

# **Expression and Engineering of Biosynthetic Enzymes in Fungi**

Von der Naturwissenschaftlichen Fakultät der  
Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

**Doktor der Naturwissenschaften (Dr. rer. nat.)**

genehmigte Dissertation

von

**Jin Feng, Master of Science (China)**

**2021**

Referent: Prof. Dr. Russell J. Cox

Korreferent: Prof. Dr. Mark Brönstrup

Tag der Promotion: 08.10.2021

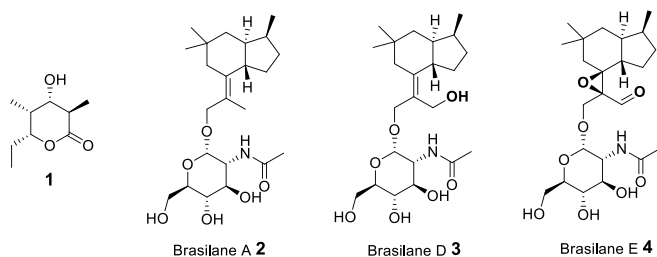
# Abstract

Key words: *Aspergillus oryzae*; Heterologous expression; DEBS1-TE; Brasilane; Biosynthesis

Fungal heterologous expression is a powerful tool for the production of natural products. In this presented work, we aim to further expand the application of fungal heterologous systems. In addition to iterative PKS, extensively studied in the past, fungal hosts are also expected to be capable of expressing modular PKS, which have never been expressed in fungi. Thus, DEBS1-TE, a minimal bimodular PKS for the triketide **1** biosynthesis, was taken as a representative model system to conduct expression in *Aspergillus oryzae* in this work.

In the re-engineered DEBS1-TE biosynthesis in *A. oryzae*, the supply pathways of propionate-related precursors were established. DEBS1-TE was codon-optimized and activated to be active in *A. oryzae*. With this engineered *A. oryzae*, the heterologous expression of a DEBS1-TE was functionally achieved to produce **1** (0.6 mg·L<sup>-1</sup>) in a filamentous fungal host for the first time. Then a set of titer optimizations of **1** was attempted, including fermentation conditions, overexpression of biosynthetic genes, and blockage of degradation pathways. It led to the 12-fold increase of the overall titer. Using the new *A. oryzae* system, we attempted to explore the possibility of fusion expression of modular and iterative PKSs. Unfortunately, there was no new product detectable.

Another aim of this work was to identify the biosynthetic pathway of brasilane-type sesquiterpenes brasilane A **2**, D **3**, and E **4** by *A. oryzae* heterologous expression. The *bra* biosynthetic pathway, including a terpene cyclase (BraA), an *N*-acetylglucosamine transferase (BraB) and a cytochrome P450 monooxygenase (BraC), were identified. With the biochemical characterization of enzymes *in vitro*, BraB displayed a broad scope of acceptor substrates *in vitro*.



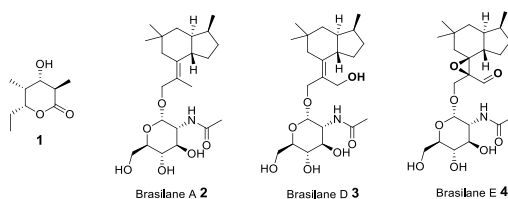
# Zusammenfassung

Schlüsselwörter: *Aspersillus oryzae*; heterologer Ausdruck; DEBS1-TE; Brasilan; Biosynthese

Die heterologe Expression in Pilzen ist ein mächtiges Werkzeug für die Herstellung von Naturprodukten. In dieser vorgestellten Arbeit wollen wir die Anwendung pilzlicher heterologer Systeme weiter ausbauen. Zusätzlich zu iterativer PKS, die in der Vergangenheit ausgiebig untersucht wurde, wird auch vermutet, dass Pilzwirte in der Lage sind, modulare PKS zu exprimieren, die noch nie in Pilzen exprimiert wurden. Daher wurde DEBS1-TE, eine minimale bimodulare PKS für die Triketid-**1**-Biosynthese, als repräsentatives Modellsystem zur Durchführung der Expression in *Aspergillus oryzae* in dieser Arbeit verwendet.

In der überarbeiteten DEBS1-TE-Biosynthese in *A. oryzae* wurden die Produktionswege von Propionat-verbundenen Vorstufen etabliert. DEBS1-TE wurde codon-optimiert und mittels eines pilzspezifischen Promotors aktiviert. Mit diesem gentechnisch veränderten *A. oryzae* wurde erstmals die heterologe Expression eines DEBS1-TE funktionell erreicht, um **1** ( $0,6 \text{ mg}\cdot\text{L}^{-1}$ ) in einem filamentösen Pilzwirt zu produzieren. Dann wurde eine Reihe von Titeroptimierungen von **1** experimentell erarbeitet, einschließlich Fermentationsbedingungen, Überexpression biosynthetischer Gene und Blockierung von Abbauwegen. Dies führte zum 12-fachen Anstieg des Gesamt-titers. Unter Verwendung des neuen *A. oryzae*-Systems versuchten wir, die Möglichkeit der Fusionsexpression von modularen und iterativen PKSs zu untersuchen. Leider war kein neues Produkt nachweisbar.

Ein weiteres Ziel dieser Arbeit war es, den Biosyntheseweg der Sesquiterpene Brasilan A **2**, D **3** und E **4** vom Brasilan-Typ durch heterologe Expression in *A. oryzae* zu rekonstruieren. Der *bra*-Biosyntheseweg, einschließlich einer Terpencyclase (BraA), einer *N*-Acetylglucosamin-Transferase (BraB) und einer Cytochrom-P450-Monooxygenase (BraC), wurde identifiziert und aufgeklärt. Bei der biochemischen Charakterisierung von Enzymen *in vitro* zeigte BraB ein breites Spektrum an Akzeptorsubstraten *in vitro*.



# Acknowledgement

First of all, I would like to thank my supervisor Prof. Russell Cox. Thank you so much for offering me a precious study opportunity in Hannover University and our group. Thank you for all your support and assistance over the past four years. During the long period of researching, you provided me a lot of helps and valuable advice. Your attitude towards scientific research also inspired me to insist on science. I feel so grateful to be your PhD student. Thank you.

I would like to thank Dr. Elizabeth Skellam. Thanks for all your efforts that you put on my project. You taught me a lot and gave so many inspiring ideas. Especially when I encountered some troubles pushing my projects forward, I appreciate your honest and sincere encouragement. Thanks for your patient instruction on my projects. I am so delighted to collaborate with you.

I want to thank you to Dr. Eric Kuhnert. You provided two very interesting projects for me. With your guide, I am pleased that we solved so many experimental problems and finally gained some achievements. I learnt a lot from your serious-minded guide and your earnest attitude. Thank you.

I would like to thank all cooperation partners. Thank Frank Surup and his colleagues from Prof. Marc Stadler group. Thanks for your important support in the brasilane project. I want to thank Maurice Hauser who took part in my two projects. Thanks for your timely and sincere assistance, and thank Anna Miller for your dedication in the brasilane work. I also thank Slawik Hrupins for your help in my project.

Thank all BMWZ media kitchen team. Especially thank Katja Körner for your patient work and responsible attitude which provide reliable support for my research. Thank all analytical departments in OCI. Thanks for your generous help.

I would like to thank all past and current Cox group members. In particular, thank Karen, Hao, Lukas and Carsten, Dongsong for teaching me experimental skills. And thank the alumni Chongqing, Steffen, Oliver, Haili, Tian, Sen, and current group members Yunlong, Mary, Lei, Katharina, Henrike, Evgeniya. Thanks.

Finally, I would like to thank my family and friends for your care and encouragement, and especially thank Su for your support and company.

# Abbreviations and Units

A	adenylation domain	IPP	isopentenyl diphosphate
aa	amino acid	<i>kanR</i>	kanamycin resistance gene
ACP	acyl carrier protein	kb	kilobase
<i>ampR</i>	ampicillin resistance gene	kDa	kilodalton
ATP	adenosine triphosphate	KR	$\beta$ -ketoreductase
AT	acyl transferase	KS	$\beta$ -ketosynthase
AU	absorption unit	LCMS	liquid chromatography-mass spectrometry
<i>A. oryzae</i>	<i>Aspergillus oryzae</i>	LB	lysogeny broth
BGC	biosynthetic gene cluster	LR	Gataway LR recombination
BLAST	basic local alignment search tool	M	mol/L
bp	base pair	MS	mass spectrometry
Carb	carbenicillin	6-MSAS	6-methylsalicylic acid synthase
<i>C-Met/CM</i>	carbon-methyltransferase	<i>m/z</i>	mass to charge ratio
<i>camR</i>	chloramphenicol resistance gene	mRNA	messenger ribonucleic acid
CoA	coenzyme A	mg	milligram
cDNA	complementary DNA	mL	milliliter
COSY	correlation spectroscopy	min	minute
d	days	M	molar
DH	dehydratase	nm	nanometer
DNA	deoxyribonucleic acid	NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
DAD	diode array detector	nrPKS	non-reducing polyketide synthase
DEBS	6-deoxyerythromycin B synthase	NRPS	nonribosomal peptide synthetase
DMSO	dimethyl sulfoxide	NMR	nuclear magnetic resonance
DMAPP	dimethylallyl diphosphate	ORF	open reading frame
DKC	Dieckmann cyclization	<i>O</i> -Met	oxygen-methyltransferase
DTT	dithiothreitol	prPKS	partial-reducing polyketide synthase
ESI	electron spray ionization	ppm	parts per million
ER	enoyl reductase	PPTase	phosphopantethinyl transferase
<i>E. coli</i>	<i>Escherichia coli</i>	PCC	propionyl-CoA carboxylase
eGFP	enhanced green fluorescent protein	PEG	polyethylene glycol
ELSD	evaporative light scattering detector	PKS	polyketide synthase
FAC-MS	fungal artificial chromosome and metabolomic scoring system	PCP	peptide carrier protein
FPP	farnesyl pyrophosphate	PCR	polymerase chain reaction
FAS	fatty acid synthase	PT	product template
FMO	flavin-dependent monooxygenase	R	reductive release domain
g	gram	RT-PCR	reverse transcription polymerase chain reaction
gDNA	genomic DNA	rpm	revolutions per minute
GGPP	geranylgeranyl diphosphate	RNA	ribonucleic acid
GPP	geranyl diphosphate	<i>S.</i>	<i>Saccharomyces cerevisiae</i>
HDAC	histone deacetylase	<i>cerevisiae</i>	
HEX	heterologous expression	SAM	<i>S</i> -adenosyl methionine
Hz	hertz	SDS-	sodium dodecyl sulphate polyacrylamide gel
HMBC	heteronuclear multiple bond correlation	PAGE	electrophoresis
HSQC	heteronuclear single quantum correlation	SAT	starter unit acyl carrier protein
HPLC	high performance liquid chromatography	T	thiolation
HRMS	high resolution mass spectrometry	TE	thiolesterase
hrPKS	highly reducing polyketide synthase	UV	ultraviolet
h	hours	UDP	uridine diphosphate
HygB	hygromycin B	<i>v/v</i>	volume per volume
		<i>w/v</i>	weight per volume
		YR	yeast recombination

# Contents

<b>Abstract</b> .....	<b>ii</b>
<b>Zusammenfassung</b> .....	<b>ii</b>
<b>Acknowledgement</b> .....	<b>iii</b>
<b>Abbreviations and Units</b> .....	<b>iv</b>
<b>Contents</b> .....	<b>v</b>
<b>1. Introduction</b> .....	<b>1</b>
<b>1.1 Fungal Natural Products</b> .....	<b>1</b>
1.1.1 Polyketides.....	2
1.1.2 Nonribosomal Peptides.....	3
1.1.3 Terpenoids.....	5
1.1.4 Alkaloids.....	6
<b>1.2 Engineered Biosynthesis of Fungal Natural Products</b> .....	<b>7</b>
1.2.1 Engineering Strategies.....	7
1.2.2 Heterologous Expression Systems.....	13
<b>1.3 Overall Aims</b> .....	<b>21</b>
<b>2. Heterologous Expression of Modular PKS DEBS1-TE in <i>A. oryzae</i></b> .....	<b>24</b>
<b>2.1 Introduction</b> .....	<b>24</b>
2.1.1 Heterologous Expression of DEBS1-TE in Bacteria.....	27
2.1.2 Heterologous Expression of Partial DEBS1-TE in Yeast.....	35
<b>2.2 Aims</b> .....	<b>38</b>
<b>2.3 Results</b> .....	<b>40</b>
2.3.1 Propionyl-CoA Production Pathway.....	40
2.3.2 Methylmalonyl-CoA Production Pathway.....	44
2.3.3 DEBS1-TE Cassette Construction.....	47
2.3.4 Product Analyses.....	52
2.3.5 Phosphopantetheinylation Pathway.....	58

2.3.6	Product Identification .....	60
2.3.7	Titer Calculation.....	65
<b>2.4</b>	<b>Discussion and Conclusion .....</b>	<b>66</b>
<b>3.</b>	<b>High-titer Production Optimization of Triketide Lactone in <i>A. oryzae</i> .....</b>	<b>67</b>
<b>3.1</b>	<b>Introduction .....</b>	<b>67</b>
3.1.1	Fermentation Conditions.....	67
3.1.2	Gene Expression Level.....	67
3.1.3	Degradation Pathways.....	68
<b>3.2</b>	<b>Aims .....</b>	<b>74</b>
<b>3.3</b>	<b>Results .....</b>	<b>75</b>
3.3.1	Optimization of Feeding Propionate.....	75
3.3.2	Removal of eGFP .....	76
3.3.3	PcsA Overexpression .....	77
3.3.4	Deletion of <i>mcsA</i> and <i>coaT</i> .....	79
<b>3.4</b>	<b>Discussion and Conclusion .....</b>	<b>90</b>
<b>4.</b>	<b>Reprogramming of DEBS1-TE in <i>A. oryzae</i> .....</b>	<b>93</b>
<b>4.1</b>	<b>Introduction .....</b>	<b>93</b>
4.1.1	DEBS1-TE Reprogramming in Bacteria.....	94
4.1.2	PKS Reprogramming in Fungi .....	97
<b>4.2</b>	<b>Aims .....</b>	<b>99</b>
<b>4.3</b>	<b>Results — Creation of Hybrid Bacterial-Fungal PKS.....</b>	<b>104</b>
4.3.1	Fusion Design .....	104
4.3.2	ACP Boundary Confirmation .....	106
4.3.1.2	Plasmid Assembly .....	107
4.3.2	Heterologous Expression in <i>A. oryzae</i> .....	111
<b>4.4</b>	<b>Results — Module Recombination.....</b>	<b>117</b>
<b>4.5</b>	<b>Discussion and Conclusion .....</b>	<b>121</b>
<b>5.</b>	<b>Expression and Engineering of the Brasilane Gene Cluster .....</b>	<b>123</b>



<b>5.1</b>	<b>Introduction .....</b>	<b>123</b>
5.1.1	Brasilane-type Sesquiterpenes in Nature.....	123
5.1.2	Fungal Glycosyltransferases and Glycosylated Sesquiterpenes .....	126
5.1.3	Bioactivities .....	128
5.1.4	Relevant Biosynthesis .....	129
<b>5.2</b>	<b>Aims .....</b>	<b>133</b>
<b>5.3</b>	<b>Results.....</b>	<b>134</b>
5.3.1	Production and Isolation of Brasilane-type Sesquiterpenes .....	134
5.3.2	Bioinformatic Analysis of Brasilanes BGC .....	136
5.3.3	Heterologous Expression of the Brasilane BGC in <i>A. oryzae</i> .....	137
5.3.4	Protein Characterization <i>In Vitro</i> .....	144
<b>5.4</b>	<b>Discussion and Conclusion .....</b>	<b>150</b>
<b>6.</b>	<b>Experimental.....</b>	<b>153</b>
<b>6.1</b>	<b>Media, Buffers and Solutions .....</b>	<b>153</b>
6.1.1	Media .....	153
6.1.2	Buffers and Solutions .....	154
<b>6.2</b>	<b>Strains, Plasmids and Primers .....</b>	<b>154</b>
6.2.1	Strains.....	154
6.2.2	Plasmids.....	155
6.2.3	Primers.....	156
<b>6.3</b>	<b>Microbiology Methods .....</b>	<b>158</b>
6.3.1	Transformation of Chemically Competent <i>E. coli</i> .....	158
6.3.2	LiOAc-mediated Transformation of Competent <i>S. cerevisiae</i> .....	158
6.3.3	PEG-mediated Transformation of <i>A. oryzae</i> NSAR1.....	159
<b>6.4</b>	<b>Molecular Biological Methods .....</b>	<b>160</b>
6.4.1	Genomic DNA, RNA and cDNA Preparation.....	160
6.4.2	Plasmid DNA Extraction .....	161
6.4.3	Gene and Plasmid Identification .....	161

<b>6.5 Biochemistry Methods.....</b>	<b>162</b>
6.5.1 Protein Expression and Purification .....	162
6.5.2 Enzyme Activity Assay.....	164
<b>6.6 Chemical Analysis.....</b>	<b>166</b>
6.6.1 Extraction of Fungal Metabolites.....	166
6.6.2 Analytical LCMS .....	167
6.6.3 GCMS .....	168
6.6.4 HRMS and NMR.....	168
<b>Reference .....</b>	<b>169</b>
<b>Appendix .....</b>	<b>178</b>
<b>NMR Spectra.....</b>	<b>178</b>
<b><i>Curriculum Vitae</i>.....</b>	<b>ix</b>
<b>List of Publications.....</b>	<b>ix</b>

# 1. Introduction

In nature, there is a considerably abundant library of natural products. Natural products are properly defined as chemical compounds produced by living organisms. They are classified as *primary metabolites* or *secondary metabolites*. Primary metabolites are involved in fundamental processes of growth, development and reproduction. Secondary metabolites are not essential for life, but often offer selective advantages, and the term *natural product* usually refers to secondary metabolites alone. Natural products originate from plants, animals, and microorganisms. Based on a wide range of sources, natural products have diverse structures, including chiral centers, ring fusions and many kinds of functional groups. Natural products have evolved over hundreds of millions of years to interact with biological systems. Therefore, compared with drugs produced by chemical synthesis, natural products generally have more significant properties in terms of recognizing and affecting biological targets.<sup>1,2</sup>

## 1.1 Fungal Natural Products

The kingdom of fungi includes an enormous diversity of organisms, which are very widely distributed in nature. To date, approximately  $10^5$  fungal species are known, but it is estimated that as many as  $2 \times 10^6$  fungal species probably exist.<sup>3,4</sup> Many fungi can produce a variety of secondary metabolites or induce the plant host to produce new compounds for surviving. Fungal secondary metabolites have been the source for new drug discovery based on their unique chemical structures. As bioactive metabolites, their therapeutic effects have been determined, such as antibacterial, antioxidant, antitumor, antidiabetic, and even insecticidal agents.<sup>5</sup> For example, penicillins **5**, one of the most successful natural products to be used as a medicine, was isolated from *Penicillium* genus. It is able to kill multiple types of bacteria.<sup>6</sup>

Some fungal natural products are recognized as model compounds and have a long history during the development of the science of fungal secondary metabolites (Fig. 1.1). For example, citrinin **6** has been used as a key model at all stages from

chemical isolation and structure elucidation, to chemical synthesis, and for the understanding of complex biosynthetic pathways.<sup>7,8</sup> Similarly, the discovery of the tropolone stipitatic acid **7** has also stimulated important developments in the field of chemical synthesis and biosynthesis.<sup>9,10</sup> According to their biosynthetic origin, fungal secondary metabolites are generally divided into four classes, including polyketides, nonribosomal peptides, terpenes, and alkaloids. Other minor classes are also known.

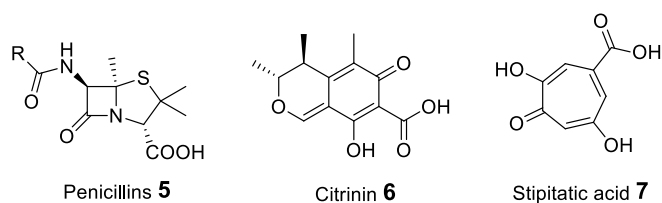


Figure 1.1 Examples of fungal natural products.

### 1.1.1 Polyketides

Polyketides are a very large and diverse family of secondary metabolites, and fungi and bacteria are a rich source of these biologically active compounds. Many polyketides are infamous toxins, such as aflatoxin B1 **8**, fumonisin **9**, pigments or virulence factors like the melanin precursor 1,3,6,8-tetrahydroxynaphthalene (THN) **10** (Fig. 1.2).<sup>11–13</sup> In addition to those, fungal polyketides are also relevant pharmaceutically, such as cholesterol-lowering statins lovastatin **11**, mevastatin **12**, and pravastatin **13**.<sup>14</sup>

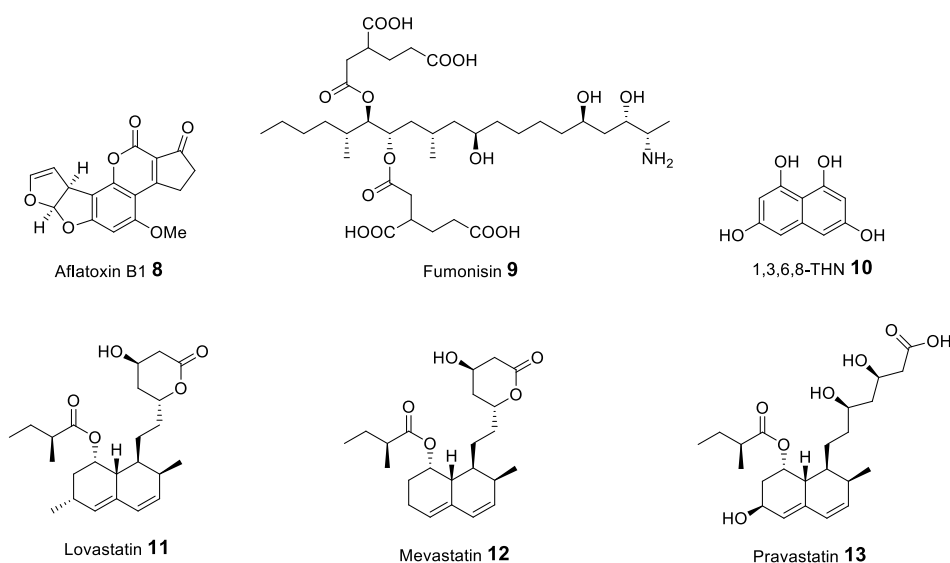
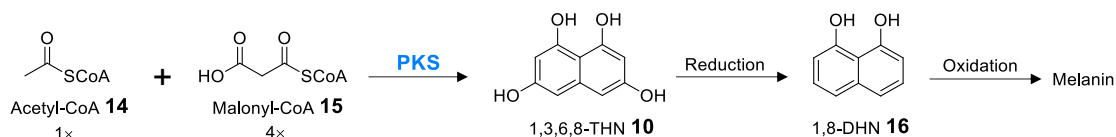


Figure 1.2 Structures of selected fungal polyketides.

Polyketide biosynthesis is closely related to fatty acid biosynthesis.<sup>15</sup> For example, the pentaketide 1,3,6,8-THN **10** is the starting precursor during the fungal melanin biosynthesis. The formation of **10** is derived from one acetyl-CoA **14** and four malonyl-CoA **15** molecules and is catalyzed by a polyketide synthase enzyme (PKS). Further, it is reduced to the intermediate 1,8-dihydroxynaphthalene (DHN) **16**, which undergoes polymerization to DHN-melanin (Scheme 1.1).

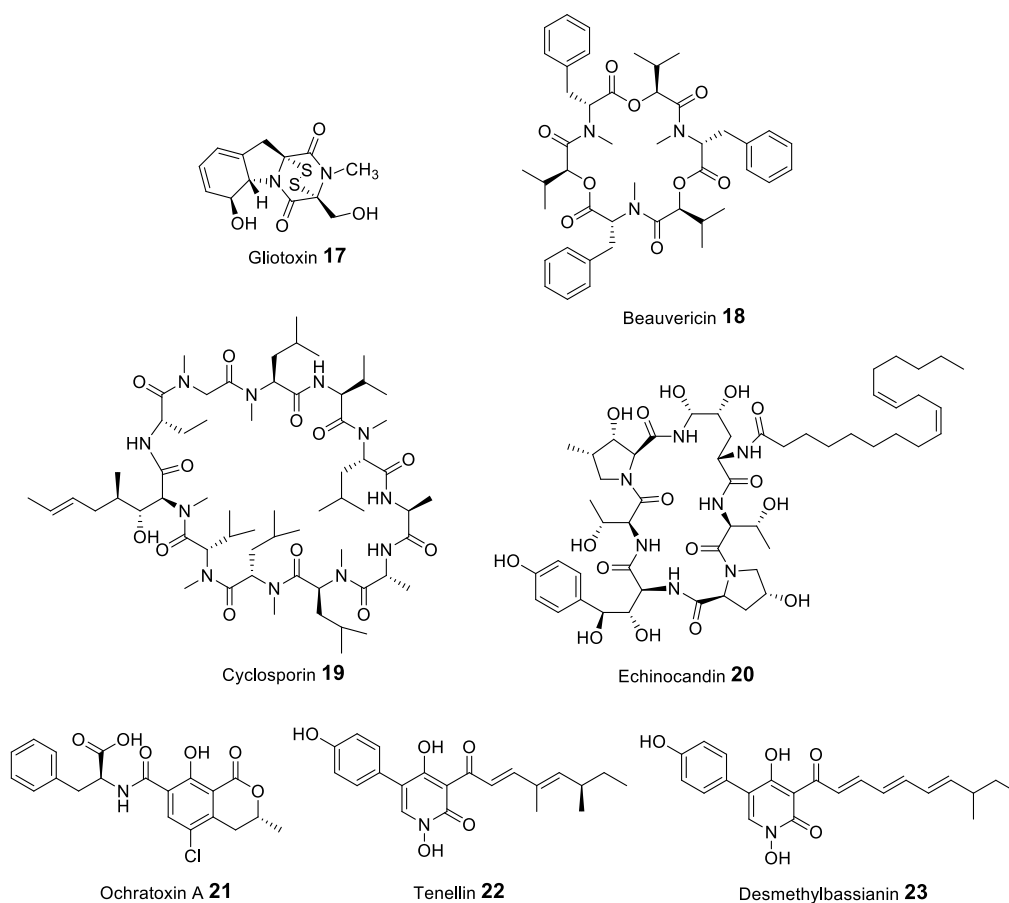


**Scheme 1.1** Biosynthetic route of fungal melanin.<sup>16</sup>

### 1.1.2 Nonribosomal Peptides

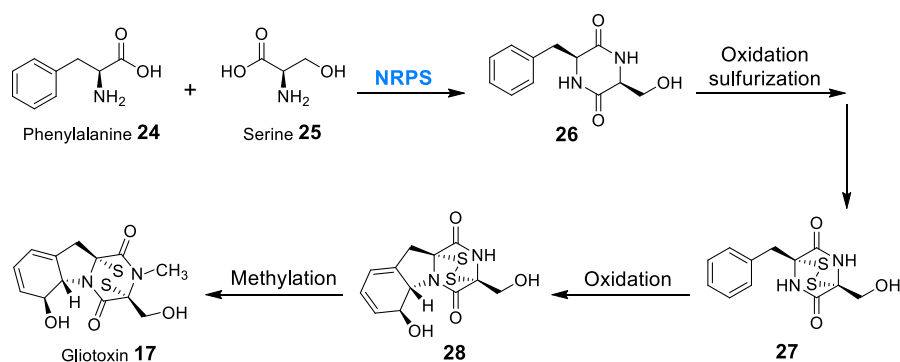
Nonribosomal peptides (NRPs) are a very important class of secondary metabolites produced in filamentous fungi. Fungal nonribosomal peptides possess remarkable structural diversity. On the basis of this, the family of nonribosomal peptide displays significant biological activities useful for pharmaceutical applications (*e.g.* antibiotics, antitumor compounds, and immunosuppressants). From the most famous drug penicillins **5** on, more and more nonribosomal peptides have been discovered to date, such as gliotoxin **17**, beauvericin **18**, cyclosporin **19**, and echinocandin **20**. In some cases, nonribosomal peptides are fused with polyketides to form hybrid polyketide–nonribosomal peptide natural products, such as ochratoxin A **21**, tenellin **22** and desmethylbassianin **23** (Fig. 1.3).<sup>17</sup>

Biosynthesis of nonribosomal peptides is catalyzed by large multifunctional enzymes called nonribosomal peptide synthetases (NRPSs) which are among the largest enzymes known in nature. In the structures of nonribosomal peptides, not only standard amino acid residues but also non-proteinogenic amino acid residues are incorporated. It can be linear, cyclic or branched structural configurations. The amino acid residues of nonribosomal peptides can be modified by oxidation, sulfurization, or *N*-methylation.<sup>18</sup>



**Figure 1.3** Structures of selected fungal nonribosomal peptides and polyketide-nonribosomal peptides.

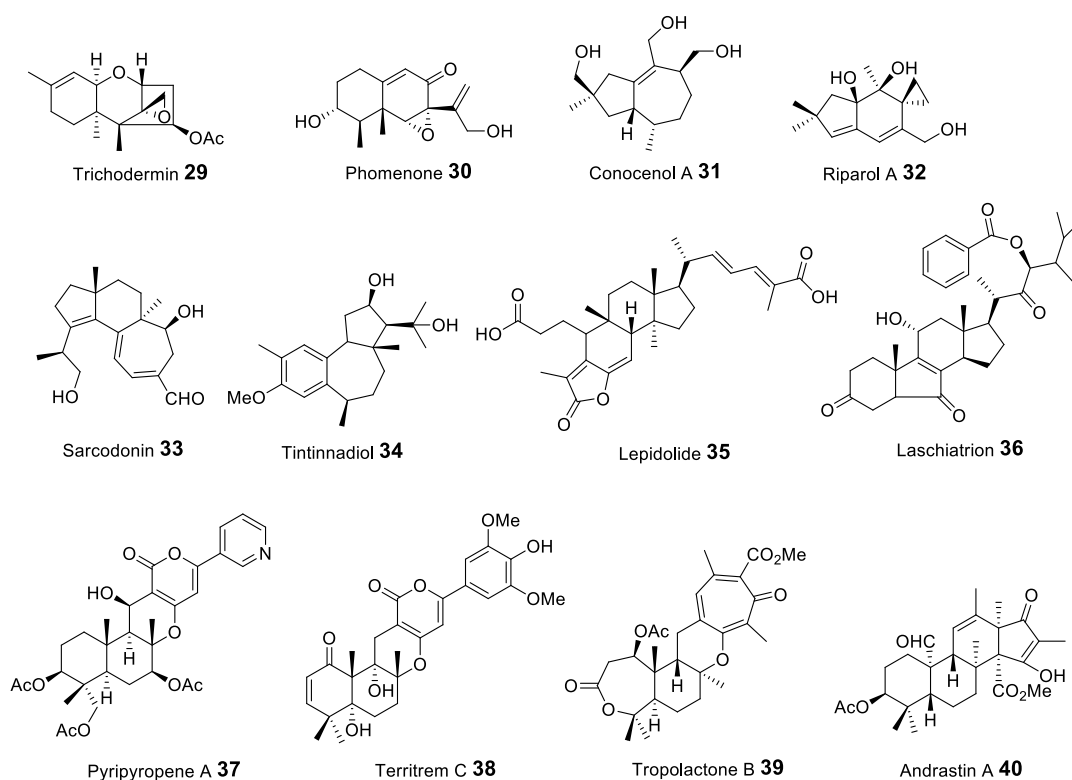
For example, gliotoxin **17** is a typical nonribosomal peptide. Its biosynthetic pathway has been proposed.<sup>19,20</sup> The initial step of the biosynthetic pathway is likely to involve condensation of phenylalanine **24** and serine **25** catalyzed by a non-ribosomal peptide synthetase (NRPS) to form **26**. Subsequently, through a series of oxidation and sulfurization, a disulphide bridge is generated in **27**. Further oxidation gives **28** and finally *N*-methylation yields gliotoxin **17** (Scheme 1.2).



**Scheme 1.2** Proposed reaction route of gliotoxin **17** biosynthesis.

### 1.1.3 Terpenoids

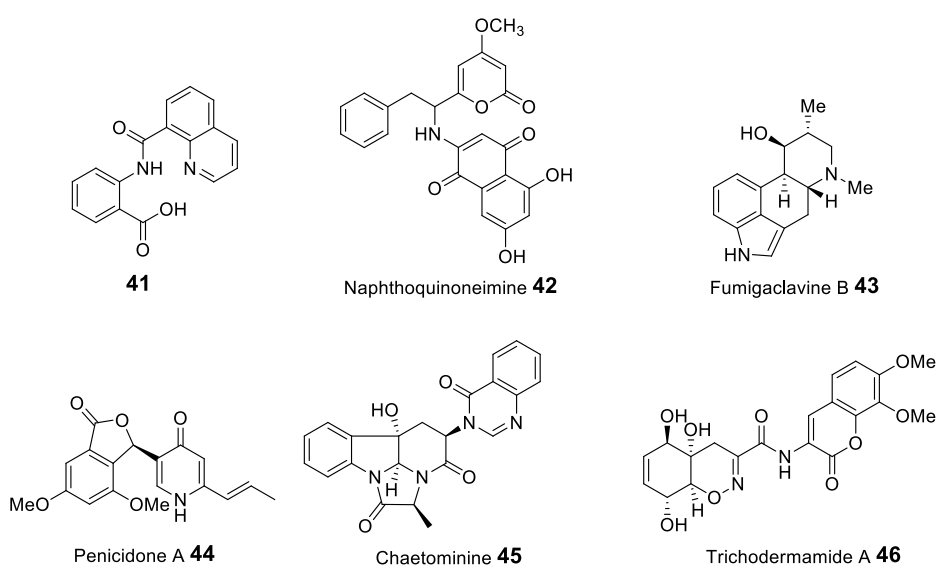
Terpenoids are among the most prominent metabolites of fungi. All terpenoids are derived from the common five-carbon isoprenyl diphosphate intermediates isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which are synthesized from acetyl-CoA **14**. Successive head-to-tail condensation of one to three IPP to DMAPP gives geranyl diphosphate (GPP), farnesyl diphosphate (FPP), or geranylgeranyl diphosphate (GGPP). These linear molecules are specifically cyclized by terpene cyclases or incorporated to other molecules by prenyl transferases, thereby generating many different kinds of terpenoids,<sup>21</sup> such as sesquiterpenoids (*e.g.* trichodermin **29**, phomenone **30**, conocenol A **31** and riparol A **32**), diterpenoids (*e.g.* sarcodonin **33** and tintinnadiol **34**), triterpenoids (*e.g.* lepidolide **35** and laschiatrion **36**), and meroterpenoids (*e.g.* pyripyropene A **37**, territrem C **38**, tropolactone B **39** and andrastin A **40**, Fig. 1.4).



**Figure 1.4** Structures of selected fungal terpenoids.

### 1.1.4 Alkaloids

Alkaloids are the fourth major secondary metabolite family found in fungi. They are an important source of leads of drugs. A number of alkaloids exhibit potent biological properties such as antimicrobial, insecticidal, cytotoxic, and anticancer activities.<sup>22</sup> Many have important pharmaceutical uses. Many alkaloids have been discovered in fungi. According to their structural diversity, alkaloids are divided into many types, such as quinoline and isoquinoline (*e.g.* **41**), amines and amides (*e.g.* naphthoquinoneimine **42**), indole derivatives (*e.g.* fumigaclavine B **43**), pyridines (*e.g.* penicidone A **44**), quinazolines (*e.g.* chaetominine **45**), and others (*e.g.* trichodermamide A **46**, Fig. 1.5).



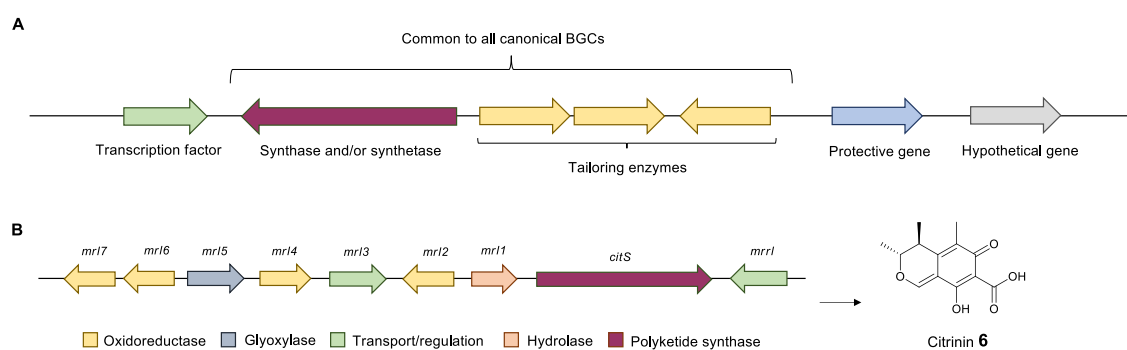
**Figure 1.5** Structures of selected fungal alkaloids.

Thus, natural product-based drug discovery has become a major exploration for scientists. Although original natural products are still discovered and directly applied, this trend of isolating natural products from nature is becoming more and more difficult. In contrast, their analogues usually play a more efficient role in the application process because of lower toxicity and better physicochemical properties. To obtain fungal natural products that are difficult to isolate or synthesize chemically, the engineered biosynthesis of fungal natural products is becoming an effective avenue. Additionally, it can also expand the diversity of natural product molecules, biosynthesizing some “unnatural” natural products.



## 1.2 Engineered Biosynthesis of Fungal Natural Products

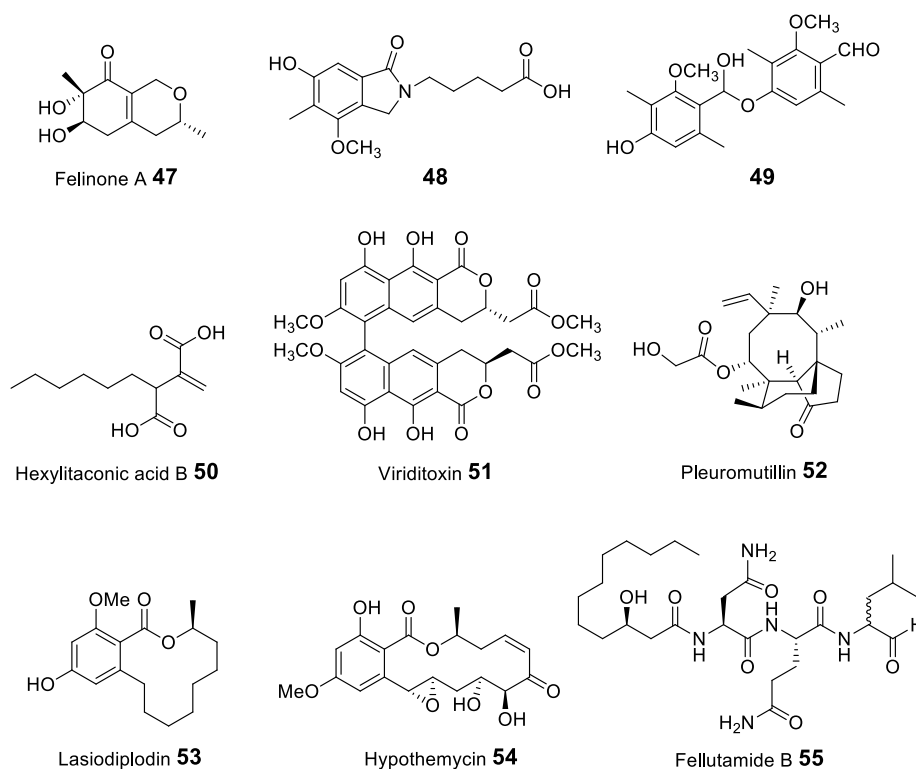
Fungal natural products are mainly derived from central metabolic pathways and primary metabolite pools where acyl-CoAs are used to synthesize the critical initial building blocks for the synthesis of polyketides (*e.g.* aflatoxin B1 **8**), nonribosomal peptides (*e.g.* penicillins **5**), terpenes (*e.g.* trichodermin **29**), and alkaloids (*e.g.* **41**).<sup>23</sup> In contrast to the genes that encode the synthesis of primary metabolites, which are dispersed through the fungal genome, the synthesis of secondary metabolites is normally encoded by genes that are arranged contiguously in a biosynthetic gene cluster (BGC, Fig. 1.6 A). Biosynthetic gene clusters are minimally composed of genes that encode a chemically defining synth(et)ase, some tailoring enzymes, transcription factors, transporters, sometimes resistance mechanisms, and often hypothetical proteins with unknown function. BGCs can range from two genes to over 20 genes. For example, the fungal gene cluster for the citrinin **6** biosynthesis is a typical biosynthetic BGC of polyketide (Fig. 1.6 B).<sup>8</sup>



**Figure 1.6** The general composition of fungal biosynthetic gene clusters (BGCs): **A**, the minimal components of a fungal biosynthetic gene cluster; **B**, the fungal gene cluster for citrinin **6** biosynthesis.

### 1.2.1 Engineering Strategies

Based on the architecture of fungal BGCs, it becomes possible to engineer the biosynthesis of fungal natural products *in vivo* using synthetic biology techniques, such as bioinformatic tools to predict and identify BGCs, genetic regulation in chassis, and introduction to heterologous hosts for expression. Here, some engineering strategies for the biosynthesis of fungal natural products (Fig. 1.7) are reviewed.



**Figure 1.7** Structures of fungal natural product examples by production of engineering strategies.

### 1.2.1.1 Manipulating Global Transcription

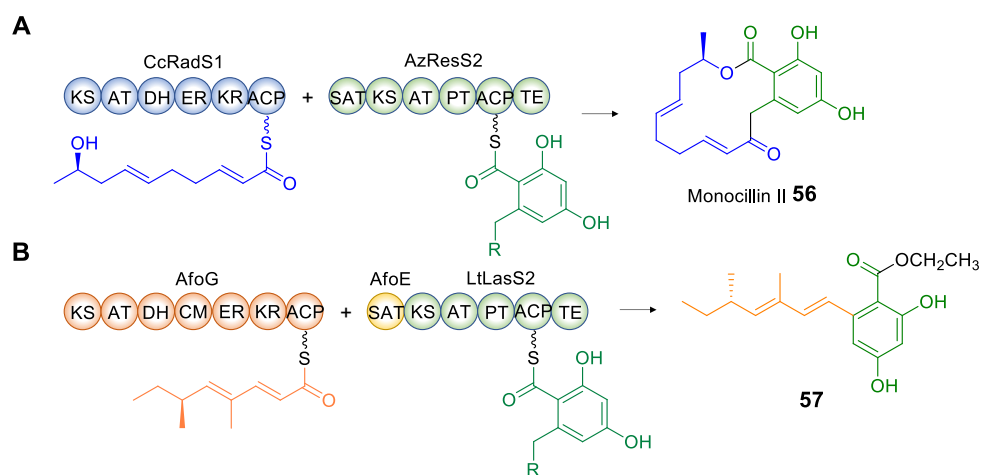
In fungi, many BGCs are located within heterochromatic regions. The transcription of BGCs is repressed in these regions, displaying silent *in vivo*. It has been found that histone deacetylases (HDACs) can negatively regulate the production of secondary metabolites. Therefore, deletion of HDAC genes or inhibition of HDAC enzymes is sometimes effective for the manipulation of global transcription. For example, the inactivation of a histone H3 deacetylase in *Calcarisporium arbuscula* led to an overall upregulation transcriptionally.<sup>24</sup> More than 75 % of biosynthetic genes in *C. arbuscula* were overexpressed. Compared to the predominant production of two secondary metabolites in the wild type, ten compounds with novel structures were overproduced in the histone deacetylase inactivation mutant.

Another example is the discovery of *mcrA*, which encodes a transcription factor in *A. nidulans*.<sup>25</sup> It was revealed that *mcrA* regulates at least ten gene clusters of secondary metabolites. Deletion of *mcrA* stimulated the production of secondary metabolites, leading to discovery of the antibiotic felinone A **47** and two novel phytotoxin cichorine derivatives **48** and **49** (Fig. 1.7). Therefore, manipulating

global transcription provides an efficient engineered strategy to mine the chemical diversity of natural products. However, such methods are highly unpredictable and often very specific to particular fungal species and do not offer a general method for activating specific BGCs in specific organisms.

### 1.2.1.2 Enzyme Engineering

Enzyme-based engineering is an extensively used strategy. Fungi possess an abundant source of enzymes, but natural enzymes have limitations such as poor productivity, low catalytic efficiency, and weak selectivity or specificity. To address these issues, engineering strategies of enzymes have been commonly employed, including directed evolution, site-directed mutagenesis, truncation, and terminal fusion.<sup>26</sup>



**Figure 1.8** Combinatorial biosynthesis of different sources of PKSs: **A**, the pairing of CcRadS1-AzResS2 produced monocillin II **56**; **B**, the pairing of AfoG-LtLasS2-SAT<sub>AfoE</sub> produced **57**. Abbreviations: KS, ketosynthase; AT, acyl-CoA transferase; DH, dehydratase; ER, enoyl reductase; KR, ketoreductase; ACP, acyl carrier protein; SAT, starter acyl transferase; CM, C-methyltransferase; PT, product template.

The engineering manipulations of fungal PKS and PKS-NRPS enzymes are a convenient point due to their modular or iterative feature. Many enzyme-based engineering of PKSs and PKS-NRPSs have exemplified the rational design of novel fungal natural products. Expression of random pairs of iterative PKS subunits from benzenediol lactone PKSs in a yeast heterologous host created a diverse library of benzenediol lactone congeners. For example, pairing CcRadS1 (from radicicol highly-reducing PKS) and AzResS2 (from resorcylic non-reducing PKS) afforded

the production of monocillin II **56** in a good yield (6 mg·L<sup>-1</sup>, Fig. 1.8 A).<sup>27</sup> The combinatorial biosynthesis of AfoG (from asperfuranone highly reducing PKS) and LtLasS2 (from lasiodiplodin non-reducing PKS) was also designed. The SAT (LtLasS2) was replaced by SAT (AfoE). AfoG-derived dimethyl tetraketide was extended by three additional ketide units by LtLasS2 to give the heptaketide **57** (0.2 mg·L<sup>-1</sup>, Fig. 1.8 B).<sup>28</sup>

The fusion expression between PKS and NRPS modules has also been achieved. PKS and NRPS modules from five different PKS-NRPSs (EqxS, FsdD, CpaS, PsoA, LovB) were fused together in a total of 57 different combinations, comprising 34 distinct module swaps.<sup>29</sup> The resulting fusions produced six new compounds. Thus, these techniques of enzyme engineering provide opportunities for the creation of more diverse unnatural products on the basis of natural enzymes. However, the yield of five unpredictable compounds from 57 cloning experiments is highly inefficient and more rational and effective methods are required.

### 1.2.1.3 Pathway Engineering

Pathway engineering (or metabolic engineering) is a powerful strategy, focusing on the manipulation of genetic or regulatory processes to produce metabolites of interest. Many techniques have been developed for pathway engineering so far.<sup>30</sup>

#### Overexpression

In fungi, there are a large number of BGCs in a silent state. By overexpression of the pathway transcription factor, the silent BGCs can be sometimes awaked to produce specific products. Alternatively, the native promoters can be replaced by strong inducible or constitutive promoters to improve titer of natural products. For example, in the biosynthetic study of alkylcitric acids in *A. niger*, the transcriptional regulator gene *akcR* was overexpressed. The production of hexylitaconic acid **50** (Fig. 1.7) was 200-fold higher than that from natural production in other strains of *A. niger*.<sup>31</sup> However, these types of experiment are not general because only between 35 % and 50 % of BGCs contain a local transcriptional regulator, and many of these

are repressors - and it not yet certain to predict if any given regulator is an activator or a repressor.

### Gene knockout

In biosynthetic studies, an uncharacterized BGC first needs to correlate with a known specific natural product. The gene knockout technique can build the correlation between BGCs and natural products. By deleting the core gene of BGCs, the relevant natural products should be abolished. In addition, the gene knockout technique can also be used to terminate a biosynthetic pathway at a specific point. The resulting intermediate metabolites can facilitate to reveal the elucidation of mechanism of the entire biosynthetic pathway. Several efficient gene-knockout methods have been developed so far.<sup>32-34</sup> Numerous biosynthetic pathways have been fully elucidated using multiple gene-knockout strategies. For example, an eight-membered gene cluster encoding viriditoxin **51** (Fig. 1.7) biosynthesis was characterized by targeted gene disruptions, revealing the roles of each gene and establishing the definite biosynthetic pathway.<sup>35</sup> However, a major disadvantage of this strategy is that many fungi which produce interesting or useful natural products cannot be manipulated genetically, because the molecular tools do not exist or because the native organism is intransigent for genetic transformation.

### Heterologous Expression

Heterologous expression of gene/gene cluster into a highly versatile surrogate host is the most successful method for characterizing the function of BGCs. Compared to secondary metabolite-producing native strains, heterologous hosts possess more effective promoters, regulatory factors, and other gene-controlling elements, leading to higher productivity of natural products.<sup>36</sup> There are many chassis strains which can be used as an excellent heterologous host, such as *Escherichia coli*, yeast, and filamentous fungi. In turn, the development of heterologous expression systems has accelerated the rate at the discovery, characterization, and engineering of fungal natural products. For example, a seven-membered gene cluster responsible for pleuromutilin **52** (Fig. 1.7) biosynthesis was identified in the producing strain *Clitopilus passeckerianus*. However, the production of pleuromutilin **52** could not be

increased significantly in this native host by various targeted approaches. When the gene cluster was reconstructed in the fungal host *Aspergillus oryzae*, a remarkable increase (2106 %) in production was achieved.<sup>37</sup>

Furthermore, as heterologous expression of BGCs is largely limited to expression of single gene/gene cluster, scalable platforms have been developed for the rapid cloning and heterologous expression of numerous BGCs to identify novel natural products.<sup>30</sup> For example, a fungal artificial chromosome and metabolomic scoring (FAC-MS) system was established. Up to 56 secondary metabolite BGCs from diverse fungal species were subject to this system and then expressed in *A. nidulans*, leading to discovery of 15 new metabolites.<sup>38</sup> Another example is the establishment of HEx (Heterologous expression) platform in *Saccharomyces cerevisiae*. 41 fungal biosynthetic gene cluster from diverse fungal species were applied on this platform, 22 of which produced detectable natural products.<sup>39</sup>

### **Combinatorial Biosynthesis**

Combinatorial biosynthesis is a strategy to generate novel natural products through reconstructing novel BGCs with genes from different biosynthetic pathways. Based on this, a combinatorial biosynthetic pathway can direct the production of the targeted product. For example, two orthologous *O*-methyltransferases (*O*-MeTs) LtO-MeT and HsO-MeT were identified in the biosynthetic gene clusters for two fungal benzenediol lactones lasiodiplodin **53** and hypothemycin **54** (Fig. 1.7). These *O*-MeTs displayed considerable substrate promiscuity and tunable regioselectivity. By combinatorial tailoring of LtO-MeT and HsO-MeT, eight unnatural *O*-methylated benzenediol lactone polyketides were generated.<sup>40</sup> Furthermore, the combinatorial biosynthesis of components even from different kingdom of life (*i.e.* bacterial, fungal and plant biosynthetic enzymes) is also compatible.

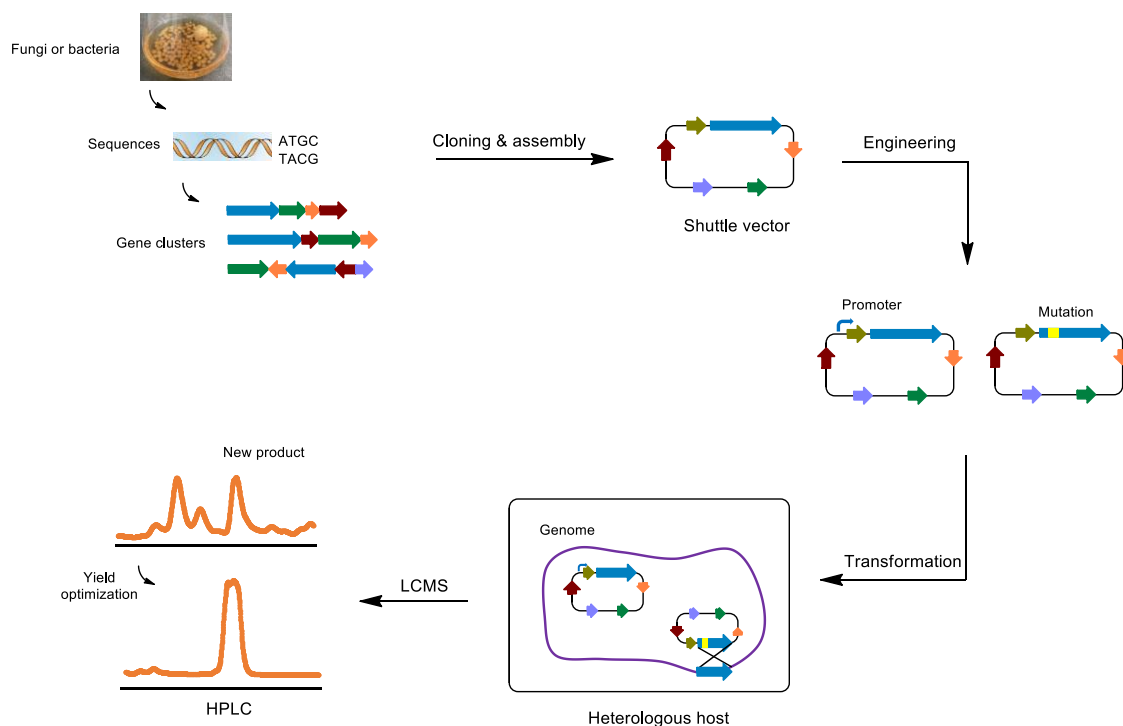
### **Other Techniques**

In addition, some other pathway engineering techniques are also used widely. For example, resistance-guided genome mining is a method to discover novel natural products by identification of a self-resistance gene encoded in the BGC. For example,

the *inp* gene cluster *A. nidulans* was identified, having a gene *inpE*, that encodes a proteasome subunit. It was speculated that the *inp* gene cluster may produce a proteasome inhibitor. Based on this hypothesis, the *inp* gene cluster was activated and produced the proteasome inhibitor fellutamide B **55** (Fig. 1.7).<sup>41</sup> Moreover, genome editing using CRISPR-Cas9 has also been well developed for pathway engineering.<sup>42</sup>

### 1.2.2 Heterologous Expression Systems

As mentioned before, heterologous expression is an effective strategy of pathway engineering. A general route of heterologous expression of biosynthetic gene clusters was showed in Scheme 1.3. First, the genomic DNA of microorganisms is sequenced and then annotated as biosynthetic gene clusters. By cloning and assembly, gene fragments are inserted into a shuttle vector. To engineered the gene cluster, promoters or mutations can be introduced. Next, the shuttle plasmid harboring the engineered gene cluster was transferred into a heterologous host for homologous recombination with the genome. Lastly, new products are generated and further optimized in production.



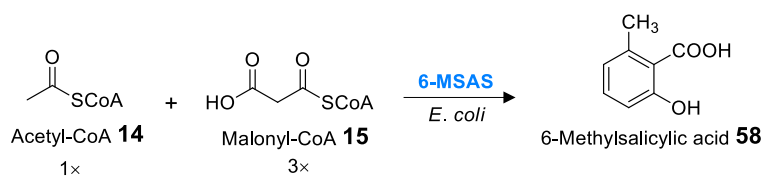
**Scheme 1.3** General route of heterologous expression of biosynthetic gene clusters.

The choice of heterologous host is a crucial step because its productivity determines the yield of natural products. The heterologous host should be well-characterized. The cloning, transformation, and cultivation methods and parameters have been established. Furthermore, the heterologous host should possess rapid growth rate. Most importantly, the gene expression mechanisms and metabolism support biosynthesis of heterologous gene clusters.<sup>43</sup> In this section, two commonly used bacterial and fungal systems are briefly reviewed on the heterologous expression of fungal natural products.

### 1.2.2.1 Bacterial Systems

Bacterial system is a highest priority to choose for the heterologous expression due to its technical convenience. Organisms such as *E. coli* are of great advantage simply because they are well-characterized and rapid to culture. Numerous genetic transformation techniques have been developed to facilitate the introduction of foreign DNA into bacterial heterologous systems. Furthermore, through optimization of pathway engineering, the production of natural products can be significantly increased, making the analysis and isolation of products much easier.

The common bacterial systems include *E. coli*, *Streptomyces* species, *Bacillus subtilis*, etc. Most researches, however, have focused on *E. coli* since it is the best model prokaryotic host for protein expression. Compounds that have been heterologously produced in *E. coli* are usually of bacterial origin. Until 1998, the first fungal secondary metabolite 6-methylsalicylic acid **58** was produced in *E. coli* (Scheme 1.4).<sup>44</sup> Nevertheless, the expression of large enzymes such as PKSs and NRPSs are still limited, and prokaryotic hosts cannot splice introns for example, do not contain complex organelles such as peroxisomes, cannot target proteins to particular intracellular compartments, and cannot post-translationally modify many proteins.



**Scheme 1.4** Heterologous production of 6-methylsalicylic acid **58** in *E. coli*. 6-MSAS: 6-methylsalicylic acid synthase.



*E. coli* has several restrictions that prevent the use of heterologous expression of fungal biosynthetic gene clusters.<sup>45</sup> First, *E. coli* has mainly been used as a model organism for protein expression and purification, not often as a producer of secondary metabolites. Many proteins need post-translational modifications after synthesis, such as the phosphopantetheinyl transferase for activation of fungal PKSs and NRPSs, and some redox partners of cytochrome P450 oxidoreductases. But *E. coli* usually cannot support it due to lack of relevant cofactors. Additionally, compared to fungal species, *E. coli* has a different codon usage bias and a higher GC content. Thereby, codon optimization of the heterologous gene is generally required before expressing in *E. coli*. Sometimes, production of fungal antibiotics in *E. coli* is also restricted due to high susceptibility.

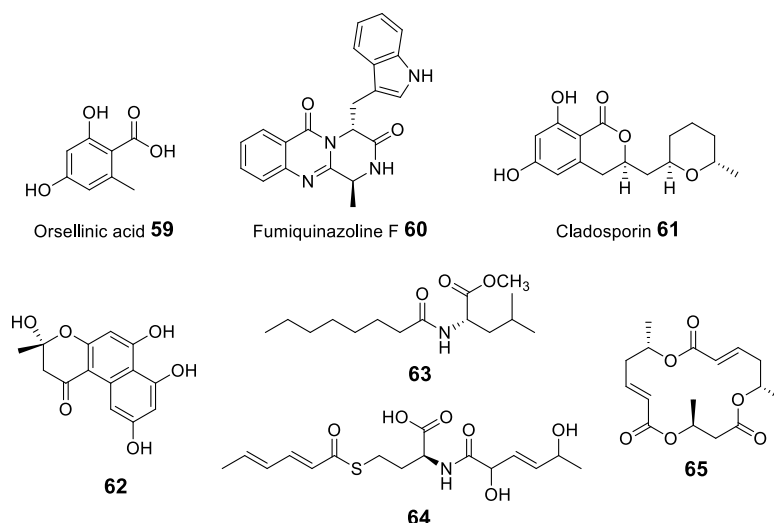
### 1.2.2.2 Fungal Systems

#### ***Saccharomyces cerevisiae***

Enzymes or secondary metabolites of fungal origin are also produced using the yeast *Saccharomyces cerevisiae*. It has various advantages in terms of heterologous expression. Similar to *E. coli*, yeast is a unicellular organism with rapid growth rate. Many genetic manipulation techniques have been developed for pathway engineering. In addition, yeast is generally regarded as a safe organism, as no toxic secondary metabolites was endogenously generated.<sup>45</sup> However this is also a disadvantage as yeast is not naturally optimized for the production of secondary metabolites.

Many gene clusters of fungal-origin have been employed in yeast for heterologous production (Fig. 1.9). For example, an engineered *Saccharomyces cerevisiae* host strain was established by Ishiuchi and coworkers. Fungal polyketide synthase or nonribosomal peptide synthetase genes in size up to 20 kb could be successfully expressed in this strain. Using this approach, secondary metabolites from various fungal species were produced (*e.g.* orsellinic acid **59** and fumiquinazoline F **60**).<sup>46</sup> Cladosporin **61**, derived from the fungus *Cladosporium cladosporioides*, is another excellent example of heterologous production in *S. cerevisiae*, confirming the identity of the putative gene cluster.<sup>47</sup> As the

aforementioned in pathway engineering, a yeast-based heterologous expression platform HEx was developed for the rapid and scalable discovery of fungal natural products.<sup>39</sup> By this high-throughput heterologous expression strategy in yeast, up to 22 detectable secondary metabolites were found (*e.g.* **62-65**), including novel compounds with unexpected biosynthetic origins, particularly from poorly studied fungal species.



**Figure 1.9** Examples of fungal natural products by heterologous production in yeast.

As a eukaryote, yeast can typically support more complex protein folding. However, in some cases post-translational modifications are still necessary when expressing proteins from higher eukaryotes. As with *E. coli*, yeast is not naturally geared for production of secondary metabolites. The lack of many building blocks also limits its application and intron splicing is problematic.

## Filamentous Fungi

Filamentous fungi are often chosen as heterologous hosts for production of natural products. Their cultivation conditions are very simple. The large-scale fermentation and isolation of natural products can also be achieved easily. Additionally, various genetic manipulations, transformation methods and selection systems are available. Since the discovery of penicillins **5**, the genus *Penicillium* has been deeply studied to produce a wide range of natural products.<sup>48</sup> For example, *P. chrysogenum* was applied as a heterologous host to express the gene cluster of mevastatin **12** from *P.*

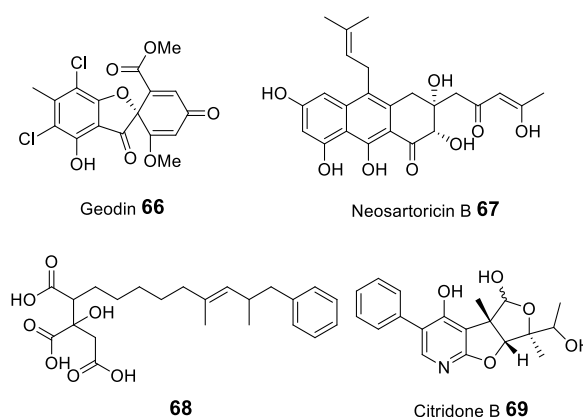
*citrinum*. With the gene *CYP105AS1*, which encodes a cytochrome P450, mevastatin **12** was converted to pravastatin **13** in *P. chrysogenum*.<sup>49</sup> Additionally fungi can often grow in complex media, or even on waste material, requiring little specialized equipment. However, a disadvantage of the use of fungi is that they grow and reproduce more slowly than bacteria and yeast, making engineering campaigns potentially time-consuming.

### ***Aspergillus* species**

*Aspergillus* species are the most extensively used fungal heterologous hosts.<sup>50,51</sup> Multiple efficient genetic transformation systems have been developed. For example, the protoplast-mediated transformation is commonly used for *A. oryzae*, *A. niger*, *A. sojae*, and *A. terreus* and the *Agrobacterium*-mediated transformation for *A. awamori*. Versatile selection marker systems including antibiotic markers *hph*, *ble*, *oliC3*, auxotrophic markers *pyrG*, *pyrE*, *argB*, *adeA*, *adeB*, *niaD*, *trpC*, and nutritional markers *amdS* and *ptrA* are available for high flexibility of genetic manipulation in *Aspergillus* strains. It was found that the homologous recombination frequencies can reach up to 100 %. This finding was confirmed in many *Aspergillus* species, such as *A. oryzae*, *A. sojae*, and *A. niger*. To maximize the heterologous production of natural products, many promoters are usable in *Aspergillus* species. The promoters used are either inducible promoters (*e.g.* *P<sub>amyB</sub>*, *P<sub>glaA</sub>*, and *P<sub>alcA</sub>*) or constitutive promoters (*e.g.* *P<sub>gpdA</sub>*, *P<sub>adh</sub>*, and *P<sub>eno</sub>*). All these factors assure *Aspergillus* a dominant place as platforms of heterologous production of natural products.

*A. nidulans* is one of the most efficient systems in filamentous fungi (Fig. 1.10).<sup>52,53</sup> For example, a simple PCR-based approach for heterologous reconstitution of intact gene clusters in *A. nidulans* by Nielsen and coworkers.<sup>54</sup> It allowed the transfer of 13 genes to *A. nidulans*, enabling the heterologous production of geodin **66**. By genome mining, a conserved gene cluster for biosynthesis of an immunosuppressive polyketide was found among dermatophytes. The heterologous expression of the gene cluster in *A. nidulans* led to the formation of neosartoricin B **67**, revealing its possible role in the pathogenesis of dermatophytes.<sup>55</sup> In a reconstitution study of biosynthetic pathway, it was found that only three enzyme, HRPKS Clz14, citrate synthase Clz17 and Clz11 could

achieve the heterologous production of a structurally complex compound **68** in *A. nidulans*.<sup>56</sup> The biosynthetic pathway of compounds in the citridone family (e.g. citridone B **69**) was elucidated through heterologous reconstitution of the *pfp* pathway in *A. nidulans*.<sup>48</sup> However *A. nidulans* also has disadvantages because it already produces many of its own secondary metabolites. This can complicate the purification of newly engineered compounds. In addition, the high availability of biosynthetic enzymes in this host means that shunt pathways are likely. Finally, the production of competing secondary metabolites also uses up valuable metabolic building blocks, reducing titers of desired engineered compounds.



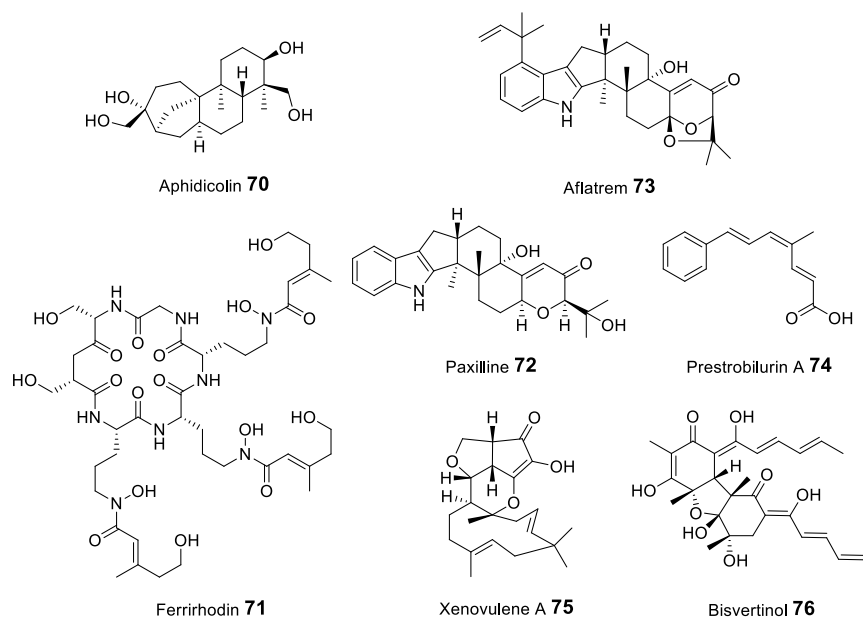
**Figure 1.10** Structures of fungal natural products heterologously produced in *A. nidulans*.

### ***Aspergillus oryzae***

*A. oryzae* has been widely used in many cases for heterologous expression. It is a GRAS organism, therefore much safer as a heterologous host for laboratorial manipulations. *A. oryzae* produces almost no secondary metabolites in standard fermentation conditions, avoiding the competition problems observed for *A. nidulans*, for example.

Many compounds have been produced by gene reconstitution in *A. oryzae* (Fig. 1.11). For example, the biosynthesis of the polyketide tenellin **22** from *Beauveria bassiana* was reconstituted in *A. oryzae*. It was the first example of heterologous expression of an entire fungal secondary metabolism gene cluster by promoter replacement.<sup>57</sup> The whole gene cluster for the biosynthesis of diterpene aphidicodin **70** was introduced into *A. oryzae*, achieving the heterologous production of **70**.<sup>58</sup> The heterologous expression of a nonribosomal peptide synthetase gene *FSN1* from

*Fusarium sacchari* in *A. oryzae* resulted in the accumulation of ferrirhodin **71**.<sup>59</sup> The indole-diterpene paxilline **72** is also a heterologous production example in *A. oryzae*. By stepwise introduction of *pax* genes into *A. oryzae*, the biosynthetic pathway of paxilline **72** was elucidated.<sup>60</sup> Using this stepwise determination strategy of biosynthetic pathways, a rapid introduction method of heterologous genes was established in *A. oryzae*, enabling the total biosynthesis of aflatrem **73**.<sup>61</sup>

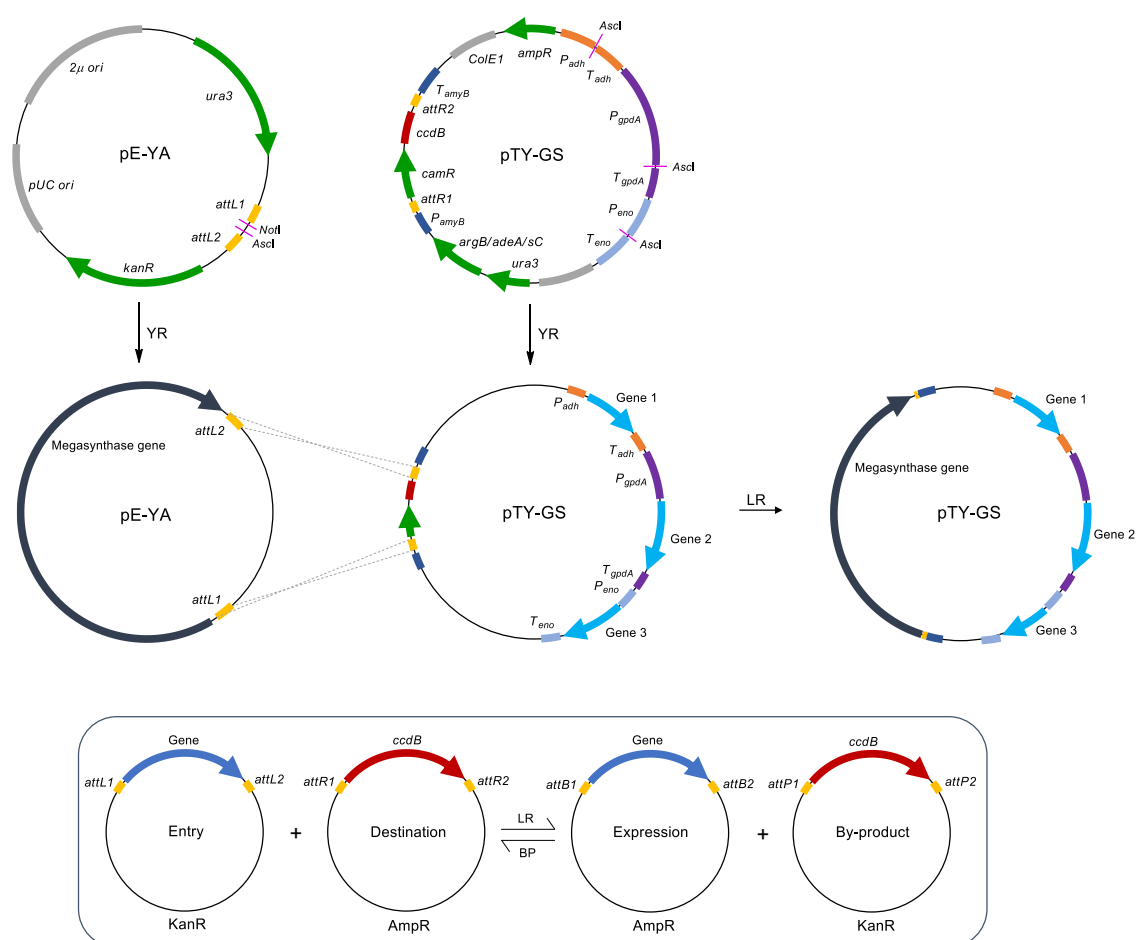


**Figure 1.11** Structures of fungal natural products heterologously produced in *A. oryzae*.

## Lazarus System

Notably, a versatile heterologous expression toolkit was described in 2012 by Lazarus and coworkers.<sup>62</sup> As shown in Scheme. 1.5, it comprises an assembly vector pE-YA for the construction of megasynthase genes and three pTY-GS vectors for multigene expression. The reconstruction of one megasynthase and three tailoring enzyme genes by yeast recombination<sup>63</sup> and Gateway<sup>®</sup> LR recombination in the expression vector can be achieved in a little more than one week. Three selection markers *argB*, *adeA*, *sC* and *bar* are available on separate pTY-GS vectors. If all markers are used simultaneously, it allows up to 16 genes to be coexpressed. The heterologous host is a quadruple auxotrophic strain *A. oryzae* NSAR1 in which *argB*, *adeA*, *sC* and *niaD* are all deleted. The corresponding auxotrophic selection medium lacking arginine, adenine, methionine, and nitrate was used for transformant screening. This toolkit facilitates the rapid assembly and reconstitution of the whole

biosynthetic pathway. Meanwhile, the use of strong promoters, including the inducible promoter  $P_{amyB}$  and the constitutive promoters  $P_{adh}$ ,  $P_{gpdA}$ , and  $P_{eno}$ , promotes high-level expression of all heterologous genes in *A. oryzae*.



**Scheme 1.5** Reconstruction of multigene expression vector for heterologous expression of natural products in *A. oryzae* NSAR1. The inset sketches the Gateway® reactions. Abbreviations: YR, yeast recombination; LR, Gateway® LR recombination; KanR, kanamycin resistance; AmpR, ampicillin resistance.

Using this heterologous expression toolkit, more natural products have been produced in *A. oryzae* (Fig. 1.11). For example, the reconstruction of the strobilurin gene cluster in *A. oryzae* led to the formation of prestrobilurin A **74**.<sup>64</sup> Heterologous reconstruction of the entire xenovulene A **75** biosynthetic pathway in *A. oryzae* allowed to elucidate the biosynthetic steps of **75**.<sup>65</sup> The biosynthetic pathway of sorbicillinoids from *Trichoderma reesei* was also reconstituted in *A. oryzae*. It resulted in the formation of a series of dimeric sorbicillinoids (e.g. bisvertinol **76**).<sup>66</sup>

*A. oryzae* can not only express heterologous complex enzymes (e.g. PKS and NRPS), but also generate products in high yields (Table 1). Therefore, *A. oryzae* is a

potent heterologous host. Additionally, it is a powerful host for refactoring of metabolic pathways and consequent production of novel compounds by combinatorial biosynthesis.<sup>45</sup>

**Table 1** Reported yields of bioactive fungal natural products by heterologous production in *A. oryzae*.<sup>45</sup>

Natural product	Native host	Bioactivity	Reported yield
Citrinin <b>6</b>	<i>Monascus purpureus</i>	Antibacterial	1.48 mg·L <sup>-1</sup>
Tenellin <b>22</b>	<i>Beauveria bassiana</i>	Inhibitor of erythrocyte membrane ATPase activity; iron chelator	234 mg·L <sup>-1</sup>
Pleuromutilin <b>52</b>	<i>Clitopilus passeckerianus</i>	Antibacterial	84 mg·L <sup>-1</sup>
Aphidicodin <b>70</b>	<i>Phoma betae</i>	Inhibitor of DNA polymerase $\alpha$	0.33 mg·L <sup>-1</sup>
Ferrirhodin <b>71</b>	<i>Fusarium sacchari</i>	Iron chelator	2.4 mg·L <sup>-1</sup>
Paxilline <b>72</b>	<i>Penicillium paxilli</i>	Inhibitor of the high-conductance calcium-activated potassium channel; antibacterial	35 mg·L <sup>-1</sup>
Aflatrem <b>73</b>	<i>Aspergillus flavus</i>	Tremorgenic	54 mg·Kg <sup>-1</sup>
Prestrobilurin A <b>74</b>	<i>Strobilurus tenacellus</i>	-	20 mg·L <sup>-1</sup>

### 1.3 Overall Aims

It is known that fungi mainly produce four types of secondary metabolites, including polyketide, nonribosomal peptides, terpenoids and alkaloids. Based on the bioactivities these metabolites have displayed, more structurally diverse natural products are worth to mining further in fungi. The engineering biosynthesis *in vivo* using synthetic biology techniques is an effective strategy. Compared to global transcription regulation and enzyme engineering, the pathway engineering is more predictable and specific. Among the pathway engineering strategies, heterologous expression is a relatively more efficient approach, especially such as the filamentous fungus *A. oryzae*, which has been widely used in many cases for heterologous expression of fungal iterative PKSs.

While the use of fungal hosts for the expression of fungal metabolites is well established, the scope of fungal hosts has not been investigated further. For example, fungal heterologous expression systems have not been investigated for the expression of bacterial modular PKSs. Thus, the main focus of this thesis concentrates on further expanding the application of *A. oryzae* as a heterologous

host, achieving the discovery of more diverse natural products. Specifically, we aimed to express a modular PKS from bacteria in a fungal host.

In particular, workflows pioneered in the Prof. Cox group, involving gene cloning and homologous recombination in yeast, followed by rapid transformation and expression in *A. oryzae* allow efficient engineering of complex genes. Several iterative PKSs have been dramatically re-engineered, such as PKSs responsible for biosynthesis of tenellin **22** and desmethylbassianin **23**.<sup>67</sup> Fungi are eukaryotic and use monocistronic operons in which each gene can be independently controlled. This may offer advantages over the use of polycistronic operons in bacteria where multiple genes are controlled from a single promoter. All these advantages provide the support for the expression of a heterologous modular PKS in *A. oryzae*.

