Biosynthesis of Xenovulenes in *Acremonium strictum*

Von der Naturwissenschaftlichen Fakultät der

Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von

Raissa Schor, M. Sc.

2018

Referent: Prof. Dr. Russell Cox

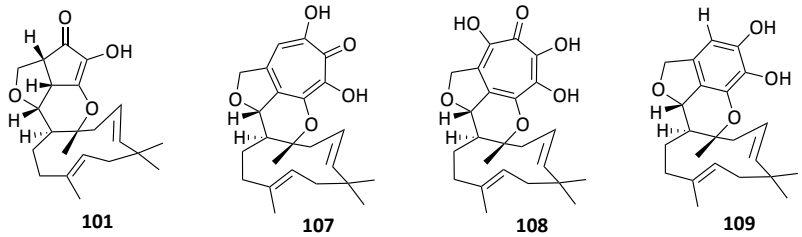
Korreferent: Prof. Dr. Andreas Kirschning

Tag der Promotion: 16. März 2018

Abstract

Keywords: natural products, biosynthesis, ascomycetes, meroterpenoids

Xenovulenes (**101**, **107**, **108** and **109**) are a family of fungal meroterpenoids produced by the ascomycete *Acremonium strictum*. Xenovulene A **101** binds the γ-butyric acid receptor *in vitro*. The compound exhibits an unusual tetrahydrofuro-cyclopentenone attached to a humulene ring, which aroused interest in its biosynthesis. In a combined genetic and chemical approach, the biosynthesis of xenovulene A **101** and co-metabolites **107**-**109** was investigated in this work.



The *A. strictum* genome as well as the transcriptome under **101** producing and non-producing conditions were sequenced. It allowed the *in silico* prediction of biosynthetic gene clusters (BGC) and identified the *aspks1* BGC as a good candidate for **101** biosynthesis.

This was confirmed by knockout of *aspks1* in *A. strictum.* Heterologous co-expression of eight genes of the *aspks1* BGC led to the production of **101** in *Aspergillus oryzae*. Four of the eight genes (*Aspks1*: NR-PKS, *asL1*: FAD dependent monooxygenase, *asL3*: non-heme Fe^{II} dependent dioxygenase, *asR2*: cytochrome P450) were shown to encode proteins homologous to the tropolone pathway in *Talaromyces stipitatus*.

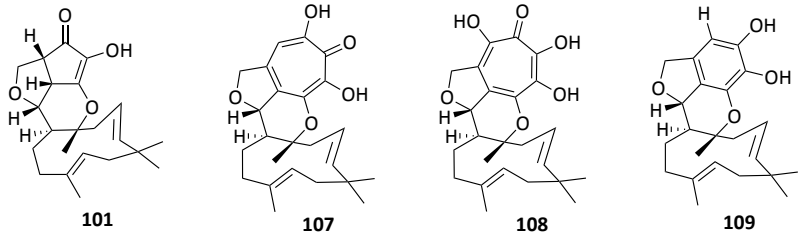
Two genes (*asR5*, *asR6*) encoding proteins of unknown function were identified to be crucial for meroterpenoid production. *In vitro* assays of AsR6 identified its role as an unprecedented humulene synthase, without homologies to any known terpene cyclase. *In vivo* evidence suggested AsR5 to be involved in the joining of polyketide and terpene precursors.

For the two ring contraction steps, two FAD dependent monooxygenases encoded by *asL4* and *asL6* were discovered. *In vivo* evidence suggested these to be aromatic hydroxylases with a distinct regioselectivity.

Zusammenfassung

Schlagwörter: Naturstoffe, Biosynthese, Ascomyceten, Meroterpenoide

Die pilzlichen Naturstoffe der Xenovulene (**101**, **107**, **108** und **109**), produziert von *Acremonium strictum*, gehören zu den Meroterpenoiden. Für Xenovulen A **101** wurde eine *In-vitro*-Bindungsaffinität an den γ-Aminobuttersäure Rezeptor gezeigt. Der Naturstoff besitzt einen seltenen Tetrahydrofurancyclopentenon Bizyklus, welcher über ein Tetrahydropyran mit Humulen verbunden ist. Diese ungewöhnlichen Strukturelemente erweckten Interesse an der Biosynthese dieses Naturstoffes, welche in dieser Arbeit mit chemischen und molekularbiologischen Methoden untersucht wurde.



Zunächst wurden das *A. strictum* Genom und Transkriptom (unter **101** produzierenden und nicht-produzierenden Bedingungen) sequenziert und analysiert. *In silico* wurde ein Biosynthesecluster (*aspks1* BGC) identifiziert, das für essentielle Proteine der **101** Biosynthese kodiert.

Ein Genknockout von *aspks1* in *A. strictum* bestätigte die bioinformatische Hypothese und das BGC. Die heterologe Koexpression von acht Genen des *aspks1* BGC in *Aspergillus oryzae* führte zur Produktion von **101** in *A. oryzae*. Für vier der acht Proteine (*aspks1*: NR-PKS, *asL1*: FAD Monooxygenase, *asL3*: Häm-unabhängige Fe^{II} Dioxygenase, *asR2*: Cytochrome P450) konnten Homolge in der *Talaromyces stipitatus* Tropolonbiosynthese identifiziert werden.

Zwei (*asR5*, *asR6*) der acht Gene kodieren für Proteine mit unbekannter Funktion. *In vitro* Untersuchungen von AsR6 zeigten dessen Rolle als neue Humulensynthase. Im Vergleich mit anderen Terpenzyklasen konnten keine Ähnlichkeiten mit AsR6 festgestellt werden. Damit stellt AsR6 eine neue Klasse von Terpenzyklasen dar. *In-vivo*-Experimente weisen darauf hin, dass AsR5 an der Kopplung von Polyketid und Humulen beteiligt ist.

Für die zwei Ringkontraktionen zu Cyclopentenon wurden zwei FAD abhängige Monooxygenasen (AsL4 und AsL6) identifiziert. *In-vivo*-Experimente weisen darauf hin, dass beide Enzyme Aromaten mit unterschiedlicher Regioselektivität hydroxylieren.

Acknowledgment

I would like to thank Prof Russell Cox for his support and professional supervision throughout my time as a PhD student. I enjoyed working on such an interesting project for the last four years and am very grateful for having had the opportunity to pursue my own scientific ideas. Thank you for your trust, your feedback and inspiring discussions.

Especially, I would like to thank Prof Andreas Kirschning and Prof Tom Simpson for taking the time to be the co-refereers of my PhD thesis. No less I thank Prof Thomas Scheper for being the chair of the examination board and examiner.

Our cooperation partners at the CeBiTEC Bielefeld, especially Prof Jörn Kalinowski and Dr Daniel Wibberg, I would like to thank for all their efforts with sequencing and helpfulness with bioinformatic matters. I am also grateful to Prof Uffe Mortensen from the DTU Denmark for generous vector donation.

At the Leibniz University I would like to thank the AK Braun especially Dr Jennifer Senkler, the AK Brüser especially Prof Thomas Brüser and Dr Patrick Stolle and the AK Turgay especially Heinrich Schäfer and Ingo Hantke.

This work would not have been possible without the BMWZ media kitchen team and all the analytical departments at the OCI, thank you all very much. Especially, I thank Dr Jörg Fohrer and Dr Gerald Dräger for their help with NMR and mass related questions. And Katja, Tjorven, Monika and Anne: thank you so much for making almost everything possible.

I would like to thank all past and current Cox group members, in particular Doug, Liz, Christoph, Eman, Claudio, Katherine, Steffen, Sen, Carsten, Miriam, Lukas and all my students.

Beyond that, I thank Nina, Francesco, Karen, Verena, Gesche, Janina, Janna and Anna for their time and support during the last four years. I also would like to thank Mike, Thea, Regina, Olli, Wojciech and Stefan from the former OCI 3rd floor.

Finally, I would like to thank my family, friends and Jan Christoph, for their support throughout my studies. This thesis would not have been possible without you.

Abbreviations and Units

A	adenine
ACP	acyl carrier protein
Ago	argonaute
asRNA	anti sense RNA
aspks1	*A. strictum* PKS1 gene
ASPM	*A. strictum* production medium
AMA1	autonomous maintaining of plasmids in *Aspergillus*
AMT	*A. tumefaciens* mediated transformation
ASSM	*A. strictum* seed medium
AU	absorption unit
BGC	biosynthetic gene cluster
BLAST	basic local alignment search tool
BLASTp/n	BLASTprotein/nucleotide
AT	acyl transferase
bp	base pair
C	cystosine
cDNA	complementary DNA
CD	conserved domain
CLC	Claisen cyclase
C-MeT	*C*-methyltransferase
CoA	Co-enzyme A
COSY	correlation spectroscopy
CRISPR	clustered interspaced short palindromic repeats
d	day(s)
DA(ase)	Diels Alder(ase)
DAD	diode array detector
DAPI	4,6-diamidino-2-phenlyindole
dATP	deoxyadenosine triphosphate
ddH_2O	double distilled H_20
dDNA	donor DNA
DH	dehydratase
DNA	deoxyribonucleic acid
DMAPP	dimethylallyl pyrophosphate
DMOA	3,5-dimethylorsellinic acid
DTT	dithioeritol
DSB	double strand break
dsRNA	double strand RNA
EDTA	ethylenediaminetetraacetic acid
ER	enoyl reductase
eGFP	enhanced green fluorescent protein
ESI	electron spray ionisation
ELSD	evaporative light scattering detector
$FAD(H_2)$	flavin adenine dinucleotide
FAS	fatty acid synthase
FPP	farnesyl pyrophosphate
G	guanine
GABA	y-ammino buturic acid
GC	gas chromatography
gen^R	geneticin resistance cassette
GFPP	geranylfarnesyl pyrophosphate
GGPP	geranylgeranly pyrophosphate
GPP	geranyl pyrophosphate
gDNA	genomic DNA
GOI	gene of interest
gRNA	guide RNA
hDA(ase)	hetero Diels Alder(ase)
HH	hammer head
HDV	hepatitis delta virus ribozyme
HMBC	heteronuclear multiple bond correlation
HPLC	high performance liquid chromatography
HR	highly reducing
HR	homologous recombination
h	hour(s)
HRMS	high resolution mass spectrometry
HSQC	heteronuclear single quantum correlation
hyg^R	hygromycin B resistance cassette
Hz	hertz
IPP	isopentenyl diphosphate
kb	kilo base pairs
KR	β-ketoreductase
KS	β-ketosynthase
KO	knock out
KOe	knock out by expression
L	litre
LC	liquid chromatography
M	molar
m/z	mass to charge ratio
min	minute
mg	milligram
mL	millilitre
MOS	3-methylorcinaldehyde synthase
mRNA	messenger RNA
MS	mass spectrometry
MS^2	MS/MS
µL	microliter
NAD(P)(H)	nicotinamide adenine dinucleotide (phosphate)
NHEJ	non-homologous end-joining
nm	nanometre
NMR	nuclear magnetic resonance
NOE	nuclear overhauser effect
NR	non-reducing
NRPS	non-ribosomal peptide synthetase
nt	nucleotide(s)
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PAM	protospacer adjacent motif
PCR	polymerase chain reaction
PEG	polyethylene glycol
PKS	polyketide synthase
PP	pyrophosphate
ppm	parts per million
PR	partially reducing
PT	product template
Q-TOF	quadrupole time-of-flight
R	reductive release domain
RT-PCR	*reverse transcriptase* PCR
revGOI	reverse gene of interest
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
rpm	revolutions per minute
SAM	*S*-adenosyl methionine
SAT	starter unit acyl carrier protein transferase
SDR	short chain dehydrogenase/reductase
SDS	sodium dodecyl sulphate
sec	second (s)
siRNA	short interfering RNA
sp.	species
USER™	uracil specific excision reagent
T	thymine
TAE	tris-acetate-EDTA
TE buffer	tris-EDTA buffer
TE	thiolesterase
THF	tetrahydrofuran
TMS	trimethylsilyl
U	uracil
UPLC	ultra-performance liquid chromatography
UV	ultra violet
WT	wild-type

Table of Contents

1 Introduction

In the widest definition, natural products are all substances produced by living organisms. Organic chemists, however, define natural products as small molecules relevant for primary and secondary metabolism. While primary metabolites are essential for the survival of an organism, secondary metabolites are regarded as an accessory that often provides a benefit in competition with other organisms. In contrast to primary metabolites, which are ubiquitously produced over species and kingdoms, secondary metabolites are often exclusively produced by one or a family of related organisms.[1]

Secondary metabolites often display antifungal, antibacterial or even more specific biological activity.[1] In fact, 65% of all newly approved small molecule drugs between 1981 and 2014 are derived from nature or have a natural origin.[2] Typical examples are the bacterial derived vancomycin **1** (antibiotic),[3] the fungal compound lovastatin **2** (cholesterol lowering agent),[4,5] and plant metabolites like morphine **3** (painkiller)[6,7] and menthol **4** (cough suppressant)[8] (Figure 1.1) all of which are of pharmacological significance.

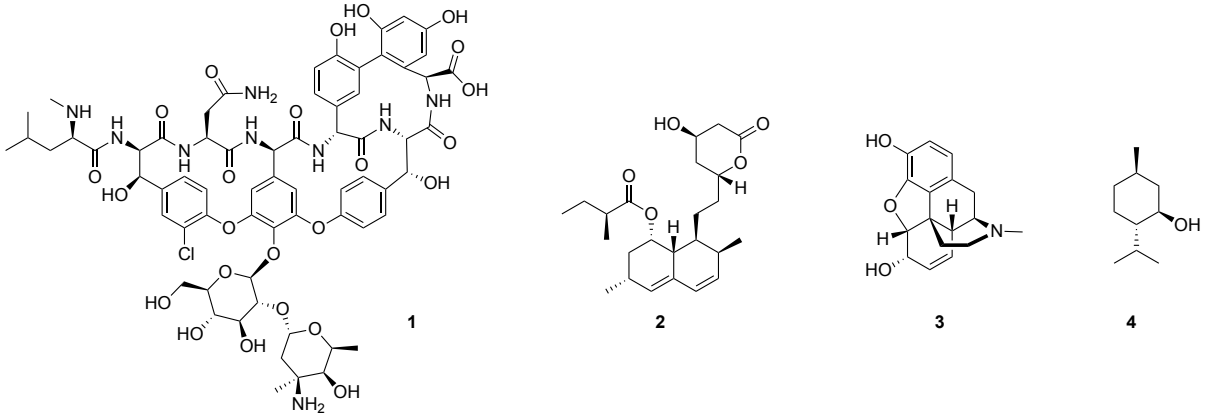
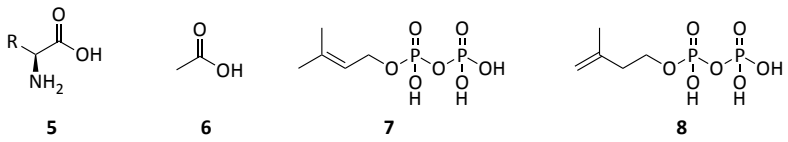


Figure 1.1 Natural products vancomycin **1**, lovastatin **2**, morphine **3** and menthol **4** of bacterial, fungal and plant origin.

These four substances also represent the four major classes of natural products: non-ribosomal peptides (**1**), polyketides (**2**), alkaloids (**3**) and terpenes (**4**).[1] These groups are classified according to the biosynthetic precursors that build up the complex molecules and the enzymes involved. The precursors of non-ribosomal peptides are amino acids **5** which are connected by non-ribosomal peptide syntethases (NRPS).[9,10] The polyketide backbone consists of formal acetate **6** units which are elongated by polyketide synthase (PKS) enzymes.[11,12] The two precursors dimethylallyl pyrophosphate **7** (DMAPP) and isopentenyl pyrophosphate **8** (IPP) combine to form different chain length terpenes which are cyclised to diverse structures by terpene cyclases.[13,14] Alkaloids are nitrogen containing substances normally synthesised from amino acids. The carbon

skeleton often originates from prenylation with the terpene precursor DMAPP **7** such as in indole alkaloids.[15]



5 **6** **7** **8**

Figure 1.2 Small building blocks for natural products.

1.1 Fungal Natural Products in the Genomic Age

Early analysis of fungal biosynthetic pathways was mainly based on feeding studies with labelled precursor molecules. Even though labelling and novel nuclear magnetic resonance (NMR) spectroscopy techniques still contribute and facilitate the understanding of natural products today, it was the development of recombinant deoxyribonucleic acid (DNA) techniques in the late 1980s which promoted the elucidation of biosynthetic pathways at the genetic level.[1] Today molecular biology techniques including gene knockout, silencing and heterologous expression are used to elucidate the function of encoded proteins and link them to the corresponding steps in the biosynthesis of secondary metabolites.[12]

With the development of fast and cheap whole genome sequencing and computational methods for data processing, the number of sequenced fungal genomes has increased rapidly.[16,17] Identification of secondary metabolite genes has thus been greatly simplified over the last decade. Genes required for the biosynthesis of one or a family of metabolites are very frequently co-located in biosynthetic gene clusters (BGC) within fungal genomes.[18,19] Usually such a BGC consists of a core synthase gene, encoding *e.g.* a PKS, NRPS or terpene cyclase, and a number of tailoring genes. The latter can encode a series of proteins such as cytochrome P450s or methyltransferases for further tailoring modifications (Chapter 1.1.3). In addition, genes encoding transporters and transcription regulators are also often clustered. *In silico* analysis of fungal genomes with the antibiotics and secondary metabolites analysis shell (antiSMASH) enables preliminary identification of those BGC.[20–22] However, limited knowledge especially in fungi, restricts the usefulness of purely bioinformatic methods, and significant experimental work is still often required.

1.1.1 Polyketide and Fatty Acid Biosynthesis

Polyketides form a major class of fungal metabolites. Very diverse structural polyketide scaffolds such as lovastatin **2**, but also small aromatic compounds such as 3-methylorsellinic acid **9** are known

(Figure 1.3). Although diverse in structure, these molecules are built up by one class of enzymes, *polyketide synthases*, deploying a restricted subset of reactions.[12]

In the 1950s Birch and co-workers demonstrated for the first time that polyketides are formed through the formal condensation of acetates **6**.[23] This confirmed Collie's hypothesis from 1907 of the eponymous *polyketone* intermediate (**10**) built from C_2 precursors and it also led to Birch realising the relationship between fatty acid and polyketide biosynthesis.[24]

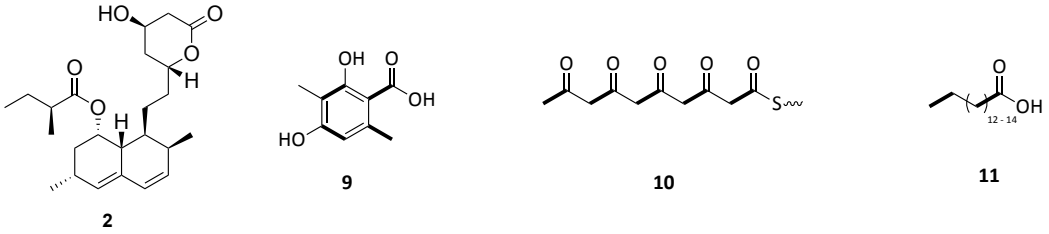


Figure 1.3 Polyketides lovastatin **2** and 3-methylorsellinic acid **9**. Pentaketone **10**, as an example for a polyketone, and fatty acids **11**. Bold bonds resemble acetate C_2 units.

Fatty acids are synthesised by condensation and reduction of starter acetate **6** and elongation malonate **13** units through multifunctional fatty acid synthase (FAS) enzymes. The acetate and malonate units are activated as Co-enzyme A (**18**, CoA) thiolester **12** and **13a**. To initiate the fatty acid synthesis these activated precursors are transferred onto the acyl carrier protein (ACP) of the FAS by an acyltransferase (AT) enzyme activity (Scheme 1.1 Step 0).[25]

The key step to build up the carbon backbone of fatty acids is a decarboxylative Claisen condensation of an acyl **12a** and a malonyl thiolester **13b** catalysed by the ketosynthase (KS) enzyme activity (Scheme 1.1 Step I). After chain extension the growing carbon backbone is covalently bound to the ACP *via* a flexible phosphopantetheine arm (Figure 1.4). This enables the bound β-keto ester (**14**) to access all active sites of the FAS protein for further β-carbonyl modifications (Figure 1.5).[25]

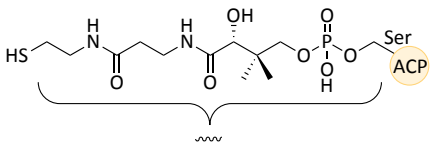
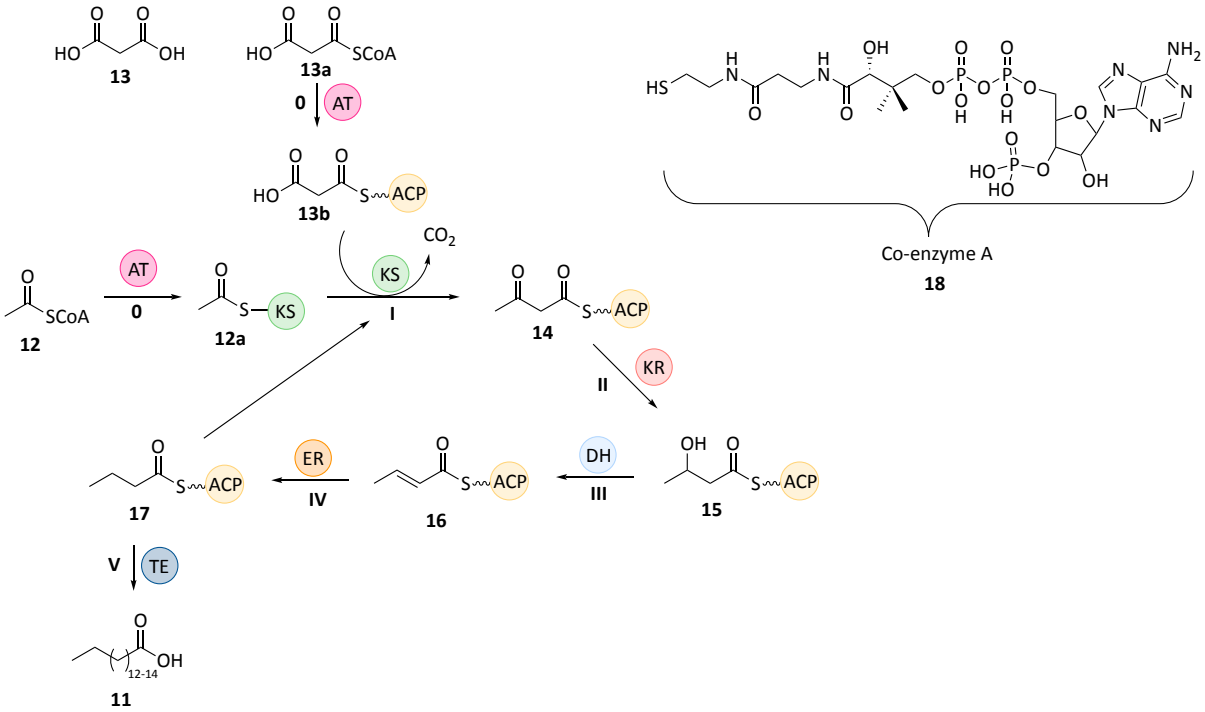


Figure 1.4 Phosphopantetheine arm of ACP.

Ketoreductase (KR), dehydratase (DH) and enoylreductase (ER) catalytic domains then further modify the β-carbonyl of **14** to give the fully saturated thiolester **17** (Scheme 1.1 Steps II-IV). The carbon chain is further elongated (and reduced) by C_2 units until it reaches a length of 16 to 18

carbon atoms. The fully saturated carbon backbone is then released as a free acid **11** by a thiolesterase (TE) domain (Scheme 1.1 Step V).[25]

All known FAS produce fatty acids in an iterative process. This means the individual β-processing domains are reused in each elongation step until the full length fatty acid is produced.[26]



Scheme 1.1 Catalytic cycle of fatty acid chain elongation and structure of Co-enzyme A **18**.

Depending on the organism, FAS proteins differ in structure and are thus divided into type I and II systems. Type I FAS are large multifunctional proteins containing all domain activities for fatty acid biosynthesis in one megasynthase protein.[26,27] In contrast, type II FAS are formed by a set of dissociated monofunctional proteins, which most likely assemble non-covalent complexes similar to type I FAS. However, both FAS architectures catalyse the same set of reactions to build up fatty acids **11** (Scheme 1.1).[28]

In the homodimer crystal structure of mammalian type I FAS a spatial division into condensing and modifying parts of the enzyme is observed (Figure 1.5).[26] KS and AT activities are essential for carbon-carbon bond formation. For further processing of the elongated carbon backbone KR, DH and ER domain activities are required. The essential ACP and TE domains are disordered in the model and hence not visible.[26]

Polyketide synthase (PKS) enzymes not only employ identical chemistry as FAS but are also closely related in structure.[12,31] PKS proteins are also divided in type I (multidomain megasynthases) and

type II systems (sets of monofunctional proteins).[31,32] In addition, ACP-independent type III PKS are reported, which do not resemble a known class of FAS.[33]

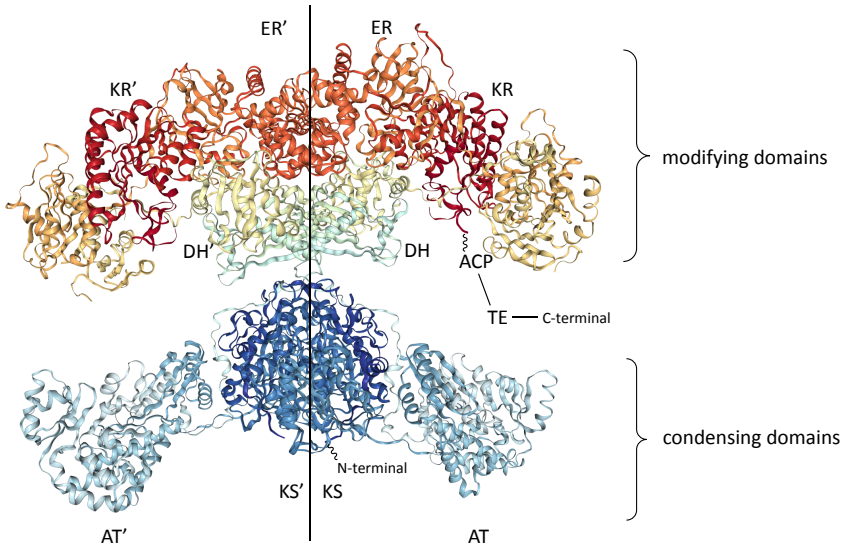


Figure 1.5 Homodimer crystal structure of mammalian FAS at 3.22 Å resolution. PDB entry 2VZ8. Image created with NGL viewer.[29,30]

Most fungal PKS are iterative type I megasynthases.[34] Iterative PKS reuse the same set of domains (like FAS) in each round of elongation (Figure 1.6A). In FAS all β-processing domains are active after each extension step. However, in PKS the β-processing is variable after each extension so that ketones, alcohols, alkenes and methylenes can be formed. This flexibility is *programmed* in the PKS protein itself (Figure 1.6A). This means that based on domain organisation only, no prediction of chain length, degree of saturation and stereochemistry of the final product can be made. A second important class of type I PKS are bacterial modular PKS.[35] The major difference to iterative PKS is that domains are not reused. For each elongation cycle the required subset of domains is organised in one module and multiple modules are present (Figure 1.6B).

Fungal iterative type I PKS are further divided into three functional classes based on the presence or absence of particular catalytic domains.[12,34] PKS producing highly reduced polyketides (highly reducing PKS, HR-PKS) possess the full set of modifying domains (KR, ER, DH) and are known to produce compounds such as squalestatin S1 tetraketide **19** (Figure 1.7A and B).[36] The different degree of saturation (alkene at C-2/3, methylene at C-5, C-7) reflects the programmed activity of domains in each individual elongation cycle.[37] In addition, a common modification in fungal polyketides, which is not observed in fatty acid biosynthesis, is the methylation at α-position by a *S*-adenosyl methionine (SAM, **20**) dependent *C*-methyltransferase (*C*-MeT) (C-4 and C-6 in **19**).[12,34]

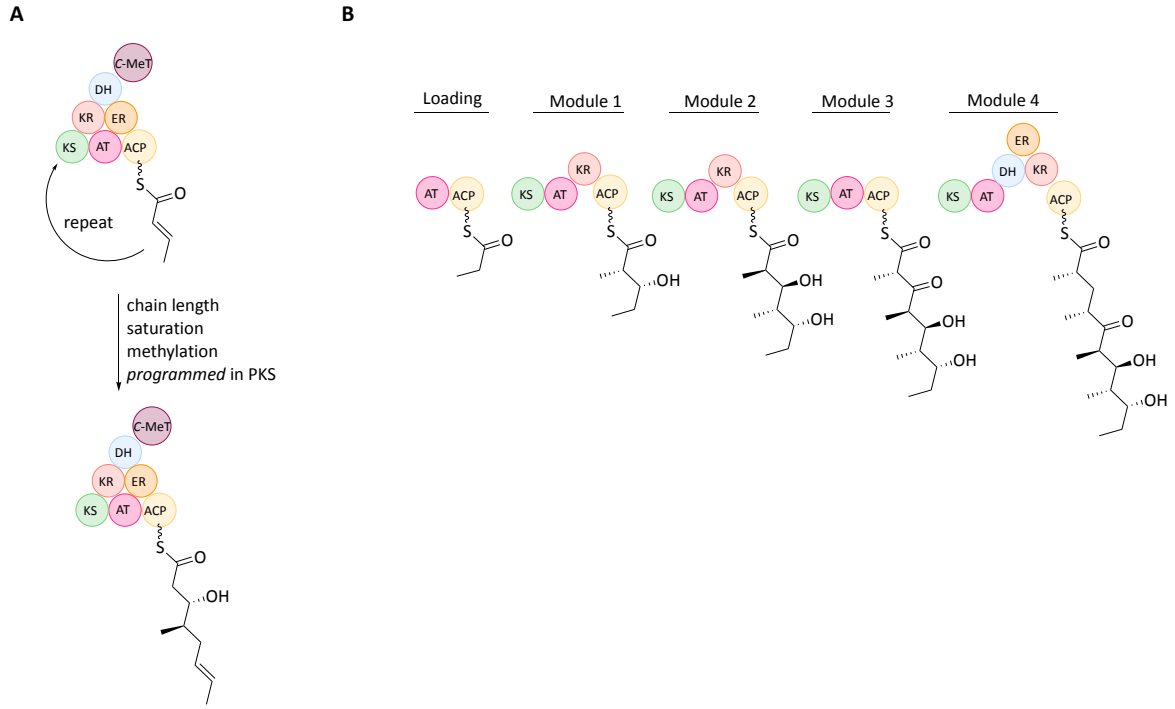


Figure 1.6 Type I PKS. **A**, Iterative PKS. **B**, Modular PKS.

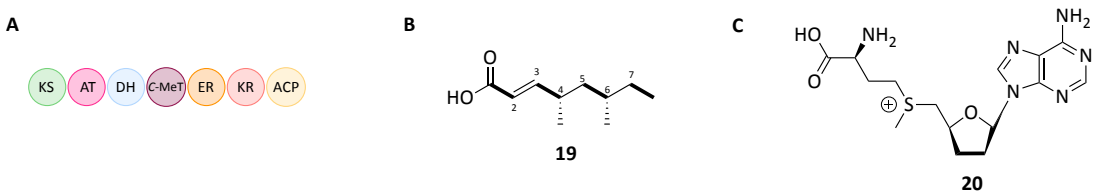
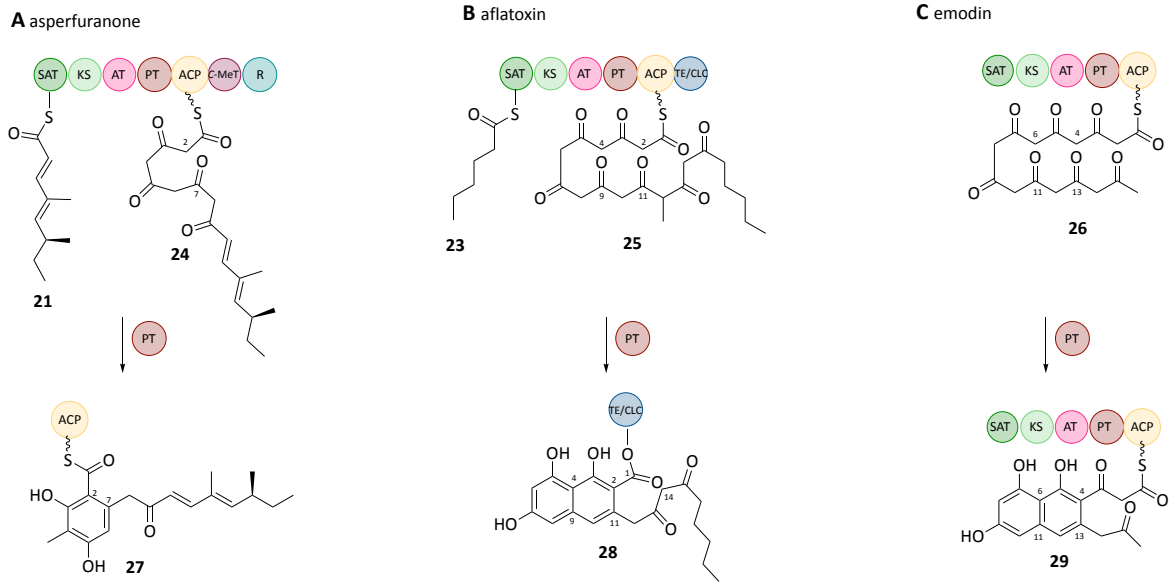


Figure 1.7 HR-PKS and *C*-MeT domain. **A**, Domain organisation of squalestatin tetraketide synthase (SQTKS). **B**, Its product tetraketide **19**. **C**, SAM **20**.

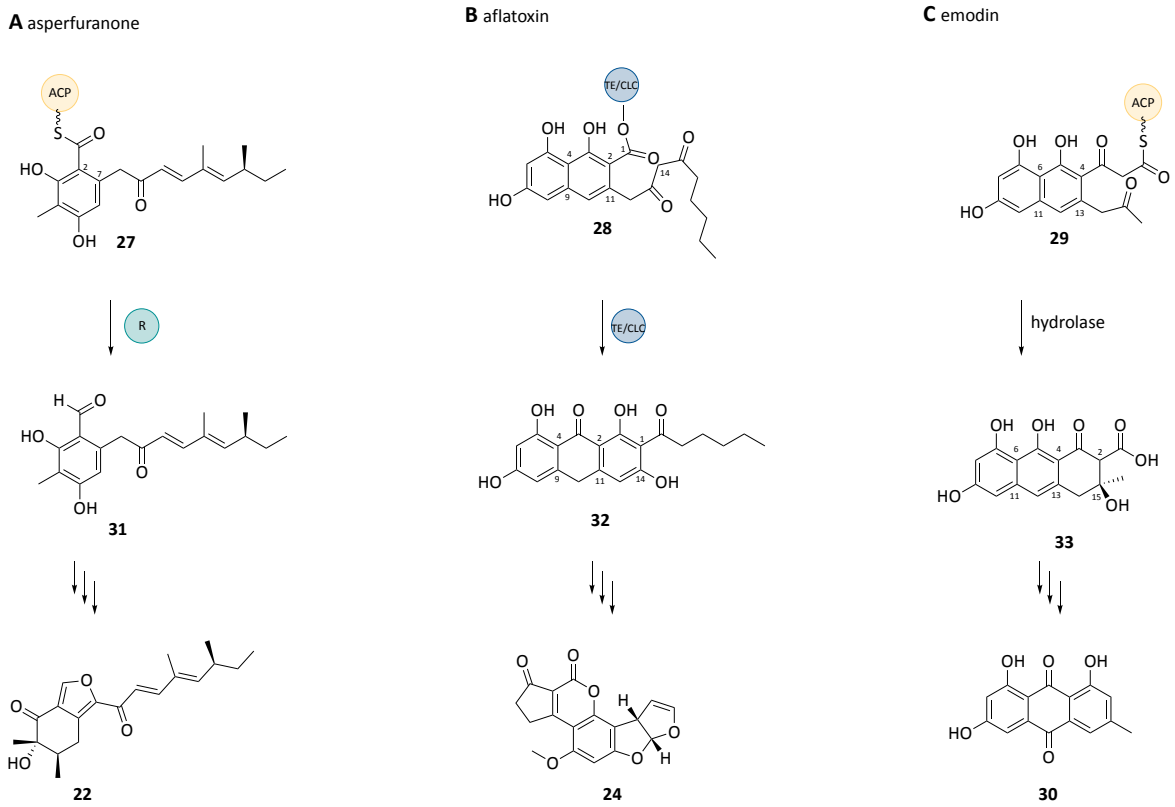
PKS producing non reduced polyketides (non-reducing PKS, NR-PKS) are more different to FAS than HR-PKS as they are lacking the whole set of modifying domains (KR, ER, DH). Instead they usually possess a starter unit acyl transferase (SAT) and a product template (PT) domain.[12,34] The SAT domain transfers more advanced precursors such as the tetraketide **21** in asperfuranone **22** and hexanoic acid **23** in aflatoxin **24** biosynthesis to the ACP (Schemes 1.2 and 1.**3**).[38,39] The PT domain was shown to control the regioselectivity of the poly-β-keto chain cyclisations (Scheme 1.2A: C-2/7, 1.2 **B**: C-4/9 **27**, 1.2 **C**: C-6/11) towards mono/mulitcyclic aromatic compounds such as intermediates **27**, **28** and **29**.[40,41]

The offloading of polyketides from PKS can occur very similar as in FAS by hydrolysis through a TE domain. However, in NR-PKS a variety of different reactions are reported. Often a domain showing similarities to known TE domains but catalysing an intramolecular Claisen cyclisation (TE/CLC domain) is involved. This usually results in phenolic compounds such as the observed intermediate **32** in aflatoxin biosynthesis.[42] Reductive release domains (R), are also known. R-domains produce

aldehydes such as the intermediate **31** on the biosynthesis towards asperfuranone **22**.[39] In emodin **30** biosynthesis the NR-PKS lacks an off-loading domain. Instead a *trans* acting hydrolase, belonging to the β-lactamase family, releases the polyketide upon formation of the cyclohexone intermediate **33**.[43]

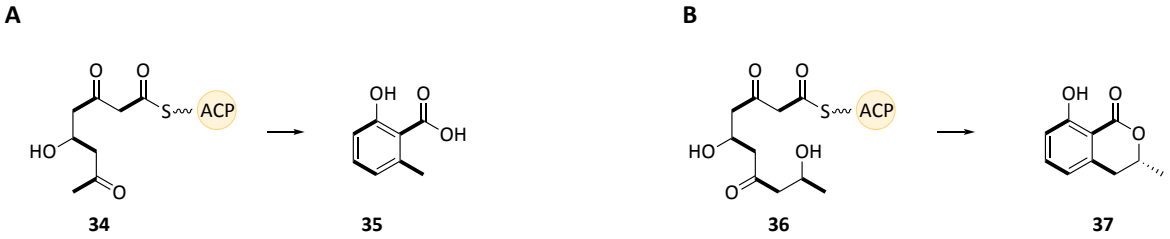


Scheme 1.2 Regioselectivity of product template domains. **A**, In asperfuranone **22**, **B**, aflatoxin **24** and **C**, emodin **30** biosynthesis.



Scheme 1.3 NR-PKS product release. **A**, Reductive, **B**, Claisen-like cyclisation and **C**, by a *trans* hydrolase.

Only a few PKS which produce partially reduced polyketides are reported (Scheme 1.4, 3-methylsalicylic acid **35**, mellein **37**).[44,45] Partially reducing PKS (PR-PKS) possess an incomplete set of modifying domains and often no known off-loading domain.[12] Phylogenetic analysis linked fungal PR-PKS more closely to bacterial modular type I PKS rather than fungal iterative PKS.[46,47]



Scheme 1.4 Partially reduced polyketides. **A**, 3-Methylsalicylic acid **35**. **B**, Mellein **37**.

1.1.2 Terpene Biosynthesis and Sesquiterpenes

A variety of terpene scaffolds of different size and structural complexity such as menthol[8] **4**, aristolochene[48] **38** and *ent*-kaurene[49] **39** are known.[13,14] Although very diverse in structure, all compounds are built up from C_5 isoprene **40** building blocks (Figure 1.8). According to the number of isoprene units incorporated in the terpene backbone these are divided into mono- (C_{10}), sesqui- (C_{15}), di- (C_{20}) and sesterterpenes (C_{25}).[13,14]

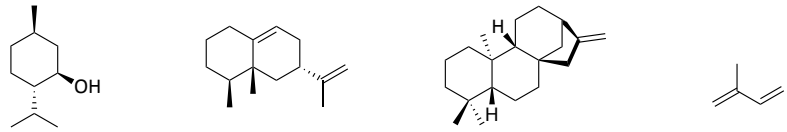
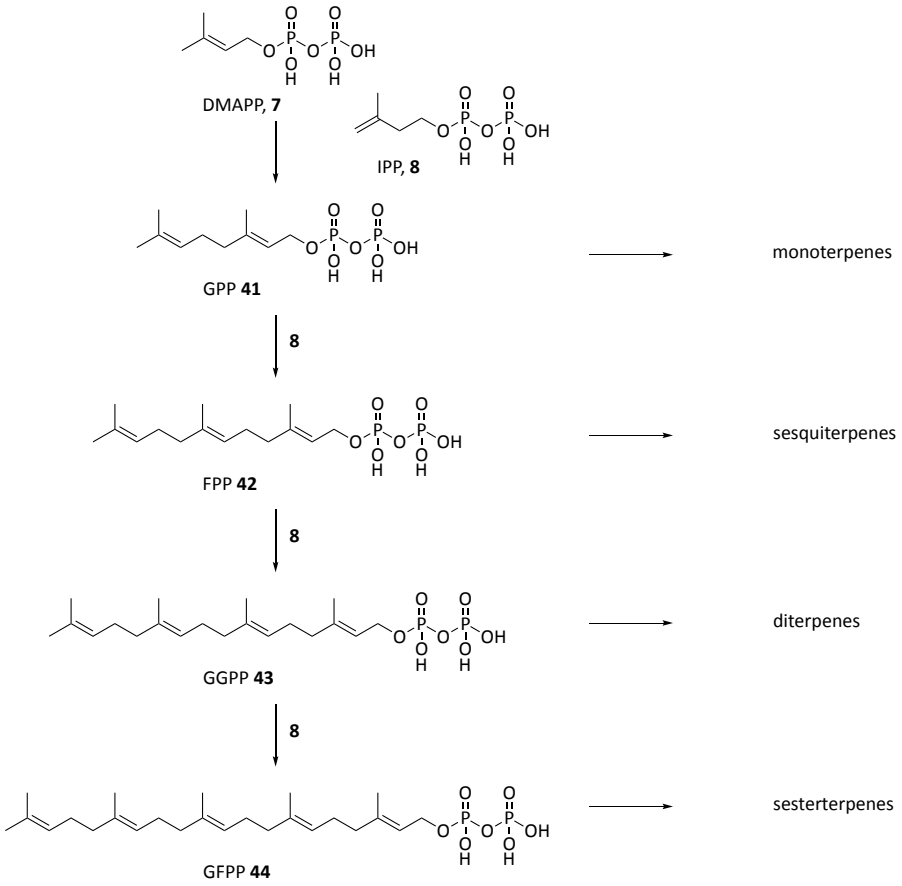


Figure 1.8 Monoterpene menthol **4**, sesquiterpene aristolochene **38**, diterpene *ent*-kaurene **39** and isoprene **40**.

The formal C_5 units are built up by the two precursors DMAPP **7** and IPP **8**, which are produced by the mevalonate pathway in fungi.[50] Isoprenyl pyrophosphate synthases catalyse multiple head-to-tail condensations of IPP **7** and DMAPP **8** to yield the linear precursor molecules: geranyl pyrophosphate **41** (GPP, C_{10}), farnesyl pyrophosphate **42** (FPP, C_{15}), geranylgeranyl pyrophosphate **43** (GGPP, C_{20}) and geranylfarnesyl pyrophosphate **44** (GFPP, C_{25}) (Scheme 1.5).[13,14]

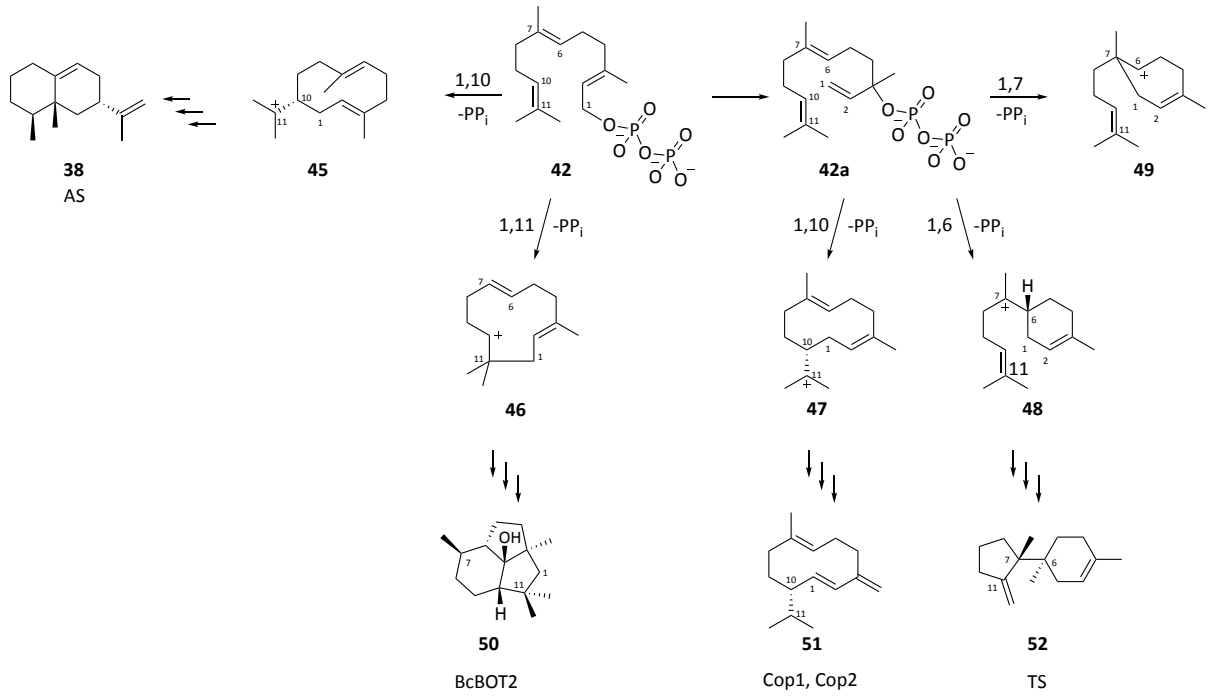
Enzymes known as *terpene cyclases* initiate dephosphorylation of the linear precursors **41**-**44** and induce a cyclisation cascade yielding very structurally diverse terpene scaffolds. Depending on the employed cyclisation mechanism, terpene cyclases are divided into two types. Class I terpene cyclases use an ionisation dependent mechanism and are reported to cyclise GPP **41**, FPP **42** as well as GGPP **43**. Class II terpene cyclases, in contrast, employ a deprotonation dependent mechanisms and use GGPP **43** or GFPP **44** as substrate.[13,14]



Scheme 1.5 Biosynthesis of different length linear terpene precursors.

C_{15} sesquiterpenes are usually formed by class I terpene cyclases. Four initial cyclisations are possible (Scheme 1.6). Direct formation of the carbocation from FPP allows 1,10 or 1,11 ring formation to **45** and **46,** respectively (Scheme 1.6). Some enzymes, however, catalyse a *trans-cis* isomerisation of the 2,3-alkene first, to give the FPP isomer **42a** and enable an initial 1,6 or 1,10 cyclisation to carbocations **47** and **48** (Scheme 1.6). Although 1,7 cyclisations are possible (**49**), no natural product arising from it has been reported.[14] Formation of diversified structures such as aristolochene **38**,[51,52] presilphiperfolan-8β-ol **50**,[53,54] (-)-germacrene D **51**[55,56] and trichodiene **52**[57] is achieved through hydride and methyl shifts as well as ring rearrangements.

Conserved active site motifs of class I terpene cyclases are DDxx(D,E) and NSE/DTE which are essential for the coordination of Mg^{2+}.[58–60] The divalent ions stabilise the pyrophosphate leaving group within the active site cavity and induce the ionisation of the substrate. Aromatic residues neighbouring the active site cavity of all sesquiterpene synthases stabilise reaction intermediates before the reaction sequence is completed either by attack of water or deprotonation.[14]

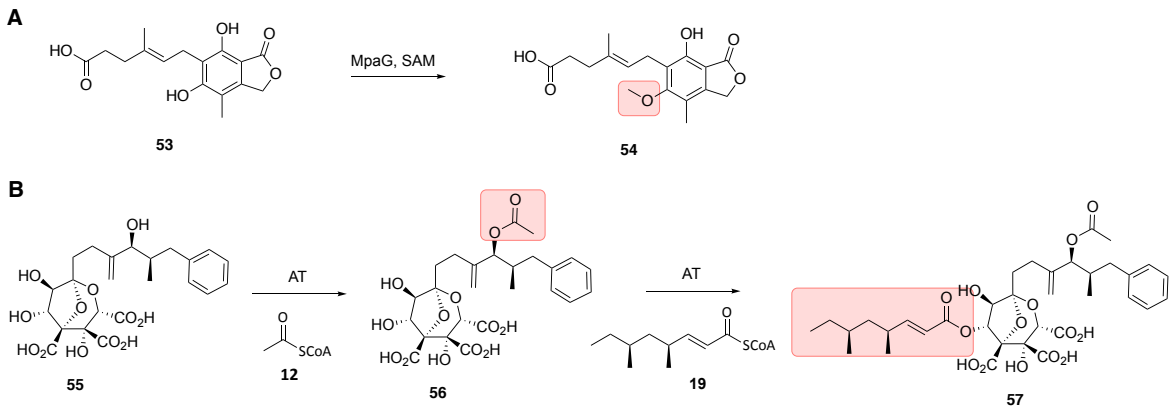


Scheme 1.6 Possible initial FPP **42** cyclisations and examples for products **38**, **50**, **51** and **52**. AS = aristolochene synthase, BcBOT2 = *Botrytis cinera* botrydial synthase, Cop1/4 = *Coprinus cinerus* terpene cyclases 14/4, TS = trichodiene synthase.

1.1.3 Tailoring of Secondary Metabolites

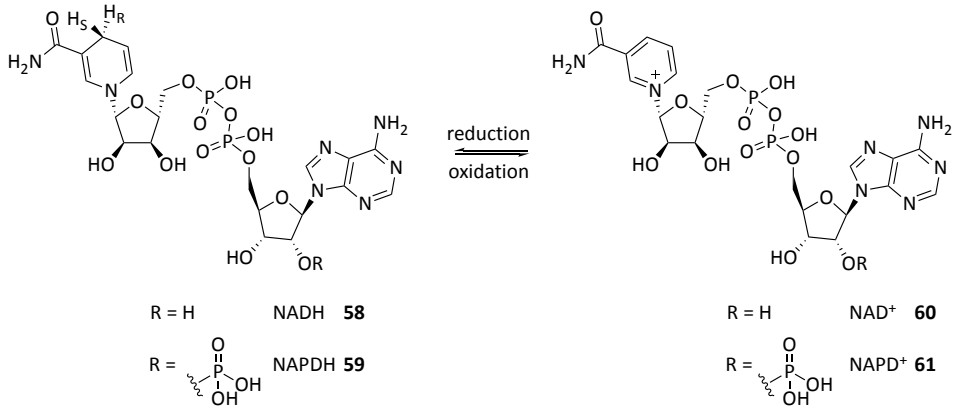
The basic natural product scaffolds produced by core secondary metabolite synthases are almost always substrates of further enzymatic reactions. These transformations are referred to as *tailoring* and involve a wide scope of reactions.

Often *O*-methylation by SAM dependent methyltransferases, as in the biosynthesis of mycophenolic acid **54** (Scheme 1.7A),[61] or *O*-acetylation by acyltransferases, as in squalestatin S1 **57** (Scheme 1.7B) biosynthesis, are observed. More complex compounds such as the squalestatin tetraketide **19** can also be transferred by acyltransferases.[62]



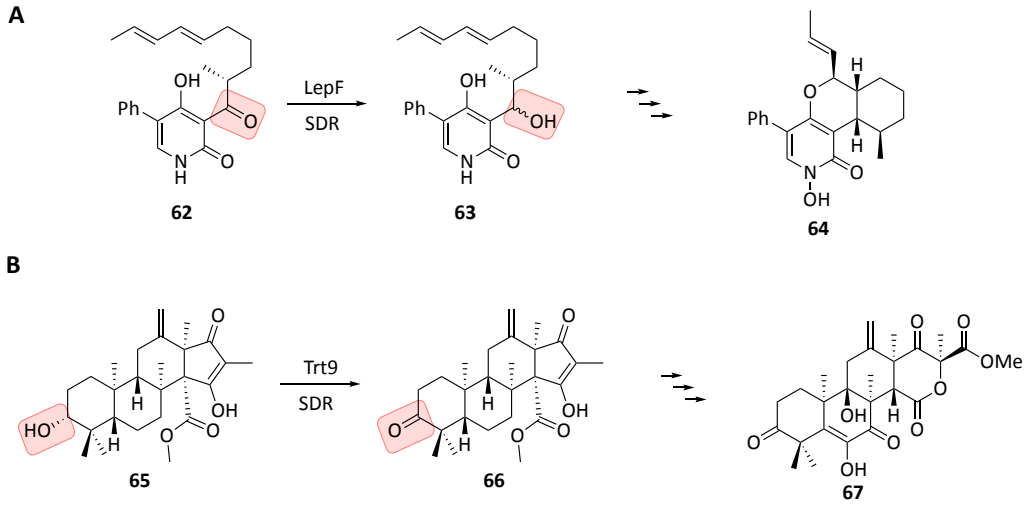
Scheme 1.7 Methylation and acetylation. **A**, In mycophenolic acid **54**, and **B**, squalestatin S1 **57** biosynthesis.

Another frequently seen class of enzymes are nicotinamide adenine dinucleotide (phosphate) (NAD(P)H **58**, **59**/NAD(P)$^+$ **60**, **61**) dependent proteins. These possess a characteristic Rossman fold to bind the co-factor which serves as a hydride donor/acceptor (Scheme 1.8).[63] Depending on the substrates and conditions, these enzymes can catalyse reductions or oxidations of a variety of substrates.



Scheme 1.8 Reduction and oxidation of NADPH/NADH **59**/**58.**

Often carbonyl to alcohol reductions and oxidations are catalysed. In the biosynthesis of leporin A **64** the short chain dehydrogenase/reductase (SDR) LepF reduces **62** to **63** (Scheme 1.9A).[64] In the biosynthesis of terretonin **67**, in contrast, a SDR (Trt9) is reported to catalyse the oxidative reaction of **65** to **66** (Scheme 1.9A).[65]



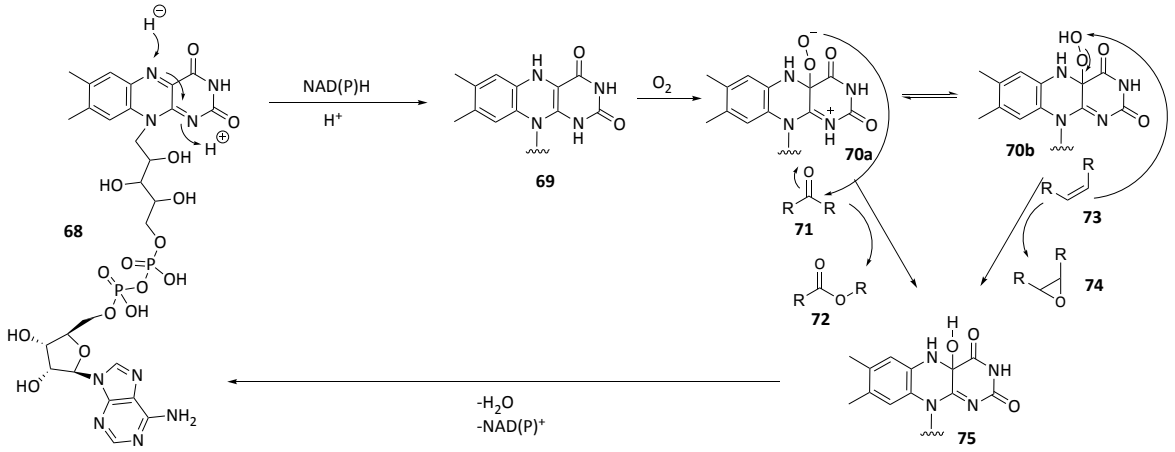
Scheme 1.9 Two examples of NAD(P)H dependent SDR catalysing the respective **A**, Reduction of **62** to **63** and **B**, oxidation of **65** to **66.**

Activation of Atmospheric Oxygen

Oxidations are an important tool of fungal tailoring reactions. Enzymes that typically catalyse these types of reactions are mono- or dioxygenases, classified according to the number of oxygen atoms which are incorporated in the substrate. As atmospheric O_2 is a diradical a direct reaction with paired electron systems is spin-forbidden. Thus, oxygenases usually deploy a co-factor to activate the oxygen for further reactions. Often metal ions such as iron or copper, as well as flavin adenine diphosphate (FAD) **68** are used for this purpose.[66]

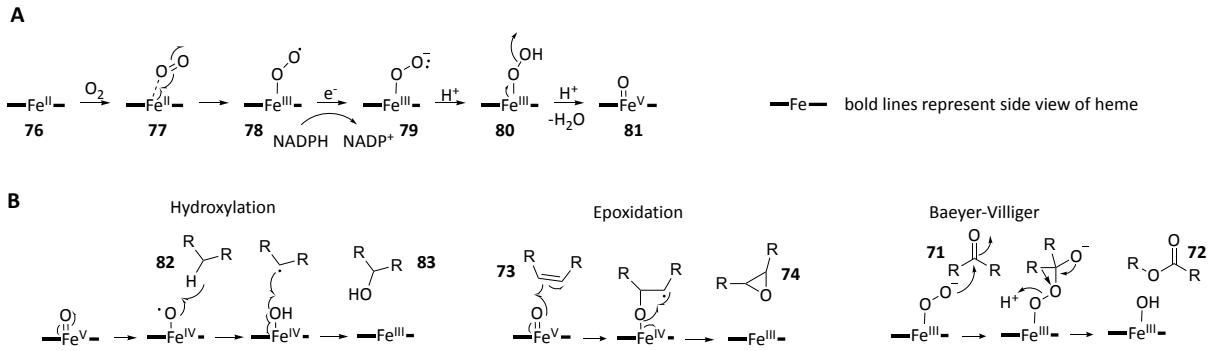
In FAD dependent monooxygenases $FADH_2$ **69** binds O_2 and forms the reactive (hydro)peroxyflavin **70a(b)** species (Scheme 1.10). **70a** and **70b** can transfer oxygen either by a nuclephilic attack (**71** to **72**) to the substrate or by a nucleophilic attack (**73** to **74**) of the substrate. After elimination of water from **75**, FAD **68** is reduced to $FADH_2$ **69** by consuming reduced NAD(P)H.[67]

Heme dependent monooxygenases, known as cytochrome P450s, use the metal Fe^{II} to activate oxygen (Scheme 1.11A).[68] Upon addition of O_2 to Fe^{II} (**76**) it is oxidised to Fe^{III} and the radical intermediate **78** is formed. Reduction by an electron (**79**) and addition of two protons, induce cleavage of the oxygen bond upon release of water. This forms the reactive oxo-Fe^{V} **81** species, which can catalyse hydroxylations (**82** to **83**) and epoxidations (**73** to **74**) of alkanes and alkenes (Scheme 1.11B).[69] The peroxo-Fe^{III} species can also react in a Baeyer-Villiger mechanisms (**71** to **72**) (Scheme 1.11B).[68]



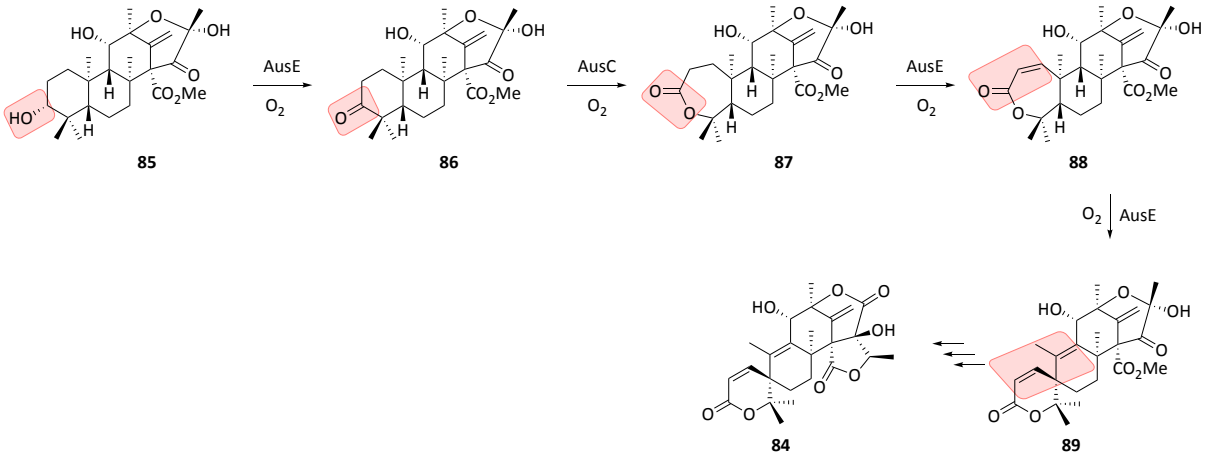
Scheme 1.10 Oxygen activation by $FADH_2$ **69** and regeneration.

Heme provides an electron transfer system for cytochrome P450 enzymes, however also a multitude of non-heme oxygenases are known. These use different co-factors such as Rieske-type ferredoxin or α-ketoglutarate to form the active iron species.[68]



Scheme 1.11 Catalytic activity of cytochrome P450s. **A**, Formation of the active oxo-Fe^{V} species. **B**, Typical reaction catalysed by cytochrome P450s.

The scope of oxygenase reactions includes simple C-H hydroxylations or alkene epoxidations (Scheme 1.11B). However, also more complex desaturations and skeletal rearrangements are often observed.[70] The biosynthesis of austinol **84**, for example, involves two oxygenases: a Bayer-Villiger monooxygenase (AusC) and a non-heme iron dependent dioxygenase (NHI, AusE). These two proteins catalyse four key reactions towards **84** (Scheme 1.12).[65] After AusE oxidises **85** to **86**, AusC induces the formation of the cyclic lactone **87**. This is followed by two additional oxidations by AusE, including the rearrangement of the cyclic lactone **88** to the final spirolactone **89** (Scheme 1.12).



Scheme 1.12 Selected steps of austinol **84** biosynthesis. AusC, Baeyer-Villiger, monooxygenase; AusE, Non-heme Fe^{II} dependent dioxygenase.

1.2 Fungal Meroterpenoids

The Greek prefix *mero* means “part, partial”. Thus meroterpenoids are natural products of combined biosynthetic origin but partially derived from terpenoids. Approximately 81% of 2009 reviewed fungal meroterpenoids possessed a polyketide part, 11% a shikimate moiety and the origin of 8% remained miscellaneous.[71]

Most polyketide terpenoids (60%) are derived (or proposed to derive) from an aromatic tetraketide precursor such as orsellinic acid **90**, 3-methylorsellinic acid **9**, 5-methylorsellinic acid **91** or 3,5-dimethylorsellinic acid **92** (DMOA) (Figure 1.9A).[71] Often oxidations and complex rearrangements occur *en-route* to the final structures, such as the rearrangements observed during biosynthesis of andrastin **94** or terretonin **95**.[72–75] In these examples the original polyketide structure cannot be deduced easily (Figure 1.9B).

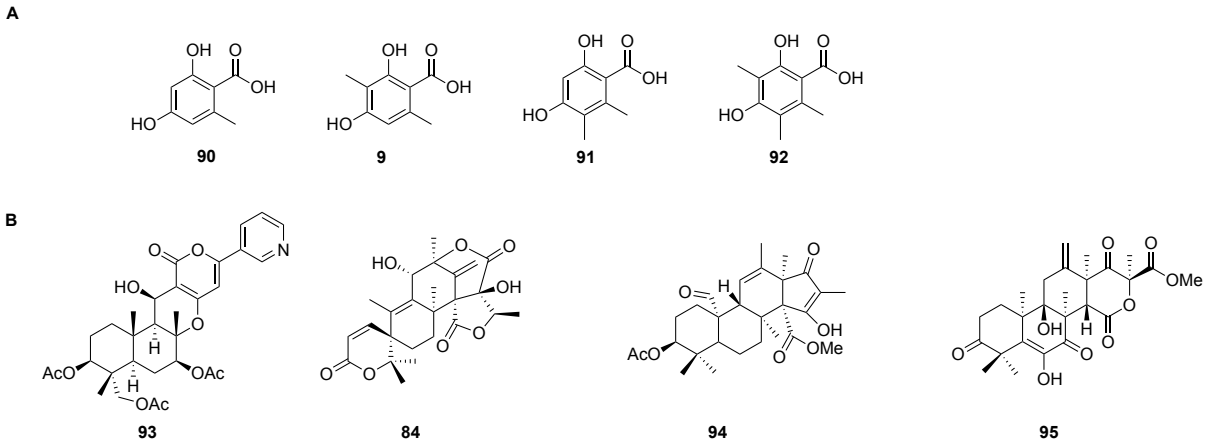
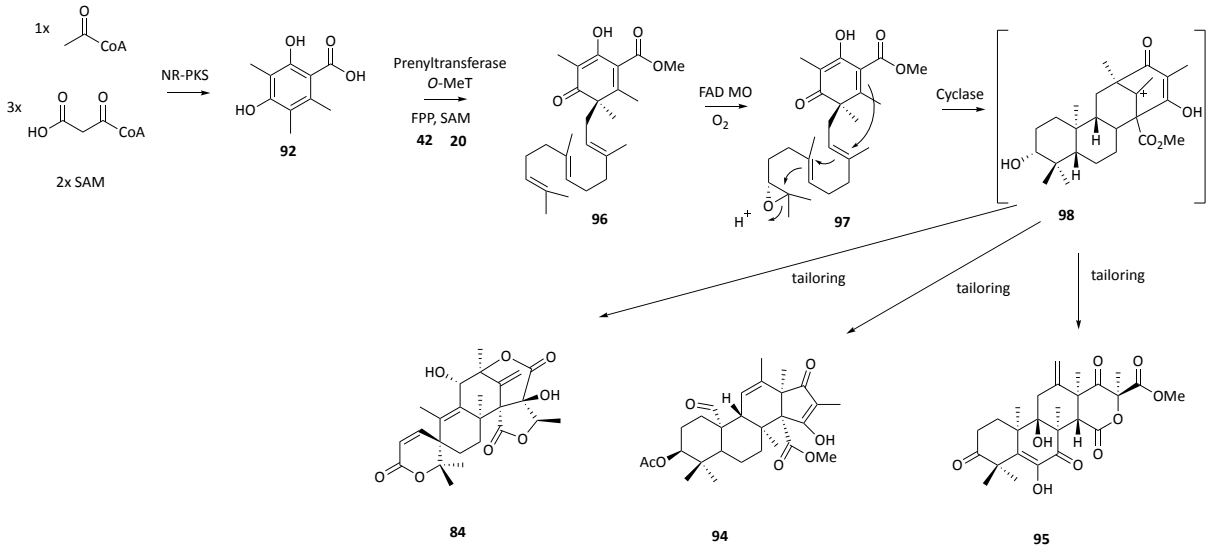


Figure 1.9 Polyketide meroterpenoids. **A**, Tetraketide precursors for meroterpenoids; **B**, Pyripyropene **93**, austinol **84**, andrastin **94** and terretonin **95**.

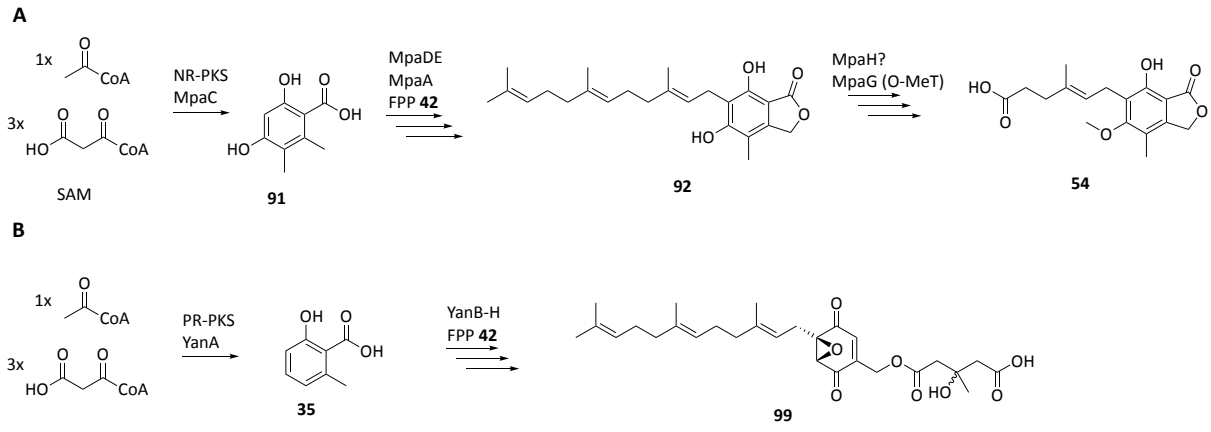
The early-stage biosynthetic steps of DMOA **92** class meroterpenoids (*e.g.* austinol **84**,[76] andrastin[72] **84** and terretonin[73–75] **95**) are very similar and BGC have been reported for a number of these compounds (Figure 1.9). A typical meroterpenoid BGC encodes a NR-PKS that produces the aromatic polyketide (**92**) and an aromatic prenyltransferase which attaches FPP **42** onto the polyketide moiety (**92** to **96**). In a subsequent reaction the terminal alkene of the sesquiterpene moiety is epoxidised (**96** to **97**) by an FAD dependent monooxygenase. Cylisation is then catalysed by an integral membrane protein. This class of terpene cyclase was first reported in the biosynthesis of pyripyropene A **93** (Scheme 1.13).[77] Further diversification is catalysed by the set of clustered genes encoding proteins for tailoring reactions.

A NR-PKS/PR-PKS and an aromatic prenyltransferase are also encoded in the BGC for the biosynthesis of compounds with non-cyclic terpene moieties such as mycophenolic acid[61,78] **54** and yanuthone D **99** (Scheme 1.14).[79] As no cyclisation of the sesquiterpene scaffold is catalysed in these cases, the respective cyclase genes are lacking in the BGC. For the biosynthesis of **54,** however, further degradation of the C-15 to C-7 terpene chain is thought to be catalysed by MpaH.[61]



Scheme 1.13 Early steps of DMOA **92** derived meroterpenoids austinol **84**, andrastin **94** and terretonin **95**. FAD dependent monooxygenase (FAD MO).

The majority of fungal meroterpenoids (reviewed in 2009 and 2014)[71,80] are formed through the attachment of a linear terpene precursor to the polyketide moiety. Few examples (fumagillin **100**,[81] xenovulene A **101**[82] and related structures, Figures **1.10** and **1.11**) are known where the terpene scaffold is proposed to cyclise first. Fumagillin **100** has been linked to a BGC, but as its terpene moiety is attached via an ester linkage rather than a C-C bond its biosynthesis is likely to be different to xenovulene A **101**.[81]



Scheme 1.14 Biosynthesis of meroterpenoids with non-cyclic terpene moiety. **A**, Biosynthesis of mycophenolic acid **54**. MpaDE: chimeric P450 and hydrolase, MpaA: prenyltransferase. **B**, Biosynthesis of yanuthone D **99**. YanB: decarboxylase, YanC: P450, YanD: putative dehydrogenase, YanE: unknown, YanF: dehydrogenase, YanG: prenyltransferase, YanH: P450.

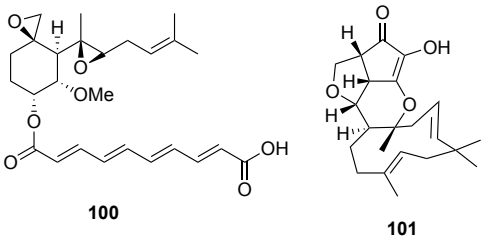
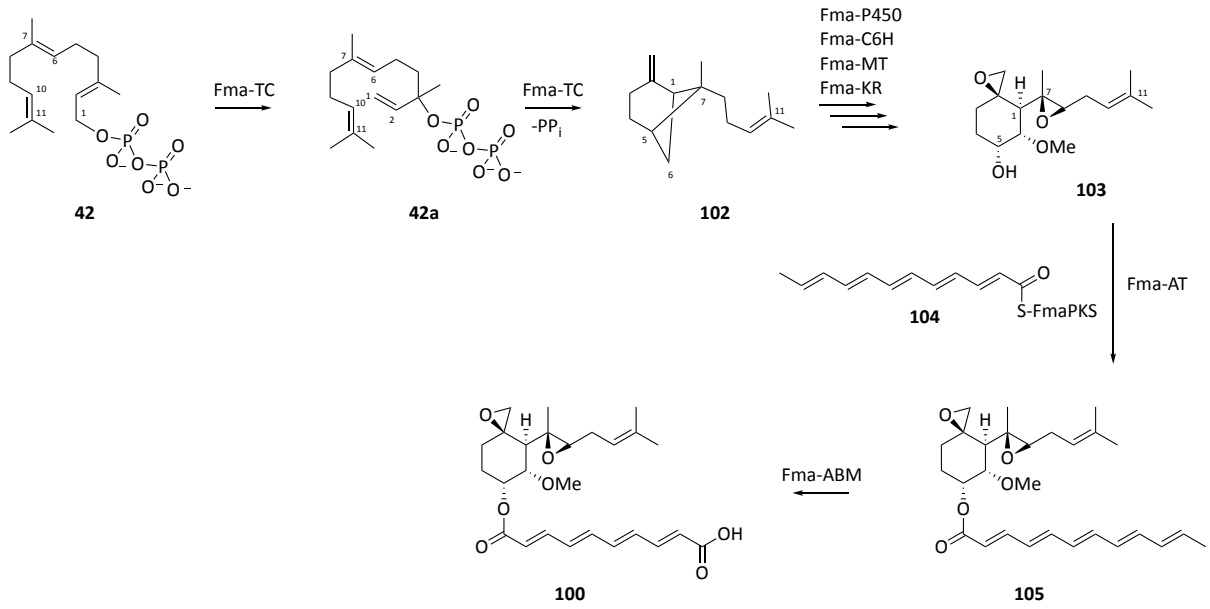


Figure 1.10 Fumagillin **100** and Xenovulene A **101**.

During the biosynthesis of fumagillin **100**, an unusual membrane bound FPP cyclase forms β-*trans*-bergamotene **102** from FPP **42** (Scheme 1.15). This protein shows homologies to the prenyltransferase UbiA,[83] but not to class I terpene cyclase or to the membrane bound terpene cyclase reported in pyripyropene A **93** biosynthesis.[77] Further oxidation and rearrangements yield the fumagillol **103** scaffold which is attached to the HR-PKS derived dodecapentenoyl skeleton (**104**) by *trans*-esterification yielding **105**. Further oxidation gives the final product fumagillin **100**.[81]



Scheme 1.15 Fumagillin **100** biosynthesis. Fma-TC, terpene cyclase; Fma-C6H, dioxygenase; Fma-MT, methyltransferase; Fma-KR, ketoreductase; Fma-AT, acyltransferase; Fma-ABM, monooxygenase.

1.3 Xenovulene A and Related Meroterpenoids

Xenovulene A **101** is a polyketide meroterpenoid that was isolated from *Acremonium strictum* in 1995. *A. strictum* is still the only known producer of this compound (Figure 1.11).[82] It contains an unusual polyketide-derived *tetrahydro-furocyclopentenone* moiety (red) that is fused via a *tetrahydropyran* (blue) to an 11-membered cyclic sesquiterpene (green) which corresponds to

humulene **106**. Related tropolone and phenolic meroterpenoids **107**, **108** and **109** have been reported to be produced alongside **101**.[84]

Similar meroterpenoids have also been reported from a series of different fungi (Figure 1.12). These include pycnidione **110** (*Phoma sp.*),[85] eupenifeldin **111** (*Eupenicillium brefeldianum*)[86] and noreupenifeldin **112** (unidentified fungus),[87] epolones A **113** and B **114** (unidentified fungus),[88] ramiferin **115** (*Kionochaeta ramifera*),[89] pughiinin A **116** (*Kionochaeta pughii*)[90] and phomanolides A **117** and B **118** (*Phoma sp.*).[91] These compounds all share the structural feature of the tetra/dihydropyran that bridges the terpene and polyketide derived moieties. Disubstitution and hydroxylation of the humulene ring appear to be common modifications of the system but have not been observed with xenovulene A **101** and its congeners **107**-**109**.

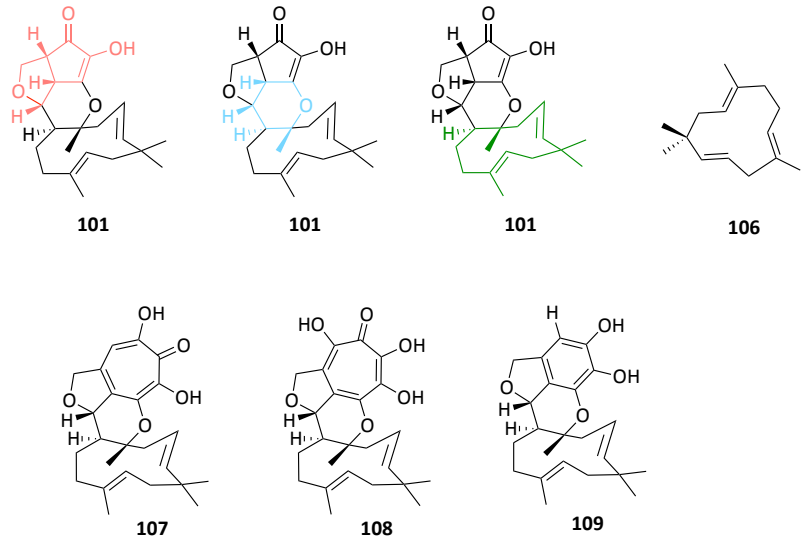


Figure 1.11 Xenovulene A **101** and co-metabolites **107**, **108** and **109**.

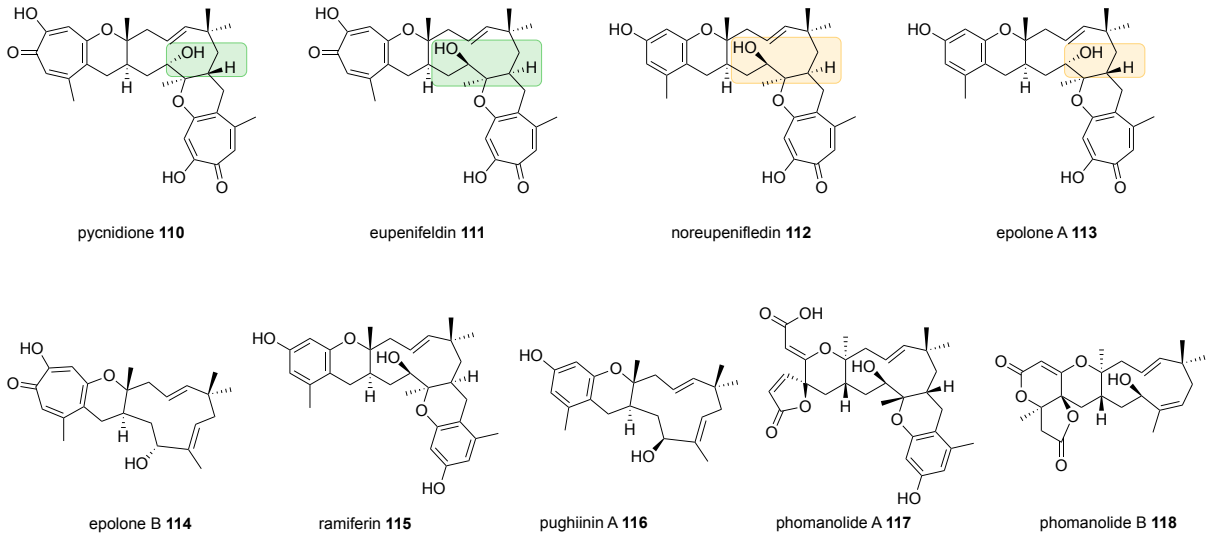


Figure 1.12 Meroterpenoids related to xenovulene A **101.**

All of these compounds show interesting biological activites. Epolones (**113**, **114**) for example induce erythropoetin expression,[88] eupenifeldin (**111**) showed antiproliferative activity against human glioma cell lines[91] and xenovulene A **101** was found to bind the gamma aminobutyric acid (GABA) receptor.[82,92]

1.3.1 Fungal Tropolone Biosynthesis in *Talaromyces stipitatus*

A shared structural feature of xenovulene type meroterpenoids (**110**-**114**) is the often observed tropolone core. Tropolones like stipitatic acid **119** are known fungal polyketide metabolites.[93]

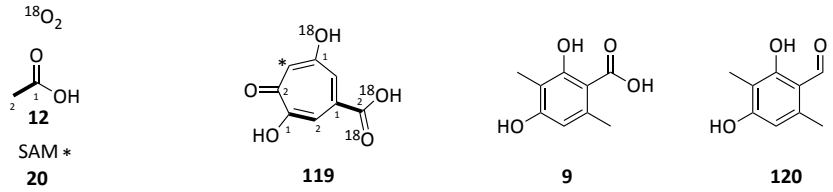
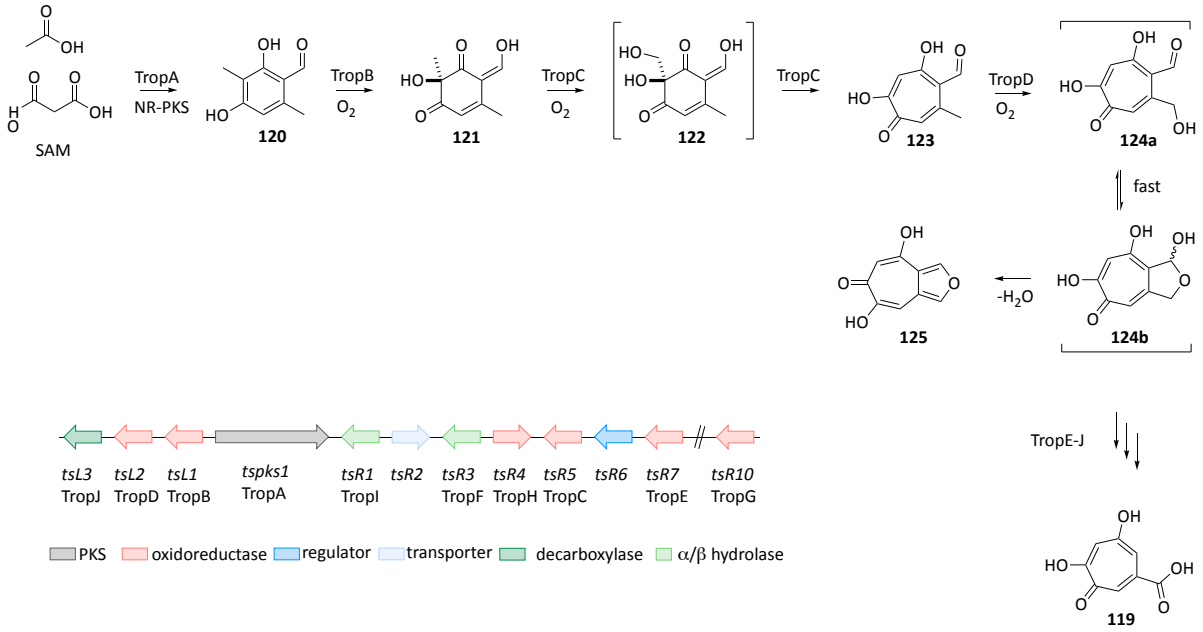


Figure 1.13 Feeding studies and deduced labelling patter of stiptitatic acid **119.**

The biosynthesis of stipitatic acid **119** has been investigated since the 1950s. Feeding studies using ^{14}C labelled precursors showed that **119** consisted of two intact C_2 units and one C_1 unit intercepting a third C_2 unit (Figure 1.13).[94] Feeding with $^{18}O_2$ showed the incorporation of one atmospheric oxygen into the tropolone scaffold.[95] Additional feeding of the advanced precursors 3-methylorsellinic acid **9** and 3-methylorcinaldehyde **120**, showed that the aldehyde **120** is converted into tropolones more effectively than the acid **9**.[96] The molecular basis for **119** biosynthesis was dissected in *Talaromyces stipitatus* where gene knockout and heterologous expression experiments identified the BGC and elucidated the pathway of **119** biosynthesis (Scheme 1.16).[97,98]

The PKS bound tetraketide precursor is reductively released by the NR-PKS TropA to give 3-methylorcinaldehyde **120**. The FAD dependent salicylate monooxygenase TropB then dearomatises the aldehyde **120** to give enone **121** by oxidation at C-3. **121** is the substrate of the non-heme Fe^{II} dependent dioxygenase TropC, which induces a ring expansion towards tropolones by hydroxylation of the methyl group at C-3 (**122**). Rearrangement leads to the first tropolone stipitaldehyde **123**. Oxidation by the cytochrome P450 TropD most likely produces **124a** in equilibrium with its hemiacetal **124b**, but this intermediate has not been directly observed. Instead, stipitafuran **125** was accumulated upon knockout of the dehydrogenase TropE. Further transformations by TropE-J convert the likely hemiacetal intermediate **124b** into stipitatic acid **119**.[97,98] Discovery of these genes has allowed the identification of parallel tropolone pathways in different fungi.[97]



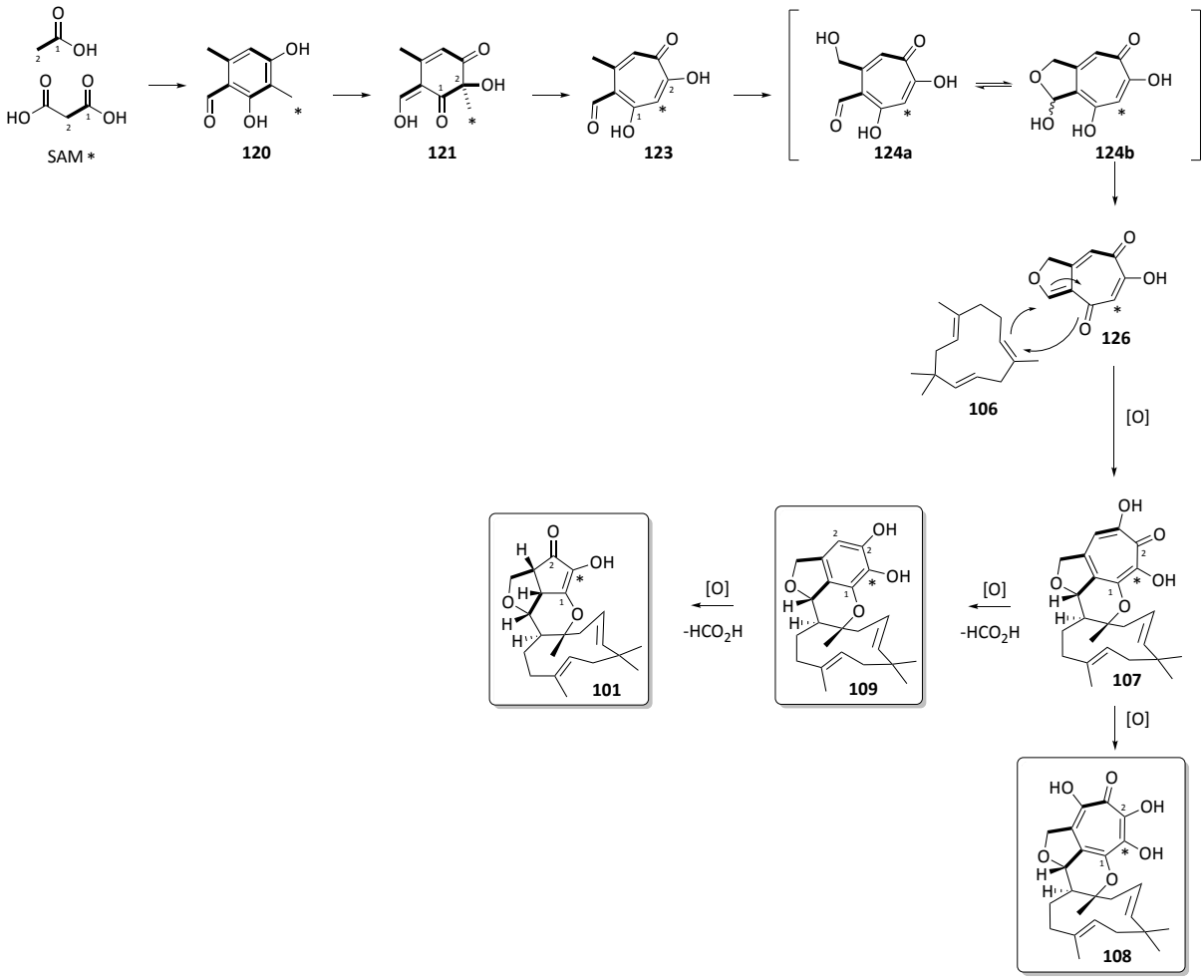
Scheme 1.16 Stipitatic acid **119** biosynthesis in *T. stipitatus* and BGC.

1.3.2 Proposed Xenovulene A Biosynthesis and Partial *aspks1* BGC

For the elucidation of xenovulene A **101** biosynthesis feeding experiments with ^{13}C labelled precursors were analysed. ^{13}C-NMR of enriched **101** showed two intact acetate units in the polyketide moiety and one acetate unit disrupted by a methionine-derived methyl group, (Scheme 1.17). The labelling pattern of the humulene ring was consistent with a terpene origin.[84]

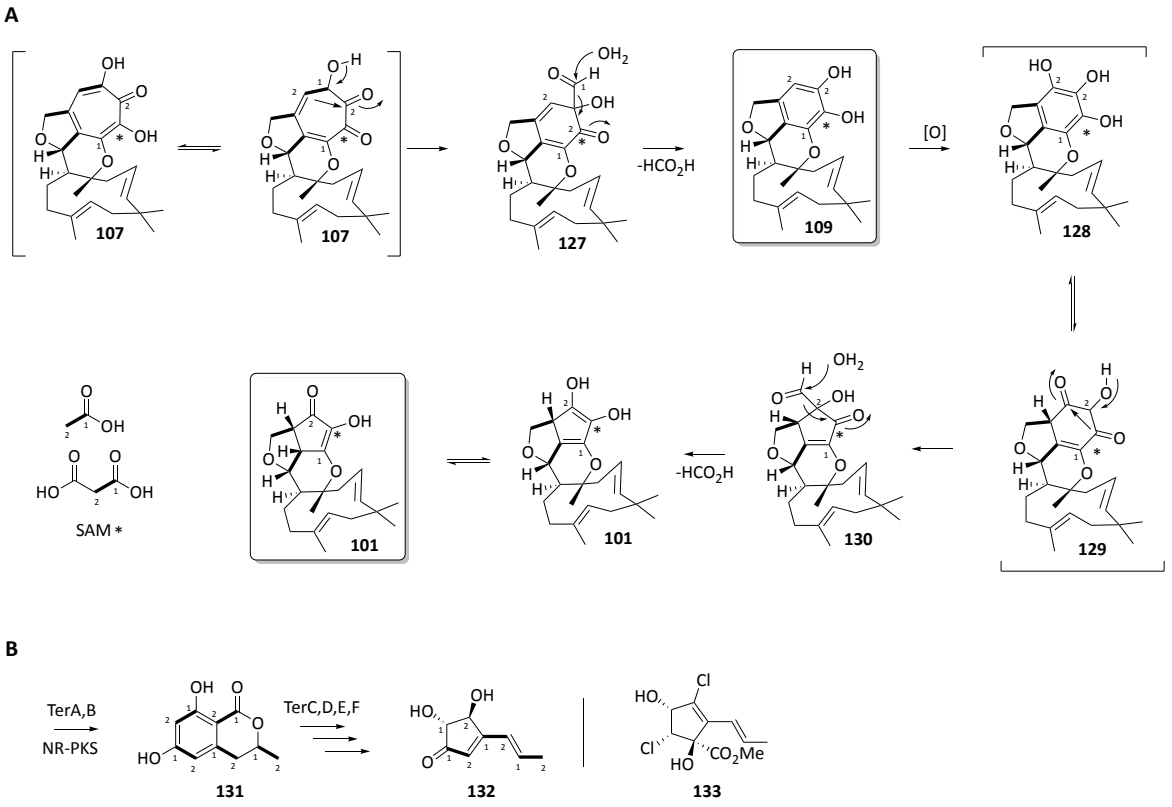
Although nowadays tropolone biosynthesis is understood on a genetic level, at the time of these feeding studies no such knowledge was available. However, from previous feeding experiments in *Penicillium stipitatum* (a tropolone producer) with ^{14}C labelled precursors it was known that the methionine derived methyl group is incorporated into the seven membered tropolone core upon disruption of an acetate unit (Chapter 1.3.1).[94] Based on the observed labelling pattern in xenovulene A **101,** and knowledge about methyl incorporation in tropolones, a ring expansion-ring contraction mechanism was proposed to generate the cyclopentenone in **101** (Scheme 1.17).[84]

For the two ring contractions from **107** to **101** an oxidation, rearrangement and deformylation cascade was proposed (Scheme 1.18A). Similar ring contractions have been proposed to happen during the biosynthesis of terrein[99] **132** and cryptosporiopsonol[100] **133** *via* a dihydroisocoumarin **131** intermediate (Scheme 1.18B). Although the terrein BGC has been identified no mechanism or proteins for this chemistry have yet been characterised.[101]



Scheme 1.17 Proposed biosynthesis based on feeding studies, heterologous gene expression of *aspks1* (encoding the NR-PKS MOS) and tropolone biosynthesis in *T. stipitatus*.

In a genetic approach, even before the tropolone BGC was discovered, a partial gDNA fragment of the xenovulene A **101** producer *A. strictum* strain was analysed for a *nrpks* gene using degenerate oligonucleotides.[102] Similar to tropolone biosynthesis, where 3-methylorcinaldehyde **120** was shown to be incorporated into stipitatic acid, it was assumed that an aromatic tetraketide produced by a NR-PKS should be a likely precursor in **101** biosynthesis.[96] The *aspks1* gene encoding a NR-PKS was identified and a ~22 kb contig was sequenced from a gDNA library and analysed for further open reading frames (ORF). In total, nine ORF were predicted and the basic local alignment search tool for proteins (BLASTp)[103] revealed the presence of two putative FAD dependent monooxygenases (AsL1, AsL4), one putative non-heme Fe^{II} dependent dioxygenase (AsL3), one short chain dehydrogenase (AsL5), one putative cytochrome P450 (AsR2), one putative transcriptional regulator (AsR3), one putative transporter (AsR1) and one hypothetical protein (AsL2) (Figure 1.14A).[102,104,105]

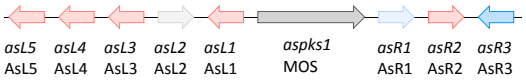


Scheme 1.18 Ring contractions. **A**, Proposed ring contraction in xenovulene A **101** biosynthesis. **B**, Ring contractions and labelling in terrein **132**. Cryptosporiopsonol **133.** TerB, DH-KR multidomain protein; TerC/TerD, FAD dependent monooxygenases; TerE, multicopper oxidase; TerF, protein with kelch motif.

Heterologous expression of *aspks1* in *A. oryzae* led to the production of 3-methylorcinaldehyde **120**, which made the NR-PKS encoded in *aspks1* the first reported PKS with a reductive release mechanism; it was named 3-methylorcinaldehyde synthase (MOS).[102] Later the *A. strictum aspks1* BGC was used to mine the available *T. stipitatus* genome for a homologous BGC that could encode proteins for stipitatic acid **119** biosynthesis. With this approach the now fully-characterised stipitatic acid BGC was identified (Chapter 1.3.1). Pairwise comparison of proteins suggested that with MOS, AsL1, AsL3 and AsR2 homologues to TropA, TropB, TropC and TropD are encoded in the *aspks1* BGC (Figure 1.14B).[106]

The 22 kb *aspks1* BGC as reported in 2007 encodes homologues for tropolone biosynthesis, which suggested it as a promising BGC for **101** biosynthesis as tropolone containing metabolites **107** and **108** were isolated alongside **101**. However, it lacks an encoded terpene cyclase to form humulene **106** from FPP **42**. It is assumed that the previously discovered BGC is incomplete and that genes encoding proteins for terpene biosynthesis are encoded nearby, but have not yet been sequenced.

A



AsL5: NAD(P)(H) dependent short-chain dehydrogenase
AsL4: FAD-dependent 2-polyprenyl 6-methoxyphenol hydroxylase
AsL3: Non-heme dependent Fe^{II} dioxygenase
AsL2: hypothetical protein
AsL1: FAD dependent salicylate monooxygenase
MOS: NR-PKS, 3-methylorcinaldehyde synthase
AsR1: MFS transporter
AsR2: cytochrome P450
AsR3: transcription factor GAL4-like

B

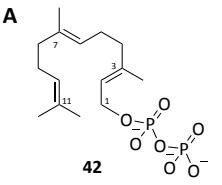
Similarity/ Identity (%aa)	TropA	TropB	TropC	TropD
MOS	51.9/38.7	-	-	-
AsL1	-	59.9/46.4	-	-
AsL3	-	-	56.2/41.6	-
AsR2	-	-	-	50.9/39.3

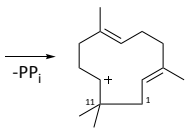
Figure 1.14 *aspks1* BGC. **A**, Partial *aspks1* BGC. (Major facilitator superfamily (MFS); Gal4, galactose induced transcription factor superfamiliy) **B**, Putative homologous proteins in the stipitatic acid **119** BGC.

1.3.3 Humulene Biosynthesis in Fungi

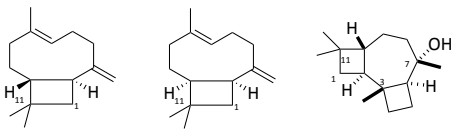
The rather simple metabolite humulene **160** is usually formed by deprotonation of the *trans*-humuly cation **46** after initial 1,11-ring closure of FPP **42** (Scheme 1.19A).[14] Although the metabolite is a common component of plant essential oils only one plant terpene cyclase (*Zingiber zerumbet* Smith) is reported that produces **106** as a major product.[107,108] Production of humulene **106** in fungi has only been reported in *Fusarium fujikorii*,[109] *Colletotrichum acutatum*[110] and *Stereum hirsutum*.[56] The responsible class I terpene cyclases Ffsc4 and CaTPS were characterised by heterologous expression. Both were found to produce **106** as well as β-caryophyllene **134**. Ffsc4 also produces epi-β-caryophyllene **135** and koraiol **136**. CaTPS additionally produces an unidentified terpene (Scheme 1.19B).[109,110] No fungal cyclase is reported to produce exclusively **106**.

A





B

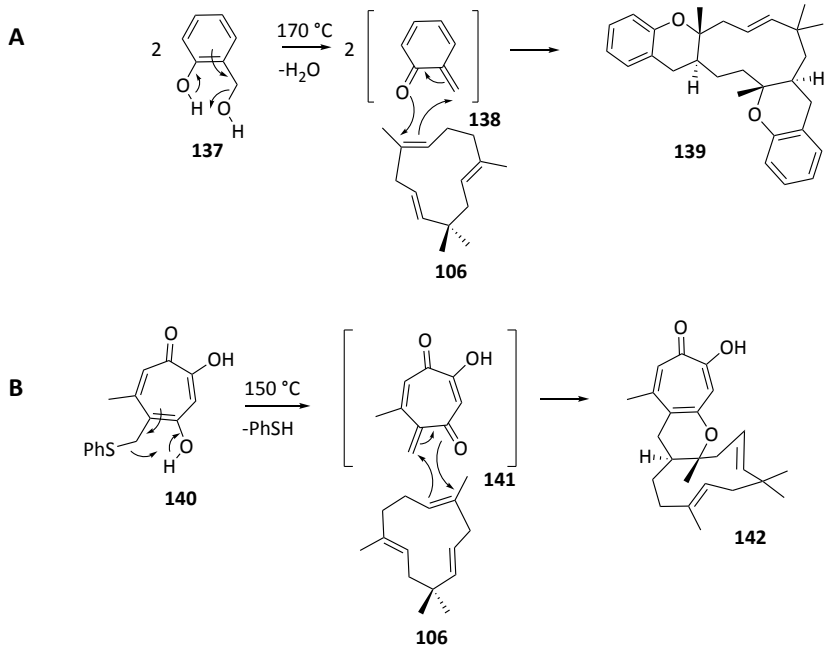


Scheme 1.19 A, Biosynthesis of humulene **106**. **B**, β-caryophyllene **134**, epi-β-caryophyllene **134** and koraiol **136**.

1.3.4 Biomimetic Experiments and Diels Alderases

Another unknown step during xenovulene A **101** biosynthesis is the attachment of the polyketide and terpene moieties. Although only tropolone (**107**, **108**) and post-tropolone phenolic (**109**) meroterpenoids have been reported, it remains elusive which polyketide precursor adds to humulene **106**. A hetero Diels Alder (hDA) cycloaddition mechanism has been proposed (Scheme 1.17), which is consistent with the relative stereochemistry. However, the enantiopurity suggest the involvement of an enzyme.[84]

In a biomimetic approach, tropolone- and benzo-*ortho*-quinone methides (**138**, **141**) were tested for their reactivity towards humulene **106** by Baldwin and co-workers (Scheme 1.20A and B).[111–113] Both quinone-methides (**138**, **141**) reacted with **106** leading to mono- and di-adducts **142** and **139** under harsh reaction conditions (150 – 170 °C, 24 h). This suggested that an inverse electron demand hetero Diels Alder (hDA) reaction is a feasible approach to form xenovulene class meroterpenoids. However, the observed reaction conditions, poor yields and low stereoselectivity suggest that an enzyme catalyses the formation of **101** *in vivo*.

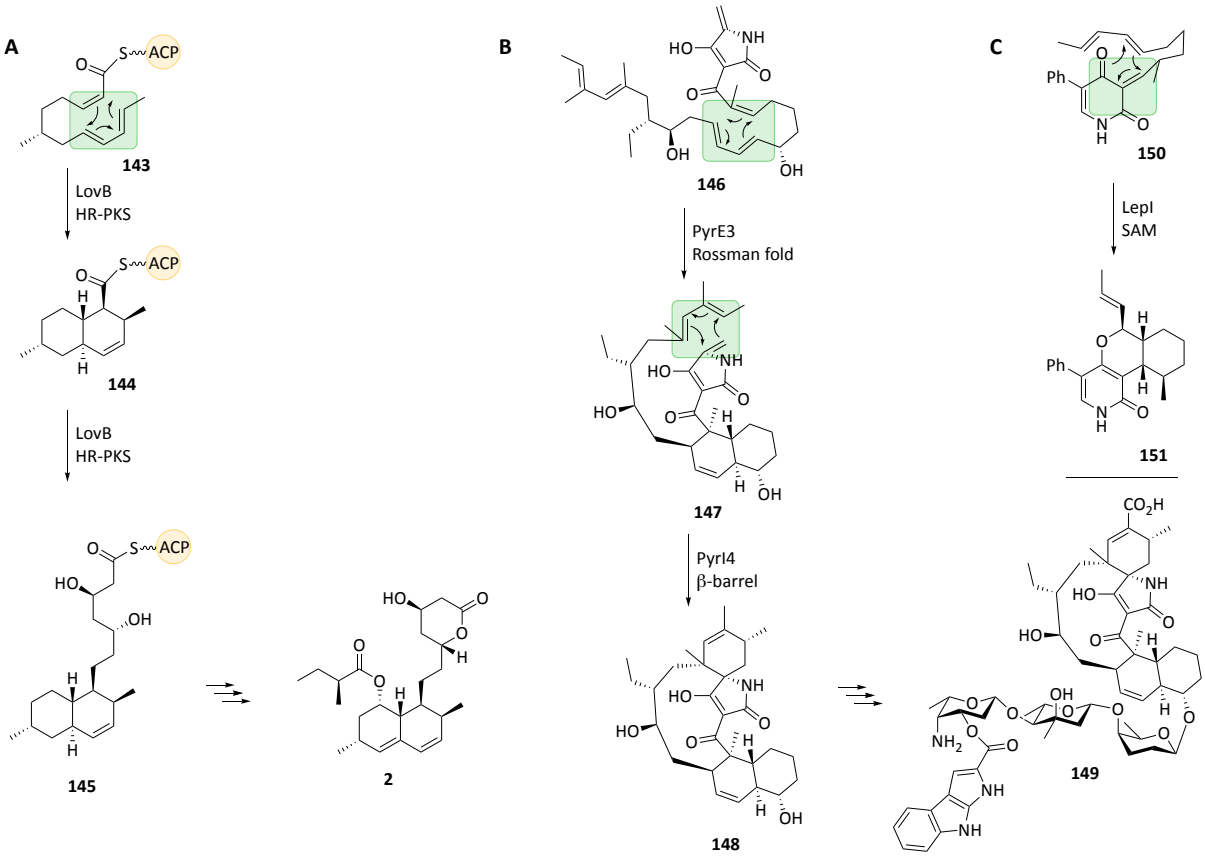


Scheme 1.20 Biomimetic studies showing benzo-*ortho*-quinone methide (**138**, **A**) and tropolone-*ortho*-quinone methides (**141**, **B**) to react with humulene **106** in a hetero Diels Alder reaction.

