

**Investigating the Biosynthesis of the Fungal
Metabolites Squalestatin S1, Strobilurin A and
SCH-642305**

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

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

**von
Karen Elisabeth Lebe, M. Sc.**

2020

Referent: Prof. Dr. Russell Cox

Korreferent: Prof. Dr. Andreas Kirschning

Tag der Promotion: 29. November 2019

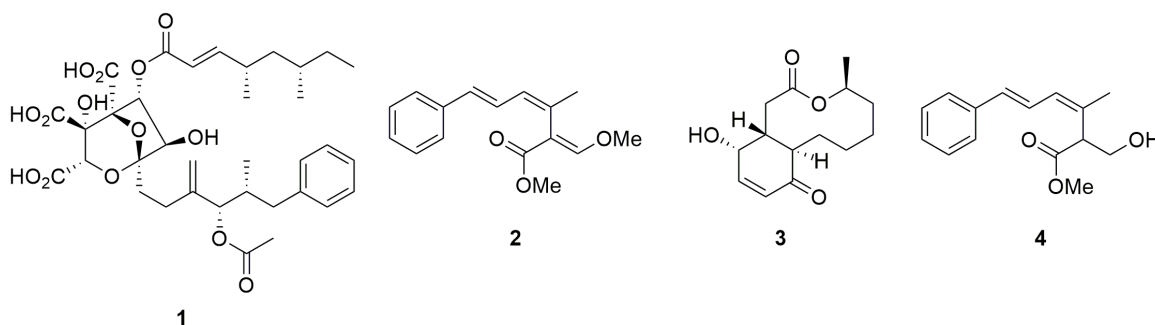
Abstract

The main focus of the presented work concentrated on understanding and engineering the biosynthesis of three polyketides produced by filamentous fungi: squalestatin S1 **1**, strobilurin A **2** and SCH-642305 **3**. In a combined genetic and chemical approach, the order and functions of genes and proteins involved in the three biosynthetic pathways were elucidated.

Using a combination of targeted gene knockout in the native organism and heterologous expression in *Aspergillus oryzae* (*A. oryzae*) three previously unknown oxygenases were identified from **1** biosynthetic gene cluster (BGC) and their role in the biosynthesis of **1** were determined. Two non-heme iron-dependent oxygenases were shown to catalyse a series of six consecutive oxidations to form the highly oxidised, bioactive core of squalestatins and an unusual copper-dependent oxygenase was found to introduce a hydroxyl required for later acetylation.

During strobilurin biosynthesis, an FAD dependent monooxygenase was identified and shown *in vitro* to catalyse an unusual oxidative rearrangement to form the core β -methoxyacrylate toxophore of this valuable class of agricultural fungicides. *In vivo* fungal expression studies revealed the biosynthetic pathway to bolineol **4** from the same BGC and uncovered the role of two *O*-methyltransferases.

In attempts to reveal each biosynthetic step in SCH biosynthesis, with special regard to formation of the 6-membered ring, the full BGC was coexpressed and re-established in *A. oryzae*. Only minor amounts of the final product were found due to a shunt pathway in *A. oryzae*. An early pathway intermediate was hydrolysed and oxidised and thus stopped the biosynthesis to **3**.



Keywords: natural products, biosynthesis, fungi, polyketides

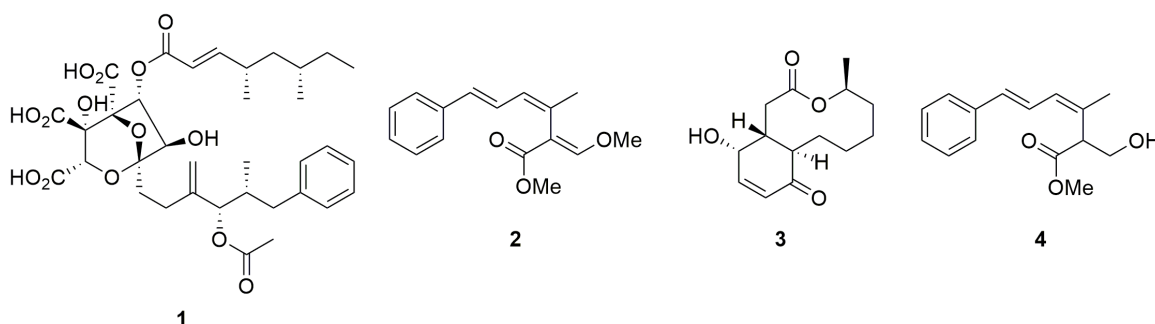
Zusammenfassung

Der Fokus der Doktorarbeit lag auf dem Verständnis und dem Engineering der Biosynthesewege von drei Polyketiden, die von filamentösen Pilzen produziert werden: Squalaestatin S1 (SQS1) **1**, strobilurin A **2** and SCH-642305 **3**. In einer Kombination aus chemischer und genetischer Analyse wurde die Funktion und Reihenfolge der an der Biosynthese beteiligten Proteine entschlüsselt.

Durch gezielte Gen-Knockouts und Heterologe Expression in *Aspergillus oryzae* (*A. oryzae*) wurden drei bisher unbekanntenen Oxygenasen aus dem SQS1 Gencluster identifiziert und ihre Rolle in der Biosynthese bestimmt. Es wurde gezeigt, dass zwei Nicht-Häm-Eisen-abhängige Oxygenasen eine Serie von sechs aufeinanderfolgenden Oxidationen katalysieren, die zu dem hochoxidierten, bioaktiven Zentrum von den Squalaestatinen führen. Eine ungewöhnliche Kupfer-abhängige Oxygenase ist verantwortlich für die Einführung einer Hydroxylgruppe, welche im weiteren Verlauf der Biosynthese acetyliert wird.

In der Strobilurin-Biosynthese wurde eine FAD-abhängige Monooxygenase identifiziert. *In vitro* Studien mit der Monooxygenase zeigten, dass sie eine ungewöhnliche oxidative Umlagerung zum β -Methoxyacrylat-Toxophor katalysiert, die dieser wertvollen Klasse von landwirtschaftlich genutzten Fungiziden ihre Wirkung verleiht. Durch *In-vivo*-Expressionsstudien in Pilzen konnte der weitere Biosyntheseweg zu Bolineol **4** aufgeklärt werden. Für die Biosynthese von Bolineol und **2** ist dasselbe Gencluster verantwortlich. Im Zuge dieser Arbeit konnte die Rolle von zwei O-Methyltransferasen in beiden Biosynthesewegen enthüllt werden.

In dem Versuch, jeden Biosyntheseschritt in der SCH-Biosynthese unter besonderer Berücksichtigung der Bildung des 6-gliedrigen Rings aufzudecken, wurden alle Gene des Biosyntheseclusters vollständig in *A. oryzae* coexprimiert und die Biosynthese rekonstruiert. Das Endprodukt konnte nur in geringen Mengen detektiert werden, da durch endogene Enzyme in *A. oryzae* ein frühes Biosyntheseintermediat hydrolysiert und oxidiert wurde, womit die Biosynthese zu **3** früh unterbrochen wurde.



Schlagwörter: Naturstoffe, Biosynthese, Pilze, Polyketide

Acknowledgement

Most importantly, I thank Prof Russell J. Cox for being the best supervisor I could have imagine. Thank you for all your assistance, support and trust throughout the last 3.5 years. I loved working on these three very interesting and challenging projects and I enjoyed discussing scientific ideas with you. I appreciate your honesty and direct way of speaking and a special thank for encouraging me to stay in science (I definitely will). I am grateful to be one of your PhD students. Thank you.

I would like to thank Prof Andreas Kirschning for being the co-referee and the chair of the examination board of my PhD thesis. You taught me the basic knowledge about natural products during my life science studies and inspired me to keep on working on polyketides. Thank you for your support throughout my scientific career and especially for giving me the opportunity to go to Cambridge. Prof Thomas Scheper I would like to thank for being the examiner of my PhD thesis. I have known you since the start of my life science studies and I am pleased that you agreed to examine my thesis.

I would like to thank all cooperation partners: the CeBiTEC Bielefeld, especially Prof Jörn Kalinowski and Dr Daniel Wibberg; Prof Rosie Bradshaw (Massey University, New Zealand); and the AK Braun, especially Dr Jennifer Senkler.

I say thank you to Dr Risa Nofiani and Dr Kate de Mattos-Shipley who worked with me on the strobilurin project and provided the genome data as well as expression vectors. Francesco Trenti, I would like to thank for sharing the Sch-642305 project.

Big thanks to the BMWZ media kitchen team, especially to Katja, who made almost everything possible in very short time. This work would not have been possible without the help of the analytical departments at the OCI, thank you all very much. I would like to give special thanks to Dr Jörg Fohrer and Dr Gerald Dräger for their help with NMR and mass related matters.

I would like to thank all past and current Cox group members, in particular Liz, Hao, Francesco, Olli, Haili, Sen, Slawik, Carsten, Erik, Lukas, Mary and all my students. Especially, I want to thank Steffen for his support with chemical questions and the OCI teaching lab.

I would like to give special thanks to my master's students Annika and Ole, who are talented students and contributed to a very nice working atmosphere. Very special thanks to my "OCI girls" Gesche, Raissa, Verena, Anna, Janina and Janna for their great support during the last years and for being real friends. I also would like to thank Egor, Duc, Helge and Christian from the Kirschning group. I would like to thank Michelangelo Marasco for help with the modelling and the good times and special thanks to Maja for always been my best science friend since the begin of my studies.

Finally, I would like to thank the most important people in my life: my parents and my sister Sunny for their great support throughout my studies and especially Alex who always believes in me and who has covered my back the last seven years. Thank you.

Submitting this thesis is only possible because of you.

Abbreviations and Units

μl	microlitre	h	hours
DEBS	6-deoxyerythromycin B synthase	Hyd	hydrolase
2MCDH	2-methyl-citrate dehydratase	HygB	hygromycin B
HMG--CoA	3-hydroxy-3-methyl-glutaryl coenzyme A	ITS	internal transcribed spacer
Ppant	4'-phosphopantetheine	Kan	kanamycin
6-dEB	6-desoxyerythronolide B	KSI	ketosteroid isomerases
6-MSA	6-methylsalicylic acid	KO	knockout
AU	absorption unit	LCMS	liquid chromatography-mass spectrometry
ACP	acyl carrier protein	L	litre
AT	acyl transferase domain	LNKS	lovastatin nonaketide synthase
ATP	adenosine triphosphate	MS	mass spectrometry
αKG	α-ketoglutarate	<i>m/z</i>	mass to charge ratio
aa	amino acid	mRNA	messenger RNA
APS	ammonium persulfate	mg	milligram
asRNA	anti sense RNA	ml	millilitre
<i>A. oryzae</i>	<i>Aspergillus oryzae</i>	min	minute
bp	base pair	M	molar
BLAST	basic local alignment search tool	SNAC	<i>N</i> -acetylcysteamine
KR	β-ketoreductase	nm	nanometer
KS	β-ketosynthase	NAD(P)(H)	nicotinamide adenine dinucleotide (phosphate)
BGC	biosynthetic gene cluster	αKG-NHFe	non-heme iron and α-ketoglutarate dependent oxygenases
Carb	carbenicillin	NHI	non-heme iron-dependent oxygenase
C-Met	carbon-methyltransferase	NHEJ	non-homologous end-joining
Cam	chloramphenicol	NR	non-reducing
CS	citrate synthase	NRPS	non-ribosomal peptide synthetase
CoA	coenzyme A	NMR	nuclear magnetic resonance
cDNA	complementary DNA	ORF	open reading frame
CDM	copper dependent monooxygenase	<i>O</i> -Met	oxygen-methyltransferase
PHM	copper dependent peptidylglycine α-hydroxylating monooxygenase	PR	partial reducing
PMO	copper dependent polysaccharide monooxygenases	ppm	parts per million
COSY	correlation spectroscopy	PAC	phage P1-derived artificial chromosomes
P450	cytochrome P450	PAL	Phenylalanine ammonia lyase
d	days	PEBP	phosphatidylethanolamine binding proteins
DH	dehydratase	PPTase	phosphopantetheinyl transferase
DNA	deoxyribonucleic acid	PAGE	polyacrylamide gel electrophoresis
DAD	diode array detector	PEG	polyethylene glycol
ddH ₂ O	double distilled H ₂ O	PKS	polyketide synthase
ESI	electron spray ionisation	PCR	polymerase chain reaction
ER	enoyl reductase	PAM	protospacer adjacent motif
<i>E. coli</i>	<i>Escherichia coli</i>	Q-TOF	quadrupole time-of-flight
EDTA	ethylenediaminetetraacetic acid	rpm	revolutions per minute
ELSD	evaporative light scattering detector	RNA	ribonucleic acid
FPP	farnesyl pyrophosphate	<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
FAS	fatty acid synthase	SAM	<i>S</i> -adenosyl methionine
FAD(H ₂)	flavin adenine dinucleotide	SCH	SCH-642305
FMO	flavin dependent monooxygenase	SDR	short chain dehydrogenase/reductase
FMN	flavin mononucleotide	SDS	sodium dodecyl sulphate
GOI	gene of interest	<i>sp.</i>	species
gDNA	genomic DNA	SQHKS	SQS hexaketide synthase
GMCO	glucose-methane-choline oxidase	SQTKS	SQS tetraketide synthase
Hz	hertz	SQS1	squalenstatin S1
HEX	heterologous expression	TEMED	tetramethylethylenediamine
HMBC	heteronuclear multiple bond correlation	TES	trace element solution
HSQC	heteronuclear single quantum correlation	TMS	trimethylsilyl
HPLC	high performance liquid chromatography	TAE	tris-acetate-EDTA
HRMS	high resolution mass spectrometry	TE buffer	tris-EDTA buffer
HR	highly reducing	UV	ultra violet
hrPKS	highly reducing polyketide synthase	wt	wildtype
HR	homologous recombination	ZA	zaragozic acid

Contents

Abstract	iii
Zusammenfassung	iv
Acknowledgement	v
Abbreviations and Units	vi
1 Introduction	1
1.1 Fungal natural products	1
1.1.1 Polyketide biosynthesis	3
1.1.2 Post-assembly line tailoring	9
1.2 Investigation of natural product biosynthesis in fungi	16
1.2.1 Isotopic labelling	16
1.2.2 Gene knockout	18
1.2.3 Fungal heterologous expression	21
1.3 Overall aims	26
2 Biosynthetic studies of Squalestatin S1	27
2.1 Introduction	27
2.1.1 Biological activity	28
2.1.2 Biosynthesis of SQS1 – <i>status quo</i>	29
2.1.3 Maleidrides biosynthesis	36
2.2 Project aims	39
2.3 Results – <i>In silico</i> analysis of putative oxygenases from SQS BGC	40
2.3.1 Putative non-heme iron-dependent oxygenases Mfr1 and Mfr2	40
2.3.2 Putative copper-dependent oxygenase Mfm1	41
2.4 Results – Analysis of squalestatin metabolites produced by wildtype fungus MF5453	43
2.5 Results – Knockout studies	45
2.5.1 Plasmid construction using yeast recombination based cloning	46
2.5.2 Bipartite method and transformation of MF5453	47
2.5.3 Knockout of non-heme iron-dependent monooxygenases Mfr1 and Mfr2	48
2.5.4 Knockout of copper dependent monooxygenase Mfm1	61

2.5.5	Knockout of acyltransferase Mfr4	75
2.5.6	Knockout of citrate synthase Mfr3	79
2.5.7	Knockout of NADP dependent dehydrogenase Mfm3	81
2.5.8	Discussion	82
2.6	Results – Heterologous expression of SQS genes in <i>A. oryzae</i> NSAR1	91
2.6.1	Yeast recombination, Gateway cloning and Transformation of <i>A. oryzae</i>	91
2.6.2	Expression of early SQS genes in <i>A. oryzae</i>	95
2.6.3	Coexpression of <i>mfpks2</i> , <i>mfm8</i> , <i>mfr3</i>	97
2.6.4	Expression of genes involved in oxidations during SQS1 biosynthesis	100
2.6.5	Discussion	104
2.7	Results – Putative SQS cluster in <i>Dothistroma septosporum</i>	106
2.8	Conclusion and Outlook	107
3	Biosynthetic studies of Strobilurin A	109
3.1	Introduction	109
3.1.1	Biosynthesis of Strobilurin A - <i>status quo</i>	110
3.1.2	Prestrobilurin production and benzoyl CoA biosynthesis in <i>S. tenacellus</i>	113
3.2	Project aims	115
3.3	Results – Heterologous expression of strobilurin A BGC in <i>A. oryzae</i> NSAR1	116
3.3.1	Vector construction and gene combinations for <i>A. oryzae</i> transformation	116
3.3.2	Heterologous expression of later pathway genes	117
3.3.3	Coexpression of <i>stPKS</i> , benzoyl-CoA producing genes and FAD monooxygenase encoding gene <i>str9</i>	120
3.3.4	Defining the minimal BGC for strobilurin A production	122
3.3.5	Individual role of O-methyltransferases Str2 and Str3	124
3.4	Results – <i>In vitro</i> activity tests with Str9	128
3.4.1	Prestrobilurin producing strain	129
3.4.2	Expression and purification of recombinant Str9	130
3.4.3	<i>In vitro</i> activity test with Str9	131
3.5	Discussion	133
3.5.1	FAD dependent oxidase Str9 accomplishes both the oxidation and the rearrangement	133
3.5.2	Role of the GMC oxidase Str4 and SDR Stl2 in regulation of strobilurin and bolineol production	136
3.5.3	Role of the O-MeTs Str2 and Str3	137
3.6	Conclusion and Outlook	140

4	Biosynthetic studies of SCH-642305	142
4.1	Introduction	142
4.1.1	Brefeldin A biosynthesis	142
4.1.2	Isolation of SCH and related compounds	145
4.1.3	Gene cluster and first biosynthetic studies of SCH	146
4.2	Project aims	150
4.3	Results – Re-establish SCH biosynthesis in heterologous host <i>A. oryzae</i> NSAR1	151
4.3.1	Biotransformation of SCH in expression host <i>A. oryzae</i> NSAR1	151
4.3.2	Yeast recombination, Gateway cloning and Transformation of <i>A. oryzae</i>	154
4.3.3	Expression of whole SCH cluster in <i>A. oryzae</i>	157
4.3.4	Expression of early genes in <i>A. oryzae</i>	168
4.3.5	Re-constructing the biosynthesis of SCH in <i>A. oryzae</i>	169
4.3.6	Expression of whole SCH cluster without hydrolase SchR1	170
4.4	Conclusion and Outlook	171
5	Overall Conclusion and Future Experiments	176
6	Experimental	178
6.1	Antibiotics, media, buffer and solutions	178
6.1.1	Media	178
6.1.2	Buffers and solutions	181
6.1.3	Antibiotics	181
6.2	Microbiology methods	182
6.2.1	Strains	182
6.2.2	<i>E. coli</i>	182
6.2.3	<i>S. cerevisiae</i>	183
6.2.4	<i>A. oryzae</i> NSAR1	183
6.2.5	Unidentified strain MF5453	184
6.3	Molecular biology methods	185
6.3.1	Polymerase Chain Reaction (PCR)	185
6.3.2	Oligonucleotides	185
6.3.3	Details for constructed vectors in this thesis	190
6.3.4	cDNA preparation	192
6.3.5	Agarose gel electrophoresis	192
6.3.6	DNA purification from gel or from PCR	192
6.3.7	Isolation of plasmid DNA from <i>E. coli</i>	192
6.3.8	Isolation of plasmid DNA from <i>S. cerevisiae</i>	193
6.3.9	Isolation of genomic DNA from fungus	193

6.3.10 Cloning procedure	193
6.4 Biochemistry methods	194
6.4.1 Recombinant protein expression and cell lysis	194
6.4.2 Protein purification by nickel affinity chromatography	194
6.4.3 SDS-PAGE	194
6.4.4 Enzyme activity assay with Str9	195
6.4.5 Bolineol and 133 toxicity test against <i>A. oryzae</i>	195
6.5 Chemical analysis	196
6.5.1 Liquid Chromatography Mass Spectrometry (LCMS)	196
6.5.2 Nuclear magnetic resonance (NMR) analysis	197
6.5.3 Extraction of MF5453 cultures	197
6.5.4 Extraction of <i>A. oryzae</i> cultures	198
6.5.5 Chemical synthesis	198
6.5.6 Isolation of strobilurin expression intermediates	200
Literature	203
7 Appendix	211
7.1 SQS1 – additional data	211
7.1.1 Genetic analysis of expression transformants	211
7.1.2 Expression of early genes	213
7.2 Strobilurin – additional data	214
7.2.1 Genetic analysis of heterologous expression transformants	214
7.2.2 ESI—MS/MS analysis of Str9	215
7.3 SCH – additional data	216
7.3.1 Expression vectors not used in this thesis	216
7.3.2 Genetic analysis of expression transformants	217
Curriculum Vitae	218
List of Publications	218

1 Introduction

In history, nature has always been the biggest source for medicine. Thousands of years ago our ancestors used plants and herbs to relieve pain or wrapped them around wounds for a better healing process. During history natural products from plants, bacteria, and fungi build an important basis for the development of new pharmaceuticals. The term "natural products" can be defined as compounds naturally produced by living organisms. A more narrow perspective defines them as small molecules relevant for primary and secondary metabolism. In contrast to primary metabolites, such as fatty acids or amino acids, secondary metabolites are not essential for survival of the organism but generally provide a benefit in competition with other organisms.^{1,2} Due to their wide range of complex structures and their high selectivity, natural products cover a broad spectrum of bioactivity and are widely used in the food, pharmaceutical and agricultural industries.³

Secondary metabolites are produced by most living organisms including animals, plants, bacteria and fungi among others. As the research project described in this thesis is about the biosynthesis of three secondary metabolites produced by fungi, the next section will focus on fungal natural products.

1.1 Fungal natural products

Fungi are key components of life on the earth. For humans they are often unimposing and it is easy to lose sight of their importance. That is why they are often referred to as the "Forgotten Kingdom".⁴ The basidiomycetes, together with the ascomycetes, make up the sub-kingdom *Dikarya* – often referred to as the "higher fungi".⁵ Strobilurin A **2**, one of the secondary metabolites discussed in this thesis, is produced by a basidiomycete: *Strobilurus tenacellus*.⁶ The two other compounds described in this work, squalestatin S1 (SQS1) **1** produced by *Phoma* sp.⁷ and SCH-642305 **3** produced by *Phomopsis* sp. CMU-LMA⁸, belong to the phylum of ascomycetes.

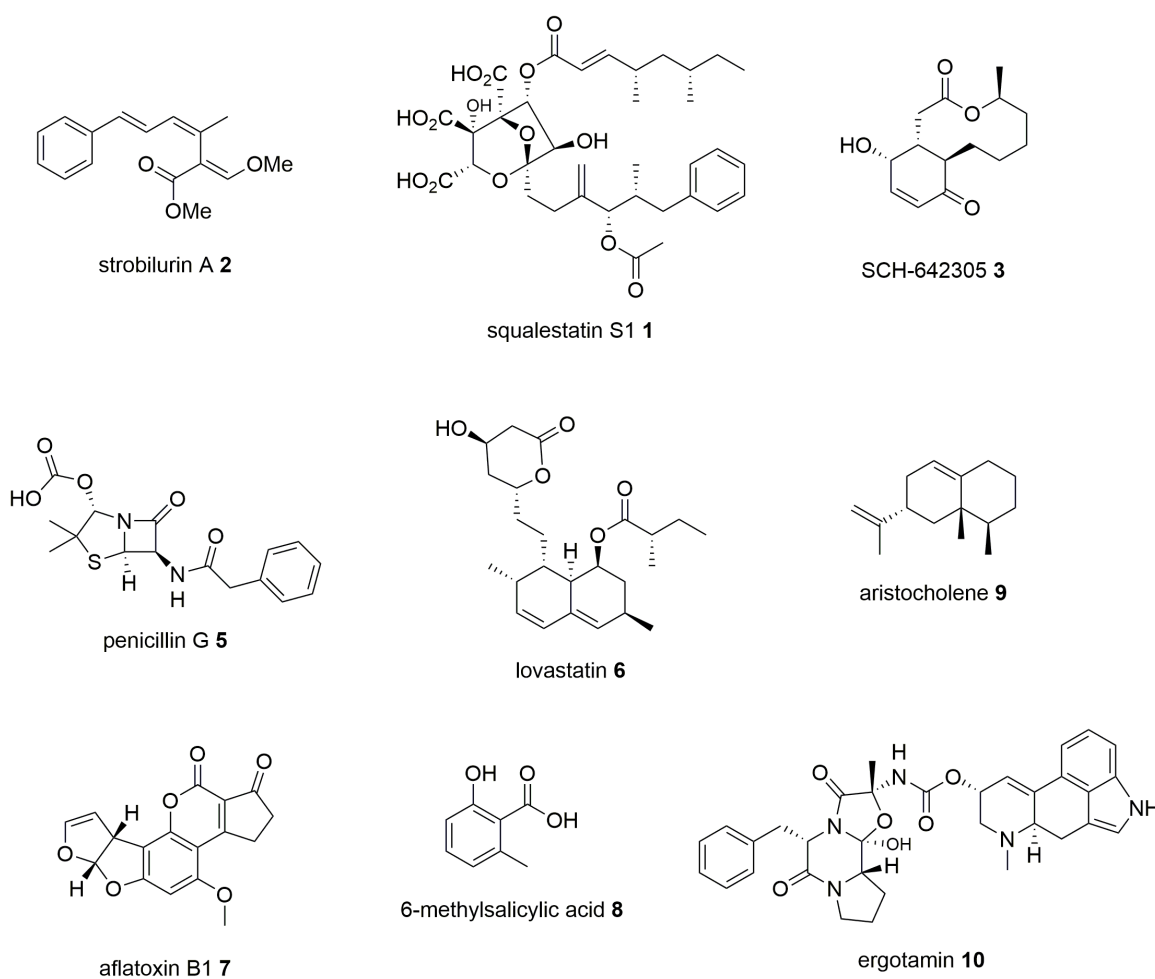


Figure 1.1 Examples of fungal natural products of different classes (polyketides, nonribosomal peptides, terpenes and alkaloids).

Throughout history, fungi have been used for preparation of alcoholic beverages, medication in traditional medicine and for cultural purposes, as drugs or food.⁹ In 1922 more than 200 fungal secondary metabolites were characterised under the leadership of Harold Raistrick.¹⁰ The most famous example of a fungal metabolite is penicillin G **5** (Fig. 1.1), which was discovered in 1929 by Alexander Fleming. The treatment with **5** saved many lives in World War II because of its antibiotic activity.²

The search for natural products has continued unabated and thousands of compounds have been subsequently isolated and characterised. Some of the most important natural products for the pharmaceutical industry derived from fungal sources.⁹ Most of them show antibacterial, antifungal or antitumour activity which makes them attractive for human interest as plant and animal toxins, growth hormones and pharmaceuticals. Besides, the enormous chemical complexity and diversity of these compounds, which arise from a limited number of precursors from primary metabolism, make this an important field of current research.¹¹

Fungal secondary metabolites are classified according to the biosynthetic pathway they derive from. Most secondary metabolites belong to the diverse class of polyketides which are made by polyketide synthases (PKS). Examples include squalestatin S1 **1**¹², inhibitor of squalene synthase, lovastatin **6**¹³, an inhibitor of 3-hydroxy-3-methyl-glutaryl (HMG) CoA reductase (both cholesterol level-lowering agents)¹⁴, strobilurin A **2**¹⁵, an agricultural fungicide, SCH-642305 **3**¹⁶, a bacterial DNA primase inhibitor, aflatoxin B1 **7**¹⁷, a carcinogenic agent, and 6-methylsalicylic acid (6-MSA) **8**.¹⁸ Other natural products, like **5**, are made by nonribosomal peptide synthetases (NRPSs) and therefore belong to the class of nonribosomal peptides. Terpenes, like aristocholene **9**¹⁹ and alkaloids, such as ergotamine **10**²⁰, also belong to the group of secondary metabolites.

1.1.1 Polyketide biosynthesis

Polyketides show an extremely large structural diversity and build up one of the major classes of fungal metabolites. The scaffold can vary from small compounds, such as 6-MSA **8**, to large molecules such as squalestatin S1 **1**. Nevertheless, they share a common pattern of biosynthesis which is similar to the biosynthesis of fatty acids. In 1955, Birch and coworkers realised that polyketides and fatty acids share similarities in *de novo* biosynthesis by discovering that polyketides are generated from the same simple building blocks (acetate units) by repeated condensation reactions, as fatty acids.²¹ Publications addressing the origin of fatty acids and polyketides have been published much earlier in the Journal of the Chemical Society in 1907.^{22,23}

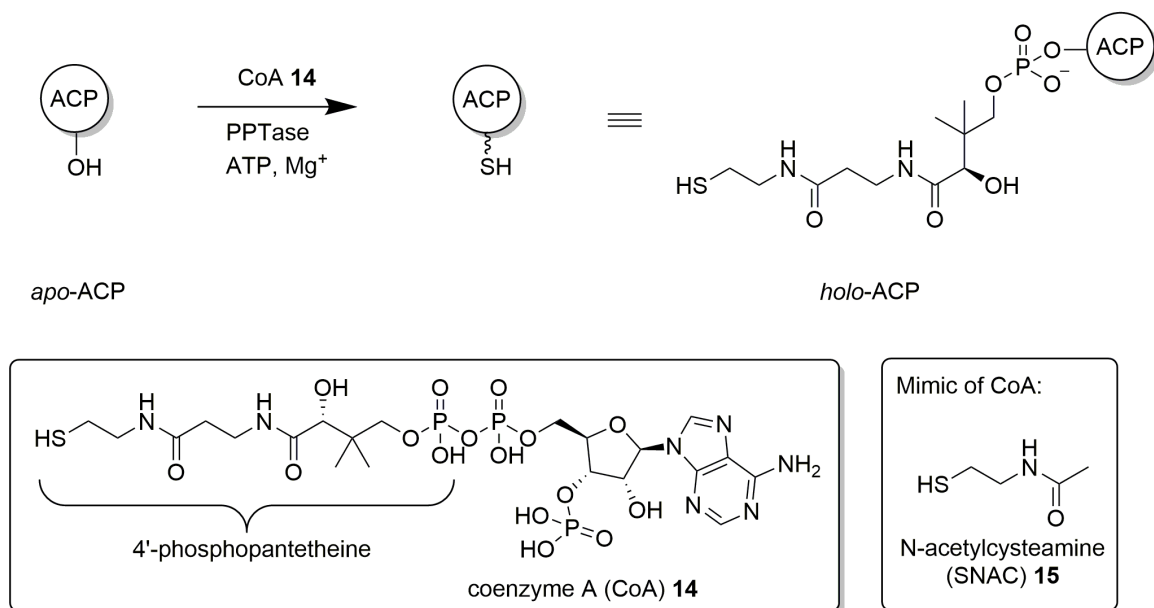
Fatty acid biosynthesis is a primary metabolic pathway. The chemical logic, the enzymatic machinery and the building blocks appear to have been recruited by the secondary pathways of polyketide assembly in living organisms.²⁴

Logic of fatty acid biosynthetic machinery

Fatty acids are primary metabolites that are synthesised by complex biosynthetic machinery, called fatty acid synthases (FAS). Polyketides are secondary metabolites which are built by polyketide synthases (PKS). The general logic for the biosynthetic process is the same for both and can be divided in three phases: loading; chain extension and β -processing (Scheme 1.2).

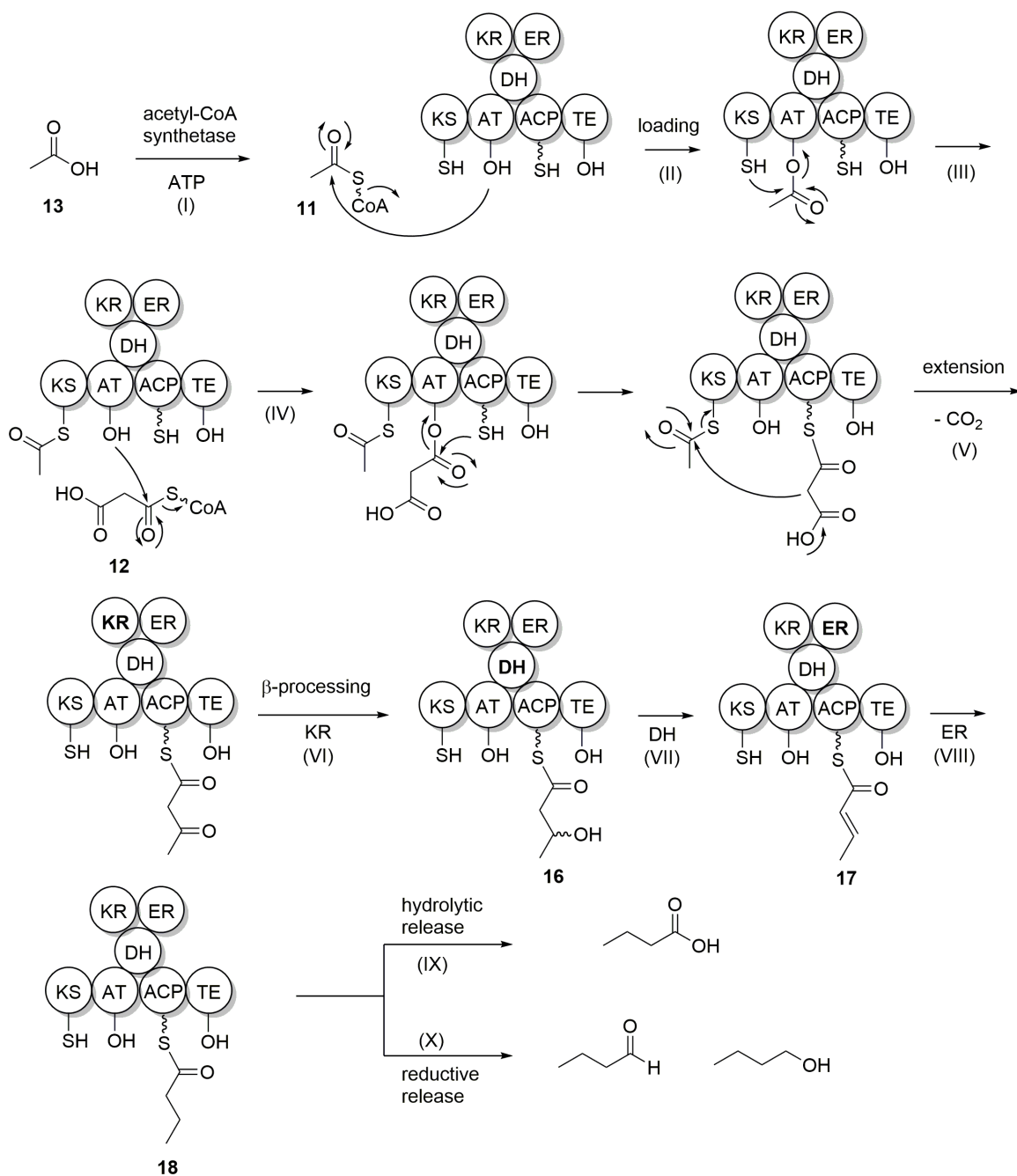
The starter unit for fatty acid and polyketide biosynthesis is usually acetyl-CoA **11** and in most cases malonyl-CoA **12** is used as elongation unit. Acetate **13** is activated as coenzyme A (CoA) thioester by the acetyl-CoA synthetase enzyme with consumption of adenosine triphosphate (ATP, Scheme 1.2, step I).²⁵

The first step in the biosynthesis of fatty acids and polyketides is the loading process, where activated acetyl-CoA **11** is transferred onto the acyl carrier protein (ACP) of the FAS/PKS by the acyl transferase (AT) (Scheme 1.2, step II). The *apo*-ACP must be activated *via* post-translational modification by a phosphopantetheinyl transferase (PPTase). The PPTase installs the CoA derived 4'-phosphopantetheine arm (PPant), forming *holo*-ACP, which enables the bound β -ketoester to access all active sites of the FAS protein (Scheme 1.1).²⁵ In biochemical studies, ACP-bound substrates can sometimes be mimicked by synthesised *N*-acetylcysteamine thioesters (SNAC) **15**, as those possess the same structural feature compared to the terminal part of **14** (Scheme 1.1).^{26,27}



Scheme 1.1 Conversion of *apo*-ACP to *holo*-ACP with consumption of CoA **14** and ATP by a phosphopantetheinyl transferase (PPTase).

The acetate is further transferred to the thiol of the β -ketosynthase (KS), leaving space at the AT for loading extender unit malonyl-CoA **12** to the ACP (Scheme 1.2, steps III, IV). The malonyl thiolester can undergo facile enzyme-mediated decarboxylation to lose CO_2 and furnish the acetyl-CoA carbanion in the biosynthetic direction (Scheme 1.2, step V). This decarboxylated thiolclaisen condensation is the key reaction type for fatty acid and all subclasses of polyketide chain elongations and is catalysed by the KS.²⁸



Scheme 1.2 Biosynthetic steps in fatty acid and polyketide biosynthesis, including loading of the starter and extender units, chain elongation and β -processing.

Further modifications of the β -carbonyl (β -processing) are catalysed by β -keto reductase (KR), which reduce the β -group with cosubstrate NAD(P)H to the β -OH-acyl-S-ACP **16**. Subsequent dehydration by dehydratase (DH), which abstracts one of the acidic α -hydrogens, results in an unsaturated thioester **17**. Because of the resulting conjugation, the terminus of the olefin is electrophilic and can accept a hydride from NADPH in the active site of the enoylreductase (ER) yielding a fully saturated thioester **18** (Scheme 1.2, steps VI, VII, VIII). After further elongation rounds by C_2 units

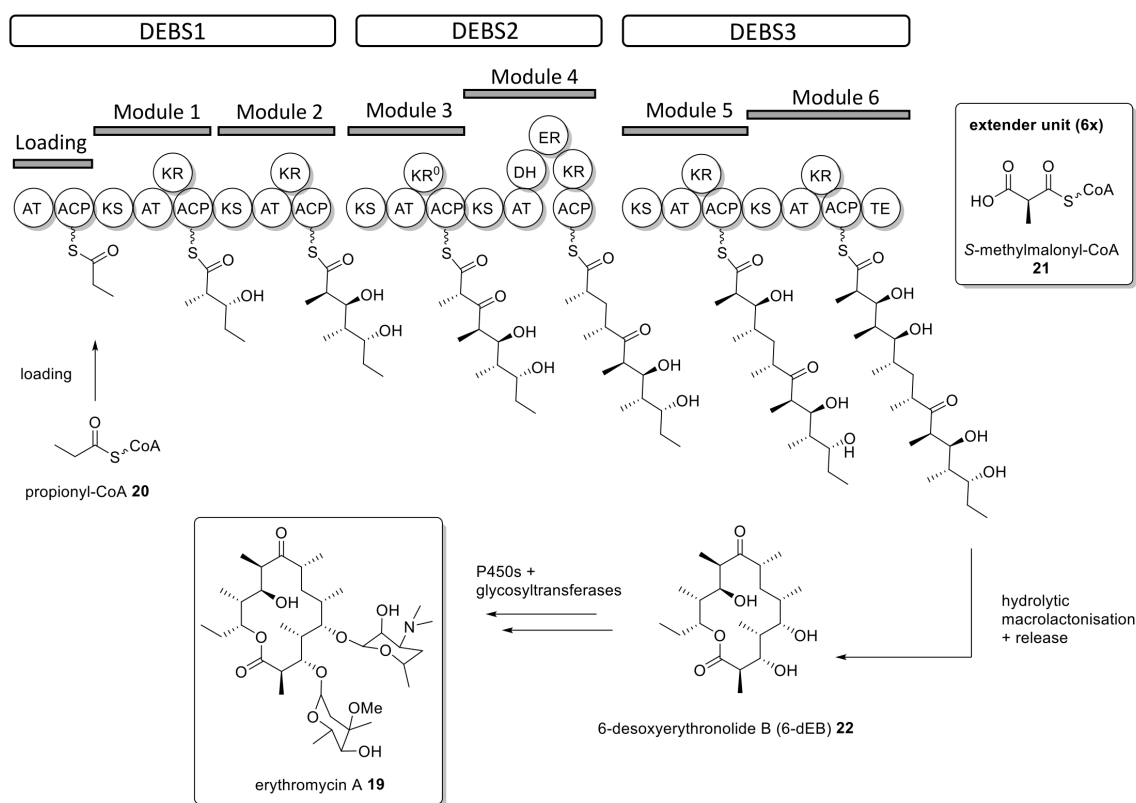
the furnished chain is released by the thioesterase (TE) domain. In the case of fatty acid biosynthesis the product is a carboxylic acid (hydrolytic release).^{24,25,28} In polyketide biosynthesis the release can either be hydrolytic or reductive, resulting in aldehydes or alcohols (Scheme 1.2, steps IX, X).

PKS use the same set and logic of chemical reactions as FAS (Scheme 1.2), but the key difference is that of programming.²⁹ Both biosynthetic pathways control the chain length by the number of extensions, but PKS can additionally control starter and extender unit selection and the extent of reduction and elimination during each β -processing cycle.^{2,30,31} In addition fungal PKS often have a functional C-MeT domain which can methylate the β -carbon at the start of the β -processing cycle.²⁹

Multidomain protein architecture

According to differences in structure, FAS proteins are divided into type I and II systems. Type I FAS are large multifunctional proteins found in animals and fungi, containing all domain activities for fatty acid biosynthesis in one megasynthase protein.³² On the contrary, type II FAS is characterised by a set of dissociated monofunctional proteins, which putatively assemble to non-covalent complexes (found in bacteria and plants). FAS consist of a single module that is used iteratively.³³

PKS are classified in three types according to enzyme organisation.³⁴ Type I PKS are multifunctional enzymes that can be divided in modular type I PKS, which occur mostly in bacteria and iterative type I PKS, which are predominantly found in fungi.^{29,35} Each single module of a modular type I PKS catalyses one cycle of polyketide chain elongation, such as the 6-deoxyerythromycin B synthase (DEBS) in the biosynthesis of erythromycin A **19** (Scheme 1.3).³⁶ DEBS comprises three proteins (DEBS 1-3), which contain seven modules with 28 domains working in line. Each module is constituted of ACP and KS domains for loading and chain extension. Further a combination of β -processing domains is included in a module: KR; DH and ER. Propionyl-CoA **20** is the starter unit for module 1. The chain is further elongated with 6 \times S-methylmalonyl-CoA **21** yielding a C₁₅-chain.³⁷ The polyketide chain is released by macrolactonisation catalysed by the *cis*-acting TE domain to furnish 6-desoxyerythronolide B (6-dEB) **22** and subsequent *post*-elongation tailoring events (oxidations by P450 cytochromes and glycosyltransferases) give **19** (Scheme 1.3).³⁷



Scheme 1.3 Modular organisation of DEBS megasynthase and biosynthetic steps to erythromycin **19**.

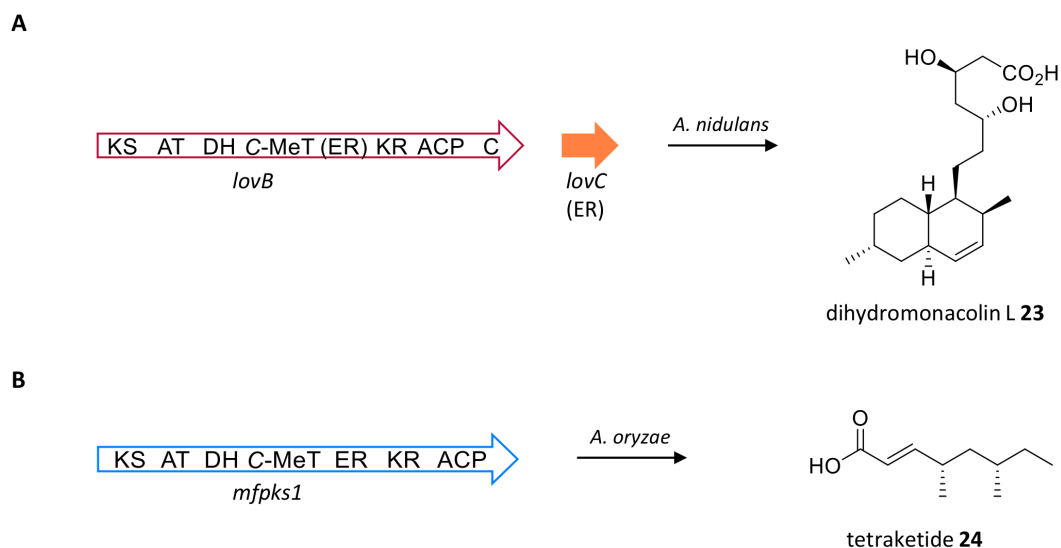
Type II PKS systems are characterised by sets of monofunctional proteins and work in an iterative fashion to build up aromatic polyketides. Furthermore, ACP-independent type III PKS are reported. Type III systems are also known as chalcone synthase-like PKS. These homodimeric enzymes act directly on acyl-CoA substrate.³⁴ All three types of PKS systems are found in bacteria. Type I PKS are also known from fungi and type III PKS are also found in plants and fungi.²⁹

Fungal PKS

In fungi, the predominant type of PKS are iterative type I megasynthases, which synthesise complex natural products using a single set of domains (one module) in a highly programmed, iterative fashion.³⁸ They can be functionally and phylogenetically organised into three main classes: nonreducing (NR), highly reducing (HR), and partial-reducing (PR) PKS.^{29,38,39} In addition to the catalytic domains described for fatty acid biosynthesis (Scheme 1.2), an iterative fungal PKS can also include further domains, *e.g.* for methylation of the α -carbon by a C-methyl transferase (C-MeT) using a methyl group from *S*-adenosylmethionine (SAM).³⁸ Polyketide synthases discussed in this thesis all belong to the class of iterative fungal type I hrPKS, including squalestatin hexaketide synthase (SQHKS), StPKS (strobilurin pathway) and SchPKS (SCH-642305 pathway). A prominent example of

iterative fungal hrPKSs is the well studied LovB (lovastatin nonaketide synthase, LNKS) of lovastatin **6** biosynthesis⁴⁰ or the SQS tetraketide synthase (SQTKS) of **1** biosynthesis (Scheme 1.4).⁴¹

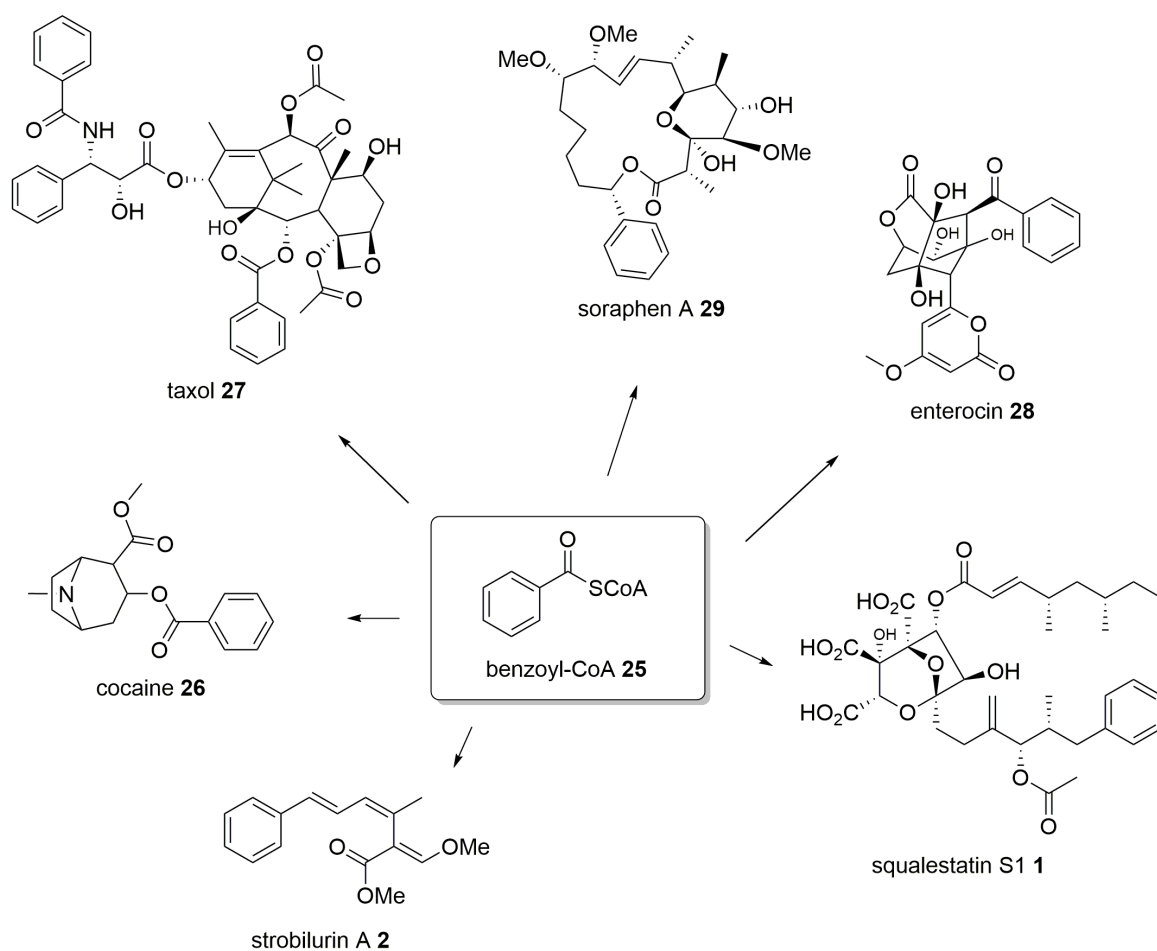
LNKS (encoded by *lovB*) was shown to possess a disfunctional ER domain within the enzyme, but coexpression experiments with an enoyl reductase encoding gene (*lovC*), located downstream of *lovB* yielded the expected PKS product dihydromonacolin L **23** (Scheme 1.4 A).⁴² Thus LovC is a *trans*-acting ER. Expression of the gene encoding SQTKS in *A. oryzae* yielded the tetraketide **24** showing it has a functional ER (Scheme 1.4 B).⁴¹



Scheme 1.4 Two examples of prominent iterative fungal hrPKS. **A**, coexpression of *lovB* and *lovC* in *A. nidulans*; **B**, domain organisation of SQTKS and expression product **24**.

The structural diversity of polyketides and accompanying diverse biological activities are the result of programmed tailoring steps catalysed by the single domains of iterative hrPKS, which are selectively used during different iterations of chain elongation, *e.g.* in lovastatin biosynthesis.⁴⁰ Furthermore, the choice of a different starter unit, such as benzoyl-CoA **25**, yields more diversity.

The use of benzoate starter unit by fungal PKS is relatively uncommon in polyketide biosynthesis. However, **25** serves as a metabolite in the biosynthesis of important plant metabolites including cocaine **26**⁴³ and taxol **27**⁴⁴. There are also bacterial metabolites, such as enterocin **28** (*Streptomyces maritimus*)⁴⁵ and soraphen A **29** (*Sorangium cellulosum*)⁴⁶, which use benzoyl CoA in their biosynthesis. In addition, two of the three fungal polyketides (squalestatin S1 **1** and strobilurin A **2**) discussed in this thesis use a benzoate starter unit (Scheme 1.5).^{7,47,48}



Scheme 1.5 Examples of natural products consisting of benzoyl-CoA building blocks.

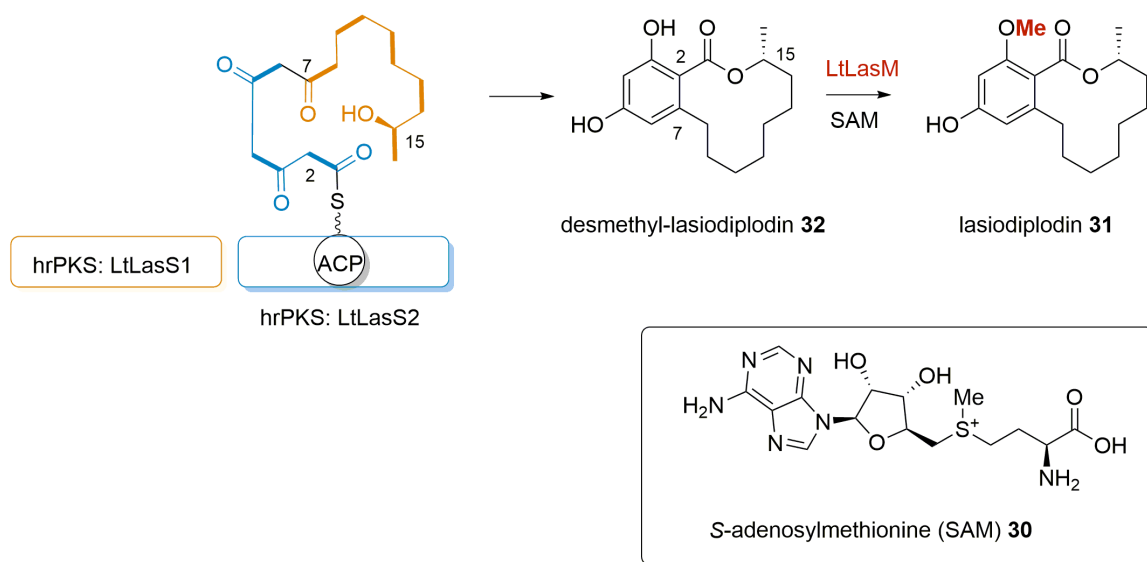
The diversity of fungal polyketides derive not only from the programmed iterative cycles of the PKS, but most importantly by post-polyketide-synthesis steps, which are discussed in section 1.1.2.²

1.1.2 Post-assembly line tailoring

Tailoring steps occur after release of the polyketide chain. The released products often do not possess the biological activities of the matured end products and need to undergo further decoration steps carried out by dedicated post-assembly line tailoring enzymes.⁴⁹

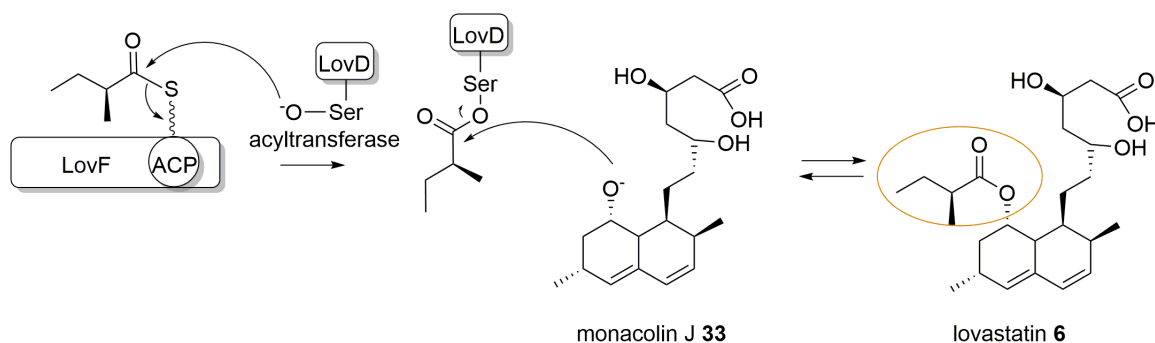
The most commonly found post-PKS modifications are catalysed by oxidoreductases. This very broad group of enzymes covers oxygenases, peroxidases, reductases, and dehydrogenases.⁵⁰ They can for example catalyse the insertion of one or two oxygen atoms from atmospheric O_2 , or the hydride transfer resulting in reduction or oxidation of alcohols to aldehydes and ketones.²⁸

Also methylations at oxygen, nitrogen or carbon centres by *S*-adenosyl methionine (SAM, **30**) dependent methyltransferases are common features of tailoring.⁵⁰ An example is the *O*-methyltransferase (*O*-MeT) LtLasM from the biosynthetic gene cluster producing the phytotoxic fungal polyketide lasiodiplodin **31**, which was shown to be responsible for the methylation of the C-3 phenolic alcohol of desmethyl-lasiodiplodin **32** to lasiodiplodin **31** by coexpression of LtLasM with two hrPKS from the BGC: LtLasS1 and LtLasS2, in yeast (Scheme 1.6).⁵¹



Scheme 1.6 Biosynthetic steps to lasiodiplodin **31** with a final *O*-methylation step catalysed by *O*-MeT LtLasM.

Often acyltransferases play a role in tailoring of the polyketide chain, such as during the biosynthesis of lovastatin **6**, where an acyltransferase LovD catalyses the transfer of a diketide to monacolin J **33** in the final step of the biosynthesis (Scheme 1.7).⁵² A similar transfer was also reported from squalestatin **1** biosynthesis, where a tetraketide chain was loaded to the core of **1** as the last step of the biosynthesis (Section 2).⁵³



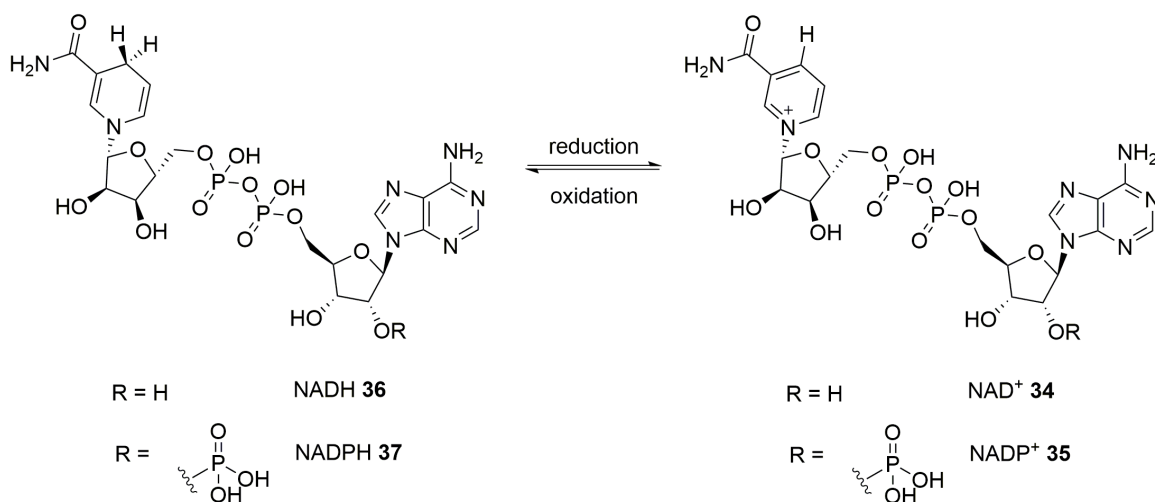
Scheme 1.7 Final step in lovastatin **6** biosynthesis catalysed by acyltransferase LovD.

Activation of atmospheric oxygen and oxygenases

As atmospheric oxygen is a diradical it is inert towards spin paired organic compounds and therefore not accessible by biological systems. Enzymes which are able to activate atmospheric O₂ are divided into oxidases and oxygenases, which incorporate oxygen atoms from dioxygen into the product(s).⁵⁴ The latter are again divided into monooxygenases that catalyse the incorporation of one atom of oxygen into the product and one atom to H₂O, while dioxygenases incorporate both atoms of oxygen into the product(s). Nature has evolved two strategies to reduce oxygen by one or two-electrons transfer. The first strategy requires active transition metal ions such as iron or copper and the second option involves the organic cofactor flavin adenine dinucleotide (FAD/FADH₂).⁵⁴ Enzyme classes catalysing these reactions are iron dependent oxygenases: heme-dependent cytochrome P450 enzymes/ non-heme iron-dependent oxygenases; and FAD dependent monooxygenases.

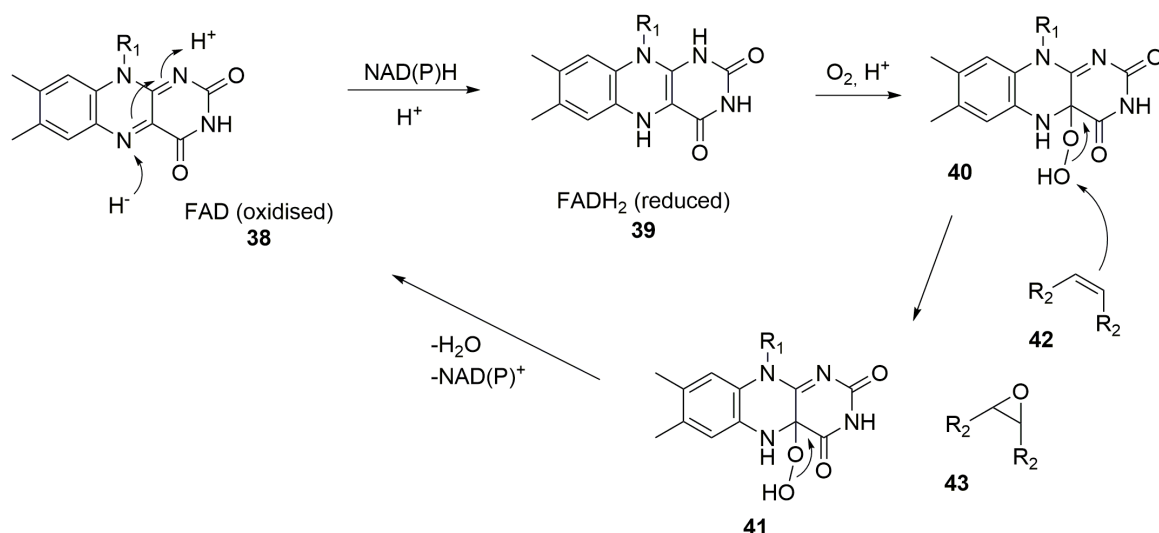
FAD dependent monooxygenases

FAD dependent enzymes, also called flavoenzymes, are redox enzymes, which utilize the covalently enzyme-bound cofactor FAD/FADH₂. Often nicotinamide adenine dinucleotide (phosphate) (NAD(P)⁺ **34(35)**/NAD(P)H **36(37)**) is also bound to the enzyme in order to reduce FAD **38** to FADH₂ **39**. NAD(P)H is in this case the reducing agent *via* two electrons transfer by hydride (H⁻) (Scheme 1.8).⁵⁵



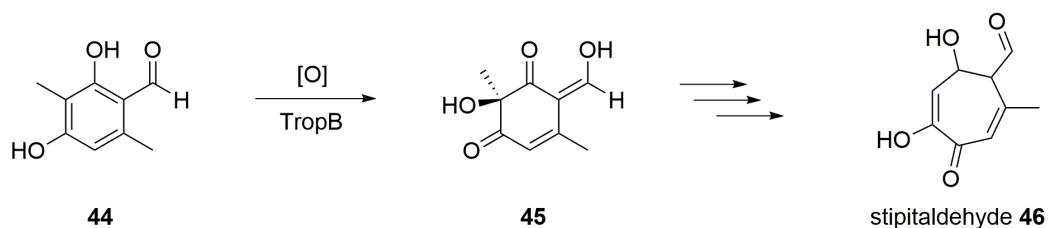
Scheme 1.8 Reduction of NADPH/NADH.

Upon binding of O₂ the reduced form **39** forms the reactive hydroperoxyflavin **40** species, which is able to transfer oxygen either by nucleophilic attack of the substrate (Scheme 1.9), or by a nucleophilic attack to the substrate. In the next step a water molecule is eliminated from **41** and the resulting **38** is further reduced to FADH₂ **39** by consuming reduced NAD(P)H (Scheme 1.9).⁵⁶



Scheme 1.9 Oxygen activation by FADH₂ and regeneration and epoxidation of an olefin **42** to the epoxide **43**.

A well known example for an FAD dependent monooxygenase (FMO) occurs in fungal tropolone biosynthesis. TropB is an enzyme shown to catalyse the step from 3-methylorcinaldehyde **44** to the hydroxylated dearomatized product **45** by *in vivo* and *in vitro* studies (Scheme 1.10).⁵⁷



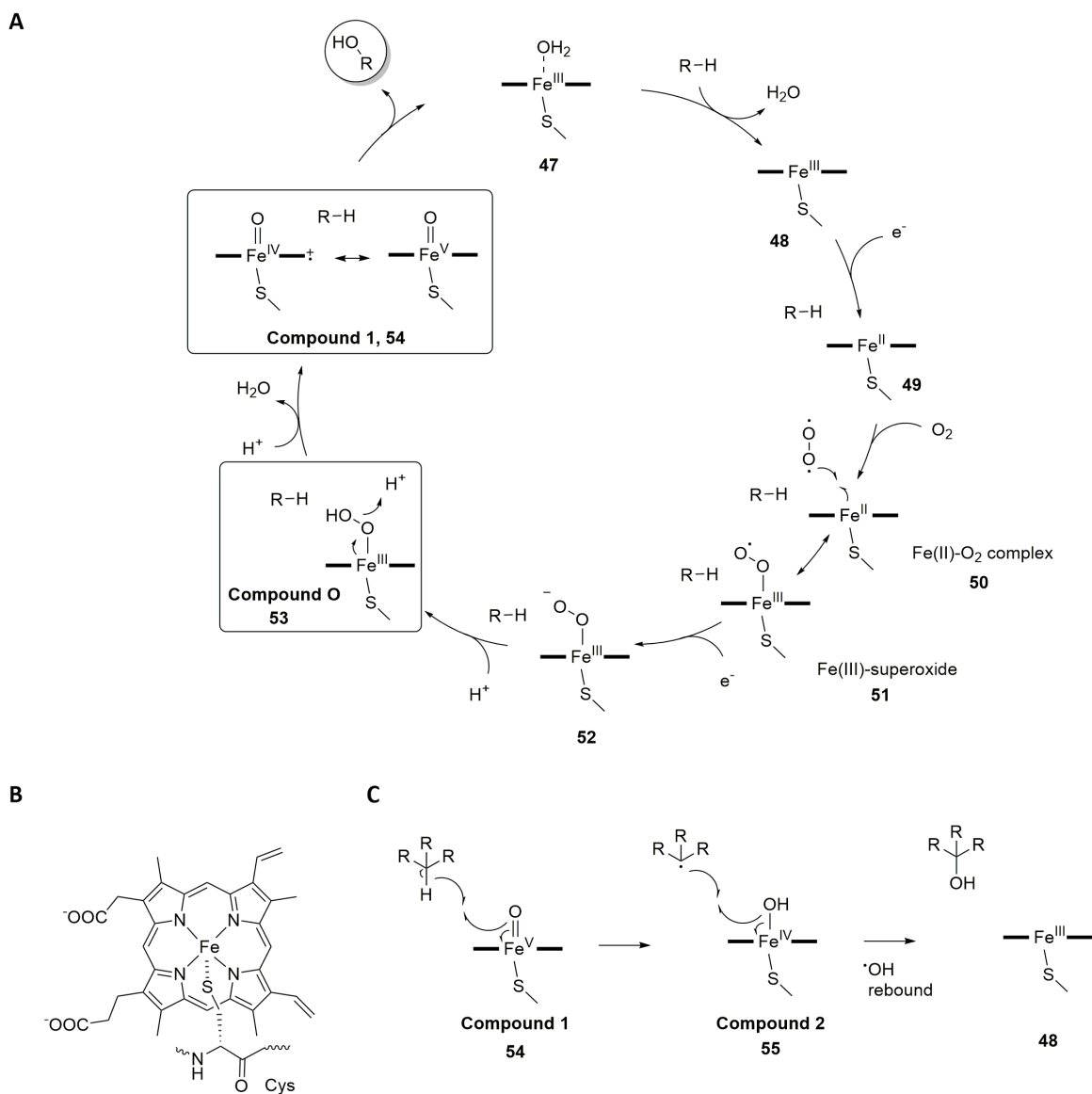
Scheme 1.10 Biosynthetic step in tropolone biosynthesis to stipitaldehyde **46** catalysed by FAD dependent monooxygenase TropB.

Cytochrome P450 monooxygenases

Cytochrome P450 monooxygenases (P450s) are heme-dependent enzymes that catalyse a diverse array of reactions and play key roles in the biosynthesis of natural products.⁵⁸ The heme (iron protoporphyrin IX) prosthetic group, which is coordinated by a thiolate ion is important for the reductive activation of molecular oxygen (Scheme 1.11 B).⁵⁹

The complex multistep mechanism starts with the heme in the resting form **47** (Scheme 1.11). Upon binding of the substrate (RH) the distal water ligand dissociates to yield **48**. A single electron reduction leads to the ferrous state **49** by a redox partner. Subsequently, the iron (II) binds dioxygen at the top axial plane of the heme ring to form a ferric peroxy state **50** (Scheme 1.11 A). The Fe(II)–O₂ complex **50** is in resonance with the Fe(III)–superoxide **51**. A second single-electron reduction

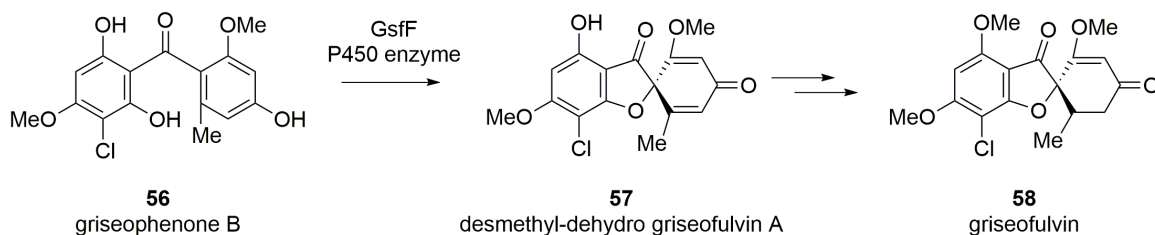
(electron comes *e.g.* from FADH₂) to **52** and further protonation furnishes the transient compound **0** (ferric hydroperoxo) species **53**. Additional protonation of this intermediate leads to cleavage of the weak O-O single bond with loss of water resulting in the highly valent oxo-iron cation radical (compound **1**) **54**, which is in resonance with Fe(V)=O oxo species (Scheme 1.11 A). The ferryl oxo compound **1** is a strong oxidising agent. It can abstract a hydrogen atom from the substrate to form compound **2** **55** and a carbon radical, prior to rebound of the hydroxyl to the substrate radical to form the hydroxylated product (Scheme 1.11 C).⁶⁰



Scheme 1.11 Catalytic mechanism of P450s adapted from Munro *et al.*⁶⁰ **A**, cytochrome P450 reaction cycle; **B**, iron atom in the equatorial plane of the heme; **C**, homolytic mechanism and hydroxyl radical rebound in heme-dependent oxygenases.

Cytochrome P450 GsfF was shown to catalyse the oxidation of griseophenone **56** to yield desmethyl-dehydro griseofulvin **57** during the biosynthesis of griseofulvin **58**.⁶¹ Computational analysis

suggests that GsfF preferentially performs phenolic O–H abstraction over benzene epoxidation during the oxidation to form the grisan structure (Scheme 1.12).⁶²



Scheme 1.12 Role of P450 GsfF during griseofulvin **58** biosynthesis.

Non-heme iron-dependent oxidative enzymes

A class of enzymes called non-heme iron-dependent oxygenases are versatile enzymes which can catalyse a variety of reactions.⁶³ Non-heme iron and α -ketoglutarate dependent oxygenases (α KG–NHFes) are the largest known subclasses of mononuclear non-heme iron enzymes and they are widely distributed in bacteria, fungi, plants and vertebrates, where they play a role in *e.g.* DNA/ RNA repair but also in biosynthesis of natural products.⁶⁴ They catalyse hydroxylation, dealkylation, desaturation, epoxidation, epimerisation, halogenation, cyclisation, peroxide formation, and ring expansion/contraction reactions by generating an Fe^{IV} -oxoiron intermediate for C–H bond activation.^{65,66} The family of α KG–NHFes possess two conserved structural features. On the one hand $\text{Fe}(\text{II})$ is ligated by two His residues and a carboxylate from either Glu or Asp. This metal motif is termed: 2–His–1–carboxylate motif. On the other hand it is located within a double-stranded β -helix fold (Fig. 1.2).⁶⁷

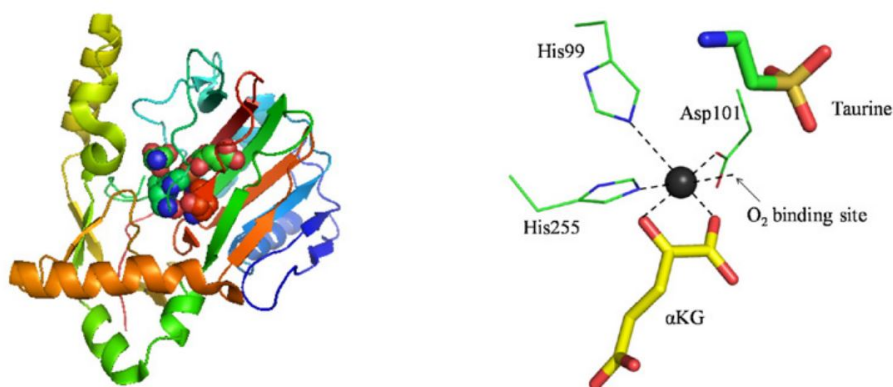
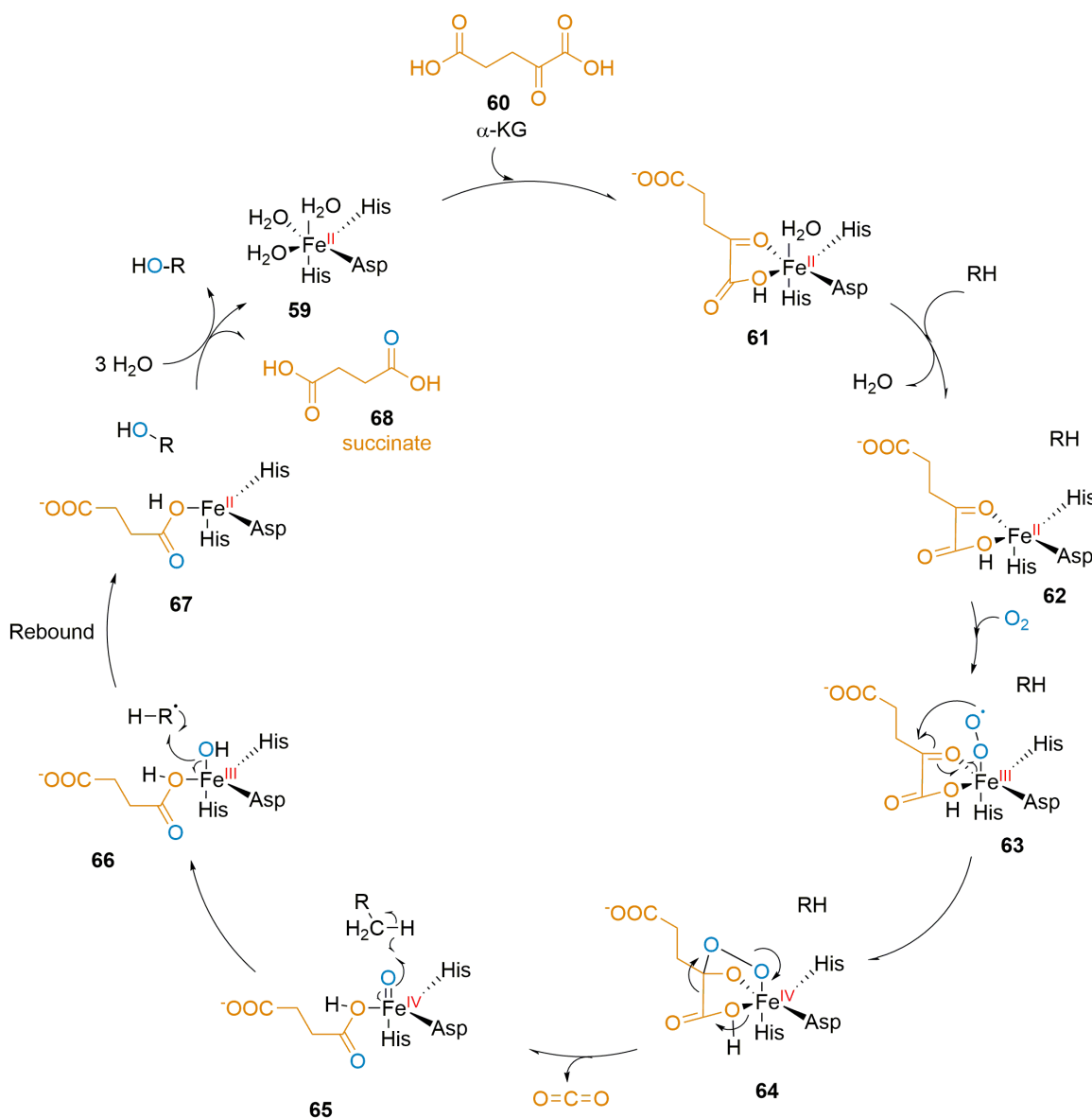


Figure 1.2 Structural information of α KG–NHFes TauD from taurine biosynthesis. Overall fold and the His–Asp–His “facial triad”, PDB ID: 1OS7, iron centre is illustrated as filled black circle, taken from Wu *et al.*⁶³

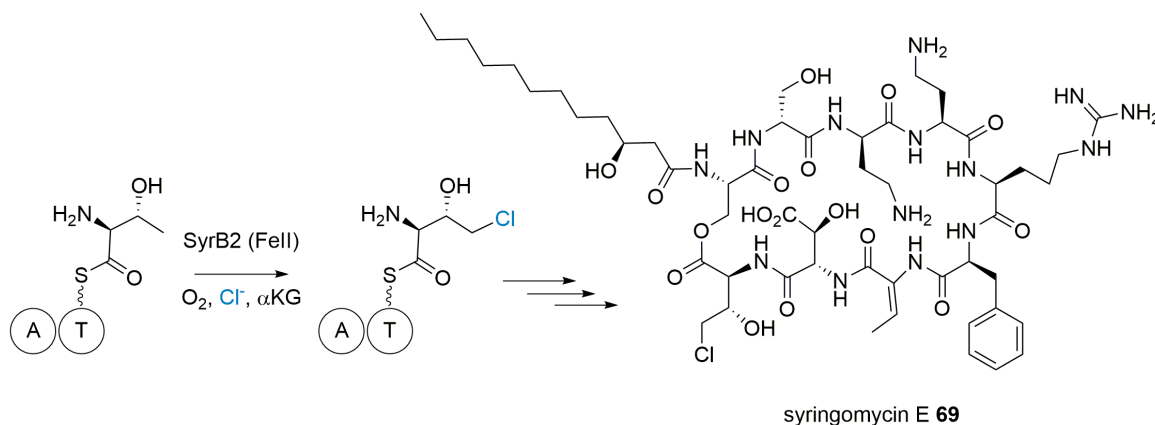
The catalytic cycle was clarified using structural and mechanistic studies. In an example for hydroxylation reaction, the cofactor Fe^{II} is coordinated in the 2–His–1–carboxylate facial triad **59**. Upon

binding of α -ketoglutarate (α KG) **60** the iron center displaces two of the three water molecules to yield **61** (Scheme 1.13). The remaining water molecule is released as soon as the substrate (R-H) binds to the active site of the enzyme (**62**) vacating space for binding the triplet oxygen molecule generating the Fe^{III}-superoxo intermediate **63**. The peroxide radical can then attack C-2 of α KG to give a peroxohemiketal bicyclic intermediate **64**, which collapses *via* oxidative decarboxylation to release CO₂ and yielding the high valence Fe^{IV}=O (ferryl) intermediate **65**. This species can abstract a hydrogen atom from the substrate to reduce the iron to the Fe^{III}-OH state **66** forming the substrate radical. In the end of the catalytic cycle the hydroxyl radical rebounds onto the substrate radical yielding the hydroxylated product (R-OH) and returning the iron back to the Fe^{II} state **67**, releasing the product and succinate **68** to give **59** (Scheme 1.13).^{65,66}



Scheme 1.13 Catalytic cycle of non-heme iron α KG-dependent oxygenase adopted from Tang *et al.*⁶⁶

Halogenases are a subset of α KG–NHF_e enzyme superfamily, which can install a halogen (Cl· or Br·) at an inactivated carbon center. The first discovered enzyme of that class was SyrB2 in 2005. It is involved in the biosynthesis of the non-ribosomal peptide syringomycin E **69**, catalysing the chlorination of the methyl group of threonine (Scheme 1.14).^{68,69}



Scheme 1.14 Chlorination in biosynthesis of non-ribosomal peptide syringomycin E **69** catalysed by SyrB2.^{68,69}

1.2 Investigation of natural product biosynthesis in fungi

There are numerous approaches known to investigate the biosynthesis of natural products. Two prominent methods are isotopic labelling and targeted gene knockouts in the native producer. A very powerful method to understand and also engineer secondary metabolite biosynthetic pathways is heterologous expression.

1.2.1 Isotopic labelling

Isotopic labelling studies have been used for many years to investigate biosynthetic pathways of polyketides. The method can reveal information about the polyketide backbone, the starter unit and also about later stages of the biosynthesis, such as oxidative or reductive modifications. The usage of isotopes is not "out of fashion", but still continues to give important insights into biosynthetic pathways of secondary metabolites.⁷⁰ In 2017, the structure of the fungal phytotoxins known as the phyllostictines has been revised using stable isotopes.⁷¹

Early labelling studies were conducted using radioactive labels, such as ¹⁴C and ³H. These are β -emitters which can be detected by a scintillation counter to a low level of incorporation due to a low background. The disadvantage of this technique is not only the radioactivity but that the exact position of incorporation in the molecule can only be determined by degradation studies.

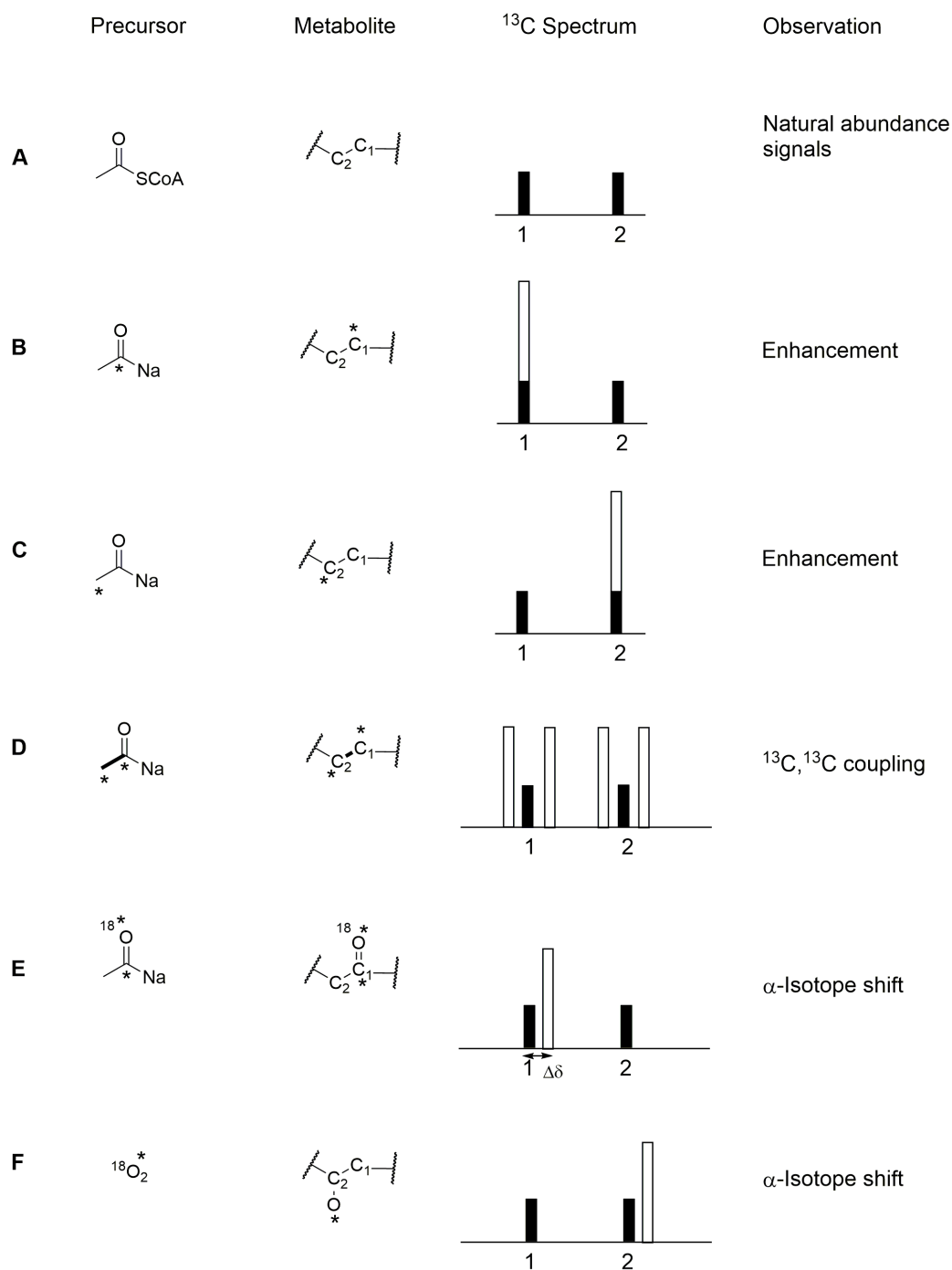


Figure 1.3 Illustration of feeding labelled precursors and subsequent signal enhancement of incorporated atoms in ^{13}C NMR spectra, adapted from Simpson.⁷²

However, pioneering work in the field of polyketide biosynthesis was done by Birch and coworkers using ^{14}C labelled acetate units to study the biosynthetic origin of 6-MSA **8** among others.²¹

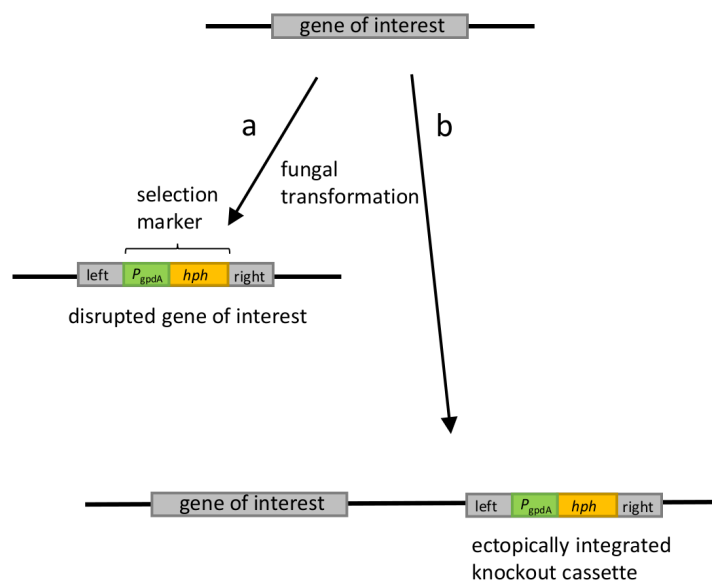
More recent labelling techniques use stable isotopes such as ^{13}C , ^2H , ^{15}N and ^{19}F . ^{13}C labelling techniques reveal information about complete labelling patterns by determining the enhancement of individual ^{13}C resonances. Examples for incorporation of ^{13}C -enriched precursors ([$1-^{13}\text{C}$]-,

[2-¹³C]-, or [1,2-¹³C₂]-acetate) into contiguous carbons of a polyketide-derived metabolite are illustrated in figure 1.3. Incorporation of a singly ¹³C-labelled precursor enhances the content of ¹³C at a particular carbon which results in an increase in intensity of the signal for that carbon compared to natural abundance of ¹³C (1.1%) (Fig. 1.3 A-C). If a doubly labelled precursor like [1,2-¹³C₂]-acetate is incorporated in a metabolite (bond stays intact), it will lead to the observation of a ¹³C-¹³C spin coupling in the corresponding ¹³C NMR spectrum (Fig. 1.3, D).⁷²

Feeding labelled ¹³C precursors allows the origin of the carbon skeleton of a compound to be determined. However, it is also important to investigate the biosynthetic origin of the oxygen atoms to understand the mechanism of the conversion during biosynthesis. The isotope ¹⁸O₂ can be introduced as doubly labelled acetate ([1-¹³C,¹⁸O₂]-acetate) or by growth in an ¹⁸O₂ environment. The isotope itself is not NMR active, but it is possible to use ¹³C as a "reporter" nucleus for oxygen to detect it indirectly by changes induced in the ¹³C NMR chemical shifts of directly-attached carbon atoms.⁷³ Substitution of a proton *alpha* or *beta* to the ¹³C by deuterium or oxygen causes an upfield shift ($\Delta\delta$) in the ¹³C chemical shift (Fig 1.3, E+F). The shift ($\Delta\delta$) depends on the atom and the hybridisation of the ¹³C atom to which it is attached.^{72,74}

1.2.2 Gene knockout

Targeted gene knockouts or gene silencing experiments are a useful tool to elucidate natural product biosynthesis. These experiments are widely used to investigate the function of encoded proteins and link pathway steps to their genes/cluster by disrupting the gene of interest (GOI) from the biosynthetic gene cluster (BGC) and analysing the changes in secondary metabolite production. A popular strategy to disrupt GOI is based on stable integration of exogenous DNA into the fungal genome by homologous recombination (HR) using a polyethylene glycol (PEG)/ CaCl₂ based transformation protocol.⁷⁵ Other transformation techniques have been shown to be successful for filamentous fungi, such as electroporation of protoplasts,⁷⁶ or *Agrobacterium tumefaciens* mediated transformation.⁷⁷ When the exogenous DNA is inside the fungal cell it is integrated into the genome by replacing the targeted gene by a marker gene (*e.g.* an antibiotic resistance gene). Thus the GOI is inactivated and not further expressed. For this purpose the transformed DNA is constructed to occupy two sequences homologous to the GOI which flank a selection marker on either side (Scheme 1.15, left and right).

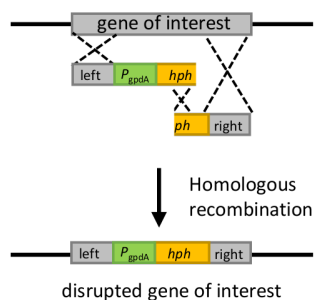


Scheme 1.15 Integration of a knockout cassette by homologous recombination [a] and ectopic integration of a knockout cassette resulting in false positive transformants [b].

In this study transformants are selected by using the *E. coli* gene *hph*, encoding a hygromycin B (*hygB*) phosphatase as selection marker for the antibiotic hygromycin B.^{75,78} However, DNA-mediated transformation experiments are a challenging task because the transferred DNA can be also integrated ectopically into the genomic DNA. In some cases ectopic integration can be predominant (Scheme 1.15 b).⁷⁹ The reason is that DNA integration in filamentous fungi is mainly driven by non-homologous end joining (NHEJ). The NHEJ-mechanism is based on random ectopic integration of DNA into the genome and site-specific recombination occurs at less than 1%.⁷⁹

In order to reduce the number of false positive transformants a split-marker technology can be used. The split-marker method was initially established for *S. cerevisiae*⁸⁰ before it has been also applied to a number of filamentous fungi. The modified gene disruption method is known as bipartite marker strategy and was developed in *A. nidulans*.⁸¹ The bipartite method ensures that substitution of a wild-type targeted gene is only possible when three cross-over events occur between two individually cotransformed DNA fragments to generate an intact resistance gene (Scheme 1.16).⁸¹ The resistance gene used for MF5453 is *hph* which is flanked by homologous gene sequences of the targeted gene that allow homologous recombination between fungal genomic DNA and the two parts of the knockout cassettes to substitute the targeted gene by the selectable marker (Scheme 1.16). A successful application of the split-marker system was shown in 2001 for *Acremonium chrysogenum* in which the *mecB* gene was inactivated. The gene *mecB* encodes for a cystathionine- γ -lyase, which is required for cephalosporin C biosynthesis. Targeted gene disruption frequency was increased to 5%.⁸² Catlett *et al.* utilized the hygromycin B resistance marker for selection and they could increase integration efficiencies to up to 100% by using flanking regions

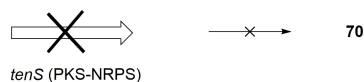
of only 285–761 bp in size.⁸³ Also in our group it was shown that the bipartite strategy was applied successfully to a variety of other fungi, resulting in frequent disruption of the GOI.^{53,84,85}



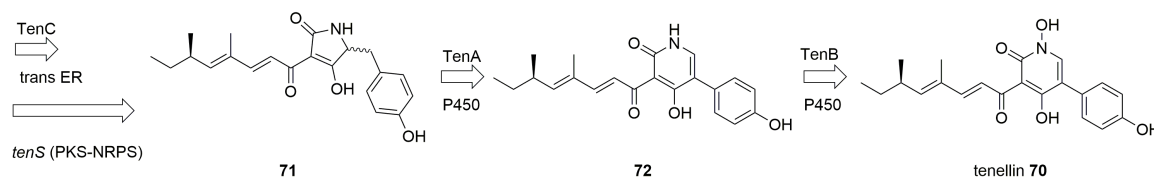
Scheme 1.16 Bipartite marker strategy: Reduction of ectopic integration by cotransformation of two DNA fragments with three homologous recombination events.

There are also other approaches for investigation of gene functions. One of them is post-transcriptional gene silencing, which is based on mRNA degradation. In comparison to gene disruption methods, silencing causes only decreased yields, but not complete vacancy of the explored secondary metabolite. Gene silencing relies on introduction of antisense RNA to the mRNA of the targeted gene. An RNA dependent RNA polymerase promotes the formation of a deviant double stranded RNA, which initiates the endogenous silencing pathway in fungi.⁸⁶

A



B



Scheme 1.17 Biosynthesis of tenellin **70**. **A**, knockout of *tenS*; **B**, biosynthetic steps catalysed by P450s TenA and TenB.