Additionally, fungi are internally more complex than bacteria. Numerous internal compartments such as peroxisomes where selective reactions could be catalyzed are present in fungi. Fungi also possess diverse tailoring enzymes and can catalyze a wide variety of oxidative rearrangements unavailable in bacteria. Thus, fungi could be very interesting hosts for the modification of bacterial polyketides. However, to-date modular PKSs are unknown in fungi, and it is unsure whether suitable precursors could be provided.

Combinatorial biosynthesis strategy will be applied within *A. oryzae* heterologous expression system. For example, the Lazarus system allows up to 16 genes to be expressed, and thereby a re-engineered biosynthetic pathway of bacterial polyketide can be artificially reconstructed. If it is feasible, the fusion expression of a modular PKS and an iterative PKS will be an interesting attempt to understand if two distinct types of PKS could be combined into one enzyme and produce more diverse derivatives. More novel natural products will be generated in this new and promising *A. oryzae* heterologous expression system.

Another aim in this thesis is to produce novel natural products and identify their biosynthetic pathways by *A. oryzae* heterologous expression. In all known brasilane-type sesquiterpenes there are only two brasilane-type glycosides. Therefore, it must have a rare type of glycosyltransferase capable of recognizing sesquiterpene aglycones as substrate, and this deserves more attention. Thus, we specifically aimed to identify a biosynthetic gene cluster responsible for brasilane-



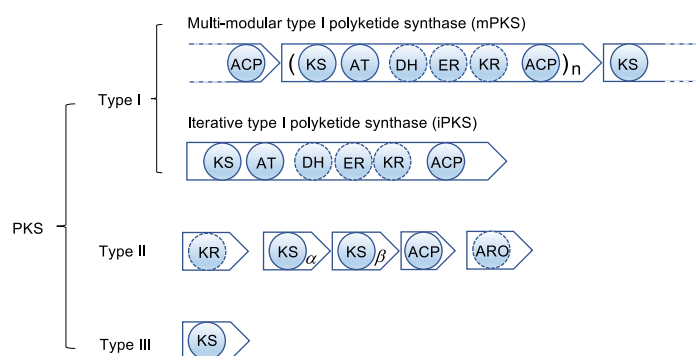
type sesquiterpene biosynthesis through the heterologous expression strategy in *Aspergillus oryzae*. For each putative gene in the gene cluster, we aimed to figure out its corresponding biochemical function during the production of brasilane-type sesquiterpene.

## 2. Heterologous Expression of Modular PKS DEBS1-TE in *Aspergillus oryzae*

### 2.1 Introduction

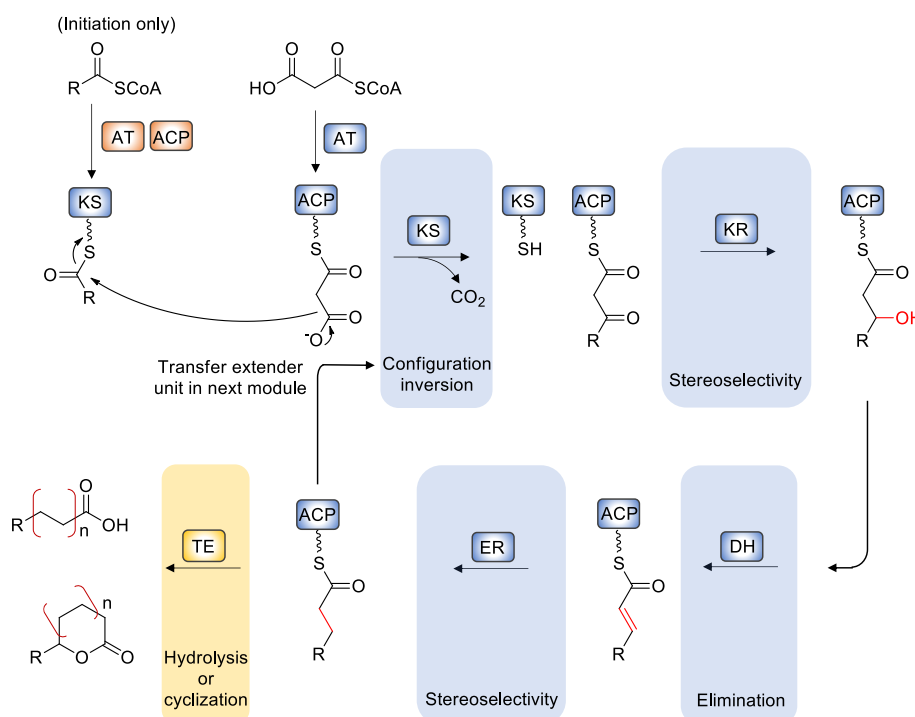
Polyketide synthases (PKS) are one of the most important class of biosynthetic enzymes. They generally catalyze the initial construction of polyketides through successive rounds of Claisen condensations of short-chain fatty acids. Types of polyketide synthase, numbers of condensation reaction, varieties of building block, and programmed  $\beta$ -processing of polyketide intermediates and post-PKS tailoring reactions (*e.g.* reduction, alkylation, or cyclization) lead to their enormous structural and functional diversities.<sup>68</sup>

Based on their structure and architecture, PKS are classified into type I, type II, and type III systems (Fig. 2.1).<sup>69</sup> Type I PKS consist of large multifunctional proteins with individual functional domains which are covalently linked. A *module* is defined as a single polypeptide which contains all the catalytic activities required to extend and process a growing polyketide chain. Type I PKS can be further classified into iterative type I PKS (iPKS) which consist of a single module, and multi-modular type I PKS (mPKS). The type I iPKS primarily exist in fungi (just occasionally in bacteria). In contrast, type I mPKS predominantly occur in bacteria (*e.g.* actinobacteria, myxobacteria, pseudomonades, and cyanobacteria). Type II PKS consist of separate proteins which form non-covalent complexes mostly found in bacteria, while type III PKS are very simple systems found in plants, bacteria and fungi but which do not use an ACP. Both Type II and Type III PKS are iterative.



**Figure 2.1** Classification of polyketide synthases (PKS).

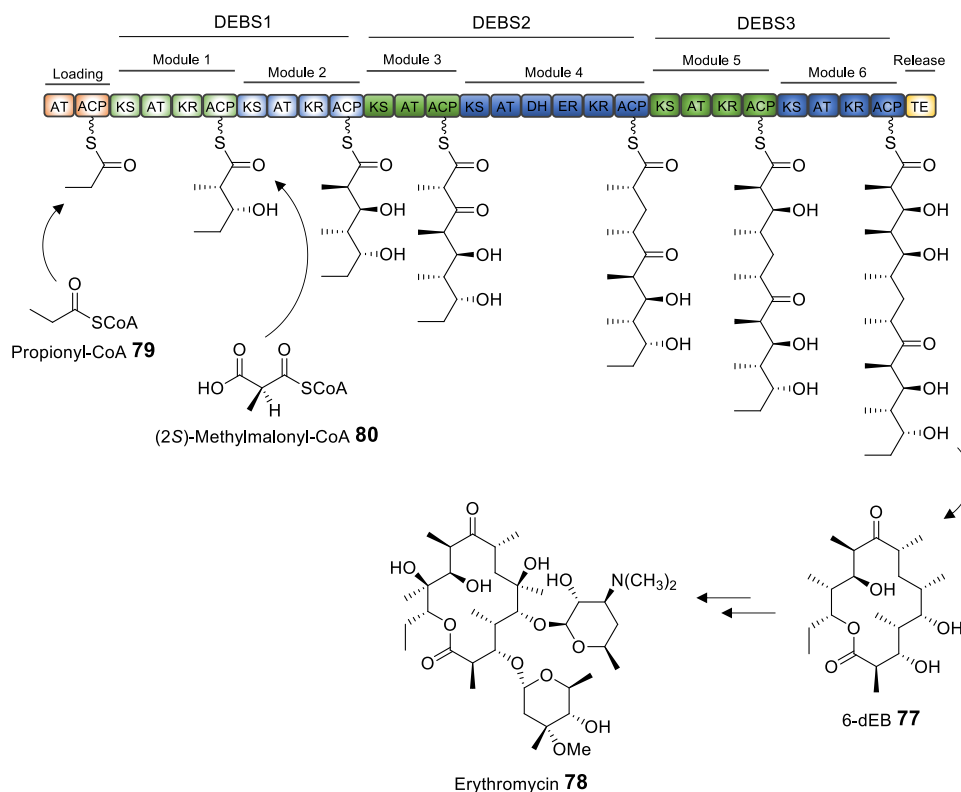
For type I PKS, one module is comprised of a minimal set of catalytic domains such as  $\beta$ -ketosynthase (KS), acyltransferase (AT) and acyl carrier protein (ACP) which achieve one round of substrate loading and condensation (Scheme 2.1).<sup>70,71</sup> Different types of PKS (and different modules of mPKS) can also contain varying numbers of enzymes that can modify the newly-formed  $\alpha$ - and  $\beta$  positions. Distinguishingly, iPKS utilize the same catalytic domains repetitively to generate the entire polyketide scaffold, while the mPKS work like assembly lines, in which one module catalyzes a single extension and processing cycle.



**Scheme 2.1** General polyketide biosynthetic cycles. Abbreviations: AT, acyltransferase; ACP, acyl carrier protein; KS,  $\beta$ -ketosynthase; KR,  $\beta$ -ketoreductase; DH, dehydratase; ER, enoyl reductase; TE, thioesterase.

A notable example of a modular PKS is 6-deoxyerythronolide B synthase (DEBS), which produces the macrolactone metabolite 6-deoxyerythronolide B (6-dEB, **77**) in *Saccharopolyspora erythraea*. The polyketide 6-dEB forms the backbone of erythromycin **78** after multiple post-polyketide modifications. DEBS consists of a *loading module* and six *extending modules* organized into three separate polypeptides known as DEBS1, DEBS2 and DEBS3 (Scheme 2.2). In DEBS there are two extending modules in each of these separate polypeptides, but in other systems this can vary considerably.

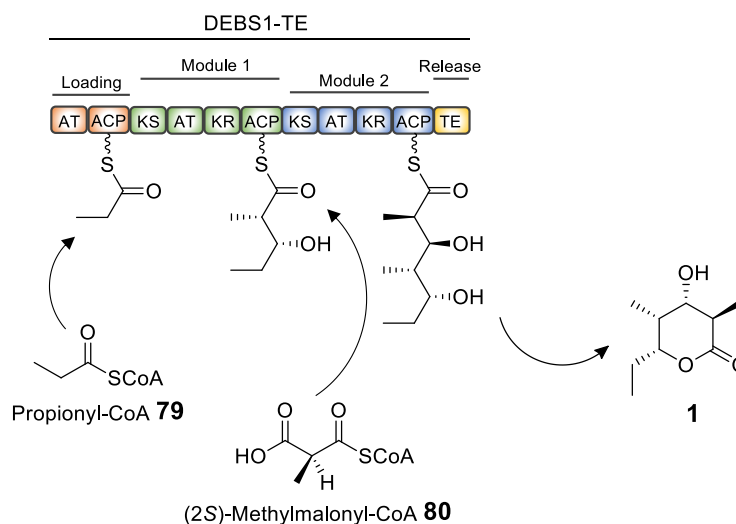
The assembly process of 6-dEB starts with loading of the *starter unit* propionyl-CoA **79**, followed by the incorporation of six elongating (2*S*)-methylmalonyl-CoA **80** *extender units*. Lastly, the assembled polyketide conducts a lactonization catalyzed by the thiolesterase (TE) domain to result in the production of 6-dEB (Scheme 2.2).<sup>72,73</sup> Notably, each module also controls the stereochemistry of the growing polyketide at the  $\alpha$ - and  $\beta$ -positions because the  $\beta$ -processing KR, DH and ER domains have their own stereoselectivity.



**Scheme 2.2** Domain organization of the 6-deoxyerythronolide B synthase (DEBS) and structures of its product 6-deoxyerythronolide B (6dEB, **77**) and the final product erythromycin **78** after post-modifications. All chain extension steps are stepwise shown attached to ACP domains, including stereochemistry of intermediates.

Since the organization of modular PKS was elucidated, repositioning of domains has become a significant way of attempting to produce various new polyketide structures. DEBS1-TE is one of the most successful examples. It is a fusion protein, which is comprised of the loading module, the first two extension modules (*i.e.* DEBS1), and the terminal TE domain of DEBS3. It had been proven that DEBS1-TE catalyzes starter-unit loading and the first two rounds of polyketide chain extension followed by TE-catalyzed lactonization. This produces a triketide lactone **1** in *Saccharopolyspora erythraea* (10~15 mg·L<sup>-1</sup>) and other bacterial and yeast hosts

(Scheme 2.3), indicating the structural integrity of this chimeric PKS is not dependent on interactions with the rest of DEBS.<sup>74,75</sup> DEBS1-TE requires propionyl-CoA **79** as the starter unit, and two equivalents of (2*S*)-methylmalonyl-CoA **80** for the extensions. DEBS1-TE is often used as a model system as its convenient size and limited chemistry make the results of engineering experiments simpler to interpret.



**Scheme 2.3** Domain organization of DEBS1-TE and the product structure of triketide lactone **1**.

Although the reconstituted DEBS1-TE could produce a triketide lactone **1** in the native strain *S. erythraea*, the challenges associated with scalable and feasible production are still the serious bottleneck, especially for the *Actinomyces* family of bacteria. Therefore, to obtain higher yield and more variety of polyketides, other well-developed heterologous hosts have been adopted. In this chapter, an overall review on heterologous expression of DEBS1-TE is summarized.

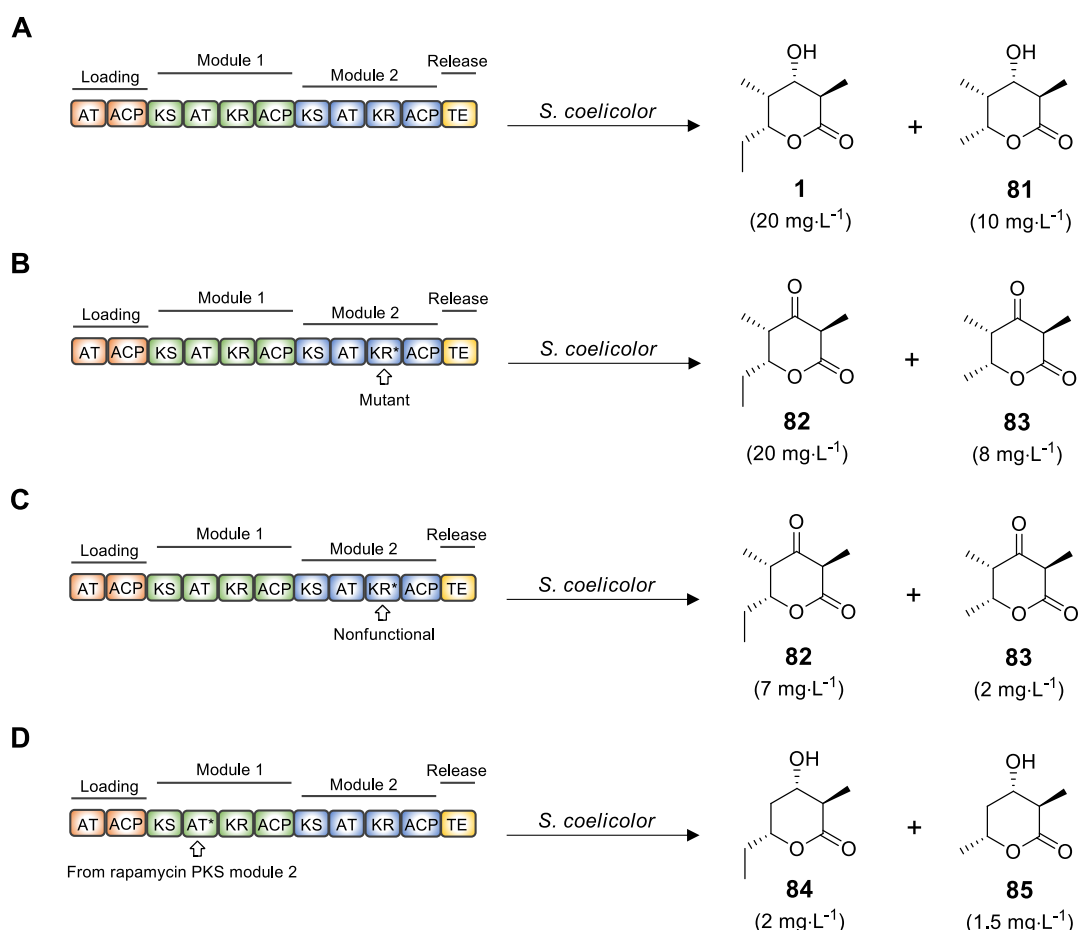
### 2.1.1 Heterologous Expression of DEBS1-TE in Bacteria

#### *Streptomyces coelicolor*

*Streptomyces coelicolor*, a well-developed heterologous host, was used to express DEBS1-TE in very early experiments. In *S. coelicolor*, the presence of propionyl-CoA synthetase and propionyl-CoA carboxylase (PCC) have been validated by isotopic labeling of propionic acid.<sup>76</sup> By this biosynthetic pathway, exogenous propionate can be used to successively generate propionyl-CoA **79** and (2*S*)-methylmalonyl-CoA **80** *in vivo*. In addition to the PCC pathway for formation of (2*S*)-methylmalonyl-CoA **80**,

a methylmalonyl-CoA mutase-epimerase pathway is also present in *S. coelicolor* to provide another source of (2*S*)-methylmalonyl-CoA **80** (from succinyl-CoA) for polyketide biosynthesis. Furthermore, *S. coelicolor* also contains an appropriate phosphopantetheinyl transferase (PPTase), which converts inactive *apo*-ACP to active *holo*-ACP prior to polyketide biosynthesis. Therefore, proper post-translational modification of DEBS1-TE can occur in the intracellular environment of the heterologous host *S. coelicolor*.<sup>76</sup>

In 1995, Kao and coworkers expressed DEBS1-TE in *S. coelicolor* for the first time.<sup>77</sup> The triketide lactone **1** was produced at a titer of 20 mg·L<sup>-1</sup> (Scheme 2.4 A). Meantime, a novel analog **81** was also produced (10 mg·L<sup>-1</sup>). Its structure was determined by NMR characterization. It revealed that **81** arises from the incorporation of an acetyl-CoA **14** starter unit instead of propionyl-CoA **79**, implying that DEBS1-TE can also recognize other acyl-CoAs as precursor.



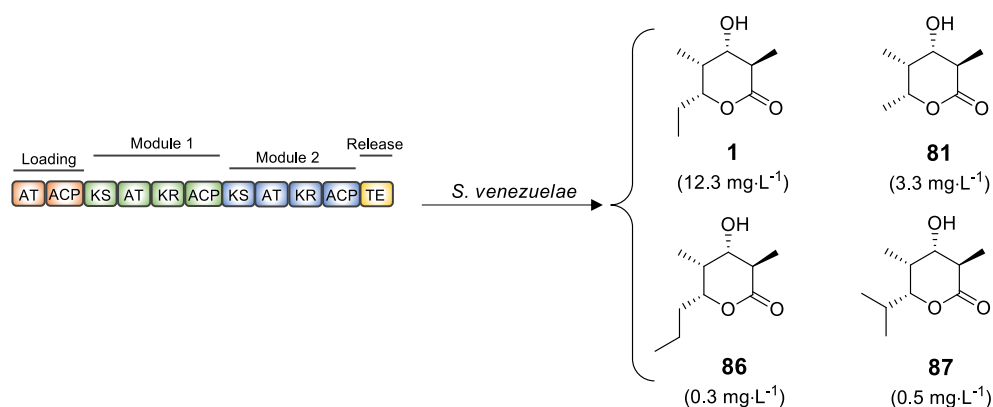
**Scheme 2.4** Expression of DEBS1-TE and its derivatives in *S. coelicolor*: **A**, DEBS1-TE and its triketide lactone products **1** and **81**; **B**, KR mutant of DEBS1-TE and its products **82** and **83**; **C**, KR3\* substitution derivative of DEBS1-TE and its products **82** and **83**; **D**, AT substitution derivative of DEBS1-TE and its products **84** and **85**. The number in brackets is their respective titer for each compound.

To alter PKS function, the domain mutation and substitution of DEBS1-TE were also conducted in *S. coelicolor*.<sup>78</sup> The reductive segment of module 2, which includes a functional  $\beta$ -ketoreductase (KR) domain, was inactivated by a random PCR mutagenesis (Scheme 2.4 B). The recombinant *S. coelicolor* strain, bearing the mutant DBES1-TE, produced the predicted triketide ketolactone **82** (20 mg·L<sup>-1</sup>) and **83** (8 mg·L<sup>-1</sup>), which differs from **1** and **81** in the degree of reduction at C-3. Similarly, when the reductive segment of module 2 was replaced by the nonfunctional KR from module 3 of DEBS, the recombinant strain *S. coelicolor* also produced **82** (7 mg·L<sup>-1</sup>) and **83** (2 mg·L<sup>-1</sup>) in smaller quantities (Scheme 2.4 C).

DEBS1-TE was also engineered in *S. coelicolor* through replacing the entire acyltransferase (AT) domain from module 1 with the malonyl-CoA specific AT domain from module 2 of rapamycin PKS (Scheme 2.4 D).<sup>79</sup> It led to the formation of two new triketide lactones **84** (2 mg·L<sup>-1</sup>) and **85** (1.5 mg·L<sup>-1</sup>), both of which specifically lack a methyl group at C-4.

### *Streptomyces venezuelae*

The pikromycin-producing strain *Streptomyces venezuelae* was also used as a heterologous host to express DEBS1-TE when exploring the ability of pikromycin PKS to simultaneously generate 12- and 14-membered ring macrolides.<sup>80</sup> DEBS1-TE was expressed in a *S. venezuelae* PikAI deletion mutant and four different triketide lactones **1**, **81**, **86**, and **87** were produced in detectable yields (Scheme 2.5). This result differs from that in *S. coelicolor*, suggesting that more kinds of starter units can be incorporated into this biosynthetic pathway in *S. venezuelae*.



**Scheme 2.5** Expression of DEBS1-TE and its products in *S. venezuelae*. The number in brackets is their respective titer for each compound.

## ***Escherichia coli***

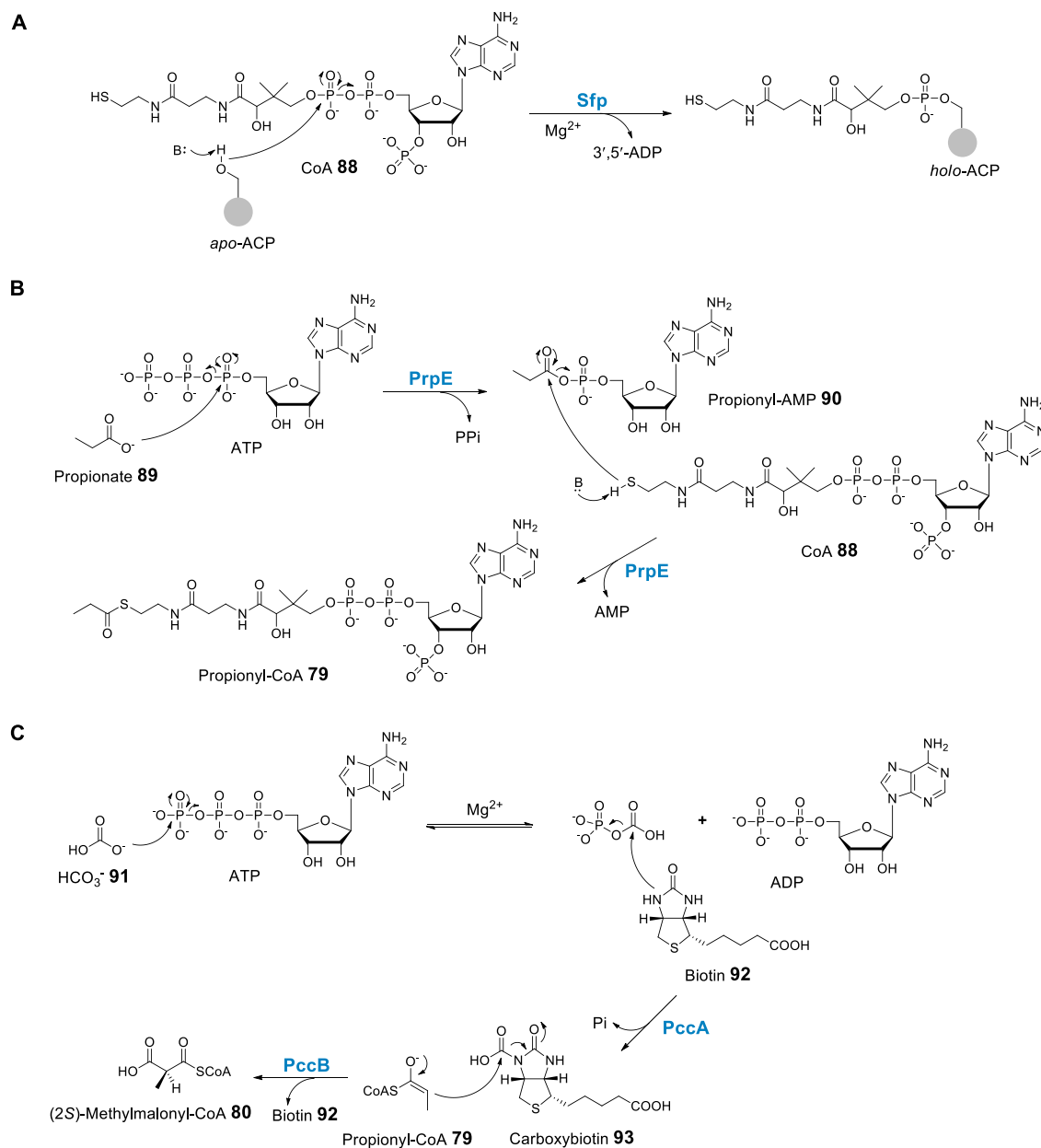
*E. coli* is widely used for gene cloning and protein expression. It has a good ability to grow quickly and molecular methods are highly developed. Many proteins can be produced in *E. coli* at high levels and in active form. However, *E. coli* encounters three major challenges for the expression of DEBS1-TE.<sup>75</sup> The first is the capability of expressing correctly folded and post-translationally modified DEBS1-TE *in vivo*. ACP domains of DEBS1-TE must be post-translationally pantetheinylated prior to functioning as an active form. Second, the building blocks for polyketide production must be available in *E. coli*. However, the formation conditions of starter unit propionyl-CoA **79** are incompatible with typical high cell density fermentation conditions. In addition, the extender unit (2*S*)-methylmalonyl-CoA **80** has not been found as a metabolite in *E. coli* to date. Thereby, new metabolic pathways are required to be reengineered to synthesize precursors *in vivo*. Lastly, in order to maximize yield, the enzymatic activity DEBS1-TE needs to be well-synchronized with precursor biosynthesis.

Despite these challenges, an engineered strain *E. coli* BAP1 for DEBS1-TE expression was established by Pfeifer and coworkers.<sup>75</sup> First, it was confirmed that *E. coli*, harboring an exogenous phosphopantetheinyl transferase Sfp (mechanism shown in Scheme 2.6 A) from *Bacillus subtilis*, has the ability of expressing active DEBS1-TE protein bearing *holo*-ACP domains.<sup>81,82</sup> The protein activity was validated by *in vitro* assays. The protein expression level of DEBS1-TE is even comparable to that in the producing strain *S. erythraea*.

In order to accumulate a high level of starter unit propionyl-CoA **79** in *E. coli*, the endogenous propionyl-CoA synthetase PrpE (mechanism shown in Scheme 2.6 B) was overexpressed in *E. coli*.<sup>83,84</sup> By integrating the *sfp* gene into the *E. coli* native *prp* operon, which also encodes propionate catabolism (PrpC, 2-methylcitrate synthase; PrpD, methylisocitrate synthase; PrpB, 2-methylisocitrate lyase), the ability of *E. coli* to utilize propionate as a carbon source was eliminated. In other words, this effort not only blocked the endogenous consumption of propionyl-CoA **79** but also increased its accumulation arising from exogenous propionate **89**. Subsequently, based on propionyl-CoA **79** production, a formation pathway of extender unit (2*S*)-methylmalonyl-CoA **80** was engineered by introducing *S.*



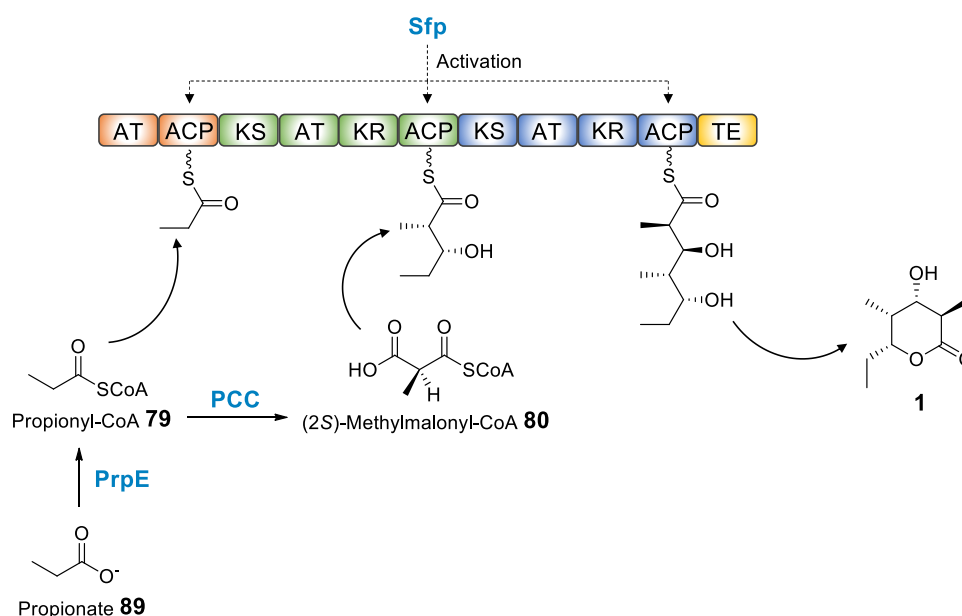
*coelicolor*-derived propionyl-CoA carboxylase (PCC). This enzyme consists of  $\alpha$  and  $\beta$  subunits, known as PccA and PccB, which together catalyze the conversion of propionyl-CoA **79** to (2*S*)-methylmalonyl-CoA **80** *in vivo* (mechanism shown in Scheme 2.6 C).<sup>85,86</sup>



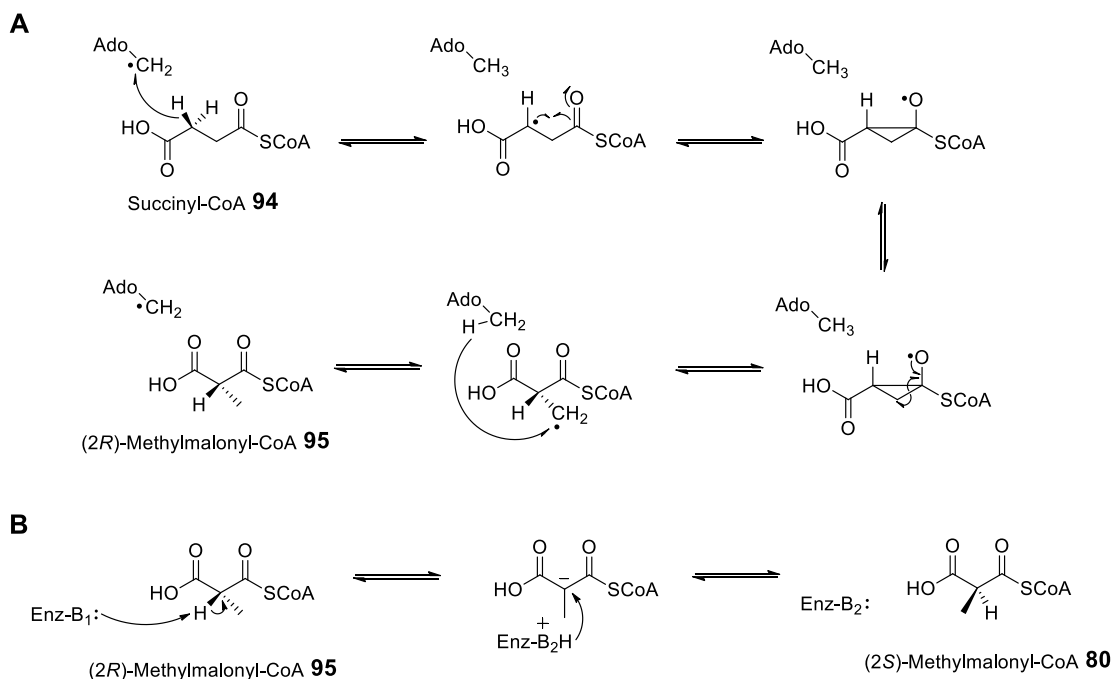
**Scheme 2.6** Enzyme-catalyzed reaction mechanisms: **A**, the activation of ACP from *apo*- to *holo*-form by PPTase Sfp with CoA **88**; **B**, the conversion from propionate **89** to propionyl-CoA **79** catalyzed by PrpE through the intermediate propionyl-AMP **90**; **C**, the (2*S*)-methylmalonyl-CoA **80** formation from propionyl-CoA **79**, carbonate **91** and ATP catalyzed by PCC through reacting with biotin **92** and the intermediate carboxybiotin **93**.

Ultimately, the engineered *E. coli* strain carrying DEBS1-TE and other ancillary genes produced the expected triketide lactone **1** (Scheme 2.7). When it was used for

DEBS expression, the cellular production concentration of 6dEB **77** reached up to  $23 \text{ mg}\cdot\text{L}^{-1}$ .<sup>75</sup>

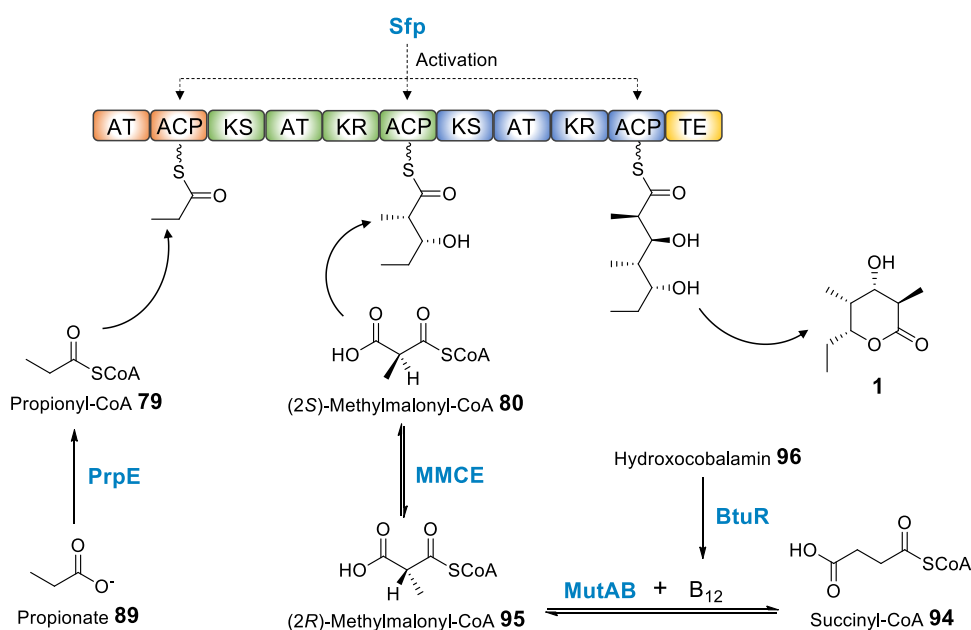


In 2002, the coenzyme B<sub>12</sub>-dependent methylmalonyl-CoA mutase-epimerase pathway was built for (2*S*)-methylmalonyl-CoA **80** production in *E. coli*.<sup>87</sup> It is an important pathway for methylmalonyl-CoA production in the polyketide-producing actinomycetes.<sup>88</sup> In contrast to the PCC pathway, in which both the starter unit and extender unit derive from exogenous propionate **89**, the extender unit of mutase-epimerase pathway stems from succinyl-CoA **94**, an intermediate in the tricarboxylic acid (TCA) cycle. Of two enzymes in this pathway, MutAB is a coenzyme B<sub>12</sub>-dependent enzyme from *Propionibacterium shermanii* and is comprised of two subunits MutA and MutB. Prior use, the B<sub>12</sub> precursor hydroxocobalamin **96** must be fed to the fermentation to produce coenzyme B<sub>12</sub> by a native adenosyltransferase BtuR in *E. coli*. MutAB along with coenzyme B<sub>12</sub> then converts succinyl-CoA **94** to (2*R*)-methylmalonyl-CoA **95** (mechanism shown in Scheme 2.8 A), which is then epimerized to the PKS extender unit (2*S*)-methylmalonyl-CoA **80** by an exogenous epimerase MMCE (mechanism shown in Scheme 2.8 B) from *P. shermanii*.<sup>89,90</sup>



**Scheme 2.8** Mechanisms of reactions catalyzed by enzymes: **A**, the conversion from succinyl-CoA **94** to (2R)-methylmalonyl-CoA **95** by the mutase MutAB (5'-deoxyadenosyl radical from coenzyme B<sub>12</sub> initiates the reaction by abstraction of a hydrogen atom from the methylene group of succinyl-CoA **94**); **B**, the interconversion of (2R)-methylmalonyl-CoA **95** and (2S)-methylmalonyl-CoA **80** catalyzed by the epimerase MMCE.

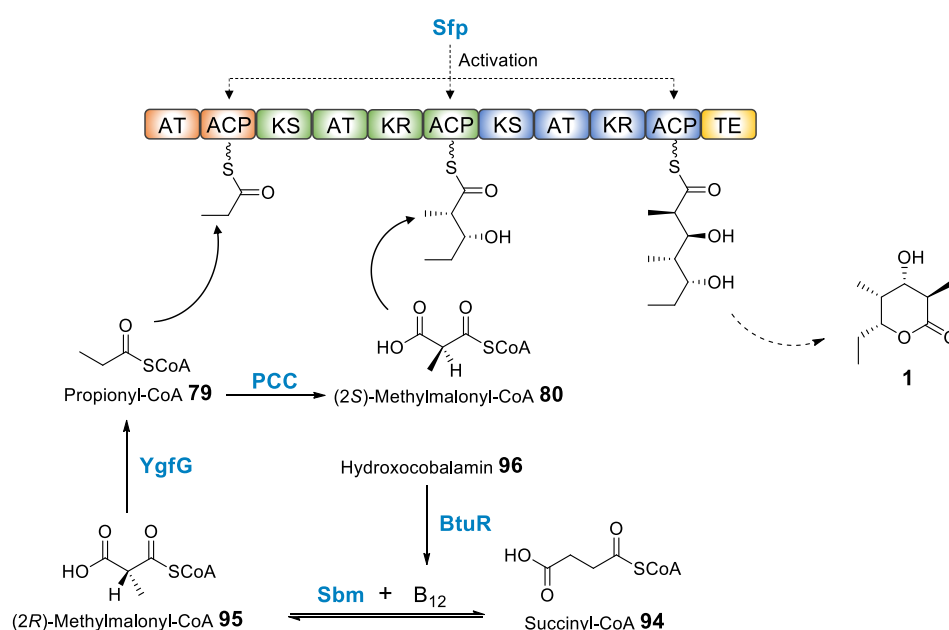
In the presence of Sfp and PrpE, triketide lactone **1** was successfully produced in *E. coli* (Scheme 2.9). When it was used to express DEBS, average 6-dEB **77** titers for this system were  $\sim 1 \text{ mg}\cdot\text{L}^{-1}$ .



**Scheme 2.9** Engineered DEBS1-TE heterologous expression via the mutase-epimerase pathway in *E. coli*.

In this pathway, there is one point of note. When the starter unit is supplied exogenously, the source of starter unit and extender unit can be decoupled, allowing unusual starter units to be incorporated into the polyketide assembly line. Therefore, through this mutase-epimerase pathway, the supply of extender unit can be independent of the start unit production, making separate optimization feasible.

In both of the pathways described above, exogenous propionate **89** was fed as the source of starter unit for polyketide biosynthesis. It appears that exogenous propionate is indispensable for heterologous polyketide production in *E. coli*. However, a later study found that *E. coli* supports low-level polyketide production without the addition of exogenous propionate.<sup>91</sup> Although DEBS was investigated as the object in the study, this proposal is also rational for DEBS1-TE. It was determined that native enzymes of *E. coli* including Sbm (a methylmalonyl-CoA mutase) and YgfG (a methylmalonyl-CoA decarboxylase) are also involved in propionyl-CoA and methylmalonyl-CoA metabolism (Scheme 2.10).



**Scheme 2.10** Proposed DEBS1-TE heterologous expression in *E. coli* in absence of exogenous propionate **89**.

In the absence of exogenous propionate, propionyl-CoA **79** is formed through a native metabolic pathway. First, succinyl-CoA **94** (from the TCA cycle) is converted to (2*R*)-methylmalonyl-CoA **95** by the native mutase Sbm and then to propionyl-CoA **79** by the native decarboxylase YgfG. Lastly, (2*S*)-methylmalonyl-CoA **80** is produced by an exogenous carboxylase PCC. Therefore, it is thought that the ability

of Sbm and YgfG to provide propionyl-CoA **79** gains dominance during polyketide production because of the lack of exogenous propionate **89**. Even though two precursors are produced *in vivo*, the yield of polyketide remains very low by this native pathway. When DEBS was expressed in this system, the titer of 6-dEB **77** was only 0.11 mg·L<sup>-1</sup>.

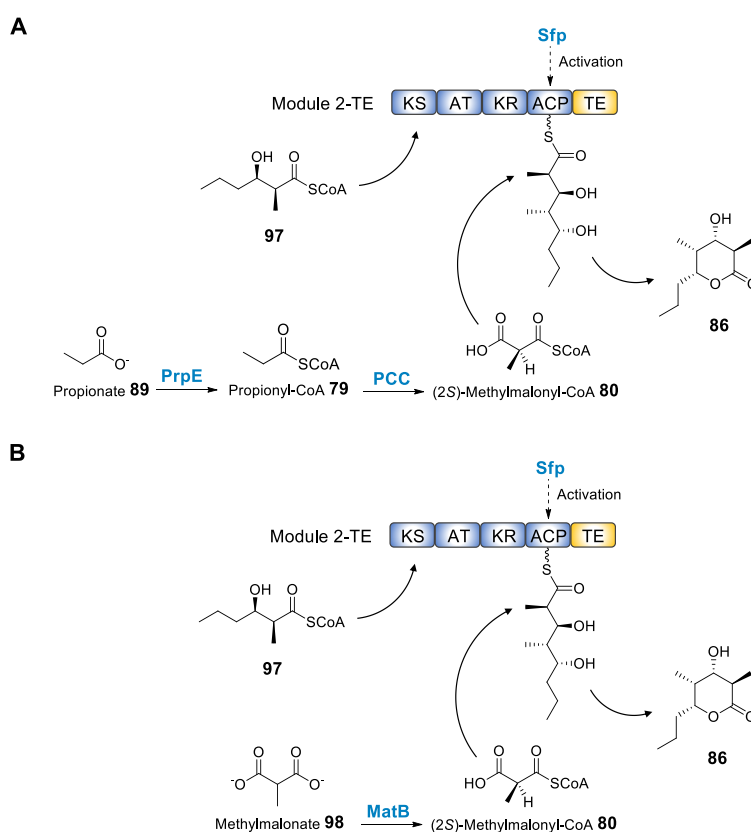
### 2.1.2 Heterologous Expression of Partial DEBS1-TE in Yeast

*Saccharomyces cerevisiae* is an ideal heterologous host for producing polyketides due to its rapid growth and fermentation as well as highly developed genetic manipulation. So far, *S. cerevisiae* has mostly been applied for heterologous expression of iterative PKS. For example, the expression of 6-methylsalicylic acid synthase (6-MSAS) along with a phosphopantetheinyl transferase Sfp in *S. cerevisiae* led to the production of a simple fungal polyketide 6-methylsalicylic acid (6-MSA, **58**) with a remarkable titer of 1.7 g·L<sup>-1</sup>.<sup>44</sup> Since that, engineering *S. cerevisiae* to produce more various types of complex polyketides has drawn attention.

Expressing modular PKS in *S. cerevisiae* is one of these challenges. However, in general, modular PKS genes derived from actinomycetes have highly rich GC content. Given that yeast codon usage is biased more towards AT, yeast may have inadequate tRNAs that are needed for modular PKS gene expression. Modular PKS and other ancillary genes may require codon optimization ahead of heterologous expression. Furthermore, *S. cerevisiae* lacks a suitable endogenous phosphopantetheinyl transferase which can activate the ACP domains of modular PKS. As shown in *E. coli*, functional expression of an exogenous modular PKS requires the Sfp coexpression for activity.<sup>75</sup> Furthermore, yeast is unable to produce some necessary acyl-CoA polyketide precursors, such as propionyl-CoA and methylmalonyl-CoA. For production of modular polyketides, heterologous acyl-CoA pathways may be required to supply sufficient levels of substrates. For example, the aforementioned expression of DBES1-TE in *E. coli* required the engineering of pathways for propionyl-CoA **79** and (2S)-methylmalonyl-CoA **80** biosynthesis.<sup>75,87</sup>

In 2006, an engineered *S. cerevisiae* strain for modular PKS expression was developed by Mutka and coworkers.<sup>92</sup> A protease-deficient strain *S. cerevisiae* BJ5464 was adopted, and genes of five rare tRNAs were introduced into it.<sup>93</sup> For

(2*S*)-methylmalonyl-CoA **80** production, two different routes were built in *S. cerevisiae*, respectively. The first one is a propionyl-CoA-dependent route which is similar to that built in *E. coli*. The propionyl-CoA synthetase gene *prpE* from *Salmonella typhimurium* and the propionyl-CoA carboxylase (PCC) pathway from *Streptomyces coelicolor* were introduced into *S. cerevisiae*.<sup>83,85</sup> The production of (2*S*)-methylmalonyl-CoA **80** stems from exogenous propionate **89** and endogenous ATP. The other route, that is independent of propionyl-CoA **79**, is the malonyl/methylmalonyl-CoA ligase (MatB) pathway from *S. coelicolor*, through which the exogenous methylmalonate **98** is converted to (2*S*)-methylmalonyl-CoA **80**.<sup>94</sup> In addition, the *sfp* phosphopantetheinyl transferase gene from *Bacillus subtilis* was also introduced into *S. cerevisiae*.<sup>81</sup> To maximize polyketide production, all genes required codon optimization before introduction into *S. cerevisiae*.

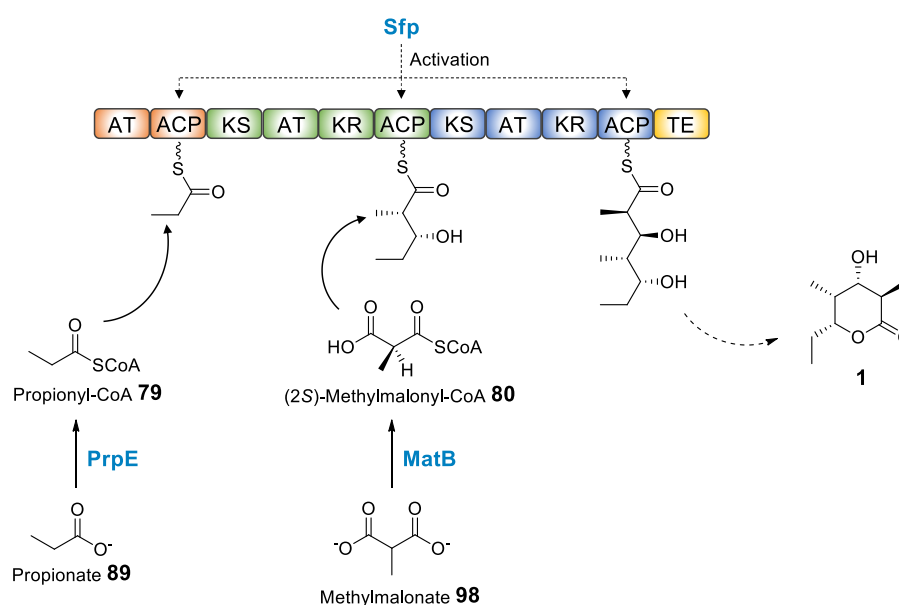


**Scheme 2.11** DEBS module 2-TE expression in engineered *S. cerevisiae* to catalyze the substrate **97** to synthesize the polyketide product triketide lactone **86**: **A**, the (2*S*)-methylmalonyl-CoA **80** production *via* the PCC pathway; **B**, the (2*S*)-methylmalonyl-CoA **80** production *via* the MatB pathway.

To assess this engineered yeast strain, DEBS module 2-TE comprised of module 2 of DBES1 fused to the thiolesterase (TE) domain of DEBS3 was expressed *in vivo*. The *N*-acetyl-cysteamine thiol (SNAC) ester of propyl-diketide **97** was fed as the

polyketide precursor. Through either the PCC pathway or MatB pathway for (2*S*)-methylmalonyl-CoA **80** supply, the final product triketide lactone **1** could be observed with a titer of about 0.5~1 mg·L<sup>-1</sup> (Scheme 2.11).

Although not DEBS1-TE expressed in the engineered *S. cerevisiae*, this experiment demonstrated that the necessary precursors for DEBS1-TE expression can be provided sufficiently in *S. cerevisiae* and the engineered metabolic pathways support the expression of a DEBS1-TE-related minimal modular PKS. Based on these results, it is deduced that the heterologous host *S. cerevisiae* is capable of expressing the modular PKS DEBS1-TE *via* either the PCC pathway or the propionyl-CoA-independent MatB pathway to produce the predicted triketide lactone **1** (Scheme 2.12).



**Scheme 2.12** Hypothesized DEBS1-TE expression in engineered *S. cerevisiae* *via* the propionyl-CoA-independent MatB pathway.

Notwithstanding, there are many remaining avenues to further explore for optimization of *S. cerevisiae* as a generic heterologous host for modular PKS production.<sup>92</sup> Further optimization of precursor supply pathways may improve the production level of polyketide by alleviating bottlenecks and thereby increasing substrate flux through the polyketide pathway. For example, MatB is flexible to accept both malonate and methylmalonate as a substrate. A racemic mixture of (2*R*, *S*)-methylmalonyl-CoA is produced by the MatB pathway. As DEBS1-TE only

recognizes the 2*S* isomer of methylmalonyl-CoA, coexpression of methylmalonyl-CoA epimerase may enhance the substrate level of the specific 2*S* isomer. In addition, due to the incompatibility of codon usage bias, it is likely that further codon modification of the PKS and/or metabolic pathway genes may significantly improve protein expression, resulting in the increase of precursor and polyketide levels.

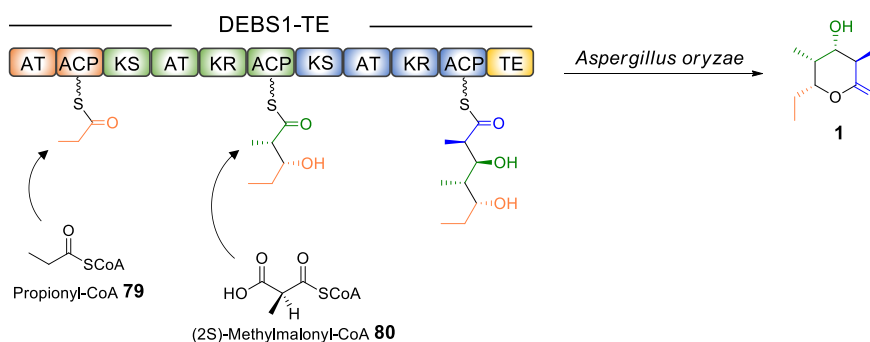
## 2.2 Aims

To date, no filamentous fungus has ever been utilized for the expression of a modular PKS. Filamentous fungi offer several advantages for polyketide expression and engineering. Particularly, workflows pioneered in the Prof. Cox group, involving cloning and recombination in yeast, followed by rapid expression in fungi allow rapid engineering of complex genes. For example, iterative PKS such as those involved in the biosynthesis of tenellin **22** and other polyketides have been dramatically altered by domain swap experiments. Fungi are eukaryotic and use monocistronic operons in which each gene can be independently controlled. This may offer advantages over the use of polycistronic operons in bacteria where multiple genes are controlled from a single promoter.

Furthermore, fungi are internally more complex than bacteria, possessing numerous internal compartments such as peroxisomes where selective reactions could be catalyzed. Fungi also possess diverse tailoring enzymes and can catalyze a wide variety of oxidative rearrangements unavailable in bacteria. Thus, fungi could be very interesting hosts for the modification of bacterial polyketides. However, to date modular PKS are unknown in fungi, and it is unsure whether suitable precursors could be provided.

In this chapter, to expand and advance the capability of fungal heterologous expression systems, expression of the modular PKS DEBS1-TE in the filamentous fungus *Aspergillus oryzae* will be attempted. DEBS1-TE, as a typical example of a modular PKS, has been well expressed in heterologous hosts bacteria and *Saccharomyces cerevisiae*. If it is feasible, triketide lactone **1** will be produced in *A. oryzae* (Scheme 2.13).





**Scheme 2.13** Project aim in this chapter.

Similar to *E. coli* and *S. cerevisiae*, *A. oryzae* also encounters many challenges, such as the difficulty of expressing highly GC-rich genes, correct protein folding, and protein post-translational modification. Moreover, the starter unit propionyl-CoA **79** exhibits toxicity to *Aspergillus* species.<sup>95</sup> The extender unit (2S)-methylmalonyl-CoA **80** is not present in fungi.<sup>96</sup> Thus, new metabolic pathways for precursor supply must be engineered in *A. oryzae*. All these problems are expected to be resolved in this chapter.

## 2.3 Results

Based on heterologous expression studies of DEBS1-TE in *E. coli* and *S. cerevisiae*, three engineering aspects were considered to be essential for the successful heterologous expression of DEBS1-TE in *A. oryzae*. These include supply of precursors, codon optimization of high GC *Streptomyces* genes, and introduction of an exogenous PPTase. Among them, precursor supply is the highest priority, followed by codon optimization. However, fungi are known to possess broadly selective PPTases such as NpgA, and therefore this element should not be a significant problem.<sup>97</sup>

### 2.3.1 Propionyl-CoA Production Pathway

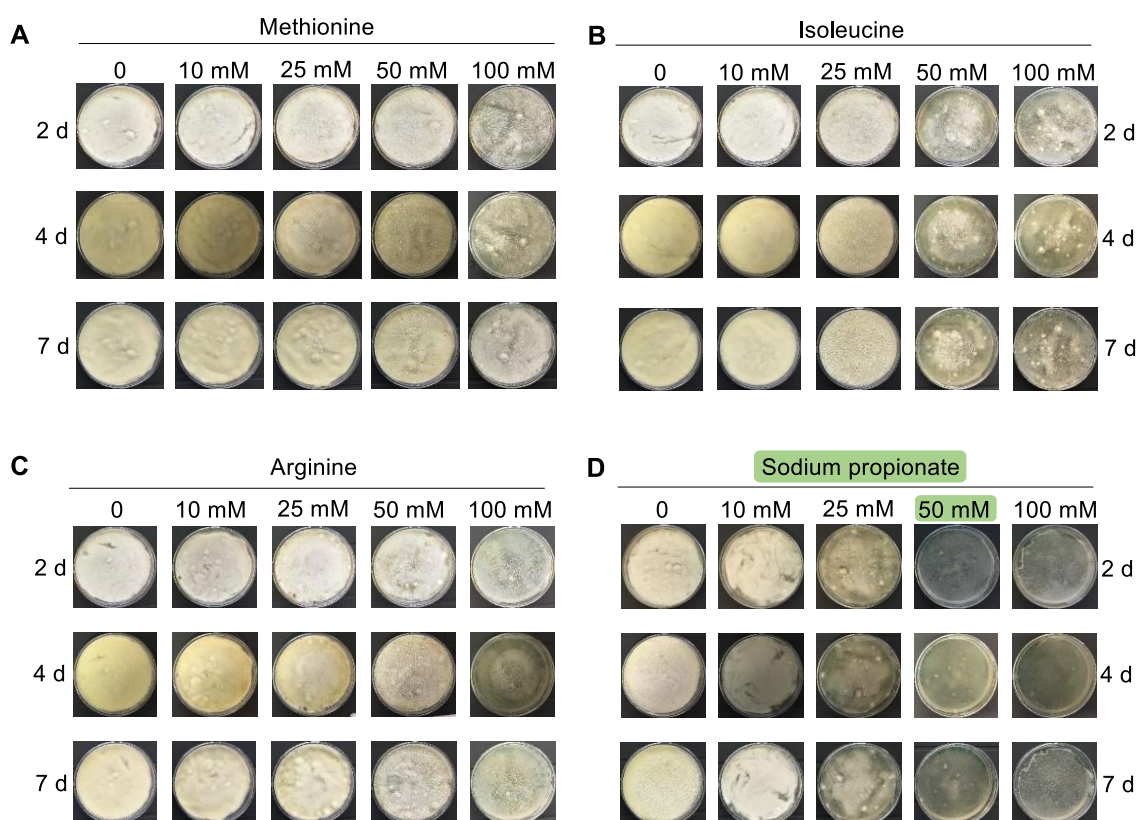
Heterologous production of the starter unit propionyl-CoA **79** is an initial and key determinant for triketide lactone **1** biosynthesis. According to reported studies, it was found that excess buildup of propionyl-CoA **79** could be able to inhibit the growth of *Aspergillus nidulans* and interfere with polyketide synthase activity.<sup>98-100</sup> High concentration of propionyl-CoA **79** *in vivo* impairs other CoA-dependent enzymes such as pyruvate dehydrogenase, succinyl-CoA synthetase, and ATP citrate lyase. Consequently, the toxicity examination of propionate is required to perform in order to figure out an optimal concentration of this building block for polyketide biosynthesis. If it were found that propionate is toxic to the host *A. oryzae* at low concentrations then the project would not be feasible.

#### 2.3.1.1 Propionyl-CoA Toxicity Assessment to *A. oryzae* NSAR1

In *Aspergillus* species, propionyl-CoA **79** mainly derives from the degradation of amino acids such as methionine, isoleucine, and arginine.<sup>101,102</sup> Besides amino acid degradation, propionyl-CoA **79** can also derive from direct activation of propionate by propionyl-CoA synthetase.<sup>100,103</sup> Here, the toxicity towards *A. oryzae* NSAR1 in the presence of propionyl-CoA **79** was assessed by supplementing amino acids or propionate.

Methionine, isoleucine, arginine, and sodium propionate solutions were prepared and then mixed into DPY agar media, respectively. For each group, five

different concentrations of supplementary were set up from 0 to 100 mM. The same amount of *A. oryzae* spore suspension was inoculated on the face of each DPY plate with supplementary inhibitor. After 2, 4, and 7 days, the growth states of each sample were recorded. The results showed that sodium propionate distinctly inhibits growth of *A. oryzae* (Fig. 2.2) at lower concentrations than three amino acids. The minimal inhibitory concentration of propionate is 50 mM, which is far less than those of other compounds that are over 100 mM. Therefore, as much as 25 mM of propionate can be fed to *A. oryzae* as the source of propionyl-CoA **79** in the case of not impairing the fungal growth.



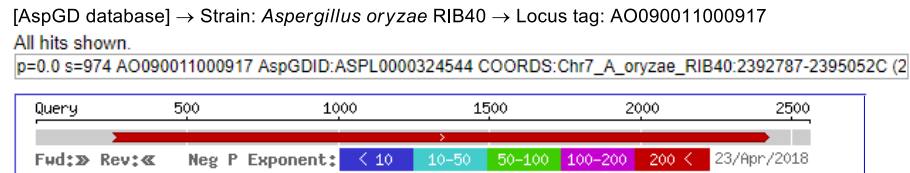
**Figure 2.2** Toxicity assessment of propionyl-CoA **79** produced from different sources to *A. oryzae* NSAR1: **A-C**, amino acids methionine, isoleucine, and arginine as the source; **D**, sodium propionate as the source.

### 2.3.1.2 *pcsA* Identification

The conversion of propionate *in vivo* requires propionyl-CoA synthetase. According to previous studies, a propionyl-CoA synthetase gene *pcsA* (accession number AY102074) was characterized in *Aspergillus nidulans*.<sup>99</sup> To confirm the existence of its homologous gene in *A. oryzae*, *A. nidulans pcsA* was used as a query sequence to search the genome of *A. oryzae* RIB40 using two database platforms: AspGD and

NCBI (Fig. 2.3). On both of them, the same sequence hit (accession number NC\_036441) was shown as the priority with 83 % sequence identity. The *A. oryzae pcsA* gene was then examined if it is present and transcribed in the *A. oryzae* NSAR1 strain used in our heterologous expression studies.

**A**



**B**

[NCBI Blastx] → Strain: *Aspergillus oryzae* RIB40 → Gene ID: 5998582 (locus tag: AO090011000917)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
unnamed protein product [Aspergillus oryzae RIB40]	<i>Aspergillus oryzae</i> RIB40	510	1099	85%	0.0	83.39%	690	XP_001826479.1
acetate-CoA ligase [Aspergillus oryzae RIB40]	<i>Aspergillus oryzae</i> RIB40	164	306	81%	6e-70	31.65%	670	XP_001820206.1
AMP-dependent synthetase and ligase [Aspergillus oryzae]	<i>Aspergillus oryzae</i>	46.6	92.4	22%	2e-12	30.61%	506	CC0008138.1

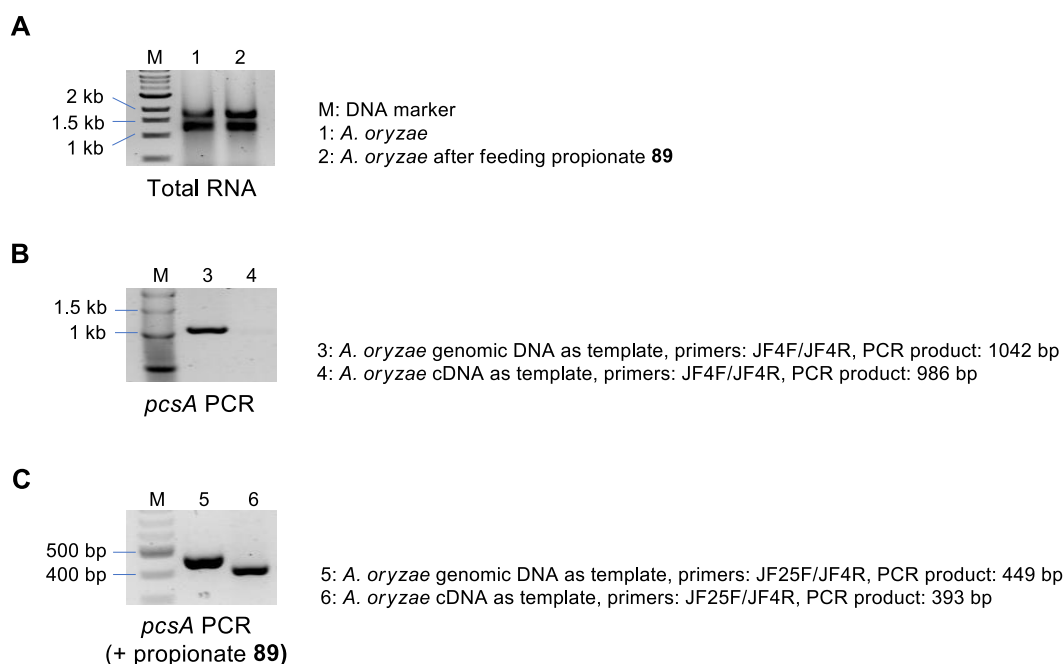
**C**

<i>A. nidulans</i> -PcsA	-MTHPQQAVHAAASLQNPFAFWSHHAQQLHWHKPSRAIGRSTKTLASGASHESWSWFPDG	59
<i>A. oryzae</i> -PcsA	MSRHPQQLVHGSSLRDPESFWSHHAQLYWHKPSHVISRHTKSLPSGTSHDHWWSWFPDG	60
<i>A. nidulans</i> -PcsA	EISTTYNCVDRHVLNGNDVAVIWDSPVIGTKKKEKYTYRQLLDEVEVLAVGLREBVGKKG	119
<i>A. oryzae</i> -PcsA	EISTTYNCVDRHVLNGNDVAVIWDSPVIGTKKKEKYTYRQLLDEVEVLAVGLREBVGKKG	120
<i>A. nidulans</i> -PcsA	DVVIYMPMPAALIGALAVARLGAITHAAVFGGFAAKSLAQRIEAARPRAILTASCGIEG	179
<i>A. oryzae</i> -PcsA	DVVIYMPMPAALIGALAVARLGAITHAAVFGGFAAKSLAQRIEAARPRAILTASCGIEG	180
<i>A. nidulans</i> -PcsA	AKGPIAYRPLVEGAIEASSFKPEKVLIVQRDQLRWNRPKLGGQRNWRNLVKSARMRGIK	239
<i>A. oryzae</i> -PcsA	SKGPVAYRPLVEGAIEASSFKPEKVLIVQRDQLRWNRPKLGGQRNWRNLVKSARMRGIK	240
<i>A. nidulans</i> -PcsA	AEPVVRSTIDGLYIYTSSTTGLPKGVVREAGGHAVGLSLSIKYLFDIHGPGDVMFCASD	299
<i>A. oryzae</i> -PcsA	AGPVPVASTIDGLYIYTSSTTGLPKGVVREAGGHAVGLSLSIKYLFDIHGPGDVMFCASD	300
<i>A. nidulans</i> -PcsA	IGVWVGHSYLYAPLLVGATTVLFEGKPVGTPDAGTFWRVVAEHKANVLFAPTALRAIR	359
<i>A. oryzae</i> -PcsA	IGVWVGHSYLYAPLLVGATTVLFEGKPVGTPDAGTFWRVVAEHKANVLFAPTALRAIR	360
<i>A. nidulans</i> -PcsA	KEDPDNKHFEVAGDNNLRHLRALFLAGERSEPSIVRAYQDLLTKHAARGALVVDNWWSS	419
<i>A. oryzae</i> -PcsA	KDDPDNKHFEVARRGKHLFRALFLAGERSEPSIVQYQDLLSRHAAPGAIVVDNWWSS	420
<i>A. nidulans</i> -PcsA	ESGSPISGLALRSVAVGRVPPRSDEYDVAPLAIRPGSAGLPMGPFVVRVVDDEGNEVAQGT	479
<i>A. oryzae</i> -PcsA	ESGSPISGLALRSTAGMTL--GDDKEVMPLAIRPGSAGLPMGPFVVRVVDDEGNEVPRGT	478
<i>A. nidulans</i> -PcsA	MGNIVMATPLAPTAFTLRFNDDERFYKGYLKRFGGRWLDTGADGMIDQDGYIHWMSRSD	539
<i>A. oryzae</i> -PcsA	MGNIVLNMPLAPTAFTLRFNDDERFYKGYLKRFGGRWLDTGADGMIDQDGYIHWMSRSD	538
<i>A. nidulans</i> -PcsA	IINVAHRFSTGGSIEQAILSHPAIGEASVVGIPDALKGHLFFAFITLKQSGGNSPARP	599
<i>A. oryzae</i> -PcsA	IINVAHRFSTG--AIEQAILSHPEVGEASVVGIPDALKGHLFFAFIQPRTASAALPATP	596
<i>A. nidulans</i> -PcsA	SAELFNSVNRVREIQGAIASLGGMIQGGMIPKTRSGKTLRRVRELVEVARGFEFEKE	659
<i>A. oryzae</i> -PcsA	TPELFNAINQRVREIQGAIASLGGMIQGGMIPKTRSGKTLRRVRELLEHGVRGDYGAP	656
<i>A. nidulans</i> -PcsA	VAVPPTVEDRGVVEVAREKVVREYFESQSGSPKAKL	694
<i>A. oryzae</i> -PcsA	VSIPPTVEDADVVEIARSKVREYFEEKQRSRAKL-	690

**Figure 2.3** Discovery of *A. oryzae pcsA* gene: **A**, the search result of AspGD database; **B**, the search result of NCBI database; **C**, the protein sequence alignment of *A. nidulans* and *A. oryzae pcsA* genes.

Given that *A. oryzae* NSAR1 is a quadruple auxotrophic strain (*argB*<sup>-</sup>, *niaD*<sup>-</sup>, *sC*<sup>-</sup> and *adeA*<sup>-</sup>) derived from its initial parent strain *A. oryzae* RIB40, *A. oryzae* RIB40 and *A. oryzae* NSAR1 should share the identical *pcsA* gene sequences.

A pair of specific 5' and 3' primers for *pcsA* identification were designed. Genomic DNA was extracted from mycelia of *A. oryzae* NSAR1 grown in DPY media and used as the PCR template. The PCR analysis showed that a targeted product was amplified as expected. The PCR product was further identified as the native *pcsA* by gene sequencing. Next, total RNA was isolated from fresh mycelia of *A. oryzae* NSAR1 grown in DPY media and reverse-transcribed into cDNA (Fig. 2.4 A). Using the cDNA as template, *pcsA* was not detectable by PCR (Fig. 2.4 B), indicating that *pcsA* is silent *in vivo* under the standard cultivation conditions.



**Figure 2.4** Gene identification of *pcsA*: **A**, extraction results of *A. oryzae* total RNA before and after feeding propionate; **B**, genomic DNA (gDNA) PCR and RT-PCR analyses of *pcsA* in the absence of propionate; **C**, gDNA PCR and RT-PCR analyses of *pcsA* in the presence of propionate.

To determine the correlation of propionate and *pcsA*, a one-day culture of *A. oryzae* NSAR1 was supplemented with 50 mM propionate **89**. After one more day of cultivation, the fresh mycelia were collected and used to isolate total RNA, followed by RT-PCR using the same primers mentioned above. A distinct band was observed at a lower position compared with that obtained from genomic DNA due to splicing of introns (Fig. 2.4 C). It means *pcsA* is activated in *A. oryzae* in the presence of propionate **89**, in accordance with previously reported results.<sup>99</sup> Based on this, it is proposed that *pcsA* is most likely the predominant gene responsible for propionate metabolism, leading to the generation of propionyl-CoA **79** *in vivo*. Therefore, the

starter unit propionyl-CoA **79** for DEBS1-TE expression can be supplied by feeding exogenous propionate **89**.

### 2.3.2 Methylmalonyl-CoA Production Pathway

According to engineered pathways for DEBS1-TE expression in *E. coli* and yeast, it was demonstrated that the production of (2S)-methylmalonyl-CoA **80** can be achieved by introducing multiple different routes, such as the PCC pathway, the mutase-epimerase pathway, or the MatB pathway. Herein, the PCC pathway was employed due to its sufficient supply of precursor as well as simplicity of genetic manipulation. Moreover, the introduction of the PCC pathway is also thought to be able to keep propionyl-CoA **79** from excess accumulation.<sup>98,100,103</sup> Propionyl-CoA carboxylase (PCC) is composed of two different subunits. The  $\alpha$  subunit (PccA) and  $\beta$  subunit (PccB) work together to carry out carboxybiotin formation and carboxyl transfer to form (2S)-methylmalonyl-CoA **80**.<sup>104</sup>

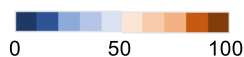
A third subunit of PCC,  $\epsilon$ , also known as PccE, was discovered from *Streptomyces coelicolor*.<sup>86</sup> The  $\epsilon$  subunit was found to specifically interact with the carboxyltransferase  $\beta$  subunit to dramatically increase the specific activity of the enzyme complex. Thus, to build a more efficient PCC pathway, all three subunits from *S. coelicolor* (PccA, PccB, and PccE) were prepared for introduction into *A. oryzae* NSAR1. The corresponding accession numbers of PCC subunit genes *pccA*, *pccB*, and *pccE* are AF113604, AF113605, and CAC21625.

#### 2.3.2.1 Condon optimization

Comparison of the codon usage of *pccABE* from *S. coelicolor* with that of *A. oryzae* from Codon Usage Database (<https://www.kazusa.or.jp/codon/>) showed that the frequencies of individual codon usages are significantly different (Table 2.1). The GC contents of *pccABE* genes are 72 %, 70 %, and 74 %, respectively, while that of *A. oryzae* genes averages at 55 %.<sup>105</sup> Therefore, it is necessary to optimize the usage of codons before heterologous expression.

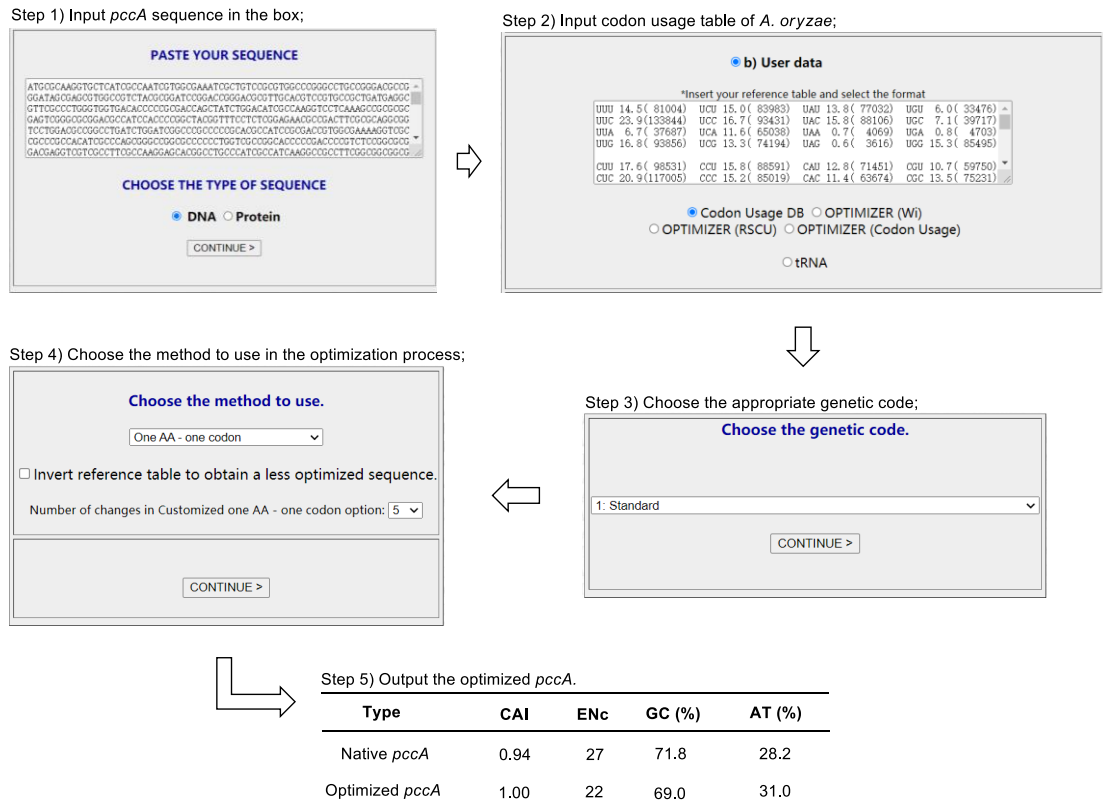
**Table 2.1** Comparison of codon usage in *A. oryzae* genes with that in native and codon-optimized *pccABE* genes.

Amino acid	Codon	Frequency (%)						
		<i>A. oryzae</i>	<i>pccA</i>		<i>pccB</i>		<i>pccE</i>	
			Native	Optimized	Native	Optimized	Native	Optimized
Ala	GCG	20	18	0	26	0	27	0
	GCA	23	1	0	2	0	9	0
	GCT	27	2	0	2	0	0	0
Arg	GCC	31	78	100	71	100	64	100
	CGT	18	7	0	6	0	0	0
	CGC	23	76	100	84	100	67	100
Asn	CGA	17	0	0	0	0	0	0
	CGG	18	14	0	6	0	17	0
	AGA	13	0	0	0	0	17	0
Asp	AGG	12	2	0	3	0	0	0
	AAT	45	14	0	0	0	0	0
	AAC	55	86	100	100	100	100	100
Cys	GAT	53	6	100	3	100	0	100
	GAC	47	94	0	97	0	100	0
	TGT	46	0	0	25	0	0	0
Gln	TGC	54	100	100	75	100	0	0
	CAA	43	8	0	11	0	0	0
	CAG	57	92	100	89	100	0	0
Glu	GAA	44	20	0	11	0	67	0
	GAG	56	80	100	89	100	33	100
	GGT	28	5	0	8	0	0	0
Gly	GGC	31	88	100	92	100	100	100
	GGA	24	0	0	0	0	0	0
	GGG	17	7	0	0	0	0	0
His	CAT	47	0	100	0	100	0	100
	CAC	53	100	0	100	0	100	0
	ATT	36	0	0	0	0	0	0
Ile	ATC	50	96	100	100	100	100	100
	ATA	14	4	0	0	0	0	0
	TTA	7	0	0	0	0	0	0
Leu	TTG	18	2	0	0	0	0	0
	CTT	19	0	0	0	0	0	0
	CTC	23	52	100	47	100	0	100
Lys	CTA	11	0	0	0	0	0	0
	CTG	22	46	0	53	0	100	0
	AAA	36	10	0	0	0	0	0
Met	AAG	64	90	100	100	100	100	100
	ATG	100	100	100	100	100	100	100
	TTT	38	0	0	0	0	0	0
Phe	TTC	62	100	100	100	100	0	0
	CCT	27	0	100	0	100	0	100
	CCC	26	79	0	62	0	22	0
Pro	CCA	25	0	0	0	0	0	0
	CCG	22	21	0	38	0	78	0
	AGT	13	0	0	0	0	0	0
Ser	AGC	18	10	0	9	0	0	0
	TCT	18	0	0	0	0	0	0
	TCC	20	72	100	77	100	67	100
Thr	TCA	14	3	0	0	0	0	0
	TCG	16	14	0	14	0	33	0
	ACT	24	0	0	3	0	0	0
Trp	ACC	32	75	100	62	100	70	100
	ACA	24	0	0	0	0	0	0
	ACG	20	25	0	35	0	30	0
Tyr	TGG	100	100	100	100	100	100	100
	TAT	47	13	0	0	0	0	0
	TAC	53	88	100	100	100	100	100
Val	GTT	27	0	0	0	0	0	0
	GTC	33	81	100	72	100	50	100
	GTA	13	2	0	0	0	0	0
	GTG	27	17	0	28	0	50	0



Based on the codon usage of *A. oryzae*, the *pccABE* sequences were optimized using the online analytical tool Optimizer (<http://genomes.urv.es/OPTIMIZER/>) and the codon usage table of *A. oryzae* from Codon Usage Database (*pccA* was taken as an

example shown in Scheme 2.14). Furthermore, the original start codon GTG was replaced with ATG, which is preferred for translation initiation in fungi. The codon-optimized sequence was then subject to Graphical Codon Usage Analyzer (<http://gcu.schoedl.de/>) for codon quality analysis. It displayed that the relative adaptiveness values of all amino acids in *A. oryzae* were up to 100 after optimization.



