstrates bioactivity: it is known to inhibit the HIV-1 integrase and is toxic for various bacterial strains.^{54,55}

In addition to hydrolysis, macrocyclization (lactonization/ lactamization) and reduction, formation of a tetramate ring could be considered as an NRPS release mechanism. Analogous to pure NRPS assembly lines, the chain termination NRPS module of a hybrid PKS-NRPS assembly line usually contains a TE domain, for hydrolysis or macrocyclization, or an R domain catalysing a reductive release. All known fungal PKS-NRPS systems appear to have a C-terminal reductive release domain. R domains exhibit sequence similarities to the short-chain dehydrogenase/reductase (SDR) superfamily with Rossmann-fold and nucleotide binding motifs.⁵⁶ R domains catalyse the NAD(P)H-dependent reductive release of the NRPS bound acyl chain as an aldehyde.

The tetramate ring formation was initially thought to always occur *via* a reductive release to generate an aldehyde that undergoes spontaneous ring closure to yield a pyrrolidine-2-one. Subsequent oxidation of the pyrrolidine-2-one by one or more cytochrome p450 monooxygenases would lead to the formation of the tetramate acid.^{48,49,57} This initial hypothesis was proven to be incorrect by Cox and co-workers when they heterologously expressed the pretenellin A **33** PKS-NRPS (*tenS*) together with its *trans*-acting enoyl reductase (*tenC*) in *Aspergillus oryzae*. TenS and TenC led to the production of pretenellin A **33** which contains a tetramic acid moiety in its scaffold (Scheme 1.6 A). For the first time it was shown that the R domain of TenS has a non-redox function and can operate as condensation catalyst. Since it catalyses a Dieckmann cyclisation reaction it was named DKC.⁵⁸ Walsh and co-workers labeled this type of R domain R*.⁴⁶

A similar fungal PKS-NRPS system responsible for the production of equisetin **32** was found in the filamentous fungi *Fusarium heterosporum*.⁴⁸ Heterologous production of EqiS (PKS-NRPS) R domain in *E. coli* followed by reactivity assays with synthetic model substrates demonstrated that EqiS_R does not perform a reducing function. Instead, a redox-independent generation of the tetramate moiety was observed where the cofactor NAD(P)H was not required. Thus, the releasing domain of EqiS also catalyses a Dieckmann condensation (Scheme 1.6 B).⁵⁹

Walsh and co-workers further investigated the question whether the reductase-like R*/DKC domain, present in fungal PKS-NRPS systems generating a tetramic acid, is involved in a non-redox reaction and does not release the PKS-NRPS hybrid chain as an aldehyde. To address this, the biosynthesis of cycloiazonic acid **31** was examined. Similar to the studies on the T and R* domains of the equisetin synthase EqiS, the T_ R*/DKC domain from the *A. flavus* PKS-NRPS CpaS was investigated *in-vitro*. Unlike EqiS_ R*/DKC, CpaS_ R*/DKC only accepted its substrate (*N*-acetoacetyl-L-tryptophan (*N*-acac-Trp) **35**) when it was natively tethered to the pantetheinyl prosthetic group of the T domain. Low molecular weight substrate mimics (*N*-acac-Trp-SNAC **37** or *S*-Coenzyme A **38**) were not accepted. In accordance with the studies with EqiS_ R*/DKC, redox-independent conversion of CpaS_T bound **35** to free *cyclo*-acetoacetyl-L-tryptophan (cAATrp) **36**, catalysed by CpaS_ R*/DKC, was observed. These findings were interpreted as providing further evidence that the tetramic acid is formed *via* a Dieckmann cyclization as release step (Scheme 1.6 C).⁴⁶

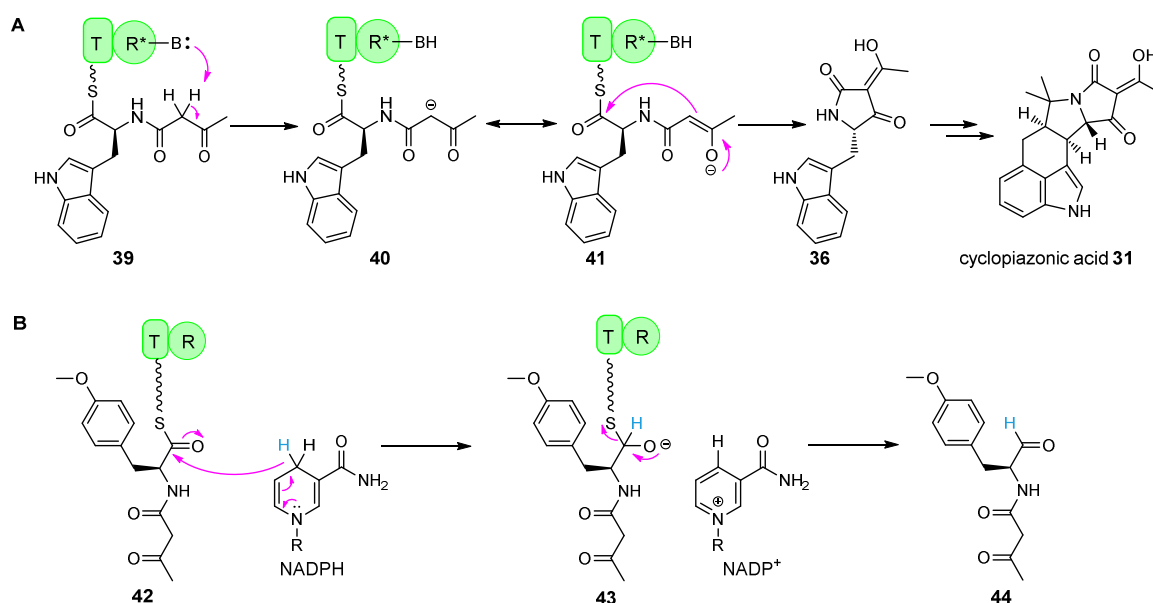


Scheme 1.6: Selected relationships within PKS-NRPS in terms of their biosynthetic release mechanism.

The release domain R*/DKC of EqiS and CpaS still belongs to the SDR superfamily, although they are redox inactive. Walsh and co-workers compared the sequences of several R and R*/DKC domains to elucidate how the R*/DKC domain might have evolved from the R domain.⁴⁶ R domains and canonical SDR enzymes feature a catalytic triad of Ser-Tyr-Lys which is essential for catalytic activity.^{60,61} The catalytic Tyr is mutated to Phe in the tenellin synthetase R*/DKC domain and to Leu in the CpaS_R*/DKC domain. However, the equisetin R*/DKC domain still contains the catalytic triad, indicating that there must be other reasons for the loss of redox activity.⁴⁶

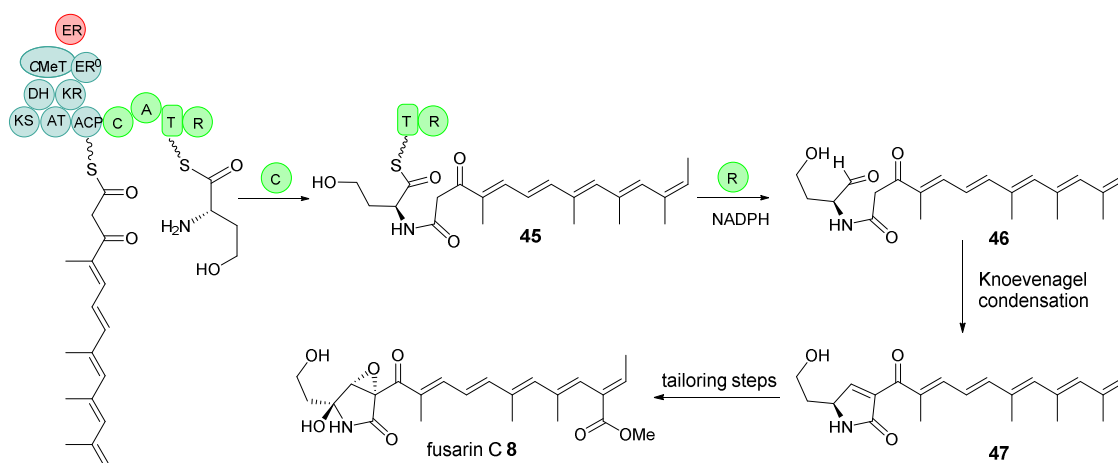
Using sequence comparison of the C-terminal R*/DKC domains, seven polar amino acid residues in CpaS_R*/DKC (C3619, D3663, H3667, S3688, D3772, D3803, H3843) were identified which are not present in redox-NRPS R domains. Mutation of D3803 to Ala reduced Dieckman condensation levels back to baseline. It was proposed that the aspartate carboxylate side chain might function as a base to generate a resonance-stabilized carbanion and to initiate the tetramate ring cyclization. This pathway can only occur for PKS-NRPS hybrid chains where the *N*-acyl substituents are 3-ketoacyl groups (Scheme 1.7 A).⁴⁶

In comparison to the Dieckmann release pathway, a reductive release catalysed by the R domain is initiated by the delivery of a hydride from NAD(P)H to the carbonyl group of the thioester to yield a tethered thiohemiacetal **43**. The negative charge of the oxygen collapses back to the carbonyl, ejecting the HS-pantetheinyl arm and releasing the aldehyde **44** (Scheme 1.7 B).⁴⁶



Scheme 1.7: Comparison of the NRPS C-Terminal R and R*/DKC Domain: **A**, redox-independent Dieckmann cyclization catalysed by the R*/DKC domain; **B**, NAD(P)H dependent reductive release catalysed by the R domain.

Reductive release catalysed by the R domain occurs for example during the biosynthesis of fusarin C **8**. The acyl bound PKS-NRPS chain is probably released as free aldehyde (Scheme 1.8).⁵⁰



Scheme 1.8: Reductive release catalysed by the R domain during the biosynthesis of fusarin C 8.⁵⁰

1.6 Cytochalasans

Cytochalasans are a diverse group of fungal polyketide-non-ribosomal peptide metabolites with a broad range of biological properties including antibacterial, antiviral and antitumor potential. The first identified and investigated cytochalasans were reported in 1966 and were named cytochalasins A (also known as dehydrophomin) **48** and B (also known as phomin) **49** (Figure 1.4).^{62,63} Both compounds exhibit toxicity against mammalian cells.⁶⁴ The name derives from the Greek word *kytos*, meaning cell, and *chalsins* meaning relaxation, due to its well-known effects on actin polymerisation. Cytochalasans are able to bind to actin filaments and thereby inhibit their polymerisation and elongation, which negatively effects cellular processes like cell division, motility and signalling.^{65,66,67} Furthermore, at higher concentration some cytochalasans interfere with glucose transport⁶⁸ and demonstrate inhibition of cholesterol transport in rat adrenals studies.⁶⁹ Current estimates suggest that over 300 different cytochalasans have been reported from various fungal sources.⁷⁰

All cytochalasans contain a tricyclic core, which consists of a macrocyclic ring fused to an isoindolone moiety. Cytochalasans derive from a highly reduced polyketide backbone and an amino acid. The differences among the cytochalasans arise mostly from variations in the size and substitution pattern of the macrocycle as well as the incorporated amino acids. The most common amino acids are phenylalanine (cytochalasins), tyrosine and *O*-methyltyrosine (pyrichalasin) and tryptophan (chaetoglobosins, Figure 1.4).

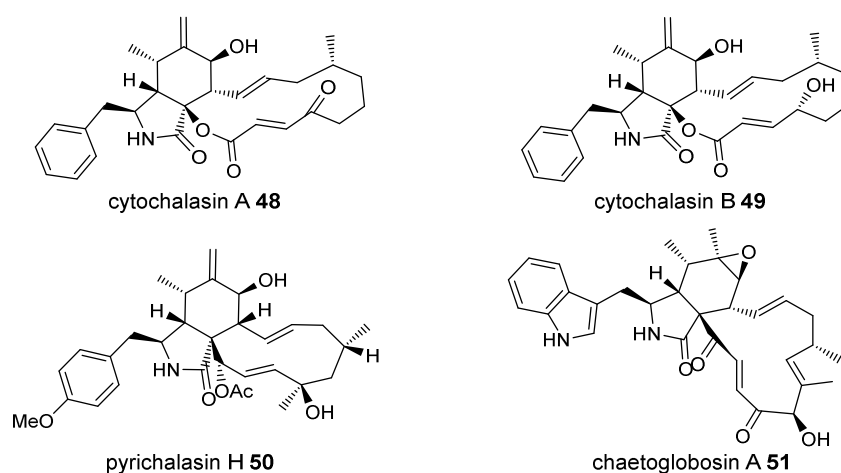


Figure 1.4: Chemical structures of cytochalasins.

The biosynthetic origin of cytochalasin B **49** was first investigated by Tamm and co-workers in *Phoma sp.* by isotopic feeding experiments (^3H , ^{13}C , ^{14}C and ^{18}O), followed by many other labelling studies.^{71–73} It was elucidated that while the polyketide backbone of cytochalasins consists of acetates and malonates, additional methyl groups are transferred from *S*-adenosylmethionine (SAM) and additional oxygen-atoms were revealed to be of atmospheric origin which is in accordance with the activity of a HR-PKS.

The first cytochalasan gene cluster to be reported was the chaetoglobosin A (*che*) cluster from *Penicillium expansum* containing a gene encoding a hybrid iterative type I polyketide synthase-non-ribosomal peptide synthetase (PKS-NRPS).⁷⁴ The cluster was presumed to be confirmed by RNA-mediated silencing of *cheA*, which led to almost complete loss of production of **51**. The gene cluster also encodes a *trans*-acting enoyl reductase (CheB), which can be observed in other fungal PKS-NRPS systems as well, such as in the biosynthesis of tenellin **9**.⁷⁵ Two putative P450 monooxygenases (CheD and CheG), two transcription factors (CheC and CheF) as well as an FAD-dependent monooxygenase (CheE) are also part of the cluster (Scheme 1.9 A).⁷⁴ Later, a second chaetoglobosin A **51** gene cluster was identified in another fungus producing **51**, *Chaetomonium globosum* and demonstrated to be the true BGC.⁷⁶ Previous silencing experiments might have caused off target effects, apparently silencing the true BGC. The function of the true cluster in *C. globosum* was confirmed by targeted gene deletion of the PKS-NRPS, encoded by *CHGG_01239* and its adjacent *trans*-acting enoyl reductase (*CHGG_01240*). Deletion mutant strains were no longer able to produce **51**. Similar to the *cheA* BGC, the *C. globosum* cluster encodes two cytochrome P450 oxygenases (*CHGG_01242-1* and *CHGG_01234*), a transcription factor (*CHGG_01237*) and an FAD-dependent monooxygenase (*CHGG_01242-2*). However, two additional genes are present in the cluster encoding proteins of (at the time) unknown function

(CHGG_01241 and CHGG_01244). CHGG_01241 exhibits no significant similarity to any characterised proteins. CHGG_01244 on the other hand has sequence similarity to $\alpha\beta$ -hydrolases. Experiments by Watanabe and co-workers have elucidated many biosynthetic steps leading to the formation of **51** from prochaetoglobosin I **56**.⁷⁶ Collectively, the late stages of the biosynthesis of cytochalasans are almost completely understood and involve predominantly oxidative tailoring steps (Scheme 1.9 B).

Another cytochalasan gene cluster, *css*, was identified in *Aspergillus clavatus* by genome mining. Deletion strains of *cssA*, encoding a PKS-NRPS knockout, did not produce cytochalasin E **52** or K **53**, confirming the association between the gene cluster and the natural products. This cluster also encodes a *trans*-acting enoyl reductase (CcsC), two P450 monooxygenases (CcsD and CcsG), a transcription factor (CcsR) and two proteins with unknown function (CcsF and CcsE). CcsF shares 51 % identity to CHGG_01241 and was previously predicted to catalyse the Diels-Alder cyclization in the biosynthetic pathway of **52** and **53**, but this hypothesis could not be confirmed experimentally so far.^{51,77} Therefore, CcsF and its homolog CHGG_01241 can be referred to as a putative Diels-Alderase (pDA).⁷⁸ CcsE shows sequence homology to $\alpha\beta$ -hydrolases; it shares 54 % identity to CHGG_01244. Unlike the chaetoglobosin gene cluster, the *css* cluster also encodes a Bayer-Villiger monooxidase (CcsB) which generates the unique vinyl carbonate moiety in **52** and **53**.⁵¹

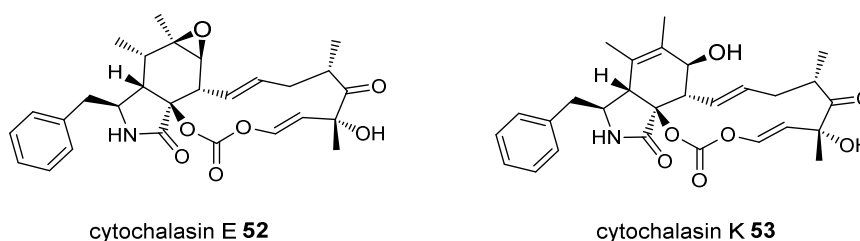
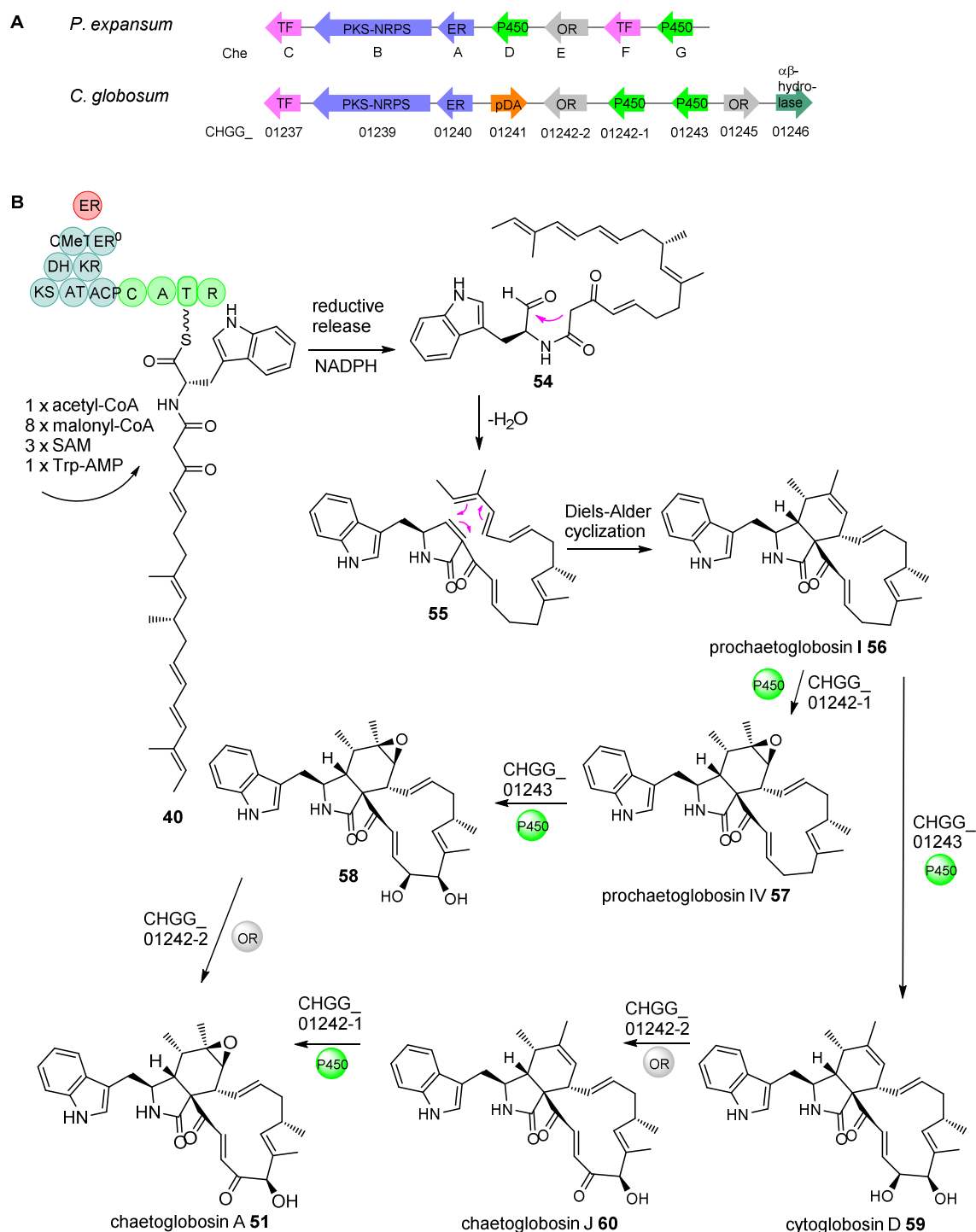


Figure 1.5: Chemical structures of cytochalasin E **52** and K **53**.



Scheme 1.9: Proposed biosynthetic pathway of chaetoglobosin A **13**.⁷⁴ KS = ketosynthase, AT = acyltransferase, DH = dehydratase, C-MeT = C-methyltransferase, ER⁰ = enoyl reductase, KR = ketoreductase, ACP = acyl carrier protein, A = adenylation, C = condensation, T = thiolation, R = release, ER = enoyl reductase, OR = oxidoreductase, TF = transcription factor, CMeT = C-methyltransferase, P450 = cytochrome P450 monooxygenases, pDA = putative Diels-Alderase.

1.7 The ACE1 Metabolite from *Pyricularia oryzae*

1.7.1 Introduction

The rice blast fungus *Pyricularia oryzae* is the most devastating fungal pathogen for rice plants as it is responsible for the destruction of rice yield, which could potentially feed 60 million people annually.⁷⁹ In order to challenge this threat, its mode of infection has been studied extensively. A spore of *P. oryzae* attaches to the surface of the rice leaf and germination occurs. A germ tube grows away from the spore and a specialised structure, the appressorium, is formed.⁸⁰ The appressorium promotes penetration into the plant epidermal cells by enzymatic and mechanical processes.⁸¹ The fungus starts to grow within the infected cell before spreading to neighbouring cells. A necrotic growth pattern occurs and results in the formation of necrotic lesions distributed on infected plant leaves. Eventually, fungal sporulation occurs from severely infected rice leaves, initiating another cycle of infection.⁸²

Certain rice strains (*Oryza sativa*) carrying the Pi33 gene are resistant to specific avirulent strains of *P. oryzae*, which encode a polyketide synthase – non-ribosomal peptide synthetase (PKS-NRPS) called the avirulence conferring enzyme 1 (ACE1).⁸³ The specific interaction between virulent and avirulent *P. oryzae* strains and the corresponding resistant or sensitive rice cultivars follow the classic gene-for-gene interaction.⁸⁴ Disease is caused for all combinations of plant and fungi except in the single case of resistant rice, carrying the resistant allele (R) which encodes a receptor and avirulent *P. oryzae*, carrying the avirulence gene (AVR) encoding an elicitor, which is recognized by the receptor. Due to the specific gene for gene interaction, the resistant rice plant is able to initiate an immune response towards the fungus (Figure 1.6).

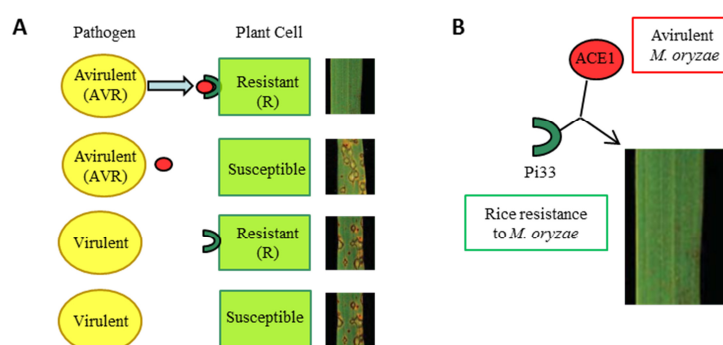


Figure 1.6: The receptor-elicitor model of gene-for-gene interactions: **A**, The resistance allele (R) of the plant encodes a receptor that recognizes an elicitor produced by the pathogen. Recognition of the pathogen elicitor by the plant receptor initiates plant defence response that lead to plant resistance. If the pathogen produces the elicitor (AVR), it is avirulent. If the pathogen does not produce the elicitor, it is virulent; **B**, Interaction between ACE1 from avirulent strain *P. oryzae* and receptor Pi33 from resistant rice through an unknown signal synthesised by ACE1.⁴⁵

For most fungi-plant-interactions the avirulence product is a protein which is directly recognized by the host plant carrying the resistance gene.⁸⁵ The *P. oryzae* *ACE1* avirulence gene differs from other fungal avirulence genes, since it encodes a polyketide synthase – non-ribosomal peptide synthetase, which is *per se* not secreted. Thus, *ACE1* is an enzyme involved in the biosynthesis of a secondary metabolite likely recognized by rice cultivars carrying the Pi33 resistance gene.⁸⁶ Mutation of the putative catalytic site of the β -ketoacyl synthase domain of the PKS-NRPS prohibits recognition of the fungus by resistant rice, suggesting that the biosynthetic activity of *ACE1* is needed for avirulence signalling. The Lebrun group demonstrated that *ACE1* is only produced temporarily and under very tight conditions during appressorial penetration. The reporter gene *egfp*, encoding the enhanced green fluorescence protein (eGFP), was fused to the *ACE1* promoter, forming the pPro*ACE1*:GFP plasmid, which was transformed into the avirulent *P. oryzae* strain Guy11. Upon excitation (450 – 490 nm), a GFP signal was detected only in the cytoplasmic compartments of the appressorium (15 h after inoculation on barley or rice leaves) and after penetration into the epidermal cells (36 h) also in infectious hyphae. Once secondary infectious hyphae began to spread within the infected leaf tissues, the GFP signal was no longer observed.⁸⁶ To date, the *ACE1* metabolite could not be successfully isolated; its structure still remains elusive.

The Lebrun group also demonstrated that *ACE1* is responsible for avirulence signalling by infection assays of rice and *Pyricularia oryzae* strains. Partial deletion of the *ACE1* gene in the avirulence fungus strain *P. oryzae* Guy11 was achieved by the plasmid p Δ *ace1*.⁸⁶ Resistant rice cultivar seedlings were infected with the mutant or with the wild type strain.⁸⁶ The rice plant contaminated with the *ace1* deletion mutant developed typical pathological symptoms such as necrosis. As a control, the plasmid *pace1_egfp*, expressing *ACE1*, was transformed into the virulent *M. grisea* 2/0/3 strain, changing it to avirulent (on resistance rice, Figure 1.7).

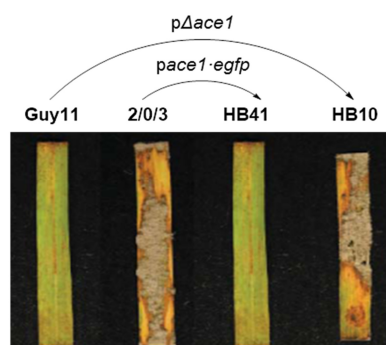


Figure 1.7: ACE1-mediated avirulence signalling. The avirulent isolate Guy11, the virulent strain 2/0/3, the transformant HB41 (2/0/3 expressing the ACE1:GFP fusion) and the Guy 11 ACE1 deletion mutant HB10 were inoculated on resistant rice cultivar. Strains expressing the ACE1 gene (native or as ACE1:eGFP fusion) conferred avirulence signalling.

1.7.2 ACE1 Gene Cluster and Proposed Biosynthesis

The avirulence conferring enzyme 1 (ACE1) is encoded by the ACE1 locus in *P. oryzae*, which was identified by map-based cloning with a genetic map constructed by restriction fragment length polymorphism and random amplified polymorphic DNA markers.⁸⁷

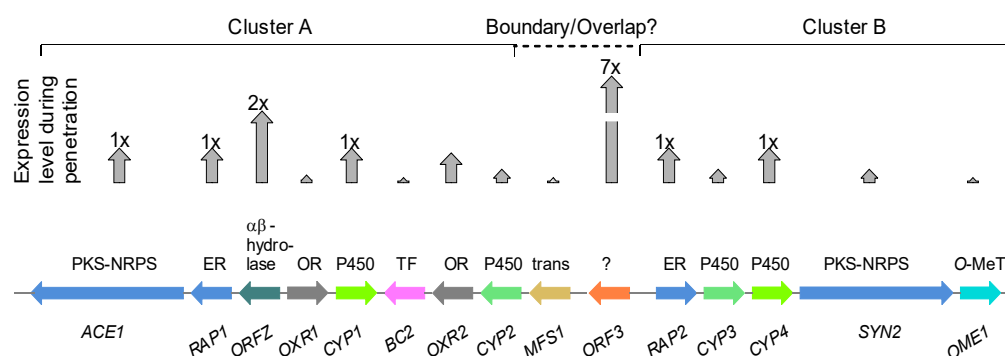


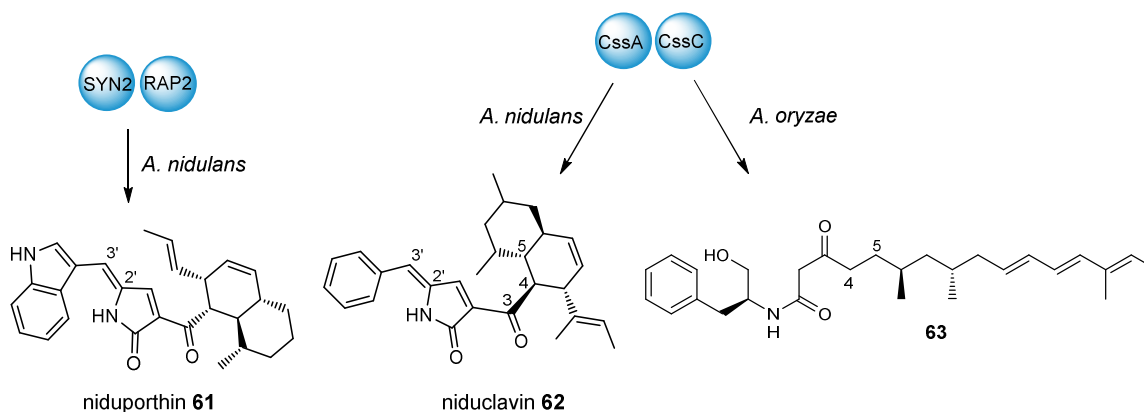
Figure 1.8: Putative ACE1 gene cluster and expression levels of ACE1 and its neighbouring genes during infection monitored by real-time RT-PCR using *ILV5* expression level as a standard to normalize the data. RNA from barley leaves infected by *P. oryzae* isolate Guy11 and RNA from mycelium grown on liquid complete medium, collected at different times after inoculation, was used.⁸⁸ ER = enoyl reductase, OR = oxidoreductase, TF = transcription factor, trans = transporter, O-MeT = O-methyltransferase, P450 = cytochrome P450 monooxygenases.

Expression of genes from the ACE1 locus was monitored by quantitative real-time RT-PCR using the *ILV5* expression level as a standard to normalize the data.⁸⁸ ACE1 is located in a cluster of 15 genes which all share the same specific and tight expression pattern during the early stage of plant infection. The cluster consists of two genes encoding PKS-NRPS (*ACE1* and *SYN2*), two trans-acting enoyl reductases (*RAP1* and *RAP2*), four cytochrome P450 monooxygenases (*CYP1-CYP4*), two oxidoreductases (*OXR1*, *OXR2*), one $\alpha\beta$ -hydrolase (*ORFZ*), one protein with unknown function

(*ORF3*), one MFS transporter (*MFS1*), one *O*-methyltransferase (*OME1*) and one putative transcription factor (*BC2*, Figure 1.8).

On the basis of extensive bioinformatic analysis done by Khaldi and co-workers, good evidence was obtained that the *ACE1* cluster can be divided into two sub-clusters named A and B.⁸⁹ It appears likely that both sub-clusters derive from an ancestral cluster by a gene duplication. Some genes may have been deleted or disrupted during this event such as *OXR1* and *ORFZ*, which are not represented in cluster B. Based on this hypothesis, at some point in the past the encoded proteins would have been identical in terms of selectivity and programmed chemistry. From this hypothetical point it is not clear, how far their respective chemical properties have diverged. Furthermore, it is not clear how independent the two sub-clusters are of each other or whether there is any cross-interaction between their gene products.

Gene inactivation of *SYN2* and *RAP2* by disruption and gene deletion experiments demonstrated that they are not required for avirulence in Pi33 resistant rice cultivars, whereas *ACE1* is essential.⁸⁸ Therefore, cluster B might not be involved in the biosynthesis of the *ACE1* metabolite. Larson and co-workers recently discovered that heterologous expression of *SYN2* and *RAP2* in *A. nidulans* led to production of niduporthin **61** (Scheme 1.10).⁹⁰ **61** contains a decalin ring and a nitrogen-containing tetramic acid unit and is built of a highly reduced polyketide chain (octaketide) fused to tryptophan. A similar compound, named niduclavin **62**, was observed when they co-expressed *ccsA* (PKS-NRPS) and *ccsC* (*trans*-acting ER) from the cytochalasin E **52** and K **53** BGC in *A. nidulans* (Figure 1.5, Scheme 1.10). Notably, niduclavin **62** consists of the same underlying polyketide chain length and amino acid as compound **63**, which was identified by Oikawa and co-workers when *cssA* and *cssC* were co-expressed in *A. oryzae* (Scheme 1.10).⁹¹ Similar to **61**, **62** contains a decalin ring system which is likely to be derived from a spontaneous intramolecular Diels-Alder reaction with the olefin at C-4/C-5 within the polyketide chain. Notably, **63** does not contain the olefin at C-4/C-5. The C-2'/C-3' double bond is believed to derive from an oxidation shunt reaction by native enzymes of *A. nidulans*. The authors speculated that the C-2'/C-3' olefin might favour the Diels-Alder reaction by activating the dienophile in the $\alpha\beta$ -position of the C-3 ketone. The similarity between **61** and **62** suggests that *SYN2* is involved in the biosynthesis of cytochalasans, albeit its final compound is unknown.⁹⁰



Scheme 1.10: Heterologous expression of *SYN2* and *RAP2* in *A. nidulans* leads to production of niduporthin **61**. Heterologous expression of *cssA* and *cssC* results in the production of niduclavin **62** (in *A. nidulans*) and **63** (in *A. oryzae*).⁹⁰

Further gene deletions in the *ACE1* cluster could not be achieved, even after inactivation of the non-homologous end joining (NHEJ) system, indicating that the *ACE1* locus has a greatly reduced ability for homologous recombination.⁹² Phylogenetic analysis of the *ACE1* cluster in filamentous fungi revealed that the *ACE1* BGC is specific to a few fungal species, always consisting of at least three genes (homologous to *ACE1*, *RAP1* and *ORF3*), suggesting a crucial role of these genes in the biosynthetic pathway of the *ACE1* metabolite.⁸⁹

Protein sequence analysis shows that the 12.4 kb *ACE1* gene encodes a hybrid HR-PKS-NRPS.³⁰ The PKS contains seven catalytic domains: KS (ketosynthase), AT (acyltransferase), DH (dehydratase), C-MeT (C-methyl transferase), ER⁰ (inactive enoyl reductase), KR (ketoreductase), and ACP (acyl carrier protein, Figure 1.9). The ER domain appears to be non-functional due to a variation in its consensus sequence. Based on this hypothesis it is presumably complemented by the *trans*-acting ER encoded by *RAP1*. This phenomenon has been observed in other biosynthetic pathways such as the tenellin **9** and lovastatin **2** systems (Scheme 1.14 and Scheme 6.9).^{49,93}



Figure 1.9: Catalytic domains of the PKS-NRPS *ACE1*. KS = ketosynthase, AT = acyltransferase, DH = dehydratase, C-MeT = C-methyl transferase, ER⁰ = inactive enoyl reductase, KR = ketoreductase, ACP = acyl carrier protein, C = condensation domain, A = adenylation domain, T = thiolation domain and R = reductive release domain.

A comparison of *ACE1* with other well-known fungal PKS-NRPS enzymes revealed that many of them produce acyl tetramic acids including pretenellin A **33**⁹⁴ and preaspyridone A **64**,⁹⁵ as well as pyrrol-2-ones such as prefusarin C **47**⁵⁰ and cytochalasans like chaetoglobosin A **51** and

cytochalasin E **52**⁵¹ (Figure 1.10). Therefore, it appears likely that *ACE1* encodes a biosynthetic protein which produces a small molecule consisting of a polyketide fused to an amino acid.

The presence of a reductive release domain in the *ACE1*-NRPS indicates the release of the acyl bound PKS-NRPS chain as a free aldehyde as observed in the biosynthesis of fusarin C **8** (Scheme 1.8).⁵⁰ The same reductive release mechanism of cytochalasins was proposed for compound **63**, due to its comparable structure.⁹¹ However, without any further experimental evidence it cannot be excluded that the R domain functions as Dieckmann cyclase as observed in the biosynthesis of tenellin **9** and equisetin **32** (Scheme 1.6 B and C).^{58,59} However, the presence of genes encoding an $\alpha\beta$ -hydrolase and a pDA in the *ACE1* BGC also suggest a strong link to cytochalasan biosynthesis. Other tailoring genes such as *CYP1* - *CYP4* and *OXR1/2* are also similar to tailoring genes from known cytochalasan BGC.

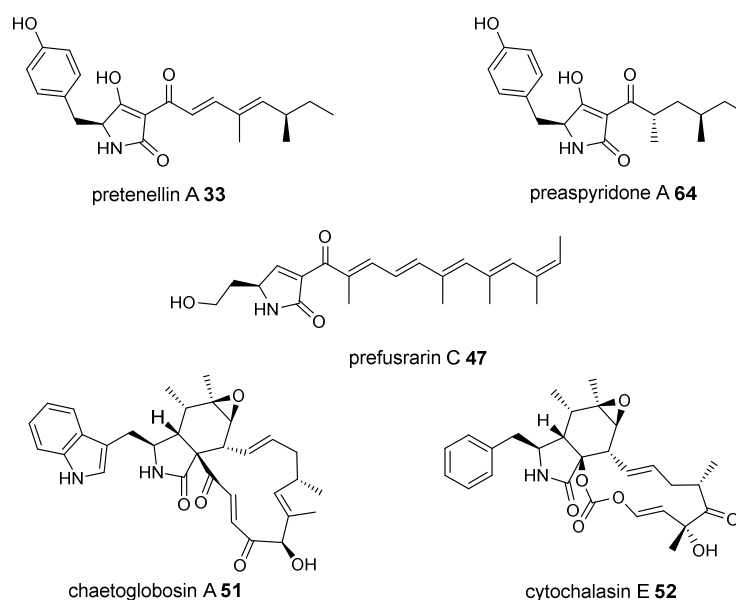
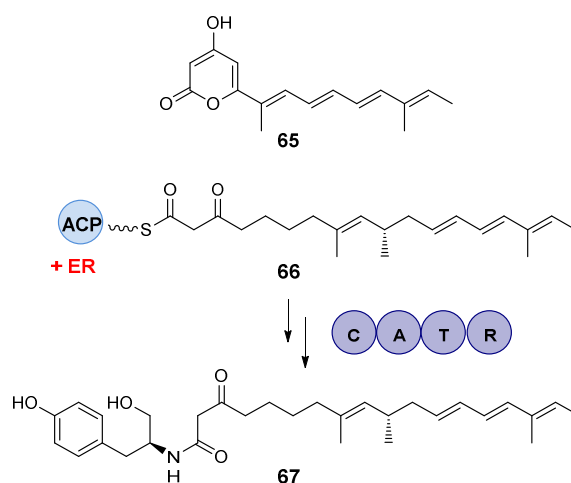


Figure 1.10: Structures of representative fungal PKS-NRPS metabolites.

Moreover, Cox and co-workers demonstrated that the heterologous expression of *ACE1* (introns were removed from the sequence) in *A. oryzae* led to the formation of compound **65** (Scheme 1.11), which is exclusively of polyketide origin, as proven by isotopic labelling studies. Co-production of the *trans*-acting ER (*RAP1*) is required to generate the full PKS product **66**, before the amino acid can be attached by the NRPS forming compound **67**.⁹⁶ Similar results were obtained for the biosynthesis of lovastatin **2**, where the lovastatin nonaketide synthase (LNKS) also requires its cognate *trans*-ER, encoded by *lovC*, for the correct processing of the growing polyketide chain.⁹³



Scheme 1.11: Deduced biosynthesis of ACE1 products in the presence and absence of RAP1. ACP = acyl carrier protein ACP (domain of ACE1), ER: enoyl reductase (RAP1); C = condensation domain, A = adenylation domain, T = thiolation domain and R = reductive release domain.

1.7.3 Melanization and Appressorium Formation

Fungi often produce various dark pigments known as melanins. Melanins are compounds of high molecular weight which are formed by oxidative polymerization of phenols. They are not essential for maintaining normal cell growth and development; moreover they seem to be involved in various protective processes against environmental stress such as UV radiation.⁹⁷ In addition, it was observed that the biosynthesis of melanin in fungi often correlates with virulence.⁹⁸

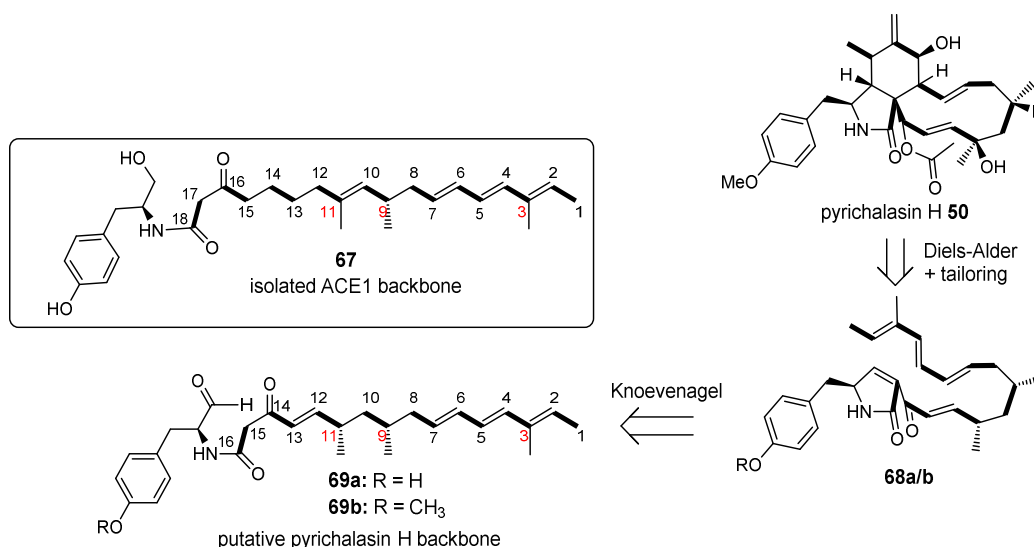
In the rice blast fungus *P. oryzae* melanization was demonstrated to be essential for maintaining the high turgor pressure present in the appressorium. A thick 1,8-dihydroxynaphthalene (DHN)-melanin layer is formed between the cell wall and the plasma membrane. The appressorium, a specialized cellular apparatus, is formed by the plant-pathogenic fungus and is needed for penetration into the cell wall. Due to a high turgor pressure within the appressorium, a hole is physically forced through the plant cell wall.^{99,100,101} Fungicides which inhibit melanin biosynthesis have been demonstrated to prevent penetration of the plant tissue.^{102,103}

Most fungal melanins derive from polymerization of the precursor 1,8-dihydroxynaphthalene **106** (DHN). The biosynthesis of DHN-melanin from pentaketides was originally discovered in *Verticillium dahliae*.^{104,105} Since then, the biosynthetic route has been investigated extensively (see Section 5.5). However, whether or how the biosynthesis of the ACE1 metabolite is connected to the formation of DHN-melanin in appressorial cells is unknown.

1.8 The Pyrichalasin H Biosynthetic Gene Cluster

Pyrichalasin H **50** is a phytotoxin produced by *Magnaporthe grisea*, a fungus closely related to *P. oryzae*. Compound **50** causes pathogenicity on weed species of grass (crabgrass, *Digitaria sanguinalis*).¹⁰⁶ In clinical studies in mice it was demonstrated that **50** leads to a strong defect in cell morphology in C3H-2K cells as well as a strong inhibition of lymphocyte capping.¹⁰⁷ The metabolite belongs to the family of cytochalasans with the speciality of a *p*-methoxyphenyl group being substituted at C-7 (Scheme 1.12).¹⁰⁶

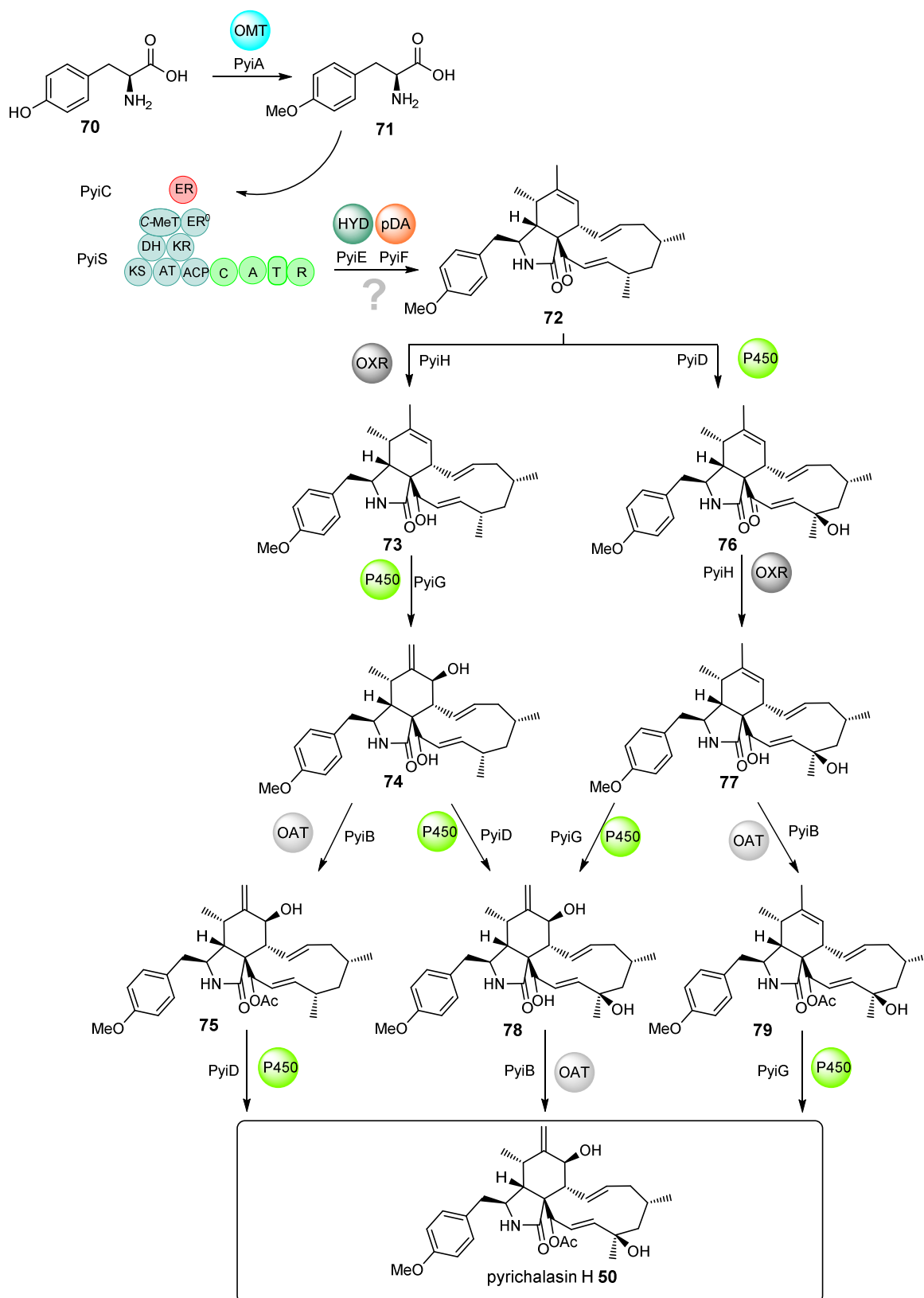
Similar to the identified ACE1 backbone **67**, tyrosine is fused to the polyketide chain to form the backbone of **50** (Scheme 1.12). Diels-Alder cyclisation, *O*-methylation and oxidative steps are thought to lead to the final product **50**. A closer look reveals that the putative pyrichalasin H backbone **69a/b** is methylated at the same positions relative to the direction of chain extension as the isolated ACE1 backbone **67**, albeit the pyrichalasin H backbone **69a/b** is two carbons (*e. g.* one extension) shorter (Scheme 1.12). In addition, **69b** could differ in its *O*-methylation pattern, whose timing is still elusive. Since **50** is structurally very similar to the ACE1 metabolite, their respective biosynthetic pathways are likely to be closely related as well.



Scheme 1.12: Proposed biosynthetic relationship between pyrichalasin H **50** and the ACE1 PKS-NRPS backbone **67**. Atom numbering highlights similarity in methylation pattern/ structures.

The genome of *M. grisea* NI981 (genomic DNA preparation done by Dr. Elizabeth J. Skellam, AK Cox, Leibniz Universität Hannover) was sequenced in collaboration with CeBiTec in Bielefeld on the MiSeq system (Illumina) in a paired-end sequencing run (estimated genome size: 41.935.773 bp). Analysis of the sequence data was done using antiSMASH⁵⁸ to identify biosynthetic gene clusters of secondary metabolites.¹⁰⁸ Three cytochalasan-like gene clusters were identified: a

putative *ACE1* gene cluster (homologous to part A of the *ACE1* gene cluster), a putative *SYN2* gene cluster (homologous to part B of the *ACE1* gene cluster) and an unknown PKS-NRPS gene cluster, which was identified by our group as the pyrichalasin H **50** BGC by targeted gene disruption of the *pyiS* gene, encoding the PKS-NRPS.¹⁰⁸ Subsequently, in parallel to this work, the late-stage biosynthetic pathway of **50** was fully elucidated using targeted gene inactivation (Scheme 1.13).¹⁰⁸ Notably, the late-stage biosynthetic pathway appeared to be not linear; rather a biosynthetic network was observed. The non-proteinogenic amino acid *O*-methyltyrosine **71** was identified as the native substrate for the A-domain of the PyiS NRPS. The biosynthetic role of PyiF (putative Diels-Alderase, homolog of ORF3) and PyiE ($\alpha\beta$ -hydrolase, homolog of ORFZ) remained unknown.



Scheme 1.13: Proposed Late-Stage Biosynthetic Pathway of Pyrichalasin H 50. OMT = *O*-methyltransferase, P450 = cytochrome P450 monooxygenase, OAT = *O*-acyltransferase, HYD = $\alpha\beta$ -hydrolase, pDA = putative Diels-Alderase, OXR = oxidoreductase.

1.9 Biomolecular Methods Used in this Study

The study of gene function in filamentous fungi is a field of research that has made great advances in the last years. The number of sequenced genomes of filamentous fungi has significantly increased, allowing the usage of protocols based on the knowledge of the gene sequence and its surrounding region.^{109,10} One approach to elucidate the function of a specific gene is by expression of a single gene or multiple genes from the gene cluster in a suitable host (see Chapter 3). Another option is the disruption or deletion of the gene of interest (GOI) in order to identify absent or newly emerged intermediates in the cellular extract of these strains compared to wild type. Structural analysis of these intermediates helps to elucidate the function of genes and to gain more details of the biosynthetic pathway.

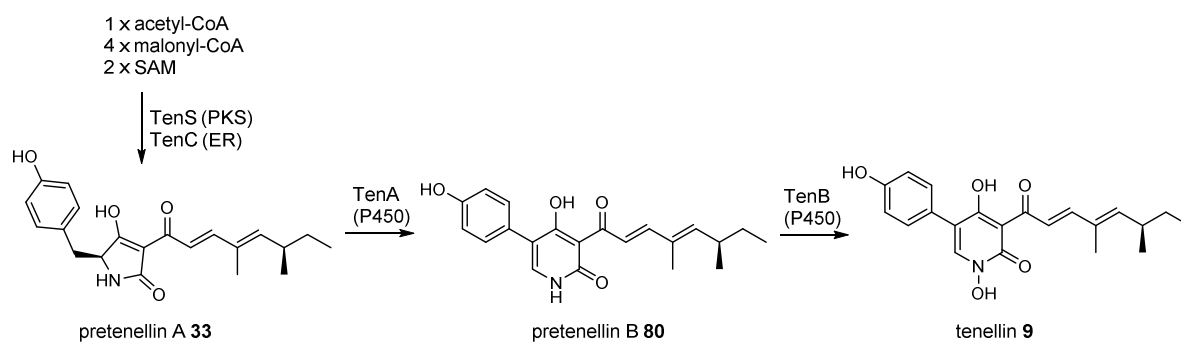
1.9.1 Targeted Gene Knockout in Filamentous Fungi

Provided that the BGC of interest is active in its native host and that the fungi is cultivable under laboratory conditions, targeted gene disruption is a useful tool to gain more insights into the biosynthetic pathway of the corresponding secondary metabolite.

To this end, the ability to perform targeted homologous recombination to delete specific genes is of fundamental importance. In the transformed deletion mutants, the target gene is replaced by a selection marker gene such as *hph* (resulting in hygromycin B resistance) and is therefore no longer expressed. At the same time this selection marker is used for screening of positive transformants. Gene deletion experiments have proven to be a useful tool for elucidation of pathways for many natural products such as chaetoglobosin A **51** and cytochalasin E **52**.^{76,110} Nevertheless, generating specific gene disruptions is a challenging task in filamentous fungi. DNA integration is mainly driven by non-homologous end joining (NHEJ) where the exogenous DNA is mainly integrated ectopically into the genome. Site-specific recombination can occur in less than 1 % of cases.¹¹¹

An alternative way to suppress gene expression is by gene silencing. Analogous to gene deletion experiments, exogenous DNA targeting the GOI is inserted into the organism producing the natural product. While gene deletions result in the disruption of the gene sequence, silencing methods are designed to lead to post-transcriptional mRNA degradation. Gene silencing is especially useful in targeting multi-copy or essential genes. The targeted gene is still present in these knock-down transformants, but the transcript level may be strongly or partially reduced.¹¹²

The biosynthesis of tenellin **9** in the filamentous fungus *Beauveria bassiana* was successfully elucidated by a combination of gene disruption and silencing methods. Targeted gene deletion of the gene *tenS*, encoding a PKS-NRPS, completely abolished production of **9**. Silencing of the two cytochrome P450 encoding genes *tenA* or *tenB* led to the accumulation of the pathway intermediates pretenellin A **33** and pretenellin B **80**, respectively. Thus, elucidation of the substrates and enzymes involved in the biosynthetic route to **9** was possible (Scheme 1.14).^{49,113}



Scheme 1.14: Biosynthesis of tenellin **9** in *Beauveria bassiana*.^{49,113}

1.9.1.1 Homologous Recombination in Fungi Using the Bipartite Method

Targeted gene replacement is one of the main strategies for functional characterization of fungal genes. A prominent strategy to disrupt genes of interest is based on stable integration of exogenous DNA using the bipartite method (also known as spilt-marker method). For this method, two constructs are required per transformation, each containing a flank of the target gene and around two thirds of a selectable marker cassette. In the case of fungi, the most commonly used selection marker is the *hph* gene (encoding a hygromycin B phosphotransferase), which confers resistance to hygromycin B. Three homologous recombination events occur, one between the overlapping regions of the selectable marker gene and one within each flanking region, resulting in a targeted gene deletion and replacement with a functional marker gene (Scheme 1.14).^{114,115} This is designed to dramatically reduce ectopic integration. Only two sequential PCRs are needed for the bipartite method. The first PCR amplifies the left and right flanking regions of the GOI and the hygromycin B selection cassette (including the *hph* gene, a constitutive promoter and a terminator sequence). 5' extensions are introduced with the respective primers, facilitating overlaps and homologous recombination of the flanks, the marker sequences and a yeast – *E. coli* shuttle vector. The three purified DNA fragments are assembled into the linearized vector in *S. cerevisiae* by homologous yeast recombination. The constructed plasmid is used as template for the second PCR to generate two DNA fragments (part 1 and part 2) for the bipartite method.

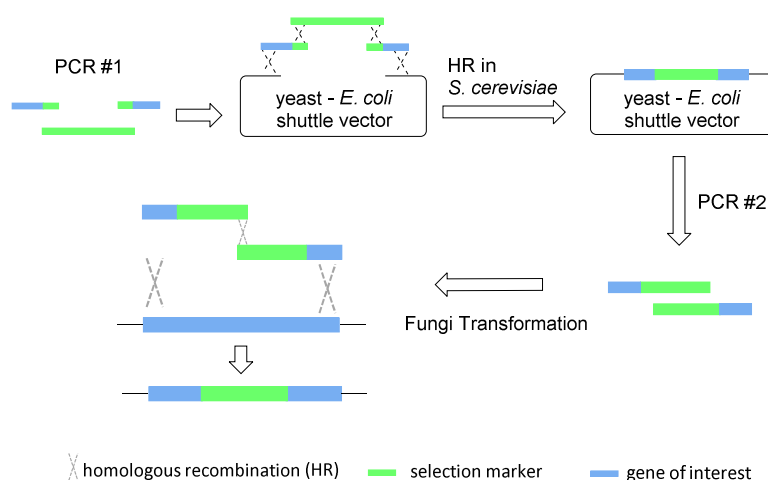


Figure 1.11: Overview of steps involved in gene disruption using the bipartite method.

A key advantage of this method is that neither fragment is able to confer resistance alone. Thus, the bipartite method significantly reduces ectopic integration of the deletion cassette by the non-homologous end joining (NHEJ) DNA repair mechanism.¹¹⁶ Ectopic integrations result in transformants which are able to grow under selective conditions, although the GOI is undisrupted (false positive transformants, Figure 1.12).

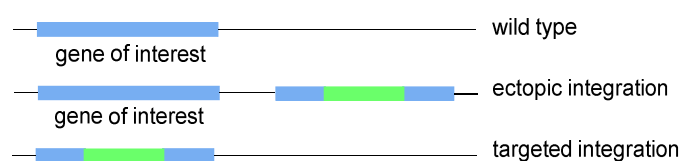


Figure 1.12: Integration of a deletion cassette by homologous recombination. Ectopic (random) integration events result in false positive transformants.

1.9.2 Complementation Studies in *M. grisea* Gene Deletion Strains

A potential aim of targeted gene deletions experiments when elucidating secondary metabolite synthesis is the loss of the final biosynthetic product and/or accumulation of biosynthetic intermediates or shunt products. In order to test if the changed secondary metabolite profile is really linked to the disrupted GOI, a functional copy of the gene can be re-introduced into the deletion strain, thereby complementing its function under control of an inducible or the natural promoter.

Furthermore, complementation studies can be used to test if a homologous gene from a different gene cluster is able to complement the deleted gene which can give insight into the substrate specificity of the gene product and the phylogenetic relationship of the respective enzymes.

Heterologous complementation might also lead to the production of new compounds and can be used for compound engineering.

1.10 Aim of the Thesis

The general aim of this work is to investigate the biosynthesis of the ACE1 metabolite which is responsible for avirulence signalling and to attempt to elucidate its structure. Previous studies and bioinformatic analysis of the *ACE1* BGC strongly suggest that the ACE1 metabolite belongs to the family of cytochalasans. Two genes of the cluster, *ORFZ* ($\alpha\beta$ -hydrolase) and *ORF3* (putative Diels-Alderase), are particularly interesting since they are highly expressed during production of the ACE1 metabolite and are likely to be involved in essential steps during the early and middle stage of biosynthesis. The main focus of this thesis will be on investigating the function of both enzymes.

Since the *ACE1* locus was previously demonstrated to have a reduced ability for homologous recombination and the ACE1 metabolite is only produced temporarily under very tight conditions, gene deletion experiments are unlikely to be useful.⁸⁸ However, the filamentous fungus *Magnaporthe grisea* produces the putative phytotoxin pyrichalasin H **50**, which belongs to the family of cytochalasans and which seems to be structurally and biosynthetically related to the ACE1 metabolite (Figure 1.13). Therefore, the **50** BGC will be used as model system for gene disruption experiments.

To elucidate the similarity of the ACE1 and pyrichalasin H **50** pathway, sequence similarities between the *ACE1* and the pyrichalsasin **50** BGC will be analysed using the ARTEMIS comparison tool. In addition, homologous protein sequences will be pairwise compared. In particular, the sequence similarity between the $\alpha\beta$ -hydrolases (*ORFZ* and *PyiE*) and the putative Diels-Alderases (*ORF3* and *PyiF*) will be determined.

Heterologous expression studies in *A. oryzae* will also be performed to elucidate the biosynthesis of the ACE1 metabolite. Preliminary results by Cox and co-workers showed that the heterologous co-expression of *ACE1* (PKS-NRPS) and *RAP1* (*trans*-acting enoyl reductase) led to the production of **67**. In addition to ACE1 and RAP1, the $\alpha\beta$ -hydrolase ORFZ and the putative DAase ORF3 are proposed to be involved in the formation of the ACE1 tricyclic core structure **81** (Figure 1.13). To elucidate whether **67** is recognized by ORFZ and/or ORF3 as a substrate or if co-expression of ORFZ and ORF3 might even prevent the formation of the alcohol **67**, will be part of this study. Moreover, expressing the whole *ACE1* BGC simultaneously will be performed with the aim to

elucidate the structure of the ACE1 metabolite and to gain more insights into its biosynthetic pathway.

Furthermore, the pathway-specific transcription factor BC2 of the *ACE1* BGC will be expressed under the control of an inducible promoter in *P. oryzae* Guy11 to test if the biosynthesis of the ACE1 metabolite can be stimulated by the presence of BC2.

The function of the $\alpha\beta$ -hydrolase will be further investigated. To this end, a *pyiE* *M. grisea* deletion strain will be constructed. Subsequently, homologous complementation studies with a functional copy of *pyiE* will be used as control. Furthermore, a point mutation within the catalytic site of *pyiE* will be designed and tested in further complementation studies. In addition, complementation studies with *ORFZ* from the *ACE1* BGC will be used to test if *ORFZ* is able to replace *PyiE*, thereby showing that the biosynthesis of the ACE1 metabolite and **50** is closely related. In addition, the function of *ORFZ* will be determined *in-vitro*. Heterologous production of *ORFZ* will be done in *E. coli*. As substrate for *ORFZ* a synthetic analogue will be used.

Analogous to the $\alpha\beta$ -hydrolase, the function of the putative Diels-Alderase (*PyiF* and/or *ORF3*) will be investigated *in-vivo* and *in-vitro*.

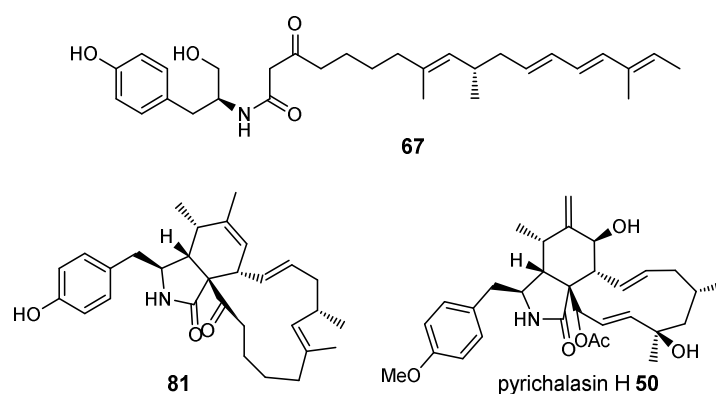


Figure 1.13: Chemical structure of **67**, **81** and pyrichalasin H **50**.

2. Bioinformatic Comparison of the *ACE1* and the Pyrichalasin H BGC

The rice blast fungus *P. oryzae* produces an avirulence secondary metabolite, the ACE1 compound. A fungus closely related to *P. oryzae*, *M. grisea* NI980, produces the cytochalasan pyrichalasin H **50**. The ACE1 compound and **50** appear to be structurally and biosynthetically related because they are both PKS-NRPS metabolites which use *O*-methyltyrosine **71** as the amino acid component (Section 1.8). Therefore their biosynthetic gene clusters (BGC) are also likely to be related. To this end, bioinformatic analysis and comparison of both BGC will be performed.

2.1 Aim of the Studies

The aim of this analysis is to elucidate if the **50** BGC and the biosynthetic pathway of **50** can be used as model system for the biosynthesis of the ACE1 metabolite. Since the *ACE1* locus was demonstrated to have a reduced ability for gene deletion by homologous recombination and the ACE1 metabolite is only produced temporarily under very tight conditions, gene deletion experiments have not proven possible.⁸⁸ However, the **50** BGC might be a good candidate for gene deletion experiments.

2.2 Results and Discussion

Sequence similarities between the *ACE1* and pyrichalasin H **50** BGC were analysed using the ARTEMIS comparison tool (Figure 2.1).¹¹⁷

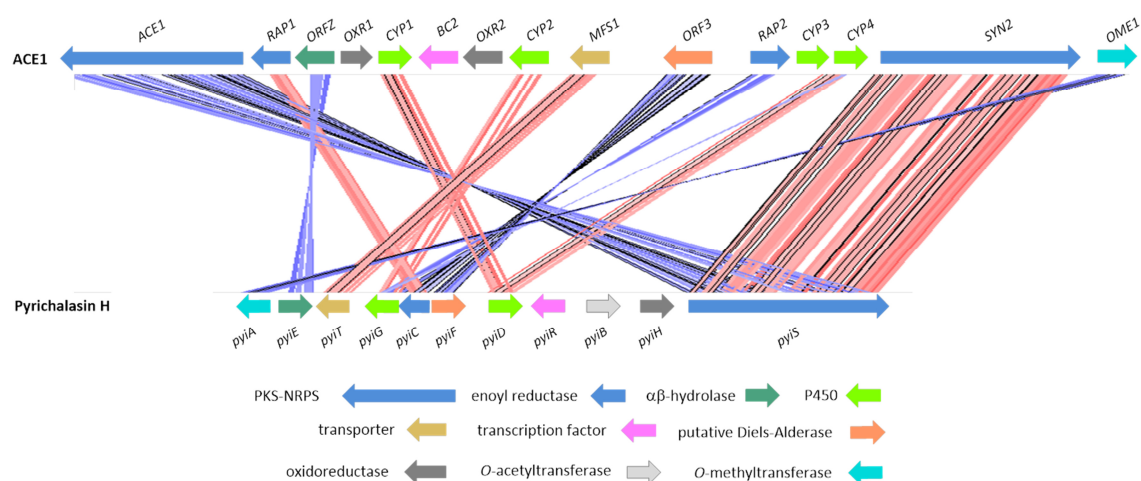


Figure 2.1: ARTEMIS comparison of the *ACE1* gene cluster from *P. oryzae* Guy 11 and the Pyrichalasin H **50** gene cluster from *M. grisea* NI980. Seven genes of the pyrichalasin H cluster exhibit high translated amino acid sequence similarity (46 - 64 % identity) with genes from the *ACE1* BGC: *OME1*, *ORFZ*, *MFS1*, *RAP2*, *ORF3*, *CYP4* and *SYN2*.

Homologous protein sequences were then pairwise compared with EMBOSS Needle (Table 2.2.1).¹¹⁸ Seven genes of the pyrichalasin BGC demonstrate high translated sequence similarity to genes from the *ACE1* BGC; they are homologous to *OME1*, *ORFZ*, *MFS1*, *RAP2*, *ORF3*, *CYP4* and *SYN2*. Since the core cluster genes of *ACE1* (homologs of *ACE1*, *RAP1*, *ORFZ* and *ORF3*) all have close homologs in the pyrichalasin BGC, the early and middle steps of biosynthesis are likely to be similar for **50** and the *ACE1* metabolite. Notably, the *pyiH* gene from the pyrichalasin BGC, encoding a NAD(P)H-dependant short-chain dehydrogenase (SDR), shares no significant sequence similarity to *OXR1* and/or *OXR2* from the *ACE1* BGC. The *OXR1* gene also encode a NAD(P)H-dependant SDR, the *OXR2* gene a FAD-dependant oxidase. Studies by our group revealed that during the biosynthesis of pyrichalasin H **50**, P*yiH* reduces the carbonyl at C-21. Subsequently, the *O*-acetyltransferase P*yiB* transfers an acetyl group to the formed hydroxyl group (Scheme 1.13). Since an *O*-acetyltransferase encoding gene is missing in the *ACE1* BGC, it is not surprising that the *ACE1* BGC also does not contain a gene homologous to *pyiH*. Notably, *ORF3* and P*yiF*, as well as *ORFZ* and P*yiE*, share a high sequence similarity. Notably, P*yiS* (PKS-NRPS) shares a slightly higher sequence similarity to *SYN2* than to *ACE1*. Also P*yiC* (enoyl reductase) shares a higher sequence similarity to *RAP2* than to *RAP1*. ORF by ORF comparison revealed that P*yiG* shares more sequence similarity to *CYP3* than to *CYP2* which was not recognized by the ARTEMIS comparison tool. One reason for that might be that bioinformatic predictions created a fusion of *pyiG* and *pyiC*. Division into two genes (*pyiG* and *pyiC*) was performed manually.

Table 2.2.1: ORF by ORF comparison of the *ACE1* and the pyrichalasin H gene clusters.

<i>Pyi</i> gene	<i>ACE1</i> gene	Putative function	Cofactor	aa % id. / sim.
<i>PyiS</i>	<i>ACE1</i>	PKS-NRPS		41.2 / 57.7
<i>PyiS</i>	<i>SYN2</i>	PKS-NRPS	NAD(P)H / SAM	49.0 / 64.8
<i>PyiC</i>	<i>RAP1</i>	enoyl reductase	NAD(P)H	41.7 / 62.0
<i>PyiC</i>	<i>RAP2</i>	enoyl reductase	NAD(P)H	57.5 / 74.0
<i>PyiF</i>	<i>ORF3</i>	putative Diels-Alderase	-	60.9 / 74.7
<i>PyiE</i>	<i>ORFZ</i>	$\alpha\beta$ -hydrolase	-	51.7 / 65.9
<i>PyiH</i>	<i>OXR1</i>	oxidoreductase/SDR	NAD(P)H	20.5 / 36.8
<i>PyiH</i>	<i>OXR2</i>	oxidoreductase OXR2: FAD/FMN-containing dehydrogenase	P <i>yiH</i> : NAD(P)H OXR2: FAD	7.7 / 10.5
<i>PyiD</i>	<i>CYP1</i>	cytochrome P450 monooxygenase	NAD(P)H	34.9 / 51.8
<i>PyiD</i>	<i>CYP2</i>	cytochrome P450 monooxygenase	NAD(P)H	21.5 / 36.7
<i>PyiD</i>	<i>CYP3</i>	cytochrome P450 monooxygenase	NAD(P)H	21.8 / 34.8
<i>PyiD</i>	<i>CYP4</i>	cytochrome P450 monooxygenase	NAD(P)H	49.4 / 69.0
<i>PyiG</i>	<i>CYP1</i>	cytochrome P450 monooxygenase	NAD(P)H	19.0 / 31.5
<i>PyiG</i>	<i>CYP2</i>	cytochrome P450 monooxygenase	NAD(P)H	34.9 / 51.2
<i>PyiG</i>	<i>CYP3</i>	cytochrome P450 monooxygenase	NAD(P)H	43.1 / 58.1
<i>PyiG</i>	<i>CYP4</i>	cytochrome P450 monooxygenase	NAD(P)H	22.7 / 37.7
<i>PyiR</i>	<i>BC2</i>	transcriptional regulator	-	15.7 / 23.3
<i>PyiT</i>	<i>MFS1</i>	MFS transporter	-	33.8 / 45.7
<i>PyiA</i>	<i>OME1</i>	<i>O</i> -methyltransferase	SAM	58.2 / 67.9
<i>PyiB</i>	-	<i>O</i> -acetyltransferase	-	-

Remarkably, P*yiR* and BC2 belong to different TF families and share only 15.7 % identity. BC2 is predicted to be a GAL4-like Zn₂Cys₆ binuclear cluster DNA-binding domain; while P*yiR* is predicted to be a C₂H₂ zinc-finger protein. BlastP searches were performed to find the 5 closest known relatives of BC2 and P*yiR*. A phylogenetic tree was created by neighbour-joining method using the Clustal Omega server.¹¹⁹ P*yiR*-like transcription factors (red in Figure 2.2) cluster more closely than the BC2-type transcription factors (black). One related protein belonged to the 5 closest hits of BC2 and P*yiR* (green). Ap*dR* involved in positive regulation of the aspyridone BGC in *Aspergillus nidulans* and Ccs*R* involved in positive regulation of the cytochalasin E BGC in *Aspergillus clavatus* are also shown (blue).^{120,121}

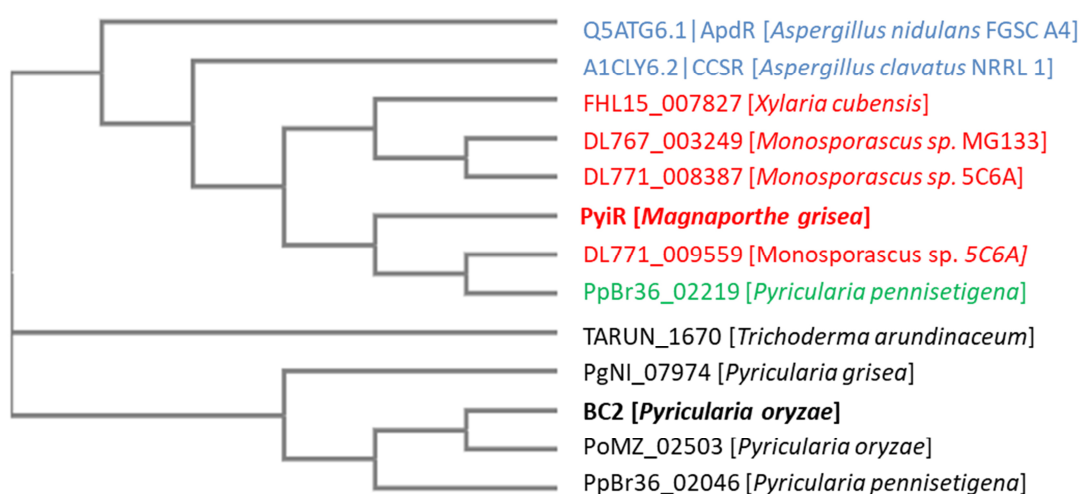


Figure 2.2: Phylogenetic relationship between BC2, P*yiR* and 9 homologous proteins identified by BlastP.

Despite minor deviations, the BGC of pyrichalasin H **50** is quite similar to the *ACE1* BGC. Therefore, it can be used as model system in this work to elucidate the function of the $\alpha\beta$ -hydrolase and the putative Diels-Alderase which will gain more insights into the biosynthetic pathway of cytochalasins.

3. Heterologous Expression of the *P. oryzae* ACE1 BGC in *A. oryzae*

Filamentous fungi have long been used for the discovery and production of biologically active and pharmaceutically useful compounds such as antibiotics.¹²² Since many biosynthetic gene clusters (BGC) are silent in their native strains under standard conditions or their respective expression levels are very low, heterologous expression of natural products has gained more interest in the last years. Beyond discovery of new compounds, heterologous expression allows the elucidation of biosynthetic pathways of microorganisms as well as genetic engineering by co-expressing genes from different pathways to generate novel analogues.

For the reconstitution of the desired biosynthetic pathway the genes can be amplified from genomic DNA (gDNA) or, if intron splicing might be a problem between distantly related species, from complementary DNA (cDNA). Gene transfer between fungal strains generally does not require codon optimization, since the codon usage is quite similar. To ensure high transcript levels it is crucial to clone each gene downstream of a strong promoter suitable for the host.

The tenellin 9 BGC, containing four genes, was one of the first to be successfully expressed in the arginine auxotrophic *A. oryzae* M-2-3 strain.¹²³ Since then, the biosynthesis of more complex secondary metabolites has been elucidated by heterologous expression studies, such as the meroterpenoid xenovulene A 82, where up to eleven genes have been expressed simultaneously (Figure 3.1).¹²⁴

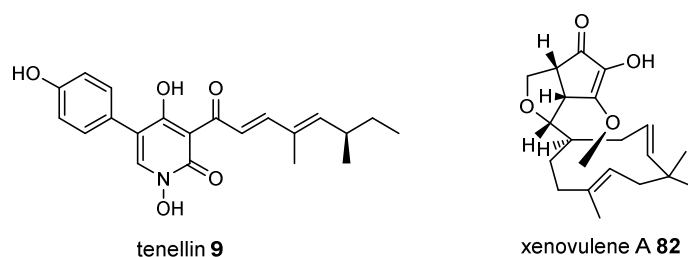


Figure 3.1: Tenellin 9 and xenovulene A 82.

3.1 Vector Construction by Homologous Recombination in *S. cerevisiae*

For large BGC the quadruple auxotrophic host *A. oryzae* NSAR1 is often used as a model host organism.¹²⁵ Rapid assembly of genes in one or a few multiple gene expression plasmids is performed using homologous recombination (HR) in yeast in order to reconstruct the whole biosynthetic pathway on a series of plasmids. The pTYGS expression vectors feature fungal selection markers for all four auxotrophies and two positive selection markers (*argB*, *adeA*, *Sc*,

niaD, *ble^R*, *bar^R*) as well as the 2μ and *colE1* origin of replication (ORI) to allow shuttling between *S. cerevisiae* and *E. coli* cells. For selection in *S. cerevisiae* the vector contains the *URA3* selection marker to allow growth in the absence of uracil. Transformant colonies contain vector DNA which can be isolated for bacterial transformation and screening. For selection in *E. coli* the vector facilitates resistance to chloramphenicol (*cam^R*) and ampicillin (*amp^R*, Figure 3.2).

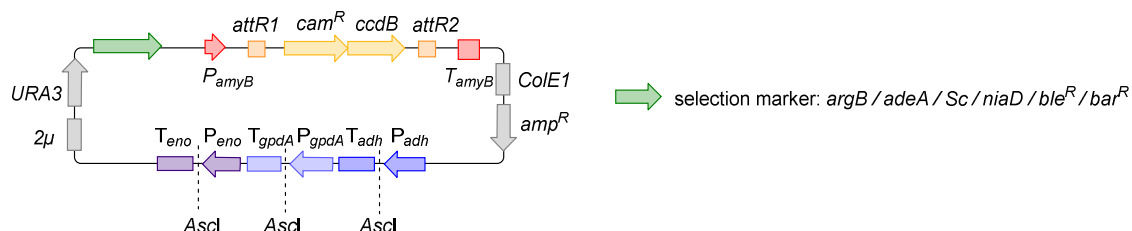


Figure 3.2: Schematic representation of the fungal expression system for *A. oryzae* NSAR1.

Furthermore, a Gateway cassette downstream of the strong inducible *amyB* promoter (P_{amyB}) is used for cloning of large fungal synthetase genes (such as PKS-NRPS) using Gateway[®]-mediated site-specific recombination *in-vitro* (Figure 3.4, B). The cloning site consists of a chloramphenicol resistance cassette (*cam^R*) and the *ccdB* gene, encoding the CcdB killer protein, which are both flanked by recombination sequences (*attR1* and *attR2*). The counterparts of the recombination sequences (*attL1* and *attL2*) are located on the entry clone pE-YA (*E. coli*-*S. cerevisiae* shuttle vector containing ORI (2μ , *pUC*) and selection marker (*URA3*, *kan^R*) for *E. coli* and *S. cerevisiae*), flanking the gene of interest (GOI). PE-YA is especially useful for assembling large genes (such as *ACE1*) from multiple PCR fragments by HR in *S. cerevisiae*. By using the LR recombinase kit (containing bacteriophage λ and integrase proteins: integrase, integration host factor and excisionase), together with the entry clone (PE-YA) containing the GOI and the destination vector (pTYGS family) two new plasmids are generated: the donor vector containing the gateway cassette (*cam^R* and *ccdB*) and the expression vector containing the GOI (Figure 3.3). Double selection of the expression clone is done by transforming the plasmid mixture into an *E. coli* TOP 10 strain, using ampicillin (for pTYGS) as selecting antibiotic. The strain is not able to survive if CcdB is produced; ensuring that only the final expression clone (pTYGS containing the GOI) is propagated.

