Natural Products Isolation from *Trichoderma reesei*

Von der Naturwissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover

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

von
Mary Labib Kamel Labib Shenouda, M.Sc. (Ägypten)

2022

Referent: Prof. Dr. Russell J. Cox

Korreferentin: Prof. Dr. rer. nat. Natalia Tschowri

Tag der Promotion: 14.02.2022

In Advance Publications of the Dissertation

Partial results obtained in the course of this PhD thesis were published in advance in the following articles approved by the corresponding mentor.

1. **Shenouda, Mary L.**, Maria Ambilika, and Russell J. Cox. "*Trichoderma reesei* Contains a Biosynthetic Gene Cluster That Encodes the Antifungal Agent Ilicicolin H." *Journal of Fungi* 7.12 (2021): 1034.

2. **Shenouda, Mary L.**, Maria Ambilika, and Russell J. Cox. "Heterologous expression of secondary metabolites in *Trichoderma reesei* for waste valorization", in preparation.

Printed and/ or published with the support of the German Academic Exchange Service.

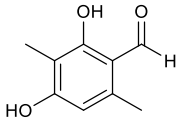
Abstract

Keywords: Natural product, Biosynthesis, *Trichoderma reesei*, Ilicicolin-H, Heterologous expression.

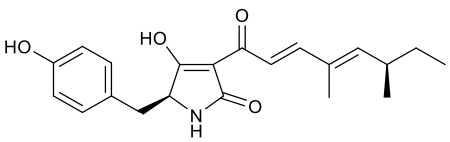
The focus of the present work is to develop *Trichoderma reesei* as a heterologous host for the production of natural products and to explore the biosynthetic potential of this fungus by expressing one of its cryptic PKS-NRPS gene clusters. Developing *T. reesei* as a heterologous host was first done by knocking out two of the main genes from the sorbicillin biosynthetic gene cluster in the auxotrophic strain *T. reesei* QM6a Δ*tmus53* Δ*pyr4* to develop a strain with cleaner metabolic background. Afterwards, a new vector with native *pdc* promoter and *pyr4* selection marker was constructed and used to express the polyketide synthase *aspks1* under the activity of the native *pdc* promoter. The transformed *T. reesei* strains were able to produce the expected natural product 3-methylorcinaldehyde **40** on different media and types of waste materials such as orange peel, banana, kiwi and potato peel as well as barley straw. Quantification of the produced compound **40** on potato peel showed the ability of the strain to produce up to 128 $mg \cdot kg^{-1}$ dry weight potato peel.

A second vector with two more native promoters (enolase and cDNA1 promoters) was constructed to allow the expression of more than one gene simultaneously in *T. reesei*. The new vector was then used to express the first two genes from the tenellin biosynthetic gene cluster, *i.e.* PKS-NRPS (encoded by *tenS*) and ER (encoded by *tenC*) under the activity of the *pdc* and *cDNA1* promoters, respectively. *T. reesei* transformants expressing *tenS* and *tenC* showed the production of the expected compound pretenellin A **77** on different media. The new strain was also able to produce **77** on banana peel, but not on barley straw under the tested cultivation conditions.

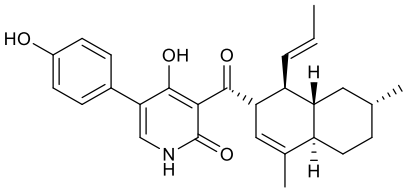
Bioinformatic analysis of the *T. reesei* genome revealed the presence of 11 PKS and two PKS-NRPS genes. Since no PKS-NRPS products have been previously reported from *T. reesei*, one of the two PKS-NRPS (TRIREDRAFT_58285) gene clusters in *T. reesei* was investigated. This was done first by overexpression of the transcription factor, which showed limited success due to the very low titre of the produced compound. Therefore, heterologous expression of this gene cluster was done in *Aspergillus oryzae* NSAR1, which confirmed that this gene cluster is responsible for the production of ilicicolin-H **72** in addition to many other intermediates and shunt products produced by the expression host.



40



77



72

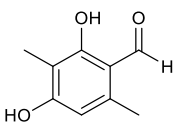
Zusammenfassung

Schlagwörter: Naturstoffe, Biosynthese, *Trichoderma reesei*, Ilicicolin-H, heterologer Ausdruck.

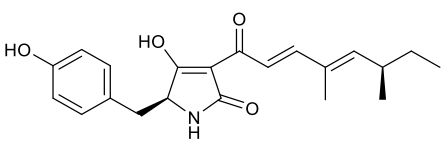
Der Schwerpunkt der vorliegenden Arbeit liegt auf der Entwicklung von *Trichoderma reesei* als heterologer Wirt für die Produktion von Naturstoffen, sowie der Erforschung des biosynthetischen Potenzials des Pilzes durch Expression eines seiner kryptischen PKS-NRPS Gencluster. Die Entwicklung von *T. reesei* als heterologer Wirt erfolgte zunächst durch den knock-out von zwei Hauptgenen des Sorbicillin-Biosynthese-Genclusters im auxotrophen Stamm *T. reesei* QM6a Δ*tmus53* Δ*pyr4*, um einen Stamm mit saubererem metabolischen Hintergrund zu entwickeln. Anschließend wurde ein neuer Vektor mit nativem *pdc*-Promotor und *pyr4* Selektionsmarker konstruiert, der zur Expression der Polyketidsynthase *aspks1* mit dem nativen *pdc* Promotor verwendet wurde. Die transformierten *T. reesei* Stämme produzierten, das erwartete Naturprodukt 3-Methylorcinaldehyde **40** auf verschiedenen Medien und auf verschiedenen Arten von Abfallmaterialien wie Orangen-, Bananen-, Kiwi-, und Kartoffelschalen sowie Gerstenstroh. Die Quantifizierung der produzierten Verbindung **40** auf Kartoffelschalen zeigte, dass der Stamm bis zu 128 $mg.kg^{-1}$ Trockengewicht Kartoffelschalen produzierte.

Ein zweiter Vektor mit zwei weiteren nativen Promotoren (Enolase- und cDNA1-Promotoren) wurde konstruiert, um die gleichzeitige Expression von mehr als einem Gen in *T. reesei* zu ermöglichen. Der neue Vektor wurde dann verwendet, um die ersten beiden Gene aus dem Tenellin-Biosynthese-Gencluster zu exprimieren, d. h. PKS-NRPS (kodiert durch *tenS*) und ER (kodiert durch *tenC*) unter der Aktivität der pdc- bzw. *cDNA1*-Promotoren. *T. reesei*-Transformanten, die *tenS* und *tenC* exprimieren, produzierten die erwartete Verbindung Pretenellin A **77** auf verschiedenen Medien. Der neue Stamm war auch in der Lage, unter den getesteten Kultivierungsbedingungen **77** auf Bananenschalen, nicht aber auf Gerstenstroh zu produzieren.

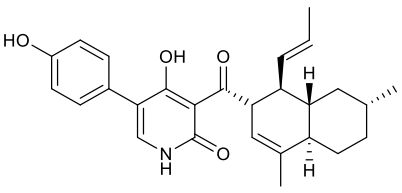
Die bioinformatische Analyse des *T. reesei*-Genoms ergab die Präsenz von 11 PKS- und zwei PKS-NRPS-Genen. Da bisher keine PKS-NRPS-Produkte aus *T. reesei* berichtet wurden, wurde eines der beiden PKS-NRPS-Gencluster (TRIREDRAFT_58285) in *T. reesei* untersucht. Dies geschah zunächst durch die Über-Expression von Transkriptionsfaktoren, die aufgrund des sehr niedrigen Titers der produzierten Verbindung nur begrenzten Erfolg zeigte. Daher wurde eine heterologe Expression dieses Genclusters in *Aspergillus oryzae* NSAR1 durchgeführt, was bestätigen konnte, dass dieser Gencluster für die Produktion von Ilicicolin-H **72** neben vielen anderen Zwischen- und Shunt-Produkten verantwortlich ist.



40



77



72

Acknowledgment

First, I would like to thank prof. Russell J. Cox for giving me the opportunity to work under his supervision. It was really a great honour to work with you. Thank you for all your advice and help over the years and for listening to me and encouraging me when I did not have the energy to continue working. I would have never wished for a better supervisor. Thank you also for having the patience to teach me and guide me through these years.

I also would like to thank Prof. Natalia Tschowri for taking the time to be the co-referee of my PhD thesis. I am also very grateful to prof. Oliver Plettenburg for being the chair of the examination board and examiner. I am also thankful for the DAAD and the Egyptian ministry of higher education for supporting my stay here in Germany.

I would also like to thank Dr. Elizabeth Skellam (Liz) for her help in the first couple of years. You were so kind to me and although I was really sad to say goodbye to you, but I am also really happy for you and I wish you all the best in your scientific carreer. I also would like to thank Karen for her help and advice in the first year in the lab, Eric for his help with bioinformatics and Kevin for his advice in the writing.

Special thanks go for the BMWZ media kitchen and technical staff, especially Katja. Thanks a lot for your help and for making everything look so easy. I would also like to thank all the people responsible for the NMR and mass spectrometry facilities in the OCI, espicially Dr. Jörg Fohrer, Dr. Lin Müggenburg and Dr. Gerald Dräger.

I would like to thank all the current and past Cox group members, especially Carsten, Katharina, Henrike, Chongqing, Tao, Maurice, Pia, Lukas, Sen, Jin, Lei, Dongsong, Yunlong, Haili, Doawan, Jenny, Ole, Slawik, Eman, Jing and Ahmed. I aslo would like to thank my Bachelor student Maria for all her help in the lab.

I would like to thank my amazing friends; Amal, Mary, Linda, Sarah, Halah, Sara, Nermin, Samira, Sarah, Mariam, Reem and Asmaa. I was told that having a PhD abroad will be very difficult emotionally and I would feel lonely, but you made it feel like home. You made this journey enjoyable with our talks, walks, late nights, trips and delicious Egyptian food.

Finally, I would like to thank the most important people in my life, my family. Mom, Dad, Mina and F. Theofan, you have always been the most loving and supporting family anybody could ask for. I can not thank you enough for all your help and support and for everything you did for me throughout my life and thank god for having you in my life. Submitting this thesis is only possible because of you.

Abbreviations and Units

µl	microliter
6-MSA	6-Methyl salicylic acid
6-MSAS	6-Methyl salicylic acid synthase
ACP	Acyl carrier protein
AcpS	Acyl carrier protein synthase
AT	Acyltransferase
BGC	Biosynthetic gene cluster
BLAST	basic local alignment search tool
BLASTp/n	BLASTprotein/nucleotide
Carb	Carbenicillin
DNA	Deoxyribonucleic acid
cDNA	complementary DNA
C-MeT	Carbon-methyltransferase
CoA	Coenzyme A
COSY	Correlation spectroscopy
DAD	Diode array detector
dd. H_2O	Double distilled water
DEBS	6-Deoxyerythronolide B
DH	Dehydratase
DMAPP	Dimethylallyl pyrophosphate
ELSD	Evaporative light scattering detector
ER	Enoyl reductase
ESI	electron spray ionisation
FA	Fatty acid
FAD	Flavin adenine dinucleotide
FAME	Fatty acid methyl ester
FAS	Fatty acid synthase
FMO	Flavin-dependant monooxygenase
gDNA	Genomic DNA
GOI	gene of interest
GRAS	Generally recognised as safe
HEx	Heterologous expression
hFAS	Human fatty acid synthase
HMG-CoA	3-hydroxy-3-methylglutaryl-CoA
HMBC	heteronuclear multiple bond correlation
HPLC	High performance liquid chromatography
HR	Homologous recombination
HRMS	High resolution mass spectrometry
hr-PKS	Highly-reducing PKS
hrs	hours
HSQC	heteronuclear single quantum correlation
hyg	hygromycin B
Hz	Hertz
iPKS	Iterative polyketide synthase
IPP	Isopentenyl diphosphate
ITS	Internal transcriber spacer
Kan	Kanamycin
kb	kilo base pairs
KO	knock out
KR	Ketoreductase
KS	Ketosynthase
M	Molar
ME	Malt extract
mFAS	Mammalian fatty acid synthase
mg	milligram
min	minute
mL	millilitre
MOS	3-methylorcinaldehyde synthase
mRNA	messenger RNA
MS	mass spectrometry
MVA	Mevalonic acid
m/z	mass to charge ratio
NAD(P)(H)	nicotinamide adenine dinucleotide (phosphate)
NHEJ	non-homologous end-joining
NMR	nuclear magnetic resonance
nm	nanometer
NP	Natural product
nr-PKS	Non-reducing PKS
NRPS	Non-ribosomal peptide synthase
P450	Cytochrome P450
P_{amyB}	AmyB promoter
PCR	Polymerase chain reaction
PPP	Pentose phosphate pathway
PEG	Polyethylene glycol
PK	Polyketide
PKS	Polyketide synthase
PKS-NRPS	Hybrid polyketide synthase-non-ribosomal peptide synthase
ppm	Part per million
PPP	allylic prenyl pyrophosphates
PPTase	Phosphopantetheine transferase
pr-PKS	Partially reducing PKS
Q-TOF	Quadrupole time-of-flight
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT-PCR	Reverse transcriptase PCR
SAM	S-adenosyl methionine
SDR	Short chain dehydrogenase/reductase
SM	Secondary metabolites
SNAC	S-acetylcysteamine
SOR	Sorbicillin
spp.	species
SS	Soysauce sucrose
Syn.	Synonym
T_{amyB}	AmyB terminator
TE	Thiolesterase
TE-buffer	Tris-EDTA buffer
TF	Transcription factor
TFD	Transcription factor decoy
TR.KO	*Trichoderma reesei* sorBC Knockout strain
T_R	Retention time
Ura	Uracil
UV	Ultraviolet
WT	Wild type
YPAD	Yeast peptone adenine dextrose

Contents

1 Introduction

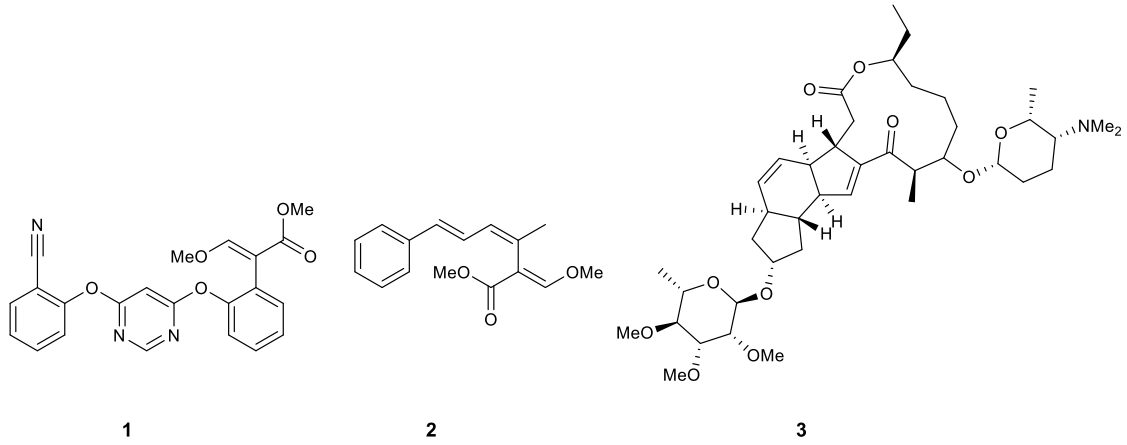
1.1 Natural Products

For all of human history, nature has served as the main source of all human needs: food, clothing, tools and medicines.[1] Natural products from plants, fungi, microorganisms and (much more rarely) animals have been a vital source of drugs that have been used for the treatment of human diseases since the dawn of medicine.[2] Until the 19th century, NPs were used mainly as crude preparations of plant material of unknown constitution to alleviate pain, diarrhoea, infection, and various other diseases. But with the revolution in chemistry during the 19th century, scientists became more interested in the identification of the structures of the active constituents of these naturally derived medicinal compounds.[3] This led to the discovery of many NPs that profoundly impacted advances in biology and inspired drug discovery and therapy.[4]

For the producing organism, these NPs may act as defence or signalling chemicals that the organism produces to fight against invading pathogens. These chemical "NPs" act either directly by inhibiting the proliferation of the invading pathogen or indirectly by disrupting chemical signal processes related to growth and development of the pathogen.[5] In that sense, these chemicals have been utilized knowingly and unknowingly by humans as a major source of agrochemicals and pharmaceuticals to fight against various diseases.[5] The successful development of the fungicide azoxystrobin **1** (inspired from the structure of natural product strobilurin A **2**) and the pesticide spinosad **3** has resulted in a renewed interest in natural product-based agrochemicals. The general public perception of the natural product-based compounds as safe and less toxic has awakened the desire for the production of more pharmaceutical compounds derived from natural sources.[5] NPs can either be used in their native form as drugs or they can serve as lead compounds for drug discovery by semisynthetic or totally synthetic modification to enhance their biological activity or physiochemical properties.[3]

NPs are broadly defined as small molecules produced naturally by any organism. These are divided into two classes: primary metabolites are the *essential* metabolites of life and are usually found in all organisms; secondary metabolites (sometimes also called specialised metabolites, SM) are often unique to a single organism or related species and

are *not essential* for life - however SMs often give an organism a selective advantage. Hence NPs with industrial applications can be produced from primary or secondary metabolism of living organisms (plants, animals or microorganisms).[6–8]



1 2 3

Figure 1.1 Chemical structures of some fungicide and pesticide natural products and derivatives.

During growth, living cells break down complex, high-molecular-weight carbon and energy sources through the action of catabolic enzymes. Reassembly of the end products of catabolism results in the production of primary metabolites such as amino acids, nucleotides, vitamins (cofactors), carbohydrates, and fatty acids. The cells then use these biosynthetic intermediates to form complex and essential metabolites that give structure and biological function to the organisms. Therefore, primary metabolites have essential roles in the growth and reproduction of the organism. On the other hand, secondary metabolism gives rise to compounds that are not essential for the growth of the organism.[9] So, SMs are generally defined as organic compounds that are not essential for the growth, development, or reproduction of an organism, but they do have an important role in the interaction of the organism with its environment, and their production is usually low.[6–8] SMs are usually more common in organisms that lack an immune system, as they increase the organism's fitness for survival by acting as alternative (chemical) defence or signalling mechanisms.[10] Biosynthesis of these SMs takes place from a limited number of precursor metabolites derived from primary metabolism (Figure 1.2).[11] This general classification of primary and secondary metabolites however still leaves a "grey area" at the boundaries, where some groups of NPs could be assigned to either division. A good example of that would be fatty acids, which are mainly classified as primary metabolites, but some representatives are extremely rare and found only in a few species, therefore can be considered as SMs.[12]

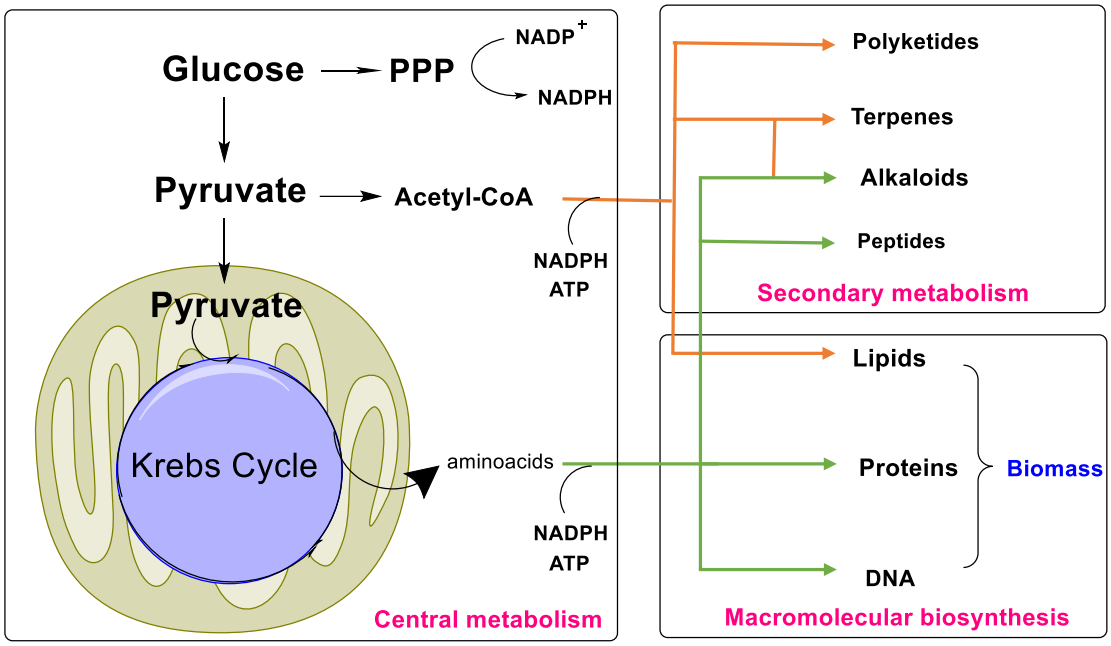


Figure 1.2 Metabolic pathway in living organisms. PPP = Pentose Phosphate Pathway and NP peptides can be ***both*** ribosomal and non-ribosomal

1.2 Importance of NPs in Medicine and Industry

Historically, SMs derived from plants were the basis of most early medicines and drugs such as aspirin **4**, morphine **5** and pilocarpine **6**.[13,14] So, by the year 1990, natural products and their analogues constituted about 80 % of drugs. More recently, almost half of the approved drugs in the market since 1994 are derived or inspired from NPs and over 100 NP-derived compounds are currently undergoing clinical trials. Antibiotics [*e.g.*, penicillin **7**, cephalosporin **8**, and erythromycin **9**], antifungals [*e.g.*, griseofulvins **10**], lipid control agents [*e.g.*, lovastatin **11** and analogues], antimalarials [*e.g.*, artemisinin **12** and quinine **13**], and anticancer drugs [*e.g.*, doxorubicin **14** and taxol **15**] have revolutionized medicine.[15,16] Furthermore, SMs have historically been used in numerous applications in agriculture such as for the control of fungal, weed and insect pests. Because of safety and environmental problems, many synthetic agricultural agents have been, and currently are being, considered for removal from the market. This creates an urgent need to find alternative ways to control farm pests and pathogens, which makes agricultural fungicidal natural products such as strobilurins and their synthetic derivatives a very appealing alternative.[17–19]

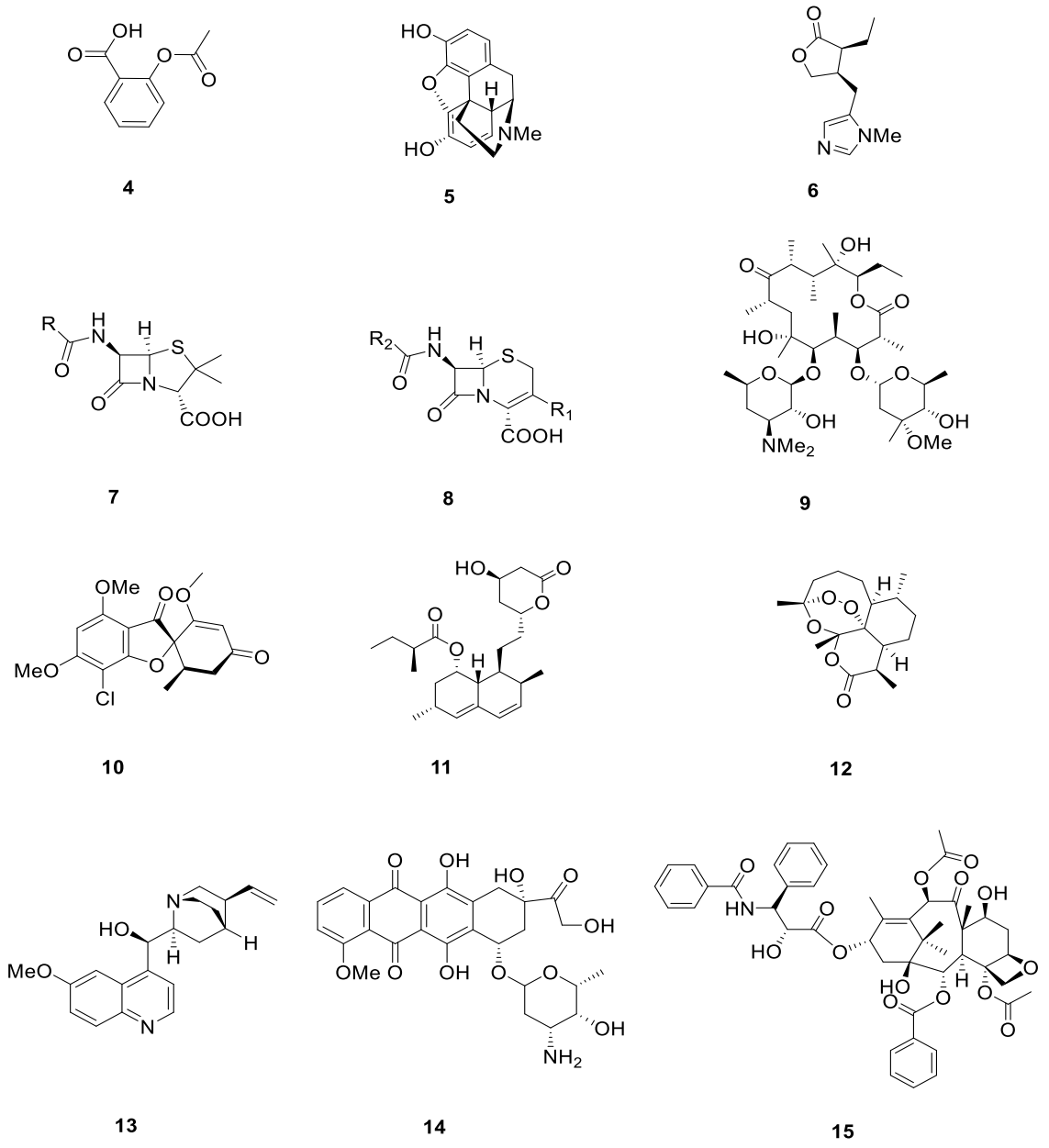


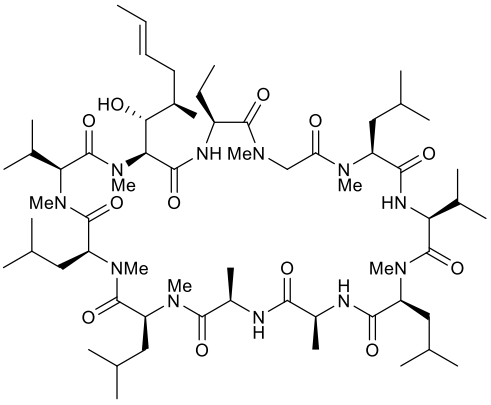
Figure 1.3 Chemical structures of some drugs derived from natural products.

Primary metabolites also play a pivotal role in industry, with amino acids, nucleotides, vitamins, solvents and organic acids being the most important primary metabolites used industrially with numerous applications in the food, chemical and nutraceutical industries.[20] They are mainly manufactured by microbial fermentation rather than chemical synthesis because the fermentations are economically competitive and produce biologically useful isomeric forms. Additional applications of primary metabolites lie in their impact as precursors of many pharmaceutical compounds. The roles of primary

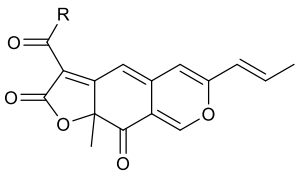
metabolites and the microbes which produce them will certainly increase in importance as time goes on.[20]

1.3 Natural Products from Fungi

In the area of microbial research, the main sources for natural products to date have been fungi and terrestrial *Actinomycetes* and hence several examples of compounds that affect our everyday life are fungal secondary metabolites.[16,21] Fungi are known for their peculiar chemistry, where they occasionally produce unusual amino acids and unlike animals and higher plants, they can use both the L- and D-enantiomers in their biosynthesis of SMs.[22] These fungal natural product can be either used as drugs or as leads for industrially important chemicals. Examples include antibiotics (*e.g*. penicillin **7**), immunosuppressant (*e.g*. cyclosporine **16**), cholesterol-lowering drugs (*e.g*. lovastatin **11**), industrial fungicides (*e.g*. strobilurin **2** and azoxystrobin **1**) and food colorants (*e.g*. azaphilones **17-18**).[23] The wide range of biological activities exhibited by fungal NPs reveals their structural diversity, which arises from their biosynthetic classification (*e.g*., fatty acids, polyketides, alkaloids, terpenes and peptides).[23] In the following section, only some selected classes of fungal NPs will be briefly discussed.



16



Rubropunctatin **17** R= C_5H_{11}
Monascorubrin **18** R= C_7H_{15}

Figure 1.4 Chemical structures of some important industrial compounds.

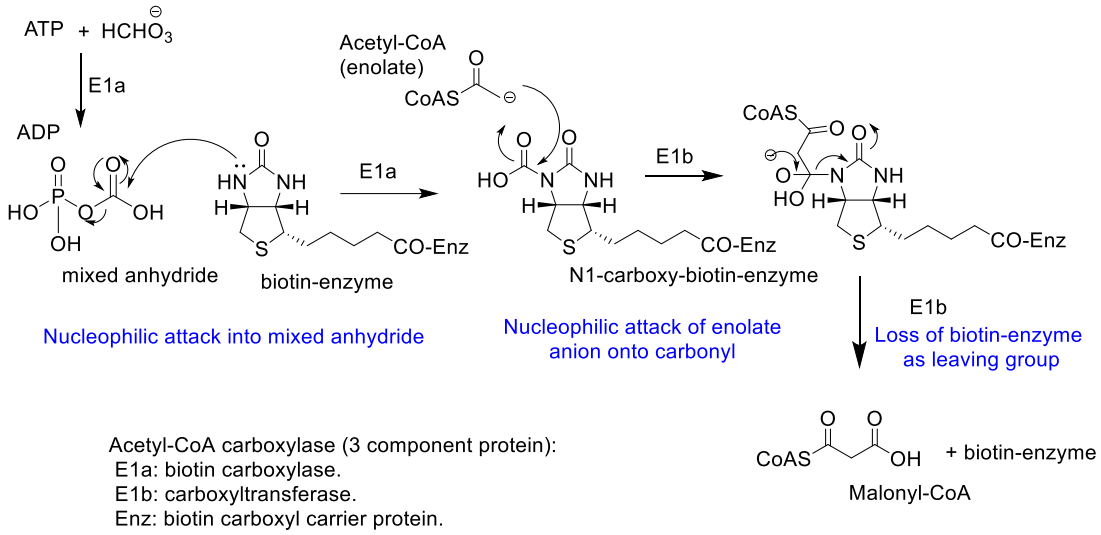
1.4 Classes of Natural Products

1.4.1 Fatty Acids

Fatty acids (FA) are considered as an important class of primary metabolites which are normally produced by all living organisms to serve as energy storage compounds or to be incorporated in the cell membrane and different cellular components.[24] Due to the close resemblance of the fatty acids and polyketides (PK) biosynthetic pathways, it is meaningful to discuss FA biosynthesis first since it has been well studied way before that of PKs. Fatty acid biosynthesis is well studied and is known to be catalysed by the enzyme fatty acid synthase (FAS).[12] FAS from various organisms show significant differences in quaternary structure.[12] For example, in bacteria and plants, fatty acid biosynthesis is carried out by individual, monofunctional enzymes in a dissociated fatty acid biosynthesis system, which is termed a type II FAS system. On the other hand, eukaryotes carry out fatty acid biosynthesis using large, multifunctional FAS enzymes, termed type I system that integrate all required enzymatic activities into unique complexes. Type I FASs can be further subdivided into fungal FASs and vertebrate FASs. [24]

Fatty Acid Biosynthesis

The biosynthesis of long-chain fatty acids occurs in two distinct steps. The first step is the conversion of acetyl-CoA to malonyl-CoA. This reaction is catalyzed by the three-component enzyme acetyl-CoA carboxylase (Scheme 1.1). This carboxylation step is mainly done to make the Claisen reaction involving acetyl-CoA more favourable. This reaction requires CO_2 (solubilized as bicarbonate), ATP and the coenzyme biotin. The second step is the conversion of acetyl-CoA and malonyl-CoA to palmitate in the presence of NADPH, a reaction catalyzed by the fatty acid synthetase (FAS).[12,25,26]



Scheme 1.1 The conversion of Acetyl-CoA into malonyl-CoA by Acetyl-CoA carboxylase enzyme.[12]

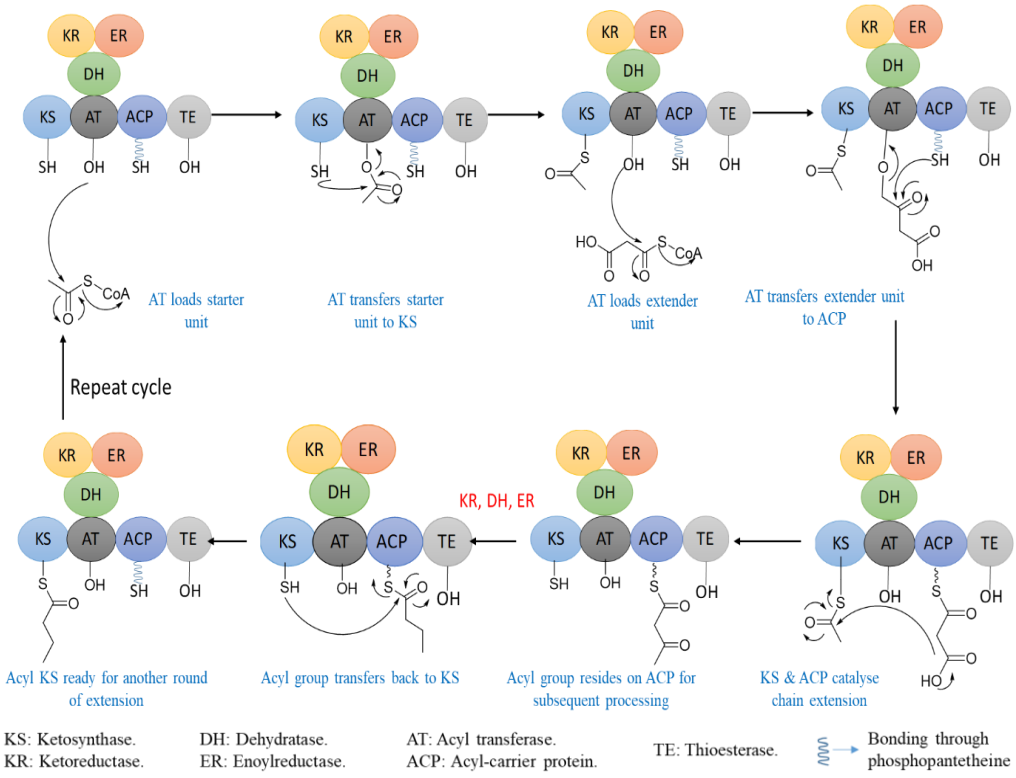
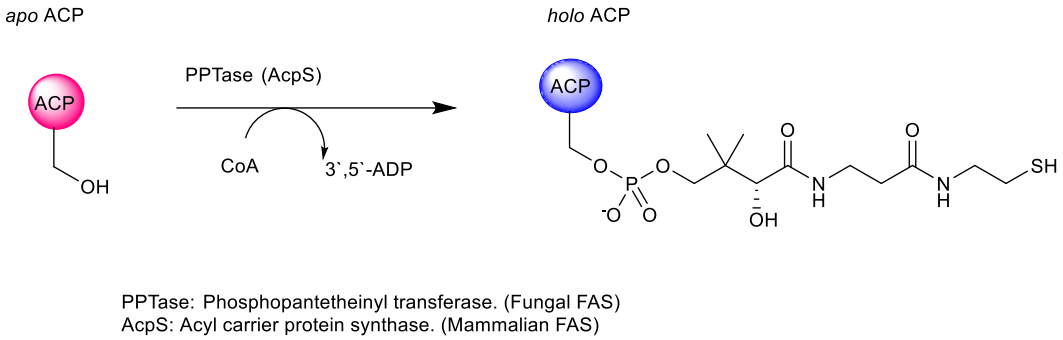


Figure 1.5 Reaction mechanism of Fatty acids biosynthesis by FAS.

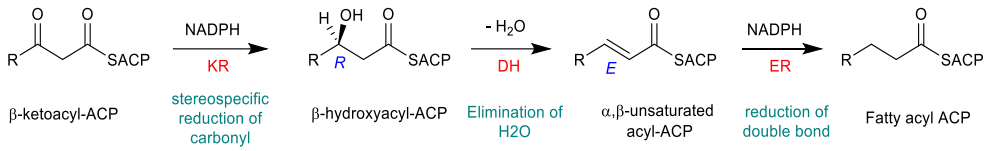
Palmitate synthesis by FAS is an iterative process that requires an orderly involvement of seven different enzymes, five of which participate sequentially eight times each during the synthesis of one molecule of palmitate.[25] The cycle starts with the acyl transacylase (AT) which loads the starter and extender units in turn onto an acyl carrier protein (ACP). The ACP then carries the intermediates through the whole catalytic cycle (Figure 1.5).[27]

The identification of a small, heat-stable protein, later named ACP, was and still is considered a key discovery in the understanding of fatty acid biosynthesis.[28] This is due to the central role of ACP in fatty acid biosynthesis as it transports and presents the growing acyl chain to appropriate catalysts for elongation and fatty acid production.[29] The ACP gets post-translationally modified with a phosphopantetheine prosthetic group derived from coenzyme A (Scheme 1.2). This phosphopantetheine group provides a long flexible arm, allowing the growing fatty acid chain to reach the active site of each enzyme in the complex, so the different chemical reactions can be performed without the need to release the intermediates from the enzyme.[12,24,28] The acyl intermediates of fatty acid biosynthesis are bound to ACP through a thiolester linkage attached to the terminal sulfhydryl of the 4`-phosphopantetheine prosthetic group.[30]



Scheme 1.2 Post-translational modification of ACP.[29,30]

Decarboxylative condensation of the acetyl and malonyl moieties to an ACP-bound ß-ketoacyl intermediate is then catalysed by the KS. The ß-carbon position is then modified by sequential action of ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) enzymes to yield a two carbon units longer saturated acyl product. This cycle is known as β-processing and involves three-stages: stereoselective reduction to the corresponding β-hydroxy ester (KR) with the consumption of NADPH in the reaction; dehydration to the *E* α,β-unsaturated ester (DH); then reduction of the alkene (ER, Scheme 1.3). The resulting saturated acyl group then functions as a starter substrate for the next round of elongation, each cycle leads to an increase in the chain length by two carbon atoms, until the growing fatty acid chain reaches the desired length of carbon atoms. The growing fatty acid chain can then be either released in the form of fatty acyl-CoA as in fungi or as a free fatty acid chain by the action of a thiolesterase (TE) as in vertebrate FAS (vFAS).[12,31]



Scheme 1.3. β–processing of FA by the action of KR, DH and ER domain.

Although both fungal and mammlian FAS are type I multifunctional proteins, there are some substantial structural differences between them (Table 1.1).[12,24,32]

Table 1.1. Comparison between fungal and mammalian Fatty acid synthase genes.

	Fungal FAS	***Mammalian FAS***
Domain organization	The functional domains are arranged on two separate polypeptides.	The functional domains are arranged on a single protein chain.
	Hetero-dodecamer	Homodimer
Size	2.6 MDa.	540 KDa
The Loading of acetyl and malonyl CoA	AT domain: loads acetyl-CoA. MPT domain: loads malonyl-CoA.	MAT domain: loads both Acetyl-CoA and Malonyl-CoA.
The activation of ACP domain	PPT domain: activate ACP by inserting a phophopantetheine arm	*holo-ACP synthase*: a separate enzyme that activates ACP by adding the phosphopantetheine arm
The release of the fatty acid chain	Released as Fatty acid-CoA by MPT domain.	Released as free fatty acid by TE domain.

1.4.2 Polyketides

Polyketides constitute a large class of natural products that have long been recognized as one of the most important classes of SMs mainly because of their outstanding range of functional and structural diversity.[12,33,34]

Polyketide Biosynthesis

Polyketide and fatty acid biosynthesis are closely related, regarding both the nature of the chemistry used in chain extension from a common pool of simple precursors and also in the overall architectural design of the enzymes that are used for chain assembly (Figure 1.6).[12,26,33–35] This striking similarity between the two enzymatic systems resulted in a concerted development of the two fields, with advances in each discipline often having immediate impact and application to the other.[26] FAS and PKS differ mainly in the starter

and extender units they utilize. For FAS, the starter unit is typically acetyl-CoA and the extender unit is malonyl-CoA and they are both loaded by a single AT domain. On the other hand, the starter unit for PKS can be acetyl or other acyl-CoAs (in bacteria) and the extender unit can be malonyl or methylmalonyl-CoA. Both starter and extender groups are loaded by specialized ATs. Another difference is the degree of reduction of the growing ß-ketoacyl moiety. In FASs the ß-ketoacyl moiety produced after each extension step is completely reduced to a saturated carbon chain prior to the next condensation step. However, in PKSs, the ß-ketoacyl moiety may be completely or partially reduced, so that the product retains keto, hydroxyl or enoyl groups along the acyl chain.[26]

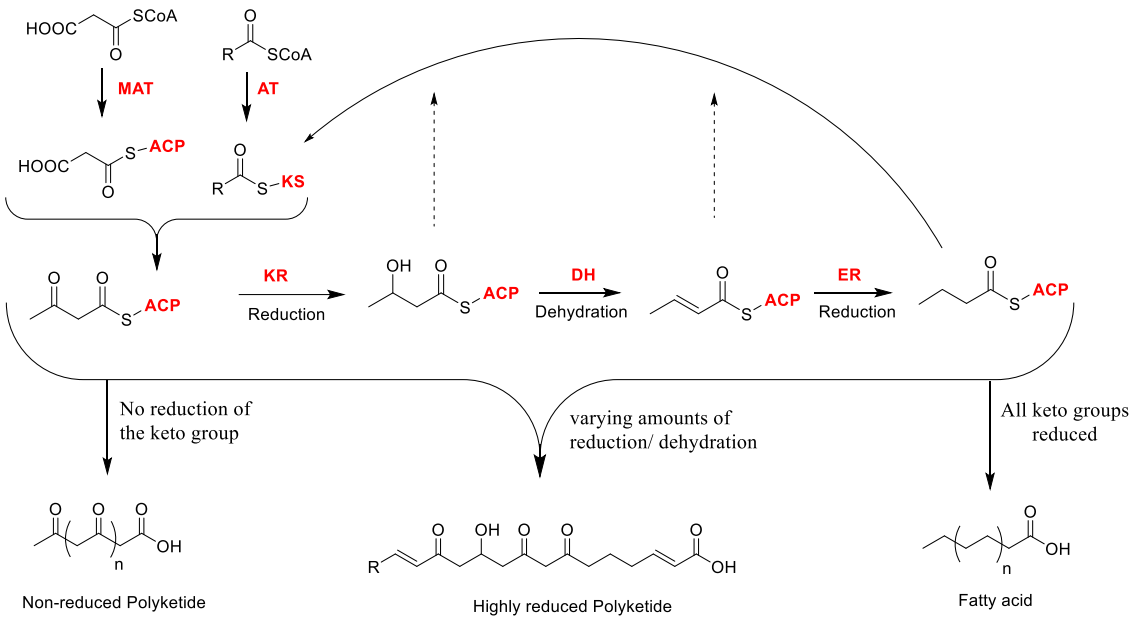


Figure 1.6. Generic reaction scheme applicable for both FAS and modular PKS.[26]

There are 3 main known architectures of PKS. Like type I and type II FAS, **type I PKS** are very large multifunctional proteins with individual functional domains, whilst **type II PKS** are composed of a complex of individual monofunctional proteins. **Type III PKS** are homodimeric proteins that utilize coenzyme A esters rather than ACPs and they employ a single active site to perform a series of decarboxylation, condensation, cyclization, and aromatization reactions (Table 1.2).[12]

Table 1.2 Comparison between the different classes of Polyketide synthase genes.

Type I	Type II	Type III
Large multifunctional proteins with individual functional domains	A complex of individual monofunctional proteins	Homodimeric proteins
use ACP to activate acyl-CoA substrates and to channel the growing polyketide intermediates	use ACP to activate acyl-CoA substrates and to channel the growing polyketide intermediates	utilize coenzyme A esters rather than ACPs
found in bacteria and fungi	restricted to bacteria	Found in plants, bacteria, and fungi.
Subdivided into 'iterative' (i.e. repeating) and 'noniterative' categories.	They are all of the iterative type	They employ a single active site to perform a series of decarboxylation, condensation, cyclization, and aromatization reactions.
Iterative systems (like the FASs) use their functional domains repeatedly to produce a particular polyketide. Non-iterative systems possess a distinct active site for every enzyme-catalysed step.	Catalyse the production of Aromatic Polyketides.	Catalyse the production of flavonoids and stilbenes.

Type I PKS can be further subdivided into iterative and non-iterative (modular) PKS (Table 1.2). The iterative PKSs use the same domains repeatedly to produce the desired polyketide, while non-iterative PKSs use distinct active site for every individual enzyme-catalysed step.[12] In spite of the apparent differences in the molecular architecture of the different PKSs, the chemical features of chain construction are basically the same in all PKSs.

Modular type I PKS, such as the erythromycin PKS, is usually discussed as a model for understanding all modular PKSs, since polyketide synthesis by this group of enzymes is probably the easiest to envisage and understand.[12] Each of the DEBS proteins consists of two functional modules, hence the name 'modular' PKS. Each module contains the three main domains, KS, AT and ACP, in addition to other variable domains (Figure 1.7). The main three domains mentioned before are essential to catalyse one cycle of chain extension, while the other set of domains such as DH, ER or KR are mainly associated with keto group modification. The loading didomain consists of AT and ACP

domains, that accepts the starter unit propionate from propionyl-CoA. Throughout the biosynthetic pathway, the polyketide chain remains bound to the PKS and the cycle is finally terminated by the action of TE. TE catalyses the off-loading and cyclisation of the final heptaketide intermediate to give 6-Deoxyerythronolide B **19**.[36]

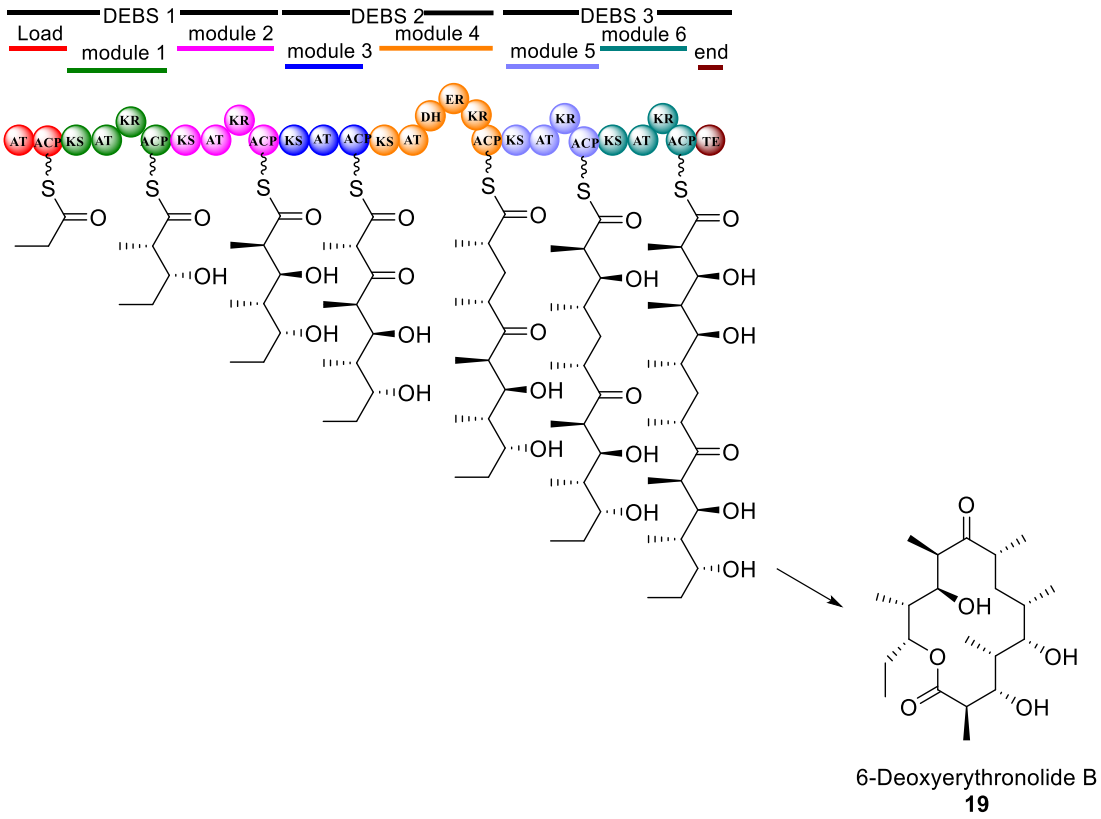
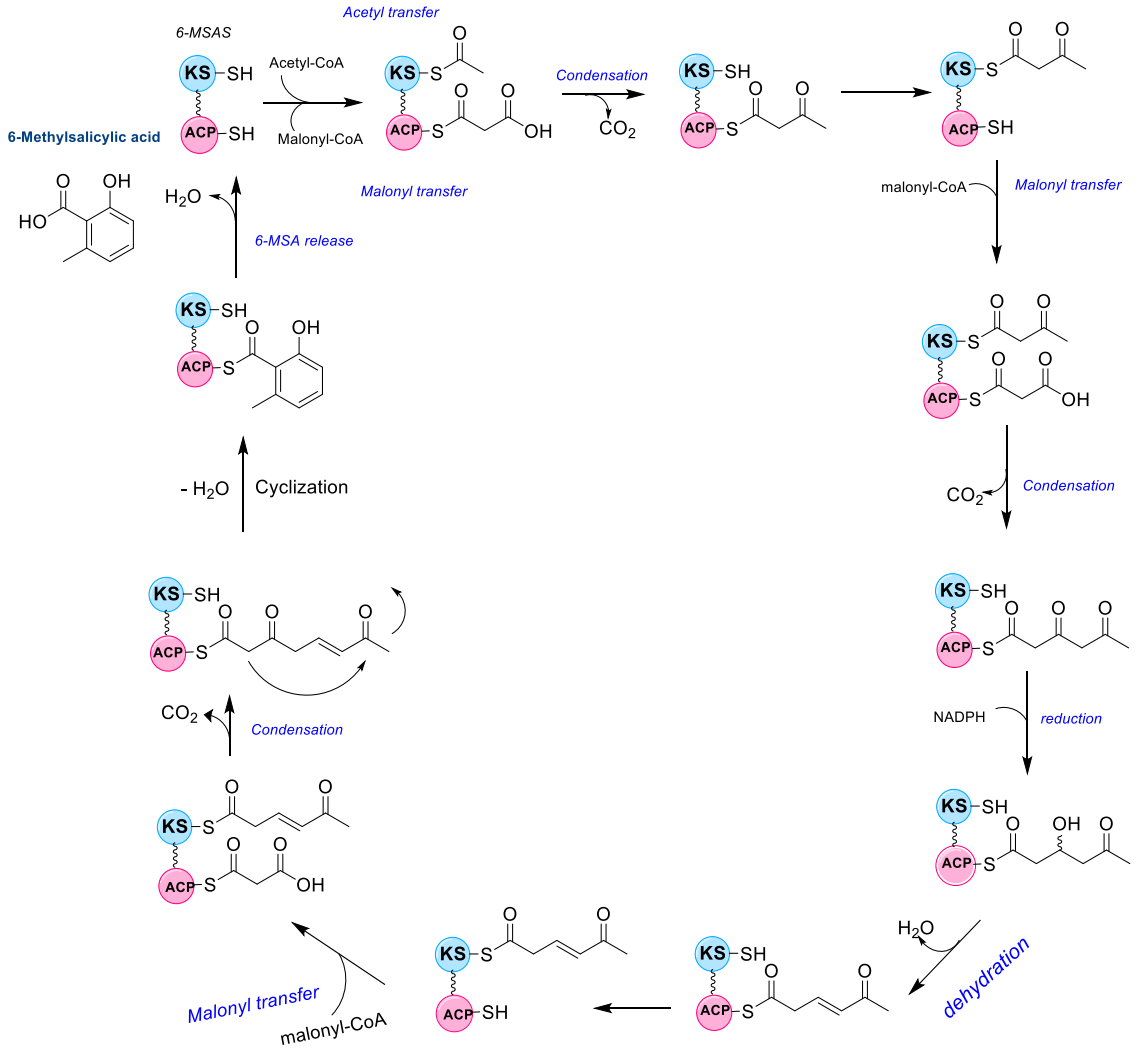


Figure 1.7. Domain organisation of the erythromycin polyketide synthase and the biosynthesis of 6-Deoxyerythronolide **19**.

Fungal polyketides are mainly produced by very large multifunctional iterative type I PKSs which are closely related to mammalian FAS.[37] Despite the huge structural variability of polyketides of fungal origin, they are predominantly produced by single iterative type I polyketide synthases (iPKS) that use a single set of domains (one module) to synthesize complex natural products in a highly programmed and iterative manner.[38] Iterative type I PKS is usually divided into three main classes: non-reducing (NR), highly-reducing (HR) and partially-reducing (PR).

6-Methyl salicylic acid synthase (6-MSAS) is an example of pr-PKS, which unlike DEBS, consists of only one single module that is used iteratively to produce 6-Methyl salicylic acid (6-MSA) and the PKS domain organization is : KS-AT-DH-KR-ACP (Scheme 1.4).[12,36]



Scheme 1.4 Biosynthesis of 6-MSA by the activity of 6-MSAS.

A well-studied example of iterative type I hr-PKS is LovB (also known as lovastatin nonaketide synthase, LNKS). This enzyme consists of one set of domains: KS-AT-DH-*C*-MeT-ER^0-KR-ACP-C (Figure 1.8). However, ER domain of LNKS (encoded by lovB) was shown to be dysfunctional and hence coexpression of the *trans*-acting ER (lovC) together with LovB in *Aspergillus nidulans* results in the production of dihydromonaclin L **20**.[39]

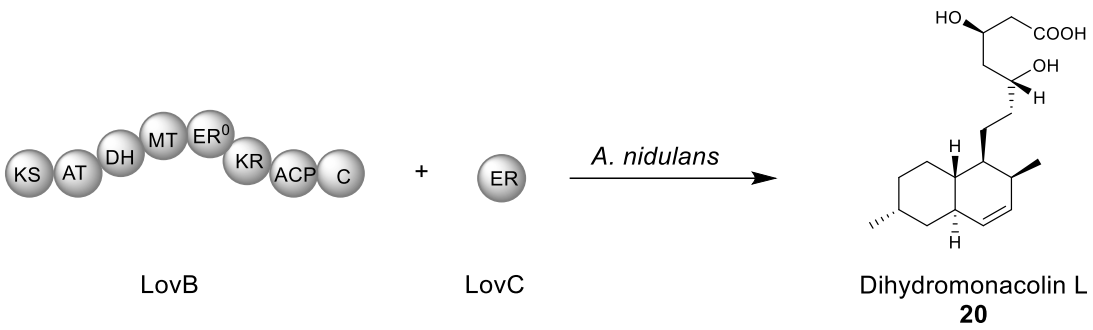
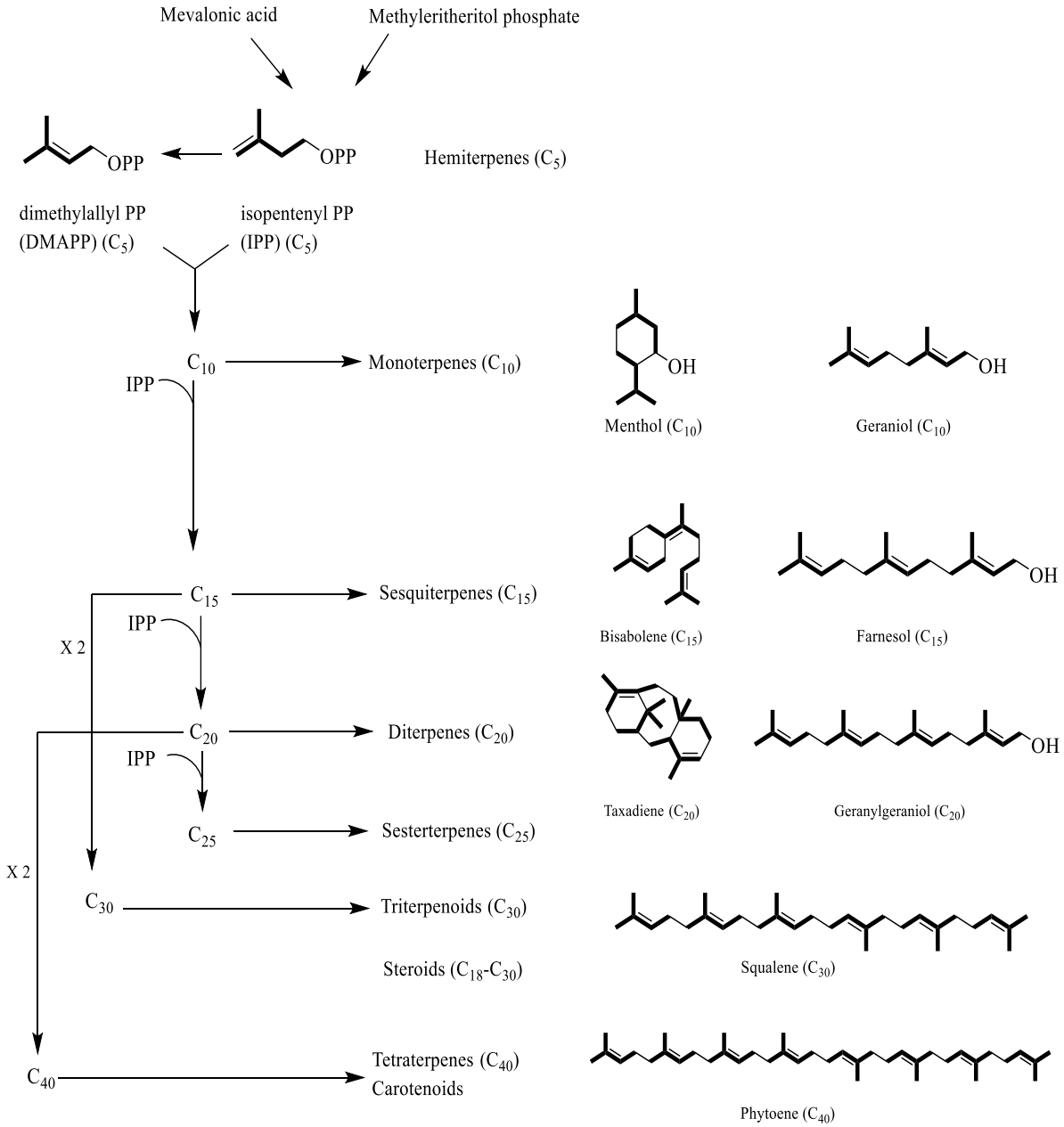


Figure 1.8 Domain structure of hr-PKS (LovB) and *trans*-acting ER (LovC) and biosynthesis of dihydromonaclin L **20** in *A. nidulans*.

1.4.3 Terpenes and Terpenoids

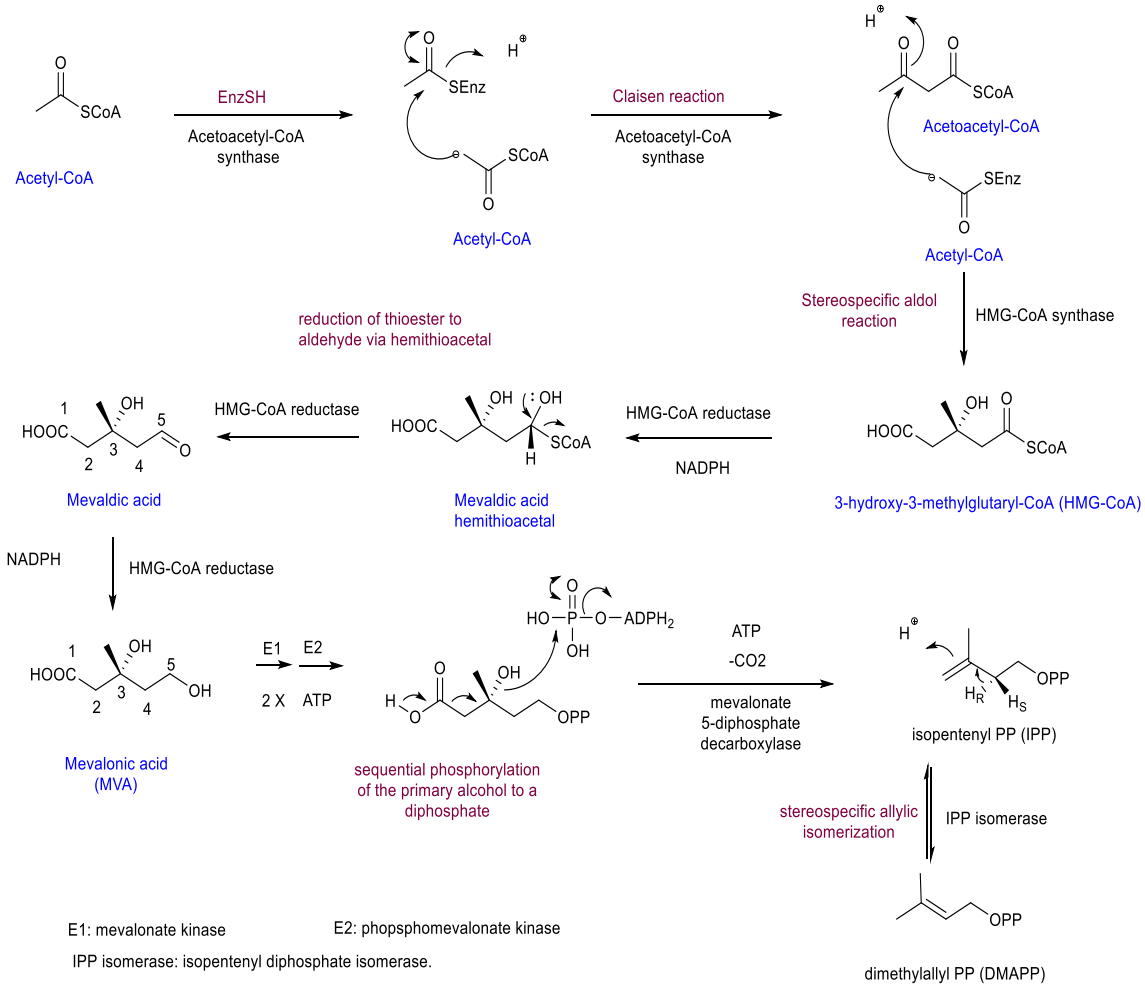
Terpenes and Terpenoids constitute a large and structurally diverse class of SMs that is formed by head-to-tail condensation of C_5 isoprene units. They are classified into hemiterpenes (C_5), monoterpenes (C_{10}), sesquiterpenes (C_{15}), diterpenes (C_{20}), sesterterpenes (C_{25}), triterpenes (C_{30}) and tetraterpenes (C_{40}, Scheme 1.5). Terpenes (compounds lacking heteroatoms) and terpenoids (compounds with heteroatoms) have a wide range of activities in plants where they serve as hormones (*e.g*. gibberellins, abscisic acid), photosynthetic pigments (phytol, carotenoids), fragrance and taste to defend themselves from organisms that feed off on them and structural components of membranes (phytosterols). In addition to that, they provide a range of commercially useful products, including solvents, flavourings, fragrances, adhesives and coatings. Some of them even have important applications in the pharmaceutical industry (such as artemisinin **12**, taxol **15**) and the agrochemical industry (such as pyrethrins, azadirachtin).[12,40,41]



Scheme 1.5 Terpenes and terpenoids biosynthetic pathway with some examples.

The biochemically active isoprene units can be derived from two pathways: either the mevalonate pathway or methylerythritol phosphate pathway. In the mevalonate pathway, three molecules of acetyl-coenzymeA are used to form mevalonic acid (MVA). Two molecules of acetyl-CoA are first combined together in a claisen condensation reaction to form acetoacetyl-CoA, which react later with the third molecule of acetyl-CoA via a stereospecific aldol addition to give the branched-chain ester 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). This is followed by a reduction of the thiolester group of HMG-CoA into aldehyde then into a primary alcohol to form MVA by the action of HMG-CoA reductase enzyme (Scheme 1.6). This is essentially an irreversible and rate-

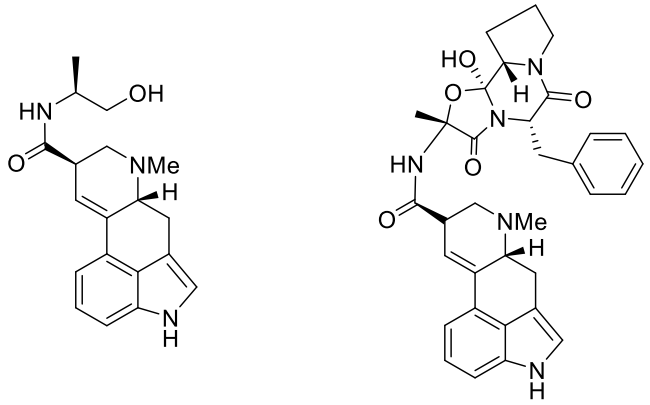
limiting step in the biosynthesis of mevalonate and steroids. Therefore drugs that inhibit HMG-CoA reductase enzyme, commonly called statins, are widely used in the treatment of hyperlipidaemia and in the management of cholesterol levels.[12]



Scheme 1.6 Mevalonic acid pathway.

1.4.4 Alkaloids

Alkaloids are low molecular weight nitrogen containing compounds found mainly in plants, but also known in bacteria and more commonly in fungi.[12] A simple and general definition of alkaloids was suggested in 1983 by Pelletier: he defined an alkaloid to be any cyclic compound containing nitrogen in a negative oxidation state.[42] According to this definition, alkaloids are not only those NPs that have a nitrogen in a heterocyclic ring but also the many other exceptions that have an extracyclic bound nitrogen such as capsaicin and colchicine.[42] An alternative and widely used definition is compounds derived from amino acids that contain a *basic* nitrogen. Some important examples of alkaloids in fungi include ergot alkaloids from *Claviceps purpurea* such as the oxytocic ergometrine **21** and ergotamine **22**, which is used in the treatment of migraine.[12]



Ergometrine
21

Ergotamine
22

Figure 1.9 Chemical structures of some alkaloids of fungal origin.

1.4.5 Peptides

Bioactive peptide natural products are categorized based on their mode of construction into either ribosomally synthezied and post-translationally modified peptides (RiPPs) or peptides syntheized by multimodular non-ribosomal peptide synthase (NRPS). RiPPs are ribosomally synthesized precursor peptides (20 - 110 amino acids) that undergo post-translational modification and processing. They consists of two main parts; leader and core peptides. Another two optional parts can also be present in the protein sequence, namely signal and recognition sequences. The leader peptide facilitates recognition by the post translational modification enzymes that modify the core peptide. Finally, proteolytic cleavage of the leader peptide results in the production of the mature peptide

(Figure 1.10). Many RiPP undergoes macrocyclization, which increases their metabolic stability and improve their cellular uptake and hence contribute to their bioactivity.[43]

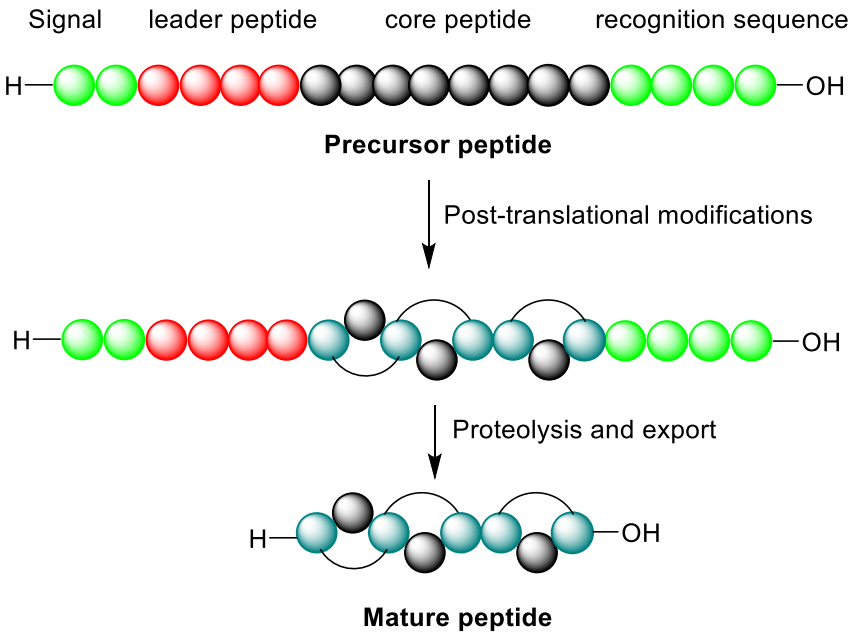


Figure 1.10 General representation of the biosynthesis of RiPPs.[43]

Non-ribosomal peptides are, on the other hand, synthesized by multifunctional, modular NRPS. The minimal domain set of NRPS is adenylation domain (A-domain), thiolation or peptide carrier protein (T-domain or PCP) and condensation domain (C-domain). The adenylation domain specifically recognizes and activates amino acids and directs their covalent binding to the 4`-phosphopantetheine arm of the PCP-domain. Then elongation of the peptide chain is done in the condensation domain by amide formation between different amino acids. Other domains for additional structural modifications can also be present such as *N*-methylation (*N*-Me domain), cyclization (Cy), epimerization (E), reductase (R) and thiolesterase domains (TE).[43,44] Despite their structural diversity, RiPP are syntheized by ribosomes, which can only employ proteinogenic amino acids. On the other hand, non-ribosomal peptides are synthesized by NRPS that can accept both proteinogenic and non-proteinogenic amino acids.[43]

1.5 Metabolic Engineering of Microorganisms

With advances in fractionation techniques to isolate and purify natural products and in the analytical techniques to determine their structures, pure bioactive compounds can now be isolated from fermentation broths in less than two weeks and the structures of more than 90% of these new compounds can be elucidated within two weeks. However, recently, natural product research became less popular in many major drug companies and, in some cases, have been completely replaced by combinatorial chemistry, which is the automated synthesis of structurally related small molecules.[19] The complexity of fungal secondary metabolism is considered one of the main reasons for this decline in interest in fungal natural products. For example, fungi often produce NPs at very low levels or in response to specific environmental factors that are difficult to reproduce in the laboratory.[23] Aiming at surpassing all these challenges, substantial advances in strategies for engineering fungal NP biosynthesis have been made in recent years. Metabolic engineering is now used to convert these reluctant bio-factories into highly efficient, focused machines capable of generating huge quantities of a molecule of interest.[23,45] Strategies for metabolic engineering of NPs have been applied with two main objectives in mind: 1) Strain improvement to increase the titre of the target compound, and 2) to elucidate the NP biosynthetic pathway or to modify the NP scaffold for improved pharmacological properties, in other words, to produce novel compounds with enhanced biological activities.[46]

1.5.1 Metabolic Engineering Strategies for Strain Improvement

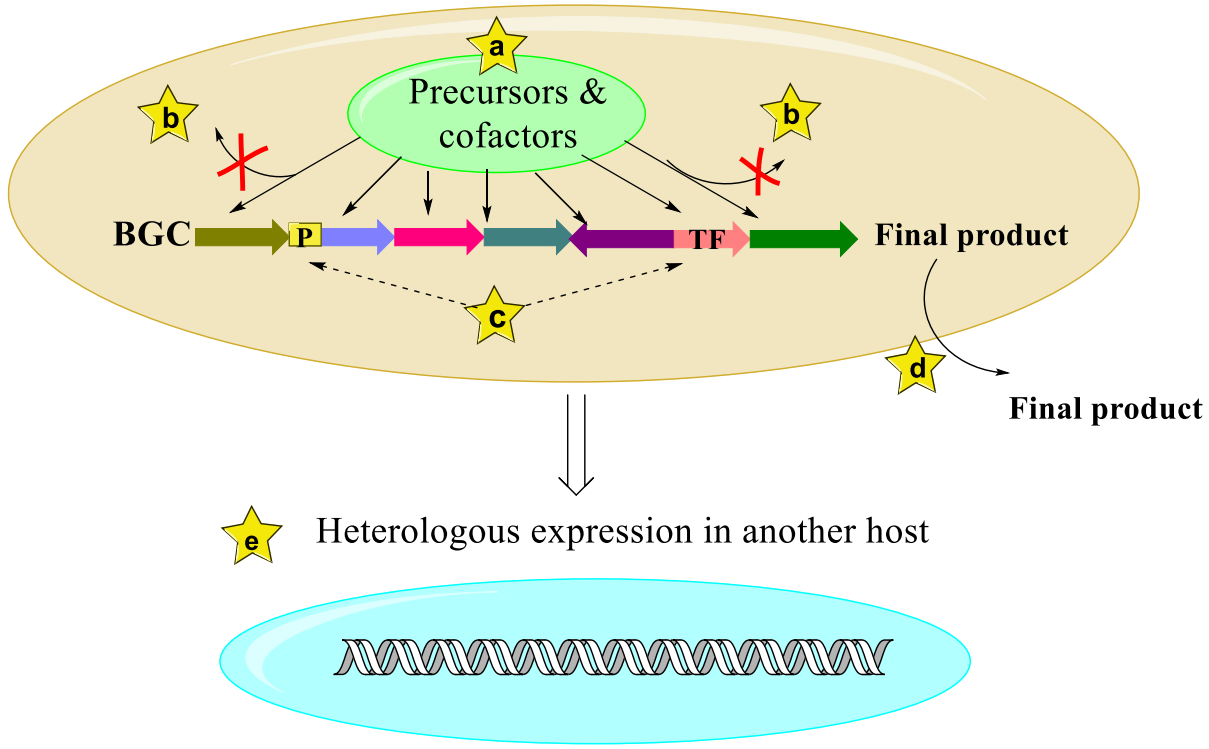
In the early years of fermentation processes, repeated random mutations followed by screening and selection were basically used for the development of better producing strains. But recently, with the significant advances in molecular biology and genetic engineering, substantial advances in strategies for engineering fungal NP biosynthesis have been introduced.[20,23] The used strategies merely depend on the genes being engineered, but generally enhancing the production of NPs can be done by the following strategies (Figure 1.11).[46,47]

- Overexpression of the genes involved in the biosynthesis of the metabolite of interest. Unfortunately, this strategy does not always lead to elevated production,

since each biosynthetic pathway is believed to consist of many sequential enzyme reactions, one of which is considered as the rate limiting step and functions to regulate the rate of the whole pathway. Overexpression of any gene other than the rate limiting step gene, would have almost no effect on the metabolite production.[47]

- Blocking competing pathways by knocking-out a reaction that degrades or converts the target metabolites. For example, in *E. coli*, a large amount of fatty acids was successfully produced by a combination of overexpression of genes involved in fatty acid synthesis and knockout of the genes involved in degradation.[46–48]
- Enhancing the production of the coenzymes required for the synthesis of the target primary metabolites as limited amounts of coenzymes such as ATP, NADH, and NADPH can severely affect the production of the metabolite. Overexpression of the malic enzyme gene in *Mucor circinelloides* was used to elevate the production yield of fatty acids by increasing the intracellular NADPH molecules available for use in fatty acid synthesis.[47,49]
- Discharging the final metabolites out of the cells can also result in enhanced production of the primary metabolite since intracellular accumulation of the final metabolite could stress the producing microorganism.[47]
- Heterologous expression of the BGC in another host which is usually desirable if the native producer does not grow well under industrial fermentation conditions, is genetically intractable, or is subjected to long growth periods.[46]

While the same strategies can be used for enhancing the production of both primary and secondary metabolites, there are still some fundamental differences between them. First, the biosynthetic pathways for some SMs are not fully understood unlike primary metabolites, which their metabolic pathways are almost fully elucidated. In addition, SMs are usually final products and therefore not susceptible to conversion or degradation by the producing microorganisms. Furthermore, they are usually secreted out of cells, so they do not accumulate in the cell. For these reasons, strategies for overproduction of primary metabolites might be more complicated than those of SM.[47]



a: Increasing the precursor and Cofactor supply.
b: Blocking competing pathways.
c: Overexpression of the genes by controlling the transcription factors(TF) and promoters (P).
d: Discharging the final metabolite out of the cell.
e: Heterologous expression of the biosynthetic gene cluster (BGC) in another host.

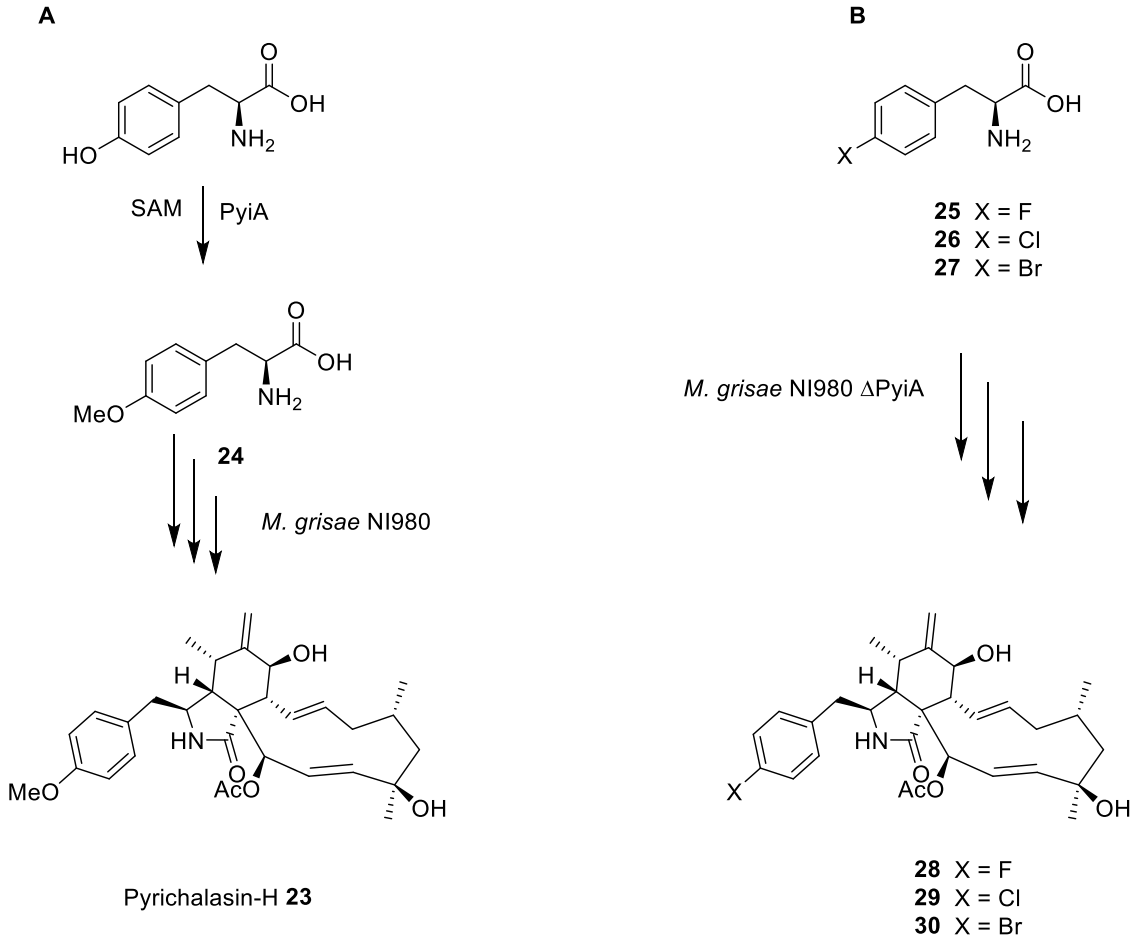
Figure 1.11 Summary of the general metabolic engineering strategies for strain improvement to enhance the production of fungal NPs.

1.5.2 Metabolic Engineering Strategies for Structural Diversification and Pathway Elucidation

With the massive advancement in genome sequencing and the rapid accumulation of fully sequenced genomes, it has become clear that many BGC in fungi are either orphan (*i.e.* unlinked to a specific NP) and/or silent (*i.e.* inactive under normal laboratory conditions).[50,51] Many metabolic engineering strategies have been developed to link novel NPs to genes, and to map novel biosynthetic pathways.[51] Metabolic engineering strategies have not been just used in the activation of silent BGC but also in the modification of the known BGC to produce new compounds with enhanced pharmacological activities.[46] Especially for highly complex NPs, where selective modifications through chemical means can be a bit challenging, metabolic engineering offers a great tool to perform these modifications with ease.[46] Strategies used are summarized below (Figure 1.12).[23,45,46,50,52,53]

- *Gene disruption through gene knockout and silencing.* A straight-forward strategy for introducing a structural change is *via* disruption of a certain gene that acts downstream in a pathway (usually a tailoring enzyme) either by gene knockout or knockdown (*i.e.* silencing). This method was proved useful in the elucidation of the biosynthetic pathway of NP such as stipitatic acid, where the knockout of genes encoding downstream-tailoring reactions helped reveal the catalytic steps leading to ring expansion to the 7-carbon aromatic tropolone nucleus.[50] Another example for pathway elucidation is the use of targeted RNA silencing in *Penicillium expansum* to link the biosynthesis of cytochalasans to a PKS–NRPS gene cluster.[50,54]

- *Precursor-directed biosynthesis and Mutasynthesis.* Another seemly simple technique to generate new analogs of a well-known NP is the precursor feeding experiment or Mutasynthesis, which exploits the promiscuity of some enzymes to accept substrates similar to their natural substrate. A good example is in the production of pyrichalasin H **23** (Scheme 1.7), where the PKS-NRPS (PyiS) uses *O*-methyltyrosine **24** produced by the action of the *O*-methyltransferase (PyiA) as a sterter unit. Feeding 4`-flouro, 4`-chloro, 4`- bromophenylalanine (**25**-**27**) to *Magnaporthae oryzae* NI980 pyiA knockout strain resulted in the production of the corresponding 4`-halo cytochalasin (**28**-**30**) in titres comparable to that of pyrichalasin-H in the WT strain. [46,50,55]



Scheme 1.7 A, Production of pyrichalasin H in *M. grisea* NI980; **B**, Production of 4`-halo cytochalasin derivatives by mutasynthesis.[55]

- *Pathway engineering and combinatorial biosynthesis.* In this approach reshuffling and swapping the genes and modules between closely related NP pathways are used to create new combinations and chimeric molecules. This can be done either by engineering modular megasynthases through mixing and matching their subunits, modules or domains or by the combinatorial engineering of the tailoring enzymes. For example, co-expression of a glycosyl transferase gene pair from *Beauveria bassiana* with various fungal PKS resulted in the production of 20 glycosylated polyketides with improved solubility and cytotoxicity.[46,56]

- *Protein engineering.* Simple swapping of functional domains sometimes results in non-functional or heavily impaired chimeric enzymes, occasionally due to the inability of the enzymes downstream in the pathway to accept the modified substrate. Therefore protein engineering either by rational design or directed

evolution has been used to increase the substrate promiscuity or to alter the substrate specificity or programming fidelity of the biosynthetic enzymes.[46,57] For example,chimeric syntheatases were constructued by exchanging sub-fragments of *C*-methyltransferase (*C*-MeT), pseudo-ketoreductase (ΨKR) and ketoreductase (KR) catalytic domain of the tenellin PKS-NRPS (TENS) with homologous fragments from desmethylbassianin (DMBS) and militarinone (MILS) PKS-NRPS. These chimeric synthetases showed altered programming fidelity and resulted in the production of mixtures of compounds with different methylation patterns and chain length.[46,57]

- *Epigenetic remodelling and activation of silent BGC.* As mentioned before, many BGC in Fungi are orphan and/or silent under certain growth conditions and therefore variation of fermentation conditions can result in the production of different compounds.[50] This phenomena has been attributed to the tight regulation of the production of these NPs in their native host, which means that many of these BGCs are located in transcriptionally repressed regions of the genome.[23,53] Therefore, direct manipulation of those BGCs' regulators may result in production of novel NPs. This can sometimes be achieved by different ways including epigenetics and chromatin remodelling, manipulation of global regulators and/or pathway-specific transcription factors or by using transcription factor decoys (TFD).[23,53] TFDs are DNA molecules designed to interfere with gene regulation through sequestration of repressors, thus rendering the BGC de-repressed. However, at present these methods are entirely empiracle.[53]

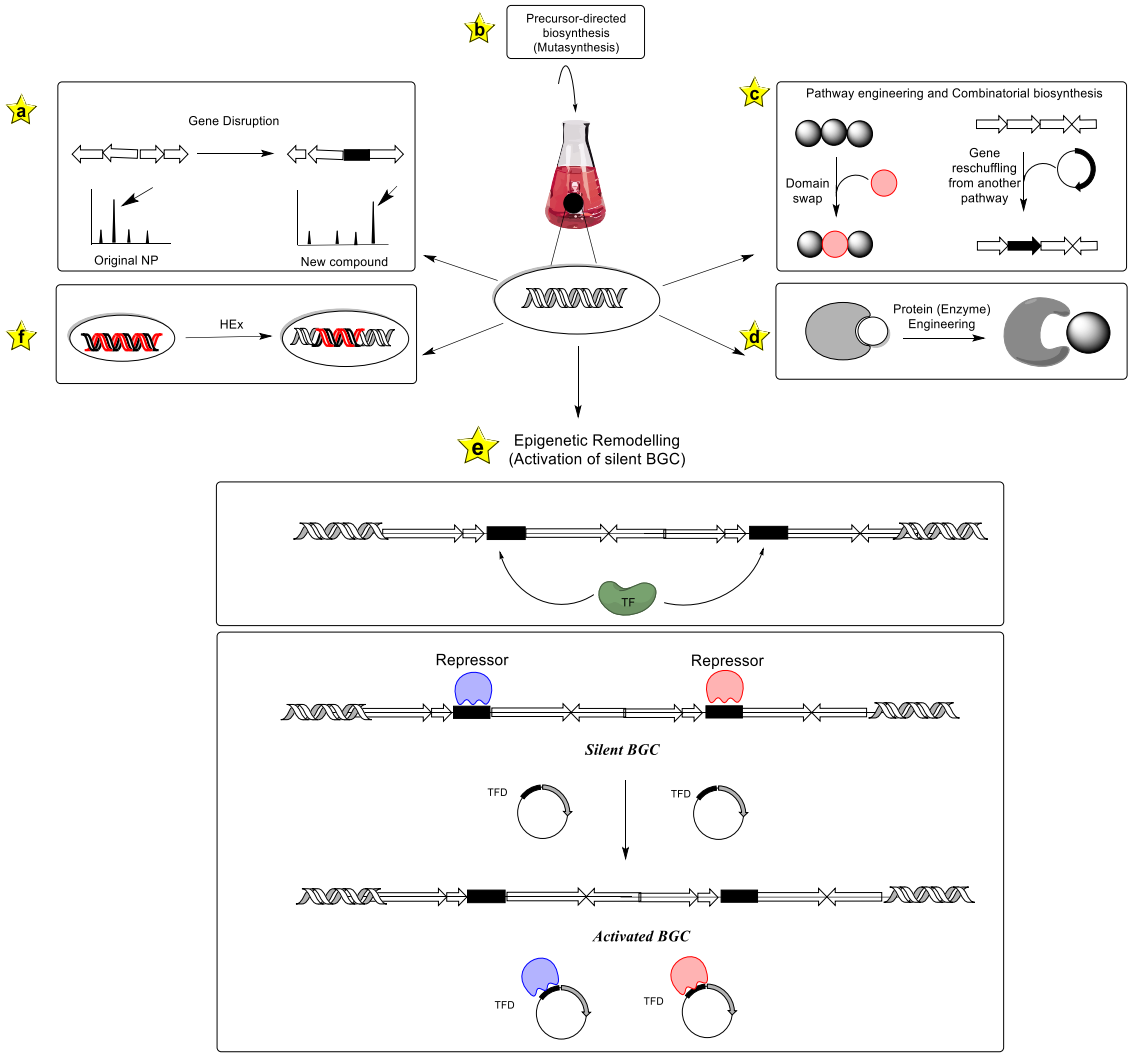


Figure 1.12 Summary of the general metabolic engineering strategies for NPs structural diversification and pathway elucidation.

Heterologous expression of BGC in another host. Although filamentous fungi produce many SM, analyses of fully sequenced fungi show that the genetic chemical diversity of sequenced BGCs is much larger than the number of metabolites actually isolated. Many of the BGCs must be silent under laboratory conditions, which further complicates product discovery, production, and characterization.[51] Heterologous expression (HEx) of the biosynthetic pathway in model systems or cell factories has facilitated product discovery and production in addition to elucidation of the functions of the genes involved in the biosynthesis. But the differences in intron splicing and in codon bias between less related organisms make the process of finding a suitable host for HEx challenging. Therefore using a filamentous fungal host for the HEx of fungal BGC is generally more successful than using bacterial or yeast strain.[58] Many of the widely used filamentous fungal species in HEx belong to the genus *Aspergillus* such as *A. oryzae* and

A. nidulans. *A. oryzae* is a species of particular interest due to its long history in food technology and its GRAS (Generally recognized as safe) status. Molecular tools to manipulate *A. oryzae* , such as selectable markers and promoters, have also been well established, making it good host for HEx of fungal NP.[21]

1.6 *Trichoderma*

Members of the fungal genus *Trichoderma* are able to survive in different geographical habitats mainly due to their high reproductive and competitive capabilites.[44,59,60] *Trichoderma* spp. are considered the most potent biocontrol agents in use today and around 90% of fungal biocontrol agents used to combat pathogenic microorganisms belong to the genus *Trichoderma*.[61,62] *Trichoderma* species secrete a multitude of biologically active natural products into their environment while having minimal nutritional needs. Therefore, *Trichoderma* spp. have gained much attention lately as promising sources of new antifungals for the management of plant pathogenic fungi. Many secondary metabolites of *Trichoderma* spp. have been reported to be active against plant pathogenic fungi. Examples include gliotoxin **31**, gliovirin **32**, harzianolide **33** and harzianopyridone **34**. In addition to their antimicrobial activities, some natural products isolated from *Trichoderma* spp. have been reported to have plant growth promotion activities such as 6-pentyl-2-H-pyran-2-one (6-PP) **35** and harzianic acid **36**.[61,63]

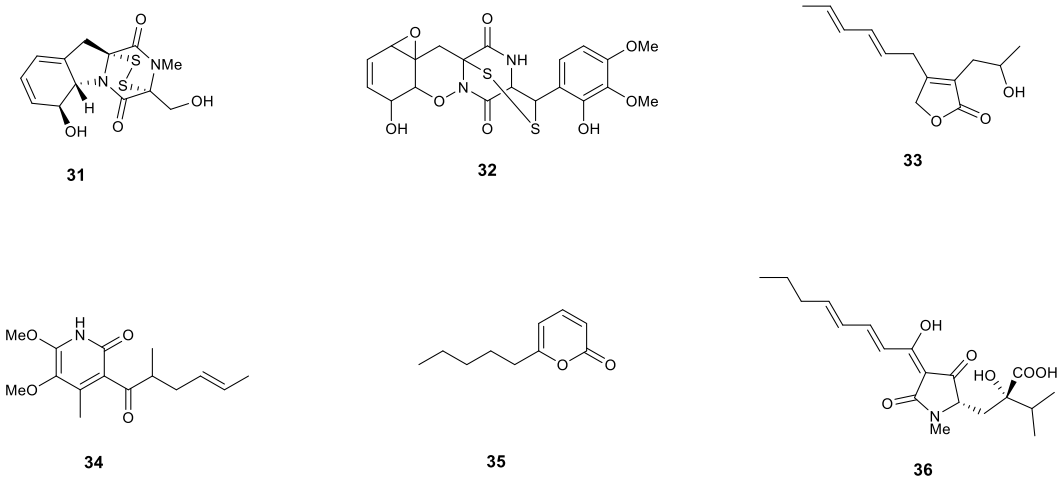


Figure 1.13 Some compounds isolated from *Trichoderma* spp.

The mechanism of the reported biocontrol activity of different *Trichoderma* spp. includes direct interaction by secretion of cell-wall degrading enzymes that degrade the cell wall

of invading microorganisms, competition with microorganisms on nutrients and by the secretion of biologically active secondary metabolites.[64,65] Indirect interactions involve activation of the plant systemic induced resistance.[65] Nevertheless, the ability of *Trichoderma* spp. to secrete cell-wall degrading enzymes had been the focus of attention for a long time.[64,66,67] With the increased interest in the identification of new antifungal and natural pesticides, exploring the biosynthetic potential of *Trichoderma* spp. and identification of the secondary metabolites responsible for this biocontrol activities have gained much attention recently.[61,63,68–71]

1.7 Aim of the Work

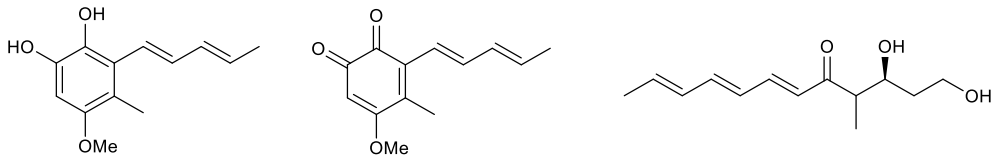
The general aim of this study is to improve *Trichoderma reesei* as a heterologous expression system for the production of high value chemicals. In order to do that, knocking-out the main biosynthetic gene cluster in *T. reesei* (*i.e.* sorbicillin BGC) would be done to produce a strain with relatively "clean" background to facilitate the heterologous expression. Also construction of expression vectors to facilitate heterologous expression of secondary metabolites in the new heterologous host. Then exploiting the inherent ability of *T. reesei* to produce high levels of cellulase enzymes that can convert plant biomass containing cellulose into glucose. These enzymes allow *T. reesei* to grow easily on waste materials and plant biomass such as wheat straw and municipal solid waste.[72,73] So *T. reesei* could then be used as a platform for the production of high value chemicals from waste materials.

Another aim of the thesis is to explore the biosynthetic potential of *T. reesei*. This could be done first through bioinformatics analysis of all PKS and hybrid PKS-NRPS gene clusters in *T. reesei*. Then the most promising BGC would be selected for further investigation by activation of the cluster selective transcription factor and also by heterologous expression of the BGC in a different heterologous host such as *Aspergillus oryzae* to confirm the function of this BGC.

2 Developing *Trichoderma reesei* as a Heterologous Host for Secondary Metabolite Production

2.1 Introduction

Recent advances in genome sequencing and bioinformatic analysis tools have revealed a multitude of new fungal biosynthetic gene clusters (BGC). Many fungal BGC remain silent under normal laboratory conditions due to suppressed transcription.[58,74] Therefore, heterologous expression in model organisms has offered a very useful platform for genome-driven natural product discovery as well as for the elucidation of biosynthetic pathways of known natural products. Many examples of cryptic and silent BGC have been successfully expressed in different fungal and bacterial heterologous hosts such as tricholignans **37**-**38** from *Trichoderma harzianum* t-22 (renamed *T. afroharzianum* t-22) in *Aspergillus nidulans*[71] and lavendiol **39** from *Streptomyces lavendulae* FRI-5 in *Streptomyces avermitilis* SUKA22.[75]



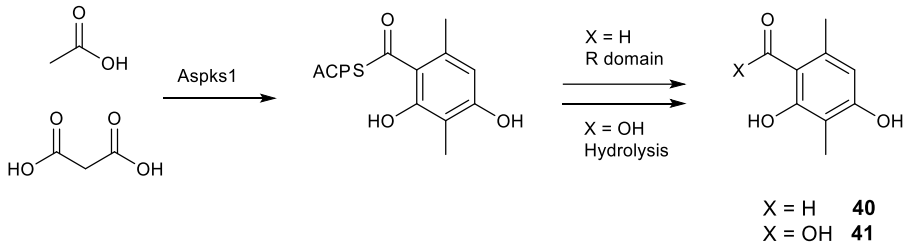
37 **38** **39**

Figure 2.1 Chemical structures of some compounds produced from cryptic gene clusters

2.1.1 Available Heterologous Hosts and their Limitations

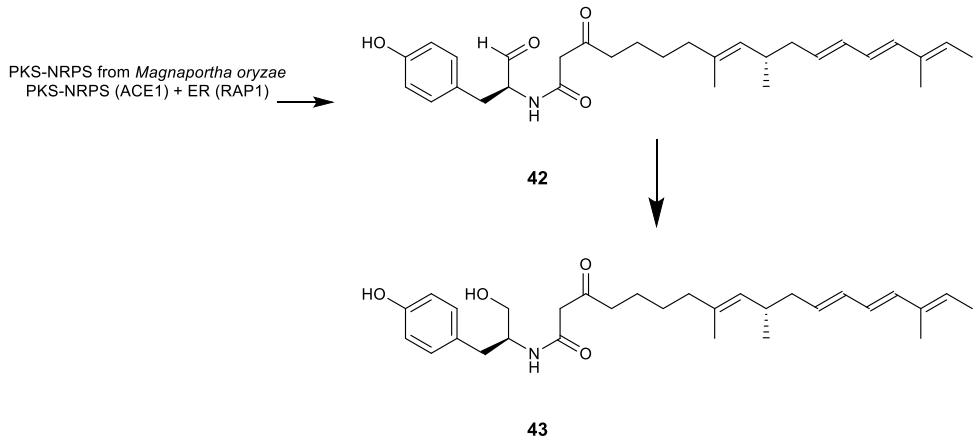
Despite the establishment of various fungal and bacterial heterologous hosts such as *E. coli, S. cerevisaea, Aspergillus nidulans* and *Aspergillus oryzae,* finding a suitable host for heterologous expression is still a challenging task for a number of reasons. For example, despite their great advantages as heterologous hosts, *E. coli* and *S. cerevisiae* have found limited application in elucidating fungal biosynthetic gene clusters mainly due to their inability to process introns that are commonly found in fungal genes, differences in codon bias and incorrect protein folding. Hence, filamentous fungal hosts have been more useful in heterologous expression of fungal BGC.[76] Well-established filamentous fungal hosts include different *Aspergillus* species such as *A. niger, A. oryzae* and *A. nidulans*.[77] However, there are still some limitations to their use such as the incorrect intron splicing especially when expressing BGC from a distantly related fungi.