**Scheme 2.14** Optimization process of *pccA* using Optimizer. Abbreviations: CAI, codon adaptation index; ENc, effective number of codons; GC (%), G+C percentage; AT (%), A+T percentage.

### 2.3.2.2 Plasmid construction

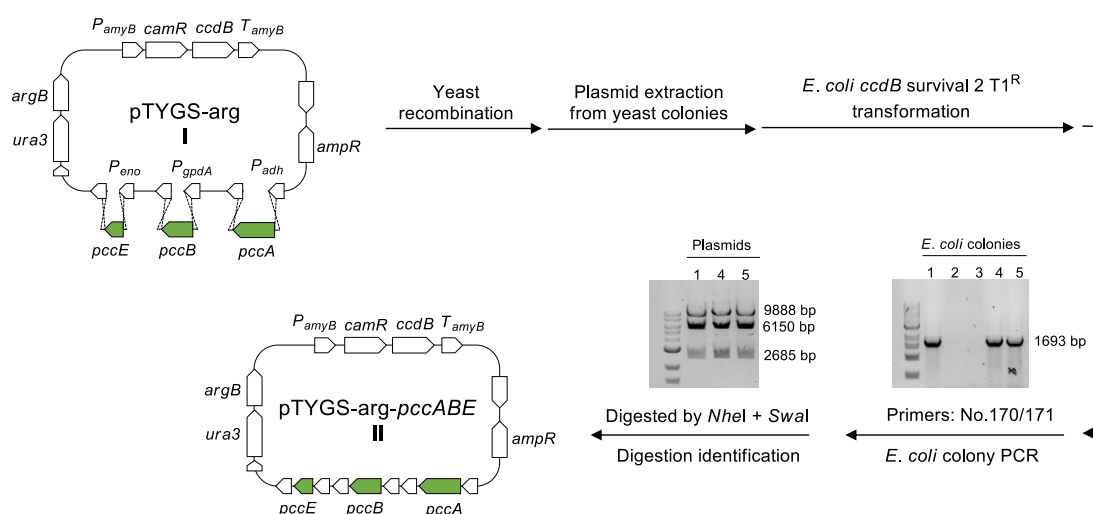
The vector pTYGS-arg I was chosen for plasmid construction. This vector carries two auxotrophic genes, *ura3* for yeast and *argB* for *A. oryzae* selection. It also possesses four different promoter/terminator pairs, including *amyB*, *adh*, *gpdA*, and *eno* for simultaneous expression of multiple genes. Between  $P_{amyB}$  and  $T_{amyB}$ , it contains a Gateway cassette consisting of LR recombination sites (*attR1* and *attR2*), the *ccdB* gene, encoding the CcdB killer protein, and the chloramphenicol resistance *camR*.<sup>62</sup>

All *pcc* gene fragments flanked with overlaps (by ca 30 bp) were synthesized commercially. Using an *AscI*-digested vector pTYGS-arg I and recovered *pccABE*



gene fragments after double restriction enzymatic digestions (*pccA-EcoRI/BamHI*, *pccB-XbaI/BamHI*, *pccE-BamHI/XbaI*), the yeast recombination was performed as previously published protocols.<sup>63</sup>

Yeast colonies were obtained and then gathered together to extract plasmids. Subsequently, using standard heat-shock protocols, the extracted plasmid mixture was introduced into *E. coli ccdB survival 2 T1<sup>R</sup>*, which is resistant against the CcdB killer protein. A large number of *E. coli* colonies were generated, and then 5 random colonies were identified by colony PCR and restriction enzyme digestion. Lastly, the expression plasmid was successfully constructed (Scheme 2.15). Three subunit genes were inserted downstream of promoters *P<sub>adh</sub>*, *P<sub>gpdA</sub>*, and *P<sub>eno</sub>*, respectively.



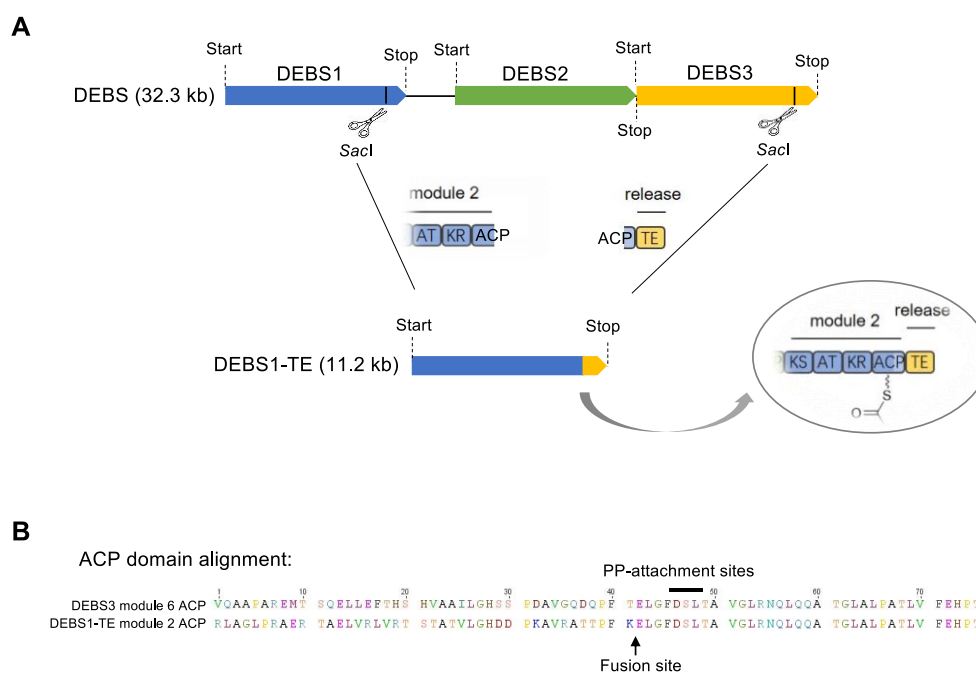
**Scheme 2.15** Construction of the plasmid pTYGS-arg-*pccABE II*. For *E. coli* colony PCR, the *pccB* was chosen to perform PCR identification. Primers No.170/171 are located in the *P<sub>gpdA</sub>* and *T<sub>gpdA</sub>*, respectively. For plasmid digestion identification, restriction enzymes *NheI*+*SmaI* were chosen. All of DNA bands above were shown as positive.

## 2.3.3 DEBS1-TE Cassette Construction

### 2.3.3.1 Design *in silico*

As DEBS1-TE is a non-native enzyme, its gene sequence is not available in the NCBI database, and consequently it has to be designed *in silico* using data from its original construction by Leadlay and coworkers.<sup>106</sup> Here, it was constructed *in silico* as shown in Fig. 2.5 A. Two *SacI* restriction sites were chosen as splice sites within ACP domains. The 5' one is from module 2 while the 3' one is from module 6 in DEBS. The fragment containing DEBS2 and DEBS3 was first deleted between two splice

sites. Then, the remaining fragment containing TE domain was ligated to the cut end of DEBS1 to form a new and complete sequence DEBS1-TE.



**Figure 2.5** *In silico* design of DEBS1-TE: **A**, DEBS was spliced at *SacI* sites and reassembled to form DEBS1-TE; **B**, the domain alignment between the newly formed DEBS1-TE module 2 ACP domain and the original DEBS3 module 6 ACP domain. Abbreviation: PP, phosphopantetheine.

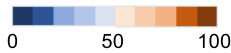
Notably, a hybrid ACP domain is formed in DEBS1-TE instead of the original ACP domain of module 2. In order to determine the integrity of the newly-formed ACP domain structurally and functionally, its amino acid sequence was aligned with that of the original ACP domain from module 6 of DEBS3 (Fig. 2.5 B). The result showed that two protein sequences align well with 72 % identity. The conserved ACP phosphopantetheine DSL attachment site is not interrupted by the fusion, implying that the newly-formed ACP domain is likely to be functional.

### 2.3.3.2 Condon optimization

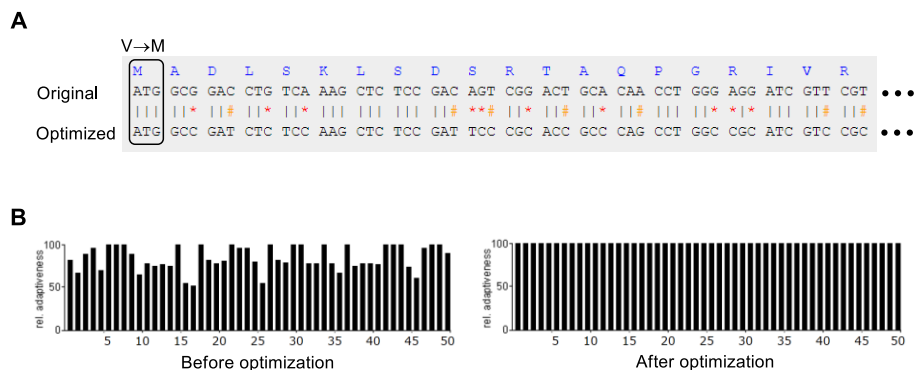
Comparison of the codon usage of DEBS1-TE from *S. erythraea* with that of *A. oryzae* from Codon Usage Database (<https://www.kazusa.or.jp/codon/>) showed that the frequencies of individual codon usages are significantly different (Table 2.2).

**Table 2.2** Comparison of codon usage in *A. oryzae* genes with that in native and codon-optimized DEBS1-TE.

Amino acid	Codon	Frequency (%)			Amino acid	Codon	Frequency (%)		
		<i>A. oryzae</i>	DEBS1-TE				<i>A. oryzae</i>	DEBS1-TE	
			Native	Optimized				Native	Optimized
Ala	GCG	20	50	0	Leu	CTT	19	3	0
	GCA	23	3	0		CTC	23	31	100
	GCT	27	2	0		CTA	11	0	0
Arg	GCC	31	45	100	CTG	22	59	0	
	CGT	18	6	0	Lys	AAA	36	0	8
	CGC	23	48	100	AAG	64	0	92	
	CGA	17	4	0	Met	ATG	100	100	100
	CGG	18	36	0	Phe	TTT	38	1	0
	AGA	13	0	0	TTC	62	99	100	
Asn	AGG	12	6	0	Pro	CCT	27	4	100
	AAT	45	8	0	CCC	26	35	0	
Asp	AAC	55	92	100	CCA	25	2	0	
	GAT	53	5	100	CCG	22	59	0	
Cys	GAC	47	95	0	Ser	AGT	13	2	0
	TGT	46	10	0	AGC	18	22	0	
Gln	TGC	54	90	100	TCT	18	1	0	
	CAA	43	7	0	TCC	20	27	100	
	CAG	57	93	100	TCA	14	2	0	
Glu	GAA	44	23	0	TCG	16	46	0	
	GAG	56	77	100	Thr	ACT	24	3	0
Gly	GGT	28	19	0	ACC	32	60	100	
	GGC	31	52	100	ACA	24	1	0	
	GGA	24	9	0	ACG	20	36	0	
	GGG	17	20	0	Trp	TGG	100	100	100
His	CAT	47	7	100	Tyr	TAT	47	5	0
	CAC	53	93	0	TAC	53	95	100	
Ile	ATT	36	1	0	GTT	27	4	0	
	ATC	50	97	100	GTC	33	44	100	
Leu	ATA	14	2	0	GTA	13	0	0	
	TTA	7	0	0	GTG	27	52	0	
	TTG	18	7	0					



A codon-optimized DEBS1-TE sequence was designed using the online analytical tool Optimizer (<http://genomes.urv.es/OPTIMIZER/>) and the codon usage table of *A. oryzae* from Codon Usage Database as the aforementioned method (Fig. 2.6 A). Meanwhile, the original start codon GTG was replaced with ATG. The codon-optimized DEBS1-TE sequence was then subject to Graphical Codon Usage Analyzer (<http://gcua.schoedl.de/>) for codon quality analysis.



**Figure 2.6** Codon optimization of DEBS1-TE: **A**, sequence alignment of the original DEBS1-TE before optimization and the optimized DEBS1-TE; **B**, the relative adaptiveness analysis of DEBS1-TE amino acids (only the first 50 aa shown).

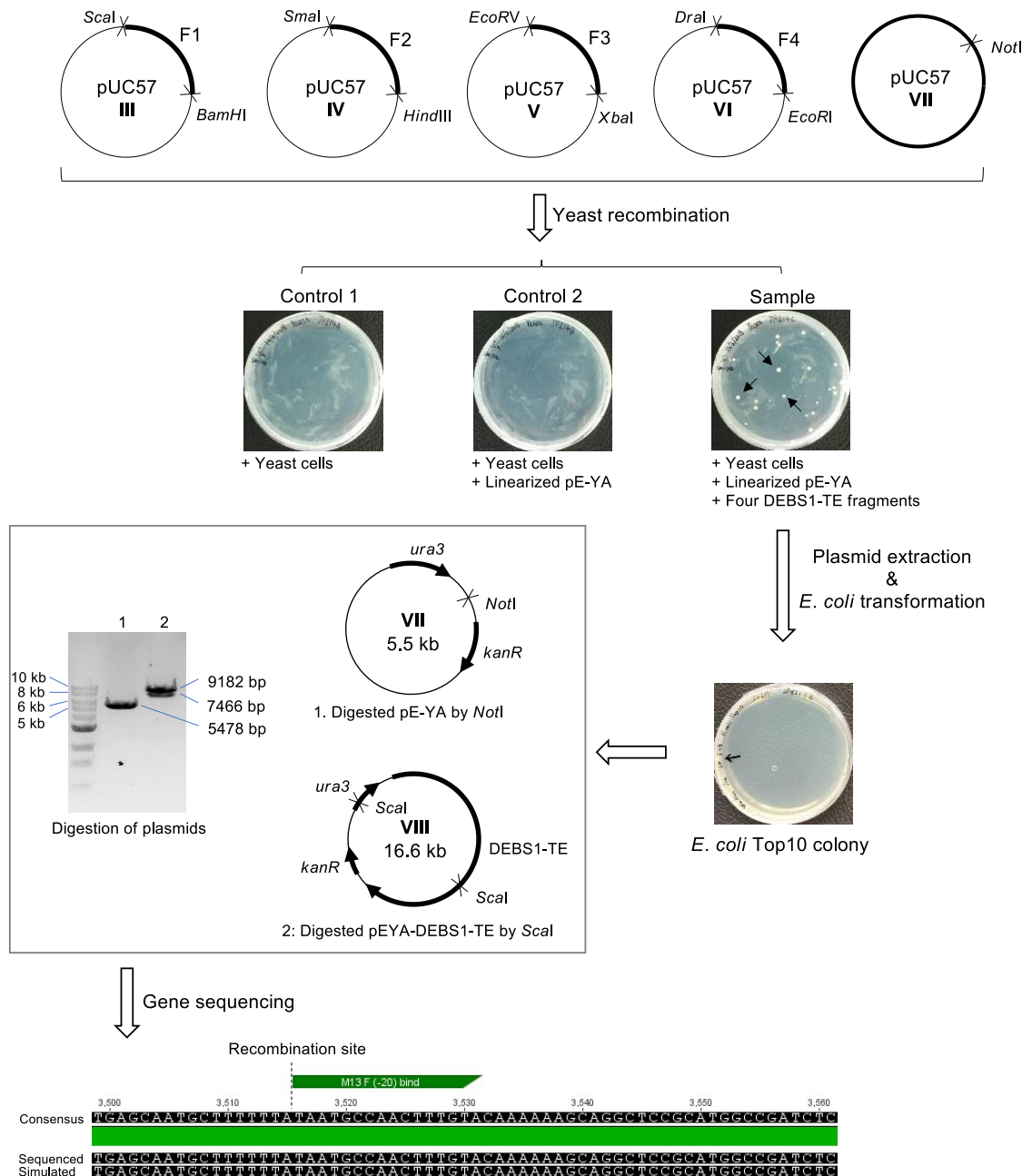
It displayed that the relative adaptiveness values of all amino acids in *A. oryzae* were up to 100 after optimization (Fig. 2.6 B). The GC content of the optimized DEBS1-TE is about 73 %. It is comparable to that of the original DBES1-TE sequence before optimization, while the average GC content of *A. oryzae* genes is ~ 55 %.<sup>105</sup>

### 2.3.3.3 Synthesis and Assembly

For the synthetic convenience, the entire DEBS1-TE sequence (~ 11 kb) was designed as four fragments (~ 2.8 kb each). Each fragment was designed to overlap with adjacent fragments or the expression vector at both ends (by *ca* 30 bp) for the subsequent yeast recombination. The initiation end of the first fragment and the termination end of the fourth fragment were respectively attached with homologous fragments, which aim for yeast recombination with the linearized pE-YA vector. To avoid the inaccessibility of PCR product due to high GC content, restriction enzyme digestion was required to obtain all desired fragments. Therefore, two different unique restriction enzymes were adhered to ends of each fragment.

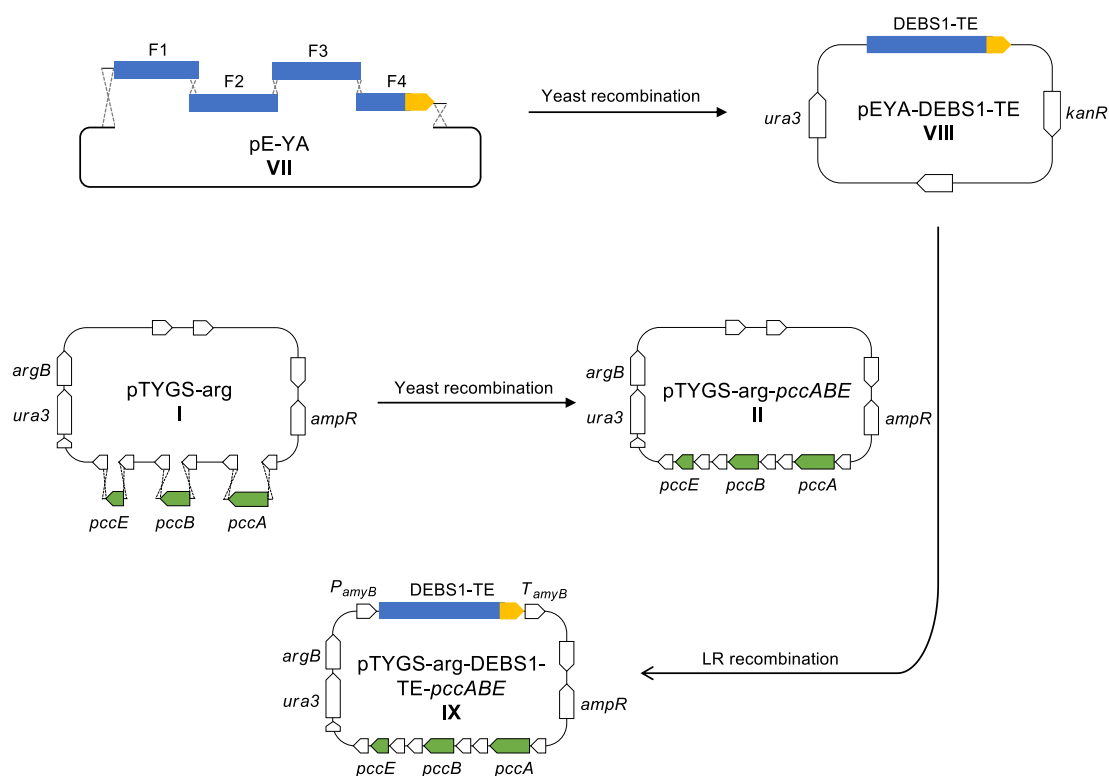
The four DEBS1-TE fragments were synthesized commercially in four separate pUC57 vectors (**III** - **VI**). All desired DEBS1-TE fragments and the linearized pE-YA vector **VII** were prepared by restriction-enzyme digestion and then introduced into competent yeast for homologous recombination (Scheme 2.16).<sup>63</sup> Two negative controls were set up. After 2 days of incubation, yeast colonies were produced only on the experimental sample. Then, plasmids were extracted from yeast colonies and immediately introduced into *E. coli* Top10 competent. *E. coli* colonies were screened by PCR identification. In the resulting positive *E. coli* colonies, plasmids were extracted and then identified by restriction enzyme digestion (Scheme 2.16). The empty vector pE-YA **VII** was used as the control. The extract plasmids were digested by the restriction enzyme *ScaI* and formed two bands (*i.e.* 9182 bp and 7466 bp) which match the predicted sizes. Ultimately, the obtained plasmid was confirmed to be pEYA-DEBS1-TE **VIII** by gene sequencing.

### 2.3.3 Results – DEBS1-TE Cassette Construction



**Scheme 2.16** Recombination, screening, and identification of the plasmid pEYA-DEBS1-TE VIII.

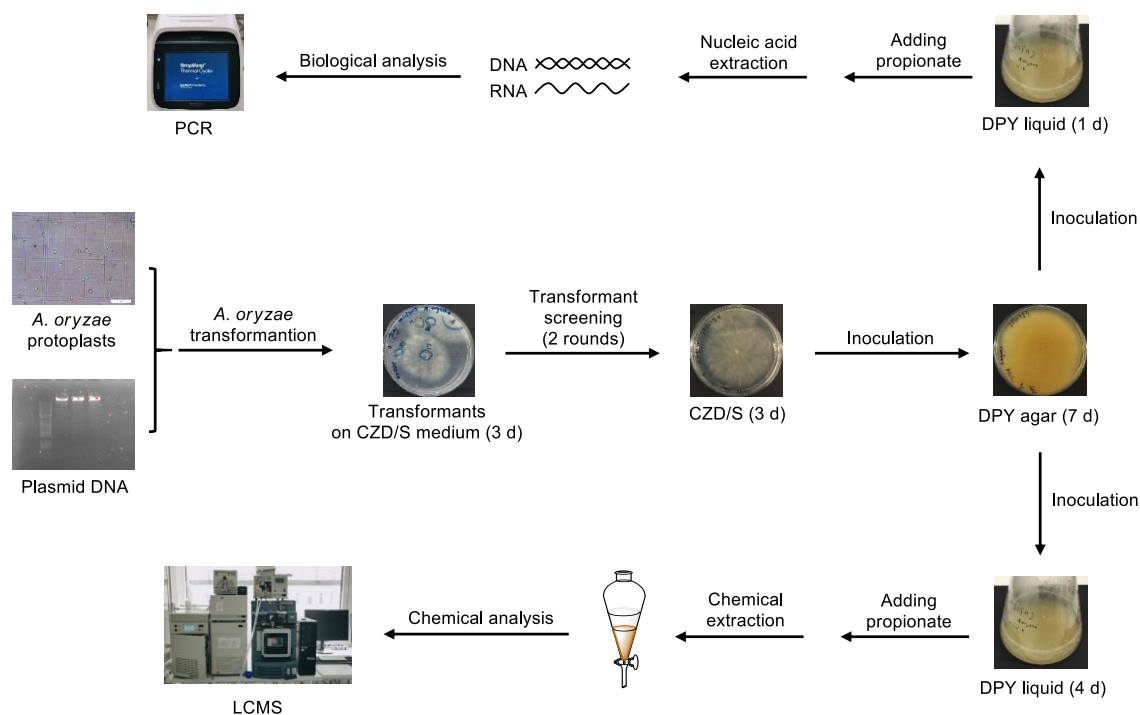
To construct the final plasmid for DEBS1-TE expression, LR recombination was carried out using pEYA-DEBS1-TE VIII as the entry clone and pTYGS-arg-*pccABE* II as the destination vector.<sup>62</sup> It led to the formation of the plasmid pTYGS-arg-DEBS1-TE-*pccABE* IX harboring *pccABE* genes and DEBS1-TE simultaneously, of which DEBS1-TE locates downstream of the promoter *P<sub>amyB</sub>* (Scheme 2.17).



**Scheme 2.17** Overall construction strategy of pTYGS-arg-DEBS1-TE-*pccABE* IX.

### 2.3.4 Product Analyses

The polyethylene glycol (PEG)-mediated fungal transformation was performed using freshly prepared *A. oryzae* NSAR1 protoplasts and the plasmid pTYGS-arg-DEBS1-TE-*pccABE* IX as previously described protocols.<sup>65</sup> The resulting *A. oryzae* DEBS1-TE-*pccABE* transformants were transferred and underwent two rounds of screening on CZD/S selection media. Pure transformants were inoculated on DPY agar media for subsequent analyses (Scheme 2.18). On the one hand, transformants were inoculated further into liquid DPY media. After 4 days of cultivation, 50 mM sodium propionate were added into cultures. At the next day, the cultures were extracted with ethyl acetate, and the extracts were chemically analyzed by LCMS. On the other hand, *A. oryzae* DEBS1-TE-*pccABE* transformants were inoculated into liquid DPY media for one day of cultivation. At the second day, 50 mM sodium propionate were added into cultures, and the cultivation was proceeded overnight. The mycelia of transformants were collected for nucleic acid extraction. The isolated nucleic acids such as genomic DNA or RNA were biologically analyzed by gDNA-PCR or RT-PCR. All analytical results of transformants are shown in the next two sections.



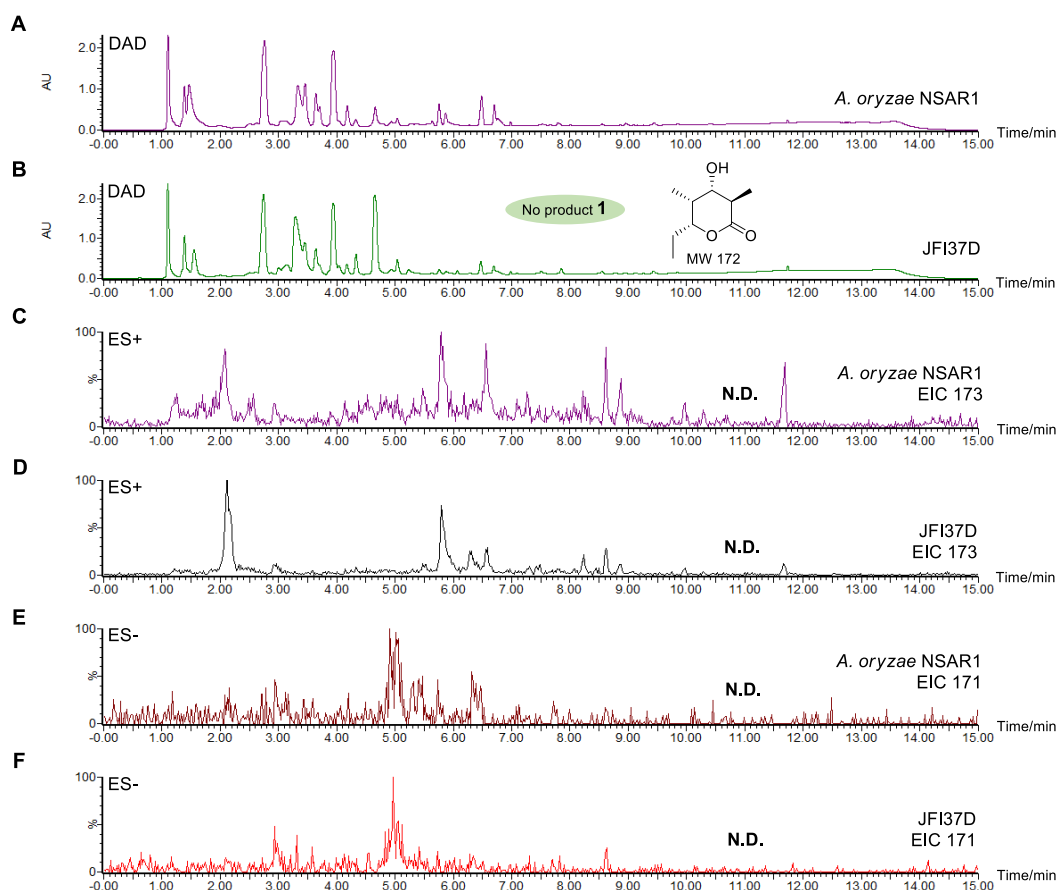
**Scheme 2.18** Workflow of *A. oryzae* transformation and transformant analyses.

### 2.3.4.1 Chemical Analyses

After fungal transformation, 24 *A. oryzae* DEBS1-TE-*pccABE* transformants (JFI37A-X) were obtained and then individually analyzed by LCMS. *A. oryzae* NSAR1 as the control was also cultivated under the same conditions. Meanwhile, the equal concentration of propionate was also added into the culture of *A. oryzae* NSAR1. The transformant JFI37D was taken as an example. By comparing with *A. oryzae* NSAR1, the product triketide lactone **1** was not observed on the diode array detection (DAD) chromatogram (Fig. 2.7 A and B). The comparison of total ion chromatogram (TIC) between the transformant JFI37D and the control could not find any significant difference either. Then, the molecular ion peaks of the targeted product **1** were extracted on TIC. No possible peak was observed on the extracted ion chromatogram (EIC) of all transformants (Fig. 2.7 C, D, E, and F). In addition, the optimization of cultivation media made no further improvement. It implied that the product **1** might not be produced in the transformant yet.

To diagnose this problem, the precursors propionyl-CoA **79** and (2*S*)-methylmalonyl-CoA **80** were analyzed to figure out whether the supply routes of building blocks are integrated or not. First of all, a reliable extraction method for

CoA thioesters was established as previously reported.<sup>107</sup> The mycelia of two-day-cultivated *A. oryzae* transformants were harvested by filtering through sterile Miracloth. Mycelia were added into screw cap tubes containing the prechilled quenching solution (prechilled acetonitrile/methanol/water (2:2:1)) and the equal volume of glass beads. Mycelia were fully homogenized and then centrifuged at  $14,000 \times g$  for 30 min at 4 °C to separate the supernatant from the glass beads and cell debris. The supernatant was immediately analyzed by LCMS.

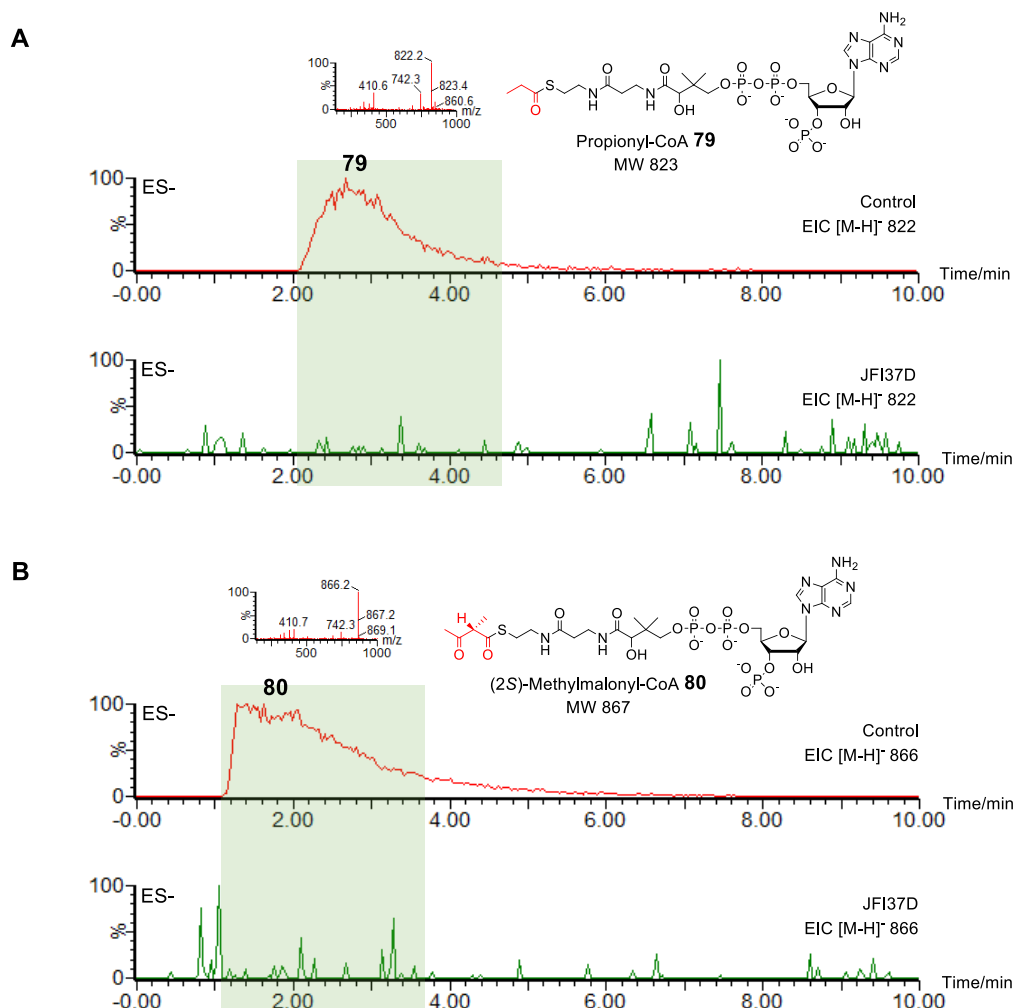


**Figure 2.7** LCMS comparative analysis of *A. oryzae* DEBS1-TE-*pccABE* transformant JFI37D and the control *A. oryzae* NSAR1: **A**, DAD chromatogram of *A. oryzae* NSAR1; **B**, DAD chromatogram of the transformant JFI37D; **C**, ES+ EIC  $[M+H]^+$  173 chromatogram of the control; **D**, ES+ EIC  $[M+H]^+$  173 chromatogram of the transformant JFI37D; **E**, ES- EIC  $[M-H]^-$  171 chromatogram of the control; **F**, ES- EIC  $[M-H]^-$  171 chromatogram of the transformant JFI37D. Abbreviation: N. D., not detected.

When the authentic substrate propionyl-CoA **79** was added into the culture at the day before harvesting mycelia, it could still be detected on LCMS. This indicates that propionyl-CoA **79** is not degraded during the extraction and analysis procedure. Propionyl-CoA **79** is observed as a wide peak on the extracted-ion chromatogram (EIC) of ES- TIC. That was also the same to the (2*S*)-methylmalonyl-CoA **80**



extraction result. However, the extraction results of *A. oryzae* DEBS1-TE-*pccABE* transformants after feeding propionate **89** did not give rise to any relevant peaks (JFI37D as an example in Fig. 2.8). Therefore, it was speculated that the engineered biosynthetic routes are not integrated, or those heterologous genes have not been expressed effectively *in vivo*.



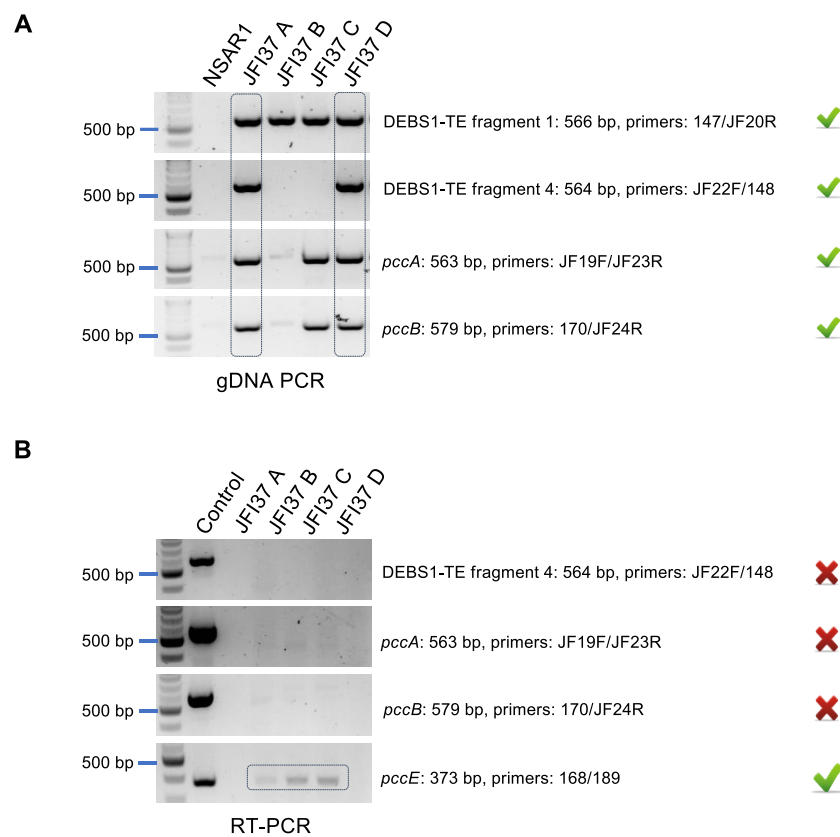
**Figure 2.8** Detection of precursors propionyl-CoA **79** and (2S)-methylmalonyl-CoA **80**: **A**, the comparison of EIC [M-H]<sup>-</sup> 822 of propionyl-CoA **79** between the control and *A. oryzae* DEBS1-TE-*pccABE* transformant JFI37D; **B**, the comparison of EIC [M-H]<sup>-</sup> 866 of (2S)-methylmalonyl-CoA **80** between the control and the transformant JFI37D.

### 2.3.4.2 Biological Analyses

In order to test the *in vivo* expression of DEBS1-TE and *pccABE*, *A. oryzae* DEBS1-TE-*pccABE* transformants were biologically analyzed, including by genomic DNA PCR and RT-PCR. First, to ensure there were no errors from the *in vivo* homologous recombination, genomic DNA of individual transformants were extracted and analyzed using multiple pairs of different primers. In some of the transformants, for

example JFI37A and JFI37D, all test fragments could be replicated simultaneously (Fig. 2.9 A). These results indicated that all genes had been recombined into the genome with integrity.

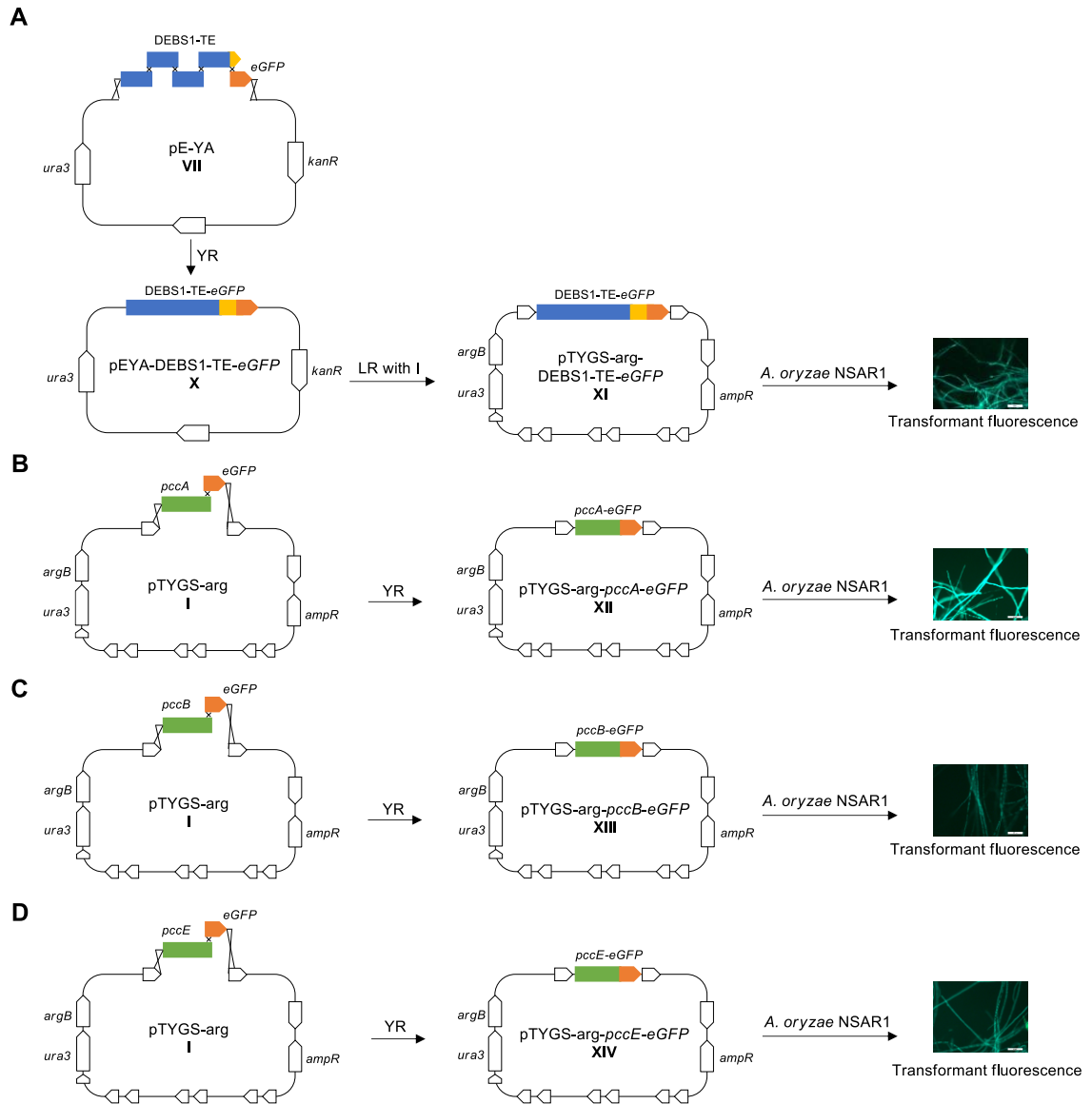
Next, the RNAs of all transformants were extracted and reverse-transcribed to cDNAs, respectively. The cDNAs were diagnosed with the same primers as above. The positive control used genomic DNA as the PCR template. It turned out that only the shortest gene *pccE* was vaguely observable by RT-PCR, while neither other longer *pcc* genes nor DEBS1-TE could be observed (Fig. 2.9 B). Nevertheless, it could be possible that RT-PCR fails due to the high GC content of these genes. Genes with high GC content might affect the amplification efficiency of PCR due to the tendency to fold into complex mRNA secondary structures.<sup>108</sup>



**Figure 2.9** Identification of gene transcription by PCR: **A**, genomic DNA PCR results of DEBS1-TE and *pcc* genes in *A. oryzae* DEBS1-TE-*pccABE* transformants JFI37A – JFI37D; **B**, RT-PCR results of DEBS1-TE and *pcc* genes in *A. oryzae* DEBS1-TE-*pccABE* transformants JFI37A – JFI37D. Green marks: observable; red markers: unobservable.

Thus, an eGFP fusion approach was chosen to verify gene expression.<sup>109</sup> In this method, a fused protein is artificially assembled by ligating a green fluorescent reporter gene *eGFP* to the 3' terminal of a gene. After translation and folding, the

expressed protein would be visible by fluorescent imaging. Here, DEBS1-TE and *pccABE* genes were fused with *eGFP* at their 3' termini and formed the plasmids **XI–XIV**, respectively. Each fused gene cassette was placed under the promoter  $P_{amyB}$  by yeast recombination, followed by being introduced into *A. oryzae* NSAR1.



**Scheme 2.19** Green fluorescent imaging of transformants with biosynthetic genes fused to *eGFP*: **A**, *eGFP* was fused with DEBS1-TE; **B**, *eGFP* was fused with *pccA*; **C**, *eGFP* was fused with *pccB*; **D**, *eGFP* was fused with *pccE*.

After a series of screening and cultivation procedures, the mycelia were observed under the fluorescent microscope. Transformants with the intense green fluorescence were found in all experimental groups (Scheme 2.19). It confirmed that all heterologous genes, even those with high GC content, had been expressed in *A.*

*oryzae*. Therefore, the fact that the engineered *A. oryzae* did not produce **1** was very likely to be caused by the other reason, that is the deficiency of biosynthetic routes.

### 2.3.5 Phosphopantetheinylation Pathway

The deficiency of phosphopantetheinylation might be the reason for the biosynthetic problem because phosphopantetheinyl transferases (PPTases) play an essential role in the biosynthesis of polyketides by activating the ACP domains of PKS.<sup>110-113</sup> A few Sfp-type PPTases have been discovered in *Aspergillus* species to date, such as NpgA from *A. nidulans*, PptA from *A. niger*, and PptB from *A. fumigatus*.<sup>97,114-119</sup> Based on NCBI blastp search, one cognate PPTase (accession number XP\_023093666) was found in *A. oryzae* RIB40. It shares 54 % identity with NpgA (accession number AAF12814).<sup>115,120</sup> However, its function to modular PKS ACPs has yet to be identified. The native fungal PPTase in *A. oryzae* might not be capable of activating ACP domains of bacterial modular PKS. Therefore, a heterologous PPTase from bacteria is required for the activation of modular PKS.

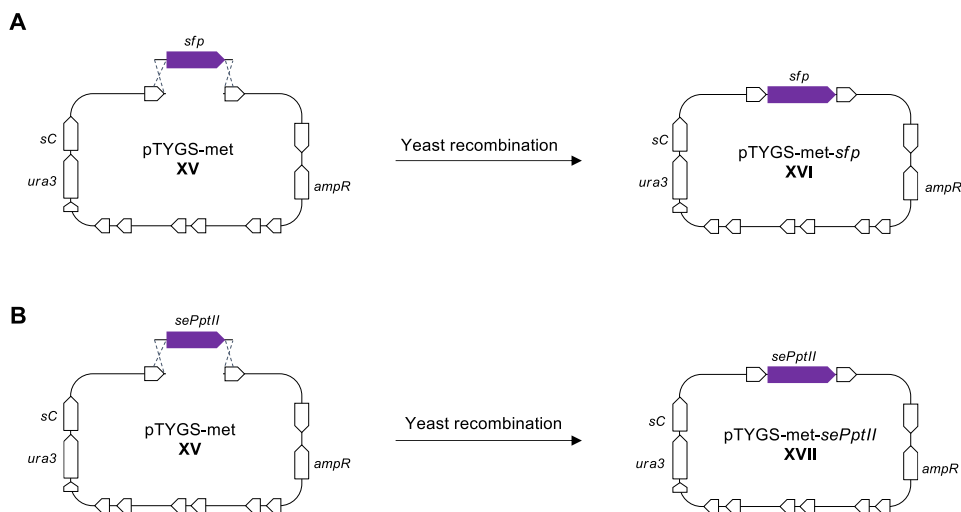
Two bacterial PPTases were adopted here. The first one was Sfp (accession number CAA44858) from *Bacillus subtilis*, which is the most widely-used PPTase so far.<sup>82</sup> It had also been used in the heterologous expression of DEBS1-TE in *E. coli* and *S. cerevisiae*.<sup>75,92</sup> The other one was SePptII (accession number AAR92400) from the erythromycin-producing strain *Saccharopolyspora erythraea*.<sup>121</sup> Both of them should be active for ACP domains of DEBS1-TE. Nevertheless, the efficiency of their activities in a fungal host was unknown.

The alignment analysis of protein sequences was then carried out (Fig. 2.10). It illustrated that PPTases from *Aspergillus* share very low identity (about 20 %) with either Sfp or SePptII from bacteria, though they all belong to the Sfp-type PPTase family.<sup>111</sup> This may explain why the native fungal PPTase may not be functional for DEBS1-TE. The comparison between SePptII and Sfp also showed the distinctive difference in conserved areas. According to the difference of conserved motifs, SePptII was classified into the FxxKESxxK subtype, while Sfp belongs to the WxxKEAxxK subtype. Therefore, the heterologous function of these two PPTases for DEBS1-TE in *A. oryzae* is worthwhile to explore so as to make up the lack of necessary biosynthetic routes.

S. erythraea-SePptII	-----miervlpegatwveafddpaeatlfee	28
B. subtilis-Sfp	-----mkiygiydm-rpls-qeenerfmtfispekrekrrfyh-----ke	39
A. oryzae-PPTase	--mqppqdessnmcvrwyidtrdltatstslpilletlqppdqesakrfyh-----lk	50
A. nidulans-NpgA	mwqdtssastspilrtwyidtrplstastaalpilletlqpdaqisvqkyyh-----lk	52
	: : * . . . :	
S. erythraea-SePptII	eaaiaravdkrrrefrtvrhcarramaelgvppapllpgergapqw-----	74
B. subtilis-Sfp	dahrtllgdvivrsv-isrqqyl-dksdirfst---q--eygkpci-----	78
A. oryzae-PPTase	dkhmslasnllkylf-ihrtrci-pwnqitiser---tpaphhrpyfnaagfiqtaa--td	103
A. nidulans-NpgA	dkhmslasnllkylf-vhrncrri-pwssiviser---tpdphrrpocyippsgsqedsfkdg	107
	: : : . * . . *	
S. erythraea-SePptII	-----pagvvgsmthcagy-----raavvpgagtvvtigidae	107
B. subtilis-Sfp	pdlpdahfnishsgrwvigafd-----sqpigidie	109
A. oryzae-PPTase	kpipniefnvshqaslvalagtilppssnndsiaptnvitnnpntstpassipqvgidit	163
A. nidulans-NpgA	ytginvefnvshqasmvaiagtaftpnsggds-----klkpevgidit	150
	. . . * :***	
S. erythraea-SePptII	pheplpggvldavslpe-----erarlrel-----stqdkvhwdrv1	145
B. subtilis-Sfp	ktkp-----isl-----eiakrffskteysdl-----lakdk----deqtdy	142
A. oryzae-PPTase	cvnerrntpetrqaedlhgvp-----shigndedglveygfr1	202
A. nidulans-NpgA	cvnerqgrngeersleslrqyidifsevfstaemaniirrlgdvssslsadrldygyr1	210
	: : *	
S. erythraea-SePptII	fsckesvykawfpltgawldfeqadltfdaaggtfharllkpgg-qaegrp1teftgrw1	204
B. subtilis-Sfp	fy-----hlwsnkesfikqegkglslpldsfsvr1hqdgqvsielpdshspcyikt ye	195
A. oryzae-PPTase	fy-----tywalkeayikmtgeallapw1relvftnvlapep---agrhlhtwgepyt	252
A. nidulans-NpgA	fy-----tywalkeayikmtgeallapw1relefsnvvapaavaesgdsgagdfgepyt	263
	* : : : : * . . : :	
S. erythraea-SePptII	aadgvvsai-----vrlrer-----	220
B. subtilis-Sfp	vdpgykmavcaahpdfpedi--tmvsye--ell-----	224
A. oryzae-PPTase	---gvktwly---gkevedvr1evvafendyliataarggg1gwrsegd-----ga	297
A. nidulans-NpgA	---gvrttly---kn1vedvr1evaalggdylfataarggg1gassrpgggpdgsgirsq	317
	* :	
S. erythraea-SePptII	-----	220
B. subtilis-Sfp	-----	224
A. oryzae-PPTase	dpwqrleki dl ekdvrcatgvqcqlk	324
A. nidulans-NpgA	dpwrrpfkkl1dierdi qp catgvoncls	344

**Figure 2.10** Sequence alignment of different sources of PPTases (the conserved motifs are highlighted in grey).

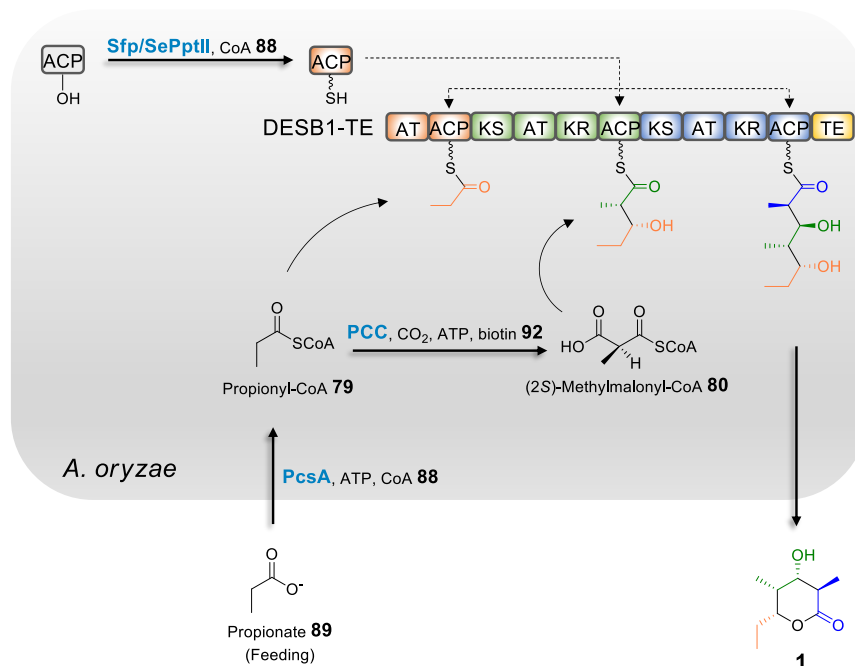
Before the heterologous expression of PPTase genes *sfp* and *sePptII*, their nucleotide sequences were codon-optimized by synthesis. Two homologous fragments from the vector pTYGS-met **XV** were then linked to ends of each gene by PCR. Subsequently, the resulting fragments were inserted downstream of the promoter *P<sub>amyB</sub>* on the plasmid pTYGS-met **XV** by yeast recombination. Ultimately, two plasmids **XVI** and **XVII** carrying different PPTases genes were prepared for the coexpression with DEBS1-TE (Scheme 2.20).



**Scheme 2.20** Construction of plasmids containing different bacterial PPTase genes by yeast recombination: **A**, the construction of the plasmid **XVI** bearing *sfp*; **B**, the construction of the plasmid **XVII** bearing *sePptII*.

### 2.3.6 Product Identification

With the supplementation of the phosphopantetheinylation pathway, the complete engineered pathway for the DEBS1-TE heterologous expression in *A. oryzae* NSAR1 was established as shown in Scheme 2.21.

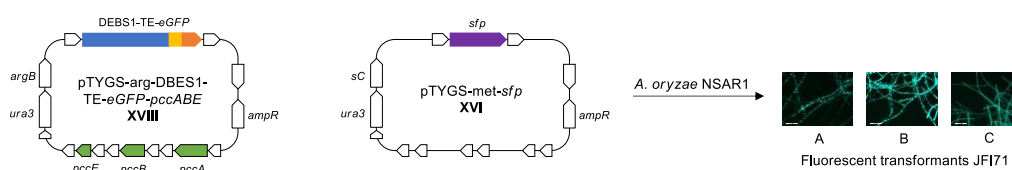


**Scheme 2.21** Engineered pathway for DEBS1-TE heterologous expression in *A. oryzae* NSAR1.

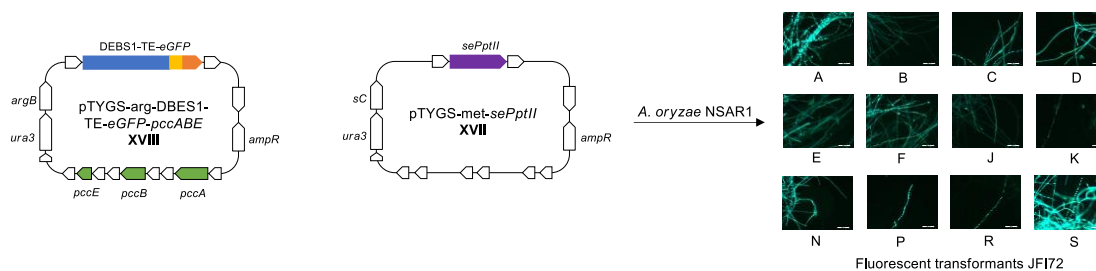
The plasmid pTYGS-arg-DEBS1-TE-*eGFP-pccABE XVIII* was constructed using the plasmid pEYA-DEBS1-TE-*eGFP X* as the entry clone and pTYGS-arg-*pccABE II* as the

destination vector by LR recombination. **XVIII** was then introduced into *A. oryzae* NSAR1 along with the pTYGS-met plasmid bearing a single PPTase gene. Two groups of fungal transformation were conducted simultaneously (Scheme 2.22), and transformants were grown on the CZD/S selection media that lack arginine and methionine at the same time. After screening and cultivation of transformants, the group with *sePptII* obtained more transformants (20 per transformation) than that of the group with *sfp* (3 per transformation). All transformants were cultivated in DPY media for 2 days, and then the mycelia were inspected under the fluorescence microscope. In the group with *sfp*, all three transformants (JFI71A-C) were observable with different extents of green fluorescence (Scheme 2.22 A). In the other group, 12 transformants (JFI72A-F, J-K, N, P, R and S) with *sePptII* were green fluorescence-positive (Scheme 2.22 B). By comparative analysis, SePptII is more efficient due to the more transformant number and higher intensity of green fluorescence.

A



B



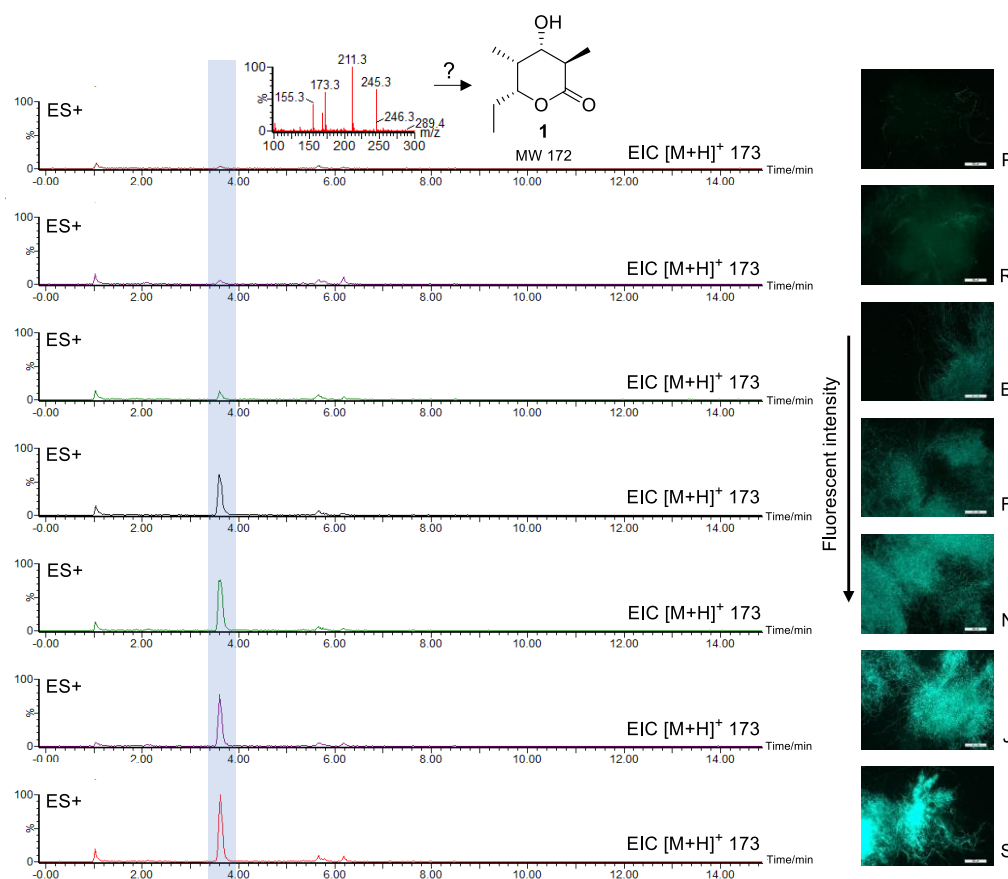
**Scheme 2.22** Cotransformation of DEBS1-TE, PCC and PPTase genes and the fluorescent imaging of transformants:

**A**, the PPTase gene is *sfp*; **B**, the PPTase gene is *sePptII*.

Only green fluorescence-positive transformants remained to proceed cultivation in DPY media. In the fifth day, 50 mM propionate **89** were fed into each culture and incubated with shaking overnight. Then, the cultures were collected and extracted with ethyl acetate. Lastly, the extracts were analyzed by LCMS. On neither DAD nor TIC chromatogram, any new peaks related to **1** can be observed. Additionally, the

EIC chromatogram of  $[M-H]^-$  171 had no relevant signals either. However, gratifyingly and for the first time, the EIC chromatogram of  $[M+H]^+$  173 showed a high and single peak in some transformants (Fig. 2.11). The fragmentation pattern contains an ion with  $m/z$  155, that appears to be caused by losing one molecule of water.

Apparently, different transformants produced different levels of this new  $[M+H]^+$  173 product. Given that different intensity of green fluorescence of transformants, there might be a correlation between the yield of this unknown product and fluorescence intensity. The alignment of these two sets of data validated this hypothesis. As the increase of fluorescence intensity, the production of this unknown product was enhanced as well. It was plausible that this new product is closely relevant to DEBS1-TE in *A. oryzae*.

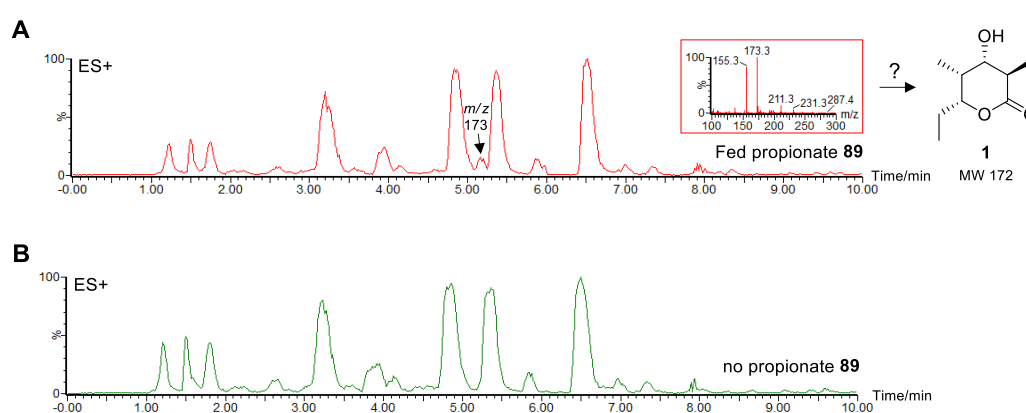


**Figure 2.11** EIC chromatograms of transformants JF172P, R, E, F, N, J, S and alignment with the fluorescence intensity.

To further clarify the correlation of the newly-formed product and DEBS1-TE, another feeding experiment was performed using the transformant with highest fluorescence intensity. 50 mM propionate **89** were fed into the culture when it was



cultivated until the day before extraction. Meanwhile, the control group without propionate supplementation was set up. The extraction and analytical procedures were done in parallel. In addition, to make comparison clear, the analytical conditions of LCMS were optimized with a shallower elution gradient. This led to the newly-formed product being visible in the ES+ TIC chromatogram (Fig. 2.12 A). In contrast, at the retention time around 5.15 min, the product was absent on the chromatogram of the control group (Fig. 2.12 B). The production of the  $m/z$  173 product is only attributed to the addition of propionate. Consequently, it suggested that the newly-formed product is very likely to be the expected product triketide lactone **1**.

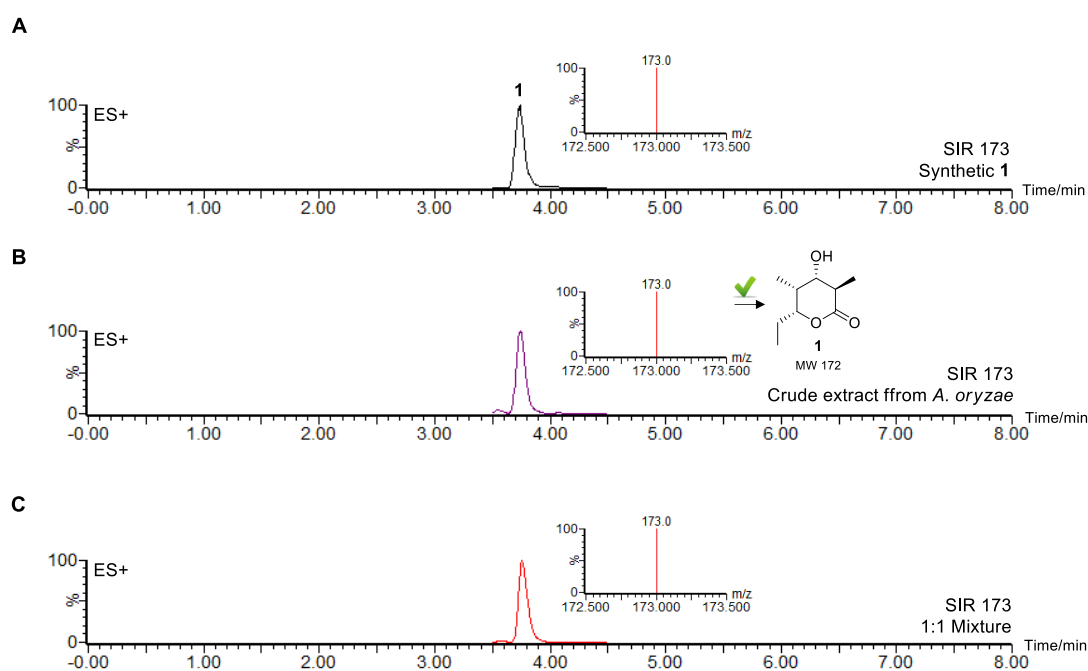


**Figure 2.12** Determination of the correlation of the formation of the  $m/z$  173 product and the supplementation of propionate **89**: **A**, ES+ TIC chromatogram of transformant with feeding propionate; **B**, ES+ TIC chromatogram of transformant without feeding propionate.

NMR characterization is generally the most precise way to confirm the structure of a compound. In this case, the isolation of the product was preformed from 2 L of the culture. However, less than 1 mg of the product was obtained finally due to its low yield. Furthermore, the relatively high polarity of the product resulted in its low purity (less than 50 %). Therefore, NMR characterization for structure elucidation was not possible.

The alternative method is to compare with the authentic compound. The authentic triketide lactone **1** was chemically synthesized by Maurice Hauser. Using it as the control, the crude extract from *A. oryzae* was comparatively analyzed with the method of selected ion recording (SIR) by LCMS (Fig. 2.13). Only the peak with  $[M+H]^+$  173 was set to record on ES+ TIC chromatogram. It turned out that peaks

recorded on both of SIR chromatograms had the identical retention time (Fig. 2.13 A and B). Subsequently, the authentic **1** with the concentration of  $20\ \mu\text{g}\cdot\text{mL}^{-1}$  was mixed with the crude extract from *A. oryzae* at the volume ratio of 1:1, followed by SIR analysis on LCMS. It also demonstrated that it was a homogenous peak, not splitting into two peaks (Fig. 2.13 C). Moreover, the fragmentation pattern of the authentic **1** was also consistent with that of the newly-formed product from *A. oryzae* transformant. The parent ion  $m/z$  173 and the daughter ion  $m/z$  155 were predominantly present on both of them.

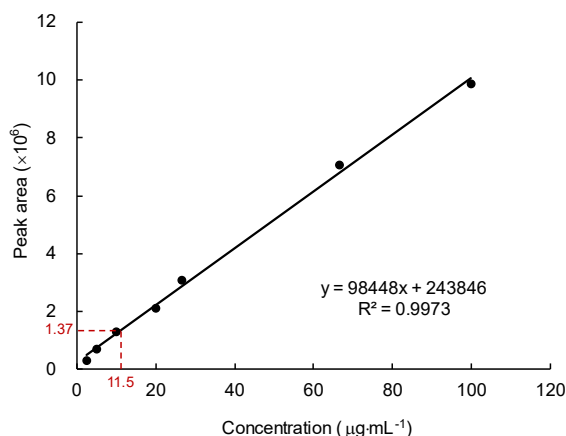


**Figure 2.13** Comparative analysis of the synthetic **1** and the crude extract from *A. oryzae* transformant: **A**, ES+ SIR chromatogram of synthetic **1**; **B**, ES+ SIR chromatogram of crude extract from *A. oryzae* transformant; **C**, ES+ SIR chromatogram of 1:1 mixture of samples above.

Taking all results above into consideration, it was concluded with certainty that the newly-formed product is the expected triketide lactone **1**. That is, the entire engineered biosynthetic pathway for the DEBS1-TE heterologous expression was properly established in the fungus *A. oryzae*.

### 2.3.7 Titer Calculation

A standard method was established for the titer measurement of triketide lactone **1** using the authentic compound. First, the compound **1** was dissolved in methanol and diluted to different concentrations from 2.5  $\mu\text{g}\cdot\text{mL}^{-1}$  to 200  $\mu\text{g}\cdot\text{mL}^{-1}$ . Each concentration of **1** was analyzed and the corresponding product peak on SIR chromatogram was integrated. Then, a calibration curve was made based on the concentrations and peak areas of compound **1** (Fig. 2.14).



**Figure 2.14** The establishment of the calibration curve for the titer calculation of the product **1**. The red numbers are the calculated values of **1** produced by *A. oryzae* transformant.

Then, a standard extraction workflow was determined. The transformant was inoculated in 100 mL DPY media. After 4 days of cultivation, 50 mM propionate was added into the culture. The culture with propionate was incubated overnight. Next, the five-day culture was extracted with the same volume of ethyl acetate twice. The organic phase was gathered and evaporated to dryness. The concentrated extract was re-dissolved in 1.5 mL methanol. Subsequently, the extract suspension was diluted 10 times with methanol. The sample was centrifuged at  $14,000 \times g$  for 5 min. Lastly, the supernatant was collected and analyzed by LCMS using the SIR detection method.

The extract from 100 mL of the transformant culture were applied to the quantitative analysis based on the calibration curve (Fig. 2.14). By calculation, the titer of the triketide lactone product **1** in *A. oryzae* was approximately  $0.6 \text{ mg}\cdot\text{L}^{-1}$ .

## 2.4 Discussion and Conclusion

Fungi possess considerable potential as hosts for the heterologous production of natural products. In this project, the core aim was to further expand the application of the fungal heterologous expression systems, especially of the *Aspergillus* genus from filamentous fungi. Except for iterative PKS, the fungal systems are also expected to be capable of expressing modular PKS, which have never been expressed in fungi. Thus, DEBS1-TE, a minimal bimodular PKS, was taken as a representative model system to conduct expression in *Aspergillus oryzae* in this work.

During the construction process of the heterologous pathway, several different factors had to be considered. In terms of precursors, propionate could be fed and subsequently converted to the starter unit propionyl-CoA **79** by a native propionyl-CoA synthetase which is the starter unit for DEBS1-TE. Further, the pathway was extended by a propionyl-CoA carboxylase system to lead to formation the extender unit (2*S*)-methylmalonyl-CoA **80**. Despite the different codon usage bias and GC content, the transcription and expression of DEBS1-TE in *A. oryzae* were not influenced markedly. The eGFP-fused expression of all four transformed bacterial genes demonstrated the high abundance of full-length DEBS1-TE enzyme in *A. oryzae*. We also showed that the native fungal PPTase seems incapable of activating the bacterial modular PKS, but this was easily solved by expressing the appropriate bacterial enzyme.

With this engineered *A. oryzae*, the heterologous expression of a modular PKS was functionally achieved in a filamentous fungal host for the first time. This fungal system incorporates independent engineered pathways so that a numerous number of modular PKS can be applied to express in *A. oryzae*. Moreover, this engineered pathway in combination with the potential of abundant tailoring enzymes in fungi will produce more various types of natural products.

## 3. High-titer Production Optimization of Triketide Lactone in *A. oryzae*

### 3.1 Introduction

In the last chapter, a new fungal heterologous expression system was established in *A. oryzae*. DEBS1-TE has been the first modular PKS expressed in the fungal host. However, the titer of the final product triketide lactone **1** is only  $0.6 \text{ mg}\cdot\text{L}^{-1}$ , that largely limits the further expansion and application of fungal heterologous expression of modular PKS. Therefore, some essential factors influencing the titer of production are reviewed in this chapter.

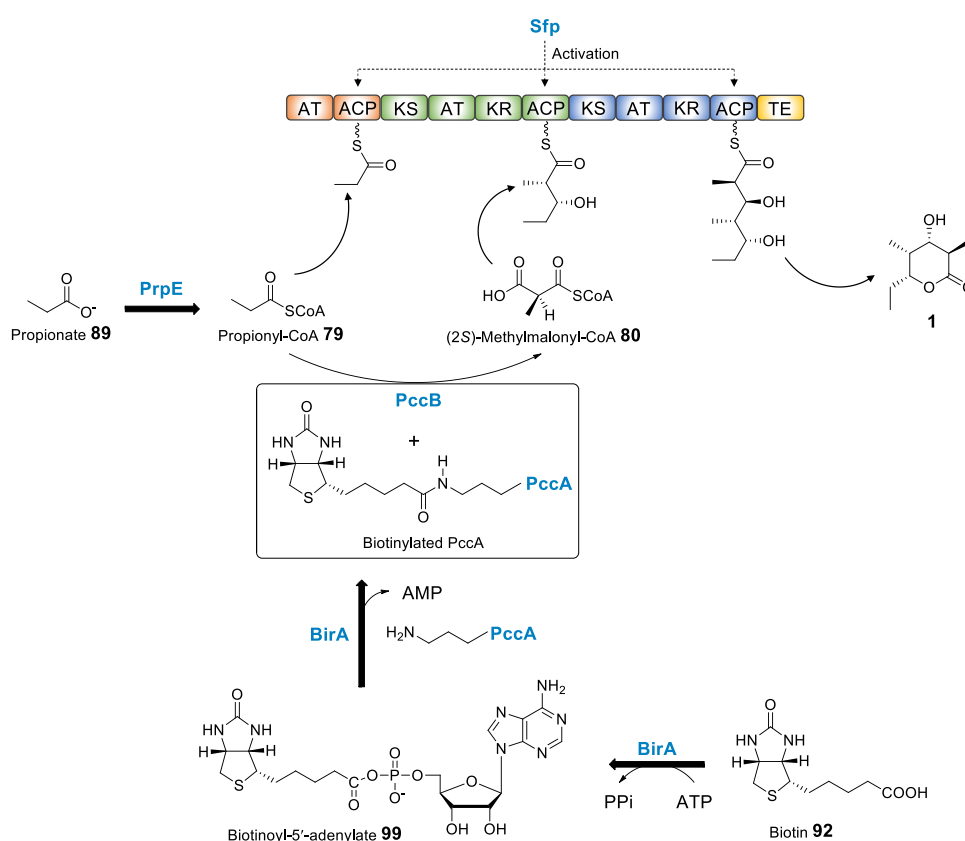
#### 3.1.1 Fermentation Conditions

Fermentation is one of the most important determinants for production of natural products in a heterologous host. A number of parameters (*e.g.* temperature, pH, cultivation time, and shaking speed) can affect the production level of desired products during fermentation. For example, in the heterologous host *E. coli* the DEBS1-TE expression level was found to be comparable to that in *S. erythraea*.<sup>122</sup> However, the activity of DEBS1-TE was not detectable at 30 °C. Upon lowering the expression temperature to 22 °C, DEBS1-TE restored its activity in the lysates of *E. coli*. The corresponding triketide lactone **1** was also produced in detectable amounts.<sup>75</sup>

#### 3.1.2 Gene Expression Level

The expression level of the targeted gene is also a fundamental factor in biosynthesis of natural products. In particular, some key pathway enzymes determine the yield of the final product. The transcriptional control for a gene generally takes place at promoter elements that drive gene expression. Endogenous promoters sometimes do not enable to maximize the transcription level of the targeted gene within the native host.<sup>123</sup> Therefore, some strategies for increasing the gene expression level may be necessary in metabolic engineering.

Gene overexpression is one of the most widely applied avenues. For example, the DEBS1-TE expression in *E. coli* was also engineered by gene overexpression (Scheme 3.1). The native propionyl-CoA synthetase gene *prpE* was overexpressed to supply sufficient amounts of precursor propionyl-CoA **79**.<sup>75</sup> In addition to that, the *E. coli* biotin ligase gene *birA* was also overexpressed. It involves the activation of biotin **92** to form the biotin-propionyl-CoA carboxylase complex, through which the carboxyl group is transferred. The overexpression of both genes facilitated the production of the final product **1**.



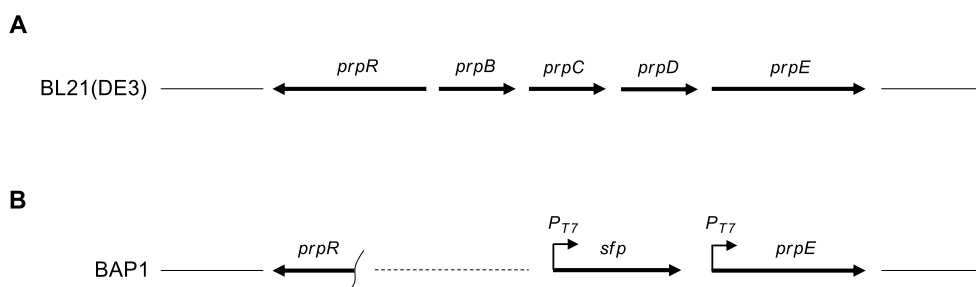
**Scheme 3.1** Overexpression of genes *prpE* and *birA* in DEBS1-TE expression in *E. coli*. Bold arrows represent the overexpressed pathways.

### 3.1.3 Degradation Pathways

#### *E. coli*

During the biosynthesis of natural products, the degradation of precursors and intermediates also has a significant influence on the titer of the final product. Therefore, deleting or blocking the degradation pathway to enhance substrate flux through the polyketide pathway is an effective approach.