An example of a multi-function Diels Alderase (DAase) is the HR-PKS lovastatin **2** synthase (LovB). It catalyses carbon chain elongation to build up the polyketide backbone and the subsequent intramolecular decalin **144** formation by an additional DA activity (Scheme 1.21A).[114] Two individual bacterial stand-alone DAase (PyrE3, PyrI4) each catalysing a [4+2] cyclisation in the biosynthesis of pyrroindomycin **149** (Scheme 1.21B) were reported.[115,116] PyrE3 is a Rossman fold protein while the crystal structure of PyrI4 showed it to be a β-barrel.

The first and only stand-alone hetero Diels Alderase (hDAase) reported shows homologies to a SAM dependent methyltransferase. This enzyme (LepI) forms the dihydropyran ring in the fungal metabolites leporin C **151** and B **64** (Scheme 1.21C).[64] A similar dihydropyran ring is also found in the xenovulene class meroterpenoids (Figures 1.11 and 1.12).

The structural divergence of these proteins (LovB, NR-PKS; PyrE3, Rossmann fold; PyrI4, β-barrel; LepI, SAM dependent) indicates that nature uses different tertiary architectures to induce these [4+2] cycloadditions, which results in difficulties predicting DAases or hDAases *in silico*.[116,117]



Scheme 1.21 Examples for catalysed Diels Alder reactions by a **A**, the multifunctional DAase LovB; **B**, Two stand-alone DAase and **C**, one stand-alone hDAase.

1.4 Project Aims

The biosynthesis of xenovulene A **101** in *A. strictum* is the key focus of this project. For this reason, the secondary metabolite profile of *A. strictum* will be analysed in detail. Previously reported **107**, **108** and **109** (Figure 1.11) and putative novel compounds will be identified, purified and their structure elucidated by subsequent full NMR analysis.

Sequencing of the whole *A. strictum* genome will be attempted to enable a full bioinformatic analysis, prediction of all BGC and completion of the partial *aspks1* BGC. Comparison of the *A. strictum* transcriptome under xenovulene A **101** producing and non-producing conditions will help to identify the BGC responsible for **101** biosynthesis, to set neat cluster boundaries and to verify the gene model as well as exon/intron predictions.

The best candidate BGC for **101** biosynthesis will be further analysed through gene knockout, editing and silencing methods in *A. strictum*. Targeted gene inactivation has not previously been achieved in this organism and thus different methods will be tested. In addition, heterologous expression of selected genes of the candidate BGC in *A. oryzae* and/or *E. coli* will be performed to characterise their role in **101** biosynthesis. The identification of a humulene **106** synthase, the proposed hDA catalyst and the enzymes involved in the two ring contractions to cyclopentenone are of special interest.

Early tropolone biosynthesis in *A. strictum* has already been proposed to be similar to stipitatic acid biosynthesis in *T. stipitatus*. It will be of interest whether the entire tropolone pathway is identical in *A. strictum*. Apart from 3-methylorcinaldehyde **120**, no tropolone precursors have previously been identified in *A. strictum*.[104] Therefore, crude extracts will be analysed for the presence of possible tropolone polyketide precursors.

2 Analysis of *A. strictum* Secondary Metabolite Production

The *A. strictum* xenovulene A producing strain was first reported for the production of **101** and co-metabolites **107**-**109** (Figure 2.1).[82,84] Extensive efforts were made to optimise the culture conditions and yield for **101**.[118] Several time course experiments within our group[104,105] and also by Blackburn and co-workers[118] showed initial production of **101** after 1-3 days (d) of liquid culture followed by stagnation after 6-8 d. Similar results were obtained in a Master thesis by Miriam Streeck, but as often observed with fungi, yields and composition of actually produced secondary metabolites were inconsistent.[119]

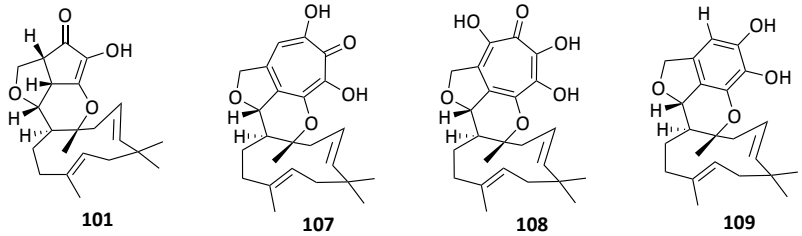


Figure 2.1 Metabolites known to be produced by *A. strictum*.

A. strictum was typically sub-cultured in **101** production media (ASPM) for 6-8 d (25 °C, 200 rpm). Cells were removed by centrifugation and the supernatant was acidified to pH 2 with 2 M HCl prior to extraction with ethyl acetate:hexane (1:1). Solvents were removed *in vacuo* and dried extracts dissolved in methanol or acetonitrile:water (9:1) to a concertation of 10 mg/mL for analytical purposes, and 50 mg/mL for purification by high performance liquid chromatography (HPLC). Compounds purified by Miriam Streeck (**107a** and **107b**, **108**) were obtained from combined extracts of a time course experiment, where extraction was carried out at ten consecutive days.[119]

A typical LCMS chromatogram of an *A. strictum* extract was analysed for the production of xenovulene A **101** and known metabolites **107**, **108** and **109**. Although structures for **107**, **108** and **109** have been previously reported in the literature,[84] no spectroscopic (NMR or UV) characterisation was available.[84] Extracted ion chromatograms using $[M]H^+$ or $[M-H]^-$ corresponding to **101**, **107**, **108** or **109** led to the identification of compounds satisfying the nominal mass of all previously reported compounds (Figure 2.2). Additionally, two compounds satisfying the nominal mass of **107** and **109** with different retention times (**109a** t_R = 9.2 and **109b** 9.4 min, **107a** t_R = 9.9 and **107b** 11.6 min) were observed for the first time. Another compound (**154**) eluting at t_R = 11.3 with a nominal mass of 382 and very similar UV absorption compared to the other tropolone compounds (**107** and **108**) was identified as a probably related meroterpenoid (Figure 2.2). All compounds were purified by HPLC and analysed by NMR spectroscopy (Chapters 2.1-2.7).

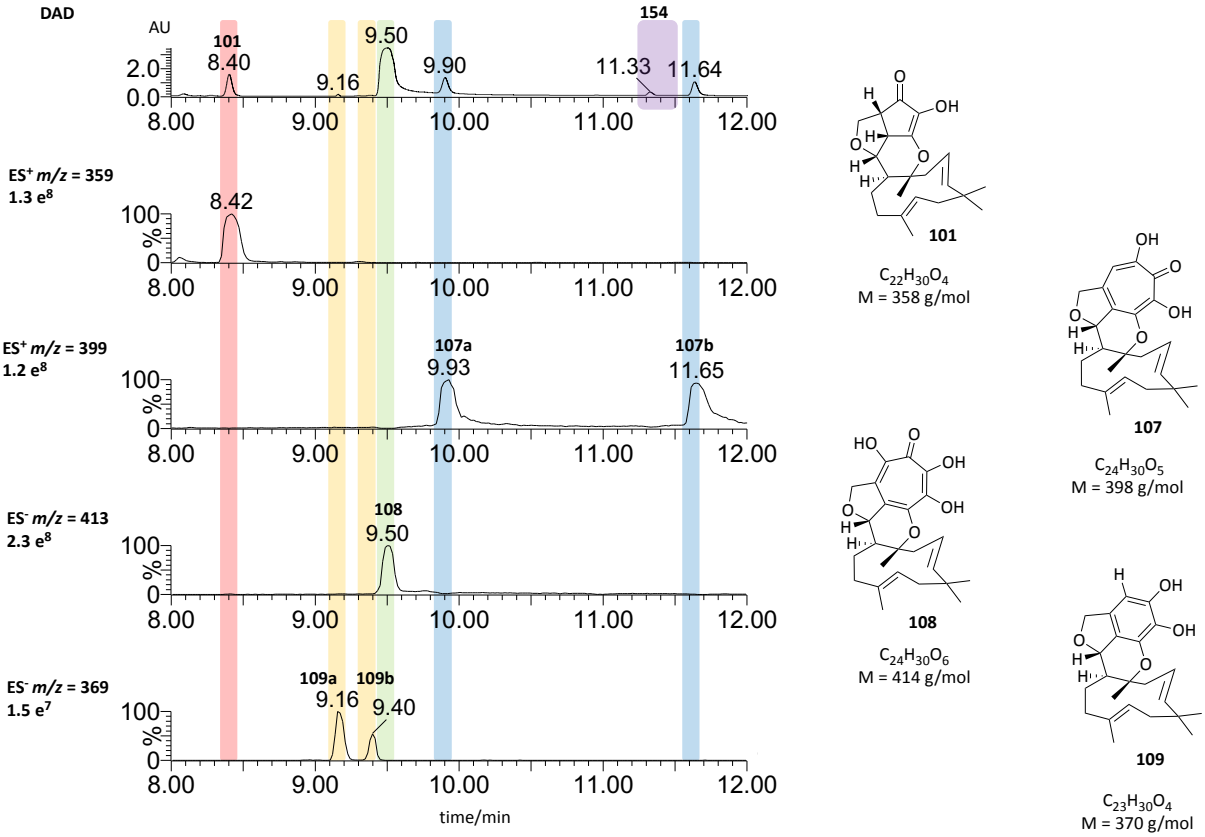


Figure 2.2 Analysis of *A. strictum* extract for known metabolites previously observed with xenovulene A **101**. Analysed by LCMS with analytical gradient A1. Top trace shows Diode array detector (DAD) chromatogram. Other traces show the indicated extracted ion chromatograms corresponding to $[M]H^+$ or $[M\text{-}H]^-$ for compounds **101, 107-109**.

2.1 Identification of Xenovulene A 101

The UV spectrum of the compound eluting at t_R = 8.4 with a nominal mass of 358 showed a maximum absorption at 276 nm, which is consistent with literature data for **101** (Figure 2.3).[82] Subsequent purification of 4 mg enabled full NMR analysis.

The acquired ^1H NMR data were identical to the literature data of **101** (Table **2.1**). However, chemical shift of all carbons in the ^{13}C NMR were observed to be shifted by 0.3 ppm (C-2 by 0.4 ppm) upfield. High resolution mass spectrometry (HRMS) confirmed the molecular formula as $C_{22}H_{30}O_4$ ($[M]Na^+$ $C_{22}H_{30}O_4Na$ calculated 359.2222, found 359.2219).

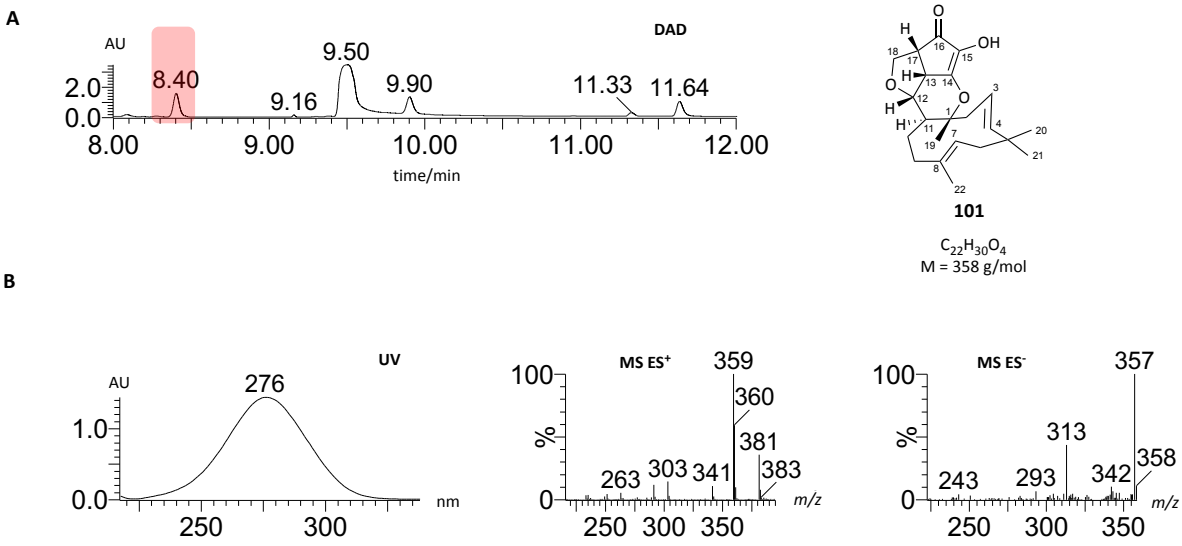


Figure 2.3 Xenovulene A **101**. **A**, Typical DAD chromatogram of *A. strictum* extracts obtained under **101** producing conditions. **B**, Characterisation of xenovulene A **101**.

Table 2.1 NMR data for **101** in CD_3OD (500 MHz) referenced to CD_3OD. Literature data was measured at 400 MHz.

Position	δ_C/ppm 101 referenced to 49.0 ppm measured/ppm	δ_H/ppm (*J* in Hz) 101 referenced to 3.31 ppm measured	δ_C/ppm 101 referenced to 49.0 ppm literature/ppm[82]	δ_H/ppm (*J* in Hz) 101 referenced to 3.4 ppm literature/ppm[82]
1	88.6	-	88.9	-
2	44.8	2.26 (dd, 1H, 14.6, 10.3)	45.2	2.38 (dd, 1H, 14.5, 9.9)
		2.61 (ddd, 1H, 14.7, 2.3, 1.6)		2.70 (br ddd, 1H, 14.5, 2.2, 1.6)
3	121.6	5.22 (ddd, 1H, 15.9, 10.2, 2.4)	121.9	5.32 (ddd, 1H, 15.9, 10.0, 2.2)
4	144.4	5.15 (dd, 1H, 16.0, 1.6)	144.7	5.24 (dd, 1H, 15.9, 1.5)
5	39.1	-	39.4	-
6	42.8	1.76 (br dd, 1H, 12.8, 4.8)	43.1	1.85 (br dd, 1H, 13.0, 4.5)
		2.23 (br dd, 1H, 12.1, 12.0)		2.33 (br dd, 1H, 12.7, 12.7)
7	124.4	5.08 (m, 1H)	124.7	5.17 (m, 1H)
8	137.1	-	137.4	-
9	37.1	1.99 (br dd, 1H, 12.5, 7.3)	37.4	2.08 (br dd, 1H, 12.5, 7.1)
		2.12 (m, 1H)		2.23 (br dd, 1H, 11.6, 11.6)
10	31.9	1.29 (m, 1 H)	32.2	1.37 (m, 1H)
		1.49 (br ddd, 1H, 13.5, 10.2, 7.1)		1.58 (br ddd, 13.2, 10.1, 6.8)
11	44.7	2.13 (br dd, 1H, 10.5, 4.6)	45.0	2.22 (br dd, 1H, 9.8, 5.2)
12	83.4	3.74 (br dd, 1H, 6.0, 4.9)	83.7	3.83 (br dd, 6,0, 6.0)
13	41.9	3.59 (dd, 1H, 5.9, 5.9)	42.2	3.67 (dd, 1H, 5.8, 5.8)
14	166.7	-	167.0	-
15	135.0	-	135.3	
16	200.7	-	201.0	-
17	48.7	2.93 (ddd, 1H, 7.9, 5.6, 1.5)	49.0	3.02 (ddd, 1H, 7.8, 5.6, 1.5)
18	70.6	3.72 (dd, 1H, 9.4, 7.7)	70.9	3.80 (dd, 1H, 9.3, 7.9)
		3.92 (br dd, 1H, 9.3, 1.5)		4.03 (br dd, 9.3, 1.5)
19	23.4	1.42 (s, 3H)	23.7	1.51 (s, 3H)
20	24.6	1.05 (s, 3H)	24.9	1.14 (s, 3H)
21	30.6	1.04 (s, 3H)	30.9	1.13 (s, 3H)
22	17.4	1.63 (s, 3H)	17.7	1.72 (s, 3H)

2.2 Characterisation of the Dihydroxytropolone Meroterpenoid 108

Although the structure of the xenovulene co-metabolite **108** with the molecular formula of $C_{24}H_{30}O_6$ has been previously reported in the literature, no chemical characterisation was available.[84] The purified compound eluting at t_R = 9.5 min was observed as major product in *A. strictum* liquid fermentation, and 113 mg were isolated (by M. Streeck) from 2 L extracted culture. This compound

showed a nominal mass of 414, determined by low resolution mass spectrometry and two UV absorption maxima (273 nm, 365 nm) (Figure 2.4).

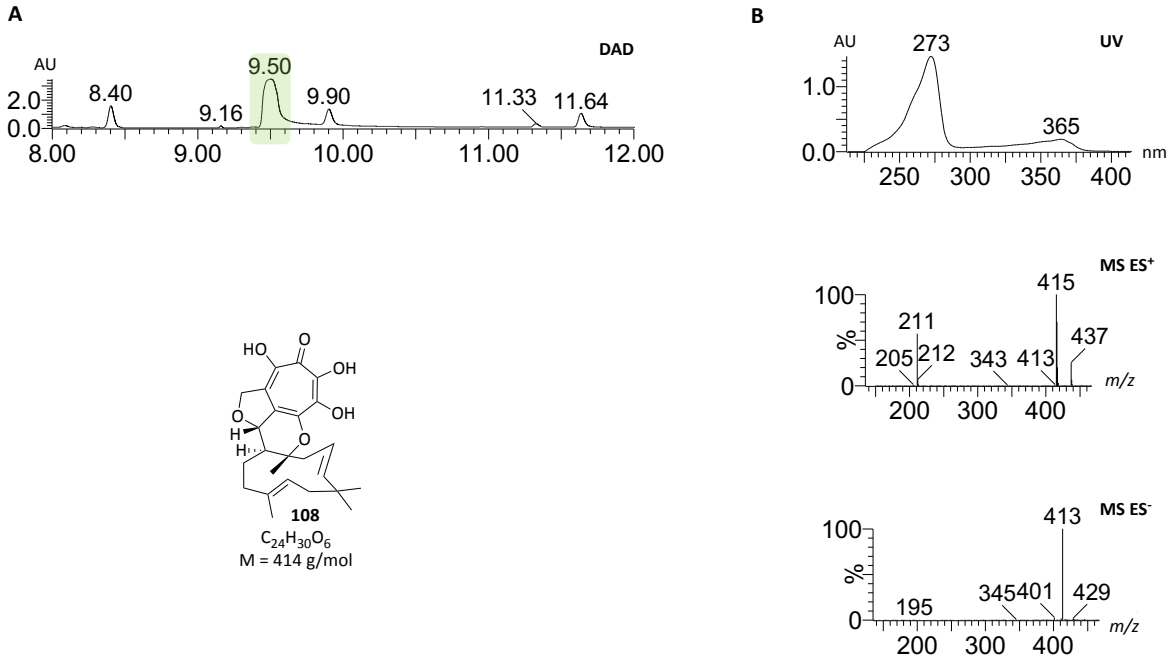


Figure 2.4 LCMS analysis for **108**. **A**, Typical DAD chromatogram of *A. strictum* extracts obtained under **101** producing conditions. **B**, Characterisation of **108**.

Analysis of ^{1}H and ^{13}C NMR data identified 24 carbon atoms and 27 protons (Table 2.2). HRMS confirmed a molecular formula of $C_{24}H_{30}O_6$ ($[M]H^+$ calculated 415.2121, found 415.2119) which suggested three exchangeable protons. The HSQC spectrum revealed four $–CH_3$, five diastereotopic $–CH_2$, five –CH and ten quaternary carbon atoms.

Evaluation of HMBC and $^{1}H,^{1}H$ COSY spectra showed:

- one CH (δ_C 40.9, δ_H 2.05);
- four aliphatic diastereotopic CH_2 (δ_C 43.6, δ_H 2.40, 2.74; δ_C 41.6, δ_H 1.75, 2.16; δ_C 37.9, δ_H 2.13, 2.31; δ_C 28.4, δ_H 1.38, 1.70);
- four CH_3 groups (δ_C 23.0, δ_H 1.33; δ_C 29.7, δ_H 1.02; δ_C 24-4, δ_H 0.93; δ_C 17.2, δ_H 1.60);
- two quaternary C (δ_C 38.5, 87.1);
- two double bond systems (δ_C 119.0, δ_H 4.91; δ_C 144.1, δ_H 5.17; δ_C 123.9, δ_H 5.04; δ_C 136.2).

These make up the typical humulene **106** scaffold (Figure 2.5, Carbon positions 1-11 and 21-24; key HMBC couplings).

The chemical shift of the quaternary C-1 (δ_C 87.1) indicated direct attached to an oxygen atom. By analysis of HMBC and $^{1}H,^{1}H$ COSY, the two double bond systems were determined at positions 3/4 (3: δ_C 119.0, δ_H 4.91; 4: δ_C 144.1, δ_H 5.17) and 7/8 (δ_C 123.9, δ_H 5.04; δ_C 136.2). The $^{3}J_{HH}$ value of 15.8 Hz indicates a *trans* alkene between C-3 and C-4. The geometry of the C-6 and C-7 alkene could

not be determined, but it was assumed to be *trans* as in humulene **106** and xenovulene A **101**. This was supported by very similar δ_C/δ_H values.

Table 2.2 NMR data for **108** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

Position	δ_C/ppm 108	δ_H/ppm (*J* in Hz) 108	HMBC (H to C)
1	87.1	-	-
2	43.6	2.40 (dd, 1H, 14.8, 10.3)	3, 4, 21
		2.75 (d, 1H, 14.8)	1, 3, 4, 11, 21
3	119.0	4.91 (m, 1H)	1, 2, 5
4	144.1	5.17 (dd, 1H, 13.0, 4.9)	2, 5, 22, 23
5	38.5	-	-
6	41.8	1.75 (dd, 1H, 13.0, 4.9)	4, 5, 7, 8, 23
		2.16 (m, 1H)	4, 5, 7, 8, 22, 23
7	123.9	5.04 (m, 1H)	5, 6, 24
8	136.2	-	-
9	37.9	2.13 (m, 1H)	7, 8, 10, 11, 24
		2.31 (m, 1H)	7, 8, 10, 11, 24
10	28.4	1.38 (m, 1 H)	1, 8, 9, 11, 12
		1.70 (m, 1H)	8, 9, 11,12
11	40.9	2.05 (br dd, 1H, 10.8, 8.2)	1, 2, 9, 10, 12, 13, 21
12	84.5	4.91 (m, 1H)	1, 11, 10, 13, 14, 15, 18, 19
13	123.5	-	-
14	142.4	-	-
15	144.6	-	-
16	150.6	-	-
17	160.3	-	-
18	146.4	-	-
19	127.4	-	-
20	72.6	5.05 (m, 1H)	13, 14, 17, 18, 19
		5.21 (dd, 1H, 13.7, 1.1)	12, 13, 16, 17, 18, 19
21	23.0	1.33 (s, 3H)	1, 2, 11
22	29.7	1.02 (s, 3H)	4, 5, 6 23
23	24.4	0.93 (s, 3H)	4, 5, 6, 22
24	17.2	1.60 (s, 3H)	7, 9, 8

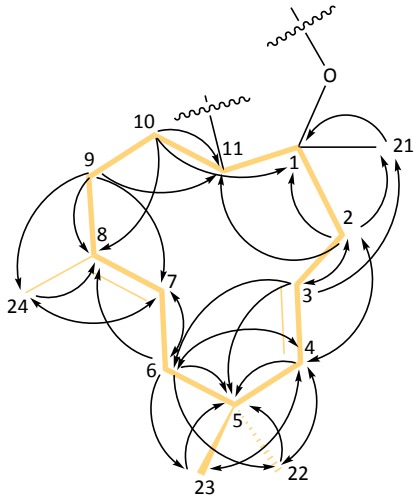


Figure 2.5 HMBC correlations in the humulene ring in **108**. Arrow direction indicates H to C.

Consideration of chemical shift values also proposed a direct attachment to oxygen for C-20 (δ_C 72.6, δ_H 5.05, 5.23) and C-12 (δ_C 84.5, δ_H 4.91). $^3J_{CH}$ coupling between C-12 (δ_C 84) and H-20 (δ_H 5.23) indicated an ether like linkage at that position. Analysis of HMBC correlations of H-12 (δ_H 4.91) identified the humulene carbons C-1, C-10, C-11 as well as C-13 (δ_C 123.5) and C-19 (δ_C 127.4) in coupling distance. $^3J_{HH}$ of H-12 to H-11 (δ_H 2.05) confirmed the direct attachment to

humulene and determined position C-12 for this -CH-O- group. Further HMBC correlation showed that the ether linked CH_2 group (δ_C 72.6, δ_H 5.05, δ_H 5.23) is in coupling distance to quaternary carbons C-13 (δ_C 123.5), C-14 (δ_C 142.4), C-17 (δ_C 160.3), C-18 (δ_C 146.4) and C-19 (δ_C 127.4). In the $^1H,^1H$ COSY spectrum only geminal $^2J_{HH}$ coupling of δ_H 5.23 and δ_H 5.05 was detected which proposed its position at C-20. The relative stereochemistry at positions C-1, C-11 and C-12 were assumed to be identical to xenovulene A **101**, but not further determined.

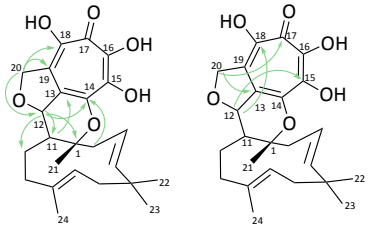


Figure 2.6 HMBC correlations of **108**.

Chemical shifts of the seven remaining quaternary carbons suggested their involvement in an aromatic system and for five of them (δ_C 142.4.0, δ_C 144.6, δ_C 146.4, δ_C 150.6 and δ_C 160.3) the attachment to an oxygen. HMBC correlations of the carbon at δ_C 123.5 to protons at C-12, C-20 and C-11 suggested its position at C-13. Chemical shift and $^3J_{CH}$ couplings of the carbon at δ_C 127.4 to both protons (δ_H 5.23, δ_H 5.05) at position C-20 and to δ_H 4.9 at C-12 suggested it to be at position C-19. The carbon at δ_C 142.4 showed a $^3J_{CH}$ coupling to the proton at C-12 and $^4J_{CH}$ couplings to δ_H 5.05 at C-20 as well as δ_H 2.40 at C-2. It was thus assigned at position C-14. The carbon at δ_C 144.6 displays a $^4J_{CH}$ correlation to proton δ_H 4.91 at C-12, but not to protons attached at C-20 and was thus assigned at C-15. The carbon at δ_C 146.4 couples to both protons at C-20 and to the one proton at C-12. It was assigned at C-18. The carbon at δ_C 160.3 with the highest chemical shift was assigned as a carbonyl. A $^4J_{CH}$ coupling to δ_H 5.23 and δ_H 5.05 at C-20 suggested its position at C-17. For the carbon at δ_C 150.6 no correlations were identified in the HMBC spectrum which strongly indicates its position at C-16 (Figure 2.6 for key HMBC correlations).

For further structure validation 30 mg of **108** were methylated with an excess of trimethylsilyl (TMS) diazomethane in tetrahydrofuran (THF) / methanol (Figure 2.7A).[120,121] LCMS analysis showed the formation four major compounds with t_R = 9.9 (I), 10.1 (II), 10.4 (III) and 10.7 (IV) min (Figure 2.7B).

The nominal mass of all four peaks was determined by low resolution mass spectrometry to be 456, which corresponds to **108** methylated at three sites. Two distinct UV absorption maxima (I = 265, 356, II = 270, 342, III = 264, 354, IV = 261, 351) were observed for all compounds (Figure 2.8).

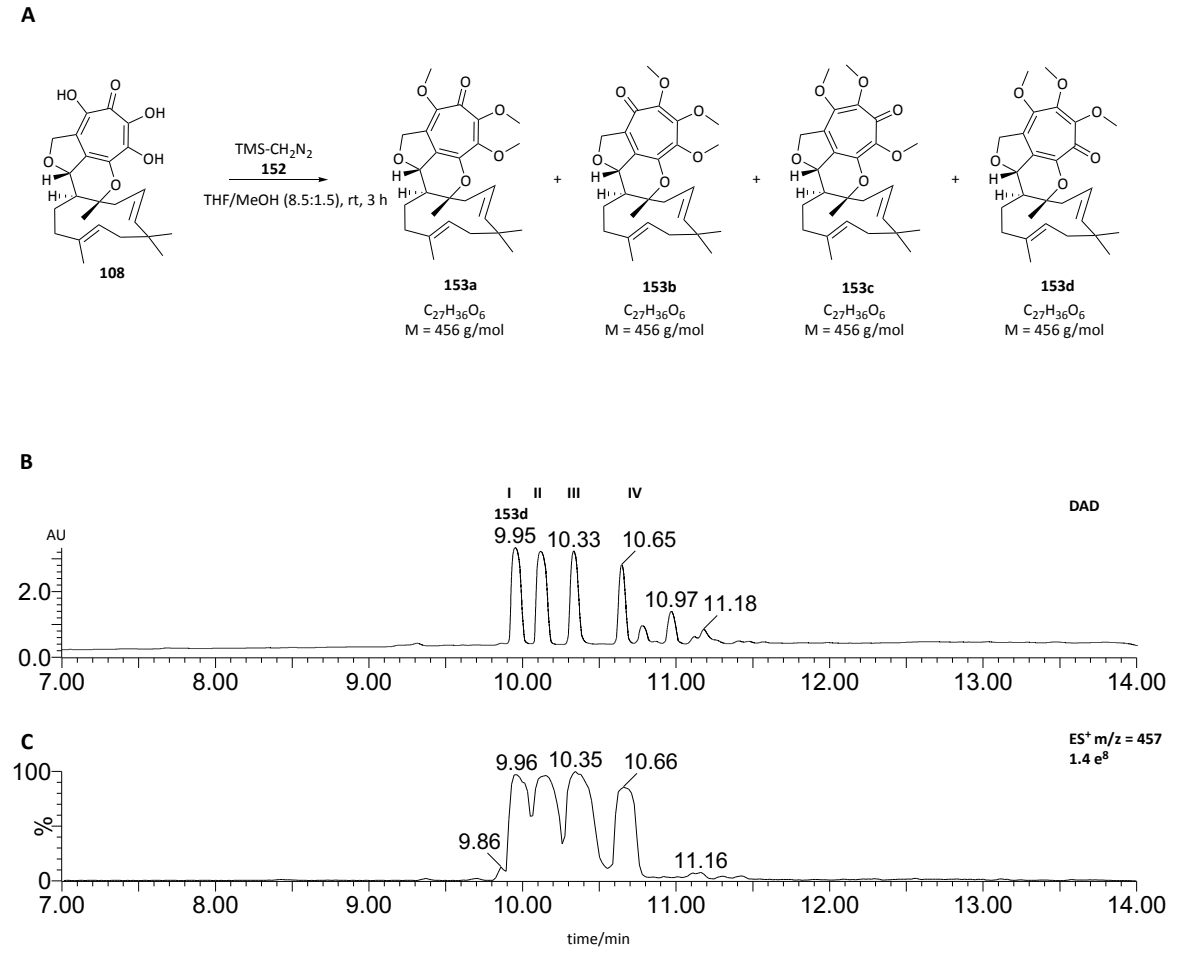


Figure 2.7 Methylation of **108** with TMS-diazomethane **152**. **A**, Formation of four **153** isomers. **B**, DAD chromatogram of reaction mixture after 3 h. **C** Extracted ion chromatogram at $[M]H^+$ = 457.

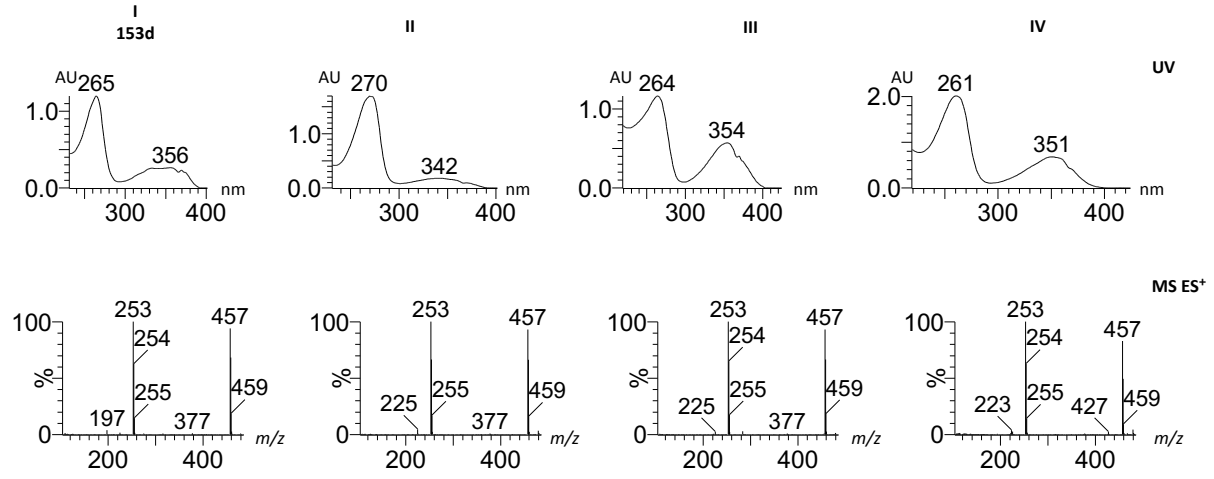


Figure 2.8 Detection of four products I-IV upon methylation of **108**.

The four peaks were purified by LCMS (I: 4.1 mg, II: 2.5 mg, III: 1.4 mg, IV: 2.6 mg). The ^{1}H NMR spectrum of compound **I** showed two new singlets with a chemical shift of δ_H 3.85 and δ_H 3.94 integrating for respectively three and six protons. ^{1}H NMR spectra for compounds II-IV showed each three new singlets between δ_H 3.89 and δ_H 3.99 integrating for three protons each (Figure 2.9). This

indicated that methylation of the aromatic alcohols trapped the four tautomers of compound **108** (Figure 2.7A).

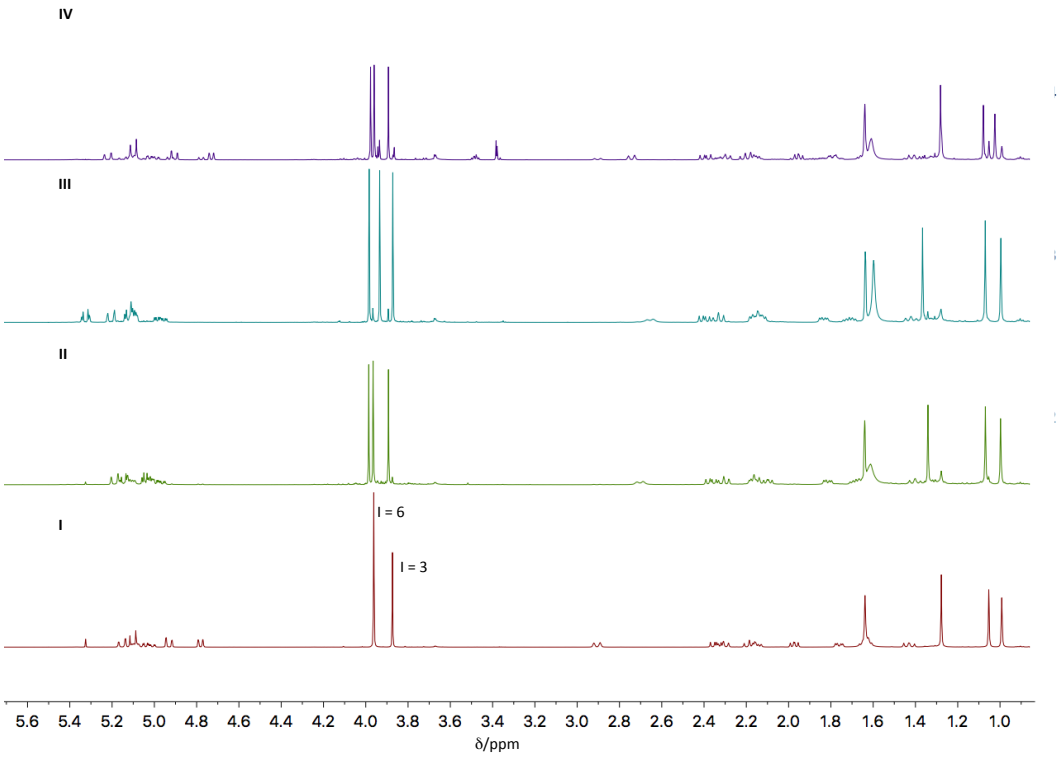
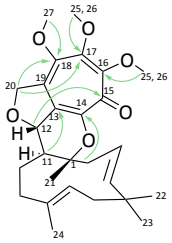


Figure 2.9 ^{1}H NMR of compounds **I (153d), II, III** and **IV**.

Full NMR characterisation and HRMS ($C_{27}H_{37}O_6$, $[M]H^+$ calculated 457.2590, found 457.2589) was carried out for compound **I** (compounds **II**, **III** and **IV** degraded before data was acquired). Protons and carbons involved in the humulene, dihydropyran and the dihydrofuran rings were assigned according to **108**, as chemical shifts and HMBC couplings were very similar (Table 2.3). Carbons within $^{3}J_{CH}$ distance of the methyl groups were identified in the HMBC (δ_C 157.3, δ_C 157.1, δ_C 146.8). Carbon at δ_C 146.8 showed $^{3}J_{CH}$ to δ_H 3.85 (at δ_C 61.2) and δ_H 5.07 and δ_H 4.91 (at C-20). An additional $^{4}J_{CH}$ to δ_H 4.75 (at C-12) proposed its position at C-18. At least one of the carbons at δ_C 157.1 and δ_C 157.3 couples with δ_H 5.07 and δ_H 4.91 (at C-20), which indicates it at C-17. Due to close proximity of the chemical shifts the correlating carbon could not be assigned. However, it was assumed that the other carbon has no correlations, which would propose it at position C-16. This is further underpinned by the $^{4}J_{CH}$ coupling of carbon δ_C 173.4 to δ_H 4.75 (at C-12) which indicates its position at C-15. Taken together these data strongly suggest that compound **I** is the 15-keto isomer **153d** (Figure 2.10).



153d

Figure 2.10 Key HMBC correlations for assignments of tropolone associated carbon atoms in **153d**.

Table 2.3 NMR data for **153d** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

Position	δ_C/ppm 153d	δ_H/ppm (*J* in Hz) 153d	HMBC (H to C)
1	85.4	-	-
2	43.0	2.31 (m, 1H)	3, 4, 14
		2.88 (ddd, 1H, 14.7, 2.2, 2.2)	1, 3, 4, 10, 11, 21
3	119.9	4.99 (ddd, 1H, 15.9, 10.4, 2.5)	4, 5, 6
4	143.2	5.12 (dd, 1H, 15.9, 1.7)	3, 5, 6, 22, 23
5	38.4	-	-
6	41.8	1.73 (dd, 1H, 12.9, 1.7)	4, 5, 7, 8, 23
		2.16 (m, 1H)	5, 6, 7, 22, 23
7	123.9	5.06 (m, 1H)	5, 24
8	136.3	-	-
9	38.1	2.11 (m, 1H)	1, 10, 24
		2.28 (m, 1H	8, 10, 11, 13, 24
10	28.5	1.40 (m, 1H)	1, 8, 9, 11, 12
		1.61 (m, 1H)	1, 8, 9, 11, 12
11	39.8	1.95 (dd, 1H, 10.6, 8.1)	1, 2, 9, 10,12, 13, 21
12	84.5	4.75 (dd, 1H, 10.6, 1.4)	10, 11, 13, 14, 17, 18, 19
13	123.8	-	-
14	152.6	-	-
15	173.4	-	-
16,17	157.1, 157.3	-	-
18	146.8	-	-
19	132.0	-	-
20	71.8	4.91 (dd, 1H, 13.9, 1.4)	13, 14, 16/17, 18, 19
		5.07 (d, 1H, 13.9)	12, 13, 16/17, 18, 19
21	22.6	1.25 (s, 3H)	1, 2, 3, 11
22	30.1	1.03 (s, 3H)	4, 5, 6, 23
23	24.3	0.96 (s, 3H)	4, 5, 6, 22
24	17.1	1.61 (s, 3H)	7, 8, 9
25,26	60.3, 61.9	3.93, 3.94 (s, 6H)	16, 17
27	61.2	3.85 (s, 3H)	18

2.3 Characterisation of Hydroxytropolone Meroterperpenoid Isomers 107a/b

Analysis of extracted ion chromatograms of *A. strictum* extracts for the known intermediate **107**, with the molecular formula $C_{24}H_{30}O_5$, identified two compounds satisfying the nominal mass of 398 (determined by low resolutions mass spectroscopy). Both compounds showed characteristic UV absorption maxima (**107a** 271, 363 and **107b** 249, 330, 389) but different retention times (**107a** t_R = 9.9 and **107b** t_R = 11.6) (Figure 2.11A and B).

Although the structure of **107a** has been published no chemical characterisation was available.[84] Purification (by M. Streeck) of 18 mg **107a** and 9 mg **107b** from 2 L of *A. strictum* culture extracts enabled full NMR characterisation. Analysis of ^1H and ^{13}C NMR data identified 24 carbon atoms and 28 protons for each **107a** and **107b** (Table 2.4). HRMS confirmed a molecular formula of $C_{24}H_{30}O_6$

for both compounds ([M]-H^+ calculated 399.2127, found **107a**: 399.2168 and **107b**: 399.2166) which suggested two exchangeable protons.

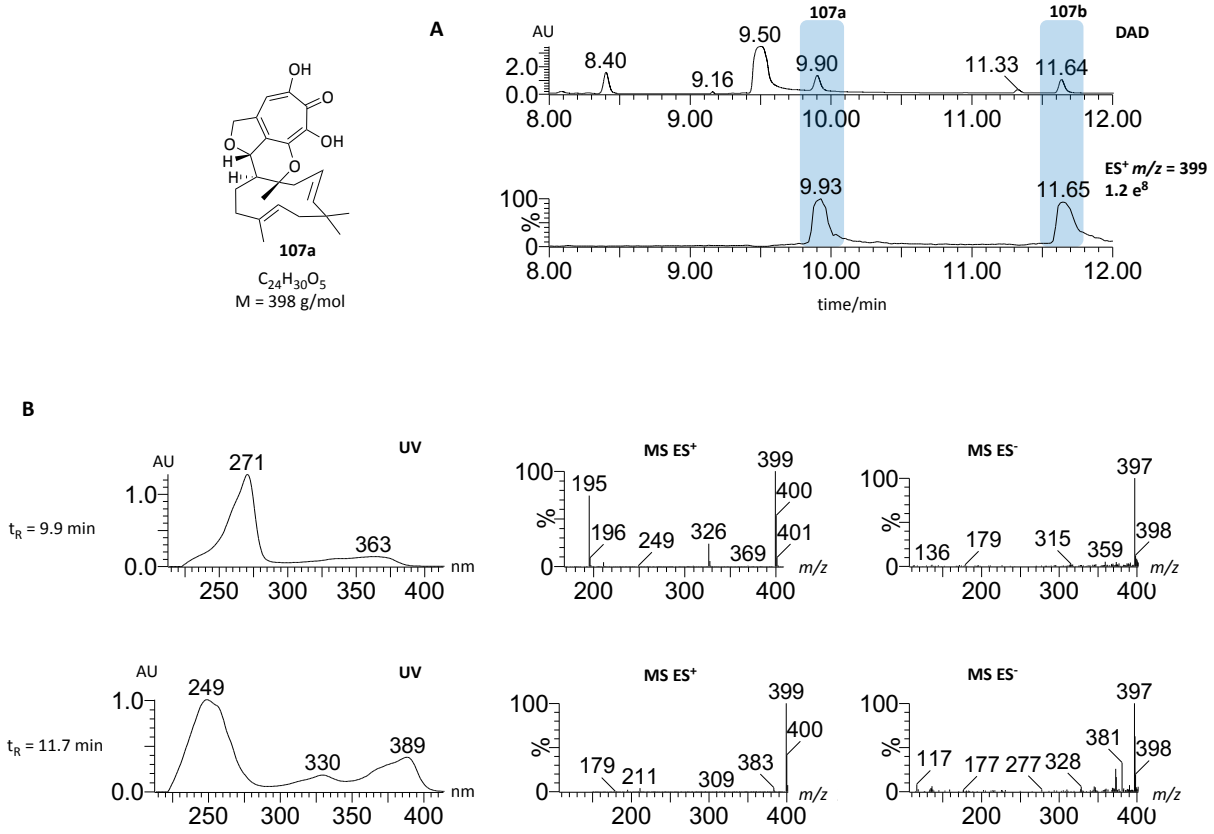


Figure 2.11 Compounds **107a** and **b**. **A**, Typical DAD chromatogram of *A. strictum* extracts obtained under **101** producing conditions and extracted ion chromatograms corresponding to [M]H^+ = 399. **B**, Chemical characterisation of two compounds with the nominal mass of 398 eluting at t_R = 9.9 min and 11.7 min.

The HSQC spectrum revealed four $-CH_3$, five diastereotopic $-CH_2$, six –CH and nine quaternary carbon atoms for both compounds (Table 2.4). Comparison with ^{1}H NMR, ^{1}H,^{1}H COSY and HMBC spectra obtained for **108** enabled rapid verification and assignment of the humulene, dihydropyran and dihydrofuran proton and carbon signals to the respective positions in **107a** and **107b**.

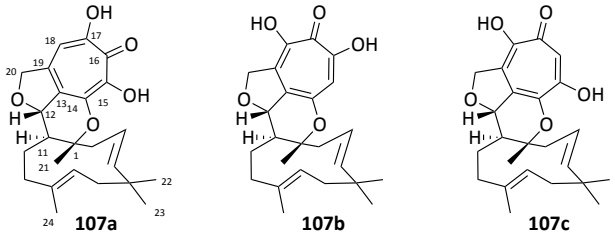


Figure 2.12 Possible structures to satisfy the HRMS and ^{1}H NMR results.

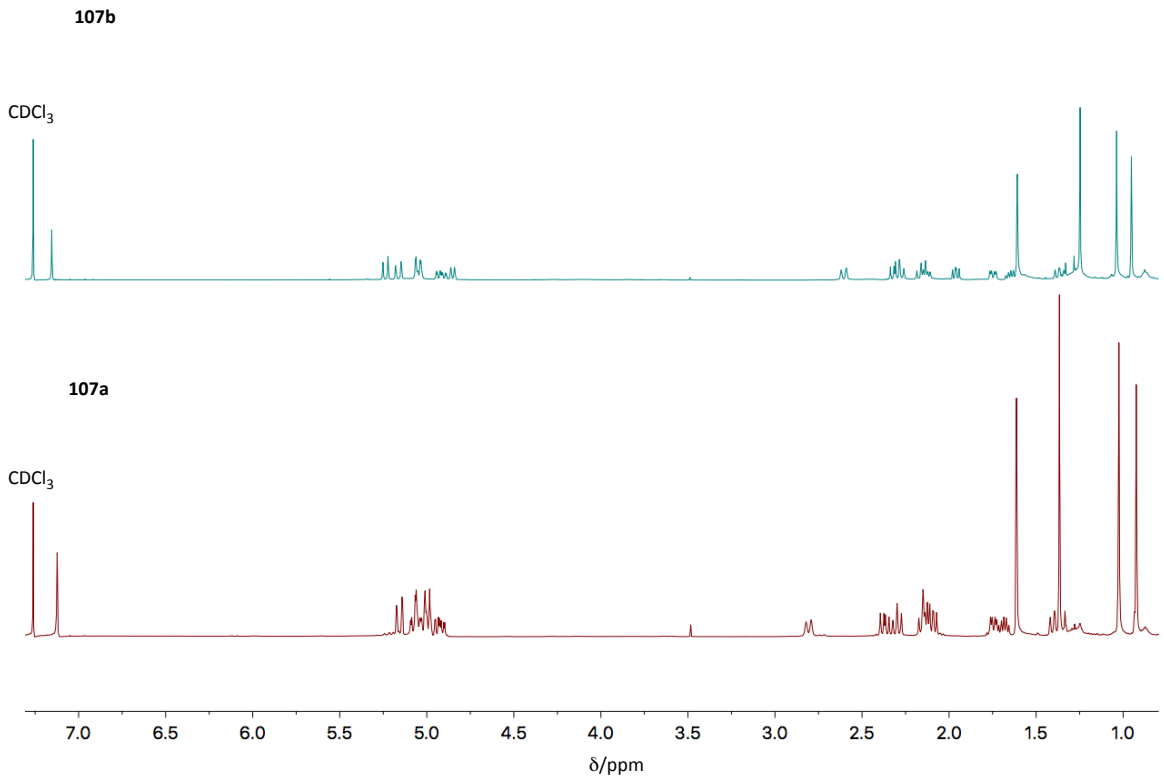
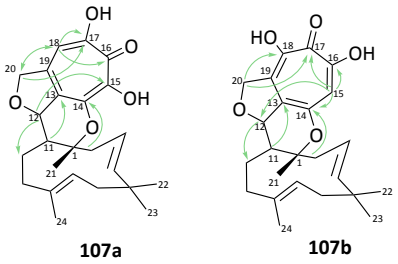


Figure 2.13 1H NMR of purified compounds **107a** and **107b**.

A sharp singlet integrating for one proton was identified in the ^{1}H NMR data for **107a** (δ_H 7.12 at δ_C 108.9) and **107b** (δ_H 7.16 at δ_C 112.3) (Figure 2.13). Consideration of chemical shifts proposed the attachment to an aromatic system in both compounds, which led to the suggested possible structures **107a**, **107b** or **107c** (Figure 2.12).

Further analysis of HMBC correlation data of **107a** showed a $^{3}J_{CH}$ coupling of proton δ_H 7.12 with C-20 (δ_C 72.5) and thus the proton δ_H 7.12 was proposed to be at position C-18, as for positions C-15 and C-16 no correlations to C-20 were expected. HMBC data of compound **107b** showed no such correlation between proton δ_H 7.16 and C-20 (δ_C 72.5). Instead a $^{4}J_{CH}$ interaction between proton δ_H 7.16 and C-12 (δ_C 84.3) was detected which suggested its attachment at position C-15 (Figure 2.14). Quaternary tropolone carbons were assigned based on indicated HMBC correlations in analogy to compound **108**. The relative stereochemistry at positions C-1, C-11 and C-12 were assumed to be identical to xenovulene A **101**, but not further determined.



107a **107b**

Figure 2.14 Key HMBC correlations for **107a** and **107b**.

Table 2.4 NMR data for **107a** and **107b** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

Position	δ_C/ppm 107a	δ_H/ppm (*J* in Hz) 107a	HMBC (H to C 107a)	δ_C/ppm 107b	δ_H/ppm (*J* in Hz) 107b	HMBC (C to H) 107b
1	87.5	-	-	85.8	-	-
2	43.7	2.37 (dd, 1H, 14.7, 10.4)	3, 4, 14	43.3	2.31 (m, 1H)	3, 4, 14
		2.81 (d, 1H, 15.0)	1, 3, 4, 5, 11, 21		2.60 (ddd, 1H, 14.9., 3.9, 1.9)	1, 3, 4, 5, 11, 21
3	119.4	4.93 (ddd, 1H, 15.8, 10.3, 2.8)	1, 2, 5	119.2	4.91 (ddd, 1H, 15.8, 10.5, 2.6)	1, 2, 5
4	143.9	5.16 (dd, 1H, 15.8, 1.7)	1, 5, 22, 23	143.8	5.16 (dd, 1H, 15.8, 1.8)	2, 5, 22, 23
5	38.5	-	-	38.5	-	-
6	41.7	1.75 (br dd, 1H, 13.0, 4.9)	4, 5, 7, 8, 23	41.7	1.75 (dd, 1H, 13.0, 4.7)	4, 5, 7, 8, 23
		2.16 (m, 1H)	4, 5, 7, 8, 22, 23		2.16 (m, 1H)	5, 7, 8, 22, 23
7	124.0	5.05 (m, 1H)	5, 24	123.8	5.04 (m, 1H)	5, 6, 24
8	136.2	-	-	136.3	-	-
9	38.0	2.13 (m, 1H)	7, 8, 10, 11, 24	38.0	2.14 (m, 1H)	7, 8, 10, 11, 24
		2.30 (br dd, 1H, 12.2, 12.2)	7, 8, 10, 11, 24		2.28 (m, 1H)	7, 8, 10, 11, 24
10	28.3	1.41 (m, 1 H)	1, 8, 9, 11, 12	28.3	1.37 (m, 1 H)	1, 8, 9, 11, 12
		1.68 (m, 1H)	8, 9, 11, 12		1.66 (m, 1H)	8, 9, 11, 12
11	41.1	2.09 (m, 1H)	1, 2, 9, 10, 12, 13, 21	40.5	1.96 (br dd, 1H, 10.7, 8.2)	1, 2, 9, 10, 12, 13, 21
12	85.2	4.99 (m, 1H)	10, 11	84.3	4.85 (br d, 1H, 10.7)	10, 11, 13, 15, 18, 19
13	127.7	-	-	122.8	-	-
14	148.2	-	-	152.6	-	-
15	149.3	-	-	112.3	7.16 (s, 1H)	12, 13, 14, 16, 17, 19
16	164.2	-	-	159.4	-	-
17	159.8	-	-	165.9	-	-
18	108.9	7.12 (s, 1H)	13, 14, 16, 17, 19, 20	147.1	-	-
19	138.2	-	-	134.5	-	-
20	75.6	5.00 (m, 1H)	12, 13, 16, 18, 19	72.5	5.05 (dd, 1H, 14.1, 2.3)	13, 14, 18
		5.08 (dd, 1H, 13.3, 3.3)	13, 19		5.21 (dd, 1H, 14.1, 1.0)	12, 13, 17, 18, 19
21	23.1	1.37 (s, 3H)	1, 2, 11	22.8	1.33 (s, 3H)	1, 2, 11
22	29.7	1.02 (s, 3H)	4, 6, 5, 23	30.0	1.02 (s, 3H)	4, 6, 5, 23
23	24.4	0.92 (s, 3H)	4, 6, 5, 22	24.3	0.93 (s, 3H)	4, 6, 5, 22
24	17.2	1.61 (s, 3H)	7, 9, 8	17.1	1.60 (s, 3H)	7, 9, 8

2.4 Characterisation of the Tropolone Meroterpenoid 154

An unknown compound with the nominal mass of 382 (determined through low resolution mass spectrometry) was frequently detected by LCMS of *A. strictum* culture extracts alongside xenovulene A **101**. The compound eluted at t_R = 11.3 and showed two UV absorption maxima (257 nm, 361 nm), similar to compounds **107**, **108** and typical for tropolones (Figure 2.15).[122] Purification of 5 mg of this compound from 400 mL *A. strictum* culture extract enabled full NMR characterisation.

Analysis of ^1H and ^{13}C NMR data identified 24 carbon atoms and 29 protons. HRMS confirmed a molecular formula of $C_{24}H_{30}O_4$ ($[M-H]^-$ $C_{24}H_{29}O_4$ calculated 381.2066, found 381.2062) which proposed one exchangeable proton. The HSQC spectrum revealed four $-CH_3$, five diastereotopic –

CH_2, seven –CH and eight quaternary carbons (Table 2.5). Comparison with ^{1}H NMR, ^{1}H,^{1}H COSY and HMBC spectra obtained for **108** enabled rapid verification and assignment of humulene, dihydropyran and dihydrofuran proton and carbon signals to the respective positions in **154**.

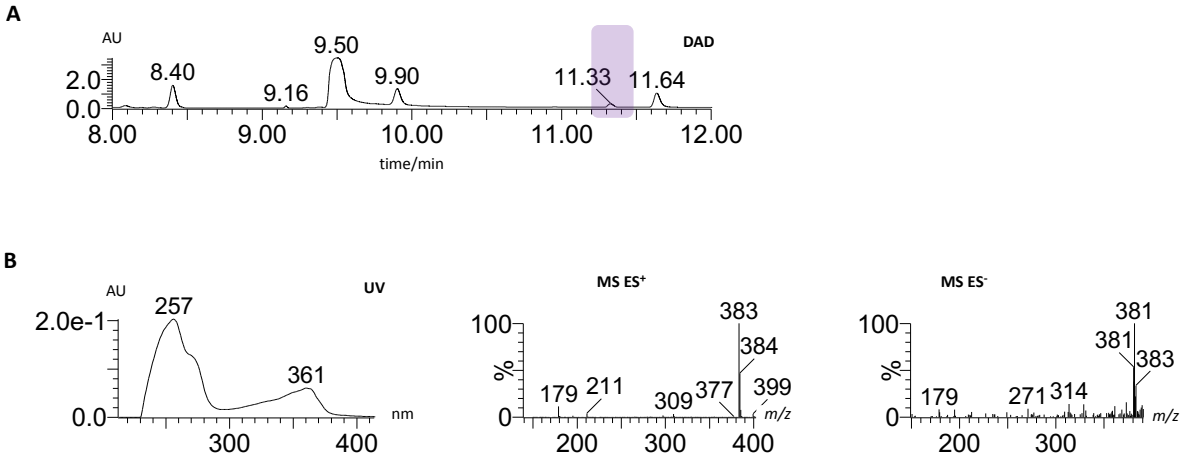


Figure 2.15 Compound **154**. **A,** Typical DAD chromatogram of *A. strictum* extracts obtained under **101** producing conditions. **B,** Characterisation of compound **154** eluting at t_R = 11.3 min.

Two sharp singlets integrating for one proton each were identified in the ^{1}H NMR data (δ_H 6.96 at δ_C 110.1, δ_H 6.91 at δ_C 113.1). Consideration of chemical shifts indicated the attachment to an aromatic system, which led to the proposal of a less oxidised tropolone scaffold (**154**, **154a** and **154b**). The three proposed structures could satisfy the molecular formula (Figure 2.16A).

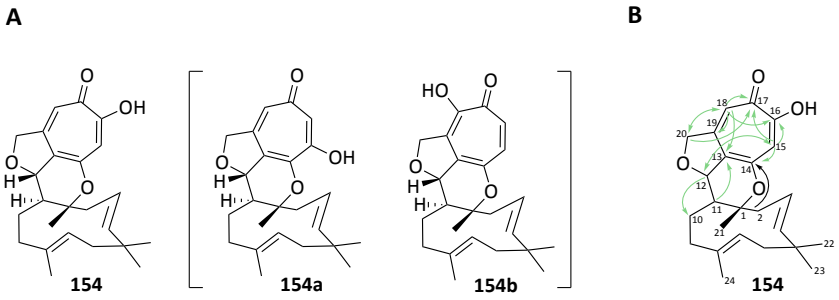


Figure 2.16 Compound **154**. **A** Proposed structures to satisfy the nominal mass. **B** Key HMBC correlations of **154**.

However, possible structure **154b** was discarded as vicinal protons would couple and result in doublets in the ^{1}H NMR spectrum. In **154** and **154a** a $^3J_{CH}$ coupling between position 18 and 20 was expected. And indeed for proton δ_H 6.91 a coupling to δ_C 75.2 (C-20) was observed, which suggested its attachment at C-18. However, to satisfy **154a** further $^3J_{CH}$ coupling between position 18 and 16 were expected, but not observed. Instead a $^4J_{CH}$ interaction between proton δ_H 6.96 and δ_C 84.2 (C-12) suggested its attachment at position C-15. Further $^3J_{CH}$ correlation of the two aromatic protons led to the assignment of tropolone carbons as indicated (Figure 2.16B).

Table 2.5 NMR data for **154** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

Position	δ_C/ppm 154	δ_H/ppm (*J* in Hz) 154	HMBC (H to C)
1	86.7	-	-
2	43.1	2.31 (dd, 1H, 14.6, 10.4)	3, 4, 14
		2.63 (ddd, 1H, 14.7, 3.9, 1.9)	1, 3, 4, 5, 11, 21
3	119.2	4.91 (m, 1H)	2, 5
4	143.9	5.17 (dd, 1H, 15.8, 1.8)	1, 2, 5, 22, 23
5	38.5	-	-
6	41.7	1.75 (dd, 1H, 12.9, 4.7)	4, 5, 7, 8, 22
		2.16 (m, 1H)	5, 7, 8, 22, 23
7	123.8	5.05 (m, 1H)	5, 6, 24
8	136.3	-	-
9	38.0	2.12 (m, 1H)	7, 8, 10, 11, 24
		2.26 (m, 1H)	7, 8, 10, 11, 24
10	28.3	1.35 (m, 1 H)	1, 8, 9, 11, 12, 13
		1.64 (m, 1H)	8, 9, 11, 12
11	40.5	1.94 (br dd, 1H, 10.7, 8.2)	1, 2, 9, 10, 12, 13, 21
12	84.2	4.80 (dd, 1H, 10.6, 2.1)	10, 11, 13, 14, 15, 18, 19
13	121.3	-	-
14	158.9	-	-
15	111.0	6.96 (s, 1H)	12, 13, 14, 16, 17
16	166.1	-	-
17	173.5	-	-
18	113.1	6.91 (s, 1H)	13, 16, 17, 19, 20
19	152.1	-	-
20	75.2	4.90 (m, 1H)	12, 13, 17, 18, 19
		4.97 (m, 1H)	13, 14, 17, 18, 19
21	22.7	1.26 (s, 3H)	1, 2, 11
22	30.0	1.04 (s, 3H)	4, 6, 5, 23
23	24.3	0.97 (s, 3H)	4, 6, 5, 22
24	17.1	1.61 (s, 3H)	7, 9, 8

2.5 Characterisation of Phenolic Meroterpenoids 109a and 109b

LCMS analysis of *A. strictum* extracts for the known intermediate **109** identified two compounds satisfying the nominal mass of 370 (determined by low resolutions mass spectrometry) (Figure 2.17).

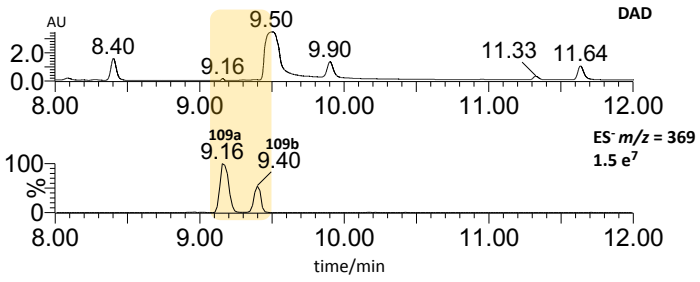


Figure 2.17 Typical DAD chromatogram of *A. strictum* extracts obtained under **101** producing conditions and extracted ion chromatogram corresponding to $[M-H]^-$ = 369.

Both compounds showed similar, weak UV absorption (Figure 2.18) but different retention times (**109a** t_R = 9.2 and **109b** t_R = 9.4). Although the structure of **109a** has been published no chemical characterisation was available.[84] Purification of 1.1 mg **109a** and 0.7 mg **109b** from 400 mL *A. strictum* culture extracts enabled full NMR characterisation for **109a** but only ^{1}H NMR for **109b**. Analysis of ^{1}H and ^{13}C NMR data of **109a** identified 23 carbon atoms and 28 protons (Table 2.6).

HRMS of **109a** confirmed a molecular formula of $C_{23}H_{30}O_4$ for ([M-H]$^-$ calculated 369.2066, found 369.2065) which suggested two exchangeable protons.

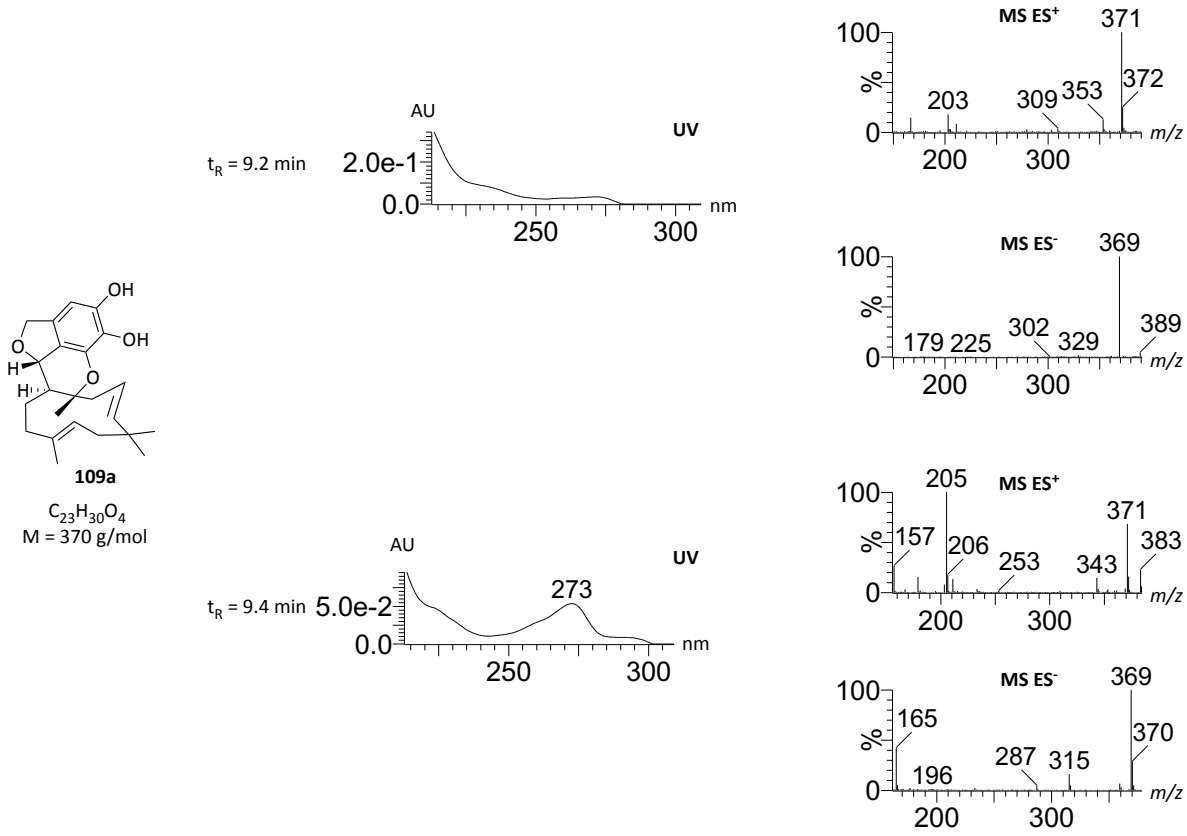


Figure 2.18 Characterisation of two compounds eluting at t_R = 9.2 min and t_R = 9.4 min with a nominal mass of 370.

The HSQC spectrum revealed four –CH_3, five diastereotopic –CH_2, six –CH and eight quaternary carbon atoms for **109a** (Table 2.6). Comparison with ^{1}H NMR, ^{1}H,^{1}H COSY and HMBC spectra obtained for **108** enabled rapid verification and assignment of the humulene, dihydropyran and dihydrofuran proton and carbon signals to the respective positions in **109a**.

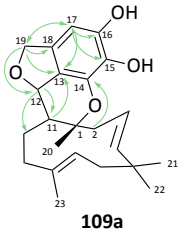


Figure 2.19 Key HMBC correlations for compound **109a**.

Table 2.6 NMR data for **109a** and **b** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$.

Position	δ_C/ppm 109a	δ_H/ppm (*J* in Hz) 109a	HMBC (H to C)	δ_H/ppm (*J* in Hz) 109b
1	86.8	-	-	-
2	43.1	2.34 (dd, 1H, 14.7, 10.5)	3, 4, 5, 14	2.28 (m)
		2.60 (ddd, 1H, 14.7, 2.3, 2.3)	1, 2, 3, 4, 21	2.50 (ddd, 14.6, 2.3, 2.3)
3	119.7	4.93 (ddd, 1H, 15.8, 10.5, 2.5)	1, 2, 5	4.96 (m)
4	143.2	5.13 (dd, 1H, 15.8, 1.8)	2, 3, 5, 20, 21	5.10 (dd, 1H, 15.9, 1.7)
5	38.4	-	-	-
6	41.6	1.71 (m, 1H)	4, 5, 7, 8, 20, 21	1.72 (m)
		2.15 (m, 1H)	5, 7, 8, 21	2.07-2.22 (m)
7	123.4	5.00 (m, 1H)	5, 23	4.98-5.03 (m)
8	136.5	-		-
9	37.9	2.11 (m, 1H)	7, 8, 10, 11, 23	2.07-2.22 (m)
		2.22 (1H)	7, 8, 10, 11, 23	2.07-2.22 (m)
10	29.1	1.29 (m, 1 H)	1, 8, 9, 11, 12, 13	contamination at same δ_H
		1.65 (m, 1H)	8, 9, 11, 12	1.69 (m)
11	43.4	1.73 (m, 1H)	1, 2, 9, 10, 12, 13	contamination at same δ_H
12	82.3	4.78 (m, 1H)	10, 11, 13, 14, 18	4.75 (br d)
13	118.7	-	-	-
14	138.5	-	-	-
15	128.8	-	-	
16	145.3	-	-	-
17	100.7	6.43 (s, 1H)	13, 15, 16, 19	-
18	131.6	-	-	-
19	73.4	4.80 (m, 1H)	12, 13, 17, 18	4.98-5.03 (m)
		5.01 (m, 1H)	13, 17, 18	5.08
20	22.8	1.25 (s, 3H)	1, 2, 11	1.25 (s)
21	30.3	1.03 (s, 3H)	4, 6, 5, 22	1.04 (s)
22	24.4	0.95 (s, 3H)	4, 6, 5, 21	0.95 (s)
23	17.1	1.59 (s, 3H)	7, 9, 8	1.59 (s)

A sharp singlet integrating for one proton was identified in the ^{1}H NMR data for **109a** (δ_H 6.43 at δ_C 100.7). Consideration of the chemical shift proposed the attachment to an aromatic scaffold. HMBC correlation of proton δ_H 6.43 was observed with δ_C 73.4 (C-19), δ_C 118.7 (C-13), δ_C 145.3 (C-16) and δ_C 128.8 (C-15), which suggested its position at C-17. Chemical shift values of three quaternary carbons (δ_C 128.8, δ_C 138.5 and δ_C 145.3) indicated a direct attachment to an oxygen. For two of these (δ_C 128.8, δ_C 145.3) a correlation to the proton (δ_H 6.43) at C-17 was identified. Consideration of HMBC correlations only did not allow differentiation between positions C-15 and C-16 for these carbons. However, simulation of chemical shifts (nmrdb.org)[123–125] suggested δ_C 128.8 (simulated at δ_C 131.5) at C-15 and δ_C 145.3 (simulated δ_C 146.8) at C-16. One carbon (δ_C 138.5) showed no correlation to the proton at position 17, instead showing a correlation to a proton (δ_H 2.34) at C-2 and was thus assigned at C-14. The quaternary carbon at δ_C 131.6 only couples with protons at C-19 and was thus assigned at position C-18.

The second isolated compound with a nominal mass of 370 was only analysed by ^{1}H NMR. The collected spectroscopic information was relatively poor, but based on comparison, **109b** is proposed to be a related compound due to the presence of key humulene signals (Figure 2.20).

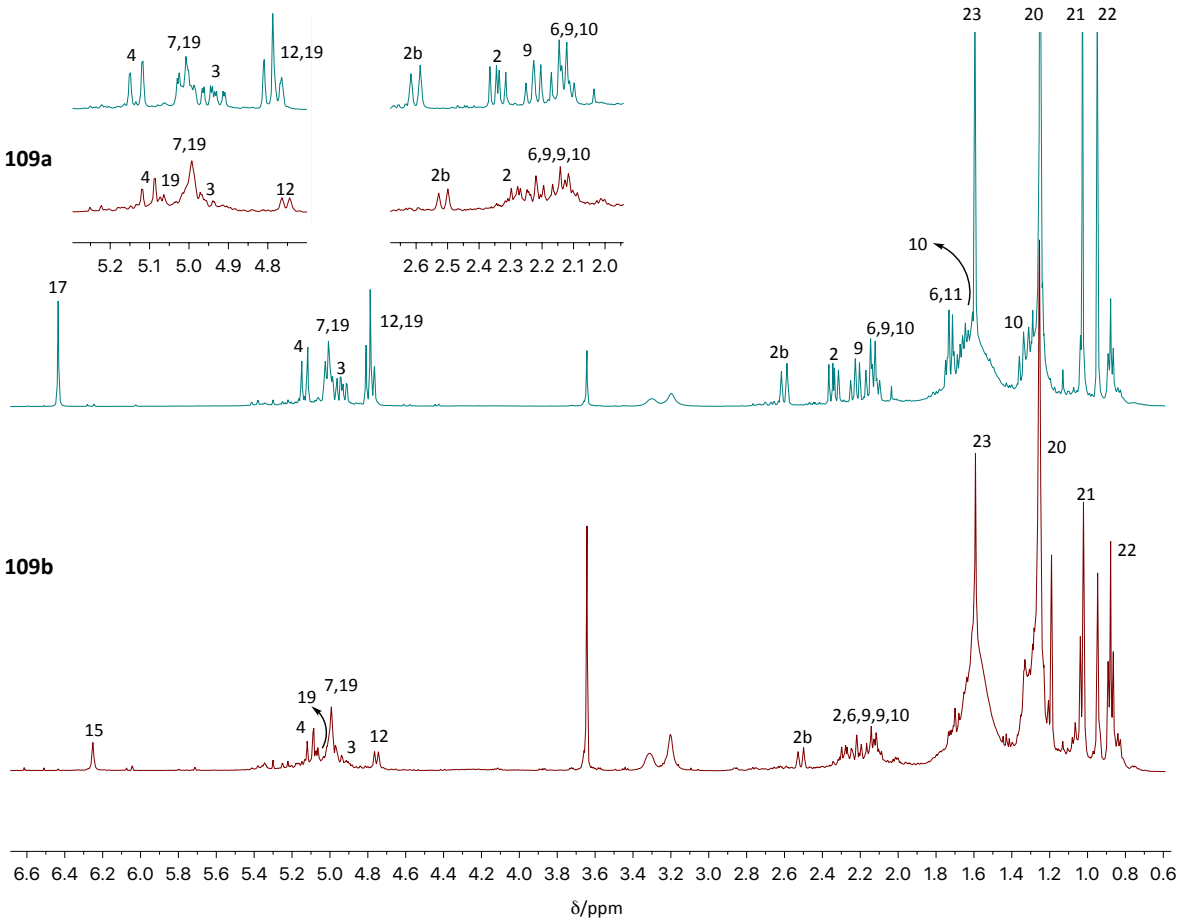


Figure 2.20 ^{1}H NMR of compounds **109a** and **109b**.

Further comparison of the ^{1}H NMR identified changes in chemical shifts of the two protons 17 and 2b. The aromatic proton (17) in **109b** has a chemical shift of δ_H 6.25, instead of δ_H 6.43 in **109a**. This could indicate a differently substituted aromatic ring, similar to previously characterised compounds **107a** and **107b** (Chapter 2.3). In addition, the chemical shift of proton 2b (δ_H 2.5) is shifted ~0.1 ppm upfield in **109b** in comparison to **109a** (δ_H 2.6). A similar, but ~0.2 ppm upfield shift of proton 2b (2b δ_H 2.60) is observed in compound **107b**, where position C-15 is substituted with a proton instead of a hydroxyl function in **107a** (2b δ_H 2.81) (Chapter 2.3). In addition, for the chemical shift of methylene protons at position C-19 a downfield shift can be observed (δ_H 4.8, δ_H 5.01 in **107a** and δ_H 4.98-5.03, δ_H 5.08 in **107b**). Again here a similar shift is observed in NMR data of the structural isomers **107a** and **107b** (δ_H 5.00, δ_H 5.08 in **107a** and δ_H 5.05, δ_H 5.21 in **107b**) (Table 2.7).

Table 2.7 Comparison of selected ^{1}H NMR data.

Position	δ_H/ppm **109a**	δ_H/ppm **10b**	δ_H/ppm **107a**	δ_H/ppm **107b**
2b	2.60	2.50	2.81	2.60
19/20	4.8	4.98-5.03	5.00	5.05
	5.01	5.08	5.21	5.21

HRMS of **109b** suggested a molecular formula of $C_{23}H_{30}O_4$ ([M-H]$^-$ calculated 369.2066, found 369.2066), which would satisfy the proposed structure **109b** and **109c** (Figure 2.21). ^{1}H NMR spectra were simulated for compounds **109b** and **109c**.[124–126] A chemical shift of δ_H 6.24 was predicted for the isomer **109b**, and δ_H 6.14 for **109c**. Comparison of predicted and measured chemical shift of δ_H 6.25 suggested that the isolated compound is more likely the **109b** structural isomer of **109a**.

For further comparison MS^2 data was gathered on the [M-H]$^-$ *m/z* = 369 peak in **109a** and **109b/c**. For both compounds one major fragmentation peak at *m/z* = 165 was observed. This could correspond to the separation of the polyketide and humulene moieties (Figure 2.21). Full NMR characterisation could prove this theory, but **109a** and **109b** are not always produced by *A. strictum* and often only traces are observed.

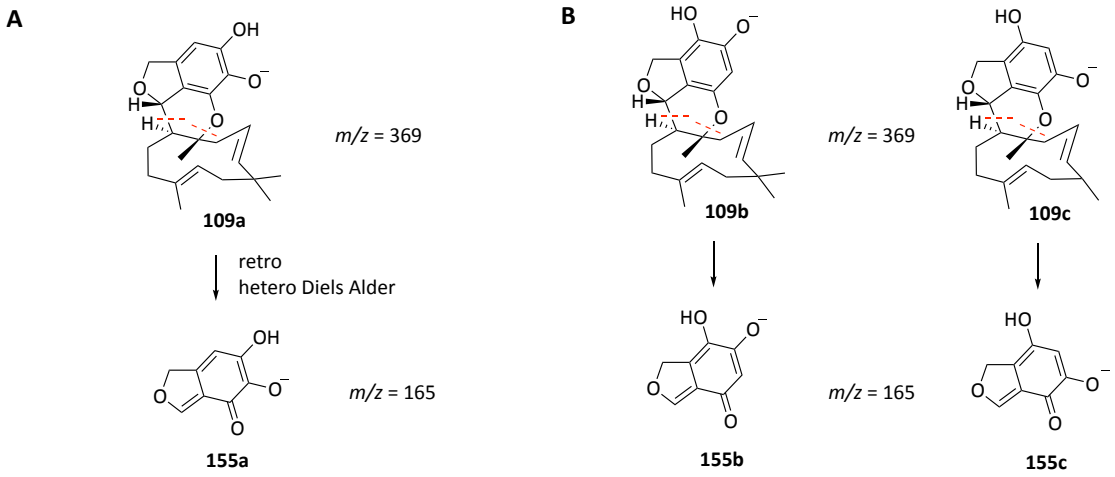


Figure 2.21 Proposed fragmentation ions of **109a** (**A**) and **109b** or **109c** (**B**).

2.6 Characterisation of Meroterpenoid 156

A compound with the nominal mass of 362 (determined through low resolution mass) was sometimes detected by LCMS in *A. strictum* culture alongside xenovulene A **101**. The compound (**156**) eluted at t_R = 8.4 and showed a maximal UV absorption at 211 nm (Figure 2.22). Purification of 1.7 mg from 2 L *A. strictum* wild-type (WT) culture extract enabled full NMR characterisation.

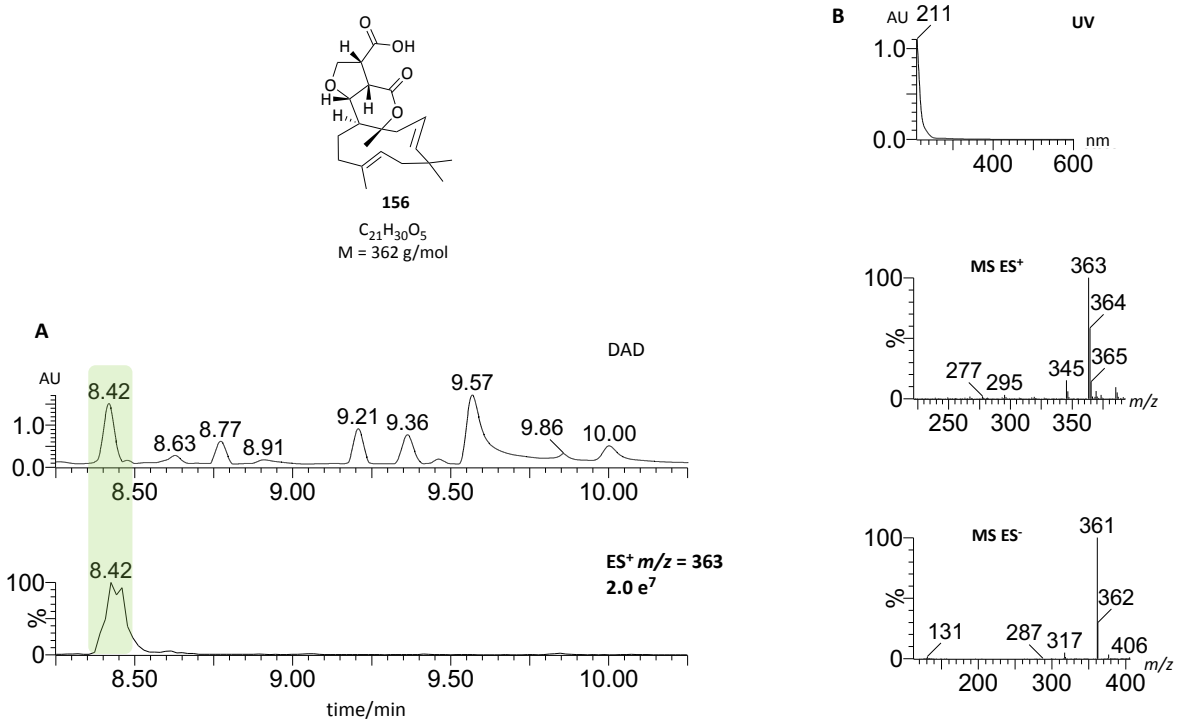


Figure 2.22 Compound **156**. **A** Structure and t_R = 8.4 min of purified compound **156**. **B,** Characterisation of compound **156**.

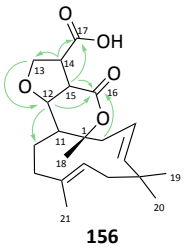
Analysis of ^{1}H and ^{13}C NMR data of **156** identified 21 carbon atoms and 29 protons (Table 2.8). HRMS of **156** confirmed a molecular formula of $C_{21}H_{30}O_5$ ($[M]H^+$ calculated 363.2171, found 363.2179) which suggested one exchangeable proton.

Table 2.8 NMR data for **156** in CD_3OD (500 MHz) referenced CD_3OD.

Position	δ_C/ppm 156	δ_H/ppm (*J* in Hz) 156	HMBC (H to C)
1	88.6	-	-
2	44.1	2.33 (m, 1H)	1, 3, 4, 5
		2.46 (ddd, 1H, 14.7, 2.2, 2.2)	1, 3, 4, 11
3	120.7	5.11 (m, 1H)	1, 2, 5
4	144.8	5.23 (dd, 1H, 15.9, 1.8)	2, 5, 19, 20
5	39.2	-	-
6	42.5	1.79 (br dd, 1H, 13.0, 4.9)	4, 5, 7, 8, 19
		2.25 (m, 1H)	4, 5, 7, 8, 19
7	124.1	5.09 (m, 1H)	5, 21
8	137.7	-	
9	37.9	2.04 (br dd, 1H, 12.8, 7.2)	7, 8, 10, 11, 21
		2.28 (m, 1H)	7, 8, 10, 11, 21
10	31.4	1.42 (m, 1 H)	1, 8, 9, 11, 12, 13
		1.56 (m, 1H)	8, 9, 11, 12
11	43.4	1.96 (dd, 1H, 8.3, 8.3)	1, 2, 9, 10, 12, 13, 18
12	85.8	4.21 (m, 1H)	10, 16
13	70.7	3.73 (m, 1H)	12, 17
		4.23 (m, 1H)	12, 14
14	49.1	3.55 (ddd, 1H, 8.1, 8.1, 5.3)	13, 15, 16, 17
15	47.8	3.74 (m, 1H)	16, 17
16	174.7	-	-
17	175.9	-	-
18	22.1	1.37 (s, 3H)	1, 2
19	24.6	1.08 (s, 3H)	4, 6, 5, 20
20	30.5	1.07 (s, 3H)	4, 6, 5, 19
21	17.3	1.65 (s, 3H)	7, 9, 8

The HSQC spectrum showed four –CH_3, five diastereotopic –CH_2, seven –CH and five quaternary carbon atoms (Table 2.8). Comparison with ^{1}H NMR, ^{1}H,^{1}H COSY and HMBC spectra obtained for **108** and **101** enabled rapid verification and assignment of humulene, tetrahydropyran and tetrahydrofuran proton and carbon signals to the respective positions in **156**. Relative stereochemistry at C-1, C-11 and C-12 was assumed to be identical to xenovulene A **101**, but not further determined.

The two newly identified carbon signals (δ_C 174.7, δ_C 175.9) suggested the presence of carboxylic acid and/or ester functional group. HMBC correlations led to the proposal of structure **156**, which satisfies the HRMS (Figure 2.23). Strong $^3J_{CH}$ correlations observed in HMBC, suggested the relative stereochemistry at C-12, C-15 and C-14 to be identical to xenovulene A **101**. However, for further validation of the stereochemistry a NOE experiment should be acquired.



156

Figure 2.23 Key HMBC correlations of compound **156**.

2.7 Characterisation of the Merotperpenoids 157 and 158

A compound with a nominal mass of 360 (determined through low resolution mass spetrometry) was rarely detected by LCMS of *A. strictum* culture alongside xenovulene A **101**. 2 mg of the compound eluting at t_R = 9.2 were purified by LCMS from 1 L *A. strictum* WT culture extract (Figure 2.24). ^{1}H NMR spectrum showed characteristic humulene signals and thus full spectroscopic information was acquired (^{13}C, ^{1}H,^{1}H COSY, HMBC, HSQC). Further evaluation of NMR data suggested the presence of two co-eluting meroterpenoid compounds (**157** and **158**) (Table 2.9). The major compound **157** was determined to have a chemical formula of $C_{21}H_{30}O_4$ ([M]H^+ $C_{21}H_{30}O_4$ calculated 347.2222, found 347.2220) and for the minor **158** $C_{21}H_{28}O_5$ was found ([M]Na^+ $C_{21}H_{28}O_5Na$ calculated 383.1834, found 383.1827).

The HSQC spectrum of compound **157** showed four –CH_3, six diastereotopic –CH_2, five –CH and six quaternary carbon atoms (Table 2.9). The HSQC spectrum of compound **158** in comparison showed four –CH_3, five diastereotopic –CH_2, five –CH and seven quaternary carbon atoms (Table 2.9). Comparison with ^{1}H NMR, ^{1}H,^{1}H COSY and HMBC spectra obtained for **108** enabled rapid

verification and assignment of humulene (C-1 to C-11; and C-18 to C-21) and bridging proton and carbon signals (C-12, C-16) to the respective positions in **157** and **158**.

In contrast to previous compounds no di- or tetrahydrofuran ring system was identified, but the ether linkage between C-12 and C-13 was still intact. $^{3}J_{CH}$ and $^{2}J_{CH}$ correlation thus strongly proposed the presence of a six membered ring with an α,β-unsaturated carbonyl functional group (δ_C 202.5 in **157** and δ_C 200.7 in **158**) in place of the former tetrahydrofuran. *Vice versa*, the original six membered di- or tetrahydropyran ring could not be detected, but a furan was proposed based on HMBC correlations (Figure 2.25).

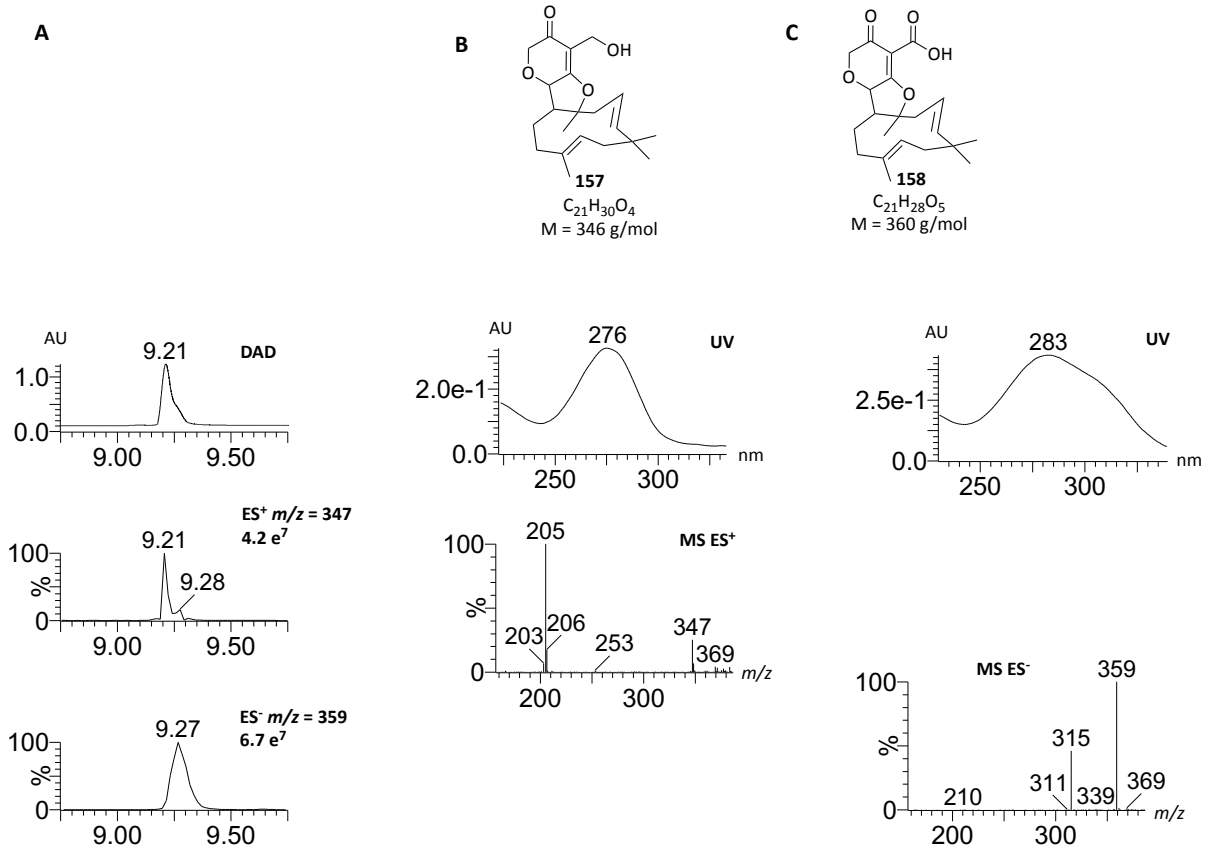


Figure 2.24 Compounds **157** and **158**. **A,** Purified compounds **157** and **158** co-eluting at t_R = 9.2 min. **B,** Characterisation of compound **157**. **C,** Characterisation of compound **158**.

The NMR data for compound **157** showed an additional -CH_2-O group (δ_H 4.39, δ_H 4.60 at δ_C 58.9). A geminal coupling is observed for the two protons attached to C-17, which suggested its attachment next to a quaternary carbon such as C-15. Close analysis of the multiplet of proton δ_H 4.39 suggested a doublet of doublets (dd) with a large ($^{2}J_{HH}$ = 13.8) geminal coupling and a small (J_{HH} = 1.4) long range coupling. However, in the ^{1}H, ^{1}H COSY no correlating proton was identified, which suggested the presence of two conformers rather than a long range coupling. The minor compound **158** lack these methylene protons, but instead a weak carbon signal at δ_C 174.3 was

observed, which proposed the presence of a carboxylic acid. No $^{3}J_{CH}$ correlation was observed for this carbon, which justifies it at position C-17.

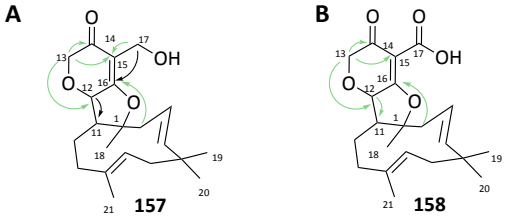


Figure 2.25 Key HMBC correlations **A,** for compound **157**. **B,** Key HMBC correlations for compound **158**.

These two compounds display a novel arrangement of the original xenovulene scaffold, thus stereochemistry at positions C-1, C-11 and C-12 is not indicated. It could be that the xenovulene A **101** stereochemistry at this positions are not retained during putative 5 and 6 ring rearrangements.

Table 2.9 NMR data for **157** and **158** in CD_3OD (500 MHz) referenced to CD_3OD.

Position	δ_C/ppm 157	δ_H/ppm (*J* in Hz) 157	HMBC (H to C) 157	δ_C/ppm 158	δ_H/ppm (*J* in Hz) 158	HMBC (H to C) 158
1	87.9	-	-	87.9	-	-
2	43.2	2.33 (m, 1H)	3, 4, 16, 18	43.4	2.33 (m, 1H)	3, 4, 16, 18
		2.58 (m, 1H)	1, 3, 4, 11, 12, 18		2.58 (m, 1H)	1, 3, 4, 11, 12, 18
3	120.7	5.10 (ddd, 1H, 12.6, 10.8, 2.3)	2, 4, 5	121.0	4.91 (ddd, 1H, 15.8, 10.5,	2, 4, 5
4	144.5	5.20 (dd, 1H, 15.9, 1.9)	3, 5, 6, 19, 20	144.2	2.6)	3, 5, 6, 19, 20
5	39.2	-	-	39.2	5.16 (dd, 1H, 15.8, 1.8)	-
6	42.6	1.75 (br dd, 1H, 12.8, 4.4)	4, 5, 7, 8, 19	42.5	-	4, 5, 7, 8, 19
		2.23 (m, 1H)	5, 7, 8, 19, 20		1.75 (dd, 1H, 13.0, 4.7)	5, 7, 8, 19, 20
7	124.2	5.08 (m, 1H)	5, 21	124.2	2.16 (m, 1H)	5, 21
8	137.7	-		137.6	5.04 (m, 1H)	
9	39.0	2.09 (br dd, 1H, 12.2, 12.2)	7, 8, 10, 11, 21	39.1	-	7, 8, 10, 11, 21
		2.34 (m, 1H)	7, 8, 10, 11, 21		2.14 (m, 1H)	7, 8, 10, 11, 21
10	29.6	1.44 (m, 1 H)	1, 8, 9, 11, 12	30.0	2.28 (m, 1H)	1, 8, 9, 11, 12
		1.57 (m, 1H)	8, 9, 11, 12		1.37 (m, 1 H)	8, 9, 11, 12
11	41.9	1.83 (m, 1H)	1, 2, 9, 10, 12, 15, 18	41.6	1.66 (m, 1H)	1, 2, 9, 10, 12, 18
12	82.7	4.57 (m, 1H)	10, 11	82.8	1.87 (m, 1H)	10, 11
13	74.1	4.01 (m, 1H)	12, 14	74.1	4.57 (m, 1H)	12, 14
		4.06 (m, 1H)	12, 14, 15		3.96 (m, 1H)	12, 14, 15
14	202.5	-	-	200.7	4.03 (m, 1H)	-
15	109.0	-	-	105.0	-	-
16	164.5	-	-	159.8	-	-
17	58.9	4.39 (dd, 1H, 13.8, 1.4)	15, 16	*174.3*	-	-
		4.60 (m, 1H)	15, 16	-	-	-
18	21.8	1.20 (s, 3H)	1, 2, 11	21.4	-	1, 2, 11
19	24.4	1.04 (s, 3H)	4, 5, 6, 20	30.5	1.21 (s, 3H)	4, 5, 6, 20
20	30.5	1.09 (s, 3H)	4, 5, 6, 19	24.6	1.04 (s, 3H)	4, 5, 6, 19
21	17.1	1.61 (s, 3H)	7, 9, 8	17.1	1.08 (s, 3H)	7, 9, 8

2.8 Analysis of the Production of Polyketide Intermediates

In addition to the described meroterpenoids (Chapters 2.1-2.7), the *A. strictum* extracts were analysed for possible polyketide precursors. The proteins encoded in the *aspks1* BGC are predicted homologues of those known from tropolone biosynthesis (Chapter 1.3.1). Thus extracted ion chromatograms were analysed for the presence of key tropolone biosynthesis intermediates such as: 3-methylorcinaldehyde **120**; the enone **121**; stipitaldehyde **123** and stipitafuran **125;** as well as cordytropolone **159** (Figure 2.26A) using $[M]H^+$ or $[M\text{-}H]^-$. As no standards for these compounds

were available, the initial assignments were tentative (Figure 2.26B). However, later gene expression in *A. oryzae* (Chapter 5) enabled the isolation and detailed characterisation of **120**, **121**, **123**, **125** and **159**. Comparison of retention time t_R, mass and UV spectra of *A. strictum* and compounds purified from *A. oryzae* confirmed their identity (Chapter 5.2 for detailed characterisation of these compounds).

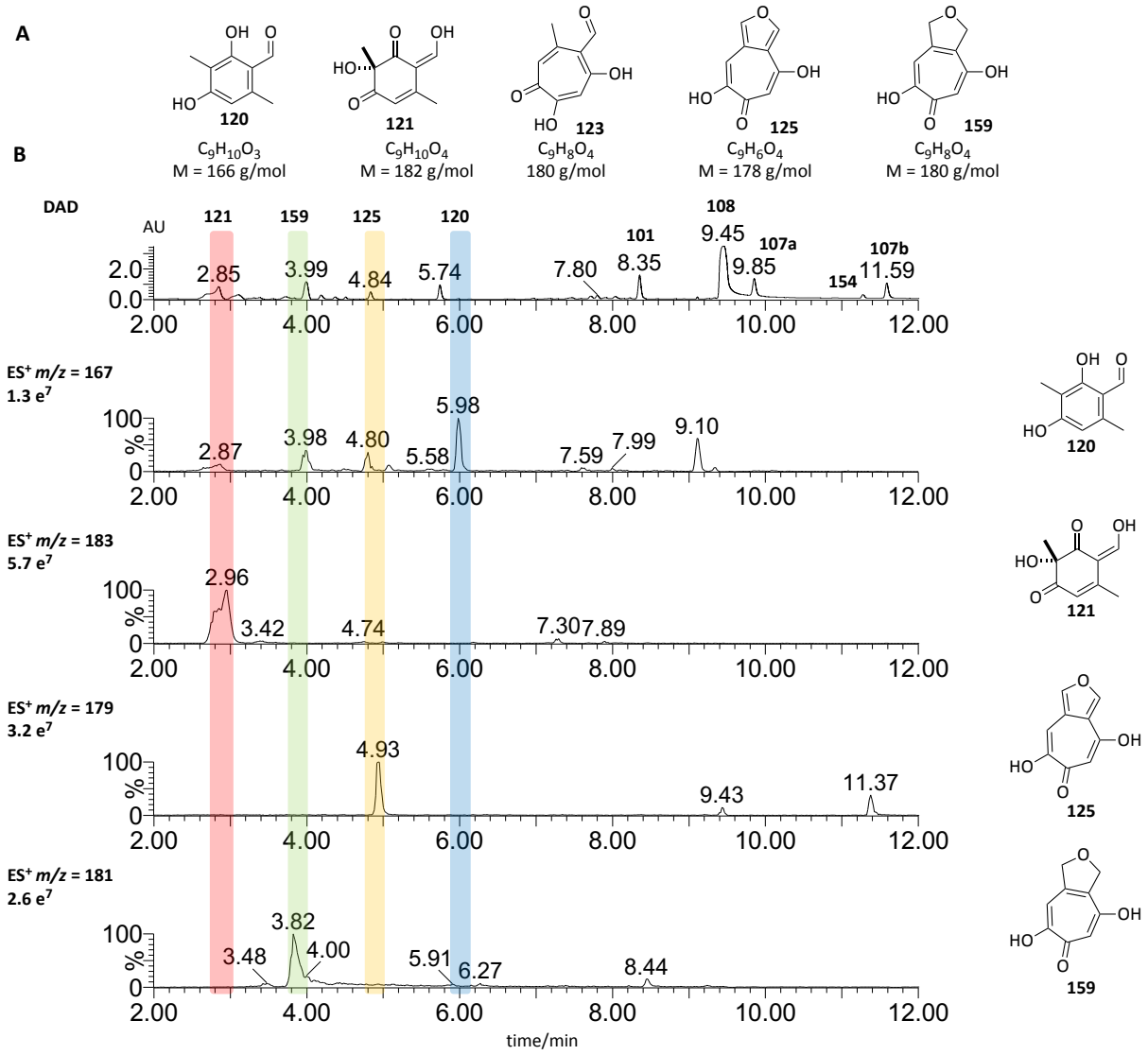


Figure 2.26 Polyketide intermediates in *A. strictum* WT extracts. **A,** Proposed polyketide precursor of xenovulene A **101**. **B,** Extracted ion chromatograms for the analysis of production of the indicated intermediates.

All key intermediates (**120**, **121** and **125**) except stipitaldehyde **123** were observed in *A. strictum* extracts. It is likely that **123** is a reactive intermediate which is prevented from accumulation due to fast processing. A similar observation was made upon heterologous gene expression (**Chapter 5**): accumulation of stiptialdehyde **123** was only observed when downstream processing genes were not co-expressed in *A. oryzae* NSAR. In addition, HRMS analysis of the *A. strictum* compound eluting

at t_R = 4.9 min confirmed its molecular formula to be identical to stipitafuran **125** ([M-H]$^-$ $C_9H_5O_4$ calculated 177.0188, found 177.0188).

2.9 Non-Producing Conditions

For the differential expression analysis (Chapter 3.2) *A. strictum* was grown under xenovulene A producing (ASPM) and non-producing conditions (DPY). In order to establish non producing conditions the ascomycete was sub-cultured in a series of different liquid media (PDB, YMG, ME, DPY, CM, ASSA; Figure 2.27). All media expect DPY showed at least trace production of **101**. However, the usual secondary metabolites can only be observed in *A. strictum* seed medium (ASSA) (**101**, **109a**, **108**, **107b**). In three other cultures a compound eluting at t_R = 9.6 (YMG, DPY, CM) was observed. In PDB medium another a new metabolite is produced eluting at t_R = 9.5 min.

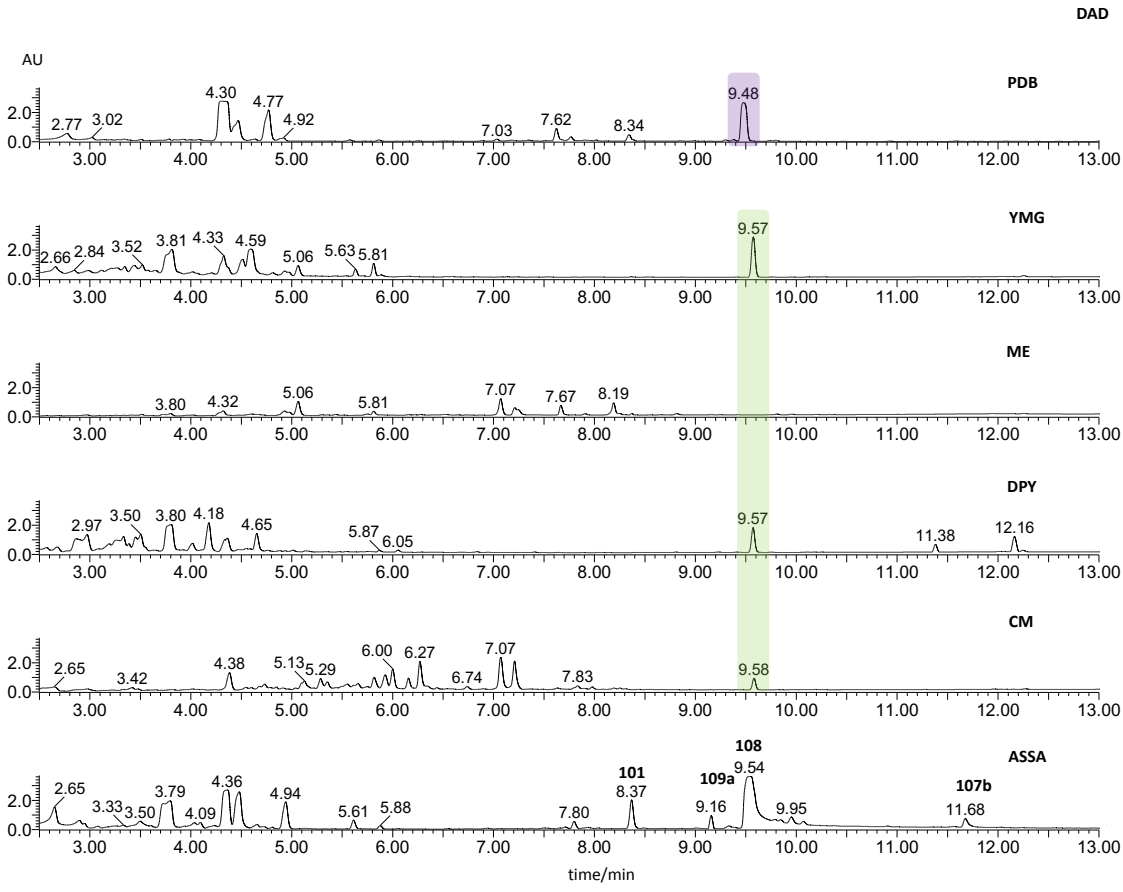


Figure 2.27 DAD chromatogram of extracts obtained from *A. strictum* liquid cultures grown for 6 d in the indicated different media (PDB, YMG, ME, DPY, CM, ASSA).

The *A. strictum* xenovulene producing strain was exploited for **101** production, but has not yet been further analysed for the production of metabolites under different growth conditions. This first screen showed that there is the potential to isolate more putative new metabolites from this

fungus. Analysis of UV and MS of the newly observed compounds, suggested no relation to xenovulenes **101**, **107**-**109**, but no further analysis was carried out.