Examples include the *ACE1* gene from *Magnaporthae oryzae*, where only two of the three introns are correctly spliced in *A. oryzae*.[78] Another example is the *aspks1* gene from *Acremonium strictum* (renamed as *Sarocladium schorii*). Analysis of the gene in the native producer *A. strictum* showed that one of the two introns in this gene could be spliced in two different ways; one to give a PKS with a reductive release mechanism to yield the aldehyde product 3-methylorcinaldehyde **40**, while the alternative splicing gave a truncated PKS that yielded 3-methylorsellenic acid **41** as the product.[76] However, heterologous expression of *aspks1* in *A. oryzae* yielded only one of the transcripts that resulted in the production of 3-methylorcinaldehyde **40** as the main product (Scheme 2.1).[76,79]



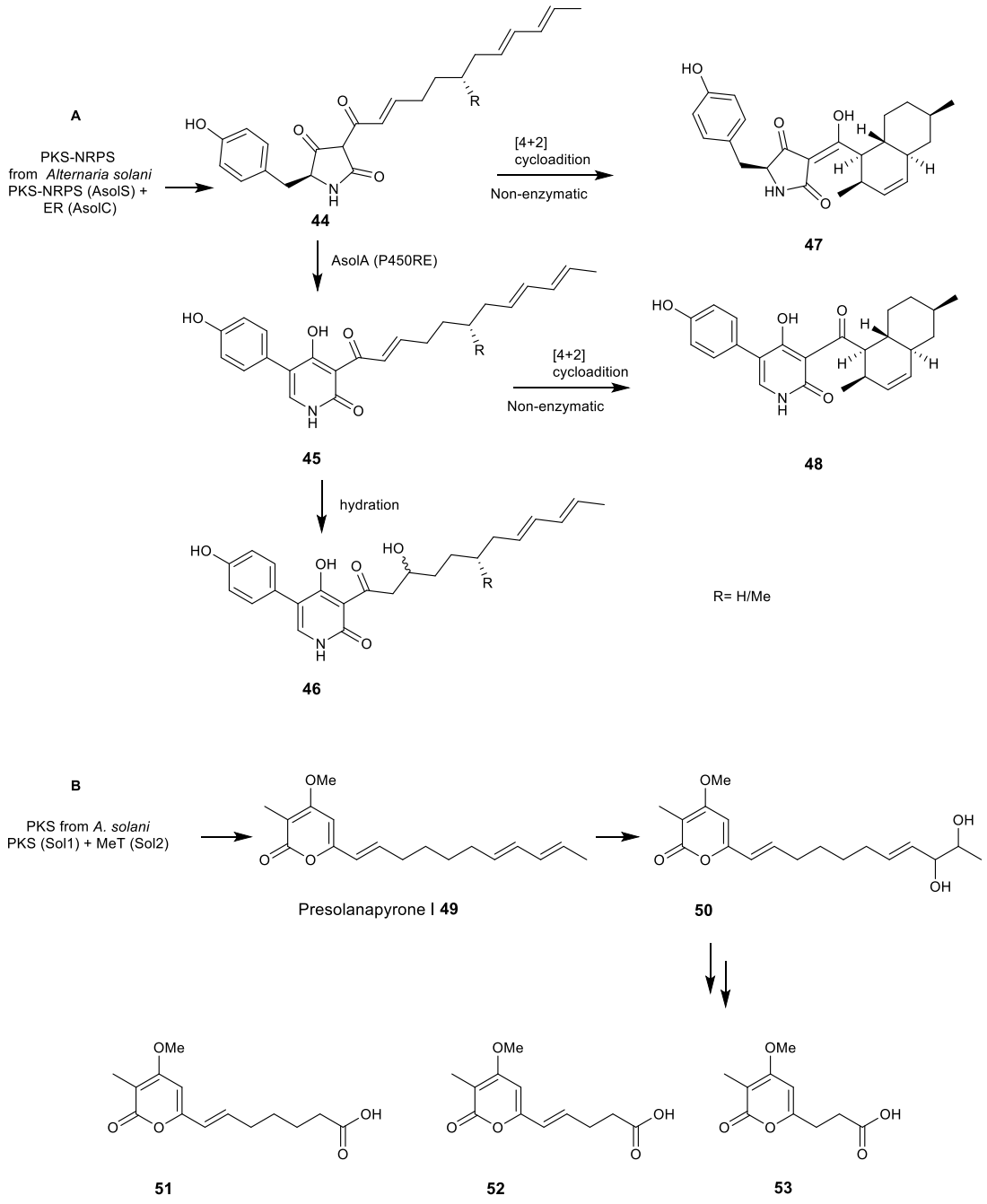
Scheme 2.1 Products of Aspks1.

Another major limitation is the unwanted activity of the native catabolic enzymes in some fungal hosts which can lead to oxidation or degradation of the produced natural product. Examples include the oxidation of the ACE1 product in *A. oryzae* after heterologous expression of the PKS-NRPS and ER genes. In this example, reduction of the expected aldehyde product **42** resulted in production of the shunt product **43**, which was not accepted by the later tailoring genes (Scheme 2.2).[78,80]



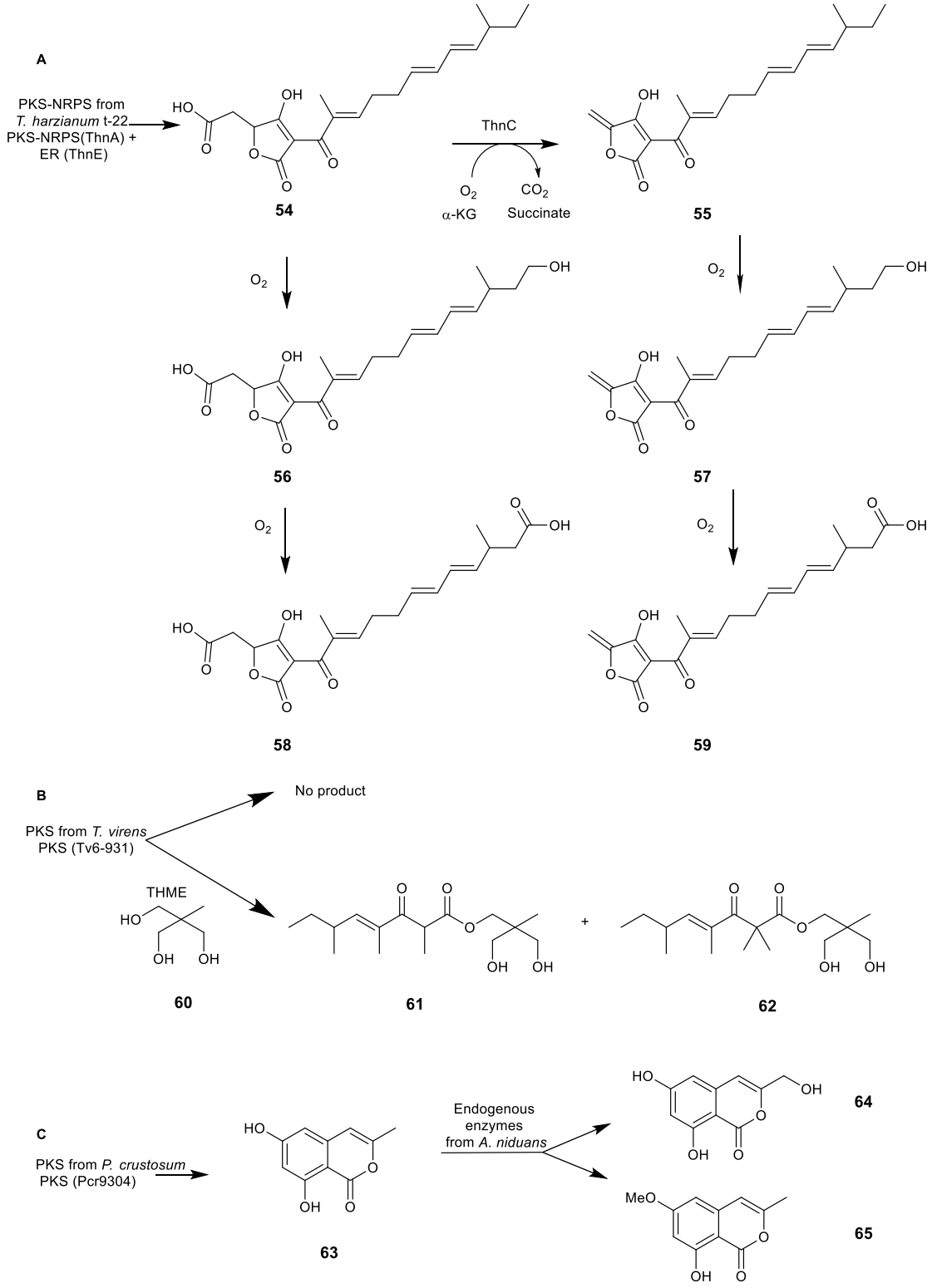
Scheme 2.2 The oxidized product of *ACE1* gene after heterologous expression in *A. oryzae*.

Other less drastic examples of oxidation or hydration by the host were also observed in *A. oryzae*, where heterologous expression of the megasynthases PKS-NRPS (AsolS) with ER (AsolC) and the PKS (Sol1) with MeT (Sol2) from *Alternaria solani* resulted in the production of the expected products **44**-**45** and **49** together with the modified structures **46**-**48** and **50**-**53** (Scheme 2.3).



Scheme 2.3 Shunt products observed during heterologous expression of some megasynthases in the heterologous host *A. oryzae.*[81,82]

A. nidulans has also shown some limitations, especially when used with BGC from *Trichoderma* spp.

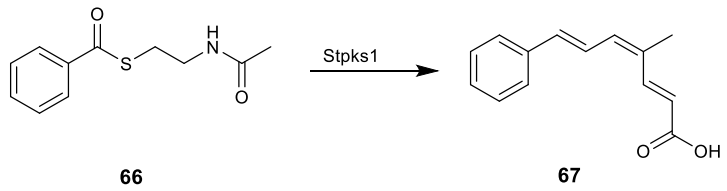


Scheme 2.4 Some limitations of the heterologous host *A. nidulans*[74,83,84]

For example, expression of a cryptic PKS-NRPS from *T. harzianum* t-22 (renamed *T. afroharzianum* t-22)[74] with its *tran*s acting ER and the non-heme iron and α-KG dependant oxygenase in *A. nidulans* resulted in the production of **54** and **55** in addition to their hydroxylated (**56- 57**) and carboxylated (**58**-**59**) analogues under aerobic culture conditions, which was attributed to the activities of endogenous cytochrome P450 oxygenases from *A. nidulans* (Scheme 2.4-A).[74]

In another example, expression of the PKS TV6-931 from *T. virens* in *A. nidulans* did not yield any new product. This was discovered to be due to the absence of the proper releasing substrate. Repeating the heterologous expression with the addition of a releasing substrate **60** resulted in the production of the two new compounds **61**-**62** (Scheme 2.4-B).[84] Furthermore, the expression of the PKS (Pcr9304) from *Penicillium crustosum* in *A. nidulans* resulted in the production of the expected isocoumarin product **63** together with some other oxidated and methylated products **64**-**65** (scheme 2.4-C).[83]

Another challenge in finding a suitable host is the absence of certain primary metabolites that may be required as starter or extender units for the biosynthesis. For example, expressing the first gene from the strobilurin biosynthesis pathway (*stpks1*) in *A. oryzae* was only successful in the presence of benyzoyl-SNAC **66** or after adding the genes responsible for benzoyl-CoA production (scheme 2.5). Therefore, there is still a need for developing new filamentous fungal heterologous hosts to overcome the limitations of some of the available hosts. [76,78,81]



66 **67**

Scheme 2.5 Expression of strobilurin PKS (Stpks1) in *A. oryzae* and feeding benzoyl-SNAC **66**.[18]

2.1.2 Developing Fungal Heterologous Hosts.

Developing a fungal heterologous host is a challenge due to the fact that, unlike *E. coli* and yeast, fungal hosts have the ability to produce numerous secondary metabolites of their own. These secondary metabolites can interfere with the heterologous expression either by complicating the detection and purification of the target compound or by competing with the heterologously expressed gene for building blocks.[85,86] This has been solved in most of the available hosts simply by deleting whole gene clusters or

megasynthases to create hosts with cleaner secondary metabolic backgrounds. For example in developing *A. nidulans* as a host, Chiang *et al*.[87] deleted the whole BGC for asperfuranone **68** to create a strain with cleaner background. In developing *Penicillium rubens* as a platform for secondary metabolite production, Pohl and Polli *et. al*.[86] knocked out the four highly expressed BGC in *P. rubens*, namely those responsible for the biosynthesis of penicillin **7**, roquefortine **69**, chrysogine **70** and fungisporin **71**. *A. oryzae* is a popular host strain because it is a 'domesticated' relative of *Aspergillus flavus* that has been develolped by unwitting human selection over a millenium for food processing. Presumably, randomly mutated strains were preferentially selected for their inability to produce toxins and bitter flavour compounds, resulting in the RIB40 strain known today that grows rapidly but produces no significant secondary metabolites under most growth conditions.[76]

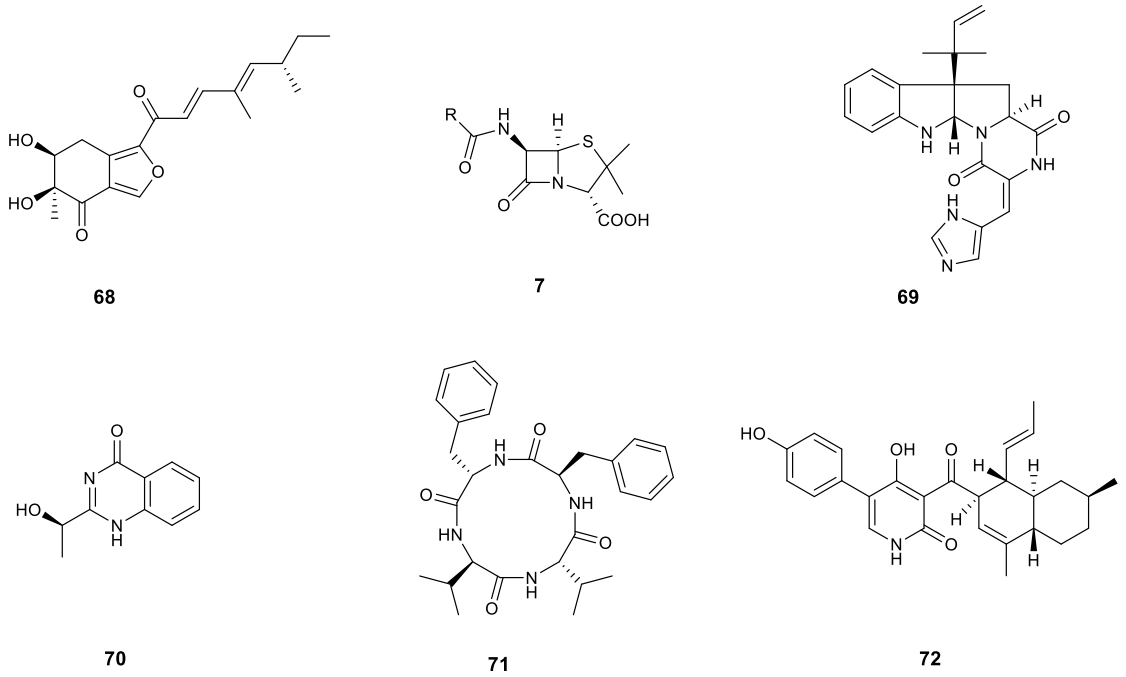


Figure 2.2 Examples of some fungal natural products.

Transformation of filamentous fungi is inefficient when compared to that of yeast and *E. coli* mainly due to the presence of fungal cell wall and the lack of natural plasmids.[58] Therefore, another important challenge with developing a fungal host is developing a transformation protocol and a toolkit to facilitate the heterologous expression of multiple genes in the designated host. This includes choosing efficient promoters, methods to insert the coding regions into the expression cassette and developing strains with appropriate selection markers and vectors with the complementing genes.[76,85] Some

successful examples have been reported, where entire gene clusters have been transferred *en masse* with its native promoters into the heterologous host without the need for choosing strong promoters suitable to the host. These examples include the expression of penicillin **7** BGC in *Neurospora crassa*[88] and ilicicolin-H **72** BGC in *A. nidulans*.[89] However, the use of native promoters is not optimal as they are usually relatively weak and sometimes are only expressed under certain conditions and therefore, the yield can be low.[90] Hence, replacing the native promoters with strong ones that will function in the heterologous host is usually more effective.[76]

The presence of specific selection-marker genes is an important requirement for efficient transformation and selection of the positive transformants. A variety of marker genes are available for different filamentous fungi. However, the use of these markers usually depends on the availability of compatible host strains that either have inactive genes or lack specific genes to allow for successful selection. Selection markers are usually categorized as *auxotrophic* or *dominant*. While auxotrophic markers rely on the existence of an auxotrophic strain with an inactive gene that can be complemented by the marker gene, dominant markers usually impart antibiotic resistance and therefore do not necessarily need a starting mutant strain.[91] Examples of dominant markers include the *E. coli* hygromycin B resistance gene *hph*[92] and the phleomycin resistance gene *ble*[93] from *Streptoalloteichus hindustanus*.[91] Many auxotrophic markers have been developed for the well-studied model organisms, *Neurospora crassa*, *A. nidulans* and *A. oryzae* such as *pyrG and argB* selection markers.[94]

Construction of most fungal natural product biosynthetic pathways have been performed in *Aspergillus* hosts[76] and therefore the development of *Aspergillus oryzae* as an efficient heterologous host will be discussed in more details to explain the required steps to develop a fungal heterologous host. Several auxotrophic strains of *A. oryzae* with their corresponding selectable markers were developed over the years such as the arginine (*argB*) auxotroph *A. oryzae M-2-3* that was developed by UV irradiation and can be complemented with *argB* from *A. nidulans*.[95] Similarly, nitrate assimilation mutants are known that can be complemented with the *A. oryzae niaD* gene.[58,96] Many vectors have been developed to facilitate the efficient expression of genes in *A. oryzae* using these selection markers.

One of the earliest developed vectors for heterologous gene expression in *A. oryzae* is pTAex3 that consists of the *amyB* expression cassette (promoter P_{amyB} and terminator T_{amyB}) together with the *A. nidulans argB* selectable marker.[97] This vector has

been used in combination with the *argB* auxotrophic strain *A. oryzae* M-2-3. The vector also contains genes for replication (*ColE1*) and selection (ampicillin resistance, *AmpR*) in *E. coli* to allow for plasmid amplification.

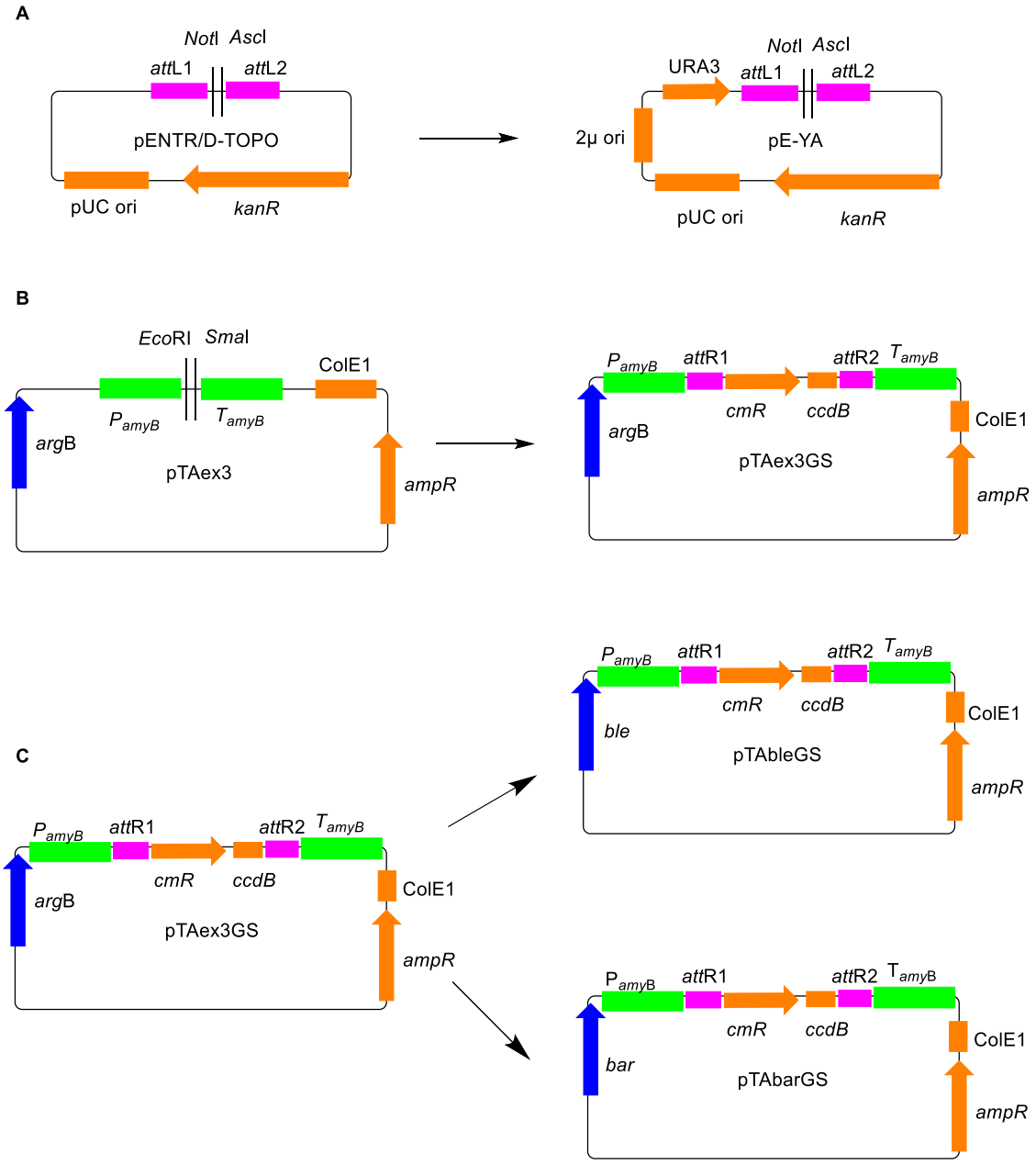


Figure 2.3 Construction of vectors used for *A. oryzae* transformation: **A**, construction of pEYA vector from the Gateway entry plasmid pENTR/D-TOPO; **B**, construction of pTAex3GS vector from pTAex3 plasmid; **C**, construction of pTAbarGS and pTAbleGS vectors from pTAex3GS.

The gene of interest is integrated in the expression site by ligation of the coding region at unique *Eco*RI and/or *Sma*I recognition sites located between P_{amyB} and T_{amyB}. Although the vector showed success in expression of fungal natural product genes, it was still quite

difficult to express large genes by this system as it required several steps of DNA fragment isolation and ligation to assemble the coding region.[58,76]

This deficiency prompted the introduction of the Gateway *in vitro* recombination system for gene transfer into the expression site of pTAex3. This was done by adding a Gateway® destination cassette, which includes the chloramphenicol resistance gene (*cmR*)[98] and the toxic *ccdB* gene flanked by *att*R1 and *att*R2 sites (Invitrogen) between *amyB* promoter and terminator.[58] The resulting vector is named pTAex3GS·*arg* (Figure 2.3-B). Additionally, the Gateway entry vector (pENTR/D-TOPO) was converted into a yeast-*E. coli* shuttle vector by insertion of the yeast origin of replication (*2μ* ori) and selection marker *URA3*[99] to allow amplification and selection of the plasmid in yeast, since it allows the successful yeast transformants to grow in the absence of uracil. The resultant plasmid was then named pEYA (Figure 2.3-A). This facilitated the assembly of megasynthases coding regions from fragments of less-than full length by yeast homologous recombination. The full length megasynthase can then be easily transferred from the entry vector pEYA to the expression vector pTAex3GS·*arg* by *in vitro* LR (Gateway®) recombination. Another two variants of pTAex3GS plasmid were constructed by replacing the *argB* marker with either phleomycin (*ble*) or glufosinate ammonium (*bar*)[100] resistance markers. These plasmids are named, pTAGS·*ble* and pTAGS·*bar*, respectively (Figure 2.3-C).

One of the first biosynthetic gene clusters that was elucidated in *A. oryzae* using pEYA and pTAex3GS vectors was the tenellin BGC. Tenellin **73** is a 2-pyridone produced by *Beauveria bassiana*. The tenellin BGC encodes a hybrid PKS-NRPS (TenS), a *trans*-acting enoyl reductase (ER, TenC) and two cytochrome P450 monooxygenasess (TenA and TenB, Figure 2.5-A). The coding sequence of *tenS* was assembled in pEYA by yeast homologous recombination, then it was transferred to the expression vector pTAex3GS·*arg* by Gateway® LR recombination. The resultant vector pArgB·*tenS* was then used to transform *A. oryzae*-M-2-3 and the transformants showed production of compounds **74** - **76** related to tenellin (Figure 2.4).[58,76,101]

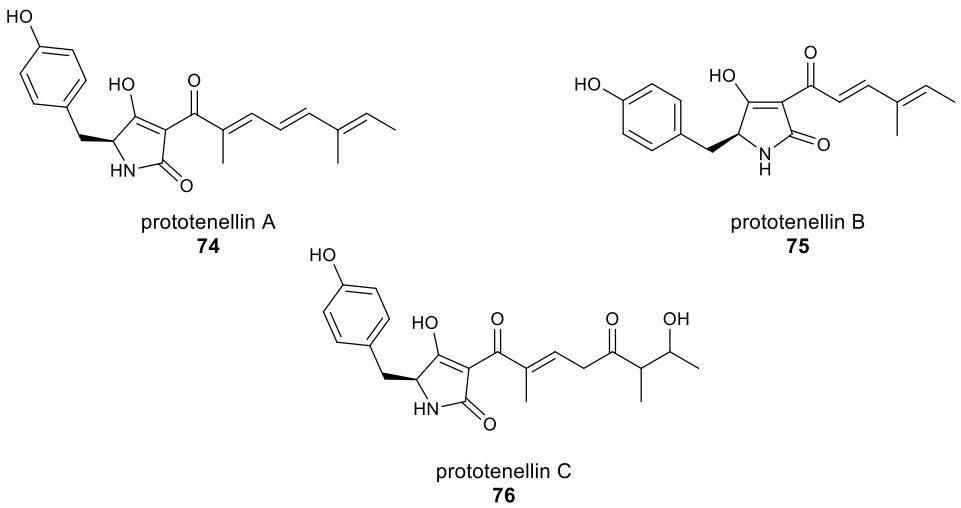


Figure 2.4 Products of expression of *tenS* gene alone in *A. oryzae* M-2-3.[101]

To restore the correct programming of *tenS*, *tenC* (encoding the trans-ER) was cloned into pENTR/D-TOPO and transferred to pTAGS·*bar* by Gateway® LR recombination to yield the vector pBar·*ten*C, which was transformed into the *A. oryzae* M-2-3 *tenS* transformant. Co-expression of *tenS* and *tenC* resulted in the production of the first authentic precursor of tenellin, pretenellin A **77**. The remaining two genes in the cluster, *tenA* and *tenB*, were cloned individually into the entry vectors and transferred to pTAGS·*ble* plasmids to yield pBle·*tenA* and pBle·*tenB* plasmids, respectively (Figure 2.5-B). The expression plasmids were then used to transform the *A. oryzae* M-2-3 TenS+TenC transformant. Expression of *tenB* with *tenS* and *tenC* resulted in no new product, however, expression of *tenA* with *tenS* and *tenC* resulted in the production of pretenellin B **78**, which is the result of ring expansion of pretenellin A. With the absence of more selection markers, the coding region of *tenB* together with P_{amyB} and T_{amyB} was amplified from pBle·*tenB* plasmid with overhang and ligated into pBle·*tenA* by yeast homologous recombination. The resultant vector pBle·*tenA*·*tenB* was then used to transform *A. oryzae* M-2-3 TenS + TenC transformant and the resultant transformants were able to produce tenellin **73** as the final product of the BGC. Heterolgous expression of tenellin BGC in *A. oryzae* M-2-3 resulted in the elucidation of the biosynthetic pathway of tenellin (Figure 2.5-C).[76,101,102]

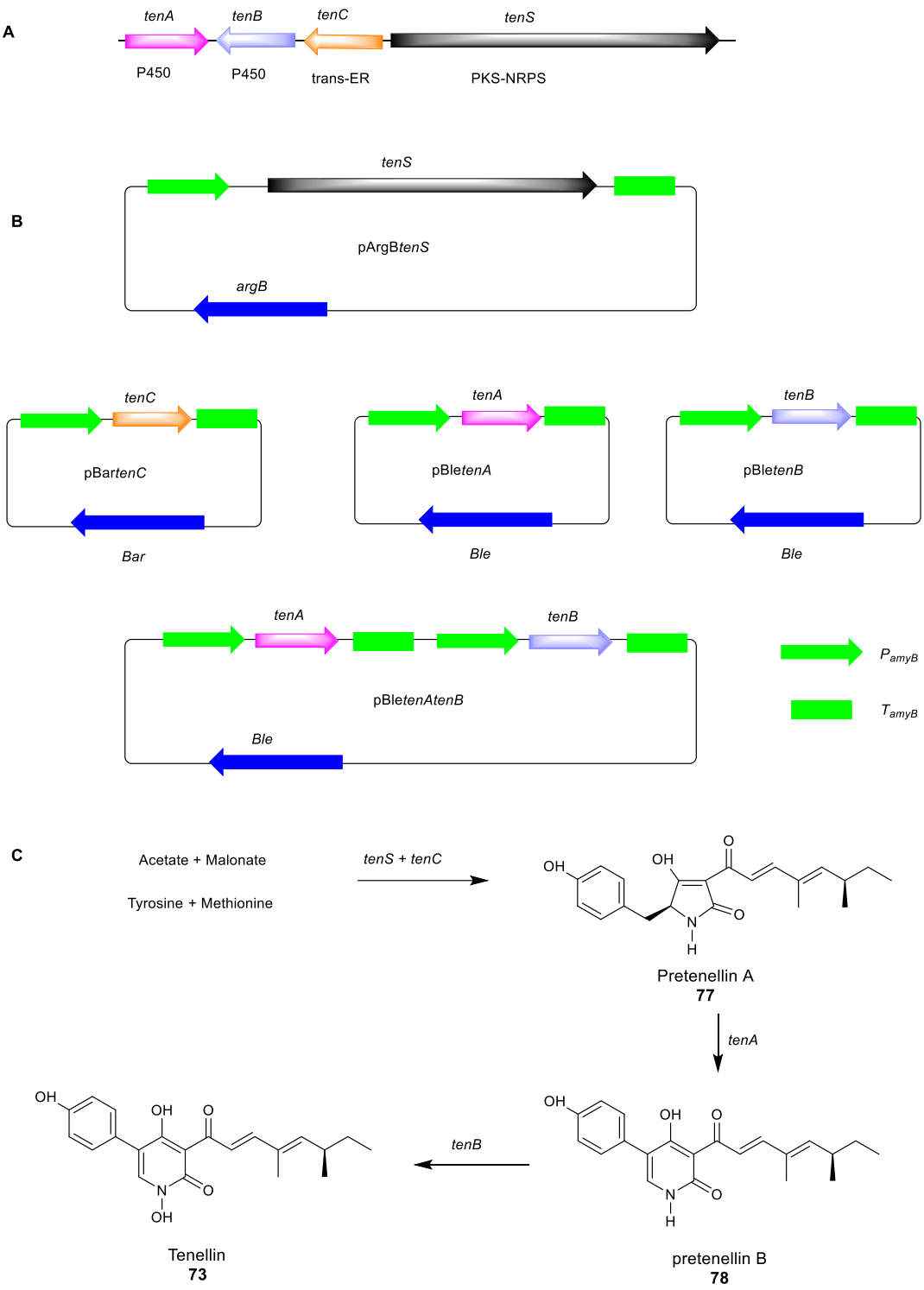


Figure 2.5 Expression of tenellin pathway in *A. oryzae* M-2-3: **A**, tenellin biosynthetic gene cluster; **B**, plasmids used in the heterologous expression of tenellin BGC in *A. oryzae* M-2-3; **C**, tenellin **73** biosynthetic pathway.[102]

Due to the limited availability of selection markers in *A. oryzae* M-2-3, new vectors for multiple gene expression were developed. First, pTAexGS·*arg* plasmid was converted into a yeast-*E. coli* shuttle vector by insertion of *2µ* ori and *URA3* to allow amplification and selection of the plasmid in yeast and the resultant vector was named pTAYAGS·*arg*.

Next, three different promoters from *A. nidulans* were chosen to be tested in the new system, namely P_{gpdA}, P_{trpC} and $P_{oliC.}$ However, only P_{gpdA} was proven to be strong enough in the expression of *eGFP* in *A. oryzae*. Therefore, another two promoters from *A. oryzae* were chosen based on *A. oryzae* EST analysis, namely promoters of alcohol dehydrogenase (P_{adh}) and enolase genes (P_{eno}). The *A. oryzae* promoters were assumed to consist of sequences up to around 500 bp upstream of the coding region of the respective genes. Using yeast homologous recombination, the three promoters P_{gpdA}, P_{adh} and P_{eno}, were added to the plasmid pTAYAGS with an *Asc*I restriction site downstream of each promoter. The resultant vector was named pTAYAGS·*arg*·3P (Figure 2.6-A) and was used to introduce the four genes of the tenellin pathway into *A. oryzae* in a single transformation step. Variants of the plasmid were made by replacing the *argB* gene with *ble* and *bar* genes to construct the vectors, pTAYAGS·*ble*·3P and pTAYAGS·*bar*·3P, respectively.[76,103]

Further development of the multigene expression vectors resulted in the addition of terminators in all expression sites (plasmid pTYGS·*argB*). The quadruply auxotrophic strain *A. oryzae* NSAR1[104] was later developed with four different auxotrophic selection markers (*argB, adeA, niaD* and *sC*).[76,104] The pTYGS·*argB* plasmid was then adjusted by replacing the *argB* selection gene with the three other auxotrophic marker genes (*niaD*, *adeA, sC*, Figure 2.6-B) in addition to the other two pTYGS vectors using the two dominant markers of *A. oryzae* (*bar* and *ble*). These combined systems allow the expression of up to 16 genes simultaneously in *A. oryzae* NSAR1 strain using auxotrophic selection alone, and up to 24 genes if the *ble* and *bar* dominant selection markers are used in addition.[76]

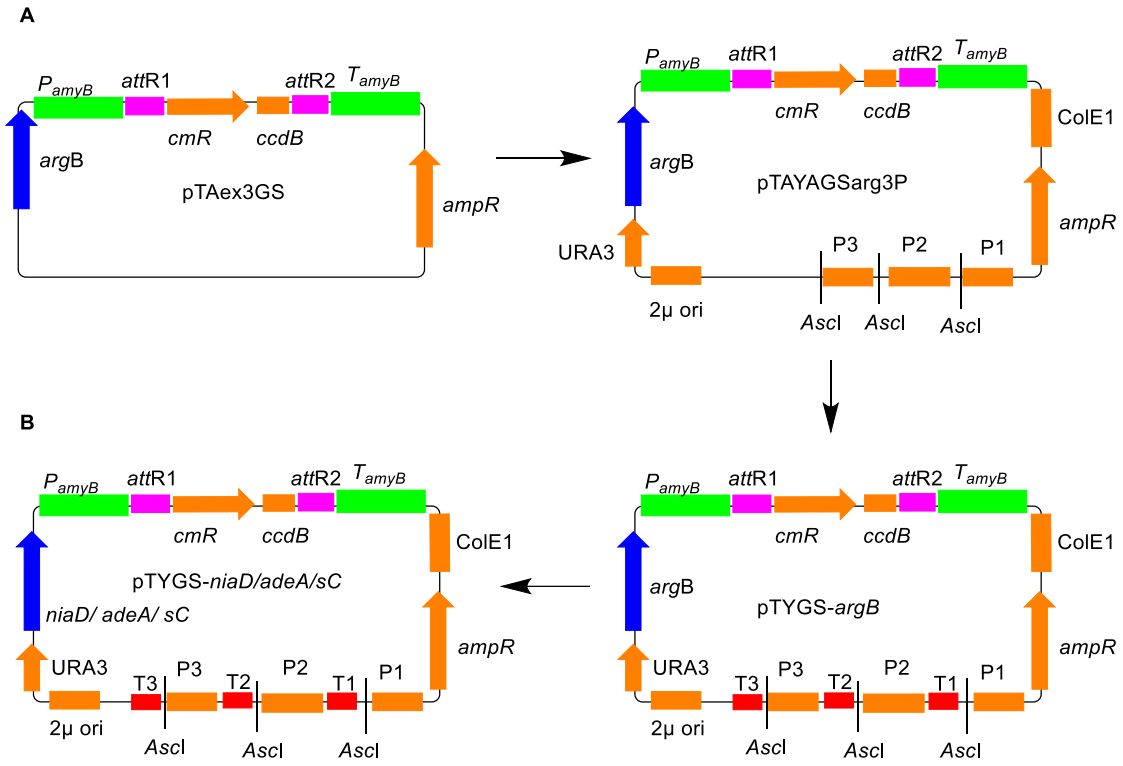


Figure 2.6 Construction of multigene expression plasmids for the heterologous expression in *A. oryzae*.

2.1.3 *Trichoderma reesei* as a Heterologous Host

Members of the fungal genus *Trichoderma* are ubiquitous inhabitants of soils, decaying wood, and plant debris.[59] Due to their metabolic diversity, high reproductive capacity and competitive capabilities in nature, they can survive in many different geographical habitats.[44,59,60] *Trichoderma reesei* (*Hypocrea jecorina*), the biotechnological workhorse of the genus, has a significant role in industry due to its production of large amounts of extracellular enzymes, especially cellulases and hemicellulases. These polymer-degrading enzymes are important industrial products, especially with respect to production of second-generation biofuels from cellulosic waste.[105] *T. reesei* is able to produce all the enzymes required for extensive degradation of crystalline cellulose and is probably the most studied fungus among the cellulase-producing organisms.[106]

The US Army *Quarter Master Research and development center at Natick, Massachusetts* initiated a research program during the Second World War in an attempt to identify the organisms responsible for the rapid deterioration of textiles in the tropical regions and to determine the mechanisms by which they operate. The research has resulted in the discovery of *T. reesei QM6a, where "QM" is derived from Quarter Master. This fungus* was detected to be the decomposing agent of U.S. Army cotton

fabrics. The strain QM6a was originally classified as *Trichoderma viride*, but was later recognized as a separate species and named according to the principal investigator Elwyn T. Reese.[107–112]

Soon after their discovery it was realized that the cellulolytic action of *T. reesei* is due to the secretion of cellulose and hemicellulose-degrading enzymes and these enzymes have attracted biotechnological interest ever since. This ability also allows *T. reesei* to grow on low-value substrates such as soybean hulls,[113] wheat straw,[72] *etc*. Various attempts to increase cellulase production by *T. reesei* have been achieved: by optimizing culture conditions; by addition of surfactants to the medium; and by mutating the fungus. Systemic mutation has resulted in the cellulase-hyperproducing strain *T. reesei* Rut-C30[112] among others. The original isolate of QM6a is the parent of all commercial *T. reesei* strains used today and due to their industrial applications, the genetic toolkit for this fungus is the most extensive of the genus.[107,108,112,114,115]

Now, *Trichoderma reesei* is an attractive host for heterologous and homologous protein expression, mainly due to its excellent capabilities in protein production which for some *T. reesei* strains has reached up to 100 $g \cdot L^{-1}$ in industrial fermentation conditions.[116] *T. reesei* has many advantages that makes it a potential host for heterologous expression of proteins: the presence of strong promoters such as P_{cbhI} that has been used for the successful production of heterologous proteins of fungal and mammalian origins; its safe use in the production of several enzymes used in food and pharmaceutical products (GRAS); and its good secretion capacity carrying out 'mammalian-like' protein modifications.[111,117] *T. reesei* also seems to have a low propensity for the production of mycotoxins, which is a great advantage that allows its safe use in research and industry.[118] Advantages of *T. reesei* as a host for heterologous expression is summarized in table 2.1.

Although *T. reesei* has been extensively developed as a host for protein production, we are aware of no reports of its use for the heterologous production of secondary metabolites.

Table 2.1 Advantages of *T. reesei* as a host for heterologous expression

Historical	• Generally recognised as safe (GRAS) • Widely used in protein production.
In Biotechnology and molecular biology	▪ good secretion capacity up to 100 $g.L^{-1}$ ▪ Can carry out 'mammalian-like' protein modifications ▪ Contains strong promoters such as *cbhI* that can be used in heterologous expression of proteins. ▪ Secretion of biomass degrading enzymes. ▪ Can grow on cheap carbon sources. ▪ Fully sequenced genome. ▪ Efficient transformation system and some molecular tools available.

2.1.4 Biotechnological Toolbox for *Trichoderma reesei*

Due to all these advantages, *T. reesei* has been used extensively as a host for heterologous and homologous protein expression and therefore substantial research have been done on *T. reesei* as a host organism. Several host strains have been developed with useful traits for protein expression and promoters have been identified and described for *Trichoderma reesei*. The following sections summarize the currently available genetic tools for metabolic engineering of *T. reesei*.

2.1.4.1 Available *T. reesei* Strains

Over the years, many strains for *T. reesei* have been developed that showed better characteristics to facilitate protein production. These characteristics include higher cellulase production, high secretion capabilities and lower production of proteases.[112] The most common method for strain development has been random mutagenesis by UV radiation or chemical mutagenesis followed by strain selection and screening for the advantageous characteristics.[119] In addition, many auxotrophic and dominant markers have been tested in *T. reesei* to facilitate gene insertion and heterologous expression. These includes *argB,*[106] *pyr4,*[120] *pyr2,*[121] *asl1,*[120] *amdS,*[106] hygromycin (*hph*)[122] and

bleomycin (*ble*)[123] resistance. However, only few of these markers have been proven to be promising in heterologous expression, for example *amdS*, *pyr4*, *hph*. Table 2.2 summarizes some of the developed strains for *T. reesei* and their selection marker.

Table 2.2 Different *T. reesei* selection markers available for heterologous protein expression.

No.	Selection marker	Abb.	Remarks	Ref
1	Hygromycin B resistance	*hph*	Very effective selection marker and mostly used in *T. reesei.* Antibiotic resistance	[122]
2	benomyl resistance	*ben*	Antibiotic resistance	[124]
3	Bleomycin/ phleomycin resistance	*ble*	• Antibiotic resistance • Singh et al.[125] reported that *T. reesei* strains showed poor selection on Zeocin, the most common form of phleomycin.	[123,126]
4	Phosphinothricin resistance	*bar*	Antibiotic resistance	[122]
5	Pyrithiamine resistance	*ptrA*	Dominant marker, resistance to pyrithiamine, which is the analog of thiamine that prevent the growth of microorganisms requiring intact thiamine for growth.	[127]
6	Invertase (sucrose utilization)	*sucI*	Dominant nutritional marker	[128]
7	Ornithine transcarbamylase	*argB*	• Auxotrophic marker (arginine). Not commonly used in *T. reesei*	[106]
8	Orotate phosphoribosyl transferase	*pyr2*	• Auxotrophic marker (Uridine). Bidirectionally selection marker.	[121]
9	Orotidine-5'-phosphate decarboxylase	*pyr4*	• Auxotrophic marker (Uridine). Bidirectionally selection marker.	[122,129]
10	Acetamidase (Acetamide utilization)	*amdS*	Bidirectionally selection marker	[106]
11	Arginiosuccinate lyase	*asl1*	Auxotrophic marker (Arginine).	[120]
12	Hexokinase	*hxk1*	Not commonly used in transformations.	[123]
13	Mitochondrial homoaconitate hydratase	*hah1*	The strains could grow on minimal media and did not exhibit a phenotype that can be easily distinguished from the parental strain.[120]	[120]

2.1.4.2 Promoter Toolbox for *T. reesei*

Many promoters for *T. reesei* have been identified and tested for their activity in protein expression in *T. reesei*. Examples include Promoters of pyruvate decarboxylase (P_{pdc}), cellobiohydrolase I (P_{cbhI}) and many other genes (table 2.3). These promoters are divided based on their activity into constitutive and tunable promoters. Constitutive promoters are

active under almost all cultivation conditions while tunable promoters can be induced or repressed by the presence of certain molecules in the media. The main advantage of constitutive promoters is the persistent expression of the protein regardless of the media composition. However, the expression levels are much less than that of other tunable promoters in *T. reesei* such as P_{cbhI}. The tunable promoters offer the advantage of controlling the production level by the addition of inducer or repressor molecules in the media. The best tunable promoters in *T. reesei* are the promoters for cellulose production such as CBHI (also named *cel7a* promoter) and CBH2 (also named *cel6a*).[130]

Table 2.3 Different promoters identified and tested for *T. reesei.*[130]

No.	Promoter type	Promoter	abbreviation	activity	Ref
1	Constitutive promoters	Pyruvate decarboxylase	P_{pdc}	• Strong promoter • Activity increase on D-glucose	116
2		Glyceraldehyde dehydrogenase	P_{gpd1}	• Activity stable on D-glucose	116
3		Enolase	P_{eno1}	• Strong promoter • Activity increase on D-glucose	116
4		cDNA1 (Gene with unknown function)	P_{cDNA1}	The strongest constitutive promoter in *T. reesei*	131
5		Pyruvate kinase	P_{Pki}	Medium strength promoter	132
6		Transcription elongation factor 1α	P_{tef1}	Medium strength promoter	131
7	Tunable promoters	Cellobiohydrolase CBH1 /CEL7A	P_{Cbh1}/ P_{cel7a}	• Strong promoters • Induced by cellulose, sophorose and lactose • Repressed by D-glucose	133
8		Cellobiohydrolase CBH2 /CEL6A	P_{Cbh2} / P_{Cel6a}		134
9		Endoglucanase EG2 / CEL5A	P_{Eg2} / P_{egl2} / P_{Cel5a}		134
10		Endoglucanase EG3 / CEL12A	P_{Eg3} / P_{egl3} /P_{Cel12a}		133

2.1.4.3 Transformation Methods and DNA Integration in *T. reesei*

Many transformation methods have been developed and used in *T. reesei* such as Agrobacterium-mediated transformation, PEG-mediated transformation of protoplasts, electroporation and particle bombardment.[120]

Cotransformation of two plasmids, one bearing the gene of interest, while the other bears the marker gene was commonly used in *T. reesei* with almost 80 % success rates although successful integration of the marker gene does not ensure the integration of the gene of interest. This necessitates laborious investigation of the gDNA of the transformants to ensure the insertion of the gene of interest in the selected transformants. Therefore, many new techniques have been developed for gene insertion into the gDNA of *T. reesei*. This includes traditional methods such as transformation of the plasmid bearing the gene of interest and the selection marker and its random or site-specific integration into gDNA.[120] Also more sophisticated and advanced methods for gene integration such as CRISPR/Cas9 and TALEs (transcription activator-like effector) methods have been developed for *T. reesei*.[135,136]

A prerequisite for the site-specific integration of DNA into the fungal chromosomal DNA is the high efficiency of homologous integration, which was achieved in different fungal strains by deletion of genes involved in non-homologous end-joining mechanism (NHEJ). For example, in *Neurospora crassa* knocking out of genes involved in NHEJ *mus-51, mus-52* and *mus-53* (homologous to human *KU70, KU80* and *Lig4*, respectively) resulted in efficient homologous integration with rates up to 100 % depending on the site of integration.[120,122] Attempts to knock out the genes of NHEJ in *T. reesei* QM6a also resulted in strains with highly efficient homologous integration such as *T. reesei* QM6a Δ*tku70* Δ*pyr2* and *T. reesei* QM6a Δ*tmus53*.[121,122] Derntl *et al*.[120] constructed two new auxotrophic strains of *T. reesei* in the QM6a Δ*tmus53* strain by knocking out the arginine-succinate lyase (*asl1*) and the orotidine-5`-phosphate decarboxylase (*pyr4*) genes in *T. reesei* QM6a Δ*tmus53*. The two new strains named *T. reesei* QM6a Δ*tmus53* Δ*asl1 and T. reesei* QM6a Δ*tmus53* Δ*pyr4* showed high rates of positive transformation reaching 100 % and 86 %, respectively.

2.1.5 Waste Valorization

Enormous amount of waste is produced annually with an estimated 2 billion tonnes of municipal solid waste and over 170 billion metric tons of biomass that could be harvested annually worldwide.[137,138] Biomass consists of three types of polymers; cellulose (30 - 40 wt%), hemicellulose (20 - 30 wt%) and lignin (10 - 25 wt%).[138] Many conventional methods for treatment and management of waste have been developed, which includes animal feeding, incineration, landfill disposal, composting and anaerobic digestion.[139] However, all these conventional techniques have shown major drawbacks to the economy and the environment.[140] For example, disposal of the waste in landfills results in many environmental issues such as greenhouse gas emission, leachate formation and odour production.[139] Whereas incineration of waste causes air pollution and is an inefficient and energy demanding process.[139] Although anaerobic digestion and fermentation are more environmentally friendly approaches for waste management, they usually result in converting organic materials to low-value added products. This results in losing the complexity of the organic materials that needs to be restored later on to produce compounds and products of economic value. Therefore, new technologies and techniques have been developed to convert waste to value-added chemicals, polymers, enzymes, fuels and energy.[139] The process of converting waste material to more useful products such as fuels and chemicals is usually called waste valorization.[141]

Lately, waste valorization has gained much attention in both the scientific and public opinion. A variety of novel techniques and strategies to extract and recover high-value added compounds from waste and to produce fuels and other chemicals using waste materials have been published. Examples include extraction of bioactive compounds and nutrients from different vegetable and fruit waste such as phenolic compounds from citrus peel residues and berries. Also, the production of enzymes such as cellulases, lipases and glucoamylases by solid state fermentation of food waste by fungal strains and bioethanol production by microbial bioconversion.[139]

Nevertheless, all these techniques still face some challenges. For example, conversion of lignocellulose to ethanol is still a complex and expensive process. Therefore the price of the cellulosic ethanol cannot compete with that of the fossil-derived fuels.[142] Also, using paper and lignocellulosic waste for solid state fermentation by fungal or microbial strains to obtain biofuels usually requires pre-treatment to convert the insoluble lignin and cellulose into simple sugars that can be processed by the microbial strains. The conventional pretreatment step is either by treatment with mineral acids or

enzymes to breakdown the lignocellulosic waste. Treatment with mineral acids results in copious amounts of inorganic salts as a waste, while treating the waste with enzymes is limited by the relatively high cost of the enzymes, making the process less cost-effective.[138,139]

Many pretreatment methods have been developed to overcome the barriers for valorization of lignocellulosic biomass imposed by the structural complexity and recalcitrance of biomass. These methods include biological, chemical, physical and physico-chemical processes such as hammer milling, fluid energy milling, colloid milling and alkali treatment.[138,143]

2.1.6 Waste Valorization by *T. reesei*

T. reesei is a saprophytic wood-decaying filamentous fungus that is well-known for its high production capacity of cellulases.[144] Due to this high cellulase secretion capacity, *T. reesei* has long been used for the industrial production of cellulolytic enzyme cocktails that are used for de-inking, pulp and paper refining, in textile industry and also in biofuel production.[145] Using *T. reesei*, the researchers in the US army research center in Natick, Mary Mandels and Elwyn Reese, aspired to developing a commercial process to convert waste cellulose to glucose. This idea practically made them the pioneers in the area of using biological processes to convert cellulosic biomass to fuel and chemicals.[143] Their main idea was to convert cellulose, which is the most abundant organic compound that is also present in waste from forests products, agriculture, and vegetables and fruits processing, into glucose that can be used later in the production of fuel alcohols or can be converted to proteins and chemicals by feeding them to appropriate microorganisms.[143] Also celluloses of various origins were tested to identify the best conditions for enhancing enzymatic conversion of waste cellulosic biomass into soluble sugars that could be fermented later on to ethanol or other liquid biofuel.[112,146–148] High levels of protein production has been attained and recently industrial-scale cellulosic ethanol plants have started to operate.[149]

In a recent attempt for the valorization of de-inking paper and slaughter house wastes, *T. reesei* was able to grow on Mandel's medium where paper sludge was used as carbon source and pasteurized blood as nitrogen source.[145] Another attempt to use *T. reesei* for textile waste valorization was also successful, however, pretreatment of the waste was always crucial in all these attempts to increase the availability of cellulose.[150]

Hence, *T. reesei* represents a very promising model organism for valorization of waste and biomass for greener and sustainable production of chemicals.[138,145,150]

2.2 Aims

For years, *T. reesei* has been used as an efficient host for heterologous protein expression, however, no reports on using it as a heterologous host for the production of secondary metabolites. Therefore, the aim of this work is to develop *T. reesei* as a heterologous expression platform for fungal natural products and to use the inherent ability of *T. reesei* to grow on waste materials to enable the direct conversion of waste materials to high value secondary metabolites.

The first aim would be testing the available plasmids, strains, promoters and transformation protocols for *T. reesei* for their usability. The plasmids available in the Cox group are mainly used for *A. oryzae* transformation and therefore contain *Aspergillus* promoters. Testing these *Aspergillus* promoters in *T. reesei* and comparing their activity to native *T. reesei* promoters would give an idea about the best promoters to be used later to construct new plasmids for *T. reesei*.

The second aim would be constructing a better heterologous host by knocking out the main BGC from *T. reesei* to develop a strain with a cleaner background to facilitate production and isolation of the expressed natural product. The sorbicillinoid BGC is the well-studied BGC in *T. reesei* that is involved in the production of the yellow-coloured pigments, sorbicillins (Section 3.1.3.) The production of this class of compounds usually hamper the downstream purification of secondary metabolites in industrial filamentous fungi and was therefore previously eliminated in some strains such as *T. reesei* and *A. chrysogenum*.[119,151] Previous attempts to knock out this gene cluster from *T. reesei* were successful but they were mainly done in the wild type strain.[152,153] Therefore, the aim here is to knockout this BGC in an auxotrophic strain to construct a strain with cleaner background that still have a selection marker for later insertion of other genes.

The next step would be construction of new vectors with native and constitutive promoters to facilitate the expression of multiple genes simultaneously in *T. reesei*. The new system can then be tested with megasynthases with well-established function to confirm the activity of the chosen promoters and constructed vectors. First, the aim is to construct a plasmid with one native promoter to test the applicability of *T. reesei* as a heterologous host for natural product biosynthesis. For testing the new system, the very-well studied methylorcinaldehyde synthase *(aspks1)* gene from *Acremonium strictum*

(renamed as *Sarocladium schorii*) will be used since it is one of the most studied PKS in our group.

With *T. reesei's* ability to grow on different waste materials, the new host could be exploited as a microbial cell factory to convert waste directly into useful natural products and chemicals. Therefore, another aim of this study is to test the ability of the new transformants of *T. reesei* to grow on different waste materials and their ability to produce the expected compounds in detectable amounts. If possible, quantification of the produced compounds on waste material would be helpful in judging the usefulness of the constructed system.

In a fifth aim, more native promoters could be added to the vector and the system could be tested for its ability to express more than one gene simultaneously, then testing the ability of the system to produce the expected natural product on waste materials. For testing this multigene expression vector, the very-well studied PKS-NRPS tenellin synthetase (TenS) together with its *trans*-acting ER (TenC) from *Beauveria bassiana* would be ideal since the tenellin pathway is one of the most-studied biosynthetic pathways in the group (Section 2.1.2).

Finally, the sixth aim would be testing the ability of the transformants produced in this work to grow and produce the expected compounds on lignocellulosic waste materials, such as barley or wheat straw. The reported global production of plant biomass amounts for about 200 * 10^9 tons annually and lignocellulosic materials represent over 90% of it.[154] Thus the ability of the produced fungal strains to grow and synthesize organic compounds on these types of waste material would represent not only a sustainable source for production, but also a greener alternative to chemical synthesis of such compounds.

2.3 Results

2.3.1 Testing Strains, Promoters and Plasmids Available in the Group

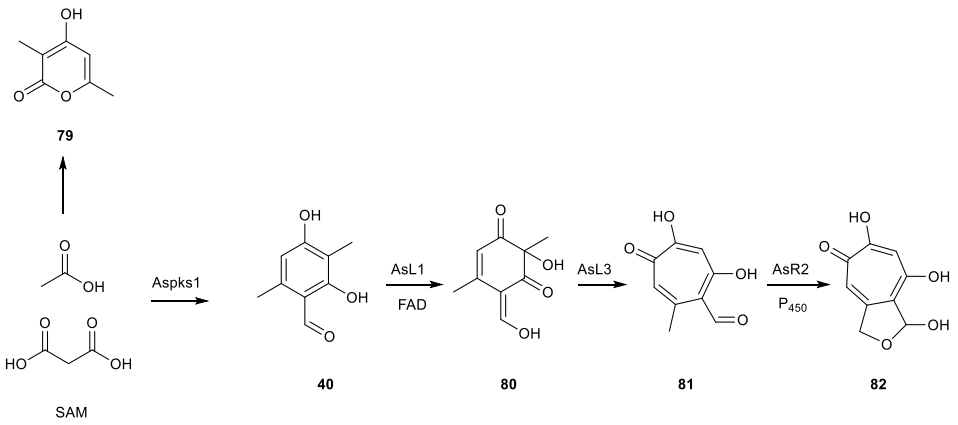
T. reesei strains used in this thesis were kindly donated by the Mach-Aigner group at the technical university of Vienna, Austria. The strains available and their selection markers are summarized in table 2.4.[120]

Table 2.4 Different *T. reesei* strains available in the group.

Strain	Phenotype		
	Hygromycin B	**5-FOA/ uridine**	**Arginine**
T. reesei* QM6a Δ*tmus53	-	-	-
T. reesei* QM6a Δ*tmus53* Δ*pyr4	-	R/A	-
T. reesei* QM6a Δ*tmus53* Δ*asl1	R	-	A

R, resistance; A, auxotrophy.

To establish *T. reesei* as a new heterologous host, the first step would be testing the system with well-established biosynthetic pathway. As a proof of concept, a well-known and tested megasynthase was used as a test to measure and establish the activity of the promoters and the expression host. Therefore, one of the most studied PKS in our group, methylorcinaldehyde synthase from *Acremonium strictum* (renamed as *Sarocladium schorii*) *aspks1*, was chosen to test the new system. In the group, a vector was already developed for *T. reesei* transformation with the native pyruvate decarboxylase promoter (P_{pdc}) and the *asl1* selection marker (pTYGS·*asl1*·*pdc*). This vector was used later on in the experiments to test the activity of the native P_{pdc} from *T. reesei*. The original plasmid used to express *aspks1* with the other three genes from the xenovulene pathway in *Aspergillus oryzae* (RSI96-1-6, pTYGS·*argB*·P_{amyB}·*aspks1*·*asL1*·*asL3*·*asR2*) used the *A. oryzae amyB* promoter (P_{amyB}).[155] It is not known if P_{amyB} is active in *T. reesei*. Vector RSI96-1-6 also contains another three genes from the xenovulene pathway (Figure 2.7) under the other three constitutive *Aspergillus* promoters and therefore could theoretically be used to test the activity of all the promoters from the pTYGS-*argB* plasmid in *T. reesei*. The expression of this plasmid in *A. oryzae* resulted in the elucidation of the first steps in the biosynthesis of the xenovulene pathway and in the production of compounds **79**-**82** (Scheme 2.6). Therefore, this plasmid was used to construct a new plasmid with suitable selection marker for *T. reesei* to test P_{amyB} as well as the other *Aspergillus* promoters P_{adh}, P_{eno} and P_{gpdA}.



Scheme 2.6 Tropolone biosynthesis in *A. oryzae* NSAR1 by expressing the first four genes in the *A. strictum aspks1* BGC.[155]

2.3.1.1 Plasmid Construction Using Yeast Homologous Recombination and LR Recombination

The two previously mentioned vectors, pTYGS·*asl1*·P_{pdc1} and RSI96-1-6 (pTYGS·*argB*·P_{amyB}·*aspks1*·*asL1*·*asL3*·*asR2*), were used to construct the two new vectors: MSIII69 (pTYGS·*asl1*·P_{pdc1}·*aspks1*) [1] and MSIII84 (pTYGS·*pyr4*·P_{amyB}·*aspks1*·*asL1*·*asL3*·*asR2*) with *aspks1* driven from P_{pdc1} or P_{amyB}, respectively (Figure 2.7). For the construction of MSIII84 (pTYGS·*pyr4*·P_{amyB}·*aspks1*·*asL1*·*asL3*·*asR2*), the original vector RSI96-1-6 (pTYGS·*argB*·P_{amyB}·*aspks1*·*asL1*·*asL3*·*asR2*, Dr. Raissa Schor, Cox group) with the *aspks1* megasynthase under the control of P_{amyB} and the other three genes from xenovulene pathway was cut using the restriction enzyme *Swa*I according to the manufacturer`s protocol. The *pyr4* gene was amplified from the gDNA of *T. reesei* using primers with 30 bp flanking regions identical to the vector backbone surrounding the *argB* gene (Figure 2.7). Yeast homologous recombination of the cut vector with the *pyr4* gene resulted in the construction of vector MSIII84 (pTYGS·*pyr4*·P_{amyB}·*aspks1*·*asL1*·*asL3 asR2*).

On the other hand, vector MSIII69 (pTYGS·*asl1*·P_{pdc1}·*aspks1*) was constructed by LR recombination of the pTYGS·*asl1*·P_{pdc} vector with the pEYA·*aspks1* vector (Dr. Raissa Schor, Cox group). The correct insertion of the genes in the corresponding vectors were confirmed by PCR and partial sequencing in all cases.

[1] Two ***different*** *T. reesei* P_{pdc} systems are used and compared here. These are annotated at P_{pdc1} and P_{pdc2}. See section 2.3.3 for a full discussion of the difference between these systems and the reason for their use.

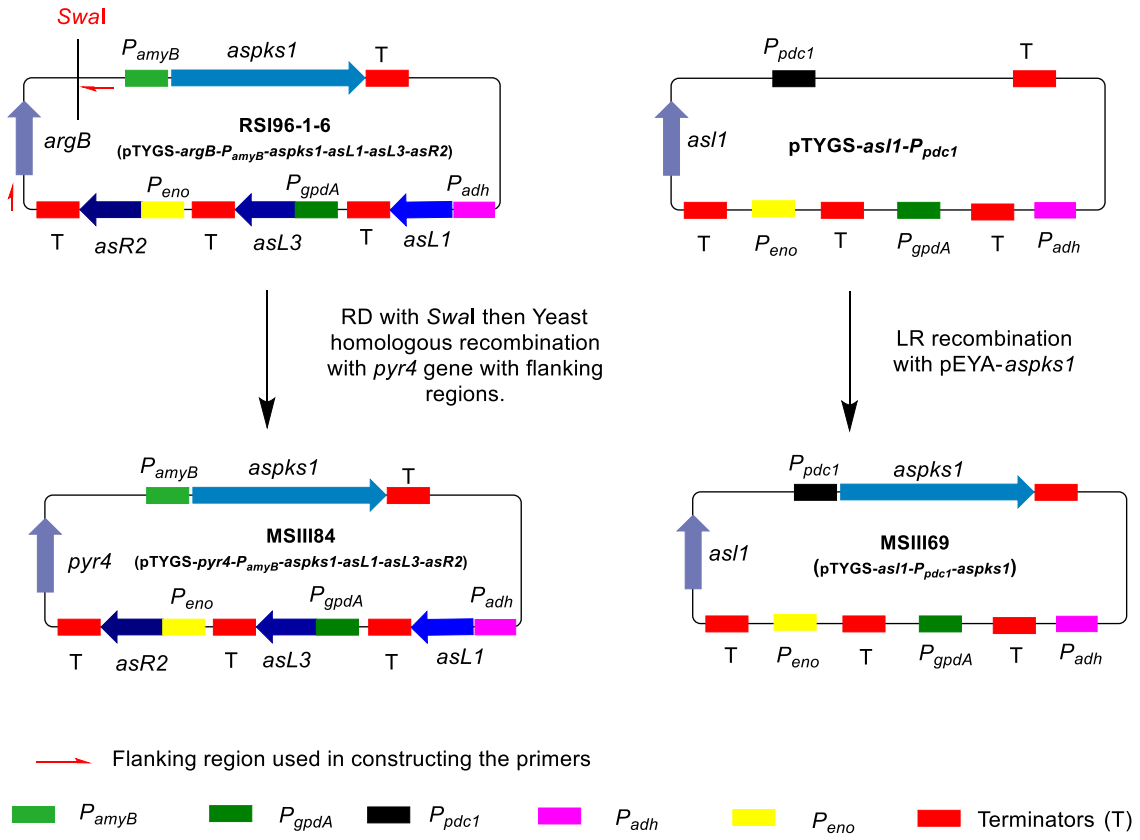


Figure 2.7 Plasmid construction using yeast homologous recombination and LR recombination.

2.3.1.2 Testing the Native *T. reesei* Promoter P_{pdc1}

Plasmid MSIII69 (pTYGS·*asl1*·P_{pdc1}·*aspks1*) was used for the transformation of the arginine auxotrophic strain *T. reesei* QM6a·*Δtmus53*·*Δasl1*. Fifteen successful transformants were chosen after three rounds of selection on minimal media lacking arginine, and the transformants were named MSIII86 A-O. The fifteen transformants were cultivated in PDB media for 7 days together with the original auxotrophic strain *T. reesei* QM6a·*Δtmus53*·*Δasl1* at 28 °C and 110 rpm. The transformants were then extracted twice with ethyl acetate and the extracts were dried over magnesium sulphate and evaporated under reduced pressure. The extracts were then submitted to LCMS analysis and the LCMS chromatograms were compared to the wild type (WT) *T. reesei* QM6a·*Δtmus53*·*Δasl1*. Only one transformant, named MSIII86-B, showed a prominent new peak (T_R = 2.8 min) when compared to the WT in addition to some other minor peaks with UV maxima corresponding to sorbcillin-related compounds (Figure 2.8).

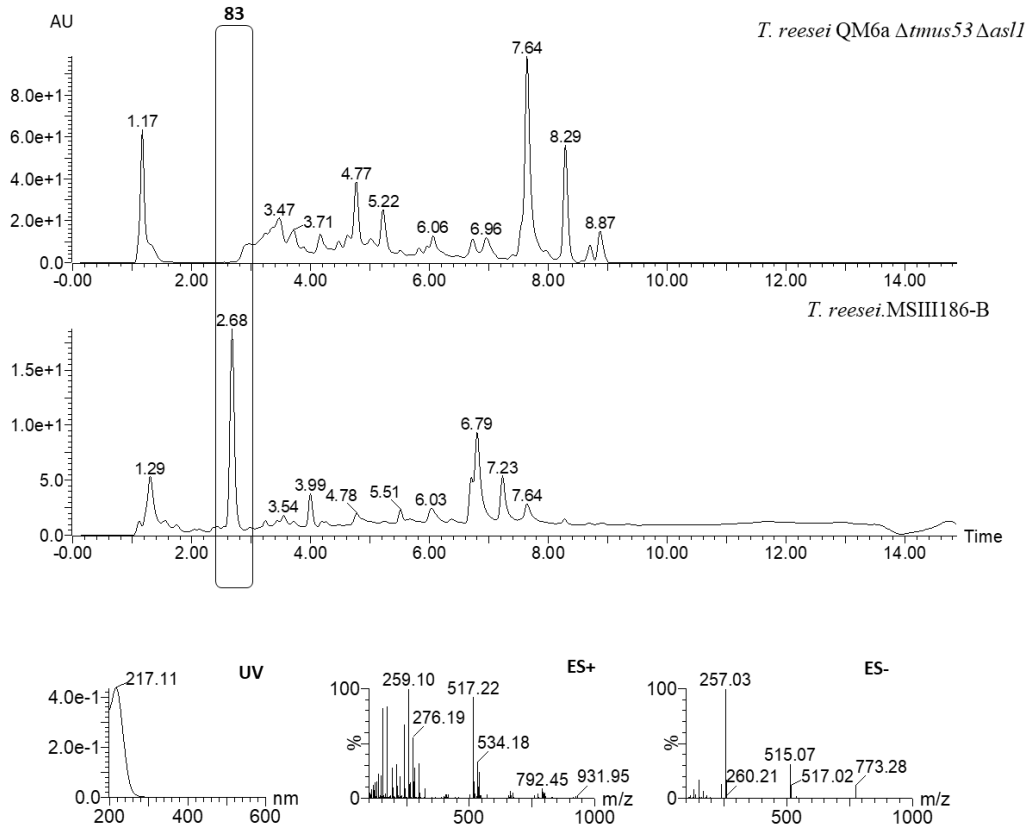


Figure 2.8 LCMS traces of MSIII86-B transformant (*T. reesei* QM6a·*Δtmus53*·P_{pdc1}·*aspks1* colony B*)* with the WT (*T. reesei* QM6a·*Δtmus53*·Δ*asl1)* showing the production of a new compound **83**.

The transformant MSIII86-B (*T. reesei* QM6a·*Δtmus53*·P_{pdc1}·*aspks1* colony B*)* was then cultivated in 1 L PDB media and grown at 28 °C and 110 rpm for seven days. This was followed by extraction of the transformant as mentioned earlier and the extract was subjected to preparative LCMS to purify the new peak with nominal mass 258. The isolated compound was then submitted to NMR and HRMS analysis. NMR data showed that there are at least three different isomers of closely related compounds present in the analysis. The main compound of these three isomers was predicted to be the hydrated scytolide **83**. However, no compounds related to the expected product of expression of the *aspks1* gene were detected. Analysis of the gDNA of thirteen different transformants showed the correct insertion of the gene in their gDNA (Figure 2.9).

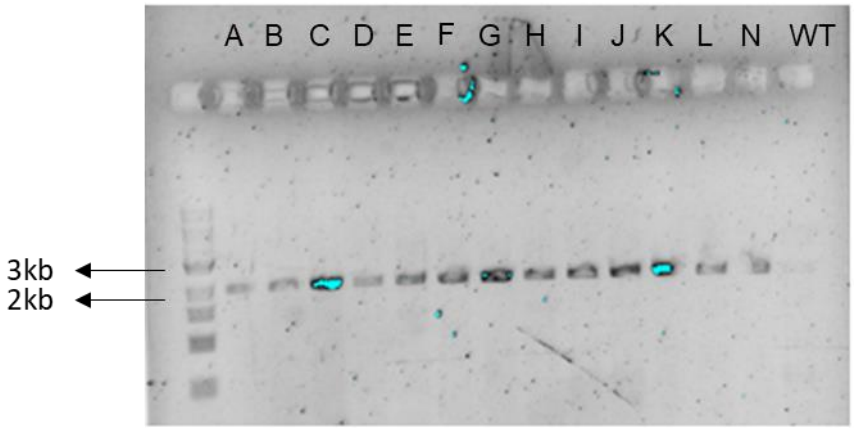
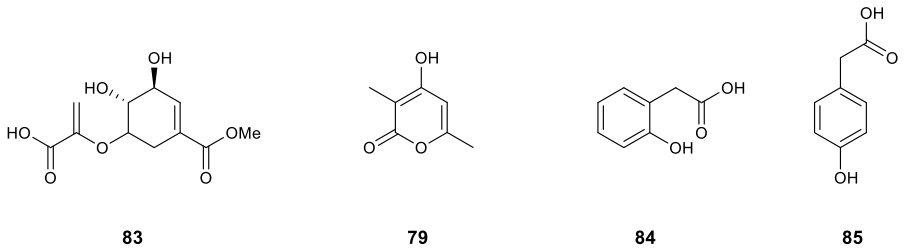


Figure 2.9 Genetic analysis of different MSIII86 transformants (*T. reesei* QM6a *Δtmus53. P_{pdc1}.aspks1*) with *aspks1* primers 83-322 (expected PCR fragment 2165 bp).

2.3.1.3 *T. reesei* Transformation of *aspks1* gene under P_{amyB}

The vector MSIII84 (pTYGS·*pyr4*·P_{amyB}·*aspks1*·*asL1*·*asL3*·*asR2*) was used for the transformation of the uridine auxotrophic strain *T. reesei* QM6a·*Δtmus53*·*Δpyr4*. After three rounds of selection on media lacking uridine, 14 different transformants were obtained, they were named MSIII93 A-N. Comparing the extracts of all the 14 transformants with the WT (*T. reesei* QM6a·*Δtmus53*·*Δpyr4)* showed that transformant MSIII93-L produced three new compounds, one with nominal mass 140 and two other compounds with nominal mass 152 (Figure 2.11). Although the compound with the nominal mass 140 showed similar UV and the same expected mass to the *aspks1* product **79**, the amount of the compound produced by the culture was too low to be isolated by preparative HPLC. Isolation of the other compounds with nominal mass 152 was possible and 1D and 2D NMR showed the compounds to be 2-hydroxy- and 4-hydroxyphenylacetic acid **84**-**85**.



83 **79** **84** **85**

Figure 2.10 Natural products isolated or detected in *T. reesei* transformants.

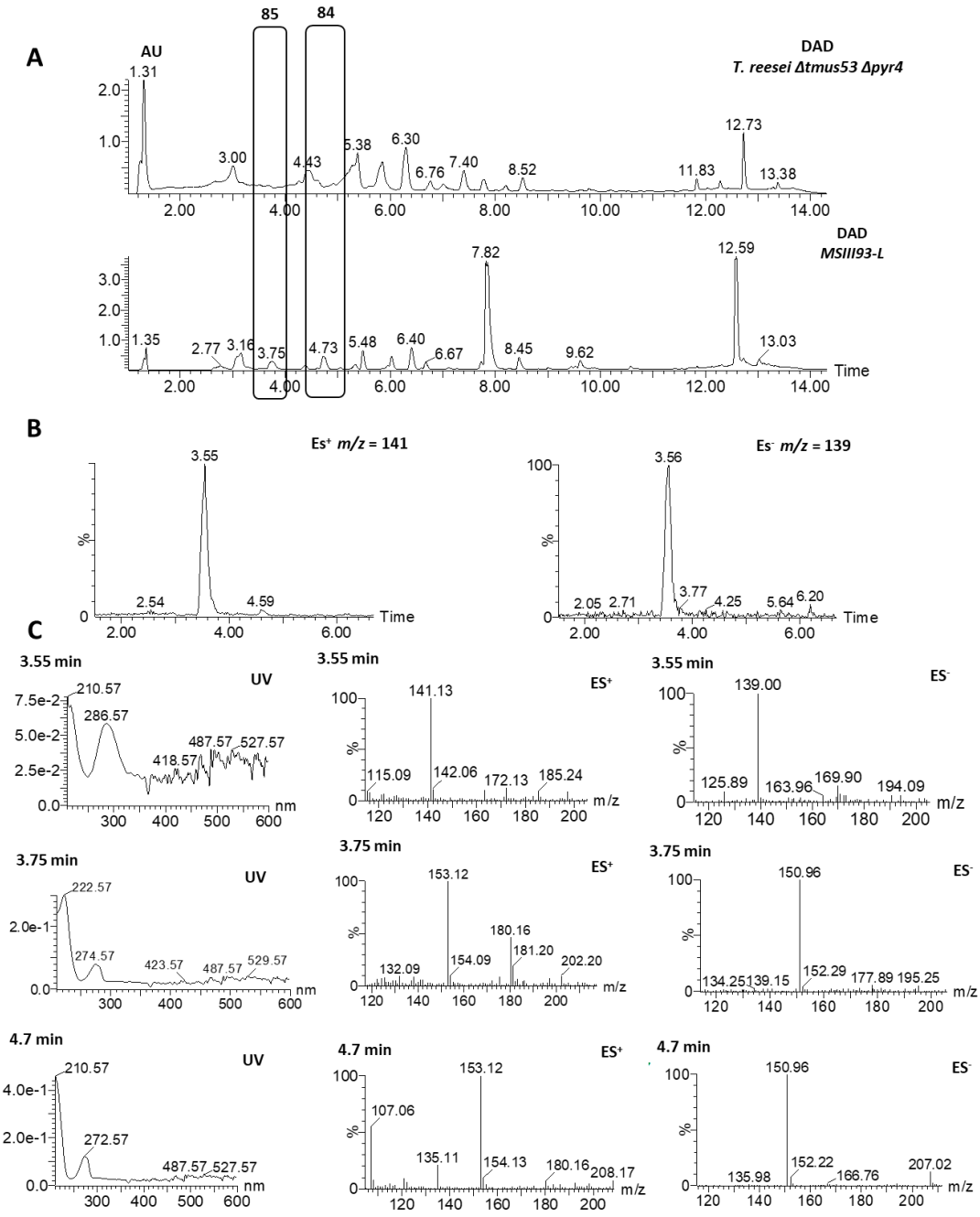


Figure 2.11 A, LCMS chromatograms of MSIII93-L (*T. reesei* QM6a·Δ*tmus53*·P_{amyB}·*aspks1 colony* L) against the WT (*T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*) using the polar gradient A2 showing the production of three new compounds; **B**, Extracted ion chromatogram (ES^+ at *m/z* = 141 and ES^- at *m/z* = 139 showing the production of compound **79** at 3.55 min in very low concentration; **C**, UV and mass data (ES^+ and ES^-) for the three compounds.

PCR analysis of four different MSIII93 transformants *(T. reesei* QM6a·Δ*tmus53*·Δ*pyr4* transformed with the pTYGS·*pyr4*·P_{amyB}·*aspks1*·*asL1*·*asL3*·*asR2* vector) showed the correct insertion of *aspks1* in all of them. However, the other three genes in the pathway

(*asL1*, *asL3* and *asR2*) were not inserted in all the colonies. Only colony N showed a band corresponding to *asL1* which is the second gene in the pathway (Figure 2.12).

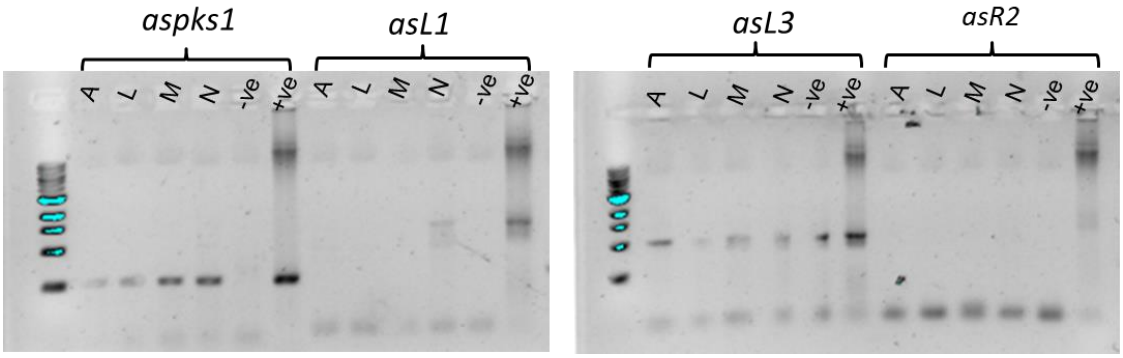


Figure 2.12 Genetic analysis of different MSIII93 transformants (*T. reesei* QM6a·Δ*tmus53*·Δ*pyr4* transformed with the pTYGS·*pyr4*·P_{amyB}·*aspks1*·*asL1*·*asL3*·*asR2* vector) with different primers for the genes *aspks1, asL1, asL3* and *asR2*.

However, the first compound in the pathway 3-methylorcinaldehyde was not produced in any of the colonies, which will prevent the activity of the later tailoring genes in the pathway. Careful inspection of extracts of all the colonies showed no production of any related compounds to the xenovulene pathway except for the 2-pyranone **79** that was produced in trace amounts by colonies A, C, G, I, J, K, L. On the other hand, almost all the colonies after transformation showed increased production of sorbcillin-related compounds (Figure 2.13).

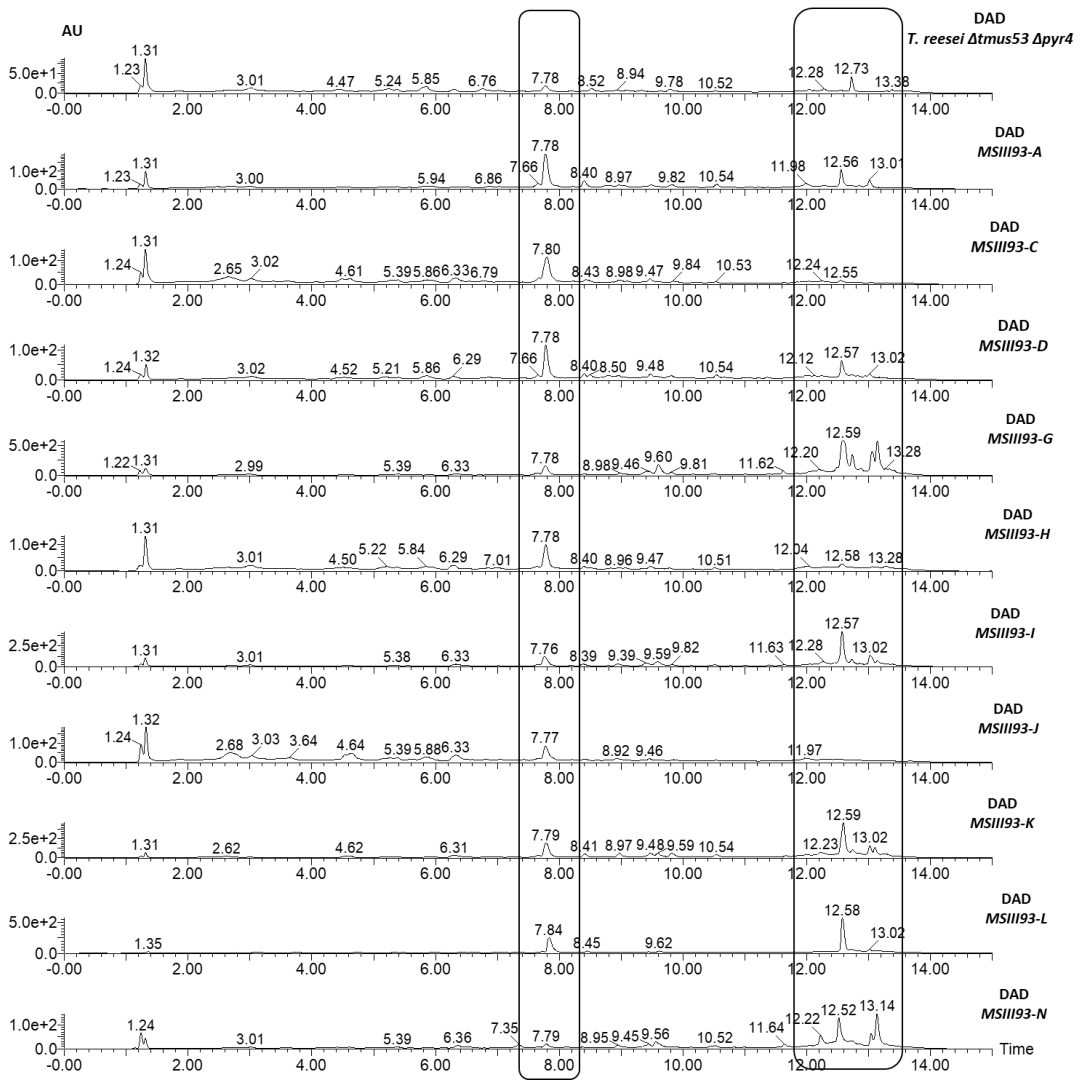


Figure 2.13 LCMS chromatograms of different MSIII93 transformants (*T. reesei* QM6a·Δ*tmus53*·Δ*pyr4* transformed with the pTYGS·*pyr4*·P_{amyB}·*aspks1*·*asL1*·*asL3*·*asR2* vector) against the WT (*T. reesei* QM6a·Δ*tmus53*·Δ*pyr4* strain) using the polar gradient A2 showing the increased production of sorbicillin-related compounds in most of the transformants.

2.3.1.4 RNA Extraction to Detect *aspks1* Transcription

To understand the reasons for the inability of *T. reesei* transformants to produce the expected product of *aspks1*, 3-methylorcinaldehyde **40**, we extracted the RNA of the two strains MSIII86-B (*T. reesei* QM6a·*Δtmus53*·P_{pdc1}·*aspks1* colony B) and MSIII193-L (*T. reesei* QM6a·*Δtmus53*·P_{amyB}·*aspks1* colony L). This was followed by reverse transcription of the extracted mRNA to obtain the cDNA and PCR analysis to check the transcription of *aspks1* using primers 1672 and 1673 (Figure 2.14). PCR analysis showed that the gene was only transcribed in transformant MSIII93-L and not in MSIII86-B, indicating that the *amyB* promoter is most probably active in *T. reesei* while P_{pdc1} appears to be inactive, which is unexpected since the native P_{pdc1} was reported to be a very strong

constitutive promoter in *T. reesei*.[116] However, these results correspond with the LCMS traces of the two transformants, where MSIII93-L (*T. reesei* QM6a·*Δtmus53*·P_{amyB}·*aspks1* colony L) showed the production of trace amounts of *aspks1*-related compound **79**, while the MSIII86-B (*T. reesei* QM6a·*Δtmus53*·P_{pdc1}·*aspks1* colony B) transformant did not produce any related compounds.

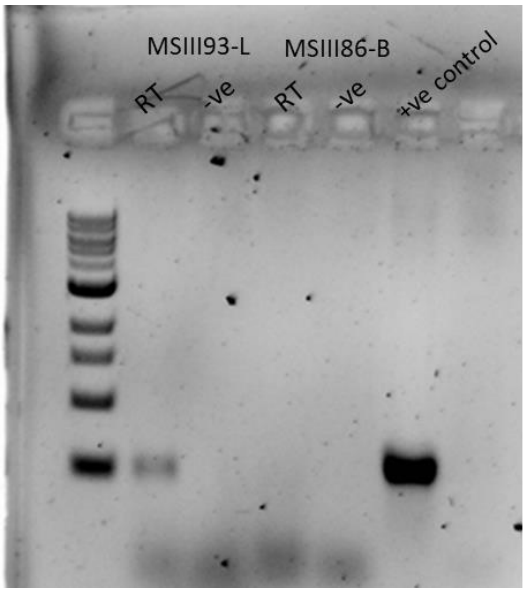


Figure 2.14 RT-PCR results of MSIII93-L (*T. reesei* QM6a·Δ*tmus53*·P_{amyB}·*aspks1* colony L) and MSIII86-B (*T. reesei* QM6a·Δ*tmus53*·P_{pdc1}·*aspks1* colony B). RT = the extracted RNA sample treated with the reverse transcriptase enzyme, -ve = the RNA sample without reverse transcriptase treatment (to ensure the absence of gDNA contamination in the extracted RNA sample), +ve control = MSIII84 (pTYGS·*pyr4*·P_{amyB}·*aspks1*·*asL1*·*asL3*·*asR2*) vector as control.

2.3.2 Knocking out the Sorbicillin Gene Cluster to Develop a Transformation Host with Cleaner Background

Production of high levels of sorbcillin-related compounds in all of the transformants was presenting a major hinderance in using *T. reesei* as a host. Firstly, because the sorbicillin gene cluster might be consuming the available acetyl and malonyl-CoA resulting in decreased levels of available building blocks for the expressed PKS. Secondly, the high levels of sorbcillin-related compounds in the extract could obscure the other peaks of other compounds produced and complicate their isolation. Therefore, it was clear that knocking out sorbicillinoid production could be important in developing *T. reesei* as a heterologous host.

2.3.2.1 Plasmid Construction Using Yeast Homologous Recombination.

The sorbicillin BGC has been identified.[152] Sorbicillin is the product of two polyketide synthases. Deletion of one of these two PKS should be enough to abolish sorbicillinoid production. However, the next enzyme in the biosynthetic pathway, SorC, has been reported to have a promiscuous activity.[156–158] The gene is able to oxidize several substrates not just sorbcillin.[156] Therefore, the presence of this gene might affect the heterologous system by performing unintended reactions and deletion might be essential to develop a better transformation host. Hence, a plasmid was constructed using yeast homologous recombination to knock out the adjacent genes *sorB* and *sorC* simultaneously. The vector was constructed using the pEYA backbone and the hygromycin resistance cassette (*hph* between P_{gpdA} and T_{trpC}) with ≈ 1 kb homologous arms to *sorB* and *sorC* genes. The plasmid was then confirmed using PCR with appropriate primers and partial sequencing and the correct plasmid was named, MSIII132 (Figure 2.15).

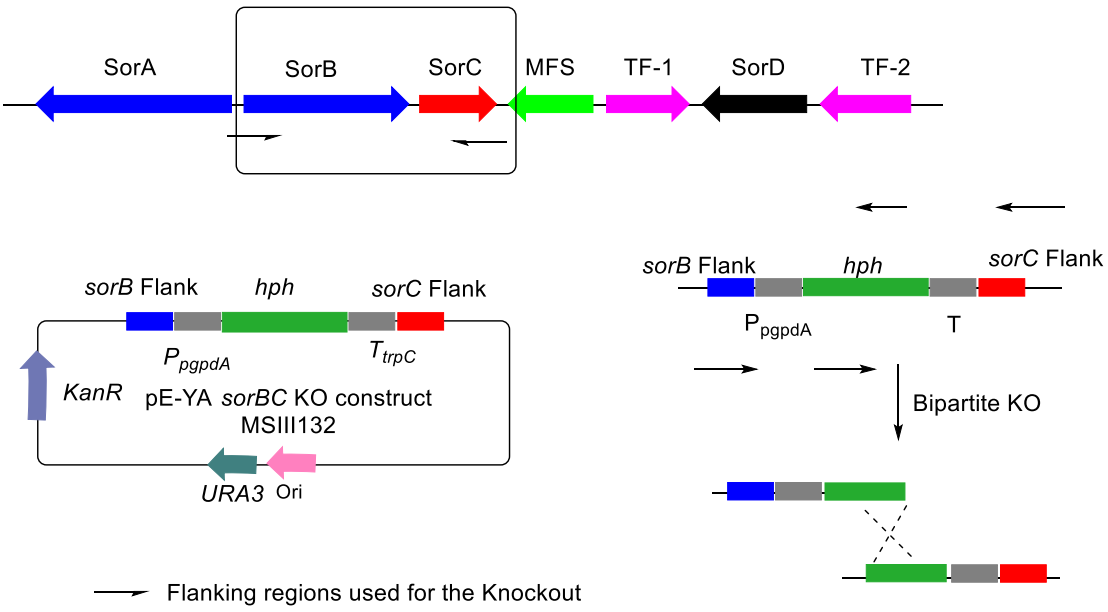


Figure 2.15 Plasmid construction to knockout the *sorB* and *sorC* genes in *T. reesei* by bipartite method.

2.3.2.2 Bipartite Method and Transformation of *T. reesei*

Bipartite gene knockout is an effective method in which a selection marker is inserted in the target gene and that also decreases the chances of ectopic integration of the DNA fragment into the gDNA of the fungus.[120] This dramatically decreases the chances of false positive transformants. In this technique, the fungus is transformed with two DNA fragments with overlapping sequence. Each fragment contains part of the hygromycin

(*hph*) resistance cassette. Neither fragment is capable of giving hygromycin resistance alone, so random insertion of individual fragments cannot give hygromycin-resistant colonies. Only strains in which three recombination events that result in correct insertion of the *hyg* cassette will be able to grow on antibiotic selection media and any ectopic integration of one of the pieces of the DNA will not result in antibiotic resistance (Figure 2.15).

Four rounds of transformations were done using the bipartite method aiming to knock out the two key genes in the sorbicillin pathway, *sorB* and *sorC*. After three rounds of selection on PDA containing hygromycin, 18 hygromycin-resistant transformants were selected. Chemical analysis of six of them showed that only one transformant, *T. reesei* QM6a *Δtmus53·Δpyr4·ΔsorBC-5B*, showed no production of any sorbicillin-related compounds compared to the WT after cultivation for seven days in DPY + 1 % glucose (Figure 2.16).

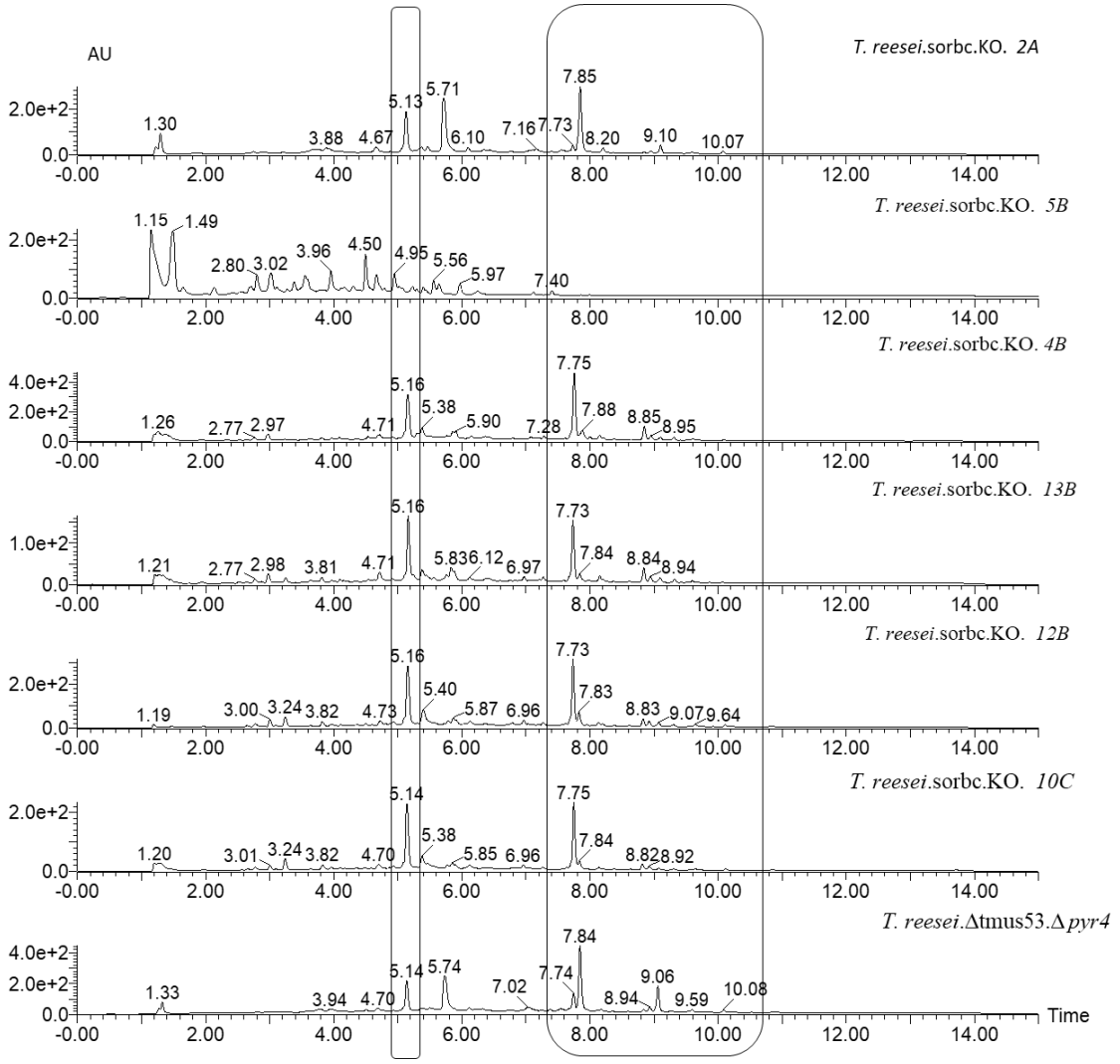


Figure 2.16 LCMS traces of the *sorBC*-KO transformants against wild type showing the absence of any sorbicillin-related compounds in only one out of the six transformants tested (*T. reesei*.sorbc.KO.5B).