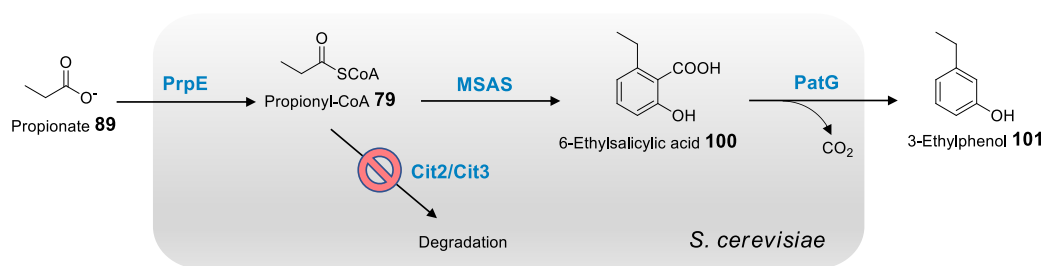
For example, in the *E. coli* strain BL21(DE3) for DEBS1-TE expression, the propionate catabolism operon *prp*, which encodes 2-methylisocitrate lyase (PrpB), 2-methylcitrate synthase (PrpC), and methylisocitrate synthase (PrpD), was interrupted by inserting an exogenous *sfp* gene to form a mutant strain *E. coli* BAP1 (Fig. 3.1).<sup>75,124</sup> The result was that the ability of *E. coli* to utilize propionate **89** as a carbon source was eliminated.



**Figure 3.1** Interruption of the propionate catabolism *prp* operon in *E. coli*: **A**, the *prp* operon in the general heterologous expression host strain *E. coli* BL(DE3); **B**, the interrupted *prp* operon in the engineered strain *E. coli* BAP1. *prpR*, transcription factor; *P<sub>T7</sub>*, T7 promoter; *prpBCDE* and *sfp* encode 2-methylisocitrate lyase, 2-methylcitrate synthase, and methylisocitrate synthase, propionyl-CoA synthetase, and phosphopantetheinyl transferase, respectively.

## *S. cerevisiae*

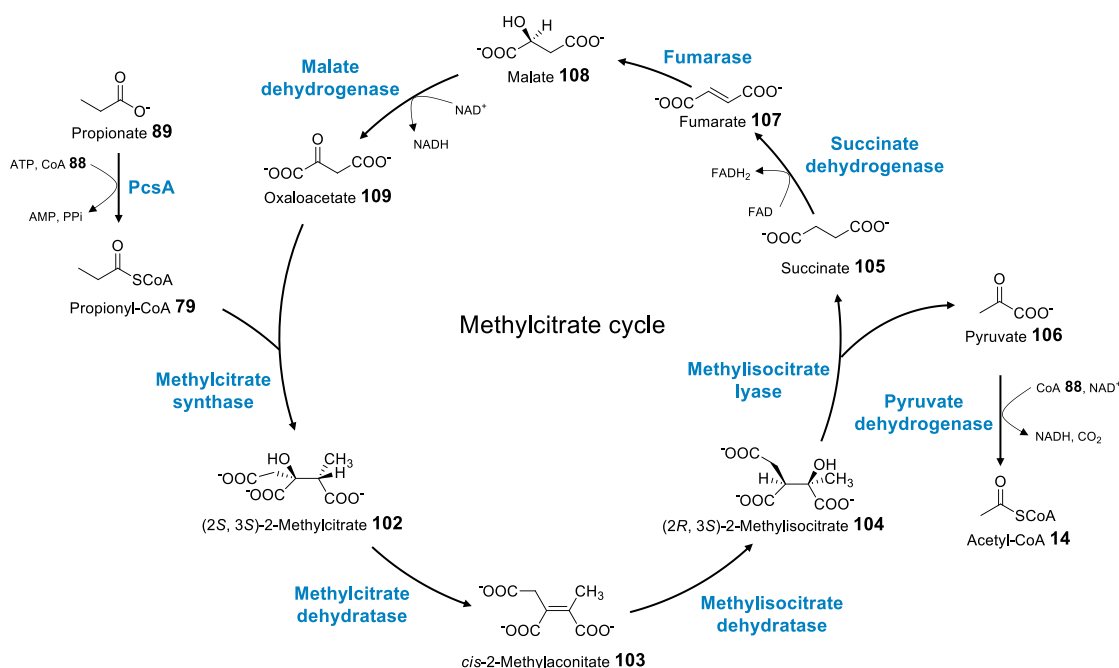
Although the current DEBS1-TE expression studies in yeast have not involved engineering the degradation pathway of precursors, the relevant degradation pathways are also inevitable due to the presence of known propionate catabolism pathways. The degradation of intracellular propionyl-CoA **79** is mediated by a 2-methylcitrate synthase. It was shown that the abolition of 2-methylcitrate synthase activity in a  $\Delta cit2 \Delta cit3$  deletion strain prevented propionate degradation.<sup>125</sup> In *S. cerevisiae*, *cit2* and *cit3* encode peroxisomal citrate synthase and mitochondrial citrate synthase, respectively. These enzymes possess dual activity as both citrate synthase and methylcitrate synthase. In a recent study on the heterologous production of a tsetse fly attractant 3-ethylphenol **101** in *S. cerevisiae*, *cit2* and *cit3* were deleted to abolish the degradation of the precursor propionyl-CoA **79**, which derives from feeding exogenous propionate **89** converted by overexpressed PrpE. Feeding propionate **89**, blocking intracellular propionyl-CoA **79** degradation and simultaneous overexpression PrpE led to an increase in the titer of 3-ethylphenol **101** from 0.18 mg·L<sup>-1</sup> to 12.5 mg·L<sup>-1</sup> in *S. cerevisiae* (Scheme 3.2).<sup>126</sup>



**Scheme 3.2** Metabolic pathways for 3-ethylphenol **101** production in *S. cerevisiae*. PrpE, propionyl-CoA synthetase; MSAS, 6-methylsalicylic acid synthase; PatG, 6-methylsalicylic acid decarboxylase; Cit2/Cit3, citrate synthases.

## Filamentous Fungi

Propionate is commonly used as a food preservative because it is toxic to many filamentous fungi and also inhibits the production of some mycotoxins.<sup>127</sup> Propionate **89** is converted to propionyl-CoA **79** *in vivo*, which can inhibit the growth of *Aspergillus* species and impair polyketide toxin production, such as sterigmatocystin in *Aspergillus nidulans*.<sup>98–100</sup> Despite its toxicity, propionate **89** is still able to be utilized by fungi as the sources of carbon and energy.



**Scheme 3.3** Methylcitrate cycle for oxidation of propionate **89** to pyruvate **106**.

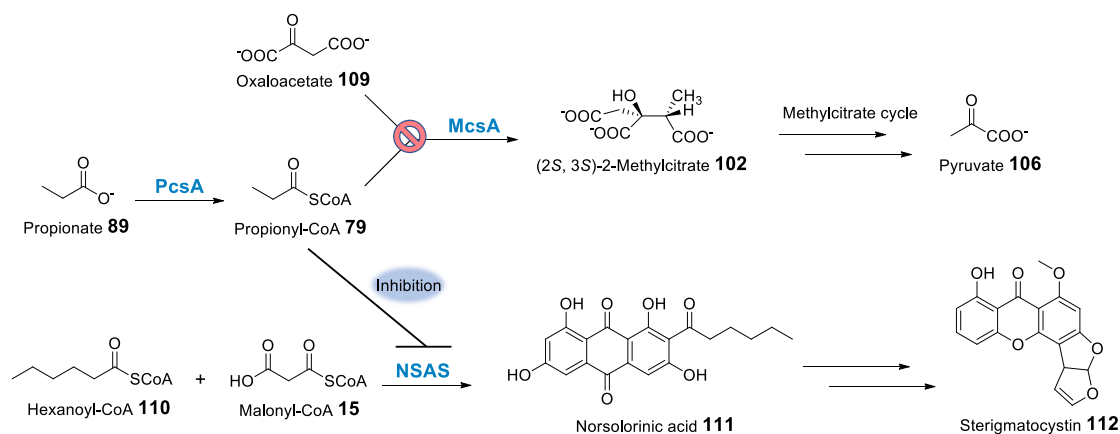
A methylcitrate cycle for  $\alpha$ -oxidation of propionate **89** to pyruvate **106** is known to be present in fungi (Scheme 3.3).<sup>128</sup> It starts from methylcitrate synthase, which condenses propionyl-CoA **79** with oxaloacetate **109** to give (2S, 3S)-2-methylcitrate



**102**. (2*S*, 3*S*)-2-methylcitrate **102** is then isomerized to (2*R*, 3*S*)-2-methylisocitrate **104** via *cis*-2-methylaconitate **103**. (2*R*, 3*S*)-2-methylisocitrate **104** is cleaved to succinate **105** and pyruvate **106** by a methylisocitrate lyase. Succinate **105** is successively degraded by succinate dehydrogenase, fumarase and malate dehydrogenase to regenerate oxaloacetate **109**, which enters the methylcitrate cycle again. Pyruvate **106** can also be further converted to acetyl-CoA **14** or serve directly as a building block for biosynthesis.

In *Aspergillus* species, propionate **89** metabolism was first investigated in *Aspergillus nidulans*. The key enzyme methylcitrate synthase was identified in *A. nidulans* (McsA, accession number CAB53336).<sup>95</sup> Its protein sequence shares more than 50 % identity to those of other eukaryotic citrate synthases such as *Homo sapiens* (accession number AAC25560) and *Saccharomyces cerevisiae* (accession number P00890), but only 14 % identity to that of *E. coli* (accession number AAB18057).<sup>124,129,130</sup> The *mcsA* deletion strain was unable to grow on propionate as a sole carbon source due to the accumulation of propionyl-CoA **79**. Furthermore, the polyketide-derived pigmentation was abolished in the *mcsA* mutant at a lower concentration of 20 mM propionate compared with 100 mM propionate for the wild-type strain.

In another mutagenesis study, it was found that an *A. nidulans* strain defective in production of the polyketide sterigmatocystin **112** was a spontaneous *mcsA* mutant. The presence of propionate **89** was able to inhibit the production of sterigmatocystin **112** and induce the transcription of *mcsA* in the wild-type strain at the same time. However, in the *mcsA* mutant, the blockage of the methylcitrate cycle led to the accumulation of excess propionyl-CoA **79** and indirectly inhibited the production of sterigmatocystin **112** (Scheme 3.4). In contrast, the overexpression of *mcsA* in the wild-type strain relieved inhibition of sterigmatocystin **112** biosynthesis by propionate **89**.<sup>100</sup> All of the above results support the hypothesis that methylcitrate synthase (McsA) is present in *Aspergillus nidulans* to direct propionyl-CoA **79** to enter the methylcitrate cycle.

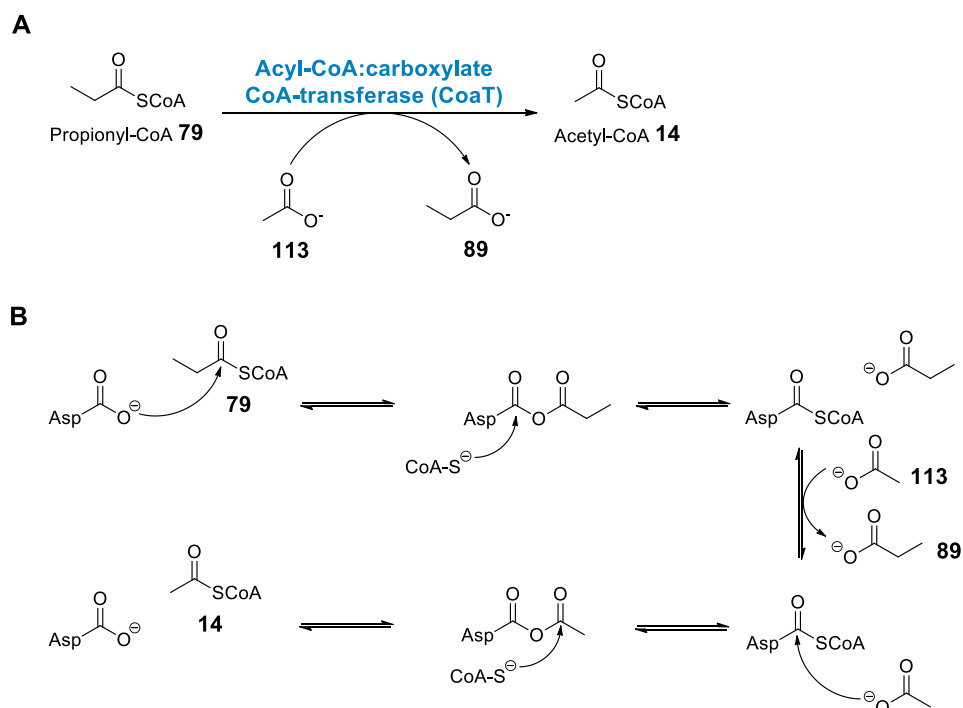


**Scheme 3.4** Biosynthesis of sterigmatocystin **112** and inhibition mechanism of propionyl-CoA **79**. NSAS: norsolorinic acid synthase.

In addition to *mcsA* from *A. nidulans*, homologous methylcitrate synthases have been identified in *A. fumigatus* and *A. niger*, respectively (accession numbers CAI61947 and BAM84220).<sup>103,131,132</sup> Both of them displayed not only methylcitrate synthase activity but also citrate synthase activity. Similarly, the deletion of these methylcitrate synthase genes resulted in the inability of these organisms to grow on propionate as the sole carbon source. In addition to these confirmed genes encoding methylcitrate synthase in *Aspergillus* species, two methylcitrate synthases FsmCsA (accession number CAZ64274) and FvMcsA (accession number CAZ64275) were discovered in *Fusarium solani* and *Fusarium verticillioides*, respectively, both of which appear adapted to growth on propionate as the carbon source.<sup>133</sup> FsmCsA and FvMcsA shares 96 % identity to each other and both exhibited low  $K_M$  values of propionyl-CoA **79**, that indicates the ability of propionyl-CoA detoxification.

However, it was also found that when acetate was used as the main carbon source, the growth of the wild-type *A. fumigatus* strain was not negatively influenced by addition of propionate. Despite the accumulation of propionyl-CoA, acetate appears competitive with propionate for activation.<sup>131</sup> Based on this finding, an acyl-CoA:carboxylate CoA transferase enzyme (CoaT, accession number CAP49328) was characterized in *A. nidulans*.<sup>134</sup> It is able to transfer the CoASH moiety from succinyl-CoA **94** and propionyl-CoA **79** to propionate **89** and acetate **113** (Scheme 3.5 A). According to the substrate type of CoA donor and acceptor, CoaT is classified as a Class III CoA-transferase. Its mechanism is shown in Scheme 3.5 B.<sup>135</sup> A propionyl-enzyme intermediate is formed by reaction of an enzyme-bound aspartate with the

donor propionyl-CoA **79**, followed by the formation of an enzyme-bound aspartyl-CoA thiolester intermediate. Subsequently, the thiolester reacts with the acceptor acetate **113** to give a new acetyl-enzyme anhydride from which the acetyl group is transferred to CoA.



**Scheme 3.5** CoA-transferase (CoaT) for propionyl-CoA detoxification and its mechanism: **A**, the conversion from propionyl-CoA **79** to acetyl-CoA **14** by CoaT; **B**, the mechanistic proposal of CoaT.

In the presence of both acetate **113** and propionate **89**, the *coaT* single mutant of *A. nidulans* displayed the inhibited phenotype, which is weaker than that of the *mcsA* single mutant under the same conditions. The growth of an *mcsA/coaT* double-deletion mutant was completely abolished on acetate/propionate medium. Therefore, it was proposed that in the wild-type situation propionyl-CoA **79** detoxification is conducted by CoaT without the urgent need of McsA. The inhibitory effort of propionate is attenuated because of addition of acetate. It is consistent with the inhibition of phenotype in *mcsA* mutant, which produced a minor phenotype than the wild-type. In contrast, in the *coaT* single mutant, it leads to the sufficient activation of McsA to remove propionyl-CoA **79**. The weakest visible phenotype is produced. The double deletion of both *mcsA* and *coaT* in *A. nidulans* therefore completely blocks the consumption pathways of propionyl-CoA **79**, leading to the inability to grow in the presence of propionate.<sup>134</sup>

*Aspergillus oryzae*, the expression host for our DEBS1-TE experiments, is closely related to the *Aspergillus* species discussed above. For DEBS1-TE expression in *A. oryzae*, we therefore expect the presence of both a methylcitrate cycle and CoA transferase for propionyl-CoA detoxification. The blockage of these degradation pathways is likely to be important for increasing the concentration of propionyl-CoA **79** and thereby improving the titer of the final polyketide product **1**.

## 3.2 Aims

As described in the last chapter, the modular PKS was expressed in filamentous fungi for the first time. It is the bimodular PKS DEBS1-TE, which was expressed in the heterologous host *A. oryzae*. However, the titer of the final polyketide product **1** is quite low. Here, some possible factors affecting the titer are analyzed and summarized, including fermentation conditions, gene expression level, and interference of precursor degradation pathways. For some similar features, many strategies used in engineered *E. coli* and yeast for DEBS1-TE expression are worthwhile for adoption in *A. oryzae*. However, there are also different aspects, such as dual propionyl-CoA **79** degradation pathways in *Aspergillus* species, which need to be specifically investigated.

In this chapter, the aim is to enhance production of precursors by optimizing biosynthetic pathway in *A. oryzae*. Through maximizing the supply of substrates and blocking their degradation, a higher titer of the final polyketide **1** could be achieved.

### 3.3 Results

#### 3.3.1 Optimization of Feeding Propionate

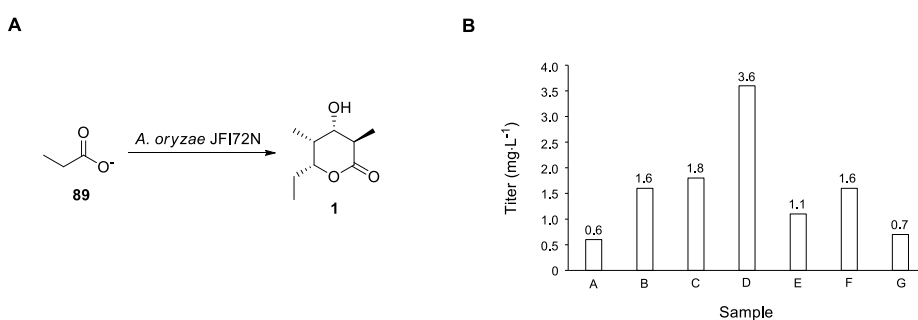
To improve the yield of the final polyketide **1** formed by DEBS1-TE, the feeding of propionate **89** was first optimized. In the prior propionate toxicity experiments, the minimal inhibition concentration of sodium propionate for *A. oryzae* was determined to be 50 mM. In other words, the maximal feeding amount of sodium propionate should not exceed 480 mg in 100 mL cultures. Based on this limitation, the optimization was performed under the defined total amount of propionate. Notwithstanding, the feeding rate is still adjustable. A slower feeding rate can reduce the toxic effort on the growth of *A. oryzae*, but it also may extend the detoxification time, leading to consumption of precursor propionyl-CoA **79**. Therefore, the feeding rate needs to be optimized so that a balanced and optimal way could be found.

In order to test if the rate of addition of propionate can affect the production of **1**, parallel experiments were set up in which 100 mL cultures of *A. oryzae* NSAR1 or *A. oryzae* transformant JFI72N were supplemented with a series of lower propionate concentrations on consecutive days (Table 3.1). In an alternative scenario a controllable-rate syringe pump was used for constant slow addition. As samples C and E show in Table 3.1, *A. oryzae* transformant was inoculated at the first day. The feeding process was initiated at the second day and then continuously until the day of extraction. Samples C and D were set up with 4 and 5 feeding days, respectively.

**Table 3.1** List of samples (A-G) with different sodium propionate-feeding conditions. Abbreviations: Ino., inoculation of *A. oryzae* transformant; Cul., cultivation; Ext., extraction of fermentation cultures. Feeding time is highlighted in green.

Sample	Time (day)											
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	12th
A	Ino.	Cul.	Cul.	Cul.	50 mM	Ext.						
B	Ino.	12.5 mM	12.5 mM	12.5 mM	12.5 mM	Ext.						
C	Ino.	Fed 50 mM with pump				Ext.						
D	Ino.	10 mM	10 mM	10 mM	10 mM	10 mM	Ext.					
E	Ino.	Fed 50 mM with pump				Ext.						
F	Ino.	6.25 mM	6.25 mM	6.25 mM	6.25 mM	6.25 mM	6.25 mM	6.25 mM	6.25 mM	Ext.		
G	Ino.	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	Ext.

The production of **1** derived from exogenous sodium propionate **89** was detected and quantified in *A. oryzae* JFI72N using the selected ion recording (SIR) method by LCMS (Fig. 3.2 A). The titers of all samples were calculated and analyzed comparatively. The results show that sample D has the highest titer with up to 3.6 mg·L<sup>-1</sup>, which is 6 times of that of sample A (Fig. 3.2 B) and indicate that extending the feeding time is beneficial for *A. oryzae* to fully take the compound and convert it to the product. However, a longer feeding time, such as 8 or 10 days, also led to the reduction of the titer, possibly due to the propionyl-CoA **79** detoxification process or the degradation of the final product. Unexpectedly, the feeding with a syringe pump over 5 days gave rise to a poor titer. It appears that propionyl-CoA **79** tends to predominantly flow into the detoxification pathway when it is present in low concentrations in *A. oryzae*.



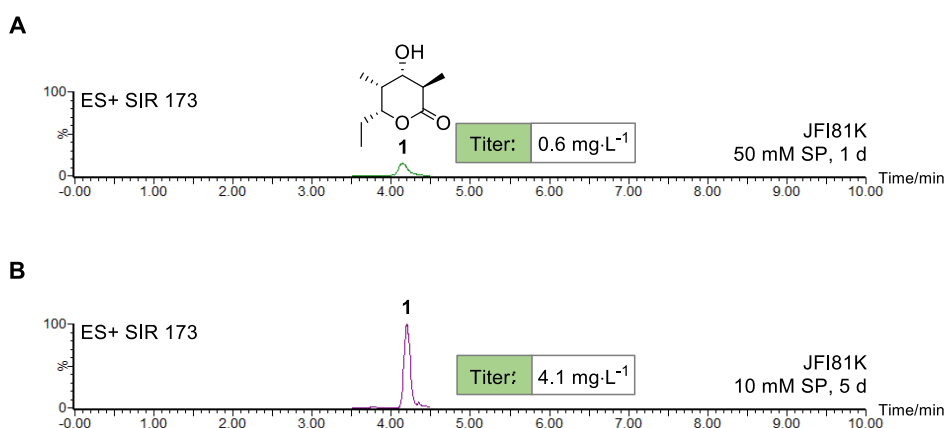
**Figure 3.2** Titer comparison of the product **1** at different feeding conditions: **A**, the conversion from propionate **89** to **1** by *A. oryzae* transformant JFI72N; **B**, titer comparison under different sodium propionate-feeding conditions. The titer for each sample is marked on the top of the bar.

### 3.3.2 Removal of eGFP

As described formerly, eGFP fused to the C-terminal of DEBS1-TE was coexpressed in *A. oryzae* for screening transformants with highly expressed-DEBS1-TE. Although the protein size of eGFP is only 27 kDa, it might affect the release of the product from TE domain of DEBS1-TE, leading to the reduction of titer of the final product **1**. Therefore, the removal of eGFP from DEBS1-TE expression cassette was attempted. After fungal transformation of DEBS1-TE without *eGFP*, 12 *A. oryzae* transformants (JFI81A-L) were obtained. All transformants were cultivated in DPY media, and at the day before extraction 50 mM sodium propionate **89** was added into cultures. Transformants were extracted and then analyzed by LCMS. It showed that 7 transformants produced the final product **1** on different degrees. Among them, the

transformant JFI81K yielded the product **1** most. The titer of **1** in JFI81K was calculated as  $0.6 \text{ mg}\cdot\text{L}^{-1}$  (Fig. 3.3 A), that is the same as that of the transformant JFI72N bearing eGFP.

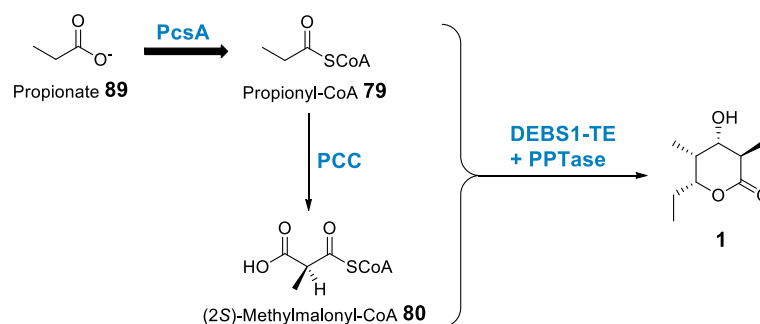
The feeding rate was also optimized for the transformant JFI81K. Sodium propionate **89** was added into the culture at the concentration of 10 mM per day over 5 days. It turned out that the titer of **1** was  $4.1 \text{ mg}\cdot\text{L}^{-1}$  (Fig. 3.3 B), which is slightly increased compared with  $3.6 \text{ mg}\cdot\text{L}^{-1}$  before removing eGFP. Nevertheless, it is plausible that eGFP is not a decisive factor for the titer improvement. The fusion of eGFP to DEBS1-TE does not significantly affect the release of the final product **1**.



**Figure 3.3** Titer calculation of **1** in *A. oryzae* transformant JFI81K without eGFP using the SIR method by LCMS: **A**, JFI81K was fed with 50 mM SP in one day; **B**, JFI81K was fed with 10 mM SP per day over 5 days. Abbreviation: SP, sodium propionate **89**.

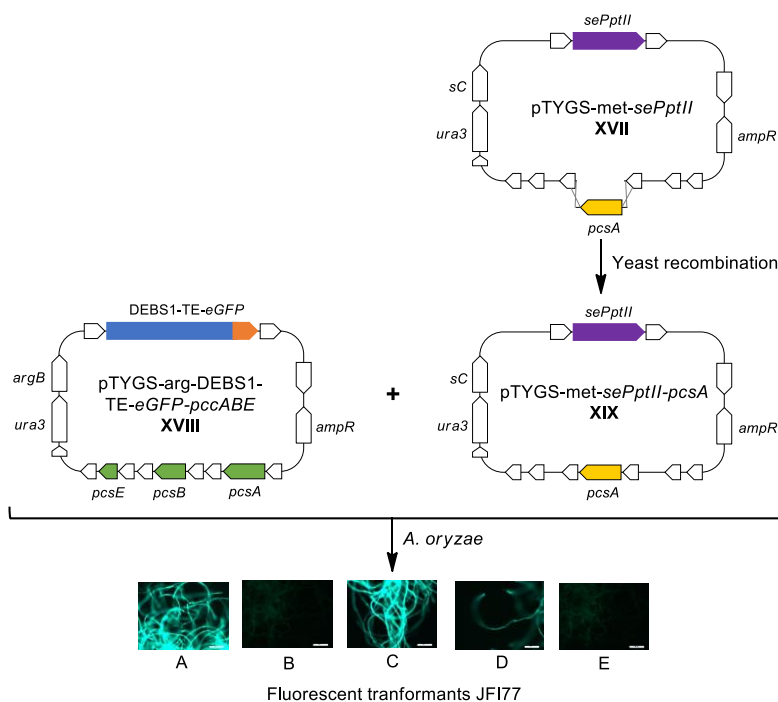
### 3.3.3 PcsA Overexpression

In the last chapter (2.3.1.2), a native propionyl-CoA synthetase gene *pcsA* has been identified in *A. oryzae* NSAR1. The transcription and expression of *pcsA* is induced by addition of exogenous propionate **89**, and then PcsA converts propionate **89** to propionyl-CoA **79** *in vivo*. Given that both the starter unit propionyl-CoA **79** and extender unit (2*S*)-methylmalonyl-CoA **80** stem from exogenous propionate **89**, the conversion efficiency of propionate **89** is speculated to be a crucial factor during the biosynthesis of triketide lactone **1**. Therefore, the native *pcsA* was attempted to overexpress in *A. oryzae* (Scheme 3.6).



**Figure 3.6** PcsA overexpression to increase the yield of 1. The grey bold arrow represents gene overexpression.

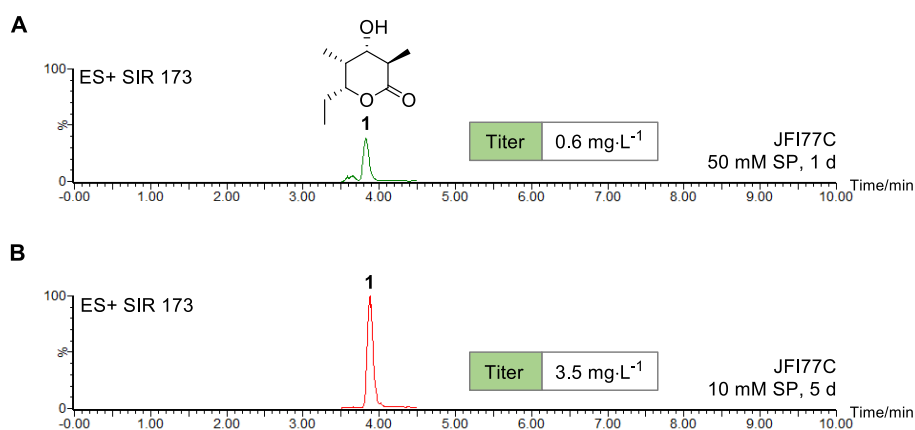
Here, the original *pcsA* in *A. oryzae* genome was not manipulated genetically. Instead, a new *pcsA* expression cassette was rebuilt *in vitro* using a copy of *pcsA* and a strong promoter followed by introduction into *A. oryzae* for overexpression. Furthermore, the introns within *pcsA* were not deleted due to the homology to *A. oryzae*. Therefore, *pcsA* was directly replicated from *A. oryzae* genome by PCR, and then inserted into the vector pTYGS-met-*sePptII* XVII by yeast recombination.<sup>63</sup> The strong constitutive promoter *P<sub>gpdA</sub>* is designated to drive the transcription of *pcsA*. The resulting plasmid pTYGS-met-*sePptII*-*pcsA* XIX along with DEBS1-TE and *pccABE* were introduced into *A. oryzae* (Scheme 3.7). Five transformants (JFI77A-E) were obtained. JFI77A, JFI77C, and JFI77D were detected as fluorescence-positive.



**Scheme 3.7** Construction of the *pcsA* cassette for gene overexpression and subsequent transformant generation.



Fluorescent transformants were cultivated in DPY, and then fed 50 mM sodium propionate **89** to the culture at the day before extraction. After cultivation, the extract was analyzed by LCMS. Lastly, the titer of the product **1** was determined as  $0.6 \text{ mg}\cdot\text{L}^{-1}$  (Fig. 3.4 A). When the feeding of sodium propionate was extended to 10 mM per day over five days, the production of **1** was improved with the titer of  $3.5 \text{ mg}\cdot\text{L}^{-1}$  (Fig. 3.4 B). Compared to the results before *pcsA* overexpression, the titer of the product was not increased significantly as expected. In the last chapter 2.3.1.2, we know that PcsA is induced by the addition of propionate, so it may be that it is already at a high enough level. Adding more PcsA *via* the *P<sub>gpdA</sub>* route here makes no additional difference. Therefore, the removal of the interference caused by propionyl-CoA degradation pathways is necessary for improving the titer of **1**.



**Figure 3.4** Titer determination of **1** in the *pcsA*-overexpressed transformant JFI77C using the SIR method by LCMS: **A**, JFI77C was fed with 50 mM SP in one day; **B**, JFI77C was fed with 10 mM SP per day over 5 days. Abbreviation: SP, sodium propionate **89**.

### 3.3.4 Deletion of *mcsA* and *coaT*

#### 3.3.4.1 Gene Identification

The degradation pathways of propionyl-CoA **79** have been proven present in *Aspergillus* species. These pathways are considered as detoxification pathways due to the toxicity of propionyl-CoA **79** to the growth of fungi and its inhibition of biosynthesis of some secondary metabolites. One key pathway proceeds *via* methylcitrate synthase (McsA) that is known to be involved in the metabolism of propionyl-CoA **79** to pyruvate **106** in several *Aspergillus* species, such as *A. nidulans*, *A. fumigatus* and *A. niger*. A second pathway is *via* acyl-CoA:carboxylate CoA transferase (CoaT). Despite only having been characterized in *A. nidulans*, the role

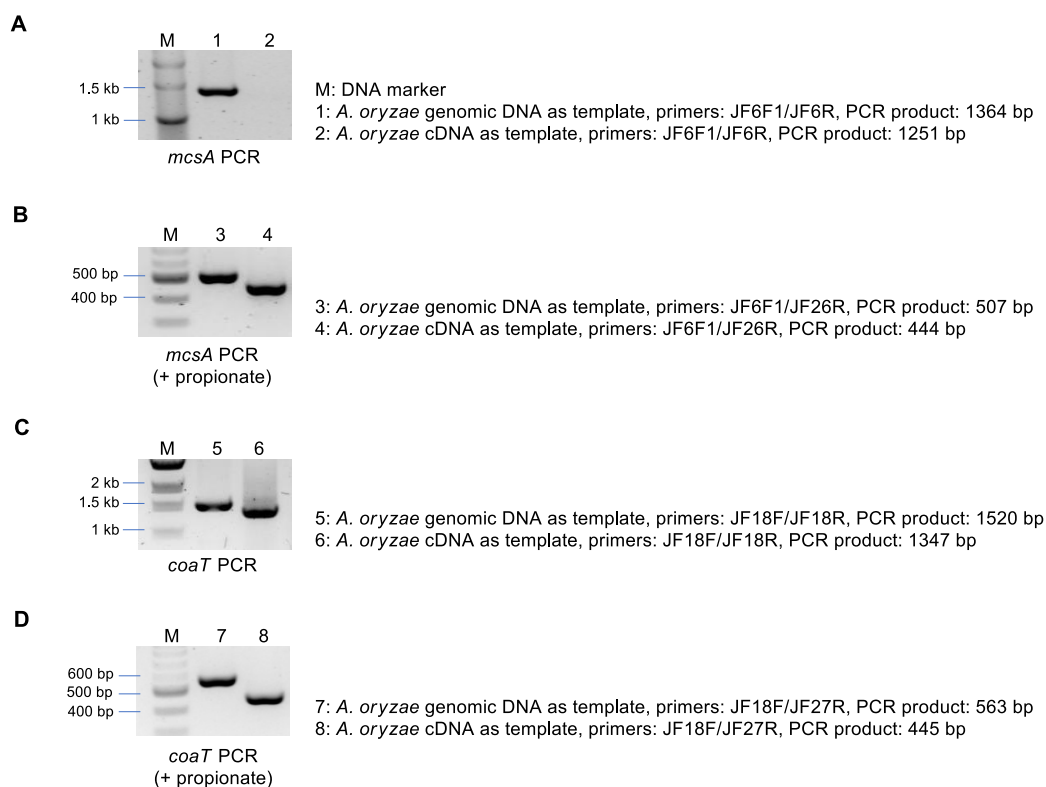
of CoaT for propionyl-CoA metabolism is also likely to be significant in *A. oryzae*. Therefore, in this section, the presence of the *mcsA* and *coaT* metabolic pathways in *A. oryzae* NSAR1 would be identified first.

The characterized fungal McsA proteins were used as the reference to search for a homolog in the genome of *A. oryzae* RIB40 using both the AspGD and NCBI databases. An *A. oryzae* homolog was found that shares 89 % identity to the *A. nidulans* McsA. The identities of other homologous genes are lower than 56 %. Similarly, one CoaT homolog was found in the *A. oryzae* RIB40; it shares 94 % identity to the characterized CoaT. The corresponding accession numbers of McsA and CoaT are XP\_001817006 and XP\_001817633, respectively. Based on these findings, primers were designed for gene identification of *mcsA* and *coaT* in *A. oryzae* NSAR1.

Genomic DNA and total RNA were extracted from the two-day-cultivated mycelia of *A. oryzae* NSAR1 grown in DPY liquid media, respectively. The total RNA was reverse-transcribed into cDNA. Using genomic DNA and cDNA as template individually, the gene identification of *mcsA* was conducted by PCR. It turned out that the homologous *mcsA* gene is present in *A. oryzae* NSAR1 genome, but under the conditions tested the gene is not transcribed (Fig. 3.5 A). Therefore, it does not express under the standard cultivation conditions. To determine the effect of propionate on *mcsA* expression, 50 mM sodium propionate was added into the *A. oryzae* culture at the day before genomic DNA or total RNA extraction. After extraction, the gene identification of *mcsA* was performed again by genomic DNA PCR and RT-PCR. It was found that *mcsA* is highly expressed under these conditions (Fig. 3.5 B). Consequently, it is deduced that the addition of propionate switches on the expression of *mcsA*. This result is accordance with published studies on *mcsA*.<sup>100</sup> The transcription analysis of *mcsA* in *A. nidulans* also showed it to be highly induced by propionate, thus supporting the presence of an active propionate metabolic pathway in *A. oryzae* NSAR1.

Similarly, a homologous *coaT* gene was also identified in *A. oryzae* NSAR1 by genomic DNA PCR and RT-PCR. Distinguishingly, the expression of *coaT* remains an active state in *A. oryzae* NSAR1 under the standard cultivation conditions without additional propionate (Fig. 3.5 C and D). This result also appears consistent with

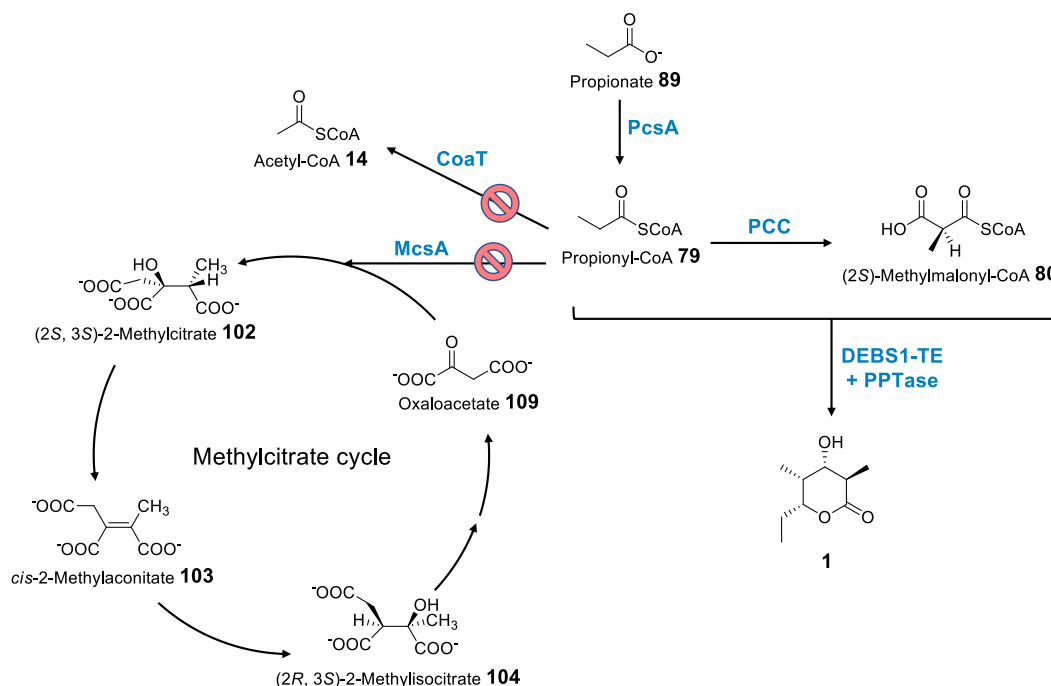
previous studies on *coaT* in *A. nidulans*.<sup>134</sup> In *A. nidulans* wild-type situation, the addition of acetate is able to attenuate the inhibitory effect of propionate. Propionyl-CoA **79** detoxification tends to be first conducted by CoaT without the urgent need of McsA. It may be because the expression of *coaT* is always active while the production of McsA needs to be induced by propionate. This probably makes sense as the CoaT is a general enzyme transferring CoA between various acids.



**Figure 3.5** Gene identification of *mcsA* and *coaT* by PCR: **A**, genomic DNA PCR and RT-PCR analyses of *mcsA* in the absence of propionate; **B**, gDNA PCR and RT-PCR analyses of *mcsA* in the presence of propionate; **C**, genomic DNA PCR and RT-PCR analyses of *coaT* in the absence of propionate; **D**, gDNA PCR and RT-PCR analyses of *coaT* in the presence of propionate.

Combining with the DEBS1-TE heterologous expression pathway, a more complete metabolic pathway can now be proposed (Scheme 3.8). Initially, the production of triketide lactone **1** starts from the feeding of propionate **89** exogenously. Propionate **89** is converted to propionyl-CoA **79** catalyzed by the native propionyl-CoA synthetase PcsA. The producing propionyl-CoA **79** then flows into four different pathways. It can be utilized as the starter unit for polyketide synthesis by DEBS1-TE. It can also be used as the intermediate to produce the extender unit (2*S*)-methylmalonyl-CoA **80**. Additionally, propionyl-CoA **79** can be hydrolyzed back to

propionate through the native CoA transferase CoaT in *A. oryzae* in the presence of acetate. Finally, propionate **89** can induce the expression of methylcitrate synthase McsA leading to degradation of propionyl-CoA **79**. Considering the poor titer of **1**, it is suggested that propionyl-CoA **79** might be significantly degraded *via* the McsA and CoaT metabolic pathways. Therefore, the blockage of these degradation pathways is necessary for the titer improvement of **1**.

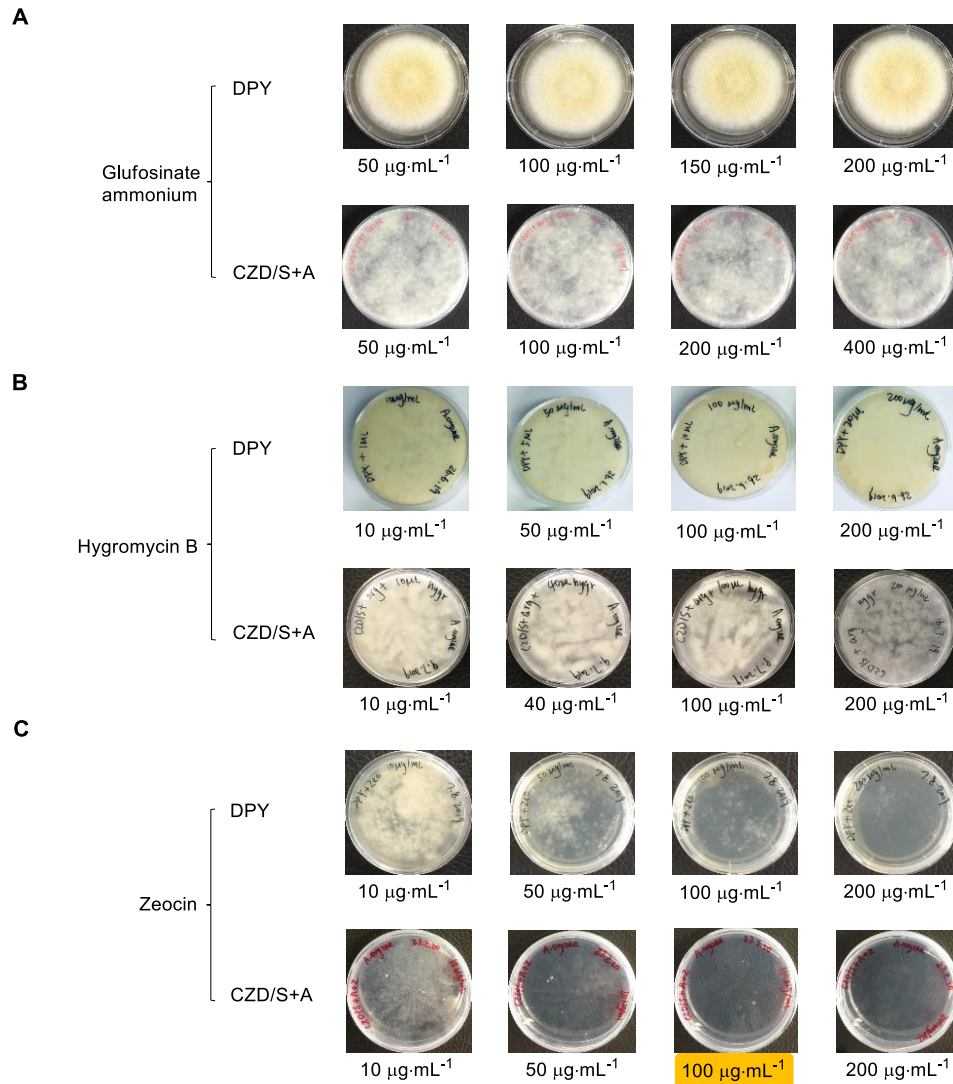


**Scheme 3.8** Metabolic pathways of propionyl-CoA **79** in engineered *A. oryzae* transformant. The red mark represents the degradation pathway of **79**.

### 3.3.4.2 Gene Knockout

Although *A. oryzae* has been extensively used for the heterologous production of natural products, gene knockout studies in *A. oryzae* are relatively limited because of high drug resistance.<sup>136</sup> For example, hygromycin B resistance genes *hph* and *hygr*, and the bleomycin/phleomycin/zeocin resistance gene *ble* have been used as selection markers in *A. oryzae*.<sup>137–139</sup> Therefore, the choice of an effective selection marker is of great significance. It can reduce false positive rates and the workload of screening. Antibiotic resistance markers are the most widely used markers. After fungal transformation with drug resistance genes, the host strain can grow in the presence of a specific antibiotic.

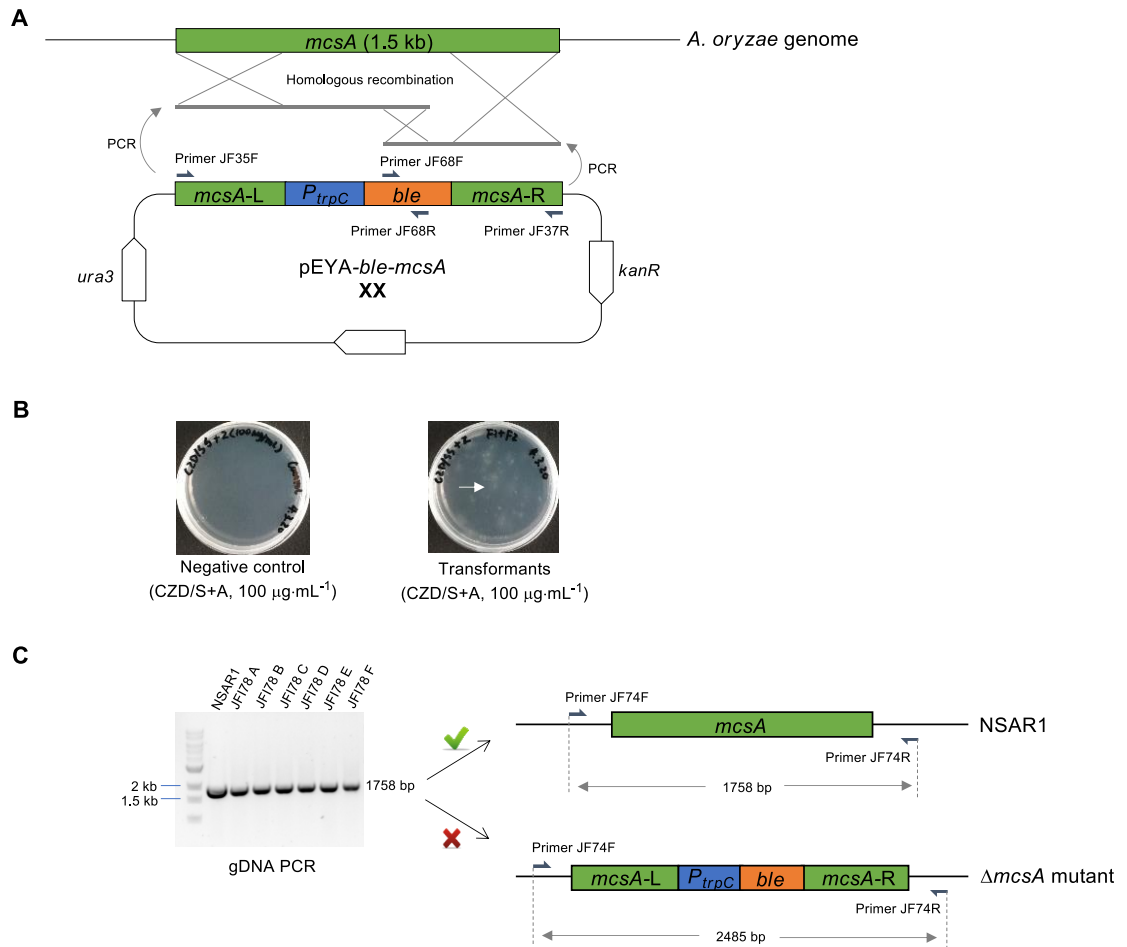
Before gene knockout, a suitable drug which can inhibit the growth of *A. oryzae* NSAR1 should be screened. Here, glufosinate ammonium (basta), hygromycin B, and zeocin were chosen as candidates. Zeocin is an alternative of phleomycin, and glufosinate ammonium had been used as selection marker in *A. nidulans*.<sup>140</sup> Next, the susceptibility assessment of three drugs to *A. oryzae* NSAR1 were conducted.



**Figure 3.6** Sensitivity test of different antibiotics to *A. oryzae* NSAR1 on different media: **A**, the glufosinate ammonium test; **B**, the hygromycin B test; **C**, the zeocin test. The orange highlighted the optimal antibiotic concentration.

Different concentrations of drugs were individually added into the media, including DPY media and CZD/S+A media. The equal volume of spore suspension was spread on the selection plates with supplementary inhibitor. All samples were then incubated at 28 °C for 3 days. The results showed neither glufosinate ammonium

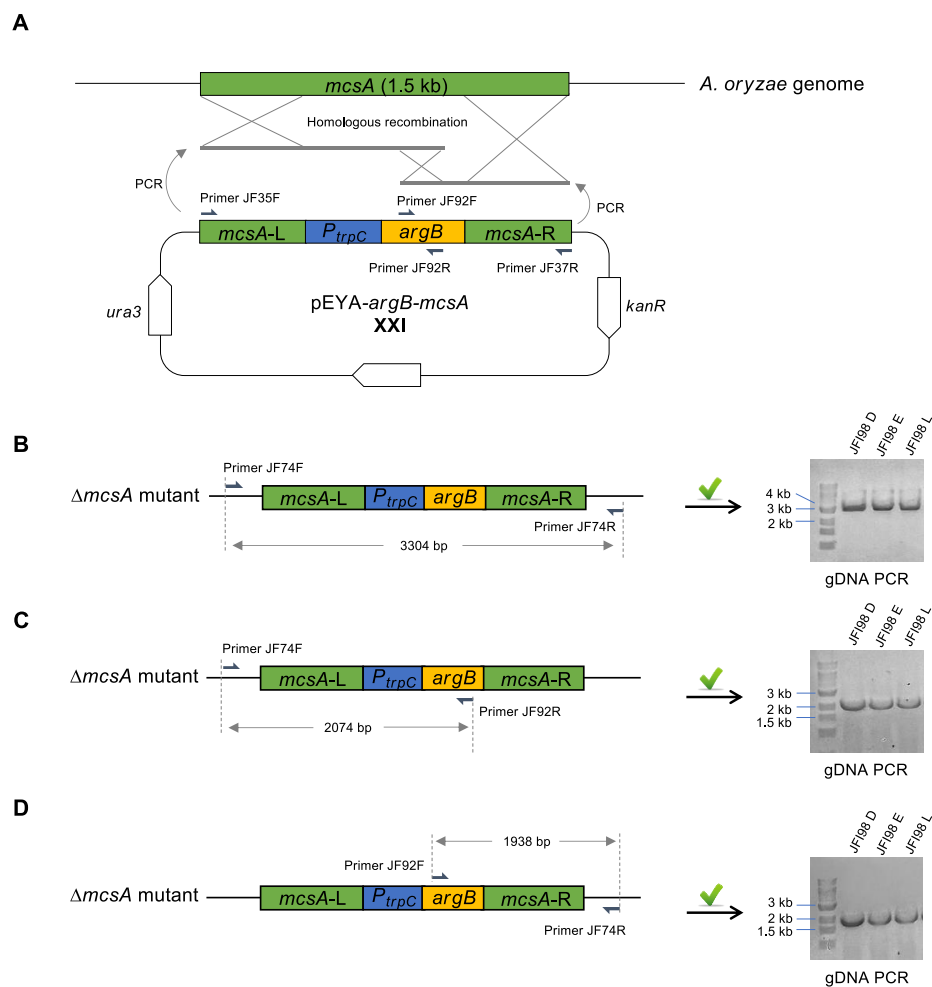
nor hygromycin B could inhibit the growth of *A. oryzae* NSAR1 at concentrations up to 200  $\mu\text{g}\cdot\text{mL}^{-1}$ . Zeocin, at 200  $\mu\text{g}\cdot\text{mL}^{-1}$  on DPY media, exhibited the inhibitory effort on *A. oryzae*, but could not completely inhibit the growth of *A. oryzae*. On CZD/S+A media, zeocin displayed significant inhibitory effect at above 100  $\mu\text{g}\cdot\text{mL}^{-1}$ , leading to the inability of *A. oryzae* to grow. Thus, zeocin resistance gene *ble* was chosen as the selection marker, and the transformant-screening conditions were determined as 100  $\mu\text{g}\cdot\text{mL}^{-1}$  zeocin on CZD/S+A media (Fig. 3.6).



**Figure 3.7** Knockout workflow of the *mcsA* target using the selection marker *ble*: **A**, the construction of the *mcsA* gene knockout cassette and the homologous recombination with *A. oryzae* genome; **B**, *A. oryzae* transformants after fungal transformation; **C**, the selection of the positive *mcsA* mutant by gDNA PCR.

We then attempted to use the bipartite knockout method to insert the *ble* gene into the *mcsA* target. First, a gene knockout cassette consisting of two *mcsA* homologous arms (750 bp each), the promoter *P<sub>trpC</sub>*, and resistance gene *ble* was constructed by yeast recombination.<sup>63</sup> Next, based on the bipartite gene knockout strategy, two separate fragments were amplified by PCR from *mcsA* gene cassette and then

introduced into *A. oryzae* for *mcsA* gene knockout as in previous protocols (Fig. 3.7 A).<sup>141</sup> The empty plasmid without the *mcsA* cassette was used as the negative control. More than 20 *A. oryzae* transformants for each plate were observed on CZD/S+A selection media supplemented with 100  $\mu\text{g}\cdot\text{mL}^{-1}$  zeocin (Fig. 3.7 B). Transformants were then screened, and genomic DNA was extracted individually. All genomic DNA samples were identified by PCR using primers JF74F and JF74R, that are located outside of *mcsA*. As calculated, in *A. oryzae* NSAR1 the PCR product size should be 1758 bp, while in the *mcsA* mutant it should be 2485 bp. Ultimately, it found that in all transformants the complete native *mcsA* gene still remained (transformants JFI78A-F shown in Fig. 3.7 C), demonstrating that the native *mcsA* gene was not disrupted.



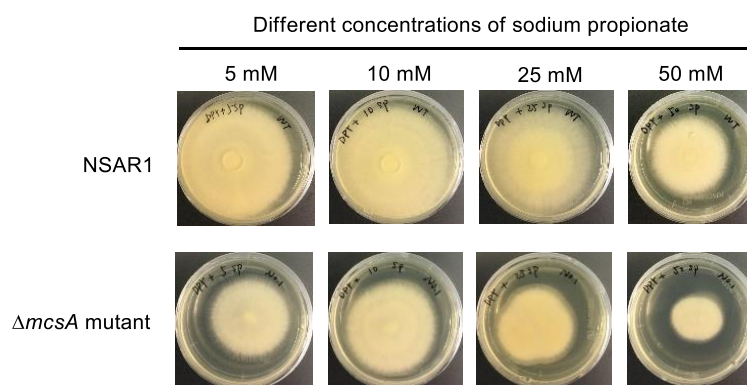
**Figure 3.8** Gene knockout of the *mcsA* target using the selection marker *argB*: **A**, the construction of the *mcsA* gene knockout cassette and the homologous recombination with *A. oryzae* genome; **B**, the identification of the whole gene cassette in the *mcsA* transformants by gDNA PCR using primers JF74F/JF74R; **C**, the identification of the left segment of the *mcsA* gene cassette using primers JF74F/JF92R; **D**, the identification of the right segment of the *mcsA* gene cassette using primers JF92F/JF74R.

To improve the rate of positive transformants, auxotrophic markers were considered instead of drug resistance markers. The auxotroph can restore the same phenotype as the wild type after introduction with the corresponding auxotrophic markers.<sup>136</sup> In filamentous fungi, several auxotrophic markers are often exploited for transformation systems of *A. oryzae*, such as *pyrG* for uridine/uracil biosynthesis, *niaD* for nitrate assimilation, *argB* for arginine biosynthesis, *adeA* for adenine biosynthesis, and *sC* for methionine biosynthesis.<sup>142-146</sup> Given that *A. oryzae* NSAR1 is a quadruple auxotrophic strain (*niaD*<sup>-</sup>, *argB*<sup>-</sup>, *adeA*<sup>-</sup>, and *sC*<sup>-</sup>), the selection marker *argB* was first used for *mcsA* gene knockout.

The new *mcsA* gene knockout cassette with *argB* was constructed, and the bipartite gene knockout-based fungal transformation was carried out as above (Fig. 3.8 A). After transformant screening, 48 transformants were obtained followed by genomic DNA extraction. Next, all genomic DNA samples were identified by PCR using multiple primers. 3 out of 48 transformants were determined as positive *mcsA* mutants (JFI98D, JFI98E and JFI98L). When using primers JF74F and JF74R, an expected PCR band was observed with the size of 3304 bp, that is distinctively different from 1758 bp in the NSAR1 strain (Fig. 3.8 B). When one of primers was located within *argB*, the accurate PCR bands were also obtained in three positive transformants (Fig. 3.8 C and D). Evidently, it demonstrated that *mcsA* gene had been deleted in this mutant strain.

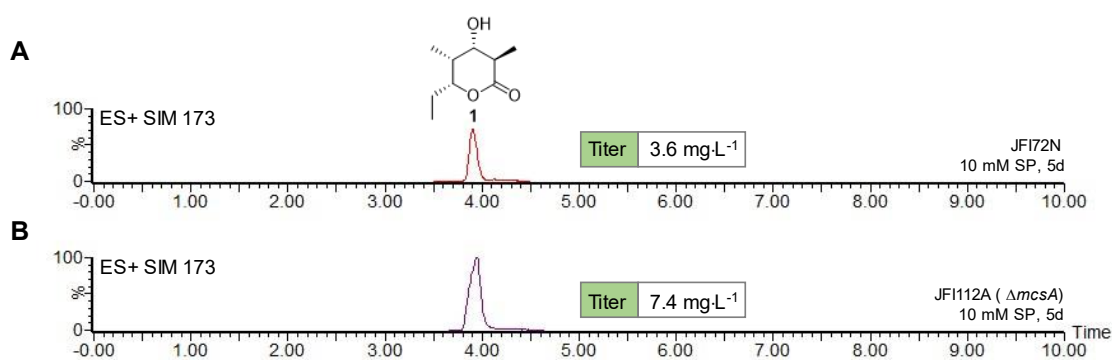
As described above, once *mcsA* is deleted, the phenotype of the corresponding  $\Delta mcsA$  mutant would get weaker than that of the wild-type. In other words, the *A. oryzae*  $\Delta mcsA$  strain would be expected to be more susceptible to propionate compared with the wild-type.<sup>95,99,100,134</sup> Here, the susceptibility of propionate to *A. oryzae*  $\Delta mcsA$  mutant strain was assessed. *A. oryzae* NSAR1 was used as the control. The NSAR1 and  $\Delta mcsA$  mutant were inoculated on DPY media supplemented with different concentrations of sodium propionate. After 5 days of incubation at 28 °C, it was observed that the phenotypes of *mcsA* mutant were weaker than that of the NSAR1 strain (Fig. 3.9), that is in accordance with studies in *A. nidulans*.<sup>95</sup> Based on this result, it also validated that the propionyl-CoA degradation *via* the methylcitrate cycle had been interrupted since *mcsA* was deleted.





**Figure 3.9** Susceptibility test of *A. oryzae*  $\Delta mcsA$  mutant to different concentrations of sodium propionate. *A. oryzae* NSAR1 was used as the control.

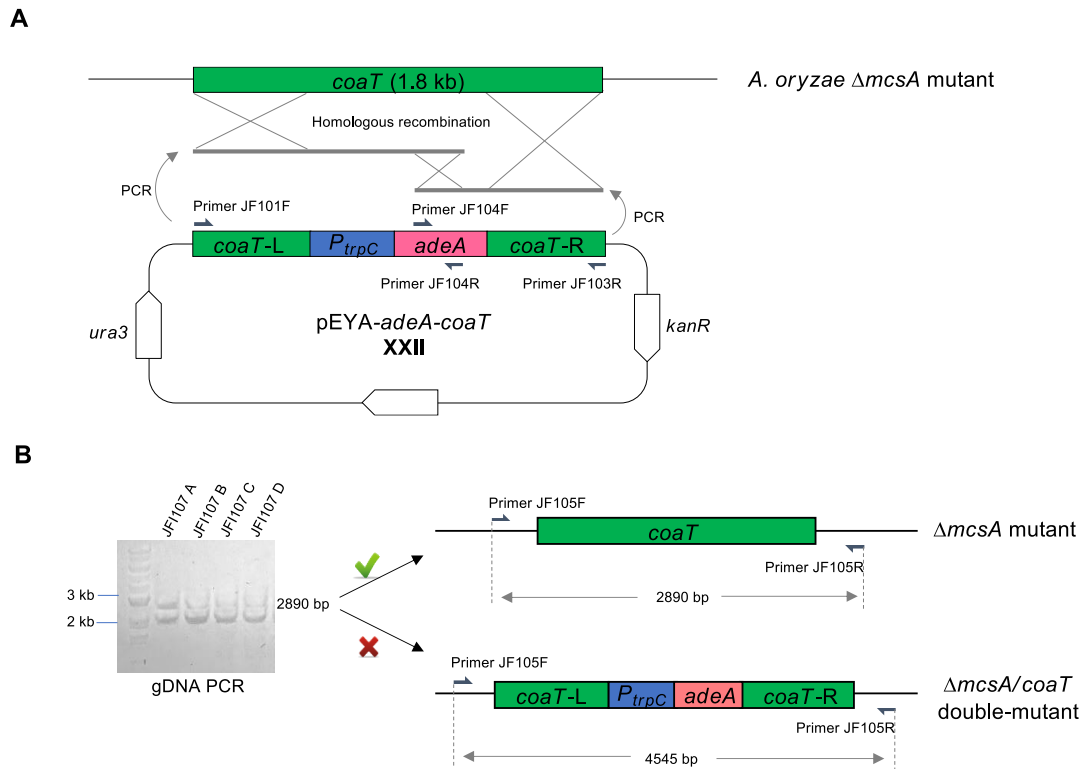
Additionally, the productivity of *A. oryzae*  $\Delta mcsA$  mutant for the production of the triketide lactone **1** was examined. All relevant genes with DEBS1-TE high expression containing DEBS1-TE, *eGFP*, *pccABE*, *sePptII*, and *pcsA* were introduced into *A. oryzae*  $\Delta mcsA$  mutant. Only one fluorescent transformant (JFI112A) was obtained. During the cultivation, 50 mM sodium propionate was added into the culture in 5 days (10 mM per day) before extraction, followed by LCMS analysis. The calculation showed that the titer of **1** was improved significantly to be  $7.4 \text{ mg}\cdot\text{L}^{-1}$  (Fig. 3.10). It indicates that *mcsA* is certainly the degradation gene for propionyl-CoA. However, the other degradation pathway *coaT* is still functional *in vivo*. Thus, in order to improve the titer of **1** further, it requires to delete *coaT* in the *A. oryzae*  $\Delta mcsA$  mutant.



**Figure 3.10** The product **1** titer comparison before and after deleting *mcsA*: **A**, the highest titer of **1** in *A. oryzae* transformant after feeding optimization; **B**, the improved **1** titer in *A. oryzae* transformant after deleting *mcsA*.

For the *coaT* gene knockout, the auxotrophic gene *adeA* was chosen as the selection marker. The *coaT* gene knockout cassette was constructed by yeast recombination and introduced into the *A. oryzae*  $\Delta mcsA$  mutant based on a bipartite gene knockout

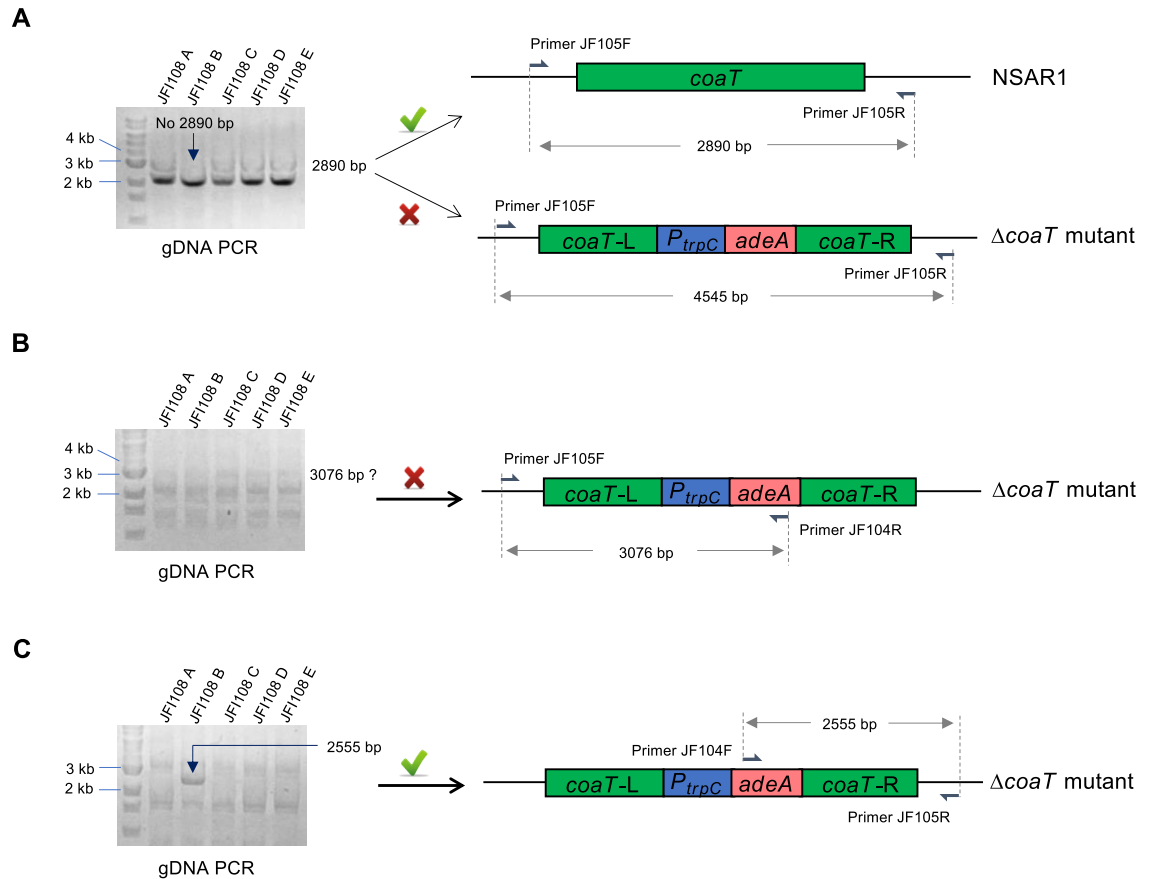
strategy (Fig. 3.11 A).<sup>63,141</sup> Over 100 transformants were obtained and then their genomic DNAs were individually identified by PCR using primers JF105F and JF105R. The results of transformants JFI107A-D were shown (Fig. 3.11 B). The expected PCR band with the size of 4545 bp was not observed in any transformant. Instead, a PCR band at 2890 bp was observed, which is matchable with the expected value from the  $\Delta mcsA$  strain. It indicated that *coaT* gene was not deleted in *A. oryzae*  $\Delta mcsA$  mutant strain yet.



**Figure 3.11** Gene knockout of the *coaT* target using the selection marker *adeA*: **A**, the construction of the *coaT* gene knockout cassette and the homologous recombination with *A. oryzae* genome; **B**, the identification of the whole gene cassette in the *coaT* transformants by gDNA PCR using primers JF105F/JF105R.

The alternative way is to change the order of gene deletion. The *coaT* gene is the first deletion in the NSAR1, followed by *mcsA* deletion in  $\Delta coaT$  mutant. Thereby, two overlapped PCR fragments were amplified from the *coaT* gene knockout cassette and then introduced into *A. oryzae* NSAR1 based on a bipartite gene knockout strategy.<sup>141</sup> After fungal transformation, 31 transformants were obtained and screened by genomic DNA PCR. Lastly, only one  $\Delta coaT$  mutant transformant JFI108B was found, though *coaT* was not completely deleted in *A. oryzae* NSAR1. The gene identification showed differences from other transformants by PCR. Using primers

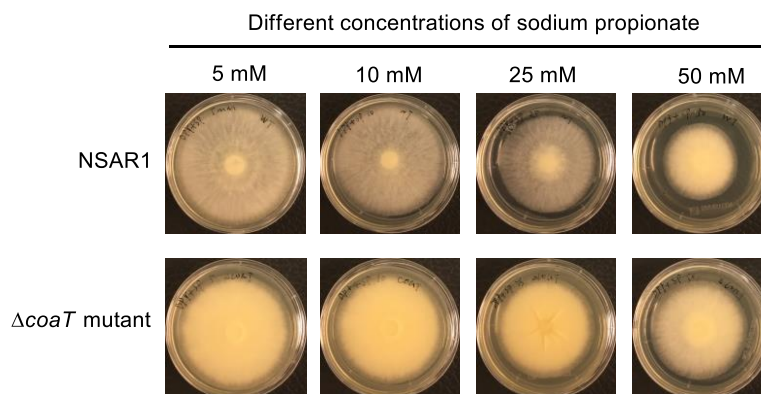
JF105F and JF105R, the native *coaT* gene at 2890 bp was absent compared with other transformants (Fig. 3.12 A). The identification of the first half of *coaT* gene cassette was ambiguous while the second half was determined in the transformant JF1108B (Fig. 3.12 B and C). Therefore, it is speculated that the integrity of *coaT* gene was disrupted. Correspondingly, it might be non-functional *in vivo*.



**Figure 3.12** Genomic DNA PCR analyses of the *A. oryzae coaT* transformants: **A**, the identification of the whole gene cassette in the *coaT* transformants using primers JF105F/JF105R; **B**, the identification of the left segment of the *coaT* gene cassette using primers JF105F/JF104R; **C**, the identification of the right segment of the *coaT* gene cassette using primers JF104F/JF105R.

The *A. oryzae*  $\Delta$ *coaT* mutant strain was also assessed for its susceptibility to propionate. *A. oryzae* NSAR1 was used as the control. The NSAR1 and  $\Delta$ *coaT* mutant strains were inoculated on DPY media supplemented with different concentrations of sodium propionate. After 4 days of incubation at 28 °C, it was found that the phenotype of the  $\Delta$ *coaT* mutant is stronger than those of NSAR1 (Fig. 3.13). It seems to be consistent with the previously published results which show that CoaT is not essential for propionate resistance.<sup>134</sup> The  $\Delta$ *coaT* mutant appears to grow slightly

more strongly in the presence of propionate than the parent strain. Thus, we speculate that the induction of *mcsA* in the presence of propionate must compensate for the lack of CoaT. The expression level of *mcsA* might be relatively higher than in *A. oryzae* NSAR1. For this reason, the inhibitory effect of propionate appears to be attenuated in the *A. oryzae*  $\Delta$ *coaT* mutant.



**Figure 3.13** Susceptibility test of *A. oryzae*  $\Delta$ *coaT* mutant to different concentrations of sodium propionate. *A. oryzae* NSAR1 was used as the control.

Although the *mcsA* deletion was attempted in  $\Delta$ *coaT* mutant many times, the construction of a  $\Delta$ *mcsA/coaT* double-mutant has not been achieved so far. It is still worth to trying in the further.

### 3.4 Discussion and Conclusion

*Aspergillus oryzae* had been engineered for the first time to express a modular PKS DEBS1-TE. However, the titer of the final polyketide **1** was low at 0.6 mg·L<sup>-1</sup>. It limits the further application of *A. oryzae* heterologous expression system. Therefore, the optimization of high-titer production of the modular PKS was attempted.

Optimization of fermentation conditions was the first priority. In the engineered biosynthetic pathway for DEBS1-TE expression, the feeding of propionate is a very crucial and initial step. It supplies the source of both starter unit propionyl-CoA **79** and extender unit (2*S*)-methylmalonyl-CoA **80**. Presumably, the feeding rate of propionate may determine the accumulation of building blocks.

Based on this, the feeding time was elongated up to five days. It led to the 6-fold increase of the overall titer.

Next, the gene expression level was also considered. The fusion gene *eGFP* at the 3' terminal of DEBS1-TE was removed. No significant influence on the product titer was found. Moreover, in the last chapter 2.3.1.2 we know that PcsA is induced in *A. oryzae* by the addition of propionate. Thereby, the native propionyl-CoA synthetase gene *pcsA* in *A. oryzae* was overexpressed. Unexpectedly, the titer of the final product **1** still remained at a similar level as before. We speculate that PcsA may be already at a high enough level. Adding more PcsA *via* the overexpression route makes no additional difference.

We therefore concluded that degradation of propionyl-CoA **79** is most likely to be the reason for low production rates of the DEBS1-TE triketide lactone **1**. Two propionyl-CoA-detoxifying genes *mcsA* and *coaT* were characterized in *A. nidulans*. They can relieve the inhibitory effort of propionate to *A. nidulans*. Using *mcsA* and *coaT* as references, two corresponding homologs were found in the *A. oryzae* genome, and then identified by PCR, and their induction investigated by RT-PCR. The gene *mcsA* is induced by propionate, but *coaT* appears to be constitutive.

To block the consumption of propionyl-CoA **79**, the first degradation gene *mcsA* was deleted using the selection marker *argB*. The *A. oryzae*  $\Delta mcsA$  strain is consistently less healthy than *A. oryzae* NSAR1, and grows significantly less well in the presence of propionate, indicating the importance of McsA for propionate detoxification. Expectedly, *A. oryzae*  $\Delta mcsA$  strain was able to increase the titer of the triketide lactone **1** significantly, suggesting that a major pathway for propionyl-CoA degradation is *via* McsA.

Next, *coaT* was knocked out. The *A. oryzae*  $\Delta coaT$  strains are much healthier than the *A. oryzae*  $\Delta mcsA$  strains. This is probably because McsA is more effective at using propionyl-CoA. In the McsA pathway propionyl-CoA is used to provide the cell with both carbon and NAD(P)H *via* the methylcitrate pathway; but in the CoaT pathway a futile cycle is set up in which ATP is consumed (*via* PcsA) and in which propionate is not effectively used as a carbon source, and no NAD(P)H is generated. It seems that the constitutive CoaT pathway is probably used as a 'first line of defence' and can detoxify low concentrations of propionate; but if propionate

concentrations rise, then the McsA pathway is induced and this can much more effectively remove propionate. Unfortunately, due to time constraints, it was not possible to transform the *A. oryzae*  $\Delta$ *coaT* strains with the triketide lactone **1** producing genes.

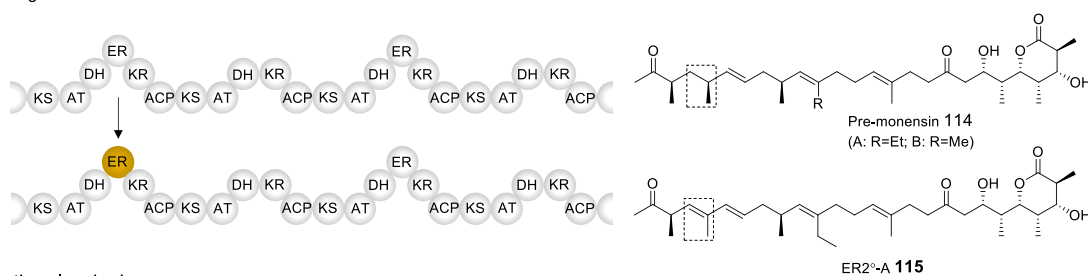
Ideally, a  $\Delta$ *mcsA/coaT* double-mutant is required for DEBS1-TE expression. Unfortunately, such a mutant has not been successfully built so far. Our attempts to create *A. oryzae*  $\Delta$ *mcsA/coaT* double-mutant were not successful, possibly because this leaves the organism without a mechanism to detoxify propionate. In the future, there are still possibilities to improve the titer of modular PKS product. Some native pathways could be manipulated or altered, and heterologous pathways could be introduced as well. For example, the generation of (2*S*)-methylmalonyl-CoA **80** could be achieved from succinate *via* methylmalonyl-CoA mutase, thus avoiding the requirement of high propionyl-CoA **79** concentrations. Possessing a higher-titer productivity, the use of *A. oryzae* will be expanded further in the biosynthesis field of natural products.

## 4. Reprogramming of DEBS1-TE in *A. oryzae*

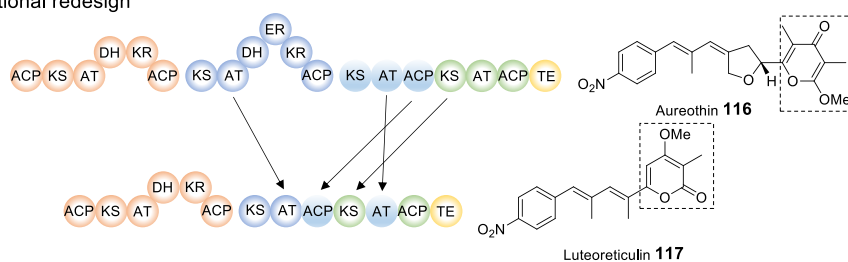
### 4.1 Introduction

To create more structural diversity, reprogramming complex polyketide synthases (PKS) at the genetic level has been considered to be an attractive avenue to generate structurally diverse polyketides since modular PKS were first reported. Many reprogramming strategies for polyketide diversifications have been widely used, such as domain modification, exchanging modules or reengineering subunits.