Both methods, gene disruption and silencing were reported to be successful during the investigation of tenellin **70** biosynthesis. The core gene, *tenS* (PKS-NRPS), was knocked out and showed abolition of **70** production and a silencing method targeting the two P450 encoding genes (*tenA* and *tenB*) was reported to result in accumulation of pathway intermediates pretenellin-A **71** and pretenellin-B **72** (Scheme 1.17).^{87,88}

1.2.3 Fungal heterologous expression

Heterologous expression of single genes, partial or entire gene clusters is a powerful tool to elucidate the biosynthetic steps of secondary metabolites in bacteria or fungi. In addition to the analysis potential, heterologous expression also bears numerous engineering possibilities and promises progress in the discovery of novel compounds or optimisation of product yields.⁸⁹ For some fungi targeted gene knockout or transformation methods are ineffective, since the GOI is mainly incorporated ectopically. One example is *Sarocladium schorii*, the producer of the fungal meroterpenoid xenovulene A **73**. Here, heterologous expression was the preferred method to study the pathway.⁹⁰

In order to uncover the steps of a biosynthetic pathway, several genes have to be expressed simultaneously. This involves the transfer of all relevant genes from a donor strain to a suitable, foreign host to reconstitute the biosynthetic pathway *in vivo*. In general, many polyketide-synthesising strains show poor laboratory growth or possess inefficient transformation and homologous recombination abilities. It is therefore desirable to transplant entire BGC into model strains which are more tractable to genetic manipulation.⁹¹

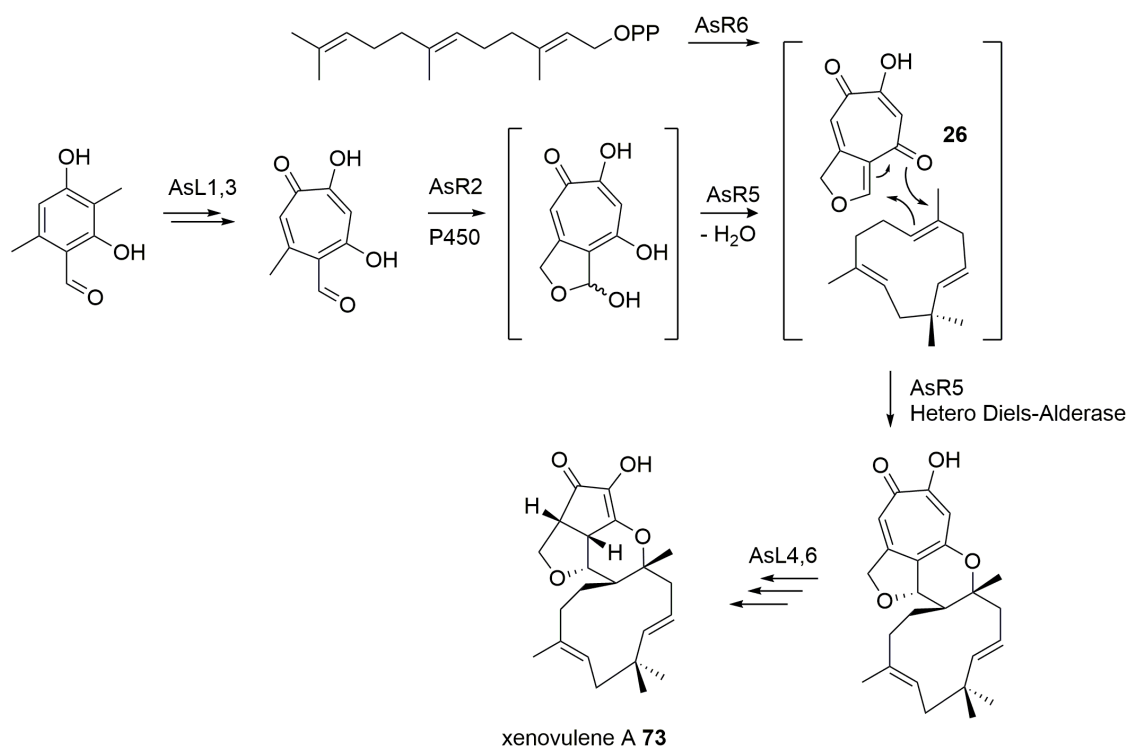
For bacterial BGC this has been achieved *via* bifunctional phage P1-derived artificial chromosomes (PAC) vectors.⁹² More recently it was shown that large gene clusters can be directly cloned into *E. coli* expression vectors. Using this approach, ten PKS and NRPS pathways from *Photorhabdus luminescens* were directly transferred into *E. coli*.⁹³ Expression of fungal genes in bacteria such as *E. coli* or actinomycetes is difficult as it bears several problems. Bacterial hosts cannot process eukaryotic introns and often possesses a significant codon bias. Furthermore, bacteria show difficulty in correctly folding fungal polypeptides.⁹⁴ However, *E. coli* is often a suitable platform for recombinant purification of single proteins which enables analysis of individual enzymes *in vitro* such as TropB, a FAD-dependent monooxygenase during stipitaldehyde **46** biosynthesis (Scheme 1.10).⁵⁷

Another host used for heterologous expression of fungal BGC is the yeast strain *S. cerevisiae*. Advantages of this host are that homologous recombination is highly efficient and that yeast itself has an insignificant endogenous secondary metabolism, minimising the risk of side reactions and cross chemistry between the expressed pathway and endogenous biosynthesis. Nevertheless, the latter advantage can sometimes be a limiting factor, *e.g.* lacking relevant building blocks.⁹⁵ Last year, HEx (Heterologous EXpression) synthetic biology platform was developed for rapid expression of fungal biosynthetic gene clusters in *S. cerevisiae* to overcome the problem of accessing metabolites of uncultivable organisms or silent gene clusters. Harvey and coworkers showed that 22 out of 41 fungal BGC produced detectable compounds (by liquid chromatography–mass spectrometry, LCMS analysis) while expressed in yeast.⁹⁶

In addition to *E. coli* and *S. cerevisiae*, *A. oryzae* spp. are often chosen as heterologous host for studying fungal biosynthetic pathways. This was shown to be successful in many examples and was summarised in multiple reviews.^{89,97,98,99,100} There are a number of advantages of filamentous fungi as heterologous hosts in contrast to *E. coli* and yeast: strains are usually stable and easily cultured in the lab, they can be modified genetically and it is not essential to coexpress a foreign PPTase gene to provide the prosthetic group for the ACP domains of PKS. Furthermore the genetic systems of the producing fungi and the host are in some cases compatible and therefore cDNA preparation or engineering intron removal is not always required.⁹⁸ Most commonly, *A. oryzae* is selected as host, as it has been used for a long time for fermentation to get food products like sake, miso and soy sauce.¹⁰⁰ Moreover, this species has been given the status of "Generally Regarded As Safe" (GRAS) organism, accordingly safe to be used for food and pharmaceuticals.¹⁰¹

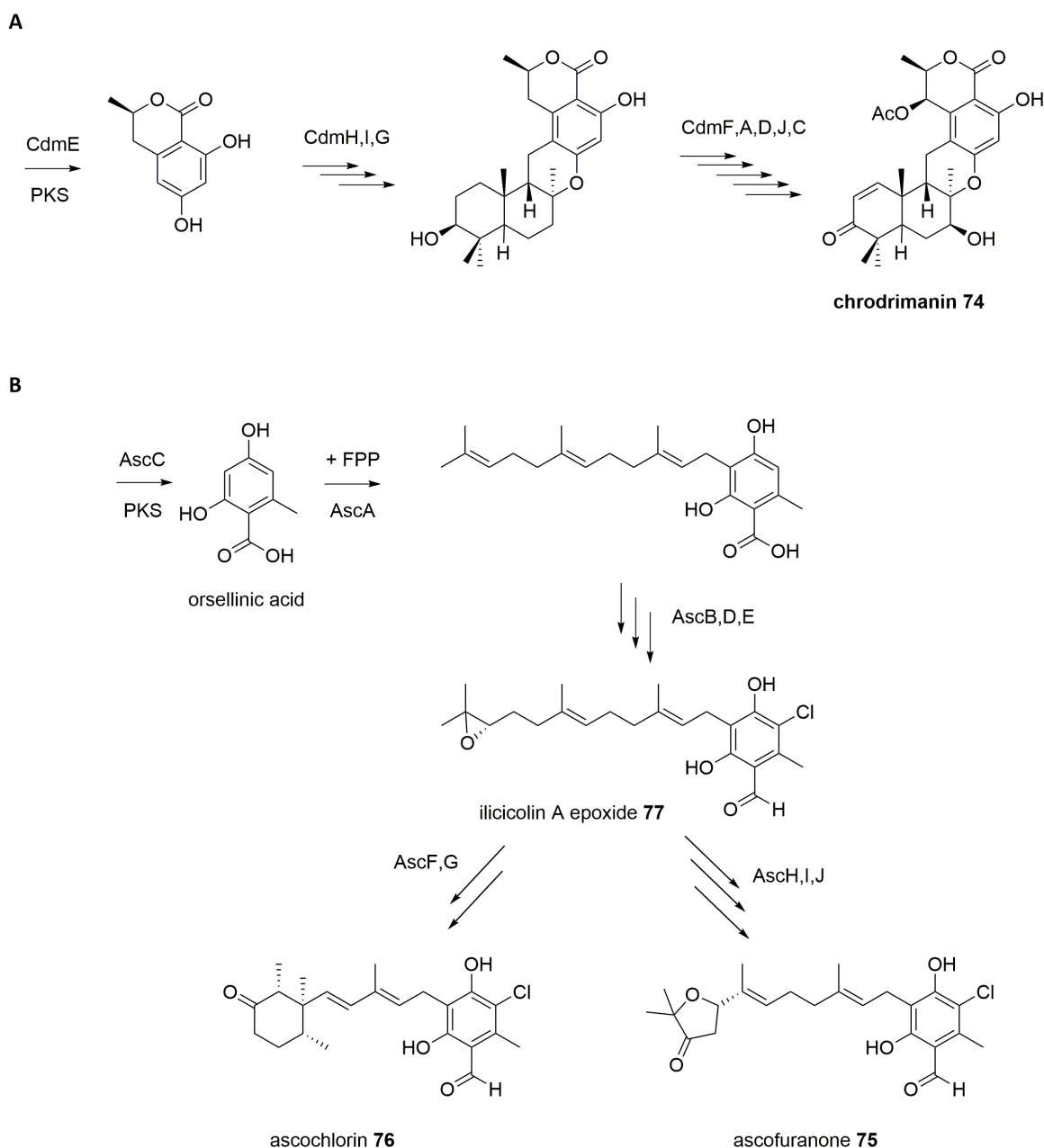
In 2010, heterologous expression was first used to reconstitute a complete fungal biosynthetic multigene cluster. The biosynthesis of tenellin **70** was reconstituted by expressing four genes (*tenS*, *tenC*, *tenA*, *tenB*) of the pathway from *Beauveria bassiana* in arginine auxotrophic *A. oryzae* M-2-3 strain with each gene under the control of the starch-inducible *amyB* promotor (Scheme 1.17).⁹⁴

The biosynthesis of more complex fungal metabolites containing a higher number of genes, such as byssochlamic acid, agnestadride A⁸⁴ and xenovulene A **73** have also been successfully investigated through heterologous expression experiments.⁹⁰ The entire biosynthetic pathway of the complex fungal meroterpenoid **73** was reconstituted in heterologous host *A. oryzae*. This allowed the elucidation of each chemical step and enzyme involved in the biosynthesis, including the initial formation of tropolone intermediates, a putative hetero Diels—Alder enzyme, a new class of terpene cyclases and two enzymes involved in oxidative ring contractions (Scheme 1.18).⁹⁰



Scheme 1.18 Basic steps in the biosynthesis of xenovulene A 73.

Recently, the Abe group reported the elucidation of very complex biosynthetic steps to a variety of other fungal meroterpenoids using heterologous expression in *A. oryzae*.^{102,103,104,105,106} They elucidated the complete biosynthetic pathway of fungal meroterpenoid chrodriamanin B 74 (*Penicillium verruculosum* TPU1311) by heterologous reconstitution of its biosynthesis in *A. oryzae*. In this study 13 different set of genes were coexpressed with up to nine genes simultaneously (Scheme 1.19 A).¹⁰³ Using the same approach Abe and coworkers also elucidated the biosynthetic pathway to the two fungal meroterpenoids ascofuranone 75 and ascochlorin 76, which branch from a common precursor, ilicicolin A epoxide 77 (Scheme 1.19 B).¹⁰⁵ All genes (*ascABCDEFG*) involved in 76 biosynthesis and a transcriptional factor *ascR* are clustered together. However, the three genes (*ascHIJ*) required for the steps to 75 are located in a distant BGC elsewhere in the genome.¹⁰⁵

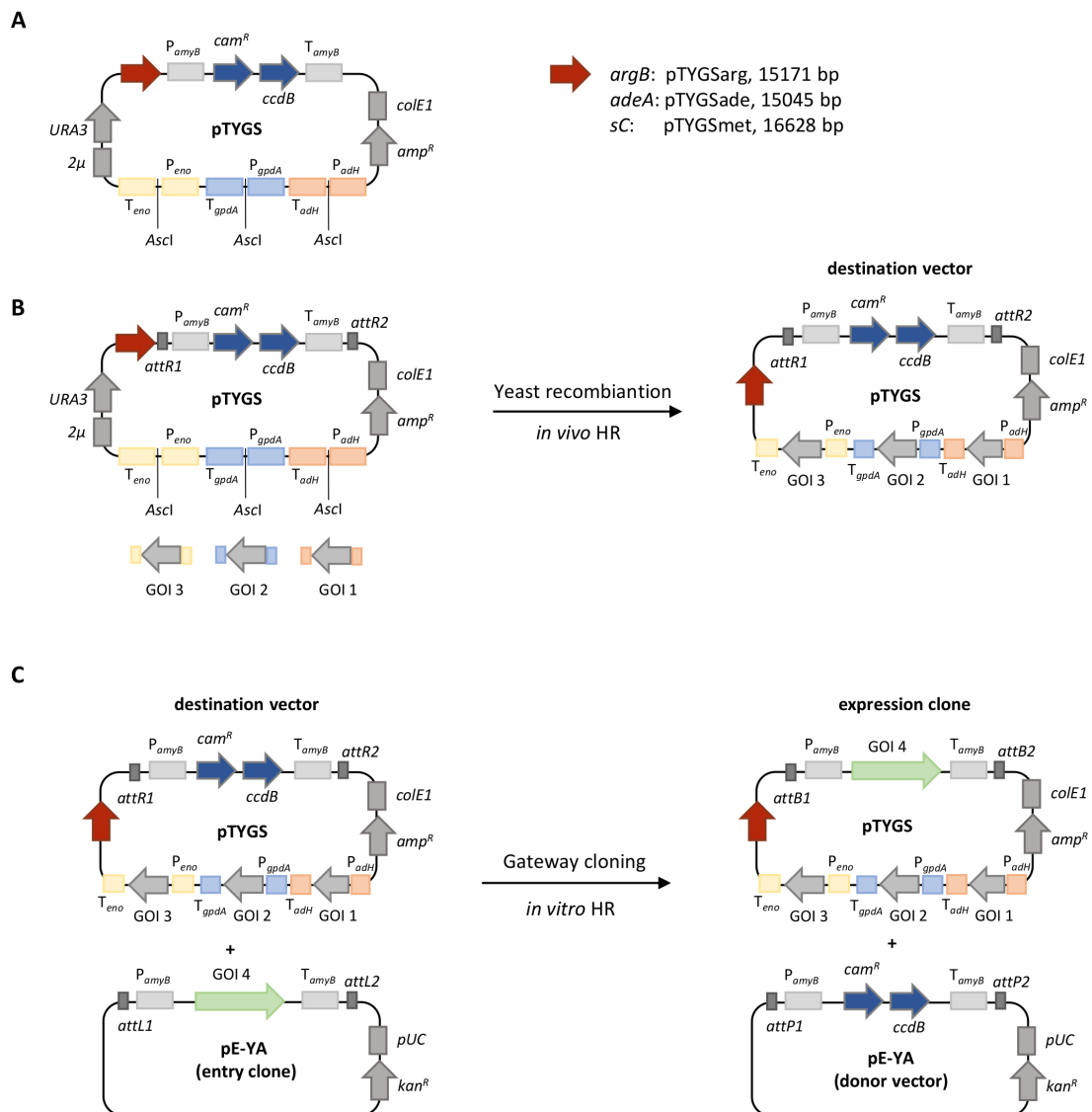


Scheme 1.19 Biosynthesis of fungal secondary metabolites elucidated *via* heterologous expression in *A. oryzae*. **A**, biosynthetic steps to chrodrimanin **74**; **B**, pathway to the two fungal meroterpenoids ascoclorin **76** and ascofuranone **75**.

In all examples discussed above and also in this thesis, a multigene expression vector system was used. The quadruply auxotrophic *A. oryzae* NSAR1 is deficient in arginine ($\Delta argB$), adenine ($adeA^-$), methionine (sC^-) and ammonium ($niaD^-$) metabolism. Additionally, the natural sensitivity of *A. oryzae* NSAR1 towards the antibiotics bleomycin and the herbicide glufosinate ($bleR$, bar) enables the use of two more selection markers. Using designed fungal expression vectors containing different selection markers ($argB$, $adeA$, sC , $niaD$, $bleR$, bar) enables the reconstitution of larger biosynthetic pathways by simultaneous expression of up to 26 genes.⁹⁸

1.2 Investigation of natural product biosynthesis in fungi

Each plasmid possesses four fungal promoter/ terminator pairs for gene cloning (P/T_{adh}, P/T_{gpd}, P/T_{eno} and P/T_{amyB}) (Scheme 1.20 A). The first three are flanked by an *Ascl* site that can be used to simultaneously add up to three tailoring genes *in vivo* by yeast homologous recombination, which include at least 30 bp homologous overlap sequence that is usually introduced by PCR through tails at the designed oligonucleotides (Scheme 1.20 B). Furthermore the fungal expression vectors contain an origin of replication *colE1* for replication in *E. coli* and an ampicillin resistance gene *amp^R* as well as an chloramphenicol gene (*cam^R*). For selection and propagation in uracil deficient *S. cerevisiae* strains the vectors contain the *ura3* gene, encoding orotidine 5'-phosphosphate decarboxylase, and the 2μ *ori* (Scheme 1.20 A). A eukaryotic origin of replication is not necessary because *A. oryzae* integrates the plasmid ectopically in its genome, thereby causing a stable transformation.⁷⁵



Scheme 1.20 Fungal expression vectors used in this work and gateway cloning procedure: **A**, pTYGS vector family; **B**, cloning of up to 3 genes into pTYGS vectors by *in vivo* homologous recombination (HR) in *S. cerevisiae*; **C**, Gateway cloning by LR recombination between pE-YA (entry clone) and pTYGS (destination vector) yielding in expression clone (pTYGS) and donor vector (pE-YA).

One additional gene can be cloned after the *amyB* promoter (P_{amyB}) which is induced by starch and maltose. This cloning site is suitable for *in vitro* Gateway cloning, as it contains recombination sites (*attR1* and *attR2*) which flank a Gateway cassette consisting of the *ccdB* gene, encoding the CcdB killer protein, and chloramphenicol resistance *cam^R* (Scheme 1.20 C). For gateway cloning a LR recombinase kit (Invitrogen) is used containing a mixture of bacteriophage λ integrase proteins (integrase, integration host factor, excisionase) which interchange DNA flanked by *attR* sites with DNA flanked by *attL* sites (e.g. *E. coli*-*S. cerevisiae* shuttle vector pE-YA). The destination vector with the gene of interest (GOI 4) can be selected by using an *E. coli* TOP10 strain as well as ampicillin as antibiotic. This is possible, as destination vectors possessing the Gateway cassette have to be propagated in *E. coli ccdB* survival cells, which are resistant to the *ccdB* gene product.

1.3 Overall aims

The main focus of this thesis concentrates on understanding and engineering the biosynthesis of three polyketides produced by filamentous fungi: squalestatin S1 **1**, strobilurin A **2** and SCH-642305 **3**.

Using most up to date methods such as targeted gene knockout, heterologous expression in *Aspergillus oryzae* and *in vitro* studies with purified enzyme should uncover the order and functions of genes involved in the three biosynthetic pathways. The particular aim for squalestatin biosynthesis is to discover the oxygenases which are involved in constructing the highly oxidised core structure of **1**. The major focus of the biosynthetic studies of strobilurin is based on investigating the unusual oxidative rearrangement to form the core β -methoxyacrylate moiety of this valuable class of agricultural used fungicides. In addition, the aim for SCH biosynthesis is to reveal the molecular mechanism of each chemical step *in vivo*, with special attention on the formation of the 6-membered ring.

2 Biosynthetic studies of Squalestatin S1

Partial results of the presented work have been published in *Chem. Sci.*, **2019**, *10*, 1227–1231.¹⁰⁷

2.1 Introduction

Squalestatin S1 (SQS1, **1**), also known as zaragozic acid (ZA) A is a fungal metabolite, discovered in 1991 and 1992 by three independent groups: Researchers from Merck¹², Glaxo¹⁰⁸ and Tokyo Noko University/Mitsubishi Kasei Corporation.¹⁰⁹ Glaxo isolated **1** from *Phoma sp.* C2932, the Tokyo group from *Setosphaeria khartoumensis* and Merck from cultures of a the non-sporulating unidentified strain MF5453. The name 'zaragozic acid' originated from its producing strain MF5453 which was isolated from a water sample taken from the Lalon river in Zaragoza, Spain.¹²

SQS1 **1** belongs to a family of squalestatins that are only produced by fungi (*Ascomycota*). Over 20 different squalestatins have been isolated from at least eleven different taxa of fungi.¹¹⁰ The first members reported from this family were ZA A **1**, ZA B **79**¹¹¹ and ZA C **80**¹¹² (Fig. 2.1). All of which contain the characteristic 2,8-dioxabicyclo[3.2.1]octane-3,4,5-tricarboxylic acid 'core' which is decorated with two lipophilic side chains. The variation is mostly due to differences in the ester side chain at C-6 and the alkyl side chain at C-1.¹¹³

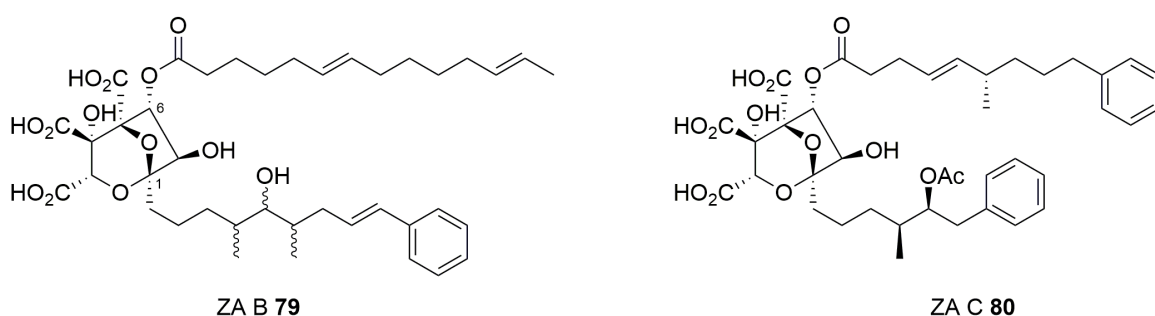


Figure 2.1 Structures of ZA B **79** and ZA C **80**.

The squalestatins are potent pico-molar inhibitors of squalene synthase, an enzyme of sterol biosynthesis. The discovery of squalestatins resulted from a screen of natural products for novel inhibitors

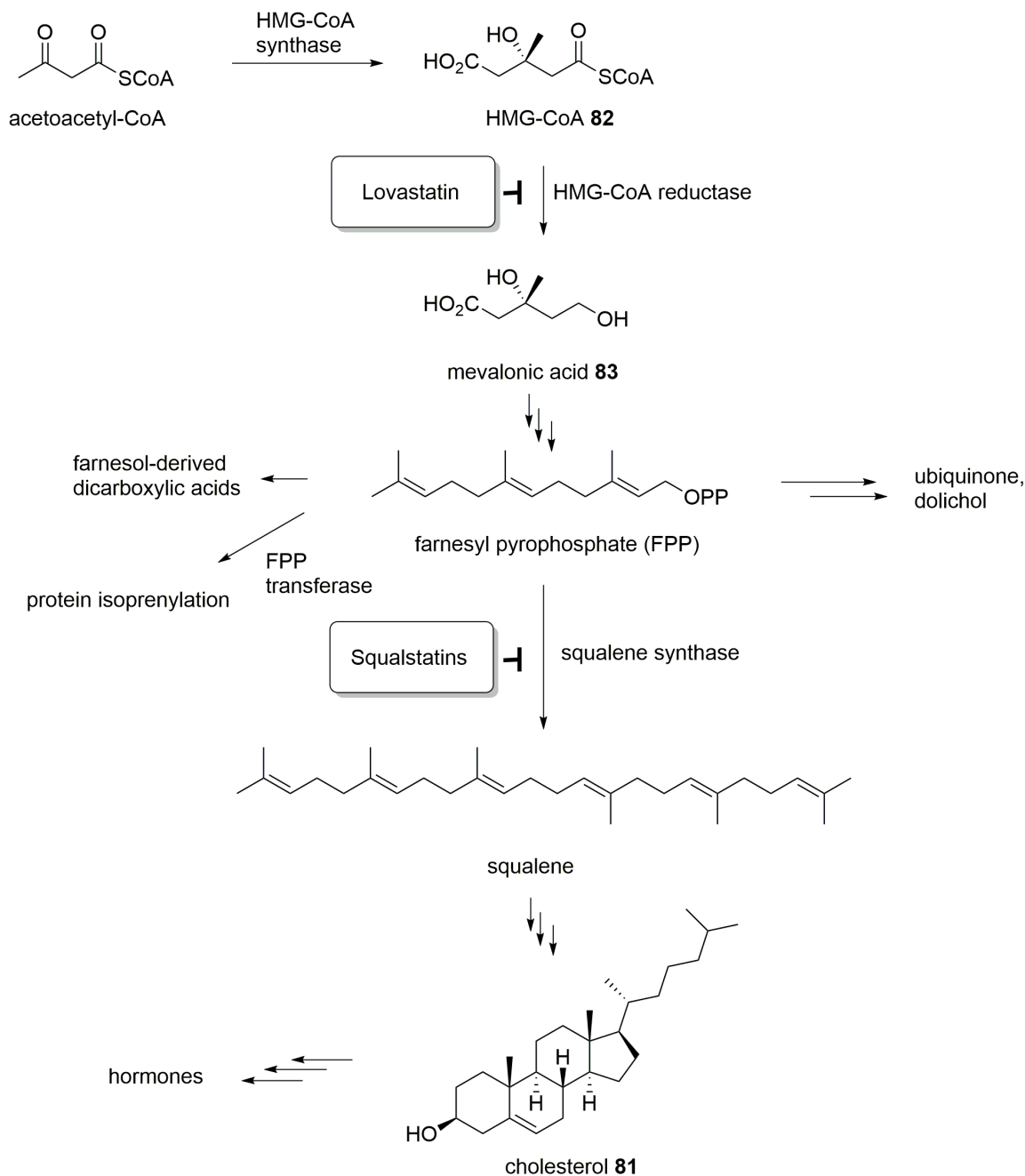
of cholesterol biosynthesis.¹² Inhibitors of cholesterol biosynthesis could be therapeutic agents with potential for the treatment of hypercholesterolemia.

2.1.1 Biological activity

SQS1 **1** shows broad antifungal activity and is used as a lead structure for cholesterol-lowering drugs.¹⁰⁸ Cholesterol biosynthesis in mammals starts with acetyl-CoA **11** which is converted via over 20 enzymatic reactions to the final product cholesterol **81** (Scheme 2.1). Among these, two enzymes: 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase and squalene synthase play key roles.¹¹⁴

HMG-CoA reductase is a target for cholesterol lowering drugs. This enzyme catalyses the conversion of HMG-CoA **82** to mevalonic acid **83**. HMG-CoA reductase inhibitors, also traditionally known as statins, reduce cardiovascular disease morbidity and mortality with a high level of safety. One example is lovastatin **6**, which was shown to be an active inhibitor of cholesterol biosynthesis in rats and dogs.¹³ HMG-CoA reductase inhibitors also show disadvantages as depletion of non-sterol products lead to toxic side effects and some people do not tolerate statins or the drug is not effective and therefore critical as treatment for coronary heart diseases.¹¹⁵

For this reason an alternative promising strategy for the development of novel cholesterol lowering agents was to target the enzyme squalene synthase, which catalyses the first step in sterol biosynthesis (Scheme 2.1). Due to the fact that squalene synthase inhibitors act at a relatively late step in cholesterol biosynthesis they do not affect the mevalonate-farnesyl diphosphate pathway.¹¹⁴ SQS1 is a potent pico-molar inhibitor of mammalian and fungal squalene synthase.¹⁰⁹ Moreover, **1** shows broad spectrum antifungal properties and was shown to lower serum cholesterol levels (up to 75 %) in marmosets *in vivo*, which have a similar lipoprotein profile compared to humans.^{108,116}



Scheme 2.1 Simplified sterol biosynthetic pathway modified after Nadin and Nicolaou¹¹⁷; lovastatin inhibits HMG-CoA synthase and squalstatins inhibit squalene synthase.

2.1.2 Biosynthesis of SQS1 – *status quo*

Since the early 1990s, the investigation of the biosynthesis of squalstatins was an area of intense interest and research. The biosynthetic origin of squalstatins was discovered independently by Merck and Glaxo.^{7,47} All squalstatins share the same unique 2,8-dioxabicyclo[3.2.1]octane-4,6,7-

trihydroxy-3,4,5-tricarboxylic acid ring system regardless of the producing fungus. Using labelling experiments with isotopically labelled substrates the origin of **1** was discovered (Fig. 2.2).

Isotopic labelling studies

Feeding experiments using isotope labelled precursors were conducted in 1992 and 1993 to determine the biosynthetic origin of the heavy atoms in squalstatin S1.^{7,47} Significant incorporation of acetate **13**, succinate **84**, L-phenylalanine **85**, benzoic acid **86** and cinnamic acid **87** as well as the methyl group from L-methionine into **1** were shown by feeding ¹⁴C-labelled precursors.⁴⁷ The exact sites of incorporation of the precursors and the direction of PKS skeleton were investigated by feeding ¹³C-labelled substrates and carbon NMR analysis.⁴⁷

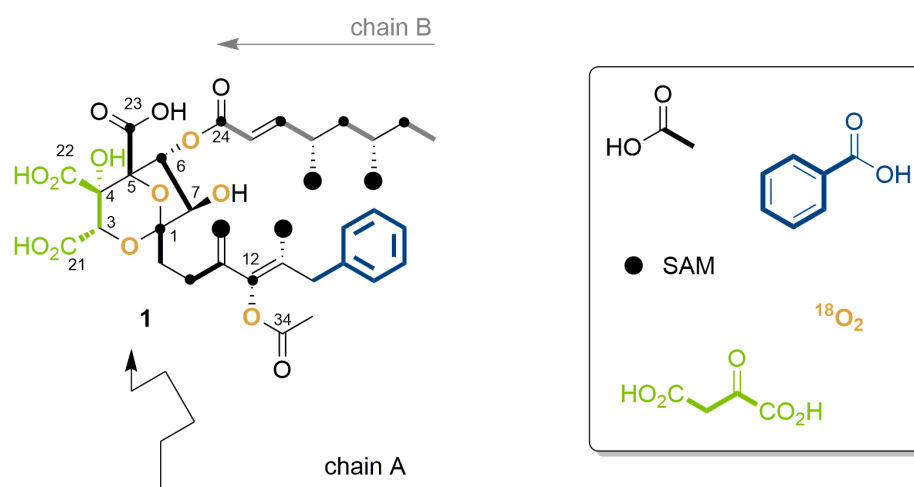


Figure 2.2 Labelling patterns of **1**.

The feeding experiments indicated that **1** is made of two polyketide chains: a main chain (A), composed of a benzoate-derived starter unit and five acetate-derived extender units resulting in a hexaketide, and a side chain (B), composed of three acetate-derived units forming a tetraketide (Fig. 2.2, chain A + B).

The biosynthesis of chain A starts with an unusual starter unit, a benzoate unit derived from phenylalanine **85** *via* benzoic acid **86**. The incorporation of [^{3-¹³C}]-phenylalanine, [^{U-¹⁴C}]-cinnamic acid, and different labelled ¹³C-benzoic acid precursors revealed that the nine carbons of phenylalanine are metabolised losing the first two carbon atoms to retain the seven carbons of benzoic acid.^{7,47} The most likely starter molecule of chain A is an activated form of benzoic acid, *e.g.* benzoyl CoA. Recent feeding studies conducted by Tang and coworkers showed that cinnamoyl CoA **88** is formed early in the SQS precursor biosynthesis.¹¹⁸

Both chains (A+B) are formed of acetate units and possess four additional methyl groups which were shown to derive from the methyl group of *L*-methionine transferred by *S*-adenosylmethionine (SAM). [1-¹³C]-propionate was not incorporated into **1**, whereas supplementation with *L*-[methyl-¹³C]-methionine resulted in high levels of enrichment at all four methyl groups (Fig. 2.2, bold circles).

Carbons C-3, C-4, C-21 and C-22 of **1** derive from incorporation of an intact C₄-unit from the citric acid cycle. The precursor is likely to be an oxidised metabolite of succinate, such as oxaloacetate **89** (Fig. 2.2, green).⁴⁷

In order to investigate the origin of the oxygen atoms of **1**, the SQS1 producing strain was incubated in the presence of [1¹³C, ¹⁸O₂]-acetate and ¹⁸O₂ in separate experiments.⁷ The two ester carbonyl groups of chain A and B were shown to derive from acetate oxygens (Fig. 2.3 A). Moreover, the closed fermentation of the producing fungus in a ¹⁸O₂ environment produced SQS1 that displayed characteristic ¹⁸O₂ carbon NMR shifts for five carbons on chain A, suggesting oxygens on carbons C-1/-5, C-3, C-6, C-7 and C-12 to be derived from atmospheric oxygen (Fig. 2.2, orange + Fig. 2.3 B). This outcome was supported by mass spectrometry of labelled **1**, as this sample showed a number of molecular ions +2, +4, +6, +8 and +10 mass units greater than unlabelled O₂. This ion pattern indicated the presence of up to five ¹⁸O₂ atoms per molecule of **1**. The lack of carbon shift at C-4 under both labelling strategies suggested that the hydroxyl group at this position is not derived from oxygen (Fig. 2.3).⁷

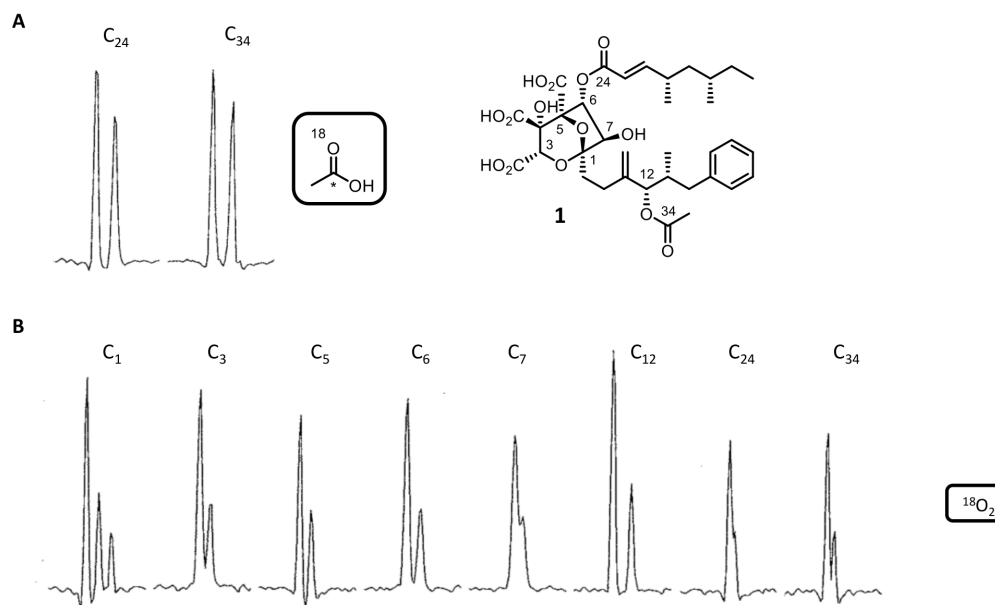
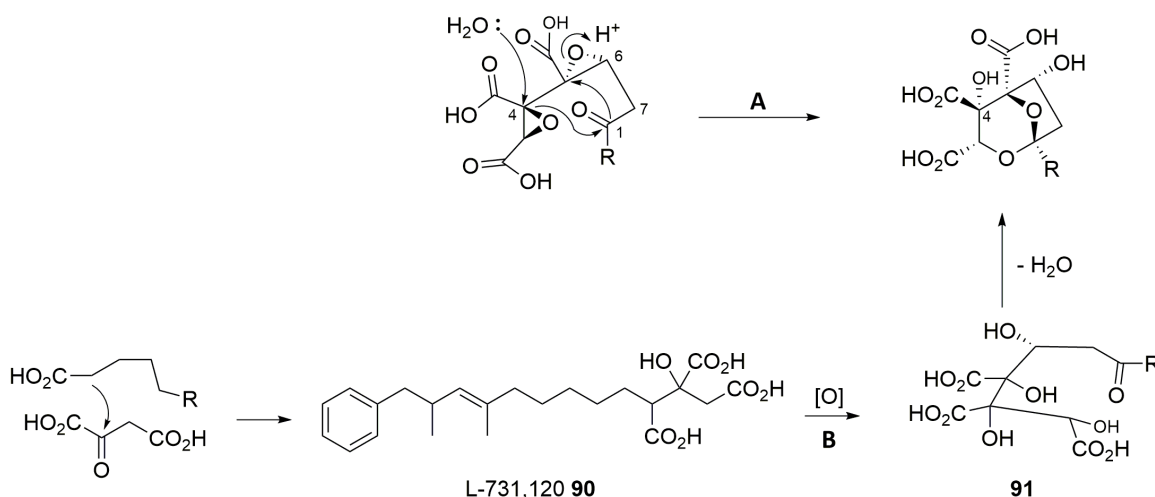


Figure 2.3 Sections of ¹³C NMR spectrum of **1** labelled by [1¹³C, ¹⁸O₂]-acetate (A) and ¹⁸O₂ gas (B) showing α -isotope shifts, original spectrum taken from Jones *et al.*, 1992.⁷

On the basis of the oxygen labelling results the group of Glaxo proposed a mechanism where a nucleophilic attack by a water molecule at C-4 triggered the ring formation through C-1 (Scheme 2.2, route A).⁷ Another proposal for the ring closing mechanism is mediated by α - and β -oxidations of an alkylcitrate **90** to give two hydroxyl groups at C-3 and C-5 as well as at C-6 to furnish intermediate **91**. The carbonyl group at C-1 must arise from aliphatic oxidation and vanished during bicyclic ring formation from the cyclic hemiketal (Scheme 2.2, route B).¹¹³ In 1995, the group of J. D. Bergstrom and K. E. Wilson isolated the alkyl citrate, L-731,120 **90**, as a minor compound of the fermentation broth of fungus MF5453 and showed that **90** is a micro-molar inhibitor of squalene synthase.¹¹⁹



Scheme 2.2 Proposed mechanisms (route A and B) for bicyclic ring formation by Glaxo and Merck, adapted from Bergstrom *et al.*, 1995.¹¹³

Gene cluster and proposed biosynthesis of SQS1

The investigation of the SQS1 gene cluster started in our group in 2000 when isolation and sequence of the gene encoding squalestatin tetraketide synthase (*pks1*) were reported.¹²⁰ A few years later the *pks1* gene was cloned and heterologously expressed in *A. oryzae* and the produced tetraketide was isolated from liquid culture.⁴¹ Recently, an important step was done in understanding SQS1 biosynthesis by generating information using full genome sequencing which are reported by Bonsch *et al.*, 2016.⁵³

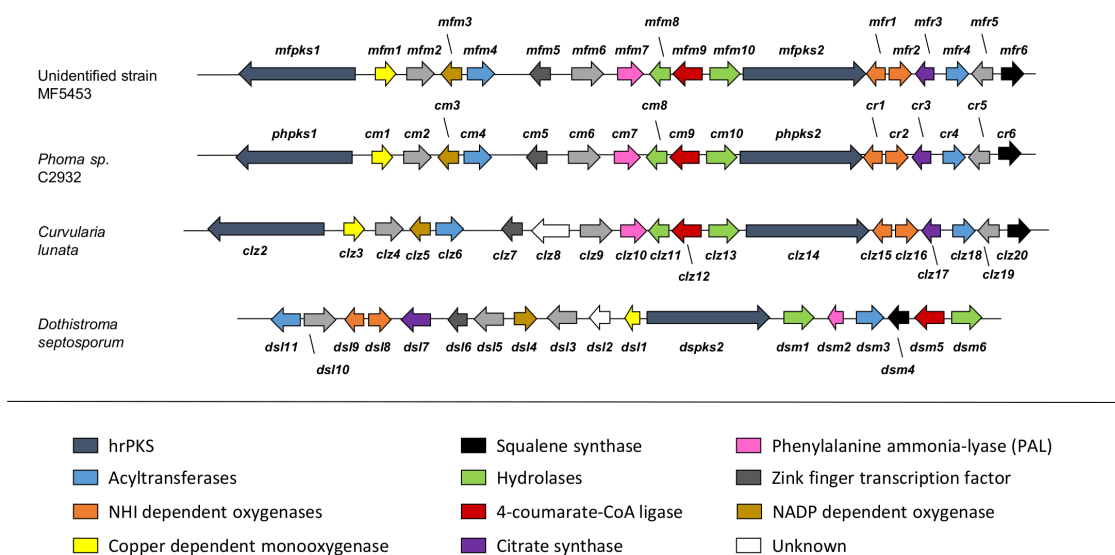


Figure 2.4 Three SQS1 gene clusters from unidentified strain MF5453, *Phoma sp.* C2932 and *Curvularia lunata* in comparison to a similar cluster found in *Dothistroma septosporum*; putative function as non-heme iron-(NHI) dependent oxygenases and copper dependent oxygenase were not determined before start of the project (Section 2.4).

The genome of two closely related fungal producers of **1** (*Phoma sp.* C2932 and unidentified strain MF5453) were sequenced. For both fungi a putative SQS1 gene cluster was found by analysing the genomes for the previously characterised *pks1* gene (Fig. 2.4).¹²¹ Both SQS1 clusters from MF5453 and C2932 show highly similar composition of genes and a high identity (74–93%) between proteins with the same predicted function (Table 2.1). For further studies the sequence of MF5453 was chosen. Recently, Tang and coworkers reported a BGC for production of **1** from fungal pathogen *Curvularia lunata* which also shows nearly identical gene organisation. The only difference is displayed as an additional gene *clz8*, which encodes for a protein with unknown function (Fig. 2.4).¹¹⁸

Using an NCBI database search a fourth putative squalenstatin BGC in pine pathogen *Dothistroma septosporum* was discovered.⁵³ In contrast to the highly similar order of genes in the first three clusters, the genes in *D. septosporum* are disordered and a corresponding gene for the tetraketide synthase (*pks1*) is missing (Fig. 2.4). Nevertheless, the composition of genes is similar and the identity between proteins with the same predicted function is still up to 73% (Table 2.1).

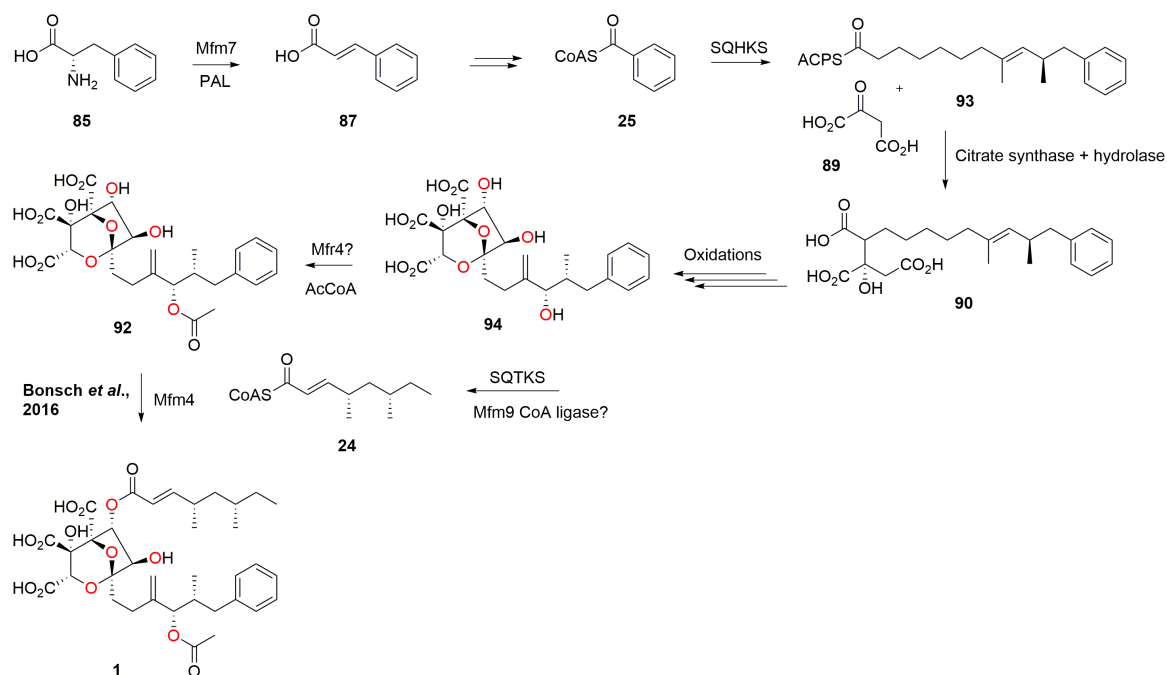
In order to link the clusters with the biosynthesis of **1** the gene for the tetraketide synthase was knocked out in the unidentified strain MF5453, resulting in loss of SQS1 production and accumulation of an intermediate **92**, lacking the tetraketide chain.⁵³

Unsurprisingly, two genes encoding highly reducing PKS were found in the first three SQS1 gene clusters: *pks1* and *pks2* (Table 2.1). It was shown before that the tetraketide synthase is responsible for the biosynthesis of the side chain of **1** by reconstruction of the 8 kb gene in a fungal expression vector and heterologous expression in *A. oryzae* M-2-3 (*argB*).^{41,123}

Table 2.1 Gene composition with putative function of SQS clusters of three producing fungus (MF5453, C2932, *C. lunata*) and of one putative SQS cluster of *D. septosporum* - identity is referred to MF5453 genes; putative function as non-heme iron-(NHI) dependent oxygenases and copper dependent oxygenase were not determined before start of the project (Section 2.4)

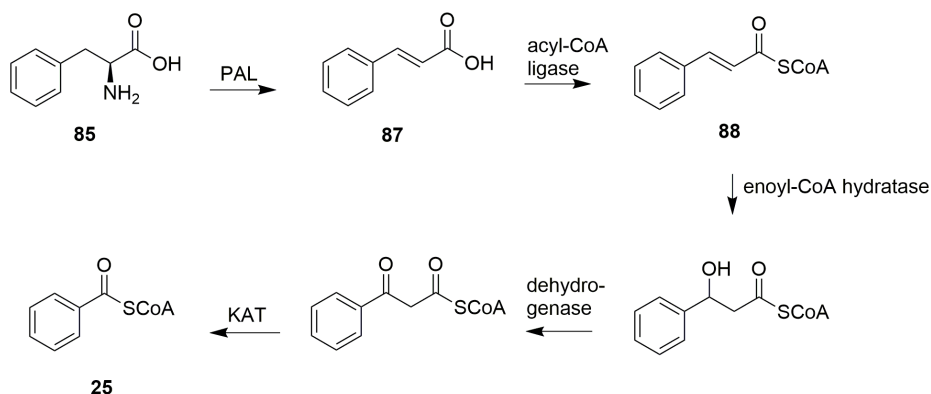
MF5453 genes ⁵³	C2932 genes ⁵³	aa % identity	<i>Curvularia lunata</i> genes ¹¹⁸	aa % identity	<i>Dothistroma septosporum</i> NZE10 ¹²²	aa % identity	Putative function
<i>mfpks1</i>	<i>phpks1</i>	87	<i>clz2</i>	88	N/A	N/A	Polyketide synthase (SQTKS)
<i>mfn1</i>	<i>cm1</i>	93	<i>clz3</i>	90	<i>dsL1</i>	32	Copper dependent oxygenase
<i>mfn2</i>	<i>cm2</i>	91	<i>clz4</i>	86	<i>dsL3</i>	64	MFS transporter
<i>mfn3</i>	<i>cm3</i>	80	<i>clz5</i>	86	<i>dsL4</i>	57	NADP dependent dehydrogenase
<i>mfn4</i>	<i>cm4</i>	89	<i>clz6</i>	88	<i>dsL11</i>	47	Acyltransferase (AT1)
<i>mfn5</i>	<i>cm5</i>	82	<i>clz7</i>	85	<i>dsL6</i>	13	Zinc Finger Transcription Factor (TF)
<i>mfn6</i>	<i>cm6</i>	81	<i>clz9</i>	94	<i>dsL5</i>	61	MFS transporter
<i>mfn7</i>	<i>cm7</i>	88	<i>clz10</i>	90	<i>dsM2</i>	17	Phenylalanine ammonia lyase (PAL)
<i>mfn8</i>	<i>cm8</i>	90	<i>clz11</i>	93	<i>dsM1</i>	62	Hydrolase
<i>mfn9</i>	<i>cm9</i>	78	<i>clz12</i>	92	<i>dsM5</i>	61	4-coumarate-CoA ligase
<i>mfn10</i>	<i>cm10</i>	78	<i>clz13</i>	84	<i>dsM6</i>	53	Hydrolase
<i>mfpks2</i>	<i>phpks2</i>	93	<i>clz14</i>	91	<i>dspks2</i>	67	Polyketide synthase (HRPKS)
<i>mfr1</i>	<i>cr1</i>	93	<i>clz15</i>	68	<i>dsL9</i>	73	Possible non-heme iron dependent oxygenase
<i>mfr2</i>	<i>cr2</i>	92	<i>clz16</i>	93	<i>dsL8</i>	61	Possible non-heme iron dependent oxygenase
<i>mfr3</i>	<i>cr3</i>	82	<i>clz17</i>	88	<i>dsL7</i>	66	Citrate synthase
<i>mfr4</i>	<i>cr4</i>	76	<i>clz18</i>	76	<i>dsM3</i>	37	Acyltransferase-3 superfamily (AT2)
<i>mfr5</i>	<i>cr5</i>	86	<i>clz19</i>	74	<i>dsL10</i>	49	MFS transporter
<i>mfr6</i>	<i>cr6</i>	74	<i>clz20</i>	87	<i>dsM4</i>	56	Squalene synthase

The second PKS gene is therefore likely to encode the squalenstatin hexaketide synthase, which builds the main polyketide chain. As shown by earlier isotopic labelling experiments, the hexaketide is formed from benzoate, an unusual starter unit for hrPKS (Fig. 2.2). The active site residues of the AT domain were shown to differ from hexaketide synthase AT domains which select acetate as the starter unit.^{53,118}



Scheme 2.3 Proposed biosynthetic pathway for squalestatin S1 **1**. Oxygen atoms marked in red derived from atmospheric oxygen.

Moreover, the SQS1 cluster was found to contain genes putatively involved in benzoate production: a phenylalanine ammonia lyase (PAL), encoded by *mfm7*, most likely catalyses the first step in the degradation of phenylalanine **85** (Scheme 2.3). This is supported by former feeding studies as phenylalanine **85** and benzoic acid **86** are known precursors of **1** (Section 2.1.2).⁷ This first step is known from other microbial and plant systems during the biosynthesis of benzoyl-CoA. In plants, cinnamyl CoA **88** is hydrated, oxidised and then subject to a retro-Claisen reaction (Scheme 2.4).^{124,125} However, recent feeding studies conducted by Tang and coworkers suggests that cinnamoyl CoA is formed early in the SQS precursor biosynthesis.¹¹⁸



Scheme 2.4 Biosynthetic steps from benzoic acid **86** to benzoyl CoA **25** in plants, KAT: 3-ketoacyl-CoA thiolase, PAL: Phenylalanine ammonia lyase.^{124,125}

Other proteins encoded by genes of the cluster may also be involved in the benzoyl CoA synthesis; for example Mfm3, a NADP-dependent dehydrogenase and CoA ligase Mfm9.

The produced hexaketide chain **93** is then predicted to be linked to an oxaloacetate moiety by a gene encoding a citrate synthase-like protein (encoded by *mfr3*, Scheme 2.3). The linkage of a polyketide chain to an oxaloacetate moiety by a citrate synthase has already been observed in fungal maleidride biosynthesis (Section 2.1.3). From alkylcitrate **90** a series of oxidation steps are necessary to form **1**, but significantly the BGC appears to encode no oxygenases by BLASTp analysis.⁵³

The cluster contains two putative acyltransferases (AT) genes. Previous work in our group concentrated on analysis of the two acyltransferases Mfm4 and Mfr4. *In vitro* analysis of Mfm4 by Verena Belt showed that this enzyme has a broad substrate selectivity in terms of its CoA substrate which was loaded onto the squalestatin core as the final step of biosynthesis (Scheme 2.3).⁵³ The other acyltransferase Mfr4, encoded by *mfr4*, could not be isolated in soluble form by expression in *E. coli*.⁵³ It presumably plays a role in attaching the acetate moiety to the hexaketide chain of **94** to form **92** (Scheme 2.3). As the *in vitro* experiment showed that the tetraketide chain could only be transferred to the core of **1** as CoA activated thiolester, the CoA ligase Mfm9 from the cluster (encoded by *mfm9*) may play a role in formation of the required tetraketide CoA. It is also possible that Mfm9 catalyses a step in the benzoate precursor biosynthesis. As Mfm9 is the only CoA ligase in the cluster it is possible, that this enzyme catalyses both reactions. In addition the squalestatin BGC encodes two possible hydrolases (Mfm8 and Mfm10) which could be involved in benzoate biosynthesis, off-loading of the hexaketide chain from the hexaketide synthase or in releasing the activated tetraketide chain from its synthase.

In addition to the biosynthetic genes, the cluster also contains genes for transporters (*mfm2*, *mfm6*, *mfr5*), an unknown gene *mfm5*, possibly a transcription factor in accordance to structure homology searches, and a putative resistance gene, *mfr6*, encoding a squalene synthase (SS) which showed mutations in substrate binding residues compared to housekeeping SS suggesting a role in self-resistance (Fig 2.4).⁵³

2.1.3 Maleidrides biosynthesis

Maleidrides, previously known as the family of carbocyclic maleic anhydrides, have been known for almost 90 years.¹²⁶ The first isolated carbocyclic maleic anhydride was glauconic acid **95** (*Penicillium purpurogenum*) in 1931.¹²⁷ In 1933, Raistrick and Smith reported byssochlamic acid **96** from *Byssochlamys fulva*.¹²⁸ Most family members contain two maleic anhydride moieties and were classified on the basis of their central ring. For example nonadrides such as **96**, octadrides such as

zopfiellin **97**¹²⁹ and heptadrides such as agnestadride A **98**¹³⁰ have been discovered with 9-, 8- and 7-membered central rings respectively (Fig. 2.5).

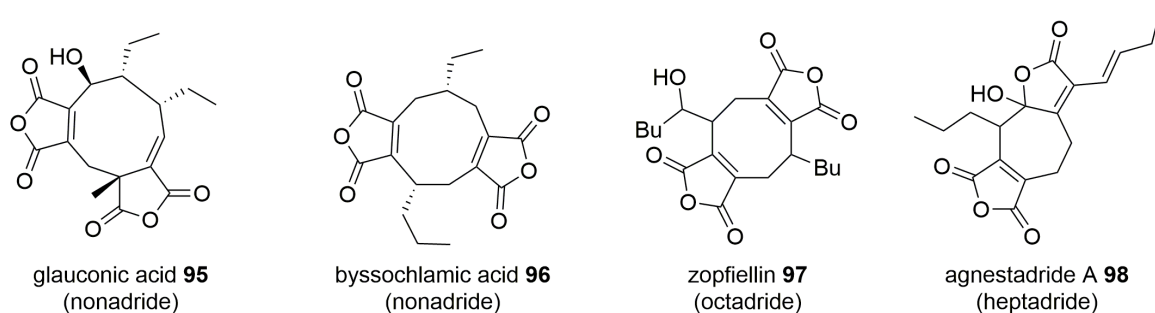
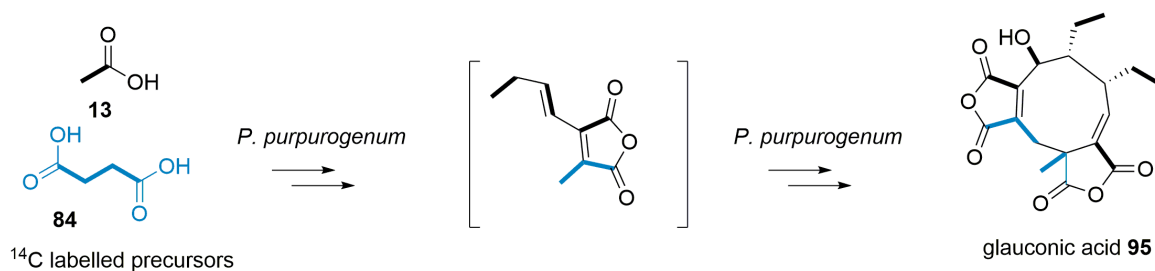


Figure 2.5 Structures of key maleidrides.

An initial proposal of the biosynthesis of byssochlamic acid suggests that a polyketide or fatty acid is linked to oxaloacetate to yield a citrate-like intermediate. This step is also proposed in SQS biosynthesis (Scheme 2.3). In maleidrides biosynthesis, the alkylcitrate dehydrates to give the observed maleic anhydride byssochlamic acid.¹³¹ Feeding studies with [1-¹⁴C]- and [2-¹⁴C]-acetate by Sutherland and coworkers showed incorporation of three acetates in each C₉ monomer in a head-to-tail fashion during biosynthesis of **95**. In addition, feeding of [2,3-¹⁴C₂]-succinate suggests that the remaining C₃ unit derives from oxaloacetate (Scheme 2.5).¹³²

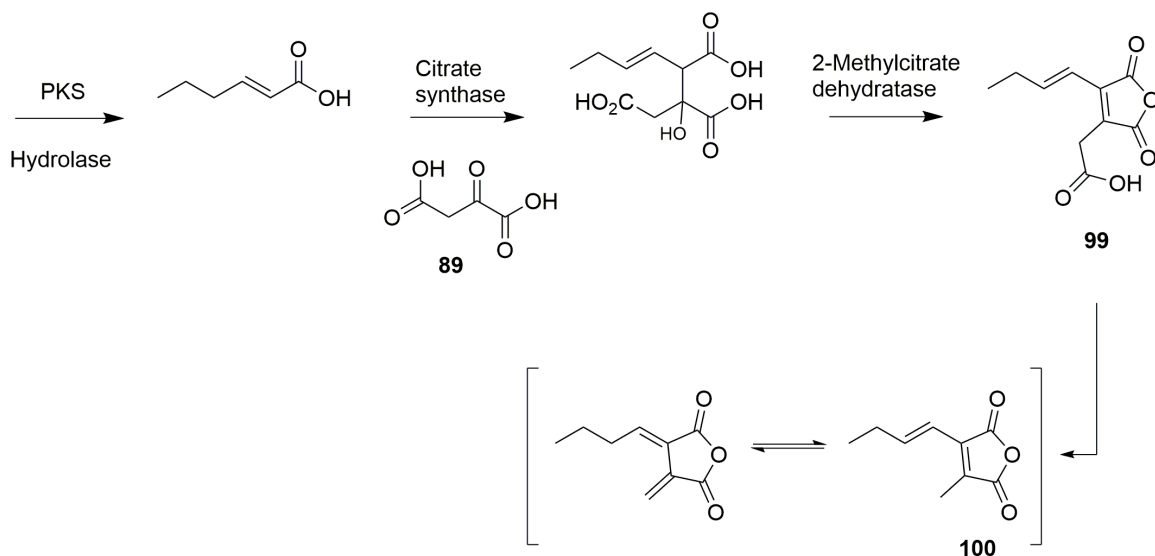


Scheme 2.5 Biosynthetic routes to maleidrides by Sutherland.¹³²

Reinvestigation of the original *B. fulva* strain showed the presence of several related compounds to **95**, such as heptadride **98**, which was shown to derive from the same pathway as byssochlamic acid. Production of a carboxymethyl maleic anhydride **99** was also reported which was shown to be unstable to spontaneous decarboxylation in solution to yield the methyl maleic anhydride **100** (Scheme 2.6).¹³⁰ In accordance to squalestatins, maleidrides, such as **96** derive from condensation of a polyketide and oxaloacetate in a citrate synthase-like process.¹²⁶

Full genome sequencing and analysis of the transcriptomic data under producing and non-producing conditions revealed a candidate BGC encoding a fungal hr-PKS and a citrate synthase (CS), a hydrolase, a 2-methyl-citrate dehydratase (2MCDH) and several proteins of unknown function. A

targeted knockout experiment of the core PKS gene linked the BGC to the production of **96**, **98** and also to production of **100**.⁸⁴



Scheme 2.6 First steps in the biosynthesis to maleidrides.⁸⁴

Our group and Oikawa and coworkers simultaneously reported that coexpression of the PKS, hydrolase, CS and 2MCDH encoding genes in a stepwise manner in fungal host *A. oryzae* led to production of the carboxymethyl maleic anhydride **99**. In detail, the PKS produces polyketides with various chain length and pattern of reduction, is then presumably released from the PKS and linked to an oxaloacetate catalysed by the CS-like enzyme. Subsequent dehydration by the 2MCDH homologue results in maleic anhydride decarboxylated monomers **100** (Scheme 2.6).^{84,133} Further coexpression with genes of unknown function from the BGC (two putative ketosteroid isomerases (KSI) and two putative phosphatidylethanolamine binding proteins (PEBP)) showed an increased titre of **99** and **100** and also produced **96** (with addition of two KSIs) and **98** (with addition of both KSIs and PEBPs) for the first time.⁸⁴

2.2 Project aims

Although the biosynthetic precursors of SQS1 have been determined by labelling studies, nothing is known of the precise molecular steps of its construction. The precursor of the hexaketide is benzoate, derived in-turn from phenylalanine, but the molecular steps are unknown. Moreover, the oxidation pattern of the hexaketide synthase product is not yet known. One of the major aims is to identify gene candidates encoding oxygenases involved in the oxidative steps to form the complex core of squalenolactone and find a gene responsible of introducing the hydroxyl group at the hexaketide chain for further acetylation. One of the two acyltransferases of the cluster, Mfr4, could not be characterised *in vitro* and therefore should be investigated *in vivo* to verify its proposed role in attachment of the acetyl side chain onto the hexaketide chain.

Initial *in silico* analysis of the unknown genes in SQS BGC should reveal putative candidates for oxygenases by using a structural homology server Phyre2.¹³⁴ Using targeted knockout experiments in the SQS producing fungus MF5453 should uncover the oxygenases responsible for building the characteristic core of **1** and introducing the hydroxyl group at the hexaketide chain for later acetylation. In addition the knockout experiments should delineate the key steps of the biosynthesis and determine the role of the second acyltransferase.

Further, *in vivo* reconstitution of the biosynthesis using heterologous expression experiments should then be deployed to examine these transformations in more detail using *A. oryzae* as the host. Oxygenated congeners of **1** accumulating upon knockout and expression experiments should be isolated and analysed using high resolution mass (HRMS), MSMS analysis and nuclear magnetic resonance (NMR) techniques in order to elucidate the structures and unveiling the steps to form squalenolactone.

2.3 Results – In silico analysis of putative oxygenases from SQS BGC

The BGC of SQS in MF5453 consists of 22 contiguous genes. Initial bioinformatic studies revealed only one gene (*mfm3*) encoding a oxidoreductase despite the fact that several oxidative steps must be required to form the core of **1** (see labelling experiments, Fig. 2.2).

2.3.1 Putative non-heme iron-dependent oxygenases Mfr1 and Mfr2

The gene *mfm3* encodes an NAD(P)H-dependent short-chain dehydrogenase/reductase (SDR) which is unlikely to be able to oxygenate.⁵³ BLASTp analysis of Mfr1 and Mfr2 showed no significant homology to proteins of known function.¹³⁵ Nevertheless, more detailed structural analysis using the protein fold recognition server PHYRE2¹³⁴ suggested that *mfr1* might encode a non-heme iron protein as it shows some structural similarity (although less than 20% sequence identity) to CytC3 which is characterised as non-heme Fe(II) and α -ketoglutarate-dependent halogenases involved in halogenation of aminobutyric acid in soil *Streptomyces*.¹³⁶

The gene *mfr2* is 41.1% identical to *mfr1* but showed even lower similarity to other known proteins. But due to its high sequence homology to Mfr1, this pair of enzymes might be involved in oxygenation to form the core of SQS1.

R1	--MATAILPSTSGVIGLWDGTTDGKEGFMDYANGDTNVKQPKEYEIQVHDIRKLDPQPTL	58
R2	MATATTTLHSTTGTVYVADGTTDGKVGYYNHTDDSTNVI-RKPIPIEVEDARTLSKSPTT	59
	: * **:*.: : ** *: : : : : *** * * * : * . * * . * . * . * . *	
R1	LKNGYELVDIPTVVTDEQFIESGKSDEGNAYIKDVYFAECKRIIEEVSGGVDLIIPVSFR	118
R2	KAEGYQLVNFHTKIPEEHFLNSK-LPENKELIEEVYFDECRRLVQEVTGAAEA-YPYVYR	117
	:**:*:*.: * : :*:***:* *.: * :*:*** **:*:***:***.*.: * :*	
R1	MREQKGEKESTTKLGNIESRYAPRPVAHLLDRDTPAITVLEETVGKEKAQELLSKHKRW	178
R2	VRNQEQNAKES--NKSNFH--TDFVPIVHVDRDDVTAPQRLRASLGAEKADMLLSKYKSY	173
	:*:*:* : :.: : .*: . * : * : * * * * * * . : : * * * * : * * * * : * : *	
R1	AQVNVWRPIGNPATMWPLCFLNHDRIPTWNYDTHVGHVWSLNDPRVSDRGQKTYDCVVKH	238
R2	GSINVWRPVKNMVQKWPMLMLVDHKSIEDWDYSTHMFTLHSSNDERVATRGAKEHETILTH	233
	..:*****: * . *** :*:*. * * :*:***: : * * * * * : * * * : : : : *	
R1	DDRYDYHYVSDLRPEECLVFCSFDSIPKYAMPHSAFWDNNVPADAPNRRSIEVRSLVFF-	297
R2	DKRYRYIYASDMTPEEAWLFFAFHSDPALGIPHGAFWDDSTKEEALTRCSIEVRIWVFFD	293
	*.*** * *.**:* ** . :* :*. * * . :*:.*****:.. :* . * ***** ** *	

Figure 2.6 Protein sequence alignment of Mfr1 and Mfr2 using Clustal Omega server¹³⁷, Fe-binding sites are marked in yellow.

By protein sequence alignment of Mfr1 and Mfr2, using Clustal Omega^{137,138}, a correlating triad (H149, D151, H240) was observed (Fig. 2.6, shown in yellow). Nevertheless, an alignment in

comparison with other known non-heme iron-dependent oxygenases (e.g. AndA and AndF from the biosynthesis of meroterpenoid anditomin¹³⁹ (Section 2.5.8) was not possible. The sequences seemed to be remarkably different, although the four proteins were built from nearly the same number of amino acids (aa) (R1 = 297 aa, R2 = 293 aa, AndA = 293 aa, AndF = 285 aa).

As other sequence homology comparisons with Mfr1 and Mfr2 only revealed proteins of unknown function it would be likely that both might be new types of enzymes, related to non-heme iron-dependent oxygenases, with oxidative function in SQS1 biosynthesis. Therefore Mfr1 and Mfr2 are possible targets for further knockout studies to uncover the oxidation steps in **1** biosynthesis.

2.3.2 Putative copper-dependent oxygenase Mfm1

Initial analysis using BLASTp¹³⁵ indicated no sequence homology of Mfm1 to a protein with known function was found. More sophisticated studies using a structural server (PHYRE2¹³⁴) showed that Mfm1 possess modest structural homology (15% identity) to a peptidylglycine α -hydroxylating monooxygenase (PHM) from *Rattus norvegicus* (PDB: 1OPM)^{53,140} The central region of Mfm1 (51% of the residues) were modelled with over 90% confidence and showed putative copper binding sites.

PHM (EC 1.14.17.3) is one of the catalytic domains of a peptidylglycine α -amidating monooxygenase (PAM), next to a peptidyl- α -hydroxyglycine amidating lyase domain. PAM is normally found in higher eukaryotes and catalyses the carboxy-terminal α -amidation of bioactive peptides.¹⁴¹ The PHM domain is a dicopper dependent monooxygenase, responsible for the α -hydroxylation of the peptidylglycine substrate.¹⁴²

In previous attempts to express Mfm1 heterologously in *E. coli* during a master's project of Annika Stein, *in silico* analysis of Mfm1 indicated a signal peptide (1–30 AA) and a transmembrane domain in the C-terminal part of the protein (339–386 AA).¹⁴³ For further *in silico* analysis only the cytoplasmic part (31–338 AA) of Mfm1 was considered. A multiple sequence alignment (MSA) with HHpred¹⁴⁴ using known protein sequences with characterised tertiary structure elements from the PDB database and our target sequence, resulted in PHM from *Rattus norvegicus* (PDB:1YI9)¹⁴⁵ being the best template with 97.1% probability and an E-value of 0.0063 for further modelling analysis (Fig. 2.7 A). The structural model of Mfm1 was generated in cooperation with Michelangelo Marasco, using the MODELLER program (version mod9.15)¹⁴⁶ with a loop refinement script. The best model was selected for further analysis and is displayed in figure 2.7 B. Based on the MSA and the structural model, conserved motifs were shown for the first active site of Mfm1 that occupies two histidines (His-203, His-205) and a methionine (Met-282) in proximity to the copper ion, similar to the CuM site of the PAM (Fig. 2.7 C).¹⁴⁷ The second active site was found to possess two conserved histidines (His-71, His-72) as ligands (Fig. 2.7 C).¹⁴³

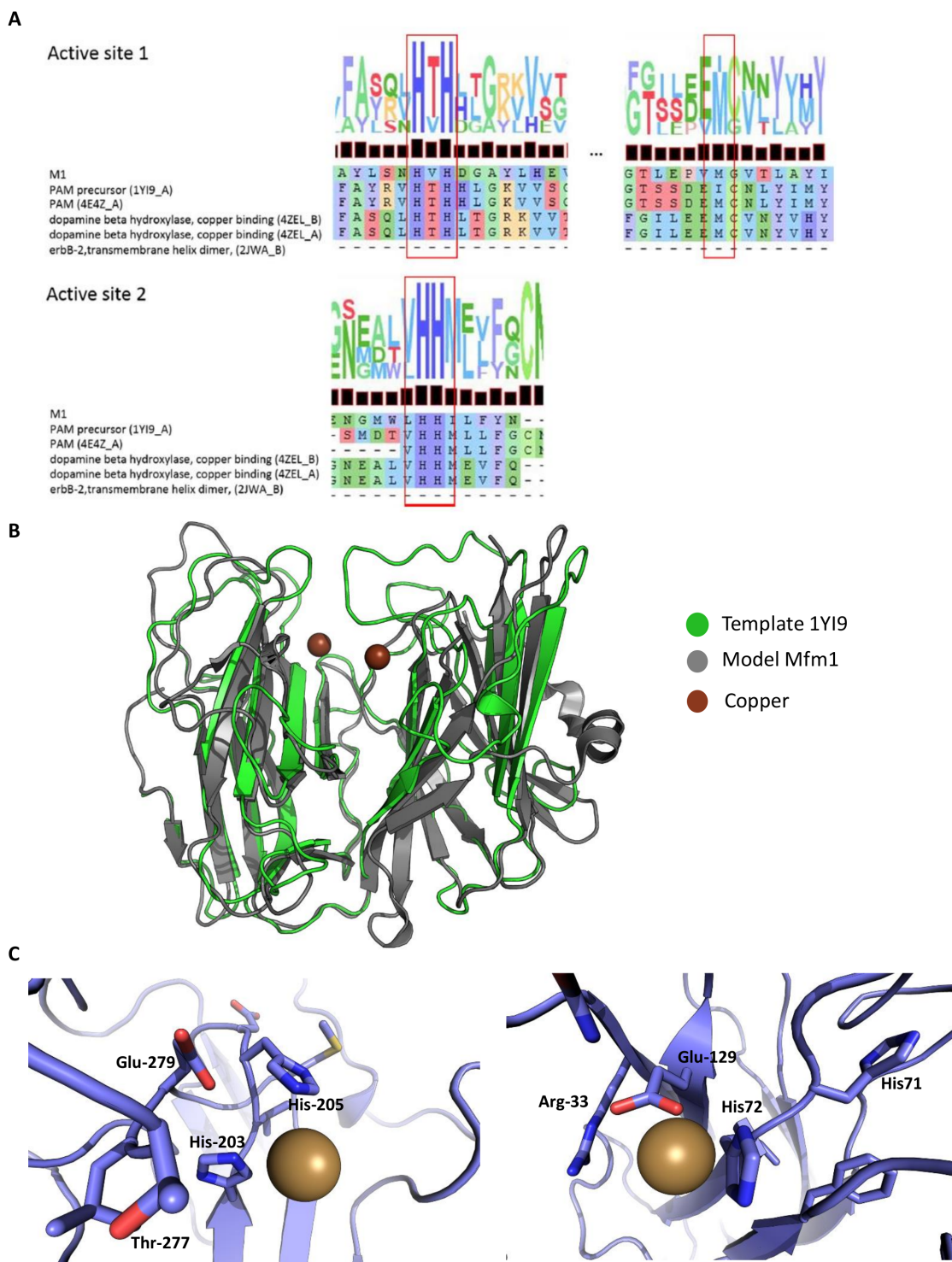


Figure 2.7 *In silico* analysis of Mfm1: **A**, MSA of Mfm1 with known copper-binding enzymes from the PDB database using web-based server HHpred¹⁴⁴; **B**, structural model of Mfm1 using MODELLER program (version mod9.15)¹⁴⁶ in cooperation with Michelangelo Marasco, with a loop refinement script (template: PDB:1YI9); **C**, catalytic sites showing His-203 and His-205 (left) and His-72 and His-73 (right) coordinating the copper ion (brown balls).

Overall, the *in silico* results suggests that Mfm1 is a copper dependent oxygenase which is possibly involved in one of the oxidation steps in **1** biosynthesis. Therefore the encoding gene *mfm1* is also a promising target for knockout studies.

2.4 Results – Analysis of squalstatin metabolites produced by wildtype fungus MF5453

The MF5453 producing strain was first reported to produce SQS1 in 1992 by the Merck group.^{12,47} A time course experiment showed the highest **1** production after 7 days of fermentation.⁴⁷ Similar results were also obtained by a former PhD student of our group (Beate Bonsch, unpublished data). It was shown that **1** was produced in YMG medium and CM2 medium, but not in LB medium (non-producing conditions).⁵³ Previous fermentation experiments of MF5453 in YMG medium showed not only production of **1** ($t_R = 8.6$ min), but also production of SQS metabolite **92** ($t_R = 5.4$ min) in our lab.⁵³

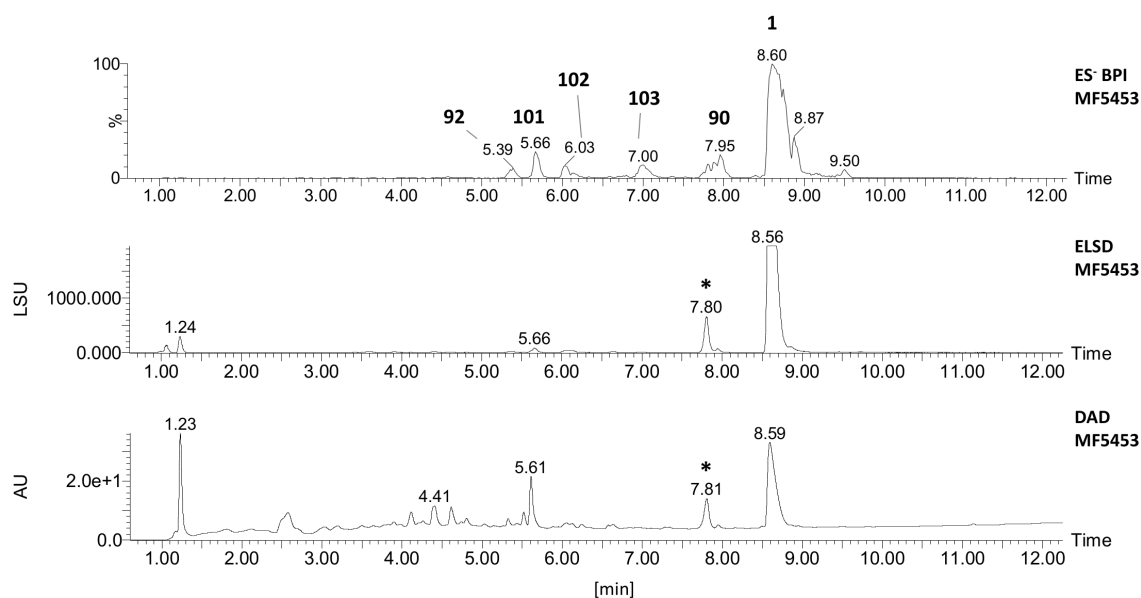


Figure 2.8 LCMS analysis of MF5453 wildtype extract obtained from fermentation experiments in YMG+TES, 7 days: DAD; ELSD; and ES⁻ BPI chromatograms showing production of major compound **1** and minor compounds related to **1**.

In this thesis the wildtype (wt) SQS1 producing strain MF5453 was commonly grown in YMG medium with addition of a trace element solution (TES), as it was shown to produce **1** in higher titres, compared to CM2 medium or YMG medium alone. Squalstatin S1 **1** was shown to be the major product of the extract obtained after 7 days of fermentation (Fig. 2.8) and was confirmed by retention time $t_R = 8.6$ min, the characteristic UV maximum at 210 nm and its distinctive ion 689.6 (negative

2.4 Results – Analysis of squalstatin metabolites produced by wildtype fungus MF5453

mode, ES^-), as previously reported by Bonsch *et al* (Fig. 2.9).⁵³ In general **1**, and all intermediates were known to ionise better in the negative mode and were usually not, or weakly, detected in the UV trace (DAD).

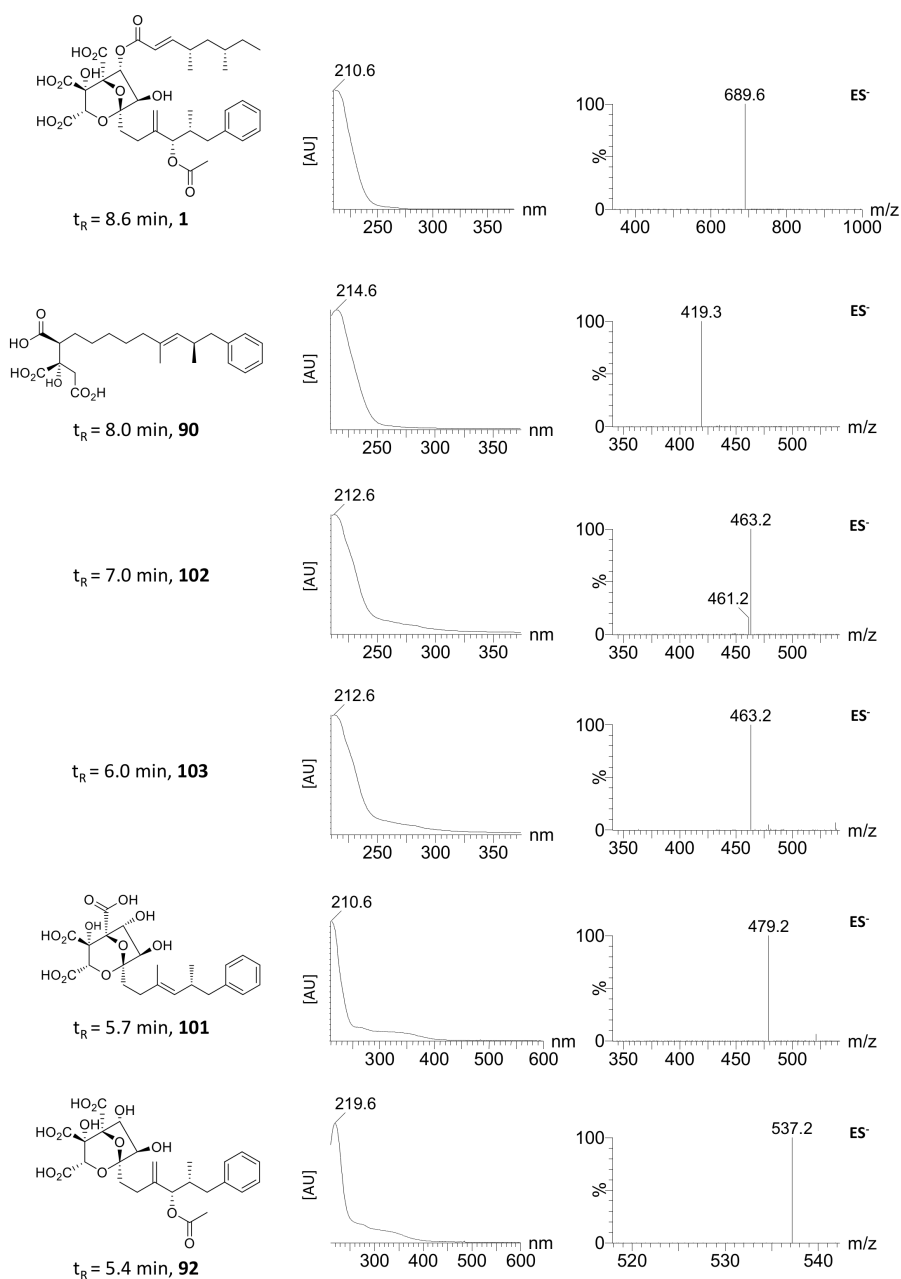


Figure 2.9 Chemical analysis of wt metabolites of MF5453 including retention times, UV and ES^- BPI chromatograms for all compounds produced from MF5453 wt fungus.

Further analysis of the wt extract also revealed production of some minor compounds **90**, **92**, **101** and **102** + **103**, which are related to **1** (Fig. 2.8). All compounds share a common UV spectrum with an absorption maximum of 210–220 nm (Fig. 2.9). Compound **92** and **90** were previously isolated and characterised.^{53,119} However, due to low titres of **90** in the wt extract (Fig. 2.8, t_R = 8.0 min, not

observed in ELSD trace), the compound was first isolated and identified as alkyl citrate L-731,120 **90** from the later-created *mfr1+2* knockout strain (Section 2.5.3).

Similarly, compound **101** was first isolated from an *mfm1* knockout strain and its structure elucidated (Section 2.5.4). Compounds **102** and **103** were also detected in extracts of the same *mfm1* knockout strain. As they share the same mass and UV spectrum they are most likely isomers (Fig. 2.9). Full structure elucidation was not possible for these compounds due to instability problems and low titres. However, MSMS analysis and some structural features such as the phenyl ring and the dimethylated triketide were observed in the ^1H NMR (Section 2.5.4).

Another compound coeluting with **90** ($t_R = 7.8$ min, Fig. 2.8) was shown to be unrelated to SQS1 on the basis of its divergent UV absorption maxima (224 and 275 nm) and its even mass in the positive mode suggesting a nitrogen atom in the chemical formula (Fig. 2.10).

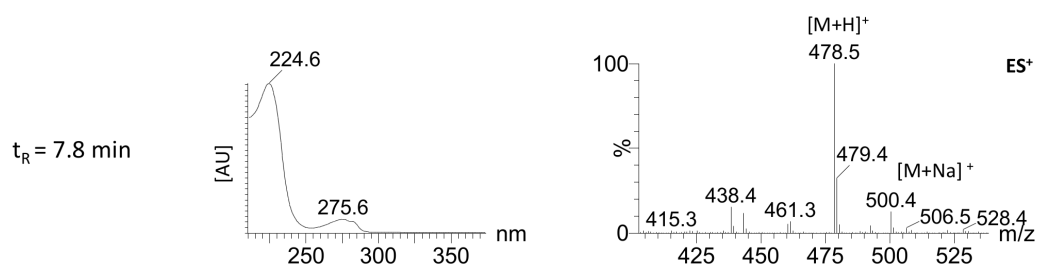


Figure 2.10 Chemical analysis of an unrelated compound produced by the SQS1 wt producing fungus MF5453; retention time, UV and ES⁺ BPI chromatograms.

2.5 Results – Knockout studies

The SQS1 producing strain MF5453 was shown to be successfully transformed within our group using a protoplast-based PEG transformation method and hygromycin B as selection. The knockout of the tetraketide gene (*mfpks1*) led to loss of SQS1 production and enhanced concentration of SQS1 precursor **92**. Applying the bipartite marker strategy three out of ten transformants showed the expected gene disruption.⁵³

In order to investigate the biosynthetic steps of SQS1 biosynthesis and to uncover the role of each gene, the following genes were targeted for disruption: *mfr3* - encoding a citrate synthase; *mfm3* - encoding an NADP dependent dehydrogenase; *mfr4* - encoding an acyltransferase; *mfm1* - encoding a putative copper dependent oxygenase; and the two similar genes *mfr1* and *mfr2* together as well as individually, which encode putative non-heme iron-dependent oxygenases.

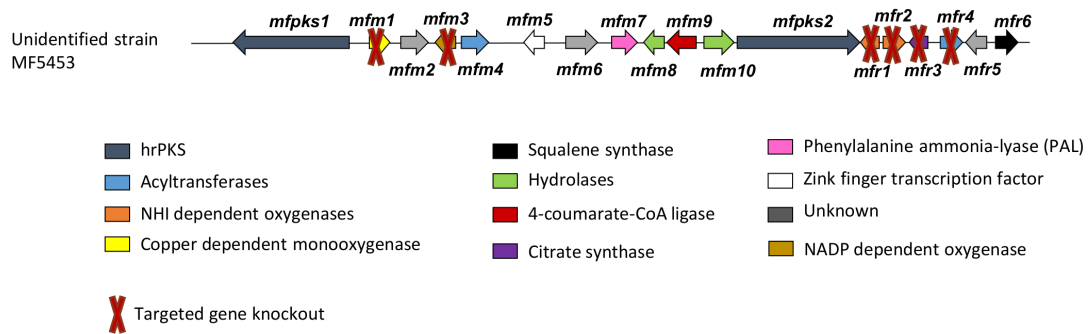
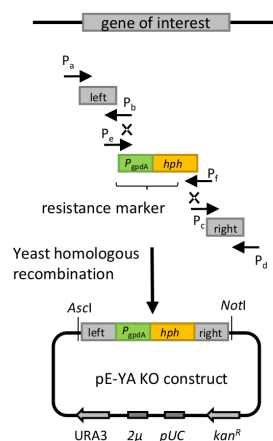


Figure 2.11 SQS1 gene cluster from unidentified strain MF5453 showing targeted gene disruption.

2.5.1 Plasmid construction using yeast recombination based cloning

In order to disrupt a GOI from SQS BGC, a knockout cassette consisting of a selection marker flanked by arms homologous to the targeted gene has to be assembled. For each GOI, a plasmid containing a knockout cassette was built using yeast recombination. *S. cerevisiae* is able to recombine DNA fragments with at least 30 bp sequence overlap by homologous recombination. Due to this ability all vectors were constructed by cotransformation of DNA fragments and a vector backbone into *S. cerevisiae* by the LiAc/SS carrier DNA/PEG method reported by Gietz *et al.*¹⁴⁸

The knockout plasmids were constructed by amplifying two homologous arms (left and right, ~500 bp each) of the target gene from genomic DNA of MF5453. The resistance marker, consisting of a fungal promotor P_{gpdA} from *A. nidulans* and the *hph* gene of *E. coli* which confers resistance to hygromycin B (*hygB^R*), were amplified from plasmid pTH-GS-eGFP (Scheme 2.7).



Scheme 2.7 Yeast homologous recombination based cloning strategy with oligonucleotides P_a – P_f used for PCR.

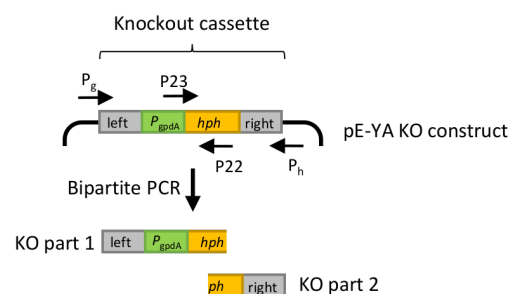
Sequence overlap of 30 bp for homologous recombination of the individual DNA fragments was introduced by PCR through designed oligonucleotides (P_a – P_f). All constructs were made using

combinations of oligonucleotides listed in Chapter 6 (Table 6.9). Corresponding oligonucleotide sequences are found in Table 6.8.

In this work entry vector pE-YA⁹⁸ was used to construct all KO vectors. Vector pE-YA consists of a 2μ origin of replication (*ori*), a *ura3* gene (encoding an orotidine 5'-phosphosphate decarboxylase) to select uracil auxotroph *S. cerevisiae* in uracil and uridine free medium. Furthermore it contains an additional *ori* (*pUC*) and selection marker (*kan^R*) for propagation in *E. coli*. pE-YA was linearised by restriction enzymes *NotI* and *AscI* and assembled with purified, amplified DNA fragments using a LiAc/SS carrier DNA/PEG method and *S. cerevisiae* as a host (Scheme 2.7).¹⁴⁹ Assembled vector DNA was isolated from *S. cerevisiae* using a commercial kit and was directly used to transform competent *E. coli* Top 10 cells. *E. coli* colonies were selected on LB medium supplemented with kanamycin (pE-YA also occupies the *nptII* gene, encoding the neomycin phosphotransferase II, which confers resistance to kanamycin, *kan^R*). Obtained colonies were analysed by colony PCR and confirmed by sequencing (Eurofins, Ebersberg).

2.5.2 Bipartite method and transformation of MF5453

The knockout plasmids constructed by yeast recombination were used as template for generating the two DNA fragments (KO part 1 and KO part 2) for the bipartite method (Scheme 2.8). The transformation protocol for strain MF5453 was developed previously by Beate Bonsch.⁵³ In this work, the protocol was optimised by changing minor components. Instead of growing the fungus for three days the time was reduced to 48 hours in order to produce finer mycelium resulting in a higher amount of protoplasts. Furthermore the washing steps for protoplasts were omitted in order to improve the yield. For each gene knockout 8 to 24 transformants could be generated using the optimised transformation protocol. Positive, proven knockout transformants were achieved with a frequency of 10 to 20% depending on the construct (Table 2.2).



Scheme 2.8 Bipartite targeted knockout strategy showing amplified parts (KO part 1 and KO part 2) for fungal transformation.

Table 2.2 Overview of transformants obtained by bipartite KO method and chemical or genetic analysis

Experiment ID	GOI	transformants	chemically analysed	genetically analysed	confirmed KO
KELI28	<i>mfr1+mfr2</i>	8	8	3	1
KELI38	<i>mfr1</i>	24	1	11	1
KELI39	<i>mfr2</i>	18	1	10	1
KELI27	<i>mfm1</i>	15	10	5	1
KELI29	<i>mfr3</i>	7	7	3	2
KELI30	<i>mfr4</i>	17	10	2	2
KELI48	<i>mfm3</i>	20	2	8	2

2.5.3 Knockout of non-heme iron-dependent monooxygenases Mfr1 and Mfr2

In silico analysis suggests that *mfr1* and *mfr2* encode non-heme iron-dependent oxygenases (Section 2.3). Due to the fact that *mfr1* and *mfr2* show high sequence similarity (51.6%, EMBOSS Needle Nucleotide Alignment¹³⁸) to each other and are closely located on the genome (403 bp in between) we decided to knockout these genes both together and individually using the bipartite method described in previous section 2.5.2.

Dual knockout of Mfr1 and Mfr2

Eight putative dual Mfr1 and Mfr2 knockout transformants were obtained from three independent transformations. For chemical analysis the eight transformants were grown under SQS1 producing conditions in parallel to a wt control. Each culture was extracted following a small scale (1 ml) protocol and analysed by analytical LCMS. Production of **1** was abolished in 7 out of 8 transformants, suggesting that in all but one transformant the targeted genes were disrupted. This implies that at least one of the two genes is essential for SQS1 production.

To confirm that the non-producing transformants are true knockout strains, the genomic DNA of three transformants (a1, c1, c2) was isolated. Oligonucleotides (P578+579) were used to amplify a region from the left part of *mfr1* to the right part of *mfr2*. The positive wt control as well as transformants c1 and c2 showed a band for *mfr1+r2* around the 2 kb standard ladder band, which corresponds to the size of intact *mfr1+r2* genes (Fig. 2.12), therefore suggesting that the gene disruption was not successful for transformant c1 and c2. Nevertheless, the *hph* gene must have integrated and recombined ectopically, as the strains were selected on hygromycin. The integration of the resistance gene might have taken place on a position that affects production of SQS1 and caused the abolition. It is also feasible that insufficient SQS1 was recovered due to the small scale extraction protocol and thus prevented its detection by LCMS analysis, but the wt control showed SQS production. Only transformant a1 was shown to be a true knockout transformant because of a higher band matching the size of the whole knockout cassette (5.2 kb) and no band could be observed corresponding to the intact genes (Fig. 2.12).

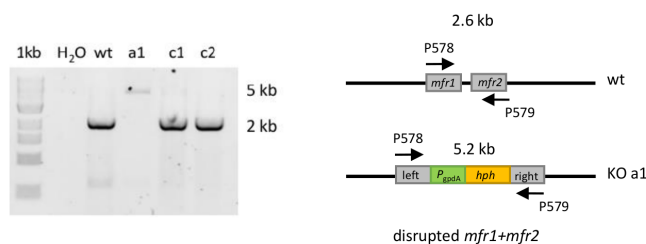


Figure 2.12 Genetic analysis of KO transformants a1, c1, c2, H₂O control and wt positive control; 1% agarose gel, 1 kb ladder; amplified *mfr1+2* gene (2 kb).

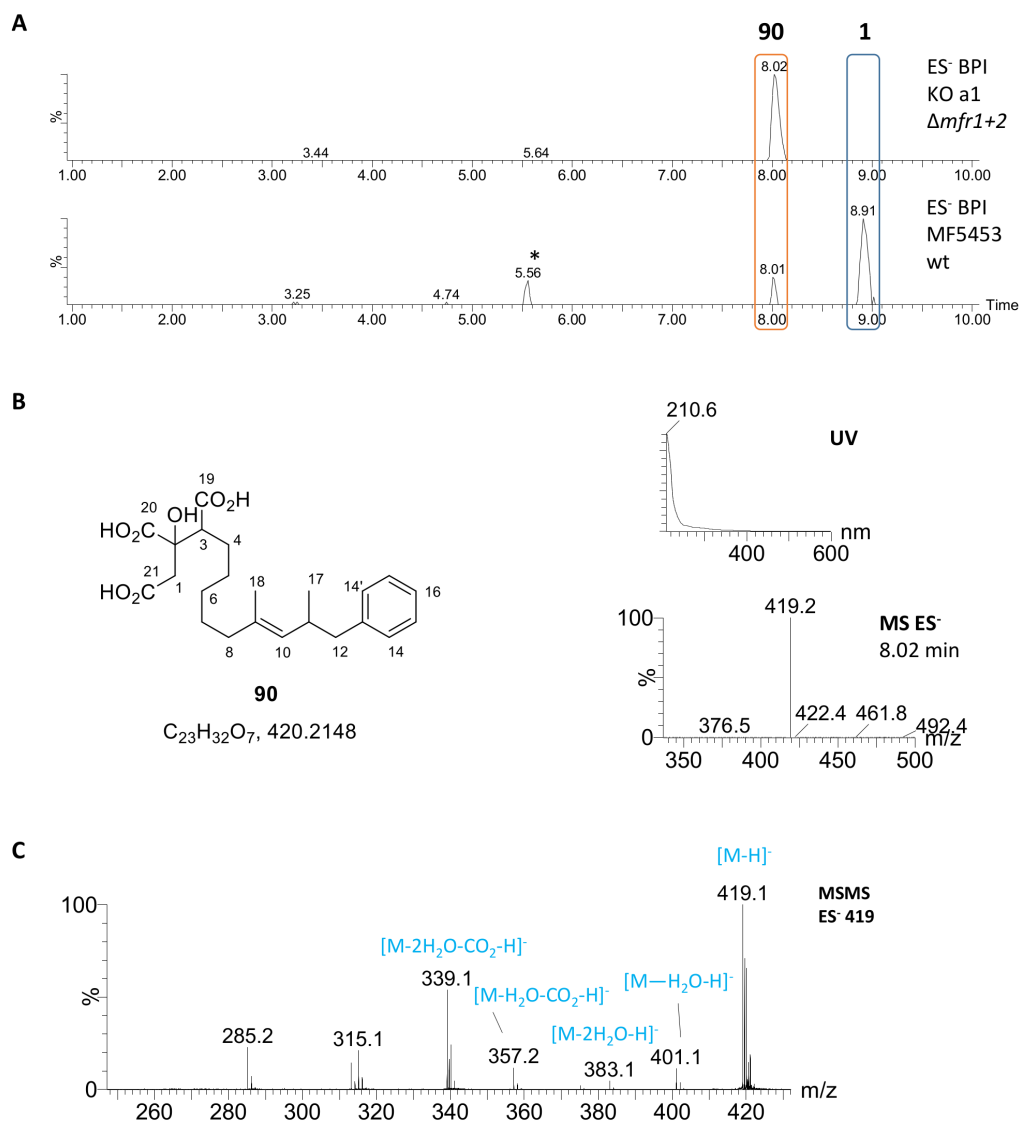


Figure 2.13 Chemical analysis of *mfr1+2* KO transformant a1: **A**, ES⁻ traces of extracts of KO transformant a1 and a wt control, in blue frame: SQS1 **1**, in orange frame: new compound **90**, *unrelated compound *m/z* 263, extracts were obtained from cultures grown under SQS1-producing conditions; and **B**, chemical characterisation of **90**: *t_R* = 8.0 min, UV spectrum, and mass spectrum; **C**, MSMS analysis.

Transformant a1 was grown in a larger scale (50 ml) in SQS producing medium. Squalestatin production was again shown to be abolished in the KO strain, as no peak at 8.9 min with the m/z of 689 (negative mode) corresponding to SQS1 was detected. In the KO strain, a compound eluting at $t_R = 8.0$ min with a nominal mass of 420 was accumulated in comparison to the wt control, where it was also detected in smaller amounts (Fig. 2.13 A, Section 2.4).

Structure elucidation of compound 90 In order to determine the structure of accumulated compound isolated from MF5453 $\Delta mfr1+2$, the fermentation of KO transformant a1 was scaled up to 8 x 100 ml, combined and extracted. From 270 mg crude extract, 5 mg of KO compound (m/z 419) was purified using preparative LCMS (peak eluting at $t_R = 8.5$ min, Fig. 2.14 A). Initial attempts to acquire suitable NMR data for structure elucidation was unsuccessful, as the compound was not pure enough. Impurities were identified by analysing the corresponding mass traces of the collected ELSD peak. Another unrelated compound "a" with $m/z = 478$ (ES^+) and $m/z = 420$ (ES^-) coeluted with desired compound 90 (Fig. 2.14 B).