2.3.2.3 Knockout of *sorB* and *sorC* in *T. reesei* and Establishing the Chemical Profile of the Knockout Strain in Different Media

PCR analysis of colony 5B and another colony (2A) together with the wild type *T. reesei* QM6a *Δtmus53·Δpyr4* using appropriate primers showed the correct insertion of the knockout cassette in the two transformants, although transformant 2A can still produce sorbicillin-related compounds (Figure 2.17). Transformant 2A may consist of mixed WT and transformed organisms, but transformant 5B appears to be a *bone fide* knockout. ITS sequencing of the correct transformant (5B) confirmed that the strain indeed is *T. reesei* and therefore the newly produced KO strain was named *T. reesei* QM6a·*Δtmus53·Δpyr4·ΔsorBC* or *TR.KO* for short.

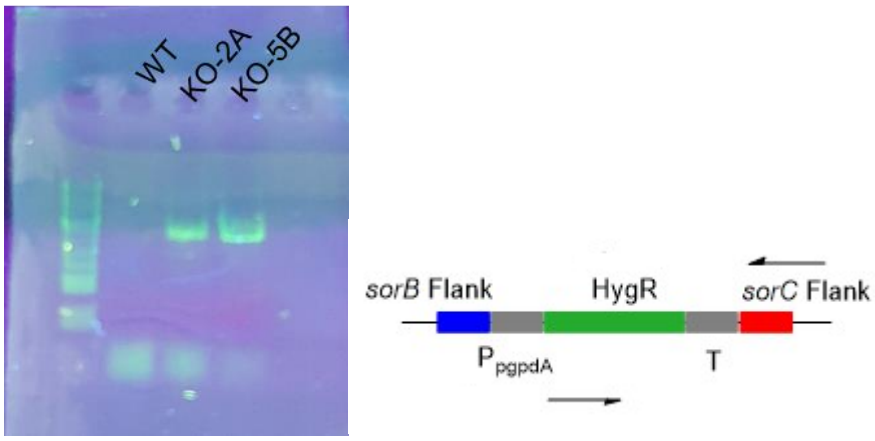


Figure 2.17 Genetic analysis of two different *T. reesei* Δ*tmus53*·Δ*pyr4*·Δ*sorBC* transformants (KO-2A and KO-5B) and the wild type *T. reesei* Δ*tmus53*·Δ*pyr4* showing the correct insertion of the KO cassette.

The correct knockout transformant 5B was then cultivated in different media to establish its chemical profile and to ensure the absence of any sorbicillin-related compounds under different cultivation conditions. The transformant was cultivated on nine different media: PDB; ME (BF); DPY; DPY+1% glucose; Mandel-Anderiotti + 1 % glucose; CM; CMP; Soya-soy-sucrose (SSS); and GNB (Figure 2.18). The new knockout strain showed no production of any sorbicillin related compounds under all the conditions tested in this experiment.

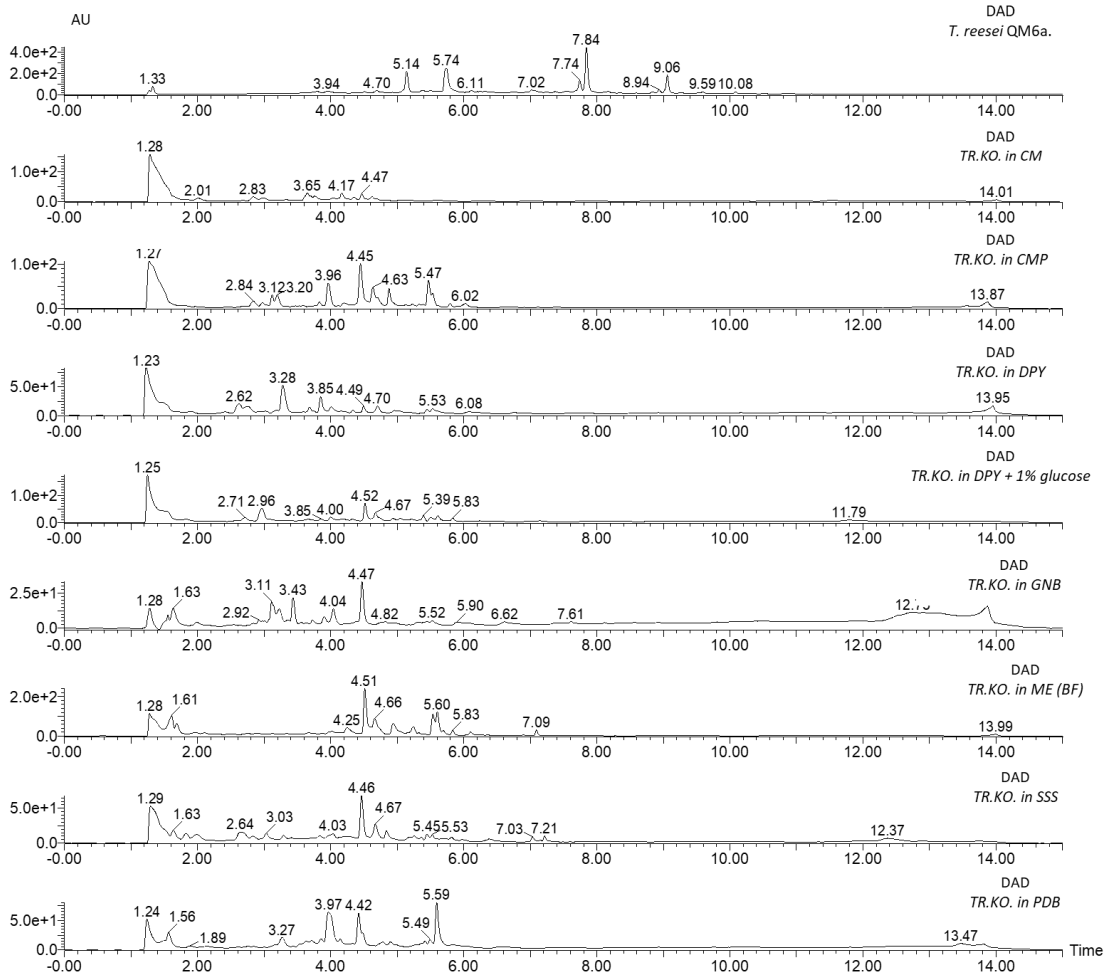


Figure 2.18 LCMS traces of *T. reesei* QM6a *Δtmus53* Δ*pyr4* Δ*sorBC* colony 5B *(*TR.KO strain) in different media showing no production of any sorbicillin-related compounds.

2.3.2.4 Isolation of New Compounds from the SorBC Knockout Strain

Although the new knockout strain did not produce any sorbicillin-related compounds, new compounds were produced that were not previously observed in *Trichoderma reesei* (Figure 2.19). This might be due to the increased availability of acetyl-CoA after knocking out the sorbicillin pathway or due to the clear analytical background that allowed other compounds to be detected in the absence of the major sorbicillin-related compounds.

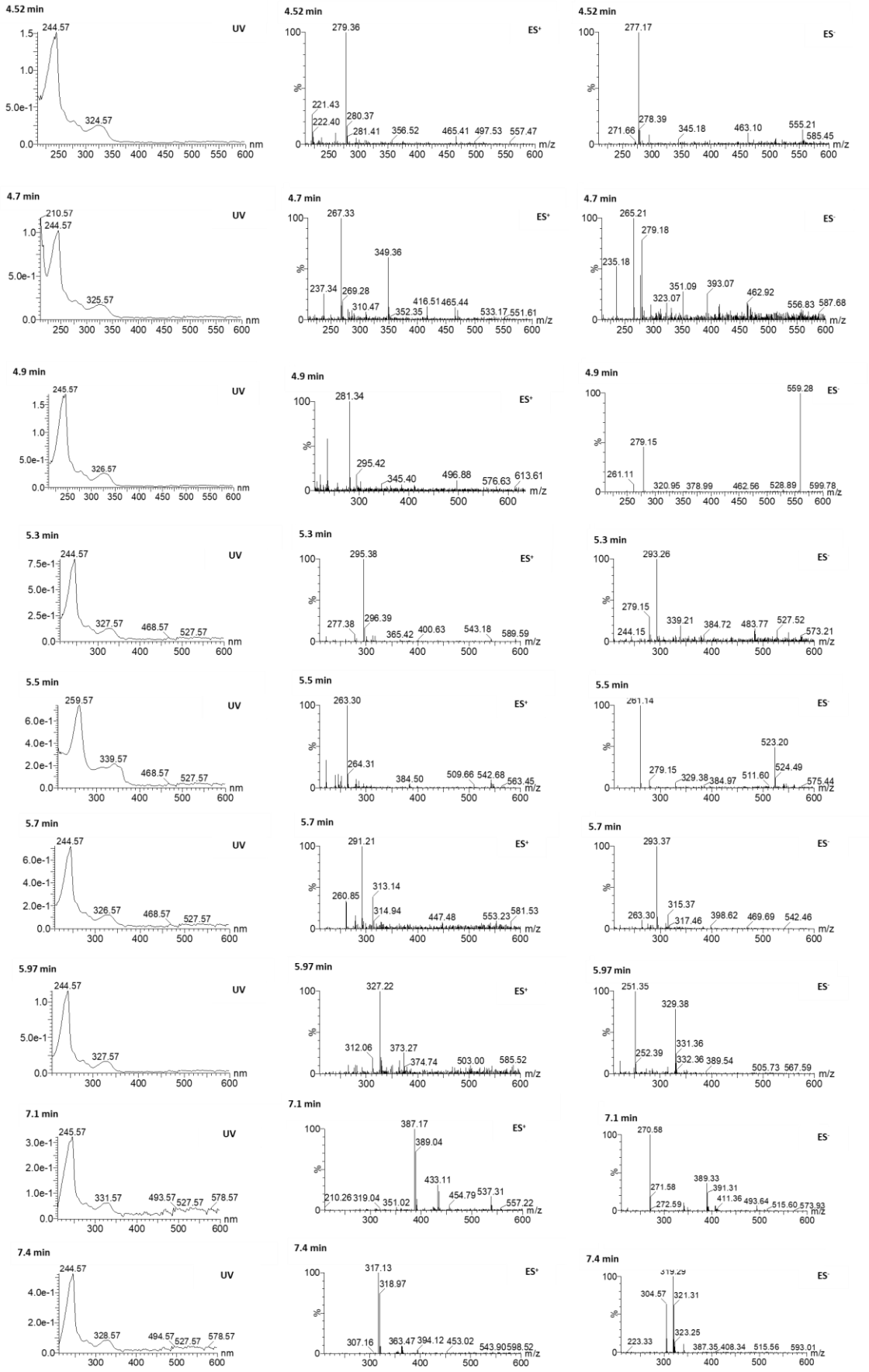
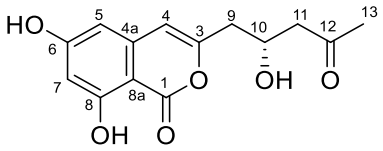


Figure 2.19 A; ES^+, ES^- and DAD spectrum for different compounds with similar UV spectrum produced by the sorBC-KO strain (*T. reesei* QM6a*Δtmus53*Δ*pyr4*Δ*sorBC* colony 5B*)* in ME (BF) media.

Only one of these compounds was present in concentration high enough to allow its isolation and structure elucidation using NMR and comparison to published data (Table 2.5). HRMS analysis of the compound confirmed a molecular formula of $C_{14}H_{15}O_6$ $[M+H]^+$ cal. 279.0869 and found 279.0887. The compound was identified to be citreoisocoumarin **86**. Although, many other compounds with UV spectra related to citreoisocoumarin were produced under different cultivation conditions, their isolation and identification were not successful due to very low titres and time constrains (Figure 2.19). Citreoisocoumarin **86** and 6-*O*-methyl-citreoisocoumarin **90** were previously reported from *Trichoderma* sp. HPQJ-34.[159]



Chemical Formula: $C_{14}H_{14}O_6$
Exact Mass: 278,08
86

Table 2.5 Chemical shifts of citroisocoumarin **86** (in comparison to literature data).

Position	Found		Reported from *Trichoderma* sp. HPQJ-34[159]		Reported for (+) citreoisocoumarin[160]	
	σ_H / *ppm*	σ_C / *ppm*	σ_H / *ppm*	σ_C / *ppm*	σ_H / *ppm*	σ_C / *ppm*
1		167.8		167.8		167.6
3		155.6		155.6		155.6
4	6.39, s	107.4	6.26, s	107.4	6.39, s	107.4
4a		141.1		141.1		141.1
5	6.32, s	103.8	6.34, s	102.7	6.32, s	103.8
6		167.4		167.3		167.2
7	6.32, s	102.7	6.26, s	99.9	6.32, s	102.7
8		164.9		164.8		164.8
8a		99.9		103.8		99.8
9	2.61 (dd, *J*= 14.5 and 8.0 Hz) 2.69 (dd)	41.9	2.55 (2H, dd, *J* = 14.3, 8.0 Hz)	41.9	2.63, m	41.9
10	4.45 (m)	66.4	4.40 (1H, dt, *J* = 12.1, 6.4 Hz)	66.4	4.45, m	66.4
11	2.69 (d, *J* = 6.3 Hz , 2H)	51.1	2.64 (2H, dd, *J* = 13.5, 5.4 Hz)	51.1	2.69, m	51.1
12		209.8		209.8		209.7
13	2.18, s	30.6	2.14, s	30.7	2.18, s	30.7

2.3.3 Building a New Vector for *T. reesei* Transformation Using the Correct P_{pdc} Promoter (P_{pdc2})

All trials to transform the genes under the activity of the native P_{pdc} were unsuccessful despite the literature reports on the strong activity of this native promoter in *T. reesei*.[116] Therefore, the plasmid pTYGS-*asl1*-P_{pdc1}, that was already available in the group was reinvestigated. The plasmid showed the correct insertion of the promoter of TRIREDRAFT_59267 gene that was annotated on NCBI as the *pdc* gene. However, the gene reported in literature to have a strong promoter in *T. reesei* is TRIREDRAFT_ 121534.[116] This gene is also annotated by NCBI as pyruvate decarboxylase (*pdc*).

Analysis of the genome of *T. reesei* on the JGI database showed that there are at least two different genes annotated as pyruvate decarboxylases; TRIREDRAFT_59267 on scaffold 6 and TRIREDRAFT_ 121534 on scaffold 8. The promoter of *pdc* reported by the literature to be active in *T. reesei* is located on scaffold 8, ~1.5 kb upstream of TRIREDRAFT_121534. Unfortunately, a former group member chose the TRIREDRAFT_59267 promoter to construct the plasmid without consulting the literature to ensure that the chosen promoter was the one proved experimentally to be a strong constitutive one. This explains the inactivity of this promoter in all the experiments performed previously. Therefore, it was necessary to construct a new plasmid with the correct P_{pdc} sequence. To avoid any confusion, the inactive *pdc* promoter is named P_{pdc1} and the active one is named P_{pdc2}.

2.3.3.1 Plasmid Construction Using Yeast Homologous Recombination.

In the construction of the auxotrophic strain *T. reesei* QM6a·*Δtmus53·Δpyr4*, Derntl *et. al*.[120] used the flanking regions to the *pyr4* gene to knockout the *pyr4* gene instead of using an antibiotic resistance cassette. The successful knockout strain was then selected using their ability to grow on the toxic 5-flouroorotic acid (5-FOA). The applicability of using these flanking regions to insert genes in a site-specific manner was tested using some reporter genes as a proof of concept.[120] The gene was successfully inserted and expressed, which could be a good strategy for site-specific integration of genes into the genomic DNA of *T. reesei*. However, Derntl *et. al.* also reported that back transformation of *pyr4* into the auxotrophic strain *T. reesei* QM6a *Δtmus53 Δpyr4* resulted in eight transformants, six of which were prototrophic, but the *pyr4* locus was re-established in only five of them. They concluded that using the fully functional *pyr4* expression cassette

can result in re-gaining prototrophy by integration of the marker elsewhere in the genome and not exclusively at the *pyr4* locus.[120]

Based on this strategy, a plasmid was constructed for easy, fast, and site-specific integration of the genes in *T. reesei* QM6a·*Δtmus53·Δpyr4* gDNA. Due to low availability of vectors and plasmids for gene integration in *T. reesei* in comparison to *Aspergillus* spp., a plasmid for *A. oryzae* transformation was used as the backbone to construct a new plasmid for *T. reesei* transformation. Due to all the advantages offered by the pTYGS-*argB* plasmid that were discussed in detail in section 2.1.2, the plasmid was used as a backbone to be modified to be used in *T. reesei*. The plasmid was constructed based on the pTYGS·*argB* plasmid, where the *Aspergillus* P_{amyB} was replaced with P_{pdc2} from *T. reesei* and the *argB* selection marker was replaced with the *pyr4* marker.

The pTYGS·*argB* plasmid was linearized using *Swa*I and the *amyB* promoter and *argB* selection marker were replaced with the native P_{pdc2} and the *pyr4* selection marker (Figure 2.20) using homologous recombination in yeast. The strain of *T. reesei* used in this experiment is a strain with a KO in the gene responsible for the ectopic integration of genes (non-homologous end-joining). Hence, the *pyr4* flanking regions of the *pyr4* locus were added between the expression cassette and the *pyr4* gene for two main reasons. First, to increase the frequency of integration of the plasmid at the *pyr4* locus in the genome. Second, to facilitate later recycling of the *pyr4* and *Ura3* genes by excision of these genes from the gDNA later by homologous recombination without affecting the expression cassette. Recycling of the *pyr4* and *Ura3* genes would allow indefinite number of transformations in the gDNA of *T. reesei* strain.

A *Pac*I restriction site was also inserted into the vector between the 3`- and 5`-*pyr4*-flanking regions by yeast homologous recombination to facilitate later linearization of the vector before transformation and further modification of the plasmid. The native promoter P_{pdc2}, 3´-downstream region from the *pyr4* gene (*pyr4*-3´-flanking region) and 5´-upstream region of the *pyr4* gene (*pyr4*-5´-flanking region)[120] were amplified from the gDNA of *T. reesei* QM6a·*Δtmus53·Δpyr4* and the *pyr4* selection marker was amplified from the gDNA of *T. reesei* QM6a using primers with 30 bp flanking arms to each other and to the pTYGS·*argB* plasmid backbone (Figure 2.20). The resulting vector was named MSIII158 and the correct insertion of the DNA fragments in the vector was confirmed by PCR and partial sequencing.

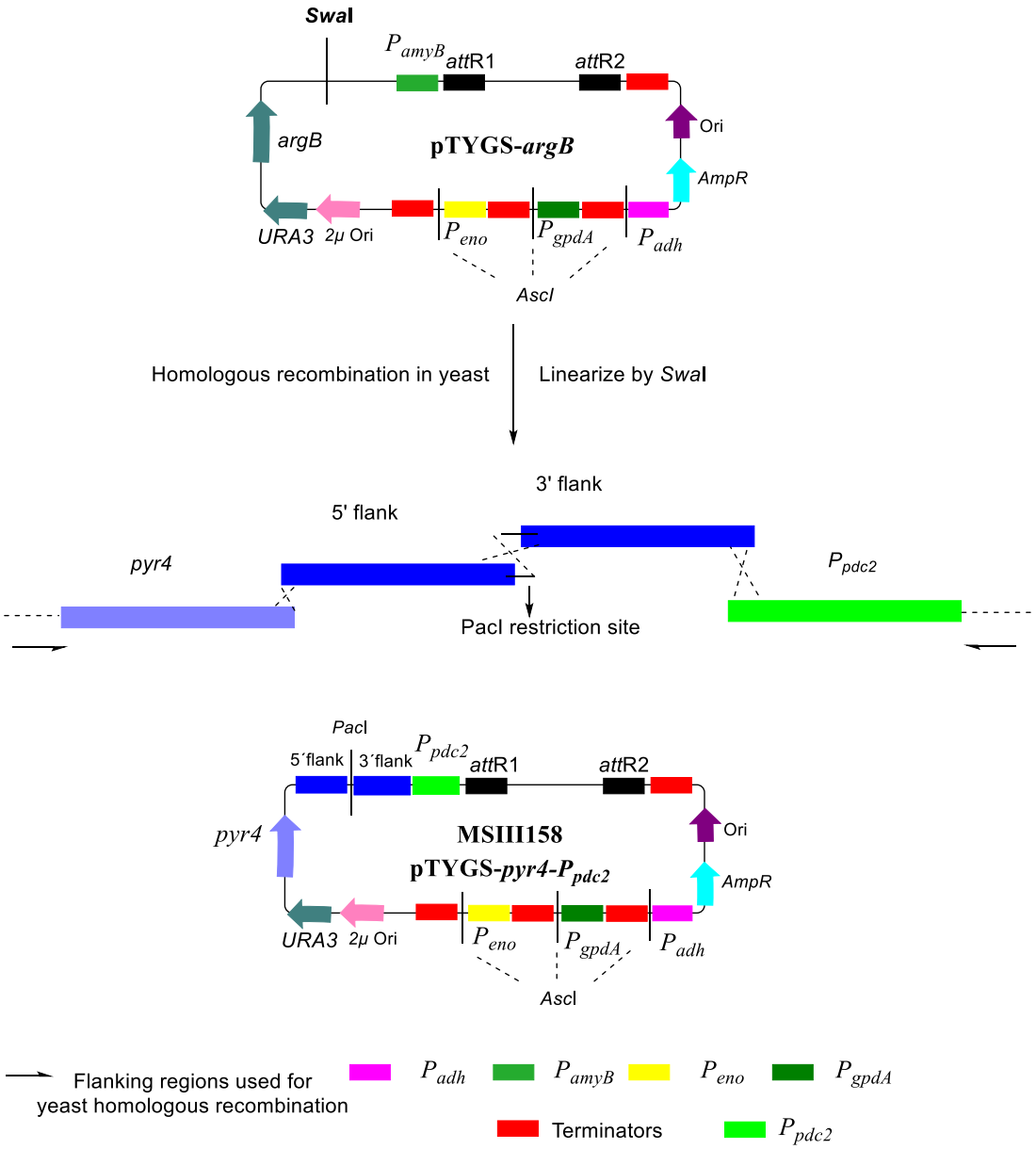


Figure 2.20 Construction of the plasmid with the correct *pdc* promoter (P_{pdc2}) using yeast homologous recombination.

2.3.3.2 LR Recombination of *aspks1* with P_{pdc2} and *T. reesei* Transformation

The constructed plasmid MSIII158 has the *att* sites that allow Gateway® recombination of any megasynthase gene present on a Gateway® entry vector. Therefore, the entry vector pEYA·*aspks1* was used together with MSIII158 (pTYGS·*pyr4*·P_{pdc2}) and the LR recombination kit to insert *aspks1* in the cloning site downstream of P_{pdc2} (Figure 2.21). The resulting vector, named MSIII163 (pTYGS·*pyr4*·P_{pdc2}·*aspks1*), was confirmed by PCR and sequencing.

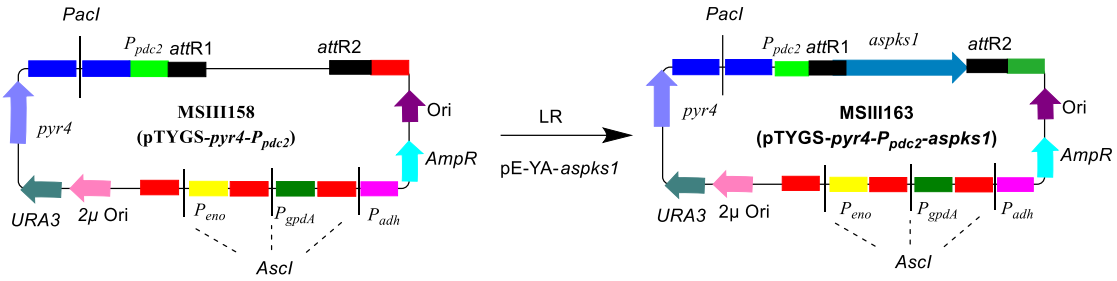


Figure 2.21 Construction of the plasmid MSIII163 (pTYGS·*pyr4*·P_{pdc2}·*aspks1*) with *aspks1* under the activity of the P_{pdc2} using LR recombination with pEYA-*aspks1* plasmid.

The vector MSIII163 (pTYGS·*pyr4*·P_{pdc2}·*aspks1*) was then linearized with *Pac*I before being transformed into *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* (*TR.KO* strain) using PEG-mediated transformation. After three rounds of selection on minimal media, eight different transformants were obtained. Five transformants were then cultivated on ME (BF) media for 6 days at 28 °C and 110 rpm. The cultures were extracted with ethyl acetate separately, evaporated under reduced pressure and submitted to LCMS. The chromatograms showed the production of two new compounds, one with nominal mass 140 at T_R = 3.3 min which matches compound **79** and another compound at 5.1 min with nominal mass of 182, which is 16 mass units higher than the expected mass of 3-methylorcinaldehyde. Careful inspection of the LCMS data showed that only trace amounts of 3-methylorcinaldehyde **40** with nominal mass of 166 is produced after 6 days of growth (Figure 2.22).

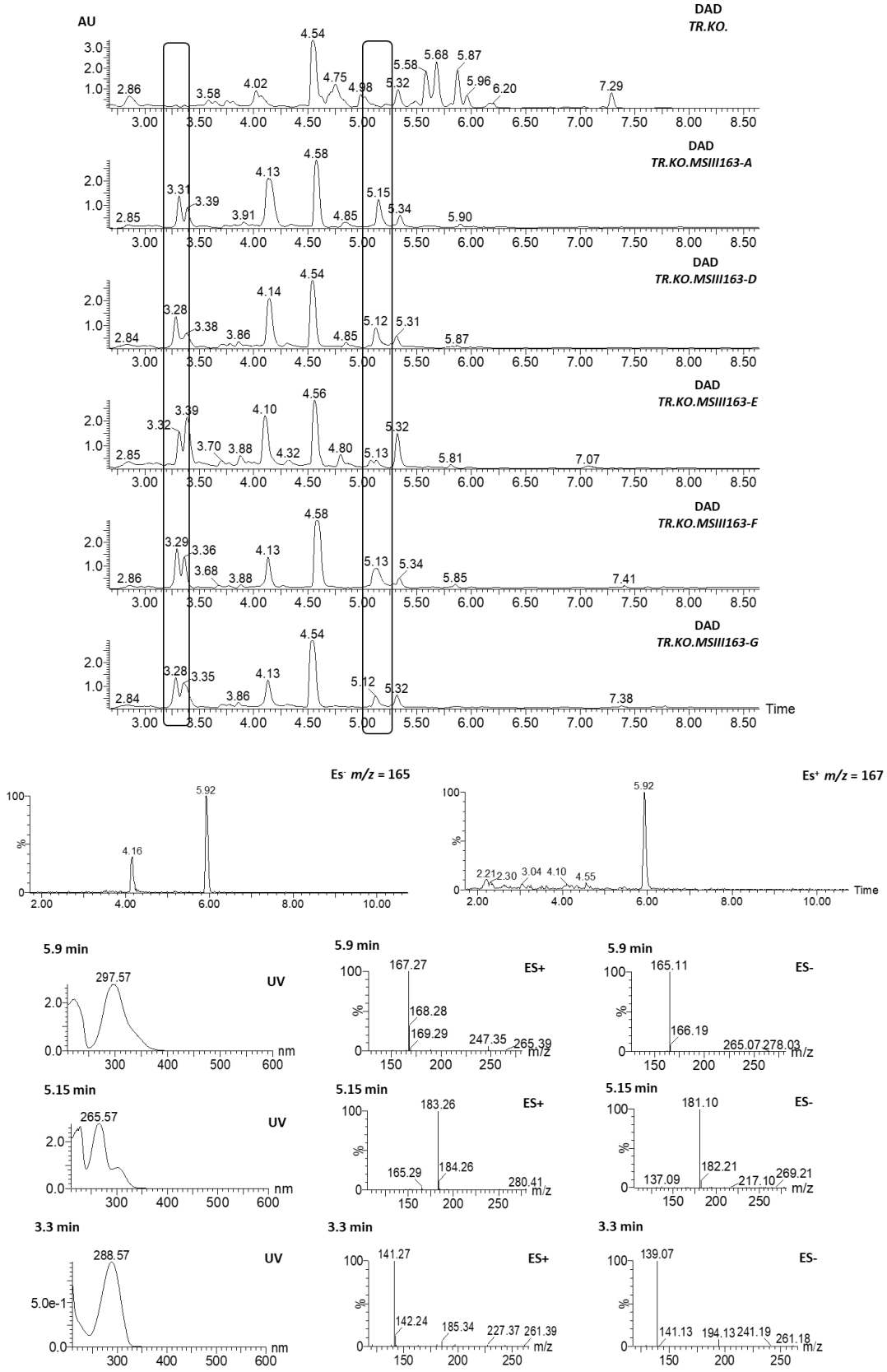


Figure 2.22 DAD LCMS traces of different *TR.KO*.MSIII163 (*T. reesei* QM6a・Δ*tmus53*・Δ*sorBC*・P_{pdc2}・*aspks1*) transformants against *TR.KO* (*T. reesei* QM6a・Δ*tmus53*・Δ*pyr4*・Δ*sorBC*) as WT after six days of cultivation showing the production of three new compounds at T_R =3.3, 5.15 and 5.9 min.

P_{pdc2} is reported to be strongly activated by glucose[130] and a former study on this promoter established that the activity of the promoter increases after 24 hours of cultivation and decreases after 72 hours of cultivation due to the total consumption of the glucose in the media.[116] Therefore, three of the successful transformants were cultivated on ME (BF) media under the same conditions but only for three days. LCMS chromatograms of the transformants showed a prominent increase in the yield of 3-methylorcinaldehyde **40** in comparison to six-day cultures, with colony C showing the best production levels (Figure 2.23).

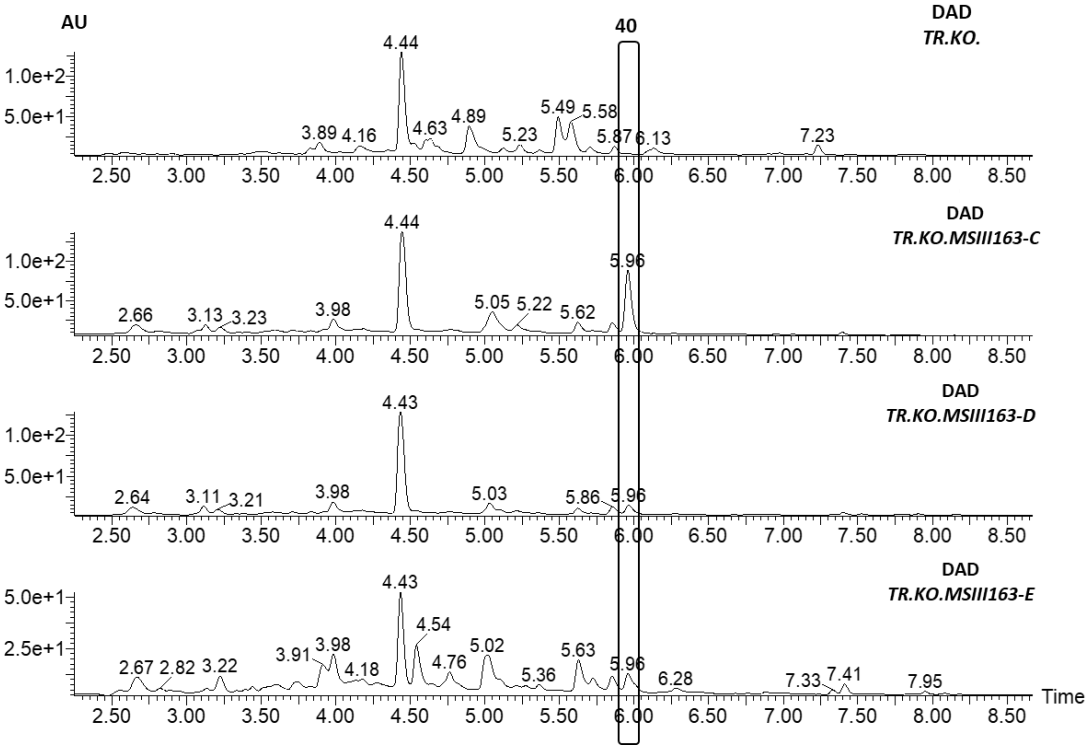


Figure 2.23 DAD LCMS traces of different TR.KO.MSIII163 (*T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1*) transformants against *TR.KO* (*T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC*) as WT after only three days of cultivation showing the production of 3-methylorcinaldehyde **40**.

To establish the time course of the production of 3-methylorcinaldehyde from *TR.KO*.MSIII163 (*T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1*) transformants, six different 500 ml baffled flasks each containing 100 ml DPY + 1 % glucose media were inoculated with the best producing transformant *TR.KO*.MSIII163-C. The flasks were incubated in the same shaker at 28 °C and 110 rpm for one to six days and were extracted at 24 hr intervals. The strain showed production of 3-methylorcinaldehyde **40** starting from the second day of cultivation (after 24 hours) and the production increased until the fourth day of cultivation. Afterwards, a decline in the production of 3-

methylorcinaldehyde was accompanied by increase in the production of a new compound with nominal mass of 182 **41** (Figure 2.24).

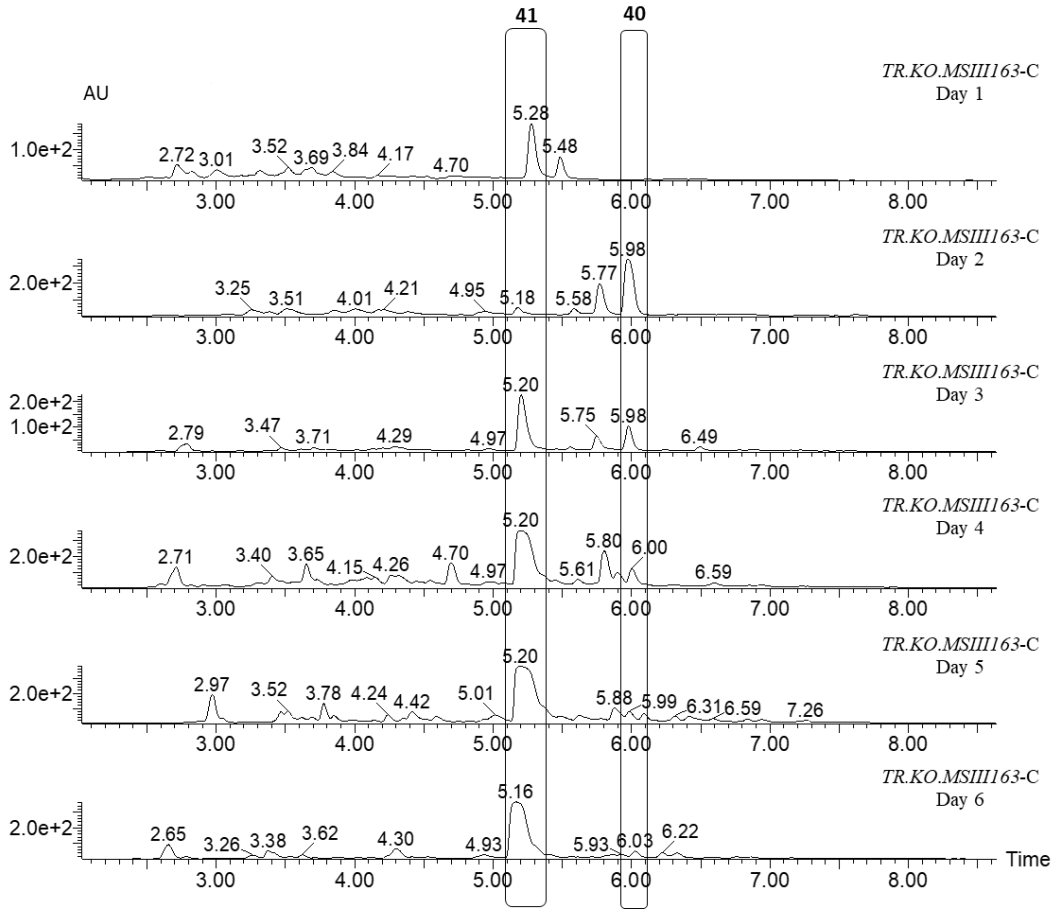
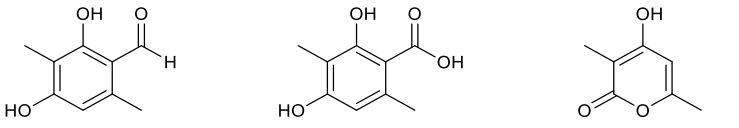


Figure 2.24 Time course for the production of 3-methylorcinaldehyde **40** and 3-methylorsellenic acid **41** from *TR.KO.MSIII163*-C (*T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1 colony* C)

The best producing strain *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1 colony* C was then cultivated on 1 L DPY + 1 % glucose media (10 x 500 ml flasks) for 48 h at 28 °C and 110 rpm. The extract was used to isolate the three main products of *aspks1* expression, compound **79** (0.7 mg), 3-methylorcinaldehyde **40** (4 mg) and the other compound with nominal mass 182 (19 mg). The chemical structures of the compounds were confirmed by HRMS and NMR analysis and comparison to published data and the compound with nominal mass 182 was identified as 3-methylorsellenic acid **41**.[161,162]



Exact Mass: 166.06 | Exact Mass: 182.06 | Exact Mass: 140.05

40 | **41** | **79**

Figure 2.25 Compounds produced in *T. reesei. P_{pdc2}.aspks1* transformants.

To check the success rate of the transformation, all of the eight transformants obtained were cultivated on PDB media for 48 h. Extraction of all of the transformants showed that seven out of the eight transformants produce 3-methylorcinaldehyde **40**, which represents ~ 87 % success rate (Figure 2.26).

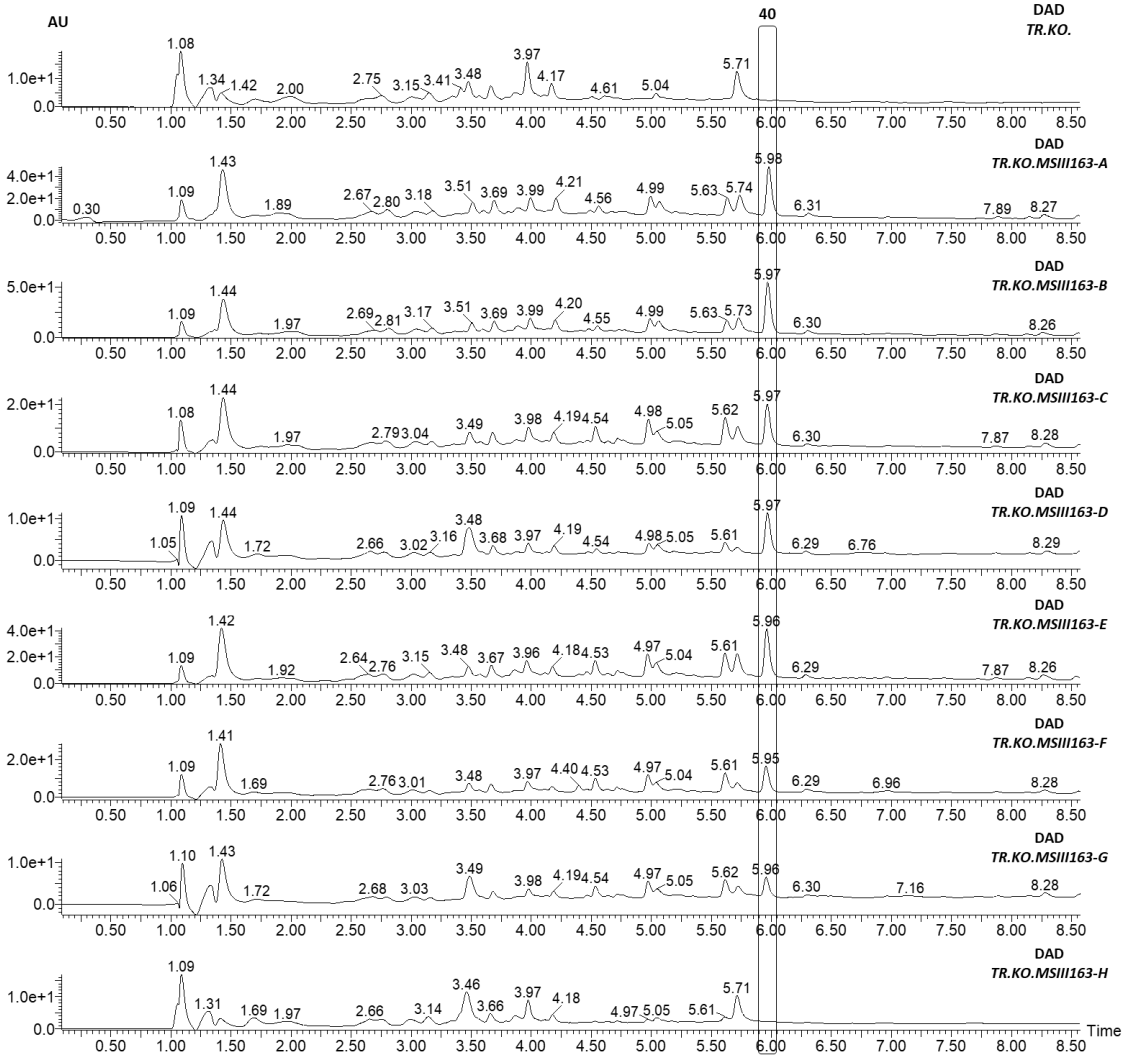


Figure 2.26 LCMS traces of eight strains of *TR.KO.MSIII163* (*T. reesei* QM6a Δ*tmus53* Δ*sorBC* P_{pdc2}-*aspks1*) transformants with the *TR.KO* (*T. reesei* QM6a･Δ*tmus53*･Δ*pyr4*･Δ*sorBC*) as WT cultivated for 48 hours on PDB media showing the production of **40** in seven out of the eight transformants.

Analysis of the gDNA of all of the eight transformants, showed the correct insertion of the *aspks1* gene in the gDNA of seven out of the eight transformants (Figure 2.27), which is in agreement with the LCMS chromatograms of the transformants.

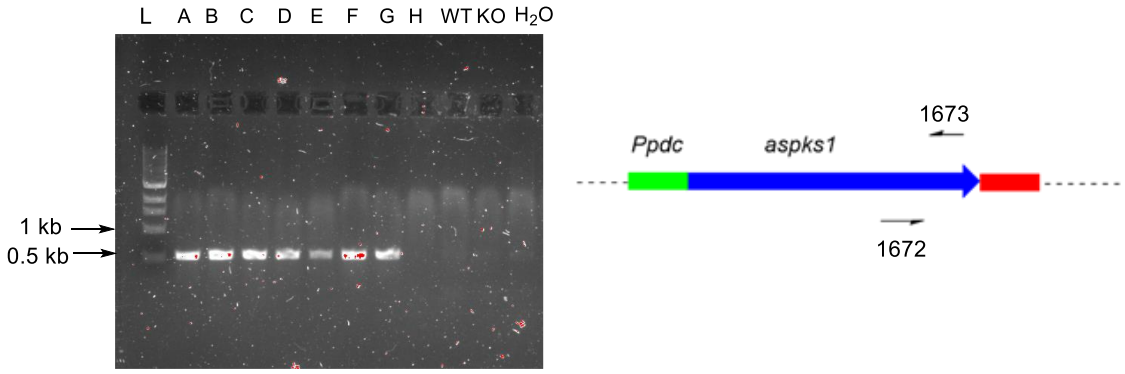


Figure 2.27 PCR analysis for the correct integration of *aspks1* gene in the gDNA of *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1* colonies A-H using primers 1672 and 1673. The gDNA of *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4* as WT (WT) and *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* (KO) strain and H_2O were used as negative controls.

2.3.3.3 Testing the Best *T. reesei.* P_{pdc2}*.aspks1* Transformant on Different Media

The best producing strain *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1 colony C* (*TR.KO.MSIII163-C*) was then used for further investigation to test the activity of the P_{pdc2} in different media, where it was cultivated on five different media to establish the effect of the media on the production of 3-methylorcinaldehyde **40**. The media used were DPY, DPY + 1 % glucose, PDB, ME (BF) and MA + 1 % glucose. On the third day of cultivation (28 °C, 110 rpm), LCMS chromatograms of the transformants showed different levels of production of 3-methylorcinaldehyde **40** and other metabolites of the *aspks1* gene in different media (Figure 2.28). The highest production level of **40** was in DPY + 1 % glucose media, but all the media showed some level of production of *aspks1* metabolites which is expected since P_{pdc} is a constitutive promoter and therefore active under different cultivation conditions.

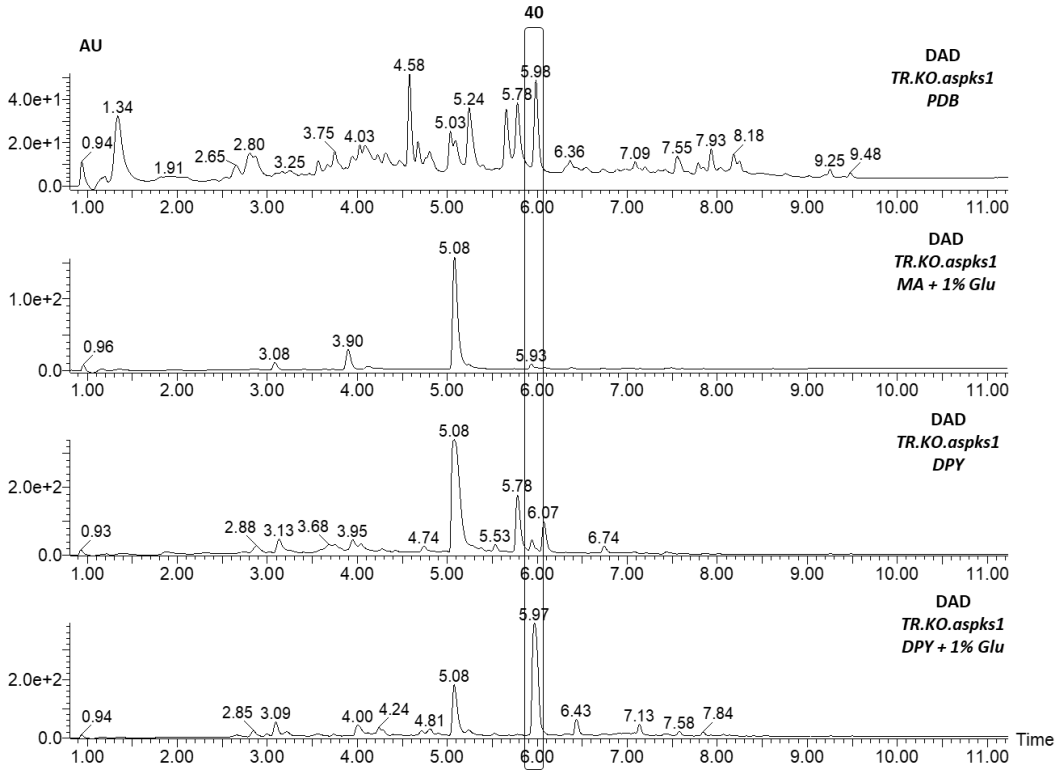


Figure 2.28 *TR.KO.MSIII163*-C strain (*T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1 colony C*) cultivated on different media showing different levels of production of 3-methylorcinaldehyde **40**.

2.3.3.4 Construction of the Calibration Curves of 3-Methylorcinaldehyde and 3-Methylorsellenic Acid

In order to quantify the amount of 3-methylorcinaldehyde **40** and 3-methylorsellenic acid **41** produced by the *aspks1* transformants, calibration curves for the two compounds were established. Serial dilutions of 3-methylorcinaldehyde **40** and 3-methylorsellenic acid **41** were made independently, with 8-9 different dilutions for each compound.

For the analysis of 3-methylorcinaldehyde, the following concentrations were prepared and 30 µl of each concentration were injected in the LCMS (150, 100, 50, 25, 12.5, 5, 2, 1, 0.5 $\mu g \cdot ml^{-1}$). For each concentration, the peak at 5.9 min which corresponds to 3-methylorcinaldehyde was integrated at a single wavelength (λ_{max} = 295 nm). The area under the curve of each peak (the average of three technical duplicates) was plotted against the concentration. As shown in figure 2.29, the resulting calibration curve had an R^2 value of 0.9999 and the equation was used to calculate the concentration of 3-methylorcinaldehyde in the media.

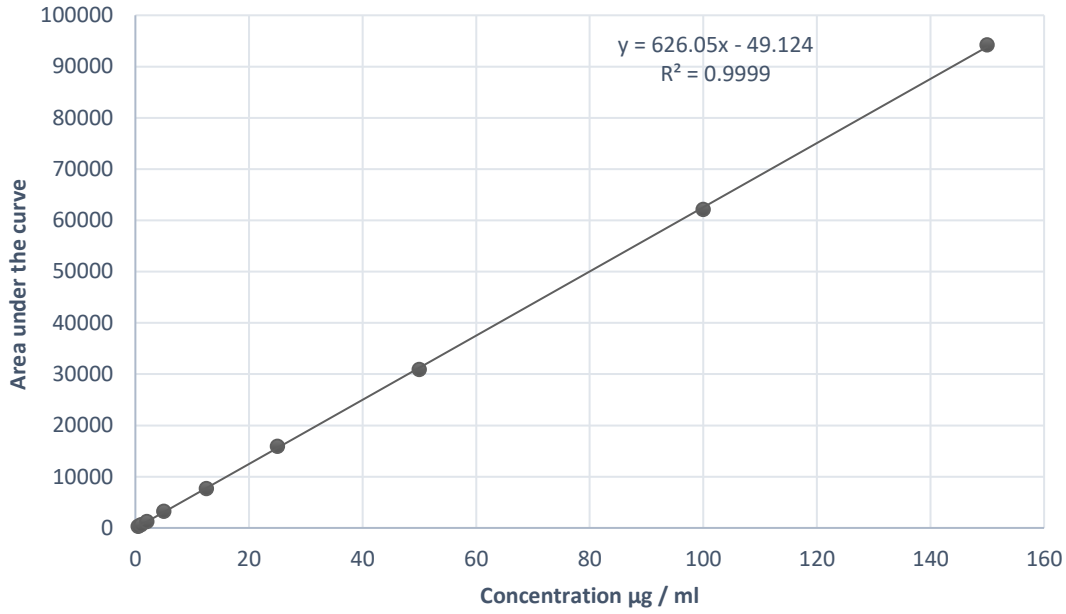


Figure 2.29 Calibration curve of 3-methylorcinaldehyde **40**.

For the analysis of 3-methylorsellenic acid, the following concentrations were prepared and 30 µl of each concentration were injected in the LCMS (150, 100, 50, 25, 15, 10, 5, 2 $\mu g \cdot ml^{-1}$). For each concentration, the peak at 5.1 min which corresponds to 3-methylorsellenic acid was integrated at a single wavelength (λ_{max} = 265 nm). Plotting the area under the curve against concentration resulted in a straight line with calculated R^2 of 0.9989 (Figure 2.30).

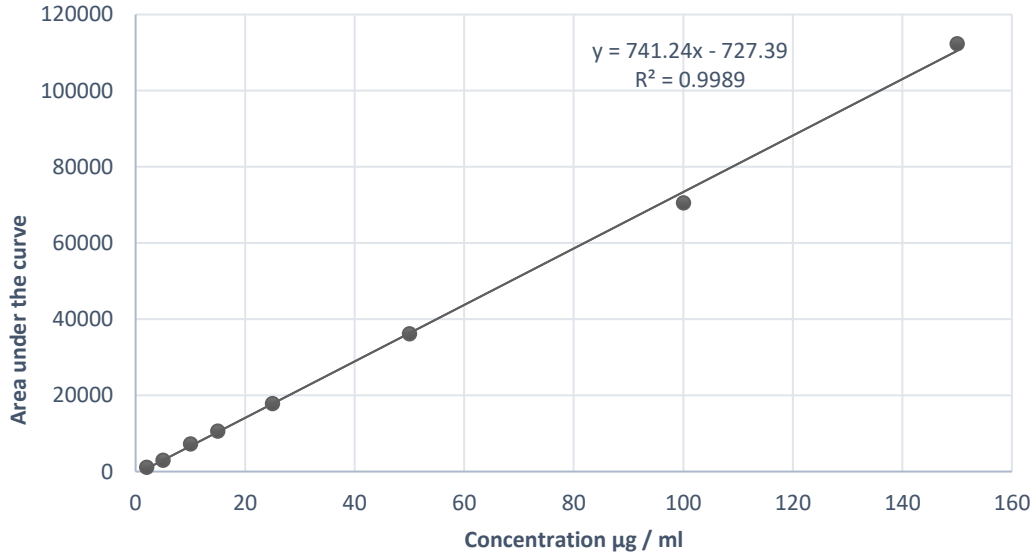


Figure 2.30 Calibration curve of 3-methylorsellenic acid **41**.

2.3.4 Testing the Productivity of the New strain on Waste Materials

Since *T. reesei* is well known for its production of cellulases and other cell-wall degrading enzymes, the ability of the new transformant, *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1 colony C (TR.KO.MSIII163-C)*, to grow on different agricultural waste material was tested.

2.3.4.1 Testing the Strain Productivity on Different Waste Materials

Different waste materials were used for this experiment such as potato peel, orange peel, banana and kiwi peel. Fifty grams of each of the these materials were weighed, cut into small pieces, and placed in 500 ml flasks with 100 ml pure water. The flasks were then autoclaved at 120 °C for 15 minutes and left to cool. The flasks were cultivated with the best producing strain, *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1 colony C (TR.KO.MSIII163-C)* and *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* (*TR.KO)* as a control, separately. The flasks were then incubated at 28 °C for 14 days in the dark (Figure 2.31).

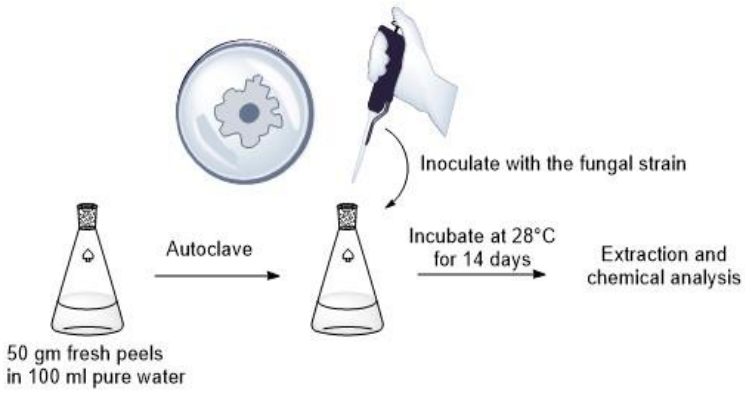


Figure 2.31 Preparation of agricultural waste media for fungal cultivation.

Except for orange peels, the fungi grew on the other three media quickly and hence each media together with the fungal mycelia was homogenised after 14 days of cultivation using a hand blender and then filtered before extraction twice with ethyl acetate. LCMS analysis of the extracts showed that the best producing strain was able to produce the two major products of *aspks1* gene **40** and **41** on the three different media but in different ratios (Figure 2.32).

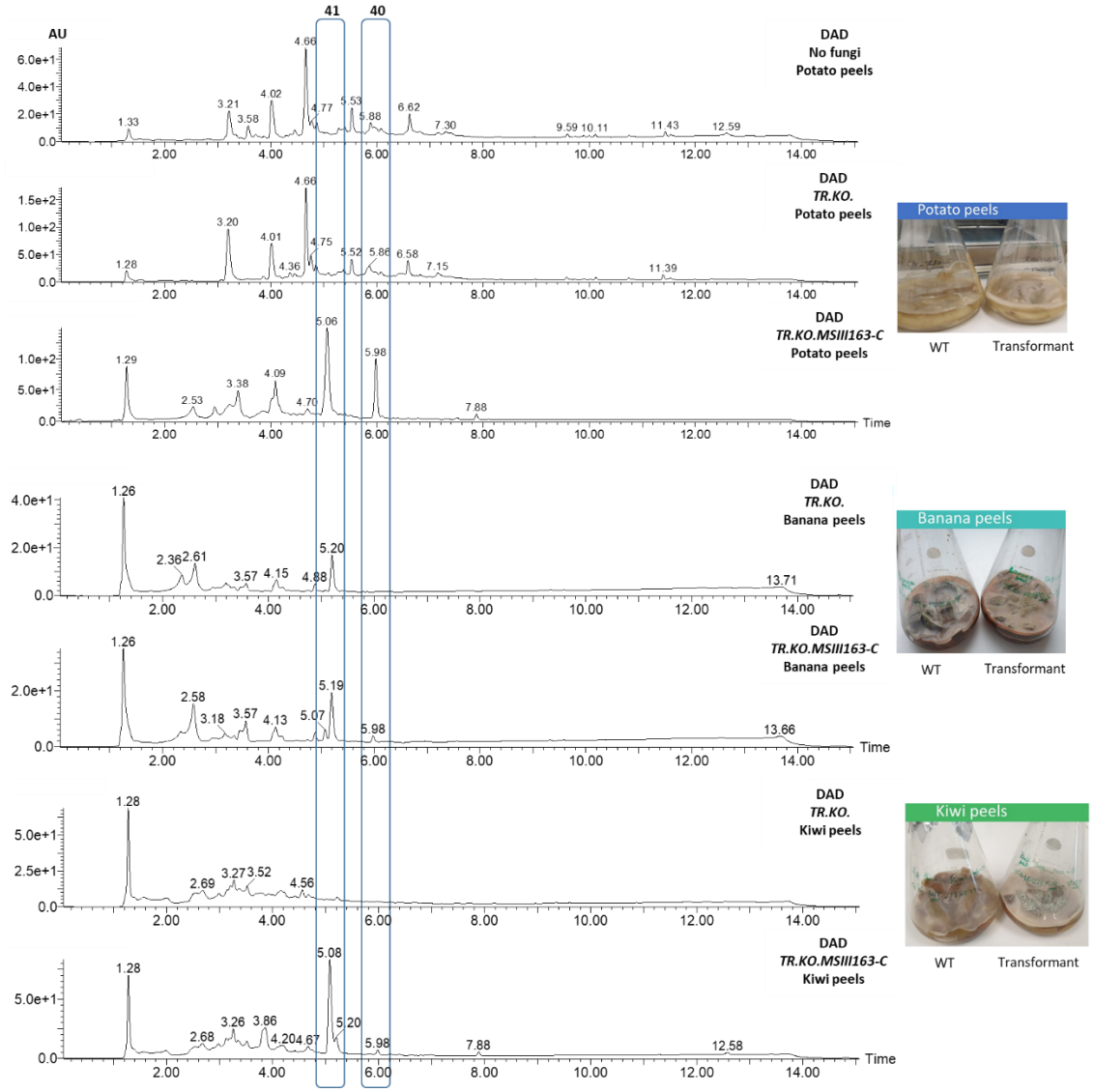


Figure 2.32 DAD LCMS traces of *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1 colony C (TR.KO.MSIII163-C)* transformant growing on different waste materials showing production of 3-methylorcinaldehyde **40** and 3-methylorsellenic acid **41**. *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* (*TR.KO)* strain was used as a control in each media.

The fungal strains were not able to grow on orange peels for over a month, therefore, the inoculated flasks were left in the incubator for longer period of time together with un-inoculated flask containing orange peels treated in the same manner. After four months of cultivation, the fungal strains were able to grow abundantly on the orange peels, while the negative control showed no growth (Figure 2.33). The fungal mycelia was then homogenised with the orange peels using a hand blender before filtration and extraction with ethyl acetate. The negative control, with the media alone, was also treated in the same way to provide a clear comparison. As shown in figure 2.33, the *aspks1* transformant was able to produce mainly 3-methylorsellenic acid **41** and small amount of 3-

methylorcinaldehyde **40** in addition to another two compounds that were also produced in the WT strain but not in the orange peels extract.

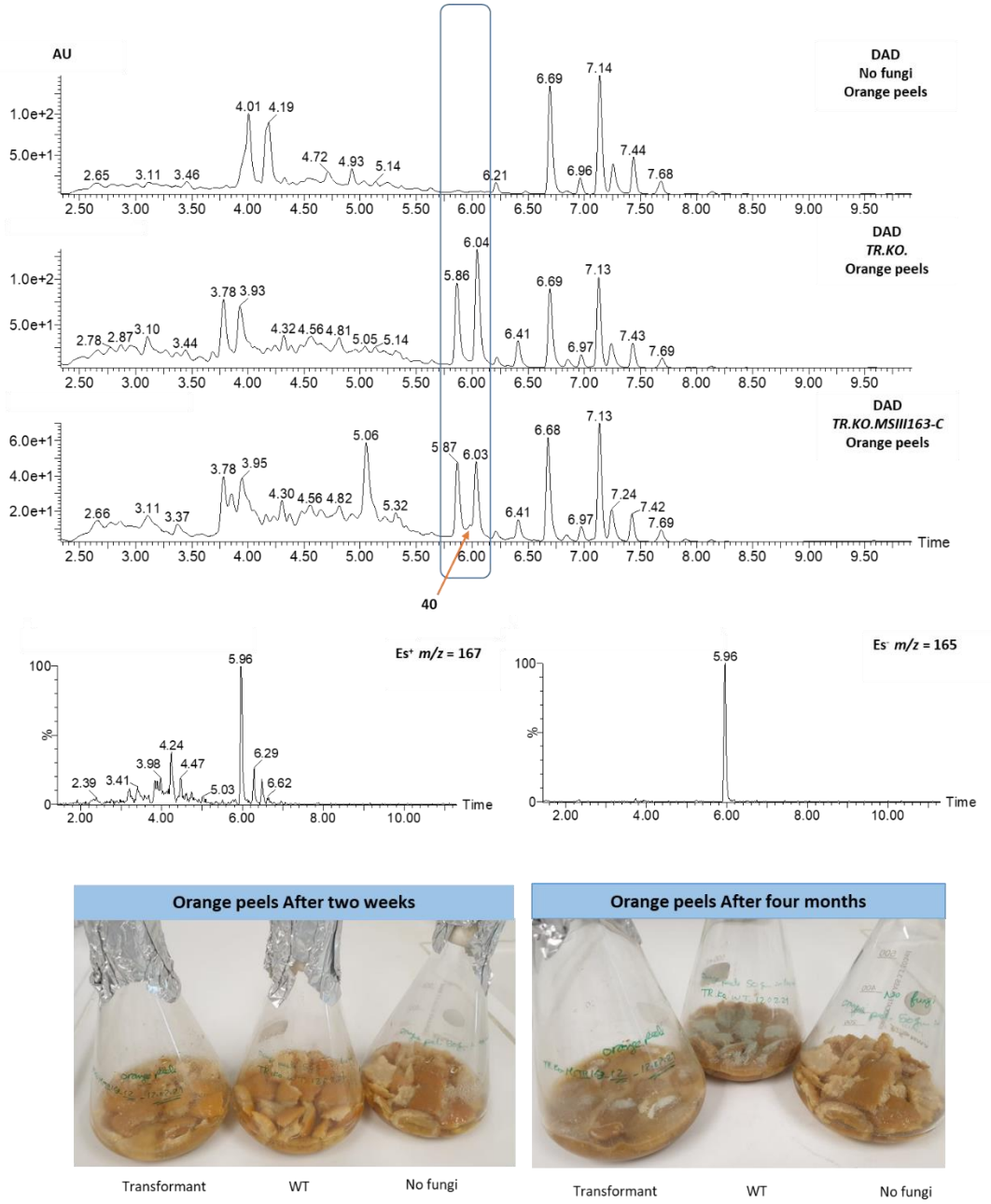


Figure 2.33 LCMS chromatograms of *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1 colony C* (*TR.KO.MSIII163*-C), *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* (*TR.KO)* and the negative control (No fungi) showing the production of 3-methylorcinaldehyde **40** by *TR. KO. MSIII163*-C strain on orange peels.

2.3.4.2 Quantification of 3-Methylorcinaldehyde and 3-Methylorsellenic Acid on Potato Peels

The best production level of 3-methylorcinaldehyde **40** was shown to be on potato peel media (Figure 2.32) and hence cultivation of the fungal strain *TR.KO.MSIII163-C (T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1* col.C) was repeated on 50 g potato peel media prepared as previously described. Another two flasks with 50 g potato peel media were used as controls, one with the *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* (*TR.KO*) and one as negative control without any fungal innoculation. After incubation in the dark at 28 °C for 14 days, all the flasks were extracted separately and each extract was dissolved in 20 ml acetonitrile. For the quantification of 3-methylorcinaldehyde and 3-methylorsellenic acid in potato peel media, 30 µl of the extract were injected into the LCMS. The peak at 5.1 min which corresponds to 3-methylorsellenic acid **41** was integrated at a single wavelength (λ_{max} = 265 nm) and the one at 5.9 min, which corresponds to 3-methylorcinaldehyde **40** was also integrated at a single wavelength (λ_{max} = 295 nm). Quantification was done by applying the equation from the two calibration curves for 3-methylorsellenic acid **41** and 3-methylorcinaldehyde **40** (section 2.3.3.4). The results showed that the strain could produce ~ 1 mg of 3-methylorcinaldehyde **40** and ~ 2 mg of 3-methylorsellenic acid **41** per flask (Table 2.6). In order to calculate the concentration of each compound per kg dry wt, 50 g of fresh potato peels were freeze-dried and weighed. Fifty grams fresh potato peels yielded 8.36 g dry wt.

Table 2.6 Calculation of the amount of the compounds in potato peels extract.

	Amount per L / mg	Amount per flask / mg	$mg \cdot Kg^{-1}$ dry weight
3-methylorsellenic acid	101.8	2.036	243.5
3-methylorcinaldehyde	53.6	1.072	128.2
Total	155.4	3.108	371.7

2.3.5 Expression of a Multi-Gene Pathway in *T. reesei*

After the successful integration and expression of the *aspks1* gene in *T. reesei* and the successful production of 3-methylorcinaldehyde on different waste materials, we were interested to test the ability of *T. reesei* to express more genes. Therefore, more native promoters were chosen to be inserted in the pTYGS plasmid to allow for the successful expression of more genes. In order to allow the fungus to express the genes under most conditions without the need for induction, constitutive promoters from *T. reesei* were chosen to be inserted in the plasmid. Based on table 2.3 (section 2.1.4.2), the strongest known constitutive promoters are P_{pdc}, P_{cDNA1} and P_{eno} (P_{TReno}). Therefore, the two promoters, P_{cDNA1} and P_{TReno} were chosen to expand the plasmid MSIII158. According to the published literature, P_{TReno} is located from 102421 bp to 103910 bp on scaffold 4[116] and P_{cDNA1} is located from 43726 bp to 44652 bp on scaffold 23 (Upstream of the *cDNA1* gene TRIREDRAFT_110879).[131]

P_{cDNA1} and P_{TReno} coding regions were amplified from gDNA of *T. reesei* QM6a *Δtmus53 Δpyr4* using primers with overhang homologous to MSIII158 backbone and to the terminators T_{adh} and T_{eno}, respectively. The primers were designed in a way to insert a *Swa*I restriction site after each of the promoters, P_{cDNA1} and P_{TReno}, to facilitate later gene insertion by yeast homologous recombination. A patch for the P_{gpdA} promoter was amplified by PCR using MSIII158 (pTYGS·*pyr4*·P_{pdc2}) vector as the template.Since the vector MSIII158 contained another site for *AscI* restriction in the Pyr4-3`flanking region, a patch for this *AscI* restriction site was also amplified by PCR using the vector MSIII158 as the template. The vector was then linearized using *Asc*I and the new vector was constructed by yeast homologous recombination between the linearized vector MSIII158 (pTYGS·*pyr4*·P_{pdc2}), the coding sequence of P_{cDNA1} and P_{TReno} and the patch for the P_{gpdA} promoter and for the AscI restriction site. The resultant vector was named MSIII172 (pTYGS·*pyr4*·P_{pdc2}·P_{cDNA1}·P_{TReno}, Figure 2.34).

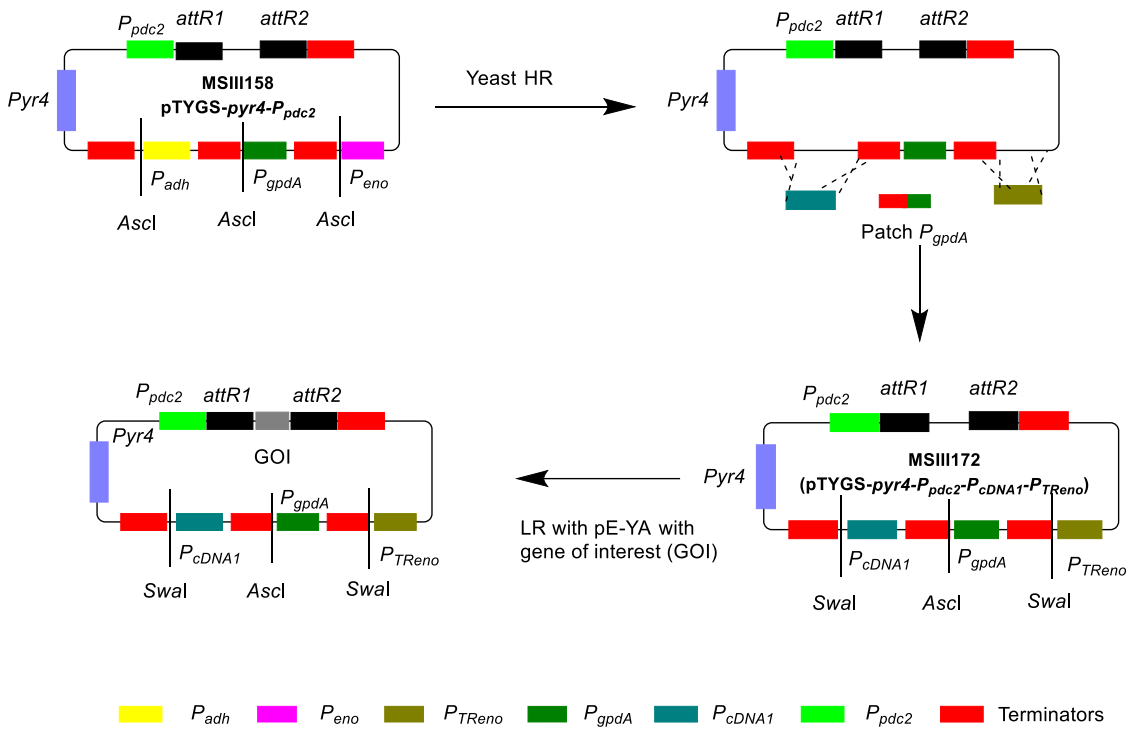


Figure 2.34 Construction of plasmid MSIII172 using yeast homologous recombination.

To test the ability of the system to express more than one gene, other more complicated secondary metabolites such as the products from the PKS-NRPS megasynthetases was used. Based on previous experiments done with different PKS-NRPS genes, most PKS-NRPS genes contain an inactive ER domain. Therefore, usually another *trans*-acting ER gene should be expressed in the heterologous host to process the secondary metabolite produced by this megasynthetase.

The vectors containing the *tenS* and *tenC* gene sequences were kindly provided by Katharina Schmidt, Cox group. LR recombination between the vector MSII172 and pEYA·TenS resulted in the construction of MSIII173 (pTYGS·*pyr4*·P_{pdc2}·*tenS*·P_{cDNA1}·P_{TReno}) with *tenS* under the activity of P_{pdc2}. The coding sequence for *tenC* was amplified from the vector pTYGS·*argB*·*tenS*·*tenC* using primers with flanking regions to P_{cDNA1} and T_{adh}. A patch for P_{TReno} was also amplified by PCR using MSIII172 (pTYGS·*pyr4*·P_{pdc2}·P_{cDNA1}·P_{TReno}) as the template. The vector MSIII173 was linearized by *Swa*I and yeast homologous recombination between the linearized vector, *tenC* gene and patch P_{TReno} resulted in the construction of MSIII174 (pTYGS·*pyr4*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC*·P_{TReno}, Figure 2.35).

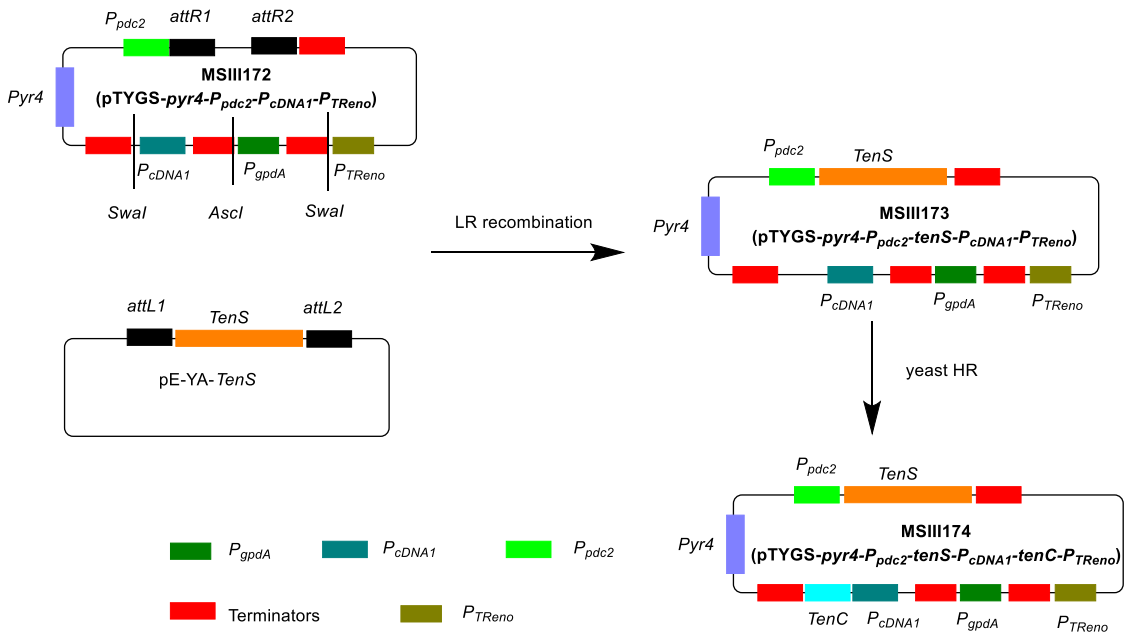


Figure 2.35 Construction of plasmid MSIII174 (pTYGS·*pyr4*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC*·P_{TReno}) using LR and yeast homologous recombination.

Sequencing of MSIII174 (pTYGS·*pyr4*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC*·P_{TReno}) showed two point mutations in the PKS-NRPS (*tenS*) gene, where one base was removed and one base was added (Figure 2.36). Although this should not result in a frame shift of the gene, the changes in the sequence would result in a change of two amino acids in the resulting protein sequence of TenS. However, based on the protein sequence published in NCBI, the two changed amino acids are not present in an active site of the protein. Sequencing of the original pEYA·TenS plasmid used for the LR recombination showed the same pattern of mutation (Figure 2.36). Since this plasmid was used to make the pTYGS·*argB*·*tenS* plasmid that was proven before to be active in *A. oryzae* strains, the transformation of TR.KO strain with this plasmid was continued.

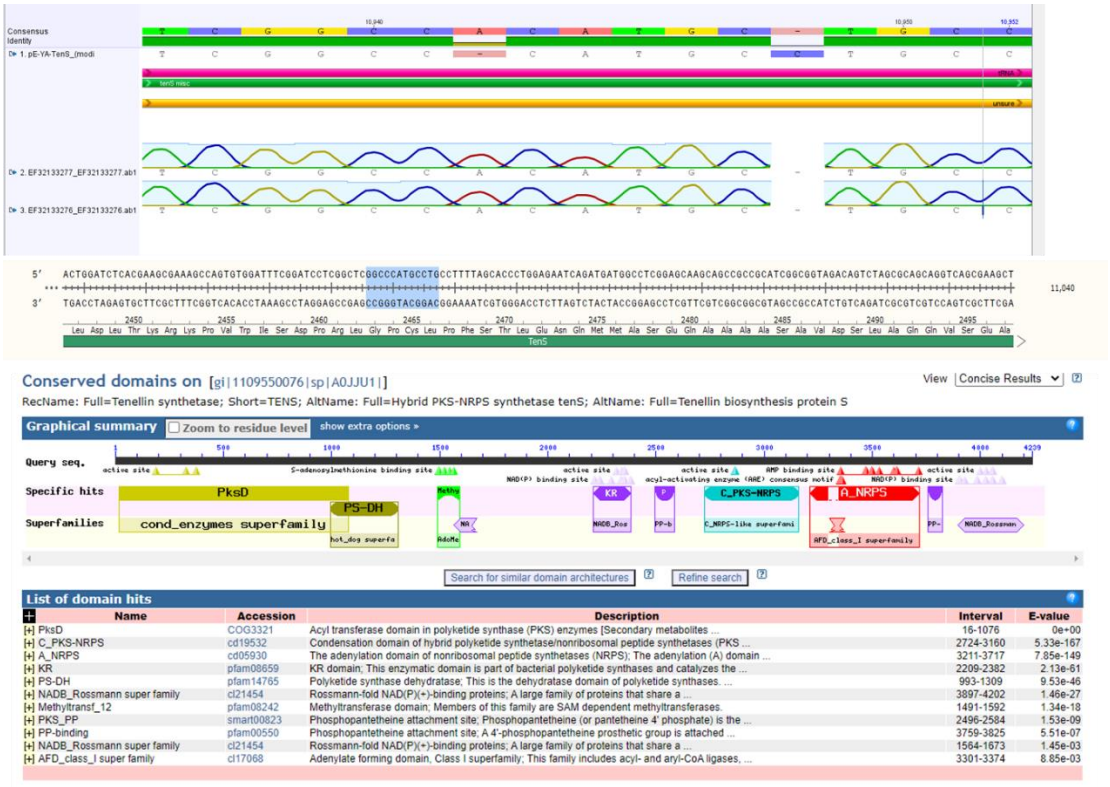


Figure 2.36 Sequencing of the KR domain in the plasmid showed two different bases between the two plasmids (MSIII174 (pTYGS·*pyr4*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC*·P_{TReno}) and pEYA·TenS) and the original sequence of the gene. Conserved domain of *tenS* as published in the NCBI database showing these mutations are not in any active site of the gene.

Transformation of *TR.KO* (*T. reesei*·QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC*) with MSIII174 (pTYGS·*pyr4*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC*·P_{TReno}) was done using the PEG-mediated transformation of protoplasts (Section 5.2.4) and resulted in the production of nine different transformants (A-I). The transformants were selected three times on minimal media and finally transferred into PDA plates as a rich media. Cultivation of five out of the nine transformants on DPY + 1 % glucose for three days and extraction showed the production of compound pretenellin A **77** in all of the tested transformants (Figure 2.37). However, the production levels were very low compared to the production level of the compound in *A. oryzae* transformants of the same genes.[102]

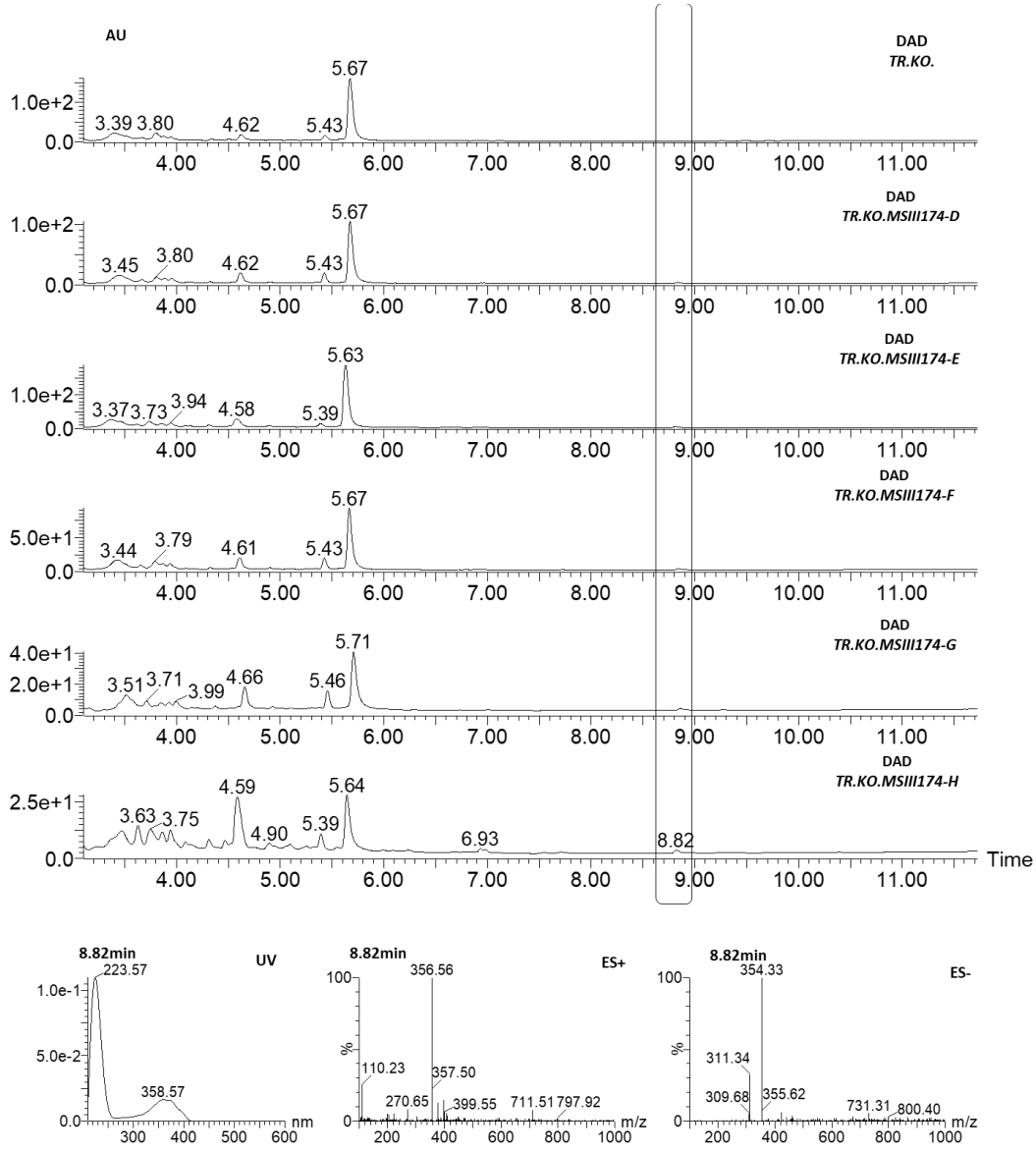


Figure 2.37 LCMS traces of different colonies of *TR.KO.MSIII174* (*T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC*) transformants on DPY+ 1% glucose after 3 days of cultivation.

A time course for the production of pretenellin A **77** from the best producing strain *TR.KO.MSIII174*-H (*T. reesei*·QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC colony H*) in DPY+ 1% glucose media, showed that the production of **77** starts after 72 hours of growth. After 96 h of growth, the production level of **77** decreases and a new peak at T_R = 7.4 min with nominal mass of 371 increases in intensity, which is 16 mass units higher than pretenellin A **77** (Figure 2.38). This might indicate an oxidised product of pretenellin A **77**. However, due to the very low production levels, the isolation of pretenellin A **77** and the other compound with nominal mass of 371 was not feasible.

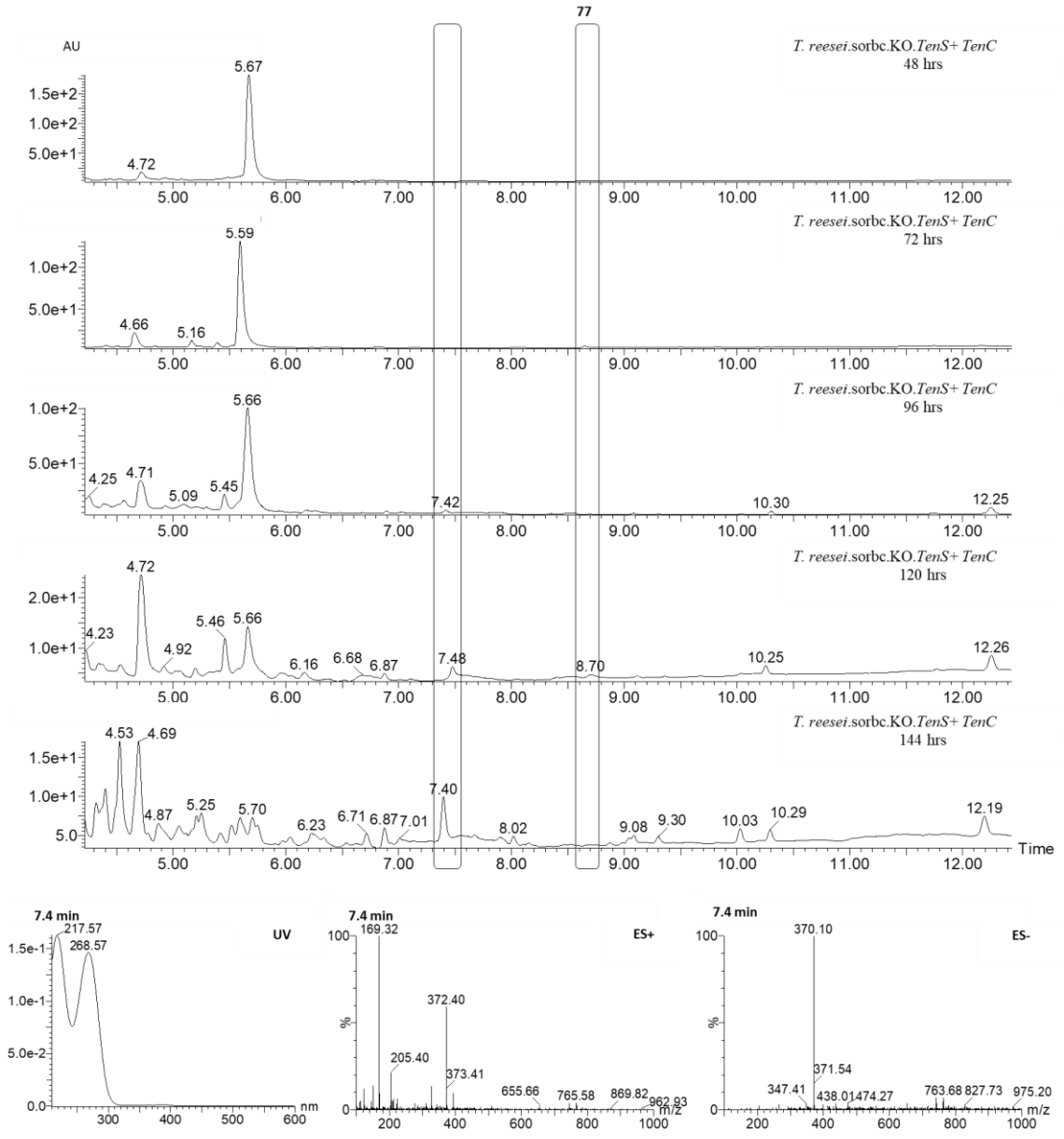


Figure 2.38 Time course of *TR.KO.MSIII174-H (T. reesei* QM6a Δ*tmus53* Δ*sorBC-P*$_{pdc2}$*-tenS-P*$_{cDNA1}$*-tenC colony H*) transformant in DPY+ 1% glucose showing that the production of pretenellin A **77** starts after 72 hrs and decreases after 96 hrs with an increase in the production of a compound with nominal mass 371.

In an attempt to optimize the cultivation conditions to increase the production of pretenellin A **77**, cultivation of the best producing strain *TR.KO.MSIII174*-H (*T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·*P*$_{pdc2}$·*tenS*·*P*$_{cDNA1}$·*tenC colony H)* was done on five different media and extracted after four days of growth (Figure 2.39). Pretenellin A **77** was produced in all the tested media but in very low concentration. The best producing media as shown in Figure 2.39 are DPY + 1 % glucose and CMP media. In comparison

to DPY + 1 % glucose, the DPY media showed higher concentration of the other compound with nominal mass 371 (T_R = 7.5 min), which might indicate that the presence of glucose in the media could decrease the level of oxidation of the produced compound **77**.

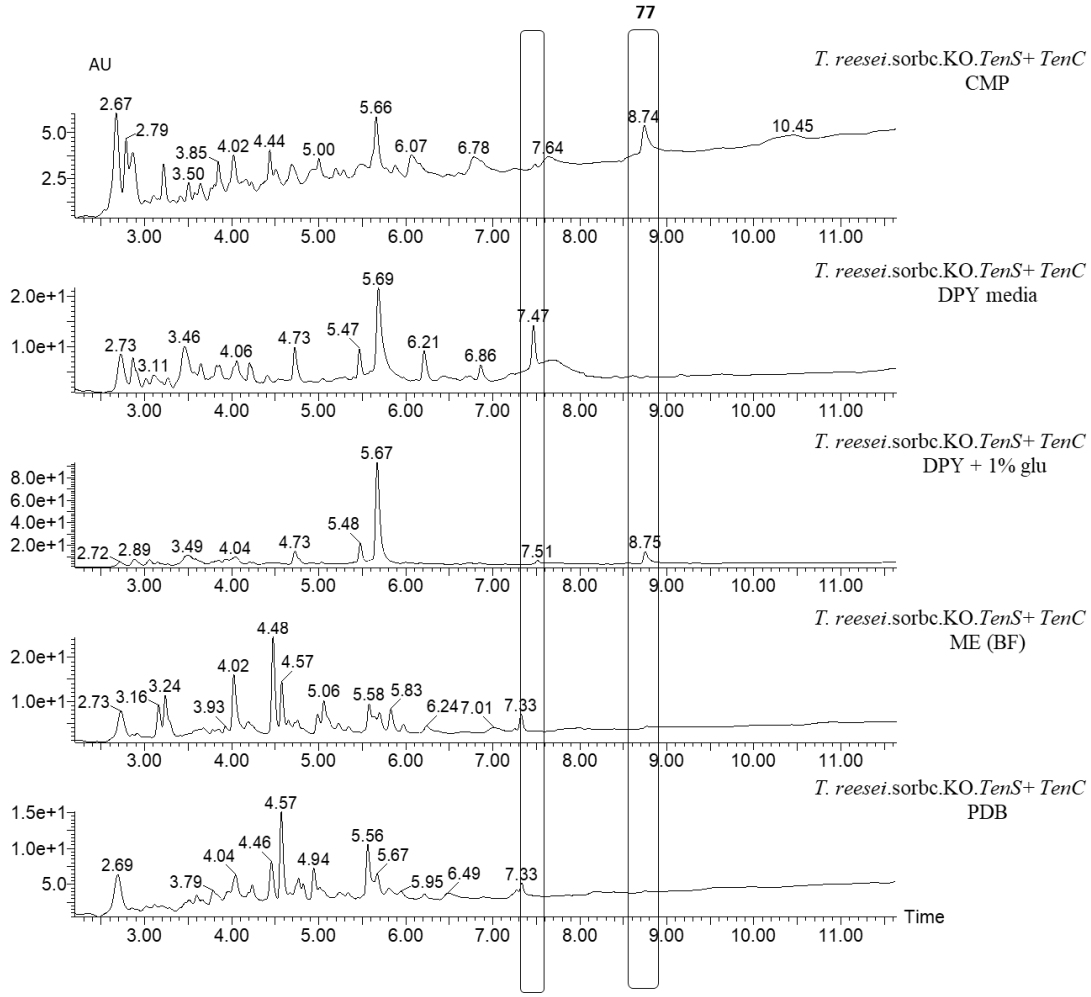


Figure 2.39 DAD chromatograms of the best producing strain of *TR.KO.MSIII174*-H (*T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC colony H*) on different media.

PCR analysis of the gDNA of *TR.KO.MSIII174* (*T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC*) transformants showed the correct insertion of *tenS* and *tenC* in all of the tested transformants (Figure 2.40).

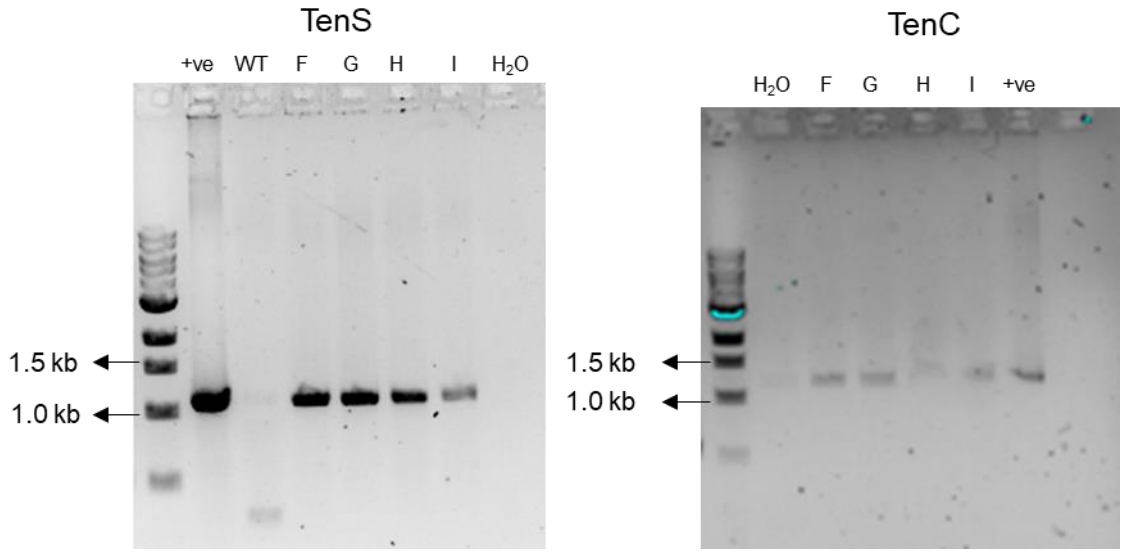


Figure 2.40 Genetic analysis of gDNA of *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC* (TR.KO.MSIII174-Colonies F-I) showing the correct integration of *tenS and tenC* genes in all of the tested transformants. Plasmid MSIII174 (*pTYGS*·*pyr4*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC*·P_{TReno}) was used as positive control and H_2O and/or TR.KO (WT) as a negative control.

The best producing transformant *TR.KO.MSIII174*-H (*T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC colony H*) was then cultivated on autoclaved banana peels (Section 2.3.4.1) to test the ability of the transformed strain to grow on waste materials and produce the expected compound. The LCMS chromatogram of the transfromant after 12 days of growth showed the production of the expected compound **77** (Figure 2.41). Nevertheless, further optimization of the conditions is still required to identify the best producing conditions and optimize the yield of the target compound.