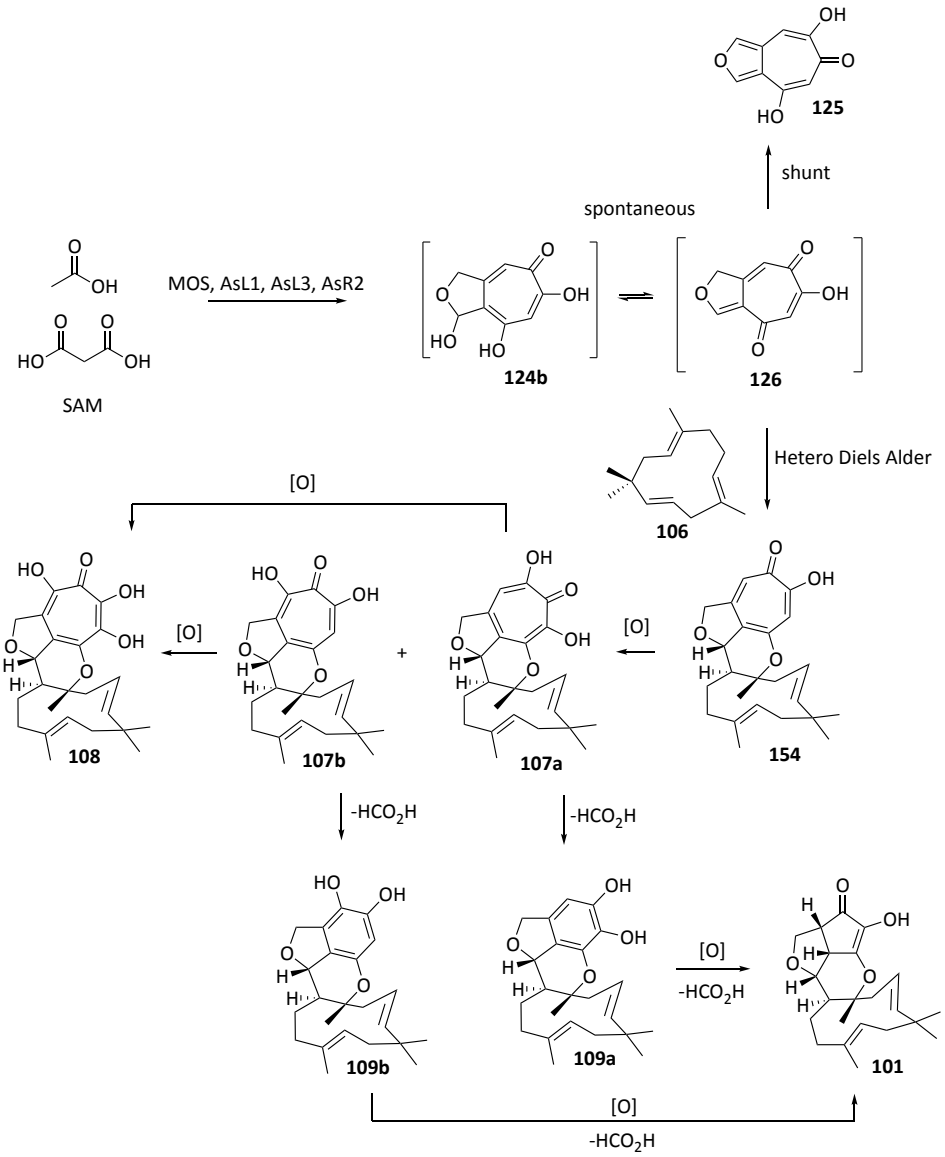
2.10 Discussion and Outlook

Xenovulene A **101** was reported alongside co-metabolites **107**, **108** and **109**. However, only full characterisation for **101** was available. Purification and NMR analysis of compounds **107a**, **108** and **109a** enabled assignment of carbon and proton signals for the first time. In addition, with **107b**, **109b** and **154** three similar but new compounds were isolated and described.

As it is not yet known at which stage the polyketide and terpene moiety in **101** are fused, the identification of the less hydroxylated tropolone **154** gives the first indication that the fusion step could happen to a tropolone moiety (**126**) (Scheme 2.1). No pre-tropolone (*e.g.* benzenoid) meroterpenoids have been observed in *A. strictum*, and neither reported in fungi producing the related compounds eupenifeldin **111** or epolone A **113**.[86,88] Comparing the level of oxidation in **154** and **111** or **113**, the oxidation pattern of stipitaldehyde **123** or stipitafuran **125** can be recognized. Stipitaldehyde **123** should also be considered as final polyketide because in **111** and **113** no further hydroxylation of the 6-methyl group is required to yield meroterpenoids. This, and the identification of **125** in extracts of *A. strictum* suggests either **123** or **125** as the most likely polyketides for humulene **106** attachment.

For both compounds **107a** and **109a** structural isomers (**107b** and **109b**) were identified. This raises the question whether these compounds are true intermediates or shunts on the xenovulene pathway. The proposed mechanism for the two consecutive ring contractions requires an oxidative rearrangement/derformylation cascade (Scheme 1.18A).[84] However, starting with meroterpenoid **154** oxidation followed by ring contraction at position C-15 or C-18 seems equally possible. Later gene expression experiments showed the involvement of two FAD dependent monooxygenases in the ring contraction mechanism (Chapter 5.7.4 for discussion of a more detailed mechanism).

Isolation and structure elucidation of three further meroterpenoids **156**, **157** and **158** from *A. strictum* suggested these to be post-xenovulene products, which most likely arise through an unknown degradation process.



Scheme 2.1 Proposed biosynthesis for **101** in regard to newly isolated compounds.

3 Genome and Transcriptome of *A. strictum*

In preliminary work, a potential partial xenovulene A **101** BCG was identified (*aspks1* BGC, Chapter 1.3.2).[102,127] The BGC encodes the core tropolone forming homologues (*aspks1*/*tspks1*: NR-PKS, *asL1*/*tsL1*: FAD dependent monooxygenase, *asL3*/*tsR5*: NHI dioxygenase, *asR2*/*tsL2*: P450).[97,98,102] In addition, two oxidoreductases are encoded (*asL4*/*asL5)* (Figure 3.1). These could be involved in the two oxidative ring contractions from seven membered tropolones to cyclopentenone in xenovulene A **101** (Scheme 1.17). However, the partial 22 kb BGC does not encode a terpene cyclase which would be expected for humulene **106** biosynthesis. In addition, no genes encoding proteins of unknown function for a putative hetero Diels Alder type fusion of polyketide and terpene (Scheme 1.17) are present. Thus to complete the partial BGC and also to enable a genome wide search for genes that could be involved in **101** biosynthesis, the *A. strictum* genome was sequenced.

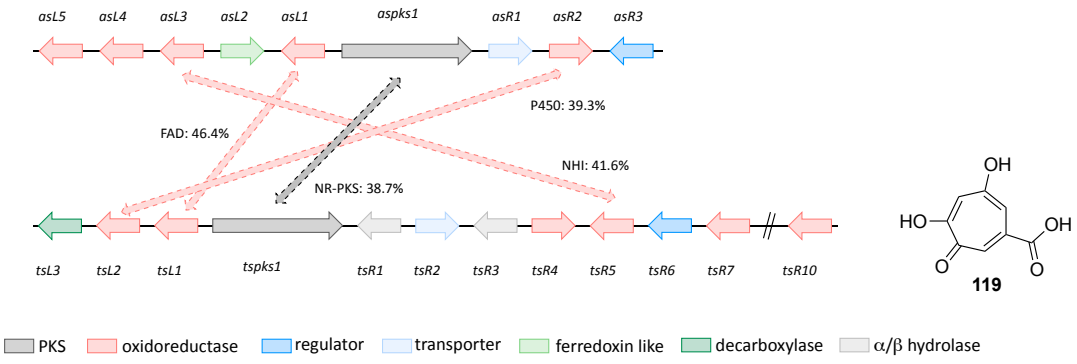


Figure 3.1 Comparison of partial *aspks1* BGC and stiptitatic acid **19** BGC from *T. stipitatus*. Homologous genes are marked by arrows, % indicates protein identity.

For further analysis of gene expression under xenovulene A **101** producing and non-producing conditions the respective transcriptomes (total of all expressed messenger ribonucleic acid, mRNA) were sequenced. Fungal genomes usually encode several BGC but not all are expressed ubiquitously. In fact, BGC activation and secondary metabolite production is regulated by environmental circumstances. For fungi, different secondary metabolites are often observed when grown in different media or static/liquid culture.[1] In the end enzymes such as PKS or NRPS synthesise the secondary metabolite, but these proteins are only present in cells after gene expression. Although attempts to correlate high mRNA levels with high protein concentrations have not been consistent,[128] increased levels of mRNA indicate active gene expression. By comparison of transcriptomes collected under xenovulene A producing and non-producing conditions it should be possible to identify genes associated with xenovulene A **101** production.

3.1 Whole Genome Sequencing and antiSMASH Analysis

Genomic DNA (gDNA) of *A. strictum* was obtained by phenol chloroform extraction of freeze-dried fungal mycelia (taken from 2.5 d old PDB liquid culture) and purified by caesium chloride gradient centrifugation (Figure 3.2A).

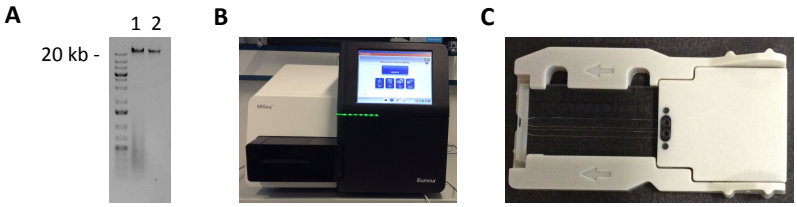


Figure 3.2 gDNA purification and sequencing. **A**, *A. strictum* gDNA 1) after phenol chloroform extraction and 2) after caesium chloride centrifugation. **B**, Illumina MiSeq System at the CeBiTec in Bielefeld and **C**, Illumina flow cell for sample.

The sample was submitted to the Center for Biotechnology (CeBiTec, University of Bielefeld) for Illumina MiSeq paired-end sequencing (Figure 3.2B and C). The raw data were processed by Dr Daniel Wibberg using an in-house software platform based on CASAVA 1.8.2 (Illumina).[129] The sequenced 300 bp reads were assembled with gsAssembler 2.8[129] with default settings. A draft genome of 33.8 Mb in 51 scaffolds with a N50 scaffold length of 1.3 Mb was obtained. Gene prediction and annotation was performed with AUGUSTUS 3.0.3 and the GenDBE platform.[130,131] A total number of 10622 genes were predicted (Table 3.1).

Table 3.1 Statistics of draft genome assembly and gene prediction in *A. strictum* genome.

Assembly statistics		Gene prediction	
Bases in Scaffolds (bp)	33,842,462	# of predicted genes	10622
# of scaffolds	51	avg. gene length (bp)	1788
# of contigs	615	sense strand	5268
GC content (%)	52.28	antisense strand	5394
Avg. scaffold (bp)	663,577	avg. # of exon per gene	2.99
N50 scaffold (bp)	1,366,384	avg. exon length (bp)	535
largest scaffold (bp)	2,269,511	largest exon (bp)	11915
avg. scaf. contig (bp)	90,075	smallest exon (bp)	2
avg. contig (bp)	55,283	avg. intron length (bp)	93
N50 contig (bp)	196,596	largest intron (bp)	1887
Largest contig (bp)	618,099	smallest intron (bp)	29

AntiSMASH is a platform which enables a genome wide search for BGC within a genome sequence. It cross-links the output with either characterised or *in silico* predicted BGC.[20–22] A version optimised for fungal genomes (fungiSMASH) predicted 118 BGC in the *A. strictum* draft genome. Only 39 of these BGC were matched to a core secondary metabolite encoding gene (PKS, NRPS or terpene) and these were further investigated. The twelve predicted PKS containing BGC formed the largest group, but also eight terpene and five NRPS-containing as well as six combined BGC were identified.

Seven of the predicted clusters showed between 14-83% genes similar to previously characterised BGC (Figure 3.3).[97,98,132–137]

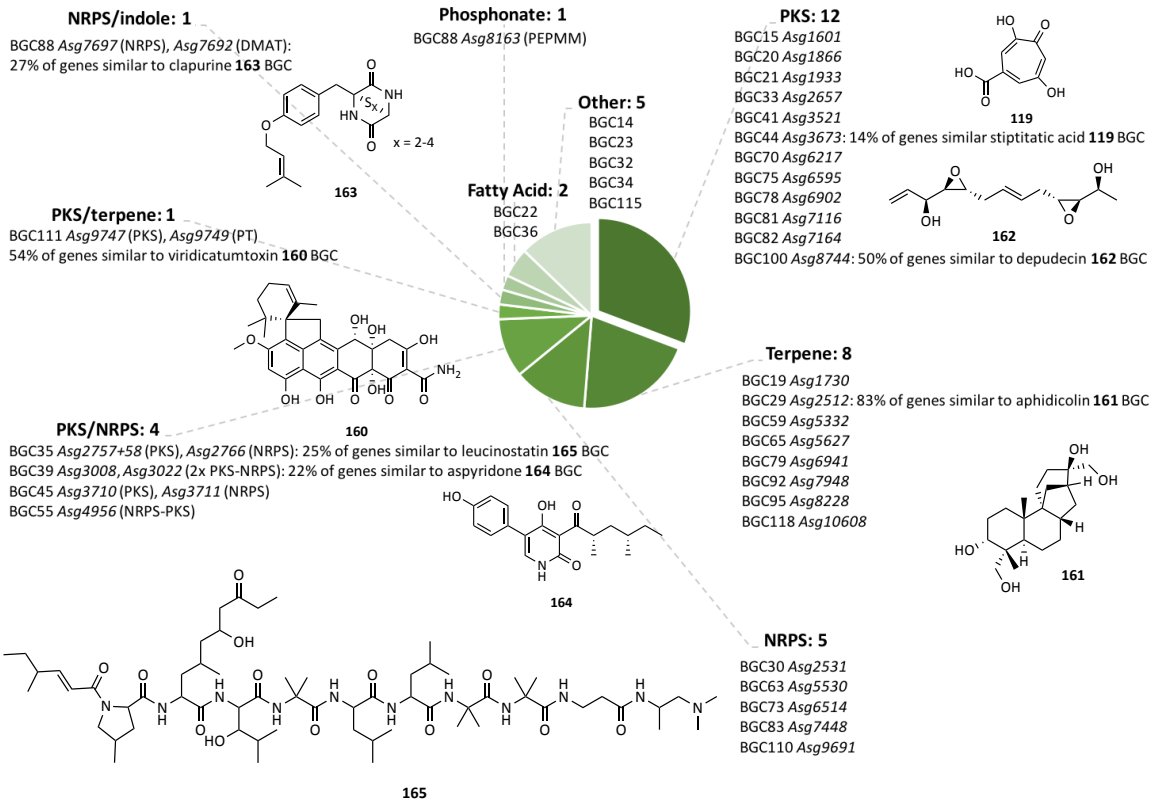


Figure 3.3 Pie chart of BGC predicted by fungiSMASH with a match to PKS, NRPS or terpene gene. The predicted BGC assigned in 'other' clade were revised and all were found to encode an adenylation domain containing protein. BGC numbers are denoted as well as the core gene (encoding PKS, NRPS and terpene) is listed with gene number. Structures of natural products are shown for BGC which show similarity to characterised BGC.

The domain organisations of all predicted PKS genes were further analysed with the conserved domain (CD) analysis tool. Ten HR-PKS, one-PR-PKS, three NR-PKS and one type III PKS were annotated (Figure 3.4). [138–141] Only one (*Asg3673*) of the three encoded NR-PKS was found to have a *C*-Met and R domain, which, according to the biosynthetic hypothesis (Chapter 1.3.2), are required to release the polyketide precursor 3-methylorcinaldeyhde **120**. Sequence alignment of *Asg3673* and *aspks1* (Figure 3.1) confirmed its identity (100%) and gave first *in silico* evidence that BGC44 (referred to as *aspks1* BGC) is the best candidate for xenovulene A **101** biosynthesis.

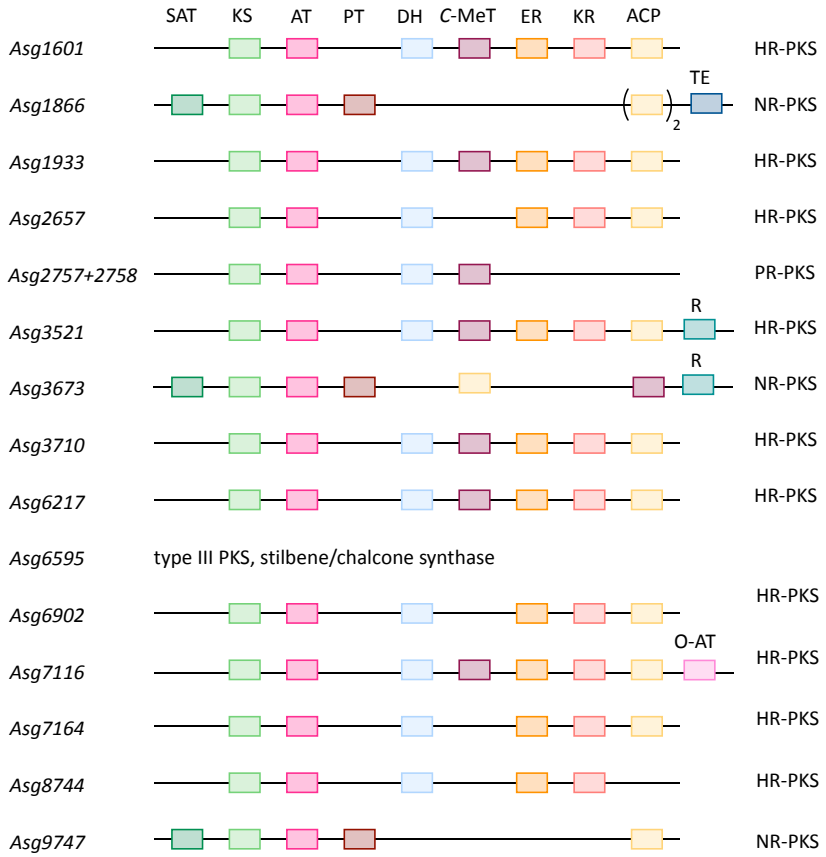


Figure 3.4 Conserved domain analysis of PKS genes in the *A. strictum* genome. SAT = starter unit acyl transferase, KS = ketosynthase, AT = acyltransferase, PT = product template, DH = dehydratase, *C*-Met = *C*-methyl transferase, ER = enoylreductase, KR = ketoreductase, ACP = acyl carrier protein, TE = thiolesterase, R = reductive release, *O*-AT = *O*-acyltransferase.

3.2 Transcriptome Analysis under Producing and Non-Producing Conditions

Xenovulene A **101** production medium (ASPM) is well known from preliminary experiments.[118] To identify **101** non-producing conditions the secondary metabolite production was analysed in different liquid media (Chapter 2.9). DPY medium was found to suppress **101** production but did not inhibit the growth of *A. strictum* (Figure 3.5). Therefore, it was selected as non-producing medium for comparative transcriptome analysis.

Liquid cultures for both conditions (ASPM, DPY) were inoculated with an *A. strictum* single colony and cultured for nine days. Samples of two biological replicates for each condition were taken at day 3, 6 and 9 and combined for total RNA isolation by kit (Zymo Research). The RNA samples were submitted to CeBiTec for cDNA library preparation and Illumina paired-end sequencing.

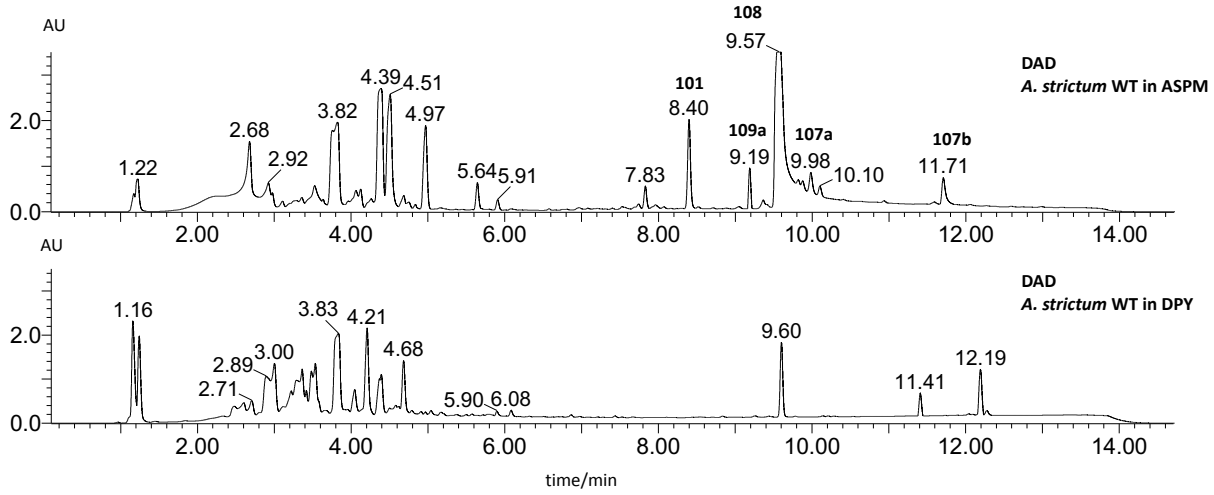


Figure 3.5 DAD chromatograms of *A. strictum* extracts grown under producing (upper) and non-producing (lower) conditions. Peaks at t_R = 8.4, 9.2, 9.6, 10.0 and 11.7 correspond to xenovulenes **101**, **107a**, **107b**, **108** and **109a**.

The raw data were again processed by Dr Daniel Wibberg using an in-house software platform based on CASAVA 1.8.2 (Illumina).[142] The 75 bp reads were mapped on the *A. strictum* draft genome with tophat2,[143] after further processing by the FASTX[144] and trimmomatic[145] toolkits. A total of ~80/120 million reads were sequenced under non-producing/producing conditions. On average 95% of those were mapped to the *A. strictum* draft genome (Table 3.2).

Table 3.2 Read count after quality control and unmapped reads.

conditions	non-producing1	non-producing2	producing1	producing2
read counts after filter and trimming	39.366.754	43.664.326	80.873.578	47.509.594
unmapped reads	2.431.118	2.731.955	4.407.226	1.223.507
%	6.2%	6.3%	5.0%	2.6%

The sequenced transcriptomes for non-producing (A) and producing (B) conditions were used to perform a differential expression sequence (DESeq) analysis by Dr Daniel Wibberg.[146] DESeq is a software package that uses statistical testing and a negative binominal distribution to compare differences in observed read counts. It evaluates whether the observed differences in two biological samples are significant or in the range of natural variation. The mean expression (meanA or meanB) of non-producing (A) and producing (B) conditions is calculated from the normalized read counts of biological replicates (in this work two per condition). The normalized read count is then used to calculate the relative gene expression (fold change, meanB over meanA) (Figure 3.6). Comparing the $\log_2$-fold change of selected genes their up- ($\log_2$-fold change>0) and downregulation ($\log_2$-fold change<0) can be determined. If a $\log_2$-fold change of >2 and an adjusted p-value >0.05 (significant differential expression determined by Benjamini-Höchberg[147] testing with a false discovery rate of 5%) was observed, genes were regarded as differentially expressed. These genes are displayed in

red (Figure 3.6); genes displayed in black were regarded as not differently expressed. Genes with large or infinite log_2-fold change are displayed at the borders of the scatter diagram.

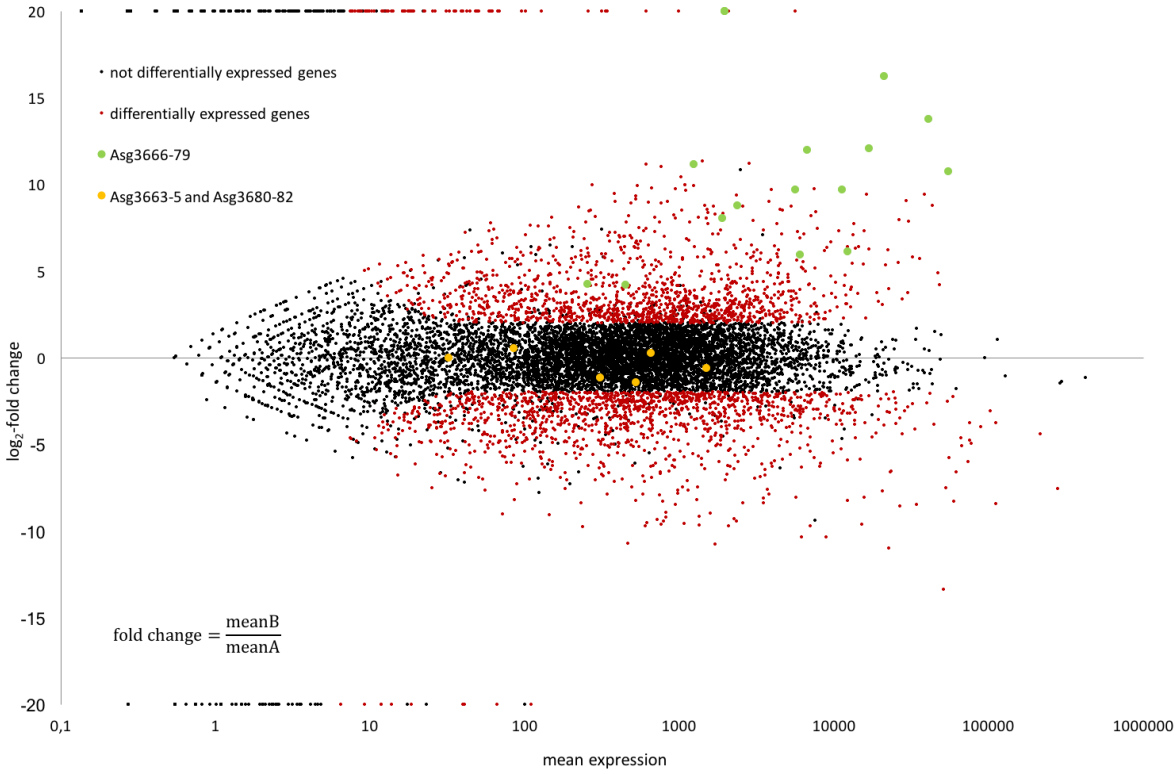


Figure 3.6 Scatter diagram shows mean expression under both conditions (mean(A+B)/2) over log_2-fold change. Every dot correlates to one gene, green dots represent *Asg3666-79* and orange dots represent *Asg3663-5* + *Asg3680-82*.

The expression of the *aspks1* gene cluster (green dots, Figure 3.6) was analysed by comparing the log_2-fold change of genes (*Asg3663-Asg3682*) encoded ~25 kb up- and downstream of *aspks1* (Figure 3.6 and Table 3.3). The log_2-fold change of genes between *Asg3666-3679* was found to be between 6 and 16. For *Asg3674* and *Asg3678*, no reads under non-producing conditions were assembled (meanA = 0), which resulted in an infinite log_2-fold change. This analysis allowed to set cluster boundaries of the *aspks1* BGC to a 16 (*Asg3666-Asg3679*) genes encoding, 48922 bp spanning genomic region. Three investigated genes adjacent to either 5' (*Asg3663-65*) or 3' (*Asg3680-82*) cluster borders (orange dots, Figure 3.6), were found to be not differentially expressed.

Table 3.3 Data calculated with DESeq for *aspks1* BGC and boundaries genes.

gene	mean expression (0.5*meanA+B)	meanA (non-producing)	meanB (producing)	Log_2-fold change
Asg3663	1496,91	1817,58	1176,24	-0,63
Asg3664	307,23	426,31	188,15	-1,18
Asg3665	64,10	600,77	707,42	0,24
Asg3666	5635,53	13,79	11257,28	9,67
Asg3667	16885,47	7,78	33763,17	12,08
Asg3668	11284,07	27,13	22541,01	9,70
Asg3669	1899,60	14,33	3784,87	8,04
Asg3670	6046,69	196,08	11897,30	5,92
Asg3671	2385,73	10,81	4760,64	8,78
Asg3672	1231,79	1,09	2462,48	11,14
Asg3673	12300,95	345,26	24256,64	6,13
Asg3674	1960,75	0,00	3921,51	∞
Asg3675	6674,83	3,28	13346,38	11,99
Asg3676	54766,45	63,89	109469,01	10,74
Asg3677	21058,44	0,55	42116,34	16,24
Asg3678	1962,28	0,00	3924,56	∞
Asg3679	40698,49	5,85	81391,12	13,76
Asg3680	450,85	46,27	855,43	4,21
Asg3681	254,08	25,19	482,98	4,26
Asg3682	32,24	32,40	32,09	-0,01

3.3 *aspks1* BGC Analysis of *A. strictum*

Analysis of the *A. strictum* genome and comparison of the transcriptome under producing and non-producing conditions strongly suggested that the *aspks1* BGC is involved in xenovulene A **101** biosynthesis. Predicted ORFs were further analysed by BLASTp[103] and annotation as well as intron prediction was manually adjusted based on the transcriptomic data (Appendix Table 9.1).

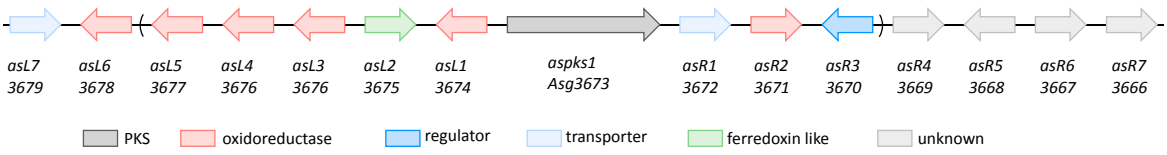
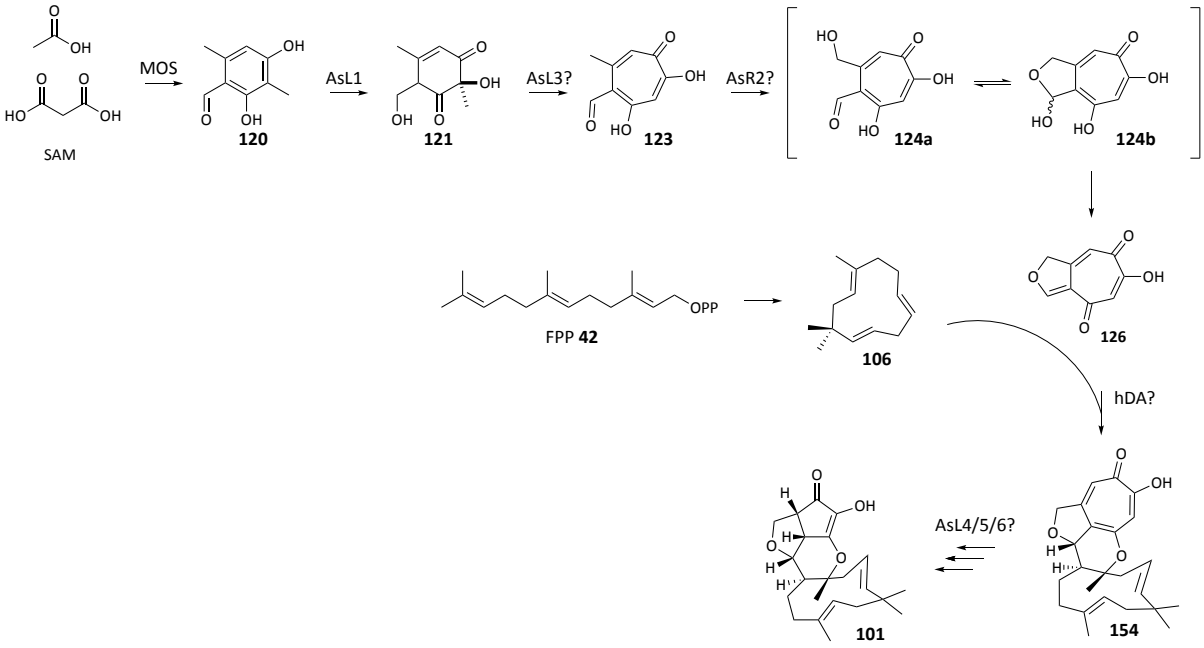


Figure 3.7 The ~49 kb *aspks1* BGC. Brackets indicate partial BGC as published 2007.[102] Not to scale.

Based on the protein homology to tropolone biosynthetic enzymes from stipitatic acid **119** biosynthesis in *T. stipitatus*, *aspks1*, *asL1*, *asL3* and *aR2* are assumed to encode homologues in *A. strictum*.[97] It is proposed that three additional transformations have to take place to form **101** from a tropolone or phenolic polyketide core:

- cyclisation of FPP to humulene **106**,
- fusion of **106** with the polyketide moiety
- and two ring contraction steps to convert the seven membered tropolone to the cyclopentenone in **101**.

The identified ORFs also encode putative transporters (*asL7*, *asR1*), one transcription factor (*asR3*), two FAD dependent oxidoreductases (*asL4*, *asL6*), and one SDR (*asL5*). In addition, a ferredoxin like protein is encoded by *asL2* and several proteins of unknown function are encoded by *asR4*, *asR5*, *asR6* and *asR7* (Table 3.4). A structural prediction of AsR5 using Phyre2[148] revealed its homology to a Ca-dependent phosphotriesterase.



Scheme 3.1 Proposed pathway for **101** formation.

Table 3.4 *Aspks1* gene cluster analysis. Transcriptome analysis confirmed exon and intron positions. Proposed function and CD, BLASTp[103] or Phyre2 analysis. Grey font colour indicates weak alignment and Phyre2 prediction.

#	gene, protein	bp, aa	Putative function	BLASTp[a], Phyre2[b] or CD[c]
Asg3679	*asL7*, AsL7	4350, 1449	transporter	ABC transporter[a]
Asg3678	*asL6*, AsL6	1272, 423	oxidoreductase	FAD binding, 2-polyprenyl 6-methoxyphenol hydroxylase[a,c]
Asg3677	*asL5*, AsL5	750, 249	oxidoreductase	NAD(P)-dependent short-chain dehydrogenase[a,c]
Asg3676	*asL4*, AsL4	1293, 430	oxidoreductase	FAD binding, 2-polyprenyl 6-methoxyphenol hydroxylase[a,c]
Asg3676	*asL3*, AsL3	1023, 340	oxidoreductase	Non-heme Fe^{II} dependent dioxygenase[a,c]
Asg3675	*asL2*, AsL2	348, 115	hypothetical	ferredoxin like, $\alpha+\beta$ barrel[b]
Asg3674	*asL1*, AsL1	1443, 480	oxidoreductase	FAD binding, salicylate monooxygenase[a,c]
Asg3673	*aspks1*, MOS	8190, 2729	NR-PKS	3-methylorcinaldehyde synthase[a,c]
Asg3672	*asR1*, AsR1	1308, 435	transporter	MFS transporter[a,c]
Asg3671	*asR2*, AsR2	1551, 516	oxidoreductase	Cytochrome P450[a,c]
Asg3670	*asR3*, AsR3	2565, 854	regulation	GAL4-like Zn(II)2/Cys6 binuclear cluster DNA-binding [a,c]
Asg3669	*asR4*, AsR4	1899, 632	hypothetical	Zn(II)2/Cys6 binuclear cluster DNA-binding domain[2]
Asg3668	*asR5*, AsR5	1206, 401	hypothetical	Ca-dependent phosphotriesterase (six bladed propeller)[b]
Asg3667	*asR6*, AsR6	1293, 430	hypothetical	unknown
Asg3666	*asR7*, AsR7	2469, 822	hypothetical	Serine protease like[b]

A gene encoding a terpene cyclase was expected to be clustered with *aspks1* to catalyse the cyclisation of FPP **42**. However, analysis of all encoded proteins did not identify a protein with significant homology to known terpene cyclases. The intermolecular fusion of polyketide and terpene precursor is proposed to be catalysed by a hetero Diels Alderase type of enzyme. A few examples of intramolecular Diels Alderase chemistry have been characterised (Chapter 1.3.4), but proteins show little structural homology and thus BLASTp[103] results ore often inconclusive.[116] However, within the *aspks1* BGC two genes encoding proteins of unknown function (*asR5*, *asR6*) and two additional genes (*asR4*, *asR7*) with weak protein homology alignments were identified. These were regarded as possible candidates to be involved in a hetero Diels Alder type reaction.

The ring contraction is proposed to happen via an oxidative rearrangement/deformylation cascade.[84] Thus it is proposed to require enzymes that oxidise aromatic ring structures, such as tropolones or phenols. Two genes (*asL4*, *asL6*) were found to encode two putative FAD dependent monooxygenases similar to the 2-polyprenyl 6-methoxyphenol hydroxylase UbiH (COG0654), which are possible candidates for this kind of chemistry.

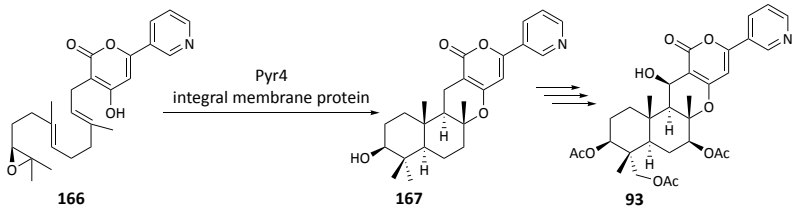
Additionally, a gene (*asL5*) encoding a putative short-chain dehydrogenase was predicted. These kinds of enzymes are known to catalyse epimerisations, decarboxylations and carbonyl-alcohol oxidoreductions (cd05233).[138–141] However, the proposed biosynthesis of **101** does not include a step in which an SDR is obviously engaged. Due to its co-expression it is likely to be involved in either **101** biosynthesis or in a further transformation step of **101**. However, no intermediate proposing such a reaction of **101** has yet been observed in *A. strictum* extracts.

The *in silico* analysis of the *aspks1* BGC showed that in addition to the core tropolone forming enzymes (MOS, AsL1, AsL3 and AsR2), promising candidates for the ring contraction (AsL4, AsL6) as well as several hypotheticals (AsR4, AsR5, AsR6, AsR7) which could be involved in the hetero Diels Alder chemistry are encoded.

The absence of a terpene cyclase encoding gene in the *aspks1* BGC led to the hypothesis of a split BGC as it was observed in the biosynthesis of the meroterpenoid austinol **84** in *A. nidulans*. In this example the PKS encoding gene is located with four other genes in one BGC and the prenyltransferase encoding gene plus ten additional genes are encoded on a different chromosome.[149] In order to identify the terpene cyclase involved in **101** biosynthesis the whole genome was screened for homologues and the transcriptome data was analysed for co-expression under **101** producing conditions (Chapter 3.3.1).

3.3.1 Terpene Cyclase Expression Analysis

Cyclisation of FPP precursors in fungi is performed by two types of terpene cyclases. 'Traditional' class I ionization dependent enzymes with a characteristic DDxxD motif for coordination of metal ions[57,150,151] (Mg^{2+}). In addition, a class of integral membrane proteins, which were first described in the biosynthesis of pyripyropene **93,** are also known.[77] For the latter, a protonation-dependent cyclisation mode is proposed, but has not yet been fully elucidated (Scheme 3.2).[77] Classical terpene cyclases use free FPP **42** as the substrate. The integral membrane proteins, however, cyclase a sesquiterpene epoxide **166** which is pre-attached to a cyclic polyketide precursor.[80]



Scheme 3.2 Cyclisation of **166** by the integral membrane protein PyrG4 in pyripyropene A **93** biosynthesis.

The cyclase involved in **93** biosynthesis, Pyr4,[77] was used as a BLASTp[103] template and one putative homologue encoded by *Asg9222* was identified in the *A. strictum* genome. Three further putative class I terpene cyclases (Asg5463, Asg5501, Asg10609) were identified in an additional BLASTp[103] homology search using basidiomycetes FPP **42** cyclases (Stehi164702, Stehi173029, Stehi125180)[56] as templates. These four BLASTp-identified together with nine other terpene synthases predicted by fungiSMASH were further analysed for their absolute expression and co-expression with the *aspks1* BGC by the DESeq tool (orange dots, Figure 3.8).[146]

In comparison to the *aspks1* BGC (green dots, Figure 3.8) the terpene cyclases (orange dots, Figure 3.8) show a low mean expression and no differential gene expression under xenovulene A producing and non-producing conditions is observed. For only one of the investigated genes (*Asg7948*) the calculated mean expression (1630) is comparable to the lower range of the *aspks1* BGC, where the lowest mean expression was observed for *Asg3672* (1231) (Tables 3.3 and 3.5). However, a BLASTp[103] homology search suggests it to be a squalene synthase rather than an FPP processing terpene cyclase.

The analysis of all putative *A. strictum* terpene cyclases by DESeq analysis showed that none of the 13 genes is co-regulated under **101** producing conditions.[146] However, some genes are constitutively expressed (*Asg1730*, *Asg7948*, *Asg10608*) but, compared to the *aspks1* BGC, at low levels. Thus with the DESeq analysis it was not possible to propose a candidate terpene cyclase which could be involved in **101** biosynthesis.

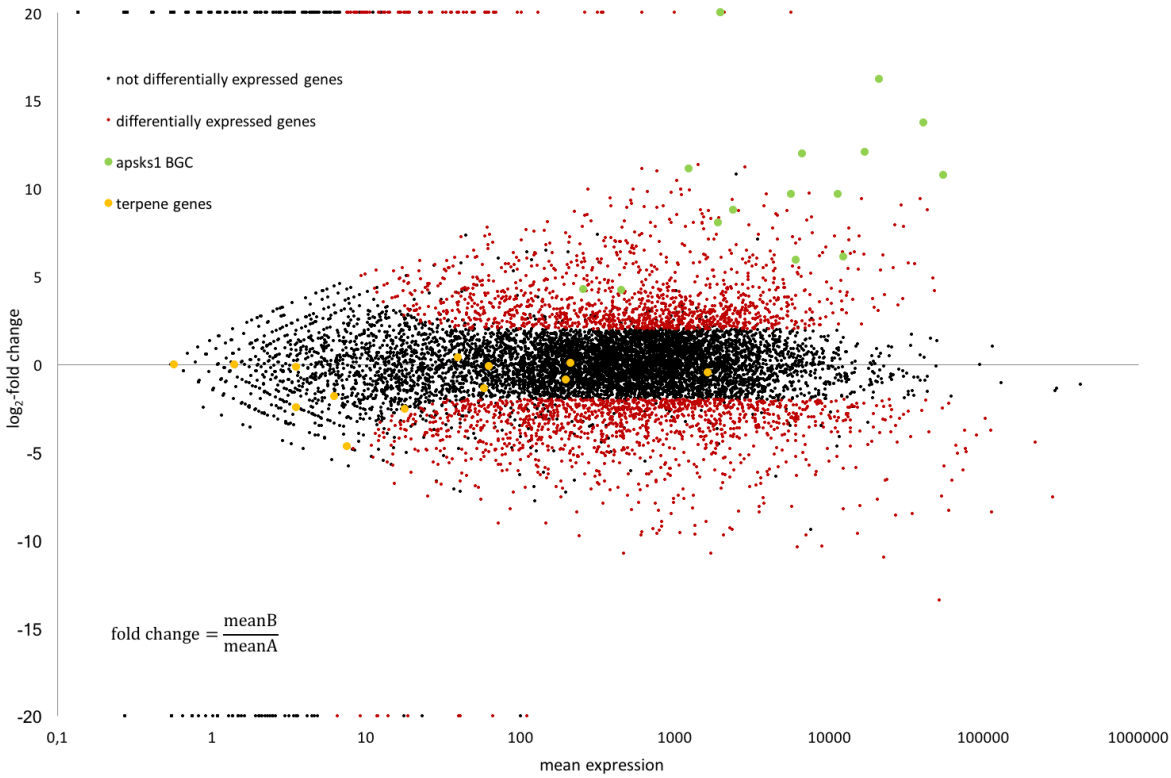


Figure 3.8 Scatter diagram shows mean expression under both conditions over log_2-fold change. Every dot correlates to one gene, green dots represent the *aspks1* BGC and orange dots represent terpene genes.

Table 3.5 Data calculated with DESeq for terpene genes. Grey font colour indicates BLASTp[103] alignment scores between 80-200.

Gene	BLASTp	mean expression	meanA	meanB	log_2-fold change
Asg1730	GGPP synthase	211,35	206,43	216,27	0,07
Asg2512	GGPP synthase	1,39	0,00	2,79	0,00
Asg5501	Terpenoid synthase	3,51	5,94	1,08	-2,45
Asg5332	Aristolochene like cyclase	7,53	14,48	0,57	-4,67
Asg5463	Terpene synthase metal binding domain	6,20	9,66	2,74	-1,82
Asg5627	Pentalene like cyclase	17,86	30,44	5,28	-2,53
Asg6941	Persilphiperfolan-8beta-ol like synthase	57,82	83,59	32,04	-1,38
Asg7948	Squalene synthase	1630,66	1897,05	1364,27	-0,48
Asg8228	Lycopene cyclase/phytoene synthase	39,28	33,86	44,69	0,40
Asg9222	Integral membrane protein	39,28	64,47	59,56	-0,11
Asg9750	FPP synthase	0,57	0,00	1,14	0,00
Asg10608	Pentalene like cyclase	196,88	254,42	139,34	-0,87
Asg10609	Terpenoid synthase	3,50	3,67	3,33	-0,14

3.3.2 Homologous BGC in *Aspergillus thermomutatus*

During the time of this work the genome of *Aspergillus thermomutatus*, which was isolated from a human nasal sinus abscess, became available (NCBI: GCA_002237265.1). No secondary metabolites are reported from this isolated strain. The species is also known as *Neosartorya pseudofischeri* and

a strain isolated from a marine source is reported to produce several metabolites such as pyripyropene A **93**, several diketopiperazines and the alkaloid neosartin C.[152] However, no xenovulene type meroterpenoids have been reported.

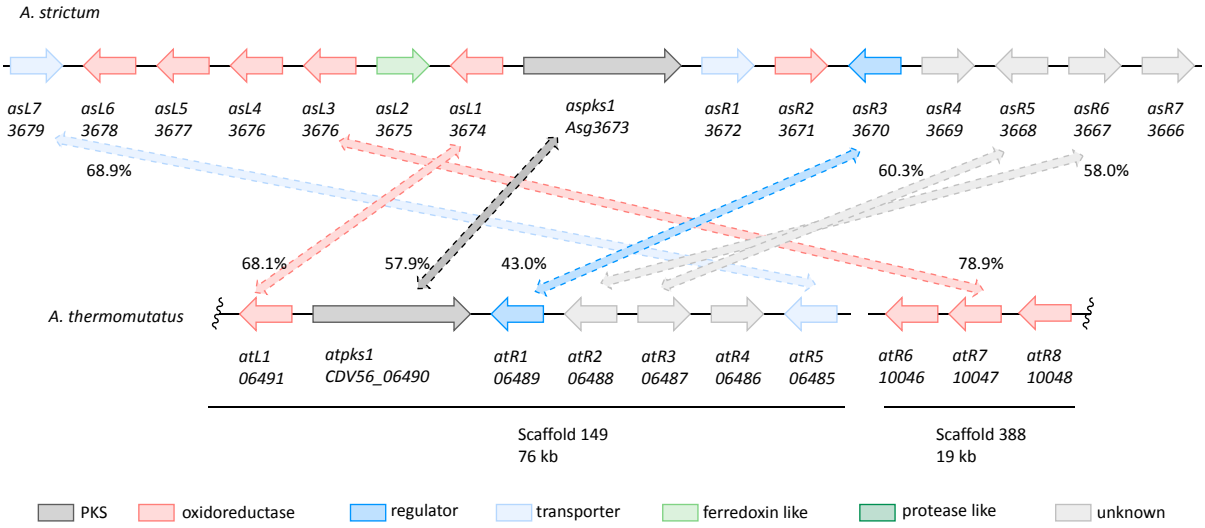


Figure 3.9 Cluster comparison of *aspks1* BGC and one unknown homologous BGC from *A. thermomutatus*. Protein homologies were calculated with EMBOSS Needle.[153–155]

BLASTp[103] homology searches with the hypothetical proteins AsR5 and AsR6 of the *aspks1* BGC as templates were carried out. The search identified two putative homologous proteins (60.3% and 58.0% amino acid identity) encoded by *CDV56_06487* and *CDV56_06488* which are located adjacent to each other in the genome of *A. thermomutatus* (Figure 3.9). Further analysis of this genomic locus also revealed three genes encoding a NR-PKS, an FAD dependent monooxygenase and a transcription factor which showed 57.9%, 68.1% and 43% amino acid identity to *aspks1* BGC proteins (Figure 3.9). This homologous BGC was found to be on a scaffold with a total length of 76 kb. Analysis of further ORF on that scaffold only identified hypothetical (encoded by *atL2-atL6*, *atL12*) and ion transport involved proteins (encoded by *atL7-atL11*) (Table 3.6). However, the whole genome of *A. thermomutatus* was analysed for further homologues of *aspks1* BGC proteins. On a second scaffold, with the total length of 19 kb, an AsL3 homologous protein (encoded by *CDV56_10047*) with 78.9% amino acid identity was predicted. The genes on its left (*CDV56_10046*) and right (*CDV56_10049*) are predicted to encode a cytochrome P450 and an SDR (Table 3.6). Comparison to *aspks1* BGC proteins (AsR2: P450, AsL5: SDR), however, showed only minor amino acid identities of 12.9% and 23.3% (Table 3.6). Three more ORF are predicted on the scaffold which encode hypothetical proteins, without homologies to further proteins encoded by the *aspks1* BGC.

In total six proteins of the *aspks1* BGC were identified to have a homologue in *A. thermomutatus*. The fact that one gene (*CDV56_10047*) is located on a different, very short scaffold indicates that

the BGC is most likely not sequenced completely. It is suggested that both scaffolds should overlap, but it is likely that due to missing sequencing data the assembly is incomplete at that genomic locus.

However, with homologues of the core tropolone-forming enzymes (NR-PKS, FAD dependent monooxygenase, NHI dioxygenase) the *A. thermomutatus* BGC encodes all necessary enzymes for tropolone (stipitaldehyde **123**) production (Scheme 1.16). In addition, the two hypothetical proteins AsR5 and AsR6 are highly conserved in both BGC and were thus analysed in more detail.

Table 3.6 Analysis of *A. thermomutatus* scaffolds 149 (79 kb) and 388 (19 kb).

#	gene	Putative function	*aspks1* BGC	BLASTp[a], Phyre2[b] or CD[c]
CDV56_06502	*atL12*	hypothetical	-	hypothetical[a]
CDV56_06501	*atL11*	transporter	-	vacuolar calcium ion transporter[a]
CDV56_06500	*atL10*	transporter	-	calcium/proton exchanger protein[a]
CDV56_06499	*atL9*	transporter	-	vacuolar calcium ion transporter[a]
CDV56_06498	*atL8*	transporter	-	sodium/calcium exchanger protein[a]
CDV56_06497	*atL7*	resistance	-	cadmium resistance transporter
CDV56_06496	*atL6*	hypothetical	-	hypothetical[a]
CDV56_06495	*atL5*	hypothetical	-	hypothetical[a]
CDV56_06494	*atL4*	hypothetical	-	hypothetical[a]
CDV56_06493	*atL3*	hypothetical	-	hypothetical[a]
CDV56_06492	*atL2*	hypothetical	-	hypothetical[a]
CDV56_06491	*atL1*	oxidoreductase	AsL1	FAD binding, salicylate monooxygenase[a,c]
CDV56_06490	*atpks1*	NR-PKS	MOS	3-methylorcinaldehyde synthase[a,c]
CDV56_06489	*atR1*	regulation	AsR3	GAL4-like Zn(II)2/Cys6 binuclear cluster DNA-binding[a,c]
CDV56_06488	*atR2*	hypothetical	AsR6	hypothetical[a]
CDV56_06487	*atR3*	hypothetical	AsR5	hypothetical[a]
CDV56_06486	*atR4*	hypothetical	-	hypothetical[a]
CDV56_06485	*atR5*	transporter	AsL7	ABC transcporter[a]
CDV56_10046	*atR6*	oxidoreductase	-	cytochrome P450 monooxygenase[a,c]
CDV56_10047	*atR7*	oxidoreductase	AsL3	Non-heme Fe^{II} dependent dioxygenase[a,c]
CDV56_10048	*atR8*	oxidoreductase	-	short-chain dehydrogenase[a,c]
CDV56_10049	*atR9*	hypothetical	-	hypothetical[a]
CDV56_10050	*atR10*	hypothetical	-	Ribonuclease like[a,c]
CDV56_10051	*atR11*	hypothetical	-	hypothetical[a]

AsR5 is a protein of 401 aa which shows very few good (alignment score >200, 3 hits) BLASTp[103] results, but was found to be a structural homologue of a Ca^{2+} dependent phosphotriesterase with high confidence by Phyre2.[148] A tertiary structure model was built with SWISS-MODEL[155-157] for AsR5 and its *A. thermomutatus* homologue (AtR3). This predicted a six-bladed propeller tertiary structure for both proteins (only shown for AsR5) (Figure 3.10).

AsR6 is a protein of 430 aa which also shows very few good (alignment score >200, 11 hits) BLASTp[103] results. Apart from the *A. thermomutatus* homologue AtR3 no other protein was found to be co-located with an AsR5 homologue. No conserved domains or structural homologues were identified using standard tools for AsR6 or its *A. thermomutatus* homologue AtR3 (CD,[138–140]

Phyre2[148], SWISS-MODEL[156–158]). Later in this work the *in vitro* activity of AsR6 was investigated and the protein was shown to cyclise FPP **42** to humulene **106** in the presence of Mg^{2+} (Chapter 5.5). Typically, class I terpene cyclases possess two conserved aspartate rich motifs to coordinate Mg^{2+}: DDxxD/E and (N)DxxT/SxxxD/E. The alignment of the plant humulene **106** synthase from *Z. zerumbet* (AB247331.1), the *F. fujikorii* koraiol **136** synthase Ffsc4 (CCT72694.1), *C. acutatum* CaTPS (KP398851.1) and aristolochene **38** synthase from *A. terreus* (AAF13264.1) shows the conserved Mg^{2+} binding sites, but no other regions of significant homology (Figure 3.11).

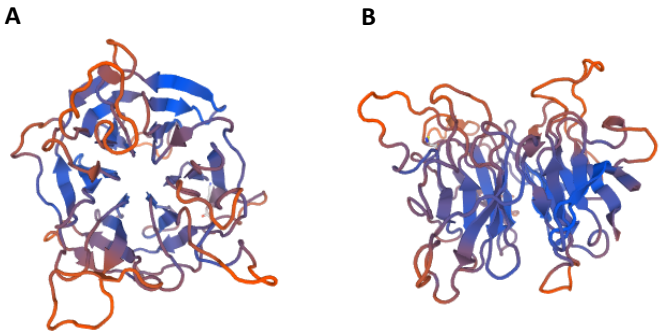


Figure 3.10 Model of AsR5 computed with SWISS-MODEL and template 2p4o.1.A (hypothetical protein of *Nostoc punctiforme* PCC 73102, sequence homology 21.22% to AsR5)**.** QMEAN of -3.79 indicates reliable quality of the model (models with QMEAN <-4 are regarded as low quality). **A**, Bottom view of predicted six-bladed propeller. **B**, Side view of predicted six-bladed propeller. Red low – blue high confidence.

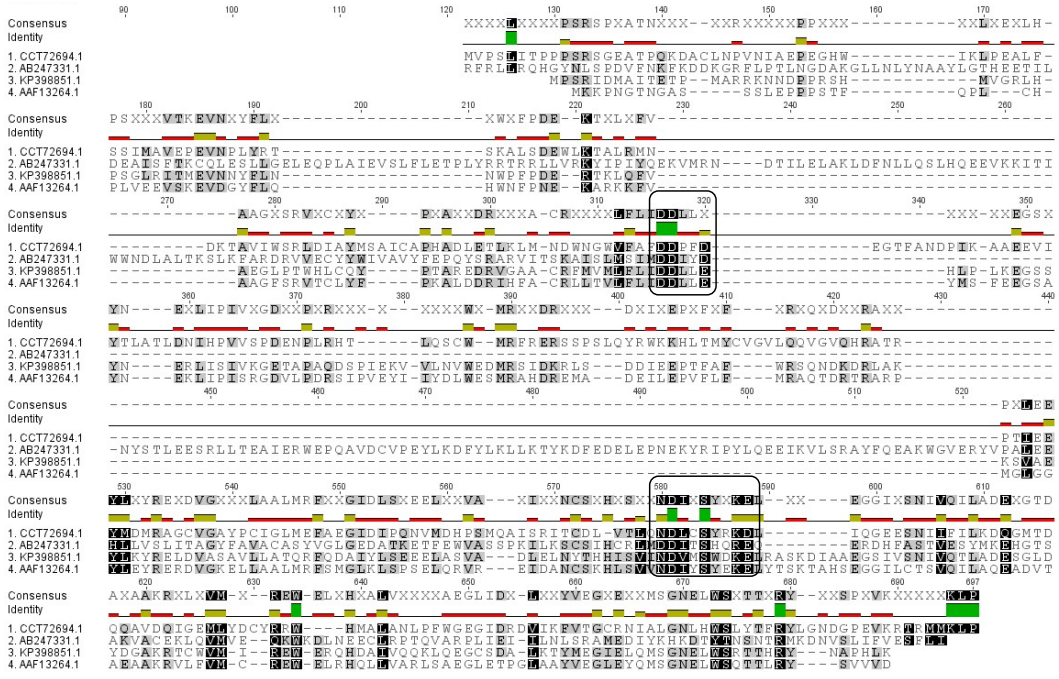


Figure 3.11 Alignment of amino acid sequences of class I terpene cyclases to show conserved motifs DDxxD/E and (N)DxxT/SxxxD/E. Humulene synthase from *Z. zerumbet* (AB247331.1), *F. fujikorii* koraiol **136** synthase Ffsc4 (CCT72694.1), *C. acutatum* CaTPS (KP398851.1) and aristolochene **38** synthase from *A. terreus* (AAF13264.1). Geneious 7.1.9 was used for the alignment.

A sequence alignment of AsR6 and selected homologues could not identify either of these conserved motifs (Figure 3.12). Bioinformatic analysis thus showed that although AsR6 was shown to produce humulene **106** from FPP **42** *in vitro* (Chapter 5.5), it has no conserved domain or sequence homologies to any other known terpene cyclase.

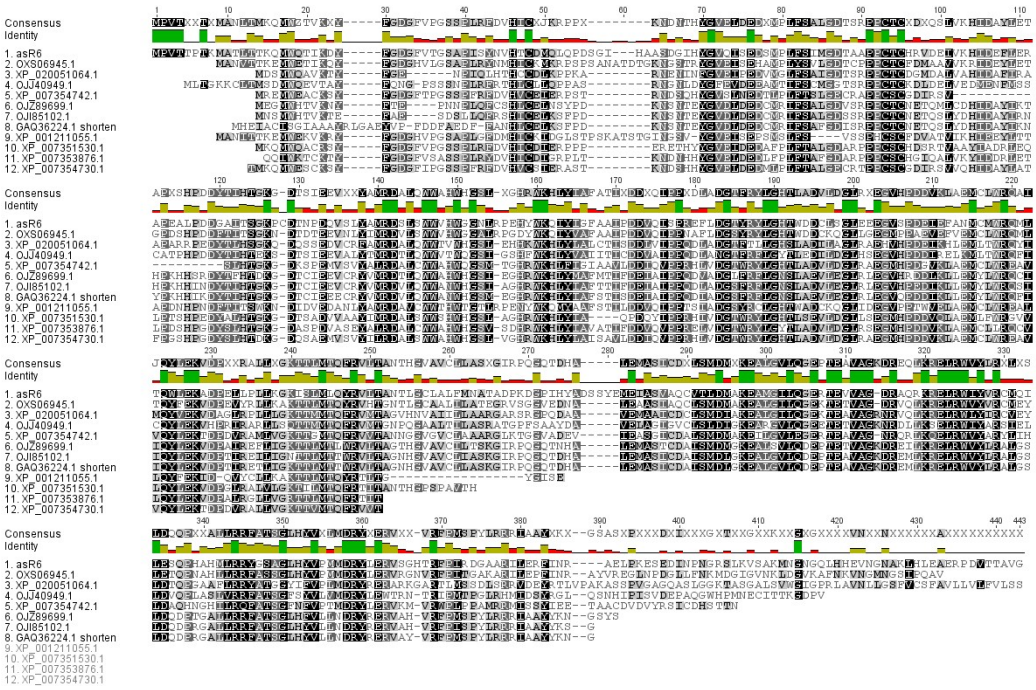


Figure 3.12 Protein alignment of AsR6 and homologous sequences identified by BLASTp.[103] Geneious 7.1.9 was used for the alignment.

3.4 Other BGC in *A. strictum*

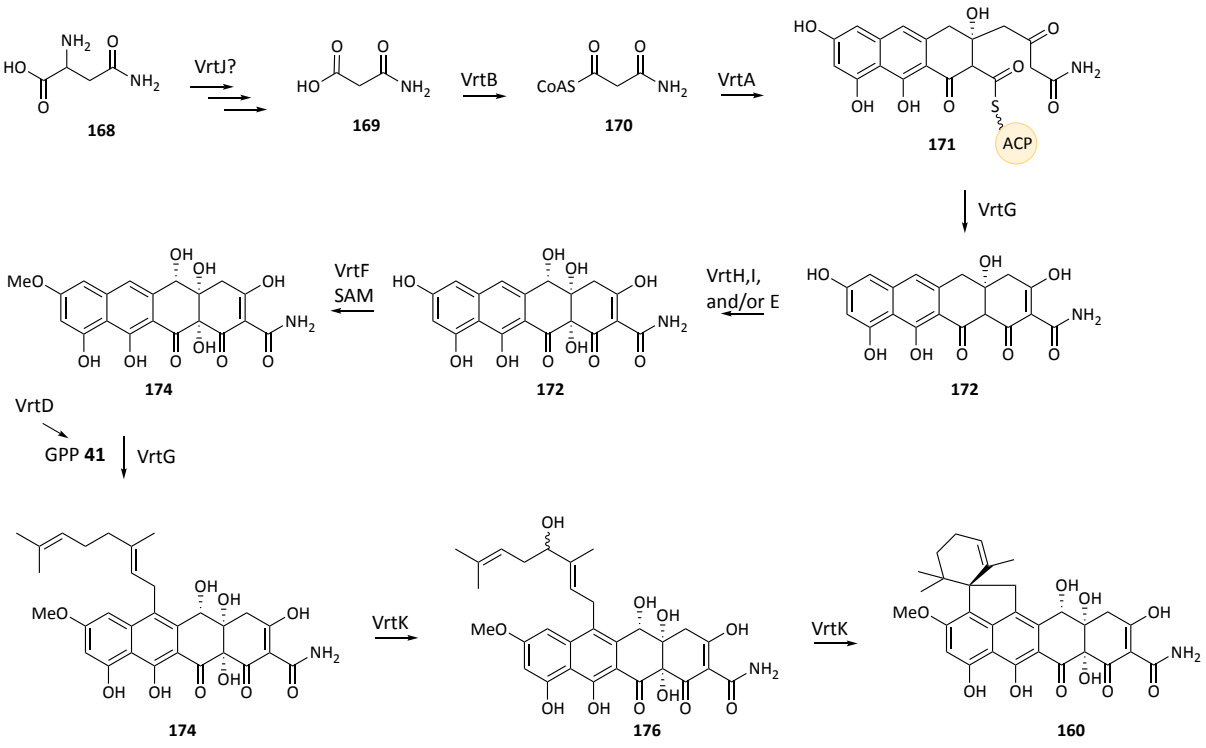
The antiSMASH analysis predicted another six BGC in the *A. strictum* genome with homologies to literature reported BGC (Chapter 3.1, Figure 3.3). These were further analysed with the artemis comparison tool.[159] Homologous protein sequences were pairwise compared using EMBOSS Needle.[160] Two of the analysed *A. strictum* BGC showed high homologies to the literature BGC and are further described in more detail (viridicatumtoxin **160** and aphidicolin **161**). The remaining four BGC (depudecin **162**, clapurine **163**, aspyridone **164**, leucinostaatin **165**: Appendix Tables 9.2-9.5) showed low homologies (22-50%) to the literature BGC and are not further discussed.

3.4.1 Viridicatumtoxin 160 like BGC

The viridicatumtoxin **160** BGC from *Penicillium aethiopicum* was identified and a possible biosynthesis was proposed by Tang and co-workers in 2010 (Scheme 3.3).[133,161] Comparison of the

A. strictum homologous BGC with the original BGC showed the presence of all genes. Amino acid identities of protein sequences were found to be between 39.1 – 77.9% (Table 3.7). This strongly suggested that the *A. strictum* BGC is likely to encode proteins involved in viridicatumtoxin **160** biosynthesis. However, the putative transcription factor *vrtR2* (*Asg9751*) is translocated. While in the *P. aethiopicum* BGC it is located at the cluster boarders, it is found in the middle (between *vrtD* and *vrtE* homologues) of the BGC in *A. strictum*. In addition, this putative transcription factor shows the lowest protein identities (39.1%) observed for the whole BGC. Analysis of the transcriptome data generated under xenovulene A **101** producing and non-producing conditions showed no abundant mRNA levels for the BGC. Attempts to identify virdicatumtoxin **160** in extracts of *A. strictum* grown in different liquid media (Chapter 2.9) were unsuccessful.

These findings suggest that the *A. strictum* putative viridicatumtoxin **160** BGC is inactive as transcription is repressed under the investigated conditions. However, it is also possible that due to the transcription factor translocation the cluster is destroyed and cannot be expressed.



Scheme 3.3 Proposed biosynthesis of viridicatumtoxin **160** in *Penicllium aethiopicum*.

Table 3.7 Viridicatumtoxin **160** like *A. strictum* BGC with 54% similarities (antiSMASH).

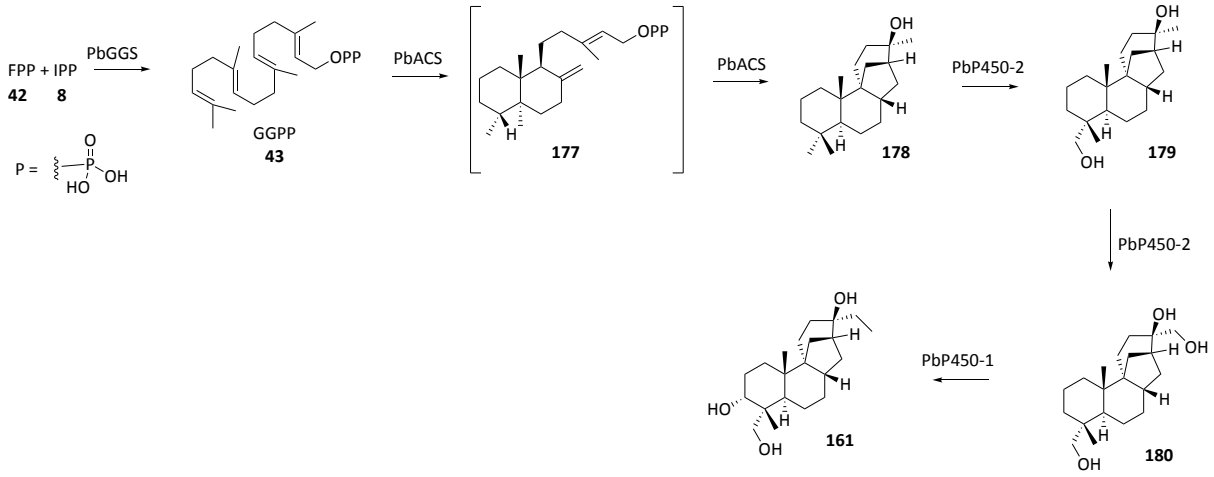
#	Putative function	Homologue	Identity/similarity %	BLASTp and CD
Asg9743	transporter	VrtL	73.9/84.9	MFS transporter
Asg9744	oxidoreductase	VrtK	77.0/85.9	cytochrome P450
Asg9745	transferase	VrtJ	66.9/77.9	PLP dependent transferase
Asg9746	oxidoreductase	VrtI	53.6/62.0	Fe dependent oxygenase
Asg9747	NR-PKS	VrtA	77.9/86.7	polyketide synthase, SAT-KS-AT-PT-ACP
Asg9748	acetoacetyl-CoA synthetase	VrtB	58.4/71.5	acetoacetyl CoA synthetase
Asg9749	prenyltransferase	VrtC	65.6/77.1	aromatic prenyltransferase
Asg9750	FPPS	VrtD	57.9/67.3	FPP synthase
Asg9751	regulation	VrtR2	39.1/49.1	transcription factor
Asg9752	oxidoreductase	VrtE	67.0/73.8	cytochrome P450
Asg9753	*O*-methyltransferase	VrtF	75.5/89.2	SAM-dependant *O*-methyltransferase
Asg9754	hypothetical	VrtG	63.8/75.3	metallo-β-lactamase superfamily
Asg9755	oxidoreductase	VrtH	71.1/84.7	FAD dependant monooxygenase
Asg9756	regulation	VrtR1	53.2/67.8	transcription factor

3.4.2 Aphidicolin 161 like BGC

The 15 kb BGC of the diterpene aphidicolin **161** was identified and characterised by Oikawa and co-workers from 2001 – 2011 (Scheme 3.4).[137,162,163] The homologous BGC found in *A. strictum* encodes the full set of genes necessary for aphidicolin **161** biosynthesis (Table 8.3). Pairwise protein identities were found to be between 25.2 – 64.8%. Similar to the viridicatumtoxin **160** homologous BGC (Chapter 3.4.1), the transcription factor here also shows the lowest homology of any protein encoded by the BGC. In comparison to the original BGC from *Phoma betae* only two genes are translocated. The PbGGS and Pb450-1 homologues encoding genes exchanged positions in the *A. strictum* BGC. Analysis of the transcriptome data suggested poor to no expression of the aphidicolin **161** BGC genes under the tested *A. strictum* fermentation conditions. However, this was not further investigated and *A. strictum* extracts were not analysed for the presence of aphidicolin **161**.

Table 3.8 Aphidicolin **161** like *A. strictum* BGC with 83% similarities (antiSMASH).

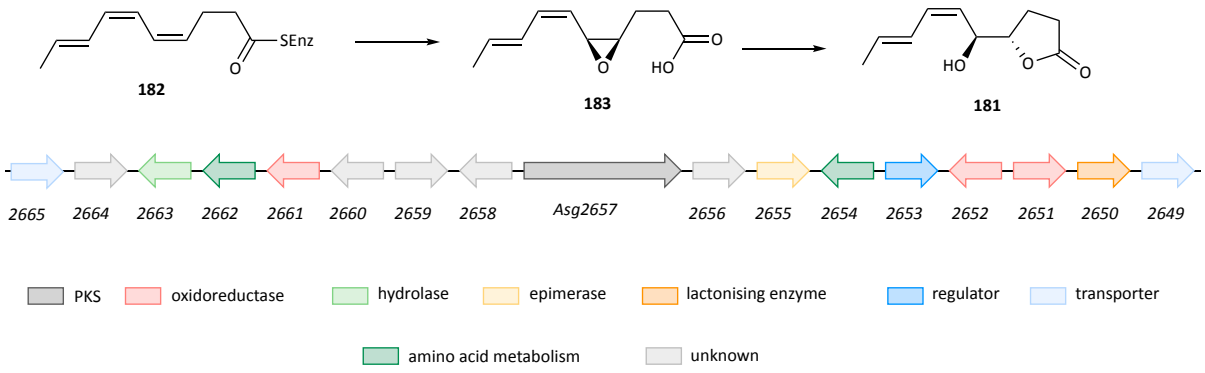
#	Putative function	Homologue	Identity/similarity %	BLASTp and CD
Asg2509	regulation	PbTF	25.2/38.8	transcription factor, GAL4
Asg2510	oxidoreductase	PbP450-2	64.8/72.9	cytochrome P450
Asg2511	transporter	PbTP	60.7/74.5	MFS transporter
Asg2512	GGPP synthase	PbGGS	57.9/68.9	polyprenyl synthase
Asg2513	terpene synthase	ACS	62.8/76.0	copalyl diphosphate synthase
Asg2514	oxidoreductase	PbP450-1	61.0/72.9	cytochrome P450



Scheme 3.4 Diterpene aphidicolin **161** biosynthesis in *Phoma betae*.

3.4.3 Lactone 181 Biosynthesis in *A. strictum*

In addition to xenovulene A **101** and related compounds, another unrelated putative polyketide derived natural product was isolated from *A. strictum* cultures grown under **101** producing conditions by Simpson and co-workers.[164] Attempts to identify lactone **181** in *A. strictum* extracts analysed in this work were unsuccessful. The pentaketide lactone **181** is proposed to be the product of a HR-PKS without *C*-MeT. Further conversion into the (*Z*,*Z*,*E*)-triene by a *cis-trans* double bond isomerase like enzyme could yield a cytochrome P450 substrate **182**. Epoxide formation and spontaneous opening could give the final lactone **181** (Scheme 3.5).[164]



Scheme 3.5 Proposed biosynthesis and putative BGC 33 for lactone **181** from *A. strictum*.

Mining the *A. strictum* genome identified 4 HR-PKS encoded by *Asg2657*, *Asg6902*, *Asg7194* and *Asg8744* with a matching domain architecture (Figure 3.4). Analysis of the BGC predicted by fungiSMASH[20–22] showed that only one BGC (33, *Asg2657*) also encodes a putative P450 (*Asg2661*) and a putative epimerase (*Asg2655*). In total, 17 ORF are predicted in this BGC. BLASTp[103] homology search of encoded proteins further identified a putative epoxide hydrolase (*Asg2663*), a lactonising

enzyme (*Asg2650*), an SDR (Asg2652) and an aldolase (Asg2652) (Table 3.9) which might be involved in further biosynthetic steps.

Table 3.9 Analysis of predicted ORF of BGC 33 by BLASTp.[103]

#	Putative function	BLASTp and CD
Asg2665	transporter	glycosyl hydrolase, MFS transporter
Asg2664	hypothetical	Zn peptidase family
Asg2663	hydrolase	epoxide hydrolase
Asg2662	amino acid metabolism	asparagine synthase, glucosamine 6-phosphate synthetase
Asg2661	oxidoreductase	cytochrome P450
Asg2660	hypothetical	enamine deaminase like
Asg2659	hypothetical	hypothetical
Asg2658	hypothetical	pimeloyl-ACP methyl ester carboxylesterase
Asg2657	HR-PKS	polyketide synthase, KS-AT-DH-ER-KR-ACP
Asg2656	hypothetical	hypothetical
Asg2655	*cis-trans* isomerase	NAD dependent epimerase/dehydratase family
Asg2654	amino acid transport and metabolism	Serine dehydratase
Asg2653	regulation	GAL4 Zn(2) Cys(6) transcription factor
Asg2652	Oxidoreductase, aldolase	short chain dehydrogenase/reductase, aldolase
Asg2651	reductase	NAD(P)(H) Rossmann fold 2-dehydrpantoate 2-reductase
Asg2650	Racemase/lactonising enzyme	glucarate dehydratase
Asg2649	transporter	MFS/sugar transporter

3.5 Classification of *A. strictum*

An ascomycete was first described in 1995 as the producer of xenovulene A **101**. Based on its morphology it was classified as *A. strictum*.[82] The genomic era has changed the classification of fungal species. Nowadays a barcoding sequence is additionally used to determine the taxonomy of newly discovered strains. For fingerprinting of fungi the internal transcribed spacer (ITS), a region between 18S and 28S ribosomal RNA (rRNA) including 5.8S rRNA, is compared to other known sequences. This method provides rapid confirmation of strain identities in laboratory use.[165] The ~600 bp ITS sequence of the **101** producer strain was amplified from gDNA with oligonucleotides P11+12, sequenced (GATC, Konstanz) and used as a BLAST nucleotide (BLASTn) search template in a homology [103] analysis. Best alignments were found to be with the *Sarocladium kiliense* ITS sequence, rather than the *A. strictum* ITS sequences, with a sequence identity of 98%. The genomes of *S. kiliense*[166] and *A. strictum* (isolate DS1bioAYav1.0) became available at the National Center for Biotechnology Information (NBCI) and the Joint Genome Institute (JGI) during the period of this work. Pairwise comparison of average nucleotide identities (ANI) with the *A. strictum* **101** producing strain by Dr Daniel Wibberg was used to analyse their relationship. The comparison showed ANI values to be ~82% for all three compared strains. Isolates representing identical species, like two strains of *Aspergillus niger* (SH2 and ATCC2), *Candida albicans* (WO-1 and A20) or *Fusarium oxysporum* (f.sp. cubense Race1 and Race4), were found to have ANI value of 97-99%.[167]

The calculated low pairwise identities of 81% and 83% of the **101** producer to *S. kiliense* and *A. strictum* indicate that it is a novel species of the *Acremonium*/*Sarocladium* family.

Table 3.10 Pairwise compared ANI for *A. strictum* **101** producer, *S. kiliense* and *A. strictum* DS1bioAYav1.0.

ANI in %	*S. kiliense*	*A. strictum* DS1bioAYav1.0	*A. strictum* 101 producer
A. strictum **101** producer	81.22	83.13	100.00
A. strictum DS1bioAYav1.0	83.18	100	
S. kiliense	100		

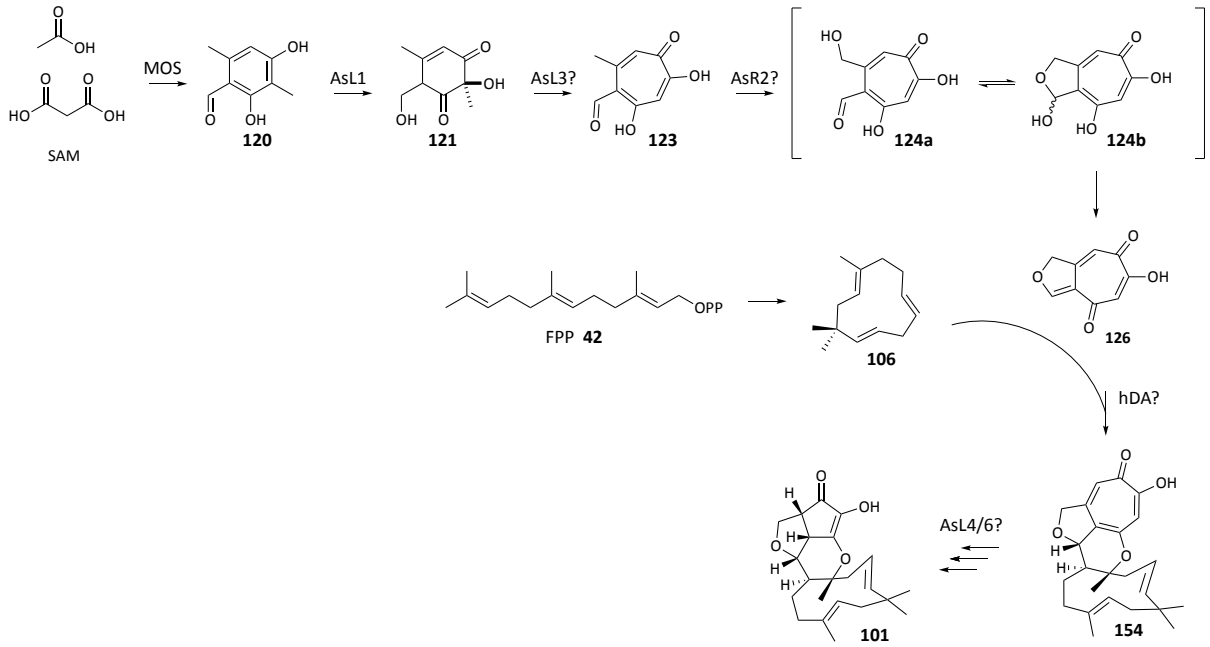
3.6 Discussion

Analysis of the *A. strictum* whole genome identified a BGC which could be involved in xenovulene A **101** biosynthesis. In addition, the **101** producing *A. strictum* strain was found to be wrongly classified as *A. strictum* and probably represents a novel member of the *Acremonium*/*Sarcoladium* family. For the time of this work it will nevertheless be referred to as *A. strictum*, but for the future the proposal of a novel name should be considered.

The *A. strictum* draft genome of 33.8 Mb was assembled in 51 scaffolds by Dr Daniel Wibberg. 10622 genes with an average number of 2.99 exons per ORF were predicted. In a comparative study of 82 ascomycete draft genomes an average genome size of 36.8 Mb was calculated and predicted to encode an average of 11129.45 genes with an average 2.58 exons per gene. The genomes compared in this publication were assembled on 7-3164 scaffolds, but 43% of those genomes consist of <100 scaffolds.[168] The *A. strictum* genome as proposed in this work thus represents an average ascomycete draft genome, with a reliable assembly in 51 scaffolds which enabled further antiSMASH analysis for secondary metabolite BGC.[20–22] For the transcriptome analysis under producing/non-producing conditions ~125/80 million reads were sequenced and ~95% could be mapped to the *A. strictum* draft genome. For comparison, the published transcriptome analysis of *Fusarium oxysporum* (*F. oxysporum*) resulted in ~63 million reads, with~91% located on the genome. This shows that quality and quantity of the transcriptomic data sequenced for either condition of *A. strictum* is in the expected range. It was used to identify differential gene expression with the DESeq software package.[146,169] The program was previously used to successfully analyse fungal mRNA expression in samples from *Rhizoctonia solani* taken during interaction with its plant host lettuce.[142]

CD[138–141] analysis of PKS encoding genes showed that only one NR-PKS (MOS encoded by *aspks1*)[102] with a domain architecture suitable for production of the predicted precursor aldehyde **120** of xenovulene A **101** is encoded in the entire *A. strictum* genome. Comparative gene expression

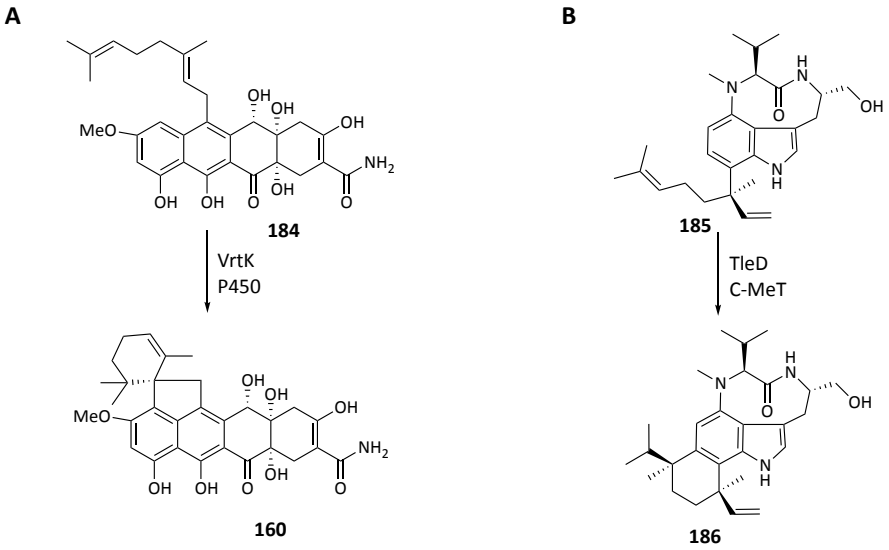
analysis (DESeq[146]) confirmed a distinct co-regulation of 15 genes accompanying *aspks1* under xenovulene A **101** producing conditions. Further BLASTp[103] homology searches of these co-located, upregulated genes confirmed tropolone core genes (*aspks1*, *asL1*, *asL3* and *asR2*)[97,102] and identified candidate genes for further transformation steps. Two of these genes encode putative aromatic hydroxylases which could be involved in the two ring contraction steps (AsL4, AsL6). Several (*asR4*, *asR5*, *asR6*, *asR7*) genes encode proteins of unknown function which are candidates to be involved in the fusion of polyketide and terpene moiety (Scheme 3.6). A predicted short-chain dehydrogenase (AsL5) was not directly assigned on the biosynthetic route towards xenovulene A **101**, but due to its expression it is assumed to further modify **101**.



Scheme 3.6 Proposed biosynthesis of **101**.

A known gene encoding a terpene cyclase to synthesise humulene **106** is not predicted to be part of the upregulated 16 genes. It was first assumed that the *aspks1* BGC could be split, as observed in the biosynthesis of the meroterpenoid austinol **84**.[76] However, DESeq analysis did not yield terpene cyclase which is co-regulated or constitutively expressed but located at a different genomic locus under **101** producing conditions. At this point it was further hypothesised that the terpene cyclisation could occur similarly to viridicatumtoxin **160**, where a cytochrome P450 oxygenase was shown to serve as a terpene cyclase (Scheme 3.7A).[133] A similar phenomenon has also been observed in the biosynthesis of the indol terpenoid teleocidin **186**, where a *trans* located *C*-methyl transferase initiates terpene cyclisation (Scheme 3.7B).[170] However, both unusual terpene cyclases have in common that the precursor terpene is attached to the polyketide precursor by a

prenyltransferase before the cyclisation occurs (similar to integral membrane protein mediated FPP cyclisation, Chapter 1.2). Although a cytochrome P450 oxygenase is encoded in the *aspks1* BGC no prenyltransferase is predicted. Additionally, the fusion pattern (Scheme 3.6) of humulene and polyketide in **154** indicates a terpene cyclisation reaction to happen first followed by a [4+2] cycloaddition to fuse humulene and polyketide (Scheme 3.6). This mechanism is also supported by biomimetic *in vitro* studies towards epolones **113**/**114** (Scheme 1.20).[111,113]

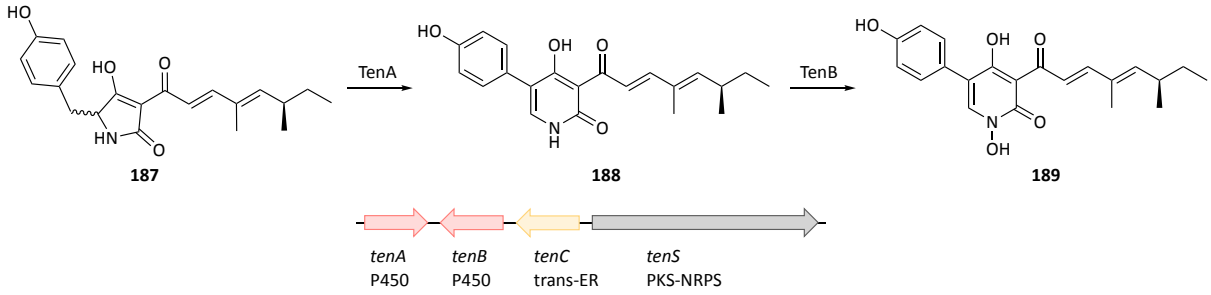


Scheme 3.7 Unusual terpene cyclisation in the biosynthesis of **A**, viridicatumtoxin **160** and **B**, teleocidin **186**.

Although the obtained genome and transcriptome data provided a good platform for the *in silico* analysis of **101** biosynthesis, the identification of a humulene synthase remained elusive. It was first assumed that *A. strictum* provides humulene **106** through one of the weakly expressed (Chapter 3.3.1) terpene synthase genes. However, later in this work AsR6 was shown to produce humulene **106** *in vitro* (Chapter 5.5). Bioinformatic analysis of AsR6 and known terpene cyclases did not reveal homologies. Thus, with AsR6, the *aspks1* BGC encodes an unknown type of terpene cyclase.

4 Targeted Gene Knockout and Knockdown in *A. strictum*

Gene disruption or silencing methods are widely used to investigate functions of encoded proteins and thus are a practical tool to elucidate natural product biosynthesis. Disrupting or silencing genes from a BGC and analysing the changes in secondary metabolite production links genes and proteins with biosynthetic steps. For example, during the elucidation of tenellin **189** biosynthesis in the filamentous fungus *Beauveria bassiana* both strategies were used successfully. Targeted knockout (KO) of *tenS*, encoding a PKS-NRPS, showed total deficiency in **189** production. Silencing of *tenA* or *tenB*, encoding two P450s, led to the accumulation of intermediates pretenellin-A **187** and pretenellin-B **188**, respectively. Thus, it was possible to assign substrates, enzymes and the catalysed reactions in the biosynthetic route **189**.[171,172]



Scheme 4.1 Biosynthesis and BGC of tenellin **189** in *B. bassiana.*

The two strategies have in common that exogenous DNA targeting the gene of interest (GOI) is inserted into the natural product producing organism. However, gene disruption results in KO of the targeted gene, whereas silencing methods rely on post-transcriptional mRNA degradation.

Different techniques for transformation of DNA into fungal cells have been proven successful in a wide range of filamentous fungi: *Agrobacterium tumefaciens* mediated transformation (AMT),[173] $CaCl_2$/polyethylene glycol (PEG) treatment;[174] electroporation of protoplasts;[175] or particle bombardment (biolistics),[176] which can be used if formation of protoplast is insufficient.

Once the exogenous DNA is inside the fungal cell it is integrated into the genome. In order to disrupt the targeted gene, the transformed DNA is designed to contain two sequences homologous to the GOI which flank a selection marker on either side (Figure 4.1). By homologous integration of this KO cassette the targeted gene is inactivated. Transformants are then selected using appropriate conditions, for example use of the *E. coli* gene *hph*, encoding a hygromycin B phosphotransferase[177] as selection marker for the antibiotic hygromycin B.[178]

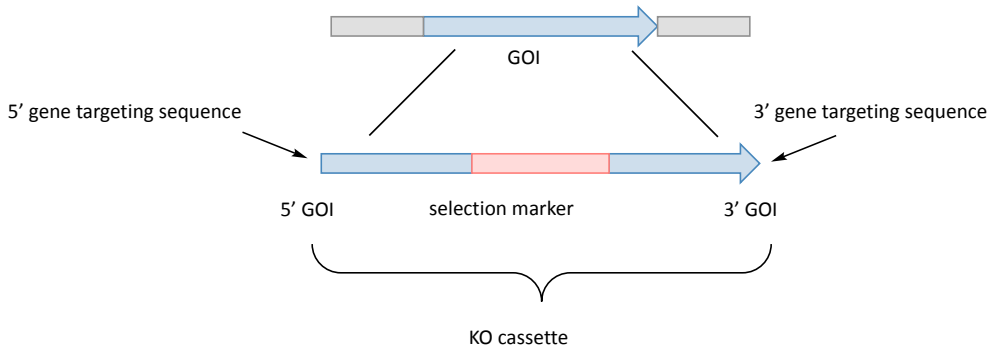


Figure 4.1 Gene disruption by homologous recombination of selection marker flanking sites and GOI.

KO methods exploit a naturally occurring DNA repair mechanism of the cell, which mends DNA lesions caused by oxidative and mechanical stress or ionizing radiation.[179] The enzyme Rad51 and a number of associated proteins are involved in *homologous recombination* in yeast *Saccharomyces cerevisiae*. This organism was reported to efficiently recombine DNA with 30 bp homology sequence (Figure 4.2).[180,181,182,183]

Although homologous recombination is the preferred mechanism of DNA repair in *S. cerevisiae*, it is not frequently observed in most ascomycetes and homologous sequences of several hundred bp are often required for effective recombination.[184] Instead, a second DNA repair mechanism, *non-homologous end joining* (NHEJ), which is independent of sequence homology, is used ubiquitously.[185] The ku heterodimer, consisting of two DNA binding proteins (ku70 and ku80), recognizes DNA double-strand breaks and recruits further enzymes such as a DNA-dependent protein kinase and the DNA ligase IV-XRCC4 complex.[182,186,187] NHEJ mediated DNA repair is error prone and the integration of exogenous DNA occurs at random genetic loci as no sequence homology is required (Figure 4.2). This ectopic integration of KO cassettes by NHEJ results in transformants with undisrupted GOI, which are able to grow under selective conditions. Often, homologous recombination frequencies of <20% are observed in filamentous fungi, which increases the number of transformants that have to be analysed. It was shown that KO of *ku70* or *ku80* results in an increase of homologous recombination rates up to 100% for *N. crassa*[188] and other fungal species.[184] However, generating Δ*ku*70 or Δ*ku*80 strains must first be achieved, and this can be difficult in itself.

A modified gene disruption method, known as the bipartite marker strategy, was developed by Nielsen *et al*. in *A. nidulans*. For this method two individual DNA fragments, each composed of a gene targeting sequence and a non-functional part of a selection marker are co-transformed. Only when the selection marker is recombined by three homologous recombination events will the resistance cassette be reconstructed and colonies be formed on selective medium (Figure 4.3). The

selection marker, usually consisting of a structural resistance gene for the appropriate antibiotic and a constitutive fungal promoter/terminator, is divided within the structural gene onto the two DNA fragments. This way the possibility of the 3' bipartite substrate being functional if integrated downstream of a native promoter is prevented. Thus ectopic integration of either DNA fragment alone cannot result in the formation of selectable transformants. It is hypothesized that this initial essential homologous recombination event channels the entire recombination towards homologous recombination rather than ectopic integration. This method was shown to increase gene targeting in *A. nidulans* by ~40%.[189]

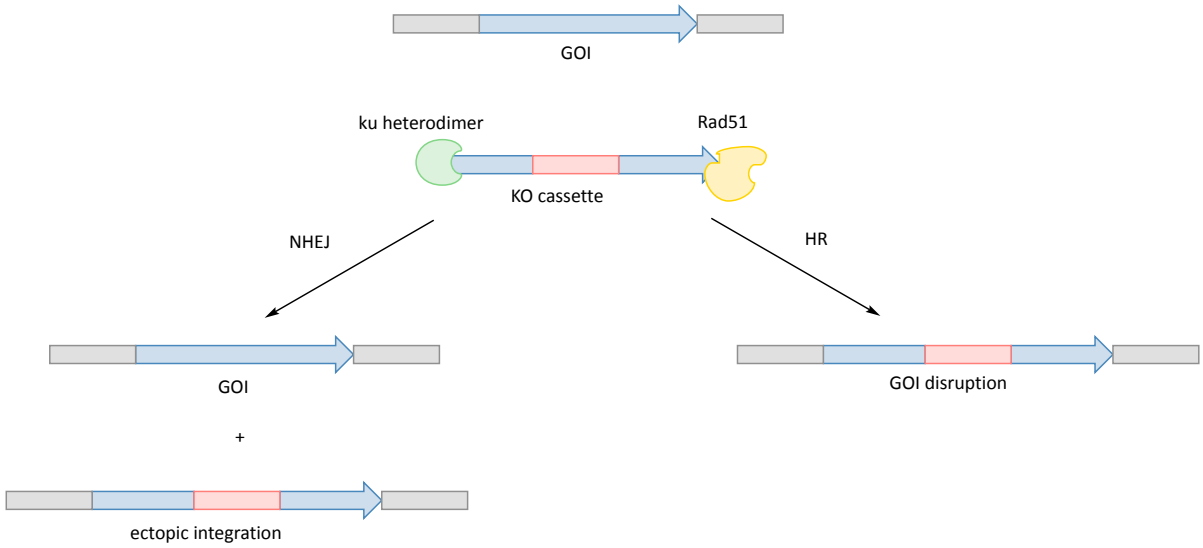


Figure 4.2 Ectopic DNA integration by non-homologous end joining (NHEJ) and targeted gene disruption by homologous recombination (HR).

The most recent developments in targeted gene disruption are genome-editing tools where engineered nucleases are deployed to insert, remove or replace DNA in a very specific manner. These include meganucleases, zinc-finger nucleases, transcription activator-like nucleases and the currently most popular CRISPR/Cas9 (type II clustered regularly interspaced palindromic repeats with the CRISPR-associated protein 9). [190,191] The multidomain nuclease Cas9 from *Streptocccus pyogenes* induces a DNA double-strand break at the locus base pairing a 20 bp guide RNA (gRNA). gRNA sequences can be adapted for the GOI, but must be adjacent to a downstream 5'-NGG motif (protospacer adjacent motif, PAM). The double-strand breaks can be repaired by NHEJ resulting in deletions or insertions of variable length. In the presence of a homologous donor DNA (dDNA) the DNA double strand break can also be repaired by homologous recombination (Figure 4.4).[190] In 2015 this technique was first reportedly used in *Trichoderma reesei*, where it was shown to improve homologous recombination efficiency to almost 100%.[192]

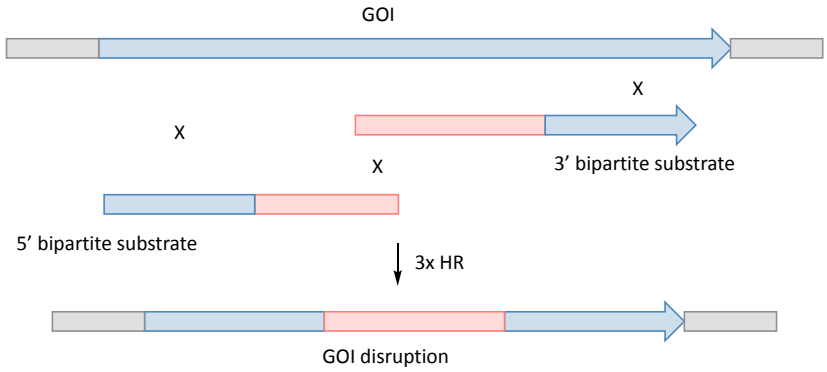


Figure 4.3 Integration of two DNA fragments with three homologous recombination (HR) events to reduce ectopic integration by the bipartite marker strategy.

A different method for the exploration of gene function is post-transcriptional gene silencing which often results in decreased yields rather than complete absence of the investigated secondary metabolite. Introduction of antisense RNA (asRNA) to the mRNA of the GOI induces the formation of aberrant double stranded RNA (dsRNA) by an RNA dependent RNA polymerase (RdRP). This initiates the endogenous silencing pathway in fungi which is involved in gene regulation and genome stability.[193,194] The dsRNA is degraded into small interfering RNA (siRNA) of ~22 nt by the endoribonuclease Dicer. This in itself results in the destruction of much mRNA, but in a second, and amplifying, process these ~22 nt fragments are recruited to the protein Argonaute (Ago) and assembled into an RNA-induced silencing complex (RISC).

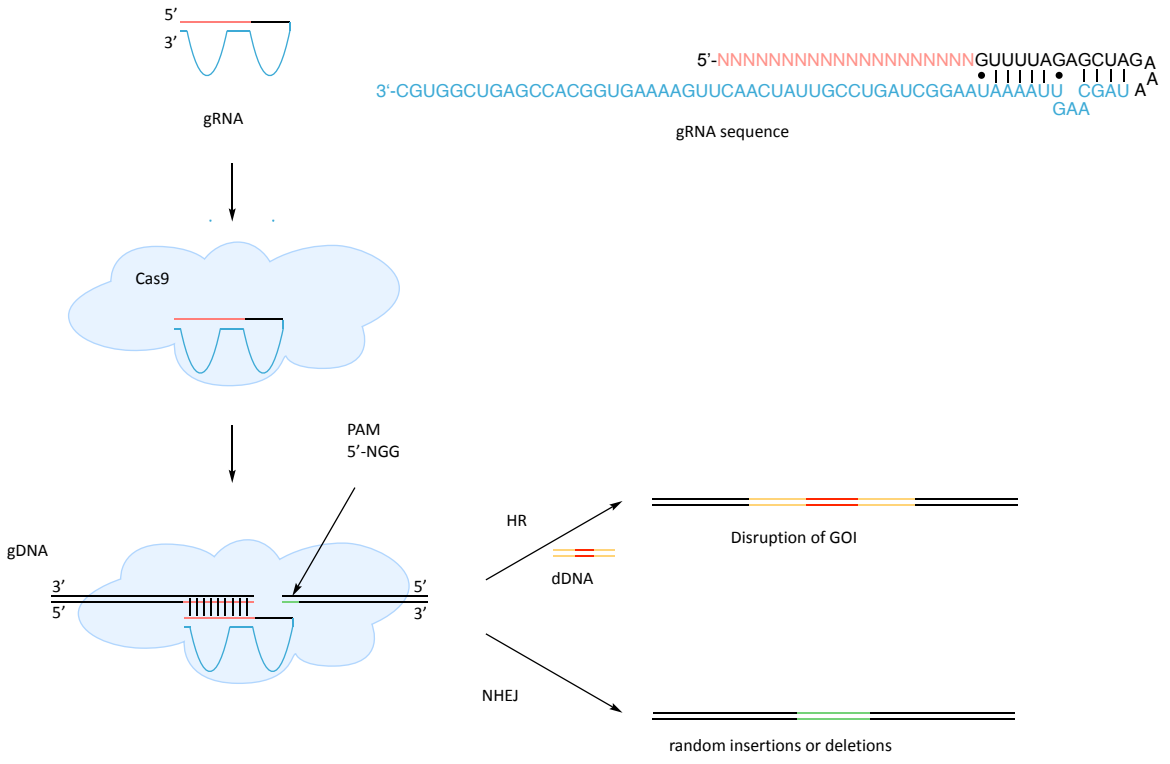


Figure 4.4 CRISPR/Cas9 mediated gene editing.

The siRNA guides RISC to the remaining mRNA which is then degraded by Ago, resulting in low levels of mRNA and thus reducing translation to protein (Figure 4.5).[193,195] This approach is independent of homologous recombination, because the reverse GOI (revGOI) can be introduced at a random locus on the genome, while the actual targeting occurs in the cytosol. Gene silencing in *B. bassiana* yielded more positive transformants and resulted in the same phenotype of metabolite production as generated KO strains. This indicates that asRNA mediated gene silencing, in this case, induced the degradation of almost 100% mRNA.[172]

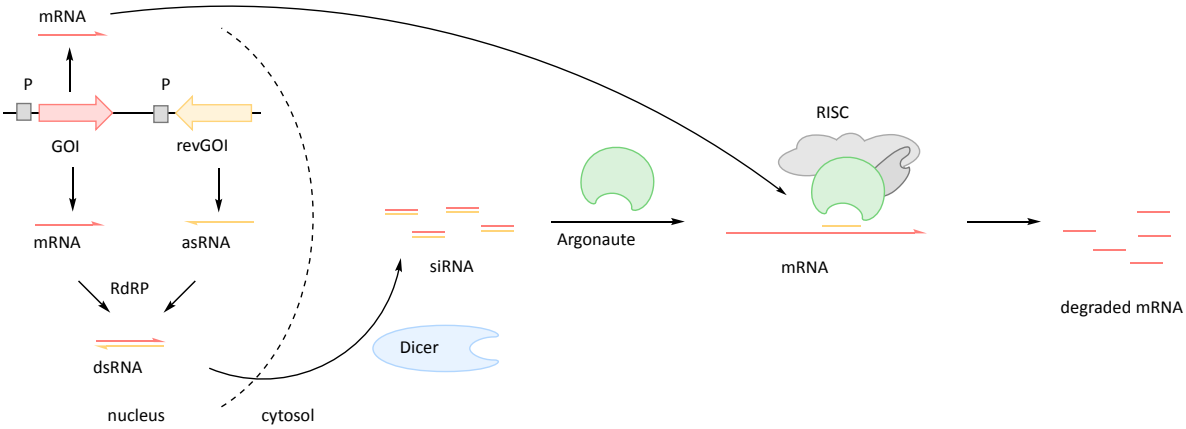


Figure 4.5 siRNA induced silencing mechanism in eukaryotes.

4.1 Transformation of *A. strictum*

The xenovulene producing *A. strictum* strain has been shown to efficiently produce protoplasts and was first successfully transformed with a $CaCl_2$/PEG method by Dr Kate Harley from our group.[127] The strain was found to be sensitive to the aminoglycoside antibiotic hygromycin B in concentrations of 50 µg/mL and a hygromycin resistance cassette was shown to be effective as selection marker during transformation.[127]

4.1.1 Antibiotics Screening

During transformation the selection marker is stably integrated into the fungal genome and cannot be recycled. For multiple rounds of transformation and selection it is beneficial to have more than one antibiotic to which the investigated microorganism shows sensitivity.

To analyse whether *A. strictum* shows sensitivity to an additional antibiotic it was sub-cultured on CD with sorbitol (CD+S) agar containing different concentrations (10, 50, 100, 200 µg/mL) of glufosinat, geneticin, neomycin or zeocin, respectively. Observation of the colonies at 25 °C for 16 d and 105 d showed that *A. strictum* is able to grow on all antibiotics at a concentration of 10 µg/mL.

However, geneticin effectively inhibits the growth at a concentration of 50 µg/mL for 16 d, whereas colonies were observed for all other tested selection reagents. At a concentration of 100 µg/mL geneticin induced growth inhibition is stable for at least 105 d (Figure 4.6). These observations suggest that geneticin could be used as an antibiotic for selection in a second round of *A. strictum* transformations.

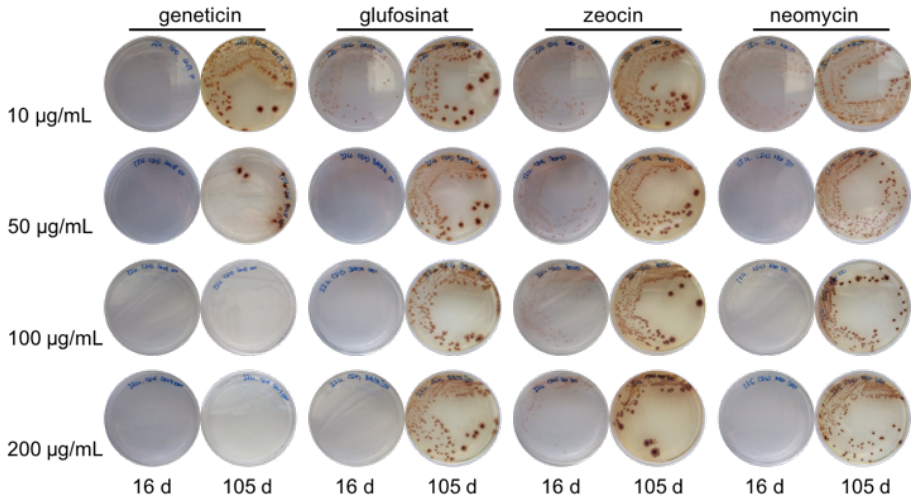


Figure 4.6 *A. strictum* CD+S agar plates with different concentrations of antibiotics at two time points.