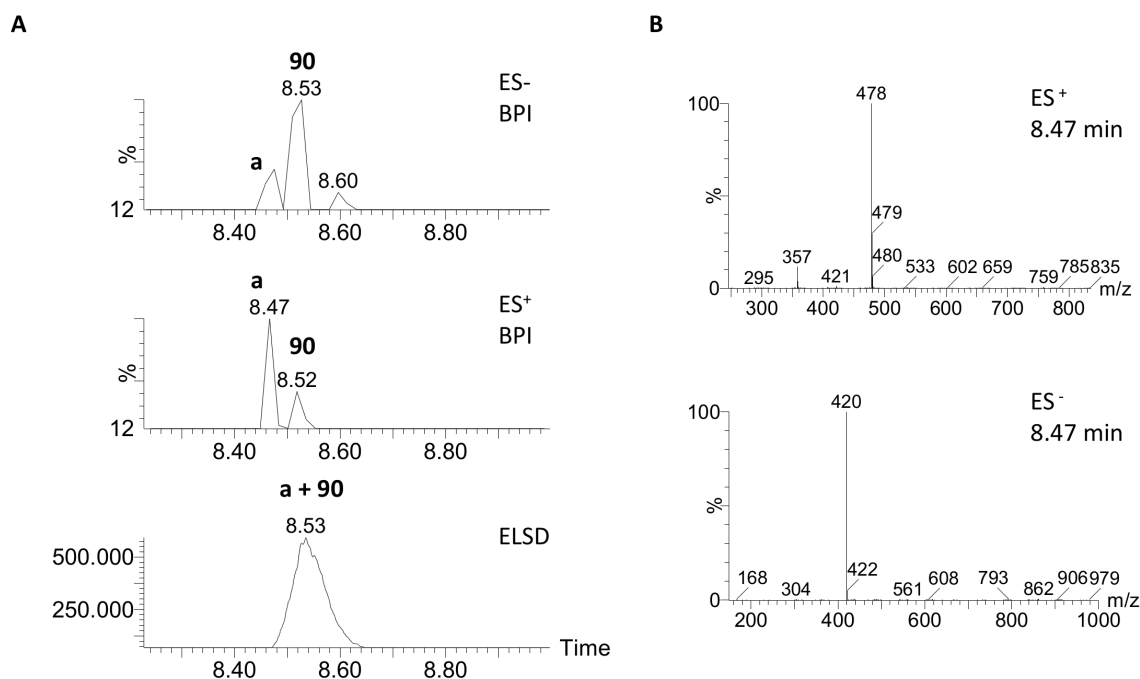
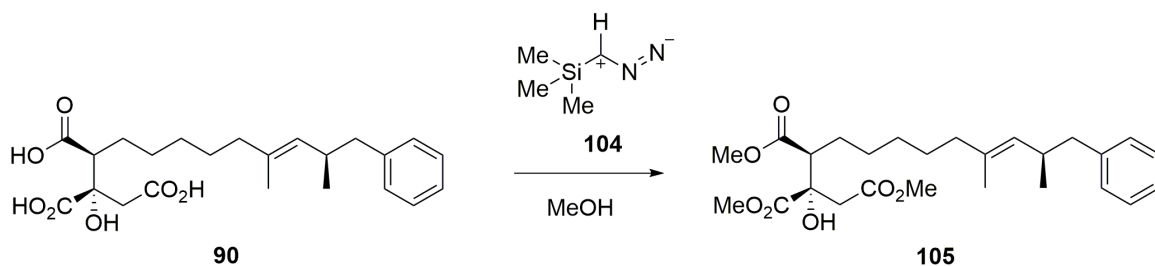


Figure 2.14 LCMS data obtained from crude extract of *mfr1+2* KO transformant: **A**, ELSD and mass traces showing major compound and impurity "a" eluting at $t_R = 8.5$ min on preparative column; **B**, mass spectra of impurity coeluting with 90.

In order to acquire good NMR data suitable for structure elucidation the mix was treated with trimethylsilyl (TMS) diazomethane **104** in methanol. Predicted carboxyl groups were converted to their corresponding methyl esters (Scheme 2.9).¹⁵⁰ The methylated compound **105** eluted at $t_R = 9.4$ min and could be separated from the impurity (eluting at $t_R = 8.5$ min) using preparative LCMS. Due to the preparative column the retention times vary from previous LCMS results obtained from

analytical runs. Methylation was confirmed by corresponding mass (Fig. 2.15). NMR data was acquired and analysis identified the structure of compound **105** was identical to the previously reported alkyl citrate L-731,120 (plus methyl esters) which was isolated as a minor component from fungus MF5453 by the group of J. D. Bergstrom and K. E. Wilson in 1995 (Section 2.4, Assignment in figure 2.16).¹¹⁹ This compound corresponds to the proposed (Scheme 2.3) product of SQHKS linked to citrate and is thus likely to be an early intermediate.



Scheme 2.9 Methylation of **90** with TMS diazomethane (**104**) in methanol yielded product **105**.

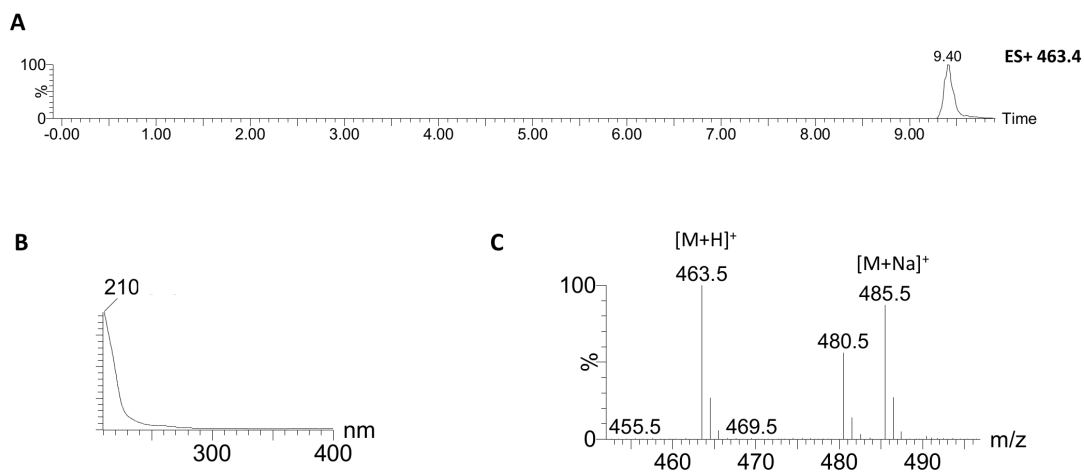
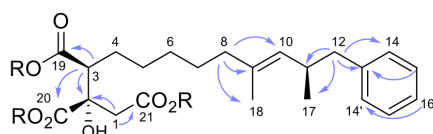


Figure 2.15 LCMS data for compound **105**: A, Extracted ion chromatogram (ES⁺ 463.4); B, UV chromatogram; C, mass spectrum (ES⁺).

A

**90** R = H**105** R = CH₃

B

Pos	Compound 90 Literature ¹¹⁹ CD ₃ CN		Compound 105 CDCl ₃				
	δ_c / ppm (125 MHz)	δ_H / ppm (500 MHz)	δ_c / ppm (125 MHz)	δ_H / ppm (500 MHz)	<i>J</i> / Hz	¹ H- ¹ H COSY	HMBC
1	41.8	2.65 (d, 16.5) & 3.02 (d, 16.5)	40.9	2.72 & 3.10	d, 16.4 & d, 16.4	-OH	2, 20, 21
2	76.5	-	76.1	-	-	-	1, 3, OH
3	53.9	2.60 (m)	53.9	2.69	dd, 11.8, 3.1	-	2, 19, 20
4	27.9	1.40 (m) & 1.66 (m)					
5	28.1	1.22 (m)					
6	29.3	1.22 (m)	27.7 / 27.8 / 29.8 / 29.9	1.18-1.34	m, 8H	-	-
7	28.2	1.22 (m)					
8	40.0	1.86 (brt, 7)	39.6	1.88	m	1.29 ppm, 10	27.7 ppm, 9, 10, 18
9	134.9	-	134.1	-	-	-	8, 18
10	131.3	4.94 (dq, 1.2, 9.0)	130.5	4.93	dd, 1.5, 9.3	8, 11, 18	8
11	35.3	2.64 (m)	34.6	2.60	m	10, 12, 17	
12	44.5	2.46 (dd, 8.1, 12.8) & 2.57 dd 6.2, 12.3)	44.2	2.51 & 2.53	m	11	11, 13, 14, 17
13	142.3	-	141.3	-	-	-	12, 15
14	130.2	-	129.4	7.13	m	-	12, 15, 16
15	126.5	-	128.1	7.24	m	-	13, 14
16	128.9	-	125.7	7.13	m	-	14
17	21.3	0.92 (d, 6.4)	21.0	0.93	d, 6.5	11, 12	10, 11, 12
18	16.0	1.36 (d, 1.9)	16.0	1.39	d, 1.3	10	8, 9, 10
19	174.1	-	173.0	-	-	-	22
20	174.6	-	174.2	-	-	-	3, 23
21	172.1	-	171.2	-	-	-	1, 3, 4, 24
22	-	-	52.1	3.71	s	-	21
23	-	-	53.3	3.82	s	-	3
24	-	-	52.1	3.67	s	-	19

Figure 2.16 Structure elucidation of **105**: **A**, displayed structure of **90** and **105**; **B**, with corresponding NMR assignment compared to literature known compound L-731,120 **90**.¹¹⁹

The double knockout of non-heme iron-dependent oxygenases Mfr1 and Mfr2 led to abolition of **1** biosynthesis and formation of compound **90**. This shows that Mfr1 and Mfr2 must act early in the pathway, and that the squalestatin hexaketide synthase (SQHKS) fully reduces at the carbon chain.

Single knockout of Mfr1 and Mfr2

The double knockout of the similar genes *mfr1* and *mfr2* led to the accumulation of intermediate **90**. In an attempt to determine the function of the individual genes, both were knocked out independently by the bipartite marker strategy as described in section 2.5.2. Because both genes are in close proximity on the genome, the left and right part of the genes for homologous yeast recombination cloning were reduced from 500 bp to ~350 and 400 bp. For the disruption of *mfr1*, 24 transformants were achieved and for the disruption of *mfr2*, 18 putative KO transformants were generated with two rounds of transformations. The reduction of the size of the recombination sites was shown to have no negative effect on the number of generated transformants.

Genetic analysis of knockout transformants In an initial experiment the genomic DNA of eleven putative *mfr1* KO transformants and ten putative *mfr2* KO transformants were isolated and tested for disruption of the GOI, using oligonucleotides P767+768 (for amplifying *mfr1* = 1020 bp) and P769+770 (for amplifying *mfr2* = 904 bp).

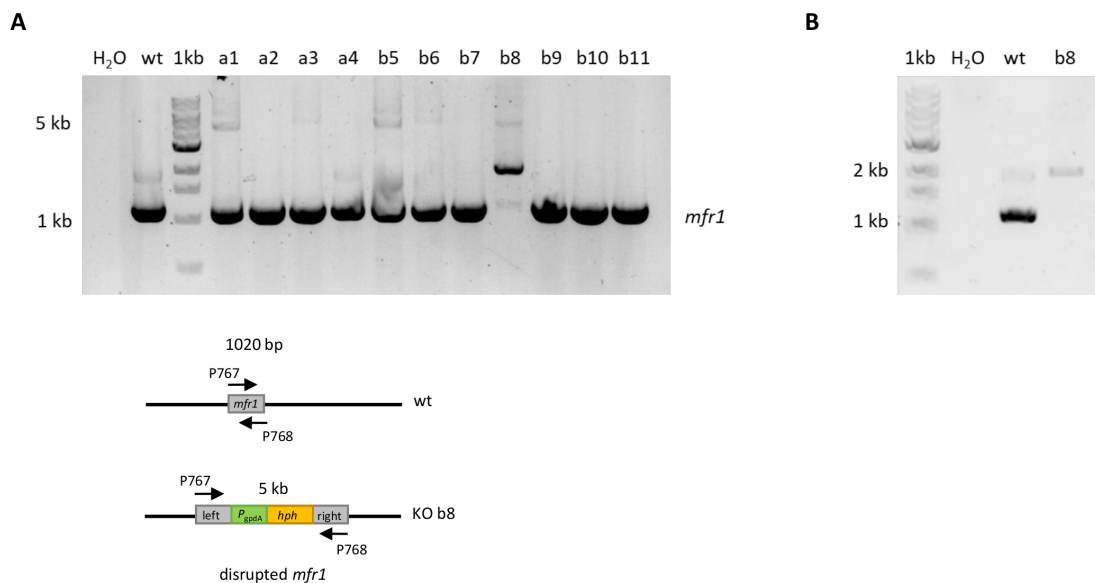


Figure 2.17 Genetic analysis of putative *mfr1* KO transformants: **A**, amplified *mfr1* gene (1020bp) with P767 + P768; **B**, repetition of PCR shown in "A" for transformant b8; H₂O control and wt positive control; 1% agarose gel; 1 kb ladder.

All, except one (b8), putative KO candidates for Mfr1 were shown to contain an intact *mfr1* gene (Fig. 2.17 A). As a second band (around 2 kb) was detected for transformant b8, the amplification was repeated and it is likely that this band is due to non-specific binding to a place somewhere else in the genome as it is also present in wt positive control (Fig. 2.17 B). Nevertheless, transformant b1 is likely to be a true knockout strain, as the gene *mfr1* could not be amplified from the genome.

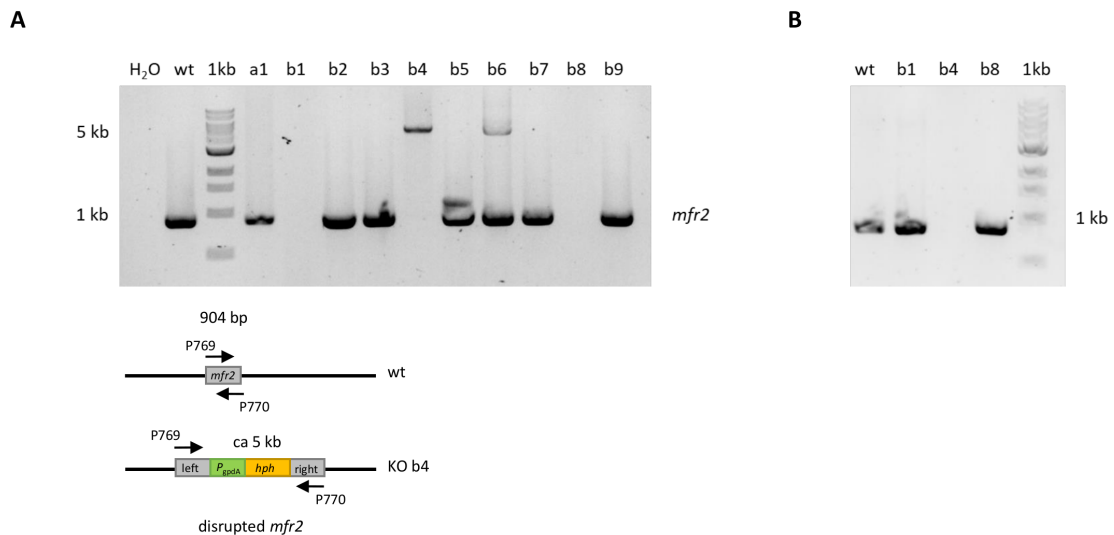


Figure 2.18 Genetic analysis of putative *mfr2* KO transformants: **A**, amplified *mfr2* gene (904 bp) with P769 + PP770; **B**, repetition of PCR shown in "A" for putative *mfr2* KO transformants b1, b4, b8; H₂O control and wt positive control; 1% agarose gel; 1 kb ladder.

Genetic analysis of putative *mfr2* KO transformants revealed that transformants b1, b4 and b8 did not show a band for the intact *mfr2* gene in an initial PCR. Instead a band with correct size of the whole knockout cassette was observed for transformant b4 indicating a true KO (Fig. 2.18 A). The PCR was repeated as for b4 and b8 no higher band was observed. Instead intact *mfr2* bands for b1 and b8 were detected. This contradicts the previous result and implies that both transformants are not true KOs (Fig. 2.18 B).

In general, the genetic analysis showed that the attempt to individually disrupt the two similar genes was more difficult than the double knockout. For some transformants a band with the size of the whole KO cassette (5 kb) was observed next to a band corresponding to the intact *mfr1/2* gene (transformant a1 and b5 of putative *mfr1* KO transformants and transformant b6 of putative *mfr2* KO transformant, Fig. 2.17+2.18). Either the cassette integrated ectopically, probably in the other non-heme iron encoding gene from the cluster due to high sequence homology, or they are not genetically pure single KO mutants but a mixture of wt and mutant.

Chemical analysis of *mfr1* KO transformant b8 Individual KO of *mfr1* alone showed an identical chemotype as the double KO, after extraction and LCMS analysis (Fig. 2.19).

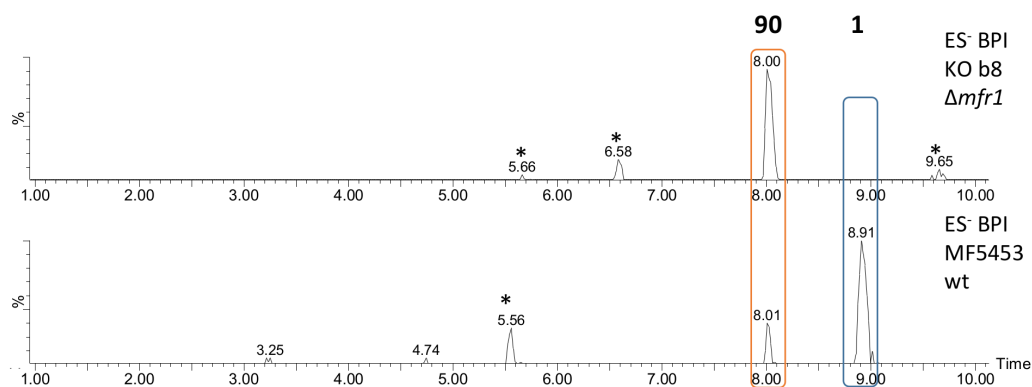


Figure 2.19 Chemical analysis of *mfr1* KO transformant b8: ES⁻ traces of extracts of *mfr1* KO transformant b8 and a wt control; in blue frame: SQS1 **1**; in orange frame: compound **90**; extracts were obtained from cultures grown in YMG+TES medium for 7 days; *unrelated compounds.

Chemical analysis of *mfr2* KO transformant b4 Disruption of *mfr2* individually resulted in a mixture of **90** plus several new compounds in very low titres with m/z values ($[M-H]^-$) of 431.2, 433.2, 435.2 (very weak) and 447.2, which likely corresponded to putative oxygenated congeners of **90** (Fig. 2.20).

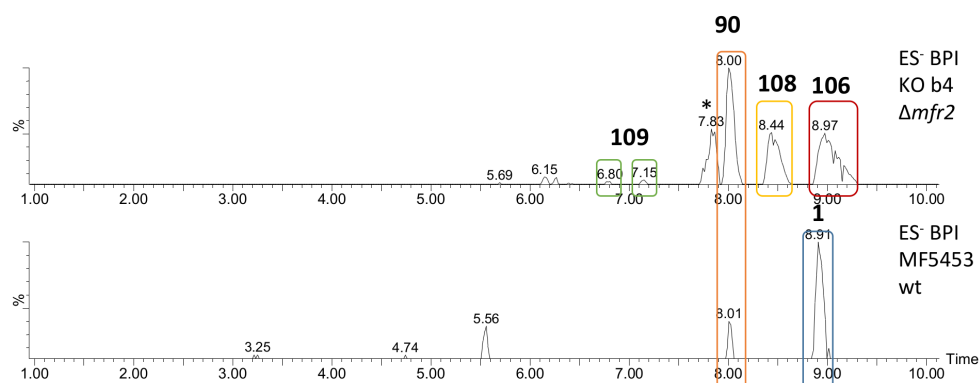


Figure 2.20 Chemical analysis of *mfr2* KO transformant b4: ES⁻ traces of extracts of *mfr2* KO transformant b4 and a wt control; in blue frame: SQS1 **1**; in orange frame: compound **90**; in green frame: compound **109**; in red frame: compound **106**; and in yellow frame: compound **108**; extracts were obtained from cultures grown in YMG+TES medium for 7 days; *unrelated compound and traces of **107**.

For structure elucidation of the new compounds, the *mfr2* KO strain was grown again. For inoculation of the main culture (14 x 500 ml flasks with 100 ml YMG+TES medium, 7 days), a seed medium (CM2) was inoculated from glycerol stocks and grown for 4 days. Isolation of the new compounds by preparative LCMS resulted in less than 1 mg product for each putative oxidative congener of **90**.

The acquired NMR data did not enable a full characterisation of **106**. However some key signals from ¹H NMR data could be assigned to structural features including the phenyl and dimethylated triketide (Table 2.3), but further structural information could not be determined by NMR for these

compounds. Furthermore, NMR analysis showed that impurities were excessive in each sample. Due to very low titres, further structure elucidation by NMR was prevented. Multiple attempts to isolate higher amounts of desired products failed and proved purification of these oxygenated congeners of **90** as extremely difficult.

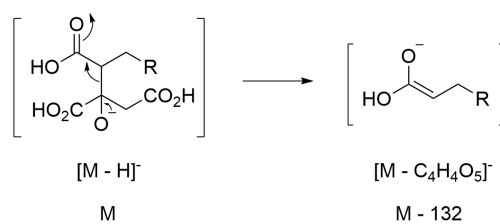
Table 2.3 Partial NMR assignment of compound **106** (CD₃OD); C-1–C-9, C-11 and C-19 to C-21 could not be assigned, corresponding putative structure in figure 2.22

Position	δ_C / ppm (125 MHz)	δ_H / ppm (500 MHz)	<i>J</i> / Hz
10	129.5	5.36	m
11	-	-	m
12	43.7	2.47–2.54	m
13	141.0	q	-
14	129.2	7.17	m
15	127.7	7.28	m
16	125.8	7.17	m
17	20.6	0.92	m
18	15.4	1.37	m

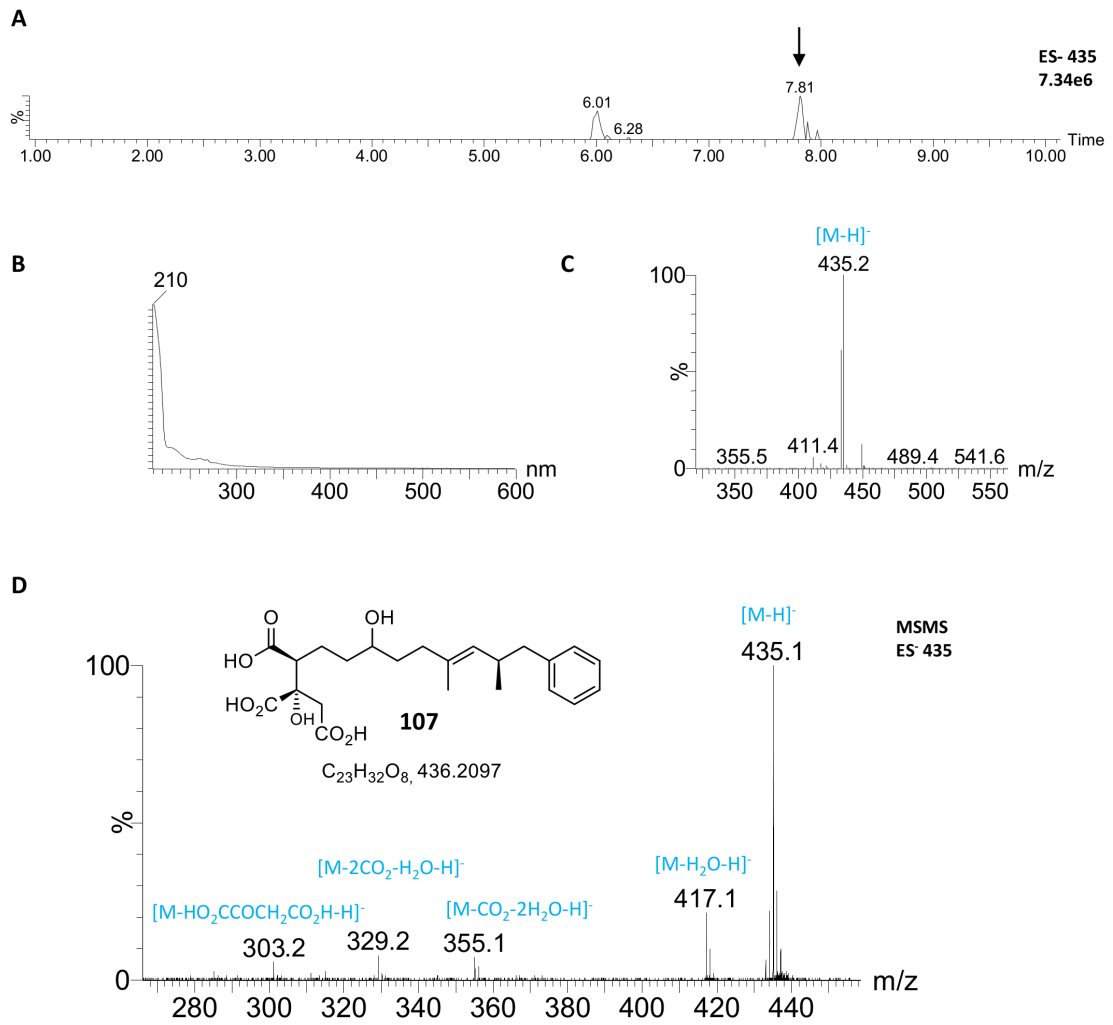
However, LCMS analysis of the new compounds revealed similar retention times and UV spectra compared to **90**. Further evidence for the existence of oxygenated congeners of **90** were obtained by HRMS and MSMS analysis of *m/z* 435, 433, 431 and 447 compounds (Table 2.4). The fragmentation pattern in mass spectra (ES⁻ mode) showed losses of CO₂ and H₂O. Besides facile loss of the oxaloacetate moiety (*m/z* = 132) by retro aldol fragmentation proving the oxidations to have occurred on the hexaketide backbone, as described in scheme 2.10 (Fig. 2.21, 2.22, 2.23, 2.24).

Table 2.4 HRMS analysis of oxygenated congeners of **90**

Compounds	HRMS results [M-H] ⁻		
	calculated	measured	ion formula
107	435.2019	435.202	C ₂₃ H ₃₂ O ₈
106	433.1862	433.1861	C ₂₃ H ₃₀ O ₈
108	431.1706	431.1702	C ₂₃ H ₂₈ O ₈
109	447.1655	447.1655	C ₂₃ H ₂₈ O ₉



Scheme 2.10 Characteristic loss of oxaloacetate (M = 132) in the ES⁻ mode for oxygenated congeners of **90**.



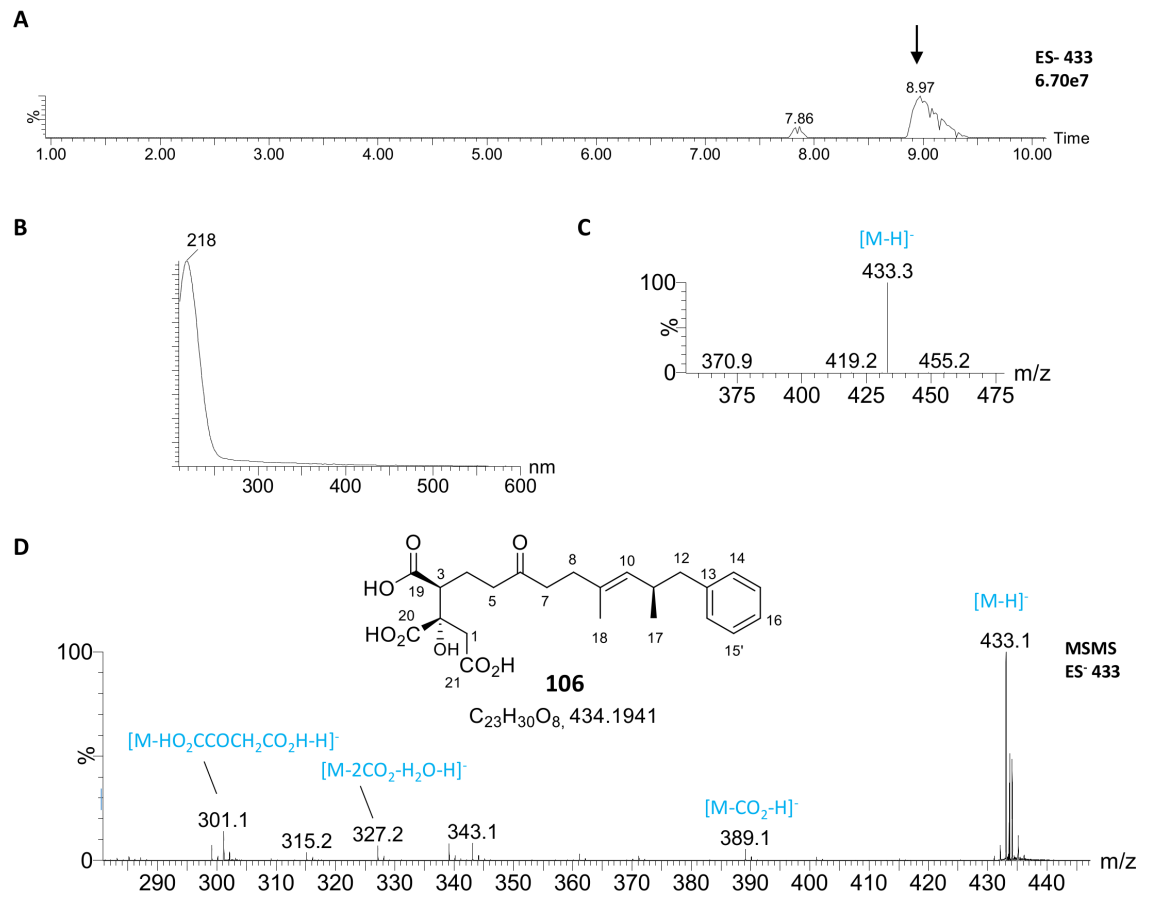


Figure 2.22 Chemical analysis of **106** isolated from *mfr2* KO transformant b4: **A**, extracted ion chromatogram (ES^- 433); **B**, UV chromatogram; **C**, mass spectrum (ES^-); **D**, MSMS fragmentation pattern of compound **106**.

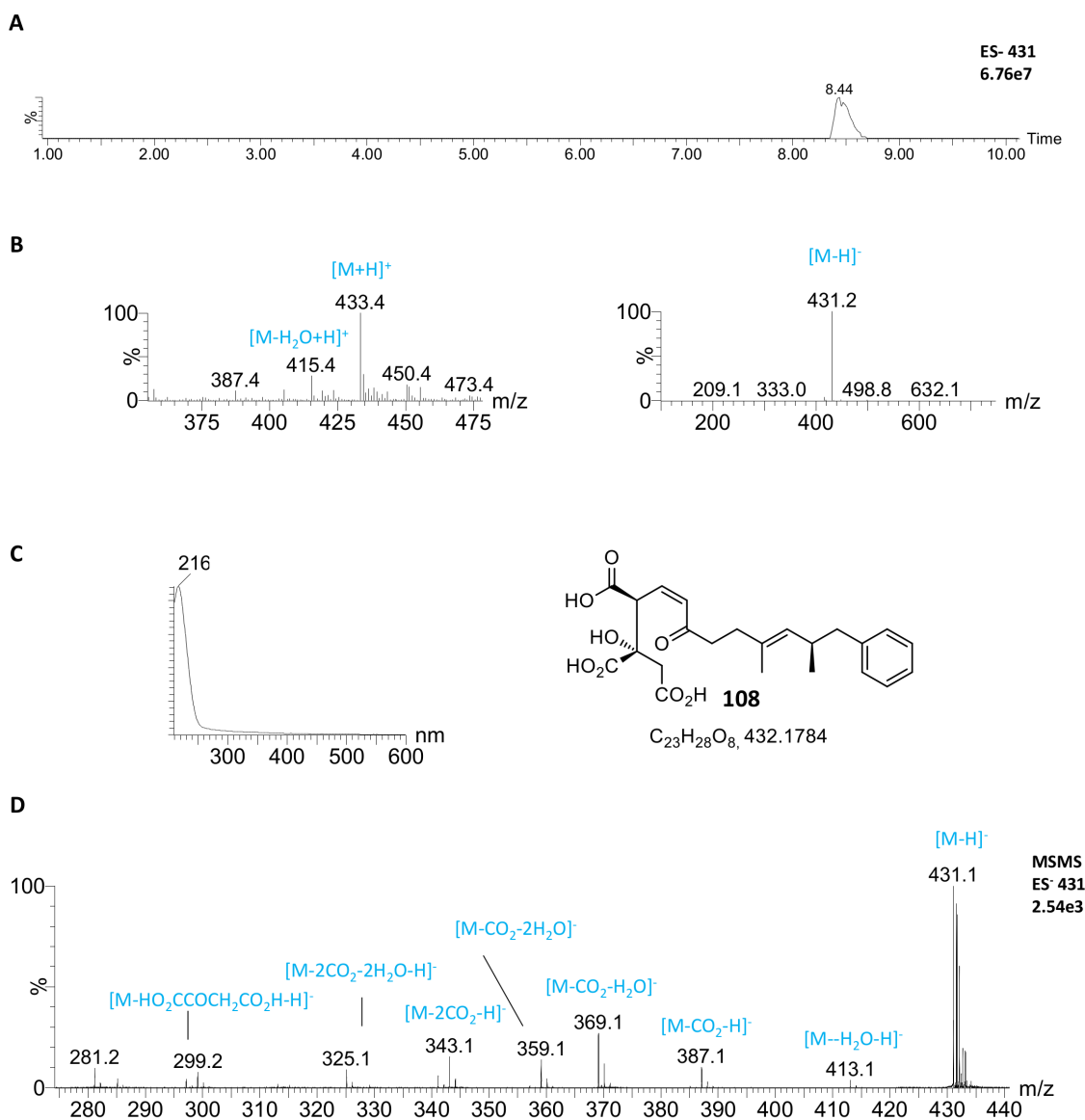


Figure 2.23 Chemical analysis of **108** isolated from *mfr2* KO transformant b4: **A**, extracted ion chromatogram (ES⁻ 431); **B**, mass spectrum (ES⁺ / ES⁻); **C**, UV chromatogram, **D**, MSMS fragmentation pattern of compound **108**.

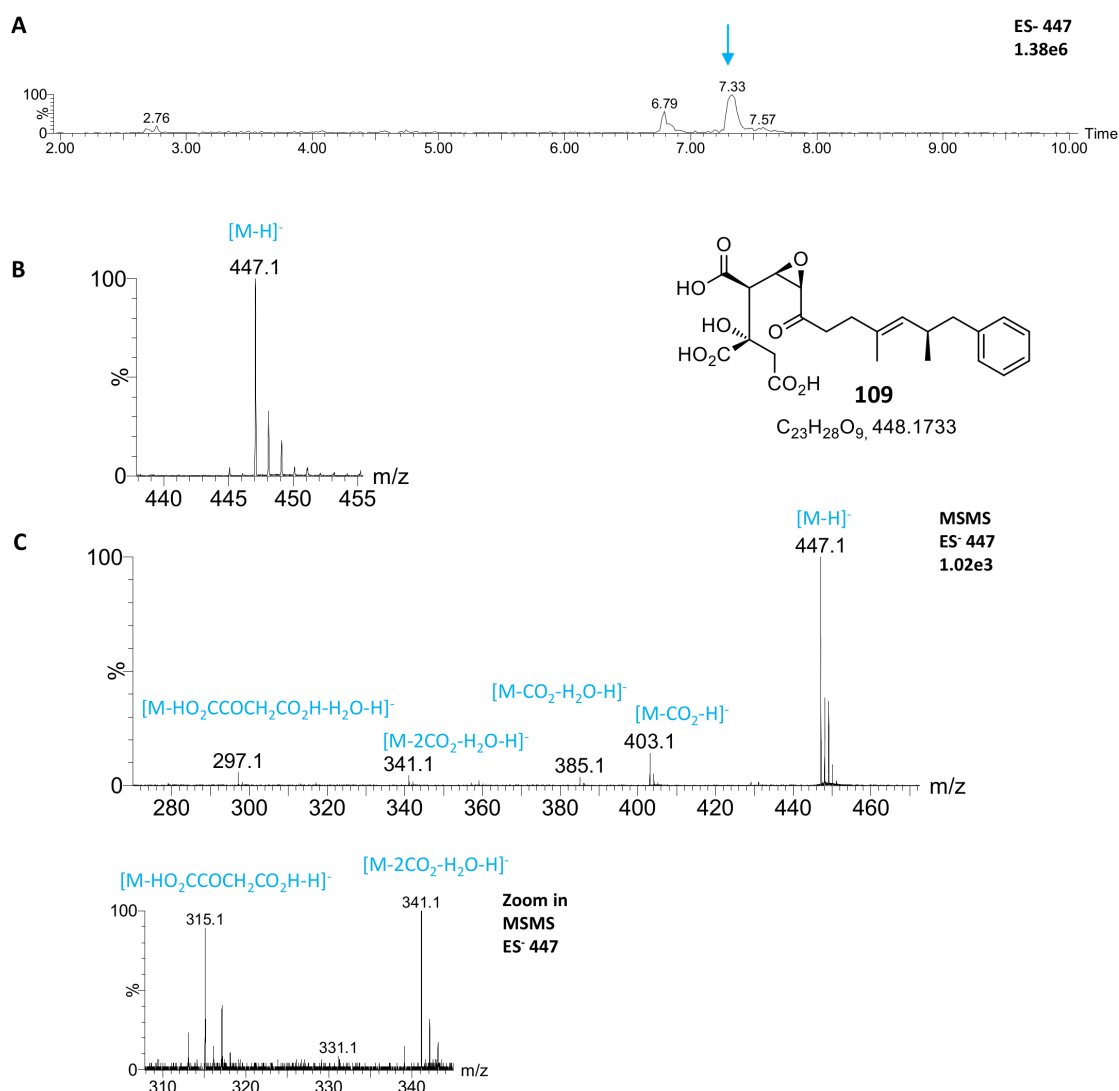


Figure 2.24 Chemical analysis of **109** isolated from *mfr2* KO transformant b4: **A**, extracted ion chromatogram (ES⁻ 447); **B**, mass spectrum (ES⁻); **C**, MSMS fragmentation pattern of compound **109**.

In conclusion, the experiments showed that both genes (*mfr1*, *mfr2*) are involved in SQS1 biosynthesis and that Mfr1 acts first, and appeared to be involved in oxygenations. It was not possible to determine any structure of congener of **90**. Likely due to the low titres and formation of shunt products as a result of their electrophilic nature. Nevertheless, the identity of the compounds formed in *mfr2* knockout experiment was confirmed by *t_R*, UV, HRMS and MSMS analysis and give confidence that these are true intermediates of the SQS1 biosynthesis. Probably **107** (only observed in traces, fig. 2.20) is built first and then undergoes further oxidations to form **106**, **108** and finally **109**. In order to form the core ring structure of SQS further oxidation steps are required.

2.5.4 Knockout of copper dependent monooxygenase Mfm1

The gene *mfm1* is predicted to encode a putative oxygenase. Further analysis using the web-based structural homology server PHYRE2¹³⁴ showed that Mfm1 belongs to the class of copper dependent monooxygenases (Section 2.3).⁵³

In order to investigate the function of Mfm1 in the biosynthesis of **1** the gene *mfm1* was disrupted. Using the bipartite method described before (Section 2.5.2) 15 transformants were generated. The extracts of ten putative knockout transformants were analysed by LCMS and compared to a wt control. Eight transformants showed no SQS1 production, whereas two of them (b4, b8) still produced **1** (Fig. 2.25).

In comparison to the wt chromatogram, five transformants (b1, b2, b3, b9, c1) showed an unambiguous accumulation of a new compound eluting at 5.7 min with m/z 479 (ES^-) and no SQS1 production (Fig. 2.25). Traces of the accumulated compound were also detected in the wt chromatogram (Section 2.4). A second compound eluting at $t_R = 8.0$ min with a nominal mass of 420, was shown to be the previously characterised **90** (Section 2.5.3). Another compound eluting at $t_R = 5.6$ min was shown to be unrelated as it was also detected in wt control and its corresponding mass (m/z 263, ES^-) is too small to belong to a SQS intermediate. Furthermore, the corresponding UV spectrum showed maxima of 419, 266 and 346 nm, which is not characteristic for SQS intermediates (UV maxima ~210 nm, Fig. 2.25).

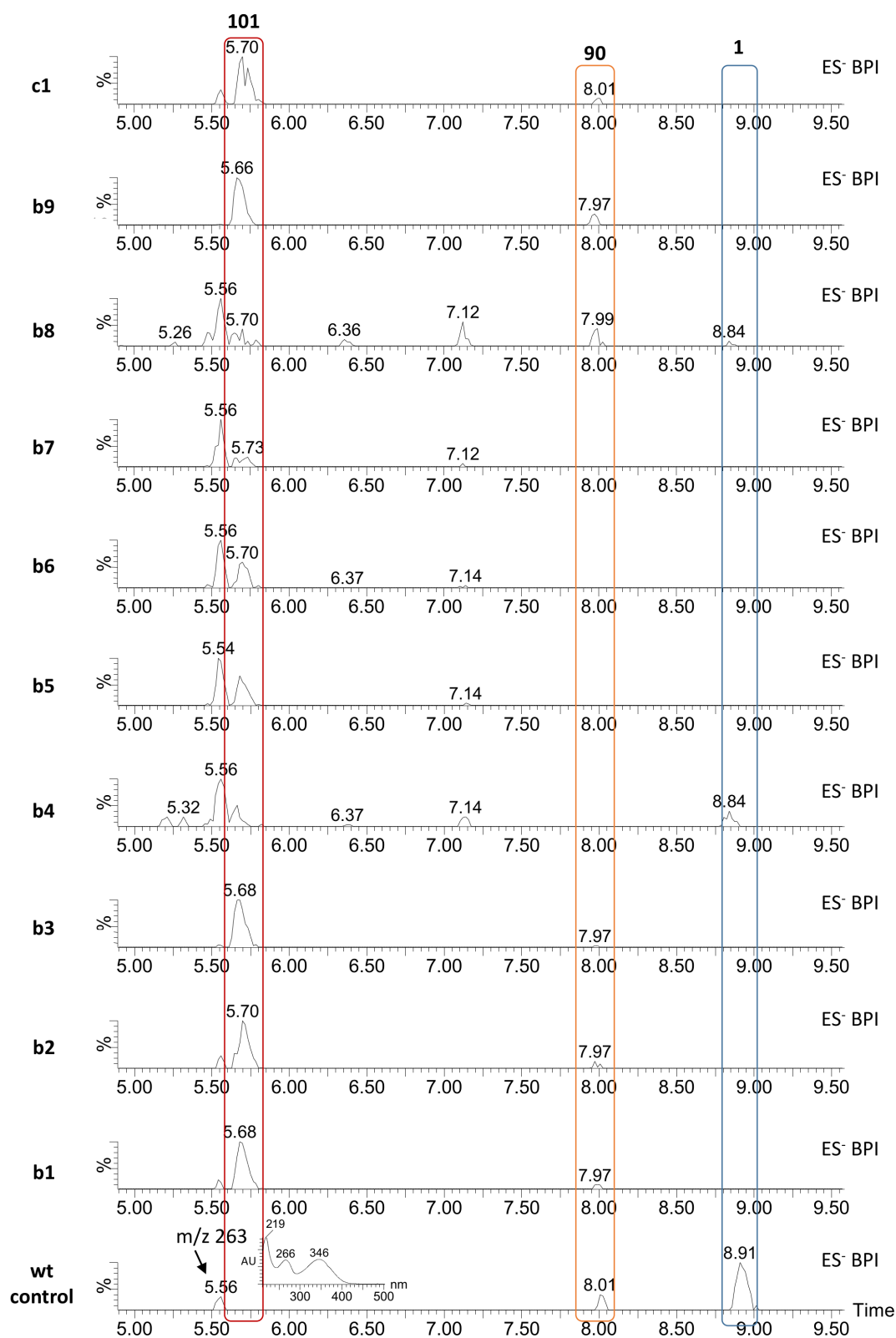
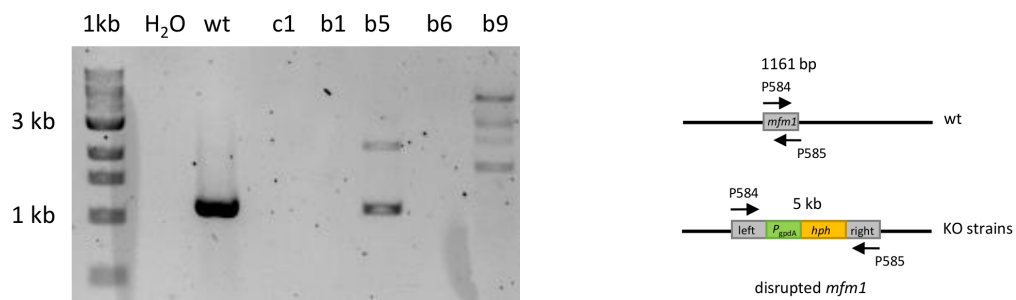


Figure 2.25 ES⁻ BPI chromatograms of extracts of ten putative *mfm1* KO transformants (b1–b9, c1) and a wt control; extracts were obtained from cultures grown in SQS production medium for 7 days and were extracted according to a small scale (1 ml) procedure; in blue frame: SQS1 1; in orange frame: compound 90; in red frame: new accumulated compound 101; unrelated compound eluting at 5.6 min, *m/z* 263 (ES⁻): UV spectrum with maxima at 419, 266 and 346 nm.

To confirm that the transformants which are not able to produce **101** are real knockout strains, the genomic DNA of transformant c1, b1, b5, b6 and b9 was isolated and tested by PCR. Oligonucleotides P584 and P585 were used to amplify the gene *mfm1* which should be disrupted in the knockout transformants. The wt positive control showed a band for *mfm1* above the 1 kb marker band (size of *mfm1* = 1161 bp). In addition transformants b1, b6, c1 and b9 showed no band for *mfm1* suggesting that the gene was disrupted and the transformants are real knockout strains. In contrast transformant b5 was shown to possess an intact *mfm1* gene and is therefore a false positive transformant (Fig. 2.26 A). Transformant b1 was further analysed as no band for the whole KO cassette was shown in gel A. Conditions of PCR "A" were repeated with extended elongation time in order to be able to detect a larger band corresponding to *mfm1* KO cassette (5 kb). The second amplification indeed showed a band for the whole KO construct in the genome of transformant b1 (Fig. 2.26 B). Further analysis using a forward primer (P411) binding the 3' end of *mfpks2* (gene upstream of *mfm1* on the genome) and the reverse primer (585) of *mfm1* gave a desired 2.6 kb fragment for the wt. For a correct KO transformant the predicted size would be around 8 kb, which is very unlikely to be amplified by PCR in this study (Fig. 2.26 B).

A



B

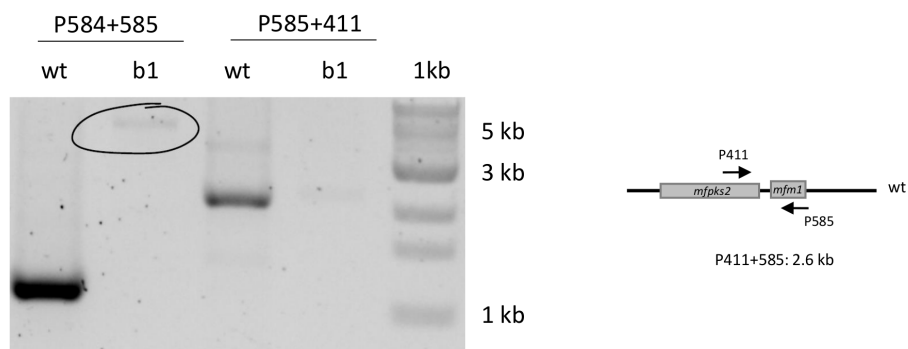


Figure 2.26 Genetic analysis of *mfm1* KO transformants c1, b1, b5, b6, b9 compared to wt control: **A**, amplified *mfm1* gene (1161 bp, P584+585); **B**, amplified *mfm1* gene (1161 bp) and KO cassette of transformant b1 (5 kb) and amplified section outside the gene (wt: 2.6 kb); KO cassette would be almost 8 kb; H₂O control and wt positive control; 1% agarose gel; 1 kb ladder; amplified *mfm1* gene (1161 bp, P584+585).

Isolation and structure elucidation of 101

In order to elucidate the structure of the major compound ($t_R = 5.7$ min) accumulating upon disruption of *mfm1* (Fig. 2.25), MF5453 $\Delta mfm1$ (transformant b1) was grown in 8 x 100 ml YMG medium. Extraction of supernatant and cells yielded 240 mg raw extract. 4.8 mg of the accumulated compound were purified using preparative LCMS. An analytical run of the extract obtained from large extraction of *mfm1* KO transformant b1 showed in addition to the major compound **101**, three new compounds eluting at $t_R = 6.2$ min (**102**), $t_R = 7.1$ min (**103**) and $t_R = 9.3$ min (**110**) (Fig. 2.27). Isolation and structure elucidation of these three compounds is discussed in section 2.5.4.

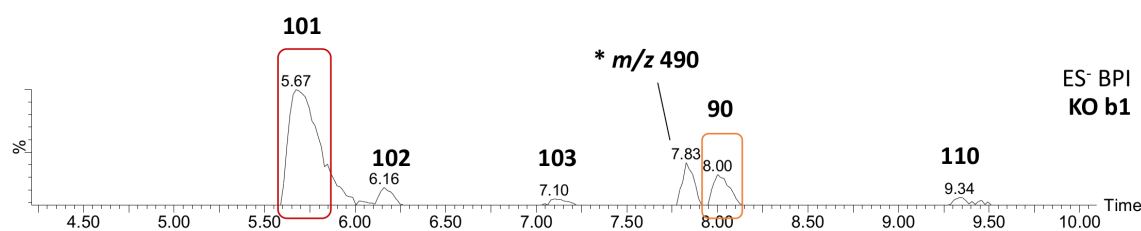


Figure 2.27 ES⁻ BPI chromatogram of extract of *mfm1* KO transformant b1: extract was obtained from a large scale fermentation (800 ml) grown in YMG medium for 7 days; in orange frame: compound **90**; in red frame: major compound **101**; unrelated compound eluting at 7.83 min, m/z 490 (ES⁻); and new compounds **102** ($t_R = 6.2$ min); **103** ($t_R = 7.1$ min); and **110** ($t_R = 9.3$ min).

Initial chemical characterisation of major compound **101** was carried out by LCMS analysis (Fig. 2.28). HRMS analysis of the accumulated intermediate gave 479.1552 [M-H]⁻ which suggested C₂₃H₂₇O₁₁ as the ion formula (calculated 479.1553) of **101**. This indicates that hydroxylation is missing from the fully hydrolysed SQS1 core **94**, which was previously characterised by Bonsch *et al.* (Scheme 2.11 A). NMR data revealed that the chemical shift of H-10 is $\delta = 5.06$, in comparison to H-10 of **94** ($\delta = 3.93$)⁵³, indicating a loss of oxygen at this position. The C-9/C-10 region isomerised in **101** (Fig. 2.29). The structure of **101** was elucidated using 1D and 2D NMR data, but full assignment of quaternary carbons was not possible, as almost no HMBC correlation was detected (Fig. 2.29, marked in blue).

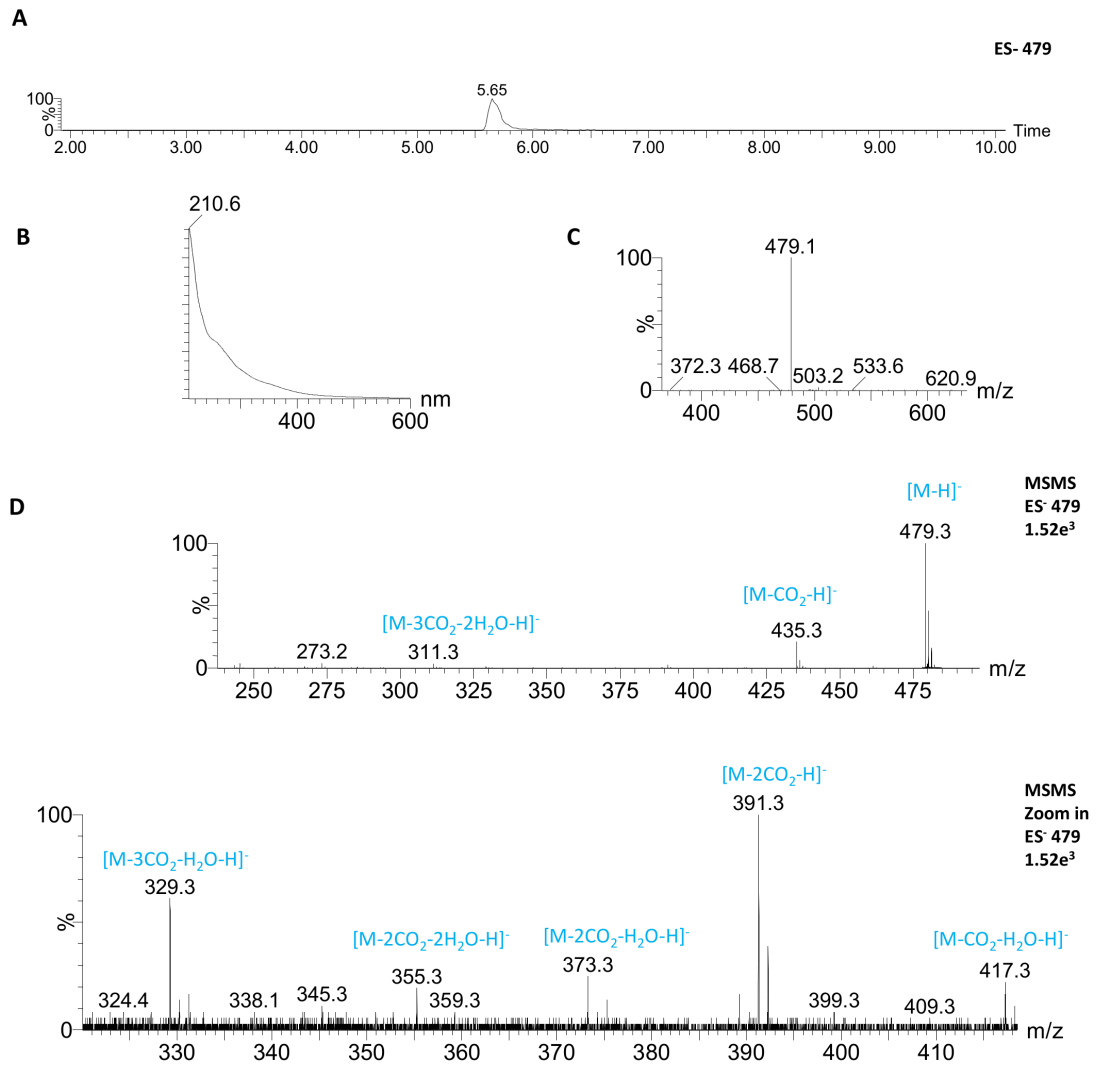
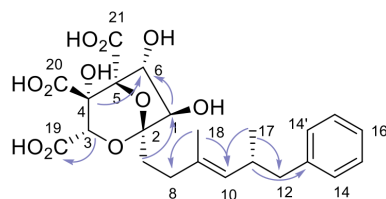


Figure 2.28 Chemical analysis of **101** isolated from *mfm1* KO transformant b1: **A**, extracted ion chromatogram (ES⁻ 479); **B**, UV chromatogram; **C**, mass spectrum (ES⁻); **D**, MSMS fragmentation pattern of compound **101**.

A

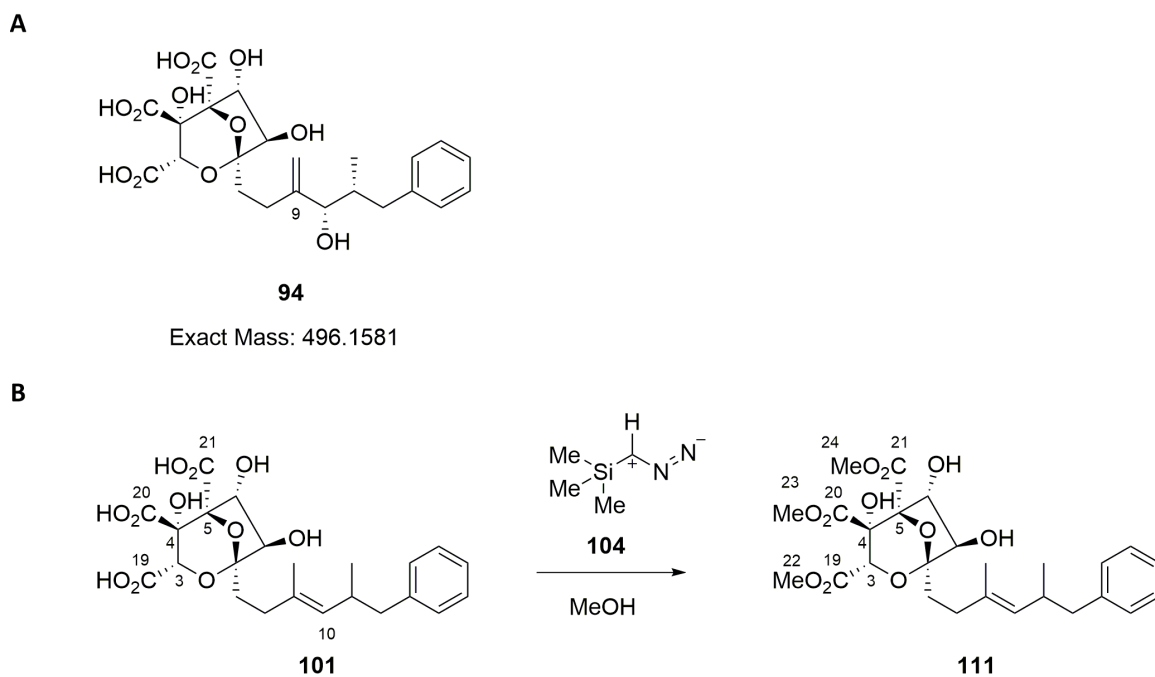
C₂₃H₂₈O₁₁, 480.1632**101**

B

Compound 101 (CD ₃ OD)					
Pos.	δ _c / ppm (125 MHz)	δ _H / ppm (500 MHz)	J / Hz	¹ H- ¹ H COSY	HMBC
1	82.4	4.07	d, 2.2	6	6, 7
2	105.3	-	-	-	7
3	75.3	5.14	s	-	19
4	74.6	-	-	-	6
5	91.8	-	-	-	-
6	78.0	5.17	d, 2.2	1	1
7	34.2	1.94	m	8	1, 2, 8
8	32.2	2.17 - 2.26 & 2.27 - 2.35	m	7	7, 9, 10, 18
9	133.9	-	-	-	18
10	129.8	5.06	dd, 9.3, 1.3	11, 18	8, 11, 17, 18
11	34.5	2.65	ddt, 9.2, 7.9, 6.4	10, 12, 17	9, 10, 12, 17
12	43.7	2.53	m	11	10, 11, 13, 14, 17
13	141.0	-	-	-	12, 15
14	128.9	7.13	d, 7.2	15	12, 15, 16
15	127.6	7.22	dd, 8.1, 6.9	14	13, 14
16	125.2	7.13	d, 7.2	-	14
17	19.9	0.97	d, 6.6	11	10, 11, 12
18	14.9	1.43	d, 1.3	10	8, 9, 10
19	169.7	-	-	-	3
20	172.1	-	-	-	-
21	168.7	-	-	-	-

Figure 2.29 Structure elucidation of **101**: **A**, key HMBC correlations observed for **101** marked with purple arrows; **B**, corresponding NMR assignment of **101**; marked positions were assigned later after methylation with TMS diazomethane **104** (blue).

To assign the five missing quaternary carbon atoms (C-4, C-5, C-19, C-20 and C-21), purified **101** was treated with TMS diazomethane **104** in methanol following the protocol of Presser and Hufner.¹⁵⁰ Product **111** was re-purified using preparative LCMS (Scheme 2.11 B). Chemical analysis of the compound using HRMS showed successful methylation of compound **101** (HRMS (ESI⁺) *m/z* calc. for C₂₆H₃₄O₁₁Na [M+Na]⁺ 545.1999, found 545.1998, Fig. 2.30). 1D and 2D NMR data was used to elucidate the structure of **111**, showing key characteristic signals for the three methylesters at 3.76 (C-22), 3.91 (C-23) and 3.81 ppm (C-24). Associated HMBC signals revealed the missing carbon atoms of **111** (Fig. 2.31 + 2.32).



Scheme 2.11 Structure elucidation of **101**: **A**, fully hydrolysed SQS1 core **94**⁵³; **B**, elucidated structure of intermediate **101**, isolated from *mfm1* KO strain treated with TMS diazomethane **104** revealed methylated compound **111**.

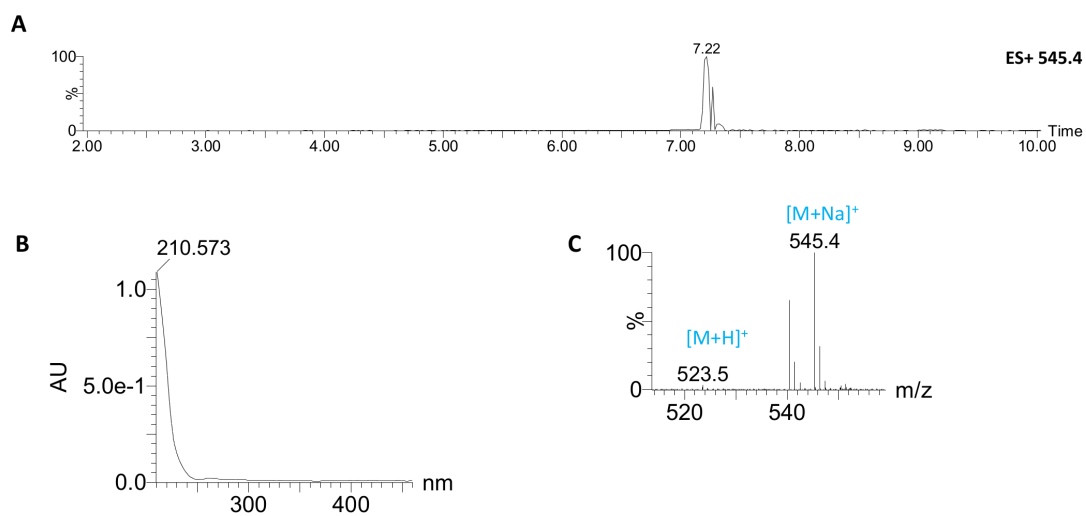
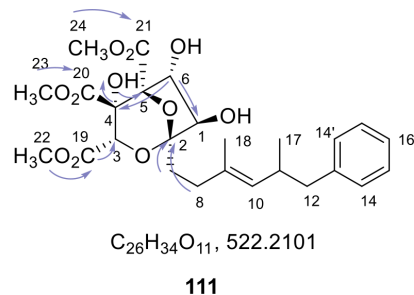


Figure 2.30 Chemical analysis of **111**: **A**, extracted ion chromatogram (ES⁺ 545.4); **B**, UV chromatogram; **C**, mass spectrum (ES⁺).

A



B

Compound 111 (CDCl ₃)					
Pos.	δ_C / ppm (125 MHz)	δ_H / ppm (500 MHz)	<i>J</i> / Hz	¹ H- ¹ H COSY	HMBC
1	82.5	4.15	d, 2.2	6	6
2	106.3	-	-	-	7, 8
3	75.6	5.16	s	22	19
4	74.8	-	-	-	6, OH
5	91.7	-	-	-	OH
6	78.6	5.14	d, 2.1	1	1
7	33.9	2.06	m	8	1, 6, 8
8	32.7	2.28	m	7	7, 9, 10, 11, 18
9	133.6	-	-	-	8, 18
10	131.1	5.07	dq, 9.2, 1.3	11, 18	8, 12, 18
11	34.6	2.64	m	10, 12, 17	9, 10, 12, 17
12	44.0	2.53	d, 7.2	11	11, 13, 14, 17
13	141.1	-	-	-	12
14	129.4	7.13	d, 7.2	12, 15	12, 13, 16
15	128.2	7.22	dd, 8.1, 6.9	14	14, 16
16	125.8	7.13	d, 7.2	-	15
17	20.8	0.93	d, 6.6	11	10, 11, 12
18	16.3	1.49	d, 1.3	10	8, 9, 10
19	167.2	-	-	-	3, 22
20	169.8	-	-	-	23
21	167.0	-	-	-	24
22	52.8	3.76	s	-	19
23	53.8	3.91	s	-	20
24	53.2	3.81	s	-	21

Figure 2.31 Structure elucidation of **111**: **A**, displayed structure of **111** with key HMBC correlation; **B**, corresponding NMR assignment.

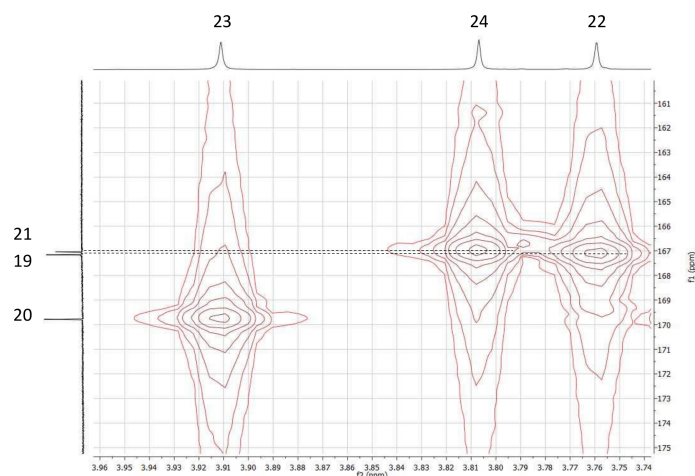
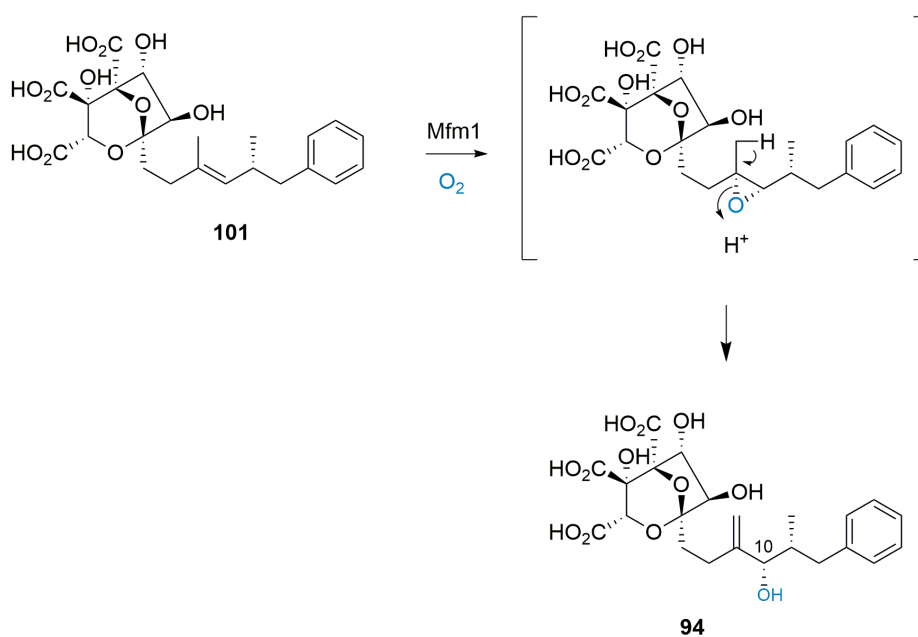


Figure 2.32 Zoomed in HMBC of **111** showing missing quaternary carbon atom correlations.

On the basis of the structure of **101** it is likely that *mfm1* encodes an enzyme which is an oxygenase. In comparison to **1** the intermediate **101** lacked the hydroxyl group at C-10 as well as the 10-*O*-acetate unit and the 4-*O*-tetraketide moiety. Therefore Mfm1 does not act on the core ring of SQS1 but after cyclisation on the hexaketide chain on C-10. Oxidation of the 9,10-olefin of **101** is proposed to involve epoxidation catalysed by the copper-dependent oxygenase Mfm1 and rearrangement to give the allylic alcohol **94** (Scheme 2.12).



Scheme 2.12 Oxidation of the 9,10-olefin of **101** by Mfm1 and rearrangement to give the allylic alcohol **94**.

Isolation and structure elucidation of 103 and 110

In addition to **101**, compound **102** and **103** (different t_R , same mass) were also produced in lower yields (less than 1 mg/l each) and purified using preparative LCMS. Initial LCMS analysis showed that the two new compounds eluting at different retention times (6.2 min and 7.1 min), showed the same HRMS (ESI- m/z calc. for $C_{23}H_{27}O_{10}$ $[M-H]^-$ 463.1604, found 463.1605) and very similar MSMS fragmentation patterns (Fig. 2.33).

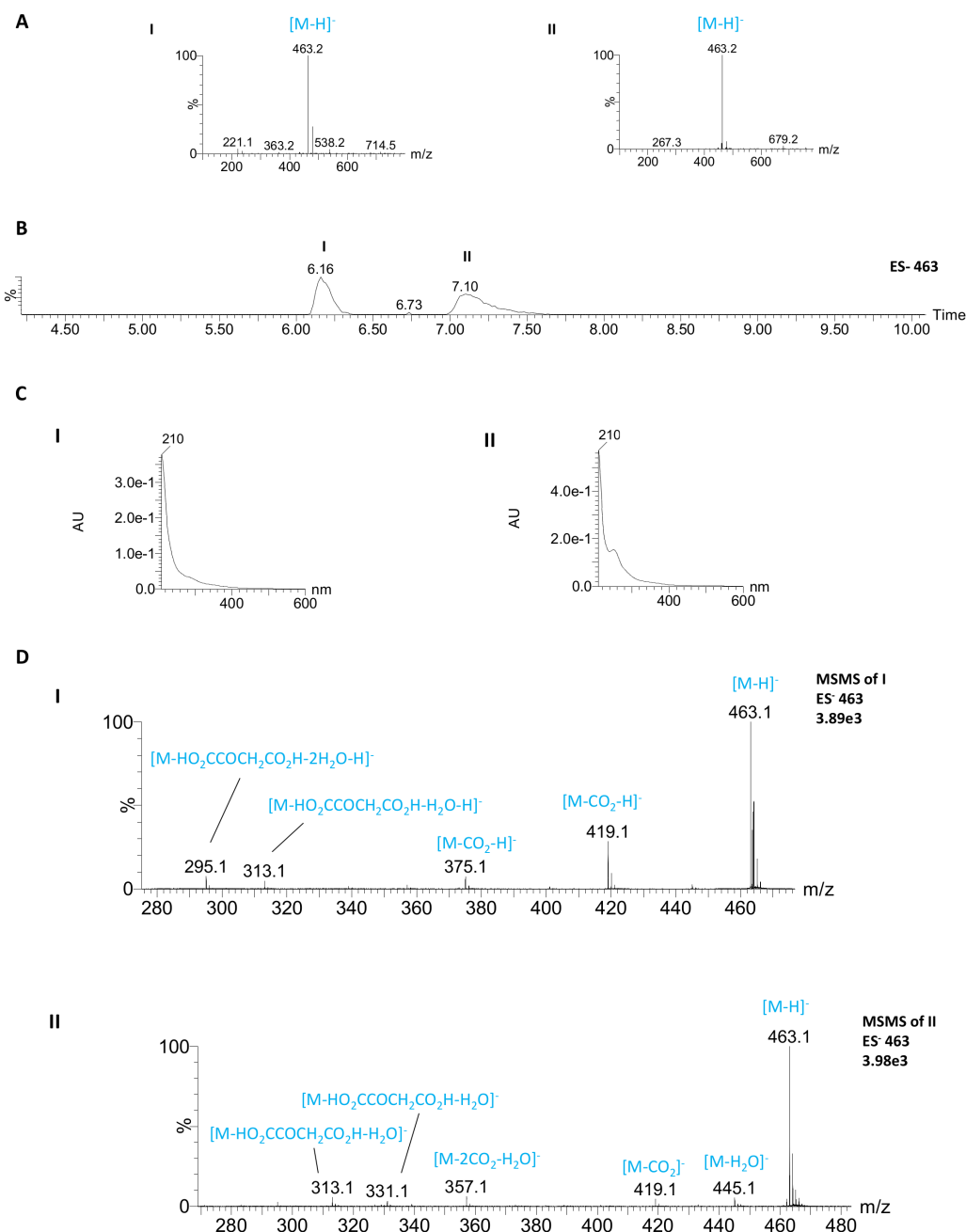


Figure 2.33 Chemical analysis of **102** = I and **103** = II: **A**, mass spectrum (ES^-) of I and II; **B**, extracted ion chromatogram (ES^- 463); **C**, UV chromatogram of I and II; **D**, MSMS analysis of I and II.

Evidence that they are both oxygenated congeners of **90** were displayed by HRMS analysis and corresponding chemical formula. Losses of CO₂, H₂O, and facile loss of the oxaloacetate moiety found by MSMS analysis proving the oxidations to have occurred on the hexaketide backbone. They are probably isomers (Fig. 2.33 D).

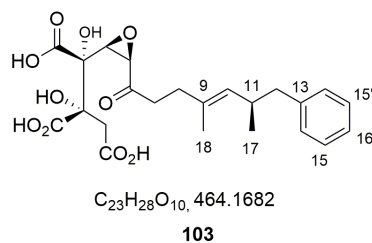


Figure 2.34 Proposed structure of **103**, corresponding partly NMR assignment shown in Table 2.5.

Table 2.5 Partial NMR data (500 MHz, CD₃OD) for compound **103** isolated from *mfm1* KO strain, C-1 to C-8 and C-19 to 21 could not be assigned

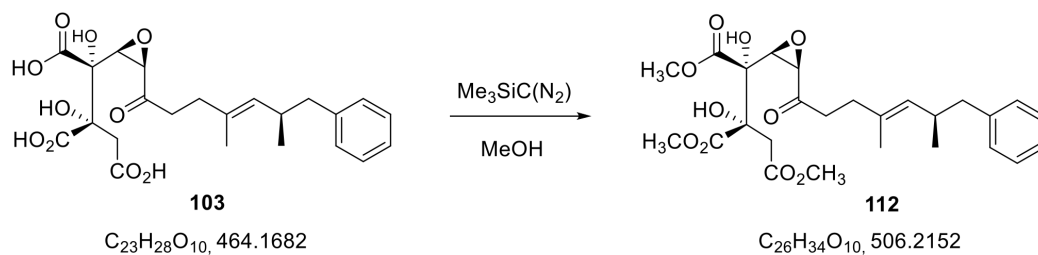
Atom	δ_C / ppm	δ_H / ppm	J / Hz	¹ H- ¹ H COSY	HMBC
9	135.2	q	-	-	18
10	131.3	5.02-5.06	m	11, 18	8, 11, 17, 18
11	36	2.66	m	10, 17	10, 12, 17
12	45.1	2.46-2.58	m	11	11, 13, 14, 14', 17
13	142.3	q	-	-	12, 15, 15'
14, 14'	130.2	7.14	m	15, 15'	15, 15', 16
15, 15'	128.8	7.24	m	14, 14'	14, 14', 13
16	126.4	7.14	m	-	14, 14', 15, 15'
17	21.6	0.97	m	11	10, 11, 12
18	16.3	1.43	d	-	8, 9, 10

Full NMR characterisation was difficult due to the low amount of substance isolated and impurities. However, structural features including the characteristic phenyl signals at 7.14 ppm (C-14/14'+C-16) and 7.24 ppm (C-15/15') and dimethylated triketide with distinctive methyl signals at 16.3 ppm (C-18) and 21.6 ppm (C-17) could be assigned (Table 2.5).

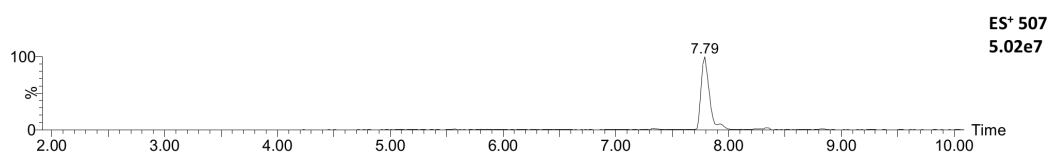
Further indication for a SQS-like structure showed successful methylation of putative carboxyl groups of **103** with TMS diazomethane, resulting in trimethylester **112** (Scheme 2.13 A). Compound **112** was re-purified by preparative LCMS, resulting in 0.4 mg compound (Scheme 2.13 B–D), but full characterisation by NMR was still prevented due to multiple coeluting compounds. Again ¹H NMR showed characteristic structural features including the phenyl and dimethylated triketide moiety as well as distinctive signals for the three methylesters (C-22 to C-24) at 3.96, 3.74, 3.73 ppm (Fig. 2.35).

Compound **102** could not be purified without another unrelated compound coeluting (containing N, odd mass) thus NMR structure elucidation was not possible.

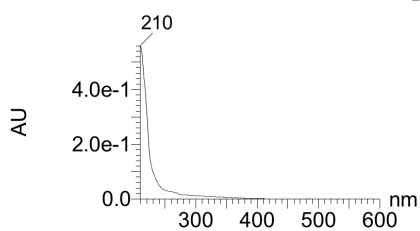
A



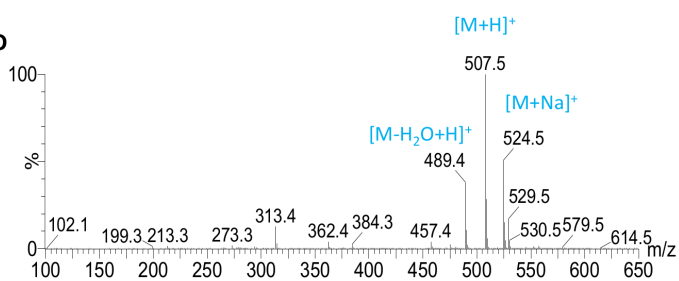
B



C



D



Scheme 2.13 Structure elucidation of **103**: **A**, reaction of **103** with TMS diazomethane to form its trimethylester **112**; **B**, extracted ion chromatogram of **112** (ES⁺ 507); **C**, UV chromatogram; **D**, mass spectrum (ES⁺).

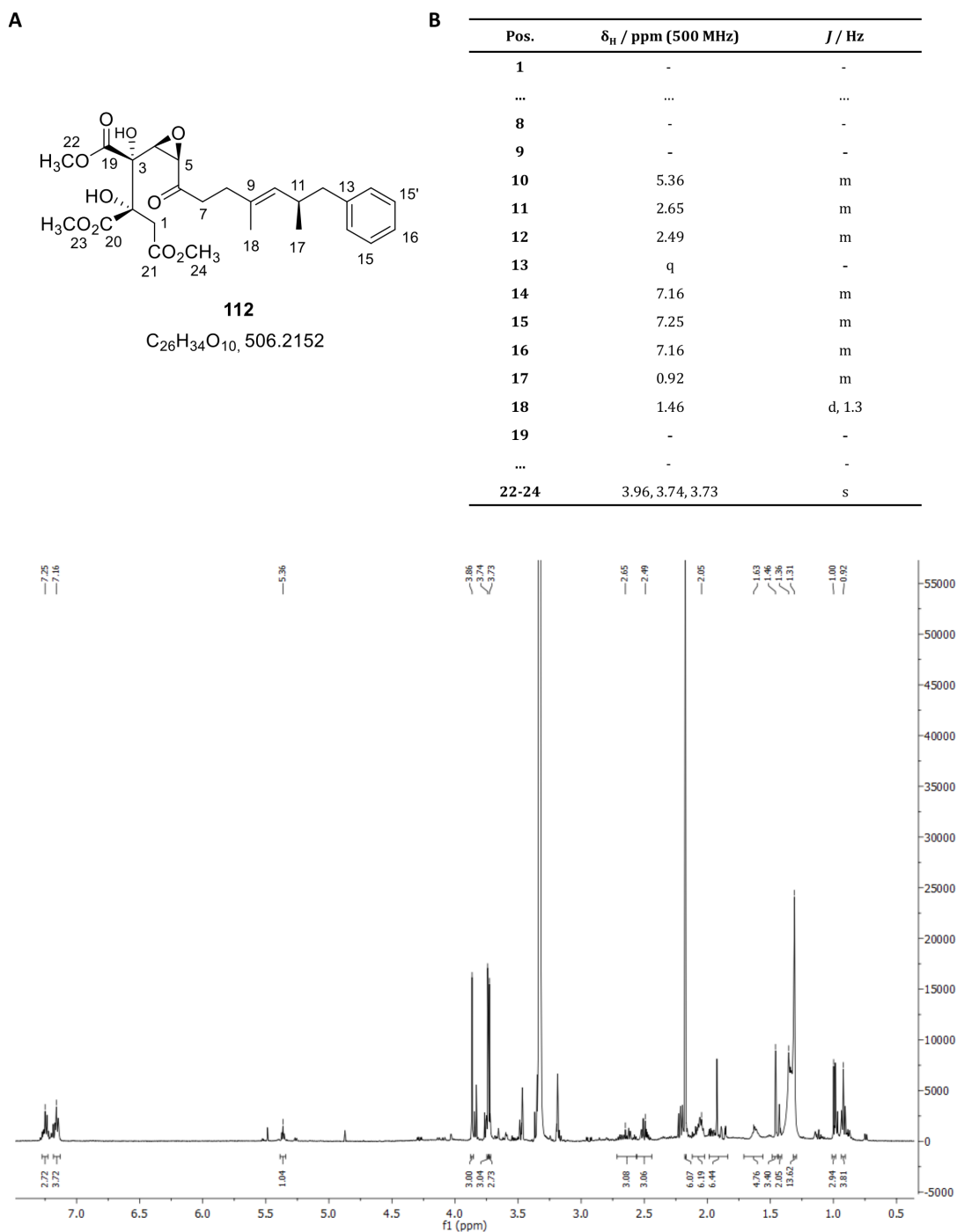


Figure 2.35 Structure elucidation of **112**: **A**, displayed structure of **112**; **B**, corresponding partly NMR assignment, C-1 to C-9 and C-19 to C-21 could not be assigned and ¹H NMR spectrum of **112**.

The *Δmfm1* strain showed also production of a new compound **110** with a nominal mass of 632 (titre 2.7 mg/l), eluting at 9.34 min (Fig. 2.27). Isolation and full characterisation by 1D and 2D NMR analysis confirmed the structure to be **110**. Compound **110** has identical structural features as **101**, but the tetraketide chain at C-6 was shown to be attached (Fig. 2.36).

Compound **110** is most likely a shunt intermediate as former *in vitro* studies showed that the tetraketide **24** was attached to the core of **1** as CoA activated compound by acyltransferase Mfm4 as the final step of the biosynthesis.⁵³ In this *in vivo* experiment the tetraketide chain was attached before hydroxylation and acetylation of the hexaketide chain as the gene expressing Mfm1 was blocked. The accessibility of **101** instead of natural substrate **94** probably led to attachment of the tetraketide by Mfm4, as a broad substrate selectivity was previously observed for this enzyme.⁵³

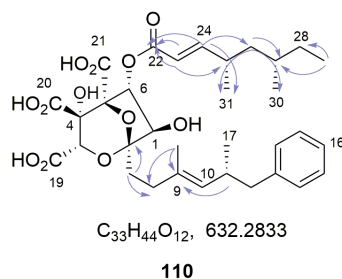


Figure 2.36 Structure of isolated compound **110** from *mfm1* KO strain, NMR assignment shown in table 2.6.

Table 2.6 NMR data (500 Mhz, CD₃OD) of compound **110** isolated from *mfm1* KO strain

Atom	δ_C / ppm	δ_H / ppm	J / Hz	¹ H- ¹ H COSY	HMBC
1	83	4.06	m	6	-
2	106.6	q			
3	130.5	5.35	m	-	-
4	q	q	-	-	-
5	q	q	-	-	-
6	81.9	6.43	m	1	1
7	35.6	1.92-1.96	m	1, 8	2, 8
8	33.7	2.19-2.26 & 2.27-2.32	m, m	7	9, 10, 18
9	135.4	q	-	-	18
10	131.2	5.02-5.06	m	11, 18	8, 11, 17, 18
11	36	2.65	m	10, 17	10, 12, 17
12	45.1	2.43-2.59	m	11	11, 13, 14, 14', 17
13	142.3	q	-	-	12, 15, 15'
14, 14'	130.3	7.13	m	15, 15'	15, 15', 16
15, 15'	129	7.22	m	14, 14'	14, 14', 13
16	126.5	7.13	m		14, 14', 15, 15'
17	21.3	0.96	d, 6.6	11	10, 11, 12
18	16.3	1.43	d		8, 9, 10
19	q	-	-	-	
20	q	-	-	-	
21	q	-	-	-	
22	167.8	-	-	-	23, 24
23	120.6	5.85	d, 15.6	24	22, 25
24	157	6.86	dd, 15.7, 8.4		22, 25, 26, 31
25	35.6	2.39-2.49	m	24, 26, 31	23, 24, 26, 31
26	44.5	1.11-1.17 & 1.39	m	25, 27	25, 27, 30, 31
27	33.2	1.27-1.39	m	26, 28, 30	26, 28, 30
28	30.8	1.10-1.15 & 1.24-1.38	m	27, 29	26, 27, 29, 30
29	11.6	0.88	m	28	27, 28
30	19.3	0.87	m	27	26, 27, 28
31	20.6	1.04	d, 6.6	25	24, 25, 26

Mfm1 cofactor dependency

Initial analysis using pBLAST showed no sequence homology of Mfm1 to a protein with known function. More sophisticated studies using a structural server (PHYRE2¹³⁴) showed that Mfm1 possess modest structural homology (15% identity) to a PHM from *Rattus norvegicus* (PDB: 1OPM, section 2.3).^{53,140}

In a fermentation experiment in which the wt strain was grown in YMG medium without trace elements (no copper), SQS1 was only observed in traces compared to compound **101** (Fig. 2.37). The *mfm1* KO strain showed a very similar chemo-type apart from no production of **1**. Fermentation experiments underlined the hypothesis of Mfm1 being a copper dependent monooxygenase as it was shown that the MF5453 wt strain produced almost exclusively SQS1 **1** by growing in YMG medium supplemented with trace element solution (containing copper amongst others).

In further experiments^a to investigate the cofactor dependency of Mfm1, 29 fermentation experiments were designed, each containing a different combination of trace elements in YMG medium. All experiments containing copper or a mixture of trace elements including copper in the medium (13 experiments), showed an improved SQS1/**101** ratio. Therefore, and in accordance with the *in silico* prediction showing conserved histidine residues in the active site, it can be assumed that copper is the cofactor of Mfm1.

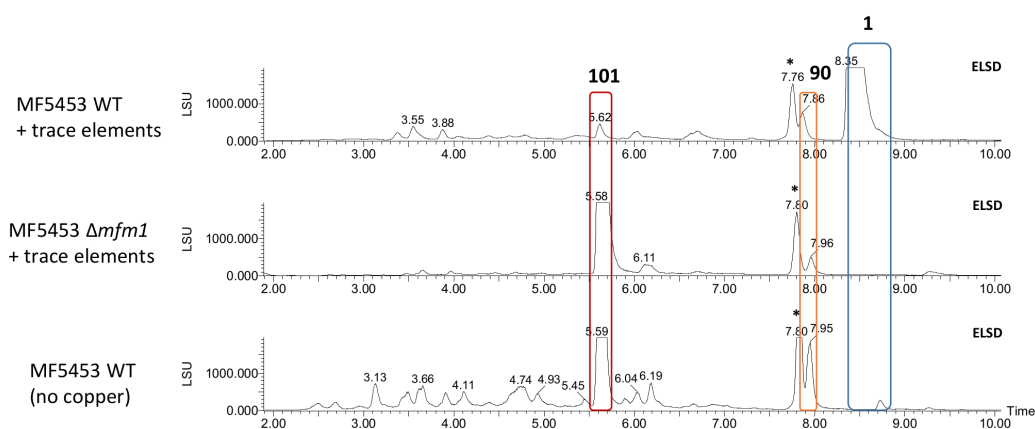


Figure 2.37 ELSD chromatograms (1-15 min) of wt extract from YMG medium and YMG medium supplemented with trace element solution (TES) and $\Delta mfm1$ mutant.

2.5.5 Knockout of acyltransferase Mfr4

The SQS1 gene cluster contains two genes encoding putative acyltransferases (ATs). One of them (Mfm4) was already investigated by expression and purification from *E. coli* and its activity

^aconducted by Simon Blazy, master student

was tested in *in vitro* assays. Mfm4 was shown to be responsible for loading acyl groups from CoA onto the squalastatin core as the final step of biosynthesis.⁵³ In previous studies the attempt to obtain soluble protein from expression of *mfr4* in *E. coli* was unsuccessful (Section 2.1).⁵³ In order to investigate the second AT from the SQS1 BGC the gene *mfr4* was disrupted in this work using the bipartite method described in section 2.5.2.

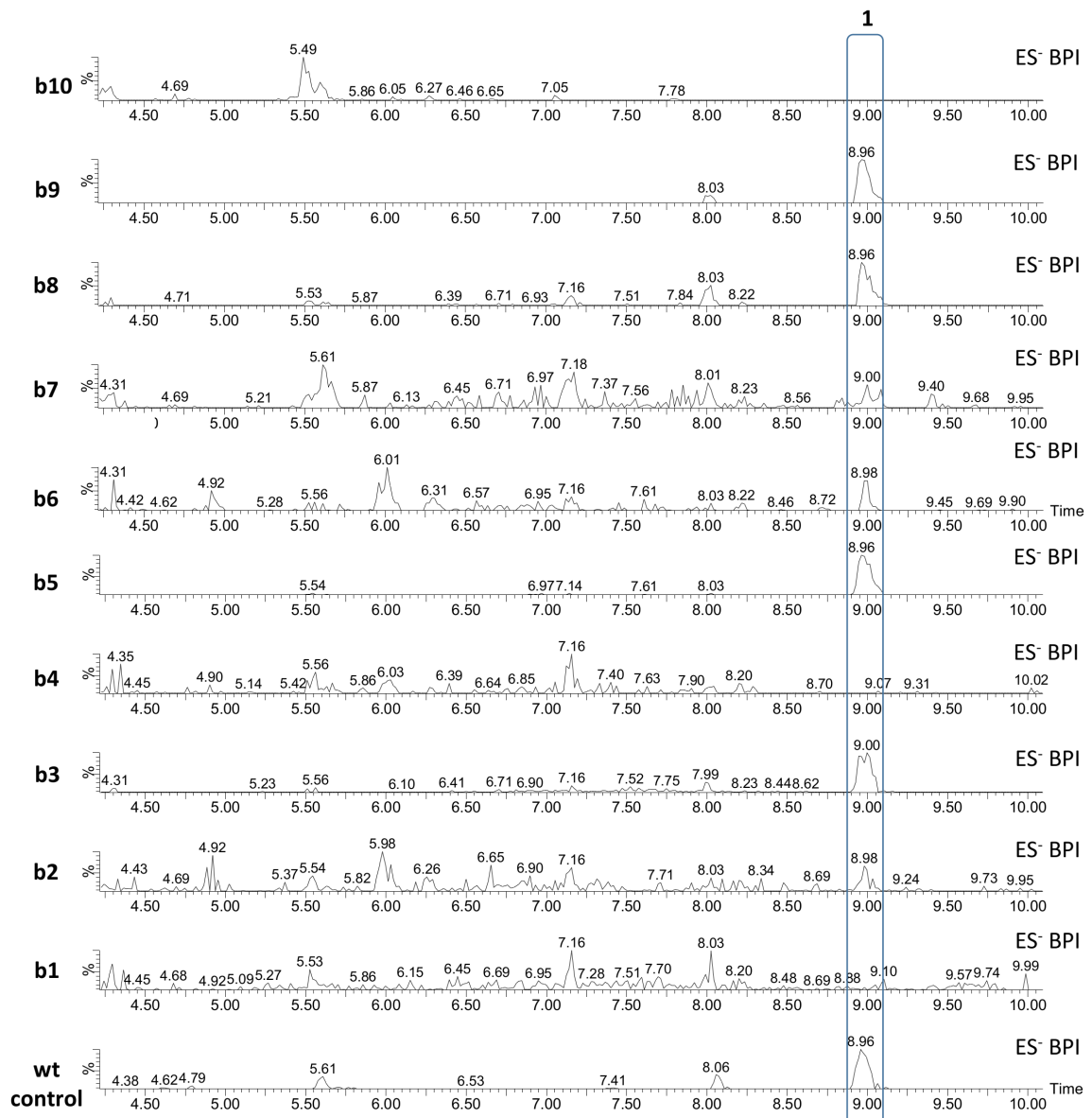


Figure 2.38 ES⁻ BPI chromatograms of initial small scale extraction of ten *mfr4* KO transformants and a wt control.

Gene disruption of *mfr4* resulted in 17 putative KO transformants. Initially 10 transformants were grown in small scale under SQS1-producing conditions in order to detect possible KO strains. LCMS analysis showed that three mutants (b1, b4 and b10) led to complete loss of SQS1 production

(Fig. 2.38). In the remaining extracts **1** was still detected suggesting no disruption of the gene. Genetic analysis revealed that the *mfr4* gene of transformant b1 and b4 was disrupted (Fig. 2.39).

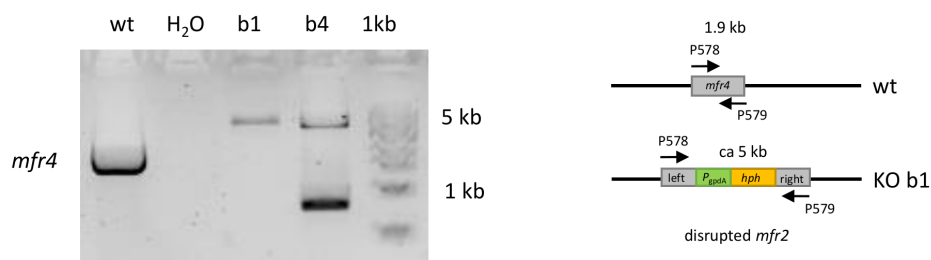


Figure 2.39 Genetic analysis of *mfr4* KO transformants b1 and b4, H₂O control and wt positive control, amplified *mfr4* gene (1.9 kb, P578+579); 1% agarose gel, 1 kb ladder.

For closer analysis of metabolites produced by the $\Delta mfr4$ strain, a 50 ml culture, grown under producing conditions was extracted. Three previously characterised intermediates of SQS1 biosynthesis were found in the extract of transformant b1. Compound **90** which was highly accumulated in $\Delta mfr1+2$ strain (proved by HRMS, calc. 419.2070, found 419.2070) and compound **101** which was identified by HRMS [calc. 479.1548, found 479.1550] (Fig. 2.40). A third compound was detected with a nominal mass of 496 eluting at $t_R = 4.4$ min, which has not been previously observed in any other KO mutant. Analysis by LCMS including HRMS and MSMS indicated that new compound **94** is the fully hydrolysed SQS1 core. This was confirmed by comparison to a synthetic standard^b (Fig. 2.41).

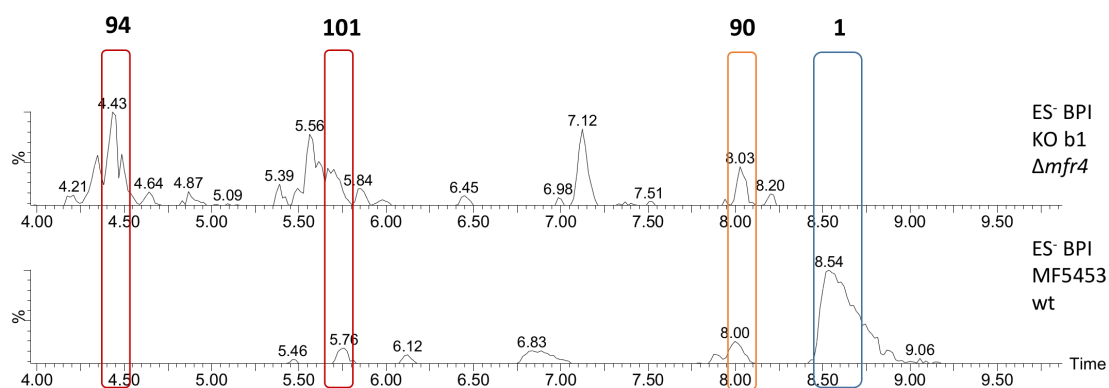


Figure 2.40 ES⁻ BPI chromatograms of *mfr4* KO transformant (b1) and a MF5453 wt control.