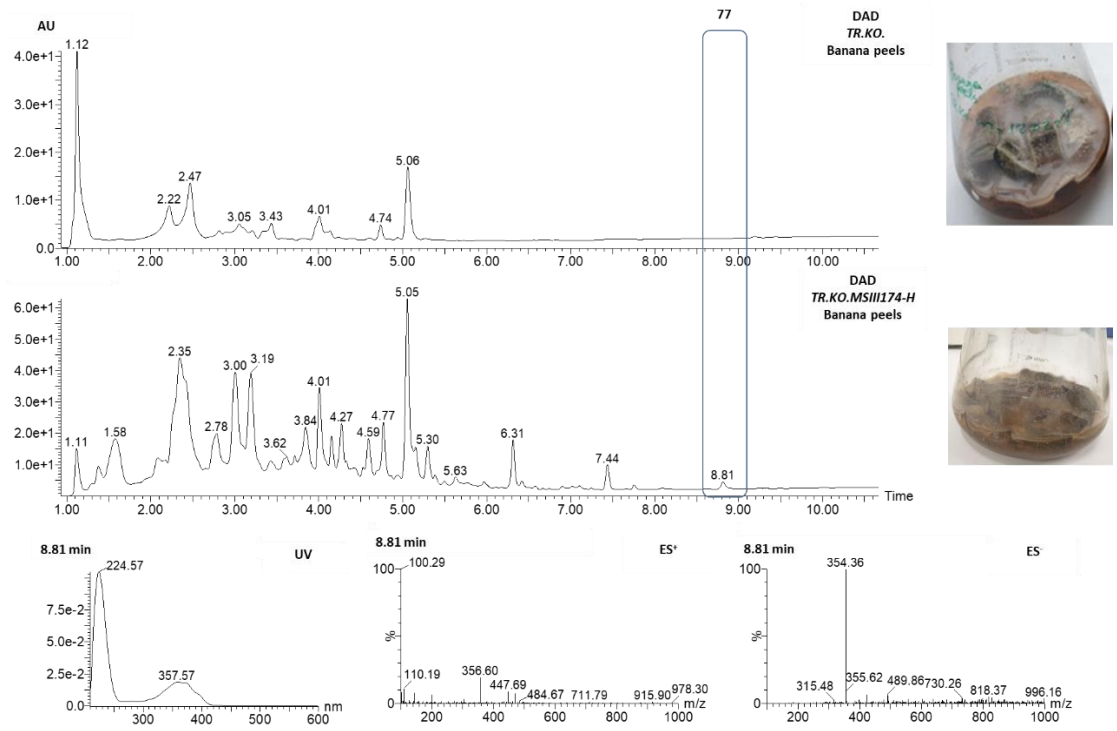


Figure 2.41 DAD chromatogram of transformant *TR.KO.MSIII174*-H (*T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC colony H*) in banana peels in comparison to TR.KO as the wild type showing the production of the expected compound pretenellin A **77**.

2.3.6 Testing the New Transformants on Straw

Straw is a form of agricultural waste that could be used as animal feed, although not the primary feed choice due to its high cellulosic and phytic acid content.[163] Attempts to use wheat straw as a substrate to grow *T. reesei* for the production of proteins were successful.[164] Therefore, the two main transformants produced in this work: *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1 colony C* (*TR.KO. MSIII163*-C) and *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*tenS*·P_{cDNA1}·*tenC colony H* (*TR.KO.MSIII174*-H) were grown on straw as a sole carbon source to check the ability of the strains to produce the expected natural product on straw. Barley traw was donated from a horse stable in Hannover and 100-200 ml of pure water were added to 5 g of the cut straw in 500 ml flasks. The flasks were then autoclaved and left to cool before cultivation of the fungal strains. *TR.KO.MSIII163*-C strain was cultivated on 5 g straw in 100 ml pure water and *TR.KO.MSIII174*-H was cultivated on 5 g straw in 200 ml water. The flasks were then incubated at 28 °C for 14 days before extraction with ethyl acetate. The extract from *TR.KO.MSIII163*-C showed the production of the expected compound 3-methylorcinaldehyde **40**, but the *TR.KO.MSIII174*-H extract showed no production of pretenellin A (Figure 2.42). This might be either due to the low production level of

pretenellin A **77** by the strain or due to the use of 200 ml pure water instead of only 100 ml in the preparation of the media.

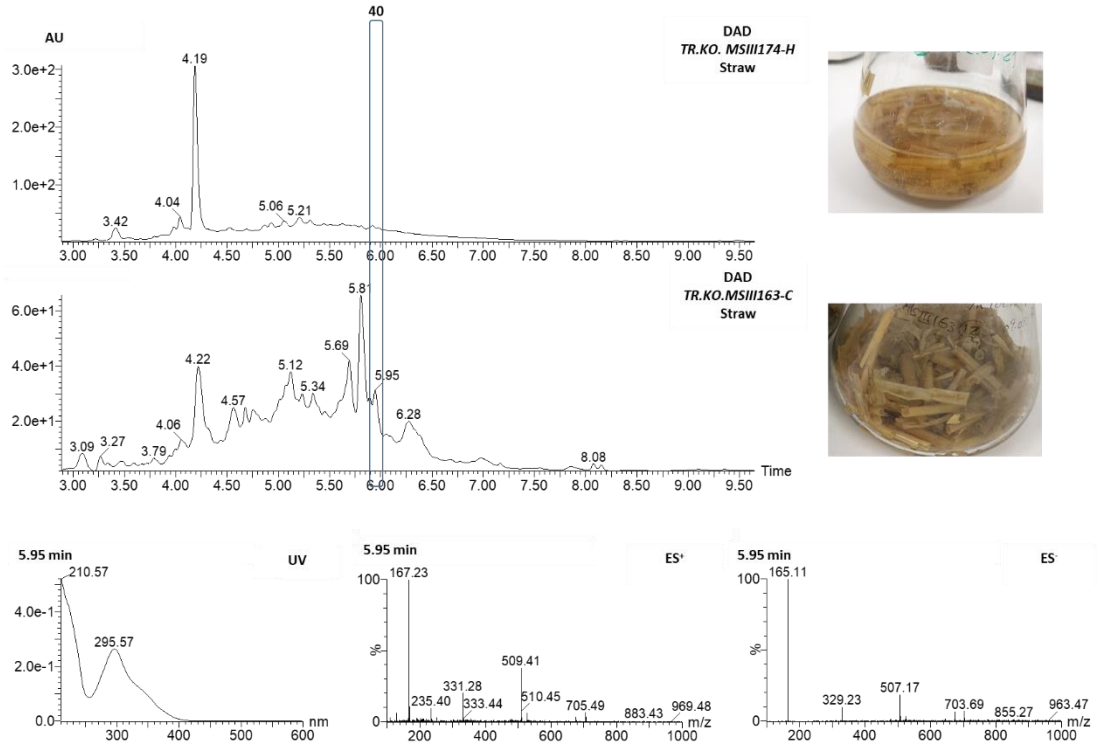


Figure 2.42 DAD chromatograms of the two *T. reesei* transformants on straw as media showing the production of **40** from *TR. KO.MSIII163*-C (*T. reesei* QM6a・Δ*tmus53*・Δ*sorBC*・P_{pdc2}・*aspks1 colony C*) strain.

2.4 Discussion

Due to its impressive ability to produce high amounts of cellulases, many researchers have developed *T. reesei* as a heterologous host for protein production. Its safety and high production capacity were very appealing and led to an increased attention of *T. reesei* in the last decades, especially after the publication of its full genome sequence in 2008.[165] Therefore, many toolkits for transforming *T. reesei* and for the expression of proteins and reporter genes have been investigated and published. This includes the development of different auxotrophic strains to facilitate the transformation, identification of many native promoters to increase and control protein expression and different methods for transformation.

2.4.1 Testing the Available Strains, Plasmids and Promoters

Despite this great interest in *T. reesei* as a heterologous host for protein production, almost no research has been done on *T. reesei* as a heterologous host for natural product production. Therefore, in this research, the suitability of *T. reesei* as a heterologous host for the expression of fungal genes for the production of various natural products was tested. The first step was testing the different strains, plasmids and promoters available in the group by expressing the nr-PKS *aspks1* in *T. reesei* using two main promoters: P_{amyB} from *Aspergillus oryzae*; and P_{pdc} from *T. reesei*. The plasmid with the native *pdc* promoter was already constructed in the group, while the plasmid with P_{amyB} was constructed by changing the selection marker of the plasmid RSI96-1-6 available in the group with the *pyr4* selection marker from *T. reesei*.

The resulting two plasmids were used to transform *T. reesei Δtmus53·Δpyr4* strain and *T. reesei Δtmus53·Δasl1*. Unfortunately, no production of any *aspks1*-related products was observed when the gene was expressed under P_{pdc}, while only very minor production of the **79** was observed when the gene was expressed under P_{amyB} (Figure 2.43). This was further confirmed by RNA extraction and reverse transcription, where *aspks1* was shown to be only transcribed under the activity of the P_{amyB} and no transcription when it was under P_{pdc1}.

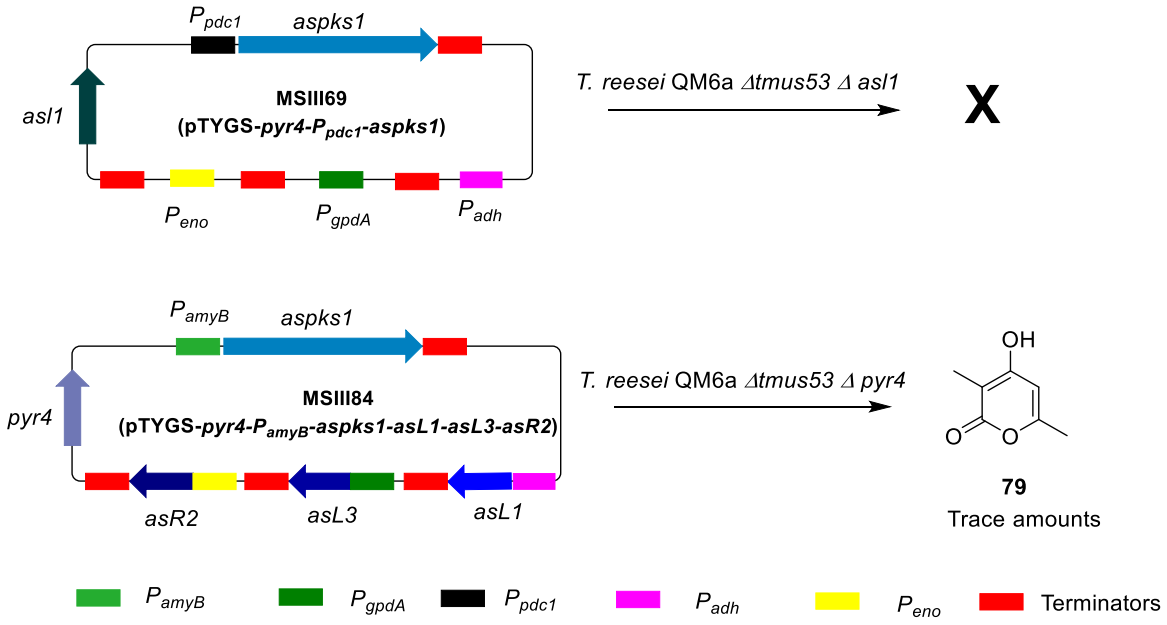


Figure 2.43 Overview of the activity of the *aspks1* gene under the native promoter P_{pdc1} and the *A. oryzae* P_{amyB}.

2.4.2 Knock Out of the Sorbcillin Gene Cluster to Construct a Strain with Cleaner Background

For P_{amyB}, since the inducing media (DPY) was used for all the experiments done with this promoter, the very low production levels could be explained since P_{amyB} comes from *Aspergillus oryzae* genome and therefore might not be a good promoter for *T. reesei*. Another reason might be the high levels of production of sorbicillins, the main class of natural products produced by *T. reesei*. This high level of production might lead to consumption of the acetyl-CoA and malonyl-CoA building units and hence lower the amounts available for other PKS such as *aspks1*.[152] The core structure of the sorbicillinoids family of NP is sorbicillin **87a**, which is the product of two polyketide synthases (*sorA* and *sorB*). Deletion of only one of these two PKS should be enough to abolish sorbicillinoid production (Figure 2.44).

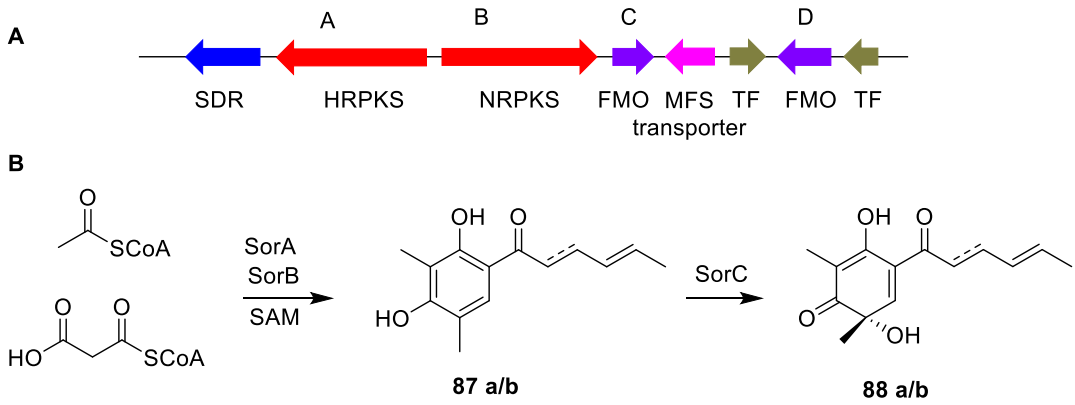


Figure 2.44 Sorbicillin biosynthesis: **A**, the biosynthetic gene cluster of sorbicllin; **B**, the first steps in the biosynthesis of the sorbicillinods family of natural products.[152]

However, the next enzyme in the biosynthetic pathway, SorC, has been reported to have a promiscuous activity,[156–158] as the gene was able to oxidize several substrates with some structural similarity to sorbicillin (Figure 2.45). Therefore, the presence of this gene might affect the heterologous system by performing unintended reactions and deletion is important to develop a better transformation host.

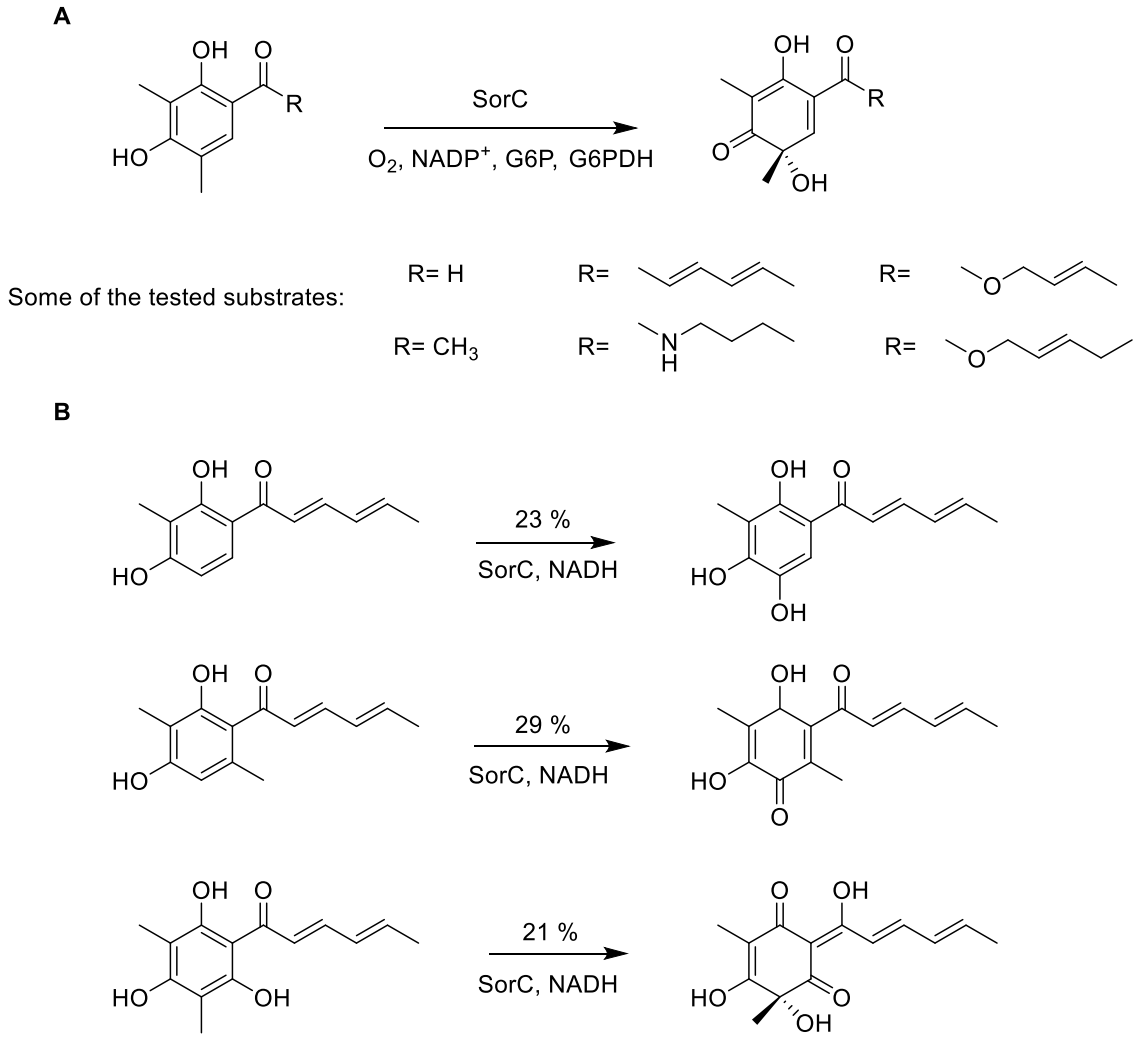
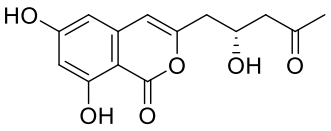


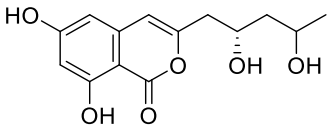
Figure 2.45 SorC activity on different substrates: **A**, SorC ability to dearomatize different substrates with different substituents on C-1;[156] **B**, SorC ability to oxidize different sorbicillin derivatives with different substitution pattern on the aromatic ring system.[158]

Hence, a plasmid to knockout the *sorB* and *sorC* genes was constructed using the hygromycin resistance cassette and flanking regions of the two genes. The two genes were knocked out using the bipartite method and one successful knockout transformant was obtained and was named *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* strain (abbreviated as *TR.KO*). The knockout was then confirmed genetically and chemically, where the TR.KO transformant was cultivated on nine different media to establish its chemical profile and to ensure the absence of any level of production of sorbcillin under different cultivation conditions. Many new compounds were produced in trace amounts from the new KO strain, probably due to the increased availability of acetyl-CoA and other building blocks. However, due to the very low concentration of all these compounds, their isolation and structure elucidation were not feasible. Nevertheless, only one compound

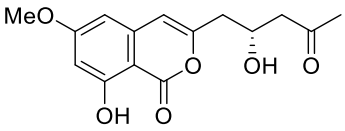
was produced in good concentration and was therefore isolated and identified to be citreoisocoumarin **79**. Although this compound was not previously reported from *T. reesei*, it was reported to be isolated from the sponge derived fungus *Trichoderma* HPQJ-34 and is a common NP isolated from different fungal species.[159] It is usually produced together with its closely related isocoumarin derivatives, such as citreoisocoumarinol **89** and 6-methyl-citreoisocoumarin **90**.[159,166] Bioinformatic anaylsis of all PKS genes from *T. reesei* (Section 3.3.1) showed the presence of a PKS with high similarity (> 67%) to the nr-PKS (PkgA) from *A. nidulans*, which was reported to produce citreoisocoumarin and related compounds.[167]



Chemical Formula: $C_{14}H_{14}O_6$
Exact Mass: 278,08
86



Chemical Formula: $C_{14}H_{16}O_6$
Exact Mass: 280.09
89



Chemical Formula: $C_{15}H_{16}O_6$
Exact Mass: 292.09
90

Figure 2.46 Some examples of citreoisocoumarin derivatives produced in fungi.

The resultant strain *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* still contains the *pyr4* auxotrophy and therefore was used for all later heterologous expression experiments. The strain also had a seemingly clean background as no sorbicillin-related compounds were produced under all tested cultivation conditions. Even the newly produced isocoumarin-related compounds were produced in very low concentrations and their level of production changed by changing the media.

2.4.3 Construction of a New Vector with the Correct P_{pdc} Sequence (P_{pdc2})

The lack of transcription under the originally used P_{pdc1} shown earlier was surprising, especially since this promoter was previously reported to be a strong constitutive promoter for *T. reesei*.[116] Therefore, the whole *aspks1* gene in the plasmid MSIII69 was sequenced to exclude the presence of any point mutation that might have caused a frame shift in the gene and resulted in a stop codon. The sequencing showed the correct sequence of the gene with no point mutation. After careful inspection of the P_{pdc1} and comparison between its sequence and the published sequence, it was obvious that P_{pdc1} used in the plasmid is different from the published one. Apparently, the *T. reesei* genome contains at least two different genes annotated as encoding pyruvate decarboxylases, both

on the NCBI and JGI sequences. Unfortunately, the *pdc1* gene chosen by a former group member to use its promoter was not the one reported in literature. This explains the inability of the gene to be transcribed under the chosen native P_{pdc1}.

Hence, a new vector with the correct P_{pdc2} (MSIII158) was constructed by adding the active native P_{pdc2} to allow good levels of transcription of the heterologously expressed gene. The two flanking arms of the *pyr4* gene with a *Pac*I restriction site between them were added to the vector to increase the chances of gene insertion in the *pyr4* locus and to separate the *pyr4* gene from the *pdc* promoter to facilitate the recycling of the *pyr4* gene later. *Pyr4* gene recycling could be done later by excision of the gene with homologous recombination without affecting the expression cassette. The pTYGS·*argB* vector was used as the backbone for the construction of MSIII158 and therefore, MSIII158 has all the advantages of the pTYGS plasmids including the Gateway® recombination system and the ability to shuttle between yeast and *E. coli* to facilitate gene insertion by yeast homologous recombination. The *aspks1* gene was used to test the new system, by construction of MSIII163 (pTYGS·*pyr4*·P_{pdc2}·*aspks1*) through Gateway® LR recombination between the entry vector pEYA·*aspks1* and the newly constructed MSIII158 (pTYGS·*pyr4*·P_{pdc2}). Transformation of *TR. KO* strain with the linearized expression vector MSIII163 (pTYGS·*pyr4*·P_{pdc2}·*aspks1*) resulted in the construction of eight different transformants. Chemical analysis of the transformants showed the production of three products of *aspks1* gene in seven out of the eight transformants but in different levels of production. The three compounds produced where isolated and their structures confirmed by HRMS and NMR to be compounds **79**, **40**, and **41** (Figure 2.47).

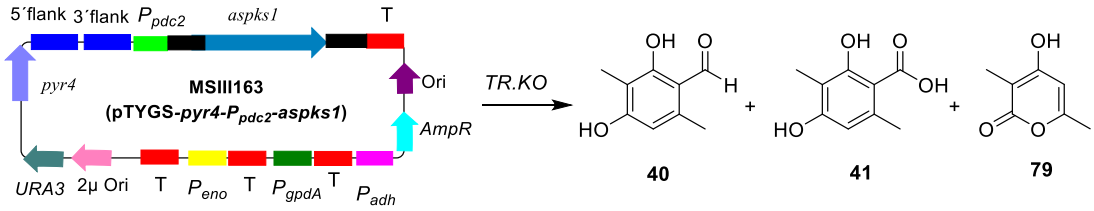


Figure 2.47 Overview of the compounds produced by the expression of *aspks1* gene under the activity of the correct *pdc* promoter in *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* strain.

However, site-specific integration of the gene usually results in a comparable levels of production of the substrate.[120] Therefore, the significant difference in the production levels of 3-methylorcinaldehyde between different transformants indicated that the gene

might not be inserted in a site specific manner. PCR analysis of the best producing transformant *TR.KO.MSIII163*-C to check for site-specific integration of the gene in the *pyr4* locus indicated that the gene is not inserted in the expected locus and was probably integrated elsewhere in the gDNA by ectopic integration. However, seven out of the eight transformants showed production of 3-methylorcinaldehyde **40**, which represents ~87% success rate, which is comparable to the one reported by Derntl *et. al*.[120] for the site-specific integration into *T. reesei* QM6a·*Δtmus53*·*Δpyr4* (86 ± 7) and also for ectopic integration of genes in *T. reesei* genome (76 ± 14).

Careful inspection of the *pyr4* flanking regions used in the plasmid showed that after linearizing the plasmid with *Pac*I, the two flanking regions would be in the wrong directions and therefore would prevent site-specific integration. This would explain the difference in gene expression levels in different transformants, as the ectopic integration could lead to gene integration randomly in the gDNA and in different copy numbers. This can lead to gene insertion in different locus with different levels of expression. However, the flanking regions were still of importance for further experiments as they separate the *pyr4* gene and the expression cassette and therefore can be used later on for recycling of the marker *pyr4* without affecting the expression of the already transformed genes.

The production of 3-methylorsellenic acid **41** by *T. reesei aspks1* transformants might be due to either different splicing of the second intron of *aspks1* gene (discussed in section 2.1.1) since the gene was amplified from the gDNA of *S. schorii*[155] and therefore has the two introns. Another reason could be the oxidation of 3-methylorcinaldehyde **40** by native *T. reesei* genes. However, different production levels of 3-methylorcinaldehyde **40** in different media might indicate the implication of native genes in the oxidation of the product 3-methylorcinaldehyde **40** to 3-methylorsellenic acid **41**. Nevertheless, addition of glucose to the media seems to inhibit the oxidation of 3-methylorcinaldehyde (as shown in the difference between the two media, DPY and DPY + 1 % glucose). This might be explained by previous reports on the activity of the transcription factor Xpp1 (Xylanase promoter binding-protein 1) that controls the balance between primary and secondary metabolites production in *T. reesei*. This gene acts as a positive regulator of primary metabolism and also as a repressor of secondary metabolism.[168] *Xpp1* was also reported to be up-regulated in the presence of D-glucose as carbon source.[169] Hence, up-regulation of this gene in the presence of glucose in the media can result in down regulation of different genes involved in secondary metabolism and hence inhibit the oxidation of 3-methylorcinaldehyde.

2.4.4 Testing the Ability of *T. reesei* Transformants to Produce Natural Products on Waste Materials

Since much research has shown the ability of *T. reesei* to produce high levels of cellulases and other cell wall degrading enzymes, the ability of the transformed *T. reesei* strains to grow on different waste materials and to produce chemical compounds was tested. The best producing 3-methylorcinaldehyde strain (*TR.KO.MSIII163*-C) was cultivated on different agricultural waste material, such as potato peel, orange peel, banana peel and kiwi peel. The transformant showed good production levels of 3-methylorcinaldehyde **40** and 3-methylorsellenic acid **41** in all the media after 14 days of growth except for orange peel, where no growth was observed for more than one month. Quantification of the two chemical compounds **40** and **41** showed that in potato peel media the transformant could produce up to ~128 $mg \cdot kg^{-1}$ dry weight of 3-methylorcinaldehyde **40** and ~ 243 $mg \cdot kg^{-1}$ dry weight of 3-methylorsellenic acid.

The strains cultivated on orange peel was left in the incubator at 28 °C for longer period of time together with a negative control flask containing un-inoculated orange peels treated in the same manner. After almost four months, the fungi showed growth on orange peels where the negative control flask showed no growth or contamination. Extraction of the transformant *T. reesei* QM6a·Δ*tmus53*·Δ*sorBC*·P_{pdc2}·*aspks1* colony C (*TR.KO.MSIII163*-C) showed the production of the expected compounds; 3-methylorcinaldehyde **40** and 3-methylorsellenic acid **41** in the extract. Research on the antifungal activity of essential oils of orange peels showed that increasing essential oil concentration inhibited the growth of *Aspergillus flavus, Penicilllium digitatum* and *P. italicum*.[170,171] Hence, the delayed growth of the fungal strains on orange peels could be attributed to the antifungal activity of the essential oils of the orange peels. So, air-drying of the orange peels prior to autoclaving might be beneficial to allow for the evaporation of all essential oils and decrease the time of cultivation of the fungal strains.

The system produced shows great opportunity for waste valorization, where agricultural waste materials such as potato peel, kiwi and banana peel can be simply used as a cultivation media for the producing strain. The waste media did not require any major pretreatment, as only adding pure water to the peel and autoclaving them was enough and the fungal strain was able to grow and produce the required chemical compounds without any further requirements such as adding trace elements or salts to the media.

2.4.5 Adding More Promoters to the Plasmid to Allow for the Expression of More Than One Gene Simultaneously

To expand the new system, more native promoters of *T. reesei* were added to the plasmid using yeast homologous recombination. The new vector (MSIII172) contains three native constitutive promoters (P_{TReno}, P_{cDNA1} and P_{pdc2}) and one *A. nidulans* constitutive promoter (P_{gpdA}). The constitutive promoters were used instead of the stronger inducible promoters involved in cellulase production to allow for constitutive production of the produced compounds under different cultivation conditions. The constructed plasmid was used to express the megasynthase PKS-NRPS gene (*tenS*) together with its *trans*-acting ER (*tenC*) in *T. reesei* by adding the *tenS* gene under the P_{pdc2} and *tenC* gene under the P_{cDNA1} (plasmid MSIII174). The ability of the system to express more genes efficiently regardless of the integration site in the DNA were proved by transforming the vector MSIII174 without linearization. Transformation with the circular vector yielded many transformants. After three rounds of selection, nine transformants were chosen for further analysis. LCMS analysis of six of these transformants after three days of cultivation showed the production of pretenellin A **77** (Figure 2.48) in all of the transformants but in a very low concentration. A time course for the production of pretenellin A **77** from the best producing transformant showed that the best production level is after 72 hrs. Therefore, the transformant was then cultivated for more than 72 hrs in different media and the extracts showed that all the media resulted in production of pretenellin A, but the best production levels were in CMP and DPY + 1 % glucose media.

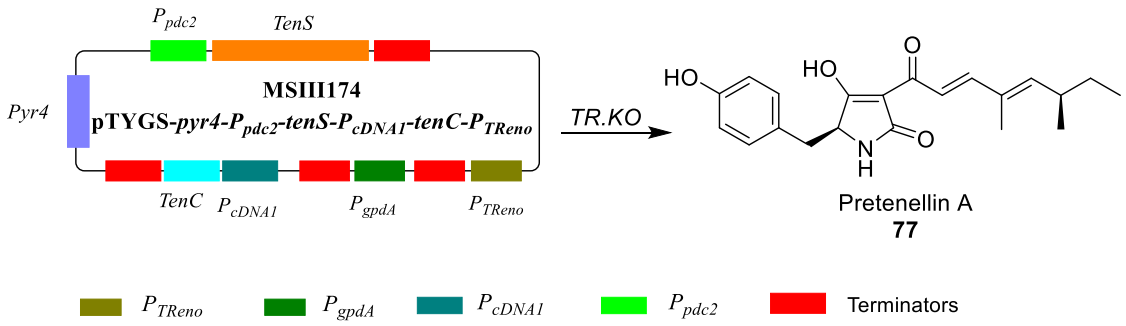


Figure 2.48 Overview of the expression of *tenS* and *tenC* genes under the activity of the native promoters P_{pdc} and P_{cDNA1} in *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* strain.

The best producing transformant was then cultivated on banana peels to test the ability of the system to produce the expected compound on waste material. The strain showed indeed the production of pretenellin A **77** but in very low concentration. Due to time

restriction, further experiments to optimize the cultivation conditions and to upscale the production of compound **77** on waste materials were not conducted. However, the ability of the system to produce a more complicated compound such as pretenellin A **77** shows that *T. reesei* is a promising fungal host for heterologous expression of BGC.

2.4.6 Testing the Ability of the New Strains to Produce the Expected Natural Products on Straw

Production of lignocellulosic-based chemicals and biofuels through biorefinaries have gained much attention lately. However, the main obstacle that prevent the advancement of biorefinaries to the commercial scale is the high cost of the enzymes used to convert the biomass into fermentable sugars.[172] *T. reesei* is one of the main industrial producers of such biomass-degrading enzymes and therefore research has focused on regulation of secretion of these biomass-degrading enzymes to reduce the cost of its production.[173] Therefore, producing *T. reesei* strains that could produce specified natural products on waste material would provide a cheap, safe, sustainable and greener alternative to production of such compounds by chemical synthesis. The newly produced strains (*T. reesei. aspks1* and *T. reesei. TenS + TenC*) were able to grow and produce the expected compounds on different types of agricultural wastes such as potato peels and banana peels.

However, its ability to grow and produce compounds on more challenging type of waste material, such as straw was still to be investigated. Previous reports have shown the ability of *T. reesei* to grow and produce proteins on pre-treated wheat and barley straw.[72,174] Hence, barely straw was used as a media to test the ability of the new strains to produce the expected natural products on lignocellulosic waste materials. The strain expressing Aspks1 was able to produce the expected compound 3-methylorcinaldehyde **40** on barely straw, however, the other strain expressing TenS and TenC was unable to produce pretenellin A under the tested condition. However, further optimization of the cultivation conditions and testing different methods of straw pre-treatment to optimize the production was not possible due to time constrains. Nevertheless, the barley straw used was only autoclaved with water without any other pretreatment and still one of the new strains was able to produce the heterologously expressed natural product on this media. This represents a great potential for the use of *T. reesei* as heterologous host for natural product production and for waste valorization.

2.5 Conclusion and Outlook

Overall, this study paves the way for the use of *T. reesei* as a microbial cell factory for natural product production and to establish the use of synthetic biology in converting low-value waste materials into valuable chemicals. The constructed system offers a high throughput and low-cost way of assembling large and multiple genes into vectors using yeast homologous recombination and Gateway® system and expressing them in *T. reesei*. To apply the constructed platform system, two types of megasynth(et)ases were tested; PKS and hybrid PKS-NRPS. All the extractions were done with ethyl acetate which is considered as one of the least environmentally harmful organic solvents (environmentally preferred solvent).[175,176] The production of natural products by this system offers a more economic and greener alternative to chemical synthesis of complex natural products.

The system constructed in this work can theoretically express up to four genes simultaneously, however, only two promoters in this plasmid have been used and proved experimentally. Future work to test the other two promoters in the plasmid could result in strains expressing up to four genes in *T. reesei*. Recycling of the *pyr4* selection marker can also be done by construction of plasmids with flanking regions to the *pyr4* and *URA3* genes to knockout these genes from the gDNA of the transformants. Later selection of the true *pyr4* KO strains can then easily be done by selecting the resultant strains on 5-FOA.

To increase the number of selectable markers in the *TR.KO* strain, constructing plasmids with flanking regions to the knocked out genes *sorB* and *sorC* can not only lead to site-specific integration into a well-transcribed locus, but will also remove the hygromycin resistance cassette and hence lead to a strain with extra dominant marker. Using *A. oryzae* plasmid system *pTYGS-ble* to create new plasmid systems to *T. reesei* using other native or non-native promoters can result in building a heterologous system that can take up to eight genes simultaneously.

Recent attempts to grow *T. reesei* on different waste materials such as textile waste, paper sludge and many other waste materials have shown the ability of the fungus to grow on a wide range of substrate.[150,177] Pretreatment of all of these waste materials usually results in better growth of the fungus. Therefore, further experiments on the system would be testing the ability of the new fungal system to produce different compounds from different pathways on a wider range of waste materials including rice

husks, sugarcane bagasse, papers, textile waste, *etc*. Also testing different solvents for extracting natural products from the waste to be more efficient and Eco-friendly. Further optimization of the system to enhance the production level of compounds on waste either by pretreatment of the waste material or by choosing different inducible promoters could also be beneficial.

3 BGC Activation In *Trichoderma reesei*

3.1 Introduction

Trichoderma spp. have found wide application in agriculture as biocontrol agents due to their ability to combat different plant pathogens. They exert their biocontrol activity mainly by secreting cell-wall degrading enzymes, secretion of antimicrobial and antifungal secondary metabolites and by competing for nutrients with the pathogen.[44] They have been used as biofertilizers and biopesticides and have even been commercialized such as Binab-TF.[44,178] However, until recently, most of the research done on *Trichoderma* spp. has mainly focused on their ability to secrete cell-wall degrading enzymes such as cellulases and hemicellulases especially in *T. reesei*, the work horse of this genus.[44] These enzymes have found different applications in biofuel production and different industries such as textiles and pulp industry. However, the increasing need to find new alternative for synthetic and chemical pesticides has raised the interest in the secondary metabolites from different *Trichoderma* spp. as efficient biopesticides.

3.1.1 Secondary Metabolites from *Trichoderma* spp

Recently, interest in secondary metabolites from *Trichoderma* spp. has been rising and many new natural products with fascinating chemistry have been isolated from different marine and terrestrial *Trichoderma* spp. These compounds include a wide variety of polyketides, terpenes, polypeptides and PKS-NRPS products.

Many PKS-NRPS products have been isolated from different *Trichoderma* spp. (Figure 3.1) with wide range of biological activities such as the antimycobacterial hirsutellone B **91**,[179] the antimicrobial trichodin A **92**, B **93**,[180] trichosetin **94**,[181] and the antifungal harzianopyridone **34**.[182] Some of them presented unusual chemistry such as trichodermone **95** with an unprecedented 7/5/6/5 tetracyclic nucleus.[183] Some other PKS-NRPS products have also been reported, such as deacetyl-18-deoxy cytochalasin H **96**[184] and harzianic acid **36.** Bioinformatic analysis of different *Trichoderma* spp. showed the presence of many PKS and PKS-NRPS BGC, however, only a few of them have been fully investigated biochemically. These examples include; harzianopyridone **34**,[185]

harzianic acid[186] **36** and the tetronate trihazone D **55**[74] BGC from *T. afroharzianum* t-22 (previously classified as *T. harzianum* t-22).

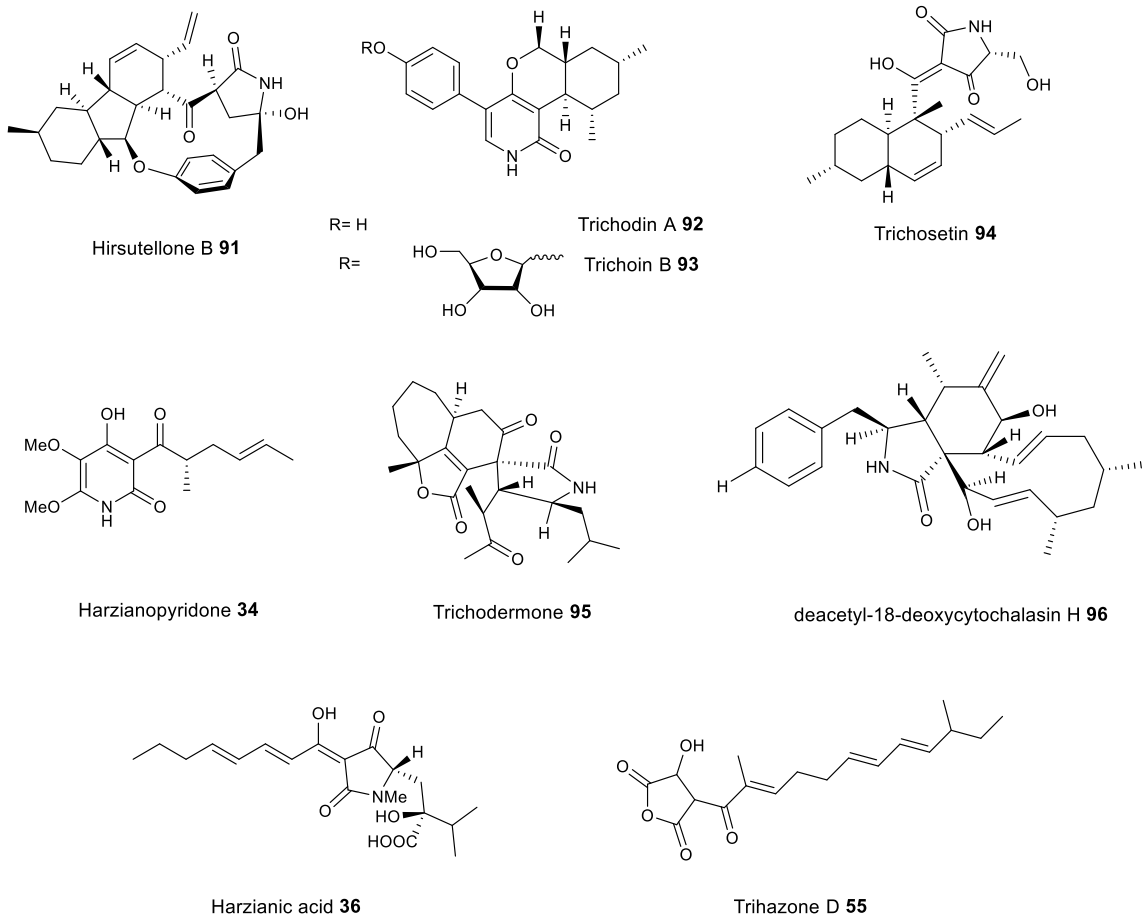


Figure 3.1 PKS-NRPS related compounds isolated from different *Trichoderma* spp.

3.1.2 Secondary Metabolites from *Trichoderma reesei*

Despite all the attention that *T. reesei* has gained and all the research done on this fungus as a model organism, most of the research done has mainly focussed on the protein expression system and the cellulases and other cell-wall degrading enzymes. Neverthelesss, some research has been done on the secondary metabolites and biosynthetic gene clusters in *Trichoderma reesei*. For example, *T. reesei* CSR-T-3 showed a biocontrol activity against *Fusarium* wilt disease of banana caused by the infection of the banana with *Fusarium oxysporum*. *T. reesei* was proven to have a positive effect on the fruit yield and weight in addition to the suppression of the pathogen (*Fusarium oxysporum*).[187] The mechanism of the biocontrol activity of *T. reesei* against the pathogenic fungus was reported to be the result of increase in the production of cell wall degrading enzymes and the secretion of antifungal secondary metabolites.[187]

Therefore, exploring the biosynthetic potential of *T. reesei* might lead to the identification of new compounds that can be used as leads for pesticides and fertilizers.

Based on the Reaxys search for all natural products isolated from *T. reesei*, 30 low Mwt compounds have been reported so far (**97-126**), of which more than 80 % are sorbicillin-related compounds (Figure 3.2-3.4)

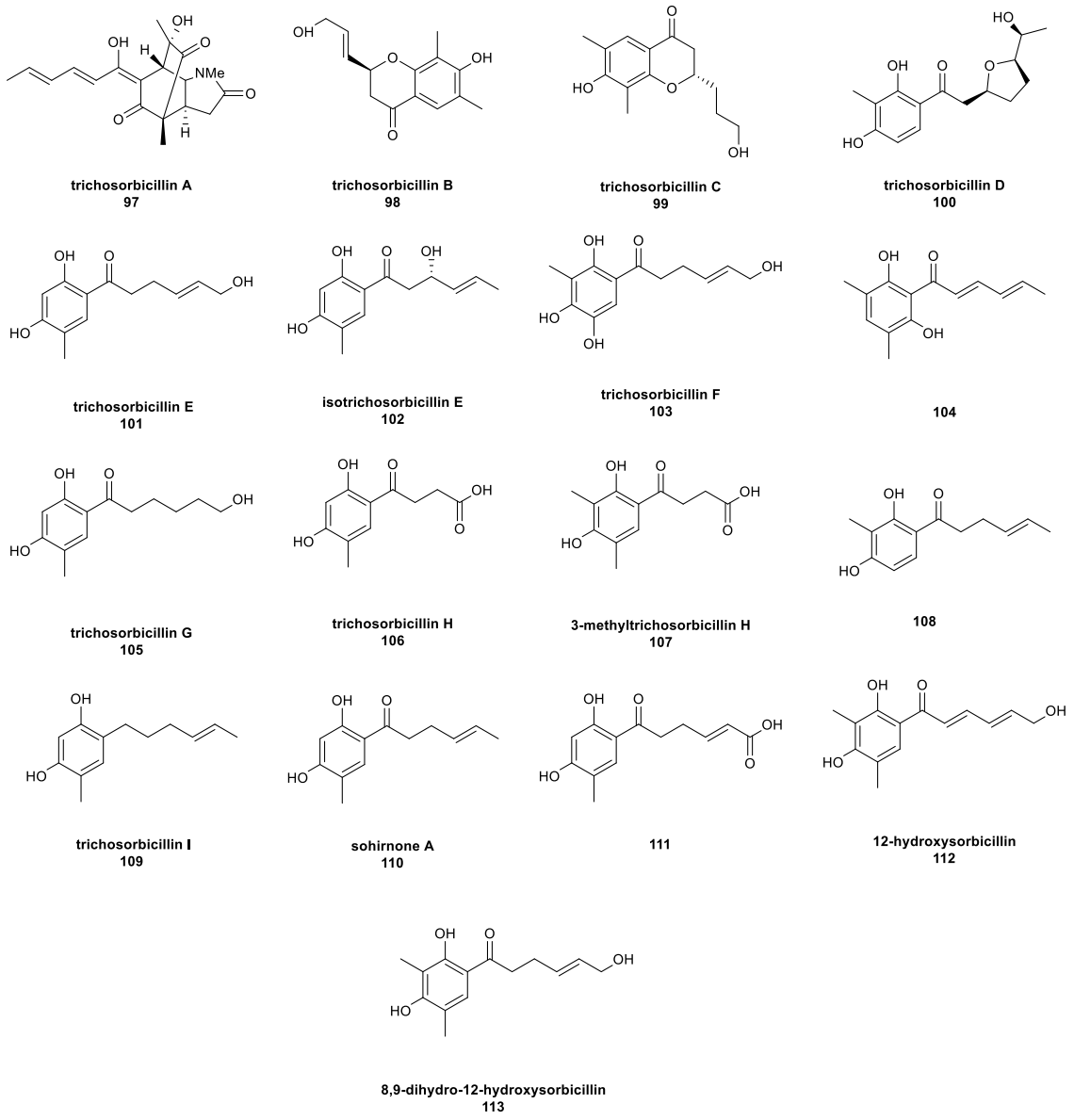


Figure 3.2 Sorbicillin-related compounds reported from *T. reesei* (monosorbicllinoids) according to reaxys.

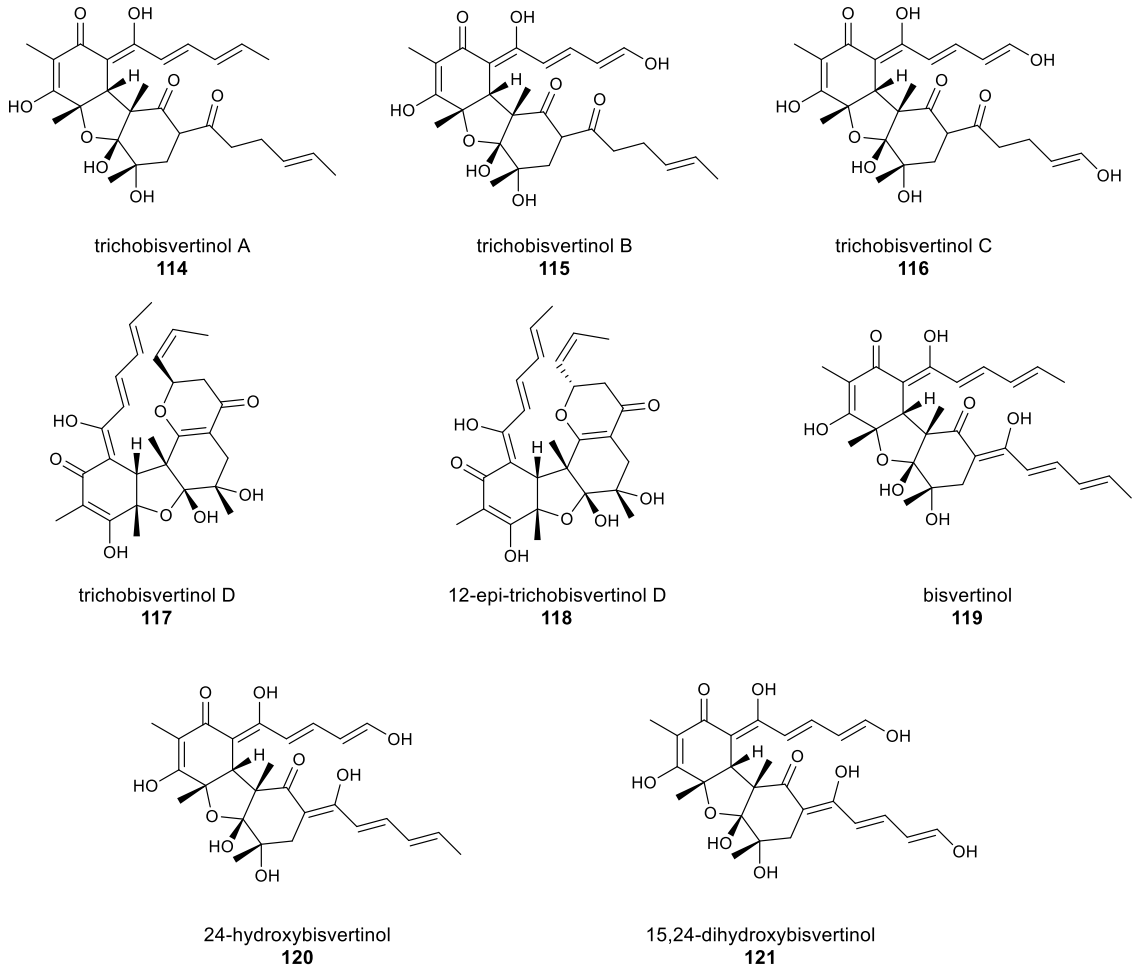


Figure 3.3 Sorbicillin-related compounds reported from *T. reesei* (bisorbicllinoids) according to reaxys.

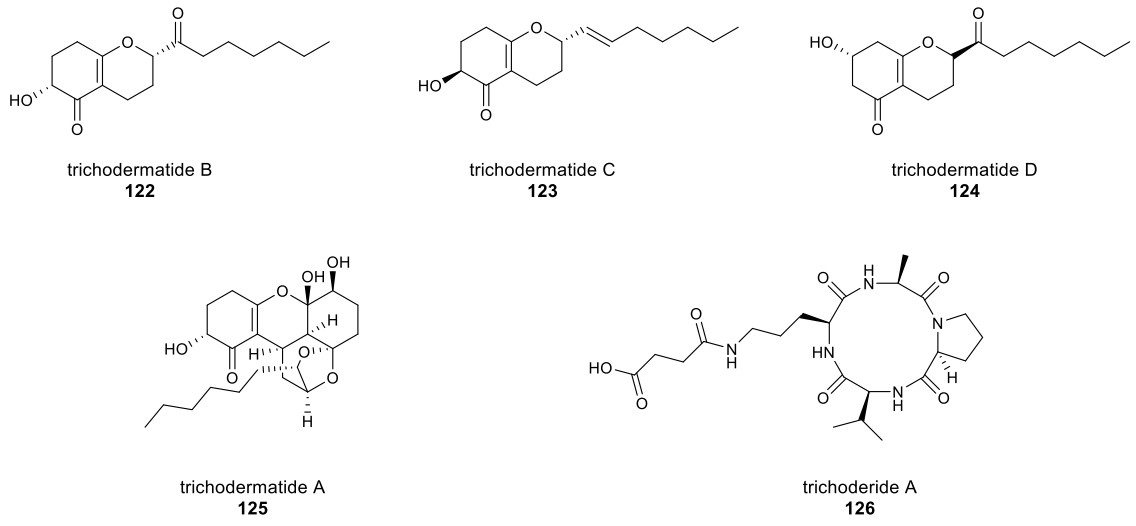
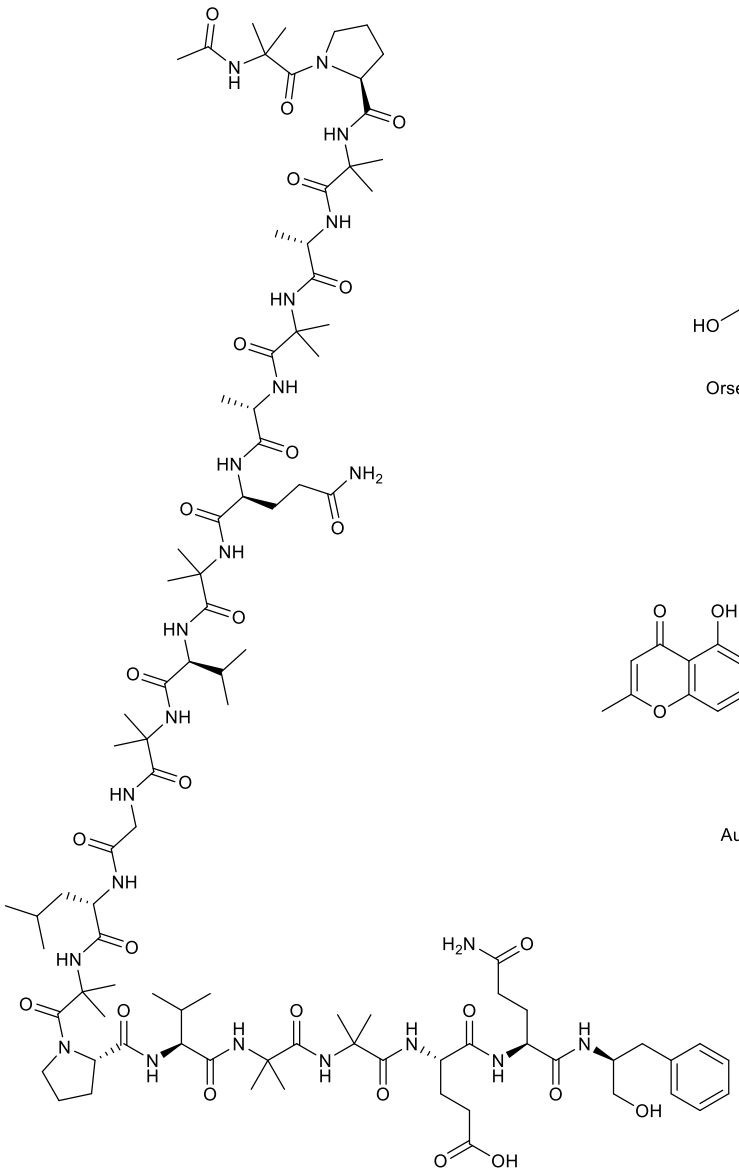


Figure 3.4 Other natural products reported from *T. reesei* according to reaxys.

Several methods have been previously used to link the BGC in fungi to their metabolites, which include transcription factor (TF) activation, heterologous expression, and knock-out experiments. Activation of the transcription factors has gained much attention lately especially in BGC activation in *Trichoderma* species, that includes activation of both global and pathway-specific TF.[44] However, only one BGC from *T. reesei*, *i.e.* the sorbicillin BGC (*sor* BGC) has been proven biochemically. One of the main challenges in exploring the full biosynthetic potential of different *Trichoderma* spp., especially *T. reesei*, is the tight control over secondary metabolism in this fungus.[44] Derntl *et. al*.[168] reported that xylanase repressor transcription factor Xpp1 (xylanase promoter binding protein 1) acts as a switch between primary and secondary metabolism in *T. reesei* and it was reported to be upregulated by the presence of glucose in the media. Therefore, adding glucose to the media can lead to a down-regulation of different genes involved in secondary metabolism.

Not only changing the media composition can affect the production of secondary metabolites, but also minor changes in other cultivation conditions such as light and pH. The effect of the light on *Trichoderma* spp., especially *T. reesei* has shown the tight control on secondary metabolism genes by the effect of different genes and transcription factors.[188–190] Also many so-called global transcription factors have been reported to have effects on the secondary metabolism in *T. reesei* such as CRE1, LAE1, Vel1 and Vel2. Even some TF that were thought to be pathway-specific TF was shown later to have a more generalized function such as *ypr2* in the sorbicillin (*sor*) BGC. YPR2 is one of the two TF present in the sorbicillin BGC that was found later to affect the levels of orsellenic acid **127** and alamethicin **128** production in *T. reesei* and its expression was regulated by light and carbon source.[189] Other genes in the vicinity of the *sor* BGC such as *usk1* (Unique sor cluster kinase) and *gpr8* (class II G-protein coupled receptors) were found to have an impact not only on the production levels of sorbicillinoids but also on the production of other secondary metabolites.[44,188,191]



Orsellinic acid **127**

Aurofusarin **129**

Alamethicin **128**

Figure 3.5 Examples of some natural products isolated or predicted to be produced by *Trichoderma* spp.

These findings indicate the complexity and very tight control over secondary metabolites production in *T. reesei*. Therefore, attempts to explore the secondary metabolite potential in *T. reesei* have been challenging and not always successful. For example, after the activation of 11 different pathway-specific TF in *T. reesei*, only two TF (TRIREDRAFT_102497 and 102499) resulted in the production of 78 new compounds when compared to the WT, most of which were chemically related to sorbicillinoids.[192] Later experiments on the two PKS in the *sor* BGC confirmed the function of these two PKS in the production of sorbicillin and sorbicillin-related compounds.[151,152,192,193]

However, all the other TF activation attempts resulted in no change in the chemical profile of the transformants in comparison to the wild type.

Attempts to heterologously express 23 different PKS, NRPS and hybrid PKS-NRPS genes from *T. reesei* in the heterologous host *Aspergillus nidulans* were not very successful. Although the correct insertion of 22 out of the 23 genes in different transformants were confirmed genetically, no new secondary metabolites were detected. Only eight of these transformants showed change in the pigmentation and colony morphology, but chemical analysis of these transformants did not reveal the production of any new compounds.[192] In this heterologous expression experiment, the two hybrid PKS-NRPS genes from *T. reesei* were heterologously expressed in *A. nidulans* and resulted in no new metabolites.

In other work, an attempt to knock out *pks4* (TRIREDRAFT_82208) in *T. reesei*, which is orthologous to the pigment-forming PKS associated with aurofusarin **129** biosynthesis in *Fusarium graminearum*, resulted in abolition of the production of green conidial pigments, but the chemical structure of the secondary metabolite produced by this cluster was not elucidated.[194] Also in an attempt to knock out the gene TRIREDRAFT_121121, which encodes a candidate fungal regulatory protein located next to the β-glucosidase gene (*cel3d*), resulted in a sorbicllinoid hyper-producing *T. reesei* strain. The gene 121121 was presumed to play a role in cellulase production in *T. reesei* and therefore a knock out experiment was planned to study the effect of this gene on cellulase production. However, off-target mutagenesis resulted in a strain capable of constitutive hyper-production of sorbicillinoids independent of culture conditions, pH, temperature and carbon source.[195]

3.1.3 Sorbicillin Biosynthetic Gene Cluster

The best studied BGC in *T. reesei* is the sorbicillin gene cluster. Sorbicillinoids (also called vertinoids)[196] are complex cyclic polyketides that were isolated for the first time by Cram and Tishler in 1948 as a yellow pigment contaminant during clinical penicillin production from the mold *Penicillium notatum (syn. P. chrysogenum)*.[44,197] This family of secondary metabolites now includes over 90 sorbicillinoids isolated from different fungal strains from both terrestrial and marine environments.[196,198,199] In 2002, Abe *et al* proposed the term “sorbicillinoid” to describe the hexaketide compounds isolated from

bisorbicillinoid-producing strains and having a sorbyl chain.[44,200] Sorbicillins are biosynthesized by a NRPKS/HRPKS hybrid cluster. Following the release of the linear hexaketide chain, a putatively non-PT-mediated Knoevenagel cyclization of the free aldehyde results in sorbicillin (Figure 3.6), the aromatic precursor to members of this class. Oxidative dearomatization by SorC then results in the production of sorbicilinol, a hydroxylated derivative of sorbicillin. Finally, SorD, a putative Diels-Alderase, catalyses the intermolecular Diels-Alder and Michael type oxidative dimerization reactions, resulting in the production of bisorbicillinoids **119 a/b** and **130 a/b/c** and other dimeric derivatives such as spirosorbicillinols **131 a/b**. SorD was also reported to catalyse epoxidation of sorbicillinol **88 a** to produce epoxysorbicillinol **132 a**.[152]

An attempt to understand the function of the flavin-containing dehydrogenase (*sorD*) in the *sor* BGC in *T. reesei*, a knock out experiment was performed in the native host. This experiment resulted in the accumulation of dihydrosorbicillinol **88 b** and reduction in the levels of sorbicillinol **88 a**. Therefore, the authors suggested that SorD might catalyse the reduction of the 2´,3´-olefin in the sorbyl side chain of dihydrosorbicillinol **88 b** to yield sorbicillinol **88 a**.[153] Although the authors noticed that no later sorbicillinoids including dimeric sorbicillinoids were produced in the *sorD*-Knockout strain, they did not link these findings to the function of SorD.[44,153] However, further studies on SorD by heterologous expression in *A. oryzae* showed that SorD catalyses both the intermolecular Diels-Alders and Michael reactions to form sorbicillinoids.[152]

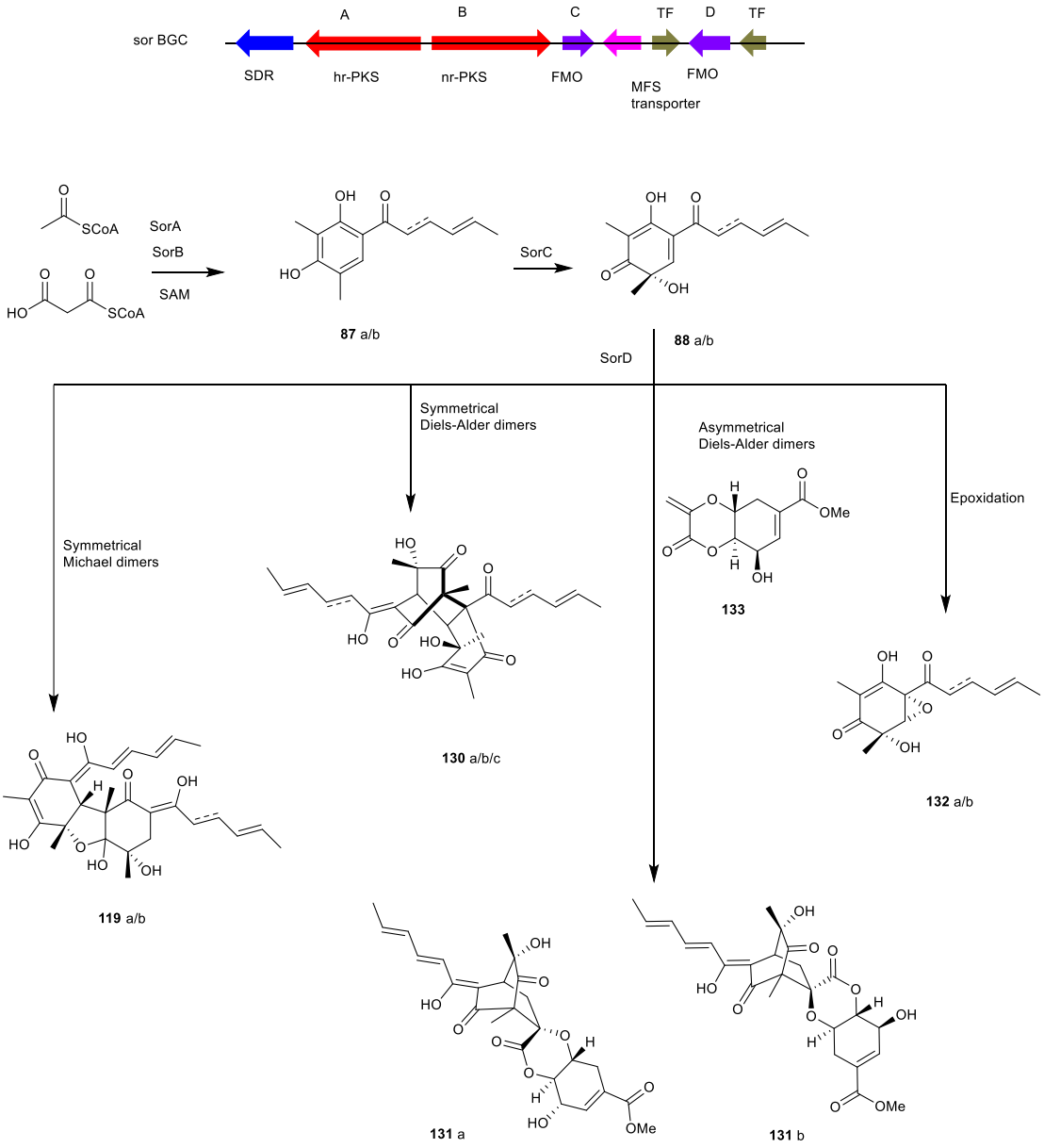


Figure 3.6 The sorbicillin BGC and the proposed biosynthetic pathway of sorbicllinoids.

3.2 Aim of the Work

The aim of this work is to unravel the biosynthetic potential of *Trichoderma reesei* by performing a bioinformatic analysis to identify all potential PKS and hybrid PKS-NRPS genes and gene clusters in *T. reesei*. Then choosing one of the cryptic or silent BGC to identify its secondary metabolite by performing experiments in the native host and also by heterologous expression experiments in *Aspergillus oryzae*.

3.3 Results

3.3.1 Bioinformatic Analysis to Identify All PKS and Hybrid PKS-NRPS Genes in *T. reesei*.

Bioinformatic analysis was done using a manual search strategy based on the Blastp algorithm[201] using a template protein sequence of known non-reducing PKS (AMS38559.1) from *Penicillium chrysogenum* against the protein database of *T. reesei* QM6a.[202] The protein database for *T. reesei* QM6a was downloaded from NCBI (Refseq assembly accession number GCF_000167675.1, V2.0). Blastp search was done using Geneious Prime® 2021.2.2 using standard parameters (BLOSUM62 matrix, Max E-value 0.05, gap cost 11 1). The Blastp search resulted in 33 hits representing 21 different proteins (Section 7.2). Careful inspection of the conserved domains on NCBI of 21 proteins showed only 13 proteins to have the minimum catalytic domains for PKS (KS and AT domains), which include 11 PKS genes and two hybrid PKS-NRPS genes (Figure 3.7).

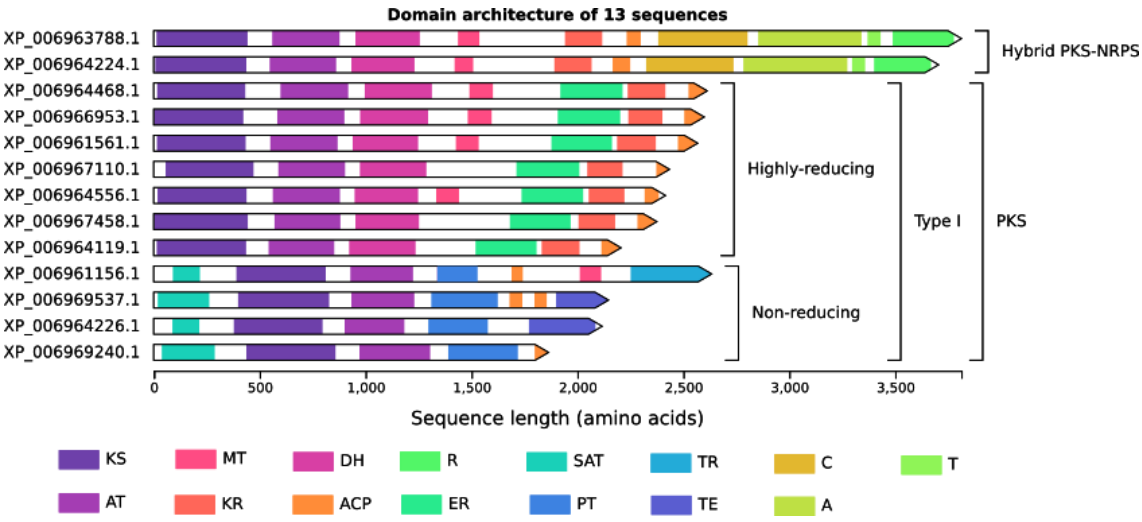
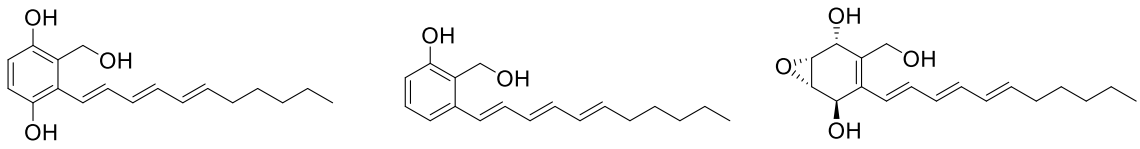


Figure 3.7 Domain architecture of all the predicted PKS and hybrid PKS-NRPS from *T. reesei* represented by synthaser software.

Table 3.1 A list of all PKS and hybrid PKS-NRPS in *T. reesei* genome and the closest transcription factors:

	Protein accession No.	Gene No. (TRIREDRAFT)	Function	Length of the protein (aa)	Closest hit by Blastp[201] (% identity at protein level, swissprot database)	Closest TF and distance to it	Ref.
1	XP_006961561	73618	hr-PKS	2567	SorA (100%)	102497 (13.76 kb) 102499 (~20kb)	152,193
2	XP_006964119	59482	hr-PKS	2205	T-toxin PKS (34.6%)	121164 (25.62 kb) 105917 (49.18 kb)	203
3	XP_006964224	59315	PKS-NRPS	3704	Fusarin C (61.2%)	105784 (9.88 kb)	204
4	XP_006964468	106272	hr-PKS	2612	Alternapyrone (32.6%)	121310 (27.9 kb) 106259 (34.67 kb)	205
5	XP_006964556	60118	hr-PKS	2415	Beauveriolides (*cm3B*, 33.6 %)	77234 (~ 15 kb).	206
6	XP_006966953	65172	hr-PKS	2598	6-hydroxymellein (*cdmE*, 36%) Ochratoxin (37%)	112783 (4.3 kb) 65070 (12.9 kb) 109230 (18.6 kb) 26871 (25.1 kb) 79725 (14.06 kb) 71823 (21.2 kb)	207,208
7	XP_006967458	65891	hr-PKS	2374	Radicicol (*radS1*, 32%)	--------	209
8	XP_006963788	58285	PKS-NRPS	3812	Illicicolin-H (*iliA*, 66.4%)	72993 (11.6 kb)	89
9	XP_006967110	65116	hr-PKS	2434	Trichoxide Vir A (87.1 %)	--------	210
10	XP_006969537	82208	nr-PKS	2146	conidial pigment synthase (76%)	112129 (48.86 kb)	211
11	XP_006961156	73621	nr-PKS	2633	SorB (100%)	102497 (~5kb) 102499 (11.36)	152,193
12	XP_006969240	81964	nr-PKS	1863	PkgA (67.1 %)	111742 (2.456 kb) 111755 (24.4 kb)	167
13	XP_006964226	105804	nr-PKS	2116	(AscC) ascofuranone/ascochlorin BGC	105805 (2.72 kb) 121130 (45.6 kb)	212

Out of all these megasynthases, only the two PKS genes responsible for the production of sorbicillinoids have been identified and confirmed biochemically in *T. reesei* (TRIREDRAFT_73618 and 73621, Table 3.1).[151,152,193] The hr-PKS XP_006967110 encoded by the gene TRIREDRAFT_65116 showed 87.1 % identity at the amino acid level to the recently investigated hr-PKS (VirA) from *Trichoderma virens* GC29-8 which is a part of a gene cluster that mediates the production of virensols **134- 135** and trichoxide **136**.[210]



Virensol A **134** Virensol B **135** Trichoxide **136**

Figure 3.8 Compounds produced by vir BGC from *Trichoderma virens* GC29-8

Although the *T. reesei* genome contains two hybrid PKS-NRPS genes, no PKS-NRPS related secondary metabolites have been reported so far from *T. reesei*. As shown in table 3.1, the two PKS-NRPS proteins show relatively high identity (> 60%) to proteins involved in the biosynthesis of Fusarin C and Ilicicolin-H **72**. Nevertheless, Ilicicolin-H **72** has not been reported to be produced by any *Trichoderma* spp. so far.[213] However, heterologous expression experiments done by the Tang group to explore the biosynthetic pathway of ilicicolin-H **72** in *penicillium variabile* (*syn. Talaromyces variabilis*) reported that different *Trichoderma* spp. possess gene clusters that show high similarity to the ilicicolin-H BGC.[214] Ilicicolin-H **72** is a broad-spectrum antifungal metabolite that was isolated for the first time in 1971 from the mycelia of the fungus *Cylindrocladium ilicicola*.[215,216] Although different *Trichoderma* spp. have wide application in biocontrol, the main active constituents behind their antifungal activity are still under investigation. Therefore, this cryptic BGC in *Trichoderma reesei* was chosen for further investigation.

Based on the NCBI prediction of the published *T. reesei* QM6a genome, the chosen BGC includes genes that encode a hybrid PKS/NRPS, a *trans*-acting ER, a putative Cytochrome P450 monooxygenase, an oxidoreductase, a TF and some other genes with unknown functions. The PKS/NRPS protein shows 66.4% identity to the ilicicolin-H PKS/NRPS protein from *Neonectria* sp. DH2.[89,201] However, ilicicolin-H was not reported before to be isolated from any *Trichoderma* spp.[213] In addition, the BGC of ilicicolin-H should contain a putative Diels-Alderase, which does not appear to be present in this BGC. Manual annotation of the genes in this BGC to determine the correct start and end of each gene in the cluster was first done by FungiSMASH,[217] which showed that the BGC might contain an unpredicted gene and showed also that the predicted PKS/NRPS has no introns, which is in contrast to the NCBI prediction that shows that the PKS/NRPS has two introns.

The genes from this cluster showed very high similarity to a closely-related cluster from *Trichoderma citriniviride*. Hence, the closely related BGC in *Trichoderma*

citrinoviride, for which the genome sequence was published in 2016, was used as a reference for comparison to re-annotate the cluster in *Trichoderma reesei*. The Artemis Comparison Tool[218] was first used to compare the two closely related BGC (Figure 3.9).

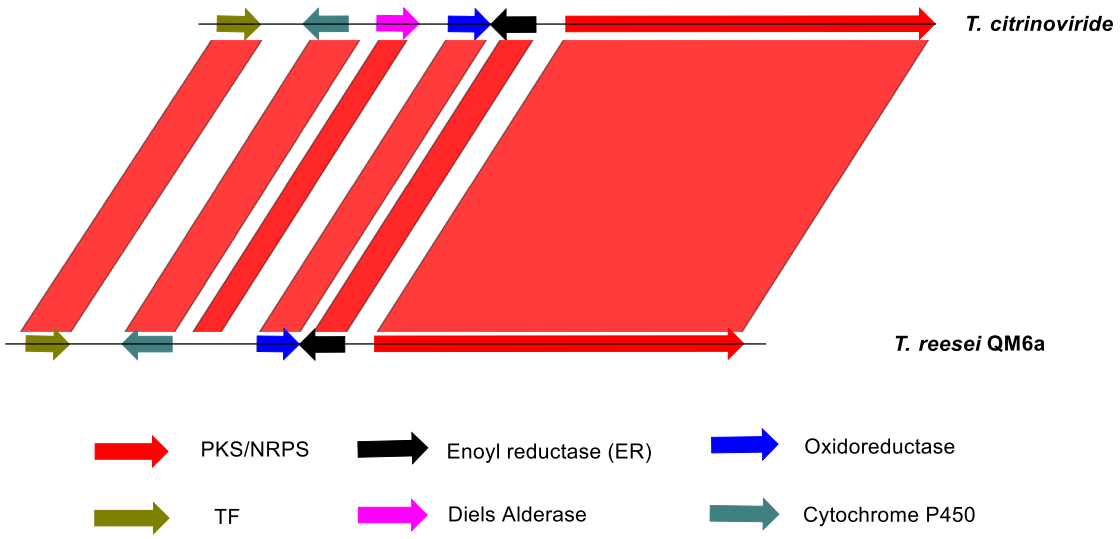


Figure 3.9 Comparison of homologous ilicicolin-H BGC from *T. reesei* and *T. citrinoviride*. Artemis ACT[218] was used for the pairwise comparison. Red lines indicate high similarity (> 80%).

The clusters indeed showed very high similarity and showed that one of the genes in the cluster had not been annotated in *Trichoderma reesei* QM6a BGC. Manual reannotation of all the genes using Geneious Prime® 2021.2.2 led to the identification of all the genes in the cluster. Protein blast[201] of the respective proteins showed high similarity (55-70%) to proteins from the ilicicolin-H BGC from *Penicillium variabile* and *Neonectria* sp.DH2 (Table 3.2).[89,214]

Table 3.2 Prediced function of each protein in the chosen BGC and their % identity to proteins with known functions in the swissprot database using Blastp.[201]

Gene number (TRIREDRAFT-number)r	***Protein name***	***Predicted Function (based on T. citrinoviride)***	***Closest hit by Blastp[201] (Swissprot)***	***% identity***	***Ref***
58285	XP_006963788.1	PKS-NRPS	*iliA* (ilicicolin-H)	66.4%	[89]
58289	XP_006963601.1	*Trans*-enoyl reductase	*iccB* (ilicicolin-H)	63.6%	[214]
76204	XP_006963789.1	NADH-dependant flavin oxidoreductase	*iccE* (ilicicolin-H)	54.9%	[214]
XX[2]	XXX[3]	*S*-adenosyl-L-methionine-dependent methyltransferase (*T.citrinoviride*)	Diels alderase *iccD* (ilicicolin-H)	55.9%	[214]
58953	XP_006963602.1	Cytochrome P450	*iliC* (iliciolin-H)	70.4%	[89]

The re-annotated *T. reesei* PKS/NRPS BGC was then used to identify homologous gene clusters from other fungal species using cblaster[219] (Figure 3.10). The results showed a highly similar BGC present in some other *Trichoderma* spp. such as the biocontrol fungus *Trichoderma virens* GV29-8, although ilicicolin-H was not reported before from any *Trichoderma* spp.[213]

[2] This gene was not annotated in the published genome of *T. reesei* QM6a and was identified by manual annotation.

[3] The BLASTp for this protein was done for the corresponding *T. citrinoviride* protein.

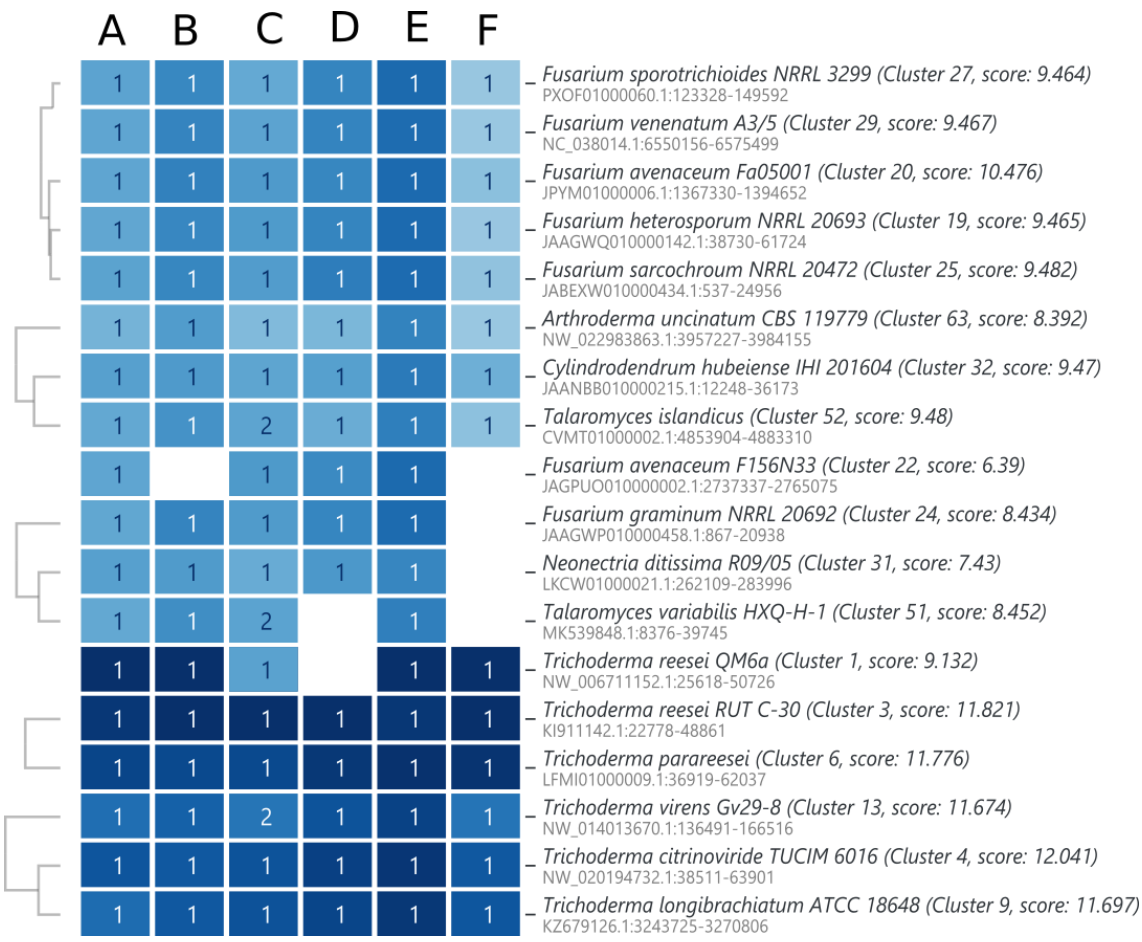


Figure 3.10 cblaster[219] results for clusters from other fungal species that are closely related to the identified PKS-NRPS cluster from *T. reesei*, Gene codes: **A**, PKS-NRPS; **B**, ER; **C**, oxidoreductase; **D**, *C*-MeT; **E**, P450; **F**, TF.

3.3.2 Gene Cluster Activation by Ectopic Expression of the Pathway-Specific TF

A transcription factor (TRIREDRAFT_72993) is present in the vicinity of the chosen biosynthetic gene cluster. Activation of this TF in *T. reesei* was facilitated by the developed toolkit for the expression of heterologous genes in *T. reesei* (Chapter 2).

3.3.2.1 Vector Construction for the Chosen TF and *T. reesei* Transformation

The chosen transcription factor; TRIREDRAFT_ 72993 is located 11.6 kb away from the PKS-NRPS (TRIREDRAFT_58285). The TF was amplified from the gDNA of *T. reesei* with flanking regions to pEYA plasmid using primers 1539 and 1540 by PCR using Q5 polymerase. The pEYA plasmid was cut using *Asc*I/NotI and yeast homologous recombination was used to construct pEYA·TRIREDRAFT_72993 (MSIII168A). LR recombination of the constructed entry plasmid with the expression vector

pTYGS·*pyr4*·P_{pdc} (MSIII158, Section 2.3.3) resulted in the construction of the expression vector pTYGS·*pyr4*·P_{pdc}·TF72293 (MSIV05, Figure 3.11).

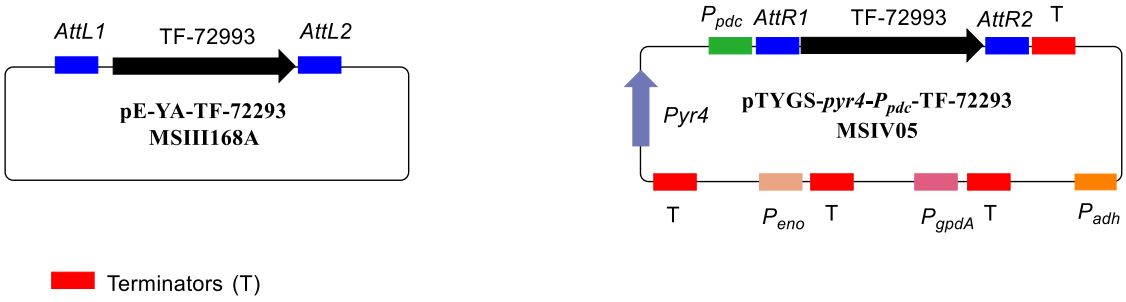


Figure 3.11 Fungal vectors used in the transcription factor activation of the cryptic cluster in *T. reesei*.

3.3.2.2 Homologous expression of TRIREDRAFT_72993 in *T. reesei*

The vector pTYGS·*pyr4*·P_{pdc}·TF72293 (MSIV05) was transformed into *T. reesei Δtmus53·Δpyr4·ΔsorBC* (Section 2.3.2) using the protoplast PEG-mediated transformation (Section 5.2.4). After three rounds of selections, eight transformants were chosen for further analysis. Five of the eight transformants were grown on DPY + 1% glucose media for 3 days at 28 °C and 110 rpm, extracted and then submitted to LCMS. When compared to the untransfomed *T. reesei* host strain, the LCMS traces of the transformants showed one minor new peak at 10.5 min and three other peaks corresponding to compounds present in very low concentration in the untransformed strain that showed enhanced production in the transformed strain (Figure 3.12-A). However, the compounds were produced in very low concentrations and that prevented their isolation and structure elucidation. Nevertheless, the new peak at 10.5 min showed Mwt corresponding to that of iliciclolin-H but different wavelength (Figure 3.12-B). Therefore, a heterologous expression experiment was planned for this gene cluster to identify the secondary metabolites produced by this BGC.

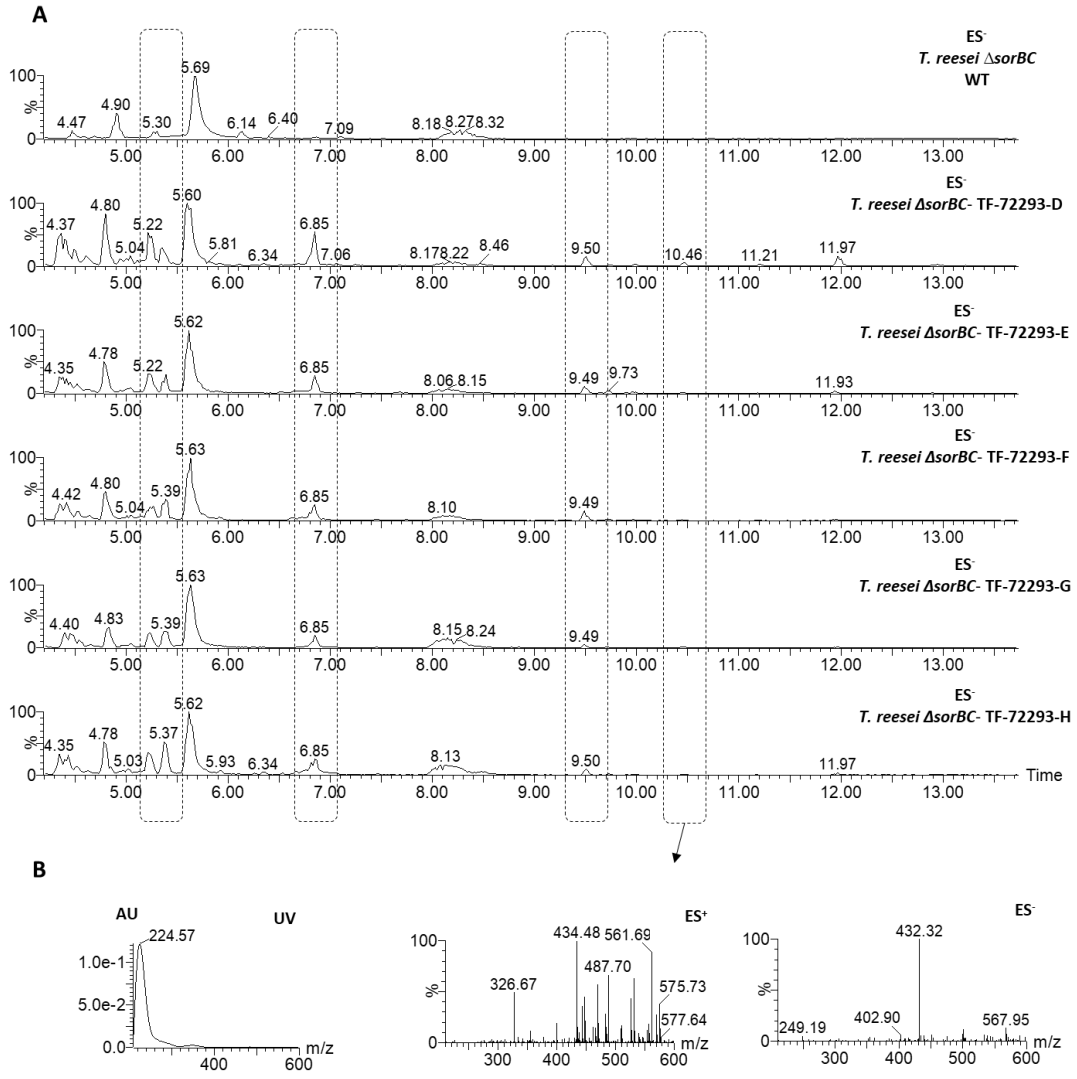


Figure 3.12 LCMS analysis of transcription factor activation experiment: **A**, ES$^-$ traces of a control (*T. reesei*.Δ*sorBC* as WT) and ES$^-$ traces of different transcription factor transformants (colonies D-H); **B**, UV chromatogram and Mass spectra (ES$^+$ and ES$^-$) of the new compound produced by the transfromants at 10.5 min.

3.3.3 Heterologous Expression of the Chosen BGC in *A. oryzae* NSAR1

3.3.3.1 Heterologous Expression of the PKS-NRPS and ER in *A. oryzae NSAR1*

In order to investigate the activity of the PKS-NRPS, a heterologous expression of the megasynthase and the accompanying enoyl reductase gene (ER) in the quadruple auxotrophic host *A. oryzae* NSAR1[104] was done. The gDNA of *T. reesei* QM6a Δ*tmus53* Δ*pyr4*[120] was used as a template to amplify genes for yeast homologous recombination (without removal of introns). The vector pEYA was used to assemble the PKS-NRPS

(*TRIREDRAFT_58285*) gene (~12 kb), in four different DNA fragments of around 3 kb each using yeast homologous recombination. This was followed by LR recombination between the entry plasmid pEYA·PKS/NRPS-58285 (MSIII139) and the expression vector pTYGS·*arg* to yield vector pTYGS·*argB*·PKS/NRPS-58285 (MSIII144, Figure 3.13). The enoyl reductase gene (*TRIREDRAFT_58289*) was also amplified from the gDNA of *T. reesei* using primers with overhang to P_{gpdA} and T_{gpdA}. Using yeast homologous recombination, the final vector pTYGS·*argB*·PKS/NRPS-58285·ER-58289 (MSIII152) was constructed which contains the PKS-NRPS (*58285*) and the ER (*58289*) encoding genes from the specified cluster (Figure 3.13).

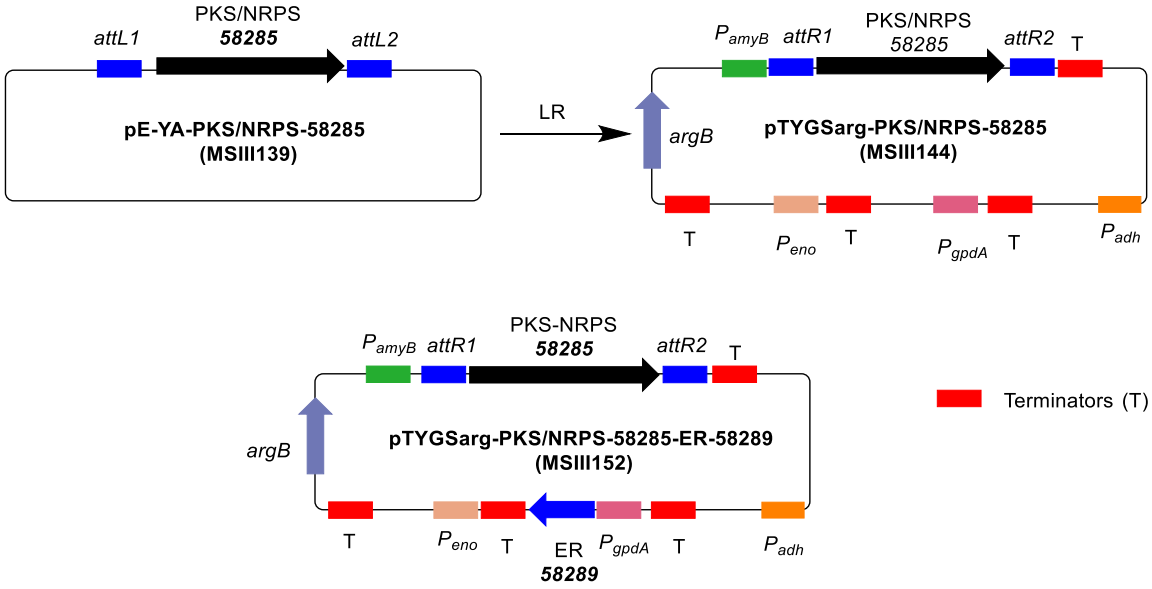


Figure 3.13 Fungal vectors used in heterologous expression of the first two genes in the cryptic BGC.

Heterologous expression of the vector pTYGS·*argB·PKS/NRPS-58285·ER-58289* (MSIII152) in *A. oryzae* NSAR1 strain was done using the PEG-mediated transformation (Section 5.2.5). After three rounds of selection, eleven transformants were obtained. The eleven transformants were cultivated in DPY media at 28 °C and 110 rpm for six days. The transformants were then extracted twice with ethyl acetate and analysed by LCMS. In the extract of five out of the eleven transformants (colony B, C, D, J and K), a new major peak eluting at T_R = 6.7 min and two minor peaks eluting at T_R = 8.1 and 11.0 min were observed compared to the WT control *A. oryzae* NSAR1 (Figure 3.14). The three compounds were shown to have closely-related UV spectra (Figure 3.15).

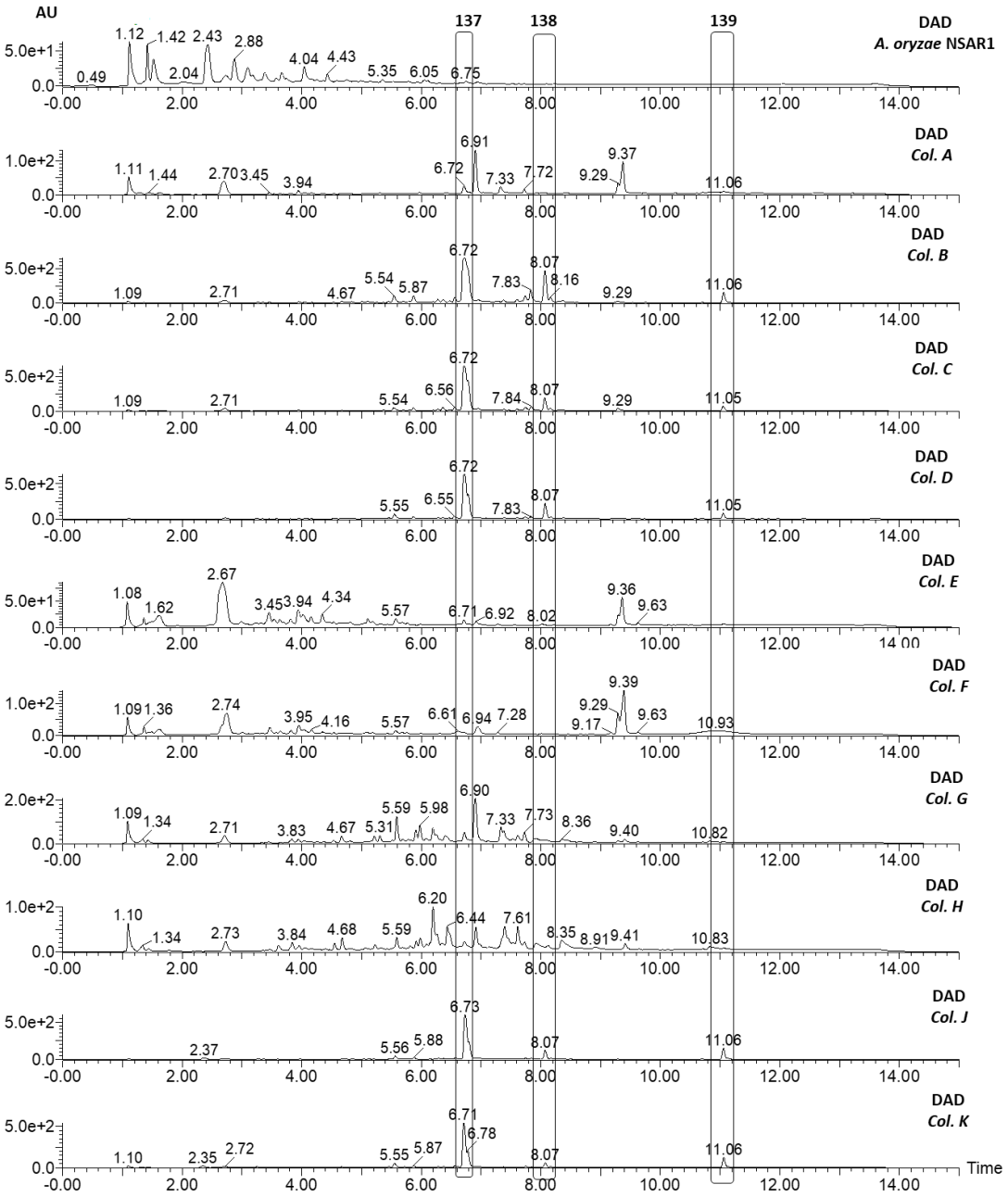


Figure 3.14 LCMS traces of *A. oryzae* NSAR1 and different *A. oryzae* + *PKS/NRPS-58285* + *ER-58289* transformants showing the DAD chromatograms of colonies A-K.