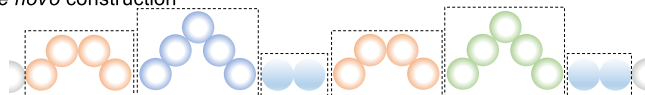
#### A Mutagenesis



#### B Rational redesign



#### C *de novo* construction



**Figure 4.1** Engineering avenues of polyketide synthases: **A**, mutagenesis of individual PKS domain; **B**, rational redesign of polyketide assembly lines; **C**, *de novo* construction of PKS based on knowledges.

The most common engineering strategy is to manipulate individual PKS domains. By specific modification, active sites can be inactivated or changed. For example, changes in KR domains can alter stereochemical activities. Therefore, the mutagenesis of individual PKS domain can lead to formation of various new derivatives. For example, ER<sup>2°</sup>-A **115** is a remarkable product arising from ER domain mutagenesis of the monensin PKS (Fig. 4.1 A). It shows more potent antibacterial activity compared to the parent compound.<sup>147</sup>

The redesign of polyketide assembly lines is another effective avenue. It can not only create a large number of valuable derivatives, but also combine features of multiple pathways or alter one pathway into another, which is analogous to evolutionary process in nature. For example, the aureothin **116** biosynthetic pathway was artificially redesigned through truncating and rearranging domains (Fig. 4.1 B). By rational redesign, the encoded PKS was reprogrammed, resulting in the alteration of molecular backbone of luteoreticulin **117**.<sup>148</sup> Additionally, knowledge-based *de novo* construction is also promising, especially for the *trans*-AT PKS systems (Fig. 4.1 C).<sup>149</sup>

In fact, almost all strategies practically have the similar principle. By recombination, alteration, or exchange of catalytic activities, a new reaction cascade is built to yield desirable compounds.<sup>150</sup> A relatively viable way for reprogramming PKS is to exchange domains as a whole to maintain the genuine multienzyme complexity. Many studies exemplified that it is more favorable to dissect the PKS into functioning domains, that can be repositioned by homologous recombination, ligation or other cloning approaches. In this section, several notable successes of PKS reprogramming from bacteria and fungi are summarized.

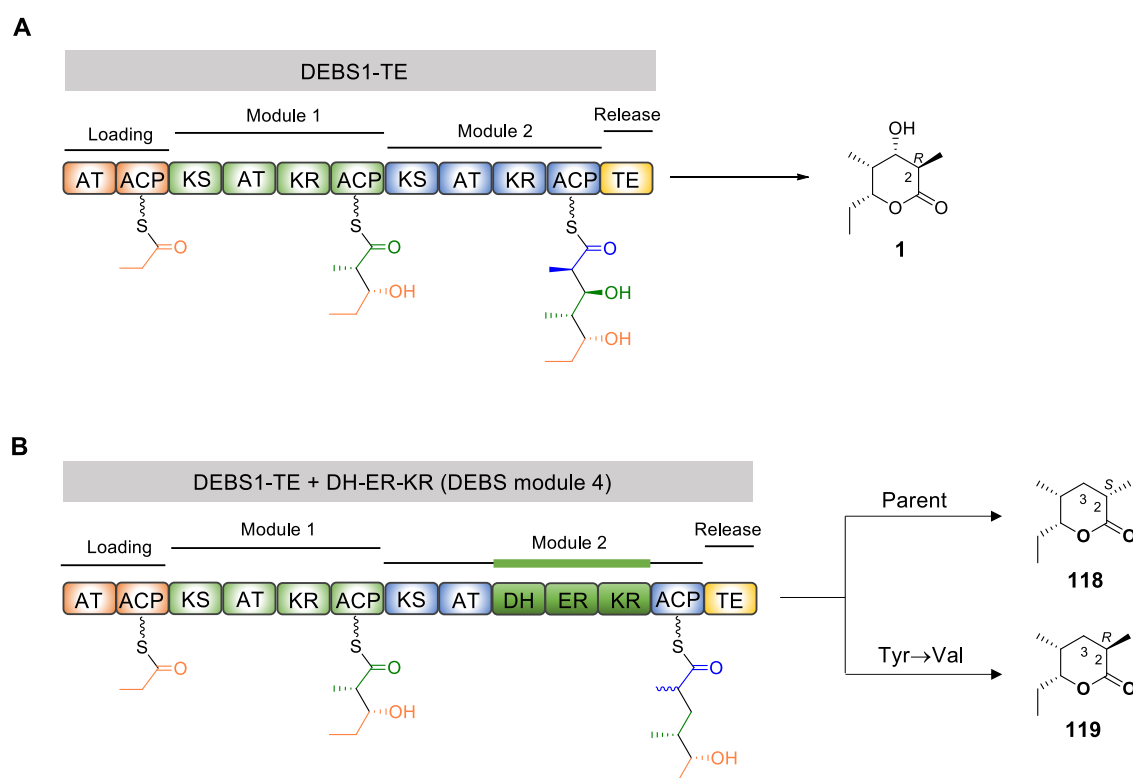
#### 4.1.1 DEBS1-TE Reprogramming in Bacteria

DEBS1-TE is one of the most widely engineered multienzymes. Due to its clear background and mechanism, it has been frequently used for PKS genetic manipulation. DEBS1-TE is a convenient 'test-bed' for the development of new engineering strategies. Multiple reprogramming avenues have been applied for DEBS1-TE to produce more polyketide analogues. Here, several key reprogramming experiments of DEBS1-TE are reviewed in detail.

Domain modification has been extensively used in different ways.<sup>151</sup> For example, the site-directed mutagenesis inactivation of specific domains can disrupt cofactor binding in KR and ER, making the domains non-functional. Another inactivation strategy is to target a specific AT domain within PKS, and to complement the missing activity with a functional alternative AT domain. In addition to that, the specificity of substrates can also be changed by targeting the potential specificity-determining residues within AT domains.



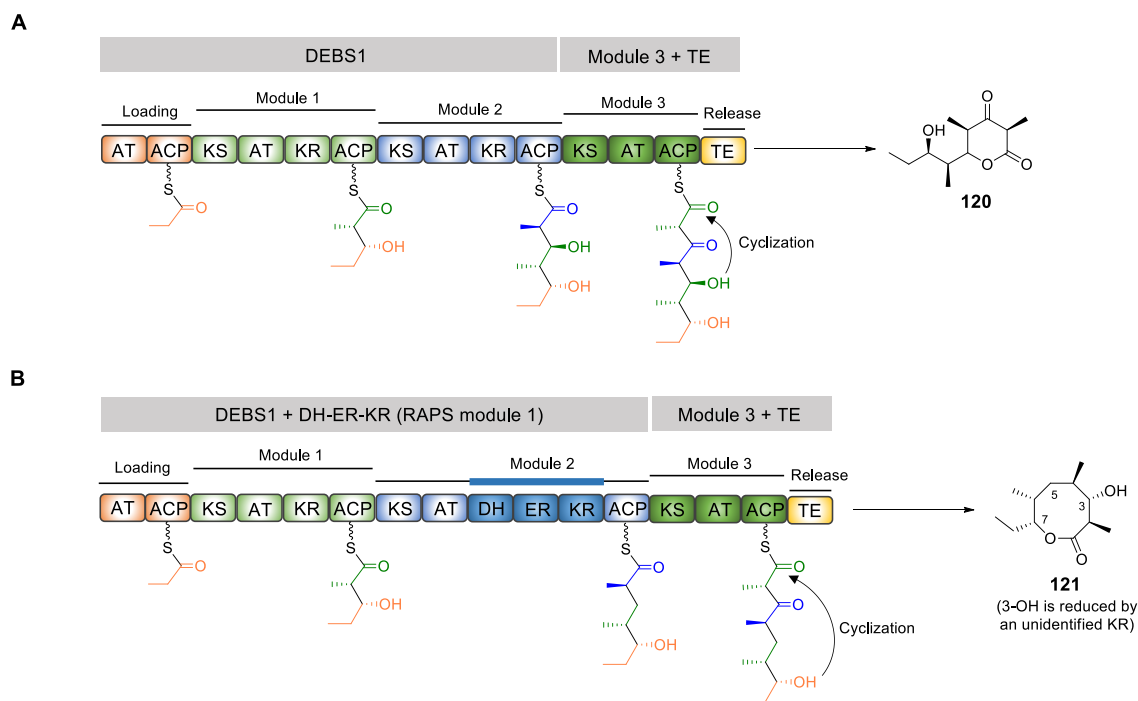
For example, the native DEBS1-TE without any domain modification yields a triketide lactone **1** (20 mg·L<sup>-1</sup>, Scheme 4.1 A). When the KR mono-domain of module 2 was replaced with the DH-ER-KR tridomain from module 4 of DEBS, a different triketide lactone **118** was produced. In the structure of **118**, the hydroxyl group at C-3 is removed. Meanwhile, the stereochemistry of C-2 is altered to 2*S*. In the parental ER, the active site tyrosine (Tyr) is highly conserved based on comparative sequence analysis.<sup>152</sup> It has a close correlation with the direction of enoyl reduction, generating a (2*S*)-methyl configuration. In contrast, the configuration 2*R* often arises from a valine (Val) at the equivalent position. In this PKS, the mutation of the active site Tyr to Val resulted in formation of **119** (1~2 mg·L<sup>-1</sup>) with the opposite configuration at C-2 (Scheme 4.1 B).<sup>153</sup>



**Scheme 4.1** Domain modification to control the stereochemistry of products: **A**, DEBS1-TE and its product triketide lactone **1**; **B**, the introduction (green domains in module 2) resulted in production of **118** with *S* configuration at C-2, while the C-2 configuration in **119** is changed to *R* after ER mutation from Tyr to Val.

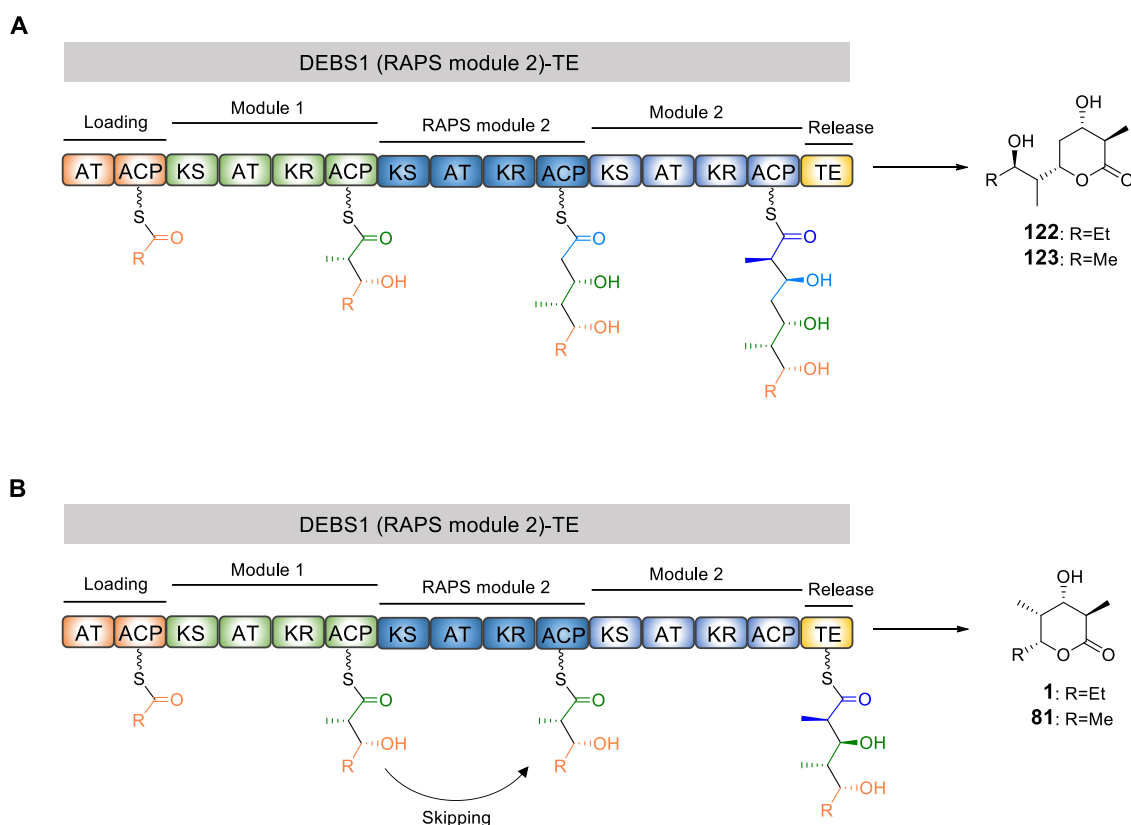
Another example is also individual KR domain exchange for the tridomain DH-ER-KR. The TE domain from the end of DEBS3 was relocated to the terminal of module 3 to form a new tetraketide synthase, leading to production of a corresponding tetraketide lactone **120** (20 mg·L<sup>-1</sup>, Scheme 4.2 A).<sup>154</sup> The hydroxyl group at C-5 as

nucleophile resulted in cyclization to form a 6-membered lactone. After introduction of a DH-ER-KR segment from the rapamycin PKS (RAPS) module 1 to replace KR domain of module 2, the hydroxyl group at C-5 was eliminated because of the reduction of DH and ER domains. Consequently, the terminal hydroxyl was used by TE domain to form an 8-membered lactone **121** (20 mg·L<sup>-1</sup>, Scheme 4.2 B).<sup>153</sup>



**Scheme 4.2** Production of different lactone rings by domain exchange: **A**, a tetraketide synthase consisting of DEBS1, module 3 and TE domain produced a 6-membered tetraketide lactone **120**; **B**, the PKS (DEBS1+module 3+TE) was further modified by introduction a reductive DH-ER-KR segment from RAPS module 1, and an 8-membered lactone ring **121** was formed.

The exchange of a whole module has also been attempted to increase the length of the final polyketide chain by one extender unit. For example, in the middle of module 1 and module 2 of DEBS1-TE, the whole module 2 from RAPS PKS was inserted.<sup>155</sup> The resulting hybrid PKS produced carbon chain-extended polyketides **122** and **123** with a total titer of 1~2 mg·L<sup>-1</sup> as expected (Scheme 4.3 A). However, this construct also gave rise to triketide lactones **1** and **81** with a total titer of 20~40 mg·L<sup>-1</sup>. It appears resulting from skipping of the internal modules (Scheme 4.3 B).



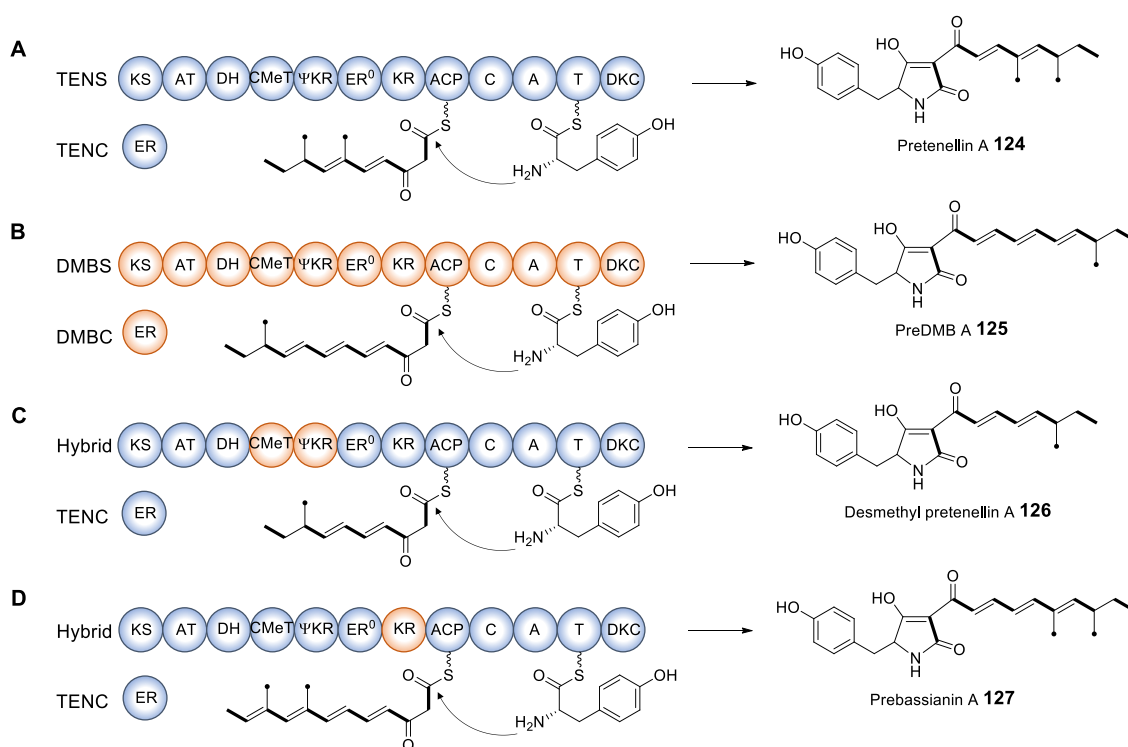
**Scheme 4.3** Addition of the whole module to produce a chain-extended polyketide: **A**, RAPS module 2 was inserted into DEBS1-TE and the predicted products **122** and **123** were produced; **B**, the triketide lactones **1** and **81** were still produced due to direct ACP-to-ACP transfer of the intermediate.

### 4.1.2 PKS Reprogramming in Fungi

Rational reprogramming has also been applied in the field of fungal highly reducing iterative PKSs. Compared with modular PKSs, the programming process of iterative PKSs is much less well understood. Iterative PKSs are composed of a single multidomain module. Despite the responsibility for all extension cycles, each extension cycle might be different from others. Therefore, the complete understanding of programming of iterative PKSs is still a challenge so far.

Notwithstanding, many studies have exemplified the reprogramming of fungal PKSs. For example, in previous studies of our group a rational reprogramming between two fungal PKSs was achieved using domain swapping. Through this way, the programming process of a fungal highly reducing PKS (hr-PKS) was deciphered.<sup>67</sup> In this study, two fungal PKS-NRPSs TENS and DMBS were focused, one for the biosynthesis of pretenellin A **124** (Scheme 4.4 A) and the other one for the biosynthesis of predesmethylbassianin A (preDMB A **125**, Scheme 4.4 B). These

two PKS-NRPSs show very similar location and order of catalytic domains. However, TENS and DMBS differ in programming with respect to both methylation pattern and chain length.



**Scheme 4.4** Reprogramming of PKS-NRPSs by rational domain swapping: **A**, the hr-PKS TENS produced the compound **124**; **B**, the hr-PKS DMBS produced the compound **125**; **C**, the hybrid hr-PKS with the replacement of C-MeT-ΨKR didomain produced the compound **126**; **D**, the hybrid hr-PKS with the replacement of KR domain produced the compound **127**. Abbreviations: KS, ketosynthase; AT, acyl transferase; DH, dehydratase; C-MeT, C-methyltransferase; ΨKR, structural β-ketoreductase; ER<sup>0</sup>, inactive enoyl reductase; ACP, acyl carrier protein; C, condensation; A, adenylation; T, thiolation; DKC, Dieckmann cyclase.

To explore the mechanism of programming of iterative hr-PKSs, the reprogramming strategy of domain swapping was conducted between TENS and DMBS. An elegant series of experiments were attempted. The segments with different size within TENS was replaced by the equivalent segments from DMBS, including KS-AT, KS-AT-DH swap, KS-AT-DH-C-MeT-ΨKR swap, KS-AT-DH-C-MeT-ΨKR-ER<sup>0</sup> swap, KS-AT-DH-C-MeT-ΨKR-ER<sup>0</sup>-KR swap, KS-AT-DH-C-MeT-ΨKR-ER<sup>0</sup>-KR-ACP swap, C-MeT-ΨKR swap, C-MeT-ΨKR-ER<sup>0</sup>-KR swap, and KR swap. Ultimately, it was found that the donor domain from DMBS was introduced into TENS, leading to formation of desmethylpretenellin A **126** (5 mg·L<sup>-1</sup>), which is monomethylated like **125** (Scheme 4.4 C), indicating that the C-MeT-ΨKR didomain strongly controls the methylation

pattern of the PKSs. **126** is the major product, not the sole product, that means *C*-MeT-ΨKR lacks complete fidelity in controlling methylation pattern. Likewise, a KR-only swap resulted in the production of prebassianin A **127** (3 mg·L<sup>-1</sup>), suggesting that the chain length should be controlled by KR domain from the DMBS and the remained demethylation by the TENS *C*-MeT domain (Scheme 4.4 D).

By domain swapping, in the case of TENS and DMBS the differences in programming associated with chain length and methylation pattern were linked to KR and *C*-MeT-ΨKR domains, respectively. More significantly, the reprogramming strategy of iterative PKSs by rational domain swapping also provide insights into the generation of diverse or targeted products by similar synthases.

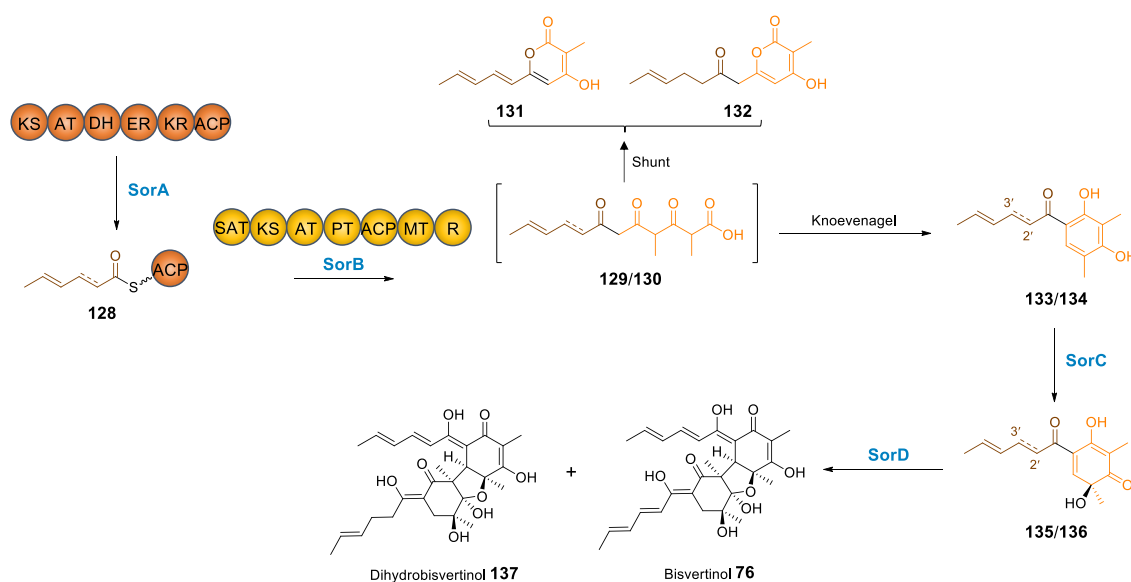
## 4.2 Aims

The reprogramming of DEBS1-TE has been widely investigated and produced a variety of novel derivatives. All reprogramming experiments of DEBS1-TE were conducted with internal modules in DEBS or foreign modules from other modular PKSs in bacteria. Likewise, iterative PKSs has also proven to be able to reprogrammed. It only can be achieved among iterative PKSs in fungi.

Since the modular PKS DEBS1-TE has been expressed in *A. oryzae* for the first time, it is plausible that the reprogrammed DEBS1-TE analogues should be able to be expressed in *A. oryzae* as well. On the other hand, iterative PKSs themselves can be expressed in fungi. Therefore, the fusion expression hypothesis of a modular PKS and an iterative PKS was proposed in this project. It is an interesting attempt to understand if two distinct types of PKS could be combined into one enzyme and produce more diverse derivatives.

Based on this, a suitable iterative PKS which could be combined with DEBS1-TE must be chosen. The selected PKS should be able to generate a triketide product, which is similar to that DEBS1-TE produces. Here, a highly reducing iterative PKS SorA and a non-reducing iterative PKS SorB involved in biosynthesis of sorbicillinoids (*e.g.* bisvertinol **76** and dihydrobisvertinol **137**) were chosen.<sup>66</sup>

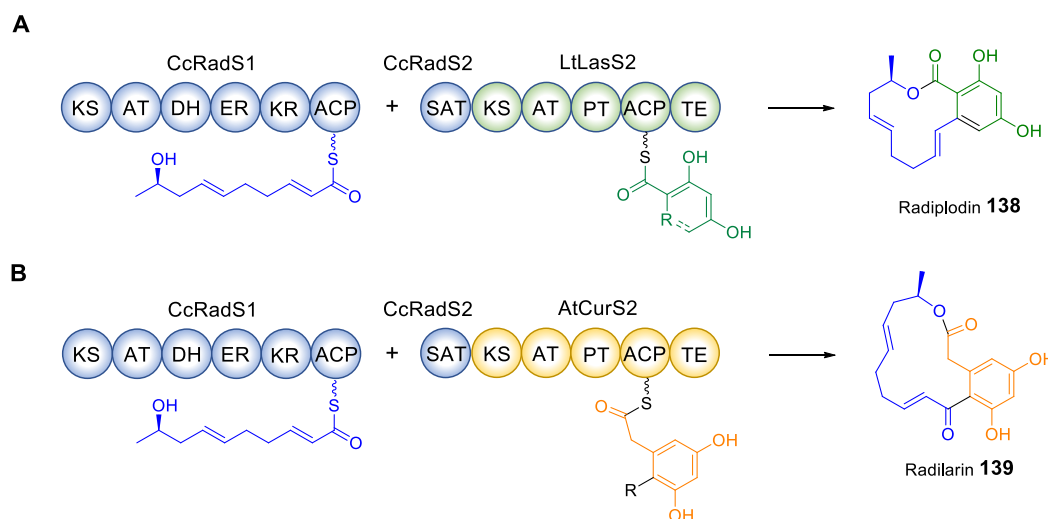
SorA is a hrPKS known to synthesize the triketide intermediate **128** which is kept tethered to the SorA ACP domain. Then, **128** is recognized by the starter acyl transferase (SAT) domain of the SorB nrPKS and then elongated by SorB for three more cycles. During this elongation process, two methyl groups are introduced by the C-methyltransferase (C-MeT) domain to form **129/130**. Then, **129/130** are reductively released, and it is thought that a Knoevenagel condensation then produces sorbicillin **133** and 2', 3'-dihydrosorbicillin **134**. These released polyketides **133/134** are then oxidized by SorC (an FMO) to **135/136**. Lastly, SorD catalyzes a Michael-type dimerization to give sorbicillinoids **76** and **137** (Scheme 4.5). When SorA and SorB were together expressed in the heterologous host *A. oryzae*, two shunt products **131** and **132** were produced, which appear to be off-loaded ahead of being mature.



**Scheme 4.5** Biosynthesis pathway of sorbicillinoids **76** and **137** by SorA, SorB, SorC and SorD. Abbreviations: SAT, starter acyl transferase; PT, product template; MT, C-methyltransferase; R, reductive release domain.

The key feature in this system which is of use in this project is that the SorA ACP-linked triketide **128** is recognized by the SorB SAT domain and transferred to SorB itself, where it can be further modified. In effect SorA makes the triketide *starter unit* for SorB. Others have shown that these types of ACP-SAT systems can be re-engineered to provide different starter units. For example, an ACP-SAT system from radicicol PKS CcRadS1 and CcRadS2 was re-engineered to different nrPKSs (e.g. LaLasS2 and AtCurS2) to provide a different starter unit. The transfer of

heterologous starter unit between noncognate enzyme pairs resulted in the formation of new compounds, such as **138** and **139** (Scheme 4.6), revealing that ACP-SAT domains are poly-specific.

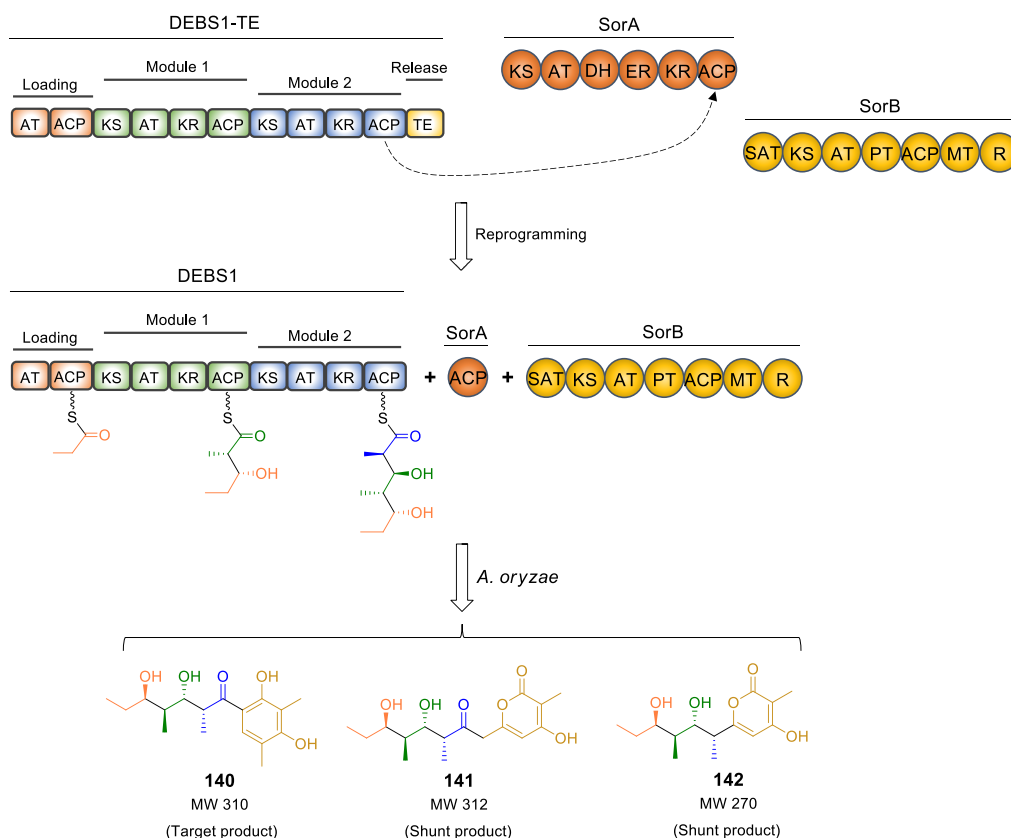


**Scheme 4.6** Fusion expression of re-engineered PKSs using the connection of an ACP-SAT system: **A**, the coexpression of hybrid PKS consisting of CcRadS1, CcRadS2-SAT and LtLasS2; **B**, the coexpression of hybrid PKS consisting of CcRadS1, CcRad2-SAT and AtCurS2.

One way of combining bacterial and fungal PKS systems to make new architectures could be to use the modular bacterial PKS such as DEBS1 to synthesize and provide a *starter unit* for the fungal PKS. The ACP-SAT system could be ideal for this. For this to be possible the bacterial DEBS1-TE would need to be engineered to remove the TE domain (or ACP-TE) and replace it with a fungal ACP-SAT system.

Thus, in this chapter, we aim to build chimeric systems in which the bacterial modular PKS provides a triketide which can then be extended by a fungal iterative nrPKS which might be able to make new benzophenones such as **140** or new pyrones such as **141** and **142** (Scheme 4.7). Three PKS components, including DEBS1, SorA-SAT and SorB, are required for hybrid PKS fusion. There are two different ways of fusion to choose. One is the fusion to be a complete enzyme using DEBS1, SorA-SAT and SorB. That means the intermediate is transferred within the whole PKS. It is more convenient for SAT to recognize the triketide provided by DEBS1. However, the size of the whole enzyme might be difficult for cloning manipulation and have a low recombination rate. Therefore, the other one is to divide the hybrid PKS into two separate enzymes: DEBS1 fused with SorA-SAT as

the first enzyme; SorB as the second enzyme. Both of the above ways have a tandem ACP domain: ACP2 from module 2 of DEBS1 and SorA-ACP. We assumed that the intermediate could be transferred between these ACPs. Another possibility is that the assembly line of PKS might be terminated at the DEBS1-ACP2. Therefore, the PKS fusion after deleting DEBS1-ACP2 would also be attempted. It implies that the triketide provided by DEBS1 could be directly attached to SorA-ACP. Given the above, four different combinations of hybrid PKS would be designed.

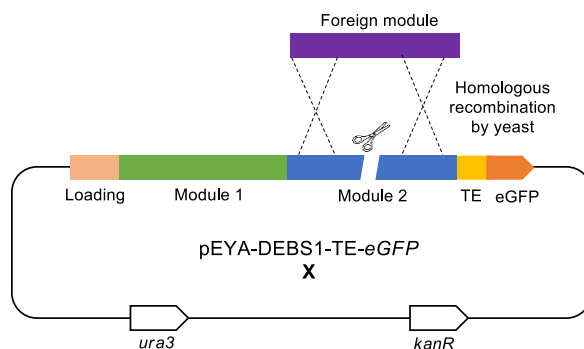


**Scheme 4.7** Possible design of a chimeric PKS consisting of a modular PKS component from DEBS1-TE and iterative PKS components from SorAB.

In addition, the other aim is the DEBS1-TE reprogramming with foreign modular PKSs. However, the traditional reprogramming process, such as a site-directed domain or module swap, was not adopted. Here, a homologous recombination by yeast will be used to replace the module 2 of DEBS1-TE with a foreign modular PKS module (Fig. 4.2). Due to the high DNA (and therefore protein) homology at some sites, it is very likely to automatically select positions for the homologous recombination in yeast.



In turn, these hybrid systems generated by yeast recombination might have less 'damage' at the recombination sites and thereby offer advantages over (human-selected) ligation sites. Furthermore, the homologous recombination process is very likely to be the process by which these systems can evolve. So, here using homologous recombination as the reprogramming tool provides many possible advantages.



**Figure 4.2** Modular PKS DEBS1-TE reprogramming driven by yeast homologous recombination.

## 4.3 Results — Creation of Hybrid Bacterial-Fungal PKS

### 4.3.1 Fusion Design

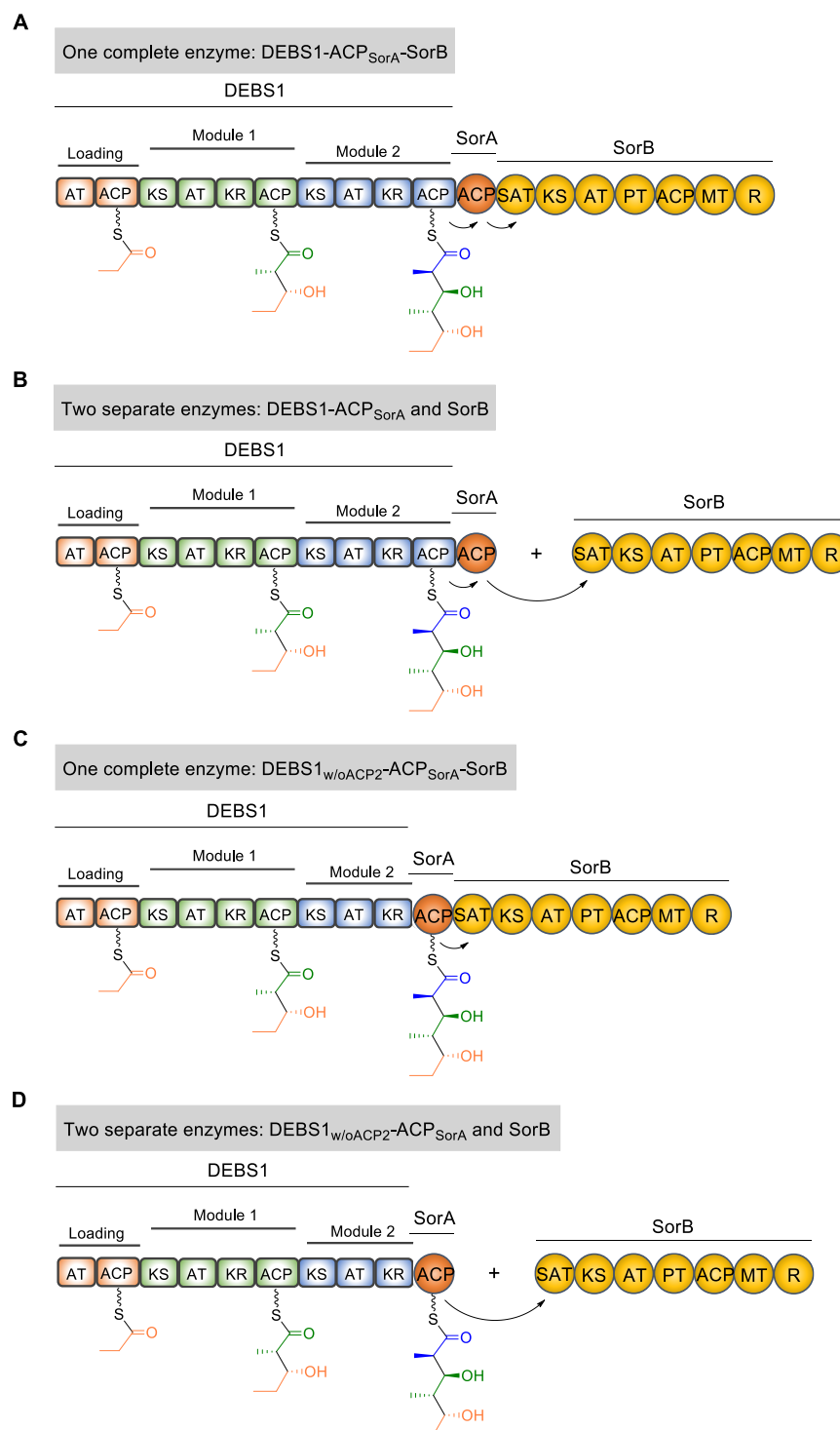
Four different hybrid PKS enzymes were designed *in silico*. We planned to use DEBS1 to create a starter unit for later extension by the fungal PKS SorB. Since DEBS1 uses a TE to off-load its product as a lactone the TE domain of DEBS1-TE should be first removed. We planned to use the SorA-ACP and the SorB-SAT domains to manage the transfer of the DEBS1 triketide to the fungal PKS. Normally a triketide is assembled by SorA on the SorA-ACP and this is recognised by the SorB-SAT and transferred to its active site hydroxyl group before extension by the SorB KS. Therefore, we planned to connect the SorA-ACP to DEBS1.

In the first design, the ACP domain of SorA was fused to the 3' terminal of DEBS1 to form DEBS1-ACP<sub>SorA</sub>. Subsequently, the entire SorB PKS was fused again to the 3' terminal of the ACP<sub>SorA</sub> domain to give rise to a hybrid PKS DEBS1-ACP<sub>SorA</sub>-SorB as a single construct (Fig. 4.3 A). In this protein we hoped that DEBS1 would work normally and assemble a triketide attached to the module 2 ACP. This could be transferred by transthioesterification to ACP<sub>SorA</sub>, and then onwards to SAT (by intramolecular transfer) and then potentially be extended by the SorB PKS. However, the gene length of the newly formed DEBS1-ACP<sub>SorA</sub>-SorB is about 19 kb, which is inconvenient for cloning manipulations and even affects the efficiency of homologous recombination.

Therefore, in the second design (Fig. 4.3 B), a stop codon TGA was added at the 3' terminal of the ACP<sub>SorA</sub>. The resulting ACP<sub>SorA</sub> was fused to the 3' terminal of DEBS1 to form the first enzyme DEBS1-ACP<sub>SorA</sub>. In this design SorB will act as a separate protein and hopefully collect the triketide starter unit from ACP<sub>SorA</sub> in the usual intermolecular way.

During the assembly line of polyketide in the two designs above, we assumed that the triketide intermediate attached to ACP domain of module 2 of DEBS1 can be transferred to the ACP domain of SorA spontaneously like the skipping mechanism shown in Scheme 4.3,<sup>155</sup> and then be processed by the downstream iterative fungal biosynthetic pathway. However, failure of this ACP-ACP transfer would prevent complete biosynthesis. Although it is easy to clone the fungal SorA-ACP downstream

of the DEBS module 2 ACP, it would be risky to rely on spontaneous intramolecular transthioesterification. Consequently, two further hybrid PKSs were designed as alternatives, in which the DEBS module 2 ACP is deleted and replaced by the fungal SorA ACP.



**Figure 4.3** Design of hybrid PKSs consisting of modular and iterative PKS components: **A**, the complete hybrid PKS consisting of DEBS1, ACP domain of SorA, and SorB; **B**, the separate hybrid PKS consisting of DEBS1-ACP<sub>SorA</sub> as the first enzyme and SorB as the second one; **C**, the complete hybrid PKS consisting of DEBS1 without ACP2, ACP<sub>SorA</sub> and SorB; **D**, the separate hybrid PKS consisting of DEBS1<sub>w/oACP2</sub>-ACP<sub>SorA</sub> as the first enzyme and SorB as the second one.

Thus, in the third design of hybrid PKS the DEBS1 ACP-TE didomain is deleted and replaced with the SorA-ACP and SorB is fused directly downstream of this. So, this new hybrid PKS was named as DEBS1<sub>w/oACP2</sub>-ACP<sub>SorA</sub>-SorB (Fig. 4.3 C). Finally, in the fourth design, DEBS1 module 2 ACP-TE is replaced by the SorA-ACP, followed by a stop codon, and SorB is expressed as a separate protein (Fig. 4.3 D). Again, this construct will rely on the natural transfer of an acyl group from the SorA ACP to the SorB SAT.

### 4.3.2 ACP Boundary Confirmation

The ACP domain of SorA is the key protein upon which all of the above designs rest. The SorA ACP will be located at the pivotal position, linking to the DEBS1 at the 5' terminal and the SorB at the 3' terminal (Fig. 4.3). Therefore, first of all, the boundaries of the SorA ACP domain should be determined. The protein sequence (from P2506 to S2624) containing the ACP domain of SorA was submitted to the Phyre2.0 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) for sequence alignment and structural analysis.<sup>156</sup> The result indicated that the query sequence shows 26 % structural identity with the single highest scoring template c6h0jA. A total of 94 residues (79 % of the sequence) were modelled with 99.8 % confidence by c6h0jA. In addition to that, the expected DSL motif is highly conserved in all homologs (Fig. 4.4 A). Meanwhile, a PDB formatted 3D model of the query protein was also obtained and visualized by EzMol (<http://www.sbg.bio.ic.ac.uk/ezmol/>). It clearly displayed the SorA ACP boundary from R2536 to R2624 (Fig. 4.4 B).

Furthermore, the protein sequence alignment was conducted between the SorA ACP and the characterized ACP domain from LovB, which is functional during the biosynthesis of lovastatin nonaketide.<sup>157</sup> The results indicated that both ACP domains share similar boundaries (Fig. 4.4 C). To ensure the effectiveness of the ACP domain, the linker sequence from I2530 to S2535 was also included (Fig. 4.4 B). Therefore, the ACP boundary of SorA was finally determined as from I2530 to S2624 in SorA and then attached to the 3' terminal of DEBS1.



**Figure 4.4** Boundary confirmation of the ACP domain of SorA: **A**, the protein alignment of SorA with homologs using Phyre2.0; **B**, the display of the protein 3D model structure by EzMol (the KR domain (blue), the linker (green), the ACP domain (yellow), the conserved region (red), and the tail sequence (grey)); **C**, the protein alignment with the characterized LovB ACP domain.

#### 4.3.1.2 Plasmid Assembly

The assembly of all hybrid PKSs was carried out by yeast recombination and *in vitro* LR recombination.<sup>62,63</sup> First of all, according to the principle of homologous recombination, each gene should be flanked with two homologous overlaps (30~50 bp). Based on this, a series of specific primers were designed. Then, all DEBS1-related gene fragments (Fig. 4.5) were amplified by PCR using the previously constructed plasmids pTYGS-arg-DEBS1-TE-*eGFP-pccABE XVIII* (in Chapter 2.3.3.3) and pTYGS-met-*sePptII XVII* (in Chapter 2.3.5) as templates. The *sorAB*-related gene fragments were replicated by PCR using *sorAB* vectors provided by Lukas Kahlert as templates.

The construction route is shown in Scheme 4.8. The assembly of the first design of hybrid PKS started from the vector pE-YA **VII**. The DEBS1 and ACP<sub>SorA</sub> fragments were inserted into pE-YA **VII** by yeast recombination to form pEYA-DEBS1-ACP<sub>SorA</sub> **XXIII**. A unique restriction enzymatic site *HindIII* was introduced at the 3' terminal of ACP<sub>SorA</sub>. By *HindIII* digestion and further yeast recombination, *sorB* was fused to the 3' terminal of ACP<sub>SorA</sub> to give pEYA-DEBS1-ACP<sub>SorA</sub>-*sorB* **XXIV**.

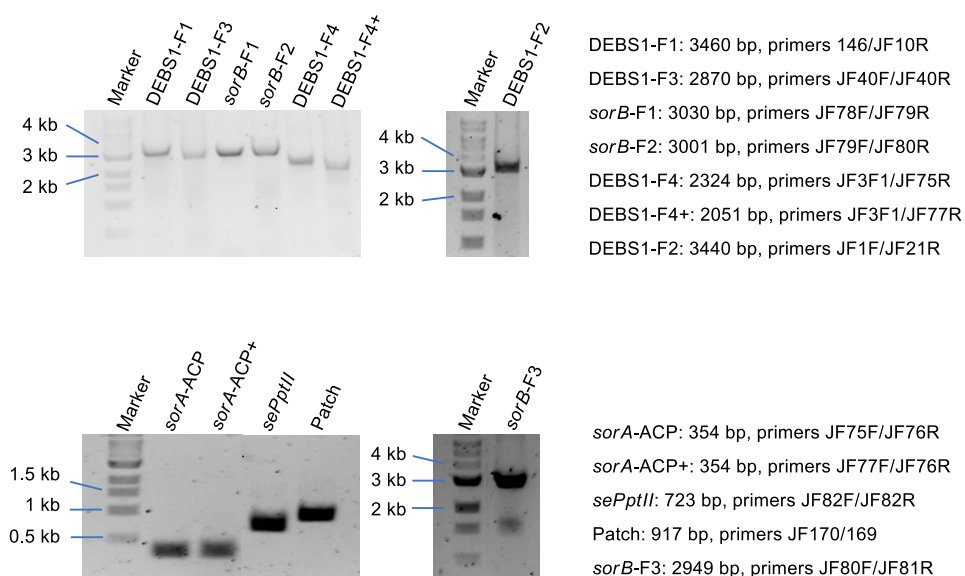
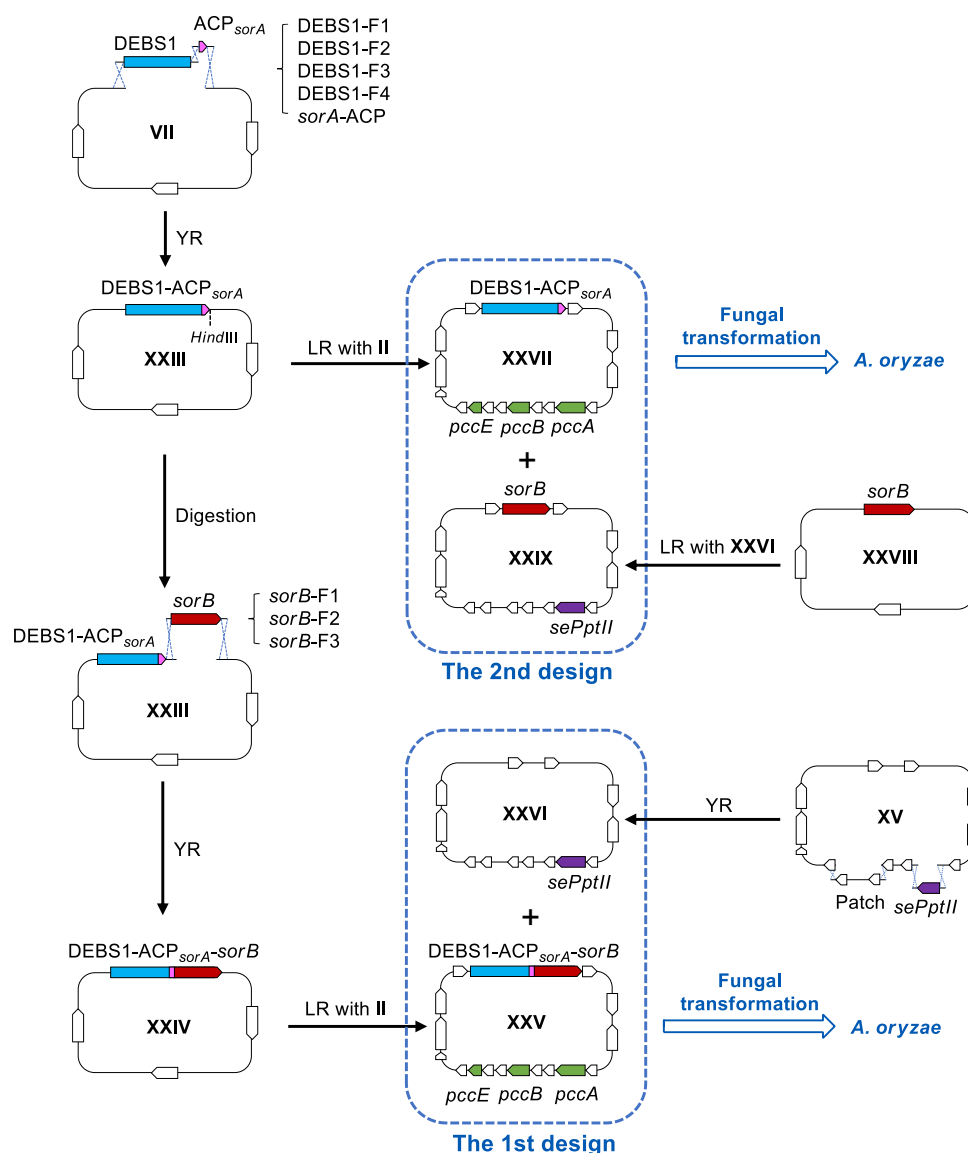


Figure 4.5 PCR fragments used for plasmid assembly.

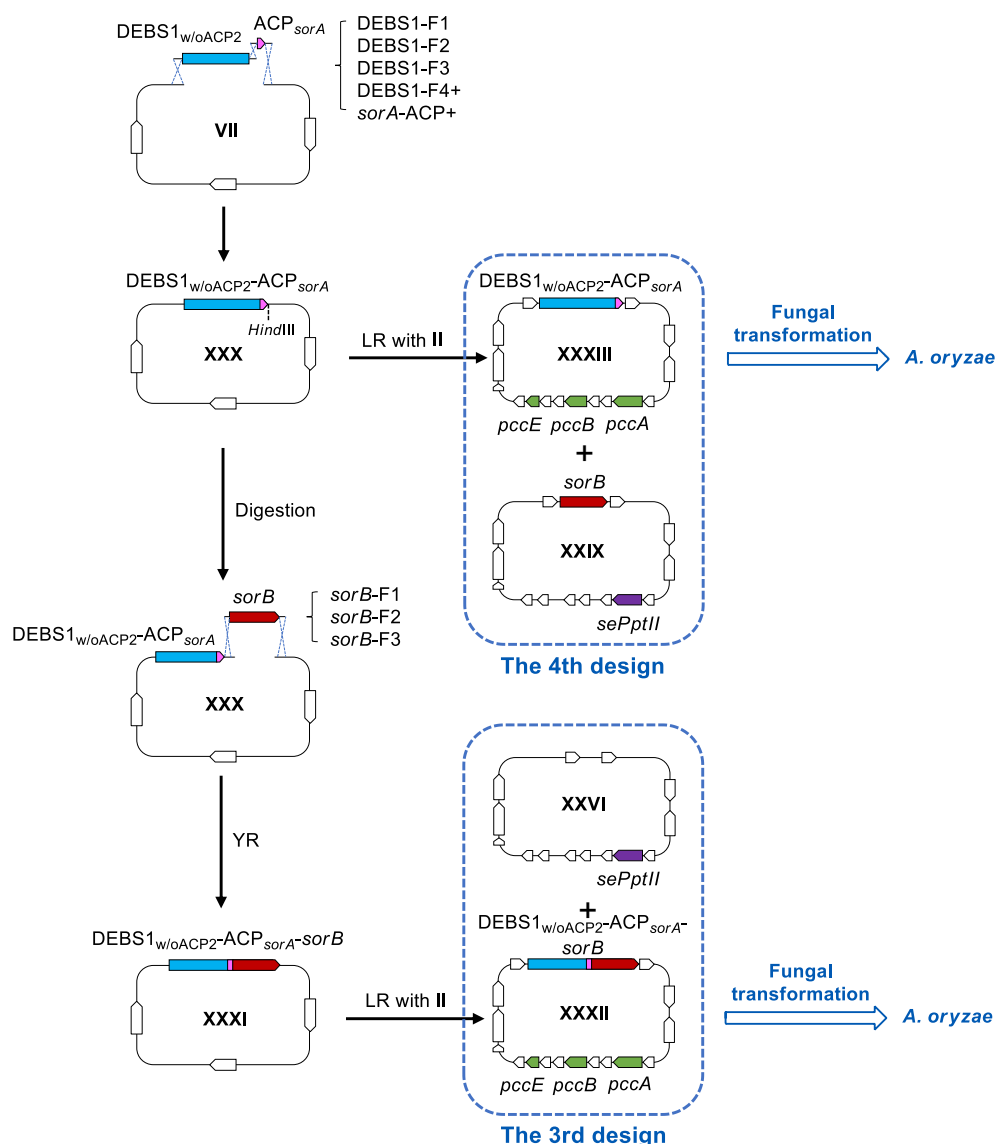
Then, the DEBS1-ACP<sub>sorA-sorB</sub> on the plasmid **XXIV** was transferred to the pTYGS-arg-*pccABE II* containing the carboxylase gene *pccABE* by LR recombination. The expression plasmid pTYGS-arg-DEBS1-ACP<sub>sorA-sorB-pccABE</sub> **XXV** was built. On the other side, the plasmid pTYGS-met-*sePptII*\* **XXVI** bearing the PPTase gene *sePptII* was prepared from the plasmid pTYGS-met **XV** by yeast recombination. Plasmids **XXV** and **XXVI** would be cotransformed into *A. oryzae* NSAR1.

For the first hybrid PKS design, the ACP<sub>sorA</sub> domain was fused together with the whole *sorB* segment on one plasmid. To make the cloning work more simply and ensure the recombination rate, a separate plasmid construction of DEBS1-ACP<sub>sorA</sub> and *sorB* was produced as the second design (Scheme 4.8). The gene DEBS1-ACP<sub>sorA</sub> on the plasmid **XXIII** was directly transferred by LR recombination, and then the plasmid pTYGS-arg-DEBS1-ACP<sub>sorA</sub> **XXVII** was generated. The *sorB* gene was transferred into the plasmid **XXVI** through LR recombination with pEYA-*sorB* **XXVIII** (provided by Lukas Kahlert), resulting in the formation of the coexpression plasmid pTYGS-met-*sorB-sePptII* **XXIX**. Similarly, plasmids **XXVII** and **XXIX** would be cotransformed into *A. oryzae* NSAR1 for heterologous expression.



**Scheme 4.8** Construction route of plasmids containing hybrid PKSs. Abbreviations: YR, yeast recombination; LR, Gateway LR recombination.

Next, the plasmid assembly of the third fusion design was performed (Scheme 4.9). The DEBS1 (lacking ACP2) fragment was first prepared and then incorporated along with ACP<sub>sorA</sub> into the plasmid pE-YA VII to form the pEYA-DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub> XXX. By the successive *Hind*III digestion and yeast recombination, the *sorB* fragment was fused to the 3' terminal of ACP<sub>sorA</sub> to form a whole gene DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-*sorB* on the plasmid XXXI. Subsequently, the whole fused gene was transferred into the expression vector II containing the carboxylase gene *pccABE* by LR recombination. The final expression plasmid pTYGS-arg-DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-*sorB*-*pccABE* XXXII was generated. The PPTase gene-containing pTYGS-met-*sePptII*\* XXVI was used as the coexpression plasmid for *A. oryzae* transformation.



**Scheme 4.9** Construction route of plasmids containing hybrid PKSs (without ACP2 of DEBS1). Abbreviations: YR, yeast recombination; LR, Gateway LR recombination.

Likewise, a separate construction of DBES1<sub>w/oACP2</sub>-ACP<sub>sorA</sub> and *sorB* was also conducted. The fused gene DBES1<sub>w/oACP2</sub>-ACP<sub>sorA</sub> was directly transferred from **XXX** into the expression vector **II** harboring the carboxylase gene *pccABE* by LR recombination. The plasmid pTYGS-arg-DBES1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-*pccABE* **XXXIII** was formed. The other PPTase gene-containing plasmid **XXIX** was used as the coexpression plasmid. Plasmids **XXXIII** and **XXIX** would be cotransformed into *A. oryzae* NSAR1 for heterologous expression.



All constructed plasmids were verified as positive by PCR identification and restriction enzyme digestion (data not shown). All fusion sites on plasmids were confirmed by gene sequencing.

### 4.3.2 Heterologous Expression in *A. oryzae*

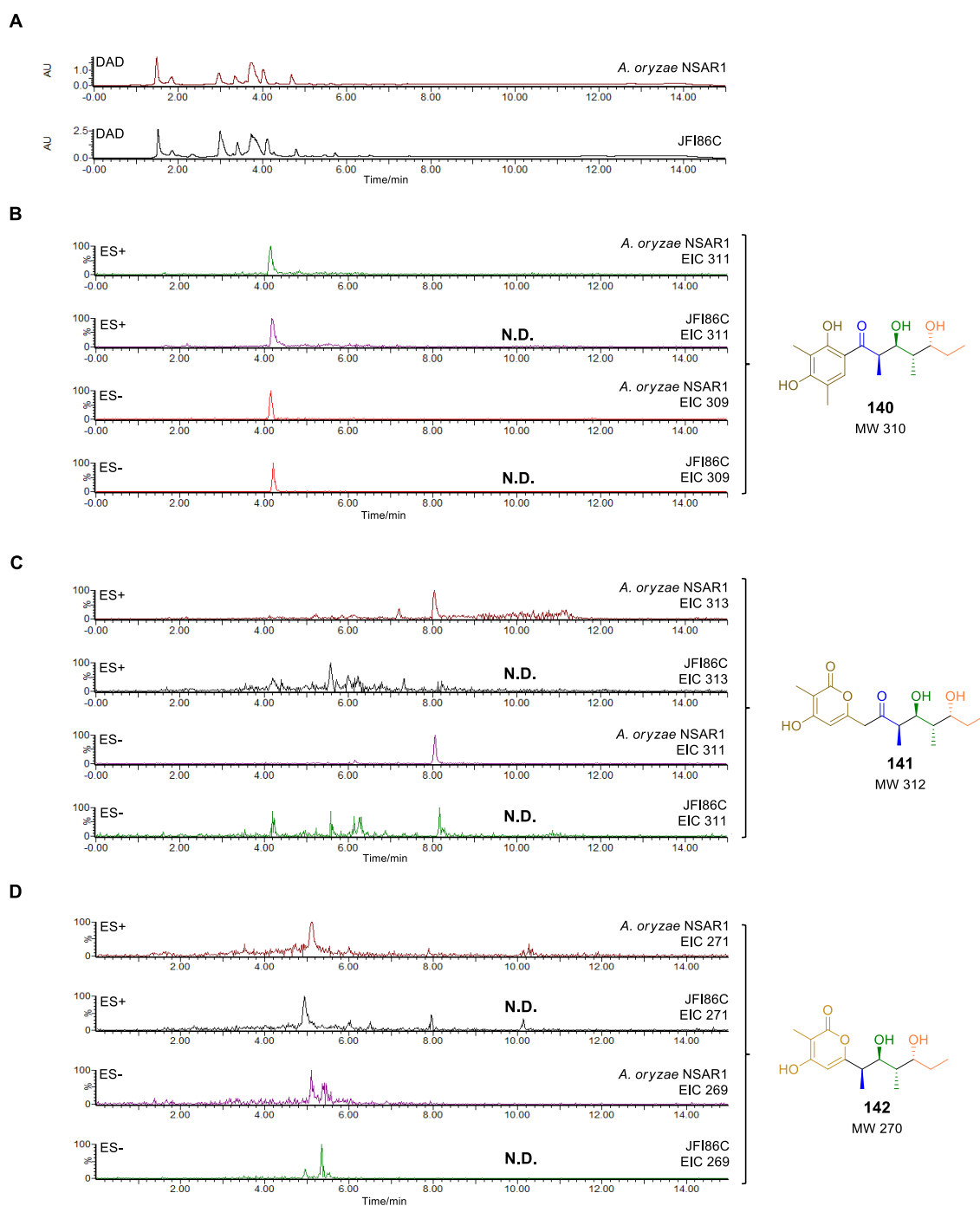
#### 4.3.2.1 LCMS Analyses of Transformants

Four different groups of plasmids were individually introduced into *A. oryzae* NSAR1 by fungal transformation,<sup>62</sup> including pTYGS-arg-DEBS1-ACP<sub>sorA</sub>-sorB-pccABE **XXV** & pTYGS-met-sePptII\* **XXVI**, pTYGS-arg-DEBS1-ACP<sub>sorA</sub> **XXVII** & pTYGS-met-sorB-sePptII **XXIX**, pTYGS-arg-DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-sorB-pccABE **XXXII** & pTYGS-met-sePptII\* **XXVI**, and pTYGS-arg-DBES1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-pccABE **XXXIII** & pTYGS-met-sorB-sePptII **XXIX** (Table 4.1).

**Table 4.1** Different gene combinations for hybrid PKS expression.

Group	Expressed plasmid pTYGS-arg			Expressed plasmid pTYGS-met			The number of transformants	New product
	No.	Gene 1	Gene 2	No.	Gene 3	Gene 4		
JFI85	<b>XXV</b>	DEBS1-ACP <sub>sorA</sub> -sorB	pccABE	<b>XXVI</b>	-	sePptII	18 (A-R)	no
JFI86	<b>XXVII</b>	DEBS1-ACP <sub>sorA</sub>	pccABE	<b>XXIX</b>	sorB	sePptII	18 (A-R)	no
JFI88	<b>XXXII</b>	DEBS1 <sub>w/oACP2</sub> -ACP <sub>sorA</sub> -sorB	pccABE	<b>XXVI</b>	-	sePptII	7 (A-G)	no
JFI89	<b>XXXIII</b>	DEBS1 <sub>w/oACP2</sub> -ACP <sub>sorA</sub>	pccABE	<b>XXIX</b>	sorB	sePptII	16 (A-P)	no

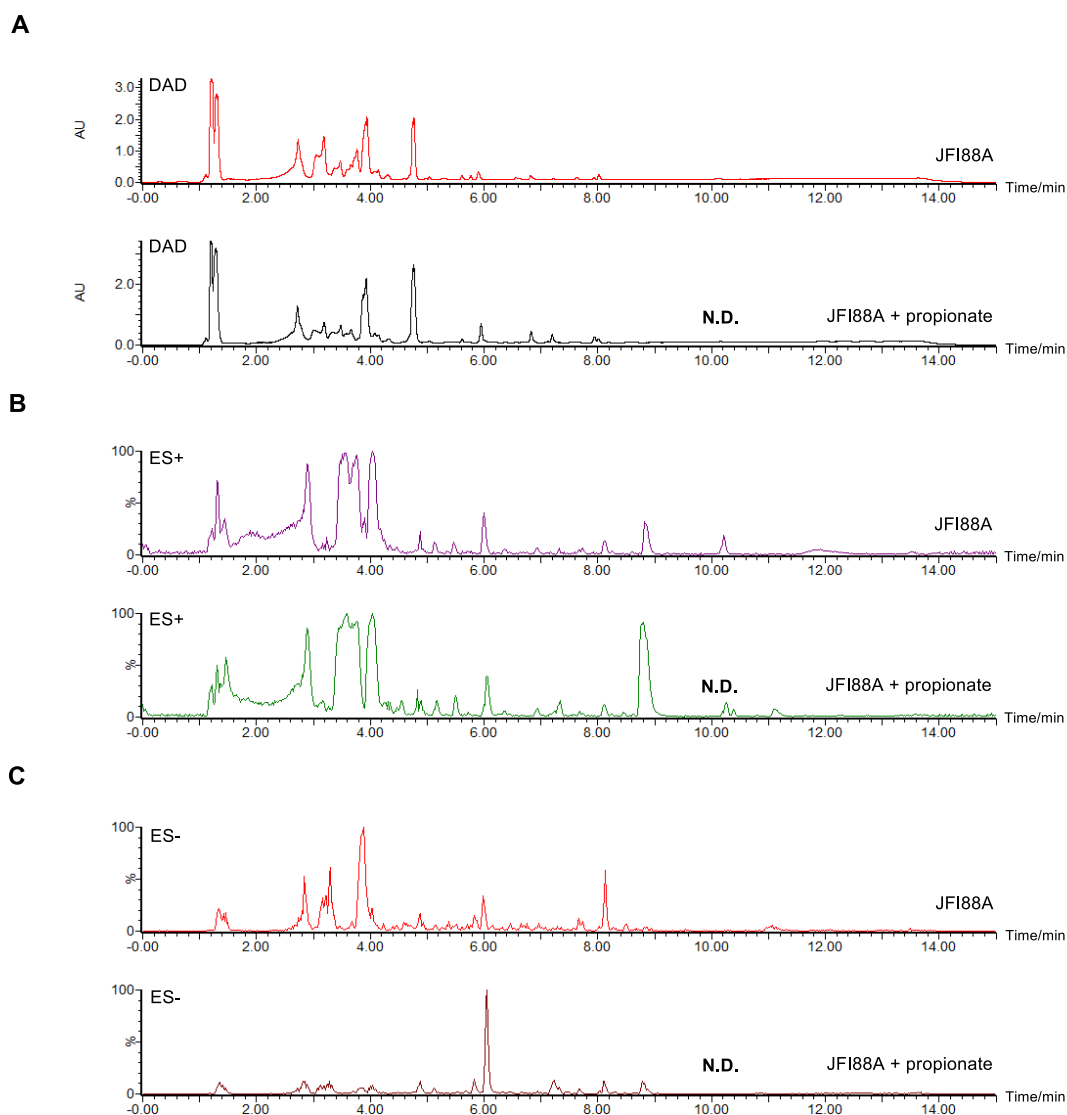
Four groups of transformants (*i.g.* JFI85, JFI86, JFI88 and JFI89) were obtained and screened on the selection media CZD/S without methionine. Then transformants were individually inoculated and cultivated in DPY liquid media for 5 days. At the day before extraction, 50 mM propionate **89** was added into the culture. After chemical extraction, the crude extracts of all transformants were analyzed by LCMS.



**Figure 4.6** LCMS analysis of *A. oryzae* transformants bearing hybrid PKSs: **A**, the comparison of DAD chromatograms between *A. oryzae* NSAR1 and transformant JFI86C; **B**, the EIC chromatograms searching for the expected product **140**; **C**, the EIC chromatograms searching for the expected product **141**; **D**, the EIC chromatograms searching for the expected product **142**. Abbreviation: N.D., not detected.

Taking the coexpression of pTYGS-arg-DEBS1-ACP<sub>SorA</sub>-pccABE **XXVII** and pTYGS-met-sorB-sePptII **XXIX** as an example, 18 transformants (JFI86A-R) were obtained. But there was no new peak observed on the DAD chromatogram compared with that of *A. oryzae* NSAR1 (Fig. 4.6 A). According to the heterologous production result of

sorbicillins by SorB in *A. oryzae*, **140** was predicted to produce as the major product and the related pyrones **141** and **142** as the minor products. The molecular ion peaks of **140**, **141** and **142** were searched on LCMS, respectively. The targeted peak was not found in the EIC ES+ nor ES- chromatograms (Fig. 4.6 B, C and D). For other groups of hybrid PKS coexpression, no new peak was found either (data not shown).



**Figure 4.7** Comparative LCMS analysis of the transformant JFI88A bearing hybrid PKSs: **A**, the DAD comparison of the transformant JFI88A with and without feeding propionate **89**; **B**, the ES+ comparison; **C**, the ES- comparison.

Abbreviation: N.D., not detected.

The predicted products **140**, **141** and **142** were absent in all transformants. It was speculated that the structures of products might not be those as expected. Products may be structurally modified by *A. oryzae*. Therefore, to find the new product with an unknown structure, the comparative LCMS analysis of transformants with and

without adding propionate was performed. Taken the transformant JFI88A as an example, which contains the hybrid PKS gene DEBS1<sub>w/oACP2</sub>-*sorAB*, the carboxylase gene *pccABE*, and the PPTase gene *sePptII*, 50 mM sodium propionate **89** was added to the culture of transformant at the day before chemical extraction, followed by LCMS analysis. The transformant in the absence of propionate was used as the control. Compared to the control, the LCMS results showed that no new peak was detectable on the DAD, TIC ES+, and TIC ES- chromatograms (Fig. 4.7). Likewise, in other groups no new peak was found either (data not shown).

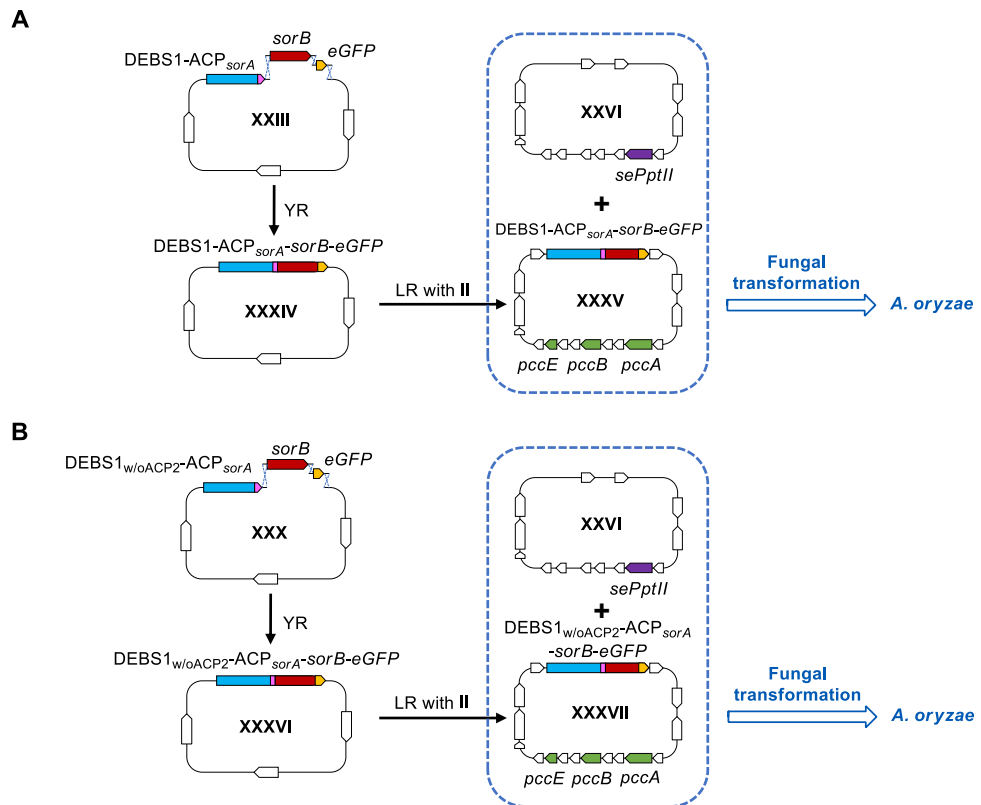
#### 4.3.2.2 Fusion Expression with eGFP

Considering the huge size (~ 20 kb) of the hybrid PKS gene containing DEBS1 and *sorAB*, the homologous recombination of the whole gene in *A. oryzae* may be difficult, leading to the production of DEBS1-*sorAB*-defective transformants. Consequently, to achieve the expression of the complete hybrid PKS, the reporter gene eGFP was fused to the 3' terminal of DEBS1-*sorAB* to screen transformants with highly-expressed PKS.

Two different groups of plasmid construction were constructed individually (Scheme 4.10 A). The first one started from the plasmid pEYA-DEBS1-ACP<sub>sorA</sub> **XXIII**. **XXIII** was digested by the unique restriction enzyme *HindIII*, and then two fragments *sorB* and *eGFP* were inserted by yeast recombination to form the plasmid pEYA-DEBS1-ACP<sub>sorA</sub>-*sorB*-*eGFP* **XXXIV**. The newly fused gene DEBS1-ACP<sub>sorA</sub>-*sorB*-*eGFP* was transferred to the expression plasmid **II** harboring the carboxylase gene *pccABE* by LR transformation, leading to the formation of the expression plasmid pTYGS-arg-DEBS1-ACP<sub>sorA</sub>-*sorB*-*eGFP*-*pccABE* **XXXV**. The plasmid pTYGS-met-*sePptII*\* **XXVI** containing the PPTase gene *sePptII* was the coexpression plasmid in this group for the next *A. oryzae* transformation.

For the second group, the construction was performed using *sorB* and *eGFP* (Scheme 4.10 B) based on the initiation plasmid pEYA-DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub> **XXX**. By yeast recombination, the plasmid pEYA-DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-*sorB*-*eGFP* **XXXVI** was formed. Then the whole gene fused to *eGFP* was transferred to the expression plasmid **II** by LR recombination to give the final plasmid pTYGS-arg-DEBS1<sub>w/oACP2</sub>-

ACP<sub>sorA-sorB-eGFP-pccABE</sub> XXXVII. Another plasmid XXVI containing the PPTase gene *sePptII* was chosen to coexpress in *A. oryzae*.



**Scheme 4.10** Construction route of plasmids containing the report gene *eGFP*. **A**, *eGFP* was fused to the 3' terminal of DEBS1-ACP<sub>sorA-sorB</sub>; **B**, *eGFP* was fused to the 3' terminal of DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA-sorB</sub>, in which ACP2 is deleted.

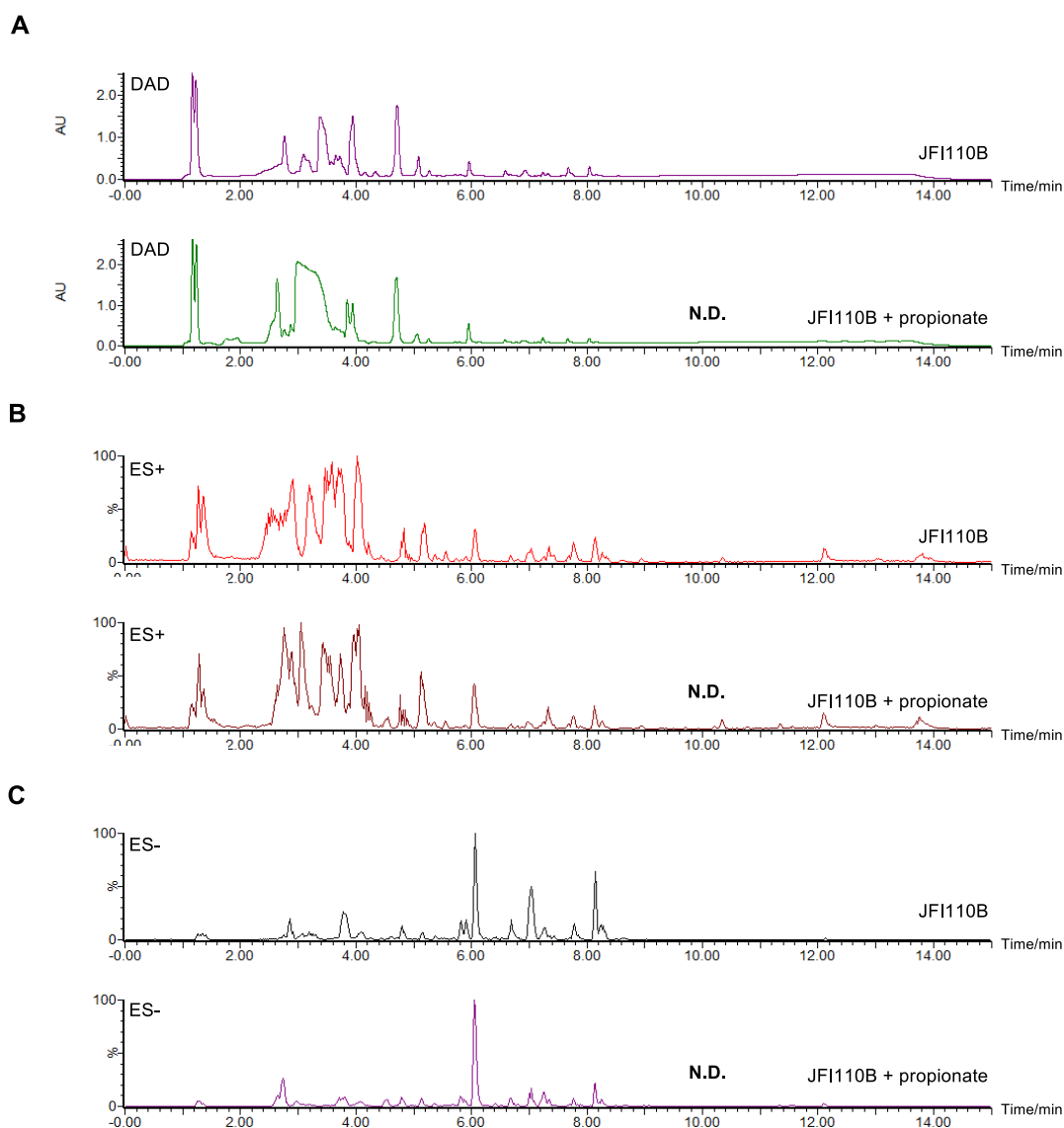
Two groups of plasmids were individually introduced into *A. oryzae*, including pTYGS-arg-DEBS1-ACP<sub>sorA-sorB-eGFP-pccABE</sub> XXXV & pTYGS-met-*sePptII*\* XXVI and pTYGS-arg-DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA-sorB-eGFP-pccABE</sub> XXXVII & pTYGS-met-*sePptII*\* XXVI (Table 4.2).