^bkindly provided by Dr Christoph Bartel

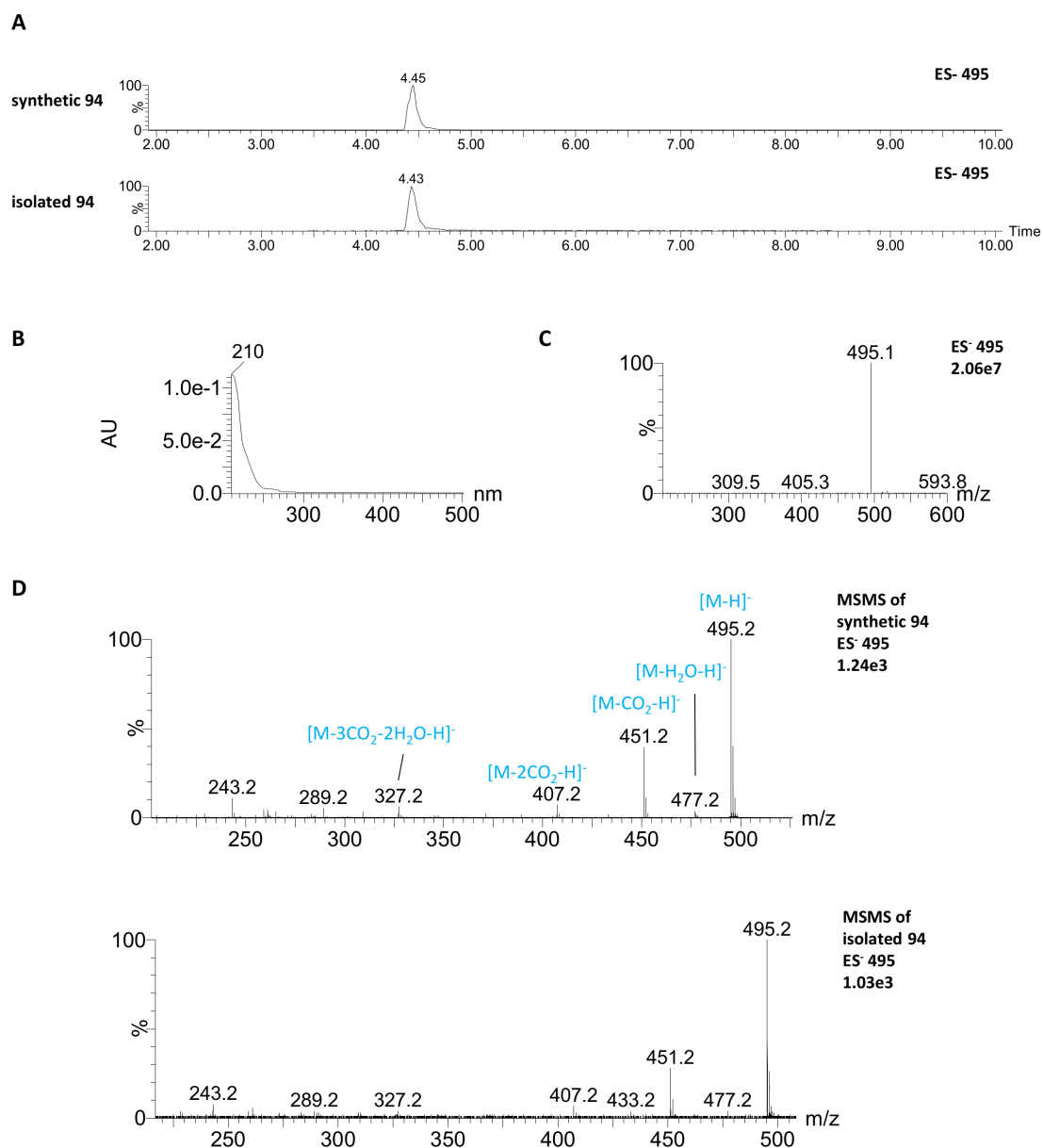
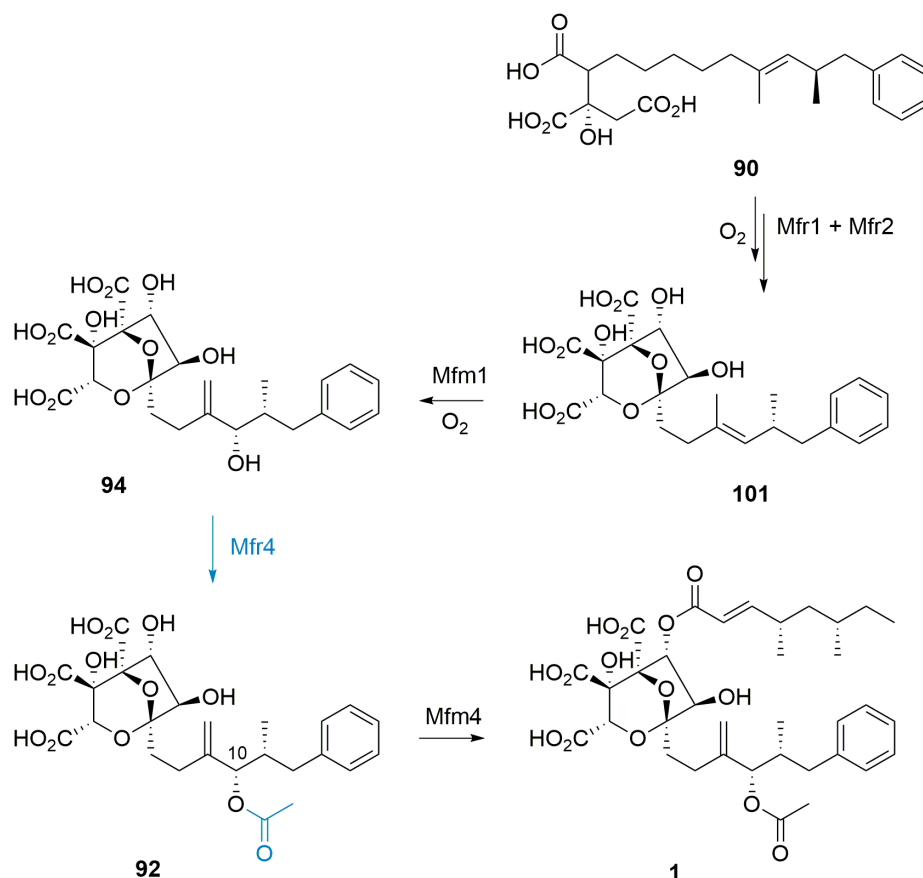


Figure 2.41 LCMS data for compound **94** in comparison to a synthetic standard of **94**: **A**, extracted ion chromatograms (ES⁻ 495); **B**, UV chromatogram; **C**, mass spectrum (ES⁻); **D**, MSMS fragmentation pattern of compound **94** in comparison to a synthetic standard of **94**.

As three of the late intermediates of SQS1 biosynthesis were found, it is likely that the transfer of the acetate to the hydroxyl group at position C-10 of the hexaketide occurs after installation of hydroxyl group by Mfm1 and before the tetraketide is attached to the core of **1** (Scheme 2.14).



Scheme 2.14 Final steps during SQS1 biosynthesis, Mfr4 catalysed reaction marked in blue.

2.5.6 Knockout of citrate synthase Mfr3

The formation of the 4,8-dioxo-bicyclo[3.2.1]octane core involves linkage of the hexaketide to oxaloacetate by citrate synthase Mfr3, which is proposed to act early in biosynthesis of **1**. Mfr3 is a good candidate for a knockout experiment, as timing and precise substrates of this step are unknown. By disrupting *mfr3*, associated KO strains were expected to produce a polyketide produced by the SQHKS.

In comparison to the other constructs described previously it was more difficult to generate $\Delta mfr3$ mutants. After five transformation procedures seven putative KO transformants were generated using the bipartite protocol described in section 2.5.2. By comparison to a wt control it was shown that three of the seven transformants led to loss of SQS1 (Fig. 2.42). However, genetic analysis showed that disruption of *mfr3* was only successful in two transformants a4 and b1 (Fig. 2.43).

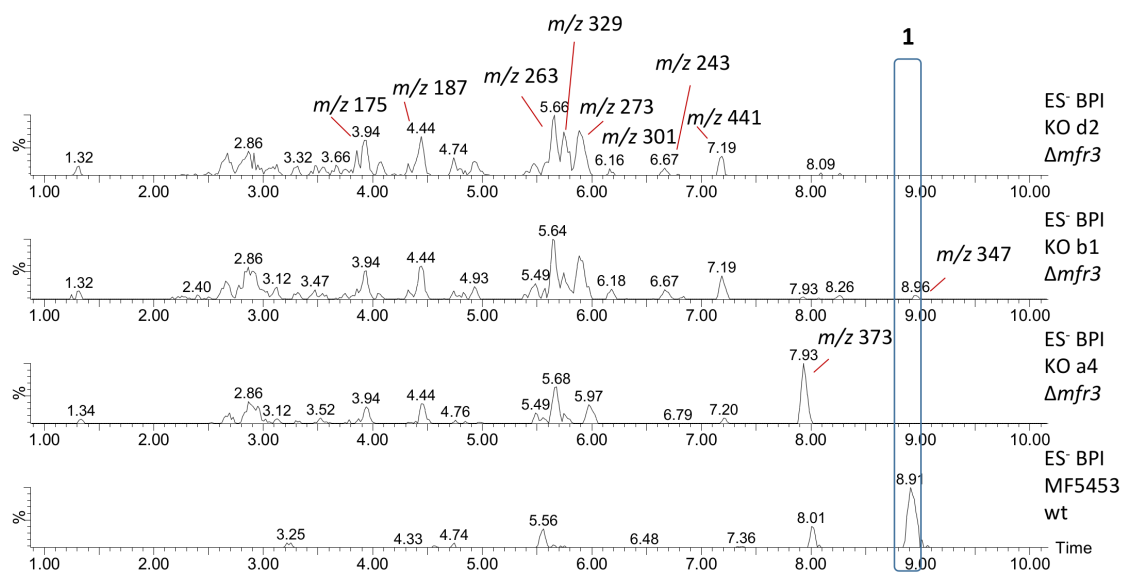


Figure 2.42 ES⁻ BPI chromatograms of extracts obtained from three $\Delta mfr3$ mutants compared to a wt control, m/z values of unrelated compounds are indicated in the chromatogram of mutant d2.

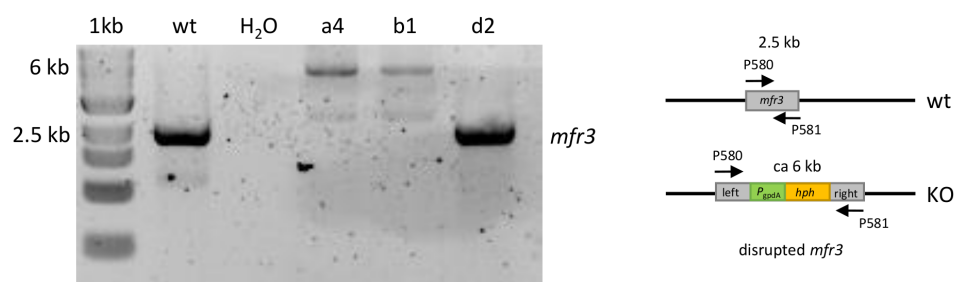


Figure 2.43 Genetic analysis of KO transformants a4, b1, d2, H₂O control and wt positive control; 1% agarose gel, 1 kb ladder; amplified *mfr3* gene (2.5 kb, P580+581).

Although it was expected to find a hexaketide related peak accumulating with a nominal mass around 300, depending on oxidation states of the polyketide chain, no new compounds were identified under "normal" conditions. Even further analysis, such as fermentation in different media (CM2, YMG and YMG+TES) and extracting cells and supernatant separately as well as at different time points (after 3, 5 and 7 days), did not result in new compounds. This may indicate that the CS is possibly involved in chain-release from the PKS, or the free polyketide is rapidly degraded. Unmethylated squalestatin tetraketide was previously shown to be rapidly digested.¹²³

2.5.7 Knockout of NADP dependent dehydrogenase Mfm3

Since it is unknown in which step of the biosynthesis the predicted NADP dependent dehydrogenase Mfm3 plays a key role, or if the enzyme is involved at all, the gene *mfm3* was disrupted using the bipartite protocol described in section 2.5.2.

Using the bipartite method with targeting sequences of 500 bp one transformant (KELI33d1) was shown to be a true KO by genetic analysis. However, the $\Delta mfm3$ mutant was shown to produce 1. In order to verify this result, more transformants should be generated, this time the targeting sides were elongated to 1 kb at each side of *mfm3*. Using the bipartite method 20 putative KO transformants were generated. Genomic DNA of 9 mutants were isolated and tested by PCR using oligonucleotides P822+823 ($t_A = 1$ min). Genetic analysis confirmed that 7 of the 9 investigated putative KO transformants were no real KO transformants because the gene for *mfm3* was still intact (band around 1 kb, *mfm3* = 1.1 kb). Only transformant 5 and 7 showed no bands at 1 kb (Fig. 2.44 A). A second round of PCR using the same set of oligonucleotides, but longer elongation time ($t_A = 3$ min), showed bands corresponding to the size of the whole 5 kb–KO cassette (Fig. 2.44 B), suggesting that transformant 5 and 7 are true KO transformants.

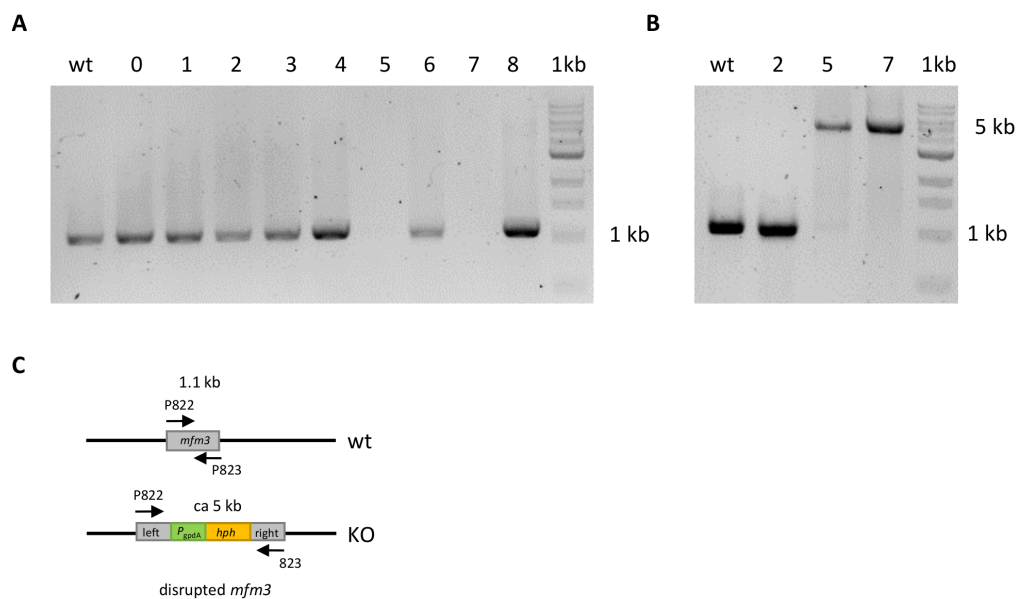


Figure 2.44 Genetic analysis of putative *mfm3* KO transformants, and wt positive control, amplified *mfm3* gene (1.1 kb, P822+823), 1% agarose gel, 1 kb ladder: **A**, first PCR, transformant 0 to 8, $t_A = 1$ min; **B**, second PCR, transformant 2, 5, 7, $t_A = 3$ min; **C**, legend for PCR including oligonucleotide positions.

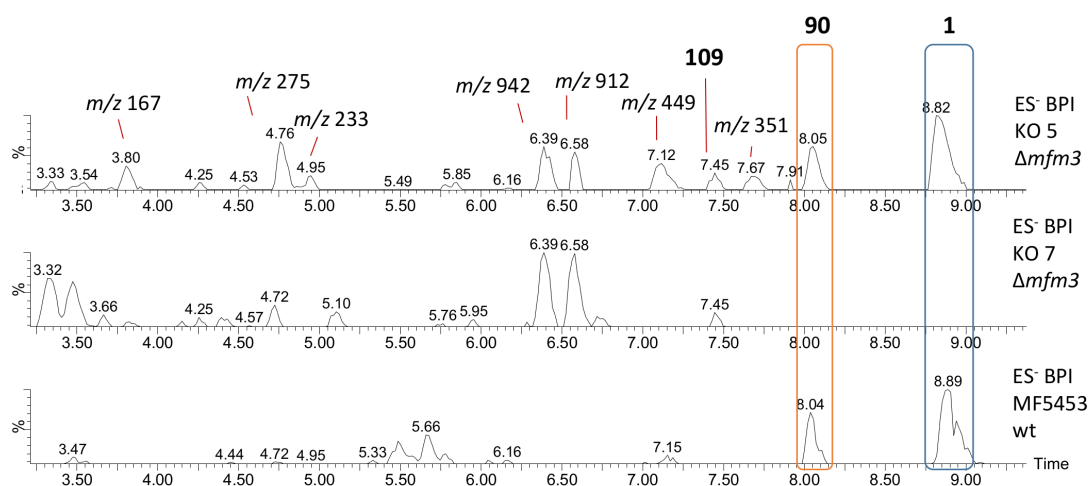


Figure 2.45 ES⁻ BPI chromatograms of extracts obtained from two $\Delta mf3$ mutants (KO 5, KO 7), compared to a wt control, m/z values of unrelated compounds are indicated in the chromatogram of mutant KO 5.

The $\Delta mf3$ mutants 5 and 7 were grown in SQS1-producing conditions and analysed by LCMS. Extract of transformant 5 showed production of **1**, **90**, as well as small amounts of **109**. In addition some unrelated peaks were observed. Also transformant 7 showed production of **1** and **90**, but only in traces. However, SQS1 intermediate **109** was produced (Fig. 2.45). The chemical analysis revealed that SQS1 is still observed notwithstanding that genetic analyses proved the disruption of *mf3*.

2.5.8 Discussion

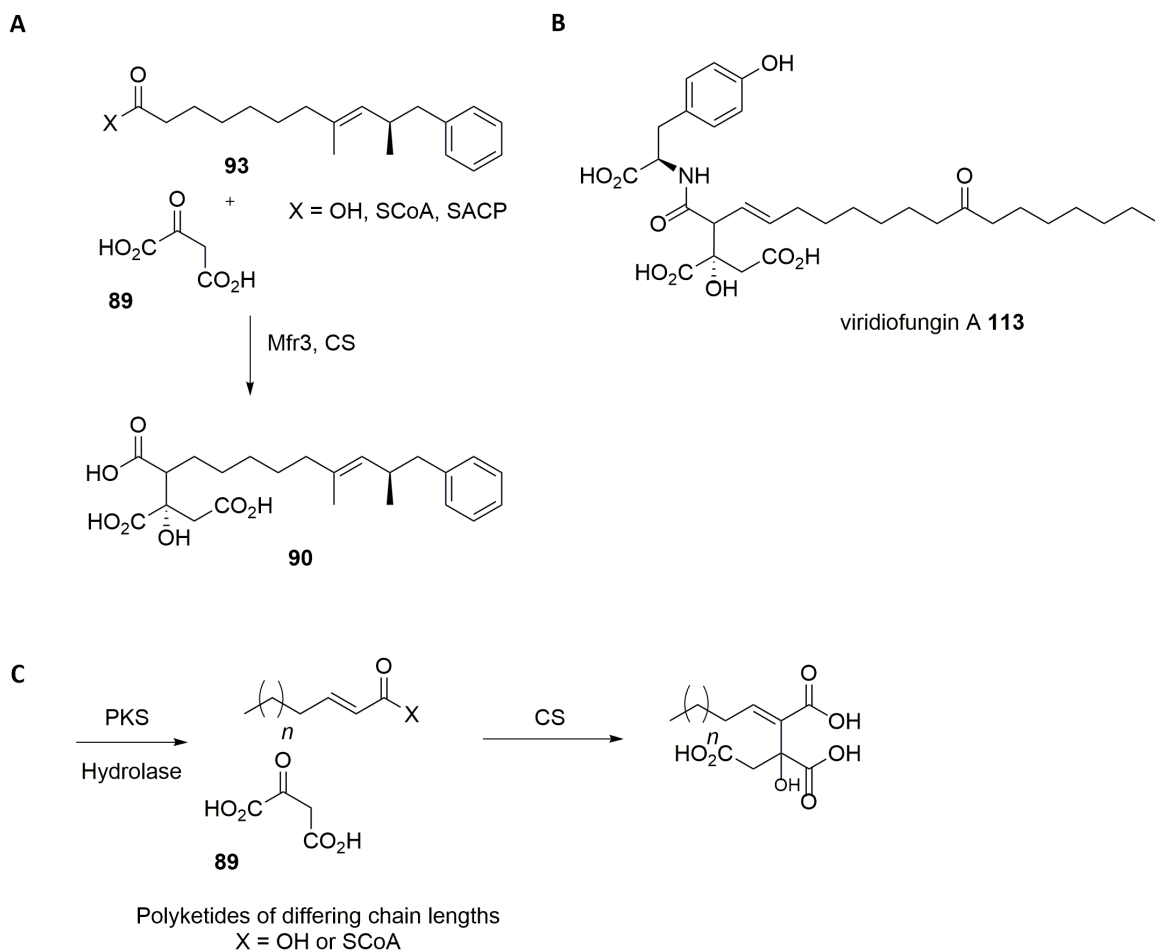
In previous biosynthetic studies of **1** Bonsch and coworkers have already shown, that the producing fungus Mf5453 is genetically tractable. A transformation protocol was established and by using the bipartite method the gene encoding the tetraketide synthase was successfully disrupted.⁵³

In this work seven more targeted gene knockouts of other genes from SQS1 BGC were generated. The Bonsch transformation was optimised.⁵³ The efficiency to obtain a true KO mutant applying the bipartite method was around 10 to 20% (Table 2.2). The gene-targeting efficiency obtained with bipartite substrates was shown to be two- to threefold higher compared to those obtained with the continuous substrates.⁸¹ However, generated KO mutants and subsequent structure elucidation of isolated compounds accumulated upon these KOs, provided the order and role of each analysed enzyme involved in SQS1 biosynthetic pathway and identified oxidative enzymes, which build the highly oxidised core of **1**.

Early steps of SQS1 biosynthesis

Feeding studies suggested a benzoate precursor as substrate for SQHKS, as benzoic acid and phenylalanine are incorporated (Section 2.1.2).⁷ The SQS1 cluster possesses genes which could be involved in benzoate production. The gene *mfm7*, encoding a PAL is predicted to catalyse the first step in the degradation of phenylalanine **85**. This step is known from other microbial and plant systems during the biosynthesis of benzoyl–CoA **25**, where cinnamate **87** is hydrated, oxidised and then subject to a retro–Claisen reaction (Chapter 1, Fig. 2.4). In order to test whether NADP dependent dehydrogenase Mfm3 is involved in this precursor biosynthesis the corresponding gene *mfm3* was disrupted. Chemical analysis of genetically confirmed mutant revealed that SQS1 is still observed, indicating that NADP dependent dehydrogenase Mfm3 is not essential for the biosynthesis of **1**. Therefore if it is involved in benzoate precursor biosynthesis it must also be complemented by an enzyme of primary metabolism. Further evidence that Mfm3 is involved in SQS1 biosynthesis is given by the fact that the encoding gene *mfm3* is conserved in all 4 SQS1 BGC (Section 2.1.2, Fig. 2.4).

In addition to the benzoate precursor pathway, production of the hexaketide **93** and linkage with oxaloacetate **89** are predicted as the first steps in the biosynthesis of **1** (Scheme 2.15 A). A knockout of putative citrate synthase encoding gene *mfr3* from the cluster was shown to abolish SQS1 production, but no new putative precursors/ intermediates of the pathway were detected, possibly due to instability of the precursor PKS chain. It has been shown by *in vitro* biomimetic synthesis that linear polyketide chains are very unstable and cyclise spontaneously in an aqueous environment.¹⁵¹ Citrate synthases are known in nearly all living cells, as part of the citric acid cycle, which plays a key role in the central metabolic pathway of aerobic organisms.¹⁵² In natural product biosynthesis citrate synthase genes are relatively rare. However, a similar reaction proposed for the first step in SQS1 biosynthesis was found in maleidride biosynthesis (Section 2.1.3). A citrate synthase–like enzyme was shown to catalyse a reaction between a hrPKS and oxaloacetate (Scheme 2.15 C).⁸⁴ Moreover, in recent investigations of the biosynthesis of phomoidride A (another gene cluster containing a hrPKS), a CS–like enzyme together with other genes was found to be likely involved in maleidrides production.¹³³



Scheme 2.15 Linkage of a polyketide chain to oxaloacetate **89** in polyketide biosynthesis: **A**, proposed first step in **1** biosynthesis; **B**, structure of viridiofungin A **113**; **C**, first step in maleidrides biosynthesis.

Furthermore isolation of alkylcitrate **90** from *mfr1+mfr2* double KO mutant suggests the linkage of hexaketide **93** to oxaloacetate **89**, presumably by Mfr3 (citrate synthase), to occur early during biosynthesis. Due to structural similarity to another fungal metabolite, viridiofungin A **113**, the linkage of the HKS product and oxaloacetate was predicted to be the first step in the biosynthetic pathway to **1** (Scheme 2.15 B).^{153,154}

Evidence for new oxygenases Mfr1 and Mfr2 in SQS1 biosynthesis

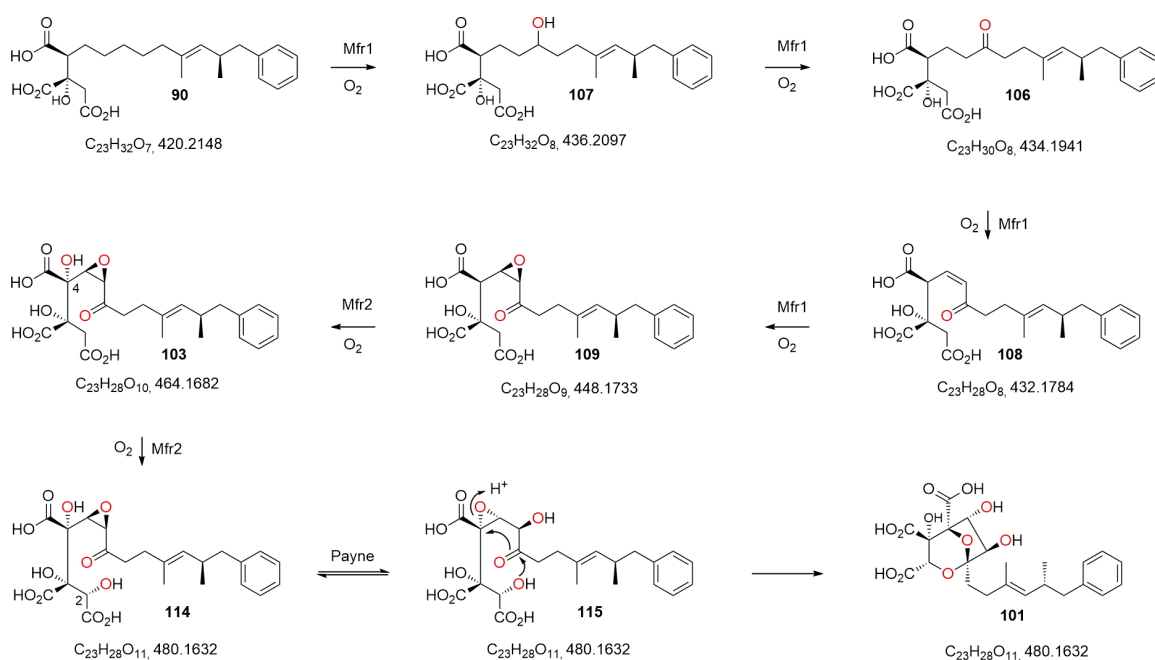
The genetic co-disruption of *mfr1* and *mfr2* as well as the individual knockouts led to loss of SQS1 production. This confirms that both are involved in SQS1 biosynthesis. Subsequent isolation of accumulated compound **90** from $\Delta mfr1+2$ strain and homology studies suggested that both enzymes play a role in oxidation of the open hexaketide–citrate **90** prior to cyclisation, or take part in cyclisation itself. Furthermore, with the identification of **90** it was shown that the linkage of the hexaketide to oxaloacetate or another intermediate, presumably by Mfr3 (citrate synthase) occurs before the

Mfr1 and Mfr2 catalysed steps. Structure elucidation of SQS precursors **90** was not possible prior to methylation of the carboxyl groups. Methylation using TMS diazomethane **104** was shown to be a powerful tool in order to purify SQS1 related compounds and to elucidate their structure, particularly to determine the carboxyl groups and quaternary C-atoms surrounding the core of SQS. The isolated intermediate **90** was shown to be a known compound, L-731,120, first isolated as a minor co-metabolite of **1** from the SQS1 producing strain MF5453 (Section 2.4). **90** is an acyclic analogue of SQS1 containing the alkyl side chain and the tricarboxylic acid groups. It represents the entire backbone of **1** prior to extensive oxidation and cyclisation.

It is possible, but rather unlikely, that **90** is not a real intermediate of SQS1 biosynthesis but a shunt product which perchance misses one or more oxidations. Shunt products are often observed in knockout or heterologous expression experiment, *e.g.* during citrinin biosynthesis.¹⁵⁵ Due to the fact that it was accumulated in high amounts in $\Delta mfr1+2$ strain and was also detected in the individual KO strain of *mfr1* and *mfr2* as well as in wt (Section 2.4), it is likely that **90** is a true biosynthetic intermediate of **1** and would invalidate the proposed ring closure mechanism proposed by the Glaxo group (Section 2.1, Scheme 2.2, route A).⁷ The ring closure could also proceed through the intermediate **91** after oxidation of **90** which is expected to lead to SQS1 acid bicyclic ring (Scheme 2.2, route B).

Further experiments using *in vitro* assays with purified enzymes (Mfr1, Mfr2) and purified substrate **90** could answer the question about shunt or not shunt product and might additionally uncover possible products and specific mode of action of each enzyme, in particular the products of Mfr2.

Individual KO of *mfr1* resulted in the same chemo-type as the dual KO of *mfr1* and *mfr2* showing Mfr1 comes first (Section 2.5.3). But upon individual KO of *mfr2* a series of new putative oxygenated congeners of **90** were detected by LCMS. HRMS and MSMS data showed that **90** is the substrate for stepwise oxidation by the putative non-heme iron oxygenase Mfr1, most likely to an alcohol **107**, a ketone **106** and then an unsaturated ketone **108** (Scheme 2.16). Further oxidation could result in the 3,4-epoxide **109**. Two peaks corresponding to the calculated mass of **109** may correspond to the keto and hemiacetal forms. These oxidised compounds were present in very low titres in the MF5453 KO strain and proved extremely difficult to purify. This is likely caused by facile shunts and degradations due to their electrophilic nature. However MSMS data convincingly showed that the oxidative modifications performed by Mfr1 leading to **107** to **109** occur on the hexaketide and not on the oxaloacetate moiety of **90**, as retro-aldol loss of unmodified oxaloacetate (m/z 132) was observed for all of them (Section 2.5.3, Fig. 2.21 to 2.24).



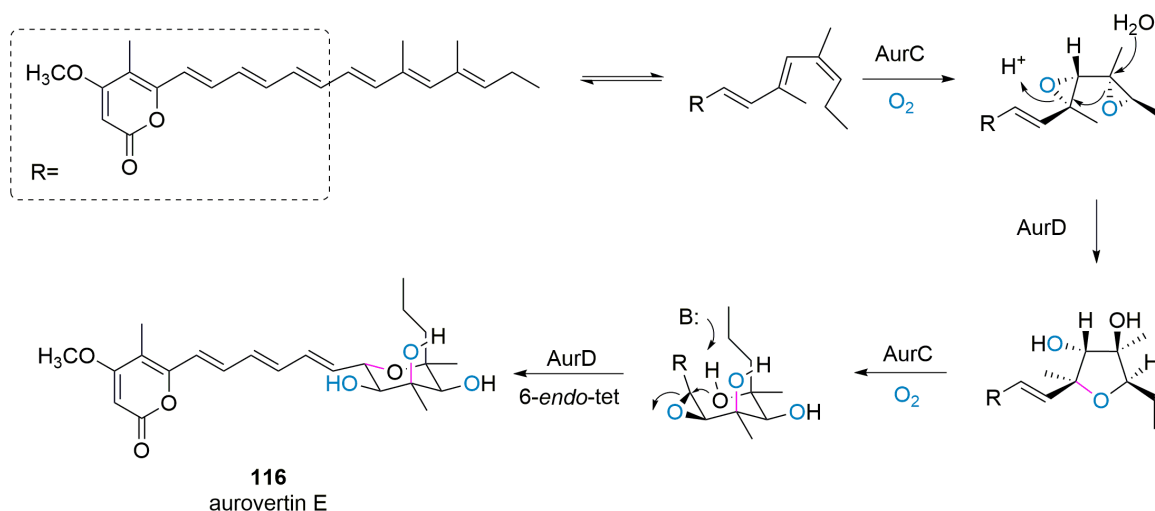
Scheme 2.16 Proposed sequence of oxidative steps catalysed by non-heme iron-dependent oxygenases Mfr1 and Mfr2 during **1** biosynthesis.

From KO of *mfm1*, encoding a copper-dependent monooxygenase two further putative oxygenated congeners were identified **103** and **101**. MS analysis of these compounds is consistent with Mfr2 hydroxylating at C-4 and C-2 to result in **103** and **114** although low titres and instability of these compounds did not allow full structural elucidation by NMR. Retro-aldol loss of unmodified oxaloacetate (m/z 132) was not observed for the major compound accumulated upon Mfm1 KO **101**. This indicates, that hydroxylation at C-2 of **114** is probably the final oxidative step catalysed by Mfr2. Facile Payne rearrangement of **114** would furnish **115** which could undergo acetal formation and concomitant transannular epoxide opening to give the observed intermediate **101** (Scheme 2.16).

In fungal PKS biosynthesis, many oxygenases have been found to work iteratively catalysing multiple similar reactions on different sites of the substrate, *e.g.* in lovastatin **6**, chaetoglobosin A, cytochalasin E, and aurovertin E **116** biosynthesis.^{156,157,158} The definition for an iterative enzyme in natural product biosynthesis can be defined as a protein with active sites that can catalyse multiple rounds of structural modifications.¹⁵⁹ In particular, many multifunctional Fe(II)/ α KG-dependent oxygenases have been found in fungal meroterpenoid biosynthesis, where a single enzyme was shown to catalyse a variety of chemistry ranging from simple hydroxylations to remarkable skeletal rearrangements (Section 1.2.3).^{102,139,160,161,162}

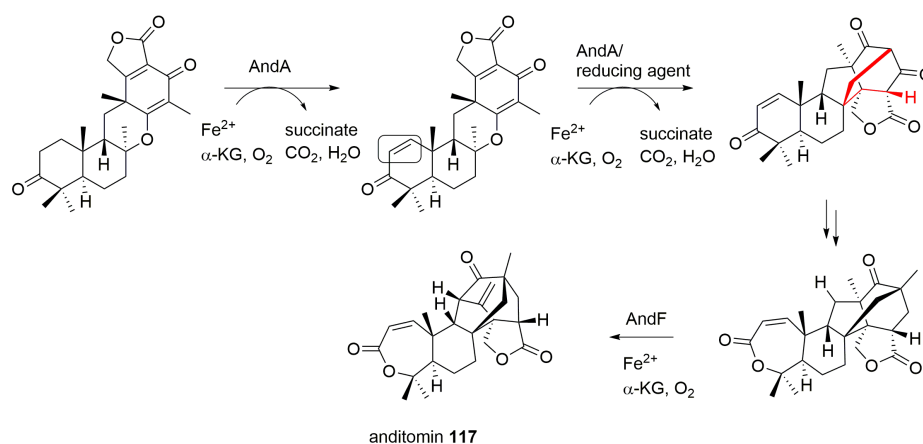
Interestingly, aurovertin E **116**, a fungal polyketide of the family of aurovertins, possess a 2,6-dioxabicyclo[3.2.1]-octane ring, similar to SQS1 core structure. In the case of **116**, it is attached to a methylated α -pyrone and derived through iterative epoxidation of a polyene precursor and a cascade

of concerted epoxide openings. AurC plays a multifunctional role. It is an iterative oxygenase, which catalyses the epoxidation of three double bonds in two separate steps in the biosynthesis of **116** (Scheme 2.17).¹⁵⁸



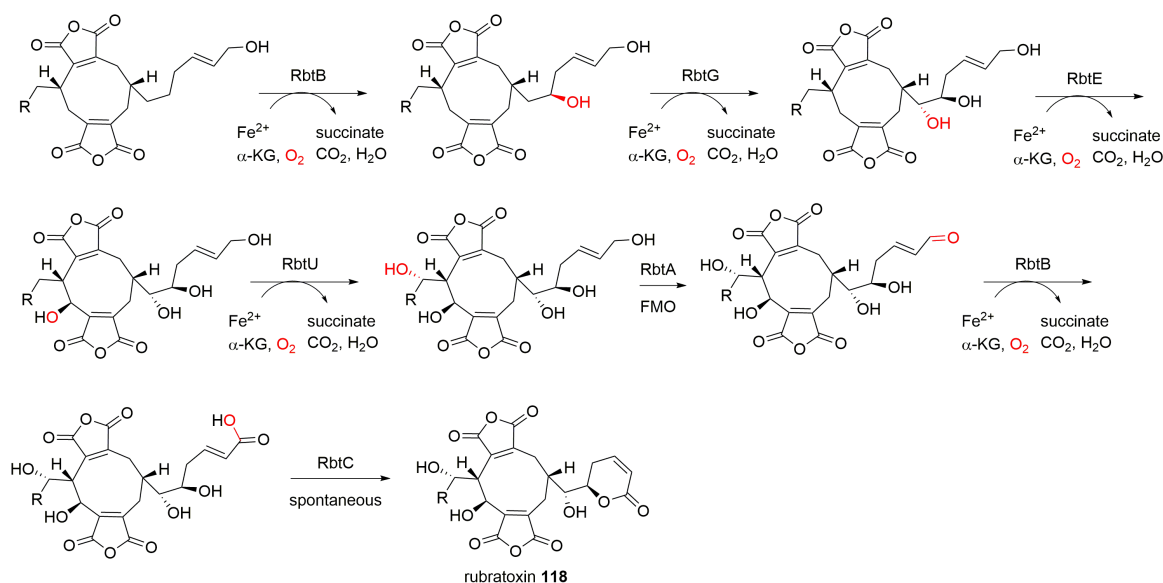
Scheme 2.17 Multifunctional role of AurC in aurovertin E **116** biosynthesis, adapted from Tang *et al.*¹⁶³.

During anditomin **117** biosynthesis two multifunctional non-heme iron-dependent dioxygenase (AndA and AndF) were found to be part of the pathway (Scheme 2.18). Anditomin possesses a unique, highly oxygenated bridged-ring system, like SQS1. The dioxygenase AndA was shown to synthesise the bicyclo[2.2.2]octane core of **117** by an unprecedented rearrangement.¹³⁹ It maintains a conserved 2-His-1-carboxylate iron-binding facial triad (H135, D137, H213), which is conserved in non-heme iron-dependent dioxygenases. An iron at this site would play a key role in the enzyme-catalysed oxygenation.¹⁶⁴ Another dioxygenase, AndF, generates the product **117** by an oxidative rearrangement (Scheme 2.18).¹³⁹



Scheme 2.18 Biosynthetic steps in anditomin **117** biosynthesis catalysed by two non-heme iron-dependent dioxygenases AndA and AndF.

Four α -ketoglutarate and iron(II)-dependent dioxygenases in the rubratoxin **118** pathway were recently shown to hydroxylate similarly unactivated carbons, but not in an iterative fashion (Scheme 2.19).¹⁶⁵



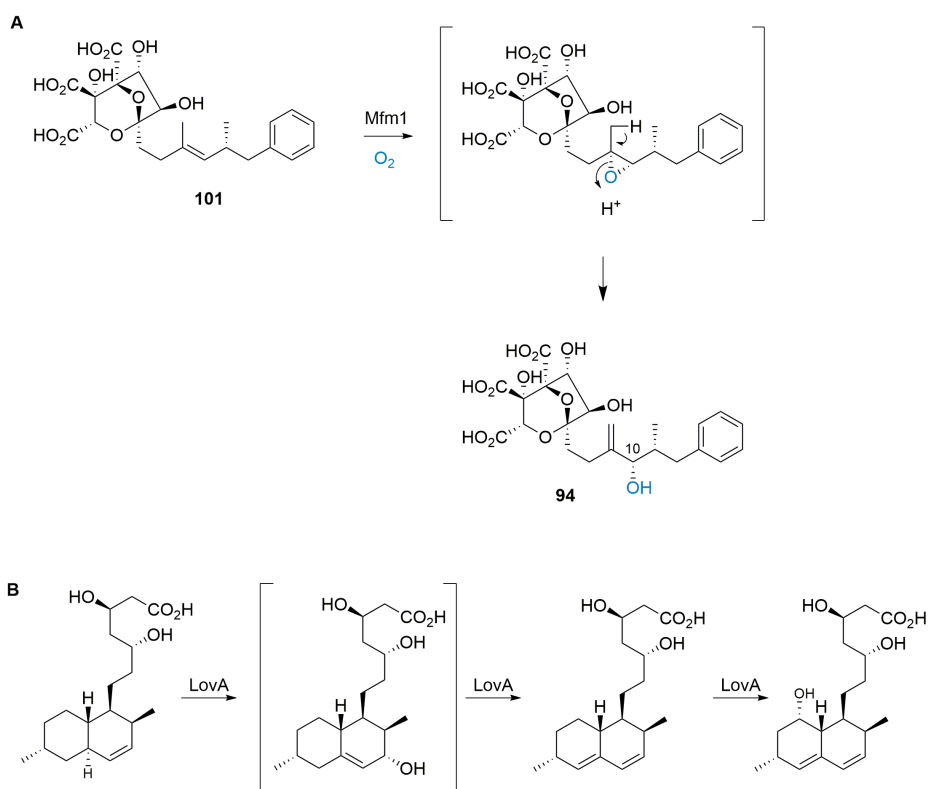
Scheme 2.19 Biosynthetic steps in rubratoxin **118** biosynthesis catalysed by four α -ketoglutarate and iron(II)-dependent dioxygenases.

Identification of a copper dependent oxygenase Mfm1 in SQS1 biosynthesis

By disrupting *mfm1*, which encodes a putative copper dependent oxygenase, it was shown that Mfm1 is part of the SQS1 cluster as no **1** was detected in the KO transformant. Furthermore structure elucidation of the accumulated compound **101** of KO strain, lacking the hydroxyl group at C-10, led to the conclusion that Mfm1 is involved in oxidation of the C-9,10-olefin of **101** in order to give the hydroxylated, unacetylated SQS precursor **94**. The mechanism probably involves epoxidation by Mfm1 and subsequent rearrangement to allylic alcohol **94** (Scheme 2.20 A). Over 20 years ago several squalestatin analogues were found that lack the oxygen at C-10 and contain a 9,10-vinyl methyl group at C-9 instead of a methylene.¹⁶⁶ The existence of these compounds already led to the proposal that, after ring formation, an oxidation of the C-9,10 olefin results in a hydroxyl group at C-10 and a double bond at C-9,18.¹¹³ This proposed biosynthetic step was elucidated and confirmed in this work. Literature reports a similar epoxidation/rearrangement sequence during the biosynthesis of lovastatin **6**, which is performed by the cytochrome P450 oxygenase LovA (Scheme 2.20 B).

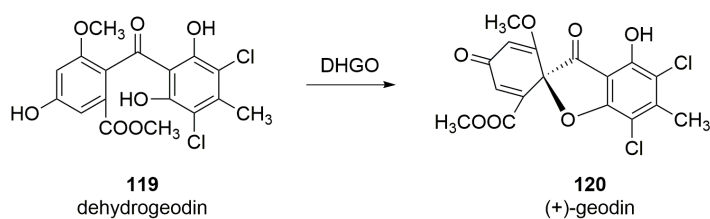
The $\Delta mfm1$ strain showed also production of a new compound **110**. NMR data confirmed the structure to be **101** with the tetraketide chain already attached. Former *in vitro* studies showed that the tetraketide **24** was attached to the core of **94** as CoA activated compound by acyltransferase Mfm4 in the final step of the biosynthesis (Section 2.1).⁵³ In this *in vivo* experiment the tetraketide chain was attached before hydroxylation and acetylation of the hexaketide chain as the gene expressing

Mfm1 was blocked. Presumably **101** was used as substrate of acyltransferase Mfm4 as its natural substrate was not present and **101** was available in high titres.



Scheme 2.20 A, Biosynthetic step from **101** to **94** catalysed by Mfm1 in SQS1 biosynthesis; B, Biosynthetic step from lovastatin biosynthesis catalysed by LovA.

Moreover, fermentation studies using different trace elements, strongly suggested copper as cofactor for Mfm1, as predicted from structure homology studies. Copper is a rarely used metal cofactor in enzyme catalysis.¹⁶³ It has relevance to metabolism in a few examples such as cytochrome c oxidase,¹⁶⁷ laccase¹⁶⁸ and tyrosinase.¹⁶⁹ There are also known fungal-derived types, like copper dependent polysaccharide monooxygenases (PMOs), that catalyze the O_2 -dependent oxidative cleavage of recalcitrant polysaccharides.¹⁷⁰ The oxidative chemistry catalysed by fungal PMOs likely leads to incorporation of molecular oxygen into C–H bonds.¹⁷¹ The metal center of these enzyme class contains a conserved type II copper binding site. Bound copper is coordinated by a bidentate N-terminal histidine residue through the amino terminus and $N\delta$ and the $N\epsilon$ of another histidine residue and a water molecule together with a buried proximal tyrosine.¹⁷⁰ The same residues were also found in Mfm1 structure prediction using PHYRE2¹³⁴ (Section 2.3). Another copper dependent enzyme which is known from fungal secondary metabolism is dihydrogeodin oxidase (DHGO). The purified enzyme showed an intense blue color and had an absorption maximum at 600 nm. It was shown to catalyse the reaction of dehydrogeodin **119** to (+)-geodin **120** in the final step of the biosynthesis in *A. terreus* (Scheme 2.21).¹⁷²



Scheme 2.21 Last step of (+)-geodin **120** biosynthesis catalysed by DHGO.

The PHYRE2 calculated model of Mfm1 showed structural homology to two oxidoreductases. The first one is a copper dependent peptidylglycine alpha-hydroxylating monooxygenase2 (PHM) from *Rattus norvegicus*¹⁴² with 15% identity and 90.9% confidence in the model. The second is an human dopamine beta-hydroxylase, which is the enzyme which converts dopamine to norepinephrine,¹⁷³ with 13% identity and 89.1% confidence. Both enzymes possess two copper-binding sites in the closed active site structure (Section 2.3).^{140,142,173}

Accompanying the bioinformatic analysis, the studies with different trace salts in the medium led to the conclusion that Mfm1 is a copper-dependent oxygenase which catalyses one of the later steps in SQS1 biosynthesis by conversion of **101** to **94** (Scheme 2.20).

Last steps of SQS1 biosynthesis are unravelled – KO of Mfr4 acyltransferase

Previous studies of SQS biosynthesis have already suggested that modifications of the acyl and alkyl chain of **1** occur late in biosynthesis as all minor components reported for squalestatins vary in the degree of oxidation and methylation of the acyl side chain and alkyl chain.¹⁶⁶

Moreover, the appearance of 4'-desacetyl- and 4'-desacetoxy-ZA-A and -C suggests that the acetylation at the hexaketide chain may be one of the final steps in SQS1 biosynthesis.¹¹³ Bonsch *et al.* showed by *in vitro* studies that acyltransferase Mfm4 is responsible for loading acyl groups from CoA onto the squalestatin core at O-6 as the final step of the biosynthesis.⁵³ Knocking out the gene for the other transferase Mfr4 led to loss of **1** and identification of intermediates **90**, **101** and **94**.

These findings suggested that the Mfr4-catalysed step is the penultimate step of biosynthesis. The results from Bonsch *et al.* that Mfm4 catalyses the last step was confirmed, as no intermediate was found with added acyl chain in the extract of Mfr4 KO transformant.

2.6 Results – Heterologous expression of SQS genes in *A. oryzae* NSAR1

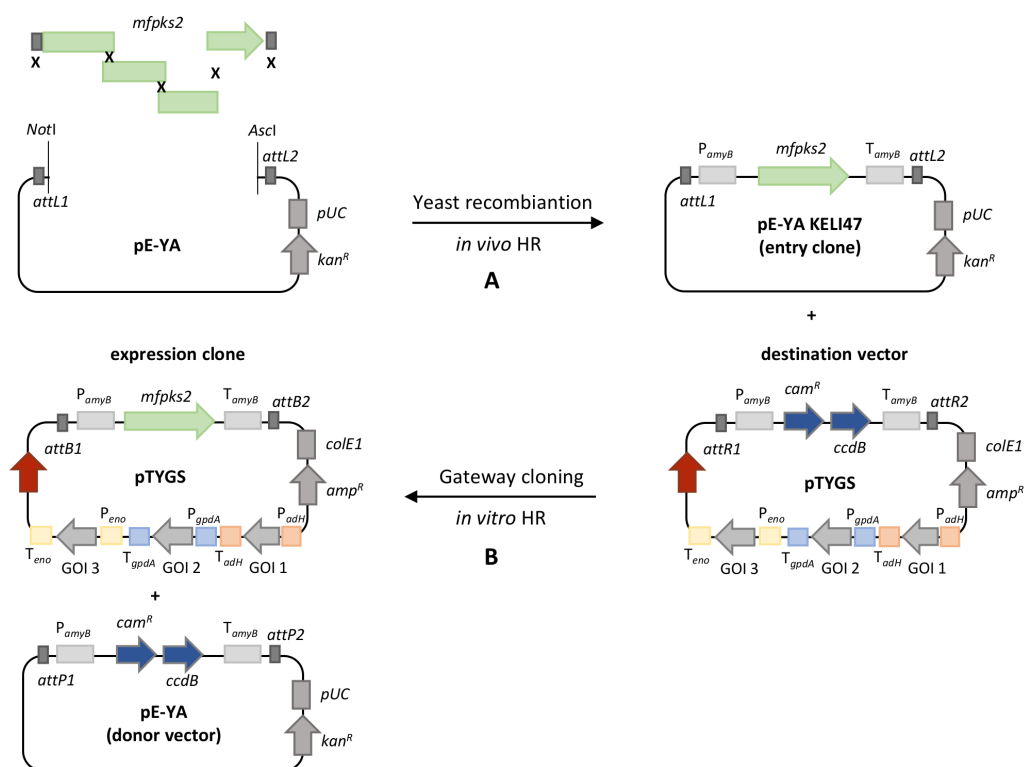
Previous knockout studies revealed the order of genes involved in SQS1 biosynthesis (Section 2.5). In order to gain more information, heterologous expression experiments were deployed using quadruply auxotrophic *A. oryzae* as the host and the modular fungal expression system described by Lazarus and coworkers (Section 1.2.3).⁹⁸

2.6.1 Yeast recombination, Gateway cloning and Transformation of *A. oryzae*

In this study the fungal expression vectors pTYGSarg, pTYGSade and pTYGSmet were used, which each contain a corresponding gene (*e.g. argB*) to complement one of the NSAR auxotrophies, allowing selection on minimal media lacking, *e.g.* arginine (Section 1.2.3).

As shown for construction of knockout cassettes (Section 2.5), pE-YA can be used for the assembly of large DNA fragments by yeast homologous recombination. Here, the vector pE-YA was used to assemble the large *mfpks2* gene (7.7 kb), from four DNA fragments of around 2 kb (Scheme 2.22 A). The assembled PKS gene was then transferred to a fungal expression vector by Gateway cloning (Section 1.2.3).

Therefore the entry vector (pE-YA) containing *mfpks2* (KELI47) and one vector of the pTYGS family (destination vector) were incubated with the LR recombinase enzyme mixture, resulting in two new plasmids. The desired expression clone contains *mfpks2* and the donor vector occupies the Gateway cassette *ccdB*, *cam*^R (Scheme 2.22 B).



Scheme 2.22 Fungal expression vectors used in this work and gateway cloning: **A**, assembly of the large PKS gene *mfpks2* by *in vivo* HR recombination (*S. cerevisiae*) in the gateway entry clone pE-YA KELI47; **B**, Gateway cloning by LR recombination between pE-YA KELI47 and pTYGS (destination vector).

Cloning overview of fungal expression vectors

According to the knockout results and the proposed biosynthesis of SQS1, the first gene involved in the biosynthesis is *mfpks2* encoding the hexaketide synthase (SQHKS). It also appears likely that there must be a system for creation of benzoyl CoA from phenylalanine. As SQHKS possesses no thiolesterase (TE) domain for off-loading of the PKS chain, one of the two genes encoding hydrolases Mfm8 and Mfm10 could also be involved in the early steps, as well as the citrate synthase Mfr3.

For initial expression experiments *mfpks2* was assembled from four cDNA fragments (RNA was isolated by Beate Bonsch) in pE-YA using homologous recombination in *S. cerevisiae* (Scheme 2.22 A, KELI47) and was further cloned into the fungal expression vector pTYGSarg *via* gateway cloning (Scheme 2.22 B, KELI47LR). Genes from MF5453 SQS1 gene cluster with predicted function for encoded proteins used in heterologous expression experiments are summarised in table 2.7. For further sequential coexpression of the early genes from the pathway different combinations (*mfpks2*+*mfm10*, *mfpks2*+*mfm10*+*mfm8*, *mfpks2*+*mfm10*+*mfr3*, *mfpks2*+*mfm8*+*mfr3*) were cloned in pTYGSarg (Fig. 2.46). The three genes involved in putative oxidations encoded by *mfr1*, *mfr2*, *mfm1* were cloned into the pTYGSmet vector in different combinations (*mfr1*, *mfr2*, *mfr1*+*mfr2*, *mfr1*+*mfr2*+*mfm1*, Fig 2.46). All

2.6 Results – Heterologous expression of SQS genes in *A. oryzae* NSAR1

Table 2.7 Genes from MF5453 SQS1 gene cluster with predicted function for encoded proteins used in heterologous expression experiments

MF5453 Gene	Putative function
<i>m1</i>	Copper dependent oxygenase
<i>m8</i>	Hydrolase
<i>m10</i>	Hydrolase
<i>mfpks2</i>	Hexaketide synthase
<i>r1</i>	Possible non-heme iron-dependent
<i>r2</i>	Possible non-heme iron-dependent
<i>r3</i>	Citrate synthase

pTYGS vectors for expression of the SQS1 BGC were constructed by *in vivo* homologous recombination in *S. cerevisiae* and *in vitro* recombination (Section 1.2.3, Scheme 1.20). Dedicated genes were amplified by PCR using Mf5453 cDNA or gDNA as template with denoted oligonucleotides (Table 2.8), which were designed to introduce the 30 bp homologous overlaps (oligonucleotide sequences: Chapter 6, Table 6.8).

Table 2.8 Construction details for fungal transformation vectors used for heterologous expression of SQS genes in *A. oryzae* NSAR1

Construct ID	Vector backbone	Template	Oligonucleotides for construction in <i>S. cerevisiae</i>
KELI47	pE-YA	cDNA	<i>sqhks</i> part1: P57+861; <i>sqhks</i> part2: P862+863; <i>sqhks</i> part3: P864+60; <i>sqhks</i> part4: P61+62;
KELI47LR	pTYGSarg	-	LR with KELI47
KELI31	pTYGSarg	cDNA	<i>mfm10</i> : P607+608; <i>mfr3</i> : P609+P610
KELI31LR	pTYGSarg	-	LR with KELI47
KELI34 (BB)	pTYGSarg	cDNA	Beate Bonsch
KELI34LR	pTYGSarg	-	LR with KELI47
KELI42	pTYGSarg	cDNA	<i>mfm10</i> : P607+608; <i>mfm8</i> : P841+P842
KELI42LR	pTYGSarg	-	LR with KELI47
KELI50	pTYGSarg	cDNA	<i>mfr3</i> : P609+610; <i>mfm8</i> : P886+P887
KELI50LR	pTYGSarg	-	LR with KELI47
KELI52A	pTYGSmet	cDNA	<i>mfr1</i> : P613+614; Patch_PadH: P89+90; Patch_Peno: P87+88
KELI52B	pTYGSmet	cDNA	<i>mfr2</i> : P615+616; Patch_PadH: P89+90; Patch_PgdA: P91+92
KELI52C	pTYGSmet	cDNA	<i>mfr1</i> : P613+614; <i>mfr2</i> : P615+616; Patch_PadH: P89+90
KELI52D	pTYGSmet	cDNA	<i>mfr1</i> : P613+614; <i>mfr2</i> : P615+616; <i>mfm1</i> : P611+612

Correct cloning and splicing was confirmed by sequencing (Eurofins, Ebersberg). An additional vector pTYGSade-*str11-str8-str10* (provided by Dr. Risa Nofiani from Bristol) contains the benzoyl CoA biosynthesis-encoding genes from *Strobilurus tenacellus* (Chapter 3). Coexpression in *A. oryzae* enables production of benzoyl CoA and complements the panel of expression clones used for subsequent *A. oryzae* NSAR1 transformation (Fig. 2.46).

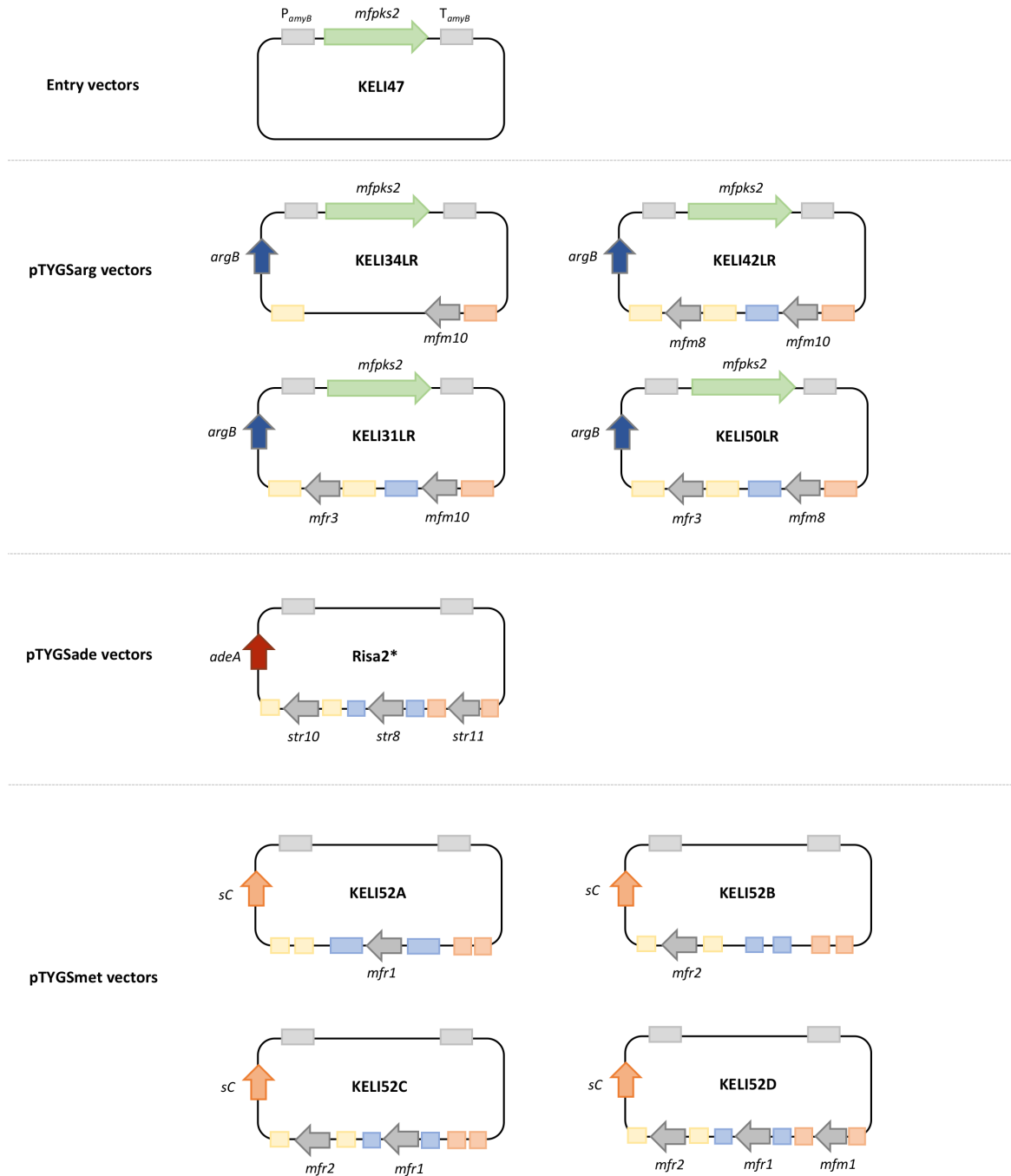


Figure 2.46 Overview of constructed vectors used for fungal expression (pTYGSarg, met, ade) and gateway entry vector pE-YA (KELI47).

Gene combinations, *A. oryzae* NSAR1 transformation and genetic analysis of transformants

In order to transform *A. oryzae* NSAR1 with combinations of constructed vectors (Section 2.6.1) a CaCl_2 /PEG mediated protoplast protocol was used. Cotransformation of one, two or three vectors were carried out in a single transformation step using selection medium which was adapted to the

vectors used in the individual transformation. The resulting colonies were picked onto secondary selection plates and further separated by streaking for single colonies on a third selection plate. Overall 10 combinations of genes from SQS1 BGC on different fungal expression vectors were used to cotransform *A. oryzae* NSAR1, resulting in 10 strains carrying genes from 1 BGC (Table 2.9).

Table 2.9 Overview of *A. oryzae* NSAR1 transformations with different combinations of genes from SQS1 BGC, ^a Coexpression with genes from strobilurin biosynthesis (*str10*, *str8*, *str11*) encoding benzoyl-CoA producing enzymes; Hyd: Hydrolase, CS: citrate synthase, NHI: non-heme iron-dependent oxygenase, CDM: copper dependent monooxygenase

Exp. No	Exp. ID	Selection marker	Transformed expression vector(s)	Genes mf							Fed benzoyl SNAC
				pks2 PKS	m10 Hyd	m8 Hyd	r3 CS	r1 NHI	r2 NHI	m1 CDM	
1	KELI47	arg	KELI47LR	X	-	-	-	-	-	-	X
2	KELI34	arg	KELI34LR	X	X	-	-	-	-	-	X
3	KELI42	arg	KELI42LR	X	X	X	-	-	-	-	X
4	KELI31	arg	KELI31LR	X	X	-	X	-	-	-	X
5	KELI5-	arg	KELI5-LR	X	-	X	X	-	-	-	X
6	KELI52	arg, ade	KELI5-LR, Risa2	X	-	X	X	-	-	-	-
7	KELI52A ^a	arg, ade, met	KELI52A, Risa2	X	-	X	X	X	-	-	-
8	KELI52B ^a	arg, ade, met	KELI52B, Risa2	X	-	X	X	-	X	-	-
9	KELI52C ^a	arg, ade, met	KELI52C, Risa2	X	-	X	X	X	X	-	-
1-	KELI52D ^a	arg, ade, met	KELI52D, Risa2	X	-	X	X	X	X	X	-

The genomic DNA of *A. oryzae* NSAR1 expression strains was isolated and analysed by PCR for the intended integration of genes using gene specific oligonucleotides. Subsequent bands corresponding to DNA fragment sizes of amplified genes confirmed the integration of genes within the *A. oryzae* NSAR1 genome (Appendix, Fig. 7.1). In order to analyse the secondary metabolite produced by genetically confirmed transformants, the colonies were grown in induction medium for P_{amyB} (DPY + 0.15% methionine) for 6 days at 28 °C and 110 rpm. The extracts were analysed by LCMS.

2.6.2 Expression of early SQS genes in *A. oryzae*

From previous knockout studies no hexaketide product was observed in extracts of $\Delta mfr3$ (citrate synthase) mutants (Section 2.5.6). In order to investigate the function and products of the early genes of the cluster, SQHKS encoding gene (*mfpks2*) was cloned into four pTYGSarg expression vectors (arg) possessing different sets of tailoring genes (Table 2.9). The first constructed expression vector contained the PKS gene only (KELI47LR), whereas the second one additionally contained the hydrolase gene *mfm10* from the cluster (KELI34LR). The hydrolase gene was included into the expression system as SQHKS possesses no thiolesterase domain. Therefore Mfm10 could be necessary for off-loading of the PKS chain from the enzyme. The third expression vector consists of both hydrolase genes of 1 BGC (*mfm10* and *mfm8*) in addition to *mfpks2*. Either hydrolase could

possibly be involved in off-loading of the PKS chain. In favour of eliminating the opportunity that the PKS product on its own is unstable or degraded, the putative next acting enzyme of the biosynthesis, *mfr3* (citrate synthase) was included into the expression system with *mfpks2* and *mfm10* (KELI31LR). The expected product would be a compound like **90**, which was the major intermediate of the double *mfr1+mfr2* KO strain (Section 2.5.3).

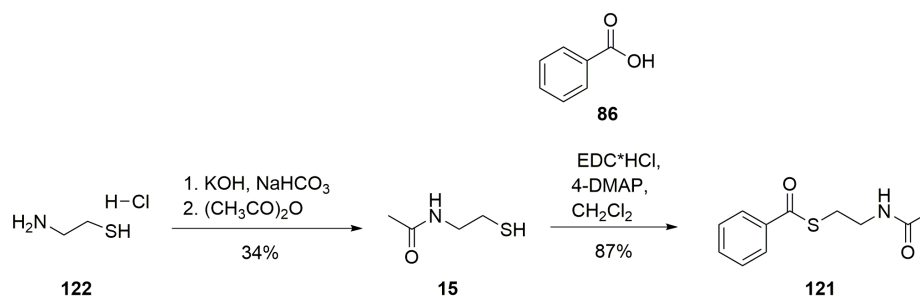
Genetically confirmed transformants and an *A. oryzae* NSAR1 wt control were grown in liquid P_{amyB} induction medium. Because of missing genes for the predicted benzoyl-CoA **25** precursor in *A. oryzae*, benzoyl-SNAC **121** was synthesised and fed to the medium [final concentration 1 mM] on days 2, 3 and 4 of fermentation, as a mimic of benzoyl-CoA **25** (Section 1.1.1). The synthesis of benzoyl-SNAC **121** is described in this section. After 5–7 days of fermentation, extracts were submitted to LCMS and subsequent chromatograms analysed for early intermediates of SQS1 pathway. In comparison to *A. oryzae* wt control grown under identical conditions, no new peaks were detected for all four experiments (Appendix, Figure 7.3). As genetic analysis revealed that all desired genes were integrated successfully, it is likely that the synthetic precursor benzoyl-SNAC **121** did not reach SQHKS or is not accepted by the synthase. A stability problem of the hexaketide product can be almost certainly excluded as coexpression of the citrate synthase did not lead to the putative product **90**, which was shown before to be stable in MF5453.

While conducting these experiments, Tang and coworkers showed that intermediate **90** is the first enzyme-free intermediate of SQS1 biosynthesis and only three enzymes (SQHKS, citrate synthase and one hydrolase) are needed to produce **90** heterologously in *A. nidulans*.¹¹⁸ They identified a homologous BGC of **1** in the fungal pathogen *Curvularia lunata* (also known as *Cochliobolus lunatus* ATCC 74067), which was previously identified as a producer of **1** (Section 2.1.2).¹¹⁰ In comparison to our expression experiment with *A. oryzae*, their expression host *A. nidulans* was shown to produce benzoyl-CoA **25** endogenously. Furthermore, they showed that in addition to the PKS (Clz14) and CS (Clz17), the hydrolase homologue to Mfm8 (Clz11) is also required to produce **90**.¹¹⁸ The homologous gene combination from MF5453 **1** BGC (*sqhks*, *mfm8* and *mfr3*) was not tested in our experiments and therefore, in agreement with results of the Tang group, initial studies to reconstitute the SQS1 biosynthesis resulted in no intermediates of the pathway. However, Tang and coworkers isolated **90** in very low yields (titre of around 0.1 mg/L) which might arise from using the low natural pool of benzoyl-CoA **25** occurring in *A. nidulans*.

Our aim was to repeat Tang's expression experiment using homologous genes from SQS producing strain MF543 in order to reproduce their result and moreover achieving better yields for **90** to have a good starting point for adding more genes involved in oxidation to the expression system.

Synthesis of benzoyl–SNAC 121

Benzoyl–SNAC 121 was synthesised, in order to provide a synthetic CoA–mimic of naturally precursor benzoyl–CoA which can be fed to the expression medium of SQHKS transformants. Benzoyl–SNAC 121 was synthesised from its corresponding carboxylic acid 86 in two steps with an overall yield of 30% (Scheme 2.23). The formation of HSNAC 15 was carried out following the protocol of Piasecki *et al.*¹⁷⁴ Corresponding experimental details and NMR data are shown in chapter 6.5.5.



Scheme 2.23 Synthesis of benzoyl–SNAC 121: first step, synthesis of *N*-acetylcysteamine (15) from cysteamine hydrochloride 122 and acetic anhydride yielded in 121 (34%); second step, synthesis of 121 from benzoic acid 86 (87% yield).

2.6.3 Coexpression of *mfpks2*, *mfm8*, *mfr3*

In order to reproduce Tang’s result, a fungal expression plasmid was constructed containing *mfpks2*, *mfr3* and *mfm8* using yeast recombination and LR recombination (KELI50LR, exp. 5/6). The desired plasmid was confirmed by sequencing and used for transformation of *A. oryzae* NSAR1 resulting in 5 transformants. As *A. oryzae* is not capable of producing benzoyl–CoA 25¹⁷⁵, a mimic of the starter unit, benzoyl–SNAC 121 was pulse fed to the fermentation (day 2, 3, 4; 1 mM final concentration). On day six the culture broth was extracted using ethyl acetate after acidifying to pH 4. As in previous experiments, no new product was detected by LCMS compared to a NSAR1 wt control (Fig. 2.47).

2.6 Results – Heterologous expression of SQS genes in *A. oryzae* NSAR1

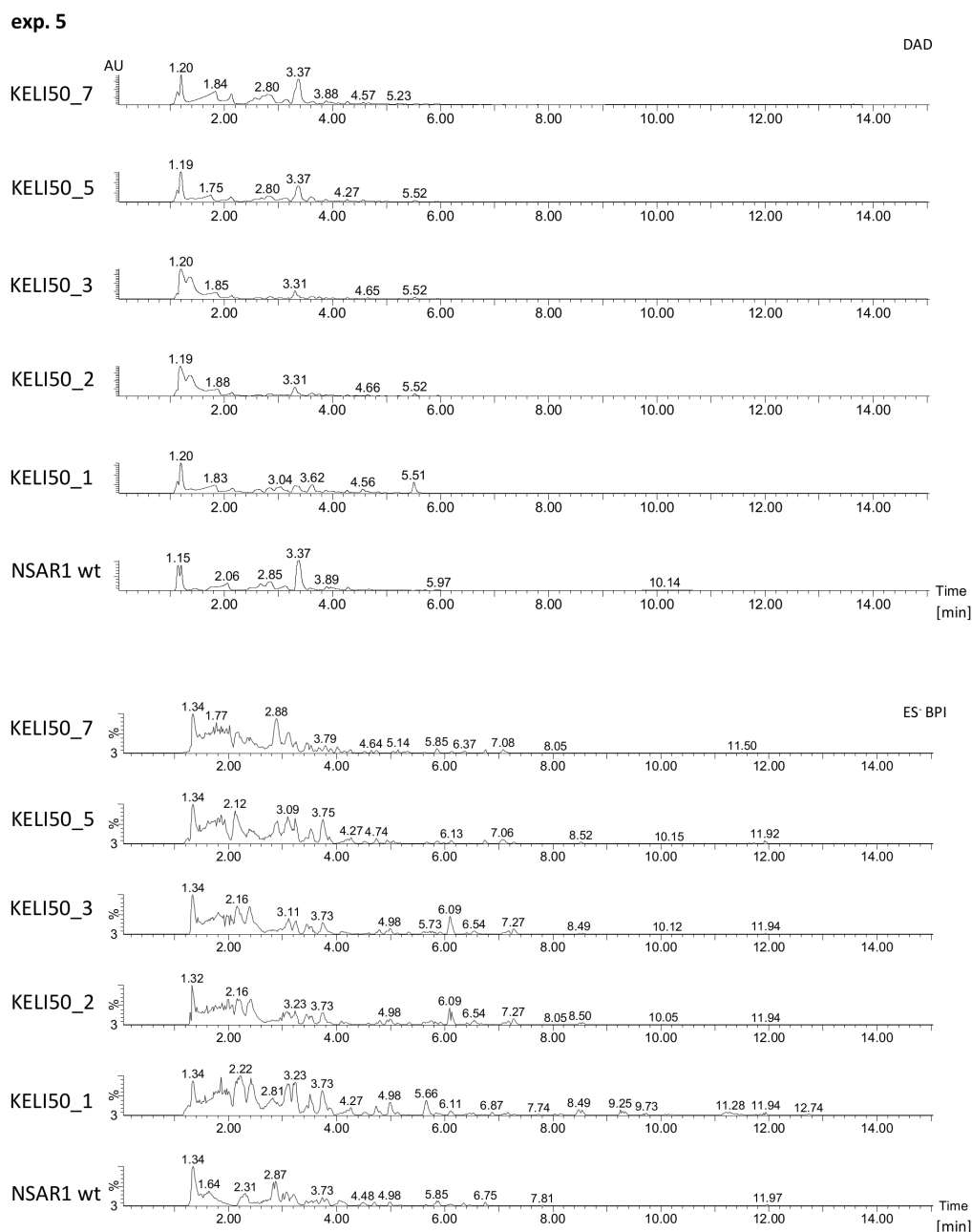


Figure 2.47 LCMS analysis of transformants containing 3 genes (*mfpks2*, *mfr3* and *mfm8*) in comparison to NSAR1 wt (exp. 5, table 2.9), DAD chromatograms (top) and ES⁻ traces (bottom), *unrelated compound ($t_R = 5.5$ min), m/z 163 (ES⁻), bottom: peak ($t_R = 6.1$ min) also in wt control (traces), extracts were obtained from cultures grown in DPY + 0.15% methionine and pulse fed with 1 mM benzoyl-SNAC **121** and grown for 6 days.

To overcome the possibility that **121** is not accepted as starter unit for SQHKS and therefore no expression product could be detected in previous experiments, a different strategy was used. In the following expression experiments 6 to 10 (Table 2.9), the precursor **25** was produced *in situ* by coexpression of benzoyl-CoA biosynthesis-encoding genes (*str8*, *str10*, *str11*) from *Strobilurus tenacellus* (Chapter 3). The plasmid containing the **25** producing genes was provided by Risa Nofiani

(Cox group, University of Bristol) on an pTYGSade expression vector (Fig. 2.46, Risa2). Additional information about the benzoate pathway in *S. tenacellus* is given in chapter 3. Cotransformation of *A. oryzae* NSAR1 with pTYGSarg-*mfpks2*-*mfr3*-*mfm8* (KELI50LR) and pTYGSade-*str8*-*str10*-*str11* (Risa2) resulted in 8 transformants. Transformant 6 was shown to have all desired genes integrated (Appendix, Fig. 7.2). Four strains (Transformant 4, 5, 6 and 7 (traces)) were shown to produce a new compound eluting at $t_R = 8.0$ min (Fig. 2.48). The retention time corresponded to the known compound **90** which was previously isolated from *mfr1+2* KO mutant. Further chemical characterisation by comparing the MSMS fragmentation pattern to a synthetic standard of **90** proved the new peak to be hexaketide citrate **90** (Fig. 2.49). Purification of compound **90** from 1 L fermentation of transformant KELI50_6 yielded 1 mg of pure **90**. The titre of **90** was around ten times higher than in Tang's experiment (titre ca 0.1 mg/L).

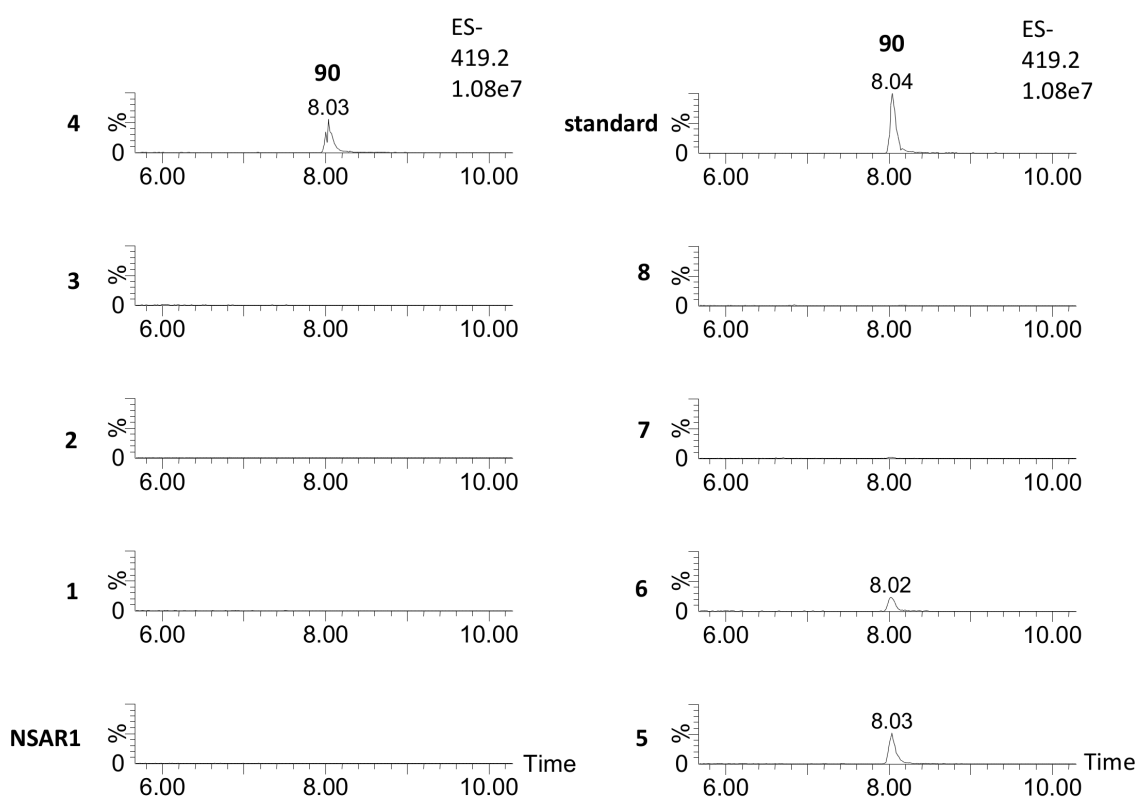


Figure 2.48 LCMS analysis of transformants containing 6 genes (*mfpks2*, *mfr3*, *mfm8*, *str8*, *str10*, *str11*) in comparison to NSAR1 wt and an analytical standard of **90** obtained from *mfr1+2* KO mutant (exp. 6, table 2.9), extracted ion chromatograms of 419.2, extracts were obtained from cultures grown in DPY + 0.15% methionine.

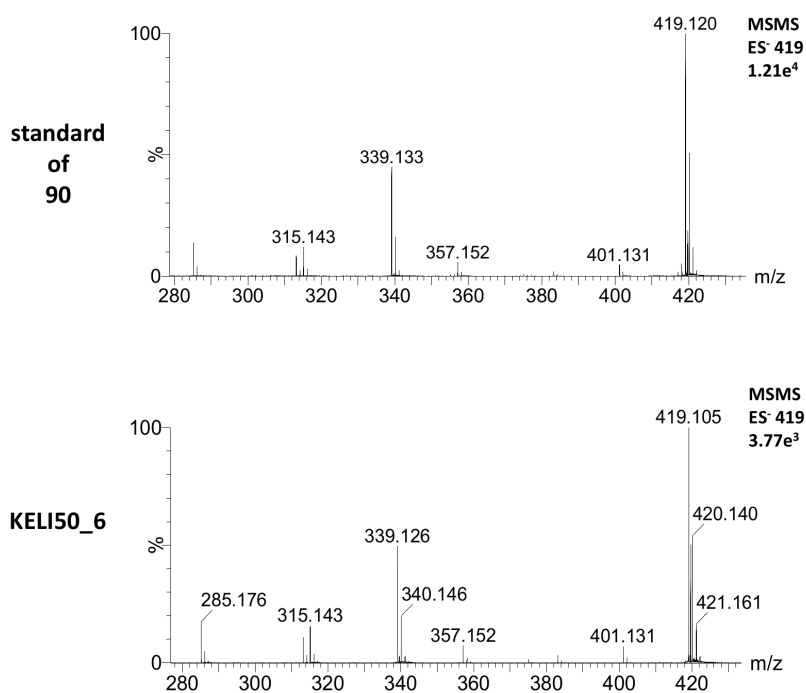


Figure 2.49 MSMS data of compound produced in transformant KELI50_6 ($t_R = 8.0$ min) in comparison to an analytical standard of 90 obtained from *mfr1+2* KO mutant.

2.6.4 Expression of genes involved in oxidations during SQS1 biosynthesis

Previous knockout studies revealed that Mfr1 and Mfr2, two putative non-heme iron dependent oxygenases, must be involved in the several oxidative steps which are required to form the core of **1**. A copper-dependent oxygenase Mfm1 was then shown to oxidise the hexaketide chain after cyclisation to give intermediate **94**. In order to investigate the oxidative steps in more detail, *mfr1*, *mfr2* and *mfm1* were added to the expression system in various combinations (Table 2.9, exp. 7 to 10).

Yeast recombination was used to construct four different constructs for expression in *A. oryzae*, pTYGSmet-*mfr1* (KELI52A), pTYGSmet-*mfr2* (KELI52B), pTYGSmet-*mfr1-mfr2* (KELI52C) and pTYGSmet-*mfr1-mfr2-mfm1* (KELI52D) (Fig. 2.46, Table 2.7). The vector pTYGSmet was chosen for all constructs as it can be coexpressed with the expression system producing precursor **90** (pTYGSarg-*sqhks-mfr3-mfm8* + pTYGSade-*str8-str10-str11*) and transformants can be selected on medium lacking arginine, adenine and methionine.

Addition of *mfr2* to the expression system producing **90** resulted in no new compound (Fig. 2.50, exp. 8), but expressing *mfr1* in presence of **90** resulted in a mixture of several new compounds in very low titres with m/z values of 432.2, 434.2, 435.2 and traces of 447.2 (Fig. 2.50, exp. 7). HRMS data confirmed these to be oxidised congeners of **90**, which were also found in *mfr2* knockout experiment

(Section 2.5). An overview of compared HRMS and MSMS data of new compounds isolated from $\Delta mfr2$ strain and expression experiment 7 is summarised in figure 2.51. Retention times of new expression compounds slightly vary (max. 0.2 min) from former KO studies, as measurements were carried out using a new LCMS column.

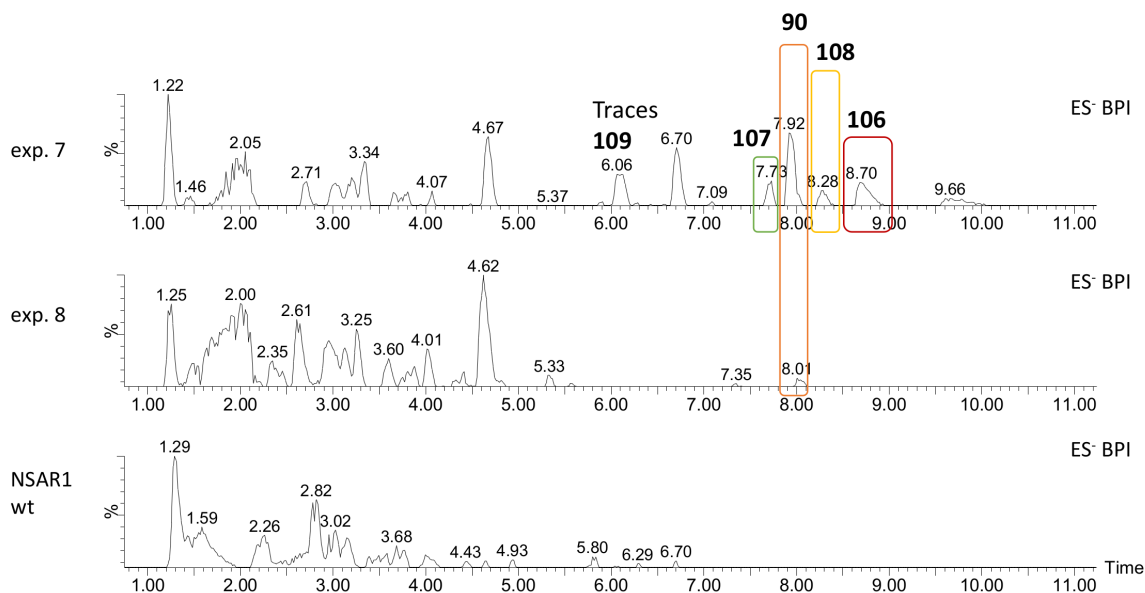


Figure 2.50 Chemical analysis of transformant KELI52_B1x1 (exp. 8, table 2.9) and KELI52_A3x4 (exp. 7, table 2.9) in comparison to an *A. oryzae* wt control by LCMS (ES⁻ BPI chromatograms shown); in orange, 90; in yellow, 108; in red, 106; in green, 107; and in traces 109; retention times are not identical to former KO studies (max. variation = 0.2 min), because measurements were carried out using a new LCMS column.

When both putative non-heme iron expressing genes (*mfr1* and *mfr2*) were coexpressed with the established expression system for 90 production, two additional new compounds 101 (m/z 479.2, ES⁻), 103 (m/z 463.2, ES⁻) were detected in low titres (Fig. 2.52, exp. 9). Using HRMS and MSMS analysis they were shown to be identical to the compounds previously isolated from *mfm1* knockout strain (Fig. 2.53). Subsequent addition of *mfm1* to the system did not result in any new compound (Fig. 2.52, exp. 10). This outcome was not unexpected as KO experiments with *mfr4* revealed the expected compound 94 only to be produced in low titres. This suggests it to be unstable and therefore not possible to detect in the extract of the expression system. In addition the titre was already low for expression products of coexpression of up to 9 genes.

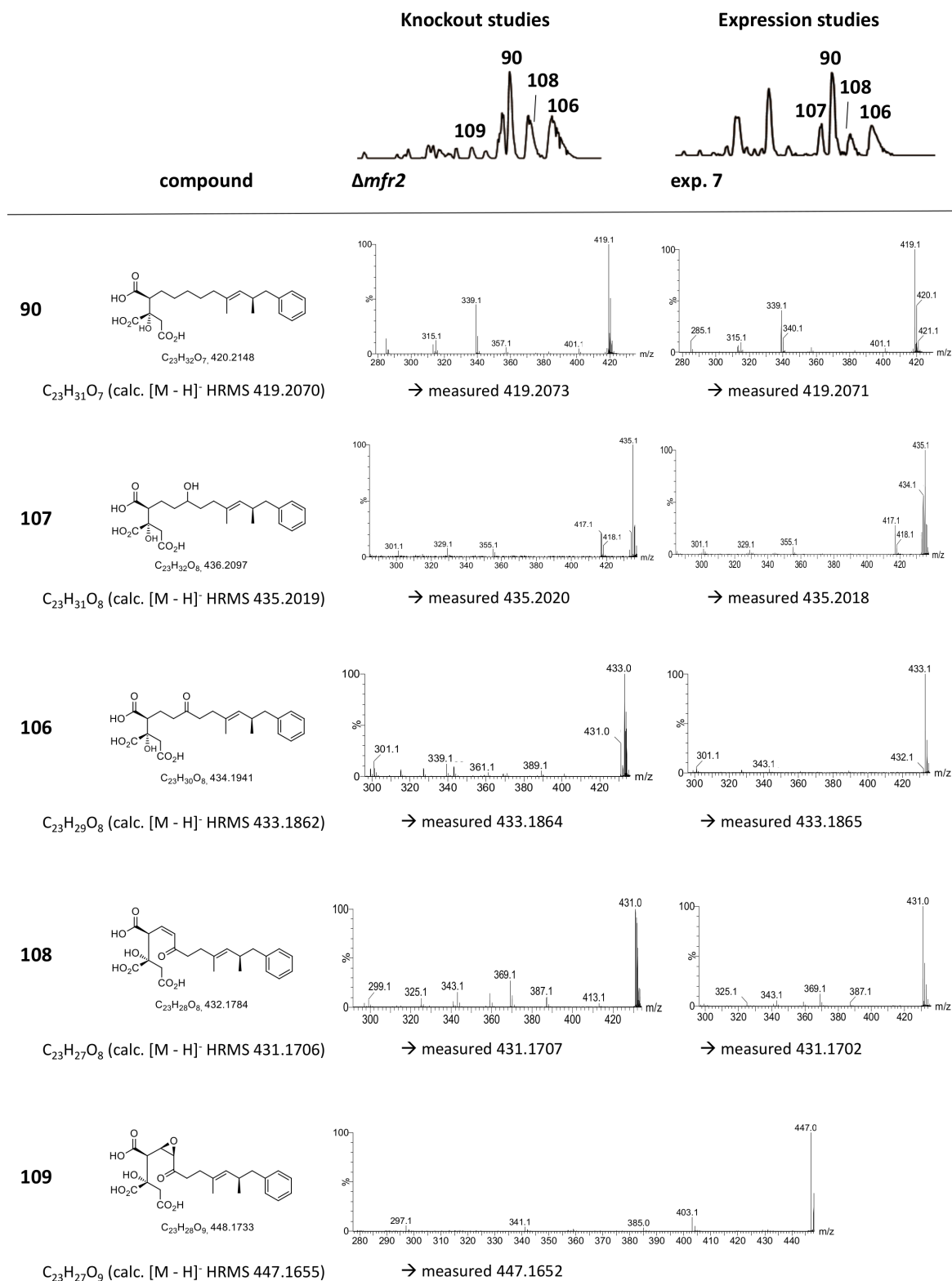


Figure 2.51 Comparison of compounds isolated from *Δmfr2* mutant and expression experiment 7 (*mfr-pks2+mfr3+mfm8+str10+str8+str11+mfr1*). HRMS and MSMS fragmentation pattern of new compounds, HRMS and MSMS data of **109** from expression strain is missing due to very low amount produced.

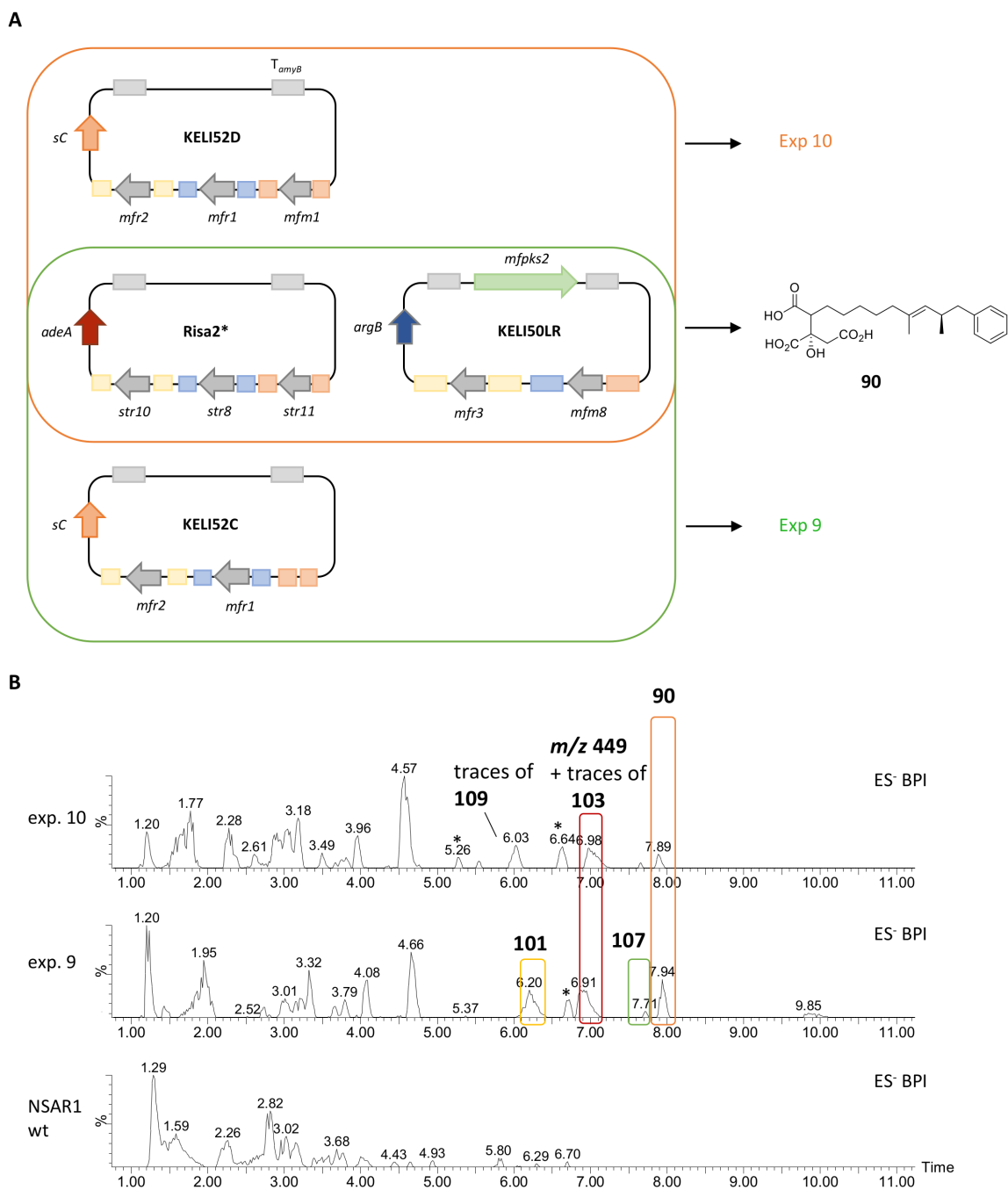


Figure 2.52 A, coexpression overview; B, chemical analysis of transformant KELI52_C1x2 (exp. 9) and KELI52_D3x3 (exp. 10) in comparison to an *A. oryzae* wt control by LCMS (ES⁻ BPI chromatograms shown); in orange, **90**; in green, **107**; in yellow, **101**; in red: **103** (coelution with m/z 449 ES⁻); and traces of **109**; retention times are not identical to former KO studies (max. variation = 0.2 min), because measurements were carried out using a new LCMS column.