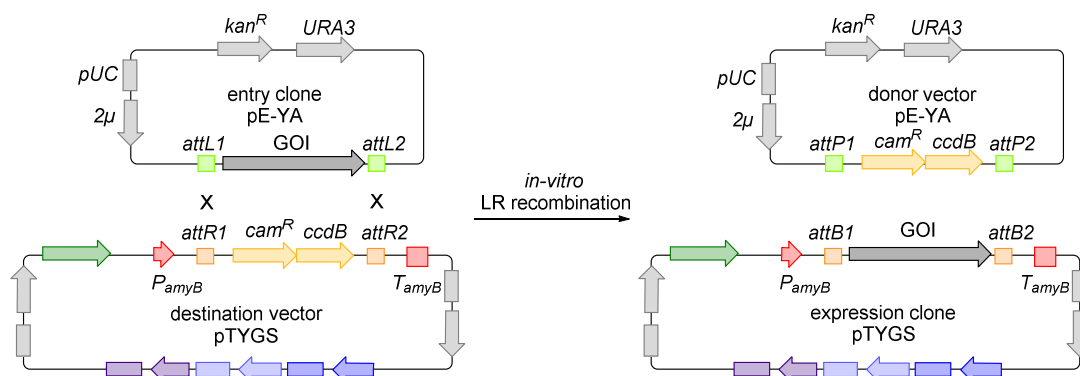


Figure 3.3: *In-vitro* (Gateway) cloning by recombination. Gateway-mediated site-specific LR recombination *in-vitro* between entry clone (pE-YA) and destination vector (pTYGS) to generate donor vector and expression clone containing the gene of interest (GOI).

In addition to the Gateway cassette, three fungal promoter/terminator (P/T_{adh} , P/T_{gpdA} and P/T_{eno}) gene cloning sites, each containing an *Ascl* restriction sequence, are included in the pTYGS vectors in order to allow *in vivo* homologous recombination by *S. cerevisiae* of up to three tailoring genes. Therefore, the vector is cut by *Ascl* into three parts and co-transformed with up to three purified PCR fragments (genes) into *S. cerevisiae*. 5' Extensions (at least 30 bp long) are included in the primers, facilitating overlaps and homologous recombination of the genes and the fungal expression vector (Figure 3.4).

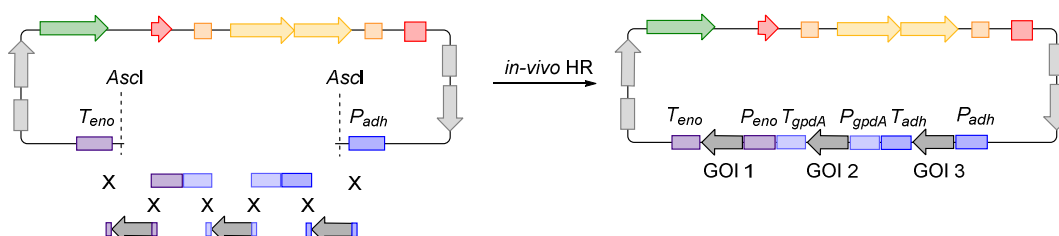


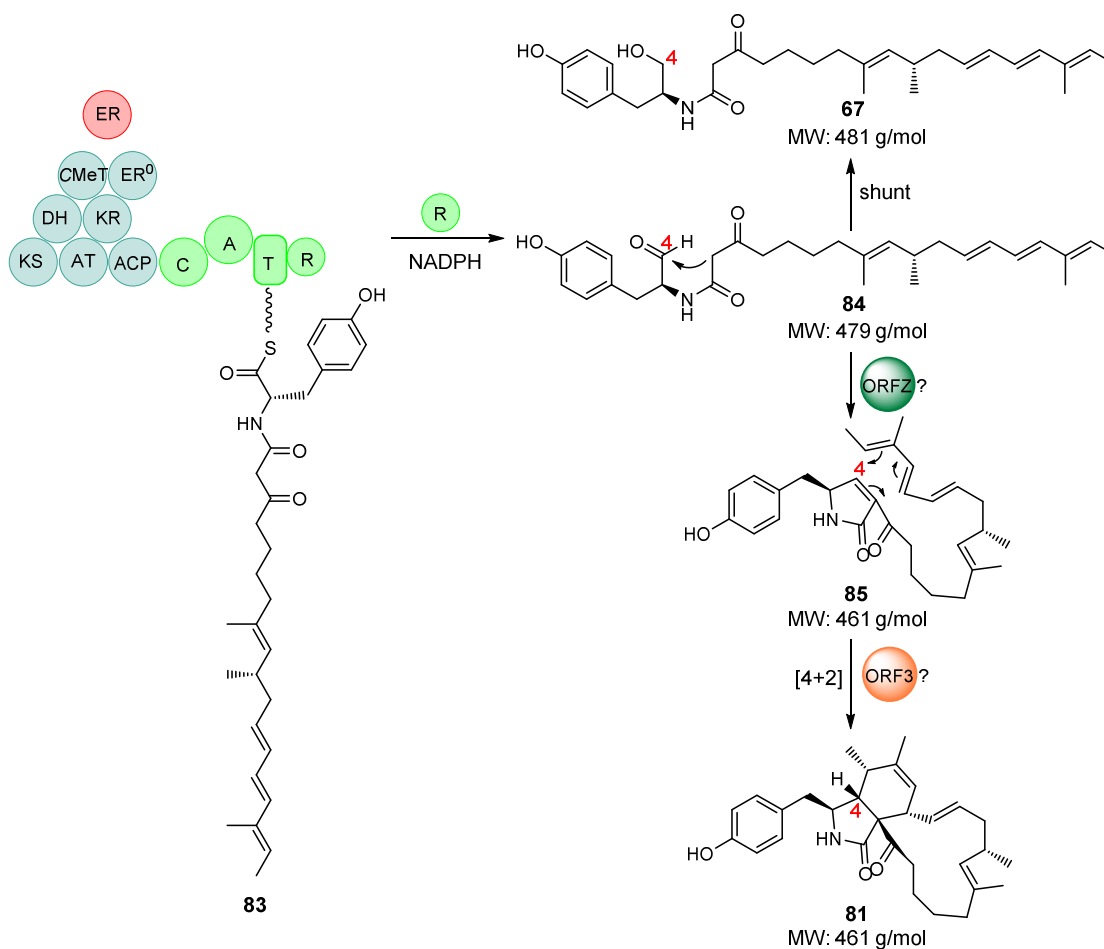
Figure 3.4: Homologous recombination. Up to three genes can be inserted into pTYGS by HR in *S. cerevisiae*. GOI = gene of interest.

3.2 Aim of the Heterologous Expression Studies

By using the fungal expression system the biosynthesis of the ACE1 metabolite will be investigated in *A. oryzae*. Preliminary results by Cox and co-workers showed that the heterologous co-expression of ACE1 (intron free), encoding a PKS-NRPS and RAP1 (intron free), encoding a *trans*-acting enoyl reductase, in *A. oryzae* M-2-3 led to the production of compound **67**, a trimethylated nonaketide fused to tyrosine (Scheme 3.1).⁹⁶ Compound **67** contains a C-4 hydroxyl group instead of the expected C-4 aldehyde. It is not clear if this reduction is caused by over-reduction by the ACE1 R-domain, or by a native *A. oryzae* enzyme. Notably, double reducing R-domains are known in other systems.^{126,127} However, Oikawa and co-workers also noticed that

expression of the PKS-NRPS CssA and its *trans*-ER CssC in *A. oryzae* also gives an aberrant alcohol product **63** (Scheme 1.10).⁹¹

In addition to ACE1 and RAP1, the $\alpha\beta$ -hydrolase ORFZ and the putative DAase ORF3 are proposed to be involved in the formation of the ACE1 tricyclic core structure **81**.^{128,51,78} ORFZ is believed to be involved in the formation of the pyrrolinone **85**, probably *via* a Knoevenagel condensation starting from aldehyde **84**. Subsequently, ORF3 might catalyse a [4+2] intramolecular cyclisation between the double bond of the pyrrolinone ring (dienophile) and the triene moiety (Scheme 3.1).



Scheme 3.1: Hypothetical early biosynthetic steps of the ACE1 metabolite.

To elucidate whether co-expression of ORFZ and ORF3 might prevent the formation of the alcohol **67**, was part of this study. The preconditions for this are that the aldehyde **84** is formed in *A. oryzae* but is subsequently reduced by a native *A. oryzae* enzyme to the alcohol **67**. If the aldehyde **84** has a long enough life-time, ORFZ and/or ORF3 might be able to drive the biosynthesis forward to the expected intermediates **85/81** before the reduction occurs.

Furthermore, expressing the whole ACE1 BGC simultaneously was also investigated with the aim to elucidate the structure of the final ACE1 metabolite and to gain more insights into its biosynthetic pathway.

3.3 Results

3.3.1 Early Steps of ACE1 Biosynthesis: Expression of the ACE1 Core Genes

3.3.1.1 Co-Expression of ACE1, RAP1

In order to further elucidate the early/middle steps of biosynthesis of the avirulence signalling compound, the quadruple auxotrophic (arginine, adenine, methionine and nitrate) host *A. oryzae* NSAR1 was used instead of the *A. oryzae* M-2-3. In previous work, an *A. oryzae* NSAR1 transformant expressing the genes ACE1 and RAP1 (*A. oryzae* ACE1 + RAP1) was already constructed using the expression vector pTYGSargB-ACE1-RAP1 and provided for further studies (Dr. Elizabeth J. Skellam, Leibniz Universität Hannover, Institute of Organic Chemistry, AK Cox). The vector pTYGSargB-ACE1-RAP1 contains the ACE1 gene (intron free) under the control of P_{amyB} and the RAP1 gene under the control of P_{adh} .

The strain *A. oryzae* ACE1 + RAP1 was cultivated for 3-4 days in DPY media (28 °C, 110 rpm) before cells and growth media were individually extracted twice with ethyl acetate and analysed by LCMS (Section 9.5). Production of several (at least 8) non-polar compounds were observed in the organic cell extracts (Figure 3.6 B). These compounds were not produced by untransformed *A. oryzae* NSAR1. Based on their characteristic UV absorption spectra between 250–300 nm, which was already observed for alcohol **67** (Figure 3.5), all compounds appeared to contain the polyketide triene moiety (indicated by red star, Figure 3.6 B). Unfortunately, only compound **67**, which was also observed in *A. oryzae* M-2-3 ACE1 + RAP1, as well as compound **86** (structure was elucidated by Dr. Elizabeth Skellam, unpublished data) could be characterised (Figure 3.6). The expected PKS-NRPS hybrid **84** and/or the pyrrolinone **85** could not be observed ($m/z = 479$ g/mol and 461 g/mol). Since all compounds are only produced in small amounts (about 1 mg/L) and most of them co-eluted with fatty acids, it was not possible to identify further compounds.

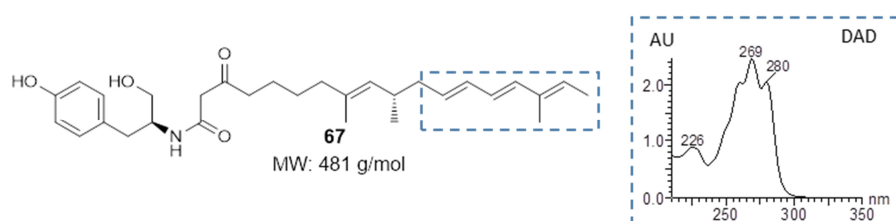


Figure 3.5: Characteristic UV absorption spectra for the triene moiety of the linear PKS-NRPS backbone.

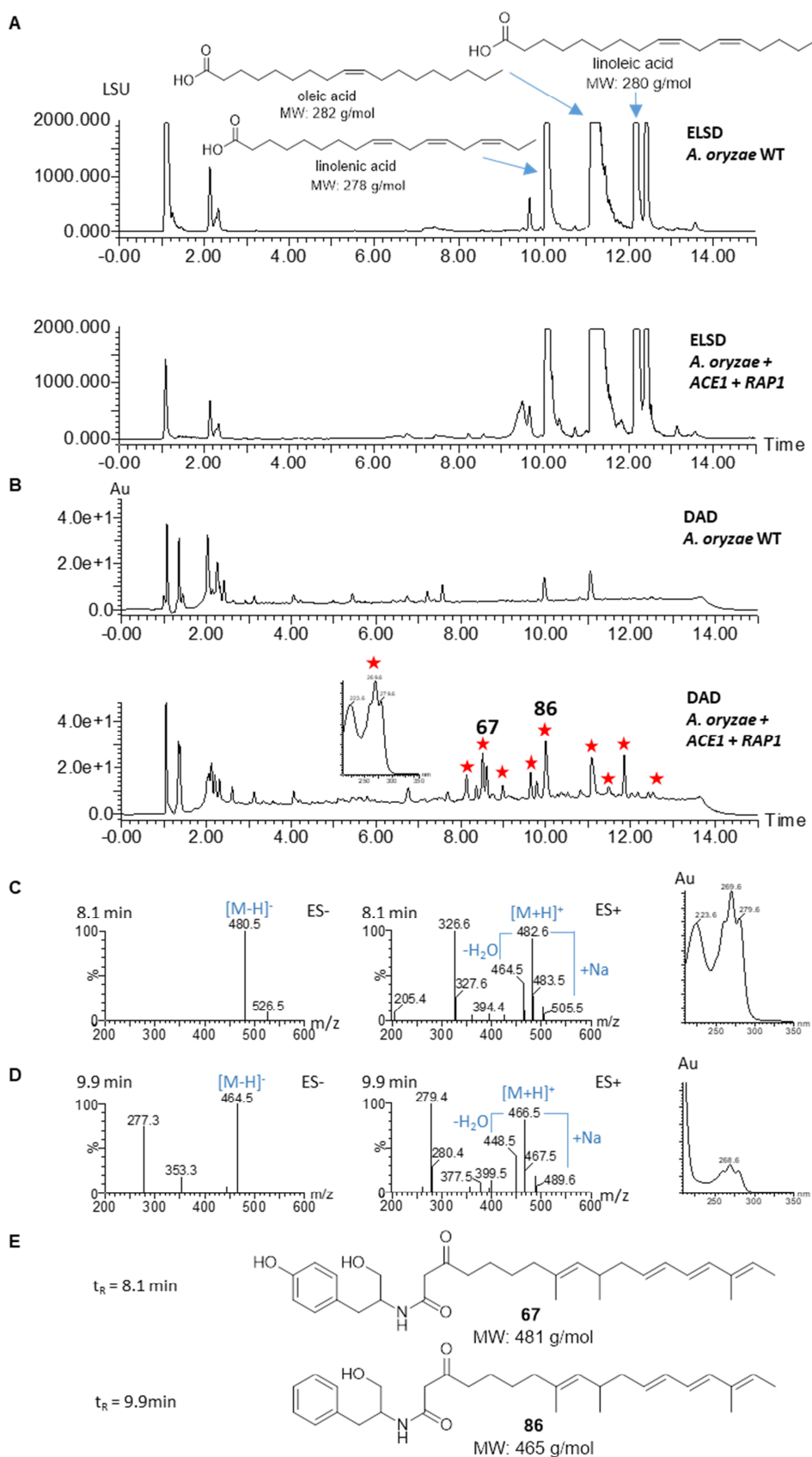


Figure 3.6: Organic extracts of *A. oryzae* NSAR1 wild type and *A. oryzae* + ACE1 + RAP1: **A**, Evaporative Light Scattering Detector (ELSD); **B**, Diode array detector (DAD), Red star: Peaks corresponding to compounds with characteristic spiky UV spectrum (triene structure). ES and UV spectrum of compound **C**, **67**; **D**, **86**; **E**, structure of compound **67** and **86**.

3.3.2 Co-Expression of ACE1, RAP1, ORF3 and ORFZ

Co-expression of ACE1 and RAP1 led to the production of **67** and **86**, as well as several related unidentified compounds. Neither the aldehyde **84** ($m/z = 479$ g/mol), nor the pyrrolinone **85** ($m/z = 461$ g/mol), was observed and also the tricyclic core structure **81** of the putative ACE1 metabolite was not formed. One aim of this work was to investigate if the $\alpha\beta$ -hydrolase ORFZ and the putative Diels-Alderase ORF3 are involved in the early/middle steps of biosynthesis - the formation of the core structure **81**. To this end, ORFZ and ORF3 were cloned together and individually (only ORF3) into the fungal expression vector pTYGSade and transformed into *A. oryzae* ACE1 + Rap1. A strain co-expressing ACE1, RAP1 and ORFZ was already generated in previous studies and provided for this work (Dr. Elizabeth J. Skellam, Leibniz Universität Hannover, Institute of Organic Chemistry, AK Cox). Two vectors pTYGSade-ORFZ-ORF3 and pTYGSade-ORF3 were generated (Section 9.4.3, Figure 3.7).

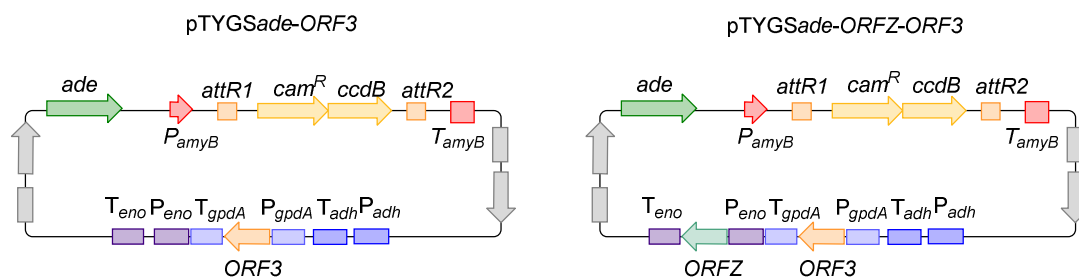


Figure 3.7: Fungal expression vector pTYGSade-ORFZ-ORF3 and pTYGSade-ORF3.

Transformation of *A. oryzae* ACE1 + RAP1 was done with the constructed plasmids using a PEG-mediated protoplast protocol. Fifteen transformants were generated for each construct. In order to confirm that the genes were integrated successfully, gDNA of transformants was extracted and analysed by PCR for the presence of ORFZ and ORF3 (Figure 3.8 A). Transformants VBI03-2-1, 2-4, 2-6, 2-16, 2-17, 2-20 and 2-23 contained both genes. Transformants VBI03-3, 6, 8, 9 and 18 contained only ORF3 (Figure 3.8 B). Correct expression of ORFZ and ORF3 was confirmed for transformant VBI03-2-4 by reverse transcriptase (RT) PCR (Section 9.4.4). Both genes were successfully expressed (Figure 3.8 C).

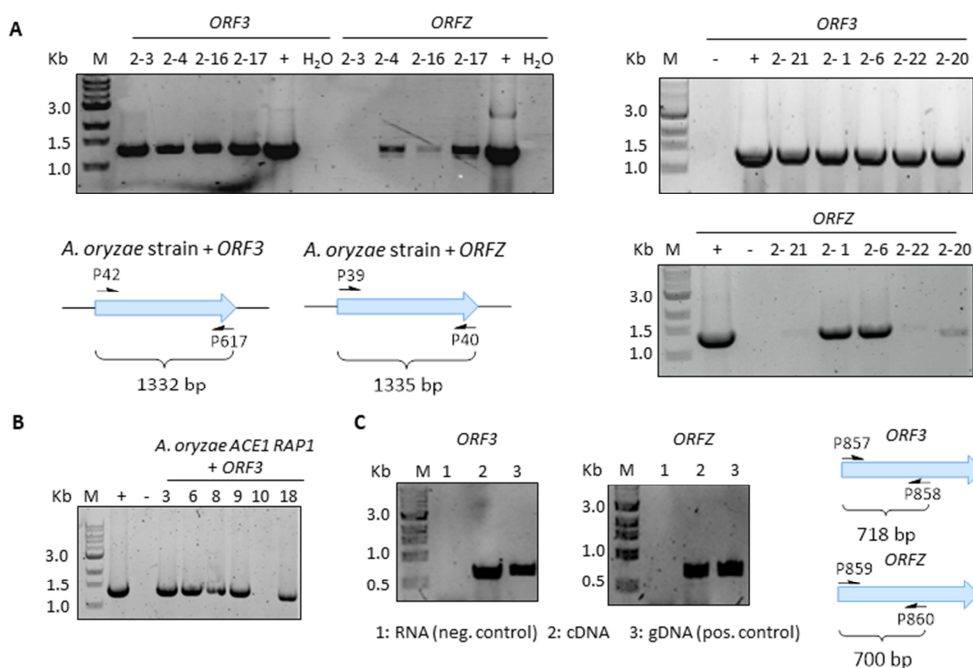


Figure 3.8: Genetic analysis of *A. oryzae* transformants: **A**, *A. oryzae* ACE1 + RAP1 + ORF3+ ORFZ; H₂O as negative control and gDNA as positive control (+); **B**, *A. oryzae* ACE1 + RAP1 + ORF3; **C**, Reverse transcriptase (RT) PCR analysis of transformant VBI03-2-4.

The strain *A. oryzae* ACE1 + RAP1 (control) and the transformants were cultivated for 3-4 days in DPY media (28 °C, 110 rpm) before cells and growth media were individually extracted with ethyl acetate and analysed by LCMS (Section 9.5). Only the organic cell extract exhibited heterologous production of secondary metabolites. Strains expressing *ORF3* in addition to *ACE1* and *RAP1* showed the same secondary metabolite profile as *A. oryzae* ACE1 + RAP1. Notably, co-expression of *ORFZ* led to the production of further shunt products (Figure 3.14, green box). Unfortunately, due to low production levels, isolation and structure elucidation by NMR was not possible. However, formation of the expected pyrrolinones **85** and **87** ($m/z = 461$ g/mol and 445 g/mol) could not be observed by LCMS or HRMS (Figure 3.9).

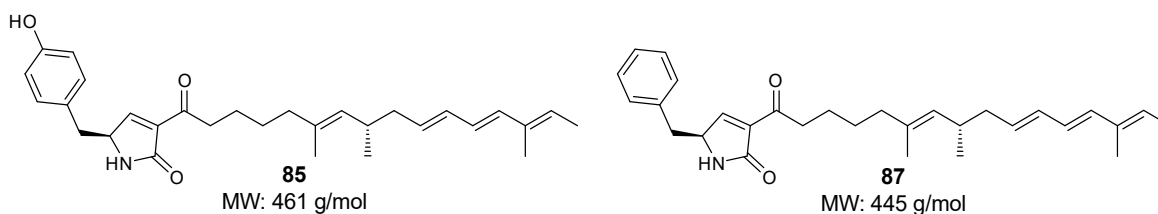


Figure 3.9: Expected but not formed pyrrolinones **85** and **87**.

Co-expression of *ORFZ* and *ORF3* with *ACE1* + *RAP1* showed the same secondary metabolite profile as co-expression of *ACE1* + *RAP1* + *ORFZ* alone. All produced compounds displayed the same characteristic UV spectrum as **67** and **86**, indicating that the polyketide triene moiety is

present. Therefore, the DA reaction did not occur. Compounds **67** and **86** were present in all organic extracts; they were not accepted as substrates by ORF3 and/or ORFZ. According to the hypothetical early biosynthetic steps of the biosynthesis of ACE1, both compounds should contain an aldehyde instead of a C-4 hydroxyl group (Scheme 3.1). Without the aldehyde the proposed Knoevenagel condensation cannot take place. Subsequently, without the formation of the pyrrolinone **85** the intramolecular DA reaction cannot occur. The biosynthesis is stalled at this stage.

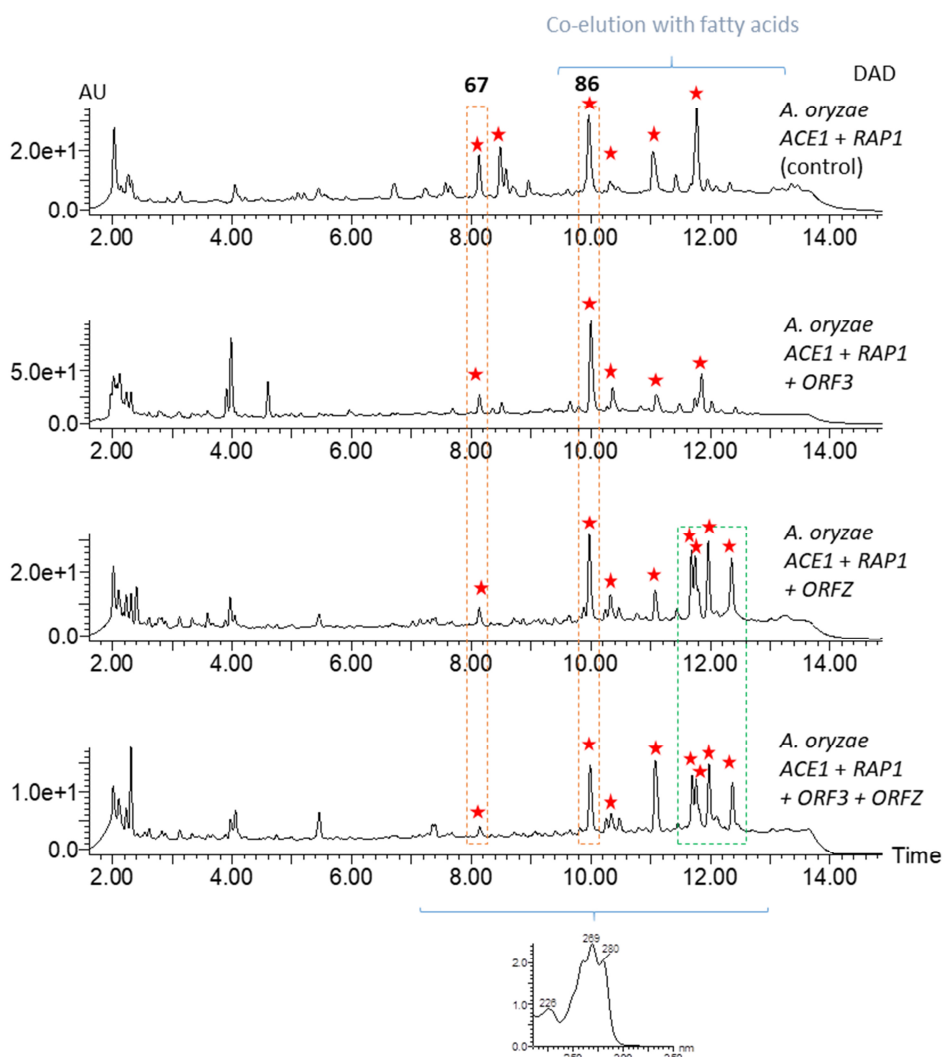


Figure 3.10: Organic extracts of *A. oryzae* transformants expressing the core gene cluster of ACE1. Red star: Peaks corresponding to compounds with characteristic spiky UV spectrum (triene structure), orange box: compound **67** and **86**, green box: shunt products observed in *A. oryzae* strains expressing ACE1 + RAP1 + ORFZ.

3.3.3 Feeding Studies with *O*-Methyl-L-Tyrosine

Gene deletion experiments of genes from the pyrichalasin H **50** BGC from *M. grisea* NI980 revealed that *O*-methyltyrosine **71** is likely to be the real substrate for the NRPS module of the PKS-NRPS PyiS (Chapter 7). By disrupting *pyiF* in *M. grisea* NI980 (encoding the putative Diels-

Alderase), early biosynthetic intermediates were isolated which contained *O*-methyltyrosine **71**, suggesting that the methyl group is transferred to tyrosine at a very early stage. Furthermore, disruption of *pyiA* (encoding an *O*-methyltransferase) abolished production of **50** (done by Chongqing Wang, Leibniz Universität Hannover, AK Cox).¹⁰⁸ An *O*-methyltransferase, encoded by *OMET*, is also present in the ACE1 BGC (56 % amino acid identity to *PyiA*), indicating that *O*-methyltyrosine **71** is probably also needed for efficient biosynthesis of the ACE1 metabolite. Therefore, commercial *O*-methyl-L-tyrosine **71** (Alfa Aesar) was fed to *A. oryzae* cultures co-expressing the core genes *ACE1*, *RAP1*, *ORF3* and *ORFZ*. A new main compound with a *m/z* of 495 g/mol was produced in all transformants. Isolation and structure elucidation by NMR and HRMS allowed the identification of this compound as the *O*-methyltyrosine nonaketide **88** (structure published in Wang, Hantke *et al*, 2019, Organic Letters).¹⁰⁸ Compared to the previously produced compounds **67** and **86**, the yield of **88** was significantly higher (2 mg/L). **67** and **86** were not produced anymore.

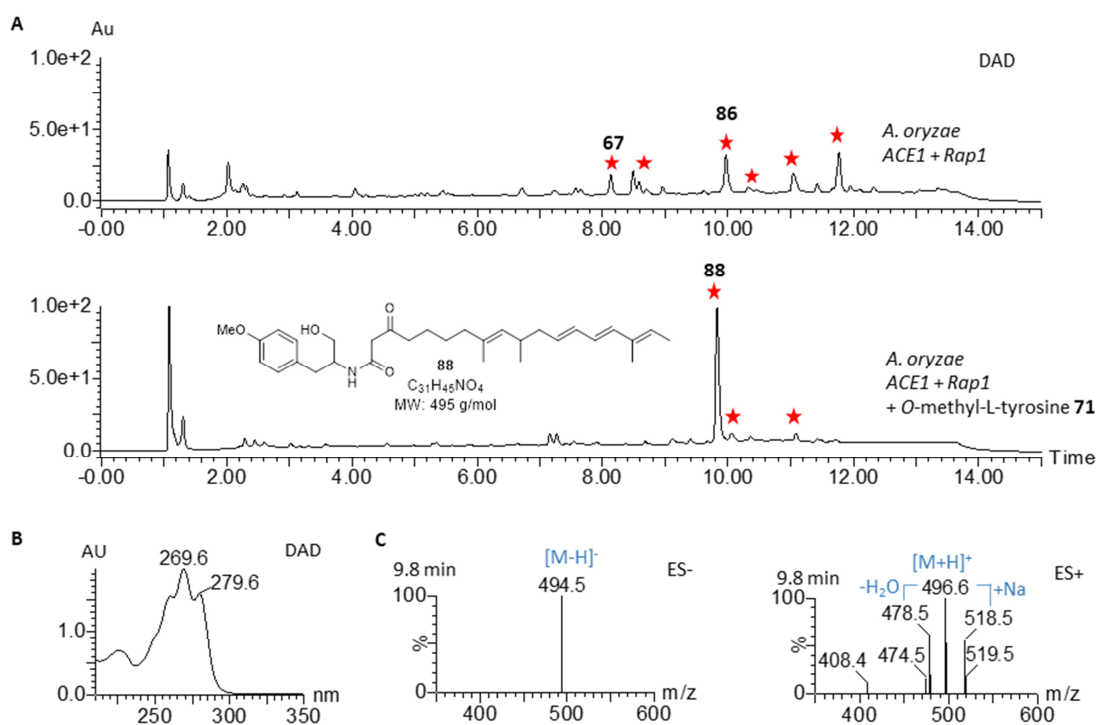


Figure 3.11: Feeding *O*-methyl-L-tyrosine **71** to *A. oryzae* ACE1 + RAP1 led to the formation of **88**: **A**, Organic extracts of *A. oryzae* ACE1 + RAP1 and *A. oryzae* ACE1 + RAP1 fed with *O*-methyl-L-tyrosine **71**; Diode array detector (DAD); **B**, UV absorption spectra of **88**; **C**, ES of compound **88**. Red star: Peaks corresponding to compounds with characteristic spiky UV spectrum (triene structure).

Co-expression of *ORFZ* and *ORF3* with *ACE1* + *RAP1* and co-feeding **71** again resulted in the production of further non-polar shunt products (Figure 3.12 A). Due to higher production levels it was possible to purify small amounts (0.5 mg) of a compound with *m/z* of 477 g/mol and to partially elucidate its structure by NMR (Figure 3.14 D, Table 3.3.1).

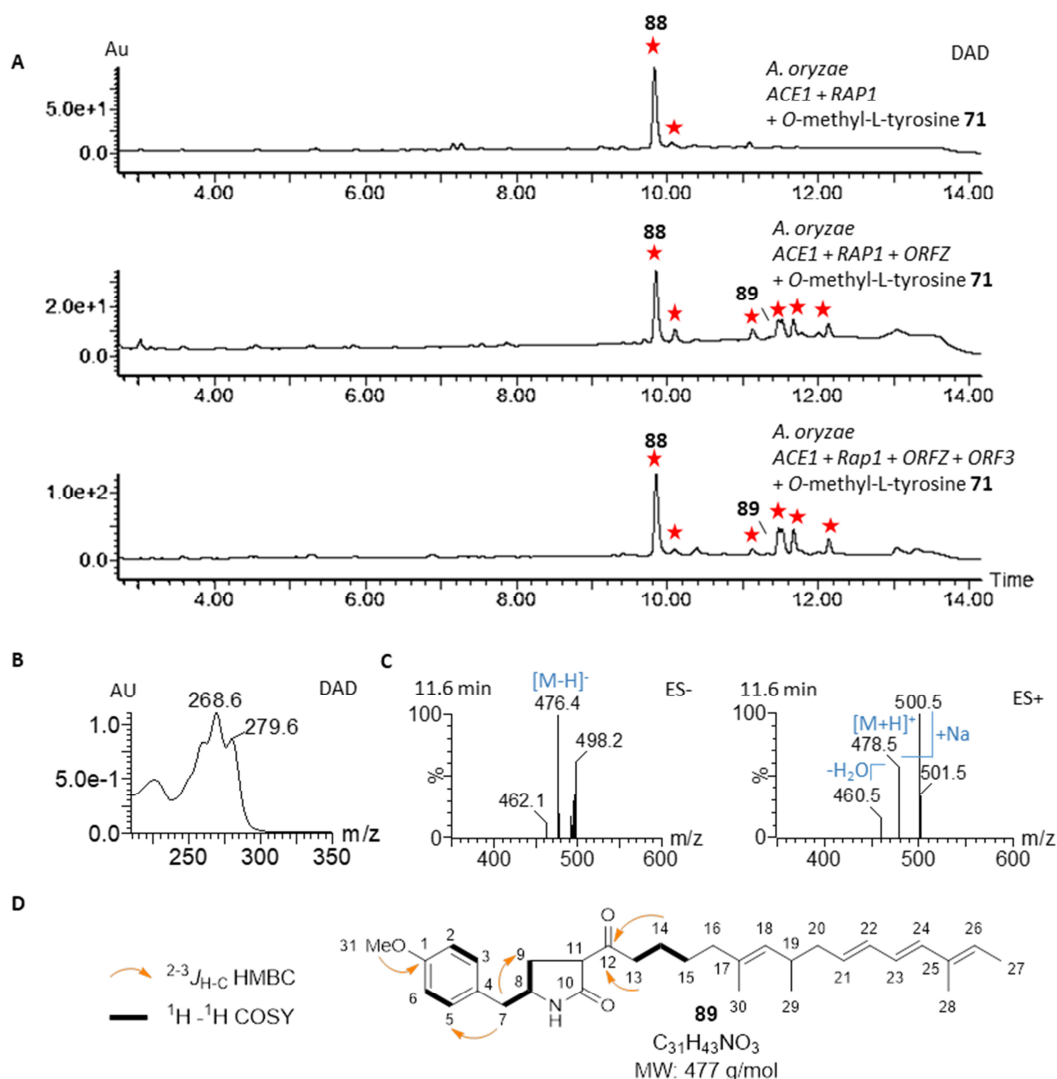


Figure 3.12: Feeding studies with *O*-methyl-L-tyrosine **71**: **A**, LCMS chromatograms (DAD trace) of organic extracts of *A. oryzae* ACE1 RAP1 and *A. oryzae* ACE1 RAP1 ORF3 ORFZ; **B**, UV-spectrum of **89**; **C**, ES spectrum of **89**;

The structure of this new compound would be consistent with the structure of compound **89**. NMR data confirmed that the compound derived from *O*-methyl-L-tyrosine **71** (Figure 3.13 A). Based on the ¹H-¹H COSY experiment signals for *O*-methyltyrosine **71** were assigned (at 6.86, 6.86; 6.86, 7.14; 7.14, 6.86 and 7.14, 7.14 ppm). In addition, a characteristic HMBC signal between the protons of the methyl group (H-31, singlet at 3.8 ppm) and the C-1 carbon of the benzoyl ring (159.5 ppm) was observed. Furthermore, HMBC signals indicated that the compound contains a pyrrolidine instead of the expected pyrrolinone (Figure 3.13 B). Protons of C-7 correlate to carbons of the aromatic ring (C-3/ C-5, 131.3 ppm) and the pyrrolidine ring (C-8 and C-9, 55.6 ppm and 28.0 ppm). One part of the polyketide side chain could also be identified. Based on the HMBC experiment, H-13 correlates to one carbonyl group (C-12 at 177.8 ppm), and two methylene groups (C-14 at 25.8 ppm and C-15 at 30.2 ppm, Figure 3.13 C).

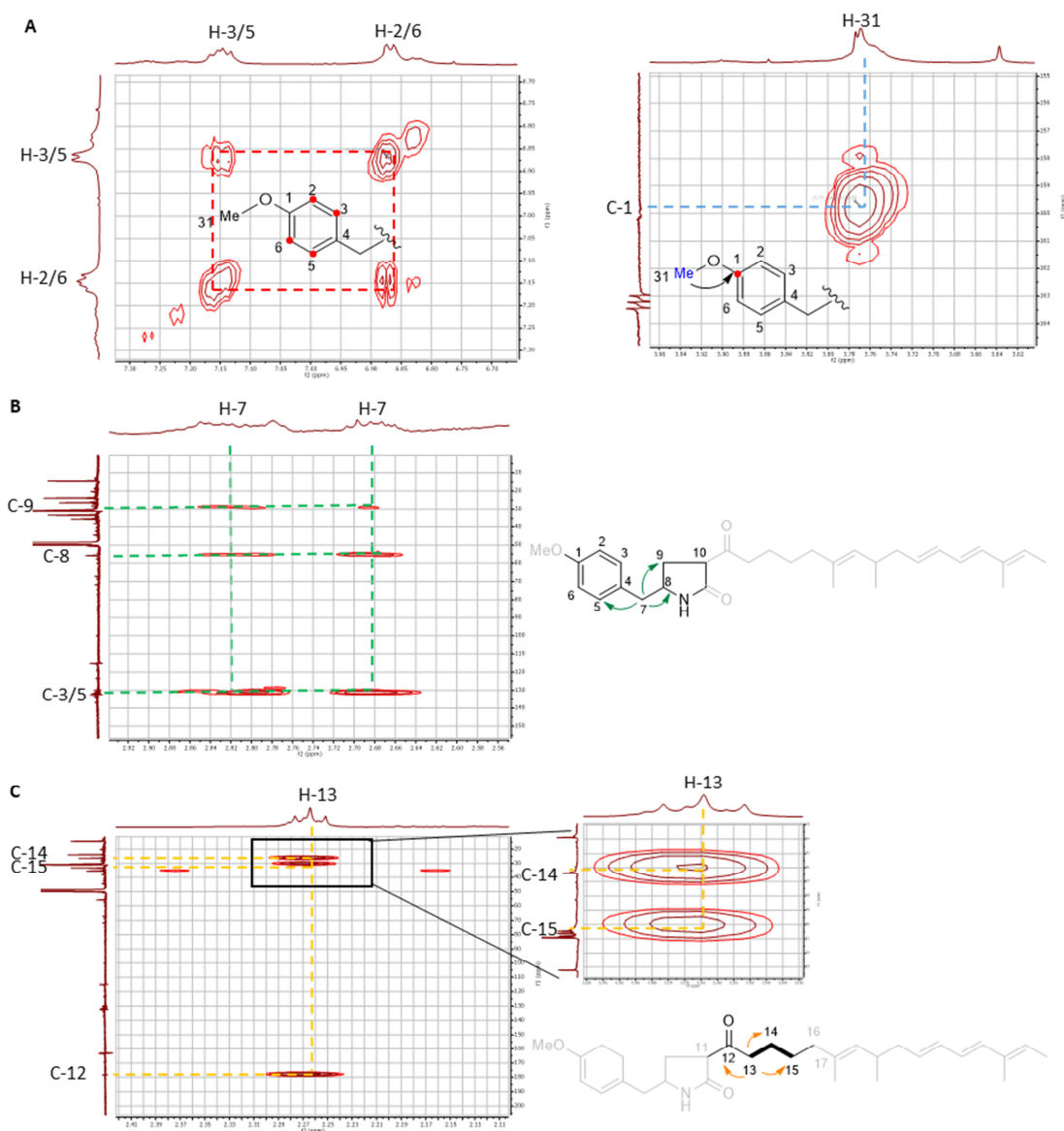


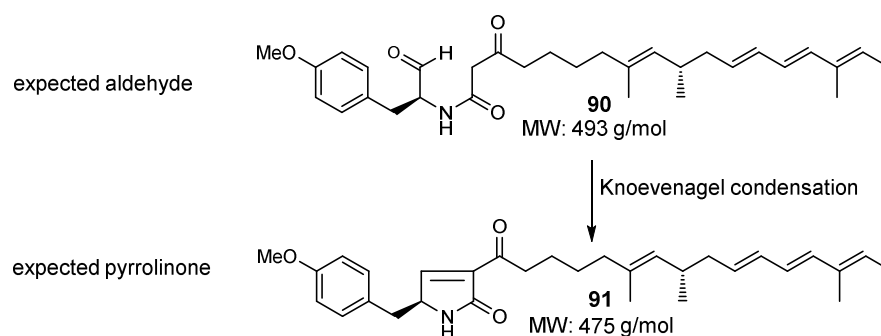
Figure 3.13: Characteristic ^1H - ^1H COSY and HMBC signals of compound **89**. **A**, ^1H - ^1H COSY of the aromatic protons; **B**, HMBC collation of H-7; **C**, HMBC correlation of H-13.

Unfortunately, further NMR signals could not be assigned to their corresponding atom. Although signals corresponding to methyl groups are present in the ^1H and ^{13}C -NMR spectrum, it was not possible to confidently assign them to the side chain methyl groups of compound **89** (C-27 to C-30). However, the characteristic UV spectrum is further evidence for the presence of a triene moiety, again indicating no Diels-Alder reaction, even though this part of the structure could not be confirmed by NMR due to weak data.

Table 3.3.1: Partial NMR data of **89**. Not all ^1H signals could be integrated due to overlap. Solvent: CD_3CN .

Atom	$\delta(^{13}\text{C})$ ppm	$\delta(^1\text{H})$ ppm	Multiplicity (J, Hz)	COSY	HMBC	HSQC
1	159.5	-				
2/6	114.6	6.86	m	3	1, 4	CH
3/5	131.3	7.14	m	2	1, 7	CH
4	130.1	-				
7	42.0	2.68 2.80	m m	8	3, 8, 9	CH_2
8	55.6	3.90	m			CH
9	28.0	2.12 2.15	m m	8	7, 8	CH_2
10	N/A					
11	N/A					
12	177.8	-				
13	35.4	2.27	t (7.41, 7.41)	14	12, 14, 15	CH
14	25.8	1.60	m		12, 13, 15	CH_2
15	30.2	1.32	m	14		CH_2
16-30	N/A					
31	55.8	-	s		1	CH_3

HRMS confirmed a molecular formula of $\text{C}_{31}\text{H}_{43}\text{NO}_3$ ($[\text{M}]\text{H}^+$ calculated 478.3323, found 478.3321). Based on partial NMR data, HRMS and UV absorption the isolated compound would be consistent with the structure of **89**. However, formation of the expected pyrrolinone **91** was not confirmed by HRMS (Figure 3.14).

**Figure 3.14:** Proposed biosynthetic intermediates of the ACE1 metabolite. Compounds were not observed.

3.3.4 Simultaneous Expression of 11 Genes of the ACE1 BGC

Heterologous expression of the core genes of the ACE1 BGC (*ACE1*, *RAP1*, *ORF3* and *ORFZ*) in *A. oryzae* NSAR1 led mainly to the formation of **67** and **86** or, if the culture was fed with *O*-methyl-L-tyrosine **71**, to the formation of **88** and **89**. These are presumably shunt products of the real PKS-NRPS intermediate. To test if the biosynthesis of the ACE1 metabolite is further catalysed by the other enzymes encoded within the ACE1 BGC, a cloning strategy was designed to transform eleven genes (*ACE1*, *RAP1*, *ORF3*, *ORFZ*, *CYP1*, *CYP2*, *CYP3*, *CYP4*, *OXR1*, *OXR2* and *OME1*) simultaneously into *A. oryzae* NSAR1 (Figure 3.15). SYN2 and RAP2 were already shown to be dispensable for the biosynthesis of the ACE1 metabolite.⁴⁵ Further, the transcription factor BC2

and the transporter MSF1 were considered to be important only for regulatory purposes within the original host and therefore non-essential for the actual biosynthetic product. For simplicity, all genes were amplified from *P. oryzae* Guy 11 gDNA and cloned directly into the fungal expression vectors pTYGS*met*, pTYGS*ade* and pTYGS*niaD* (Figure 3.15, Section 9.4.3).

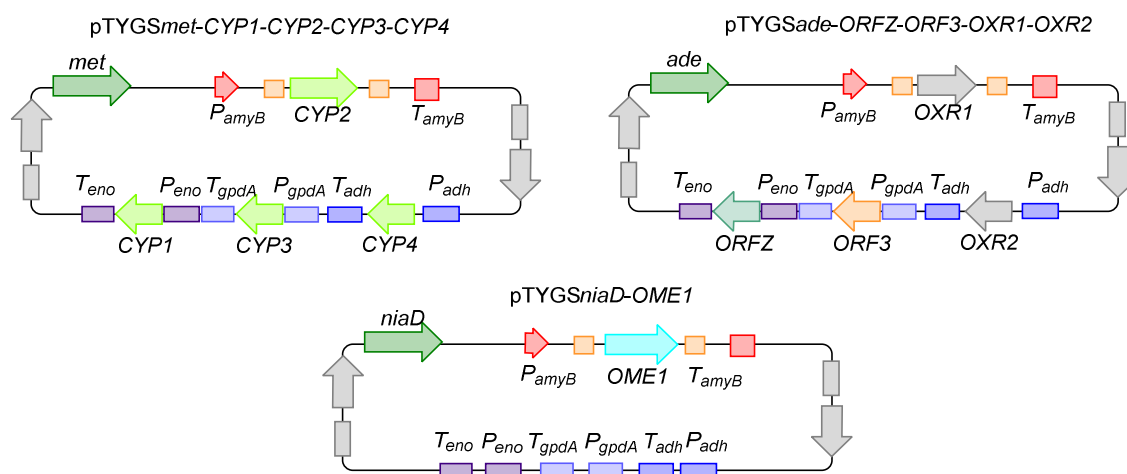


Figure 3.15: Constructed fungal expression vectors.

Transformation of *A. oryzae* NSAR1 was performed with the constructed plasmids using a PEG-mediated protoplast protocol. Since *niaD* complementation was observed to be a poor selection marker, pTYGS*niaD-OMET* was co-transformed with the other three plasmids using an excess of pTYGS*niaD-OMET* (ratio 3:1) and a selection medium without methionine, arginine and adenine. Fifteen transformants were generated and further analysed genetically by PCR and chemically by LCMS (Section 9.4 and 9.5). Integration of all genes was only confirmed for three transformants (VBI23-8, VBI23-9 and VBI23-12, Table 3.3.2).

Table 3.3.2: Genes confirmed by PCR in each transformant.

ID	RAP1	ACE1	ORF3	ORFZ	OXR1	OXR2	CYP1	CYP2	CYP3	CYP4	OME1
2	x	x	x	x	x	x	x		x	x	
3	x	x	x	x	x	x					x
4											
5	x		x	x					x		x
6	x	x		x	x						x
7	x	x							x	x	
8	x	x	x	x	x	x	x	x	x	x	x
9	x	x	x	x	x	x	x	x	x	x	x
10	x	x	x	x	x	x	x	x	x	x	
11	x	x							x	x	x
12	x	x	x	x	x	x	x	x	x	x	x
13	x		x	x	x	x					x
14	x		x	x	x	x			x	x	

Transformants VBI23-8, 9 and 12 were cultivated for 3-4 days in DPY media (28 °C, 110 rpm) before cells and growth media were individually extracted with ethyl acetate and analysed by LCMS (Section 9.5). Only the organic cell extract exhibited heterologous production of secondary metabolites. The main compound produced by transformant VBI23-8 and VBI23-12 was compound **88**, which was previously observed when *ACE1* was co-expressed with *RAP1* and *O*-methyl-L-tyrosine **71** was fed to the growth medium. Methylation of tyrosine to *O*-methyltyrosine **71** must have been catalysed by the *O*-methyltransferase *OME1*. To this end, *OME1* is part of the *ACE1* BGC. Also compound **89** could be observed in organic extract of VBI23-8. The organic extract of VBI23-9 contained compound **67** and **86**, indicating that expression of *OME1* was not successful.

Overall, the same secondary metabolite profile was observed during co-expression of *ACE1* and *RAP1* alone. Formation of **89** might be catalysed by *ORFZ*. Nevertheless, all observed compounds were not accepted as substrates by the putative DAase *ORF3* or other tailoring enzymes of the *ACE1* BGC. These findings strongly suggest that the observed compounds are shunt products.

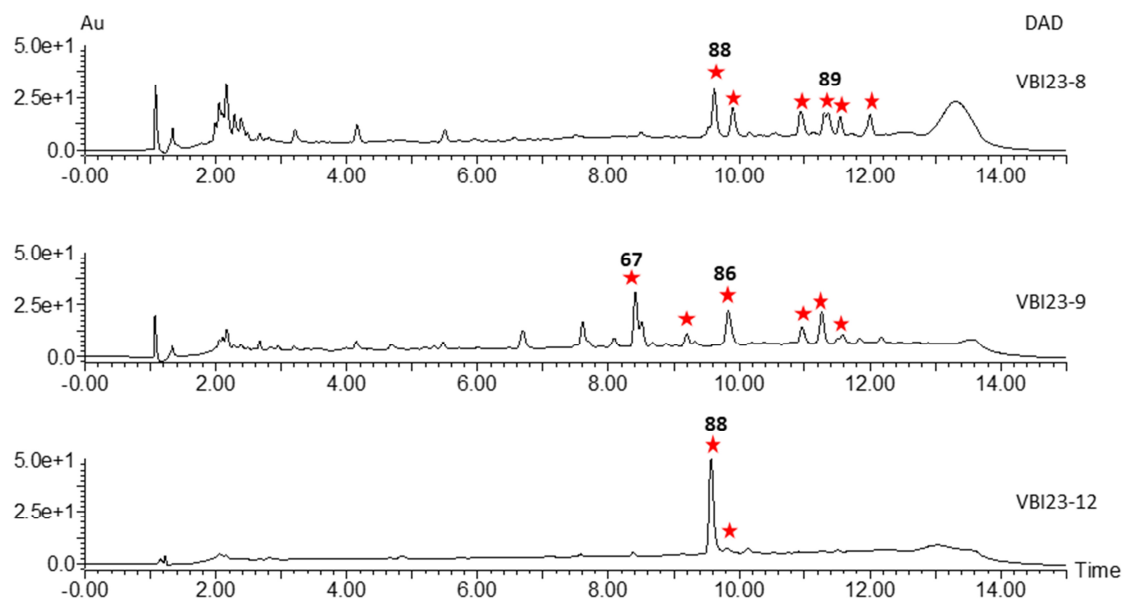


Figure 3.16: Co-expressing 11 genes of the *ACE1* BGC. LCMS chromatograms (DAD trace) of organic extracts of VBI23-8, VBI23-9 and VBI23-12. Red star: Peaks corresponding to compounds with characteristic spiky UV spectrum (triene structure).

3.4 Discussion

The successful heterologous expression of partial or complete biosynthetic pathways in *A. oryzae* has been demonstrated for several fungal natural products like citrinin **146** and tenellin **9**.^{129,123} Heterologous expression of *ACE1* and *RAP1* in *A. oryzae* led to the formation of several new compounds which all shared a UV absorption spectra characteristic for a triene moiety. Notably,

only organic extracts of the cells contained these compounds. Therefore, low production levels and accumulation of co-eluting fatty acids prevented the straightforward elucidation of their respective structures. Nevertheless, two compounds could be identified as **67** and **86**. Production of **67** was previously observed in similar experiments using *A. oryzae* M-2-3 as heterologous host.⁹⁶

Co-expression of the putative DAase, encoded by *ORF3*, did not affect the secondary metabolite profile. On the other hand co-expression of the $\alpha\beta$ -hydrolase, encoded by *ORFZ*, led to production of additional non-polar shunt products containing a triene moiety. Nevertheless, yields were too low to further analyse these compounds. The expected pyrrolinone **85** could not be detected by HRMS.

Co-expression of *ORFZ* and *ORF3* resulted in the same secondary metabolite profile in comparison to the expression of *ORFZ* alone. Reverse transcriptase (RT) PCR confirmed the transcription of *ORFZ* and *ORF3* in *A. oryzae*, suggesting that **67** and **86** are not the native substrate for these enzymes.

Feeding *O*-methyl-L-tyrosine **71** to a culture of *A. oryzae* ACE1 + *RAP1* led to the production of **88** in significant higher yields compared to the previous produced compounds **67** and **86** (2 mg/L). Hence, *O*-methyl-L-tyrosine **71** is preferably accepted as a substrate for the A-domain of the ACE1 NRPS. During biosynthesis of the ACE1 metabolite the hydroxyl group of tyrosine is probably methylated by the *O*-methyltransferase OME1, which is encoded within the ACE1 BGC.

Unlike in other known PKS-NRPS systems, the amino acid carbonyl of **88** (as well as **67** and **86**) has been reduced to an alcohol rather than undergoing the more usual Dieckmann cyclization to form a tetramic acid (biosynthesis of tenellin **9**, Scheme 1.6) or to undergo a Knoevenagel condensation after reductive release, forming the pyrrolinone **91** as proposed for the biosynthesis of fusarin C **8** (Scheme 1.8).^{58,50} The C-terminal R-domain of the ACE1 NRPS might be responsible for the observed double reduction.

Indicating otherwise, formation of the pyrrolinone **91** proceeding from the unreactive alcohol **88** is not possible. More likely, reduction of the aldehyde **90** is caused by a native *A. oryzae* enzyme, probably an oxidoreductase. Oikawa and co-workers also observed the alcohol **63** when they expressed the cytochalasin PKS–NRPS *ccsA* and its *trans*-acting ER *ccsC* in *A. oryzae*. Thus, **63** is probably a shunt metabolite from the cytochalasin pathway (Figure 3.17).⁹¹

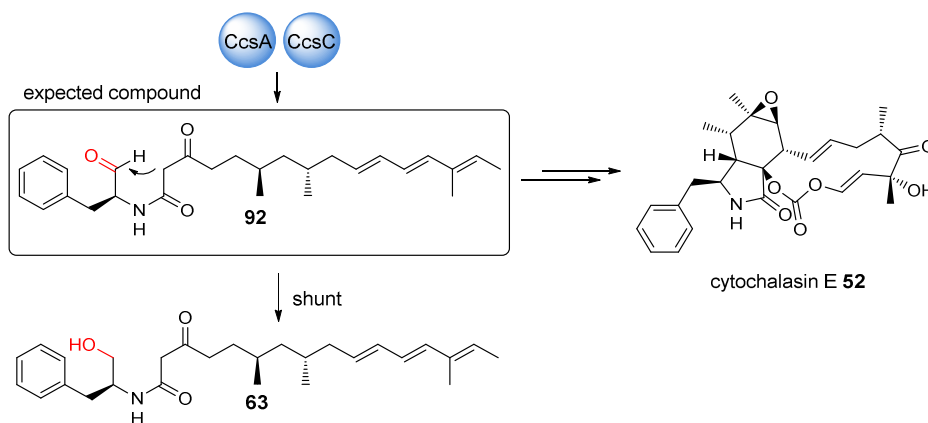
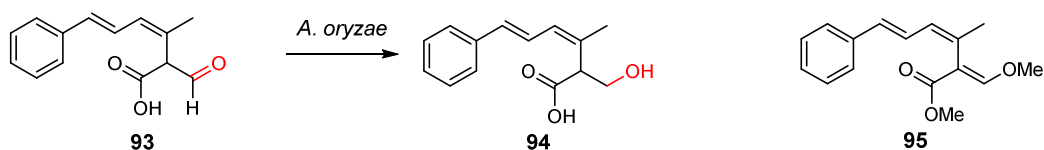


Figure 3.17: Isolated PKS-NRPS compound **63** from *A. oryzae* by Oikawa.⁹¹

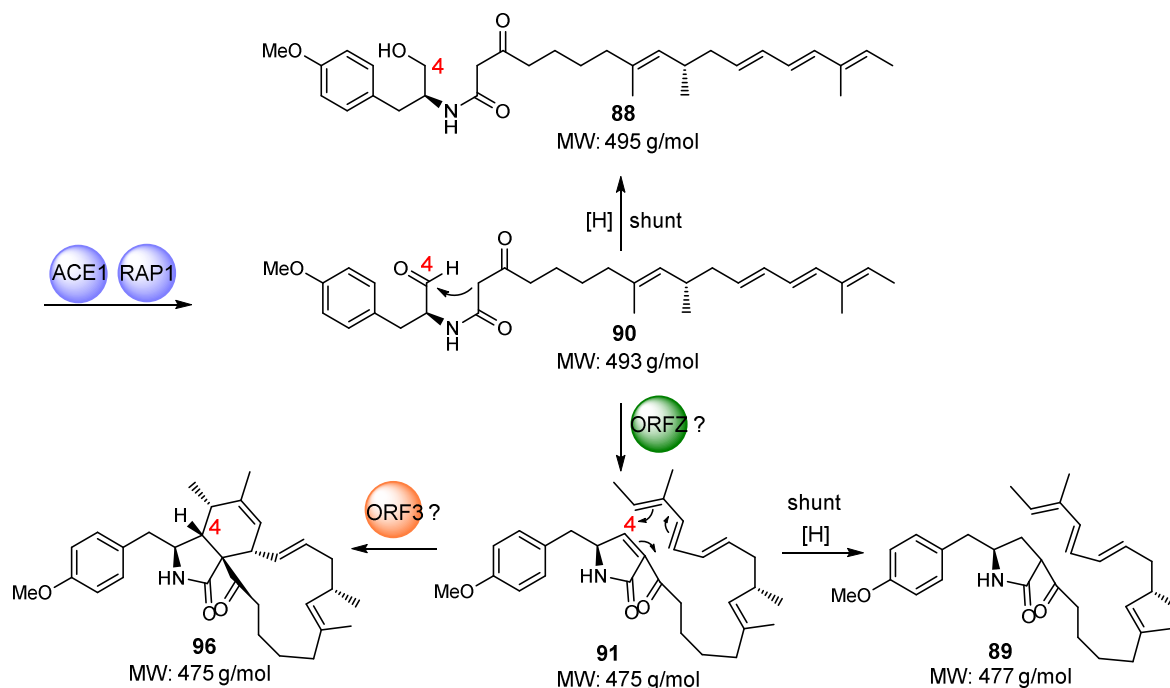
Notably, various shunt pathways that contained such a reduction of aldehydes have already been reported in *A. oryzae*. Intermediates of the citrinin **146** biosynthetic pathway containing highly reactive aldehydes are reductively shunted to numerous byproducts, including azaphilones and their aminated products in the absence of the final pathway reduction catalysed by CitE.¹²⁹ Another example is the biosynthesis of strobilurin **95**, where the reduction of aldehyde **93** to **94** is also apparently catalysed by a native *A. oryzae* enzyme (Scheme 3.2).¹³⁰



Scheme 3.2: Within the pathway to strobilurin **95**, the aldehyde **93** is shunted to **94** by a native *A. oryzae* enzyme.

Co-expression of *ORFZ* together with *ACE1* and *RAP1* (+ *O*-methyl-L-tyrosine **71**) also led mainly to the production of **88**. In addition, small amounts of a new compound were observed whose structure is consistent with **89**. In the proposed biosynthetic pathway of the ACE1 metabolite, the $\alpha\beta$ -hydrolase *ORFZ* is predicted to be involved in the formation of the pyrrolinone ring, which would lead to compound **91** (Scheme 3.3). Compound **89** is closely related to **91** and might have been formed by a reductive step. Notably, compounds similar to **89** were identified in Δ *pyiF* *M. grisea* deletion mutants (Chapter 7), providing more evidence that **89** might be the correct structure of the observed compound ($m/z = 477$ g/mol). However, at this point it remains unclear whether **89** might be a shunt product of **91** formed by a native *A. oryzae* enzyme, or if it is a true pathway intermediate. If *ORFZ* is involved in the formation of **91** or **89** also remains elusive.

Co-expression of the putative DAase ORF3 did not alter the metabolite profile, indicating that **89** is not the correct substrate. Since the *ORF3* gene does not contain introns, it is unlikely that it is transcribed incorrectly in *A. oryzae*, suggesting that ORF3 is active. Regardless of these observations, the bottleneck for constitutive production of the ACE1 core structure **81** is the formation of the alcohol **88** (or **86** and **67**). Resolving this issue should eventually result in higher yields of **91** by ORFZ, which might subsequently be used as a substrate for ORF3, before **91** is finally converted into **89** (Scheme 3.3)



Scheme 3.3: Proposed middle-stage biosynthetic pathway of the ACE1 metabolite.

Another approach to overcome the stalled biosynthesis of the ACE1 metabolite was the co-expression of the whole ACE1 BGC (besides *SYN2*, *RAP2*, *MSF1* and *BC2*) in *A. oryzae* NSAR1. This was based on the hypothesis that lack of one of the tailoring enzymes might hinder the formation of **88**. However, no significant changes were observed in the metabolite profiles of transformed strains expressing almost the whole BGC versus transformed strains expressing ACE1, RAP1 and ORFZ alone. Compound **88** was still the main produced compound, strongly suggesting that **88** is a shunt product.

O-Methyl-L-tyrosine **71** was demonstrated to be the preferred substrate of the ACE1 NRPS A-domain. Co-expression of ORFZ led to the formation of a new compound **89** which might be a shunt product of the expected pyrrolinone **91**. This is the first indication that ORFZ is involved in the formation of the pyrrolinone **91**. Probably the ACE1 PKS-NRPS backbone is released from the

R-domain as aldehyde **90**. Subsequently, a native *A. oryzae* enzyme probably efficiently catalyses its reduction to the alcohol **88**. However, small amounts of the released aldehyde **90** are probably converted to **91** by ORFZ. Reduction of the pyrrolinone double bond would lead to the formation of **89**. This step might also be catalysed by a native *A. oryzae* enzyme. Whether both reductive steps are catalysed by the same *A. oryzae* enzyme, or if two enzymes are involved in these steps is not clear. However, a fast reduction in *A. oryzae* reduces the aldehyde **90** and also the formed pyrrolinone **91** before the DA reaction can occur - the biosynthesis of the ACE1 metabolite cannot proceed.

To this end, *A. oryzae* is an unsuitable host for the heterologous production of the ACE1 metabolite due to cross-chemical reactions. One way to overcome this problem would be to identify the enzyme responsible for the reduction of the aldehyde **90** and to disrupt its function. For this purpose, an *in-vitro* screening system was established (Chapter 4).

Remarkably, not all *Aspergillus* hosts exhibit the observed reduction activity. The responsible enzyme might therefore be unique to *A. oryzae* and might be involved in secondary metabolism and not in the essential primary metabolism. Larson and co-workers recently discovered that heterologous expression of *ccsA* (PKS-NRPS) and *ccsC* (*trans*-acting ER) from the cytochalasin E **52** and K **53** BGC in *A. nidulans* led to production of niduclavin **62** (Scheme 1.10).⁹⁰ Compound **62** contained a decalin ring and a nitrogen-containing tetramic acid unit and was built of a highly reduced polyketide chain fused to phenylalanine. Notably, niduclavin **62** consists of the same underlying polyketide chain length and amino acid as compound **63**, which was identified by Oikawa and co-workers when *cssA* and *cssC* were co-expressed in *A. oryzae* (Scheme 1.10).⁹¹ Reduction of the aminoaldehyde **92** to alcohol **63** did not occur in *A. nidulans*; instead a tetramic acid was formed.

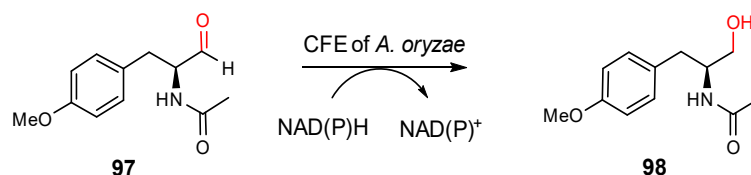
Nevertheless, more insights into the early/middle biosynthetic steps of the biosynthesis of the ACE1 metabolite were obtained. *O*-Methyl-L-tyrosine **71** was identified as the preferred substrate for the A-domain of the ACE1 NRPS. The *O*-methyltransferase OME1 appears to transfer the methyl group to the hydroxyl group of tyrosine at an early stage of biosynthesis. Furthermore, first indications were obtained indicating that the $\alpha\beta$ -hydrolase ORFZ might be involved in the formation of the pyrrolinone ring. Unfortunately, the putative DAase ORF3 did not exhibit any activity by performing the expected ring closing reaction *in-vivo*. Based on the fact that the expected substrate of ORF3, the pyrrolinone **91**, was not observed in any organic extracts, this result is not surprising.

4. In-vitro Screening for an Oxidoreductase from *A. oryzae*

Aspergillus oryzae is often used as host for the heterologous expression of single genes or whole biosynthetic gene cluster (BGC). In some cases, in addition to the expected metabolite, shunt by-products are also produced. This has been observed for the heterologous expression of the citrinin **146** and the strobilurin **95** BGC.^{129,130} The formation of these shunt products results from cross-chemical reactions catalysed by endogenous *Aspergillus* enzymes.

Heterologous expression of the *ACE1* BGC led mainly to the formation of alcohol **88**, which is most likely a shunt product of aldehyde **90** (Chapter 3). A native *A. oryzae* enzyme is believed to be responsible for this undesirable reduction step.

To test whether any native host enzyme is responsible for the formation of the shunt alcohol **88**, an *in-vitro* enzyme assays was designed containing the cell free extract (CFE) of *A. oryzae* NSAR1, a substrate analogue **97** (provided by Haili Zhang and Franck Siakeu, AK Cox Leibniz Universität Hannover) and the cofactors NADH or NADPH. Conversion of **97** to **98** would prove the presence of an oxidoreductase which reduces aldehydes related to **90** and **97** (Scheme 4.1).



Scheme 4.1: Proposed enzymatic conversion of **97** to **98**.

4.1 Results

To obtain CFE containing the protein of interest a culture of *A. oryzae* expressing *ACE1* and *RAP1* was grown for 3-4 days in 100 mL DPY media (28 °C, 180 rpm) before the cells were collected by filtration. As a control, one part of the cells was extracted with ethyl acetate and analysed by LCMS to confirm the formation of the alcohol **67**, indicating the presence of the protein of interest (Figure 4.1).

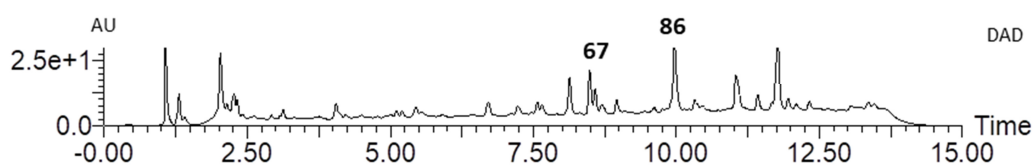


Figure 4.1: Organic extract of *A. oryzae* *ACE1* + *RAP1* transformant.

The rest of the cells were disrupted by a mortar and pestle under liquid nitrogen and resuspended in reaction buffer (Section 9.4.14). Cell debris were removed by centrifugation and the supernatant was directly used for activity assays. Aldehyde **97** ($m/z = 221$ g/mol) was used as a substrate analogue. LCMS analysis shows that this exists as an equilibrating mixture of hydrate **99** ($t_R = 3.1$ min) and aldehyde **97** ($t_R = 3.8$ min, Figure 4.2).

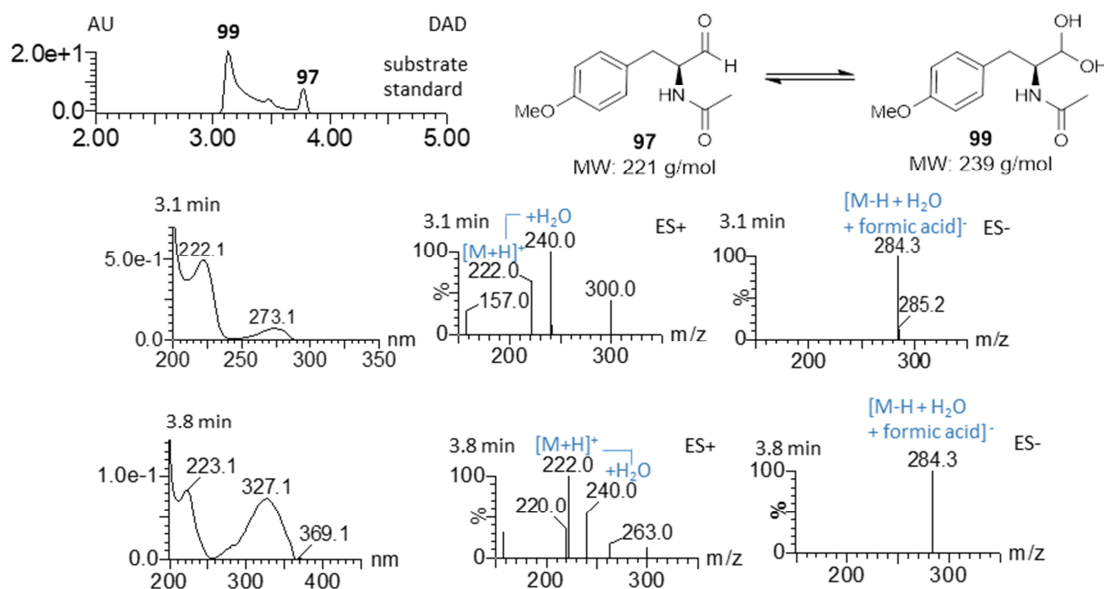


Figure 4.2: Substrate analogue **97**.

In-vitro enzyme assays were performed in a total volume of 100 μ l of 50 mM Sodium Phosphate buffer (pH 7.0) with a final concentration of 2.5 - 5 mM of the substrate model **97**, 2.5 mM NADPH, 2.5 mM NADH and 0.5 - 1 mg/ml of CFE. The reaction was incubated at 25 $^{\circ}$ C or 30 $^{\circ}$ C for 30-60 min. After this time the protein was precipitated with acetonitrile and removed from the reaction mixture by centrifugation prior to LCMS analysis (Section 9.4.14). Boiled CFE served as negative control. A standard of alcohol **98** was used as positive control. As expected, reduction of aldehyde **97** to alcohol **98** was observed (Figure 4.3).

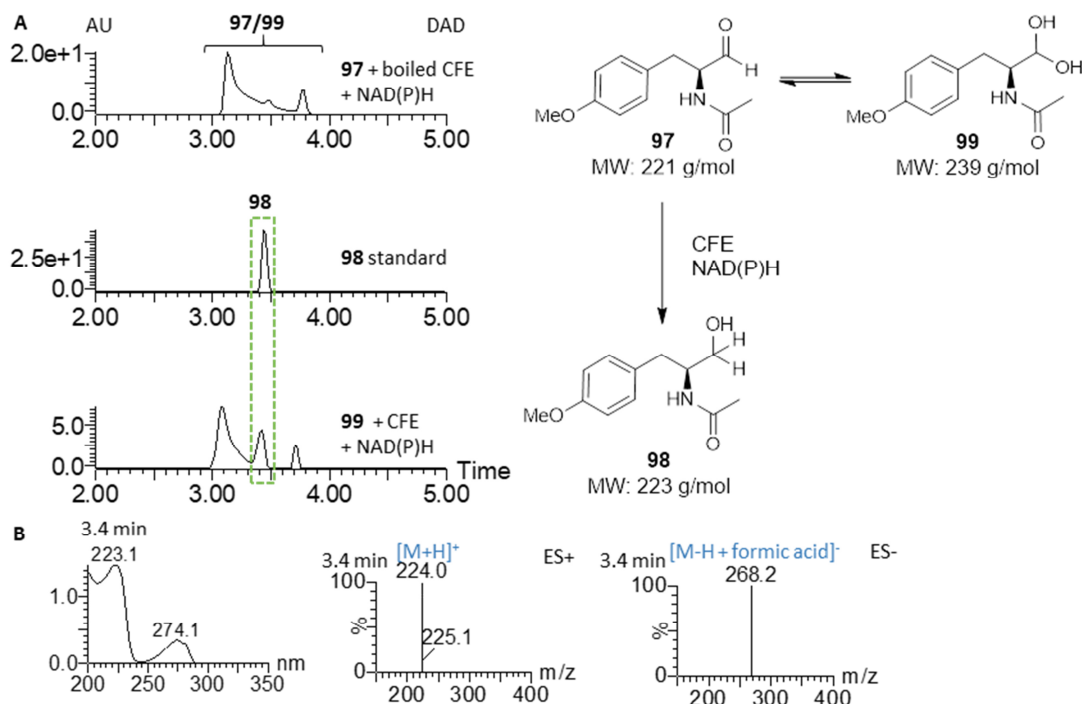


Figure 4.3: **97** is converted to **98** by the CFE of *A. oryzae*: **A**, Enzyme assays followed by LCMS; **B**, ES and UV spectra of **98**.

As a next step, the responsible enzyme should be identified. To this end, proteins of the CFE were separated by size exclusion chromatography (SEC) using a Superdex™ 75 10/300 GL column. Sample preparation and column chromatography were performed according to manufacturer using an ÄKTA™ pure chromatography system (Figure 4.4). 1 mL fractions were collected over 35 mL.

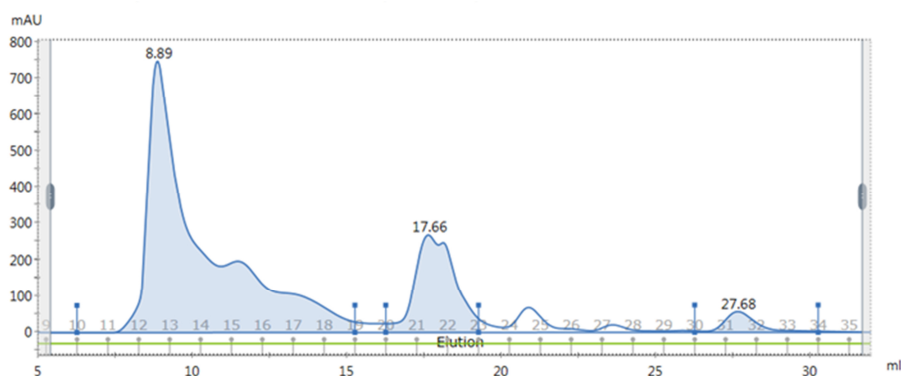


Figure 4.4: Results of the size exclusion chromatography using a Superdex™ 75 10/300 GL column.

Each collected 1 mL fraction was tested with the assay mixture for activity. The *in-vitro* assays were performed as previously described. 0.15 mg/ml enzyme mix (fraction) was used. Three fractions (14, 15 and 16) exhibited a strong activity, manifested by the conversion of the aldehyde **97** to **98** (Figure 4.5). Fraction 13 and 17 demonstrated much lower activity and the other fractions no activity.

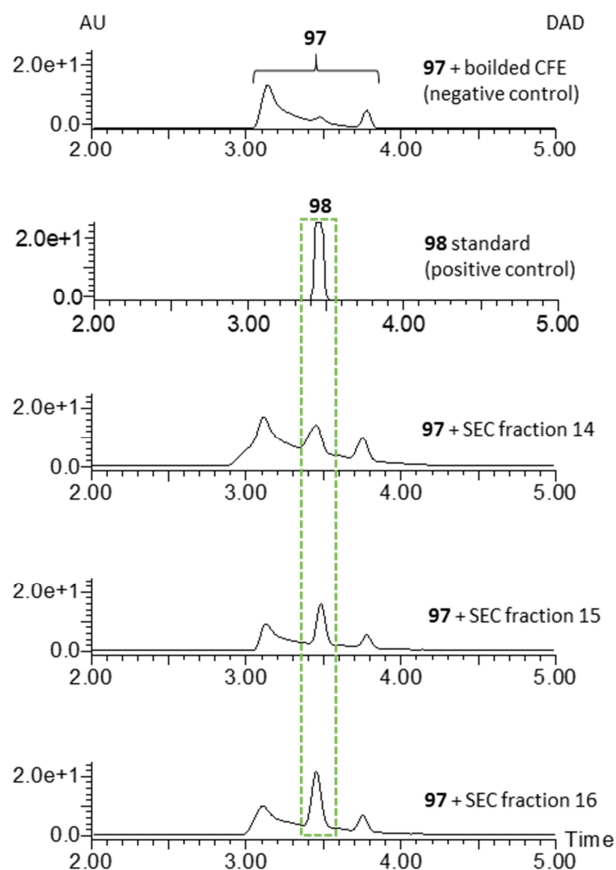


Figure 4.5: Results of *in-vitro* assays of SEC fractions. SEC fraction 14, 15 and 16 showed enzymatic activity.

SEC fraction 16 was further examined in more detail to test whether the enzyme is NADH or NADPH dependent. To this end, assays were performed containing only NADH or NADPH, respectively. This showed that the enzyme of interest is NADPH dependent (Figure 4.6). Four activity assays were performed simultaneously. The first assay contained 2.5 mM NADPH and 2.5 mM NADH (in addition to the enzyme mix and the substrate **97**). As expected, conversion of the aldehyde **97** to **98** was observed. This assay served as positive control. The second assay did not contain any cofactors and served as negative control. As expected, formation of **98** was not observed. The other two assays contained either NADPH or NADH as cofactor. Aldehyde **97** was only converted to **98** in the presence of NADPH. NADH was not accepted as cofactor, the reaction outcome was the same as for the negative control.

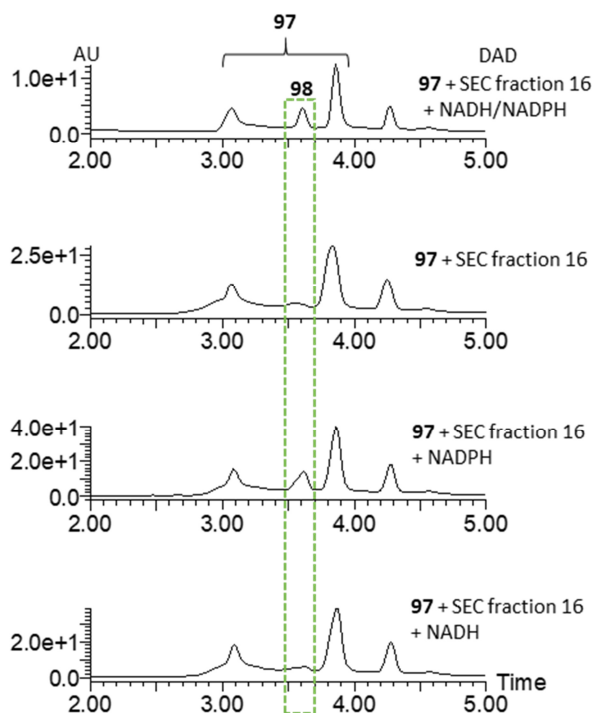


Figure 4.6: Cofactor determination. NADPH is needed to catalyse the formation of **97** to **98**.