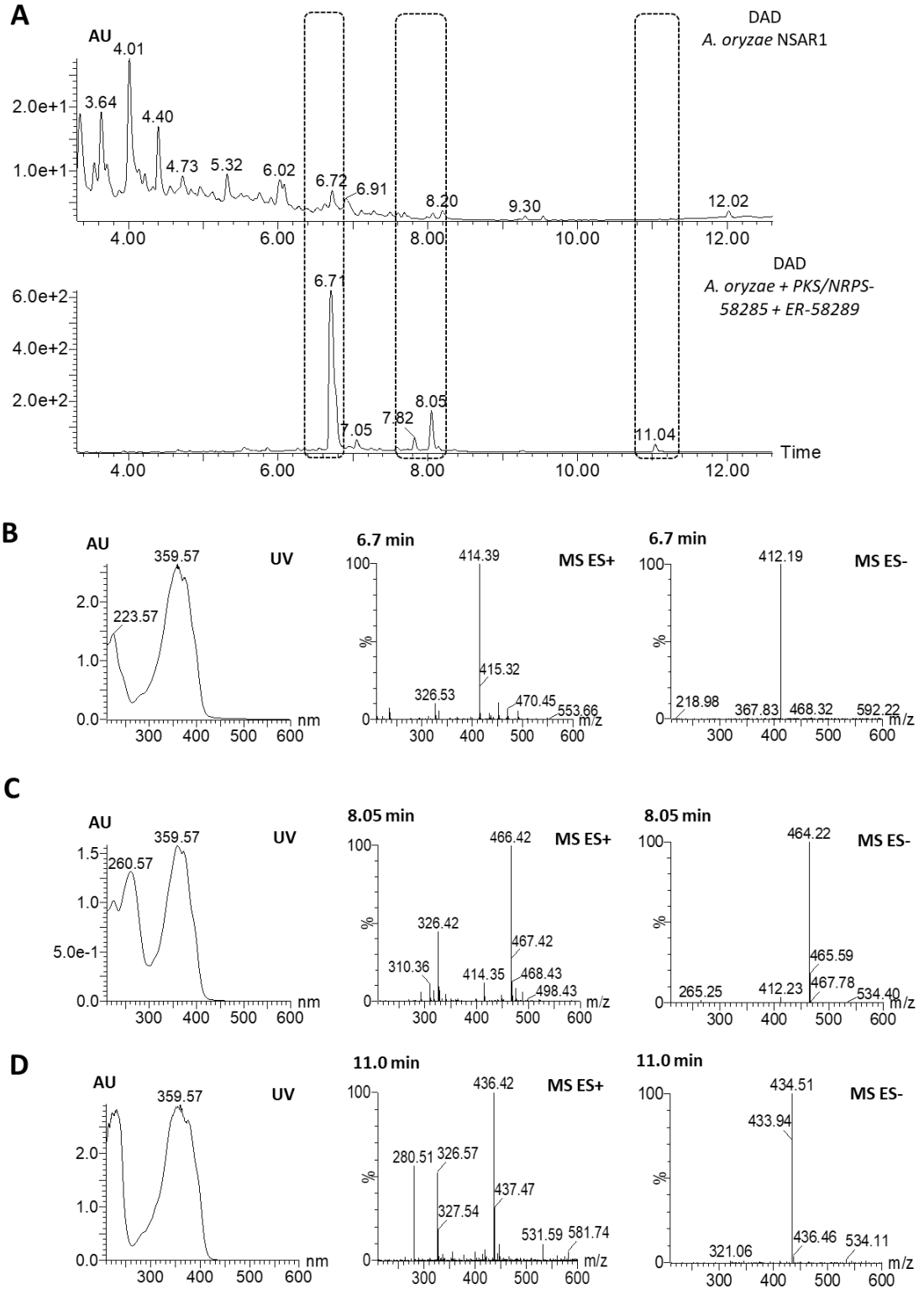


Figure 3.15 Chemical analysis of *A. oryzae* transformant expressing the first two genes in the cluster: **A**, LCMS traces of *A. oryzae* + *PKS/NRPS-58285* + *ER-58289* and *A. oryzae* NSAR1, showing the DAD chromatogram; **B-D**, UV chromatograms and mass spectra (ES^+ and ES^-) of the three major compounds produced by the *A. oryzae* + *PKS/NRPS-58285* + *ER-58289* transformant.

3.3.3.2 Compound Isolation and Structure Elucidation

For structure elucidation of the new compounds, the producing strain *A. oryzae + PKS/NRPS-58285 + ER-58289 colony* B was cultivated in 10 x 100 ml DPY media for six days. The crude extract was then submitted to preparative LCMS and 30 mg of the major compound **137** (T_R = 6.7 min) and 4 mg of the minor compound **138** (T_R = 8.05 min) were isolated. However, isolation of compound **139** (T_R = 11.0 min) was not possible due to the very low concentration of the compound in the extract. HRMS analysis of the major compound **137** confirmed a molecular formula of $C_{23}H_{28}NO_6$ ([M] + H^+ calculated 414.1917, found 414.1917). ^{13}C-NMR analysis of this compound in deuterated methanol showed the presence of only 18 carbon resonances corresponding to 20 carbon atoms, suggesting that three of the expected carbon resonances may be too weak to be directly observed. Eight peaks were in the range of 114-156 ppm indicating aromatic or olefinic carbons (Figure 3.16-A).

HSQC analysis of the compound showed that the two ^{13}C resonances at 116.1 and 131.8 ppm correspond to the two proton doublet signals at 6.67 and 6.99 ppm respectively, each integrating to two protons. This indicated the presence of a *para*-substituted benzene ring (Figure 3.16-B). HSQC spectra also showed the presence of two methyl carbons at 12.1 and 19.9 ppm, four methylene carbons at 27.7, 36.8, 37.7 and 42.4 ppm, five methine carbons at 31.3, 64.0, 116.7, 146.7 and 150.7 ppm and five quaternary carbons at 127.8, 135.6, 157.3, 176.2 and 176.9 ppm.

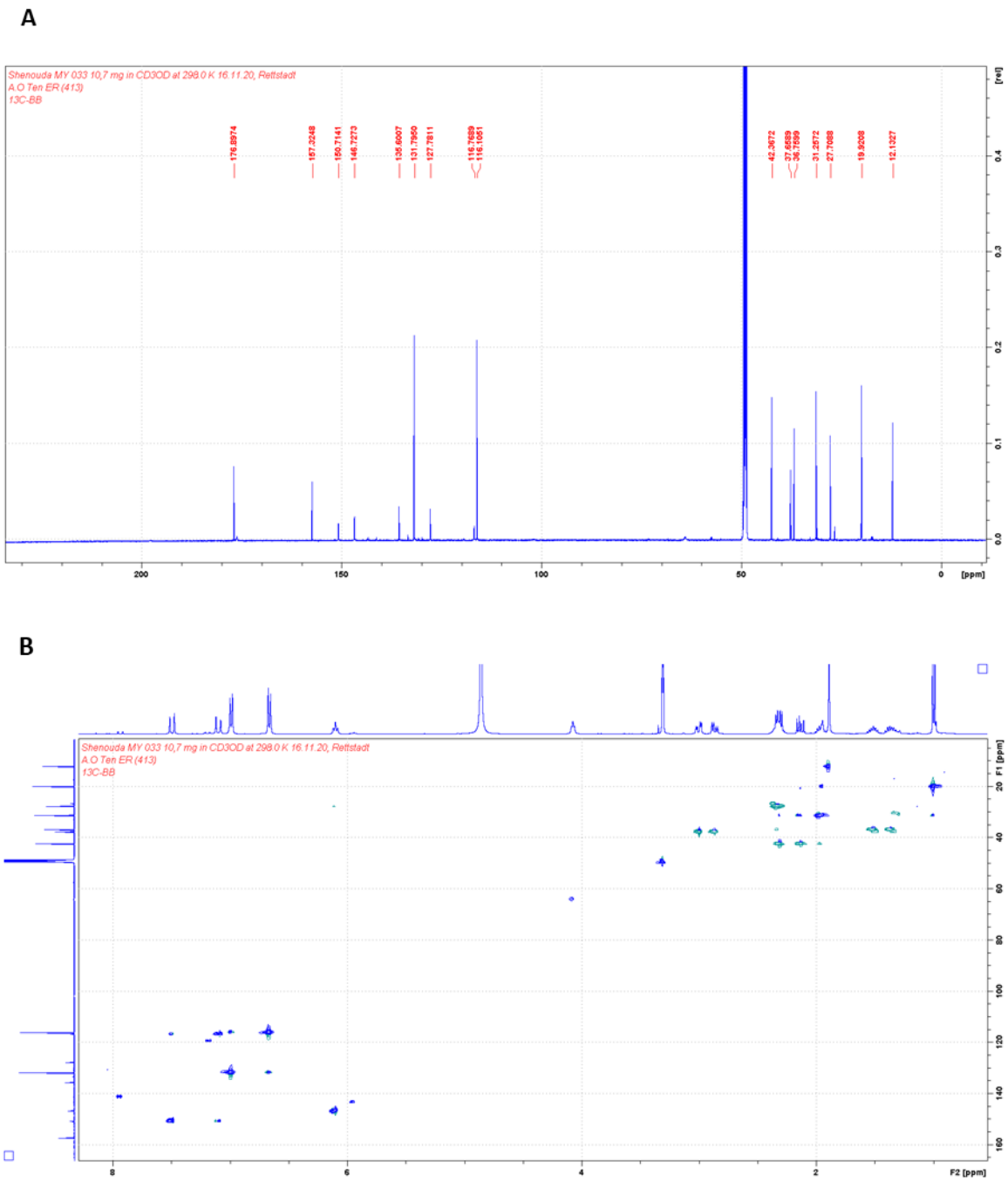


Figure 3.16 NMR spectra of compound **137** in methanol-d_4: **A**, ^{13}C-NMR spectrum of compound **137**; **B**, HSQC spectrum of compound **137**.

Proton signals for the hydroxyl groups are missing, presumably due to proton exchange in the deuterated solvent methanol. This can also explain the missing signals for three quaternary carbons in the ^{13}C-NMR spectrum as ^{2}H-exchange in deuterated solvent can result in delayed relaxation in the ^{13}C-NMR and therefore, very broad signals that are blurred with baseline noise. To identify the missing carbon signals and hydroxyl protons, the deuterated methanol was evaporated and re-protonation of compound **137** was done

by dissolving the compound in pure methanol and evaporation several times. Finally, 7 mg of the pure compound was dissolved in deuterated DMSO-d_6 and submitted for NMR analysis. ^{13}C-NMR spectrum of compound **137** in DMSO-d6 showed the presence of 21 peaks corresponding to 23 carbon atoms (Figure 3.17).

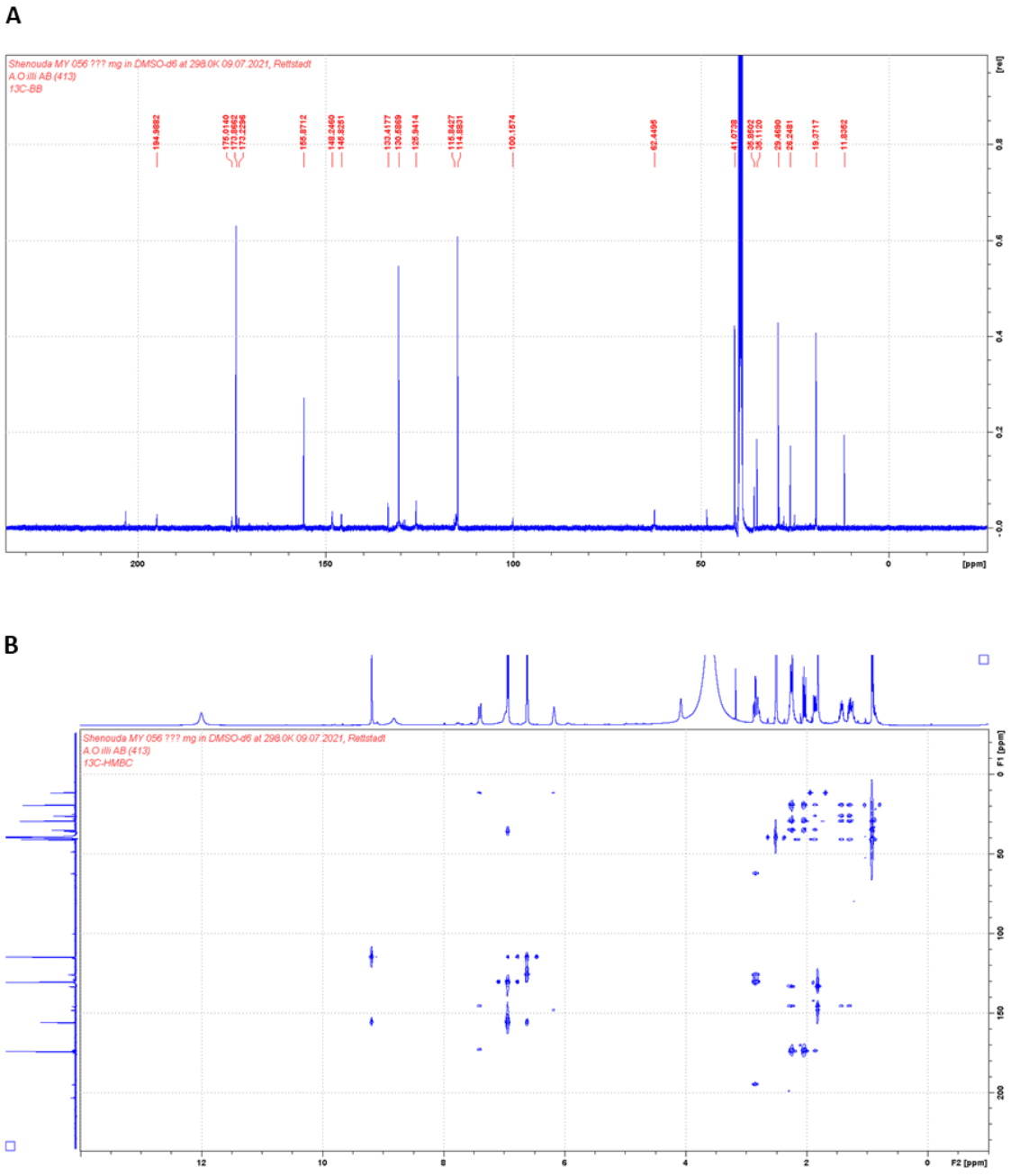


Figure 3.17 NMR spectra of compound **137** in DMSO-d_6: **A**, ^{13}C-NMR spectrum of compound **137**; **B**, HMBC spectrum of compound **137**.

Although the ^{1}H-NMR spectrum of the compound in DMSO-d_6 showed the missing hydroxyl groups, some proton signals were poorly separated and overlapped with the water and solvent peaks (Figure 3.18). Therefore, the data obtained from the two solvents were used together to fully elucidate the structure (Table 3.3).

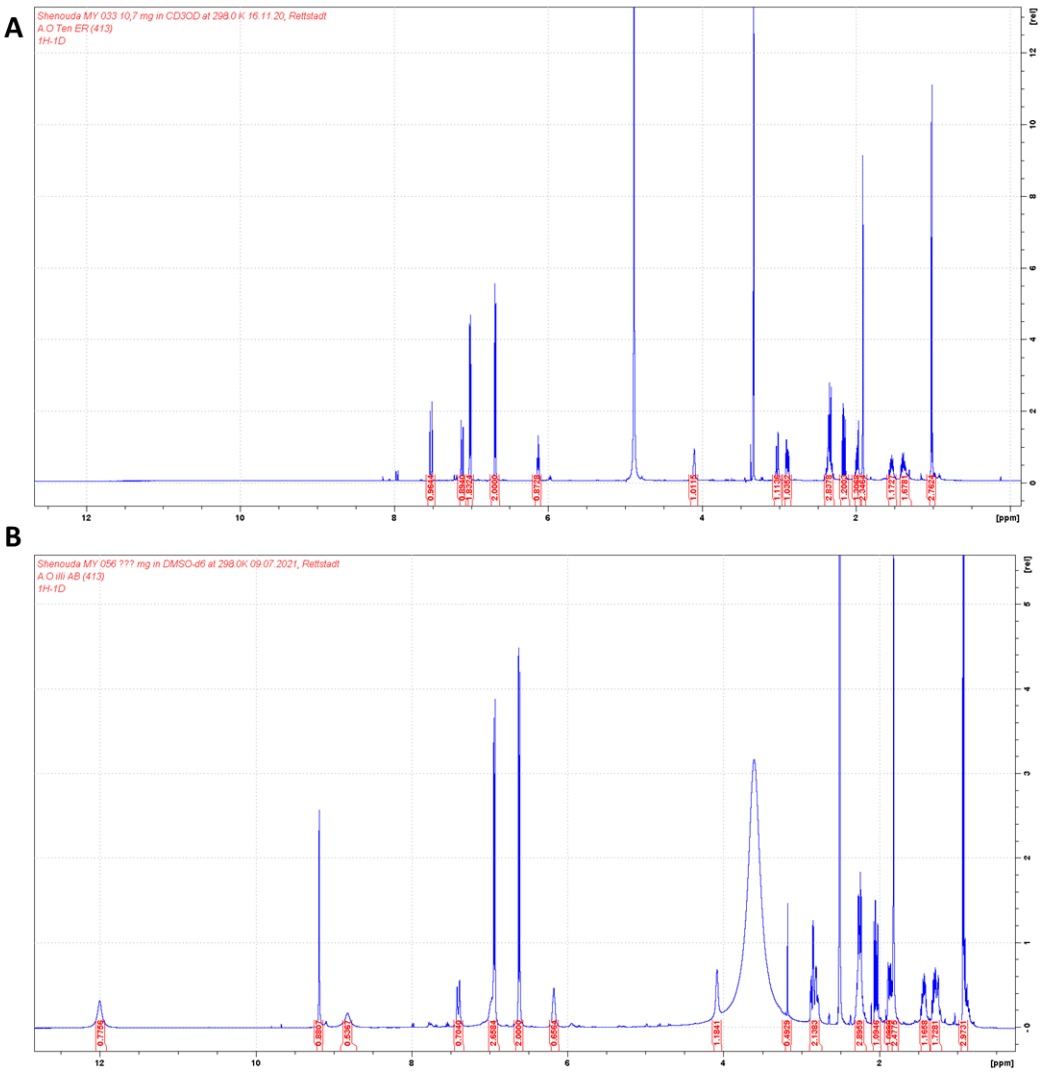


Figure 3.18 NMR spectra of compound **137**: **A**, ^{1}H-NMR spectrum of compound **137** in Methanol-d_4; **B**, ^{1}H-NMR spectrum of compound **137** in DMSO-d_6.

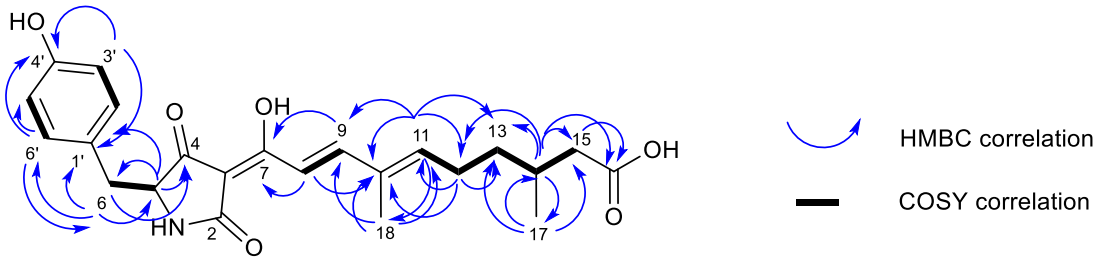


Table 3.3 Chemical shifts of compund **137** in Methanol-d_4 and DMSO-d_6.

	In methanol-d_4				In DMSO-d_6		
Position	**δ_C**	**δ_H (mult., int., *J*)**	**COSY**	**HMBC**	**δ_C**	**δ_H**	**HMBC (H to C)**
2	---	---	---	---	175.0	---	---
3	---	---	---	---	100.2	---	---
4	---	---	---	6	195.0	---	---
5	64.0	4.08 (t, 1H, *J* = 5 Hz)	6, 1`	6, 1`	62.5	4.08 (1H, broad peak)	4, 1`(very weak)
6	37.7	(H_a) 2.87 (dd, 1H, *J* = 6.0, 14.0 Hz) (H_b) 3.00 (dd, 1H, *J* = 4.5, 14.0 Hz)	5	5, 1`, 2`, 6`	35.9	(H_a) 2.80 (dd, 1H, *J* = 5.3, 14.0 Hz) (H_b) 2.86 (dd, 1H, *J* = 4.7, 14.0 Hz)	4, 5, 1`, 2`, 6`
7	176.9	---	---	---	173.2	---	---
8	116.7	7.10 (d, 1H, *J* = 15.6 Hz)	9	7, 9, 10, 11,	115.8	6.98 (overlap with 2`,6`)	---
9	150.7	7.50 (d, 1H, *J* = 15.6 Hz)	8	8, 11, 12, 18	148.2	7.41 (d, 1H, *J* = 15.6 Hz)	7, 10, 11, 18
10	135.6	---	---	8, 9, 11, 12, 14	133.4	---	---
11	146.7	6.11 (t, 1H, *J* = 7.5 Hz)	12, 14	8, 9, 12, 13, 14	145.8	6.17 (1H, broad peak)	9, 12, 13, 18
12	27.7	2.33 (2H, m)	11	---	26.3	2.26 (overlap with H-15b)	10, 11, 13
13	36.8	(H_a) 1.36 (1H, m) (H_b) 1.51 (1H, m)	12	11, 12, 14, 15, 17	35.1	(H_a) 1.28 (1H, m) (H_b) 1.43 (1H, m)	11, 12, 14, 15, 17
14	31.3	1.97 (1H, m)	15, 17	12, 13, 15, 16, 17	29.5	1.86 (1H, m)	12, 13, 15, 16, 17
15	42.4	(H_a) 2.14 (dd, 1H, *J* = 7.8, 15.0 Hz) (H_b) 2.32 (dd, 1H, *J* = 6.0, 15.0 Hz)	14	13, 14, 16, 17	41.1	(H_a) 2.04 (dd, 1H, *J* = 8.0, 15.0 Hz) (H_b) 2.25 (dd, 1H, *J* = 5.8, 15.0 Hz)	13, 14, 16, 17
16	176.2	---	---	14, 15	173.9	---	---
17	19.9	0.99 (d, 3H, *J* = 6.7 Hz)	14	13, 14, 15	19.4	0.91 (d, 3H, *J* = 6.7 Hz)	13, 14, 15
18	12.1	1.89 (s, 3H)	11	9, 10, 11	11.8	1.81 (s, 3H)	9, 10, 11
1`	127.8	---	---	5, 6, 3`, 5`	125.9	---	---
2`, 6`	131.8	6.99 (d, 2H, *J* = 8.5 Hz)	3`, 5`	6, 3`, 4`, 5`	130.6	6.95 (d, 2H, *J* = 8.5 Hz)	5, 6, 3`, 4`, 5`
3`, 5`	116.1	6.67 (d, 2H, *J* = 8.5 Hz)	2`, 6`	2`, 4`, 6`	114.9	6.62 (d, 2H, *J* = 8.5 Hz)	1`, 4`
4`	157.3	---	---	---	155.9	---	---
16-OH	---	---	---	---	---	12.00 (s, 1H)	---
4`-OH	---	---	---	---	---	9.19 (s, 1H)	2`, 3`, 4`, 5`, 6`

The ^{1}H-^{1}H COSY spectrum revealed the coupling between the methylene protons (H-6a at 2.87 ppm and H-6b at 3.00 ppm) and the methine proton (H-5 at 4.08 ppm, Figure 3.19-A). Based on that and the HMBC correlations (Figure 3.19 B-D), it was evident that the structure has a tetramic acid moiety.

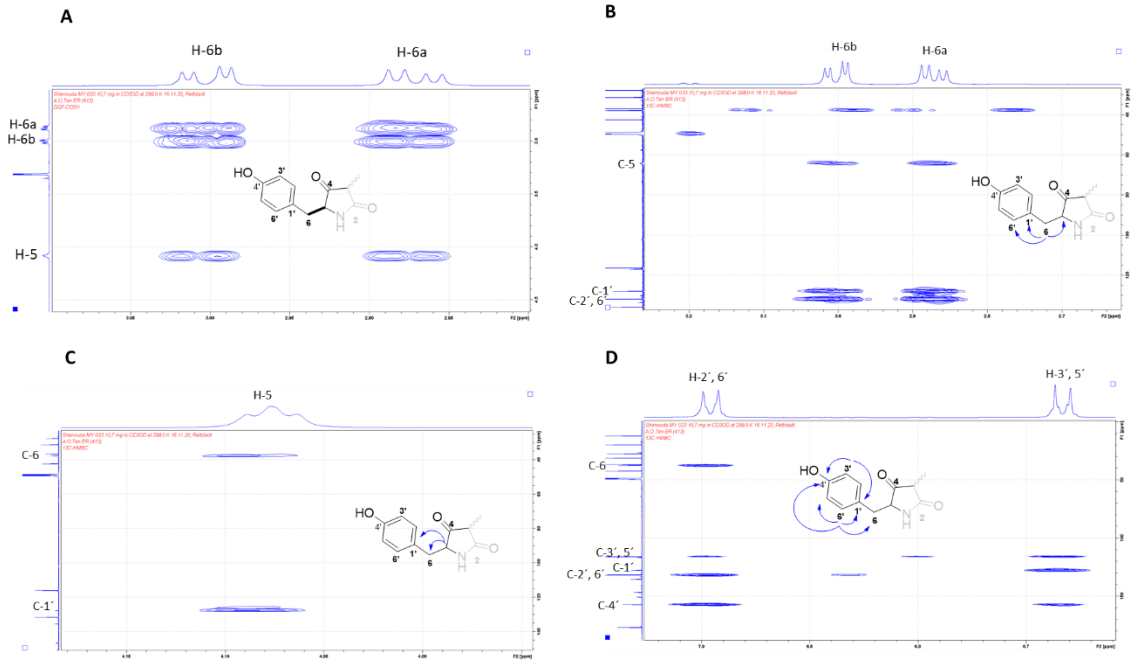


Figure 3.19 NMR spectra of compound **137**: **A**, COSY correlation of H-6; **B**, HMBC correlation of H-6; **C**, HMBC correlation of H-5; **D**, HMBC correlation of H-3´,5´and 2´,6´.

The side chain of the structure showed signals for four carbons in the olefin range; three methine carbons (at 116.7, 150.7 and 146.7) and one quaternary carbon at 135.6 ppm. The olefin at C-8/C-9 was shown to be *trans*, as the corresponding doublets at 7.10 and 7.50 ppm had the same coupling constant of 15.6 Hz. Based on COSY and HMBC correlations, the rest of the structure was elucidated (Figure 3.20). Notably, although the signal for the methyl group (H-18 at 1.89 ppm) is an apparent singlet, it demonstrated a coupling to the olefin hydrogen (H-11 at 6.11 ppm) which is 4-bonds away that might be attributed to the W-coupling between these protons.

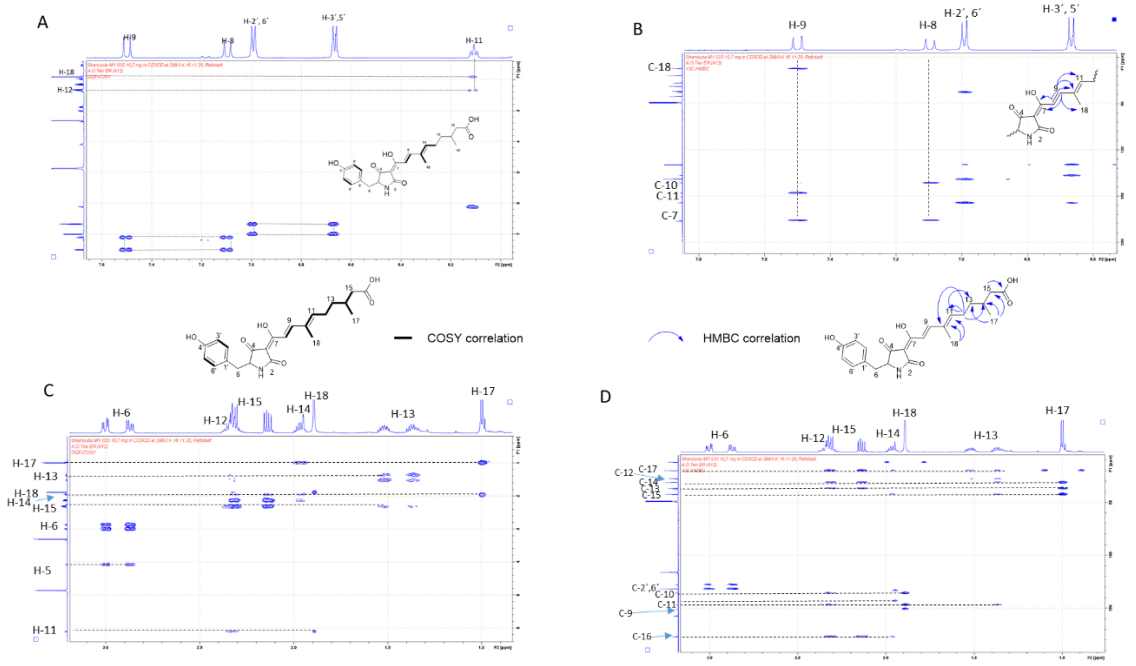


Figure 3.20 NMR spectra of compound **137**: **A**, COSY correlations of H-8, 9, 11, 2`, 6`and 3`,5`; **B**, HMBC correlation of H-8 and H-9; **C**, COSY correlation of H-6, 12, 13, 14, 15, 17 and 18; **D**, HMBC correlation of H-6, 12, 13, 14, 15, 17 and 18.

The chemical structure of the minor compound with nominal mass of 465 (T_R = 8.05 min) was identified based on HRMS, 1D and 2D NMR data. HRMS analysis of the minor compound **138** confirmed a molecular formula of $C_{27}H_{32}NO_6$ ([M] + H^+ calculated 466.2230, found 466.2237). When compared to the major compound **137**, the NMR signals for compound **138** in the tetramic acid region is quite similar. However, the rest of the structure showed some differences in the NMR data. For the side chain of the structure, the ^{1}H-NMR spectrum exhibited eight methines (including three vinylic), three methylene and two methyl groups. Based on 1D and 2D NMR, the compound was identified to be compound **138**, which is the oxidized Diels-Alder product of the ilicicolin-H tetramic acid (Table 3.4).

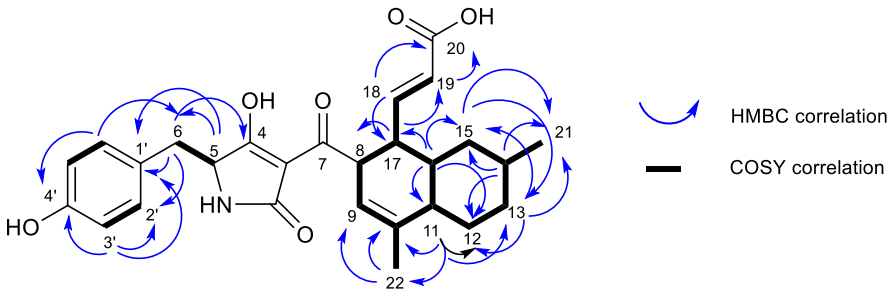


Table 3.4 Chemical shifts of compound **138** in Methanol-d_4

Position	δ_c	δ_H (mult., int., *j*)	COSY	HMBC
2	173.8	----------	----------	----------
3	101.5	----------	----------	----------
4	196.8	----------	----------	----------
5	64.0	4.00 (t, 1H, *J* = 5 Hz)	6	1`, 6
6	37.5	2.98 (dd, 1H, *J* = 4, 14.4 Hz) 2.86 (m, 1H)	5	4, 1`, 2`, 6`
7	190.0	----------	----------	----------
8	45.5	4.29 (broad, 1H)	9, 17, 22	----------
9	119.4	5.11 (broad, 1H)	8, 22	----------
10	141.5	----------	----------	----------
11	45.8	1.61 (m, 1H)	12, 22	10, 12, 13, 15, 22
12	30.6	Ha 1.04 (m, 1H) Hb 2.08 (dd, 1H, *J* = 3, 12.3 Hz)	11, 13	11, 13
13	36.7	Ha 0.99 (m, 1H) Hb 1.79 (d, 1H, *J* = 12.4 Hz)	12	12, 15, 21
14	34.0	1.44 (m, 1H)	21	12, 15, 21
15	41.1	Ha 0.54 (q, 1H, *J* = 12.2 Hz) Hb 1.62 (m, 1H)	14, 16	13, 21
16	40.8	1.97 (m, 1H)	11, 15	11, 12, 15, 17
17	46.8	2.41 (m, 1H)	8, 18	8, 19
18	151.5	6.84 (dd, 1H, *J* = 10.4, 15.4 Hz)	17, 19	17, 20
19	124.3	5.82 (d, 1H, *J* = 15.5 Hz))	18	17, 20
20	169.5	----------	----------	----------
21	23.0	0.89 (d, 3H, *J* = 6.5 Hz)	14	15
22	21.3	1.70 (s, 3H)	8, 9, 11	9, 10, 11, 13
1`	127.6	----------	----------	----------
2`, 6`	131.8	6.99 (d, 2H, *J* = 8.4 Hz)	3`, 5`	6, 4`
3`, 5`	116.1	6.67 (d, 2H, *J* = 8.4 Hz)	2`, 6`	1`, 2`, 6`, 4`
4`	157.4	----------	----------	----------

The side chain of compound **138** showed a number of COSY correlations (Figure 3.21-A-C) proving the predicted decalin structure.

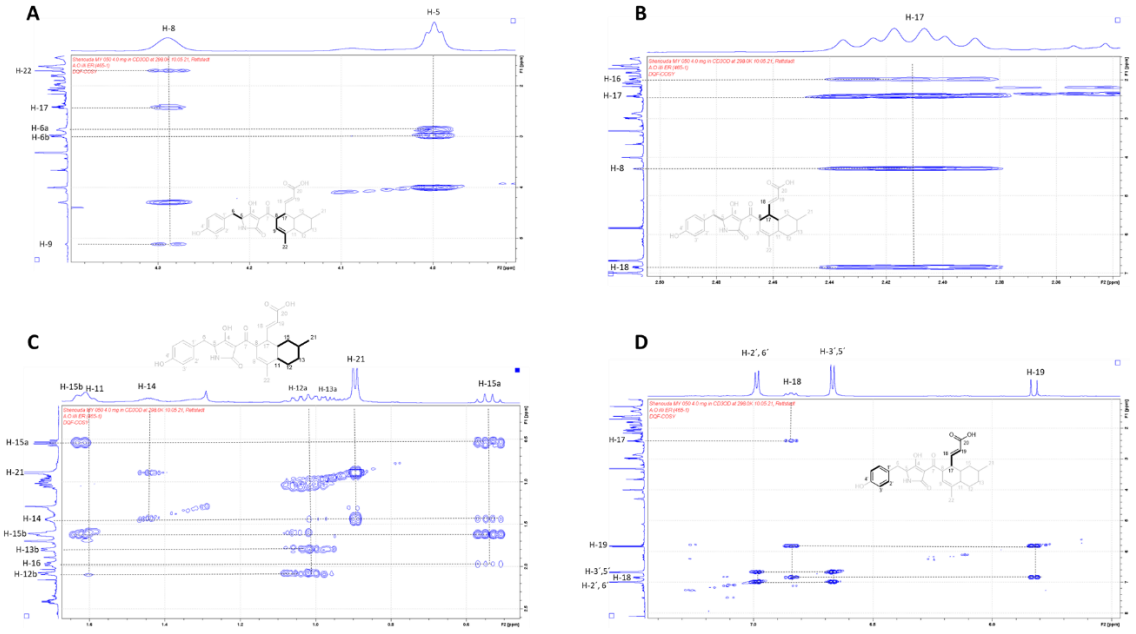


Figure 3.21 COSY correlations of compound **138**: **A**, COSY correlation of H-5 and H-8; **B**, COSY correlations of H-17; **C**, COSY correlation of H-11, 12, 14, 15 and 21; **D**, COSY correlation of H-18, 19 and H2`, 3`, 5`, 6`

The predicted decalin structure was also confirmed by HMBC correlations (Figure 3.22). This indicated that the formation of the decalin rings could occur spontaneously by intramolecular Diels-Alder reaction. The rest of the structure was also confirmed by HMBC correlations (Figure 3.23).

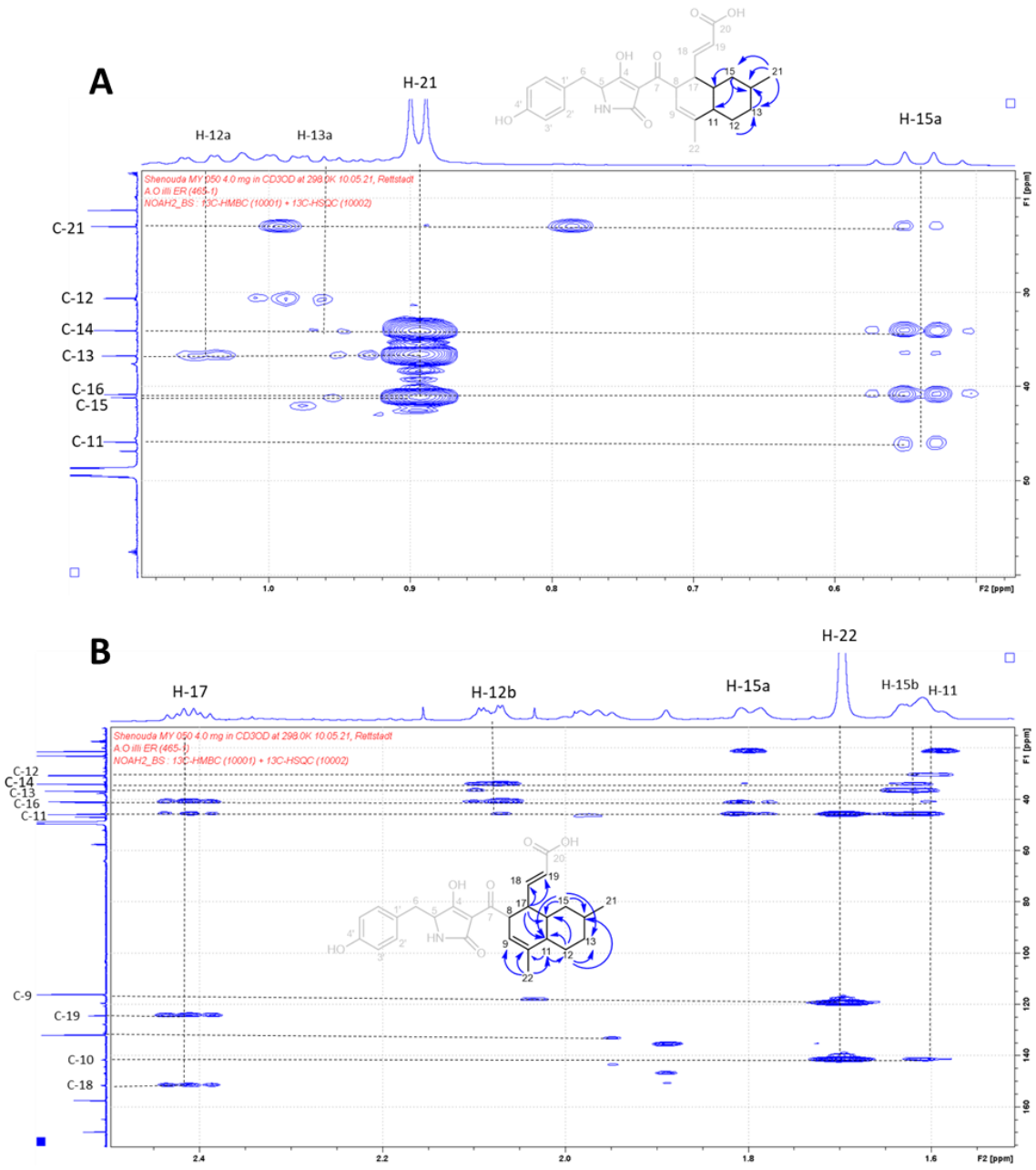


Figure 3.22 HMBC correlations of the decalin structure of compound **138**.

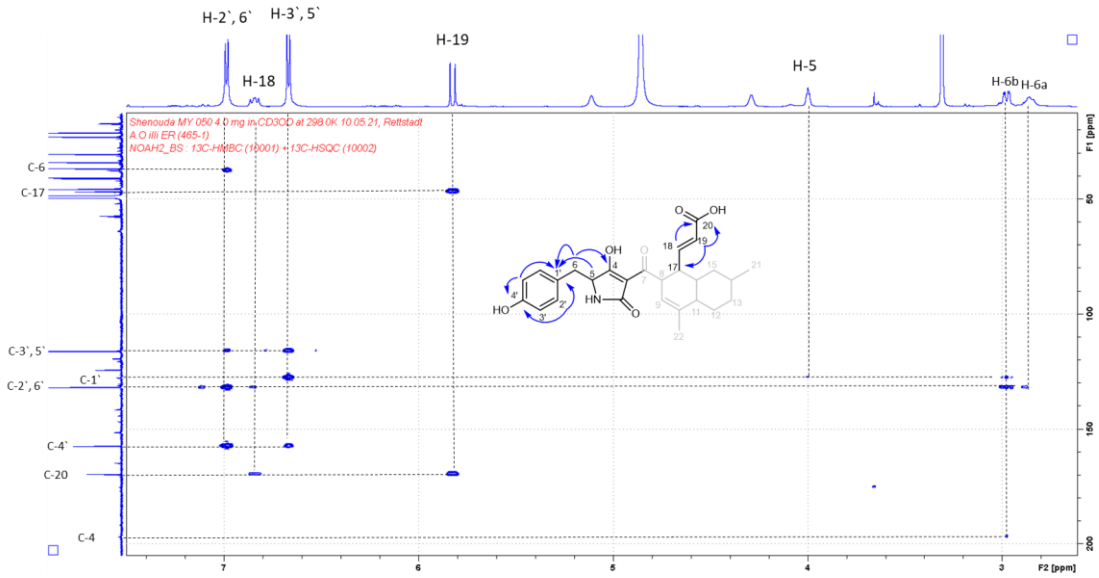


Figure 3.23 HMBC correlations of compound **138** confirming the rest of the structure.

However, the major compound **137** lacks a double bond at the end of the side chain that could interact with the double bond at C-8/C-9 to form this kind of decalin structure. This might indicate the formation of a compound with longer side chain similar to that of ilicicolin-H tetramic acid. This compound could then form the new compound **138** by intramolecular Diels-Alder reaction followed by oxidation of the methyl group at the end of the carbon chain. Also the fact that the major product **137** of the megasynthase has very similar chemical structure to the ilicicolin-H tetramic acid **139** with only four carbon less was quiet intriguing, especially with the unsaturation at the end of the chain that makes the structure vulnerable to oxidation. With many previous records of *A. oryzae* oxidation products produced from the heterologous expression, a time course of the transformant was required to ensure the stability of the produced compounds in *A. oryzae* culture.

3.3.3.3 Time course Study

The tranformant *A. oryzae +PKS/NRPS-58285 + ER-58289* was cultivated in seven different 500 ml baffled flasks, each containing 100 ml DPY media. The flasks were then incubated at 28 °C and 110 rpm and extraction of one flask was done every 24 hrs to construct the time course. Time course of *A. oryzae +PKS/NRPS-58285 + ER-58289* (Figure 3.24) showed that starting from the second day of cultivation a new compound with nominal mass of 435 is produced as a major product (compound **139**). Starting from

the third day of cultivation, compound **137** was also produced in almost equal amounts to that with nominal mass 435. However, on the fourth day of cultivation, production of compound **137** started to increase, while that of compound **139** started to decrease. Compound **138** was only produced on the fifth day of cultivation.

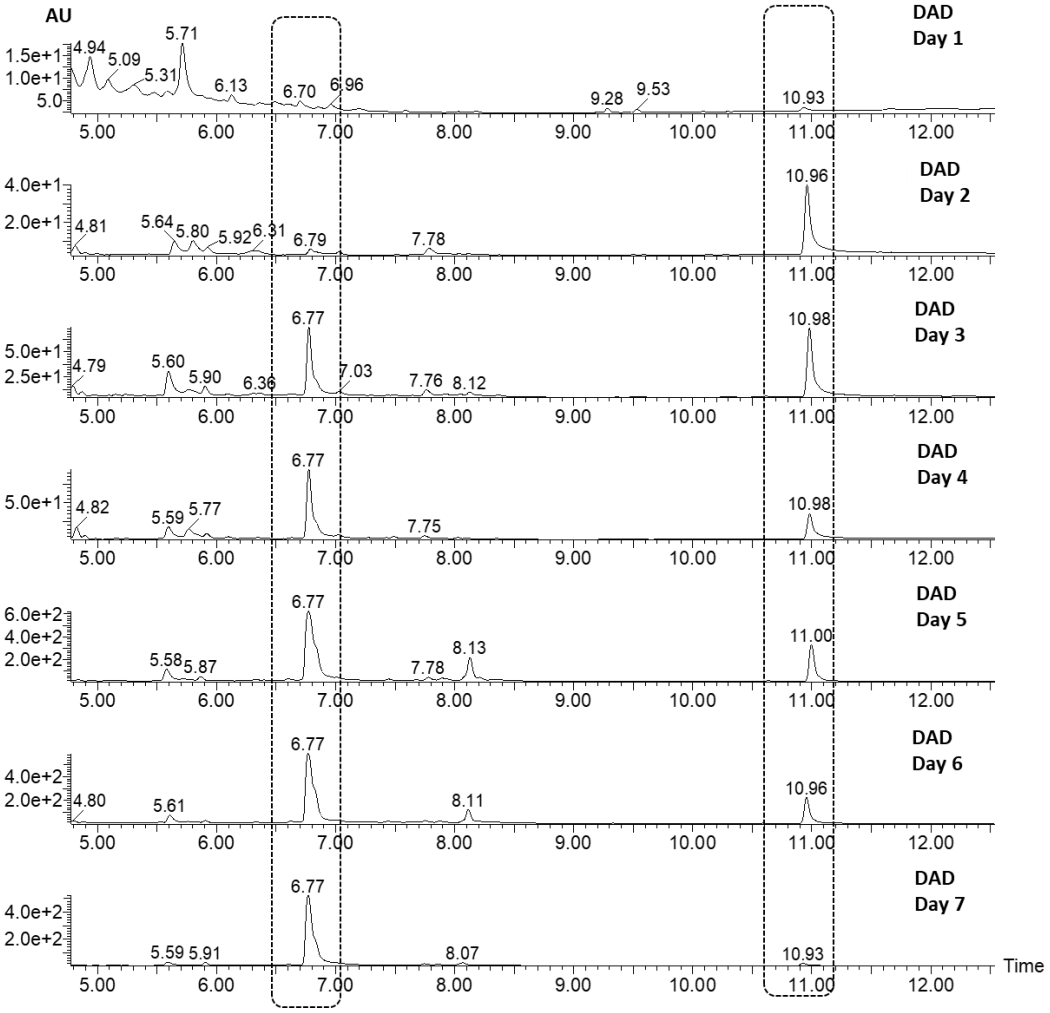


Figure 3.24 Time course of *A. oryzae* + *PKS-NRPS-58285* + *ER-58289* transformant, showing the DAD chromatogram of the transformant.

3.3.3.4 Heterologous Expression of the First Three Genes of the BGC

A vector for the expression of the first three genes of the cluster; PKS/NRPS (TRIREDRAFT_58285), ER (TRIREDRAFT_58289), $P450_{RE}$ (*TRIREDRAFT_58953*) was constructed using yeast homologous recombination using gDNA of *T. reesei* QM6a·Δ*tmus53*·Δ*Pyr4* as template for the amplification of the genes (Figure 3.25).

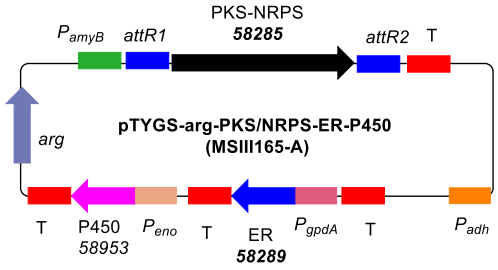


Figure 3.25 Vector map for the vector used for expression of the first three genes from the cluster.

Transformation of *A. oryzae* NSAR1 was done with the constructed vector using a PEG-mediated transformation protocol and this resulted in the generation of four transformants. The strain *A. oryzae* NSAR1 (control) and the transformants were cultivated for 6 days in DPY media (28 °C, 110 rpm). The transformants were then extracted with ethyl acetate and submitted to LCMS for analysis. The four transformants with the three genes showed a new peak with *m/z* = 433 that has the same mass and UV spectrum as the 2-pyridone compound produced in ilicicolin-H biosynthetic pathway **140**. Another two minor new compounds with *m/z* = 411 and *m/z* = 451 were also produced (Figure 3.26).

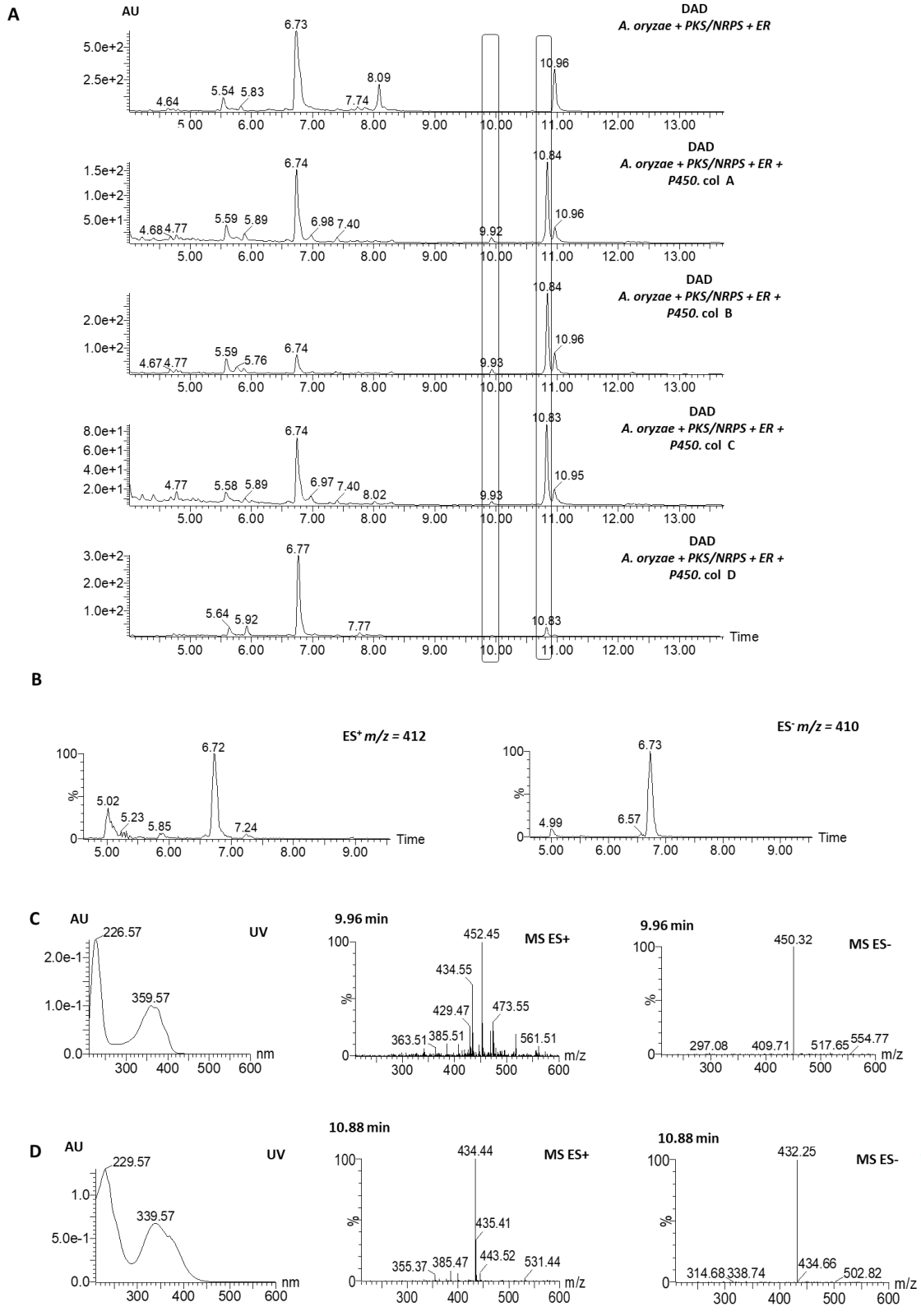


Figure 3.26 Co-expression of three genes in *A. oryzae* NSAR1: **A**, LCMS chromatograms of different colonies of *A. oryzae* + *PKS/NRPS-58285* + *ER-58289* + *P450-58953* against *A. oryzae* + *PKS/NRPS-58285* + *ER-58289*; **B**, Extracted ion chromatogram (ES^+ at *m/z* = 412 and ES^- at *m/z* = 410); **C-D**, UV, ES^+ and ES^- of the two new compounds produced by *A. oryzae* + *PKS/NRPS-58285* + *ER-58289* + *P450-58953* transformants.

For the structure elucidation of the new compounds, the producing strain *A. oryzae* + *PKS/NRPS-58285* + *ER-58289* + *P450-58953* was cultivated in 1 L DPY media for six days at 28 °C and 110 rpm.

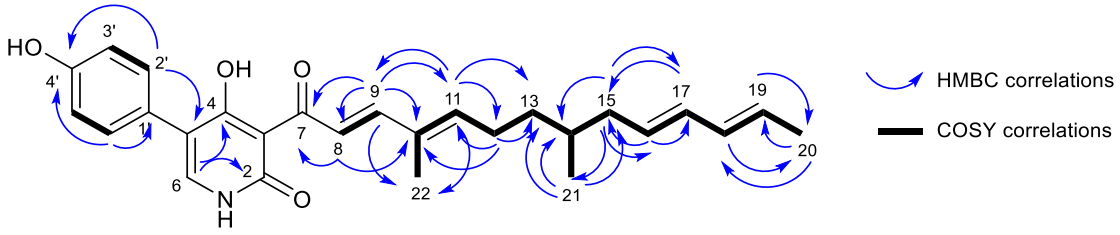
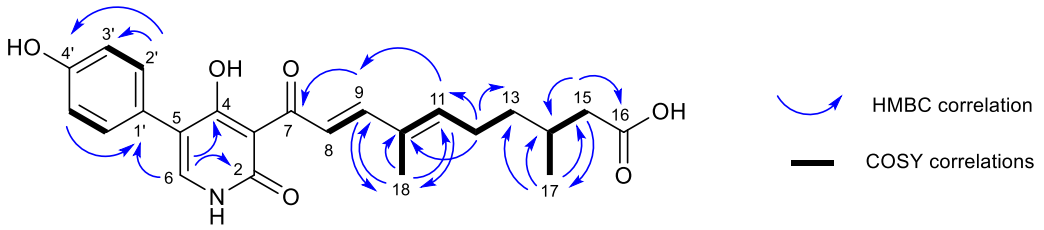


Table 3.5 Chemical shifts of compound **140** (Comparison with reported NMR data[214])

	Reported NMR (in $CDCl_3$: acetone-d_6 = 3:1)[214]		Found NMR (in Acetone-d_6)	
	δ_C	δ_H **(MULT., INT., *J*)**	δ_C	δ_H **(MULT., INT., *J*)**
2	162.9	-------	162.7	---------
3	106.7	---------	107.2	---------
4	178.0	--------	179.0	---------
5	114.8	----------	114.6	---------
6	139.1	7.44 (s, 1H)	140.8	7.61 (obscured, 1H)
7	194.6	7	195.2	
8	123.3	8.01 (d, 1H, *J* = 15.5 Hz)	124.2	8.15 (d, 1H, *J* = 15.5 Hz)
9	150.1	7.53 (d, 1H, *J* = 15.5 Hz)	150.4	7.59 (d, 1H, *J*= 15.5 Hz))
10	134.5	---------	135.1	---------
11	144.8	6.01 (t, 1H, *J* = 7.5 Hz)	145.5	6.15 (t, 1H, *J* = 7.5 Hz)
12	26.8	2.22 (m, 2H,)	27.4	2.32 (m)
13	35.5	Ha 1.22 (m, 1H) Hb 1.42 (m, 1H)	36.4	1.30 1.51
14	33.1	1.50 (m, 1H)	33.8	1.58 (m, 1H)
15	40.0	Ha 1.88 (q, 1H, *J* = 13.0, 7.0 Hz) Hb 2.02 (q, 1H, *J* = 13.0, 6.0 Hz)	40.5	1.94 (m) 2.11 (m)
16	130.0	5.45 (m, 1H)	130.8	5.55 (m, 1H)
17	131.4	5.88 (m, 1H)	132.8	6.00 (m, 1H)
18	131.6	5.94 (m, 1H)	132.8	6.04 (m, 1H)
19	126.7	5.50 (m, 1H)	127.3	5.56 (m, 1H)
20	17.9	1.65 (d, 3H, *J* = 6.5 Hz)	18.1	1.69 (d, , 3H, *J* = 6.5 Hz)
21	19.4	0.85 (d, 3H, *J* = 7.0 Hz)	19.7	0.924 (d, 3H, *J* = 7.0 Hz)
22	12.4	1.84 (s, 3H)	12.5	1.90 (s, 3H)
1'	124.4	---------	125.4	---------
2'/6'	130.3	7.26 (d, 1H, *J* = 8.5 Hz)	131.1	7.34 (d, 1H, *J* = 8.5 Hz)
3'/5'	115.4	6.83 (d, 1H, *J* = 8.5 Hz)	115.9	6.87 (d, 1H, *J* = 8.5 Hz)
4'	156.9	---------	157.8	---------

The crude extract was then submitted to preparative LCMS and 1 mg of each of the two compounds with nominal mass 433 and 411 were purified. However, isolation of the minor compound with nominal mass 451 was not possible due to the very low concentration in the extract. The two isolated compounds were then submitted to HRMS, 1 and 2D NMR analysis. HRMS analysis of compound **140** with nominal mass 433 confirmed a molecular formula of $C_{27}H_{30}NO_4$ ([M] - H^+ calculated 432.2175, found 432.2175). ^{13}C- and ^{1}H-NMR data of this compound in deuterated acetone was comparable to that reported for compound **140** from the ilicicolin-H BGC (Table 3.5).[214]



xx

Table 3.6 Chemical shifts of compound **141** in deuterated Methanol-d_4.

No.	δ_C	δ_H **(MULT., INT., *J*)**	COSY	HMBC
2	164.5	----------	---------	---------
3	------	----------	---------	---------
4	178.1	----------	---------	---------
5	-------	----------	---------	---------
6	140.6	7.47 (s, 1H)	---------	2, 4, 1`
7	195.9	----------	---------	---------
8	124.5	7.97 (d, 1H, *J* = 15.5 Hz)	9	--------
9	151.1	7.59 (d, 1H, *J* = 15.5 Hz)	8	7, 18
10	135.8			
11	145.6	6.10 (t, 1H, *J* = 7.5 Hz)	12, 18	9, 18
12	27.6	2.33 (m, 2H)	11	10, 11, 13
13	37.0	Ha 1.36 (m, 1H) Hb 1.53 (m, 1H)	12	---------
14	31.5	1.97 (m, 1H)	15, 17	---------
15	43.4	Ha 2.11 (q, 1H, *J* = 14.0, 8.0 Hz) Hb 2.30 (m, 1H)	14	14, 16, 17
16	178.2	----------	---------	---------
17	20.1	1.00 (d, 3H, *J* = 7.0 Hz)	14	13, 14, 15
18	12.5	1.92 (s, 3H)	11	9, 10, 11
1´	125.4	----------	---------	---------
2´, 6´	131.4	7.29 (d, 1H, *J* = 8.5 Hz)	3', 5'	3', 4', 5'
3´, 5´	116.1	6.82 (d, 1H, *J* = 8.5 Hz)	2', 6'	1'
4´	158.3	----------	---------	---------

HRMS analysis of compound **141** with nominal mass 411 confirmed a molecular formula of $C_{23}H_{24}NO_6$ ([M] - H^+ calculated 410.1604, found 410.1608). The molecular weight of this compound is two mass units less than compound **137**, which might indicate a structure with 2-pyridone ring instead of tetramic acid ring due to oxidative ring expansion through the action of the expressed P450 enzyme. The structure of the compound **141** was elucidated based on 1 and 2D NMR (Table 3.6).

^{13}C- and ^{1}H-NMR signals for the side chain of the compound were comparable to those of the tetramic acid compound **137** in deuterated methanol but showed some down-field chemical shift that could be attributed to the conjugation with the phenyl-2-pyridone rings. ^{1}H-NMR of compound **141** showed three signals in the aromatic range 6.8-7.5 ppm corresponding to five hydrogens. The singlet peak at 7.47 ppm corresponding to H-6 showed no COSY correlations (Figure 3.27-B) and HMBC correlations to two carbonyl carbons at 164.5 and 178.1 ppm and to a quaternary carbon at 125.4 (Figure 3.28-B). This indicated the expansion of the tetramic acid 5-membered ring to the 2-pyridone 6-membered ring. The rest of the structure was confirmed by COSY and HMBC correlations (Figure 3.27 and 3.28).

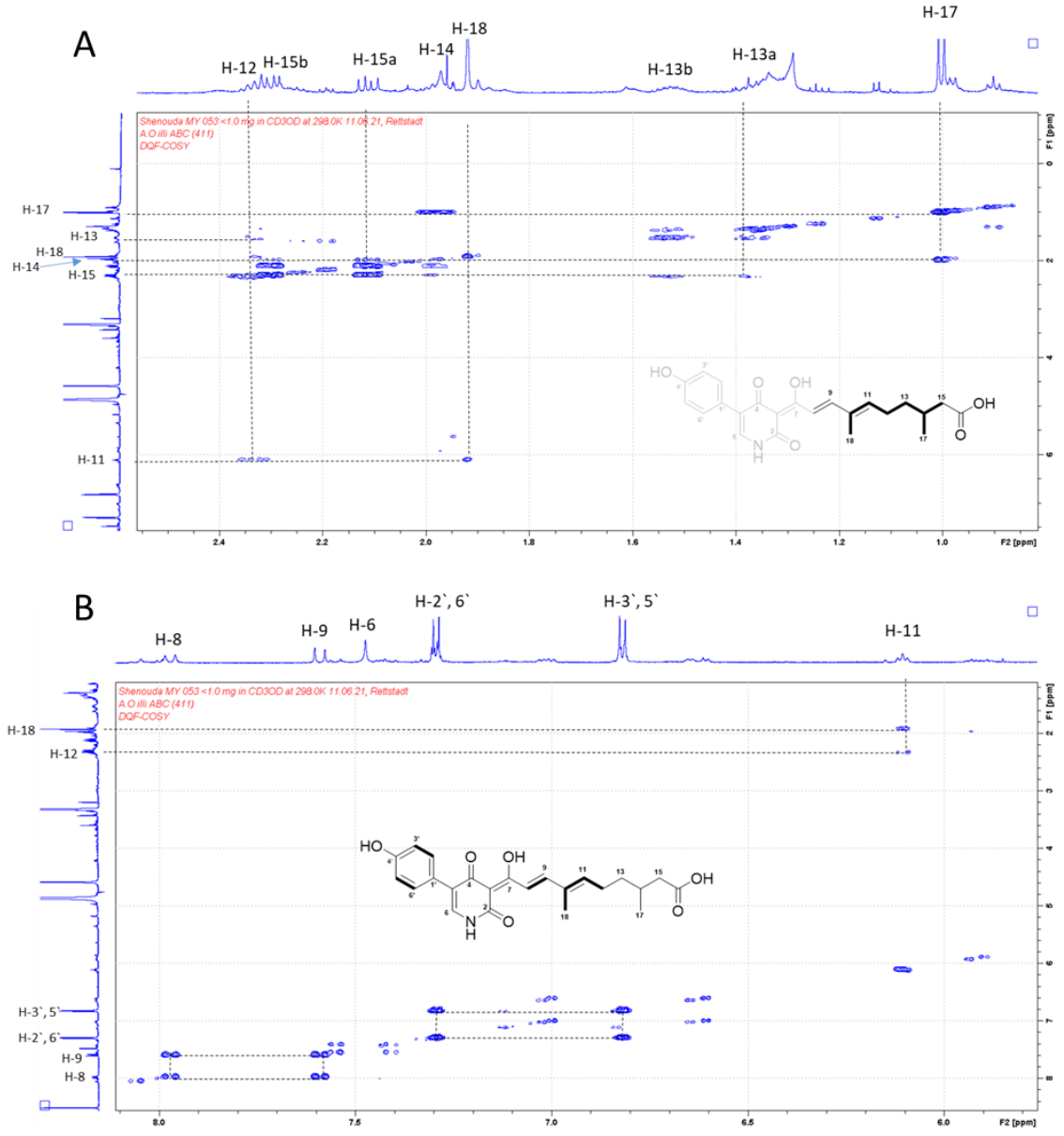


Figure 3.27 COSY correlations of compound **141**.

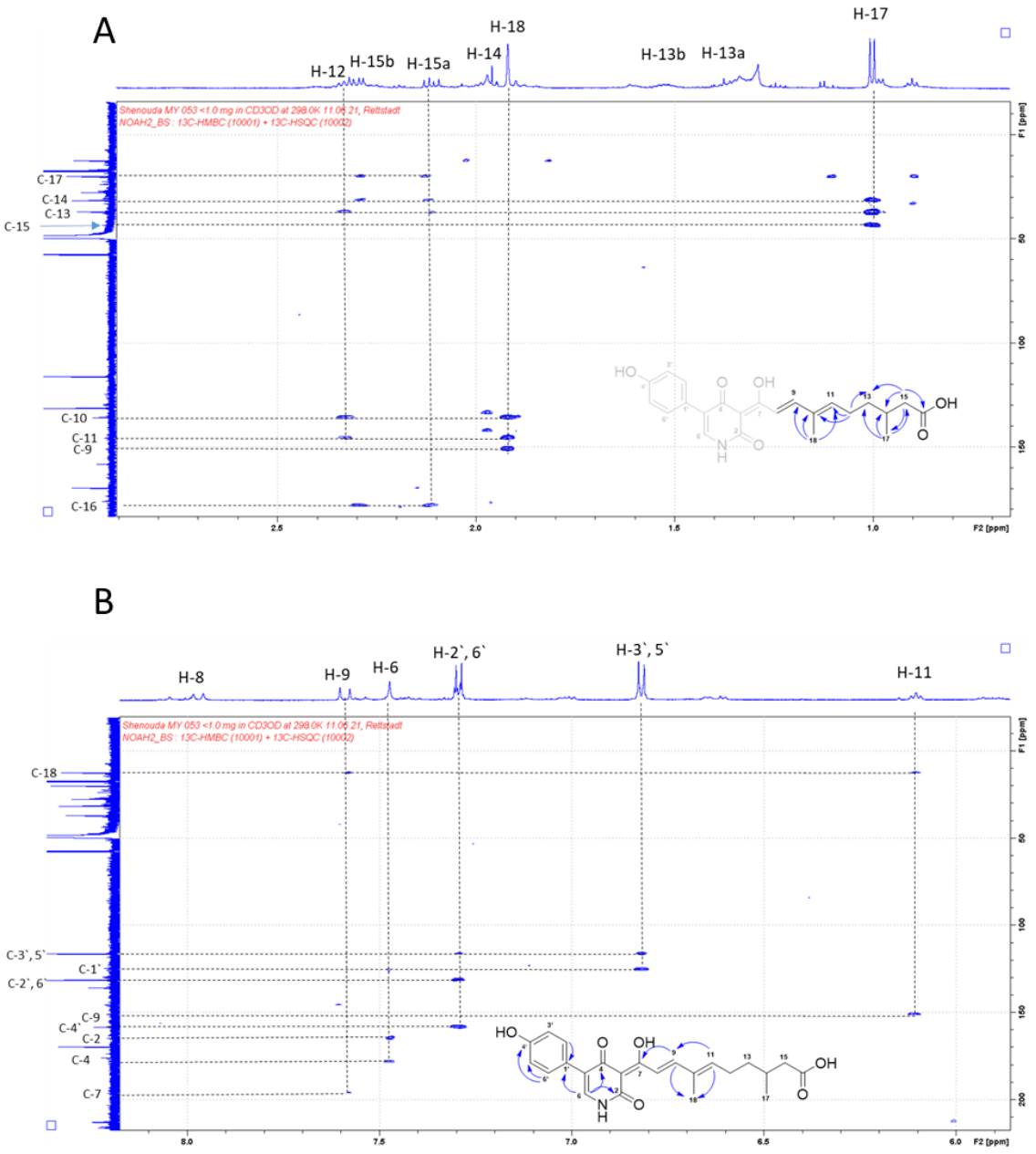


Figure 3.28 HMBC correlations of compound **141**.

3.3.3.5 Heterologous expression of the first four genes from the cluster

A vector for the expression of the first four genes of the cluster; PKS/NRPS (TRIREDRAFT_58285), ER (TRIREDRAFT_58289), P450$_{RE}$ (*TRIREDRAFT_58953*) and *C*-methyl transferase (*C-MeT*) was constructed using yeast homologous recombination using gDNA of *T. reesei* QM6a·Δ*tmus53*·Δ*Pyr4* as template for the amplification of the genes (Figure 3.29).

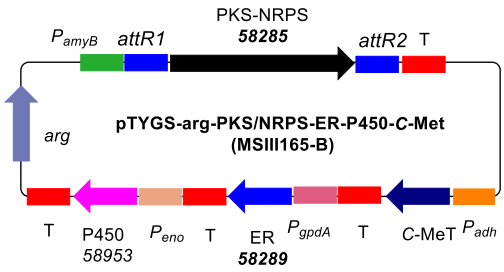


Figure 3.29 Vector map for the vector used for expression of the first four genes from the cluster.

Transformation of *A. oryzae* NSAR1 was done with the constructed vectors using PEG-mediated transformation protocol and resulted in the generation of six transformants with the four genes (PKS/NRPS + ER + P450 + *C*-MeT). The strain *A. oryzae* NSAR1 (control) and the transformants were cultivated for 6 days in DPY media (28 °C, 110 rpm). The transformants were then extracted with ethyl acetate and submitted to LCMS for analysis. Out of the six transformants, only three transformants showed the production of a new compound that have the same mass and UV as that of ilicicolin-H **72** and 8-*epi*-ilicicolin-H **142** (Figure 3.30). Another new minor compound with same mass as ilicicolin-H, but different UV, was also produced at 10.5 min (Figure 3.30-C), with data corresponding to that of the minor compound produced by activation of the TF in *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* (Section 3.3.2).

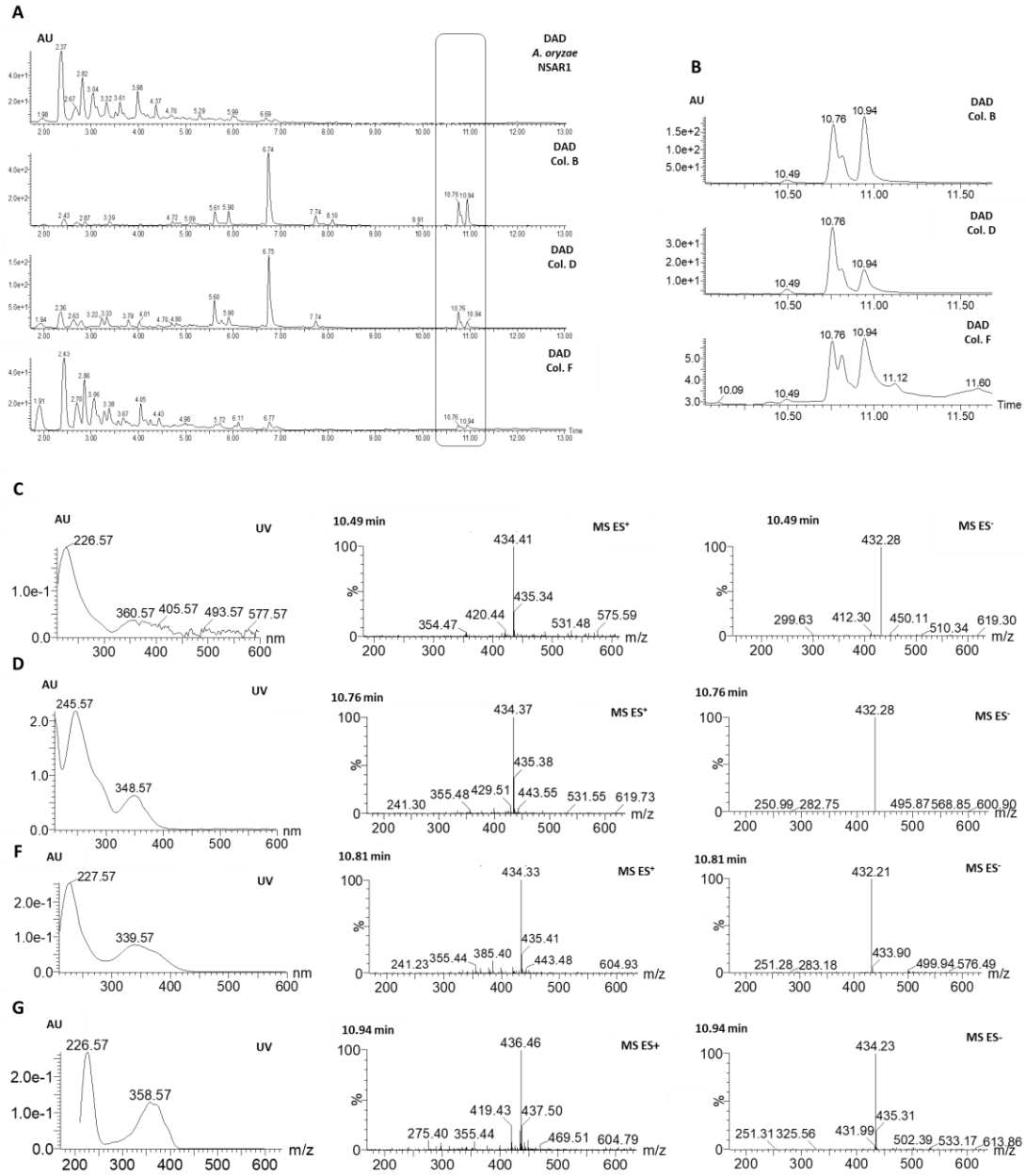
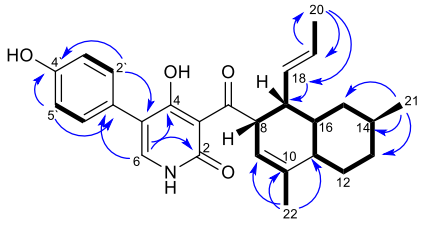


Figure 3.30 Co-expression of four genes in *A. oryzae* NSAR1: **A-B**, LCMS chromatograms of different colonies of *A. oryzae* + *PKS/NRPS58285·ER58289·P45058953·C-MeT* against *A. oryzae* NSAR1 as the WT; **C-G**, UV, ES^+ and ES^- chramtograms of the four compounds produced by *A. oryzae* + *PKS/NRPS-58285* + *ER-58289* + *P450-58953* + *C-MeT* transformants.

Isolation of this compound with molecular mass of 433 was done by cultivation of *A. oryzae* + *PKS/NRPS58285·ER58289·P45058293·C*-MeT colony B in 1 L DPY media for 5 days then extraction and isolation of the compound using preparative HPLC. This afforded ~ 1 mg of compound **142** that was submitted for NMR analysis. 1 and 2D NMR analysis and comparison with published data (Table 3.7) showed that the compound isolated was 8-*epi*-ilicicolin-H **142**.



8-*epi*-Ilicicolin-H **142**

Table 3.7 Chemical shifts of compound **142** (Comparison with reported NMR data[214])

	Reported NMR for 8-epi-ilicicolin-H (in acetonitrile-d_3)[214]		Found NMR for 8-epi-ilicicolin-H (in acetonitrile-d_3)	
	δ_C	δ_H **(MULT., INT., *J*)**	δ_C	δ_H **(MULT., INT., *J*)**
2	162.9	----------	162.4	----------
3	108.0	----------	----------	----------
4	177.6	----------	177.7	----------
5	114.8	----------	114.3	----------
6	140.8	7.42 (s, 1H)	140.7	7.40 (s, 1H)
7	209.8	----------	209.5	----------
8	50.8	5.07 (m, 1H)	50.7	5.07 (m, 1H)
9	120.1	5.47 (m, 1H)	120.1	5.45 (m, 1H)
10	141.2	----------	141.2	----------
11	46.0	1.54 (m, 1H)	45.9	1.55 (m, 1H)
12	30.4	Ha 1.00 (m, 1H) Hb 2.03 (m, 1H)	30.4	Ha 0.99 (m, 1H) Hb 2.07 (m, 1H)
13	36.3	Ha 0.93 (m, 1H) Hb 1.77 (m, 1H)	36.3	Ha 0.95 (m, 1H) Hb 1.78 (m, 1H)
14	33.7	1.41 (m, 1H)	33.7	1.42 (m, 1H)
15	40.7	Ha 0.44 (q, 1H, *J* = 24.5, 12.0 Hz) Hb 1.74 (m, 1H)	40.7	Ha 0.46 (q, 1H, *J* = 24.5, 12.0 Hz) Hb 1.77 (m, 1H)
16	40.6	1.93 (m, 1H)	40.7	1.94 (obscured by solvent peak, 1H)
17	47.4	2.17 (m, 1H)	47.5	2.18 (Obscured by H_2O peak, 1H)
18	133.4	5.45 (m, 1H)	133.5	5.46 (m, 1H)
19	127.3	5.44 (m, 1H)	127.2	5.46 (m, 1H)
20	18.1	1.54 (d, 3H)	18.0	1.56 (d, 3H, *J* = 4.5 Hz)
21	23.0	0.88 (d, 3H, *J* = 6.5 Hz)	23.1	0.89 (d, 3H, *J* = 6.5 Hz)
22	21.5	1.66 (s, 3H)	21.6	1.68 (m, 3H)
1`	125.4	----------	125.5	----------
2`, 6`	131.4	7.26 (d, 1H, *J* = 9.0 Hz)	131.4	7.27 (d, 1H, *J* = 8.7 Hz)
3`, 5`	116.0	6.83 (d, 1H, *J* = 9.0 Hz)	116.1	6.84 (d, 1H, *J* = 8.7 Hz)
4`	157.5	10.09 (s, OH)	157.5	----------

However, the NMR data showed that the compound was not 100 % pure and some very minor peaks could be seen that matches that of ilicicolin-H (Table 3.8). This might indicate that the compound could epimerise spontaneously, but in very low amount (Figure 3.31).

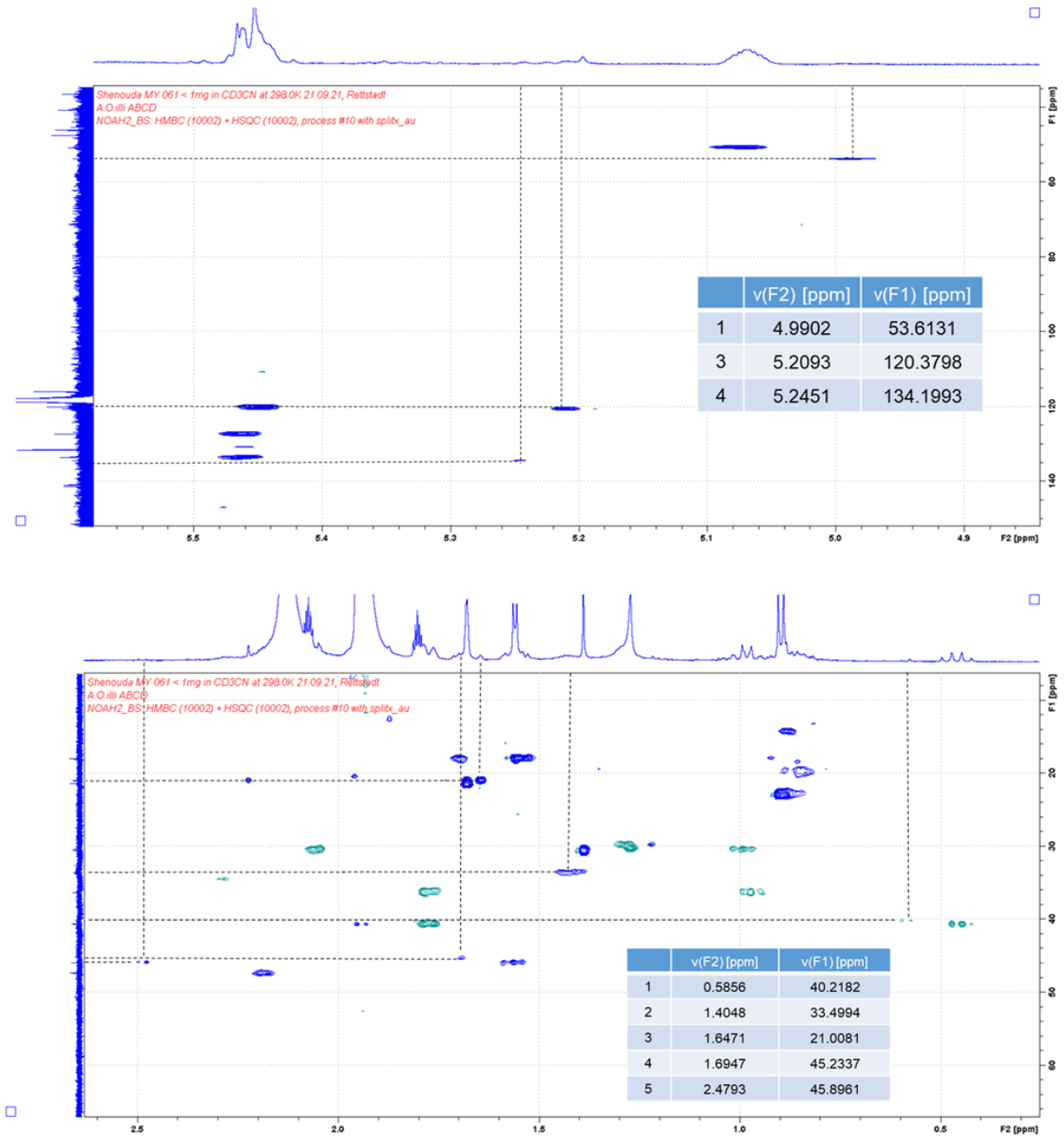
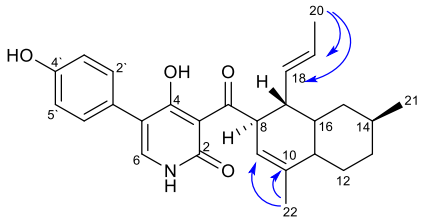


Figure 3.31 HSQC spectrum showing minor peaks corresponding to ilicicolin-H **72**.



Ilicicolin-H **72**

Table 3.8 Chemical shifts of compound **72** (Comparison with reported NMR data[214])

	Reported NMR for ilicicolin-H (in acetonitrile-d_3)[214]		Found NMR for ilicicolin-H (in acetonitrile-d_3)	
	δ_C	δ_H **(MULT., INT., *J*)**	δ_C	δ_H **(MULT., INT., *J*)**
2	162.8	----------	162.4	----------
3	107.8	----------	----------	----------
4	177.6	----------	177.7	----------
5	114.6	----------	114.3	----------
6	140.9	7.43 (s, 1H)	140.7	7.41 (s, 1H)
7	210.7	----------	----------	----------
8	53.8	4.98 (m, 1H)	53.6	4.99 (m, 1H)
9	120.4	5.21 (m, 1H)	120.4	5.21 (m, 1H)
10	139.1	----------	139.1	----------
11	45.1	1.67 (m, 1H)	45.2	1.69 (m, 1H)
12	30.6	Ha 0.97 (m, 1H) Hb 2.03 (m, 1H)	30.4	Ha 1.02 (m, 1H) Hb 2.05 (m, 1H)
13	36.2	Ha 0.97 (m, 1H) Hb 1.76 (m, 1H)	36.3	Ha 0.99 (m, 1H) Hb 1.76 (m, 1H)
14	33.4	1.38 (m, 1H)	33.5	1.40 (m, 1H)
15	40.7	Ha 0.57 (q, 1H) Hb 1.76 (m, 1H)	40.2	Ha 0.59 (q, 1H) Hb 1.77 (m, 1H)
16	44.2	1.22 (m, 1H)	44.3	1.20 (m, 1H)
17	45.8	2.48 (m, 1H)	45.9	2.48 (m, 1H)
18	134.2	5.21 (m, 1H)	134.2	5.24 (m, 1H)
19	126.8	5.32 (m, 1H)	126.6	5.32 (m, 1H)
20	18.0	1.52 (dd, 3H, *J* = 6.5, 1.3 Hz)	18.0	1.56 (d, 3H, *J* = 4.5 Hz)
21	23.0	0.88 (d, 3H, *J* = 6.5 Hz)	23.0	0.89 (d, 3H, *J* = 6.5 Hz)
22	21.0	1.62 (s, 3H)	21.0	1.65 (m, 3H)
1`	125.3	----------	125.5	----------
2`, 6`	131.4	7.26 (d, 1H, *J* = 8.6 Hz)	131.4	7.27 (d, 1H, *J* = 9.0 Hz)
3`, 5`	116.0	6.83 (d, 1H, *J* = 8.6 Hz)	116.1	6.84 (d, 1H, *J* = 9.0 Hz)
4`	157.5	----------	157.5	----------

HMBC correlations also showed the presence of two compounds in the mixture; the major compound was 8-*epi*-ilicicolin-H **142**, while the minor one was ilicicolin-H **72** (Figure 3.32).

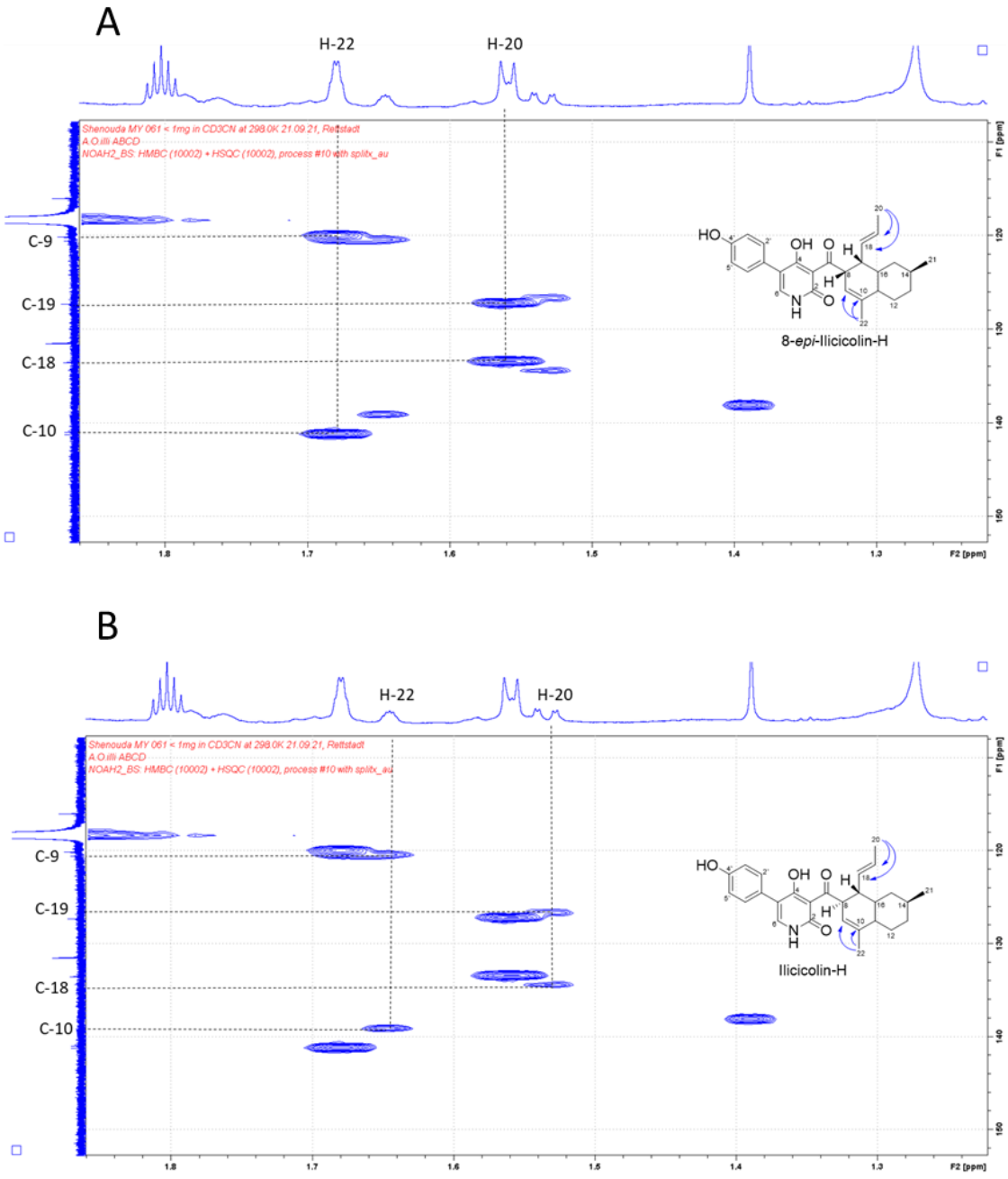


Figure 3.32 HMBC correlation of the compound with nominal mass 433 isolated from *A. oryzae* trasnformants expressing the four genes from the investigated BGC: **A**, HMBC spectrum showing major peaks corresponding to the major compound 8-*epi*-ilicicolin-H **142**; **B**, HMBC spectrum showing minor peaks corresponding to the minor compound ilicicolin-H **72**.

3.5 Discussion

Although *Trichoderma reesei* has been used as a fungal host for protein production and as an excellent producer of cellulase and hemicellulose enzymes, research on the secondary metabolites of *T. reesei* is not well developed. This might be attributed to the limited availability of molecular methods to express and activate the BGC of *T. reesei.* Previous attempt to activate the BGC in *T. reesei* showed limited success.[192,194] Therefore, the aim of this work is to explore the biosynthetic potential of *T. reesei* by heterologous expression of one or more of the cryptic or silent biosynthetic gene clusters from *T. reesei* in *A. oryzae*.

3.5.1 Bioinformatic Analysis to Identify all PKS and Hybrid PKS-NRPS from *T. reesei* Genome

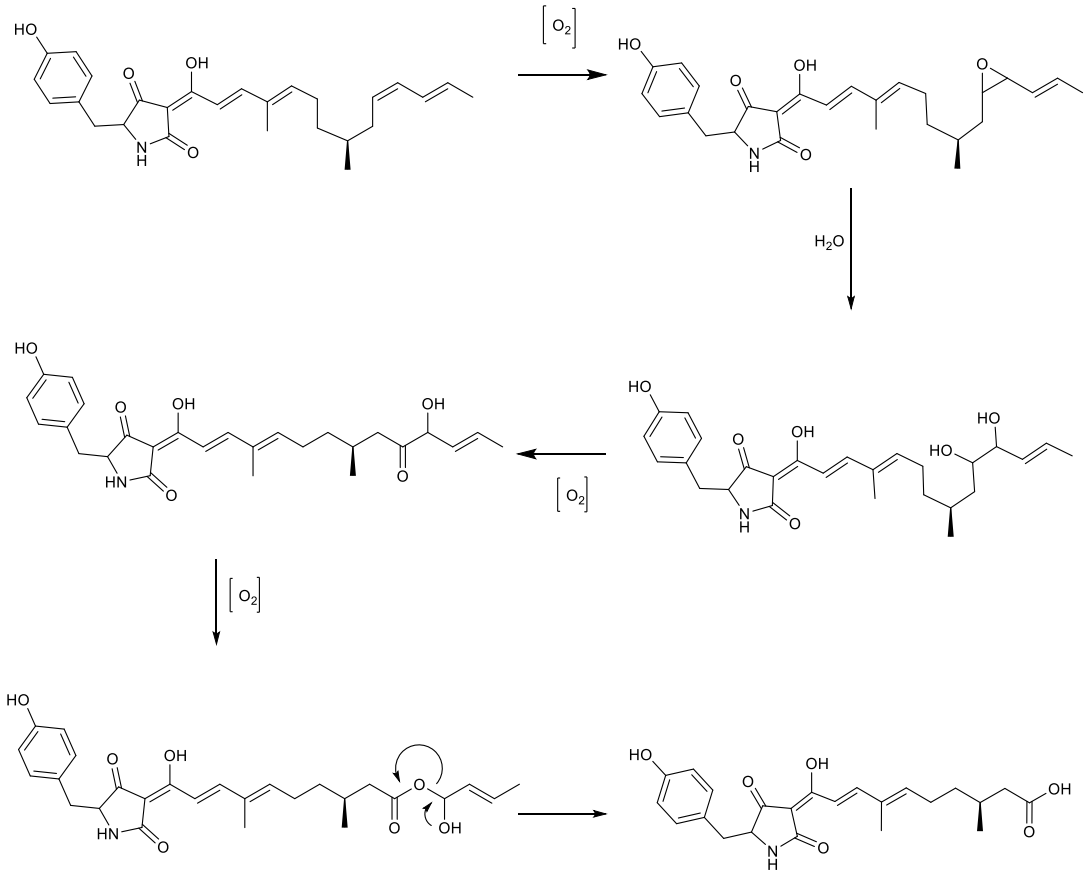
Bioinformatic anaylsis of all BGC in *T. reesei* showed the presence of 11 PKS and two PKS-NRPS gene clusters in *T. reesei.* As no PKS-NRPS products have been previously reported form *T. reesei*, one of the two cryptic hybrid PKS-NRPS BGC was chosen for further investigation.

3.5.2 Transcription Factor Activation.

One of the two PKS/NRPS BGC in *T. reesei*, which showed high similarity to the antifungal ilicicolin-H BGC, was chosen for further investigation. First, expression of the transcription factor TRIREDRAFT_72993 present in the vicinity of this cluster in *T. reesei* QM6a·Δ*tmus53*·Δ*pyr4*·Δ*sorBC* strain resulted in the production of some compounds that were also present in minor concentration in the WT strain and only one new compound with odd molecular mass indicating the presence of an odd number of Nitrogen atoms. This compound showed a molecular mass of 433, which is the same as that of ilicicolin-H, but different UV from that of ilicicolin-H. However, the new compound was produced in very low titer which prevented its isolation and structure elucidation. Therefore, a heterologous expression experiment of this gene cluster in *A. oryzae* NSAR1 was planned to identify the product of this BGC.