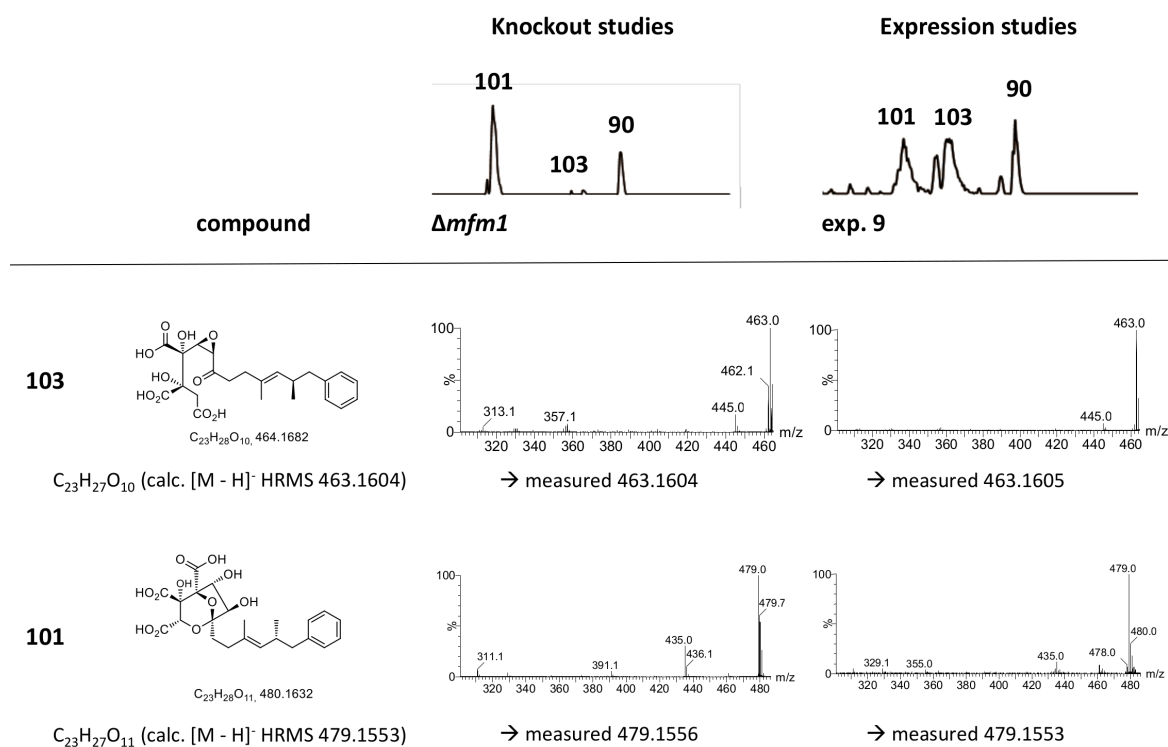


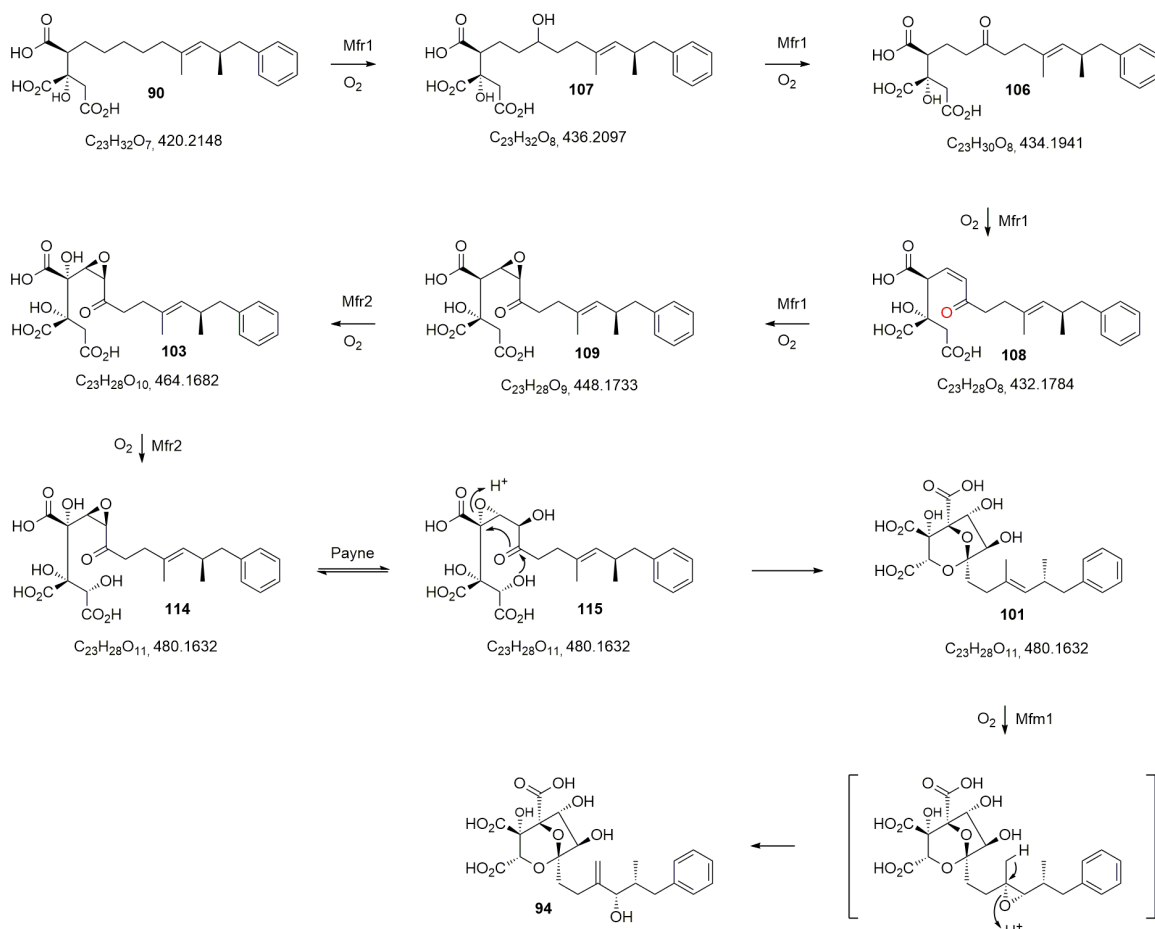
Figure 2.53 Comparison of compounds isolated from *Δmfm1* mutant and expression experiment 9 (*mfpks2+mfr3+mfm8+str10+str8+str11+mfr1+mfr2*). HRMS and MSMS fragmentation pattern of new compounds.

Isolation and therefore full structure elucidation of new produced oxidative congeners of **90** was prevented due to very low titres in expression strains. Nevertheless, it was shown that heterologous expression of up to 6 genes from SQS BGC, coexpressed with benzoyl-CoA producing genes from strobilurin biosynthesis, yielded in several previously characterised oxygenated intermediates of squalastatin biosynthetic pathway. The expression results reinforce the outcome of the knock-out studies and together showing the order and function of oxidative enzymes involved in SQS1 biosynthesis.

2.6.5 Discussion

Expression results confirmed the outcome of the knockout studies and showed that **90** is the substrate for stepwise oxidation on C-5 by putative non-heme iron dependent oxygenase Mfr1, most likely to an alcohol **107**, a ketone **106**, and then an unsaturated ketone **108**. Subsequent oxidation would result in the 3,4-epoxide **109** (Scheme 2.24). The detected intermediates were also observed in extracts of the *Δmfr2* strain. Both, the MF5453 KO strain and the *A. oryzae* expression strain (KELI52_A3x4), produced these compounds in very low titres. Therefore, the oxygenated congeners of **90** were proven extremely difficult to isolate, most likely because of facile shunts and degradations due to their electrophilic nature. However, obtained intermediates from both experiments were confirmed by t_R ,

HRMS and MSMS analysis and give confidence that these are true intermediates of the biosynthetic pathway to **1**. MSMS analysis definitively showed that the oxidative modifications performed by Mfr1 leading to **107**, **106**, **108** and **109** occur on the hexaketide and not on the oxaloacetate moiety of **90**, as retro–aldol loss of unmodified oxaloacetate (m/z 132) was detected, which was not observed for **101** indicating that hydroxylation at C-2 of **114** is probably the final oxidative step catalysed by Mfr2.



Scheme 2.24 Proposed sequence of oxidative steps during the formation of **94**.

The HRMS analysis suggests hydroxylating at C-2 and C-4 by Mfr2 to give **103** and **114**. Both intermediates were produced in low titres and due to instability of **103** full structural elucidation by NMR was prevented. An alternative pathway could involve hydroxylation at C-4 first followed by hydroxylation at C-2 (Scheme 2.24). In both cases, facile Payne rearrangement of **114** would yield **115** which could form the acetale and concomitant transannular epoxide opening would give the stable intermediate **101** (Scheme 2.24). This series of oxidation events is in accordance with previous $^{18}\text{O}_2$ feeding experiments (Section 2.1.2)⁷ and with the knockout studies (Section 2.5). Oxidation of the 8,9–olefin probably involves epoxidation catalysed by the copper–dependent oxygenase Mfm1 and rearrangement to give the allylic alcohol **94** (Scheme 2.24).

Knockout as well as expression experiments revealed that oxygenated intermediates on the pathway to form the core of **1** are highly unstable (apart from **101**). The question to be answered is how such a pathway could have evolved as the oxidised congeners of **90** appear to be transient species. However, intermediate **90** was isolated before from producing organisms as a minor cometabolite of **1** and it is known to be a micro-molar inhibitor of squalene synthase.¹¹⁹ Accordingly, the pathway to SQS1 benefits by Mfr1 and Mfr2 (possibly evolved by duplication), as these two enzymes significantly increases the potency of the pathway product, as **1** is a pico-molar inhibitor of SS.¹²

Synthetic studies towards the squalestatins demonstrated the requirement for protecting groups, when nucleophilic and electrophilic compounds are in close proximity.¹⁷⁶ This fact affirm the instability of oxidative intermediates of **90**. Nature, anyhow, appears to bypass this problem by using catalysts which can guide unstable intermediates rapidly from **90** towards stable products such as **101**. Hence, the three oxidative enzymes, Mfr1, Mfr2 and Mfm1 catalyse a remarkably efficient series of seven sequential oxidations to convert polyketide citrate **90** to the highly functionalised squalestatin precursor **94** (Scheme 2.24).

2.7 Results – Putative SQS cluster in *Dothistroma septosporum*

Previous bioinformatic studies of the SQS cluster by Beate Bonsch revealed a putative squalestatin BGC in forest pathogen *Dothistroma septosporum*, which showed similar composition of genes in comparison to the other three known BGC of **1** (Section 2.1.2).^{53,118} As no SQS-like compound was known to be produced by this fungus, a master's student (under my supervision), Ole Böhl, was aiming to isolate a new member of the SQS family. The project was carried out in collaboration with the group of Rosie E. Bradshaw from New Zealand, who kindly provided the *Dothistroma septosporum* wt strain and a PKS knockout mutant with corresponding transcriptome data.

More detailed *in silico* analysis of the genome and transcriptome data of the *Dothistroma septosporum* wt strain and the PKS knockout mutant in comparison with the MF5453 BGC revealed that the *D. septosporum* is missing a gene corresponding to *mfpks1* (tetraketide synthase). Further, transcriptome data for *D. septosporum* showed that most of the corresponding biosynthetic genes are upregulated. The only exceptions were found to be *dsM1* (hydrolase, homologous to *mfm8*) and *dsM2* (PAL, homologous to *mfm7*), which showed up to hundred times lower expression rates.¹²²

From the expression results (described in section 2.6) and Tang and coworkers results¹¹⁸, it is known that both genes are involved in the early steps of the biosynthesis to SQS1. The PAL is involved in benzoyl CoA precursor biosynthesis and the hydrolase Mfm8 catalyses the hydrolyses to give the first enzyme-free intermediate **90**. The *in silico* analysis of SQS BGC of *D. septosporum* suggested production of a SQS metabolite lacking the tetraketide chain. The fact, that two crucial

genes (encoding the PAL and hydrolase) are weakly expressed could prevent production of a SQS intermediate. However, it might also be possible that *D. septosporum* has endogenous benzoyl CoA producing genes or the precursor of the PKS is not benzoyl-CoA.

Although gene composition and transcriptome data provided evidence for SQS genes being expressed by *D. septosporum*, no new SQS compound could be isolated trying different production conditions.^{122,177}

2.8 Conclusion and Outlook

The aim of this project was to elucidate the biosynthesis of squalestatin and to delineate enzymes involved in oxidative steps to form the complex core structure of **1**. Therefore seven targeted knockout strains were generated, disrupting single genes, or two genes simultaneously, of **1** BGC.

Production of new intermediates of **1**, accumulating upon knockout experiments, revealed the order of enzymes involved in the biosynthesis of **1** and elucidated oxidative enzymes Mfr1 and Mfr2 (non-heme iron-dependent oxygenases), as well as copper dependent monooxygenase Mfm1 (Section 2.5). In addition, heterologous expression of SQS1-producing genes in *A. oryzae* NSAR1 and re-constitution of the biosynthetic pathway revealed further detailed information about the role and function of enzymes involved and is supported by structure elucidation of intermediates isolated from KO experiments.

In conclusion, three previously unknown oxygenases were elucidated, which belong to the classes of copper-dependent monooxygenases (Mfm1) and to non-heme iron-dependent oxygenases (Mfr1 and Mfr2) and catalyse a remarkably efficient series of seven sequential oxidations to convert polyketide citrate **90** to the highly functionalised squalestatin precursor **94** in order to install the bioactive core of **1**. These results, combined with previous studies by Bonsch *et al.*,⁵³ as well as those of Tang and coworkers,¹¹⁸ provide a full biosynthetic pathway for the production of **1** in which the order of enzymes encoded by the squalestatin BGC is now elucidated.

In the future, *in vitro* studies of the powerful non-heme iron-dependent oxygenases Mfr1 and Mfr2 and copper dependent monooxygenase Mfm1 would shed light on their unusual chemistry and could even provide opportunities to engineer these biocatalysts to expand nature's catalytic versatility and develop production of medicinally important molecules for drug discovery. Previous extensive attempts to express copper dependent monooxygenase encoding gene *mfm1* in *E. coli* failed.¹⁴³ Also expression of another gene from **1** BGC, CoA-ligase encoding gene *mfm9*, in *E. coli* was unsuccessful, indicating that enzymes from **1** BGC are not easily expressed in heterologous host *E. coli*. Apart from *in vitro* tests, ongoing investigation of the biosynthetic pathway of simple alkyl

citrate, viridifungin A **113**, could be the starting point for a rational genetic engineering strategy where oxidative genes of the SQS1 pathway are coexpressed with those of **113** in order to give novel, highly oxidised, viridifungins. Mix and match experiments with oxidative genes from SQS1 BGC could expand the chemical diversity of simpler natural products and may lead to access of novel bioactive compounds.

3 Biosynthetic studies of Strobilurin A

Partial results of the presented work contributed to a publication in *Nature Communications*, **2018**, *9*.¹⁷⁵

3.1 Introduction

Strobilurin A **2** (also known as mucidin) belongs to the family of strobilurins which are produced by basidiomycete fungi.⁶ Strobilurin A was first isolated in 1977 from the mycelium of *Strobilurus tenacellus* strain (No. 21602) and was shown to be highly active against yeasts and filamentous fungi.⁶ A closely related compound was isolated from *Oudemansiella mucida* and named oudemansin A **123** (Fig. 3.1). Both compounds showed complete inhibition of respiration in fungi and investigation with Ehrlich carcinoma ascitic (ECA) cells revealed that the location of action is exclusively the respiratory chain.¹⁷⁸

The common structural element, the β -methoxyacrylate, is the key toxophore which targets the Qo site of complex III of the mitochondrial electron transport chain and subsequently prevents ATP synthesis.¹⁵ Other effects such as complete inhibition of protein, DNA and RNA synthesis is due to intracellular ATP deficiency.¹⁷⁹ As a result of their simple structure and strong bioactivity against fungi, strobilurins are used as lead compounds for a number of agricultural fungicides, such as azoxystrobin **124** (Syngenta) and Kresoxim methyl **125** (BASF) (Fig. 3.1).^{179,180} Both compounds show an increased photo-stability and selectivity compared to **2** and they are among the most widely used agricultural fungicides worldwide. The strobilurin fungicides are estimated to have been worth \$ 3.4 billion per annum in 2015 and make up 25% of the fungicide market and 6.7% of the total crop protection market world-wide.

Various functionalised natural strobilurins are known, including strobilurin F **126**¹⁸¹ (hydroxylated aromatic ring) and strobilurin B **127**⁶ (chlorinated, Fig. 3.1). Related compounds such as bolineol¹⁸² **4** have also been reported, which possesses the same carbon skeleton to strobilurin A but differs in oxidation level (Fig. 3.1).¹⁸³

Even though there is great interest in strobilurins as agricultural fungicides, very little is known about their detailed biosynthesis and the genes and enzymes involved.

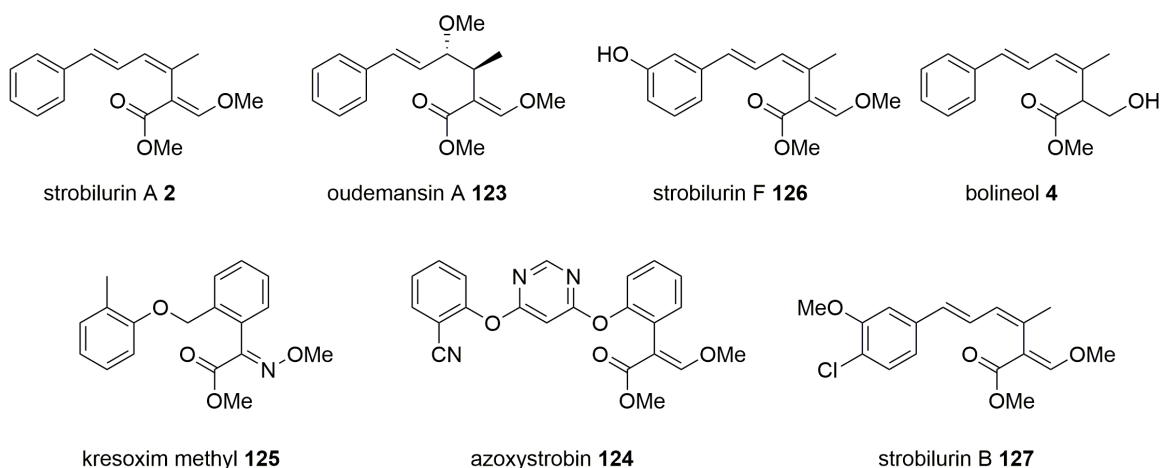


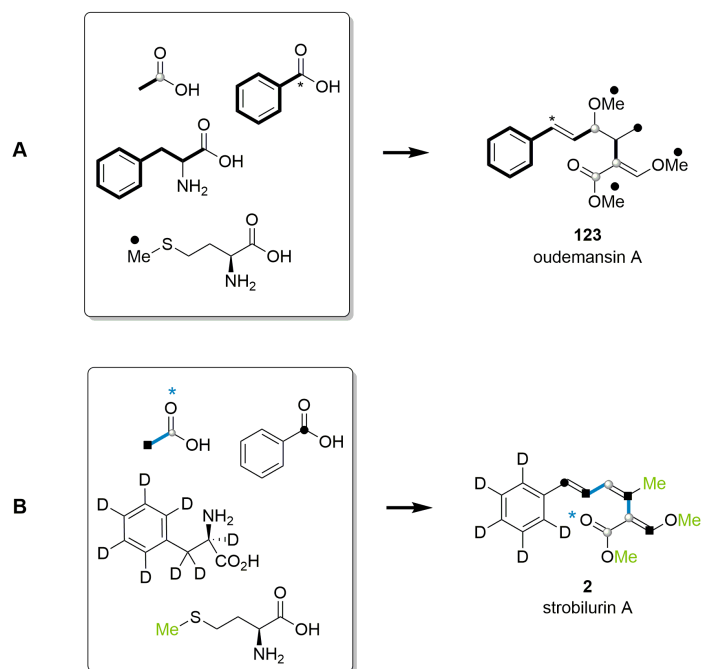
Figure 3.1 Structures of key natural and synthetic strobilurins.

3.1.1 Biosynthesis of Strobilurin A - *status quo*

Feeding studies with labelled precursors and proposed biosynthesis

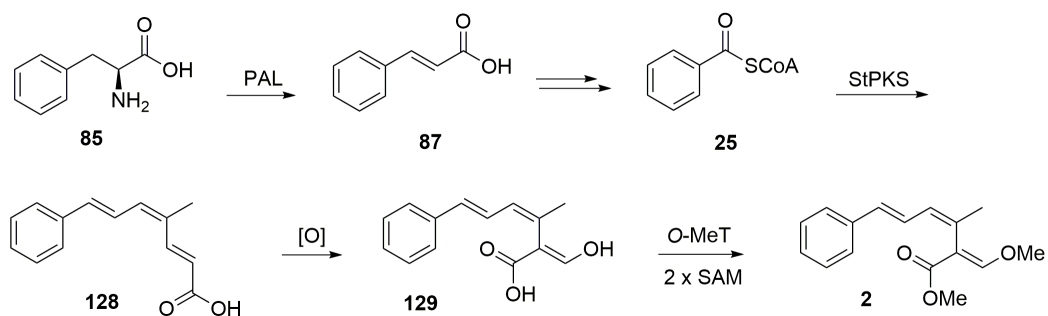
In the 1990s feeding experiments using isotopic labelled precursors showed that oudemansin A **123** is a polyketide using benzoate as starter unit. The benzoate itself is derived from degradation of phenylalanine **85** *via* cinnamate **87**. The polyketide chain is extended three times and C-methylated once. For generation of the key β -methoxyacrylate structure an unusual oxidative rearrangement must occur and further SAM-dependent O-methylations led to oudemansin A **123** (Scheme 3.1 A).⁴⁸

Further investigation of the biosynthetic origin of **2** was carried out by our group. Feeding singly and doubly labelled [¹³C]-acetates, [methyl-¹³C]-methionine, [²H₈]-, [3-¹³C]- and [2,3-¹³C₂]-phenyl-alanines to the **2** producing fungus *S. tenacellus* showed that incorporation of precursors was consistent with earlier feeding studies with **123** (Fig. 3.1 B).¹⁸⁴



Scheme 3.1 Incorporation of biosynthetic precursors: **A**, incorporation of ^2H -, ^{13}C - and ^{14}C -labelled precursors into oudemansin A **123**; **B**, incorporation of ^2H - and ^{13}C -labelled precursors into strobilurin A **2**.

Further investigations with advanced intermediates confirmed that **2** is produced *via* a linear tetraketide, called prestrobilurin A **128**. The latter is proposed to be the substrate for an oxidative rearrangement to form the core β -methoxyacrylate moiety and there is evidence for the intermediacy of an epoxide in the key rearrangement (Scheme 3.2).¹⁸⁴



Scheme 3.2 Proposed biosynthetic steps to strobilurin A **2**.

Gene cluster of strobilurin A

In previous studies in our group the producing strain of **2**, *S. tenacellus*, was sequenced and bioinformatic as well as transcriptomic analyses revealed a 50 kb BGC for strobilurin A production. The cluster contains a gene which encodes a highly reducing PKS (*stpks1*), genes predicted to encode pro-

teins involved in phenylalanine metabolism (such as a PAL) and further numerous genes encoding tailoring proteins involved in methylation and redox reactions (Fig. 3.2 A).^{175,185}

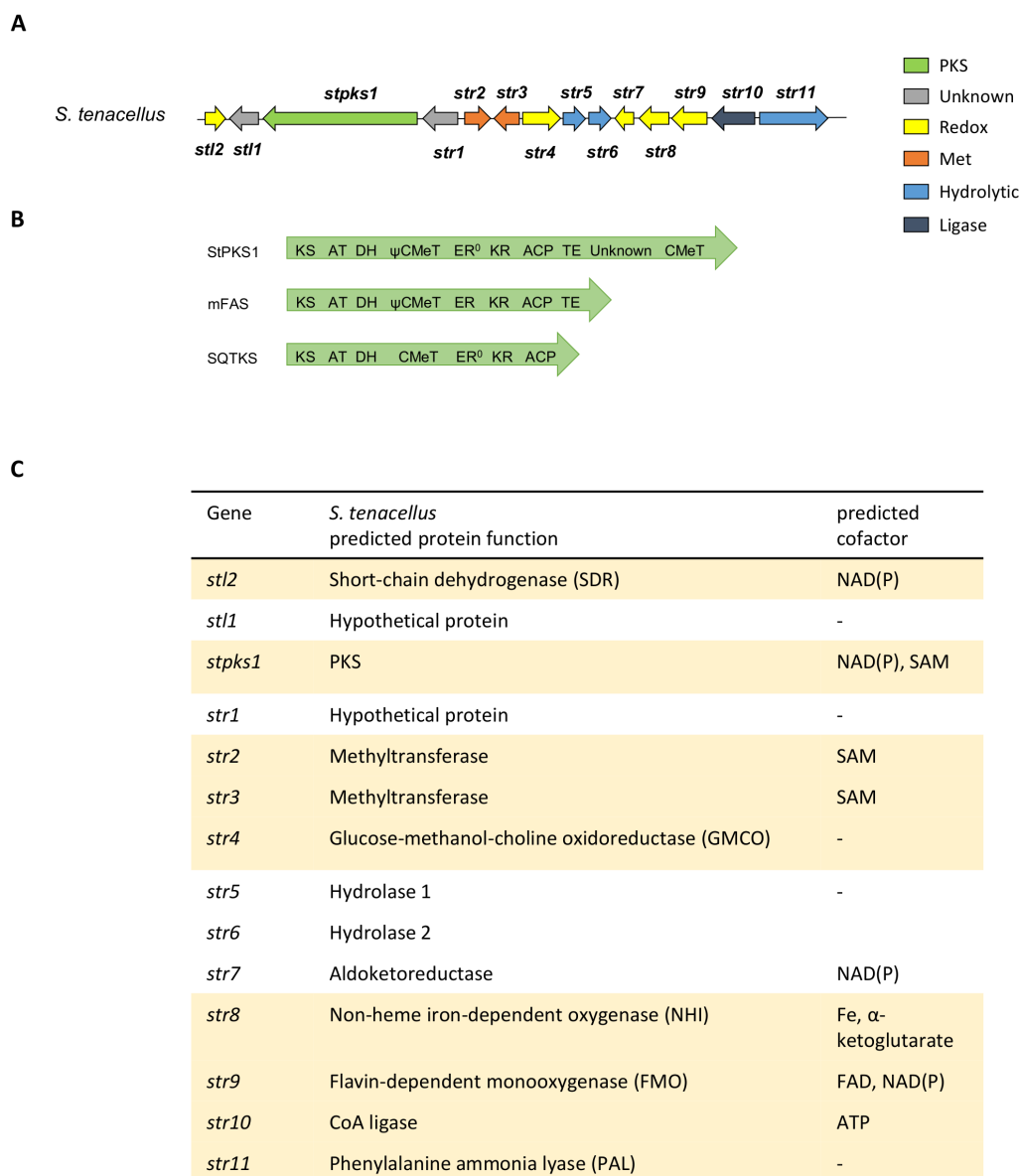


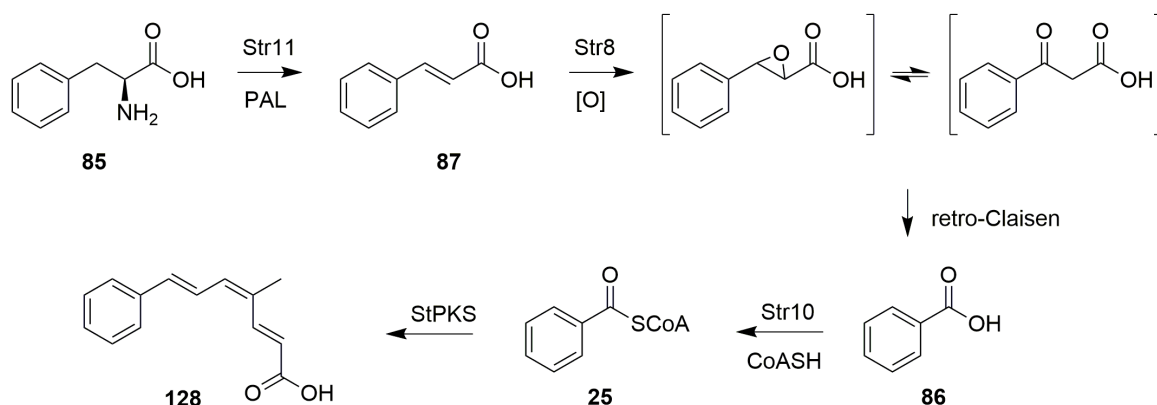
Figure 3.2 **A**, organisation of genes in 2 BGC; **B** detailed analysis of StPKS domain architecture; **C**, overview of predicted protein functions of encoded genes in 2 BGC, yellow shaded rows indicate genes shown to be involved in the biosynthesis of strobilurin A.

It is known that resistance to strobilurins is induced by mutations in the mitochondrial cytochrome b, encoded by *CYTB*.^{186,187} Nevertheless no homologues of *CYTB* were found in the 2 BGC or nearby, but the mitochondrial *CYTB* of *S. tenacellus* shows characteristic motifs of being strobilurin resistant.^{175,185}

Closer observations of StPKS revealed an unusual domain architecture. It was shown to contain domains usually observed in hrPKS, such as KS, AT, DH, a probably inactive C-MeT, ER, KR and ACP, but it also possesses three unusual domains in the C-terminal part of the enzyme. The ACP is followed by a partial $\alpha\beta$ -hydrolase which is homologous to the mammalian FAS thiolesterase (33% identical), an unknown domain and a second C-MeT domain (Fig. 3.2 B). In contrast to the probably inactive C-MeT domain, located in between the DH and KR domains, the C-terminal C-MeT domain is likely to be active, as the SAM-binding site was shown to be intact. Moreover, the ER domain is also likely to be inactive due to mutations in the active site, compared to SQTKS.^{175,185}

3.1.2 Prestrobilurin production and benzoyl CoA biosynthesis in *S. tenacellus*

Initial studies to elucidate the biosynthesis by *in vivo* heterologous expression experiments were conducted by Risa Nofiani (Cox group, University of Bristol). As targeted gene knockouts in *S. tenacellus* and *S. lutea* were not successful, key genes from the cluster were expressed in *A. oryzae*. The 8.5 kb intron-free PKS was reconstructed using yeast homologous recombination and cloned into the fungal expression vector pTAGS by Risa Nofiani.^{175,185} Expression of *stpks1* in *Aspergillus oryzae* did not lead to the production of new compounds. But when the fermentation was supplemented with benzoyl SNAC **121**, a mimic of the putative starter unit benzoyl CoA **25** (Section 1.1.1), production of a strobilurin precursor prestrobilurin A **128** was detected.^{175,185} Even higher yields of **128** production were achieved when benzoyl-CoA producing genes (*str11*, *str8*, *str10*) from the BGC were coexpressed in *A. oryzae*. The PAL (encoded by *str11*) is able to convert endogenous phenylalanine **85** to cinnamate **87** which undergoes oxygenation by a non-heme iron-dependent oxygenase (encoded by *str8*) and then retro-Claisen reaction to form benzoic acid **86**. This is then activated to benzoyl CoA **25** by a CoA ligase (encoded by *str10*). Benzoyl CoA **25** is the substrate for StPKS and yields prestrobilurin A **128** (Scheme 3.3).^{175,185}



Scheme 3.3 Benzoyl CoA biosynthetic pathway in *S. tenacellus* elucidated by heterologous expression experiments in *A. oryzae* NSAR1.^{175,185}

In general, benzoate is a rare starter unit in fungal polyketide biosynthesis (Chapter 1.1.1). However, recent results of our group and the Tang group have shown that benzoyl CoA **25** also forms the starter unit for SQHKS in squalastatin biosynthesis (Section 2.1).^{53,118} Both fungal pathways share a common PAL enzyme but the SQS pathway involves formation of cinnamoyl CoA **88** early in the pathway, whereas strobilurin pathway involves direct oxidative conversion of cinnamic acid **87** to benzoic acid **86**.^{175,185}

The expression experiments conducted by Risa Nofiani showed that the first enzyme free intermediate of StPKS is prestrobilurin A **128**.^{175,185} It is possible that the TE domain downstream of the ACP domain in StPKS is involved in chain release and the C-terminal C-MeT could be responsible for attachment of the C-4 methyl group, as the more common upstream C-MeT domain seemed to be inactive due to mutations in the active site.^{175,185}

The analysis of genomic and transcriptomic data of genes clustered with *stpks1* revealed multiple putative gene candidates involved in catalysing the oxidative step, including *str4* (GMC oxidoreductase), *str7* (aldoketoreductase), *str9* (FAD dependent monooxygenase) and *stl2* (alcohol dehydrogenase). The two methyltransferase encoding genes *str2* and *str3* were likely involved in O-methylation to yield strobilurin.¹⁸⁵ As four possible candidate genes for oxidations were identified from the cluster, initially all putative genes which were predicted to be part of **2** biosynthesis (*stpks1+stl2+str2+str3+str4+str7+str8+str9+str10+str11*) were coexpressed in *A. oryzae* (Fig. 3.3). Apart from prestrobilurin A no new related product could be observed.¹⁸⁵ Later genetic analysis revealed that *str4* has not been incorporated into the genome of the heterologous host. Furthermore, updated transcriptomic data revealed a mis-prediction for the start and stop codon of *str9*.¹⁸⁵ Dr Kate de Mattos-Shiple (Cox group, Bristol University) repaired the gene *str9* and cloned it together with *stl2* into expression vector pTYGSmet.

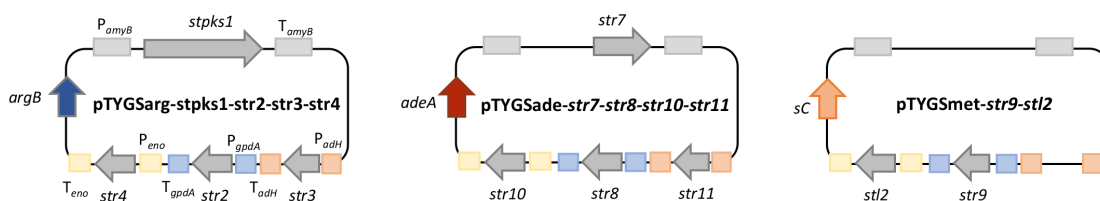


Figure 3.3 Plasmid maps of genes from **2** BGC (*stpks1+stl2+str2+str3+str4+str7+str8+str9+str10+str11*) coexpressed in initial heterologous expression experiments in *A. oryzae* NSAR1 conducted by Dr Risa Nofiani.¹⁸⁵

The characteristic *E,Z,E* triene in all strobilurins is highly unusual and the origin of this motif has been hitherto unexplained. In polyketide biosynthesis *Z*-olefins are rare. Nevertheless there are examples known, such as in borrelidin **130** PKS, where a DH domain creates a typical *E*-olefin, which is later isomerised (Fig. 3.4).¹⁸⁸ Other examples were described in bacterial *trans*-AT PKS (bongkrelic acid **131**), where the KR sets up a β -alcohol *anti* to an α -proton and subsequent *syn*

dehydration catalysed by the DH domain results in the *Z*-olefin directly (Fig. 3.4).¹⁸⁹ It is possible that in strobilurin biosynthesis the KR/DH domain is also able to control different stereoselectivity during the three iterations perhaps in response to methylation, or the unknown domain of the StPKS has a role in formation of the *Z*-olefin.¹⁷⁵

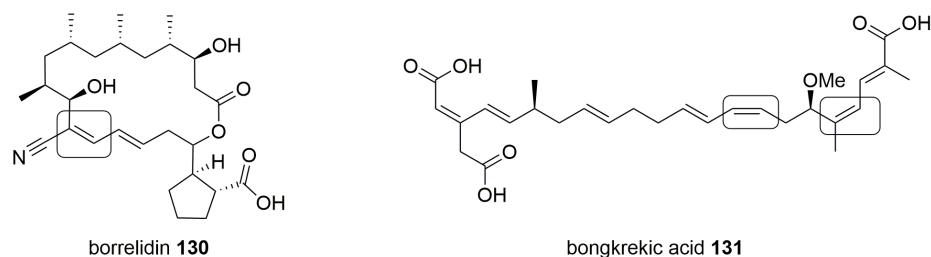


Figure 3.4 Two examples of *Z*-olefins in polyketide biosynthesis: borrelidin **130** and bongkreic acid **131**.

3.2 Project aims

Prestrobilurin A **128** was postulated to be the substrate which undergoes oxidative rearrangement to form the core β -methoxyacrylate moiety, but until now there is no direct evidence for this step (Scheme 3.2).

Earlier isotopic feeding studies proposed an oxidative rearrangement possibly involving an epoxide intermediate and two further *O*-methylation steps to yield **2**.^{48,184} Initial attempts to investigate the biosynthetic steps from prestrobilurin A **128** to strobilurin A **2** (referred to as later steps) by heterologous expression of genes from **2** BGC in *A. oryzae* NSAR1, conducted by Risa Nofiani, have not been successful yet.¹⁸⁵

The aim of this chapter was to explore the later steps of strobilurin A biosynthesis particularly the unusual oxidative rearrangement to form the core β -methoxyacrylate moiety. Therefore, all putative genes involved in **2** biosynthesis (including repaired *str9*) should be coexpressed with prestrobilurin producing genes in the heterologous host *A. oryzae* NSAR1 in order to re-establish strobilurin A production.

In addition the number of genes coexpressed should be narrowed down to a minimal gene cluster for **2** biosynthetic pathway. The gene involved in the oxidative rearrangement should be identified and investigated further by heterologous expression in *E. coli* and subsequent *in vitro* studies with the substrate prestrobilurin **128** should be conducted to study this step in detail.

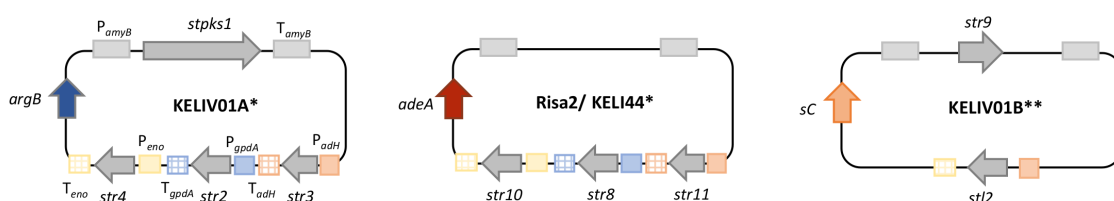
In vivo characterisation of the *O*-MeT-encoding genes *str2* and *str3* by heterologous expression in *A. oryzae* NSAR1 should reveal the order of methylation steps in strobilurin and bolineol biosynthesis. Using the same methodology should reveal the role of short-chain dehydrogenase (SDR) StI2 and glucose-methane-choline oxidase (GMCO) Str4 in the biosynthesis of strobilurin and bolineol.

3.3 Results – Heterologous expression of strobilurin A BGC in *A. oryzae* NSAR1

3.3.1 Vector construction and gene combinations for *A. oryzae* transformation

In order to investigate the later steps of strobilurin biosynthesis, the provided vectors from Bristol (KELIV01A, KELI44, KELIV01B) were used. Together these contain all genes predicted to be essential for strobilurin A production (Fig. 3.5). For testing different set of genes and defining a minimal gene set for production of **2**, five additional fungal expression vectors (KELIV07, KELIV04, KELIV11, KELIV12A, KELIV12B) were constructed using yeast homologous recombination, described earlier in section 2.6.1 (Fig. 3.5). Correct cloning was confirmed by partial sequencing (Eurofins, Ebersberg) and details for the used cloning strategy and corresponding oligonucleotides are summarised in chapter 6, table 6.10.

Provided vectors by Bristol group



* provided by Dr Risa Nofiani, ** provided by Dr Kate de Mattos-Shiplee

Vectors constructed in this work

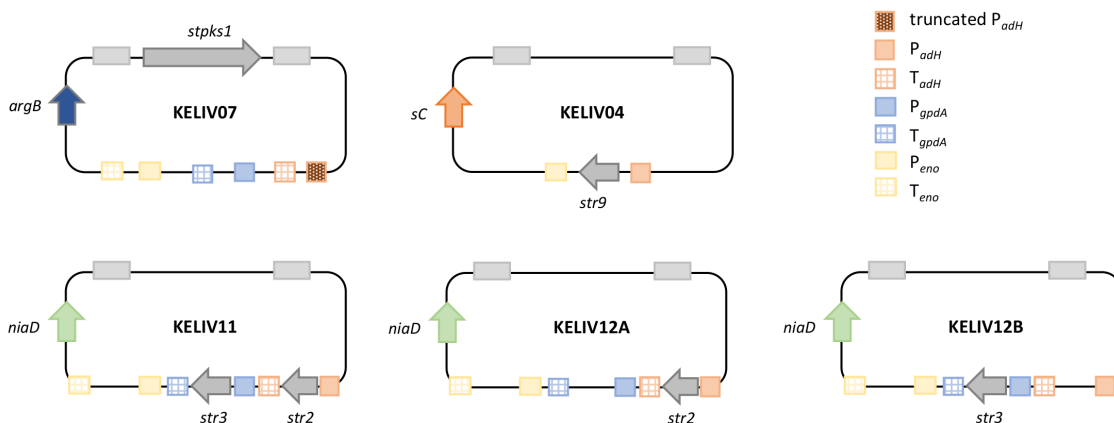


Figure 3.5 Overview of kindly provided fungal expression vectors from Bristol by Dr Risa Nofiani (*) and Dr Kate de Mattos-Shiplee (**) (both: Cox group, University of Bristol) and constructed vectors in this work used for fungal expression (pTYGSarg, met, ade, niaD).

In order to transform *A. oryzae* NSAR1 with combinations of constructed vectors (Fig. 3.5) a CaCl_2 /PEG mediated protoplast protocol was used, which has already been described in section 2.6.1.

Cotransformation of two, three or four vectors were carried out in a single transformation step using selection medium which was adapted to the vectors used in the individual transformation. Overall 10 combinations of genes from strobilurin BGC on different fungal expression vectors were used to cotransform *A. oryzae* NSAR1, resulting in 10 strains occupying genes from BGC of 2 (Fig. 3.6). Genetic analysis showed the correct insertion of desired genes into *A. oryzae* (Appendix, Fig. 7.4).

Summary of heterologous expression experiments in *A. oryzae* designed to probe the biosynthesis of 2

Exp. ID	Selection marker	Transformed expression vector(s)	Genes with protein functions										Feed 121	Exp. No.
			<i>stpks1</i>	<i>str11</i>	<i>str8</i>	<i>str10</i>	<i>str9</i>	<i>str4</i>	<i>stl2</i>	<i>str2</i>	<i>str3</i>			
			PKS	PAL	NHI	CoA ligase	FMO	GMCO	SDR	OMT1	OMT2			
KELIV02	arg, ade, met	KELIV01A, KELI44, KELIV01B	X	X	X	X	X	X	X	X	X	X		1
KELIV04	arg, ade, met	KELIV01A, KELI44, KELIV04	X	X	X	X	X	X		X	X		2	
KELIV05 feed	arg	KELIV01A	X					X		X	X	X	3	
KELIV04_4	arg, ade	KELIV01A, KELI44	X	X	X	X		X		X	X		4	
KELIV07	arg, ade, met	KELIV07, KELI44, KELIV01B	X	X	X	X	X						5	
KELIV08	arg, ade, met	KELIV07, KELI44, KELIV04	X	X	X	X	X		X				6	
KELIV11A	arg, ade, met, niaD	KELIV07, KELI44, KELIV04, KELI11	X	X	X	X	X			X	X		7	
KELIV11B	arg, ade, met, niaD	KELIV07, KELI44, KELIV01B, KELI11	X	X	X	X	X		X	X	X		8	
KELIV12A	arg, ade, met, niaD	KELIV07, KELI44, KELIV01B, KELI12A	X	X	X	X	X			X			9	
KELIV12B	arg, ade, met, niaD	KELIV07, KELI44, KELIV04, KELI12B	X	X	X	X	X				X		10	

Figure 3.6 Overview of *A. oryzae* NSAR1 transformations with different combinations of genes from BGC of 2.

3.3.2 Heterologous expression of later pathway genes

From previous feeding studies, *in silico* analysis of the gene cluster and initial expression experiments leading to prestrobilurin A production (Section 3.1.1), it was reasoned that the next acting gene must be involved in the oxidative rearrangement. This may be achieved by either the GMCO encoded by *str4* or the flavin-dependent monooxygenase (FMO) encoded by *str9*. Further *str2* and *str3* (two SAM-dependent methyltransferases) were included in the first expression experiment containing the full complement of genes, as they are probably responsible for the addition of the two *O*-methyl groups of 2. Finally, the gene *stl2* encoding a SDR is also coexpressed as production of bolineol 4 requires a reduction step.

In an initial heterologous expression experiment (exp. 1) the full complement of genes, including the benzoyl CoA-production system (*str8*, *str10* and *str11*), were coexpressed in *A. oryzae* NSAR1 (Table 3.6, exp. 1). LCMS analysis showed two main products. These were shown to be prestrobilurin A 128 ($t_R = 7.6$ min) and a peak at $t_R = 8.4$ min corresponded to strobilurin A 2 production (Fig. 3.7 a).

3.3 Results – Heterologous expression of strobilurin A BGC in *A. oryzae* NSAR1

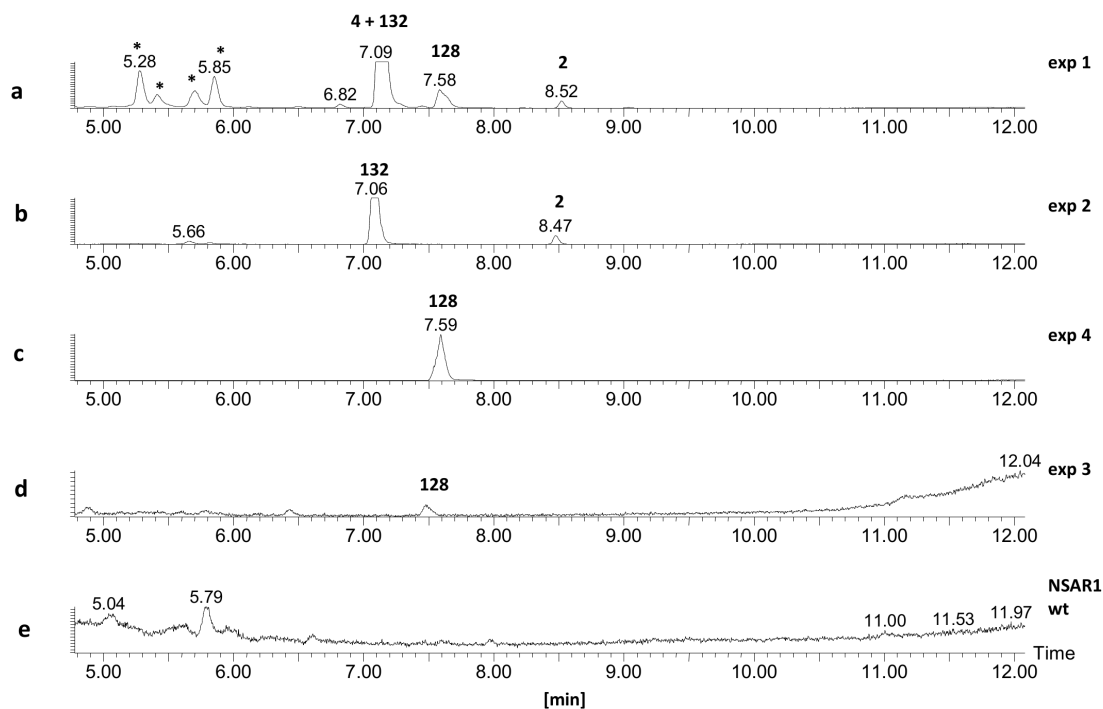


Figure 3.7 LCMS evaporative light-scattering (ELS) chromatograms of organic extracts of *A. oryzae* NSAR1 expression strains (arbitrary units). **a**, *A. oryzae* NSAR1 + *stpks1* + *str2* + *str3* + *str4* + *str11* + *str8* + *str10* + *str9* + *stl2*; **b**, *A. oryzae* NSAR1 + *stpks1* + *str2* + *str3* + *str4* + *str11* + *str8* + *str10* + *str9*; **c**, *A. oryzae* NSAR1 + *stpks1* + *str2* + *str3* + *str4* + *str11* + *str8* + *str10*; **d**, *A. oryzae* NSAR1 + *stpks1* + *str2* + *str3* + *str4* + benzoyl SNAC **121**; **e** untransformed *A. oryzae* NSAR1. *Unrelated compounds.

The mass and UV spectrum as well as the HRMS (ESI⁺, C₁₆H₁₈O₃Na calc. [M+Na]⁺ 281.1154, measured 281.1157) was characteristic for **2** (Fig. 3.8). Further isolation of 2.6 mg of **2** from one litre of culture medium and subsequent NMR analysis proved it to be **2** by comparison to literature data (Chapter 6, Table 6.14).⁴⁸

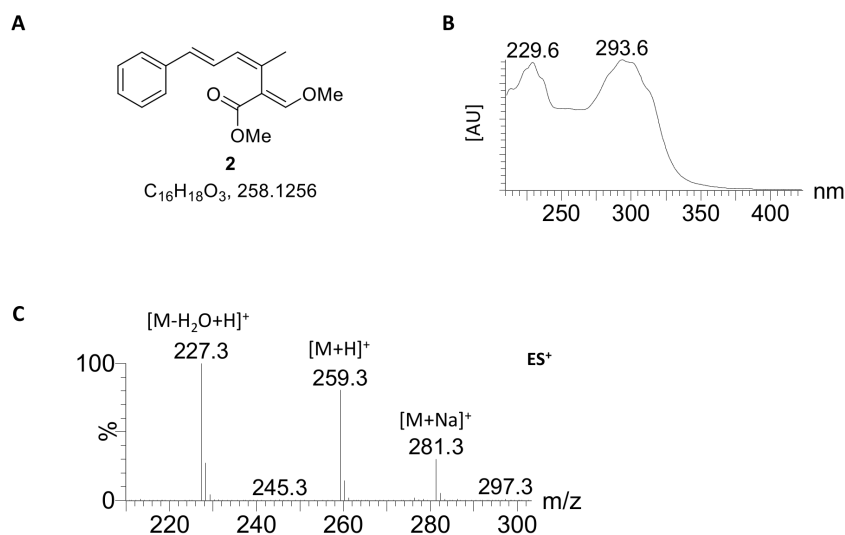
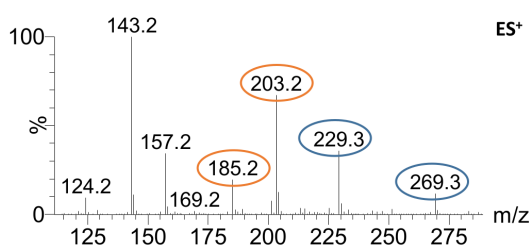


Figure 3.8 A, structure of **2**; B, UV chromatogram; C, ES⁺ spectrum.

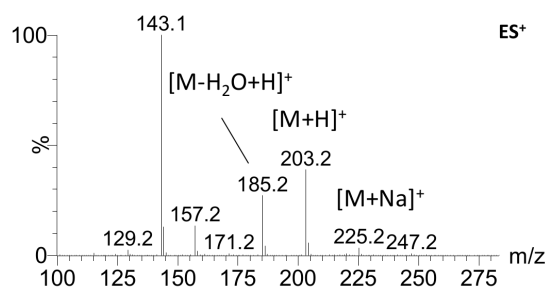
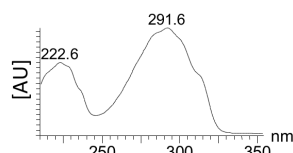
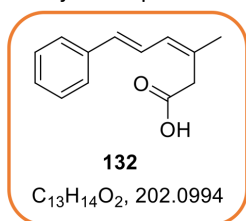
The major product produced in exp. 1 was found to be a new compound ($t_R = 7.1$ min, $C_{13}H_{15}O_2$ $[M+H]^+$ calc. 203.1072, measured 203.1071, Fig. 3.7 a). It was produced in high titres (90 mg/L culture) and was isolated, submitted to NMR analysis and shown to be the *5E,3Z* carboxymethyldiene **132**, according to coupling constants of $J = 15.5$ Hz (H-5/H-6) and $J = 10.8$ Hz (H-4), respectively. Compound **132** has been previously synthesised during the total synthesis of strobilurin A **2** (Fig. 3.9).¹⁹⁰ Corresponding NMR assignment is displayed in chapter 6, section 6.5.6. A second, minor compound was detected to coelute with **132**. It was shown to be bolineol **4**, because of its distinctive $[M-H_2O+H]^+$ (229.3) and $[M+Na]^+$ (269.3) ions (comparison to a standard of bolineol, Fig. 3.9).

A exp 1, $t_R = 7.1$ min



B

major compound:



C

minor compound:

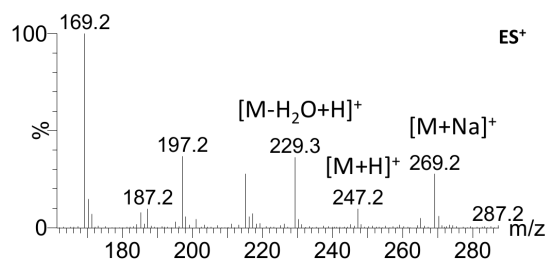
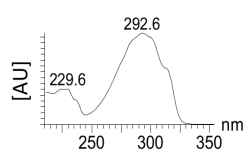
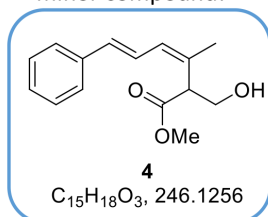


Figure 3.9 LCMS analysis of exp. 1 (Fig. 3.6): **A**, ES^+ spectrum of exp. 1 at $t_R = 7.1$ min corresponding to **132** and **4**; **B**, analytical data for standard of **132**; **C**, analytical data for bolineol **4**.

In a second experiment (Fig. 3.6, exp. 2) the SDR encoding gene *stl2* was omitted from the system by using methionine vector KELIV04 for coexpression in *A. oryzae* NSAR1, lacking *stl2*. LCMS analysis revealed a very similar result, with production of **2** and major compound **132**, but lacking bolineol **4** (Fig. 3.7 b, exp. 2). Omission of the FMO encoded by *str9* (Fig. 3.6, exp. 4) led to the production of only prestrobilurin A **128** (Fig. 3.7 c, exp. 4). The same result was achieved if the benzoate genes

(*str11*, *str8*, *str10*) were neglected from expression system and the fermentation was supplemented with benzoyl SNAC **121** (Fig. 3.6, exp. 3), although **128** is produced in much lower titre in this case (Fig. 3.7 d, exp. 3).

3.3.3 Coexpression of *stPKS*, benzoyl–CoA producing genes and FAD monooxygenase encoding gene *str9*

Coexpression of a minimal set of genes (*stpk1+str10+str8+str11+str9*) without GMCO- and O–MeT-encoding genes (Fig. 3.6, exp. 5) led to the formation of major compound **132**, and also a new compound **133**, eluting at $t_R = 5.9$ min, 30 mg/L culture (Fig. 3.10 B, exp. 5). The new compound was shown to have characteristic absorption maxima (229.6, 294.6 nm) compared to **2** and HRMS analysis revealed chemical formula $C_{14}H_{16}O_3$ ($[M-H]^-$ calc. 231.1021, measured 231.1021, Fig. 3.10 C).

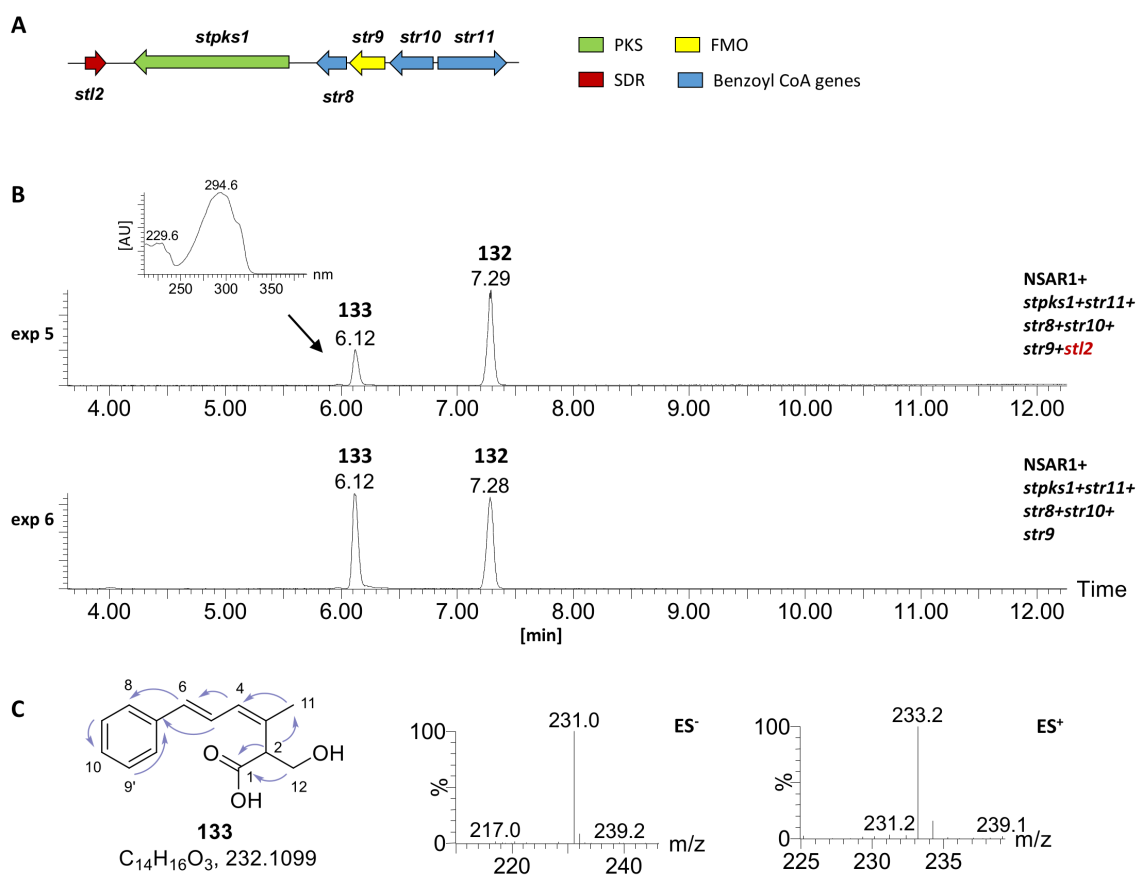


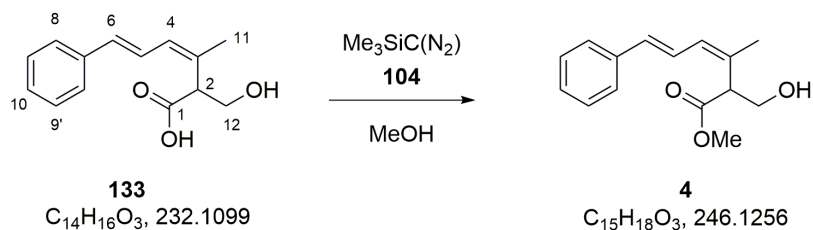
Figure 3.10 A, LCMS evaporative light-scattering (ELS) chromatograms of organic extracts of *A. oryzae* NSAR1 expression strains (arbitrary units): exp. 5, *A. oryzae* NSAR1 + *stpk1* + *str11* + *str8* + *str10* + *str9* + *stl2* with corresponding UV spectrum of compound **133**; exp. 6, *A. oryzae* NSAR1 + *stpk1* + *str11* + *str8* + *str10* + *str9*; B, structure of compound **133** with key HMBC correlations, corresponding NMR assignment is shown in table 3.1, and mass spectra (ES^+ and ES^-).

Isolation and full structure determination by NMR confirmed the new compound **133** to be desmethylbolineol (Table 3.1, Fig. 3.11).

Table 3.1 Chemical shifts of 3.8 mg **133** in CDCl₃ (400 MHz)

Atom	δ_C / ppm	δ_H / ppm (J / Hz)	COSY	HMBC
1	177.9	q	-	2, 12
2	49.9	4.08 (m, 1H)	12	1, 3, 4, 11, 12
3	130.9	q	-	11
4	131.1	6.24 (d, $J = 11.4$, 1H)	5, 11	2, 6, 11
5	123.6	7.02 (dd, $J = 15.3$, 1H)	4, 6	4, 7
6	133.5	6.56 (d, $J = 15.3$, 1H)	5	4, 5, 8
7	137.3	q	-	5, 6, 8
8/ 8'	128.8	7.31 (m, 2H)	9	9, 10
9/ 9'	126.6	7.41 (m, 2H)	8, 10	8, 10
10	127.9	7.20 (m, 1H)	9	8, 9
11	21.3	1.86 (m, 3H)	4	2, 4
12	61.9	3.71 (dd, 1H), 4.08 (m, 1H)	2	1

Desmethylbolineol **133** was smoothly converted to bolineol **4** by treatment with trimethylsilyldiazomethane **104** (Scheme 3.4) and the spectroscopic data were shown to be identical to literature data of bolineol (NMR data in chapter 6, section 6.5.6).¹⁸²



Scheme 3.4 Treatment of compound **133** with trimethylsilyldiazomethane **104** yielded bolineol **4**.

Omission of the SDR encoding gene *stl2* (Fig. 3.6, exp. 6) gave the same result, showing that formation of **133** does not require the SDR and must represent a shunt pathway in *A. oryzae* in the absence of the late-acting methyltransferases Str2 and Str3 (Fig. 3.10, exp. 6). Compound **133** showed the rearranged structure, but in comparison to strobilurin A lacks the O-methyl groups. In addition, NMR data showed that the C-2/C-12 olefin of **133** and the second carboxyl group are reduced indicated by the diastereomeric protons at C-12 (3.71, 4.08 ppm) in proximity to the alcohol group at C-12 (61.9 ppm, Table 3.1, Fig. 3.11).

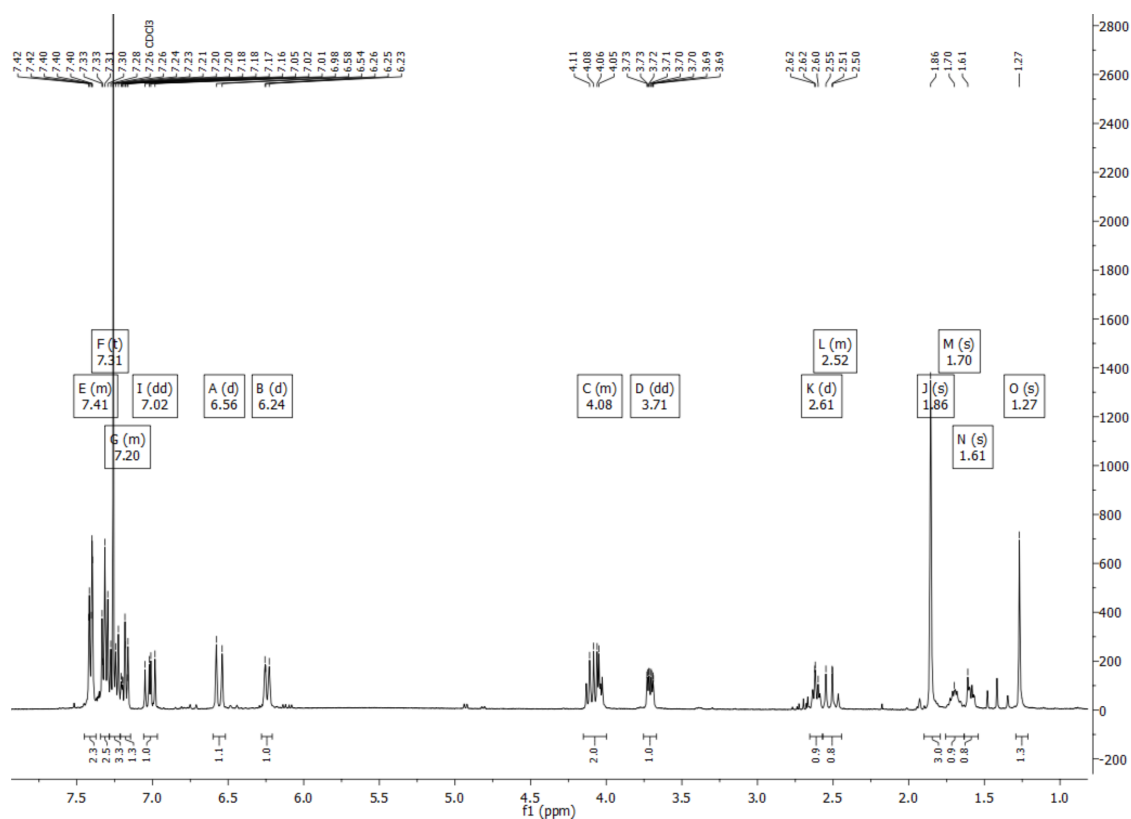


Figure 3.11 ^1H NMR data of compound 133 in CDCl_3 (400 MHz) referenced to CDCl_3 ; see table 3.1 for assignment.

These results show that FAD dependent oxidase Str9 is responsible for the unusual rearrangement to form the toxophore of **2**, as the second putative oxidation gene *str4* (GMCO) was omitted in these experiments.

3.3.4 Defining the minimal BGC for strobilurin A production

Adding the two SAM dependent *O*-methyltransferases (*O*-MeTs) encoding genes *str2* and *str3* from **2** BGC to the previous expression system in *A. oryzae* NSAR1 (*stpks1* + *str11* + *str8* + *str10* + *str9*, Fig. 3.6, exp. 7) resulted in extremely unhealthy *A. oryzae* cultures, which were difficult to select and grow (Fig. 3.13 a). Extraction of these cultures revealed the formation of bolineol **4** as the major compound, together with lesser amounts of **132** and **133**. Strobilurin A **2** was detected by its distinctive mass spectrum, but in very low titre (Fig. 3.12 B, exp. 7).

3.3 Results – Heterologous expression of strobilurin A BGC in *A. oryzae* NSAR1

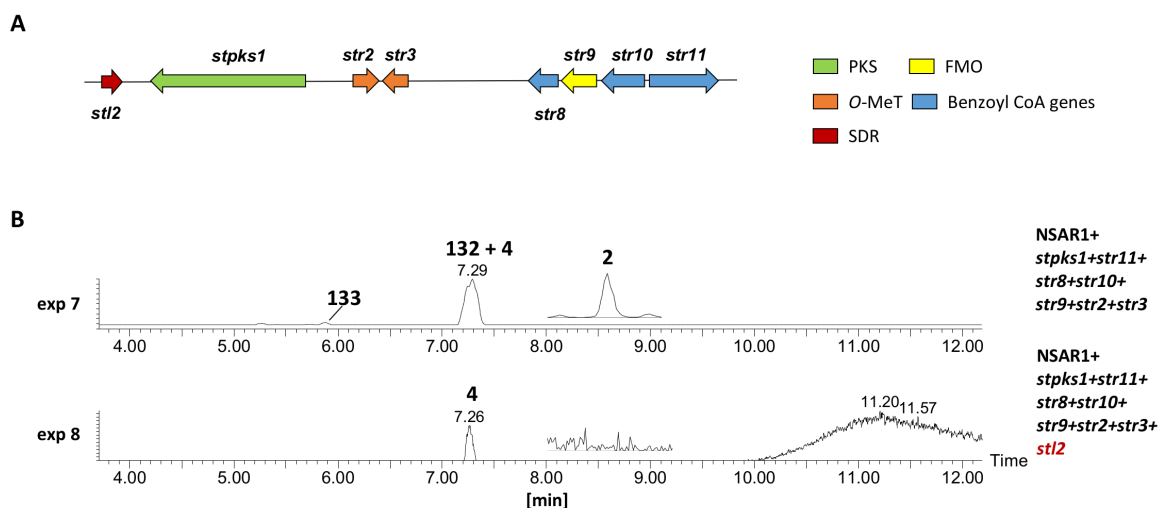


Figure 3.12 A, coexpression of genes from 2 BGC; B, LCMS evaporative light-scattering (ELS) chromatograms of organic extracts of *A. oryzae* NSAR1 expression strains (arbitrary units): **exp. 7**, *A. oryzae* NSAR1 + *stpks1* + *str11* + *str8* + *str10* + *str9* + *str2* + *str3*; **exp. 8**, *A. oryzae* NSAR1 *stpks1* + *str11* + *str8* + *str10* + *str9* + *str2* + *str3* + *stl2*, inset traces shows extracted ion chromatogram for *m/z* 259.1 corresponding to 2.

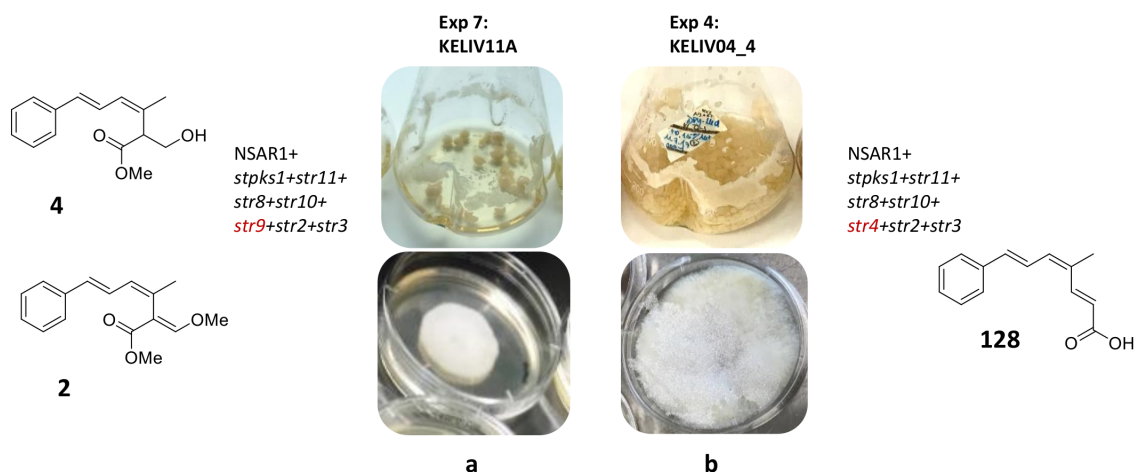


Figure 3.13 Pictures of two different *A. oryzae* NSAR1 expression strains (7 days in DPY medium): **a**, sick phenotype of exp. 7 (Fig. 3.6, exp. 7, KELIV11A) NSAR1+*stpks1*+*str11*+*str8*+*str10*+*str9*+*str2*+*str3* producing 2 and 4; in comparison with **b**, healthy phenotype of exp. 4 (Fig. 3.6, exp. 4, KELIV04_4) NSAR1+*stpks1*+*str11*+*str8*+*str10*+*str4*+*str2*+*str3* producing 128.

Production of strobilurin A 2 and also bolineol 4, showed that the *O*-MeTs Str2 and Str3 are responsible for the addition of the two *O*-methyl groups. Further, these result showed that expression of *stpks1* (PKS) and *str11* + *str8* + *str10* (benzoyl-CoA production) + *str9* (FMO) and *str2* + *str3* (*O*-MeTs) is sufficient to produce strobilurin A in heterologous host *A. oryzae* NSAR1 and therefore demonstrate the minimal BGC for production of 2.

Reinclusion of the SDR Stl2 (Fig. 3.6, exp. 8) produced 4 only, with no production of 2 (Fig. 3.12 B, exp. 8). These results also showed that without coexpression of GMCO-encoding gene *str4* only trace amounts of 2 were produced and that bolineol 4 is the major product.

3.3.5 Individual role of O-methyltransferases Str2 and Str3

In order to investigate the individual role of the two O-MeTs in strobilurin 2 and bolineol 4 biosynthesis, two expression vectors pTYGS-*niaD*-*str2* and pTYGS-*niaD*-*str3* were constructed (Fig. 3.5) and used together with the previous system (*arg*, *ade*, *met*) to cotransform *A. oryzae* NSAR1 yielding in 7 transformants for exp 9 and 6 transformants for exp 10 (Fig. 3.6, exp. 9+10).

Heterologous expression of O-MeTs Str2 and Str3 individually

In order to investigate the function of Str3 in strobilurin A 2 and bolineol 4 biosynthesis, *str2* was omitted from the expression system, by expression of the minimal BGC for 2 production lacking *str2* (Fig. 3.6, exp. 9). This resulted in production of 132 and 133 (Fig. 3.14, exp. 9). The same panel of compounds was already found in previous expression experiments, where no O-MeT genes were coexpressed (Fig. 3.10, exp. 5+6). All compounds were identified by comparing retention times and mass fragmentation patterns to previously isolated standards. Detailed analysis of the mass spectrum (ES+ mode) of the peak eluting at 7.23 min showed that only the unmethylated compound 132 was detected by its distinctive $[M-H_2O+H]^+$ (185.2), $[M+H]^+$ (203.2) and $[M+Na]^+$ (225.1) ions (Fig. 3.15 a).

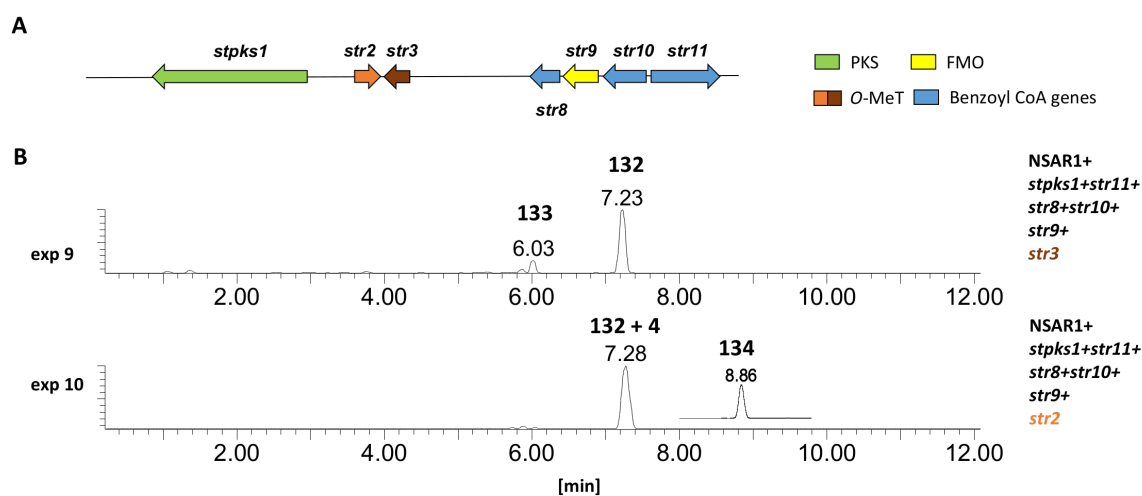


Figure 3.14 A, coexpression of genes from 2 BGC; B, LCMS evaporative light-scattering (ELS) chromatograms of organic extracts of *A. oryzae* NSAR1 expression strains (arbitrary units): **exp. 9**, *A. oryzae* NSAR1 + *stpks1* + *str11* + *str8* + *str10* + *str9* *str3*; **exp. 10**, *A. oryzae* NSAR1 *stpks1* + *str11* + *str8* + *str10* + *str9* + *str2*, inset trace shows DAD trace.

3.3 Results – Heterologous expression of strobilurin A BGC in *A. oryzae* NSAR1

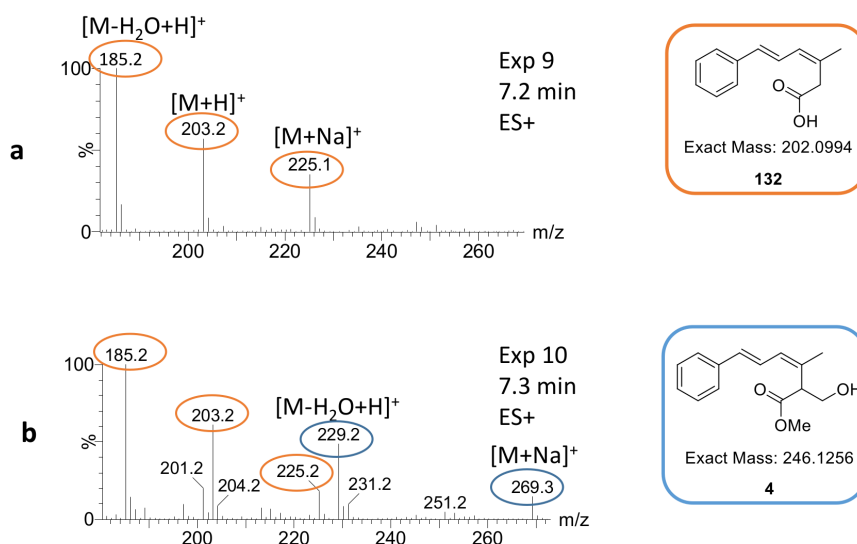


Figure 3.15 Mass spectra in ES⁺ mode: a, exp. 9 at 7.2 min; b, exp. 10 at 7.3 min.

Bolineol **4** or strobilurin A **2** were not produced. As both **133** and **132** are lacking *O*-methylations, we reasoned that either Str3 is not functional or, more likely, can only catalyse a second *O*-methylation step. Interestingly, some of the transformants showed a slightly healthier phenotype than others when growing on DPY agar. After LCMS analysis of the extracts, the healthier phenotype was shown to correlate to a higher production of **133** (Fig. 3.16 A). Growing *A. oryzae* NSAR1 on DPY agar plates containing various concentrations of **133**, revealed a MIC of 0.4 mg/ml for **133** against *A. oryzae* NSAR1 (Fig. 3.16 B).

In order to investigate the function of Str2 in strobilurin A **2** and bolineol **4** biosynthesis, *str3* was omitted from the expression system, by expression of the minimal BGC for **2** production lacking *str3* (Fig. 3.6, exp. 10). This led to one major peak in the ELSD traces of the extracts of transformants (Fig. 3.14, exp. 10). According to the retention time ($t_R = 7.3$ min) and the mass spectrum, the major compounds produced corresponded to **132** and **4**. Bolineol **4** was detected to coelute with **132** by its distinctive $[M-H_2O+H]^+$ (229.2) and $[M+Na]^+$ (269.3) ions (Fig. 3.15 b).

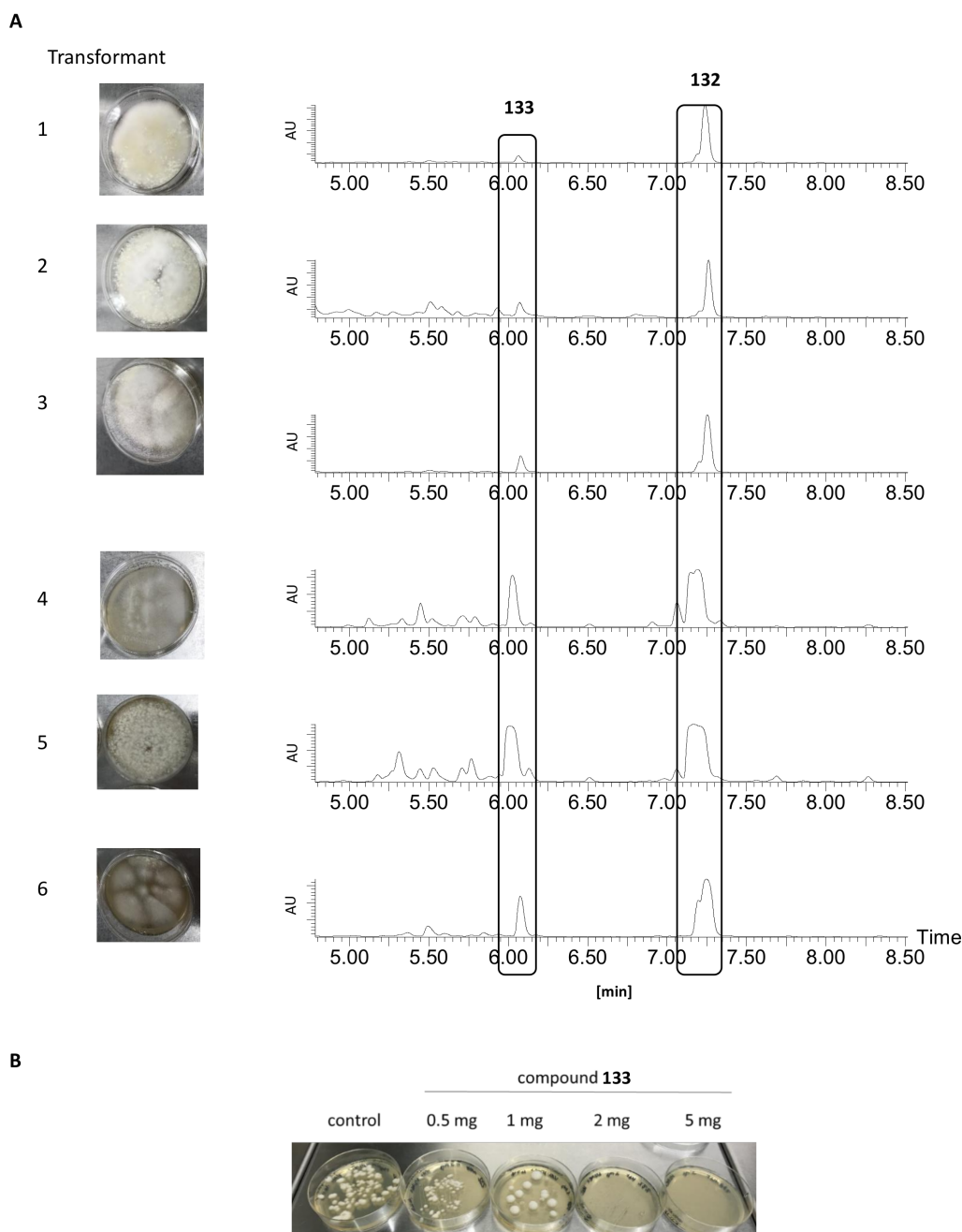


Figure 3.16 A, phenotype and LCMS (DAD) traces of NSAR1 *stPKS1 + str11 + str8 + str10 + str9 + str3* transformants 1–6; B, toxicity test with compound **133** against *A. oryzae* NSAR1; DPY agar (5ml) was supplemented with various concentrations (0.5, 1, 2, 5 mg) of compound **133** diluted in 200 μ l DMSO, control only 200 μ l DMSO.

A new minor compound **134** eluting at 8.9 min was also detected in the DAD trace of the producing fungus (Fig. 3.14, exp. 10, inset trace). The amount produced was too little to purify the compound and to elucidate the structure by NMR. Nevertheless, the mass spectrum correlates to an *O*-methylated **132**, matching its ions: $[M+H]^+$ (217.2) and $[M+Na]^+$ (239.2) ions (Fig. 3.17 A). Further support for this structure is given by its characteristic UV spectrum (maxima at 212, 292 nm), which matches UV spectrum for compound **132** (Fig. 3.17 B). In comparison to production of **133** in

previous experiment (exp. 9), in this study with *str2* being expressed, compound **133** could not be detected, and neither could production of **2** be observed.

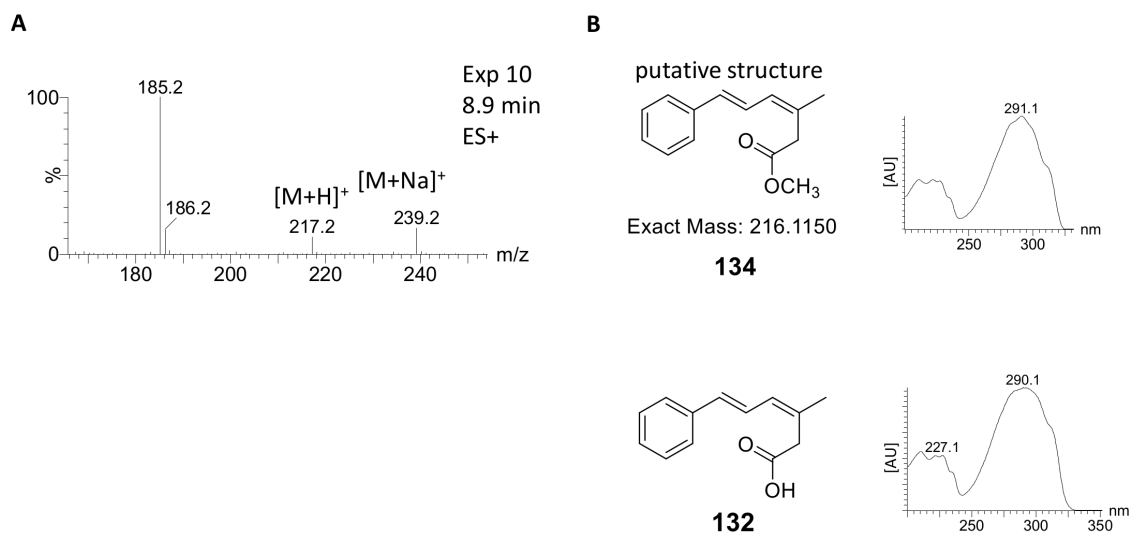


Figure 3.17 Analytical data of compound **134**: **A**, mass spectra in ES⁺ mode of exp 10 at 8.9 min; **B**, UV spectra of compound **134** in comparison with compound **132**.

Toxicity of bolineol

Overexpression of *str2* in the absence of *str3* resulted in extremely unhealthy *A. oryzae* cultures on DPY agar. This was reasoned to be caused by bolineol **4** production. In order to test this hypothesis, a toxicity test with **4** towards untransformed *A. oryzae* was conducted. Bolineol could not be isolated from the expression host, due to coelution with **132**. Therefore, desmethylbolineol **133** was isolated from *A. oryzae* NSAR1 transformant of exp. 6 (30 mg/L culture) and was smoothly converted to **4** by treatment with trimethylsilyldiazomethane (see section 6.5.6).¹⁸² The agar toxicity test revealed a MIC of 0.1 mg/ml for **4** against *A. oryzae* NSAR1 (Fig. 3.18).

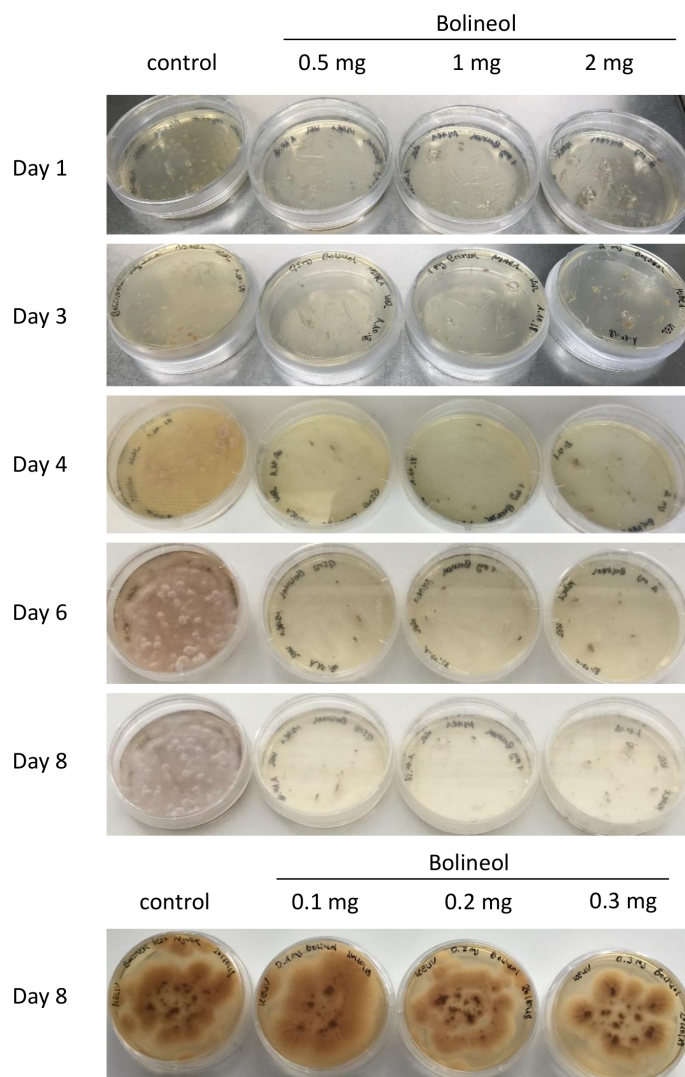
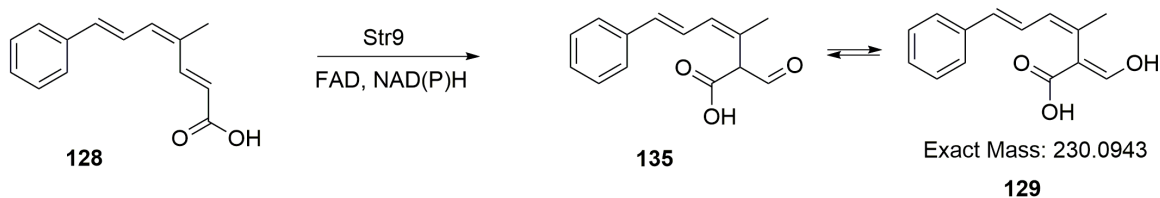


Figure 3.18 Toxicity test with bolineol 4 against *A. oryzae* NSAR1. DPY agar (5ml) was supplemented with various concentrations (0.1, 0.2, 0.3, 0.5, 1, 2 mg) of compound 4 diluted in 200 μ l DMSO, control only 200 μ l DMSO.

3.4 Results – *In vitro* activity tests with Str9

FAD dependent oxidase Str9 was found to catalyse the unusual rearrangement to form the toxophore of **2** by heterologous expression experiments in *A. oryzae*. The rearranged product **133** was shown to be a reduced congener of the expected aldehyde product **135** *in vivo* (Scheme 3.5 A). In order to verify the *in vivo* results and to find the real product of the rearrangement, the sequence of Str9 should be codon-optimised and expressed in *Escherichia coli* in N-terminal his-tagged form and purified as soluble protein. Its activity should be tested in *in vitro* assays using purified prestrobilurin **128** as substrate and essential cofactors FAD and NAD(P)H (Scheme 3.5 B).



Scheme 3.5 Rearrangement catalysed by Str9: **A**, transformation catalysed by Str9 *in vivo* (heterologous expression in *A. oryzae* NSAR1); **B**, postulated reaction catalysed by Str9 *in vitro*.

3.4.1 Prestrobilurin producing strain

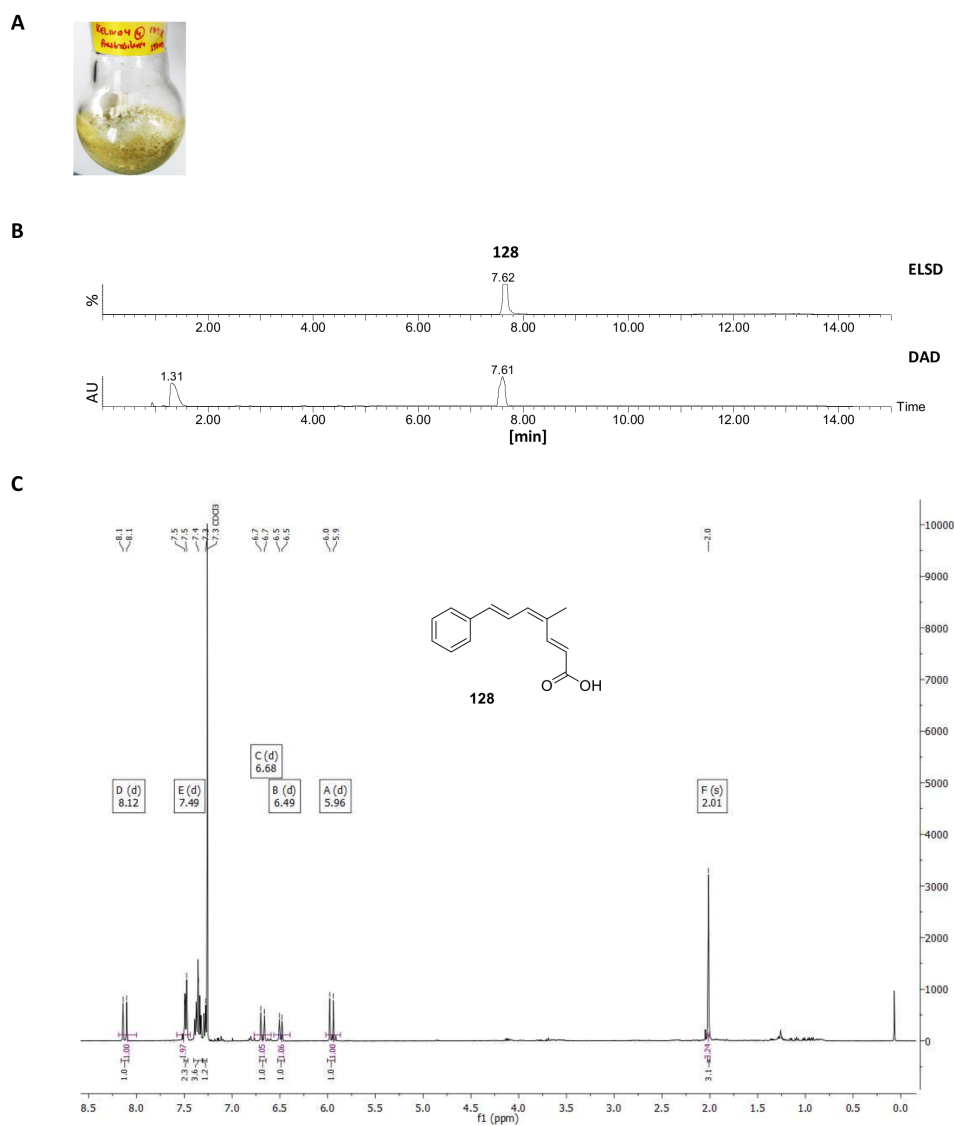


Figure 3.19 Analysis of crude extract of *A. oryzae* NSAR1+stpks1+str8+str10+str11+str2+str3+str4 (KELIV04_4, Fig. 3.6, exp. 4): **A**, corresponding yellow crude extract; **B**, LCMS analysis (DAD and ELSD trace) of the crude extract; **C**, ^1H NMR in CDCl_3 showing only 128 signals.

The appropriate substrate for *in vitro* assays with Str9 is prestrobilurin A **128** (Scheme 3.5).

The previously engineered heterologous expression strain *A. oryzae* NSAR1 containing *stpks1+str8+str10+str11+str2+str3+str4* (KELIV04_4, Fig. 3.6, exp. 4) showed production of very high titres of prestrobilurin A. Extraction of one litre of the supernatant of fermented KELIV04_4, using the common extraction method for *A. oryzae* (Section 6.5.4), yielded 150 mg of bright yellow crude extract (Fig. 3.19 A). Prestrobilurin A was previously described to be a yellow powder.^{175,185} LCMS and proton NMR analysis of the crude extract showed that the producing fungus almost exclusively produced the substrate **128**, as no contaminating peaks could be detected (Fig. 3.19 B + C). Therefore no further purification step was required.

3.4.2 Expression and purification of recombinant Str9

The *str9* gene was obtained as *E. coli* codon optimised N-terminal his-tagged clones, expressed in its soluble form in *E. coli* BL21 (16 °C, 20 h, 0.1 mM IPTG, section 6.4.1) and purified by gravimetric Ni-NTA affinity purification (Fig. 3.20 A, section 6.4.2).

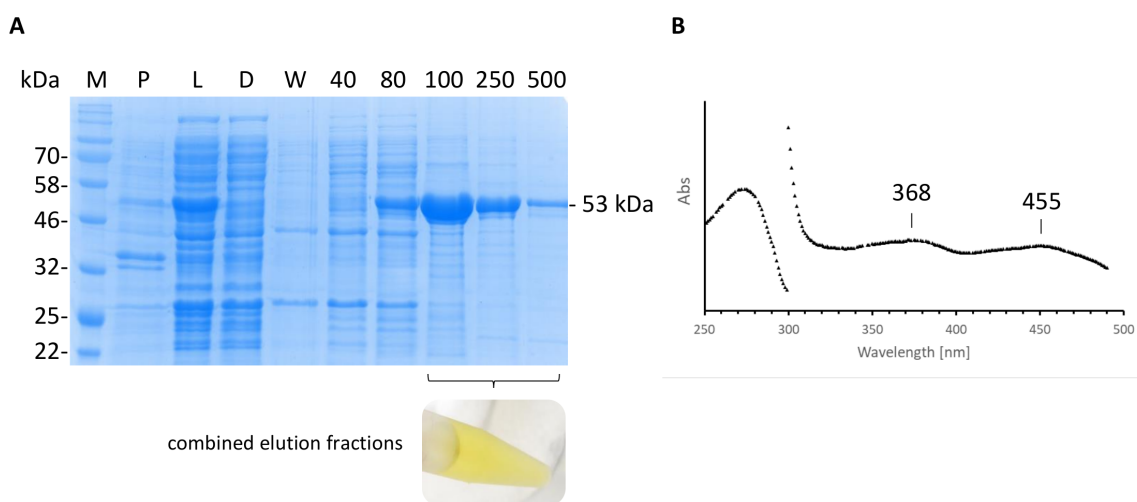


Figure 3.20 A, SDS-PAGE of Ni-NTA purified His₆-Str9, 12%- polyacrylamide gel, protein pre-stained marker (M), lysate (L), cell pellet (P), Flow-through (D), Elution fractions in [mM] imidazole with combined elution fractions after Ni-NTA of His₆-Str9 showing its yellow colour; B, UV/VIS spectrum of purified His₆-Str9 with FAD specific wavelength maxima at 368 and 455 nm.

The recombinant protein His₆-Str9 was shown to have a bright yellow colour suggesting its *holo*-form with cofactor FAD bound to the enzyme (Fig. 3.20 A). This was verified by ultraviolet spectroscopy showing the characteristic absorption maxima for FAD at 368 and 455 nm (Fig. 3.20 B). The protein sequence of Str9 was confirmed by mass spectrometry (Dr. Jennifer Senkler, Braun group, Leibniz Universität Hannover), using the methods of Klodmann and coworkers (Appendix 7.2.2 Fig. 7.5).¹⁹¹

3.4.3 *In vitro* activity test with Str9

Initial *in vitro* assays using purified prestrobilurin **128**, recombinant, purified His₆-Str9 and NAD(P)H in various buffers in air were performed, but no oxidised and rearranged product **135** could be observed by LCMS analysis (Fig. 3.21). The reason for this was most likely the observed insolubility of the substrate **128** in the aqueous buffer system.

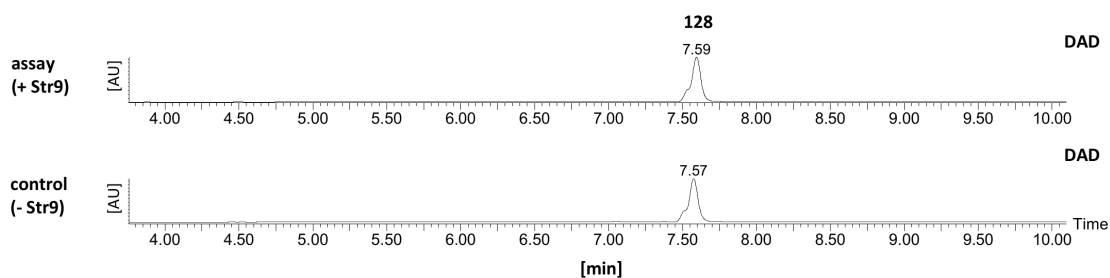


Figure 3.21 LCMS analysis of *in vitro* assay with substrate **128** and purified enzyme Str9 in comparison to a control without Str9 (DAD traces).