4.1.2 Transformation of *A. strictum* with pTH-GS-*egfp*

A reliable transformation method enables further genetic manipulation of *A. strictum* and was re-established and optimised using the vector pTH-GS-*egfp*. The vector contains a resistance cassette (hyg^R) for selection with hygromycin B and the *egfp* gene, encoding an enhanced green fluorescent protein (eGFP), as visible marker. The *egfp* gene is expressed under the control of the starch or maltose inducible P_{amyB} promoter and terminator from *A. oryzae* (Figure 4.7).[196,197]

A. strictum PEG/$CaCl_2$ mediated protoplast (Figure 4.7B) transformation with pTH-GS-*egfp* (Figure 4.7A) resulted in 12 colonies growing on tertiary selection plates. For fluorescence analysis 6 transformants were sub-cultured in ASPM liquid medium, which contains maltose as the carbon source for the induction of P_{amyB}. Samples of 4 d old transformants and A. strictum WT were microscopically analysed. Upon excitation (450 – 490 nm) the transformants displayed a green fluorescence correlating to the expression of *egfp* and translation into protein, whereas wildtype control does not emit green light (Figure 4.7C I-IV). This confirms the successful transformation and selection method as well as the effective initiation of transcription downstream of P_{amyB} in *A. strictum*. These findings confirm previous results of Dr Jack Davison from our group, who previously showed that *efpg* was expressed in *A. strictum* downstream of P_{amyB}.[104]

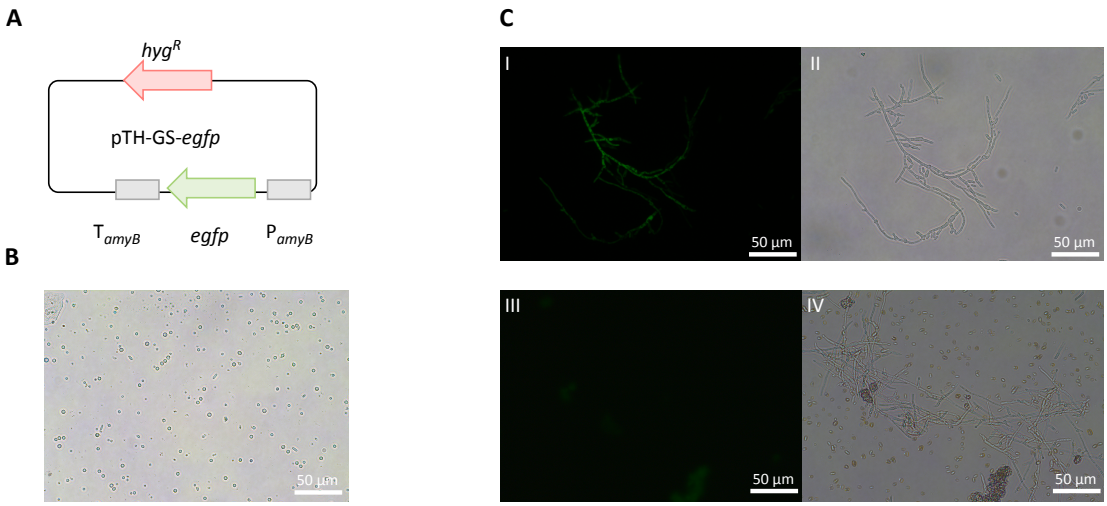


Figure 4.7 *A. strictum* transformation. **A,** Vector map of pTH-GS-*egfp*; **B**, *A. strictum* protoplasts; **C**, *A. strictum* transformed with pTH-GS-*egfp* (I+II) and WT (III+IV), I+III) upon excitation and II+IV) bright-field.

Although the transformation method was successfully re-established with pTH-GS-*egfp*, recovering *A. strictum* transformants was difficult. Of 18 further transformations targeting *aspks1* (Chapter 4.3.2) only four yielded a reasonable number of transformants (8 to 14), indicating a major selection issue, as 14 transformations yielded no colonies at all. Two main observations were made: either the transformation plates were overgrown after three to five days; or no colonies could be recovered at all. To address this, several changes to the original transformation protocol were made. Firstly, to remove spores causing a high background, cells were collected by miracloth instead of centrifugation. Secondly, different molecular weights PEG (3350, 4000) were tested, but gave inconclusive results. And finally, instead of overlaying the plates 16 h after transformation, the protoplasts were directly plated onto solid media containing 50-100 µg/mL hygromycin B.

These changes improved the process but still did not result in a reliable transformation efficiency of *A. strictum*. In a comparison experiment, selection agar was prepared with tap, millipore or VE water and used for transformation. Tap water resulted in overgrown plates after three days, millipore water yielded no colonies and VE water resulted in colonies growing after 10 d. This sensitivity towards the water quality does not affect the positive control, where untransformed protoplasts are plated on selection agar without antibiotics. For those protoplasts recovery is observed after 16 h under all tested conditions. This experiment showed that *A. strictum* protoplast recovery in the presence of hygromycin B is depending on the water quality and has a great impact on the success rate of the transformation.

4.1.3 Nuclei Staining of *A. strictum*

Fungal protoplasts are generated through the incubation of freshly grown mycelia with a cell wall degrading enzyme mixture. Usually one protoplast contains an uncertain number of nuclei.

Therefore, primary transformants are often heterokaryons and have to be genetically purified by re-streaking of conidiaspores. Depending on the number of nuclei per conidiaspore the efficiency of this process can vary.[198] The fluorescent dye 4',6-diamidino-2-phenylindole (DAPI) binds AT rich regions of DNA in the minor groove and can be visualized upon excitation (358 nm) by microscopy.[199] It was used to stain a 2 d old sporulating *A. strictum* sample obtained from liquid culture. Excitation and microscope analysis showed that every conidiaspore harbours a single nucleus (Figure 4.8B). These findings facilitated genetic purification of *A. strictum* in further experiments.

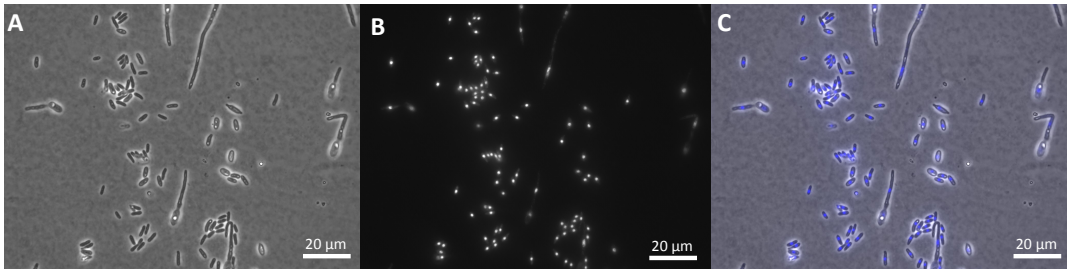


Figure 4.8 *A. strictum* spores with DAPI stained nuclei, **A,** bright-field **B,** DAPI and **C,** overlay of A+B.

4.1.4 Assembly of Vector Constructs by Homologous Recombination in *S. cerevisiae*

For targeted gene disruption a KO cassette consisting of a selection marker (*hyg*^{R} or the geneticin resistance cassette *gen*^{R}) flanked by arms homologous to the GOI has to be assembled. *S. cerevisiae* can recombine DNA fragments by homologous recombination with an overlap sequence of 30 bp. It has been used extensively to build up vector constructs through PEG mediated co-transformation of individual DNA fragments and a vector backbone (Figure 4.9).[180,200]

In this work pE-YA[201] was used as vector for the assembly of all KO and silencing constructs of Chapter 4**.** This vector contains the *2µ* origin of replication (*ori*) and *ura3* gene (encoding an orotidine 5'-phopsphate decarboxylase) to select uracil auxotroph *S. cerevisiae* in uracil and uridine free medium. All DNA fragments for yeast assembly were amplified by PCR using gDNA of *A. strictum* as gene specific template. For amplification of the fungal selection marker appropriate vector DNA (pTH-GS-*egfp*, pTYG-*gen*^{R}) was used as the PCR template (Figure 4.10). The necessary 30 bp sequence overlap for homologous recombination of the individual DNA fragments was introduced by PCR through tails on the designed oligonucleotides. pE-YA was linearized by restriction hydrolysis with either *Asc*I, *Not*I or both enzymes prior to PEG mediated co-transformation with the PCR amplified DNA fragments. Assembled vector DNA was isolated from *S. cerevisiae* by kit and directly used to transform competent *E. coli*. For propagation in *E. coli* pE-YA has an additional *ori* (*pUC*) and the *nptII* gene, encoding the neomycin phosphotransferase II, which

also confers resistance to kanamycin (*kan*^{R}). Obtained *E. coli* colonies were screened by PCR and confirmed by sequencing (Eurofins, Ebersberg).

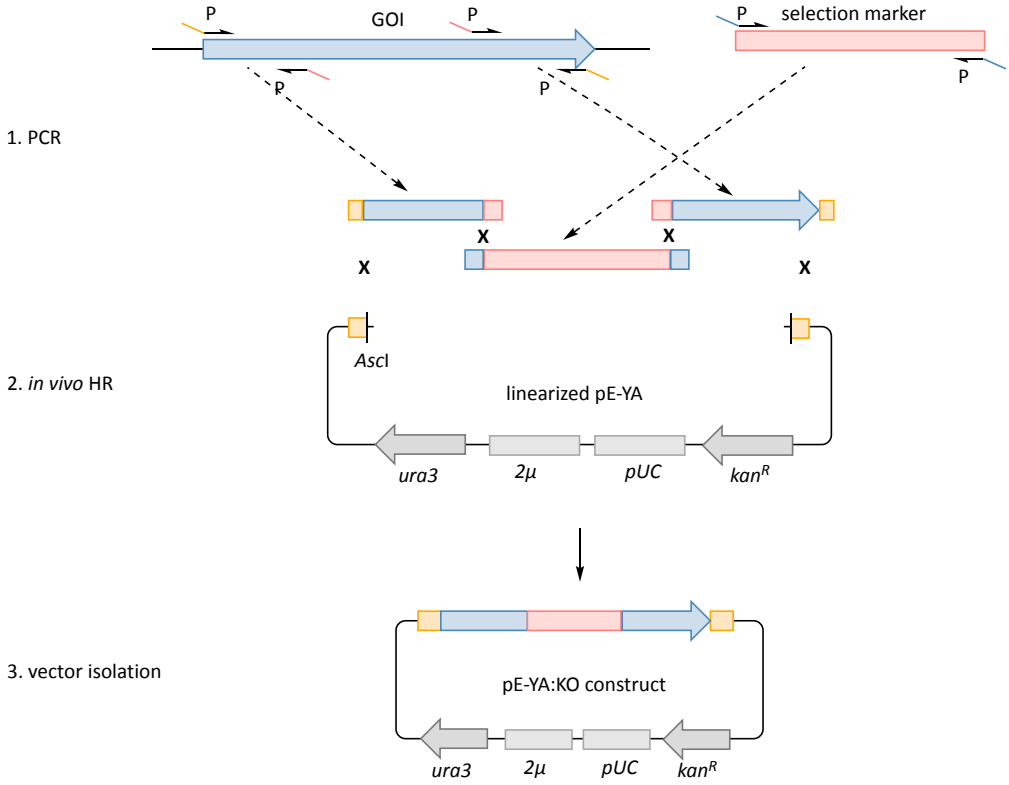


Figure 4.9 Yeast homologous recombination (HR) cloning strategy.

The selection marker for hygromycin B (*hyg*^{R}) and geneticin (*gen*^{R}) consists of the respective bacterial resistance gene *hph* or *nptII* (also confers resistance to geneticin) (Figure 4.10). The resistance genes are expressed under the control of the *A. nidulans* promoter/terminator P_{gpdA}/T_{trpC}.

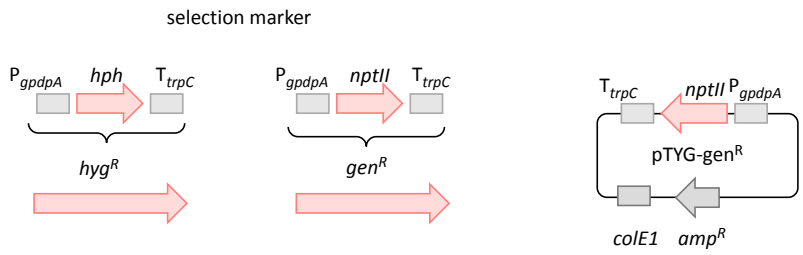


Figure 4.10 Selection genes under the control of the fungal P_{gpdA} and T_{trpC} and vector map of pTYG-*gen*^{R}.

4.2 Attempted Targeting of *asl4*, *asL5* and *asL6*

All previous attempts for gene disruption or silencing (Chapter 4.3) in *A. strictum* within our group were targeting *aspks1*. It has never been attempted to target any of the clustered genes for disruption. In other filamentous fungi, e.g. *Magnaporthe grisea* and *A. nidulans* it was shown that

gene disruption is locus dependent.[202,203] In *M. grisea* the secondary metabolite *ACE1* gene cluster, involved in avirulence signalling during rice infection, was targeted for different gene disruptions. However, it was only possible to disrupt 3 of 15 genes at this genetic locus.[202] A similar phenomenon was observed in *A. nidulans*, where disruption of two genes (*niaD*, encoding a nitrate reductase, and *amdS*, encoding an acetamidase) showed that targeting of one (*niaD*) genetic locus was five-fold more efficient than the other.[203] By targeting *aspks1* clustered genes it should be investigated whether disruption of genes other than *aspks1* is more efficient. In addition, successful KO of a putative tailoring enzyme encoding gene could elucidate its role during the biosynthesis of xenovulene A **101**.

4.2.1 Attempted Bipartite Knockouts of *asL4*, *asL5* and *asL6*

The genes *asL4* and *asL6*, both encoding putative FAD dependent monooxygenases, and *asL5*, encoding a putative short-chain dehydrogenase/reductase, were selected for gene knockouts with the bipartite method (Figure 4.11).[189]

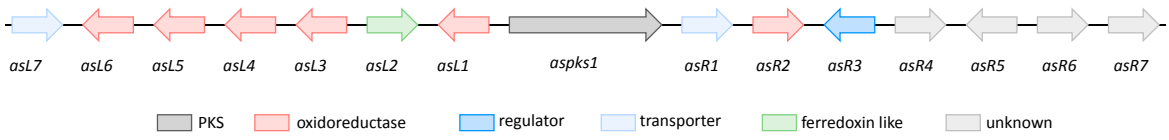


Figure 4.11 *aspks1* BGC.

The KO cassettes (RSI77 1-6) for each gene were constructed by *in vivo* homologous recombination in *S. cerevisiae* with selection markers for hygromycin B (RSI77 2, 4 and 6) and geneticin (RSI77 1, 3 and 5), respectively (Figure 4.12 and Table 7.7). DNA fragments for assembly were amplified by PCR using oligonucleotides P_a-P_f (Figure 4.12 and Table 7.7, Chapter 4.1.4).

Table 4.1 Overview of transformants obtained by bipartite substrate KO and chemical or genetic analysis.

Transformation Experiment	GOI	Resistance cassette	Plasmid ID for PCR	#transformants	# of chemically analysed transformants	# of genetically analysed transformants
RSI93 3	*asL5*	gen^R	RSI77 3	8	8	-
RSI93 5	*asL6*	gen^R	RSI77 5	8	8	-
RSI103/104 2	*asL4*	hyg^R	RSI77 2	1	-	1
RSI103/104 4	*asL5*	hyg^R	RSI77 4	3	-	3
RSI103/104 6	*asL6*	hyg^R	RSI77 6	18	-	18

The 5’ and 3’ bipartite substrates for *asL4*, *asL5* or *asL6* were amplified by PCR using oligonucleotides P_{g+h} and P_{i+j} from the corresponding template vector DNA (RSI77 1–6, Figure 4.12). The DNA was purified and directly used for PEG/$CaCl_2$ mediated protoplast transformation of

A. strictum. In total 1 transformant targeting *asL4*, 11 transformants targeting *asL5* and 26 transformants targeting *asL6* were obtained from three rounds of transformations (Table 4.1).

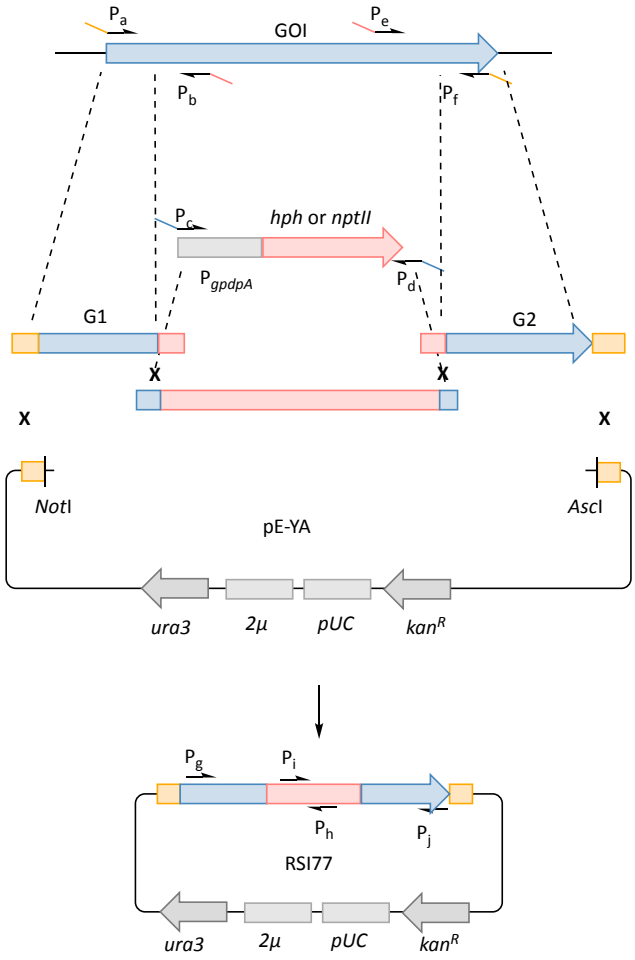


Figure 4.12 General method for assembly of KO cassettes for *asL4*, *asL5* and *asL6* in pE-YA by yeast homologous recombination (HR).

For chemical analysis the 16 transformants obtained from transformation RSI93 (Table 4.1) were grown under xenovulene A **101** producing conditions, extracted following a small scale (1 mL) protocol and analysed by analytical LCMS. Production of xenovulene A **101** was observed by UV absorption, mass and t_R in all chromatograms (Figure 4.13).

For genetic testing of the 22 transformants obtained from transformations RSI103 and RSI104 (Table 4.1) gDNA was extracted and used as PCR template. Disruption of genes was analysed with oligonucleotides specific for the targeted genes and compared to *A. strictum* WT gDNA (Figure 4.14). A DNA fragment corresponding to the size of non-disrupted GOI was amplified for all tested transformants, rather than the ~3 kb longer KO cassette.

All chemically analysed transformants still produced **101** and all genetically tested transformants possessed a non-disrupted GOI. This indicates that the bipartite substrate cassette has ectopically

integrated within the genome for all targeted genes (*asL4*, *asL5*, *asL6*). These findings suggets that either the genetic locus of the whole *aspks1* BGC is protected from gene disruptions by homologous recombination or non-homologous end joining is much more effective in *A. strictum*.

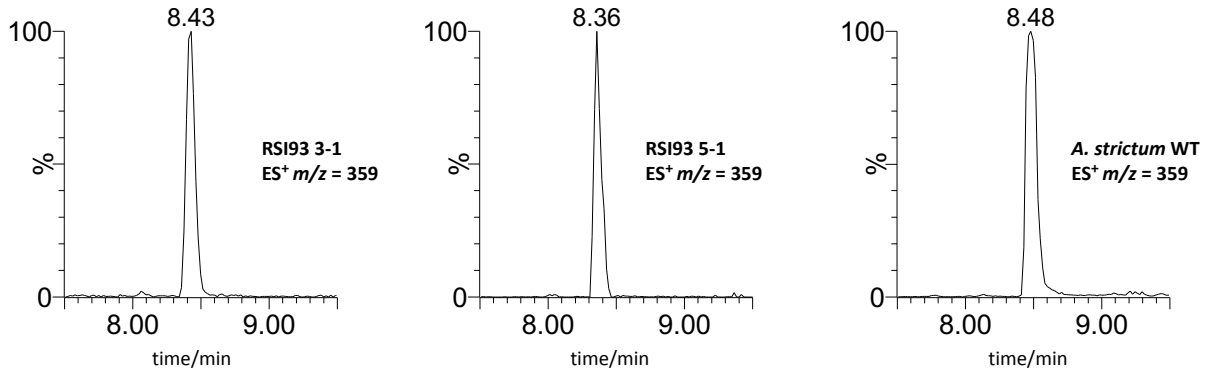


Figure 4.13 LCMS analysis of two transformants obtained with the bipartite marker strategy by analytical LCMS. Extracted ion (ES+) chromatograms at *m/z* = 359 and t_R = 8.4 min show the presence of **101** in the extracts.

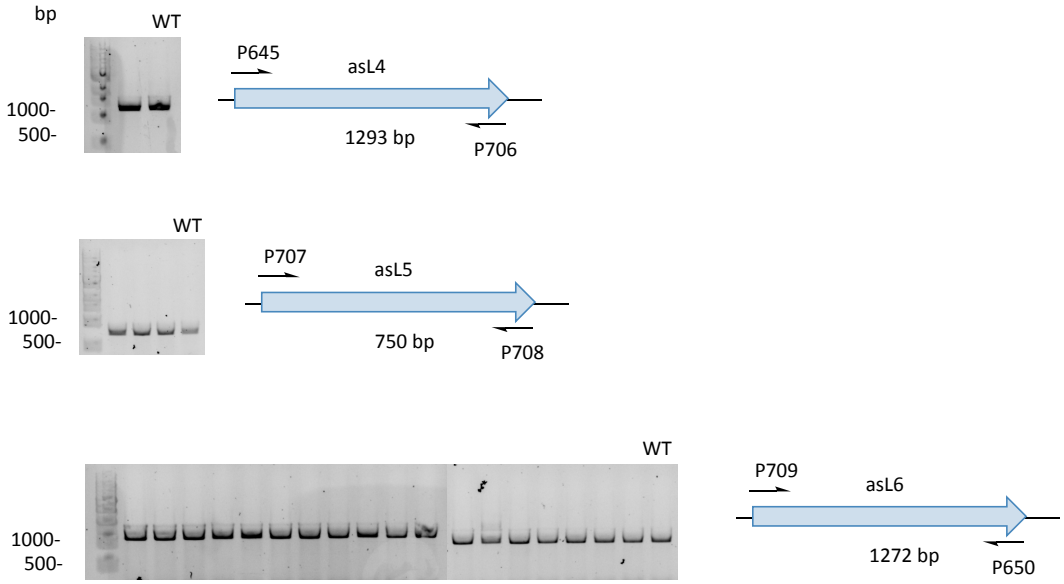


Figure 4.14 Genetic analysis of transformants obtained with bipartite marker strategy by PCR.

4.2.2 Attempted Silencing of *asL4*, *asL5* and *asL6*

The attempted targeted knockouts (*asL4*, *asL5*, *asL6*; Chapter 4.2.1) showed that *A. strictum* is proficient in ectopic integration but not in homologous recombination of DNA. This suggests that targeted gene knockouts which rely on gene disruption by homologous site specific integration are not promising in this organism. However, ectopic integration of exogenous DNA has been shown to be successful (Chapters 4.1.2, 4.21) and should hence be used to silence *asL4*, *asL5* or *asL6*. Previous attempts to silence *aspks1* by Dr Elizabeth Skellam and Dr Jack Davison using exogenous promoters

to drive transcription of different length asRNA were unsuccessful (Chapter 4.3).[104,105] However, in these experiments heterologous promoters were used and it was not clear if they promoted effective gene expression in *A. strictum*. The genome sequence of *A. strictum* enabled the use of an endogenous promoter to drive the expression of asRNA. For this experiment the promoter of the strongly expressed primary metabolism gene *gpdA* (encoding glyceraldehyde-3-phosphate dehydrogenase A) was selected.

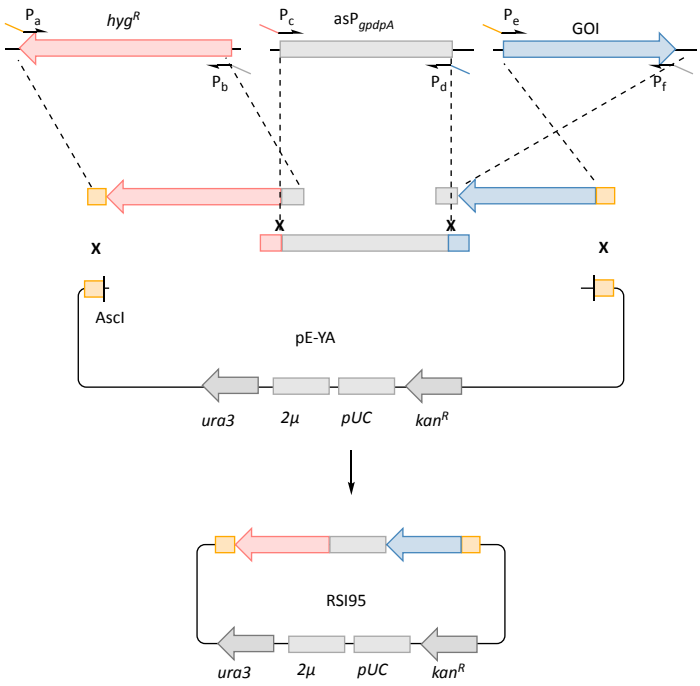


Figure 4.15 Assembly of silencing cassettes for *asL4*, *asL5* and *asL6* in pE-YA by yeast homologous recombination (HR).

Assembly of the silencing vectors containing *as*RNA of the GOI downstream of *A. strictum* P_{gdpA} (*as*P_{gdpA}) and next to a hygromycin resistance cassette was achieved by *in vivo* homologous recombination in *S. cerevisiae* (RSI95 2: *asL4*, RSI95 4: *asL5* or RSI95 6: *asL6*). DNA fragments for assembly were amplified with the denoted oligonucleotides P_a–P_f by PCR (Table 7.8 for used oligonucleotides and Chapter 4.1.4 for experimental details).

The silencing vectors for *asL4*, *asL5* or *asL6* (RSI95 2, RSI95 4 and RSI95 6) were directly used for PEG/$CaCl_2$ mediated protoplast vector transformation of *A. strictum*. In total, 25 transformants targeting *asL4*, 16 targeting *asL5* and 12 targeting *asL6* were obtained from two rounds of transformations (RSI106, RSI101, Table 9.5). For chemical analysis 11 *asL4*, 5 *asL5* and 8 *asL6* transformants were selected, grown under **101** producing conditions and extracted following a large scale (50 mL) protocol. Analytical LCMS analysis and interpretation of the chromatograms did not show any distinct changes in the production of secondary metabolites (Figure 4.16).

To confirm the vector integration two transformants of each targeted gene (*asL4*: RSI101B2-5 and B2-6, *asL5*: RSI101B4-3 and B4-4, *asL6*, RSI106 6-1 and 6-2) were selected for gDNA isolation. Transformants were genetically analysed by PCR using oligonucleotides P686+P_a and the respective gDNA templates (Figure 4.17). A ~2800 bp (*asL4*, *asL6*) or a ~2300 bp (*asL5*) DNA fragment were amplified from gDNA of both (*asL4*, *asL5*) and from one (*asL6*) tested transformant, but not from *A. strictum* WT gDNA (Figure 4.17). This confirms the ectopic integration of the silencing vector and the presence of *as*GOI downstream of *as*P_{gpdA} in all but one (RSI106 6-2) of the tranformants transformants (Figure 4.17). Further PCR analysis for the integration of the hygromycin resistance cassette using P419+420, resulted in the amplification of ~600 bp for all gDNA templates, but not for *A. strictum* WT gDNA. This indicates that upon integration of the silencing construct in transformant RSI106 6-2 (targeting *asL6*) *as*P_{gdpA} was separated from *as*GOI.

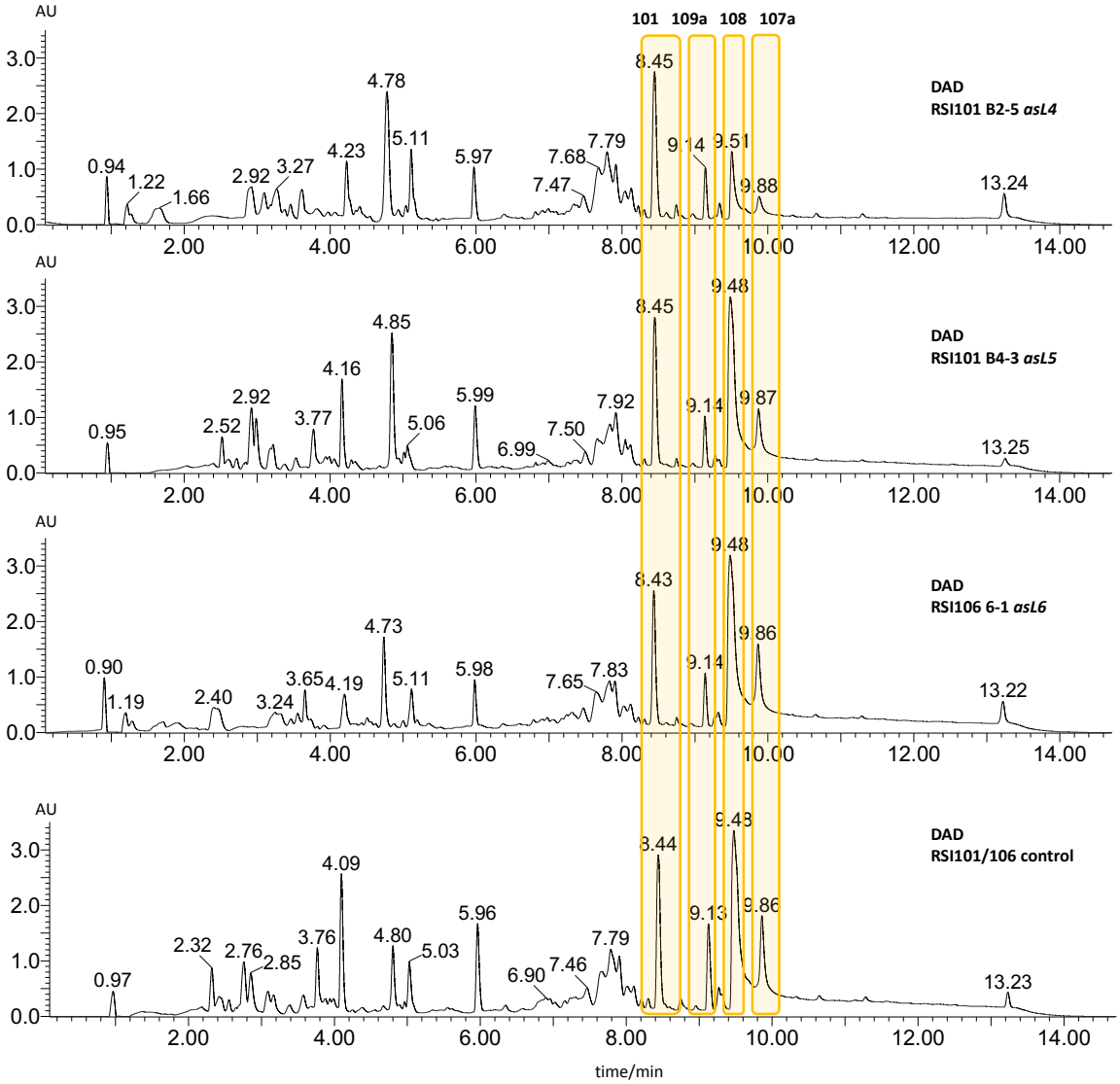


Figure 4.16 Chemical analysis of *A. strictum* transformants obtained through transformation of silencing vector (RSI95) by analytical LCMS (Gradient A1). DAD of selected transformants RSI101 B2-5, 101 B4-3 and 106 6-1.

Although genetic analysis confirmed the correct integration of the silencing construct in 5 out of 6 transformants tested, the secondary metabolite production did not show distinct changes. These findings suggest that either the attempted gene silencing in *A. strictum* is unsuccessful or that the targeted genes of the *aspks1* BGC are not involved in **101** biosynthesis. However, the *in silico* evidence (Chapter 3.2) strongly suggests that the *aspks1* BGC is responsible for **101** biosynthesis.

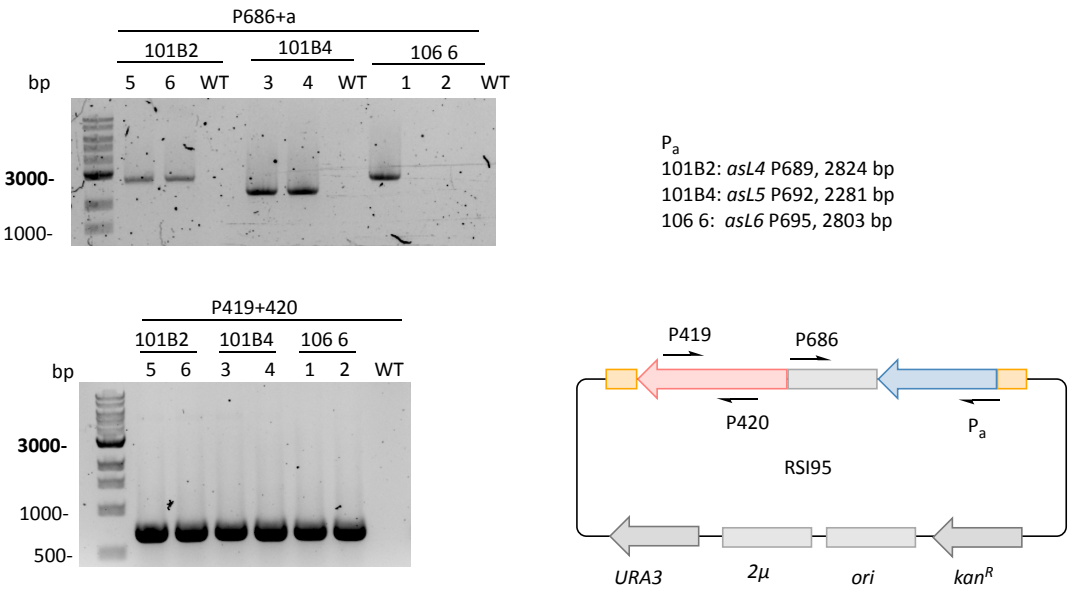


Figure 4.17 Genetic analysis of transformants obtained with transformation of silencing vectors (RSI95) by PCR.

4.3 *aspks1* Knockout

All previous unsuccessful attempts in our group to link the *aspks1* BGC and xenovulene A **101** production targeted the *aspks1* gene. Traditional KO as well as RNA induced silencing experiments by Dr Kate Harley,[127] Dr Elizabeth Skellam[105] and Dr Jack Davison[104] showed no loss or significant decrease in xenovulene A **101** production. For the disruption of *aspks1* a vector (AsPKS1KO-2, Figure 4.18A) containing a KO cassette with the hygromycin marker hyg^R was constructed by Dr Kate Harley and transformed by Dr Elizabeth Skellam. In a total of 14 transformations 84 transformants were generated, extracted and analysed for their secondary metabolite production by LCMS. All showed production of **101** and the genetic analysis by Southern Blots showed non-disrupted *aspks1* (Figure 4.18). Additional gene silencing experiments were conducted by Dr Elizabeth Skellam and Dr Jack Davison using different promoters and antisense RNA (AsSil and pUC18+*hph*, pAsSilPKS, pTAex3-*hph*-aspks1AS; Figure 4.18B-D), but none of the 42 analysed transformants showed a significant decrease in xenovulene A **101** production.

The development of the bipartite marker strategy[189] as well as the implementation of a CRISPR/Cas9[192,198] system for ascomycetes provided two new approaches which should be used to generate targeted gene knockouts of *aspks1*. Although the KO of *aspks1* proved difficult in previous experiments it was selected as targeted gene as its disruption would give clear evidence that the BGC is involved in **101** biosynthesis.

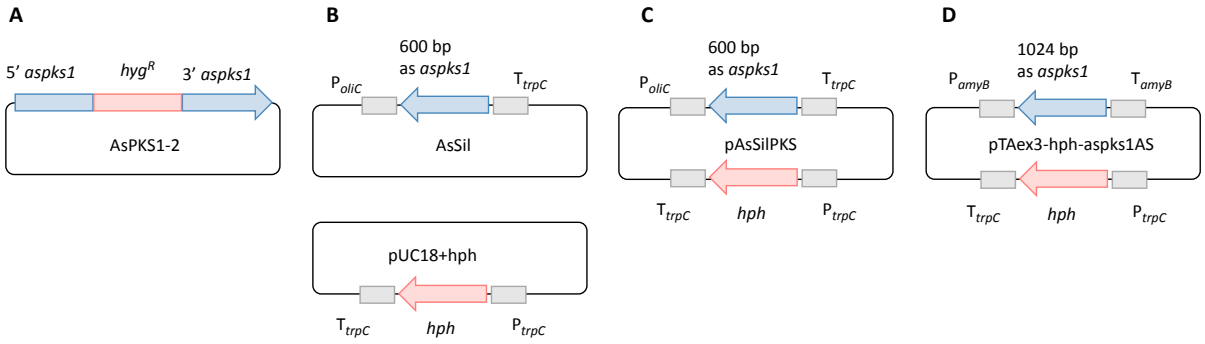


Figure 4.18 Vectors used to transform *A. strictum* in previous experiments **A**, traditional KO of *aspks1*, **B-D**, silencing attempts with different length of asRNA and promoters.

4.3.1 Attempted Gene Editing of *aspks1* by CRISPR/Cas9

During the time of this work several strategies to apply CRISPR/Cas9 technology in filamentous fungi have been published.[192,198,204] Amongst the first is a single vector based method developed by Nødvig *et al*. This strategy was selected for application in *A. strictum* due to facile cloning and transformation methodology.[198]

A single fungal CRISPR vector (pFC332) is used to introduce a codon optimized *cas9* with an *SV40* nuclear localization signal and a gRNA cassette (Figure 4.19) into the fungal host in one transformation step. The gene encoding the Cas9-SV40 fusion protein is expressed under the control of the *A. nidulans* constitutive promoter and terminator P_{tef1}/T_{tef1}. Stable integration and expression of *cas9* might cause off target effects. This is prevented through a nucleotide sequence that induces autonomous maintaining of plasmids in *Aspergillus* (AMA1).[205,206] The gRNA cassette (Figure 4.19 and 4.20) to target individual genes can be introduced by uracil-specific excision reagent cloning (USER™ Enzyme, New England Biolabs) into a *Pac*I/*Nt.Bbv*CI site.[207]

Short RNA without a cap structure and poly(A) tail, such as gRNA, are transcribed by RNA polymerase III in other CRISPR/Cas9 systems. However, RNA Polymerase III promoters are poorly defined in filamentous fungi and hence Nødvig *et al.* embedded the gRNA in a longer transcript which is transcribed by RNA Polymerase II under the control of the constitutive P_{gpdA}/T_{trpC} (Figure 4.20). This transcript encodes two ribozyme sequences, a 5' hammerhead (HH) and a 3' hepatitis delta virus (HDV) (Figure 4.20, black). The autocatalytic RNA cleaving activity of these

ribozymes frees the gRNA in the fungal host. The gene targeting gRNA leads the Cas9-SV40 fusion protein to the targeted sequence by nucleotide base pairing (Figure 4.20, blue and red).[198,208]

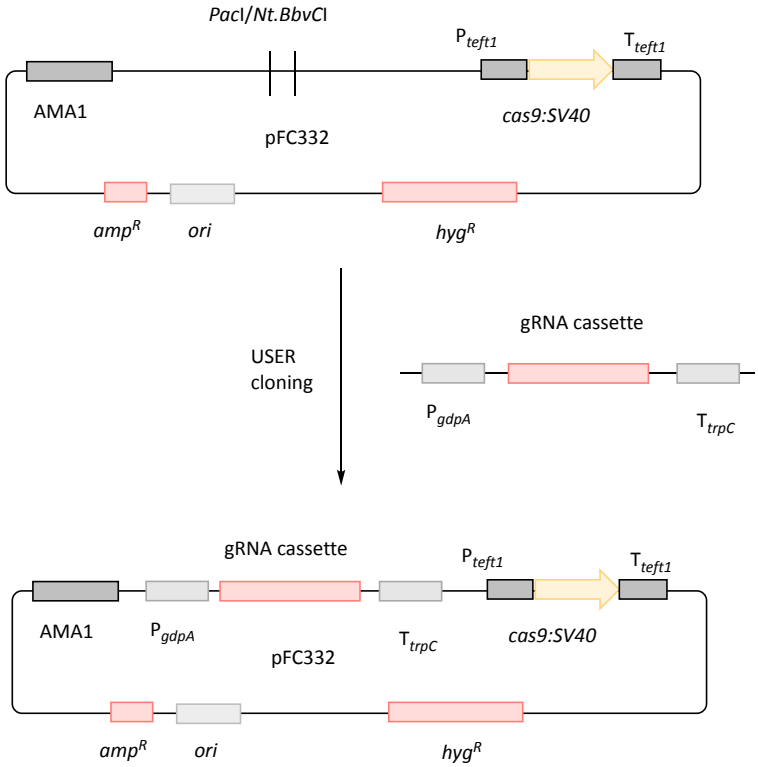


Figure 4.19 Fungal CRSIPR/Cas9 plasmid pFC332.

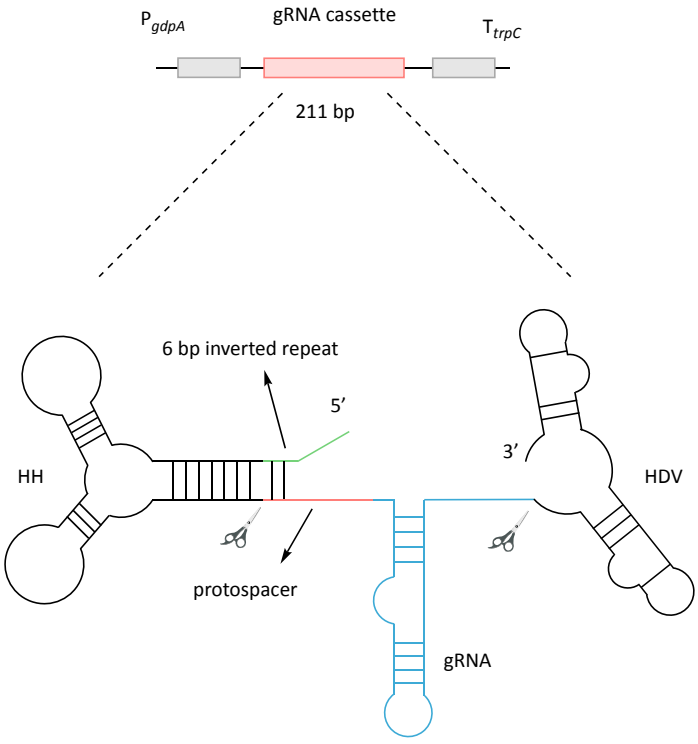


Figure 4.20 gRNA cassette design used for the CRISPR/Cas9 system developed by Nodvig *et al*.

The individual gRNA sequence as well as the 6 bp inverted repeat (Figure 4.20, green) for HH cleavage are introduced through oligonucleotides in PCR (P_a and P_b) using pFC334 as template (Figure 4.21). pFC334 is a vector containing a gRNA cassette with a protospacer (Figure 4.20, red) to target *yA* (p-diphenol oxidase) in *A. nidulans*, but can be used as a PCR template for any gRNA cassette. The two amplified DNA fragments (F1 andF2) can then be directionally fused into the fungal CRISPR (pFC) vector by USER cloning (Figure 4.21 and 4.25).

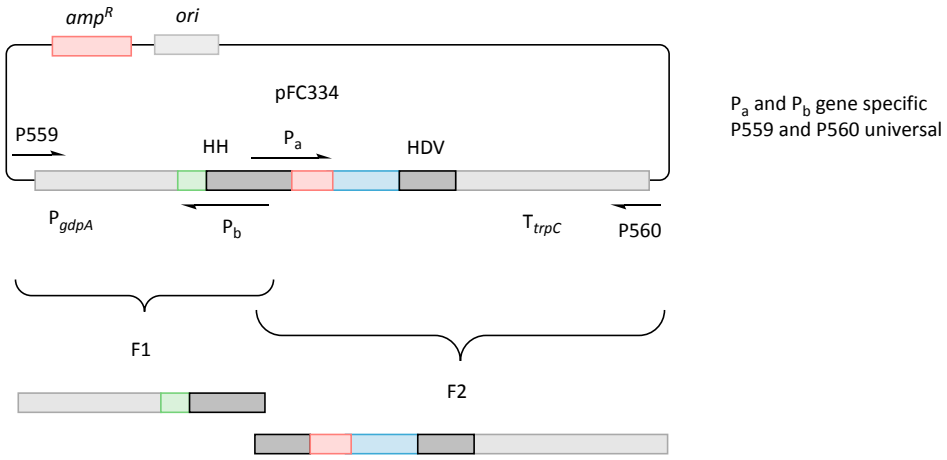


Figure 4.21 Amplification of gRNA cassette fragments F1 and F2 from pFC334 for USER cloning.

To target *aspks1* the CRIPSR/Cas9 plasmid pFC332 was sub-cloned with a gRNA cassette to induce a DSB within *aspks1* after transformation of the vector into *A. strictum*. Repair of the DSB should lead to an insertion or deletion at the cut DNA site through non-homologous end joining and result in a frameshift within the *aspks1* gene (Figure 4.22).

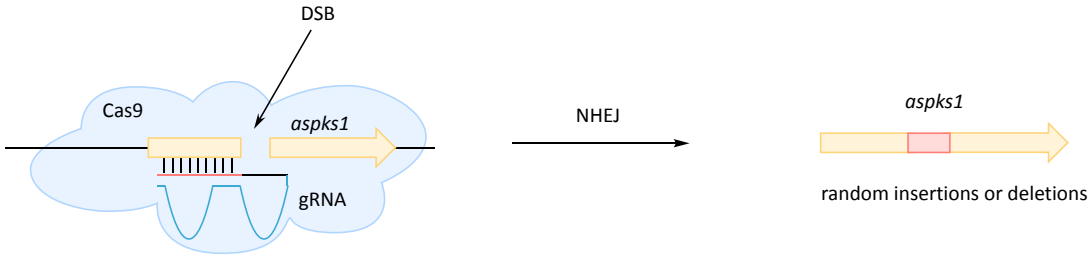


Figure 4.22 CRISPR/Cas9 induced targeted DSB and repair by non-homologous end joining (NHEJ) resulting in random insertions or deletions within *aspks1*.

Therefore, the gRNA cassette targeting *aspks1* was amplified in two individual PCRs using pFC334 as template and oligonucleotides P559+562 (F1) and P561+560 (F2) (Figure 4.20 and 4.25). USER cloning was used to fuse *Pac*I/*Nt.Bbv*CI hydrolysed pFC332, F1 (544 bp) and F2 (424 bp) in a directional manner. Transformants were screened with colony PCR and positives were confirmed by sequencing (Eurofins, Ebersberg).

The protospacer was selected on exon 1 of *aspks1* adjacent to a 5'-TGG PAM motif (Figure 4.23). It was chosen to be close to the 5' end of *aspks1* so the intended frameshift would affect the function of the entire translated protein.

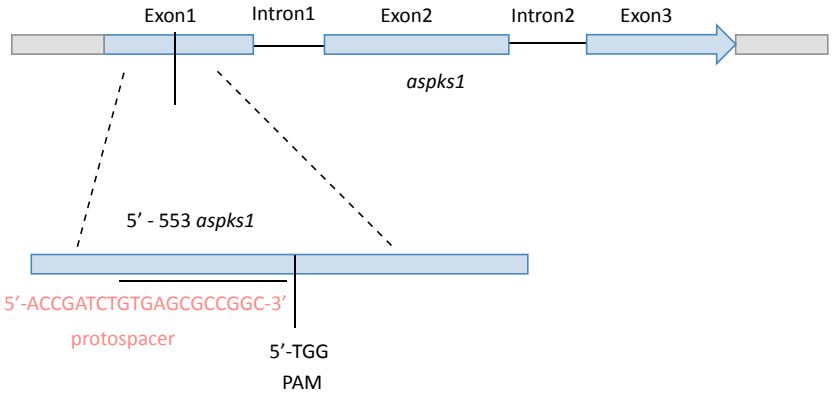


Figure 4.23 Protospacer position on exon1 directly upstream to 5'-TGG PAM at *aspks1*.

The obtained plasmid (RSI59B2) was used to transform *A. strictum* by $CaCl_2$/PEG mediated protoplast transformation. This yielded 45 colonies from two rounds of transformations. 15 colonies were selected for analysis of secondary metabolite production under **101** producing conditions (ASPM) in liquid fermentation. Cell free cultures were extracted following a small scale protocol (1 mL) and analysed by LCMS. The chromatograms showed production of **101,** deduced by UV absorption, mass and t_R in all screened transformants (Figure 4.24). Further genetic analysis was not carried out, as the production of **101** clearly indicated that the applied CRISPR/Cas9 system in combination with the selected gRNA is not effective in *A. strictum*.

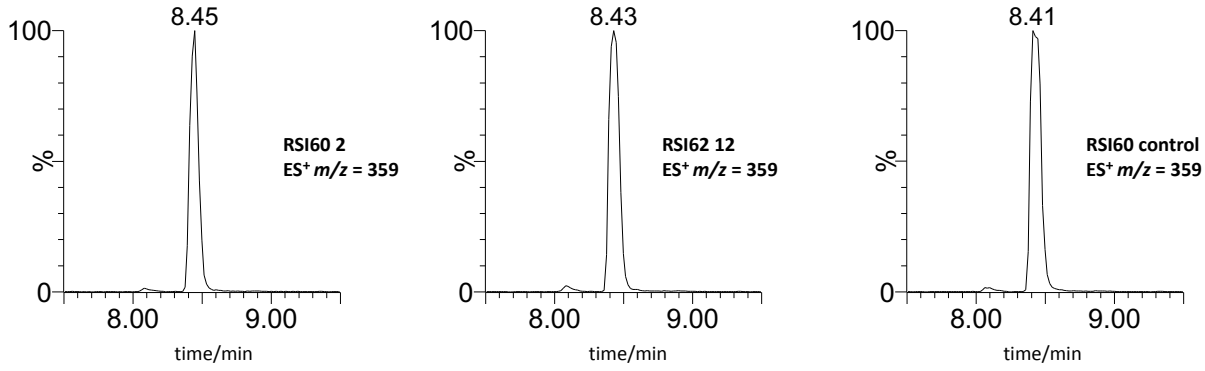


Figure 4.24 Chemical analysis of transformants obtained through transformation of *A. strictum* with the CRIPSR plasmid targeting *aspks1* (RSI59B2) on analytical LCMS. Extracted ion (ES^+) chromatograms at *m/z* = 359 and t_R = 8.4 min show the presence of **101** in the extracts.

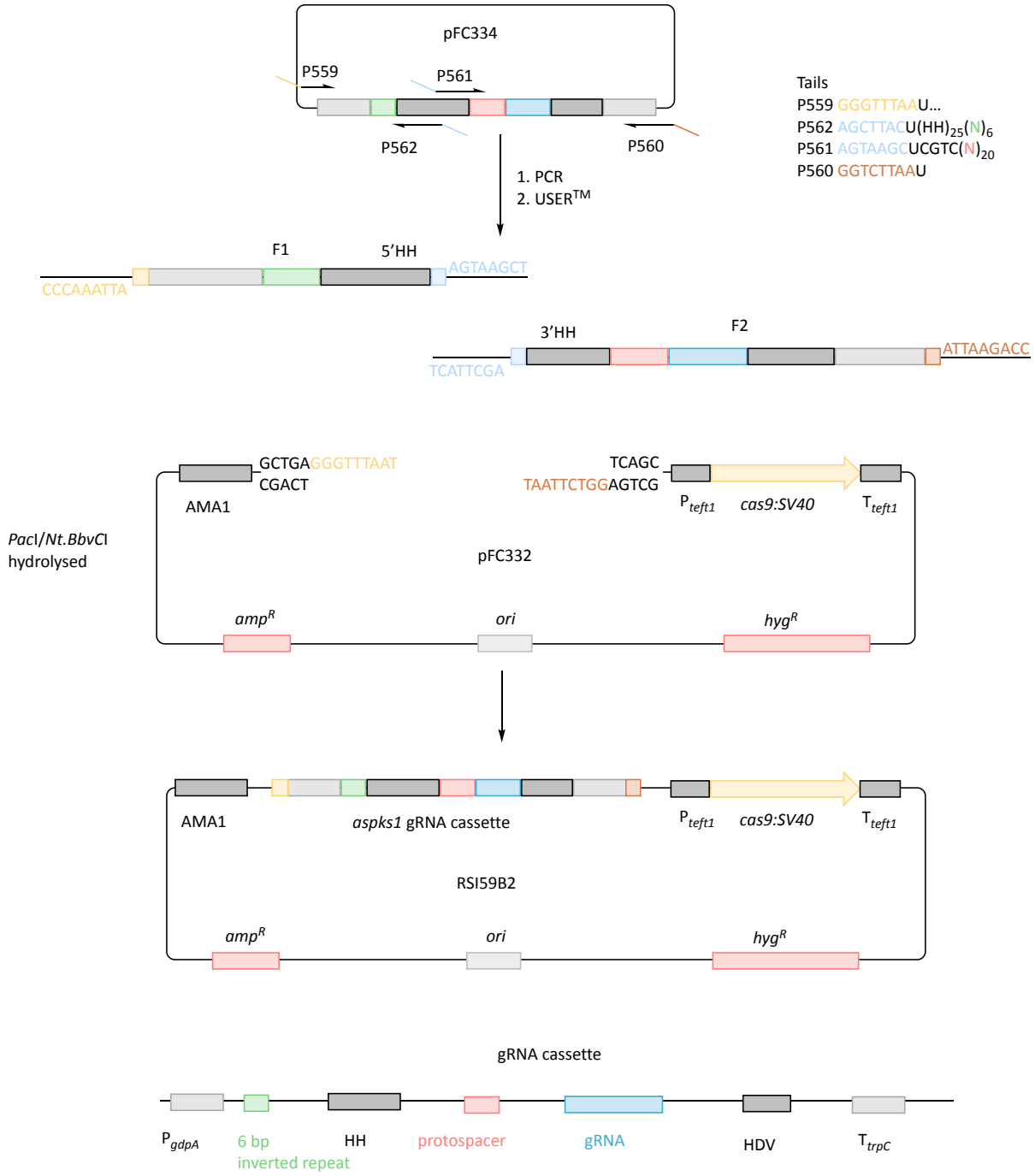


Figure 4.25 gRNA cloning strategy and used oligonucleotides for USER cloning. Uracil containing tails were introduced trough tails at the oligonucleotides (P559-562). Overlapping and complementing sequences are coloured accordingly.

4.3.2 *aspks1* Disruption by Bipartite Marker Strategy

As a second approach to disrupt *aspks1* in *A. strictum* the bipartite marker strategy was applied.[189] The 5' and 3' bipartite substrate were assembled in two separate constructs (RSI21A1, RSI21B5) by *in vivo* homologous recombination in *S. cerevisiae* (Figure 4.26, Chapter 4.1.4 for experimental details). The *aspks1* targeting sequences and the two hygromycin resistance cassette containing fragments were amplified by PCR using the indicated oligonucleotides (Figure 4.26).

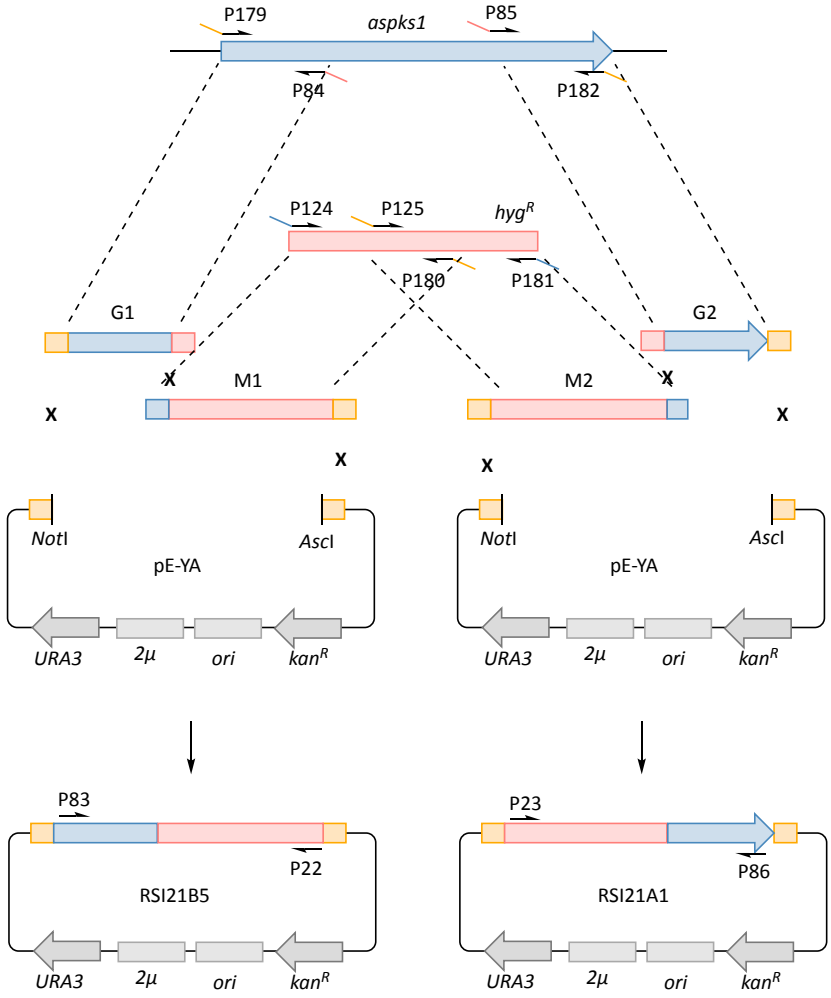


Figure 4.26 Assembly of bipartite substrates for *aspks1* in pE-YA by yeast HR.

The bipartite substrates were PCR amplified with P83+22 and P23+86 using RSI21B5/RSI21A1 as PCR template. The DNA fragments were purified and used directly for $CaCl_2$/PEG mediated protoplast transformation of *A. strictum*. In 4 rounds of transformations 43 colonies were generated and analysed for their secondary metabolite production (Appendix Table 9.6).

The 43 putative Δ*aspks1* colonies were grown under xenovulene A **101** producing conditions (ASPM), cell free culture broth was extracted following small scale extraction (1 mL) and analysed by analytical LCMS. The chromatograms showed production of **101**, deduced by UV absorption, mass and t_R in all but one (RSI58 1-1) analysed transformants. The remaining cell free culture broth (49 mL) of RSI58 1-1 was extracted and LCMS analysis confirmed the lack of **101** or related compounds **107-108** (Figure 4.27).

Genomic DNA of transformants RSI58 1-1 – 1-9 was isolated and used for further genetic characterisation by PCR. Correct integration of the KO cassette was investigated at the 5' and 3' site of *aspks1*. For this purpose, five different sets of oligonucleotides were designed (Figure 4.28A):

1) P323+322 amplifying a 2400 bp fragment at 5' *aspks1* from non-disrupted *A. strictum* gDNA only. With gDNA of RSI58 1-1 – 1-9 as template a 2400 bp DNA fragment was amplified for all used templates, but RSI58 1-1, indicating a 5' disruption of *aspks1* (Figure 4.28A1).

2) P323+618 amplifying a 1291 bp fragment at 5' *aspks1* if the 5' bipartite substrate integrated correctly. With gDNA of RSI58 1-1 as template a 1291 bp DNA fragment was amplified, but not from *A. strictum* WT gDNA, suggesting correct integration of the 5' bipartite substrate (Figure 4.28B2).

3) P419+420 amplifying a 710 bp fragment only if the hygromycin cassette (*hyg*^{R}) recombined accordingly. With gDNA of RSI58 1-1 a 710 bp DNA fragment was amplified, confirming the recombination of the hygromycin selection marker (Figure 4.28B3).

4) P81+557 amplifying a 2047 bp fragment at 3' *aspks1* from non-disrupted *A. strictum* gDNA only. With gDNA of RSI58 1-1 as well as *A. strictum* WT a 2047 bp DNA fragment was amplified, indicating no disruption of 3' *aspks1* (Figure 4.28B4).

5) P619+557 amplifying a 1652 bp fragment at 3' *aspsk1* if the 3' bipartite substrate integrated correctly. With gDNA of RSI58 1-1 no DNA fragment of the expected size (1652 bp) was amplified, but instead a ~4 kb band was detected. The ~4 kb DNA fragment was not amplified from *A. strictum* WT gDNA confirming an unpredicted integration of the 3' bipartite substrate. Any attempts to sequence the ~4 kb DNA fragment were unsuccessful (Figure 4.28B5).

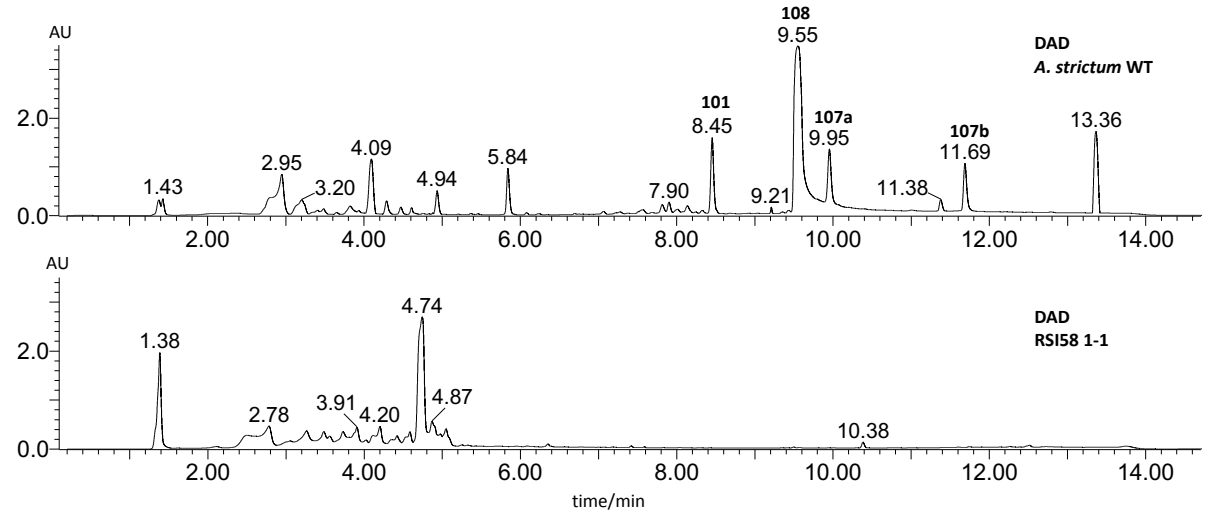


Figure 4.27 Chemical analysis of RSI58 1-1 by analytical LCMS. DAD of RSI58 1-1 and *A. strictum* WT extracts.

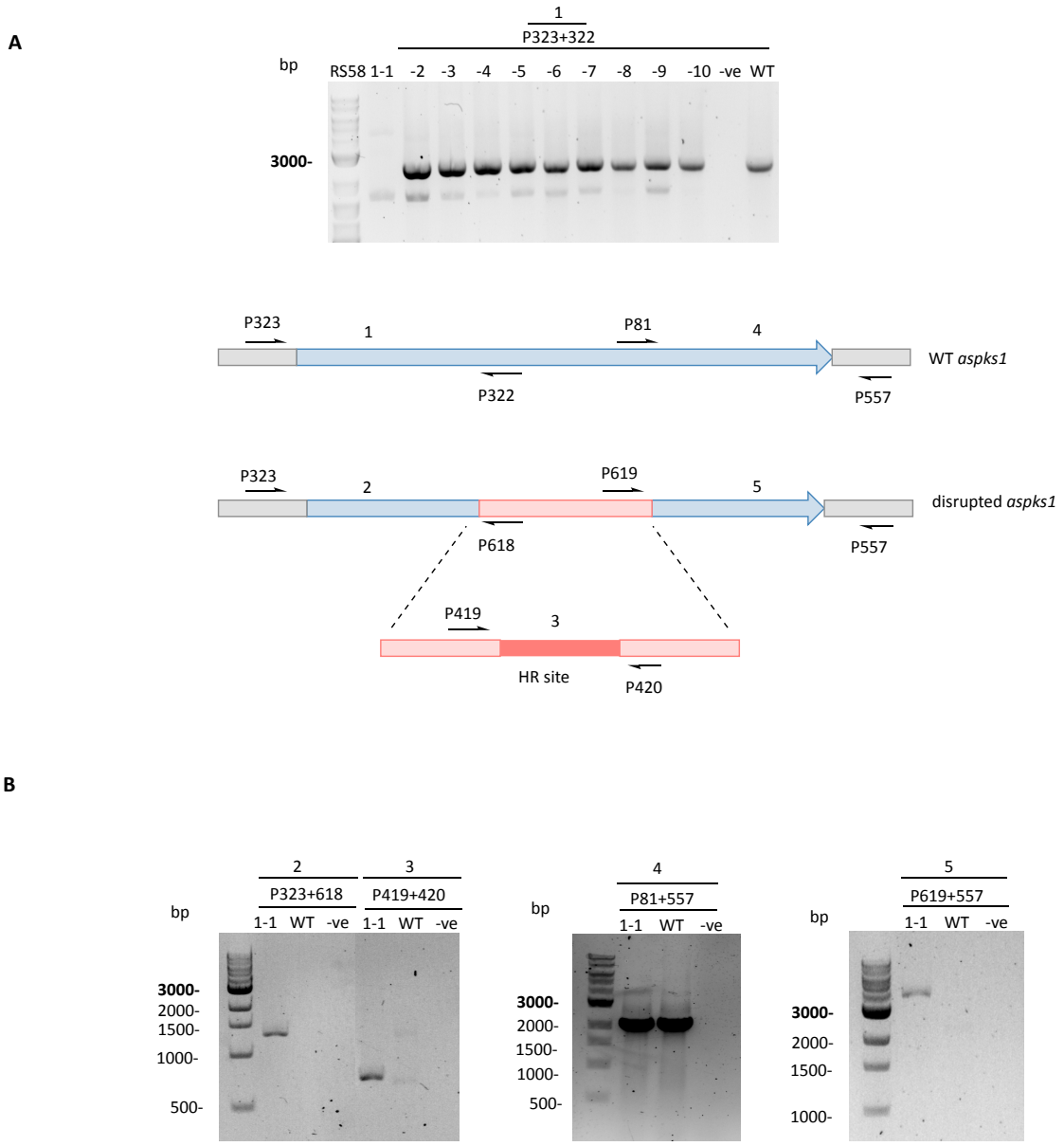


Figure 4.28 Genetic analysis of **A**, RSI58 1-1 – 1-9 and **B**, RSI58 1-1 by PCR. Negative controls are labelled with –ve.

Further analysis of 5' substrate integration in RSI58 1-1 with P83+322 (Figure 4.29) resulted in the amplification of a 2163 bp DNA fragment. The amplification of a DNA fragment with this oligonucleotide combination was unexpected, as correct integration of the 5' KO substrate should have removed the annealing site of P322 (Figure 4.29). At this point it was hypothesized that 5' and 3' substrates were recombined by homologous recombination first and random integration at the 5' site of *aspks1* led to the displacement of the promoter region and the inability to initiate downstream transcription at that site. Oligonucleotides P826+899 were designed to amplify a 139 bp DNA fragment spanning 3' *aspks1* (introduced through bipartite substrates) and endogenous 5' *aspks1* (Figure 4.29). Gel electrophoresis of the DNA fragment amplified by PCR using P826+899 suggested it to be approximately 600 bp larger than predicted (Figure 4.29).

Sequencing confirmed an insertion of 590 bp between 3' bipartite substrate and 5' endogenous *aspks1*. By sequence alignment the inserted fragment was found to be the 590 bp homologous recombination site (red) of the hygromycin resistance gene *hph* with three T→C point mutations (Figure 4.29).

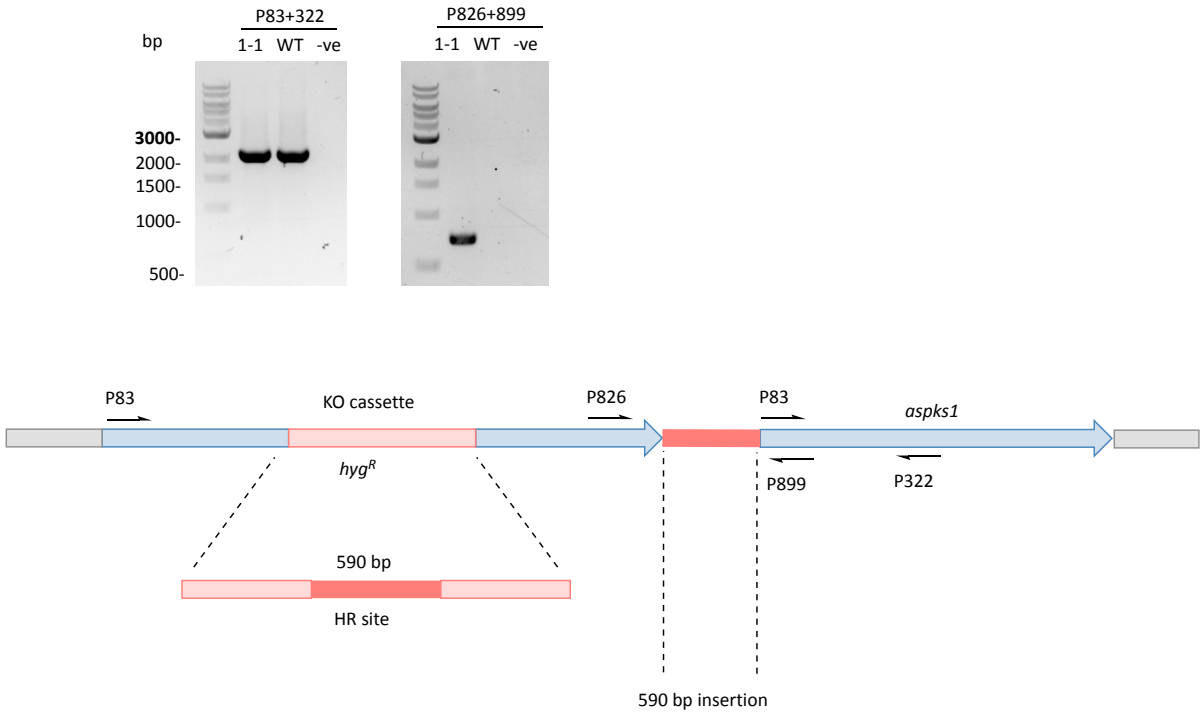


Figure 4.29 Further genetic analysis RSI58 1-1 by PCR. Negative controls are labelled with –ve.

Genetic analysis of RSI58 1-1 showed that *aspks1* was displaced from its promoter by the recombined bipartite substrates and a 590 bp insertion. This probably led to the inability to express *aspks1* which would explain the lack of xenovulene A **101** and related compounds **107**-**109** production. This strongly suggests that *aspks1* is involved in **101** biosynthesis (Figures 4.27 and 4.28). However, as the targeted gene disruption did not occur as intended it is also posssible that disruption of *aspks1* and deletion of **101** production is still a coincidence. The bipartite substrates could have integrated elsewhere in the genome and disrupted a random gene at the same time. A complementation experiment, where *aspks1* is reintroduced in the Δ*aspks1* strain (RSI58 1-1) should re-establish **101** biosynthesis and give final confirmation that *aspks1* encodes the protein that produces the polyketide precursor for xenovulene A **101**.

Therefore, *aspks1* was inserted into the fungal expression vector pTYG-*gen*^{R} (Figure 4.10) by LR recombination (RSI74, Figure 4.31, for experimental details Chapter 5.1.1). The vector contains the geneticin resistance marker *gen*^{R} and a fungal expression cassette. Gene expression under the control of the inducible promoter P_{amyB}, which was shown to induce downstream transcription in *A. strictum* (Chapter 4.1.2).[104] $CaCl_2$/PEG mediated protoplast transformation (RSI79) of the

A. strictum Δ*aspks1* (RSI58 1-1) strain with the vector RSI74 (*aspks1* in pTYG-*gen*^{R}) resulted in one *A. strictum* Δ*aspks1*+*aspks1* (RSI79 6) colony growing on tertiary plates with hygromycin B and geneticin. The transformant was grown under **101** producing conditions, extracted following a large scale protocol (50 mL) and analysed by analytical LCMS. The chromatogram showed low production of **101**, deduced by UV absorption, mass and t_R = 8.4 min. HRMS ($C_{22}H_{31}O_4$, calculated 359.2214, found 359.2222 $[M-H]^+$) confirmed the production of **101**.

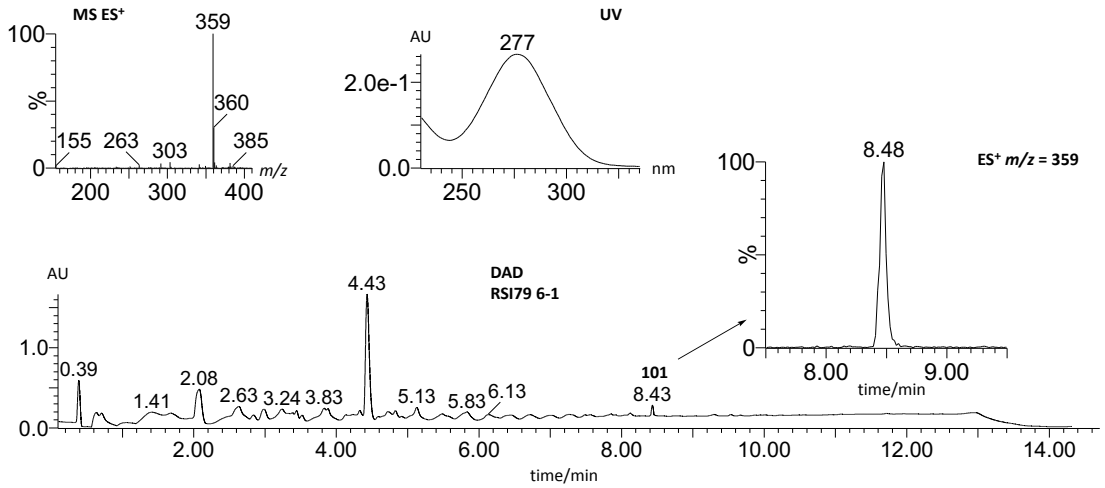


Figure 4.30 Chemical analysis of RSI79 6-1 by analytical LCMS. Extracted ion (ES^+) chromatogram at *m/z* = 359, t_R = 8.4 min and DAD of RSI79 6-1 show the presence of **101** in the extracts.

Genomic DNA of *A. strictum* Δ*aspks1*+*aspks1* (RSI79 6) was isolated and analysed for the integration of *aspks1* (RSI74) by PCR. As controls also gDNA of *A. strictum* WT and *A. strictum* Δ*aspks1* (RSI58 1-1) were analysed using four pairs of oligonucleotides (Figure 4.31):

1) P660+659 amplifying 543 bp of the geneticin resistance gene if the vector integrated successfully. Only with gDNA of *A. strictum* Δ*aspks1*+*aspks1* (RSI79 6) as PCR template a 543 bp fragment was amplified. This confirms the integration of the geneticin resistance (*gen*^{R}) (Figure 4.31 1).

2) P147+322 amplifying a 2279 bp of *aspks1* under the control of P_{amyB}. Only with *A. strictum* Δ*aspks1*+*aspks1* (RSI79 6) gDNA as PCR template a 2279 bp was amplified. This shows that *aspks1* downstream P_{amyB} is present after transformation of RSI74 (*aspks1* in pTYG-*gen*^{R}) (Figure 4.31 2).

3) P323+618 amplifying a 1291 bp fragment at 5' *aspks1* if the 5' bipartite substrate integrated correctly to prove that the Δ*aspks1* strain RSI58 1-1 was further transformed. A 1291 bp DNA fragment was amplified from *A. strictum* Δ*aspks1*+*aspks1* (RSI79 6) and *A. strictum* Δ*aspks1* (RSI58 1-1) gDNA. This confirms the disruption of endogenous *aspks1* in *A. strictum* Δ*aspks1*+*aspks1* (RSI79 6) (Figure 4.31 3).

4) P323+322 amplifying a 2400 bp fragment at 5' *aspks1* from *A. strictum* WT gDNA. This PCR is designed to identify a possible *A. strictum* WT contamination. Only a 2400 bp DNA fragment from *A. strictum* WT gDNA was amplified. This confirms the lack of endogenous *aspks1* and thus contamination free transformants *A. strictum* Δ*aspks1*+*aspks1* (RSI79 6) and *A. strictum* Δ*aspks1* (RSI58 1-1) (Figure 4.31 4).

Genetic and chemical analysis of *aspks1* deletion *A. strictum* Δ*aspks1* (RSI58 1-1) and complementation *A. strictum* Δ*aspks1*+*aspks1* (RSI79 6) strains confirmed that *aspks1*, encoding the NR-PKS MOS, is essential for the biosynthesis of meroterpenoids **101, 107**-**109**. However, the unexpected integration of the bipartite substrates, which resulted in the displacement of *aspks1* from its promoter, indicates that the applied bipartite method is not efficient in the non-homologous end joining proficient *A. strictum* strain.

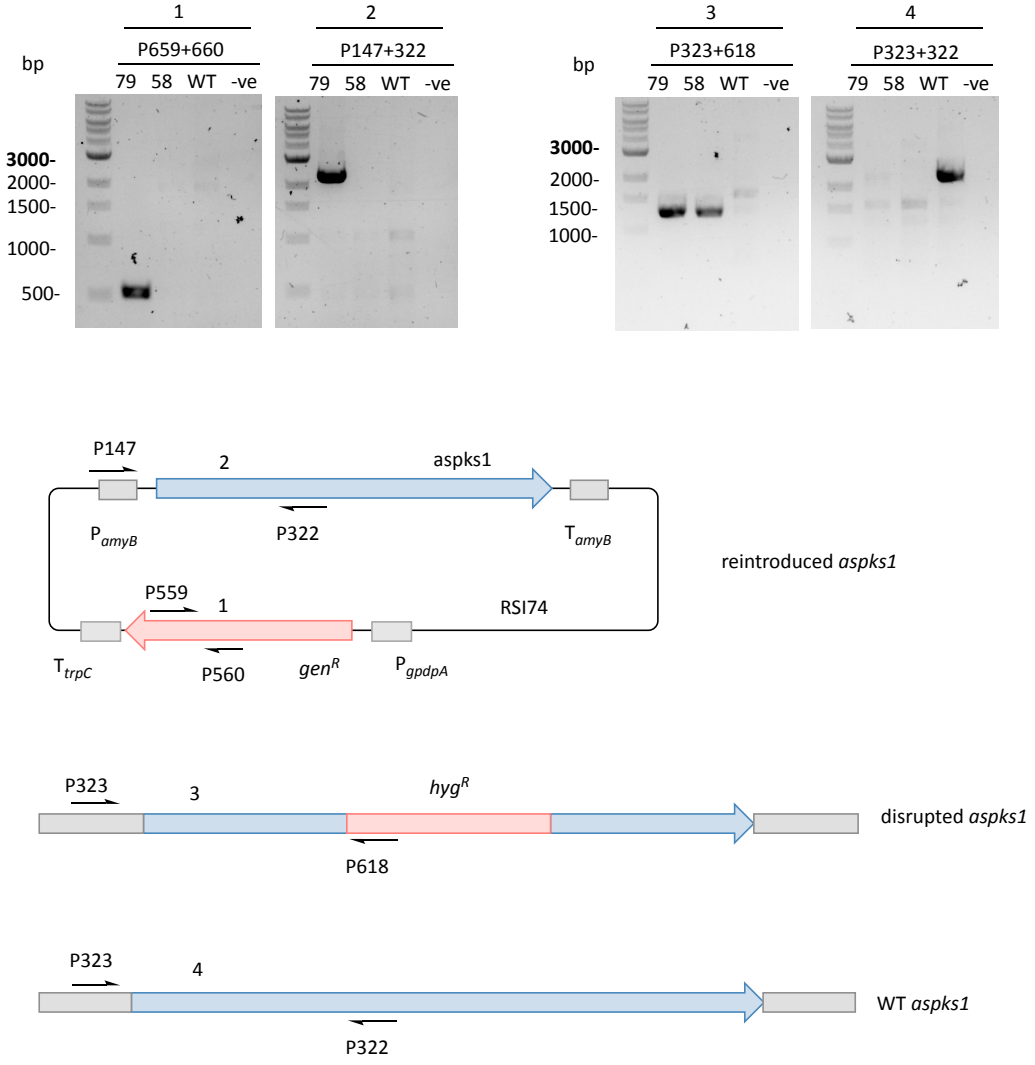


Figure 4.31 Genetic analysis RSI79 1-6 by PCR. Negative controls are labelled with –ve.

4.4 Discussion and Outlook

Targeted gene KO and isolation of accumulating intermediates is a powerful tool to elucidate the biosynthesis of natural products. Xenovulene A **101** belongs to a rare family of secondary metabolites from ascomycetes and the *A. strictum* producer is the only strain with an available sequenced genome that can be exploited for genetic manipulation. Abundant protoplasts formation as well as successful vector transformation provided a basis for further genetic manipulations.[104,105,127]

The previously used traditional KO strategy[104,105,127] as well as the bipartite strategy[189] applied in this work showed that targeted gene disruption in *A. strictum* at the *aspks1* BGC genetic locus is very inefficient. Of 90 transformants generated with bipartite substrates targeting different genes (*aspks1*, *asL4*, *asL5* and *asL6*) none was shown to have integrated the KO cassette as intended. Reconstitution of the selection gene (*hph* or *nptII*) proved that *A. strictum* is able to homologously recombine DNA fragments, but appears to be barely deploying it. However, with RSI58 1-1 (*A. strictum* Δ*aspks1*), a transformant unable to produce **101** was generated by transformation of the bipartite substrates targeting *aspks1*. Genetic analysis by PCR revealed an unexpected integration/insertion event (Figure 4.29) that separated the promoter from *aspks1* and resulted in loss of mRNA transcription. A subsequent complementation experiment with *aspks1* (RSI79, *A. strictum* Δ*aspks1+aspks1*) confirmed that the *aspks1* BGC is involved in xenovulene A **101** production. These findings are further supported by heterologous gene expression of the *aspks1* BGC genes which led to the isolation of **101** from *A. oryzae* NSAR1 (Chapter 5). This work demonstrated that although the bipartite strategy improves targeted gene disruptions in a variety of other fungi,[62,209,210] it is not applicable for the *aspks1* BGC genetic locus of *A. strictum*. KO of the non-homologous end joining involved genes *ku*80 or *ku*70 in *A. strictum* could improve gene targeting as ectopic integration should be less efficient in the mutant strain. However, although Δ*ku*80 mutants were shown to improve gene targeting rates in *N. crassa* and *M. ruber*,[188,211]it did not improve gene targeting at the *ACE1* genetic locus in *M. grisea*.[202]

With attempted gene silencing and CRISPR/Cas9 two strategies independent of homologous recombination were deployed to explore further gene functions of the *aspks1* BGC. The LCMS analysis of 24 putative silenced and 15 transformants generated with CRISPR/Cas9 gene editing showed that both methods did not have the intended effect in *A. strictum*.

The gene silencing strategy was based on previous experiments where silencing of *tenA* or *tenB* was achieved through the introduction of full length *tenA* or *tenB* antisense RNA in *B. bassiana*.[172] Full length *as*GOI (*asL4*, *asL5* or *asL6)* was assembled 3' of ~1.5 kb *A. strictum* endogenous promoter

asP_{gpdpA}. It was assumed that 1.5 kb promoter length are sufficient, however it has not been analysed whether the promoter sequence is long enough to drive gene expression. In addition, expressing full length asRNA was shown to be successful in *B. bassiana*,[172] but in several experiments hairpin RNA (hpRNA)[212–214] with targeting sequence was found to be more efficient in stable gene silencing.

In the model organism *N. crassa* two silencing pathways have been described: meiotic silencing during the sexual cycle and quelling during the vegetative cycle. For each pathways one set of RNA silencing genes is conserved (Table 4.2).[194] A BLASTp[103] analysis of the *A. strictum* genome using the corresponding RdRP, Ago and Dicer protein sequences as search templates identified *A. strictum* homologues (Table 4.2). Dicer DCL-2 involved in meiotic silencing was found to have one homologue in *A. strictum*, but not its quelling homologue DCL-1. A similar phenomenon was observed in *A. nidulans*, where only one set of homologues of RNA silencing proteins (involved in quelling) is encoded. Upon KO of those genes no changes in viability was observed. This suggested that RNA silencing proteins have a 'non-housekeeping' function such as defence against viral dsRNA rather than being involved in gene regulation and genome stability.[215] If this was also the case for *A. strictum* a less abundant expression of genes involved in silencing would be expected, which could explain the observed inefficiency of gene silencing. To analyse the efficiency of RNA silencing in *A. strictum* a reporter gene such as *egfp* could be used to visualize the silencing effect. A construct where antisense *egfp* (as*egfp)* is located downstream of asP_{gpdpA} could be inserted into *A. strictum* transformed with pTH-GS-*egfp* (RSI16, Chapter 4.1.2). Subsequent microscopic analysis should give evidence about the functionality of gene silencing in *A. strictum* and the corresponding vectors designed in this work.

Table 4.2 Overview of protein homologues predicted to be involved in RNA silencing in *A. strictum*.

N crassa		**NCBI Accession**	***A. strictum***	**E value**
meiotic silencing	QDE-1 (RdRP)	XM_953954.2	Asg2314	5e-147
	QDE-2 (Ago)	AF217760.1	Asg4864	0
	DCL-2 (Dicer)	XM_958445.3	Asg5496	0
quelling	SAD-1 (RdRP)	XM_959155.3	Asg850	0
	SMS-2 (Ago)	AF508210.1	Asg4581	0
	DCL-1 (Dicer)	XM_956805.2	Asg5496	4e-40

For gene editing with CRISPR/Cas9[198] a vector based strategy optimised for the model organism *Aspergillus* was selected due to facile cloning and transformation. Despite successful

transformation of *A. strictum*, no Δ*aspks1* strain could be generated with this method. With the used vector (RSI59B2), *cas9* is transcribed from *A. nidulans* constitutive P_{teft}, but this promoter has not previously been tested for its activity in *A. strictum*. To analyse the ability of P_{teft} to drive transcription in *A. strictum cas9* should be expressed alongside *egfp*. The resulting fusion protein could be visualized by microscopy and the cellular localisation could be determined. If the gene expression and nuclear protein localisation can be confirmed, different protospacer sequences should be tested to improve gene targeting. In this work the protospacer was selected manually. In order to reduce off target effects, which have been observed with CRISPR/Cas9,[216] a software package, such as sgRNAcas[217], should be used. At the time when the experiment was set up for *A. strictum* only two CRIPSR/Cas9 strategies were available for filamentous fungi, but in the meanwhile further strategies have been published and should be considered for future experiments.[218]

5 Heterologous Expression of the *A. strictum aspks1* BGC in *A. oryzae*

Heterologous expression of single genes or entire BGC is used to analyse biosynthetic routes of microorganisms where targeted gene disruption or transformation methods are ineffective. Beyond that it offers many engineering possibilities such as co-expression of genes from different biosynthetic pathways to produce novel hybride secondary metabolites. Gene expression in *Escherichia coli* and recombinant protein purification allows the study of reactions *in vitro*.

The arginine auxotroph *A. oryzae* M2-3 strain has previously been used successfully to express single genes or entire BGC. For example, the four genes which encode the tenellin **189** BGC make up the first complete fungal BGC to be expressed heterologously.[219] The biosynthesis of more complex fungal metabolites, such as the meroterpenoids pyripyropene **93** and anditomin **190**, where up to eleven genes have been expressed simultaneously, have also been elucidated by heterologous expression studies (Figure 5.1).[77,220]

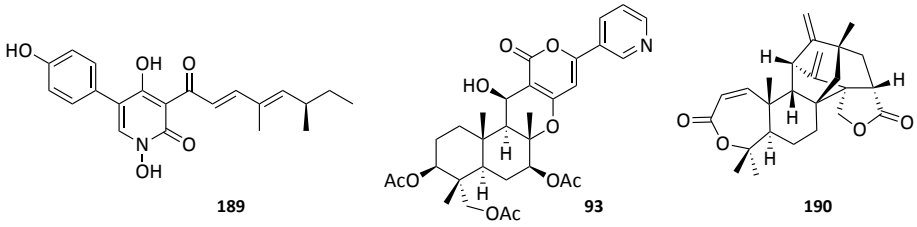


Figure 5.1 Tenellin **189**, pyripyropene **93** and anditomin **190**.

For the reconstitution of the desired biosynthetic pathway the genes can be directly amplified by PCR from g- or cDNA of the producing ascomycete. For closely related species intron splicing is usually not problematic, however when expressing genes from distantly related fungi intron free clones are advantageous. Due to similar codon usage between the fungal strains no optimisation is generally required for gene expression in *A. oryzae*. However, to activate transcription it is important to insert the GOI downstream of an *A. oryzae* promoter.

5.1 Yeast Recombination, Gateway Cloning and Transformation of *A. oryzae*

The quadruple auxotroph *A. oryzae* NSAR1 is deficient in arginine (Δ*argB*), methionine (*sC*$^-$), adenine (*adeA*$^-$) and ammonium (*niaD*$^-$) metabolism which enables its use as a host-vector system.[221] The four auxotrophies can be exploited for transformation of fungal expression vectors by complementation with the four different selection markers (*argB*, *sC*, *adeA*, *niaD*). Additionally,

the natural sensitivity of *A. oryzae* NSAR1 towards the antibiotics bleomycin and the herbicide glufosinate (*ble*^{R}, *bar*) enables the use of two more selection markers.

The pTYGS fungal expression vectors, which are available with selection markers for all auxotrophies and sensitivities (*argB*, *sC*, *adeA*, *niaD*, *ble*^{R}, *bar*), contain four fungal promoter/terminator (P/T$_{amyB}$, P/T$_{adh}$, P/T$_{gpdA}$, P/T$_{eno}$) gene cloning sites (Figure 5.2A). In theory, the simultaneous expression of 24 genes is possible.[201] Three of these cloning sites include an *Asc*I restriction sequence (P/T$_{adh}$, P/T$_{gpdA}$, P/T$_{eno}$) and are designed for *in vivo* homologous recombination by *S. cerevisiae*. For this, the desired vector is cut into three parts and co-transformed with up to three PCR amplified DNA fragments (genes) (Figure 5.2B). *S. cerevisiae* assembles the vector and genes by homologous recombination in the desired order as long as the DNA fragments include at least 30 bp unique homologous overlap sequence. These can be introduced by PCR through tails at the designed oligonucleotides.[180] The vector also contains *ura3*, encoding orotidine 5'-phopsphate decarboxylase, and the *2µ ori* for selection and propagation in uracil deficient *S. cerevisiae* strains. Colonies able to grow on uracil- and uridine-free medium contain vector DNA which can be isolated and further transformed into *E. coli* for facile screening of colonies by PCR. For selection and propagation in *E. coli* the vector contains an appropriate *ori* (*colE1*) and confers resistance to chloramphenicol (*cam*^{R}) as well as ampicillin (*amp*^{R}).

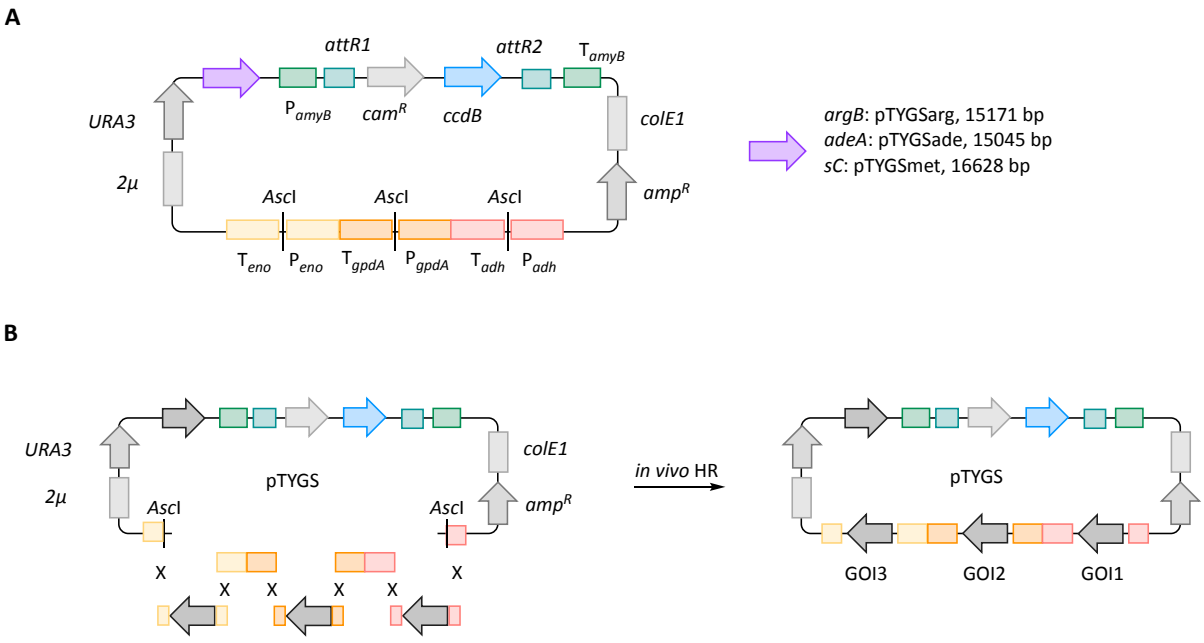


Figure 5.2 Fungal expression vectors used in this work. **A**, pTYGS vector family. **B**, Cloning of up to 3 genes into pTYGS vectors by *in vivo* HR in *S. cerevisiae*.

The P/T$_{amyB}$ cloning site contains recombination sequences (*attR1* and *attR2*) flanking a gateway cassette (*ccdB* gene, encoding the CcdB killer protein, and chloramphenicol resistance, *cam*^{R}) for use with an *in vitro* gateway cloning kit (LR recombinase, Invitrogen). The kit consists of an enzyme

mixture containing bacteriophage λ integrase proteins (integrase, integration host factor, excisionase) which interchange DNA flanked by *attR* sites with DNA flanked by *attL* sites (Figure 5.2A).[222] *attL* recombination sequences are included in the *E. coli-S. cerevisiae* shuttle vector pE-YA (entry clone), which also contains ORI (*2μ*, *pUC*) and selection markers (*ura3*, *kan*^{R}) for selection and propagation in *S. cerevisiae* and *E. coli* (Figure 5.3A).

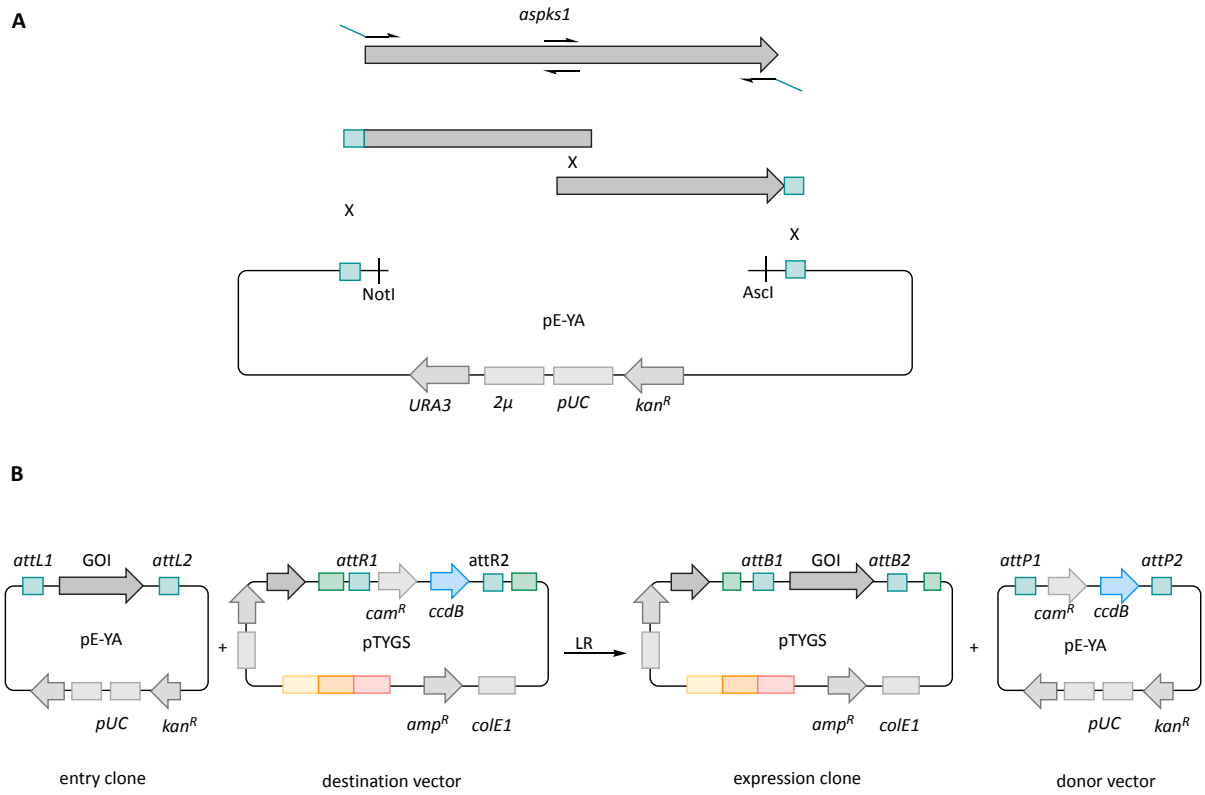


Figure 5.3 Gateway cloning. **A**, Assembly of large synthase gene (*aspks1*) by *in vivo* HR recombination (*S. cerevisiae*) in the gateway entry clone pE-YA. **B**, Gateway cloning by LR recombination between pE-YA (entry clone) and pTYGS (destination vector) yielding in expression clone (pTYGS) and donor vector (pE-YA).

pE-YA can be used for HR by *S. cerevisiae* to assemble large synthase genes, such as *pks*, in multiple fragments, but it is also useful for small genes (Figure 5.3A). By incubation of the entry vector (pE-YA) containing a GOI and a vector of the pTYGS family (destination vector) with the LR recombinase enzyme mixture two new plasmids are formed: the expression clone containing the GOI and the donor vector containing the gateway cassette (*ccdB*, *cam*^{R}) (Figure 5.2B). Double selection is used to isolate the expression clone only. Destination vectors with gateway cassette have to be propagated in *E. coli ccdB* survival cells, which are resistant to the *ccdB* gene product. Thus for selection of the expression clones only, an *E. coli* TOP10 strain as well as a different antibiotic (ampicillin) is used.

5.1.1 Gene Selection and Vector Construction

KO of *aspks1* (Chapter 4.3.2) in *A. strictum* confirmed the involvement of the *aspks1* BGC in **101** biosynthesis. Differential expression analysis (Chapter 3.2) showed the upregulation of 15 genes (*asL1-7*, *asR1-7* and *aspks1*). For heterologous co-expression 12 genes were selected, omitting only transporters (*asR1*, *asL7*) and transcription factors (*asR3*) (Figure 5.4).

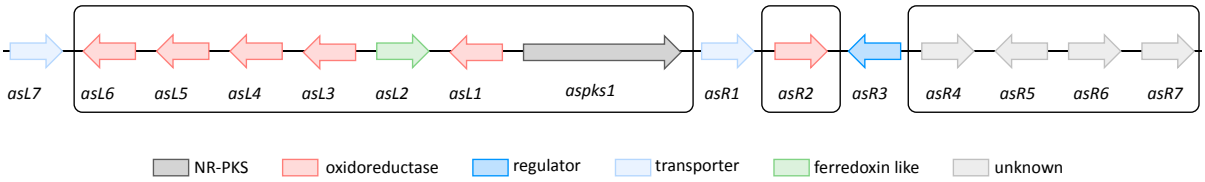


Figure 5.4 *aspks1* BGC, framed genes were selected for heterologous expression in *A. oryzae* NSAR1.

The biosynthetic hypothesis suggested that early tropolone formation is similar to *T. stipitatus* (Chapter 1.3.2). Thus, for co-expression of *aspks1*, *asL1*, *asL3* and *asR2*, the genes were cloned into the pTYGSarg vector. For sequential co-expression of these genes, different combinations (*aspks1*, *aspks1*+*asL1*, *aspks1*+as*L*1+*asL3* and *aspks1*+*asL1*+*asL3*+*asR2*) were also cloned into pTYGSarg. The three additional oxidoreductases encoded by *asL4*, *asL5* and *asL6*, which could be involved in the ring contraction, were cloned into the pTYGSmet vector. The genes encoding proteins of unknown function (*asL2*, *asR4*, *asR5* and *asR6*) were cloned into the pTYGSade vector. For co-expression of different sub-sets of genes, combinations of two genes each (*asL4*, *asL5* and *asL6* or *asL2*, *asR4* and *asR6*) were cloned into pTYGSmet and pTYGSade (Figure 5.5).

All pE-YA entry and pTYGS(arg/ade/met) vectors for expression of the *aspks1* BGC were constructed by *in vivo* homologous recombination in *S. cerevisiae* and *in vitro* recombination (Figure 5.5). Genes were amplified by PCR using *A. strictum* c- or gDNA as template with denoted oligonucleotides (Table 7.9), which were designed to introduce the 30 bp homologous overlaps (Chapter 5.1 and 4.1.4 for experimental details). Only the 8190 bp *aspks1* gene was amplified in two fragments (each ~4100 bp) with 70 bp overlap from *A. strictum* gDNA and assembled in the gateway entry vector pE-YA (Figure 5.3A).

Genes amplified from cDNA (*asL1*, *asL3* and *asR2*) were initially sub-cloned in pCR2.1 by the TOPO TA cloning kit (Invitrogen). Correct splicing was confirmed by sequencing (Eurofins, Ebersberg) and the vector used as template in further PCR reactions. Assembly of 1 or 2 genes in pTYGS expression vectors was achieved by designing oligonucleotides with an overlap to T_{adh}/P_{gpdA} or T_{gpdA}/P_{eno} and P_{eno}, which resulted in the removal of the 2480 bp T_{adh}/P_{gpdA} or the respective 800 bp T_{gpdA}/P_{eno} DNA fragment. All entry and expression clones used for LR recombination and subsequent *A. oryzae* NSAR1 transformation are displayed in Figure 5.5. Additional vectors which

were constructed in this work but have not yet been used for fungal transformations are displayed in Appendix Figure 9.41.

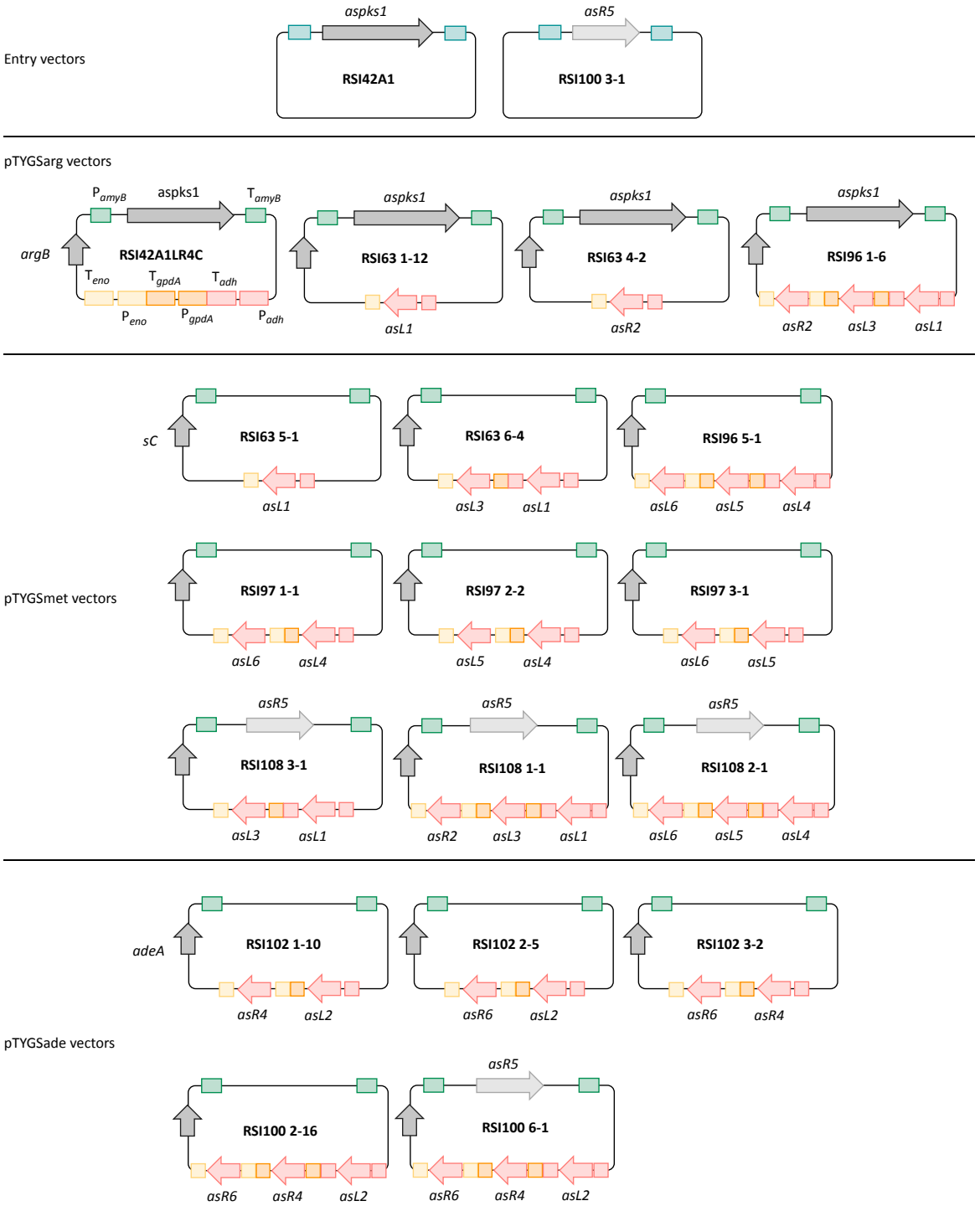


Figure 5.5 Overview of constructed in this work used gateway entry and fungal expression (pTYGSarg, met, ade) vectors.