3.5.3 Heterologous Expression of the PKS/NRPS BGC in *A. oryzae* NSAR1

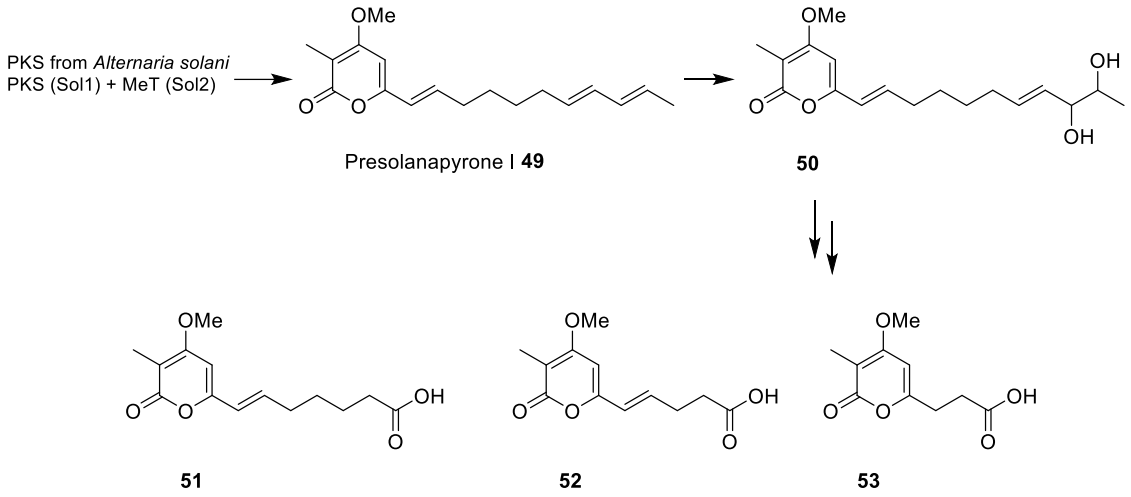
Heterologous expression of the PKS-NRPS (*TRIREDRAFT_58285*) and *trans*-acting ER (*TRIREDRAFT_58289*) in *A. oryzae* NSAR1 resulted in the production of one major compound and two minor ones with odd Mwt indicating the presence of an odd number of nitrogen atoms. The major compound and one of the minor compounds were isolated by preparative LCMS and the structures were identified based on the HRMS and NMR data. Structures of the two compounds were identified to be **137** and **138** and they showed high structure similarity to iliciolin-H tetramic acid. Although one of the minor compounds produced showed same mass and UV as the one reported for ilicicolin-H tetramic acid, the compound was produced in very low titers and its structure could not be confirmed by NMR.



Scheme 3.1 Proposed oxidation of illicicolin-H tetramic acid in *A. oryzae*.

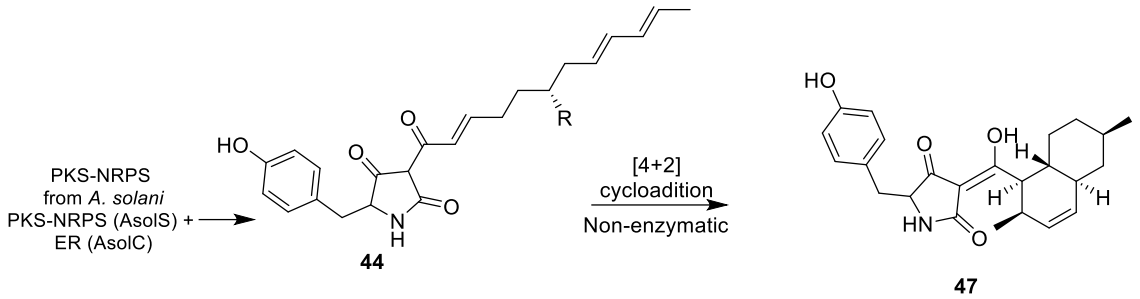
Time course of the *A. oryzae* + *PKS/NRPS-58285* + *ER-58289* transformant showed that after two days of cultivation the predicted ilicicolin-H tetramic acid **139** was the major compound produced. Compound **137** was only produced after three days of cultivation. Starting from day four, the production level of compound **139** decreased and that of compound **137** increased. This could indicate oxidative cleavage of the diene of compound **139**, which resulted in the production of the acid **137** with a shorter carbon chain (Scheme 3.1).

Aspergillus oryzae has been reported to perform native oxidation of heterologously expressed BGC. A similar oxidation process was reported in the expression of PKS *sol1* and the methyl trasnferase *sol2* from *Alternaria solani* in *A. oryzae*. The PKS product presolanapyrone I **49** was oxidated and this oxidation resulted in shunt products with shorter chain length **50** – **53** (Scheme 3.2).



Scheme 3.2 Shunt products produced during the expression of *sol1* and *sol2* form *A. solani* in *A. oryzae*.

Compound **138** showed a decalin structure resembling that of illcicolin-H and was only produced after longer periods of cultivation. This might indicate a slow spontaneous Diels-Alder reaction occurring in *A. oryzae* to yield the decalin structure, which was followed by oxidation of the terminal methyl group to yield compound **138**. A similar Diels-Alder reaction was also reported to occur spontaneously during the expression of the PKS-NRPS *AsolS* and the ER *AsolC* from *A. solani* in *A. oryzae* (Scheme 3.3).



Scheme 3.3 Shunt product produced by spontaneous Diels-Alder reaction of the tetramic acid produced by the expression of AsolS and AsolC from *A. solani*.

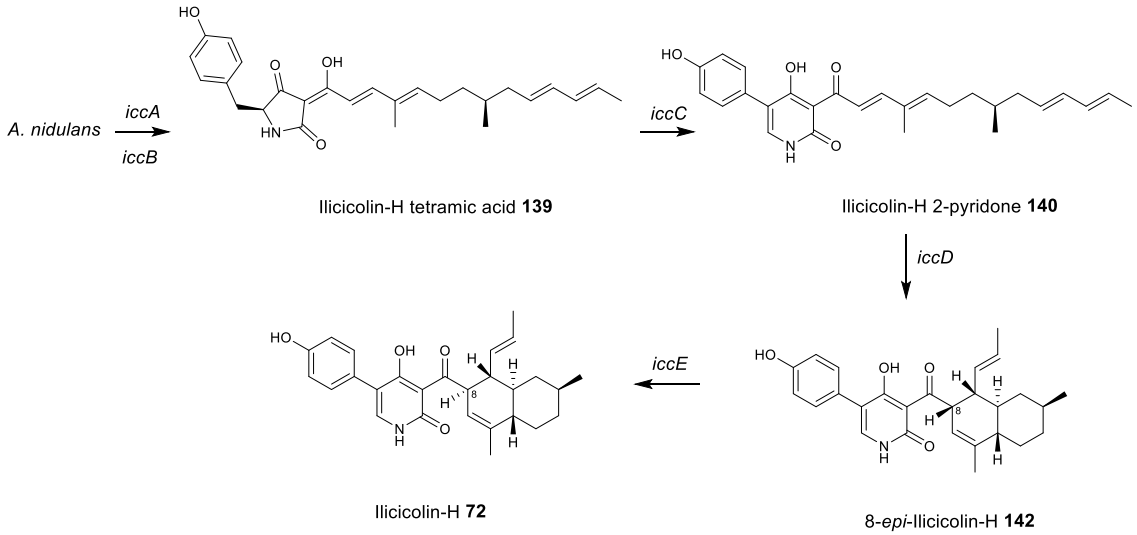
Expression the third gene in the BGC (*TRIREDRAFT_58953*), which is predicted to be a putative ring expansion P450, with the PKS-NRPS and ER genes resulted in the production of four strains. Analysis of the chemical profile of the *A. oryzae* + *PKS/NRPS-58285* + *ER-58289* + *P450-58953* strain showed the production of two new compounds. Isolation of the two compounds and structure elucidation by HRMS and NMR analysis showed the two compounds to be **140** and **141**, which are the 2-pyridone intermediate of illicicolin-H BGC and 2-pyridone shunt product.

Expression of the fourth gene in the cluster (*C*-MeT), which is predicted to be a Diels-Alderase, resulted in the production of two new compounds; one major compound with same mass and UV of illicicolin-H **72** and 8-*epi*-ilicicolin-H **142** and another very minor compound with the same mass, UV and retention time as that produced by the activation of the TF in *T. reesei*. After isolation of the major compound, NMR analysis showed that the compound is a mixture of the two compounds illicicolin-H **72** and 8-*epi*-ilicicolin-H **142**, with 8-*epi*-ilicicolin-H **142** being the major product. However, isolation of the minor compound was not possible due to the very low titres and the time constraints.

Illicicolin-H **72** is a broad spectrum antifungal agent that was isolated for the first time from *Cylindrocladium ilicicola* MFC-870.[215] It exerts its antifungal activity by inhibiting the yeast cytochrome bc1 complex at the Qn site.[89] The biosynthetic gene cluster of ilicicolin-H was identified by heterologous expression of candidate gene clusters from producing strains *Neonectria* sp.DH2[89] and *Penicillium variabile*[214] in the heterologous host *A. nidulans*.

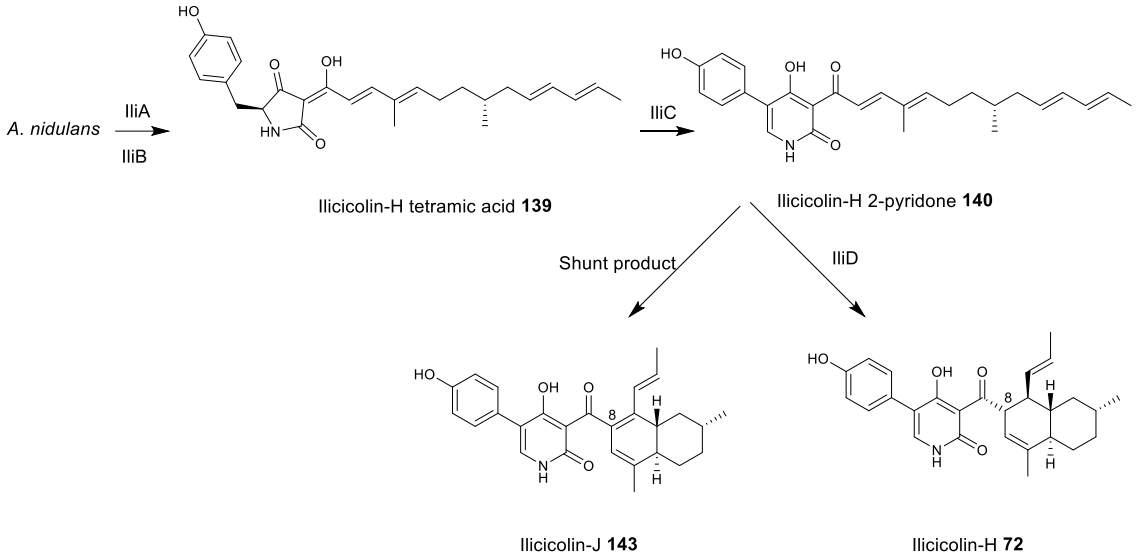
The ilicicolin-H BGC is reported to consist of genes encoding for a hybrid PKS-NRPS (IliA /IccA) and *trans*-acting ER (IliB /IccB) which act together to assemble the backbone for ilicicolin-H and then the structure is offloaded with a tetramic acid group.

The tetramic acid is then converted to 2-pyridone by a putative ring expansion P450 enzyme (IliC/ IccC). Finally, the decalin part of the structure is constructed by the action of intramolecular Diels-Alderase (IliD/IccD).[89,214] According to Zhang *et al*.[214] the structure produced by the Diels-Alderase is 8-*epi*-ilicicolin-H **142**, which is then converted by the epimerase (IccE) to illcicolin-H **72**. They also reported the spontaneous pH-dependent non-enzymatic conversion of 8-*epi*-ilicicolin-H **142** to ilicicolin-H **72**, in which increasing the pH of the reaction increased the epimerization rate (Scheme 3.4). The antifungal activity of Ilicicolin-H is reported to be 100 times that 8-*epi*-ilicicolin-H **142**, and therefore, previous derivatization efforts of ilicicolin-H to make antifungal analogues have always preserved the C-8 stereochemistry.



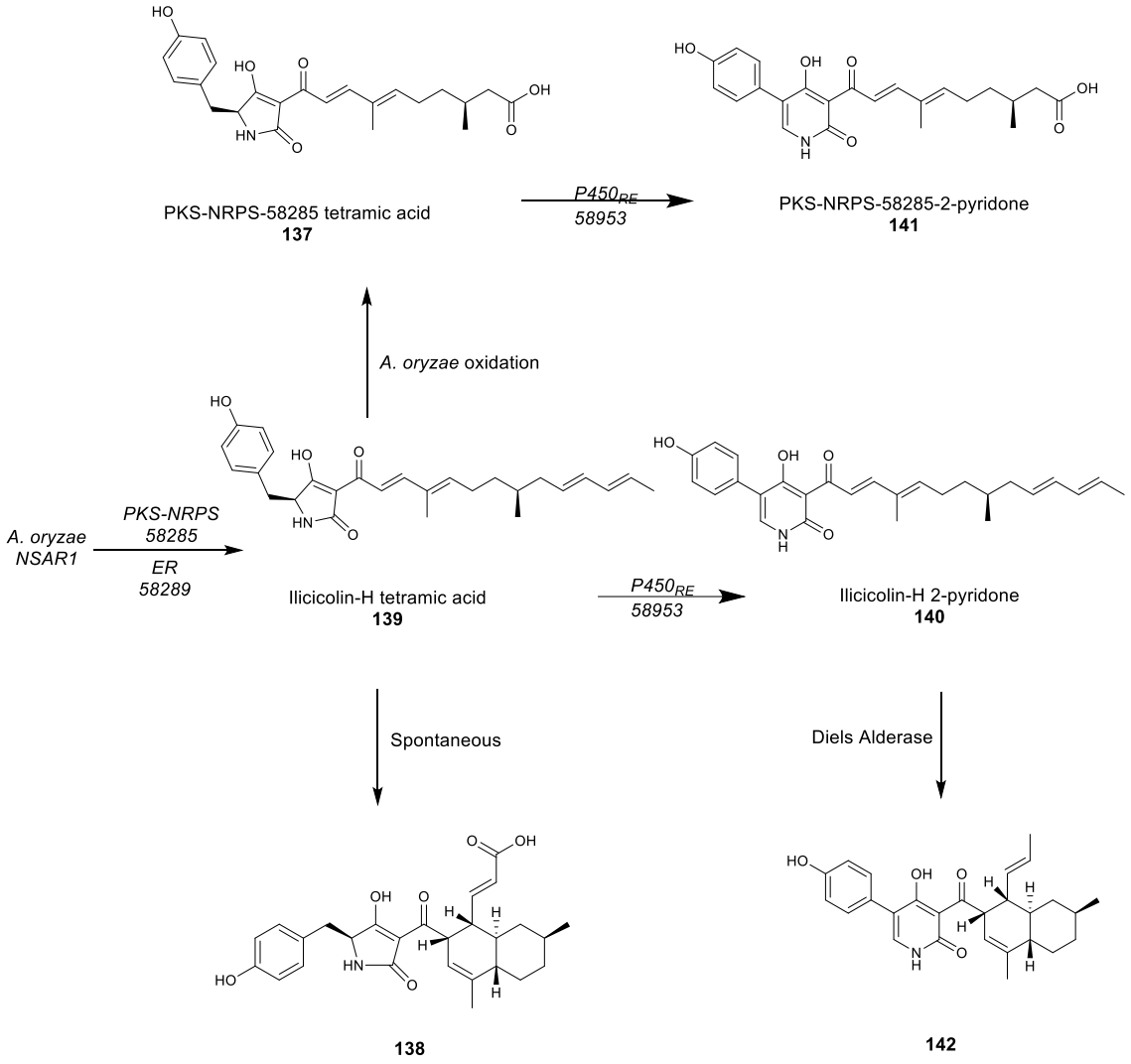
Scheme 3.4 Proposed biosynthetic pathway of ilicicolin-H **72** in *Penicillium variable*.[214]

However, Lin et al.[89] reported the production of iliciolin-H **72** and ilicicolin-J **143** by only expressing the first four genes in the cluster *(iliA-D*), showing that the oxidoreductase IliE is not directly involved in iliciolin-H **72** biosynthesis. They also showed that the new compound produced ilicicolin-J **143** has comparable antifungal activity to ilicicolin-H **72** despite the elimination of the stereochemistry at C-8 with the doube bond (Scheme 3.5).[89]



Scheme 3.5 Proposed biosynthetic pathway of ilicicolin-H **72** in *Neonectria* sp.DH2.[89]

In this work, heterologous expression of the first four genes in the cluster resulted in the production of 8-*epi*-ilicicolin-H **142** as major products in addition to ilicicolin-H **72** and other ilicicolin-H precursors **139**-**140** as minor compounds. Other shunt products **137**-**138 and 141** were also produced by the action of the native *A. oryzae* enzymes, which provide another example for the limitation of using some of the available fungal heterologous hosts and shows the importance of identifying and developing more fungal hosts for the expression of fungal natural products (Scheme 3.6).



Scheme 3.6 The proposed biosynthetic pathway of PKS/NRPS-58285 BGC in *A. oryzae* NSAR1.

3.6 Conclusion and Outlook

Although many PKS-NRPS products with very interesting chemistry have been reported from different *Trichoderma* spp., no PKS-NRPS products have been reported before from *T. reesei*. Therefore, in this work, one of the two cryptic PKS-NRPS gene clusters in *T. reesei* (TRIREDRAFT_58285) was chosen to be studied first by activation of the transcription factor TRIREDRAFT_79932 found 11.6 kb from the PKS-NRPS which resulted in the production of one new compound with odd molecular mass. This compound showed the same mass but different UV spectrum of that of ilicicolin-H **72**. However, isolation of this compound was not possible due to the very low titres. Hence, heterologous expression of this BGC was performed in *A. oryzae* NSAR1 strain to identify the product of this BGC.

Expression of the first two genes in the cluster (PKS-NRPS and the *trans*-acting ER) led to the production of two new compounds that were later identified to be two shunt products **137-138** probably produced by the activity of native *A. oryzae* enzymes on the ilicicolin H tetramic acid compound **139**. Addition of later genes in the cluster (*i.e.* P450 and Diels-Alderase) showed the production of ilicicolin-H biosynthetic intermediates **140** and **142** in addition to new shunt compound **141** (Figure 3.33).

This experiment shows the ability of the cryptic BGC from *T. reesei* to produce the antifungal ilicicolin-H. This gene cluster is also present in some other *Trichoderma* spp. despite the absence of any record on its isolation form *Trichoderma* spp. This compound might be one of the active constituents responsible for the biopesticidal activity of some *Trichoderma* spp. Therefore, expression of more BGC from *T. reesei* might lead to the isolation of new compounds with diverse biological activity that can be used as drug leads. This study also provides another example, in which the heterologous host enzymes can further modify the main products of the gene cluster to make new compounds.

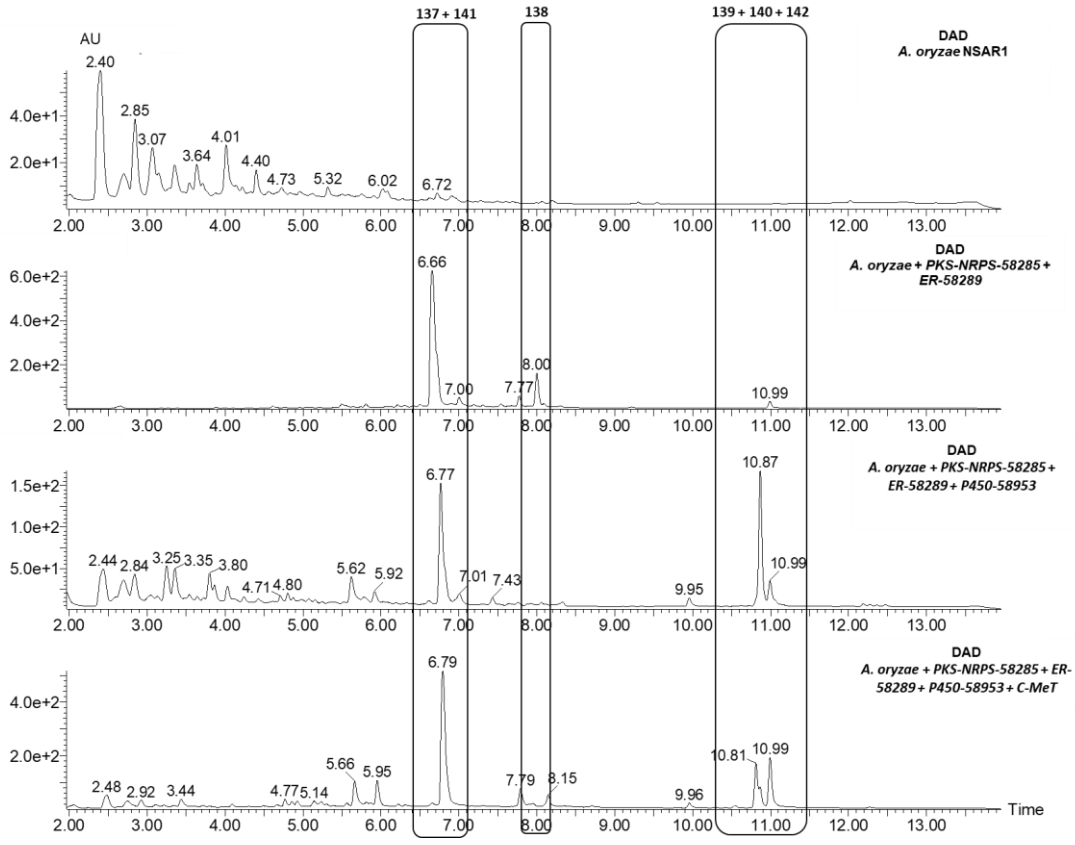


Figure 3.33 LCMS chromatograms of different *A. oryzae* transformants with ilicicolin-H **72** BGC against the WT *A. oryzae* NSAR1, showing the DAD chromatograms.

4 Overall Conclusion and Outlook

Overall, the main focus of the work described in this thesis has been to develop *Trichoderma reesei* as a heterologous host for the production of natural products and also to explore the inherent biosynthetic potential of this fungus. Due to its ability to grow on different carbon sources and waste material, *T. reesei* represent a promising platform for the production of chemicals on different waste materials. Although *T. reesei* has been used as a heterologous host for the production of cellulases and other proteins, its use as a heterologous host for the production of natural products has not yet been reported.

Therefore, in this work, *T. reesei* was developed to be a platform for natural product production. This was done first by knocking out two of the main genes in the sorbicillin pathway (*sorB* and *sorC*) to develop a strain with a cleaner secondary metabolic background. Then new vectors were developed to allow for the expression of up to three heterologous genes under native promoters. The new vectors were tested by expressing two different classes of megasynthases; PKS (Aspks1) and PKS-NRPS (TenS). Expression of *aspks1* under the activity of the *pdc* promoter (P_{pdc2}) resulted in the production of the expected 3-methylorcinaldehyde **40** and its oxidised congener 3-methylorsellinic acid **41**. The developed *T. reesei* strain expressing *aspks1* was cultivated on different waste materials such as potato peel, orange peel, banana peel, kiwi peel and barley straw. The strain was able to produce detectable amounts of 3-methylorcinaldehyde **40** on all tested waste material, with potato peels producing up to 128 $mg \cdot kg^{-1}$ dry wt.

Expressing more than one gene in *T. reesei* was also successful by the expression of the *tenS* encoded PKS-NRPS under the control of P_{pdc2} with its *trans*-acting ER (encoded by *tenC*) under the control of the native *cDNA1* promoter (P_{cdna1}). The *T. reesei* strain expressing TenS and TenC was able to produce the expected natural product pretenellin A **77**. Cultivation of the pretenellin A **77** producing strain on waste material (*e.g.* banana peel) also resulted in the production of pretenellin A **77**. Proving that *T. reesei* is a promising cell factory for natural product production.

This research paves the way for the use of *T. reesei* as a microbial cell factory for waste valorization. It represents a very promising, cheap, fast, sustainable and green alternative for synthetic chemistry in the production of natural products. The promoters

used in this work were mainly constitutive promoters to allow the production of natural products under all cultivation conditions. Nevertheless, these constitutive promoters used are reported to be less active than the tunable cellulase promoters are[130] and they also showed enhanced production levels in the presence of glucose in the media. Therefore, future experiments would be further enhancement of the system by the construction of new vectors with stronger promoters such as the tunable cellulase promoters *cbh1* and *cbh2*.[130] These promoters would allow the system to produce higher concentration of the expected compounds on lignocellulosic waste and other biomass.

The development of host strains with more selection markers could allow the expression of more vectors in parallel and hence more genes. For example, new selection marker for *T. reesei* has been recently reported from Derntl *et. al*.[220] with histidine auxotrophy in addition to the uridine auxotrophy (*T. reesei* QM6a·Δ*pyr4*·Δ*his1*). Furthermore, the use of other selection markers such as the auxotrophic *argB* and the dominant nutritional marker *amdS* were reported to be successful in *T. reesei*.[106] With the increasing interest in *T. reesei*, more and more selection markers and promoters are being reported and the toolbox for metabolic engineering of this fungus is expected to grow and expand widely in the near future.

Another aspect that should be considered in future work is the pre-treatment of the waste materials to increase the accessibility of cellulose and hemicellulose in the media. In this work, only minimal pre-treatment was used by autoclaving the waste materials with water. Although the fungal strains were still able to grow and produce the expected compounds on such waste materials with minimal pre-treatment, proper pre-treatment of the waste media could result in better growth and higher production levels of the compounds. Many pre-treatment methods for lignocellulosic waste have already been reported.[221] In addition, more extensive optimization of the cultivation conditions could allow higher production levels in shorter period and could result in a system that is able to produce natural products from waste in an industrial scale.

Furthermore, *T. reesei* was reported to grow on many different types of waste such as pre-treated paper waste, wheat straw, *etc*.[72,143] Therefore, testing the ability of the strain to grow and produce natural products on other types of waste materials could open many new avenues for waste management and valorization. For example, according to Eurostats on waste, paper and cardboard was the main packaging waste material in the EU, with over 30 million tonnes in 2018.[222] Also in this research, all extractions were

done using the environmentally benign solvent ethyl acetate. However, testing other greener methods and solvents for the extraction of natural products from the waste would also be beneficial in developing even more efficient and eco-friendly system for natural product production.

The other main aim of this thesis was to explore the biosynthetic potential of *T. reesei*. A bioinformatic analysis of all PKS and PKS-NRPS genes in *T. reesei* showed the presence of 11 PKS and two PKS-NRPS genes in the *T. reesei* genome. Because no PKS-NRPS-related compounds were previously reported from *T. reesei*, one of the PKS-NRPS genes (*TRIREDRAFT_58285*) was chosen for further investigation. Activation of the chosen PKS-NRPS gene cluster by ectopic overexpression of the pathway transcription factor showed limited success, resulting in the production of a very minor compound that could not be isolated due to its very low titre. However, heterologous expression of this gene cluster in *A. orzyae* resulted in the production of ilicicolin-H biosynthetic intermediates **139**-**140** and **142** in addition to other new shunt products **137**-**138** and **141**. Iliciolin-H **72**, which is an antifungal drug, was not previously reported from *T. reesei* or any other *Trichoderma* spp. The presence of an active BGC for ilicicolin-H in *T. reesei* might contribute in the understanding of the biocontrol activity reported for this fungus.

Heterologous expression of *T. reesei* BGC in *A. oryzae* was proved to be an efficient method for exploring and understanding the biosynthetic potential of this fungus. The *T. reesei* genome contains 11 PKS genes, with only four of them either explored or have high similarity to already explored gene clusters from other fungal strains. Future experiments to express other BGC from *T. reesei* in other heterologous hosts can result in discovery of new natural products and expand the knowledge and understanding of the biosynthetic ability of this fungus.

5 Experimental

All chemicals and media ingredients used in this work were purchased from Duchefa Biochemie, Roth, VWR, Fisher scientific, Sigma Aldrich, abcr and Formedium.

5.1 Growth Media, Buffers, Solutions and Antibiotics.

All media, buffers, solutions and antibiotics were prepared with double distilled water (dd.H_2O), where deionized water was further purified by GenPure Pro UV/UF Milipore device from Thermo scientific. All growth media and sterile solutions were autoclaved at 121 °C for 15 min using Systec VX150 or a classic Prestige Medical 2100 autoclave. Antibiotic solutions were sterilized using a sterile syringe filter (0.45µm pore size, Roth).

5.1.1 Media

Table 5.1 Growth media used in this work.

Media	Composition [% (*w/v*)]	Ingredients
GNB	2.00	D(+)-Glucose monohydrate
	3.00	Nutrient broth Nr. 2 from Oxoid
LB	0.50	Yeast extract
	1.00	Tryptone
	0.50	NaCl
SOC	93.75 (*v/v*)	SOB
	1.25 (*v/v*)	MgCl2 x 6 H2O
	5.00 (*v/v*)	D(+)-glucose 20 %
YPAD	1	Yeast extract
	2	Tryptone
	2	D(+)-Glucose Monohydrate
	0.03	Adenine
ME (BF)	2.00	Malt extract
	0.10	Peptone ex Soya
	2.00	D(+)-Glucose monohydrate
Mandels Andreotti (MA) + 1% glucose	0.14	$(NH_4)_2SO_4$
	0.20	KH_2PO_4
	0.03	$CaCl_2$. 2(H_2O)
	0.03	$MgSO_4$. 7(H_2O)
	0.03	Urea
	0.1	Peptone (casein)
	0.0005	$FeSO_4$. 7(H_2O)
	0.0016	$MnSO_4$. H_2O
	0.0014	$ZnSO_4$. 7(H_2O)
	0.0002	$CoCl_2$. 2(H_2O)
	1.00	Glucose monohydrate

Continue Table 5.1 Growth media used in this work.

DPY	2.00	Dextrin from potato starch
	1.00	Polypeptone
	0.5	Yeast extract
	0.5	KH_2PO_4
	0.05	$MgSO_4$. 6(H_2O)
DPY + 1% glucose	2.00	Dextrin from potato starch
	1.00	Polypeptone
	0.5	Yeast extract
	0.5	KH_2PO_4
	0.05	$MgSO_4$. 6(H_2O)
	1.00	D(+)-Glucose Monohydrate
SS	5.00	Soya sauce
	5.00	Sucrose
PDB	2.40	Potato Dextrose broth
CM	2.00	Cottonseed flour
	10.00	Lactose monohydrate
CMP	3.50	Czapex Dox broth
	2.00	D(+)-Maltose monohydrate
	1.00	Polypeptone

Table 5.2 Solid media used in this work.

Agar	Composition [% (*w/v*)]	Ingredients
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
	2.50	Agar
LB agar	0.5	Yeast extract
	1.00	Tryptone
	0.5	NaCl
	1.5	Agar
ME agar	1.28	Malt extract
	0.08	Peptone ex Soya
	0.24	Glycerol
	0.28	Dextrin from potato starch
	1.5	Agar
ME (BF) agar	2.00	Malt extract
	0.10	Peptone ex Soya
	2.00	D(+)-Glucose monohydrate
	1.50	Agar
YPAD agar	1.00	Yeast extract
	2.00	Tryptone
	2.00	D(+)-Glucose Monohydrate
	0.03	Adenine
	1.50	Agar
CZD/S agar	3.50	Czapek Dox broth
	18.22	D-sorbitol (1M)
	0.10	$(NH_4)_2SO_4$
	0.05	Adenine
	0.15	L-Methionine
	1.50	Agar
CZD/S soft agar	3.50	Czapek Dox broth
	18.22	D-sorbitol (1M)
	0.10	$(NH_4)_2SO_4$
	0.05	Adenine
	0.15	L-Methionine
	0.80	Agar

Continue table 5.2 Agar used in this work.

Agar	Composition [% (*w/v*)]	Ingredients
CD1 agar	3.50	Czapek Dox broth
	4.68	NaCl (0.8 M)
	0.10	$(NH_4)_2SO_4$
	0.05	Adenine
	0.15	L-Methionine
	1.50	Agar
Mandel's Agar w/o Peptone, w Glycerol, w Glucose	0.03	Urea
	0.14	$(NH_4)_2SO_4$
	0.02	KH_2PO_4
	0.03	$CaCl_2$. $2(H_2O)$
	0.03	$MgSO_4$. $6(H_2O)$
	0.01	Tween 80
	1.00	Carboxymethyl cellulose
	0.0005	$FeSO_4$. $7(H_2O)$
	0.0016	$MnCl_2$. $4(H_2O)$
	0.0017	$ZnCl_2$
	0.0002	$CoCl_2$. $6(H_2O)$
	1.00	Glycerol
	1.00	D(+)-Glucose Monohydrate
	1.5	Agar
Mandel's Agar w/o Peptone, w Glycerol, w Glucose, w sorbitol	0.03	Urea
	0.14	$(NH_4)_2SO_4$
	0.02	KH_2PO_4
	0.03	$CaCl_2$. $2(H_2O)$
	0.03	$MgSO_4$. $6(H_2O)$
	0.01	Tween 80
	1.00	Carboxymethyl cellulose
	0.0005	$FeSO_4$. $7(H_2O)$
	0.0016	$MnCl_2$. $4(H_2O)$
	0.0017	$ZnCl_2$
	0.0002	$CoCl_2$. $6(H_2O)$
	1.00	Glycerol
	1.00	D(+)-Glucose Monohydrate
	18.22	Sorbitol (1M)
	1.5	Agar
DPY agar	2.00	Dextrin from potato starch
	1.00	Polypeptone
	0.5	Yeast extract
	0.5	KH_2PO_4
	0.05	$MgSO_4$. $6(H_2O)$
	2.50	Agar
PD agar	2.40	Potato Dextrose broth
	1.50	Agar
DPY + 1% glucose agar	2.00	Dextrin from potato starch
	1.00	Polypeptone
	0.5	Yeast extract
	0.5	KH_2PO_4
	0.05	$MgSO_4$. $6(H_2O)$
	1.00	D(+)-Glucose monohydrate
	2.50	Agar

5.1.2 Buffers and solutions

Table 5.3 Solutions:

Solutions	Components
Protoplasting solution for *T. reesei*	1.2 M Sorbitol, 100 mM KH_2PO_4, pH 5.6
Washing solution for *T. reesei*	1.2 M Sorbitol, 10 mM Tris-HCl, pH 7.5
Resuspending solution for *T. reesei*	1 M Sorbitol, 10 mM $CaCl_2$, 10 mM Tris-HCl pH 7.5
PEG 6000 solution for *T. reesei*	50 mM $CaCl_2$, 10 mM Tris-HCl pH 7.5, 25 % (w/v) PEG6000, pH 7.5
Protoplasting solution for *A. oryzae*	0.8 M NaCl, 10 mg/mL T*richoderma harzianum* lysing enzyme
Transformation solution I for *A. oryzae*	0.8 M NaCl, 10 mM CaCl2, 50 mM Tris-HCl, pH 7.5
Transformation solution II for *A. oryzae*	0.8 M NaCl, 10 mM CaCl2, 50 mM Tris-HCl, pH 7.5, 60 % (w/v) PEG3350
10 x TE buffer	100 mM Tris HCl pH 8, 10 mM EDTA
TRIzol	Purchased from Thermo Fischer Scientific
Loading-dye (6 x)	0.25 % (*w/v*) bromophenol blue, 30 % (*v/v*) glycerine, 0.25 % (w/v) xylene cyanol
50 X TAE buffer	2 M Tris acetate, 0.05 M EDTA, pH 8.3

5.1.3 Enzymes and Antibiotics

The enzyme and antibiotics used in this work are summarized table 5.4 and 5.5.

Table 5.4 Enzymes.

Name	Function	Reference
AscI	Restriction enzyme	NEB
PacI	Restriction enzyme	NEB
SwaI	Restriction enzyme	NEB
NotI	Restriction enzyme	NEB
High-capacity RNA-to-cDNA™ Kit	reverse transcriptase	Thermo Fisher Scientific
Gateway™ LR Clonase™ II Enzyme Mix	contains integrase, excisionase and integration host factor	Thermo Fisher Scientific
OneTaq® 2X Master Mix with Standard Buffer	DNA polymerase with agarose gel running buffer	NEB
Proteinase K	Proteinase	Thermo Fisher Scientific
Q5® High-Fidelity DNA Polymerase 2x Master Mix	DNA polymerase with enhanced proofreading ability	NEB
***Trichoderma harzianum* lysing enzyme**	Cell Dissociation and cell lysis; preparation of protoplasts	Sigma
Driselase	Cell Dissociation and cell lysis; preparation of protoplasts	Sigma

Table 5.5 Antibiotics.

Name	Stock concentration (mg/mL in H_2O)	Final concentration (µg/mL)
Carbenicillin disodium salt (Carb)	50	50
Kanamycin (Kan)	50	50
Hygromycin B (hyg)	50	100-150

5.2 Microbiological Methods

5.2.1 Bacterial and Fungal Strains

Bacterial and fungal strains used in this thesis are summarized in table 5.6.

Table 5.6 Bacterial and fungal strains.

Strain	Genotype	Origin
E. coli 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	Thermo Fisher Scientific
E. coli ccdB survival cells	F-*mcrA* 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	Thermo Fisher Scientific
S. cerevisiae CEN.PK2	*MATa/α ura3-52/ura3-52 trp1-289/trp1-289 leu2-3_112/leu2-3_112 his3 Δ1/his3 Δ1MAL2-8C/MAL2-8C SUC2/SUC2*	Hahn group, Hannover
Trichoderma reesei-QM6a Δtmus53 Δpyr4	*Δtmus53, Δ pyr4*	Mach-Aigner group, Vienna [120]
Trichoderma reesei-QM6a Δtmus53 Δasl1	*Δtmus53, Δ asl1*	Mach-Aigner group, Vienna[120]
A. oryzae NSAR1	*ΔniaD, ΔsC, ΔadeA, ΔargB*	Lazarus group, Bristol [104]
Trichoderma reesei-QM6a Δtmus53 Δpyr4 ΔsorBC	*Δtmus53, Δ pyr4 ΔsorBC*	This work

5.2.2 *Escherischia coli*

Growth and Maintenance

E. coli strains were cultivated on liquid and solid LB medium with the approperiate antibiotic. Cultures were incubated at 37°C and 200 rpm for 16-24 hours. Glycerol stocks (25% glycerol) were stored at -80°C for long term storage of the strains.

Transformation of Chemically Competent *E. coli*

Competent *E. coli* cells stored at -80°C were thawed on ice for 5 min. 1-10 µL vector DNA were added and cells were stored on ice for 20-30 min. Heat shock was conducted at 42°C for 30-45 s and cells were immediately placed on ice for 2 min. 250 µL SOC medium were then added and samples were incubated at 37°C with 300 rpm shaking for 1 hr. The cells were spinned down for 15 s and 250 µL of the supernantatnt was discarded. The bacterial cells were then suspended in the remaining supernatant and spread on LB agar plates containing the appropriate antibiotic and incubated over night at 37°C.

5.2.3 *Saccharomyces cerevisiae*

Growth and Maintenance

S. cerevisiae stain was cultivated on YPAD agar and incubated at 30°C for three to five days. A single colony was used to inoculate 10 ml liquid YPAD media and was then grown overnight at 30°C and 200 rpm. After transformation with *ura3* containing plasmids, *S. cerevisiae* strains was grown on SM-Ura agar plates at 30°C for three to five days.

Preparation and transformation of competent cells

The high efficiency version of LiAc/SS carrier DNA/PEG protocol developed by Gietz and Schiestl[223] was used for the preparation and transformation of competent *S. cerevisiae* cell. A single colony was transferred into 10 mL YPAD medium and incubated over night at 30°C with 200 rpm shaking. This starter culture was added to 40 mL YPAD medium in a 250 mL Erlenmeyer flask and incubated for another 4.5 h at 30°C with 200 rpm shaking. Culture was harvested by centrifuging at 3.000 x g for 5 min. The cell pellet was first washed with 25 mL dd.H_2O before it was resuspended in 1 ml dd.H_2O, transferred into a 1.5 mL reaction tube and centrifuged at 13.000 xg for 30 s and supernatant discarded. The pellet was suspended in 400 µL water and each 100 µL were transferred into a separate 1.5 mL reaction tube

For yeast recombination based cloning using *S. cerevisiae*, 100 µL aliquots of yeast competent cells were centrifuged for 30 s at 13.000 x g and the supernatant was discarded. The following components were added to the pellet in order: 240 µL PEG solution (50 % (*w/v*) polyethylene glycol 3350), 36 µL LiAc (1 M), 50 µL denaturated salmon testis DNA (2 mg.mL^{-1}), 34 µL DNA containing the linearized plasmid and desired inserts obtained by PCR in equimolar concentration. The PCR fragments contain each 30 bp overlap at both 5` and 3` with the cut sites of the vector fragments to facilitate homologous recombination. The uncut plasmid was used for positive and linearized plasmid or dd.H_2O was used for negative control. The pellet was resuspended in the transformation mix by vigorous vortexing and incubated for 40 min at 42°C. Cells were pelleted by centrifugation at 13.000 xg for 30 s and supernatant was removed. The pellet was resuspended in 500 µL dd.H_2O and each 250 µL were spread on selective SM-Ura plates, which were incubated for three to four days at 30°C. Yeast plasmid was extracted

using a Zymoprep™ Yeast Plasmid Miniprep II kit (Zymo reserach, Orange, California, USA).

5.2.4 *Trichoderma reesei*

Growth and Maintenance

T. reesei was grown on ME or PD agar at 28°C. Transformed *T. reesei* strains were grown either on minimal media (Mandel`s agar without peptone + 1% glucose + 1% glycerol) or DPY+ 1% glucose agar and incubated at 28°C for 3-7 days. For glycerol stocks 5 mL of dd.H_2O were added to the plate and spores were scraped off using a sterile spatula. 800 µL of spore solution were supplemented with 800 µL glycerol (50%) and stored in a cryovial at -80°C. For liquid cultures 100 mL DPY + 1 % glucose medium or ME (BF) media was inoculated with *T. reesei* spore solution and incubated at 28°C for three to six days with 110 rpm shaking.

PEG-Mediated Transformation

The conidia from sporulating plate or spores stored as glycerol stocks were inoculated in 100 ml PDB or DPY media and incubated at 28°C with shaking at 110 rpm for 2 days. The germinated conidia were separated from the liquid medium by filtration through sterile Miracloth. 20 ml of filter-sterilized protoplasting solution containing 12 mg/ml *Trichoderma* lysing enzyme or 10 mg/ml Driselase™ enzyme was used to resuspend the mycelia. The mixture was then incubated at 28°C with shaking at 110 rpm for 2-3 hours. Protoplasts were separated from undigested mycelial debris by filtering through Miracloth and then centrifugation at 3500 x g for 3 minutes to separate protoplasts. Additional washing step (optional) was done by suspending the protoplasts in 15 ml washing solution for *T. reesei* then centrifugation at 3500 x g for 3 min. Protoplasts were then resuspended in 300 µl resuspension solution.

The concentration of the protoplasts was assessed microscopically. 10 µL of plasmid DNA was mixed with 100 µl of the protoplast suspension and 200 µl of PEG 6000 solution and the mixture was incubated on ice for 20 min. 0.5 ml of PEG 6000 solution and 1 ml resuspension solution were added to the transformation mixture and then the whole mixture was incubated at room temperature for 5 min. Aliquots of 300 ml of the solution were mixed with 20 ml of 50°C Mandel's Agar without peptone and with 1% glycerol, 1% glucose and 1 M sorbitol. The plates were incubated at 28°C for 3-7

days until the colonies appeared. For selection on uridine (or arginine) deficient media, the selection was repeated three times by picking the colonies using sterile tooth picks and transferring them into new minimal media plates (Mandel's Agar without peptone and with 1% glycerol and 1% glucose).

5.2.5 *Aspergillus oryzae NSAR1*

Growth and Maintenance

A. oryzae NSAR1 was grown on DPY agar at 28°C. For glycerol stocks, 5 ml of dd.H_2O were added to the plate and spores were scraped off using a sterile woody spatula. 800 µl of spore solution were supplemented with 800 µl glycerol (50 %) and stored in a cryovial at -80°C. For liquid cultures 100 ml DPY medium were inoculated with *A. oryzae* spore solution and incubated at 28°C for two to seven days with 110 rpm shaking.

PEG-mediated Transformation of *A. oryzae* NSAR1

A. oryzae NSAR1 conidia from sporulating plates were inoculated into 100 ml DPY media and incubated overnight at 28°C with shaking at 110 rpm. Germinated conidia were collected by filtrating with the sterile Miracloth and the pellet was resuspended in 15 mL of filter-sterilized protoplasting solution (10 mg/mL *Trichoderma* lysing enzyme in 0.8 M sodium chloride). The conidia was then incubated with the protoplasting solution for three hours at 28°C with gentle mixing. The protoplasts were released from hyphal strands by gentle pipetting with wide-bore pipette and filtered through sterile Miracloth. The filtrate was then centrifuged at 3000 x g for 5 min to pellet the protoplasts which were resuspended in solution 1 (100 µl per transformation) and the concentration of the protoplasts was assessed microscopically. 10 µl of the plasmid DNA was added to 100 µl of the protoplast suspension and incubated on ice for 2 min.

Then 1 ml of solution 2 was added and the transformation mixture was incubated at room temperature for 20 min. 5 ml of molten CZD/S soft agar at 50°C was added to the transformation mixture and overlaid onto pre-prepared CZD/S and the plates were incubated at 28°C for 3-5 days. For selection on arginine–deficient media, the selection has to be repeated twice to prevent false positives due to consuming arginine from dead cell material. A few spores of the grown colony were picked with a tooth pick and transferred to a CD1 plate. The plates were incubated at 28°C for 3–4 days and the step

was repeated by streaking out single colonies. Finally, a single colony was grown on DPY agar for one week at 28°C before fermentation.

5.3 Molecular Biology Methods

All enzymes used was purchased from NEB and Thermo Fisher scientific and was used according to the manufacturer's instructions using the supplied buffers.

5.3.1 Vectors Used in this Work

The vectors used in this thesis are summarized in table 5.7 and 5.8.

Table 5.7 Vectors kindly provided for this work

Name	Origin
pTYGS-*asl1-pdc*	Eman Bassiony
pTYGS-*argB*	Lazarus group, Bristol
pEYA-*aspks1*	Dr. Raissa Shor
RSI96-1-6	Dr. Raissa Shor
pEYA-*TenS*	Katharina Schmidt
pTYGS-*argB-TenS-TenC*	Katharina Schmidt

Table 5.8 Vectors constructed in this work

vector	ID	Vector backbone	Oligonucleotides for construction in *S. cerevisiae.*
pTYGS-*asl1-*P_{pdc1}*-aspks1*	MSIII67	pTYGS-*asl1-pdc*	LR with pE-YA-*aspks1*
pTYGS-*pyr4-*P_{amyB}*-aspks1-asL1-asL3-asR2*	MSIII84	RSI96-1-6	*Pyr4*: 1659 + 1660
pE-YA-*sorBC* Knockout	MSIII132	pE-YA	P_{a+b}: 1678 + 1679 P_{c+d}: 11680 +1681 P_{e+f}: 1682 + 1683
pTYGS-*pyr4-*P_{pdc2}	MSIII158	pTYGS-argB	*Pyr4*: 1690 + 1691 *Pyr*-5`flank: 1684 + 1685 *Pyr4*-3`flank: 1686 + 1687 P_{pdc2}: 1688 + 1689
pTYGS-*pyr4-*P_{pdc2}*-aspks1*	MSIII163	MSIII158	LR with pE-YA-aspks1
pTYGS-*pyr4-*P_{pdc2}*-*$P_{TRcDNA1}$*-*P_{TReno}	MSIII172	MSIII158	P_{cDNA1}: 1694 + 1695 P_{TReno}: 1696 + 1697 Patch P_{gpdA}: 91 + 92 Patch AscI: 1692 + 1693
pTYGS-*pyr4-*P_{pdc2}*-TenS*	MSIII173-A	MSIII172	LR with pE-YA-*TenS*
pTYGS-*pyr4-*P_{pdc2}*-TenS-*$P_{TRcDNA1}$*-TenC*	MSIII174	MSIII173-A	*tenC*: 1700 + 1701 Patch P_{TReno}: 1698 + 1699

Continue table 5.8 Vectors constructed in this work

vector	ID	Vector backbone	Oligonucleotides for construction in *S. cerevisiae.*
pE-YA-*PKS-NRPS-58285*	MSIII139	pE-YA	P_{a+b}: 1543+ 1544 P_{c+d}: 1545 +1546 P_{e+f}: 1547 + 1548 P_{g+h}: 1549 + 1550
pTYGS-*argB-PKS-NRPS-58285*	MSIII144	pTYGS-*argB*	LR with MSIII139
pTYGS-*argB-PKS-NRPS-58285-ER*	MSIII152	MSIII144	ER-58289: 1551 + 1552 Patch adh: 89 + 90 Patch Peno: 87 + 88
pTYGS-*argB-PKS-NRPS-58285-ER-P450*	MSIII165-A	MSIII144	ER-58289: 1551 + 1552 P450-58953: 1553 + 1554 Patch Peno: 87 + 88
pTYGS-*argB-PKS-NRPS-58285-ER-P450-C-MeT*	MSIII165-B	MSIII144	ER-58289: 1551 + 1552 P450-58953: 1553 + 1554 C-MT: 1555 + 1556
pE-YA-*TF-72993*	MSIII168-A	pE-YA	TF-72993: 1539 + 1540
pTYGS-*pyr4-P_{pdc2}-TF-72993*	MSIV05	MSIII158	LR with MSIII168-A

5.3.2 Oligonucleotides used in this work

The oligonucleotides used in this work are summarized in table 5.9.

Table 5.9 Primers and oligonucleotides.

NO.	Name	sequence
11	ITS1	TCCGTAGGTGAACCTGCGG
12	ITS4	TCCTCCGCTTATTGATATGC
83	P1asPKS1HygF	ATGGCAGCTCATGGGCAAAC
87	Peno plugF	CTTCTTAAATATCGTTGTAACTGTTCCTGA
88	Peno plugR	CGAAGTATATTGGGAGACTATAGCTACTAG
89	Padh plugF	ATTCACCACTATTATTCCCACCCTATAATA
90	Padh plugR	GAGACGAAACAGACTTTTTCATCGCTAAAA
91	PgdpA plugF	CTTTTCTTTTCTCTTTCTTTTCCCATCTTC
92	PgpdA plugR	TGACCTCCTAAAACCCCAGTG
322	ASPKS12095R	CTGCTGGCTTAACACGTGC
421	asL1 –P1	TTCTTTCAACACAAGATCCCAAAGTCAAAGATGGACAGCCCAGAAGTAT
423	asL1-P2	TTTCATTCTATGCGTTATGAACATGTTCCCTTAAAGAGTATAGCCGCC
424	asL3-P1	ACAGCTACCCCGCTTGAGCAGACATCACCGATGGGCAGCCTCACTGAT
426	asL3-P2	TACGACAATGTCCATATCATCAATCATGACCTATGGTAGCACTACTGGC
427	asR2-P1	CGACTGACCAATTCCGCAGCTCGTCAAAGGATGGCTCTCGCACAGCAA
428	asR2-P2	CAGGTTGGCTGGTAGACGTCATATAATCATTCATTCTTTGTCCGAGCG
682	TenS-TDS-F	CAAGCTCAGTCACTCGCTCA
683	TenS-TDS-R	CACTATCCTCGCCGATGCTT
1327	Hyg LF RP	CGTCAGGACATTGTTGGAG
1328	Hyg RF FP	GCTTTCAGCTTCGATGTAGG
1539	TF-788- PF	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGGAGTCGAGCACCACGAC
1540	TF-788- PR	TGCCAACTTTGTACAAGAAAGCTGGGTCGGCTAAACAACCGTGCACCGCT
1541	Ten check PF	AGCTGGCAGTCGACCCATCC
1542	Ten Check PR	CGTCTGGATCGGGCCTGTGT

Continue table 5.9 Primers and oligonucleotides.

NO.	Name	sequence
1543	IlliA-P1	TAATGCCAACTTTGTACAAAAAAGCAGGCTATGGATCAACAAAGACAGCA
1544	IlliA- P2	ACACCTTGTCGTGGTCCCACGAGTATAAAGGCAATCCCTTGGGGATTTTG
1545	IlliA- P3	CTTTATACTCGTGGGACCACGACAAGGTGTTCTGGCGCGAGGGACGTCTA
1546	IlliA- P4	TCTCCGGGCTGCAATCTTGCTCATCAATGGGGATGACGTGTGTGCCATGA
1547	IlliA- P5	CCATTGATGAGCAAGATTGCAGCCCGGAGAGGCTCTCTGCTATTGCAAAT
1548	IlliA- P6	CCAGAACAACGCCCAGCCCAATGACATCCATGACAATGTGATGGTATGCG
1549	IlliA- P7	TGGATGTCATTGGGCTGGGCGTTGTTCTGGACGACTTGAACAATGCTTAC
1550	IlliA- P8	TATAATGCCAACTTTGTACAAGAAAGCTGGTTAATCGCTTCCCATATAAG
1551	ER-58289-gpdA-PF	ACAGCTACCCCGCTTGAGCAGACATCACCGATGACTGTCATTGATGTGCT
1552	ER-58289-gpdA-PR	TACGACAATGTCCATATCATCAATCATGACTCAAGCCGATACAGCAATGG
1553	P450-58953-adh-PR	TTTCATTCTATGCGTTATGAACATGTTCCCTCACTTGTAGAAAGCTATCT
1554	PF-long P450-adh	TTCTTTCAACACAAGATCCCAAAGTCAAAGATGTTGTTGCCCGAGTACGT
1555	PF C-MT eno	CGACTGACCAATTCCGCAGCTCGTCAAAGGATGTCTTCTACTCAAACAAC
1556	PR C-MT eno	CAGGTTGGCTGGTAGACGTCATATAATCATCTACTTGCTTTTGTGGTCTT
1659	Pyr4 cloning P1	GAAATAATCCTGTGGGTCTTTGTCCGACTCTTGCTAAATGCCTTTCTTTA
1660	Pyr4 cloning P2	GGCCTCATTTGACTATAATTTACATAAATTACGAAAAAGGTCTGACTGGG
1661	Pyr4 seq P1	CCAAACTTGCGGTCCTCGAA
1662	Pyr4 seq P2	TGTTCGGCTCAGGTCGCTCA
1663	Aspks1-P1-F	CGCATTACTCCCATCATCAC
1664	aspks1-seq-P2	CGCCAACTGCAAAAAGAAGG
1665	aspks1-seq-P3	CACAGCAACCGGAAACCTCC
1666	aspks1-seq-P4	CAGCCGGGATCAAGTTTCCT
1667	aspks1-seq-P5	ACACGCTTTCCACTCGGTTC
1668	aspks1-seq-P6	CAACGTTAGTGGACAATGGA
1669	aspks1-seq-P7F	AGCAAGTTGGTACAATGGAG
1670	aspks1-seq-P7R	GCTTCGAGCCAGTTTTGCTA
1671	aspks1-seq-P8	ACTAGCCTCGAACGAGTGCG
1672	aspks1-seq-P9	ACTGGGACAGATCGCAGGCA
1673	aspks1-seq-P10R	TCACAACAAGAACCCCACCT
1674	Asl seq P1	GGTAGTAAGTCGGCGTCTTG
1675	Asl1 seq P2	TTGGCCGAAAGCACGCCCTG
1676	Asl1 seq P3	GATCCAAGGATATCTTCCTT
1677	Asl1 seq P4	TAGACTAGACTAGTAATTGT
1678	SorBC-KO-long-RF	GCCAACTTTGTACAAAAAAGCAGGCTCCGCATGGCGGCCTCAAGTACACG
1679	SorBC-KO-long-RR	ACGTATTTCAGTGTCGAAAGATCCACTAGAATGGCTCAGACTTGCCTTGG
1680	SorBC-KO-long-HygF	ATCAAACGCGCCAAGGCAAGTCTGAGCCATTCTAGTGGATCTTTCGACAC
1681	SorBC-KO-long-HygR	CAACATCCGCCTTCGCAGTTGACCCATCGCCAGGTCGAGTGGAGATGTGG
1682	SorBC-KO-long-LF	AGCGCCCACTCCACATCTCCACTCGACCTGGCGATGGGTCAACTGCGAAG
1683	SorBC-KO-long-LR	AACTTTGTACAAGAAAGCTGGGTCGGCGCGCTAATGCTTCTCTAACACCT
1684	PF-Pyr4-5‘flank	AAGAAAGAAGTAAAGAAAGGCATTTAGCAAAGATCTCGAGATAGTATCTC
1685	PR-Pyr4-5‘flank	GACAAGAGCTCTTTCTCAGCTCTTAATTAATCCGGAGTAGCTCTTCACTG
1686	PF-pyr4-3‘ flank	ACCAGTGAAGAGCTACTCCGGATTAATTAAGAGCTGAGAAAGAGCTCTTG
1687	PR-pyr4-3‘ flank	TCTGCCGCGCCAAGTAGCCCTGGAAGTCCTCGTGCGTCTCGTTGTGCTCG
1688	PF-Ppdc	CACCGCCCGCGAGCACAACGAGACGCACGAGGACTTCCAGGGCTACTTGG
1689	PR-Ppdc	GCTTTTTTGTACAAACTTGTGATATCCAATGATTGTGCTGTAGCTGCGCT
1690	Primer1- pyr4	GATGCCCTCGTTCTGTCCACAATTTCTTTCACGAAAAAGGTCTGACTGGG
1691	Primer2- pyr4	ATTACTGAATGAGATACTATCTCGAGATCTTTGCTAAATGCCTTTCTTTA

Continue Table 5.9 Primers and oligonucleotides.

NO.	Name	sequence
1692	AscI plug F	TGCCATCGAAGGAAGCAAAAGGTCAGGCTT
1693	AscI plug R	CTTCGTCCGTCCTCGAAAGATGGAATTCAT
1694	PcDNA1-PF	CGTGAACCATCACCCAAATCAAGTTTTTTATTTCCCCAGACAATGATGGT
1695	PcDNA1-PR	TATGCGTTATGAACATGTTCCCATTTAAATTTGAGAGAAGTTGTTGGATT
1696	TR-Peno-PF	GACCCACTGGGGTTTTAGGAGGTCAATTGTTTTGAAGCTATTTCAGGTGG
1697	TR-Peno-PR	CTGGTAGACGTCATATAATCATATTTAAATTGATTCCGTCCTGGATTGCC
1698	Plug-TReno PF	TCCATCTCGCCGAGAAGTGG
1699	Plug-TReno-PR	TATAGCTACTAGCGACGGTC
1700	TenC-PcDNA-PF	CTTTTTGATCAATCCAACAACTTCTCTCAAATGGCAGCCATCTCTTCCCC
1701	TenC-PcDNA-PR	TTTCATTCTATGCGTTATGAACATGTTCCCTCAGGGCAGCGCCTCCTCTG
1702	Check pdc TF-788 PF	GGTGCCGCTGCCCATTTGAG
1703	Check pdc TF-788 PR	GCGAGTGCGTCGAGGTCCAA
1704	Check illiA PF	CCGAGTTCGAGACGTGGGCT
1705	Check illiA PR	CCACCAGTACCGGCTCCGAT

5.3.3 Plasmid DNA Extraction from *E. coli*

Plasmid DNA was generally extracted from overnight culture of *E. coli* using the NucleoSpin® Plasmid Kit (Machery-Nagel), according to the manufacture's protocol. The plasmid DNA concentration of each isolation was measured using Nanodrop spectrophotometer.

5.3.4 Plasmid DNA Extraction from *S. cerevisiae*

Plasmid DNA was isolated from the 2-3 days culture of *S. cerevisiae* using Zymoprep™ Yeast Plasmid Miniprep II Kit (Zymo Research, USA), according to the manufacture's protocol. All the colonies grown on one SM-URA agar plate after the yeast recombination were collected with a tooth pick and dissolved in 200 µl of solution 1 of the yeast miniprep kit. Then yeast DNA was isolated according to manufacturers' instructions and 10 µl dd.H_2O was used for the final elution. Generally, 3-10 µl of the isolated plasmid was transformed immediately to competent *E. coli* cells.

5.3.5 Fungal Genomic DNA and RNA extraction and conversion into cDNA

Fungal genomic DNA was isolated from about 100 mg of mycelia using the GeneElute™ Plant Genomic DNA Miniprep Kit (Sigma-Aldrich). Fungal total RNA was isolated using TRIzol method (Thermo Fischer Scientific) followed by RNA concentration using RNA clean and concentrator™ kit (Zymo Research). RNA was converted to cDNA using the High-Capacity RNA-to-DNA™ kit (Thermo Fischer Scientific). All the extractions and reactions were done according to the manufacture's protocols.

5.3.6 Polymerase Chain Reaction (PCR)

PCR was used to amplify DNA fragments for both screening and cloning purposes. OneTaq® 2× Master Mix (NEB) was used for screening purposes, while Q5® High–Fidelity 2× Master Mix (NEB) was used for precise fragment amplification for cloning purposes. Both enzymes were used according to the manufacture's protocols.

5.3.7 Cloning Procedure

Restriction digest

All of restriction enzymes were purchased from NEB. The enzymatic reactions were conducted according to restriction digest protocols. The digested plasmids were then evaluated by agarose gel electrophoresis.

Gateway Cloning

Invitrogen Gateway® LR Clonase II enzyme mix was used to perform in vitro recombination between pE-YA entry plasmids and pTYGS expression plasmids. The enzyme was used according to the manufacturer's protocol with reduced reaction scale (reduced to half).

DNA Sequencing

DNA samples were sequenced by Eurofins Genomics (Ebersberg). 2µL of 10µM primer solution was added to templates consisting of at least 15 µg of purified DNA. The mixture was then sent for sequencing.

5.3.8 Agarose Gel Electrophoresis

Agarose gel electrophoresis was used to analyse DNA fragments and Rota safe DNA stain was used to visualize the DNA fragments. Generally, 1% agarose solution was prepared by dissolving 1g Agarose in 100 mL of 0.5 x TAE buffer. DNA samples were mixed with 6 X DNA Loading Dye and loaded on 1% (w/v) agarose gels. 1 kb DNA Ladder (NEB) was used as a molecular DNA size marker. Electrophoresis was carried out in horizontal gel tanks (BioRad) at 90–110 V for 20-35 min using 0.5 xTAE as running buffer. DNA

was visualised under UV (254 nm) and photographed using Gel Doc™ XR+ Gel Documentation System (BioRad).

5.3.9 DNA recovery from PCR

NucleoSpin® Gel and PCR Clean–up Kit was used to directly purify PCR mixture according to the manufacture's protocol. The isolated DNA fragments were then evaluated by agarose gel electrophoresis.

5.3.10 *E. coli* colony PCR

To screen positive colonies of *E. coli*, part of each single colony was picked from the LB agar plate using a toothpick and dipped into 10 µl TE buffer in PCR tube. The cells were then disrupted by heating to 80°C for 15 minutes followed by 55°C for another 15 min using a thermocycler. Then 1 µl of the cell suspension was used for PCR (using OneTaq® 2× Master Mix, NEB), according to the manufacturers protocol.

5.4 Chemical Analysis

5.4.1 Solvents and Chemicals

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

5.4.2 Liquid Chromatography Mass Spectrometry (LCMS)

Analytical

Analytical LCMS data of the organic extracts were obtained using a Waters LCMS system that consisted of a Waters 2767 autosampler, a Waters 2545 pump and Phenomenex Kinetex column (2.6 µm, C_{18}, 100 Å, 4.6 x 100 mm) equipped with a Phenomenex Security Guard precolumn (Luna, C_5, 300 Å). Detection was performed by a diode array detector from 210-600 nm (DAD; Waters 2998 or Waters 996), an electron light scattering detector (ELSD; Waters 2424) and together with a mass spectrometry, Waters SQD-2 mass detector, operating simultaneously in ES^+ and ES^- modes between 150 and 1000 m/z.. Gradient was run over 15 min starting at 10 % acetonitrile/ 90 %

HPLC grade water (+ 0.05 % formic acid) and ramping to 90 % acetonitrile. Flowrate was 1 mL/min and 20-30 µL of the sample were injected. Two different gradients was used for the analysis (Table 5.10), where the standard gradient was used in almost all of the experiments unless stated otherwise. Data were displayed using the software MassLynx.

Table 5.10 Different programs used for analytical LCMS.

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

Preparative

Single compounds were purified from the raw organic extracts by a Waters mass–directed autopurification system. It compromises of a Waters 2767 autosampler, a Waters 2545 Pump and a Phenomenex Kinetex Axia column (5 µm, C_{18}, 100 Å, 21.2 x 250 mm) equipped with a Phenomenex Security Guard precolumn (Luna, C_5, 300 Å). A water/acetonitrile gradient depending on compound polarity was run over 15 minutes with a flowrate of 20 mL/min and a post-column flow split of 100:1. The minority flow was applied for simultaneous analysis by a diode array detector (Waters 2998) in the range 210 to 600 nm, an evaporative light scattering detector (ELSD; Waters 2424) and a Waters SQD-2 mass detector, operating in ES^+ and ES^- modes between 100 and 1000 *m/z*. Selected peaks were collected into test tubes and solvent was evaporated under reduced pressure.

High Resolution Mass Spectrometry (HRMS)

Compound was dissolved in methanol (1 mg/mL). High Resolution Mass Spectrometry was performed on a Q-Tof Premier mass spectrometer (Waters) coupled to an Acquity UPLC–domain (Waters). Electron spray Ionisation (ESI) mass spectroscopy was measured in positive or negative mode depending on the compound.

5.4.3 Nuclear Magnetic Resonance (NMR) analysis

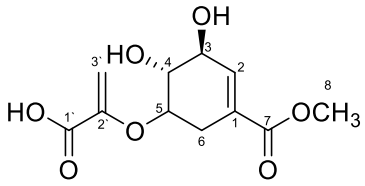
Bruker Ascend 400 MHz, Bruker DRX 500 MHz or a Bruker Ascend 600 MHz Spectrometer (Bruker) were used for NMR measurements of the samples. Raw data were then analyzed using the software Bruker TopSpin 3.5. Chemical shifts are expressed in parts per million (ppm) in comparison to the Tetramethylsilane (TMS) standard and are referenced to the deuterated solvent.

5.4.4 Extraction of Fungal Cultures

For *A. oryzae* and *T. reesei* transformants, a small cell sample was dried by Büchner filtration or gravity filtration and used for gDNA analysis. Mycelia were ground using a hand blender and removed from the culture supernatant by Büchner filtration. Supernatant was acidified with 2 M HCL to pH 3-4 and extracted twice with ethyl acetate. Combined organic layers were dried over $MgSO_4$ and solvents were removed under vacuum. Organic residue was dissolved in methanol or acetonitrile to a concentration of 5-10 mg/mL (analytical) or 50 mg/mL (preparative) and filtered over glass wool.

5.4.5 Isolation of compounds from *T. reesei* transformants

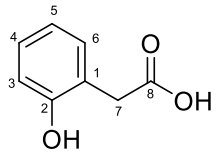
Compounds were isolated from different *T. reesei* transformants using preparative LCMS (15 min gradient: 10 to 90% water–acetonitrile). The NMR data of the isolated compounds is shown in ppm and compared to literature values.



83

8 mg of compound **83** was purified as white powder from the extract of *T. reesei* QM6a Δ*tmus53* P_{pdc1}-*aspks1* colony B. NMR analysis of the compound showed that the compound was not pure and the data contain at least three different isomers of the same compound in addition to other impurities. The compound was predicted to be hydrated scytolide based on NMR data, however, it was not possible to verify the stereochemistry of the compound as the sample was not pure so it was assumed to be as that of scytolide. **HRMS** (ES^+) *m/z* calc. for $C_{11}H_{14}O_7Na$ $[M+H]^+$ 281.0637, found 281.0641.

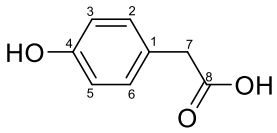
NMR data of compound **83**: ^{1}H NMR (500 MHz, CD_3OD), δ_H: 6.73 (1H, t, *J* = 2 Hz, H-2), 5.46 (1H, d, *J* = 2.5 Hz, Ha-3`), 4.91 (1H, d, *J* = 2.5 Hz, Hb-3`), 4.63 (1H, m, H-3), 4.23 (1H, m, H-5), 4.20 (1H, m, H-4), 3.75 (3H, s, H-8), 3.04 (1H, dd, *J* = 6, 18 Hz, Ha-6), 2.22 (1H, m, Hb-6). ^{13}C-NMR (125 MHz, CD_3OD), δ_C: 167.9 (*C*-7), 167.2 (*C*-1`), 151.4 (*C*-2`), 140.5 (*C*-2), 128.3 (*C*-1), 96.7 (*C*-3`), 79.9 (*C*-3), 77.3 (*C*-4), 72.3 (*C*-5), 52.5 (*C*-8) and 29.3 (*C*-6).



84

2 mg of compound **84** was purified as white powder from the extract of *T. reesei* QM6a Δ*tmus53 P*$_{amyB}$*-aspks1* colony L.

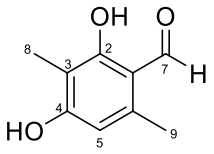
NMR data of compound **84**:[224] ^{1}H NMR (500 MHz, CD_3OD), δ_H: 7.07-7.10 (2H, m, H-3/6), 6.77-6.78 (2H, m, H-4/5) and 3.57 (2H, s, H-7). ^{13}C-NMR (125 MHz, CD_3OD), δ_C: 177.1 (*C*-8), 156.9 (*C*-2), 132.0 (*C*-6), 129.2 (*C*-3), 123.3 (*C*-1), 116.2 and 120.5 (*C*-4/5) and 37.7 (*C*-7).



85

7 mg of compound **85** was purified as white powder from the extract of *T. reesei* QM6a Δ*tmus53 P*$_{amyB}$*-aspks1* colony L.

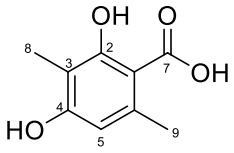
NMR data of compound **85**: ^{1}H NMR (500 MHz, CD_3OD), δ_H: 7.08 (2H, d, *J* = 8.6 Hz, H-2/6), 6.72 (2H, d, *J* = 8.6 Hz, H-3/5) and 3.48 (2H, s, H-7). ^{13}C-NMR (125 MHz, CD_3OD), δ_C: 176.3 (*C*-8), 157.4 (*C*-4), 131.3 (*C*-2/6), 126.9 (*C*-1), 116.2 (*C*-3/5) and 41.2 (*C*-7).



40

1 mg of compound **40** was purified from the extract of *T. reesei* QM6a Δ*tmus53* Δ*sorBC* P_{pdc2}-*aspks1* colony C. **HRMS** (ES^-) *m/z* calc. for $C_9H_9O_3$ $[M+H]^-$ 165.0552, found 165.0549.

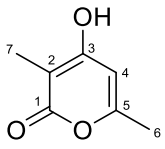
NMR data of compound **40**:[225] ^{1}H-NMR (500 MHz, CD_3OD), δ_H: 10.04 (1H, s, H-7), 6.23 (1H, s, H-5), 2.47 (3H, s, H-9) and 1.98 (3H, s, H-8).



41

15 mg of compound **41** was purified from the extract of *T. reesei* QM6a Δ*tmus53* Δ*sorBC* P_{pdc2}-*aspks1* colony C. **HRMS** (ES^-) *m/z* calc. for $C_9H_9O_4$ $[M+H]^-$ 181.0501, found 181.0502.

NMR data of compound **41**:[226] ^{1}H-NMR (500 MHz, CD_3OD), δ_H: 6.20 (1H, s, H-5), 2.45 (3H, s, H-9) and 2.00 (3H, s, H-8). ^{13}C-NMR (125 MHz, CD_3OD), δ_C: 174.2 (*C*-7), 163.4 (*C*-2), 159.9 (*C*-4), 140.2 (*C*-6), 109.9 (*C*-5), 108.3 (*C*-3), 103.5 (*C*-1), 22.9 (*C*-9) and 6.5 (*C*-8).



79

0.7 mg of compound **79** was purified from the extract of *T. reesei* QM6a Δ*tmus53* Δ*sorBC* P_{pdc2}-*aspks1* colony C.

NMR data of compound **79**: ^{1}H NMR (500 MHz, CD_3OD), δ_H: 5.99 (1H, s, H-4), 2.20 (3H, s, H-6) and 1.85 (3H, s, H-7). ^{13}C-NMR (125 MHz, CD_3OD), δ_C: 169.1 (*C*-1), 167.9 (*C*-3), 161.5 (*C*-5), 101.5 (*C*-4), 98.7 (*C*-2), 19.5 (*C*-6) and 8.2 (*C*-7).

6 Bibliography

1. P. Hunter, *EMBO Rep.*, 2008, **9**, 838–40.
2. F. E. Koehn and G. T. Carter, *Nat. Rev. Drug Discov.*, 2005, **4**, 206–220.
3. W. O. Foye, T. L. Lemke and D. A. Williams, *Foye's principles of medicinal chemistry: Seventh edition*, Wolters Kluwer Health/Lippincott Williams & Wilkins, 2013.
4. B. Shen, *Cell*, 2015, **163**, 1297–300.
5. S. J. Cutler and H. G. Cutler, *Biologically active natural products : pharmaceuticals*, CRC Press, 2000.
6. R. H. Baltz, *J. Ind. Microbiol. Biotechnol.*, 2019, **46**, 281–299.
7. K.-M. Oksman-Caldentey and D. Inzé, *Trends Plant Sci.*, 2004, **9**, 433–440.
8. A. L. Demain and S. Sanchez, *J. Antibiot. (Tokyo).*, 2009, **62**, 5–16.
9. S. W. Drew and A. L. Demain, *Ann. Rev. Microbiol*, 1977, **31**, 343–356.
10. R. A. Maplestone, M. J. Stone and D. H. Williams, *Gene*, 1992, **115**, 151–157.
11. J. C. Nielsen and J. Nielsen, *Synth. Syst. Biotechnol.*, 2017, **2**, 5–12.
12. P. M. Dewick, *Medicinal natural products : a biosynthetic approach*, Wiley, 2002.
13. D. J. Newman, G. M. Cragg and K. M. Snader, *Nat. Prod. Rep.*, 2000, **17**, 215–234.
14. A. L. Harvey, R. Edrada-Ebel and R. J. Quinn, *Nat. Rev. Drug Discov.*, 2015, **14**, 111–129.
15. J. W.-H. Li and J. C. Vederas, *Science*, 2009, **325**, 161–5.
16. A. L. Harvey, *Drug Discov. Today*, 2008, **13**, 894–901.
17. M. J. Rice, M. Legg and K. A. Powell, *Pestic. Sci.*, 1998, **52**, 184–188.
18. R. Nofiani, K. de Mattos-Shipley, K. E. Lebe, L.-C. Han, Z. Iqbal, A. M. Bailey, C. L. Willis, T. J. Simpson and R. J. Cox, *Nat. Commun.*, 2018, **9**, 3940.
19. G. Strobel, B. Daisy, U. Castillo and J. Harper, *J. Nat. Prod.*, 2004, **67**, 257–268.
20. S. Sanchez and A. L. Demain, *Microb. Biotechnol.*, 2008, **1**, 283–319.
21. F. Alberti, G. D. Foster and A. M. Bailey, *Appl. Microbiol. Biotechnol.*, 2017, **101**, 493–500.
22. L. Bohlin, U. Göransson, C. Alsmark, C. Wedén and A. Backlund, *Phytochem. Rev.*, 2010, **9**, 279–301.
23. E. Skellam, *Trends Biotechnol.*, 2019, **37**, 416–427.
24. T. Maier, M. Leibundgut, D. Boehringer and N. Ban, *Q. Rev. Biophys.*, 2010, **43**, 373–422.
25. S. J. Wakil, J. K. Stoops and V. C. Joshi, *Annu. Rev. Biochem.*, 1983, **52**, 537–

579.

26. S. Smith and S. C. Tsai, *Nat. Prod. Rep.*, 2007, **24**, 1041–1072.
27. J. M. Crawford and C. A. Townsend, *Nat. Rev. Microbiol.*, 2010, **8**, 879–89.
28. S. Smith, A. Witkowski and A. K. Joshi, *Prog. Lipid Res.*, 2003, **42**, 289–317.
29. J. Beld, D. J. Lee and M. D. Burkart, *Mol. Biosyst.*, 2015, **11**, 38–59.
30. S. W. White, J. Zheng, Y.-M. Zhang and C. O. Rock, *Annu. Rev. Biochem.*, 2005, **74**, 791–831.
31. T. Maier, M. Leibundgut and N. Ban, *Science*, 2008, **321**, 1315–22.
32. D. A. Herbst, C. A. Townsend and T. Maier, *Nat. Prod. Rep.*, 2018, **35**, 1046–1069.
33. R. J. Cox, *Org. Biomol. Chem.*, 2007, **5**, 2010–2026.
34. J. Staunton and K. J. Weissman, *Nat. Prod. Rep.*, 2001, **18**, 380–416.
35. M. L. Vickery and B. Vickery, in *Secondary Plant Metabolism*, Macmillan Education UK, London, 1981, pp. 56–87.
36. J. Staunton and K. J. Weissman, *Nat. Prod. Rep.*, 2001, **18**, 380–416.
37. R. J. Cox and T. J. Simpson, *Methods Enzymol.*, 2009, **459**, 49–78.
38. I. Fujii, *J. Antibiot. (Tokyo).*, 2010, **63**, 207–218.
39. B. D. Ames, C. Nguyen, J. Bruegger, P. Smith, W. Xu, S. Ma, E. Wong, S. Wong, X. Xie, J. W.-H. Li, J. C. Vederas, Y. Tang and S.-C. Tsai, *Proc. Natl. Acad. Sci.*, 2012, **109**, 11144–11149.
40. D. J. McGarvey and R. Croteau, *Plant Cell*, 1995, **7**, 1015.
41. D. Cox-Georgian, N. Ramadoss, C. Dona and C. Basu, in *Medicinal Plants: From Farm to Pharmacy*, Springer, Cham, 2019, pp. 333–359.
42. Margaret F. Roberts. and Michael Wink, *Alkaloids: biochemistry, ecology, and medicinal applications.*, Plenum Press, 1998.
43. T. Dang and R. D. Süssmuth, *Acc. Chem. Res.*, 2017, **50**, 1566–1576.
44. M. L. Shenouda and R. J. Cox, *RSC Adv.*, 2021, **11**, 3622–3635.
45. J. W. Lee, D. Na, J. M. Park, J. Lee, S. Choi and S. Y. Lee, *Nat. Chem. Biol.*, 2012, **8**, 536–546.
46. L. B. Pickens, Y. Tang and Y.-H. Chooi, *Annu. Rev. Chem. Biomol. Eng.*, 2011, **2**, 211–36.
47. K. Tamano, *Front. Microbiol.*, 2014, **5**, 718.
48. E. J. Steen, Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McClure, S. B. del Cardayre and J. D. Keasling, *Nature*, 2010, **463**, 559–562.
49. Y. Zhang, I. P. Adams and C. Ratledge, *Microbiology*, 2007, **153**, 2013–2025.
50. R. J. Cox and K. Williams, in *Natural Products: Discourse, Diversity, and Design*, 2014, pp. 245–260.
51. D. C. Anyaogu and U. H. Mortensen, *Front. Microbiol.*, 2015, **6**, 77.

52. D. H. Scharf and A. A. Brakhage, *J. Biotechnol.*, 2013, **163**, 179–183.

53. B. Wang, F. Guo, S.-H. Dong and H. Zhao, *Nat. Chem. Biol.*, 2019, **15**, 111–114.

54. J. Schümann and C. Hertweck, *J. Am. Chem. Soc.*, 2007, **129**, 9564–9565.

55. C. Wang, C. Lambert, M. Hauser, A. Deuschmann, C. Zeilinger, K. Rottner, T. E. B. Stradal, M. Stadler, E. J. Skellam and R. J. Cox, *Chem. – A Eur. J.*, 2020, **26**, 13578–13583.

56. D. A. Hopwood, F. Malpartida, H. M. Kieser, H. Ikeda, J. Duncan, I. Fujii, B. A. M. Rudd, H. G. Floss and S. Ōmura, *Nature*, 1985, **314**, 642–644.

57. Xiao-Long Yang, Steffen Friedrich, Sen Yin, Oliver Piech, Katherine Williams, T. J. Simpson and R. J. Cox, *Chem. Sci.*, 2019, **10**, 8478–8489.

58. K. a. K. Pahirulzaman, K. Williams and C. M. Lazarus, *A toolkit for heterologous expression of metabolic pathways in aspergillus oryzae*, Elsevier Inc., 1st edn., 2012, vol. 517.

59. C. P. Kubicek, M. Komon-Zelazowska and I. S. Druzhinina, *J. Zhejiang Univ. Sci. B*, 2008, **9**, 753–763.

60. V. K. Gupta, *Biotechnology and biology of trichoderma*, Elsevier, 2014.

61. R. A. A. Khan, S. Najeeb, S. Hussain, B. Xie and Y. Li, *Microorganisms*, 2020, **8**, 817.

62. M. Sood, D. Kapoor, V. Kumar, M. S. Sheteiwy, M. Ramakrishnan, M. Landi, F. Araniti and A. Sharma, *Plants*, 2020, **9**, 762.

63. F. Vinale, K. Sivasithamparam, E. L. Ghisalberti, M. Ruocco, S. Woo and M. Lorito, *Nat. Prod. Commun.*, 2012, **7**, 1934578X1200701133.

64. S. Djonović, G. Vittone, A. Mendoza-Herrera and C. Kenerley, *Mol. Plant Pathol.*, 2007, **8**, 469–480.

65. E. Sharon, I. Chet and Y. Spiegel, in *Biological Control of Plant-Parasitic Nematodes:*, Springer, Dordrecht, 2011, vol. 11, pp. 183–201.

66. Q. Migheli, L. González-Candelas, L. Dealessi, A. Camponogara and D. Ramón-Vidal, *Phytopathology*, 1998, **88**, 673–677.

67. S. L. Woo, B. Donzelli, F. Scala, R. Mach, G. E. Harman, C. P. Kubicek, G. Del Sorbo and M. Lorito, *Mol. Plant-Microbe Interact.*, 1999, **12**, 419–429.

68. L. K. T. Al-Ani, in *Secondary Metabolites of Plant Growth Promoting Rhizomicroorganisms*, Springer Singapore, 2019, pp. 125–143.

69. C. Keswani, K. Bisen, M. K. Chitara, B. K. Sarma and H. B. Singh, in *Agro-Environmental Sustainability*, Springer International Publishing, 2017, vol. 1, pp. 63–79.

70. N. Sahebani and N. Hadavi, *Soil Biol. Biochem.*, 2008, **40**, 2016–2020.

71. M. Chen, Q. Liu, S. S. Gao, A. E. Young, S. E. Jacobsen and Y. Tang, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 5499–5504.

72. R. Doppelbauer, H. Esterbauer, W. Steiner, R. M. Lafferty and H. Steinmüller, *Appl. Microbiol. Biotechnol.*, 1987, **26**, 485–494.

73. S. Silva, B. B. Elmore and H. K. Huckabay, *Appl. Biochem. Biotechnol.*, 1995,

51, 145–153.

74. Y. Zhu, J. Wang, P. Mou, Y. Yan, M. Chen and Y. Tang, *Org. Biomol. Chem.*, 2021, **19**, 1985–1990.

75. I. G. U. Pait, S. Kitani, F. W. Roslan, D. Ulanova, M. Arai, H. Ikeda and T. Nihira, *J. Ind. Microbiol. Biotechnol.*, 2018, **45**, 77–87.

76. C. M. Lazarus, K. Williams and A. M. Bailey, *Nat. Prod. Rep.*, 2014, **31**, 1339–1347.

77. D. Lubertozzi and J. D. Keasling, *Biotechnol. Adv.*, 2009, **27**, 53–75.

78. Z. Song, W. Bakeer, J. W. Marshall, A. A. Yakasai, R. M. Khalid, J. Collemare, E. Skellam, D. Tharreau, M. H. Lebrun, C. M. Lazarus, A. M. Bailey, T. J. Simpson and R. J. Cox, *Chem. Sci.*, 2015, **6**, 4837–4845.

79. K. M. Fisch, E. Skellam, D. Ivison, R. J. Cox, A. M. Bailey, C. M. Lazarus and T. J. Simpson, *Chem. Commun.*, 2010, **46**, 5331–5333.

80. Verena Hantke, E. J. Skellam and R. J. Cox, *Chem. Commun.*, 2020, **56**, 2925–2928.

81. R. Fujii, T. Ugai, H. Ichinose, M. Hatakeyama, T. Kosaki, K. Gomi, I. Fujii, A. Minami and H. Oikawa, *Biosci. Biotechnol. Biochem.*, 2016, **80**, 426–431.

82. T. Ugai, A. Minami, K. Gomi and H. Oikawa, *Tetrahedron Lett.*, 2016, **57**, 2793–2796.

83. P. Xiang, L. Ludwig-Radtke, W. B. Yin and S. M. Li, *Org. Biomol. Chem.*, 2020, **18**, 4946–4948.

84. L. Hang, M.-C. Tang, C. J. B. Harvey, C. G. Page, J. Li, Y.-S. Hung, N. Liu, M. E. Hillenmeyer and Y. Tang, *Angew. Chemie*, 2017, **129**, 9684–9688.

85. U. Galm and B. Shen, *Expert Opin. Drug Discov.*, 2006, **1**, 409–437.

86. C. Pohl, F. Polli, T. Schütze, A. Viggiano, L. Mózsik, S. Jung, M. de Vries, R. A. L. Bovenberg, V. Meyer and A. J. M. Driessen, *Sci. Rep.*, 2020, **10**, 1–16.

87. Y. M. Chiang, C. E. Oakley, M. Ahuja, R. Entwistle, A. Schultz, S. L. Chang, C. T. Sung, C. C. C. Wang and B. R. Oakley, *J. Am. Chem. Soc.*, 2013, **135**, 7720–7731.

88. I. Kjærbølling, U. H. Mortensen, T. Vesth and M. R. Andersen, *Fungal Genet. Biol.*, 2019, 130, 107–121.

89. X. Lin, S. Yuan, S. Chen, B. Chen, H. Xu, L. Liu, H. Li and Z. Gao, *Molecules*, 2019, **24**, 2267.

90. Y. He and R. J. Cox, *Chem. Sci.*, 2016, **7**, 2119–2127.

91. P. J. Punt and C. A. M. J. J. van den Hondel, *Methods Enzymol.*, 1992, **216**, 447–457.

92. L. Gritz and J. Davies, *Gene*, 1983, **25**, 179–188.

93. D. Drocourt, T. Calmels, J.-P. Reynes, M. Baron and G. Tiraby, *Nucleic Acids Res.*, 1990, **18**, 4009–4009.

94. X. Su, G. Schmitz, M. Zhang, R. I. Mackie and I. K. O. Cann, *Adv. Appl. Microbiol.*, 2012, **81**, 1–61.

95. K. Gomi, Y. Iimura and S. Hara, *Agric. Biol. Chem.*, 1987, **51**, 2549–2555.

96. S. E. Unkles, E. I. Campbell, Y. M. J. T. de Ruiter-Jacobs, M. Broekhuijsen, J. A. Macro, D. Carrez, R. Contreras, C. A. M. J. J. van den Hondel and J. R. Kinghorn, *Mol. Gen. Genet. MGG*, 1989, **218**, 99–104.

97. T. Fujii, H. Yamaoka, K. Gomi, K. Kitamoto and C. Kumaga, *Biotechnol. Biochem.*, 1995, **59**, 1869–1874.

98. I. W. Nilsen, I. Bakke, A. Vader, Ø. Olsvik and M. R. El-Gewely, *J. Bacteriol.*, 1996, **178**, 3188–3193.

99. A. M. Myers, A. Tzagoloff, D. M. Kinney and C. J. Lusty, *Gene*, 1986, **45**, 299–310.

100. C. J. Thompson, N. R. Movva, R. Tizard, R. Crameri, J. E. Davies, M. Lauwereys and J. Botterman, *EMBO J.*, 1987, **6**, 2519–2523.

101. L. M. Halo, J. W. Marshall, A. A. Yakasai, Z. Song, C. P. Butts, M. P. Crump, M. Heneghan, A. M. Bailey, T. J. Simpson, C. M. Lazarus and R. J. Cox, *ChemBioChem*, 2008, **9**, 585–594.

102. M. N. Heneghan, A. a. Yakasai, L. M. Halo, Z. Song, A. M. Bailey, T. J. Simpson, R. J. Cox and C. M. Lazarus, *ChemBioChem*, 2010, **11**, 1508–1512.

103. K. A. K. Pahirulzaman, K. Williams and C. M. Lazarus, *Methods Enzymol.*, 2012, **517**, 241–260.

104. F. J. Jin, J. Maruyama, P. R. Juvvadi, M. Arioka and K. Kitamoto, *FEMS Microbiol. Lett.*, 2004, **239**, 79–85.

105. A. V. Gusakov, *Trends Biotechnol.*, 2011, **29**, 419–425.

106. M. Penttilä, H. Nevalainen, M. Rättö, E. Salminen and J. Knowles, *Gene*, 1987, **61**, 155–164.

107. M. Saloheimo, T. Pakula, N. Aro and J. J. Joensuu, in *Biotechnology and Biology of Trichoderma*, Elsevier, 2014, pp. 81–88.

108. A. Schuster and M. Schmoll, *Appl. Microbiol. Biotechnol.*, 2010, **87**, 787–99.

109. Reese and E.T., *Biotechnol. Bioeng. Symp.*, 1976, **6**, 9–20.

110. T. Nakari-Setälä, M. Paloheimo, J. Kallio, J. Vehmaanperä, M. Penttilä and M. Saloheimo, *Appl. Environ. Microbiol.*, 2009, **75**, 4853–4860.

111. H. Nevalainen, P. Suominen and K. Taimisto, *J. Biotechnol.*, 1994, **37**, 193–200.

112. R. Peterson and H. Nevalainen, *Microbiology*, 2012, **158**, 58–68.

113. S. Ellilä, L. Fonseca, C. Uchima, J. Cota, G. H. Goldman, M. Saloheimo, V. Sacon and M. Siika-aho, *Biotechnol. Biofuels*, 2017, **10**, 1–17.

114. E. T. Reese and H. S. Levinson, *Physiol. Plant.*, 1952, **5**, 345–366.

115. M. Mandels, J. M. Nystrom and R. K. Andren, *Science*, 1974, **184**, 524.

116. J. Li, J. Wang, S. Wang, M. Xing, S. Yu and G. Liu, *Microb. Cell Fact.*, 2012, **11**, 84.

117. V. Sewalt, D. Shanahan, L. Gregg, J. La Marta and R. Carrillo, *Ind. Biotechnol.*, 2016, **12**, 295–302.

118. J. C. Frisvad, L. L. H. Møller, T. O. Larsen, R. Kumar and J. Arnau, *Appl. Microbiol. Biotechnol.*, 2018, **102**, 9481–9515.

119. J. Kluge, D. Terfehr and U. Kück, *Appl. Microbiol. Biotechnol.*, 2018, **102**, 6357–6372.

120. C. Derntl, D. P. Kiesenhofer, R. L. Mach and A. R. Mach-Aigner, *Appl. Environ. Microbiol.*, 2015, **81**, 6314–6323.

121. M. S. Jørgensen, D. A. Skovlund, P. F. Johannesen and U. H. Mortensen, *Microb. Cell Fact.*, 2014, **13**, 1–9.

122. M. G. Steiger, M. Vitikainen, P. Uskonen, K. Brunner, G. Adam, T. Pakula, M. Penttilä, M. Saloheimo, R. L. MacH and A. R. Mach-Aigner, *Appl. Environ. Microbiol.*, 2011, **77**, 114–121.

123. L. Hartl and B. Seiboth, *Curr. Genet.*, 2005, **48**, 204–211.

124. C. K. Peterbauer, E. Heidenreich, R. T. Baker and C. P. Kubicek, *Can. J. Microbiol.*, 1992, **38**, 1292–1297.

125. A. Singh, L. E. Taylor, T. A. Vander Wall, J. Linger, M. E. Himmel, K. Podkaminer, W. S. Adney and S. R. Decker, *Biotechnol. Adv.*, 2015, **33**, 142–154.

126. A. Gatignol, H. Durand and G. Tiraby, *FEBS Lett.*, 1988, **230**, 171–175.

127. T. Kubodera, N. Yamashita and A. Nishimura, *Biosci. Biotechnol. Biochem.*, 2002, **66**, 404–406.

128. T. Bergès, C. Barreau, J. F. Peberdy and L. M. Boddy, *Curr. Genet.*, 1993, **24**, 53–59.

129. F. Gruber, J. Visser, C. P. Kubicek and L. H. de Graaff, *Curr. Genet.*, 1990, **18**, 71–76.

130. E. Fitz, F. Wanka and B. Seiboth, *Front. Bioeng. Biotechnol.*, 2018, **6**, 135.

131. F. Uzbas, U. Sezerman, L. Hartl, C. P. Kubicek and B. Seiboth, *Appl. Microbiol. Biotechnol.*, 2012, **93**, 1601–1608.

132. W. Kurzatkowski, A. Törrönen, J. Filipek, R. L. Mach, P. Herzog, S. Sowka and C. P. Kubicek, *Appl. Environ. Microbiol.*, 1996, **62**, 2859–2865.

133. Z. Rahman, Y. Shida, T. Furukawa, Y. Suzuki, H. Okada, W. Ogasawara and Y. Morikawa, *Appl. Microbiol. Biotechnol.*, 2009, **82**, 899–908.

134. S. Miyauchi, V. S. Te'o, P. L. Bergquist and K. M. H. Nevalainen, *N. Biotechnol.*, 2013, **30**, 523–530.

135. R. Liu, L. Chen, Y. Jiang, Z. Zhou and G. Zou, *Cell Discov.*, 2015, **1**, 1–11.

136. P. Liu, W. Wang and D. Wei, *J. Ind. Microbiol. Biotechnol.*, 2017, **44**, 1367–1373.

137. S. Nanda and F. Berruti, *Environ. Chem. Lett.*, 2020, **19**, 1433–1456.

138. M. Garedew, F. Lin, B. Song, T. M. DeWinter, J. E. Jackson, C. M. Saffron, C. H. Lam and P. T. Anastas, *ChemSusChem*, 2020, **13**, 4214–4237.

139. S. Maina, V. Kachrimanidou and A. Koutinas, *Curr. Opin. Green Sustain. Chem.*, 2017, **8**, 18–23.

140. G. Hamer, *Biotechnol. Adv.*, 2003, **22**, 71–79.

141. R. A. D. Arancon, C. S. K. Lin, K. M. Chan, T. H. Kwan and R. Luque, *Energy Sci. Eng.*, 2013, **1**, 53–71.

142. O. Rosales-Calderon and V. Arantes, *Biotechnol. Biofuels*, 2019, **12**, 240.

143. M. Mandels, L. Hontz, J. Nystrom and I. by Lee R. Lynd, *Biotechnol. Bioeng.*, 2010, **105**, 1–25.

144. F. Gao, Z. Hao, X. Sun, L. Qin, T. Zhao, W. Liu, H. Luo, B. Yao and X. Su, *Biotechnol. Biofuels 2018 111*, 2018, **11**, 1–13.

145. R. Weiss, A. Eischer, T. Tadic, S. M. Gritsch, M. Ortner, K. Prall, E. Neunteufel, R. F. Putz, G. M. Guebitz and G. S. Nyanhongo, *J. Clean. Prod.*, 2020, **275**, 122882.

146. M. Mandels, J. Weber and R. Parizek, *Enhanced Cellulase Production by a Mutant of Trichoderma viride*, 1971, vol. 21.

147. M. Mandels and E. T. Reese, *J. Bacteriol.*, 1957, **73**, 269–278.

148. M. Mandels and D. E. Eveleigh, *Biotechnol. Biofuels*, 2009, **2**, 20.

149. I. S. Druzhinina and C. P. Kubicek, *Microb. Biotechnol.*, 2017, **10**, 1485–1499.

150. H. Wang, G. Kaur, N. Pensupa, K. Uisan, C. Du, X. Yang and C. S. K. Lin, *Process Saf. Environ. Prot.*, 2018, **118**, 143–151.

151. C. Derntl, A. Rassinger, E. Srebotnik, R. L. Mach and A. R. Mach-Aigner, *Appl. Environ. Microbiol.*, 2016, **82**, 6247–6257.

152. L. Kahlert, E. F. Bassiony, R. J. Cox and E. J. Skellam, *Angew. Chemie*, 2020, **132**, 5865–5871.

153. C. Derntl, F. Guzmán-Chávez, T. M. Mello-de-Sousa, H. J. Busse, A. J. M. Driessen, R. L. Mach and A. R. Mach-Aigner, *Front. Microbiol.*, 2017, **8**, 1–12.