However, usage of a system containing a high concentration of ethanol (50%) led to conversion of trace amounts of the substrate. A product corresponding to **135** could clearly be observed by LCMS analysis showing characteristic UV maxima for strobilurins and corresponding mass spectra (Fig. 3.22, 3.23, 3.24).

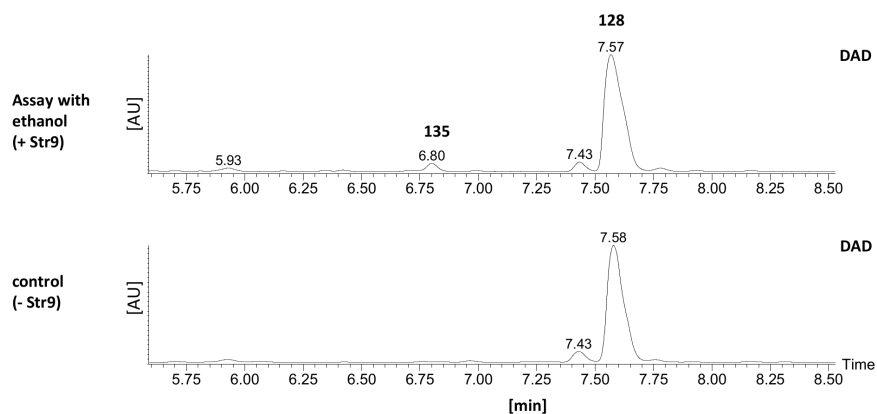


Figure 3.22 LCMS analysis of *in vitro* assay with substrate **128** and purified enzyme Str9 in comparison to a control missing Str9 with an excess of ethanol in the assay mixture (DAD traces).

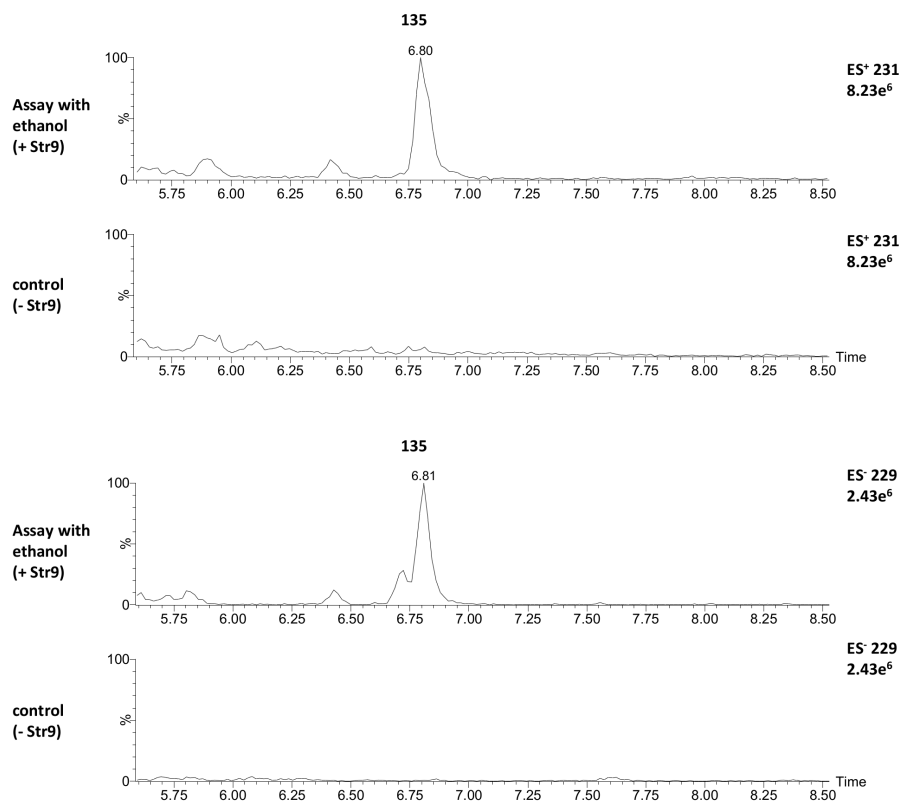


Figure 3.23 ES⁻ and ES⁺ extracted ion chromatograms of *in vitro* assay with substrate 128 and purified enzyme Str9 in comparison to a control missing Str9 with an excess of ethanol in the assay mixture.

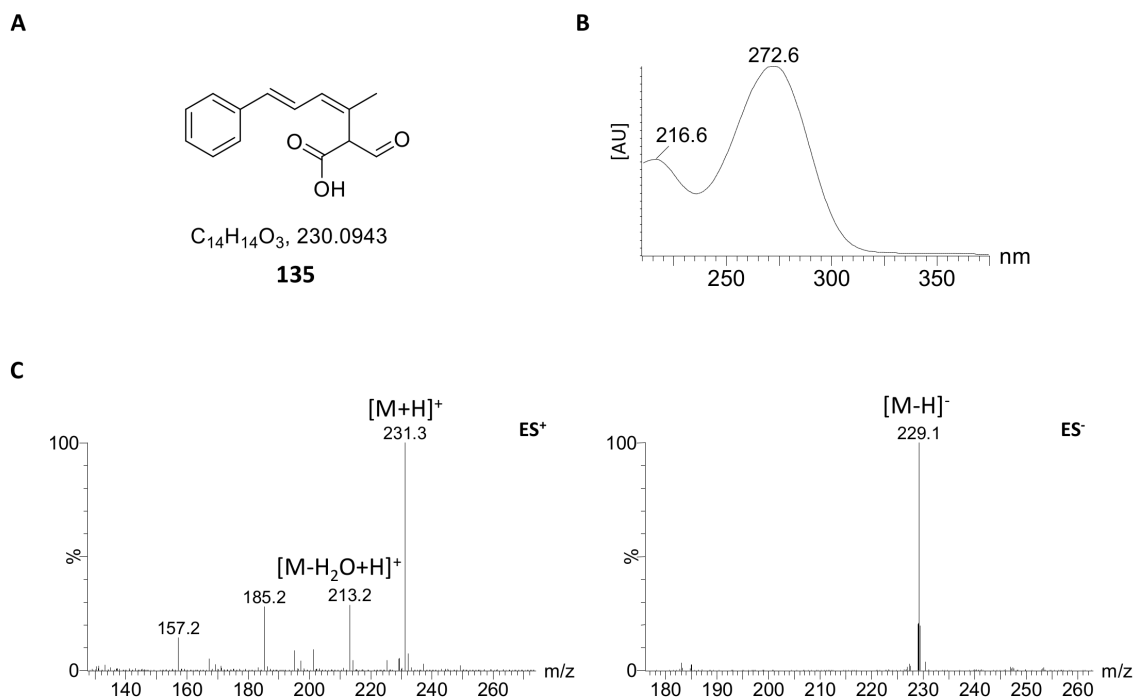
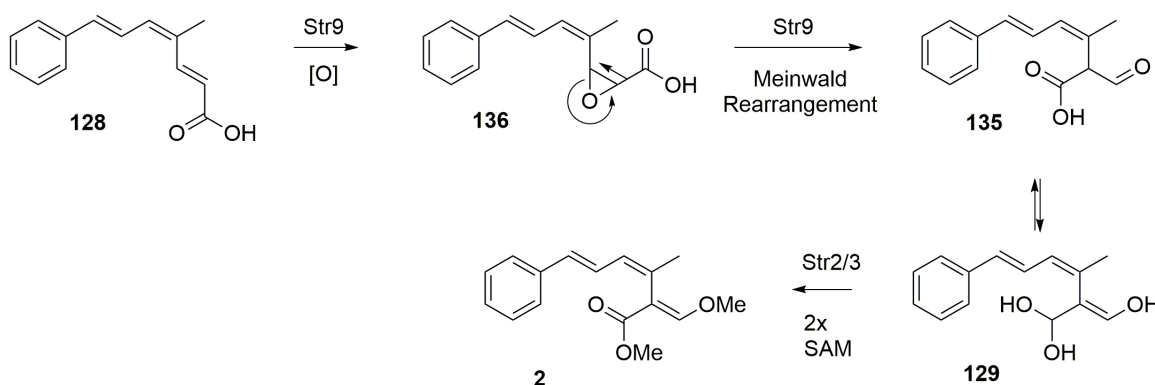


Figure 3.24 Chemical analysis of product 135 of *in vitro* assay with Str9: **A**, structure of product 135; **B**, UV chromatogram; **C**, corresponding ES⁻ and ES⁺ spectra.

3.5 Discussion

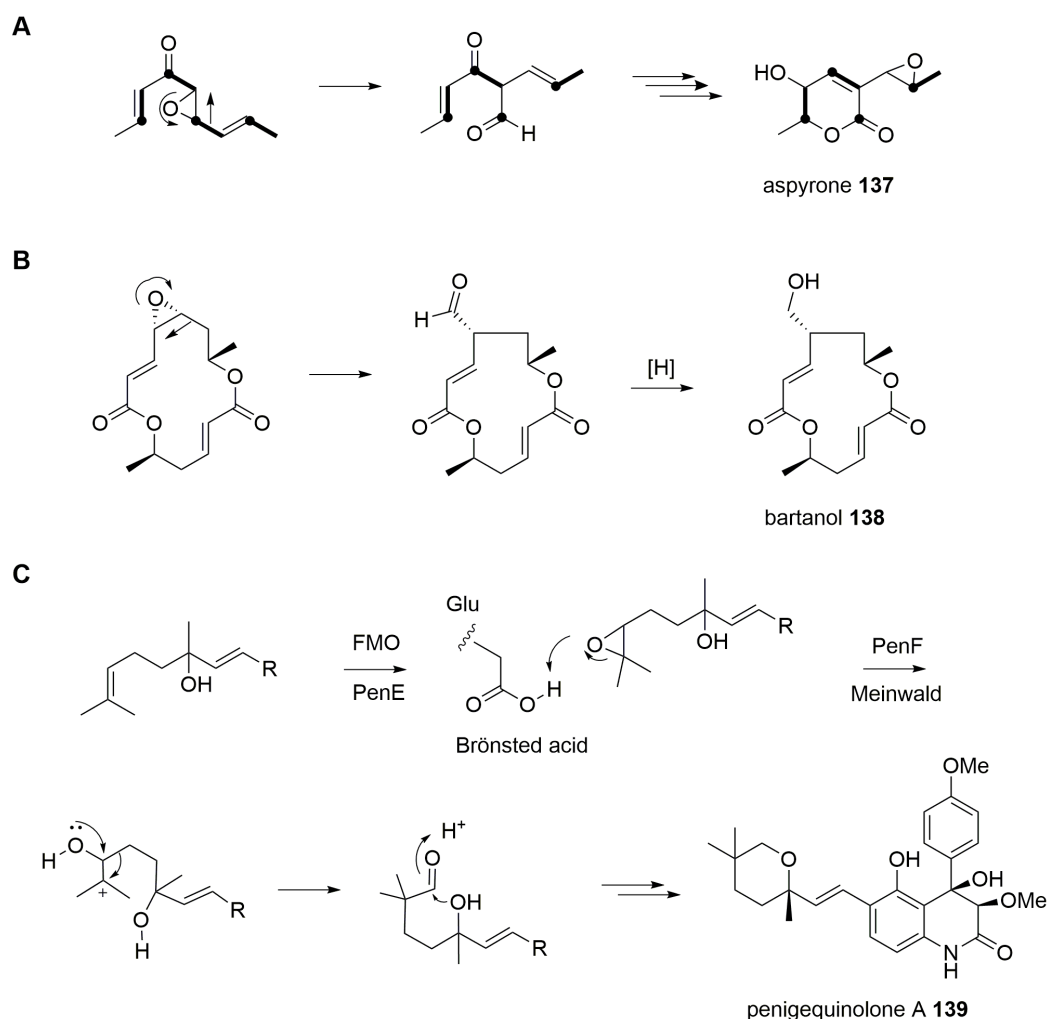
3.5.1 FAD dependent oxidase Str9 accomplishes both the oxidation and the rearrangement

Previous isotopic labelling experiments postulated that the released polyketide prestrobilurin **128** requires oxidation and rearrangement to form the key acrylate moiety of the strobilurins (Scheme 3.6).^{48,184} Heterologous expression of the FAD-dependent oxidase encoded by *str9* in the presence of **128** catalysed the oxidative rearrangement *in vivo* involving the highly unusual migration of a carboxylate. The same reaction was also observed in *in vitro* assays with the recombinant, purified His₆-Str9. Based on the *in vivo* and *in vitro* results it is proposed that epoxidation of the 2,3 olefin of **128**, followed by Meinwald rearrangement of **136** would give the aldehyde intermediate **135** (Scheme 3.6).



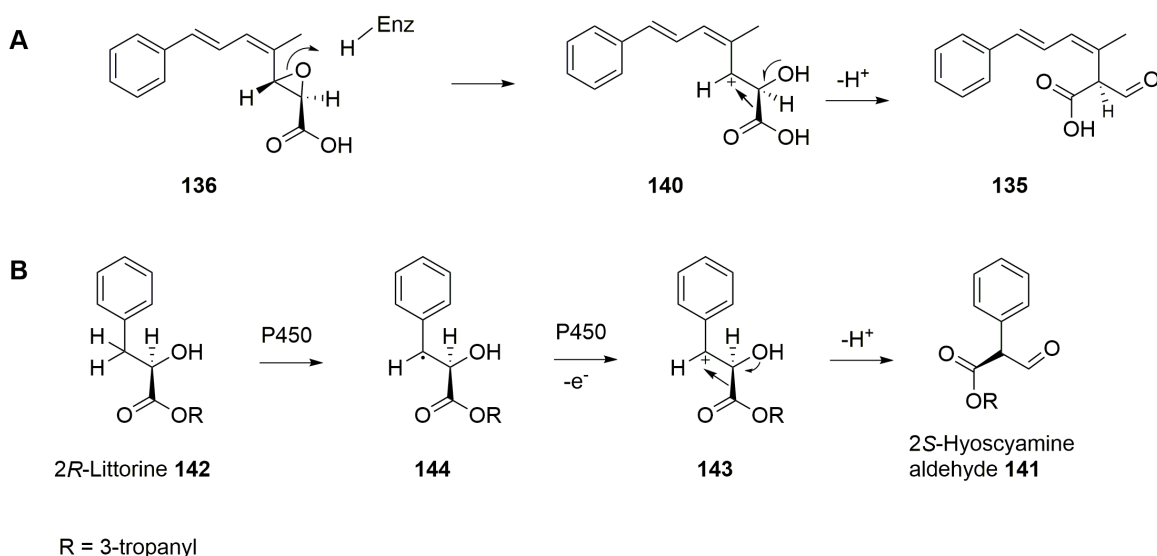
Scheme 3.6 Biosynthetic steps to strobilurin A 2.

Related oxidative rearrangements involving an epoxide intermediate are proposed for the biosynthesis of aspyrone **137**⁷⁴ and bartanol **138**¹⁹² (Scheme 3.7 A+B). Last year, Tang and coworkers found a family of CrtC-type (carotenoid 1,2-hydratase) enzymes, which can catalyse similar rearrangements of epoxides. For the particular CrtC-enzyme PenF this rearrangement is a Meinwald rearrangement on the pathway to penigequinolone A **139** (Scheme 3.7 C).¹⁹³



Scheme 3.7 Proposed oxidative rearrangements similar to Str9-catalysed reaction in strobilurin biosynthesis: **A**, biosynthesis of aspyrone **137**; **B**, biosynthesis of bartanol **138**; **C**, biosynthesis of penigequinolone A **139**.

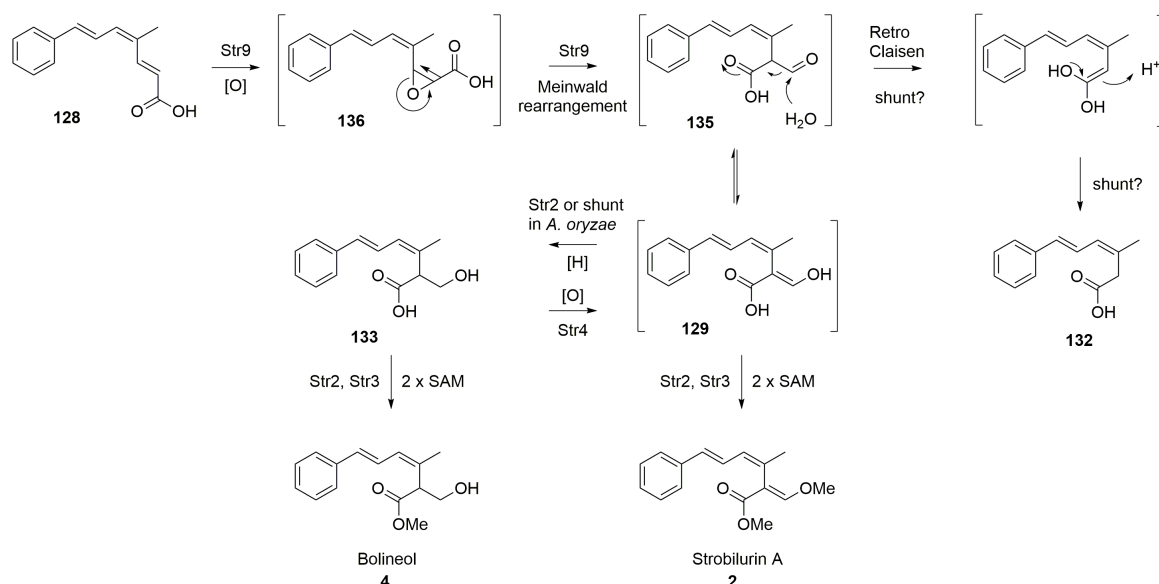
In the strobilurin rearrangement, formation of radical intermediates, as involved in similar processes during the mechanism of methylmalonate mutase¹⁹⁴, seems unlikely. A cationic mechanism *via e.g.* **140** is consistent with other Meinwald processes¹⁹⁵ and is proposed because it shows remarkable structural similarities to the biosynthesis of the tropane alkaloid hyoscyamine **141** from littorine **142**, in which a cytochrome P450 enzyme is known to catalyse the key rearrangement of **142**.¹⁹⁶ Calculations suggest that the lowest-energy pathway for this conversion involves rearrangement of a cationic species such as **143** obtained by facile oxidation of an intermediate radical **144**, with acceleration promoted by partial deprotonation of the alcohol and partial protonation of the carboxyl (Scheme 3.8).¹⁹⁶



Scheme 3.8 Proposed mechanism of the strobilurin rearrangement (A) and comparison with a proposed cationic rearrangement mechanism of 2*R*-littorine **142** (B).

In the *in vivo* expression experiments the product of the unusual rearrangement was shown to be **133**, which is probably a shunt product in *A. oryzae*, as *in vitro* studies with Str9 resulted in aldehyde **135** and *A. oryzae* is known to reduce aldehydes easily.¹⁵⁵ In addition, intermediate **135** seems to be highly reactive and in the absence of other enzymes it appears to undergo rapid retro-Claisen reaction in *A. oryzae* to give the observed carboxylic acid **132** (Scheme 3.9). This step is linked to the proposed oxidative ring contractions involved during xenovulene A **73** biosynthesis which are likewise catalysed by FAD-dependent enzymes AsR4 and AsR6 in *Sarocladium schorii*.⁹⁰ However, Str9 showed no significant sequence homology to AsR4, ASR6 and PenF, suggesting that Str9 accomplishes both the oxidation of **128** and the rearrangement to form **135**.

It is likely that in *Strobilurus* species rapid enolisation of **135** would give the β -methoxyacrylate skeleton **129** and methylation (catalysed by Str2 and Str3) would give strobilurin A **2** directly (Scheme 3.9).



Scheme 3.9 Biosynthetic steps to strobilurin A **2** and bolineol **4** including shunt pathways in *A. oryzae* observed in expression experiments.

3.5.2 Role of the GMC oxidase Str4 and SDR Stl2 in regulation of strobilurin and bolineol production

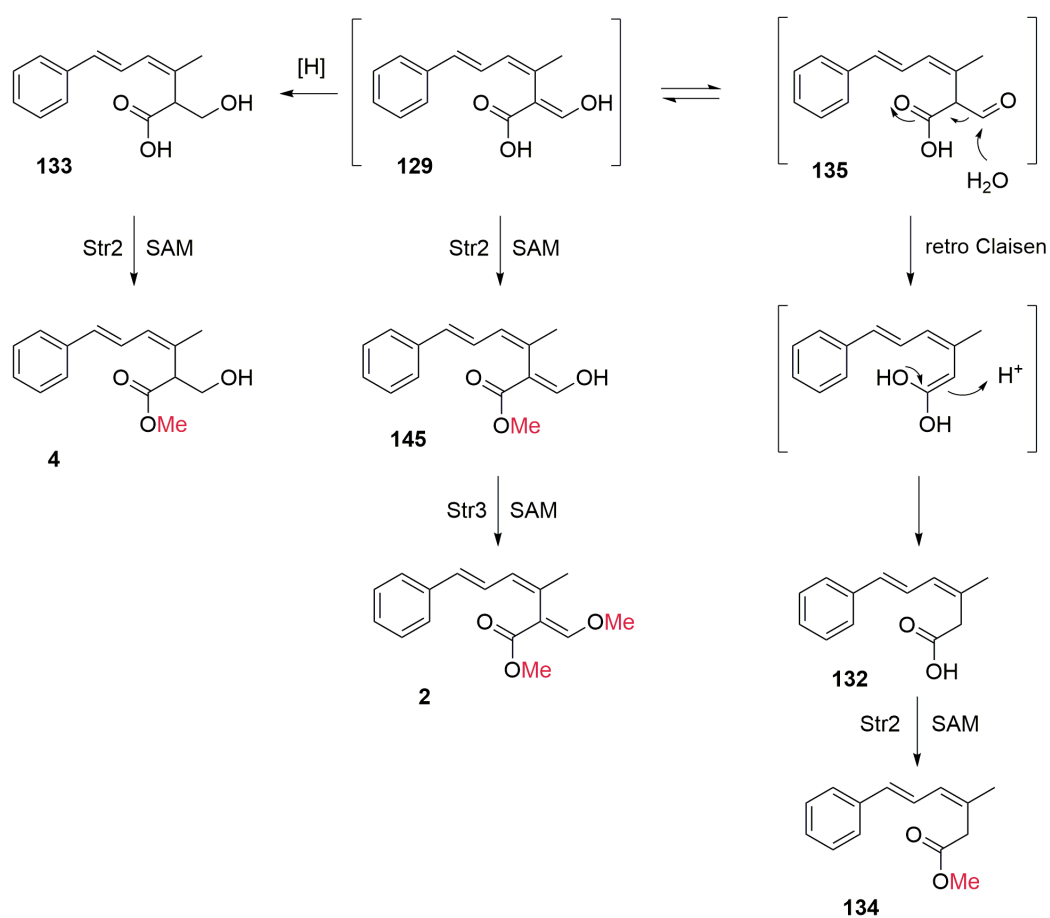
Heterologous expression experiments in *A. oryzae* revealed the reductive pathway to **133** and bolineol **4**, in which the SDR is possibly catalysing the reduction (Scheme 3.9). As mentioned before, the reduction occurs both in the presence of SDR Stl2 (exp. 1, exp. 8) and, in absence of Stl2 in heterologous host *A. oryzae* (exp. 7).

Together with oxidase Str4, Stl2 could be involved in controlling the selective production of bolineol or strobilurin in *S. tenacellus*. Str4 belongs to the GMCO superfamily, which are enzymes catalysing the oxidation of primary alcohols to aldehydes. It was shown that in the presence of Str4 and in the absence of the reductase Stl2, the pathway produces strobilurin A **2** and the shunt **132** (exp. 2). In the absence of Str4, and the presence of Stl2, the pathway produces bolineol **4** instead (exp. 8). This suggests that the selective production of **2** and **4** is possibly regulated by control of the individual promoters. In the *in vivo* expression experiments using *A. oryzae* as host the more crude control of the promoters (*e.g.* exp. 1) produces both compounds.

Another possible pathway from bolineol to strobilurin would include alcohol oxidation and subsequent *O*-methylation, but there was no LCMS evidence for a mono-methyl-intermediate.

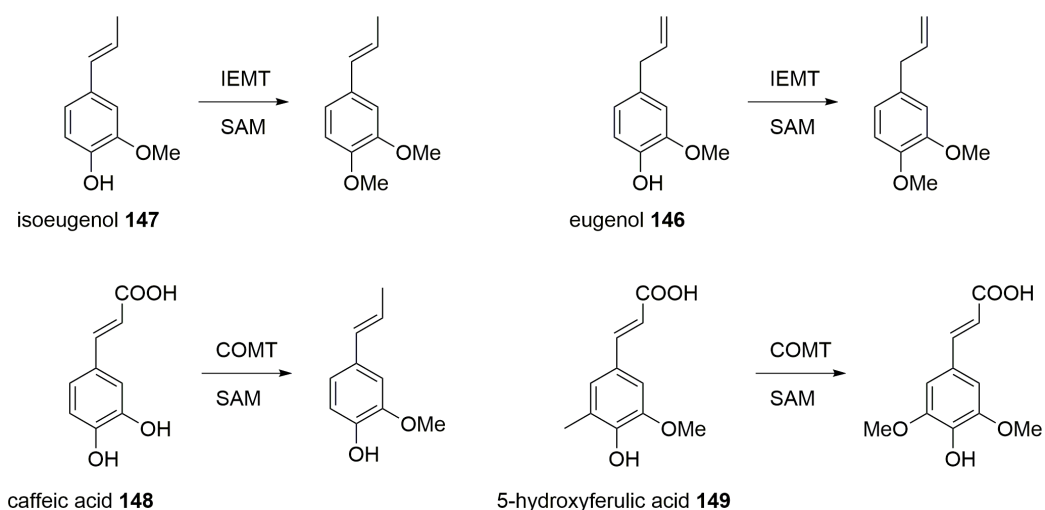
3.5.3 Role of the O-MeTs Str2 and Str3

Expression results showed that O-methylation in strobilurin biosynthesis is regiospecific. Both O-MeTs Str2 and Str3 catalyse unique steps in the biosynthesis to **2**. Even though they share 66% amino acid identity, these two O-MeTs showed different reactivity towards their substrates. In strobilurin A biosynthesis the initial O-methylation is selectively catalysed by Str2 at the carboxyl group of **129** to form **145**. Further O-methylation by Str3 gives **2** (Scheme 3.10). As in *A. oryzae* the aldehyde **135** is reduced to the alcohol **133**, also in absence of the reductive enzyme Stl2, we could not show existence of neither **129** nor **145** *in vivo*. Nevertheless, the expression experiments showed that Str2 is the O-MeT which catalyses the transfer of the methyl group to the carboxyl group of **133** as the final step of the biosynthesis to bolineol **4** (Scheme 3.10).



Scheme 3.10 Biosynthetic steps to strobilurin A **2** and bolineol **4** including methylation steps catalysed by Str2 and Str3.

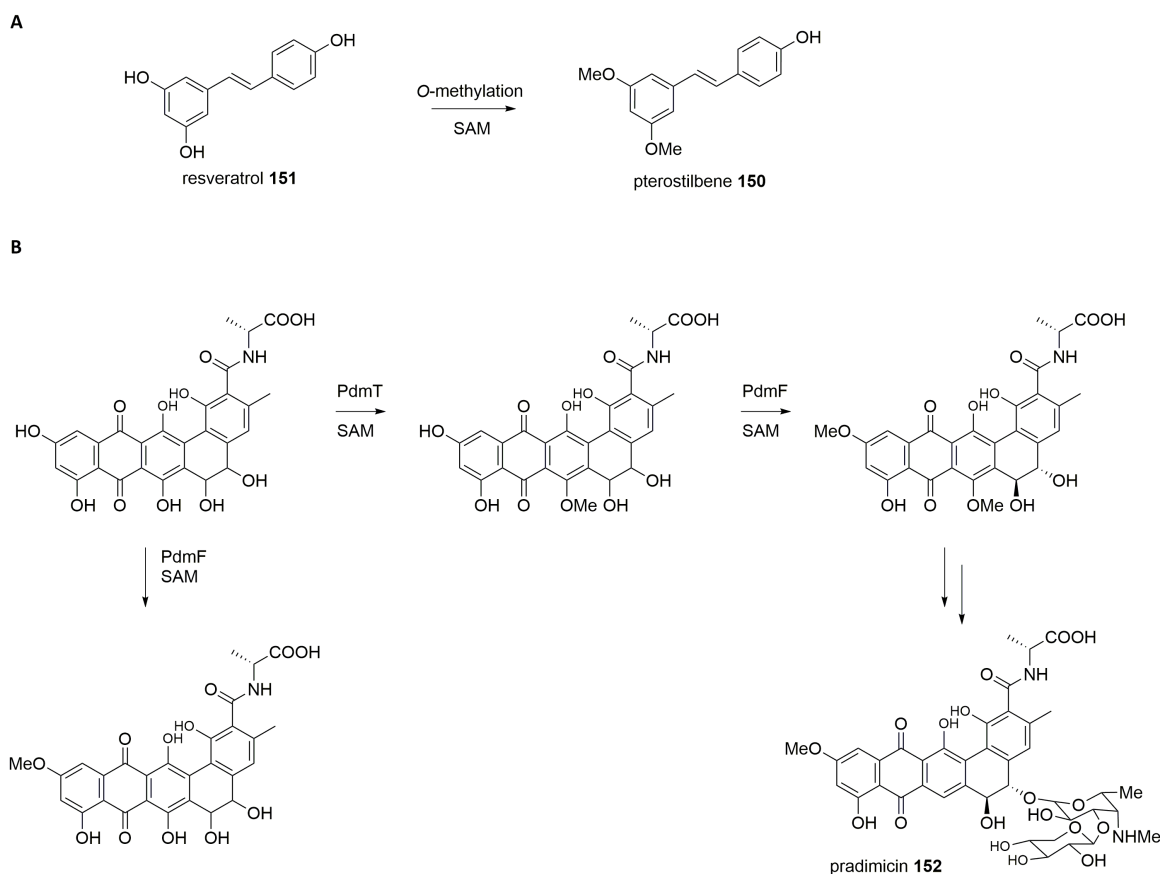
In nature, there are several examples for O-MeTs which share a very similar amino acid sequence but show different substrate selectivity or regioselectivity. For example, two O-MeTs from *Clarkia breweri* which share 83% identity: IEMT and COMT, which methylate eugenol **146**/ isoeugenol **147** and caffeic acid **148**/ 5-hydroxyferulic acid **149**, respectively (Scheme 3.11).¹⁹⁷



Scheme 3.11 O-methylation of eugenol **146**/ isoegenol **147** and caffeic acid **148**/ 5-hydroxyferulic acid **149**.

In general, O-methyltransferases have a huge influence of diversity of natural products. Methylation of hydroxyl groups lead to a decreased water solubility and favors excretion of a metabolite from the cell which affecting the biological activity.¹⁹⁸

According to plate toxicity tests with **4** and **133**, O-methylation by Str2 caused four times higher toxicity toward *A. oryzae*. Similar experiments testing strobilurin A fungicidal activity revealed a MIC of 25 µg/ml against a variety of plant pathogenic fungi growing on agar.¹⁹⁹ Therefore, a second O-methylation leads to 16 times higher toxicity against fungi compared to the unmethylated compound **133**, and four times higher compared to the singly methylated **4**. On account of these results, it can be assumed, that the toxophore group of strobilurins is more potent in its methylated form. This is in agreement with other natural products, like pterostilbene **150**, which bears a higher antifungal activity due to its two methoxyl groups, than its desmethylated precursor resveratrol **151** (Scheme 3.12 A).^{200,201}



Scheme 3.12 *O*-methylation secondary metabolite biosynthesis: **A**, *O*-methylation of resveratrol **151** to pterostilbene **150**; **B**, *O*-methylation in pradimicin **152** biosynthesis.

In experiment 10, a new compound **134** was detected. The amount produced was enough to detect it, but too little to isolate it for structure elucidation. Nevertheless, it is likely that *O*-MeT Str2 also transfers a methyl group to the Claisen product **132**, as its structure is very similar to the natural substrate **133** and is produced in high titres in *A. oryzae*. The amounts of compounds produced in the expression host can vary significantly from the natural producer, probably due to wrong balance of expression rate of genes, based on random choice of promoters.

When *O*-MeT Str3 is expressed in the absence of Str2 no methylation product was found. Therefore, we reasoned that *O*-MeT Str3 requires *O*-methylation by Str2 prior to its own methylation step to **2**, as production of strobilurin A was detected when both *O*-MeTs were expressed (exp. 1+2), but not by expressing them individually (exp. 9+10). This is in contrast to observations made in pradimicin **152** biosynthesis, where one of the two regiospecific *O*-MeTs, PdmF, was shown to catalyse *O*-methylation regardless of a first methylation step by PdmT (Scheme 3.12 B).²⁰²

3.6 Conclusion and Outlook

Overall, these results contribute to the understanding of how nature has generated the valuable class of strobilurin agricultural used fungicides. On the basis of *in vivo* heterologous expression experiments in *A. oryzae* a minimal gene set for the production of bolineol **4** and strobilurin **2** was elucidated. From previous studies by Risa Nofiani it is already known, that production of PKS precursor benzoyl CoA from phenylalanine involves expression of *str11* (PAL), *str8* (NHI) and *str10* (CoA ligase). Coexpression with the core biosynthetic gene *stpks1* (PKS) revealed production of prestrobilurin **128**.^{175,185}

In this thesis, it was shown that addition of *str9* (FMO) and *str2/3* (*O*-MeTs) to the previous expression system was sufficient in order to produce **2** and **4** *in vivo* (exp 7). The FAD dependent oxidase Str9 was found to catalyse the key oxidative rearrangement of a polyketide backbone (**128**) to form the crucial β -methoxyacrylate toxophore of the strobilurin fungicides, most likely in a Meinwald rearrangement. This was verified *in vitro* by conversion of the substrate **128** to the rearranged product **135** by the purified enzyme (His₆-Str9). However, to understand the mechanism, further optimisation of the *in vitro* assay and experiments with isotopic labelled prestrobilurin derivatives are required.

The expression results further revealed that regiospecific *O*-methylation, by two *O*-MeTs from **2** BGC, Str2 and Str3, contributes to a better understanding of the detailed biosynthetic steps to form the pharmacological toxophore of the strobilurins. Str2 was shown to methylate prior to methylation by Str3 to give **2**. In the reductive pathway to bolineol, which is probably catalysed by Stl2 (SDR) in *S. tenacellus*, the *O*-MeT Str2 is responsible for methylation as the final step of the biosynthesis. Activity tests revealed that the rearranged, unmethylated intermediate **133**, the singly methylated bolineol and the doubly methylated **2** are toxic towards *A. oryzae*. With increased number of methyl groups the toxicity towards *A. oryzae* was shown to be increased.

In addition to the elucidation of the biosynthetic pathway to **2** and **4**, two possible shunt pathways were discovered in *A. oryzae*. The product of the oxidative rearrangement was shown to be highly active in the expression host and was rapidly shunted in a *retro*-Claisen reaction to the carboxylic acid **132**, missing the β -arm of the molecule. It is possible that this "shunt" pathway is a kind of self defence mechanism in *A. oryzae* to protect itself from production of **2**, which is highly toxic. Also the reductive pathway to bolineol, which was catalysed by an endogenous reductase from *A. oryzae* in the absence of Stl2 (SDR), could display the same kind of self defence mechanism. In activity tests it was shown that **4** is less toxic than **2**.

There are still a number of unsolved questions about the biosynthesis of strobilurins, which could be addressed in the future. *S. tenacellus* produce a range of modified strobilurins, such as the chlorinated strobilurin B **127**. However, the cluster encodes no obvious halogenase gene that could be involved in

its production. It is possible that unknown proteins encoded in, or near, the BGC (e.g. *str1*) may form a new class of halogenases. It is also possible that proteins encoded by *str5/str6* (hydrolases) which appear to have no chemical role in the biosynthesis of **2** and **4**, could be involved in tailoring events in the biosynthesis of other, modified strobilurins. Furthermore, the domain organisation of StPKS differs from other highly reducing iterative fungal PKS, as it appears to include very unusual partial hydrolase and C-MeT domains downstream of the usually terminal ACP. It would be interesting to study these domains on the molecular base, also in order to understand how the highly unusual *E,Z,E* triene **128** is built.

4 Biosynthetic studies of SCH-642305

4.1 Introduction

The unusual bicyclic 10-membered macrolide, SCH-642305 (SCH 3, Fig. 4.1), was first isolated from fermentation broth of the fungus *Penicillium verrucosum* (culture ILF-16214) in 2003 as a novel bacterial DNA primase inhibitor (EC₅₀ of 70 μM).¹⁶ SCH is structurally related to the fungal metabolite brefeldin A **153** which has been previously investigated. It is therefore useful to begin with a review of what is known of its biosynthesis.

4.1.1 Brefeldin A biosynthesis

Brefeldin A **153** is a bicyclic polyketide isolated from various filamentous fungi such as *Eupenicillium brefeldianum* ATCC 58665.^{203,204,205} **153** consists of a cyclopentane fused to a 13-membered macrolide which is structurally related to SCH 3 (Fig. 4.1). Brefeldin A **153** was shown to have antiviral, antifungal and antitumor properties.²⁰⁶ Nevertheless, **153** is mainly used to study protein transport among eukaryotes as it is a potent inhibitor for vesicle formation and transport between the endoplasmic reticulum and Golgi apparatus.^{207,208,209}

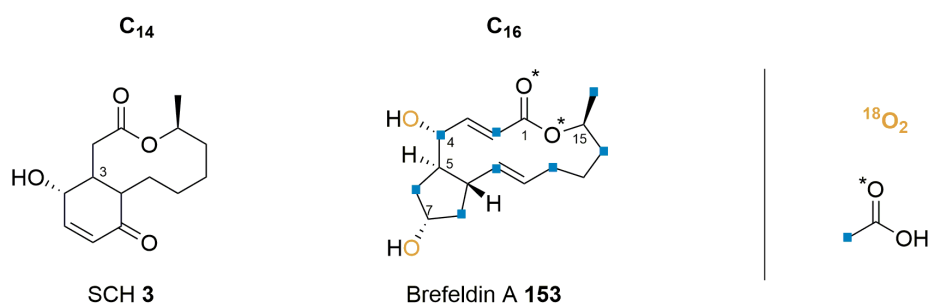
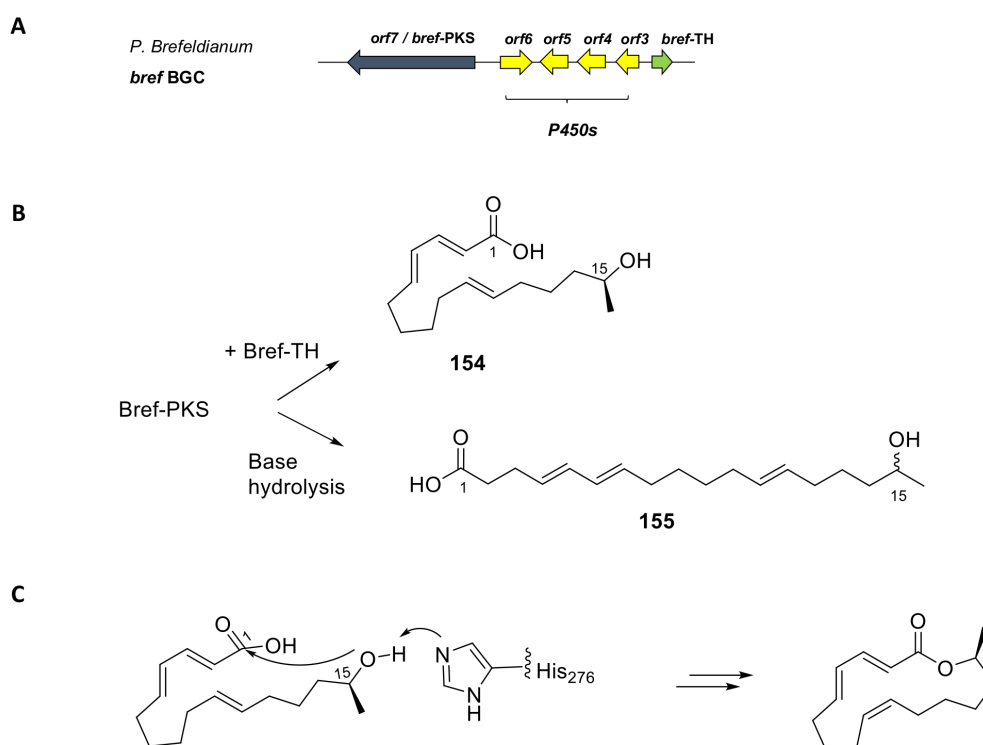


Figure 4.1 Incorporation of labelled sodium [¹⁸O₂, 2-³H₃]-acetate and ¹⁸O₂ into **153**.

The biosynthetic origin of **153** was studied by labelling studies, that revealed it to be an octaketide.^{210,211,212} Feeding sodium [¹⁸O₂, 2-³H₃]-acetate and ¹⁸O₂ revealed that oxygen atoms at C-1

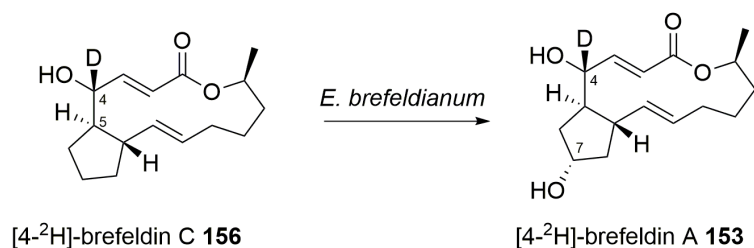
and C-15 are derived from acetate and the two hydroxyl groups at C-4 and C-7 are derived from atmospheric oxygen (Fig. 4.1).²¹²

Tang and collaborators found a likely *bref* BGC in *Eupenicillium brefeldianum*, which consists of a hrPKS, a thiohydrolase and four P450 encoding genes (Fig. 4.3). Reconstitution of the hrPKS (Bref-PKS) and the partnering thiohydrolase (Bref-TH) in yeast and *in vitro* led to production of a linear triene octaketide **154**, which possesses the same length (16 carbons) and expected β -reduction pattern for formation of **153**. Expression of only Bref-PKS in absence of Bref-TH yielded a linear nonaketide **155** (Scheme 4.1 A).²⁰⁵ These results indicated that Bref-TH has a critical role in chain release control of the octaketide product from Bref-PKS. It is important to notice that the polyketide starter unit at C-15 remains hydroxylated and that all products of Bref-PKS are *E* configured trienes.



Scheme 4.1 A, BGC for brefeldin A **153** (*bref* BGC) in *P. Brefeldianum*; B, chain-length control of the hrPKS product of **153** biosynthesis; C, initial proposal for macrolactonisation in **153** biosynthesis.²⁰⁵

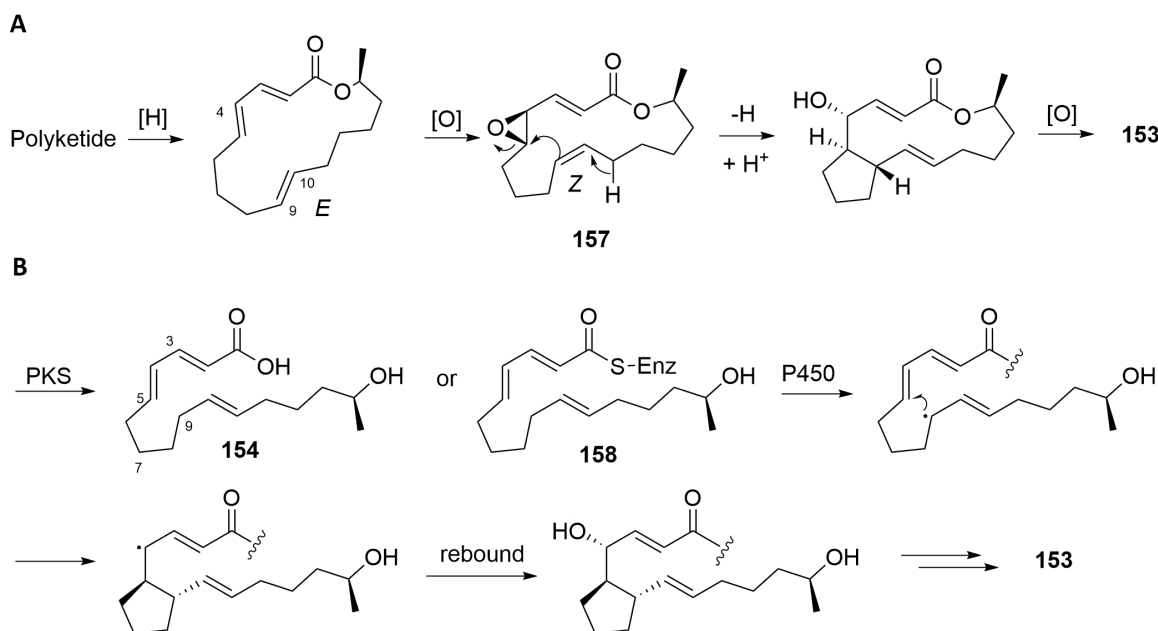
The macrolactonisation is not yet elucidated for **153**. Feeding a labelled brefeldin derivative ([4-²H]-brefeldin C **156**) to *E. brefeldianum* led to biotransformation of **156** to [4-²H]-brefeldin A, demonstrating the C-7 hydroxylation to be a late stage event, that does not contribute to the cyclopentane ring closure (Scheme 4.2). This suggests that the ring formation in **153** biosynthesis does not rely on C-7 oxygenation.²¹⁰



Scheme 4.2 Conversion of [4-²H]-brefeldin C to [4-²H]-brefeldin A by *E. brefeldianum*.

Tang and coworkers proposed a mechanism where the Bref-TH performs off-loading and lactonisation using its catalytic His-276 to deprotonate the OH-15 of **154** yielding a suitable nucleophile for lactonisation (Scheme 4.1 B).²⁰⁵ However, all observed *in vitro* products, **154**, **155** and some minor compounds (not shown) were linear.²⁰⁵ This suggests the necessity of closing the cyclopentane ring first by bringing the nucleophile closer to favour the attack to perform lactonisation (Scheme 4.3 B).

The mechanism of the cyclopentane ring formation is not yet elucidated. Two groups have made different proposals. Mabuni *et al.* suggested a mechanism which involves a C-4/C-5 epoxide intermediate **157** and subsequent ring opening (Scheme 4.3 A). The latter requires an unnatural Z C-9/C-10 olefin which has to be put in place by post tailoring steps.²¹² Tang and coworkers suggested a radical-mediated cyclisation catalysed by one of the four P450 enzymes encoded in *bref* BGC (Scheme 4.3 B).^{163,205} Either the PKS-bound intermediate **158** or the isolated linear acid **154** can be the substrate for the P450-catalysed C-C coupling (Scheme 4.3 B).²⁰⁵



Scheme 4.3 Two proposed mechanism for macrocyclisation in **153** biosynthesis: **A**, mechanism proposed by Mabuni *et al.*²¹²; **B**, radical mechanism by Tang *et al.*, adapted from Tang *et al.*, 2016.¹⁶³

4.1.2 Isolation of SCH and related compounds

In 2011, the endophytic filamentous ascomycete *Phomopsis* sp. CMU-LMA was shown to produce **3** alongside a variety of other polyketides (Fig. 4.2).⁸ This strain showed a 20 times higher production of SCH (50 mg/l) than reported in previous studies with *Penicillium verrucosum*.¹⁶ In addition to **3**, the other produced compounds were classified in four groups: **A**, substituted 10-membered lactones SCH-642305 **3**, LMA-P1 **159**; **B**, 14-membered lactone benquoine **160**; **C**, six-membered lactones LMA-P2 **161**, DHTO **162**; and **D**, linear derivatives LMA-P3 **163**, benquinol **164** and DHTTA **165** (Fig. 4.2).⁸

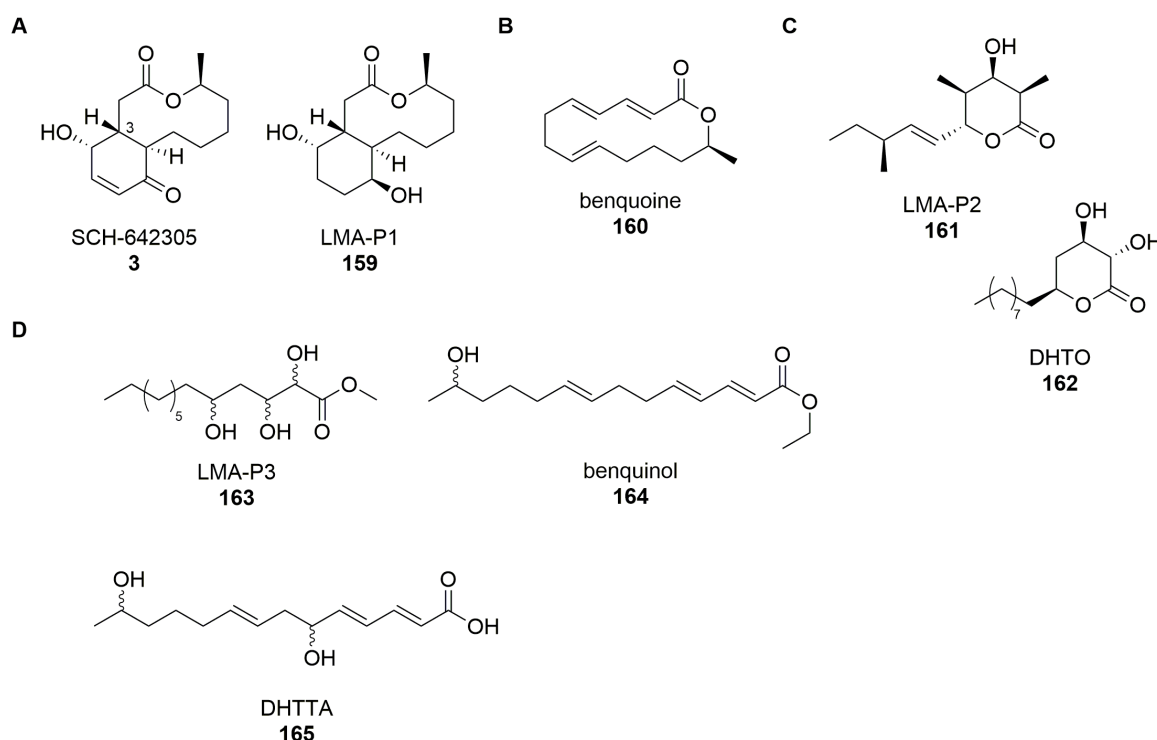


Figure 4.2 Compounds isolated from *Phomopsis* sp. CMU-LMA.⁸

In this study, the focus is on the major compound **3**. SCH possesses not only antibacterial properties but also antiviral features (against HIV-1 Tat protein) and a limited fungicide activity.^{213,214} The structure of SCH was determined by MS/NMR analysis and absolute configuration of the asymmetric centers was determined by X-ray crystallographic analysis of the *para*-bromobenzoate derivative of **3**.¹⁶ Due to its possible usage as an antibiotic against multi drug resistance *Staphylococcus aureus* and as a bioactive compound in AIDS therapy, a number of total syntheses were published between 2005 and 2009.^{213,215} SCH **3** appears to be related to brefeldin **153** because it is a bicyclic macrolactone in which the lactone is formed to an alcohol at the penultimate chain position. However, unlike **153** (C_{16}) the chain of **3** is shorter (C_{14}), the carboxylic ring is 6-membered *versus* 5-membered and ring junction is at C-3 rather than C-5 in brefeldin. However, it is easy to imagine how a brefeldin-like

thiohydrolase could provide a shorter polyketide in the case of SCH, and how a similar ring-forming process could form the 5- and 6-membered carbocycles.

4.1.3 Gene cluster and first biosynthetic studies of SCH

In silico analysis of *sch* BGC

In previous studies, the genome of *Phomopsis* CMU-LMA was sequenced (by Jamal Ouzzani and Emilie Adelin, 2005, ICSN Gif-Sur-Yvette France) and 10 years later the raw data was assembled by Daniel Wibberg (CeBiTec, Bielefeld). Bioinformatic studies, conducted by Francesco Trenti (Cox group, Hannover), revealed a putative *sch* BGC by homology searches using *brefPKS*.⁸⁵ The *Phomopsis sch* BGC was shown to be homologous to *bref* BGC, and was shown to contain genes for a hrPKS (*schPKS*), a hydrolase (*schR1*), two cytochrome P450s (*schR2/R3*), a serine-hydrolase (*schR4*), a NAD dependent oxidoreductase (*schR5*), a transporter (*schR6*), and an FAD dependent oxidase (*schR7*, Fig. 4.3). *SchPKS* and *Bref-PKS* were shown to share 56% sequence identity and the *trans*-partnering hydrolases *Bref-TH* and *SchR1* were found to share 66% sequence identity at amino acid level.⁸⁵ Additional *in silico* analysis displayed another homologous cluster (*PVSCH* BGC) in a known SCH producing strain *Penicillium verrucosum*, which consists of only four genes, including a hrPKS (*PVSCH-PKS*), a thiohydrolase (*R1*), a P450 encoding gene (*R3*) and an FAD dependent oxidase encoding gene (*R2*, Fig. 4.3).¹⁶

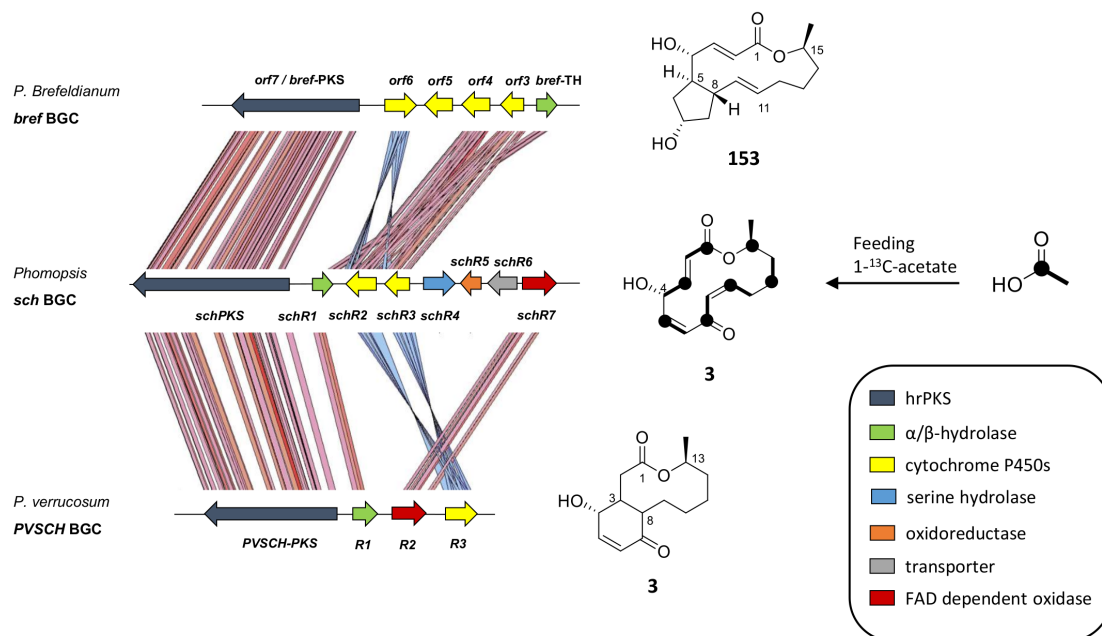


Figure 4.3 Comparison of homologous *bref* BGC from *P. brefeldianum*, *sch* BGC from *Phomopsis* sp. and *sch* BGC from *P. verrucosum*. Artemis double ACT was used for the pairwise comparison at translated nucleotide level, setting a threshold at 20. Red lines imply genes on the same DNA strand, blue lines for genes on opposite strands; done by Francesco Trenti.⁸⁵

The comparison of all three homologous clusters revealed a similar composition of genes. All BGC contain a conserved hrPKS encoding gene and a gene encoding an α/β -hydrolase.²⁰⁵ An FAD dependent oxidase encoding gene was only found in both *sch* BGC (*schR7*, *R2*), but not in *bref* BGC (Fig. 4.3). The number of P450 genes differ in the compared clusters from one (*sch* BGC, *P. verrucosum*), two (*sch* BGC, *Phomopsis*) to four in *bref* BGC. This might indicate gene duplication for the P450 genes (Fig. 4.3). As in *P. verrucosum* only one copy of a P450 gene is present in the *PV SCH* BGC, suggesting that one P450 would be sufficient for the biosynthesis of **3**. The putative *sch* BGC was later linked to production of SCH by disruption of the core hrPKS gene conducted by Francesco Trenti.⁸⁵

Brefeldin **153** and SCH **3** share a common structure and highly homologous genes including the core hrPKS and partnering hydrolase, which suggests a similar chemistry for the biosynthesis. According to **153** biosynthesis, also biosynthetic production of **3** needs a PKS product with a terminal hydroxyl group (O-13), an all *E* configured triene system and chain-length control most likely by SchR1 (hydrolase, in accordance to Bref-TH). However, the composition of post tailoring genes in both clusters (*sch* BGC + *bref* BGC) differs. *Sch* BGC occupies two genes encoding for cytochrome P450s whereas *bref* BGC contains four P450 encoding genes (Fig. 4.3). Furthermore the genes *schR4*, *schR5* and *schR7* show no homologue in the *bref* BGC, suggesting a distinctive pathway for SCH biosynthesis.

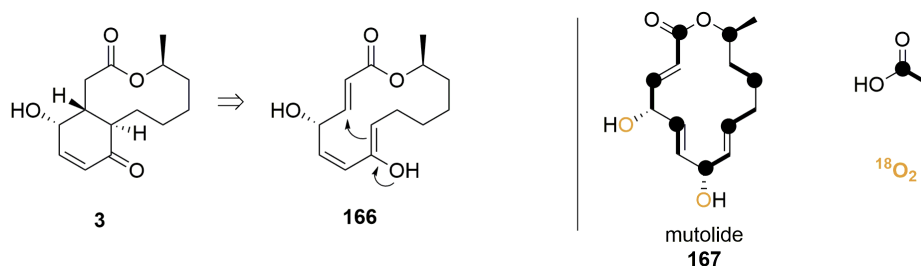
Feeding studies of SCH

Feeding [1-¹³C]-acetate confirmed **3** to be a highly reduced heptaketide, with a consistent alternating pattern of incorporation of the label (Fig. 4.3).⁸⁵ This result excluded atom rearrangement during the formation of the 6-membered ring and the oxidation pattern of **3** suggested post-elongation oxidative events, especially regarding the hydroxyl group at position C-4, which is probably derived from atmospheric oxygen.⁸⁵

Proposed biosynthetic pathway for SCH

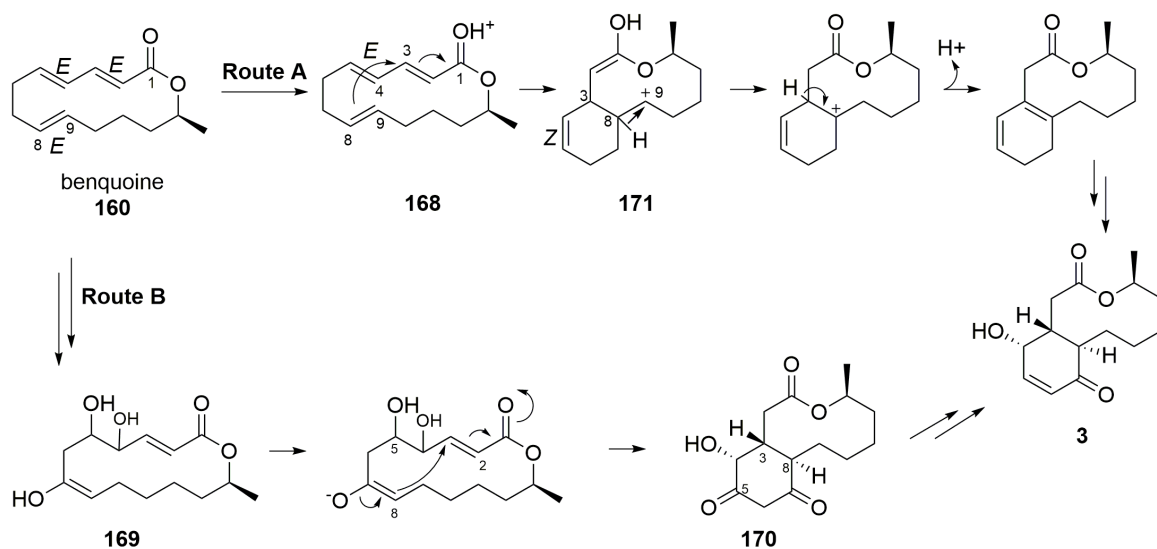
Until 2018 there was not much known about the biosynthesis of SCH. In 2006, Snider and Zhou reported a total synthesis towards **3**. Their strategy was based on transannular Michael cyclisation of a hypothetical 14-membered lactone **166** which is structurally related to mutolide **167** (Scheme 4.4).²¹⁵ Mutolide is a polyketide which was isolated in 2000 from fungus F-24'707.²¹⁶ The biosynthetic origin of the carbon skeleton of **167** is derived from seven acetate/malonate units which was shown by feeding experiments with sodium [1-¹³C]-acetate (Scheme 4.4). Cultivation of the producing strain in a closed ¹⁸O₂ atmosphere indicated the incorporation of both hydroxyl oxygens by monooxygenases.²¹⁶ The hypothetical biosynthetic route to SCH **3** proposed by Snider and Zhou on the basis of

the fungal metabolite mutolide **167** is supported by their synthesis. However, **167** has never been isolated from any 3-producing strain, and it requires a 5Z-configured olefin.



Scheme 4.4 A hypothetical biosynthetic route to SCH **3** proposed by Snider and Zhou on the basis of the fungal metabolite mutolide **167**.^{215,216}

In contrast, the isolation of the all *E*-configured benquoine **160** revealed a different hypothesis for the biosynthetic pathway of SCH in which **160** would be the precursor. Adelin *et al.* proposed two routes to SCH **3** starting from benquoine **160** (Scheme 4.5). Route A involves carbonyl protonation to give **168** followed by intramolecular rearrangements, whereas route B starts with a series of oxidoreductions to yield **169** and further deprotonation and subsequent enzyme-catalysed Michael cyclisation results in **170**. Reduction of the carbonyl at C-5 would yield **3** (Scheme 4.5).⁸ An indication that route A is probably wrong is given by the C-4/C-5 olefin geometry switch, where the *E* configuration of **168** is transformed to the *Z* configuration of **171** without any reason. Furthermore, the proposed cationic chemistry of route A is more common in terpene biosynthesis and typically not found in fungal polyketide biosynthesis.



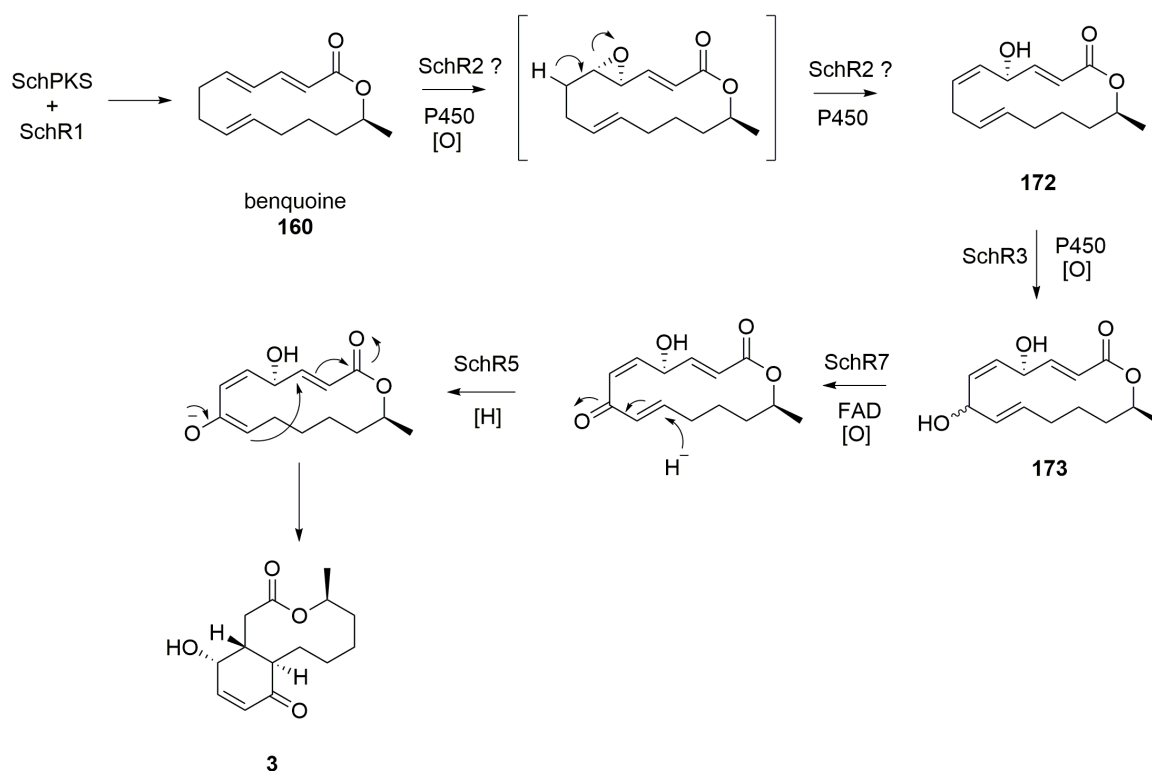
Scheme 4.5 Two different biosynthetic routes (route A + B) to SCH **3** proposed by Adelin *et al.* starting from benquoine **160**.⁸

Knockout studies revealed a new biosynthetic route to SCH

Previous knockout studies conducted by Francesco Trenti (Cox group, Hannover) confirmed the *sch* BGC by deletion of the core hrPKS and resulting abolition of SCH **3** production. As production of DHTTA **165** and LMA–P1 **159** were also abolished due to knockout of *schPKS* in the wt fungus, it seemed that these compounds may all be produced by the same PKS, but not necessarily on the same biosynthetic pathway. Evidence is given by occupying the same number of carbons (14 carbons) and a similar double bond system. Further the C-13 hydroxylation involved in the macrolactonisation is conserved among the mentioned compounds and compound LMA–P1 **159** shows the same bicyclic system as **3**.⁸⁵ In contrast, DHTO **162**, LMA–P3 **163** were still be observed from PKS KO extract suggesting them to arise from different pathways.⁸⁵

In *P. verrucosum* only one copy of P450 encoding gene (*R3*) might be sufficient to produce **3** (Fig. 4.3). Individual gene knockout of both of the P450 encoding genes from the *sch* BGC of *Phomopsis* led to abolition of SCH suggesting that both P450s are essential for **3** biosynthesis in *Phomopsis*.⁸⁵ Deletion of one of the P450s (*schR3*) resulted in accumulation of new compound **172** and knockout of FAD dependent oxidase encoding gene *schR7* led to observation of another new compound **173**. Therefore the SCH oxidations at C-4 and C-7 are likely inserted by the two P450s, which is in agreement with the labelling studies for **153** biosynthesis (Section 4.1.1). Oxidation at C-4 is possibly introduced by SchR2 and hydroxylation of C-7 is probably catalysed by SchR3.⁸⁵ Disruption of the other P450 encoding gene *schR2* led to abolition of **3** production but did not produce any new intermediate, as also observed for disruption of *schR5*. However, the abolition of **3** production for all KO mutants (*schPKS*, *schR2*, *schR3*, *schR5*, *schR7*) showed that all these genes were essential for SCH biosynthesis. Attempts to knockout the serine hydrolase encoding gene *schR4* were not successful as the gene *schR4* was not disrupted in the obtained mutants.⁸⁵

In accordance to the knockout studies of SCH, the proposed biosynthesis of **3** would start with benquoine **160** being the product of the highly reducing PKS in combination with thiohydrolase SchR1, leaving the 2*E*, 4*E*, 8*E* triene system in place (Scheme 4.6). The macrolactonisation might arise spontaneously during or after the chain release. Epoxidation by cytochrome P450 SchR2 at the 4*E* olefin and opening to the observed compound **172**, with subsequent double bond migration could be followed by action of the second P450 SchR3 inserting an hydroxyl group on C-7, yielding observed intermediate **173**. Thus the knockout experiments conducted by Francesco Trenti proved the C-5–olefin to be *Z*. The FAD dependent oxidase SchR7 could convert the C-7 hydroxyl group to a keto group, suggested by *in silico* studies. SchR7 was found to share 28% sequence identity to a known flavoprotein converting hydroxyl groups to ketone from ascomycete *Neurospora crassa*.⁸⁵ Putative reduction of the C-8/ C-9 olefin by oxidoreductase SchR5 would lead to the enolate which could perform a Michael–type cyclisation generating the 6–membered ring (Scheme 4.6).⁸⁵



Scheme 4.6 Proposed biosynthesis of SCH-642305.

Although **3** and **153** share a very similar structure, as well as homologue genes for the hrPKS and hydrolase, the biosynthesis appears to be different. Both compounds showed similar P450-mediated oxidation at C-7 after chain release. For **153** biosynthesis it was found that the C-7 oxidation happened after the 5-membered ring formation (shown by feeding studies, Section 4.1.1).²¹⁰ The occurrence of **172** and **173** in knockout studies showed that hydroxylation in SCH biosynthesis occurs before the 6-membered ring is arranged, followed by oxidation to the ketone.⁸⁵ This suggests a different mechanism of ring formation in the biosynthesis of **153** and **3**.

4.2 Project aims

Although SCH has a rather simple structure, the molecular basis of its biosynthetic steps have not been elucidated. Initial knockout studies by Francesco Trenti elucidated a gene cluster responsible for SCH production and discovered genes essential for production of **3**. The characterisation of two metabolites **172** and **173** accumulating upon those knockouts gave initial details about the order of steps in the biosynthesis. However, targeted gene deletion was found to be not efficient in the producing fungus due to unreliable production of metabolites. In addition, newly produced

intermediates were rapidly shunted or degraded by the cell and therefore making subsequent analysis complex.

To overcome the KO difficulties, *A. oryzae* NSAR1 should be used as a host organism to heterologously express genes from the *sch* BGC *in vivo*. Heterologous expression is a more reliable method to resolve the biological mechanisms that happen *in vivo* during its biosynthesis, with particular interest to answer the question about the chemistry behind the 6-membered ring formation and in-turn form a better hypothesis about brefeldin A biosynthesis. With this knowledge new hybrid metabolites could potentially be engineered. In order to test if SCH biosynthesis can be re-established in *A. oryzae* NSAR1, **3** was fed to the host organism in initial biotransformation experiments.

4.3 Results – Re-establish SCH biosynthesis in heterologous host *A. oryzae* NSAR1

4.3.1 Biotransformation of SCH in expression host *A. oryzae* NSAR1

For expression experiments, heterologous host *A. oryzae* NSAR1 was chosen. In 2011, Adelin and coworkers reported that incubation of Sch-642305 **3** with *Aspergillus ochraceus* ATCC 1009 resting cells led to three new SCH derivatives through an oxidoreduction of the six-membered ring of the molecule. Reduction of the double bond gave compound **174**, which subsequently undergoes carbonyl reduction to **175** and ring hydroxylation to **176** (Fig. 4.4).²¹⁷

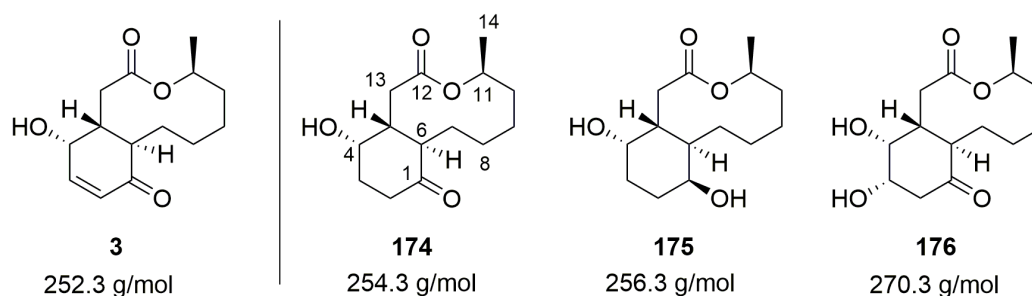


Figure 4.4 Structure of Sch-642305 **3** and its three major derivatives (**174**, **175** and **176**) obtained from biotransformation of **3** by *Aspergillus ochraceus* ATCC 1009.

In order to verify if *A. oryzae* also transforms SCH **3** to the previously observed derivatives, a feeding experiment was conducted. To a growing culture (2 days) of *A. oryzae* in DPY medium (4 x 100 ml), 7 mg of **3** were fed (diluted in 1 ml of DMSO). Extraction of the culture on day 5 of fermentation and subsequent submission to LCMS analysis revealed two new compounds eluting at similar retention times ($t_R = 5.0$ min and $t_R = 5.5$ min) as SCH ($t_R = 5.3$ min, Fig. 4.5 A). SCH was only observed in

traces. The nominal mass of both detected compounds was 254, two mass units higher than the mass of SCH (252 g/mol). Observed mass fragmentation in the positive mode showed the same fragmentation pattern for all three compounds. The sodium adduct, the proton adduct as well as fragments corresponding to two losses of water (for **174** and **177**) and three losses of water (for **3**) were observed (Fig. 4.6). This indicates that the new compounds **174** and **177** are most likely derivatives of SCH.

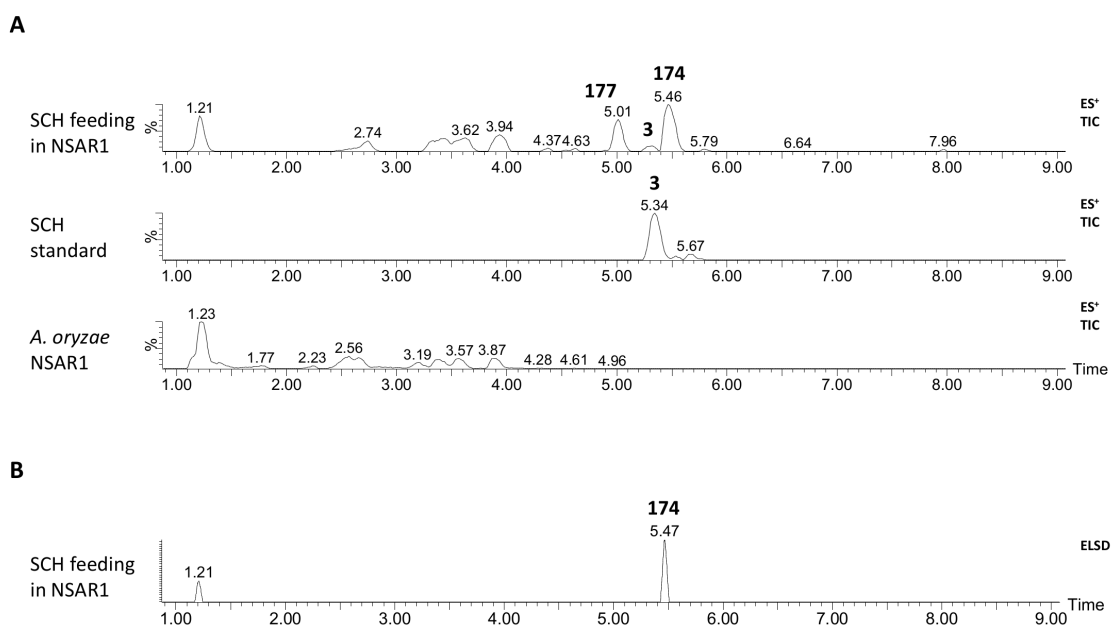


Figure 4.5 LCMS analysis of feeding experiment (**3** to *A. oryzae* NSAR1): **A**, ES⁺ traces of a control (*A. oryzae* NSAR1 wt), a standard of **3** and fed **3** to NSAR1; **B**, ELSD trace of **3** and fed **3** to NSAR1.

Further evidence is given by a very similar UV maxima for the three compounds ($UV_{max} = 216\text{--}219$, Fig. 4.6). Additional HRMS analysis of compound **174** revealed the chemical formula (ES⁻ C₁₄H₂₁O₄ calc. 253.1440, measured 253.1440). According to the ELSD trace of the feeding extract, high amounts of **174** were produced, but neither SCH nor **177** could be observed, suggesting very low yields (Fig. 4.5 B). Using preparative LCMS, 1.2 mg of **174** were isolated. Compound **177** could not be purified due to very low titre. Also HRMS analysis of this compound was prevented due to the same reason.

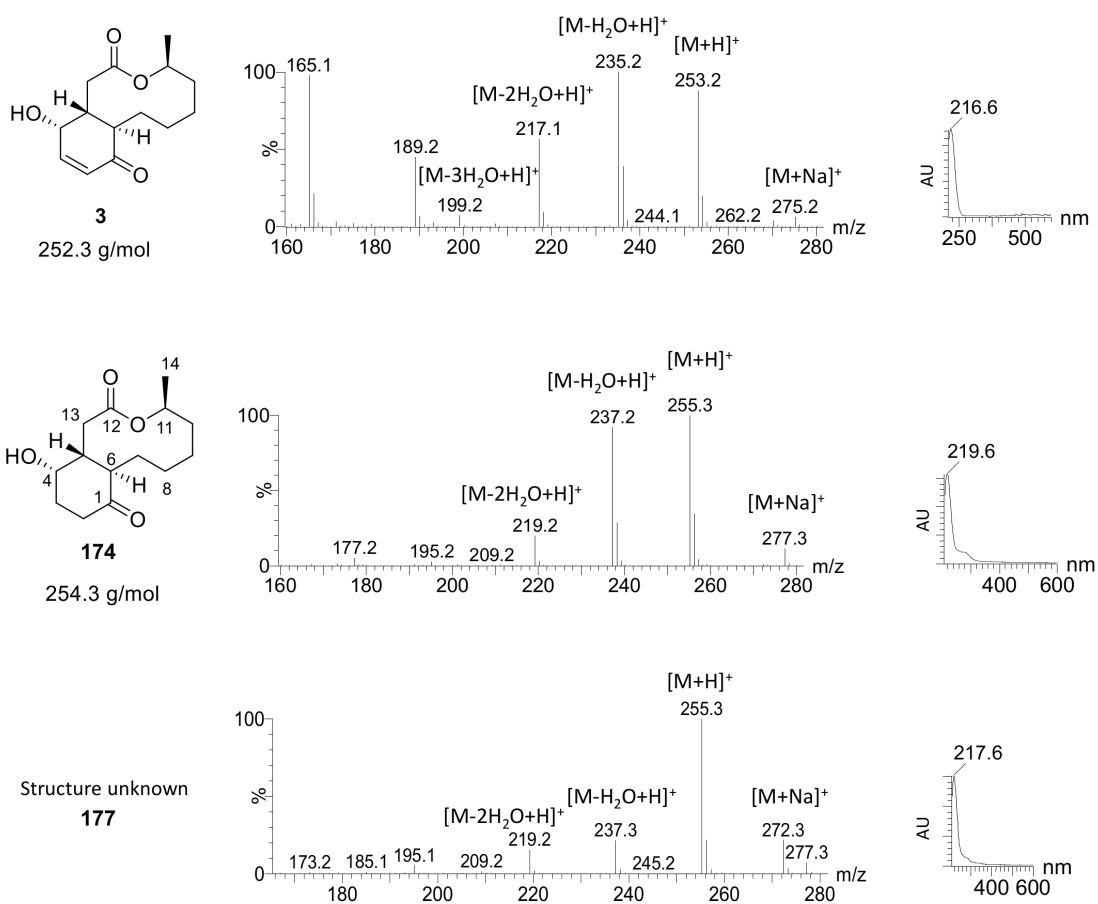


Figure 4.6 Mass spectra (ES⁺) and UV chromatograms of 3, 174 and 177.

NMR analysis (500 MHz) of 1.2 mg of **174** in CDCl₃ provided good quality data. The elucidated structure was found to be the same reduced SCH analogue **174**, previously observed in Adelin and coworkers experiment.²¹⁷ The ¹³C as well as the proton signals showed to be almost identical, with our data giving additional information (Table 4.1). The second newly observed compound **177** could be an isomer of **174** as it showed the same mass fragmentation pattern and UV maximum, but a slightly different retention time.

In conclusion, these results show that our heterologous expression host, *A. oryzae* NSAR1, also reduces SCH to **174** and a similar compound **177** through biocatalysis. The structure of the second observed minor compound **177** could not be elucidated due to very low titres. Also Adelin and coworkers found a minor compound with $m/z = 254$ eluting prior to SCH, but were not able to characterise its structure due to the same issue.²¹⁷ In our experiment only compound **174** and **177** were observed, whereas biotransformation experiments with SCH in *A. ochraceus* yielded two additional derivatives of SCH: **175** and **176** (Fig. 4.4). One explanation for the difference could be the different species used (here we used *A. oryzae* instead of *A. ochraceus*) or compounds **175** and **176** were built after day 5 of fermentation which was suggested by kinetic monitoring studies of the

biotransformation conducted by Adelin *et al.*²¹⁷ They showed that only **174** was built after 5 days and compound **175** was only started to be formed from day 5 on. Compound **176** even started to appear at day 7 suggesting it to be a direct product of **174**.²¹⁷

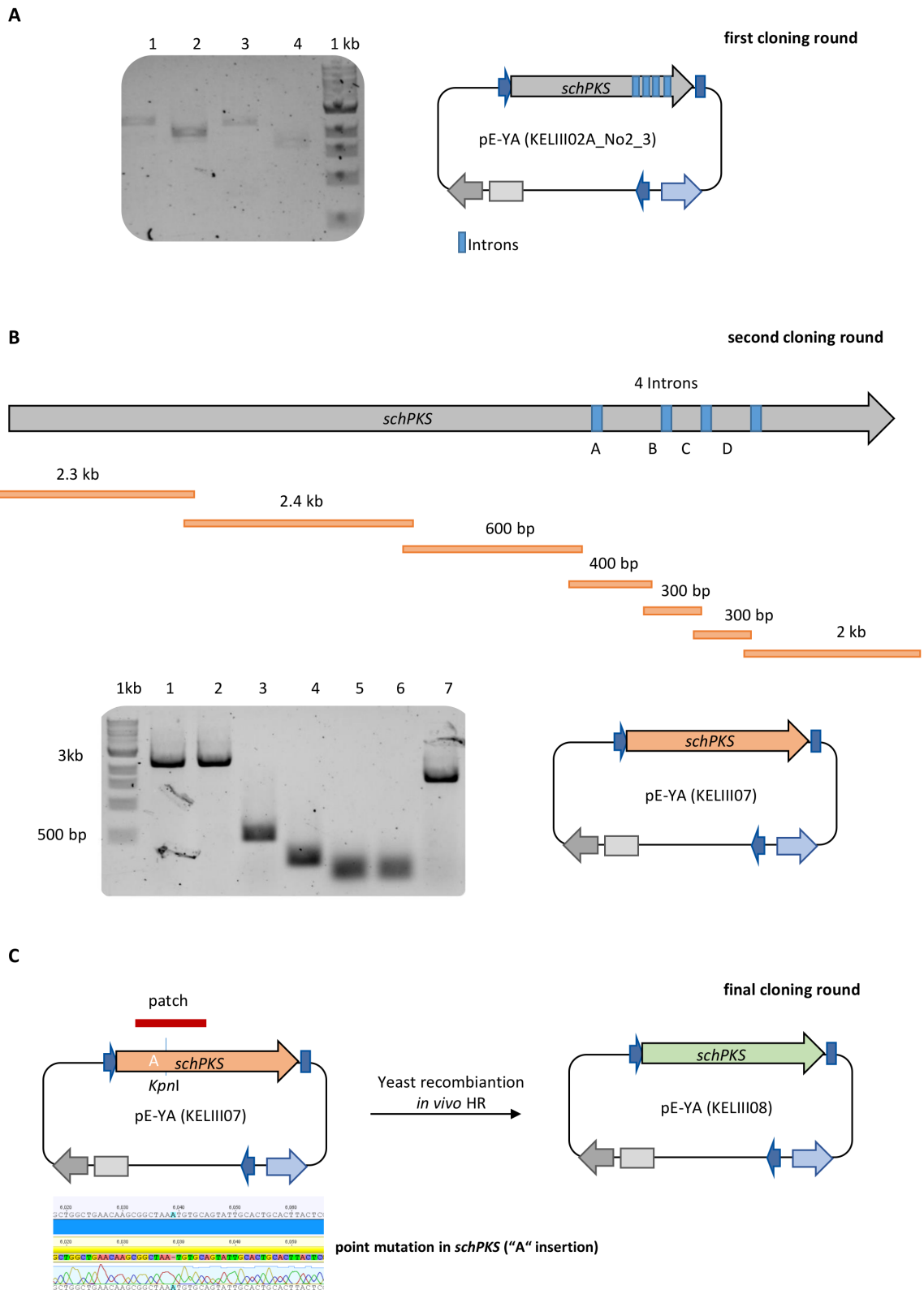
Table 4.1 Chemical shifts of 1.2 mg **174** in CDCl₃

Atom	Compound 174 (1.2 mg, CDCl ₃ , 500 MHz)			Adelin <i>et al.</i> (2011), CDCl ₃		
	δ_C / ppm	δ_H / ppm	<i>J</i> / Hz	δ_C / ppm	δ_H / ppm	<i>J</i> / Hz
1	210.5	-	-	211	-	-
2a	35.8	2.24	dddd (14.3, 4.8, 2.4, 0.7)	35.8	2.2	m
2b	35.8	2.6	ddt (14.3, 6.4, 0.7)	35.8	2.66	m
3a	32.5	1.85	ddd (14.1, 4.8, 2.2)	32.5	1.78	m
3b	32.5	2.15	dddd (14.0, 6.1, 3.5, 2.3)	32.5	1.94	m
4	71.0	4.13	q (2.7)	70.9	4.1	m
5	41.8	2.37	dtd (12.1, 6.6, 2.2)	41.7	2.36	q (7.3)
6	47.0	2.66	m	47.8	2.66	m
7	22.1	1.47	m	21.2	1.35	m
8	22.7	1.20, 1.76	m	22.2	1.35	m
9	25.2	1.40, 1.60	m	22.7	1.35	m
10	33.2	1.35	m	33	1.35	m
11	74.1	5.1	m	74.1	5.07	m
12	173.2	-	-	173.2	-	-
13	39.7	2.60, 2.76	m, d (5.4)	39.8	2.66	m
14	19.7	1.26	d (6.5)	19.7	1.23	d (6.4)

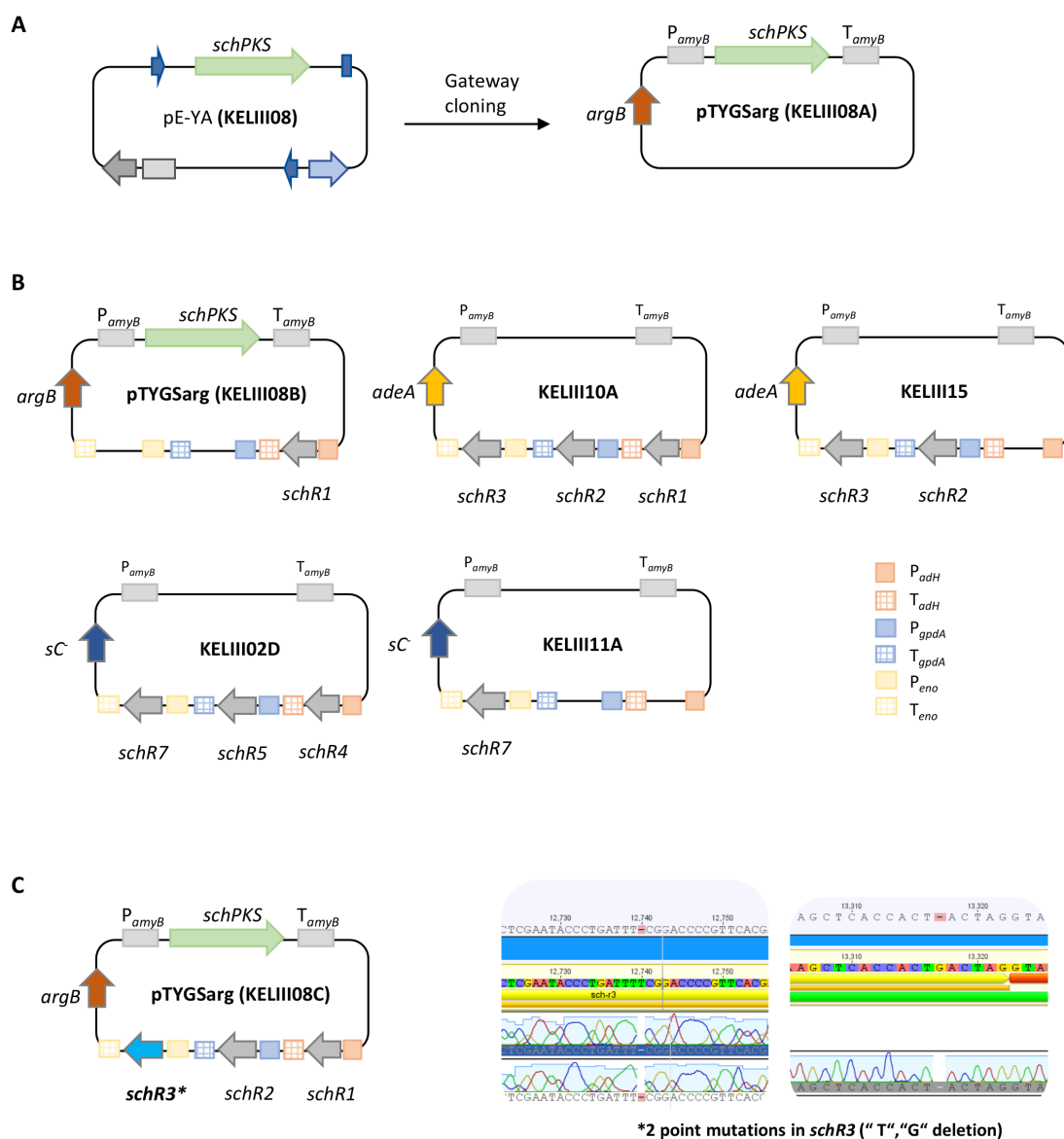
4.3.2 Yeast recombination, Gateway cloning and Transformation of *A. oryzae*

For successful expression of SCH genes from *Phomopsis sp.* CMU-LMA in the heterologous host *A. oryzae* NSAR1, RNA was isolated and transcribed to cDNA in order to eliminate all intron positions. The SCH producing strain was grown for 5 days in liquid PDB medium without shaking at 30 °C prior to RNA isolation. The synthesised cDNA was used as a template to amplify genes for yeast homologous recombination. As described previously for cloning of *mfpk2* (Section 2.6.1), the vector pE-YA was used to assemble the *schPKS* gene (8 kb), in four DNA fragments of around 2 kb each. Synthesised cDNA was used as a template for amplification of four parts of *schPKS*, resulting in faint bands on the gel (Scheme 4.7 A). Sequencing results of the constructed pE-YA-*schPKS* (KELIII02A(No2_3)) indicated that four predicted introns were still in place probably due to contaminated cDNA with gDNA. As RNA isolation was difficult and cDNA synthesis not very efficient for the large PKS transcript, the remaining introns were removed by a manual yeast recombination based strategy. The gene was amplified in 7 parts from vector KELIII02A(No2_3), in order to eliminate the introns by using oligonucleotides with a 30 bp overlap to the next exon positions (Scheme 4.7 B). Sequencing results of the constructed vector (KELIII07) revealed a point mutation ("A" insertion at position 2490 of *schPKS*) in the N-terminal part of *schPKS* (Fig. 4.7 C). An insertion would cause a frame shift and has to be repaired for successful expression. Again yeast homologous expression was used to eliminate the extra base by cutting plasmid KELIII07 with *KpnI* and repairing the vector with a patch amplified from KELIII02A(No2_3) (Scheme 4.7 C).

4.3 Results – Re-establish SCH biosynthesis in heterologous host *A. oryzae* NSAR1



Scheme 4.7 Cloning strategy of *schPKS* using three rounds of yeast recombination: **A**, first cloning round, amplifying four parts of *schPKS* from cDNA; **B**, second cloning round, manually removing 4 remaining introns by amplifying seven DNA fragments from template KELIII02A(No2_3) using yeast recombination; **C**, final cloning round, repairing of "A" insertion using *KpnI* digest and amplifying a patch from KELIII02A(No2_3), resulting in KELIII08.



Scheme 4.8 Constructed expression vectors containing *schPKS* and tailoring genes from *sch* BGC: **A**, Gateway cloning resulting in pTYGSarg–*schPKS* (KELIII08); **B**, constructed plasmids containing *sch* tailoring genes; **C**, vector KELIII08C occupying point mutations in *schR3*.

The repaired pE–YA vector containing *schPKS* (KELIII08) was checked by sequencing and used for *in vitro* recombination (Gateway cloning) with an expression vector for fungal transformation (pTYGSarg), resulting in pTYGSarg–*schPKS* (Scheme 4.8 A, KELIII08A). All tailoring genes (*schR1*, *schR2*, *schR3*, *schR4*, *schR5*, *schR7*) could be amplified from cDNA and after purification using a kit a suitable amount of DNA (40 ng/ml) was achieved for each gene in order to perform the yeast recombination. Constructed expression vectors for expression in *A. oryzae* NSAR1 are displayed in scheme 4.8 B and oligonucleotides and templates used for vector construction are summarised in chapter 6, table 6.11. Vectors which were constructed in this work, but not used in this study are displayed in figure 7.6 (Appendix).

Expression vector KELIII08C containing *schPKS*, *schR1*, *schR2* and *schR3* was found to possess two point mutations ("T", "G" deletion) in *schR3* (Scheme 4.8 C). These mutations would cause a frame shift and therefore leading to an inactive *schR3* gene in expression experiments. Nevertheless, the constructed vector was used for experiments lacking *schR3*.

In order to transform *A. oryzae* NSAR1 with combinations of constructed vectors (Fig. 4.8) a CaCl₂/PEG mediated protoplast protocol was used, which has already been described in section 2.6.1. Cotransformation of one, two or three vectors were carried out in a single transformation step using selection medium which was adapted to the vectors used in the individual transformation. Overall 8 combinations of genes from *sch* BGC on different fungal expression vectors were used to cotransform *A. oryzae* NSAR1, resulting in 8 strains occupying genes from *sch* BGC (Table 4.2). Genetic analysis showed the correct insertion of desired genes into *A. oryzae* (Appendix, Fig. 7.7).

Table 4.2 Overview of *A. oryzae* NSAR1 transformations with different combinations of genes from *sch* BGC; TH: thioesterase, Hyd: Hydrolase, Oxi: Oxidase

Exp. No	Exp. ID	Selection marker	Transformed expression vector(s)	Genes <i>sch</i>						
				<i>pks</i> PKS	<i>r1</i> TH	<i>r2</i> P450	<i>r3</i> P450	<i>r4</i> Hyd	<i>r5</i> Oxi	<i>r7</i> FMO
1	KELIII12	arg, ade, met	KELIII08A, KELIII10A, KELIII02D	X	X	X	X	X	X	X
2	KELIII08A	arg	KELIII08A	X	-	-	-	-	-	-
3	KELIII08B	arg	KELIII08B	X	X	-	-	-	-	-
4	KELIII08C	arg	KELIII08C	X	X	X	-	-	-	-
5	KELIII13	arg, ade	KELIII08A, KELIII10A	X	X	X	X	-	-	-
6	KELIII14	arg, ade, met	KELIII08A, KELIII10A, KELIII11A	X	X	X	X	-	-	X
7	KELIII16	arg, ade, met	KELIII08A, KELIII15, KELIII02D	X	-	X	X	X	X	X

4.3.3 Expression of whole SCH cluster in *A. oryzae*

In an initial experiment (exp. 1, table 4.2), all genes from the *sch* BGC were coexpressed in *A. oryzae* NSAR1. Gene *schR6* was not included, as it was predicted to encode a transporter. Cotransformation of three expression vectors, containing 7 genes (*schPKS*, *schR1*, *schR2*, *schR3*, *schR4*, *schR5*, *schR7*), resulted in 16 transformants. For analysis of their secondary metabolite profile 11 of the transformants were sub-cultured in P_{amyB} induction medium (DPY). The obtained extracts were submitted to LCMS and the chromatograms were analysed. From previous biotransformation experiments it is known, that the expected product would not be **3**, but its reduced analogue **174**, eluting at t_R = 5.5 min (Section 4.3.1).