Fraction 14, 15 and 16 were combined, concentrated and further separated by a strong anion exchange column (QFF, HiTrap® IEX Selection Kit, GE Healthcare, potassium phosphate buffer pH 7, elution was performed with a linear salt gradient (from 0.15 M to 0.5 M NaCl) which separates proteins based on their charge. 1 mL fractions were collected and tested for enzyme activity. Only the void fraction showed enzymatic activity, suggesting that the protein of interest was neutral or positively charged at pH 7. However, separation and visualization of the void fraction by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) revealed that the two purification steps were sufficient (Figure 4.7).

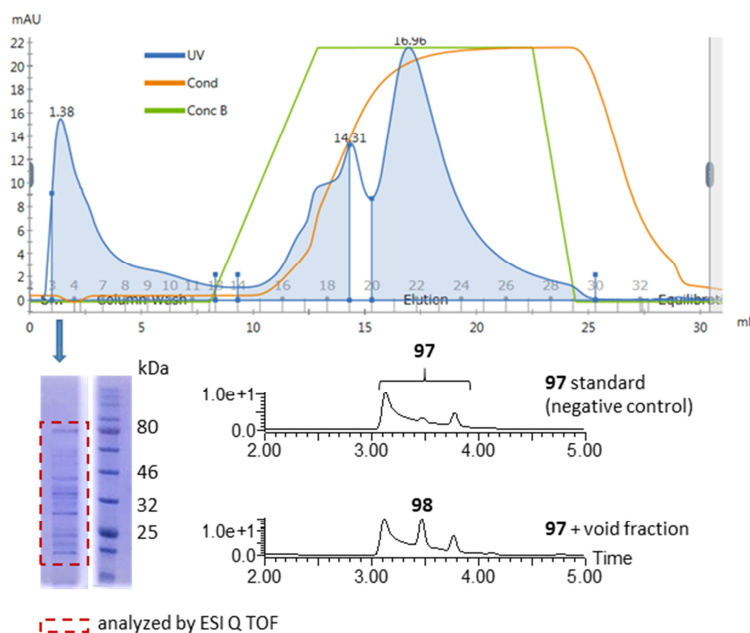


Figure 4.7: Results of the ion exchange chromatography using a QFF (Q Sepharose Fast Flow) 1 mL column. Only the void fraction showed activity towards **97**. The protein mixture of the void fraction was analysed by SDS-PAGE and ESI Q TOF.

4.2 Protein Identification by ESI Q TOF

The protein mixture of the void fraction was enzymatically digested by trypsin before the peptide fragments were further analysed by ESI Q TOF (Dr. Jennifer Senkler work group Prof. Braun, Institute of Plant Genetics, Leibniz University of Hannover, Table 4.2.1). Identified peptide sequences were searched against the NCBI database considering only *A. oryzae* protein hits. Based on the previous results, the expected protein was proposed to belong to NAD(P)H-dependant oxidoreductases.

Table 4.2.1: Results of ESI Q TOF. Positive hits are indicated in green.

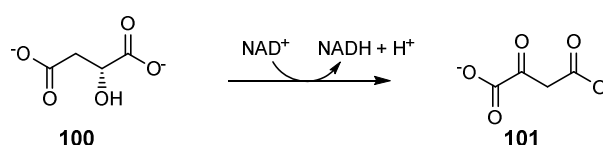
Accession	Description	Score	Coverage	AAs	MW [kDa]	calc. pI
169766066	malate dehydrogenase	412	28,2	340	35,7	8,76
439867	alcohol dehydrogenase	402	26,7	349	36,9	6,54
169779451	triosephosphate isomerase	249	20,3	251	27,4	5,55
169781716	phosphoglycerate kinase	195	10,6	417	44,4	7,94
169786471	glucoamylase	175	4,9	612	65,4	5,17
530721724	citrate synthase, partial	139	5,9	460	50,6	8,28
60760031	manganese superoxide dismutase, partial	134	23,9	134	14,3	7,05
145256869	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	110	3,2	774	87,0	6,68
134058618	unnamed protein product	72	4,7	212	23,3	5,43
83767456	unnamed protein product	67	1,5	779	84,9	6,76
83772095	unnamed protein product	62	4,1	291	30,3	5,99
134082236	unnamed protein product	57	1,8	502	55,8	6,43

Two potential NAD(P)H-dependant enzymes were identified as potentially responsible for the conversion of **97** to **98**: a malate dehydrogenase (AC: 169766066) and an alcohol dehydrogenase (AC: 439867, Table 4.2.1)

4.3 Discussion and Outlook

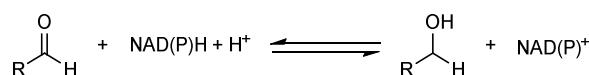
Heterologous production of the ACE1 metabolite in *A. oryzae* was only partially successful. It appears that a native *A. oryzae* enzyme reduced the aldehyde intermediate **90** to the alcohol **88** at an early biosynthetic step, thereby stalling the biosynthetic pathway at this point. The substrate analog **97** was reduced to **98** when incubated with the CFE of *A. oryzae*, confirming the idea that a native *A. oryzae* enzyme was responsible for this step. The designed screening assay to identify this enzyme was successfully established and applied to analyse the CFE. Two enzymes were identified to be probably responsible for the observed interruption of the ACE1 biosynthesis: a malate dehydrogenase and an alcohol dehydrogenase.

Malate dehydrogenases (MDH) reversibly catalyse the oxidation of malate **100** to oxaloacetate **101** as part of the citric acid cycle while reducing NAD^+ to $\text{NADH} + \text{H}^+$ (Scheme 4.2).¹³¹ As part of the primary metabolism it has a very specific and essential role and might therefore not be involved in the observed reduction of the aldehyde **90**. Also deletion experiments would not be possible since disruption of the gene would most likely be lethal. It also uses NAD(H) which was shown to be not involved. Nevertheless, since its enzyme activity against **97** was not tested, its role in this reduction step cannot be excluded.



Scheme 4.2: Catalysed reaction by malate dehydrogenase.

Alcohol dehydrogenases are enzymes that catalyse the reduction of aldehydes (or ketones) to alcohols or the oxidation of alcohols, respectively, in a NAD(P)H-dependant manner (Scheme 4.3).¹³²



Scheme 4.3: Catalysed reaction by alcohol dehydrogenases.

To this end, a reduction of aldehyde **90** to alcohol **88** might be catalysed by the identified alcohol dehydrogenase. Disrupting its function, *e. g.* by targeted gene deletion, might recover *A. oryzae* as

suitable host for the heterologous expression of the *ACE1* BGC as well as for other biosynthetic pathways where an aldehyde intermediate is formed. Eventually, it might be possible to investigate if the aldehyde **90** is recognized by the $\alpha\beta$ -hydrolase ORFZ to form the pyrrolinone **91**. Subsequently, the role of the putative DAase ORF3 could be investigated. The proposed [4+2] cyclisation of **91** would lead to the formation of the tricyclic core structure **96** of the *ACE1* metabolite. With this, the middle biosynthetic steps and the formation of the core structure of cytochalasans would finally be revealed.

Based on a more general view, a suitable host for the heterologous expression of cytochalasan gene clusters would substantially improve and accelerate research regarding novel bioactive cytochalasans by engineering their BGCs (*e. g.* domain swaps) or by feeding various amino acid analogues as shown for the biosynthesis of pyrivalasin H **50** in *M. grisea*.¹⁰⁸

5. Overexpression of the Transcription Factor *BC2* from the *ACE1* BGC in *P. oryzae*

5.1 The Transcription Factor *BC2*

During the infection cycle of rice leaves the avirulent *P. oryzae* Guy11 strain produces the *ACE1* metabolite, which is only produced temporarily and under very tight conditions in the appressorium during fungal penetration into rice leaves (Section 1.7).⁸⁶ The structure of the *ACE1* metabolite is therefore unknown, since it is only momentarily produced in an inaccessible organ.

Due to its biological activity and its interesting role in fungi-plant interactions, the *ACE1* biosynthetic gene cluster (BGC) responsible for the biosynthesis of the *ACE1* metabolite, is of great interest (see Section 1.7). The *ACE1* BGC consists of 15 genes including the *BC2* gene encoding a putative transcription factor (TF), which belongs to the $Zn(II)_2Cys_6$ binuclear zinc finger family (see Section 1.2 for information about TF and 1.7.2 for information about the *ACE1* BGC). Efforts to disrupt or silence the *BC2* gene failed, therefore its role in the regulation of the *ACE1* cluster remained elusive.⁸⁸ However, expression levels of the *ACE1* gene cluster generated during fungal penetration into the host plant demonstrated that *BC2* is (weakly) expressed and therefore might act as activator rather than as repressor. Since its expression is appressorium-specific, it might be controlled by a regulatory network within the appressorium.⁸⁸

A putative *ACE1* BGC was also identified in the fungus *M. grisea* NI980.¹⁰⁸ Transcriptome analysis of *M. grisea* revealed that neither genes of the *ACE1* nor of the *SYN2* BGC were properly expressed (done by Chongqing Wang and Elizabeth Skellam, AK Cox, Leibniz Universität Hannover, the fungi was cultivated for 5 days, unpublished data). These results further suggest that *BC2* may work as activator and not as a repressor.

5.2 Activation Strategies for Silent Fungal BGC

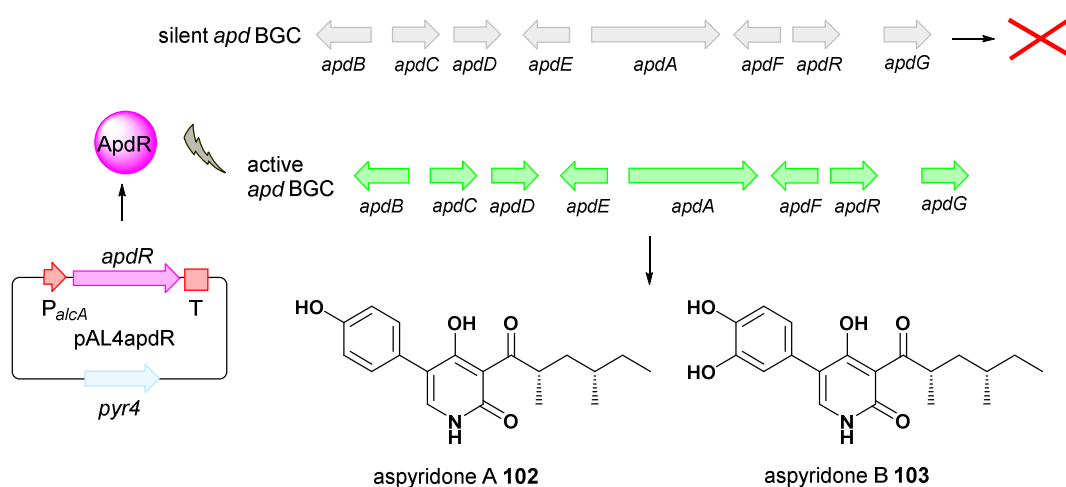
Fungal species are a good source of structurally diverse bioactive metabolites. However, especially under laboratory growth conditions, most BGC are silent, since their activation relies on environmental cues. Consequently, in the last years, many efforts have been made on activating silent gene clusters in fungal species.

Changing the secondary metabolite production of a strain can be sometimes achieved by changing the culture conditions, such as the composition of the growth media, the temperature

or the light conditions. Under changing cultivation parameters one strain is usually able to produce various compounds.¹³³ In addition to the culture conditions, external cues have been demonstrated to change the secondary metabolite production. The fungus *Phomopsis asparagi* for example produced three new compounds, chaetoglobosin-510, -540, and -542, when confronted with the F-actin inhibitor jasplakinolide.¹³⁴ Co-cultivation of two or more species can also stimulate the production of new secondary metabolites. The coral bacterium *Bacillus amyloliquefaciens* GA40 for example produced lipopeptide antifungal metabolites when co-cultivated with *A. fumigatus* and *A. niger*.¹³⁵ However, such metabolites are highly unpredictable and cannot be used to systematically activate any given BGC.

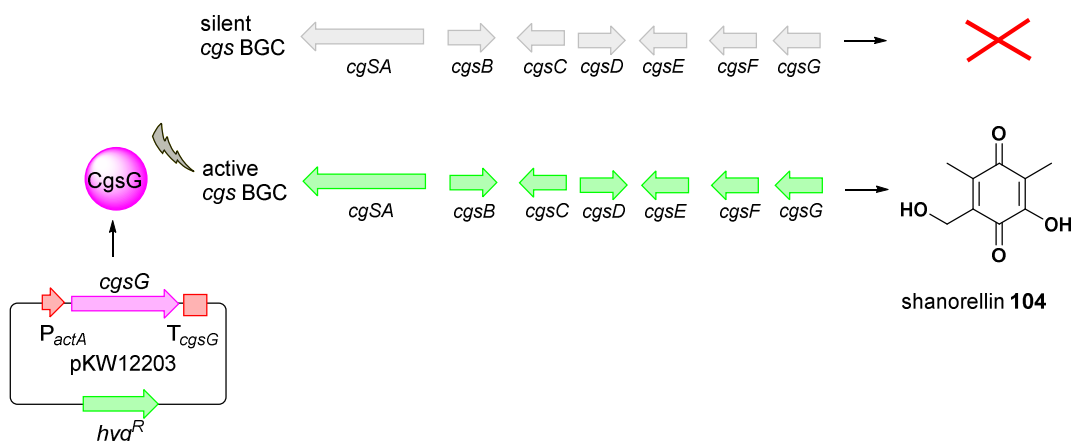
The stimulation of a specific BGC can be more challenging if the mode of activation is unknown. Heterologous expression of the whole BGC in a suitable host organism using defined promoters was demonstrated to successfully lead (in some cases) to the production of the desired compound (see Chapter 3, introduction).

As another promising approach it was demonstrated that in some cases the overproduction of the native pathway-specific TF of the BGC of interest can lead to the production of the corresponding secondary metabolite (for discussion of how pathway-specific TF are thought to work see Section 1.2). For example, Hertweck and co-workers demonstrated for the first time that ectopic expression of a pathway-specific TF in its native host *A. nidulans* under the control of an inducible promoter successfully led to the activation of the corresponding silent BGC. Thus, expression of the *apdR* gene, encoding a putative activator gene, resulted in the production of aspyridones A **102** and B **103** by initiating the expression of the genes from the *apd* BGC (Scheme 5.1).¹³⁶



Scheme 5.1: Activation of the silent *apd* BGC lead to the production of aspyridones A **102** and B **103** in *A. nidulans*.¹³⁶

Similar results were obtained in *Chaetomium globosum* when the *cgsG* gene, encoding a TF, present in the silent *cgs* BGC was expressed under the control of an inducible promoter. The *cgsG* gene is part of a BGC of seven genes including a PKS encoding gene; its activation led to the biosynthesis of shanorellin **104** (Scheme 5.2).¹³⁷



Scheme 5.2: Activation of the silent *cgs* BGC lead to the production of shanorellin **104** in *Chaetomium globosum*.¹³⁷

Transcription factors can not only promote (act as activator) the expression of a gene; they can also repress the expression by blocking the recruitment of the RNA-polymerase II (enzyme performs transcription of DNA into mRNA).¹³⁸ Therefore, expression of the TF does not always lead to increased expression levels. In addition, in some cases the expression of the BGC is not only controlled by a single regulator but by a combination of several signals.¹³⁹

5.3 Aim of the Studies

The first aim will be to determine if production of the TF *BC2* initiates the expression of the *ACE1* BGC in *P. oryzae* Guy11. To this end, a transformation method for *P. oryzae* Guy11 will be established. Subsequently, expression of *BC2* under the control of a defined inducible promoter will be performed, thereby showing if the biosynthesis of the *ACE1* metabolite can be stimulated by the presence of *BC2*. In addition, provided that the *ACE1* BGC is activated by *BC2*, the structure of the *ACE1* metabolite and/or other novel secondary metabolites will be elucidated.

5.4 Results

5.4.1 Transformation of *P. oryzae* Guy11 with pTH-GS-*egfp*

The *P. oryzae* Guy11 strain is known to be sensitive to the antibiotics hygromycin B and BASTA (glufosinate ammonium).¹⁴⁰ In order to develop a reliable transformation method, genetic manipulation of *P. oryzae* Guy11 was established using the vector pTH-GS-*egfp*. The vector features a hygromycin B resistance cassette (*hyg^R* consisting of the *E. coli hph* gene driven by *A. nidulans P_{gpdA}*) as well as the *egfp* gene (encoding an enhanced green fluorescent protein (eGFP)), as visible marker. The *egfp* gene is expressed under the control of the strong inducible (by starch or maltose) *amyB* promoter (*P_{amyB}*) from *A. oryzae* (Figure 5.1).^{141,142}

Transformation of *P. oryzae* Guy 11 was performed with vector pTH-GS-*egfp* using a PEG-mediated protoplast protocol (Section 9.3.4) resulting in twelve hygromycin B resistant transformants. Six transformants and the wild type were cultivated for 5 days in DPY medium before the samples were analysed by fluorescence microscopy. Upon typical GFP excitation and emission conditions (450 - 490 nm) the successful production of eGFP was ascertained in all transformants, which was not observed for the wild type control (Figure 5.1, B I-IV). Thus, the strain *P. oryzae* Guy11 was successfully transformed using hygromycin B as the selection marker. Furthermore, *P_{amyB}* efficiently initiates expression of *egfp* demonstrating that it can be used in further studies to express the gene of interest in *P. oryzae* strains.

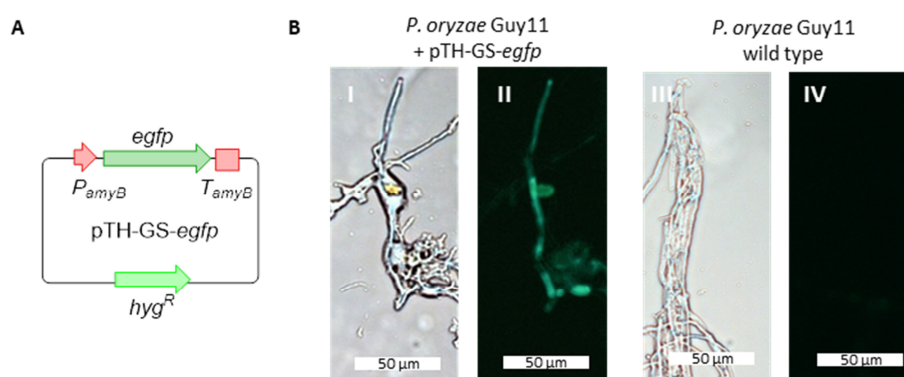


Figure 5.1: *P. oryzae* Guy11 transformation: **A**, Vector map of pTH-GS-*egfp*; **B**, *P. oryzae* transformed with pTH-GS-*egfp* (I+II) and WT (III+IV), I+III) upon excitation and II+IV) bright-field.

5.4.2 Homologous Overexpression of *BC2* *P. oryzae* Guy11

A fungal expression vector (pTGYS-*hyg^R*-*BC2*) for the expression of *BC2* was constructed by *in-vivo* homologous recombination in *S. cerevisiae* using a hygromycin B resistance cassette as selection marker (Figure 5.2 A). DNA fragments for assembly were amplified by PCR. For amplification of

BC2, *P. oryzae* Guy11 gDNA was used as template. The selection marker was amplified from vector DNA (pTH-GS-*egfp*).

P. oryzae PEG/CaCl₂ mediated protoplast transformation with pTGYS-*hyg*^R-*BC2* resulted in the isolation of fourteen hygromycin resistant transformants growing on tertiary selection plates (Figure 5.2).

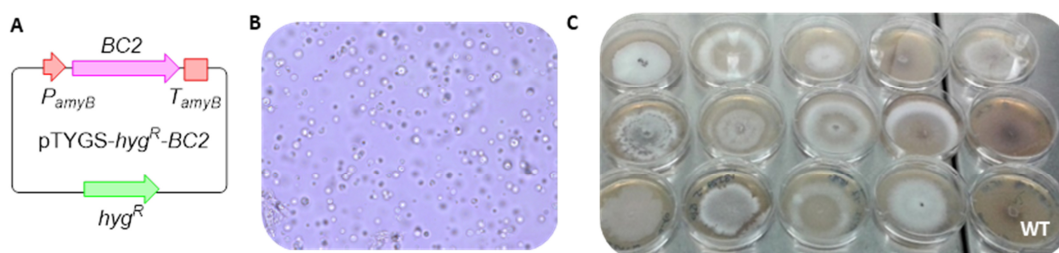


Figure 5.2: *P. oryzae* Guy11 transformation: **A**, Vector map of pTYGS-*hyg*^R-*BC2*; **B**, *P. oryzae* protoplasts isolated after 2 h of incubation in protoplasting solution (40x); **C**, *P. oryzae* Guy11 transformants and wild type (WT) on DPY plate after 12 days. Previous selection plates contained 250 µg/ml hygromycin B.

All transformants were cultivated for 7 days in DPY media (25 °C, 110 rpm) before cells and growth media were individually extracted twice with ethyl acetate and analysed by LCMS (Section 9.5). In five transformants (VBI27-3, 4, 5, 12) production of a new, dark brown compound **105** (MW = 348 g/mol) was observed in extracts of the media (Figure 5.4 A – C). The compound **105** was not observed in extracts of *P. oryzae* Guy11 wild type strain.

As part of the *ACE1* BGC the *BC2* gene is present in the *P. oryzae* Guy 11 WT strain and therefore also in all transformants. To test whether an extra copy of the *BC2* gene (which is fused to *P*_{amyB}) is present in the transformants, a PCR experiment was designed with the aim to amplify the *P*_{amyB}-*BC2* junction. Unfortunately, also the WT control (negative control) yielded a PCR fragment of the expected size. Probably, unspecific interactions of the forward primer with the native promoter of *BC2* led to the observed fragment (data not shown). To overcome this problem, the junction between *BC2* and the terminator sequence (*T*_{amyB}) was amplified using P1521 and P1522. This time no DNA fragment was observed in the WT control. As a positive control vector DNA (pTGYS-*hyg*^R-*BC2*) was used. A copy of *BC2* was integrated into the genome of the four genetically analysed transformants (VBI27-3, 4, 5, 12, Figure 5.3).

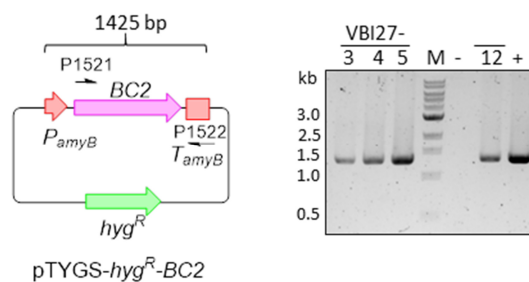


Figure 5.3: Genetic analysis of *P. oryzae* Guy11 transformants. M = marker.

In all transformants production of metabolites related to the *ACE1* metabolite or pathway intermediates was not observed. Metabolites related to the *ACE1* compound would have a presumable mass greater than 465 g/mol (the identified *ACE1* backbone **67** has a mass of 481 g/mol, the expected pyrrolinone **85** of 465 g/mol, see Section 3.3.1). In addition, the UV absorption spectra were used for screening. As in the biosynthesis of pyrichalsasin H **50**, *O*-methyl tyrosine **71** is fused to the polyketide backbone **88** of the *ACE1* metabolite. Compounds derived from tyrosine usually demonstrate characteristic UV absorption spectra at 275 nm (see Section 7.3.1).

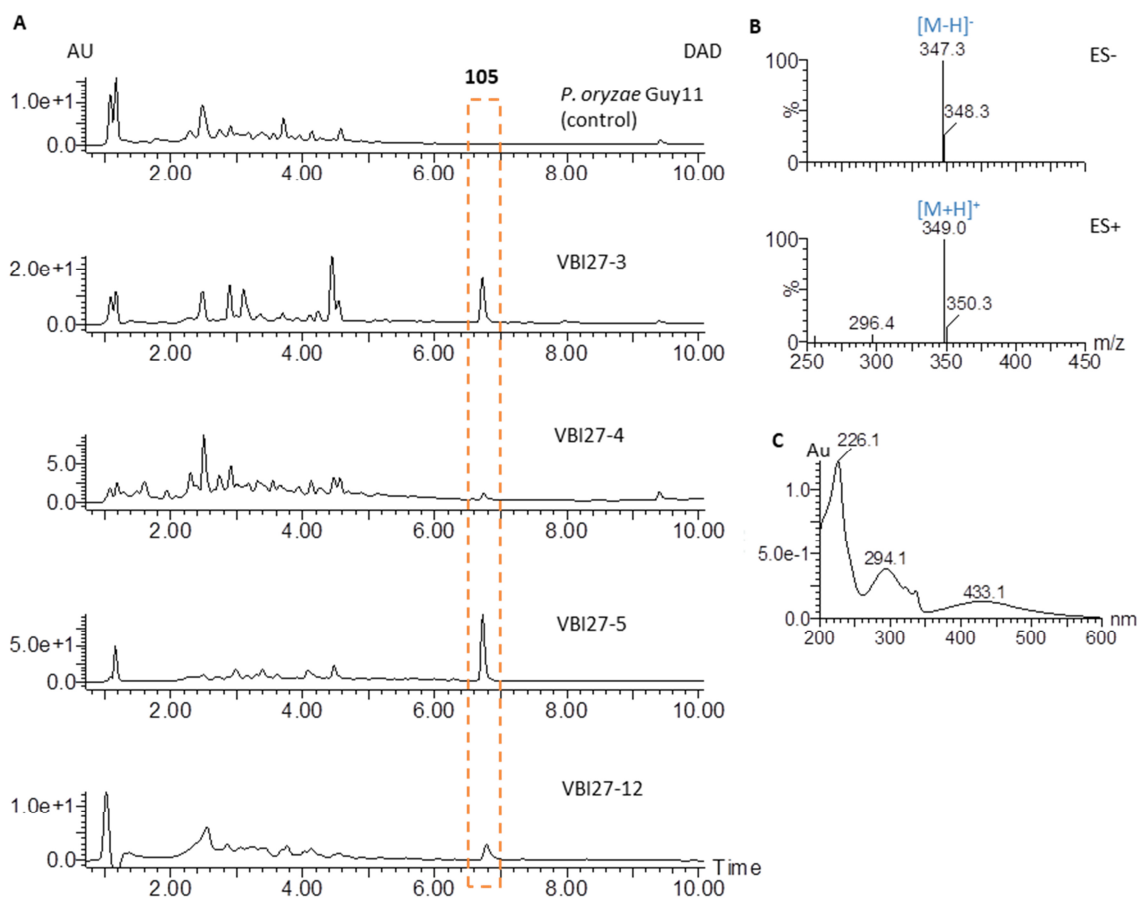


Figure 5.4: Organic extracts of *P. oryzae* transformants: **A**, Organic extracts of *P. oryzae* + pTYGS-*hygR-BC2* and wild type; Diode array detector (DAD); **B**, ES of compound **105**; **C**, UV absorption spectra of **105**.

1.3 mg of **105** were purified by preparative LCMS and the structure was elucidated by NMR (Table 5.4.1). HRMS confirmed a molecular formula of $C_{20}H_{12}O_6$ ($[M]-H^+$ calculated 347.0542, found 347.0529). Compound **105** was identified as the already known hinnulin A. This compound was previously isolated from *Nodulisporium hinnuleum* and belongs to the class of 1,2'-binaphthyl natural products.¹⁴³ Further compounds related to **105** were not observed.

Couplings in the $^1H-^1H$ COSY spectrum revealed characteristic signals for aromatic regions (Figure 5.5 A). Compound **105** contains 8 aromatic protons (H-1, H-7, H-8, H-12, H-13, H-18, H-19 and H-20). Notably, the signal for H-1 is a singlet (6.24 ppm); it demonstrated no coupling signal in the COSY spectrum. Based on the HMBC experiment, H-1 correlates to two carbonyl groups (C-6 at 193.5 ppm and C-3 at 182.6 ppm), one hydroxyl group (C-2 at 161.6 ppm) and one alkene (C-5 at 115.8 ppm) and is part of the first aromatic ring (Figure 5.5 B). Key HMBC signals for the adjacent aromatic ring (ring 2) derived from the hydroxyl group at C-10 (Figure 5.5 C). Notably, only for 10-OH a signal (singlet at 12.86 ppm) was observed. Proton signals for the remaining three phenolic protons are missing (or too broad and blurred with baseline noise), presumably due to proton exchange in deuterated solvent. Starting from H-8, HMBC data revealed that the second and the third aromatic ring are connected by a carbon-carbon bond of C-9 and C-11 (Figure 5.5 D).

Table 5.4.1: NMR data of **105**. Not all 1H signals could be integrated due to overlap. Solvent: CD_3CN .

Atom	δ_C (ppm)	δ_H (ppm)	Multiplicity (J, Hz)	COSY	HMBC	HSQC	Structure
1	111.4	6.24	s	-	2, 3, 5, 6	CH	<p>105 $C_{20}H_{12}O_6$ MW: 348 g/mol</p>
2	161.6	-	-	-	-	-	
	182.6	-	-	-	-	-	
4	131.0	-	-	-	-	-	
5	115.8	-	-	-	-	-	
6	193.5	-	-	-	-	-	
7	120.1	7.71	d (7.6)	8	3, 5, 8, 9	CH	
8	138.4	7.58	d (7.6)	7	4, 10, 11	CH	
9	139.6	-	-	-	-	-	
10	160.6	-	-	-	-	-	
11	127.4	-	-	-	-	-	
12	130.0	7.24	d (7.7)	13	9, 14, 16	CH	
13	109.8	6.87	d (7.8)	12	11, 14, 15	CH	
14	155.9	-	-	-	-	-	
15	115.4	-	-	-	-	-	
16	136.3	-	-	-	-	-	
17	155.9	-	-	-	-	-	
18	110.4	6.82	d (7.6)	19	15, 17, 20	CH	
19	128.6	7.23	t (8.1, 8.1)	18, 20	16, 17	CH	
20	119.2	6.97	d (8.5)	19	11, 15, 16, 18	CH	
10 - OH	-	12.86	s	-	5, 9, 10	-	

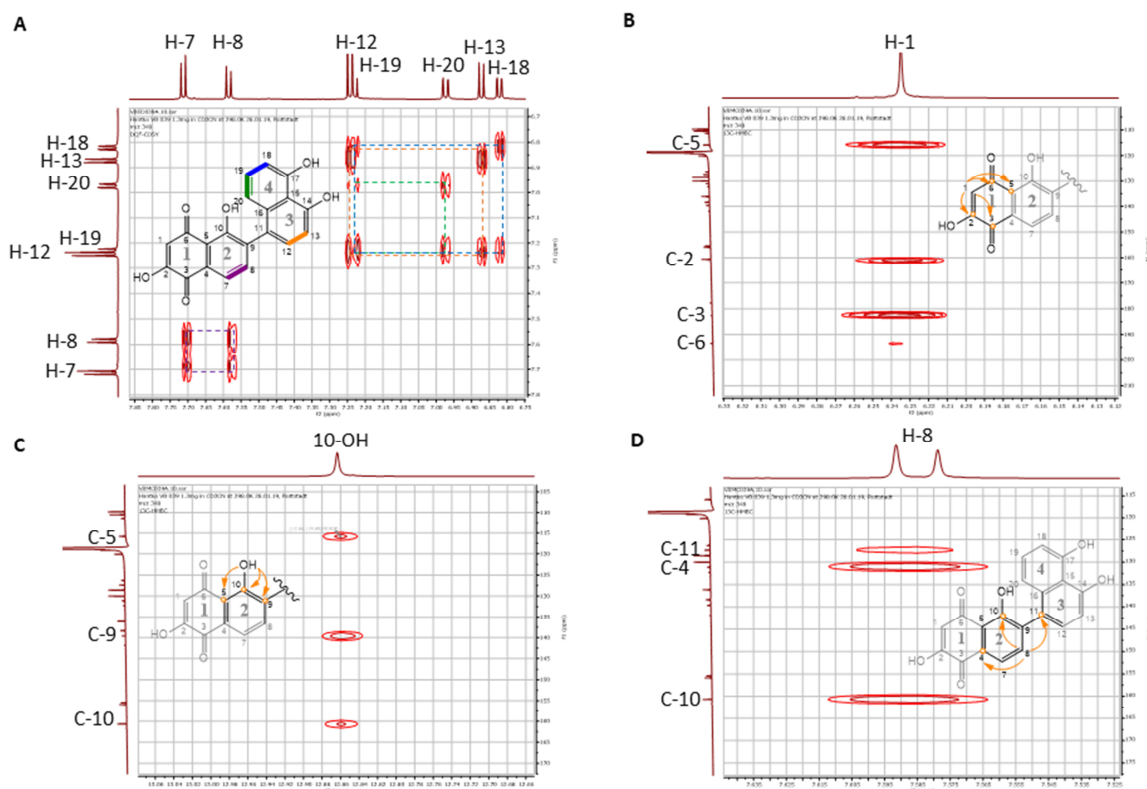


Figure 5.5: NMR spectra of the purified compound **105**: **A**, ^1H - ^1H COSY of the aromatic protons; **B**, HMBC correlation of H-1; **C**, HMBC correlation of 10-OH; **D**, HMBC correlation of H-8.

The strain VBI27-5 was extracted at different time points (between 2 and 7 days) to further screen for biosynthetic intermediates of **105** and compounds potentially related to the *ACE1* metabolite. After 2 to 3 days the organic extracts demonstrated almost no secondary metabolite production (data not shown). However, the extract collected on day 4 displayed a new compound ($m/z = 334$ g/mol, Figure 5.6). From day 5, only the production of **105** was observed.

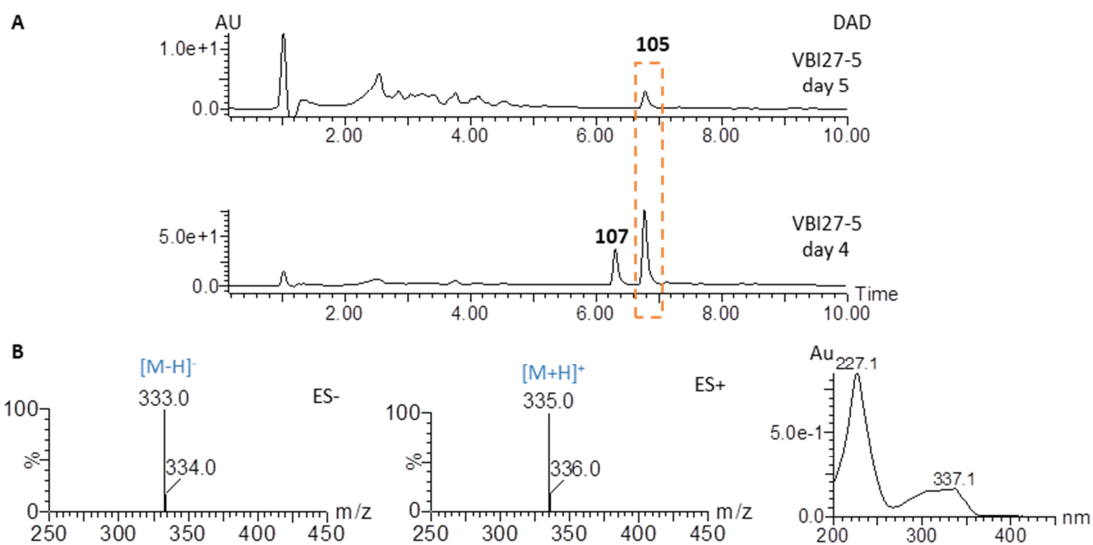
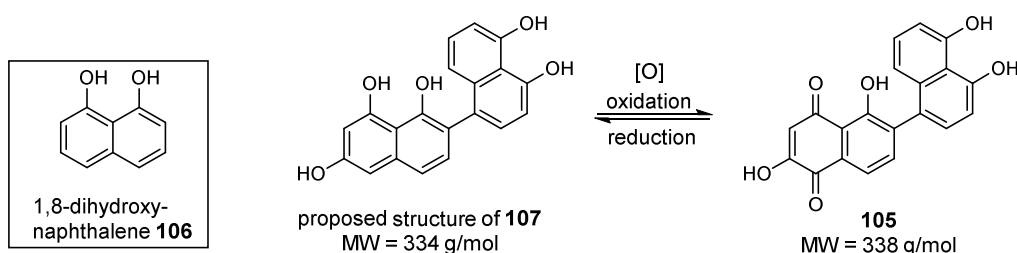


Figure 5.6: Organic extracts of VBI27-7 cultured for 4 and 5 days: **A**, Organic extracts; Diode array detector (DAD); **B**, ES and UV absorption spectra of **107**.

The mass difference between **105** and **107** is 14 g/mol, which would be consistent with the absence of a carbonyl group. Since **107** was only observed on day 4 together with **105**, it is likely to be a pathway intermediate or shunt of **105**. One should keep in mind, that compounds such as **105** derive from 1,8-dihydroxynaphthalene **106** (DHN) monomers.^{104,105} Based on this knowledge, a structure was proposed for **107** (Scheme 5.3). The UV absorption spectra of 337 nm (Figure 5.6 C) is similar to the UV absorption spectra of other known compounds derived from DHN **106**, further suggesting that **107** might be a biosynthetic intermediate of **105**.^{144,145} Oxidation of **107** would result in the formation of **105**.



Scheme 5.3: Proposed structure of the observed compound **107**

5.4.3 Reverse Transcriptase (RT) Analysis of Transformant VBI27-5

Successful expression of the transcription factor *BC2* was determined by reverse transcriptase (RT) analysis of transformant VBI27-5. To this end, messenger RNA (mRNA) of VBI27-5 and the *P. oryzae* Guy11 wild type strain (used as a control) was extracted using the ZR Fungal Bacterial RNA MiniPrep kit (Zymo Research) kit and converted into the corresponding complementary DNA (cDNA) before the samples were analysed by PCR (Section 9.4.4). RNA was used as negative control to ensure that no gDNA remained in the sample. In addition, a control PCR was performed to amplify the house-keeping gene *MGG_06650*, encoding tubulin alpha-B chain. Since the gene *MGG_06650* contains introns, the amplification product from gDNA is longer than the product from mRNA (Figure 5.7 A). As expected, expression of *BC2* was confirmed for VBI27-5, but was absent in the WT (Figure 5.7 B).

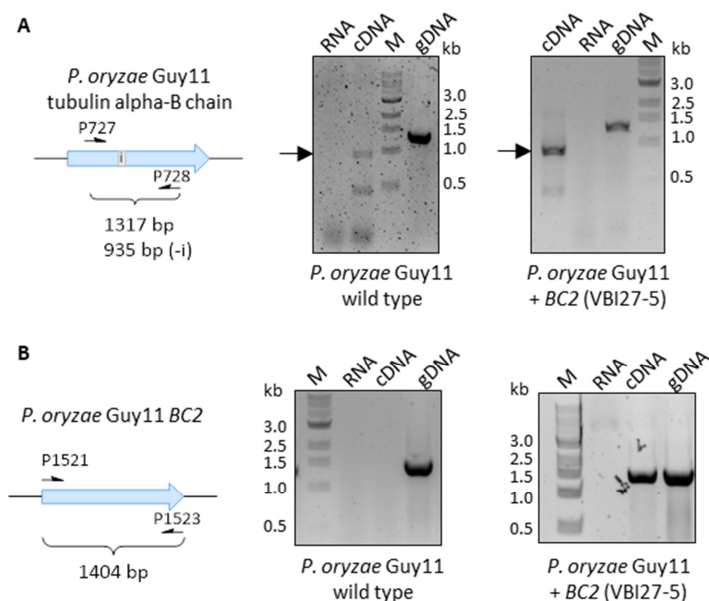


Figure 5.7: Reverse Transcription (RT) analysis of transformant VBI27-5: **A**, The gene encoding for tubulin alpha-B chain was used as control for gDNA contamination; **B**, Amplification of *BC2* was observed in VBI27-5, M = marker.

In addition, expression of genes from the *ACE1* BGC (*ACE1*, *RAP1*, *ORFZ* and *ORF3*) was tested (Figure 5.8). In accordance with the observed secondary metabolite profile, expression of all four genes was not observed. Apparently, expression of *BC2* did not activate the biosynthetic pathway of the *ACE1* metabolite under the tested conditions.

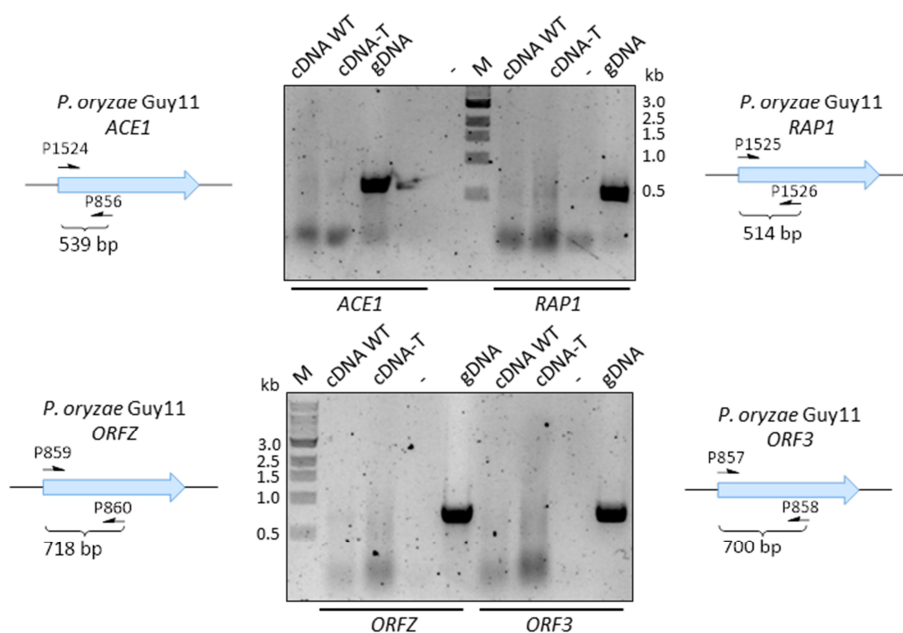


Figure 5.8: RT-PCR results of *ACE1*, *RAP1*, *ORFZ* and *ORF3*. WT = wild type, T = transformant VBI27-5, RNA isolated from VBI27 was used as negative control (-), M = marker.

Remarkably, expression of *BC2* appeared to have influenced expression of another BGC, leading to the production of **105**. Based on the dark-brown color and its structure, **105** seems to be related to fungal 1,8-dihydroxynaphthalene (DHN)-melanin. Therefore, the **105** BGC is likely to contain a gene encoding a PKS homologous to the *alb1* gene (AFUA_2G17600) known to be involved in the synthesis of 1,3,6,8-tetrahydroxynaphthalene **108** (T₄HN), the precursor of DHN **106** in fungi.¹⁴⁶ The *Colletotrichum lagenarium* *PKS1* gene encodes a non-reducing-PKS (nrPKS) which is known to produce T₄HN **108**.^{147,148} A BLAST search was performed with the *PKS1* gene to identify a “melanin-like” BGC in *P. oryzae* Guy11. Two potential genes were identified (*MGG_07219*, 72 % identity and *MGG_00428*, 35 % identify). *MGG_07219* also shows 48% identity to the *A. fumigatus alb1* PKS and *MGG_00428* 42 % identity. Genes adjacent to *MGG_07219* encode two transcription factors, an iron transport multicopper oxidase, a T₄HN reductase and several hypothetical uncharacterised proteins. A BLAST search with the *MGG_07220* gene (encoding iron transport multicopper oxidase) exposed its sequence similarity to a laccase-like multicopper oxidase (79.1 % identity) in *Colletotrichum orbiculare*. Both enzymes, the T₄HN reductase and the laccase, are known to be involved in the biosynthesis of DHN-melanins.^{149,150,151,152} The genes adjacent to *MGG_00428* encode a short-chain dehydrogenase/reductase SDR, an *O*-methyltransferase, a 1,2-dioxygenase and several hypothetical proteins unlikely to be involved in the biosynthesis of melanin or **105**. The genetic borders of the respective locus of both clusters are unknown. Based on these findings, the *MGG_07219* BGC is the best bioinformatic hit and might be involved in the biosynthesis of **105**. However, further experiments were conducted to experimentally validate this hypothesis.

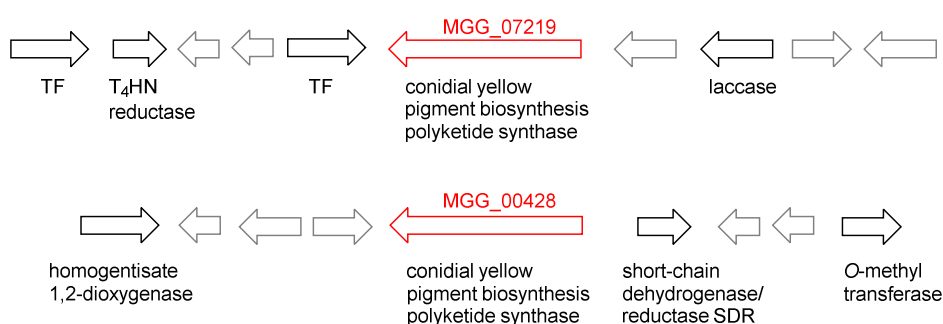


Figure 5.9: “Melanin-like” BGC in *P. oryzae* Guy11.

Expression of both genes, *MGG_07219* and *MGG_00428*, was tested by PCR of the previously prepared cDNA template. Remarkably, expression of *MGG_07219* was observed in the transformant VBI27-5 and not in the wild type control (Figure 5.10). Expression of *MGG_00428* was not observed; neither in the transformant VBI27-5 nor in the wild type control (Figure 5.10).

These findings support the assumption that *MGG_07219* might be involved in the biosynthesis of **105**.

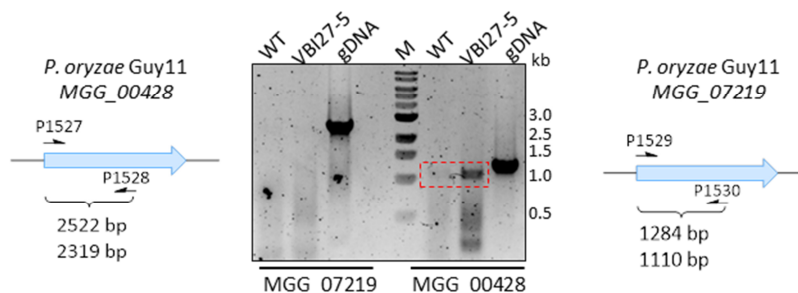


Figure 5.10: RT-PCR results of *MGG_07219* and *MGG_00428*. WT = wild type, M = marker.

To gain more insight into the identified BGC containing *MGG_07219*, expression of the 6 adjacent genes was tested. These encode a T₄HN reductase, two transcription factors (TF1 and TF2), two hypothetical proteins (named in this work L2 and R3) and a laccase-like multicopper oxidase (Figure 5.11). The remaining genes (encoding hypothetical proteins named in this work L1, R1 and R2) were rather small (< 200 bp) and unlikely encode any functional enzymes. Therefore, these genes were excluded from further examinations. Expression of the T₄HN reductase gene (*MGG_07216*) and the gene encoding TF1 (*MGG_07215*) was observed for the wild type and the transformant strain. However, expression of the other genes was not observed; neither in the wild type strain nor in the transformant.

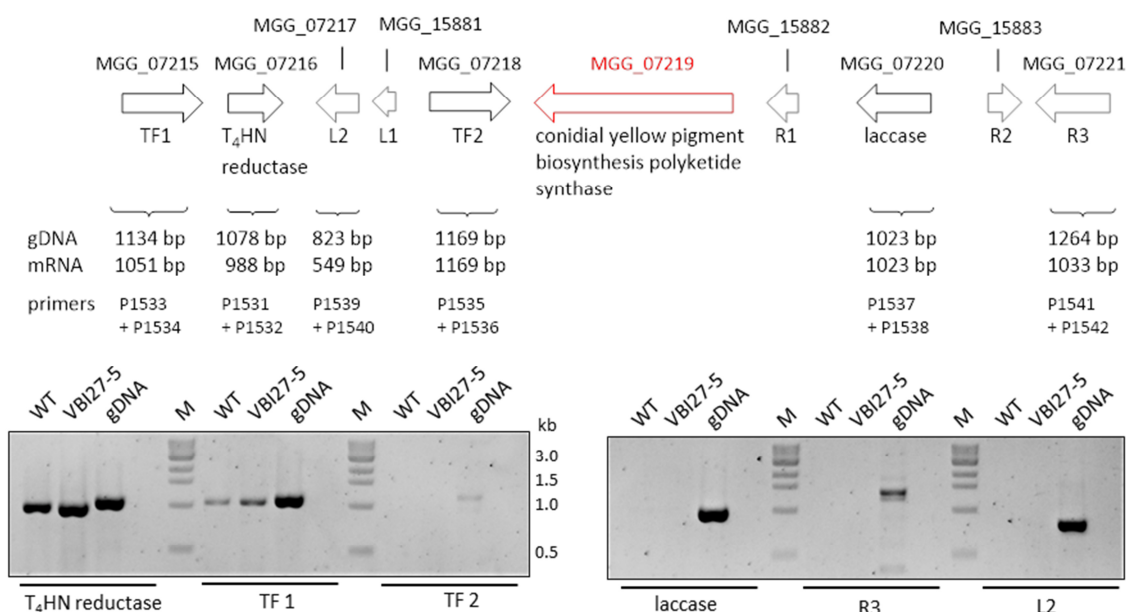


Figure 5.11: RT-PCR results of the *MGG_07219* BGC. WT = wild type, M = marker.

5.5 Discussion

During the invasion of rice plants by *P. oryzae* Guy11, the fungus is known to produce an avirulence signal - the *ACE1* metabolite. Under laboratory conditions, production of the *ACE1* metabolite was not observed; the *ACE1* BGC remained silent.^{86,153} The transcription factor *BC2* is part of the *ACE1* BGC, however, its mode of action is still unknown.⁸⁸

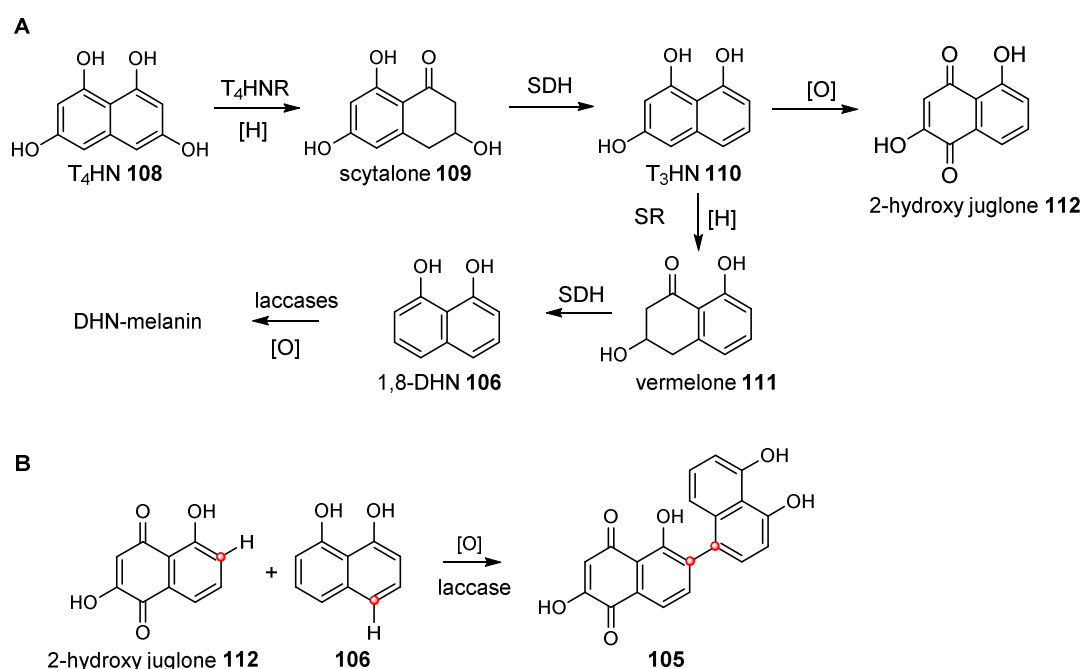
Our experiments showed that overexpression of *BC2* in *P. oryzae* Guy11 does not lead to expression of *ACE1*, *RAP1*, *ORF3* or *ORFZ*. Consistent with this no *ACE1*-related compounds were observed. Apparently, this pathway-specific TF was not able to activate the *ACE1* BGC under tested conditions. One reason for this might be that the expression of *BC2* is activated together with the remaining genes of the *ACE1* BGC by an unknown signal. Another explanation might be that *BC2* was not able to activate the *ACE1* BGC, since the locus of this cluster is inaccessible in *P. oryzae* Guy11.

Eukaryotic DNA is organized within the nucleus by a complex of DNA and histone proteins (H2A, H2B, H3 and H4) called the nucleosome. Units of nucleosomes form a higher-ordered structure – the chromatin.¹⁶ Post-translational modifications of amino acid tails of the histones (such as acetylation and methylation) influence the DNA accessibility by altering the chromatin conformation. Subsequently, the chromatin structure has an impact on the expression pattern by allowing or restricting, amongst other things, TF to access the gene locus.¹⁵⁴ In 2007 Keller and co-workers demonstrated for the first time that activation of secondary metabolites gene clusters is connected to the chromatin structure. Deletion of the *hdaA* gene, encoding a histone deacetylase, in *A. nidulans* caused transcriptional activation of two SM gene clusters producing sterigmatocystin and penicillin.¹⁵⁵ Histone acetylation was demonstrated to play an essential role in regulating the activity of SM gene cluster in many fungi. In the aflatoxin **25** BGC acetylation of histone H4 in the promoter region of genes was demonstrated to correlate with transcriptional activation.¹⁵⁶ Production of fusarin C **8** in *F. fujikuroi* was exposed to be associated with an enrichment of acetylated histone H3 at the active **8** BGC.^{157,158} In addition, histone methylations often mark silent BGC and are another way to control transcription of genes.¹⁵⁹

Based on these findings, the histone-state of the *ACE1* cluster is likely to be crucial for the expression of the genes. The chromatin structure at the *ACE1* locus is probably linked to the inability to generate deletion mutants for most of the genes (except for *ACE1*, *SYN2* and *RAP2*). In addition, marker genes (*e. g.* resistance to hygromycin) integrated at this locus are likely not expressed, and therefore transformants cannot be selected. The chromatin structure could also

explain why attempts of producing the TF *BC2* failed to activate the expression of *ACE1* cluster genes. One way to circumvent this problem might be to change the conformation of the chromatin from heterochromatin (tightly packed form of chromatin) to euchromatin (lightly packed form of chromatin). To address this, inhibiting the methylation of histones and/or preventing their deacetylation, either genetically or chemically, might help to loosen the chromatin structure. Future experiments could be to feed epigenetic modifiers (*e.g.* DNA methyltransferase and histone deacetylase inhibitors) to *P. oryzae* strains expressing *BC2*.

Notably, overexpression of *BC2* in *P. oryzae* Guy11 led to the production of hinnulin A **105**. Compound **105** was first identified in the fungus *Nodulisporium hinnuleum* and belongs to the class of 1,2'-binaphthyl natural products.¹⁴³ Based on its colour and structure, the biosynthesis of **105** is probably connected to the DHN-melanin pathway. Most fungal melanins are derived from T₄HN **108**. The melanin pathway involves reduction of T₄HN **108** to scytalone **109** by T₄HN reductase (T₄HNR) and then dehydration of scytalone **109** by scytalone dehydratase (SDH) to form 1,3,8-trihydroxynaphthalene **110** (T₃HN). This is then reduced again to form vermellone **111**, at which point SDH acts again to form 1,8-DHN **106**. 1,8-DHN **106** is then polymerised by various laccases to form melanin (Scheme 5.4 A). Oxidation of T₃HN **110** is also known to lead to 2-hydroxy juglone **112** although the molecular basis for this step is not known.^{160–164} It seems likely that hinnulin A **105** could be formed by oxidative crosslinking of 1,8-DHN **106** and 2-hydroxyjuglone **112** by a laccase enzyme (Scheme 5.4 B).



Scheme 5.4: **A**, Known pathway to DHN melanin in fungi and **B**, proposed biosynthesis of **105**. T₄HNR = T₄HN reductase; SDH = scytalone dehydratase; SR = scytalone reductase.

Based on this hypothesis, a BlastP search was performed using the *Colletotrichum lagenarium* *PKS1* gene encoding a non-reducing-PKS (nrPKS) which is known to produce T₄HN as query sequence.^{147,148} Two putative “melanin-like” BGC in *P. oryzae* Guy11 were identified; both clusters contained a gene encoding for a conidial yellow pigment PKS (*MGG_07219* and *MGG_00428*). Moreover, the *MGG_07219* BGC contained genes encoding a T₄HN reductase and a laccase-like multicopper oxidase related to ABR1 and ABR2 which are proteins known to be involved in melanin formation in *Aspergillus fumigatus*.¹⁶⁵ This laccase could potentially be involved in the proposed formation of **105**, but its expression could not be confirmed. However, genes encoding SDH and SR and the potential oxidation of T₃HN **110** to 2-hydroxyjuglone **112** do not appear to be encoded by this BGC, although candidate genes are present elsewhere in the genome of *P. oryzae* Guy11. Collectively, the cluster was a good candidate and likely to be involved in the biosynthesis of **105**. RT-PCR analysis revealed that the *MGG_07219* gene was expressed in transformant VBI27-5, whereas expression in the *P. oryzae* Guy11 wild type control strain could not be confirmed. However, the expression profile of genes located next to *MGG_07219* was identical for VBI27-5 and the wild type strain.

Due to these results, expression of *MGG_07219* might be involved in the biosynthesis of **105**. However, one should keep in mind that the *MGG_07219* gene might not have been detected in the wild type cDNA sample, due to low mRNA quality caused by inefficient removal of RNAses during purification. To address this, isolation of wild type mRNA (and subsequent conversion into cDNA) could be repeated to confirm that *MGG_07219* is not expressed under tested conditions. Moreover, targeted gene deletion of *MGG_07219* could be performed in the strain VBI27-5 to test if the production of **105** is linked to the expression of *MGG_07219*.

In 2010 it was reported for the first time that cross-talk between fungal BGC occurs. Induced overexpression of the pathway-specific TF *scpR* in *A. nidulans* led to the activation of its corresponding *inp* BGC. However, the produced compound was not observed. In addition, a second BGC, located on a different chromosome, was activated and led to the production of asperfuranone **114**. It was demonstrated that ScpR activated the expression of *afmA*, the pathway-specific TF of the **114** BGC. Subsequently, expression of the **114** BGC was activated by AfoA (Figure 5.12).¹⁶⁶

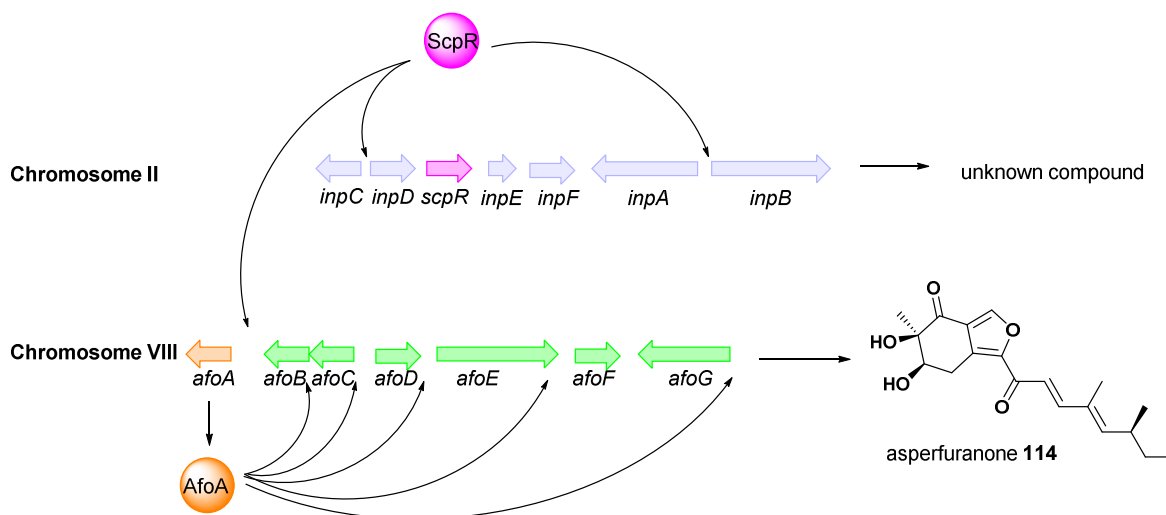


Figure 5.12: Model cross-talk between two fungal BGC present in *A. nidulans*: the *inp* and the *afo* BGC.¹⁶⁶

Collectively, two main observations were made: BC2 is not able to activate the expression of the other genes from the *ACE1* BGC under tested conditions; instead BC2 appears to activate a gene cluster responsible for the production of hinnulin A **105**, a compound related to DHN-melanin (Figure 5.13). DHN-melanin plays an essential role during invasion of the plant, and like the *ACE1* compound, it is located in the appressorium. The appressorium is a structure which mediates penetration of the fungus into the plant through the generation of turgor. Deposition of melanin in the cell wall of the appressorium was demonstrated to be essential for maintaining turgor pressure (for more information see Section 1.7.3).^{167,101} To this end, expression of the *ACE1* BGC might be directly connected to the production of **105**.

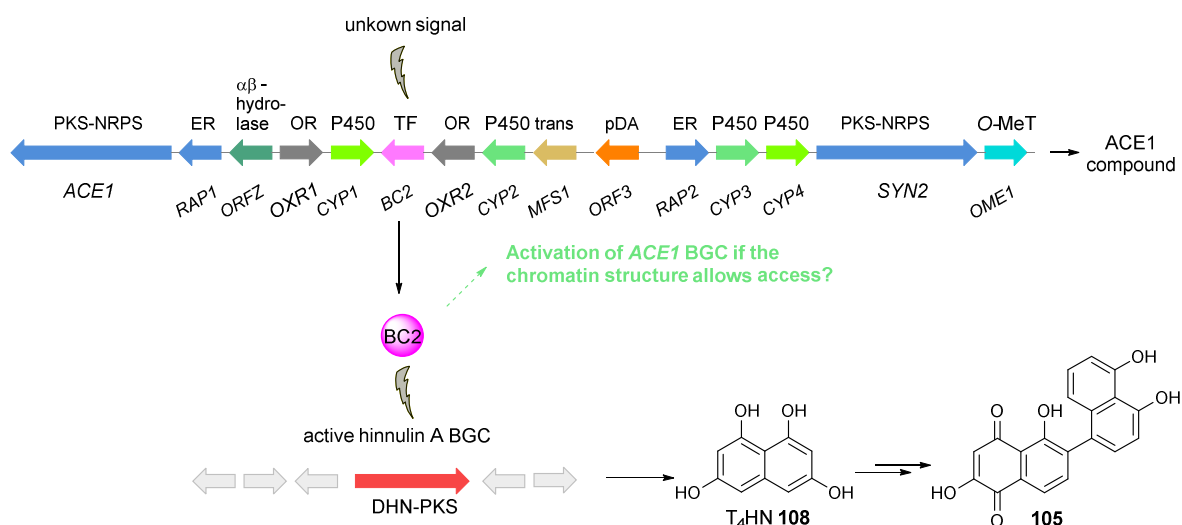
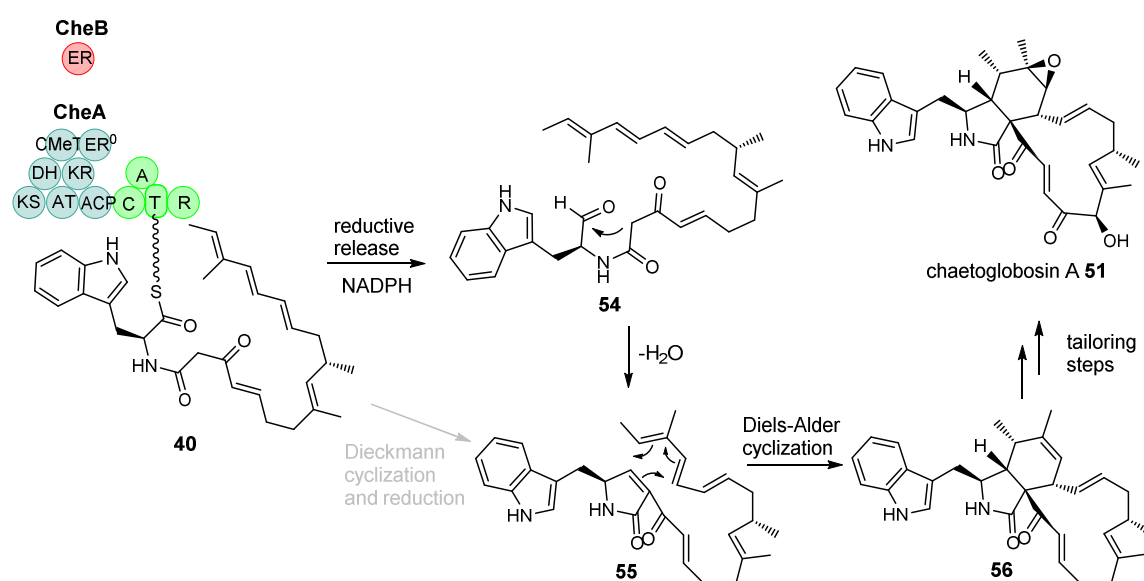


Figure 5.13: Proposed interaction between the transcription factor BC2 from the *ACE1* BGC and the hinnulin A **105** BGC in *P. oryzae* Guy11.

6. Investigations of the $\alpha\beta$ -Hydrolases ORFZ and PylE

6.1 Introduction

All cytochalasan BGC have a common gene that encodes an $\alpha\beta$ -hydrolase. However, the precise function of this hydrolase during the biosynthetic pathway is still unknown, but it is likely to be involved in the early/middle steps of biosynthesis. Based on the proposed biosynthetic pathway of chaetoglobosin A **51**, a biosynthetic route for cytochalasans was proposed (Scheme 6.1).⁷⁶ Within this pathway, a reductive off-loading mechanism was proposed due to the presence of a reductive-release domain at the C-terminus of the NRPS, as observed in the biosynthesis of fusarin C **8**.⁵⁰ The released aminoaldehyde **54** might undergo a Knoevenagel condensation to form a pyrrolinone **55**, which could function as a dienophile and subsequently undergo an intramolecular Diels-Alder (DA) reaction to form the core structure of cytochalasans. A second route was proposed where the release domain acts as Dieckmann cyclase as observed during the biosynthesis of tenellin **9**.⁵⁸ But this would require a later reduction.



Scheme 6.1: Proposed early biosynthetic steps of chaetoglobosin A **51**.⁷⁶

During the biosynthesis of cytochalasans it is unknown whether the formation of the pyrrolinone **55** occurs spontaneously or is enzyme catalysed. To this end, one aim of this work was to elucidate the function of the $\alpha\beta$ -hydrolase ORFZ from the *ACE1* biosynthetic gene cluster (BGC) in *P. oryzae* Guy11 and its homolog PylE from the pyrichalasin H **50** BGC in *M. grisea* NI980. *ORFZ/pylE* belong to the core gene cluster of cytochalasans (in addition to PKS-NRPS, enoyl reductase and putative Diels-Alderase) and are believed to play an essential role in the formation

of the core structure of the ACE1 metabolite and **50**, respectively. However, no strong evidence for their role has been disclosed.

6.1.1 The Superfamily of $\alpha\beta$ -Hydrolases

The superfamily of $\alpha\beta$ -hydrolase-fold proteins is a large group of structurally related enzymes containing highly preserved catalytic residues. On the sequence level they usually do not share any significant sequence similarity. In addition, they often do not catalyse the same reactions or react on the same substrates, making it hard to predict the catalytic function of $\alpha\beta$ -hydrolases.

The canonical $\alpha\beta$ -hydrolase fold consists of a mainly parallel eight-stranded β -sheet in the order 12435678 which is surrounded by six α -helices forming α -turn- β motifs (only strand β_2 is antiparallel, Figure 6.1).^{168–171}

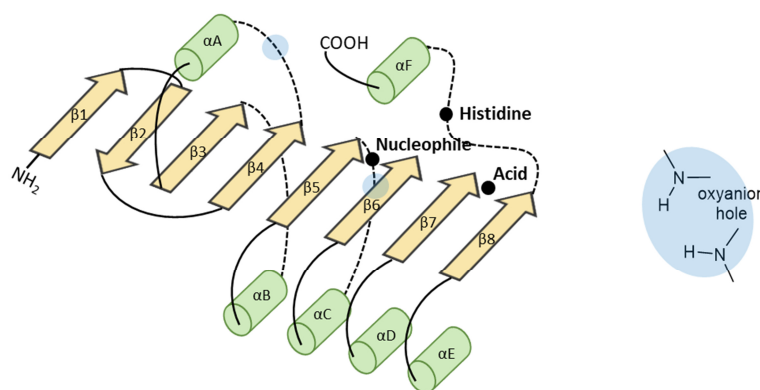


Figure 6.1: Canonical $\alpha\beta$ -hydrolase fold.¹⁷²

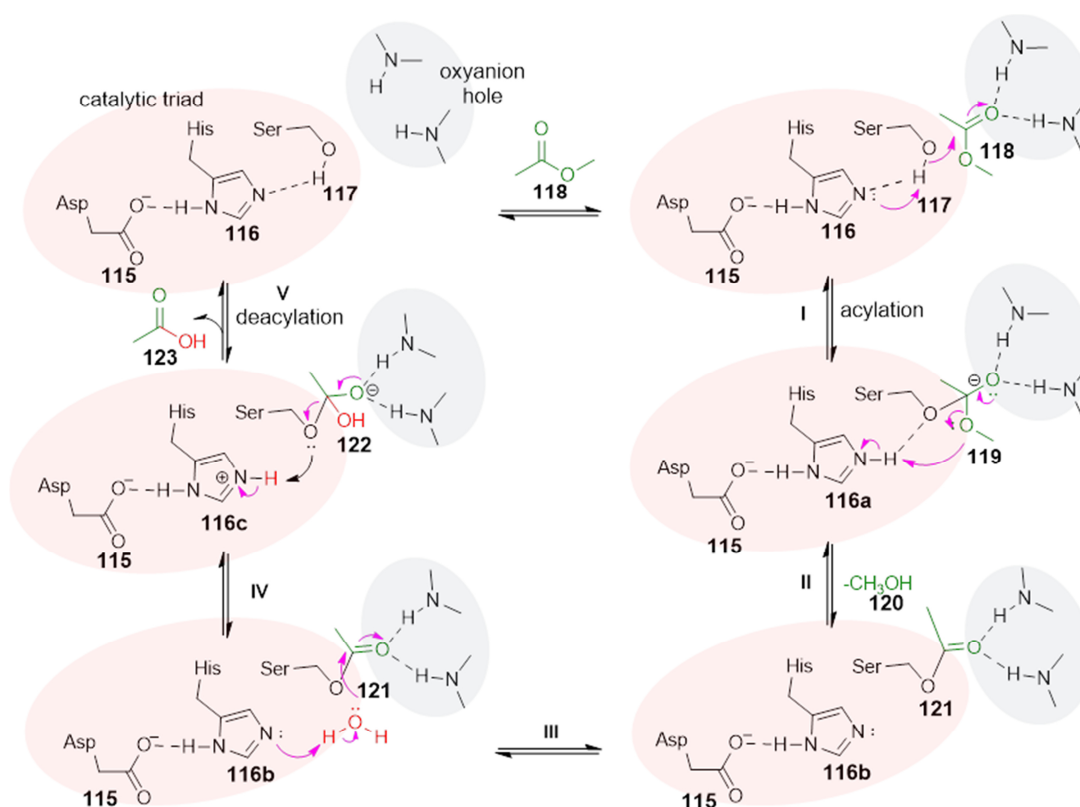
In its three dimensional folded structure, the first and the last strand are orthologous to each other due to a left-handed super-helical twist.^{172,168,169,173,174} The degree of twisting, as well as the number and position of the α -helices can differ, in order to accommodate different substrates. However, α C seems to play a strategic role for the positioning of the nucleophile. Insertions and/or deletions of residues especially shape the substrate-binding site. These modifications range from minor changes of only a few residues to the insertion of whole domains, working as a movable lid that seals the catalytic pocket.

The $\alpha\beta$ -hydrolase fold provides a robust scaffold for the catalytic triad; a nucleophile positioned after strand β_5 (serine, cysteine or aspartate), an acid moiety mostly located after β_7 (aspartate or glutamate) and a highly conserved histidine residue located after the last β strand (Figure 6.1).¹⁶⁸ Furthermore, the architecture of the fold leads to the formation of an 'oxyanion hole' formed by two main-chain nitrogen atoms to stabilize the negatively charged tetrahedral-

intermediate that occurs during hydrolysis.¹⁷⁵ The nucleophile is located in a sharp turn, the 'nucleophilic elbow', and embedded within the conserved consensus sequence 'Gly-X-Nu-X-Gly' (X = any residue, Nu = nucleophile). The geometry of this nucleophilic elbow forces the nucleophile residue to obtain energetically unfavourable main chain torsion angles, inducing steric restriction to adjacent residues and contributing to the formation of the oxyanion hole. The first backbone nitrogen of the oxyanion hole is located next to the nucleophile, the second between strand β 5 and helix α A (Figure 6.1, highlighted in blue).^{168,175}

6.1.2 The Catalytic Triad

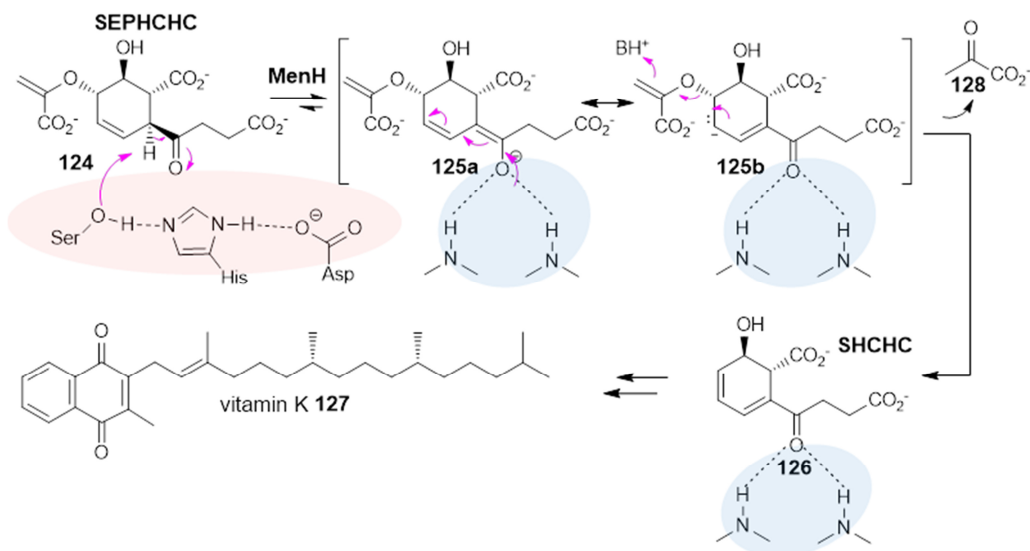
Since the underlying principles are comparable, the general mechanism of catalytic hydrolysis can be explained on the example of a serine-esterase (Scheme 6.2).¹⁷⁶ To increase the reactivity of the nucleophile **117**, the proton of the hydroxyl group (in the case of Ser as nucleophile) is extracted by a conserved histidine residue **116**. The histidine residue **116** is forced into the correct conformation by an acid moiety **115**, which additionally stabilises its protonated form (Scheme 6.2). The activated nucleophile **117** is now able to attack the carbonyl carbon of its substrate **118**, forcing the carbonyl oxygen to accept a lone electron pair forming a negatively charged tetrahedral-intermediate **119**, which is stabilized by the oxyanion hole (Scheme 6.2, I). The negative charge of the oxygen collapses back to the carbonyl, ejecting one part of the substrate **118** as an alcohol **120** which leads to an acyl-enzyme-complex **121** (Scheme 6.2, II). Ejection of the leaving group is often supported by donation of a proton by histidine **116a**. Next (i) a water molecule (hydrolysis) or (ii) a nucleophile (transesterification) attacks the acyl-intermediate **121** forming a second tetrahedral-intermediate **122** (Scheme 6.2, III + IV), which resolves by ejecting the enzyme-bound nucleophile **117**. Thereby, the free enzyme and the product **123** are released (Scheme 6.2, V).



Scheme 6.2: General esterase mechanism.¹⁷⁶

Apart from the traditional hydrolytic function, the catalytic triad of some enzymes of the $\alpha\beta$ -hydrolase superfamily acts as a general base, such as in the biosynthesis of vitamin K **127**.¹⁷⁷ The $\alpha\beta$ -hydrolase MenH of the vitamin K **127** BGC deprotonates SEPHCHC **124** ((1*R*,2*S*,5*S*,6*S*)-2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate) leading to the elimination of pyruvate **128** to form SHCHC **126** (1*R*, 6*R*-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate). It was proposed that MenH abstracts a proton from the α -carbon of the succinyl carbonyl group to generate an enolate intermediate **125a** followed by elimination in an E1cb mechanism of the enolpyruvyl group (Scheme 6.3).^{178–180}

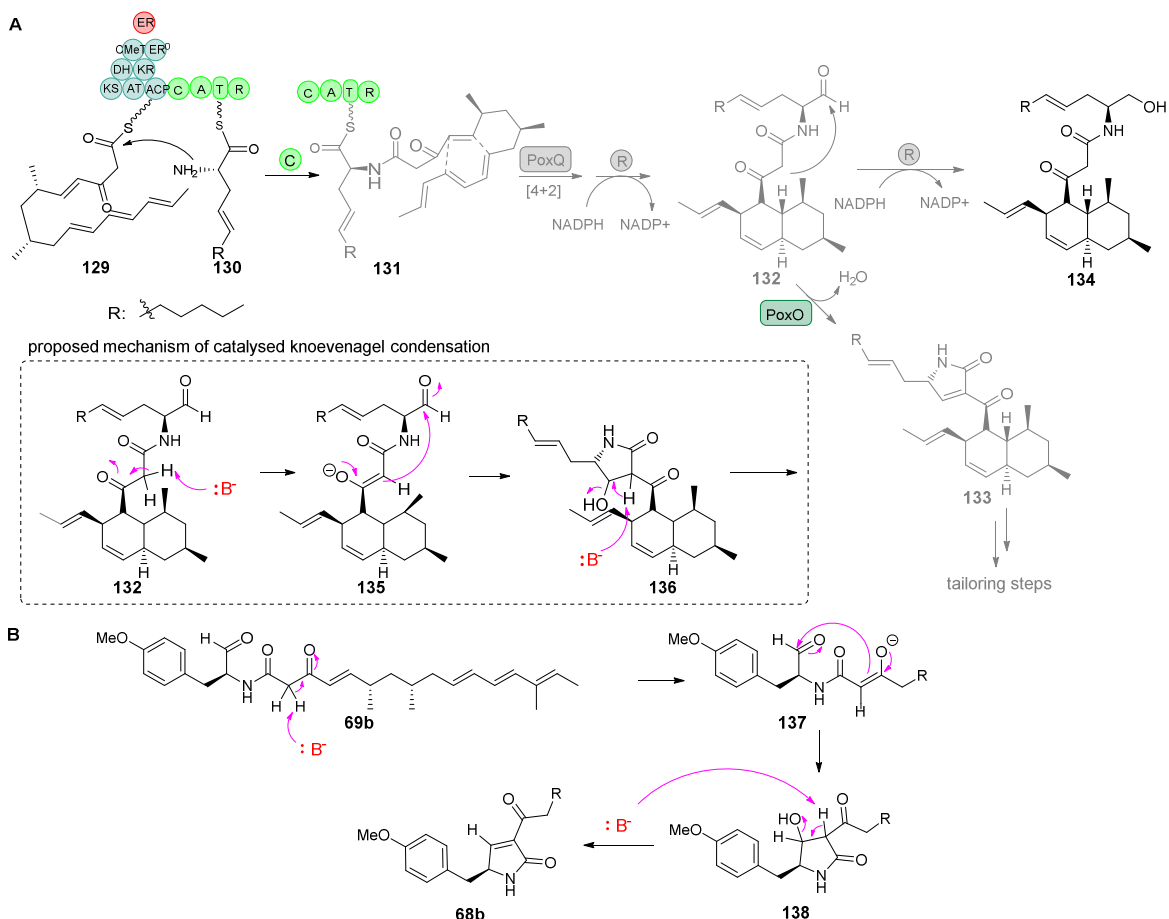
Formation of the active catalytic triad and oxyanion hole is often associated with conformational changes induced by substrate-enzyme interactions: the free enzyme is inactive, enabling high substrate specificity and a lack of hydrolytic activity towards non-native substrates.¹⁷⁷



Scheme 6.3: Early steps of the vitamin K **127** biosynthetic pathway.¹⁷⁷

Recently, Tang and co-workers demonstrated that an $\alpha\beta$ -hydrolase (encoded by *poxO*) from a PKS-NRPS (*pox*) gene cluster in *Penicillium oxalicum* might catalyse a Knoevenagel condensation to form compound **133** which is a proposed intermediate during the biosynthesis of oxaleimides (Scheme 6.4 A).¹⁸¹ In a *poxO* knockout strain they identified compound **134**, which might be a shunt product of the expected aldehyde **132**. The authors hypothesized that **132** is reduced, probably by the R-domain of the NRPS or by a native host enzyme as observed in *A. oryzae* (see Section 4), to the shunt product **134**. However, the aldehyde **132** was not observed. PoxQ might, similar to MenH, serve as a base, extracting the acidic α -hydrogen of the diketone moiety of **132**, thus forming an enol intermediate **135** (Scheme 6.4 A). Compound **135** might then react intramolecularly with the aldehyde moiety, which subsequently results in a base-induced elimination of water forming **136**. However, beside the deletion experiment of *poxO*, no further experimental evidence was obtained to verify that PoxO catalyses the Knoevenagel condensation. However, a biosynthetic pathway to the formation of **133** was suggested by the authors (Scheme 6.4).