154. J. K. Saini, R. Saini and L. Tewari, *3 Biotech*, 2015, **5**, 337–353.

155. R. Schor, C. Schotte, D. Wibberg, J. Kalinowski and R. J. Cox, *Nat. Commun.*, 2018, **9**, 1–9.

156. S. Dockrey, A. Baker, C. E. Suh, A. R. Benítez, T. Wymore, C. L. Brooks III and A. R. H. Narayan, *ACS Cent. Sci.*, 2019, **5**, 1010–1016.

157. S. A. Baker Dockrey, A. L. Lukowski, M. R. Becker and A. R. H. Narayan, *Nat. Chem.*, 2017, **10**, 119–125.

158. A. Sib and T. A. M. Gulder, *Angew. Chemie Int. Ed.*, 2018, **57**, 14650–14653.

159. F. Fang, J. Zhao, L. Ding, C. Huang, C. B. Naman, S. He, B. Wu, P. Zhu, Q. Luo, W. H. Gerwick, X. Yan, Q. Wang, Z. Zhang and W. Cui, *Mar. Drugs*, 2017, **15**, 260.

160. M. Liu, M. Ohashi, Y.-S. Hung, K. Scherlach, K. Watanabe, C. Hertweck and Y. Tang, *J. Am. Chem. Soc.*, 2021, **143**, 7267–7271.

161. K. F. Nielsen, M. Månsson, C. Rank, J. C. Frisvad and T. O. Larsen, *J. Nat. Prod.*, 2011, **74**, 2338–2348.

162. A. M. Bailey, R. J. Cox, K. Harley, C. M. Lazarus, T. J. Simpson and E. Skellam,

Chem. Commun., 2007, 4053–4055.

163. Z. Shahryari, M. H. Fazaelipoor, P. Setoodeh, R. B. Nair, M. J. Taherzadeh and Y. Ghasemi, *Int. J. Recycl. Org. Waste Agric.*, 2018, **7**, 345–355.

164. J. J. Laukevics, A. F. Apsite, U. E. Viesturs and R. P. Tengerdy, *Biotechnol. Bioeng.*, 1984, **26**, 1465–1474.

165. D. Martinez, R. M. Berka, B. Henrissat, M. Saloheimo, M. Arvas, S. E. Baker, J. Chapman, O. Chertkov, P. M. Coutinho, D. Cullen, E. G. J. Danchin, I. V Grigoriev, P. Harris, M. Jackson, C. P. Kubicek, C. S. Han, I. Ho, L. F. Larrondo, A. L. de Leon, J. K. Magnuson, S. Merino, M. Misra, B. Nelson, N. Putnam, B. Robbertse, A. A. Salamov, M. Schmoll, A. Terry, N. Thayer, A. Westerholm-Parvinen, C. L. Schoch, J. Yao, R. Barabote, M. A. Nelson, C. Detter, D. Bruce, C. R. Kuske, G. Xie, P. Richardson, D. S. Rokhsar, S. M. Lucas, E. M. Rubin, N. Dunn-Coleman, M. Ward and T. S. Brettin, *Nat. Biotechnol.*, 2008, **26**, 553–560.

166. A. O. Noor, D. M. Almasri, A. A. Bagalagel, H. M. Abdallah, S. G. A. Mohamed, G. A. Mohamed and S. R. M. Ibrahim, *Molecules*, 2020, **25**, 395.

167. M. Ahuja, Y.-M. Chiang, S.-L. Chang, M. B. Praseuth, R. Entwistle, J. F. Sanchez, H.-C. Lo, H.-H. Yeh, B. R. Oakley and C. C. C. Wang, *J. Am. Chem. Soc.*, 2012, **134**, 8212–8221.

168. C. Derntl, B. Kluger, C. Bueschl, R. Schuhmacher, R. L. Mach and A. R. Mach-Aigner, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, E560–E569.

169. C. Derntl, A. Rassinger, E. Srebotnik, R. L. Mach and A. R. Mach-Aigner, *Biotechnol. Biofuels*, 2015, **8**, 1–11.

170. M. J. Velázquez-Nuñez, R. Avila-Sosa, E. Palou and A. López-Malo, *Food Control*, 2013, **31**, 1–4.

171. D. R. L. Caccioni, M. Guizzardi, D. M. Biondi, A. Renda and G. Ruberto, *Int. J. Food Microbiol.*, 1998, **43**, 73–79.

172. V. Novy, F. Nielsen, B. Seiboth and B. Nidetzky, *Biotechnol. Biofuels*, 2019, **12**, 1–16.

173. V. K. Gupta, A. S. Steindorff, R. G. de Paula, R. Silva-Rocha, A. R. Mach-Aigner, R. L. Mach and R. N. Silva, *Trends Biotechnol.*, 2016, **34**, 970–982.

174. N. Peitersen, *Biotechnol. Bioeng.*, 1975, **17**, 1291–1299.

175. T. Welton, *Proc. R. Soc. A Math. Phys. Eng. Sci.*, 2015, **471**, 20150502.

176. D. Prat, J. Hayler and A. Wells, *Green Chem.*, 2014, **16**, 4546–4551.

177. T. T. Lai, T. T. H. Pham, K. Adjallé, D. Montplaisir, F. Brouillette and S. Barnabé, *Waste and Biomass Valorization*, 2016, **8**, 1081–1088.

178. V. Mommaerts, G. Platteau, J. Boulet, G. Sterk and G. Smagghe, *Biol. Control*, 2008, **46**, 463–466.

179. M. Isaka, W. Prathumpai, P. Wongsa and M. Tanticharoen, *Org. Lett.*, 2006, **8**, 2815–2817.

180. B. Wu, V. Oesker, J. Wiese, R. Schmaljohann and J. F. Imhoff, *Mar. Drugs*, 2014, **12**, 1208–1219.

181. E. C. Marfori, S. Kajiyama, E. I. Fukusaki and A. Kobayashi, *Zeitschrift fur*

Naturforsch. - Sect. C J. Biosci., 2002, **57**, 465–470.

182. J. M. Dickinson, J. R. Hanson, P. B. Hitchcock and N. Claydon, *J. Chem. Soc. Perkin Trans. 1*, 1989, 1885–1887.

183. G. Ding, H. Wang, L. Li, B. Song, H. Chen, H. Zhang, X. Liu and Z. Zou, *J. Nat. Prod.*, 2014, **77**, 164–167.

184. H. Chen, G. Daletos, F. Okoye, D. Lai, H. Dai and P. Proksch, *Nat. Prod. Commun.*, 2015, **10**, 1934578X1501000412.

185. U. Bat-Erdene, D. Kanayama, D. Tan, W. C. Turner, K. N. Houk, M. Ohashi and Y. Tang, *J. Am. Chem. Soc.*, 2020, **142**, 8550–8554.

186. L. Xie, X. Zang, W. Cheng, Z. Zhang, J. Zhou, M. Chen and Y. Tang, *J. Am. Chem. Soc.*, 2021, **143**, 9575–9584.

187. T. Damodaran, S. Rajan, M. Manoharan, Ram Gopal, K. Yadav, S. Kumar, I. Ahmad, N. Kumari, V. K. Mishra and S. K. Jha, *Front. Microbiol.*, 2020, **11**, 2900.

188. W. Hinterdobler, S. Beier, A. A. Monroy, H. Berger, C. Dattenböck and M. Schmoll, *Front. Bioeng. Biotechnol.*, 2020, **8**, 1293.

189. E. Hitzenhammer, C. Büschl, M. Sulyok, R. Schuhmacher, B. Kluger, E. Wischnitzki and M. Schmoll, *BMC Genomics*, 2019, **20**, 211.

190. A. A. Monroy, E. Stappler, A. Schuster, M. Sulyok and M. Schmoll, *PLoS One*, 2017, **12**, 1–20.

191. S. Beier, W. Hinterdobler, A. A. Monroy, H. Bazafkan and M. Schmoll, *Front. Microbiol.*, 2020, **11**, 974.

192. M. S. Jørgensen, T. O. Larsen, U. H. Mortensen and D. Aubert, Kgs. Lyngby: Technical University of Denmark, 2013.

193. C. Derntl, F. Guzmán-Chávez, T. M. Mello-de-Sousa, H.-J. Busse, A. J. M. Driessen, R. L. Mach and A. R. Mach-Aigner, *Front. Microbiol.*, 2017, **8**, 2037.

194. L. Atanasova, B. P. Knox, C. P. Kubicek, I. S. Druzhinina and S. E. Baker, *Eukaryot. Cell*, 2013, **12**, 1499–1508.

195. C. Li, F. Lin, W. Sun, S. Yuan, Z. Zhou, F.-G. Wu and Z. Chen, *Biotechnol. Biofuels*, 2018, **11**, 291.

196. J. Meng, X. Wang, D. Xu, X. Fu, X. Zhang, D. Lai, L. Zhou, G. Zhang, J. Meng, X. Wang, D. Xu, X. Fu, X. Zhang, D. Lai, L. Zhou and G. Zhang, *Molecules*, 2016, **21**, 715.

197. O. Salo, F. Guzmán-Chávez, M. I. Ries, P. P. Lankhorst, R. A. L. Bovenberg, R. J. Vreeken and A. J. M. Driessen, *Appl. Environ. Microbiol.*, 2016, **82**, 3971–3978.

198. A. M. Harned and K. A. Volp, *Nat. Prod. Rep.*, 2011, **28**, 1790.

199. A. Al Fahad, A. Abood, K. M. Fisch, A. Osipow, J. Davison, M. Avramović, C. P. Butts, J. Piel, T. J. Simpson and R. J. Cox, *Chem. Sci.*, 2014, **5**, 523–527.

200. N. Abe, T. Arakawa, K. Yamatoto and A. Hirota, *Biosci. Biotechnol. Biochem.*, 2002, **66**, 2090–2099.

201. S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman, *Nucleic Acids Res.*, 1997, **25**, 3389–3402.

202. E. Kuhnert, J. C. Navarro-Muñoz, K. Becker, M. Stadler, J. Collemare and R. J. Cox, *Stud. Mycol.*, 2021, **99**, 100118.

203. B. SE, K. S, I. P, A. T, L. BY, S. L, Y. OC and T. BG, *Mol. Plant. Microbe. Interact.*, 2006, **19**, 139–149.

204. D. W. Brown, R. A. E. Butchko, M. Busman and R. H. Proctor, *Fungal Genet. Biol.*, 2012, **49**, 521–532.

205. I. Fujii, N. Yoshida, S. Shimomaki, H. Oikawa and Y. Ebizuka, *Chem. Biol.*, 2005, **12**, 1301–1309.

206. X. Wang, Y. Le Gao, M. L. Zhang, H. D. Zhang, J. Z. Huang and L. Li, *J. Biotechnol.*, 2020, **309**, 85–91.

207. T. Bai, Z. Quan, R. Zhai, T. Awakawa, Y. Matsuda and I. Abe, *Org. Lett.*, 2018, **20**, 7504–7508.

208. A. Gallo, B. Knox, K. Bruno, M. Solfrizzo, S. Baker and G. Perrone, *Int. J. Food Microbiol.*, 2014, **179**, 10–17.

209. S. Wang, Y. Xu, E. Maine, E. Wijeratne, P. Espinosa-Artiles, A. Gunatilaka and I. Molnár, *Chem. Biol.*, 2008, **15**, 1328–1338.

210. L. Liu, M. C. Tang and Y. Tang, *J. Am. Chem. Soc.*, 2020, **141**, 19538–19541.

211. G. Zeng, P. Zhang, Q. Zhang, H. Zhao, Z. Li, X. Zhang, C. Wang, W. Yin and W. Fang, *PLoS Genet.*, 2018, **14**, e1007472.

212. Y. Araki, T. Awakawa, M. Matsuzaki, R. Cho, Y. Matsuda, S. Hoshino, Y. Shinohara, M. Yamamoto, Y. Kido, D. Inaoka, K. Nagamune, K. Ito, I. Abe and K. Kita, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 8269–8274.

213. T. A. Rush, H. K. Shrestha, M. Gopalakrishnan Meena, M. K. Spangler, J. C. Ellis, J. L. Labbé and P. E. Abraham, *Front. Fungal Biol.*, 2021, 41.

214. Z. Zhang, C. S. Jamieson, Y. L. Zhao, D. Li, M. Ohashi, K. N. Houk and Y. Tang, *J. Am. Chem. Soc.*, 2019, **141**, 5659–5663.

215. S. Hayakawa, H. Minato and K. Katagiri, *J. Antibiot. (Tokyo).*, 1971, **24**, 653–654.

216. S. B. Singh, W. Liu, X. Li, T. Chen, A. Shafiee, D. Card, G. Abruzzo, A. Flattery, C. Gill, J. R. Thompson, M. Rosenbach, S. Dreikorn, V. Hornak, M. Meinz, M. Kurtz, R. Kelly and J. C. Onishi, *ACS Med. Chem. Lett.*, 2012, **3**, 814.

217. K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema and T. Weber, *Nucleic Acids Res.*, 2019, **47**, W81–W87.

218. T. Carver, K. Rutherford, M. Berriman, M. Rajandream, B. Barrell and J. Parkhill, *Bioinformatics*, 2005, **21**, 3422–3423.

219. C. L. M. Gilchrist, T. J. Booth, B. van Wersch, L. van Grieken, M. H. Medema and Y.-H. Chooi, *Bioinforma. Adv.*, 2021, **1**, vbab016.

220. P. Primerano, M. Juric, R. Mach, A. Mach-Aigner and C. Derntl, *Fungal Biol. Biotechnol. 2021 81*, 2021, **8**, 1–9.

221. W. L. Marsden, P. P. Gray and M. Mandels, *Crit. Rev. Biotechnol.*, 1985, **3**, 235–276.

222. Packaging waste statistics - Statistics Explained, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Packaging_waste_statistics, (accessed 12 October 2021).

223. R. D. Gietz and R. H. Schiestl, *Nat. Protoc.*, 2007, **2**, 31–34.

224. D. Yang and H. Fu, *Chem. – A Eur. J.*, 2010, **16**, 2366–2370.

225. S. A. B. Dockrey and A. R. H. Narayan, *Org. Lett.*, 2020, **22**, 3712–3716.

226. Z. Shang, Z. Khalil, L. Li, A. A. Salim, M. Quezada, P. Kalansuriya and R. J. Capon, *Org. Lett.*, 2016, **18**, 4340–4343.

7 Appendix

7.1 Chapter 2 additional data

7.1.1 Compound 83

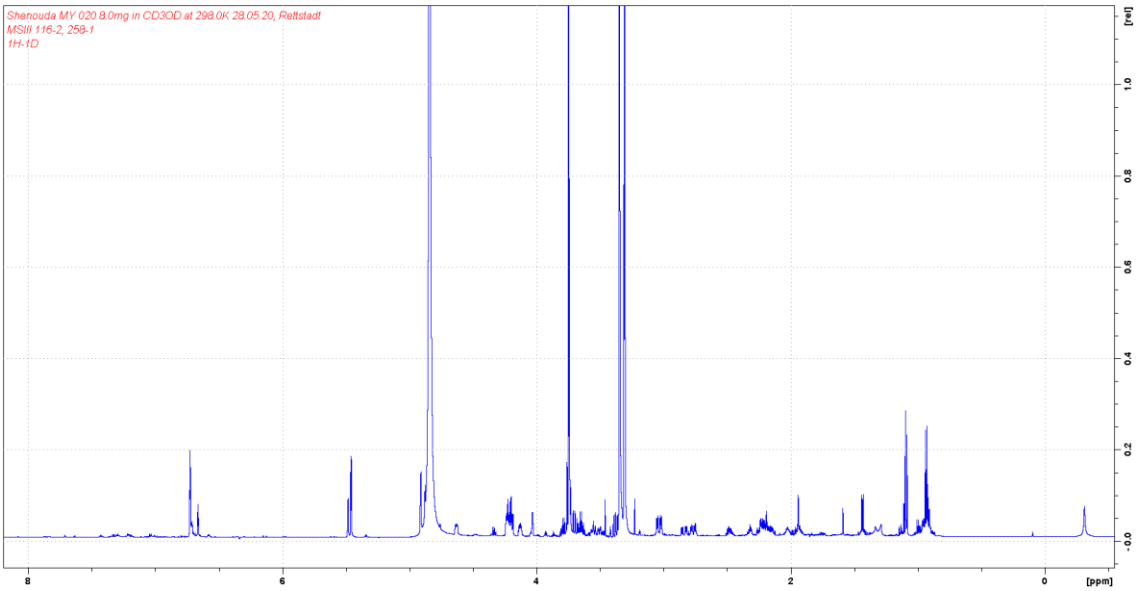


Figure 7.1 ^{1}H-NMR spectrum of compound **83** (hydrated scytolide, 500 MHz, CD_3OD).

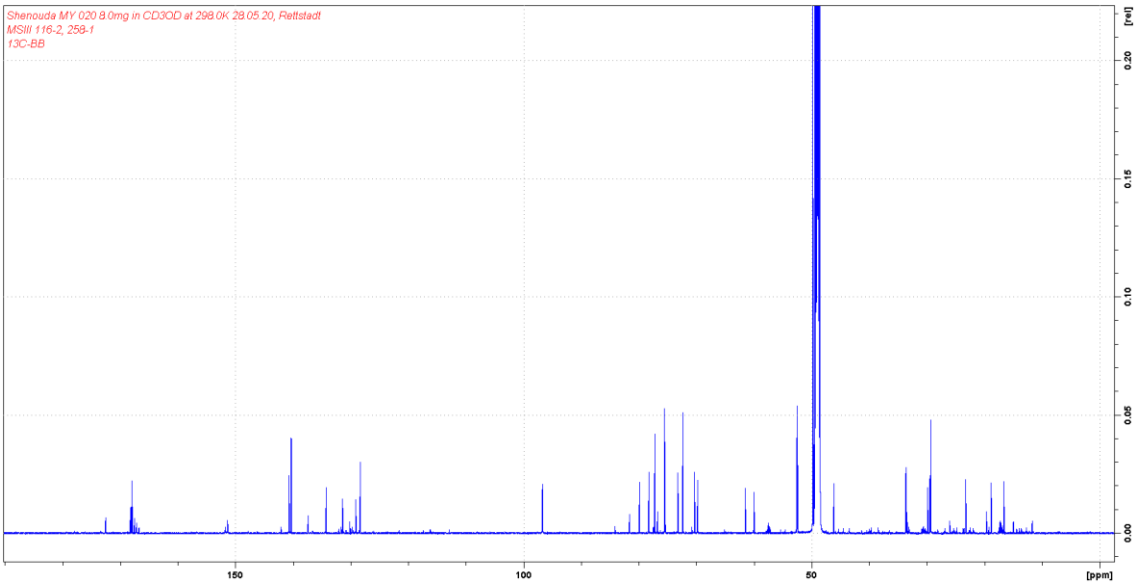


Figure 7.2 ^{13}C-NMR spectrum of compound **83** (hydrated scytolide, 125 MHz, CD_3OD).

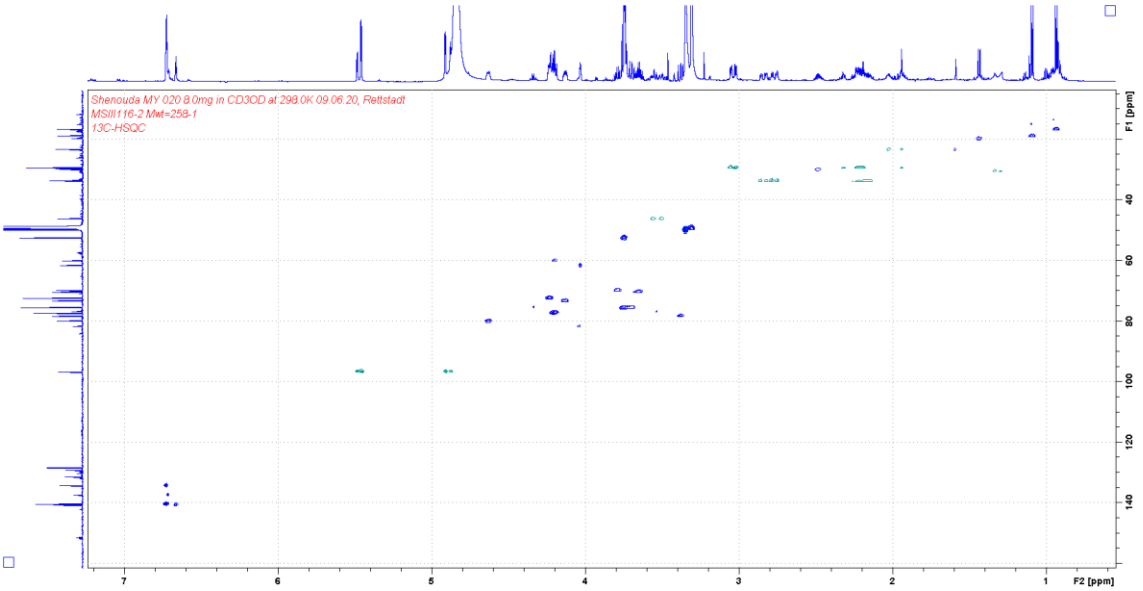


Figure 7.3 HSQC spectrum of compound **83** (hydrated scytolide).

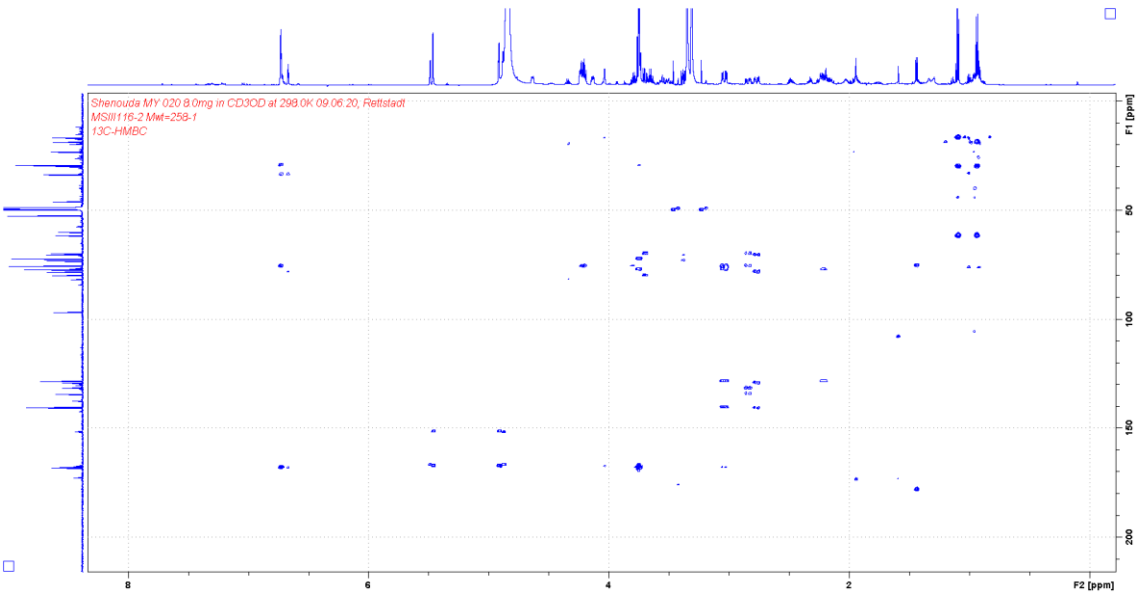


Figure 7.4 HMBC spectrum of compound **83** (hydrated scytolide).

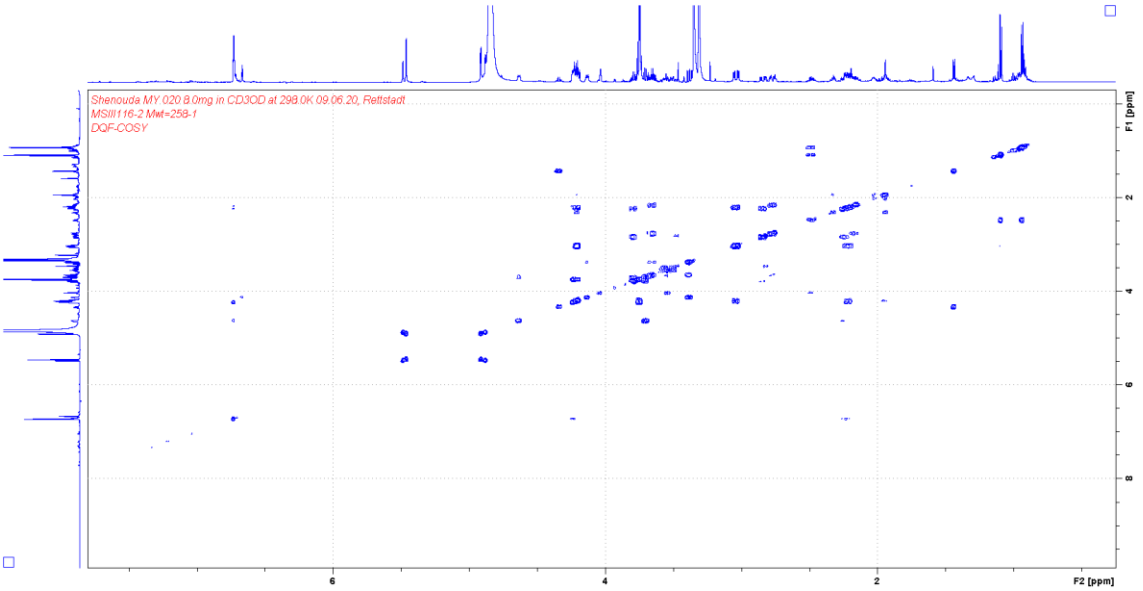


Figure 7.5 COSY spectrum of compound **83** (hydrated scytolide).

7.1.2 Compound 84

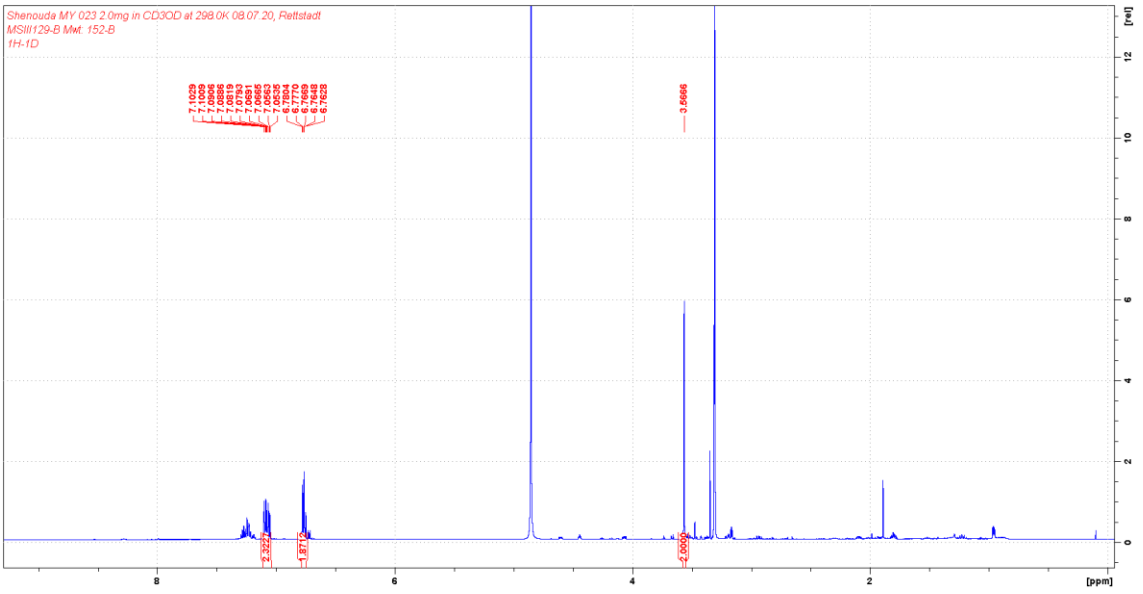


Figure 7.6 ^{1}H-NMR spectrum of compound **84** (ortho-hydroxypheny acetic acid, 500 MHz, CD_3OD).

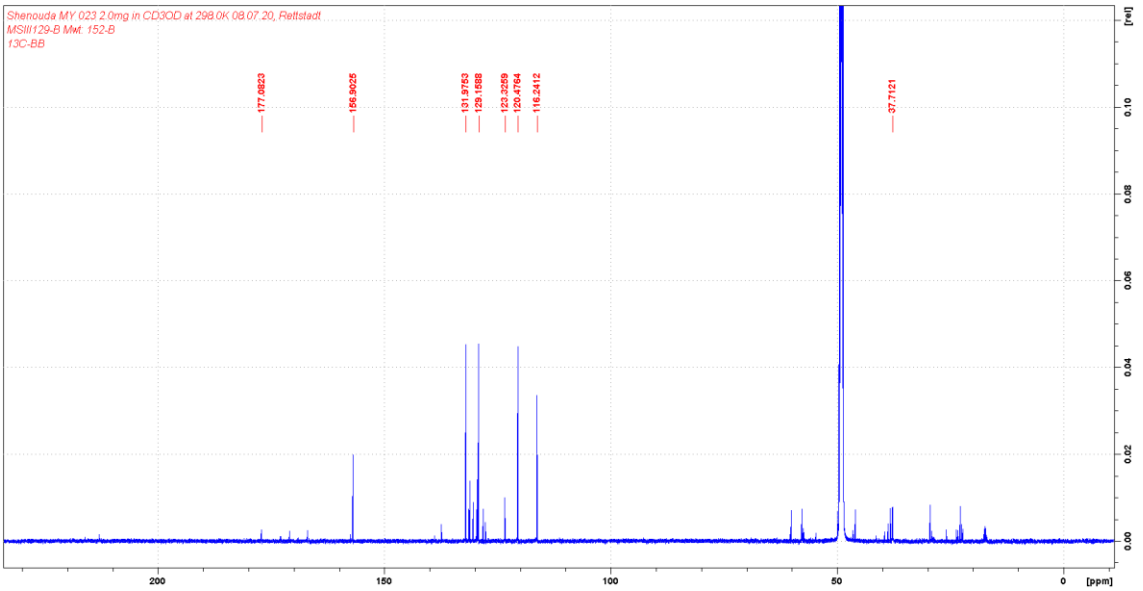


Figure 7.7 ^{13}C-NMR spectrum of compound **84** (ortho-hydroxypheny acetic acid, 125 MHz, CD_3OD).

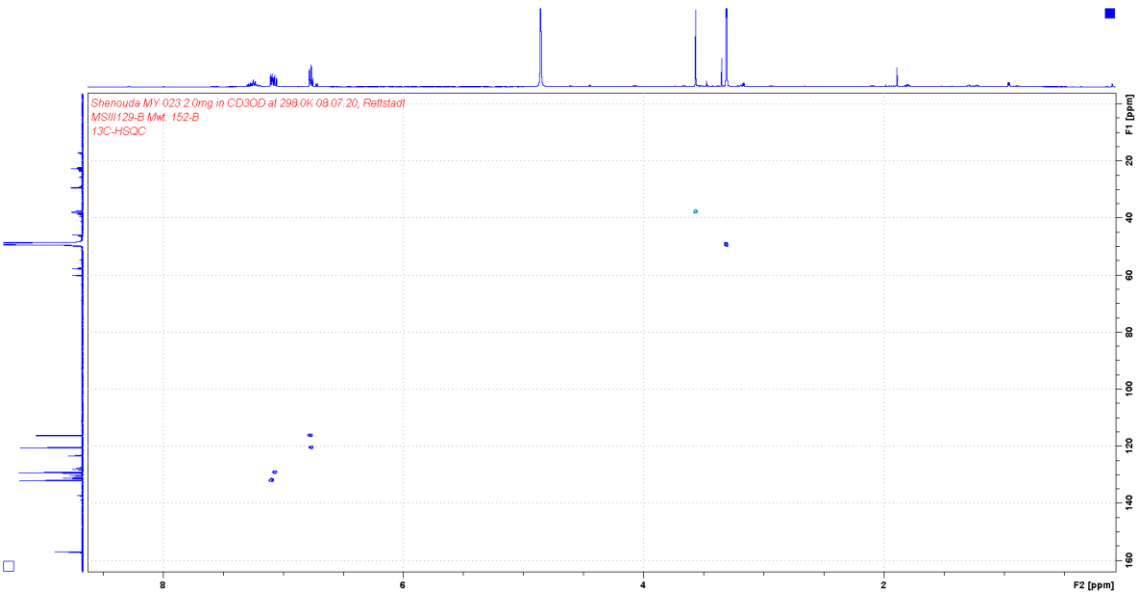


Figure 7.8 HSQC spectrum of compound **84** (ortho-hydroxyphenyl acetic acid).

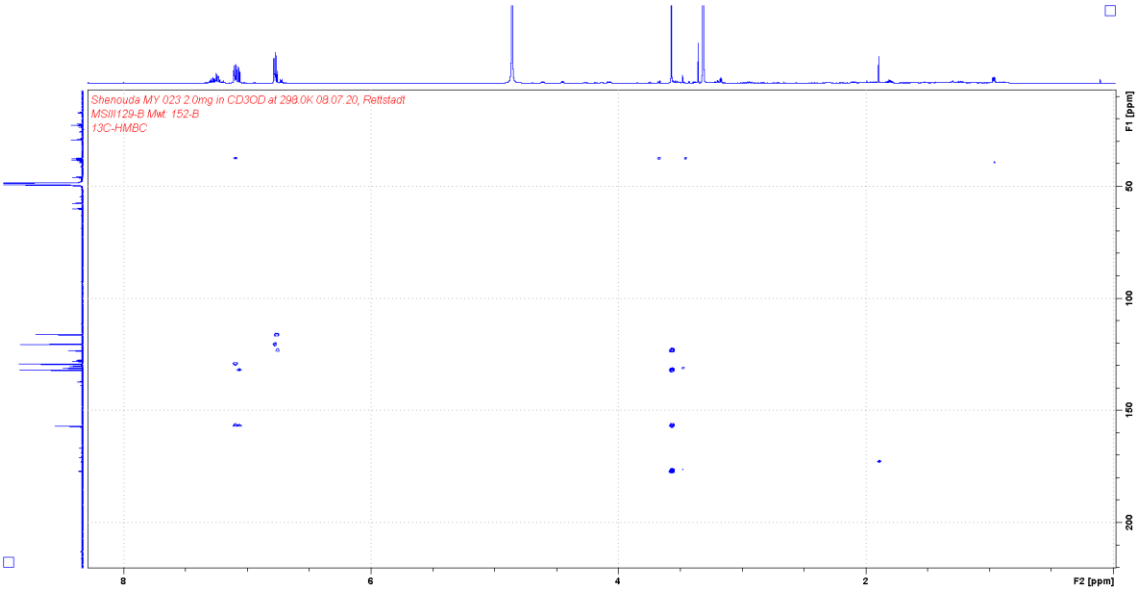


Figure 7.9 HMBC spectrum of compound **84** (ortho-hydroxyphenyl acetic acid).

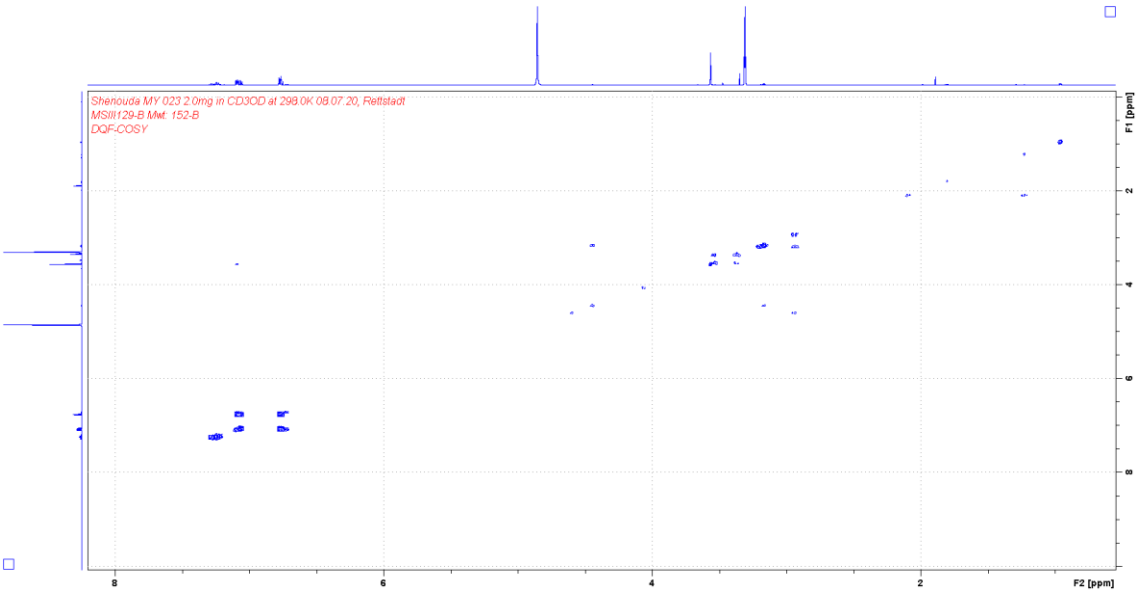


Figure 7.10 COSY spectrum of compound **84** (ortho-hydroxyphenyl acetic acid).

7.1.3 Compound 85

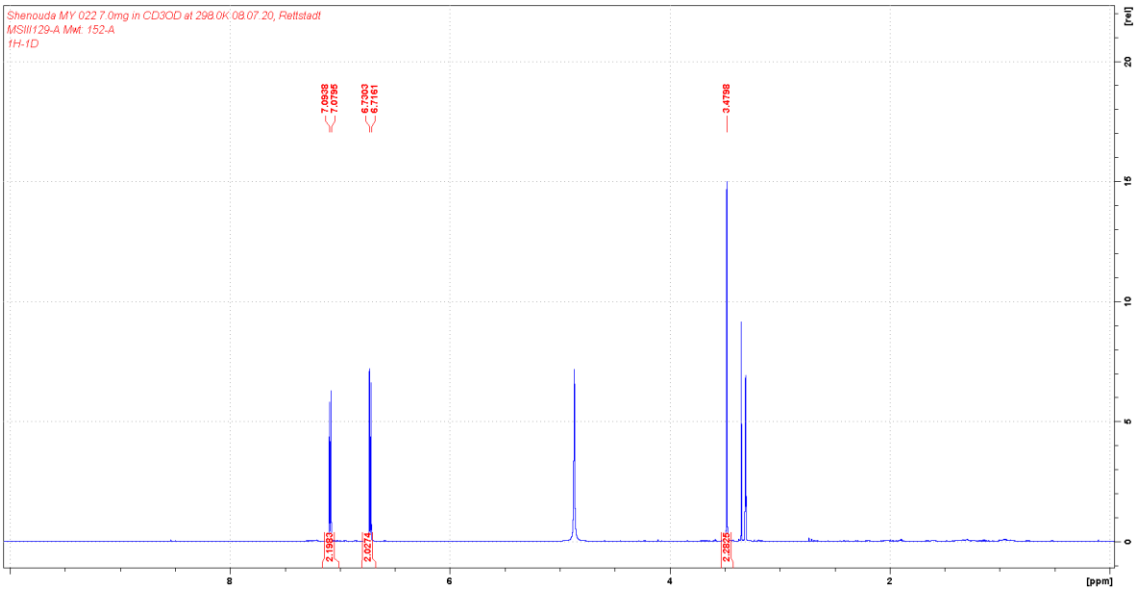


Figure 7.11 ^{1}H-NMR spectrum of compound **85** (para-hydroxypheny acetic acid, 500 MHz, CD_3OD).

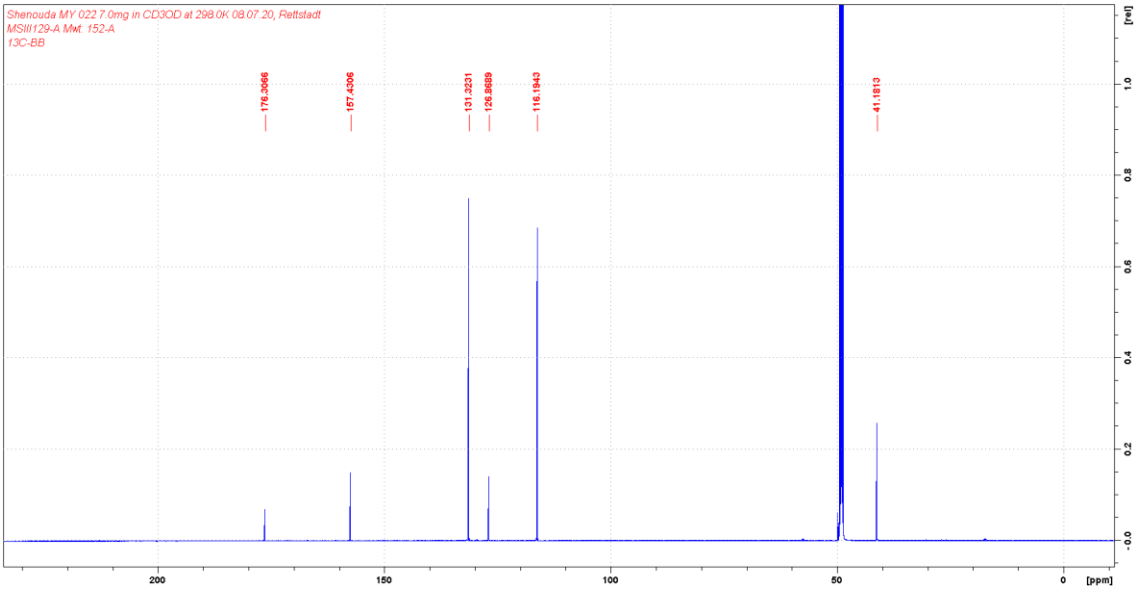


Figure 7.12 ^{13}C-NMR spectrum of compound **85** (para-hydroxypheny acetic acid, 125 MHz, CD_3OD).

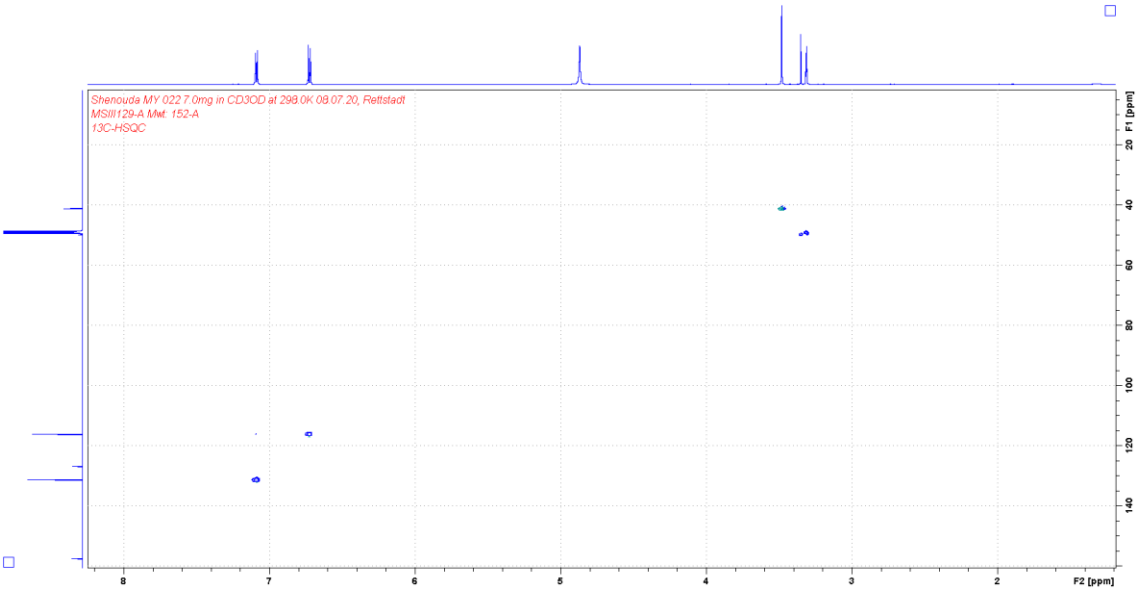


Figure 7.13 HSQC spectrum of compound **85** (para-hydroxypheny acetic acid).

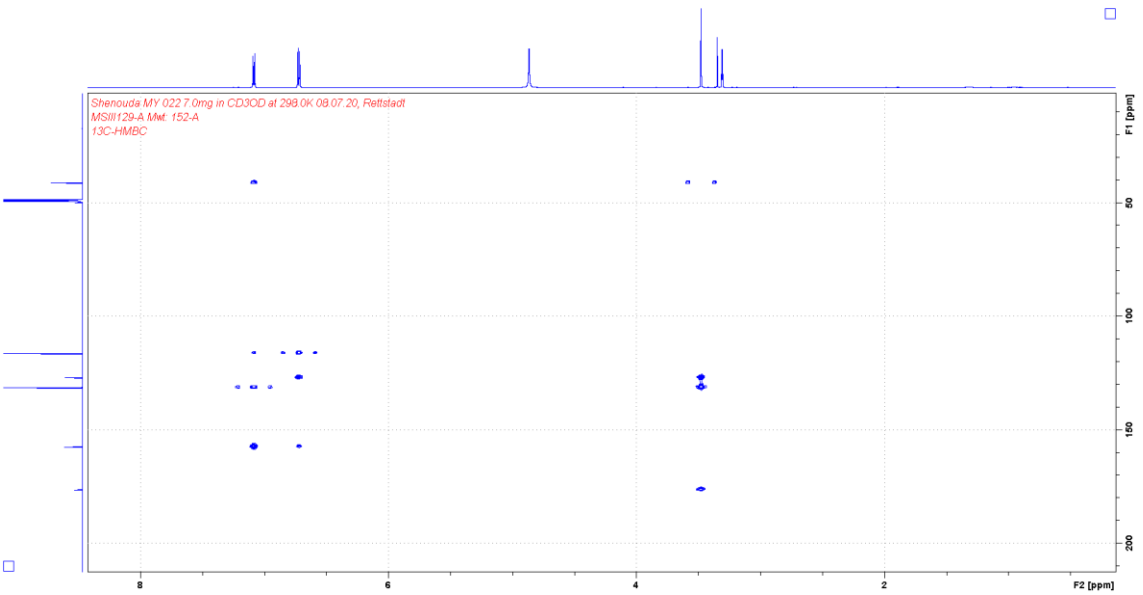


Figure 7.14 HMBC spectrum of compound **85** (para-hydroxyphenyl acetic acid).

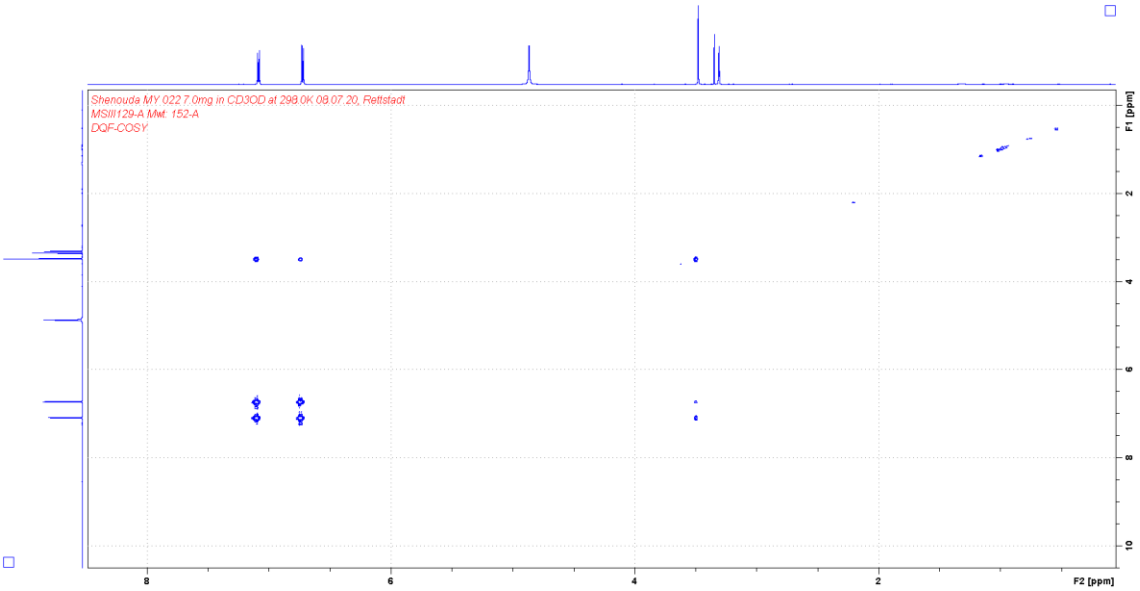


Figure 7.15 COSY spectrum of compound **85** (para-hydroxyphenyl acetic acid).

7.1.4 Compound 86

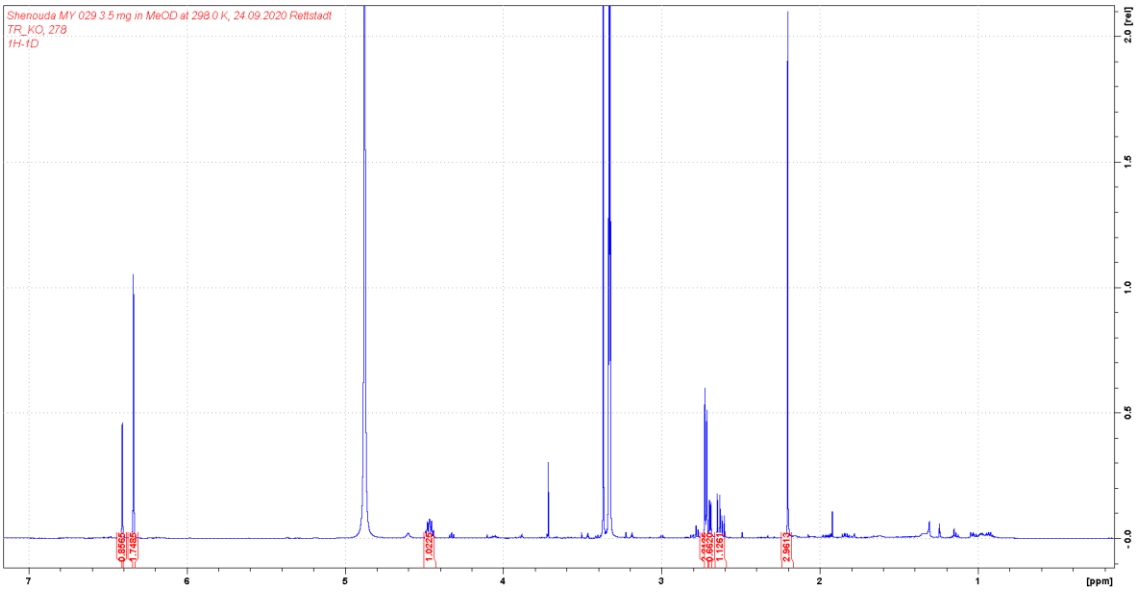


Figure 7.16 ^{1}H-NMR spectrum of compound **86** (citreoisocoumarin, 500 MHz, CD_3OD).

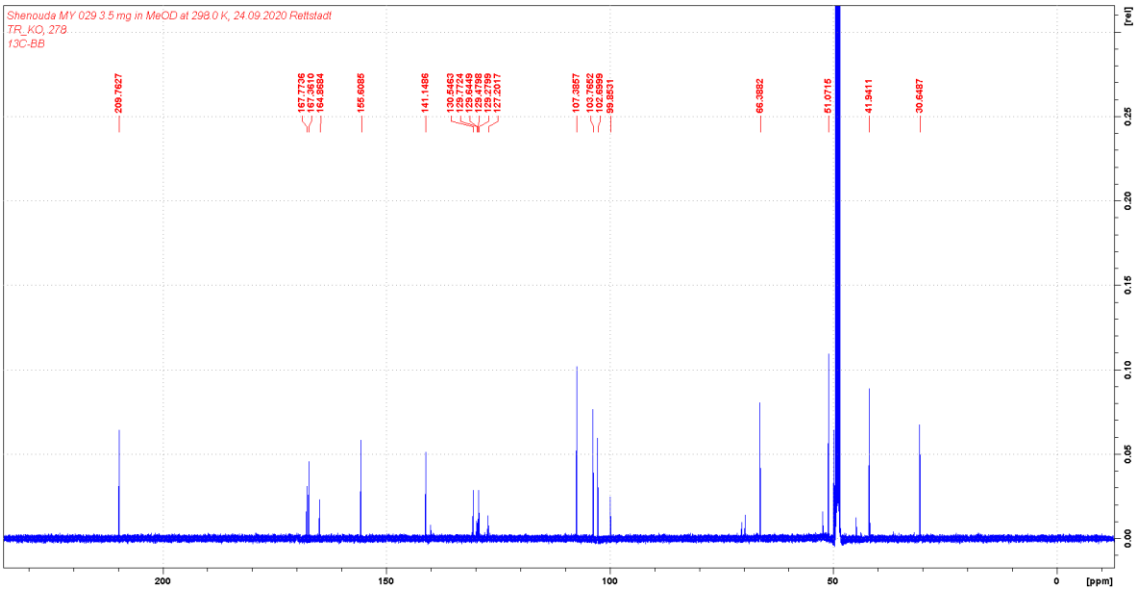


Figure 7.17 ^{13}C-NMR spectrum of compound **86** (Citreoisocoumarin, 125 MHz, CD_3OD).

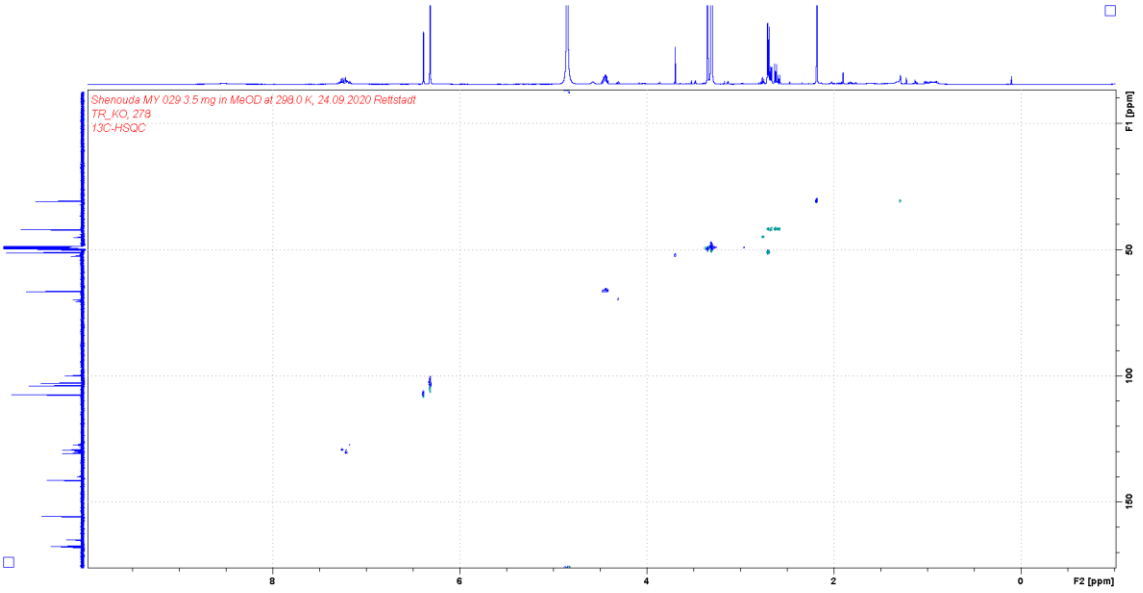


Figure 7.18 HSQC spectrum of compound **86** (Citroisocoumarin).

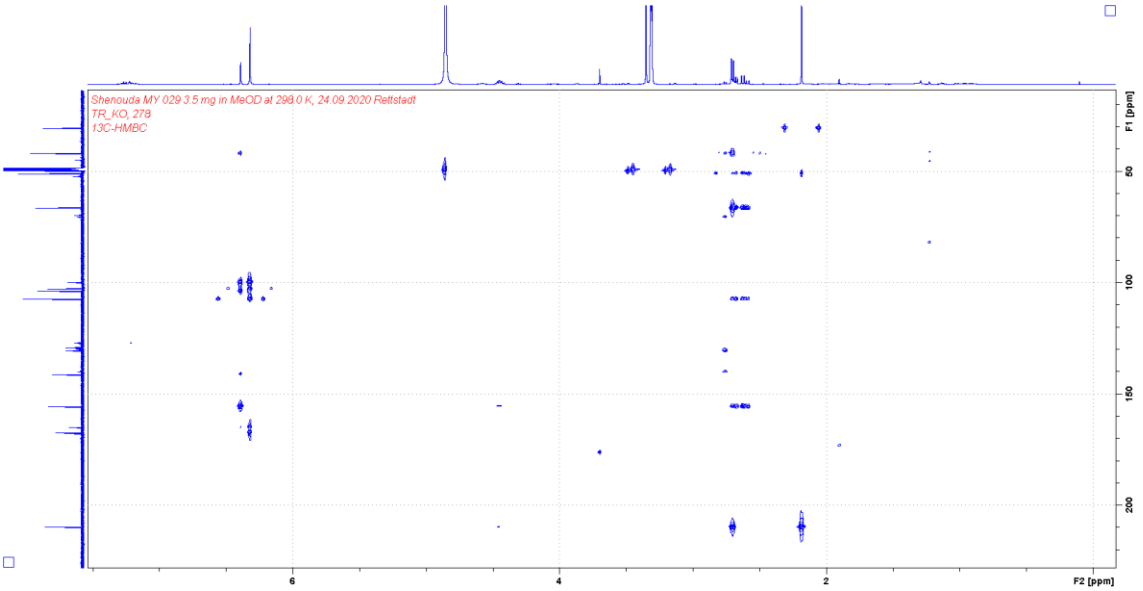


Figure 7.19 HMBC spectrum of compound **86** (Citroisocoumarin).

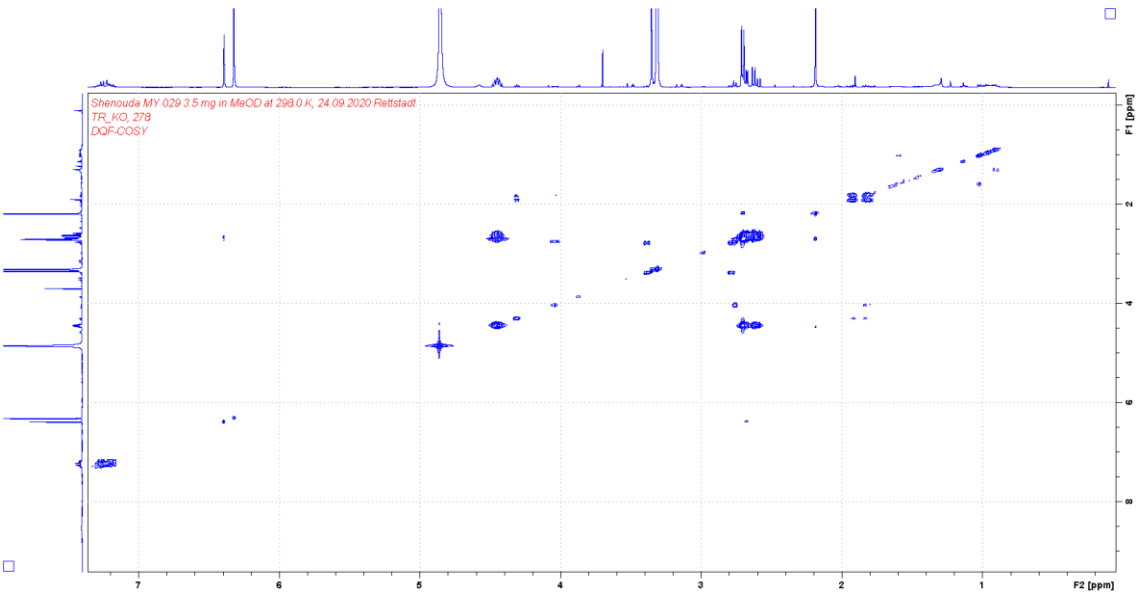


Figure 7.20 COSY spectrum of compound **86** (Citroisocoumarin).

7.1.5 Compound 40

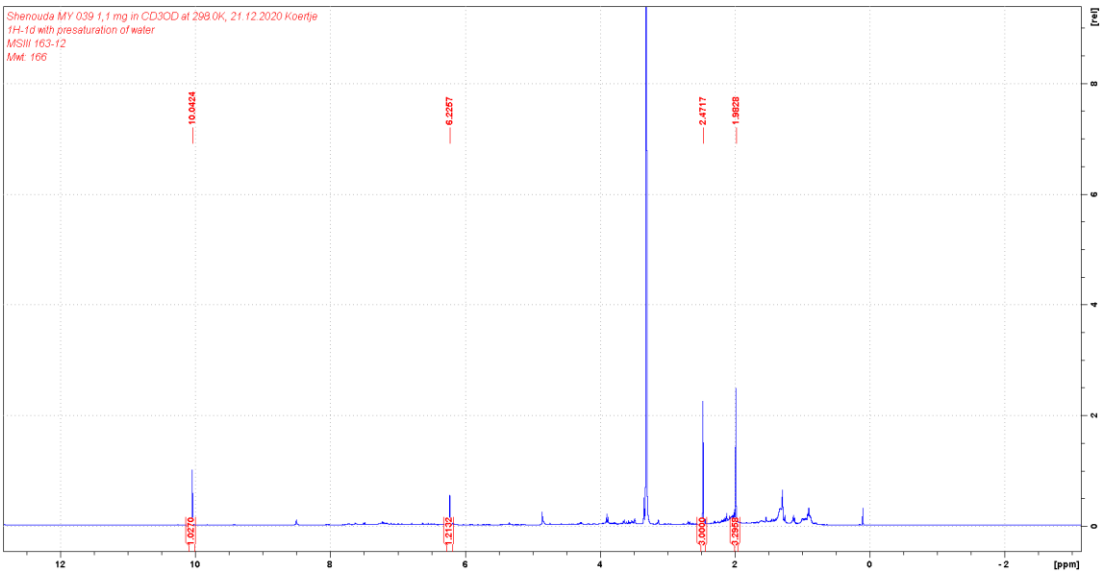


Figure 7.21 ^{1}H-NMR spectrum of compound **40** (3-methylorcinaldehyde, 500 MHz, CD_3OD).

7.1.6 Compound 41

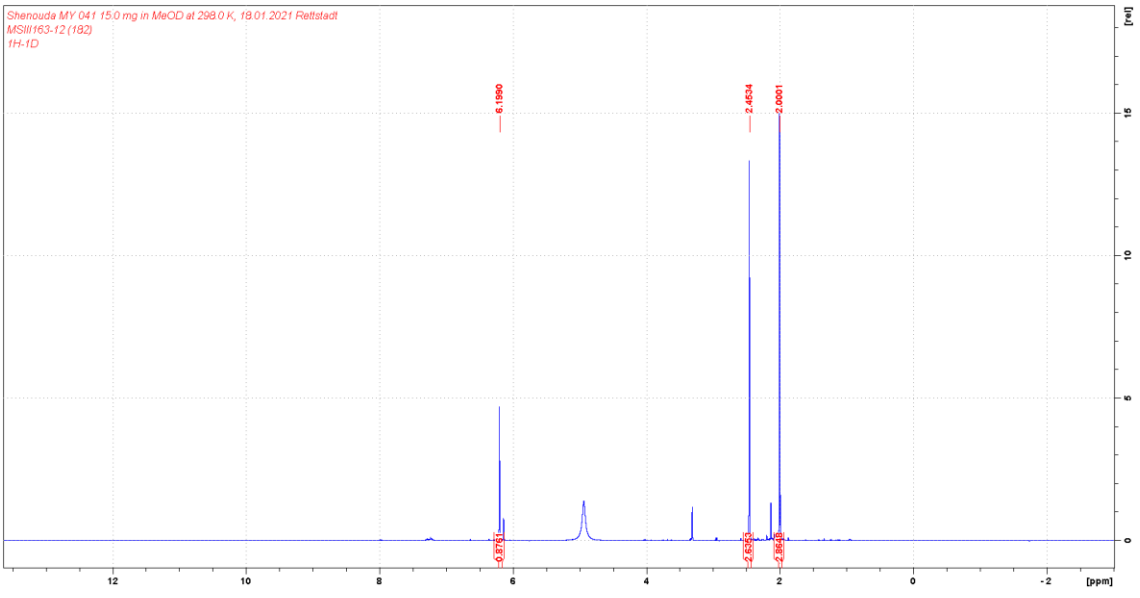


Figure 7.22 ^{1}H-NMR spectrum of compound **41** (3-methylorsellenic acid, 500 MHz, CD_3OD).

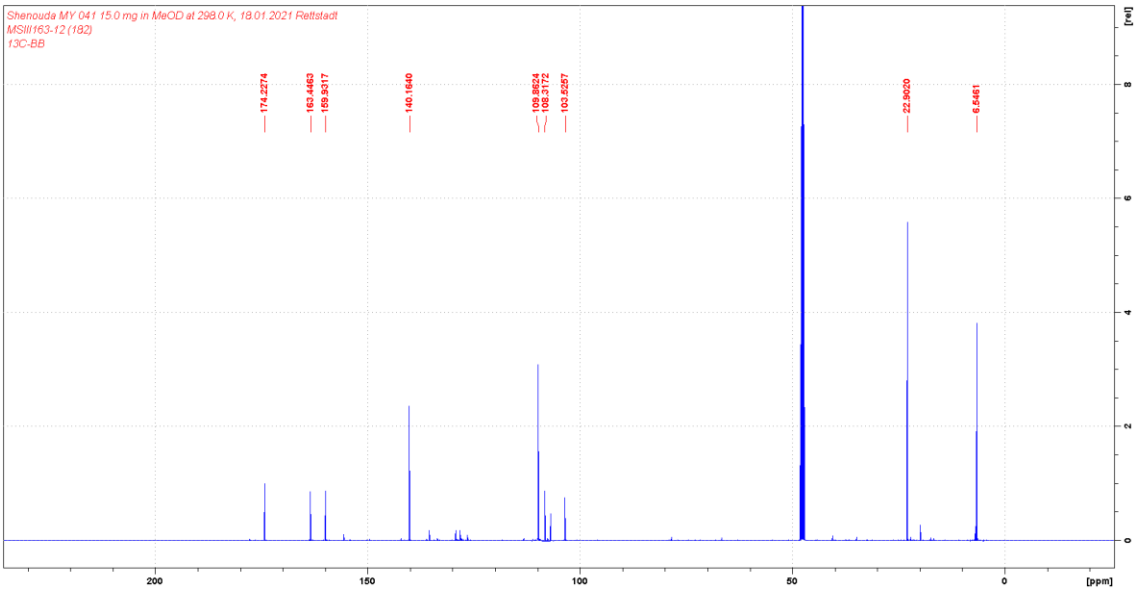


Figure 7.23 ^{13}C-NMR of compound **41** (3-methylorsellenic acid, 125 MHz, CD_3OD).

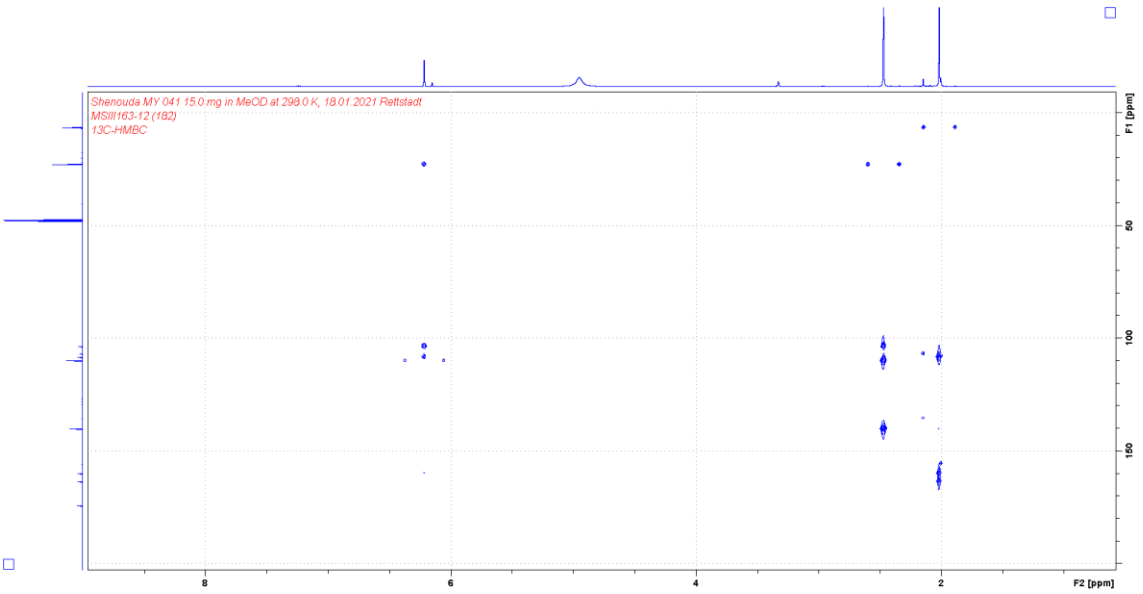


Figure 7.24 HMBC spectrum of compound **41** (3-methylorsellenic acid).

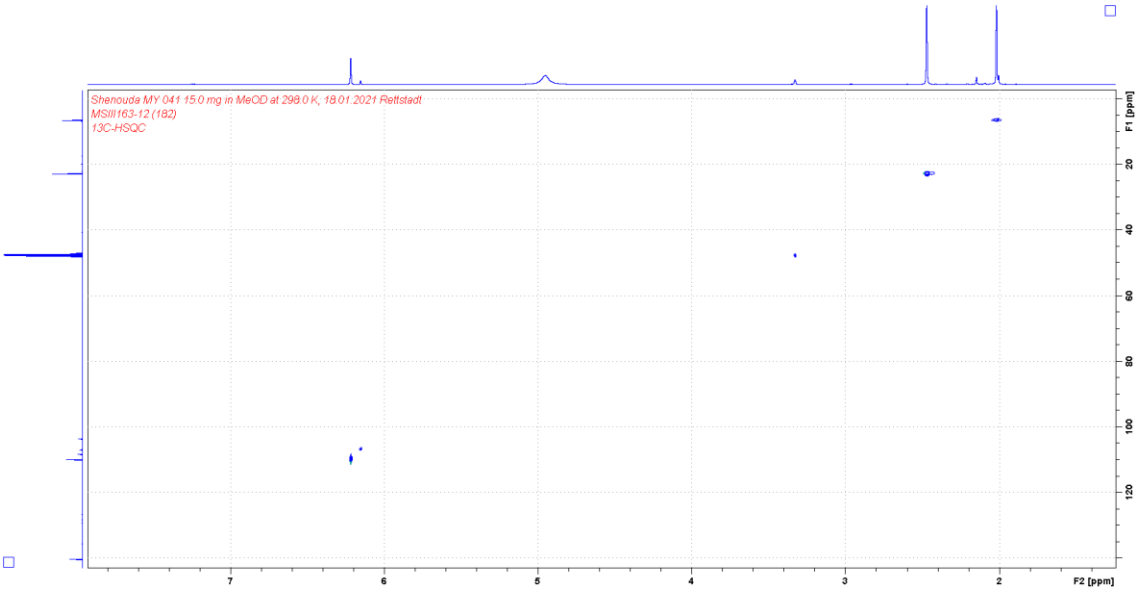


Figure 7.25 HSQC spectrum of compound **41** (3-methylorsellenic acid).

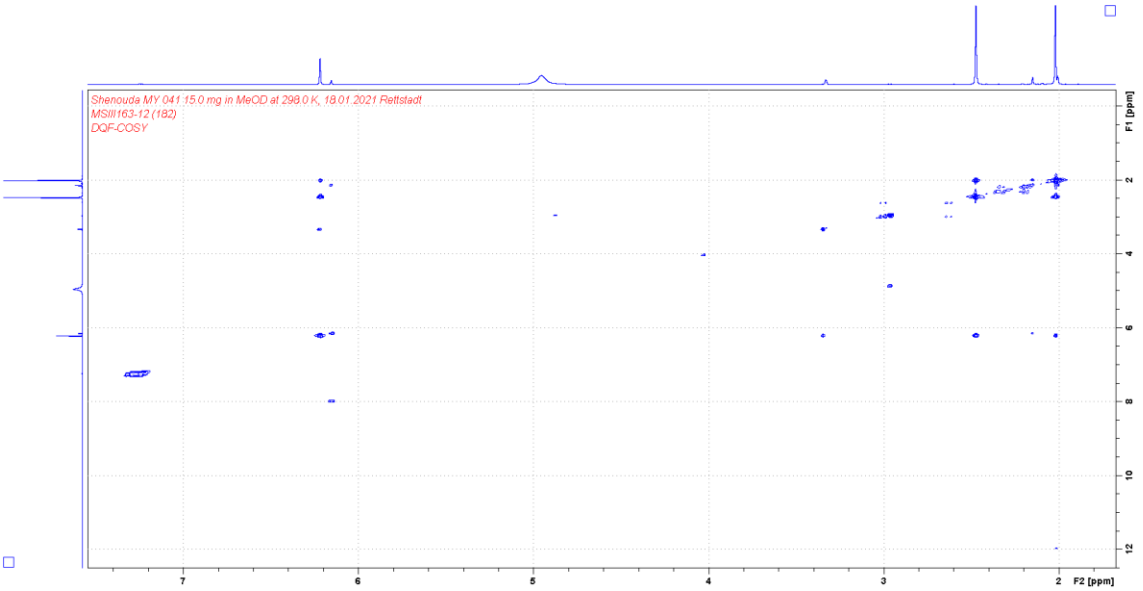


Figure 7.26 COSY spectrum of compound **41** (3-methylorsellenic acid).

7.1.7 Compound 79

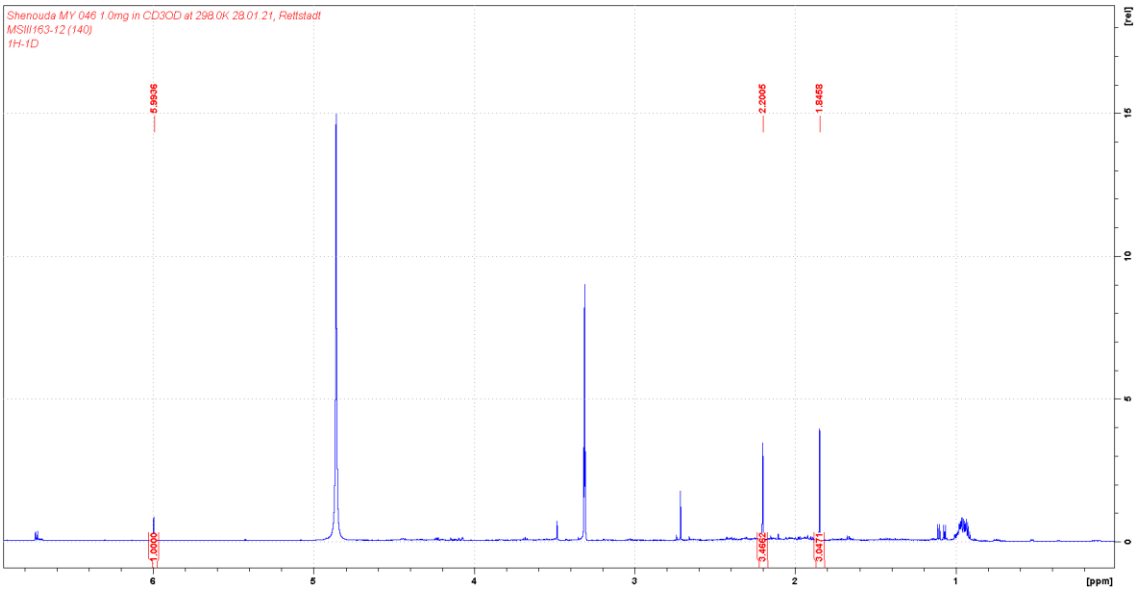


Figure 7.27 ^{1}H-NMR spectrum of compound **79** (2-pyranone 140, 500 MHz, CD_3OD).

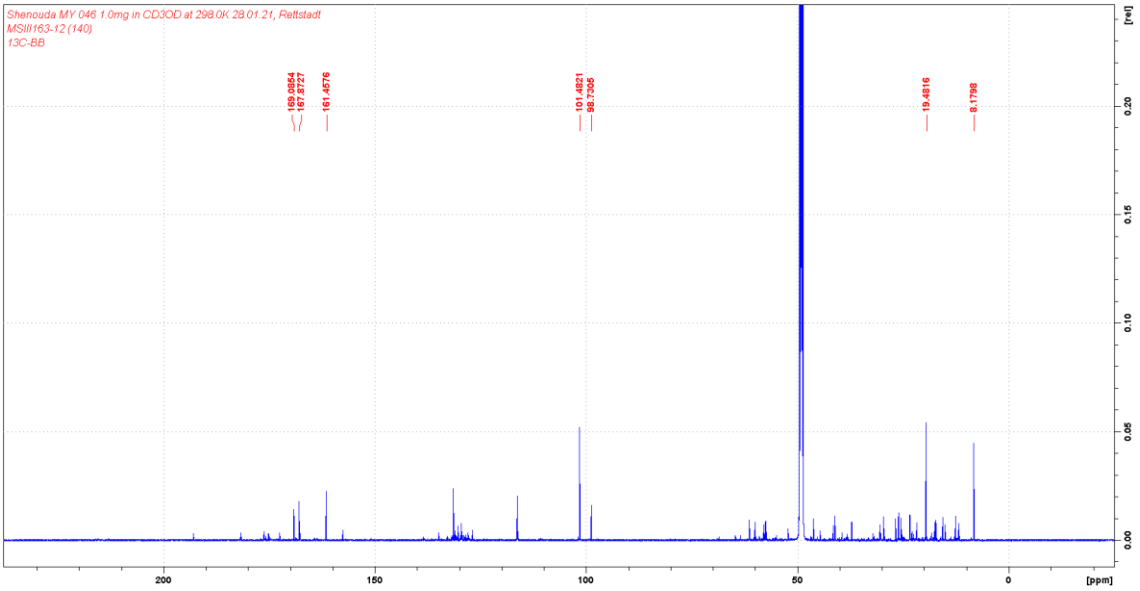
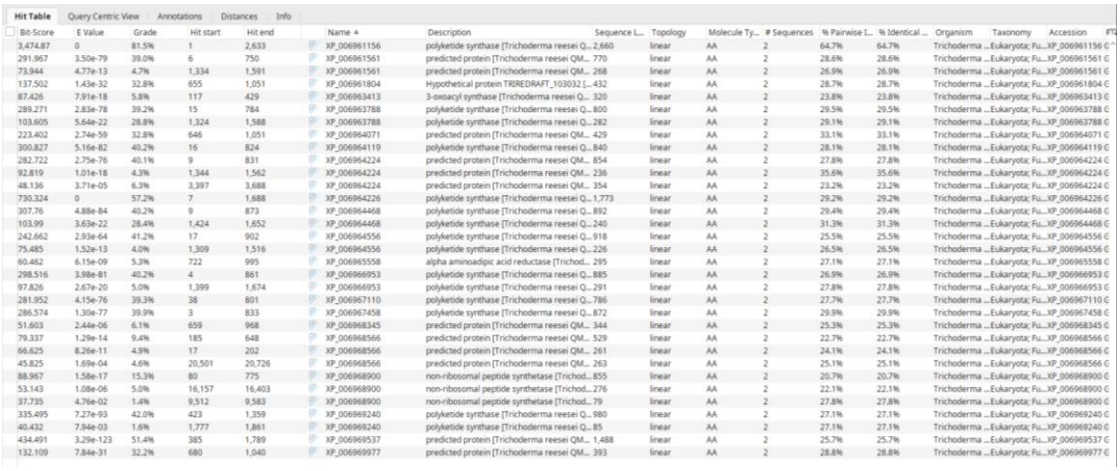


Figure 7.28 ^{13}C-NMR spectrum of compound **79** (2-pyranone 140, 125 MHz, CD_3OD).

7.2 Chapter 3 Additional data

7.2.1 Bioinformatic analysis of all PKS and PKs-NRPS genes from *T. reesei* genome



Hit Table | Query Centric View | Annotations | Distances | Info

Bit-Score	E Value	Grade	Hit start	Hit end	Name	Description	Sequence L...	Topology	Molecule Ty...	# Sequences	% Pairwise I...	% Identical ...	Organism	Taxonomy	Accession
3,474.87	0	81.5%	1	2,633	XP_006961156	polyketide synthase [Trichoderma reesei Q...	2,660	linear	AA	2	64.7%	64.7%	Trichoderma ...	Eukaryota; Fu...	XP_006961156
291.967	3.50e-79	39.0%	6	750	XP_006961561	predicted protein [Trichoderma reesei QM...	770	linear	AA	2	28.6%	28.6%	Trichoderma ...	Eukaryota; Fu...	XP_006961561
73.944	4.77e-13	4.7%	1,334	1,591	XP_006961561	predicted protein [Trichoderma reesei QM...	268	linear	AA	2	26.9%	26.9%	Trichoderma ...	Eukaryota; Fu...	XP_006961561
137.502	1.43e-32	32.8%	655	1,051	XP_006961804	Hypothetical protein TRIREDRAFT_103032 [...	432	linear	AA	2	28.7%	28.7%	Trichoderma ...	Eukaryota; Fu...	XP_006961804
87.426	7.91e-18	5.8%	117	429	XP_006963413	3-oxoacyl synthase [Trichoderma reesei Q...	320	linear	AA	2	23.8%	23.8%	Trichoderma ...	Eukaryota; Fu...	XP_006963413
289.271	2.83e-78	39.2%	15	784	XP_006963788	polyketide synthase [Trichoderma reesei Q...	800	linear	AA	2	29.5%	29.5%	Trichoderma ...	Eukaryota; Fu...	XP_006963788
103.605	5.64e-22	28.8%	1,324	1,588	XP_006963788	polyketide synthase [Trichoderma reesei Q...	282	linear	AA	2	29.1%	29.1%	Trichoderma ...	Eukaryota; Fu...	XP_006963788
223.402	2.74e-59	32.8%	646	1,051	XP_006964071	predicted protein [Trichoderma reesei QM...	429	linear	AA	2	33.1%	33.1%	Trichoderma ...	Eukaryota; Fu...	XP_006964071
300.827	5.16e-82	40.2%	16	824	XP_006964119	polyketide synthase [Trichoderma reesei Q...	840	linear	AA	2	28.1%	28.1%	Trichoderma ...	Eukaryota; Fu...	XP_006964119
282.722	2.75e-76	40.1%	9	831	XP_006964224	predicted protein [Trichoderma reesei QM...	854	linear	AA	2	27.8%	27.8%	Trichoderma ...	Eukaryota; Fu...	XP_006964224
92.819	1.01e-18	4.3%	1,344	1,562	XP_006964224	predicted protein [Trichoderma reesei QM...	236	linear	AA	2	35.6%	35.6%	Trichoderma ...	Eukaryota; Fu...	XP_006964224
48.136	3.71e-05	6.3%	3,397	3,688	XP_006964224	predicted protein [Trichoderma reesei QM...	354	linear	AA	2	23.2%	23.2%	Trichoderma ...	Eukaryota; Fu...	XP_006964224
730.324	0	57.2%	7	1,688	XP_006964226	polyketide synthase [Trichoderma reesei Q...	1,773	linear	AA	2	29.2%	29.2%	Trichoderma ...	Eukaryota; Fu...	XP_006964226
307.76	4.88e-84	40.2%	9	873	XP_006964468	polyketide synthase [Trichoderma reesei Q...	892	linear	AA	2	29.4%	29.4%	Trichoderma ...	Eukaryota; Fu...	XP_006964468
103.99	3.63e-22	28.4%	1,424	1,652	XP_006964468	polyketide synthase [Trichoderma reesei Q...	240	linear	AA	2	31.3%	31.3%	Trichoderma ...	Eukaryota; Fu...	XP_006964468
242.662	2.93e-64	41.2%	17	902	XP_006964556	polyketide synthase [Trichoderma reesei Q...	918	linear	AA	2	25.5%	25.5%	Trichoderma ...	Eukaryota; Fu...	XP_006964556
75.485	1.52e-13	4.0%	1,309	1,516	XP_006964556	polyketide synthase [Trichoderma reesei Q...	226	linear	AA	2	26.5%	26.5%	Trichoderma ...	Eukaryota; Fu...	XP_006964556
60.462	6.15e-09	5.3%	722	995	XP_006965558	alpha aminoadipic acid reductase [Trichod...	295	linear	AA	2	27.1%	27.1%	Trichoderma ...	Eukaryota; Fu...	XP_006965558
298.516	3.98e-81	40.2%	4	861	XP_006966953	polyketide synthase [Trichoderma reesei Q...	885	linear	AA	2	26.9%	26.9%	Trichoderma ...	Eukaryota; Fu...	XP_006966953
97.826	2.67e-20	5.0%	1,399	1,674	XP_006966953	polyketide synthase [Trichoderma reesei Q...	291	linear	AA	2	27.8%	27.8%	Trichoderma ...	Eukaryota; Fu...	XP_006966953
281.952	4.15e-76	39.3%	38	801	XP_006967110	polyketide synthase [Trichoderma reesei Q...	786	linear	AA	2	27.7%	27.7%	Trichoderma ...	Eukaryota; Fu...	XP_006967110
286.574	1.30e-77	39.9%	3	833	XP_006967458	polyketide synthase [Trichoderma reesei Q...	872	linear	AA	2	29.9%	29.9%	Trichoderma ...	Eukaryota; Fu...	XP_006967458
51.603	2.44e-06	6.1%	659	968	XP_006968345	predicted protein [Trichoderma reesei QM...	344	linear	AA	2	25.3%	25.3%	Trichoderma ...	Eukaryota; Fu...	XP_006968345
79.337	1.29e-14	9.4%	185	648	XP_006968566	predicted protein [Trichoderma reesei QM...	529	linear	AA	2	22.7%	22.7%	Trichoderma ...	Eukaryota; Fu...	XP_006968566
66.625	8.26e-11	4.9%	17	202	XP_006968566	predicted protein [Trichoderma reesei QM...	261	linear	AA	2	24.1%	24.1%	Trichoderma ...	Eukaryota; Fu...	XP_006968566
45.825	1.69e-04	4.6%	20,501	20,726	XP_006968566	predicted protein [Trichoderma reesei QM...	263	linear	AA	2	25.1%	25.1%	Trichoderma ...	Eukaryota; Fu...	XP_006968566
88.967	1.58e-17	15.3%	80	775	XP_006968900	non-ribosomal peptide synthetase [Trichod...	855	linear	AA	2	20.7%	20.7%	Trichoderma ...	Eukaryota; Fu...	XP_006968900
53.143	1.08e-06	5.0%	16,157	16,403	XP_006968900	non-ribosomal peptide synthetase [Trichod...	276	linear	AA	2	22.1%	22.1%	Trichoderma ...	Eukaryota; Fu...	XP_006968900
37.735	4.76e-02	1.4%	9,512	9,583	XP_006968900	non-ribosomal peptide synthetase [Trichod...	79	linear	AA	2	27.8%	27.8%	Trichoderma ...	Eukaryota; Fu...	XP_006968900
335.495	7.27e-93	42.0%	423	1,359	XP_006969240	polyketide synthase [Trichoderma reesei Q...	980	linear	AA	2	27.1%	27.1%	Trichoderma ...	Eukaryota; Fu...	XP_006969240
40.432	7.94e-03	1.6%	1,777	1,861	XP_006969240	polyketide synthase [Trichoderma reesei Q...	85	linear	AA	2	27.1%	27.1%	Trichoderma ...	Eukaryota; Fu...	XP_006969240
434.491	3.29e-123	51.4%	385	1,789	XP_006969537	predicted protein [Trichoderma reesei QM...	1,488	linear	AA	2	25.7%	25.7%	Trichoderma ...	Eukaryota; Fu...	XP_006969537
132.109	7.84e-31	32.2%	680	1,040	XP_006969977	predicted protein [Trichoderma reesei QM...	393	linear	AA	2	28.8%	28.8%	Trichoderma ...	Eukaryota; Fu...	XP_006969977

Figure 7.29 Blastp results of all proteins form *T. reesei* QM6a against PKS from *P. chrysogenum* (Geneious Prime).