4.3 Results – Re-establish SCH biosynthesis in heterologous host *A. oryzae* NSAR1

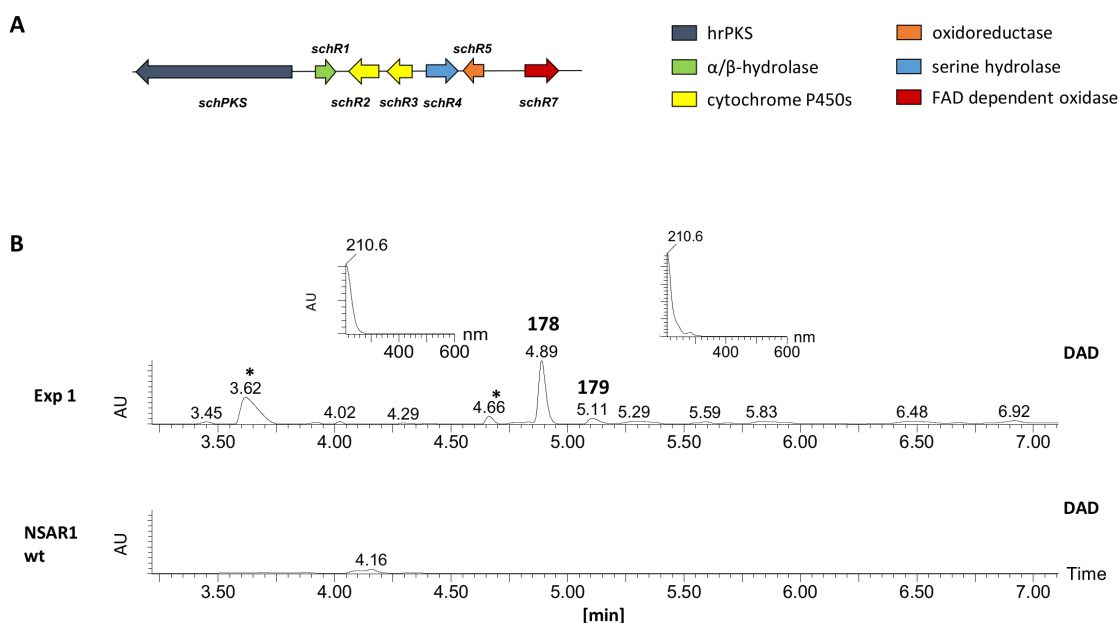


Figure 4.7 Chemical analysis of transformant containing whole *sch* BGC (exp. 1, table 4.2): **A**, Coexpression of 7 genes (*schPKS*, *schR1*, *schR2*, *schR3*, *schR4*, *schR5*, *schR7*); **B**, LCMS trace of exp. 1 compared to NSAR1 wt control, DAD chromatogram with UV spectrum of **178** eluting $t_R = 4.9$ min and **179** eluting at $t_R = 5.1$ min.

In the extract of the transformant containing the whole *sch* BGC, a new major peak eluting at $t_R = 4.9$ min and a minor compound eluting at $t_R = 5.1$ min were observed compared to a wt control (Fig. 4.7). Both compounds **178** and **179** were shown to possess a similar UV spectrum as reduced SCH **174** with a maximum absorption at 210 nm (Fig. 4.7).

Further HRMS analysis showed that major compound **178** has the identical mass and thus molecular formula as **174** ($\text{ES}^- \text{C}_{14}\text{H}_{21}\text{O}_4$ calc. 253.1440, measured 253.1440). Minor compound **179** showed the same HRMS as SCH, accordingly two hydrogen atoms less than **174** ($\text{ES}^- \text{C}_{14}\text{H}_{19}\text{O}_4$ calc. 251.1283, measured 251.1286, Fig. 4.8). More sophisticated analysis of the mass fragmentation pattern revealed loss of three water molecules for compound **178** (Fig 4.8). A similar pattern was also observed for **179**, showing significant fragments for losses of two water molecules (Fig 4.8). LCMS analysis indicated that the new compounds are both analogues of **174**, but not identical, as the retention times vary between 0.4 to 0.6 min.

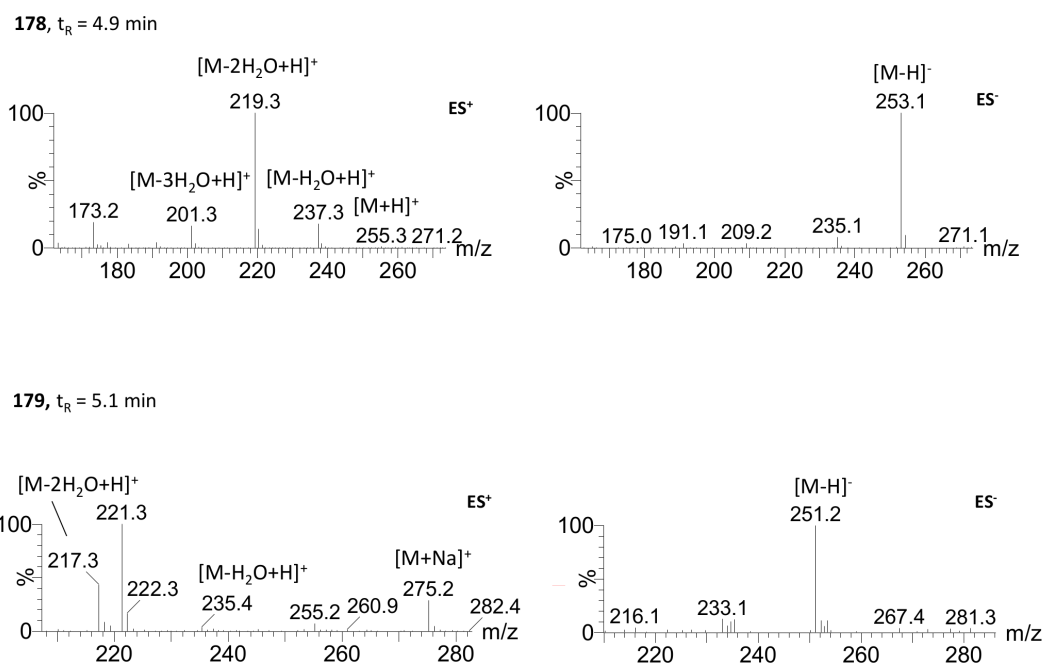
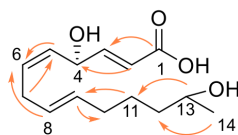


Figure 4.8 LCMS (ES^+ and ES^-) and HRMS analysis of new compounds obtained from extractes of transformant containing whole *sch* BGC (exp. 1, table 4.2).

Structure elucidation of compound 178 and 179

For structure elucidation of major compound **178**, combined analytical extracts of producing transformants were submitted to preparative LCMS analysis and 1 mg of pure **178** was isolated. Subsequent NMR analysis in deuterated DMSO yielded 1H and ^{13}C chemical shifts for C-1 to C-11 corresponding to compound **172** (Fig. 4.9 A). Compound **172** has previously been isolated from $\Delta schR3$ strain by Francesco Trenti.⁸⁵ However, compound **172** ($m/z = 236.3$ g/mol) has a lower mass than **178** ($m/z = 254.3$ g/mol), corresponding to loss of one molecule of water (Fig. 4.9 B).

A



178
254.3 g/mol

Atom	δ_c / ppm	δ_c / ppm	δ_H / ppm	δ_H / ppm	J / Hz	$^1\text{H}-^1\text{H}$ COSY	HMBC
		172 (FT)	172 (FT)				
1	167.4	166.3	-	-	-	-	2, 3
2	120.0	117.5	5.94, dd	5.86	dd, 15.4, 1.6	3	1, 3, 4
3	149.1	149.6	7.18, dd	6.67	dd, 15.4, 4.9	2, 4	1, 2, 4, 5
4	65.9	68.3	5.26, m	4.99	d, 6.2	3, 5	2, 3, 5, 6
5	131.0	131.5	5.51, m	5.29	m	6	3, 4, 6, 7
6	128.8	128.6	5.54, m	5.42	m	5, 7	4, 5, 7
7	30.4	32.1	2.82, m	2.79	m	6, 8	5, 6, 8, 9
8	130.9	132.1	5.55, m	5.39	m	7, 9, 10	6, 7, 9
9	127.5	126.6	5.43, m	5.42	m	10	10, 11
10	32.1	32.9	1.91, 2.08, m	1.95	m	9, 11a, 11b, 12	9, 11
11a	25.3	24.4	1.78, m	1.36	m	10, 12	9, 10
11b	25.3	24.4	1.19, m	1.28	m	10, 12, 13	9, 10
12	38.6	34.4	1.74, 1.51, m	1.28	m	10, 11	13, 14
13	65.6	71.8	4.82, m	3.56	m	12, 14	10, 11, 12, 14
14	23.7	20.1	1.27, d	1.02	d, 6.1	13	12, 13

B

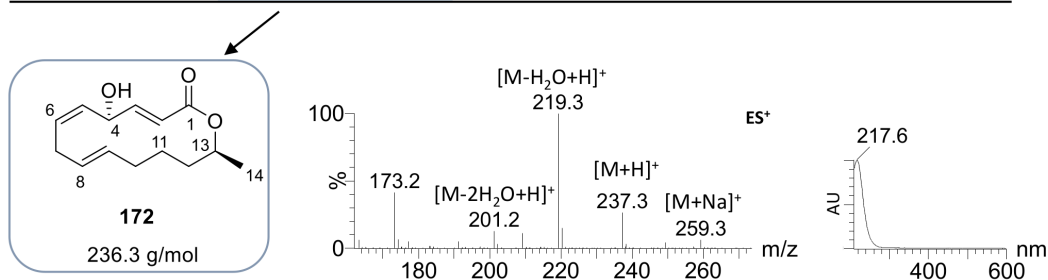


Figure 4.9 Structure elucidation of **178**: **A**, structure of **178** (1 mg, DMSO, 500 MHz) with key HMBC correlations and assigned NMR signals in comparison with **172** (1 mg, CDCl_3 , 400 MHz), isolated from $\Delta schR3$ strain (data provided by Francesco Trenti)⁸⁵; **B**, structure of **172** with corresponding ES^+ and UV spectrum.

Chemical shifts were assigned according to previously isolated **172**. However, comparison of the ^1H NMR signals identified changes in chemical shifts of the protons at C-13 and C-14. The chemical shift of methyl group protons at C-14 showed an upfield shift of 0.25 ppm. Furthermore, a larger shift was identified for protons at C-13, which shifted around 1.3 ppm upfield compared to **172**. This could indicate that the lactone is not closed in this analogue of SCH, but opened to the corresponding free acid.

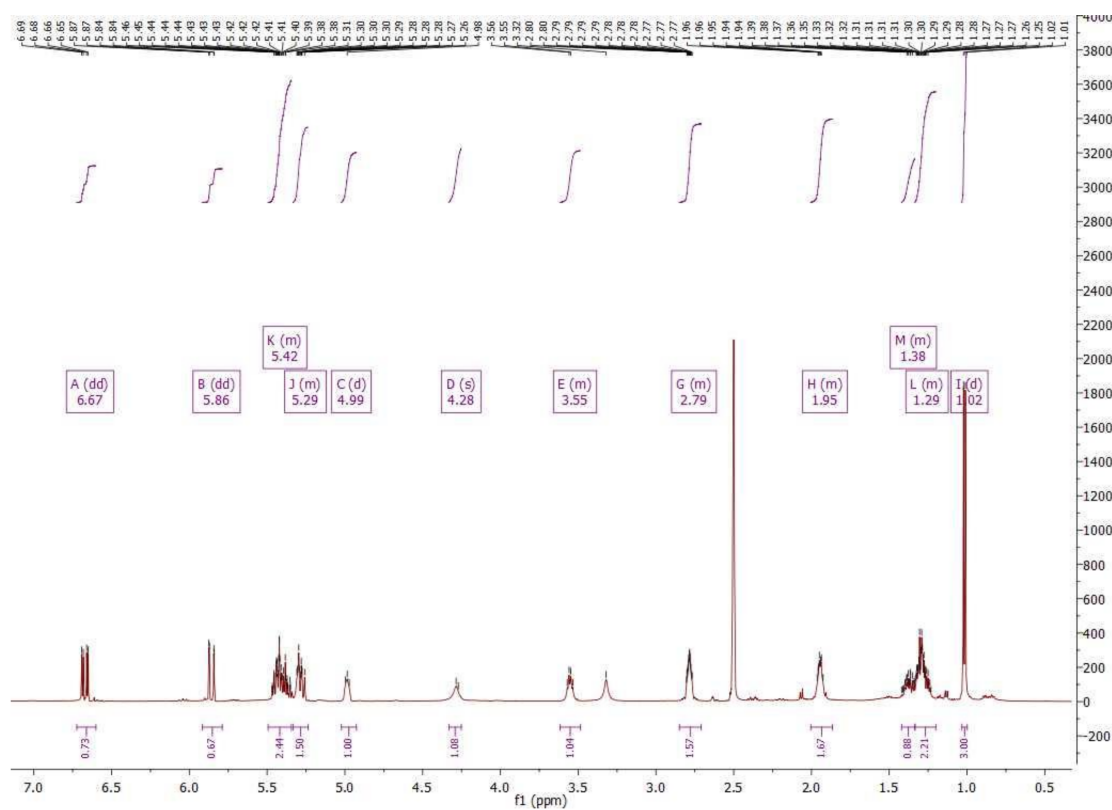


Figure 4.10 ^1H NMR of compound **178** in DMSO-d_6 (500 MHz) referenced to DMSO-d_6 .

The olefin at C-2/C-3 was shown to be *trans*, as the corresponding doublets of doublets at 5.86 and 6.67 ppm appeared to have the same coupling constant of 15.4 Hz (Fig. 4.9 A + 4.10). Corresponding signals of H-5, H-6, H-8 and H-9 were superimposed in the region between 5.29 and 5.42 ppm. In order to determine the stereoisomery of the two olefins C-5/C-6 and C-8/C-9 of **178**, ^1H -decoupling experiments were performed. H-7 decoupling allowed the determination of the coupling constant between H-5 and H-6, which was found to be 10.8 Hz, suggesting a *cis* olefin (Fig. 4.11). This is in agreement to the J_{5-6} (11.3 Hz) of **172** (2*E*,5*Z*,8*Z*, Fig. 4.9 A).⁸⁵

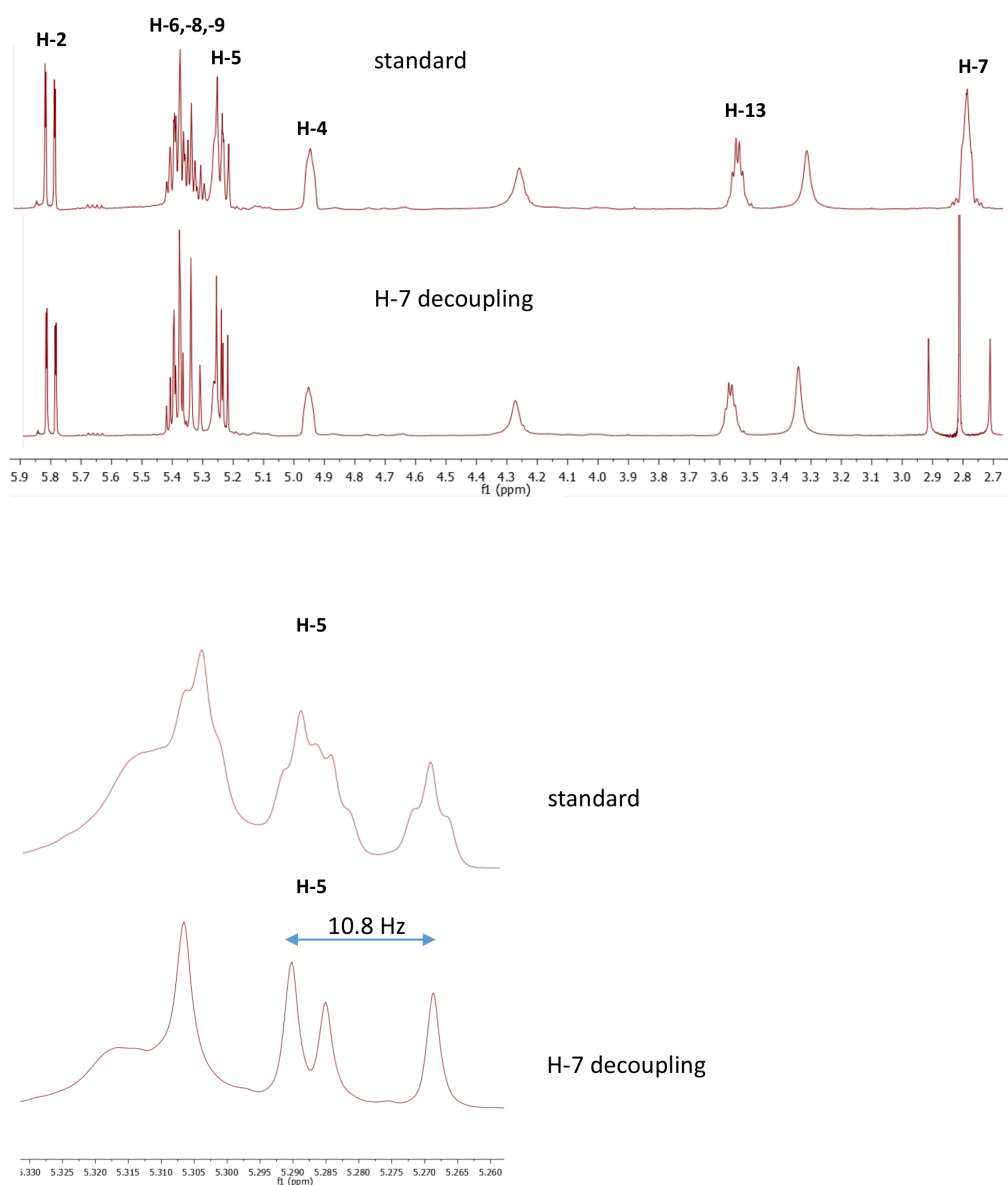


Figure 4.11 Decoupling of H-4 in 178 ^1H -NMR (500 MHz).

In order to determine the stereoisomery of the C-8/C-9 olefin, H-7 and H-10 decoupling was performed in two separate experiments. NMR decoupling measurements on a Bruker DRX 600 MHz allowed the determination of the coupling constant between H-8 and H-9, which was found to be 15.4 Hz, suggesting a *trans* olefin (Fig. 4.12), which is in agreement to the J_{8-9} (15.1 Hz) of **172** (*2E,5Z,8Z*, Fig. 4.9 A).⁸⁵

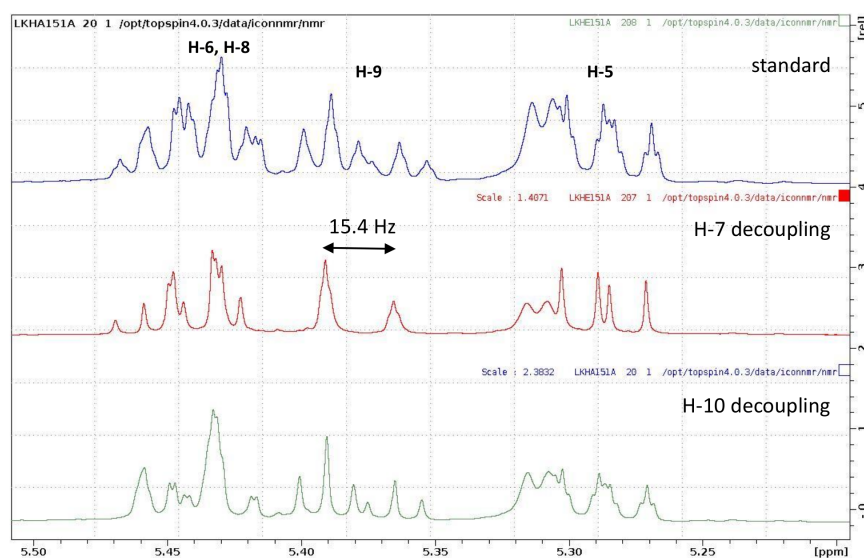


Figure 4.12 Decoupling of H-7 (red) and H-10 (green) in 178 ^1H -NMR (600 MHz)

Further evidence that **178** is the free acid and not the lactone were given by using pH indicator Bromocresol green. While **172** spotted on a TLC plate was not turning yellow upon exposure of pH indicator Bromocresol green, **178** showed a yellow spot indicating an acid (Fig. 4.13).

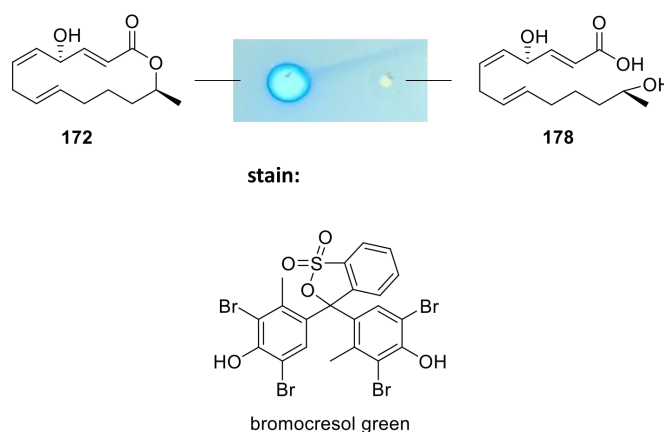
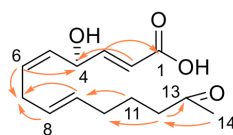


Figure 4.13 TLC plate with one drop of each **178** and **172**, stained with pH indicator Bromocresol green - a colour change to yellow indicating an acid.

Together with the NMR data, these results strongly proposed structure **178**, which is in accordance with the HRMS result. Nevertheless, expected reduced analogue of SCH **174** was not observed in the DAD trace, as no peak was detected around 5.5 min.



179
252.3 g/mol

Atom	δ_C / ppm	δ_H / ppm	Multiplicity (J, Hz)	1H - 1H COSY	HMBC
1	167.2	-	-	-	2, 3
2	119.6	5.86	dd, 15.6, 1.6	3	1, 3, 4
3	149.6	6.67	dd, 15.6, 4.9	2, 4	1, 2, 4, 5
4	65.8	4.99	m	3, 5	2, 3, 5, 6
5	131.0	5.28	ddt, 10.9, 8.2, 1.5	4, 6	3, 6, 7
6	128.7	5.44	m	5, 7	4, 5, 7
7	30.4	2.79	m	6	6, 8, 9
8	130.2	5.39	m	7, 9	7, 9
9	128.1	5.39	m	8, 10	7, 8, 10
10	30.3	1.93	m	9, 11	8, 9, 10, 11
11	23.0	1.50	p, 7.4 Hz	10, 12	9, 10, 12, 13
12	42.0	2.39	t, 7.3 Hz	11	10, 11, 13
13	208.3	-	-	-	11, 12, 14
14	29.7	2.06	s	-	12, 13

Figure 4.14 Structure elucidation of **179** (1.9 mg, DMSO, 500 MHz) with key HMBC correlations and assigned NMR signals.

For structure elucidation of minor compound **179**, possessing the same nominal mass as SCH (m/z 252), the producing transformant was re-grown in 10 x 100 ml DPY medium and extracted after 7 days. The crude extract was submitted to preparative LCMS analysis and 1.9 mg of pure **179** was isolated. Subsequent NMR analysis in deuterated DMSO (500 MHz) revealed almost identical shifts compared to those of **178** (Fig. 4.14) suggesting a similar structure and configuration pattern of the olefins. However, position C-13 was identified as quaternary C-atom and the corresponding chemical shift of C-13 (δ 208.3) indicated a carbonyl group. This assumption would also correspond to HRMS analysis of **179** ($C_{14}H_{19}O_4$ calc. 251.1283, measured 251.1286).

The DAD chromatogram of *A. oryzae* NSAR1 transformant containing genes for the whole *sch* BGC, revealed two new compounds. Major compound **178**, as well as minor compound **179**, are congeners of expected reduced form of SCH **174**. Nevertheless, the characteristic lactone structure element was not observed for either compound.

Detailed analysis of mass traces

More detailed analysis showed a minor compound with the nominal mass of 254 eluting at $t_R = 5.5$ min in the ES^+ BPI chromatograms of three transformants containing all genes of *sch* BGC, which corresponds to **174** (Fig. 4.15). Also the minor reduced SCH compound **177** previously detected in

the biotransformation experiment with SCH, was produced in traces ($t_R = 5.0$ min) (Fig. 4.15). Both peaks were shown to have the characteristic mass fragmentation pattern (namely loss of one and two H_2O), strongly suggesting them to be expected reduced congeners of SCH, **174** and **177** (Fig. 4.16).

Structure elucidation of compound **177** by NMR was not possible due two very low titres. Further HRMS analysis was also prevented for the same reason. Nevertheless, these results indicated that the 7 genes (*schPKS*, *schR1*, *schR2*, *schR3*, *schR4*, *schR5*, *schR7*), defined as the whole *sch* BGC, expressed in heterologous host *A. oryzae*, led to trace amounts of reduced forms of SCH. This confirms the cluster to be responsible for production of **3** and showed that no other genes are essential to produce SCH.

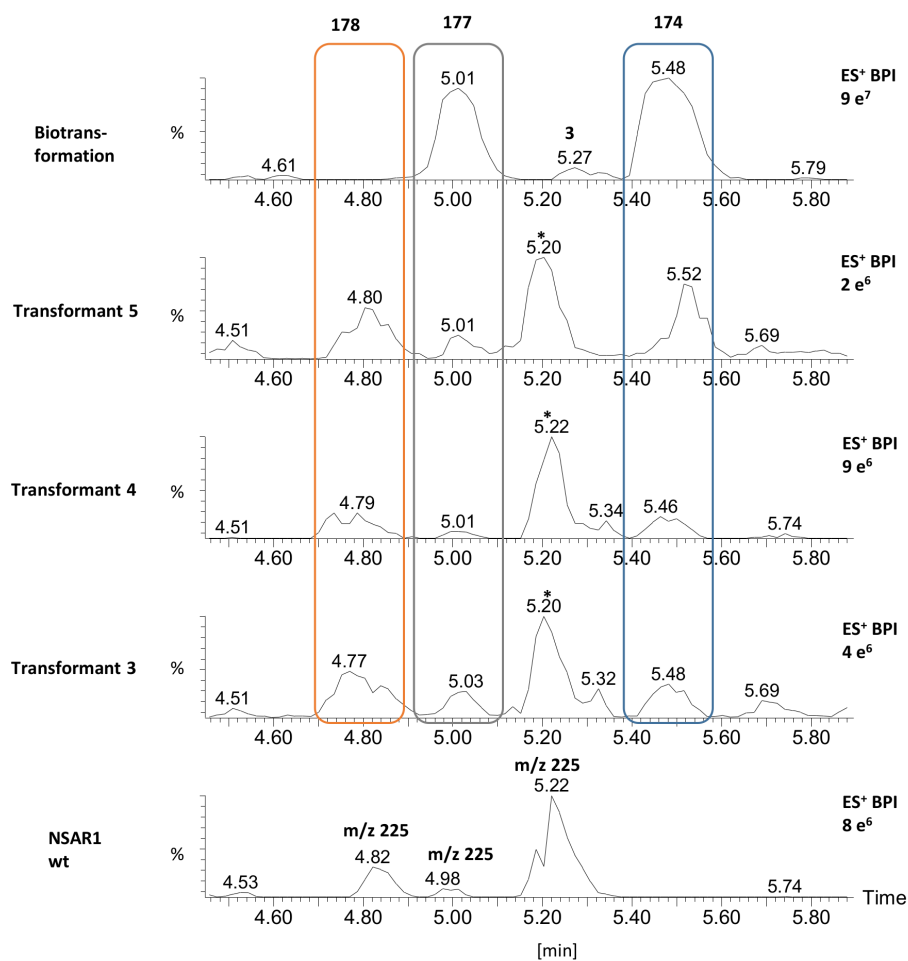
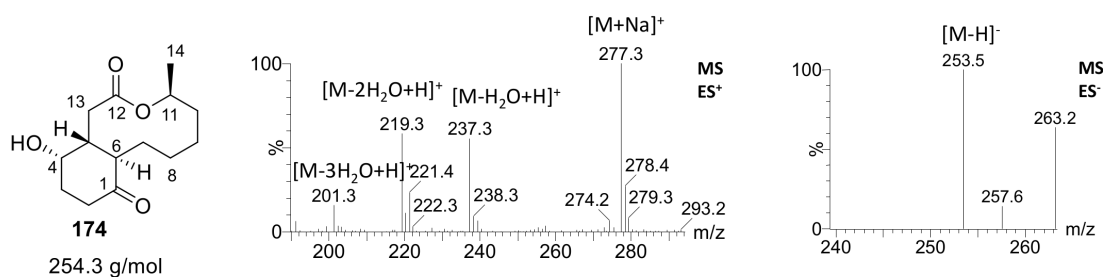


Figure 4.15 Chemical analysis of *A. oryzae* NSAR1 + 7 genes, exp. 1 (Table 4.2), transformant 3, 4, 5, in comparison to *A. oryzae* NSAR1 wt and biotransformation experiment feeding SCH to *A. oryzae* NSAR1 wt, ES⁺ traces.

4.3 Results – Re-establish SCH biosynthesis in heterologous host *A. oryzae* NSAR1

Compound **174** produced by *A. oryzae* NSAR1 + 7 genes (transformant 4), eluting at 5.5 min



Compound **177** produced by *A. oryzae* NSAR1 + 7 genes (transformant 4), eluting at 5.0 min

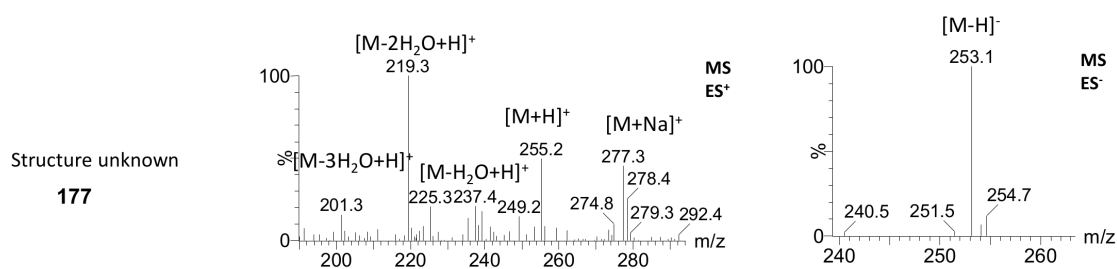


Figure 4.16 MS analysis of **174** and **177** produced by *A. oryzae* NSAR1 + 7 genes (transformant 4).

However, just trace amounts of SCH analogue **174** were observed from expression transformants. The major compound **178** was shown to be an acyclic derivative of **172**. As both intermediates accumulated upon disruption of *schR3* and *schR7* were found to be lactones suggesting that **178** and minor compound **179** were shunt products. It is possible that endogenous genes in *A. oryzae* led to hydrolysis of an early pathway product in SCH biosynthesis, probably **172** or a related compound, yielding **178**. The shunt product would then stop the pathway to SCH and adventitious oxidation of the hydroxyl group at C-13 to the ketone can occur, probably catalysed by a monooxygenase from the host organism.

Another possibility is that the programming of SchPKS resulted in no ketoreduction (by KR domain) in the first cycle of β -processing yielding a ketone function at C-13. Tang and coworkers also found that Bref-PKS revealed products with different reduction pattern *in vitro* due to limiting amounts of NADPH.²⁰⁵ A minor PKS product unlactonised 7-dehydrobrefeldin A acid **180** was shown to possess a keto group instead of an alcohol group at C-15 (analogue to C-13 in **3**, Scheme 4.17).²⁰⁵ 7-dehydrobrefeldin A acid **180** was earlier isolated from the endophytic fungal strain *Cylindrocarpon obtusisporum* together with brefeldin A **153**.²¹⁸

4.3 Results – Re-establish SCH biosynthesis in heterologous host *A. oryzae* NSAR1

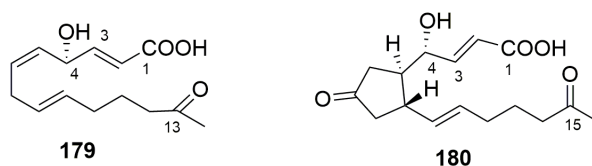
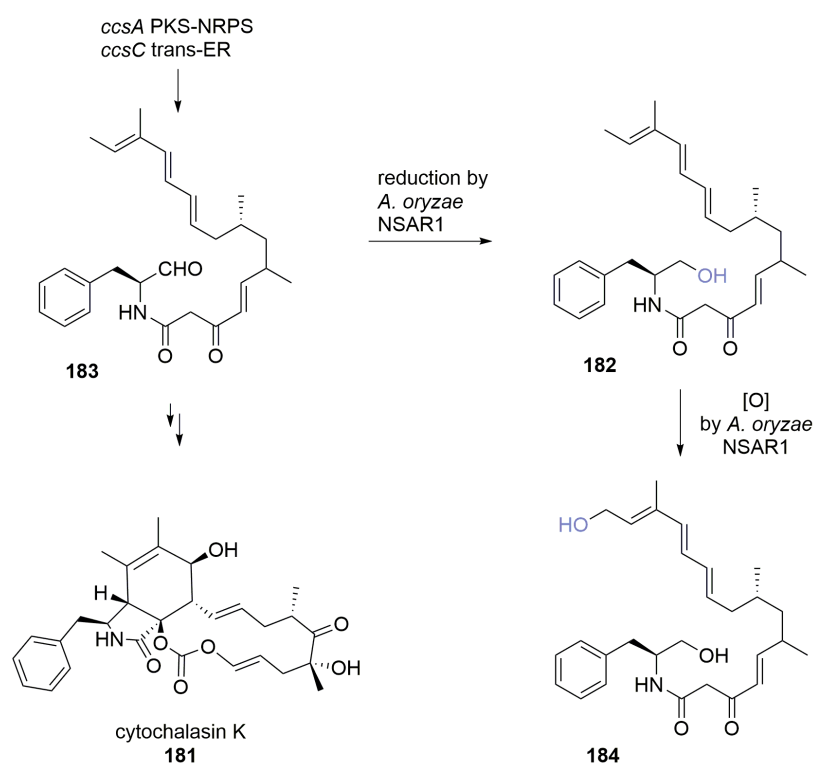


Figure 4.17 Comparison of **179** observed from *in vivo* expression experiments with *A. oryzae* NSAR1 and 7-dehydrobrefeldin A acid **180** produced by Bref-PKS.

Oikawa and coworkers previously reported unexpected side reactions by the heterologous host *A. oryzae* NSAR1 in studies on fungal polyketide cytochalasin K **181**.²¹⁹ They were able to isolate a linear intermediate analog **182** from the *A. oryzae* transformant harboring polyketide PKS-NRPS and *trans*-acting ER genes *ccsA* and *ccsC*, which is most likely derived by a two-step reduction of **183** by reductases in *A. oryzae* (Scheme 4.9).²²⁰ Moreover determination of a second by-product **184** suggesting oxidation of **182** by an unknown monooxygenase in *A. oryzae* (Scheme 4.9).²¹⁹



Scheme 4.9 Proposed biosynthetic pathway of cytochalasin and undesired reduction and random oxidation by heterologous host *A. oryzae* NSAR1 while expression of PKS-NRPS encoding gene *ccsA* and *trans*-acting ER encoding gene *ccsC*.²¹⁹

Similar oxidation products have been reported in heterologous production of desmethylbassianin²²¹ and in heterologous expression of the avirulence gene ACE1 from the fungal rice pathogen *Magnaporthe oryzae* in *A. oryzae*.²²² All observed oxidations, probably caused by endogenous enzymes in *A. oryzae*, occurred at the terminal end of linear polyketide precursors, as also shown in our experiments, where **178** was oxidised to **179**.

4.3.4 Expression of early genes in *A. oryzae*

The results of coexpression of 7 genes from the cluster showed that *schPK* was cloned correctly. Previous knockout studies could not reveal the PKS product, which was predicted to be benzoquinone 160 or a related species.⁸⁵

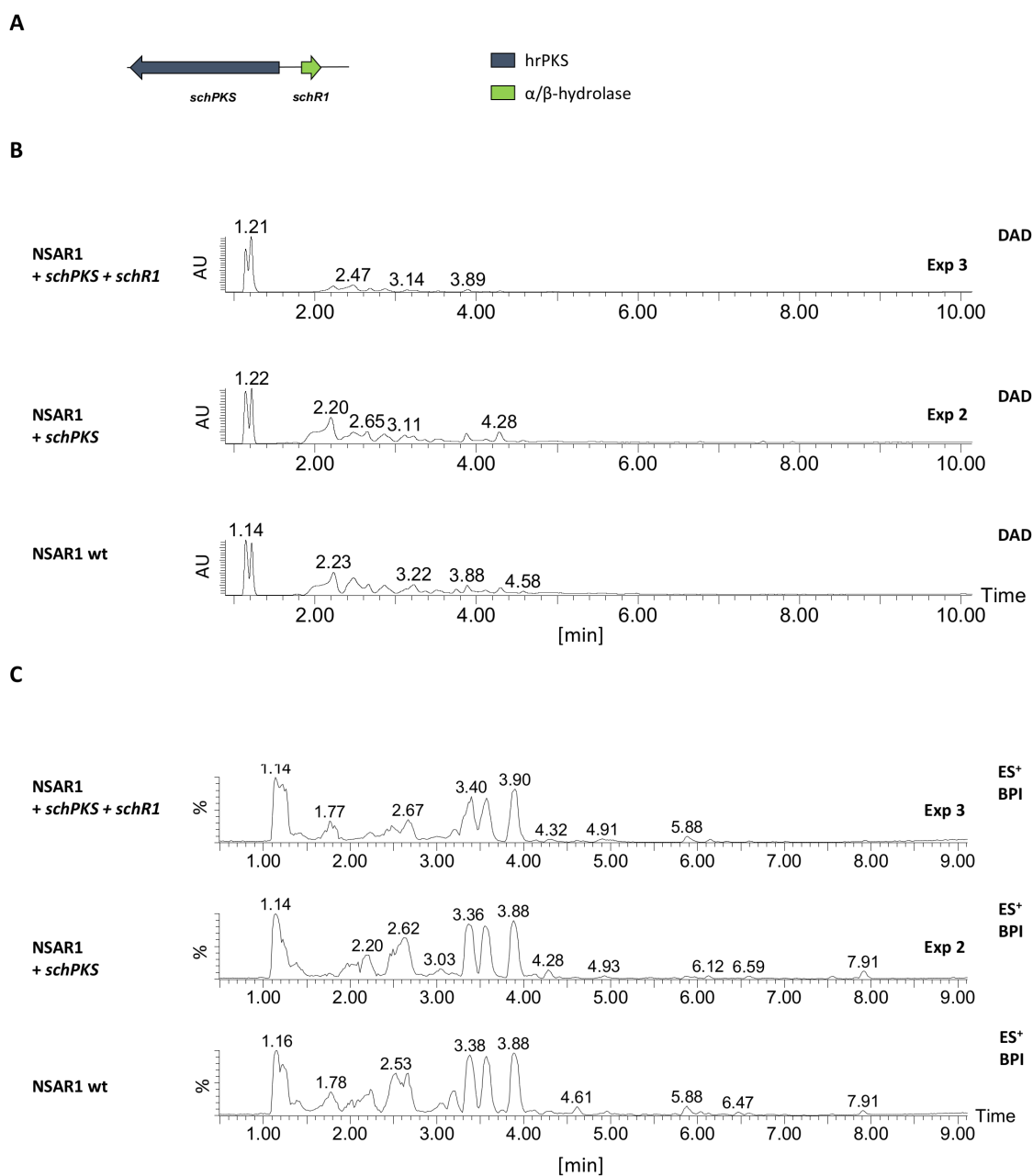


Figure 4.18 Chemical analysis of extract of expression exp. 2 (table 4.2, *schPKS*) and exp. 3 (table 4.2, *schPKS* + *schR1*) compared to *A. oryzae* NSAR1 wt control: **A**, coexpression of *schPKS* and *schR1*; **B**, DAD analysis; **C**, ES⁺ traces.

Expression of *schPKS* alone (Table 4.2, exp. 2) and coexpression with hydrolase encoding gene *schR1* in analogy to brefeldin **153** biosynthesis²⁰⁵ (Table 4.2, exp. 3) in heterologous host *A. oryzae* NSAR1 should provide the first intermediate of the pathway to SCH **3**.

In total, 7 transformants containing *schPKS* in their genome and 9 transformants containing *schPKS* and *schR1* in their genome were grown, extracted and submitted to LCMS analysis. Neither the DAD chromatograms nor the ES⁺ traces showed new peaks compared to NSAR1 wt control (Fig. 4.18). This indicates that no PKS product could be detected using standard LCMS methods. Previous results showed that SCH analogues can be produced by expression experiments, neglecting the possibility of an error in the PKS sequence. The PKS product is possibly degraded in the heterologous host.

4.3.5 Re-constructing the biosynthesis of SCH in *A. oryzae*

Cotransformation of cytochrome P450 encoding gene *schR2* with *schPKS* and *schR1* resulted in 10 transformants. Vector KELIII08C (exp. 4, table 4.2) was used to transform *A. oryzae* NSAR1, containing a mutated *schR3* gene. As the mutations lead to a frame shift, expression of *schR3* cannot result in an active protein. All 10 transformants were grown in DPY and extracted after 7 days and submitted to LCMS analysis. The chemical analysis revealed a similar metabolic profile as observed for the whole cluster expression. Compound **178** was again found to be the major compound ($t_R = 4.9$ min) and **179** the minor compound (Fig. 4.19, exp. 4). Traces of reduced SCH **174** were not observed in the extract of this transformants. Adding the second P450 encoding gene *schR3* to the expression system showed no production of any SCH metabolite (Fig.4.19, exp. 5). In total, 6 extracts of transformants containing *schPKS*, *schR1*, *schR2*, *schR3* were analysed. Upon addition of oxidase encoding gene *schR7* compounds **178** and **179** were detected again, but in lower titres (Fig. 4.19, exp. 6).

Together these results show that SchPKS, the partnering hydrolase SchR1 and one of the two P450s SchR2 are sufficient to produce **172** analogue **178**, which is probably a shunt product in *A. oryzae*. Previous knockout studies led to the disruption of *schR3* and subsequent accumulation of compound **172**, indicating hydroxylation at C-4 by SchR2.⁸⁵ In these expression experiments in *A. oryzae*, **178** is produced in high amounts no matter how many tailoring genes were added to the expression system (e.g. *schR3*, *schR7*), indicating a pathway interruption by *A. oryzae* endogenous enzymes, which possibly hydrolyse **172** to give the observed acid **178**. Further adventitious oxidation catalysed by an enzyme in *A. oryzae* would yield **179**.

4.3 Results – Re-establish SCH biosynthesis in heterologous host *A. oryzae* NSAR1

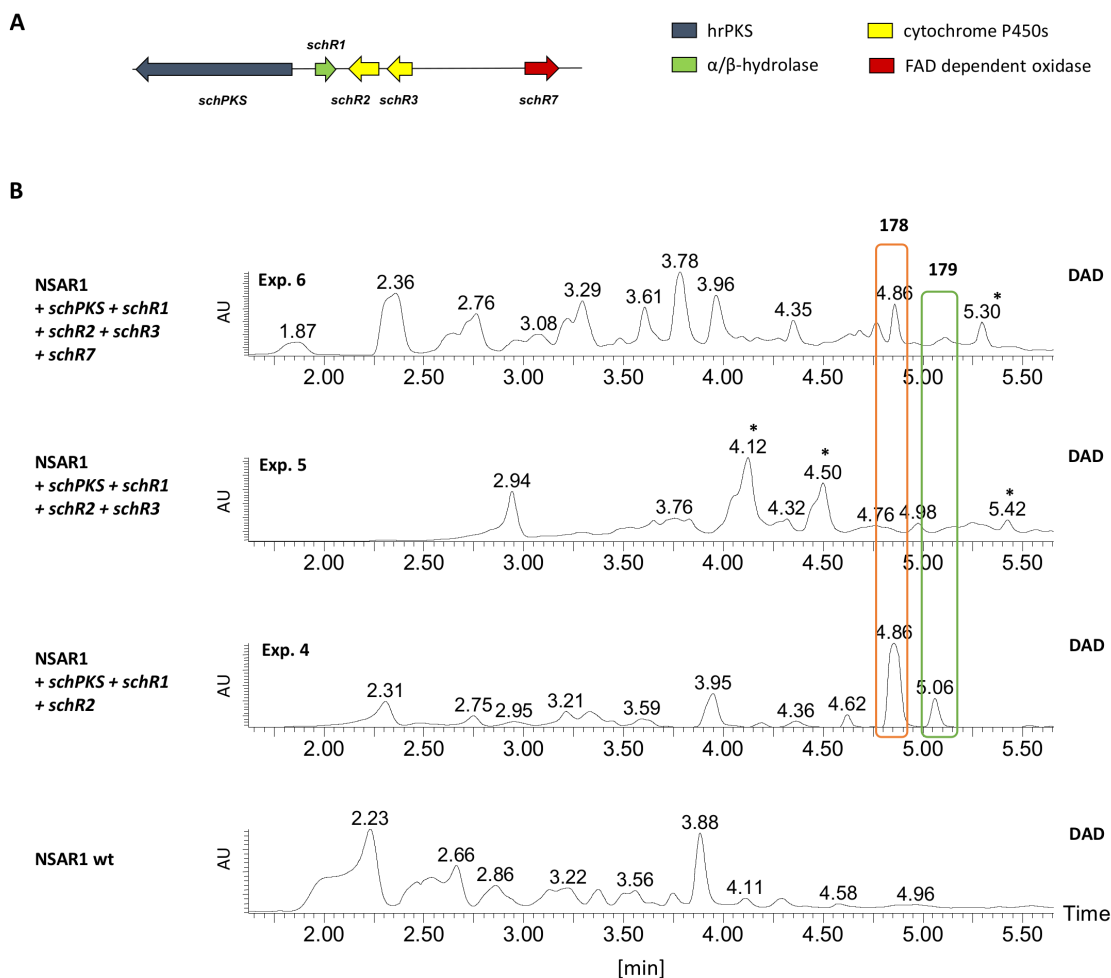


Figure 4.19 Chemical analysis of extract of expression exp. 4, exp. 5, exp. 6 (table 4.2, compared to *A. oryzae* NSAR1 wt control: **A**, expressed genes in exp. 4, 5 and 6; **B**, DAD analysis; * marked peaks correspond to unrelated compounds).

4.3.6 Expression of whole SCH cluster without hydrolase SchR1

It is also possible that over-expressed gene *schR1* encoding for the partnering hydrolase led to unwanted over-hydrolysis of SCH pathway intermediate 172. In order to overcome this possibility *schR1* was removed from the expression system and in the same way assuming that off-loading of the PKS product would be catalysed from an *A. oryzae* hydrolase. Co-expression of *schPKS*, *schR2*, *schR3*, *schR4*, *schR5* and *schR7* (exp. 7, table 4.2) in *A. oryzae* resulted in no SCH intermediate produced (Fig. 4.20). This experiment showed that hydrolase SchR1 is essential for production of SCH or intermediates from the pathway.

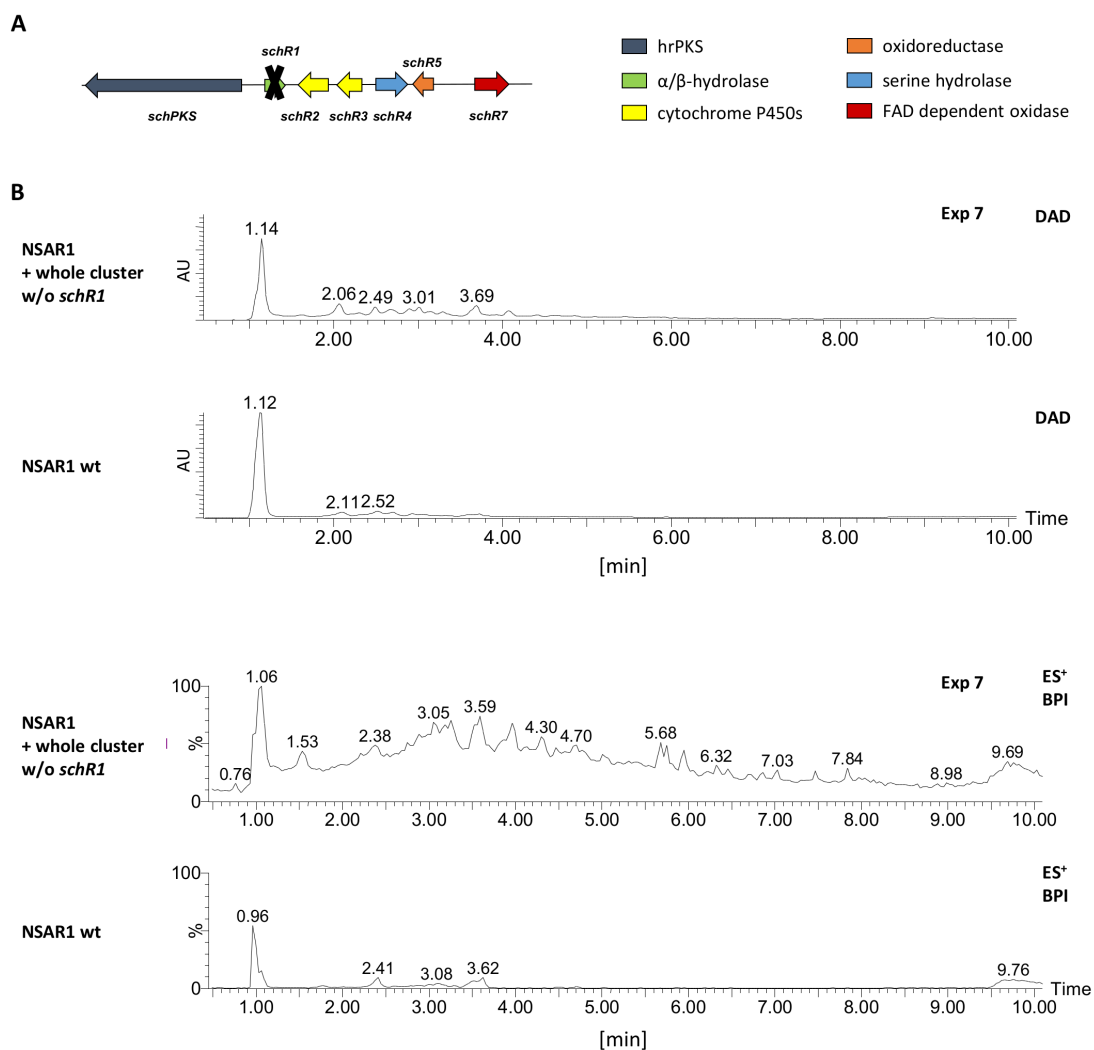
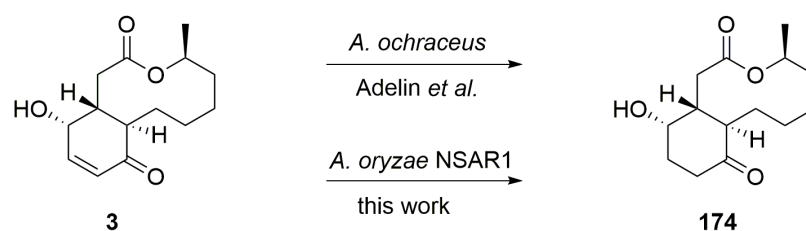


Figure 4.20 A, coexpression of *schPKS*, *schR2*, *schR3*, *schR4*, *schR5* and *schR7*; B, chemical analysis of extract of *A. oryzae* NSAR1 *sch* BGC missing *schR1* (exp. 7, table 4.2, compared to *A. oryzae* NSAR1 wt control, DAD chromatograms and ES⁺ traces.

4.4 Conclusion and Outlook

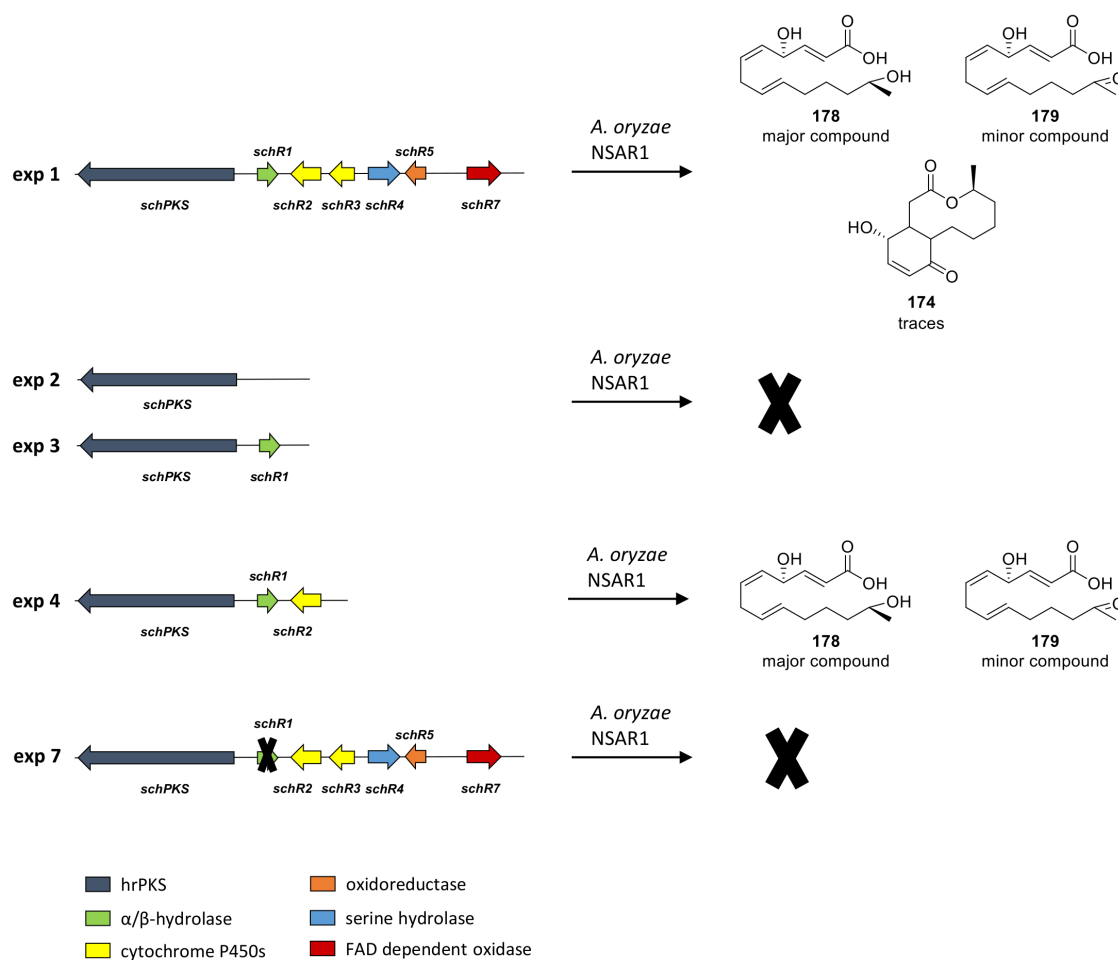
Biotransformation experiments showed that our heterologous expression host, *A. oryzae* NSAR1, reduced fed SCH 3 to 174 and a minor compound 177 through biocatalysis, as previously observed in *A. ochraceus* by Adelin and coworkers (Scheme 4.10).²¹⁷ Due to very low titres of 177 its structure could not be elucidated. Nevertheless these results indicated that heterologous expression of genes from *sch* BGC would lead to 174 instead of 3 in our host *A. oryzae* NSAR1.



Scheme 4.10 Schematic overview of biotransformation of SCH in *Aspergillus* species: *A. ochraceus*²¹⁷ and *A. oryzae* NSAR1, yielding reduced congener **174**.

For successful expression of SCH genes from **3** producing fungus *Phomopsis* sp. CMU-LMA in the heterologous host *A. oryzae* NSAR1, RNA was isolated and transcribed into cDNA in order to eliminate all intron positions. Yeast recombination based cloning was used to assemble the 8 kb *schPKS* gene. Corresponding tailoring genes of the cluster were cloned into fungal expression plasmids in different set of genes using the same method (Fig. 4.8). Overall 8 combinations of genes from *sch* BGC on various fungal expression vectors were used to transform *A. oryzae* NSAR1, applying a CaCl₂/PEG mediated protoplast protocol (Table 6.11). Genetic analysis *via* PCR of resulting strains showed the desired set of genes integrated into the genome of the heterologous host (Fig. 7.7).

Coexpression of all predicted biosynthetic genes from *sch* BGC (*schPKS*, *schR1*, *schR2*, *schR3*, *schR4*, *schR5*, *schR7*) in *A. oryzae* NSAR1 (exp. 4, table 4.2), resulted in production of trace amounts of biotransformed SCH **174**, as well as the minor compound **177** (Scheme 4.11, exp. 1). This was shown by the same UV, mass fragmentation pattern and retention time compared to a standard of **174**. Further evidence by isolation of **177** and NMR analysis was not possible due to tiny amounts produced. Nevertheless desired compound **174** was only observed in traces, whereas two new compounds were produced in higher titres: major compound **178** and a minor compound **179**. Both compounds were isolated and subsequent NMR analysis resulted in two new structures which showed similarities to previously observed **172** from knockout studies of *schR3*. Compound **178** was shown to be an acyclic derivative of **172** and **179** appeared to be an oxidised congener of **178**, as the hydroxyl group at C-13 is oxidised to the ketone.

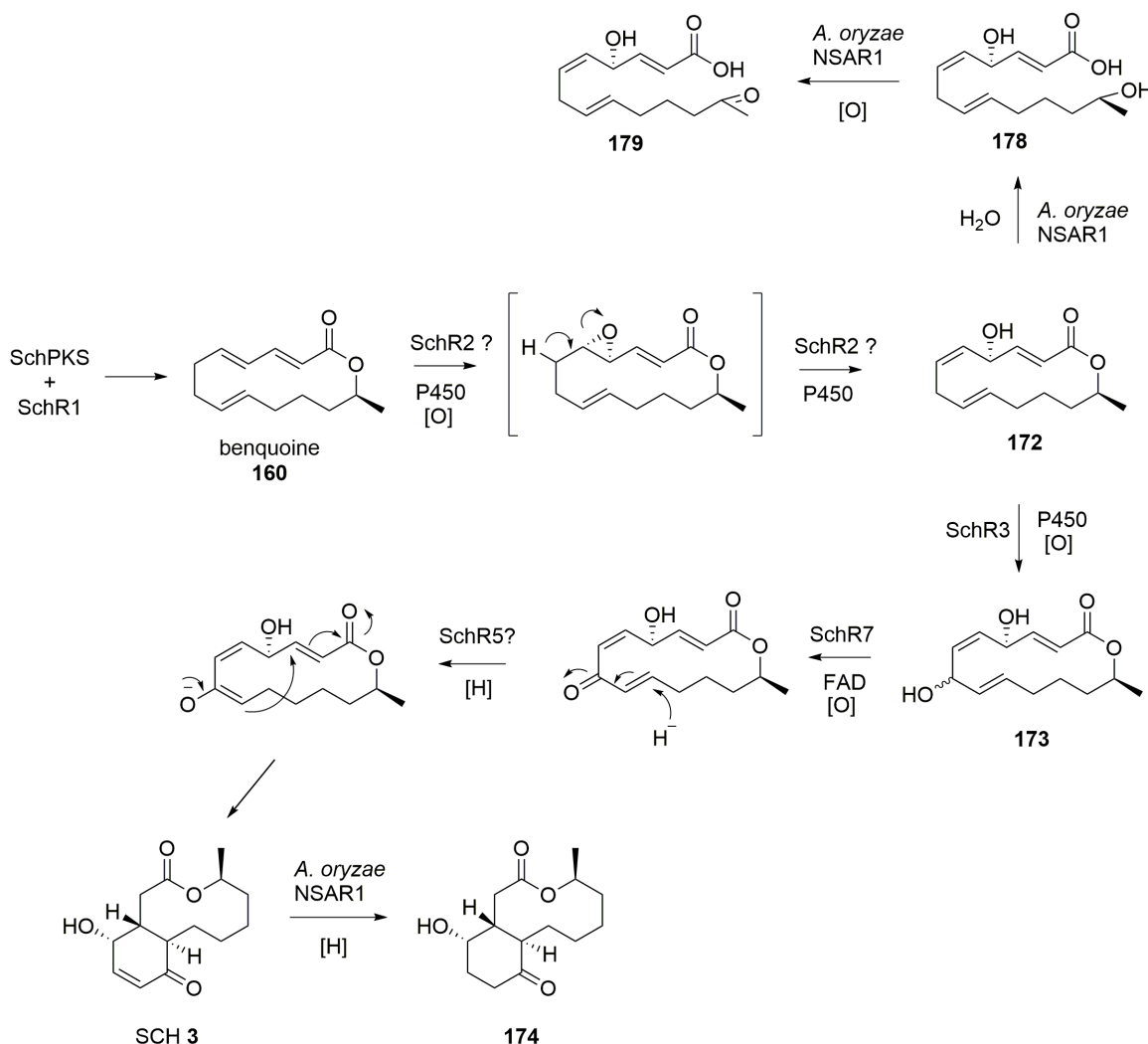


Scheme 4.11 Overview of compounds produced by heterologous expression experiments with different sets of genes from 3 BGC in *A. oryzae* NSAR1.

Previous knockout studies suggested a biosynthetic pathway, in which benquoine **160** is the first enzyme free intermediate and **172** the product of an oxidation step catalysed by cytochrome P450 SchR2. A second oxidation by P450 SchR3 is proposed to yield compound **173**, which was isolated from a *schR7* knockout experiment (Scheme 4.12).

As disruption of *schR3* and *schR7* led to lactone intermediates, it was suggested that the acyclic intermediates **178** and **179** observed in heterologous expression experiments are shunt products. Observation of very low titres of **174**, but high amounts of hydrolysed intermediates **178** and **179** in coexpression experiments with all pathway genes is consistent with a shunt pathway in *A. oryzae* (Scheme 4.12). Endogenous genes in *A. oryzae* possibly lead to hydrolysis of an early pathway product in SCH biosynthesis (possibly **172**), as the minimal gene set for production of **178** and **179** included only three genes: *schPKS*; the partnering hydrolase encoding gene *schR1*; and one of the two P450 encoding genes *schR2* (exp. 4, table 4.2, Scheme 4.11). The real product of the reaction catalysed by SchR2 is probably **172**, which is hydrolysed to **178**. The shunt product **178** would then

accumulate and block the pathway to SCH. Further random oxidation of the hydroxyl group at C-13 of **178** would yield in **179**, probably catalysed by an oxidative enzyme from *A. oryzae* metabolism (Scheme 4.12).



Scheme 4.12 Proposed biosynthesis of SCH **3** with elucidated steps and involved enzymes from knockout studies conducted by Francesco Trenti⁸⁵ and putative shunt products derived from expression experiments in *A. oryzae* NSAR1.

Attempts to identify a product of SchPKS was not possible using heterologous expression in *A. oryzae*, probably due to analytical issues detecting a very non-polar compound such as the proposed benquoine **160** (Scheme 4.11, exp. 2+3). Benquoine was also never observed in *Phomopsis* wt fermentations and therefore prevented feeding of **160** to *A. oryzae*. Previous experiments in which linear fatty acid-like compounds were expressed in *A. oryzae* resulted in very rapid degradation.¹²³

Re-construction of the pathway in *A. oryzae* was not efficient, because of the putative shunt products which occur early during biosynthesis of SCH in the heterologous host. Nevertheless, it was shown that in principal all genes from *sch* BGC were cloned correctly and expressed in *A. oryzae* NSAR1, as

trace amounts of the expected reduced form of SCH **174** and **177** could be detected (Scheme 4.11, exp. 1).

Overall, these experiments show that *A. oryzae* NSAR1 is probably not a suitable host for reconstruction of SCH biosynthesis. Compound **178** was found to represent a "bottleneck" in expression experiments, by stopping the biosynthetic pathway to SCH possibly due to hydrolysis of an early intermediate **172** (Scheme 4.12). It was shown previously that *A. oryzae* endogenous enzymes can lead to shunt products or by-products of a biosynthetic pathway expressed heterologously.^{219,222} In order to avoid the undesired reaction in the host strain, Oikawa and coworkers introduced the downstream gene in the heterologous production of solanapyrone C, which abolished undesired products.²¹⁹ In our re-constitution experiments with genes from *sch* BGC this strategy was shown to be unsuccessful. The biosynthesis was still observed to be halted with production of **178** although downstream genes as second P450 SchR3 and also oxidase SchR7 were introduced to the expression system. Sequencing results showed that SchR3 and SchR7 were correctly cloned, ruling out inactivity of these proteins.

To overcome the putative shunt problem in *A. oryzae* NSAR1, another fungal host, such as *Trichoderma reesei*, could be chosen for heterologous expression of genes from the *sch* BGC. The filamentous ascomycete *T. reesei* is highly relevant for industry due to its outstanding secretory capacities and therefore also an attractive host for expression of homologous and heterologous proteins.^{223,224} In order to use *T. reesei* as recombinant production host, it was rationally engineered and optimised for fermentation processes and a broad selection of promoters and expression systems is provided.²²⁵

5 Overall Conclusion and Future Experiments

Overall, the combination of gene deletion experiments, fungal heterologous expression, and *in vitro* enzymatic reactions was demonstrated to be a powerful tool to elucidate the fungal polyketide pathways to squalestatin S1 **1**, strobilurin A **2**, and to SCH-642305 **3**, as well as to identify key enzymes for their biosynthesis.

Using a combination of seven targeted gene knockouts and heterologous expression of squalestatin genes in host *A. oryzae* revealed the order of enzymes involved in the biosynthesis of **1** and uncovered three previously unknown oxidative enzymes involved in the pathway to **1**. Two non-heme iron-dependent oxygenases Mfr1 and Mfr2 were shown to catalyse an impressive series of six consecutive oxidation steps from alkylcitrate **90** to form the core structure of **1**. Additional hydroxylation on the hexaketide chain catalysed by an unknown copper dependent monooxygenase Mfm1 revealed **94**, which is the substrate for acyltransferase Mfr4 yielding **92**. These results, together with earlier studies of the biosynthesis of **1** by Bonsch *et al.*⁵³ and Tang and coworkers¹¹⁸, contribute to the elucidation of the full biosynthetic pathway. Future experiments with purified oxygenases (Mfr1 and Mfr2) of the pathway could reveal more information about the detailed function of these unusual tailoring enzymes (detailed conclusion in section 2.8).

The previously unknown enzyme involved in the oxidative rearrangement to form the toxophore of strobilurins was uncovered in this thesis using heterologous expression of various gene sets for production of **2**. By narrowing down the number of coexpressed genes, the FAD dependent monooxygenase Str9 was identified to catalyse the oxidative rearrangement. *In vitro* studies with the purified enzyme confirmed the *in vivo* expression results and elucidated the rearranged product **135**, which was shown to undergo two *O*-methylation steps by *O*-MeTs Str2 and Str3 to yield **2**. The formally reduced congener bolineol **4** was shown to be produced by the same BGC as **2**. Product **135** was shown to be reduced in *A. oryzae* (probably catalysed by SDR Stl2 in *S. tenacellus*) to **133**, which forms the substrate for *O*-MeTs Str2 to yield **4**. Further expression results suggested that the SDR Stl2 together with oxidase Str4 could be involved in controlling the selective production of **2** or **4** in *S. tenacellus*, possibly regulated by control of the individual promoters. These results contribute to

the understanding of how nature has generated this valuable class of agricultural used fungicides. Future experiments could regard the investigation of the unusual domains within StPKS to uncover the origin of the *E,Z,E* motif of prestrobilurin **128** (detailed conclusion in section 3.6).

The usage of fungal heterologous expression for the elucidation of biosynthetic pathways was shown to be successful for **1** and **2**. Investigation of the SCH biosynthetic pathway were shown to be difficult using heterologous expression in *A. oryzae* due to shunt pathways in *A. oryzae*. Individual expression of *schPKS* or coexpression experiments with *schPKS* and *schR1* (in analogy to brefeldin biosynthesis), showed no intermediate *in vivo*, possibly due to rapid degradation of the PKS product. Feeding **3** to the fungal host *A. oryzae* led to conversion to a reduced congener **174**. Nevertheless, re-constitution of the biosynthetic pathway of **3** was shown to be successful by expression of the full *sch* BGC in *A. oryzae* yielding traces of biotransformed SCH **174**. However, the major products of this experiment were the acyclic derivative **178** and its oxidised congener **179**. It was suggested that both are shunt products originating from **172** ($\Delta schR3$ product),⁸⁵ which are hydrolysed by *A. oryzae* and therefore blocking the pathway to SCH. Coexpression experiments with *schPKS*, *schR1* (hydrolase), and *schR2* (P450) revealed the same panel of putative shunt products **178** and **179**. To overcome the putative shunt problem in *A. oryzae* NSAR1, another fungal host, such as *Trichoderma reesei*, could be chosen for heterologous expression of genes from the *sch* BGC (detailed conclusion in section 4.4).

6 Experimental

All chemicals and media ingredients used in this work were purchased from Sigma Aldrich, Thermo Fisher Scientific, Carl Roth, VWR, Acros and Duchefa Biochemie.

6.1 Antibiotics, media, buffer and solutions

Buffers and media used in this work were sterilised for 15 min at 120 °C (Autoclave 2100 Classic, Prestige Medical) or by disposable sterile filter (0.45 µm pore size, Roth). To prepare all media, buffers and antibiotic stocks, deionised water was further purified by a GenPure Pro UV/UF milipore device from Thermo Scientific.

6.1.1 Media

Table 6.1 Media used in this work for experiments with *E. coli*, *S. cerevisiae* and MF5453

Media	Composition [% (w/v)]	Ingredients
LB	0.50	yeast extract
	1.00	tryptone
	0.50	NaCl
Super optimal broth (SOB)	0.50	yeast extract
	2.00	tryptone
	0.06	NaCl
	0.02	KCl
SOC	93.75 (v/v)	SOB
	1.25 (v/v)	MgCl ₂ × 6 H ₂ O
	5.00 (v/v)	D(+)-glucose 20 %
YPAD	1.00	yeast extract
	2.00	tryptone
	2.00	D(+)-glucose monohydrate
	0.03	adenine
CM2 (pH 6.8)	2.00	cotton seed flour
	10.00	lactose monohydrate
	2.00 (v/v)	trace element solution
YMG (+TES)	0.40	D(+)-glucose monohydrate
	0.40	yeast extract
	1.00	malt extract
optional	2.00 (v/v)	trace element solution

Table 6.2 Media used in this work for experiments with *A. oryzae*

Media	Composition [% (w/v)]	Ingredients
GNB	2.00	D(+)-glucose Monohydrate
	3.00	Nutrient broth Nr. 2 from Oxoid
CMP	3.50	Czapek Dox broth
	2.00	D(+)-maltose monohydrate
Starch M	1.00	polypeptone
	2.00	starch from potato
	1.00	polypeptone
	5.00 (v/v)	solution A
	5.00 (v/v)	solution B
DPY	2.00	dextrin from potato starch
	1.00	polypeptone
	0.50	yeast extract
	0.50	monopotassium phosphate
	0.05	MgSO ₄ × H ₂ O

Table 6.3 Agar used in this work for experiments with *E. coli*, *S. cerevisiae* and MF5453

Agar	Composition [% (w/v)]	Ingredients
LB agar	0.50	yeast extract
	1.00	tryptone
	0.50	NaCl
	1.50	agar
ME agar	1.28	malt extract
	0.08	peptone ex Soya
	0.24	glycerol
	0.28	dextrin from potato starch
	1.50	agar
Softagar	0.50	NaCl
	1.00	agar
YPAD agar	1.00	yeast extract
	2.00	tryptone
	2.00	D(+)-glucose monohydrate
	0.03	adenine
	1.50	agar
SM-URA agar	0.17	yeast nitrogen base
	0.50	ammonium sulfate
	2.00	D(+)-glucose monohydrate
	0.077	complete supplement mixture minus uracil
	1.50	agar

Table 6.4 Agar used in this work for experiments with *A. oryzae*

Agar	Composition [% (w/v)]	Ingredients
CZD/S agar	3.50	Czapek Dox broth
	18.22	D-sorbitol (1 M)
	0.10	ammonium sulfate
	0.05	adenine
	0.15	L-methionine
	1.50	agar
CZD/S1 agar (-Met)	3.50	Czapek Dox broth
	18.22	D-sorbitol (1 M)
	0.10	ammonium sulfate
	0.15	L-methionine
	1.50	agar
	optional	
CZD/S Softagar	3.50	Czapek Dox broth
	18.22	D-sorbitol (1 M)
	0.10	ammonium sulfate
	0.05	adenine
	0.15	L-methionine
	0.80	agar
CZD/S1 (-Met) Softagar	3.50	Czapek Dox broth
	18.22	D-sorbitol (1 M)
	0.10	ammonium sulfate
	0.15	L-methionine
	0.80	agar
	optional	
CD1 agar	3.50	Czapek Dox broth
	4.68	NaCl (0.8 M)
	0.10	ammonium sulfate
	0.05	adenine
	0.15	L-methionine
	1.50	agar
CD2 agar (-Met)	3.50	Czapek Dox broth
	4.68	NaCl (0.8 M)
	0.10	ammonium sulfate
	0.15	L-methionine
	1.50	agar
	optional	
DPY agar	2.00	dextrin from potato starch
	1.00	polypeptone
	0.50	monopotassium phosphate
	0.50	yeast extract
	0.05	MgSO ₄ × H ₂ O
	2.50	agar

6.1.2 Buffers and solutions

Table 6.5 Buffers and solutions used in this work with their composition

Solutions	Composition [% (w/v)]	Ingredients
Solution A	4.00	NaNO ₃
	4.00	KCl
	1.00	MgSO ₄
	0.02	FeSO ₄ × 7 H ₂ O
Solution B	2.00	K ₂ HPO ₄
Trace elements solution	0.1000	FeSO ₄ × 7 H ₂ O
	0.0758	MnSO ₄ × 4 H ₂ O
	0.0025	CuCl ₂ × 2 H ₂ O
	0.0100	CaCl ₂ × 2 H ₂ O
	0.0056	boric acid
	0.0019	ammonium molybdate tetrahydrate
	0.0200	ZnSO ₄ × 2 H ₂ O (dissolved in 0.6 M HCl)
	Composition [mM]	
TAE buffer	40	Tris-HCl
	20	acetic acid
	1	EDTA
MF5453 protoplasting solution	10	sodium phosphate buffer pH 6
	800	NaCl
	10	CaCl ₂
	20 mg/ml	lysing enzyme from <i>Trichoderma hazianum</i> (Sigma-Aldrich)
<i>A. oryzae</i> NSAR1 protoplasting solution	10 mg/ml	Driselase from <i>Basidiomycetes sp.</i> (Sigma-Aldrich)
	800	NaCl
	10 mg/ml	lysing enzyme from <i>Trichoderma hazianum</i> (Sigma-Aldrich)
	50	Tris-HCl pH 7.5
Fungal transformation solution 1 (Trafo)	800	NaCl
	10	CaCl ₂
	60 % (w/v)	PEG 3350
Fungal transformation solution 2 (Trafo)	50	Tris-HCl pH 7.5
	800	NaCl
	10	CaCl ₂

6.1.3 Antibiotics

Antibiotics stock solutions were prepared in millipore water or ethanol. They were filter sterilised through 0.45 µm syringe filter and stored at -20 °C. Stock and working concentrations are listed in table 6.6. Carbenicillin was used for ampicillin selection.

Table 6.6 Antibiotics used in this work

Antibiotic	Solvent	Stock concentration [mg/ml]	Working concentration [µg/ml]
Carbenicillin (Carb)	H ₂ O	50	50
Chloramphenicol (Cm)	Ethanol	30	30
Hygromycin (Hyg)	H ₂ O	50	150
Kanamycin (Kan)	H ₂ O	50	50

6.2 Microbiology methods

6.2.1 Strains

Table 6.7 Strains and origin used in this work

Strain	Genotype	Phylum	Origin
<i>E. coli</i> OneShot TOP10	F- <i>mcrA</i> Δ(<i>mrr</i> - <i>hsdRMS</i> - <i>mcrBC</i>) Φ80 <i>lacZ</i> Δ <i>M15</i> Δ <i>lacX74</i> <i>recA1</i> <i>araD139</i> Δ(<i>ara</i> <i>leu</i>)7697 <i>galU galK rpsL</i> (Str ^R) <i>endA1 nupG</i>	<i>Proteobacteria</i>	Thermo Fisher Scientific
<i>E. coli</i> OneShot <i>ccdB</i> survival 2T1 ^R	F- <i>mcrA</i> Δ(<i>mrr</i> - <i>hsdRMS</i> - <i>mcrBC</i>) Φ80 <i>lacZ</i> Δ <i>M15</i> Δ <i>lacX74</i> <i>recA1</i> <i>ara</i> Δ139 Δ(<i>ara</i> - <i>leu</i>)7697 <i>galU galK</i> <i>rpsL</i> (Str ^R) <i>endA1 nupG flhA::IS2</i>	<i>Proteobacteria</i>	Thermo Fisher Scientific
<i>E. coli</i> BL21 (DE3)	F <i>ompT hsdSB</i> (<i>r_B</i> <i>m_B</i>) <i>gal dcm</i> (DE3)	<i>Proteobacteria</i>	Thermo Fisher Scientific
<i>Saccharomyces cerevisiae</i> CEN.PK	MATa/α <i>ura3-52/ura3-52</i> <i>trp1-289/trp1-289</i> <i>leu2-3_112/leu2-3_112 his3</i> Δ1/ <i>his3</i> Δ1MAL2-8C/MAL2-8C SUC2/SUC2	<i>Ascomycetes</i>	Hahn group, Hannover
<i>Aspergillus oryzae</i> NSAR1	Δ <i>argB</i> sC ⁻ <i>adeA niaD</i> ⁻	<i>Ascomycetes</i>	Lazarus group, Bristol
Unidentified strain MF5453	wildtype	<i>Ascomycetes</i>	

6.2.2 *E. coli*

Growth and maintenance

All *E. coli* strains were grown on solid or liquid LB medium with appropriate antibiotics. Cultures were incubated at 37 °C and 200 rpm for 12 to 18 h. For long term storage glycerol stocks (25 % glycerol) were stored at -80 °C.

Transformation of chemically competent *E. coli* cells

An aliquot of 50 μl chemically competent cells^c was defrosted on ice. After addition of either 1 μl of purified plasmid or 10 μl for a ligation mixture, the cells were left on ice for 10 to 30 min. The heat shock was carried out at 42 °C for 30 to 45 s before the cells were immediately put on ice for 2 min, following addition of 250 μl of SOC medium and 1 h incubation at 37 °C, 350 rpm. The cells were spun down for 15 s, most of the supernatant was discarded and around 200 μl were streaked onto LB agar plates with appropriate antibiotic(s). The plates were left to grow overnight in a 37 °C incubator.

^ckindly provided by technical staff

6.2.3 *S. cerevisiae*

Growth and maintenance

A stock of *S. cerevisiae* CEN.PK2 was spread onto YPAD agar medium and incubated at 30 °C for 48 h. 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 solid or liquid SM-URA, incubated at 30 °C for up to 3 days.

Preparation and transformation of competent *S. cerevisiae* cells

Preparation of *S. cerevisiae* competent cells was done using the LiOAc/SS carrier DNA/PEG protocol developed by Gietz and Woods.²²⁶ For each transformation competent yeast cells were prepared freshly. A single colony was inoculated into 10 ml of YPAD medium and grown overnight at 30 °C, 200 rpm. The seed culture was added to 40 ml of YPAD in a 250 ml flask and incubated at 30 °C, 200 rpm, for 4–5 h. Cells were harvested by centrifugation for 5 min at 3000 x g. After washing with 25 ml water the cell pellet was resuspended in 1 ml water and transferred to a 1.5 ml tube. The cells were pelleted at 20000 x g for 15 s, the supernatant discarded and cells resuspended in 400 µl water. Afterwards the suspension was aliquoted (100 µl into 1.5 mL tubes each). For yeast recombination based cloning using *S. cerevisiae*, competent cell aliquots were pelleted 20000 x g for 15 s and the supernatant discarded, following addition of 250 µl PEG solution, 36 µl 1 M LiOAc, 50 µl ssDNA (2 mg/mL) and up to 34 µl of DNA was added to the pelleted cells. Cells were resuspended in the transformation mixture by careful pipetting. The mixture was incubated at 30 °C for 50 min and afterwards cells were pelleted at 3000 x g for 15 s. The cell pellet was gently resuspended in 1 ml water before 200 µl of the mixture was spread over SM-URA plates and incubated at 30 °C for 3 to 4 days.

6.2.4 *A. oryzae* NSARI

Growth and maintenance

Aspergillus oryzae NSARI was grown on ME or DPY agar plates at 28 °C or in DPY medium at 28 °C and 120 rpm. For long term storage glycerol stocks (25 % glycerol) were prepared from spores collected from plates and stored at –80 °C.

PEG-mediated transformation of *A. oryzae* NSAR1

A. oryzae conidia from a sporulating plate were inoculated into 100 ml GN medium in a 250 ml flask, which was incubated overnight at 30 °C, 180 rpm. The grown mycelia was collected by filtration through sterile Miracloth and was incubated in 10 ml protoplasting solution containing 20 mg/mL Lysing Enzyme, which was sterilized through a 0.45 µm syringe filter, to build protoplasts (lying falcon, 30 °C, 70 rpm, 3 h). The protoplasts were released from hyphal strands by gentle pipetting with wide-bore pipette. Afterwards the mycelia was collected by filtration through sterile miracloth and washed with NaCl (0.8 M). The supernatant was centrifuged (10 min, 3000 x g) for collecting protoplasts. Resulting supernatant was discarded and collected protoplasts were resuspended in solution 1 (100 µl per transformation) and 10 µl of purified plasmid DNA was added. The transformation mixture was incubated on ice for 2 min before adding 1 ml solution 2 and incubating again for 20 min at room temperature. Afterwards pre-warmed 5 ml CZD/S(1) softagar (-methionine) (50 °C) was added to the mixture and overlaid onto prepared CZD/S(1) plates (-methionine). Plates were incubated at 28 °C for 3–5 days until colonies appeared.

For selection on arginine-(adenine or methionine) deficient media, the selection has to be repeated twice in order to prevent false positives due to consuming arginine from dead cell material. A few spores of the grown colony were picked with a tooth pick and transferred to a CD1 plate, without sorbitol. The plates were incubated at 28 °C for 3–4 days and the step was repeated by streaking out single colonies. In preparation for fermentation, a single colony was grown on DPY agar for 1 week.

6.2.5 Unidentified strain MF5453

Growth and maintenance

Unidentified fungus MF5453 was grown on solid ME agar or in different liquid media (CM2, YMG, YMG+TES) at 25 °C for production of SQS1 (200 rpm if liquid culture). For long term storage glycerol stocks (50 % glycerol) were prepared from mycelium of liquid cultures and stored at -80 °C.

Transformation

The fungus MF5453 was grown in CM medium for 48 h at 25 °C, 200 rpm. The mycelium was washed three times by filtration through a coarse sieve (0.8 M NaCl) in order to remove particles of cotton seed flour. Around 1 g of fine mycelia was collected in a sterile falcon tube and 10 ml of filter-sterilized protoplasting solution (20 mg/ml) lysing enzyme, Sigma L1412, 20 mg/ml of driselase

in 0.8 M sodium chloride, 10 mM sodium phosphate buffer pH 6) was added. The protoplasting mixture was incubated at 30 °C for 3 hours with gentle shaking (lyng falcon, 70 rpm). The suspension was filtered through a wide-bore and cotton-filled autoclaved pipette tip (5 ml) and washed with 0.9 M NaCl. Resulting protoplast solution was pelleted at 3000 × g for 10 min. Protoplasts were resuspended in 160 µl solution 1 and 40 µl solution 2. Non-purified DNA fragments (part1 and part2) of bipartite PCR were also added (10 µl each) and the suspension was incubated on ice for 30 min. After adding 1 ml of solution 2 the mixture was incubated for another 20 min at room temperature. Pre-warmed 10 ml soft agar (45 °C, 1 % agar containing 5 % NaCl) were poured to each transformation mixture and mixed gently before distributing on to two thin MEA pates containing 150 µg/ml hygromycin B (37 °C pre-warmed plates). After incubation at 25 °C for 24 h the plates were overlaid with 10 ml MEA agar containing 150 µg/ml hygromycin B. After incubation at 25 °C for 8–10 days, growing colonies were selected and cultivated on secondary plates with 150 µg/ml hygromycin B. Growing colonies were picked to MEA plate without antibiotic before transferring to liquid medium for subsequent fermentation and analysis.

6.3 Molecular biology methods

All enzymes were purchased from New England Biolabs (NEB) and used according to manufacturer's instructions with appropriate supplied buffers.

6.3.1 Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) was used to amplify DNA fragments. *OneTaq*[®] 2X Master Mix (NEB) with standard buffer was used for screening purposes. For cloning tasks the proofreading *Q5*[®] High-Fidelity 2X Master Mix (NEB) was used. For both enzymes manufacturer's instructions were followed. In case of *E. coli* colony screening a single colony was picked with a toothpick and directly transferred to the reaction tube as template for the PCR reaction.

6.3.2 Oligonucleotides

All oligonucleotides used in this work were designed using the software Geneious (7.1.9) and purchased from Sigma Aldrich.

Table 6.8 Oligonucleotides used in this work

Cox ID number	5'-3' sequence
22	CGTCAGGACATTGTTGGAG

Table 6.8 Oligonucleotides used in this work

Cox ID number	5'-3' sequence
23	CTGTCGAGAAGTTTCTGATCG
57	GCCAACTTTGTACAAAAAGCAGGCTCCGCATGGACGTTTCCAAGGAAGA
60	TCCTCGCCATGCCTGTTATC
61	CCGCGGAATGGCTTGATAAC
62	TGCCAACTTTGTACAAGAAAGCTGGGTCCGGCTAAGAACCAGAGTTCTTCA
64	ATGCCAAGTTGCCAACAGGT
87	CTTCTTAAATATCGTTGTAAGTGTTCCTGA
88	CGAAGTATATTGGGAGACTATAGCTACTAG
89	ATTCACCACTATTATCCACCCTATAATA
90	GAGACGAAACAGACTTTTTTCATCGCTAAAA
91	CTTTCTTTTCTCTTTCTTTTCCCATCTC
92	TGACCTCCTAAAACCCCAGTG
306	TTAAGAAGGAGCCCTTACCAAGGGTGGGCCCGGGAAACCTTGACTCACTGCGG
307	ACTAGAGGATCCCCATCATGGGCATATCTCGCACTTTGGT
308	ACCAAAGTGCGAGATATGCCATGATGGGGATCCTCTAGT
310	AGGTCACGGTATAGCGCAGGCAGGTCGAGTGGAGATGTGG
313	TACGACAATCTCTCTCCACCCATGATGGGGATCCTCTAGT
314	ACTAGAGGATCCCCATCATGGGTGGAGAGAGATTGTCGTA
315	CCACATCTCCACTCGACCTGGGATCGTTGCGGAAATACAC
316	GTGTATTTCCGCAACGATCCCAGGTCGAGTGGAGATGTGG
411	AGCGAATGGATATGGTCGGG
461	ACTAGAGGATCCCCATCATGTGCATCGGGGCCGCTCCATG
462	CATGGAGCGGCCCGATGCACATGATGGGGATCCTCTAGT
463	CCACATCTCCACTCGACCTGGGGATCGAGTCATCTCTTCG
464	CGAAGAGATGACTCGATCCCCAGGTCGAGTGGAGATGTGG
567	ACTAGAGGATCCCCATCATGCGACCACAGCTTGCAACATC
568	GATGTTGCAAGCTGTGGTGCATGATGGGGATCCTCTAGT
569	AAGATGGCATTGTTTCGTGGCCAGGTCGAGTGGAGATGTGG
570	CCACATCTCCACTCGACCTGGCCACGAACAATGCCATCTT
573	ACTAGAGGATCCCCATCATGCACTACCAAGAAGCTCGGCA
574	TGCCGAGCTTCTTGGTAGTGCATGATGGGGATCCTCTAGT
575	CCTACTTTGCCGTCTGTGGTCAGGTCGAGTGGAGATGTGG
576	CCACATCTCCACTCGACCTGACCACAGACGGCAAAGTAGG
578	CTATATGGGAGAGGGATCAG
579	TGCCATCTCATGATTATCAGAC
580	TTTATATCTCACGACGTGCA

Table 6.8 Oligonucleotides used in this work

Cox ID number	5'-3' sequence
581	ATTGGCTCTTGTATGTTAC
582	GAAACCTTGACTIONACTGCGG
583	TAAAAATTTGAACTTTCGCGAG
584	TGCATTTCTCACCATCGCAC
585	TCGTCTCGAGAGCCATTTCG
586	AGAACACCAAGGAGCGAACC
587	GGTCTACGAGCATCAGAGGC
589	GCCAACTTTGTACAAAAAGCAGGCTCCGCTGCATTTCTCACCATCGCAC
590	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTTCGTCTCGAGAGCCATTTCG
591	GCCAACTTTGTACAAAAAGCAGGCTCCGCAGAACACCAAGGAGCGAACC
592	TGCCAACTTTGTACAAGAAAGCTGGGTCGGGGTCTACGAGCATCAGAGGC
593	GCCAACTTTGTACAAAAAGCAGGCTCCGCTTTATATCTCACGACGTGCA
594	TGCCAACTTTGTACAAGAAAGCTGGGTCGGATTGGCTCTTGTATGTTAC
597	GCCAACTTTGTACAAAAAGCAGGCTCCGCCTATATGGGAGAGGGATCAG
598	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTGCCATCTCATGATTATCAG
605	CCACATCTCCACTCGACCTGAGATGCACCTCTGTAAC
606	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTCACATCCTTAGACCTTCTTC
607	TTTCTTTCAACACAAGATCCCAAAGTCAAATGCGATTACTAGTACCAT
608	ACGACAATGTCCATATCATCAATCATGACCTTATCCTGTAACACTTCTT
609	GTCGACTGACCAATTCCGCAGCTCGTCAAATGGCTATCGTCAACGGCGC
610	GGTTGGCTGGTAGACGTCATATAATCATACTACAACCTGAACTTCTGCA
611	TTTCTTTCAACACAAGATCCCAAAGTCAAATGCTCAGAATGAAGAAGAT
612	TTCATCTATGCGTTATGAACATGTTCCCTCTAGCGAGCTAAGAATGCTT
613	TAACAGCTACCCCGCTTGAGCAGACATCACCATGGCGACAGCTATCCTTCC
614	ACGACAATGTCCATATCATCAATCATGACCCTAAAAGAACACCAAGGAGC
615	GTCGACTGACCAATTCCGCAGCTCGTCAAATGGCGACTGCAACAACAAC
616	GGTTGGCTGGTAGACGTCATATAATCATACTAGTCAAAGAAAACCCAGA
746	GATGTCTACTTCGCGGAGTGCAGGTCGAGTGGAGATGTGG
747	CCACATCTCCACTCGACCTGCACTCCGCGAAGTAGACATC
748	AACTTTGTACAAGAAAGCTGGGTCGGCGCGCCCGGGCAAGAGCAGCAATTGACCAT
751	AAGCCGACATGCTATTGAGCCATGATGGGGATCCTCTAGT
752	CCACACATTGATCGATCCGTCAGGTCGAGTGGAGATGTGG
753	CCACATCTCCACTCGACCTGACGGATCGATCAATGTGTGG
754	AACTTTGTACAAGAAAGCTGGGTCGGCGCGCCCGGGTCCATTGTATGTCCATGCCT
765	GCCAACTTTGTACAAAAAGCAGGCTCCGCAGTCAACAACAACACTCTAC
766	ACTAGAGGATCCCATCATGTTGGGGTGCAGTGACATCGT

Table 6.8 Oligonucleotides used in this work

Cox ID number	5'-3' sequence
767	CAAGAGCAGCAATTGACCAT
768	AGAACACCAAGGAGCGAACC
769	GACTGCAACAACAACCTCTAC
770	TCCATTGTATGTCCATGCCT
822	ATGGCGGCTACGAATGATCCC
823	GAAACCTTGACTCACTGCGGC
841	GTCGACTGACCAATTCCGCAGCTCGTCAAAATGGACTTCCCCGGGAATTC
842	GGTTGGCTGGTAGACGTCATATAATCATACTACCAATGTGTGCGAAAGGA
861	CCAATGTATACTGGTCTCACCCATCCGCGAAACAATTGGCCACTGCCTC
862	GAGGCAGTGGCCAATTGTTTCGCGGATGGGGTGAGACCAGTATACATTGG
863	TGTAGAGGTCTGCCACGGTTGCCATATGATGAGAATGATAGCCCATTGGT
864	ACCAATGGGCTATCATTCTCATCATATGGCAACCGTGGCAGACCTCTACA
867	GCCAACTTTGTACAAAAAGCAGGCTCCGCAATCCTCACATCCTTAGACC
868	ACTAGAGGATCCCCATCATGGCGAGATATGCCGAATAGTA
869	TACTATTCGGCATACTCGCCATGATGGGGATCCTCTAGT
870	GCTGCTCCACATAACTTGCTCAGGTCGAGTGGAGATGTGG
871	CCACATCTCCACTCGACCTGAGCAAGTTATGTGGAGCAGC
872	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCGATATGTGGTTCAATCATT
886	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGGACTTCCCCGGGAATTC
887	ACGACAATGTCCATATCATCAATCATGACCTACCAATGTGTGCGAAAGGA
988	ATGCCATACCCGTCTGACTTG
989	CAAGCAAATAGTCCAACGACA
966	GCCAACTTTGTACAAAAAGCAGGCTCCGCATGCCTTCGTCATACATACA
967	AGAAGTGGGTAGCACGTTC
968	AGCCTCTCGTGACTGCATTG
969	TGCTGTAGTTGTGCTTGGAG
970	CCTCTCCAACGCCACGGTCT
971	CGTGAGCTGAAGATATTGGC
972	CACATTGGCGCCGAGATATT
973	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTCACGTAACAACCATCTTGG
975	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGGCGCCAACAGTCGAACC
976	GGTTGGCTGGTAGACGTCATATAATCATACTAATTCCTTAATAAATGCT
977	TTCATTCTATGCGTTATGAACATGTCCCTCTAATTCCTTAATAAATGCT
978	TAACAGCTACCCCGCTTGAGCAGACATCACATGGCATTACACGATATTTT
979	ACGACAATGTCCATATCATCAATCATGACCCTACTTGACGTGGAAGCGAA
980	GTCGACTGACCAATTCCGCAGCTCGTCAAAATGCTGAGCATTACAATATT

Table 6.8 Oligonucleotides used in this work

Cox ID number	5'-3' sequence
981	GGTTGGCTGGTAGACGTCATATAATCATACCTAGTCAGTGGTGAGCTTTT
982	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGGTCCGAATTCTGTGTCT
983	TTCATTCTATGCGTTATGAACATGTTCCCTTCAGTCATAGTTTGCCATTG
984	AACAGCTACCCCGCTTGAGCAGACATCACCATGGAGCAACAAACCATTCT
985	ACGACAATGTCCATATCATCAATCATGACCTCATACAGCAAAAGATTGCC
986	GTCGACTGACCAATTCGCGAGCTCGTCAAAATGAACCCTTACTTTCAAAG
987	GGTTGGCTGGTAGACGTCATATAATCATACTTACAGCACTGCAGAGCTGA
990	ATGATCGTGATGTCTCGCCTG
991	CTATCGCAGCATCCTCCCTCT
992	ATGACCATCCTCCGTTCTCGC
993	ATTGTCTACGATCGGTCGCGC
1011	ATGGCCGTTGATCGCAAGA
1012	CTAGACAAGGCGTTCCTCA
1013	ATGCCTGAAGGAACCATGA
1014	TTAGGGCTTGATCGCAATT
1015	ATGTCACCTACTGCTGAAAT
1016	GAATACGAATGGACCATCAG
1036	TCCTTCCCTCTCTCCGGTGACTCGTTTTT
1037	GCAGTCCCTACTCTCGCGT
1038	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGGCCGTTGATCGCAAGAT
1039	GGTTGGCTGGTAGACGTCATATAATCATACCTAGACAAGGCGTTCCTCAA
1079	CAGGGACAATGCCCATGGTTACAGCAACGTCCTTGAAGTTCACACCCGCC
1080	AGCTGTACGCGCGGGTGTGAACCTCAAGGACGTTGCTGTAACCATGGGC
1081	TCAAGAACACACTCGCAAGCGTCGATGCCTCCTCAAATGACATTGAGTCA
1082	ACCATATTCCTGACTCAATGTCATTTGAGGAGGCATCGACGCTTGCGAGT
1083	CCCTCCCGTGGCGGAATGAATCAAGACTCGATGACCCTTTTGGGTGTTTG
1084	TACGACCTGGCAAACACCCAAAAGGGTCATCGAGTCTTGATTATTCCGC
1085	CTTCTCATCATTTCCGACGGTGGCGAATATCTCGGCGCAATGTGCTGGC
1086	ATCCAGATTTGCCAGCACATTGGCGCCGAGATATTCGCCACCGTCGAAA
1103	TGTAACCATCTGCTTTCTTATCAAAAGTGTGTCCATATTATGCTCAGGG
1104	TGTACCTCAGCCCTGAGCATAATATGGACAACACTTTTGATAAGAAAAGCA
1105	ATGTCTTCTCCTGCTGCTCA
1106	TTAGACACGCTTCTGCGACC
1107	ATGGCCGCCGAATCTGCTAA
1108	TTAGGCACTTCTTTCACCG
1109	GTCGGGCTGGGTACATCCAT

Table 6.8 Oligonucleotides used in this work

Cox ID number	5'-3' sequence
1110	GGGAGTCTCTCAAGGCCCTCC
1116	CGACCATGATGATAACCACGGC
1134	ATGGGAGCCTCGCACTCGACA
1135	TTATGCGACGCCCTTGATGAG
1251	TTTCTTTCAACACAAGATCCCAAAGTCAAATGTCTTCTCTGCTGCTCA
1252	TTCATTCTATGCGTTATGAACATGTTCCCTTTAGACACGTTCTGCGACC
1253	TAACAGCTACCCCGCTTGAGCAGACATCACATGGCCGCCGAATCTGCTAA
1254	ACGACAATGTCCATATCATCAATCATGACCTTAGGCACTTCTTTCACCG

6.3.3 Details for constructed vectors in this thesis

Combination of oligonucleotides used for amplification of squalenylase KO fragments and construction details for fungal transformation plasmids used for heterologous expression of strobilurin A genes and SCH genes in *A. oryzae* NSAR1 are displayed in tables 6.9, 6.10 and 6.11.