5.1.2 *A. oryzae* Transformation and Gene Combinations

Different combinations of the constructed vectors (Chapter 5.1.1) were used to transform *A. oryzae* NSAR1 by a $CaCl_2$/PEG mediated protoplast protocol. Usually two or three fungal expression

vectors, with different selection markers (pTYGSarg, -met, -ade) were co-transformed in a single transformation. The selection medium was adapted to the vectors used in the individual transformation.

In experiment RSI105, for example, the three vectors RSI96 1-6 (pTYGSarg), RSI96 5-1 (pTYGSmet) and RSI100 2-16 (pTYGSade) were co-transformed in *A. oryzae* NSAR1. As each of these vectors complements a different auxotrophy, the protoplasts were recovered and selected on minimal medium (Chapter 7.4.3) without arginine, methionine or adenine (Table 7.1). Colonies which grew on selection medium were picked onto secondary selection plates and further growing colonies were genetically purified by streaking for single colonies on a third selection plate.

In total, 16 different combinations of fungal expression vectors were used to co-transform *A. oryzae* NSAR1. This lead to the isolation of 16 different *A. oryzae* NSAR1 strains, each expressing a different sub-set of genes of the *A. strictum aspks1* BGC (Table 5.1).

Table 5.1 Overview of *A. oryzae NSAR1* transformations with different combinations of genes. Chromatograms of grey labelled experiments are included in the Appendix Figure 9.42, as no additional information about the biosynthesis of xenovulene A **101** was gained from these combinations of genes.

Transformation ID	Transformed vector constructs vector backbone pTYGSarg, -met, -ade	Genes *as*											# colonies
		PKS	***L1***	***L2***	***L3***	***L4***	***L5***	***L6***	***R2***	***R4***	***R5***	***R6***	
RSI80 1	RSI42A1LR4C, pTYGSmet, -	x	-	-	-	-	-	-	-	-	-	-	6
RSI70 1	RSI63 1-12, RSI63 5-1, -	x	x	-	-	-	-	-	-	-	-	-	10
RSI72 2	RSI63 1-12, RSI63 6-5, -	x	x	-	x	-	-	-	-	-	-	-	6
RSI80 2	RSI63 4-2, RSI63 6-5, -	x	x	-	x	-	-	-	x	-	-	-	13
RSI98 1	RSI96 1-6, RSI96 5-1, -	x	x	-	x	x	x	x	x	-	-	-	9
RSI105	RSI96 1-6, RSI96 5-1, RSI100 2-16	x	x	x	x	x	x	x	x	x	-	x	17
RSI107	RSI96 1-6, RSI96 5-1, RSI100 6-1	x	x	x	x	x	x	x	x	x	x	x	28
RSI109	RSI96 1-6, -, RSI108 2-1	x	x	-	x	x	x	x	x	-	x	-	24
RSI110	RSI96 1-6, -, RSI100 6-1	x	x	x	x	-	-	-	x	x	x	x	14
RSI111 1	RSI96 1-6, RSI97 1-1, RSI100 6-1	x	x	x	x	x	-	x	x	x	x	x	5
RSI111 2	RSI96 1-6, RSI97 2-2, RSI100 6-1	x	x	x	x	x	x	-	x	x	x	x	8
RSI111 3	RSI96 1-6, RSI97 3-1, RSI100 6-1	x	x	x	x	-	x	x	x	x	x	x	7
RSI113	RSI96 1-6, RSI108 1-1,	x	x	-	x	-	-	-	x	-	x	-	12
RSI114 B1	RSI96 1-6, RSI108 2-1, RSI102 1-10	x	x	x	x	x	x	x	x	x	x	-	10
RSI114 B2	RSI96 1-6, RSI108 2-1, RSI102 2-5	x	x	x	x	x	x	x	x	-	x	x	7
RSI114 B3	RSI96 1-6, RSI108 2-1, RSI102 3-5	x	x	-	x	x	x	x	x	x	x	x	10

Genomic DNA of selected *A. oryzae* NSAR1 colonies was isolated and analysed by PCR for the intended integration of genes with gene specific oligonucleotides (Figure 5.6). *A. strictum* WT gDNA served as positive, and gDNA of an *A. oryzae* NSAR1 strain (RSI74 2) transformed with pTYGSarg, pTYGSmet and pTYGSade empty vectors as negative, control (Figure 5.6). Comparison of DNA fragment sizes amplified from the different *A. oryzae* NSAR1 gDNAs and *A. strictum* WT gDNA confirmed the intended integration of genes within the respective *A. oryzae* NSAR1 genome.

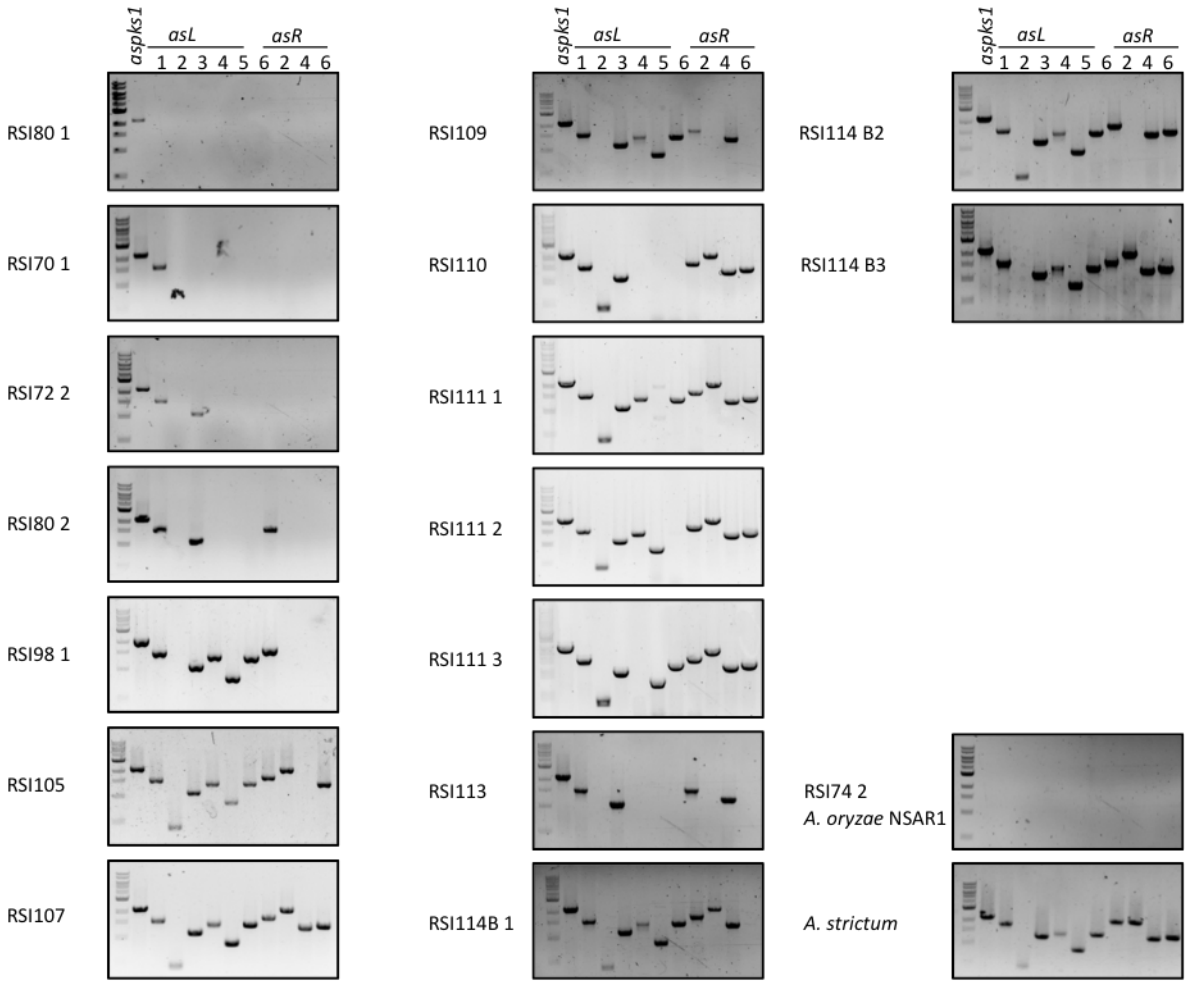


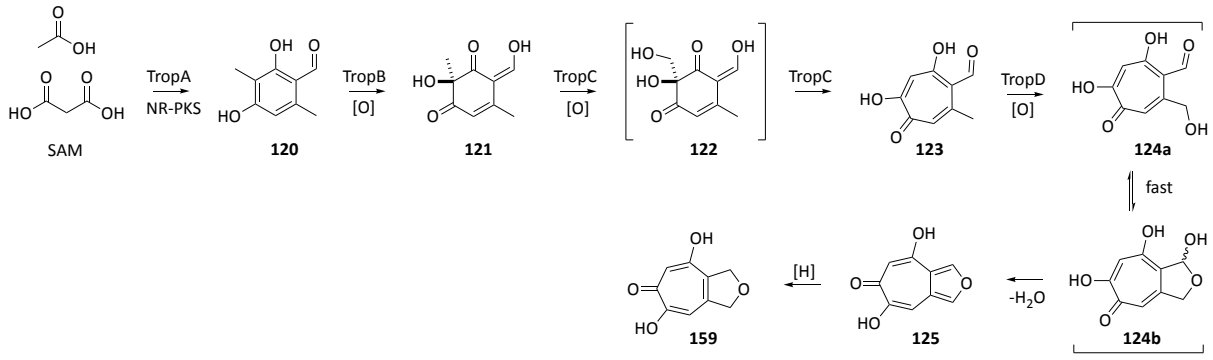
Figure 5.6 PCR analysis of *A. oryzae* NSAR1 transformants, negative control with RSI74 2 (transformed with pTYGSarg, met and ade empty vectors) and positive control with *A. strictum* WT gDNA. Oligonucleotides: *aspks1* (P83+322), *asL1* (P421+423), *asL2* (P755+756), *asL3* (P424+426), *asL4* (P645+706), *asL5* (P421+423), *asL6* (P707+708), *asR2* (P709+650), *asR4* (P757+758), *asR5* (P761+762), *asR6* (P759+760).

The secondary metabolite production of all transformants was analysed by sub-culturing the colonies in P_{amyB} induction medium (Starch M + 0.15% methionine or DPY) for 3-6 d at 28 °C and 120 rpm. The cultures were acidified with HCl (2 M), the cells homogenised and the solids removed by Buchner filtration. The cleared culture broth was extracted twice with ethyl acetate or ethyl acetate:hexane (1:1). The organics were dried over $MgSO_4$, concentrated *in vacuuo* and dissolved in methanol or acetonitrile: water (9:1) to a concentration of 10 mg/mL. The extracts were then analysed for their secondary metabolite production by LCMS.

5.2 Early steps of Xenovulene Biosynthesis: Tropolone Formation

Tropolone biosynthesis (Scheme 5.1) in *T. stipitatus* is achieved by 3 core enzymes: the NR-PKS TropA, releasing **120**; the FAD dependent monooxygenase TropB, which hydroxylates and dearomatizes **120** to **121**; and hydroxylation of **121** to **122** by the non-heme Fe^{II} dependent dioxygenase TropC followed by ring expansion to give the first tropolone **123**.[97]

In the next step the cytochrome P450 monooxygenase TropD hydroxylates **123** at *C*-9 to give the primary alcohol **124a**, which most likely is in equilibrium with its hemiacetal form **124b**. However, neither **124a** nor **124b** have been observed upon KO of *tsL2* (encoding TropD) in *T. stipitatus*, but instead the shunt products stipitafuran **125** and cordytropolone were observed.[98]



Scheme 5.1 Tropolone biosynthesis in *T. stipitatus*.

Within the *aspks1* BGC genes encoding homologous proteins were predicted (Figure 1.14 B). Previously reported heterologous expression of *aspks1* as well as co-expression of *aspks1*+*asL1* in *A. oryzae* M2-3 led to the production of **120** and **121**, respectivly.[102,104] This confirmed the homology to TropA and TropB encoding genes *tspks1* and *tsL1*. In order to express the entire *aspks1* BGC, these initial steps had to be re-established in the fungal host *A. oryzae* NSAR1. In addition, *A. strictum asL3* and *asR2* (homologues *tropC* and *tropD*) were heterologously expressed in *A. oryzae* NSAR1 for the first time.

5.2.1 Expression of *aspks1* in *A. oryzae*

Transformation of *A. oryzae* NSAR1 with *aspks1* alone yielded 6 colonies that were grown in P_{amyB} induction medium. The obtained extracts (Chapter 5.1.2) were submitted to LCMS and the chromatograms were analysed for the production of 3-methylorcinaldehyde **120**. In comparison to an *A. oryzae* NSAR1 chromatogram obtained under identical conditions, several new peaks were detected in the DAD chromatogram (Figure 5.7C) mainly eluting between t_R = 5 - 8 min.

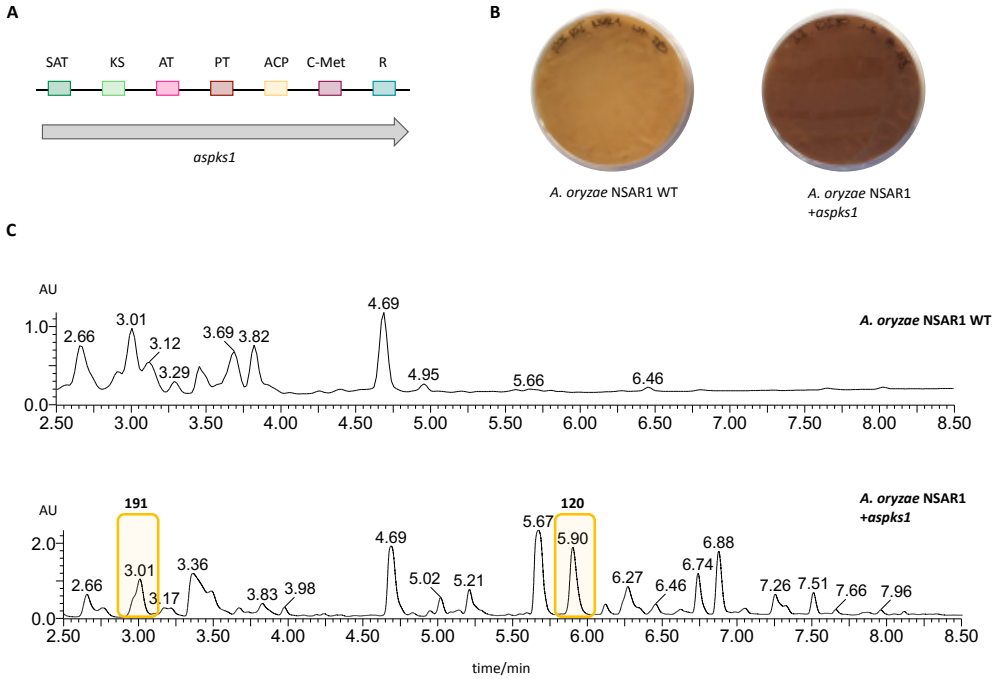


Figure 5.7 Expression of *aspks1*. **A**, Domain organisation of *aspks1* encoding the NR-PKS MOS (SAT: starter unit acyl transferase, KS: ketosynthase, AT: Acyl transferase, PT: product template, ACP: acyl carrier protein, *C*-Met: *C*-methyl transferase, R: reductive release). **B**, *A. oryzae* NSAR1 WT and *A. oryzae* NSAR1 +*aspks1* on DPY agar. **C**, DAD chromatogram of an *A. oryzae* WT and ab *A. oryzae* +*aspks1* sample measured with analytical gradient A1 on LCMS. Extracts were obtained from cultures grown in Starch M + 0.15% methionine for 6 d.

Production of a compound with the nominal mass of 166 corresponding to **120** eluting at t_R = 5.9 min was observed in all 6 transformants. The UV spectrum showed an absorption maxima at 297 nm, which is consistent with literature data for **120** (Figure 5.8).[102] A molecular formula of $C_9H_{10}O_3$ ($[M-H]^- C_9H_9O_3$ calculated 165.0552, found 165.0551) was confirmed for this peak by HRMS.

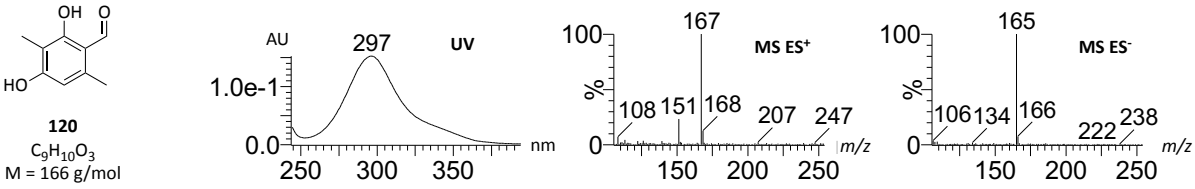


Figure 5.8 Chemical characterisation of **120** (t_R = 5.9 min) by UV and mass spectrum obtained from ES^+ and ES^-.

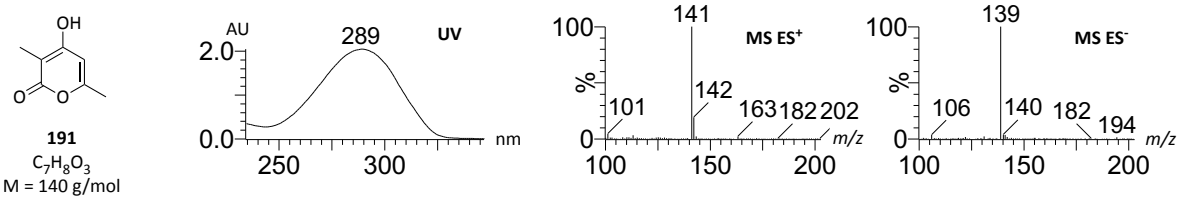


Figure 5.9 Chemical characterisation of **191** (t_R = 3.0 min) by UV and mass spectrum obtained from ES^+ and ES^-.

In addition, a compound with the nominal mass of 140 corresponding to **191**, which was previously observed upon *aspks1* expression, was identified in all chromatograms at t_R = 3.0 min (Figure 5.9).[102]

This compound (3 mg) was purified from a 1 L Starch M culture extract and the structure confirmed by ^{1}H NMR as well as HRMS ([M]Na^{+} $C_7H_8O_3Na$ calculated 163.0371, found 163.0371) (**Chapter 7.5.6** for NMR data).

5.2.2 Co-Expression of *aspks1+asL1*

Transformation of *A. oryzae* NSAR1 with *aspks1+asL1* yielded 11 colonies. For analysis of their secondary metabolite profile the transformants were sub-cultured in P_{amyB} induction medium. The obtained extracts (Chapter 5.1.2) were submitted to LCMS and the chromatograms were analysed for the production of **121** (Figure 5.10).

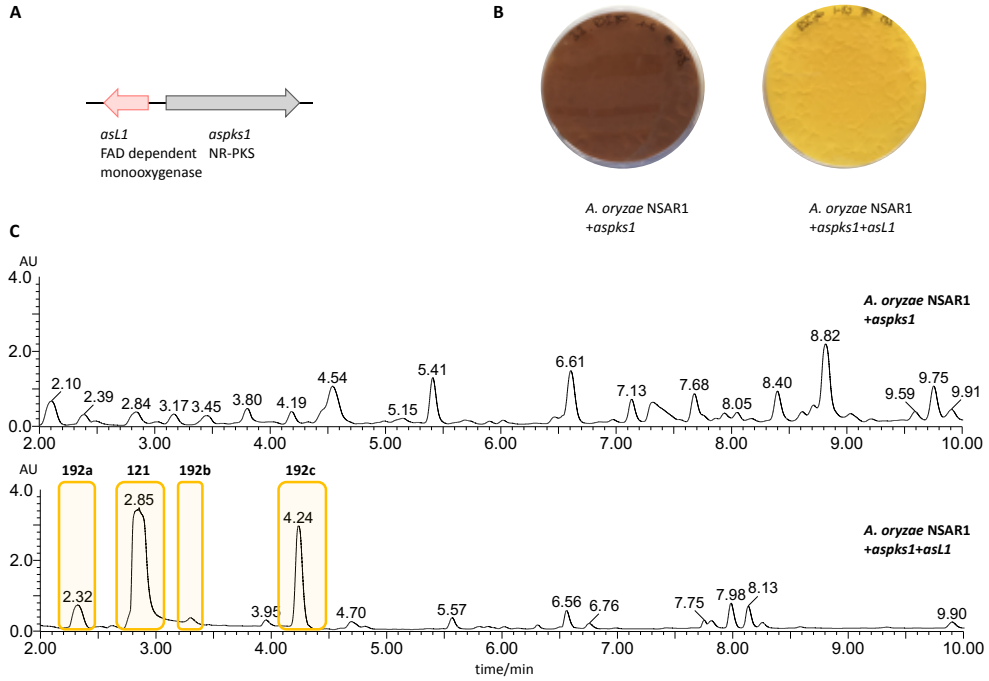


Figure 5.10 Co-expression of two genes. **A**, Co-expression of *aspks1* and *asL1*. **B**, *A. oryzae* NSAR1 +*aspks1* and *A. oryzae* NSAR1 +*aspks1+asL1* on DPY agar. **C**, DAD chromatogram of an *A. oryzae* NSAR1 +*aspks1* and an *A. oryzae* NSAR1 +*aspks1+asL1* sample measured with analytical gradient A2 on LCMS. Extracts were obtained from cultures grown in DPY for 4 d.

In ten extracts a new compound eluting at t_R = 2.9 min with a nominal mass of 182 and an UV absorption maximum of 308 nm (Figure 5.11) was detected. Purification of 15 mg as a dark brown oil from 1 L Starch M culture extract enabled subsequent ^{1}H NMR analysis (Chapter 7.5.6 for NMR data). The NMR results, as well as HRMS were consistent with literature data of enone **121** ([M]Na^{+} calculated 205.0477, found 205.0477).[97]

Three additional new peaks were observed at t_R = 2.3, 3.3 and 4.3 min which showed a maximum UV absorption of 303 nm similar to **121**. Three characteristic ions (*m/z* = 415, 197 and 179) in ES^{+} were found for all peaks (Figure 5.13A). Any attempts to purify either of these compounds yielded

enone compound **121**. This supports the idea that these are unstable methanol adducts **192a-c** of **121** that could be reversibly formed due to methanol present in the sample (Figure 5.13B).

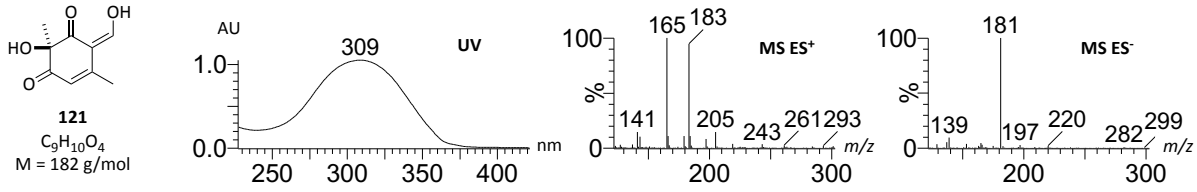


Figure 5.11 Chemical characterisation of **121** (t_R = 2.9 min) by UV and mass spectrum obtained from ES^+ and ES^-.

Another rarely detected compound was purified (4 mg) from a 1 L Starch M culture extract (Figure 5.14A). The compound **193** eluting at t_R = 7.4 min showed a characteristic UV absorption of 294 nm and a nominal mass of 210 was determined by low resolution mass (Figure 5.14B). Analysis of ^1H and ^{13}C NMR data identified 11 carbon atoms and 14 protons (Table 5.2). HRMS confirmed a molecular formula of $C_{11}H_{14}O_4$ ($[M]Na^+$ calculated 233.0790, found 233.0787).

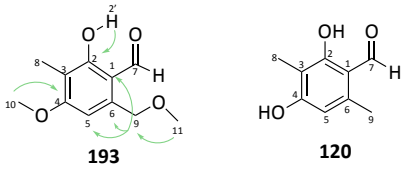


Figure 5.12 Structural proposal for **193** (arrows indicate key HMBC correlations) and **120** for comparison.

Table 5.2 NMR data for **193** in $CDCl_3$ (400 MHz) referenced to $CDCl_3$.

Position	δ_C/ppm 193	δ_H/ppm (*J* in Hz) 193	HMBC (H to C)	δ_C/ppm 120[102]	δ_H/ppm (*J* in Hz) 120[102]
1	140.7	-	-	141.4	-
2	163.2	-	-	161.0	
2′	-	12.40 (s, 1H, O-H)	2, 3, 6, 7		12.65 (s, 1H, OH)
3	113.6	-	-	108.9	-
4	163.8	-	-	164.1	-
5	104.3	6.47 (s, 1H)	1, 3, 4, 6, 7, 8, 9	109.9	6.20 (s, 1 H)
6	112.8	-	-	113.3	-
7	193.9	10.14 (s, 1H)	2, 3, 6	193.0	10.07 (s, 1H)
8	7.4	2.08 (s, 3H)	1, 2, 3, 4, 5, 6	6.8	2.08 (s, 3H)
9	72.2	4.67 (s, 2H)	1, 5, 6, 11	17.9	2.50 (s, 3H)
10	56.0	3.92 (s, 3H)	4, 5	-	-
11	58.2	3.41 (s, 3H)	9	-	-

NMR data of **193** and **120** were found to be almost identical. Apart from two additional O-CH_3 chemical shifts, they are very similar (Table 5.2, Figure 5.16)). However, the 6-CH_3 group of **120** does not appear in NMR data collected for **193**, but instead an O-CH_2-group, proposing hydroxylation and methylation at this position. Evaluation of HMBC couplings led to the structural proposal of **193** (Figure 5.12).

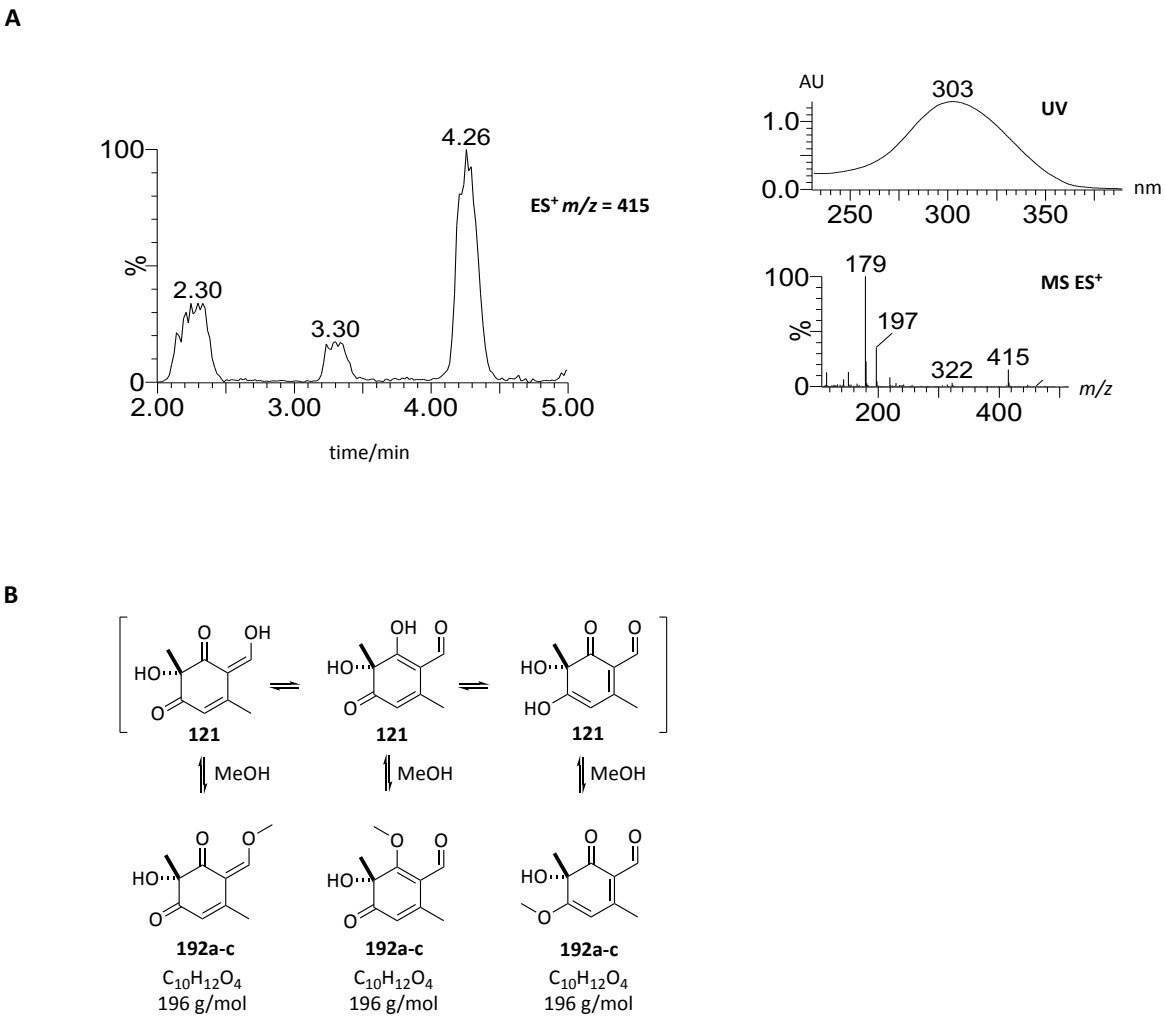


Figure 5.13 Compounds **192a-c. A**, Extracted ion chromatogram (ES+ at *m/z* = 415) (t_R = 2.3, 3.3 and 4.3 min). UV and mass spectrum obtained from ES^+ is identical for the three observed compounds. **B**, Proposed methanol adducts of **121**.

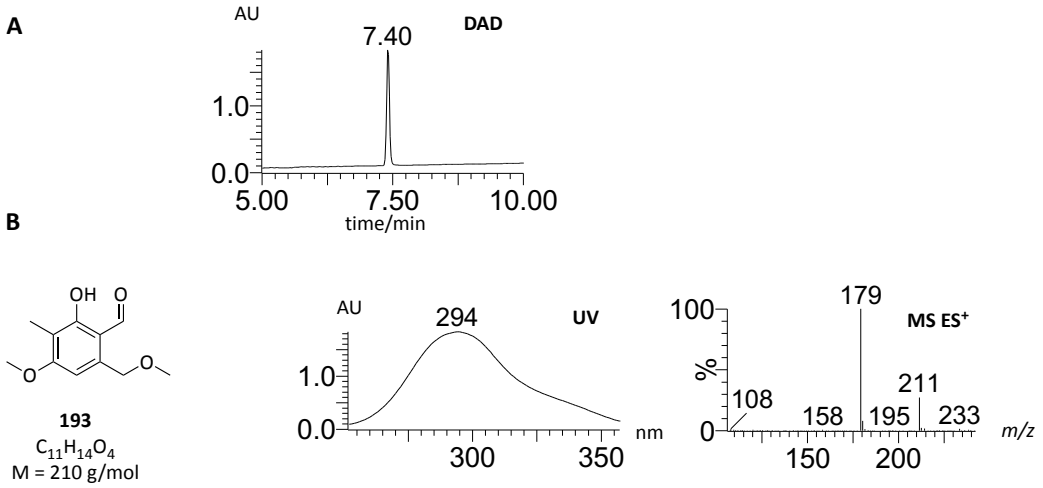


Figure 5.14 Compound **193. A**, DAD chromatogram of purified compound **193** analysed with analytical gradient A1. **B**, Chemical characterisation of **193** (t_R = 7.5 min) by UV and mass spectrum obtained from ES^+.

5.2.3 Co-Expression of *aspks1+asL1+asL3*

Transformation of *A. oryzae* NSAR1 with *aspks1+asL1+asL3* yielded 6 colonies. For analysis of their secondary metabolite profile the transformants were sub-cultured in P_{amyB} induction medium. Liquid cultures and plates of these transformants displayed a bright yellow colour (Figure 5.15B). The obtained extracts (Chapter 5.1.2) were submitted to LCMS and the chromatograms were analysed for the production of stipitaldeyhde **123** (Figure 5.15C).

In four extracts a new compound eluting at t_R = 7.4 min was identified. It was determined to have a nominal mass of 180 and displayed three characteristic UV absorption maxima (Figure 5.16) (246, 271, 377 nm). These characteristics are consistent with literature data for **123**.[97,223]

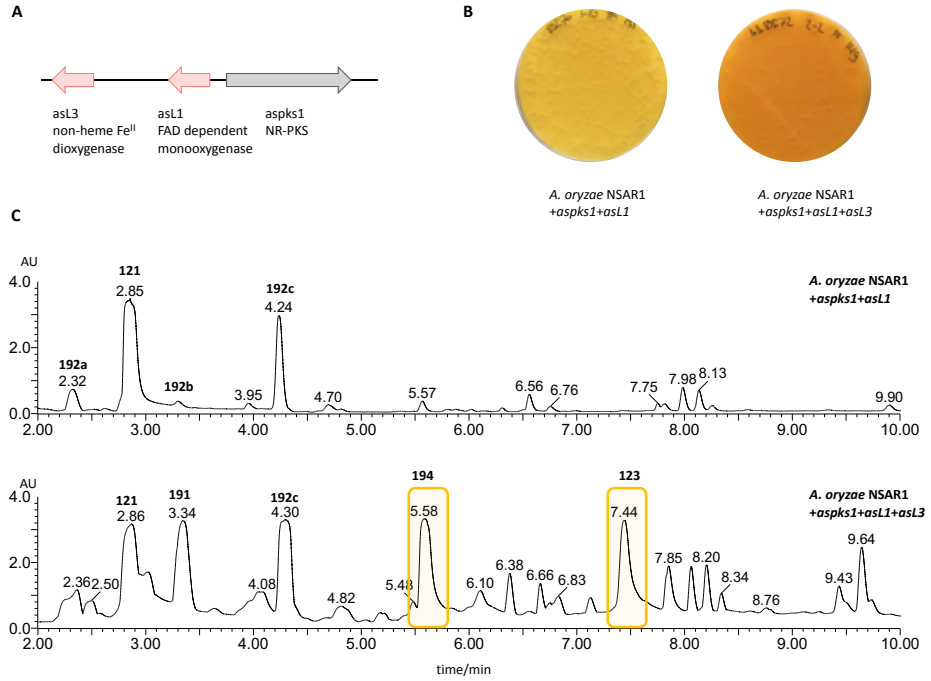


Figure 5.15 Expression of three genes. **A** Co-expression of *aspks1+asL1+asL3*. **B** *A. oryzae* NSAR1 +*aspks1+asL1* and *A. oryzae* NSAR1 +*aspks1+asL1+asL3* on DPY agar. **C** DAD chromatogram of an *A. oryzae* NSAR1 +*aspks1* and an *A. oryzae* NSAR1 +*aspks1+asL1* sample measured with analytical gradient A2 on LCMS. Extracts were obtained from cultures grown in DPY for 4 d.

Purification of this compound (0.6 mg) from 1 L Starch M culture extract, showed the compound to have ^{1}H NMR data which is highly similar to **123** (Table 5.3). The small differences in chemical shifts compared to literature data could be explained due to a pH dependent different resonance structure. HRMS confirmed a molecular formula of $C_9H_8O_4$ ($[M-H]^-$ $C_9H_7O_4$ calculated 179.0344, found 179.0344).

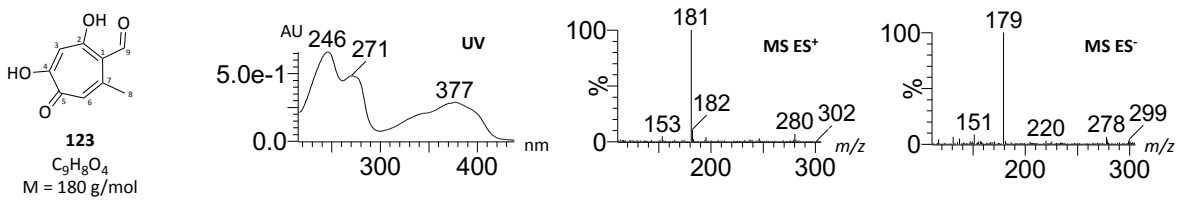


Figure 5.16 Chemical characterisation of **123** (t_R = 7.4 min) by UV and mass spectrum obtained from ES^+ and ES^-.

Table 5.3 ^{1}H NMR data for **123** in d_6-DMSO (500 MHz) referenced to d_6-DMSO.

Position	δ_H Literature[97]	δ_H Purified
3	6.75 (s, 1H)	6.28 (s, 1H)
6	6.69 (s, 1H)	6.37 (s, 1H)
8	2.57 (s, 3H)	2.46 (s, 3H)
9	10.01 (s, 1H)	9.82 (s, 1H)

In addition, a compound with the nominal mass of 152 corresponding to **194**, which previously was observed upon KO of *tropD* in *T. stipitatus*, was identified. The compound eluting at t_R = 5.6 min showed two characteristic UV absorption maxima (250 nm, 343 nm) consistent with literature data for **194** (Figure 5.17).[97,224] This compound (4 mg) was purified from a 1 L Starch M culture extract and the structure confirmed by NMR as well as HRMS ([M]-H^+ $C_8H_9O_3$ calculated 153.0552, found 153.0553) (Chapter 7.5.6 for NMR data).

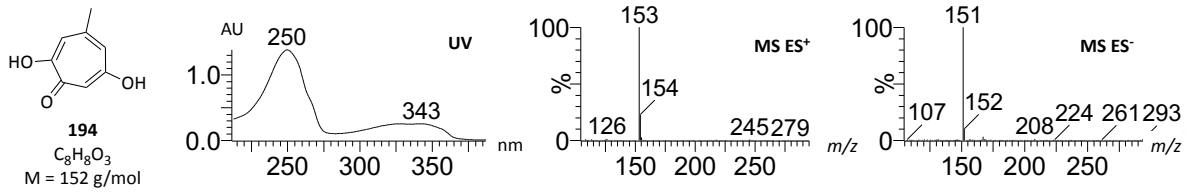


Figure 5.17 Chemical characterisation of **194** (t_R = 5.6 min) by UV and mass spectrum obtained from ES^+ and ES^-.

5.2.4 Co-Expression of *aspks1+asL1+asL3+asR2*

Transformation of *A. oryzae* NSAR1 with *aspks1*+*asL1*+*asL3*+*asR2* yielded 11 colonies. For analysis of their secondary metabolite profile the transformants were sub-cultured in P_{amyB} induction medium. Liquid cultures and plates of these transformants displayed a bright yellow colour (Figure 5.18B). The obtained extracts (Chapter 5.1.2) were submitted to LCMS and the chromatograms were analysed (Figure 5.18C). Several new compounds eluting at t_R = 4.7, 5.5, 7.3 and 8.4 min were observed in comparison to an extract of a transformant expressing *aspks1*+*asL1*+*asL3* only.

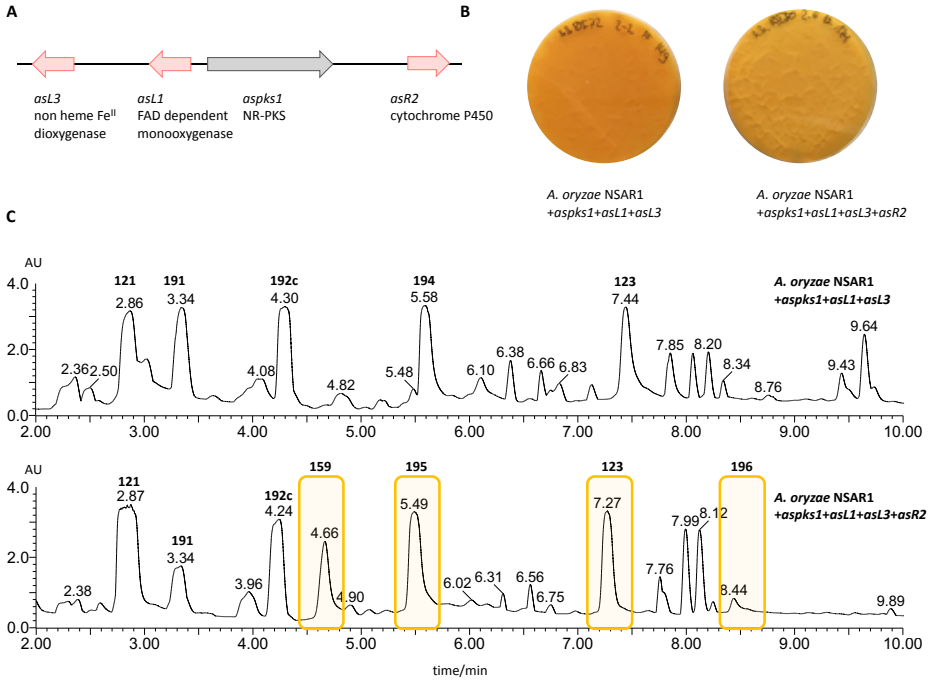


Figure 5.18 Expression of four genes. **A**, Co-expression of *aspks1+asL1+asL3+asR2*. **B**, *A. oryzae* NSAR1 *+aspks1+asL1+asL3 and A. oryzae* NSAR1 *+aspks1+asL1+asL3+asR2* on DPY agar. **C**, DAD chromatogram of an *A. oryzae* NSAR1 *+aspks1+asL1+asL3 and A. oryzae* NSAR1 *+aspks1+asL1+asL3+asR2* sample measured with analytical gradient A2 on LCMS. Extracts were obtained from cultures grown in DPY for 4 d.

The compound eluting at t_R = 7.3 min was determined to have a nominal mass of 178 and displayed three characteristic UV absorption maxima (261, 272, 317 nm) (Figure 5.19). These characteristics are consistent with literature data for **125**.[98] Purification of 4 mg from 1 L DPY culture, subsequent NMR (^{1}H, HMBC, HSQC) analysis and HRMS ([M]-H$^+$ $C_9H_6O_4$ calculated 179.0344, found 179.0344) confirmed the structure (Chapter 7.5.6 for NMR data).

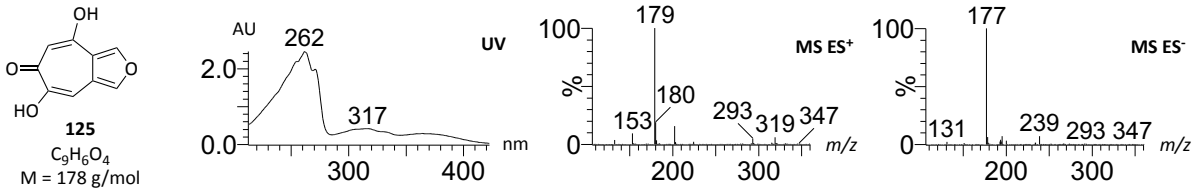


Figure 5.19 Chemical characterisation of **123** (t_R = 7.3 min) by UV and mass spectrum obtained from ES$^+$ and ES$^-$.

In addition, a compound eluting at t_R = 5.5 min with highly similar maximum UV absorption spectrum was detected alongside **125** (Figure 5.20). It was determined to have a nominal mass of 192, which could correspond to a methylated shunt, **195a** or **195b**.

Analysis of ^{1}H and ^{13}C NMR data of the isolated compound (4 mg, 1 L DPY culture) showed 10 carbon atoms and 7 protons. HRMS of **195** confirmed a molecular formula of $C_{10}H_8O_4$ ([M]H$^+$ $C_{10}H_9O_4$ calculated 193.0501, found 193.0502) which suggested one exchangeable proton. Comparison of

the NMR data obtained for **195a** or **195b** and **125**, showed highly similar signals, with one additional methyl group at δ_H 3.65 (δ_C 56.3) in the data for **195** (Table 5.4).

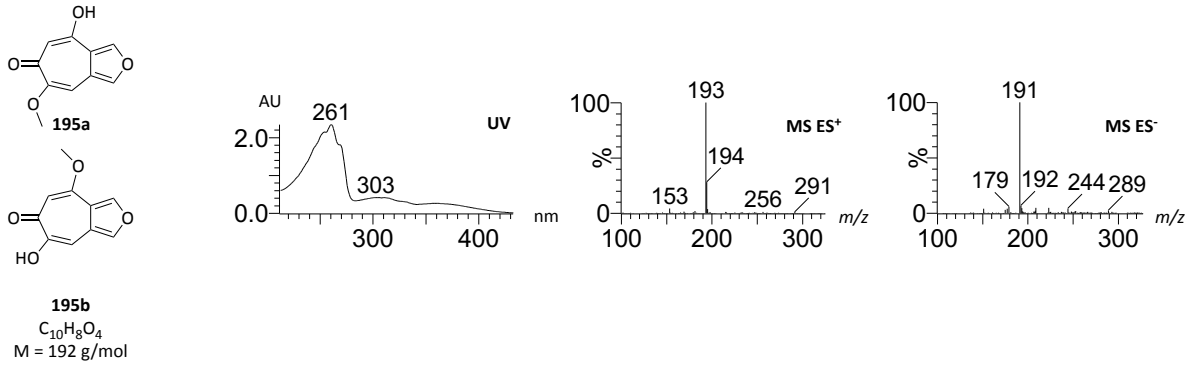


Figure 5.20 Chemical characterisation of **195** (t_R = 5.5 min) by UV and mass spectrum obtained from ES^+ and ES^-

The chemical shift of this methyl group strongly suggests the attachment to an oxygen atom. Analysis of HMBC correlations and an unusual $^5J_{H,H}$ coupling of the methyl protons at δ_H 3.65 and the proton at δ_H 6.64 suggested the methylation of the 2-OH group in **195a** (Figure 5.21). The $^5J_{H,H}$ coupling was observed in the $^1H,^1H$ COSY only. In ^1H NMR signals of protons H-2 and H-10 appeared as singlets, suggesting a $^5J_{H,H} < 1$ Hz which supports the correlation over five bonds.

Table 5.4 NMR data for **195a** in d_6-DMSO (500 MHz) referenced to d_6-DMSO in comparison to literature data for **125**.

Position	δ_C/ppm 195a	δ_H/ppm (*J* in Hz) 195a	HMBC (H to C)	δ_C/ppm 125[98]	δ_H/ppm 125[98]
1	176.9	-	-	180.0	-
2	152.6	-	-	149.3	-
3	100.8	6.64 (s, 1 H)	1, 2, 5, 8	102.0	6.86 (s, 1 H)
4	119.1	-	-	119.6	-
5	122.4	-	-	119.4	-
6	165.1	-	-	161.4	-
7	109.7	5.86 (s, 1 H)	1, 2, 5, 6, 9	106.9	6.16 (s, 1H)
8	143.0	8.14 (d, 1 H, 1.9)	2, 3, 4, 5, 9	142.7	8.34 (d, 1 H, 1.8)
9	144.0	8.17 (d, 1 H, 1.9)	4, 5, 7, 8	144.3	8.22 (d, 1 H, 1.8)
10	56.3	3.65 (s, 3H)	2, 3	-	-

Analysis of the compound eluting at t_R = 8.4 min (nominal mass 356, maximal UV absorption 257, 321) suggested it to be the previously identified talaroditroplone **196** (Figure 5.22). This compound is known from *T. stipitatus* and is thought to be formed by a spontaneous [4+2] addition, followed by rearrangement, of 2 molecules of **125**.[98] Purification of 2 mg from 1 L DPY culture extract and subsequent full NMR analysis as well as HRMS ($[M-H]^-$ $C_{18}H_{11}O_8$ calculated 355.0454, found 355.0452) confirmed this observation (Chapter 7.5.6 for NMR data).

The final newly observed peak at t_R = 4.7 showed two distinct UV absorption maxima and a nominal mass of 180, which is consistent with literature data of cordytropolone **159** (Figure 5.23).[98] This compound was observed previously in *T. stipitatus* upon KO of *tropD* and is most likely a shunt

product arising from reduction of stipitafuran **125**.[98] HRMS confirmed the molecular formula for this peak ([M]H$^+$ $C_9H_9O_4$ calculated 181.0501, found 181.0502).

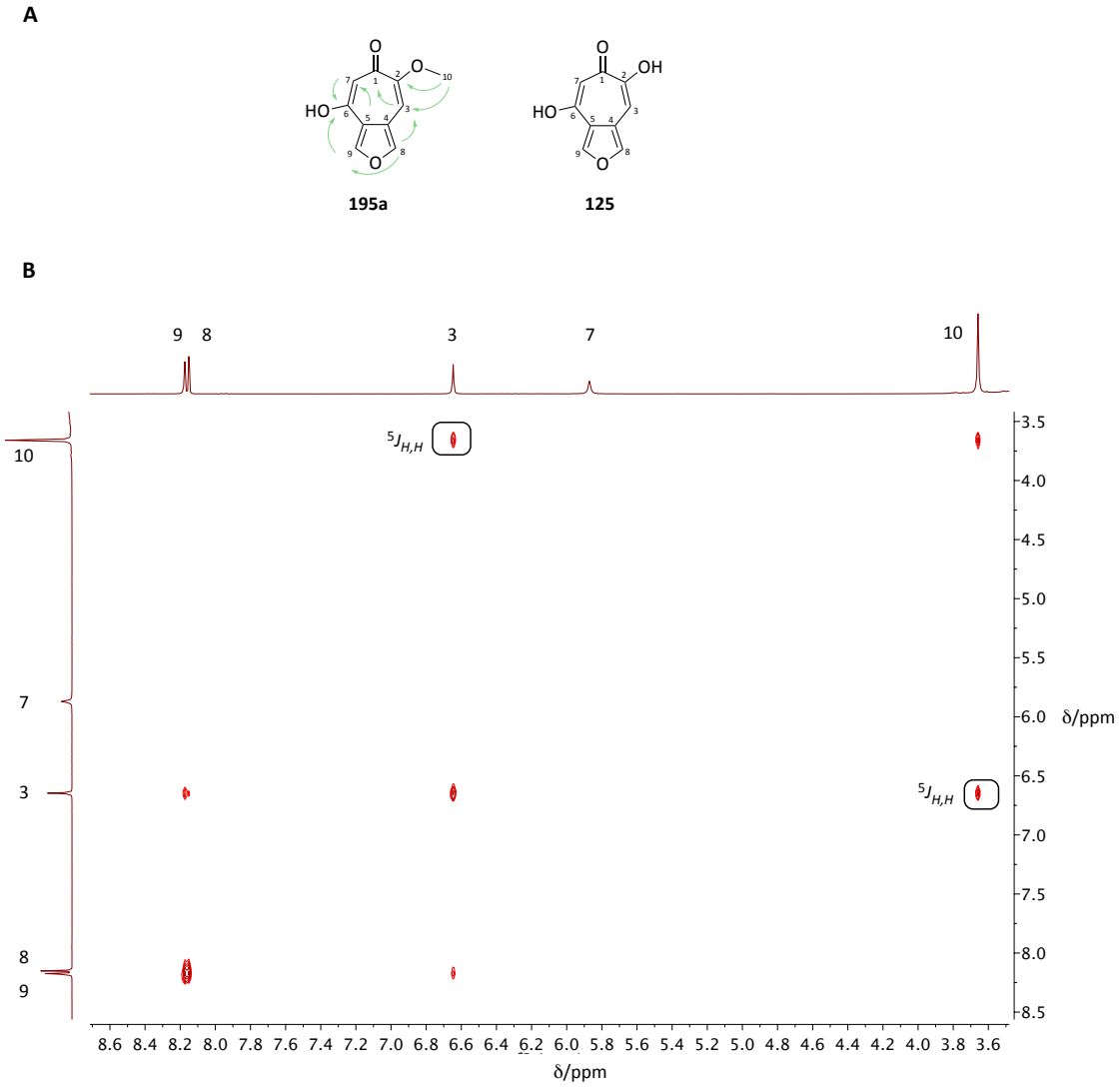


Figure 5.21 Compound **195a. A**, Key HMBC correlations observed for **195a**. **B**, ^{1}H,^{1}H COSY spectrum of **195a** showing $^5J_{H,H}$ coupling.

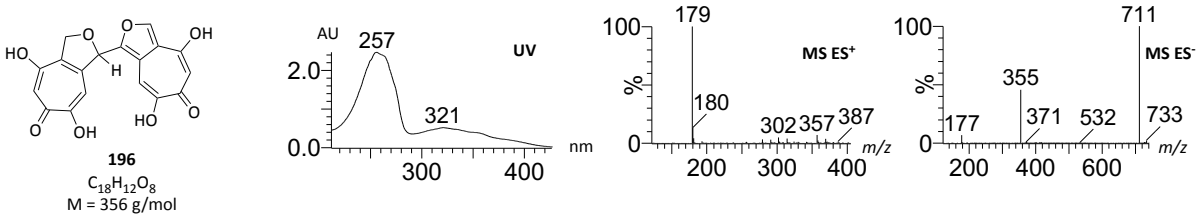


Figure 5.22 Chemical characterisation of **196** (t_R = 8.4 min) by UV and mass spectrum obtained from ES$^+$ and ES$^-$.

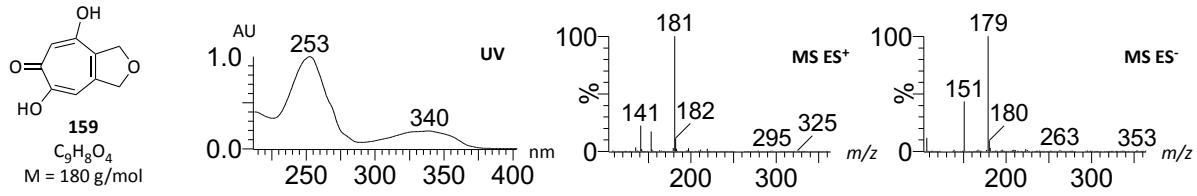
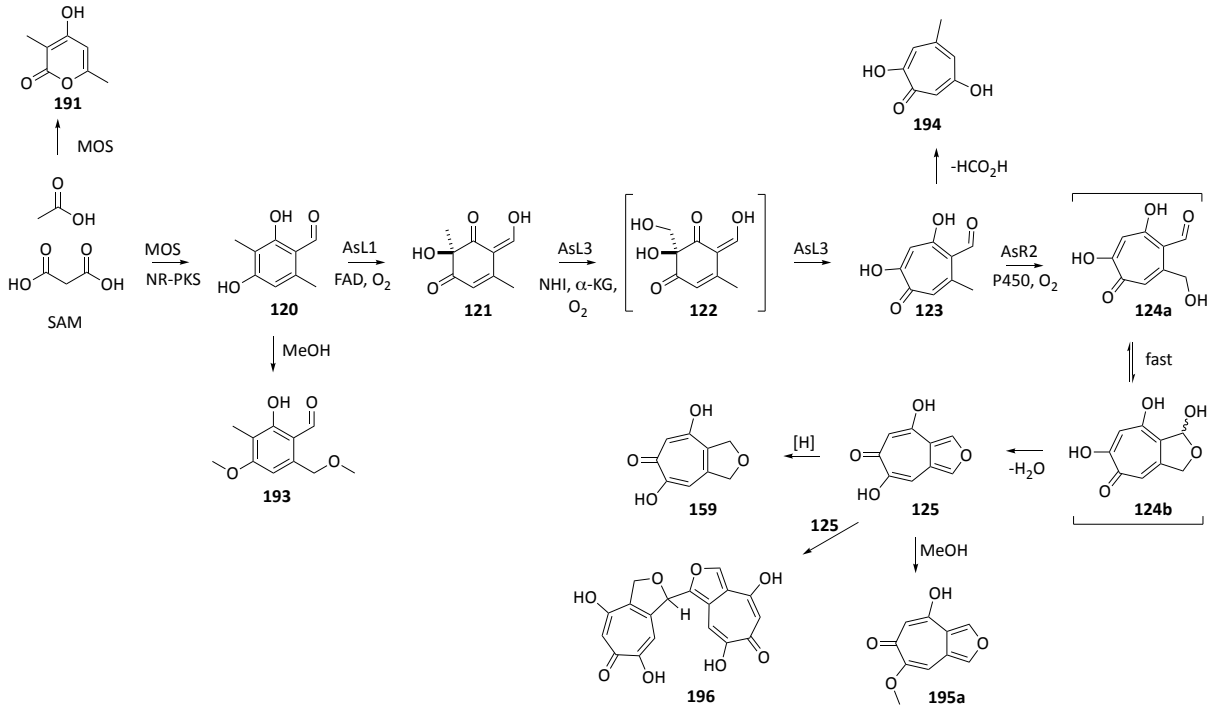


Figure 5.23 Chemical characterisation of **159** (t_R = 4.7 min) by UV and mass spectrum obtained from ES^+ and ES^-.

5.2.5 Summary of Heterologous Tropolone production in *A. oryzae*

The sequential expression of *aspks1*, *asL1*, *asL3* and *asR2* in *A. oryzae* NSAR1 led to the isolation and characterisation of the expected intermediates (**120**, **121**, **123**) which have been described previously in the biosynthesis of stipitatic acid **119**.[97,98] This showed that *aspks1*,[102] *asL1*,[104] *asL3* and *asR2* encode true tropolone forming proteins (Scheme 5.2). The direct intermediates (**120**, **121**, **123**) were observed, as well as previously observed shunt products (**191**, **194**, **125**, **159**, **196**). In addition, two compounds (**193**, **195a**) probably forming through spontaneous addition of methanol (used in the preparation of LCMS samples) were isolated and fully characterised for the first time. However, the proposed product **124a**/**b** of AsR2 was not observed.

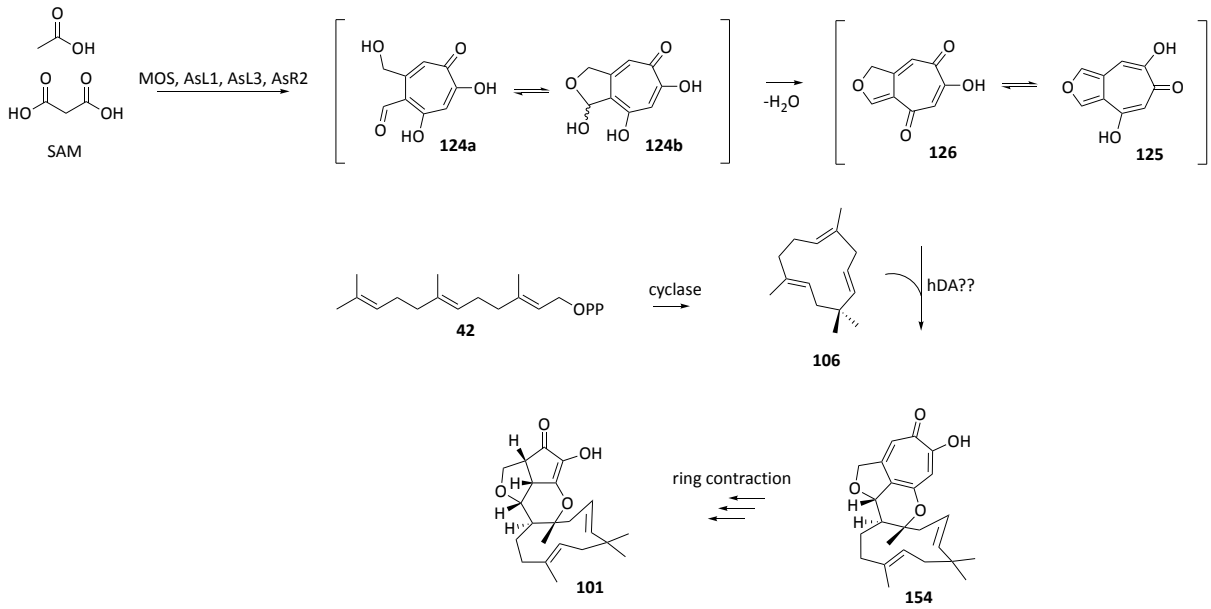


Scheme 5.2 Tropolone biosynthesis in *A. oryzae* NSAR1 by expressing *A. strictum aspks1* BGC genes. **Scheme 5.7** for proposed mechanism of formation of **193** and **195a**.

5.3 Simultaneous Expression of 11 genes of the *aspks1* BGC

The observation that four core genes of the *A. strictum aspks1* BGC encode proteins for tropolone biosynthesis confirmed the initial bioinformatic hypothesis and led to the isolation of stipitafuran **125**. For the following required three processes (humulene formation, fusion of polyketide and terpene, and the two ring contractions) no logical prediction could be made from *in silico* analysis of the BGC. This is due to the lack of a suitable terpene cyclase encoding gene and to the biosynthetically unprecedented steps of a proposed hetero Diels Alder reaction as well as two unusual ring contractions (Scheme 5.3).

In an attempt to express as many genes as possible from the *aspks1* BGC, 12 genes were initially selected (*aspks1*, *asL1-6*, *asR2*, *asR4-7*), omitting only the two transporter (*asL7*, *asR1*) and the transcription factor (*asR3*) encoding genes. However, due to initial difficulties in cloning *asR7* it was not included in the first expression experiments.



Scheme 5.3 Biosynthetic proposal for the formation of xenovulene A **101**.

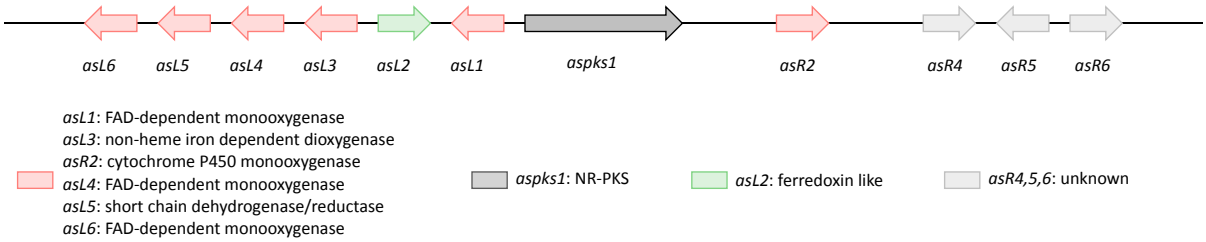


Figure 5.24 Eleven genes of the *aspks1* BGC selected for expression in *A. oryzae* NSAR1.

Transformation of *A. oryzae* NSAR1 with *aspks1*, *asL1-6, asR2* and *asR4-6* yielded 28 colonies. For analysis of their secondary metabolite profile the transformants were sub-cultured in P_{amyB} induction medium. Liquid cultures and plates of these transformants did appear slightly darker than those of previous experiments expressing tropolone forming genes only (Figure 5.25).

Figure 5.25 DPY agar plates of *A. oryzae* NSAR1 WT, *A. oryzae* NSAR1 +*aspks1+asL1+asL3+asR2* and *A. oryzae* NSAR1 +*aspks1+asL1-6+asR2+asR4-6.*

The obtained extracts (Chapter 5.1.2) were submitted to LCMS and the chromatograms were analysed. Apart from the previously characterised intermediates and shunts (Chapter 5.2), no distinct new peaks were detected in the DAD chromatogram (Figure 5.26).

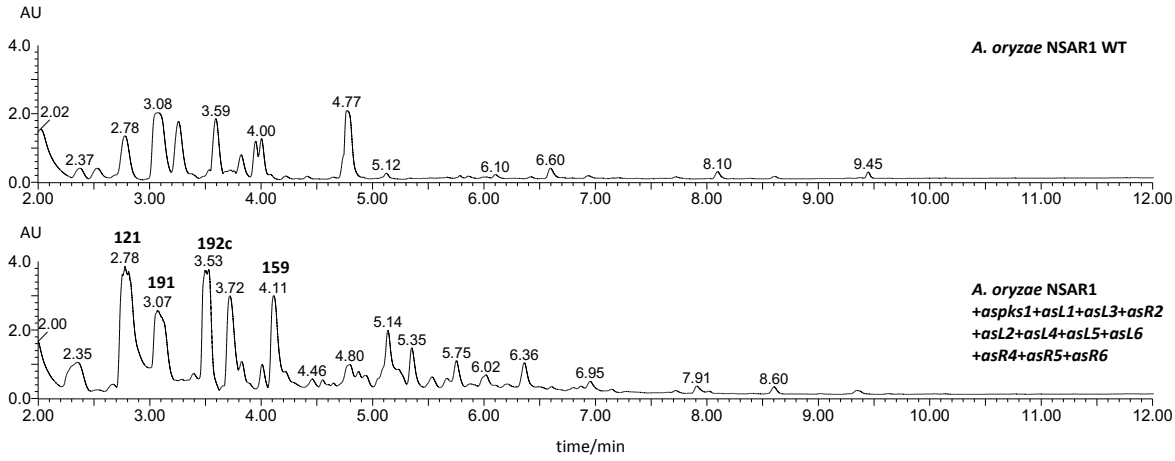


Figure 5.26 DAD chromatogram of an *A. oryzae* NSAR1 WT and an *A. oryzae* NSAR1 +*aspks1+asL1-asL6+asR2+asR4-6* sample measured with analytical gradient A1 on LCMS. Extracts were obtained from cultures grown in Starch M + 0.15% methionine for 6 d.

However, a trace compound with the nominal mass of 358 was detected in ES^- and ES^+ at t_R = 8.4 in 16 transformants, which corresponds to the retention time of xenovulene A **101**. This peak was shown to have a characteristic UV absorption maximum at 276 nm. These findings are consistent with literature data as well as previously purified (Chapter 2.1) xenovulene A **101**.[82] Further HRMS ($[M]Na^+$, $C_{22}H_{30}O_4Na$, calculated 381.2042, found 381.2041) and MS^2 (Appendix 9.4.4) of $[M]H^+$ 359 strongly suggested it to be **101**.

In order to improve the titres of **101**, four different media (ASPM, DPY, Starch M, CMP) were extracted at three different time points (day 3, 6, 10). Titres were improved in DPY medium after 3 d and **101** was detected by DAD (Figure 5.28). Purification of 1 mg of this compound from a 1 L DPY culture extract, and subsequent NMR analysis (1H, ^{1}H,^{1}H COSY, HMBC, HSQC) confirmed the structural identity of the compound produced by *A. oryzae* NSAR1 as **101** (Chapter 7.5.6 for NMR data).

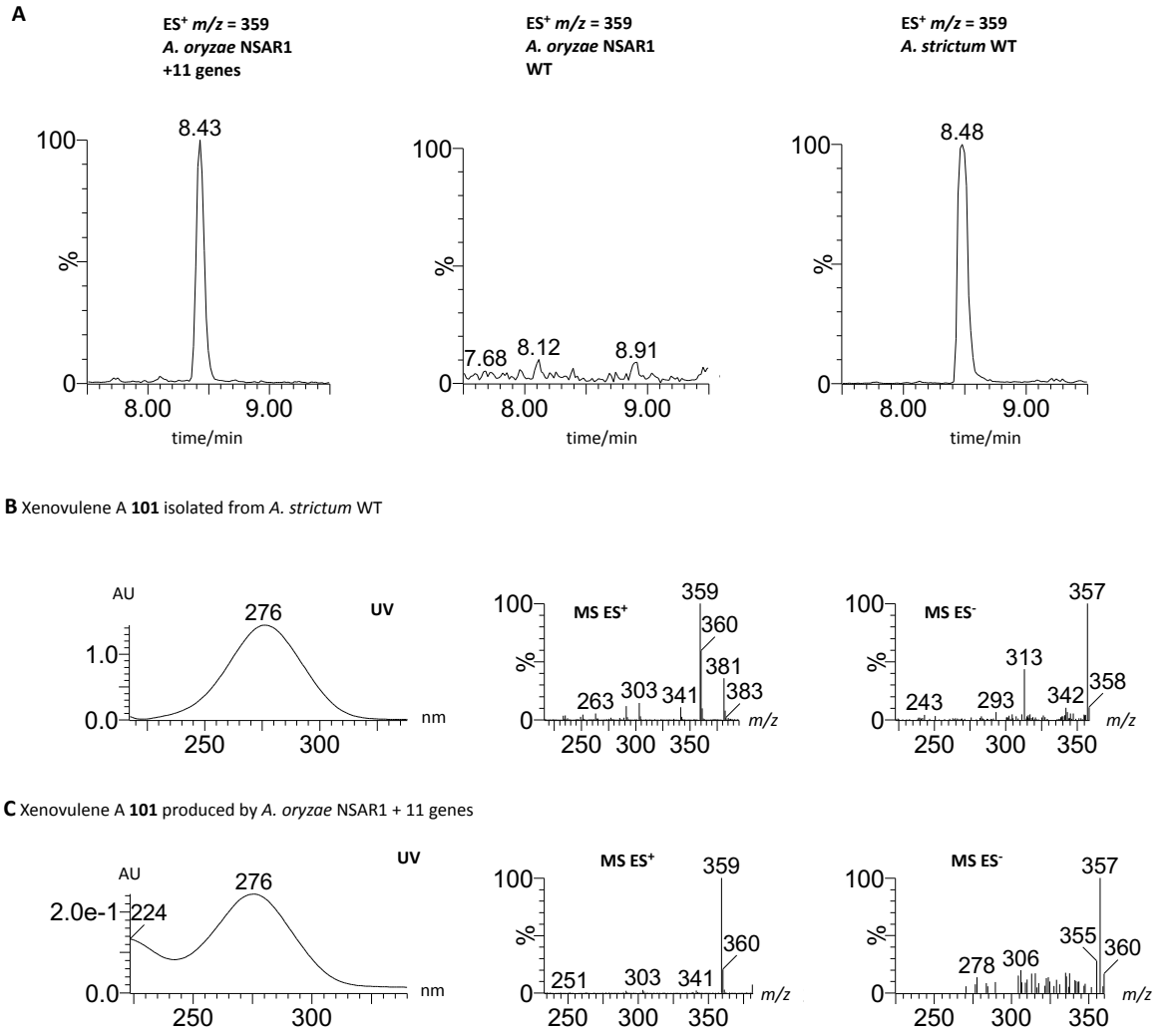


Figure 5.27 Xenovulene A **101**. **A**, Extracted ion chromatograms (ES^{+} *m/z* = 359) of samples of *A. oryzae* NSAR1 +11 genes, *A. oryzae* NSAR1 WT and *A. strictum*. **B**, Chemical characterisation of **101** isolated from *A. strictum* WT and **C**, **101** produced by *A. oryzae* NSAR1 transformed with 11 genes from the *aspks1* BGC.

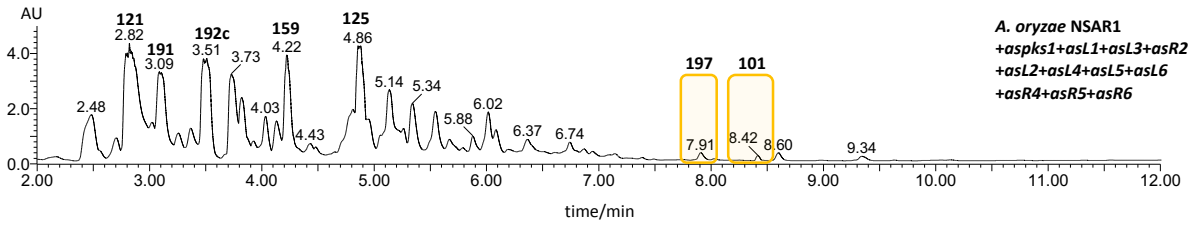


Figure 5.28 DAD chromatogram of an *A. oryzae +aspks1+asL1-asL6+asR2+asR4-6* sample measured with analytical gradient A1 on LCMS. Extracts were obtained from cultures grown in DPY for 6 d.

In the DAD chromatogram an additional peak (**197**) at t_R = 7.9 min was detected. The retention time suggested it to be a compound related to **101** and was thus targeted for purification. Approximately 1 mg was obtained from 2 L of DPY culture extracts. Subsequent ^{1}H NMR analysis showed characteristic humulene signals. Thus full NMR was acquired and analysis of ^{1}H and ^{13}C NMR data of **197** identified 23 carbon atoms and 30 protons (Table 5.5). HRMS of **197** confirmed a molecular formula of $C_{23}H_{31}O_6$ ($[M-H]^-$ calculated 403.2121, found 403.2129) which suggested two exchangeable protons.

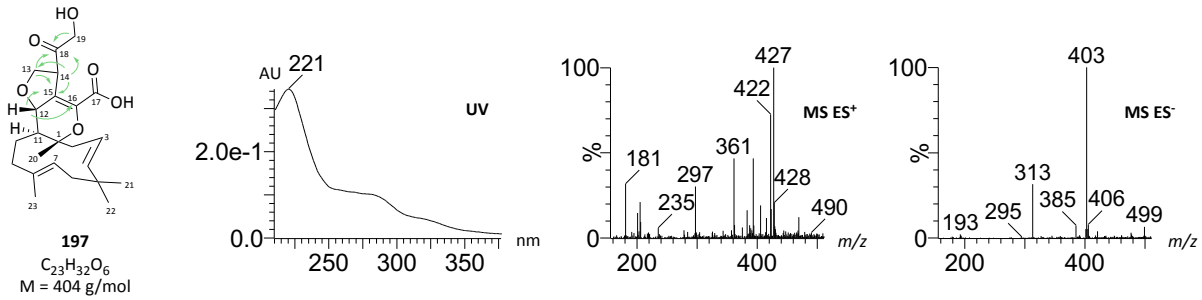


Figure 5.29 Chemical characterisation of **197** (t_R = 7.9 min) by UV and mass spectrum obtained from ES^+ and ES^-. Key HMBC correlations are indicated.

Chemical shifts of the humulene skeleton were assigned according to previously isolated *A. strictum* WT compounds (Chapter 2.1-2.7). In addition, the characteristic ether linkage of C-12 (δ_C 84.3) to C-13 (δ_C 72.5) was identified. For the usually observed tetrahydrofuran ring in xenovulene A **101** only one additional CH (δ_C 53.7) group could be detected. This proposed an unsaturated system at C-15, which was in accordance with the chemical shift of δ_C 116.9. Further consideration of the chemical shifts led to the proposal of a carbonyl (δ_C 211.2), a carboxylic acid (δ_C 170.3) and a CH_2-OH (δ_C 67.3) group. The methylene protons at δ_C 67.3 only displayed germinal coupling, which indicated the attachment next to quaternary carbon like the carbonyl functional group (δ_C 211.2). Further HMBC correlations (Figure 5.29) strongly proposed structure **197**, which is in accordance with the HRMS result. Relative stereochemistry at C-1, C-11 and C-12 was assumed to be identical to xenovulene A **101**, but not further determined.

Table 5.5 NMR data for **197** in CD_3OD (500 MHz) referenced to CD_3OD.

Position	δ_C/ppm 197	δ_H/ppm (*J* in Hz) 197	HMBC (H to C)
1	83.3	-	-
2	43.6	2.25 (m, 1H)	3, 11
		2.60 (bd, 1H, 14.2)	1, 3, 4, 11
3	122.5	5.16 (m, 1H)	2, 5, 21
4	143.1	5.15 (m, 1H)	2, 3, 5, 21
5	39.3	-	-
6	42.8	1.75 (dd, 1H, 12.6, 4.6)	4, 5, 7, 8, 21
		2.20 (m, 1H)	4, 7, 21, 22
7	124.3	5.10 (m, 1H)	6, 9, 23
8	137.7	-	-
9	39.2	2.01 (m, 1H)	7, 8, 10, 11, 23
		2.32 (m, 1H)	7, 8, 10, 11, 23
10	29.6	1.40 (m, 1 H)	8, 9
		1.49 (m, 1H)	8, 9, 12
11	41.0	1.87 (m, 1H)	1, 2, 9, 10, 12, 20
12	84.3	4.12 (dd, 1H, 10.0, 1.7)	10, 11, 15, 16
13	72.5	3.91 (m, 1H)	12, 15, 18
		3.93 (m, 1H)	12, 15, 18
14	53.7	4.09 (m, 1H)	13, 15, 18
15	116.9	-	-
16	143.2	-	-
17	170.3	-	-
18	211.2	-	-
19	67.3	4.17 (d, 1H, 18.1)	18
		4.31 (d, 1H, 18.1)	18
20	22.4	1.14 (s, 3H)	1, 2, 11
21	24.6	1.09 (s, 3H)	4, 5, 6, 22
22	30.8	1.04 (s, 3H)	4, 5, 6, 21
23	17.2	1.62 (s, 3H)	6, 7, 9

5.4 Defining the Minimal *aspks1* BGC

Although the simultaneous expression of eleven genes led to the production of xenovulene A **101**, it is possible that not all genes are essential for the biosynthesis. Four core proteins (MOS, AsL1, AsL3, AsR2) were shown to be necessary for tropolone production. But for the remaining proposed four chemical transformations (humulene formation, fusion of terpene and polyketide moiety, two ring contractions) it is likely that fewer than seven proteins are required. In a KO by expression (KOe) experiment, eight different *A. oryzae* NSAR1 strains were generated. Seven of those included sets of ten genes (Figure 5.30 1-7) and one included a set of eight genes (Figure 5.30 8).

For every set of genes, several transformants (Table 5.1) were sub-cultured in P_{amyB} induction medium. Extracts (Chapter 5.2.1) were analysed for their production of xenovulene A **101** or any related meroterpenoids observed in *A. strictum*. In all experiments the previously isolated and characterised polyketide tropolone (Chapter 5.2) intermediates **121**, **191**, **192c**, **159**, **194**, **195a**, **125** and **196** were observed.

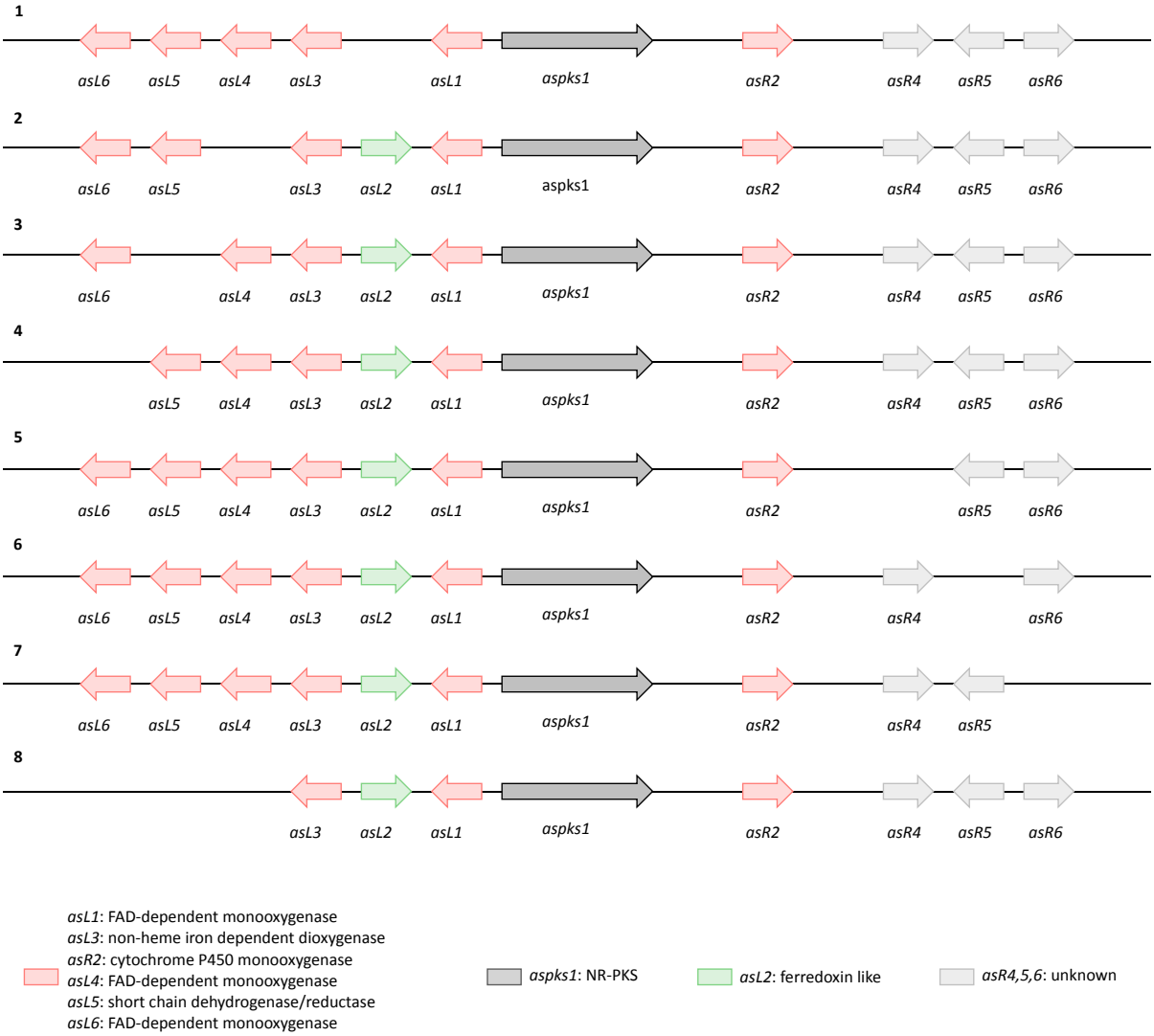


Figure 5.30 Overview of the genes simultaneously expressed in *A. oryzae* NSAR1.

For xenovulene A **101** production the extracted ion chromatogram at *m/z* = 359 $[M]H^+$ was analysed (Figure 5.33). In extracts where *asL2*, *asL5* or *asR4* were not expressed, xenovulene A **101** was clearly detected (Figure 5.33). This indicates that proteins encoded by these genes are not essential for **101** production. On the contrary, when *asR5* or *asR6* were not expressed, no xenovulene or other meroterpenoids were observed. This shows that these genes are essential for the biosynthesis (Figure 5.33). When omitting either *asL4* or *asL6* only traces of xenovulene **101** are observed (Figure 5.33). In a separate experiment, expressing *aspks1*, *asL1*, *asL2*, *asR4*, *asR5*, *asR6* in *A. oryzae* NSAR1, but neither *asL4, asL5* or *asL6*, no **101** production was observed (Figure 5.33, Chapters 5.6 for further analysis of these transformants).

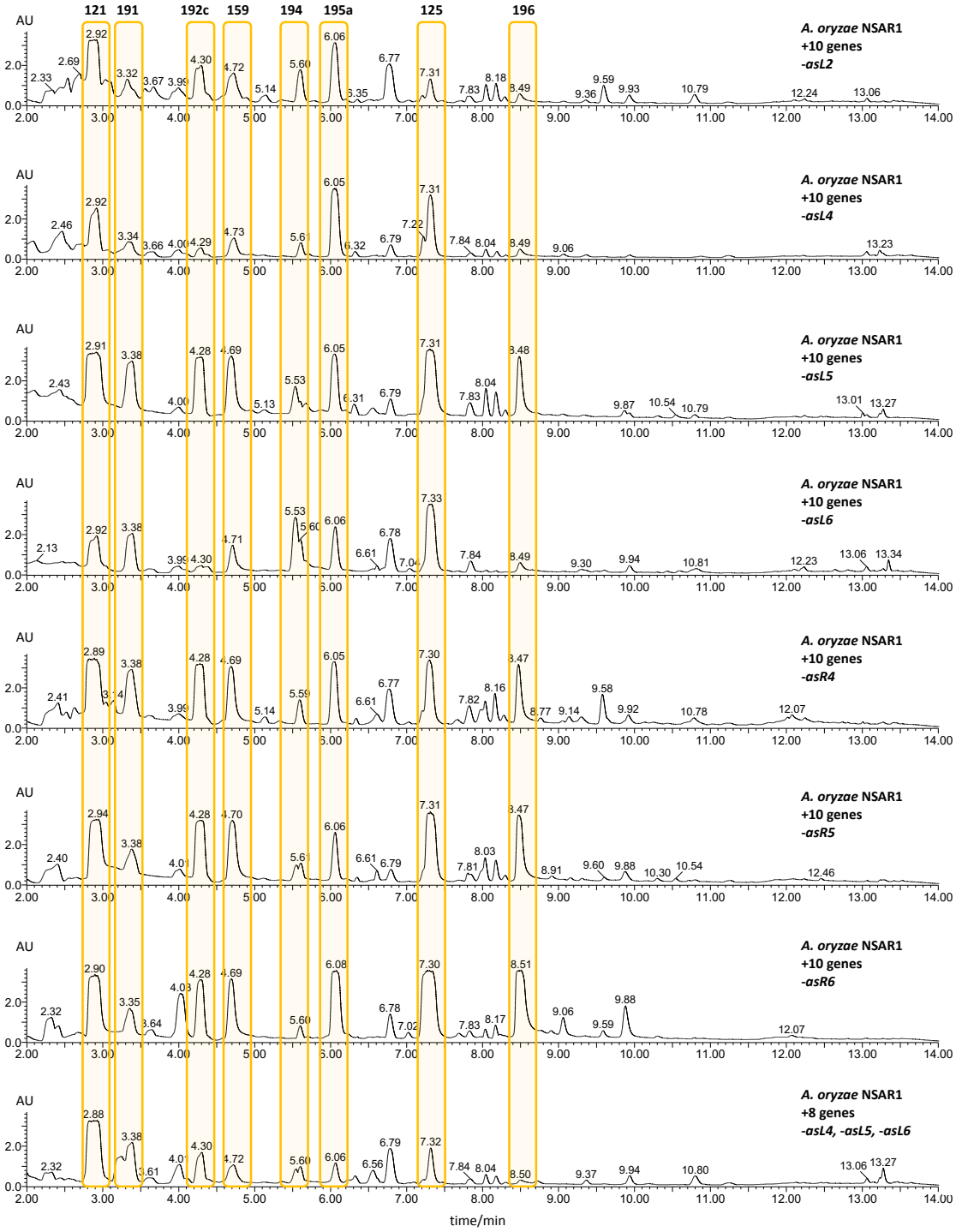


Figure 5.31 DAD chromatogram of *A. oryzae* NSAR1 extracts measured with analytical gradient A2 on LCMS. Extracts were obtained from transformants (DPY, 4 d) expressing different set of genes as indicated (Figure 5.30).

These findings strongly suggested that the minimal xenovulene A BGC is composed of eight genes (*aspks1*, *asl1*, *asL3*, *asL4*, *asL6*, *asR2*, *asR5*, and *asR6*). Subsequent transformation of *A. oryzae* NSAR 1 with these eight genes and analysis of the secondary metabolite production by Carsten Schotte showed the production of **101** (Figure 5.32). This confirmed the minimal xenovulene *aspks1* BGC.

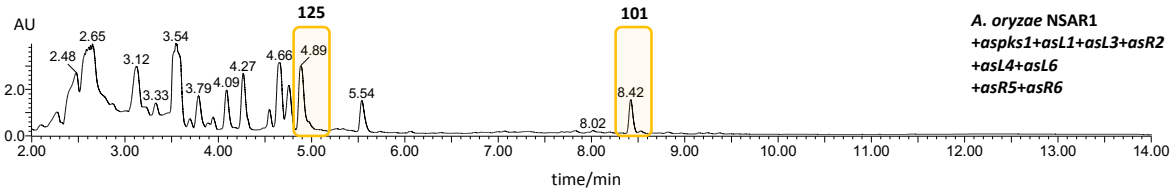


Figure 5.32 DAD chromatogram of an *A. oryzae* NSAR1 expressing the minimal BGC sample measured with analytical gradient A1 on LCMS. Extracts were obtained from cultures grown in DPY for 5 d by Carsten Schotte. Data by courtesy of Carsten Schotte.

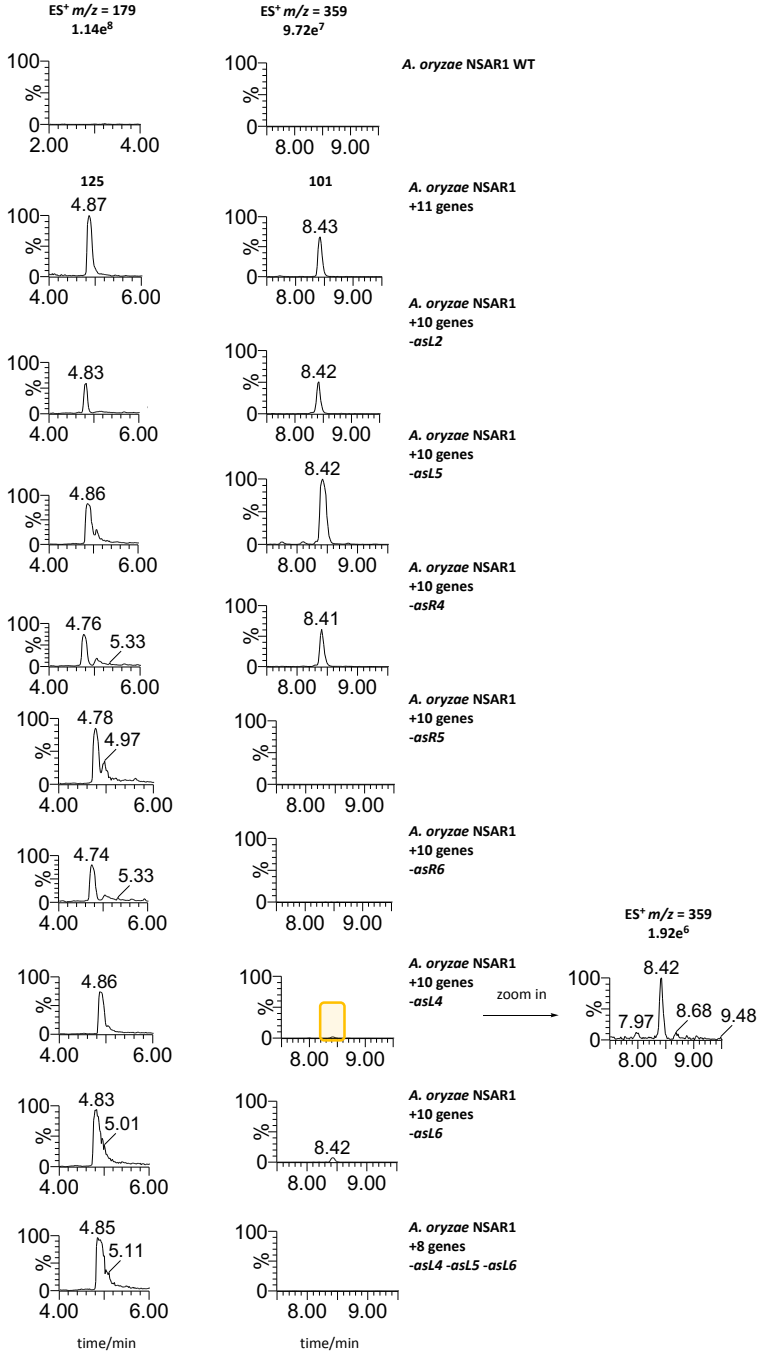
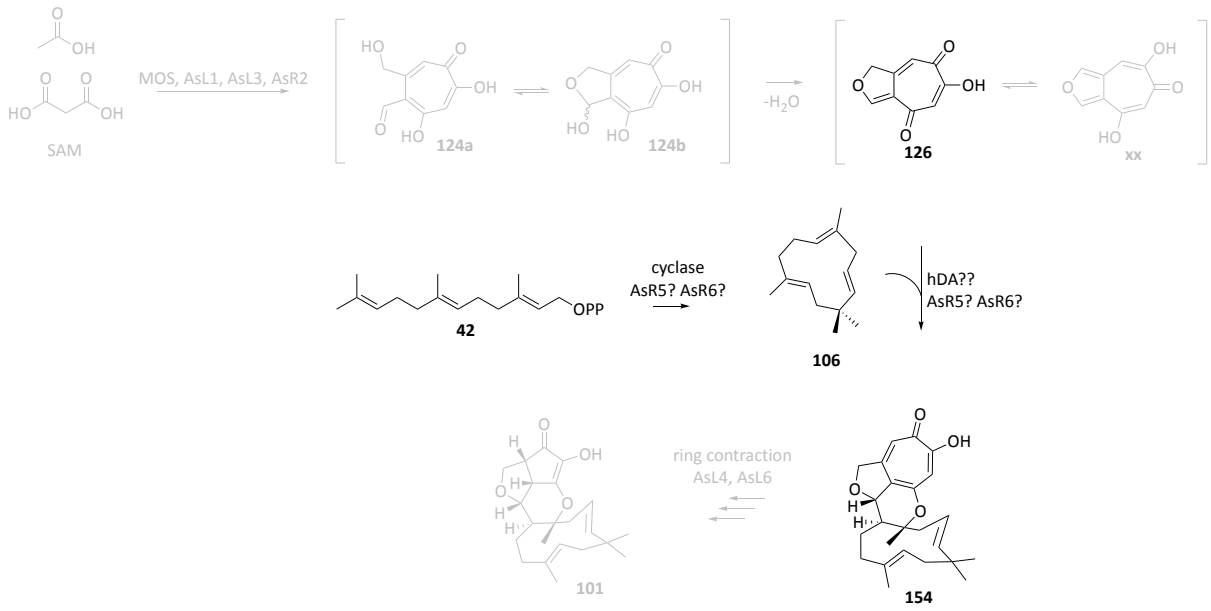


Figure 5.33 Extracted ion chromatograms ES$^+$ *m/z* = 179 for **125** and *m/z* = 359 for **101** for all extracts obtained through KOe experiments. Compound **125** shows the production of polyketide precursor and serves as an internal control. All data for *m/z* = 359 are shown at equivalent scales (9.72 e^7). Experiment *–asL4* is additionally shown at 1.92e^6 to show traces of **101**.

5.5 Formation and Attachment of Humulene

Four (*aspks1*, *asL1*, *asL3* and *asR2*) of the eight genes of the minimal BGC were assigned to their role in biosynthesis (Chapter 5.2). Of the four remaining genes, two encode putative FAD dependent oxidoreductases (*asL4*, *asL6*) and two showed very weak bioinformatic functional predictions (*asR5*, *asR6*). The KO by expression (KOe) experiments (Chapter 5.3) showed that omission of either *asL4* or *asL6* still led to the production of traces of **101** (Figure 5.33). This indicates that these transformants express all genes necessary for meroterpenoid production. It was thus concluded that proteins encoded by *asL4* and *asL6* are most likely involved in the later ring contractions (Chapter 5.6). KOe of either *asR5* or *asR6* only yielded polyketide precursors (Figure 5.33). The genes *asR5* and *asR6*, both encoding proteins of unknown function, were thus proposed to be involved in meroterpenoid formation (Scheme 5.4).



Scheme 5.4 Biosynthetic proposal for proteins of unknown function encoded in *asR5* and *asR6*.

Transformation of *A. oryzae* NSAR1 with 6 genes (*aspks1*, *asL1*, *asL3*, *asR2*, *asR5*, *asR6*) and analysis of the secondary metabolite production by Carsten Schotte showed the production of a novel compound at t_R = 11.2 min (Figure 5.34). UV absorption, t_R and the determined low resolution mass of 382 were consistent with the *A. strictum* WT meroterpenoid **154** described earlier in this thesis (Chapter 2.4). The molecular formula was confirmed by HRMS ($[M]H^+$, $C_{24}H_{31}O_4$, calculated 383.2222, found 383.2223). This gave further evidence that unknown proteins AsR5 and/or AsR6 must be involved in humulene **106** production and its fusion to the polyketide moiety.

The *asR5* and *asR6* genes were obtained as *E. coli* codon optimised N-terminal his_6 tagged clones, solubly expressed (AsR5: Raissa Schor, AsR6: Carsten Schotte) and purified by Ni-NTA affinity

purification (by Carsten Schotte) (Figure 3.35). Protein sequences of AsR5 and AsR6 were confirmed by mass spectrometry (Dr. Jennifer Senkler, AG Braun, Leibniz Universität Hannover) (Appendix 9.4.5).

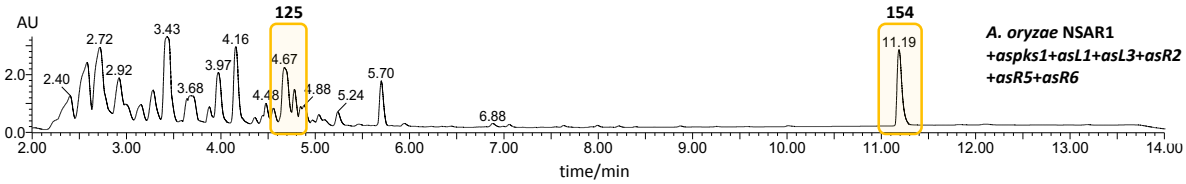


Figure 5.34 DAD chromatogram of *A. oryzae* NSAR1 transformed with six genes showing the tropolone meroterpenoid **154**. Data by courtesy of Carsten Schotte.

In vitro assays of the individual enzymes with FPP **42** and Mg^{2+} were carried out by Carsten Schotte. The assays (500 µL) were extracted with n-pentane (300 µL) and analysed by gas chromatography coupled to MS (GCMS). Incubation of FPP **42** and Mg^{2+} with AsR6, but not AsR5 (data not shown), resulted in the detection of a major novel peak eluting at t_R = 8.7 (Figure 5.36A and 5.36B). As negative control heat inactivated protein was incubated with FPP **42** and Mg^{2+} under assay conditions, which did not lead to the production of this novel peak (Figure 5.36B). This peak was identified as humulene by comparison of retention time and mass fragmentation (Figure 5.36B and 5.36C) with a humulene standard (purchased at Sigma-Aldrich, 96% purity). The observed minor peaks in assay and standard were not further characterised.

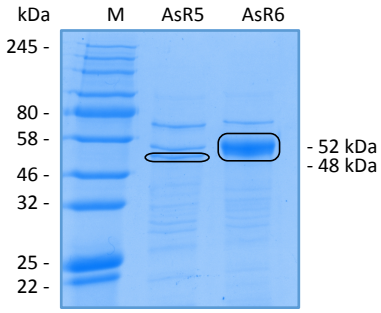


Figure 5.35 12% SDS-PAGE of Ni-NTA purified AsR5 and AsR6. Picture by courtesy of Carsten Schotte.

Further *in vitro* assays (conducted by Carsten Schotte) using AsR6, stipitafuran **125** and humulene **106** (Figure 5.5A) or AsR5, AsR6, FPP **42** and stipitafuran **125** (Figure 5.5B) were designed to produce meroterpenoid **154**. However, these experiments have not yet shown any conversion.

The heterologous expression of *aspks1*, *asL1*, *asL3*, *asR2*, *asR5* and *asR6* led to the production of meroterpenoid **154**. These results showed that AsR5 and AsR6 are involved in humulene formation and its fusion to the polyketide moiety. Further *in vitro* assays of AsR6 identified its role as a novel humulene **106** synthase. This strongly suggests that AsR5 must be involved in joining humulene to

a tropolone precursor to form **154**. However, first *in vitro* assays with stipitafuran **125** as polyketide precursor and humulene **106** did not give further evidence for this.

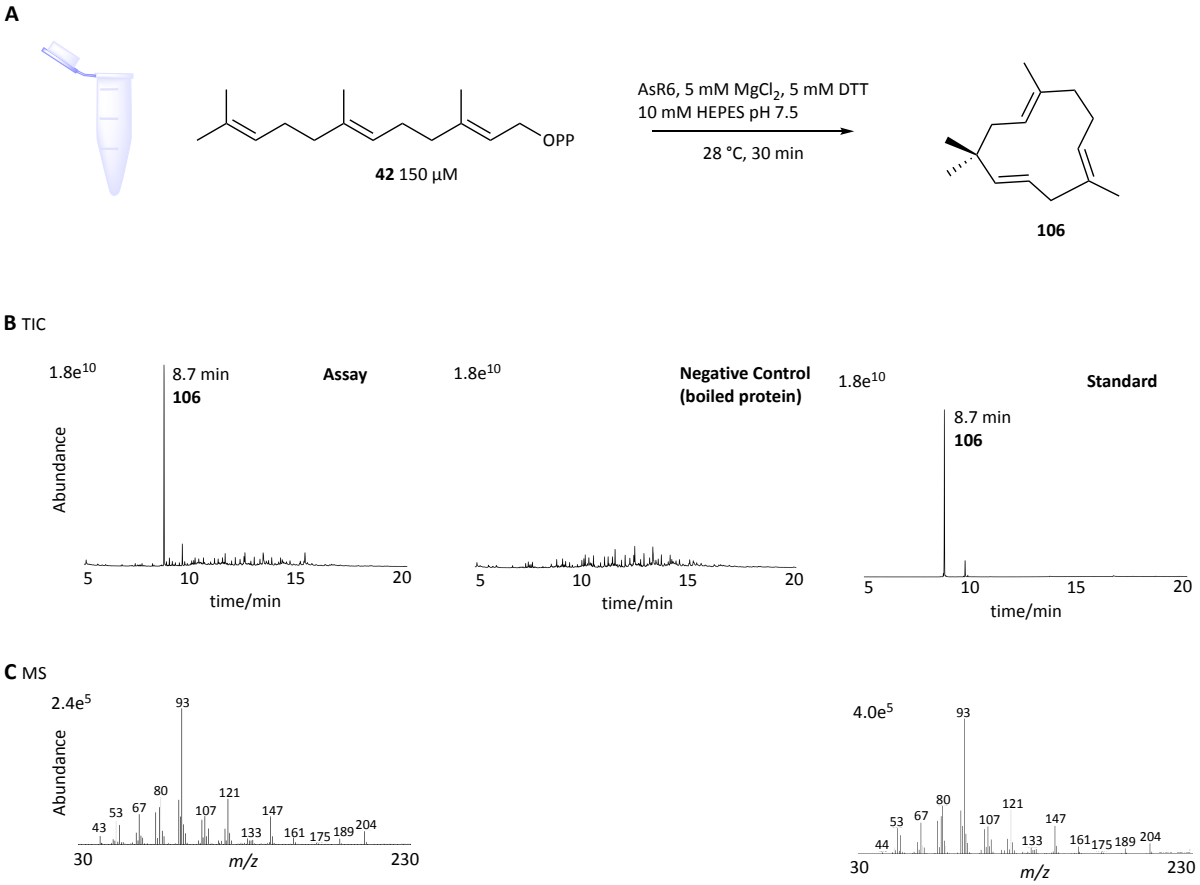
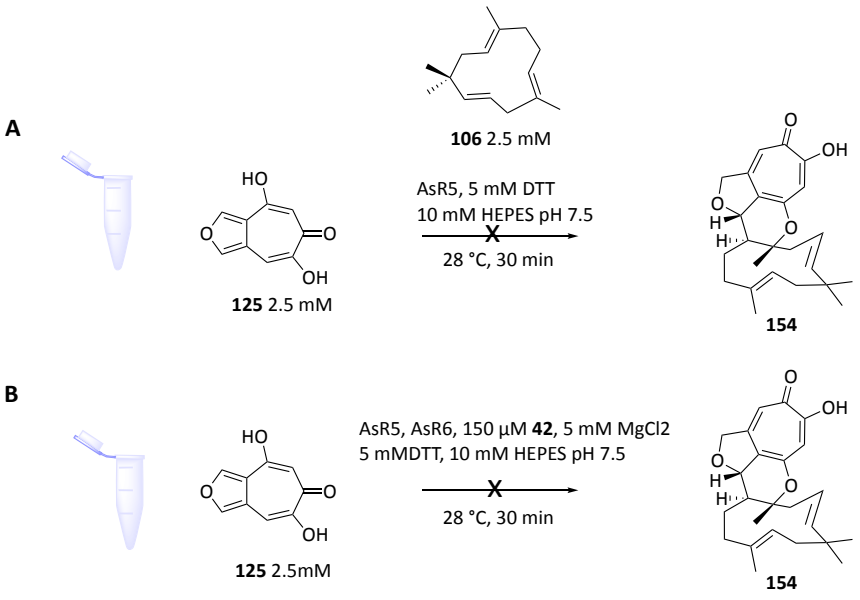


Figure 5.36 AsR6 *in vitro* assay. **A**, Assay conditions. **B**, TIC of assay extract, negative control (heat inactivated AsR6) and standard humulene **106**. **C**, MS (electron impact ionisation) of humulene **106** produced by AsR6 and standard. GCMS data by courtesy of Carsten Schotte.