Gene sequences used in the expression of the first 4 genes in PKS/NRPS-58285 BGC

PKS-NRPS (TRIREDRAFT_58285):

ATGGATCAACAAAGACAGCAGGATGGGCGGTTTCGCGAACCAATAGCGGTGATTGGATCCGCCTGTAGATTCCCA
GGAGCGTGTAGCTCCCCATCAAAGCTATGGGATCTCCTCCAACAGCCCCGTGATGTTCTGAAAGAATTTCACCCCG
AACGACTCAATCTATCCCGCTTCCACCACAAGAGTGGTGATACCCATGGCGCTACAGACGTCGGAAACAAATCCTA
CCTTCTTGAAGAGGACACTCGCCTTTTCGATGCTGCCTTCTTTGGCATCAGCCCAATCGAGGCTGCAGGCATGGAT
CCACAGCAGAGAATATTGTTAGAGGCTGTTTATGAGACATTCGAGTCCGCCGGAGTGACTCTTGATGAACTGAGAG
GCTCCTTCACATCCGTTCATGTGGGATGCATGACGTCGGATTGGGCAAACATCCAGGCCCGCGATACCGAGACGG
TGCCCAAGTACAATGCTACTGGCTCAGCCAACAGCATCTTGTCAAACCGGATTTCGTACATCTTCGACCTCAAGGG
CCCGTCTGAGACGATAGACACGGCCTGCTCGAGCTCCCTGGTGGCTTTGCACAATGCTGCGCGGGGTCTTCTCAA
CGGTGACTGCAACGCCGCTGTCGTTGCTGGTGTGAACTTGATCCTGGATCCCGCCCCATTCATCAACGAGTCTAAG
CTACACATGCTGTCTCCCGATGCCAGATCACGCATGTGGGACAAAGATGCGAATGGCTATGCGCGTGGTGAAGGT
GCTGGTGCATTGCTGCTCAAGACGCTGAGCCAAGCTCTGAAAGACGGAGACCATATCGAGGGTCTCGTAAGGGCC
ACTGGAGTCAACTCGGACGGCCAGAGCCCAGGGATCACAATGCCGTATTCTCCGACACAGACTGCTCTTATTCAAC
AGACGTACGCGCGAGCTGGGCTGGATCCCATCAAAGACCGACCGCAGTACATCGAATGCCACGGTACAGGTACCC
CCGCAGGCGACCCTGTGGAAGCACGAGCACTTAGCAATGCCTTCATCGGCCATGTTGACAAGTCGACCGAAAACC
CGATCCTCGTCGGATCAATCAAAACCGTTATCGGACACTTGGAGGGTGGCGCTGGCATTGCTGGCGTCATAAAAGT
CCTGCTCAGCATCAAGTACCGCGTTATCCCCCCCAACCTCCTCTTCAAGGACCTAAATCCGGATATCGCCCCCTAC
TATGGGCCCCTTCAAATCCCAACGACGGCCGTGCCCTGGCCAGAACTGCCACCTGGTACTCCTGCTCGCGCAAGC
GTGAACAGCTTTGGATTCGGGGGAACCAATGCTCACGCCGTCATCGAGAGTTTCGATGAGAGCTTCGGACCCCGA
GCTAGCTCTGCTGATGAAGGCGTGGTTGGTCCTTTGTTGTTCTCCGCGAAGTCTGGAGCTTCATTGCTGCGGTCTG
TCAGGGACAATTTGGAGTATCTGCAAAAGGACGCATCGCTTGACCTCCGCGACTTGAGCGCTTTTTTGCAGTCAAG
GCGTACCACACACCGTGTGCGTGCTCACTTTACCGGGTCATCTAGAGATGCAGTTCTTGAGAACATGACCAACTTC
ATTACAACTCATGCCCAGTCGGCCAGTGATCAGATCGGCCACCAGCCTCAGCTTATCAACCCGAAGGAGGTTCCC
GGCGTCCTCGGTGTATTCACGGGGCAGGGCGCTCAGTGGCCCGCCATGGGTCGTGAGCTGTACTGCAAGTCAGC
ACTCTTTCGCAAGACCATTGAGGAGTGCGAAGCAGTTCTTGATGCGCTTCCGCAGCAAGACCGCCCCGATTGGTCT
CTCACGGAAGAGCTCATGAAGGACGCTATATCTTCCCGTATTTCCGAGGCGGCCCTGTCTCAACCGCTCTGCTCAG
CCGTCCAGCTTGCATTGGTCAAGCTACTCCAGGCGGCTGGCATCGCATTTGACGCAGTTGTCGGCCACTCATCTG
GTGAGATTGCCGCCACTTTTGCGGCCGGTATCATCACGCTCCAGGGCGCCATGCAGATTGCCTATTACCGCGGTT
TACATGCGAAGTCTGCTGGTGGCGCCAATCATGCCAAGGGCGCCATGATGGCCGTTGGGTTGTCGTTCGCGGAAG
CTAAGCAATTCTGCTCACGCCCGGAATTCGACGGTAAGATCAAAGTCGCAGCTAGTAACGCCCCAAAGTCTGTGAC
GATCTCCGGCGATGTCGACGCTATTGTTCAGGCGCACGAGATCCTTCAGGCAGACAACATCTTTGCGCGCCGCTT
ACAAGTGGACACTGCATACCACTCGCATCACATGTTGCCCTGCTGCCAGCCCTATCTAAACTCGCTGCTGGCTTGC
AACATCAAAGTCATGCAGCCTACGGGCAAGTGTACTTGGATTTCCAGCGTACGAGGCGACACACAGCTCCTCAGG
GGTGATCTAAGTTCTCTCAAAGGACAATACTGGGTTGACAATATGGTTCGCACGGTTCTGTTTACTCAAGCTATTGA
GTCTTCCATCTGGCATGGAGGCCCGTTCGATCTGGCCATTGAAGTCGGCCCCCACCCGGCACTGAAGGGCCCAAC

CGAACAGACGCTCAAAGCCTCGTACGGAATTCTCCCTGTGTATACTGGCGTTCTGAAACGAGGCGAAAACAACGTC
GAAGCTTTCTCCGCGGCCGTCGGTACTACTTGGGCTCAGCTTGGGCCGGCATTTGTGGACTTTGCTGGTTATCGC
AACCTGTTCTACGAGGCTGAGACGCCTGTATTCAAAATCCCCAAGGGATTGCCTTTATACTCGTGGGACCACGACA
AGGTGTTCTGGCGCGAGGGACGTCTATCTCGCCGATTCCGTCTCGGGAATGACCAAGGCCACGAACTTCTGGGAC
GTCGCACGCTCGACGACAATGATACCGAGCTACGCTGGCGCAATGTCCTCAAGCTGGGTGAAATGCCATGGCTTC
GCGGACACGAGGTTCTTGGTGAAGTACTACTTCCTGGAGCGTCTTACGTGTCGGTCGCTGTGCAGGCTGGCCATA
ACATTGCCATGACAATGGGAAAGAATGTTCGGCTGATTGAGATTTCCAACGTCGATATCCTGCGTCCGGTTGTCGT
TCCTGAAGGCACCGATGGCGTCGAGACTCTCTTTACGACTCGCATTGTGGAGACGACCAGAGACCATATCAAAGCA
GAGTTCATGTACTACGTATGTCCCGATGAATCTCTCGGTACTATGCTCCGGACTTGCAATGGCGAAGTCATGGTCT
ATCTCGAGACACAGACGGATATCGCATCCCCAGATACGCTGCCCGCTAAAGATGCCATGCCTTTGAACCTCACCAA
GATTGACACGGAAAGGGTCTACTCTTTGTTTAAAGATATTGGCCTCAACTACTCTGGCCTTTTCCGCGGCATTTCTA
CCATTGAAAGAAAGCTTGACTATGCCAGCACTAGGAGCACCTGGGTGGATGGTCTGGATAGCGCTTACGTTGTCCA
TCCCGCCATGCTCGATGTCGCTTTCCAGACCATGTTCATTGCAAAGGCTCATCCGGCCAGTCGTCAGATCAACTCT
GCTCTCTTGCCATCCCACATCGATCGTGTCCGCATCAACCCGTCAGTGCAGTTCACACAGCCTAATGAGGTGGCTG
AGACTGCTGCCGAGTTCGAGACGTGGGCTGTCAAGCAAACGGCTAACTCGTTGGTTGGCGATCTCAACATCTACG
ATGCCGCGACTGGCAAAACCTTTCTGCAAGTGGAAGGACTTGCCGTAAACTCCGTGGGTGAGCAGGACGCCTCTT
CCGACAATTCAATGTTTTCTAGAACTCTCTGGGGTCAAGATGTCACTCTCGGCCTTCCTGATCCAGTTAGAGATCCT
GTTAAGGATGCCGAAGGGTTGCAGATGGCCCATGCTGTCGAACGCGTTGCAGTATTCTACGTCAAGAATATCCTCA
AAGAGGTCAAGAAGGAGGAGAGAGCCGACTTTCAGTGGTACCACCAGCGCATGTTTGAGGCCTTCGAGAACCACA
TCAGTATCGTCAAGAACGGTGAGCATCCCATCATCTTGCCTGAATGGTTGGATGACGAGGAGTCCATTCTTGACGA
ACTCGACTCTAAACACGGCGATACGATTGACTTTAAGCTTCTCCACGCCGTGGGCAAGGACTTGGCCCATGTCGTT
CGAGGTAACAAGCAGATGCTCGAGGTTATGACCAAGAACGACATGCTGAACCGCTTCTACATGGAAGGCTATGCCT
CTGTTCCGACGAACAAGGCGGTGGGAGATGTAATGCGGCAACTTTCTTTCAAGTTTCCTCGAGCCAAGATTCTAGA
AATCGGAGCCGGTACTGGTGGAACATCTTGGAGTGTGTTGAAGAGTATCAACGATGCGTACGAGTCGTACACTTAC
ACCGATGTCTCATCTGGGTTCTTTCATCTGGCCGAGGAGAAATTCTCCGACTTTGCCCACAAGATAATCTTTAAGGT
GCTAGACATTGAGCAGGAGCCCAAGGAGCAGGGTTTCAAAGAGCAATCCTACGACATCATCATTGCTGCCCTTGTC
CTCCACGCAACCCACGACTTGGAAAGAACAATGCGCCATGCCCGATCATTACTCAAGCCCGGTGGTTTCCTTGTCA
TGGTAGAGCTTACAGGTACCATGAGTGTCCGCGCCACTTTGGTTATGGGAGGTCTTCCAGGATGGTGGCTGGGTG
AGAATGATGGCCGACGGCTTAGCCCTCTCGTCACAGCCATCGAATGGGACAGGCTATTGCAGAATACCGGCTTCT
CTGGTGCTGATGCCGTCATTCACGACTTGGCCGACGAAGAGAAGCACTGCACTGCCCTCATCATGGGTCAGGCTG
TTGACGACGACTTTCAGCGCCTAAGATCTCCGCTGAGCACTGTCGCTGAGCTTCCTCCCCCAGATGACCCCATCCT
TGTCATTGGAGGCAAGCGATTGTCAACGTCCAAGGTGATCAGAGAGATTCAGAAGCTATTTCCAAGGAAATGGGCT
CGACACGTGCGAGTTTACAAGAGCATCGACGAAGTCAACGTGTCAGGAATTACTCCCGGTCTCGACGTTCTTTCCT
TACAGGACCTTGATGAGTCTCCATTTTCACACACCATAACGGCTCGTTGCATGACCATTCTCCAAACCCTACTAATG
AATGCTAAGAACATTCTTTGGGTTACGGGTGCCCGTGAATGTCAGAGTCCGAGAGCCAATCTCATCCACGGCATTA
TGCGAGTTGTGCCCTCTGAGCTCCCGCAGCTGAACGTTCAGGTGCTGGGACTTGAAGCCACTGCGACTCCCGTAG
CTCTCGCAAGGACCTCGGTAGAGATGTTCTTACGGCTCCGGGAGACAGGCACAGATGGAGGCAACAGTCATCGTG
AGATGCTGTGGTCCATGGAGCCGGAGCTCGACATTGTTCGTGACCGCATCATCATCCCTCGTGTGGTGCCCGACA
ATGAACTCAACGAGCTATACAATGCTTCCAAGCGCGCCATCACCAAGACCGTCGATGCAAGCAAGACTCCTGTGGA
TGTGGTACAGAAAGATGGCAGGCTGTCTCTGCAAGTGGCTGTGGCCGCCAGTAATACCGGCCGAGTTGAGTTGCT
CAAGGTTCATTACTCTCTGCGCATTCCCAGCATGAGCGGAGATGAGCTTTTCATCATGGCCGGACGTTCCGGGTCA
TCATGGGTAGCTGCTGTCACCAATGCCAACACCTCTATCGCTAGAGTTCATGGCACACACGTCATCCCCATTGATG
AGCAAGATTGCAGCCCGGAGAGGCTCTCTGCTATTGCAAATTACATCTTGTCATGGTCCATCTCGACAGTGGCCGG
CCCCCACGCCTCGGTACTTCTCTTCGAGGCTCATGATTCCCTAGCTGCATCAGTCAAAGCTCAGCTTGCGGCAGCA
GGTGGCAAGGCTTTCTTTGCTTCCACGCGGACGCATGGAACTCCGGCTAACGGCATCAAGATCCATGCTCTGGCG
TCTAAGCGCTCCATTCAGAGAGTTGTGCCACATGATGTCCAGCTCTTCATTGACTGCTCGCTGGATTCTTCCCCAG
CTGCGGCTACTCTTACGCAGAGCATGCCAAGTAATTGCATCAGCCGGCAGCTCGACGTACAACTTGTCCAAGAAGC
TCTTCGTAGTGCCCACAGGGAGTCGGTAGAGCTGCTCAAGGAGGCTTTCGCATTCCCGGGCGTCGAATCGCCTGC
TGTTCCCGTCGTCCCTGCACAGACGCTAGCCGGACGCAGCTTCATGTCTCTCAAAGCTCAAGCCTATGTCACCGAT
TGGACCACGCAGGTGATCTCGACCACTGTCCCGCCGCTCAGCCTTGAGGGTCTGTTCAGGCCTGACAAGACGTAC
TTCATGGCGGGCATGGCAGGAGGACTAGGCCTATCGATCTGTCAGTGGATGATACGCAACGGAGCCAAGCACATG
GTCATTACCAGTCGTAATCCACAGGTGAACACGGCCACCCTAGAGGAAGCTGAGCGCGTCGGAGCCACGGTCAAG
GTCCTCGCCATGGACCTGACCAACAGAGAGTCCGTCGAAAAGGTCGTGCAGCAAATTCGCGAGACTATGCCTCCC
ATCGCAGGTGTTTGCAACGCCGCCATGGTTTTGAAGGACGGCTTTTTTGTGGACATGGATGTTGATCAGTTTAACAA
TACCCTCGCCGCCAAGGTCATCGGGTCCGAAAACCTCGATTCTGTCTTCAGCAGCACTCCACTCGACTTTTTCATC
TTACTGGGCTCGGTTGCCTCCGTCATCGGAAACGTTGGACAGTCCAACTACCATGCCGCCAACTTGTTCATGGATA
GCCTTGTTCACCAACGGCGTGCGCGTGGCCTTGCAGCGTCCATTGTTCACATTGCCTATGTAACGGACGTTGGCTA
TGTTACCCGCGAGGAGCGCGACCGTCAGCTTGACTCCCACTTTCGCAAGGTCAGGCTCATGCCCACTTCCGAGAC
CGACGTTCACCACGCCTTTGCTGAGGCTGTGAAGGGTGGAAAGCCGGGGAGCACCTCGGGCTACCATGATATAAT
CATGGGCATTGAGCCGCTCAGGGAGCAGATTCCACTGGACCAGCAGCCTCTGTGGATGAAGAACCCACGCTTTGC
CCACTTTGACCAGCACGCAATCCACGCCCAGCACGAACGTGGCTCGACGGGATCTATTGATAACGTCCGAGCCTT
GGTGGAGAAGGCCGAGAAGGAAGAAGAAGCCATAGACGCAGTCATGGCAGCTTTCTGTGCGAAATTGGAATCCAT
CTTGCAGCTCACGGCCGGTTGTATCAATGTGCAACGCCCGATTACTGATTTGGGAATCGACTCGCTAGTTGCCGTG
GAAATCCGAACATGGTTCCTCAAAGAACTCGGTGCCGACGTGCCCGTCGTTAAAATTCTTGGAGGTGATACTGTCC
AGCAACTGTCTACAATCGCAACGAAGAAGCTACTGGCCAAGAATATGGAAGCCGGGGCTGAGAAGAAGTCGACAA
TTGAGAAGCCTGCCGAAGCTACAGCCCCGGTTCGTATCCCAGCCCCAACGCCTCTTGCTACTTCGCCACACATAG
CTATTGAATCGTCTCATATTGCGTCTGGTGATCCAACTCCAGGTTCAACTGATCAAAGGGGCCCCATACTGACACC
CGAGAGTGATCGGTCTTCCTTTCTCACAACTCCGGCTGCCGTGGATATTGACAGTGTGAGTGTGTTAACAGCAGCC
ACCAGCAAACAGGAGGAATCCATCTCGAGCAACAGCTCCTACACAAAGGACGATACGCCCGAGCTCGAAGATGAA
TACGAAAGCTCTGTCGGCGTCGCGGCAGAGGAGGCGGATGAGATGAGAGTTCGGCCAGGAATAATCCGTGAAGA
GCGAATGTCACCCGCCCAGGCCCGCCTCTGGTTCATGTCTCAACATCTCGAAAATTCTGCGGCGTACAACATGGCT
TTCCGTTACAAGGTCAAGGGACCTATTGGAACTGCGAGACTGAAGCACGCACTGTCGTTTACGACGCATAAGCATG
AATGTCTGAGGATGTGCTTCTATTCGCACCTTGAAGACGGCCATCCTATGCAAGGTCTCATGGCCTCCTCTCTCTG
CAACTTCAAGCACACTGTCCATGCATCCGAAGACGATGTTGCTCGGGAAATGTCGCGGCTTTCCACCTGTCAATGG
AACCTGCAGTACGGGCATACGATGGAAGTGTCACTGCTTAGCCGCGACTCCGAGGACCACGACCTGATTATCGCA
TACCATCACATTGTCATGGATGTCATTGGGCTGGGCGTTGTTCTGGACGACTTGAACAATGCTTACAATATGAAGCC
TGTGAACAAGAGTGCCGGATCATATATAGAATTCTCCGTGAAGCAGCTGGAGCAGCAAGCCCGAGGCGATTTTGAC

AAGCAGCTATCTTTCTGGGAGGCGGAGTTCAGGACAATACCTCAGAGGCTTCCGCGTCTGCCATTTGCTCGCATCG
GCAGCCCCAGCGGCGTCGCCGAGTCAGTGCCTGATACGTATCATCAATACCGCGAACTGAACGATACGCAGTTTG
CATCGCTGAAGGCGACCTGCCAAAAGCTTCGTATCAGCCCGTTCCACTTCCACCTTTCTCTTCTCCAAGTCTTGCTT
TCTCGTTATACCAACACTGATGATCTCTGCATTGGGATTGTTGACGCCAATCGCAACGATCACCGATACGCTGGTAC
TGTGGGTTGCTTCGTCAACATGGTCCCGGTTCGCCTGCATGTGCCCCCCGGCAAGGCTTCATTCGCCGACGTCGC
TCAGCAAACACGCAGGAGCGCCCTCCAGGCTCTTGAGCATTCGGCAGTGCCTTTCGACATGATACTAGACAAGATC
AAGGCTCCAAGATCCTCTAGAAGCACCCCACTGTTTCAAGCGGCACTGAACTATCGCACCGGCTCCATCTGGGAAT
TGCCCCTGGGACGGGCCAAAATGACCATGGCAGATGTTAAAGACGCCAACAATCCATATGATCTGAGCTTGGGCAT
CGCAGAGACACGCACTGGATGCATGGTAGAGGTTTATGTCTCGTCCTCTGCTTATGGTGCCGAGGCTTGCGCTAC
CATCATGGATGCGTACATGCGTCTTCTGGATGATTTCTCCAACACCCCGGATCTTCAAATCAATGACTGTAACATCT
ATGGCAAACCTGAGATTGACAGCTCATTGAGCATTGGACAGGGCCCCGTCATACACTTTGAGTGGCCGGCAACACT
TTCACAACGCTTCCTGGATATGGTCCGCCTCCACCATGGTGACCCAGCTGTGACCGATAAGACGGGCACCTTGAC
GTACTCCGAGCTTCTAGAAAGAGTCAACGGTGTGTCCGACACACTGATCCGCCATGGATGCAAGTCAGGCACCCG
TATCGCCGTGCTTTGCGAACCATCCATCGACACCATCGTGGCCATGCTCGCCATCTTACACATTGGTGCCGTTTAC
GTTCCTCTGGATGTTAGCTTGCCGACGGCGCGTCATGTTGCTATGATACAGTCCTCCAAGCCATCGTTCGTTGTGA
ACCACGCTTCTACAGACGGTCAGGCGCGGGAATGCATCGGCAAGATCGAATTCCCCATCCGGCGAGTCGTCCTTG
ATGACGTTATAGAGGAAGAGATTAGAGTGACGGTGCCCTGTTTGGCGTCACCGGATGCTTGCAGCATCGTCCTCTT
CACCAGCGGCTCGACGGGCACGCCAAAGGGCATAATGCTCTCGCAGGCCAACTTCGCCAACCATCTAGCTCTCAA
GACGCACTTGCTCGGATTGGGCAAAGAGACCGTTTTGCAACAGAGCTCCACGGGCTTTGACATGTCCATTGTCCAG
ATGTTCTGCGCTCTCGCGAACGGTGGTCGCTTGGTAATCGCCCCCTTTGACATCAGGCGCGATCCAATCGAAATG
GTGTCACTTGTGTGCAGCGAGCATATATCCCTCTCTATCGCTACGCCATCCGAATATCTCGCCTGGACCCGCTATG
GGATTTCCTCGTTGAAGGAGAACACGGCGTGGCGACATATTTGTATGGGTGGTGAGTCAGTCACTCGTCAGCTGG
TCTTGGAGCTTCGTCGTCTTGGCCTTCCAAATTTGAGAGTCACAAATTGTTACGGACCAACGGAGATCACGGCCGC
AGCTTCATTCCAACCCATCGACCTCGGGGGTCAAGAAGATGCTCATCCAGACATGGTCAAGTACAAGGTTGGCAAG
GTGCTCCCCAACTACACTGTGTCCATCCTGGATGCAAGGGGCAGCCCTCAGCCGGTCAACCATACGGGAGAAGTC
TGTATCGGAGGGGCGGGAGTGGCCTTGGGTTACATCAGCGCATCCGATGACGCCGCCAGCAAGTTCATCGTGACT
GCGCAGGGACAAAGGATGTATCGTACTGGAGATCGTGGGCGCCTCTTGTCTGACGGGACATTGCTGCTATTTGGC
CGCATTGAGGGAGATTCTCAGATCAAACTTCGGGGCATCCGGGTCGATTTGCAGGAAGTAGAATTGGCCGTACTTG
AGGCTGCAGACGGCCTCCTGTCTAATGTGGTTGTCTCTCAGCGGGGAGATGTGCTCATTGCCCATGCCACAGTCC
CTCCAGGCGAACCTGAAGCCACATTGATCACTGAAGATGATCTGAATCGCGTGCTGAGGCGACTGGATCTTCCCCA
ATACTTTATTCCAGCTCGTATCGTCATCTTGTCATCCCTGCCCACTAATGCGAACGGGAAGTTGGATCGCAAAGCAA
TCGCAGCCCTTCCGCTTCCATCATGTCAATCAAGGGCTGATTCAGGTTCAGAAGAAGAGAAGATGACGATCGAAGA
AGGCGAGTTGCGCCTTTTGTGGGAACGTGTTCTTCCCCAGGTCGCCCCCTTGACACGCATCGCGCCGTCGTCGGA
CTTTTTCCTCTTGGGAGGCAATTCACTTTTACTCATGAAGCTTCAGGCGGCGATGAAGGACTCGATGAATGTCATGA
TGTCAACCCGAAAGCTATACCAGGCGAGCACGTTACGGGAGATGGCTCGGGCGGTCGACAAGCAGCGGCGGTCG
CAAGCTGAACATGATGACCAAAGAGAGATTGACTGGGCAGCAGAGACAGCTATTCCTAAGTGGCTCTTGAACCAGA
TTCACGAACAATCTCAGTCAAAGCAGGTCGTGGGCGCAAAGTCACCAGTGGAGAAGATCAGCGTGTTGATGACCG
GTGCAACAACATTCCTGGGAGGTCACCTGTTGAAGTCTCTCCTTGAGTCGGACAAGGTGAGCAAGATTTACTGCAT
CGCCGTTCTCGCCGACGATCAGCATCTACTCCCTGAAAACGAAAAGATTGAATCCTTCAATGGAAGTCTGCTTTCAC
CCACTCTGGGGCTCGATGCAACTGAGCGAGTGCGATTAGAGAGCACGGCAGATATCATCGTTCACGCAGGAGGTA
GCGGTCACTGCCTCAATACATACGCAACATTGCGCATCCCCAATGTTATCTCTACCCATTTTCTAGCCTCAATGGCG
CTGCCGCGCTCTATTCCACTGCTGTTCCTCTCATCCAGCCGGGTCGTTCTTCTGACTGGAAACACTTCGCCCCCGG
CGGCTTCCGTCAGCGGTTTCCCTCCGGCAACAGATGGGCGGGAAGGCCATATGATGAGCAAGTGGACAAGCGAG
GTATTCTTGGAGAACCTTGTCGGACACCTACTGGCCTCATCACCCCGACATCAACAGAACCCTTGGACTGTCTCAG
TGCATCGCCCTTCGGTCATTGTCAGCGAACACGCTCCCAATTCTGATGCCCTGAATGCGATCCTTCGGTACTCTATT
TTGATGCAGTCATCGCCCCGAATGGACAATGTTGTGGGCTATCTTGACCTGGCCCAGCTGGATACCGTTATTGCGG
AGTTGCGCCAGTCGGTGATCCAGCTTGGCTCGGGCCAGGGCAAACAGGAATCTACGAATATTCTTTTTAAGCACCA
TTCCGGTGGGATCAAGGTTCCGGCGAGTGAACTTTGGTCCCATCTGGGAAAGGTATATGGCATCACTTTTGATGAG
GTTGAGCTGGATGAGTGGCTGCGCCGTGCGGCCAGGGCGGGGATCGATCCCTTGATTACGGCGTACCTGGAGGC
TATTCAGGCCAACGGGGCAAAGATGATTTTCCCTTATATGGGAAGCGATTAA.

ER (TRIREDRAFT_58289):

ATGACTGTCATTGATGTGCTGCCACAGACACAGCGGGCATTGAAGATTGTCGGGCCCAATGCCGTCTCTGTGAAC
GCAGCCGCCCCTCTGCCAGACATCGAGCCAACAGACGTCCTTGTTCGGGTCGTTTGCGTCTCCATCAACCCGGTC
GATGGCAAGGCAGCAGACATGTCACCACAACTTGGAGCTACGAGCGGCACTGACTTCTCTGGTGTTGTCGTGGCT
CTTGGCGCGGATGTCGAGGCAGATAATTGGAGAGAAGCAAACACCATGAAGCCCGTCAGGATCGGAGACCGTGTG
TTTGGAGGAATCTTCGGCAATGACCCACTGAGACCTCATAATGGCGCGTTCGCGGACTATGTTGCGGTGCCAGCA
CGCCTTGTCTGGCACATTCCGACAGGCACAGACTTTGCAACCGCCGCTACCATGGGCGCCGCAATTGCTACGGTT
GGGTTGTCACTCTTCAACTACTTGGGACTTCCGCTACCATCTAAGTCCAAGGCTGGGCTGCCAGTCGTCACGACAT
GCTCAGCTGCCTCTTCCACAAACGTTTTGCAGTTGGGAGCAGAAGCCTGGTTTGACTACAAGTCACCCACCTGCGG
CGCAGACATCCGTGAGCACACCAATGACTCTCTCGCCTTCGCGTTGGACTGTATCACCGACACTGCATCTATGGGG
ATTTGCTATGAAGCACTCGGCTCTGCTGGCGGTCGTTATGTGGCGTTAGATGCCTTTCCGGTTCGTGGCCATACTC
GCCGCAGCGTGGTACCTGAATGGGTGTGCACACCTACACAGTTCGGCAAAGCTATTCGCTGGACGCCCCCCTACG
ACCTCGAGCCACGCCCGTATGATCTCAAATGCGCTGAGCTCTGGTATGTGGTTGCGCAACGACTGATTGATGAGG
GTCTCATTGCATCTCACCCTTTAGAGAAGAGAAACGGAGGACTGTCTGCGGTTCCTGAGGGTATGGAAGAGGTGC
GAAGAGGTCAGATCAAGGGAAAGAAATTAGTCTATACCATTCTTGACAGCGAACCCATTGCTGTATCGGCTTGA.

P450 (TRIREDRAFT_58953):

ATGTTGTTGCCCGAGTACGTCAGCGACTTGGGCGGCTCACATGCGGTCACATTAGCGGCCTTAGGCGTGAGCGTC
CTGGCTGTTCTGTTCACAAGGCTGTTCAACGAGAGCAGGAAGGGTCACGAAGCGGGCTGCAAGCCTGTGACCCGC
GTCAGGCAATGGGAACCTTTCTTTGGGCTCGACATGGTGTATAGCCAGCTACAAGCTCTAAAGCGGAACTACTATC
TGGATTGGCTGAGGAATCTGCACAAGGGCAAGCCAAAAACATTTGCCATTACCTTCATGGGAGTCAAGCAGATTTG
TACCATTGAGGGCGAAAACCTCAAGGCAATTCAAGCAACCAACTTCAAAGACTTTGGGCTGGAGCCCATGCGCCG
GAAGACGAAAGGTGCGATGCCCTTTGCCGACAAGGGCATCAGCACAACGGATGGAAAGAACTGGGAGTTTGCTCG
ATTTCTCGTGAAGCCTTTCTTCTACAGAGAAGTTTACGCTAGCATTGACCGAGTCGACCCTTACGTCCGTAAGCTAT
TCGGCCTCCTGCCAGAAGAAGATGGCGTGACATTTGACATCCAGCCTCTGATTCAGCGATGGGTAGGTCTCATTTC
TGCTCGCCTCGCCCCCTGCCGTCGTCAGATGTGGCATTCCCGATGGCTGTCGGGGTGTTTGAATGAAAAGTGTCT
AACCATGGCCTAGTTCCTTGATTTAACGAGCGAGTTTATCTTTGGCAAGACTATGGATTCTATGACTTATCCCGATC
GTGCCAACATCACATGGACTATGCTAGACGTTCTACGAGGCGGTCGACTTCGCATCCAAATGTACAAGTTTCTGTG
GGCATTCAACTGGAACTGGTGGCTGAAGGCCGTTTACGAAGTCCATGACTTTGTCAACGTACACATCCGCAGCACA
TATAAAGAACTCGCAGCGCGCGAACAGCGTATCAGGGACGGGCTGCCTGTGGGCCCAGAGCGTGTTGATCTCCTG
TGGTATATGGCTACCCACGTCAGAGATGAGGAGGAACTGAGGTCTCAGCTGTGTCTGGTTTTTGTTCCAAACAACG
ACACCACGTCCATCTTCATTAGCAACTGCATATGGCACCTAGCTCGTGATCATGAAGCATGGCAGAAGCTGCGCAA
GGAAGTTCTCGACTACGGCGATCAGCCCATCACATTCGAGTCCCTAAGAAGCATGCCCTATCTGAACGGCGTGCTC
AATGAGAGTAAGATATCCCTTTTTATTTCTACCAAGGCGGAACGCCAAGTCCTGACCTCTCTCTTCTAGCACATCGG
TTGACTCCCAACAACATCGTCCAAGTGCGTGCTTGCCTCAACGATTCCGTCCTGCCTCTGGGTGGGGGTCCTGATG
GCAAGTCGCCCTTATACGTGAACAAAGGCGACCTCGTCTCAGTAACCAAGACCGTCATGTATCGCGACCCAGATAT
CTGGGGTCCCGACGTCGATGTATTCAGACCGGAGCGTTTTTTTGGTGTCCGCGGCAACTGGAACTTTCTGCCCTTT
GGGGGCGGTCCTCGTCGGTGCCCTGCGCAGATGATGGTCCAGACAGAGTCGGCCTATATGCTGTTCCAACTGGC
AAAGAGGTACTCGCGACTTGAGTGTCGTGACCCGGAGCCTTACACCGCCGTGATGCGGATCGGGCCGTCCAATAT
TAACGGCGTCAAGATAGCTTTCTACAAGTGA.

C-MeT:

ATGTCTTCTACTCAAACAACGGCGGCAGAGCCCATCATGACCGACAATGTGGCCCTGCGAGCCTACTACGAATCGT
GGGATTCACGCGTCGTTTATCAAATTATTATGGGCGGAACACAGCATTTCGGCTACTGGGATAAAGACACCTACTG
GCCATTTCCGCTGGGGTCGAAACTTCGTCGGTATGCATCCAGTTCAATAGCCTCTGTGGGCAGAAGCTCCGAATAG
CTCATACTGATTCCATCGCAAATAGCATGGAGCAAAAGTTGATGGAGATCCTGGCTCTCCCAAAGGGCTCCCGTGT
GCTAGATGCGGGTTGTGGAGTTGGCCACGTTGCGCGGTACATGGCTCAACATGGCATGCGCGTCTTTGGCATTGA
TATCATCGACTGGGCGATTGAAGACGCCCGGAAAGCGGCAAAAGATGCTGGTTTGTCAAAGGAGATGATGTCGGT
GGAGAAGATGGACTACCATCATCTCGACAGTCTCGCGAGTGAGTCCTTCGACGGGGTATACACCATGCAGGCGTT
TGGCCACGCTGTCGATCCTCAGAAAGCGATGGCCGGGTTTTTCAGAGTCGTACGGCCGGGCGGCAGAATTGCCAT
GGTCGAGGTCGAGCGCAAGACAGCTGCGAAGCATGACGATCCCAACGACAGACTCACGCAGGAGTTAAAAATGGT
CAATGACTATACTGTCATGCCAACAAATGAGGCCGCAAGCGAAGACTACTTCAAAAATCTGCTGGAAGAAGCCGGT
TTCGTCGACGTTGTAGTACGCGACTGGCAGCCGAATATTCTGCCTATTTTGCGACTCTTTTACTCGCTAGTTATGAT
TCCATATCTATTCTTTAGGTTGTTTGGGAACGAAAAGAGCTTTATCAACATGATCTGTGCCAGAAGCGGATATGCTG
GCAGATCACGATGGAGGTTTGTTGCCATTACTGCGACAAAGGCCGGGGAAAAGCTCGAAGACCACAAAAGCAAGT
AG.

7.2.2 NMR data of compounds isolated in chapter 3

7.2.2.1 Compound 137

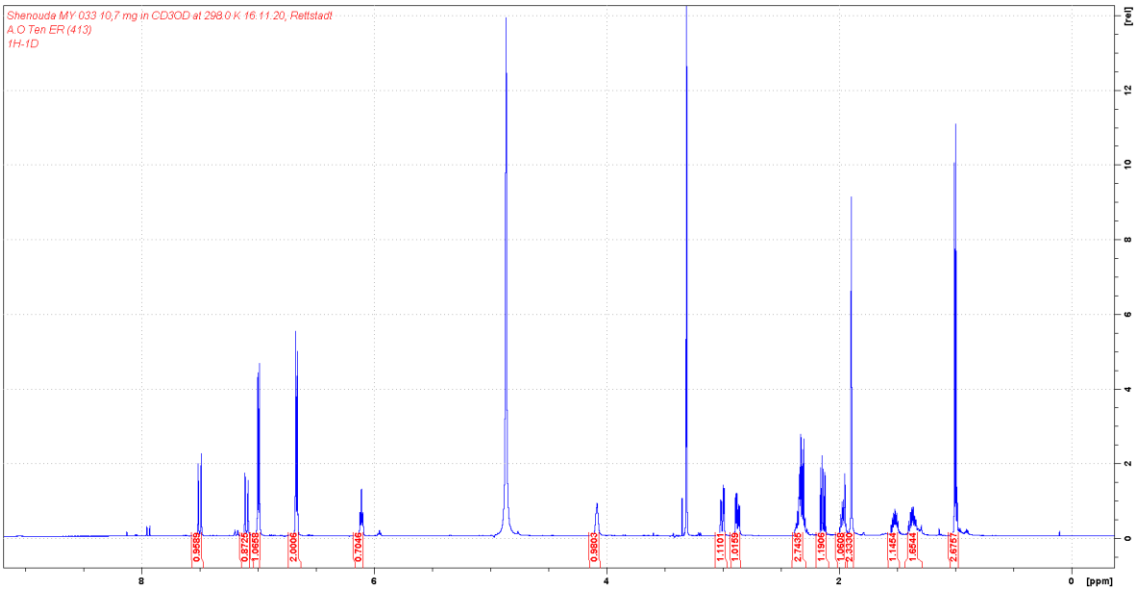


Figure 7.30 ^{1}H-NMR spectrum of compound **137** (500 MHz, CD_3OD).

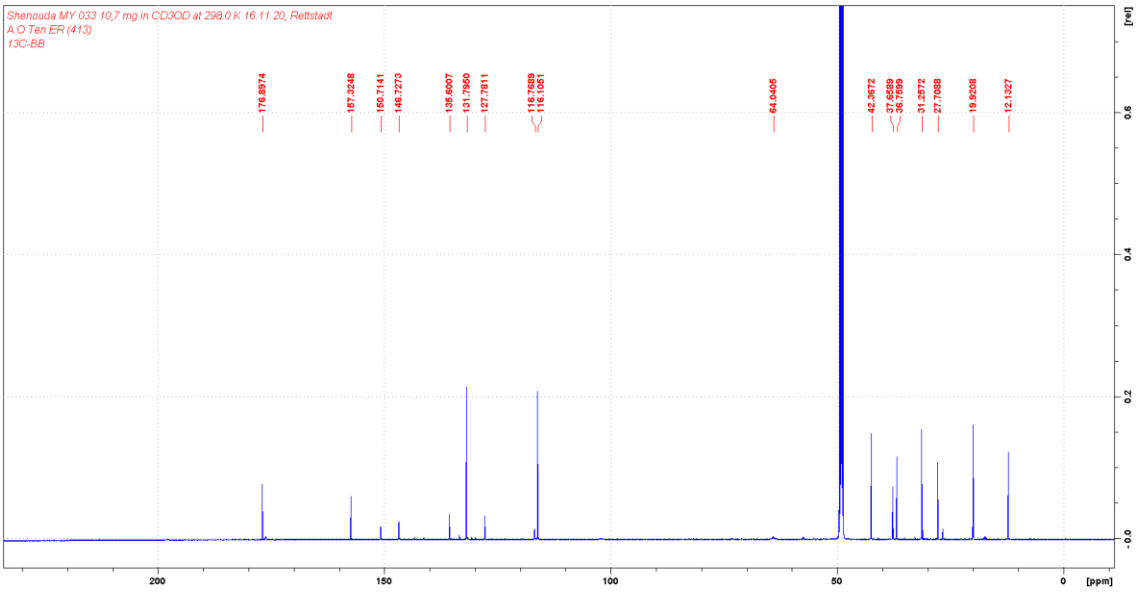


Figure 7.31 ^{13}C-NMR spectrum of compound **137** (125 MHz, CD_3OD).

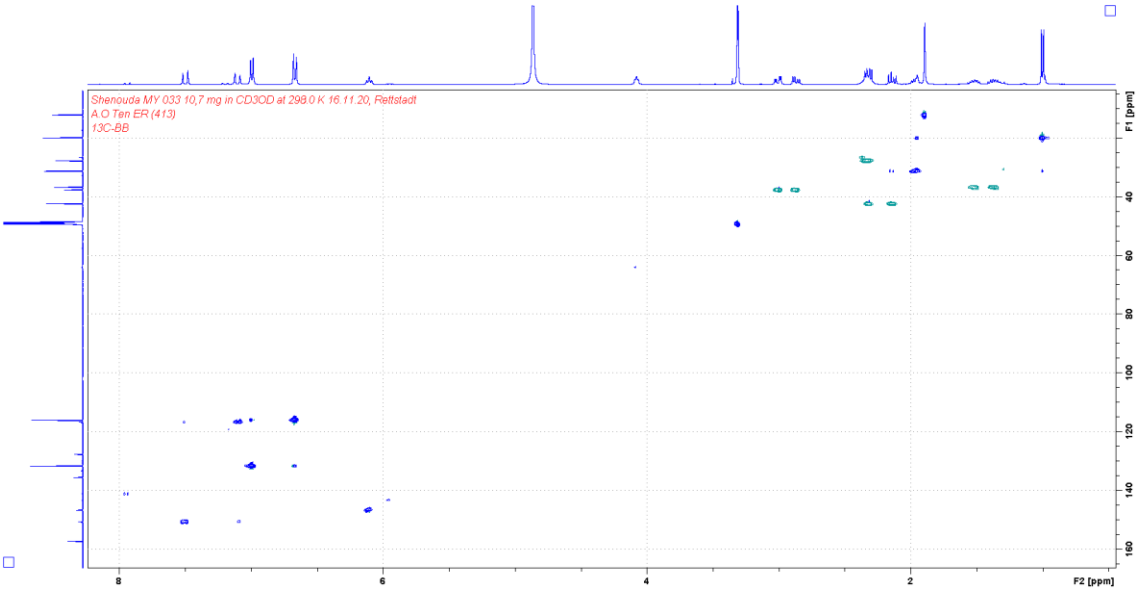


Figure 7.32 HSQC spectrum of compound **137**.

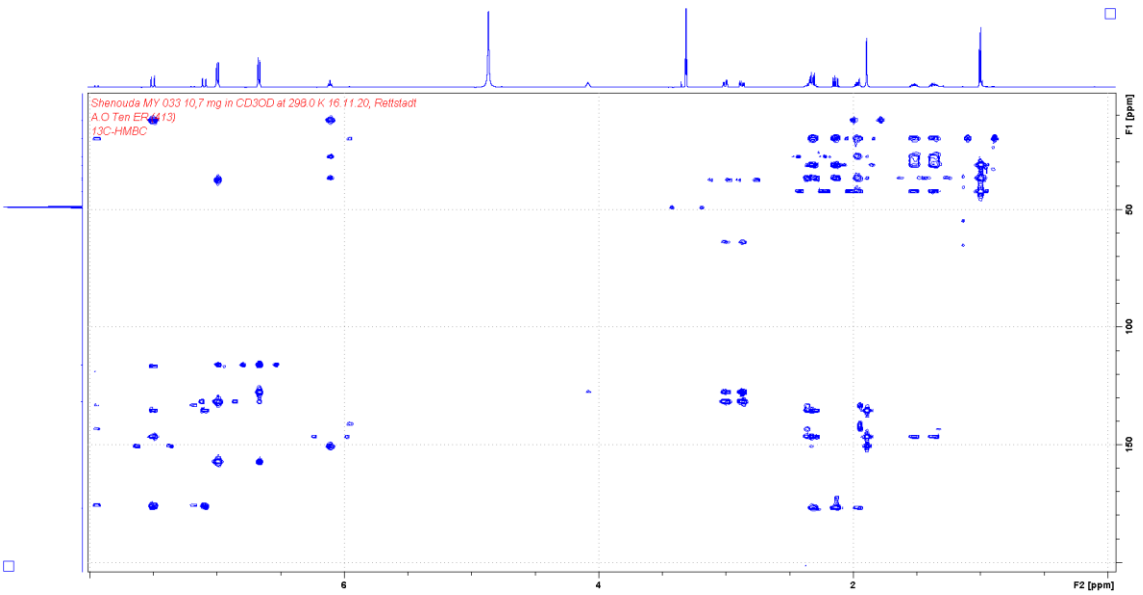


Figure 7.33 HMBC spectrum of compound **137**.

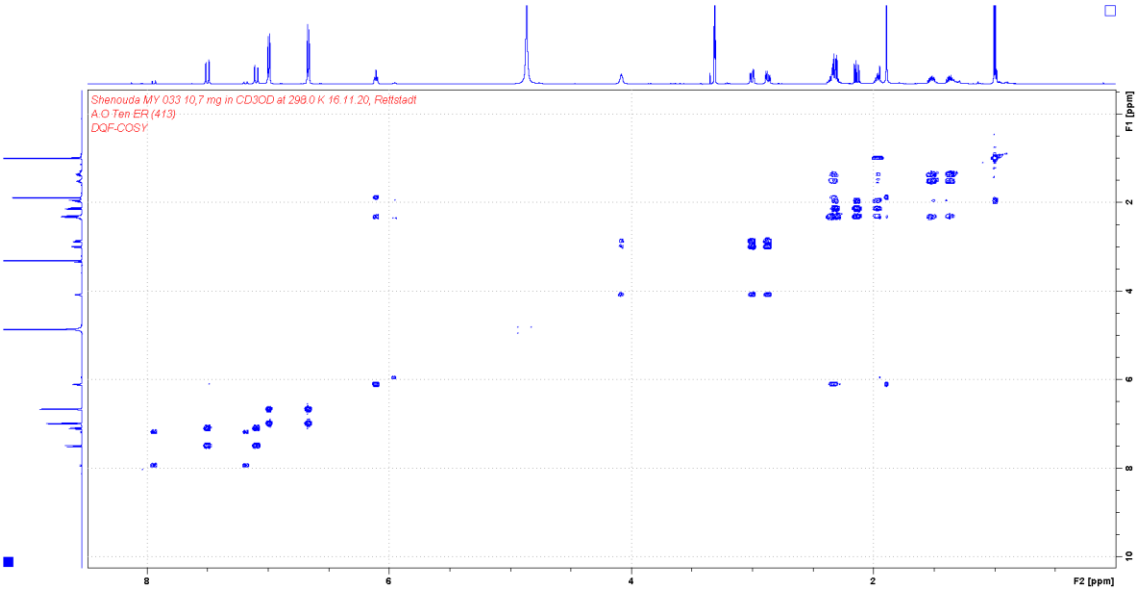


Figure 7.34 COSY spectrum of compound **137**.

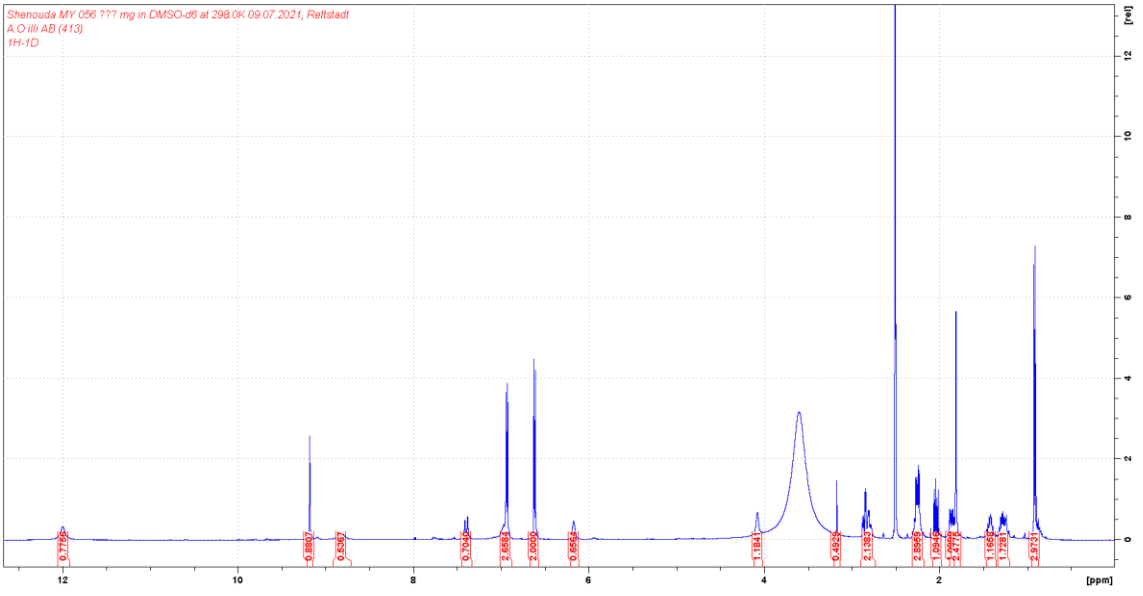


Figure 7.35 ^{1}H-NMR spectrum of compound **137** (500 MHz, DMSO-d6).

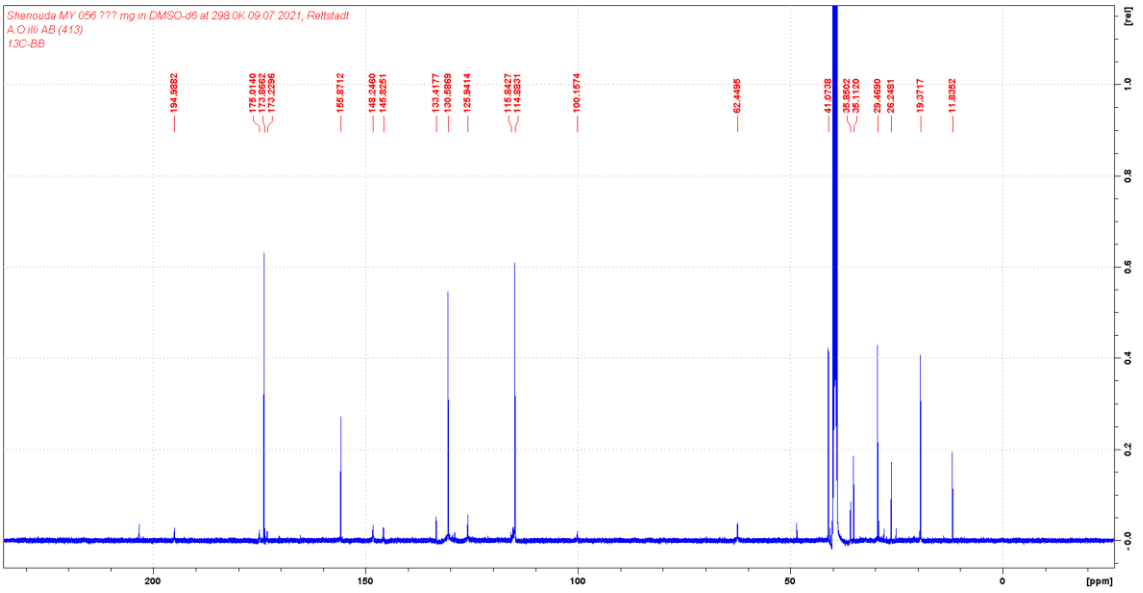


Figure 7.36 ^{13}C-NMR spectrum of compound **137** (125 MHz, DMSO-d6).

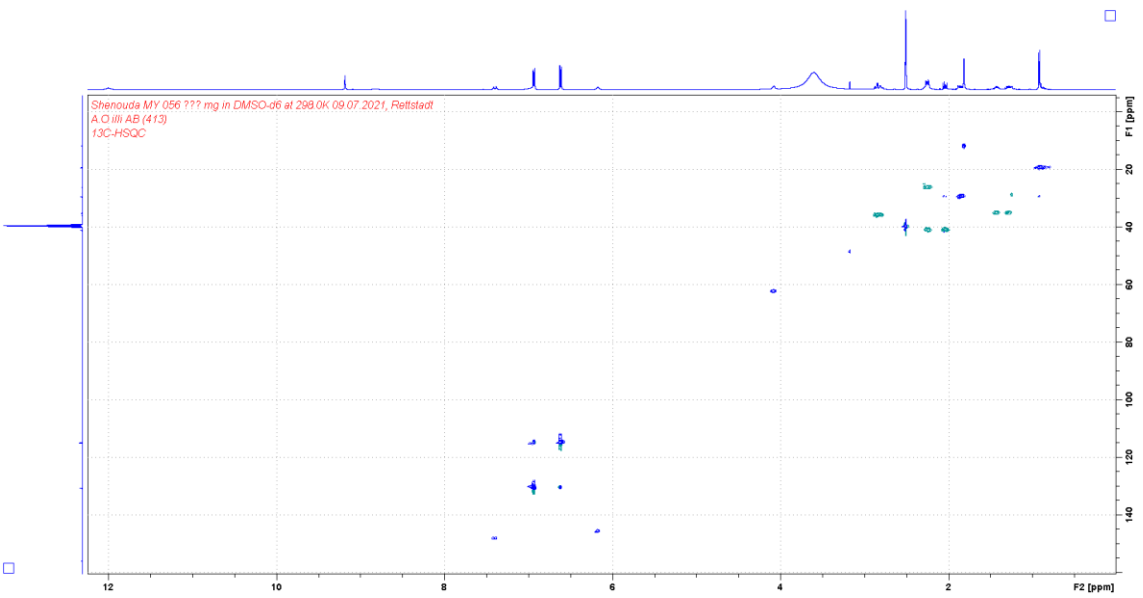


Figure 7.37 HSQC spectrum of compound **137** (DMSO-d6).

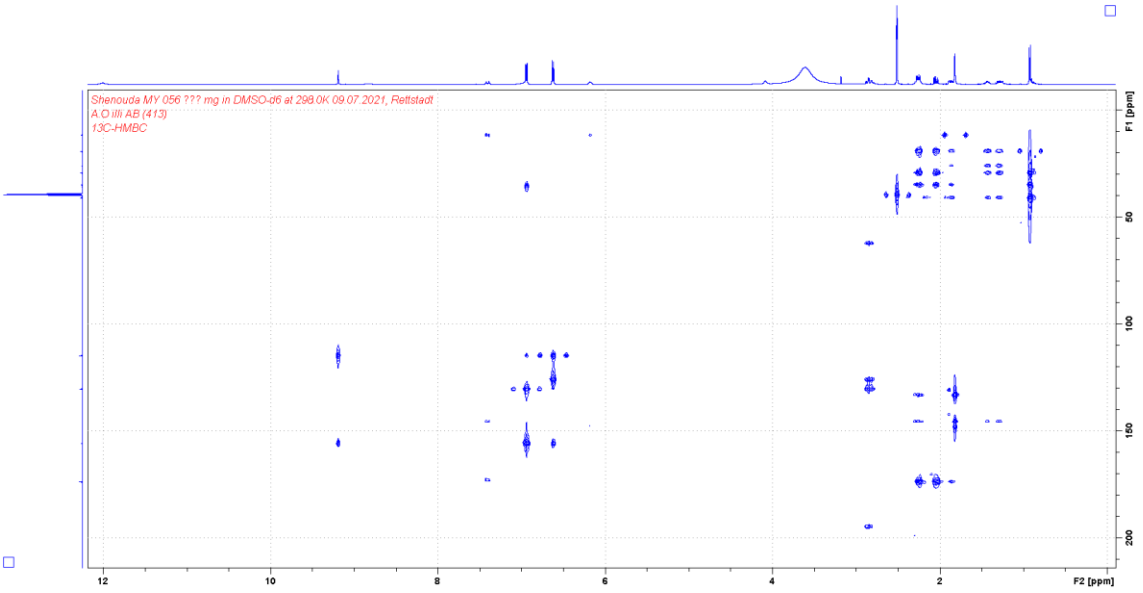


Figure 7.38 HMBC spectrum of compound **137** (DMSO-d6).

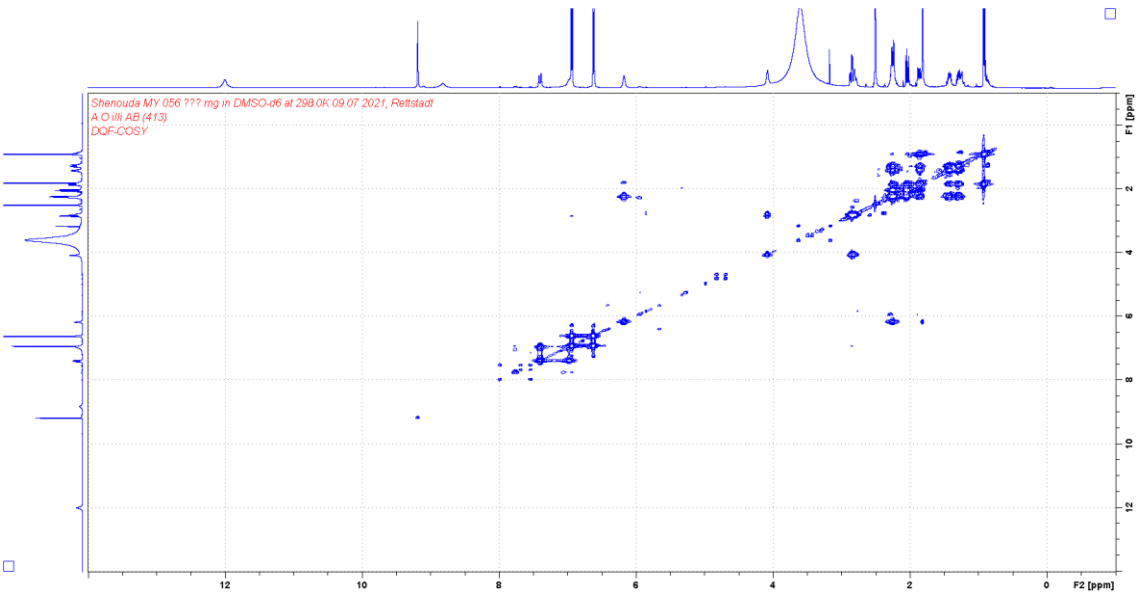


Figure 7.39 COSY spectrum of compound **137** (DMSO-d6).

7.2.2.2 Compound 138

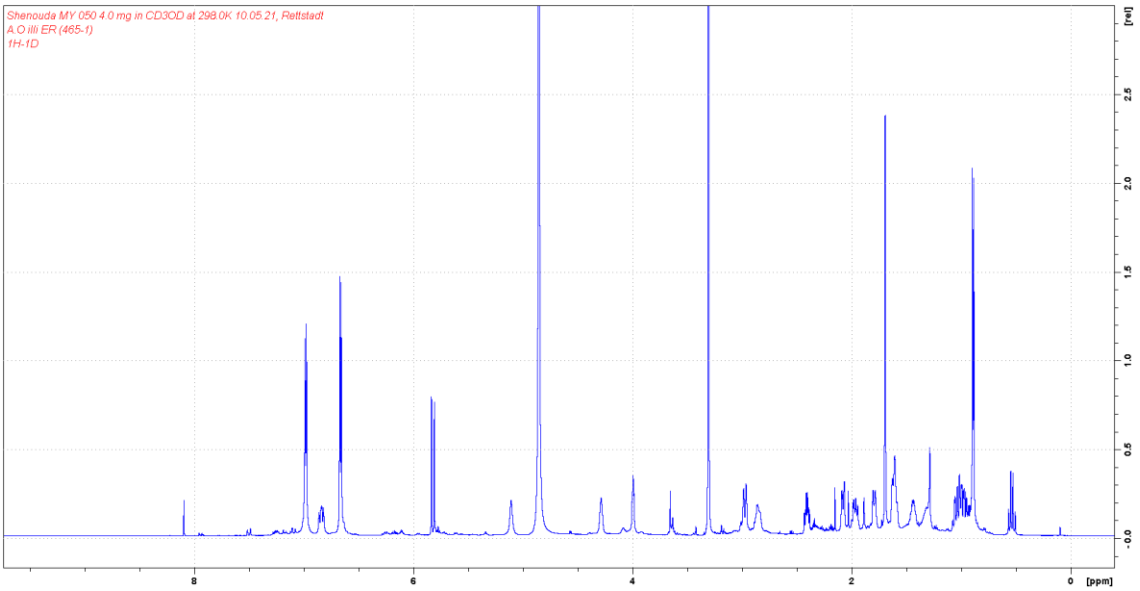


Figure 7.40 ^{1}H-NMR spectrum of compound **138** (500 MHz, CD_3OD).

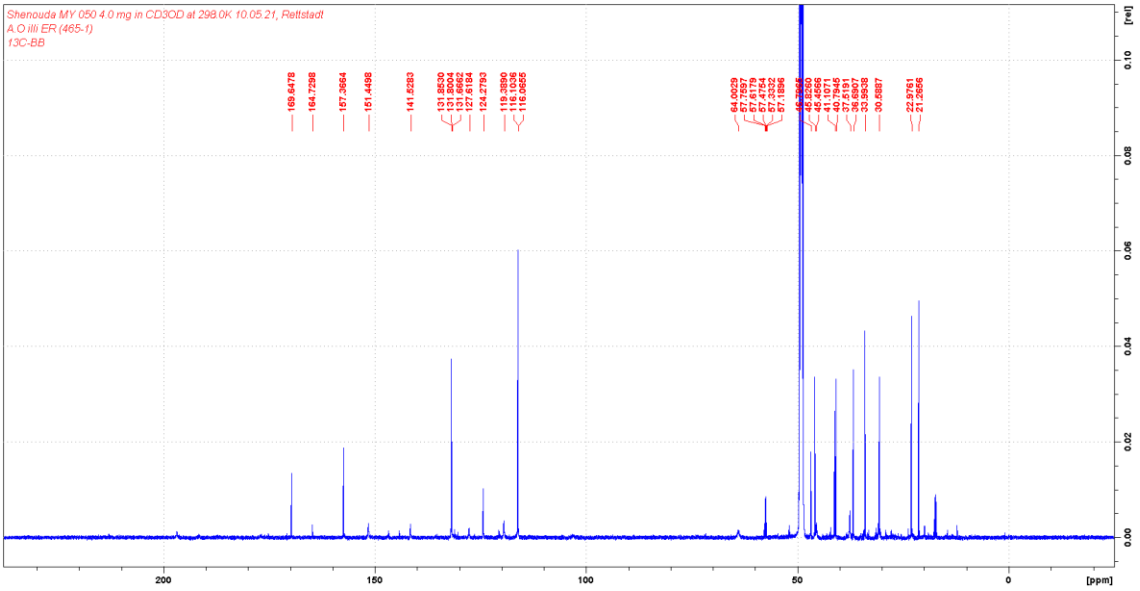


Figure 7.41 ^{13}C-NMR spectrum of compound **138** (125 MHz, CD_3OD).

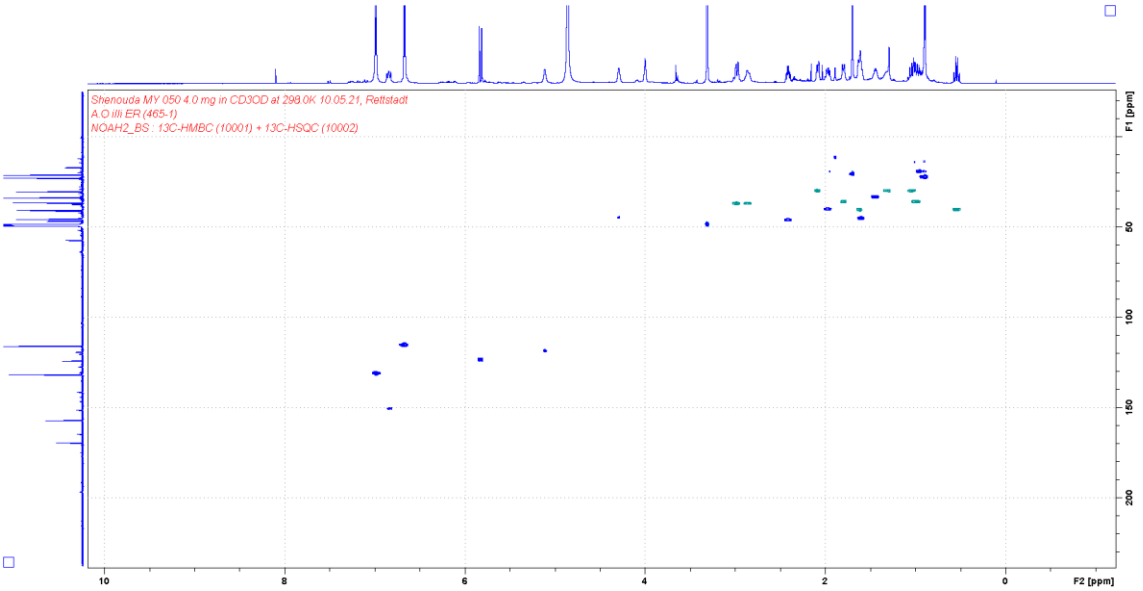


Figure 7.42 HSQC spectrum of compound **138**.

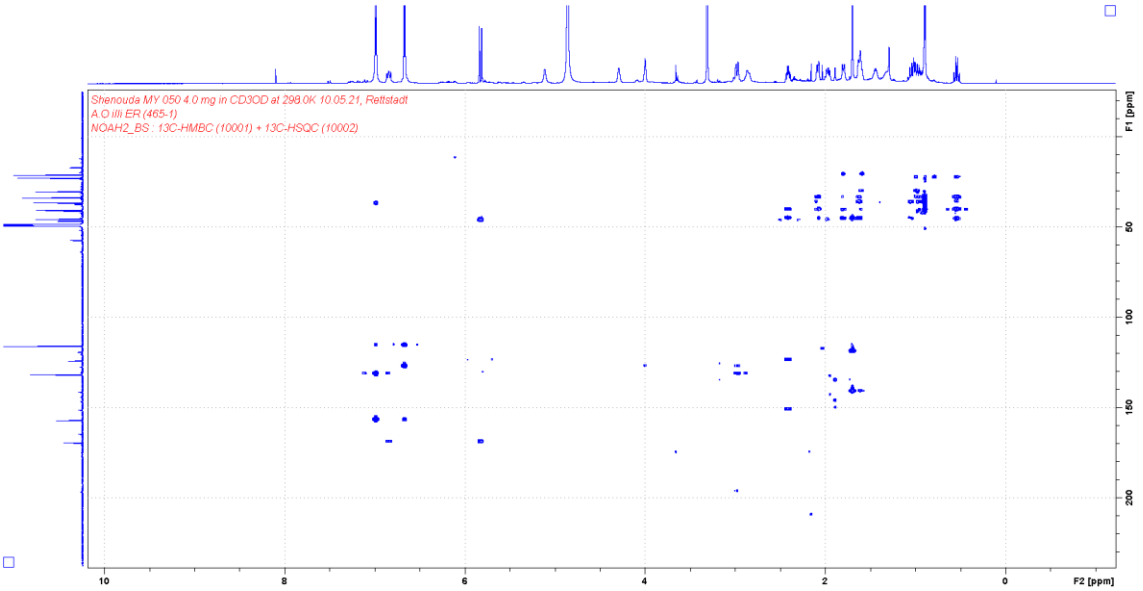


Figure 7.43 HMBC spectrum of compound **138**.

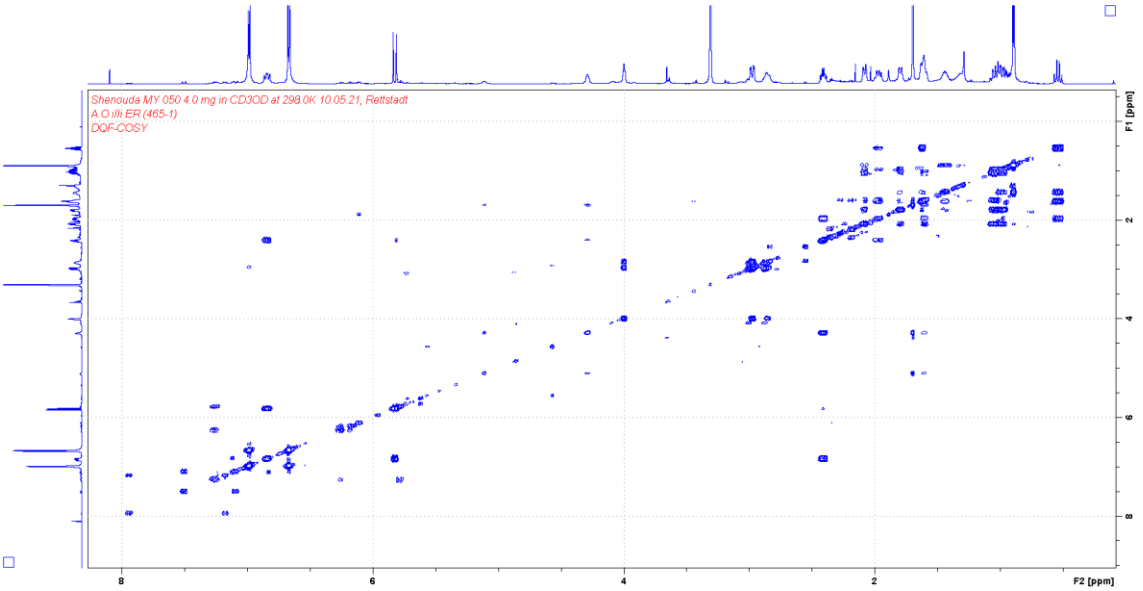


Figure 7.44 COSY spectrum of compound **138**.

7.2.2.3 Compound 140

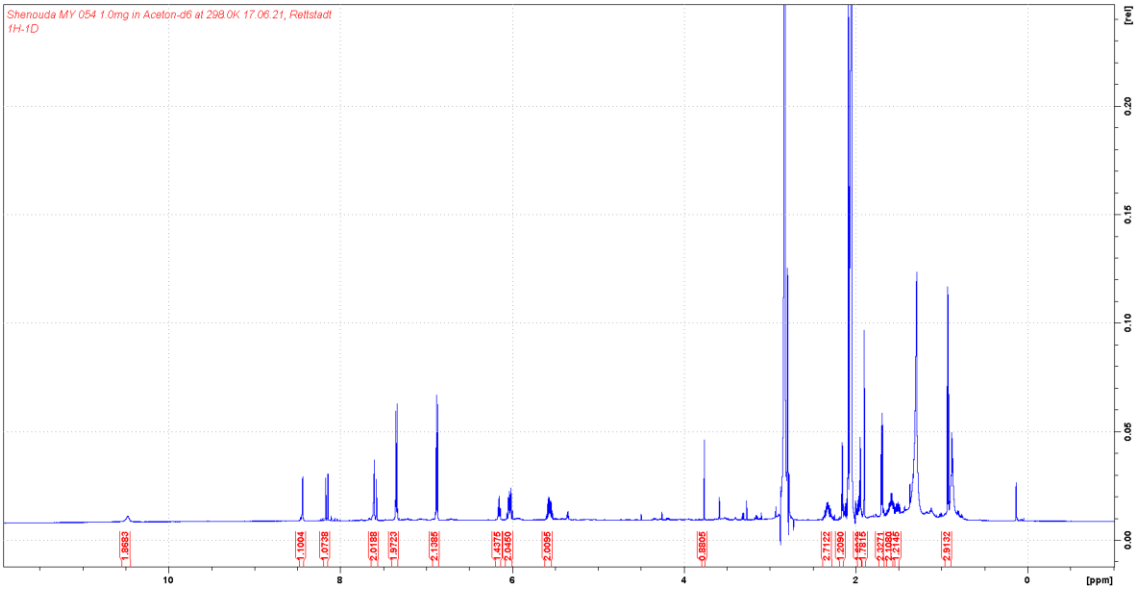


Figure 7.45 ^{1}H-NMR spectrum of compound **140** (500 MHz, Acetone-d6).

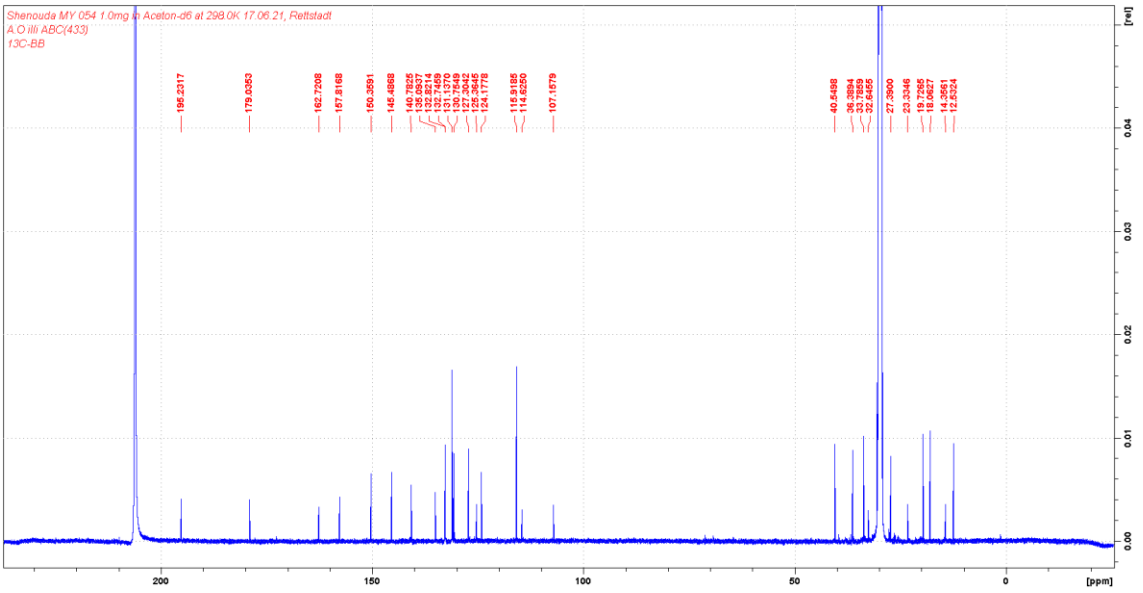


Figure 7.46 ^{13}C-NMR spectrum of compound **140** (125 MHz, Acetone-d6).

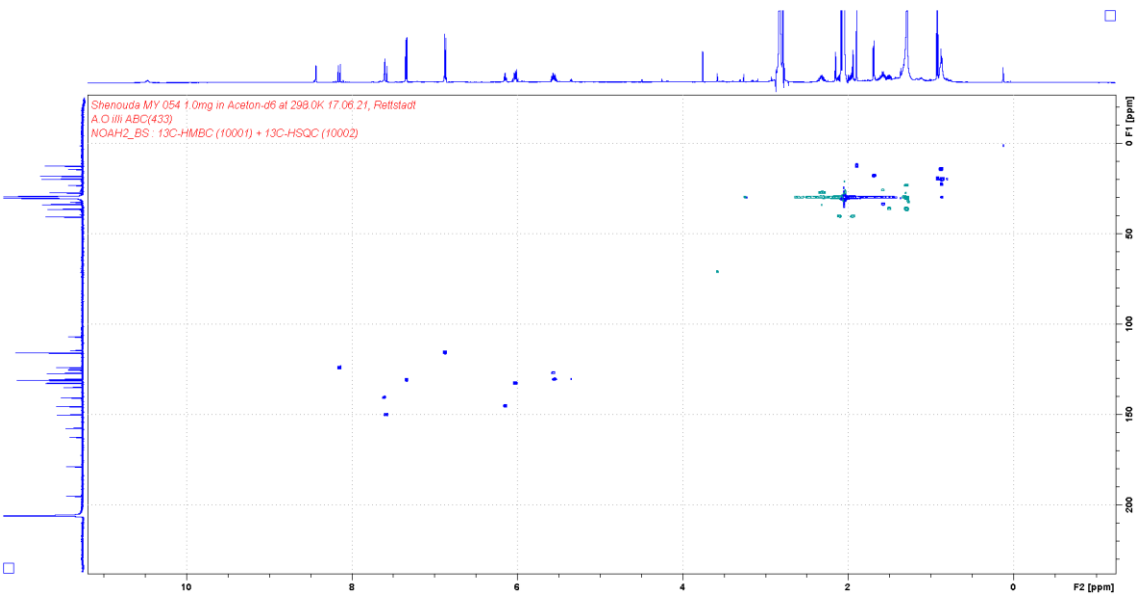


Figure 7.47 HSQC spectrum of compound **140** (Acetone-d6).

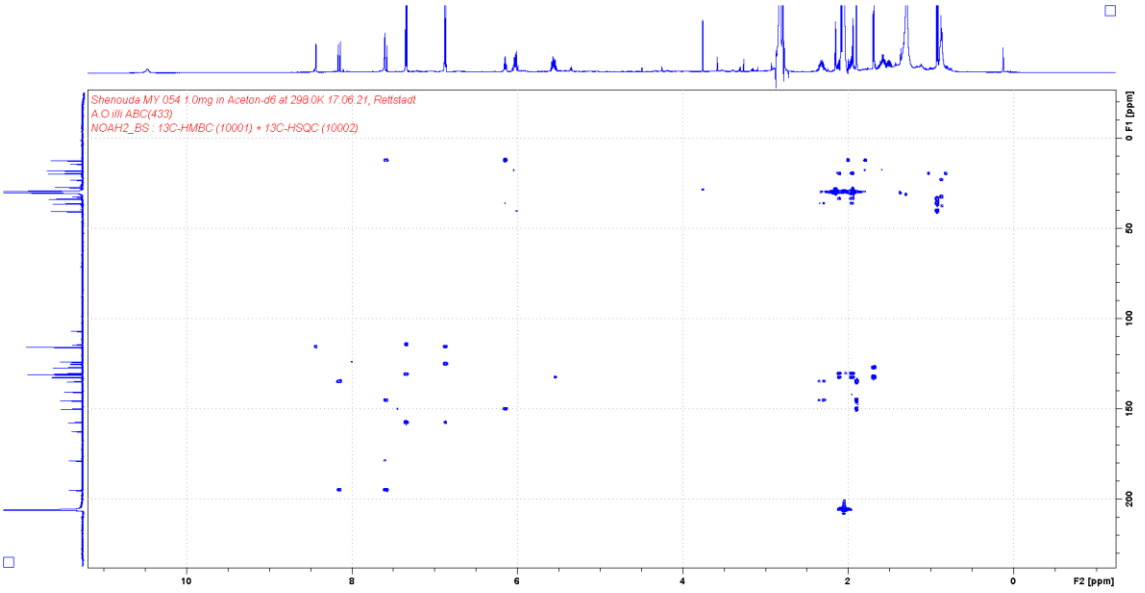


Figure 7.48 HMBC spectrum of compound **140** (Acetone-d6).

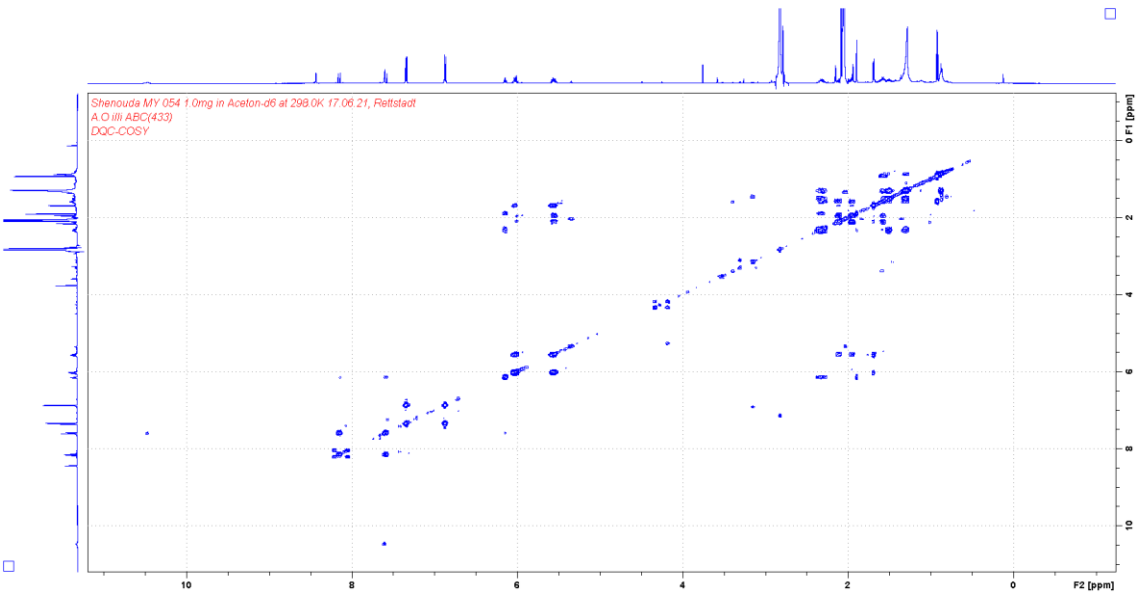


Figure 7.49 COSY spectrum of compound **140** (Acetone-d6).

7.2.2.4 Compound 141

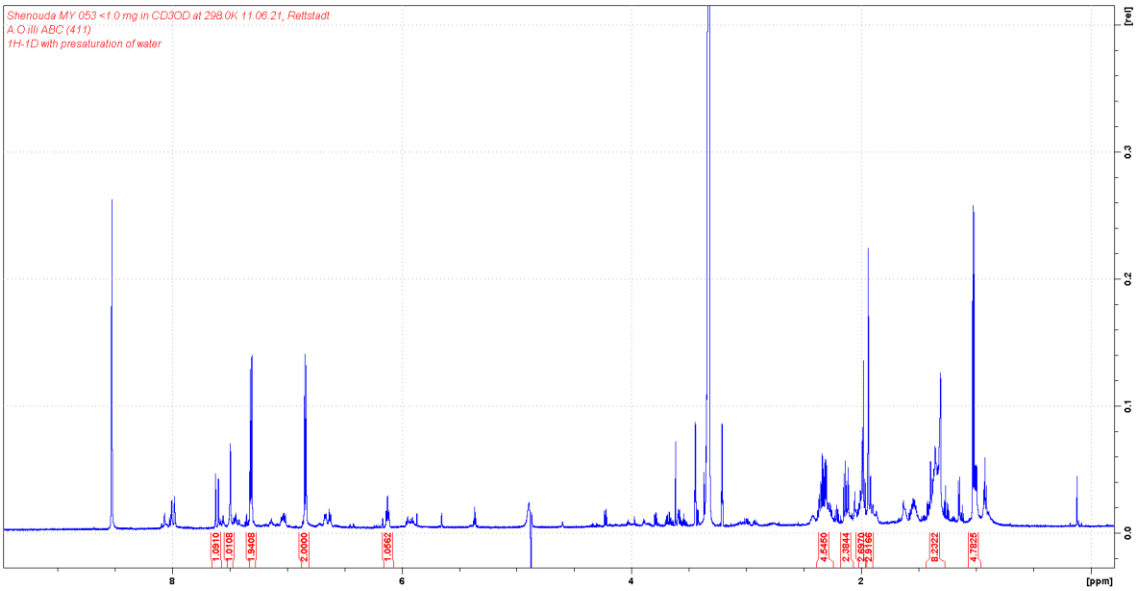


Figure 7.50 ^{1}H-NMR spectrum of compound **141** (500 MHz, CD_3OD).

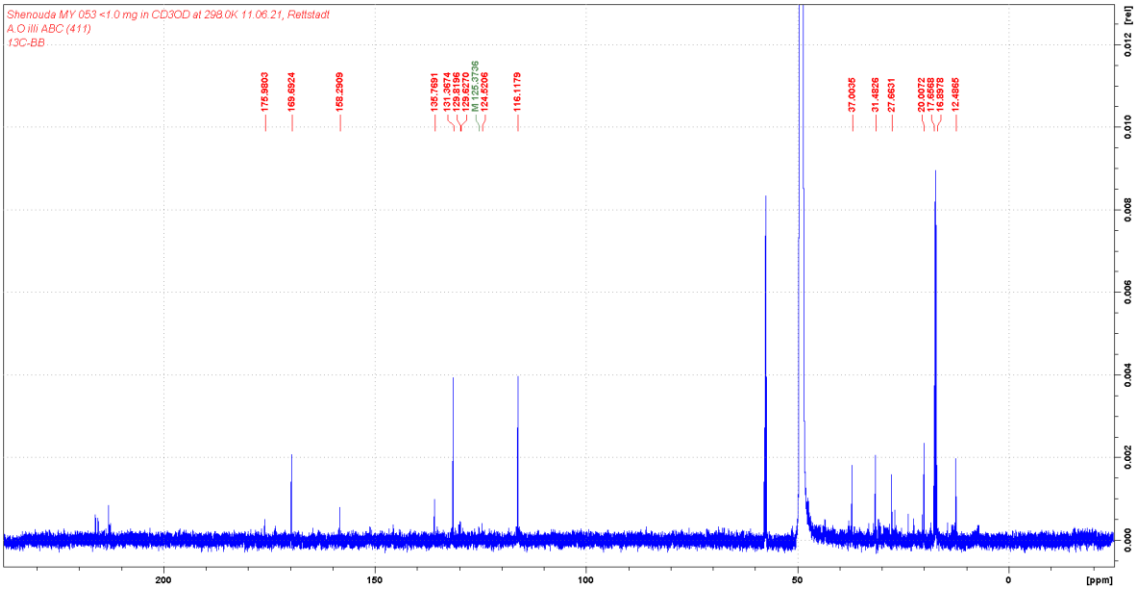


Figure 7.51 ^{13}C-NMR spectrum of compound **141** (125 MHz, CD_3OD).

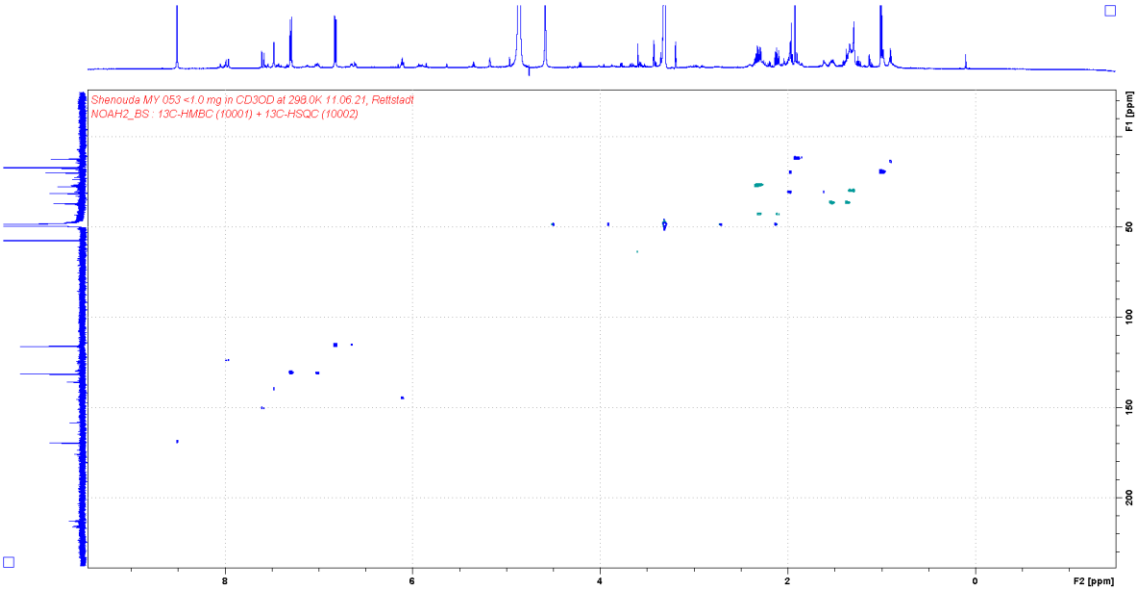


Figure 7.52 HSQC spectrum of compound **141**.

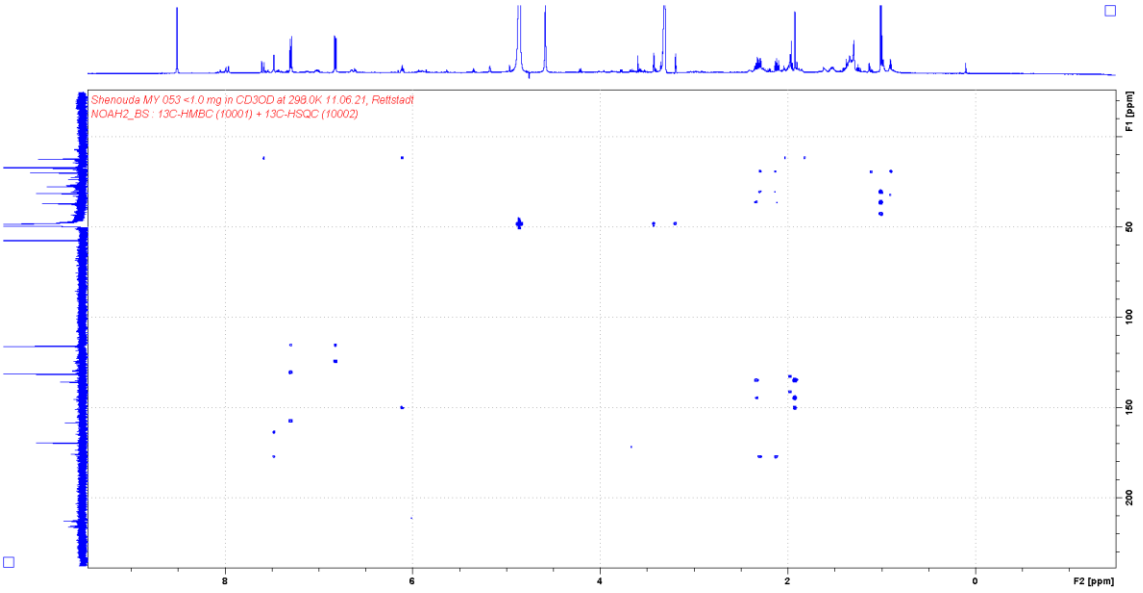


Figure 7.53 HMBC spectrum of compound **141**.

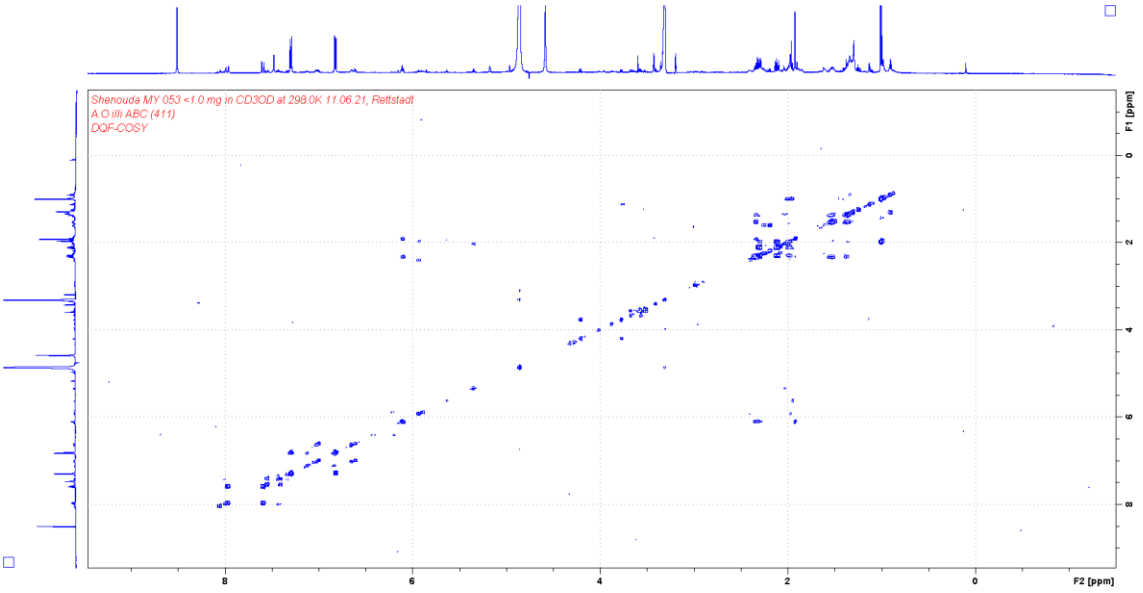


Figure 7.54 COSY spectrum of compound **141**.

7.2.2.5 Compound 142

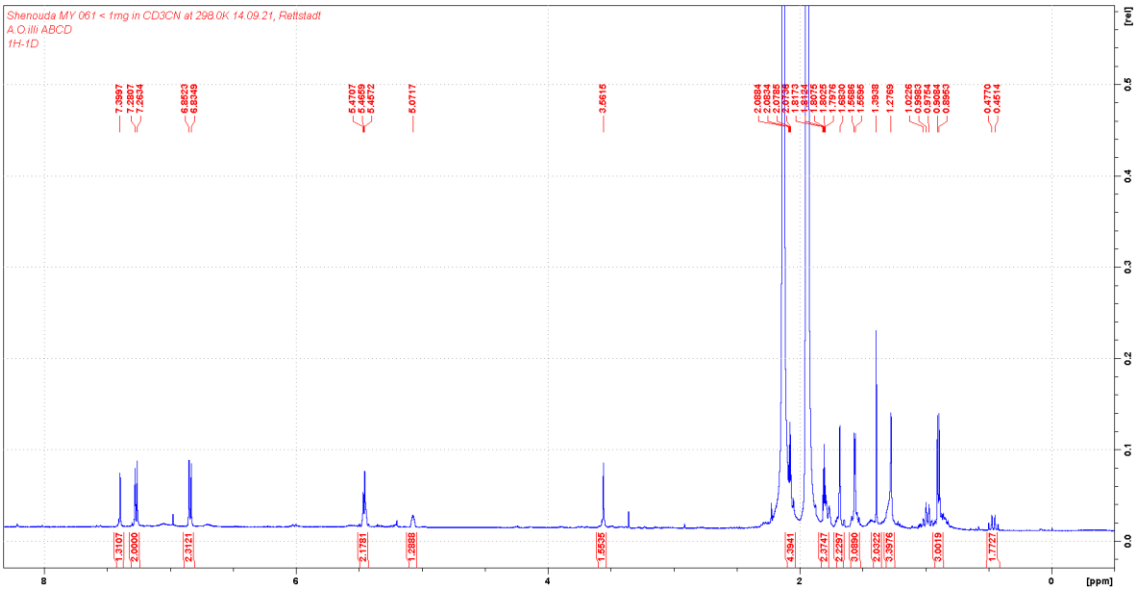


Figure 7.55 ^{1}H-NMR spectrum of compound **142** (500 MHz, Acetonitrile-d3).

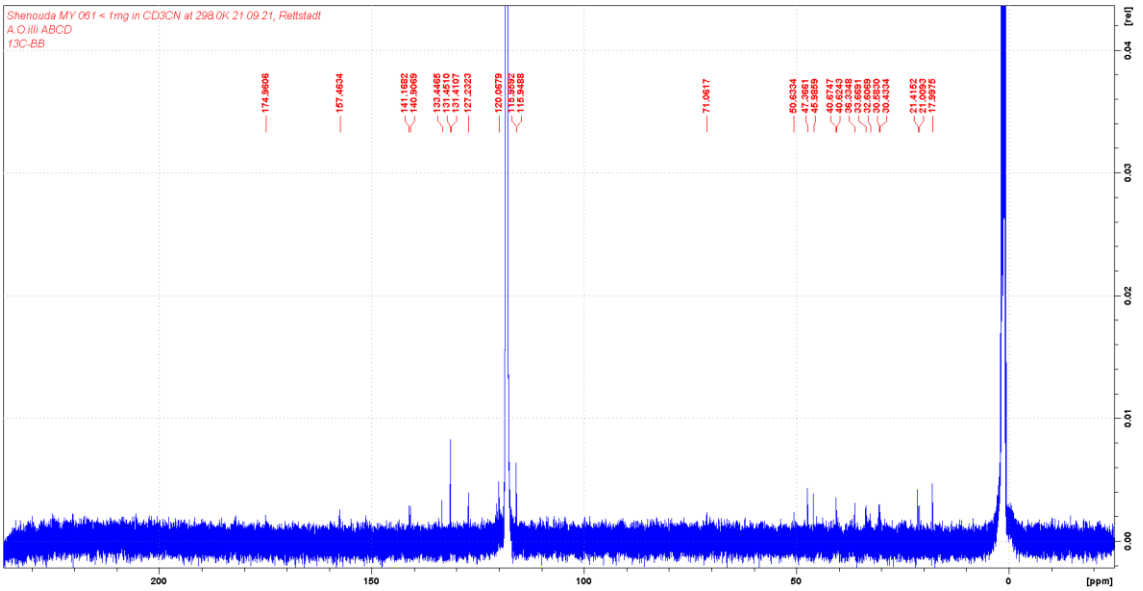


Figure 7.56 ^{13}C-NMR spectrum of compound **142** (125 MHz, Acetonitrile-d3).

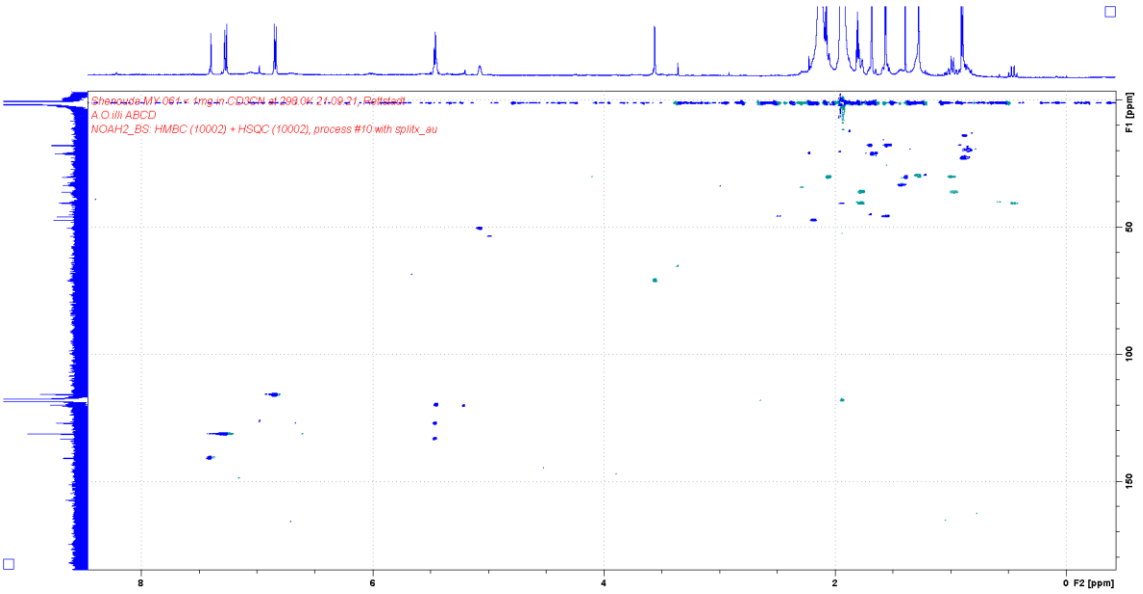


Figure 7.57 HSQC spectrum of compound **142** (Acetonitrile-d3).

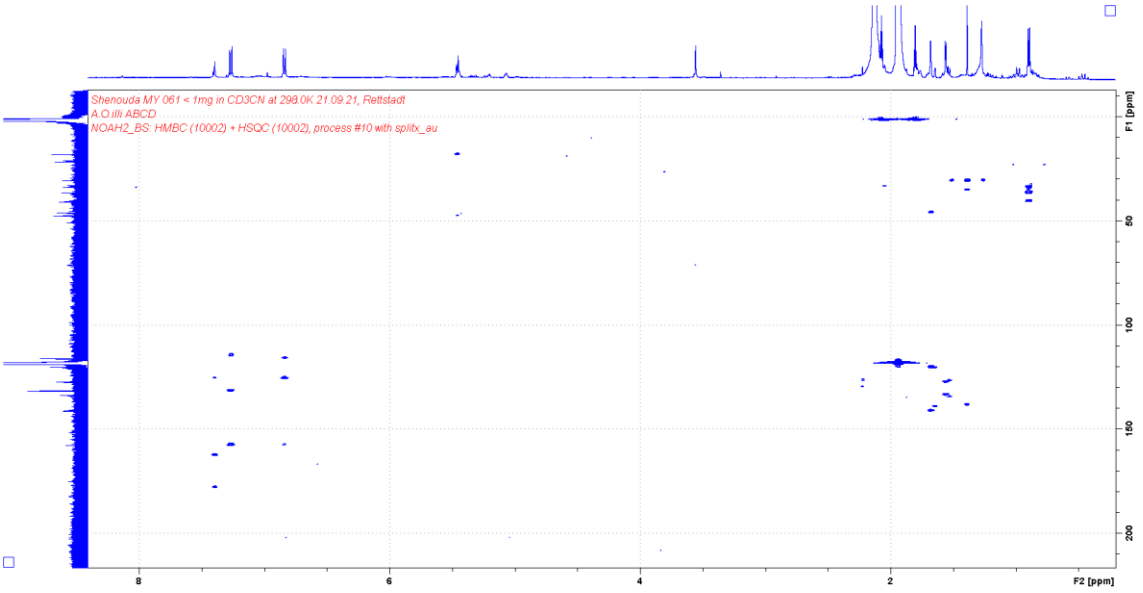


Figure 7.58 HMBC spectrum of compound **142** (Acetonitrile-d3).

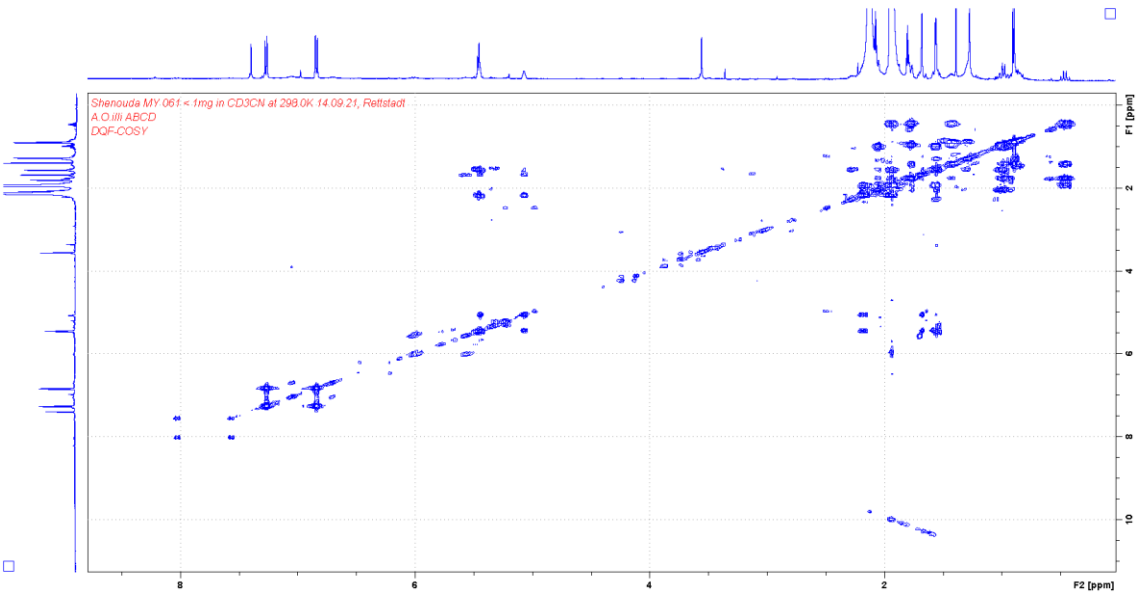


Figure 7.59 COSY spectrum of compound **142** (Acetonitrile-d3).

Curriculum vitae

Apr 2018- Feb 2022	**PhD student**, organic chemistry, Gottfried Wilhelm Leibniz University of Hannover, Hannover, Germany.
Sep 2012- Sep 2016	**Master of Pharmaceutical Sciences**, Pharmacognosy, Faculty of Pharmacy, Alexandria University, Alexandria, Egypt.
Sep 2007- Jun 2012	**Bachelor of pharmaceutical sciences,** Faculty of Pharmacy, Alexandria University, Alexandria, Egypt.
Sep 2004- Jun 2007	**High school**, Abbas Helmy secondary school for girls, Alexandria, Egypt.

List of Publications

1. Abd El-Salam, N. M., Radwan, M. M., Wanas, A. S., **Shenouda, M. L.**, Sallam, S. M., Piacente, S., ElSohly, M., & Ghazy, N. A. (2016). Phytochemical and biological evaluation of *Alcea rosea* L., growing in Egypt. *Planta Medica*, *82*(05), PC83. (Poster).

2. Abdel-Salam, N. A., Ghazy, N. M., Sallam, S. M., Radwan, M. M., Wanas, A. S., ElSohly, M., El-Demellawy, M. A., Abdel-Rahman, N. M., Piacente, S., & **Shenouda, M. L.** (2018). Flavonoids of *Alcea rosea* L. and their immune stimulant, antioxidant and cytotoxic activities on hepatocellular carcinoma HepG-2 cell line. *Natural product research*, *32*(6), 702-706.

3. Abdel Salam, N. A., Ghazy, N. M., Shawky, E., Sallam, S. M., & **Shenouda, M. L.** (2018). Validated HPTLC Method for Dihydrokaempferol-4′-O-glucopyranoside Quantitative Determination in *Alcea* Species. *Journal of chromatographic science*, *56*(6), 518-523.

4. **Shenouda, M. L.**, & Cox, R. J. (2021). Molecular methods unravel the biosynthetic potential of *Trichoderma* species, *RSC Advances*, *11*(6), 3622-3635.