Table 6.9 Combination of oligonucleotides used for amplification of squalenylase KO fragments

Construct ID	Target	Oligonucleotides for amplification of KO fragments	Oligonucleotides for construction in <i>S. cerevisiae</i>
KELI30	<i>mfr4</i>	P _g : P578 + P22 P23 + P _h : P579	P _{a+b} : P597+314 P _{c+d} : P598+315 P _{e+f} : P313+316
KELI27	<i>mfm1</i>	P _g : P584 + P22 P23 + P _h : P585	P _{a+b} : P589+567 P _{c+d} : P570+590 P _{e+f} : P568+569
KELI28	<i>mfr1+2</i>	P _g : P586 + P22 P23 + P _h : P587	P _{a+b} : P591+573 P _{c+d} : P576+592 P _{e+f} : P574+575
KELI38	<i>mfr1</i>	P _g : P767 + P22 P23 + P _h : P768	P _{a+b} : P591+573 P _{c+d} : P747+748 P _{e+f} : P574+746
KELI39	<i>mfr2</i>	P _g : P769 + P22 P23 + P _h : P770	P _{a+b} : P765+766 P _{c+d} : P753+754 P _{e+f} : P571+752
KELI29	<i>mfr3</i>	P _g : P580 + P22 P23 + P _h : P581	P _{a+b} : P593+461 P _{c+d} : P594+463 P _{e+f} : P462+464
KELI48	<i>mfm3</i>	P _g : P822 + P22 P23 + P _h : P823	P _{a+b} : P867+868 P _{c+d} : P869+870 P _{e+f} : P871+872

Table 6.10 Construction details for fungal transformation plasmids used for heterologous expression of strobilurin A genes in *A. oryzae* NSAR1

Construct ID	Vector backbone	Template	Oligonucleotides for construction in <i>S. cerevisiae</i>
KELIV01A	pTYGSarg	cDNA	<i>stPKS, str2, str3, str4</i> ; kindly provided by Dr Risa Nofiani
KELI44	pTYGSade	cDNA	<i>str8, str10, str11</i> ; kindly provided by Dr Risa Nofiani
KELIV01B	pTYGSmet	cDNA	<i>str9, stl2</i> ; kindly provided by Dr Kate de Mattos-Shiple
KELIV07	pTYGSarg	KELIV01A	template cut with <i>NdeI</i> and repaired with patch amplified from pTYGSarg: P1136+1137 to remove all genes but <i>stPKS</i>
KELIV04	pTYGSmet	KELIV01B	<i>str9</i> : P1038+1039
KELIV11	pTYGSniaD	KELIV01A	<i>str2</i> : P1251+1252; <i>str3</i> : P1253+1254; Patch_Peno: P87+88
KELIV12A	pTYGSniaD	KELIV11	<i>str2</i> : P1251+1252; Patch_PgdpA: P91+92; Patch_Peno: P87+88
KELIV12B	pTYGSniaD	KELIV11	Patch_PadH: P89+90; <i>str3</i> : P1253+1254; Patch_Peno: P87+88

Table 6.11 Construction details for fungal transformation plasmids used for heterologous expression of SCH genes in *A. oryzae* NSAR1

Construct ID	Vector backbone	Template	Oligonucleotides for construction in <i>S. cerevisiae</i>
KELIII02A(No2_3)*	pE-YA	cDNA	<i>schPKS 4 fragments</i> : P966-P973
KELIII07*	pE-YA	KELIII02A(No2_3)	<i>schPKS 7 fragments</i> : P966+1103; P1104+1109, P1110+1079, P1080+1081, P1082+1083, P1084+1085, P1086+973
KELIII08	pE-YA	KELIII02A(No2_3)	Patch: P1104+1116; KELIII07 cut with <i>KpnI</i>
KELIII02B	pTYGSarg	cDNA	<i>schR1</i> : P975+976
KELIII02C*	pTYGSmet	cDNA	<i>schR1</i> : P975+977, <i>schR2</i> : P978+979, <i>schR3*</i> : P980+981
KELIII02D	pTYGSmet	cDNA	<i>schR4</i> : P982+983, <i>schR5</i> : P984+985, <i>schR7</i> : P986+987
KELIII08A	pTYGSarg	-	LR with KELIII08 and KELIII02A
KELIII08B	pTYGSarg	-	LR with KELIII08 and KELIII02B
KELIII08C*	pTYGSarg	-	LR with KELIII08 and KELIII02C
KELIII10A	pTYGSade	KELIII08C, gDNA (<i>schR3</i>)	<i>schR1</i> : P975+977, <i>schR2</i> : P978+979, <i>schR3</i> : P980+981
KELIII10B	pTYGSade	KELIII08C	<i>schR1</i> : P975+977, <i>schR2</i> : P978+979, Patch_Peno: P87+88
KELIII10C	pTYGSade	KELIII08C	<i>schR1</i> : P975+977, Patch_PgdpA: P91+92, <i>schR3</i> : P980+981
KELIII11A	pTYGSmet	KELIII02D	Patch_PadH: P89+90, Patch_PgdpA: P91+92, <i>schR7</i> : P986+987
KELIII11B	pTYGSmet	KELIII02D	Patch_PadH: P89+90, <i>schR5</i> : P984+985, <i>schR7</i> : P986+987
KELIII11C	pTYGSmet	KELIII02D	<i>schR4</i> : P982+983, Patch_PgdpA: P91+92, <i>schR7</i> : P986+987
KELIII15	pTYGSade	KELIII10A	Patch_PadH: P89+90, <i>schR2</i> : P978+979, <i>schR3</i> : P980+981

*point mutations in gene

6.3.4 cDNA preparation

cDNA was prepared from total RNA using oligo(dT) nucleotides and the RevertAid Premium Transcriptase kit purchased from Thermo Fisher Scientific. Manufacturer's instructions were followed. In case of SCH cDNA preparation for amplification of *schPKS* Maxima Reverse Transcriptase (NEB) was used in combination with specific primer P973.

6.3.5 Agarose gel electrophoresis

Agarose gel electrophoresis was used to purify and analyse double stranded DNA fragments and was performed using horizontal gel tanks (BioRad) with 0.5 x TAE buffer. DNA samples were mixed with 6X DNA Loading Dye and loaded on 1% (*w/v*) agarose gels (electrophoresis grade agarose dissolved in 40 ml TAE buffer for two gels) which contained RedSafe DNA Stain (1 μ l in 40 ml). 1 kb DNA Ladder (NEB) was used as a molecular DNA size marker. Electrophoresis was carried out at 80–120 V for 20 min to 45 min using 0.5xTAE as running buffer. DNA was visualised under a UV-trans-illuminator (254 nm) and photographed using a UVP camera (Gel Doc XR+, Bio-Rad). NucleoSpin[®] Gel and PCR Clean-up kit purchased from Macherey–Nagel was used to purify and extract PCR amplified DNA after agarose gel purification. Manufacturer's protocol was followed as recommended.

6.3.6 DNA purification from gel or from PCR

Amplified DNA was purified for sequencing or cloning using NucleoSpin[®] Gel and PCR Clean-up kit purchased from Macherey–Nagel. Manufacturer's protocol was followed as recommended. Prior to elution with ddH₂O the column was heated to 70 °C for 5 min.

6.3.7 Isolation of plasmid DNA from *E. coli*

Isolation of plasmid DNA from an overnight culture of *E. coli* was performed using the NucleoSpin[®] Plasmid kit (Macherey–Nagel), according to the manufacturer's protocol. For larger plasmids (> 15 kb) PureLink[®] Quick Plasmid Miniprep Kit (Invitrogen) was used to obtain high concentrated plasmid.

6.3.8 Isolation of plasmid DNA from *S. cerevisiae*

Zymoprep Yeast Plasmid Miniprep II kit purchased from Zymo Research was used to isolate plasmid DNA from yeast cells. The whole amount of colonies of half a SM-URA transformation plate generated by yeast recombination were collected with a tooth pick into the first solution of the yeast miniprep kit. Yeast DNA was isolated according to manufacturer's instructions. For final elution 10 μ L ddH₂O was used.

6.3.9 Isolation of genomic DNA from fungus

Approximately 100 mg of filtered mycelium from a liquid culture was used to extract genomic DNA, following the the manufacturer's instructions of the GeneElute™ Plant Genomic DNA Miniprep Kit (SIGMA Life Science).

6.3.10 Cloning procedure

Restriction digest

Restriction enzyme digestions (single or double) were carried out according to the manufacturer's guidelines. Reaction mixtures were incubated at 37 °C in a incubator for 0.5 to 5 h and afterwards subsequently inactivated for 20 min at 65 °C or 80 °C. The resulting fragments were then separated by agarose gel electrophoresis or used directly for yeast recombination.

Dephosphorylation of linear DNA and ligation

In order to prevent self-ligation of linearised plasmids, they were treated with shrimp alkaline phosphatase (SAP). After digestion, 1 μ L of enzyme was added to the reaction mixture for 1 h at 37 °C. Ligation reactions were performed using T4 DNA ligase. The reaction typically contained 5X molar excess of DNA insert. Reaction mixtures were incubated in the fridge overnight followed by inactivation at 65 °C for 20 min.

Gateway Cloning

For *in vitro* recombination from pE-YA entry clones into pTYGS family destination vectors, the Invitrogen Gateway LR Clonase II enzyme mix was used according to manufacturer's protocol with a reduced reaction scale (reduced by half).

DNA sequencing

DNA samples were sequenced by Eurofins Genomics (Ebersberg).

6.4 Biochemistry methods

A codon optimised sequence for *str9* was purchased from Invitrogen (Darmstadt, Germany). The sequence was constructed with a N-terminal 6xhis tag in the expression vector pET100.

6.4.1 Recombinant protein expression and cell lysis

Recombinant N-terminal hexa-histidine tagged Str9 was produced by overexpression in BL21 (DE3) *E. coli* cells using the plasmid pET100-His₆-Str9, which contains a codon optimised sequence for Str9. For overproduction single colonies were used to inoculate LB medium. Cultures were grown to an OD₆₀₀ of 0.4 to 0.6 at 37 °C, induced with 0.1 mM isopropyl- β -D-thiogalactopyranoside (IPTG) and shaken at 16 °C for 20 h. Cells were harvested by centrifugation (4000 g, 4 °C, 15 min). For purification, 2 g cells were resuspended in 20 mL lysing buffer (50 mM Phosphate buffer, pH 8.0) and cell disruption was done by sonification (pulse program, 30 sec on/off, 15 min). After centrifugation (20000 g, 4 °C, 40 min) the obtained crude lysate was passed through a gravimetric Ni-NTA column (2 mL bed volume).

6.4.2 Protein purification by nickel affinity chromatography

The loaded Ni-NTA column was washed with 5 mL of lysing buffer containing 20 mM imidazole. Elution of the target protein was achieved with 3 mL of lysing buffer containing 250 to 500 mM imidazole. The elution fractions were combined and concentrated (Satorius Vivaspin 20–10000 MW cut-off) and imidazole was removed from the buffer. The purified enzyme was immediately used for activity assays or analysed using polyacrylamide gel electrophoresis (SDS-PAGE).

6.4.3 SDS-PAGE

Separation and analysis of proteins was carried out using 12 % polyacrylamide gels (Table 6.12). Protein samples were prepared with SDS loading buffer and boiled at 95 °C for 10 min before 15 μ l of the samples and 5 μ l of Color Prestained Protein Standard (Broad Range, 11–245 kDa, NEB) were loaded onto the gel. The gels were run for 60 min at 40 mA. Afterwards gels were incubated in 30 mL

Coomassie staining solution for 1 h before destaining in Coomassie bleach solution for another hour. Destained gels were scanned using the Molecular Imager Gel doc XR+ (Bio–Rad) system.

Table 6.12 Composition of a 12 % polyacrylamide gel

	12 % separating gel [ml]	5 % stacking gel [ml]
30 % Acrylamide/bisacrylamide® (Rotiphorese Gel 30 (37.5:1))	3.00	0.535
ddH ₂ O	2.45	1.70
1.5 M Tris–HCl, pH 8.8	1.90	-
0.5 M Tris–HCl, pH 6.8	-	0.25
10 % (w/v) SDS	0.075	0.02
10 % (w/v) APS	0.075	0.02
TEMED	0.003	0.002

Table 6.13 Buffer and solution used for protein purification and analysis

Solutions	Composition [% (w/v)]	Ingredients
Coomassie Dye	25	acetic acid
	10	isopropanol
	0.1	coomassie
Coomassie Bleach	25	acetic acid
	10	Isopropanol
4 x Lämmli buffer	250 mM	Tris–HCl, pH 6.8
	8	SDS
	40	glycerine
	10	β–Mercaptoethanol
	2.5 mg/l	bromophenol blue
	Composition [mM]	
10 x SDS Running buffer	25	Tris
	192	Glycine
	0.1%	SDS

6.4.4 Enzyme activity assay with Str9

Activity assays were carried out in a total volume of 400 µL containing 2 mM of prestrobilurin in 200 µL ethanol and 5 mg/mL of Str9 in 200 µL phosphate buffer (pH 8.0, 50 mM phosphate buffer plus 1.25 mM NADH/NADPH mixture and 0.3 mM FAD, 2.5 mM FMN mixture). Reactions were incubated at 30 °C for 30 min at 400 rpm and then for 60 h at room temperature without shaking. Prior to extraction with 400 µL ethyl acetate, reactions were acidified with 2 M HCl to pH 4. After evaporation of the solvent, the product was resolved in 150 µL DMSO and analysed by HPLC-MS. Reactions without enzyme served as controls.

6.4.5 Bolineol and 133 toxicity test against *A. oryzae*

Toxicity test were carried out on small agar plates filled with 5 ml DPY agar supplemented with different concentrations of bolineol 4 (0.1, 0.2, 0.3, 0.5, 1, 2 mg) or 133 (0.5, 1, 2, 5 mg). Both

compounds were diluted in 200 μ l DMSO prior to addition to the liquid DPY agar. Control plates were supplemented with 200 μ l DMSO only. *A. oryzae* NSAR1 spores were streaked on the plates on day 1.

6.5 Chemical analysis

6.5.1 Liquid Chromatography Mass Spectrometry (LCMS)

Analytical LCMS

Analytical LCMS data was obtained using a Waters LCMS system comprising of a Waters 2767 autosampler, Waters 2545 pump system, a Phenomenex Kinetex column (2.6 μ m, C18, 100 \AA , 4.6 x 100 mm) equipped with a Phenomenex Security Guard precolumn (Luna, C5, 300 \AA) 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 spectrometry, Waters SQD-2 mass detector, operating simultaneously in ES⁺ and ES⁻ modes between 150 and 1000 m/z . 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

In order to purify compounds from a raw extract or from a reaction mixture a Waters mass-directed autopurification system comprising of a Waters 2767 autosampler, Waters 2545 pump system, a Phenomenex Kinetex Axia column (5 μ m, C18, 100 \AA , 21.2 x 250 mm) equipped with a Phenomenex Security Guard precolumn (Luna, C5, 300 \AA). 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).

High Resolution Mass Spectrometry (HRMS)

High Resolution Mass Spectrometry (HRMS) was performed on a Q-ToF Premier mass spectrometer (Waters). It was coupled to an Acquity UPLC-domain. The ESI mass spectroscopy was measured in positive and negative modes.

6.5.2 Nuclear magnetic resonance (NMR) analysis

NMR measurements were acquired on a Bruker DRX 400, 500 MHz or 600 MHz and were done by the NMR department of Organic Chemistry Department, Leibniz Universität Hannover. Chemical shifts are shown in parts per million (ppm) in comparison to the TMS (Tetramethylsilane) standard. The coupling constants J were quoted in Hz throughout. Complete structural elucidation of synthesised compounds as well as purified/ isolated compounds were obtained with 2D experiments: Correlation Spectroscopy (COSY), Heteronuclear Single Quantum Coherence (HSQC), Heteronuclear Multiple Bond Correlation (HMBC) as required.

6.5.3 Extraction of MF5453 cultures

MF5453 cultures were clarified by Büchner filtration (sometimes after homogenising the cultures with a blender), the supernatant was acidified to pH 2-4 with concentrated H_2SO_4 and extracted twice with ethyl acetate. Combined organic layers were dried over MgSO_4 and solvent removed under reduced pressure. Extracts were dissolved in acetonitrile/ H_2O (1:1) to a concentration of 10 mg/ml, filtered over glass wool and analysed or purified by LCMS.

Small scale extraction (1 ml)

For screening of multiple putative KO transformants, a small scale extraction method was used. Therefore, 1 ml of 10 ml of MF5453 liquid culture (CM2 medium, grown for 7 days) was spinned down (5 min, maximum speed). The obtained supernatant was acidified with 2 M H_2SO_4 and transferred to a 2 ml centrifuge tube and extracted twice with 1 ml ethyl acetate (1:1). Solvents were removed in a vacuum centrifuge, samples were dissolved in 0.2 ml in acetonitrile/ H_2O and analysed by LCMS.

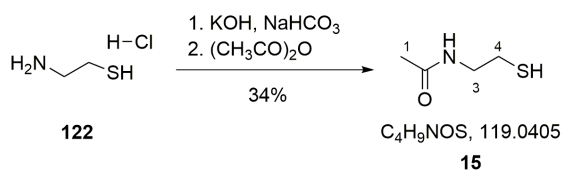
6.5.4 Extraction of *A. oryzae* cultures

A. oryzae cultures were clarified by Büchner filtration, cells were stirred in 100 ml acetone for 1 h, the solvent was removed under reduced pressure and to the remaining water fraction 100 ml H₂O was added and like the supernatant acidified to pH 2 with 2 M HCl and extracted twice with an equal amount of ethyl acetate. Combined organic layers were dried over MgSO₄ and solvent removed under vacuum. Extracts were dissolved in acetonitrile/H₂O (1:1) for squalenstatin intermediates and in methanol for SCH and strobilurin intermediates to a concentration of 10 mg/ml, filtered over glass wool and analysed or purified by LCMS.

6.5.5 Chemical synthesis

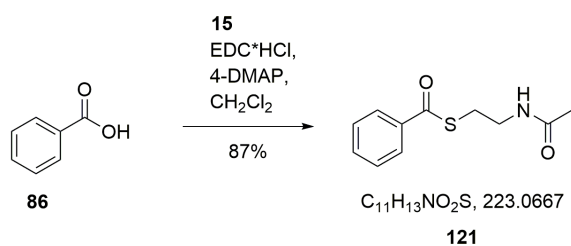
The numeration of the molecules in this thesis was defined pragmatically.

Synthesis of *N*-acetylcysteamine¹⁷⁴ (HSNAC, 15)



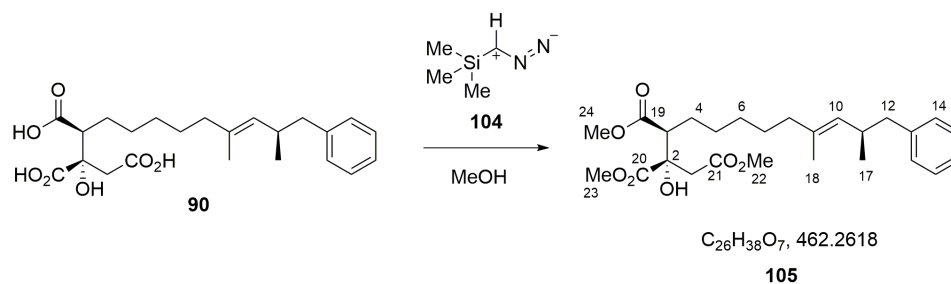
Cysteamine hydrochloride (9.01 g, 79.28 mmol, 1.0 eq.) was dissolved in water (360 ml). Sodium bicarbonate (20.00 g, 238.08 mmol, 3.0 eq.) and potassium hydroxide (4.51 g, 80.44 mmol, 1.02 eq.) were added to the stirred solution. Acetic anhydride (8.09 g, 7.49 ml, 79.26 mmol, 1.0 eq.) was added drop wise over 5 min and the reaction mixture was stirred for another 10 min. The aqueous layer was extracted with ethyl acetate (3 x 300 ml). The organic phases were combined, washed with saturated NaCl solution (100 ml), dried over MgSO₄ and concentrated *in vacuo* to yield SNAC **15** (colourless oil, 3.278 g, 27.50 mmol, 34%).

¹H NMR (400 MHz, CDCl₃) δ ppm: 1.34–1.38 (t, *J* = 8.5 Hz, 1H, S–H), 2.00 (d, *J* = 1.1 Hz, 3H, H–1), 2.66–2.69 (m, 2H, H–4), 3.39–3.44 (qd, *J* = 6.3, 1.6 Hz, 2H, H–3); 6.12 (brs, 1H, N–H). ¹³C NMR (101 MHz, CDCl₃) δ ppm: 170.4 (C–2), 42.6 (C–3), 24.7 (C–4), 23.4 (C–1); *R_f* = 0.2, 1:1 (*v/v*) CH₂Cl₂:ethyl acetate.

Synthesis of benzoyl-SNAC²⁷ (**121**)

EDC x HCl (1.39 g, 0.0073 mol, 1.1 equiv.) and a catalytic amount of 4-DMAP (0.18 g, 0.0014 mol, 0.22 equiv.) were added to a stirred solution of benzoic acid (800 mg, 0.007 mol, 1.0 equiv.) and *N*-acetylcysteamine (0.86 g, 0.007 mol, 1.1 equiv.) in anhydrous CH₂Cl₂ (15 ml) at 0 °C. After stirring at room temperature overnight the reaction was quenched with HCl (1 M, 20 ml). The aqueous layer was extracted with CH₂Cl₂. The combined organic layers were washed with saturated NaHCO₃ and NaCl-solution, dried over MgSO₄ and concentrated under reduced pressure. Purification by flash chromatography on silica gel (CH₂Cl₂:EtOAc / 1:1) yielded benzoyl-SNAC **121** (1.27 g, 87%) as a white powder.

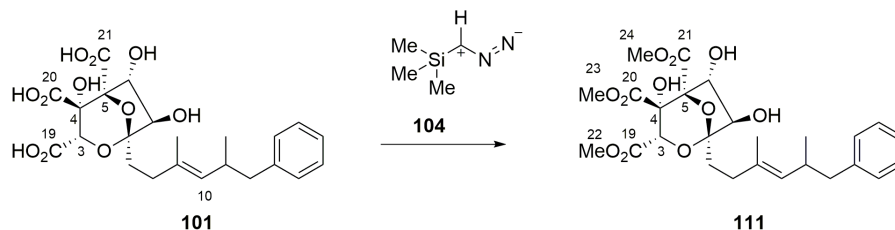
¹H NMR (400 MHz, CDCl₃) δ ppm: 1.98 (s, 3H, H-1), 3.22–3.25 (t, *J*=6.8 Hz, 2H, H-4), 3.52–3.56 (m, 2H, H-3), 5.91 (s, 1H, N-H), 7.45–7.48 (m, 2H, H-8/8'), 7.57–7.62 (m, 1H, H-9), 7.96–7.98 (m, 2H, H-7/7'). HRMS (ESI⁺) *m/z* calc. for C₁₁H₁₃NO₂SNa [M+Na]⁺ 246.0565, found 246.0562, *R*_f = 0.3 (1:1 (*v/v*) CH₂Cl₂:ethyl acetate).

Synthesis of **105**

Isolated, impure **90** (5.8 mg, 0.01 mmol, 1 equiv.) was dissolved in 1 ml MeOH and TMS (Trimethylsilyl)diazomethane solution (2.0 M in diethyl ether) (45 μl, 0.09 mmol, 9 equiv.) was added to the stirred solution. The mixture was stirred overnight at room temperature and quenched the next day by removing solvent under reduced pressure. The remaining product was purified by preparative LCMS as described in section 6.5.1 but a gradient from 70 % to 90 % acetonitrile was used which

yielded 1 mg of product **105**. For NMR assignment see section 2.5.3, figure 2.16. **HRMS** (ESI⁺) *m/z* calc. for C₂₆H₃₈O₇Na [M+Na]⁺ 485.2515, found 485.2513.

Synthesis of 111



Isolated **101** (4.8 mg, 0.01 mmol, 1 equiv.) was dissolved in 1 ml MeOH and TMS diazomethane solution (2.0 M in diethyl ether) (45 μ l, 0.09 mmol, 9 equiv.) was added to the stirred solution. The mixture was stirred overnight at room temperature and quenched the next day by removing solvent under reduced pressure. The remaining product was purified by preparative LCMS as described in section 6.5.1 which yielded 1.1 mg of product **111**. **HRMS** (ESI⁺) *m/z* calc. for C₂₆H₃₄O₁₁Na [M+Na]⁺ 522.1999, found 522.1998. For NMR assignment see section 2.5.4, figure 2.29.

6.5.6 Isolation of strobilurin expression intermediates

Compounds were isolated from raw extract of fungal cultures using preparative LCMS (15 min gradient: 10 to 90% water–acetonitrile).

Strobilurin A 2

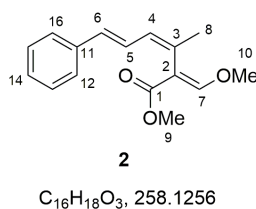
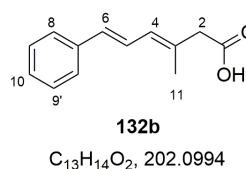
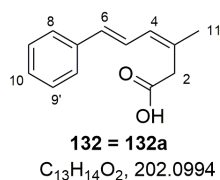


Table 6.14 NMR data of **2** compared to literature⁴⁸

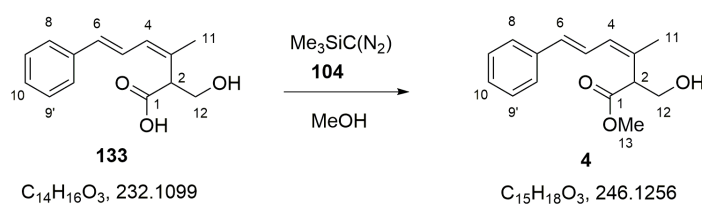
LK122			Literature		
Atom	δ_C / ppm	δ_H / ppm (J/Hz)	Atom	δ_C / ppm	δ_H / ppm (J/Hz)
1	168	q	1	167.3	q
2	111	q	2	110.3	q
3	131.5	q	3	131	q
4	129.9	6.27 m (1H)	4	129.4	6.23 dqd ($J = 9.4$ Hz, 1.2 Hz, 0.8 Hz)
5	126.7	6.62 dd ($J = 15.6, 10.7$)	5	126.2	6.67 dd ($J = 15.0$ Hz, 9.4 Hz)
6	131.3	6.49 dd ($J = 15.8, 0.8$ Hz)	6	130.7	6.44 dd ($J = 15.0$ Hz, 0.8 Hz)
7	159	7.43 s (1H)	7	158.5	7.40 s (1H)
8	23.8	1.98 m (3H)	8	23.3	1.97 d ($J = 1.2$ Hz, 3H)
9	51.8	3.74 s (3H)	9	51.1	3.69 s (3H)
10	62.1	3.85 s (3H)	10	61.4	3.74 s (3H)
11	138	q	11	137.4	q
12, 16	126.5	7.34 m (2H)	12, 16	125.9	
13, 15	128.6	7.28 m (2H)	13, 15	128.1	7.10-7.36 m (5H, phenyl)
14	127.3	7.19 m (1H)	14	126.8	

Compound 132

Isolation of **132** from raw extract (1 l fermentation) of *A. oryzae* expression strain containing *stPKS*, *str2*, *str3*, *str4*, *str11*, *str8*, *str10*, *str9* and *stI2* yielded in 90 mg of compound **132**. Structure elucidation revealed two isomers **132a** and **132b** in the ratio 3:1. HRMS (ESI⁻) m/z calc. for C₁₃H₁₅O₂ [M+H] 203.1072, found 203.1071.

NMR data of **132a** is identical to literature known data.¹⁹⁰ ¹H NMR (400 MHz, CDCl₃) of **132a** δ ppm 1.96 (s, 3 H, H-11), 3.32 (s, 2 H, H-2), 6.22 (d, $J = 10.8$ Hz, 1 H, H-4), 6.53 (d, $J = 15.4$ Hz, 1 H, H-6), 6.96 (dd, $J_1 = 10.9$ Hz, $J_2 = 15.5$ Hz, 1 H, H-5), 7.22 (m, 1H, H-10), 7.31 (m, 2 H, H-8/8'), 7.40 (m, 2H, H-9/9'); ¹³C NMR (101 MHz, CDCl₃) δ ppm 24.7 (C-11), 38.0 (C-2), 124.4 (C-5), 126.5 (C-9/9'), 127.6 (C-10), 128.7 (C-8/8'), 129.6 (C-4), 130.7 (C-3), 132.5 (C-6), 137.6 (C-7), 177.2 (C-1).

¹H NMR (400 MHz, CDCl₃) of **132b** δ ppm 1.96 (s, 3 H, H-11), 3.16 (s, 2 H, H-2), 6.16 (m, Hz, 1 H, H-4), 6.53 (d, $J = 15.4$ Hz, 1 H, H-6), 6.96 (dd, $J = 10.9$ Hz, 15.5 Hz, 1 H, H-5), 7.22 (m, 1H, H-10), 7.31 (m, 2 H, H-8/8'), 7.40 (m, 2H, H-9/9'); ¹³C NMR (101 MHz, CDCl₃) δ ppm 17.3 (C-11), 45.0 (C-2), 124.8 (C-5), 126.5 (C-9/9'), 127.6 (C-10), 128.7 (C-8/8'), 129.8 (C-4), 131.0 (C-3), 132.6 (C-6), 137.7 (C-7), 177.7 (C-1).

Methylation of **133** to bolineol **4**

To confirm the structure of compound **133** (0.011 g, 0.048 mmol, 1.0 eq.), it was methylated using TMS-CHN₂ (0.016 g, 0.14 mmol, 3.0 eq.) in MeOH at RT overnight in order to obtain bolineol **4** (2.1 mg) after purification using preparative LCMS. Comparison of the obtained ¹H NMR data (see below) to literature data¹⁸² confirmed the structure.

¹H NMR (400 MHz, CDCl₃) δ ppm 1.81 (m, 3H, H-11), 3.74 (3H, s, H-13), 3.68 (1H, dd, *J* = 11.1, 5.5, H-2), 3.99 (1H, dd, *J* = 8.7, 5.5, H-12), 4.10 (1H, dd, *J* = 11.1, 8.7, H-12), 6.22 (1H, dq, *J* = 11.1, 1.3, 4-H), 6.55 (1H, d, *J* = 15.4, H-6), 7.00 (1H, dd, *J* = 15.3, 11.0, H-5), 7.23 (1H, m, H-10), 7.32 (m, 2H, H-8/8'), 7.41 (2H, m, H-9/9').

Literature

- (1) Williams, D. H.; Stone, M. J.; Hauck, P. R.; Rahman, S. K. *J. Nat. Prod.* **1989**, *52*, 1189–1208.
- (2) Keller, N. P.; Turner, G.; Bennett, J. W. *Nat Rev Microbiol* **2005**, *3*, 937–947.
- (3) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2012**, *75*, 311–335.
- (4) De Mattos-Shiple, K.; Ford, K.; Alberti, F.; Banks, A.; Bailey, A.; Foster, G. *Stud. Mycol.* **2016**, *85*, 125–157.
- (5) Hibbett, D. S. *Mycologia* **2006**, *98*, 917–925.
- (6) Anke, T.; Oberwinkler, F.; Steglich, W.; Schramm, G. *J. Antibiot. (Tokyo)*. **1977**, *30*, 806–810.
- (7) Jones, C. A.; Sidebottom, P. J.; Cannell, R. J. P.; Noble, D.; Rudd, B. A. M. *J. Antibiot.* **1992**, *45*, 1492–1498.
- (8) Adelin, E.; Servy, C.; Cortial, S.; Lévaïque, H.; Martin, M.-t.; Retailleau, P.; Le, G.; Bussaban, B.; Lumyong, S.; Ouazzani, J. *Phytochemistry* **2011**, *72*, 2406–2412.
- (9) Dias, D. A.; Urban, S.; Roessner, U. *Metabolites* **2012**, *2*, 303–336.
- (10) Raistrick, H. *Proc. R. Soc. London* **1950**, *136*, 481–508.
- (11) Ji, H.-F.; Li, X.-J.; Zhang, H.-Y. *EMBO Rep.* **2009**, *10*, 194–200.
- (12) Bergstrom, J. D.; Kurtz, M. M.; Rew, D. J.; Amend, a. M.; Karkas, J. D.; Bostedor, R. G.; Bansal, V. S.; Dufresne, C.; VanMiddlesworth, F. L.; Hensens, O. D. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 80–84.
- (13) Alberts, A. W. et al. *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77*, 3957–3961.
- (14) Moore, R. N.; Bigam, G.; Chan, J. K.; Hogg, A. M.; Nakashima, T. T.; Vederas, J. C. *J. Am. Chem. Soc.* **1985**, *107*, 3694–3701.
- (15) Becker, W. F.; von Jagow, G.; Steglich, W. *FEBS Lett.* **1981**, *132*, 329–333.
- (16) Chu, M.; Mierzwa, R.; Xu, L.; He, L.; Terracciano, J.; Patel, M.; Gullo, V.; Black, T.; Zhao, W.; Chan, T. M.; McPhail, A. T. *J. Nat. Prod.* **2003**, *66*, 1527–1530.
- (17) Yu, J.; Bhatnagar, D.; Cleveland, T. E. *FEBS Lett.* **2004**, *564*, 126–130.
- (18) Beck, J.; Ripka, S.; Siegner, A.; Schlitz, E.; Schweizer, E. *Eur. J. Biochem.* **1990**, *192*, 487–498.
- (19) Cane, D. E.; Ha, H. J.; Pargellis, C.; Waldmeier, F.; Swanson, S.; Murthy, P. P. N. *J. Am. Chem. Soc.* **1989**, *111*, 8914–8916.
- (20) Kellner, U.; Zocher, R.; Kleinkauf, H. *J. Gen. Microbiol.* **1980**, *118*, 485–494.
- (21) Birch, A. J.; Massy-Westropp, R. A.; Moye, C. J. *Aust. J. Chem.* **1955**, *8*, 539–544.
- (22) Collie, J. J. *Chem. Soc. Trans.* **1907**, *91*, 1806.
- (23) Raper, H. J. *Chem. Soc. Trans.* **1907**, *91*, 1831.
- (24) Hopwood, D. A.; Sherman, D. H. *Annu. Rev. Genet.* **1990**, *24*, 2–3.
- (25) Beld, J.; Lee, D. J.; Burkart, M. D. *Mol. Biosyst.* **2015**, *11*, 38–59.
- (26) Hemmerling, F.; Lebe, K. E.; Wunderlich, J.; Hahn, F. *ChemBioChem* **2018**, *19*, 1006–1011.
- (27) Berkhan, G.; Merten, C.; Holec, C.; Hahn, F. *Angew. Chemie Int. Ed.* **2016**, *1–5*.
- (28) Staunton, J.; Weissman, K. J. *Nat. Prod. Rep.* **2001**, *18*, 380–416.
- (29) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5*, 2010–2026.
- (30) Hertweck, C. *Angew. Chemie - Int. Ed.* **2009**, *48*, 4688–4716.
- (31) Weissman, K. J. *Nat. Prod. Rep.* **2015**, *32*, 436–453.

- (32) Maier, T.; Leibundgut, M.; Ban, N. *Methods* **2008**, *1315*, 1–26.
- (33) Smith, S.; Tsai, S. C. *Nat. Prod. Rep.* **2007**, *24*, 1041–1072.
- (34) Shen, B. *Curr. Opin. Chem. Biol.* **2003**, *7*, 285–295.
- (35) Weissman, K. J.; Leadlay, P. F. *Nat. Rev. Microbiol.* **2005**, *3*, 925–936.
- (36) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. *Nature* **1990**, *348*, 176–178.
- (37) Cane, D. E. *J. Biol. Chem.* **2010**, *285*, 27517–27523.
- (38) Chooi, Y. H.; Tang, Y. *J. Org. Chem.* **2012**, *77*, 9933–9953.
- (39) Kroken, S.; Glass, N. L.; Taylor, J. W.; Yoder, O. C.; Turgeon, B. G. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 15670–15675.
- (40) Cacho, R. A.; Thuss, J.; Xu, W.; Sanichar, R.; Gao, Z.; Nguyen, A.; Vederas, J. C.; Tang, Y. *J. Am. Chem. Soc.* **2015**, *137*, 15688–15691.
- (41) Cox, R. J.; Glod, F.; Hurley, D.; Lazarus, C. M.; Nicholson, T. P.; Rudd, B. a. M.; Simpson, T. J.; Wilkinson, B.; Zhang, Y. *Chem. Commun.* **2004**, 2260–2261.
- (42) Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vederas, J. C.; Hutchinson, C. R. *Science* **1999**, *284*, 1368–1372.
- (43) Herbert, R. B. *Nat. Prod. Rep.* **1992**, *9*, 507–529.
- (44) Baloglu, E.; Kingston, D. G. *J. Nat. Prod.* **1999**, *62*, 1448–1472.
- (45) Piel, J.; Hertweck, C.; Shipley, P. R.; Hunt, D. M.; Newman, M. S.; Moore, B. S. *Chem. Biol.* **2000**, *7*, 943–955.
- (46) Hill, A. M.; Thompson, B. L.; Harris, J. P.; Segret, R. *Chem. Commun.* **2003**, *3*, 1358–1359.
- (47) Byrne, K. M.; Arison, B. H.; Nallin-Omstead, M.; Kaplan, L. *J. Org. Chem.* **1993**, *58*, 1019–1024.
- (48) Nerud, F.; Sedmera, P.; Zouchov, Z.; Muslilek, V.; Vondracek, M. *Collect. Czechoslov. Chem Commun.* **1982**, *47*, 1020–1025.
- (49) Olano, C.; Carmen, M.; Méndez, C.; Salas, J. A. *Nat. Prod. Rep.* **2010**, *27*, 571–616.
- (50) Rix, U.; Fischer, C.; Remsing, L. L.; Rohr, J. *J. Nat. Prod. Rep.* **2002**, *19*, 542–580.
- (51) Xu, Y.; Zhou, T.; Espinosa-Artiles, P.; Tang, Y.; Zhan, J.; Molnár, I. *ACS Chem. Biol.* **2014**, *9*, 1119–1127.
- (52) Xie, X.; Watanabe, K.; Wojcicki, W. A.; Wang, C. C. C.; Tang, Y. *Chem. Biol.* **2006**, *13*, 1161–1169.
- (53) Bonsch, B.; Belt, V.; Bartel, C.; Duensing, N.; Koziol, M.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. *Chem. Commun.* **2016**, *52*, 6777–6780.
- (54) Bugg, T. D. *Tetrahedron* **2003**, *59*, 7075–7101.
- (55) Badieyan, S.; Bach, R. D.; Sobrado, P. J. *J. Org. Chem.* **2015**, *80*, 2139–2147.
- (56) Van Berkel, W. J.; Kamerbeek, N. M.; Fraaije, M. W. *J. Biotechnol.* **2006**, *124*, 670–689.
- (57) Cai, M.; Davison, J.; Cox, R. J.; Simpson, T. J.; al Fahad, A.; Bailey, A. M.; Yehia, S. Y.; Song, Z.; Lazarus, C. M. *Proc. Natl. Acad. Sci.* **2012**, *109*, 7642–7647.
- (58) Rudolf, J. D.; Chang, C. Y.; Ma, M.; Shen, B. *Nat. Prod. Rep.* **2017**, *34*, 1141–1172.
- (59) De Montellano, P. R. *Chem. Rev.* **2009**, *110*, 932–948.
- (60) Munro, A. W.; Girvan, H. M.; Mason, A. E.; Dunford, A. J.; McLean, K. J. *Trends Biochem. Sci.* **2013**, *38*, 140–150.
- (61) Cacho, R. A.; Chooi, Y. H.; Zhou, H.; Tang, Y. *ACS Chem. Biol.* **2013**, *8*, 2322–2330.
- (62) Grandner, J. M.; Cacho, R. A.; Tang, Y.; Houk, K. N. *ACS Catal.* **2016**, *6*, 4506–4511.

- (63) Wu, L. F.; Meng, S.; Tang, G. L. *Biochim. Biophys. Acta - Proteins Proteomics* **2016**, *1864*, 453–470.
- (64) Gao, S. S.; Naowarajna, N.; Cheng, R.; Liu, X.; Liu, P. *Nat. Prod. Rep.* **2018**, *35*, 792–837.
- (65) Hausinger, R. P. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 21–68.
- (66) Tang, M. C.; Zou, Y.; Watanabe, K.; Walsh, C. T.; Tang, Y. *Chem. Rev.* **2017**, *117*, 5226–5333.
- (67) Martinez, S.; Hausinger, R. P. *J. Biol. Chem.* **2015**, *290*, 20702–20711.
- (68) Vaillancourt, H.; Yin, J.; Walsh, C. T. **2005**, *102*, 10111–10116.
- (69) Blasiak, L. C.; Vaillancourt, F. H.; Walsh, C. T.; Drennan, C. L. *Nature* **2006**, *440*, 368–371.
- (70) Rinkel, J.; Dickschat, J. S. *Beilstein J. Org. Chem.* **2015**, *11*, 2493–2508.
- (71) Trenti, F.; Cox, R. J. *J. Nat. Prod.* **2017**, *80*, 1235–1240.
- (72) Simpson, T. J. *Chem. Soc. Rev.* **1987**, *16*, 123.
- (73) Vederas, J. C. *Can. J. Chem.* **1982**, *60*, 1637–1642.
- (74) Ahmed, S. A.; Simpson, T. J.; Staunton, J.; Sutkowski, A. C.; Trimble, L. A.; Vederas, J. C. *J. Chem. Soc., Chem. Commun.* **1985**, 1685–1687.
- (75) Weld, R. J.; Plummer, K. M.; Carpenter, M. A.; Ridgway, H. J. *Cell Res.* **2006**, *16*, 31–44.
- (76) Chakraborty, B. N.; Kapoor, M. *Nucleic Acids Res.* **1990**, *18*, 6737.
- (77) Groot, M. J. A. D.; Bundock, P.; Hooykaas, P. J. J.; Beijersbergen, A. G. M. *Nat. Biotechnol.* **1998**, *16*, 839–842.
- (78) Gritz, L.; Davies, J.; A, B. S.; Acacias, R. *Gene* **1983**, *25*, 179–188.
- (79) Kück, U.; Hoff, B. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 51–62.
- (80) Fairhead, C.; Llorente, B.; Denis, F.; Soler, M.; Dujon, B. *Yeast Funct. Anal. Reports* **1996**, *12*, 1439–1457.
- (81) Nielsen, M. L.; Albertsen, L.; Lettier, G.; Nielsen, J. B.; Mortensen, U. H. *Fungal Genet. Biol.* **2006**, *43*, 54–64.
- (82) Liu, G.; Casqueiro, J.; Bañuelos, O.; Cardoza, R. E.; Gutiérrez, S.; Martín, J. F. *J. Bacteriol.* **2001**, *183*, 1765–1772.
- (83) Catlett, N. L.; Lee, B.-N.; Yoder, O. C.; Turgeon, B. G. *Fungal Genet. News* **2002**, *2*, 9–11.
- (84) Williams, K.; Szwalbe, A. J.; Mulholland, N. P.; Vincent, J. L.; Bailey, A. M.; Willis, C. L.; Simpson, T. J.; Cox, R. J. *Angew. Chemie - Int. Ed.* **2016**, *55*, 6784–6788.
- (85) Trenti, F. The Biosynthesis of Phyllostictine A and Sch-642305 from *Phyllosticta cirsii* and *Phomopsis CMU-LMA.*, Ph.D. Thesis, 2018.
- (86) Nakayashiki, H. *FEBS Lett.* **2005**, *579*, 5950–5957.
- (87) Eley, K. L.; Halo, L. M.; Song, Z.; Powles, H.; Cox, R. J.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J. *ChemBioChem* **2007**, *8*, 289–297.
- (88) Halo, L. M.; Heneghan, M. N.; Yakasai, A. A.; Song, Z.; Williams, K.; Bailey, A. M.; Cox, R. J.; Lazarus, C. M.; Simpson, T. J. *J. Am. Chem. Soc.* **2008**, *130*, 17988–17996.
- (89) He, Y.; Wang, B.; Chen, W.; Cox, R. J.; He, J.; Chen, F. *Biotechnol. Adv.* **2018**, *36*, 739–783.
- (90) Schor, R.; Schotte, C.; Wibberg, D.; Kalinowski, J.; Cox, R. J. *Nat. Commun.* **2018**, *9*, 1–9.
- (91) Weissman, K. J. *Nat. Prod. Rep.* **2016**, *33*, 203–230.
- (92) Jones, A. C.; Gust, B.; Kulik, A.; Heide, L.; Buttner, M. J.; Bibb, M. J. *PLoS One* **2013**, *8*, e69319.
- (93) Huang, F.; Fu, J.; Bian, X.; Xia, L.; Hu, S.; Stewart, A. F.; Müller, R.; Plaza, A.; Zhang, Y.; Wang, H.; Seibert, P. M. *Nat. Biotechnol.* **2012**, *30*, 440–446.

- (94) Heneghan, M. N.; Yakasai, A. A.; Halo, L. M.; Song, Z.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; Lazarus, C. M. *ChemBioChem* **2010**, *11*, 1508–1512.
- (95) Anyaogu, D. C.; Mortensen, U. H. *Front. Microbiol.* **2015**, *6*, 1–6.
- (96) Miranda, M. et al. *Sci. Adv.* **2018**, *4*, eaar5459.
- (97) Wasil, Z.; Pahirulzaman, K. A. K.; Butts, C. P.; Simpson, T. J.; Lazarus, C. M.; Cox, R. J. *Chem. Sci.* **2013**, *4*, 3845–3856.
- (98) Lazarus, C. M.; Williams, K.; Bailey, a. M. *Nat. Prod. Rep.* **2014**, *31*, 1339–1347.
- (99) Skellam, E. *Trends Biotechnol.* **2018**, *37*, 416–427.
- (100) Alberti, F.; Foster, G. D.; Bailey, A. M. *Appl. Microbiol. Biotechnol.* **2016**, 493–500.
- (101) Barbesgaard, P.; Heldt-Hansen, H. P.; Diderichsen, B. *Appl. Microbiol. Biotechnol.* **1992**, *36*, 569–572.
- (102) Matsuda, Y.; Bai, T.; Phippen, C. B.; Nødvig, C. S.; Kjærboelling, I.; Vesth, T. C.; Andersen, M. R.; Mortensen, U. H.; Gotfredsen, C. H.; Abe, I.; Larsen, T. O. *Nat. Commun.* **2018**, *9*, 1–10.
- (103) Bai, T.; Quan, Z.; Zhai, R.; Awakawa, T.; Matsuda, Y.; Abe, I. *Org. Lett.* **2018**, *20*, 7504–7508.
- (104) Cao, Z.; Li, S.; Lv, J.; Gao, H.; Chen, G.; Awakawa, T.; Abe, I.; Yao, X.; Hu, D. *Acta Pharm. Sin. B* **2019**, *9*, 433–442.
- (105) Araki, Y.; Awakawa, T.; Matsuzaki, M.; Cho, R.; Matsuda, Y.; Hoshino, S.; Shinohara, Y.; Yamamoto, M.; Kido, Y.; Inaoka, D. K.; Nagamune, K.; Ito, K.; Abe, I.; Kita, K. *Proc. Natl. Acad. Sci.* **2019**, *116*, 8269–8274.
- (106) Hu, Z.; Awakawa, T.; Ma, Z.; Abe, I. *Nat. Commun.* **2019**, *10*, 1–10.
- (107) Lebe, K. E.; Cox, R. J. *Chem. Sci.* **2019**, *10*, 1227–1231.
- (108) Dawson, M. J.; Farthing, J. E.; Marshall, P. S.; Middleton, R. F.; O'Neill, M. J.; Shuttleworth, A.; Stylli, C.; Tait, R. M.; Taylor, P. M.; Wildman, H. G.; Buss, A. D.; Langley, D.; Hayes, M. V. *J. Antibiot. (Tokyo)*. **1992**, *46*, 639–647.
- (109) Hasumi, K.; Tachikawa, K.; Sakai, K.; Murakawa, S.; Yoshikawa, N.; Kumazawa, S.; Endo, A. *J. Antibiot. (Tokyo)*. **1993**, *46*, 689–691.
- (110) Bills, G. F.; Peláez, F.; Polishook, J. D.; Diez-Matas, M. T.; Harris, G. H.; Clapp, W. H.; Dufresne, C.; Byrne, K. M.; Nallin-Omstead, M.; Jenkins, R. G.; Mojena, M.; Huang, L.; Bergstrom, J. D. *Mycol. Res.* **1994**, *98*, 733–739.
- (111) Hensens, O. D.; Dufresne, C.; Liesch, J. M.; Zink, D. L.; Reamer, R. A.; Vanmiddlesworth, F. **1993**, *34*, 399–402.
- (112) Dufresne, C. et al. **1992**, *48*, 10221–10226.
- (113) Bergstrom, J. D.; Dufresne, C.; Bills, G. F.; Nallin-Omstead, M.; Byrne, K. *Annu Rev Microbiol* **1995**, *49*, 607–639.
- (114) Goldstein, J. L.; Brown, M. S. *Nature* **1990**, *343*, 425–430.
- (115) Charlton-Menys, V.; Durrington, P. N. *Drugs* **2007**, *67*, 11–16.
- (116) Baxter, A.; Fitzgerald, B. J.; Hutson, J. L.; McCarthy, A. D.; Motteram, J. M.; Ross, B. C.; Sapra, M.; Snowden, M. A.; Watson, N. S.; Williams, R. J.; Wright, C. J. *Biol. Chem.* **1992**, *267*, 11705–11708.
- (117) Nadin, A.; Nicolaou, K. C. *Angew. Chemie Int. Ed. English* **1996**, *35*, 1622–1656.
- (118) Liu, N.; Hung, Y.-S.; Gao, S.-S.; Hang, L.; Zou, Y.; Chooi, Y.-H.; Tang, Y. *Org. Lett.* **2017**, *19*, 3560–3563.

- (119) Harris, G. H.; Dufresne, C.; Joshua, H.; Koch, L. A.; Zink, D. L.; Salmon, P. M.; Göklen, K. E.; Kurtz, M. M.; Rew, D. J.; Bergstrom, J. D.; Wilson, K. E. *Bioorganic Med. Chem. Lett.* **1995**, *5*, 2403–2408.
- (120) Nicholson, T. P.; Rudd, B. A. M.; Dawson, M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *Chem. Biol.* **2001**, *8*, 157–178.
- (121) Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. a.; Breitling, R.; Takano, E.; Weber, T. *Nucleic Acids Res.* **2013**, *41*, 204–212.
- (122) Böhl, O. Investigation of secondary metabolite biosynthetic gene clusters in fungi., Master's Thesis, Leibniz University Hanover, 2018.
- (123) Skellam, E. J.; Hurley, D.; Davison, J.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *Mol. Biosyst.* **2010**, *6*, 680–682.
- (124) Qualley, A. V.; Widhalm, J. R.; Adebesin, F.; Kish, C. M.; Dudareva, N. *Proc. Natl. Acad. Sci.* **2012**, *109*, 16383–16388.
- (125) Hertweck, C.; Moore, B. S. *Tetrahedron* **2000**, *56*, 9115–9120.
- (126) Schor, R.; Cox, R. *Nat. Prod. Rep.* **2018**, *35*, 230–256.
- (127) Wijkman, N. *Justus Liebigs Ann. Chem.* **1931**, *485*, 61–73.
- (128) Raistrick, H.; Smith, G. *Biochem. J.* **1933**, *27*, 1814–1819.
- (129) Futagawa, M.; Wedge, D. E.; Dayan, F. E. *Pestic. Biochem. Physiol.* **2002**, *73*, 87–93.
- (130) Szwalbe, A. J.; Williams, K.; O'Flynn, D. E.; Bailey, A. M.; Mulholland, N. P.; Vincent, J. L.; Willis, C. L.; Cox, R. J.; Simpson, T. J. *Chem. Commun.* **2015**, *51*, 17088–17091.
- (131) Hamor, T. A.; Paul, I. C.; Monteath Robertson, J.; Sim, G. A. *Experientia* **1962**, *18*, 352–354.
- (132) Bloomer, B. J. L.; Sutherland, J. K. J. *Chem. Soc.* **1968**, 588–591.
- (133) Fujii, R.; Matsu, Y.; Minami, A.; Nagamine, S.; Takeuchi, I.; Gomi, K.; Oikawa, H. *Org. Lett.* **2015**, *17*, 5658–5661.
- (134) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. *Nat. Protoc.* **2015**, *10*, 845–858.
- (135) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, 403–410.
- (136) Wong, C.; Fujimori, D. G.; Walsh, C. T.; Drennan, C. L. *J. Am. Chem. Soc.* **2009**, *131*, 4872–4879.
- (137) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. *Mol. Syst. Biol.* **2011**, *7*, 1–6.
- (138) Li, W.; Cowley, A.; Uludag, M.; Gur, T.; McWilliam, H.; Squizzato, S.; Park, Y. M. i.; Buso, N.; Lopez, R. *Nucleic Acids Res.* **2015**, *43*, W580–W584.
- (139) Matsuda, Y.; Wakimoto, T.; Mori, T.; Awakawa, T.; Abe, I. *J. Am. Chem. Soc.* **2014**, *136*, 15326–15336.
- (140) Prigge, S. T.; Prigge, S. T.; Kolhekar, A. S.; Eipper, B. A.; Mains, R. E.; Amzel, L. M. *Science* **1997**, *278*, 1300–1305.
- (141) Bradbury, A. F.; Finnie, M. D.; Smyth, D. G. *Nature* **1982**, *298*, 686–688.
- (142) Prigge, S. T.; Kolhekar, A. S.; Eipper, B. A.; Mains, R. E.; Amzel, L. M. *Nat. Struct. Biol.* **1999**, *6*, 976–83.
- (143) Stein, A. Cloning , expression and activity of a copper oxidase involved in the biosynthesis of Squalestatin S1., Master's Thesis, Leibniz Universität Hannover, 2017.
- (144) Hildebrand, A.; Remmert, M.; Biegert, A.; Söding, J. *Proteins Struct. Funct. Bioinforma.* **2009**, *77*, 128–132.

- (145) Siebert, X.; Eipper, B. A.; Mains, R. E.; Prigge, S. T.; Blackburn, N. J.; Amzel, L. M. *Biophys. J.* **2005**, *89*, 3312–3319.
- (146) Webb, B.; Sali, A. *Curr. Protoc. Bioinforma.* **2016**, *54*, 5.6.1–5.6.37.
- (147) Kline, C. D.; Mayfield, M.; Blackburn, N. J. *Biochemistry* **2013**, *52*, 2586–2596.
- (148) Gietz, R. In *Yeast Genet.* Smith, J., Burke, D., Eds.; Humana Press: 2014; Vol. 1205; Chapter Methods in.
- (149) Gietz, R. D.; Schiestl, R. H. *Nat. Protoc.* **2007**, *2*, 35–37.
- (150) Presser, A.; Hüfner, A. *Monatshefte für Chemie* **2004**, *135*, 1015–1022.
- (151) Harris, M.; Harris, C. M. *Pure Appl. Chem.* **1986**, *58*, 283–294.
- (152) Wiegand, G.; Remington, S. J. *Ann. Rev. Biophys. Biophys. Chem.* **1986**, 97–117.
- (153) Mandala, S. M.; Thornton, R. A.; Frommer, B. R.; Dreikorn, S.; Kurtz, M. B. *J. Antibiot. (Tokyo)*. **1997**, *50*, 339–343.
- (154) Harris, G.; Turner Jones, E.; Meinz, M.; Nallin-Omstead, M.; Helms, G.; Bills, G.; Zink, D.; Wilson, K. *Tetrahedron Lett.* **1993**, *34*, 5235–5238.
- (155) He, Y.; Cox, R. J. *Chem. Sci.* **2016**, *1*, 2119–2127.
- (156) Barriuso, J.; Nguyen, D. T.; Li, J. W. H.; Roberts, J. N.; MacNevin, G.; Chaytor, J. L.; Marcus, S. L.; Vederas, J. C.; Ro, D. K. *J. Am. Chem. Soc.* **2011**, *133*, 8078–8081.
- (157) Hu, Y.; Dietrich, D.; Xu, W.; Patel, A.; Thuss, J. A. J.; Wang, J.; Yin, W.-B.; Qiao, K.; Houk, K. N.; Vederas, J. C.; Tang, Y. *Nat Chem Biol.* **2014**, *10*, 552–554.
- (158) Mao, X. M.; Zhan, Z. J.; Grayson, M. N.; Tang, M. C.; Xu, W.; Li, Y. Q.; Yin, W. B.; Lin, H. C.; Chooi, Y. H.; Houk, K. N.; Tang, Y. *J. Am. Chem. Soc.* **2015**, *137*, 11904–11907.
- (159) Hang, L.; Liu, N.; Tang, Y. *ACS Catal.* **2016**, *6*, 5935–5945.
- (160) Watanabe, K.; Wang, C. C. C.; Noguchi, H.; Hotta, K.; Tanaka, H.; Koyama, F.; Ishikawa, N. *Angew. Chemie Int. Ed.* **2014**, *53*, 12880–12884.
- (161) Matsuda, Y.; Iwabuchi, T.; Fujimoto, T.; Awakawa, T.; Nakashima, Y.; Mori, T.; Zhang, H.; Hayashi, F.; Abe, I. *J. Am. Chem. Soc.* **2016**, *138*, 12671–12677.
- (162) Matsuda, Y.; Abe, I. *Nat. Prod. Rep.* **2016**, *33*, 1–28.
- (163) Tang, M.-C.; Zou, Y.; Watanabe, K.; Walsh, C. T.; Tang, Y. *Chem. Rev.* **2017**, *117*, 5226–5333.
- (164) Kovaleva, E. G.; Lipscomb, J. D. *Nat. Chem. Biol.* **2008**, *4*, 186–93.
- (165) Bai, J.; Yan, D.; Zhang, T.; Guo, Y.; Liu, Y.; Zou, Y.; Tang, M.; Liu, B.; Wu, Q.; Yu, S.; Tang, Y.; Hu, Y. *Angew. Chemie - Int. Ed.* **2017**, *56*, 4782–4786.
- (166) Blows, W. M.; Foster, G.; Lane, S. J.; Noble, D.; Piercey, J. E.; Sidebottom, P. J.; Webb, G. *J. Antibiot. (Tokyo)*. **1994**, *47*, 740–754.
- (167) Yoshikawa, S.; Shimada, A. *Chem. Rev.* **2015**, *115*, 1936–1989.
- (168) Rodríguez Couto, S.; Toca Herrera, J. L. *Biotechnol. Adv.* **2006**, *24*, 500–513.
- (169) Ramsden, C. A.; Riley, P. A. *Bioorganic Med. Chem.* **2014**, *22*, 2388–2395.
- (170) Quinlan, R. J. et al. *Proc. Natl. Acad. Sci.* **2011**, *108*, 15079–15084.
- (171) Beeson, W. T.; Phillips, C. M.; Cate, J. H. D.; Marletta, M. A. *J. Am. Chem. Soc.* **2012**, *134*, 890–892.
- (172) Huang, K. X.; Fujii, I.; Ebizuka, Y.; Gomi, K.; Sankawa, U. *J. Biol. Chem.* **1995**, *270*, 21495–21502.
- (173) Vendelboe, T. V.; Harris, P.; Zhao, Y.; Walter, T. S.; Harlos, K.; El Omari, K.; Christensen, H. E. M. *Sci. Adv.* **2016**, *2*, 1–10.

- (174) Piasecki, S. K.; Taylor, C. A.; Detelich, J. F.; Liu, J.; Zheng, J.; Komsoukaniants, A.; Siegel, D. R.; Keatinge-Clay, A. T. *Chem. Biol.* **2011**, *18*, 1331–1340.
- (175) Nofiani, R.; de Mattos-Shiple, K.; Lebe, K. E.; Han, L. C.; Iqbal, Z.; Bailey, A. M.; Willis, C. L.; Simpson, T. J.; Cox, R. J. *Nat. Commun.* **2018**, *9*, 1–11.
- (176) Almohseni, H. A.; Al Mamari, H. H.; Valade, A.; Sintim, H. O.; Hodgson, D. M. *Chem. Commun.* **2018**, *54*, 5354–5356.
- (177) Ozturka, K.; Dupont, P.-Y.; Chettri, P.; McDougal, R.; Böhl, O.; Cox, R. J.; Bradshaw, R. E. *Fungal Biol.* **2019**, *123*, 397–407.
- (178) Anke, T.; Hecht, H. J.; Schramm, G.; Steglich, W. **1979**, 1112–1117.
- (179) Sauter, H.; Steglich, W.; Anke, T. *Angew. Chemie Int. Ed.* **1999**, *38*, 1328–1349.
- (180) Bartlett, D. W.; Clough, J. M.; Godwin, J. R.; Hall, A. A.; Hamer, M.; Parr-Dobrzanski, B. *Pest Manag. Sci.* **2002**, *58*, 649–662.
- (181) Fredenhagen, A.; Kuhn, A.; Peter, H. H.; Cuomoand, V.; Giuliano, U. *J. Antibiot. (Tokyo)*. **1990**, *43*, 2–7.
- (182) Bedford, C. T.; Perry, D.; Sharma, R. K. *Nat. Prod. Res.* **2008**, *22*, 1535–1539.
- (183) Taylor, P.; Street, M.; Wt, L.; Bedford, C. T.; Perry, D.; Sharma, R. K. *Nat. Prod. Res.* **2008**, *22*, 1535–1539.
- (184) Iqbal, Z.; Han, L. C.; Soares-Sello, A. M.; Nofiani, R.; Thormann, G.; Zeeck, A.; Cox, R. J.; Willis, C. L.; Simpson, T. J. *Org. Biomol. Chem.* **2018**, *16*, 5524–5532.
- (185) Nofiani, R. An Investigation of Genes Involved in Strobilurin Biosynthesis., Ph.D. Thesis, University of Bristol, 2016.
- (186) Kraiczky, P.; Haase, U.; Gencic, S.; Flindt, S.; Anke, T.; Brandt, U.; Von Jagow, G. *Eur. J. Biochem.* **1996**, *235*, 54–63.
- (187) Zhou, Y.; Chen, L.; Hu, J.; Duan, H.; Lin, D.; Liu, P.; Meng, Q.; Li, B.; Si, N.; Liu, C.; Liu, X. *Sci. Rep.* **2015**, *5*, 1–10.
- (188) Kandziora, N.; Andexer, J. N.; Moss, S. J.; Wilkinson, B.; Leadlay, P. F.; Hahn, F. *Chem. Sci.* **2014**, *5*, 3563–3567.
- (189) Xie, X.; Cane, D. E. *Biochemistry* **2018**, *57*, 3126–3129.
- (190) Grigorieva, N. Y.; Popovsky, V. A.; Stepanov, A. V.; Lubuzh, E. D. **2010**, *59*, 2086–2093.
- (191) Klodmann, J.; Senkler, M.; Rode, C.; Braun, H.-P. *Plant Physiol.* **2011**, *157*, 587–598.
- (192) Hanson, K.; Neill, J. A. O.; Simpson, T. J.; Willis, C. L. *J. Chem. Soc. Perkin Trans.* **1994**, *1*, 2493–2497.
- (193) Long, M.; Tao, S.; Vega, D.; Jiang, T.; Wen, Q.; Sophia, L. *Nat Chem Biol.* **2017**, *13*, 325–332.
- (194) Zhao, Y.; Such, P.; Ritey, J. *Angew. Chem. Inf. Ed. Engl.* **1992**, *31*, 215–216.
- (195) Robinson, M. W.; Pillinger, K. S.; Graham, A. E. *Tetrahedron Lett.* **2006**, *47*, 5919–5921.
- (196) Sandala, G. M.; Smith, D. M.; Radom, L. *J. Am. Chem. Soc.* **2008**, *130*, 10684–10690.
- (197) Wang, J.; Pichersky, E. *Arch. Biochem. Biophys.* **1999**, *368*, 172–180.
- (198) Berim, A.; Hyatt, D. C.; Gang, D. R. *Plant Physiol.* **2012**, *160*, 1052–1069.
- (199) Beaument, K.; Clough, J. M.; de Fraine, P. J.; Godfrey, C. R. *Pestic. Sci.* **1991**, *31*, 499–519.
- (200) Alessandro, M.; Di Marco, S.; Osti, F.; Cesari, A. *Phytopathol. Mediterr.* **2000**, *39*, 357–365.
- (201) Pezet, R.; Gindro, K.; Viret, O.; Richter, H. *Vitis - J. Grapevine Res.* **2004**, *43*, 145–148.

- (202) Han, J. W.; Ng, B. G.; Sohng, J. K.; Yoon, Y. J.; Choi, G. J.; Kim, B. S. *J. Appl. Microbiol.* **2018**, *124*, 144–154.
- (203) Singleton, V. I.; Bohonos, N.; Ullstrup, A. J. *Nature* **1958**, *181*, 1072–1073.
- (204) Suzuki, Y.; Tanaka, H.; Aoki, H.; Tamura, T. *Agric. Biol. Chem.* **1970**, *34*, 395–413.
- (205) Zabala, A. O.; Chooi, Y.-H.; Choi, M. S.; Lin, H.-C.; Tang, Y. *ACS Chem. Biol.* **2014**, *9*, 1576–1586.
- (206) Betina, V. *Folia Microbiol (Praha)* **1992**, *1*, 3–11.
- (207) Misumi, Y.; Misumi, Y.; Miki, K.; Takatsuki, A.; Tamura, G.; Ikehara, Y. *J. Biol. Chem.* **1986**, *261*, 11398–11403.
- (208) Oda, K.; Hirose, S.; Takami, N.; Misumi, Y.; Takatsuki, A.; Ikehara, Y. *FEBS Lett.* **1987**, *214*, 135–138.
- (209) Kato, S.; Ito, S.; Noguchi, T.; Naito, H. *BBA - Gen. Subj.* **1989**, *991*, 36–43.
- (210) Yamamoto, Y.; Hori, A.; Hutchinson, C. R. *J. Am. Chem. Soc.* **1985**, *107*, 2471–2474.
- (211) Coombe, R. G.; Foss, P. S.; Jacobs, J. J.; Watson, T. R. *Aust. J. Chem.* **1969**, *22*, 1943–1950.
- (212) Mabuni, C. T.; Garlaschelli, L.; Ellison, R. A.; Hutchinson, C. R. *J. Am. Chem. Soc.* **1979**, *101*, 707–714.
- (213) Mehta, G.; Shinde, H. M. *Chem. Commun.* **2005**, 3703–3705.
- (214) Nicoletti, R.; Lopez-Gresa, M. P.; Manzo, E.; Carella, A.; Ciavatta, M. L. *Mycopathologia* **2007**, *163*, 295–301.
- (215) Snider, B. B.; Zhou, J. *Org. Lett.* **2006**, *8*, 1283–1286.
- (216) Bode, H. B.; Walker, M.; Zeeck, A. *European J. Org. Chem.* **2000**, 1451–1456.
- (217) Adelin, E.; Servy, C.; Cortial, S.; Lévaïque, H.; Gallard, J. F.; Martin, M. T.; Retailleau, P.; Bussaban, B.; Lumyong, S.; Ouazzani, J. *Bioorganic Med. Chem. Lett.* **2011**, *21*, 2456–2459.
- (218) Yu, B. Z.; Zhu, N.; Du, Z. Z. *Helv. Chim. Acta* **2010**, *93*, 324–328.
- (219) Fujii, R.; Ugai, T.; Ichinose, H.; Hatakeyama, M.; Kosaki, T.; Gomi, K.; Fujii, I.; Minami, A.; Oikawa, H. *Biosci. Biotechnol. Biochem.* **2016**, *80*, 426–431.
- (220) Fujii, R.; Minami, A.; Gomi, K.; Oikawa, H. *Tetrahedron Lett.* **2013**, *54*, 2999–3002.
- (221) Heneghan, M. N.; Yakasai, A. A.; Williams, K.; Kadir, K. A.; Wasil, Z.; Bakeer, W.; Fisch, K. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; Lazarus, C. M. *Chem. Sci.* **2011**, *2*, 972–979.
- (222) Song, Z.; Bakeer, W.; Marshall, J. W.; Yakasai, A. A.; Khalid, R. M.; Collemare, J.; Skellam, E.; Tharreau, D.; Lebrun, M. H.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. *Chem. Sci.* **2015**, *6*, 4837–4845.
- (223) Gruber, F.; Visser, J.; Kubicek, C. P.; de Graaff, L. H. *Curr. Genet.* **1990**, *18*, 71–76.
- (224) Wang, B.; Xia, L. *Bioresour. Technol.* **2011**, *102*, 4568–4572.
- (225) Fitz, E.; Wanka, F.; Seiboth, B. *Front. Bioeng. Biotechnol.* **2018**, *6*, 1–15.
- (226) Gietz, R. D.; Woods, R. A. *Methods Enzymol.* **2002**, *350*, 87–96.

7 Appendix

7.1 SQS1 – additional data

7.1.1 Genetic analysis of expression transformants

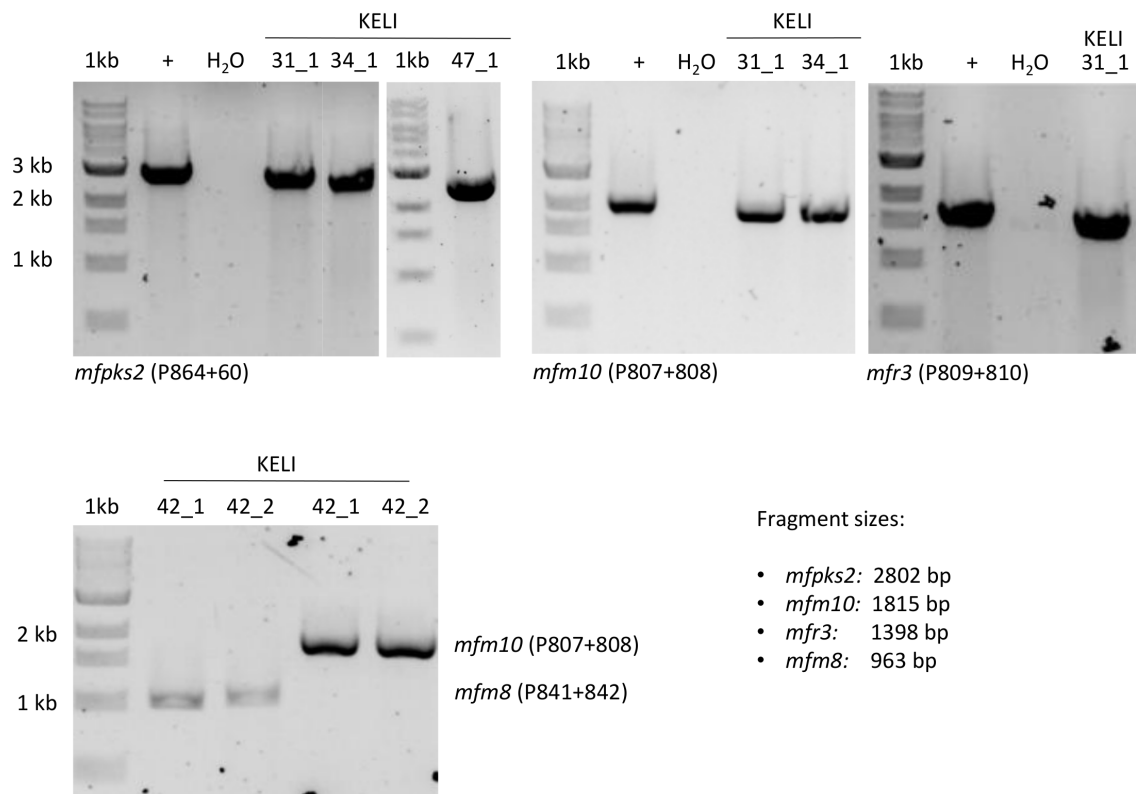


Figure 7.1 Genetic analysis of *A. oryzae* NSAR1 transformants of experiments 1 to 4.

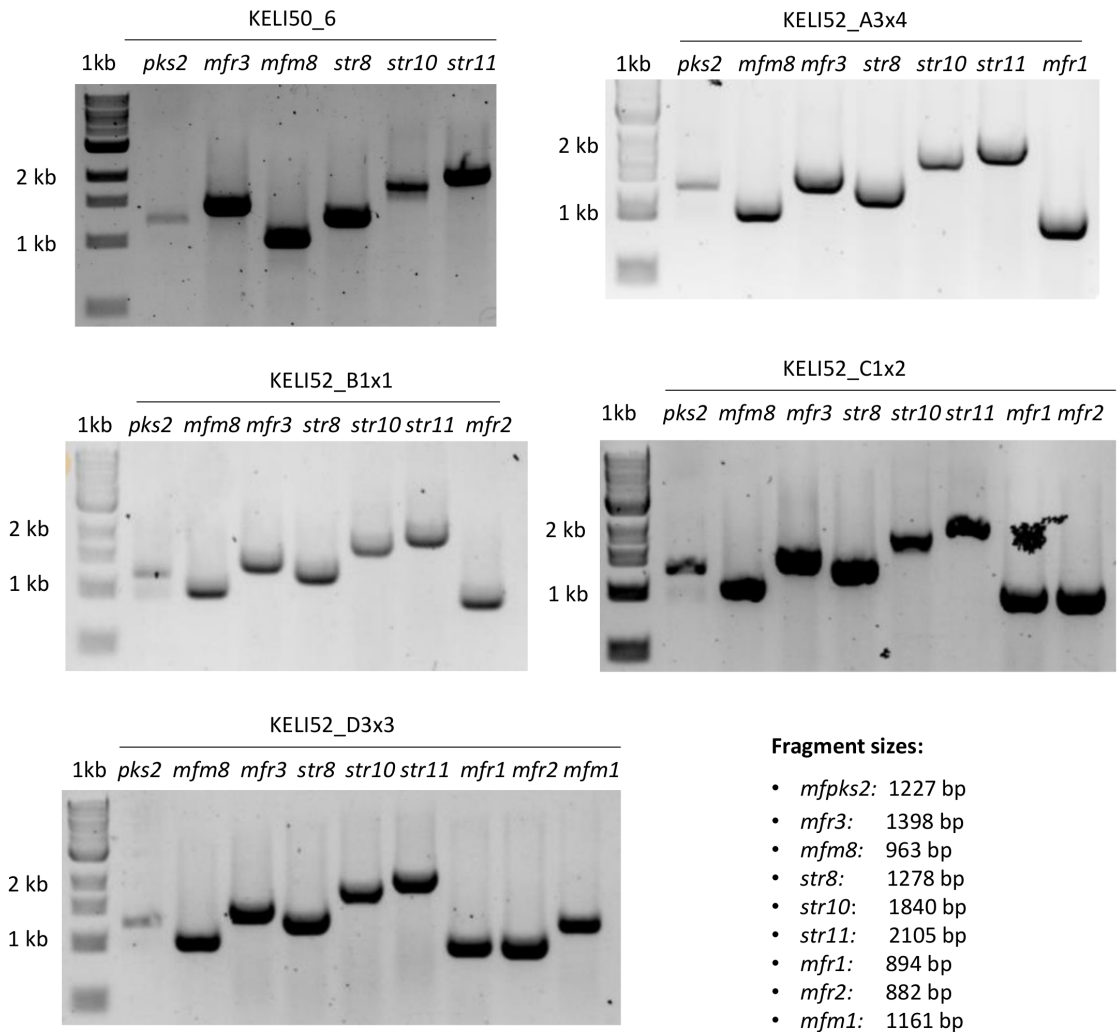


Figure 7.2 PCR analysis of *A. oryzae* NSAR1 transformants of experiments 5 to 10; oligonucleotides: *mfpks2* (P862+863), *mfm8* (P886+887), *mfr3* (P609+610), *str10* (P992+993), *str11* (P988+989), *str8* (P990+991), *mfr1* (P613+614), *mfr2* (P615+616), *mfm1* (P611+612).

7.1.2 Expression of early genes

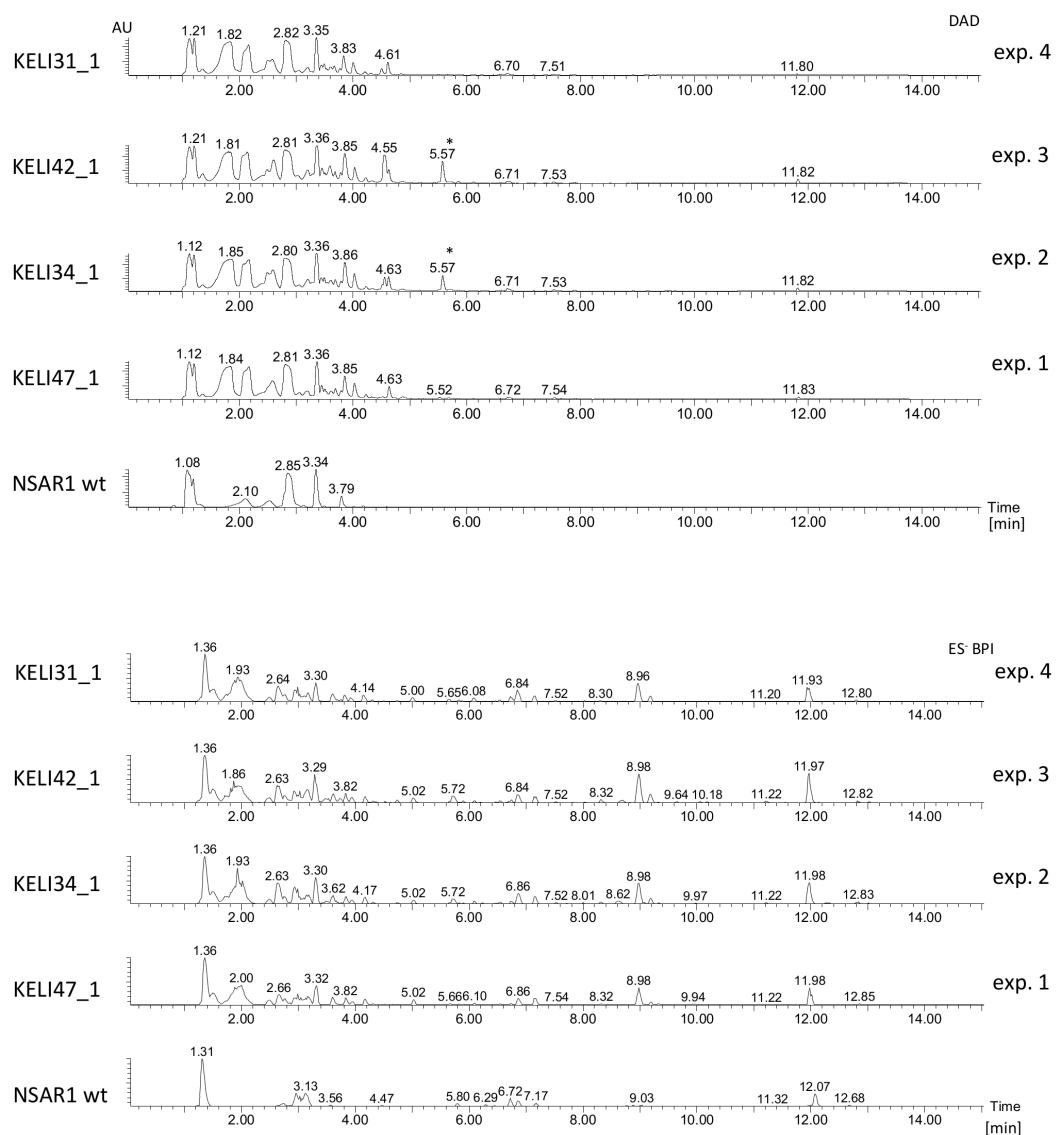


Figure 7.3 Chemical analysis of transformants of experiment 1 to 4 in comparison to *A. oryzae* NSAR1 wt control, DAD (top) and ES- traces (bottom), *unrelated compound ($t_R = 5.57$ min), m/z 163 (ES-), bottom: peak ($t_R = 8.98$ min) also in wt control ($t_R = 9.03$ min).

7.2 Strobilurin – additional data

7.2.1 Genetic analysis of heterologous expression transformants

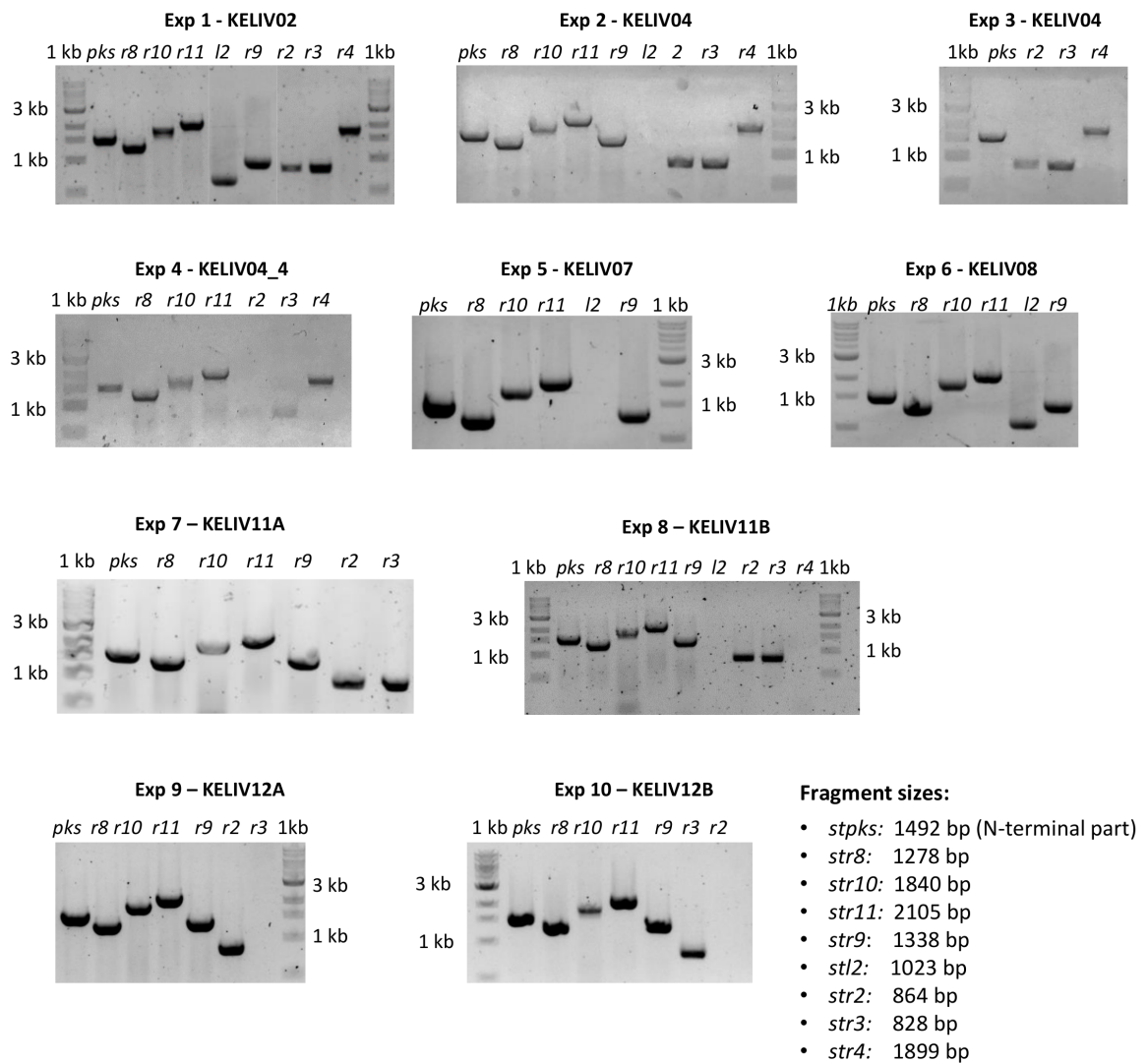


Figure 7.4 Genetic analysis of *A. oryzae* NSAR1 transformants of experiments 1 to 10 from strobilurin heterologous expression experiments.

7.2.2 ESI—MS/MS analysis of Str9

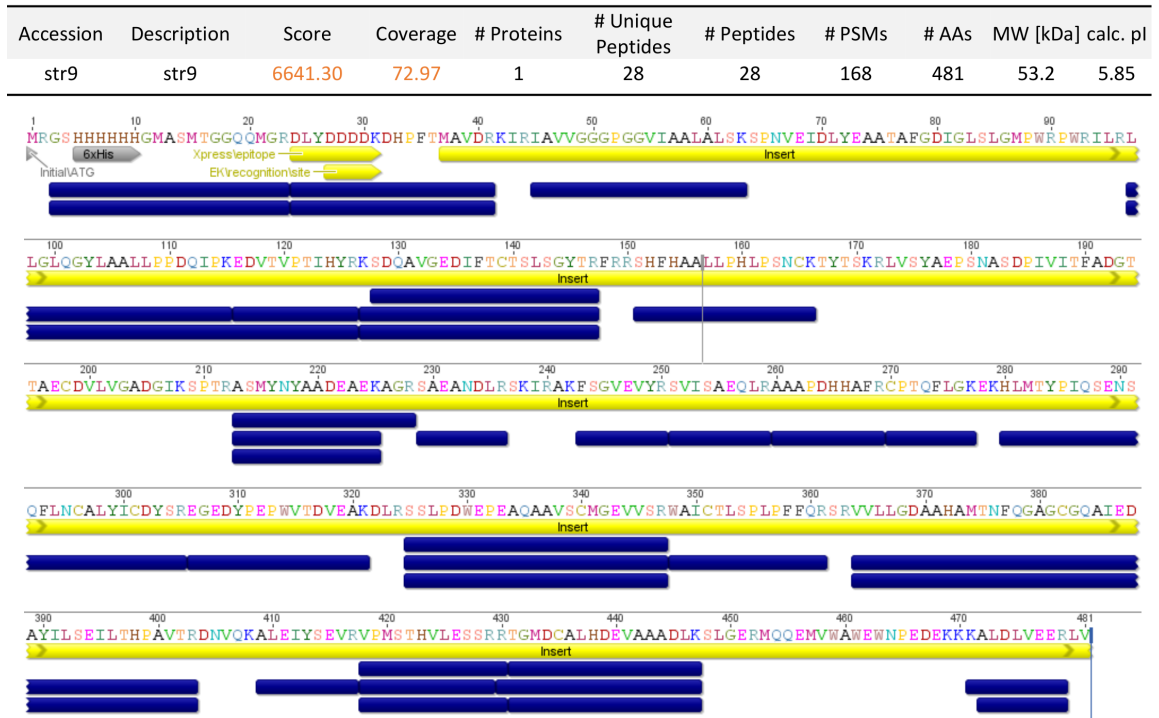


Figure 7.5 ESI—MS/MS analysis of His₆—Str9. The expressed protein band was cut from the SDS PAGE and confirmed by mass spectrometry using the methods of Klodmann and coworkers.¹⁹¹ Blue bars indicate detected peptide fragments.

7.3 SCH – additional data

7.3.1 Expression vectors not used in this thesis

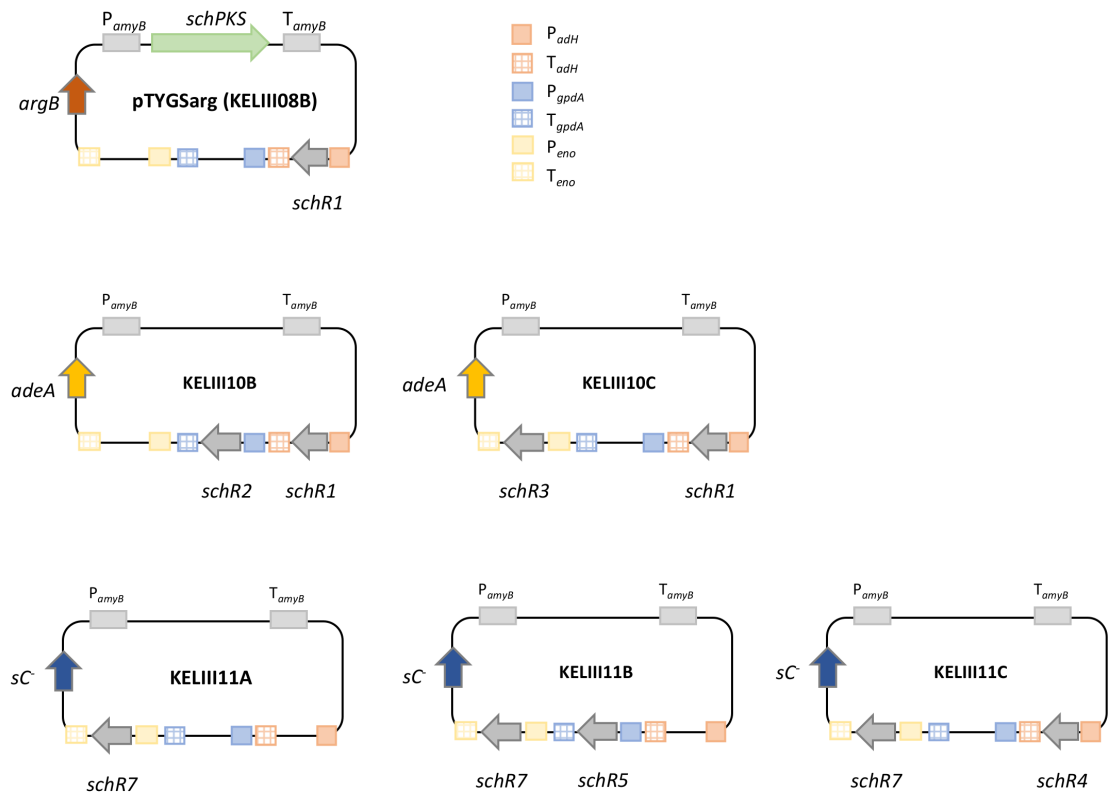


Figure 7.6 Constructed fungal expression vectors for expression of genes from *sch* BGC, which were not used in this study.

7.3.2 Genetic analysis of expression transformants

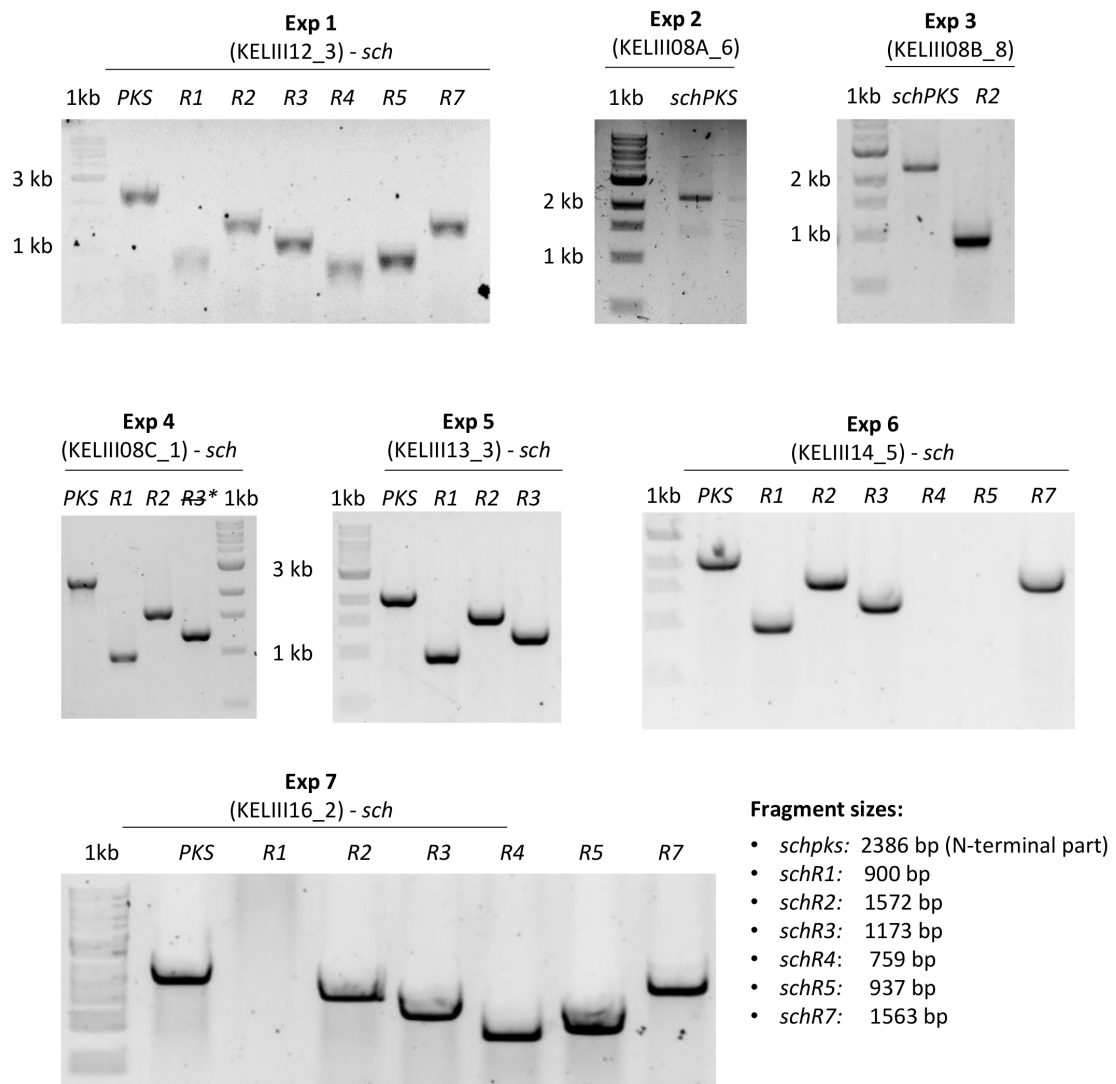


Figure 7.7 PCR analysis of *A. oryzae* NSAR1 transformants of experiments 1 to 7 with indicated gene fragment sizes; oligonucleotides: *schPKS* (P966+1103), *schR1* (P975+976), *schR2* (P978+979), *schR3* (P980+981), *schR4* (P982+983), *schR5* (P984+985), *schR7* (P986+987), *schR3**: intact gene with point mutations.

Curriculum Vitae

Dec 2015 - May 2019	PhD student in Organic Chemistry, Gottfried Wilhelm Leibniz Universität Hannover, Germany
Oct 2013 - Nov 2015	Master of Science , Life Science, Gottfried Wilhelm Leibniz Universität Hannover, Germany
Oct 2010 - Aug 2013	Bachelor of Science , Life Science, Gottfried Wilhelm Leibniz Universität Hannover, Germany
Jun 2010	Abitur , Erich Kästner Gymnasium Laatzen, Germany

List of Publications

1. Lebe, K. E.; Cox, R. J. *Chem. Sci.* **2019**, *10*, 1227-1231.
2. Nofiani, R.; de Mattos–Shiple, K.; Lebe, K. E.; Han, L. C.; Iqbal, Z.; Bailey, A. M.; Willis, C. L.; Simpson, T. J.; Cox, R. J. *Nat. Commun.* **2018**, *9*, 1–11.
3. Hemmerling, F.; Lebe, K. E.; Wunderlich, J.; Hahn, F. *Chembiochem* **2018**, *19*, 1006-1011.
4. Tao, W.; Yurkovich, M. E.; Wen, S.; Lebe, K.E.; Samborsky, M.; Liu, Y.; Yang, A.; Liu, Y.; Ju, Y.; Deng, Z.; Tosin, M.; Sun, Y.; Leadlay, P.F. *Chem. Sci.* **2016**, *7*, 376-385.