ORFZ and PylE could react in a similar way as PoxO. In the case of the biosynthesis of **50**, PylE might catalyse the Knoevenagel reaction to convert **69b** into **68b** (Scheme 6.4 B). PoxO is homologous to ORFZ (49 % identity, 64 % similarity) and PylE (47 % identity, 62 % similarity), suggesting a similar function.



6.2 Aim of the Studies

The first aim of the study was to determine whether the $\alpha\beta$ -hydrolase PylE from the pyrlichalasin H **50** BGC in *M. grisea* NI980 is involved in the biosynthesis of **50** and to elucidate its function. To this end, it was intended to generate Δ *pylE* *M. grisea* deletion strains. In addition, complementation studies with a functional copy of *pylE* and a copy of *pylE* carrying a point mutation in its (potential) catalytic site were planned. Moreover, complementation studies with ORFZ from the ACE1 BGC of *P. oryzae* Guy11 were planned, thereby demonstrating that the biosynthesis of the ACE1 metabolite and **50** are closely related.

In addition, the function of ORFZ should be determined *in-vitro* using heterologous production in *E. coli*. Regarding the assessment of enzyme activity *in-vitro*, it was intended to use a synthetic substrate analogue as a substrate for ORFZ.

6.3 Results

Notably, very high titres of pyrichalasin H **50** could be produced by *M. grisea* NI980 (up to 60 mg/L, Figure 6.2). The pyrichalasin BGC was therefore used for targeted knockout (KO) experiments to elucidate the function of the $\alpha\beta$ -hydrolase PyiE.

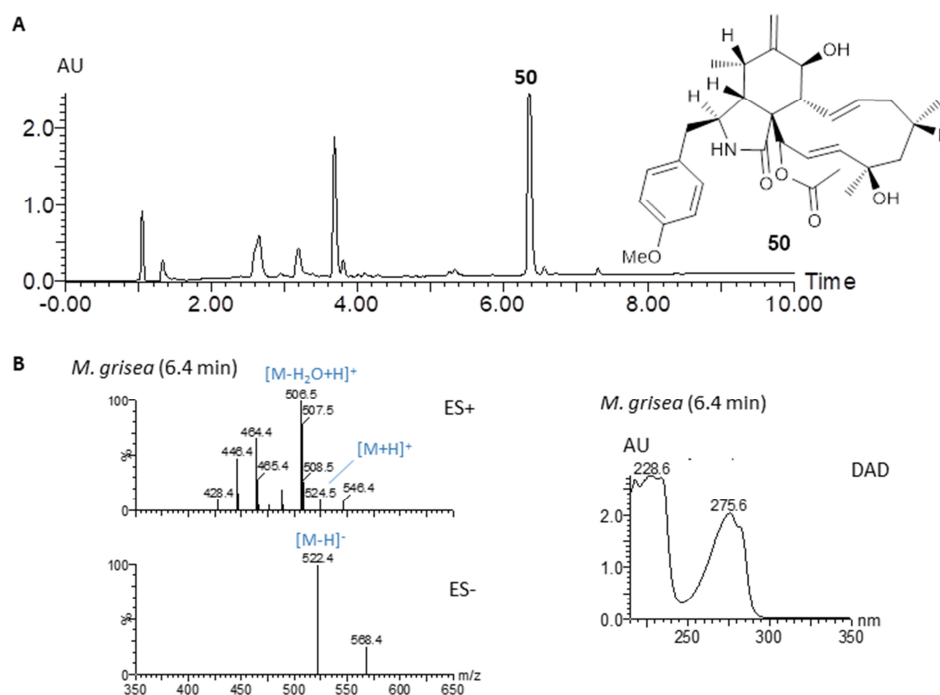


Figure 6.2: Organic extract of *M. grisea* NI980 wild type: **A**, LCMS chromatogram (DAD trace); **B**, ES and UV absorption spectrum of **50**.

6.3.1 Targeted Gene Deletion of *pyiE*

In order to investigate the function of the putative $\alpha\beta$ -hydrolase PyiE from the Pyrichalasin H **50** BGC in *M. grisea*, the gene *pyiE* was disrupted using the bipartite method (Section 1.9.1). The gene deletion cassette (pE-YA-*pyiE*-KO) was constructed by *in-vivo* homologous recombination in *S. cerevisiae* using a hygromycin B resistance cassette as selection marker (encoding a hygromycin B phosphotransferase gene, *hph*). DNA fragments for assembly were amplified by PCR. For amplification of *pyiE* *M. grisea* NI980 gDNA was used as template. The selection marker was amplified from vector DNA (pTH-GS-*egfp*). The 5' and 3' bipartite substrates for *pyiE* were amplified from the constructed vector DNA (pE-YA-*pyiE*-KO, Figure 6.3 A, Section 9.4.3).

Transformation of *M. grisea* NI980 was performed with the constructed plasmid using a PEG-mediated protoplast protocol. Eighteen transformants were selected on hygromycin B agar plates and cultivated for 7 days in DPY media (25 °C, 110 rpm). Subsequently, cells and growth media were individually extracted twice with ethyl acetate and analysed by LCMS (Section 9.5). In five

transformants (VBI09 -3, 4, 5, 12 and 13) production of **50** was almost completely abolished (Figure 6.3 B). Genetic analysis revealed that in the transformants VBI09-4, 5, 12 and 13 the complete hygromycin B resistance cassette was integrated correctly into *pyiE*, whereas in the case of VBI09-3 only the right hand part was integrated (Figure 6.3 B).

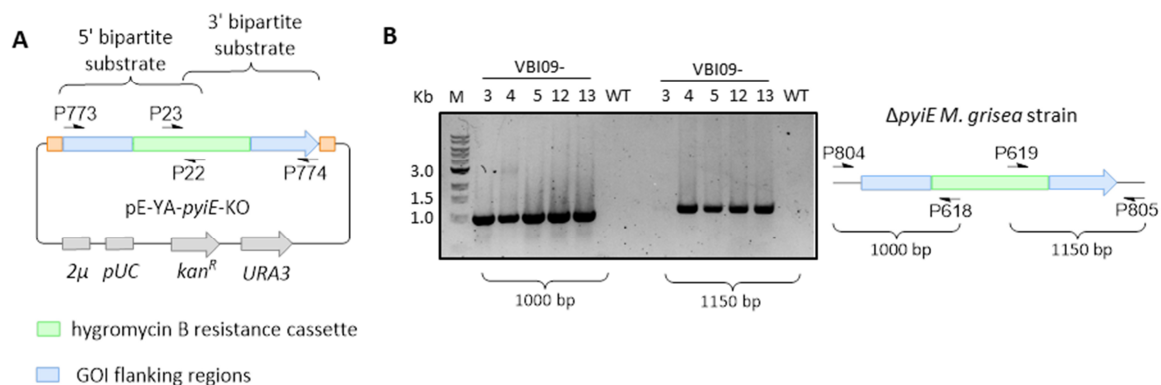
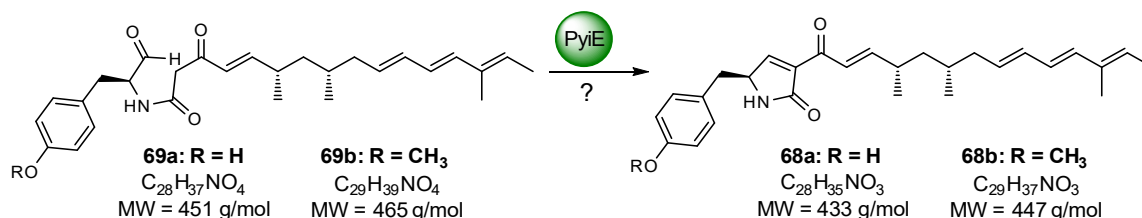


Figure 6.3: Genetic analysis of *pyiE* targeted knockout transformants; **A**, Constructed *pyiE* KO vector (VBI09) for amplifying the 5' and 3' bipartite substrates for *pyiE*; **B**, Genetic analysis of transformants obtained with bipartite marker strategy by PCR. GOI = gene of interest, M = marker.

In all Δ *pyiE M. grisea* strains (VBI09 -3, 4, 5, 12 and 13) production of **50** was almost completely abolished, suggesting an essential role of P*yiE* in the biosynthesis of **50** (Figure 6.4). Only traces of **50** were observed. Unfortunately, no metabolic precursor(s) of **50** could be identified. Plausible intermediate would have been aldehyde **69a** (derived from tyrosine **70**, $m/z = 451$ g/mol) and/or **69b** (derived from *O*-methyltyrosine **71**, $m/z = 465$ g/mol, Scheme 6.5). Since the aldehyde **69a/b** is very reactive, reduced shunt metabolites of this intermediate were likely to be expected. Alcohol shunt intermediates are known to be stable in *A. oryzae* (see 3.3.2), but it is possible that produced (alcohol) shunts or intermediates are degraded in *M. grisea*. Alternatively P*yiE* might be involved in offloading complete or aberrant intermediates from the PKS-NRPS. In this case no biosynthetic intermediates of **50** would be expected in the organic extracts.



Scheme 6.5: Proposed P*yiE* substrate: aldehyde **69a** and/ or **69b**.

Notably, in addition to **50**, small amounts of a new compound with $m/z = 493$ g/mol were detected in the Δ *pyiE M. grisea* strains.

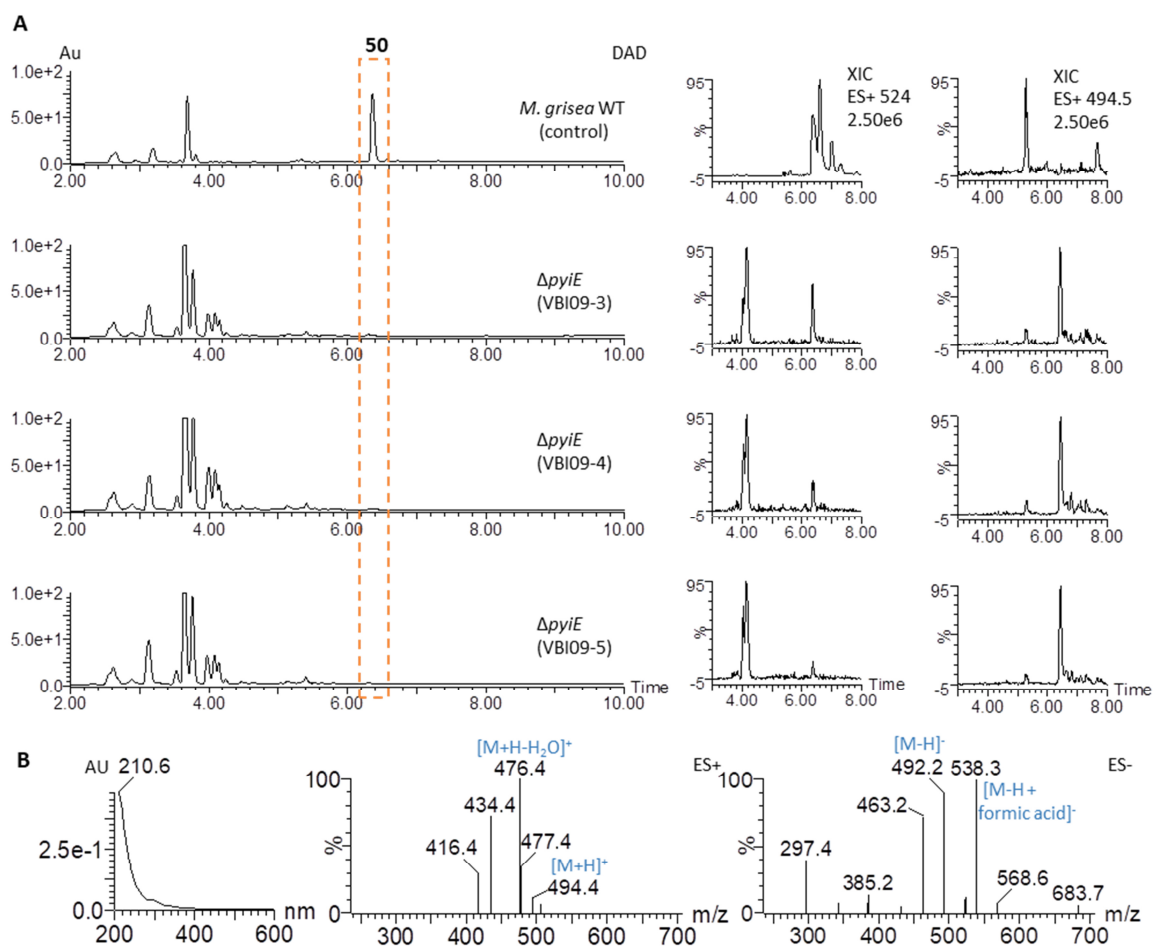


Figure 6.4: Chemical analysis of $\Delta pyiE$ *M. grisea* transformants: **A**, LCMS chromatograms (DAD trace) of organic extracts of transformants VBI09-3, VBI09-4, VBI09-5 and a *M. grisea* wild type control + extracted ion chromatogram (XIC) using $[M+H]^+$ corresponding to **50** ($m/z = 524$ g/mol) and to **139** ($m/z = 494$ g/mol); **B**, ES and UV absorption spectrum of **139**.

0.4 mg of **139** was purified by preparative LCMS from the combined extracts of the mutants and the structure was elucidated by NMR (Figure 6.5, Section 11.1, Table 11.1.1). HRMS confirmed this structure by suggesting a molecular formula of $C_{30}H_{40}NO_5$ ($[M]H^+$ calculated 494.2909, found 494.2906). The observed compound **139** was revealed to be the already known cytochalasin H **139**, a phenylalanine analogue of **50**.

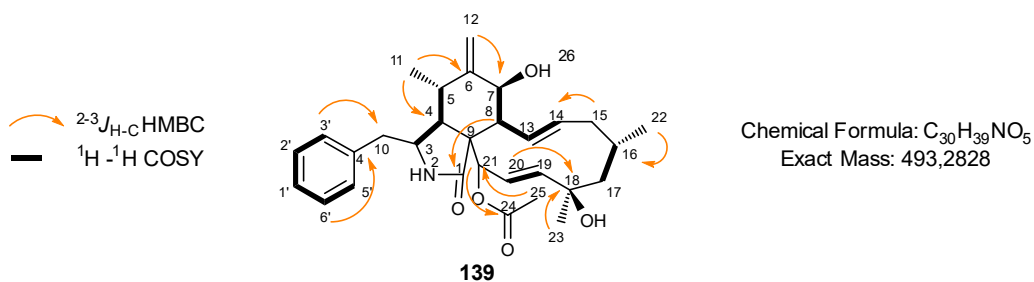


Figure 6.5: Isolated compound **139**.

Notably, formation of **139** in low titres was also observed by our group when the *pyiA* gene, encoding the *O*-methyltransferase from the pyrichalasin H **50** BCG, was disrupted.¹⁰⁸ The *pyiA* gene is located directly upstream of *pyiE* in the opposite orientation. The promoter of *pyiA* is located between *pyiA* and *pyiE*. Since the secondary metabolite profile of Δ *pyiE* *M. grisea* closely resembles the secondary metabolite profile of Δ *pyiA* *M. grisea* (*i. e.* extremely low production of **50** and observation of **139**) it is possible that the *pyiA* gene or its promoter region was damaged during the transformation procedure. To test if the *pyiA* gene is non-functional a control experiment could be performed. Addition of *O*-methyl-L-tyrosine **71** to growing Δ *pyiE* *M. grisea* cultures would complement the loss of *pyiA*. If the *pyiA* gene was really disrupted, this should lead to the true Δ *pyiE* *M. grisea* phenotype. Another control experiment would be to verify the expression of *pyiA* by RT-PCR. Another way to test if, apart from *pyiE*, other genes of the pyrichalasin BCG were disrupted is to perform complementation experiments. Expressing a functional copy of *pyiE* in Δ *pyiE* *M. grisea* strains should complement the loss of *pyiE* and restore the production of **50**.

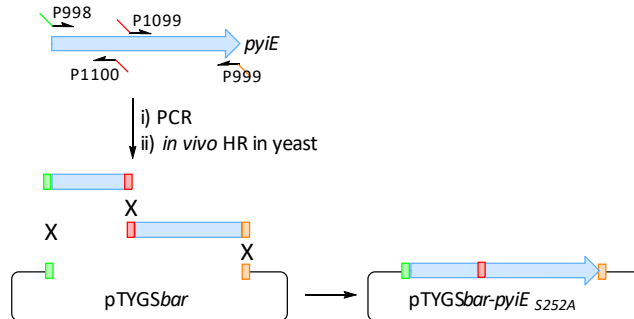
6.3.2 Complementation of *M. grisea* *pyiE* Deletion Strains

In order to test if the disruption of *pyiE* is really linked to the observed changes in biosynthesis of **50**, a functional copy of *pyiE* was re-introduced downstream of the inducible *amyB* promoter (P_{amyB}) and transformed into two Δ *pyiE* *M. grisea* strains (VBI09-4 and VBI09-5). In the same way, a *pyiE* gene carrying a point mutation in its potential catalytic site (PyiE_{S252A}) was introduced into the Δ *pyiE* *M. grisea* deletion strains. Heterologous complementation studies were also carried out to investigate if the closely related homologous protein ORFZ (49 % identity, 64 % similarity) from the *ACE1* BGC was able to restore production of **50**.

6.3.2.1 Construction of Complementation Vectors

For the complementation studies the respective gene of interest was cloned into the fungal expression vector pTYGS*bar* (which contains a glufosinate (BASTA) resistance cassette) downstream of P_{amyB} (Section 9.4.3). Amplification of *pyiE* was performed from *M. grisea* gDNA. 5' Extensions were generated using primers with specific overlaps, facilitating subsequent homologous recombination with the linearized (by the restriction enzyme *NotI*) pTYGS*bar* expression vector. Site-directed mutation of *pyiE* was generated by primer extension (Scheme 6.6). In addition to the primers generating overlaps to the vector sequence (P998 and P999), two internal primers with complementary ends (P1099 and P1100) were needed. These internal primers contained the desired base substitution. During PCR the mutation is

incorporated, replacing the original sequence. By homologous recombination both purified fragments were combined to the final product *pyiE*_{S252A} and incorporated into the vector pTYGSbar. The successful construction of the plasmid pTYGSbar-*pyiE*_{S252A} was confirmed by sequencing.



Scheme 6.6: Site-directed mutation of *pyiE* to *pyiE*_{S252A} by primer extension.

Amplification of *ORFZ* was performed from *P. oryzae* gDNA. The *ORFZ* gene was cloned into the entry vector pE-YA before it was transferred into pTYGSbar by *in-vitro* LR recombination. The purified DNA fragments were assembled into the linearized vector using homologous recombination in *S. cerevisiae*. Three fungal expression vectors were generated: pTYGSbar-*pyiE*, pTYGSbar-*pyiE*_{S252A} and pTYGSbar-*ORFZ* (Section 9.4.3)

6.3.2.2 Homologous Complementation with *pyiE* from the Pyrichalasin H BGC

Two Δ *pyiE* *M. grisea* strains (VBI09-4 and 5) were transformed with the expression vector pTYGSbar-*pyiE*. Ten transformants (VBI16-4-1 to VBI16-4-5 and VBI16-5-1 to VBI16-5-5) were generated, extracted and analysed regarding their secondary metabolite production by LCMS (Section 9.5). All strains were cultivated in starch containing complete media (DPY) to ensure the induction of *P_{amyB}*. Genomic DNA analysis revealed that a copy of *pyiE* was integrated into the genome of the transformants in all cases except VBI16-4-1 and -3 (Figure 6.6).

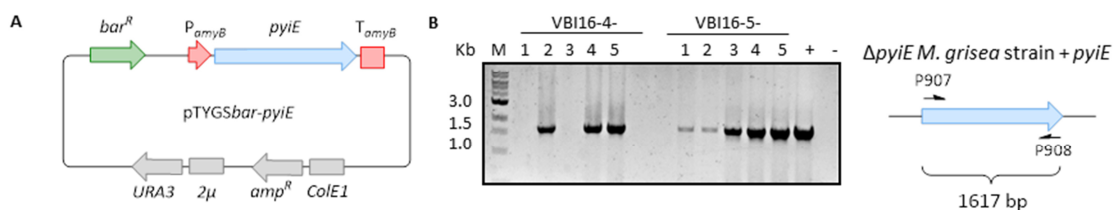


Figure 6.6: Genetic analysis of Δ *pyiE* + *pyiE* transformants: **A**, Complementation vector pTYGSbar-*pyiE*; **B**, Genomic DNA analysis using P907 and P908 (expected PCR fragment: 1617 bp), Δ *pyiE* *M. grisea* gDNA was used as negative control (-), vector DNA as positive control (+). M = marker.

As expected, production of **50** was restored to a certain level for all strains carrying a functional copy of *pyiE* (Figure 6.7). Nevertheless, compared to *M. grisea* wild type production (positive control) titers of **50** were still significantly decreased. Based on the UV absorption spectra using a single wavelength at 275 nm (characteristic UV absorption spectra of (*O*-methyl)tyrosine **71** derived compounds, see Figure 7.5) the yields of **50** of the transformants were compared relative to the yield in the *M. grisea* WT control strain using the integral of each peak (Figure 6.7, green numbers). Using the chromatographic peak integration the area under the chromatographic peak is measured and can be used to calculate unknown concentrations. During this study the area under the peak was used to approximately compare amounts of **50** in the Δ *pyiE* *M. grisea*, the *M. grisea* WT control strain and the transformants. The integral of each peak was determined using the Waters MassLynx™ Software. The amounts of **50** present in the *M. grisea* WT strain was set to 100 %. Compared to the WT levels, relative production levels of **50** in the Δ *pyiE* *M. grisea* correspond to 0.4 %. Compared to the amounts of **50** present in the Δ *pyiE* *M. grisea* WT strain (0.4 %) the yields of **50** in the transformant strains were significantly increased (production levels up to 10.9 % compared to the WT). Although the experiments were not designed for a quantitative analysis, the integrals provide enough evidence to conclude that production of **50** was restored compared to the Δ *pyiE* *M. grisea* control strain.

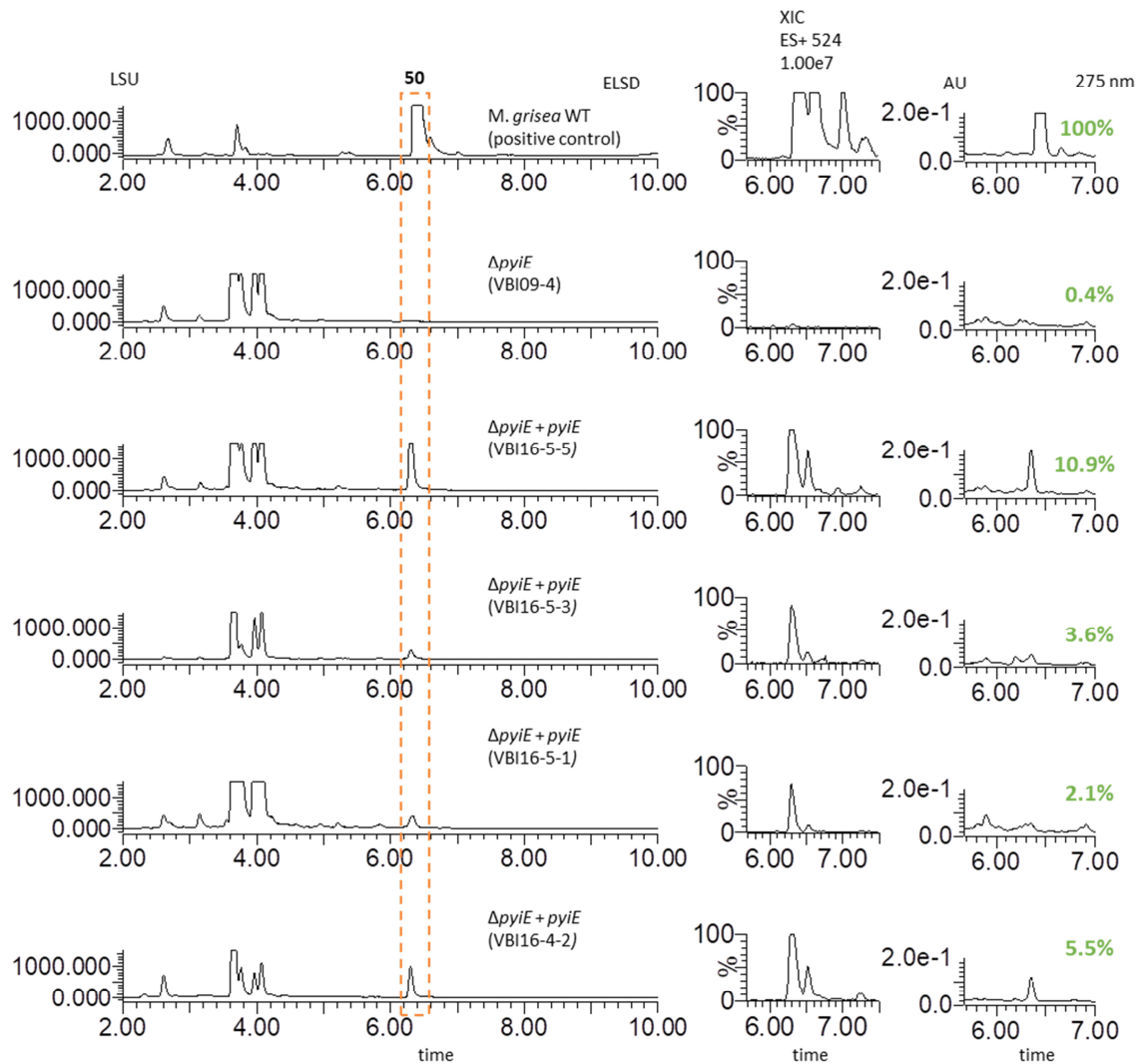


Figure 6.7: Homologous complementation of Δ *pyiE* *M. grisea* strains. LCMS chromatograms (DAD trace) and extracted ion chromatogram (XIC) using $[M+H]^+$ corresponding to **50** ($m/z = 524$ g/mol) of organic extracts. Based on the UV absorption spectra at 275 nm the yields of **50** of the transformants were compared relative to the yield in the *M. grisea* WT control strain (set 100 %) using the integral of each peak (green numbers).

Homologous complementation of Δ *pyiE* *M. grisea* strains with a functional copy of *pyiE* (under the control of P_{amyB}) restored the production of **50**, albeit the yield was decreased compared to the wild type strain. One reason for that might be that P_{amyB} is not as strong as the native promoter of *pyiE*. To address this, a copy of *pyiE* under the control of its native promoter was transformed into Δ *pyiE* *M. grisea* strains. For this purpose, the vector pTGYSbar- P_{native} -*pyiE* was constructed (Section 9.4.3).

One Δ *pyiE* *M. grisea* strain (VBI09-4) was transformed with the expression vector pTGYSbar- P_{native} -*pyiE*. Eleven transformants (VBI29-4-1 to VBI29-4-11) were generated, extracted and analysed regarding their secondary metabolite production by LCMS (Section 9.5). Genomic DNA analysis

revealed that a copy of P_{native} -*pyiE* was integrated into the genome of the transformants in all cases (Figure 6.8).

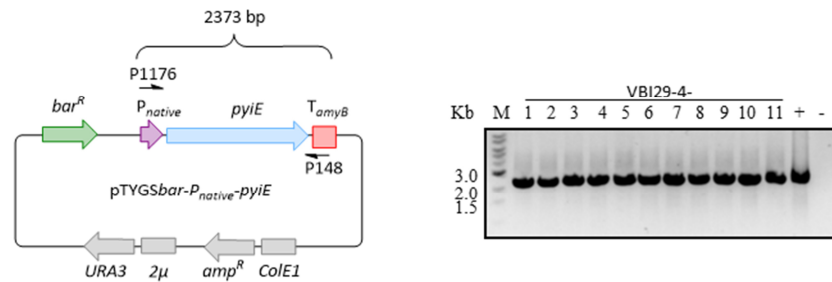


Figure 6.8: Genetic analysis of Δ *pyiE* + P_{native} -*pyiE* transformants: Left: complementation vector pTYGSbar- P_{native} -*pyiE*; right: genomic DNA analysis using P1176 and P148 (expected PCR fragment: 2373 bp), Δ *pyiE* *M. grisea* gDNA was used as negative control (-), vector DNA as positive control (+). M = marker.

Unexpectedly, production of **50** was not restored. Production levels of **50** in the transformants were again compared with the levels of **50** in the Δ *pyiE* *M. grisea* control strain using chromatographic peak integrals. Since the experiments were not designed to be quantitative, this method is just an estimation to determine if production levels of **50** are significantly changed. All organic extracts of the transformants displayed almost the same secondary metabolite profile as the Δ *pyiE* *M. grisea* control strain. Based on the UV absorption spectra using a single wavelength at 275 nm the yields of **50** of the transformants were compared relative to the yield in the *M. grisea* WT control strain (set 100 %) using the integral of each peak. Compared to the amounts of **50** present in the Δ *pyiE* *M. grisea* control strain (0.3 %), yields of **50** were not significantly changed in the transformant strains (production levels between 0.3 % - 0.9 %, Figure 6.9, green numbers). For an unknown reason the native promoter seemed to be inefficient, the *pyiE* gene was not expressed. To confirm this, RT-PCR analysis could be performed.

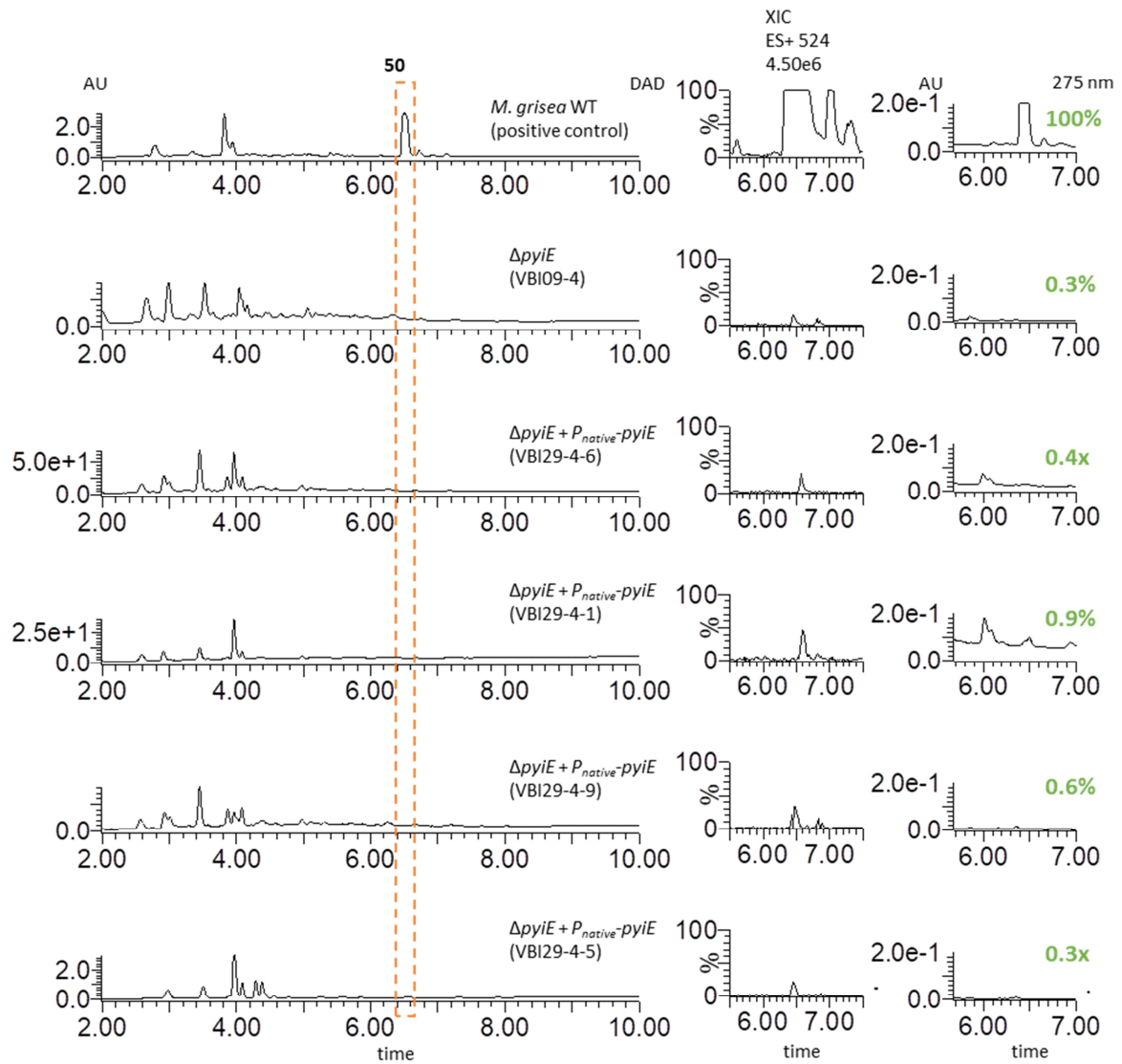


Figure 6.9: Homologous complementation of $\Delta pylE$ *M. grisea* strains with the native promoter of *pylE*. LCMS chromatograms (DAD) and extracted ion chromatogram (XIC) using $[M+H]^+$ corresponding to **50** ($m/z = 524$ g/mol) of organic extracts. Based on the UV absorption spectra at 275 nm the yields of **50** of the transformants were compared relative to the yield in the *M. grisea* WT control strain (100 %) using the integral of each peak (green numbers).

One explanation might be that unless the promoter-gene DNA fragment ($P_{native-pylE}$) is introduced into its original locus, gene expression is restricted. Genomic DNA analysis revealed that in all transformants the *pylE* deletion cassette is still present, indicating that the DNA fragment $P_{native-pylE}$ integrated ectopically (Figure 6.10).

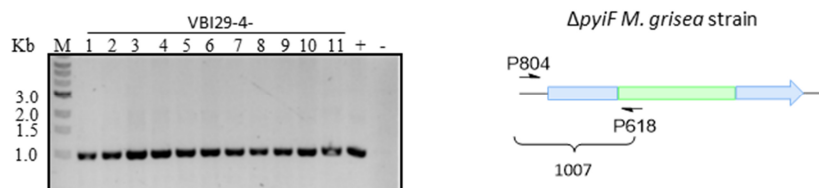
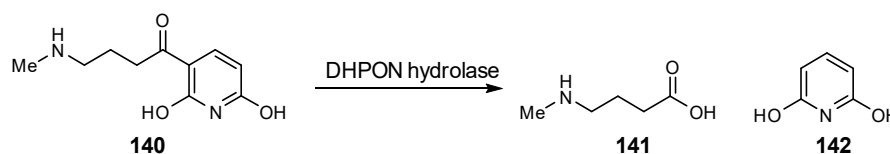


Figure 6.10: Screening for the *pyiE* deletion cassette in $\Delta pyiE + P_{native-pyiE}$ transformants. Genomic DNA analysis using P804 and P618 (expected PCR fragment: 1007 bp), *M. grisea* gDNA was used as negative control (-), $\Delta pyiE M. grisea$ gDNA as positive control (+). M = marker.

6.3.2.3 Homologous Complementation with *pyiE*_{S252A} Point Mutation

A typical feature of $\alpha\beta$ -hydrolase-fold proteins is the presence of a catalytic triad (Section 6.1.2). In order to identify catalytic residues in PyiE which might be part of the catalytic triad, a structural homology model of PyiE was generated using the PHYRE2¹⁸² server. The 2,6-dihydroxy-*pseudo*-oxynicotine (DHPON) hydrolase from the nicotine-degradation pathway of *Arthrobacter nicotinovorans* was identified as the closest (available) structural homolog with high confidence (100 %, PDB: 2JBW). The DHPON hydrolase splits DHPON **140** into γ -*N*-methyl-aminobutyrate **141** and 2,6-dihydroxy-pyridine **142** (Scheme 6.7).¹⁸³



Scheme 6.7: Enzymatic role of the DHPON hydrolase during the nicotine-degradation pathway of *A. nicotinovorans*.¹⁸³

Pairwise alignment of the amino acid sequences revealed an identity of 21 % and a similarity of 34 % between PyiE and DHPON hydrolase. Thus, based on their amino acid sequence, the homology of both proteins is rather low. In the absence of a better model, the DHPON hydrolase was used to identify catalytic residues in PyiE. The DHPON hydrolase contains a catalytic triad (Ser217, His329 and Asp300) and the associated oxyanion hole is formed by the amides 147-N and 218-N (Figure 6.11).¹⁸⁴

Based on the alignment, the amino acid Ser252 was identified in PyiE as possible candidate for the nucleophile of the catalytic triad (Figure 6.11, highlighted in pink).

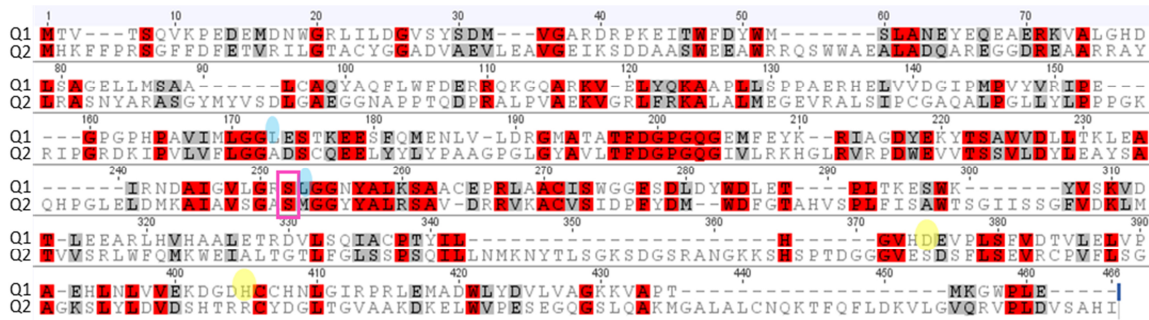


Figure 6.11: Structure-based alignment of DHPON hydrolase (Q1) with PylE (Q2). Yellow: Catalytic triad (Ser217, His329 and Asp300), blue: amides (147-N and 218-N), pink box: Serine of catalytic triad which is change to alanine in the sequence of PylE (PylE_{S252A}).

The constructed vector pTYGSbar-pylE_{S252A} was transformed into Δ pylE *M. grisea* strains (VBI09-4 and VBI09-5) using a PEG-mediated protoplast method. Twelve transformants (VBI28-4-1 to VBI28-4-6 and VBI28-5-1 to VBI28-5-6) were selected on BASTA agar plates, extracted and analysed for their secondary metabolite production by LCMS (Section 9.5). Genomic DNA analysis of eight transformants revealed that the pylE_{S252A} was integrated into the genome of the transformants VBI28-4-3/4 and VBI28-5-2/3/4 (Figure 6.12).

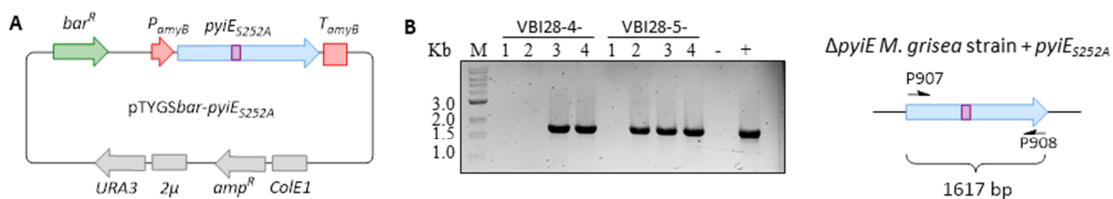


Figure 6.12: Genetic analysis of Δ pylE + pylE_{S252A} transformants: **A**, Complementation vector pTYGSbar-pylE_{S252A}; **B**, Genomic DNA analysis using P907 and P908 (expected PCR fragment: 1617 bp), Δ pylE *M. grisea* gDNA was used as negative control (-), vector DNA (VBI28) as positive control (+). M = marker.

Based on the UV absorption spectra using a single wavelength at 275 nm the yields of **50** of the transformants were compared relative to the yield in the *M. grisea* WT control strain (set 100%) using the integral of each peak. Compared to the amounts of **50** present in the Δ pylE *M. grisea* control strain (0.5%), yields of **50** were not significantly changed in the transformant strains (production levels between 0.3% - 0.6% of **50**, Figure 6.13, green numbers). Homologous complementation with native pylE under the same conditions led to increased yields (up to 10.9%) of **50** compared to the Δ pylE deletion control strain (Section 6.3.2.2). Therefore, the amino acid Ser252 might be part of the catalytic triad and therefore crucial for the catalytic properties of PylE. In addition, it could also be important for the structural architecture of the protein.

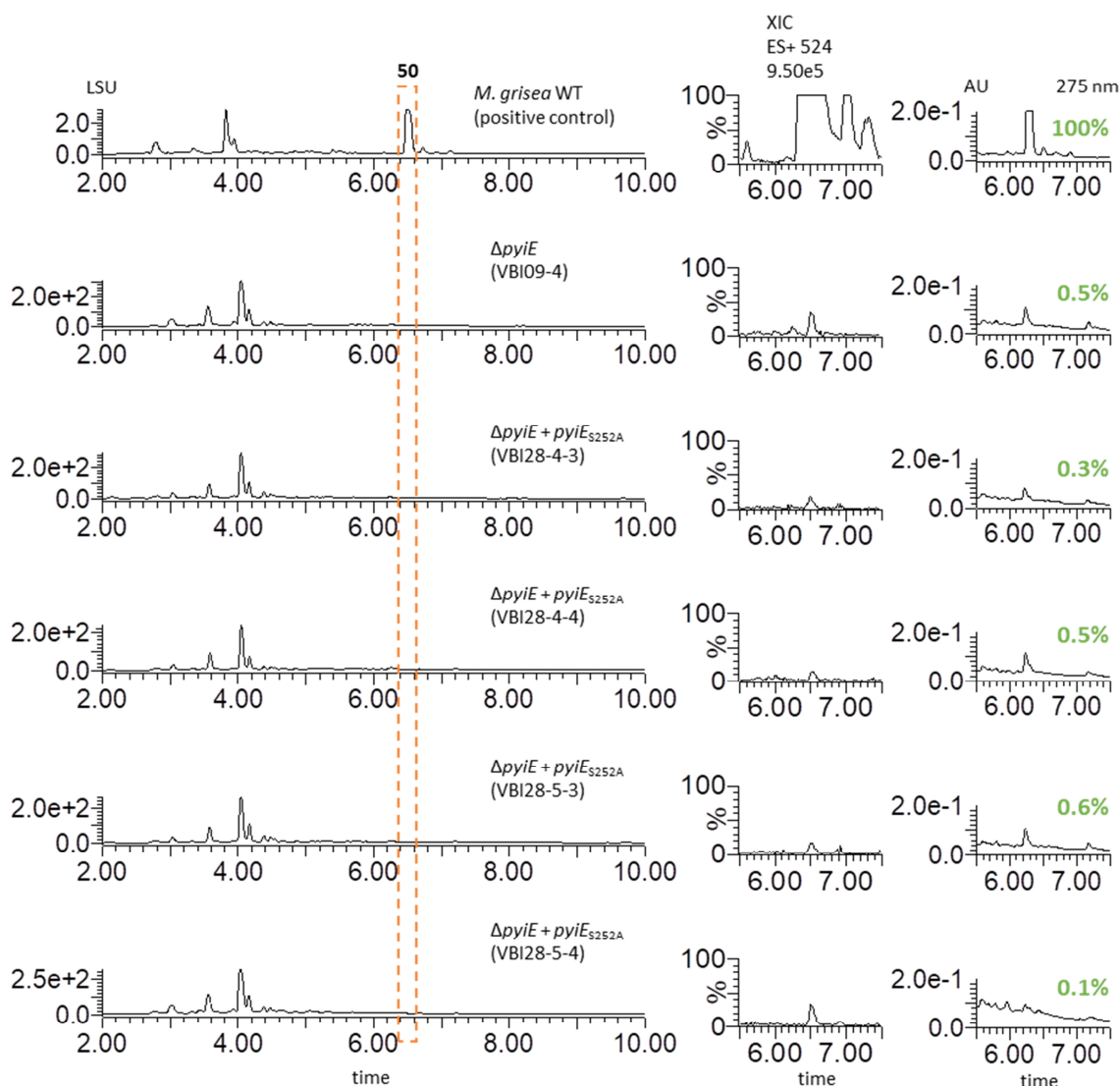


Figure 6.13: ELSD chromatograms of *M. grisea* $\Delta pyiE + pyiE_{S252A}$ complementation strains and extracted ion chromatogram (XIC) using $[M+H]^+$ corresponding to **50** ($m/z = 524$ g/mol) of organic extracts. Based on the UV absorption spectra at 275 nm the yields of **50** of the transformants were compared relative to the yield in the *M. grisea* WT control strain (set 100 %) using the integral of each peak (green numbers).

6.3.2.4 Heterologous Complementation of $\Delta pyiE$ *M. grisea* Strains with ORFZ

The $\alpha\beta$ -hydrolases PylE from the pyrichalasin H **50** BGC (*M. grisea*) and ORFZ from the ACE1 BGC (*P. oryzae* Guy11) exhibits 66 % sequence similarity (51 % identity). In order to test if ORFZ can complement the disrupted *pyiE* gene and restore efficient **50** production, the plasmid pTYGSbar-ORFZ was transformed independently into two *pyiE* *M. grisea* deletion strains (VBI09-4 and VBI09-5). Ten transformants (VBI13-4-1 to VBI13-4-6 and VBI13-5-1/ 2/ 3/ 5/ 6) were generated, extracted and analysed for their secondary metabolite production by LCMS. Genomic DNA analysis revealed that ORFZ was integrated into the genome of the transformants in almost all cases except VBI13-5-3 and 6 (Figure 6.14)

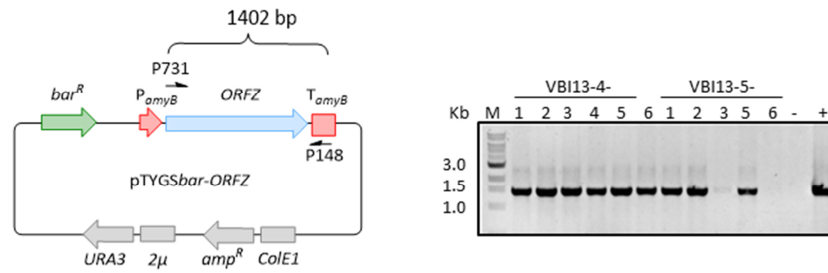


Figure 6.14: Genetic analysis of Δ *pyiE* + *ORFZ* transformants. Genomic DNA analysis using P731 and P148 (expected PCR fragment: 1402 bp), vector DNA was used as positive control (+), a Δ *pyiE* *M. grisea* gDNA was used as negative control (-). M = marker.

All strains carrying a functional copy of *ORFZ* produced comparable amounts of **50** with respect to the *pyiE* deletion strain VBI09-4 (negative control, Figure 6.15). Based on the UV absorption spectra using a single wavelength at 275 nm the yields of **50** of the transformants were compared relative to the yield in the *M. grisea* WT control strain (set 100 %) using the integral of each peak. Compared to the amounts of **50** present in the Δ *pyiE* *M. grisea* control strain (0.3 %), yields of **50** were not significantly changed in the transformant strains (production levels of **50** between 0.4 % – 0.6 %). Expression of *ORFZ* did not complement the loss of *pyiE*.

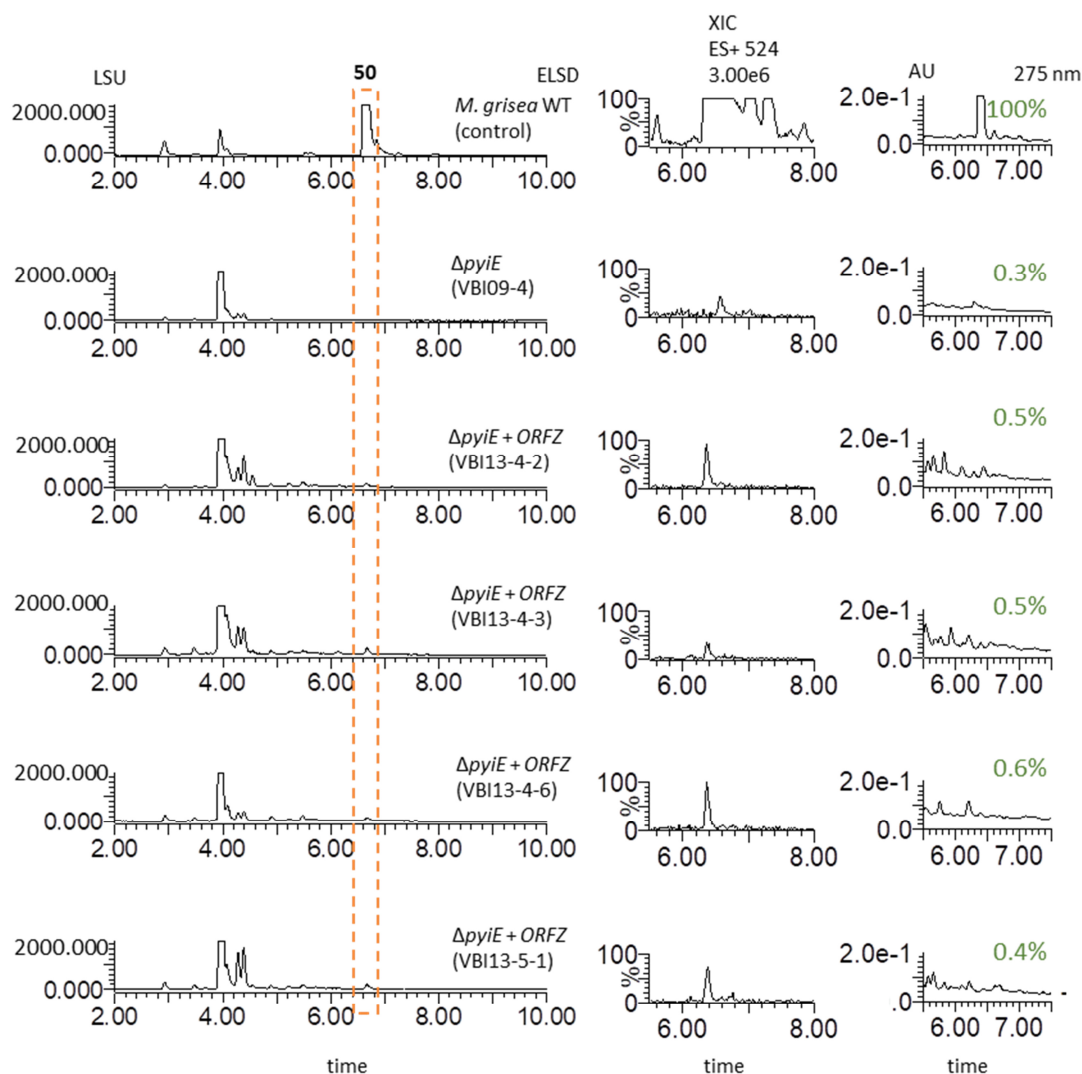


Figure 6.15: ELSD chromatograms of *M. grisea* Δ *pyiE* + ORFZ complementation strains and extracted ion chromatogram (XIC) using $[M+H]^+$ corresponding to **50** ($m/z = 524$ g/mol) of organic extracts. Based on the UV absorption spectra at 275 nm the yields of **50** of the transformants were compared relative to the yield in the *M. grisea* WT control strain (set 100 %) using the integral of each peak (green numbers).

6.3.3 *In-vitro* Studies with the $\alpha\beta$ -Hydrolase ORFZ

In parallel to the *in-vivo* studies of *pyiE* and ORFZ in *M. grisea*, the protein ORFZ from the ACE1 BGC of *P. oryzae* Guy11 was heterologously produced in *E. coli* in order to test its function by *in-vitro* enzyme assays. At the beginning of this work the pyrichalasin H **50** BGC was not available; therefore, the focus was on ORFZ.

A synthetic substrate **142** was provided by Haili Zhang (Leibniz Universität Hannover, Institute of Organic Chemistry, AK Cox), who started his synthesis in parallel to this work (Figure 6.16). LCMS analysis shows that this exists as an equilibrating mixture of hydrate **143** ($t_R = 3.8$ min) and aldehyde **142** ($t_R = 4.2$ min).

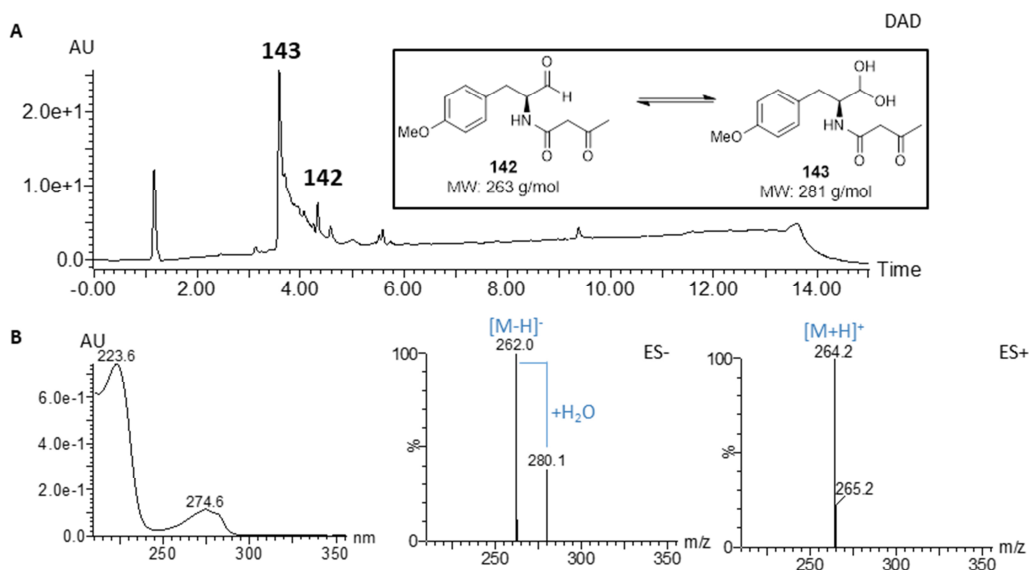
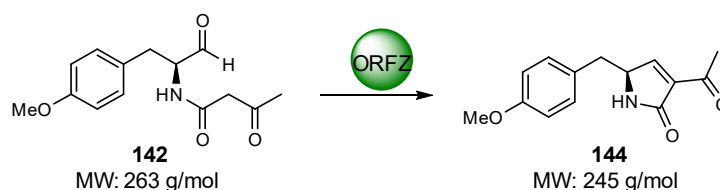


Figure 6.16: Aldehyde **142**: **A**, LCMS chromatogram (DAD trace); **B**, ES and UV absorption spectrum of **142**.

Based on the proposed biosynthetic role of ORFZ as an $\alpha\beta$ -hydrolase catalysing a Knoevenagel condensation, aldehyde **142** might be converted into pyrrolinone **144** (Scheme 6.8).



Scheme 6.8: Proposed catalytic role of ORFZ.

As a control, pyrrolinone **144** was also synthesised by Haili Zhang and used as positive control for the *in-vitro* enzyme assays (Figure 6.17).

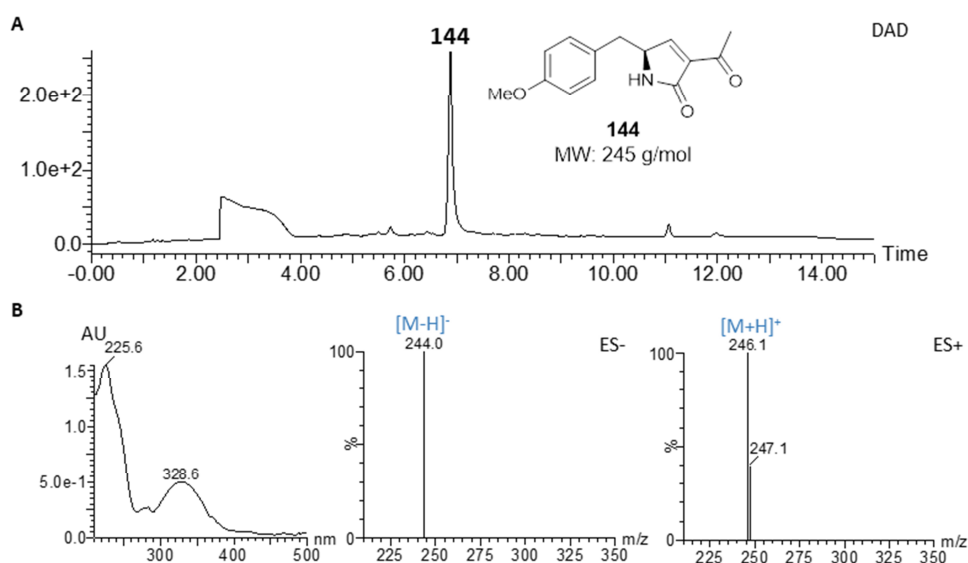


Figure 6.17: Pyrrolinone **144**: **A**, LCMS chromatogram (DAD trace); **B**, UV absorption spectrum + ES of **144**.

6.3.3.1 Cloning, Heterologous Production and Purification of ORFZ

The native *P. oryzae* ORFZ gene was cloned into the expression vector pET28a(+) via *Nde*I and *Xho*I restriction sites. The successful construction of the plasmid pET28a(+)-ORFZ was confirmed by sequencing. The plasmid was transformed into chemically competent *E. coli* BL21 (DE3) cells. Single colonies were used to inoculate 2TY medium containing 50 ng/ml of kanamycin. Cultures were grown to an OD₆₀₀ of 0.4 - 0.6 at 37 °C before adding IPTG to a final concentration of 0.1 - 1 mM according to experimental procedures. The cells grew at 25 °C for 18 h after addition of IPTG (Figure 6.18 A).

His₆-ORFZ was purified from the soluble protein fraction (supernatant) using a Ni-NTA chromatography. Elution was done by a stepwise imidazole gradient (20 mM- 500 mM). Each fraction was collected separately and analysed by SDS-Page and Coomassie staining (Figure 6.18 B). His₆-ORFZ has an expected molecular weight of 49.3 kDa. The purified protein was enzymatically digested and the peptide fragments were analysed by ESI Q-TOF (Dr. Jennifer Senkler, work group Prof. Braun, Institute of Plant Genetics, Leibniz University of Hanover) to confirm the correct sequence. The production of ORFZ could be verified with coverage of 65 % and a score of 4573.46 (Figure 6.18 C). Observation of peptide fragments corresponding to both the C- and N-terminus of the protein, suggest that the isolated protein is full-length.

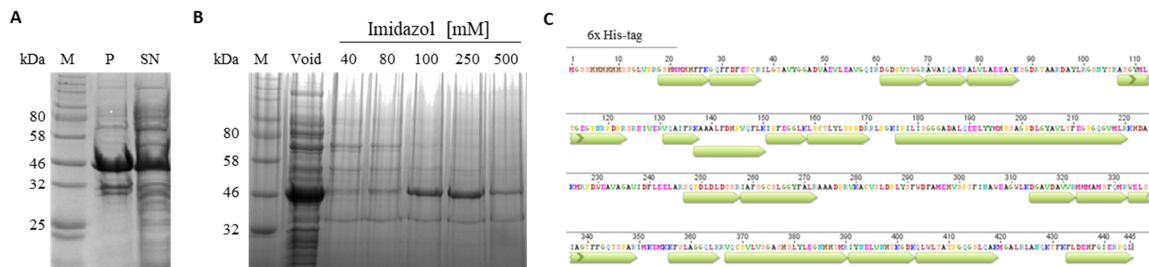


Figure 6.18: Production of recombinant His₆-ORFZ in *E. coli* BL21 and purification by Ni-NTA. Bands were visualized by Coomassie staining. M: marker (Colour Prestained Protein Standard, 11–245 kDa); **A**, The cells grew at 25 °C for 18 h after IPTG induction. P: protein extracts were prepared from cell pellets; SN: protein extracts were prepared from culture supernatant; **B**, His₆-tagged protein purification by Ni-NTA; protein sample was loaded onto Ni-NTA column, eluted by a stepwise imidazole gradient and analysed by SDS-PAGE; **C**, ESI Q-TOF analysis of ORFZ. Yellow arrows indicate identified peptide sequences.

6.3.3.2 Protein Purification using SUMO fusion systems

Recombinant protein production has to overcome many challenges including protein degradation, protein misfolding and low solubility. Enhanced protein production and solubility can sometimes be achieved using a SUMO fusion system. In eukaryotic cells, SUMO (small ubiquitin-related

modifier) proteins covalently bind to the lysine side chain of the target protein. SUMO proteases (two in yeast: ULP1 and ULP2) are able to identify and degrade those sumoylated proteins. SUMO modification was shown to play a role in many cellular processes such as nucleo-cytoplasmic transport and apoptosis.^{185,186}

Using yeast SUMO (SMT3), a SUMO fusion system was established by fusing the protein of interest at its N-terminus to a His₆-SUMO-tag. After protein purification by standard Ni-NTA, the SUMO protease 1 (Ulp1) can be used to cleave off the His₆-SUMO-tag resulting in a target protein with an untagged native N-terminus.^{187,188} Notably, this method allows the specific design of the N-terminus without the need of a methionine as the first amino acid.

ORFZ was cloned into the SUMO vector pCA528¹⁸⁹ and transformed into competent *E. coli* BL21 (DE3) cells. The successful construction of the plasmid pCA528-*ORFZ* was confirmed by sequencing. Protein production was obtained according to experimental procedures (Section 9.4.11).

The cell lysate was loaded onto a Ni-NTA column, washed with 20 mM imidazole wash buffer, eluted with elution buffer containing 250 mM imidazole and analysed by SDS-PAGE and Coomassie staining (Figure 6.19). His₆-SUMO-*ORFZ* has an expected molecular weight of 58.9 kDa. All elution fractions were combined and loaded onto a desalting column to remove imidazole. SUMO protease 1 (Ulp1) was used to cleave off the His₆-SUMO tag. Cleaved *ORFZ* was recovered in the flow-through fractions after a second Ni-NTA column whereas the His₆-SUMO and His₆-Ulp1 remained on the column (Figure 6.19 B + C).

The purified protein was enzymatically digested by trypsin before the peptide fragments were further analysed by ESI Q-TOF (Dr. Jennifer Senkler work group Prof. Braun, Institute of Plant Genetics, Leibniz University of Hannover). The production of *ORFZ* could be verified with coverage of 72.6 % and a score of 5088.02 (Figure 6.19 D). Again, observation of the C- and N-terminal peptides indicates that the protein is not truncated.

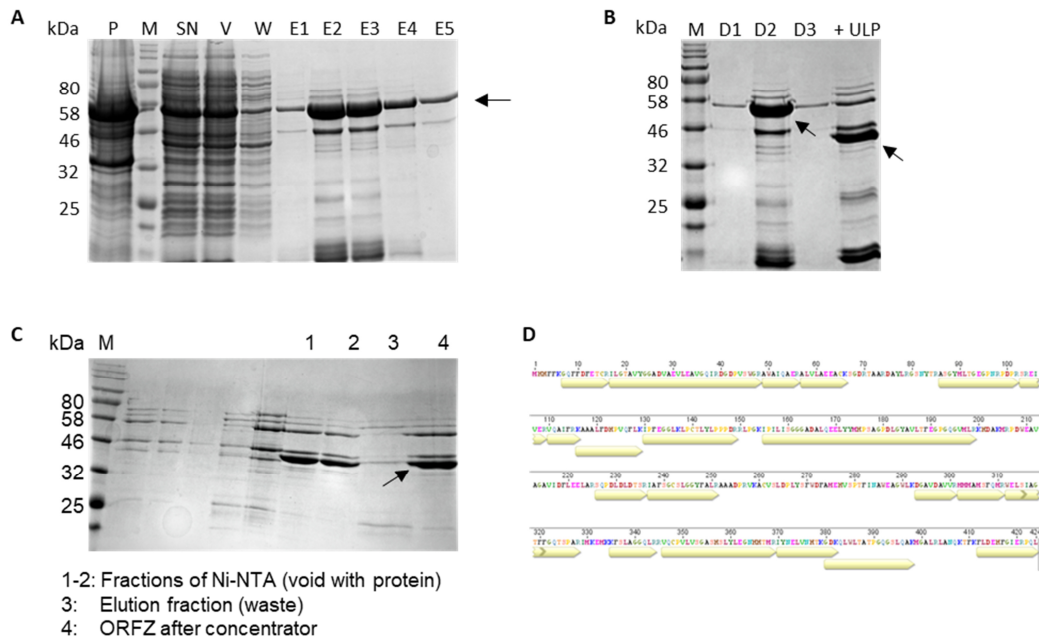


Figure 6.19: Purification of the SUMO-ORFZ fusion protein. Bands were visualized by Coomassie staining. M: marker (Colour Prestained Protein Standard, 11–245 kDa). The molecular weight of the target protein was about 46.9 kDa and of the SUMO fusion 58.9 kDa and is indicated by the black arrows. The cells grew at 16 °C for 18 h after IPTG induction; **A**, Purification of SUMO-ORFZ by Ni-NTA. Protein sample was loaded onto Ni-NTA column, washed with 20 mM imidazole wash buffer, eluted with elution buffer containing 250 mM imidazole and analysed by SDS-PAGE. P: protein extracts were prepared from cell pellets; SN: protein extracts were prepared from culture supernatant; V: void fraction, W: wash fraction; E1-E5: elution fractions; **B**, Analysis of SUMO-ORFZ fusion protein following cleavage with ULP SUMO protease; D1- D3: fractions of desalting column; +ULP (lane 5): SUMO-ORFZ cleavage products after incubation with ULP protease; **C**, second Ni-NTA to remove 6 x His-SUMO tag and 6 xHis-ULP protease from protein sample. ORFZ is now in the void fractions (lane 1 and 2), lane 3: elution fraction (250 mM inimidazol); lane 4: purified recombinant ORFZ; **D**, ESI Q-TOF analysis of ORFZ. Yellow arrows indicate identified peptide sequences.

6.3.3.3 *In-vitro* Studies with ORFZ

In order to examine ORFZ regarding the catalysis of the Knoevenagel condensation, an *in-vitro* enzyme assay was designed. Enzyme assays were performed under various conditions (pH 6 - 8, phosphate/ HEPES/ MOPS buffer, 25/ 30 °C, +/- 0.01 % triton-X) using the substrate analogue **142**. In all assays the negative control (no or boiled enzyme) demonstrated the same reaction outcome as the test tubes (+ 2.5 - 5 mM ORFZ). No activity of ORFZ was observed. Neither the pyrrolinone **144** nor other products of ORFZ were observed (Figure 6.20).

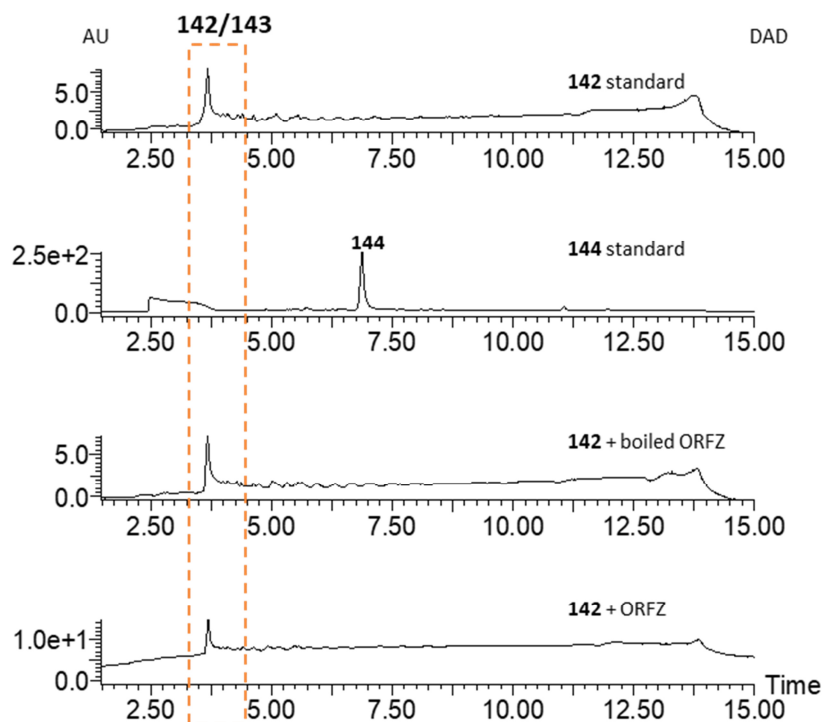


Figure 6.20: Example of enzyme assay. Buffer: HEPES pH 7, reaction conditions: 30 min at 30 °C.

Enzyme activity of ORFZ towards the synthetic substrate **142** was not observed. There are several explanations for this. One explanation would be that the protein was inactive. Heterologous production of proteins in *E. coli* can lead to misfolded, and therefore inactive, proteins. In addition, the protein can lose its activity during the purification procedure. In the absence of a suitable positive control, the activity of ORFZ could not be verified. Another reason might be that the substrate **142** was not accepted by ORFZ, since, compared to its putative native substrate **90**, it is lacking the main part of the polyketide chain. Thus, no activity of ORFZ was observed. Some enzymes demonstrate high substrate specificity; only the native substrate is accepted by the catalytic pocket.¹⁹⁰ A third possibility would be that ORFZ is not involved in the proposed Knoevenagel condensation and is therefore not able to catalyse this reaction. Collectively, the function of ORFZ remains elusive. Furthermore, the results suggest that the Knoevenagel condensation does not occur spontaneously under the tested conditions.

6.4 Discussion

The biosynthesis of pyrichalasin H **50** and other known cytochalasans requires the formation of a pyrrolinone ring in order to undergo an intramolecular Diels-Alder reaction to form the isoindole macrocycle **72**. The formation of the pyrrolinone **68b** is thought to be constructed by a Knoevenagel condensation of the released aminoaldehyde **69b**.

Disruption of the *pyiE* gene in *M. grisea* almost abolished production of **50**. Surprisingly, a new compound **139** with almost the same retention time as **50** ($t_R = 6.4$ min) was detected in traces in the deletion strains and identified as the already known cytochalasan H **139**, a phenylalanine analogue of **50**. Notably, disruption of *pyiA*, the gene encoding the *O*-methyltransferase from the **50** BCG, also led to the formation of **139** in low titres.¹⁰⁸ The *pyiA* gene is located directly upstream of *pyiE* in the opposite orientation, therefore the promoter of *pyiA* is located between *pyiA* and *pyiE*. Since the secondary metabolite profile of Δ *pyiE* *M. grisea* looks like the secondary metabolite profile of Δ *pyiA* *M. grisea* (*i. e.* extremely low production of **50** and observation of **139**) it is possible that the *pyiA* gene or its promoter region was damaged during the transformation procedure. To address this, feeding *O*-methyl-L-tyrosine **71** to Δ *pyiE* *M. grisea* cultures would ensure that it is selected by the PyiS A-domain. If the expression of the *pyiA* gene was really disrupted, supplementation with *O*-methyl-L-tyrosine **71** should reveal the true Δ *pyiE* *M. grisea* phenotype and might result in the observation of new pathway intermediates (or shunts) which can be linked to the function of PyiE.

In contrast to this hypothesis, homologous complementation of Δ *pyiE* *M. grisea* strains with a functional copy of *pyiE* (under the control P_{amyB}) restored the production of **50**, albeit the yield was dramatically decreased. One reason for that might be that P_{amyB} is not as strong as the native promoter of *pyiE*. To address this, a copy of *pyiE* under the control of its native promoter was transformed into Δ *pyiE* *M. grisea* strains. Unexpectedly, production of **50** was not restored. All organic extracts of the transformants displayed the same secondary metabolite profile as the Δ *pyiE* *M. grisea* control strain. For an unknown reason the native promoter seems to be inefficient. One explanation might be that, at least for the genetically tested transformants, the DNA fragment P_{native} -*pyiE* integrated ectopically into the genome of *M. grisea* and not at its original locus. Nevertheless, since production of **50** was restored in the first experiment (up to 10.9%; production levels in the Δ *pyiE* *M. grisea* deletion strain: 0.4%), it can be concluded (based only on this results) that apart from *pyiE* no other gene of the **50** BGC was disrupted. If this conclusion turns out to be true (*i. e.* the *pyiA* gene is correctly expressed) this would provide evidence that PyiE is involved in crucial early biosynthetic steps.

Homologous complementation of Δ *pyiE* *M. grisea* strains with a point mutation of *pyiE* (*pyiE*_{S252A}) displayed the same secondary metabolite profile as the deletion strains. Production of **50** was not restored. Thus, provided that the *pyiA* gene is functional, Ser252 must be crucial for the activity of PyiE and might be part of the catalytic triad; a typical feature of $\alpha\beta$ -hydrolase-fold proteins.

Heterologous complementation of $\Delta pyiE$ *M. grisea* strains with *ORFZ* did not restore the production of **50**.

One possible way of action of PyiE/ *ORFZ* would be that it somehow interacts with the PKS-NRPS substrate while it is still bound to the ACP. Recently it was demonstrated that the putative hydrolase CitA (15 % identity/ 26 % similarity to PyiE) from the citrinin BGC is able to remove ACP-bound acyl intermediates *in trans*; it appears to be a proof-reading hydrolase.¹⁹¹ Previous studies by Cox and co-workers demonstrated that deletion of *citA* in *Monascus ruber* substantially reduced the formation of the post-PKS aldehyde **145**.¹²⁹

Heterologous co-expression of the PKS encoded by *citS* and *citA* in *Aspergillus oryzae* resulted in significantly higher titres than from expression of PKS alone. However, the catalytic role of CitA remained unclear. It has recently been demonstrated *in-vitro* that addition of CitA to various combinations of PksCT (CitS) domains (minimal PKS, minimal PKS + C-methyltransferase, minimal PKS + C-methyltransferase + product template) resulted in a concentration-dependent decrease of the non-productive intermediates **147**, **148** (triketides), **149** (tetraketide), and **150** (pentaketide) as well as its native substrate aldehyde **145**. Decrease of **145** was suspected to be caused by inefficiencies of the synthetic construct (Figure 6.21).¹⁹¹

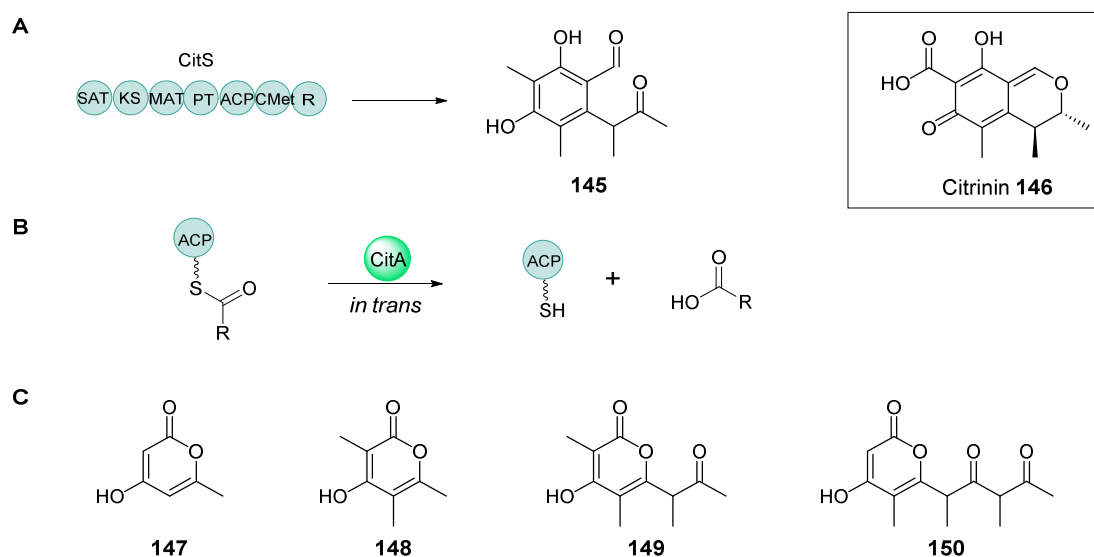


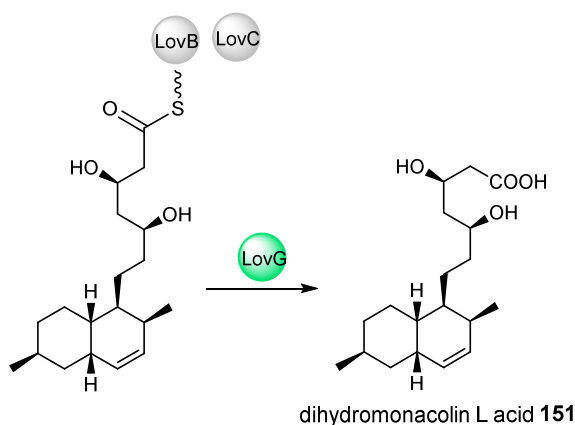
Figure 6.21: Proposed role of CitA in the biosynthesis of Citrinin **146**: **A**, CitS (PKS-NRPS) produces aldehyde **87**, which is transformed to citrinin **146** by several tailoring steps; **B**, proposed function of CitA (putative hydrolase): hydrolysis of ACP-bound acyl intermediates to free the ACP from stalled intermediates; **C**, Compounds of *in-vitro* studies.

In-vitro experiments using *N*-acetylcysteamine (SNAC) or CoA thioesters as substrate mimics for CitA exhibited only weak activity of CitA, suggesting that ACP is an essential component for CitA efficiency. Thereby the activity of CitA would be focussed on PKS-bound intermediates, leaving alone acyl-CoA compounds.¹⁹¹

CitA appears to work as a proof-reading hydrolase. P*yi*E and ORFZ might work in a similar way, since the loss of P*yi*E dramatically reduces production of **50** and no biosynthetic intermediates of **50** were observed. To address this, an *in-vitro* assay using ACP-bound substrate could be established to test whether P*yi*E/ ORFZ demonstrate activity under those conditions.