**Table 4.2** Different gene combinations for expression of hybrid PKS fused with *eGFP*.

Group	Expressed plasmid pTYGS-arg			Expressed plasmid pTYGS-met		The number of transformants	New product
	No.	Gene 1	Gene 2	No.	Gene 3		
JF1109	XXXV	DEBS1-ACP <sub>sorA-sorB</sub> -eGFP	<i>pccABE</i>	XXVI	<i>sePptII</i>	16 (A-P)	no
JF1110	XXXVII	DEBS1 <sub>w/oACP2</sub> -ACP <sub>sorA-sorB</sub> -eGFP	<i>pccABE</i>	XXVI	<i>sePptII</i>	12 (A-L)	no

Transformants were obtained (*i.g.* JF1109A-P and JF1110A-L) and then screened by fluorescence imaging as the above methods. It found that the positive rate of

fluorescent transformant is very low. In each group shown in Table. 4.2, only one fluorescent transformant was gained from more than ten transformants (16 JFI109 transformants and 12 JFI110 transformants), individually named as JFI109D and JFI110B. Subsequently, the fluorescent transformants were comparatively analyzed by LCMS under the absence and presence conditions of propionate. Unfortunately, LCMS analysis displayed that no additional peak was produced (JFI110B shown in Fig. 4.8).

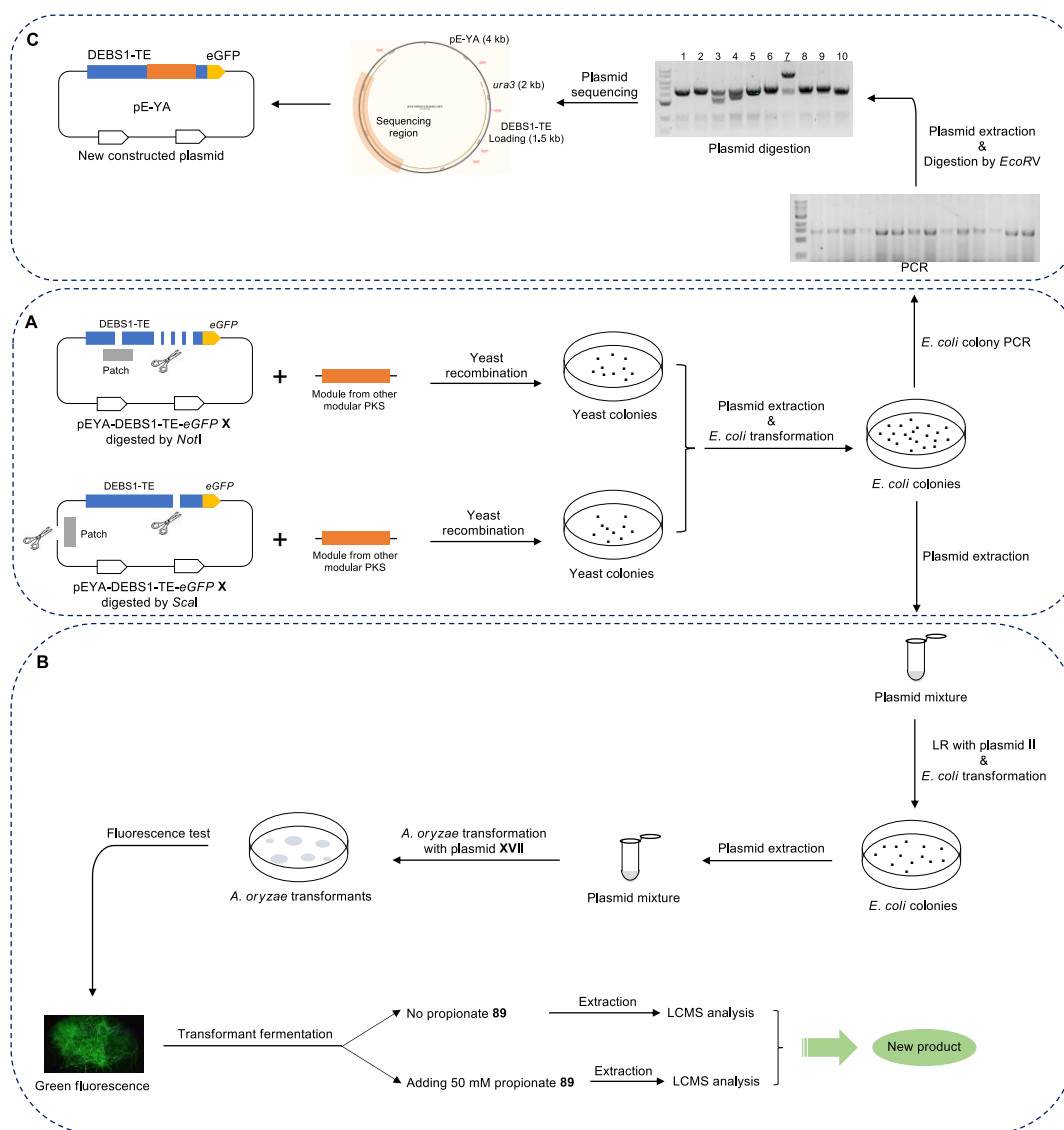


**Figure 4.8** Comparative LCMS analysis of the transformant JFI110B bearing hybrid PKSs fused with eGFP: **A**, the DAD comparison of the transformant with and without feeding propionate **89**; **B**, the ES+ comparison; **C**, the ES- comparison.

Abbreviation: N.D., not detected.

## 4.4 Results — Module Recombination

For module recombination between DBES-TE and other foreign modular PKSs, four characterized candidate modules provided by Prof. Kira J. Weissman group were chosen as donors, including: module 11 (AmpM11) from the amphotericin PKS; module 1 (LanM1) from the lankamycin PKS; module 7 (LasM7) from the lasalocid PKS; and module 11 (NysM11) from the nystatin PKS. The plasmid pEYA-DEBS1-TE-*eGFP* was chosen as the acceptor. Prior to recombination, module 2 of DEBS1-TE was cleaved by different restriction enzymes individually (Scheme 4.11 A). One enzyme *NotI* generated multiple digestion sites while the other enzyme *ScaI* only led to formation of one digestion site on module 2. The additional cuts on the plasmid were repaired by recombination using corresponding patches.



**Scheme 4.11** Workflow of module recombination and product analysis.

After introduction of digested pEYA-DEBS1-TE-*eGFP* X fragments along with a foreign module fragment to yeast, a large number of yeast colonies were obtained (Scheme 4.11 A) compared to control experiments. All colonies were harvested together, and then plasmids were extracted from the combined yeast colonies, followed by introduction into *E. coli*. The resulting *E. coli* colonies were mixed. The mixture of total plasmids was then extracted from *E. coli* colonies where both the positively-recombined plasmids and false positives should be included. Using the plasmid mixture as entry vectors and the carboxylase gene-containing plasmid pTYGS-arg-*pccABE* II as the destination vector, LR recombination was conducted *in vitro* (Scheme 4.11 B).<sup>62</sup> *E. coli* colonies were obtained, from which the plasmid mixture was extracted again. The plasmid mixture containing recombined-module expression vectors was ready for fungal transformation.

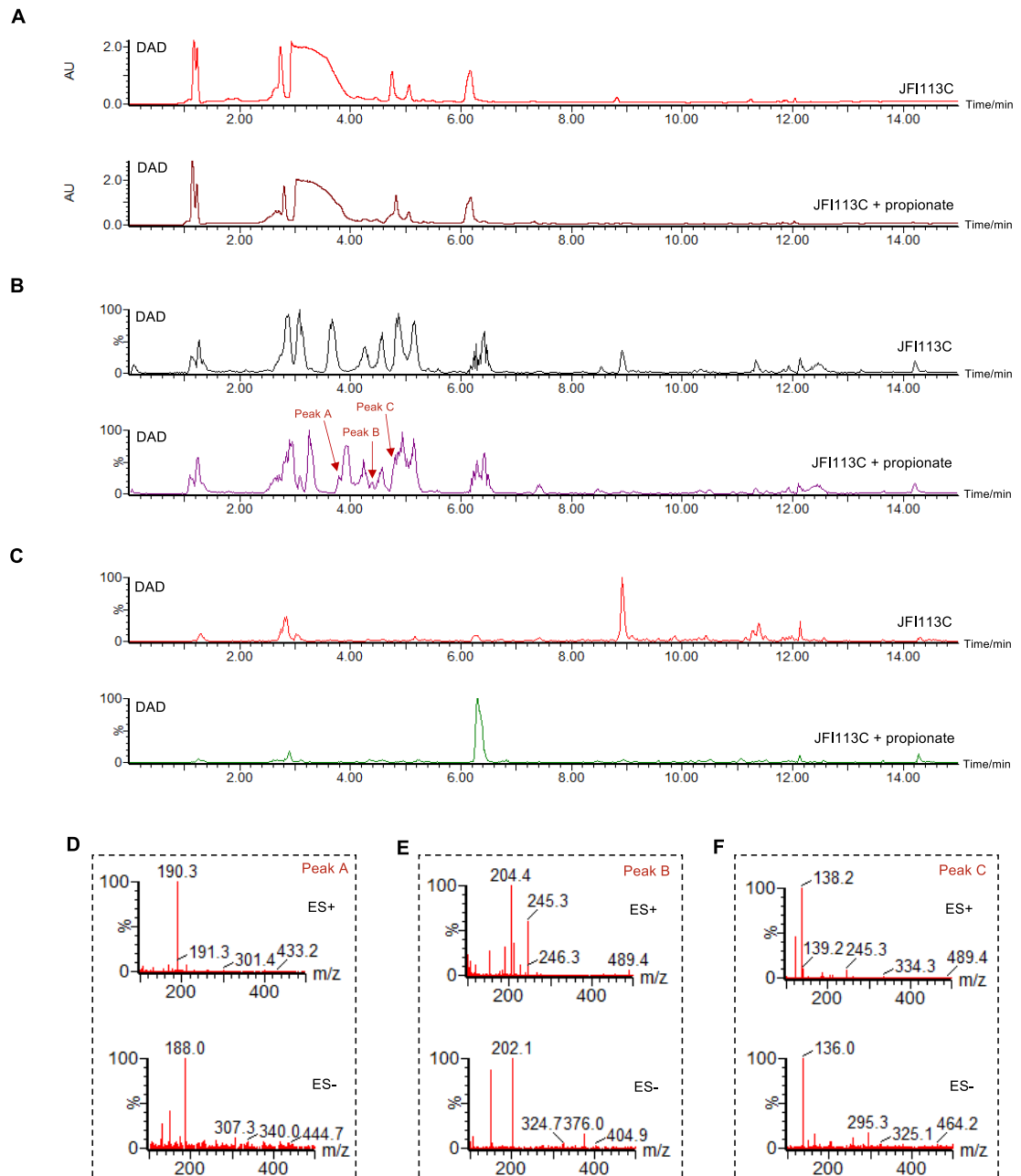
The plasmid mixture together with the PPTase gene-containing plasmid pTYGS-met-*sePptII* XVII were introduced into *A. oryzae* NSAR1. The corresponding number of transformants for AmpM11, LanM1, LasM7 and NysM11 were 3 (JFI111A-JFI111C), 4 (JFI112A-JFI111D), 5 (JFI113A-JFI111E) and 2 (JFI114A-JFI111B) respectively (Table 4.3). The resulting *A. oryzae* transformants were first tested by fluorescence imaging. Subsequently, the green fluorescent transformants (*i.g.* JFI111C, JFI112A, JFI112D, JFI113B, JFI113C and JFI114A) were fermented in DPY media. At the day before chemical extraction, 50 mM propionate **89** was added into the culture of transformants. As the methods described above, the comparative LCMS analysis was conducted under the absence and presence conditions of propionate. Lastly, the possible new products were searched from transformants.

**Table 4.3** Different module recombination between DEBS1-TE and other PKS modules.

Group	Expressed plasmid pTYGS-arg			pTYGS-met	The number of transformants	Fluorescent transformants	New product
	Gene 1	Gene 2	Gene 3	Gene 4			
AmpM11	recombined DEBS1-TE	eGFP	pccABE	sePptII	3 (JFI111A-C)	JFI111C	no
LanM1	recombined DEBS1-TE	eGFP	pccABE	sePptII	4 (JFI112A-D)	JFI112A, JFI112D	no
LasM7	recombined DEBS1-TE	eGFP	pccABE	sePptII	5 (JFI113A-E)	JFI113B, JFI113C	Peak A, B, C
NysM11	recombined DEBS1-TE	eGFP	pccABE	sePptII	2 (JFI114A-B)	JFI114A	no

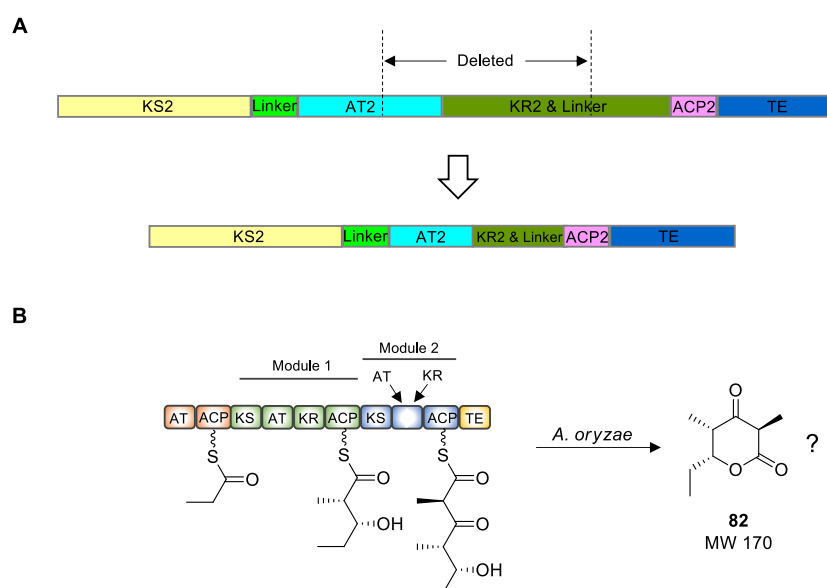


LCMS analysis showed that only in the transformant JF1113C three new peaks A, B and C were observed. Although the peaks were very small on the DAD chromatogram, they are significantly detectable on both the TIC ES+ and ES- chromatograms (Fig. 4.9 A, B and C). The molecular weight of the newly formed peaks was 189, 203 and 137, respectively (Fig. 4.9 D, E and F). In the control without feeding propionate **89**, the peaks were absent. In addition to that, it was found that the newly produced peaks do not exist in *A. oryzae* NSAR1, either (data not shown).



**Figure 4.9** LCMS analysis of the transformant JF1113C in the absence and presence of propionate: **A**, the DAD chromatogram comparison; **B**, the TIC ES+ chromatogram comparison; **C**, the TIC ES- chromatogram comparison; **D**, the ES+/- spectra of peak A; **E**, the ES+/- spectra of peak B; **F**, the ES+/- spectra of peak C.

However, odd masses mean incorporation of nitrogen, so they are unlikely to be related to DEBS1-TE. Based on this result, the earliest *E. coli* colonies derived from yeast recombination were screened by *E. coli* colony PCR (Scheme 4.11 C). Ten positive colonies harboring individual plasmid were selected, and then extracted plasmids. Next, the extracted plasmids were identified by restriction enzyme digestion using *EcoRV*. The result indicated that the plasmid No.7 was the most matchable with the predicted size. Then the plasmid was subject to gene sequencing. The recombined region of DEBS1-TE was sequenced, spliced and analyzed.



**Figure 4.10** Recombination analysis of module 2 of DEBS1-TE: **A**, the sequencing result after yeast recombination of module 2; **B**, the prediction of the product by the recombined DEBS1-TE.

The gene sequencing illustrated that the foreign module was not integrated into DEBS1-TE. Instead, a segment of DEBS1-TE itself containing AT2 domain and KR2 domain was deleted after yeast recombination (Fig. 4.10 A). Unexpectedly, a hybrid domain was formed, but the deletion of base pairs did not shift the open reading frame. However, the product of the recombined DEBS1-TE appears difficult to predict (Fig. 4.10 B). Therefore, the heterologous expression of the newly recombined plasmid in *A. oryzae* is required to do first in the future.

## 4.5 Discussion and Conclusion

*A. oryzae* has been developed to express modular PKSs for the first time, which further promotes the application of this heterologous expression system. In this chapter, we aimed to use this new *A. oryzae* system to express more diverse PKSs.

Reprogramming has been a widely used strategy to generate more diverse PKSs. Both modular PKSs and iterative PKSs have developed many examples of reprogramming. However, the combination between modular and iterative PKSs has never been investigated so far. If it is feasible, the advantage of each type of PKS may be utilized to direct the synthesis of targeted product or create more novel structures in this new *A. oryzae* system.

In this chapter, we aimed to explore the possibility of fusion expression of modular and iterative PKSs in the new *A. oryzae* system. DEBS1-TE was chosen since it had been expressed in *A. oryzae* for the first time. The PKS SorAB was a fungal example which had been successfully expressed in *A. oryzae*.<sup>66</sup> The key point is the connection of PKSs to generate a functionally active PKS. The biosynthetic segment in SorA for the earliest triketide production was replaced by DEBS1 (Scheme 4.7). In this way, the intermediate triketide produced by DEBS1 successively gets into the assembly line by SorB. However, it is unknown whether the cascade of ACP2 from DEBS1 and ACP from SorA could transfer the intermediate. Therefore, the removal of ACP2 from DEBS1 was also examined. A series of fused hybrid PKSs were designed and expressed in *A. oryzae*. Unfortunately, there was no new product detectable. Many possible factors may influence the production of product, such as the low yield of *A. oryzae* itself, the domain recognition between modular and iterative PKSs, and the efficiency of homologous recombination of huge genes.

The other attempt in this chapter was the module recombination of modular PKS by yeast. Four characterized modules from bacterial PKSs were chosen to integrate into the module 2 region by random homologous recombination in yeast. Ultimately, it was found that the foreign module is not recombined into DEBS1-TE. Instead, the self-connection of the plasmid occurred in yeast, leading to the formation of a new truncated DEBS1-TE. Meanwhile, three new products were also detected in the corresponding transformant harboring the recombined plasmid. Although the characterization of these products has not been conducted, they are

unlikely to be associated with the recombined DESB1-TE because of the odd masses. Due to the time constraint, it is impossible to proceed the product isolation and characterization.

In the future, more interactions between modular and iterative PKSs can be established in this new *A. oryzae* heterologous expression system.

## 5. Expression and Engineering of the Brasilane Gene Cluster

### 5.1 Introduction

Brasilane-type sesquiterpenes are widely distributed in nature. They generally possess two typical features in their structures, five methyl groups and a 5/6 bicyclic carbon skeleton.<sup>158</sup> These unusual structures have drawn the interest of synthetic and biosynthetic chemists in the past decades. Meanwhile, in pharmacological studies the brasilanes also display interesting bioactivities. Therefore, an overview of brasilane-type sesquiterpenes is presented in this chapter in terms of structural diversity, potential bioactivities for clinical applications, and relevant biosynthetic studies.

#### 5.1.1 Brasilane-type Sesquiterpenes in Nature

The brasilane core structure appears to be produced by many organism families, including basidiomycetes, various ascomycetes, algae, liverwort and sea hare. To date, more than 30 brasilane-type sesquiterpenes have been isolated and spectroscopically characterized (Fig. 5.1). The first compounds were reported in 1978 when brasilenol **143**, epibrasilenol **144**, and brasilenol acetate **145** were isolated from the marine opisthobranch *Aplysia brasiliana*.<sup>159</sup> They all possess the characteristic brasilane-type skeleton. In 1990, two new rearranged sesquiterpenes **146** and **147** based on the brasilenol scaffold were obtained from the Mediterranean marine alga *Laurencia obtusa* and one brasilane-type sesquiterpene alcohol conocephalenol **148** was isolated from the Scottish liverwort *Conocephalum conicum*, respectively.<sup>160-161</sup>

Subsequently, the genus *Laurencia* received more attention because of the constituent variability in the collections from different sites. In 1991, three new brasilane derivatives **149-151** were isolated from the Mediterranean marine alga *L. obtuse* as well.<sup>162</sup> The structure of **149** contains a 5,6-double bond instead of a 1,6-double bond. At the same time, three new brasilane derivatives **152-154** were isolated from *L. implicata* collected in Australia.<sup>163</sup> Among them, the absolute

configuration of **154** was revised by the isomer's total synthesis.<sup>164</sup> In 1999, three more new brasilane-type sesquiterpenes **155-157** were isolated from *C. conium* collected in southern Germany.<sup>165</sup> In 2002, it was discovered that the halogenated brasilenols **158-160** are present in the strain *L. obtusa* derived from Greek seas.<sup>166</sup> Interestingly, these metabolites contain an unprecedented 1,6-epoxy moiety as well.

In 2008, a new brasilane-type sesquiterpene xylarenic acid **161** was obtained from the endophytic fungus *Xylaria* sp. NCY2.<sup>167</sup> In 2015 altering cultivation parameters resulted in the discovery of six new brasilane-type sesquiterpenes diaporols J-O (**162-167**) from the endophyte fungus *Diaporthe* sp. collected in China.<sup>168</sup> In another investigation, Hu *et al.* isolated two new brasilane-type sesquiterpene derivatives brasilane B **168** and brasilane C **169** and a glycoside metabolite brasilane A **2** with an *N*-acetylglucosamine moiety from cultures of the basidiomycete *Coltricia sideroides*.<sup>169</sup> It is the first time to see a glycosylated form in brasilane-type metabolites. On the other hand, the secondary metabolites harboring this kind of sugar moiety are also quite rare in the entire field of natural products.

Shortly after, another brasilane-type sesquiterpene glycoside hypoxyside **170** was isolated from an endolichenic fungus *Hypoxyton fuscum*.<sup>170</sup> In this molecule, a glucose moiety is linked to the brasilane structure. When the brasilane-type biosynthesis was investigated, a new sesquiterpene trichobrasilenol **171** was isolated and identified by the heterologous expression in *Aspergillus oryzae* and the *in vitro* conversion with the precursor FPP.<sup>171</sup> This new sesquiterpene is coincidentally the aglycone of the glycoside brasilane A. In another biosynthetic study, a new sesquiterpene synthase from the fungus *Trichoderma viride* was identified biochemically.<sup>172</sup> It is responsible for the production of a new sesquiterpene **172**. Meantime, a new esterified derivative **173** was also discovered when expressed in yeast heterologously. In addition, six highly oxygenated brasilane-type sesquiterpenes with an  $\alpha$ ,  $\beta$ -unsaturated ketone unit, brasilanones A-F (**174-179**), were isolated from the marine-derived fungus *Aspergillus terreus*.<sup>158</sup> This is the first discovery of brasilane-type sesquiterpenes from the genus *Aspergillus*.

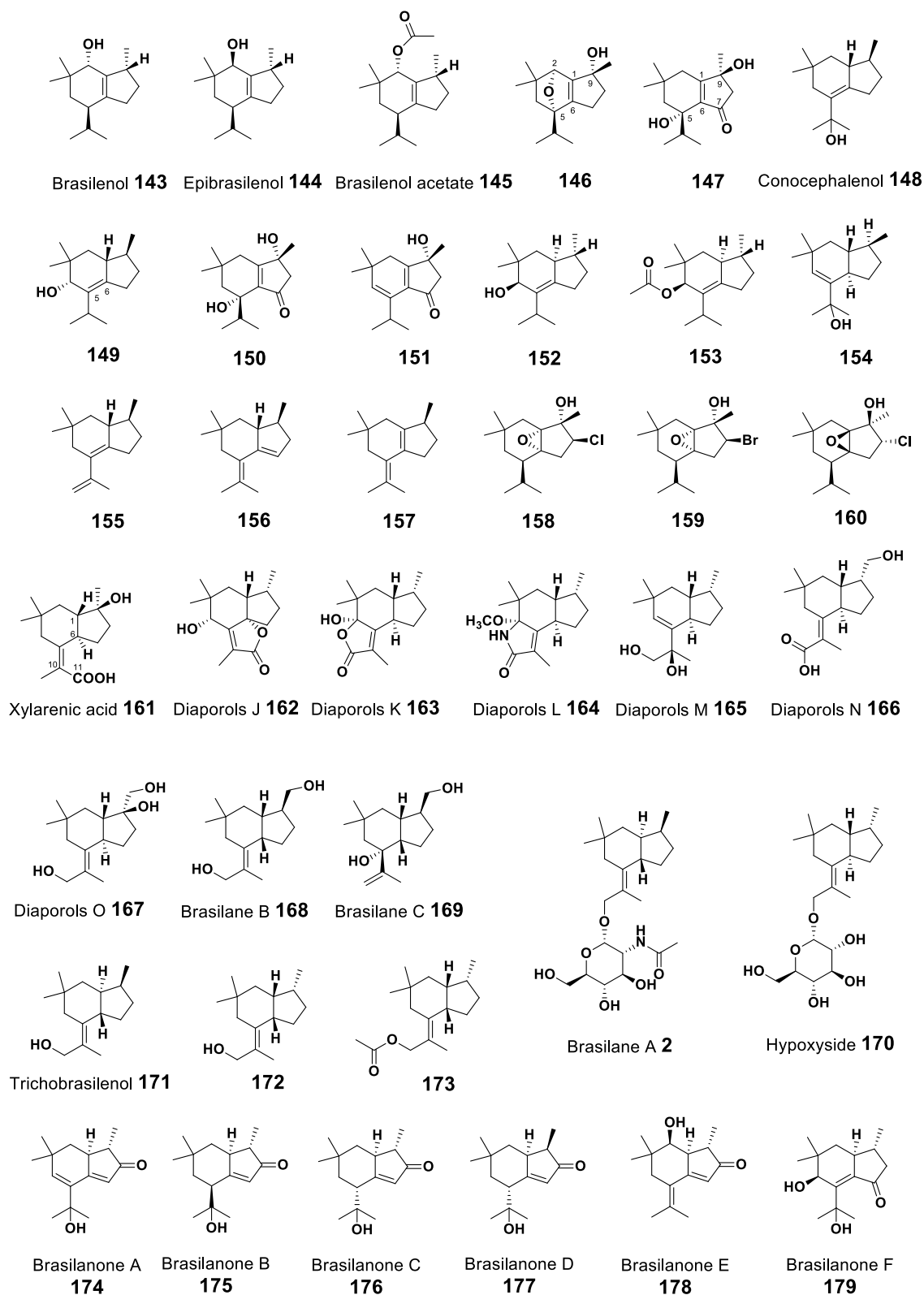
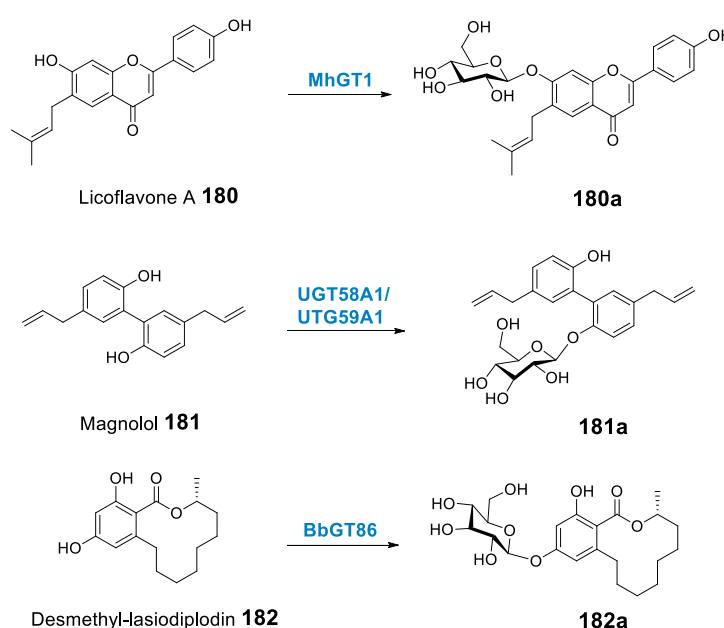


Figure 5.1 Chemical structures of brasilane-type sesquiterpenes.

### 5.1.2 Fungal Glycosyltransferases and Glycosylated Sesquiterpenes

Glycosyltransferases are the catalyst of glycosylation during the production of natural products. They are generally considered as an efficient modification approach after the skeletal scaffold of the molecule is formed. On the one hand, their scope of modification is very broad. According to some studies of microbial transformation, it has been well known that a large variety of natural products or small molecules could be glycosylated by fungi *in vivo*, especially some complex scaffolds. On the other hand, fungi also possess quite high conversion of glycosylation compared with that of bacteria. Therefore, fungal glycosyltransferases are a huge and potent resource to exploit.

However, only a few fungal glycosyltransferases have been discovered and functionally characterized so far (Scheme 5.1). These glycosyltransferases have the features in common, such as low sequence identity and less conserved motif. Some others even have transmembrane regions in their protein sequence. These disadvantages have limited or delayed the further discovery of fungal glycosyltransferases.



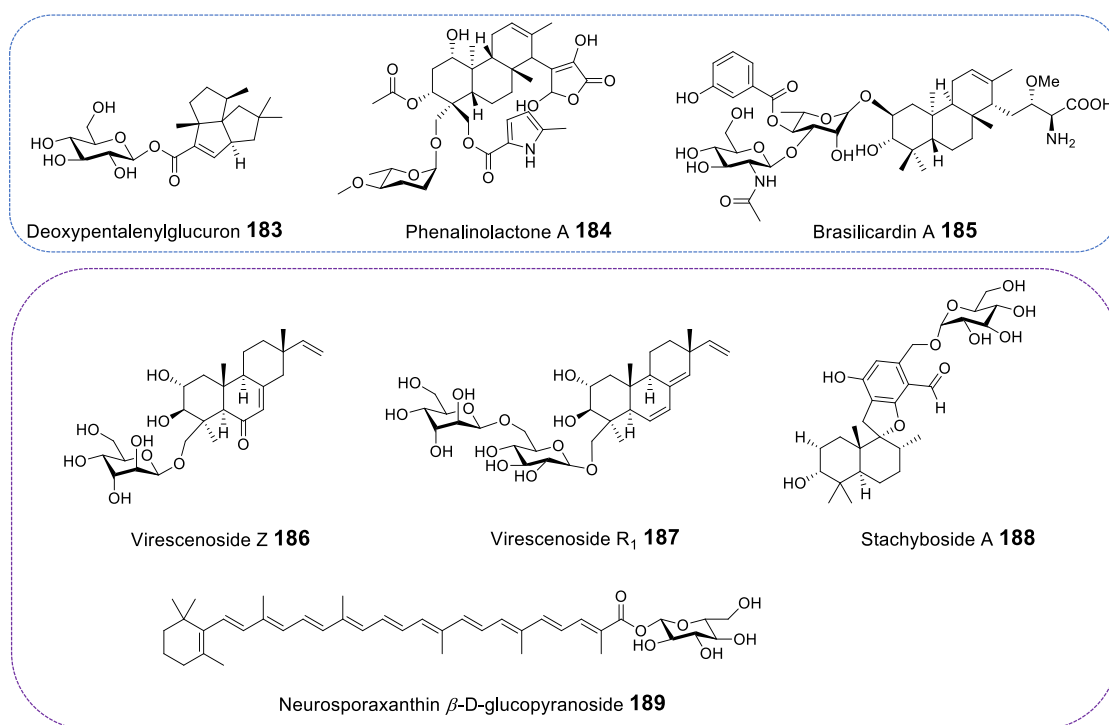
**Scheme 5.1** Characterized glycosyltransferases from fungi and their respective catalyzation.

The fungal glycosyltransferases involved in small-molecule natural products were described intensively in 2017. A versatile *O*-glycosyltransferase from the fungus *Mucor hiemalis* was biochemically characterized and explored.<sup>173</sup> It displayed a wide



substrate promiscuity and regioselectivity for six different kinds of phenolic compounds. Subsequently, two more phenolic glycosyltransferases (UGT58A1 and UGT59A1) were identified in *Absidia coerulea* and *Rhizopus japonicus*, respectively.<sup>174</sup> Both of them are membrane-bound glycosyltransferases. Another prominent example is the discovery of a glycosyltransferase from *Beauveria bassiana*.<sup>175-176</sup> It can catalyze many types of substrates, such as flavonoids, stilbenoids, anthraquinones, and benzenediol lactones. However, there is still no fungal terpene glycosyltransferases reported so far.

In addition, the glycosylated sesquiterpenes deserve a mention here. They are very important subfamily members among various sesquiterpenes as it is well-known that the glycosylation of natural products plays a significant role in altering the properties of the parent scaffold.



**Figure 5.2** Examples of glycosylated terpene metabolites from bacteria (in blue area) and fungi (in purple area).

While a large number of glycosylated terpenes have been discovered in plants, this type of natural product is relatively rare from microorganisms. In bacteria, approximately 41 glycosylated terpenes have been reported to date,<sup>177</sup> such as sesquiterpenes deoxypentalenyglucuron **183**, diterpenes phenalinolactone A **184** and brasilicardin A **185** (Fig. 5.2).<sup>178-180</sup> In fungi the glycosylated terpenes are even

rarer. Most of them derive from the marine fungi. For example, a series of diterpene virescenosides M-X, Z, R<sub>1</sub>-R<sub>3</sub> and Z<sub>4</sub>-Z<sub>8</sub> were isolated from the fungus *Acremonium striatisporum*, including both monoglycosides (*e.g.* virescenoside Z **186**) and diglycosides (*e.g.* virescenoside R<sub>1</sub> **187**).<sup>181</sup> In 2013, two phenylspirodrimane-type glucosidic meroterpenoids stachybosides (*e.g.* stachyboside A **188**) were separated from the sponge-derived fungus, *Stachybotrys chartarum* MXH-73.<sup>182</sup> In addition, the carotenoid glycosyl ester neurosporaxanthin  $\beta$ -D-glucopyranoside **189** was isolated from *Fusarium* sp. collected from the seawater surface (Fig. 5.2).<sup>183</sup>

### 5.1.3 Bioactivities

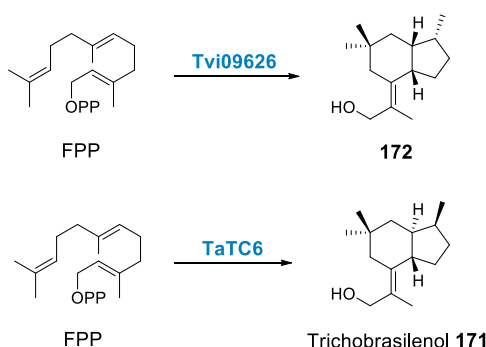
Brasilane-type sesquiterpenes are widespread in nature though, the pharmacological studies have not been extensively explored due to the insufficiency of metabolites. To date, only three brasilane derivatives have been investigated pharmacologically. It revealed that brasilane B **168** presents the weak cytotoxicity against human myeloidleukemia HL-60<sup>169</sup> and brasilanone A **174** and brasilanone E **175** show up moderate inhibitory effect with NO inhibition rates of 47.7 % ( $p < 0.001$ ) and 37.3 % ( $p < 0.001$ ) at the concentration of 40  $\mu$ M, respectively.<sup>158</sup> Hence, it is quite promising to figure out potential bioactivities from unknown brasilane derivatives.

Given the linkage of sugar moiety to the core structure of terpenes, the bioactivities of relevant glycosylated terpenes from microorganisms were searched. For example, the glycosylated diterpene phenalinolactone A **184** displays moderate growth inhibition towards Gram positive bacteria.<sup>180</sup> Brasilicardin A **185**, which possesses a rhamnose and an *N*-acetylglucosamine on the skeleton, exhibited a potent immunosuppressive activity.<sup>179</sup> Although no reports supported the role of glycosylation in enhancing bioactivity in the sesquiterpene subfamily, it is still plausible that the glycosylated sesquiterpenes, for instance the brasilane-type sesquiterpenes, are more capable of exerting their pharmacological activities after being modified by glycosylation or beneficial to improve the water solubility of sesquiterpene aglycone *in vivo*.

### 5.1.4 Relevant Biosynthesis

Terpenes, as the largest class of natural products, have been investigated extensively on their biosynthesis. The basic carbon skeleton is built from a few precursors, such as geranyl (GPP), farnesyl (FPP), geranylgeranyl (GGPP), and geranylarnesyl (GFPP) diphosphate. These precursors are converted to various terpenes by terpene cyclases (TCs) that could ionize precursors by abstracting or protonating diphosphate. In the initial process of terpene formation, the intermediate cation usually causes the cyclization of carbon skeleton. Therefore, a polycyclic skeleton is formed and meanwhile possesses a few stereochemistry centers.

The first brasilane-type sesquiterpene synthase Tvi09626 was discovered in the filamentous fungus *Trichoderma viride* J1-030 through genome mining (Scheme 5.2). Its protein sequence shares identities of 90 % and 85 % with the enzymes from the fungi *T. virens* Gv29-8 and *T. reesei* QM6a, respectively. The protein sequence alignment shows that Tvi09626 has the highly conserved regions, including motifs DDxxD/E and NSE/DTE. By the phylogenetic analysis, Tvi09626 is located to the Clade V of Class I terpene synthases. By *in vitro* study, it clearly revealed that only FPP can be recognized as substrate by Tvi09626 to form the compound **172**. This is the first case that uses a biosynthetic approach to synthesize the brasilane-type sesquiterpene.

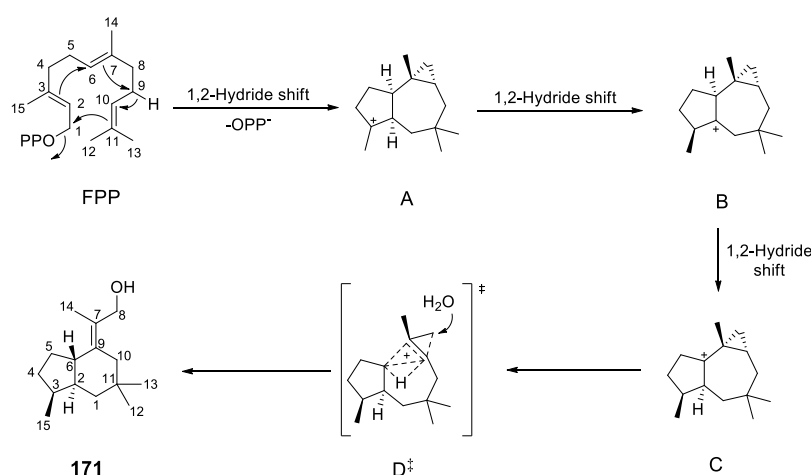


**Scheme 5.2** Production of two brasilane-type sesquiterpenes by Tvi09626 and TaTC6 *in vitro*.

The second rare example is the discovery of the biosynthetic gene *taTC6* of the brasilane-type sesquiterpene trichobrasilenol **171** from the fungi *Trichoderma atroviride* FKI-3849 (Scheme 5.2). It also belongs to the type I family of terpene synthase while it is phylogenetically distant from other characterized fungal terpene

synthases. It was identified and characterized by the heterologous expression in *Aspergillus oryzae* and the enzymatic assay *in vitro*. The *in vitro* study also exhibited that this terpene synthase only accepts FPP as substrate.

For many known terpenes, the formation mechanism of their polycyclic skeleton is relatively obvious to deduce. Isotopic labelling methods can elucidate some cases that involve additional rearrangement procedures or irregular precursors. However, in some other cases, an unusual rearrangement of polycyclic skeleton occurs. In particular, the number of methyl groups (or their equivalents) after rearrangement is even more than that of those in the precursors. A typical example is brasilane-type sesquiterpene alcohol xylarenic acid **161**.<sup>167</sup> It is assembled from FPP which has 4 methyl groups, but there is an extra methyl group equivalent (C-11) on its structure. The complex and unknown formation mechanism of this additional methyl group was addressed by Murai and his coworkers.<sup>171</sup>



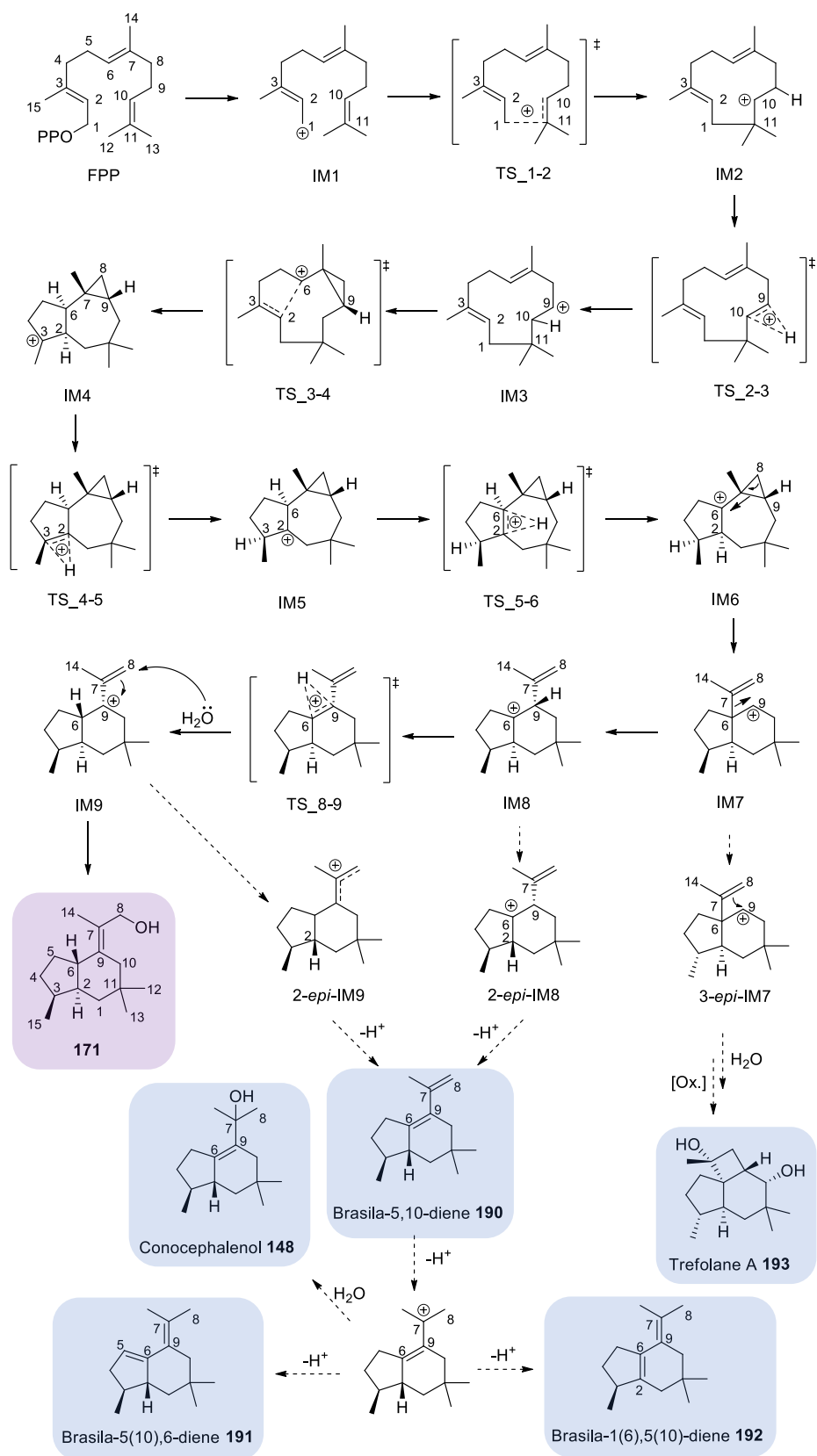
**Scheme 5.3** Cyclization mechanism from FPP to trichobrasilenol **171** (Compound numbering only shows the origin of each carbon from FPP).

By the isotopic labelling, a mechanism of the skeletal rearrangement into trichobrasilenol **171** was proposed (Scheme 5.3). The initial reaction starts by loss of diphosphate at C-1 and then a 1,11-cyclization. To prevent secondary cations as intermediates from forming side products, a simultaneous 1,2-hydride migration and double ring formation take place to **A**. Then two more 1,2-hydride migrations lead to intermediate **C**. **C** is attacked by one molecule of water to conduct the rearrangement through the transition state **D<sup>‡</sup>**, which involves another simultaneous hydride migration, cyclopropane ring opening, cyclohexane ring

formation, and double bond introduction, eventually leading to trichobrasilenol **171** with an extra methyl group equivalent (C-8).

However, this unusual skeletal rearrangement is still controversial because of its feasibility of six chemical events occurred in one step. Sato *et al.* proposed a different mechanism based on the density functional theory (DFT) calculations (Scheme 5.4).<sup>184</sup> In their study, they firstly established a powerful method that combines quantum-chemical calculations with the artificial force induced reaction (AFIR). Then it was adopted to analyze the reaction mechanism of this usual skeletal rearrangement. The results indicated that the biosynthetic pathway of trichobrasilenol **171** is involved a multistep carbocation cascade, in which cyclopropylcarbinyl cations in equilibrium with homoallyl cations play a pivotal role.

The reaction cascade starts with the dissociation of the pyrophosphate of FPP to form an allylic carbocation (IM1) and then give a humulyl cation (IM2). A 1,2-hydride shift and two more annulations take place to form the 5/7/3 tricyclic intermediate (IM4). Subsequently, two successive 1,2-hydride shifts occur within the tertiary carbocations leading to the formation of the most stable tertiary cation (IM6). The intermediate IM6 contains a highly distorted cyclopropane ring. Therefore, an unusual skeletal rearrangement with ring contraction occurs to form a 5/6 bicyclic intermediate (IM8). Afterwards, a 1,2-hydride shift takes place again to give an energetically more stable allylic carbocation (IM8). Finally, a hydration reaction proceeds at C-8 to yield trichobrasilenol **171**. Notably, according to this new mechanism, the possible biosynthetic pathways of related byproducts of trichobrasilenol could be elucidated as well, such as conocephalenol **148**, brasila-5(10)-diene **190**, brisila-5(10),6-diene **191**, brasila-1(6), 5(10)-diene **192**, and trefolane A **193**.



**Scheme 5.4** Proposed biosynthetic pathway of trichobrasilenol **171** (purple) and related byproducts (blue) based on the DFT calculations (IM stands for intermediates and TS stands for transition state).

## 5.2 Aims

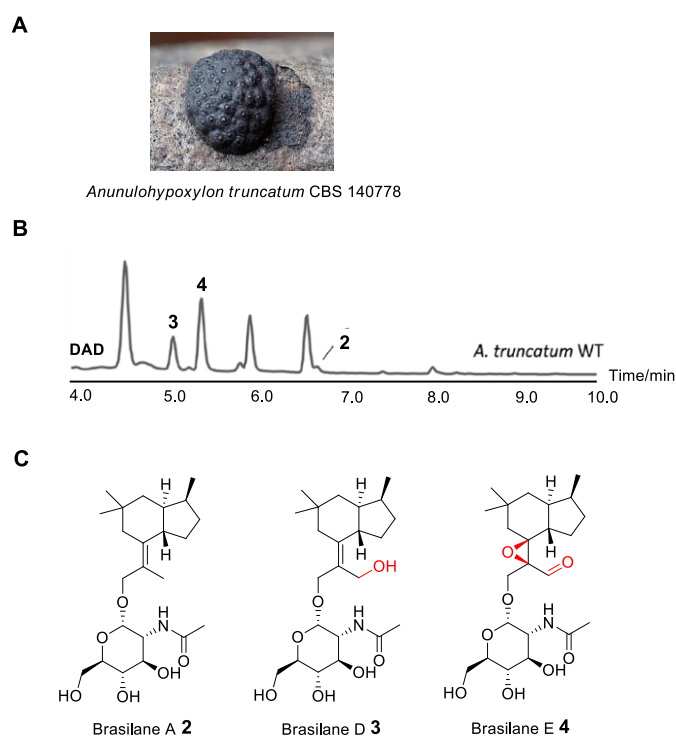
In all known brasilane-type sesquiterpenes there are only two brasilane-type glycosides, a glucoside and an *N*-acetylglucoside. Therefore, it must have a rare type of glycosyltransferase capable of recognizing sesquiterpene aglycones as substrate, and this deserves more attention.

Thus, the aim in this chapter is to identify a biosynthetic gene cluster responsible for brasilane-type sesquiterpene biosynthesis in a *Hypoxylaceae* species by bioinformatics analysis. For each putative gene in the gene cluster, it aims to figure out its corresponding biochemical function during the production of brasilane-type sesquiterpene. In this whole scheme, the heterologous expression strategy in *Aspergillus oryzae* will be adopted as an effective avenue to elucidate the biosynthetic pathway *in vivo*. In addition, the *in vitro* assay will also be utilized to characterize the crucial enzymes present in this pathway, including sesquiterpene synthase and glycosyltransferase.

## 5.3 Results

### 5.3.1 Production and Isolation of Brasilane-type Sesquiterpenes

*Annulohyphoxylon truncatum* CBS 140778 was used in this work (Fig. 5.3 A) and was obtained from Dr. E. Kuhnert. It was grown in YMG medium for 14 days through a successive two-step inoculation procedure. The culture broth was extracted and yielded *ca* 1 g of crude extract. The crude extract was fractionated by flash chromatography and preparative HPLC (Fig. 5.3 B). Three pure compounds **2**, **3** and **4** were obtained (Fig. 5.3 C).

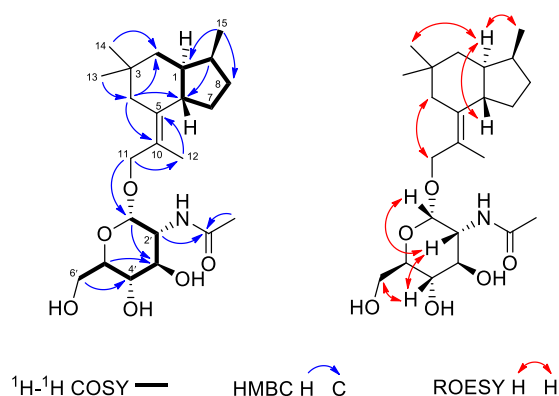


**Figure 5.3** Discovery and isolation of brasilane-type sesquiterpenes: **A**, Phenotype of *Annullohyphoxylon truncatum* CBS 140778; **B**, HPLC chromatogram of the crude extract of *A. truncatum* cultivated in YMG medium; **C**, chemical structures of three isolated metabolites **2-4**.

Metabolite **2** was isolated as a colorless oil (2.4 mg). Its molecular formula  $C_{23}H_{39}NO_6$  was deduced from its pseudomolecular ion peak cluster at  $m/z$  426.2840 in the HRESIMS spectrum. Its  $^1H$  and HSQC spectra revealed the presence of five methyls, six aliphatic methylenes and eight methines. The  $^{13}C$  spectrum indicated the additional existence of one carbonyl, one quaternary aliphatic and two olefinic carbon atoms. A database search with this data within the Chapman & Halls Dictionary of Natural Products on DVD suggested metabolite **2** being brasilane A,

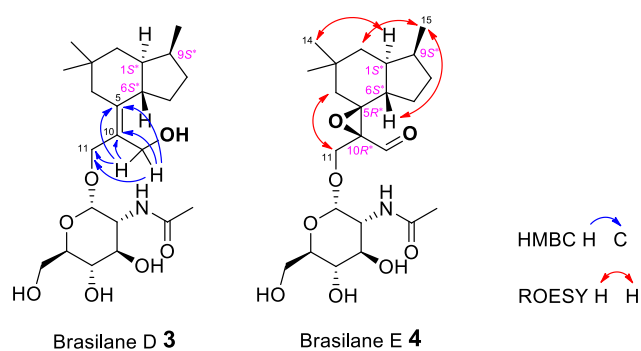


whose identity was confirmed by the comparison of  $^{13}\text{C}$  NMR data obtained in methanol- $d_4$  with those of Hu *et al.* (Fig. 5.4).<sup>169</sup>



**Figure 5.4** Key 2D NMR correlations of brasilane A **2**.

The colorless oil brasilane D **3** (2.9 mg) was analyzed for a molecular formula of  $\text{C}_{23}\text{H}_{39}\text{NO}_7$  based on its quasimolecular ion cluster at  $m/z$  442.2830 in the HRESIMS spectrum, implying the formal addition of an oxygen atom compared to brasilane A **2**. The NMR data of **3** were highly similar to those of brasilane A **2**, with the key difference being the replacement of the methyl group 12- $\text{H}_3$  by an oxymethylene ( $\delta_{\text{C}}$  61.2;  $\delta_{\text{H}}$  4.26, 4.33). HMBC correlation of 12- $\text{H}_a$  and 12- $\text{H}_b$  to C-5, C-10 and C-11 confirmed this assignment (Fig. 5.5). Since the chemical shifts and coupling constants of **3** are comparable with those of **2**, the common  $1\text{S}^*$ ,  $6\text{S}^*$ ,  $9\text{S}^*$  configuration was assigned for **3**.



**Figure 5.5** Key 2D NMR correlations of brasilane D **3** and brasilane E **4**.

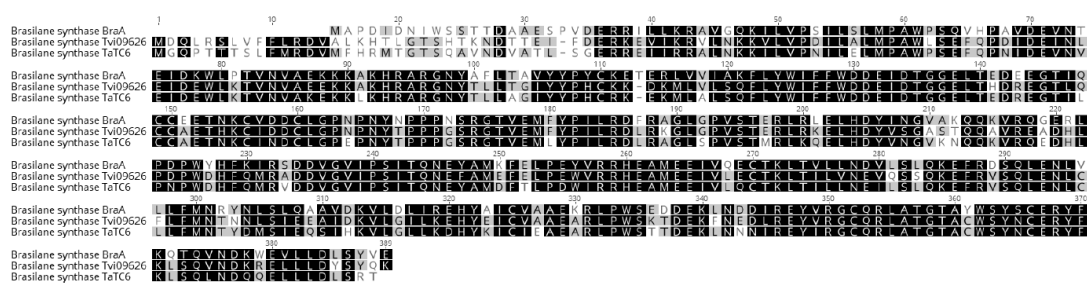
Brasilane E **4** was also isolated as a colorless oil (7.3 mg). The HRESIMS spectrum showed a quasimolecular ion cluster at  $m/z$  456.2599, implying the molecular formula  $\text{C}_{23}\text{H}_{37}\text{NO}_8$ . The NMR data of **4** were very similar to those of brasilane A **2**.

However, an aldehyde moiety was observed ( $\delta_{\text{H}}$  9.85,  $\delta_{\text{C}}$  201.7) instead of methyl 12- $\text{H}_3$ . Furthermore, carbon atoms C-5 ( $\delta_{\text{C}}$  71.5) and C-10 ( $\delta_{\text{C}}$  33) were shifted significantly to high field, suggesting the bonding of oxygen atoms to both in place of the  $\Delta^{5,10}$  double bond. Due to the molecular formula, an epoxide moiety was deduced. Analogously to brasilane **2**, the ROESY correlation between 15- $\text{H}_3$  and 6- $\text{H}/2\text{-H}_a$  in combination with the one from 14- $\text{H}_3$  and 1- $\text{H}$  revealed the relative configuration as  $1S^*$ ,  $6S^*$ ,  $9S^*$  (Fig. 5.5). ROESY correlations further enabled us to assign the two additional stereo centers of **4**. Since 4- $\text{H}_a$  and 4- $\text{H}_b$  show ROESY correlations to 11- $\text{H}_b$ , the C-10 is in equatorial position in contrast to the axial epoxide, and **4** has a  $1S^*$ ,  $5R^*$ ,  $6S^*$ ,  $9S^*$ ,  $10R^*$  configuration.

### 5.3.2 Bioinformatic Analysis of *Brasilanes* BGC

The genome of *A. truncatum* was sequenced with a combination of Oxford Nanopore and Illumina technology by colleagues at CeBiTec and the results were previously published.<sup>185</sup> Gene prediction was performed by applying Augustus version 3.2 and GeneMark-ES 4.3.6. using default settings.<sup>186-187</sup> For Augustus, species parameter sets were established based on GeneMark-ES fungal version predictions. Predicted genes were functionally annotated using a modified version of the genome annotation platform GenDB 2.0 for eukaryotic genomes as previously described.<sup>188-</sup>

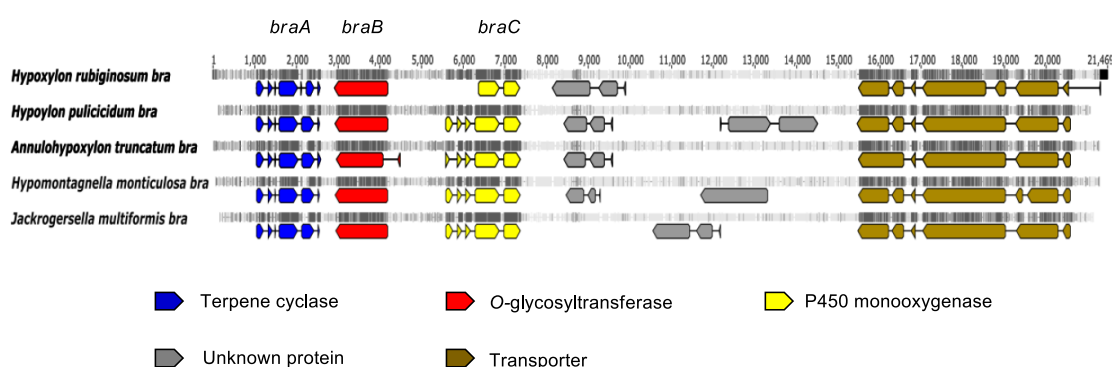
189



**Figure 5.6** Alignment of the brasilane synthase protein sequences from *Annulohyphoxylon truncatum* (BraA), *Trichoderma viride* (Tvi09626) and *T. atroviride* (TaTC6). Regions with high similarity across all sequences are highlighted in black.

Candidate gene clusters were preliminarily identified by manual blastp searches against an *A. truncatum* protein database using the Tvi09626 and TaTC6 brasilane

synthase (accession number LC484924) as template.<sup>171-172</sup> Nine terpene cyclases with homology to the latter were identified, but only one (*bra*) was associated with the predicted tailoring genes (glycosyl transferase and monooxygenase) and had the high sequence identity of terpene cyclase to both templates (Fig. 5.6). We also found near-identical gene clusters in related genome-sequenced species of the *Hypoxylaceae* (*Hypomontagnella monticulosa*, *Hypoxylon pulicidum*, *Hypoxylon rubiginosum*, and *Jackrogersella multiformis*, Fig. 5.7),<sup>185</sup> some of which were able to produce compounds brasilane A **2**, D **3**, and E **4** (data not shown). The high similarity of the *bra* cluster sequences enabled us to validate the respective coding sequences from *A. truncatum* due to the availability of transcriptome data for *Hypom. monticulosa*.<sup>185</sup>



**Figure 5.7** Alignment of the *bra* gene clusters found in different species of the *Hypoxylaceae*. Predicted coding regions of the different genes are shown below the sequences revealing putative erroneous gene predictions. Regions with high nucleotide similarity are shown in dark grey and indicate a high degree of gene conservation for the *bra* core genes.

The re-annotated *A. truncatum bra* cluster was uploaded to GenBank and can be accessed under the accession number MT383109.

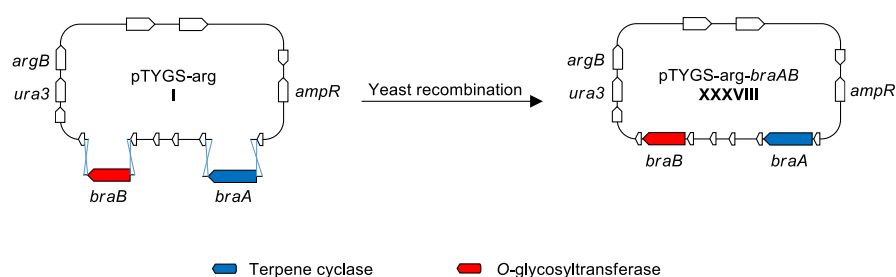
### 5.3.3 Heterologous Expression of the Brasilane BGC in *A. oryzae*

Based on the bioinformatics analysis, the *bra* gene cluster was chosen to be the objective for further investigation. To elucidate the function of each gene in the cluster, the gene knockout approach was initially used. Unfortunately, it was not possible to produce a mutant of *A. truncatum* in which *braA* gene was deleted. Therefore, the fungal heterologous expression was adopted as an alternative

strategy.<sup>62</sup> Here, a quadruple auxotrophic mutant *A. oryzae* NSAR1 (*niaD*<sup>-</sup>, *sC*<sup>-</sup>, *argB*<sup>-</sup>, *adeA*<sup>-</sup>) was chosen as a host in this work.

### *A. oryzae braAB*

Ahead of the heterologous expression study in *A. oryzae*, genomic DNA was extracted from a 5-day old liquid culture of *A. truncatum* grown in YMG medium. First, *braA* and *braB* were directly amplified from gDNA using PCR. The introns in both *braA* and *braB* were not removed. The amplified fragments were then inserted into the plasmid pTYGS-arg I with *argB* as auxotrophic marker by yeast homologous recombination as described previously,<sup>190</sup> leading to the formation of the vector pTYGS-arg-*braAB* XXXVIII (Scheme 5.5).

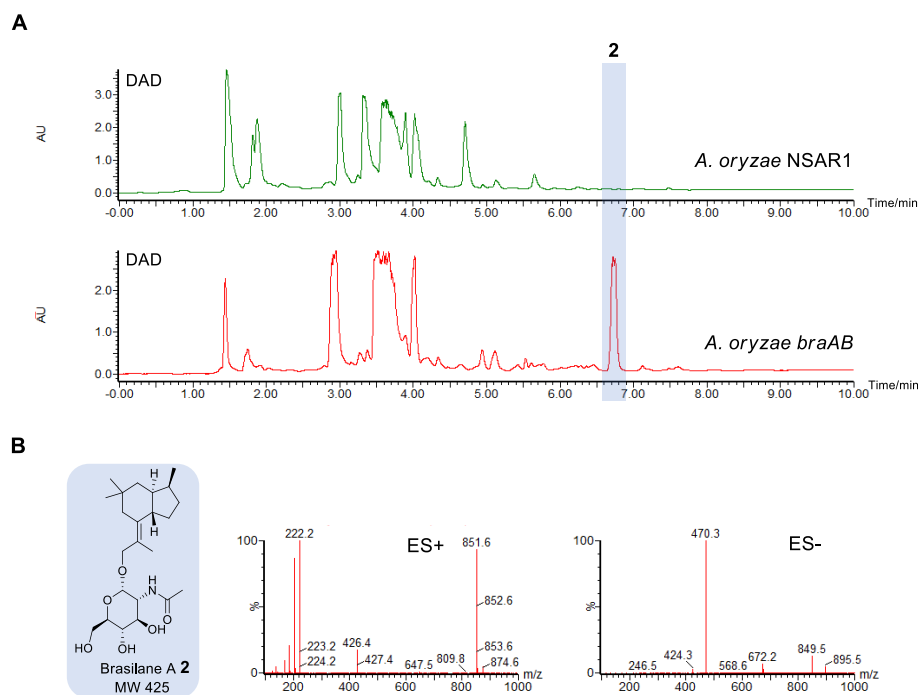


**Scheme 5.5** Construction of the plasmid pTYGS-arg-*braAB* XXXVIII.

The resulting vector was transformed into the auxotrophic host *A. oryzae* NSAR1 by using previously described methods. After two rounds of screening, 8 transformants were obtained from the selection media CZD/S and then transferred onto DPY for the following fermentation for 5 days. The culture broth of each transformant was extracted and analyzed by LCMS analysis.

The results indicated that 6 of the 8 *A. oryzae braAB* transformants produced a new peak in common at 7.1 min on the HPLC chromatogram in comparison with control *A. oryzae* NSAR1 (Fig. 5.8 A). The UV absorption of this new product fits that of brasilane A **2**. In addition,  $m/z$  426.6  $[M+H]^+$  and  $m/z$  470.4  $[M+HCOO]^-$  were observed on ES+ and ES-, respectively (Fig. 5.8 B). In order to identify its structure, this highly productive transformant was fermented in 1 L scale of DPY media. 23 mg compound was isolated by preparative HPLC. By 1D and 2D NMR comparison with data of the reference compound, this metabolite was determined to be brasilane A

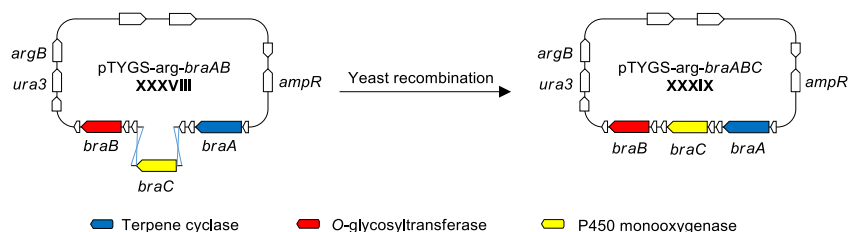
2. Therefore, the preliminary result of heterologous expression implied that these two genes *braA* and *braB* in the *bra* gene cluster are functional and responsible for brasilane A 2 production.



**Figure 5.8** LCMS analysis of the *A. oryzae braAB* transformant: **A**, HPLC comparison of control *A. oryzae* NSAR1 and transformant *A. oryzae braAB*; **B**, ES+ and ES- spectra of the product **2**.

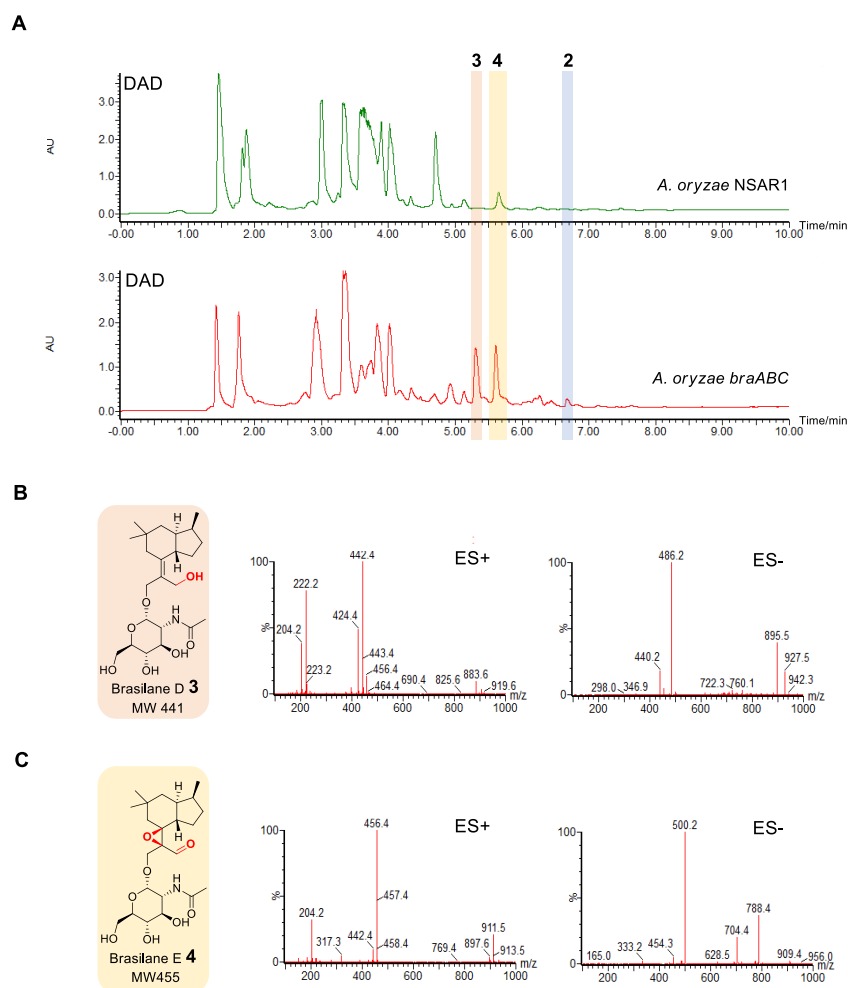
### *A. oryzae braABC*

To investigate this question whether *braC* is involved in the oxidation of brasilane A 2 to form brasilane D 3 or brasilane E 4, the additional coexpression of *braC* with *braA* and *braB* was performed in *A. oryzae*. The vector pTYGS-arg-*braAB* XXXVIII was linearized and *braC* was inserted using yeast recombination method (Scheme 5.6).<sup>63</sup> The *braC* gene was also directly amplified from the genomic DNA of *A. truncatum*.



**Scheme 5.6** Plasmid construction of pTYGS-arg-*braABC* XXXIX.

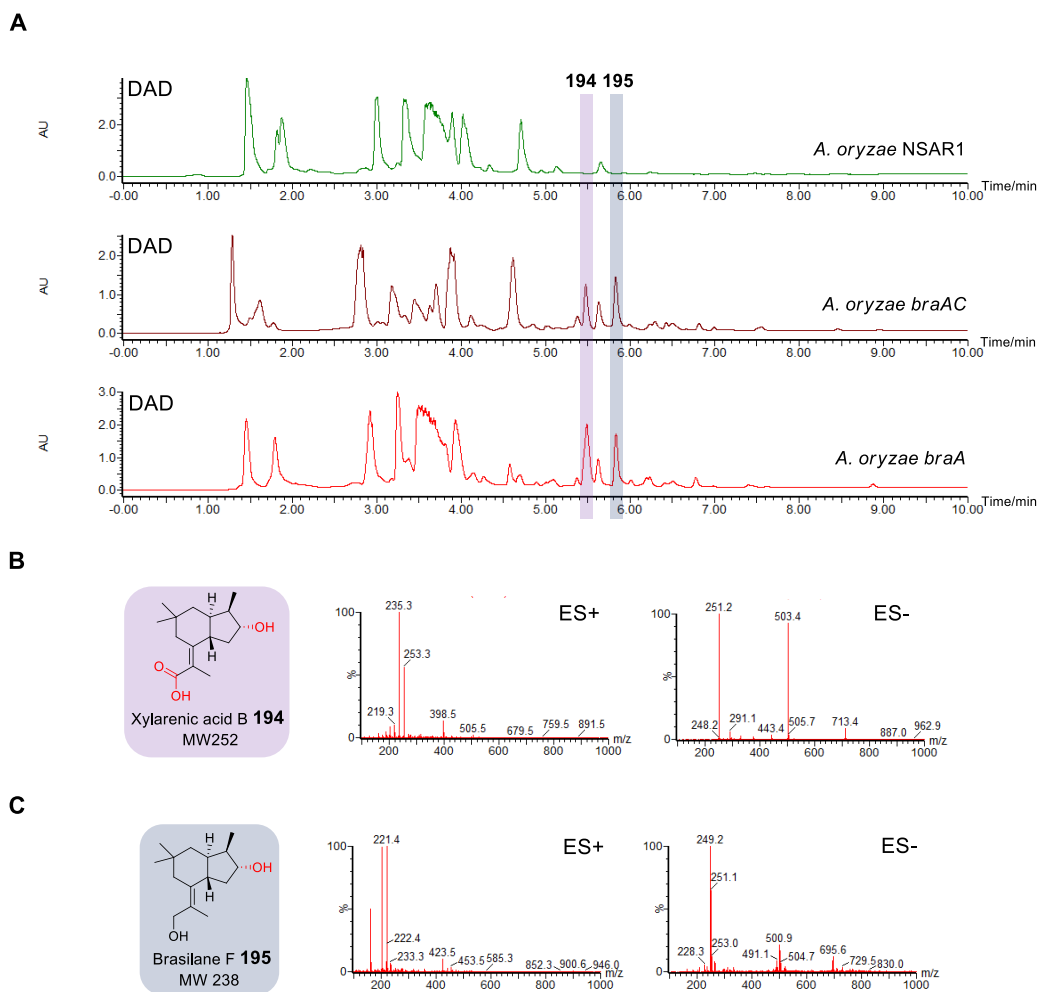
The resulting vector pTYGS-arg-*braABC* XXXIX was then transformed into *A. oryzae* NSAR1. The LCMS analysis of *A. oryzae braABC* transformants showed that two new peaks were produced between the retention time 5 min and 6 min (Fig. 5.9). Their UV and  $m/z$  on both ES+ and ES- match those of brasilane D **3** and brasilane E **4**, respectively. These two compounds were isolated and identified as brasilane D and brasilane E by NMR. Thus, the *braC* gene is functional in this pathway according to the result of the heterologous expression in *A. oryzae*. Interestingly, on the LCMS chromatogram of *A. oryzae braABC*, the titer of brasilane A **2** was apparently reduced. It is very likely that the accumulation of brasilane D **3** and brasilane E **4** led to the consumption of brasilane A **2**. In other words, the glycosylation catalyzed by BraB is supposed to occur before the oxidation of BraC in this biosynthetic pathway. Thus, it is necessary to confirm the catalytic order of these functional genes during the formation of brasilane D **3** and brasilane E **4**.



**Figure 5.9** LCMS analysis of the *A. oryzae braABC* transformant: **A**, HPLC comparison of control *A. oryzae* NSAR1 and transformant *A. oryzae braABC*; **B**, ES+ and ES- spectra of product **3**; **C**, ES+ and ES- spectra of product **4**.

***A. oryzae braAC* and *A. oryzae braA***

For the purpose of determining the biosynthetic order in this pathway, the vector pTYGS-arg-*braAC* XL was designed and heterologously expressed in *A. oryzae* NSAR1. Since the glycosyltransferase gene *braB* was absent in this construct, trichobrasilenol **171** formed by BraA would only be modified by BraC to give two different oxidized derivatives as expected, which should be consistent with the aglycone form of brasilane D **3** and brasilane E **4**.



**Figure 5.10** Production and isolation of shunt products: **A**, HPLC comparison of control *A. oryzae* NSAR1 and transformants *A. oryzae braAC* and *A. oryzae braA*; **B**, chemical structure and mass spectrum of **194**; **C**, chemical structure and mass spectrum of **195**.

However, the heterologous expression results took place unexpectedly. Two unknown compounds were observed on the LCMS chromatogram (Fig. 5.10 A). Neither of their  $m/z$  match those of desired compounds. According to the mass gap, these two new products are presumably the oxidized forms of trichobrasilenol **171**.

The emergence of these oxidation steps might arise from a native oxidase of *A. oryzae*, which played a dominant role compared to BraC.

**Table 5.1** NMR data ( $^1\text{H}$ -700 MHz,  $^{13}\text{C}$ -175 MHz) of **194** in pyridine- $d_5$ .

Atom#	C shift/ppm	H shift/ppm (J/Hz)	COSY	N/ROSEY	C to H HMBC
1	43.7, CH	2.38, m		15, 13, 2, 7, 12, 7, 8	6
2a	40.8, CH <sub>2</sub>	1.21, dd (12.8, 12.4)	2	14, 2, 6	13, 14, 3, 1, 4
2b		1.39 dd (12.4, 2.3)	2,4	14, 2, 1	13, 3, 1, 4
3	33.7, C				
4a	46.4, CH <sub>2</sub>	2.99, br d (13.8)	2	14, 13, 4, 12	13, 14, 3, 2, 6, 10
4b		2.01, s		12, 4	13
5	138.6, C				
6	46.3, CH	2.15, m	7	2, 8	7
7a	41.8, CH <sub>2</sub>	2.15, m	7, 8	15, 1, 8	6, 8, 5
7b		2.73, m	6, 7, 8	15, 9, 12, 1, 8	9, 6, 8
8	80.7, CH	4.08, m	7, 7	15, 6, 7, 1, 7	15
9	43.1, CH	2.18, m	15	7	15, 6
10	126.7, C				
11	175.8, C				
12	17.0, CH <sub>3</sub>	2.24, br s		4, 1, 7, 4	10, 11
13	26.7, CH <sub>3</sub>	1.09, s		14, 1, 4	14, 3, 2
14	32.7, CH <sub>3</sub>	0.97, d (8.0)		13, 2, 2, 4	13, 3, 2, 4
15	15.1, CH <sub>3</sub>	0.98, s	9	7, 1, 7, 8	9, 8

**Table 5.2** NMR data ( $^1\text{H}$ -700 MHz,  $^{13}\text{C}$ -175 MHz) of **195** in pyridine- $d_5$ .

Atom#	C shift/ppm	H shift/ppm (J/Hz)	COSY	N/ROSEY	C to H HMBC
1	44.6, CH	2.05, m	2, 2	8	
2a	41.5, CH <sub>2</sub>	1.37, m	2, 1, 6, 4	2	
2b		1.19, br t (12.5)	2, 1	2	14, 13, 3, 1, 6
3	34.0, C				
4a	44.7, CH <sub>2</sub>	2.44, dd (13.9, 1.6)	2, 4	14, 13, 4, 11, 11	3, 2, 6, 10, 5
4b		1.60, br d (13.9)	14, 12, 4	6, 4	
5	136.9, C				
6	47.3, CH	2.05, m	2, 7, 7	4, 7	
7a	42.1, CH <sub>2</sub>	1.70, m	6, 7, 8	7	6, 8
7b		2.55, ddd (11.9, 6.9, 5.1)	7, 6, 8	7, 12, 6, 8	1
8	81.7, CH	3.70, ddd (7.1, 6.9, 3.2)	7, 9, 7	15, 1, 7	
9	43.1, CH	1.77, m	15, 8		
10	127.8, C				
11a	64.8, CH <sub>2</sub>	3.86, d (11.6)	11	12, 4, 11	10, 5
11b		4.21, d (11.6)	11	12, 4, 11	10, 5
12	17.6, CH <sub>3</sub>	1.86, s	4	7, 11, 11	11, 10, 5
13	32.6, CH <sub>3</sub>	0.97, s		4	14, 3, 2, 4
14	26.4, CH <sub>3</sub>	0.88, s		4	13, 3, 2, 4
15	14.9, CH <sub>3</sub>	0.90, d (7.53)	9	8	9, 1, 8

To verify this hypothesis, the terpene cyclase BraA was individually expressed in *A. oryzae*. The result indicated that the BraA product trichobrasilenol **171** could not be detected by either LCMS or GCMS. However, two same major products were

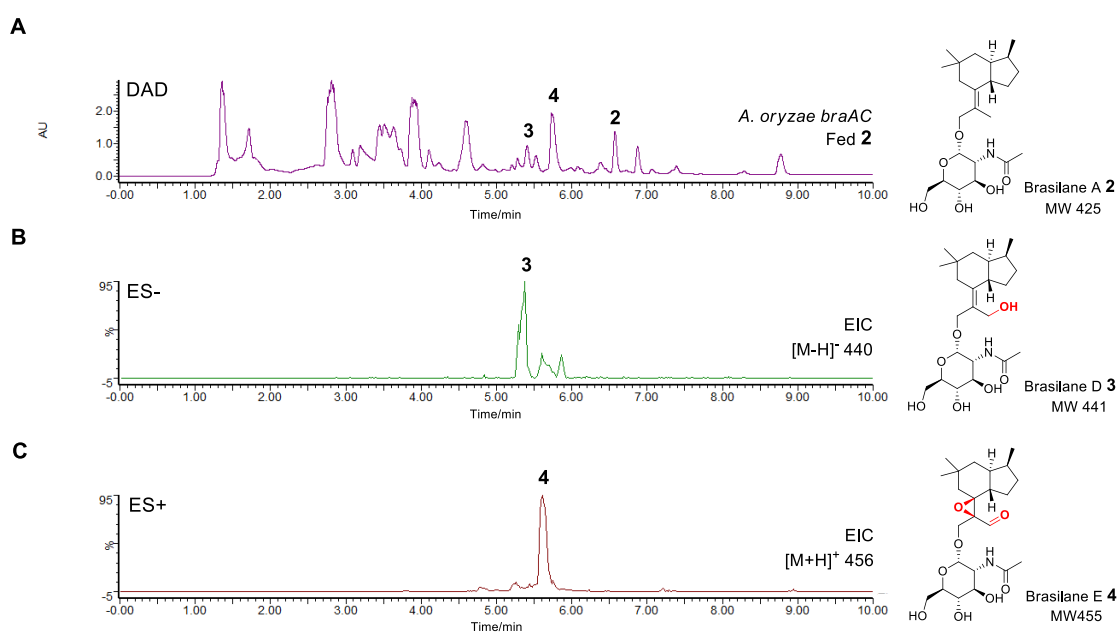


presented on the LCMS chromatograms as well (Fig. 5.10 A). These products were purified, and elucidation of their structures showed them to be two new oxygenated brasilanes designated as xylarenic acid B **194** and brasilane F **195** (Fig. 5.10 B and Table 5.1 and 5.2).