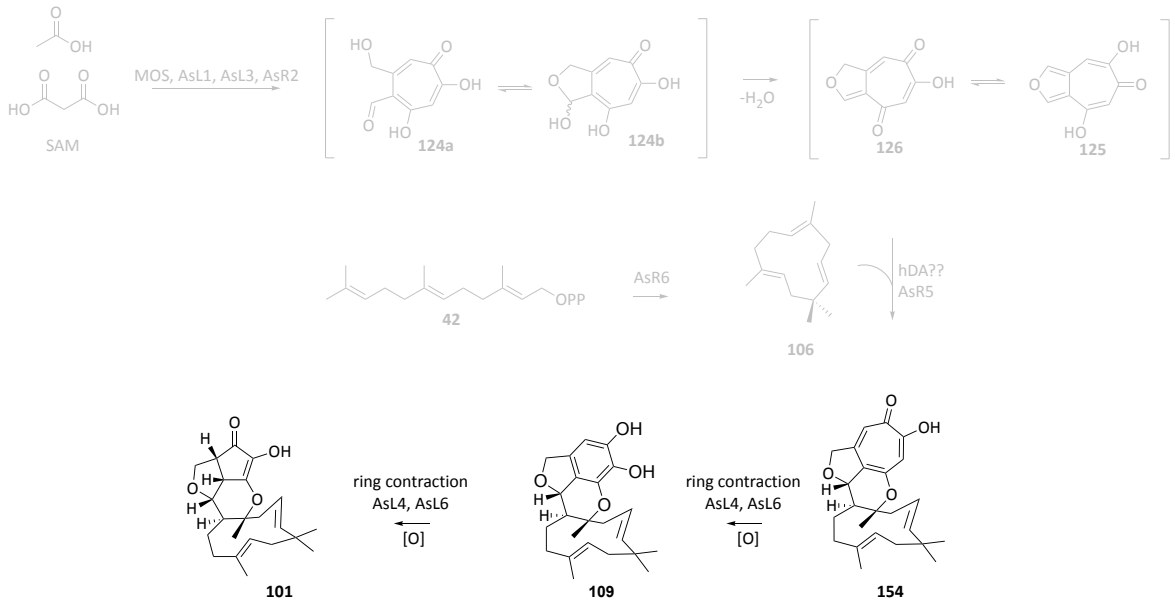
Scheme 5.5 Attempted *in vitro* assays to investigate a hetero Diels Alder activity of AsR5. **A**, Providing humulene **106**. **B**, Adding AsR6, **42** and Mg^{2+} for humulene **106** production *in vitro*.

5.6 Ring Contraction: from Tropolone towards the Cyclopentenone

For five of the eight proteins encoded by the minimal *aspks1* BGC a role during biosynthesis was assigned (Chapter 5.2: *aspks1*, *asL1*, *asL3*, *asR2*; Chapter 5.5: *asR6*). For AsR5 (Chapter 5.5) the gathered experimental data strongly proposes the involvement in the attachment of polyketide precursor and humulene.

For the two remaining putative FAD dependent monooxygenases AsL4 and AsL6 a role in the ring contraction of tropolones to cyclopentenone in **101** was proposed. The analysis of fungal extracts lacking either expression of *asL4* or *asL6* still showed trace production of Xenovulene A **101** (Figure 5.33). However, analysis of fungal extracts lacking expression of both *asL4* and *asL6* only yielded the meroterpenoid **154** (Figure 5.34). These results strongly proposed that AsL4 and AsL6 are involved in the two ring contraction steps from tropolone to xenovulene A **101** and that they can partially complement each other (Scheme 5.6).

Further analysis of the chromatograms of extracts lacking *asL4* or *asL6* expression identified the presence of two compounds with a nominal mass of 370 at t_R = 9.2 min and t_R = 9.4 min (Figure 5.37).



Scheme 5.6 Biosynthetic proposal for the ring contraction of meroterpenoid **154** to xenovulene A **101**.

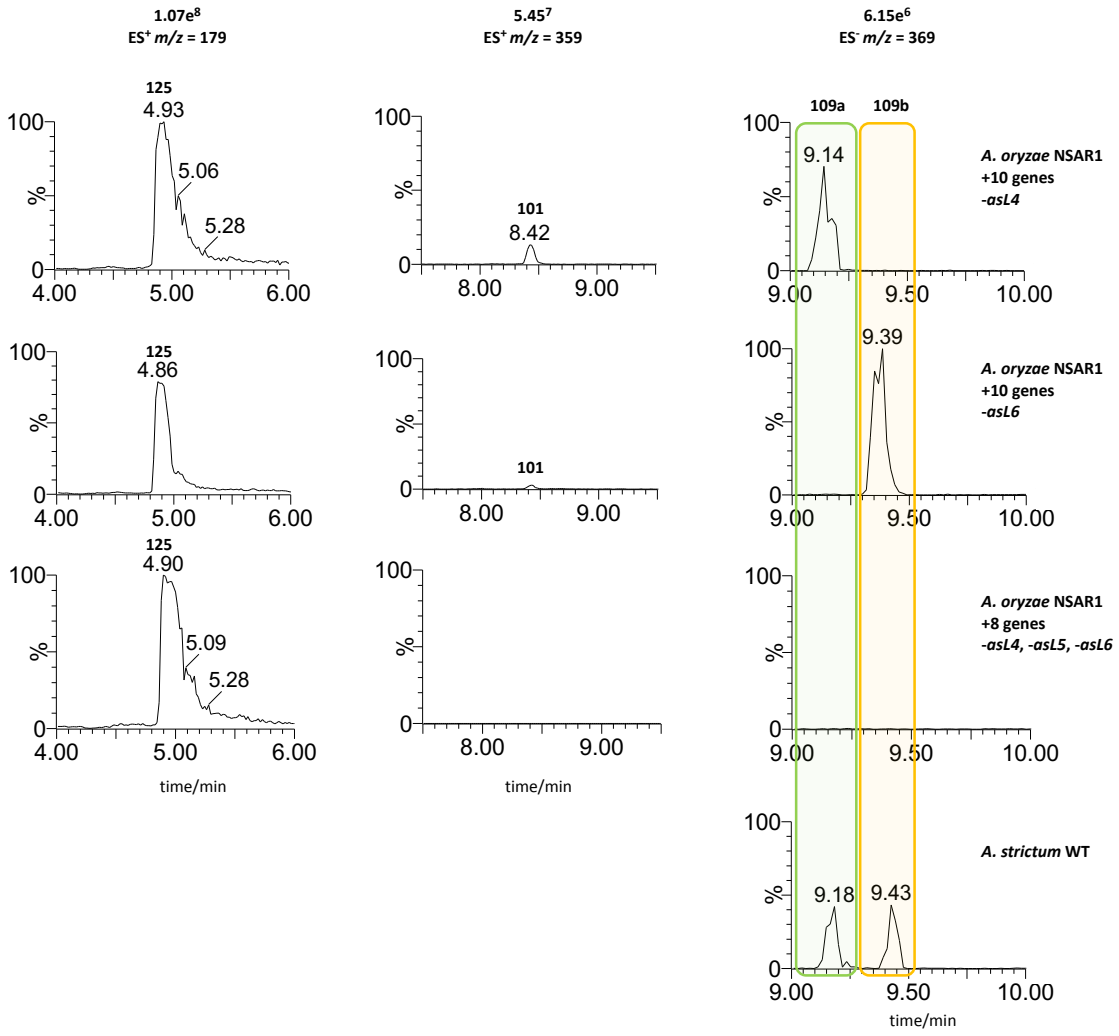


Figure 5.37 Extracted ion chromatograms ES^- *m/z* = 369 for **109a/b**, ES^+ *m/z* = 179 for **125** and *m/z* = 359 for **101** for extracts obtained through KOe experiments. Compound **125** shows the production of polyketide precursor and serves as an internal control.

The compound with t_R = 9.2 min was only observed in fungal extracts lacking *asL4* expression whereas the compound eluting at t_R = 9.4 min was only traceable in extracts lacking *asL6* expression. Retention time, HRMS and MS^2 (Appendix 9.4.6) identified the two compounds as phenolic meroterpenoids **109a** ($[M-H]^-$, $C_{23}H_{29}O_4$, calculated 369.2066, found 369.2066) and **109b** ($[M-H]^-$, $C_{23}H_{29}O_4$, calculated 369.2066, found 369.2065), which were previously observed in *A. strictum* WT (Figure 5.37, Chapter 2.5). These compounds were not observed in fungal extracts obtained from transformants lacking *asL4* and *asL6* expression at the same time (Figure 5.37).

These results strongly propose that the two FAD dependent monooxygenases AsL4 and AsL6 are involved in two the ring contraction steps. The observation of traces of xenovulene A **101** detected when either *asL4* or *asL6* is KOe, but no detection of **101** when both genes are lacking, indicates that the encoded proteins can at least partially complement one another.

5.7 Discussion

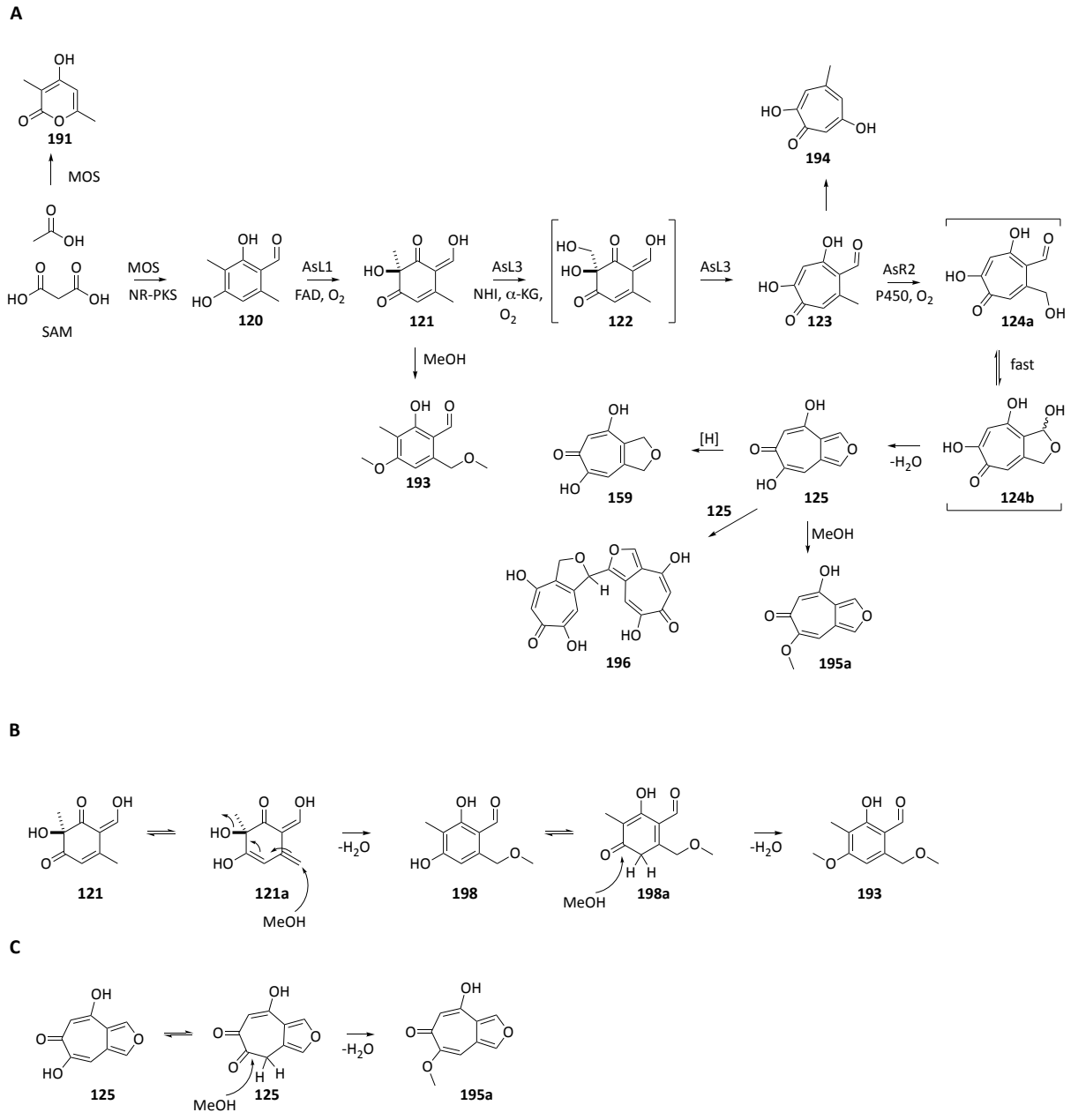
Elucidation of the biosynthesis of xenovulene A **101** in *A. strictum* with attempted KO, CRISPR/Cas9 and silencing strategies proved very difficult (Chapter 4). However, transcriptomic data (Chapter 3.2) allowed neat BGC cluster boundaries to be set and thus enabled the heterologous expression of the whole *aspks1* BGC.

All genes of the *aspks1* BGC, except transporters and transcription factors, were successfully cloned from *A. strictum* g- and cDNA into several vectors of the pTYGS family. Through subsequent protoplast transformation of *A. oryzae* NSAR1 with 16 different vector combinations the secondary metabolite production of transformants harbouring 16 different combinations of genes was analysed.

5.7.1 Tropolone Biosynthesis in *A. strictum*

Sequential transformation and expression of *aspks1*, *asL1*, *asL3* and *asR2* led to the isolation of intermediates (**120**, **121**, **123**) and shunt products (**191**, **193**, **194**, **125**, **159**, **196**) which were previously observed in tropolone biosynthesis in *T. stipitatus* (Chapter 5.2).[97,98] Proteins encoded by these four genes were predicted *in silico* to be homologous to the tropolone forming enzymes TropA (NR-PKS), TopB (salicylate monooxygenase), TropC (non-heme Fe^{II} dependent dioxygenase) and TropD (P450) from *T. stipitatus*.[97] Previous expression of *aspks1*[102] and *aspks1*+*asL1*[104] already confirmed this homology and was re-established in this work.

Further co-expression of *asL3* and *asR2* and isolation of intermediate **123** and shunt **125** confirmed these protein homologies for the first time. Thus, the *aspks1* BGC encodes a full set of tropolone forming proteins. MOS (NR-PKS, encoded by *aspks1*) releases **120**, which is oxidatively dearomatised by the salicylate monooxygenase AsL1 to give enone **121** (Scheme 5.7A). This intermediate undergoes further hydroxylation (**122**) and ring expansion to stipitaldeyhde **123** by AsL3, a non-heme Fe^{II} dependent dioxygenase. The P450 monooxygenase AsR2 hydroxylates the methyl group and most likely forms **124a**, which is in equilibrium with its hemiacetal **124b**. However, this compound was not directly observed, but instead three previously reported shunts (**125**, **159**, **196**) probably arising from it. The two newly observed compounds **193** and **195a** are most likely adducts formed with methanol as solvent (Scheme 5.7B and 5.7C).



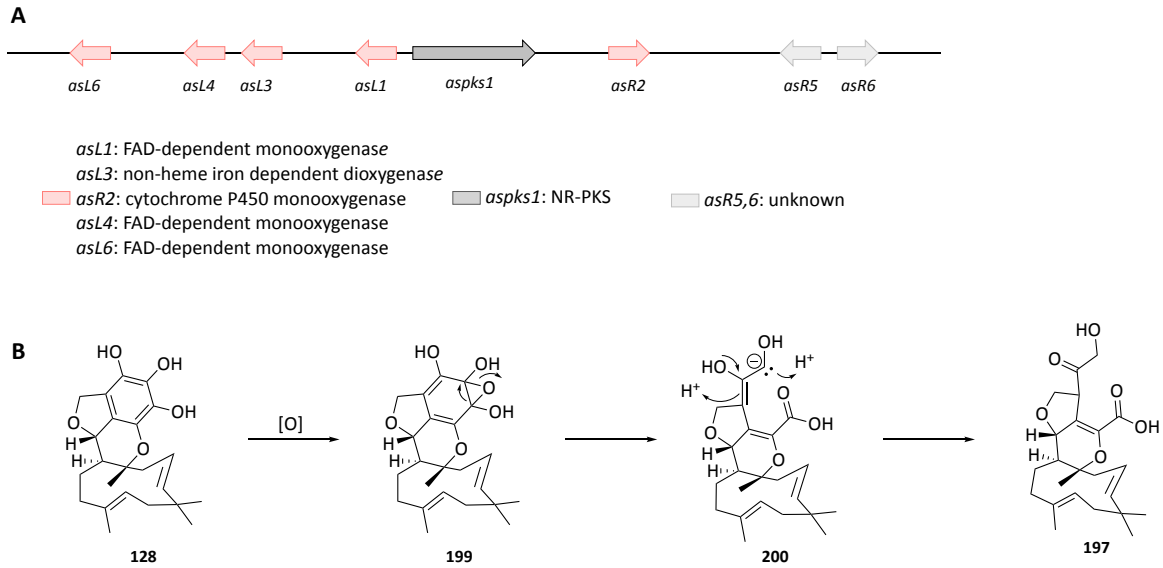
Scheme 5.7 Tropolone biosynthesis in *A. oryzae* NSAR1 by expression of genes encoded in the *aspks1* BGC from *A. strictum*. **A**, Porposed biosynthetic pathway. **B**, Proposed addition of two molecules methanol to **121**. **C**, Proposed addition of methanol to **125**.

5.7.2 Heterologous Production of Xenovulene A 101 and Minimal *aspks1* BGC

In a further experiment an additional seven genes from the *aspks1* BGC (eleven genes in total) were co-expressed which led to the production of xenovulene A **101** and compound **197** (Figure 5.38A). The latter, **197**, most likely arises as degradation product of phenolic meroterpenoids rather than as a product of xenovulene A **101** itself as it has 23 carbon atoms, one more than **101** (Figure 5.38B).

This result confirmed that although the *aspks1* BGC does not encode a known class of terpene cyclase, it encodes all proteins to form the merotperpenoid **101**. This is consistent with previous

in silico analysis of the *A. strictum* transcriptome, where no co-regulated terpene gene could be identified (Chapter 3.3.1). In order to determine the minimal *aspks1* BGC a KO by expression method was employed and this identified four additional genes which are crucial for **101** biosynthesis: *asL4*, *asL6*, *asR5* and *asR6*.



Scheme 5.8 Xenovulene A **101** biosynthesis in *A. oryzae* NSAR1. **A**, Xenovulene A **101** minimal BGC. **B**, Proposed oxidative degradation of **128** to shunt **197**, observed in *A. oryzae* NSAR1 only.

5.7.3 Humulene Production and Joining to Polyketide

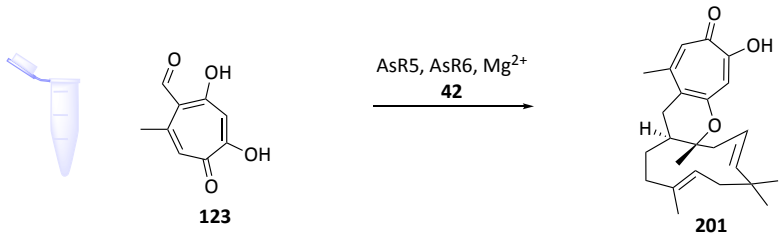
Co-expression of core tropolone encoding genes (*aspks1*, *asL1*, *asL3* and *asR2*) with *asR5* and *asR6* in *A. oryzae* NSAR1, both encoding proteins of unknown function, led to the production of meroterpenoid **154**. This strongly suggested that these novel proteins catalyse the two cryptic steps in meroterpenoid formation: cyclisation of FPP **42** to **106** and attachment of humulene **106** to the polyketide precursor.

To probe this hypothesise, both proteins (AsR5 and AsR6) were solubly expressed in *E. coli* and *in vitro* assays were conducted (by Carsten Schotte). In the presence of Mg^{2+} AsR6 was shown to cyclise FPP **42** to humulene **106**. AsR6 is a protein that possesses no DDxxD or NSE motif, which is characteristic for class I terpene cyclases. It furthermore shows no homologies to any known terpene cyclase (Chapter 3.3.2). Thus AsR6 represents an unprecedented terpene cyclase that is first described in xenovulene A **101** biosynthesis. Further mechanistic studies and identification of active site residues have to await the crystallographic analysis of the protein.

In a further attempt to characterise AsR5, another protein of unknown function, *in vitro* enzyme assays were carried out. Incubation of neither AsR5, AsR6, FPP, **125** and Mg^{2+}/Ca^{2+} or AsR5, humulene **106** and **125** showed conversion to the tropolone meroterpenoid **154** (Scheme 5.5).

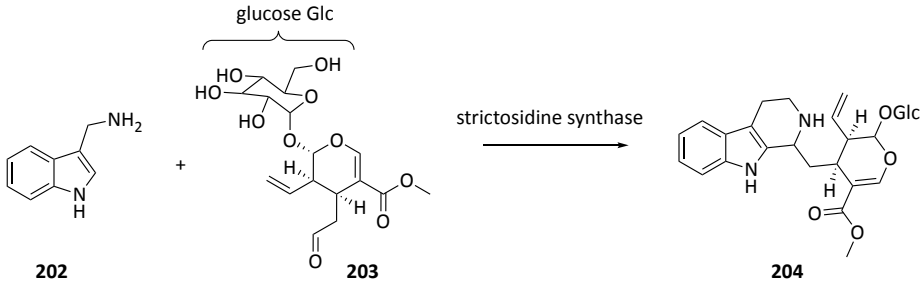
Although the protein was expressed and purified in soluble form, the obtained yields were poor and the activity could not be confirmed. This could be due to several reasons such as lacking a co-factor or the presence of incorrect substrates.

This suggests that reaction conditions for AsR5 in *in vitro* assays need further optimisation. Testing different co-factors and polyketide precursors could identify the true substrates and reconstitute the biosynthesis of meroterpenoid **154** *in vitro*. Comparison of xenovulene A **101** and epolone B **114** suggests stipitaldehyde **123** as the likely polyketide substrate. In parallel systems the related meroterpenoids **110**-**116** are formed despite the lack of the cytochrome P450 oxygenase catalysed hydroxylation at the methyl group in **123**. Further, *A. oryzae* NSAR1 expression studies lacking *asR2* (P450) could also help to elucidate the polyketide precursor.



Scheme 5.9 Proposed *in vitro* assay with **123** as substrate.

AsR5 was predicted to be a Ca^{2+} dependent phosphotriesterase homologue and a model proposed the tertiary structure of a six bladed propeller protein (Figure 3.10, Chapter 3.3.2). Strictosidine **204** synthase, a plant enzyme that catalyses an intermolecular Pictet-Spengler reaction of tryptamine **202** and secologanin **203**, was shown to crystallise as a similar six bladed propeller.[225]



Scheme 5.10 Pictet-Spengler reaction catalysed by strictosidine synthase.

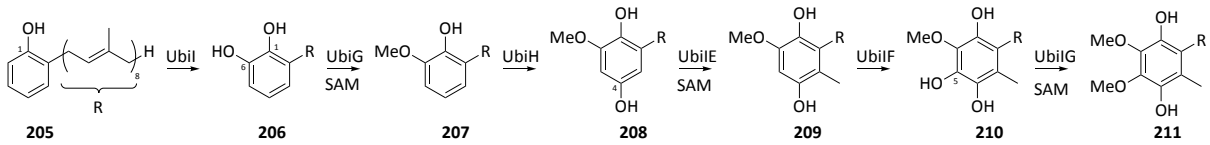
Protein homology modelling of AsR5 and strictosidine synthase using SWISS-MODEL showed that substrate bindings sites are not conserved.[156–158] However, six bladed propeller proteins are known to mediate protein-protein interaction and to display diverse functions. These proteins can trap substrates at their top face or between domains.[226] During the biosynthesis of xenovulene A **101** two advanced biosynthetic precursors have to react. It is possible that the tertiary structure of an AsR5 six bladed propeller traps humulene **106** and the required polyketide precursor in close

proximity and induces a conformational organisation to enable the proposed hetero Diels Alder reaction.

5.7.4 Ring Contraction from Tropolone to Cyclopentenone

The final steps of **101** biosynthesis are two ring contractions to form a cyclopentenone from tropolone meroterpenoids (Scheme 1.18A). The *A. oryzae* NSAR1 expression studies showed that *asL4* and *asL6*, both encoding FAD dependent monooxygenases, are crucial for this chemistry. When both are not expressed only meroterpenoid **154** can be observed. However, when either *asL4* or *asL6* are KOe, traces of **101** and the two different phenolic structural isomers **109a** and **109b** are observed. These findings strongly suggest that AsL4 and AsL6, which share 33.4% protein sequence identity, catalyse the two ring contractions and can, at least partially, complement each other.

Conserved domain analysis[141] of AsL4 and AsL6 showed a high similarity to UbiH, an FAD dependent monooxygenase involved in ubiquinone 8 **211** formation (Scheme 5.11). UbiH, UbiI and UbiF belong to a family of regioselective aromatic hydroxlyases. In *E. coli* UbiI hydroxylates the aromatic core at C-6, UbiH at C-4 and UbiF at C-5 (Scheme 5.11). Although, catalysing very similar chemistry the overall sequence identity of these proteins is only between 29 – 38%.[83,227]

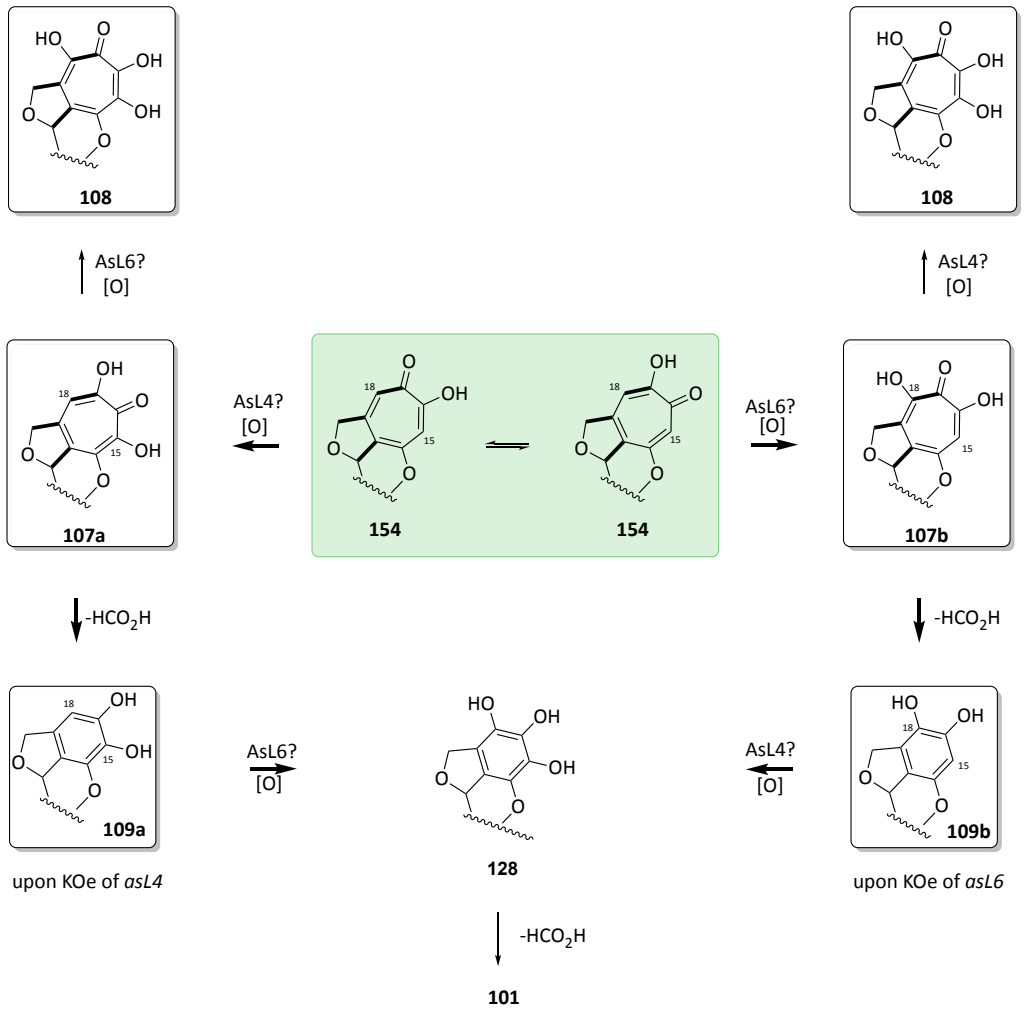


Scheme 5.11 Regioselective aromatic hydroxylations catalysed by UbiI, UBiH and UbiF.

The isolated *A. strictum* WT compounds **107a** and **107b**, as well as the exclusive observation of meroterpenoid **109a** or **109b**, when either *asL4* or *asL6* is KOe in *A oryzae* NSAR1 support a regioselective oxidation mechanism for the two monooxygenases AsL4 and AsL6 (**Scheme 5.12**). In regard to observed phenolic meroterpenoid **109a** upon KOe of *asL4*, it could be proposed that the enzymatically less preferred intermediate **109a** accumulates, which would lead to the conclusion that AsL4 could oxidise C-15. *Vice versa*, when *asL6* is KOe, phenolic compound **109b** accumulates and suggests AsL6 to oxidise C-18.

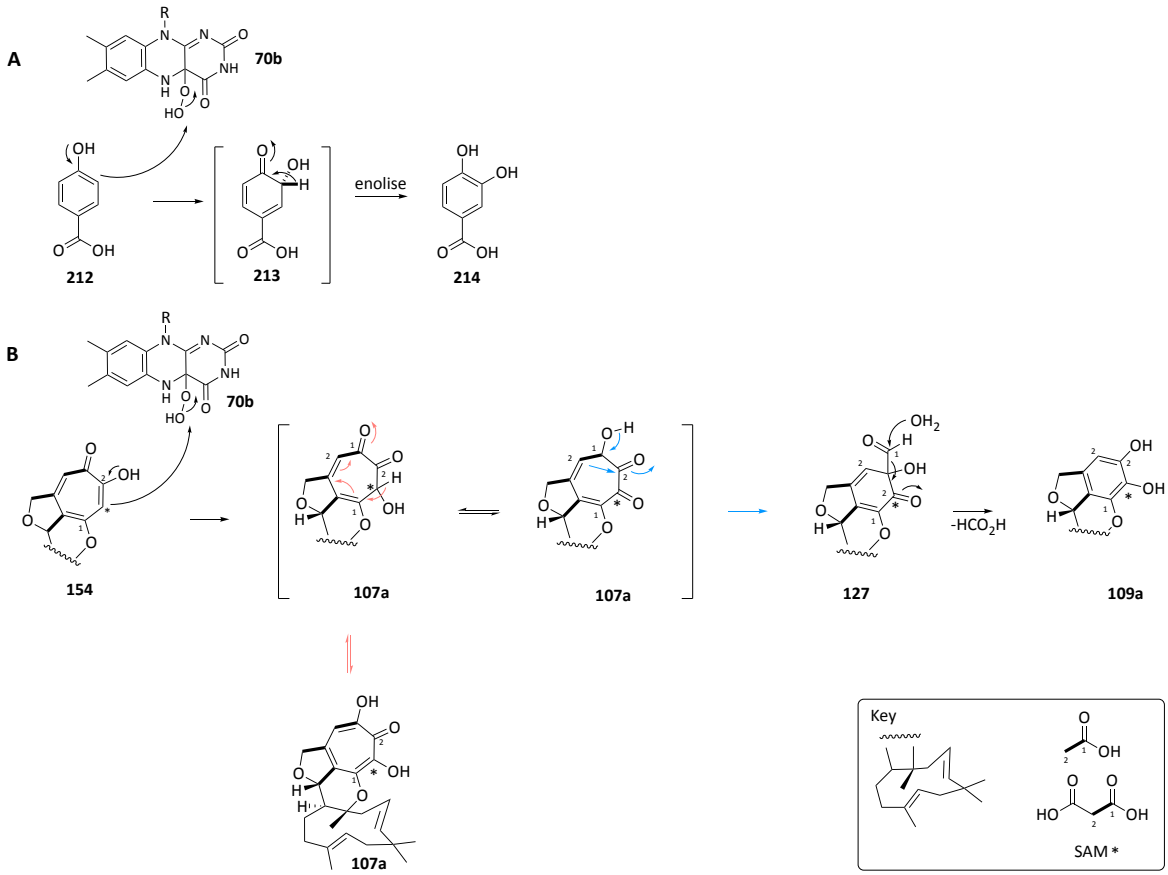
Mechanism and enzymes for similar ring contractions are unknown, but aromatic hydroxylations and mechanisms of FAD dependent monooxygenases have been reported.[228] A detailed mechanism was investigated for *p*-hydroxybenzoate 3-hydroxylase.[229] The direct hydroxylation by a nucleophilic attack of the activated *p*-hydroxybenzoic acid **212** to the hydroperoxoflavin **70b** resulted in the respective dihydroxylated compound **214** *via* a cyclohexadieneone intermediate **213**

(Scheme 5.13A).[230] This direct hydroxylation would explain the intermediates **107a** and **107b** observed in *A. strictum* WT. In addition, it is in accordance with Simpson's previously proposed ring contraction mechanism for xenovulene A **101** and supported by the feeding experiments (Scheme 5.13B).[84]



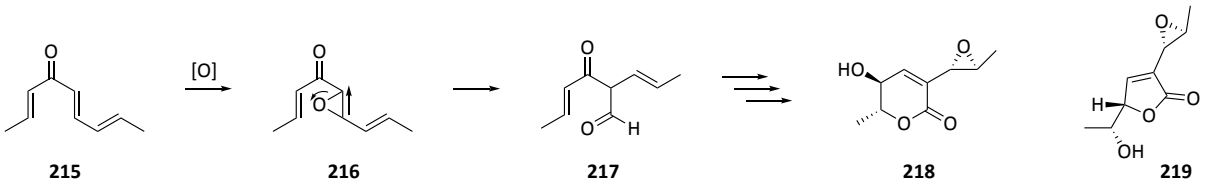
Scheme 5.12 Regioselective hydroxylation and ring contraction by AsL4 and AsL6. Framed compounds are characterised in Chapter 2.

FAD dependent monooxygenases are also often linked to epoxidations of alkenes such as reported for austinol **84** and related meroterpenoids **93**-**95** (Chapter 1.2, Scheme 1.13).[72,76,77] In the biosynthesis of aspyrone **218** and asperlactone **219** such an epoxide **216** is proposed to rearrange via an 1,2 alkyl migration to **217** (Scheme 5.14).[231]



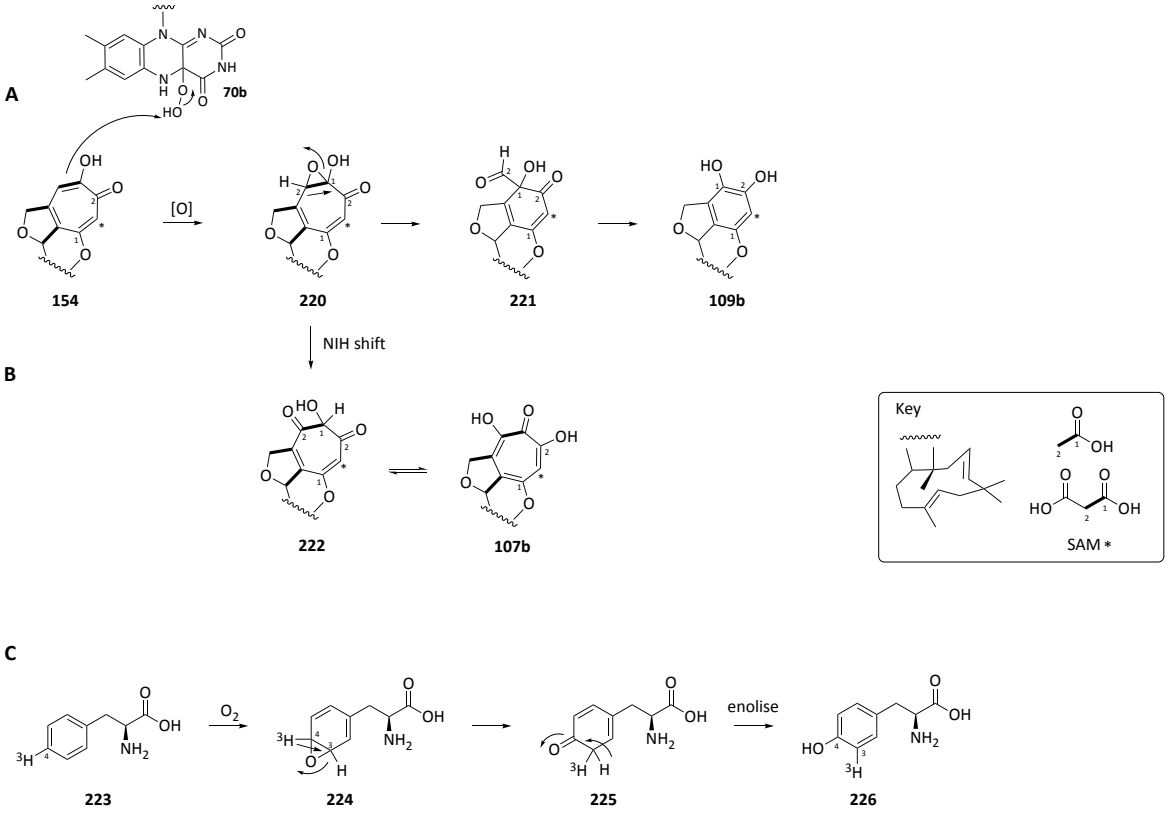
Scheme 5.13 FAD dependent aromatic hydroxylation. **A**, Hydroxylation of *p*-hydroxybenzoic acid **212**. **B**, Simpson's proposed hydroxylation and ring contraction of tropolone meroterpenoid **154**.[84]

An alternative epoxidation dependent oxidation of **154** to **107a** or **107b**, could also explain the ring contraction towards phenolic compounds **109a** and **109b** through 1,2 alkenyl migration (Scheme 5.15A). The corresponding 1,2 hydride migration, known as the NIH shift, would explain and define compounds **107a** and **107b** as shunts (Scheme 5.15B).[232] The NIH shift is observed in hydroxylations of a variety of aromatic compounds by phenylalanine-, tryptophan or tyrosine hydroxylases (Scheme 5.15C).[232,233]



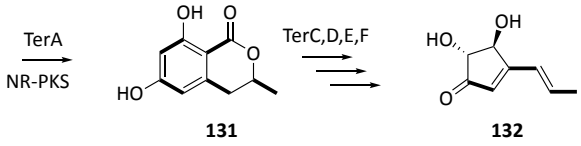
Scheme 5.14 Porposed epoxide rearrangement in the biosynthesis of aspyrone **218** and asperlactone **219**.

However, attempts to correlate a possible epoxide mechanism with the results of the feeding experiments of xenovulene A **101** resulted in either incorrect labelling of **101** or in the requirement of one or two additional oxidations.



Scheme 5.15 Proposed FAD dependent epoxidation. **A**, Ring contraction and **B**, NIH 1,2 hydride shift. **C**, Classic NIH shift of phenylalanine hydroxylase on 4-tritiophenylalanine **223**.[232]

A similar ring contraction mechanism is proposed in terrein **132** biosynthesis (Scheme 5.16). Although the molecular steps have not yet been characterised, a BGC responsible for **132** biosynthesis has been reported.[101] Protein sequences of two putative FAD dependent monooxygenases (encoded by *terC* and *terD*) essential for **132** biosynthesis, were compared to AsL4 and AsL6 using EMBOSS Needle.[153] Pairwise protein identity was between 14-17 % and similarity less then 27% of all compared sequences, thus indicating no close relation for ring contractions in **101** and **132**.



Scheme 5.16 Terrein **132** biosynthesis in *A. terreus*.

Taken together these findings suggest that AsL4 and AsL6 are regioselective aromatic hydroxylases. *In vivo* evidence showed that the order of oxidations can alter. This indicates that AsL4 and AsL6 can hydroxylate both tropolones and phenolic compounds, even though a reduced efficiency is observed. However, mechanistic studies of a possible ring contraction mechanism have to await recombinant protein production to enable *in vitro* assays and crystallisation experiments.

6 Conclusion and Outlook

The major focus of this project was the identification and investigation of a BGC involved in xenovulene A **101** biosynthesis. For this reason, genome, transcriptome and secondary metabolomes of *A. strictum* were analysed in detail.

By KO of the core PKS encoding-gene in *A. strictum* the 49 kb large *aspks1* BGC was confirmed as responsible for xenovulene A **101** production (Chapter 4). By heterologous co-expression all genes crucial for **101** biosynthesis were identified and further characterised (Chapter 5).

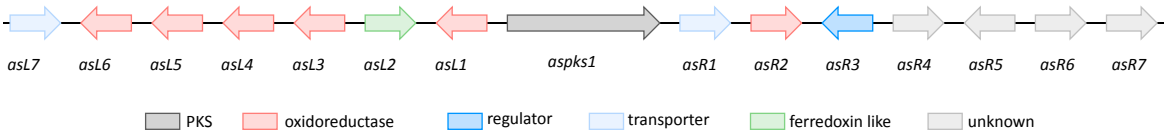
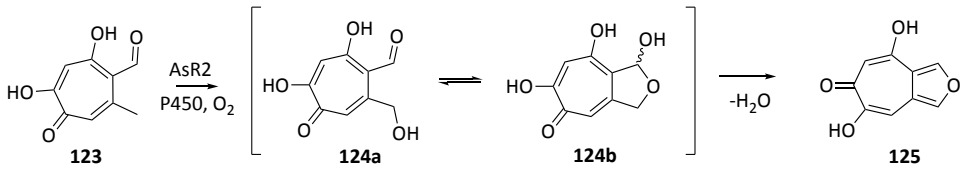


Figure 6.1 *aspks1* BGC, encoding xenovulene A **101** biosynthesis.

Heterologous co-expression of xenovulene A **101** early genes (*aspks1*, *asL1*, *asL3*, *asR2*) and also tropolone precursor **120**, **121** and **125** in *A. strictum* WT chromatograms showed for the first time that the *aspks1* BGC encodes a tropolone pathway homologous to *T. stipitatus* (Chapter 2.8 and 5). The direct intermediate produced by the cyctochrome P450 monooxygenase AsR2, which is proposed to be compound **124a** in equilibrium with its hemiacetal **124b**, was not observed. Also in the *T. stipitatus* pathway this likely intermediate is shunted to **125** probably very quickly (Scheme 6.1).[98] Expression of the *T. stipitatus* AsR2 homologue yielded insoluble protein, which made *in vitro* assays impossible.[97] However, in order to identify the true AsR2 product *in vitro* enzyme assays are most promising.



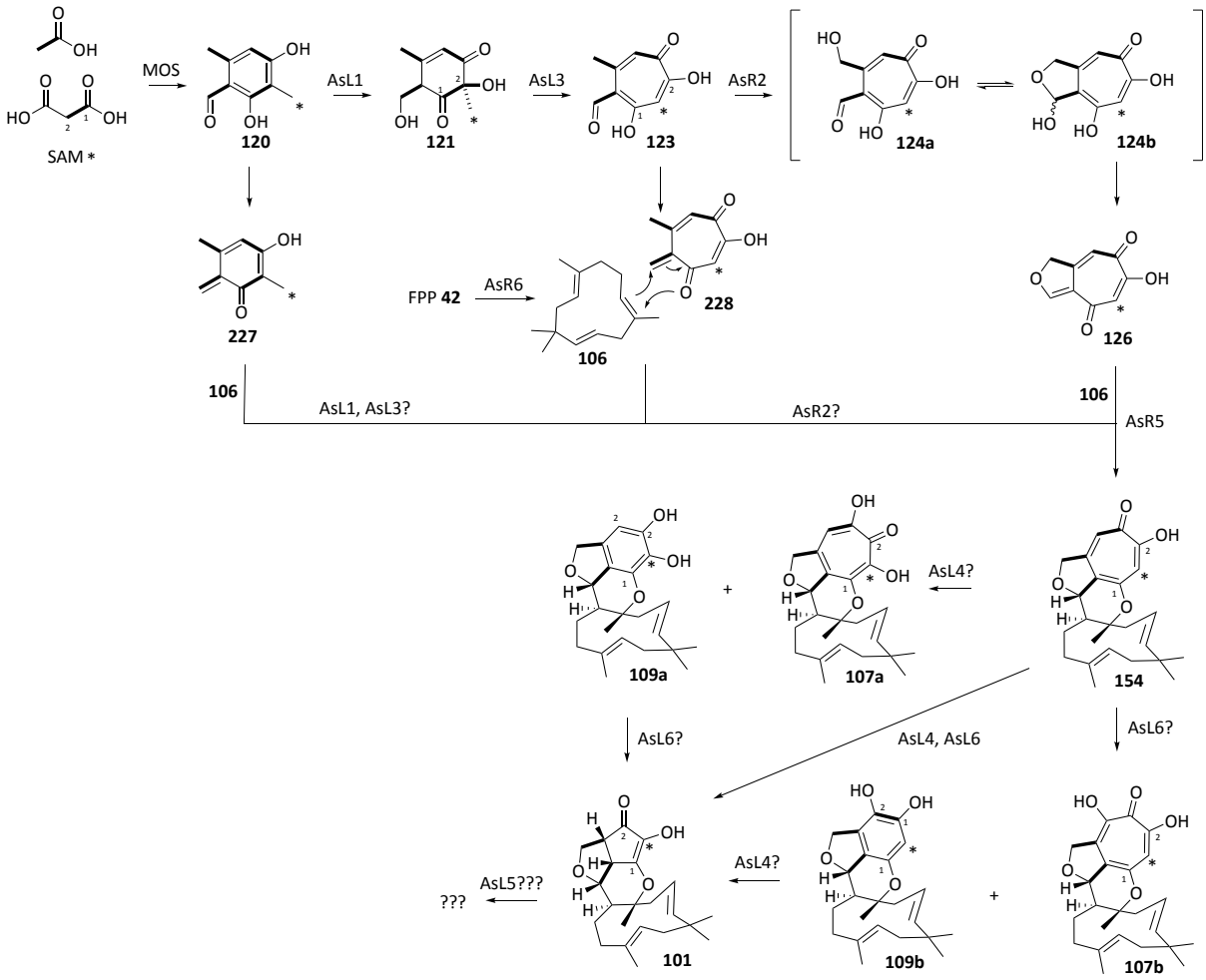
Scheme 6.1 Proposed hemiacetal **124a** intermediate upon cytochrome P450 monooxygenase hydroxylation of the methyl group.

Further heterologous expression of different gene sets of the *aspks1* BGC in *A. oryzae* NSAR1 identified two genes encoding proteins of unknown function, which are essential for meroterpenoid production. *In vitro* enzyme assay confirmed one (AsR6) as unprecendent humulene synthase (Chapter 5.5). Although, this work showed that the protein is Mg^{2+} dependent, such as classical terpene cyclases, no active site residues for Mg^{2+} coordination were identified *in silico*. Future crystallisation of this protein will identify active site residues and possibly reveal more about a structural relationship to classical terpene cyclases.

The other unknown protein (AsR5) is very likely involved in the fusion of polyketide and terpene moiety (Chapter 5.5). For this reaction a hetero Diels Alder reaction has been proposed and its feasibility was shown *in vitro* by Baldwin and co-workers (Chapter 1.3.4).[111,113] Only a few examples of enzymes catalysing a DA or hDA reaction are reported in the literature and a variety of protein tertiary structures are used.[117] However, the structural prediction of this protein is a six-bladed propeller and these are known for the capacity to trap compounds.[226] This enzyme could trap a polyketide and a terpene moiety in close proximity and conformation to induce the hDA reaction. To prove a hetero Diels Alder mechanism, the true polyketide substrate would have to be identified and the protein activity confirmed *in vitro* (Scheme 6.2). However, protein kinetics alone will not show the nature of a concerted Diels Alder reaction. Additional molecular dynamic simulations under consideration of a protein crystal structure or substrate transitions states, which were used to show the concerted Diels Alder reaction in abyssomicin and leporin B **64** biosynthesis, could finally confirm the mechanism.[64,234]

The two FAD dependent monooxygenases AsL4 and AsL6 were identified as crucial for the proposed ring contractions (Chapter 5.4 and 5.6). It is likely that both enzymes catalyse regioselective aromatic hydroxylations or epoxidations and a subsequent ring contraction each (Scheme 6.2). However, no similar mechanism is elucidated on a molecular level. It was initially proposed that the mechanism and enzymes involved in xenovulene A **101** ring contractions are similar to those from terrein **132** biosynthesis, but sequence comparison of encoded protein did not identify close homologues. For further elucidation of a mechanism, substrates, intermediates and shunt products these proteins should be recombinantly expressed and *in vitro* assays conducted.

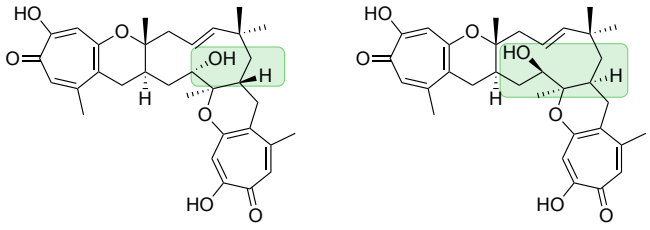
In addition, one gene, *asL5*, encoding a short chain dehydrogenase is part of the *aspks1* BGC and its expression is up-regulated under **101** producing conditions (Chapter 3.2). Nevertheless, the protein was found to be unnecessary for **101** biosynthesis. Short chain dehydrogenases can catalyse oxidations and reductions of a variety of substrates (Chapter 1.1.3).[63] Often ketone to alcohol and *vice versa* reactions are observed. In order to identify its activity *in vitro* assays with different intermediates from the xenovulene pathway could be conducted. Maybe xenovulene A **101** itself is the subject of a further transformation (Scheme 6.2).



Scheme 6.2 Proposed xenovulene A **101** biosynthesis based on results gained this work.

Taken together, we were able to identify a BGC for xenovulene type meroterpenoids for the first time. Even beyond that we could assign likely functions for the individual proteins and propose a more detailed biosynthetic route (Scheme 6.2).

These results will simplify the identification of similar BGC in different fungi as we showed by the identification of a homologous BGC in *A. thermomutaus* (Chapter 3.3.2). Genome sequencing of fungal strains producing similar meroterpenoids (Chapter 1.3) and *in silico* identification of homologous *aspks1* BGC will enable the elucidation of the biosynthesis of an entire class of meroterpenoids. Of special interest hereby is, how or whether the stereoselectivity of the addition of the second tropolone is controlled. Although, the two diastereomers pycnidione **110** and eupenifeldin **111** are reported these compounds have been isolated from different fungi, indicating an enzyme controlled reaction.



pycnidione **110** eupenifeldin **111**

Figure 6.2 The diastereomers pyndidione **110** and eupenifeldin **111**.

A rational genetic engineering strategy where genes of the xenovulene A **101** pathway are co-expressed with those of related meroterpenoids such as the epolone A **113** pathway could lead to novel compounds. A similar approach by expressing chimeric PKS-NRPS hybride system was successfully used to produced novel cytochalasins.[235] If successfully employed in the xenovulene meroterpenoids case, this could expand the chemical diversity of these natural products and may grant access to novel bioactive compounds.

7 Experimental

All chemicals and media ingredients used in this work were purchased from Carl Roth, Sigma-Aldrich, VWR, Merck, Duchefa Biochemie, Formedium, Acros, Alfa Aesar, abcr and Thermo Fisher Scientific unless denoted differently.

7.1 Media, Antibiotics, Buffers

Deionised water was further purified by a GenPure Pro UV/UF milipore device from Thermo Scientific and used to prepare all media, antibiotics and buffers unless denoted differently. Media and Buffers were sterilised at 120 °C for 15 min using a Systec VX150 or a Classic Prestige Medical 2100 autoclave. Antibiotics were prepared at concentrated stock solutions and filter sterilised using a 45 µM syringe filter with a cellulose acetate membrane.

7.1.1 Media

Table 7.1 Components of media and agar used in this work are listed.

Media and Agar	Concentration/% (w/v)	Ingredient
2 tryptone yeast medium (2TY)	1.6	tryptone
	1.0	yeast extract
	0.5	NaCl
A. strictum seed agar (ASSA)	1.10	D(+)-glucose monohydrate
	1.50	glycerol
	1.50	polypeptone
	0.30	sodium chloride
	0.50	malt extract
	0.10	Tween 80
	1.50	agar
		pH 6.0 with H_2SO_4
A. strictum seed medium (ASSM)	1.10	D(+)-glucose monohydrate
	1.50	glycerol
	1.50	polypeptone
	0.30	sodium chloride
	0.50	malt extract
	0.10	Tween 80
	0.10	Junlon PW110
		pH 6.0 with H_2SO_4
A. strictum production medium (ASPM)	4.74	D(+)-maltose monohydrate
	0.70	yeast extract
	0.10	Tween 80
		pH 6.0 with H_2SO_4
CM	2.00	Cottonseed flour
	10.0	Lactose monohydrate
Czapek Dox with sorbitol, soft agar (CD+S soft agar)	3.50	Czapek Dox broth
	18.22	D-sorbitol
	0.80	Agar
		prepare with deionised water for *A. strictum* transformation

Media and Agar	Concentration/% (w/v)	Ingredient
Czapek Dox with sorbitol (CD+S)	3.50	Czapek Dox broth
	18.22	D-sorbitol
	1.50	Agar
		prepare with deionised water for *A. strictum* transformation
Czapek Dox with supplements (CZD)	3.50	Czapek Dox broth
	4.68	sodium chloride
	0.10	ammonium sulfate
	0.05	adenine
	0.15	L-methionine
	1.50	agar
Czapek Dox with sorbitol and supplements (CZD+S)	3.50	Czapek Dox broth
	18.22	D-sorbitol
	0.10	ammonium sulfate
	0.05	adenine
	0.15	L-methionine
	1.50	agar
Czapek Dox with sorbitol and supplements soft agar (CZD+S soft agar)	3.50	Czapek Dox broth
	18.22	D-sorbitol
	0.10	ammonium sulfate
	0.05	adenine
	0.15	L-methionine
	0.80	agar
CZD without methionine (CZD1)	3.50	Czapek Dox broth
	4.68	sodium chloride
	0.10	ammonium sulfate
	0.05	adenine
	1.50	agar
CZD+S without methionine (CZD1+S)	3.50	Czapek Dox broth
	18.22	D-sorbitol
	0.10	ammonium sulfate
	0.05	adenine
	1.50	agar
CZD+S soft agar without methionine (CZD1+S soft agar)	3.50	Czapek Dox broth
	18.22	D-sorbitol
	0.10	ammonium sulfate
	0.05	adenine
	0.08	agar
CZD1 without adenine (CZD2)	3.50	Czapek Dox broth
	4.68	sodium chloride
	0.10	ammonium sulfate
	1.50	agar
CZD1+S without adenine (CZD2+S)	3.50	Czapek Dox broth
	18.22	D-sorbitol
	0.10	ammonium sulfate
	1.50	agar
CZD1+S soft agar without adenine (CZD2+S soft agar)	3.50	Czapek Dox broth
	18.22	D-sorbitol
	0.10	ammonium sulfate
	0.08	agar
DPY	2.00	dextrin from potato starch
	1.00	polypeptone
	0.50	yeast extract
	0.50	KH_2PO_4
	0.05	$MgSO_4$
	2.50	agar

Media and Agar	Concentration/% (w/v)	Ingredient
Lysogeny broth (LB)	0.50	yeast extract
	1.00	tryptone
	0.50	NaCl
LB agar	0.50	yeast extract
	1.00	tryptone
	0.50	NaCl
	1.50	agar
Malt extract agar (MEA)	1.28	malt extract
	0.08	peptone ex soya
	0.24	glycerol
	0.28	dextrin from potato starch
	1.50	agar
Malt extract media (ME)	1.28	malt extract
	0.08	peptone ex soya
	0.24	glycerol
	0.28	dextrin from potato starch
Potato dextrose broth agar (PDA)	2.40	potato dextrose broth
	1.50	agar
Potato dextrose broth (PDB)	2.40	potato dextrose broth
Supplement mixture minus uracil (SM-URA)	0.17	yeast nitrogen base
	0.50	$(NH_4)_2SO_4$
	2.00	D(+)-glucose monohydrate
	0.077	complete supplement mixture minus uracil
	1.50	agar
Super optimal broth (SOB)	0.5	yeast extract
	2.00	tryptone
	0.06	NaCl
	0.02	KCl
SOB with magnesium chloride and glucose (SOC)	93.75 (v/v)	SOB
	1.25 (v/v)	$MgCl_2$ x 6 H_2O (2M)
	5.00 (v/v)	D(+)-glucose (20%)
Solution A	4.00	$NaNO_3$
	4.00	KCl
	1.00	$MgSO_4$
	0.02	$FeSO_4$ x 7 H_2O
Solution B	2.00	K_2HPO_4
Starch M	2.00	starch from potato
	1.00	polypeptone
	5.00	solution A
	5.00	solution B
Yeast maltose glucose media (YMG)	0.40	D(+)-glucose monohydrate
	0.40	yeast extract
	1.00	malt extract
YPAD	1.00	yeast extract
	2.00	tryptone
	2.00	D(+)-glucose monohydrate
	0.03	adenine

7.1.2 Antibiotics

Table 7.2 Concentrations of stock solutions and working concentrations of used antibiotics.

Antibiotic	Stock concentration in mg/mL	Final concentration in µg/mL
Kanamycin	50 in ddH_2O	50
Carbenicillin	50 in ddH_2O	50
Chloramphenicol	34 in undenatured ethanol	34
Hygromycin B	50 in ddH_2O	50-150
Geneticin	50 in ddH_2O	50-150

7.1.3 Buffers and Solutions

Table 7.3 Components of buffers and solutions used in this work are listed.

Buffer	Concentration	Components
gDNA extraction buffer	10 mM 10 mM 0.5%	Tris-HCl, pH 8 EDTA, pH 8 SDS (w/v)
50x Tris-Acetate-EDTA (TAE)	2 M 1 M 50 mM	Tris-HCl Acetic acid EDTA
6x DNA Loading Dye	0.1% (w/v) 0.1% (w/v) 30% (v/v)	bromophenol blue xylene cyanol glycerol
Coomassie Bleach	25% (v/v) 10% (v/v)	acetic acid isopropanol
Coomassie Dying Solution	25% (v/v) 10% (v/v) 0.1% (w/v)	acetic acid isopropanol Coomassie Brilliant Blue
Fungal transformation solution I	10 mM 0.8 M NaCl 50 mM	$CaCl_2$ NaCl Tris-HCl, pH 7.5
Fungal transformation solution II	50 mM 0.8 M 50 mM 60%	$CaCl_2$ NaCl Tris-HCl, pH 7.5 PEG3350 (w/v)
Protein Elution Buffer	50 mM 150 mM 50 - 500 mM 10%	Tris-HCl, pH 8.0 NaCl imidazole glycerol
Protein Resuspension Buffer	50 mM 150 mM 20 mM 10%	Tris-HCl, pH 8.0 NaCl imidazole glycerol
Protein Storage Buffer	50 mM 150 mM 10%	Tris-HCl, pH 8.0 NaCl glycerol

Buffer	Concentration	Components
4x SDS Loading Buffer	250 mM 8% (w/v) 40% (v/v) 10% (v/v) 2.5 mg/L	Tris-HCl, pH 6.8 sodium dodecyl sulfate (SDS) glycerol β-mercaptoethanol bromophenol blue
10x SDS Running Buffer	25 mM 192 mM 1% (w/v)	Tris gylcerole SDS pH 8.4-8.9
50x TAE	2 M 50 mM	Tris acetate EDTA pH 8.3
10xTE	100 mM 10 mM	Tris-HCl, pH 8 EDTA
A. oryzae NSAR1 protoplasting solution	10 mg/mL 0.8 M	lysing enzyme from *Trichoderma harzianum* (Sigma-Aldrich) NaCl
A. strictum protoplasting solution	20 mg/mL 10 mg/mL 0.8 M 10 mM	lysing enzyme from *Trichoderma harzianum* (Sigma-Aldrich) driselase from *Basidiomycetes sp.* (Sigma-Aldrich) NaCl K_2HPO_4/KH_2PO_4, pH 6

7.2 Molecular Biology

Standard restrictions enzymes and size standards were purchased from New England Biolabs (NEB) unless denoted differently.

7.2.1 Agarose Gel Electrophoresis

Quality and quantity of DNA and RNA samples were analysed by horizontal agarose gel electrophoresis. Depending on the samples the agarose concentration was adjusted in a range of 0.8 – 2.0% (w/v) agarose in 0.5x TAE buffer. Agarose and buffer were molten in a microwave, gel stain (1 µL/25 mL, Roti-GelStain, Roth) was added and poured on a tray with comb fixed in a gel caster system (Bio-Rad). For the analysis of RNA samples, all buffers were prepared freshly with double distilled water. After solidifying of the agarose, combs were removed and the tray transferred to a Sub-Cell GT agarose gel electrophoresis system (Bio-Rad). 6x DNA loading dye was added to samples to a final concentration of 1x prior to transferring it to the gel. 1 µL of 1 kb or 100 bp DNA ladder was loaded alongside and the gel was run for 30-90 min at 100 V using 1x TAE as running buffer. DNA bands were analysed under UV light with a gel documentation system (Gel Doc XR+, Bio-Rad) and the software Image Lab (Bio-Rad).

7.2.2 PCR

For DNA amplification two types of polymerases were available. The proofreading Q5 High Fidelity 2x Master Mix was used when sequence identity was crucial, for other applications the One *Taq* Quick-Load 2x Master Mix was sufficient. Reactions were carried out in volumes of 10 – 100 µL according to the manufacturer's protocol and diluted with double distilled millipore water (ddH_2O). Depending on the sample gDNA, complementary DNA (cDNA) or plasmid DNA was used as template. For *E. coli* colony screening the PCR template was obtained by picking a single colony with a toothpick and transferring it directly to a 10 µL PCR reaction. Annealing temperatures were adjusted to the oligonucleotides. For All PCR reactions a Bio-Rad T100 thermal cycler was used.

Table 7.4 Pipetting scheme for standard PCR reactions.

Component	50 µL Reaction	Final concentration
Q5 High Fidelity 2x Master Mix	25 µL	1x
10 µM forward primer	2.5 µL	0.5 µM
10 µL reverse primer	2.5 µL	0.5 µM
template DNA	variable	> 1000 ng
ddH_2O	to 50 µL	
One *Taq* Quick-Load 2x Master Mix	25 µL	1x
10 µM forward primer	1 µL	0.2 µM
10 µL reverse primer	1 µL	0.2 µM
template DNA	variable	> 1000 ng
ddH_2O	to 50 µL	

Table 7.5 Standard PCR programs for the two different polymerases used.

Program	Step	Temperature	Time
Q5 High-Fidelity 2x Master Mix	Initial Denaturation	98 °C	2 min gDNA, cDNA; 30 sec vector DNA
	30 cycles	98 °C	5 – 10 sec
		50 – 72 °C	10 – 30
		72 °C	20 – 30 sec/kb
	Final Extension	72 °C	2 min
	Hold	8 °C	∞
One *Taq* Quick-Load 2x Master Mix	Initial Denaturation	94 °C	2 min gDNA; 30 sec vector DNA; 10 min colony
	30 cycles	94 °C	PCR
		50 – 72 °C	15 – 30 sec
		72 °C	10 – 30
	Final Extension	72 °C	20 – 30 sec/kb
	Hold	8 °C	2 min

7.2.3 Oligonucleotides

All oligonucleotides used in this work were designed using the software geneious and purchased at Sigma Aldrich.

Table 7.6 All oligonucleotides used in this work are listed with their corresponding cox group ID.

Cox group ID	Sequence 5'-3'
11	TCCGTAGGTGAACCTGCGG
12	TCCTCCGCTTATTGATATGC
22	CTAGAAAGAAGGATTACCTC
23	CTGTCGAGAAGTTTCTGATCG
37	GCAAAGTACATCTGCGAGCG
38	CAAGAACCCCACCTCCTTCC
81	GCTCAGCAAGGTTGGGTTTG
82	CGTTGTGGATGATGTGCGTC
83	ATGGCAGCTCATGGGCAAAC
84	CACTAGAGGATCCCCATCATGCGTGTTCATGTAACGAACCCTG
85	CACATCTCCACTCGACCTGGCTCTGGCCAAGATTACCG
86	TCACAACAAGAACCCCACCTCC
124	CAGGGTTCGTTACATGAACACGCATGATGGGGATCCTCTAGTG
125	CGGTAATCTTGGCCAGAGCCAGGTCGAGTGGAGATGTG
145	CAGGAAACAGCTATGACC
146	TGTAAAACGACGGCCAGT
147	TGCTTGGAGGATAGCAACCG
148	GGGGATGACAGCAGTAACGA
179	GCCAACTTTGTACAAAAAAGCAGGCTCCGCGGAATTCATGGCAGCTCATGGGCAAAC
180	TGCCAACTTTGTACAAGAAAGCTGGGTCGGGAATTCCGTCAGGACATTGTTGGAG
181	GCCAACTTTGTACAAAAAAGCAGGCTCCGCGAATTCCTGTCGAGAAGTTTCTGATCG
182	TGCCAACTTTGTACAAGAAAGCTGGGTCGGGAATTCTCACAACAAGAACCCCACCTCC
200	CGTTCCTCGTTAAATCATCACA
201	CAATCTCTACCAGCCCACGA
202	CATTGATGAACACGCCCTTGC
203	GACGAGGATTCTGGTCGAGG
204	CCATGAAGGTCAAGATCATTGC
205	CTCGTCGTACTCCTGCTTGG
206	GTCCGTCTTCGACCACTGG
207	CGTAGTAGTTCTCGACACCAGG
318	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGGCAGCTCATGGGCAAAC
319	CATGTGTCCTCGCAGAAGTTG
320	CACAGAAGCAAGCACGATTC
321	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTCACAACAAGAACCCCACCTC
322	CTGCTGGCTTAACACGTGC
323	GACTGGAGATAGCCGTGTGC
406	CCAGTTCTTCTCGGCGTTC
419	CCACTTCATCGCAGCTTGAC
420	AACATCGCCTCGCTCCAG
421	TTCTTTCAACACAAGATCCCAAAGTCAAAGATGGACAGCCCAGAAGTAT
422	CAGGTTGGCTGGTAGACGTCATATAATCATTTAAAGAGTATAGCCGCC
423	TTTCATTCTATGCGTTATGAACATGTTCCCTTAAAGAGTATAGCCGCC
425	CAGGTTGGCTGGTAGACGTCATATAATCATCTATGGTAGCACTACTGGC
424	ACAGCTACCCCGCTTGAGCAGACATCACCGATGGGCAGCCTCACTGAT

Cox group ID	Sequence 5’-3’
426	TACGACAATGTCCATATCATCAATCATGACCTATGGTAGCACTACTGGC
427	CGACTGACCAATTCCGCAGCTCGTCAAAGGATGGCTCTCGCACAGCAA
428	CAGGTTGGCTGGTAGACGTCATATAATCATTCATTCTTTGTCCGAGCG
434	GCGGCCGCATACCTAGACGCATG
435	GGTACCAGGACTTTGAAGCAGCCCT
436	GCGGCCGCATCCAACATAAGAATAA
437	GGTACCTGTGGAGGTTTGCTACTGG
555	CCTTAAAGGGACGCTGGAGG
556	CCTCCAGTTCTTCTCGGCGTTCTGG
557	CTATCGCCCATCACCCTCAC
558	CGAGACTGAGGAATCCGCTC
559	GGGTTTAAUGCGTAAGCTCCCTAATTGGC
560	GGTCTTAAUGAGCCAAGAGCGGATTCCTC
561	AGTAAGCUCGTCACCGATCTGTGAGCGCCGGCGTTTTAGAGCTAGAAATAGCAAGTTAAA
562	AGCTTACUCGTTTCGTCCTCACGGACTCATCAGACCGATCGGTGATGTCTGCTCAAGCG
563	GTGAAGGAGATGGAGCCGTC
564	GTTGCTTGATCCACAGCGTC
601	AGCTTACUCGTTTCGTCCTCACGGACTCATCAGCGCCTACGGTGATGTCTGCTCAAGCG
602	AGTAAGCUCGTCCGCCTACCTTTTCCGCCTGAGTTTTAGAGCTAGAAATAGCAAGTTAAA
603	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGGCTCTCGCACAGCAATT
604	GGTTGGCTGGTAGACGTCATATAATCATACTCATTCTTTGTCCGAGCGGC
618	GCTCGACGTATTTCAGTGTC
619	GCTCCGTAACACCCAATACG
621	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGCCGCAACTAAAGGTTCT
622	CACTAGAGGATCCCCATCATGATGAGGAGCTTGCTACCGG
623	CCGGTAGCAAGCTCCTCATCATGATGGGGATCCTCTAGTG
624	GGGTTGTCCCTGCGTGTCTCAGAAGAACTCGTCAAGAAGG
625	CCTTCTTGACGAGTTCTTCTGAGACACGCAGGGACAACCC
626	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTCACTCCTTGAGAAGCTCTG
627	CTGGGTTGTCCCTGCGTGTCCTATTCCTTTGCCCTCGGAC
628	GTCCGAGGGCAAAGGAATAGGACACGCAGGGACAACCCAG
629	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATCAAGCGATCGTTGTAATT
630	CACTAGAGGATCCCCATCATGGCCAAGGGGAGTGGCCATT
631	AATGGCCACTCCCCTTGGCCATGATGGGGATCCTCTAGTG
632	CTGCTCGCCGGTGATCTTTCAGAAGAACTCGTCAAGAAGG
633	CCTTCTTGACGAGTTCTTCTGAAAGATCACCGGCGAGCAG
634	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTTGTGTGTAGACGGACGC
635	CACTGCTCGCCGGTGATCTTCTATTCCTTTGCCCTCGGAC
636	GTCCGAGGGCAAAGGAATAGAAGATCACCGGCGAGCAGTG
637	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGACTGTGAAGATCCTTGT
638	CACTAGAGGATCCCCATCATGTCCGCTTCATCCTCAGACTT
639	AAGTCTGAGGATGAAGCGGACATGATGGGGATCCTCTAGTG
640	TTCTTGAACAGCGTGCTGTCAGAAGAACTCGTCAAGAAGG
641	CCTTCTTGACGAGTTCTTCTGACAGCACGCTGTTCAAGAA
642	TGCCAACTTTGTACAAGAAAGCTGGGTCGGTCATGCCTCAAACTCCAGCT
643	AATTCTTGAACAGCGTGCTGCTATTCCTTTGCCCTCGGAC
644	GTCCGAGGGCAAAGGAATAGCAGCACGCTGTTCAAGAATT
645	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGCCGCAACTAAAGGTTCT
646	GGTTGGCTGGTAGACGTCATATAATCATACTCACTCCTTGAGAAGCTCTG
647	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGAGCGCCATTCAAAGACT
648	GGTTGGCTGGTAGACGTCATATAATCATACTTAGATCTGAGCAAGGCCAC
649	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGACTGTGAAGATCCTTGT

Cox group ID	Sequence 5’-3’
650	GGTTGGCTGGTAGACGTCATATAATCATACTCATGCCTCAAACTCCAGCT
653	ATGCCGCAACTAAAGGTTCT
654	TCACTCCTTGAGAAGCTCTG
655	ATCAAGCGATCGTTGTAATT
656	TTGTGTGTAGACGGACGC
657	ATGACTGTGAAGATCCTTGT
658	TCATGCCTCAAACTCCAGCT
659	AGAGGCTATTCGGCTATGAC
660	CCATGATATTCGGCAAGCAG
684	GCCAACTTTGTACAAAAAAGCAGGCTCCGCCAGGTCGAGTGGAGATGTGG
685	TAATCGCTCACCTCAACAGCATGATGGGGATCCTCTAGTG
686	CACTAGAGGATCCCCATCATGCTGTTGAGGTGAGCGATTA
687	CAGAGCTTCTCAAGGAGTGATGTGAAAGATCTAGACAAGA
688	TCTTGTCTAGATCTTTCACATCACTCCTTGAGAAGCTCTG
689	TGCCAACTTTGTACAAGAAAGCTGGGTCGGATGCCGCAACTAAAGGTTCT
690	GTGGCCTTGCTCAGATCTAATGTGAAAGATCTAGACAAGA
691	TCTTGTCTAGATCTTTCACATTAGATCTGAGCAAGGCCAC
692	TGCCAACTTTGTACAAGAAAGCTGGGTCGGATGAGCGCCATTCAAAGACT
693	AGCTGGAGTTTGAGGCATGATGTGAAAGATCTAGACAAGA
694	TCTTGTCTAGATCTTTCACATCATGCCTCAAACTCCAGCT
695	TGCCAACTTTGTACAAGAAAGCTGGGTCGGATGACTGTGAAGATCCTTGT
706	TTTCATTCTATGCGTTATGAACATGTTCCCTCACTCCTTGAGAAGCTCTG
707	ACAGCTACCCCGCTTGAGCAGACATCACCGATGAGCGCCATTCAAAGACT
708	TACGACAATGTCCATATCATCAATCATGACTTAGATCTGAGCAAGGCCAC
709	CGACTGACCAATTCCGCAGCTCGTCAAAGGATGACTGTGAAGATCCTTGT
710	TACGACAATGTCCATATCATCAATCATGACTCACTCCTTGAGAAGCTCTG
711	CGACTGACCAATTCCGCAGCTCGTCAAAGGATGAGCGCCATTCAAAGACT
755	TTCTTTCAACACAAGATCCCAAAGTCAAAATGGCTGCTCCATCACTCGAA
756	TTTCATTCTATGCGTTATGAACATGTTCCCTTTACAGTCCCTTTCTAACG
756	AACAGCTACCCCGCTTGAGCAGACATCACCATGCCTGCCTCACTCCCAGG
758	ACGACAATGTCCATATCATCAATCATGACCCTATCGCCAGTCAAAATTCG
759	GTCGACTGACCAATTCCGCAGCTCGTCAAAATGCCCGTTACTACCCCCAC
760	GGTTGGCTGGTAGACGTCATATAATCATACTTACCCAACAGCAGTTGTTA
761	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGCGTCGCAGTTTTCTTAT
762	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCTAGAAGTGAAAGCCAGTCG
763	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGGGCGAGGCGTGCTCAAA
764	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCTAGGCGATTTCCGCACGCT
771	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGGGCGAGGCGTGCTCAAAG
772	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCTAGGCGATTTCCGCACGCTGC
787	TACGACAATGTCCATATCATCAATCATGACTTACAGTCCCTTTCTAACGC
788	CGACTGACCAATTCCGCAGCTCGTCAAAGGATGCCTGCCTCACTCCCAGG
789	CAGGTTGGCTGGTAGACGTCATATAATCATACCTATCGCCAGTCAAAATT
790	CGACTGACCAATTCCGCAGCTCGTCAAAGGATGCCCGTTACTACCCCCAC
791	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGCCTGCCTCACTCCCAGG
792	TACGACAATGTCCATATCATCAATCATGACCTATCGCCAGTCAAAATTCG
793	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGGGCGAGGCGTGCTCAAA
794	GGTTGGCTGGTAGACGTCATATAATCATACTAGGCGATTTCCGCACGCTG
795	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGCCTGCCTCACTCCCAGG
796	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCCTATCGCCAGTCAAAATTC
799	ATGGGCGAGGCGTGCTCAAAG
800	CTAGGCGATTTCCGCACGCTGC
816	GAGGTCAGGGACTCGGGCATCGTCATTGGATGTGGTAACG

Cox group ID	Sequence 5'-3'
817	CGTTACCACATCCAATGACGATGCCCGAGTCCCTGACCTC
824	GGTTGGCTGGTAGACGTCATATAATCATACTTACAGTCCCTTTCTAACGC
825	TTTCTTTCAACACAAGATCCCAAAGTCAAAATGCCCGTTACTACCCCCAC
826	GCTTAGATTGTTTGTTAGAAGC

Table 7.7 Overview of used oligonucleotides used for PCR amplification of DNA fragments to build KO cassettes for *asL4*, *asL5* and *asL6*. For oligonucleotides binding sites a-f Figure 4.2.

Construct ID	GOI	Selection marker and oligonucleotides for amplification of KO fragments	Oligonucleotides for construction in *S. cerevisiae*
RSI77 1	*asL4*	gen^R P_{g+h}:P653+660 P_{i+j}: P659+654	P_{a+b}: P621+622 P_{c+d}: P623+624 P_{e+f}: P625+626
RSI77 2	*asL4*	hyg^R P_{g+h}: P653+22 P_{i+j}: P23+654	P_{a+b}: P621+622 P_{c+d}: P623+627 P_{e+f}: P928+626
RSI77 3	*asL5*	gen^R P_{g+h}: P655+660 P_{i+j}: P659+656	P_{a+b}: P629+630 P_{c+d}: P631+632 P_{e+f}: P633+634
RSI77 4	*asL5*	hyg^R P_{g+h}: P655+22 P_{i+j}: P23+656	P_{a+b}: P629+630 P_{c+d}: P631+635 P_{e+f}: P636+634
RSI77 5	*asL6*	gen^R P_{g+h}: P657+660 P_{i+j}: P659+658	P_{a+b}: P637+638 P_{c+d}: P639+640 P_{e+f}: P641+642
RSI77 6	*asL6*	hyg^R P_{g+h}: P657+22 P_{i+j}: P23+658	P_{a+b}: P637+638 P_{c+d}: P639+643 P_{e+f}: P644+642

Table 7.8 Overview of used oligonucleotides used for PCR amplification of DNA fragments to build silencing cassettes for *asL4*, *asL5* and *asL6*. For oligonucleotides binding sites a-f see Figure 4.15.

Construct ID	GOI	Resistance cassette	Oligonucleotides for construction in *S. cerevisiae*
RSI95 1 RSI95 2	*asL4*	gen^R hyg^R	P_{a+b}: P684+685 P_{c+d}: P686+687 P_{e+f}: P688+689
RSI95 2 RSI95 4	*asL5*	gen^R hyg^R	P_{a+b}: P684+985 P_{c+d}: P868+690 P_{e+f}: P691+692
RSI95 5 RSI95 6	*asL6*	gen^R hyg^R	P_{a+b}: P684+985 P_{c+d}: P686+693 P_{e+f}: P694+695

Table 7.9 Overview of used oligonucleotides used for PCR amplification of DNA fragments to build fungal expression and LR entry clones.

Construct ID	Vector backbone	Template	Oligonucleotides for construction in *S. cerevisiae*
RSI63 1-12	RSI42A1LR4C (pTYGSarg)	RSI67-4	*asL1*: P421+422
RSI63 2-5	RSI42A1LR4C (pTYGSarg)	RSI67-4, RSI67-2	*asL1*: P421+423 *asL3*: P425+425
RSI63 4-2	RSI42A1LR4C (pTYGSarg)	RSI67-1	*asR2*: P603+604
RSI96 1-6	RSI42A1LR4C (pTYGSarg)	RSI67-4, RSI67-2, RSI67-1	*asL1*: P421+423 *asL3*: P424+426 *asR2*: P427+428
RSI96 2-9	RSI42A1LR4C (pTYGSarg)	gDNA	*asL4*: P645+706 *asL5*: P707+708 *asL6*: P709+650
RSI63 5-1	pTYGSmet	RSI67-4	*asL1*: P421+422
RSI63 6-4	pTYGSmet	RSI67-4, RSI67-2	*asL1*: P421+423 *asL3*: P425+425
RSI96 4-2	pTYGSmet	RSI67-4, RSI67-2, RSI67-1	*asL1*: P421+423 *asL3*: P424+426 *asR2*: P427+428
RSI96 5-9	pTYGSmet	gDNA	*asL4*: P645+706 *asL5*: P707+708 *asL6*: P709+650
RSI100 1-9	pTYGSmet	gDNA	*asL2*: P755+756 *asR4*: P757+758 *asR6*: P759+760
RSI97 1-1	pTYGSmet	gDNA	*asL4*: P645+710 *asL6*: P709+650
RSI97 2-2	pTYGSmet	gDNA	*asL4*: P645+710 *asL5*: P648+711
RSI97 3-1	pTYGSmet	gDNA	*asL5*: P647+708 *asL6*: P709+650
RSI112 1-1 1	pTYGSmet	gDNA	*asL2*: P755+824
RSI112 1-2 1	pTYGSmet	gDNA	*asR4*: P791+789
RSI112 1-3 1	pTYGSmet	gDNA	*asR6*: P825+760
RSI76 1-5	pTYGSade	gDNA	*asL4*: P645+646
RSI76 2-3	pTYGSade	gDNA	*asL5*: P647+648
RSI76 3-3	pTYGSade	gDNA	*asL6*: P649+650
RSI100 2-16	pTYGSade	gDNA	*asL2*: P755+756 *asR4*: P757+758 *asR6*: P759+760
RSI102 1-10	pTYGSade	gDNA	*asL2*: P755+P787 *asR4*: P788+789
RSI102 2-5	pTYGSade	gDNA	*asL2*: P755+787 *asR6*: P790+760
RSI102 3-5	pTYGSade	gDNA	*asR4*: P791+792 *asR6*: P790+760
RSI112 2-1 1	pTYGSade	gDNA	*asL2*: P755+824
RSI112 2-2 1	pTYGSade	gDNA	*asR4*: P791+789
RSI112 2-3 1	pTYGSade	gDNA	*asR6*: P825+760
RSI42A1LR4C	pTYGSarg	-	LR with RSI42A1
RSI100 9-1	RSI96 5-9 (pTYGSmet)	-	LR with RSI100 4-1
RSI108 1-1	RSI96 4-2 (pTYGSmet)	-	LR with RSI100 3-1
RSI108 2-1	RSI96 5-9 (pTYGSmet)	-	LR with RSI100 3-1
RSI108 3-1	RSI63 4-2 (pTYGSmet)	-	LR with RSI100 3-1
RSI100 6-1	RSI100 2-16 (pTYGSade)	-	LR with RSI100 3-1
RSI42A1	pE-YA	gDNA	*aspks1*: P318+319, P320+321
RSI100 3-1	pE-YA	gDNA	*asR5*: P761+762
RSI100 4-1	pE-YA	RSI100 7-9	*asR7*: P771+P772
RSI100 8-1	pE-YA	gDNA	*asR4*: P795+816, P817+796
RSI67 1	pCR2.1	cDNA	*asR2*: P427+428
RSI67 2	pCR2.1	cDNA	*asL3*: P424+425
RSI67 4	pCR2.1	cDNA	*asL1*: P421+423
RSI100 7-9	pCR2.1	gDNA	*asR7*: P799+800

7.2.4 Extraction, Purification, Concentration and Sequencing of DNA

For downstream processes PCR amplified DNA and vector DNA were purified by kit according to the manufacturer's protocol. For linear DNA the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) was used. Vector DNA was extracted from a 3-5 mL *E. coli* overnight culture using the NucleoSpin Plasmid kit (Macherey-Nagel). The LB overnight culture containing the appropriate antibiotic was inoculated from either a fresh plate or a 25% glycerol stock and grown at 37 °C,

180 rpm for 12-16 hours. The quality of purified DNA samples was determined on agarose gels and the concentration measured on a spectrophotometer (DeNovix DS-11+ Spectrophotometer). Sequencing of DNA samples was carried out through Eurofins Genomics with custom oligonucleotides.

7.2.5 Restriction Hydrolysis and Cloning

Standard restriction hydrolysis and ligation with T4 DNA Ligase was performed according to the manufacturer's protocol.

7.2.5.1 TOPO-TA cloning

The Invitrogen TOPO-TA cloning kit was used according to the provided protocol to directly ligate *Taq* amplified linear DNA fragments. To ligate Q5 amplified DNA it was first kit purified and then incubated 10 min at 68 °C with *Taq* polymerase to introduce A-overhangs.

Table 7.10 Pipetting scheme to introduce A-overhangs on Q5 purified amplified PCR products.

Component	10 µL reaction	Final concentration
Taq DNA Polymerase	0.2 µL	
10x Standard *Taq* reaction buffer	1 µL	1x
2.5 mM dATP	0.5 µL	125 µM
DNA fragment	to 10 µL	

7.2.5.2 Gateway Cloning

The Invitrogen Gateway LR Clonase II enzyme mix was used for *in vitro* recombination from pE-YA entry clones into pTYGS family destinations vectors. The manufacturer's protocol was optimised and a reduced reaction scale was used.

Table 7.11 Pipetting scheme of the optimised recombination reaction.

Component	5 µL reaction	Concentration
Entry clone	0.5-3.5 µL	50-150 ng
Destination vector	0.5 µL	150 ng
TE buffer, pH 8.0	to 4 µL	1x
Gateway LR Clonase II	1 µL	0.2 µg/µL
	incubation at 25 °C for 1-12 h	
2 µg/µL Proteinase K	0.5 µL	

7.2.5.3 USER Cloning

DNA for USER cloning was amplified with Phusion U polymerase (Thermo Scientific), which is not inhibited by dU containing templates. The procedure was carried out according to the provided

protocol. PCR amplified DNA was gel purified to remove template DNA. Vector DNA was hydrolysed with *Pac*I and *Nt.Bbv*CI. Restriction hydrolases were heat inactivated. USER cloning kit was used according to the provided protocol and 2-12 µL reaction mixture were used to transform competent *E. coli*.

Table 7.12 Pipetting scheme for USER cloning.

Component	12 µL reaction	Final concentration
Purified PCR product	10 µL	
linearized vector	1 µL	20 ng
USER enzyme	01 µL	1 U
	to 12 µL	

7.2.6 Extraction of fungal gDNA and RNA and Reverse Transcription into cDNA

gDNA

For small-scale gDNA isolation the GenElute Plant Genomic DNA Miniprep kit (Sigma-Aldrich) was used. Fungal material was obtained from liquid culture or agar plates and was either directly used or stored at -20 °C prior to gDNA extraction. When needed as PCR template only, grinding in liquid nitrogen was omitted and for elution 1x 100 µL TE buffer (pH 8.0) was sufficient.

For large-scale gDNA isolation 300 mg of freeze dried fungal material (obtained from a 3 d old PDB culture of *A. strictum*) was grinded with mortar and pistil in liquid nitrogen. 20 mL gDNA extraction buffer was added to the disrupted cells in a 50 mL tube and mixed gently on a vertical carousel (25 rpm, 15 min). 20 mL phenol:chloroform:isoamyl alcohol (PCI) was added and mixed gently on a vertical carousel (25 rpm, 15 min). Aqueous phase was recovered by centrifugation (10 min, maximal speed) and extraction with 20 mL PCI was repeated (6-7 times) until interphase was clear. 10 mL chloroform:isoamyl alcohol (24:1) was added to the aqueous phase and mixed on a vertical carousel (15 min, 25 rpm). Centrifugation (10 min, maximal speed) recovered the aqueous phase, which was subsequently added to 0.7-1 volumes of isopropanol (100%). The mixture was inverted carefully until gDNA precipitation was observed. gDNA was pelleted by centrifugation (5 min, 4000 rpm), washed with 75% ethanol and left to dry until the pellet was clear. 500 µL TE pH 8.0 was used to suspend the pellet over night at 4 °C. For caesium chloride gradient centrifugation 3.2 g caesium chloride was balanced in a 15 mL tube and tared. 160 µL ethidiumbromide (10 mg/mL), 500 µL gDNA in TE and TE buffer was added to 3.2 g. The solution was carefully transferred to a Beckman Coulter Polyallomer tube (326819) with a Pasteur pipette. Centrifugation was carried out at 45000 rpm for 24 h using a Beckman SW 55 Ti rotor and UZ Maxima XP-80 Beckman Coulter ultracentrifuge. gDNA band was detected under UV light. Fractions (400 µL) were collected by hand.

A 5 µL sample of each fraction was analysed on a 0.8% agarose gel (130 V, 30 min). gDNA containing fractions were combined, diluted in 2.5 volumes of water and gDNA precipitated with 0.7-1 volume of isopropanol. After incubation on ice for 5 min, the gDNA was pelleted by centrifugation (10 min, maximal speed). Pelleted DNA was dried and suspended in 100 µL ddH_2O. Sodium acetate (3 M) was added to a final concentration of 0.3 M and 2.5 volumes of 96% ethanol were used to precipitate gDNA. After incubation on ice for 10 min, gDNA was pelleted by centrifugation (10 min, maximum speed). gDNA was washed with 75% ethanol and centrifuged (10 min, maximum speed) again. Pellet was left to dry before resuspension in 50 µL ddH_2O overnight. For whole genome sequencing a concentration of 100 ng/µL in a total volume of 100 µL ddH_2O or 100 mM Tris-HCl pH 8.0 is required. Concentration was determined on a spectrophotometer prior to dilution to the required concentration.

RNA and cDNA

The ZR Fungal Bacterial RNA MiniPrep kit (Zymo Research) was used to isolate total RNA according to the provided protocol. RNase free tubes and plugged pipet tips were used to avoid degradation of RNA. cDNA was prepared from total RNA using oligo(dT) nucleotides and the RevertAid Premium Transcriptase kit (Thermo Scientific) according to protocol.

Yeast Vector DNA extraction

Yeast vector DNA was extracted with the Yeast Plasmid Miniprep II kit (Zymoprep). Cells of the quarter of a plate (Chapter 7.4.2) were collected with a toothpick and directly suspended in kit solution I. DNA was purified according to protocol and eluted in 10 µL sterile ddH_2O. 2-5 µL DNA were directly used for transformation of competent *E. coli* cells.

7.3 Biochemistry

Clones for *E. coli* codon optimised *asR5* and *asR6* were purchased with an N-terminal 6x*his* tag in pET100 at Invitrogen (Darmstadt, Germany). AsR5 was first solubly produced by Raissa Schor, but later purification optimised by Carsten Schotte. AsR6 was solublly produced and optimised by Carsten Schotte. Both vectors were used to transform *E. coli* BL21 (DE3), which was used as heterologous gene expression host for protein production.

7.3.1 Protein Expression and Cell Lysis

An *E. coli* BL21(DE3) overnight LB culture was used to inoculate 100-1000 mL 2TY medium. The culture was incubated at 37 °C and 200 rpm until an OD_{600} of 0.4-0.8 was reached. The gene expression and protein production was induced with 0.1 mM isopropyl β-D-1-

thiogalactopyranoside (IPTG) for 16 h at 16, 20 or 25 °C. The cells were used directly or stored at -20 °C for later protein purification. All following steps were carried out on ice or at 4 °C.

The pellet was suspended in 10-50 mL pre cooled protein resuspension buffer and cells lysed by sonication (pulse program 7x 30 s, 7x 30 s cool down). Cell debris were removed by centrifugation (8500 rpm, 45 min, 4 °C) and supernatant used for affinity purification.

7.3.2 Affinity Purification by Gravity Column

Up to 1 mL Ni-NTA Agarose (Macherey-Nagel, Düren) was equilibrated with protein resuspension buffer (2 x 1 mL, 2000 g, 2 min, 4 °C) prior to incubation with supernatant obtained in Chapter 7.3.1. After up to 1.5 h at 4 °C on an agitator the Ni-NTA agarose was collected by gravity, washed with protein resuspension buffer (10 mL) and protein eluted with an imidazole gradient in protein elution buffer (50-500 mM imidazole). Three fractions each 500 µL were taken per imidazole concentration. Carsten Schotte optimised protein production and purification for fast protein liquid chromatography (FPLC).

7.3.3 Polyacrylamide Gel Electrophoresis

Protein samples were analysed by vertical SDS polyacrylamide gel electrophoresis (SDS-PAGE). Gels were casted with the Mini-PROTEAN Tetra Handcast system and run in Mini-PROTEAN Tetra vertical electrophoresis cell (Bio-Rad). First the separating gel was casted and overlaid with isopropanol until solidification. In a next step isopropanol was removed, stacking gel was casted and polymerised.

15 µL protein sample and 5 µL 4x SDS loading buffer were denatured at 95 °C for 5 min. Up to 20 µL were loaded on a 12% polyacrylamide gel alongside with 5 µL prestained standard (Broad Range, NEB) and separated for 50-90 min at 20-40 mA. Gels were stained in coomassie dying solution (1 h, agitator) and destained with coomassie bleach (1 h, agitator). Gels were scanned with gel documentation system (Gel Doc XR+, Bio-Rad) and the software Image Lab (Bio-Rad).

Table 7.13 Components of a 12% polyacrylamide gel.

Component	12% separating gel	5% stacking gel
30% acrylamide/bisacryamide (37,5:1)	3.00 mL	535 µL
1.5 M Tris-HCl, pH 8.8	1.90 mL	-
0.5 M Tris-HCl, pH 6.8	-	250 µL
10% (w/v) SDS	75.0 µL	20.0 µL
10% (w/v) APS	75.0 µL	20.0 µL
TEMED	3.00 µL	2.00 µL
ddH_2O	2.45 mL	1.70 mL

7.4 Microbiology

Table 7.14 Strains and origin.

Strain	Genotype	Phylum	Origin
Acremonium strictum	wildtype	*Ascomycete*	Cox group, Bristol
Aspergillus oryzae NSAR1	Δ*argB sC*$^{-}$ *adeA niaD*$^{-}$	*Ascomycete*	Lazarus group Bristol
Saccharomyces cerevisiae CEN.PK	MATa/α ura3-52/ura3-52 trp1-289/trp1-289 leu2-3_112/leu2-3_112 his3 Δ1/his3 Δ1 MAL2-8C/MAL2-8C SUC2/SUC2	*Ascomycetes*	Lazarus group Bristol
Escherichia coli OneShot TOP10	F-*mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (Str^{R}) *end*A1 *nup*G	*Proteobacteria*	Thermo Fisher Scientific
Escherichia coli OneShot *ccd*B survival 2T1^{R}	F-*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*Δ139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^{R}) *end*A1 *nup*G *fhu*A::*IS2*	*Proteobacteria*	Thermo Fisher Scientific
Escherichia coli BL21 (DE3)	F^{-} ompT *hsd*S$_{B}$ (r$_{B}^{-}$, m$_{B}^{-}$) *gal dcm* (DE3)	*Proteobacteria*	Thermo Fisher Scientific

7.4.1 *E. coli*

Maintenance

For short term cultivation of *E. coli*, it was streaked for single colonies on LB agar containing the appropriate antibiotic with a Drigalski spatula and incubated at 37 °C for 12 h. Plates were stored for up to 1 month at 4 °C. 3-5 mL LB overnight cultures were inoculated from plates or from 25% glycerol stocks. Stocks were prepared from 500 µL overnight culture and 500 µL 50% glycerol and stored at -80 °C.

Transformation

To 50 µL of competent *E. coli* 1-3 µL vector DNA was added and incubated on ice for up to 30 min. After a 45 sec heat shock at 42 °C, the cells were placed on ice for 2 min and 300 µL of SOC medium was added. Cells were incubated at 37 °C, 350 rpm for 1 h and then collected by centrifugation (4000 g, 3 min). 300 µL SOC was discarded and the cells suspended in the remaining supernatant prior to plating out on LB agar containing the appropriate antibiotic for 12 h at 37 °C.

7.4.2 *S. cerevisiae*

Maintenance

For short-term cultivation of *S. cerevisiae,* it was streaked on YPAD and incubated at 28 °C for up to one week or at 4 °C for up to 1 month. For stock preparation 1 mL of a ddH_2O was used to suspend

a single colony from a YPAD plate. Cells were collected with a pipette and 50% glycerol was added to a final concentration of 25% glycerol. Stocks were stored at -80 °C.

Transformation

A young (3-5d) single colony of *S. cerevisiae* was used to inoculate a 10 mL YPAD pre-culture and incubated at 28 °C, 200 rpm. After 18 h 40 mL of YPAD media was added and incubated for 4 h at 28 °C, 200 rpm. Cells were collected by centrifugation (3000 g, 5 min), washed with 5 mL ddH_2O and centrifuged (3000 g, 5 min). Supernatant was discarded; cells suspended in 1 mL ddH_2O and transferred to a 1.5 mL microfuge tube. The mixture was centrifuged (14500 rpm, 15 s) and cells were suspended in 400 µL 0.1 M lithium acetate. 50 µL aliquots were prepared, centrifuged (14500 rpm, 15 s) and the supernatant discarded. 240 µL of a 50% PEG3350 solution, 36 µL 1 M lithium acetate, 50 µL carrier DNA (denatured salmon sperm DNA, 2 mg/mL in TE buffer) and up to 34 µL DNA were added to the cell pellet regarding the order. Concentration of each linear DNA fragment was approximately 0.5-1 µg. Fragments to be joined contained at least 30 bp overlap sequence and were added in a 1:1:1 ratio. Cells were suspended in the transformation mixture and incubated at 42°C for 50 min. Cells were collected by centrifugation (6000 rpm, 15 s), the supernatant was discarded and pellet was suspended in 1 mL ddH_2O. 200 µL were spread on SM-URA plates and incubated at 28 °C for 3-4 d. As negative control ddH_2O, as positive control circular plasmid and as background control linearized plasmid was used.

7.4.3 *A. oryzae*

Maintenance

For short-term cultivation of *A. oryzae* NSAR1 it was streaked on DPY and incubated at 28 °C for up to one month or at 4 °C for up to 3 months. For stock preparation 10 mL of a ddH_2O was added to a sporulating DPY plate. Mycelia and spores were scraped of using an inoculation loop and mixed vigorously. 50% glycerol was added to a final concentration of 25% glycerol to prepare a mycelia/spore mixed stock. When spore stocks were prepared mixture was gravity filtered over sterile miracloth and glycerol to a final concentration of 25% added. Stocks were stored at -80 °C.

Transformation

1 mL spore suspension from a fresh *A. oryzae* NSAR1 DPY plate was used to inoculate 100 mL (500 mL flask) of GN liquid culture and incubated for 12 h at 28 °C, 180 rpm. Cells were collected by filtration over sterile miracloth, washed with 0.8 M NaCl (50-100 mL) and suspended in 10 mL filter-sterilised *A. oryzae* NSAR1 protoplasting solution by inversion. Suspension was incubated at 30 °C and 150 rpm for 2 h. Protoplast formation was checked every hour for completion using a microscope. When majority of mycelia had been digested into spherical protoplasts, these were

released by repeated pipetting with a cut tip and gravity filtration through sterile miracloth removed remaining mycelia. Protoplasts were collected by centrifugation (2000 g, 5 min) and directly suspended in the required amount of fungal transformation solution I.

DNA (≥ 1µg, in 10 µL ddH_2O) was mixed with 100 µL protoplasts and incubated on ice for 5 min. 1 mL of fungal transformation solution II was added drop wise and the mixture incubated at 25 °C for 20 min. 14 mL molten CZD+S 0.8% agar (~37 °C) was added and the mixture poured over two plates containing CZD+S 1.5% agar. Plates were incubated at 28 °C until colonies could be observed (usually 6 d), these were transferred to secondary plates of CZD 1.5% agar. Vigorously growing colonies were transferred onto third plate (CZD 1.5% agar). Real transformants were used to inoculate liquid cultures and to prepare 25% glycerol stocks. The supplements in all CDZ media were adapted for each transformation in regard of the added DNA.

7.4.4 *A. strictum*

Maintenance

For short-term cultivation of *A. strictum* single colonies were streaked on ASSA and incubated at 25 °C for up to one month or at 4 °C for up to 3 months. For stock preparation 10 mL of a 0.01% Tween 80 solution was added to a sporulating ASSA plate. Mycelia and spores were scraped of using an inoculation loop and mixed vigorously. 50% glycerol was added to a final concentration of 25% glycerol to prepare a mycelia/spore mixed stock. When spore stocks were prepared the mixture was gravity filtered over sterile miracloth and glycerol to a final concentration of 25% added. Stocks were stored at -80 °C.

Transformation

100 µL spore suspension was used to inoculate 100 mL (500 mL flask) of ASSM liquid culture and incubated for 1.5 d at 25 °C, 200 rpm. Cells were collected by filtration over sterile miracloth, washed with 0.8 M NaCl (50-100 mL) and suspended in 10 mL filter-sterilised *A. strictum* protoplasting solution by inversion. Suspension was incubated at 30 °C and 150 rpm up to 1.5 h. Protoplast formation was checked every hour for completion using a microscope. When majority of mycelia had been digested into spherical protoplasts, these were released by repeated pipetting with a cut tip and gravity filtration through sterile miracloth removed remaining mycelia. Protoplasts were collected by centrifugation (2000 g, 5 min) and washed twice with 20 mL 0.8 M NaCl (2000 g, 5 min). 10 mL fungal transformation solution I was used to suspend protoplasts and the concentration was determined by microscopic analysis using a haemocytometer. Cells were collected in a final centrifugation (2000 g, 5 min) and suspended in solution I to a concentration of 1-9 x 10^7 protoplasts per L. Protoplasts were stored on ice and used as soon as possible.

DNA (≥ 1µg, in 10 µL ddH_2O) was mixed with 100 µL protoplasts and incubated on ice for 5 min. 1 mL of fungal transformation solution II was added drop wise and the mixture incubated at 25 °C for 20 min. 30 mL molten CD+S 0.8% agar (~37 °C) with 100 µg/mL hygromycin B was added and the mixture poured over two plates. Plates were incubated at 25 °C until resistant colonies could be observed (usually 10 d), these were transferred to secondary plates of CD+S 1.5% agar containing 100 µg/mL hygromycin B. Vigorously growing colonies were transferred onto third plate (CD+S 1.5% agar, 100 µg/mL Hygromycin B) after 7 d. Real transformants were used to inoculate liquid cultures for xenovulene production and prepare 25% glycerol stocks. All used CD+S media were prepared with deionised water, as no colonies could be recovered with milipore water.

Nuclei staining

A. strictum was grown in 100 mL ASSM liquid media for three days at 25 °C and 200 rpm. A 1 mL sample was taken and washed twice with water (2 x 1 mL). Pelleted cells (5 min, maximum speed) were suspended in 1 mL ethanol (70%) and incubated at 25 °C for 10 min. Cells were washed with water (1x 1 mL) and DAPI (stock solution 1 mg/mL) added to a final concentration of 0.1 µg/mL in water or methanol. After incubation for 10 min at 25 °C cells were washed once with water (1x 1mL) and suspended in 25% glycerol. Cells were immobilised in 1% agar in between two microscope cover slides.

7.5 Chemical Analysis of Fungal cultures

Fungal cultures for extraction were inoculated either from plate or directly from glycerol stock. When a fresh fungal plate was used, 3 mL water was added, fungal material scraped off with the tip and transferred into the flasks. 100 mL *A. oryzae* NSAR1 cultures were grown in 500 mL baffled flasks. *A. strictum* was cultivated in 50 mL and 100 mL using 250 mL/500 mL flasks respectively.

Table 7.15 Standard conditions for liquid fermentation of the cultured ascomycetes.

Strain	Medium	Cultivation conditions
A. strictum	ASPM, PDB, ME, YMG, CM	25 °C, 200 rpm, 3-10 d
A. oryzae NSAR1	CMP, Starch M + 0.15% methionine, DPY	28 °C, 110 rpm, 4-6 d, baffled flasks

7.5.1 *A. strictum*

100 mL *A. strictum* cultures were centrifuged (8000 rpm, 10 min), supernatant acidified to pH 2 with HCl (2M) and extracted twice with ethyl acetate/hexane (1:1, 2 x 100 mL). Combined organic layers were dried over $MgSO_4$ and solvents removed *in vacuo*. Extract was dissolved in methanol or

acetonitrile:water (9:1) to a concentration of 10 mg/mL, filtered over glass wool and analysed by LCMS.

For screening of multiple colonies, a small scale extraction method was used. 1 mL of *A. strictum* liquid culture was centrifuged (5 min, maximum speed), the supernatant was transferred to a 2 mL centrifuge tube and extracted twice with 1 mL ethyl acetate/hexane (1:1). Solvents were removed in a vacuum centrifuge, samples were dissolved in 200 µL methanol and directly analysed by LCMS.

7.5.2 *A. oryzae*

100 mL *A. oryzae* NSAR1 cultures were acidified to pH 2 with HCl (2M) and incubated for 30 min at 28 °C, 110 rpm. Cells were disrupted with a hand blender and removed by Büchner filtration. Supernatant was extracted twice with ethyl acetate or ethyl acetate/hexane (1:1, 2 x 100 mL). Combined organic layers were dried over $MgSO_4$ and solvents removed *in vacuo*. Extract was dissolved in methanol to a concentration of 10 mg/mL, filtered over glass wool and analysed by LCMS.

When cells and supernatant were extracted separately, cells were removed by filtration and stirred in acetone (100 mL for 100 mL culture volume) for 3 h. In a next step cells were removed by filtration and the water/acetone mixture concentrated *in vacuo*. Resulting water layer was increased to 100 mL with deionised water, acidified to pH 2 with HCl (2M) and extracted with ethyl acetate twice (2 x 100 mL). Supernatant was acidified to pH 2 with HCl (2M) and extracted twice with ethyl acetate (2 x 100 mL). Combined organic layers of both samples were dried over $MgSO_4$ and solvents removed *in vacuo*. Extracts were dissolved in methanol or acetonitrile:water (9:1) to a concentration of 10 mg/mL, filtered over glass wool and analysed by LCMS.

7.5.3 LCMS

A Waters 2545 binary gradient module and a Waters 515 HPLC pump coupled to a Waters 2998 DAD, a Waters 2420 evaporative light scattering detector (ELSD) and a Waters single quadrupole mass detector 2 (SQ detector 2) were used for analytical and preparative LCMS analysis. Column selection was managed by the Waters system fluidics organizer (SFO). The DAD measured wavelengths from 210 to 600 nm and the mass detector was adjusted depending on the sample to measure in a range of 100-1000 *m/z*. Analytical and preparative LCMS methods were operated in water and acetonitrile with the addition of 0.05% (v/v) formic acid. LCMS grade acetonitrile was purchased and water was prepared by double distillation of millipore water.

7.5.3.1 Analytical

Samples were dissolved in methanol or acetonitrile:water 9:1 to a concentration of 10 mg/mL, filtered over glass wool and 10 – 20 µL injected onto a Phenomenex Kinetex 2.6 µm C_{18} 100 Å column (size 100 x 4.6 mm). The gradient was selected depending on the sample and run with 1 mL/min.

Table 7.16 Different programs used for analytical LCMS.

Program	Time/min	Water-acetonitrile/%
A1 standard	0-1	90-10
	1-10	10-90
	10-12	10-90
	12-13	90-10
	13-15	90-10
A2 polar	0-1	90-10
	1-10	60-40
	10-12	10-90
	12-13	90-10
	13-15	90-10

7.5.3.2 Preparative

Samples were dissolved in methanol or acetonitrile:water 9:1 to a concentration of 50 mg/mL, filtered over glass wool and 50 – 200 µL injected onto a Phenomenex Kinetex 5 µm C_{18} 100 Å, AXIA column (size 250 x 21.2 mm). The gradient was selected depending on the sample and run with 20 mL/min. Fractions were collected with the Waters Sample Manager 2767 by either mass directed or time dependent trigger. Combined fractions were dried using the vacuum centrifuge Christ RVC 2-25 CD plus.