Usually, type I PKS have a C-terminal thioesterase (TE) domain which is associated with the product release.¹⁹² However, other proteins, including thioesterases, could also act *in-trans* to catalyse hydrolytic release. In recently categorized Group VII nr-PKS, a reductase (R) domain substitutes the TE domain and catalyses NADPH-dependent release of the acyl-bound product as an aldehyde as observed in the biosynthesis of citrinin **146**. An adjacent putative $\alpha\beta$ -hydrolase-encoding gene was identified in a number of BGC of Group VII PKS and demonstrated to be required for efficient product formation.^{193,129,194}

During the biosynthesis of lovastatin **2** in *A. terreus*, the hydrolase LovG (11.0 % identity/ 18.8 % similarity to P*yi*E) was identified to be responsible for product release from the lovastatin nonaketide synthase (LovB, Scheme 6.9). In Δ *lovG* mutant strains, production of **2** was dramatically reduced (< 1 % compared to the wild type). Since no simultaneous accumulation of pathway intermediates or shunts could be observed, it was proposed that LovG plays a crucial but not essential role in the early steps of the pathway. Furthermore, co-expression of *lovG* with *lovB* and *lovC* (*trans*-acting enoyl reductase) in *S. cerevisiae* resulted in the production of **151** confirming that LovG is the releasing partner of LovB in *A. terreus*.¹⁹⁵



Scheme 6.9: Early biosynthetic steps of lovastatin **2**. LovG is the releasing partner of LovB and LovC.

Similar to *pyiE*, *lovG* is located near the genes encoding the nonaketide PKS and the *trans*-acting enoyl reductase. Adjacent genes are often co-regulated by the same transcription factor, indicating an important biosynthetic role for P*yiE* and its homologs.¹⁹⁶

LovG releases its substrate as a carboxylic acid. However, during the biosynthesis of cytochalasans the released product is supposed to be an aldehyde. Therefore, it is unlikely that P*yiE* and ORFZ react in the same way as LovG. Nevertheless, it cannot be excluded that P*yiE* and ORFZ might interact with their corresponding PKS-NRPS and the enoyl reductase to increase efficient production of **50** and the ACE1 metabolite, respectively.

7. Investigations of the Putative Diels-Alderase

7.1 Introduction

Probably the most intriguing step during the biosynthesis of the ACE1 metabolite and pyrichalasin H **50**, as well as cytochalasins in general, is the formation of the isoindole core fused to a macrocycle, which is proposed to happen *via* an intramolecular Diels-Alder (DA) [4+2] cycloaddition. Whether this reaction is spontaneous or enzyme catalysed is unknown. To this end, one aim of this work was to elucidate the function of the putative Diels-Alderase (pDAase) ORF3 from the ACE1 BGC in *P. oryzae* Guy11 and its homolog PylF from the pyrichalasin H **50** BGC in *M. grisea* NI980.

The DA reaction takes place between a 1,3-diene (in *s-cis* conformation) and a moiety containing a double or triple bond (dienophile) resulting in a cyclic product *via* the formation of two new σ C-C bonds while two π C-C bonds are lost, an enthalpically favourable process. This arrangement occurs through a single pericyclic transition state.^{197,198} Up to four new chiral centres can be generated, making the DA reaction especially useful for synthetic chemistry.

7.1.1 Frontier Molecule Orbital Theory

Frontier Molecule Orbital (FMO) theory can be used to describe and understand DA reactions.^{199–201} The focus of this theory lies on the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). In the best case, the HOMO of the diene has perfect molecular orbital (MO) overlap with the LUMO of the dienophile, enabling the flow of electron to form two new C-C bonds (Figure 7.1 B).

The simplest [4+2] cycloaddition occurs between 1,3-butadiene **152** and ethylene **153** (Figure 7.1 A). Since the HOMO_{diene}–LUMO_{dienophile} orbitals are far apart in energy in this particular case, the reaction requires catalysis.^{199–201}

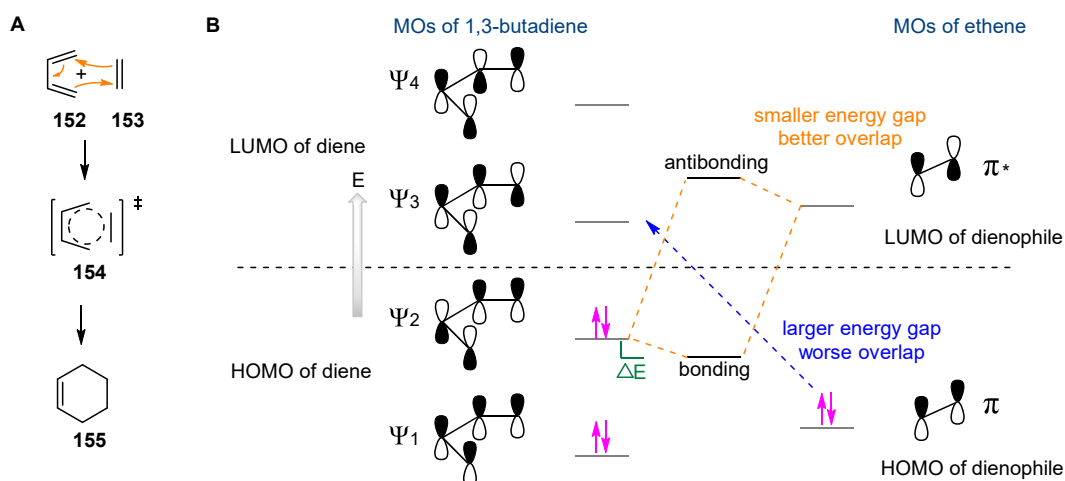
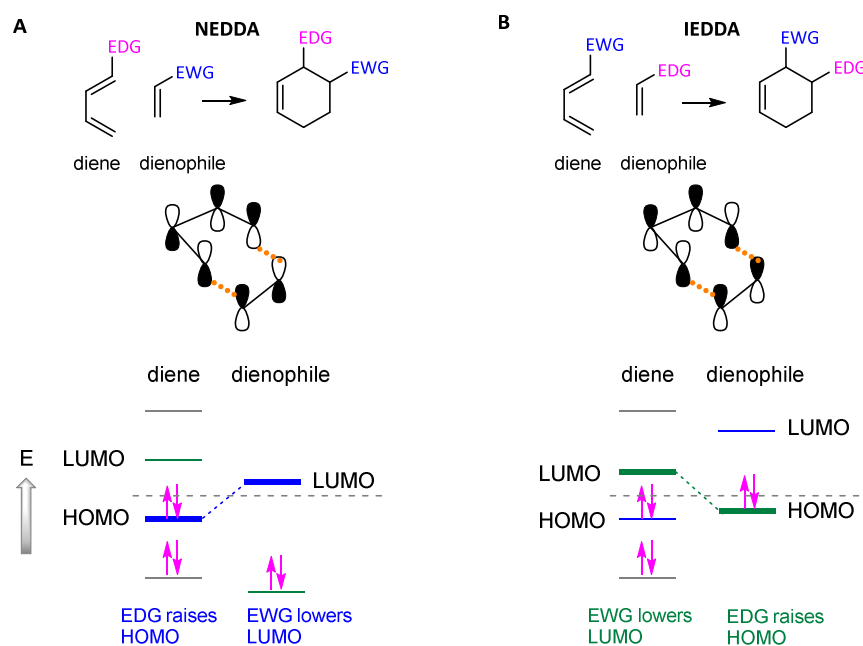


Figure 7.1: Molecular Orbitals (MO) in the Diels-Alder Reaction: **A**, DA reaction between 1, 3-butadiene **152** and ethylene **153**; **B**, Frontier molecular orbital (FMO) interaction for **152** and **153**. HOMO = highest occupied molecular orbital, LUMO = lowest unoccupied molecular orbital, E = energy.

To increase the reaction rate, either the energy level of the LUMO of the dienophile can be lowered by decreasing its electron density or the HOMO of the diene can be increased. Either or both modifications would strengthen the $\Delta E_{\text{HOMO-LUMO}}$ interaction.²⁰¹ Therefore, the dienophile is usually associated with an electron withdrawing group (EWG) and the diene with an electron donating group (EDG, Scheme 7.1 A). In addition to this normal electron demand Diels-Alder (NEDDA) pathway, the reaction can also occur *via* an inverse electron demand Diels-Alder (IEDDA) pathway in which the donating/ withdrawing groups are exchanged. (Scheme 7.1).²⁰¹⁻²⁰³



Scheme 7.1: Favourable feature of the diene and the dienophile. **A**, Normal electron demand Diels-Alder (NEDDA) reaction; **B**, inverse electron demand Diels-Alder (IEDDA) reaction. HOMO = highest occupied molecular orbital, LUMO = lowest occupied molecular orbital, EDG = electron-donating group, EWG = electron-withdrawing group, E = energy.

7.1.2 How Does a Diels-Alderase Operate?

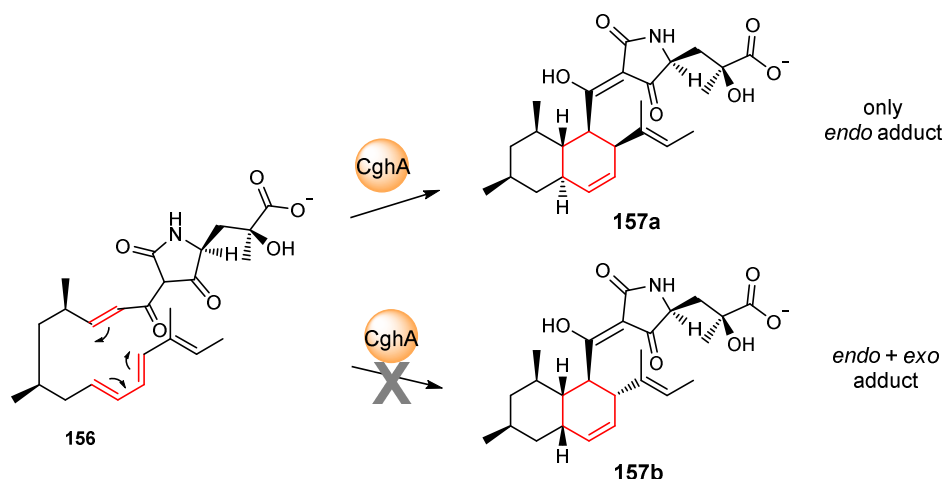
The existence of enzymes catalysing a DA reaction has long been doubted since most observed [4+2] cycloadditions also proceed at appreciable rates in the absence of any enzyme.^{204,205} In the last years numerous natural products were discovered such as the cytochalasans, lovastatin **2**, spinosyn A **171** and Sch21092 **157a/b** containing a cyclohexene ring with defined stereochemical configuration consistent with a [4+2] cycloaddition of a diene with a dienophile. These findings suggest the involvement of a DAase like enzyme.

Despite the fact that [4+2] cycloadditions have been proposed to be involved during the biosynthesis of more than 400 natural products, only a few DA enzymes have been biochemically characterized.^{206,207} This is due to the fact that it is difficult to predict DAases *in silico* due to poor sequence homology and the absence of specific catalytic motifs or cofactors. Another reason might be that most [4+2] cycloadditions were identified to also proceed at sufficient rates in the absence of enzyme, albeit the stereochemistry of the product might vary. Nevertheless, some DAases have been characterised in more detail giving the first hints on how a DAase might operate.

7.1.2.1 Conformational Catalysis

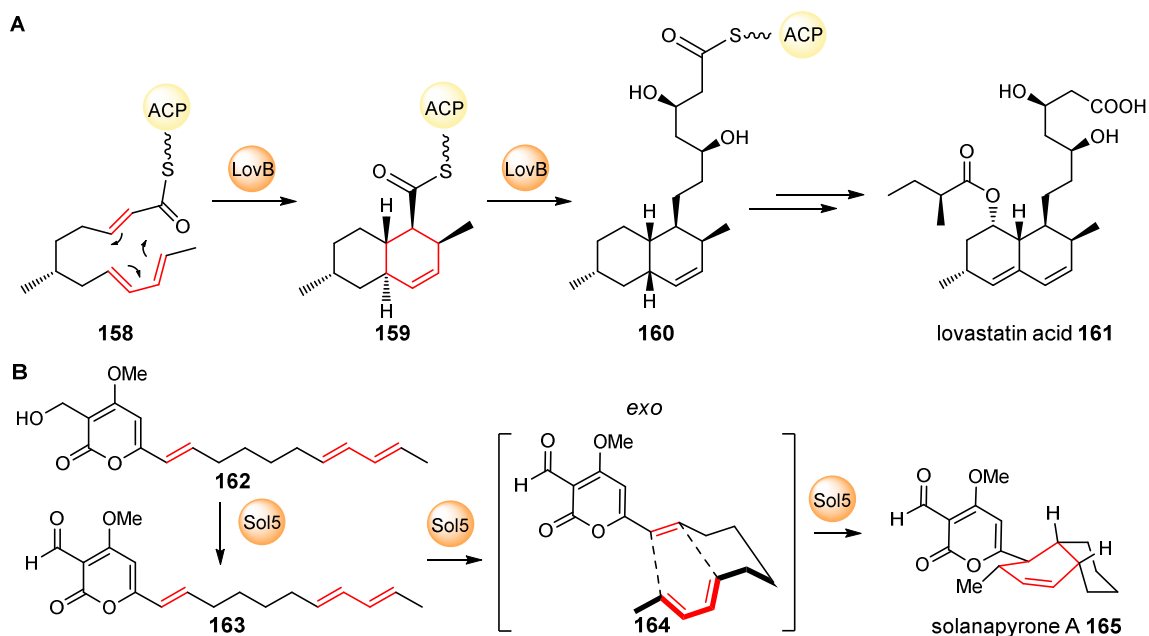
A DAase could operate in several possible ways. One possibility is conformational catalysis. The enzyme only serves as a (chiral) template encouraging the formation of conformers for productive reaction and hindering the formation of undesired conformers. This is supported by the observation that during the biosynthesis of the hybrid alkaloid-polyketide Sch210972 **157a/b**, the decalin-forming DAase CghA catalyses formation of the *endo* adduct **157a**, while in a Δ cghA deletion strain both the *endo* and *exo* forms **157a/b** are produced (Scheme 7.2).⁷⁷

However, little is known about the binding and stabilization of pericyclic transition states. Thus, the mode of action of a DAase is of great interest. The concerted simultaneous reorganization of the six-electron system of the conjugated diene and the dienophile processes without any charge development. Confirming this idea, changes in the polarity of the solvent are known to have little effect on the reaction rate of DA reactions, provided that the DA reaction only occurs between two purely hydrocarbon reactants.²⁰⁸ The rate enhancement of DA reactions in aqueous compared to organic solvents can be attributed to hydrophobic packing of the diene and the dienophile, bringing both substrates close together favouring the formation of the transition state. Thereby a DAase could simply act as a conformational template to enable the cyclization reaction.^{209,210}



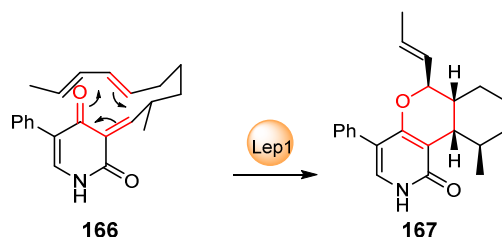
Scheme 7.2: The decalin-forming DAase CghA serves in Sch21092 formation as a chiral template to direct all the reaction flux over the endo transition state. The *exo* adduct is produced together with the *endo* adduct in a $\Delta cghA$ mutant strain.⁷⁷

Often [4+2] cyclases are multifunctional enzymes such as the Lovastatin **2** nonaketide synthase (LovB) and the solanapyrone synthase (Sol5). LovB is a PKS which catalyses the formation of the polyketide backbone **158** and subsequent arrangement of the intramolecular decalin ring by its additional DAase activity (Scheme 7.3 A).²¹¹ Sol5, encoding a flavin-dependent oxidase, catalyses the oxidation of prosolanapyrone III **162** to generate prosolanapyrone II **163** followed by an intramolecular [4+2] cycloaddition to produce solanapyrone A **165** (Scheme 7.3 B).



Scheme 7.3: Examples for catalysed Diels-Alder reactions by multifunctional DAases; A, LovB; B, Sol5

Another structurally unrelated DAase catalysing a [4+2] cycloaddition is LepI, which shows high homology to a SAM dependant methyltransferase. In the biosynthesis of leporin C **167** LepI catalyses the formation of the dihydropyran ring *via* a hetero DA reaction (Scheme 7.4).²¹²

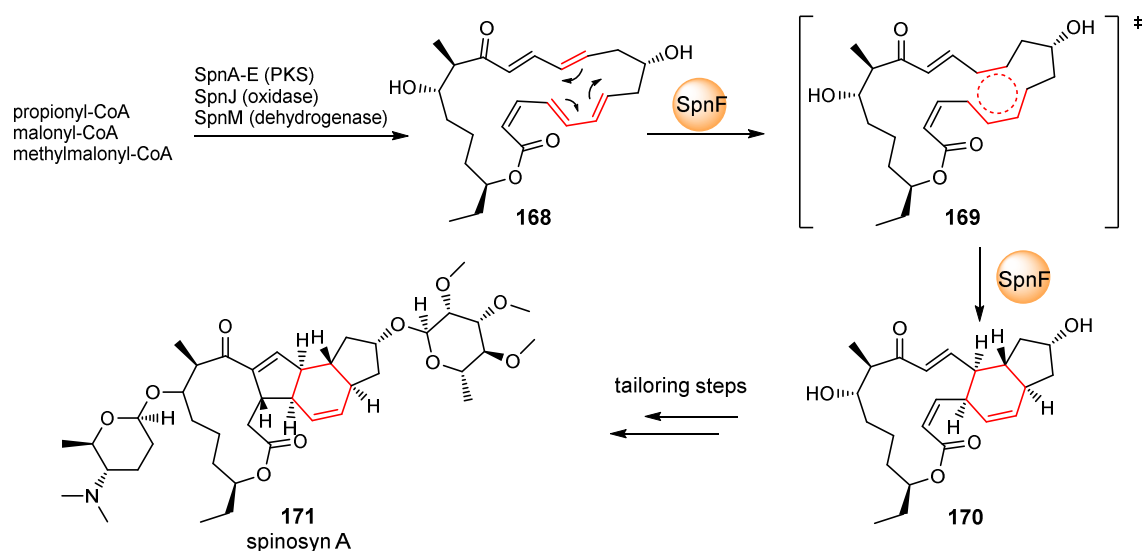


Scheme 7.4: Examples for hetero Diels-Alder reactions catalysed by Lep1.²¹²

In all these cases the DA-activity could be regarded as “accidental” reactivity of an enzyme with a different, true purpose. The DA-activity is probably due to conformational catalysis and due to the conformational restricted and chiral environment of the active site.

7.1.2.2 Chemical Catalysis

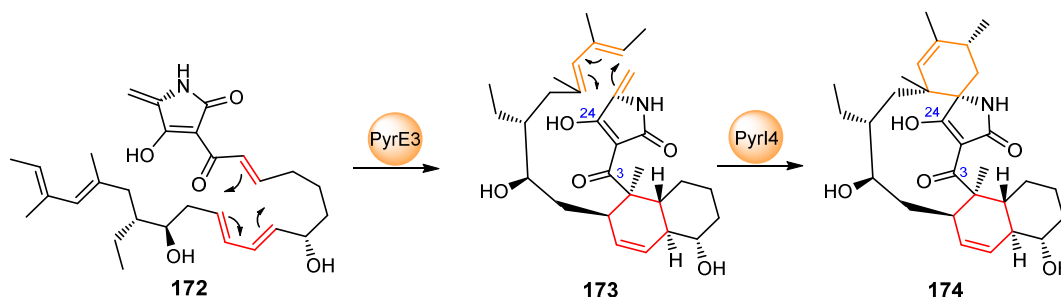
Another possible mechanism of DAases is by chemical catalysis. This could be achieved by lowering the free energy of activation for cyclization leading to a rate enhancement compared to the rate of a reference reaction. During the biosynthesis of spinosyn A **171**, for example, Liu and co-workers demonstrated that a putative DAase SpnF is able to accelerate the rate of spontaneous DA conversion of pentaene **168** to the tricyclic product **170** with an estimated 500-fold rate enhancement (Scheme 7.5).²¹³



Scheme 7.5: Biosynthesis of spinosyn A **171**. The [4+2] cycloaddition reaction is catalysed by SpnF.²¹³

During the biosynthesis of pyrroindomycin two bacterial stand-alone DAases were reported (PyrE3, PyrI4) to catalyse a [4+2] cycloaddition to form its core structure **174** (Scheme 7.6). PyrE3

belongs to the family of Rossmann fold proteins and is predicted to be FAD-dependent, while PyrI4 contains a β -barrel.^{214,215} Crystal structure analysis of PyrI4 alone, and in complex with its substrate, revealed that a lid-like interaction of the N-terminal sequence of PyrI4 with the substrate imposes conformational changes in its catalytic core, the β -barrel, conducting the *exo*-selective [4+2] cycloaddition. Enzymatic activity of PyrI4 was tested by mutagenesis-based bioassays, emphasizing that the key residue Q115 interacts with the hydroxyl at C-24 and carbonyl at C-3 of the ligand *via* two hydrogen bonds, causing an electron-withdrawing effect on the substrate. This interaction might lower the LUMO energy of the dienophile to facilitate [4+2] cycloaddition.²¹⁴ This could therefore be an example of chemical catalysis. [4+2] cyclases such as PyrE3 and PyrI4 which only have the function to catalyse a [4+2] cycloaddition could be considered as “true” DAases. The previously mentioned DAase SpnF from the spinosyn A **171** biosynthesis can also be considered as true DAase whose sole purpose is to catalyse a DA reaction (Scheme 7.5).



Scheme 7.6: Examples for stand-alone DAases PyrE and PyrI4.^{214,215}

7.1.2.3 [4 + 2] Cycloaddition during the Biosynthesis of Cytochalasans

Another putative DAase is encoded in the BGC of cytochalasans.⁷⁸ The intramolecular cycloaddition of cytochalasans is predicted to occur after the hybrid polyketide-amino acid chain is released from the PKS-NRPS (Section 1.6). Prior to the DA reaction the pyrrolinone has to be formed, which is then thought to serve as the dienophile. The pyrrolinone is possibly formed *via* a Knoevenagel condensation after the PKS-NRPS product is released as an aldehyde by the C-terminal reductase domain, as observed in the biosynthesis of fusarin C **8** (Scheme 1.8).^{50,78} An alternative route of pyrrolinone formation might be a Dieckman cyclization (DKC) catalysed by the PKS-NRPS release domain followed by reduction. DKC release has been observed during the biosynthesis of pretenellin A **33** (Scheme 1.6).⁴⁹ The pyrrolinone (usually) appears to undergo an *endo*-selective DA reaction with the polyketide terminal triene to form the isoindolone-fused macrocycle. An exception of this rule is chaetochalasin A **175** which probably undergoes an *exo*-transition state resulting in the opposite stereochemistry at C-5 and C-8.^{216,217}

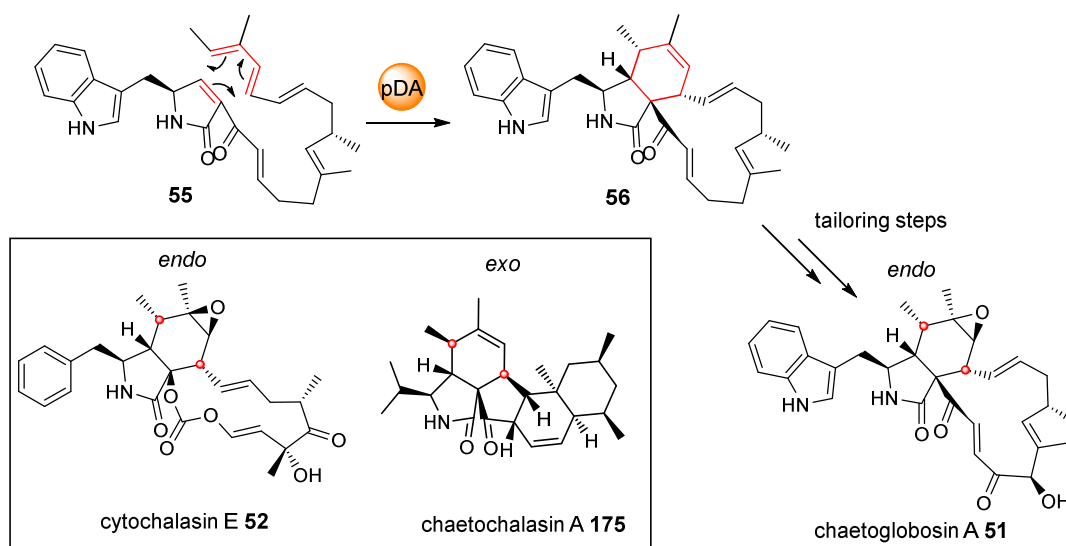
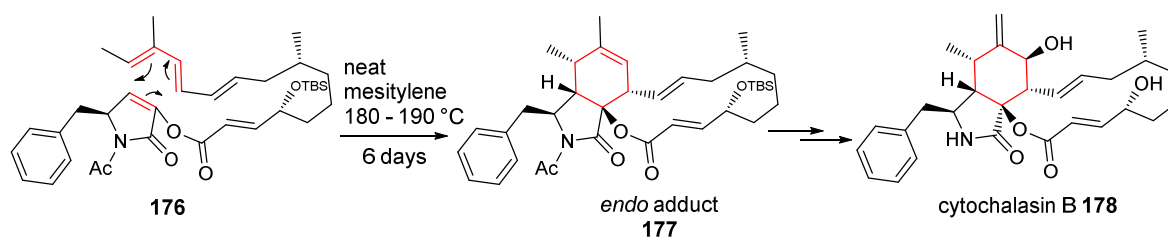


Figure 7.2: *Endo*- and *exo*-selective DA reaction of cytochalasans.

The assumption that the DA reaction during the biosynthesis of cytochalasans depends on a catalysing cyclase is supported by observations in the total synthesis of cytochalasans. Stork and Nakamura developed the first macrocyclization strategy in 1982 while synthesizing cytochalasins B **178**. To achieve the DA reaction the cytochalasin backbone **176** had to be heated in neat mesitylene (180–190 °C) for nearly a week to yield 35 % product **177** (4:1 mixture of *endo:exo*, Scheme 7.7), indicating that the DA reaction does not occur spontaneously at room temperature.²¹⁸



Scheme 7.7: DA by Stork and Nakamura in the total synthesis of cytochalasin B **178**.²¹⁸

DAases involved in the biosynthesis of cytochalasans probably belong to the stand-alone DAases with no additional function besides catalysing the [4+2] cyclisation. Whether this reaction requires conformational or chemical catalysis is unknown.

7.2 Aims of This Study

To gain more insights into the DA reaction of cytochalasans and to test whether this reaction is spontaneous or enzyme catalysed the function of the putative Diels-Alderase ORF3 from the *ACE1*

BGC in *P. oryzae* Guy11 and its homolog PyiF from the pyrichalasin H **50** BGC in *M. grisea* NI980 will be investigated by targeted gene deletion experiments and *in-vitro* studies. ORF3 and PyiF encode for a protein that exhibit no significant sequence similarity to characterised proteins and contain no conserved catalytic or cofactor-binding domains. The function of ORF3 and PyiF is therefore undetermined. Notably, all identified cytochalasan BGC contain a homolog of ORF3/PyiF, indicating that ORF3/PyiE and their homologs may play an essential post-PKS biosynthetic step - probably as a putative DAase.^{51,78} The **50** BGC was demonstrated to be a good homolog of the *ACE1* BGC (Chapter 2). ORF3 and PyiF share an amino acid identity of 60 % (75 % similarity). Therefore, the role of both proteins in the biosynthesis of **50** and the *ACE1* metabolite is likely to be similar. Gene deletion experiments are only possible in the **50** BGC in *M. grisea* NI980, since the production levels of **50** are very good (60 mg/L). The *ACE1* metabolite from the *ACE1* BGC in *P. oryzae* Guy11 cannot be observed under laboratory conditions. Therefore, targeted gene disruption of *pyiF* will be accomplished. Furthermore, the putative DAase-activity of ORF3 will be investigated *in-vitro*.

7.3 Results

7.3.1 Targeted Gene Deletion of *pyiF*

In order to investigate the function of the putative DAase PyiF in *M. grisea* NI980 and to elucidate its role in the biosynthetic pathway of pyrichalasin H **50**, its encoding gene *pyiF* was disrupted using the bipartite method (Section 1.9.1.1). The deletion cassette (pE-YA-*pyiF*-KO) was constructed by *in-vivo* homologous recombination in *S. cerevisiae* using a hygromycin B resistance cassette as selection marker (encoding a hygromycin B phosphotransferase gene, *hph*, Figure 7.3 A). DNA fragments for assembly were amplified by PCR. For amplification of *pyiF*, *M. grisea* NI980 gDNA was used as template. The selection marker was amplified from vector DNA (pTH-GS-*egfp*, section 5.4.1). The 5' and 3' bipartite substrates for *pyiF* were amplified by PCR from the constructed vector DNA (pE-YA-*pyiF*-KO, Figure 7.3 A, Section 9.4.3).

Twenty hygromycin resistant transformants were cultivated for 7 days in DPY media (25 °C, 110 rpm) before cells and growth media were individually extracted twice with ethyl acetate and analysed by LCMS (Section 9.5). In six transformants (VBI08-1, 3, 6, 9, 33, 34) production of **50** was completely abolished (Figure 7.4 A and B). Genetic analysis revealed that in transformants VBI08-6, -9 and -34 the complete hygromycin B resistance cassette was integrated correctly into *pyiF*, whereas in the case of VBI08-33 only the right-hand part was integrated (Figure 7.3 B). Transformants VBI08-1 and -3 still contained a functional copy of *pyiF*, however the organic

extract of these transformants contained almost no secondary metabolites, indicating a severe indirect and probably pleiotropic effect of the genomic integration of the resistance cassette (data not shown).

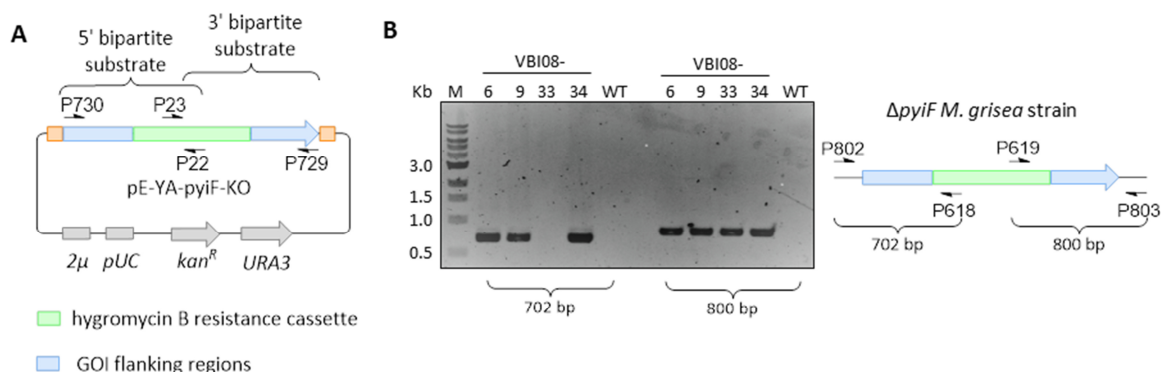
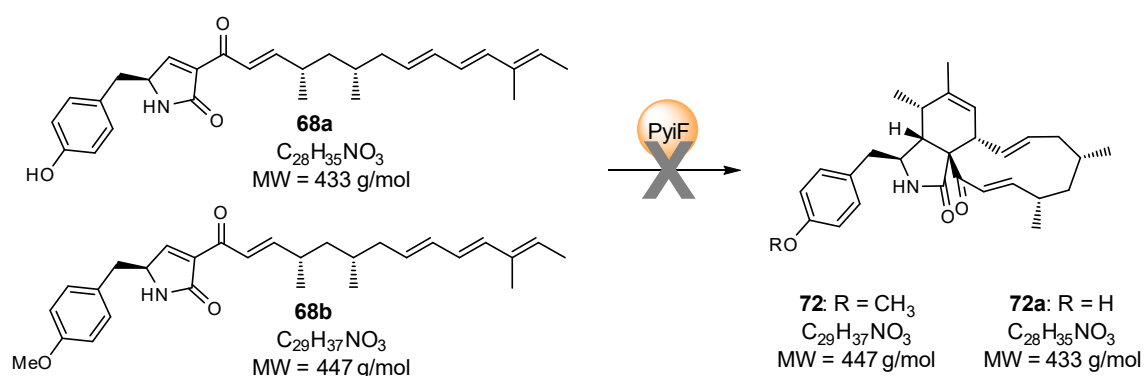


Figure 7.3: Genetic analysis of *pyiF* targeted knockout transformants: **A**, Constructed *pyiF* KO vector (VBI08) for amplifying the 5' and 3' bipartite substrates for *pyiF*; **B**, Genetic analysis of transformants obtained with bipartite marker strategy by PCR. GOI = gene of interest, M = marker

Successful disruption of *pyiF* should prevent the proposed intramolecular [4+2] DA cycloaddition to form the characteristic core; resulting in an accumulation of the pyrrolinone **68a** (derived from tyrosine **70**) and/or **68b** (derived from *O*-methyltyrosine **71**, Scheme 7.8).



Scheme 7.8: Hypothetical substrate(s) **68a/b** for the putative DAase PyiF.

LCMS analysis organic extracts of the *M. grisea* wild type and three Δ *pyiF* *M. grisea* deletion strains (VBI08 -9, 33, 34) are illustrated in Figure 7.4. The extracted ion chromatogram (XIC) using [M-H]⁻-corresponding to **50** ($m/z = 522$ g/mol) was used to observe production of **50**, which was absent in all transformants (Figure 7.4 B). However, the expected pyrrolinone(s) **68a/b** could not be observed ($m/z = 433$ g/mol and 447 g/mol). The molecular weight of the compounds which could be identified ranged between 469 – 519 g/mol (Table 7.3.1).

Remarkably, the organic extracts of the Δ pyiF *M. grisea* strains (VBI08 -9, 33, 34) contained several new compounds which, based on their UV absorption spectra at 275 nm, seemed to derive from (*O*-methyl) tyrosine **71**, (Figure 7.4 C).

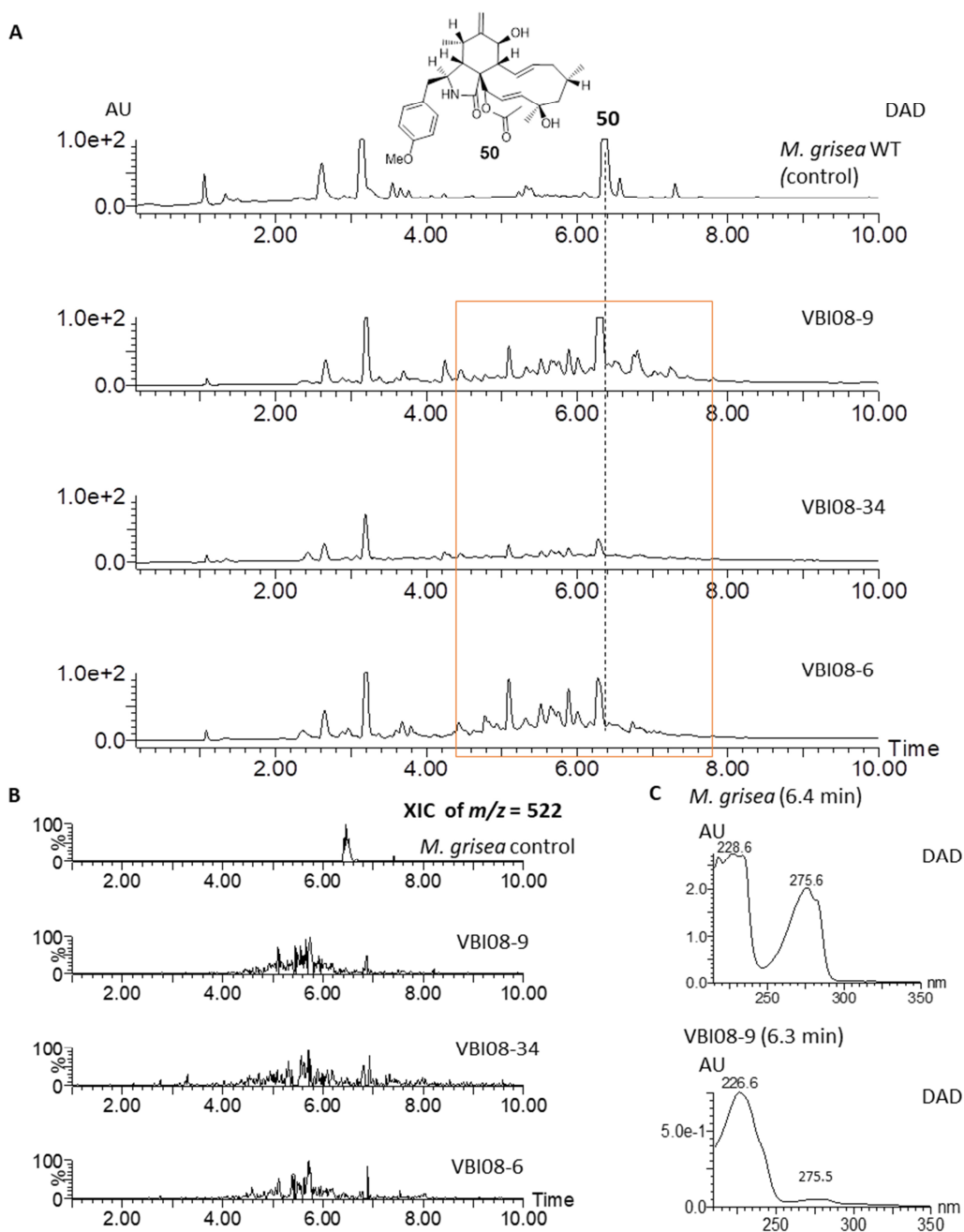


Figure 7.4: Chemical analysis of Δ pyiF *M. grisea* strains: **A**, LCMS chromatograms (DAD trace) of organic extracts from targeted deletion experiments in *M. grisea*; pyrichalasin H **50** is not produced in transformant VBI08-9, -34, and -6, instead several new related compounds are produced (orange rectangle); **B**, Extracted ion chromatogram (XIC) using $[M-H]^-$ corresponding to **50** ($m/z = 522$ g/mol) indicating that production of **50** is abolished in transformant VBI08-9, -34, and -6; **C**, UV-spectrum of **50** ($t_R = 6.4$ min) in comparison to a related compound produced in transformant VBI08-9 ($t_R = 6.3$ min).

Nevertheless, the characteristic spiky UV spectra corresponding to the triene moiety (observed for alcohol **67**; Section 3.3.1) was not detected, strongly suggesting that the triene is not present in the observed compounds (Figure 7.5).

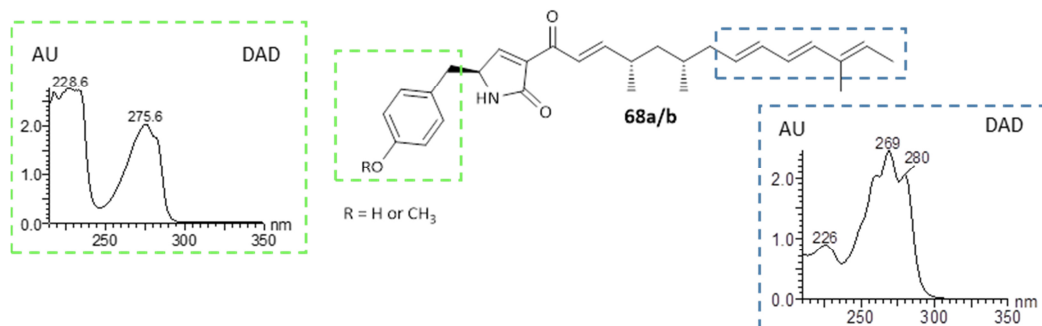


Figure 7.5: Characteristic UV absorption spectra for (*O*-methyl) tyrosine **71** (275 nm) and the triene moiety (spiky between 230 - 300 nm) of the linear hybrid PKS-NRPS backbone chain **68a/b**.

To further analyse the produced compounds, a 1 litre culture of VBI08-9, grown for 7 days at 25 °C in DPY media, was extracted. In comparison to the wild type chromatogram (evaporative light scattering detection (ELSD)) several new peaks were observed between 4.9 - 7.6 min with a main peak at $t_R = 6.3$ min (Figure 7.6).

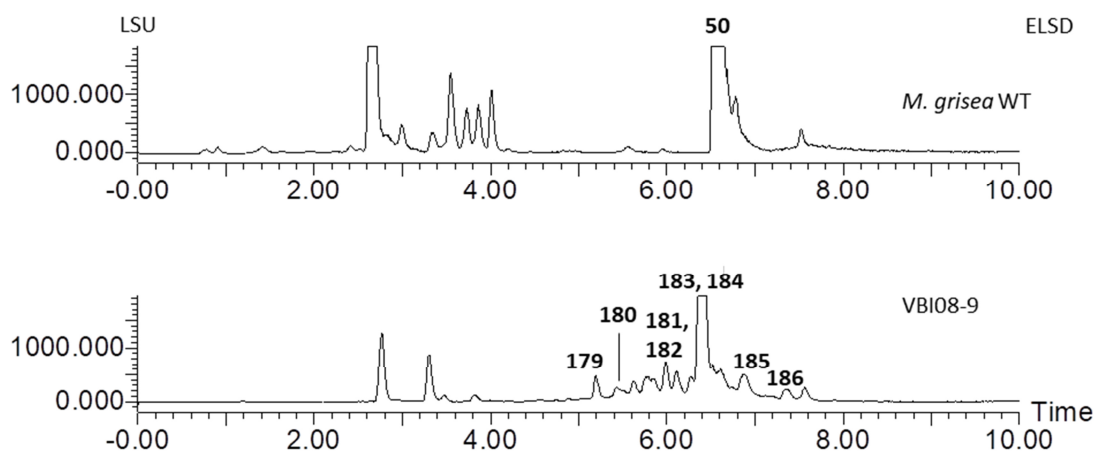


Figure 7.6: Evaporative light scattering detection (ELSD) of organic extracts of *M. grisea* wild type and VBI08-9. Eight compounds were purified by preparative LCMS and were analysed by NMR: **179–186**.

Eight compounds (**179**: 0.4 mg, **180**: 0.3 mg; **181** and **182** (same elution time): 0.5 mg; **183** and **184** (same elution time): 1.5 mg; **185**: 0.5 mg; **186**: 1 mg) were purified by preparative LCMS (Figure 7.7). All compounds displayed a UV absorption spectrum at 275 nm, suggesting that they derived from tyrosine **70** or *O*-methyltyrosine **71**. The main peak contains a mixture of compound **183** and **184**.

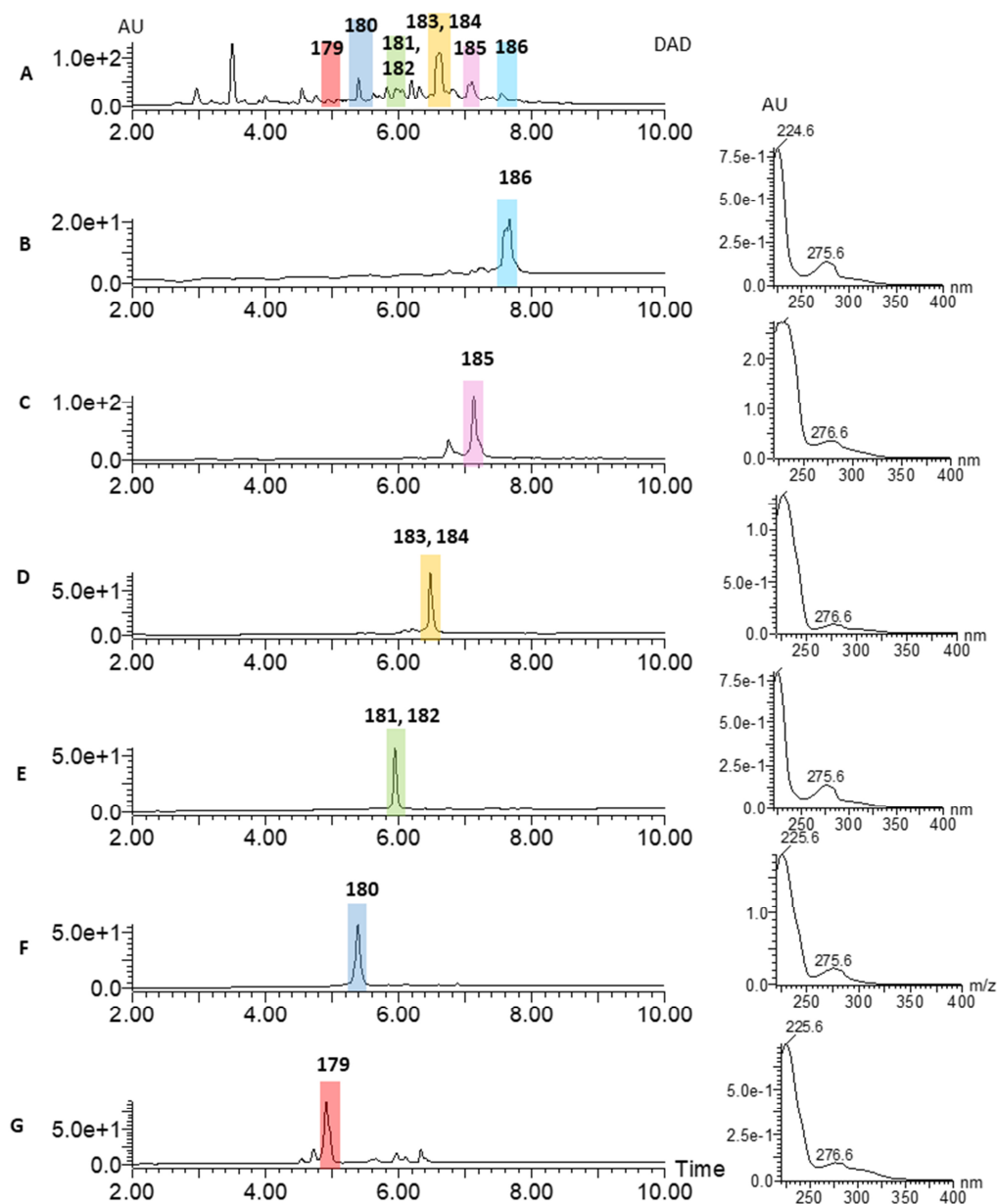


Figure 7.7: Results of preparative LCMS (DAD chromatogram): **A**, organic extract of VB108-9; **B-G**, purified and re-injected compounds **179–186** and their corresponding UV absorption spectrum.

The observed mass of each compound ($m/z = 469$ g/mol to 519 g/mol) further suggest that **179–186** might be pathway (shunt) intermediates of pyrichalasin H **50** (Figure 7.8).

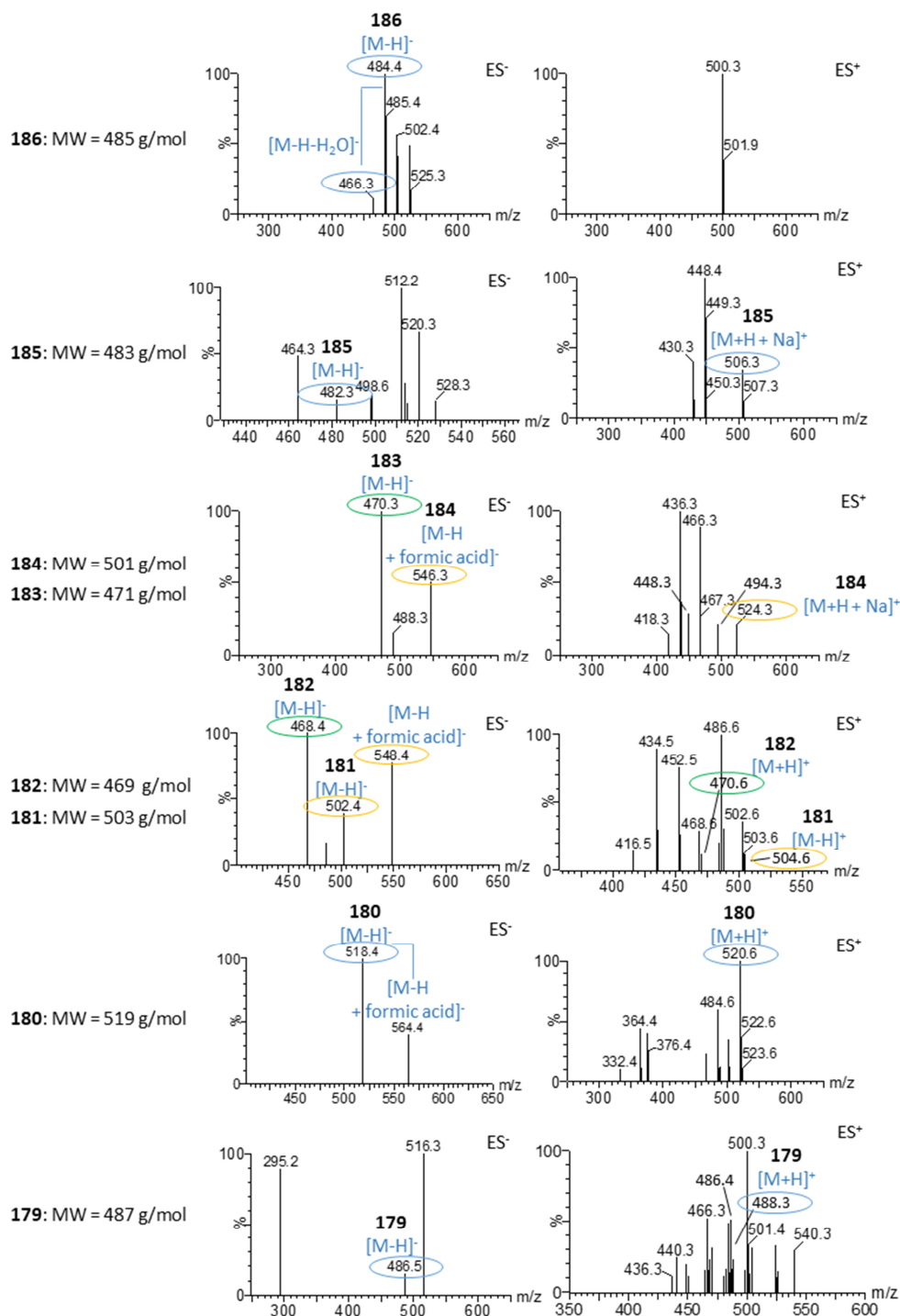


Figure 7.8: LCMS (ES⁺ and ES⁻) analysis of new compounds obtained from extracts of transformant VBI08-9 (Δ pyiF *M. grisea* deletion strain).

7.3.2 Structure Elucidation of Compound 179 - 186 by NMR

Structure elucidation of the obtained compounds by preparative LCMS was performed by nuclear magnetic resonance (NMR) and the molecular formula for each compound was determined by high resolution mass spectrometry (HRMS, Table 7.3.1, see supplementary for NMR data,

Section 11.2). ^1H - ^1H COSY data revealed characteristic signals for (*O*-methyl) tyrosine **71** related compounds in the aromatic region (Figure 7.9 A). Based on the HMBC experiment, *O*-methyltyrosine **71** derived compounds can be distinguished from tyrosine **70** derived compounds, by the presence or absence of a strong methyl (singlet at 3.8 ppm) to ring carbon correlation (C-1 carbon of the benzoyl ring, 159 ppm, Figure 7.9 B).

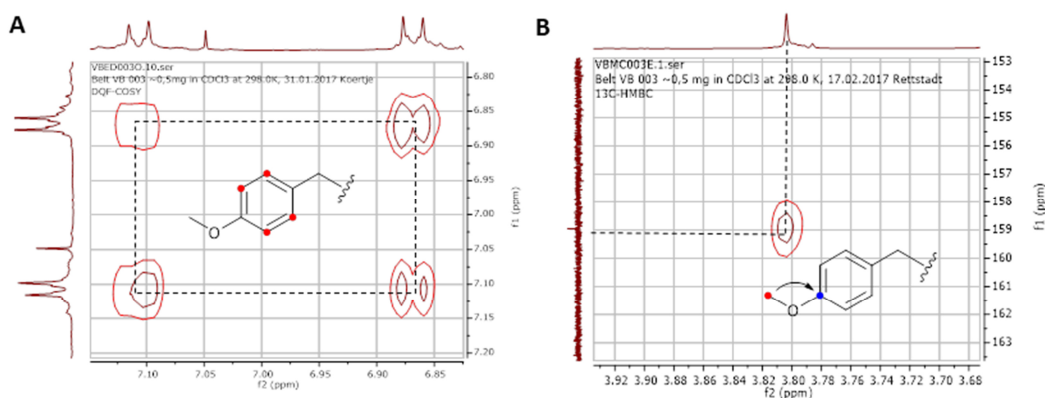


Figure 7.9: Characteristic NMR signals of *O*-methyltyrosine **71**; **A**, ^1H - ^1H COSY signals for *O*-methyltyrosine **71** (at 6.86, 6.86; 6.86, 7.14; 7.14, 6.86 and 7.14, 7.14 ppm); **B**, HMBC signal between the protons of the methyl group (singlet at 3.8 ppm) and the C-1 carbon of the benzoyl ring (159 ppm).

Compound **181/ 182** and **183/ 184** shared the same retention time and were therefore purified together. **181** and **184** are derived from *O*-methyltyrosine **71**, **180** and **183** from tyrosine **70**. A mixture of *O*-methyltyrosine **71** and tyrosine **70** derived compounds can be identified based on the proton and ^1H - ^1H COSY signals (Figure 7.10). Aromatic protons of *O*-methyltyrosine **71** display higher ppm signals. Base on the intensity of the proton signals for compound **183** and **184** (main peak, see Figure 7.6), compound **184** is the main produced compound (Figure 7.10, green rectangle: *O*-methyltyrosine **71** derived compound **184**, blue rectangle: tyrosine **70** derived compound **183**).

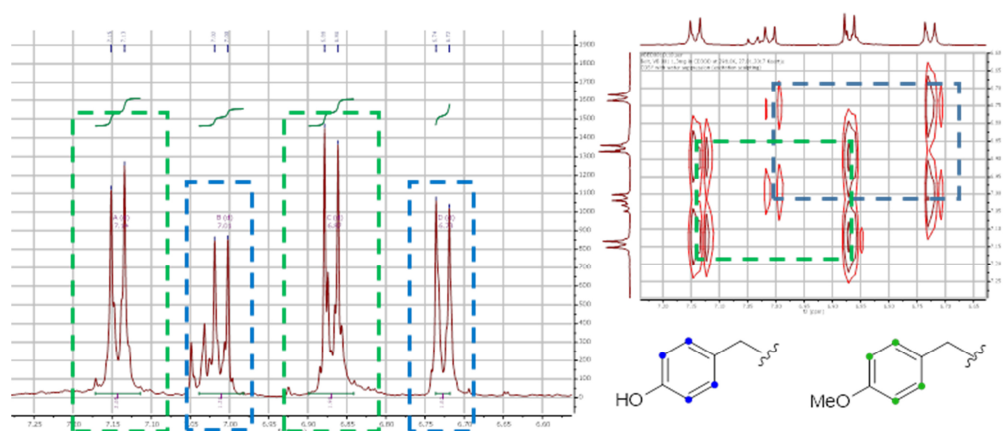
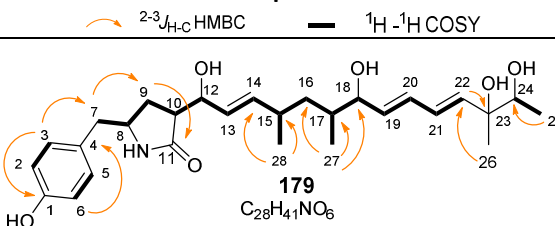
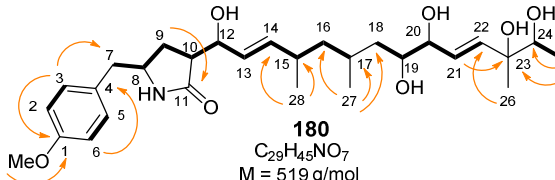
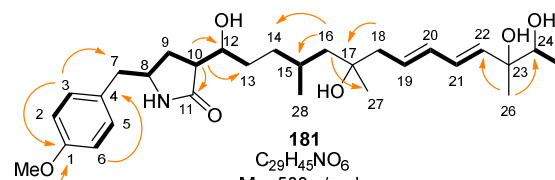
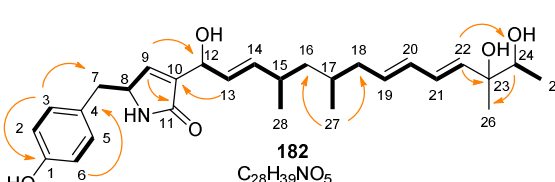
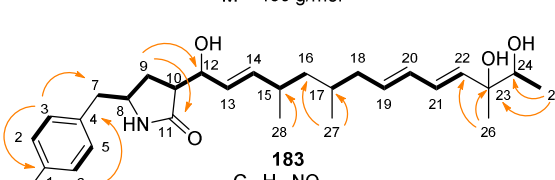
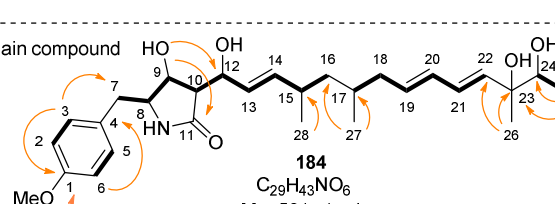
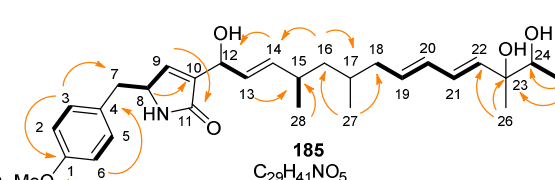
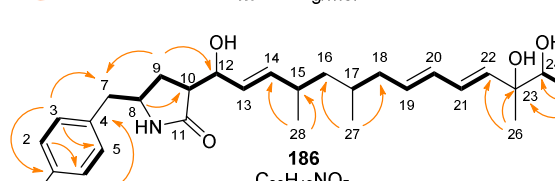


Figure 7.10: Proton and ^1H - ^1H COSY signals of compound **183** and **184**. Green rectangle: *O*-methyltyrosine **71** derived compound **184**, blue rectangle: tyrosine **70** derived compound **183**.

Table 7.3.1: Structure elucidation of purified compound I-VIII by NMR and HRMS. Solvent: MeOD.

Compound	HRMS data
 <p>179 C₂₈H₄₁NO₆ M = 487 g/mol</p>	C ₂₈ H ₄₁ NO ₆ ([M-H] ⁻ calculated 486.2854, found 486.2856)
 <p>180 C₂₉H₄₅NO₇ M = 519 g/mol</p>	C ₂₉ H ₄₅ NO ₇ ([M]+Na ⁺ calculated 542.3089, found 542.3094)
 <p>181 C₂₉H₄₅NO₆ M = 503 g/mol</p>	C ₂₉ H ₄₅ NO ₆ ([M-H] ⁻ calculated 502.3167, found 502.3169)
 <p>182 C₂₈H₃₉NO₅ M = 469 g/mol</p>	C ₂₈ H ₃₉ NO ₅ ([M-H] ⁻ calculated 468.2750, found 468.2750)
 <p>183 C₂₈H₄₁NO₅ M = 471 g/mol</p>	C ₂₈ H ₄₁ NO ₅ ([M-H] ⁻ calculated 470.2905, found 470.2906)
 <p>184 C₂₉H₄₃NO₆ M = 501 g/mol</p>	C ₃₀ H ₄₄ NO ₈ (C ₂₉ H ₄₃ NO ₆ + formic acid [M]H ⁺ calculated 546.3067, found 546.3073)
 <p>185 C₂₉H₄₁NO₅ M = 483 g/mol</p>	C ₂₉ H ₄₁ NO ₅ ([M]+Na ⁺ calculated 506.2881, found 506.2880)
 <p>186 C₂₉H₄₃NO₅ M = 485 g/mol</p>	C ₂₉ H ₄₃ NO ₅ ([M-H] ⁻ calculated 484.3066, found 484.3063)

All isolated compounds displayed an underlying linear structure of the polyketide moiety (Table 7.3.1). Therefore, the DA reaction to form the macrocycle did not occur, supporting the proposal that P*yiF* catalyses the [4+2] DA cycloaddition. The C-23/ C-24 olefin was dihydroxylated in all compounds, probably *via* epoxide formation and hydrolysis, explaining the loss of the characteristic UV spectrum corresponding to the triene. *P. oryzae* is known to produce compounds containing a diol as pyriculol **187** and pyriculariol **188**, albeit in the first case the mid-chain polyene was epoxidised by a monooxygenase (Figure 7.11).²¹⁹ The same monooxygenase responsible for the formation of the diol in **188** might have led to the observed diols in **179** - **186**.

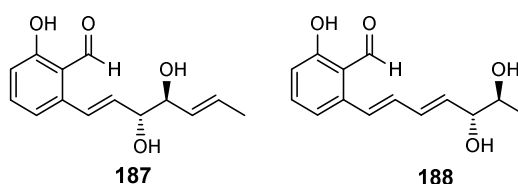


Figure 7.11: Secondary metabolites produced by *P. oryzae*.

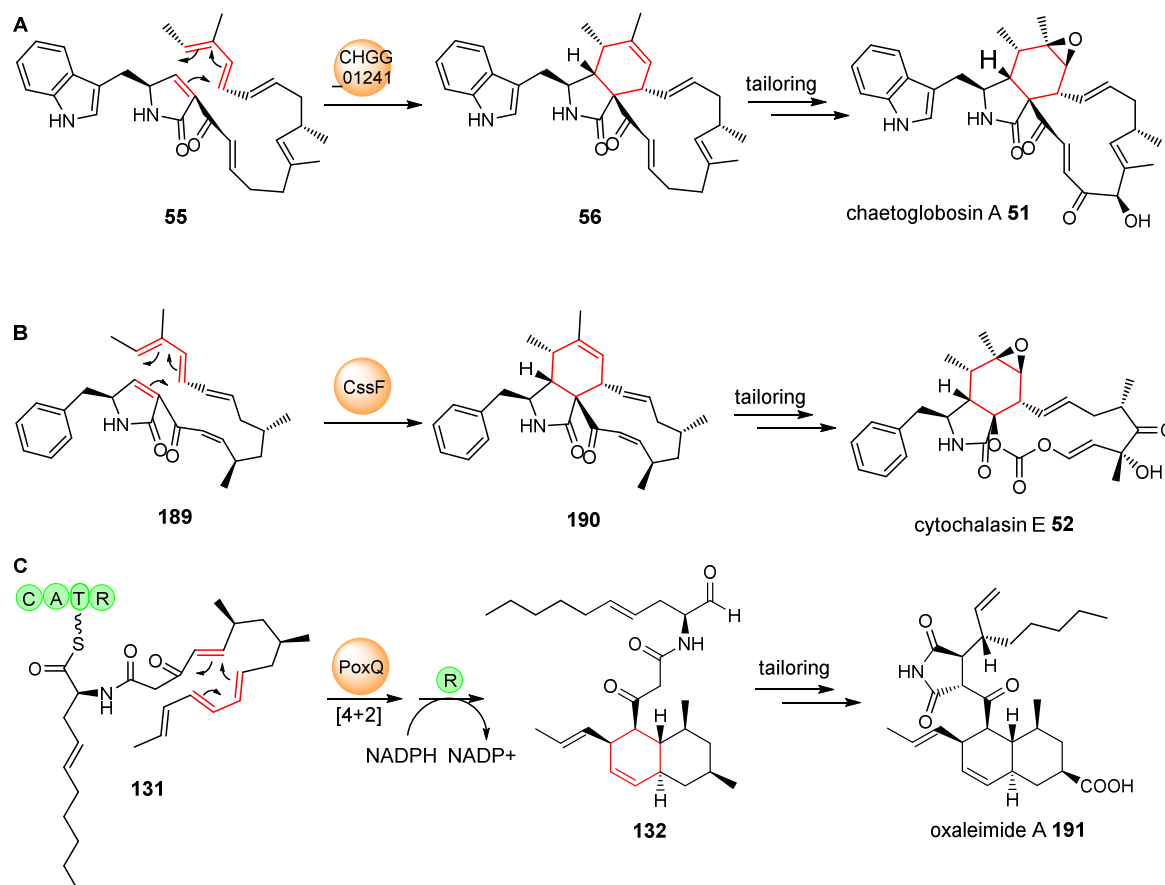
Compound **180** also features C-19/ C-20 olefin dihydroxylation. Compound **179** contains an additional hydroxyl group at C-18 and compound **181** is hydroxylated at C-17 and compound **184** is hydroxylated at C-9. Another deviation from the expected structure(s) **68a/b** is that all compounds carry a hydroxyl group at C-12 instead of a carbonyl group. Furthermore, the C-9/ C-10 olefin (part of the pyrrolinone ring) has been reduced in all compounds, excluding compound **185**. In five compounds (**180**, **181**, **184**, **185**, **186**) the hydroxyl group of the tyrosine is methylated. The observed compounds appeared to be shunt products of **68a/b** formed *via* oxidative and reductive steps. These data demonstrate that P*yiF* is needed to facilitate the [4+2] DA cycloaddition, however it still remains unclear whether **68a** and **68b** are the native substrates of P*yiF*, respectively.

7.3.3 Complementation Studies with Δ *pyiF* *M. grisea* Strains

In order to test if the disruption of *pyiF* is really linked to the observed changes in biosynthesis of **50**, a functional copy of *pyiF* was introduced downstream of the inducible *amyB* promoter (P_{amyB}) and transformed into two Δ *pyiF* *M. grisea* strains (VBI08-9 and VBI08-34). Furthermore, heterologous complementation studies were carried out to investigate if the closely related homologous protein from the *ACE1* BGC ORF3 (55 % identity, 69 % similarity) is able to restore production of **50**.

Three other homologous proteins were tested: CHGG_01241 (44 % identity, 61 % similarity) from the chaetoglobosin BGC in *Chaetomium globosum*, C*ssF* (61 % identity, 71 % similarity) from the

cytochalasin BGC in *Aspergillus clavatus* and PoxQ (28 % identity, 43 % similarity) from the PoxQ BGC in *Penicillium oxalicum*, which catalyses a decalin forming Diels-Alder reaction (Scheme 7.9).^{73,91,181}

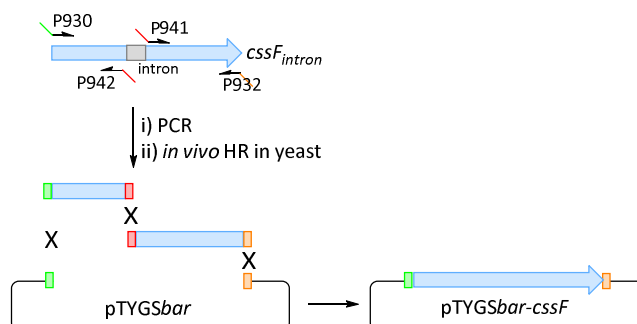


Scheme 7.9: Proposed [4+2] cycloadditions catalysed by DAases in the biosynthesis of **A**, chaetoglobosin A **51**; **B**, cytochalasin E **52**; **C**, oxaleimide A **191**.

7.3.3.1 Construction of Complementation Vectors

For the complementation studies the respective gene of interest was cloned into the fungal expression vector pTYGS $\bar{b}ar$ downstream of the strong inducible *amyB* promoter. Amplification of *pyiF* (P845 + P846) and *ORF3* (P806 + P807) was done from *M. grisea* and *P. oryzae* gDNA, respectively. Both genes were cloned into the entry vector pE-YA before they were individually transferred into pTYGS $\bar{b}ar$ by *in-vitro* LR recombination (Section 3.1). *CHGG_01241* (P933 + P934), *cssF* (P930 + P932) and *poxQ* (P935 + P936) were amplified from synthetic vector DNA (Invitrogen). 5' Extensions are generated by the primers, facilitating overlaps and homologous recombination of the flanks and the linearized (by the restriction enzyme *NotI*) pTYGS $\bar{b}ar$ expression vector. The purified DNA fragment was assembled into the linearized vector using homologous recombination in *S. cerevisiae*.

Bioinformatic analysis using FGENESH Soft Berry²²⁰ predicted that *cssF* contains one intron. To this end, *cssF* was amplified in two parts (exon 1 and exon 2, using P930 + P942 and P941 + P932) with homologous flanks for each other and the pTYGSbar expression vector, thereby removing the intron (Scheme 7.10).



Scheme 7.10: Removal of the intron by primer extension and *in vivo* homologous recombination (HR).

In total six complementation vectors were generated (Table 7.3.2, Section 9.4.3).

Table 7.3.2: Constructed complementation vectors.

Gene	Organism and BGC	Name
<i>pyiF</i>	<i>M. grisea</i> NI980, pyrichalasin H	pTYGSbar- <i>pyiF</i>
<i>ORF3</i>	<i>P. oryzae</i> Guy11, ACE1	pTYGSbar- <i>ORF3</i>
<i>CHGG_01241</i>	<i>Chaetomium globosum</i> , chaetoglobosin	pTYGSbar- <i>CHGG_01241</i>
<i>cssF</i>	<i>Aspergillus clavatus</i> , cytochalasin	pTYGSbar- <i>cssF</i>
<i>cssF_{intron}</i>	<i>Aspergillus clavatus</i> , cytochalasin	pTYGSbar- <i>cssF_{intron}</i>
<i>poxQ</i>	<i>Penicillium oxalicum</i> , <i>pox</i>	pTYGSbar- <i>poxQ</i>

7.3.3.2 Homologous Complementation with *pyiF* from the Pyrichalasin H BGC

Two Δ *pyiF* *M. grisea* strains (VBI08-9 and -34) were transformed with the expression plasmid pTYGSbar-*pyiF*. Twelve transformants (VBI15-9-1 to VBI15-9-6 and VBI15-34-1 to VBI15-34-6) were generated, extracted and analysed regarding their secondary metabolite production by LCMS (Section 9.5). Genomic DNA analysis revealed that a copy of *pyiF* was integrated into the genome of the transformants in all cases except VBI15-9-6 (Figure 7.12).

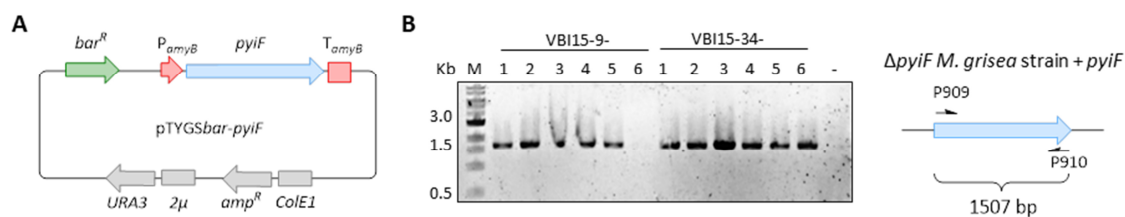


Figure 7.12: Genetic analysis of $\Delta pyiF + pyiF$ transformants: **A**, Complementation vector VBI15-I; **B**, Genomic DNA analysis using P909 and P910 (expected PCR fragment: 1507 bp), $\Delta pyiF M. grisea$ gDNA was used as negative control (-). M = marker.

As expected, in all strains carrying a functional copy of *pyiF*, the production of **50** was restored. However, the obtained yields of **50** were significantly smaller compared to the wild type as seen in the LCMS chromatograms (DAD trace) and extracted ion chromatogram (XIC) using $[M-H]^-$ corresponding to **50** ($m/z = 522$ g/mol, **Figure 7.13**). In addition, some of the shunt intermediates (**179 – 186**) observed in the $\Delta pyiF M. grisea$ deletion strain were also detected in most of the transformants (mainly compound **184**). Probably in the transformants PyiF is not able to react with the whole amount of substrate in time; therefore some of it gets shunted to the observed bi-products (**179 – 186**). One reason for that might be that the concentration of PyiF is decreased due to its non-native promoter and/or its foreign locus on the genome. In addition, the gene *pyiF* is no longer clustered with the other biosynthetic genes of the pyrichlasain H **50** BGC; therefore ribosomal production of PyiF might take place far away from the rest of the biosynthetic machinery. Until PyiF is present, its substrate gets most likely (degraded and) converted into the observed shunt products (**179 – 186**).

This experiment strongly suggests that PyiF is needed to facilitate the [4+2] DA cycloaddition. Nevertheless, since production of **50** could not be restored to 100 %, either P_{amyB} does not work as efficiently as the native promoter of *pyiF* and/or the localization of *pyiF* within the genome of *M. grisea* affects its level of transcription. The presence and heterologous expression of two selection marker genes (*hph* and *bar*) in the strains might also have an effect on the production levels.

Based on the extracted ion chromatogram (XIC) using $[M-H]^-$ corresponding to **50** ($m/z = 522$ g/mol) and to **184** + formic acid ($m/z = 546$ g/mol) chromatographic integrals were obtained. Using chromatographic peak integration the area under the chromatographic peak is measured and can be used to calculate unknown relative concentrations. During this study the area under the peak was used to approximately compare amounts of **50** in the *M. grisea* WT strain and the transformants and the amounts of **184** in the $\Delta pyiF M. grisea$ deletion strain and the

transformants. The integral of each peak was determined using the Waters MassLynx™ Software. Compared to the amounts of **50** present in the *M. grisea* WT strain (100 %) the yields of **50** in the transformant strains were significantly decreased (around 10 %, **Figure 7.13**, green numbers). Although the experiments were not designed for a quantitative analysis, the integrals provide enough evidence to conclude that production of **50** was restored compared to the Δ *pyiE* *M. grisea* control strain (0 %).

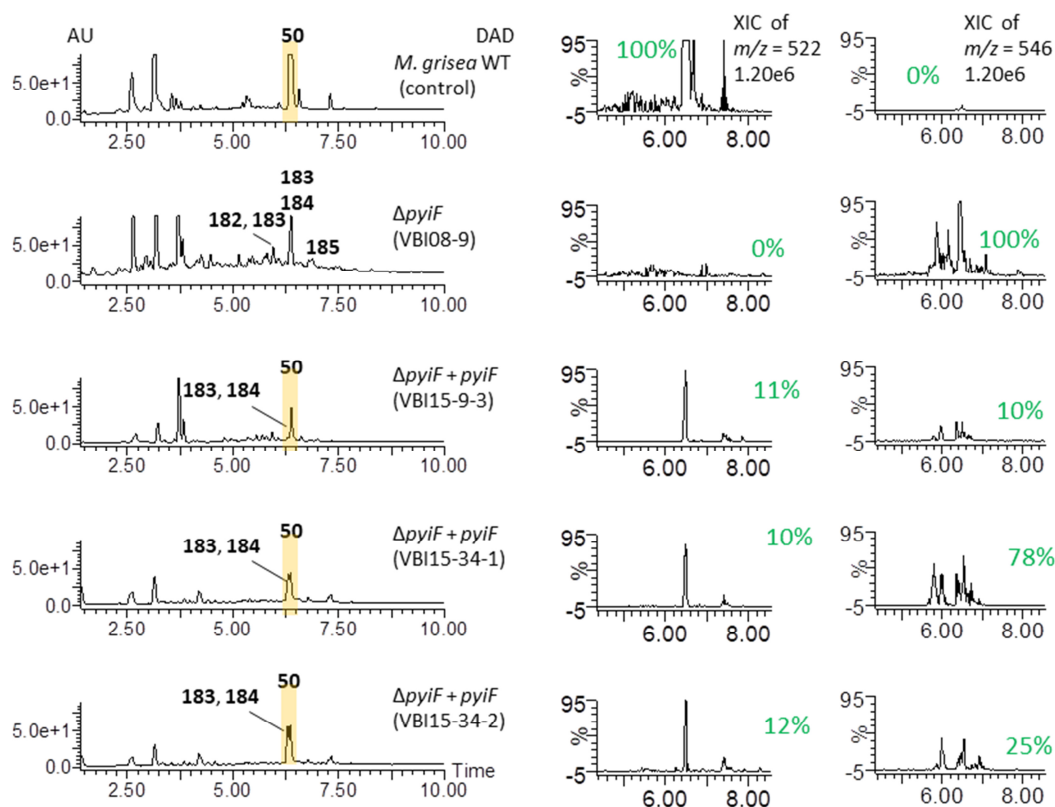


Figure 7.13: Homologous complementation of Δ *pyiF* *M. grisea* strains. LCMS chromatograms (DAD trace) and extracted ion chromatogram (XIC) using $[M-H]^-$ corresponding to **50** ($m/z = 522$ g/mol) and to **184** ($m/z = 546$ g/mol). Green numbers: relative amount of **50** (compared to the WT, set 100 %) and **184** (compared to Δ *pyiF* *M. grisea* strains, set 100 %)

7.3.4 Heterologous Complementation with Homologous Genes of *pyiF*

In order to elucidate the function of ORF3, especially regarding its substrate specificity, the Δ *pyiF* *M. grisea* strains (VBI08-9 and -34) were complemented with *ORF3* (59 % identity), a homolog of *pyiF* from the ACE1 BGC in *P. oryzae* Guy11. Both biosynthetic pathways seem to be closely related, suggesting that ORF3 might be able to replace PyiF as a DAase.

Twelve transformants (VBI11-9-1 to VBI11-9-6 and VBI11-34-1 to VBI11-34-6) were generated and analysed regarding their secondary metabolite production by LCMS. Genomic DNA analysis

revealed that the *ORF3* was successfully integrated into the genome of the transformants in all cases except VBI11-9-4, VBI11-34-4 and VBI11-34-5 (Figure 7.14).

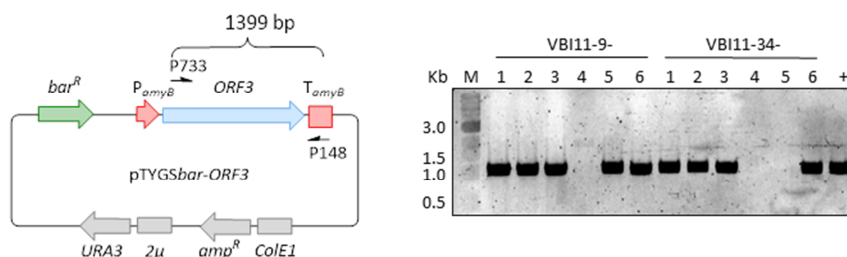


Figure 7.14: Genetic analysis of $\Delta pyiF + ORF3$ transformants. Genomic DNA analysis using P733 and P148 (expected PCR fragment: 1399 bp), vector DNA was used as positive control (+).

Strains carrying the integrated gene of *ORF3* restored the production of **50** as seen in the LCMS chromatograms (DAD trace) and extracted ion chromatograms (XIC) using $[M-H]^-$ corresponding to **50** ($m/z = 522$ g/mol, approximately from 0 % to 10 %, **Figure 7.15**). However, the yields were quite low compared to the wild type, but similar compared to the homologous complementation studies with *pyiF* (Section 7.3.3.2). No new compounds were observed.