Against the expectation neither **194** nor **195** were oxygenated at C-12, C-5 or C-10. Instead, both molecules display a hydroxyl group at C-8, and **194** has an additional carboxylic acid functionality at C-10, which is similar to xylarenic acid **161**.<sup>167</sup> While ROESY correlations and a series of 1D NOE NMR experiments confirmed the relative configuration of the new molecules including the *trans* configuration of 1-H/6-H, their absolute stereochemistry was determined by derivatization of 8-OH with Mosher's acid (MTPA). Therefore, based on these results above, BraC is proposed to be functional *in vivo* only when brasilane A **2** is supplied as substrate. In the absence of BraB, the role of BraC turned to be silent in *A. oryzae*.

### Feeding Experiment of *A. oryzae braAC*

To examine the substrate selectivity of BraC, a feeding experiment was conducted with the transformant *A. oryzae braAC*. 1.5 mg of brasilane A **2** dissolved in 100  $\mu$ L methanol were fed into the two-day culture of *A. oryzae braAC*. After two more days of cultivation, the culture was extracted and analyzed by LCMS.



**Figure 5.11** Feeding experiment of transformant *A. oryzae braAC*: **A**, HPLC chromatogram of *A. oryzae braAC* fed with brasilane A **2**; **B**, extracted ion chromatogram (EIC) of brasilane D **3**; **C**, EIC chromatogram of brasilane E **4**.

As shown in Fig. 5.11, a certain amount of brasilane A **2** was converted to two oxidized products brasilane D **3** and brasilane E **4** in the presence of BraC. It implied that BraC is capable of specifically recognizing the glycoside brasilane A **2** as substrate. Its substrate selectivity determines the direction of the *bra* biosynthetic pathway.

### 5.3.4 Protein Characterization *In Vitro*

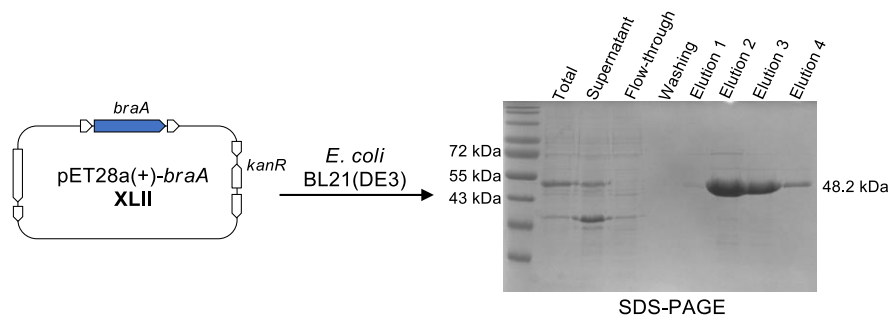
Although the product of BraA in *A. oryzae* could neither be unambiguously detected by LCMS nor GCMS, the terpene alcohol trichobrasilenol **171** is still the most likely product of BraA. On the one hand, the aglycone form of all isolated brasilanes in this study shows the same absolute configuration as reported by Murai *et al.*<sup>171</sup> On the other hand, the protein sequence of BraA also shows high similarity to that of TaTC6. Therefore, to further confirm this prediction, the protein characterization *in vitro* was adopted as an alternative method. Besides, BraB, as a rare *N*-acetylglucosamine transferase in biosynthesis of natural products, is worth to deeply explore by enzymatic assay *in vitro* as well. Thus, in this part, BraA and BraB were cloned and overexpressed in the host *E. coli* followed by a series of *in vitro* assays using purified enzymes.

#### 5.3.4.1 BraA Expression and *In Vitro* Assay

##### Protein Expression and Purification

Given the presence of introns in *braA* and the relatively low transcription level of *braA* in *A. truncatum* compared with in the heterologous host *A. oryzae* under the strong promoter, the total RNA was extracted from the freshly cultivated mycelia of *A. oryzae braA* and reverse-transcribed into cDNA. The cDNA sequence of *braA* was amplified by PCR. The obtained *braA* fragment flanked by two overlaps was inserted into the MCS region of pET-28a(+) by Quick-change method.<sup>191</sup> The resulting expression vector pET28a(+)-*braA* **XLII** was introduced into *E. coli* BL21(DE3) for overexpression. After protein isolation and purification by Ni-NTA affinity chromatography, a distinct and soluble protein band was displayed on SDS-PAGE

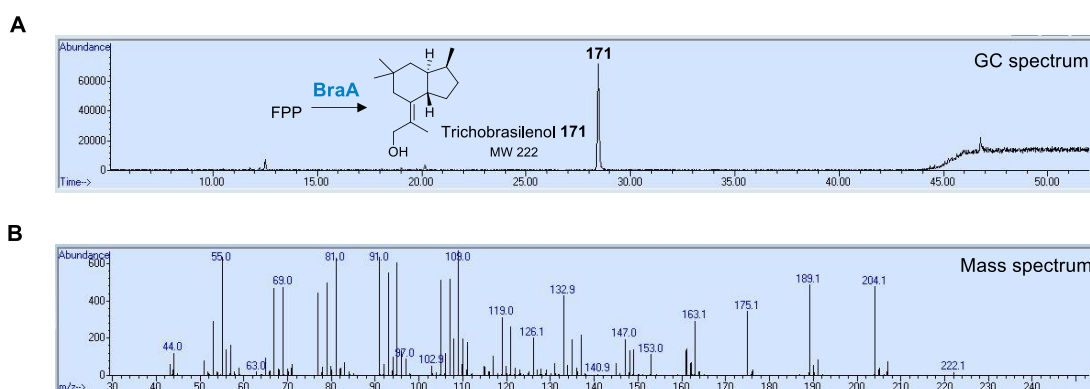
(Scheme 5.7). Its real size is consistent with the predicted 48.2 kDa. It was stored in tris buffer (pH 7.5) for the following *in vitro* assays.



**Scheme 5.7** Plasmid construction and protein purification result of BraA on SDS-PAGE.

### In Vitro Assay

To identify the product of BraA, the candidate enzyme was examined in 300  $\mu\text{L}$  reaction volume, which contains tris/HCl buffer (pH 7.5), 150  $\mu\text{g}$  BraA, 500  $\mu\text{M}$  FPP, 5 mM  $\text{MgCl}_2$  and 5 mM dithiothreitol (DTT). After incubation at 28  $^\circ\text{C}$  for 30 min, the solution was extracted with pentane. The organic phase was then analyzed by GCMS. A noticeable and single peak was shown on the GC chromatogram (Fig. 5.12). According to the mass spectrum, it revealed that the molecular weight of this product is 222.1, that is in accordance with the molecular weight of trichobrasilenol **171**. The respective fragmentation pattern ( $m/z$  204.1, 189.1 and 175.1) corresponds well with those previously reported for trichobrasilenol. Based on these results above, it is confirmed that BraA is responsible for the production of trichobrasilenol **171** in *A. truncatum*.

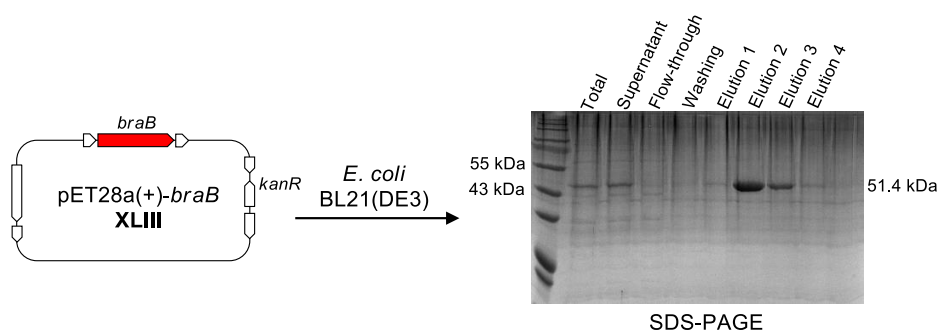


**Figure 5.12** Formation and identification of trichobrasilenol **171**. **A**, GC chromatogram of BraA *in vitro* characterization with the substrate FPP; **B**, Mass spectrum of the product **171**.

### 5.3.4.2 BraB Expression and *In Vitro* Assay

#### Protein Expression and Purification

For BraB protein expression and characterization *in vitro*, the cDNA sequence of BraB was codon-optimized and synthesized commercially based on the use of *E. coli* host. The optimized sequence was inserted downstream of the *T7* promoter in the plasmid pET-28a(+). After transformation into *E. coli* BL21(DE3), the protein purification process was conducted as described before. The obtained fractions were analyzed by SDS-PAGE. A pure protein band was displayed on SDS-PAGE located between 43 kDa and 55 kDa (Scheme 5.8). Its real molecular weight precisely matches the predicted size (51.4 kDa).

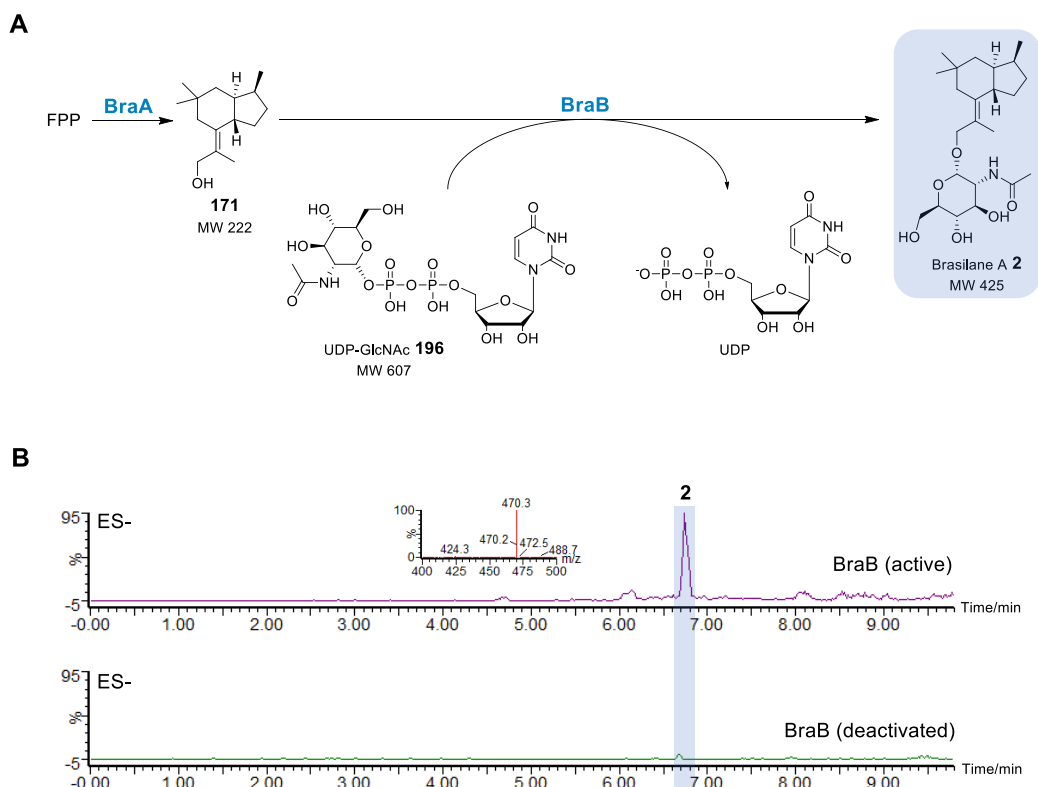


**Scheme 5.8** Plasmid construction and protein purification result of BraB on SDS-PAGE.

#### Trichobrasilenol as substrate

To test the activity of BraB, trichobrasilenol **171** as natural substrate should be prioritized to carry out *in vitro* assay. However, trichobrasilenol could not be produced and isolated individually from the heterologous host *A. oryzae* due to additional oxidative modifications. Meanwhile, this volatile compound was also difficult to sufficiently prepare by *in vitro* reaction. Therefore, a two-enzyme reaction containing both BraA and BraB was adopted (Fig. 5.13 A). In addition to components required for BraA reaction, the sugar donor UDP-*N*-acetylglucosamine (UDP-GlcNAc **196**, 2 mM) and 200  $\mu$ g candidate enzyme BraB were added into the mixture. After incubation at 37 °C overnight, the solution was extracted with ethyl acetate. The organic phase was analyzed by LCMS. For the negative control, the candidate enzyme was deactivated by heat prior to *in vitro* reaction. As shown in Fig. 5.13 B, the ionized peak representing brasilane A **2** with  $m/z$  470.3 [ $M + HCOO$ ]<sup>-</sup> was

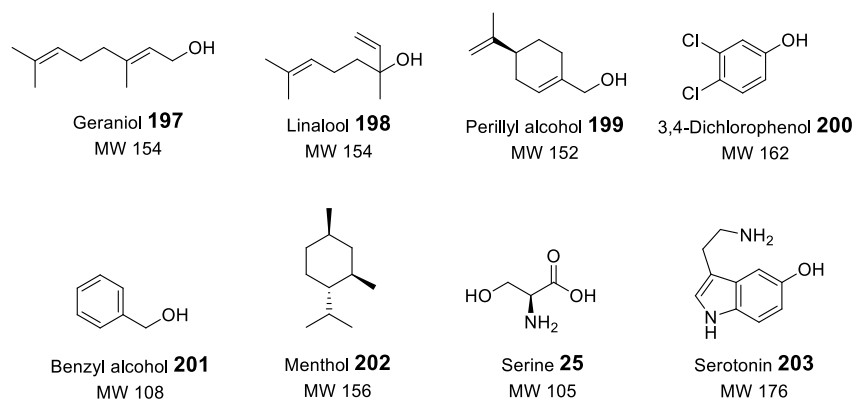
observed compared with the negative control on the ES- mode of LCMS. It suggested that BraB is active and can recognize trichobrasilenol **171** as a substrate *in vitro*.



**Figure 5.13** Two-enzyme assays using BraA and BraB: **A**, hypothetical reaction route; **B**, ES- TIC spectra of the experimental reaction using active BraB and the negative control with deactivated BraB.

## Other acceptors

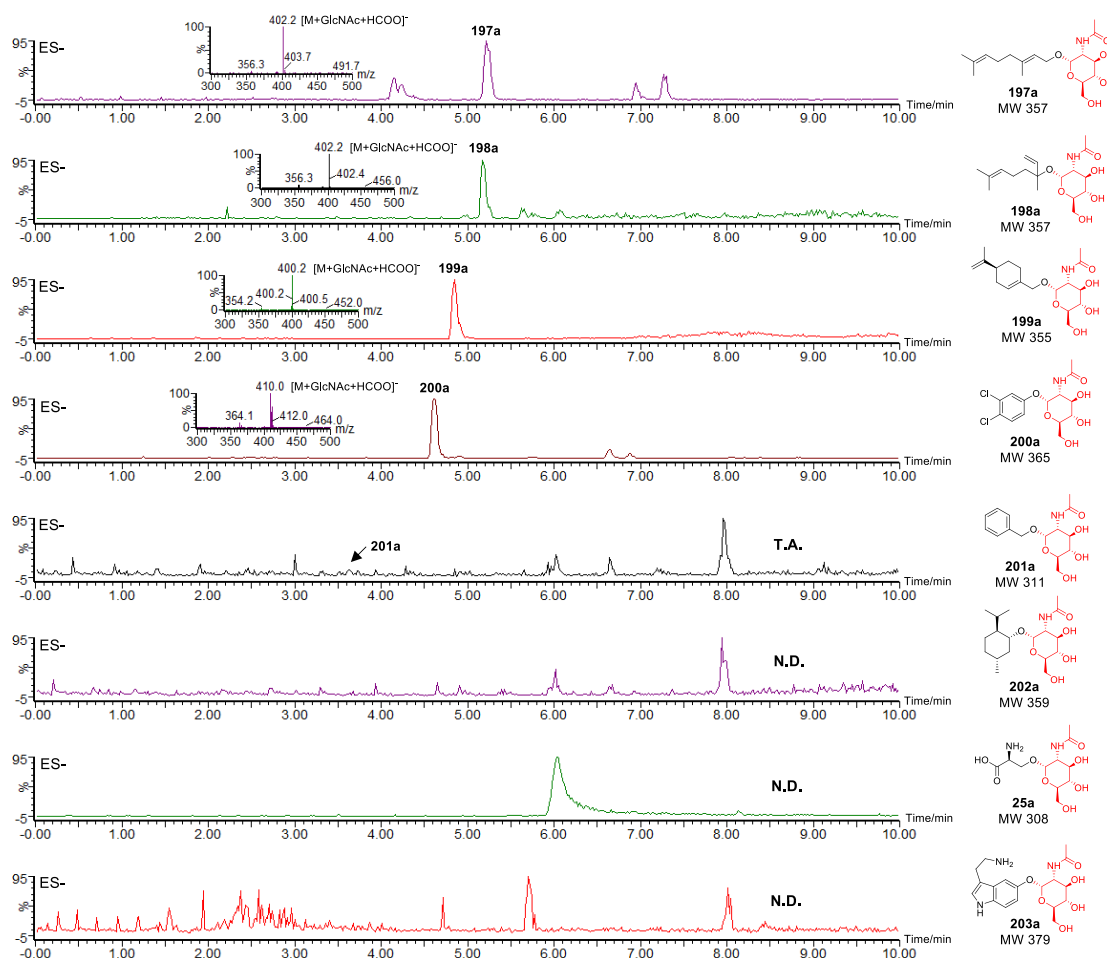
Using trichobrasilenol **171** as the natural acceptor, the glycosylation role of BraB has been confirmed *in vitro*. To further explore the substrate tolerance of BraB, more acceptor substrates were examined. As shown in Fig. 5.14, they structurally belong to primary, secondary and tertiary alcohols. Two aromatic alcohols are also included.



**Figure 5.14** Structures of all candidate acceptors for BraB *in vitro* assay.

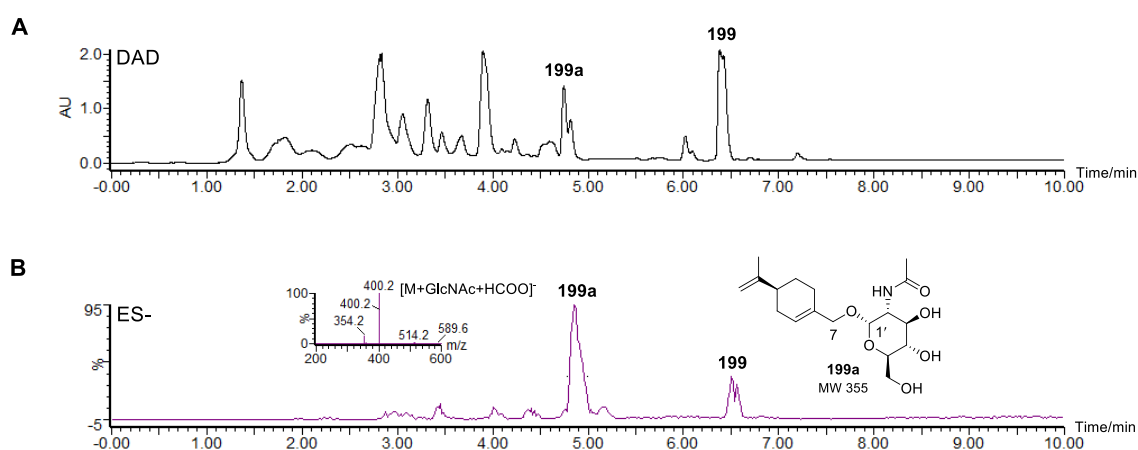
Herein, all candidate acceptors were tested with BraB in the presence of UDP-GlcNAc **196**. All reactions were conducted using 2 mM UDP-GlcNAc **196**, 200  $\mu$ g BraB, 0.5 mM substrate (geraniol **197**, linalool **198**, perillyl alcohol **199**, 3,4-dichlorophenol **200**, benzyl alcohol **201**, menthol **202**, serine **25**, or serotonin **203**) in CutSmarter Buffer (NEB). After incubation at 37 °C overnight, the solution was extracted with ethyl acetate. The organic phase was analyzed by LCMS.

The results suggested that substrates **197-200** could be converted to their respective glycosides **197a-200a** with high yield (Fig. 5.15). For example, the conversion ratio of perillyl alcohol **199** could reach up to 100 % in 1 h, indicating that BraB is a highly efficient catalyst for glycosylation of natural products. The glycosylated product **201a** of substrate **201** was detected only in trace amounts. Base on the overall results of all acceptors, it can be seen that BraB is able to glycosylate a broad scope of acceptor substrates.



**Figure 5.15** ES- spectra of LCMS chromatograms of acceptors by BraB *in vitro*. The small ES- spectrum shows the molecular weight of the respective product peak. Abbreviations: T.A., detected in trace amounts; N.D., not detected.

To demonstrate that the *in vitro* products are in agreement with the prediction and that BraB can be used for biotechnological applications, *in vivo* preparative scale reactions were conducted. 1 L of *E. coli* transformant harboring the plasmid pET28a(+)-*braB* **XLIII** were cultivated overnight and then 75 mg of perillyl alcohol **199** were fed into the culture broth. After 1 day of biotransformation, it resulted in the formation of the respective glycoside **199a** on LCMS chromatogram of the *E. coli* crude extract (Fig. 5.16). Subsequent isolation and purification led to yield 10 mg of pure compound. HMBC correlations from 7-H to C'-1 and 1'-H to C-7, respectively, confirmed the metabolite as a perillyl glycoside with a *N*-acetylglucosamine moiety.

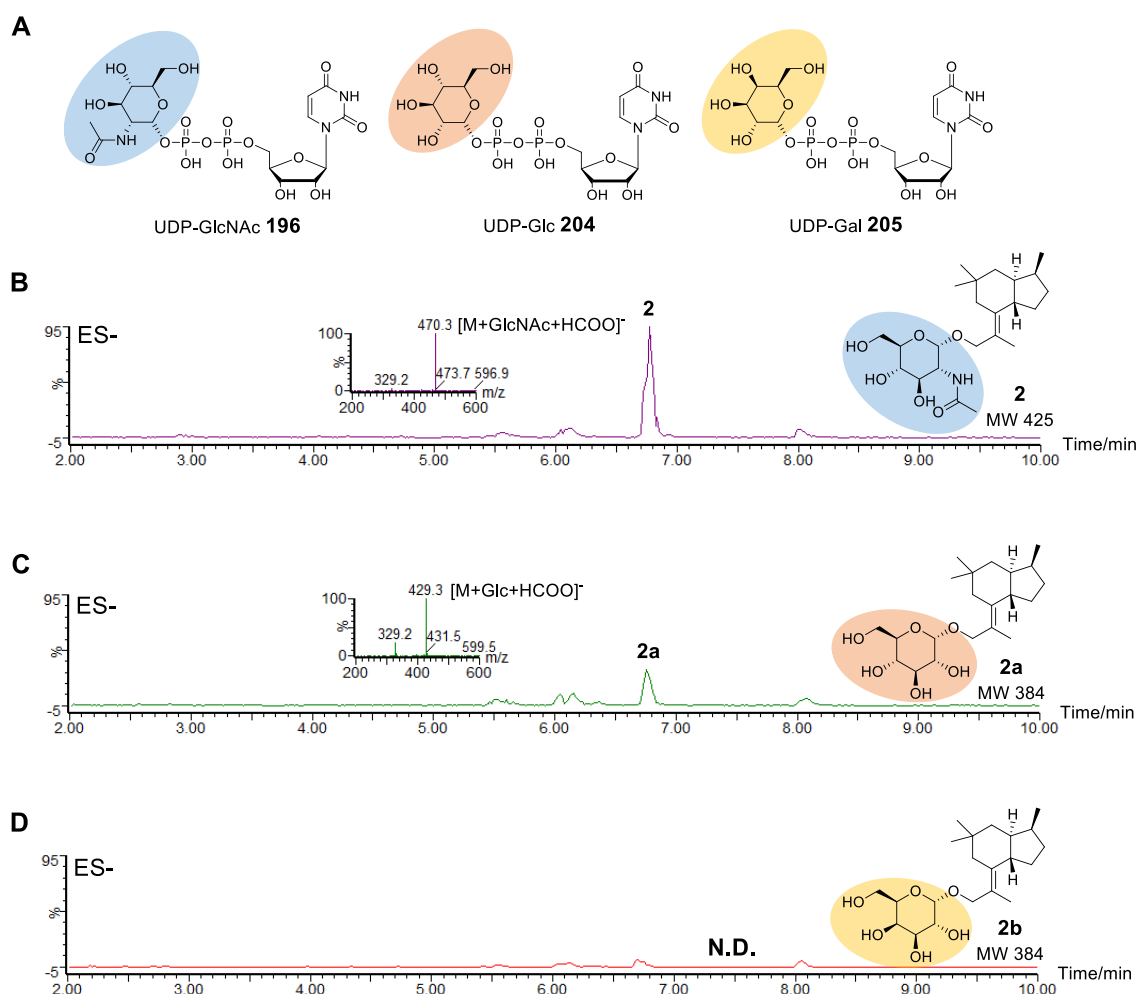


**Figure 5.16** Biotransformation of perillyl alcohol **199** by BraB within *E. coli* transformant *in vivo*: **A**, DAD chromatogram of crude extract; **B**, ES- TIC chromatogram of crude extract.

### Different UDP-sugars as donor

To test the UDP-sugar selectivity of BraB, the enzyme was incubated with different UDP-sugars such as UDP-GlcNAc **196**, UDP-Glc **204**, and UDP-Gal **205** (Fig. 5.17 A). Here, two-enzyme reactions using BraA and BraB were chosen due to lacking pure trichobrasilenol **171** as substrate. The enzymatic reactions were conducted under the same conditions in parallel. The results suggested that products with *N*-acetylglucosamine moiety and glucose moiety are detectable on LCMS chromatograms (Fig. 5.17 B and C), respectively. The substrate reacted with UDP-GlcNAc **196** has the higher conversion compared with that of UDP-Glc **204**. Therefore, this enzyme is defined as an *N*-acetylglucosamine transferase according to its strong preference towards UDP-GlcNAc **196**. This also explains why brasilane

glycosides with *N*-acetylglucosamine moiety are the predominant pathway products in *A. truncatum* and the heterologous host *A. oryzae*.



**Figure 5.17** Glycosylation of different donors by BraB: **A**, structures of different UDP-sugars **196**, **204** and **205**; **B**, ES- TIC chromatogram of enzymatic assays by BraB with UDP-GlcNAc **196** as donor; **C**, ES- TIC chromatogram of enzymatic assays by BraB with UDP-Glc **204** as donor; **D**, ES- TIC chromatogram of enzymatic assays by BraB with UDP-Gal **205** as donor. Abbreviation: N.D., not detected.

## 5.4 Discussion and Conclusion

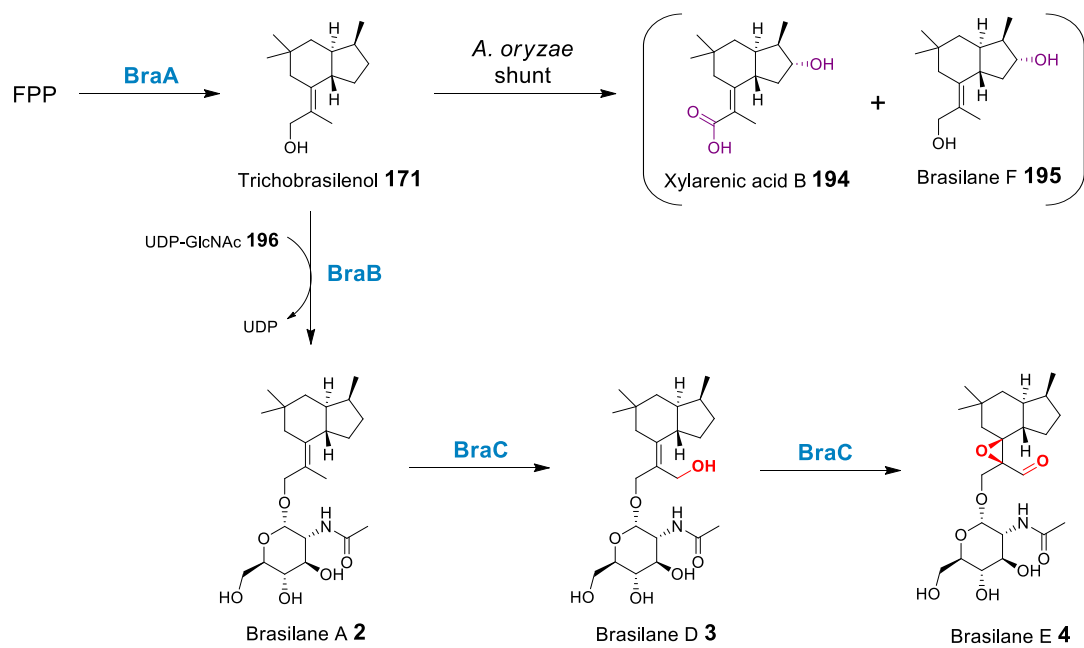
The aim of this work was to identify the biosynthetic pathway of brasilane-type sesquiterpenes in fungi. Due to the unique structure and various bioactivities, many brasilane-type sesquiterpenes have been discovered, but not deeply investigated on biosynthetic pathway. Especially, the biosynthesis of several brasilane glycosides might lead to the discovery of novel fungal tailoring enzymes.



In this work, a putative gene cluster (*bra*) for the production of brasilanes in *A. truncatum* was proposed based on bioinformatic analyses (Scheme 5.9). According to the alignment with other putative fungal gene clusters, it revealed the common presence of three core genes encoding: a terpene cyclase; an *O*-glycosyltransferase; and a cytochrome P450 monooxygenase. Subsequently, the fermentation of *A. truncatum* validated the production of brasilanes. Three brasilane glycosides were isolated and characterized spectroscopically. Therefore, linking the production of brasilanes to the *bra* gene cluster was the next objective to achieve.

Here, the well-established heterologous expression approach in *A. oryzae* was adopted. Through a series of heterologous expressions, it suggested that BraA is responsible for the formation of trichobrasilenol **171**, followed by glycosylation modification of BraB to yield brasilane A **2** both in *A. truncatum* and in *A. oryzae*. The further oxidative processes of brasilane A were achieved in successive two steps by BraC to form brasilane D **3** and brasilane E **4** (Scheme 5.9). However, in the case where BraA and BraC expressed in the heterologous host, two shunt products were produced unexpectedly. Based on these evidences, it can be deduced that in the heterologous host *A. oryzae* BraC only recognizes brasilane A **2** as substrate specifically. In other words, the glycosylation step of BraB is supposed to be prior to the oxidative steps of BraC *in vivo*. In addition, the enzymatic reactions of BraA and BraB also confirmed their functions *in vitro*, respectively. More interestingly, BraB, defined as a rare *N*-acetylglucosamine transferase, also could recognize multiple UDP-sugars and a broad scope of acceptor substrates *in vitro*.

In the future, the identification of the biosynthetic gene cluster *bra* will provide insights into the investigation of other brasilane gene clusters. Meanwhile, the potential of this *N*-acetylglucosamine transferase is worth continuing to explore on substrate promiscuity. Its broad substrate selectivity should offer significant opportunities for use in pathway engineering and other biotechnological applications.



**Scheme 5.9** Brasilane biosynthetic gene cluster (*bra*) from *A. truncatum* and biosynthetic pathway for the production of **2**, **3**, and **4** based on experimental evidence.

## 6. Experimental

General biological reagents were purchased from ThermoFisher Scientific (Waltham, MA USA) and Sigma-Aldrich (Germany). All of restriction enzymes were purchased from New England Biolabs (Massachusetts, USA). The synthetic genes involved in this work were produced by Synbio Technologies (USA) and BaseClear B.V. (The Netherlands). All comparison and analysis work of gene sequences were performed with the software Geneious. All enzymes were used as the standard methods and all molecular biological procedures were performed according to the standard protocols supplied in kits. Analytical grade chemicals and reagents were purchased from Sigma, Roth and Fischer. The solvents used for HPLC were analytical grade.

### 6.1 Media, Buffers and Solutions

#### 6.1.1 Media

**Table 6.1** Media used in the thesis.

Media	Composition (%(v/w))	Ingredients
GN	2.00	D(+)-Glucose Monohydrate
	1.00	Nutrient broth Nr. 2 from Oxoid
LB	0.50	Yeast extract
	1.00	Tryptone
	0.50	Sodium chloride
YPAD	1.00	Yeast extract
	2.00	Tryptone
	2.00	D(+)-Glucose Monohydrate
SM-URA	0.03	Adenine
	0.17	Yeast nitrogen base
	0.50	Ammonium sulfate
	2.00	D(+)-Glucose Monohydrate
	0.077	Complete supplement mixture minus Uracil
CZD/S	3.50	Czapek Dox broth
	18.22	D-Sorbitol
	0.10	Ammonium sulfate
	0.05	Adenine
	0.15	L-Methionine
CZD/S+A	3.50	Czapek Dox broth
	18.22	D-Sorbitol
	0.10	Ammonium sulfate
	0.10	Arginine
	0.05	Adenine
CZD/S w/o methionine	0.15	L-Methionine
	3.50	Czapek Dox broth
	18.22	D-Sorbitol
	0.10	Ammonium sulfate

Media	Composition ( %(v/w) )	Ingredients
DPY	0.05	Adenine
	2.00	Dextrin from potato starch
	1.00	Polypeptone
	0.50	Yeast extract
	0.50	Monopotassium phosphate
YMG	0.50	Magnesium sulfate hexahydrate
	0.40	D(+)-Glucose Monohydrate
	0.40	Yeast extract
	1.00	Malt extract

## 6.1.2 Buffers and Solutions

**Table 6.2** Buffers and solutions used in the thesis.

Buffers/Solutions	Concentration	Ingredients
Amp/Carb	50 mg·mL <sup>-1</sup>	Ampicillin/Carbenicillin
Kan	50 mg·mL <sup>-1</sup>	Kanamycin
PEG solution	50 % (w/v)	Polyethylene glycol 3350
LiOAc	1 M	Lithium acetate
Salmon testis DNA solution	2 mg·mL <sup>-1</sup>	Salmon testis DNA
Lysing solution for <i>A. oryzae</i>	10 mg·mL <sup>-1</sup>	<i>Trichoderma</i> lysing enzyme
Solution I (pH7.5)	0.8 M	NaCl
	0.8 M	NaCl
	10 mM	CaCl <sub>2</sub>
	50 mM	Tris-HCl, 7.5
Solution II (pH7.5)	60 % (w/v)	Polyethylene glycol 3350
	0.8 M	NaCl
	10 mM	CaCl <sub>2</sub>
	50 mM	Tris-HCl, 7.5
Lysis buffer	50 mM	Tris/HCl, pH 8.0
	150 mM	NaCl
	10 % (w/v)	Glycerol
Elution buffer	20 mM	Imidazole
	50 mM	Tris/HCl, pH 8.0
	150 mM	NaCl
Storage buffer	10 % (w/v)	Glycerol
	500 mM	Imidazole
	50 mM	Tris/HCl, pH 7.5
	20 % (w/v)	Glycerol

## 6.2 Strains, Plasmids and Primers

### 6.2.1 Strains

One Shot™ Top10 chemically competent *Escherichia coli* and One Shot™ *ccdB* Survival™ 2 T1<sup>R</sup> competent *E. coli* were used as hosts for the construction of general plasmids. *Saccharomyces cerevisiae* strain CEN.PK was used as the host for the

expression plasmid assembly by homologous recombination. *Aspergillus oryzae* NSAR1 was used as the heterologous host for fungal transformation and metabolite production.

**Table 6.3** Strains used in the thesis.

Strain	Genotype	Phylum	Origin
<i>E. coli</i> One Shot Top10	<i>F mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (Str<sup>R</sup>) endA1 nupG</i>	Proteobacteria	Thermo Fisher Scientific
<i>E. coli</i> One Shot <i>ccdB</i> Survival 2 T1 <sup>R</sup>	<i>F mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araΔ139 Δ(ara-leu)7697 galU galK rpsL (Str<sup>R</sup>) endA1 nupG fhuA::IS2</i>	Proteobacteria	Thermo Fisher Scientific
<i>E. coli</i> BL21 (DE3)	<i>F ompT hsdSB (rB<sup>-</sup>, mB<sup>-</sup>) gal dcm (DE3)</i>	Proteobacteria	Thermo Fisher Scientific
<i>S. cerevisiae</i> CEN.PK	<i>MATa/α ura3-52/ura3-52 trp1-289/trp1-289 leu2-3_112/leu2-3_112his3 Δ1/his3 Δ1 MAL2-8C/MAL2-8C SUC2/SUC2</i>	Ascomycetes	Hahn group, Hannover
<i>A. oryzae</i> NSAR1	<i>argB<sup>-</sup> sC ΔadeA niaD<sup>-</sup></i>	Ascomycetes	Lazarus group, Bristol
<i>Annulohyphoxylon truncatum</i> CBS 1407781	Wildtype	Ascomycetes	Dr. Eric Kuhnert

## 6.2.2 Plasmids

All plasmids used, constructed or synthesized commercially were included in the Table 6.4 below. Plasmids were generally stored in the glycerol at -80 °C.

**Table 6.4** Plasmids used and constructed in this study.

No.	Name (inserted genes)	Marker genes	Expression host
I	pTYGS-arg	<i>ampR, ura3, argB</i>	-
II	pTYGS-arg- <i>pccABE</i>	<i>ampR, ura3, argB</i>	-
III	pUC57- <i>DEBS1-F1</i>	<i>ampR</i>	-
IV	pUC57- <i>DEBS1-F2</i>	<i>ampR</i>	-
V	pUC57- <i>DEBS1-F3</i>	<i>ampR</i>	-
VI	pUC57- <i>DEBS1-F4</i>	<i>ampR</i>	-
VII	pE-YA	<i>kanR, ura3</i>	-
VIII	pEYA- <i>DEBS1-TE</i>	<i>kanR, ura3</i>	-
IX	pTYGS-arg- <i>DEBS1-TE-pccABE</i>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
X	pEYA- <i>DEBS1-TE-eGFP</i>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XI	pTYGS-arg- <i>DEBS1-TE-eGFP</i>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XII	pTYGS-arg- <i>pccA-eGFP</i>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1

No.	Name (inserted genes)	Marker genes	Expression host
XIII	pTYGS-arg- <u>pccB-eGFP</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XIV	pTYGS-arg- <u>pccE-eGFP</u>	<i>ampR, ura3, sC</i>	-
XV	pTYGS-met	<i>ampR, ura3, sC</i>	<i>A. oryzae</i> NSAR1
XVI	pTYGS-met- <u>sePptII</u>	<i>ampR, ura3, sC</i>	<i>A. oryzae</i> NSAR1
XVII	pTYGS-met- <u>sePptII</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XVIII	pTYGS-arg- <u>DEBS1-TE-eGFP-pccABE</u>	<i>ampR, ura3, sC</i>	<i>A. oryzae</i> NSAR1
XIX	pTYGS-met- <u>sePptII-pcsA</u>	<i>kanR, ura3</i>	-
XX	pEYA- <u>ble-mcsA</u>	<i>kanR, ura3</i>	-
XXI	pEYA- <u>argB-mcsA</u>	<i>kanR, ura3</i>	-
XXII	pEYA- <u>adeA-coaT</u>	<i>kanR, ura3</i>	-
XXIII	pEYA- <u>DEBS1-ACP<sub>sorA</sub></u>	<i>kanR, ura3</i>	-
XXIV	pEYA- <u>DEBS1-ACP<sub>sorA</sub>-sorB</u>	<i>kanR, ura3</i>	-
XXV	pTYGS-arg- <u>DEBS1-ACP<sub>sorA</sub>-sorB-pccABE</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XXVI	pTYGS-met- <u>sePptII*</u>	<i>ampR, ura3, sC</i>	<i>A. oryzae</i> NSAR1
XXVII	pTYGS-arg- <u>DEBS1-ACP<sub>sorA</sub>-pccABE</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XXVIII	pEYA- <u>sorB</u>	<i>kanR, ura3</i>	-
XXIX	pTYGS-met- <u>sorB-sePptII*</u>	<i>ampR, ura3, sC</i>	<i>A. oryzae</i> NSAR1
XXX	pEYA- <u>DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub></u>	<i>kanR, ura3</i>	-
XXXI	pEYA- <u>DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-sorB</u>	<i>kanR, ura3</i>	-
XXXII	pTYGS-arg- <u>DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-sorB-pccABE</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XXXIII	pTYGS-arg- <u>DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-pccABE</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XXXIV	pEYA- <u>DEBS1-ACP<sub>sorA</sub>-sorB-eGFP</u>	<i>kanR, ura3</i>	-
XXXV	pTYGS-arg- <u>DEBS1-ACP<sub>sorA</sub>-sorB-eGFP-pccABE</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XXXVI	pEYA- <u>DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-sorB-eGFP</u>	<i>kanR, ura3</i>	-
XXXVII	pTYGS-arg- <u>DEBS1<sub>w/oACP2</sub>-ACP<sub>sorA</sub>-sorB-eGFP-pccABE</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XXXVIII	pTYGS-arg- <u>braAB</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XXXIX	pTYGS-arg- <u>braABC</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XL	pTYGS-arg- <u>braAC</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XLI	pTYGS-arg- <u>braA</u>	<i>ampR, ura3, argB</i>	<i>A. oryzae</i> NSAR1
XLII	pET28a(+)- <u>braA</u>	<i>kanR</i>	<i>E. coli</i> BL21(DE3)
XLIII	pET28a(+)- <u>braB</u>	<i>kanR</i>	<i>E. coli</i> BL21(DE3)

### 6.2.3 Primers

Primers were commercially synthesized in the form of lyophilized powder by Sigma Aldrich. The synthetic concentration was 100  $\mu$ M. Before use, each was diluted to 10  $\mu$ M with deionized H<sub>2</sub>O for PCR reaction. The stock solutions of primers were stored at -20 °C. Primers used in the thesis were included in the table 6.5 below.

**Table 6.5** Primers used in the thesis.

No.	Sequence (5'-3')	Information
JF4F	CTGCGCTTCTGATATCGGGT	<i>pcsA</i> test primer
JF4R	GTCCCTGTATCATGCCTCCG	<i>pcsA</i> test primer
JF25F	CCCGATTGTTTAAATGATGACGAG	<i>pccA</i> test primer
JF20R	CCAATCGGTGACAGGCTCGAA	DEBS1-TE test primer
JF22F	CGATCGCGCCATCCTCCTCG	DEBS1-TE test primer
JF19F	ATACCCCTGCCACTCCTACC	<i>pccA</i> test primer
JF23R	CTCCTCGACGAGCTTCTGATGG	<i>pccA</i> test primer
JF24R	ACGACGAGGGAGATCTGAGG	<i>pccB</i> test primer

No.	Sequence (5'-3')	Information
JF6F1	CCATGCCCATGCGAACCACTC	<i>mcsA</i> test primer
JF6R	GGTAGAAGAGAACCACAGACGC	<i>mcsA</i> test primer
JF26R	GGTGGGACTTCTCGGCAATT	<i>mcsA</i> test primer
JF18F	GGCGGCCTTCTACCTCAAG	<i>coaT</i> test primer
JF18R	GAGGAGACAGACCACGCACG	<i>coaT</i> test primer
JF27R	GCACCAGGGATGATACCACCAT	<i>coaT</i> test primer
JF35F	ATGGCCATGCCCATGCGAAC	<i>mcsA</i> overlap L primer
JF37R	TCATTTCTCAGGCCAGGA	<i>mcsA</i> overlap R primer
JF68F	CGCGTGATTTCTGAGAGATGAC	<i>ble</i> primer
JF68R	GTCATCGCTGGGCCACTAGCG	<i>ble</i> primer
JF74F	CGACACAAAGCAGTGCTCCTGTGTTG	<i>mcsA</i> mutant test primer
JF74R	CTGGCGTCTATGCAAATGGGCATATGGC	<i>mcsA</i> mutant test primer
JF92F	ACTCCTCGCAAACCATGCC	<i>argB</i> primer
JF92R	TGGTTTGCAGAAGCTTTCTCTG	<i>argB</i> primer
JF101F	AAGATAGACGAAAAGAAAAC	<i>coaT</i> overlap L primer
JF103R	GAGGTTGTTCTTGTGGGAG	<i>coaT</i> overlap R primer
JF104F	GGTTGACTGCCTCAGGCCTGC	<i>adeA</i> primer
JF104R	CGACCGATCTCGTACGAGTCC	<i>adeA</i> primer
JF105F	GAGAGCCGAGAATACGCCGAG	<i>coaT</i> mutant test primer
JF105R	GAGATGTGGGTGCTCATGCTGG	<i>coaT</i> mutant test primer
JF10R	GAAGACAGGGGAGGTATCGAG	DEBS1-F1 primer
JF40F	CGAGATGCCTCCTGAGACCGCC	DEBS1-F3 primer
JF40R	CCTGGACGGCGGAATGGAACCTCG	DEBS1-F3 primer
JF78F	AGCAGTTTTGCCAAATTTGCAAAGTCTTCAATGGCGGCCTCAAGTACACG	<i>sorB</i> -F1 primer
JF79R	CCAGCTCATAGCACACGAGTACTGC	<i>sorB</i> -F1 primer
JF79F	GCTCTCAAAGTCTGGTGTGATG	<i>sorB</i> -F2 primer
JF80R	GCCAGCGCAGCCGTAAGTCTTG	<i>sorB</i> -F2 primer
JF3F1	CCGATGTCGAGAAGCTCCTC	DEBS1-F4(+) primer
JF75R	CGCCTCATCTCGCGACGAGGCAGACGCAATATCGAGCTGCTGGCCGATAT	DEBS1-F4 primer
JF77R	CGCCTCATCTCGCGACGAGGCAGACGCAATGGACTCGGTCTCAGGCTCGC	DEBS1-F4+ primer
JF1F	TTTCTCCCGCATGAACTCCC	DEBS1-F2 primer
JF21R	GGCGGGCATCATCGATCTCA	DEBS1-F2 primer
JF75F	CTCGCGCATCATATCGGCCAGCAGCTCGATATTGCGTCTGCCTCGTCGCG	<i>sorA</i> -ACP primer
JF76R	TGCCAACTTTGTACAAGAAAGCTGGGTGCGAAGCTTTTCATGAAGACTTTGCAAATT	<i>sorA</i> -ACP(+) primer
JF77F	GAGCGCGCCGGCGAGCCTGAGACCGAGTCCATTGCGTCTGCCTCGTCGCG	<i>sorA</i> -ACP+ primer
JF82F	TTTCTTTCAACACAAGATCCCAAAGTCAAATGATAGAGAGGGTTCTCCC	<i>sePpil'</i> primer
JF82R	TTCATTCTATGCGTTATGAACATGTTCCCTTACCAGTCTCTAAGTCGGA	<i>sePpil'</i> primer
JF80F	GTTCCACAGCCCCAGCATGTAGCC	<i>sorB</i> -F3 primer
JF81R	TGCCAACTTTGTACAAGAAAGCTGGGTGCGTTCATCGCAAAAAGCCGCTAT	<i>sorB</i> -F3 primer
AtBraA_1F	TCTTTCAACACAAGATCCCAAAGTCAAAGGATGGCTCCAGACATCGACAA	<i>braA</i> primer
AtBraA_1R	CTATGCGTTATGAACATGTTCCCTGGCGGTCATTCCACATATGACAAAT	<i>braA</i> primer
AtBraB_1F	CGACTGACCAATTCCGCAGCTCGTCAAAGGATGGTACCTTCGGTAATGGA	<i>braB</i> primer
AtBraB_1R	CTGGTAGACGTCATATAATCATACGGCGGTTACTTCTTCGGCTCCAGAA	<i>braB</i> primer
AtBraC_1F	CAGTACCCCGCTTGAGCAGACATCACC GGATGGCATCTTTACTTGCC	<i>braC</i> primer
AtBraC_1R	ATGTCCATATCATCAATCATGACCGGCGCTATCGAATCTTCTTAAATA	<i>braC</i> primer
170	ACTTCATCGCAGCTTGACTA	General primer ( <i>P<sub>gpdA</sub></i> )
171	TCTTTCATTATCTTGCGAAC	General primer ( <i>T<sub>gpdA</sub></i> )
147	TGCTTGAGGATAGCAACCG	General primer ( <i>P<sub>anyB</sub></i> )
148	GGGGATGACAGCAGTAACGA	General primer ( <i>T<sub>anyB</sub></i> )
168	CTTCCGTCCTCCAAGTTAGT	General primer ( <i>P<sub>eno</sub></i> )
189	CTTTGTGCAGCTCAGAGTT	General primer ( <i>T<sub>eno</sub></i> )
146	GTA AACGACGGCCAG	General primer (pE-YA)
169	ACCATCTTTCGATAATGTGT	General primer ( <i>T<sub>eno</sub></i> )

## 6.3 Microbiology Methods

### 6.3.1 Transformation of Chemically Competent *E. coli*

An aliquot of 50  $\mu\text{L}$  chemically competent cells (kindly provided by technical staff) was defrosted on ice. Add 30 ng of purified plasmids into the *E. coli* and gently mix by flicking the bottom of the microcentrifuge tube a couple of times. Then incubate the mixture on ice for approximately 20 min. After that, heat-shock the tube by immediately put the bottom of the tube into the 42 °C water bath for about 75 s. Place the tube back on ice and cool it for 2 min. Add 300  $\mu\text{L}$  of LB or SOC liquid media without antibiotics into the tube with *E. coli* and shake at 37 °C for 1 h. Before plating, prepare some LB agar plates with appropriate antibiotics at room temperature in advance. Spread some or all 1 h-cultures of *E. coli* onto the LB agar plates. Place plates in a 37 °C incubator overnight and then observe the colonies on the face of the agar plates.

All *E. coli* strains harboring vectors were cultured on the LB liquid or solid media in the presence of appropriate antibiotics. For some positive transformants of *E. coli*, to use them for a long term, colonies were dissolved in 25 % (w/v) glycerol and stored at -80 °C.

### 6.3.2 LiOAc-mediated Transformation of Competent *S. cerevisiae*

Yeast was streaked out on YPAD agar and incubated at 28 °C for three days. A single colony was transferred into 10 mL YPAD medium and incubated overnight at 28 °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. All cultures were harvested by centrifuging at 3000 g for 5 min. Pellet was washed with 25 mL deionized H<sub>2</sub>O, and then centrifugation was repeated. Following the pellet was resuspended in 1 mL of deionized H<sub>2</sub>O, transferred into a 1.5 mL reaction tube and centrifuged at 11,000 g for 15 s. The pellet was resuspended in 400  $\mu\text{L}$  deionized H<sub>2</sub>O and each 50  $\mu\text{L}$  was transferred into a separate 1.5 mL reaction tube. Each sample was pelleted at 11,000 g for 15 s and the pellet was used



for the yeast transformation protocol. The following components were added to the pellet in order: 240  $\mu\text{L}$  PEG solution (50 % (w/v) polyethylene glycol 3350); 36  $\mu\text{L}$  LiAc (1 M); 50  $\mu\text{L}$  denatured salmon testis DNA (2  $\text{mg}\cdot\text{mL}^{-1}$  in TE buffer); 34  $\mu\text{L}$  DNA master mix containing the linearized plasmid and desired inserts in equimolar concentration (the uncut plasmid was used as the positive control and the linearized plasmid was used as the negative control).

The pellet was resuspended in the transformation mix and first incubated for 30 min at 30 °C, then for 40 min at 42 °C. Cells were pelleted by centrifugation at 11,000 g for 15 s and supernatant was removed. The pellet was resuspended in 500  $\mu\text{L}$  deionized  $\text{H}_2\text{O}$ , and 250  $\mu\text{L}$  suspension was spread on selective SM-Ura plates, which were incubated for 3 to 4 days at 28 °C. Yeast plasmid was extracted using a Zymoprep™ Yeast Plasmid Miniprep II kit (Zymo research, Orange, California, USA).

The untransformed yeast cells were grown on YPAD liquid or solid media. All transformed yeast cells were grown on SM-Ura selection media. Both of them were incubated in a 28 °C incubator or shaker at 200 rpm.

### 6.3.3 PEG-mediated Transformation of *A. oryzae* NSAR1

*A. oryzae* NSAR1 conidia from sporulating plates were inoculated into 50 mL GN and incubated overnight at 28 °C with shaking. Germinated conidia were collected by filtrating with the sterile Miracloth. 10 mL of filter-sterilized protoplasting solution (10  $\text{mg}\cdot\text{mL}^{-1}$  *Trichoderma* lysing enzyme in 0.8 M sodium chloride) was used to resuspend the pellet. The tube was then incubated at room temperature with gentle mixing for 3 h. The protoplasts were released from hyphal strands by gentle pipetting with wide-bore pipette (5 mL) and filtered through sterile Miracloth. The filtrate was centrifuged at 3000 g for 5 min to pellet the protoplasts which were resuspended in solution I (100  $\mu\text{L}$  per transformation; 0.8 M sodium chloride, 10 mM calcium chloride, 50 mM Tris-HCl pH 7.5). The concentration of the protoplasts was assessed microscopically. 10  $\mu\text{L}$  of mini-prep DNA ( $\sim 1 \mu\text{g}$ ) was added to 100  $\mu\text{L}$  of the protoplast suspension and incubated on ice for 2 min. 1 mL of solution II (60 % (w/v) PEG 3350, 0.8 M sodium chloride, 10 mM calcium chloride, 50 mM Tris-HCl

pH 7.5) was added and the transformation mixture was incubated at room temperature for 20 min. 5 mL of molten CZD/S top medium (3.5 % Czapek Dox broth, 1 M sorbitol, 0.05 % adenine, 0.15 % methionine, 0.10 % ammonium sulphate, 0.8 % agar) at 50 °C was added to the transformation mixture and overlaid onto prepared plates (3.5 % Czapek Dox broth, 1 M sorbitol, 0.05 % adenine, 0.15 % methionine, 0.10 % ammonium sulphate, 1.50 % agar). Plates were incubated at 28 °C.

All untransformed *A. oryzae* was grown on DPY liquid or solid media. All transformed *A. oryzae* was grown on DPY, CZD/S, CZD/S+A, and CZD/S without methionine solid or liquid media for extraction of metabolites. In the fungal transformation process, all relevant CZD/S agar media were used for screening of pure transformants. All media were grown in a 28 °C incubator or shaker at a speed of 110 rpm.

## 6.4 Molecular Biological Methods

### 6.4.1 Genomic DNA, RNA and cDNA Preparation

Genomic DNA of *A. oryzae* was isolated from about 100 mg of mycelia for each extraction. The GeneElute Plant Genomic DNA Miniprep Kit (Sigma-Aldrich) was used according to the manufacture's protocol.

The total RNA of *A. oryzae* was isolated from approximately 100 mg fresh mycelia without the medium. The commercially chemical reagent TRIzol™ was used. In this method, it was used together with other organic solvents. First, 1 mL TRIzol™ reagent was added to per 100 mg fungal mycelia and homogenized using a homogenizer. The mixture was incubated for 5 min to make nucleoproteins complex dissociated completely. Then, 0.2 mL chloroform was added into the mixture, incubated for 3 min, and centrifuged for 15 min at 12,000 g at a prechilled 4 °C centrifuge. The upper aqueous phase containing RNA was transferred to a new tube. 0.5 mL of isopropanol was added to the aqueous phase, incubated for 10 min, and centrifuged for 10 min at 12,000 g at 4 °C. The supernatant was discarded, and the pellet was resuspended in 1 mL of 75 % ethanol. The sample was vortexed briefly, and then centrifuged for 2 min at 12,000 g at 4 °C. The RNA pellet was air-dried for

a few minutes, and the pellet was resuspended in 50  $\mu$ L of RNase-free water. The total RNA extracted from *A. oryzae* was stored at -80 °C for a long term.

Before using freshly prepared RNA or stored RNA at -80 °C, DNase I was needed to remove the residual genomic DNA in the total RNA. The total RNA was reversed transcriptionally to cDNA with High-Capacity RNA-to-cDNA™ Kit (ThermoFisher Scientific). The high-quality cDNA was obtained according to the manufacture's method.

#### 6.4.2 Plasmid DNA Extraction

Plasmid DNA of *E. coli* was generally extracted from 3-5 mL of overnight-cultivated *E. coli* culture using the NucleoSpin® Plasmid Kit (Machery-Nagel). All of procedures were performed according to the manufacture's protocol. The total content of each isolation can reach more than 20  $\mu$ g of plasmid DNA. The adaptable size covers all of constructed vectors in this thesis.

Plasmid DNA of yeast was isolated from the 3 d-culture of *S. cerevisiae* using Zymoprep™ Yeast Plasmid Miniprep II Kit (Zymo Research, USA). All of procedures were performed according to the manufacture's protocol. All of yeast colonies were collected from one selection agar plate and dissolved in 200  $\mu$ L solution 1 of the kit as the beginning. Ultimately, the isolated plasmid was recovered in a small volume of 15  $\mu$ L deionized H<sub>2</sub>O. Then all of the suspension was transformed immediately to competent *E. coli* for the subsequent plasmid identification.

#### 6.4.3 Gene and Plasmid Identification

To test the plasmid construction or amplify fragments from DNA templates, the polymerase chain reaction was carried out with commercial PCR polymerases. OneTaq® 2× Master Mix (NEB) was only used for screening positive colonies because of its comparatively low fidelity. For precise fragment amplification for plasmid construction, Q5® High-Fidelity 2× Master Mix (NEB) was used. Both of them were

used according to the manufacture's standard protocols.

All of restriction enzymes were provided by NEB commercially. The enzymatic reactions were conducted according to restriction digest protocols. To ensure the plasmid cut completely, the incubation time was always overnight. And then the digestion effort was evaluated by agarose gel electrophoresis.

To recover a single gene fragment amplified by PCR, NucleoSpin® Gel and PCR Clean-up Kit was used to directly purify PCR mixture according to the manufacture's protocol. In the other case, to recover one fragment from many, PCR products were separated on the agarose gel by nucleotide electrophoresis. Then the target fragment was cut off from the gel and subsequently purified with the kit above.

To screen positive colonies of *E. coli*, *E. coli* single colonies were picked out with toothpicks from the LB agar plate and dipped into PCR reaction mixture for a few seconds, which was prepared with OneTaq® 2 × Master Mix (NEB) in advance. The PCR program was performed as the manufacture's protocol.

DNA samples were sequenced by Eurofins Genomics (Ebersberg). Templates consisted of more than 1.5 µg of purified DNA. 2 µL of 10 µM primer solution was added into the mixture. The total volume of the mixture was complemented to at least 17 µL for gene sequencing.

## 6.5 Biochemistry Methods

### 6.5.1 Protein Expression and Purification

#### Expression and Purification of BraA

For protein expression of the terpene cyclase BraA in *E. coli* BL21(DE3), the *braA* gene was amplified from *A. oryzae braA* cDNA. RNA was obtained by growing *A. oryzae braA* for 4 days in DPY liquid medium at 28 °C and 110 rpm and subsequent RNA extraction of the mycelia using the Quick-RNA Fungal/Bacterial Miniprep Kit (Zymo Research, Irvine, CA, USA). RNA was transcribed into cDNA applying the High Capacity RNA-to-DNA kit (Applied Biosystems, Foster City, CA, USA). The *braA* gene

was cloned into the pET28a+ plasmid by yeast heterologous combination according to the literature.<sup>63</sup> The purified vector was transformed into *E. coli* BL21(DE3) by standard heat shock method. 5 mL of LB medium supplemented with 50  $\mu\text{g}\cdot\text{mL}^{-1}$  kanamycin was inoculated with the transformed host and grown overnight at 37 °C and 200 rpm. 1 mL of the seed culture was transferred to 1 L LB medium (+50  $\mu\text{g}\cdot\text{mL}^{-1}$  kanamycin) and equally distributed between 10 baffled 500 mL Erlenmeyer flasks. The cultures were cultivated at 37 °C and 200 rpm for 3 – 4 hours until the OD<sub>600</sub> reached approximately 0.5. To induce the protein expression 0.5 mM IPTG was added to the flasks. After overnight growth at 16 °C and 200 rpm, the cells were harvested by centrifugation (9,000 g, 3 min, 4 °C). The cell pellet was resuspended in 40 mL lysis buffer (50 mM Tris/HCl pH 8.0, 150 mM NaCl, 10 % glycerol (v/v), 20 mM imidazole) and then lysed by sonication followed by centrifugation (10,000 g, 20 min, 4 °C) to remove the debris. All purification steps were carried out on an ÄKTAFPLC (GE Healthcare) with a 5 mL nickel affinity HisTrap FF column (GE Healthcare). The target protein was eluted of the column with elution buffer (50 mM Tris/HCl pH 8.0, 150 mM NaCl, 10 % glycerol (v/v), 500 mM imidazole) and detected by UV. Fractions with strong UV signals were combined and solution was replaced by storage buffer (50 mM Tris/HCl pH 7.5, 20 % glycerol (v/v)) using an Amicon Ultra-15 centrifugal filter (Millipore, 30K) and finally concentrated in 0.5 mL volumes. The protein concentration was assessed with a spectrophotometer and the protein was stored at -80 °C. Purified protein was analyzed by SDS-PAGE.

### **Expression and Purification of BraB**

For protein expression of the *N*-acetylglucosamine transferase BraB in *E. coli* BL21(DE3) a commercially synthesized vector (BaseClear, Leiden, Netherlands) based on the pET28a+ plasmid was used. Therein, the codon-optimized gene sequence of *braB* was inserted downstream of the T7 promoter. Transformation of the vector into *E. coli* BL21(DE3) and protein purification followed the method above. The protein concentration was assessed with a spectrophotometer and the protein was stored at -80 °C. Purified protein was analyzed by SDS-PAGE.

## 6.5.2 Enzyme Activity Assay

### ***In-vitro* Enzyme Assays with BraA**

To identify the product of BraA, the enzyme was tested in 300  $\mu$ L reaction volumes containing Tris/HCl buffer, 150  $\mu$ g BraA, 500  $\mu$ M FPP, 5 mM MgCl<sub>2</sub> and 5 mM DTT. The solution was incubated at 28 °C for 30 min and extracted with 300  $\mu$ L pentane. The organic phase was analyzed by GC-MS.

### ***In-vitro* Enzyme Assays with BraB**

Acceptor substrate specificity of BraB was assessed in 100  $\mu$ L reaction volumes containing CutSmarter Buffer (NEB), 2 mM UDP-GlcNAc, 200  $\mu$ g of purified enzyme and 0.5 mM substrate (geraniol, linalool, perillyl alcohol, menthol, benzyl alcohol, serine, serotonin, or 3,4-dichlorophenol). The reaction mixtures were incubated at 37 °C overnight and extracted by adding 200  $\mu$ L ethyl acetate. The organic phase was dried and dissolved in 200  $\mu$ L MeOH for LCMS analysis. For negative control, the enzyme was deactivated by heat (95 °C for 5 min) prior to *in-vitro* reaction. Donor substrate specificity of BraB was tested in 100  $\mu$ L reaction volumes containing 10  $\mu$ L 10x CutSmarter Buffer (NEB), 10  $\mu$ M UDP-sugar (UDP-GlcNAc, UDP-glucose, or UDP-galactose), 100  $\mu$ g of BraB, 150  $\mu$ g of BraA, 100  $\mu$ M FPP and 5 mM DTT. The reaction mixtures were incubated at 30 °C overnight and extracted by adding 200  $\mu$ L ethyl acetate. The organic phase was dried and dissolved in 200  $\mu$ L MeOH for LCMS analysis.

In addition, donor substrate specificity of BraB was tested with perillyl alcohol as alternative acceptor substrate in 100  $\mu$ L reaction volumes containing CutSmarter Buffer (NEB), 2 mM UDP-sugar (UDP-GlcNAc, UDP-glucose, or UDP-galactose), 200  $\mu$ g of purified enzyme and 0.5 mM perillyl alcohol. The reaction mixtures were incubated at 37 °C overnight and extracted by adding 200  $\mu$ L ethyl acetate. The organic phase was dried and dissolved in 200  $\mu$ L MeOH for LCMS analysis.

To further assess the preferences of BraB for UDP-sugar, a competition assay was conducted. For this purpose, a reaction assay in triplicate was prepared with a mixture of UDP-GlcNAc and UDP-Glc. The reaction mixtures were composed of 10  $\mu$ L 10x CutSmarter Buffer (NEB), 200  $\mu$ g of BraB, 150  $\mu$ g of BraA, 200  $\mu$ M FPP, 5 mM

DTT and 100  $\mu\text{M}$  of each UDP-sugar. The reaction mixtures were incubated at 30 °C for 10 h and extracted by adding 200  $\mu\text{L}$  ethyl acetate. The organic phase was dried and dissolved in 200  $\mu\text{L}$  MeOH for LCMS analysis.

### ***In-vivo* Enzyme Assays with BraB**

100  $\mu\text{L}$  of *E. coli* pTYGS-arg-*braB* seed culture was transferred to 100 mL LB medium (+ 50  $\mu\text{g}\cdot\text{mL}^{-1}$  kanamycin) in a baffled 500 mL Erlenmeyer flask. The culture was cultivated at 37 °C and 200 rpm for 3-4 hours until the OD<sub>600</sub> reached approximately 0.5. To induce the protein expression 0.5 mM IPTG was added to the flask and 5 mg perillyl alcohol dissolved in 100  $\mu\text{L}$  methanol was fed to the culture simultaneously. After overnight growth at 16 °C and 200 rpm, the culture was extracted with two volumes of ethyl acetate. The organic phase was concentrated to dryness and dissolved in methanol for LCMS analysis.

For compound isolation a large-scale fermentation (1 L) was conducted. The *E. coli* transformant was cultivated in 10 baffled 500 mL Erlenmeyer flasks. The induction, feeding (7.5 mg perillyl alcohol per flask) and extraction were performed with the same conditions as described before.

Fractionation of half of the crude extract was achieved by preparative reversed-phase HPLC (Gilson GX270). A VP Nucleodur® 100-5 C18ec column (250 x 21 mm, 5  $\mu\text{m}$ ) connected to a Kromasil® 100 C18 pre-column (50 x 20 mm, 7  $\mu\text{m}$ ) served as stationary phase. The mobile phase was composed of Milli-Q water (Millipore Schwalbach, Germany; solvent A2) supplemented with 0.05 % trifluoroacetic acid (TFA) and acetonitrile (HPLC-grade; solvent B2) supplemented with 0.05 % TFA. The yellowish extract (89 mg) was dissolved in 1.2 mL of methanol. The separation was accomplished over two runs with the following elution conditions at a flow rate of 20  $\text{mL}\cdot\text{min}^{-1}$ : isocratic conditions at 5 % solvent B2 for 5 min → linear gradient of B2 from 5 % to 25 % in 10 min → linear gradient of B2 from 25 % to 50 % in 30 min → linear gradient of B2 from 50 % to 100 % in 10 min. UV detection was carried out at  $\lambda = 190, 200$  and 220 nm. The glycosylated product (5 mg) was obtained in pure amounts at a  $R_t$  of 28 min.

## 6.6 Chemical Analysis

### 6.6.1 Extraction of Fungal Metabolites

#### Production and Isolation of Brasilane-type Terpenoids

The ascospore-derived isolate *Annulohyphoxylon truncatum* CBS 1407781 was grown in a 500 mL Erlenmeyer flask containing 200 mL YMG medium (1.0 % malt extract, 0.4 % glucose, 0.4 % yeast extract, pH 6.3) as seed culture on a shaker for 5 days at 25 °C and 140 rpm. Two 5 L Erlenmeyer flasks with 2 L YMG medium (total volume 4 L) were each inoculated with 10 mL of the homogenized seed culture and incubated for 9 days under the same conditions. Afterwards, the biomass was separated from the culture broth by vacuum filtration and the supernatant was incubated with 40 g Amberlite XAD-16N resin (Sigma-Aldrich) for 4 h under shaking conditions. The filtered adsorber resin was then extracted with 1 L acetone and the organic solvent was evaporated to yield 1 g crude extract. The latter was pre-fractionated by flash chromatography using a Reveleris X2 purification system (Büchi, Flawil, Switzerland) equipped with a Reveleris Silica 40 g column and an ELS detector. A two-stage gradient (flowrate of 40 mL·min<sup>-1</sup>) with dichloromethane (DCM, solvent A1) and DCM/acetone (8:2, solvent B1) in the first stage was applied going from 0 % to 20 % B in 9.5 min. For the second stage solvent B and DCM/acetone/methanol (8:2:3, solvent C) were used starting with a linear gradient from 20 % to 60 % C in 9.5 min, followed by 60 % to 100 % C in 5 min and finished with isocratic conditions (100 % C) for 5 min. Fractions between 16 and 20 min were combined (112 mg) and forwarded to a second round of flash chromatography using a Reveleris Silica 12 g column and a linear gradient (flowrate of 30 mL·min<sup>-1</sup>) with solvent A and B from 0 % to 100 % B in 25 min. Fractions between 10 and 15 min were combined (48 mg), which contained a mixture of brasilane A **2** and E **4**. Fractions between 15 and 18 min (37 mg combined) contained semi-pure quantities of brasilane D **3**. For final purification the mixture of **2** and **4** was separated on a preparative HPLC system (Gilson, Middleton, USA) equipped with a GX-271 Liquid Handler, a 172 DAD, a 305 and 306 pump (with 50SC Piston Pump Head). As stationary phase a VP 250/21 Nucleodur 100-5 C18 ec column (Macherey–Nagel) was used and the mobile phase was composed of deionized water with 0.5 % acetic acid (solvent D) and acetonitrile with 0.5 % acetic acid (solvent E). A linear gradient



from 45 % to 80 % solvent E in 24 min with a flow rate set to 20 mL·min<sup>-1</sup> was applied to yield **4** at a retention time (Rt) of 8 min (7.3 mg) and **2** at Rt = 15 min (2.4 mg). To further purify **3** the same preparative HPLC conditions were used. **3** eluted after 6.2 min and 2.9 mg of pure compound was obtained.

### **Production and Analysis of *A. oryzae* Metabolites**

*A. oryzae* NSAR1 or transformant was cultured in DPY for a couple of days. After incubation, a small part of mycelia in each flask was collected by filtering with sterile ceramic funnels and then packed in aluminum foil for saving in the -20 °C fridge. The remaining culture was poured in a 500 mL beaker, and the mycelia were homogenized by electric stirrer. After breaking, the filtrate was collected by ceramic funnels and then saved in Erlenmeyer flasks for extraction. The filtrate and 100 mL of ethyl acetate were added to a separatory funnel, mixed and the organic level was then separated. 100 mL of ethyl acetate more was added and repeated again. About one scoop of anhydrous magnesium sulfate was added to the organic level, mixed and filtered with fold filter papers. The filtrate (weigh before and after drying) was dried, and about 1 mL of methanol was added to dissolve the extract. The solution was filtered with glass cotton, and the sample was subjected to LCMS for analysis.

#### **6.6.2 Analytical LCMS**

Analytical LCMS data was obtained using a Waters LCMS system consisting of a Waters 2767 autosampler, Waters 2545 pump system, a Phenomenex Kinetex column (2.6 μm, C<sub>18</sub>, 100 Å, 4.6 x 100 mm) equipped with a Phenomenex Security Guard precolumn (Luna, C<sub>5</sub>, 300 Å) at a flow rate of 1 mL·min<sup>-1</sup>. Detection was carried out by a diode array detector (Waters 2998) in the range 210 to 600 nm and an ELSD detector (Waters 2424) connected to a mass spectrometry, Waters SQD-2 mass detector, operating simultaneously in ES+ and ES- modes between 150 and 1000 *m/z*. The mobile phase was composed of HPLC-grade water mixed with 0.05 % formic acid (solvent A2) and HPLC-grade acetonitrile mixed with 0.045 % formic acid (solvent B2). A solvent gradient was run over 15 min starting at 10 % B2 and ramping up to 90 % B2. In case of the competition assay a shallower gradient was applied ramping from 10 to 50 % B2 in 15 min.

### 6.6.3 GCMS

GCMS data were obtained on 6890N Network GC system (Agilent Technologies, Santa Clara, CA, USA) combined with a 5973 inert Mass Selective Detector (Agilent Technologies). The GC was equipped with an Agilent HP-5MS capillary column (30 m, 0.25 mm, 0.25  $\mu\text{m}$ , 5 % Diphenyl / 95 % Dimethylpolysiloxan). The following GC parameters were used for analysis: inlet pressure 20.96 psi; split mode (5:1); split flow 10  $\text{mL}\cdot\text{min}^{-1}$ ; carrier gas helium; saver flow 30  $\text{mL}\cdot\text{min}^{-1}$ ; saver time 2 min; injection volume 1  $\mu\text{L}$ ; temperature gradient starting at 100  $^{\circ}\text{C}$  for 2 min, increasing at 2  $^{\circ}\text{C}\cdot\text{min}^{-1}$  to 180  $^{\circ}\text{C}$ , then increasing at 40  $^{\circ}\text{C}\cdot\text{min}^{-1}$  to 300  $^{\circ}\text{C}$ . Mass spectra were recorded under the following conditions: source temperature 230  $^{\circ}\text{C}$ , quadrupole temperature 150  $^{\circ}\text{C}$ , mass range 43.0–600.0.

### 6.6.4 HRMS and NMR

NMR data were recorded with a Bruker AVII-600 spectrometer equipped with a BBFO SmartProbe, a Bruker Avance III 500 MHz spectrometer equipped with a BBFO (plus) SmartProbe ( $^1\text{H}$  500 MHz,  $^{13}\text{C}$  125 MHz), and a Bruker Avance III 700 MHz spectrometer equipped with a 5 mm TCI cryoprobe ( $^1\text{H}$  700 MHz,  $^{13}\text{C}$  175 MHz). Chemical shifts  $\delta$  were referenced to methanol- $d_4$  ( $^1\text{H}$ ,  $\delta = 3.31$  ppm;  $^{13}\text{C}$ ,  $\delta = 49.15$  ppm), chloroform- $d$  ( $^1\text{H}$ ,  $\delta = 7.27$  ppm;  $^{13}\text{C}$ ,  $\delta = 77.00$  ppm), acetone- $d_6$  ( $^1\text{H}$ ,  $\delta = 2.05$  ppm;  $^{13}\text{C}$ ,  $\delta = 29.92$  ppm), acetonitril- $d_3$  ( $^1\text{H}$ ,  $\delta = 1.94$  ppm;  $^{13}\text{C}$ ,  $\delta = 1.39$  ppm), benzene- $d_6$  ( $^1\text{H}$ ,  $\delta = 7.16$  ppm;  $^{13}\text{C}$ ,  $\delta = 128.39$  ppm), DMSO- $d_6$  ( $^1\text{H}$ ,  $\delta = 2.50$  ppm;  $^{13}\text{C}$ ,  $\delta = 39.51$  ppm) and pyridin- $d_5$  ( $^1\text{H}$ ,  $\delta = 7.22$  ppm;  $^{13}\text{C}$ ,  $\delta = 123.87$  ppm). HR-ESI-MS measurements were conducted on a maXis ESI-TOF (Bruker Daltonics) mass spectrometer coupled to an Agilent 1200 Infinity Series HPLC system (Agilent Technologies). An Acquity UPLC BEH C18 column (Waters, Milford, USA; 2.1 x 50 mm, 1.7  $\mu\text{m}$ ) was used as stationary phase. Compounds were eluted with solvent A (Milli-Q water + 0.1 % formic acid) and solvent B (HPLC grade acetonitrile + 0.1 % formic acid) using a flow rate of 0.6  $\text{mL}\cdot\text{min}^{-1}$  and the following gradient: 5 % B for 0.5 min increasing to 100 % B in 19.5 min and then maintaining 100 % B for 5 min. UV/vis spectra was detected in the range of 200 – 600 nm. Mass spectra were recorded in positive ionization mode (ESI+) with a scan range of 100 – 2500  $m/z$ , a temperature of 200  $^{\circ}\text{C}$  and capillary voltage of 4500 V.