7.5.4 GCMS

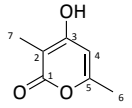
GCMS analysis was carried out on a HP 6890 gas chromatograph connected to an Agilent 5973 mass detector on an OPTIMA 5 MS capillary column (30 m, 0.32 mm i.d., 0,25 µm film). Chromatography was achieved with 1.5 mL/min helium flow, electron energy of 70 eV and 280 °C transfer line. Injection volume of 5 µL was injected with a split less method. GC program: 0-1 min / 50 °C, gradient 20 °C/min to 300°C.

7.5.5 HRMS

High resolution mass spectrums were acquired on a Waters Acquity ultra-performance liquid chromatography (UPLC) coupled to a quadrupole time-of-flight mass spectrometer (Q-TOF).

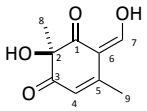
7.5.6 NMR spectroscopy

All NMR spectra were acquired on a Bruker Ascend 400 MHz or Bruker DRX 500 MHz Spectrometer respectively. The data is shown in ppm, referenced to the used deuterated solvent and coupling constant *J* are quoted in Hz. Coupling constants and multiplets are described as observed. Deviations are caused by digitisation. The software MestReNova 10.0 was used for analysis of the data.



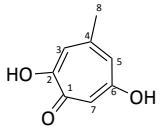
191
$C_7H_8O_3$
M = 140 g/mol

Compound **191**:[236] ^{1}H NMR (500 MHz, CD_3OD), δ_H: 5.98 (s, 1H, 4-C*H*), 2.19 (s, 3H, 6-CH_3), 1.84 (s, 3H, 7-CH_3). ^{13}C NMR (125 MHz, CD_3OD), δ_C: 169.2 (*C*-1), 168.9 (*C*-3), 161.3 (*C*-5), 102.0 (*C*-4), 98.5 (*C*-2), 19.5 (*C*-6), 8.2 (*C*-7).



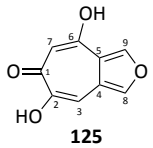
121
$C_9H_{10}O_4$
M = 182 g/mol

Compound **121**:[97] ^{1}H NMR (400MHz, *d*$_6$-DMSO), δ_H: 8.13 (s, 1H, 7-C*H*OH), 5.64 (s, 1H, 4-C*H*), 2.22 (s, 3H, 8-CH_3), 1.30 (s, 3H, 9-CH_3).



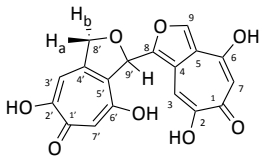
194
$C_8H_8O_3$
M = 152 g/mol

Compound **194**:[98] ^{1}H NMR (500 MHz, CD_3OD), δ_H: 6.92 (s, 1H, 3-C*H*), 6.78 (s, 1H,7-C*H*), 6.69 (s, 1H, 5-C*H*), 2.38 (s, 3H, 8-CH_3). ^{13}C NMR (125 MHz, CD_3OD), δ_C: 174.2 (1-*C*O), 169.5 (6-*C*OH), 166.4 (2-*C*OH), 150.4 (4-*C*CH3), 122.0 (5-*C*H), 118.1 (3-*C*H), 113.7 (7-*C*H), 27.8 (8-*C*H$_3$).



125
$C_9H_6O_4$
M = 178 g/mol

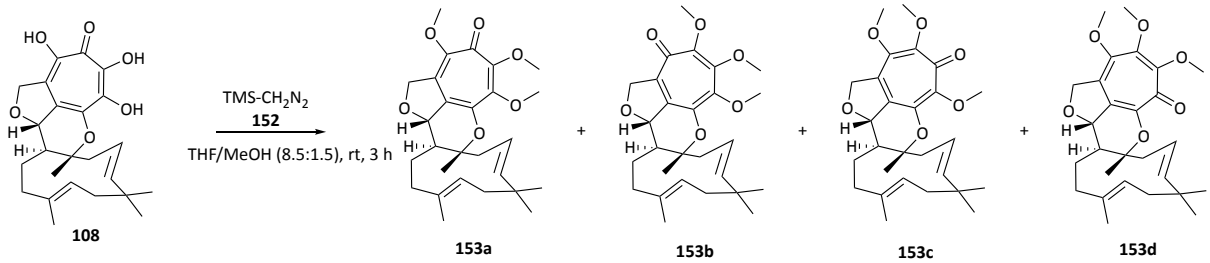
Compound **125**:[98] ^{1}H NMR (500 MHz, CD_3OD), δ_H: 8.20 (dd, 1H, $^4J_{HH}$ = 1.9, $^5J_{HH}$ = 0.8, 9-C*H*OCH), 8.01 (d, 1H, $^4J_{HH}$ = 1.9, 8-C*H*OCH), 6.89 (s, 1H, $^5J_{HH}$ = 0.8, 3-C*H*). ^{13}C NMR (125 MHz, CD_3OD), δ_C: 182.0 (1-*C*O), 168.4 (6-*C*OH), 150.8 (2-*C*OH), 145.8 (9-*C*HOCH), 143.6 (8-*C*HOCH), 121.7 (4-*C*), 121.6 (4-*C*), 107.5-107.3 (7-*C*D), 103.6 (3-*C*H).



196
$C_{18}H_{12}O_8$
M = 356 g/mol

Compound **196**:[98] ^{1}H NMR (500 MHz, CD_3OD), δ_H: 8.11 (s, 1H, 9-C*H*), 7.00 (s, 1H, 3-C*H*), 6.86 (s, 1H, 3'-C*H*), 6.60 (d, 1H, J_{HH} = 2.4, 9'-C*H*OC), 5.32 (dd, 1H, J_{HH} = 14.0, 3.6, 8'-*C*H$_2$OC, H$_a$), 5.15 (d, 1H, J_{HH} = 13.8, 8'-*C*H$_2$OC, H$_b$). ^{13}C NMR (125 MHz, CD_3OD), δ_C: 182.2 (1-*C*O), 173.4 (1'-*C*O), 167.6 (2'-*C*OH), 167.0 (6'-*C*OH), 165.6 (6-*C*), 153.0 (8-*C*), 150.7 (2-*C*OH), 148.9 (4'-*C*), 144.2 (9-*C*H), 128.6 (5'-*C*), 122.0 (5-*C*), 118.1 (4-*C*), 107.0 (3'-*C*H), 103.5 (3-*C*H), 80.6 (9'-*C*HOC), 78.80 (8'-*C*H$_2$OC). 7-CH and 7'-CH are most likely exchanged to 7-CD and 7'-CD.

7.5.7 Methylation of Compound 108



108

153a
$C_{27}H_{36}O_6$
M = 456 g/mol

153b
$C_{27}H_{36}O_6$
M = 456 g/mol

153c
$C_{27}H_{36}O_6$
M = 456 g/mol

153d
$C_{27}H_{36}O_6$
M = 456 g/mol

Scheme 7.1 Methylation of **108** with TMS-diazomethane **152**.

To a solution of 0.075 mmol **108** in THF/Methanol (8.5/1.5 v/v) 0.450 mmol TMS-diazomethane (2 M in hexane) was added. The reaction mixture was stirred at 25 °C for 3 h. Solvents were removed and mixture dissolved in methanol for LCMS analysis.[120,121]

8 Bibliography

(1) Keller, N. P.; Turner, G.; Bennett, J. W. *Nat. Rev. Microbiol.* **2005**, *3*, 937.
(2) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2016**, *79*, 629.
(3) McGuire, J. M.; Wolfe, R. N.; Ziegler, D. W. *Antibiot. Annu.***1955**, 612.
(4) Endo, A. *J. Antibiot. (Tokyo).* **1979**, *32*, 852.
(5) Endo, A. *J. Antibiot. (Tokyo).* **1980**, *33*, 334.
(6) Kirby, G. W. *Science* **1967**, *155*, 170.
(7) Ziegler, J.; Facchini, P. J. *Annu. Rev. Plant Biol.* **2008**, *59*, 735.
(8) Croteau, R. B.; Davis, E. M.; Ringer, K. L.; Wildung, M. R. *Naturwissenschaften* **2005**, *92*, 562.
(9) Stack, D.; Neville, C.; Doyle, S. *Microbiology* **2007**, *153*, 1297.
(10) Bills, G.; Li, Y.; Chen, L.; Yue, Q.; Niu, X.-M.; An, Z. *Nat. Prod. Rep.* **2014**, *31*, 1348.
(11) Staunton, J.; Weissman, K. J. *Nat. Prod. Rep.* **2001**, *18*, 380.
(12) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5*, 2010.
(13) Christianson, D. W. *Chem. Rev.* **2006**, *106*, 3412.
(14) Quin, M. B.; Flynn, C. M.; Schmidt-Dannert, C. *Nat. Prod. Rep.* **2014**, *31*, 1449.
(15) Xu, W.; Gavia, D. J.; Tang, Y. *Nat. Prod. Rep.* **2014**, *31*, 1474.
(16) Grigoriev, I. V.; Nikitin, R.; Haridas, S.; Kuo, A.; Ohm, R.; Otillar, R.; Riley, R.; Salamov, A.; Zhao, X.; Korzeniewski, F.; Smirnova, T.; Nordberg, H.; Dubchak, I.; Shabalov, I. *Nucleic Acids Res.* **2014**, *42*, 699.
(17) Shendure, J.; Aiden, E. L. *Nat. Biotechnol.* **2012**, *30*, 1084.
(18) Keller, N. P.; Hohn, T. M. *Fungal Genet. Biol.* **1997**, *21*, 17.
(19) Zhang, Y.-Q.; Wilkinson, H.; Keller, N. P.; Tsitsigiannis, D. In *Handbook of Industrial Microbiology*; Marcel Dekker: New York, 2005; pp 355–386.
(20) Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. a; Weber, T.; Takano, E.; Breitling, R. *Nucleic Acids Res.* **2011**, *39*, W339.
(21) Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. a.; Breitling, R.; Takano, E.; Weber, T. *Nucleic Acids Res.* **2013**, *41*, 204.
(22) Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Müller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H. *Nucleic Acids Res.* **2015**, *43*, W237.
(23) Birch, A.; Massy-Westrop, P.; Moye, C. *Aust. J. Chem.* **1955**, *8*, 539.
(24) Collie, J. *J. Chem. Soc. Trans.* **1907**, *91*, 1806.
(25) Beld, J.; Lee, D. J.; Burkart, M. D.; Lee, D. J.; Lee, D. J. *Mol. Biosyst.* **2014**, *11*, 38.
(26) Maier, T.; Maier, T.; Leibundgut, M.; Leibundgut, M.; Ban, N.; Ban, N. *Science* **2008**, *321*, 1315.
(27) Maier, T.; Jenni, S.; Ban, N. *Science* **2006**, *311*, 1258.
(28) White, S. W.; Zheng, J.; Zhang, Y.-M.; Rock, C. O. *Annu. Rev. Biochem.* **2005**, *74*, 791.
(29) Rose, A. S.; Hildebrand, P. W. *Nucleic Acids Res.* **2015**, *43*, W576.
(30) Rose, A.; Bradley, A.; Valasatava, Y.; Duarte, J.; Prlic, A.; Rose, P. In *ACM Proceedings of the 21st International Conference on Web3D Technology*; 2016; pp 185–186.
(31) Smith, S.; Tsai, S.-C. *Nat. Prod. Rep.* **2007**, *24*, 1041.
(32) Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A. *Nat. Prod. Rep.* **2007**, *24*, 162.
(33) Shimizu, Y.; Ogata, H.; Goto, S. *ChemBioChem* **2017**, *18*, 50.
(34) Chooi, Y.-H.; Tang, Y. *J. Org. Chem.* **2012**, *77*, 9933.
(35) Weissman, K. J. *Nat. Prod. Rep.* **2014**, *32*, 436.
(36) 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**, *20*, 2260.
(37) Roberts, D. M.; Bartel, C.; Scott, A.; Ivison, D.; Simpson, T. J.; Cox, R. J. *Chem. Sci.* **2017**, *8*, 1116.
(38) Crawford, J. M.; Dancy, B. C. R.; Hill, E. a; Udwary, D. W.; Townsend, C. a. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 16728.
(39) Chiang, Y. M.; Szewczyk, E.; Davidson, A. D.; Keller, N.; Oakley, B. R.; Wang, C. C. C. *J. Am. Chem. Soc.* **2009**, *131*, 2965.
(40) Crawford, J. M.; Thomas, P. M.; Scheerer, J. R.; Vagstad, A. L.; Kelleher, N. L.; Townsend, C. A. *Science* **2008**, *320*, 243.
(41) Li, Y.; Xu, W.; Tang, Y. *J. Biol. Chem.* **2010**, *285*, 22764.
(42) Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y. *Chem. Biol.* **2001**, *8*, 189.
(43) Awakawa, T.; Yokota, K.; Funa, N.; Doi, F.; Mori, N.; Watanabe, H.; Horinouchi, S. *Chem. Biol.* **2009**, *16*, 613.
(44) Schweizer, E.; Beck, J.; Pika, S.; Siegner, A.; Schiltz, E. *Eur. J. Biochem.* **1990**, *498*, 487.
(45) Sun, H.; Ho, C. L.; Ding, F.; Soehano, I.; Liu, X. W.; Liang, Z. X. *J. Am. Chem. Soc.* **2012**, *134*, 11924.
(46) Kroken, S.; Glass, N. L.; Taylor, J. W.; Yoder, O. C.; Turgeon, B. G. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 15670.
(47) Schmitt, I.; Martín, M. P.; Kautz, S.; Lumbsch, H. T. *Phytochemistry* **2005**, *66*, 1241.
(48) Cane, D. E.; Rawlings, B. J.; Yang, C. C. *J. Antibiot. (Tokyo).* **1987**, *40*, 1331.
(49) Tudzynski, B. *Appl. Microbiol. Biotechnol.* **2005**, *66*, 597.
(50) Miziorko, H. M. *Arch. Biochem. Biophys.* **2011**, *505*, 131.
(51) Caruthers, J. M.; Kang, I.; Rynkiewicz, M. J.; Cane, D. E.; Christianson, D. W. *J. Biol. Chem.* **2000**, *275*, 25533.
(52) Shishova, E. Y.; Costanzo, L. Di; Cane, D. E.; Christianson, D. W.; Laboratories, D. V.; Uni, V.; Pennsyl, V. *Biochemsitry* **2007**, *47*, 1941.
(53) Pinedo, C.; Wang, C.-M.; Pradier, J.-M.; Dalmais, B.; Choquer, M.; Pêcheur, P. Le; Morgant, G.; Collado, I. G.; Cane, D. E.; Viaud, M. *ACS Chem. Biol.* **2008**, *3*, 791.
(54) Wang, C.; Hopson, R.; Lin, X.; Cane, D. E. *J. Am. Chem. Soc.* **2009**, *131*, 8360.
(55) Agger, S.; Lopez-Gallego, F.; Schmidt-Dannert, C. *Mol. Microbiol.* **2009**, *72*, 1181.
(56) Quin, M. B.; Flynn, C. M.; Wawrzyn, G. T.; Choudhary, S.; Schmidt-Dannert, C. *Chembiochem* **2013**, *14*, 2480.
(57) Rynkiewicz, M. J.; Cane, D. E.; Christianson, D. W. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 13543.
(58) Cane, D. E.; Swanson, S.; Murthy, P. P.; N. *J. Am. Chem. Soc.* **1981**, *103*, 2136.
(59) Cane, D. E.; Ha, H. J.; Pargellis, C.; Waldmeier, F.; Swanson, S.; Murthy, P. P. N. *Bioorg. Chem.* **1985**, *13*, 246.
(60) Cane, D. E.; Xue, Q.; Fitzsimons, B. C. *Biochemistry* **1996**, 12369.

(61) Regueira, T. B.; Kildegaard, K. R.; Hansen, B. G.; Mortensen, U. H.; Hertweck, C.; Nielsen, J. *Appl. Environ. Microbiol.* **2011**, *77*, 3035.

(62) 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.

(63) Jörnvall, H.; Persson, B.; Jörnvall, H.; Persson, B.; Krook, M.; Atrian, S.; Gonzàlez-Duarte, R.; Jeffery, J.; Ghosh, D. *Biochemistry* **1995**, *34*, 6003.

(64) Ohashi, M.; Liu, F.; Hai, Y.; Chen, M.; Tang, M.; Yang, Z.; Sato, M.; Watanabe, K.; Houk, K. N.; Tang, Y. *Nature* **2017**, *549*, 502.

(65) Matsuda, Y.; Awakawa, T.; Wakimoto, T.; Abe, I. *J. Am. Chem. Soc.* **2013**, *135*, 10962.

(66) Bugg, T. D. H. *Tetrahedron* **2003**, *59*, 7075.

(67) van Berkel, W. J. H.; Kamerbeek, N. M.; Fraaije, M. W. *J. Biotechnol.* **2006**, *124*, 670.

(68) Torres Pazmiño, D. E.; Winkler, M.; Glieder, A.; Fraaije, M. W. *J. Biotechnol.* **2010**, *146*, 9.

(69) Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. *Chem. Rev.* **2005**, *105*, 2253.

(70) Cox, R. *Nat. Prod. Rep.* **2014**, *31*, 1405.

(71) Geris, R.; Simpson, T. J. *Nat. Prod. Rep.* **2009**, *26*, 1063.

(72) Matsuda, Y.; Awakawa, T.; Abe, I. *Tetrahedron* **2013**, *69*, 8199.

(73) Itoh, T.; Tokunaga, K.; Radhakrishnan, E. K.; Fujii, I.; Abe, I.; Ebizuka, Y.; Kushiro, T. *ChemBioChem* **2012**, *13*, 1132.

(74) Guo, C. J.; Knox, B. P.; Chiang, Y. M.; Lo, H. C.; Sanchez, J. F.; Lee, K. H.; Oakley, B. R.; Bruno, K. S.; Wang, C. C. C. *Org. Lett.* **2012**, *14*, 5684.

(75) Matsuda, Y.; Awakawa, T.; Itoh, T.; Wakimoto, T.; Kushiro, T.; Fujii, I.; Ebizuka, Y.; Abe, I. *ChemBioChem* **2012**, *13*, 1738.

(76) Lo, H.; Entwistle, R.; Guo, C.; Ahuja, M.; Szewczyk, E.; Hung, J.; Chiang, Y.; Oakley, B. R.; Wang, C. C. C. *J. Am. Chem. Soc.* **2012**, *134*, 4709.

(77) Itoh, T.; Tokunaga, K.; Matsuda, Y.; Fujii, I.; Abe, I.; Ebizuka, Y.; Kushiro, T. *Nat. Chem.* **2010**, *2* (10), 858.

(78) Hansen, B. G.; Mnich, E.; Nielsen, K. F.; Nielsen, J. B.; Nielsen, M. T.; Mortensen, U. H.; Larsen, T. O.; Patil, K. R. *Appl. Environ. Microbiol.* **2012**, *78*, 4908.

(79) Holm, D. K.; Petersen, L. M.; Klitgaard, A.; Knudsen, P. B.; Jarczynska, Z. D.; Nielsen, K. F.; Gotfredsen, C. H.; Larsen, T. O.; Mortensen, U. H. *Chem. Biol.* **2014**, *21*, 519.

(80) Matsuda, Y.; Abe, I. *Nat. Prod. Rep.* **2016**, *33*, 26.

(81) Lin, H. C.; Chooi, Y. H.; Dhingra, S.; Xu, W.; Calvo, A. M.; Tang, Y. *J. Am. Chem. Soc.* **2013**, *135*, 4616.

(82) Ainsworth, A. M.; Chicarelli-robinson, M. I.; Copp, B. R.; Fauth, U.; Hylandsn, P. J.; Holloway, J. A.; Beirnem, G. B. O.; Porter, N.; Renno, D. V; Richards, M.; Robinson, N. *J. Antibiot. (Tokyo).* **1995**, 568.

(83) Pelosi, L.; Ducluzeau, A.-L.; Loiseau, L.; Barras, F.; Schneider, D.; Junier, I.; Pierrel, F. *mSystems* **2016**, *1*.

(84) Raggatt, M. E.; Simpson, T. J.; Inês Chicarelli Robinson, M. *Chem. Comm.* **1997**, 2245.

(85) Harris, G. H.; Hoogsteen, K.; Silverman, K. C.; Raghoobar, S. L.; Bills, G. F.; Lingham, R. B.; Smith, J. L.; Dougherty, H. W.; Cascales, C.; Peuezt, F. *Tetrahedron* **1993**, *49*, 2139.

(86) Mayerl, F.; Gao, Q.; Huang, S.; Klohr, S. E.; Matson, J. A.; Gustavson, D. R.; Pirnik, D. M.; Berry, R. L.; Fairchild, C.; Rose, W. C. *J. Antibiot. (Tokyo).* **1993**, 1082.

(87) Ayers, S.; Zink, D. L.; Powell, J. S.; Brown, C. M.; Grund, A.; Bills, G. F.; Platas, G.; Thompson, D.; Singh, S. B. *J. Nat. Prod.* **2008**, *71*, 457.

(88) Cai, P.; Smith, D.; Cunningham, B.; Brown-Shimer, S.; Katz, B.; Pearce, C.; Venables, D.; Houck, D. *J. Nat. Prod.* **1998**, *3864*, 791.

(89) Bunyapaiboonsri, T.; Veeranondha, S.; Boonruangprapa, T.; Somrithipol, S. *Phytochem. Lett.* **2008**, *1*, 204.

(90) Pittayakhajonwut, P.; Theerasilp, M.; Kongsaeree, P.; Rungrod, A.; Tanticharoen, M.; Thebtaranonth, Y. *Planta Med.* **2002**, *68*, 1017.

(91) Zhang, J.; Liu, L.; Wang, B.; Zhang, Y.; Wang, L.; Liu, X.; Che, Y. *J. Nat. Prod.* **2015**, *78*, 3058.

(92) Thomas, P.; Sundaram, H.; Krishek, B. J.; Chazot, P.; Xie, X.; Bevan, P.; Brocchini, S. J.; Latham, C. J.; Charlton, P.; Moore, M.; Lewis, S. J.; Thornton, D. M.; Stephenson, F. a; Smart, T. G. *J. Pharmacol. Exp. Ther.* **1997**, *282*, 513.

(93) Birkinshaw, H. J.; Chambers, A. R.; Raistrick, H. **1942**, 242.

(94) Bentley, R. *J. Bio. Chem.* **1963**, *238*, 1895.

(95) O'Sullivan, M. C.; Schwab, J. M. *Bioroganic Chem.* **1995**, 131.

(96) Byrant, R. W.; Light, R. J. *Biochemistry* **1974**, *13*, 1516.

(97) Davison, J.; al Fahad, A.; Cai, M.; Song, Z.; Yehia, S. Y.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 7642.

(98) al Fahad, A.; Abood, A.; Simpson, T. J.; Cox, R. J. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 7519.

(99) Hill, A.; Rayner, H. C.; Staunton, J. *J.C.S Chem. Comm.* **1976**, 380.

(100) Holker, J. S. E.; Young, K. *J.C.S. Chem. Comm.* **1975**, 525.

(101) Zaehle, C.; Gressler, M.; Shelest, E.; Geib, E.; Hertweck, C.; Brock, M. *Chem. Biol.* **2014**, *21*, 719.

(102) Bailey, A. M.; Cox, R. J.; Harley, K.; Lazarus, C. M.; Simpson, T. J.; Skellam, E. *Chem. Commun.* **2007**, *39*, 4053.

(103) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. *J. Mol. Biol.* **1990**, *215*, 403.

(104) Davison, J. The Genetic, Molecular and Biochemical Basis of Fungal Tropolone Synthesis, University of Bristol, 2012.

(105) Skellam, E. Investigating the Programming of Fungal Tetraketide Synthases, University of Bristol, 2007.

(106) Davison, J.; Cai, M.; Song, Z.; Yehia, S. Y.; Lazarus, C. M.; al Fahad, A.; Cai, M.; Song, Z.; Yehia, S. Y.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. *Proc. Natl. Acad. Sci.* **2012**, *109*, 7642.

(107) Yu, F.; Okamto, S.; Nakasone, K.; Adachi, K.; Matsuda, S.; Harada, H.; Misawa, N.; Utsumi, R. *Planta* **2008**, *227*, 1291.

(108) Alemdar, S.; Hartwig, S.; Frister, T.; König, J. C.; Scheper, T.; Beutel, S. *Appl. Biochem. Biotechnol.* **2016**, *178*, 474.

(109) Brock, N. L.; Huss, K.; Tudzynski, B.; Dickschat, J. S. *Chembiochem* **2013**, *14*, 311.

(110) Amby, D. B.; Manczak, T.; Petersen, M. A.; Sundelin, T.; Weitzel, C.; Grajewski, M.; Simonsen, H. T.; Jensen, B. *Microbiology* **2016**, *162*, 1773.

(111) Baldwin, J. E.; Mayweg, A. V.; Neumann, K.; Pritchard, G. J. *Org. Lett.* **1999**, *1*, 1933.

(112) Adlington, R. M.; Baldwin, J. E.; Pritchard, G. J.; Williams, A. J.; Watkin, D. J. *Org. Lett.* **1999**, 1937.

(113) Adlington, R. M.; Baldwin, J. E.; Mayweg, A. V. W.; Pritchard, G. J. *Org. Lett.* **2002**, *4*, 1933.

(114) Auclair, K.; Sutherland, A.; Kennedy, J.; Witter, D. J.; Heever, J. P. Van Den; Hutchinson, C. R.; Vederas, J. C. *J. Am. Chem. Soc.* **2000**, 11519.

(115) Tian, Z.; Sun, P.; Yan, Y.; Wu, Z.; Zheng, Q.; Zhou, S.; Zhang, H.; Yu, F.; Jia, X.; Chen, D.; Mándi, A.; Kurtán, T.; Liu, W. *Nat. Chem.*

Biol. **2015**, 259.
(116) Zheng, Q.; Guo, Y.; Yang, L.; Zhao, Z.; Wu, Z.; Zhang, H.; Liu, J.; Cheng, X. *Cell Chem. Biol.* **2016**, *23*, 352.
(117) Hashimoto, T.; Kuzuyama, T. *Curr. Opin. Chem. Biol.* **2016**, *35*, 117.
(118) Blackburn, M.; Fauth, U.; Katzer, W.; D, R.; Trew, S. *J. Ind. Microbiol.* **1996**, 36.
(119) Streeck, M. Linking Genes and Biosynthesis of Xenovulene A production in *Acremonium strictum*, 2015.
(120) Piettre, S. R.; Ganzhorn, A.; Hoflack, J.; Islam, K.; Hornsperger, J. *J. Am. Chem. Soc.* **1997**, *119*, 3201.
(121) Kühnel, E.; Laffan, D. D. P.; Lloyd-Jones, G. C.; Martínez Del Campo, T.; Shepperson, I. R.; Slaughter, J. L. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 7075.
(122) Dauben, H. J.; Ringold, H. J.; Detert, F. L.; Nozoe, T.; Mukai, T.; Takase, K.; Nagase, T. *Tetrahedron* **1962**, *18*, 859.
(123) Steinbeck, C.; Krause, S.; Kuhn, S. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1733.
(124) Castillo, A. M.; Patiny, L.; Wist, J. *J. Magn. Reson.* **2011**, *209*, 123.
(125) Banfi, D.; Patiny, L. *Chim. Int. J. Chem.* **2008**, *62*, 280.
(126) Aires-de-Sousa, J.; Hemmer, M. C.; Gasteiger, J. *Anal. Chem.* **2002**, *74*, 80.
(127) Harley, K. Investigating Fungal Tetraketide Synthase Gene Clusters, University of Bristol, 2006.
(128) Greenbaum, D.; Colangelo, C.; Gernstein, M.; Williams, K. *Genome Biol* **2003**, *4*, 117.
(129) Wibberg, D.; Andersson, L.; Tzelepis, G.; Rupp, O.; Blom, J.; Jelonek, L.; Pühler, A.; Fogelqvist, J.; Varrelmann, M.; Schlüter, A.; Dixelius, C. *BMC Genomics* **2016**, 245.
(130) Rupp, O.; Becker, J.; Brinkrolf, K.; Timmermann, C.; Borth, N.; Pühler, A.; Noll, T.; Goesmann, A. *PLoS One* **2014**, *9*, 245.
(131) Stanke, M.; Schöffmann, O.; Morgenstern, B.; Waack, S. *BMC Bioinformatics* **2006**, *7*, 62.
(132) Dopstadt, J.; Neubauer, L.; Tudzynski, P.; Humpf, H. U. *PLoS One* **2016**, *11*, 1.
(133) Chooi, Y.-H.; Cacho, R.; Tang, Y. *Chem. Biol.* **2010**, *17*, 483.
(134) Bergmann, S.; Schümann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. *Nat. Chem. Biol.* **2007**, *3*, 213.
(135) Wang, G.; Liu, Z.; Lin, R.; Li, E.; Mao, Z.; Ling, J.; Yang, Y.; Yin, W. B.; Xie, B. *PLoS Pathog.* **2016**, *12*, 1.
(136) Wight, W. D.; Kim, K.-H.; Lawrence, C. B.; Walton, J. D. *Mol. Plant-Microbe Interact. MPMI* **2009**, *22*, 1258.
(137) Toyomasu, T.; Nakaminami, K.; Toshima, H.; Mie, T.; Watanabe, K.; Ito, H.; Matsui, H.; Mitsuhashi, W.; Sassa, T.; Oikawa, H. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 146.
(138) Marchler-Bauer, A.; Bryant, S. H. *Nucleic Acids Res.* **2004**, *32*, 327.
(139) Marchler-Bauer, A.; Lu, S.; Anderson, J. B.; Chitsaz, F.; Derbyshire, M. K.; DeWeese-Scott, C.; Fong, J. H.; Geer, L. Y.; Geer, R. C.; Gonzales, N. R.; Gwadz, M.; Hurwitz, D. I.; Jackson, J. D.; Ke, Z.; Lanczycki, C. J.; Lu, F.; Marchler, G. H.; Mullokandov, M.; Omelchenko, M. V.; Robertson, C. L.; Song, J. S.; Thanki, N.; Yamashita, R. A.; Zhang, D.; Zhang, N.; Zheng, C.; Bryant, S. H. *Nucleic Acids Res.* **2011**, *39*, 225.
(140) Marchler-Bauer, A.; Derbyshire, M. K.; Gonzales, N. R.; Lu, S.; Chitsaz, F.; Geer, L. Y.; Geer, R. C.; He, J.; Gwadz, M.; Hurwitz, D. I.; Lanczycki, C. J.; Lu, F.; Marchler, G. H.; Song, J. S.; Thanki, N.; Wang, Z.; Yamashita, R. A.; Zhang, D.; Zheng, C.; Bryant, S. H. *Nucleic Acids Res.* **2015**, *43*, D222.
(141) Marchler-Bauer, A.; Bo, Y.; Han, L.; He, J.; Lanczycki, C. J.; Lu, S.; Chitsaz, F.; Derbyshire, M. K.; Geer, R. C.; Gonzales, N. R.; Gwadz, M.; Hurwitz, D. I.; Lu, F.; Marchler, G. H.; Song, J. S.; Thanki, N.; Wang, Z.; Yamashita, R. A.; Zhang, D.; Zheng, C.; Geer, L. Y.; Bryant, S. H. *Nucleic Acids Res.* **2017**, *45*, D200.
(142) Verwaaijen, B.; Wibberg, D.; Kröber, M.; Winkler, A.; Zrenner, R.; Bednarz, H.; Niehaus, K.; Grosch, R.; Pühler, A.; Schlüter, A. *PLoS One* **2017**, *12*, 1.
(143) Kim, D.; Pertea, G.; Trapnell, C.; Pimentel, H.; Kelley, R.; Salzberg, S. L. *Genome Biol.* **2013**, *14*, R36.
(144) Pearson, W. R.; Wood, T.; Zhang, Z.; Miller, W. *Genomics* **1997**, *46*, 24.
(145) Bolger, A. M.; Lohse, M.; Usadel, B. *Bioinformatics* **2014**, *30*, 2114.
(146) Anders, S.; Huber, W. *Genome Biol* **2010**, *11*, R106.
(147) Benjamini, Y.; Hochberg, Y. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289.
(148) Kelly, L. A.; Mezulis, S.; Yates, C.; Wass, M.; Sternberg, M. *Nat. Protoc.* **2015**, *10*, 845.
(149) Lo, H.; Entwistle, R.; Guo, C.; Ahuja, M.; Szewczyk, E.; Hung, J.; Chiang, Y.; Oakley, B. R.; Wang, C. C. C. *J. Am. Chem. Soc.* **2012**, *134*, 4709.
(150) Starks, C. M.; Back, K.; Chappell, J.; Noel, J. P. **1997**, *277*.
(151) Lesburg, C. A.; Zhai, G.; Cane, D. E.; Christianson, D. W. *Sci.* **1997**, *277*, 1820.
(152) Liang, W. L.; Le, X.; Li, H. J.; Yang, X. L.; Chen, J. X.; Xu, J.; Liu, H. L.; Wang, L. Y.; Wang, K. T.; Hu, K. C.; Yang, D. P.; Lan, W. J. *Mar. Drugs* **2014**, *12*, 5657.
(153) Rice, P.; Longden, I.; Bleasby, A. *Trends Genet.* **2000**, *16*, 276.
(154) Li, W.; Cowley, A.; Uludag, M.; Gur, T.; McWilliam, H.; Squizzato, S.; Park, Y. M.; Buso, N.; Lopez, R. *Nucleic Acids Res.* **2015**, *43*, W580.
(155) McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y. M.; Buso, N.; Cowley, A. P.; Lopez, R. *Nucleic Acids Res.* **2013**, *41*, 597.
(156) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. *Bioinformatics* **2006**, *22*, 195.
(157) Guex, N.; Peitsch, M. C.; Schwede, T. *Electrophoresis* **2009**, *30*, 162.
(158) Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Cassarino, T. G.; Bertoni, M.; Bordoli, L.; Schwede, T. *Nucleic Acids Res.* **2014**, *42*, 252.
(159) Carver, T. J.; Rutherford, K. M.; Berriman, M.; Rajandream, M. A.; Barrell, B. G.; Parkhill, J. *Bioinformatics* **2005**, *21*, 3422.
(160) Rice, P.; Longden, I.; Bleasby, A. *Trends Genet.* **2000**, *16*, 276.
(161) Chooi, Y. H.; Hong, Y. J.; Cacho, R. A.; Tantillo, D. J.; Tang, Y. *J. Am. Chem. Soc.* **2013**, *135*, 16805.
(162) Oikawa, H.; Toyomasu, T.; Toshima, H.; Ohashi, S.; Kawaide, H.; Kamiya, Y.; Ohtsuka, M.; Shinoda, S.; Mitsuhashi, W.; Sassa, T. *J. Am. Chem. Soc.* **2001**, *123*, 5154.
(163) Fujii, R.; Minami, A.; Tsukagoshi, T.; Sato, N.; Sahara, T.; Ohgiya, S.; Gomi, K.; Oikawa, H. *Biosci. Biotechnol. Biochem.* **2011**, *75*, 1813.
(164) Clough, S.; Raggatt, M. E.; Simpson, T. J.; Willis, C. L.; Whiting, A.; Wrigley, S. K. *J. Chem. Soc. Perkin Trans. 1* **2000**, *15*, 2475.
(165) Schoch, C. L.; Seifert, K. a; Huhndorf, S.; Robert, V.; Spouge, J. L.; Levesque, C. A.; Chen, W. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 6241.
(166) Etienne, K. A.; Roe, C. C.; Smith, R. M.; Vallabhaneni, S.; Duarte, C.; Escandón, P.; Castañeda, E.; Gómez, B. L.; Bedout, C. De; López, L. F.; Salas, V.; Hederra, L. M.; Fernández, J.; Pidal, P.; Hormazabel, J. C.; Otaíza-O'ryan, F.; Vannberg, F. O.; Gillece, J.; Lemmer, D.; Driebe, E. M.; Engelthaler, D. M.; Litvintseva, A. P. *Emerg. Infect. Dis.* **2016**, *22*, 476.

(167) Wibberg, D.; Rupp, O.; Blom, J.; Jelonek, L.; Kröber, M.; Verwaaijen, B.; Goesmann, A.; Albaum, S.; Grosch, R.; Pühler, A.; Schlüter, A. *PLoS One* **2015**, *10*, 1.
(168) Mohanta, T. K.; Bae, H. *Biol. Proced. Online* **2015**, *17*, 8.
(169) Guo, L.; Han, L.; Yang, L.; Zeng, H.; Fan, D.; Zhu, Y.; Feng, Y.; Wang, G.; Peng, C.; Jiang, X.; Zhou, D.; Ni, P.; Liang, C.; Liu, L.; Wang, J.; Mao, C.; Fang, X.; Peng, M.; Huang, J. *PLoS One* **2014**, *9*.
(170) Awakawa, T.; Zhang, L.; Wakimoto, T.; Hoshino, S.; Mori, T.; Ito, T.; Ishikawa, J.; Tanner, M. E.; Abe, I. *J. Am. Chem. Soc.* **2014**, *136*, 9910.
(171) 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.
(172) 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.
(173) de Groot, M. J.; Bundock, P.; Hooykaas, P. J.; Beijersbergen, a G. *Nat. Biotechnol.* **1998**, *16*, 839.
(174) Fincham, J. R. *Microbiol. Rev.* **1989**, *53*, 148.
(175) Chakraborty, B. N.; Kapoor, M. *Nucleic Acids Res.* **1990**, *18*, 6737.
(176) Armaleo, D.; Ye, G. N.; Klein, T. M.; Shark, K. B.; Sanford, J. C.; Johnston, S. A. *Curr. Genet.* **1990**, *17*, 97.
(177) Gritz, L.; Davies, J. *Gene* **1983**, *25*, 179.
(178) Weld, R. J.; Plummer, K. M.; Carpenter, M. A.; Ridgway, H. J. *Cell Res.* **2006**, *16*, 31.
(179) Lieber, M. R. *Annu. Rev. Biochem.* **2010**, *79*, 181.
(180) Hua, S. B.; Qiu, M.; Chan, E.; Zhu, L.; Luo, Y. *Plasmid* **1997**, *38*, 91.
(181) Sung, P. *Science* **1994**, *265*, 1241.
(182) Haber, J. E. *Trends Genet.* **2000**, *16*, 259.
(183) San Filippo, J.; Sung, P.; Klein, H. *Annu. Rev. Biochem.* **2008**, *77*, 229.
(184) Krappmann, S. *Fungal Biol. Rev.* **2007**, *21*, 25.
(185) Watson, R. J.; Burchat, S.; Bosley, J. *Fungal Genet. Biol.* **2008**, *45*, 1348.
(186) Ramsden, D. A.; Geliert, M. *EMBO J.* **1998**, *17*, 609.
(187) Critchlow, S. E.; Jackson, S. P. *Trends Biochem. Sci.* **1998**, *23*, 394.
(188) Ninomiya, Y.; Suzuki, K.; Ishii, C.; Inoue, H. **2004**, *101*, 12248.
(189) Nielsen, M. L.; Albertsen, L.; Lettier, G.; Nielsen, J. B.; Mortensen, U. H. *Fungal Genet. Biol.* **2006**, *43*, 54.
(190) Sander, J. D.; Joung, J. K. *Nat. Biotechnol.* **2014**, *32*, 347.
(191) Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E. *Science* **2012**, *337*, 816.
(192) Liu, R.; Chen, L.; Jiang, Y.; Zhou, Z.; Zou, G. *Cell Discov.* **2015**, *1*, 15007.
(193) Nakayashiki, H. *FEBS Lett.* **2005**, *579*, 5950.
(194) Dang, Y.; Yang, Q.; Xue, Z.; Liu, Y. *Eukaryot. Cell* **2011**, *10*, 1148.
(195) Meister, G.; Tuschi, T. *Nature* **2004**, *431*, 343.
(196) Tsuchiya, K.; Tada, S.; Gomi, K.; Kitamoto, K.; Kumagai, C.; Tamura, G. *Biosci. Biotechnol. Biochem.* **1992**, *11*, 1849.
(197) Zhang, G.; Gurtu, V.; Kain, S. R. *Biochem. Biophys. Res. Commun.* **1996**, *227*, 707.
(198) Nødvig, C. S.; Nielsen, J. B.; Kogle, M. E.; Mortensen, U. H. *PLoS One* **2015**, *10*, e0133085.
(199) Wilson, W. D.; Tanious, F. A.; Barton, H. J.; Strekowski, L.; Boykin, D. W. *J. Am. Chem. Soc.* **1989**, *111*, 5008.
(200) Gietz, R. D. *Methods Mol. Biol.* **2014**, *1205*, 1.
(201) Lazarus, C. M.; Williams, K.; Bailey, A. M. *Nat. Prod. Rep.* **2014**, *31*, 1339.
(202) Collemare, J.; Pianfetti, M.; Houlle, A. E.; Morin, D.; Camborde, L.; Gagey, M. J.; Barbisan, C.; Fudal, I.; Lebrun, M. H.; Böhnert, H. U. *New Phytol.* **2008**, *179*, 196.
(203) Bird, D.; Bradshaw, R. *Mol. Gen. Genet.* **1997**, *255*, 219.
(204) Pohl, C.; Kiel, J. A. K. W.; Driessen, A. J. M.; Bovenberg, R. A. L.; Nygård, Y. *ACS Synth. Biol.* **2016**, *5*, 754.
(205) Gems, D.; Johnstone, I. L.; Clutterbuck, A. J. *Gene* **1991**, *98*, 61.
(206) Ruiz-Díez, B. *J. Appl. Microbiol.* **2002**, *92*, 189.
(207) Geu-Flores, F.; Nour-Eldin, H. H.; Nielsen, M. T.; Halkier, B. A. *Nucleic Acids Res.* **2007**, *35*, e55.
(208) Gao, Y.; Zhao, Y. *J. Integr. Plant Biol.* **2014**, *56*, 343.
(209) 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.
(210) Trenti, F.; Cox, R. J. *J. Nat. Prod.* **2017**, *80*, 7.
(211) He, Y.; Liu, Q.; Shao, Y.; Chen, F. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 4965.
(212) Goldoni, M.; Azzalin, G.; Macino, G.; Cogoni, C. **2004**, *41*, 1016.
(213) Kadotani, N.; Nakayashiki, H.; Tosa, Y.; Mayama, S. *Mol. Plant. Microbe. Interact.* **2003**, *16*, 769.
(214) Namekawa, S. H.; Iwabata, K.; Sugawara, H.; Hamada, F. N.; Koshiyama, A.; Chiku, H.; Kamada, T.; Sakaguchi, K. *Microbiology* **2005**, 151, 3669.
(215) Hammond, T. M.; Bok, J. W.; Andrewski, M. D.; Reyes-Domínguez, Y.; Scazzocchio, C.; Keller, N. P. *Eukaryot. Cell* **2008**, *7*, 339.
(216) Kim, D.; Bae, S.; Park, J.; Kim, E.; Kim, S.; Yu, H. R.; Hwang, J.; Kim, J.; Kim, J. *Nat. Methods* **2015**, *12*, 237.
(217) Xie, S.; Shen, B.; Zhang, C.; Huang, X.; Zhang, Y. *PLoS ONE* **2014**, *9*, 1.
(218) Krappmann, S. *Med. Mycol.* **2017**, *55*, 16.
(219) 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.
(220) Matsuda, Y.; Wakimoto, T.; Mori, T.; Awakawa, T.; Abe, I. *J. Am. Chem. Soc.* **2014**, *136*, 15326.
(221) Jin, F. J.; Maruyama, J. I.; Juvvadi, P. R.; Arioka, M.; Kitamoto, K. *FEMS Microbiol. Lett.* **2004**, *239*, 79.
(222) Hartley, J. L.; Temple, G. F.; Brasch, M. A. *Genome Res.* **2000**, *10*, 1788.
(223) Al Fahad, A. Tropolone and Sorbicillactone Biosynthesis in Fungi, University of Bristol, 2014.
(224) Doi, K. *Bull. Chem. Soc. Jpn.* **1962**, *35*, 67.
(225) Ma, X.; Panjikar, S.; Koepke, J.; Loris, E.; Stöckigt, J. *Plant Cell* **2006**, *18*, 907.
(226) Chen, C. K. M.; Chan, N. L.; Wang, A. H. J. *Trends Biochem. Sci.* **2011**, *36*, 553.
(227) Meganathan, R. *FEMS Microbiol. Lett.* **2001**, *203*, 131.
(228) Moonen, M. J. H.; Fraaije, M. W.; Rietjens, I. M. C. M.; Laane, C.; Van Berkel, W. J. H. *Adv. Synth. Catal.* **2002**, *344*, 1023.
(229) Entsch, B.; Berkel, J. H. *FASEB J.* **1995**, *9*, 476.
(230) Merenyi, G.; Lind, J.; Anderson, R. F. *J. Am. Chem. Soc.* **1991**, *113*, 9371.

(231) Ahmed, S. A.; Simpson, T. J.; Staunton, J.; Sutkowski, A. C.; Trimble, L. A.; Vederas J. C. *Chem. Comm.* **1985**, 1685.
(232) Guroff, G.; Daly, J. W.; Jerina, D. M.; Renson, J.; Witkop, B.; Udenfriend, S. *Science* **1967**, *157*, 1524.
(233) Carr, R. T.; Balasubramanian, S.; Hawkins, P. C. D.; Benkovic, S. J. *Biochemistry* **1995**, *34*, 7525.
(234) Byrne, M. J.; Lees, N. R.; Han, L. C.; Van Der Kamp, M. W.; Mulholland, A. J.; Stach, J. E. M.; Willis, C. L.; Race, P. R. *J. Am. Chem. Soc.* **2016**, *138*, 6095.
(235) Nielsen, M. L.; Isbrandt, T.; Petersen, L. M.; Mortensen, U. H.; Andersen, M. R.; Hoof, J. B.; Larsen, T. O. *PLoS One* **2016**, *11*, 1.
(236) Preindl, J.; Jouvin, K.; Laurich, D.; Seidel, G.; Fürstner, A. *Chem. - A Eur. J.* **2016**, *22*, 237.

9 Appendix

9.1 Chapter 2 NMR spectra

9.1.1 Compound 108

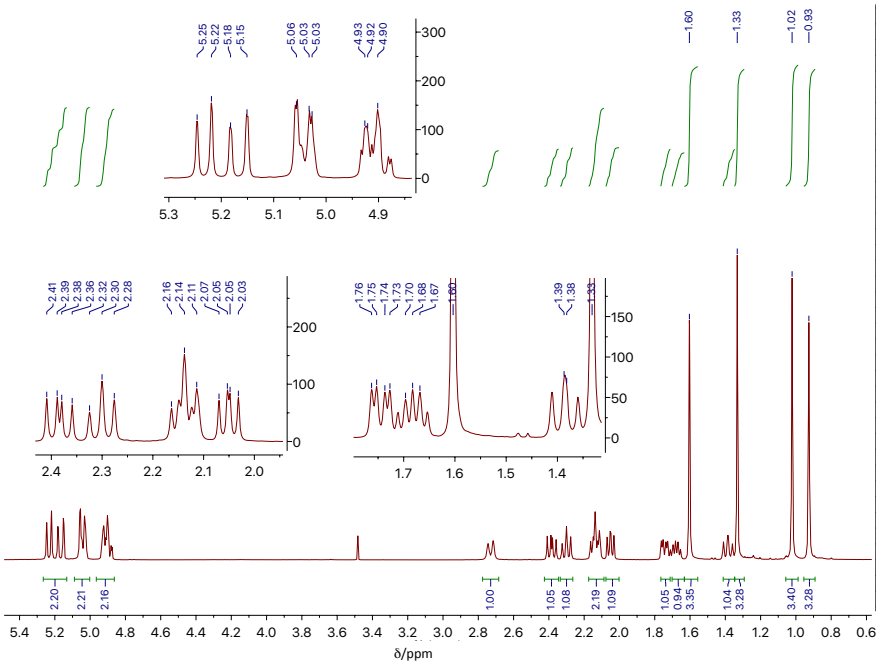


Figure 9.1 ^{1}H NMR of compound **108** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.2** for assignment.

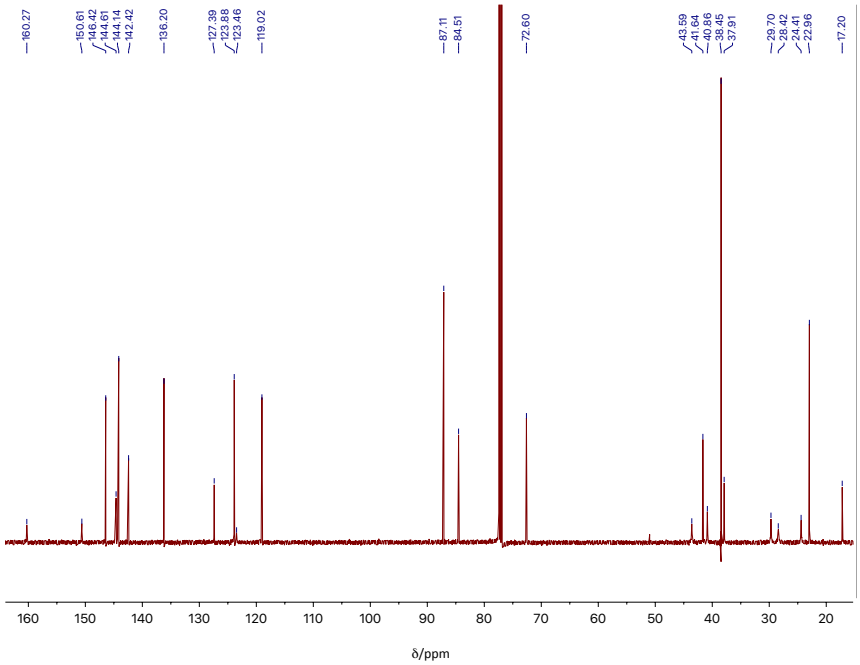


Figure 9.2 ^{13}C NMR of compound **108** in $CDCl_3$ (125MHz) referenced to $CDCl_3$. **Table 2.2** for assignment.

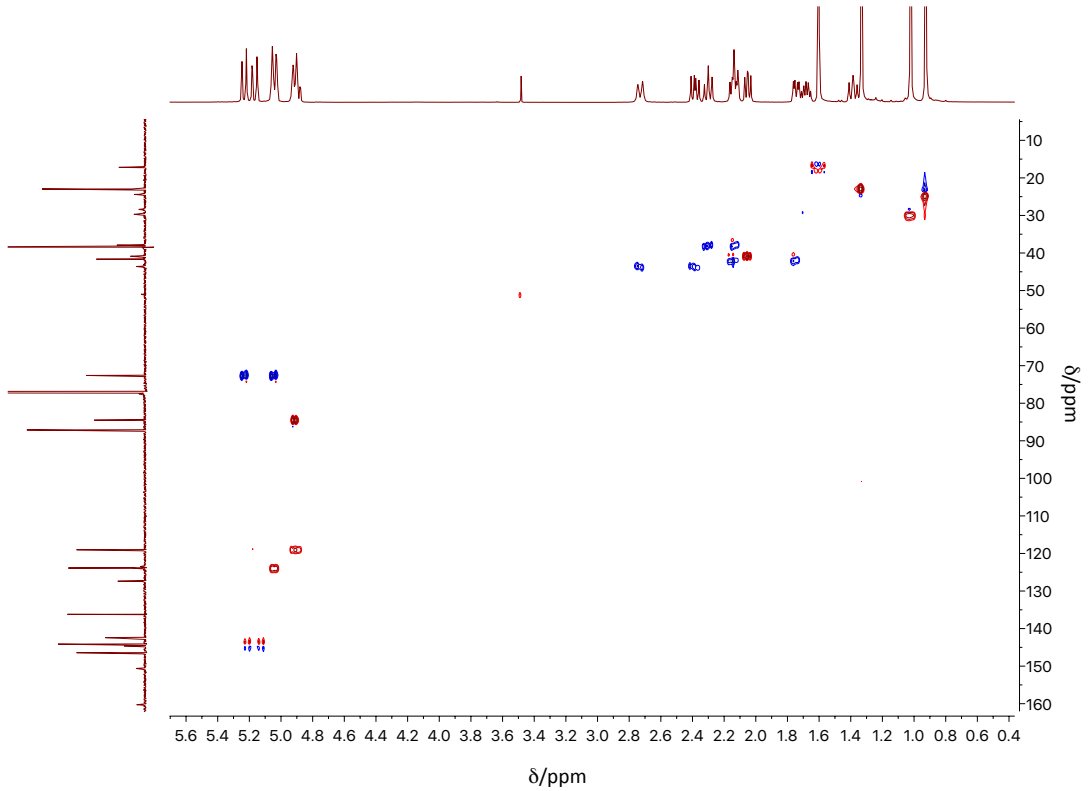


Figure 9.3 HSQC of compound **108.**

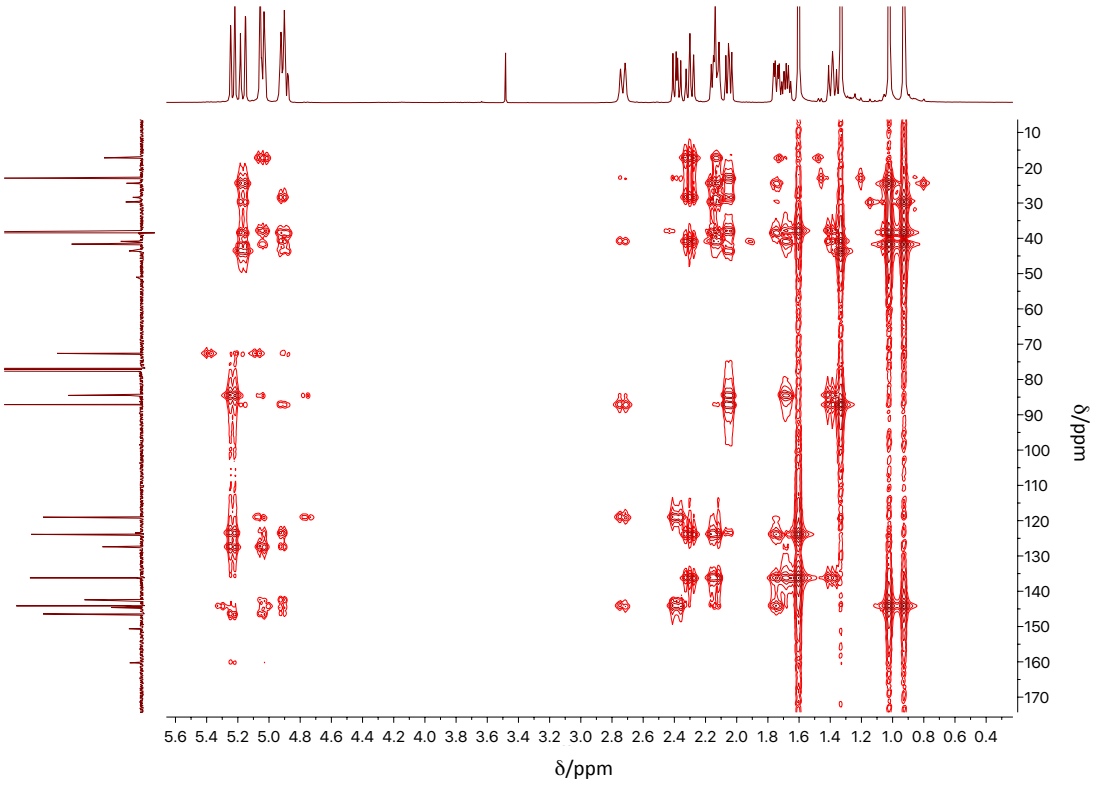


Figure 9.4 HMBC of compound **108**.

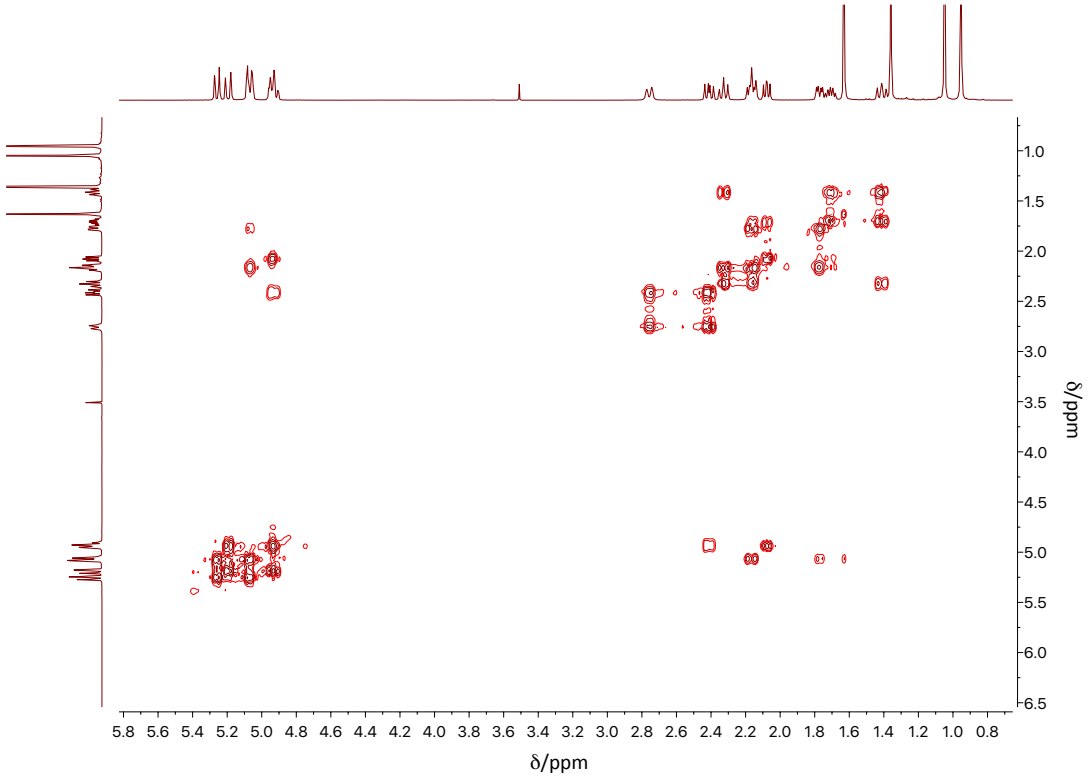


Figure 9.5 ^{1}H,^{1}H COSY of compound **108**.

9.1.2 Compound 153d

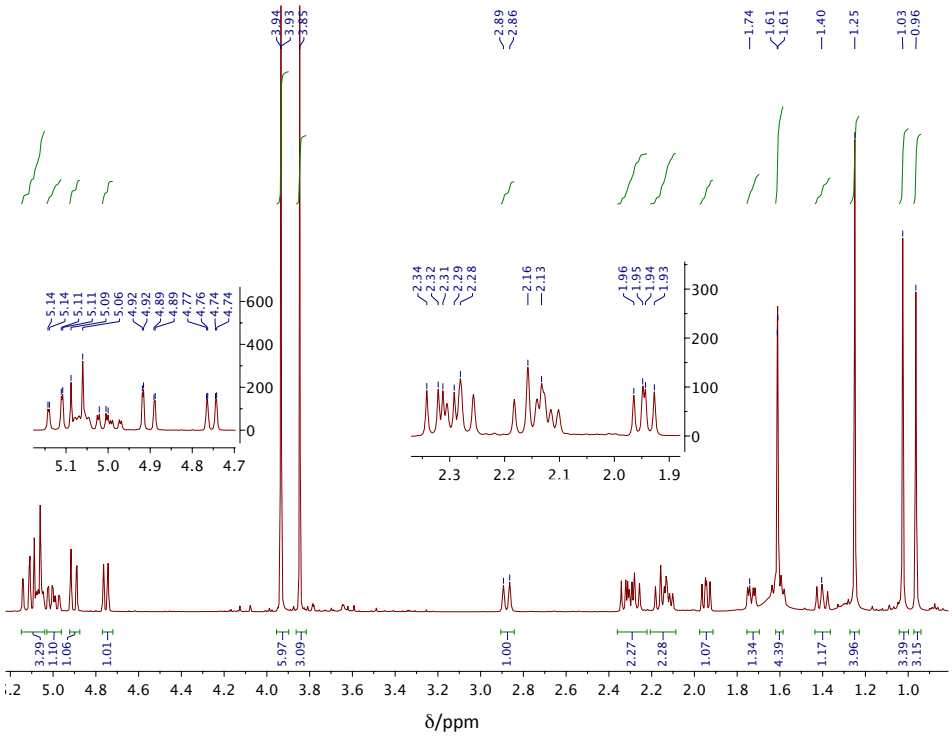


Figure 9.6 ^{1}H NMR of compound **153d** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.3** for assignment.

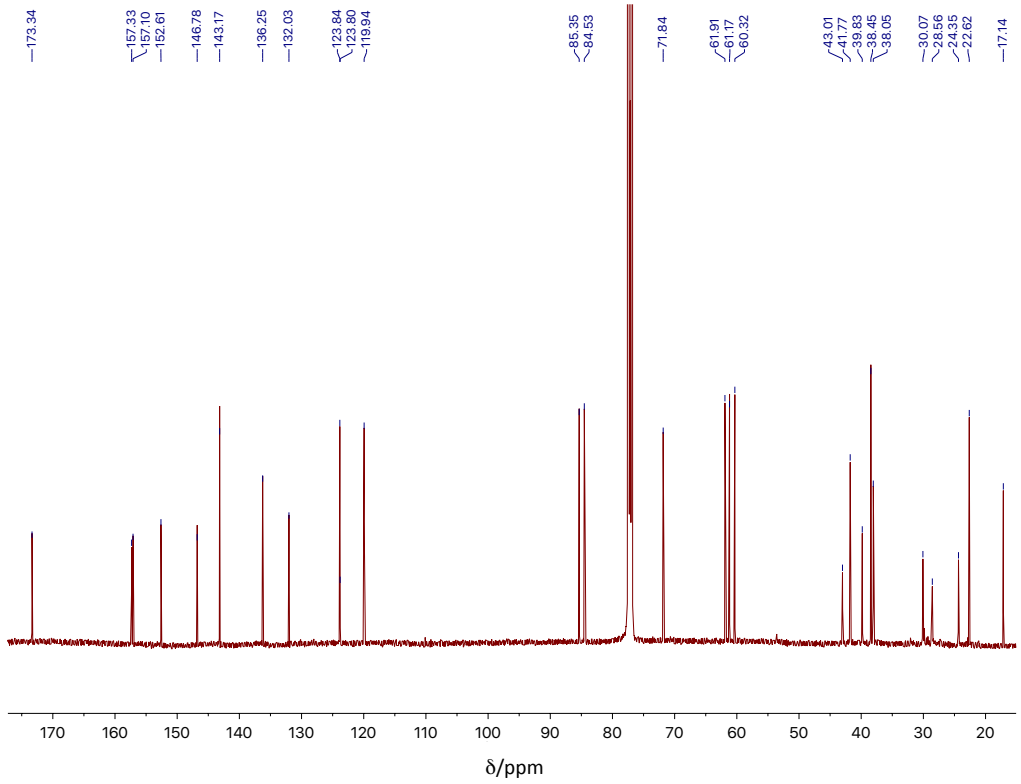


Figure 9.7 ^{13}C NMR of compound **153d** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.3** for assignment.

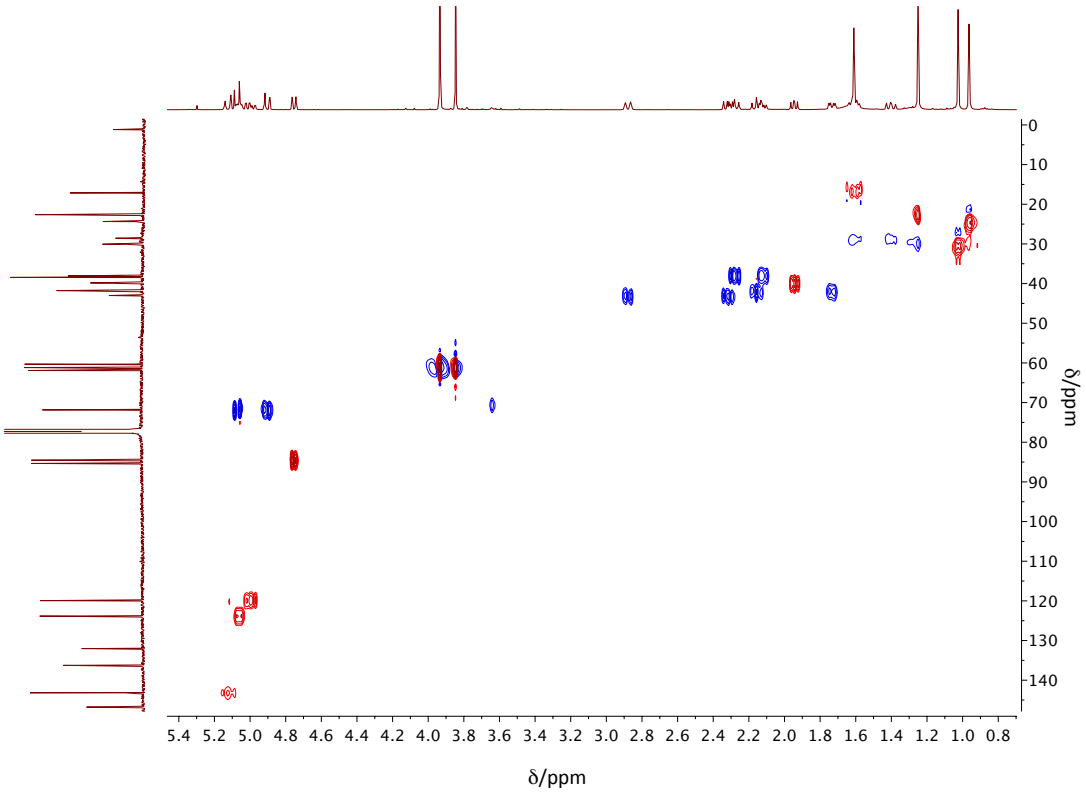


Figure 9.8 HSQC of compound **153d**.

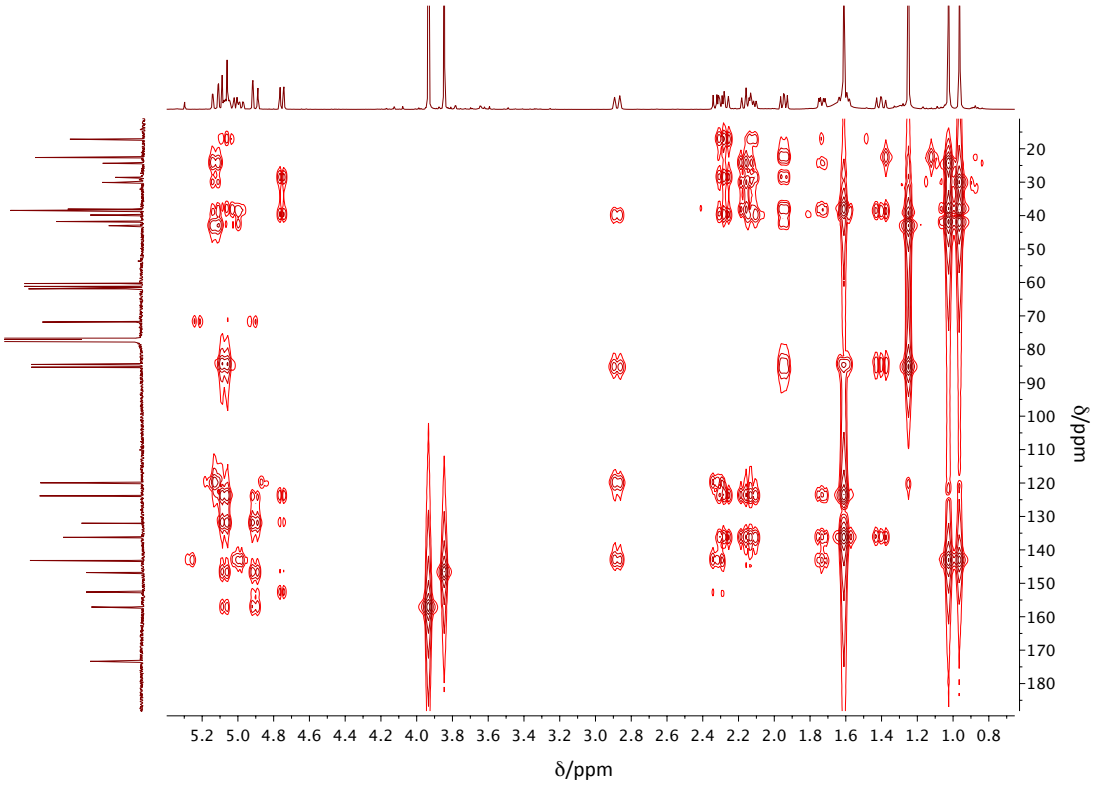


Figure 9.9 HMBC of compound **153d**.

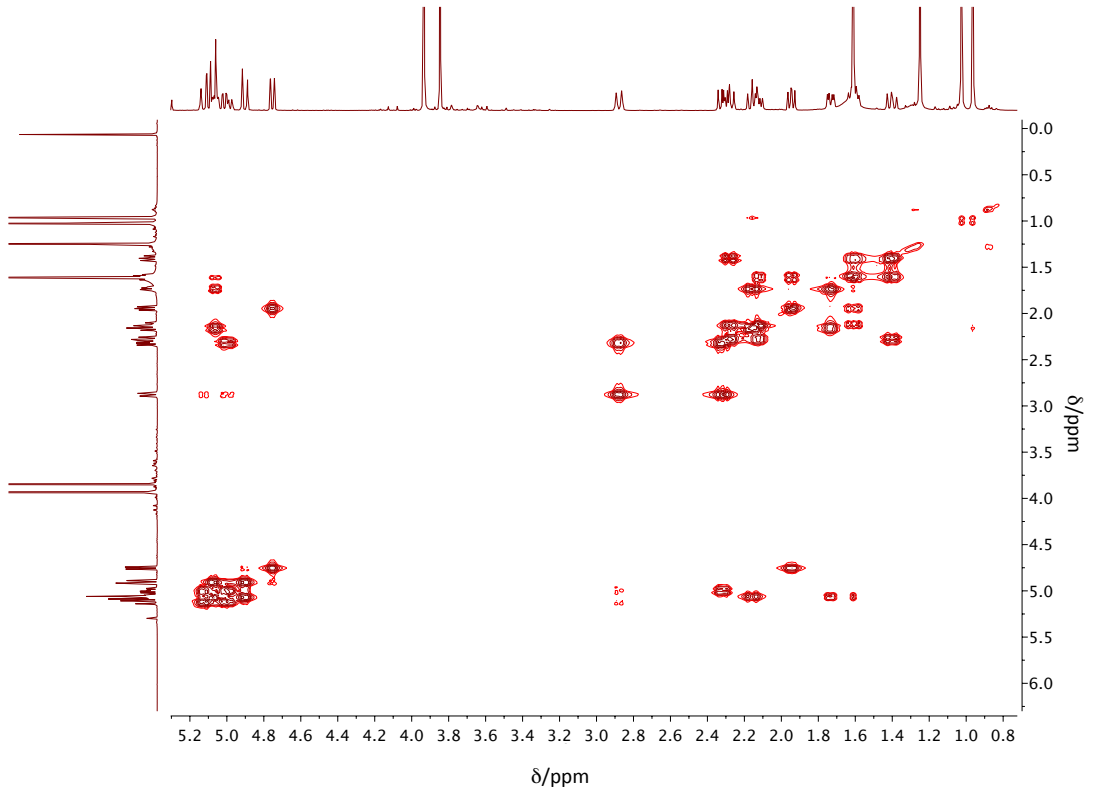


Figure 9.10 ^{1}H,^{1}H COSY of compound **153d**.

9.1.3 Compound 107a

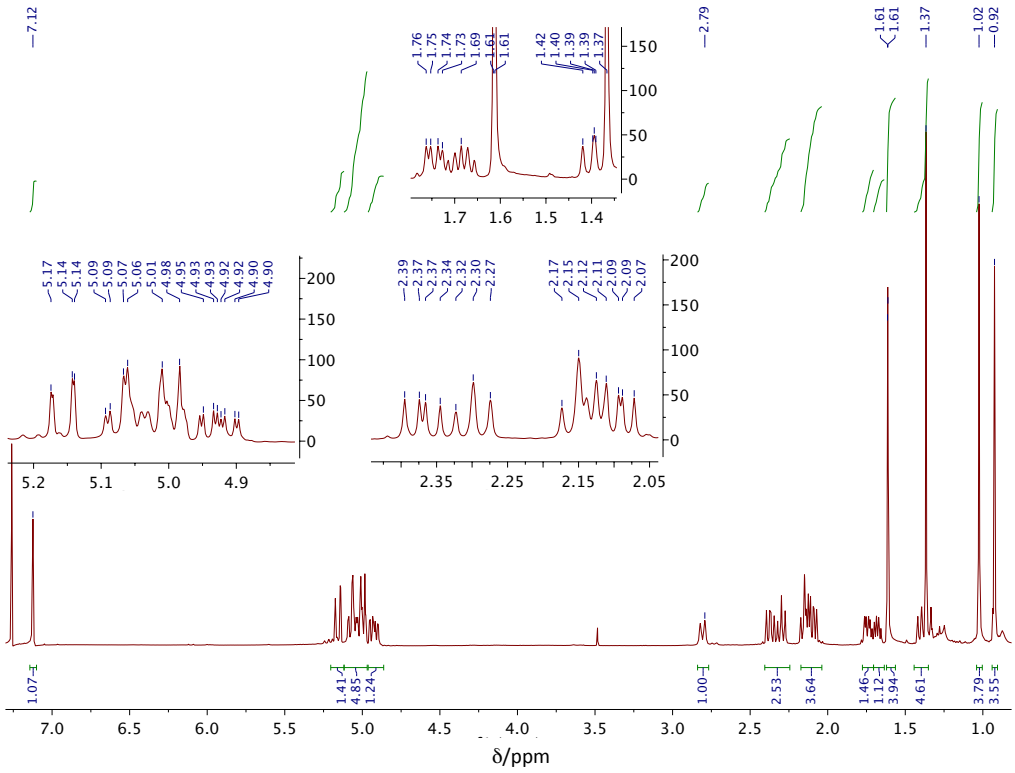


Figure 9.11 ^{1}H NMR of compound **107a** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.4** for assignment.

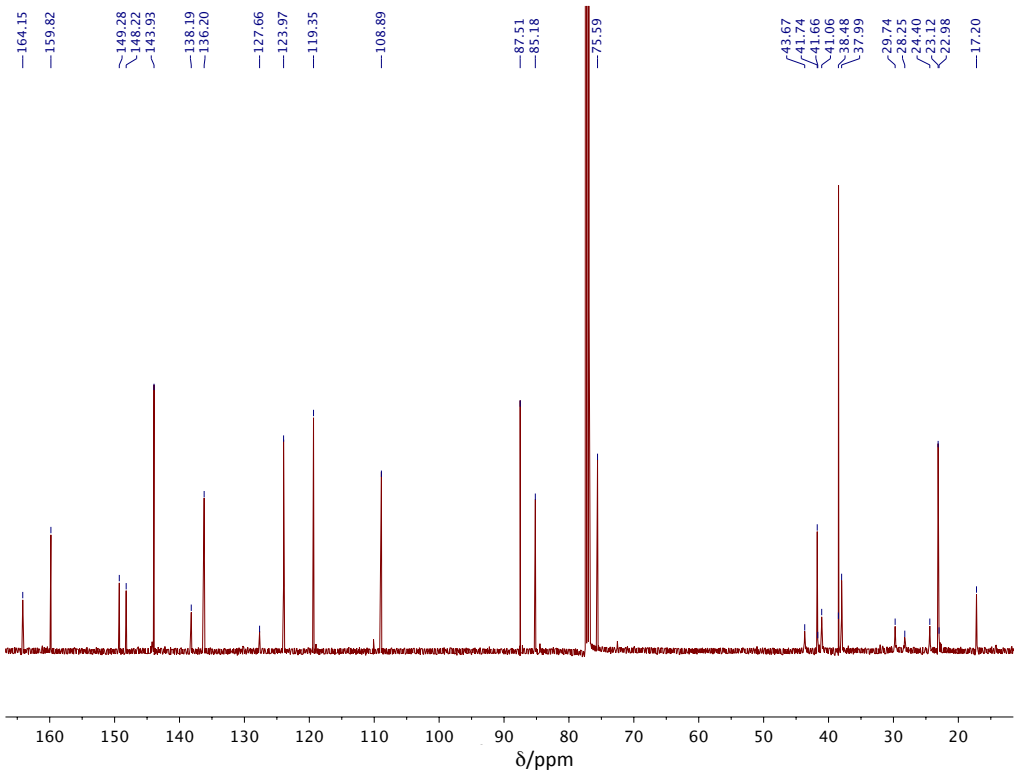


Figure 9.12 ^{13}C NMR of compound **107a** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.4** for assignment.

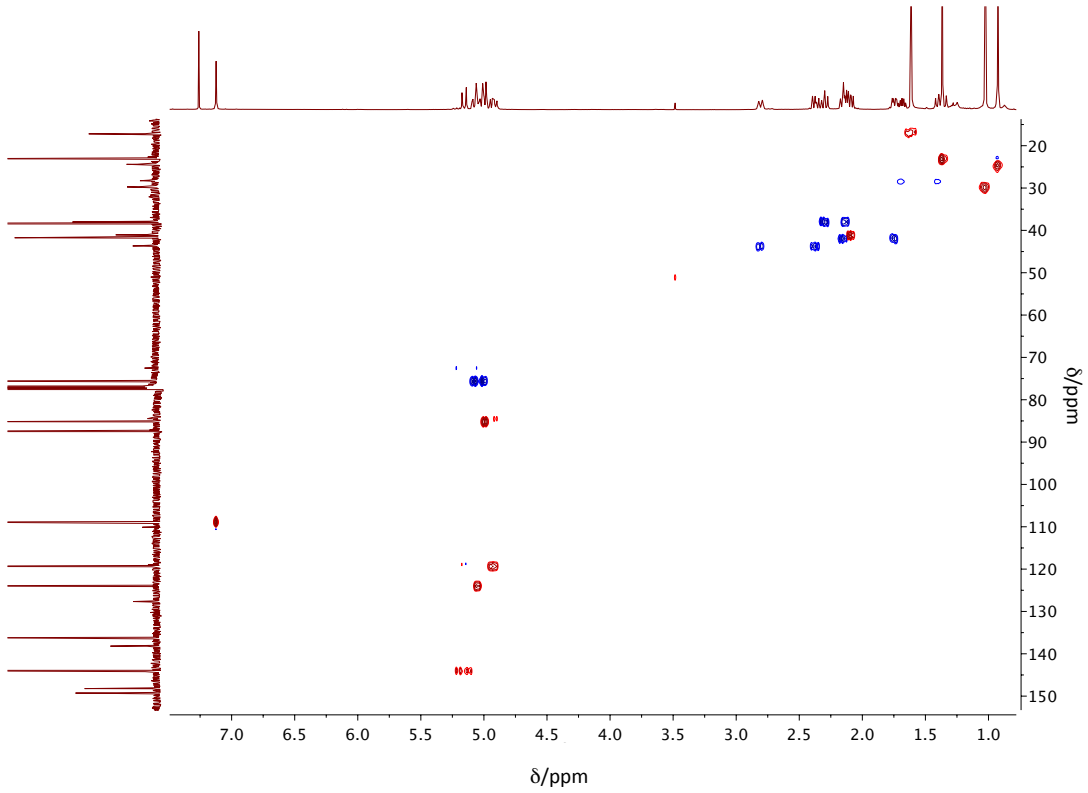


Figure 9.13 HSQC of compound **107a**.

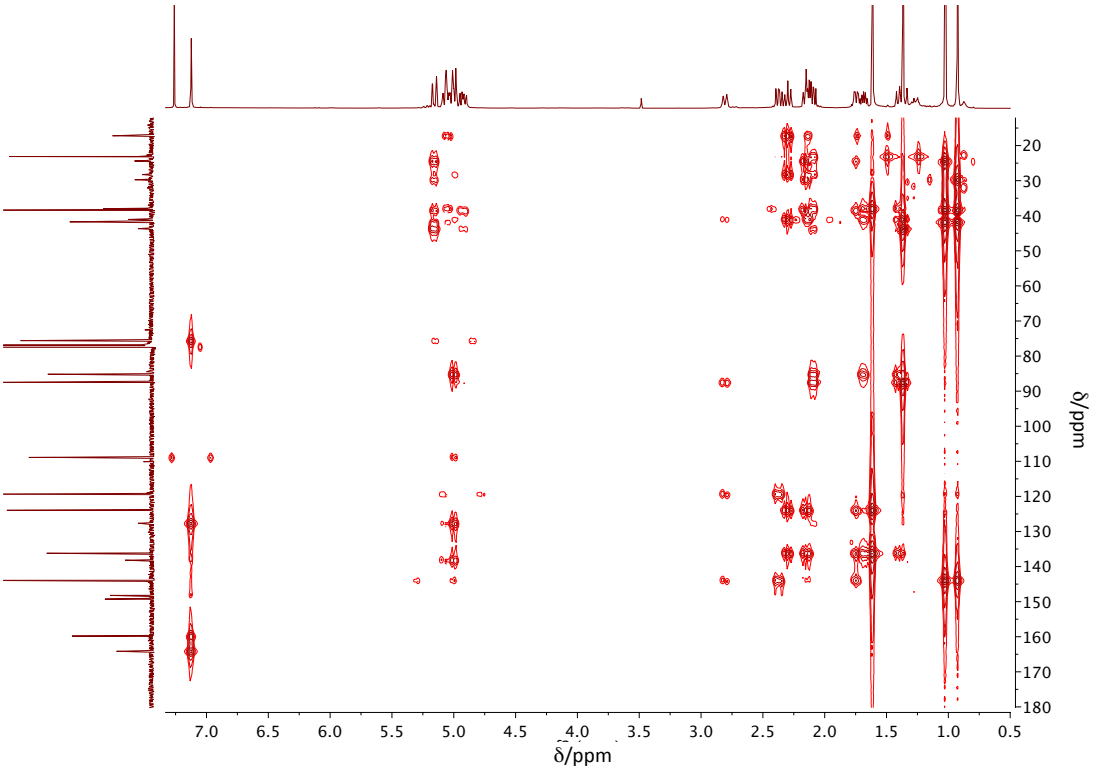


Figure 9.14 HMBC of compound **107a**.

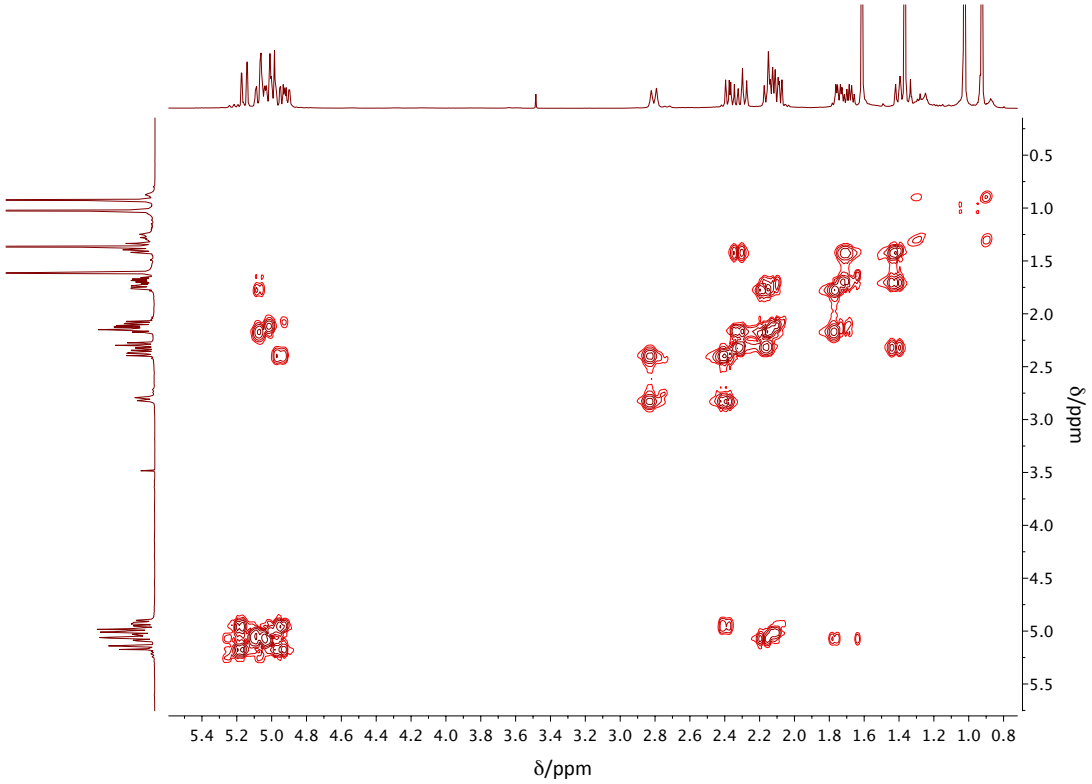


Figure 9.15 $^{1}H,^{1}H$ COSY of compound **107a**.

9.1.4 Compound 107b

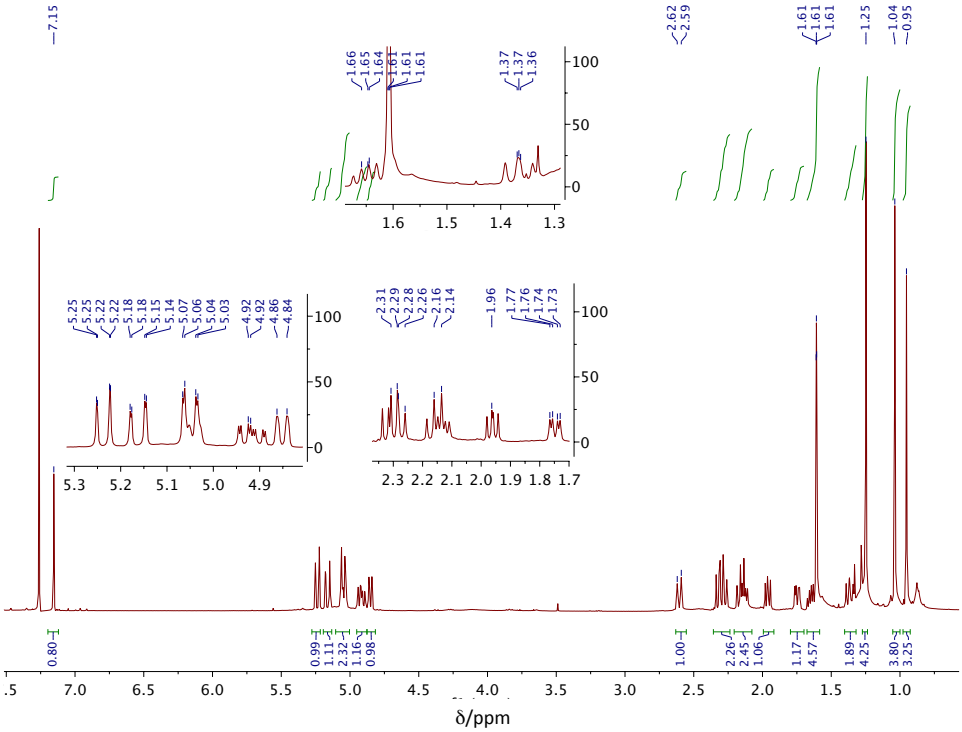


Figure 9.16 ^{1}H NMR of compound **107b** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.4** for assignment.

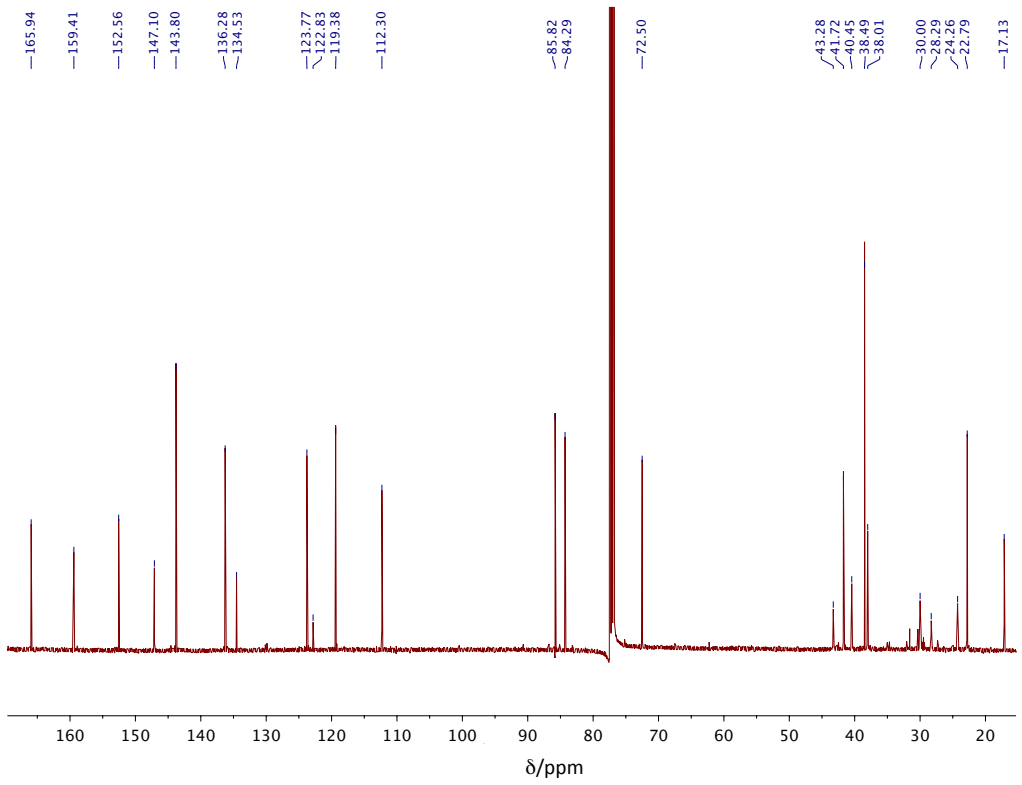


Figure 9.17 ^{13}C NMR of compound **107b** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.4** for assignment.

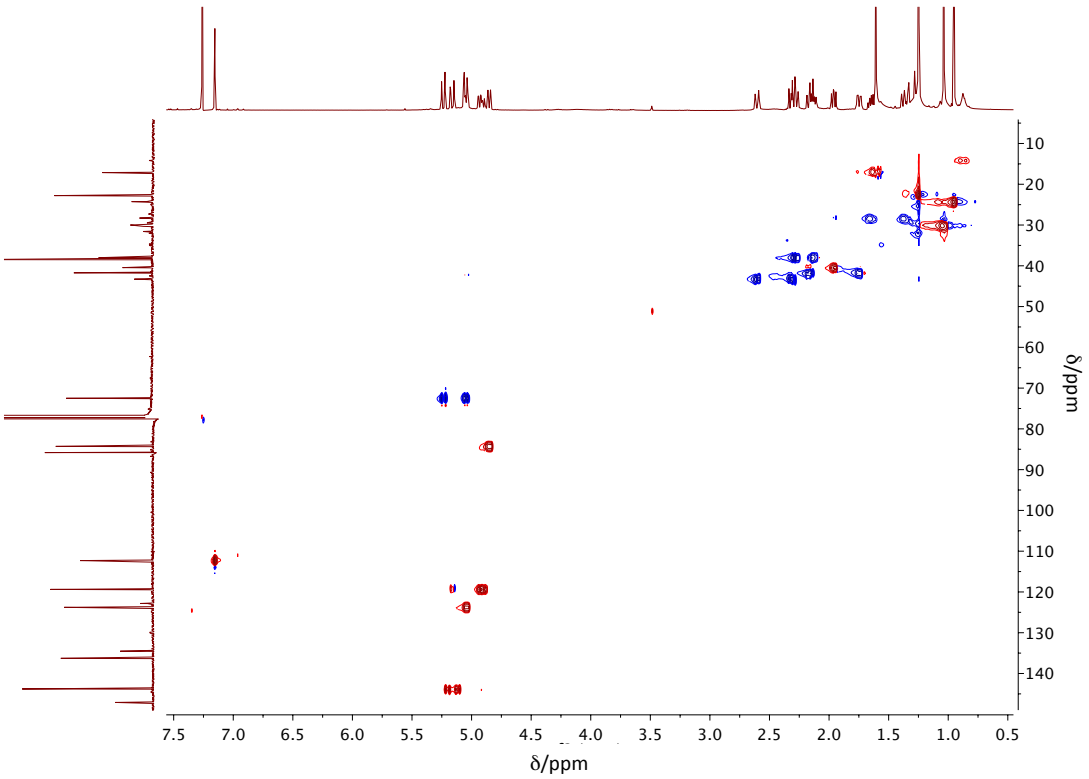


Figure 9.18 HSQC of compound **107b**.

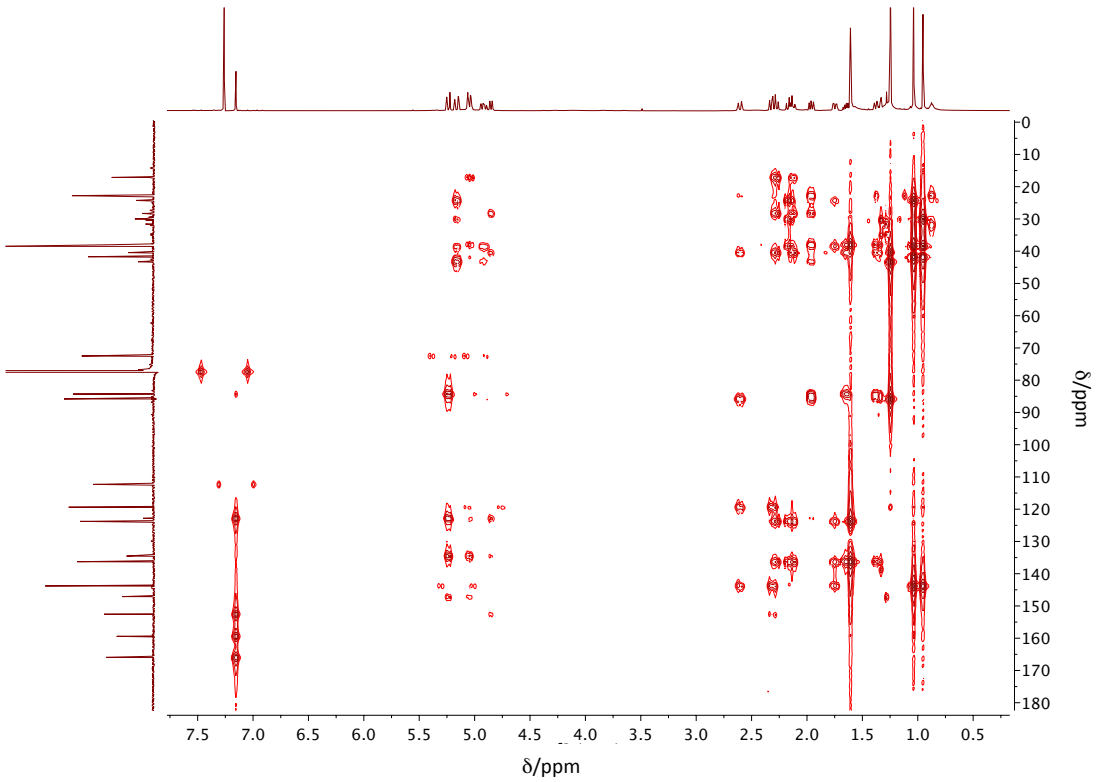


Figure 9.19 HMBC of compound **107b**.

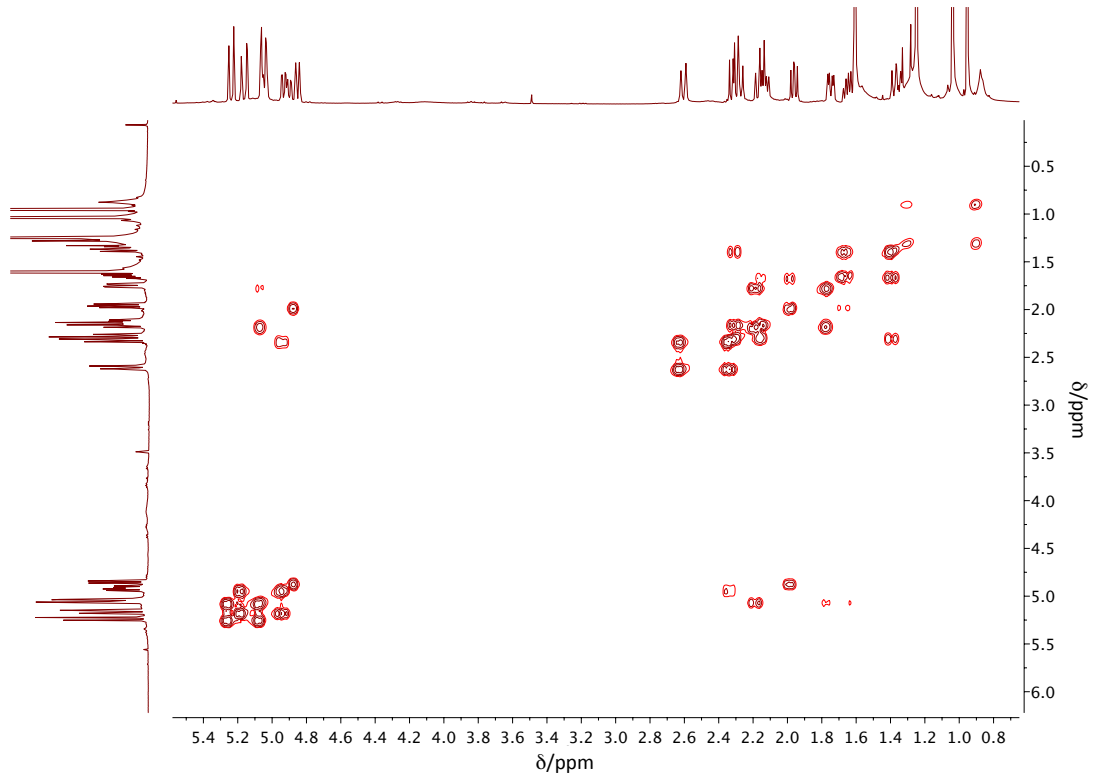


Figure 9.20 ^{1}H,^{1}H COSY of compound **107b**.

9.1.5 Compound 154

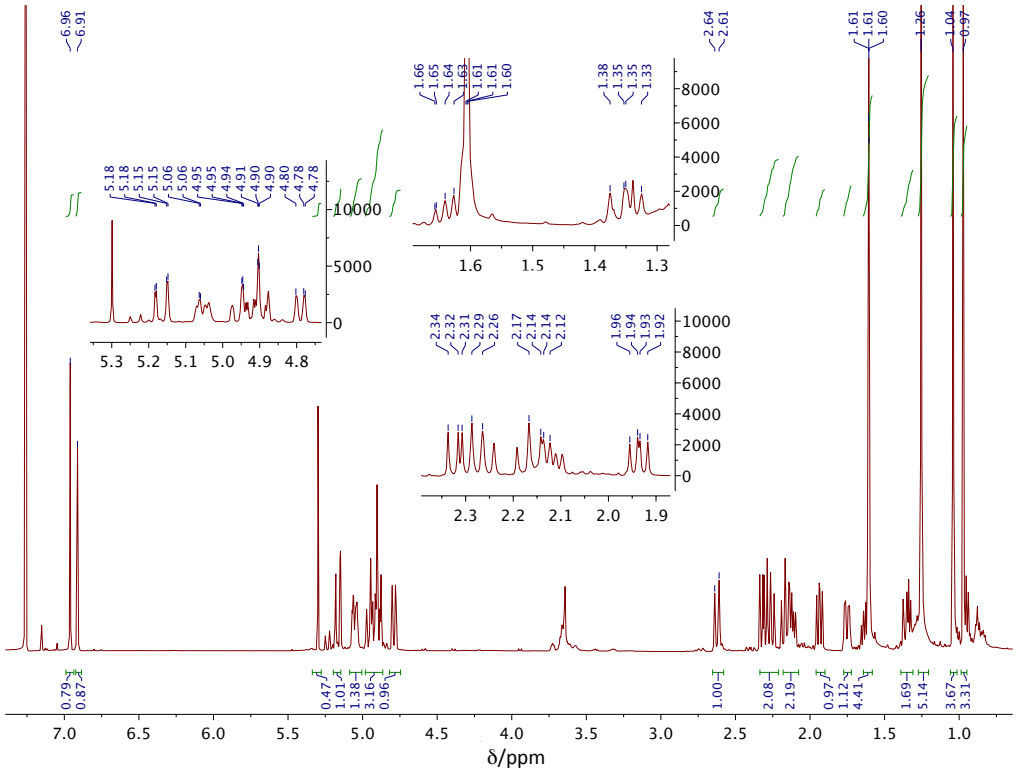


Figure 9.21 ^{1}H NMR of compound **154** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.5** for assignment.

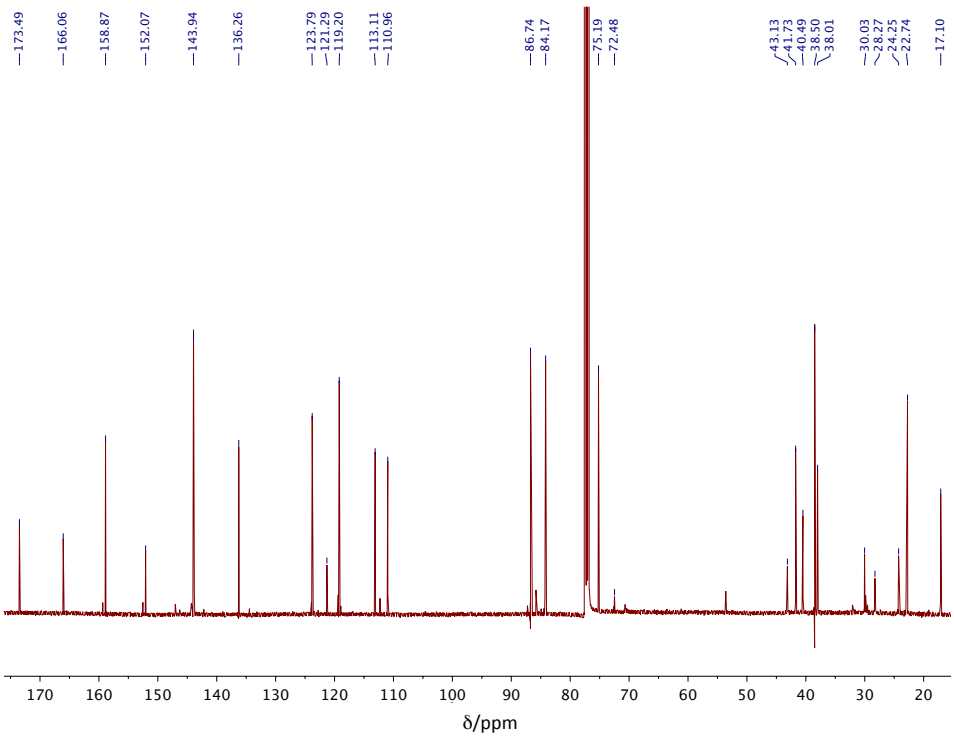


Figure 9.22 ^{13}C NMR of compound **154** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.5** for assignment.

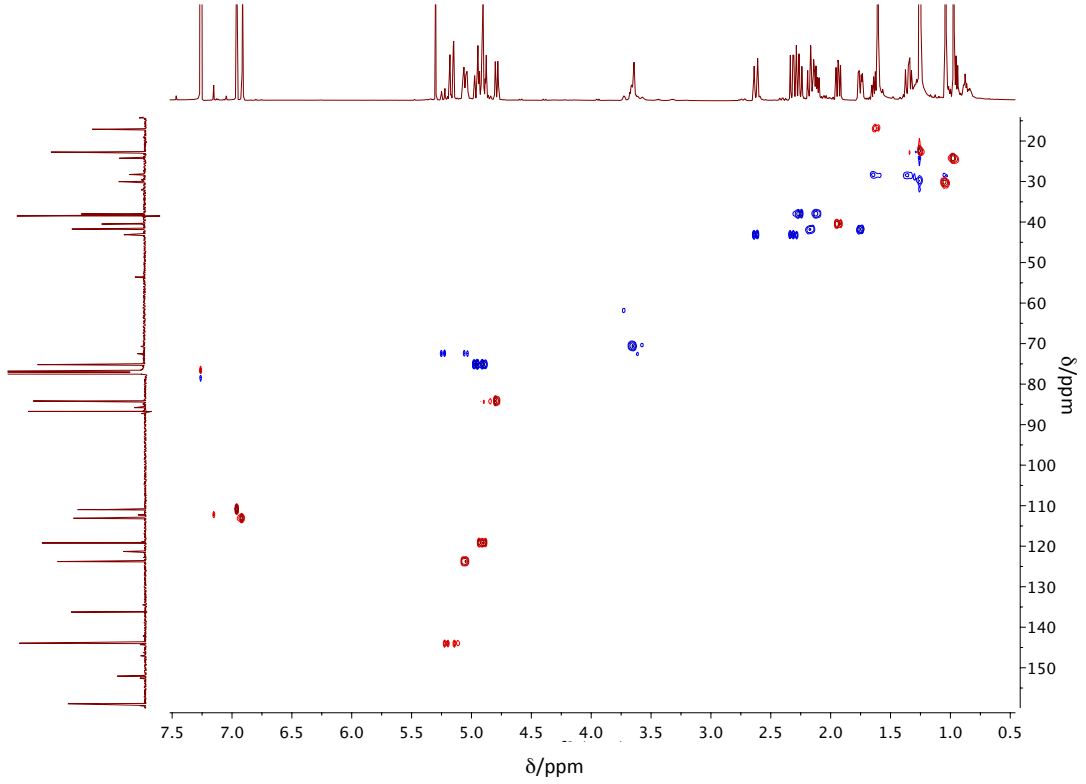


Figure 9.23 HSQC of compound **154**.