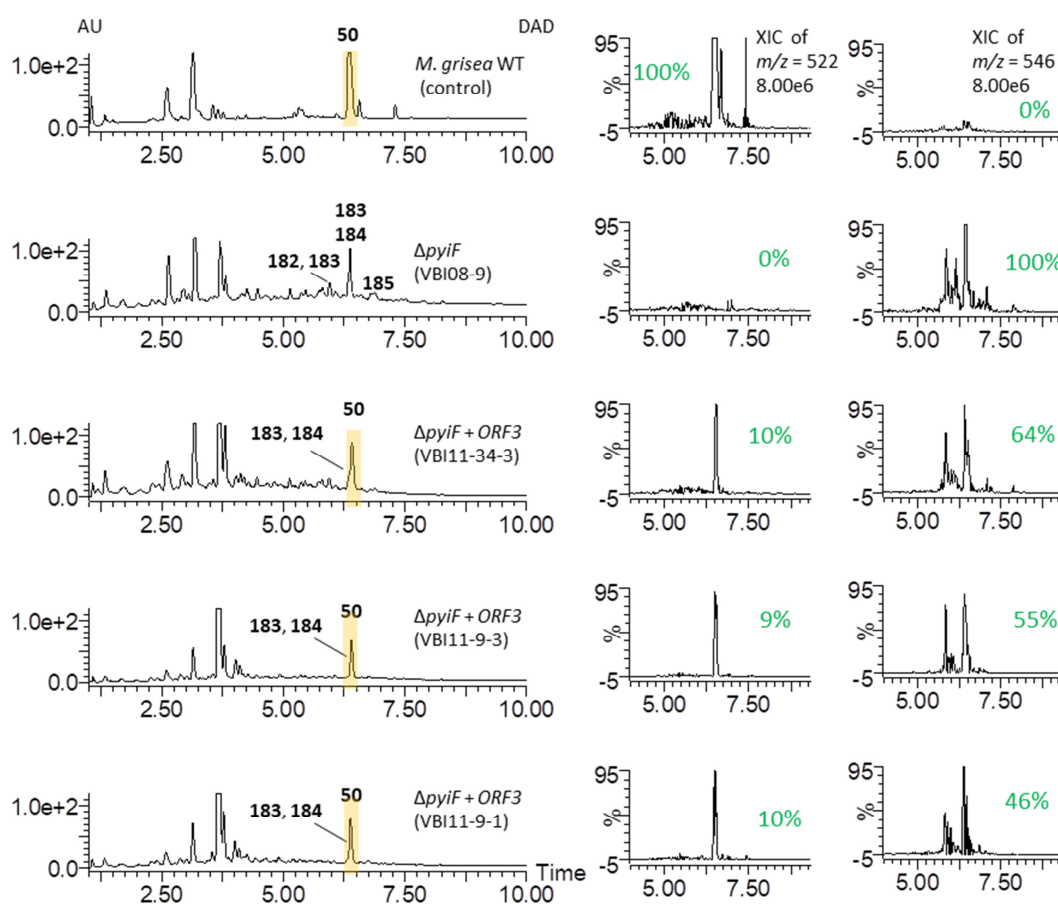


Figure 7.15: Heterologous complementation of $\Delta pyiF$ *M. grisea* strains with *ORF3*. LCMS chromatograms (DAD trace) and extracted ion chromatogram (XIC) using $[M-H]^-$ corresponding to **50** ($m/z = 522$ g/mol) to **184** ($m/z = 546$ g/mol). Green numbers: relative amount of **50** (compared to the WT, set 100 %) and **184** (compared to $\Delta pyiF$ *M. grisea* strains, set 100 %)

In order to gain further insight into their substrate specificity, analogous complementation experiments of the $\Delta pyiF$ *M. grisea* strain were carried out with various genes encoding for potentially related DAases *CHGG_01241*, *cssF*, *cssF_{intron}* and *poxQ*, respectively. For each transformation 7-10 BASTA resistant transformants were obtained and were analysed for their secondary metabolite production by LCMS. In addition, genomic DNA analysis confirmed the successful genome integration of the respective construct (Figure 7.16).

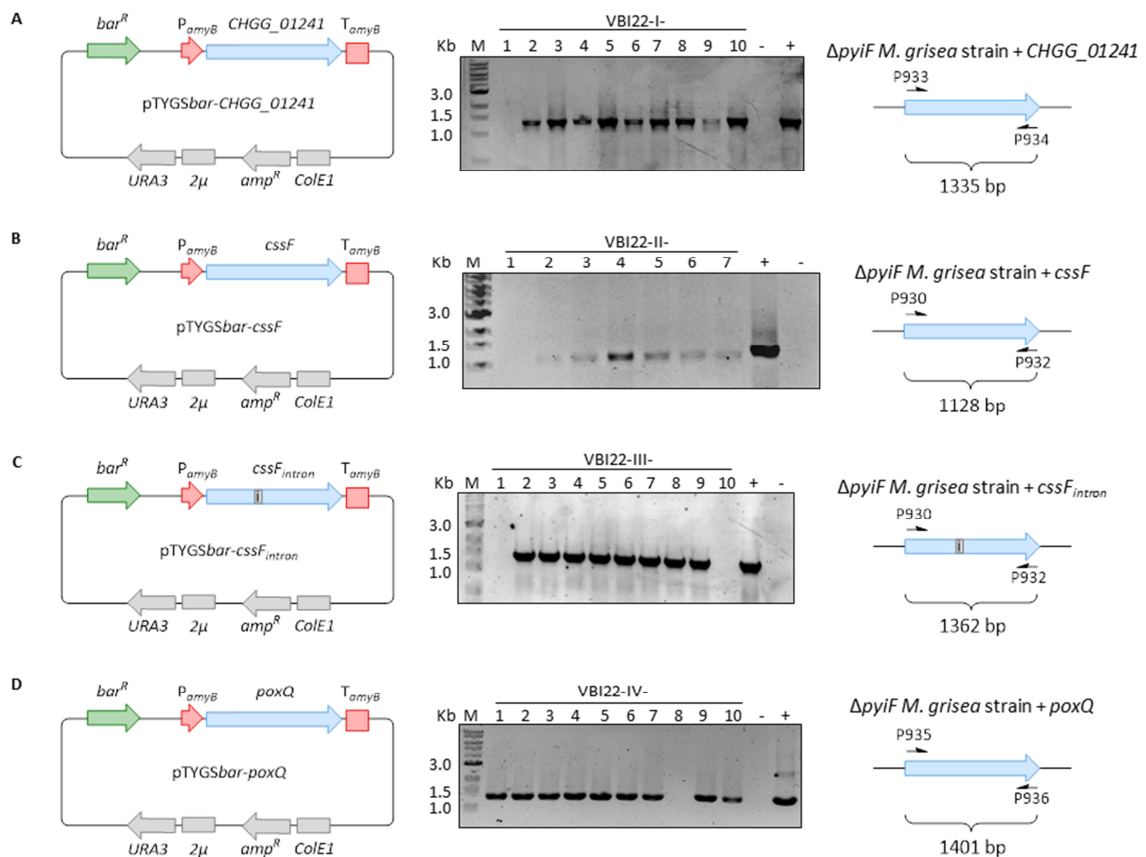


Figure 7.16: Genetic analysis of various $\Delta pyiF$ *M. grisea* transformants: **A**, $\Delta pyiF$ + *CHGG_01241* transformants; **B**, $\Delta pyiF$ + *cssF_{intron}* transformants; **C**, $\Delta pyiF$ + *cssF* transformants; **D**, $\Delta pyiF$ + *poxQ* transformants.

Figure 7.17 illustrates examples of the organic extracts of transformants VBI08-9 (negative control), VBI22-I-2 (*CHGG_01241*), VBI22-III-3 (*cssF* with intron), VBI22-II-3 (*cssF* without intron) and VBI22-IV-1 (*poxQ*). All transformants demonstrated the same secondary metabolite profile as the $\Delta pyiF$ *M. grisea* control strain. Accordingly, the production of **50** could not be restored.

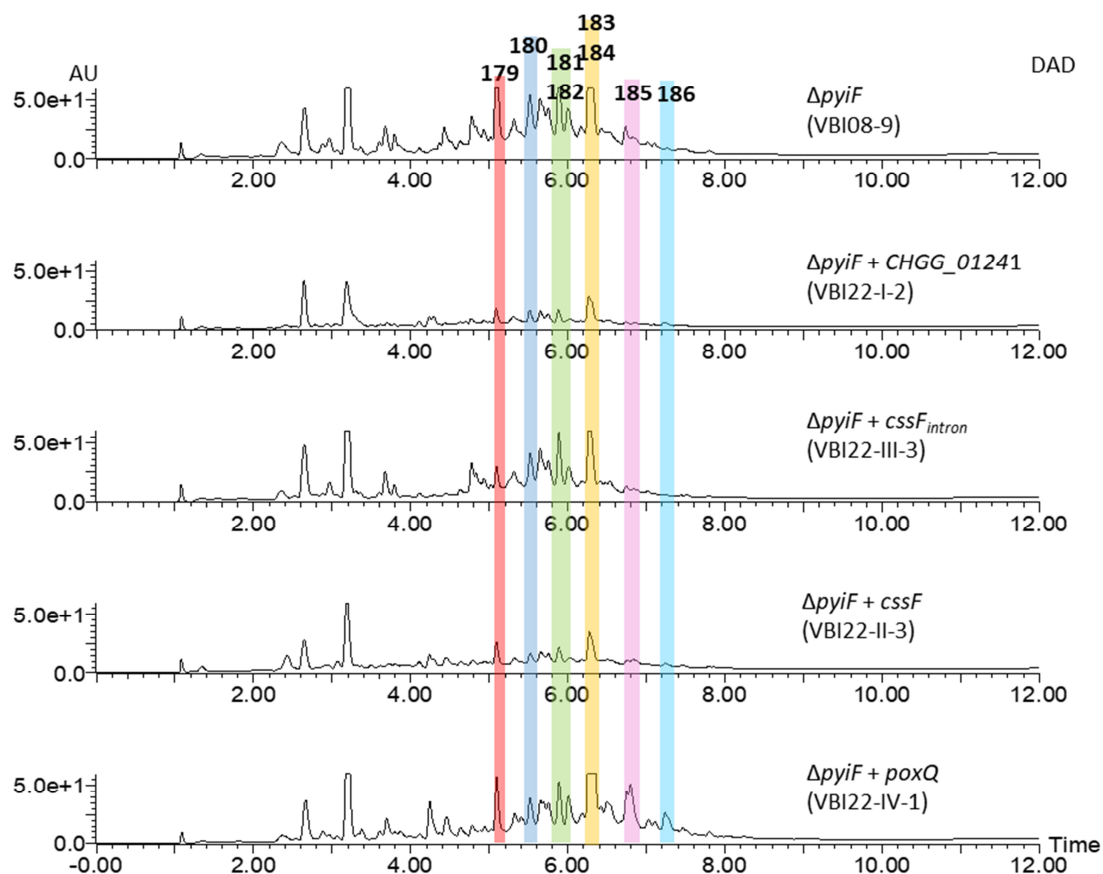
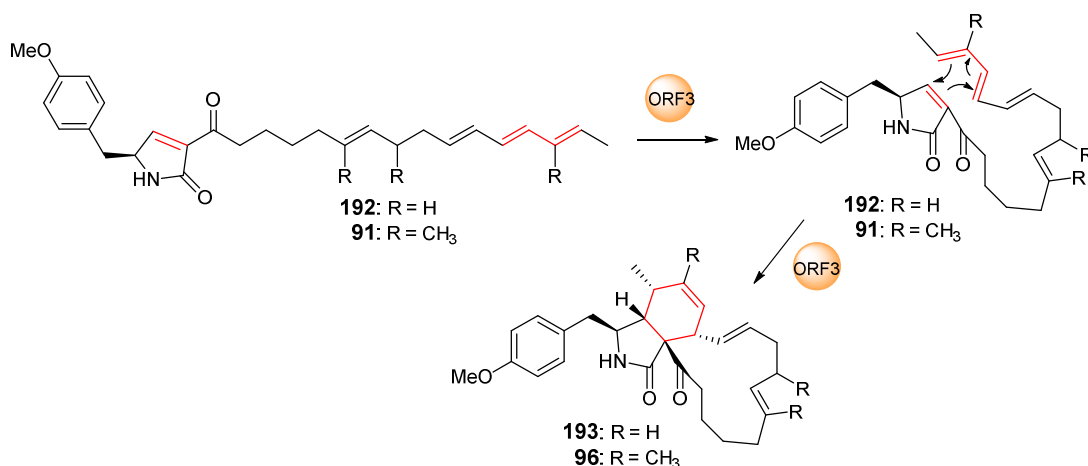


Figure 7.17: Heterologous complementation of $\Delta pyiF$ *M. grisea* strains with *CHGG_01241*, *cssF* and *poxQ*. LCMS chromatograms (DAD trace) of organic extracts from $\Delta pyiF$ *M. grisea* strains (VBI09, negative control) and generated transformants (VBI22-I-2, VBI22-II-3, VBI22-III-3 and VBI22-IV-1).

7.3.5 *In-vitro* Studies with ORF3

In parallel to the *in-vivo* studies of *pyiF* and *ORF3* in *M. grisea* by target directed gene disruption and complementation studies, the protein ORF3 from the *ACE1* BGC of *P. oryzae* Guy11 was heterologously produced in *E. coli* in order to test its function as a putative DAase *via in-vitro* enzyme assays. At the beginning of this work the pyrichalasin H **50** BGC was not available; therefore, the focus was on ORF3.

Synthetic substrates for testing (compound **91** and **192**, Scheme 7.11) should be provided by Haili Zhang (Leibniz Universität Hannover, Institute of Organic Chemistry, AK Cox), who started his synthesis in parallel to this work.



Scheme 7.11: Planned *in-vitro* enzyme assays with ORF3.

7.3.5.1 Cloning, Heterologous Expression and Purification of ORF3

The native *P. oryzae* ORF3 gene was cloned into the expression vector pET28a(+) via *Nde*I and *Xho*I restriction sites. The successful construction of the plasmid pET28a(+)-ORF3 was confirmed by sequencing. The plasmid was transformed into chemically competent *E. coli* BL21 (DE3) cells. Single colonies were used to inoculate 2TY medium containing kanamycin. Cultures were grown to an OD₆₀₀ of 0.4-0.6 at 37 °C before adding IPTG to a final concentration of 0.1 - 1 mM according to experimental procedures (Section 9.4.11). To optimize the protein production, various growth temperatures and *E. coli* expression strains (BL21 (DE3), BL21 RP codon plus, Rossetta2) were tested.

Unfortunately, even in BL21-CodonPlus RP and Rossetta2, both of which contain tRNAs rarely used in *E. coli* to enhance heterologous expression, ORF3 could not be detected in the cellular extract by SDS-PAGE (Figure 7.18 A).

These results suggested that there could be a problem with the codon usage, since the transformed sequence of ORF3 derived from the fungus *P. oryzae* Guy11. The fungal origin could cause problems in the bacterial production of ORF3 in *E. coli*. Therefore, the whole sequence was codon optimized for *E. coli* using the webserver OPTIMIZER. The fully codon optimized ORF3 sequence (ORF3_{opt.}) was synthesized and cloned into the plasmid pET100-D/TOPO (Invitrogen). The heterologous production of the codon optimized His₆-ORF3 in BL21 (DE3) was successful. However, all ORF3 protein was located in the insoluble fraction. His₆-ORF3 has an expected molecular weight of 50.7 kDa (Figure 7.18 B).

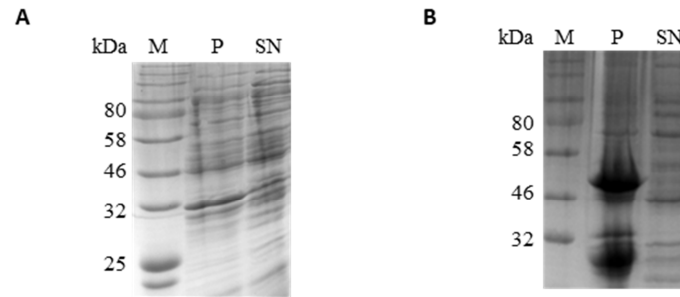


Figure 7.18: Protein production of ORF3 in *E. coli* BL21 from a fungal gene sequence and an *E. coli* codon optimised gene sequence. Bands were visualized by Coomassie staining. M: marker (Colour Prestained Protein Standard, 11–245 kDa); The cells grew at 25 °C for 18 h after IPTG induction. P: protein extracts were prepared from cell pellets; SN: protein extracts were prepared from culture supernatant; **A**, fungal *ORF3* gene sequence; **B**, *ORF3* gene sequence codon optimised for *E. coli*.

To improve the solubility of ORF3 the plasmid pET100-D/TOPO-*ORF3* was transformed into various *E. coli* strains including Arctic Express (expression at very low temperatures), BL21 pLysS (no basal expression, good for toxic proteins) and BL21 cells containing the chaperone plasmid pG-KJE8. The co-expression of chaperones is known to increase the yields of soluble protein.²²¹ Furthermore, protein production was tested at 16 °C and 25 °C and using an auto-induction medium. Collectively, no soluble protein could be observed in usable amounts under tested conditions (data not shown).

7.3.5.2 Protein Purification Using SUMO Fusion Systems

The SUMO fusion technology for enhanced protein production is described in Section 6.3.3.2. *ORF3* was cloned into the vector pCA528 according to experimental procedures. The successful construction of the plasmid pCA528-*ORF3* was confirmed by sequencing. Transformation into competent *E. coli* BL21 (DE3) cells and protein production was obtained according to experimental procedures.

The cell lysate was loaded onto a Ni-NTA column, washed with 20 mM imidazole wash buffer, eluted with elution buffer containing 250 mM imidazole and analysed by SDS-PAGE and Coomassie staining (Figure 7.19 A). His₆-SUMO-ORF3 has an expected molecular weight of 58.6 kDa. All elution fractions were combined and loaded onto a desalting column to remove imidazole. SUMO protease 1 (Ulp1) was used to cleave off the His₆-SUMO tag. Cleaved ORF3 was recovered in the flow-through fractions after a second Ni-NTA column whereas the 6x His-smt3 and 6x His Ulp1 remained on the column (Figure 7.19 B + C).

The purified protein was enzymatically digested by trypsin before the peptide fragments were further analysed by ESI Q-TOF (Dr. Jennifer Senkler work group Prof. Braun, Institute of Plant Genetics, Leibniz University of Hannover). The production of ORF3 could be verified with coverage of 79.2 % and a score of 6007.73 (Figure 7.19 D). Observation of N- and C-terminal peptides confirmed the protein was full-length.

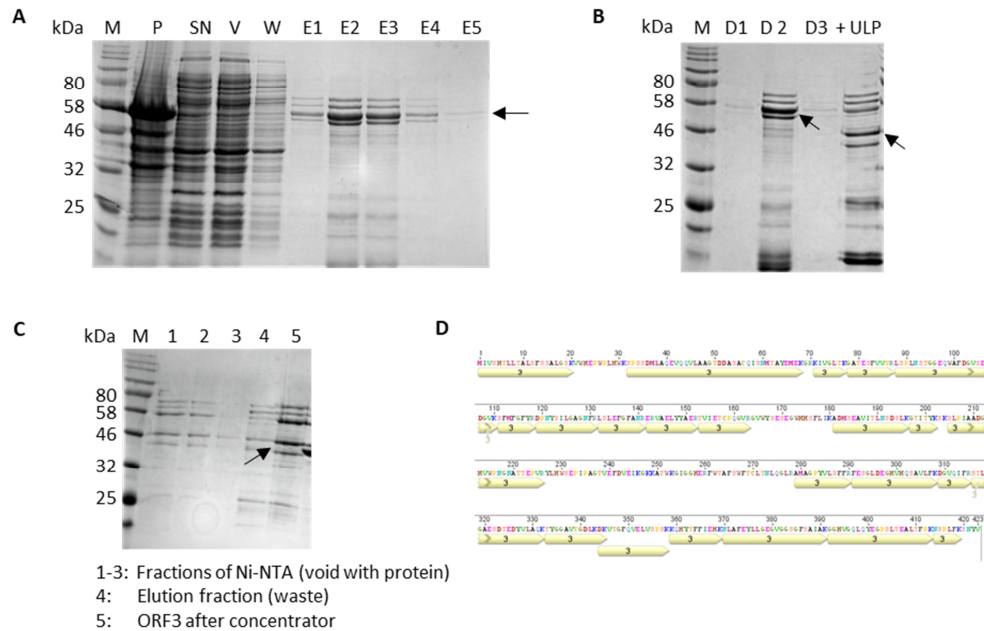


Figure 7.19: Purification of the His₆-SUMO-ORF3 fusion protein. Bands were visualized by Coomassie staining. M: marker (Colour Prestained Protein Standard, 11–245 kDa). The molecular weight of the target protein is 46.5 kDa and of the SUMO fusion 58.6 kDa and is indicated by the black arrows. The cells grew at 25 °C for 18 h after IPTG induction **A**, Purification of His₆-SUMO-ORF3 by Ni-NTA. Protein sample was loaded onto Ni-NTA column, washed with 20 mM imidazole wash buffer, eluted with elution buffer containing 250 mM imidazole and analysed by SDS-PAGE. P: protein extracts were prepared from cell pellets; SN: protein extracts were prepared from culture supernatant; V: void fraction, W: wash fraction; E1-E5: elution fractions. **B**, Analysis of SUMO-ORF3 fusion protein following cleavage with ULP SUMO protease; D1- D3: fractions of desalting column; +ULP (lane 5): His₆-SUMO-ORF3 cleavage products after incubation with ULP protease. **C**, second Ni-NTA to remove 6 x His-SUMO tag and 6x His-ULP protease from protein sample. ORF3 is now in the void fractions (lane 1-3), lane 4: elution fraction (250 mM imidazole); lane 5: purified recombinant ORF3. **D**, ESI Q-TOF analysis of ORF3. Yellow arrows indicate identified peptide sequences.

7.3.5.3 *In-vitro* Studies with ORF3

In order to examine ORF3 regarding the catalysis of the DA reaction and to prove its function as a putative DAase, an *in-vitro* enzyme assay was designed. Unfortunately, the synthetic substrates **91** and **192** were not available yet. Therefore, enzyme assays with ORF3 had to be postponed.

7.4 Discussion

The biosynthesis of **50** and other known cytochalasans requires a [4+2] cycloaddition in order to form the isoindole macrocycle. This core structure of cytochalasans is thought to be constructed by a DA reaction, mainly due to the stereochemistry of the six-membered ring and the fact that the [4+2] cycloaddition does not occur spontaneously at room temperature (Scheme 7.7). Targeted gene disruption of *pyiF* in *M. grisea* completely abolished the production of **50**, concurrently resulting in the accumulation of several new related compounds. Structure elucidation of eight compounds (**179** - **186**) revealed that all compounds shared an underlying linear structure probably derived from the hybrid PKS-NRPS backbone **68a/b**, the putative substrate of P*yiF*. P*yiF* probably catalyses the formation of **68a/b** into **72/72a**, therefore, disruption of the function of *pyiF* would result in an accumulation of **68a/b**. Accordingly, compound **179** - **186** are presumably shunt products of **68a/b** (Scheme 7.12).

The polyunsaturated compound **68a/b** (triene) is most likely epoxidised by a monooxygenase in *M. grisea* and spontaneous hydrolysis of the epoxides resulted in the observed diols (Table 7.3.1). This is supported by the observation that expression of *ACE1* in *A. oryzae* or *P. oryzae* resulted in the production of 12,13-dihydroxypyriculariapyrone **193** and the closely related 10,11-dihydroxypyriculariapyrone **194**.⁹⁶ A similar epoxidation of a polyunsaturated polyketide followed by hydrolysis was observed when *tenS* from the tenellin **9** BGC was expressed in *A. oryzae* which resulted in the formation of protenellin C **195**.²²² This suggested that *A. oryzae* contains a monooxygenase selective for terminal methylbutenyl motifs, since the triene structure of the ACE1 RAP1 product **67** (Section 3.3.1) did not serve as a substrate. *P. oryzae* is known to produce compounds containing a diol such as pyriculol **187** and pyriculariol **188**, albeit in the first case the mid-chain polyene was epoxidised by a monooxygenase.^{223,219} The same monooxygenase responsible for the formation of the diol in **188** might be responsible for the observed diols in **179** - **186**.

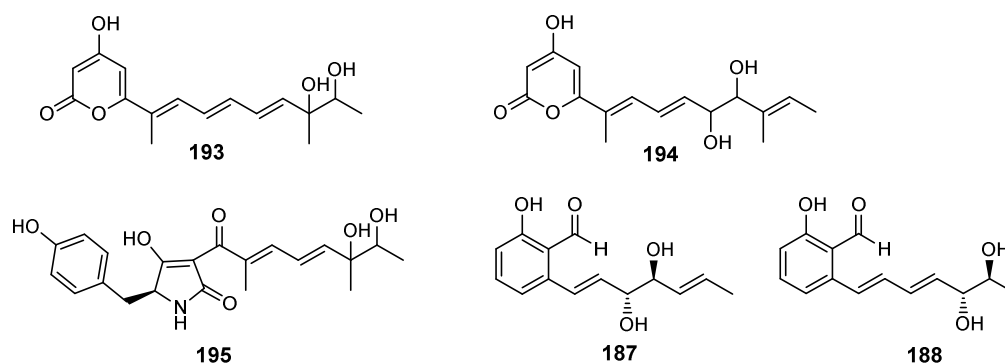
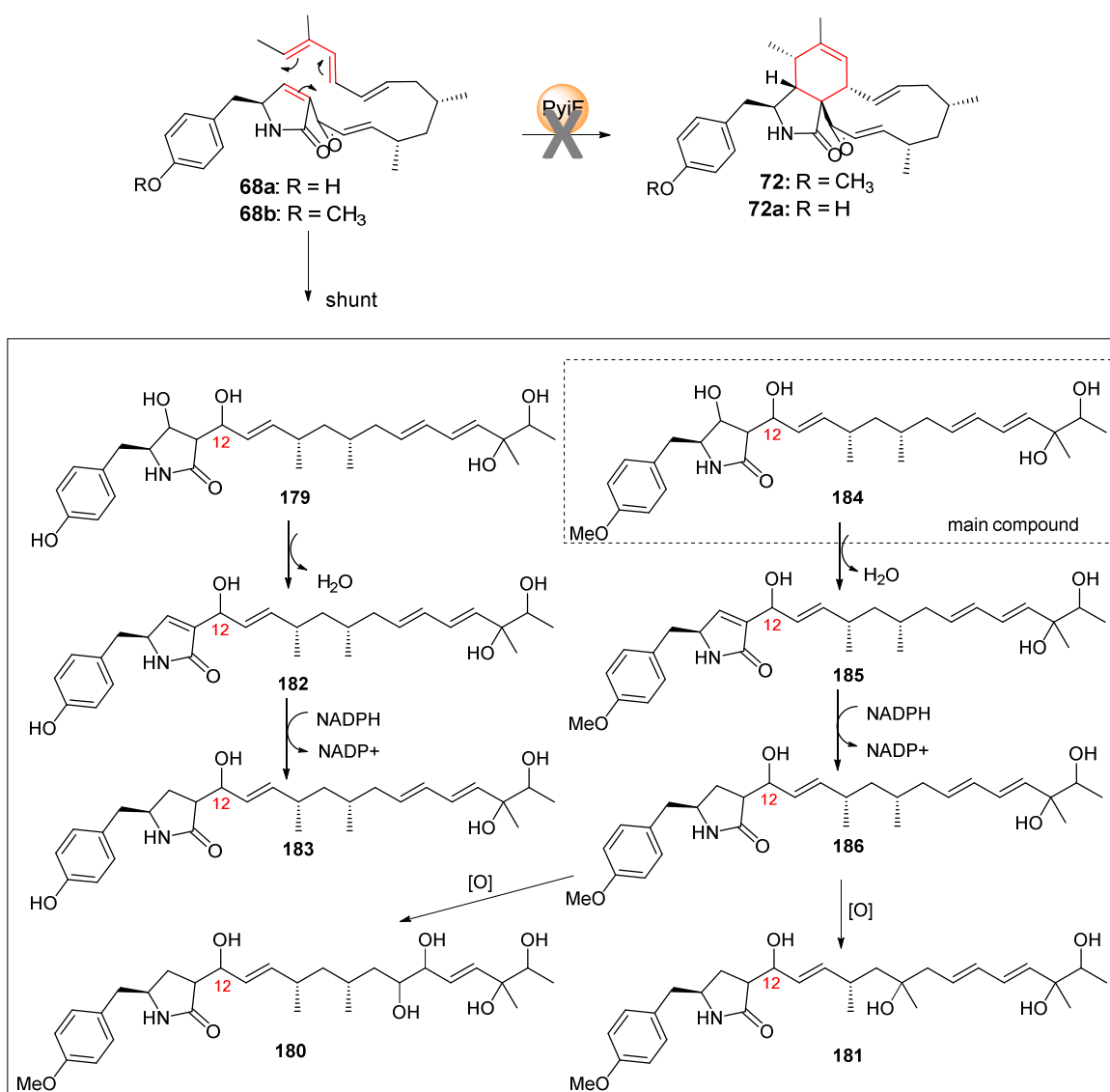


Figure 7.20: Examples of polyunsaturated compounds which were converted into diols.

Notably, in all isolated compounds (**179** - **186**) the expected C-12 ketone was reduced to an alcohol (Scheme 7.12). A carbonyl group at this position would lower the electron density of the dienophile, an energetically favourable event, pushing the DA reaction forward. Therefore, the observed C-12 hydroxyl group probably derived from a reduction step catalysed by an *M. grisea* specific oxidoreductase.

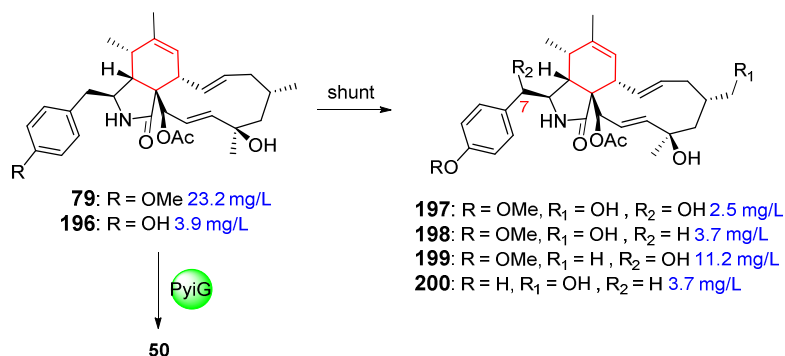
Based on the structures of the isolated compounds a possible reaction pathway can be proposed (Scheme 7.12). Compounds **182**/ **185** are the dehydration products of **179**/ **184**. Reduction of **182**/ **185** leads to compound **183**/ **186**. Compound **180** and **181** are oxidation products of **186**.



Scheme 7.12: Putative DA reaction catalysed by PyiF and isolated intermediates (**179** - **186**) purified from a Δ pyiF *M. grisea* strain.

Additional oxidative steps catalysed by *M. grisea* specific enzymes could also be observed for late-stage biosynthetic intermediates of **50**. Targeted inactivation of *pyiG*, encoding a P450

monooxygenase, abolished the production of **50**, and five new compounds **196** – **200** were observed, whereof **197** - **200** are shunt products (Scheme 7.13).¹⁰⁸ The major compound is **199** which contains an additional hydroxyl group at C-7.



Scheme 7.13: Oxidised shunt products observed in Δ *pyiG* *M. grisea* strains.¹⁰⁸

Tyrosine *O*-methylated and unmethylated compounds (**179** – **186**) were isolated. As recently demonstrated, the methyl group is transferred by the *O*-methyltransferase P*yiA*, which is encoded within the **50** BGC.¹⁰⁸ The main produced compound **184** derived from *O*-methyltyrosine **71**, suggesting that methylation of tyrosine **70** by P*yiA* occurs at an early biosynthetic step (Scheme 1.13).

As expected, homologous complementation of Δ *pyiF* *M. grisea* strains restored the production of **50**, albeit the yield was significantly decreased (around 90 %), indicating that *P_{amyB}* is not as strong as the native promoter of *pyiF* and/or the location of the gene within the genome affects its transcription or the availability of P*yiF*. To exclude these effects one could clone *pyiF* under control of its native promoter and introduce it into the original locus. Nevertheless, these findings prove that apart from *pyiF* no other gene of the **50** BGC was disrupted and that the observed change in biosynthesis is exclusively caused by the loss of P*yiF*.

Remarkably, the Δ *pyiF* *M. grisea* strains were also successfully complemented with *ORF3*, the homolog gene of *pyiF* from the *ACE1* BGC of *P. oryzae* Guy11. Both biosynthetic pathways are predicted to be closely related (see Chapter 2). Pyrichalasin H **50** as well as the *ACE1* metabolite both derived from *O*-methyltyrosine **71**. Furthermore, they share the same methylation pattern. The putative PKS-NRPS backbone **68a/b** of **50** and *ACE1* **91** (illustrated after formation of the pyrrolinone ring) only differ in the polyketide chain length: **50** derived from an octaketide, the putative *ACE1* metabolite from a nonaketide (Figure 7.21). Unfortunately, the activity of *ORF3* as a DAase could not be tested *in-vitro*. However, soluble protein was obtained and can be used in future studies.

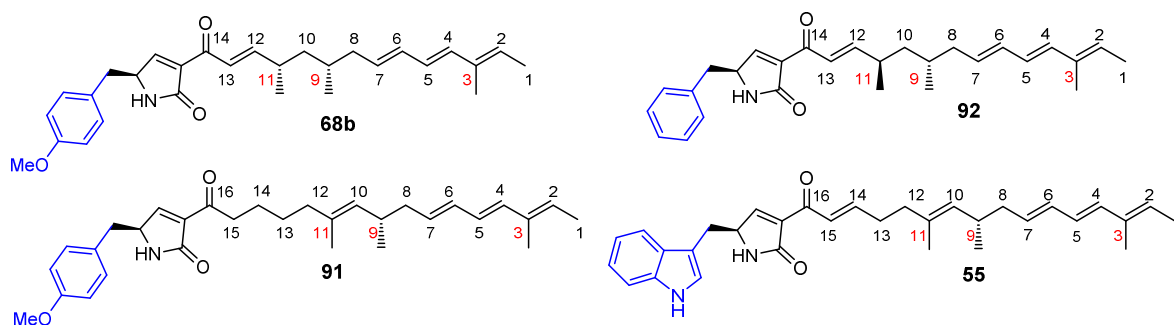


Figure 7.21: Proposed substrates for DAases involved in the biosynthesis of cytochalasans.

Heterologous complementation with *CHGG_01241*, *cssF* and *poxQ* did not restore the production of **50**. The hybrid PKS-NRPS chain of chaetoglobosin A **51** derives from tryptophan fused to a nonaketide, which displays the same methylation pattern as **68a** and **91**. The assumed formation of the pyrrolinone ring leads to the backbone **55**. The polyketide chain of **55** contains a hydroxyl group, increasing the polarity of the backbone.

In the case of cytochalasin E **52**, phenylalanine is fused to an octaketide (backbone **92**). Although the homology of *CssF* and *PyiF* is very high (61 % identity, 71 % similarity), production of **50** could not be restored by complementing the disrupted *pyiF* gene with *cssF*.

These findings indicate that the substrate binding pocket of the putative DAs might be rather specific. *CHGG_01241* and *CssF* were not able to complement the loss of *PyiF*, even though they are quite homologous. *ORF3*, on the other hand, accepted the substrate of *PyiF* (probably **68a/b**), leading to the DA reaction. Although the length of the polyketide chain of **68b** and **91** differs, the incorporated amino acid *O*-methyl tyrosine **71** and the methylation pattern of the polyketide moiety are the same. Nevertheless, during target gene inactivation experiments of the late stage tailoring enzymes of the **50** BGC, a new cytochalasin was observed, which derived from phenylalanine instead of *O*-methyl tyrosine **71**.¹⁰⁸ Therefore, at least the DAase *PyiF* appeared to be able to accept phenylalanine derived substrates, indicating that the selectivity of *PyiF* is, to certain extent, promiscuous.

PoxQ catalyses a decalin forming DA reaction which is not related to the biosynthesis of cytochalasans. Unsurprisingly, the loss of *PyiF* was not complemented in a strain expressing *poxQ*. But neither did it form decalins, probably the pyrichalasin backbone **68b** was not accepted as substrate by *PoxQ*, or the triene was destroyed by oxidation before *PoxQ* could act. In general the formation of a decalin ring starting from the pyrichalasin backbone **68b** would be possible. Larsen

and co-workers identified a decalin-containing compound, niduclavin **62**, when they co-expressed the *A. clavatus* hybrid PKS-NRPS *cssA* and its *trans*-acting ER *cssC* from the cytochalasin E **52** BGC in *A. nidulans*.⁹⁰ However, the authors speculated that the formation of the decalin ring was only favoured due to an additional double bond, generated by a native *A. nidulans* enzyme, between C-2' and C3' (Scheme 1.10). A double bond at this specific position would lead to an additional activation of the dienophile in the $\alpha\beta$ -position of the ketone (C-3). Since the pyrlichalasin backbone **68b** lacks this double bond, formation of the decalin ring is unlikely to occur.

8. Overall Conclusion and Outlook

The major focus of this project was to investigate the biosynthesis of the ACE1 metabolite produced by the rice blast fungus *P. oryzae* Guy11. The ACE1 BGC was reported to be expressed only under very specific conditions during appressorium penetration into plant cells.¹⁵³ Under these conditions, the ACE1 metabolite is produced in very low quantities for a very short period and it has therefore been impossible to isolate and identify the avirulence signal compound. Thus, gene deletion experiments for the ACE1 locus in *P. oryzae* are unlikely to be successful. Previous studies and bioinformatic analysis of the ACE1 BGC strongly suggest that the ACE1 metabolite belongs to the family of cytochalasans.^{96,78}

Another cytochalasan BGC, similar to the ACE1 BGC, was identified in the fungus *M. grisea* and demonstrated to be responsible for the production of pyrichalasin H 50.¹⁰⁸ Bioinformatic analysis of the pyrichalasin and the ACE1 BGC showed that the clusters are homologous (Figure 8.1, Chapter 2). Therefore, the pyrichalasin BGC was used as model system in this work to elucidate the function of two enzymes likely to play a key role in the biosynthesis of cytochalasans: the $\alpha\beta$ -hydrolase (PylE) and the putative Diels-Alderase (PylF). To this end, gene deletion and complementation studies were performed in *M. grisea* (Chapter 6 and 7).

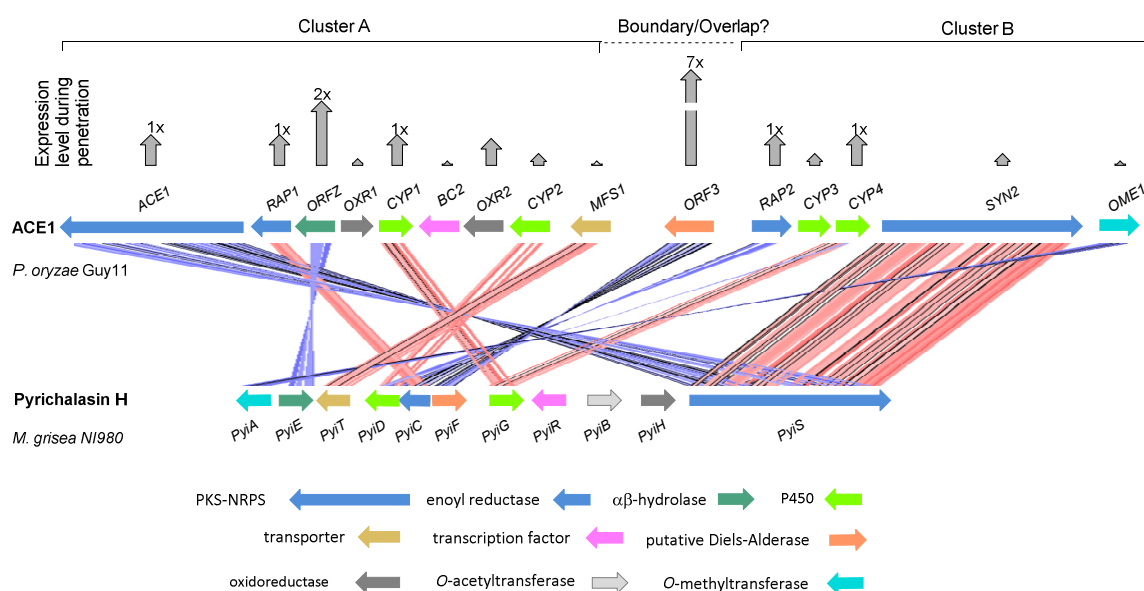


Figure 8.1: ARTEMIS comparison of the ACE1 gene cluster from *P. oryzae* Guy 11 and the pyrichalasin H gene cluster from *M. grisea* N1980. Expression levels of genes from the ACE1 BGC during fungal penetration into the host plant are demonstrated (data from Collemare *et al.*).⁸⁸

By heterologous expression studies, more insights into the early and middle biosynthetic steps of the biosynthesis of the ACE1 metabolite were obtained. *O*-Methyl-L-tyrosine **71** was identified as the native substrate for the A-domain of the ACE1 NRPS (Chapter 3). Feeding *O*-methyl-L-tyrosine **71** to a culture of an *A. oryzae* *ACE1 + RAP1* transformant led to the production of **88** in significant higher yields compared to the previous produced compounds **67** and **86** (2 mg/L). In accordance with this, *O*-methylated compounds were also observed in Δ *pyiF* *M. grisea* deletion mutants, further suggesting that the transfer of the methyl group to the hydroxyl group of tyrosine appears at an early stage of biosynthesis (Chapter 7). Feeding *O*-methyl-L-tyrosine **71** to the *A. oryzae* culture expressing *ACE1* and *RAP1* to form **88** was not required anymore when the *O*-methyltransferase *OME1* was co-expressed (Figure 8.2, Chapter 3). Thus, *OME1* belongs to the *ACE1* cluster; even so its gene locus appears to belong to the *SYN2* cluster (cluster B) in *P. oryzae* Guy11 (Figure 8.1). This assumption is further supported by genome sequencing results of *M. grisea* BR29: in *M. grisea* BR29 the *OME1* gene is located contiguous with the *ACE1* cluster.¹⁰⁸

Furthermore, the first evidence was obtained, by heterologous expression in *A. oryzae*, indicating that the $\alpha\beta$ -hydrolase ORFZ is involved in early biosynthetic steps. Upon co-expression of *ACE1*, *RAP1* and *ORFZ* a new metabolite was formed with *m/z* of 477 g/mol, which would be consistent with the structure of compound **89**. Unfortunately, the structure of **89** was not fully characterised experimentally. However, compounds similar to **89** were identified in Δ *pyiF* *M. grisea* deletion mutants (Chapter 7). Compound **89** might be a shunt product of the pyrrolinone **91**, the expected *ACE1* intermediate and possible product of *ORFZ*. These findings are strong evidence for the operation of (a) native *A. oryzae* enzymes(s) catalysing the reduction of the aldehyde **90** and the pyrrolinone **91** before the DA reaction can occur. Therefore, the biosynthetic pathway of the *ACE1* metabolite probably contains the formation of the pyrrolinone ring **91**. The $\alpha\beta$ -hydrolase might be involved in this step (Figure 8.2).

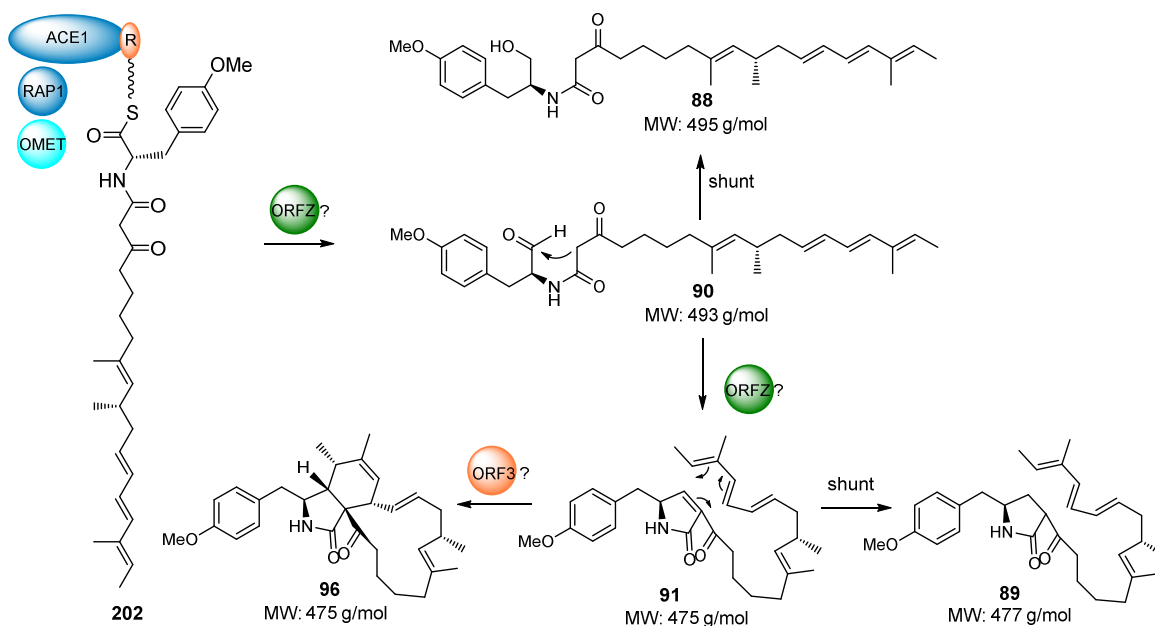


Figure 8.2: Proposed early steps in the biosynthesis of the ACE1 metabolite.

However, the main produced compound in the host *A. oryzae* was **88**, even after all biosynthetic genes of the ACE1 cluster were co-expressed (Chapter 3). Unlike in other known PKS-NRPS systems, the amino acid carbonyl of has been double reduced to an alcohol rather than undergoing the more usual Dieckmann cyclization to form a tetramic acid (biosynthesis of tenellin **9**) or to undergo a single reduction followed by a Knoevenagel condensation, forming the pyrrolinone as proposed for the biosynthesis of fusarin C **8** (Scheme 1.6).^{58,50} *In-vitro* enzyme assays confirmed that reduction of the aldehyde intermediate **90** to the alcohol **88** is catalysed by a native *A. oryzae* enzyme, probably an alcohol dehydrogenase (AC: 439867), at an early biosynthetic step, thereby stalling the biosynthetic pathway of ACE1 (Chapter 4). Disrupting its function, *e.g.* by targeted gene deletion, might recover *A. oryzae* as suitable host for the heterologous expression of the ACE1 BGC as well as for other biosynthetic pathways where an aldehyde intermediate is formed.

Ectopic expression of *BC1*, encoding a putative pathway-specific transcription factor from the ACE1 BGC, in *P. oryzae* Guy11 did not lead to the production of ACE1-related compounds (Chapter 5). However, the known compound hinnulin A **105** was formed, which is probably formed from 1,8-dihydroxynaphthalene (DHN) **106**. 1,8-DHN **106** is the known precursor of melanin in *P. oryzae* and other fungi. A putative partial gene cluster potentially involved in the biosynthesis of **105** and melanin was validated by RT-PCR and a possible biosynthetic pathway was proposed (Scheme 5.4). Interestingly, it is known that the ACE1 signal molecule is produced in the appressoria of *P. oryzae*, which are also heavily melanised. The linkage between the ACE1 BGC

and a putative melanin BGC in *P. oryzae* is therefore probably not coincidental. However, further detailed experiments will be required to unpick these relationships.

Further attempts to elucidate the function of the $\alpha\beta$ -hydrolase PyiE during the biosynthesis of **50** were performed in *M. grisea*. Disruption of the *pyiE* gene almost abolished production of **50** (Chapter 6). Remarkably, a new compound known as cytochalasan H **139** was detected in low yields in the deletion strains. Compound **139** is a phenylalanine analogue of **50**. Notably, disruption of *pyiA*, the gene encoding the *O*-methyltransferase from the **50** BCG, also led to the formation of **139** in low titres.¹⁰⁸ Besides **139**, no additional compounds related to **50** were observed. The *pyiA* gene is located directly upstream of *pyiE* in the opposite orientation, therefore the promoter of *pyiA* is located between *pyiA* and *pyiE*. Since the secondary metabolite profile of $\Delta pyiE$ *M. grisea* looks like the secondary metabolite profile of $\Delta pyiA$ *M. grisea* (*i. e.* extremely low production of **50** and observation of **139**) it is possible that the *pyiA* gene or its promoter region was damaged during the transformation procedure. The loss of PyiA might therefore be responsible for the decreased production of **50**. To test if the *pyiA* gene is non-functional a control experiment could be performed. Addition of *O*-methyl-L-tyrosine **71** to growing $\Delta pyiE$ *M. grisea* cultures would complement the loss of *pyiA*. If the *pyiA* gene was really disrupted, this should lead to the true $\Delta pyiE$ *M. grisea* phenotype. Another control experiment would be to verify the expression of *pyiA* by RT-PCR.

In this work a complementation experiment was performed as control experiment to test if other genes from the pyrichalasin cluster were disrupted. Expressing a functional copy of *pyiE* under the control of P_{amyB} in $\Delta pyiE$ *M. grisea* strains partially restored the production of **50**, albeit the yields were dramatically decreased (only up to 10.9%) compared to the wild type (100%). However, compared to the amount of **50** produced in the $\Delta pyiE$ *M. grisea* control strain (0.4%) the amount of **50** in the transformants was approximately up to 27 times higher. These findings suggest that the *pyiA* gene (and other genes of the cluster) were still functional.

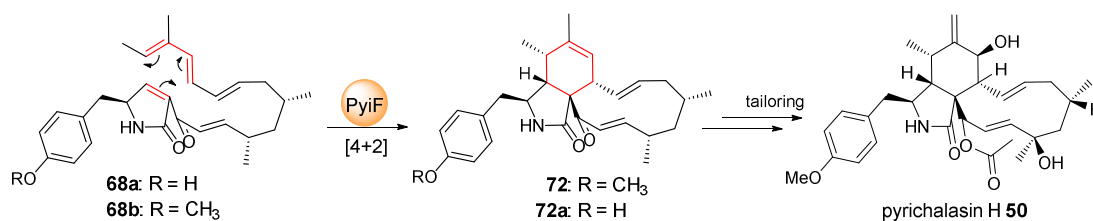
To test if the yields of **50** in the transformants can be increased, a functional copy of *pyiE* was expressed under the control of its native promoter. Surprisingly, homologous complementation with *pyiE* under the control of its native promoter did not restore the production of **50**. For an unknown reason the native promoter seems to be inefficient. One explanation might be that unless the promoter-gene fusion is introduced into its original locus, gene expression is restricted. Screening further transformants might have identified a strain where the DNA fragment $P_{native-pyiE}$ was integrated into its original locus by homologous recombination. However, at least for the

genetically tested transformants the DNA fragment $P_{native}\text{-}pyiE$ integrated ectopically. Homologous complementation of $\Delta pyiE$ *M. grisea* strains with a point mutation of *pyiE* (*pyiE*_{S252A}) under the control of P_{amyB} did also not restore the production of **50**. Thus, Ser252 must be crucial for the activity of PyiE and might be part of the catalytic triad; a typical feature of $\alpha\beta$ -hydrolase-fold proteins. However, until it was not further confirmed that apart from *pyiE* no other genes (e. g. the *pyiA* gene) was disrupted, these results cannot be strongly interpreted.

Enzyme activity of ORFZ towards the synthetic substrate **142** was not observed. There are several explanations for this. Either the protein was inactive, or the substrate **142** was not accepted by ORFZ, since, compared to its putative native substrate **90**, it is lacking the main part of the polyketide chain. A third possibility would be that ORFZ is not involved in the proposed Knoevenagel condensation and is therefore not able to catalyse this reaction. Collectively, the function of ORFZ remains elusive. Furthermore, the results suggest that the Knoevenagel condensation does not occur spontaneously under the tested conditions.

Overall, the results obtained by *in-vivo* experiments in *M. grisea* cannot be strongly interpreted because of the possible damage to the *O*-methyltransferase gene *pyiA* or its promoter. Further experiments will be needed to test if the loss of *pyiA* led to the observed $\Delta pyiE$ *M. grisea* phenotype. By *in-vitro* enzyme assays no evidence could be found for the involvement of the $\alpha\beta$ -hydrolase in the Knoevenagel condensation. Therefore, the activity of the $\alpha\beta$ -hydrolase remains elusive for now.

Targeted gene disruption of *pyiF* (putative Diels-Alderase) in *M. grisea* completely abolished the production of **50**, concurrently resulting in the accumulation of several new related compounds (Chapter 7). Structure elucidation of eight compounds (**179** - **186**) revealed that all compounds shared an underlying linear structure probably derived from the hybrid PKS-NRPS backbone **68a/b**, the putative substrate of PyiF. As expected, homologous complementation of $\Delta pyiF$ *M. grisea* mutants restored the production of **50**. Remarkably, the $\Delta pyiF$ *M. grisea* strains were also successfully complemented with *ORF3*, the gene homolog of *pyiF* from the *ACE1* BGC of *P. oryzae* Guy11. These findings confirm the idea that both pathways are closely related. Furthermore, they strongly suggest that *pyiE* and *ORF3* encode a Diels-Alderase, which catalyses a [4+2] cycloaddition in order to form the isoindole macrocycle of *ACE1* and **50**, respectively (Scheme 8.1).



Scheme 8.1: Proposed function of PylF in the biosynthesis of pyrichalasin H 50 as a Diels-Alderase catalyzing an intramolecular [4+2] cyclisation.

In-vitro enzyme assays with ORFZ and/or ORF3 using model substrates were not successful. So far, no stable conditions for the substrate analogues **142**, **192** and **91** were achieved. However, soluble protein was obtained for ORFZ and ORF3 and can be used in future studies (Chapter 6 and 7).

Outlook and Future Experiments

In future work, the ACP-R domain of the ACE1 NRPS could be heterologous produced in *E. coli*. Subsequently, its native substrate, the PKS-NRPS backbone **202**, has to be chemically attached. Thereby, the correct release mechanism could be investigated *in-vitro*. This would answer a) the question if an aldehyde **90** is released and b) whether the formation of the pyrrolinone **91** is spontaneous or if it needs enzymatic catalysis. Adding ORFZ and/or ORF3 to the reaction mixture might give further insights into their biosynthetic role.

Another promising experiment would be to use the Δ *pyiS* *M. grisea* deletion strain constructed by our group as host for co-expressing ACE1 biosynthetic genes. The *pyiS* gene encodes the pyrichalasin H 50 PKS-NRPS. Undesired modifications of the aldehyde intermediate **90** by native *M. grisea* enzymes are unlikely to occur. Since **50** is no longer produced in this mutant strain, competition for substrates with the ACE1 metabolite pathway should be avoided.

The chromatin structure at the ACE1 locus is crucial for its transcriptional activation. Investigations of the respective chromatin status will gain more insights and might help to understand the tight regulation of this cluster. Altering the conformation of the chromatin, *e. g.* by inhibiting the methylation of histones and/or preventing their deacetylation, either genetically or chemically, might help to loosen the chromatin structure. Subsequently, ectopic expression of BC2 would be more promising to result in the production of the ACE1 metabolite. In addition, gene deletion experiments might be possible after the chromatin structure has been loosened.

In-vitro enzyme assays with ORFZ and/or ORF3 are still a promising approach to elucidate their function and to gain more insights into the biosynthesis of cytochalasans. The first aim will be to achieve stable and standardized conditions for the enzyme assays. Subsequently, the activity of both enzymes can be tested using appropriate substrates.

9. Experimental

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

9.1 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 9.1.1: 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
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 ₄) ₂ SO ₄ , 2 % (w/v) D(+)-Glucose Monohydrate, 0.077 % (w/v) Complete supplement mixture minus Uracil, 1.5 % (w/v) Agar

Table 9.1.2: Media and agar used in this work for *A. oryzae*.

Media	Composition
CZD/S agar	3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) (NH ₄) ₂ SO ₄ , 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 ₄) ₂ SO ₄ , 0.15 % (w/v) L-methionine, 0.05 % adenine, 0.8 % (w/v) agar
CZD/S1 agar (-met)	3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) (NH ₄) ₂ SO ₄ , 0.15 % (w/v) L-methionine (optional), 1.5 % (w/v) agar
CZD/S1 soft gar (-met)	3.5 % (w/v) Czapek Dox broth, 18.22 % (w/v) D-sorbitol, 0.1 % (w/v) (NH ₄) ₂ SO ₄ , 0.15 % (w/v) L-methionine (optional), 0.8 % (w/v) agar
CD2 agar (-met)	3.5 % (w/v) Czapek Dox broth, 4.68 % (w/v) NaCl (0.8 M), 0.1 % (w/v) (NH ₄) ₂ SO ₄ , 0.15 % (w/v) L-methionine (optional),
DPY	2 % (w/v) Dextrin from potato starch, 1 % (w/v) Polypeptone, 0.5 % (w/v) Yeast extract, 0.5 % (w/v) KH ₂ PO ₄ , 0.05 % (w/v) MgSO ₄ * x H ₂ O
DPY Agar	DPY Medium, 2.5 % (w/v) Agar
GNB	2 % (w/v) D(+)-Glucose Monohydrate, 3 % (w/v) Nutrient broth Nr. 2 from Oxoid

Table 9.1.3: Media and agar used in this work for *P. oryzae* and *M. grisea*.

Media	Composition
Trace elements (TE)	7.9 mg/L ZnSO ₄ .7H ₂ O, 0.6 mg/L CuSO ₄ , 0.1 mg/L H ₃ BO ₃ , 0.2 mg/L MnSO ₄ .H ₂ O, and 0.14 mg/L NaMoO ₄
Vitamin solution (V)	1 mg/L thiamine, 5 µg/L biotin
CM	10 g/L glucose, 2 g/L peptone, 1 g/L yeast extract, 1 g/L casamino acids, 6 g/L NaNO ₃ , 0.5 g/L KCl, 0.5 g/L MgSO ₄ , 1.5 g/L KH ₂ PO ₄ , PH 6.5, after autoclaving add filter sterilised 0.1 % (v/v) V and TE solution
TNK-CP Agar	10 g glucose, 2 g yeast extract, 15 g agar, 2 g NaNO ₃ , 2 g KH ₂ PO ₄ , 0.5 g MgSO ₄ .7H ₂ O, 0.1 g CaCl ₂ .2H ₂ O, and 4 mg FeSO ₄ .7H ₂ O, after autoclaving add filter sterilised 0.1 % (v/v) V and TE solution
TNK-CP-Su Agar	TNK-CP Agar, 200 g/L sucrose

Table 9.1.4: Buffers and Solutions used in this work.

Media	Composition
Coomassie dye	25 % (w/v) acetic acid, 10 % (w/v) Isopropanol, 0.1 % (w/v) Coomassie
Coomassie bleach	25 % (w/v) acetic acid, 10 % (w/v) Isopropanol
IMAC loading buffer	50 mM Tris-HCl, pH 8, 150 mM NaCl, 10 mM Imidazole
IMAC elution buffer	50 mM Tris-HCl, pH 8, 150 mM NaCl, 40-500 mM Imidazole
50 xTAE buffer	2 M Tris acetate, 0.05 M EDTA, pH 8.3
4X Lämmli buffer	10 % (v/v) β-Mercaptoethanol, 2,5 mg/l Bromophenol Blue 0.25 % (w/v) Bromophenol Blue, 30 % (w/v) Glycerine, 0.25 % (w/v) Xylene Cyanol,
Loading-dye (6 x)	50 mM Na ₂ HPO ₄ , pH 8 25 mM Tris, 192 mM Glycine, 0.1 % (w/v) SDS, pH 8.3
10X SDS running buffer	25 mM Tris, 192 mM Glycine, 0.1 % (w/v) SDS, pH 8.3

Table 9.1.5: Solutions used for fungal transformations.

Media	Composition
<i>A. oryzae</i>	
Solution 1	0.8 M NaCl, 10 mM CaCl ₂ , 50 mM Tris-HCL, pH 7.5
Solution 2	60 % PEG 3350, 0.8 M NaCl, 10 mM CaCl ₂ , 50 mM Tris-HCl, pH 7.5
<i>M. grisea/ P. oryzae</i>	
STC buffer	200 g/L sucrose, 50 mM CaCl ₂ , 50 mM Tris-HCl, pH 7.5
PTC buffer	60 % PEG 3350, 1 M Sorbitol, 50 mM CaCl ₂ , 50 mM Tris-HCl, pH 7.5
TB4 buffer	200 g/L sucrose, 3 g/L yeast extract

9.2 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. Carbenicillin was used in place of ampicillin. Stock and working concentrations are listed in Table 9.2.1.

Table 9.2.1: Used antibiotics in this work.

Antibiotic	Stock [mg/mL]	Final [µg/mL]
Carbenicillin	50 in H ₂ O	50
Chloramphenicol	30 in EtOH	30
Gentamycin	20 in H ₂ O	20
Glufosinat-ammonium (BASTA)	50 in H ₂ O	50-100
Hygromycin B	50 in H ₂ O	120-240
Kanamycin	50 in H ₂ O	50
Tetracyclin	50 in H ₂ O	10

All enzymes used in this study were obtained from NEB or Thermo Scientific and used following the manufacturers' instructions

9.3 Microbiology Methods

Microbial strains which were kindly provided for this work or purchased from Thermo Fischer Scientific, Stratagene or Novagen are listed in Table 9.3.1 and Table 9.3.2.

Table 9.3.1: *E. coli* strains used in this work.

Strain	Genotype	Origin
BL21 (DE3)	F ⁻ <i>ompT hsdS_B (r_B⁻, m_B⁻) dcm gal λ(DE3)</i>	Thermo Fisher Scientific
BL21-CodonPlus RP	F ⁻ <i>ompT hsdS(rB – mB –) dcm+ Tet^R gal endA Hte [argU proL Cam^R]</i>	Stratagene
OneShot Top10	F ⁻ <i>mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (Str^R) endA1 nupG</i>	Thermo Fisher Scientific
OneShot ccdB survival 2T1 ^R	F ⁻ <i>mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araΔ139 Δ(ara-leu)7697 galU galK rpsL (Str^R) endA1 nupG fhuA::IS2</i>	Thermo Fisher Scientific
Rosetta 2	F ⁻ <i>ompT hsdS_B(r_B⁻ m_B⁻) gal dcm (DE3) pRARE2 (Cam^R)</i>	Novagen

Table 9.3.2: Fungal strains used in this work.

Fungal strain	Features	Reference
<i>A. oryzae</i> NSAR1	<i>ΔargB sC adeA niaD</i>	Lazarus group, Bristol
<i>A. oryzae</i> NSAR1 <i>ACE1 RAP1</i>	<i>ACE1 RAP1</i>	Dr. Elizabeth Skellam
<i>A. oryzae</i> NSAR1 <i>ACE1 RAP1 ORFZ</i>	<i>ACE1 RAP1 ORFZ</i>	Dr. Elizabeth Skellam
<i>M. grisea</i> NI981	wild type	Lebrun group, Thiverval-Grignon
<i>P. grisea</i> Guy11	wild type	Lebrun group, Thiverval-Grignon
<i>S. cerevisiae</i> CEN.PK2	MATa/a ura3-52/ura3-52 trp1-289/trp1-289 leu2-3_112/leu2-3_112 his3 D1/his3 D1MAL2-8C/MAL2-8C SUC2/SUC2	Hahn group, Hannover

9.3.1 *Escherichia coli* Strains

Growth and maintenance

E. coli strains were cultivated on liquid or solid LB medium/ agar with appropriate antibiotics. Cultures were incubated 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.

Heat-shock transformation of *E. coli* Strains

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.

9.3.2 *Saccharomyces cerevisiae* CEN.PK2

Growth and maintenance

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.

Transformation of competent *S. cerevisiae* cells

Preparation and transformation of competent *S. cerevisiae* cells was done using the LiOAc/SS carrier DNA/PEG protocol developed by Gietz and Woods.²²⁴ 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.

9.3.3 *Aspergillus oryzae* NSAR1

Growth and maintenance

Aspergillus oryzae NSAR1 was grown on DPY agar plates or in DPY medium at 28 °C for 3-7 days (120 rpm). For long term storage *A. oryzae* spores were collected from plates by adding 3 ml distilled water, mixed with glycerol (final concentration 50 %) and stored in cryovials at –80 °C.

PEG-mediated transformation of *Aspergillus oryzae* NSAR1

The growth medium for maintenance of the strains and for obtaining mycelium for protoplast formation was DPY. DPY plates were incubated at 28°C for 5-7 days. 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 protoplasting 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 100 µL of the protoplast suspension and incubated on ice for 2 min. After adding 1 mL of fungal transformation solution 2 the transformation mixture was incubated 20 min at room temperature. 5 mL of molten (50°C) CZD/S1 soft agar (-methionine) was added to the transformation mixture and overlaid onto CZD/S1 plates (-methionine). 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 (-methionine) (2nd and 3rd selection plate) to reduce the number of false-positive transformants. From the 3rd selection plate a single colony was picked with a toothpick and grown alone on CD2 agar (-methionine) 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 120 rpm followed by extraction and chemical analysis (Section 9.5).

9.3.4 *Magnaporthe grisea* and *Pyricularia oryzae*

Growth and maintenance

P. oryzae and *M. grisea* were grown on DPY or CM plates at 25 °C for 10-14 days. For long term storage 50 mL DPY medium was inoculated with 1 mL of a spore suspension (spores were collected by adding 3 mL distilled water to the agar plates) and incubated at 25 °C for 5-7 days, shaking at 110 rpm. The culture suspension was mixed with glycerol (final concentration 50 %) and stored in cryovials at -80 °C.

PEG-mediated transformation of *M. grisea* and *P. oryzae*

Preparation and transformation of *P. oryzae*/*M. grisea* protoplasts followed reported methods described for *M. grisea*,²²⁵ the only modification was that glass beads were used for homogenization of the cells.

In brief, *P. oryzae* was grown on complete medium (CM) plates for 10-12 days and used to inoculate 100 ml CM liquid media, which was incubated at 25 °C for 3-4 days with shaking at 150 rpm. Mycelia were homogenized using glass beads and transferred to new 100 mL CM followed by incubation at 25 °C for 1-2 days. Mycelia was collected by filtration and protoplasts produced by an enzyme mixture (*Trichoderma harzianum* lysing enzymes (20 mg/ml) and Driselase enzyme (10 mg/ml) in 0.9 M NaCl). Protoplasts were filtered through miracloth, collected by centrifugation (2000 rpm for 10 min) and resuspended in STC buffer. Previously prepared protoplasts which had been frozen and stored at -80 °C were used directly for each transformation experiment. Protoplasts were diluted to 3×10^7 /ml and approximately 5 µg (20 µl) of appropriate plasmid or DNA fragments were added to 200 µl protoplast solution. The mixture was incubated on ice for 30 min. After adding 1 ml of PTC buffer the mixture was incubated at RT for 30 min. 4 ml of TB4 recovery medium was added and the protoplast were allowed to regenerate overnight, gently shaking. The supernatant was reduced by centrifugation to 1 ml and added to 20 ml of molten TNK-SU-CP containing 250 µg/ml hygromycin B or 50 µl BASTA. Plates were incubated at 25 °C for 5-7 days until the appearance of colonies. Transformants were further selected to monoclonies on TNK-CP-Hyg/ TNK-CP-BASTA plates.

9.4 Molecular Biology Methods

9.4.1 Oligonucleotides

All oligonucleotides used in this work were purchased from Sigma Aldrich (Table 9.4.1).

Table 9.4.1: Oligonucleotides used in this work

Name	Cox ID	5'-3' sequence
HygB F1 R	22	CTAGAAAGAAGGATTACCTC
HygB F2 F	23	CTGTGCGAGAAGTTTCTGATCG
ORFZ/enoF	39	CGACTGACCAATTCGCGAGCTCGTCAAAGGATGCACCACTTTTTCAAAGG
ORFZ/enoR	40	CAGGTTGGCTGGTAGACGTCATATAATCATCTAAAGCTGGGGCCGCTCAA
ORF3/gpdR	42	TACGACAATGTCCATATCATCAATCATGACCTAGACATAATTGCTCTTGA
P5-Hyg-1F	55	CATGATGGGATCCTCTAGTG
P8-Hyg-2R	56	CAGGTCGAGTGGAGATGTG
Peno plugF	87	CTTCTTAAATATCGTTGTAAGTGTTCCTGA
Peno plugR	88	CGAAGTATATTGGGAGACTATAGCTACTAG
Padh plugF	89	ATTCACCACTATTATCCCACCCTATAATA
Padh plugR	90	GAGACGAAACAGACTTTTTTCATCGCTAAAA
PgdpA plugF	91	CTTTTCTTTCTTTTCTTTTCCCCTCTTC

Name	Cox ID	5'-3' sequence
PgpdA plugR	92	TGACCTCTAAAACCCAGTG
M13 F	146	TGTA AACGACGGCCAGT
M13 R	145	CAGGAAACAGCTATGACC
FP pTYGS PamyB	147	TGCTTGGAGGATAGCAACCG
RP pTYGS TamyB	148	GGGGATGACAGCAGTAACGA
T7-Seq-F	231	TAA TAC GAC TCA CTA TAG GG
T7term-Seq-R	232	CTAGTTATTGCTCAGCGGT
ORF3long/gdpA F	617	ACAGCTACCCCGCTTGAGCAGACATCACCGATGATTGTGTCCCACACTCT
HygB control F1 R	618	GCTCGACGTATTTCAAGTGC
HygB control F2 F	619	GCTCCGTAACACCCAATACG
pyiF-KO-5F	663	TTGTACAAAAAAGCAGGCTCCGCGGTACCGAATAAGCAACATTACATC
pyiF -KO-5R	664	ACGTATTTCAAGTGTGAAAGATCCACTAGACTACTGTGGATGCGTTCTCA
pyiF -KO-3F	665	AGCGCCCACTCCACATCTCCACTCGACCTGACTACTATTGTCGCAACCT
pyiF -KO-3R	666	TGTACAAGAAAGCTGGGTGCGCGGAAGCTTGCCACCCCTCGATTGAGAG
Tubulin F	727	TGGTTGCCAGATCGCCAATTC
Tubulin R	728	CATCACCACGGTACAGCAAGC
pyiF -KO-3R short	729	CTCGATTGAGAGGAGAAGCC
pyiF -KO-5F short	730	CCGAATAAGCAACATTACATCG
ORFZ-Ndel-F	731	TCGTGCATATGATGCACCACTTTTTC
ORFZ-XhoI-R	732	TGTACCTCGAGCTAAAGCTGGGGCC
ORF3-Ndel-F	733	TCGTACATATGATGATTGTGTCCCAC
ORF3-XhoI-R	734	TGTACCTCGAGCTAGACATAATTGC
pyiE-KO-5F	735	TGTACAAAAAAGCAGGCTCCGCGGTACCCATCAACTTTACGAAAACGGGC
pyiE -KO-5R	736	ACGTATTTCAAGTGTGAAAGATCCACTAGACTACAATGAAAACACAAGC
pyiE KO-3F	737	AGCGCCCACTCCACATCTCCACTCGACCTGTCGCCACCAGTTTAATTTGC
pyiE KO-3R	738	GTACAAGAAAGCTGGGTGCGCGGAAGCTTTAGAATGCCGCTATTTTGCA
pyiE -KO-5F short	773	CATCAACTTTACGAAAACGGGC
pyiE -KO-3R short	774	TAGAATGCCGCTATTTTGCA
pyiF-KO control F	802	CTGCCCAGCTTCATATTTG
pyiF -KO control R	803	TTGGGGAAGGTCAAGGCCT
pyiE -KO control F	804	AGTGCACACTGTAGGTTAGC
pyiE -KO control R	805	TTTGTACGCTGCCGCGAGCAT
pEYA-ORF3 F	806	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGATTGTGTCCCACACTCT
pEYA-ORF3 R	807	AACTTTGTACAAGAAAGCTGGGTGCGCGCTAGACATAATTGCTCTTGA
pEYA-ORFZ F	808	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGCACCACTTTTCAAAGG
pEYA-ORFZ R	809	AACTTTGTACAAGAAAGCTGGGTGCGCGCTAAAGCTGGGGCCGCTCAA
pEYA-pyiF F	845	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGCTGCCAGCTTCATATT
pEYA-pyiF R	846	AACTTTGTACAAGAAAGCTGGGTGCGCGCTATACATAGTTACTTCGGA
pEYA-pyiE F	847	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGCATAAGTTCTTCCCGCG
pEYA-pyiE R	848	AACTTTGTACAAGAAAGCTGGGTGCGCGCTAAATGTGTGCTGAAACAT
cDNA ACE1 F	855	CTCAAGACCTACGCCACCCTG
cDNA ACE1 R	856	CTGCCTTGATGAGGTATTAG
cDNA ORF3 F	857	CTCAACTCGGATTCGCTC
cDNA ORF3 R	858	CATAATTGCTCTTGAACAGTGG
cDNA ORFZ F	859	GATCTGGGATATGCCGTTTTG
cDNA ORFZ R	860	CAATTCCAAAGTGCTCATCGAG
pyiE F	907	ATGCATAAGTTCTTCCCGCG
pyiE R	908	CTAAATGTGTGCTGAAACATCC
pyiF F	909	ATGCTGCCAGCTTCATA
pyiF R	910	TCGGAACAGAGGTGAGTTC
ccsF Aug pEYA-F	911	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGCCTCTCACGGAGCAGTG
ccsF Aug pEYA-R	912	AACTTTGTACAAGAAAGCTGGGTGCGCGCTTATGCCACGTAGTTACTGC
ccsF ord pEYA-F	913	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGCGTACCTCCATTCTACT
CHGG_01241 pEYA F	914	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGAAACTTTGTGCTCTCTT
CHGG_01241 pEYA R	915	AACTTTGTACAAGAAAGCTGGGTGCGCGCTACAGCTCCCTTTCTCCT
PoxQ FgenesH pEYA F	916	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGGCTCGAATCCCATTGGA
PoxQ FgenesH pEYA R	917	AACTTTGTACAAGAAAGCTGGGTGCGCGCTACTGCCCCCATTCCTCC
OXR1-Padh F	924	TTTCTTTCAACAGAAGATCCCAAAGTCAAATGTCTGTGCCAAGTTTCAT
OXR1-Tadh R	925	TTCATTCTATGCGTTATGAACATGTTCCCTTCAAGTGGTGCGAGATTCAT
OXR2-Pgpd F	926	AACAGCTACCCCGCTTGAGCAGACATACCATGCGTTCAATTATCTCGGC
OXR2-Tgpd R	927	ACGACAATGTCCATATCATCAATCATGACCTCACCTCTTCTTTGAAACG
CYP1-Peno F	928	GTCGACTGACCAATTCGCGAGCTCGTCAAATGATGACTCCAACCGAAT

Name	Cox ID	5'-3' sequence
CYP1-Teno R	929	GGTTGGCTGGTAGACGTCATATAATCATAACCTATTTCTCCCTCGTTCA
PamyB ccsF long F	930	TCTGAACAATAAACCCACAGCAAGCTCCGATGCGTACCTCCATTCTACT
PamyB ccsF F	931	TCTGAACAATAAACCCACAGCAAGCTCCGATGCGTCTCACGGAGCAGTG
TamyB ccsF R	932	CTCTCCACCCTTCACGAGCTACTACAGATCTTATGCCACGTAGTTACTGC
PamyB CHGG_01241 F	933	TCTGAACAATAAACCCACAGCAAGCTCCGATGAAACTTTGTGCTCTCTT
TamyB CHGG_01241 R	934	CTCTCCACCCTTCACGAGCTACTACAGATCCTACAGCTCCCTTTCTCCT
PamyB PoxQ F	935	TCTGAACAATAAACCCACAGCAAGCTCCGATGGCTCGAATCCCATTGGA
TamyB PoxQ R	936	CTCTCCACCCTTCACGAGCTACTACAGATCCTACTGCCCCCATTCCTCC
ccsF Intron F	941	CGGTTCTGGTCTGCGTTTAGCTGGTTCACCTGTCTTACGCACCTCAGCGC
ccsF Intron R	942	GCGCTGAGGTGCGTAAGACAGGTGAACCAGCTAAACGCAGACCAGAACCG
PamyB pyiE	998	TCTGAACAATAAACCCACAGCAAGCTCCGATGCATAAGTTCTTCCCGCG
TamyB pyiE	999	CTCTCCACCCTTCACGAGCTACTACAGATCCTAAATGTGTGCTGAAACAT
CYP3 PgpdA F	1047	AACAGTACCCCGCTTGAGCAGACATCACCATGTTAGATATGTCCCAACC
CYP3 TgpdA R	1048	ACGACAATGTCCATATCATCAATCATGACCCTAAGCATTGTTGAGGTTTC
CYP4 Padh F	1049	TTTCTTTCAACACAAGATCCCAAAGTCAAATGTTGTCTCTTTTGAAAC
CYP4 Tadh R	1050	TTTCTTCTATGCGTTATGAACATGTTCCCTCTACTCAACATTCCTGGACT
OXR2 pEYA F	1051	TTATAATGCCAACTTTGTACAAAAAGCAGATGCGTTCAATTATCTCGGC
OXR2 pEYA R	1052	TTATAATGCCAACTTTGTACAAGAAAGCTGTCATATTTCCGCTTTACCTG
OMET pEYA F	1053	TTATAATGCCAACTTTGTACAAAAAGCAGATGTCAACAATGGCCCTCCA
OMET pEYA R	1054	TTATAATGCCAACTTTGTACAAGAAAGCTGTCACGAAAGGCAGACTTCCA
ORF3ACE1opt_BsmBI_F	1055	ATCGTCTCtggtggGATGATTGTGAGCCATACCTCGC
ORF3ACE1opt_XhoI_R	1056	TActcgagCTACACATAATTGCTTTTAAACAGC
ORFZACE1opt_BsmBI_F	1057	ATCGTCTCtggtggGATGCACCACTTTTTCAAAGGCCAGTTTTTCG
ORFZACE1opt_XhoI_R	1058	TActcgagCTAAAGCTGGGGCCGCTCAATTCC
BC2 PamyB F	1063	CTGAACAATAAACCCACAGCAAGCTCCGAATGTTGATGGATTGGAATTA
BC2 TamyB R	1064	CTCTCCACCCTTCACGAGCTACTACAGATCCTATGGCCCGGTGGTCTTC
hph Pgpd F	1065	TAACAGCTACCCCGCTTGAGCAGACATCACATGCCTGAACTCACCGCGAC
hph Teno R	1066	AGGTTGGCTGGTAGACGTCATATAATCATACTATTCTTTGCCCTCGGAC
amyB OMET F	1097	TCTGAACAATAAACCCACAGCAAGCTCCGATGTCAACAATGGCCCTCCA
amyB OMET R	1098	CTCTCCACCCTTCACGAGCTACTACAGATCTCAGAAAGGCAGACTTCCA
pyiE PM F	1099	CCATTGCCGTATCCGGTGCCGCATGGGTGGTACTATGCGCTGCGATCT
pyiE PM R	1100	AGATCGCAGCGCATAGTAACCACCCATGGCGGCACCCGGATACGGCAATGG
native Prom PyiE F	1176	ATCATGGTGTTTTGATCATTTTAAATTTTTAGTTAATGCAGACCAGGTCA
Cyp2 F	1299	ATGACGACAGTCCAGAAAGA
Cyp2 R	1300	CTATAACCTCCTTTCTCGAA
BC2 F	1521	ATGTTGATGGATTGGAATTA
BC2 TamyB	1522	CTTACGAGCTACTACAG
BC2 R	1523	CTATGGCCCGGTGGTCTTC
ACE1 cDNA new F	1524	CAGCATCAGCAGCTGACG
RAP1 cDNA F	1525	GGCTATCACCACTGCTCG
RAP1 R	1526	CTACAGCCGCACAACAATCT
melanin 1 F	1527	CTCGTAATTTGCGGTCGCCTC
melanin 2 R	1528	CAGCTGCGTCAGCTGTTTG
melanin 2 F new	1529	GCTCTGAGGAAAGAGATCG
melanin 2 R new	1530	GTGAACGTCTAGTCCCTTGG
T4NR F	1531	GGCACACTTTCTCTGAGCAT
T4NR R	1532	ACTAGCTCATGTCGACCCATC
TF 1 F	1533	TGGTGCCACCGCAATCTCAG
TF 2R	1534	CGGCCTCTCAATCTCCTTGA
TF 2 F	1535	CATGTACATGGTCTGGTC
TF 2 R	1536	CGCCGATGTTGCTTGACCA
laccase F	1537	GTCAAGGGCAAGACCTACCG
laccase R	1538	TCACCCATGGTAGACAGTGTC
L2 F	1539	CAGCTACAATCGCGCATGCAC
L2 R	1540	CTATCGAGGCGTAATGGAT
L3 F	1541	GTCTTTGCCGACGCATCGAG

9.4.2 Polymerase Chain Reaction (PCR)

PCR was used to amplify DNA fragments from plasmid or genomic DNA templates. For cloning the proofreading Q5[®] High-Fidelity 2X Master Mix (New England Biolabs) was used following the manufacturers' instructions. Sequences were checked by DNA sequencing (Mix2Seq kit, Eurofins). OneTaq[®] 2X Master Mix (New England Biolabs) with standard buffer was used, following the manufacturers' instructions, for screening purposes

9.4.3 Obtained and Constructed Vectors in this Thesis

Construction of fungal and bacterial expression vectors was performed using the described *in-vitro* cloning procedures (Section 9.4.10) and *in-vivo* homologous recombination in *S. cerevisiae* (Section 9.3.2).

Table 9.4.2: Vectors kindly provided or purchased for this work.

Name	Origin
pE-YA	Lazarus Group, Bristol
pTYGSargB	Lazarus group, Bristol
pTYGSade	Lazarus group, Bristol
pTYGSmet	Lazarus group, Bristol
pTYGSniaD	Lazarus group, Bristol
pTYGSniaD	Lazarus group, Bristol
pTYGSbar	Lazarus group, Bristol
pTH-GS-egfp	Lazarus group, Bristol
pET28a (+)	Hahn group, Hannover
pCA528	Turgay group, Hannover
pTAYA-ACE1-RAP1	Dr. Elizabeth Skellam
pET100/D-TOPO-ORF3 _{opt.}	Invitrogen
pMA-T <i>cssF</i>	Invitrogen
pMK-RQ CHGG_01241	Invitrogen
pMK <i>poxQ</i>	Invitrogen

Table 9.4.3: Oligonucleotides (ON) used for construction of expression vectors (*A. oryzae* and *P. oryzae*)

Name	ID	vector backbone	ON for construction in <i>S. cerevisiae</i>
pTYGSade-ORFZ-ORF3		pTYGSade	ORF3: 42 + 617; ORFZ: 39 + 40; Patch_ <i>adh</i> : 89 + 90
pTYGSade-ORF3		pTYGSade	ORF3: 42 + 617; Patch_ <i>adh</i> : 89 + 90; Patch_ <i>eno</i> : 87 + 88
pTYGSmet-CYP1-CYP3-CYP4	VBI76	pTYGSmet	CYP1: 928 + 929 ; CYP3: 1047 + 1048; CYP4: 1049 + 1050
pTYGSmet-CYP1-CYP2-CYP3-CYP4	VBI77	pTYGSmet	LR with VBI100
pTYGSade-ORFZ-ORF3-OXR1	VBI80	pTYGSade	ORF3: 42 + 617; ORFZ: 39 + 40; OXR2: 924 + 925
pTYGSade-ORFZ-ORF3-OXR1-OXR2	VBI78	pTYGSade	LR with VBI79
pTYGSniaD-OME1	VBI74	TYGSniaD	LR with VBI75
pE-YA-OME1	VBI75	pE-YA	OME1: 1053 + 1054
pE-YA-CYP2	VBI100	pE-YA	CYP2: x + x
pE-YA-OXR2	VBI79	pE-YA	OME1: 1051 + 1052
pTYGS- <i>hyg</i> ^R -BC2	VBI99	pTYGSade	BC2: 1063 + 1064; <i>hph</i> : 1065 + 1066; Patch_ <i>adh</i> : 89 + 90

Table 9.4.4: Oligonucleotides (ON) used for construction of expression vectors for *E. coli*

Name	ID	vector backbone	ON for construction in <i>S. cerevisiae</i>
pET28a(+)- <i>ORF3</i>	VBI19	pET28a	<i>ORF3</i> : 733 + 734
pET28a(+)- <i>ORFZ</i>	VBI20	pET28a	<i>ORFZ</i> : 731 + 732
pCA528- <i>ORFZ</i>	VBI109	pCA528	<i>ORFZ</i> : 1057 + 1058
pCA528- <i>ORF3_{opt}</i>	VBI110	pCA528	<i>ORF3_{opt}</i> : 1055 + 1056

Table 9.4.5: Oligonucleotides (ON) used for construction and amplification of pyrichalasin KO fragments

Name	ID	ON for amplification of KO fragments	vector backbone	ON for constructon in <i>S. cerevisiae</i>
pE-YA- <i>pyiE</i> -KO	VBI22	5' fragment: 773 + 22 3' fragment: 23 + 774	pE-YA	5' <i>pyiE</i> : 735 + 736; <i>hygR</i> : 55 + 56; 3' <i>pyiE</i> : 737 + 738
pE-YA- <i>pyiF</i> -KO	VBI13	5' fragment: 730 + 22 3' fragment: 23 + 729	PE-YA	5' <i>pyiF</i> : 663 + 664; <i>hygR</i> : 55 + 56; 3' <i>pyiF</i> : 665 + 666

Table 9.4.6: Oligonucleotides (ON) used for construction of complementation vectors (*M. grisea*)

Name	ID	vector backbone	ON for construction in <i>S. cerevisiae</i>
pE-YA- <i>pyiF</i>	VBI41	pE-YA	<i>pyiF</i> : 845 + 846
pTYGSbar- <i>pyiF</i>	VBI43	pTYGSbar	LR with VBI41
pE-YA- <i>ORF3</i>	VBI35	pE-YA	<i>ORF3</i> : 806 + 807
pTYGSbar- <i>ORF3</i>	VBI38	pTYGSbar	LR with VBI35
pE-YA- <i>cssF</i>	VBI69	pE-YA	<i>cssF</i> : 913 + 912
pE-YA- <i>cssF_{intron}</i>	VBI56	pE-YA	<i>cssF</i> exon 1: 193 + 942, <i>cssF</i> exon 2: 941 + 912
pE-YA- <i>CHGG_01241</i>	VBI57	pE-YA	<i>CHGG_01241</i> : 914 + 915
pE-YA- <i>PoxQ</i>	VBI58	pE-YA	<i>PoxQ</i> : 916 + 917
pTYGSbar- <i>cssF</i>	VBI72	pTYGSbar	LR with VBI69
pTYGSbar- <i>cssF_{intron}</i>	VBI60	pTYGSbar	LR with VBI56
pTYGSbar- <i>CHGG_01241</i>	VBI61	pTYGSbar	LR with VBI57
pTYGSbar- <i>PoxQ</i>	VBI62	pTYGSbar	LR with VBI58
pE-YA- <i>pyiE</i>	VBI42	pE-YA	<i>pyiE</i> : 847 + 848
pTYGSbar- <i>pyiE</i>	VBI43	pTYGSbar	LR with VBI42
pTYGSbar- <i>pyiE_{S252A}</i>	VBI73	pTYGSbar	<i>pyiE</i> fragment 1: 998 + 1100, <i>pyiE</i> fragment 2: 1099 + 999
pTYGSbar- <i>P_{native}-pyiE</i>	VBI86	pTYGSbar	<i>P_{native}-pyiE</i> : 1176 + 999
pE-YA- <i>ORFZ</i>	VBI36	pE-YA	<i>ORFZ</i> : 808 + 809
pTYGSbar- <i>ORFZ</i>	VBI39	pTYGSbar	LR with VBI36

9.4.4 RNA extraction and cDNA Preparation

All utilities used in handling mycelia were washed with DEPC-treated water and autoclaved. Grown mycelia were collected by filtration under vacuum and washed with DEPC-treated water. Using mortar and pestle, mycelia were then ground under liquid nitrogen to a very fine powder. 50 mg of the powdered mycelia were processed according to instructions of RNeasy Plant Mini Kit (Qiagen). To remove genomic DNA a digestion step was carried out by the DNase (Qiagen) according to manufacturer instructions. Fresh purified total RNA was used in further steps.

For the preparation of cDNA the RevertAid Premium Transcriptase kit (Thermo Fisher Scientific) was used on total RNA (using oligo(dT) nucleotides). Manufacturers' instructions were followed.

9.4.5 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® -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 20 - 30 min in a Bio-Rad gel chamber containing 0.5 x TAE running buffer. DNA fragments were visualized by the *Molecular Imager Gel doc XR+* (Bio-Rad) system under UV light (312 nm). Digital images of the gels were created with the associated software.

9.4.6 DNA Purification from Gel or PCR

Amplified DNA was purified for sequencing or cloning using NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel) according to the manufacturer's instructions. If necessary, DNA samples were separated by gel electrophoresis prior to purification under previous describes conditions. Gels containing Roti® -Safe GelStain (Roth) were visualised by UV light and fragments of the right size were cut out of the gel with a scalpel.

9.4.7 Isolation of Plasmid DNA from *E. coli*

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

9.4.8 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. The whole amount of cells was 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 ddH₂O was used.

9.4.9 Isolation of Genomic DNA from Fungi

For the genomic DNA preparation 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 used to extract genomic DNA.

9.4.10 Cloning Procedure

Restriction enzyme digestion

Restriction enzyme digestions (single or double) of plasmids or PCR products were performed following the manufacturer's guidelines of the enzymes (New England Biolabs or Thermo Scientific). DNA samples were mixed with the enzyme or enzyme mixture and the appropriate buffer. The reaction mixture was incubated at 25 °C or 37 °C (depending on the restriction enzyme) for 1 to 12 h. Enzymes were inactivated by heating the sample at 65 °C or 80 °C for 20 min. Digested DNA or plasmid were purified using the NucleoSpin® Gel and PCR Clean-up kit (Macherey–Nagel) as described previously.

Dephosphorylation and ligation

Digested plasmids were dephosphorylated using shrimp alkaline phosphatase (SAP) to prevent self-ligation. 1 µl of enzyme was added after digestion to the reaction mixture and incubated at 37 °C for 1 h. T4 DNA ligase (New England Biolabs) and its appropriate buffer were used for ligation reaction (final volume 10 µl) following the manufacturer's instructions. Molar ratio of 1:5 plasmid to insert was used. Ligation reactions were incubated at 4 °C overnight or at room temperature for 30 min followed by heat inactivation at 65 °C.

Gateway cloning

LR-recombination between the entry clone and the destination vector was performed using Gateway LR Clonase enzyme mix II kit (Invitrogen) following the manufacturer's instructions. For transformation of competent *E. coli* T10 cells 1 - 2 µl of the reaction mixture was used.

DNA sequencing

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

9.4.11 Heterologous Protein Production and Purification

Heterologous production of ORFZ and ORF3

The *ORFZ* and the *ORF3* gene were amplified from plasmid pENTR_ORFZ and pENT_ORF3, respectively, and cloned into the expression vector pET28a(+) via *NdeI* and *XhoI* restriction sites or into pCA528 (SUMO) via *BsmBI* and *XhoI* restriction sites (pCA528 was cut with *BsaI* and *XhoI*). Each constructed plasmid was transformed into chemically competent *E. coli* BL21 (DE3) cells via

heat-shock transformation. A Single colony was picked to inoculate an overnight culture (LB medium with kanamycin) which was used to inoculate kanamycin containing 2TY medium. Cultures were grown to an OD₆₀₀ of 0.4-0.6 at 37 °C before protein production was induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The cultures were further incubated at 16 or 25 °C for 20 h to increase the solubility of the produced protein. Cells were harvested by centrifugation (5000 x g, 4 °C, 15 min).

The vector pET28-*ORF3* was also transformed into Rossetta2 and CodonPlus RP cells *via* heat-shock and cultivated and tested for the production of soluble protein.

Codon optimisation of the *ORF3* gene sequence

The whole sequence of *ORF3* was codon optimized for *E. coli* using the webserver OPTIMIZER (<http://genomes.urv.es/OPTIMIZER/Form.php>). The fully codon optimized *ORF3* sequence (*ORF3*_{opt.}) was synthesized and cloned into the plasmid pET100 by Invitrogen.

Cell Lysis and Protein Extraction

The harvested cell pellets were resuspended in 10 mL cold IMAC loading 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 7 min. The cell lysate was centrifuged at 12000 rpm for 45 min at 4 °C. Samples of the supernatant and the pellet were used for SDS PAGE. The protein of interest was purified using immobilized metal ion affinity chromatography. During the whole time the cell samples were kept on ice.

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 crude lysate was passed through a Ni-NTA column (2 mL bed volume) Firstly, the column was washed with 10 mL loading buffer. Elution was performed with a stepwise gradient of imidazole (20 - 500 mM). 2 mL fractions were collected over the course of elution and analysed using SDS-PAGE. The eluate was concentrated and the binding buffer was replaced by storage buffer (50 mM potassium phosphate buffer, pH 8.0) using amicon ultra centrifugal filter units, cut off 10000 MW, from Merck. The purified enzyme was stored at -20 °C or immediately used for enzyme assays.

9.4.12 Protein Analysis by SDS-PAGE, Bradford and ESI Q-TOF

Protein Separation Using SDS - Polyacrylamide Gel Electrophoresis

Separation and analysis of proteins was performed using 12 % polyacrylamide gels (Table 9.4.7). 15 µL of each protein sample were mixed with 5 µL SDS loading buffer and boiled at 95 °C for 5 min. 10 to 20 µL of the samples and 5 µL of Colour Prestained Protein Standard (Broad Range, 11–245 kDa, NEB) were loaded onto the gel. The gels ran at 20 - 40 mA for 50 - 90 min. Gels were incubated in 30 mL Coomassie staining solution for 2 - 3 h. Destaining was done using Coomassie bleach solution. The incubation time varied between 30 - 45 min. To increase the quality of the gels they were kept in water after destaining until they were scanned using the *Molecular Imager Gel doc XR+* (Bio-Rad) system. Digital images of the gels were created with the associated software.

Table 9.4.7: Composition of a 12 % separating gel and a 5% stacking gel.

Composition	separating gel	stacking gel
30 % Acrylamid/Bisacrylamid (Rotiphorese® Gel 30 (37,5:1) (Roth))	3.00 mL	0.535 mL
ddH ₂ O	2.45 mL	1.70 mL
1,5 M Tris-HCl, pH 8,8	1.90 mL	-
0,5 M Tris-HCl, pH 6,8	-	0.250 mL
10 % (w/v) SDS	0.075 ml	0.002 mL
10 % (w/v) APS	0.075 ml	0.002 mL
TEMED	0.0003 ml	0.0002 mL

Protein Identification by ESI Q-TOF

The purified protein (protein band was cut out of SDS-PAGE) was enzymatically digested by trypsin before the peptide fragments were further analysed by ESI Q-TOF (Dr. Jennifer Senkler work group Prof. Braun, Institute of Plant Genetics, Leibniz University of Hannover).

Protein Quantification Using Bradford Method

Protein concentration was determined using a Bradford assay. Protein samples were diluted 1:5 with water to a final volume of 100 µL, mixed with 900 µL of the Bradford reagent (Roti®Nanoquant (ROTH)) and incubated for 15 min. A standard curve was developed using a series of Bovine Serum Albumin (BSA) concentrations in a range from 0.1 mg/ml to 1.5 mg/ml. The concentration of the sample was determined photometrically at 595 nm and calculated to the standard curve.

9.4.13 Enzyme Activity Assay with ORFZ

The *in-vitro* enzyme assays were performed in a total volume of 100 μ l in 50 mM Na-Phosphate buffer (pH 7.0) with a final concentration of 2.5-5 mM of the substrate model **142**, and 0.15 -0.5 mg/ml of enzyme. The reaction was incubated at 25 °C or 30 °C for 30-90 min. The protein was precipitated with acetonitrile and removed from the reaction mixture by centrifugation prior to LCMS analysis. Boiled enzyme served as negative control.

9.4.14 Enzyme Activity Assay with *A. oryzae* Cell Free Extract

Preparation of cell free extract (CFE)

To obtain cell free extract (CFE) containing the protein of interest, a culture of *A. oryzae* expressing *ACE1* and *RAP1* was grown for 3-4 days in 100 mL DPY media (28°C, 180 rpm) before the cells were collected by filtration. As a control, one part of the cells was extracted with ethyl acetate and analysis by LCMS to confirm the formation of the alcohol **67**, indicating the presence of the protein of interest. The rest of the cells were disrupted by a mortar and pestle under liquid nitrogen and resuspended in 50 mM Na-Phosphate buffer (pH 7.0). Cell debris were removed by centrifugation, the supernatant was filter through a 0.22 μ m filter and 500 μ L (20 mg/ml) were injected into the sample loop of the ÄKTA-FPLC system.

Size exclusion chromatography (SEC)

Proteins were separated by their size using a Superdex™ 75 10/300 GL (GE) column following manufacture's introductions. 1 mL fractions were collected and tested for reductive activity.

Ion exchange chromatography (IEX)

Fraction of interest were combined, concentrated and further separated by a strong anion exchange column (QFF, HiTrap® IEX Selection Kit, GE Healthcare) which separates proteins based on their charge (potassium phosphate buffer pH 7, elution was performed with a linear salt gradient (from 0.15 M to 0.5 M NaCl). The column was used following manufacture's introductions. 1 mL fractions were collected and tested for enzyme activity.

In-vitro enzyme assays

The *in-vitro* enzyme assays were performed in a total volume of 100 μ l in 50 mM Na-Phosphate buffer (pH 7.0) with a final concentration of 2.5-5 mM of the substrate model **97**, 2.5 mM NADPH, 2.5 mM NADH and 0.5 -1 mg/ml of CFE. The reaction was incubated at 25°C or 30 °C for 30-60 min. The protein was precipitated with acetonitrile and removed from the reaction mixture by

centrifugation prior to LCMS analysis. Boiled enzyme served as negative control. A standard of alcohol **98** was used as positive control.

9.5 Chemical Analysis

9.5.1 Extraction of *A. oryzae* Cultures

Grown *A. oryzae* cultures were segregated into mycelia mass and aqueous phase by vacuum filtration. The aqueous phase was acidified with 2 M HCl to pH 3-4 and extracted twice with equal volume of ethyl acetate. The mycelia were homogenised with a blender and stirred for 1-3 h in ethyl acetate. The extract was separated from mycelia by filtration. Organic layers were dried with anhydrous MgSO₄ and the solvent removed with rotary evaporator. Extracts were dissolved in methanol to a concentration of 10 mg/ml, filtered over glass wool and analysed by LC/MS. All compounds produced by the transformants were mainly present in mycelia.

9.5.2 Extraction of *M. grisea* and *P. oryzae* Cultures

Grown *P. oryzae* or *M. grisea* cultures were segregated into mycelia mass and aqueous phase by vacuum filtration. The aqueous phase was acidified with 2 M HCL to pH 3-4 and extracted twice with an equal volume of ethyl acetate. The mycelia were homogenized with a blender and stirred for 1-3 h in ethyl acetate. The extract was separated from mycelia by filtration. Combined organic layers were dried with anhydrous MgSO₄ and the solvent removed with rotary evaporator. Extracts were dissolved in methanol to a concentration of 10 mg/ml, filtered over glass wool and analysed or purified by LCMS. All compounds produced were mainly present in the aqueous phase.

9.5.3 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 μ m, C₁₈, 100 Å, 4.6 x 100 mm) equipped with a Phenomenex Security Guard precolumn (Luna, C₅, 300 Å) and a flow rate of 1 ml/min. 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+ and ES- modes between 150 and 1000 *m/z*. A solvent gradient was run over 15 min starting at 10 % acetonitrile/ 90 % HPLC grade water (0.05 % formic acid) and ramping to 90 % acetonitrile.

Preparative LCMS was used to purify compounds from a raw extract or from a reaction mixture. This consisted of a Waters mass-directed autopurification system comprising of a Waters 2767 autosampler, Waters 2545 pump system, a Phenomenex Kinetex Axia column (5 μm , C_{18} , 100 \AA , 21.2 x 250 mm) equipped with a Phenomenex Security Guard precolumn (Luna, C_5 , 300 \AA). A solvent gradient was run over 10 min or 15 min starting at 10 % acetonitrile/ 90 % HPLC grade water (0.05 % formic acid) and ramping to 90 % acetonitrile. The flow was set to 20 ml/min and 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) in the range 210 to 600 nm and an ELSD detector (Waters 2424) together with 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 tubes were evaporated (vacuum centrifuge).

9.5.4 High Resolution Mass Spectrometry (HRMS)

High Resolution Mass Spectrometry (HRMS) was performed on a Q-ToF Premier mass spectrometer (Waters) coupled to an Acquity UPLC-domain. Positive and negative ions were measured for accurate mass analysis.

9.5.5 Nuclear Magnetic Resonance (NMR) Analysis

NMR measurements were done by the NMR department of Organic Chemistry Department, Leibniz University Hannover. ^1H and ^{13}C NMR spectra were recorded on either on a Bruker DRX 400 or 500 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).

10. Bibliography

- 1 W. B. Turner, *Fungal Metabolites*, Academic Press, London, 1971.
- 2 J. D. Bergstrom, M. M. Kurtz, D. J. Rew, A. M. Amend, J. D. Karkas, R. G. Bostedor, V. S. Bansal, C. Dufresne, F. L. Van Middlesworth and O. D. Hensens, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 80–84.
- 3 A. W. Alberts, J. Chen, G. Kuron, V. Hunt, J. Huff, C. Hoffman, J. Rothrock, M. Lopez, H. Joshua, E. Harris, A. Patchett, R. Monaghan, S. Currie, E. Stapley, G. Albers-Schonberg, O. Hensens, J. Hirshfield, K. Hoogsteen, J. Liesch and J. Springer, *Proc. Natl. Acad. Sci. U. S. A.*, 1980, **77**, 3957–3961.
- 4 K. Yu, B. Ren, J. Wei, C. Chen, J. Sun, F. Song, H. Dai and L. Zhang, *Mar. Drugs*, 2010, **8**, 2744–2754.
- 5 E. S. and E. S. J. Beck, S. Ripka, A. Siegner, *Eur. J. Biochem.*, 1990, **498**, 487–498.
- 6 W. C. A. Gelderblom, W. F. O. Marasas, P. S. Steyn, P. G. Thiel, K. J. Van Der Merwe, R. Van Rooyen, P.H. Vlegaar and P. L. Wessels, *J. Chem. Soc. Chem. Commun.*, 1984, 122–124.
- 7 A. G. McInnes, Donald G. Smith and C.-K. Wat, *J. Chem. Soc.*, 1974, **158**, 145–146.
- 8 N. P. Keller and T. M. Hohn, *Fungal Genet. Biol.*, 1997, **21**, 17–29.
- 9 H. C. Lo, R. Entwistle, C. J. Guo, M. Ahuja, E. Szewczyk, J. H. Hung, Y. M. Chiang, B. R. Oakley and C. C. C. Wang, *J. Am. Chem. Soc.*, 2012, **134**, 4709–4720.
- 10 H. Sierra, M. Cordova, C. S. J. Chen and M. Rajadhyaksha, *J. Invest. Dermatol.*, 2015, **135**, 612–615.
- 11 S. Bergmann, J. Schümann, K. Scherlach, C. Lange, A. A. Brakhage and C. Hertweck, *Nat. Chem. Biol.*, 2007, **3**, 213.
- 12 W. C. Yaegashi J, Oakley BR, *J Ind Microbiol Biotechnol*, 2015, **41**, 433–442.
- 13 J. Ma, *Protein Cell*, 2011, **2**, 879–888.
- 14 L. D. S. Carvalho, L. K. Teixeira, N. Carrossini, A. T. N. Caldeira, K. M. Ansel, A. Rao and J. P. B. Viola, *Cell Cycle*, 2007, **6**, 1789–1795.
- 15 H. U. Van Der Heul, B. L. Bilyk, K. J. McDowall, R. F. Seipke and G. P. Van Wezel, *Nat. Prod. Rep.*, 2018, **35**, 575–604.
- 16 R. D. Kornberg, *Science*, 1974, **184**, 868–871.
- 17 S. L. Berger, *Nature*, 2007, **447**, 407.
- 18 T. Kouzarides, *Cell*, 2007, **128**, 693–705.
- 19 A. J. Birch and F. W. Donovan, *Aust. J. Chem.*, 1953, **6**, 360–368.
- 20 A. J. Birch, R. A. Massy-Westropp and C. J. Moye, *Aust. J. Chem.*, 1955, **8**, 539–544.
- 21 J. Staunton and K. J. Weissman, *Nat. Prod. Rep.*, 2001, **18**, 380–416.
- 22 D. K. Liscombe, G. V. Louie and J. P. Noel, *Nat. Prod. Rep.*, 2012, **29**, 1238–1250.
- 23 C. Hertweck, *Angew. Chemie - Int. Ed.*, 2009, **48**, 4688–4716.
- 24 J. Staunton and B. Wilkinson, *Chem. Rev.*, 2002, **97**, 2611–2630.
- 25 R. A. Cacho, J. Thuss, W. Xu, R. Sanichar, Z. Gao, A. Nguyen, J. C. Vederas and Y. Tang, *J. Am. Chem. Soc.*, 2015, **137**, 15688–15691.
- 26 L. M. Halo, J. W. Marshall, A. A. Yakasai, Z. Song, C. P. Butts, M. P. Crump, M. Heneghan, A. M. Bailey, T. J. Simpson, C. M. Lazarus and R. J. Cox, *ChemBioChem*, 2008, **9**, 585–594.
- 27 B. Bonsch, V. Belt, C. Bartel, N. Duensing, M. Koziol, C. M. Lazarus, A. M. Bailey, T. J. Simpson and R. J. Cox, *Chem. Commun.*, 2016, **52**, 6777–6780.
- 28 J. M. Crawford, P. M. Thomas, J. R. Scheerer, A. L. Vagstad, N. L. Kelleher and C. A. Townsend, *Science*, 2008, **320**, 243–246.
- 29 J. M. Crawford, B. C. R. Dancy, E. A. Hill, D. W. Udvary and C. A. Townsend, *Proc. Natl. Acad. Sci.*, 2006, **103**, 16728–16733.
- 30 R. J. Cox, *Org. Biomol. Chem.*, 2007, **5**, 2010–26.
- 31 J. F. Barajas, G. Shakya, G. Moreno, H. Rivera, D. R. Jackson, C. L. Topper, A. L. Vagstad, J. J. La Clair, C. A. Townsend, M. D. Burkart and S.-C. Tsai, *Proc. Natl. Acad. Sci.*, 2017, **114**, E4142–E4148.
- 32 Y. Li, W. Xu and Y. Tang, *J. Biol. Chem.*, 2010, **285**, 22764–22773.
- 33 H. Sun, C. L. Ho, F. Ding, I. Soehano, X. W. Liu and Z. X. Liang, *J. Am. Chem. Soc.*, 2012, **134**, 11924–11927.
- 34 I. Soehano, L. Yang, F. Ding, H. Sun, Z. J. Low, X. Liu and Z.-X. Liang, *Org. Biomol. Chem.*, 2014, **12**, 8542–8549.
- 35 R. J. Cox, *Org. Biomol. Chem.*, 2007, **5**, 2010–2026.
- 36 Y.-H. Chooi and Y. Tang, *J Org Chem*, 2012, **22**, 9933–9953.
- 37 O. McAuliffe, R. P. Ross and C. Hill, *Chem. Rev.*, 2001, **105**, 633–683.
- 38 M. A. Marahiel, T. Stachelhaus and H. D. Mootz, *Chem. Rev.*, 1997, **97**, 2651–2674.
- 39 E. A. Felnagle, E. E. Jackson, Y. A. Chan, A. M. Podevels, D. Berti, M. D. McMahon and M. G. Thomas, *Mol Pharm.*, 2011, **5**, 191–211.
- 40 G. L. Challis and J. H. Naismith, *Curr Opin Struct Biol.*, 2012, **14**, 748–756.
- 41 A. L. Demain and R. P. Elander, *Antonie Van Leeuwenhoek*, 1999, **75**, 5–19.
- 42 G. Banko, A. L. Demain and S. Wolfe, *J. Am. Chem. Soc.*, 1987, **109**, 2858–2860.
- 43 H. van Liempt, H. von Döhren and H. Kleinkauf, *J. Biol. Chem.*, 1989, **264**, 3680–3684.
- 44 M. Juguët, S. Lautru, F. X. Francou, Š. Nezbedová, P. Leblond, M. Gondry and J. L. Pernodet, *Chem. Biol.*, 2009, **16**, 421–431.

- 45 J. Collemare, A. Billard, H. U. Böhnert and M. H. Lebrun, *Mycol. Res.*, 2008, **112**, 207–215.
- 46 X. Liu and C. T. Walsh, *Biochemistry*, 2009, **48**, 8746–8757.
- 47 P. K. Chang, B. W. Horn and J. W. Dorner, *Fungal Genet. Biol.*, 2009, **46**, 176–182.
- 48 J. W. Sims, J. P. Fillmore, D. D. Warner and E. W. Schmidt, *Chem. Commun.*, 2005, 186–188.
- 49 K. L. Eley, L. M. Halo, Z. Song, H. Powles, R. J. Cox, A. M. Bailey, C. M. Lazarus and T. J. Simpson, *ChemBioChem*, 2007, **8**, 289–297.
- 50 Z. Song, R. J. Cox, C. M. Lazarus and T. J. Simpson, *ChemBioChem*, 2004, **5**, 1196–1203.
- 51 K. Qiao, Y. H. Chooi and Y. Tang, *Metab. Eng.*, 2011, **13**, 723–732.
- 52 M. Laursen, M. Bublitz, K. Moncoq, C. Olesen, J. V. Møller, H. S. Young, P. Nissen and J. P. Morth, *J. Biol. Chem.*, 2009, **284**, 13513–13518.
- 53 K. Moncoq, C. A. Trieber and H. S. Young, *J. Biol. Chem.*, 2007, **282**, 9748–9757.
- 54 S. B. Singh, D. L. Zink, M. A. Goetz, A. W. Dombrowski, J. D. Polishook and D. J. Hazuda, *Tetrahedron Lett.*, 1998, **39**, 2243–2246.
- 55 R. F. Vesonder, L. W. Tjarks, W. K. Rohwedder, H. R. Burgmeister and J. A. Laugal, *J. Antibiot.*, 1979, **32**, 759–761.
- 56 K. L. Kavanagh, H. Jörnvall, B. Persson and U. Oppermann, *Cell. Mol. Life Sci.*, 2008, **65**, 3895–3906.
- 57 S. Maiya, A. Grundmann, X. Li, S. M. Li and G. Turner, *ChemBioChem*, 2007, **8**, 1736–1743.
- 58 L. M. Halo, J. W. Marshall, A. A. Yakasai, Z. Song, C. P. Butts, M. P. Crump, M. Heneghan, A. M. Bailey, T. J. Simpson, C. M. Lazarus and R. J. Cox, *ChemBioChem*, 2008, **9**, 585–594.
- 59 J. W. Sims and E. W. Schmidt, *J. Am. Chem. Soc.*, 2008, **130**, 11149–11155.
- 60 C. Filling, K. D. Berndt, J. Benach, S. Knapp, T. Prozorovski, E. Nordling, R. Ladenstein, H. Jörnvall and U. Oppermann, *J. Biol. Chem.*, 2002, **277**, 25677–25684.
- 61 R. Ladenstein, J. O. Winberg and J. Benach, *Cell. Mol. Life Sci.*, 2008, **65**, 3918–3935.
- 62 D. C. Aldridge, J. J. Armstrong, R. N. Speake and W. B. Turner, *J. Chem. Soc.*, 1967, **1667**, 1667–1676.
- 63 W. Rothweiler and C. Tamm, *Experientia*, 1966, **22**, 750–752.
- 64 S. B. Carter, *Nature*, 1967, **213**, 261–264.
- 65 S. MacLean-Fletcher and T. D. Pollard, *Cell*, 1980, **20**, 329–341.
- 66 J. F. Casella, M. D. Flanagan and S. Lin, *Nature*, 1981, **293**, 302–305.
- 67 A. Ghosh, J. Maniloff and D. A. Gerling, *Cell*, 1978, **13**, 57–64.
- 68 A. L. Rampal, H. B. Pinkofsky and C. Y. Jung, *Biochemistry*, 1980, **19**, 679–683.
- 69 J. F. Crivello and C. R. Jefcoate, *J. Biol. Chem.*, 1980, **255**, 8144–8151.
- 70 G. Wei, D. Tan, C. Chen, Q. Tong, X.-N. Li, J. Huang, J. Liu, Y. Xue, J. Wang, Z. Luo, H. Zhu and Y. Zhang, *Sci. Rep.*, 2017, **7**, 42434.
- 71 M. Binder, J.-R. Kiechel and C. Tamm, *Helv. Chim. Acta*, 1970, **53**, 1797–1812.
- 72 A. Probst and C. Tamm, *Helv. Chim. Acta*, 1981, **64**, 18–19.
- 73 H. Oikawa, Y. Murakami and a Ichihara, *J. Chem. Soc. Trans. 1*, 1992, 2955–2959.
- 74 J. Schümann and C. Hertweck, *J. Am. Chem. Soc.*, 2007, **129**, 9564–9565.
- 75 L. M. Halo, A. A. Yakasai, Z. Song, R. J. Cox, T. J. Simpson, M. N. Heneghan, K. Williams, A. M. Bailey and C. M. Lazarus, *J. Am. Chem. Soc.*, 2008, **130**, 17988–17996.
- 76 K. Ishiuchi, T. Nakazawa, F. Yagishita, T. Mino, H. Noguchi, K. Hotta and K. Watanabe, *J. Am. Chem. Soc.*, 2013, **135**, 7371–7377.
- 77 D. M. Sato, D. F. Yagishita, P. D. T. Mino, D. U. Nahoko, D. A. Patel, D. Y.-H. Chooi, D. Y. Goda, D. W. Xu, P. D. H. Noguchi, T. Yamamoto, P. D. K. Hotta, P. D. Kendall, N. Houk, P. D. Y. Tang, And and P. D. K. Watanabe, *ChemBioChem*, 2016, **16**, 2294–2298.
- 78 E. Skellam, *Nat. Prod. Rep.*, 2017, **34**, 1252–1263.
- 79 E. C. Oerke, *Agric. Syst.*, 1996, **51**, 367–369.
- 80 S. L. Tucker and N. J. Talbot, *Annu. Rev. Phytopathol.*, 2001, **39**, 385–417.
- 81 R. J. Howard and B. Valent, *Annu. Rev. Microbiol.*, 1996, **50**, 491–512.
- 82 R. A. Wilson and N. J. Talbot, *Nat. Rev. Microbiol.*, 2009, **7**, 185–195.
- 83 R. Berruyer, H. Adreit, J. Milazzo, S. Gaillard, A. Berger, W. Dioh, M.-H. Lebrun and D. Tharreau, *Theor. Appl. Genet.*, 2003, **107**, 1139–47.
- 84 H. H. Flor, *Annu. Rev. Phytopathol.*, 1971, **9**, 275–296.
- 85 W. Knogge, in *Kempken F. (eds) Agricultural Applications. The Mycota XI (Springer-Verlag)*, 2002, pp. 289–310.
- 86 Böhnert, H. U., I. Fudal, W. Dioh, D. Tharreau, J. L. Notteghem and M. H. Lebrun, *Plant Cell*, 2004, **16**, 2499–2513.
- 87 W. Dioh, D. Tharreau, J. L. Notteghem, M. Orbach and M. Lebrun, *Mol Plant Microbe Interact.*, 2000, **13**, 217–227.
- 88 J. Collemare, M. Pianfetti, M. Vial, A. Houille, D. Morin, L. Camborde, M. Gagey, M. H. Lebrun and H. Bohnert, *New Phytol.*, 2008, **179**, 196–208.
- 89 N. Khaldi, J. Collemare, M.-H. Lebrun and K. H. Wolfe, *Genome Biol.*, 2008, **9**, R18.
- 90 M. L. Nielsen, T. Isbrandt, L. M. Petersen, U. H. Mortensen, M. R. Andersen, J. B. Hoof and T. O. Larsen, *PLoS One*, 2016, **11**, 1–18.
- 91 R. Fujii, A. Minami, K. Gomi and H. Oikawa, *Tetrahedron Lett.*, 2013, **54**, 2999–3002.
- 92 F. Villalba, J. Collemare, P. Landraud, K. Lambou, V. Brozek, B. Cirer, D. Morin, C. Bruel, R. Beffa and M. H.

- Lebrun, *Fungal Genet. Biol.*, 2008, **45**, 68–75.
- 93 J. Kennedy, K. Auclair, S. G. Kendrew, C. Park, J. C. Vederas and C. R. Hutchinson, *Science*, 1999, **284**, 1368–1372.
- 94 L. M. Halo, M. N. Heneghan, A. A. Yakasai, Z. Song, K. Williams, A. M. Bailey, R. J. Cox, C. M. Lazarus and T. J. Simpson, *J. Am. Chem. Soc.*, 2008, **130**, 17988–17996.
- 95 Z. Wasil, K. A. K. Pahirulzaman, C. Butts, T. J. Simpson, C. M. Lazarus and R. J. Cox, *Chem. Sci.*, 2013, **4**, 3845–3856.
- 96 Z. Song, W. Bakeer, J. Marshall, A. Yakasai, R. Mohd-Khalid, J. Collemare, E. Skellam, D. Tharreau, M.-H. Lebrun, C. M. Lazarus, A. M. Bailey, T. J. Simpson and R. J. Cox, *Chem. Sci.*, 2015, **6**, 4837–4845.
- 97 H. C. Eisenman and A. Casadevall, *Appl. Microbiol. Biotechnol.*, 2012, **93**, 931–940.
- 98 E. S. Jacobson, *Clin. Microbiol. Rev.*, 2000, **13**, 708–717.
- 99 J. C. De Jong, B. J. McCormack, N. Smirnov and N. J. Talbot, *Nature*, 1997, **389**, 244–245.
- 100 R. J. Howard and M. A. Ferrari, *Exp. Mycol.*, 1989, **13**, 403–418.
- 101 R. J. Howard and B. Valent, *Annu. Rev. Microbiol.*, 1996, **50**, 491–512.
- 102 H. D. Sisler, *Crop Prot.*, 1986, **5**, 306–313.
- 103 C. P. Woloshuk, H. D. Sisler and E. L. Vigil, *Physiol. Plant Pathol.*, 1983, **22**, 245–259.
- 104 R. D. Stipanovic and A. A. Bell, *J. Org. Chem.*, 1976, **41**, 2468–2469.
- 105 A. A. Bell, J. E. Puhalla, W. J. Tolmsoff and R. D. Stipanovic, *Can. J. Microbiol.*, 1976, **22**, 787–799.
- 106 M. Nukina, *Agric. Biol. Chem.*, 1987, **51**, 2625–2628.
- 107 T. Hirose, Y. Izawa, K. Koyama, S. Natori, K. Iida, I. Yahara, S. Shimaoka and K. Maruyama, *Chem. Pharm. Bull.*, 1990, **38**, 971–974.
- 108 C. Wang, V. Hantke, R. J. Cox and E. Skellam, *Org. Lett.*, 2019, **21**, 4163–4167.
- 109 I. V. Grigoriev, R. Nikitin, S. Haridas, A. Kuo, R. Ohm, R. Otilar, R. Riley, A. Salamov, X. Zhao, F. Korzeniewski, T. Smirnova, H. Nordberg, I. Dubchak and I. Shabalov, *Nucleic Acids Res.*, 2014, **42**, 699–704.
- 110 Y. Hu, D. Dietrich, W. Xu, A. Patel, J. A. J. Thuss, J. Wang, W. Yin, K. Qiao, K. N. Houk, J. C. Vederas and Y. Tang, *Nat Chem Biol.*, 2015, **10**, 552–554.
- 111 U. Kück and B. Hoff, *Appl. Microbiol. Biotechnol.*, 2010, **86**, 51–62.
- 112 D. Müller and U. Stahl, in *Progress in Botany: Genetics Physiology Systematics Ecology*, eds. K. Esser, U. Lüttge, W. Beyschlag and J. Murata, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 31–49.
- 113 L. M. Halo, M. N. Heneghan, A. A. Yakasai, Z. Song, K. Williams, A. M. Bailey, R. J. Cox, C. M. Lazarus and T. J. Simpson, *J. Am. Chem. Soc.*, 2008, **130**, 17988–17996.
- 114 C. Fairhead, B. Llorente, F. Denis, S. O. L. E. R. Maria, B. Dujon, U. De Gknitique, L. Ufr, M. Curie and C. N. R. S. Ura, *Yeast Funct. Anal. Rep.*, 1996, **12**, 1439–1457.
- 115 M. L. Nielsen, L. Albertsen, G. Lettier, J. B. Nielsen and U. H. Mortensen, *Fungal Genet. Biol.*, 2006, **43**, 54–64.
- 116 R. J. Watson, S. Burchat and J. Bosley, *Fungal Genet. Biol.*, 2008, **45**, 1348–1363.
- 117 T. J. Carver, K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell and J. Parkhill, *Bioinformatics*, 2005, **21**, 3422–3423.
- 118 P. Rice, I. Longden and A. Bleasby, *Trends Genet.*, 2000, **16**, 276–277.
- 119 F. Sievers and D. G. Higgins, *Methods Mol. Biol.*, 2014, **1079**, 105–116.
- 120 Z. Wasil, K. A. K. Pahirulzaman, C. Butts, T. J. Simpson, C. M. Lazarus and R. J. Cox, *Chem. Sci.*, 2013, **4**, 3845–3856.
- 121 K. Qiao, Y.-H. Chooi and Y. Tang, *Metab. Eng.*, 2011, **13**, 723–732.
- 122 H. He, H. Y. Yang, R. Bigelis, E. H. Solum, M. Greenstein and G. T. Carter, *Tetrahedron Lett.*, 2002, **43**, 1633–1636.
- 123 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.
- 124 R. Schor, C. Schotte, D. Wibberg, J. Kalinowski and R. J. Cox, *Nat. Commun.*, 2018, **9**, 1963.
- 125 K. A. Pahirulzaman, K. Williams and C. M. Lazarus, in *Methods Enzymol.*, 2012, vol. 512, pp. 241–260.
- 126 N. Gaitatzis, B. Kunze and R. Müller, *Proc Natl Acad Sci U S A*, 2001, **98**, 11136–11141.
- 127 Y. Li, K. J. Weissman and R. Müller, *J. Am. Chem. Soc.*, 2008, **130**, 7554–7555.
- 128 E. M. Niehaus, K. Kleigrew, P. Wiemann, L. Studt, C. M. K. Sieber, L. R. Connolly, M. Freitag, U. Güldener, B. Tudzynski and H. U. Humpf, *Chem. Biol.*, 2013, **20**, 1055–1066.
- 129 Y. He and R. J. Cox, *Chem. Sci.*, 2016, **7**, 2119–2127.
- 130 R. Nofiani, K. de Mattos-Shiple, K. E. Lebe, L. C. Han, Z. Iqbal, A. M. Bailey, C. L. Willis, T. J. Simpson and R. J. Cox, *Nat. Commun.*, 2018, **9**, 3940.
- 131 P. Minarik, N. Tomaskova, M. Kollarova and M. Antalík, *Gen. Physiol. Biophys.*, 2002, **21**, 257–265.
- 132 J. S. Deetz, C. A. Luehr and B. L. Vallee, *Biochemistry*, 1984, **23**, 6822–6828.
- 133 H. B. Bode, B. Bethe, R. Höfs and A. Zeeck, *ChemBioChem*, 2002, **3**, 619–627.
- 134 O. E. Christian, J. Compton, K. R. Christian, S. L. Mooberry, F. A. Valeriote and P. Crews, *J. Nat. Prod.*, 2005, **68**, 1592–1597.
- 135 W. J. Moree, J. Y. Yang, X. Zhao, W.-T. Liu, M. Aparicio, L. Atencio, J. Ballesteros, J. Sanchez, R. G. Gavilan, M. Gutierrez and P. C. Dorrestein, *J. Chem. Ecol.*, 2013, **39**, 1045–1054.
- 136 S. Bergmann, J. Schümann, K. Scherlach, C. Lange, A. A. Brakhage and C. Hertweck, *Nat. Chem. Biol.*, 2007, **3**, 213–217.

- 137 Y. Tsunematsu, S. Ichinoseki, T. Nakazawa, N. Ishikawa, H. Noguchi, K. Hotta and K. Watanabe, *J. Antibiot. (Tokyo)*, 2012, **65**, 377–380.
- 138 E. Batlle, E. Sancho, C. Francí, D. Domínguez, M. Monfar, J. Baulida and A. G. De Herreros, *Nat. Cell Biol.*, 2000, **2**, 84–89.
- 139 A. J. Berk and M. C. Schmidt, *Genes Dev.*, 1990, **4**, 151–155.
- 140 W. I. M. M. Bakeer, *Thesis Univ. Bristol, Sch. Chem.*, 2011.
- 141 G. Zhang, V. Gurtu and S. R. Kain, *Biochem. Biophys. Res. Commun.*, 1996, **227**, 707–711.
- 142 K. Tsuchiya, S. Tada, K. Gomi, K. Kitamoto, C. Kumagai and G. Tamura, *Biosci. Biotechnol. Biochem.*, 1992, **56**, 1849–1853.
- 143 G. Schlingmann, S. Luckman, L. A. McDonald and F. E. Koehn, *Rec. Nat. Prod.*, 2011, **5**, 252–260.
- 144 D. C. Allport and J. D. Bu'Lock, *J. Chem. Soc.*, 1956, 4090–4093.
- 145 N. N. Stepanichenko, L. N. Ten, A. A. Tyshchenko, V. M. Shevtsova, S. Z. Mukhamedzhanov and A. G. Kas'yanenko, *Chem. Nat. Compd.*, 1985, **21**, 455–460.
- 146 H. F. Tsai, Y. C. Chang, R. G. Washburn, M. H. Wheeler and K. J. Kwon-Chung, *J. Bacteriol.*, 1998, **180**, 3031–3038.
- 147 Y. Takano, Y. Kubo, K. Shimizu, K. Mise, T. Okuno and I. Furusawa, *Mgg Mol. Gen. Genet.*, 1995, **249**, 162–167.
- 148 A. L. Vagstad, E. A. Hill, J. W. Labonte and C. A. Townsend, *Chem. Biol.*, 2012, **19**, 1525–1534.
- 149 M. H. Wheeler, *Trans. Br. Mycol. Soc.*, 1983, **81**, 29–36.
- 150 A. A. Bell and M. H. Wheeler, *Annu. Rev. Phytopathol.*, 1986, **24**, 411–451.
- 151 M. J. Butler and A. W. Day, *Can. J. Microbiol.*, 1998, **44**, 1115–1136.
- 152 W. A. Edens, T. Q. Goins, D. Dooley and J. M. Henson, *Appl. Environ. Microbiol.*, 1999, **65**, 3071–3074.
- 153 I. Fudal, J. Collemare, H. U. Böhnert, D. Melayah and M. H. Lebrun, *Eukaryot. Cell*, 2007, **6**, 546–554.
- 154 J. Collemare and M. F. Seidl, *FEMS Microbiol. Rev.*, 2019.
- 155 E. K. Shwab, W. B. Jin, M. Tribus, J. Galehr, S. Graessle and N. P. Keller, *Eukaryot. Cell*, 2007, **6**, 1656–1664.
- 156 L. V. Roze, A. E. Arthur, S. Y. Hong, A. Chanda and J. E. Linz, *Mol. Microbiol.*, 2007, **66**, 713–726.
- 157 L. Studt, F. J. Schmidt, L. Jahn, C. M. K. Sieber, L. R. Connolly, E. M. Niehaus, M. Freitag, H. U. Humpf and B. Tudzynski, *Appl. Environ. Microbiol.*, 2013, **79**, 7719–7734.
- 158 E. M. Niehaus, K. Kleigrew, P. Wiemann, L. Studt, C. M. K. Sieber, L. R. Connolly, M. Freitag, U. Güldener, B. Tudzynski and H. U. Humpf, *Chem. Biol.*, 2013, **20**, 1055–1066.
- 159 Y. Reyes-Dominguez, J. W. Bok, H. Berger, E. K. Shwab, A. Basheer, A. Gallmetzer, C. Scazzocchio, N. Keller and J. Strauss, *Mol. Microbiol.*, 2010, **76**, 1376–1386.
- 160 A. Andersson, D. Jordan, G. Schneider and Y. Lindqvist, *Structure*, 1996, **4**, 1161–1170.
- 161 M. Nakasako, T. Motoyama, Y. Kurahashi and I. Yamaguchi, *Biochemistry*, 1998, **37**, 9931–9939.
- 162 D. I. Liao, G. S. Basarab, G. S. Gatenby, B. Valent and D. B. Jordan, *Structure*, 2001, **9**, 19–27.
- 163 T. Lanišnik Rižner and M. H. Wheeler, *Can. J. Microbiol.*, 2003, **49**, 110–119.
- 164 M. H. Wheeler and R. D. Stipanovic, *Arch. Microbiol.*, 1985, **142**, 234–241.
- 165 H. F. Tsai, M. H. Wheeler, Y. C. Chang and K. J. Kwon-Chung, *J. Bacteriol.*, 1999, **181**, 6469–6477.
- 166 S. Bergmann, A. N. Funk, K. Scherlach, V. Schroeckh, E. Shelest, U. Horn, C. Hertweck and A. A. Brakhage, *Appl. Environ. Microbiol.*, 2010, **76**, 8143–8149.
- 167 R. J. Howard, M. A. Ferrari, D. H. Roach and N. P. Money, *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 11281–11284.
- 168 D. L. Ollis, E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman and J. Schrag, *Protein Eng.*, 1992, **5**, 197–211.
- 169 A. G. Murzin, S. E. Brenner, T. Hubbard and C. Chothia, *J. Mol. Biol.*, 1995, **247**, 536–540.
- 170 C. A. Orengo, A. D. Michie, S. Jones, D. T. Jones, M. B. Swindells and J. M. Thornton, *Structure*, 1997, **5**, 1093–1108.
- 171 T. Hotelier, *Nucleic Acids Res.*, 2003, **32**, 145–147.
- 172 M. Nardini and B. W. Dijkstra, *Curr. Opin. Struct. Biol.*, 1999, **9**, 732–737.
- 173 C. A. Orengo, A. M. Martin, G. Hutchinson, S. Jones, D. T. Jones, A. D. Michie, M. B. Swindells and J. M. Thornton, *Acta Crystallogr.*, 1998, **54**, 1155–1167.
- 174 X. Cousin, T. Hotelier, K. Giles, J. P. Toutant and A. Chatonnet, *Nucleic Acids Res.*, 1998, **26**, 226–228.
- 175 U. G. Wagner, M. Hasslacher, H. Griengl, H. Schwab and C. Kratky, *Cell Press*, 1996, **4**, 811–822.
- 176 A. Rauwerdink and R. J. Kazlauskas, *ACS Catal.*, 2015, **5**, 6153–6176.
- 177 M. Jiang, X. Chen, X. H. Wu, M. Chen, Y. D. Wu and Z. Guo, *Biochemistry*, 2009, **48**, 6921–6931.
- 178 F. G. Bordwell, *Acc. Chem. Res.*, 1972, **5**, 374–381.
- 179 F. M. Bickelhaupt, G. J. Buisman, L. J. de Koning, N. M. M. Nibbering and E. J. Baerends, *J. Am. Chem. Soc.*, 1995, **117**, 9889–9899.
- 180 S. Gronert and S. R. Kass, *J. Org. Chem.*, 1997, **62**, 7991–8000.
- 181 M. Sato, J. E. Dander, C. Sato, Y. Hung, S.-S. Gao, M.-C. Tang, L. Hang, J. M. Winter, N. K. Garg, K. Watanabe and Y. Tang, *J. Am. Chem. Soc.*, 2017, **139**, 5317–5320.
- 182 L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass and M. J. E. Sternberg, *Nat. Protoc.*, 2015, **10**, 845–858.
- 183 P. Sachelar, E. Schiltz, G. L. Igloi and R. Brandsch, *J. Bacteriol.*, 2005, **187**, 8516–8519.
- 184 C. Schleberger, P. Sachelar, R. Brandsch and G. E. Schulz, *J. Mol. Biol.*, 2007, **367**, 409–418.
- 185 J. S. Seeler, O. Bischof, K. Nacerddine and A. Dejean, *Curr. Top. Microbiol. Immunol.*, 2007, **313**, 49–71.
- 186 I. Meinecke, A. Cinski, A. Baier, M. A. Peters, B. Dankbar, A. Wille, A. Drynda, H. Mendoza, R. E. Gay, R. T. Hay,

- B. Ink, S. Gay and T. Pap, *Proc. Natl. Acad. Sci. USA*, 2007, **104**, 5073–5078.
- 187 M. P. Malakhov, M. R. Mattern, O. A. Malakhova, M. Drinker, S. D. Weeks and T. R. Butt, *J. Struct. Funct. Genomics*, 2004, **5**, 75–86.
- 188 R. J. Peroutka III, S. J. Orcutt, J. E. Strickler and T. R. Butt, in *Heterologous Gene Expression in E.coli: Methods and Protocols*, eds. T. C. Evans Jr. and M.-Q. Xu, Humana Press, Totowa, NJ, 2011, pp. 15–30.
- 189 C. Andréasson, J. Fiaux, H. Rampelt, M. P. Mayer and B. Bukau, *J. Biol. Chem.*, 2008, **283**, 8877–8884.
- 190 B. M. Dunn and S. Hung, *Biochim. Biophys. Acta*, 2000, **1477**, 231–240.
- 191 P. A. Storm and C. A. Townsend, *Chem. Commun.*, 2017, **54**, 50–53.
- 192 A. L. Vagstad, S. B. Bumpus, K. Belecki, N. L. Kelleher and C. A. Townsend, *J. Am. Chem. Soc.*, 2012, **134**, 6865–6877.
- 193 M. Ahuja, Y. M. Chiang, S. L. Chang, M. B. Praseuth, R. Entwistle, J. F. Sanchez, H. C. Lo, H. H. Yeh, B. R. Oakley and C. C. C. Wang, *J. Am. Chem. Soc.*, 2012, **134**, 8212–8221.
- 194 B. Balakrishnan, R. Chandran, S.-H. Park and H.-J. Kwon, *J. Microbiol. Biotechnol.*, 2015, **25**, 1648–1652.
- 195 W. Xu, Y. H. Chooi, J. W. Choi, S. Li, J. C. Vederas, N. A. Da Silva and Y. Tang, *Angew. Chemie (Int. ed.)*, 2013, **52**, 6472–6475.
- 196 R. Hershey, E. Yeger-Lotem and H. Margalit, *Trends Genet.*, 2005, **21**, 138–142.
- 197 B. R. Beno, S. Wilsey and K. N. Houk, *J. Am. Chem. Soc.*, 1999, **121**, 4816–4826.
- 198 R. D. Bach, J. J. W. McDouall, H. B. Schlegel and G. J. Wolber, *J. Org. Chem.*, 1989, **54**, 2931–2935.
- 199 K. N. Houk, *J. Am. Chem. Soc.*, 1973, **95**, 4092–4094.
- 200 K. N. Houk and R. W. Strozier, *J. Am. Chem. Soc.*, 1973, **95**, 4094–4096.
- 201 J. Sauer and R. Sustmann, *Angew. Chemie Int. Ed. English*, 1980, **19**, 779–807.
- 202 Z. Zhang, C. S. Jamieson, Y.-L. Zhao, D. Li, M. Ohashi, K. N. Houk and Y. Tang, *J. Am. Chem. Soc.*, 2019, **141**, 5659–5663.
- 203 B. L. Oliveira, Z. Guo and G. J. L. Bernardes, *Chem. Soc. Rev.*, 2017, **46**, 4895–4950.
- 204 H. Oikawa, K. Katayama, Y. Suzuki and A. Ichihara, *J. Chem. Soc., Chem. Commun.*, 1995, 1321–1322.
- 205 D. J. Witter and J. C. Vederas, *J. Org. Chem.*, 1996, **61**, 2613–2623.
- 206 H. Oikawa and T. Tokiwano, *Nat. Prod. Rep.*, 2004, **21**, 321–352.
- 207 H. Oikawa, in *Comprehensive Natural Products II Chemistry and Biology*, eds. H.-W. Liu and L. Mander, Elsevier, Oxford, 2010, pp. 277–314.
- 208 C. Reichard and T. Welton, *Solvent and Solvent Effects in Organic Chemistry (Wiley-VCH)*, VCH, Cambridge, 2011.
- 209 R. Breslow, *Acc. Chem. Res.*, 1991, **24**, 159–164.
- 210 W. Blokzijl, M. J. Blandamer and J. B. F. N. Engberts, *J. Am. Chem. Soc.*, 1991, **113**, 4241–4246.
- 211 K. Auclair, A. Sutherland, J. Kennedy, D. J. Witter, J. P. Van den Heever, C. R. Hutchinson and J. C. Vederas, *J. Am. Chem. Soc.*, 2000, **122**, 11519–11520.
- 212 M. Ohashi, F. Liu, Y. Hai, M. Chen, M. Tang, Z. Yang, M. Sato, K. Watanabe, K. N. Houk and Y. Tang, *Nature*, 2017, **549**, 502–506.
- 213 S. Liu and C. Cheng, *Wiley Interdiscip Rev RNA*, 2015, **4**, 547–566.
- 214 Q. Zheng, Y. Guo, L. Yang, Z. Zhao, Z. Wu, H. Zhang, J. Liu, X. Cheng, J. Wu, H. Yang, H. Jiang, L. Pan and W. Liu, *Cell Chem. Biol.*, 2016, **23**, 352–360.
- 215 Z. Tian, P. Sun, Y. Yan, Z. Wu, Q. Zheng, S. Zhou, H. Zhang, F. Yu, X. Jia, D. Chen, A. Mándi, T. Kurtán and W. Liu, *Nat. Chem. Biol.*, 2015, **11**, 259.
- 216 H. Oikawa and T. Tokiwano, *Nat. Prod. Rep.*, 2004, **21**, 321–352.
- 217 H. Oh, D. C. Swenson, J. B. Gloer, D. T. Wicklow and P. F. Dowd, *Tetrahedron Lett.*, 1998, **39**, 7633–7636.
- 218 G. Stork and E. Nakamura, *J. Am. Chem. Soc.*, 1983, **105**, 5510–5512.
- 219 S. Jacob, T. Grötsch, A. J. Foster, A. Schüffler, P. H. Rieger, L. P. Sandjo, J. C. Liermann, T. Opatz and E. Thines, *Microbiol. (United Kingdom)*, 2017, **163**, 541–553.
- 220 A. A. Salamov and V. V. Solovyev, *Genome Res.*, 2000, **10**, 516–522.
- 221 A. de Marco, E. Deuerling, A. Mogk, T. Tomoyasu and B. Bukau, *BMC Biotechnol.*, 2007, **7**.
- 222 M. N. Heneghan, A. A. Yakasai, K. Williams, K. A. Kadir, Z. Wasil, W. Bakeer, K. M. Fisch, A. M. Bailey, T. J. Simpson, R. J. Cox and C. M. Lazarus, *Chem. Sci.*, 2011, **2**, 972–979.
- 223 S. Iwasaki, S. Nozoe, S. Okuda, Z. Sato and T. Kozaka, *Tetrahedron Lett.*, 1969, **10**, 3977–3980.
- 224 R. D. Gietz and R. A. Woods, *Methods Enzymol.*, 2002, **350**, 87–96.
- 225 N. J. Talbot, D. J. Ebbale and J. E. Hamer, *Plant Cell*, 1993, **5**, 1575–1590.

11. Appendix

11.1 NMR Data Chapter 6

Compound 139

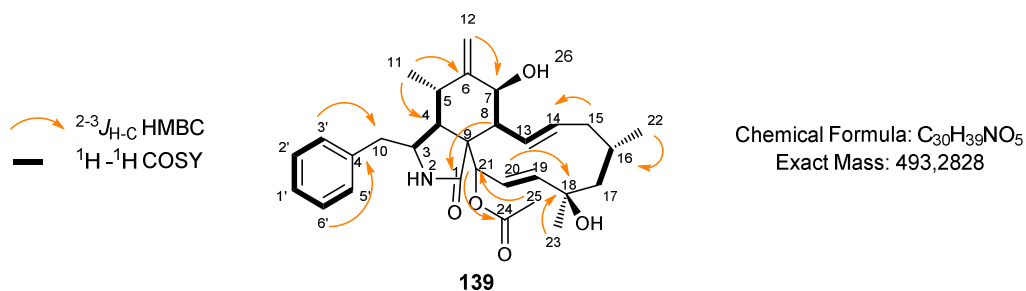


Figure 11.1: Isolated compound 139.

Table 11.1.1: NMR data of 139. Not all ¹H signals could be integrated due to overlap. Solvent: MeOD.

Atom	δ_c ppm	δ_H ppm	Multiplicity (J, Hz)	COSY	HMBC	HSQC
1'	127.9	7.23	d (7.4)		3'	CH
2'	129.6	7.30	d (7.6)	6'	6', 4'	CH
3'	130.9	7.19	d (8.2)	5'	1', 5', 10	CH
4'	134.4	-				CH
5'	130.9	7.19	d (8.2)	3'	1', 3', 10	CH
6'	129.6	7.30	d(7.6)	2'	2', 4'	CH
1	177.8	-				
2	-	-				
3	55.1	3.29	m	4, 10	1, 5, 4'	CH
4	50.0	2.18	dd (2.9, 5.4)	3, 5	1, 5, 6, 8, 9, 21	CH
5	33.6	2.63	m	4, 11	3, 4, 6, 12	CH
6	151.2	-				CH
7	72.4	3.8	d (10.5)	8	5, 6, 8, 12, 13	CH
8	47.8	2.94	t (10.0, 10.0)	7, 13	1	CH
9	53.8					
10	45.2	2.88, 2.72	dd (5.5, 13.2) dd (8.2, 13.3)	3	3, 3', 4'	CH ₂
11	13.6	0.57	d (6.8)	5	4, 5, 6	CH ₃
12	113.4	5.20 5.00	br. s br. s	5, 7	7, 5	CH ₂
13	129.3	5.64	m	14	7, 8, 15	CH
14	137.7	5.30	m	13, 15	8, 15	CH ₂
15	44.5	2.03 1.75	m m	14	13, 14, 16	CH ₂
16	29.2	1.77	m	17, 22		CH
17	55.0	1.54 1.78	m m	16	16, 18, 19, 22	CH ₂
18	74.8	-				
19	138.9	5.53	d (2.4)	20		CH
20	126.9	5.74	dd (2.4, 16.6)	19	18, 21	CH
21	78.4	5.41	t (2.5, 2.5)	20	8, 19, 20, 24,	CH
22	26.5	1.02	d (6.2)	16	16	CH ₃
23	30.5	1.28	s		17, 18, 19	CH ₃
24	172.0	-				CH
25	20.7	20.70	s		21	CH ₃

11.2 NMR Data Chapter 7

Characteristic ^1H - ^1H -COSY and HMBC signals for compounds **179** – **186** on the example of compound **185**, **179** and **184** (Figure 11.2, Figure 11.3, Figure 11.4, Figure 11.5).

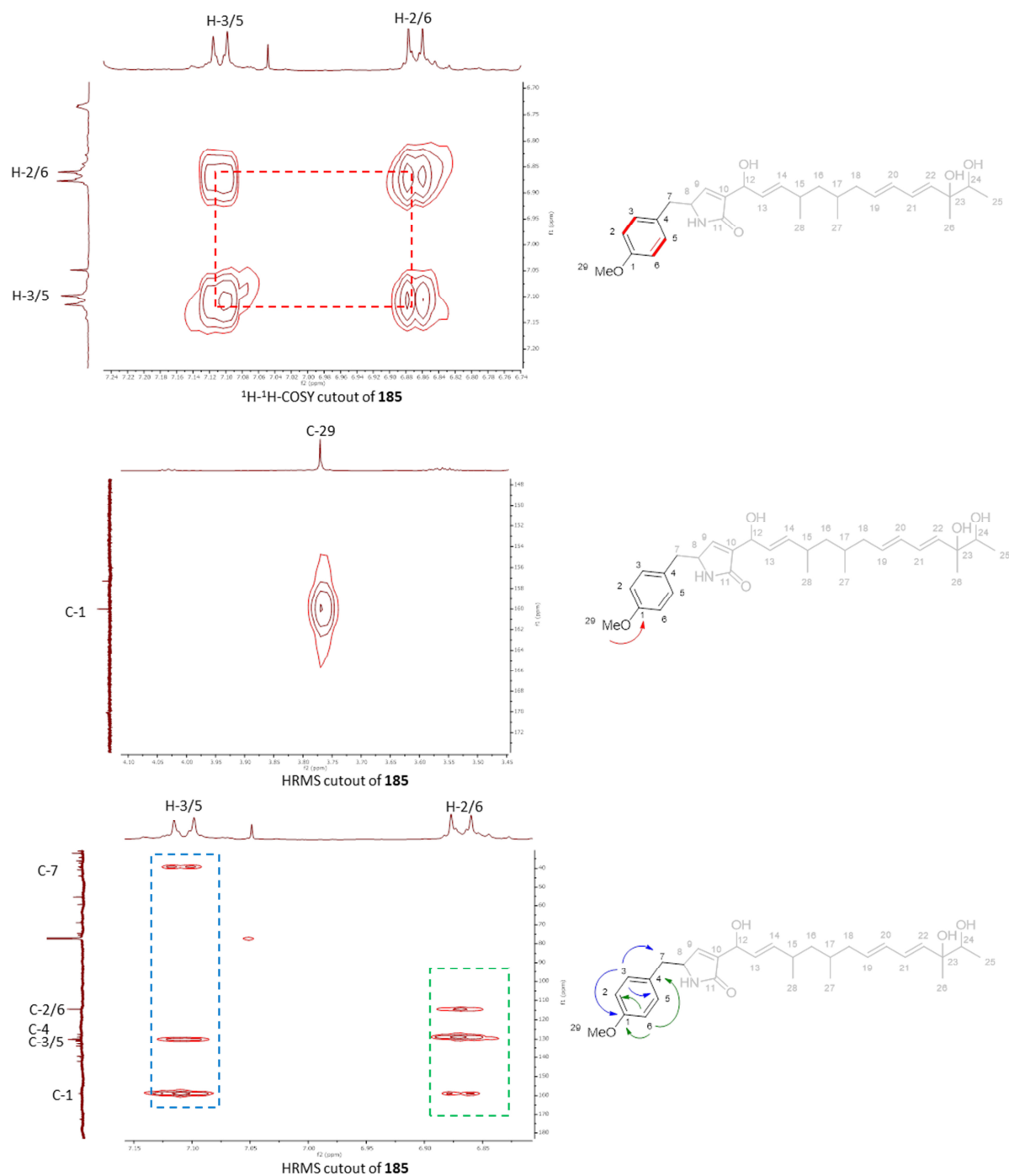


Figure 11.2: ^1H - ^1H -COSY and HMBC signals for **185**.

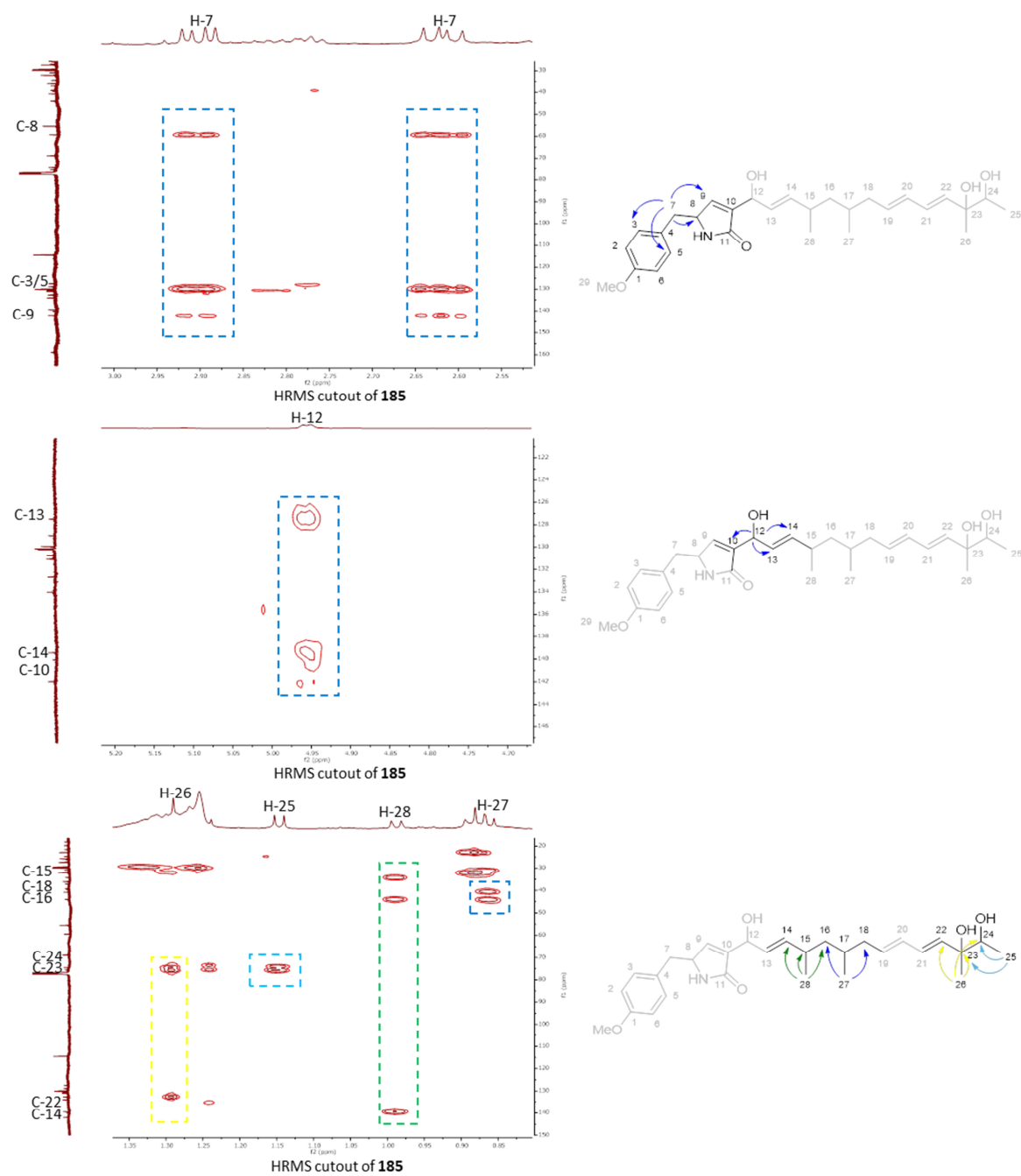


Figure 11.3: ^1H - ^1H -COSY and HMBC signals for **185**.

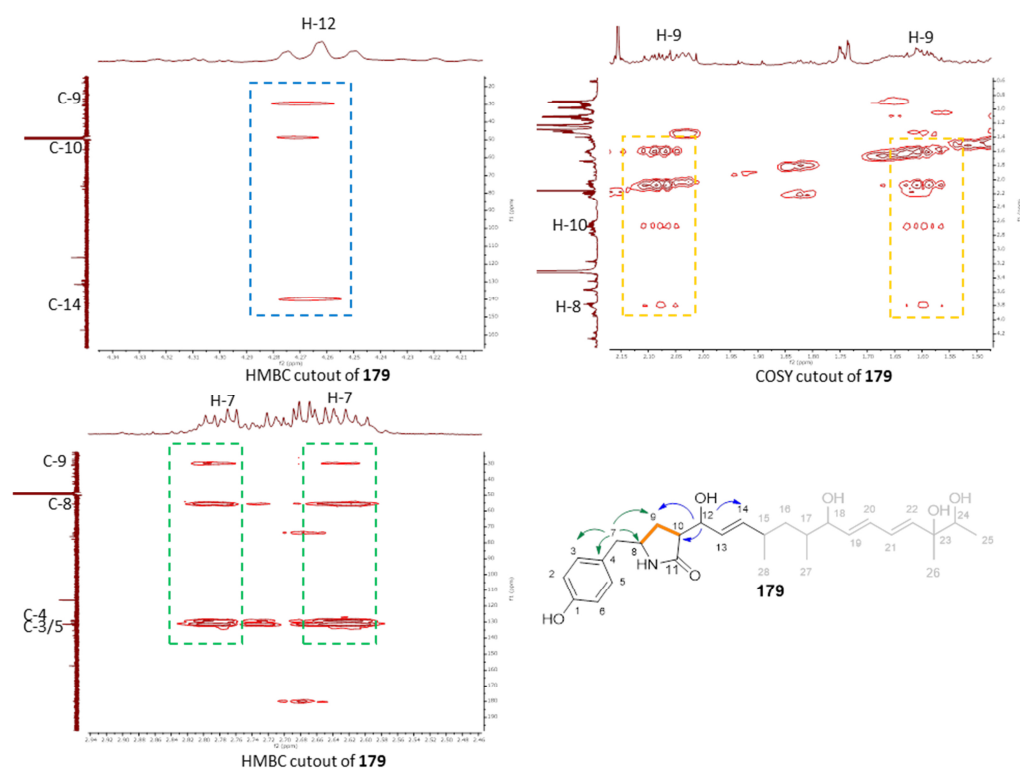


Figure 11.4: Characteristic ^1H - ^1H -COSY and HMBC signals of the pyrrolidin-2-one ring on the example of compound **179**.

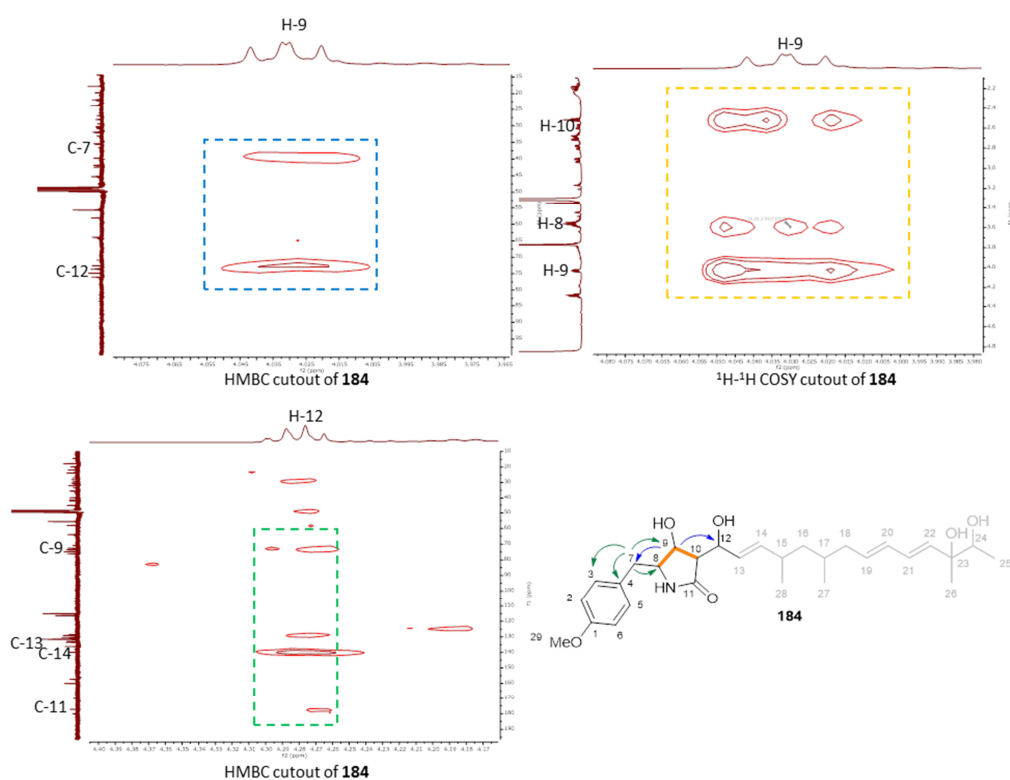


Figure 11.5: Characteristic ^1H - ^1H -COSY and HMBC signals of the 4-hydroxypyrrolidin-2-one ring on the example of compound **184**.

Compound 179

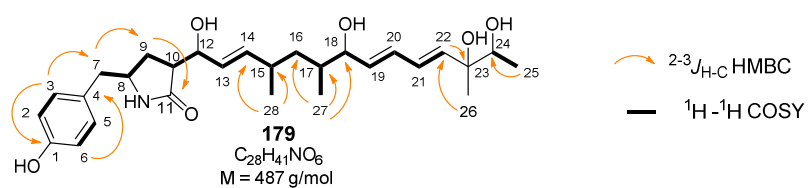
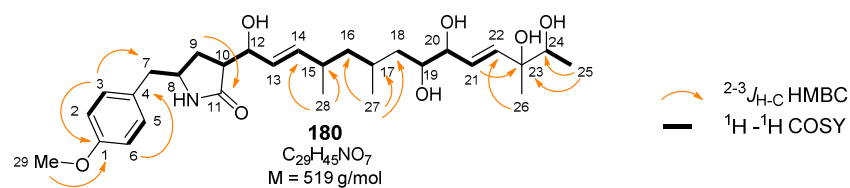


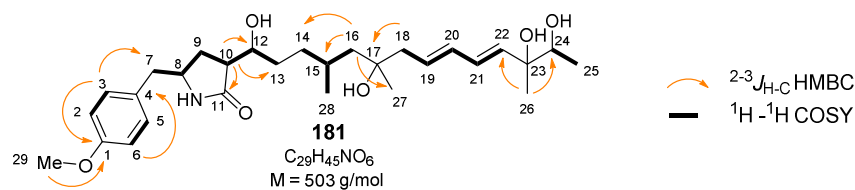
Figure 11.6: Isolated compound 179.