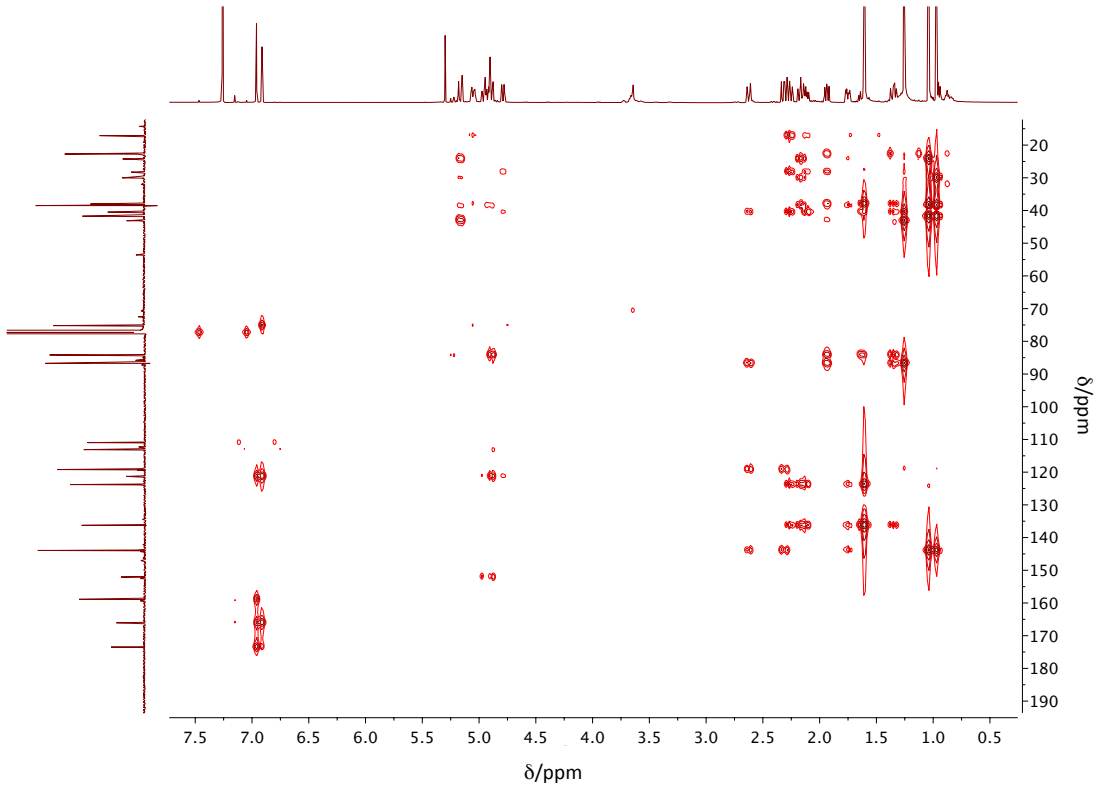


Figure 9.24 HMBC of compound **154**.

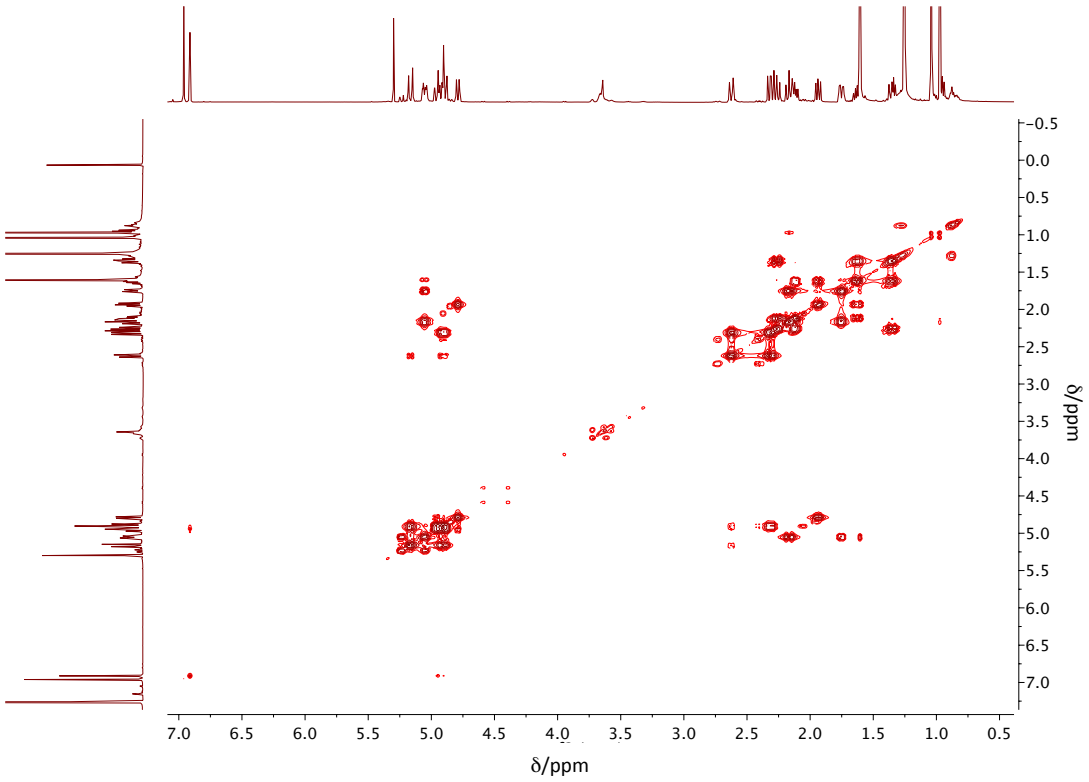


Figure 9.25 $^{1}H,^{1}H$ COSY of compound **154**.

9.1.6 Compound 109a

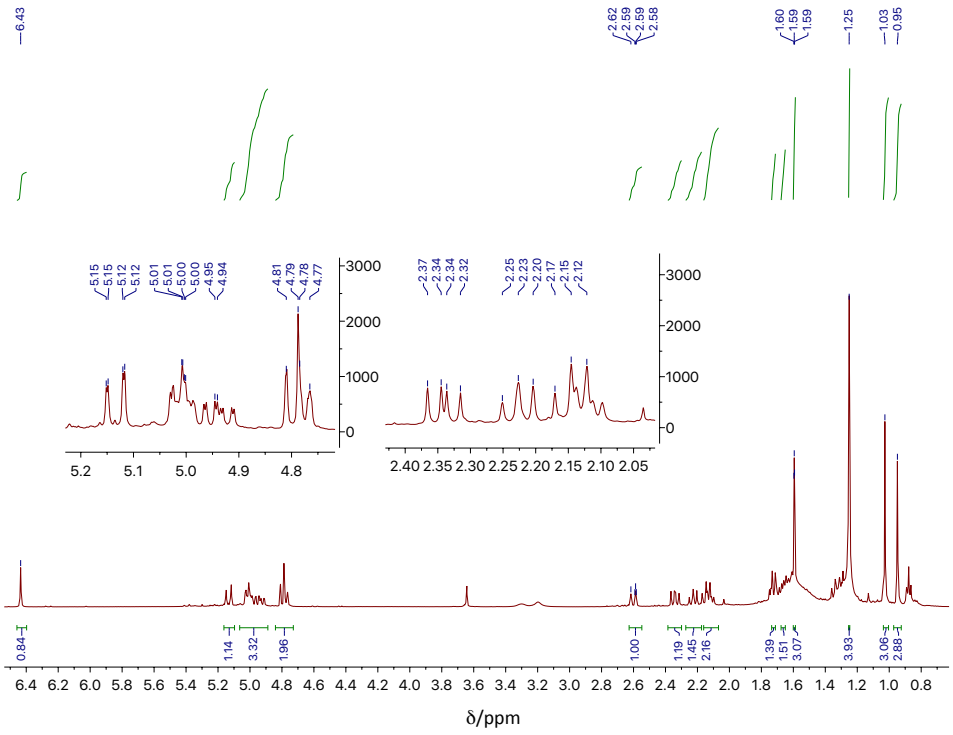


Figure 9.26 ^{1}H NMR of compound **109a** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.6** for assignment.

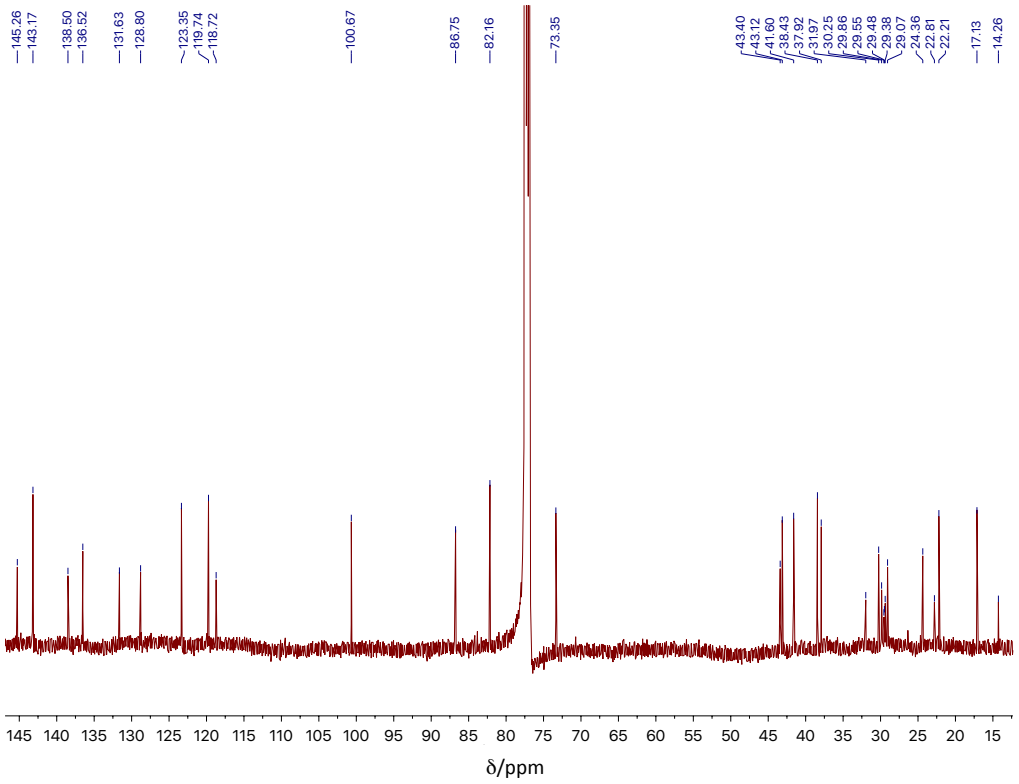


Figure 9.27 ^{13}C NMR of compound **109a** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 2.6** for assignment.

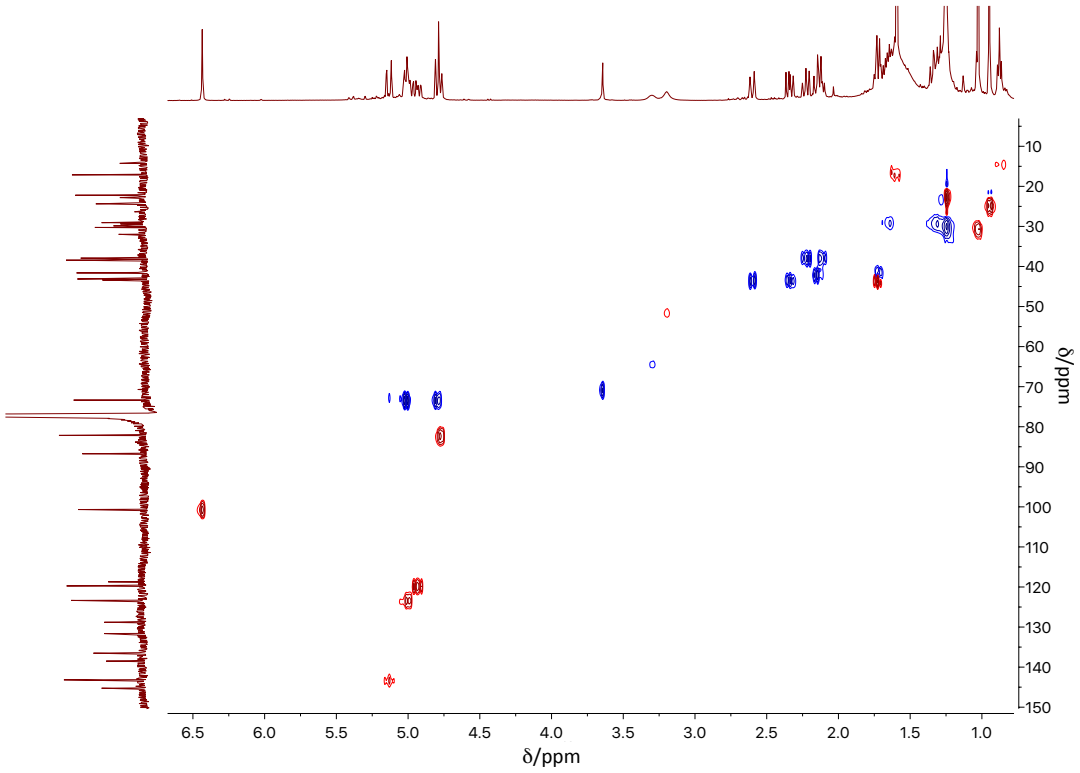


Figure 9.28 HSQC of compound **109a**.

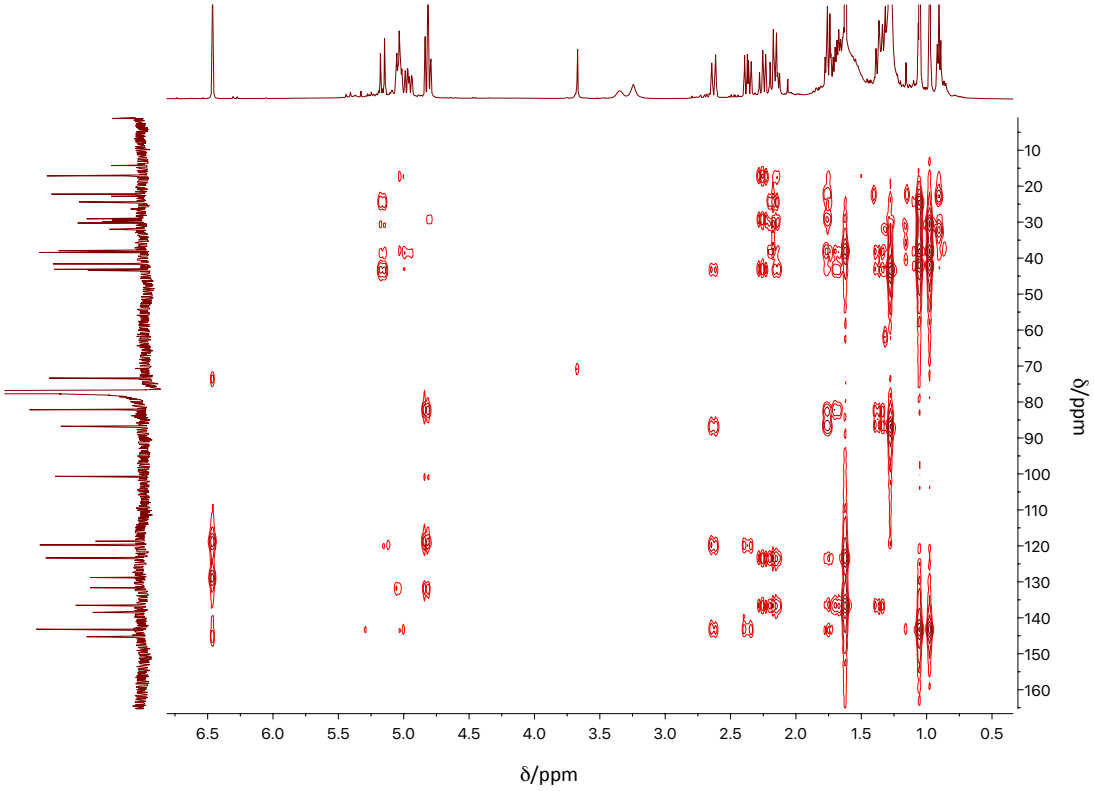


Figure 9.29 HMBC of compound **109a**.

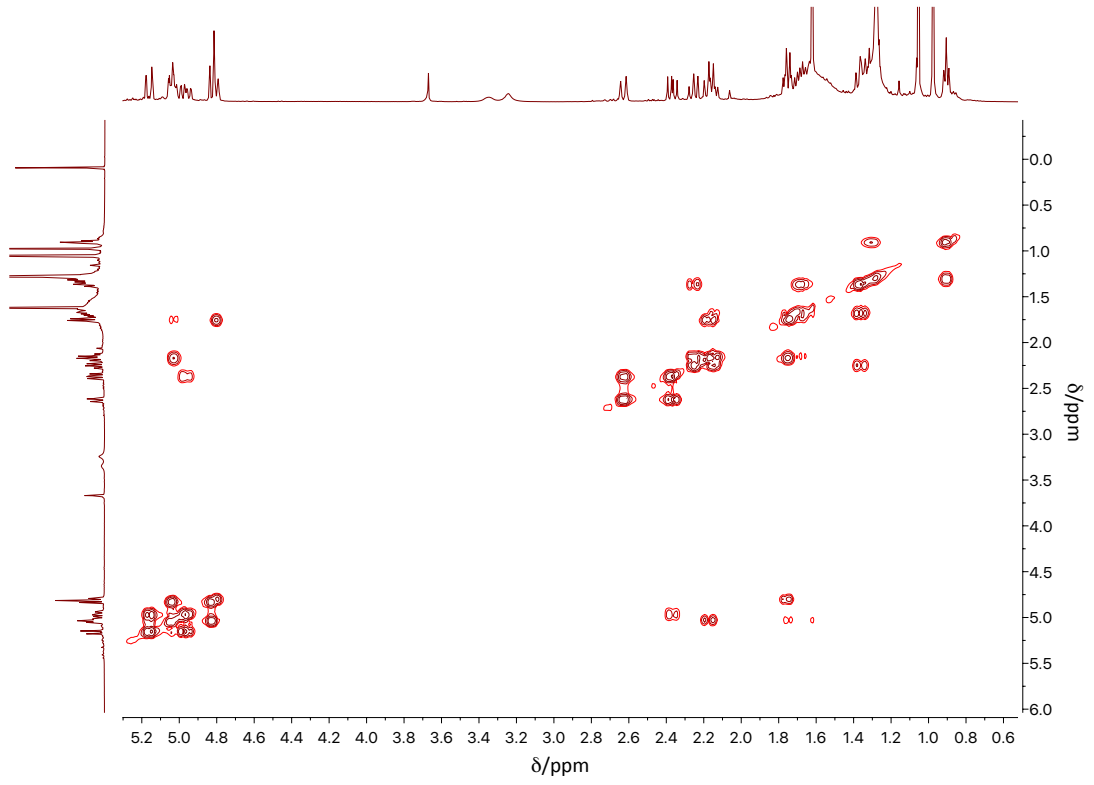


Figure 9.30 1H,1H COSY of compound **109a**.

9.1.7 Compound 156

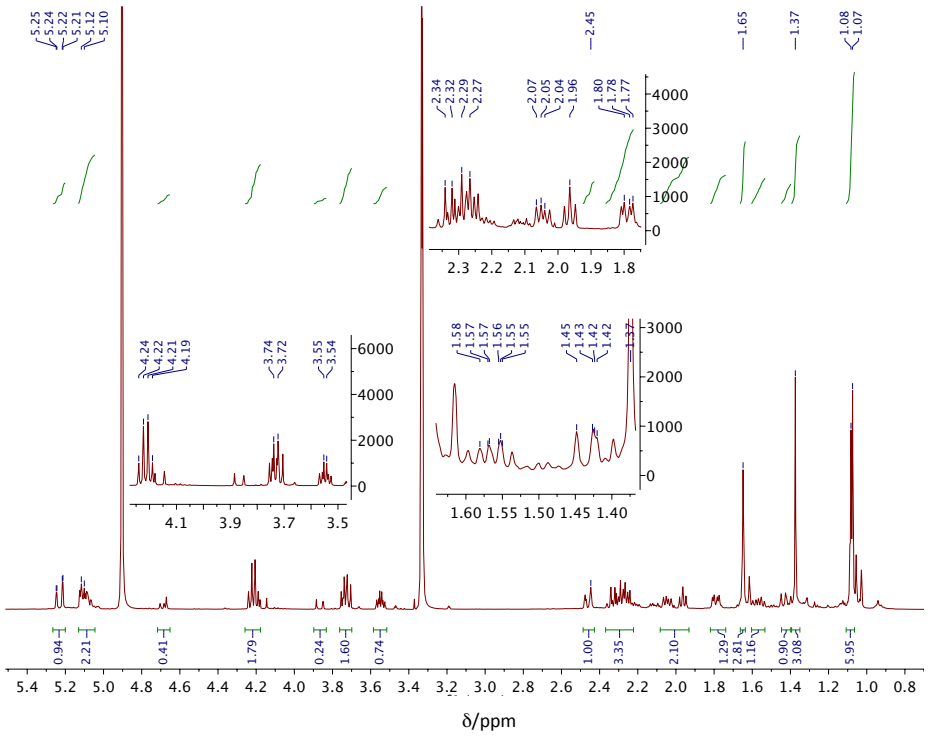


Figure 9.31 ^{1}H NMR of compound **156** in CD_3OD (500 MHz) referenced to CD_3OD. **Table 2.8** for assignment.

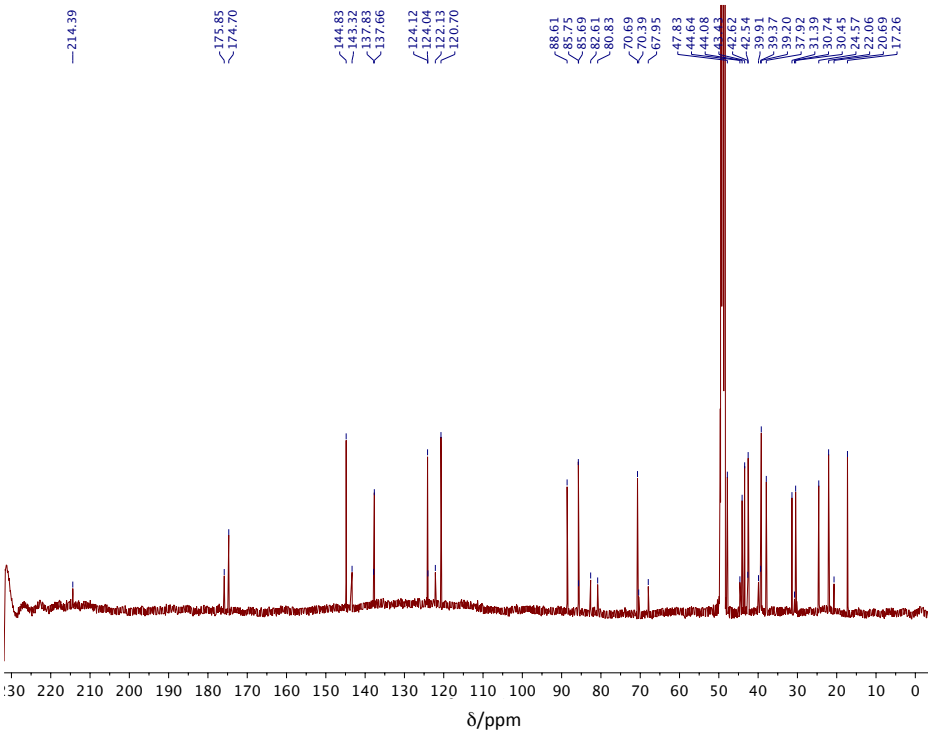


Figure 9.32 ^{13}C NMR of compound **156** in CD_3OD (500 MHz) referenced to CD_3OD. **Table 2.8** for assignment.

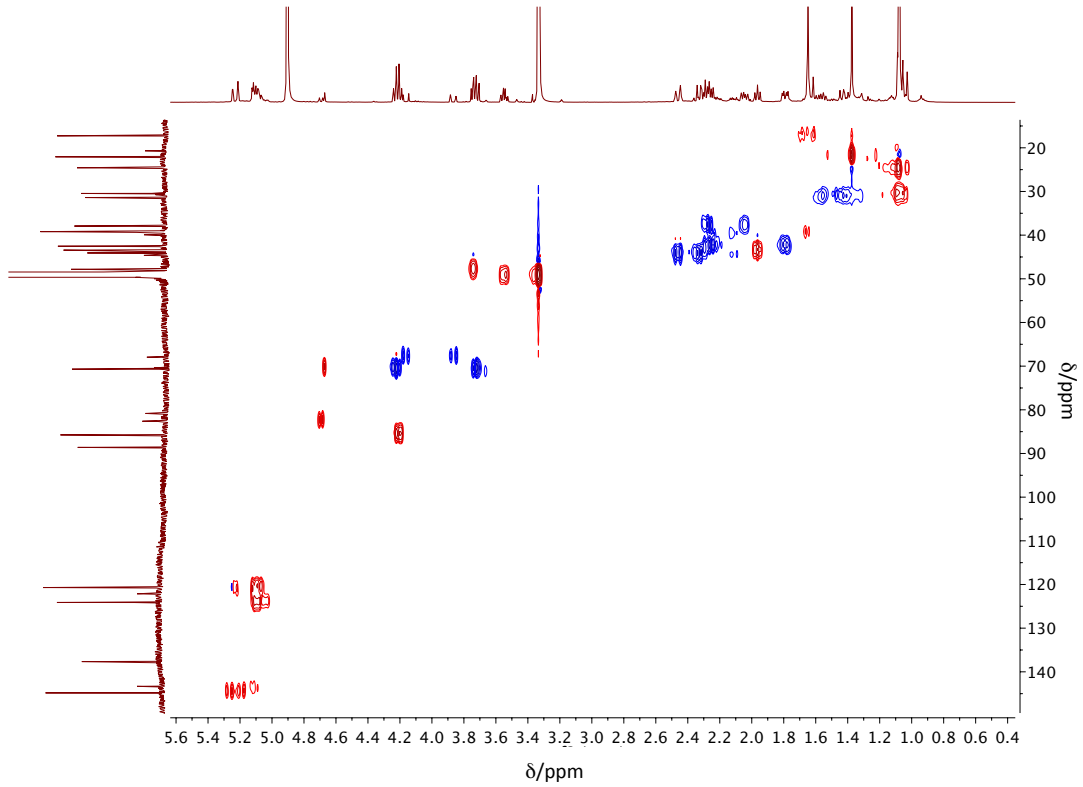


Figure 9.33 HSQC of compound **156**.

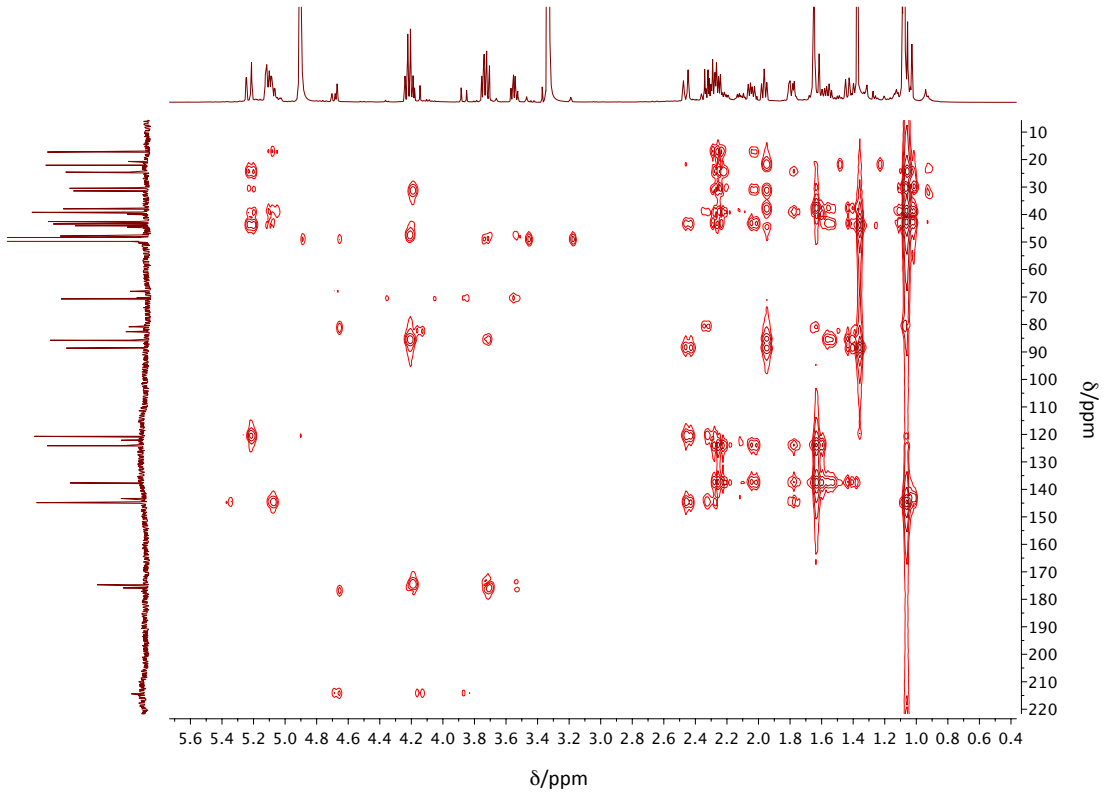


Figure 9.34 HMBC of compound **156**.

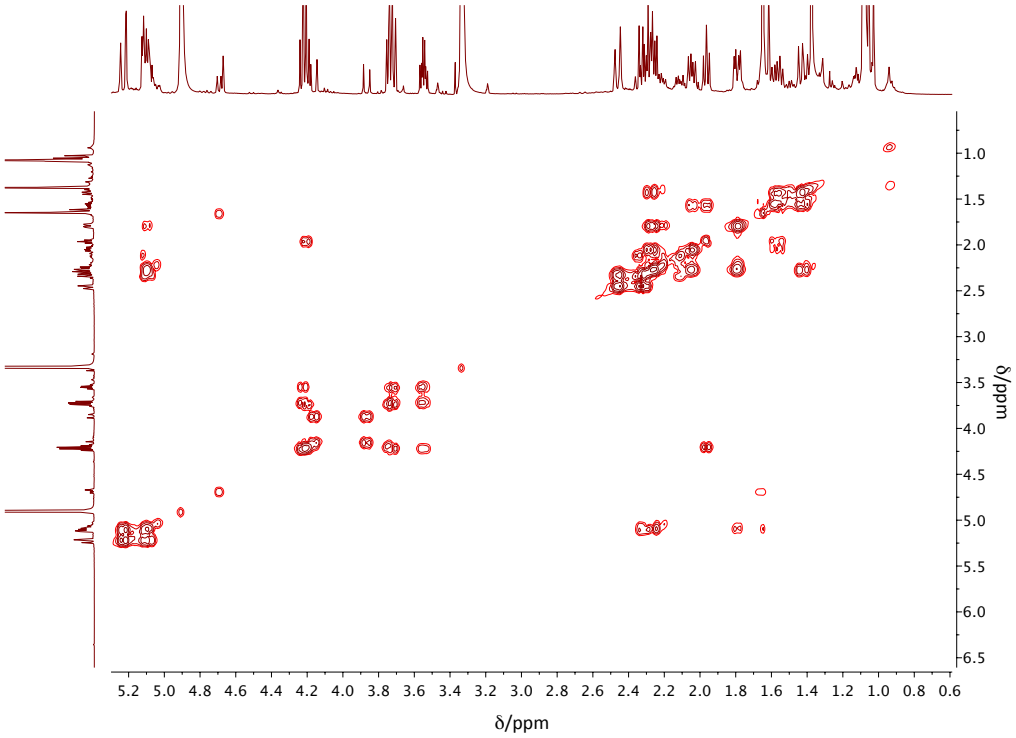


Figure 9.35 $^{1}H,^{1}H$ COSY of compound **156**.

9.1.8 Compound 157 and 158

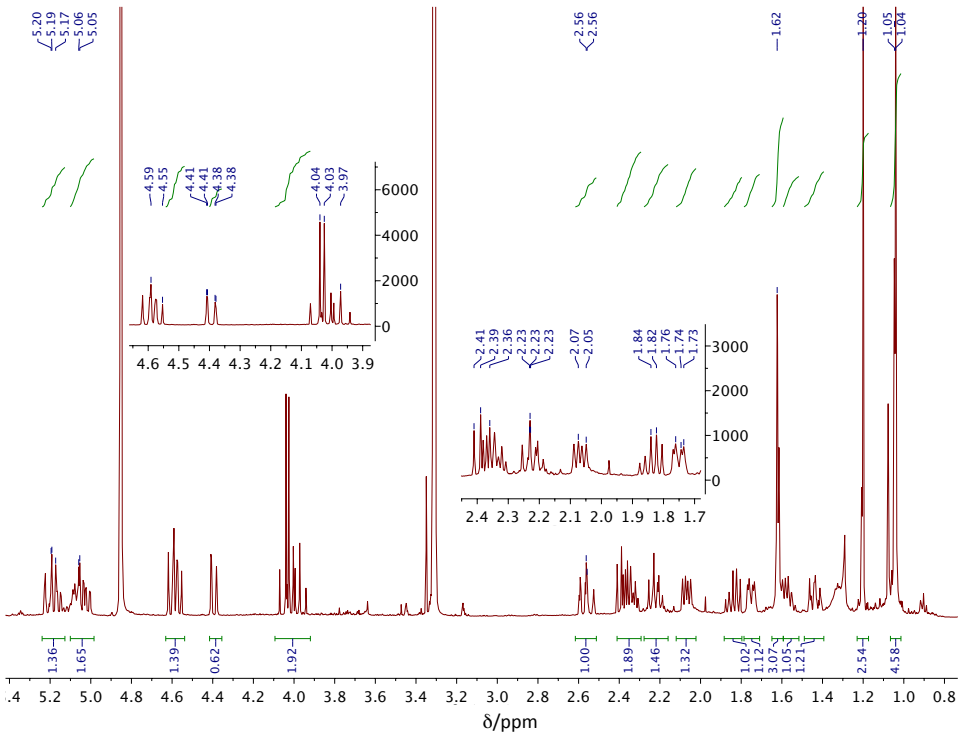


Figure 9.36 ^{1}H NMR of compounds **157** and **158** in CD_3OD (500 MHz) referenced to CD_3OD. **Table 2.9** for assignment.

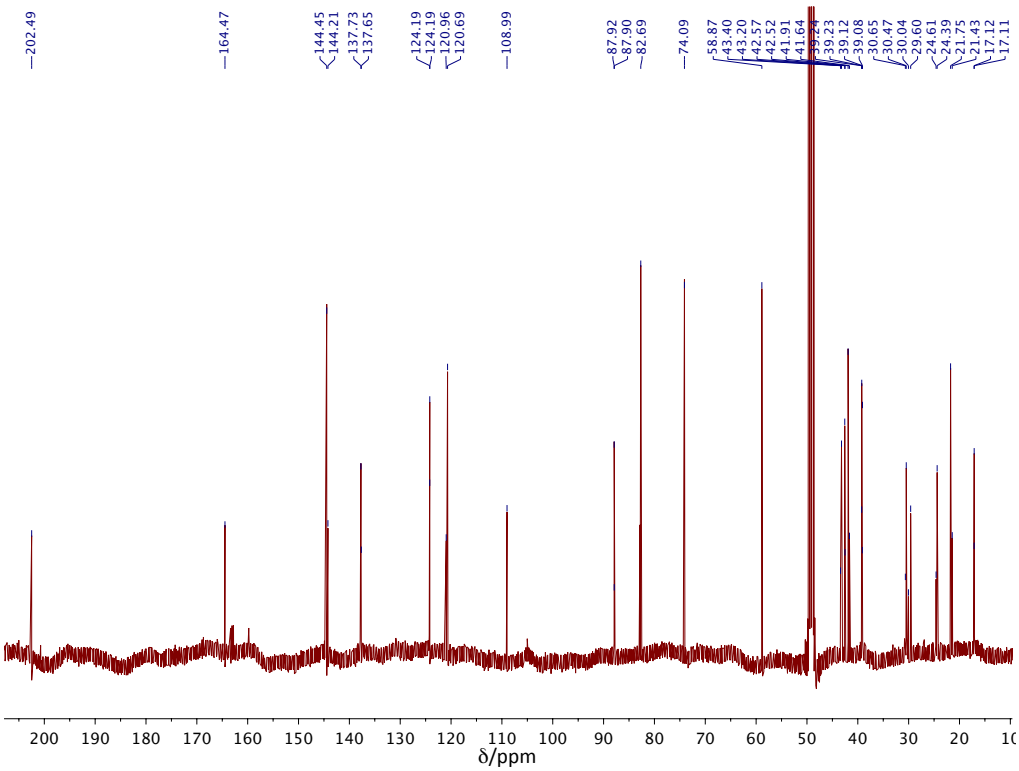


Figure 9.37 ^{13}C NMR of compounds **157** and **158** in CD_3OD (500 MHz) referenced to CD_3OD. **Table 2.9** for assignment.

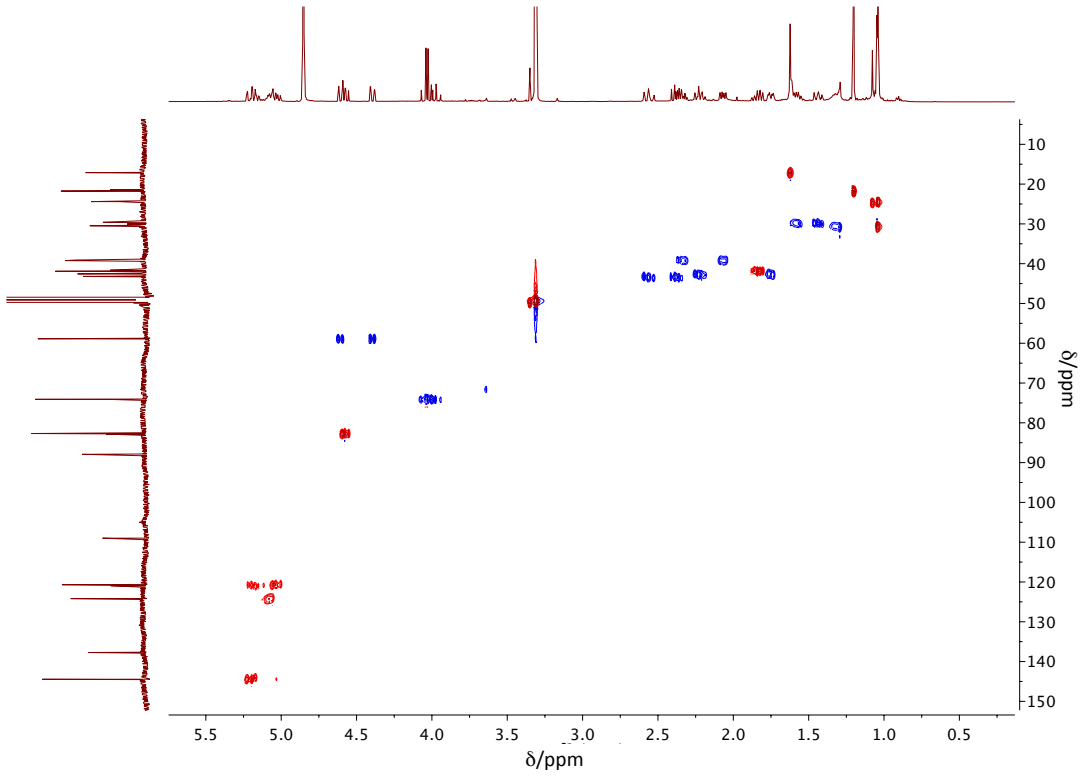


Figure 9.38 HSQC of compounds **157** and **158**.

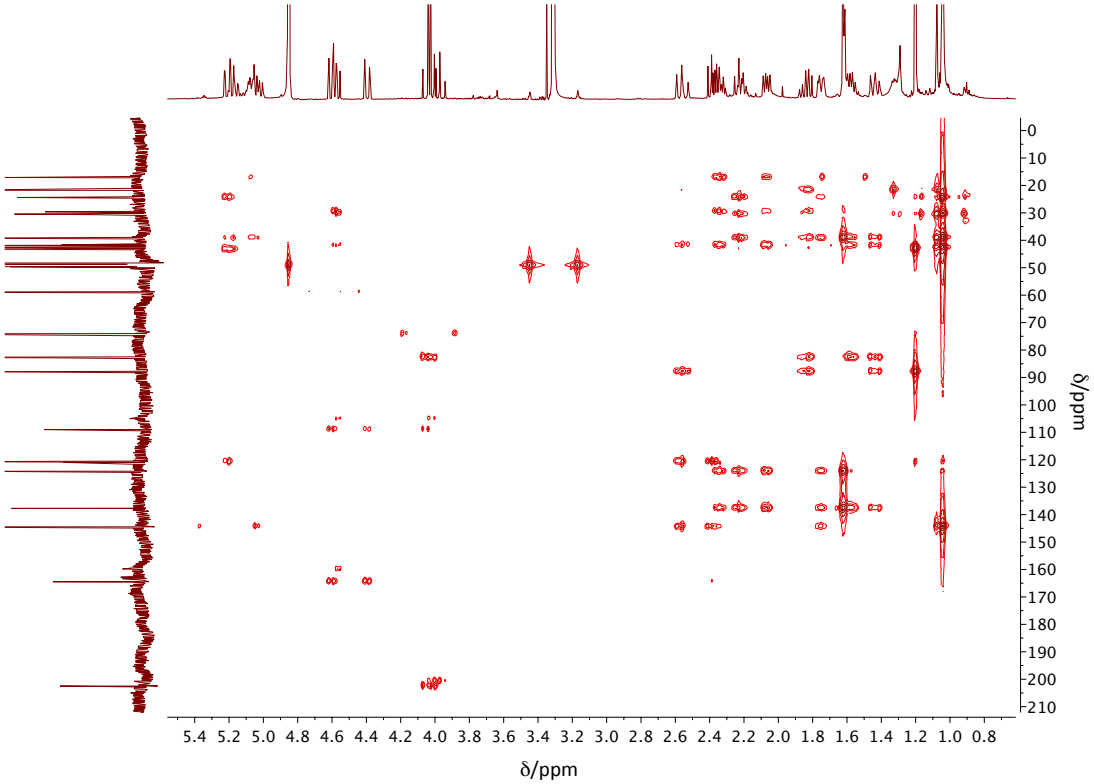


Figure 9.39 HMBC of compounds **157** and **158**.

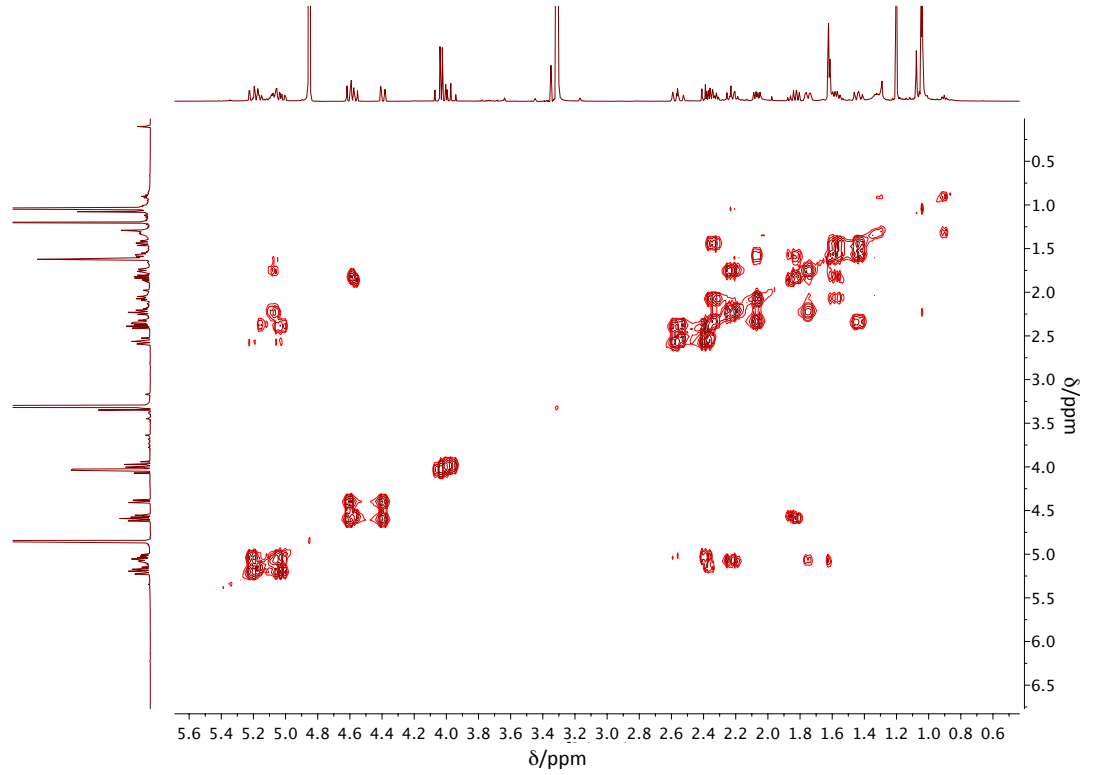


Figure 9.40 ^{1}H,^{1}H COSY of compounds **157** and **158**.

9.2 Chapter 3

9.2.1 *aspks1* BGC

Table 9.1 Exon and intron positions in the *aspsk1* BGC.

#	gene, protein	bp (mRNA), aa	Exons on gDNA
Asg3679	*asL7*, AsL7	4350, 1449	1-626, 697-846, 938-1884, 1944-2841, 2932-4638
Asg3678	*asL6*, AsL6	1272, 423	1-1272
Asg3677	*asL5*, AsL5	750, 249	1-750
Asg3676	*asL4*, AsL4	1293, 430	1-1293
Asg3676	*asL3*, AsL3	1023, 340	1-327, 390-693, 764-1155
Asg3675	*asL2*, AsL2	348, 115	1-348
Asg3674	*asL1*, AsL1	1443, 480	1-205, 305-637, 716-1212, 1288-1695
Asg3673	*aspks1*, MOS	8190, 2729	1-630, 699-6883, 6910-8334
Asg3672	*asR1*, AsR1	1308, 435	1-173, 239-826, 891-986, 1062-1349, 1410-1572
Asg3671	*asR2*, AsR2	1551, 516	1-305, 387-745, 854-1087, 1173-1628, 1719-1807, 1887-1990
Asg3670	*asR3*, AsR3	2565, 854	1-2565
Asg3669	*asR4*, AsR4	1899, 632	1-746, 813-1964
Asg3668	*asR5*, AsR5	1206, 401	1-1206
Asg3667	*asR6*, AsR6	1293, 430	1-1293
Asg3666	*asR7*, AsR7	2469, 822	1-2469

9.2.2 Other BGC

Table 9.2 Depudecin **162** like *A. strictum* BGC with 50% similarities (antiSMASH).

#	Putative function	Homologue	Identity/similarity %	BLASTp and CD
Asg8737	transporter	-	-	Mitochondrial Fe^{II} transporter
Asg8738	hypothetical	-	-	hypothetical
Asg8739	oxidoreductase	-	-	FAD dependent oxidoreductase
Asg8740	oxidoreductase	-	-	alcohol dehydrogenase
Asg8741	transporter	-	-	MFS transporter
Asg8742	oxidoreductase	-	-	NAD dependent
Asg8743	oxidoreductase	-	-	epimerase/dehydratase
Asg8744	HR-PKS	DEP5	49.3/65.3	FAD dependent oxidoreductase
Asg8745	oxidoreductase	DEP4	45.7/64.4	HR-PKS, KS-AT-DH-ER-KR
Asg8746	transporter	DEP3	23.9/32.7	FMN dependent monoxygenase
Asg8747	hypothetical	-	-	MFS transporter
Asg8748	oxidoreductase	-	-	β-barrel
Asg8749	regulation	-	-	shor-chain dehydrogenase
Asg8750	hypothetical	-	-	transcription factor
Asg8751	transporter	-	-	hypothetical

Table 9.3 Leucinostatin **165** like *A. strictum* BGC with 25% similarities (antiSMASH).

#	Putative function	Homologue	Identity/similarity %	BLASTp and CD
Asg2757	PKS	-	-	PKS, AT-DH-*C*-MeT
Asg2758	keto synthase	-	-	keto synthase
Asg2759	hypothetical	-	-	hypothetical
Asg2760	regulation	-	-	transcription factor
Asg2761	hypothetical	-	-	hypothetical
Asg2762	thiolesterase	VFPBJ_02535	25.7/31.9	thiolesterase
Asg2763	hypothetical	-	-	class II aldolase/adducin
Asg2764	transporter	VFPBJ_02529	48.1/60.4	ABC transporter
Asg2765	transporter	VFPBJ_02522	55.5/70.6	ABC transporter
Asg2766	NRPS	-	-	NRPS, A-A-PCP-C-PCP-TD
Asg27647	hypothetical	-	-	hypothetical

Table 9.4 Aspyridone **164** like *A. strictum* BGC with 22% similarities (antiSMASH).

#	Putative function	Homologue	Identity/similarity %	BLASTp and CD
Asg3006	regulation	-	-	transpcription factor
Asg3007	regulation	-	-	transpcription factor
Asg3008	PKS-NRPS	AN8412	43.7/60.5	PKS-NRPS, KS-AT-DH-MeT-KR-ACP-
		Asg3222	33.5/51.0	C-A-PLP-TD
Asg3009	enoyl reductase	AN8409	40.5/43.7	enoyl reductase
Asg3010	regulation	-	-	transcription factor
Asg3011	oxidoreductase	-	-	FAD/NAD(P) binding protein
Asg3012	methyltransferase	-	-	*O*-methyltransferase
Asg3013	transport	-	-	MFS transporter
Asg3014	oxidoreductase	-	-	cytochrome P450
Asg3015	regulation	-	-	transcription factor
Asg3016	hypothetical	-	-	hypothetical
Asg3017	oxidoreductase	AN8411	53.4/67.8	cytochrome P450
Asg3018	oxidoreductase	-	-	cytochrome P450
Asg3019	hypothetical	-	-	hypothetical
Asg3020	hypothetical	-	-	AMP dependent synthetase/ligase
Asg3021	hypothetical	-	-	hypothetical
Asg3022	PKS-NRPS	AN8412	33.1/50.3	PKS-NRPS, KS-AT-DH-MeT-KR-ACP-
		Asg3008	33.5/51.0	C-A-PLP-TD
Asg3024	enoyl reductase	-	-	enoyl reductase
Asg3025	hypothetical	-	-	PEP kinase
Asg3026	methyl transferase	-	-	*O*-methyltransferase
Asg3027	oxidoreductase	-	-	cytochrome P450, FMN, NAD/FAD

Table 9.5 Clapurine **163** like *A. strictum* BGC with 27% similarities (antiSMASH).

#	Putative function	Homologue	Identity/similarity %	BLASTp and CD
Asg7689	hypothetical	-	-	hypothetical
Asg7690	oxidoreductase	CPUR_02681	51.4/66.3	Pyridine nucleotide-disulphide
Asg7691	aminotransferase	CPUR_02679	37.6/47.5	oxidoreductase
Asg7692	prenyltransferase	CPUR_02678	36.0/51.4	PLP dependent transferase
	oxidoreductase	CPUR_02677	29.1/44.5	aromatic prenyltransferase (DMATS)
Asg7693	hypothetical	CPUR_02676	51.4/69.7	cytochrome P450
Asg7694	transporter	-	-	glutathione-S-transferase
Asg7695	hypothetical	-	-	MFS transporter
Asg7696	regulation	-	-	hypothetical
Asg7697	NRPS	CPUR_02680	30.0/40.5	transcription factor
Asg7698	hypothetical	-	-	NRPS, A-PCP-C-A-PCP-A
Asg7699	hypothetical	-	-	hypothetical
Asg7700	hypothetical	-	-	hypothetical
Asg7701	hydrolase	-	-	β-glucosidase
Asg7702	regulation	-	-	fumarylacetoacetate hydrolase
Asg7703	hypothetical	-	-	transcription factor
Asg7704	transport	-	-	peptidase
Asg7705	hypothetical	-	-	MFS transporter

9.3 Chapter 4

Table 9.6 Overview of transformants obtained by bipartite substrate KO targeting *aspks1*.

Transformation	GOI	Resistance cassette	ID of transformed vector	#transformants	#transformants
RSI101/106 2	*asL4*	*hyg*^{R}	RSI95 2	25	11
RSI101/106 4	*asL5*		RSI95 4	16	5
RSI101/106 6	*asL6*		RSI95 6	12	8

Table 9.7 Overview of transformants obtained by bipartite substrate KO targeting *aspks1*.

Transformation	GOI	Resistance cassette	Plasmid ID for PCR	#transformants
RSI32	*aspks1*	*hyg*^{R}	RSI21B5 and RSI21A1	12
RSI47	*aspks1*	*hyg*^{R}		8
RSI58	*aspks1*	*hyg*^{R}		14
RSI66	*aspks1*	*hyg*^{R}		9

9.4 Chapter 5

9.4.1 Additional Constructed Vectors

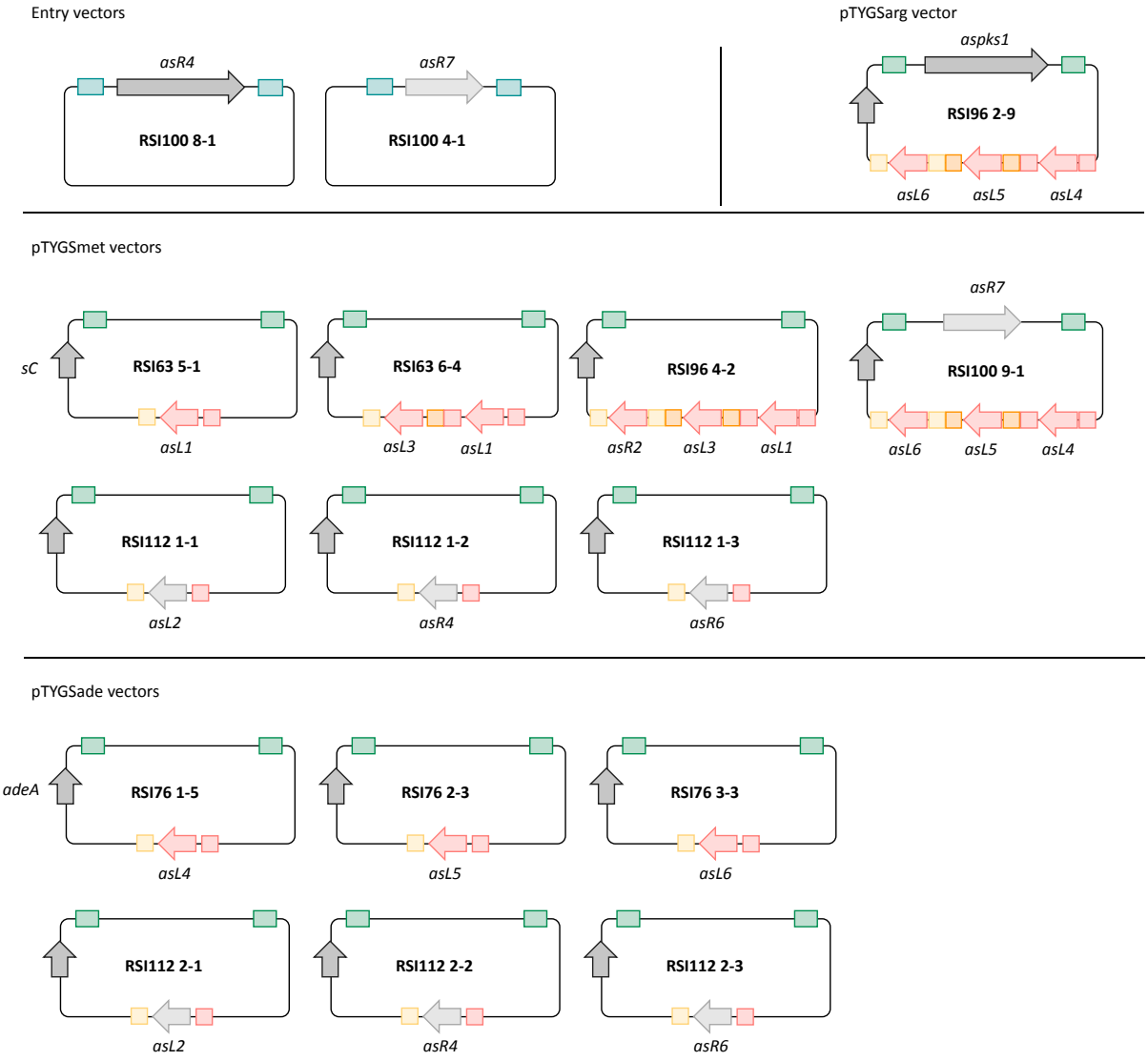


Figure 9.41 Additional vectors constructed in this work.

9.4.2 LCMS Analysis of *A. oryzae* Transformations RSI98 and RSI113

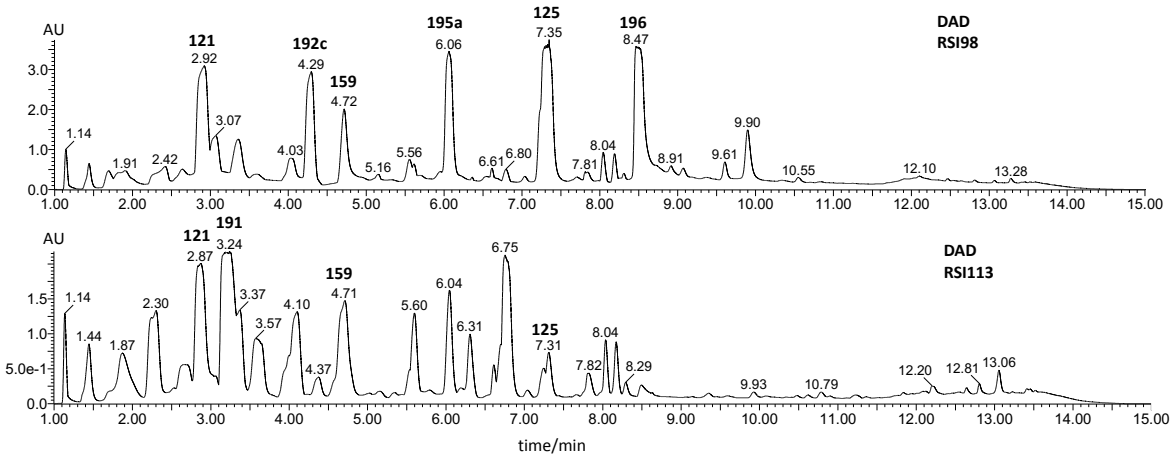


Figure 9.42 DAD chromatogram of *A. oryzae* NSAR1 transformants RSI98 1-11 and RSI113 measured with analytical gradient A2 on LCMS. Extracts were obtained from cultures grown in DPY for 5 d.

9.4.3 NMR of novel Compounds

9.4.3.1 Compound 193

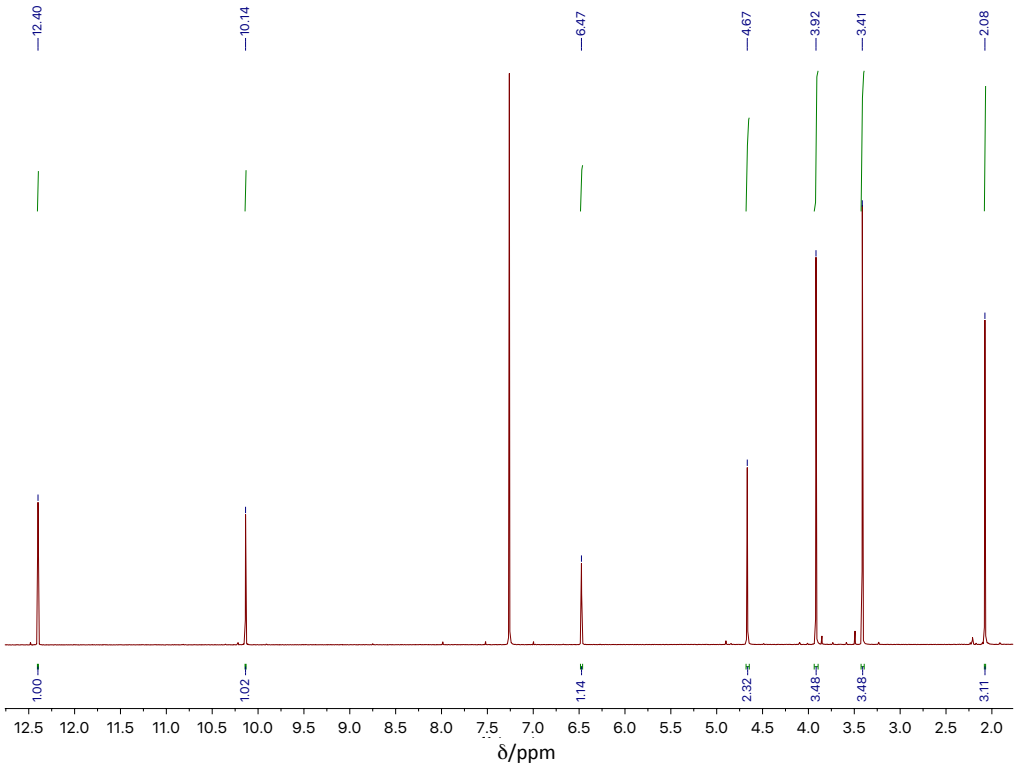


Figure 9.43 ^{1}H NMR of compound **193** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 5.2** for assignment.

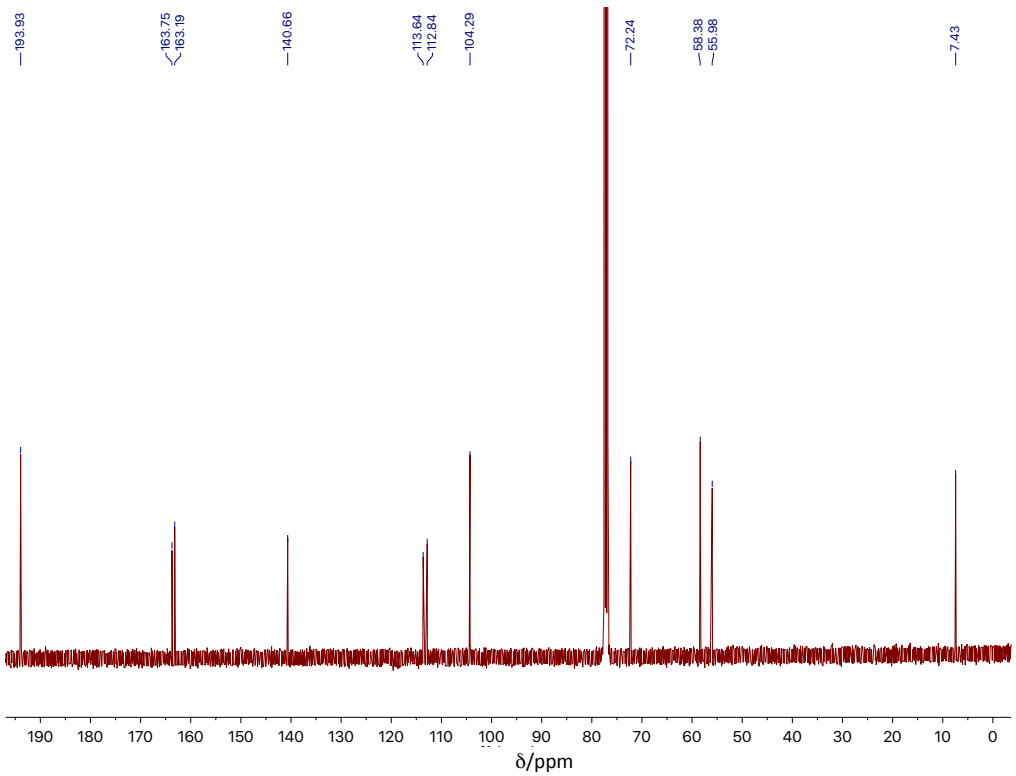


Figure 9.44 ^{13}C NMR of compound **193** in $CDCl_3$ (500 MHz) referenced to $CDCl_3$. **Table 5.2** for assignment.

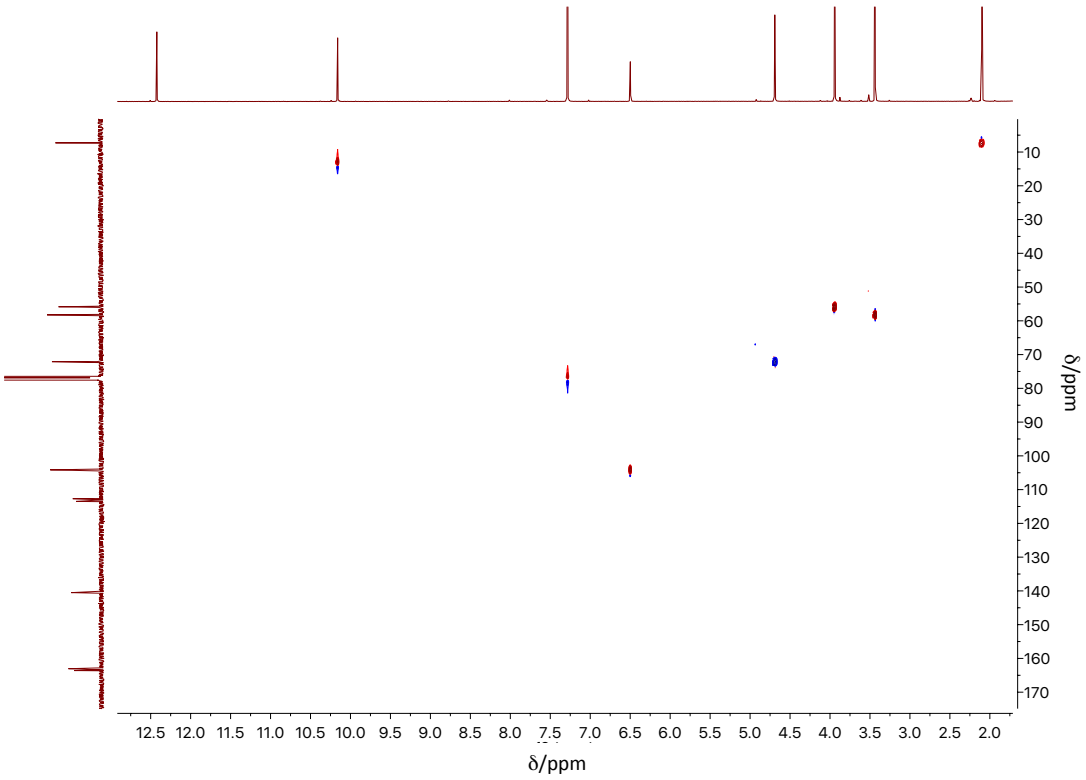


Figure 9.45 HSQC of compound **193**.

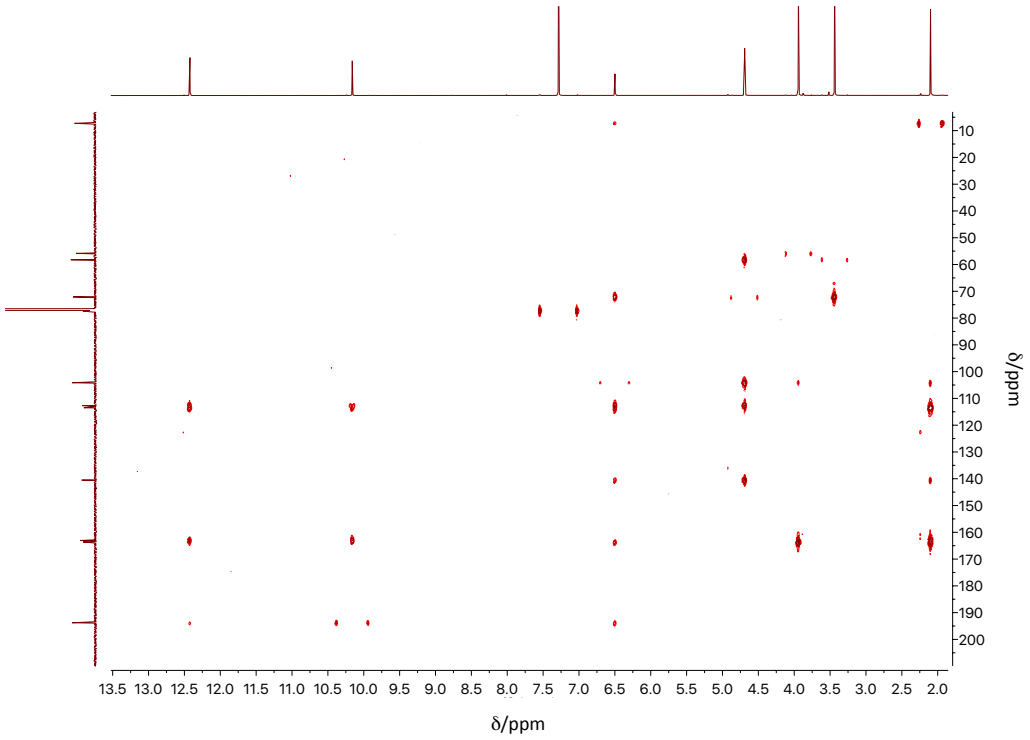


Figure 9.46 HMBC of compound **193**.

9.4.3.2 Compound 195a

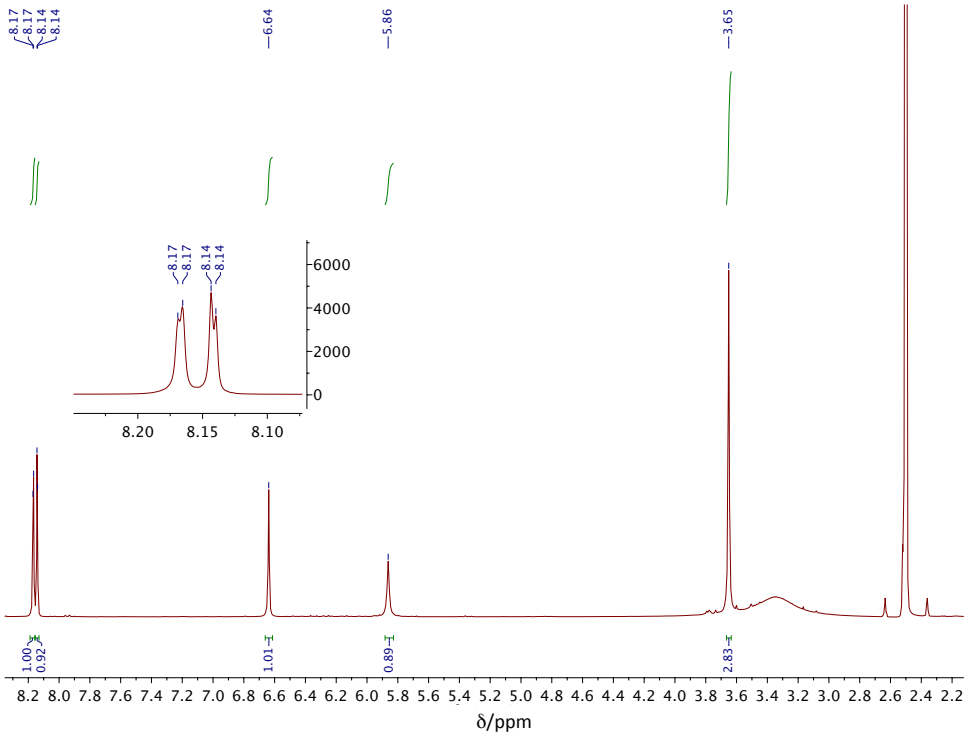


Figure 9.47 ^{1}H NMR of compound **195a** in d_6-DMSO (500 MHz) referenced to d_6-DMSO **Table 5.4** for assignment.

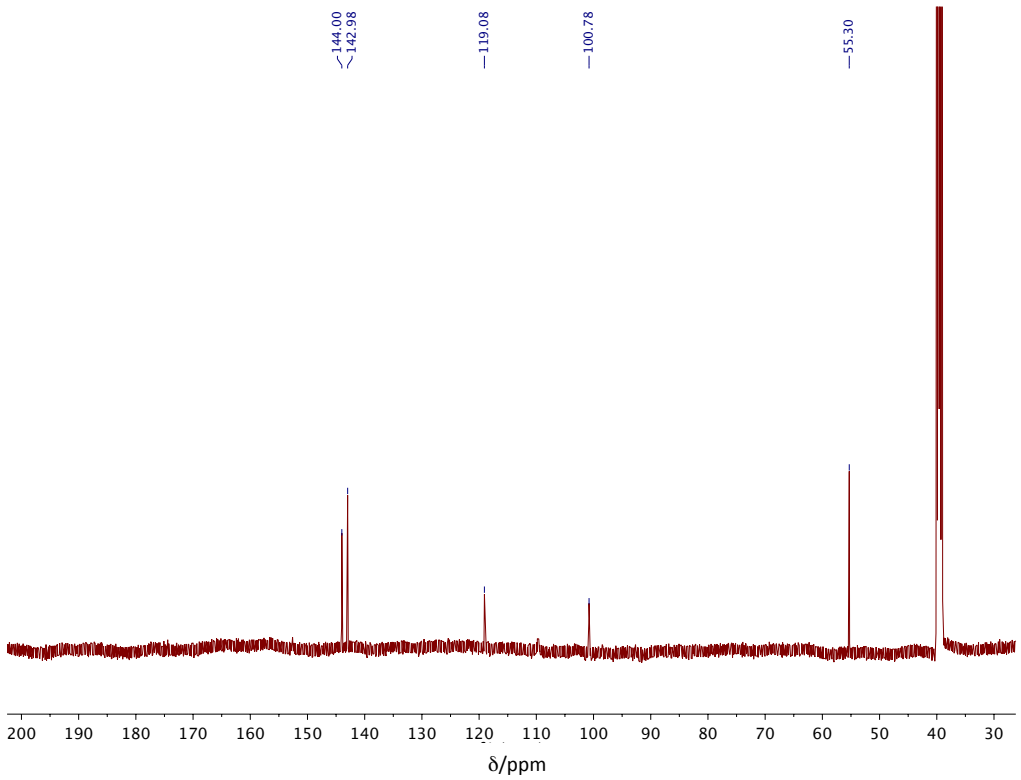


Figure 9.48 ^{13}C NMR of compound **195a** in d_6-DMSO (500 MHz) referenced to d_6-DMSO. **Table 5.4** for assignment.

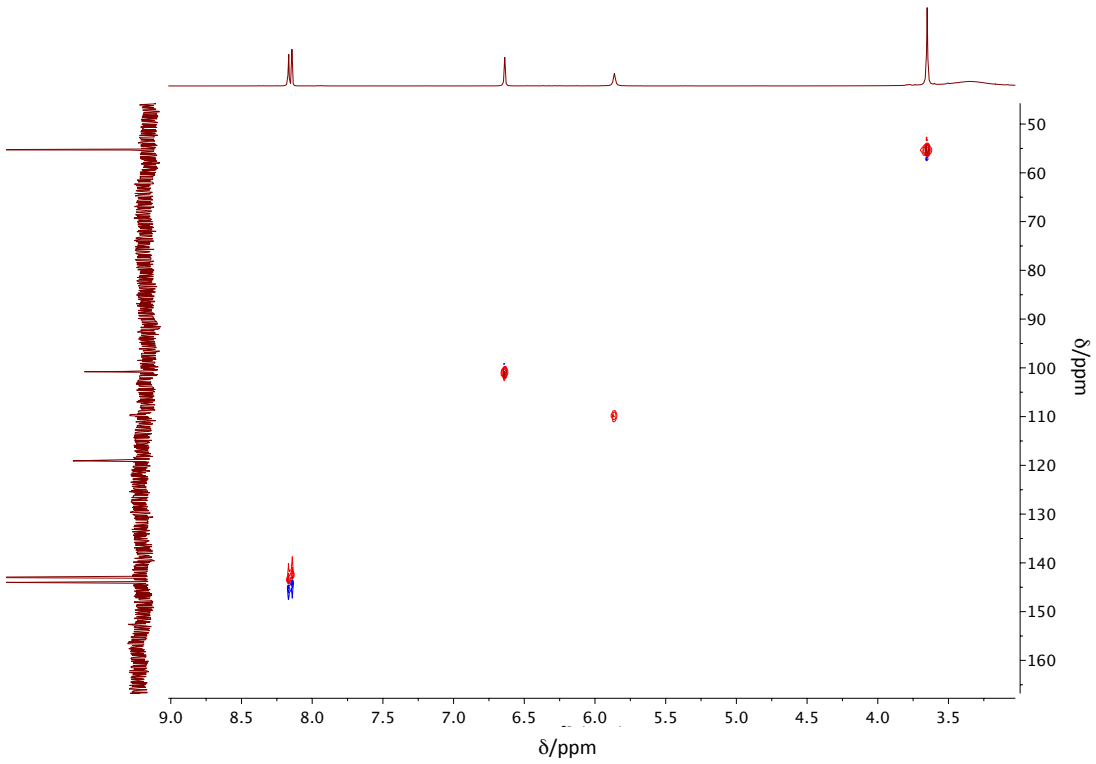


Figure 9.49 HSQC of compound **195a**.

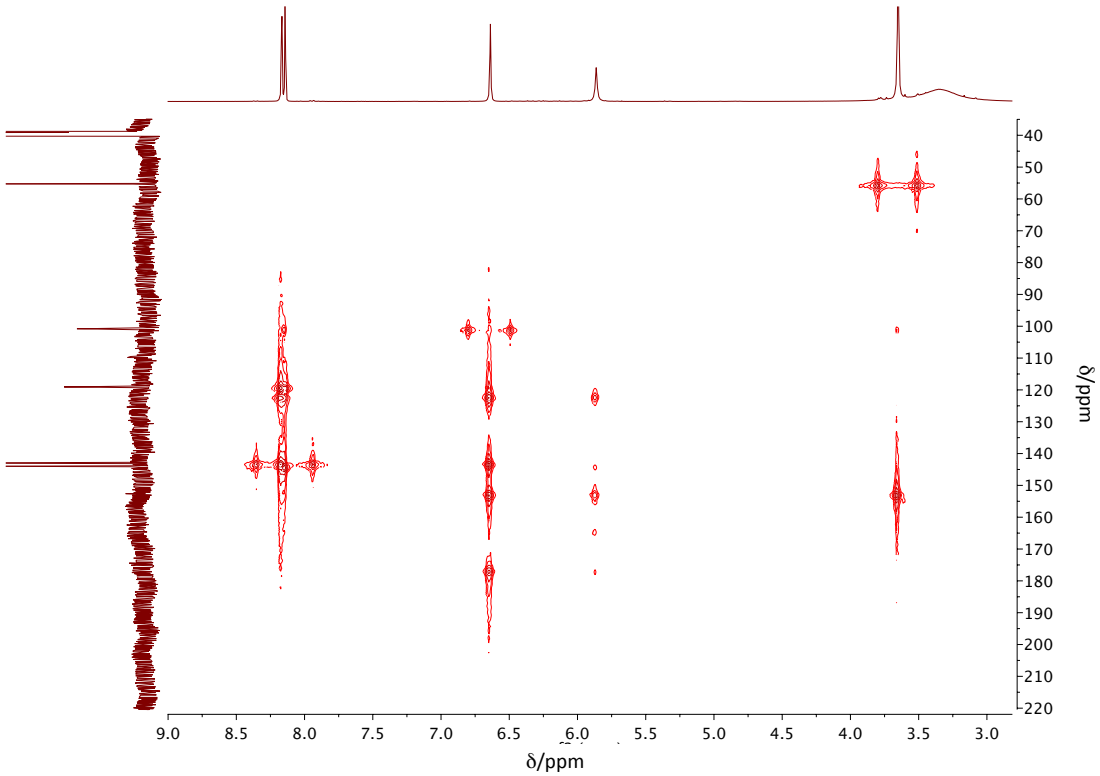


Figure 9.50 HMBC of compound **195a**.

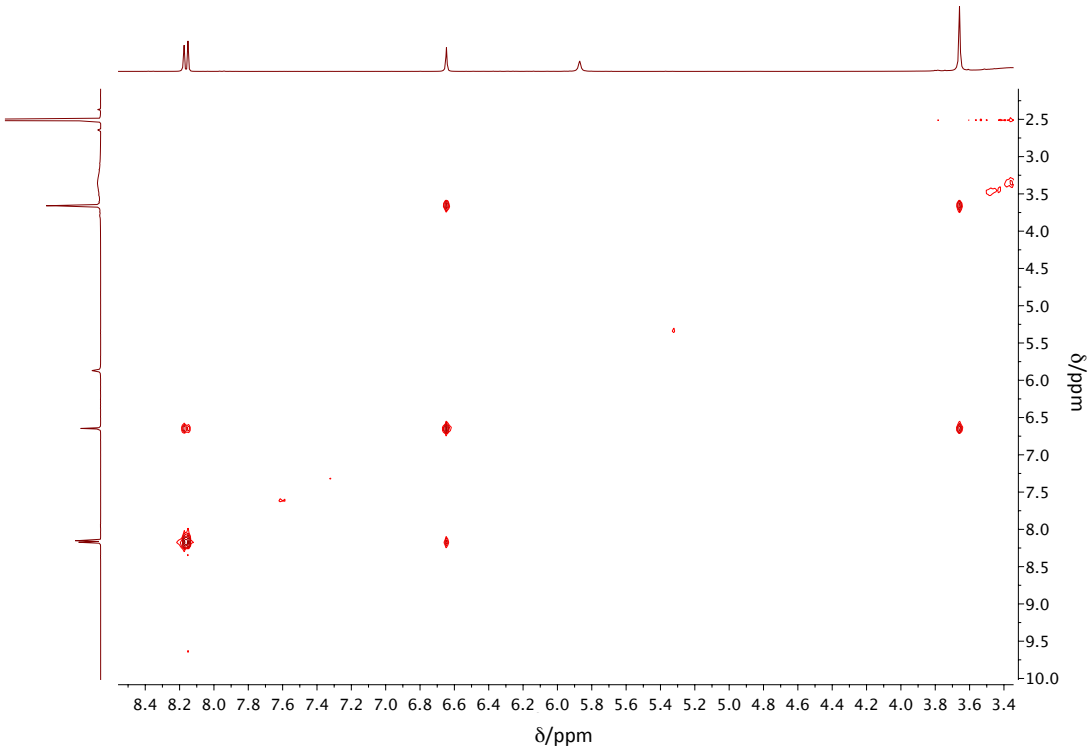


Figure 9.51 ^{1}H,^{1}H COSY of compound **195a**.

9.4.3.3 Compound 197

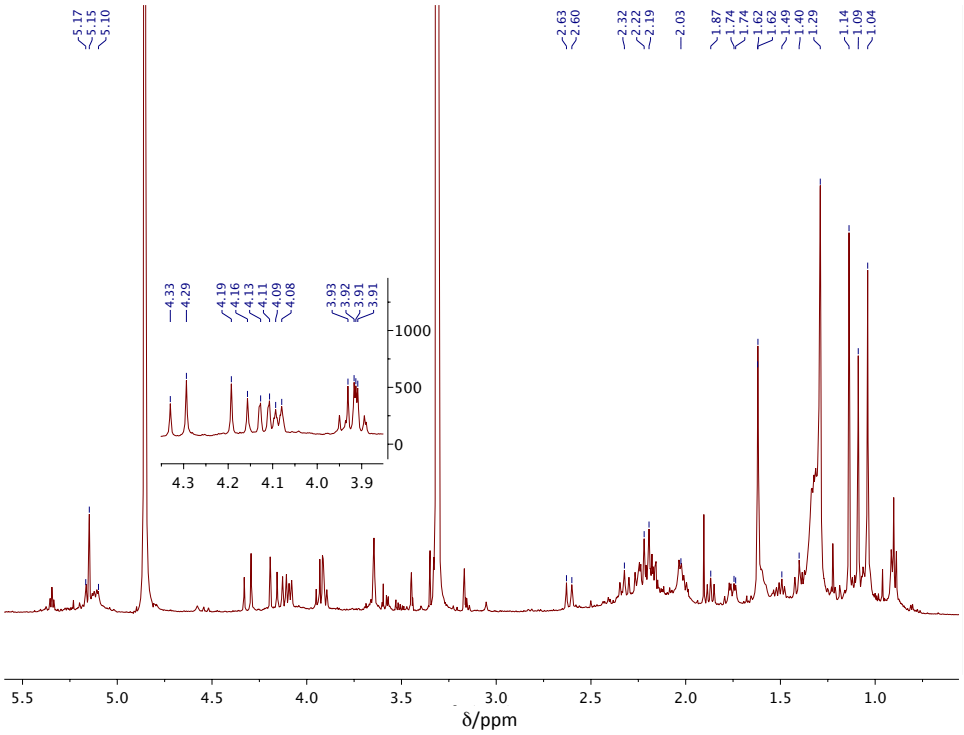


Figure 9.52 ^{1}H NMR of compound **197** in CD_3OD (500 MHz) referenced to CD_3OD. **Table 5.5** for assignment.

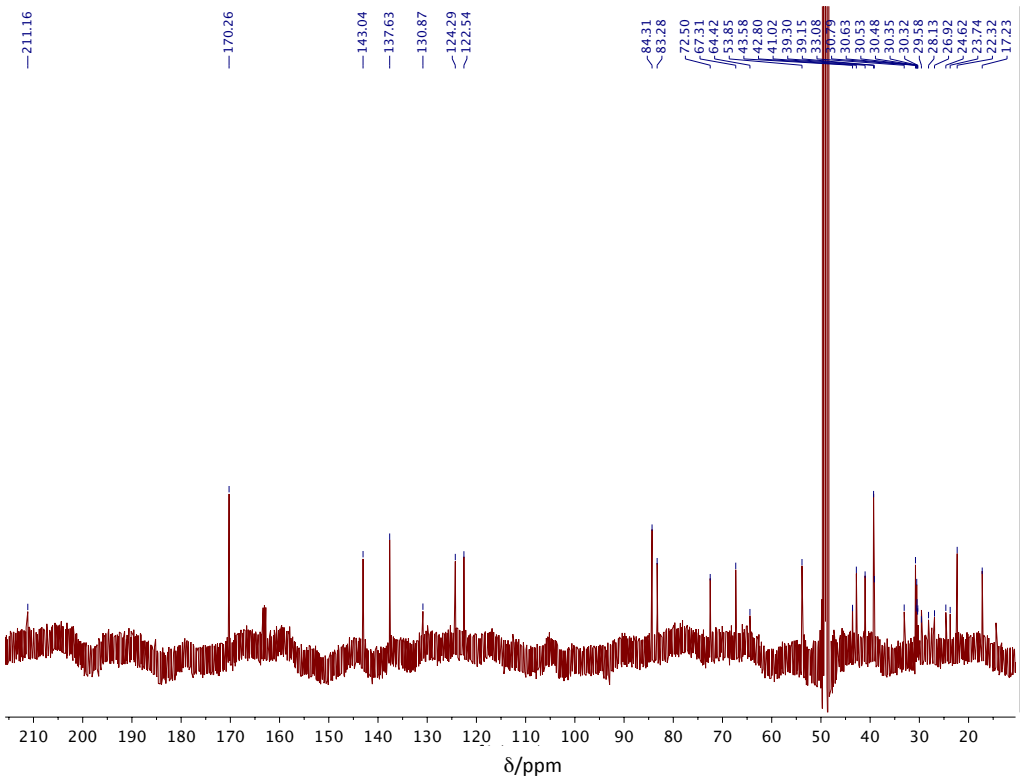


Figure 9.53 ^{13}C NMR of compound **197** in CD_3OD (500 MHz) referenced to CD_3OD. **Table 5.5** for assignment.

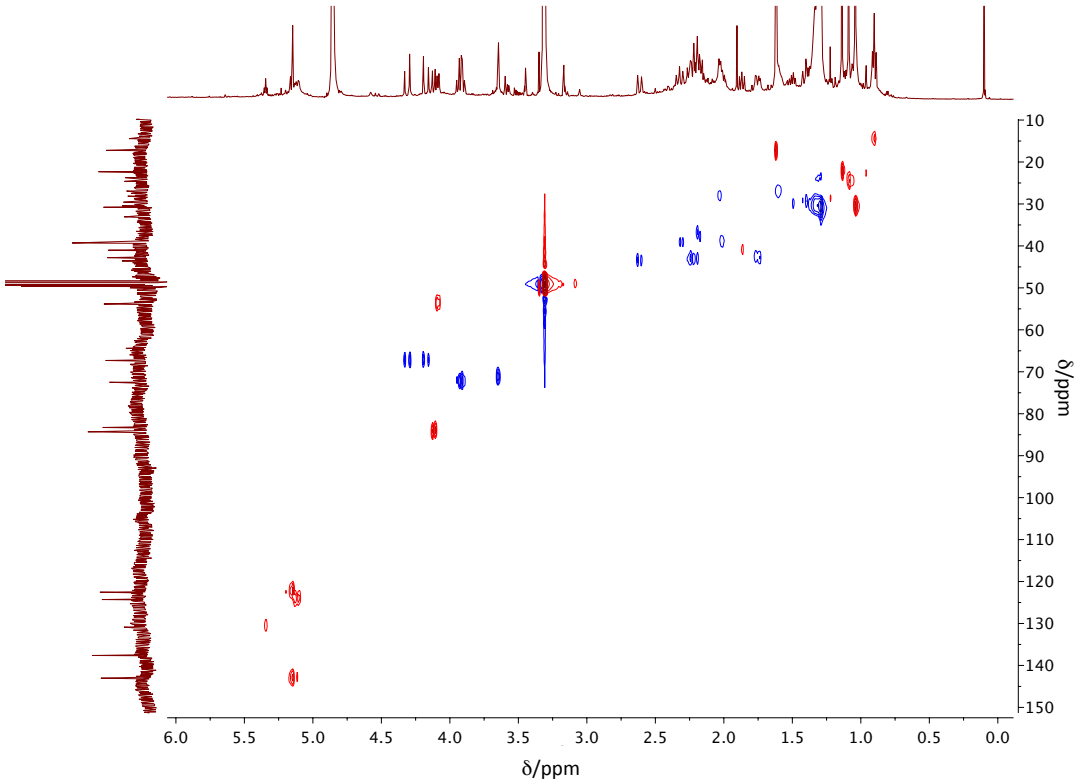


Figure 9.54 HSQC of compound **197**.

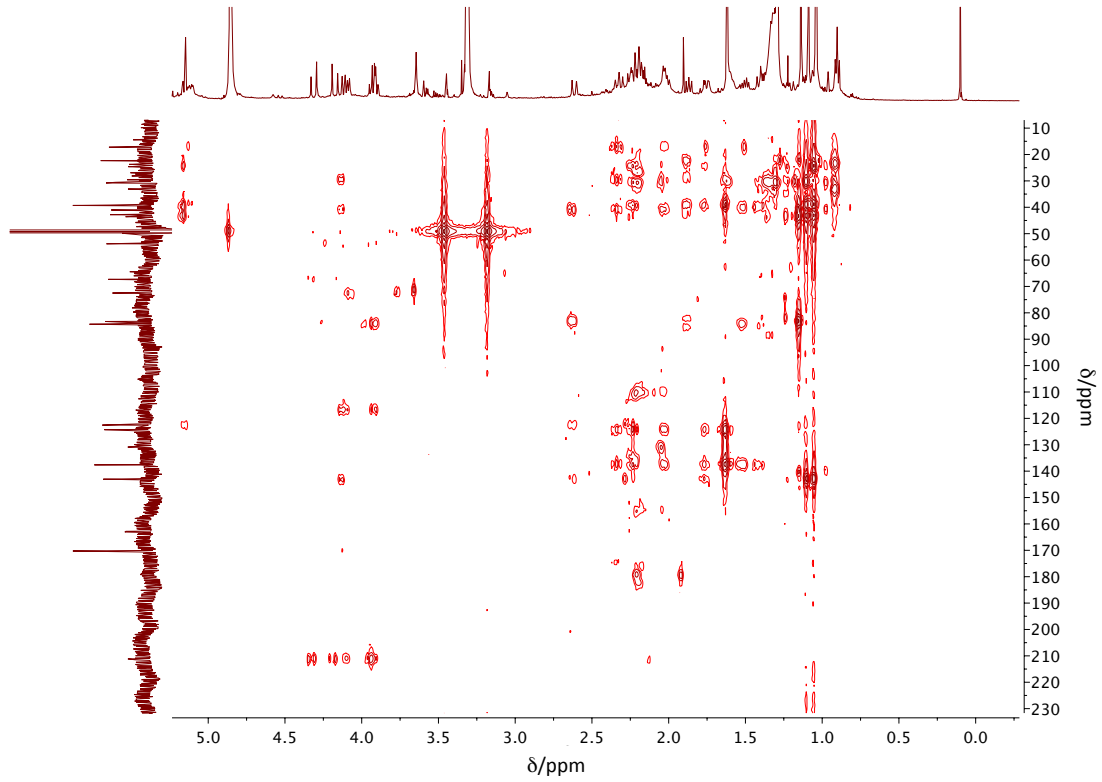


Figure 9.55 HMBC of compound **197**.

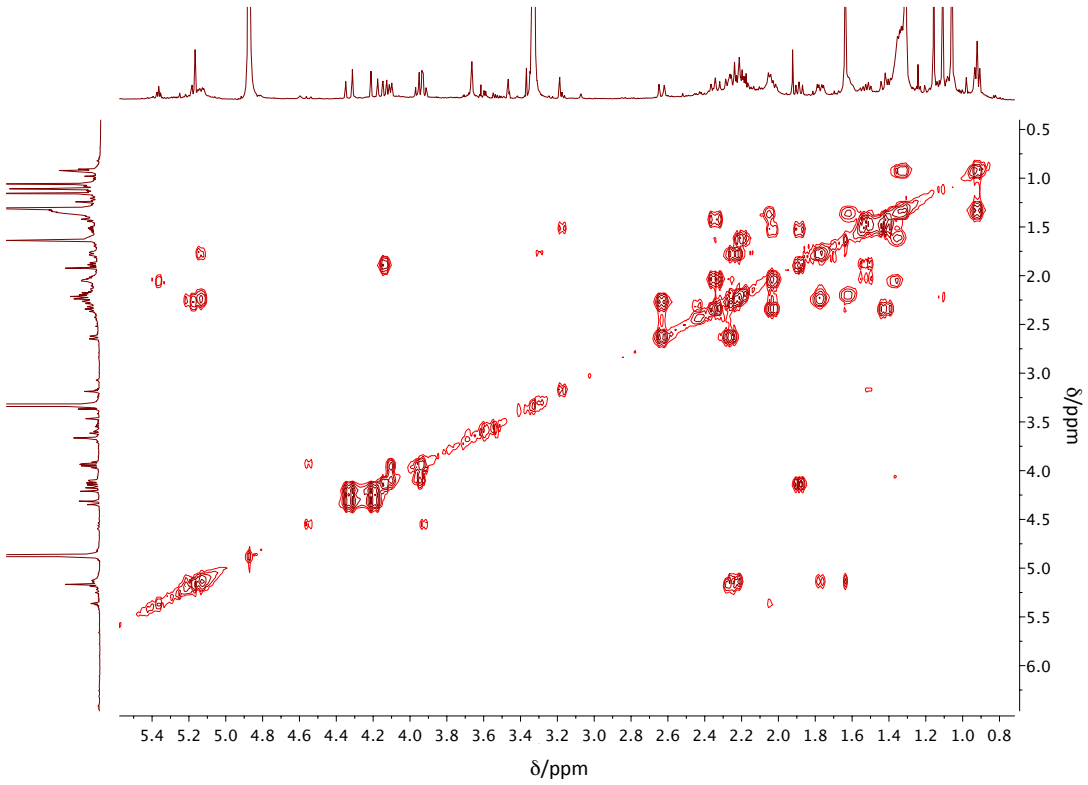


Figure 9.56 $^{1}H,^{1}H$ COSY of compound **197**.

9.4.4 MS^2 Analysis of Xenovulene A 101

MS^2 of *m/z* = 359 $[M]H^+$ of xenovulene A **101** from *A. strictum* WT (collision energy 5-15 eV): 359, 341, 303, 285, 235, 233, 215, 181, 155, 123, 109, 95.

MS^2 of *m/z* = 359 $[M]H^+$ of xenovulene A **101** from *A. oryzae NSAR1* RSI107 B7 (collision energy 5-15 eV): 359, 341, 303, 285, 235, 233, 215, 181, 155, 123, 109, 95.

9.4.5 *E. coli* Codon Optimised Sequences and ESI-Q-TOF-MS

AsR5:5'-ATGCGTCGTAGCTTTCTGATTAGCGCAGCACTGGGTCTGAGCATGAGCACACCGGCACTGGCAGCAAGCATTCAGAGCGTTCTGGGTTATC
TGCGTCCGACCAGCCATCATCATGCACCGTGTGCCGATGATGTTGTTCTGAAACAGAGCGCAGGTAGCGATAGTGCAGCACCGGATCCGCTGCCGAG
CCGTGTTGTTCATAATTGGCCGAATGGCACCTGGATTGAAAACATTAGCGTTCGTCCGAATGGTAATCTGCTGGTTAGCCAGAGTACACCGCGTGGTC
GTGTTTGGCAGGTTAAAGAACCGTGGCTGGATGAACCGAAAGTTGAACTGGCATATGATTTTGATGAATGGGTTGATCGCATTATTGGTATTGGTGA
AACCACACCGGATAAATATGTTGTTGTTGGTAGCCGTTTTTATAGCCTGGATCCGCAGAGCAGCCAGGTTGAACGTACCTTTTGTGCAATGGAACTGG
ATTTTACCAAAGGTGAAAAACCGAGCGCACGTCTGGTTGCACGTTTTCCGCATGCAAATCTGCTGCAGAGCGTTAGCGCACTGCCGTGGGATCGTAGC
GTGGTTCTGATTTCAGATCAGTACCTGCTGCATCCGCGTGCAGATTGGGAAGATCTGACCCCTGGTCCGGGTCAGATTTGGCGTCTGGATACCAAAAC
CGGTCATCATGAAATTGTGATGACCAACTATGCCGAAATGAACACCACCTATAATCATGGTCTGGATGTGGGTATTAACGGCATTAAAATCCATGGCG
ATCATCTGTATTGGATCAATATGGATACCGGTGGTGCGTATCGTGTTCGCATTGATAAATACGGTTATCCGACACCGCTGAATGCAGTTCCGGAAACA
CTGGGTGTTGCCGAAGATGCACTGTGGGATGATTTTGCAATGCATGGCACCCGCATTGGTGAAGAAAGTGATGATACCACCATGTTTGCAACCAGCA
TTGTTAATCTGATGGCAATTAGTCCGGAAAATGGCACCATTGTTCCGCTGGCAGGCGTTGGCACCAGCGAACCGATGGGTTTTCCGGGTCCGACCTCA
GCACAGTTTGGTCGTACCGAAAAAGATAGCCATATTCTGTATGTGACCGGCAAACTGTTTAATGTTCCGCCTAGCATTCGTGATGTTGTGATTCAAGGT
TGGGTTCGTGCAATTGATACCACGGGTTTTCACTTTTAA-3'

AsR6:5'-ATGCCGGTTACCACACCGACCAAAATGGCAACCCTGACCACCAAACAAATGTGGCAGACCATCAAAGATTATTTTGGTGATGGTTTTGTTAC
CGGTAGCGCACCGATTAGCTATAACGTTCATACCTGTGATATGCAACTGCAGCCGGATAGCGGTATTCATGCAGCAAGTGATGGTATTCATTATGGTG
TTCAGATTAGCGAAGATAGCATGCCGCTGTTTAGCATTATGGGTGATACCGCAGCACCGCCTTGTACCTGTCATCGTGTTGATGAAATTGTGAAACAC
ATCGACGAATTTCTGGAACGTGCACCGGAAGCACTGCCGGATGATGGTGCAATTACCAGCGGTAAACCGTGTGATACCAATCCGGATCAGGTTAGCC
TGTATGCAATGCGTGATAGCCTGAGTTGGTGGGTTCATTGGGGTGGTAATCTGCGTCCGGAACATTATTGGAAGCAGATTTATATCGGTTTTGCAGCC
ATTCCTGATGATGTGCAGATTAGTCCGCGTGAATTCCTGGATGGCACCTATCGTTATCTGGGTCATACCTGGGATGATTGTCTGAGCGGTCTGGAAGA
AGAGGGCGTTAGTCCGGATGAAATCGAATTTGCCAATATGTGTATGTGGCGTCAGATGCTGACCCAGTGGCTGGAAAAAGCAGATCCTGAACTGCTG
CCGCTGCTGAAAGGTAAAATTAGCCTGATGCTGCAGTATCGTGTTCTGACCGCAAATACCCTGGGTTGTCTGGCACTGTTTATGAATGCAACCGCAGA

TCCGAAAGATGGTCCGATCCATTATGCAGATAGCAGCTATGAAATGGAAATTGCAAGCGTTGCACAGTGTGTTACCCTGGATATGGCAAAAGAAGCA
ATGGGTATTCTGCAGGGTGAACGTACCGAAGTTGTTGCCGGTGATCGTGCACAGCGTAAACGTGAACTGCGTTGGATTTATGTTCGTTGTATGCAGAT
TCTGGAAAGCCAGCCGCATGCACATATGCTGCGTCGTTATGGTAGCGCAGGTCTGCATTATGTTCCGATGATGGATCGCTATCTGGAACGCGTTAGCG
GTCATACCCGTTTTCCGATTCGTGATGGTGCCGCACGTATCCTGGAACGTTTTATCAATCGTGCAGAACTGCCGAAAGAAAGCGAGGATATTAATCCG
AATGGTCGTAGCCTGAAAGTTAGCGCAAAAATGAATGGTAATGGTCAGCTGCATCATGAAGTTAATGGCAATGCAAAACTGCATCTGGAAGCCGAAC
GTCCGGATGTTACCACCGCAGTTGGTTAA-3’

AsR5, ESI-qTOF-MS, Coverage 70.32%, identified peptides:

MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDHPFTMRRSFLISAALGLSMSTPALAASIQSVLGYLRPTSHHHAPCADDVVLKQSAGSDSAAPDPLP
SRVVHNWPNGTWIENISVRPNGNLLVSQSTPRGRVWQVKEPWLDEPKVELAYDFDEWVDRIIGIGETTPDKYVVVGSRFYSLDPQSSQVERTFCAMELDF
TKGEKPSARLVARFPHANLLQSVSALPWDRSVVLISDQYLLHPRADWEDLTPGPGQIWRLDTKTGHHEIVMTNYAEMNTTYNHGLDVGINGIKIHGDHLY
WINMDTGGAYRVRIDKYGYPTPLNAVPETLGVAEDALWDDFAMHGTRIGEESDDTTMFATSIVNLMAISPENGTIVPLAGVGTSEPMGFPGPTSAQFGRT
EKDSHILYVTGKLFNVPPSIRDVVIQGWVRAIDTTGFHF

AsR6, ESI-qTOF-MS, Coverage 65.45%, identified peptides:

MRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDHPFTMPVTTPTKMATLTTKQMWQTIKDYFGDGFVTGSAPISYNVHTCDMQLQPDSGIHAASD
GIHYGVQISEDSMPLFSIMGDTAAPPCTCHRVDEIVKHIDEFLERAPEALPDDGAITSGKPCDTNPDQVSLYAMRDSLSWWVHWGGNLRPEHYWKQIYIGF
AAIPDDVQISPREFLDGTYRYLGHTWDDCLSGLEEEGVSPDEIEFANMCMWRQMLTQWLEKADPELLPLLKGKISLMLQYRVLTANTLGCLALFMNATAD
PKDGPIHYADSSYEMEIASVAQCVTLDMAKEAMGILQGERTEVVAGDRAQRKRELRWIYVRCMQILESQPHAHMLRRYGSAGLHYVPMMDRYLERVSG
HTRFPIRDGAARILERFINRAELPKESEDINPNGRSLKVSAKMNGNGQLHHEVNGNAKLHLEAERPDVTTAVG

9.4.6 MS^2 Analysis of Phenolic Meroterpenoids 109a and 109b

t_R = 9.2 min: MS^2 of *m/z* = 369 $[M-H]^-$ of **109a** from *A. strictum* WT (collision energy 10-20 eV): 369, 341, 165, 137; MS^2 of *m/z* = 369 $[M-H]^-$ of **109a** from *A. oryzae* NSAR1 RSI111 2-2 (collision energy 10-20 eV): 369, 341, 165, 137.

t_R = 9.4 min: MS^2 of *m/z* = 369 $[M-H]^-$ of **109b** from *A. strictum* WT (collision energy 5-15 eV): 369, 165. MS^2 of *m/z* = 369 $[M-H]^-$ of **109b** from *A. oryzae* NSAR1 RSI111 3-3 (collision energy 5-15 eV): 369, 165.

Dec 2013 – Dec 2017	**PhD student** in Organic Chemistry, Gottfried Wilhelm Leibniz Universität Hannover, Germany
Oct 2010 – Jul 2013	**Master of Science**, Chemistry, Julius-Maximilans-Universität Würzburg, Germany, and Heriot Watt University Edinburgh, Scotland
Oct 2007 – Aug 2010	**Bachelor of Science**, Chemistry and Management, Universität Ulm, Germany
Jun 2007	**Abitur**, Johannes-Scharrer Gymnasium Nürnberg, Germany