Table 11.2.1: NMR data of 179. Not all ¹H signals could be integrated due to overlap. Solvent: MeOD

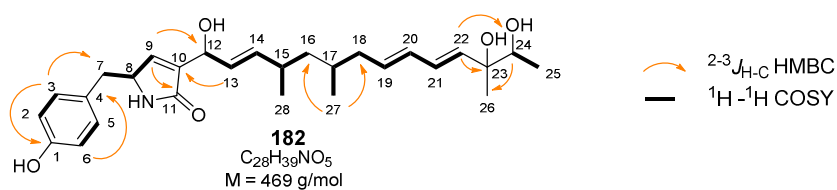
Aatom	δ _c ppm	δ _H ppm	Multiplicity (J, Hz)	COSY	HMBC	HSQC
1	157.30	-				
2	116.3	6.73	m	3	1, 4, 6	CH
3	131.4	7.02	d (8.5)	2	1, 2, 5, 7	CH
4	129.3	-				
5	131.4	7.02	d (8.5)	6	1, 2, 5, 7	CH
6	116.3	6.73	m	5	1, 4, 6	CH
7	42.4	2.78	m	8	3, 5, 4, 8, 9	CH ₂
		2.63	m			
8	55.6	3.81	m	7, 9	4,	CH
9	29.4	1.60	m	8, 10		CH
		2.10	m			
10	49.0	2.58	m	7, 8, 9	11, 12	CH
11	180.0	-				
12	73.6	4.27	m	10, 13	9, 10, 14	CH
13	130.9	5.36	m	12, 14		CH
14	139.9	5.50	m	13,15		CH
15	35.6	2.24	m	14, 16, 28	13, 14	CH
16	41.0	1.45	m	15, 17		CH ₂
		1.10	m			
17	37.9	1.65	m	27		
18	77.7	3.92	m	17, 19	19, 20	CH ₂
19	135.5	5.67	m	18, 20	21, 23, 24	CH
20	132.2	6.22	m	19	18, 21, 22	CH
21	129.7	6.28	m	22		CH
22	138.4	5.83	d (15.1)	21	20, 23, 26	CH
23	76.0	-				
24	74.8	3.57	d (6.5)	25		CH
25	17.8	1.11	d (6.5)	24	24, 26	CH ₃
26	23.7	1.23	s		22, 24	CH ₃
27	15.1	0.89	d (6.7)	17	16, 17, 18	CH ₃
28	22.5	0.98	d (6.6)	15	14, 15, 16	CH ₃

Compound **180****Figure 11.7:** Isolated compound **180**.**Table 11.2.2:** NMR data of **180**. Not all ^1H signals could be integrated due to overlap. Solvent: MeOD

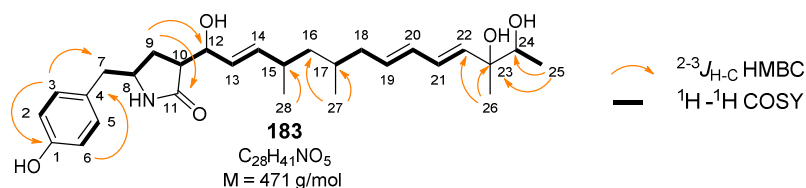
Atom	δ_{C} ppm	δ_{H} ppm	Multiplicity (J , Hz)	COSY	HMBC	HSQC
1	160.0	-				
2	115.0	6.87	d (8.6)	3	1, 4, 6	CH
3	131.3	7.15	d (8.6)	2	1, 2, 5, 7	CH
4	130.5	-				
5	131.3	7.15	d (8.6)	6	1, 2, 5, 7	CH
6	115.0	6.87	d (8.6)	5	1, 4, 6	CH
7	42.3	2.78	m	8	8	CH ₂
		2.63	m			
8	55.3	3.83	m	7, 9		CH
9	29.3	1.60	m	8, 10		
		2.10	m			CH
10	49.1	2.66	m	9, 12	11	CH
11	180.0	-				
12	74.0	4.22	m	10, 13	9, 10, 11, 14	CH
13	131.3	5.34	m	12, 14	12, 15	CH
14	140.3	5.55	m	13, 15	12, 15, 28	CH
15	35.3	2.26	m	16, 28	17	CH
16	46.1	1.29	m	17		CH ₂
		1.16	m			
17	27.8	1.76	m	16		CH
18	49.0	1.40	m			CH ₂
		1.27	m			
19	74.9	3.60	m	18, 20		CH
20	77.6	3.92	m	19, 21		CH
21	129.8	5.80	m	20	23	CH
22	137.7	5.80	m			CH
23	76.0	-				
24	74.9	3.58	m	25	23, 25, 26	CH
25	17.8	1.11	d (6.5)	24	24	CH ₃
26	23.8	1.24	s		22, 24, 25	CH ₃
27	20.6	0.89	d (6.6)		15, 16	CH ₃
28	22.2	0.96	d (6.6)	15	14, 15, 16	CH ₃
29	55.7	3.77	s		1	CH ₃

Compound **181****Figure 11.8:** Isolated compound **181****Table 11.2.3:** NMR data of **181** main. Not all ¹H signals could be integrated due to overlap. Solvent: MeOD.

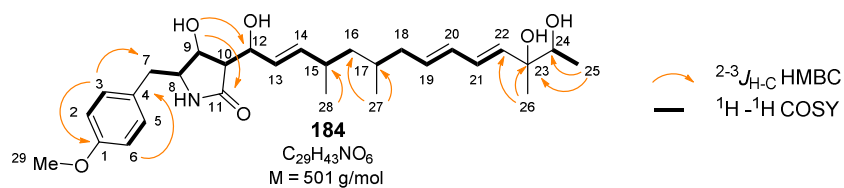
Atom	δ_c ppm	δ_H ppm	Multiplicity (J, Hz)	COSY	HMBC	HSQC
1	160.0	-	-			
2	114.9	6.86	d (8.6)	3	1, 4, 6	CH
3	131.6	7.14	d (8.6)	2	1, 2, 5, 7	CH
4	130.4	-	-			
5	131.6	7.14	d (8.6)	6	1, 2, 5, 7	CH
6	114.9	6.86	d (8.6)	5	1, 4, 6	CH
7	41.9	2.83	dd (5.6, 13.7)	8	3, 8, 4, 9	CH ₂
		2.71	dd (7.0, 13.6)			
8	55.2	3.84	m	7, 9	4, 9	CH
9	29.8	1.58	m	8, 10	10	CH ₂
		2.13	m			
10	48.5	2.57	m	9	9, 11, 12, 13	CH
11	180.8	-	-			
12	73.5	3.65	m	10, 13	9, 11, 14	CH
13	32.1	1.31	m	12		CH
14	35.7	1.15	m	15		CH ₂
		1.56	m			
15	29.9	1.63	m	14		CH
16	49.5	1.46	dd (4.0, 14.4)	15	14, 15, 17, 18, 27, 28	CH ₂
		1.27				
17	74.1	-	-			
18	47.0	2.25	d (1.2)	19	16, 17, 19, 20, 27	CH ₂
19	130.7	5.73	m	18, 20		CH
20	134.4	6.09	m	21	18, 21, 22	CH
21	130.3	6.28	m	20, 22		CH
22	136.8	5.72	m	21	20, 21, 23, 26	CH
23	76.0	-	-			
24	74.8	3.55	m	25	23, 25, 26	CH
25	17.8	1.10	d (6.5)	24	24	CH ₃
26	23.8	1.23	s		22, 24	CH ₃
27	27.3	1.14	s			CH ₃
28	22.6	0.94	d (6.6)	15	14, 15, 16	CH ₃
29	55.7	3.77	s		1	CH ₃

Compound **182**Figure 11.9: Isolated compound **182**.Table 11.2.4: NMR data of **182**. Not all ¹H signals could be integrated due to overlap. Solvent: MeOD.

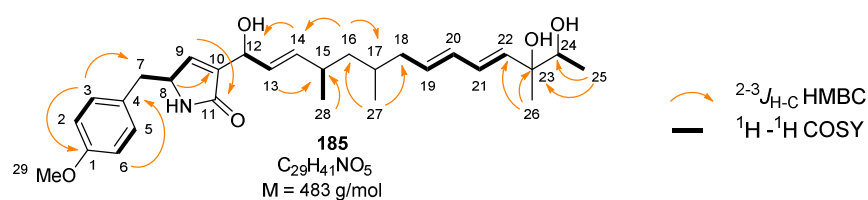
Atom	δ_c ppm	δ_H ppm	Multiplicity (J, Hz)	COSY	HMBC	HSQC
1	157.3	-				
2	116.1	6.69	d (8.0)	3	1, 4, 6	CH
3	131.6	7.00	d (8.5)	2	1, 2, 5, 7	CH
4	128.2	-				
5	131.6	7.00	d (8.5)	6	1, 2, 5, 7	CH
6	116.1	6.69	d (8.0)	5	1, 4, 6	CH
7	39.1	2.90	dd (5.6, 13.7)	8	3, 4, 8, 9	CH ₂
		2.76	m			
8	60.2	4.34	m	7, 9	4, 7, 9, 10	CH
9	144.5	6.91	t (1.6, 1.6)	8	8, 11, 12	CH
10	142.5	-				CH
11	174.5	-				
12	67.9	4.76	m	13	9, 10, 13, 14	CH
13	129.6	5.4	dd (6.0, 15.5)	12	10, 12, 15	CH
14	138.9	5.34	m	15	12, 15	CH
15	35.5	2.21	m	16, 28		CH
16	45.1	1.31	m			CH ₂
17	31.9	1.61	m		18	CH
18	41.8	2.03	m	17, 19	17, 19, 27	CH ₂
		1.92	m			
19	133.8	5.62	m	18, 20	17, 18	CH
20	132.7	6.02	m	19, 21	18, 21, 22	CH
21	130.4	6.25	m	20, 22	19, 23	CH
22	136.7	5.72	m	21	21, 23, 26	CH
23	76.2	-				
24	74.9	3.55	m	25	22, 23, 25, 26	CH
25	17.8	1.09	d (6.4)	24	24	CH ₃
26	23.8	1.24	s		23	CH ₃
27	19.7	0.86	d (6.6)	17	16, 17, 18	CH ₃
28	21.9	0.93	d (6.7)	15	14, 16	CH ₃

Compound **183**Figure 11.10: Isolated compound **183**.Table 11.2.5: NMR data of **183**. Not all ¹H signals could be integrated due to overlap. Solvent: MeOD.

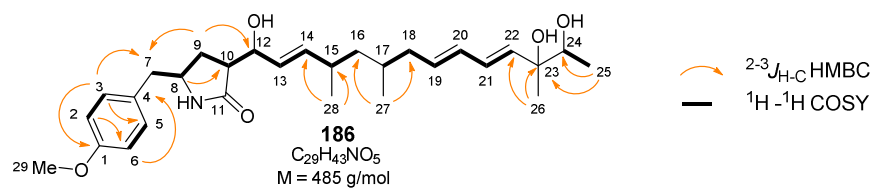
Atom	δ_c ppm	δ_H ppm	Multiplicity (J, Hz)	COSY	HMBC	HSQC
1	157.3	-				
2	116.4	6.73	d (8.5)	3	4, 6	CH
3	131.4	7.01	d (8.5)	2	1, 5, 7	CH
4	128.6	-				
5	131.4	7.01	d (8.5)	6	1, 5, 7	CH
6	116.4	6.73	d (8.5)	5	4, 6	CH
7	42.5	2.79	dd (5.6, 13.5)	8	8, 9	CH ₂
		2.59	dd (7.4, 13.5)			
8	55.6	3.77	m	7, 9	4, 9	CH
9	29.4	2.03	m	8, 10	7, 12	CH ₂
		1.57	m		12	
10	49.9	2.68	m		11, 12, 13	CH
11	180.0	-				
12	73.5	4.28	m	10	9, 11, 13, 14	CH
13	129.2	5.32	m	12, 14	12, 14, 15	CH
14	139.9	5.52	m	15	12, 13, 15	CH
15	35.6	2.23	m	16, 28		CH
16	45.5	1.35	m	15		CH ₂
17	32.1	1.61	m	18		CH
18	42.0	2.06	m	17, 19	16, 17, 20, 27	CH ₂
		1.95	m			
19	133.6	5.66	m	20		CH
20	132.9	6.05	dd (10.3, 15.1)	19, 21		CH
21	130.3	6.24	m	20, 22	22, 23	CH
22	136.2	5.71	d (15.6)	21	21, 23, 26	CH
23	76.2	-				
24	74.9	3.55	m	25	23, 24, 26	CH
25	17.8	1.08	d (5.4)	24	23, 24	CH ₃
26	23.8	1.22	d (4.7)		22, 23, 24	CH ₃
27	19.9	0.89	d (6.8)	17	16, 17, 18	CH ₃
28	22.2	0.97	d (5.0)	15	14, 15, 16	CH ₃

Compound **184**Figure 11.11: Isolated compound **184**.Table 11.2.6: NMR data of **184**. Not all ¹H signals could be integrated due to overlap. Solvent: MeOD.

Atom	δ_c ppm	δ_H ppm	Multiplicity (J, Hz)	COSY	HMBC	HSQC
1	160.0	-				
2	115.0	6.87	d (8.7)	3	4, 6	CH
3	131.6	7.14	d (8.7)	2	1, 5, 7	CH
4	130.3					
5	131.6	7.14	d (8.7)	6	1, 5, 7	CH
6	115.0	6.87	d (8.7)	5	4, 6	CH
7	39.7	2.92	dd (5.4,14.0)	8	8, 9	CH ₂
		2.70	m			
8	63.9	3.57	m	7, 9	4, 9	CH
9	73.6	4.03	dd (4.7, 6.0)	8, 10	7, 12	CH
10	58.0	2.52	dd (5.8, 5.8)		9, 11, 12, 13	CH
11	177.2	-				
12	72.6	4.28	m	10	9, 11, 13, 14	CH
13	129.1	5.35	dd	12, 14	12, 14, 15	CH
14	139.9	5.48	dd (6.5,9.6)	15	12, 13, 15	CH
15	35.4	2.23	m	16, 28		CH
16	45.3	1.34	m	15		CH ₂
17	31.9	1.61	m			CH
18	41.8	2.06	m	17, 19	16, 17, 20, 27	CH ₂
		1.95	m			
19	133.6	5.66	m	20		CH
20	132.9	6.05	dd (10.3,15.1)	19, 21		CH
21	130.3	6.24	m	20, 22	22, 23	CH
22	136.2	5.71	d (15.6)	21	21, 23, 26	CH
23	76.0	-				
24	74.8	3.55	m	25	23, 24, 26	CH
25	17.8	1.10	d (5.4)	24	23, 24	CH ₃
26	23.8	1.22	d (4.7)		22, 23, 24	CH ₃
27	19.8	0.87	d (6.8)	17	16, 17, 18	CH ₃
28	21.9	0.96	d (5.0)	15	14, 15, 16	CH ₃
29	55.7	3.77	s		1	CH ₃

Compound **185**Figure 11.12: Isolated compound **185**Table 11.2.7: NMR data of **185**. Not all ^1H signals could be integrated due to overlap. Solvent: CDCl₃.

Atom	δ_{C} ppm	δ_{H} ppm	Multiplicity (J , Hz)	COSY	HMBC	HSQC
1	159.0	-				
2	114.5	6.87	d (8.6)	3	1, 4, 6	CH
3	130.1	7.11	d (8.6)	2	1, 2, 5, 7	CH
4	128.9	-				
5	130.1	7.11	d (8.6)	6	1, 2, 5, 7	CH
6	114.5	6.87	d (8.6)	5	1, 4, 6	CH
7	39.2	2.9	dd (5.8, 13.6)		8, 3, 5, 9	CH ₂
		2.62	dd (8.9, 13.6)			
8	59.4	4.25	m	7, 9	4, 7, 9, 10	CH
9	142.0	6.73	d (1.6)	8	8, 11	CH
10	140.1	-				
11	172.8	-				
12	68.9	4.95	d (4.9)	13	10, 13, 14	CH
13	127.5	5.58	m	12	15	CH
14	139.5	5.58	m		12, 18	CH
15	34.3	2.27	m	28		CH
16	44.1	1.32	m		14, 15, 17, 27	CH ₂
		1.07	m			
17	30.9	1.43	m			CH
18	40.8	2.03	m	19	16, 19, 27	CH ₂
		1.94	dd (7.2, 14.3)			
19	134.0	5.67	m	18, 20		CH
20	131.0	6.03	m	19, 21	18, 22	CH
21	130.7	6.28	m	22	19, 20, 23	CH
22	132.7	5.67	m	21	23, 24, 26	CH
23	75.5	-				
24	74.6	3.65	m	25	22	CH
25	18.1	1.15	d (6.5)	24	23, 24	CH ₃
26	24.7	1.29	s		23, 24, 22	CH ₃
27	19.8	0.86	d (5.9)		16, 18	CH ₃
28	21.6	0.99	d (6.7)	15	14, 15, 16	CH ₃
29	55.5	3.8	s		1	CH ₃

Compound **186**Figure 11.13: Isolated compound **186**.Table 11.2.8: NMR data of **186**. Not all ¹H signals could be integrated due to overlap. Solvent: CDCl₃.

Atom	δ_c ppm	δ_H ppm	Multiplicity (J, Hz)	COSY	HMBC	HSQC
1	158.8	-				-
2	114.5	6.86	m	3	4, 6	CH
3	130.2	7.09	m	2	1, 5, 7	CH
4	129.3	-				-
5	130.2	7.08	m	6	1, 5, 7	CH
6	114.5	6.86	m	5	4, 6	CH
7	42.5	2.78	m	8	8, 9	CH ₂
		2.61	m			
8	54.3	3.78	m	7, 9	4, 9	CH
9	27.7	2.02	m	8, 10	7, 12	CH ₂
		1.88	m			
10	46.5	2.64	m		9, 12, 13	CH
11	179.0	-				-
12	73.4	4.42	dd (3.5, 6.4)	10, 13	9, 11, 13, 14	CH
13	127.9	5.37	m	12, 14	12, 14, 15	CH
14	139.3	5.58	m	15	12, 13, 15	CH
15	34.3	2.24	m	16, 28		CH
16	44.4	1.31	m	15		CH ₂
17	31.0	1.59	m			CH
18	40.8	2.05	m	17, 19	16, 17, 20, 27	CH ₂
		1.93	m			
19	132.6	5.67	m	20		CH
20	130.7	6.28	m	19, 21		CH
21	131.1	6.05	m	20, 22	22, 23	CH
22	132.7	5.65	m	21	21, 23, 26	CH
23	75.5	-				-
24	74.5	3.65	m	25	23, 24, 26	CH
25	18.1	1.15	d (6.5)	24	23, 24	CH ₃
26	24.7	1.29	s		22, 23, 24	CH ₃
27	19.7	0.85	d (6.6)	17	16, 17, 18	CH ₃
28	21.8	0.96	d (6.7)	15	14, 15, 16	CH ₃
29	55.1	3.80	s		1	CH ₃

Compound 179

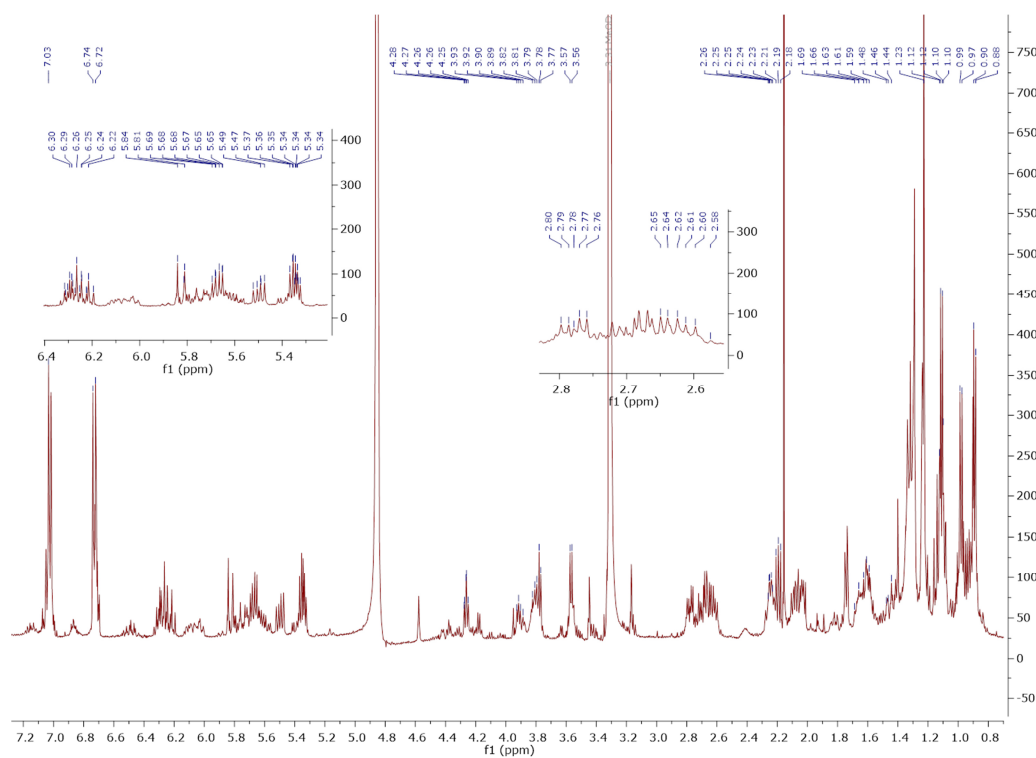


Figure 11.14: ^1H NMR of compound 179 in MeOD (500 MHz) referenced to MeOD.

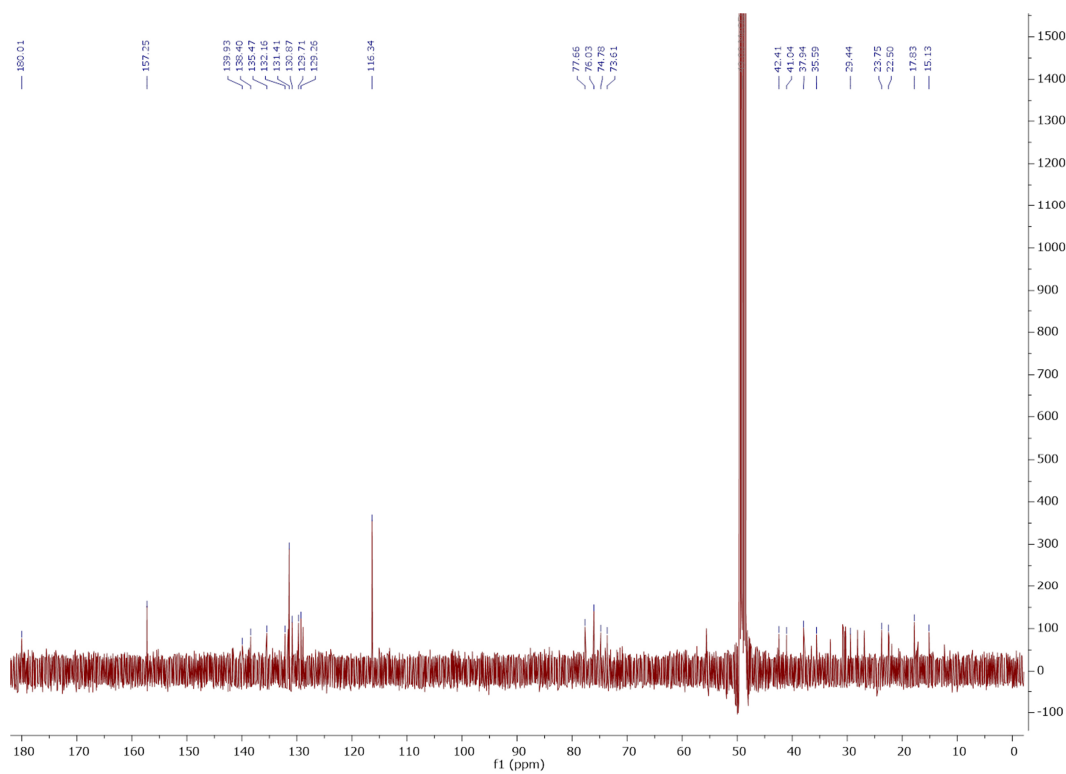


Figure 11.15: ^{13}C NMR of compound 179 in MeOD (500 MHz) referenced to MeOD.

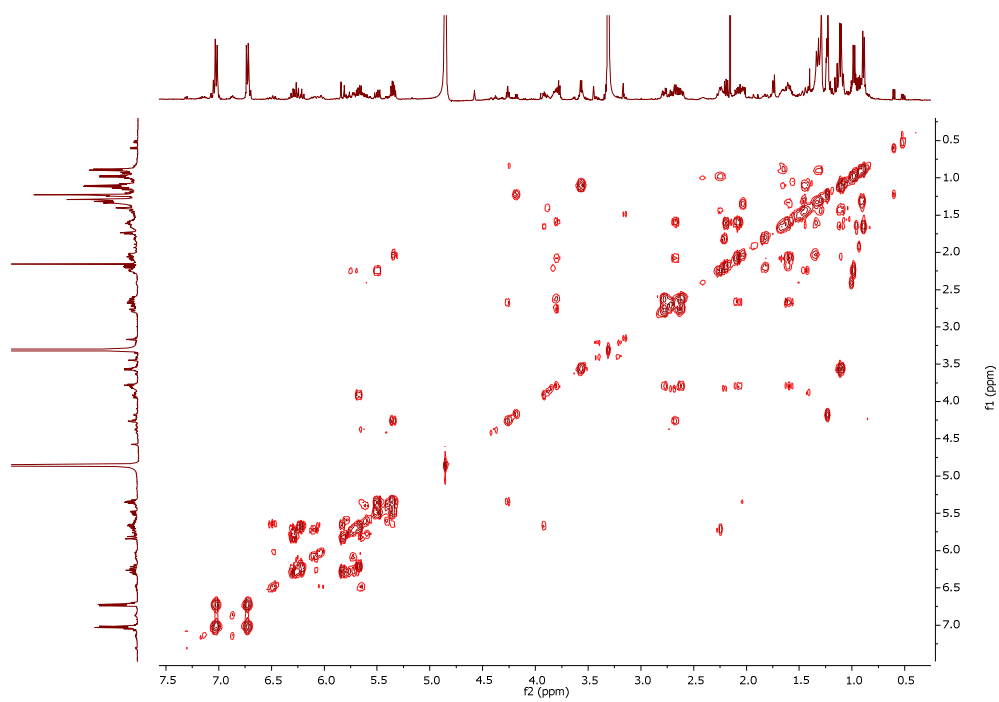


Figure 11.16: COSY of compound 179.

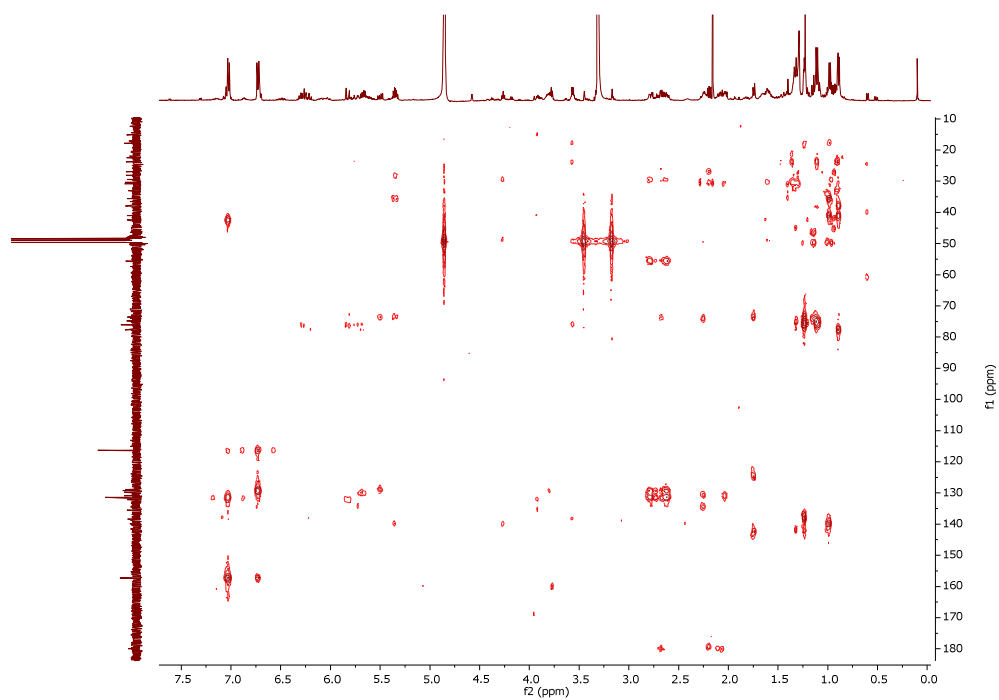


Figure 11.17: HMBC of compound 179.

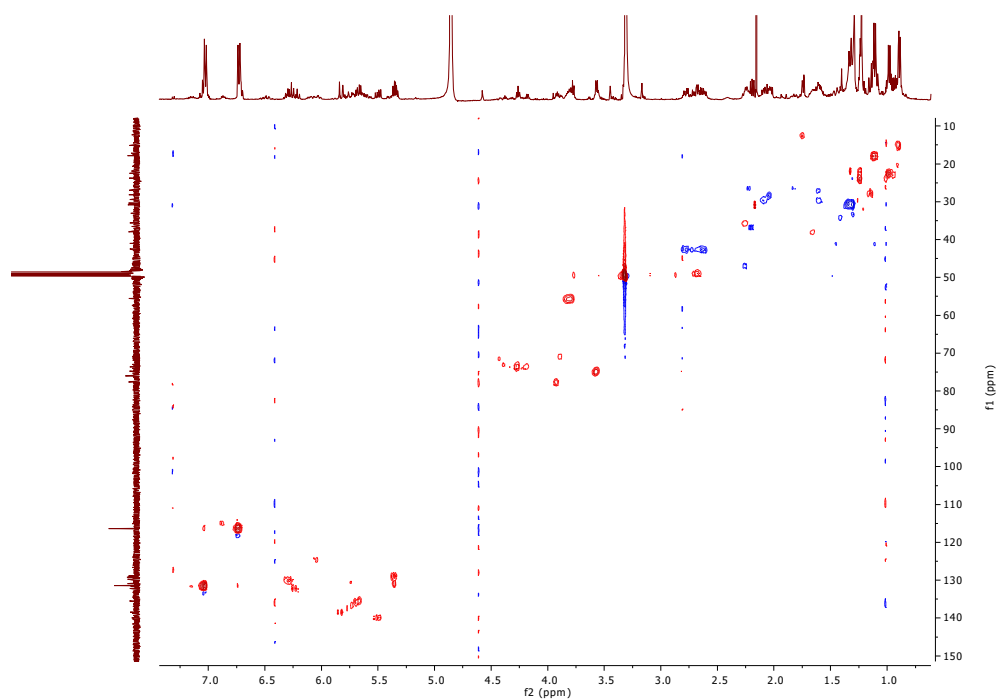
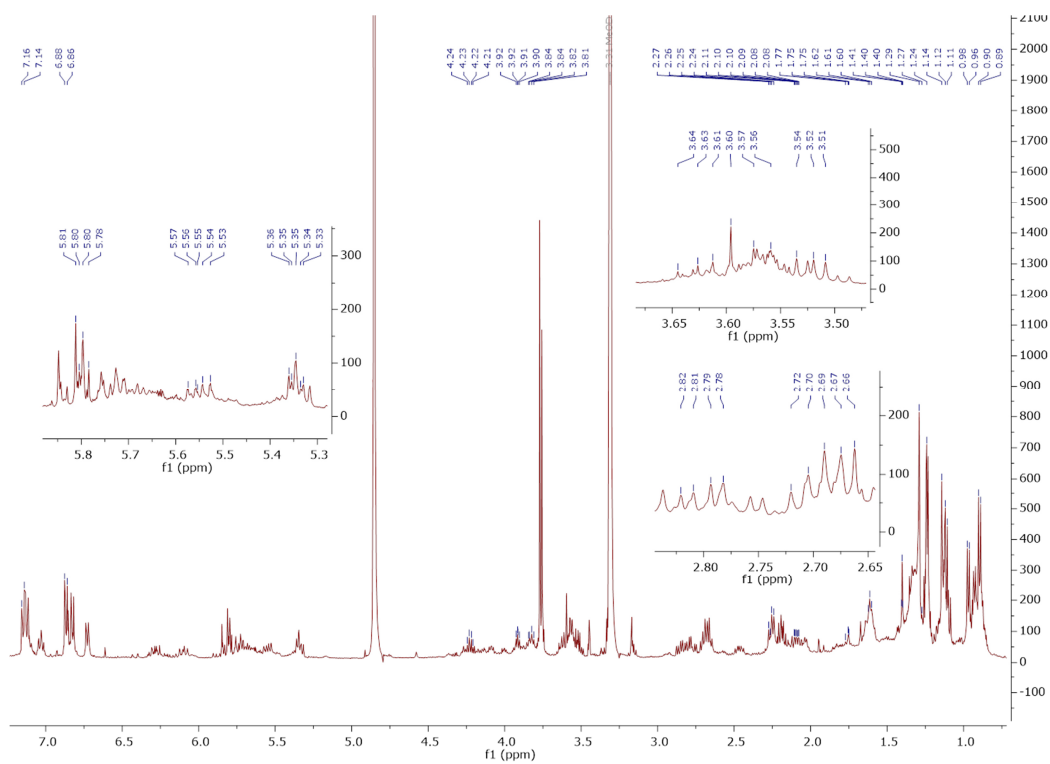


Figure 11.18: HSQC of compound 179.

Compound 180

Figure 11.19: ^1H NMR of compound 180 in MeOD (500 MHz) referenced to MeOD.

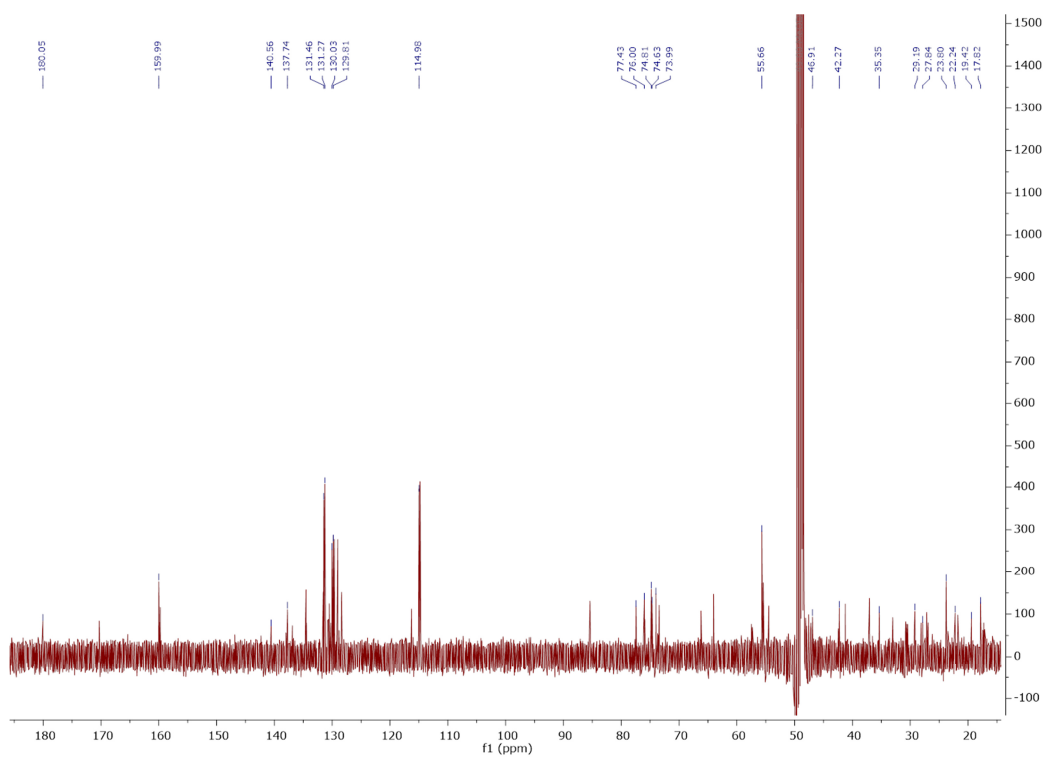


Figure 11.20: ^{13}C NMR of compound **180** in MeOD (500 MHz) referenced to MeOD.

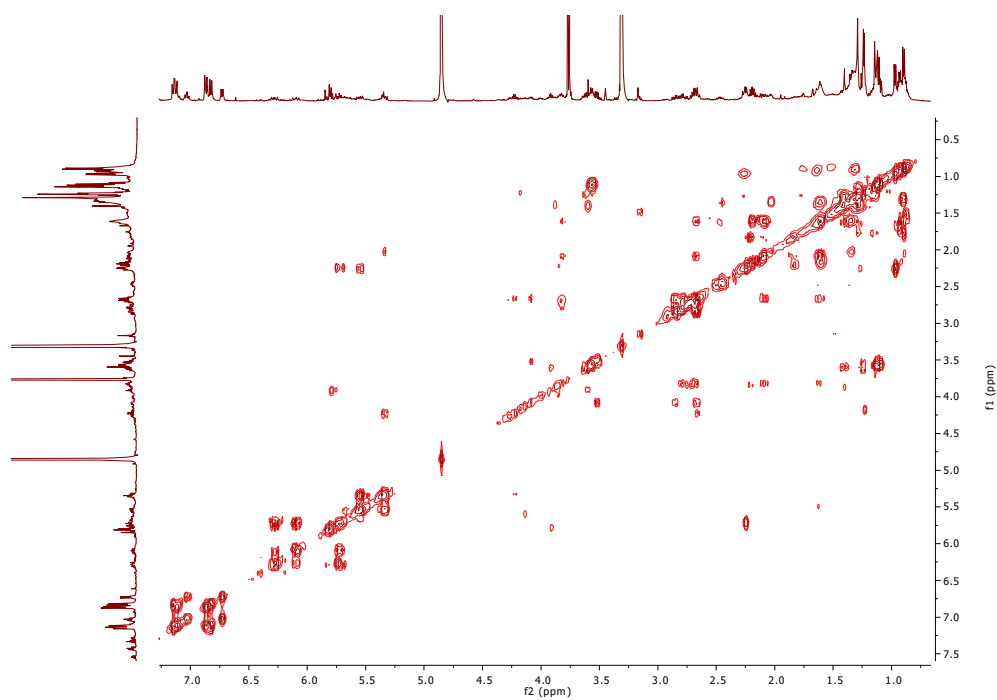


Figure 11.21: COSY of compound **180**.

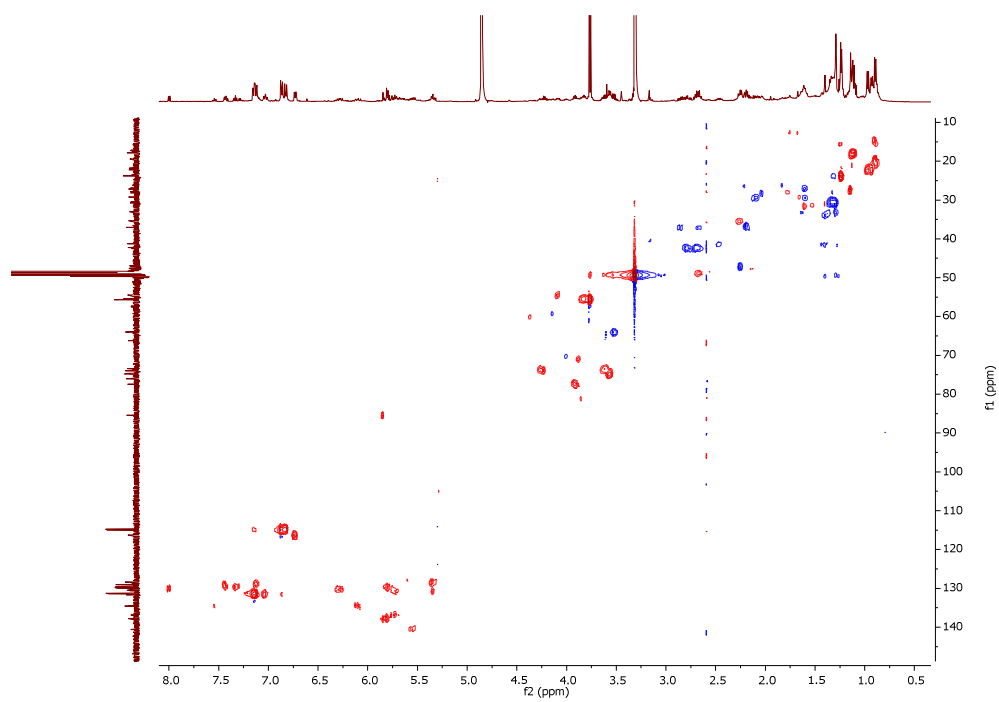


Figure 11.22: HSQC of compound 180.

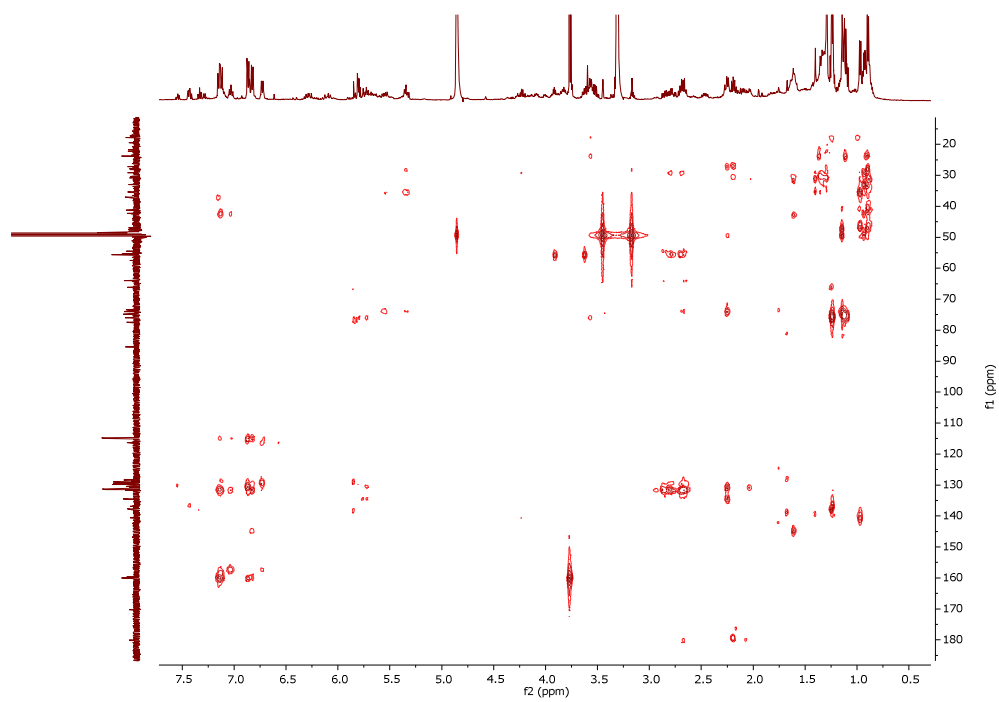


Figure 11.23: HMBC of compound 180.

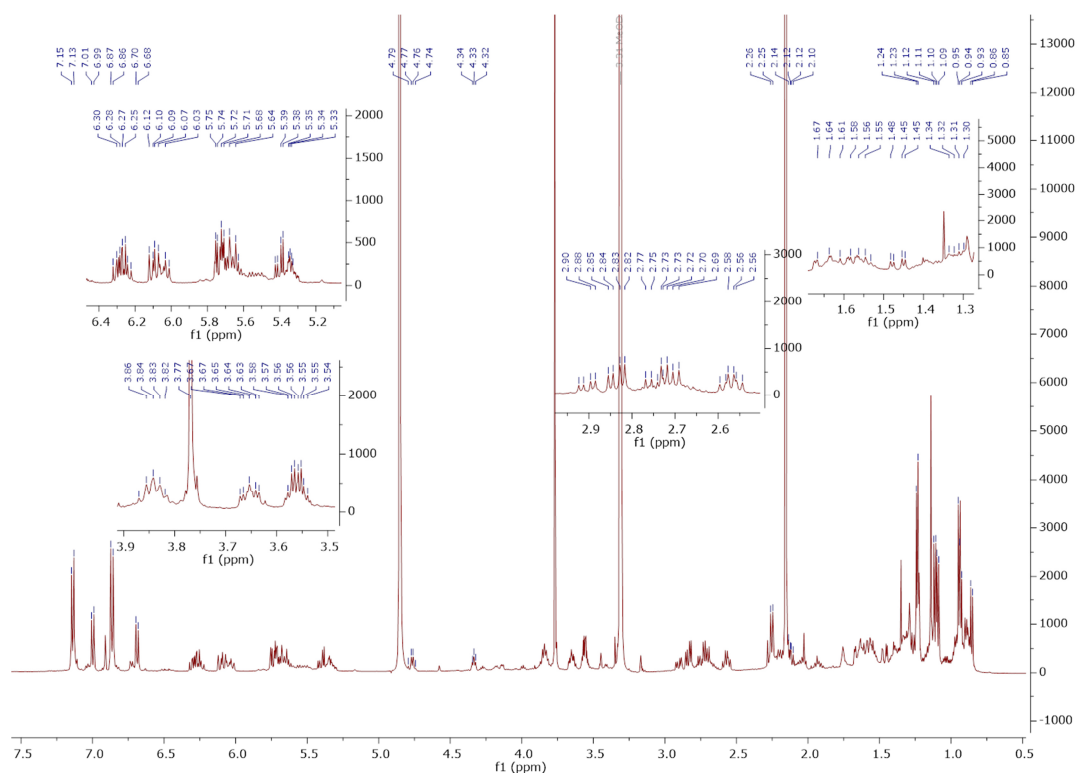
Compound **181** and **182**

Figure 11.24: ^1H NMR of compound **181** and **182** in MeOD (500 MHz) referenced to MeOD.

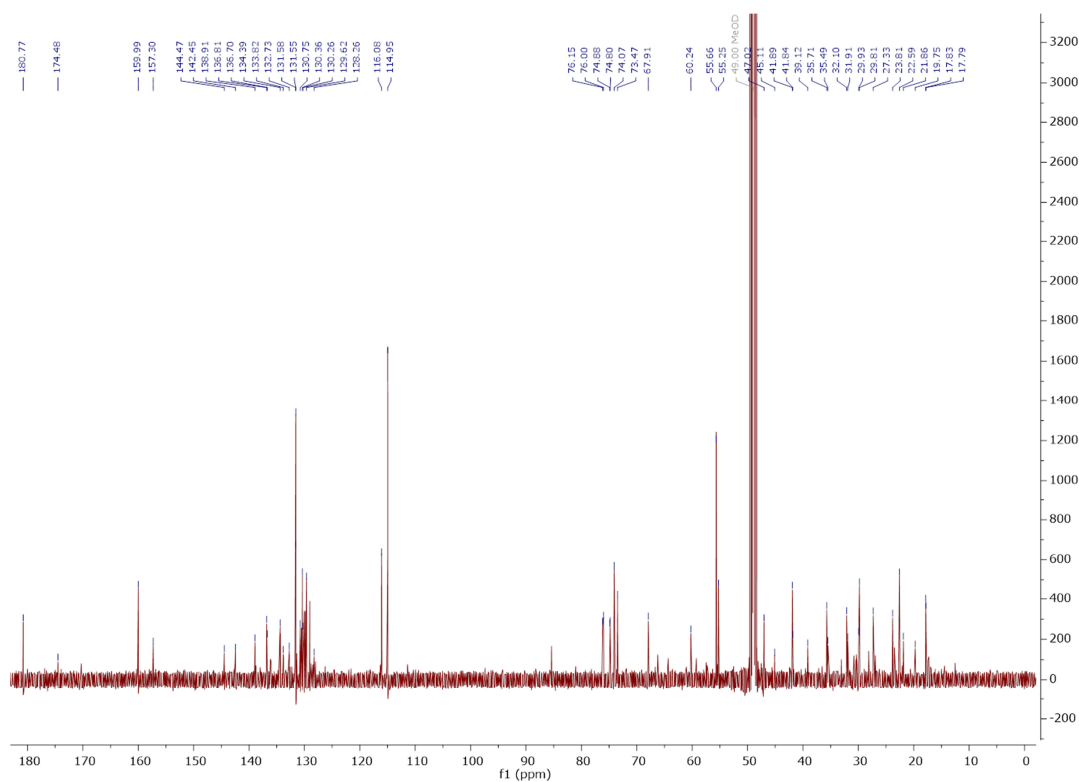


Figure 11.25: ^{13}C NMR of compound **181** and **182** in MeOD (500 MHz) referenced to MeOD.

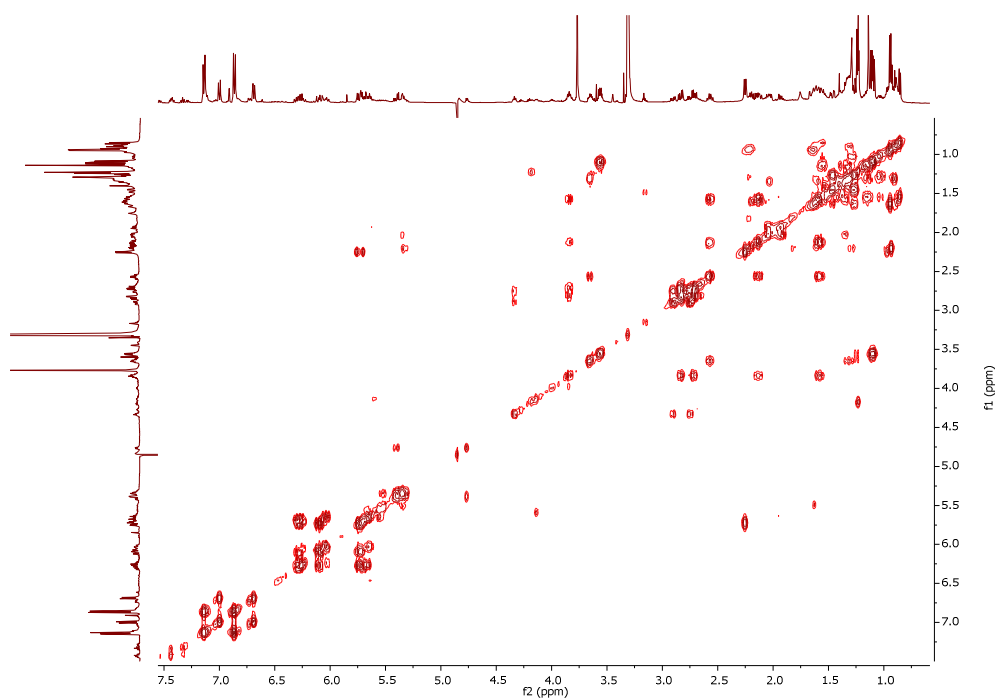


Figure 11.26: COSY of compound 181 and 182.

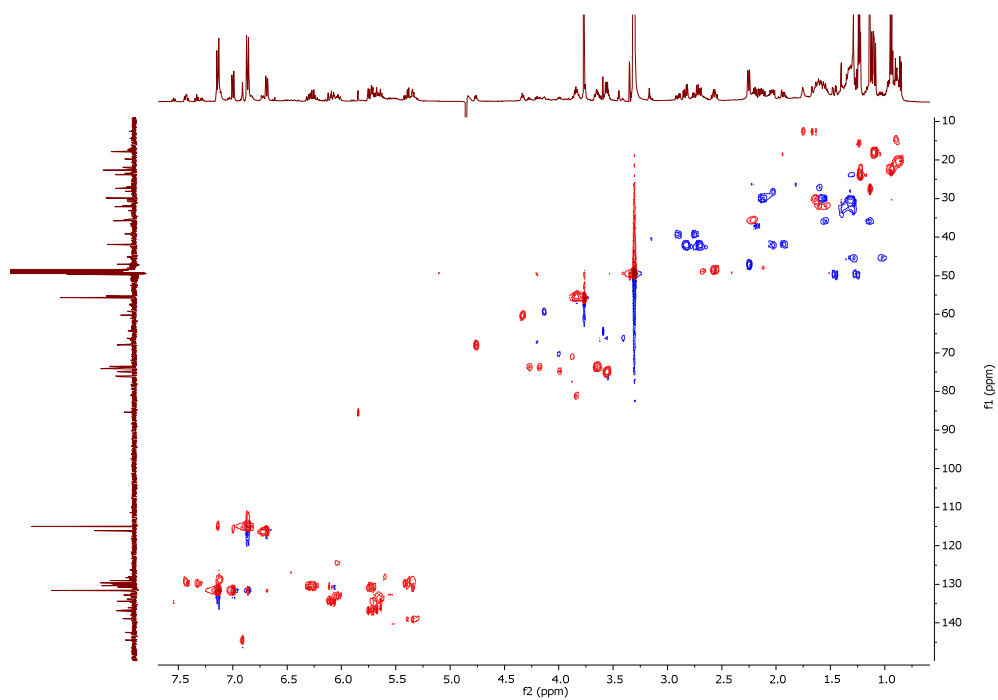


Figure 11.27: HSQC of compound 181 and 182.

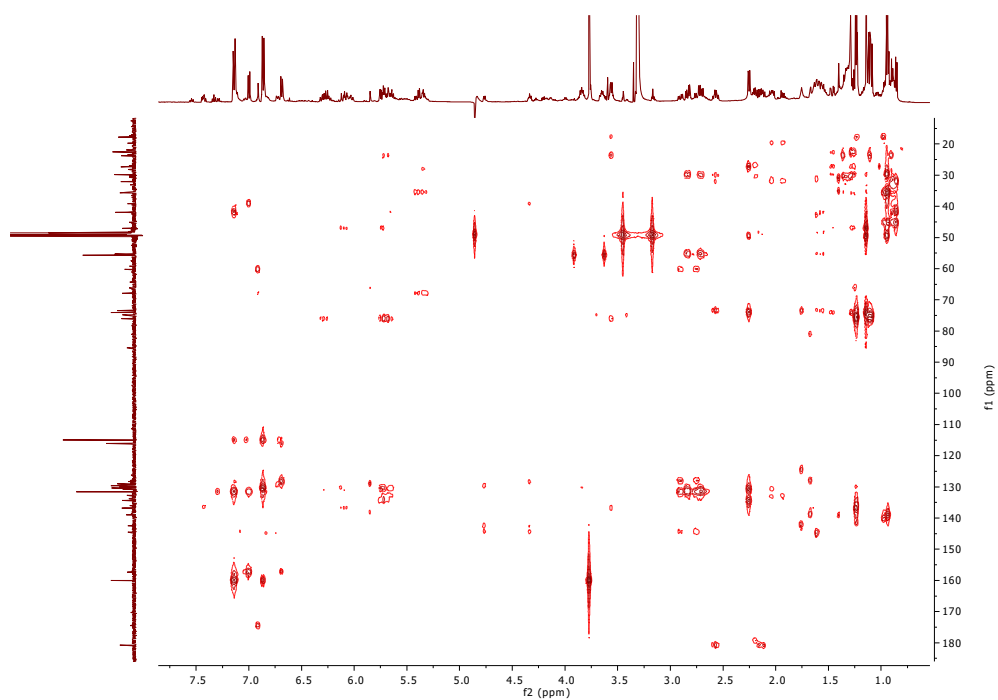


Figure 11.28: HMBC of compound **181** and **182**.

Compound **183** and **184**

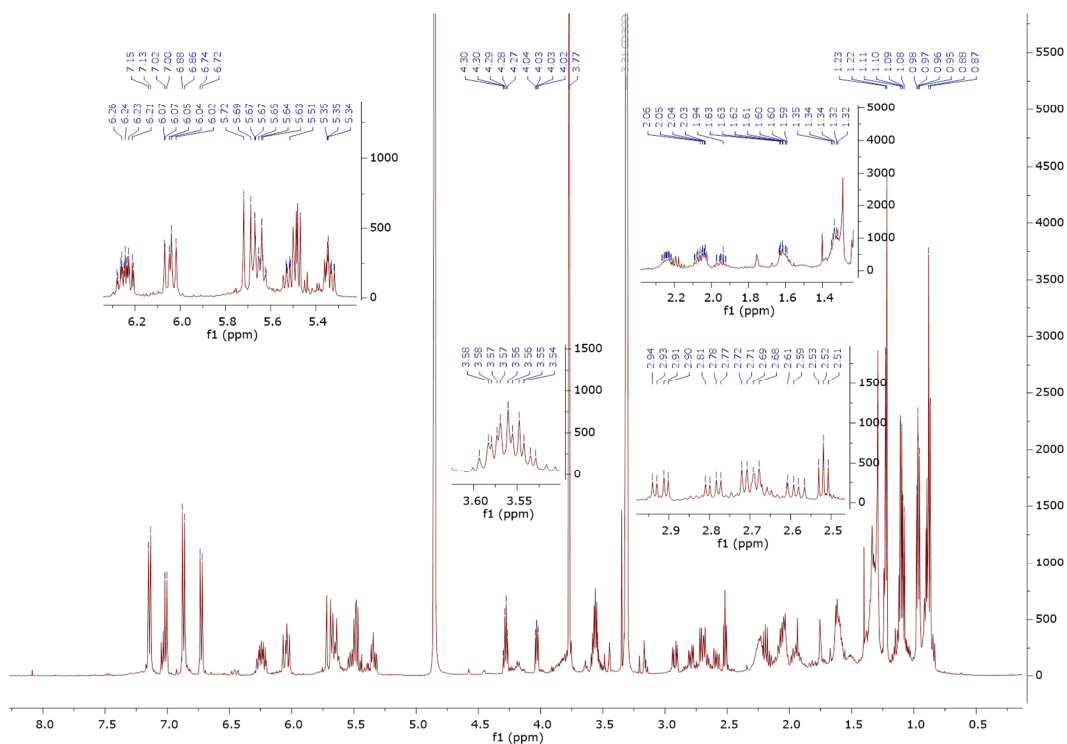


Figure 11.29: ^1H NMR of compound **183** and **184** in MeOD (500 MHz) referenced to MeOD.

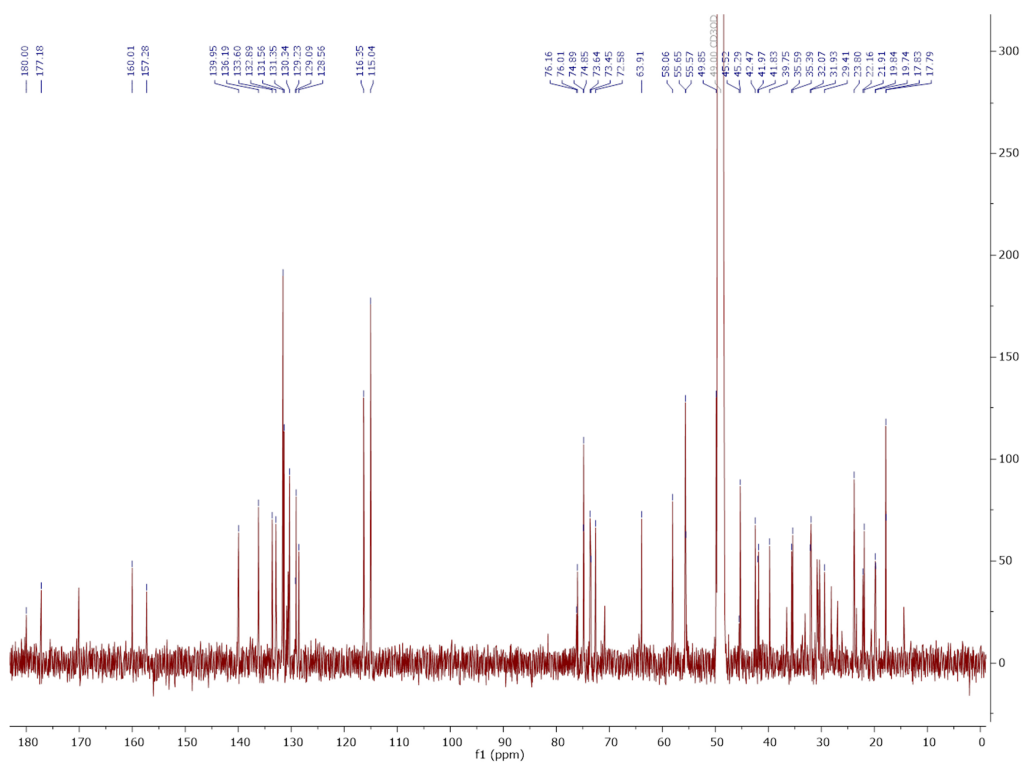


Figure 11.30: ^{13}C NMR of compound **183** and **184** in MeOD (500 MHz) referenced to MeOD.

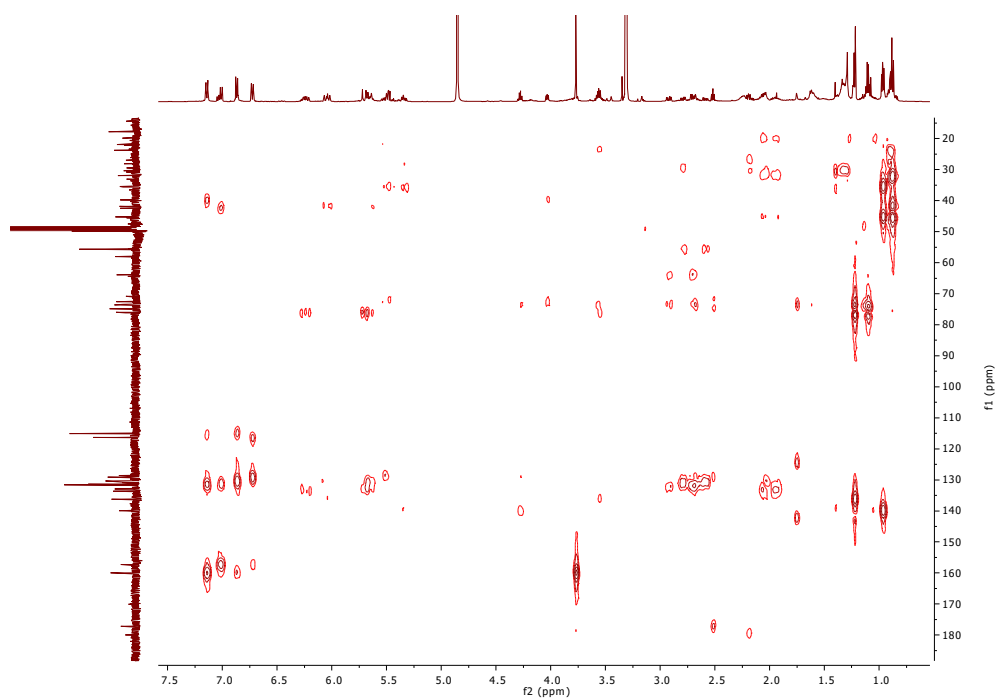


Figure 11.31: HMBC of compound **183** and **184**.

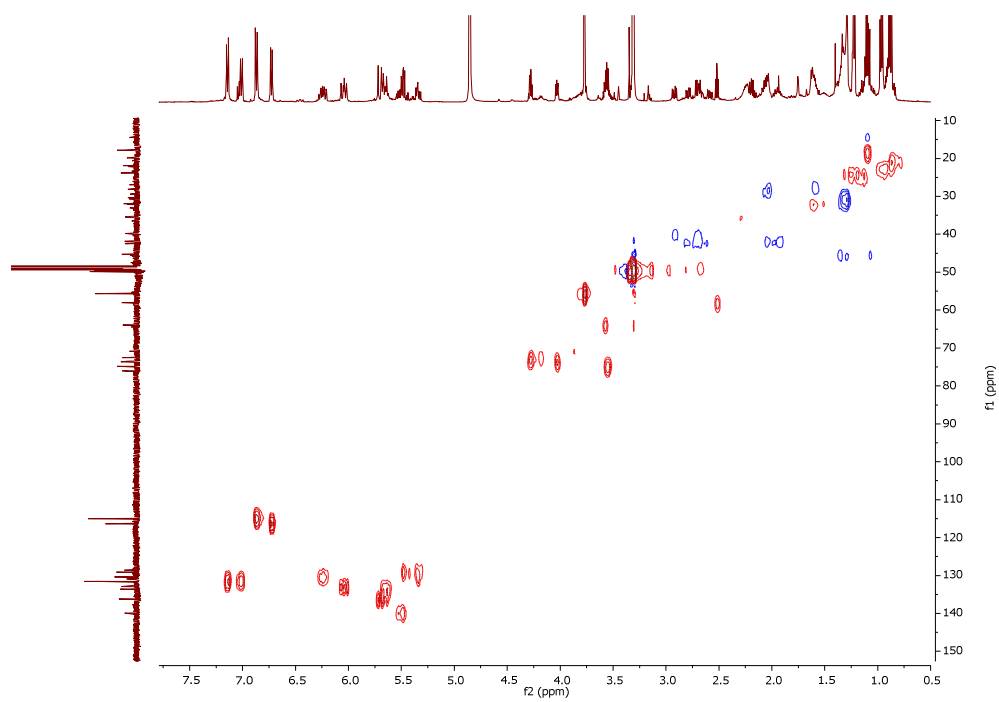


Figure 11.32: HSQC of compound **183** and **184**.

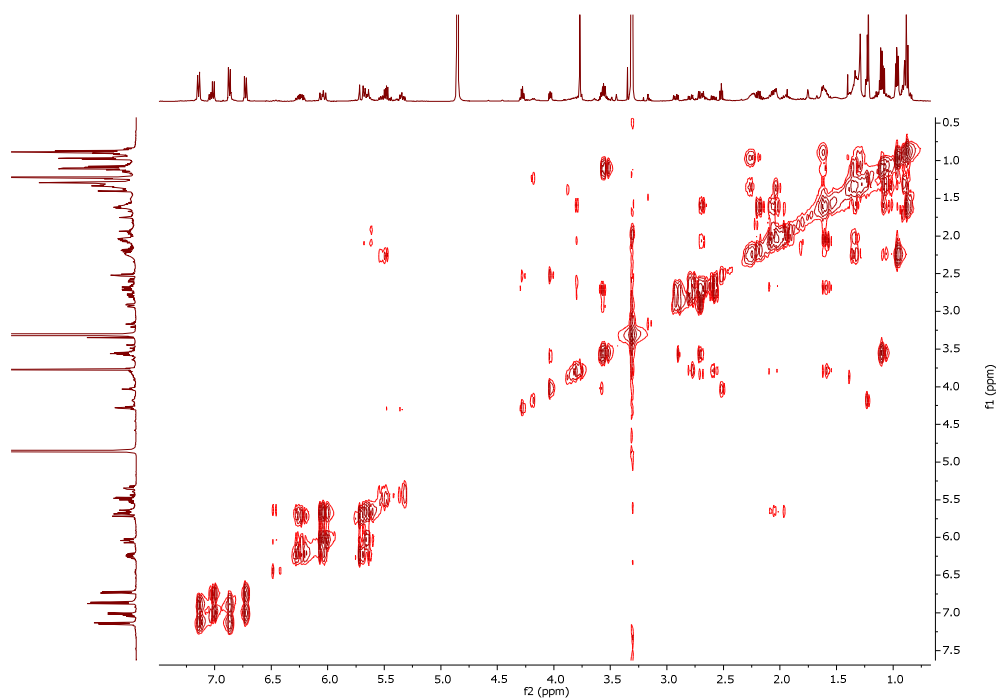


Figure 11.33: COSY of compound **183** and **184**.

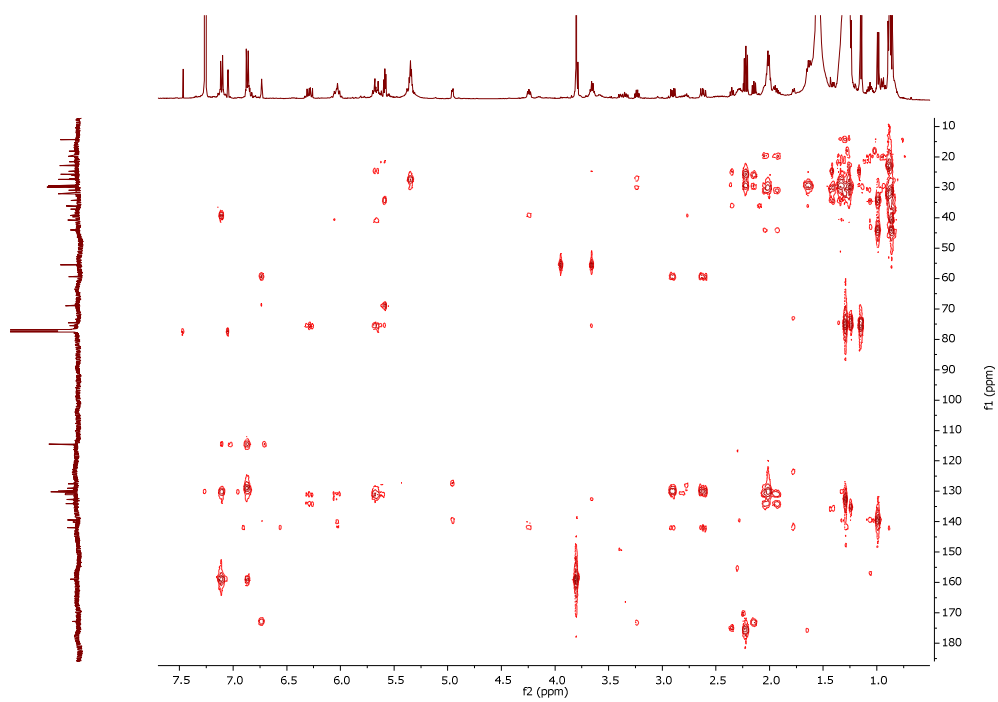


Figure 11.36: HMBC of compound 185.

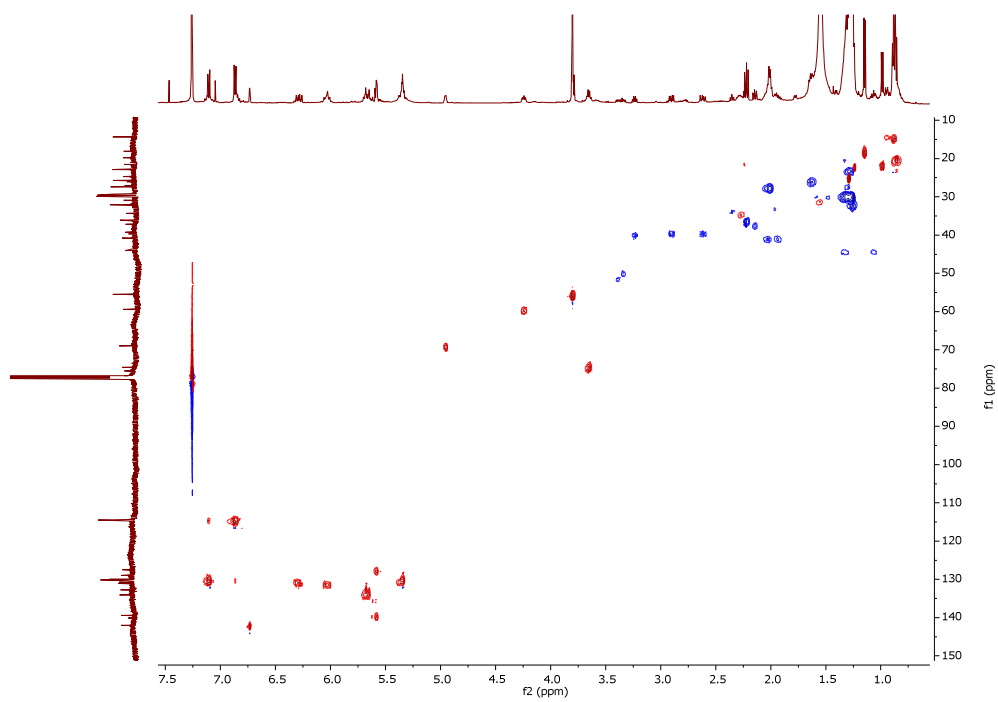
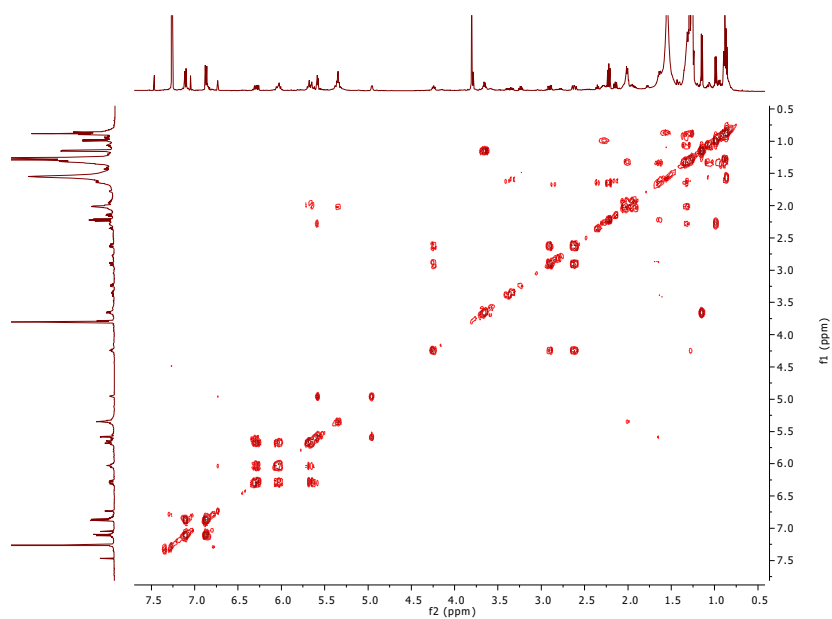
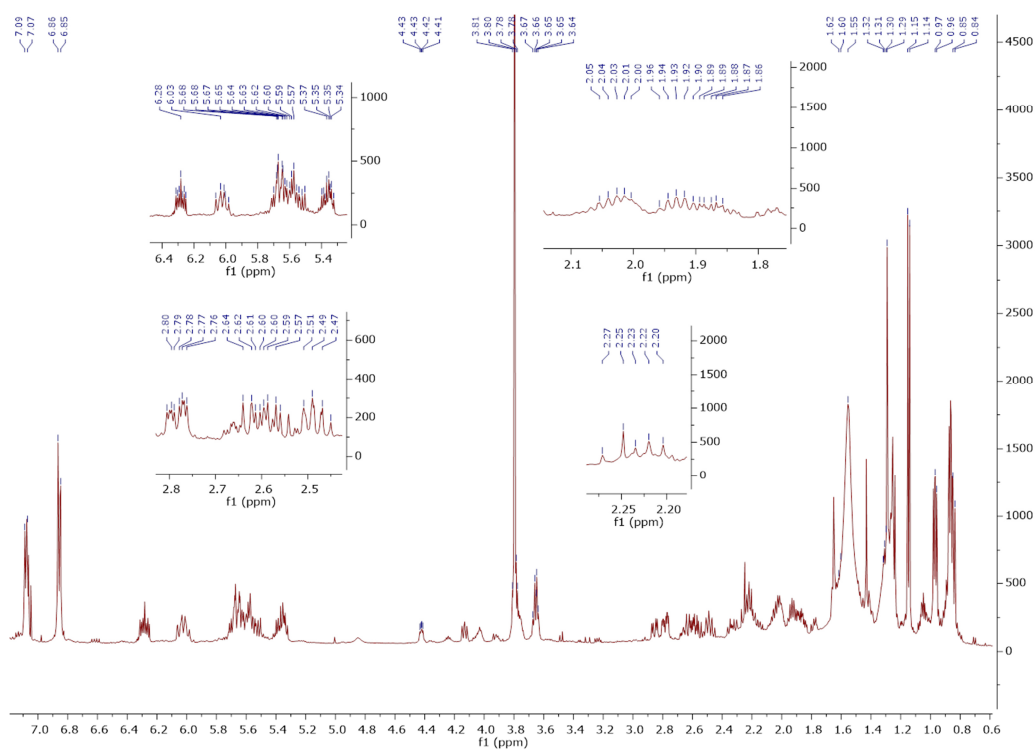


Figure 11.37: HSQC of compound 185.

Figure 11.38: COSY of compound **185**.Compound **186**Figure 11.39: ^1H NMR of compound **186** in CDCl_3 (500 MHz) referenced to CDCl_3 .

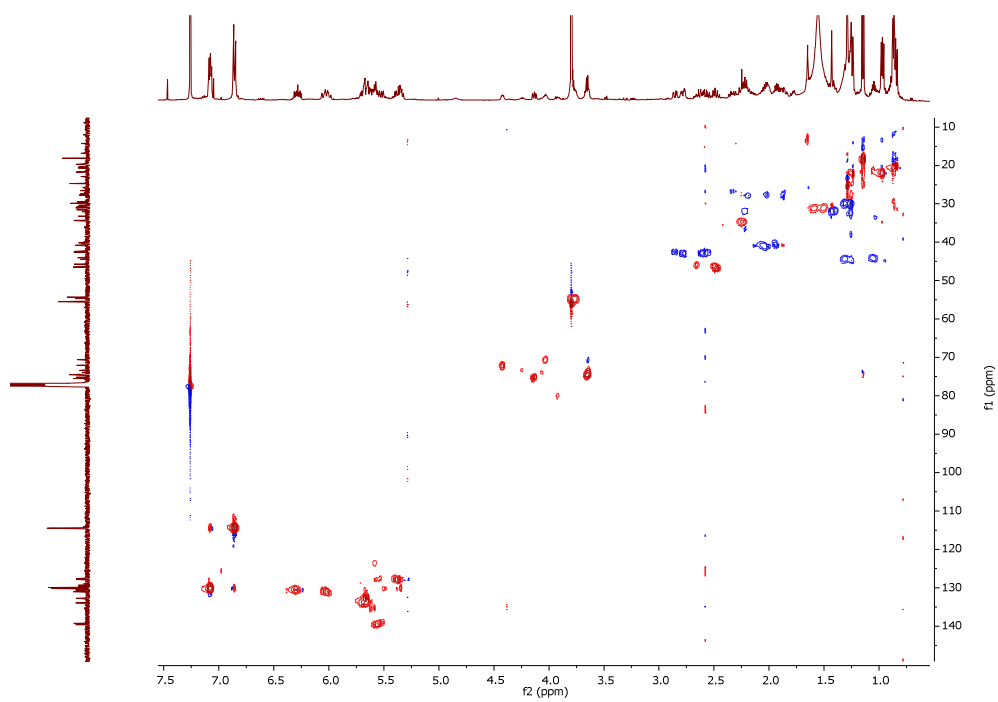


Figure 11.42: HSQC of compound 186.

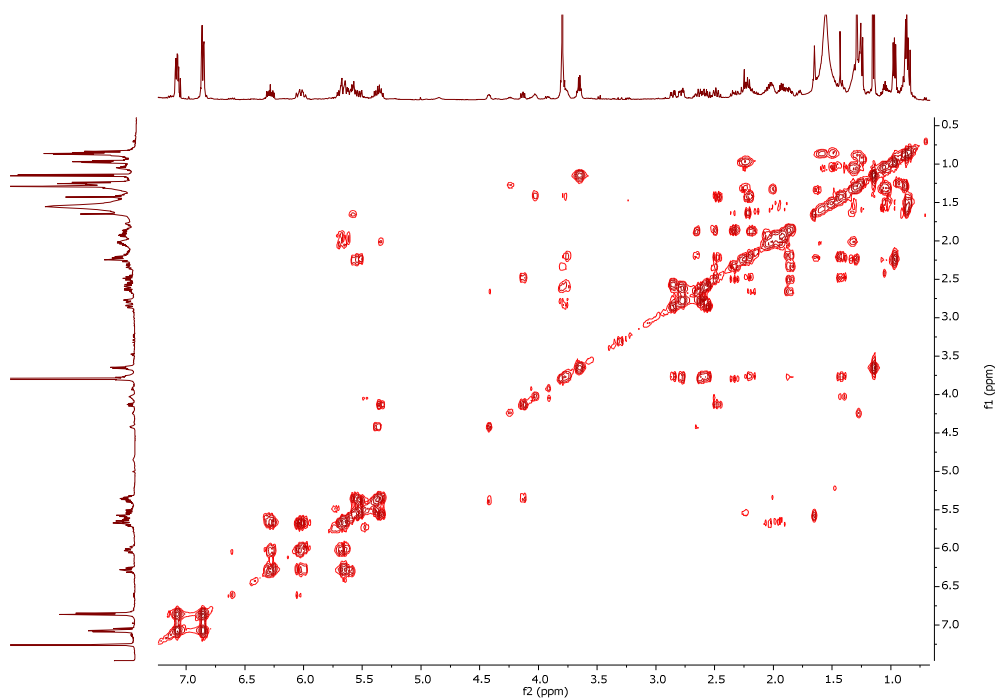


Figure 11.43: COSY of compound 186.

Curriculum Vitae

Apr 2016 – Dez 2019	PhD student in Organic Chemistry, Gottfried Wilhelm Leibniz Universität Hannover, Germany
Oct 2013 – Mar 2016	Master of Science , Life Science, Gottfried Wilhelm Leibniz Universität Hannover, Germany
Oct 2010 – Sep 2013	Bachelor of Science , Life Science, Gottfried Wilhelm Leibniz Universität Hannover, Germany
Jul 2002 – Jul 2009	Abitur , Gymnasium in Großburgwedel

List of Publications

1. V. Hantke, C. Wang, E. Skellam and R. J. Cox, *RSC Adv.*, **2019**, 9, 35797–35802
2. C. Wang, V. Hantke, R. J. Cox, and E. Skellam, *Organic Letters*, **2019**, 21, 4163–4167.
3. T. F. Tashjian, I. Lin, V. Belt, T. M. Cafarelli and V. G. Godoy, *Frontiers in Microbiology*, **2017**, 8, Article 288
4. B. Bonsch, V. Belt, C. Bartel, N. Duensing, M. Konziol, C.M. Lazarus, A.M. Bailey. T. J. Simpson and R. J. Cox, *Chemical Communications*, **2016**, 52, 6777-6780