## Reference

- (1) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2016**, *79* (3), 629–661.
- (2) Ali, S. M.; Siddiqui, R.; Khan, N. A. *J. Pharm. Pharmacol.* **2018**, *70* (10), 1287–1300.
- (3) Blackwell, M. *Am. J. Bot.* **2011**, *98* (3), 426–438.
- (4) Schueffler, A.; Anke, T. *Nat. Prod. Rep.* **2014**, *31* (10), 1425–1448.
- (5) Daley, D. K.; Brown, K. J.; Badal, S. *Pharmacogn. Fundam. Appl. Strateg.* **2017**, 413–421.
- (6) Rodriguez, R. J.; White, J. F.; Arnold, A. E.; Redman, R. S. *New Phytol.* **2009**, *182* (2), 314–330.
- (7) Schor, R.; Cox, R. *Nat. Prod. Rep.* **2018**, *35* (3), 230–256.
- (8) He, Y.; Cox, R. J. *Chem. Sci.* **2016**, *7* (3), 2119–2127.
- (9) Dewar, M. J. S. *Nature* **1945**, *155* (3), 50–51.
- (10) Davison, J.; Al Fahad, A.; Cai, M.; Song, Z.; Yehia, S. Y.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (20), 7642–7647.
- (11) Kutchan, T. M.; Dittrich, H.; Bracher, D.; Zenk, M. H. *Tetrahedron* **1991**, *47* (31), 5945–5954.
- (12) Sweeney, M. J.; Dobson, A. D. W. *FEMS Microbiol. Lett.* **1999**, *175* (2), 149–163.
- (13) Bell, A. A.; Wheeler, M. H. *Annu. Rev. Phytopathol.* **1986**, *24* (1), 411–451.
- (14) Hendrickson, L.; Davis, C. R.; Roach, C.; Nguyen, D. K.; Aldrich, T.; McAda, P. C.; Reeves, C. D. *Chem. Biol.* **1999**, *6* (7), 429–439.
- (15) Schuemann, J.; Hertweck, C. *Physiol. Genet.* **2009**, 331–351.
- (16) Butler, M. J.; Day, A. W. *Can. J. Microbiol.* **1998**, *44* (12), 1115–1136.
- (17) Oide, S.; Turgeon, B. G. *Mycoscience* **2020**, *61* (3), 101–110.
- (18) Eisfeld, K. *Physiol. Genet.* **2009**, 305–330.
- (19) Gardiner, D. M.; Howlett, B. J. *FEMS Microbiol. Lett.* **2005**, *248* (2), 241–248.
- (20) Scharf, D. H.; Chankhamjon, P.; Scherlach, K.; Dworschak, J.; Heinekamp, T.; Roth, M.; Brakhage, A. A.; Hertweck, C. *ChemBioChem* **2021**, *22* (2), 336–339.
- (21) Schmidt-Dannert, C. *Adv. Biochem. Eng. Biotechnol.* **2015**, *148* (November 2014), 19–61.
- (22) Zhang, Y.; Han, T.; Ming, Q.; Wu, L.; Rahman, K.; Qin, L. *Nat. Prod. Commun.* **2012**, *7* (7), 963–968.
- (23) Keller, N. P. *Nat. Rev. Microbiol.* **2019**, *17* (3), 167–180.
- (24) Mao, X. M.; Xu, W.; Li, D.; Yin, W. B.; Chooi, Y. H.; Li, Y. Q.; Tang, Y.; Hu, Y. *Angew. Chemie - Int. Ed.* **2015**, *54* (26), 7592–7596.

- (25) Oakley, C. E.; Ahuja, M.; Sun, W. W.; Entwistle, R.; Akashi, T.; Yaegashi, J.; Guo, C. J.; Cerqueira, G. C.; Russo Wortman, J.; Wang, C. C. C.; Chiang, Y. M.; Oakley, B. R. *Mol. Microbiol.* **2017**, *103* (2), 347–365.
- (26) Sharma, A.; Gupta, G.; Ahmad, T.; Mansoor, S.; Kaur, B. *Food Rev. Int.* **2021**, *37* (2), 121–154.
- (27) Xu, Y.; Zhou, T.; Zhang, S.; Espinosa-Artiles, P.; Wang, L.; Zhang, W.; Lin, M.; Gunatilaka, A. A. L.; Zhan, J.; Molnár, I. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (34), 12354–12359.
- (28) Bai, J.; Lu, Y.; Xu, Y. M.; Zhang, W.; Chen, M.; Lin, M.; Gunatilaka, A. A. L.; Xu, Y.; Molnár, I. *Org. Lett.* **2016**, *18* (6), 1262–1265.
- (29) Kakule, T. B.; Lin, Z.; Schmidt, E. W. *J. Am. Chem. Soc.* **2014**, *136* (51), 17882–17890.
- (30) Skellam, E. *Trends Biotechnol.* **2019**, *37* (4), 416–427.
- (31) Palys, S.; Pham, T. T. M.; Tsang, A. *Front. Microbiol.* **2020**, *11*, 1378.
- (32) Delmas, S.; Llanos, A.; Parrou, J. L.; Kokolski, M.; Pullan, S. T.; Shunburne, L.; Archer, D. B. *Appl. Environ. Microbiol.* **2014**, *80* (11), 3484–3487.
- (33) Arazoe, T.; Ogawa, T.; Miyoshi, K.; Yamato, T.; Ohsato, S.; Sakuma, T.; Yamamoto, T.; Arie, T.; Kuwata, S. *Biotechnol. Bioeng.* **2015**, *112* (7), 1335–1342.
- (34) Mizutani, O.; Arazoe, T.; Toshida, K.; Hayashi, R.; Ohsato, S.; Sakuma, T.; Yamamoto, T.; Kuwata, S.; Yamada, O. *J. Biosci. Bioeng.* **2017**, *123* (3), 287–293.
- (35) Urquhart, A. S.; Hu, J.; Chooi, Y. H.; Idnurm, A. *Fungal Biol. Biotechnol.* **2019**, *6* (1), 1–13.
- (36) Bhattarai, K.; Bastola, R.; Baral, B. *Antibiotic Drug Discovery: Challenges and Perspectives in the Light of Emerging Antibiotic Resistance*, 1st ed.; Elsevier Inc., 2020; Vol. 105.
- (37) Bailey, A. M.; Alberti, F.; Kilaru, S.; Collins, C. M.; De Mattos-Shiple, K.; Hartley, A. J.; Hayes, P.; Griffin, A.; Lazarus, C. M.; Cox, R. J.; Willis, C. L.; O'Dwyer, K.; Spence, D. W.; Foster, G. D. *Sci. Rep.* **2016**, *6* (February), 1–11.
- (38) Clevenger, K. D.; Bok, J. W.; Ye, R.; Miley, G. P.; Verdan, M. H.; Velk, T.; Chen, C.; Yang, K. H.; Robey, M. T.; Gao, P.; Lamprecht, M.; Thomas, P. M.; Islam, M. N.; Palmer, J. M.; Wu, C. C.; Keller, N. P.; Kelleher, N. L. *Nat. Chem. Biol.* **2017**, *13* (8), 895–901.
- (39) Harvey, C. J. B.; Tang, M.; Schlecht, U.; Horecka, J.; Fischer, C. R.; Lin, H. C.; Li, J.; Naughton, B.; Cherry, J.; Miranda, M.; Li, Y. F.; Chu, A. M.; Hennessy, J. R.; Vandova, G. A.; Inglis, D.; Aiyar, R. S.; Steinmetz, L. M.; Davis, R. W.; Medema, M. H.; Sattely, E.; Khosla, C.; Onge, R. P. S.; Tang, Y.; Hillenmeyer, M. E. *Sci. Adv.* **2018**, *4* (4).
- (40) Wang, X.; Wang, C.; Duan, L.; Zhang, L.; Liu, H.; Xu, Y. M.; Liu, Q.; Mao, T.; Zhang, W.; Chen, M.; Lin, M.; Gunatilaka, A. A. L.; Xu, Y.; Molnár, I. *J. Am. Chem. Soc.* **2019**, *141* (10), 4355–4364.
- (41) Yeh, H.; Ahuja, M.; Chiang, Y.; Oakley, C. E.; Yoon, O.; Hajovsky, H.; Bok, J.; Keller, N. P.; Clay, C. C.; Oakley, B. R. **2019**, *11* (8), 2275–2284.

- (42) Van Campenhout, C.; Cabochette, P.; Veillard, A. C.; Laczik, M.; Zelisko-Schmidt, A.; Sabatel, C.; Dhainaut, M.; Vanhollebeke, B.; Gueydan, C.; Kruys, V. *Biotechniques* **2019**, *66* (6), 295–302.
- (43) Zhang, H.; Boghigian, B. A.; Armando, J.; Pfeifer, B. A. *Nat. Prod. Rep.* **2011**, *28* (1), 125–151.
- (44) Kealey, J. T.; Liu, L.; Santi, D. V.; Betlach, M. C.; Barr, P. J. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (2), 505–509.
- (45) Alberti, F.; Foster, G. D.; Bailey, A. M. *Appl. Microbiol. Biotechnol.* **2017**, *101* (2), 493–500.
- (46) Ishiuchi, K.; Nakazawa, T.; Ookuma, T.; Sugimoto, S.; Sato, M.; Tsunematsu, Y.; Ishikawa, N.; Noguchi, H.; Hotta, K.; Moriya, H.; Watanabe, K. *ChemBioChem* **2012**, *13* (6), 846–854.
- (47) Cochrane, R. V. K.; Sanichar, R.; Lambkin, G. R.; Reiz, B.; Xu, W.; Tang, Y.; Vederas, J. C. *Angew. Chemie* **2016**, *128* (2), 674–678.
- (48) Zhang, Z.; Qiao, T.; Watanabe, K.; Tang, Y. *Angew. Chemie - Int. Ed.* **2020**, *59* (45), 19889–19893.
- (49) McLean, K. J.; Hans, M.; Meijrink, B.; Van Scheppingen, W. B.; Vollebregt, A.; Tee, K. L.; Van Der Laan, J. M.; Leys, D.; Munro, A. W.; Van Den Berg, M. A. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (9), 2847–2852.
- (50) Meyer, V.; Wu, B.; Ram, A. F. J. *Biotechnol. Lett.* **2011**, *33* (3), 469–476.
- (51) Anyaogu, D. C.; Mortensen, U. H. *Front. Microbiol.* **2015**, *6* (FEB), 1–6.
- (52) Chiang, Y.; Oakley, C. E.; Ahuja, M.; Entwistle, R.; Schultz, A.; Chang, S.; Sung, C. T.; Wang, C. C. C.; Oakley, B. R. **2014**, *135* (20), 7720–7731.
- (53) Yaegashi, J.; Oakley, B. R.; Wang, C. C. C. *J. Ind. Microbiol. Biotechnol.* **2014**, *41* (2), 433–442.
- (54) Nielsen, M. T.; Nielsen, J. B.; Anyaogu, D. C.; Holm, D. K.; Nielsen, K. F.; Larsen, T. O.; Mortensen, U. H. *PLoS One* **2013**, *8* (8).
- (55) Yin, W.; Chooi, Y. H.; Smith, A. R.; Cacho, R. A.; Hu, Y.; White, T. C.; Tang, Y.; Angeles, L.; Angeles, L.; City, K. *ACS Synth. Biol.* **2014**, *2* (11), 629–634.
- (56) Liu, N.; Hung, Y. S.; Gao, S. S.; Hang, L.; Zou, Y.; Chooi, Y. H.; Tang, Y. *Org. Lett.* **2017**, *19* (13), 3560–3563.
- (57) Heneghan, M. N.; Yakasai, A. A.; Halo, L. M.; Song, Z.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; Lazarus, C. M. *ChemBioChem* **2010**, *11* (11), 1508–1512.
- (58) Fujii, R.; Minami, A.; Tsukagoshi, T.; Sato, N.; Sahara, T.; Ohgiya, S.; Gomi, K.; Oikawa, H. *Biosci. Biotechnol. Biochem.* **2011**, *75* (9), 1813–1817.
- (59) Munawar, A.; Marshall, J. W.; Cox, R. J.; Bailey, A. M.; Lazarus, C. M. *ChemBioChem* **2013**, *14* (3), 388–394.
- (60) Tagami, K.; Liu, C.; Minami, A.; Noike, M.; Isaka, T.; Fueki, S.; Shichijo, Y.; Toshima, H.; Gomi, K.; Dairi, T.; Oikawa, H. *J. Am. Chem. Soc.* **2013**, *135* (4), 1260–1263.

- (61) Tagami, K.; Minami, A.; Fujii, R.; Liu, C.; Tanaka, M.; Gomi, K.; Dairi, T.; Oikawa, H. *ChemBioChem* **2015**, *16* (17), 2076–2080.
- (62) Pahirulzaman, K. A. K.; Williams, K.; Lazarus, C. M. *A Toolkit for Heterologous Expression of Metabolic Pathways in Aspergillus Oryzae*, 1st ed.; Elsevier Inc., 2012; Vol. 517.
- (63) Gietz, R. D.; Schiestl, R. H. *Nat. Protoc.* **2007**, *2* (1), 35–37.
- (64) Nofiani, R.; de Mattos-Shiple, K.; Lebe, K. E.; Han, L. C.; Iqbal, Z.; Bailey, A. M.; Willis, C. L.; Simpson, T. J.; Cox, R. J. *Nat. Commun.* **2018**, *9* (1), 1–11.
- (65) Schor, R.; Schotte, C.; Wibberg, D.; Kalinowski, J.; Cox, R. J. *Nat. Commun.* **2018**, *9* (1).
- (66) Kahlert, L.; Bassiony, E. F.; Cox, R. J.; Skellam, E. J. *Angew. Chemie - Int. Ed.* **2020**, *59* (14), 5816–5822.
- (67) Fisch, K. M.; Bakeer, W.; Yakasai, A. A.; Song, Z.; Pedrick, J.; Wasil, Z.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. *J. Am. Chem. Soc.* **2011**, *133* (41), 16635–16641.
- (68) Cox, R. J. *Org. Biomol. Chem.* **2007**, *5* (13), 2010.
- (69) Fujii, I. *Nat. Prod. Rep.* **2009**, *26* (2), 155–169.
- (70) Staunton, J.; Weissman, K. J. *Nat. Prod. Rep.* **2001**, *18* (4), 380–416.
- (71) Du, L.; Lou, L. *Nat. Prod. Rep.* **2010**, *27* (2), 255–278.
- (72) Chan, Y. A.; Podevels, A. M.; Kevany, B. M.; Thomas, M. G. *Nat. Prod. Rep.* **2009**, *26* (1), 90–114.
- (73) Yuzawa, S.; Backman, T. W. H.; Keasling, J. D.; Katz, L. J. *Ind. Microbiol. Biotechnol.* **2018**, *45* (7), 621–633.
- (74) Cortes, J.; Wiesmann, K. E.; Roberts, G. a; Brown, M. J.; Staunton, J.; Leadlay, P. F. *Science* **1995**, *268* (5216), 1487–1489.
- (75) Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. *Science (80-. )*. **2001**, *291* (5509), 1790–1792.
- (76) Kao, C. M.; Katz, L.; Khosla, C. **1994**, *265* (July), 28–31.
- (77) Kao, C. M.; Khosla, C.; Luo, G.; Cane, D. E.; Katz, L. *J. Am. Chem. Soc.* **1995**, *117* (35), 9105–9106.
- (78) Bedford, D.; Jacobsen, J. R.; Luo, G.; Cane, D. E.; Khosla, C. *Chem. Biol.* **1996**, *3* (10), 827–831.
- (79) Oliynyk, M.; Brown, M. J. B.; Cortes, J.; Staunton, J.; Leadlay, P. F. *Chem. Biol.* **1996**, *3* (10), 833–839.
- (80) Kim, B. S.; Cropp, T. A.; Florova, G.; Lindsay, Y.; Sherman, D. H.; Reynolds, K. A. *Biochemistry* **2002**, *41* (35), 10827–10833.
- (81) Gokhale, R. S.; Tsuji, S. Y.; Cane, D. E.; Khosla, C. *Science (80-. )*. **1999**, *284* (5413), 482–485.
- (82) Mofid, M. R.; Finking, R.; Essen, L. O.; Marahiel, M. A. *Biochemistry* **2004**, *43* (14), 4128–4136.

- (83) Horswill, A. R.; Escalante-Semerena, J. C. *Microbiology* **1999**, *145*, 1381–1388.
- (84) Horswill, A. R.; Escalante-Semerena, J. C. *Biochemistry* **2002**, *41* (7), 2379–2387.
- (85) Rodriguez, E.; Gramajo, H. *Microbiology* **1999**, *145* (11), 3109–3119.
- (86) Diacovich, L.; Peirú, S.; Kurth, D.; Rodríguez, E.; Podestá, F.; Khosla, C.; Gramajo, H. *J. Biol. Chem.* **2002**, *277* (34), 31228–31236.
- (87) Dayem, L. C.; Carney, J. R.; Santi, D. V.; Pfeifer, B. A.; Khosla, C.; Kealey, J. T. *Biochemistry* **2002**, *41* (16), 5193–5201.
- (88) Zhang, W.; Reynolds, K. A. *J. Bacteriol.* **2001**, *183* (6), 2071–2080.
- (89) Frey, P. A. *Compr. Nat. Prod. II Chem. Biol.* **2010**, *7*, 501–546.
- (90) Fuller, J. Q.; Leadlay, P. F. *Biochem. J.* **1983**, *213* (3), 643–650.
- (91) Zhang, H.; Boghigian, B. A.; Pfeifer, B. A. *Biotechnol. Bioeng.* **2010**, *105* (3), 567–573.
- (92) Mutka, S. C.; Bondi, S. M.; Carney, J. R.; Da Silva, N. A.; Kealey, J. T. *FEMS Yeast Res.* **2006**, *6* (1), 40–47.
- (93) Hani, J.; Feldmann, H. *Nucleic Acids Res.* **1998**, *26* (3), 689–696.
- (94) An, J. H.; Kim, Y. S. *European Journal of Biochemistry.* 1998, pp 395–402.
- (95) Brock, M.; Fischer, R.; Linder, D.; Buckel, W. *Mol. Microbiol.* **2000**, *35* (5), 961–973.
- (96) Ledley, F. D.; Crane, A. M.; Klish, K. T.; May, G. S. *Biochem. Biophys. Res. Commun.* **1991**, *177* (3), 1076–1081.
- (97) Keszenman-Pereyra, D.; Lawrence, S.; Twfieg, M. E.; Price, J.; Turner, G. *Curr. Genet.* **2003**, *43* (3), 186–190.
- (98) Brock, M.; Buckel, W. *Eur. J. Biochem.* **2004**, *271* (15), 3227–3241.
- (99) Zhang, Y. Q.; Brock, M.; Keller, N. P. *Genetics* **2004**, *168* (2), 785–794.
- (100) Zhang, Y. Q.; Keller, N. P. *Mol. Microbiol.* **2004**, *52* (2), 541–550.
- (101) Sokatch, J. R.; Sanders, L. E.; Marshall, V. P. *J. Biol. Chem.* **1968**, *243* (10), 2500–2506.
- (102) Tashiro, S. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74* (11), 4947–4950.
- (103) Ibrahim-Granet, O.; Dubourdeau, M.; Latgé, J. P.; Ave, P.; Huerre, M.; Brakhage, A. A.; Brock, M. *Cell. Microbiol.* **2008**, *10* (1), 134–148.
- (104) Rodriguez, E.; Gramajo, H. *Microbiology* **1999**, *145* (11), 3109–3119.
- (105) Tokuoka, M.; Tanaka, M.; Ono, K.; Takagi, S.; Shintani, T.; Gomi, K. *Appl. Environ. Microbiol.* **2008**, *74* (21), 6538–6546.
- (106) Cortes, J.; Wiesmann, K. E. H.; Roberts, G. A.; Brown, M. J. B.; Staunton, J.; Leadlay, P. F. *Science (80-. )*. **1995**, *268* (5216), 1487–1489.
- (107) Krink-Koutsoubelis, N.; Loechner, A. C.; Lechner, A.; Link, H.; Denby, C. M.; Vögeli, B.; Erb, T. J.; Yuzawa, S.; Jakociunas, T.; Katz, L.; Jensen, M. K.; Sourjik, V.; Keasling, J. D. *ACS Synth. Biol.* **2018**, *7* (4), 1105–1112.
- (108) Parret, A. H.; Besir, H.; Meijers, R. *Curr. Opin. Struct. Biol.* **2016**, *38*, 155–162.

- (109) Song, Z.; Bakeer, W.; Marshall, J. W.; Yakasai, A. A.; Khalid, R. M.; Collemare, J.; Skellam, E.; Tharreau, D.; Lebrun, M.-H.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. *Chem. Sci.* **2015**, *6* (8), 4837–4845.
- (110) Yang, G.; Zhang, Y.; Lee, N. K.; Cozad, M. A.; Kearney, S. E.; Luesch, H.; Ding, Y. *Sci. Rep.* **2017**, *7* (1), 1–9.
- (111) Kim, J. H.; Komatsu, M.; Shin-ya, K.; Omura, S.; Ikeda, H. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (26), 6828–6833.
- (112) Horbach, R.; Graf, A.; Weihmann, F.; Antelo, L.; Mathea, S.; Liermann, J. C.; Opatz, T.; Thines, E.; Aguirre, J.; Deising, H. B. *Plant Cell* **2009**, *21* (10), 3379–3396.
- (113) Beld, J.; Sonnenschein, E. C.; Vickery, C. R.; Noel, J. P.; Burkart, M. D. *Nat. Prod. Rep.* **2014**, *31* (1), 61–108.
- (114) Oberegger, H.; Eisendle, M.; Schrettl, M.; Graessle, S.; Haas, H. *Curr. Genet.* **2003**, *44* (4), 211–215.
- (115) Márquez-Fernández, O.; Trigos, Á.; Ramos-Balderas, J. L.; Viniestra-González, G.; Deising, H. B.; Aguirre, J. *Eukaryot. Cell* **2007**, *6* (4), 710–720.
- (116) Jørgensen, T. R.; Park, J.; Arentshorst, M.; van Welzen, A. M.; Lamers, G.; vanKuyk, P. A.; Damveld, R. A.; van den Hondel, C. A. M.; Nielsen, K. F.; Frisvad, J. C.; Ram, A. F. J. *Fungal Genet. Biol.* **2011**, *48* (5), 544–553.
- (117) Neville, C.; Murphy, A.; Kavanagh, K.; Doyle, S. *ChemBioChem* **2005**, *6* (4), 679–685.
- (118) Stack, D.; Neville, C.; Doyle, S. *Microbiology* **2007**, *153* (5), 1297–1306.
- (119) Allen, G.; Bromley, M.; Kaye, S. J.; Keszenman-Pereyra, D.; Zucchi, T. D.; Price, J.; Birch, M.; Oliver, J. D.; Turner, G. *Fungal Genet. Biol.* **2011**, *48* (4), 456–464.
- (120) Kim, J. M.; Song, H. Y.; Choi, H. J.; So, K. K.; Kim, D. H.; Chae, K. S.; Han, D. M.; Jahng, K. Y. *J. Microbiol.* **2015**, *53* (1), 21–31.
- (121) Weissman, K. J.; Hong, H.; Oliynyk, M.; Siskos, A. P.; Leadlay, P. F. *ChemBioChem* **2004**, *5* (1), 116–125.
- (122) Wiesmann, K. E. H.; Cortés, J.; Brown, M. J. B.; Cutter, A. L.; Staunton, J.; Leadlay, P. F. *Chem. Biol.* **1995**, *2* (9), 583–589.
- (123) Blazeck, J.; Alper, H. S. *Biotechnol. J.* **2013**, *8* (1), 46–58.
- (124) Textor, S.; Wendisch, V. F.; De Graaf, A. A.; Müller, U.; Linder, M. I.; Linder, D.; Buckel, W. *Arch. Microbiol.* **1997**, *168* (5), 428–436.
- (125) Graybill, E. R.; Rouhier, M. F.; Kirby, C. E.; Hawes, J. W. *Arch. Biochem. Biophys.* **2007**, *465* (1), 26–37.
- (126) Hitschler, J.; Grininger, M.; Boles, E. *Sci. Rep.* **2020**, *10* (1), 1–11.
- (127) Macfarlane, S.; Macfarlane, G. T. *Proc. Nutr. Soc.* **2003**, *62* (1), 67–72.
- (128) Tabuchi, T.; Serizawa, N.; Uchiyama, H. *Agric. Biol. Chem.* **1974**, *38* (12), 2571–2572.
- (129) Goldenthal, M. J.; Marin-Garcia, J.; Ananthakrishnan, R. *Genome* **1998**, *41* (5), 733–738.
- (130) Suissa, M.; Suda, K.; Schatz, G. *EMBO J.* **1984**, *3* (8), 1773–1781.



- (131) Maerker, C.; Ronde, M.; Brakhage, A. A.; Brock, M. *FEBS J.* **2005**, *272* (14), 3615–3630.
- (132) KOBAYASHI, K.; HATTORI, T.; HONDA, Y.; KIRIMURA, K. *Biosci. Biotechnol. Biochem.* **2013**, *77* (7), 1492–1498.
- (133) Domin, N.; Wilson, D.; Brock, M. *Microbiology* **2009**, *155* (12), 3903–3912.
- (134) Fleck, C. B.; Brock, M. *Mol. Microbiol.* **2008**, *68* (3), 642–656.
- (135) Strauss, E. *Compr. Nat. Prod. II Chem. Biol.* **2010**, *7*, 351–410.
- (136) He, B.; Tu, Y.; Jiang, C.; Zhang, Z.; Li, Y.; Zeng, B. *Microorganisms* **2019**, *7* (4).
- (137) Fernandez, E. Q.; Moyer, D. L.; Maiyuran, S.; Labaro, A.; Brody, H. *Fungal Genet. Biol.* **2012**, *49* (4), 294–301.
- (138) Suzuki, S.; Tada, S.; Fukuoka, M.; Taketani, H.; Tsukakoshi, Y.; Matsushita, M.; Oda, K.; Kusumoto, K. I.; Kashiwagi, Y.; Sugiyama, M. *Biochem. Biophys. Res. Commun.* **2009**, *383* (1), 42–47.
- (139) Zhang, S.; Ban, A.; Ebara, N.; Mizutani, O.; Tanaka, M.; Shintani, T.; Gomi, K. *J. Biosci. Bioeng.* **2017**, *123* (4), 403–411.
- (140) Bromann, K.; Toivari, M.; Viljanen, K.; Vuoristo, A.; Ruohonen, L.; Nakari-Setälä, T. *PLoS One* **2012**, *7* (4).
- (141) Wang, C.; Hantke, V.; Cox, R. J.; Skellam, E. *Org. Lett.* **2019**.
- (142) Nguyen, K. T.; Ho, Q. N.; Do, L. T. B. X.; Mai, L. T. D.; Pham, D. N.; Tran, H. T. T.; Le, D. H.; Nguyen, H. Q.; Tran, V. T. *World J. Microbiol. Biotechnol.* **2017**, *33* (6), 1–11.
- (143) Sun, Y.; Niu, Y.; He, B.; Ma, L.; Li, G.; Tran, V. T.; Zeng, B.; Hu, Z. *J. Microbiol. Biotechnol.* **2019**, *29* (2), 230–234.
- (144) Nguyen, K. T.; Ho, Q. N.; Pham, T. H.; Phan, T. N.; Tran, V. T. *World J. Microbiol. Biotechnol.* **2016**, *32* (12), 1–9.
- (145) Nemoto, T.; Watanabe, T.; Mizogami, Y.; Maruyama, J. I.; Kitamoto, K. *Appl. Microbiol. Biotechnol.* **2009**, *82* (6), 1105–1114.
- (146) Jin, F. J.; Katayama, T.; Maruyama, J. ichi; Kitamoto, K. *Appl. Microbiol. Biotechnol.* **2016**, *100* (21), 9163–9174.
- (147) Kushnir, S.; Sundermann, U.; Yahiaoui, S.; Brockmeyer, A.; Janning, P.; Schulz, F. *Angew. Chemie - Int. Ed.* **2012**, *51* (42), 10664–10669.
- (148) Sugimoto, Y.; Ding, L.; Ishida, K.; Hertweck, C. *Angew. Chemie - Int. Ed.* **2014**, *53* (6), 1560–1564.
- (149) Piel, J. *Nat. Prod. Rep.* **2010**, *27* (7), 996–1047.
- (150) Hertweck, C. *Trends Biochem. Sci.* **2015**, *40* (4), 189–199.
- (151) Weissman, K. J. *Nat. Prod. Rep.* **2016**, *33* (2), 203–230.
- (152) Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Clay, A. T. *Nat. Chem. Biol.* **2012**, *8* (7), 615–621.

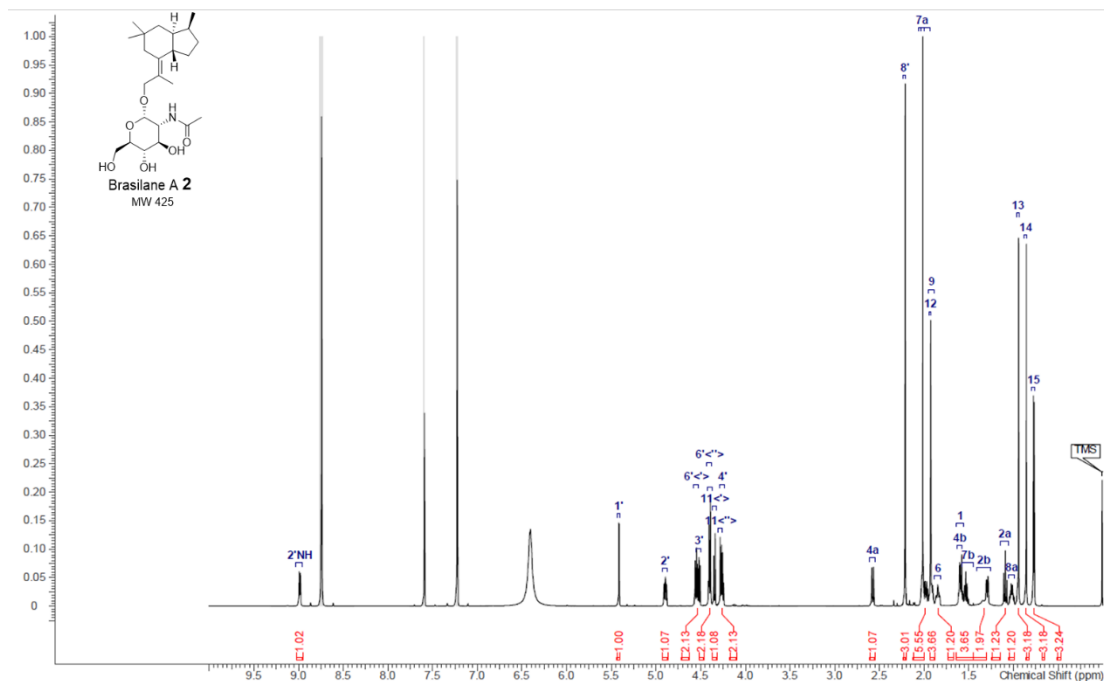
- (153) Kwan, D. H.; Sun, Y.; Schulz, F.; Hong, H.; Popovic, B.; Sim-Stark, J. C. C.; Haydock, S. F.; Leadlay, P. F. *Chem. Biol.* **2008**, *15* (11), 1231–1240.
- (154) Kao, C. M.; Luo, G.; Katz, L.; Cane, D. E.; Khosla, C. *J. Am. Chem. Soc.* **1996**, *118* (38), 9184–9185.
- (155) Rowe, C. J.; Böhm, I. U.; Thomas, I. P.; Wilkinson, B.; Rudd, B. A. M.; Foster, G.; Blackaby, A. P.; Sidebottom, P. J.; Roddis, Y.; Buss, A. D.; Staunton, J.; Leadlay, P. F. *Chem. Biol.* **2001**, *8* (5), 475–485.
- (156) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. *Nat. Protoc.* **2016**, *10* (6), 845–858.
- (157) Ma, S. M.; Tang, Y. *FEBS J.* **2007**, *274* (11), 2854–2864.
- (158) Wu, Z.; Li, D.; Zeng, F.; Tong, Q.; Zheng, Y.; Liu, J.; Zhou, Q. *Phytochemistry* **2018**, *156*, 159–166.
- (159) Stallard, M. O.; Fenical, W.; Kittredge, J. S. *Tetrahedron* **1978**, *34* (14), 2077–2081.
- (160) Caccamese, S.; Amico, V.; Neri, P. *J. Nat. Prod.* **1990**, *53* (5), 1287–1296.
- (161) Zinsmeister, H. D.; Mues, R. In *Bryophytes Their Chemistry and Chemical Taxonomy*; 1990; p 41.
- (162) Amico, V.; Caccamese, S.; Neria, P.; Russo, G.; Fotia, M. *Phytochemistry* **1991**, *30* (6), 1921–1927.
- (163) Wright, A. D.; König, G. M.; Sticher, O. *J. Nat. Prod.* **1991**, *54* (4), 1025–1033.
- (164) Tori, M.; Nakashima, K.; Seike, M.; Asakawa, Y.; Wright, A. D.; König, G. M.; Sticher, O. *Tetrahedron Lett.* **1994**, *35* (19), 3105–3106.
- (165) Melching, S.; König, W. A. *Phytochemistry* **1999**, *51* (4), 517–523.
- (166) Iliopoulou, D.; Vagias, C.; Galanakis, D.; Argyropoulos, D.; Roussis, V. *Org. Lett.* **2002**, *4* (19), 3263–3266.
- (167) Hu, Z.-Y.; Li, Y.-Y.; Huang, Y.-J.; Su, W.-J.; Shen, Y.-M. *Helv. Chim. Acta* **2008**, *91* (1), 46–52.
- (168) Chen, C. J.; Liu, X. X.; Zhang, W. J.; Zang, L. Y.; Wang, G.; Ng, S. W.; Tan, R. X.; Ge, H. M. *RSC Adv.* **2015**, *5* (23), 17559–17565.
- (169) Hu, D.; Zhang, S.; He, J.; Dong, Z.; Li, Z.; Feng, T.; Liu, J. *Fitoterapia* **2015**, *104*, 50–54.
- (170) Basnet, B. B.; Chen, B.; Suleimen, Y. M.; Ma, K.; Guo, S.; Bao, L.; Huang, Y.; Liu, H. *Planta Med.* **2019**, *85* (13), 1088–1097.
- (171) Murai, K.; Lauterbach, L.; Teramoto, K.; Quan, Z.; Barra, L.; Yamamoto, T.; Nonaka, K.; Shiomi, K.; Nishiyama, M.; Kuzuyama, T.; Dickschat, J. S. *Angew. Chemie - Int. Ed.* **2019**, *58* (42), 15046–15050.
- (172) Sun, X.; Cai, Y.; Yuan, Y.; Bian, G.; Ye, Z.; Deng, Z. *J. Org. Chem.* **2019**, *15*, 2052–2058.
- (173) Feng, J.; Zhang, P.; Cui, Y.; Li, K.; Qiao, X.; Zhang, Y.-T.; Li, S.-M.; Cox, R. J.; Wu, B.; Ye, M.; Yin, W.-B. *Adv. Synth. Catal.* **2017**, *359*, 995–1006.
- (174) Xie, K.; Dou, X.; Chen, R.; Chen, D.; Fang, C.; Xiao, Z.; Dai, J. *Appl. Environ. Microbiol.* **2017**, *83* (8), e03103-16.

- (175) Xie, L.; Zhang, L.; Wang, C.; Wang, X.; Xu, Y. ming; Yu, H.; Wu, P.; Li, S.; Han, L.; Gunatilaka, A. A. L.; Wei, X.; Lin, M.; Molnár, I.; Xu, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (22), E4980–E4989.
- (176) Xie, L.; Zhang, L.; Bai, J.; Yue, Q.; Zhang, M.; Li, J.; Wang, C.; Xu, Y. *J. Agric. Food Chem.* **2019**, *67* (31), 8573–8580.
- (177) Elshahawi, S. I.; Shaaban, K. A.; Kharel, K.; Thorson, J. S. *Chem. Soc. Rev.* **2015**, *44*, 7591.
- (178) Takahashi, S.; Takeuchi, M.; Arai, M.; Seto, H.; Otake, N. *J. Antibiot. (Tokyo)*. **1983**, *36* (3), 226–228.
- (179) Shigemori, H.; Komaki, H.; Yazawa, K.; Mikami, Y.; Nemoto, A.; Tanaka, Y.; Sasaki, T.; In, Y.; Ishida, T.; Kobayashi, J. N. I. *J. Org. Chem.* **1998**, *63* (20), 6900–6904.
- (180) Gebhardt, K.; Meyer, S. W.; Schinko, J.; Bringmann, G.; Zeeck, A.; Fiedler, H.-P. *J. Antibiot. (Tokyo)*. **2011**, *64*, 229–232.
- (181) Li, K.; Cai, J.; Su, Z.; Yang, B.; Liu, Y.; Zhou, X.; Huang, J.; Tao, H. *Front. Chem.* **2020**, *7*, 879.
- (182) Ma, X.; Li, L.; Zhu, T.; Ba, M.; Li, G.; Gu, Q.; Guo, Y.; Li, D. *J. Nat. Prod.* **2013**, *76* (12), 2298–2306.
- (183) Sakaki, H.; Kaneno, H.; Sumiya, Y.; Tsushima, M.; Miki, W.; Kishimoto, N.; Fujita, T.; Matsumoto, S.; Komemushi, S.; Sawabe, A. *J. Nat. Prod.* **2002**, *65* (11), 1683–1684.
- (184) Sato, H.; Hashishin, T.; Kanazawa, J.; Miyamoto, K.; Uchiyama, M. *J. Am. Chem. Soc.* **2020**, *142* (47), 19830–19834.
- (185) Wibberg, D.; Stadler, M.; Lambert, C.; Bunk, B.; Spröer, C.; Rückert, C.; Kalinowski, J.; Cox, R. J.; Kuhnert, E. *Fungal Divers.* **2020**, No. 0123456789.
- (186) Stanke, M.; Diekhans, M.; Baertsch, R.; Haussler, D. *Bioinformatics* **2008**, *24* (5), 637–644.
- (187) Ter-Hovhannisyan, V.; Lomsadze, A.; Chernoff, Y. O.; Borodovsky, M. *Genome Res.* **2008**, *18* (12), 1979–1990.
- (188) Meyer, F.; Goesmann, A.; McHardy, A. C.; Bartels, D.; Bekel, T.; Clausen, J.; Kalinowski, J.; Linke, B.; Rupp, O.; Giegerich, R.; Pühler, A. *Nucleic Acids Res.* **2003**, *31* (8), 2187–2195.
- (189) Rupp, O.; Becker, J.; Brinkrolf, K.; Timmermann, C.; Borth, N.; Pühler, A.; Noll, T.; Goesmann, A. *PLoS One* **2014**, *9* (1).
- (190) Gietz, R. D.; Schiestl, R. H. *Nat. Protoc.* **2007**, *2* (1), 35–37.
- (191) Bok, J. W.; Keller, N. P. *Methods Mol Biol.* **2012**, *944*, 163–174.

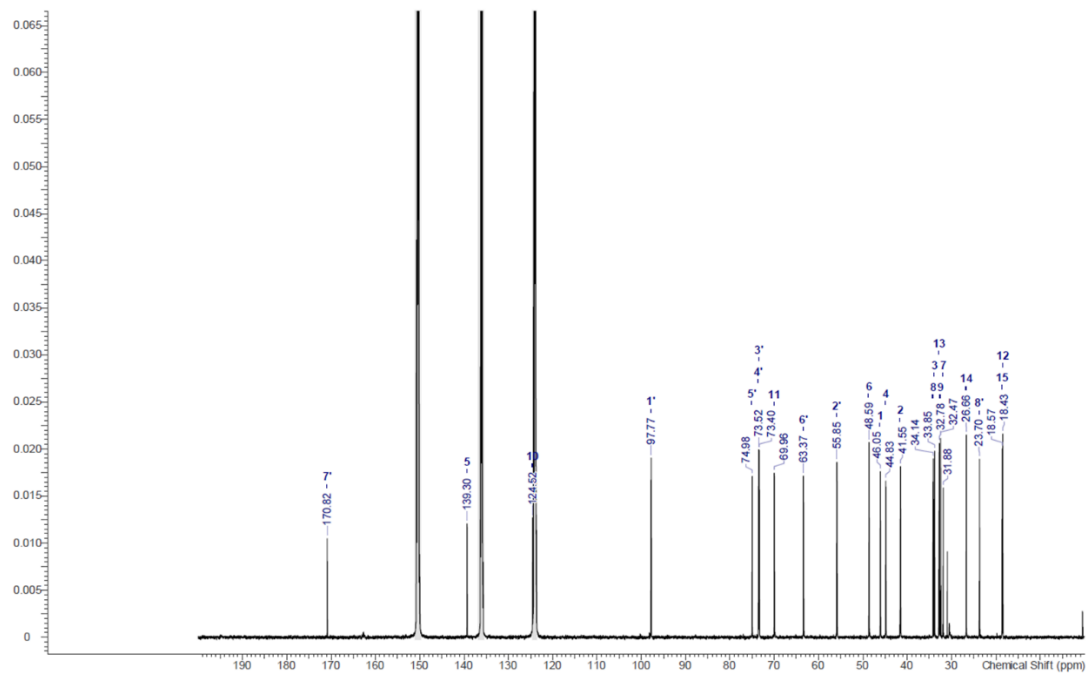
# Appendix

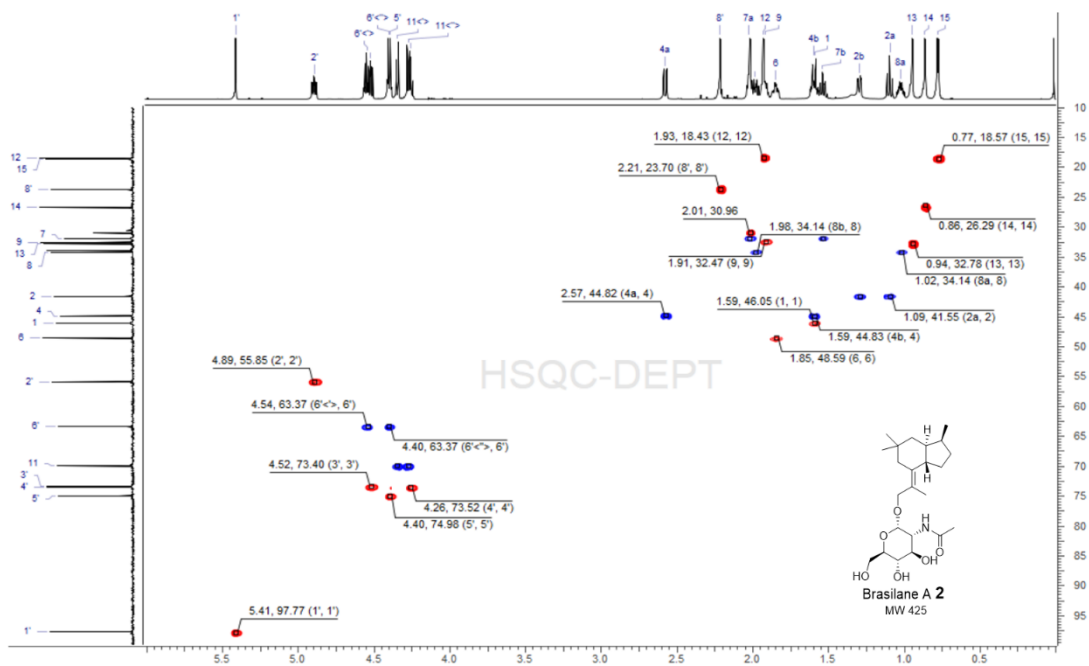
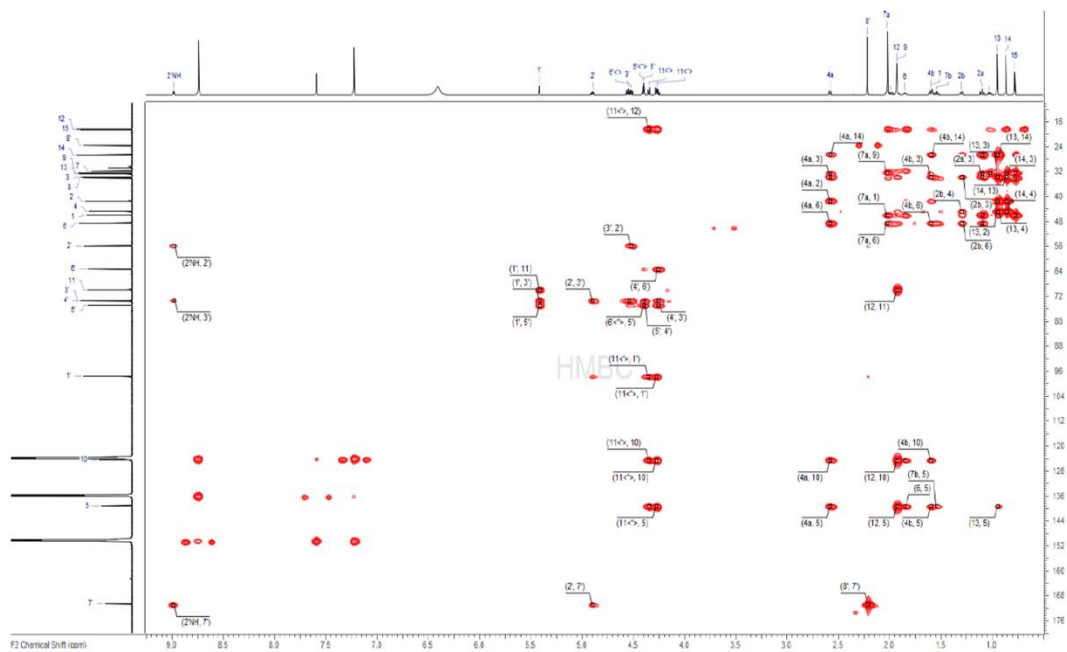
## NMR Spectra

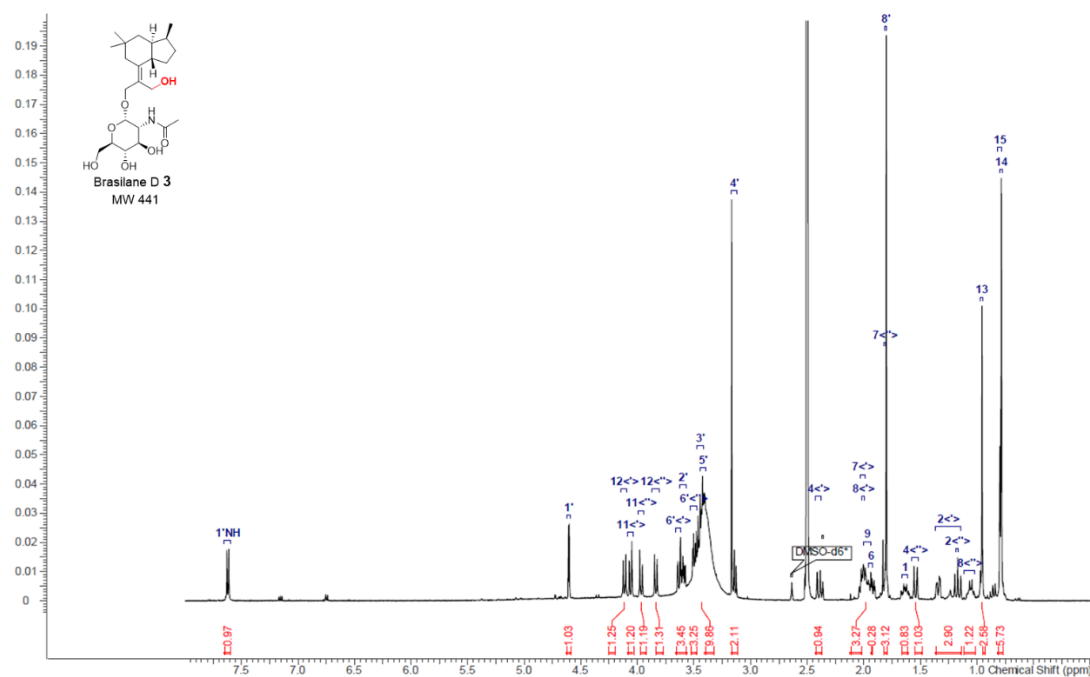
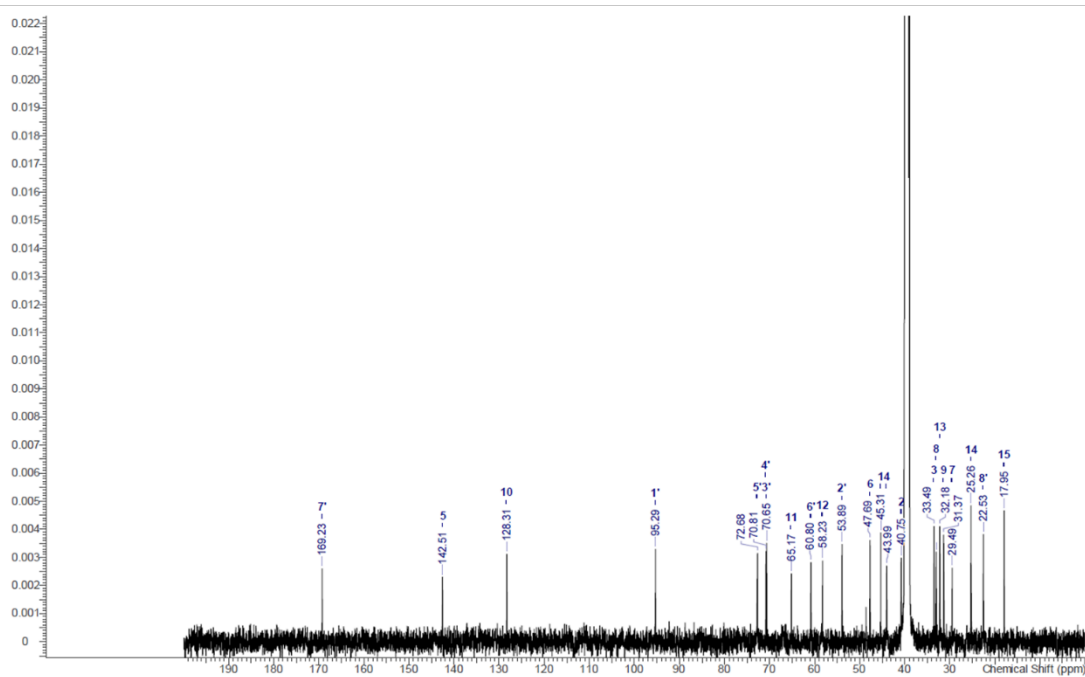
$^1\text{H}$  NMR spectrum (700 MHz, pyridine- $d_5$ ) of brasilane A 2.

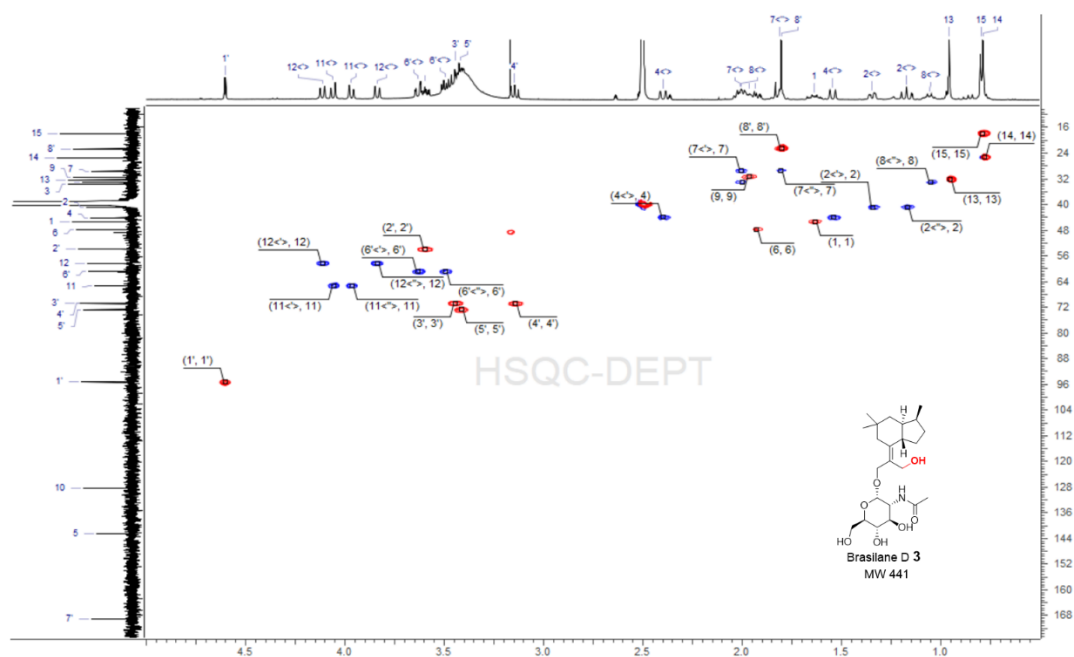
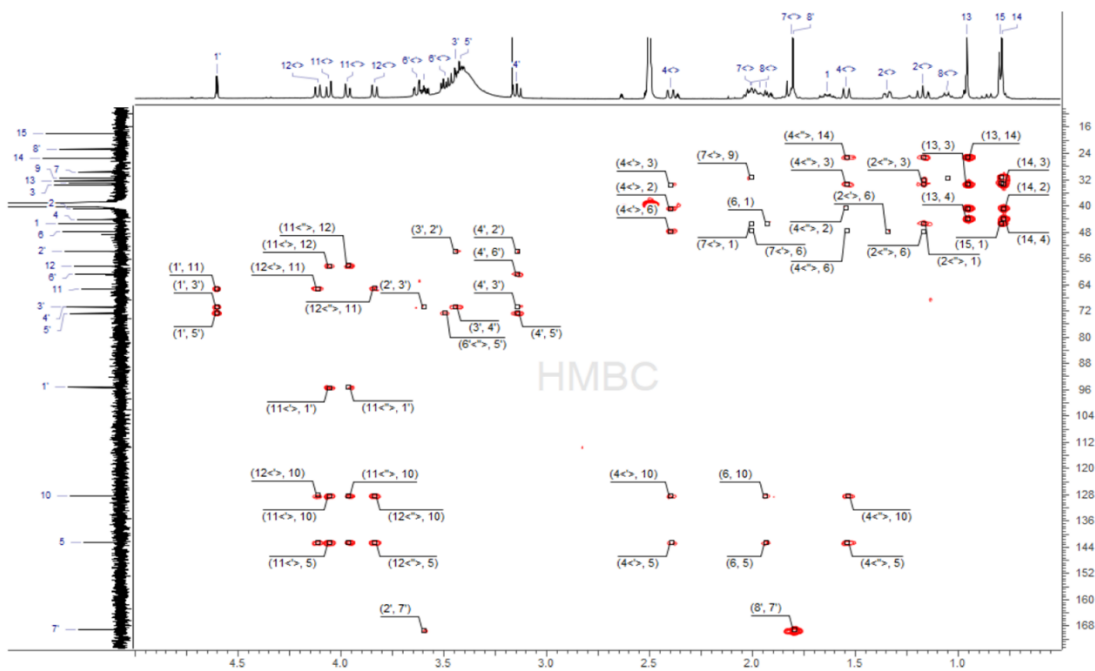


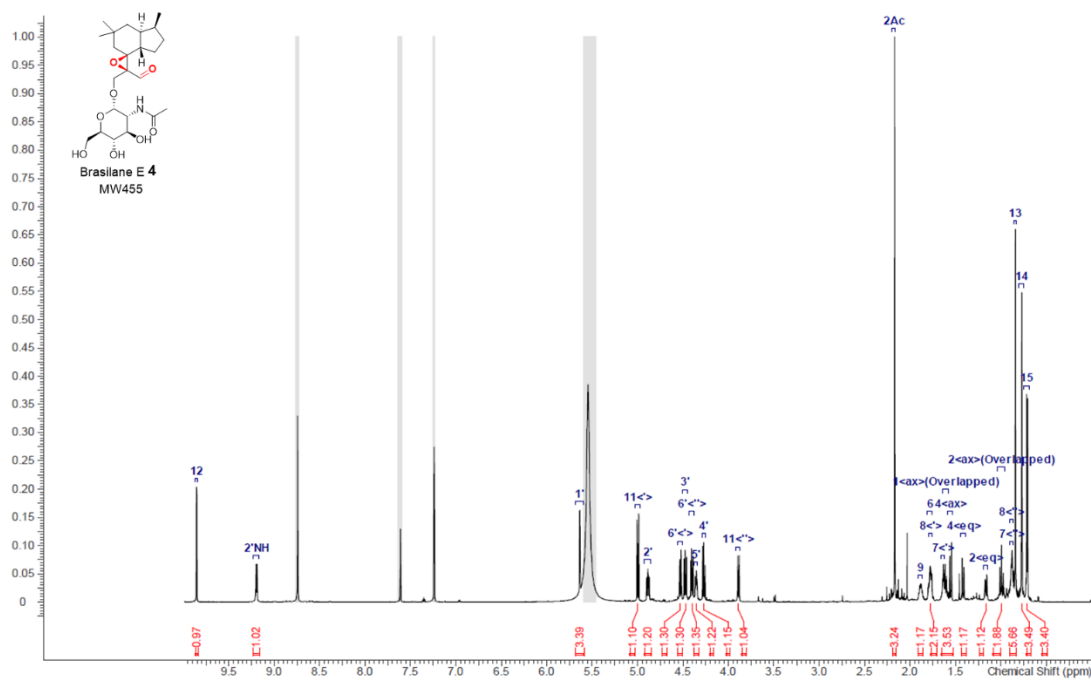
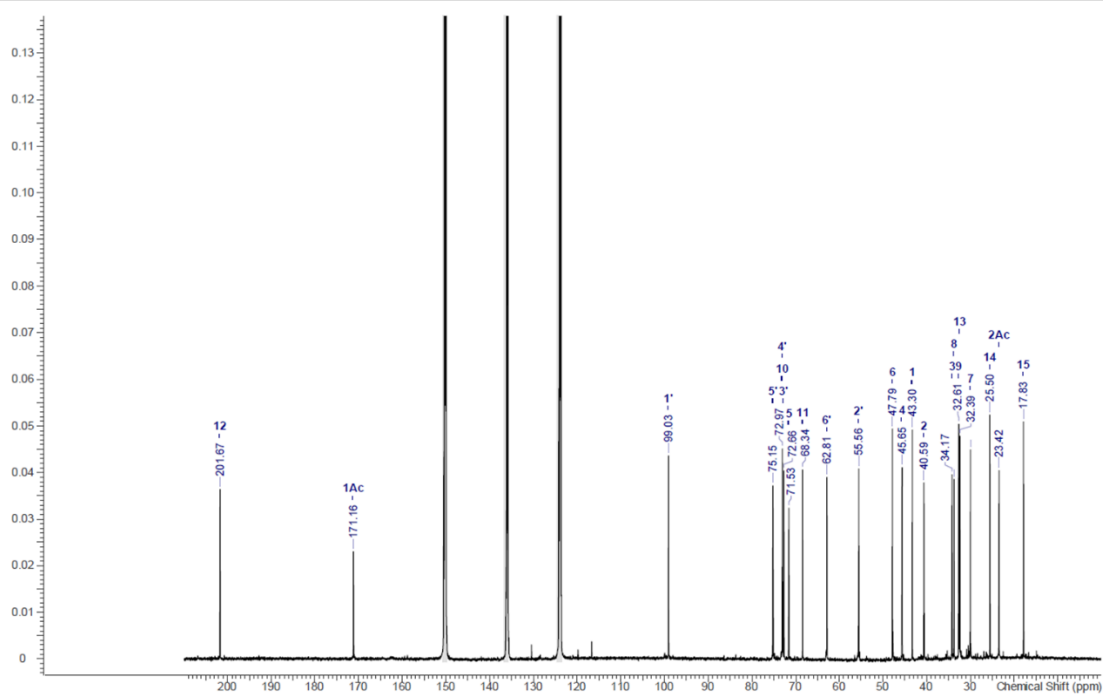
$^{13}\text{C}$  NMR spectrum (125 MHz, pyridine- $d_5$ ) of brasilane A 2.



HSQC NMR spectrum (500 MHz, pyridine- $d_5$ ) of brasilane A 2.HMBC NMR spectrum (700 MHz, pyridine- $d_5$ ) of brasilane A 2.

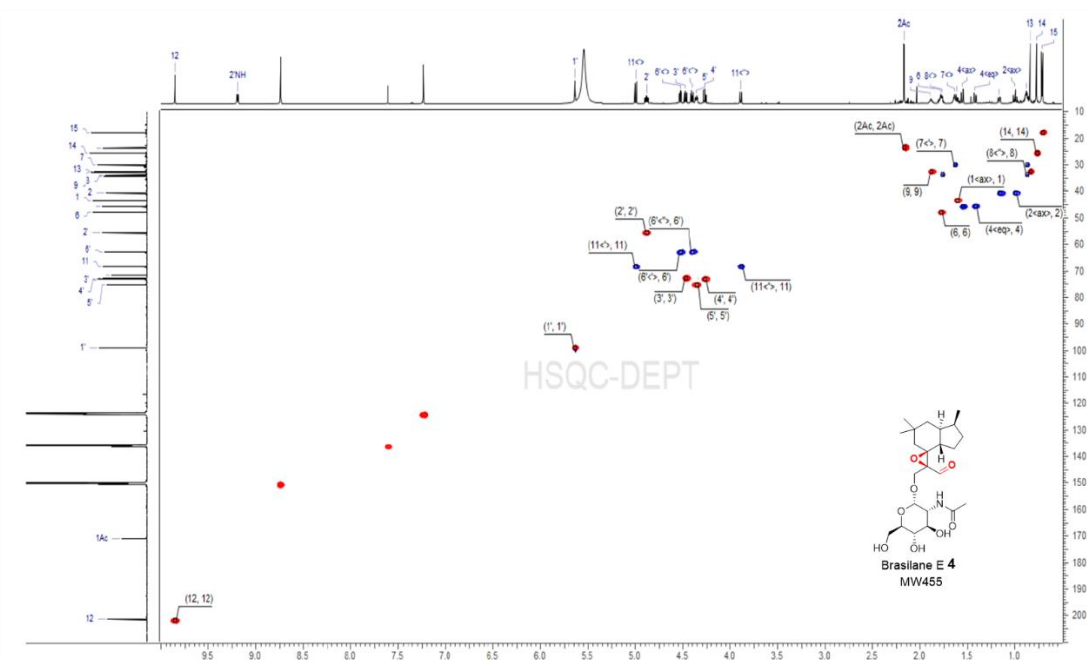
<sup>1</sup>H NMR spectrum (500 MHz, DMSO-*d*<sub>6</sub>) of brasilane D 3.<sup>13</sup>C NMR spectrum (125 MHz, DMSO-*d*<sub>6</sub>) of brasilane D 3.

HSQC NMR spectrum (500 MHz, DMSO- $d_6$ ) of brasilane D 3.HMBC NMR spectrum (500 MHz, DMSO- $d_6$ ) of brasilane D 3.

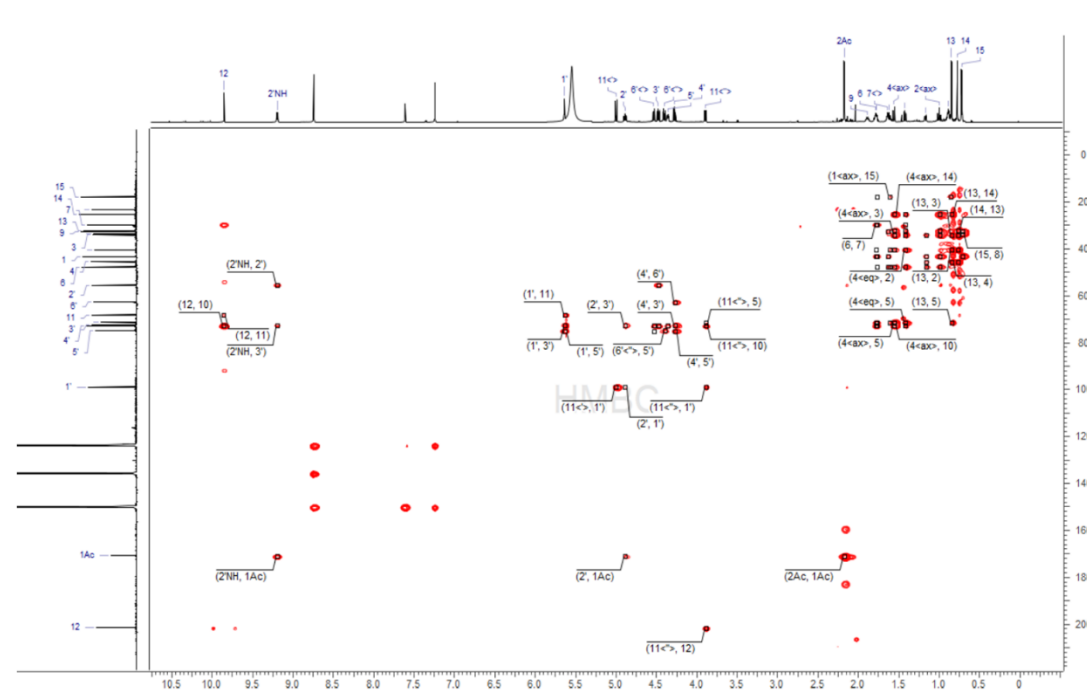
<sup>1</sup>H NMR spectrum (500 MHz, pyridine-*d*<sub>5</sub>) of brasilane E 4.<sup>13</sup>C NMR spectrum (175 MHz, pyridine-*d*<sub>5</sub>) of brasilane E 4.

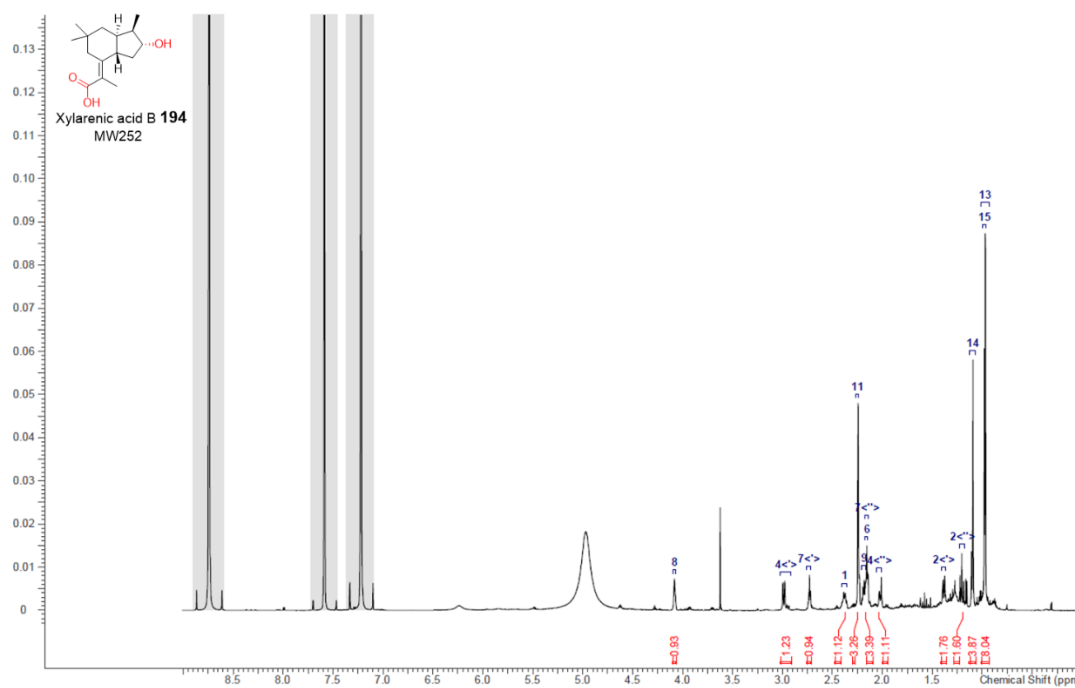
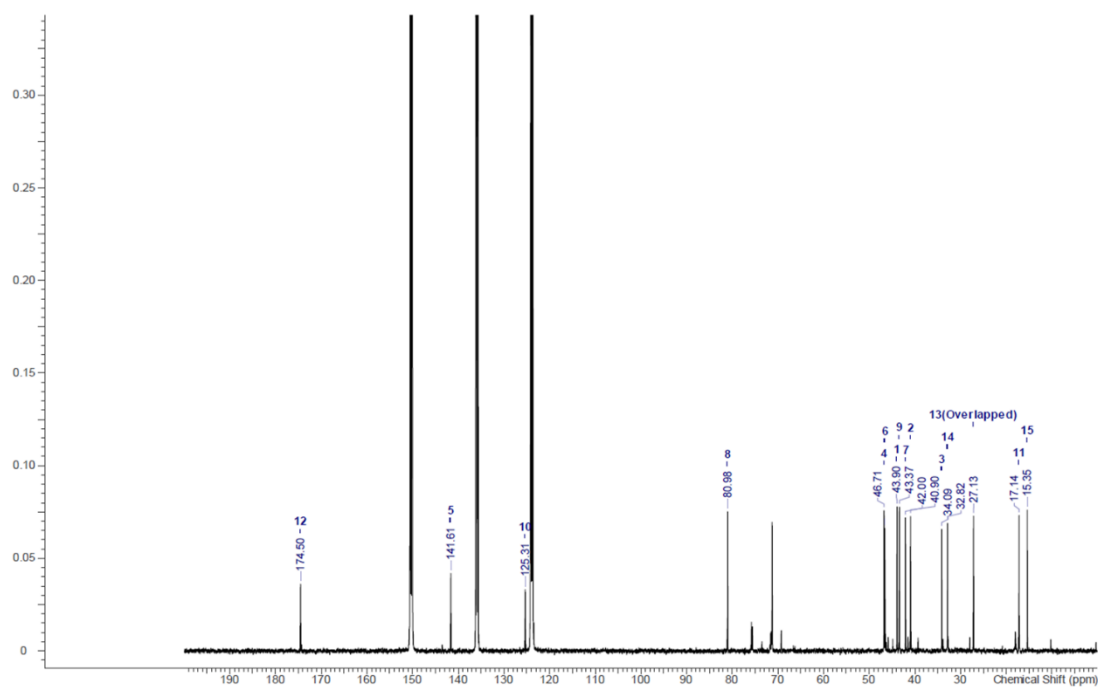


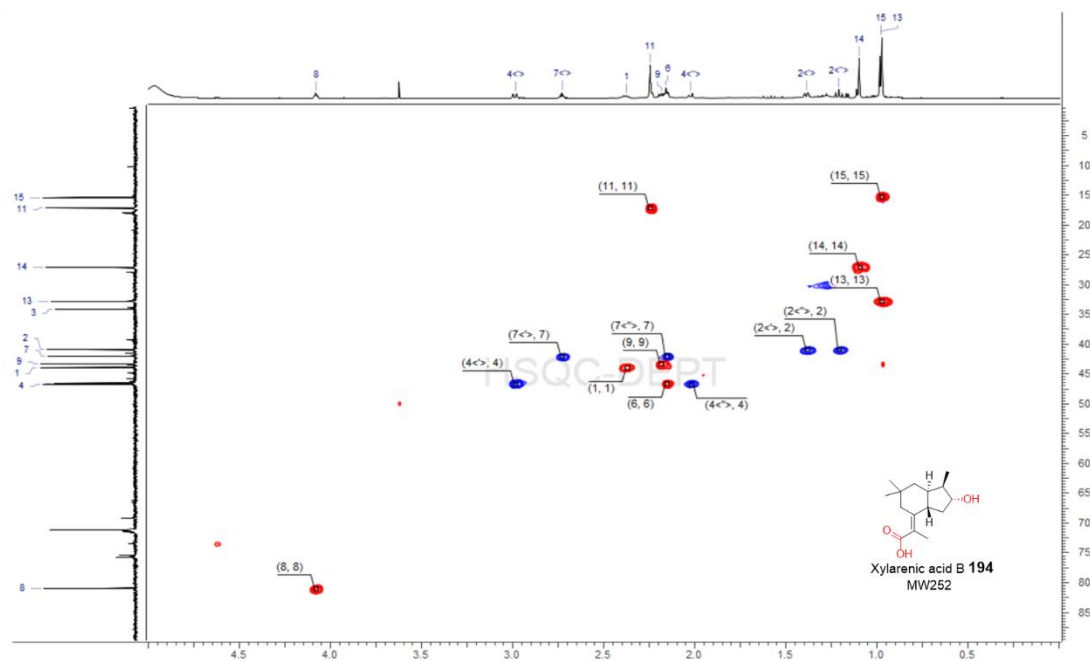
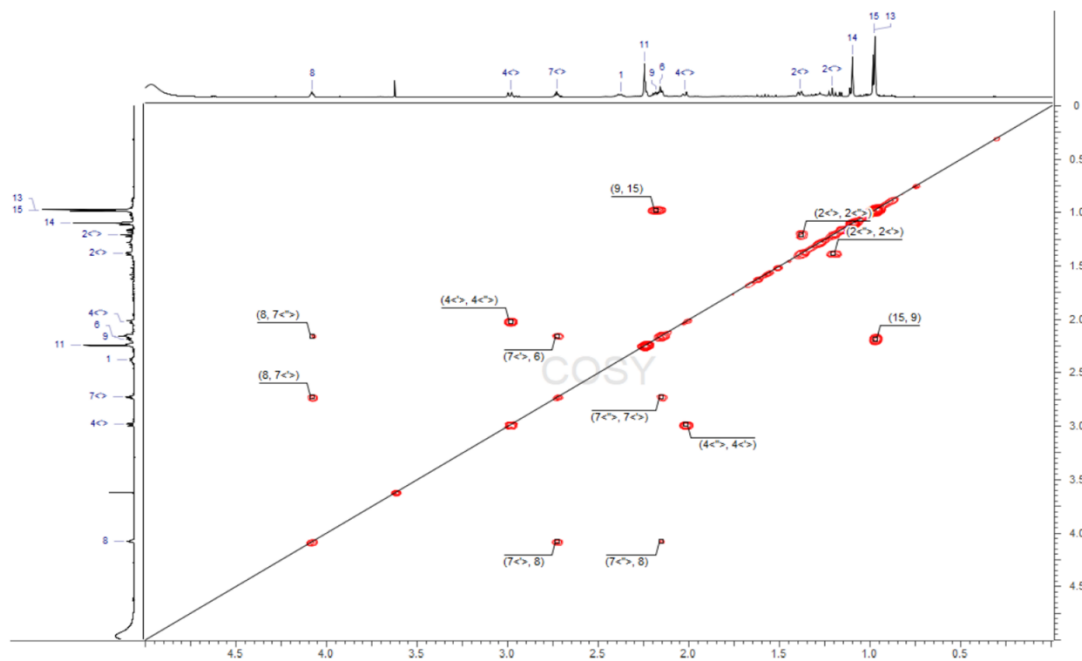
HSQC NMR spectrum (700 MHz, pyridine-*d*<sub>5</sub>) of brasilane E 4.

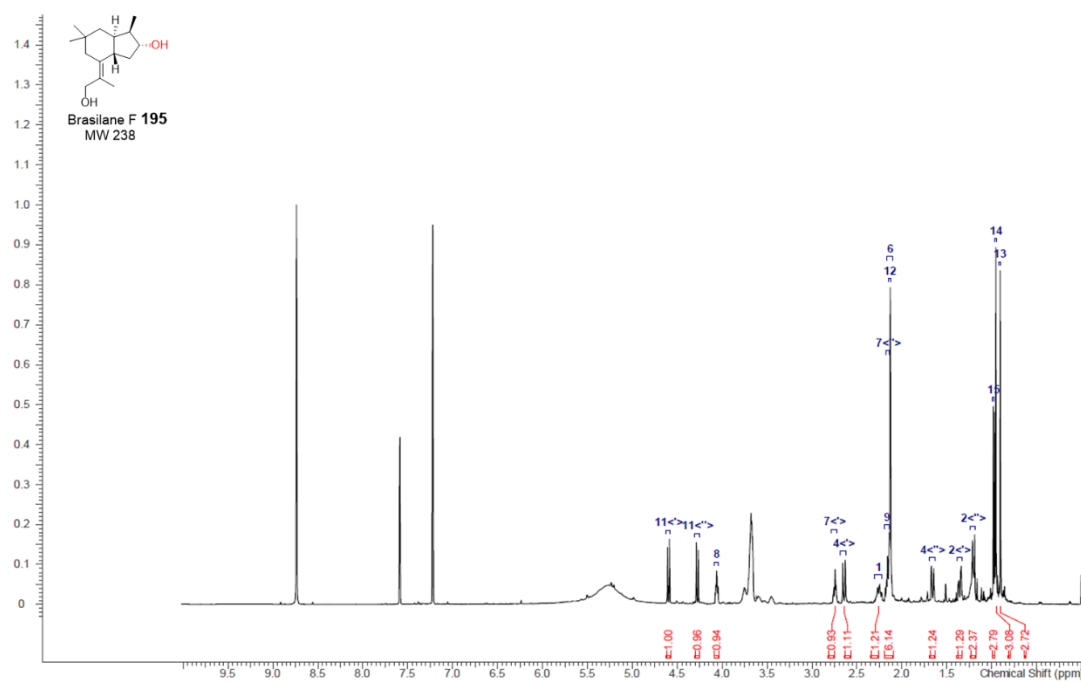
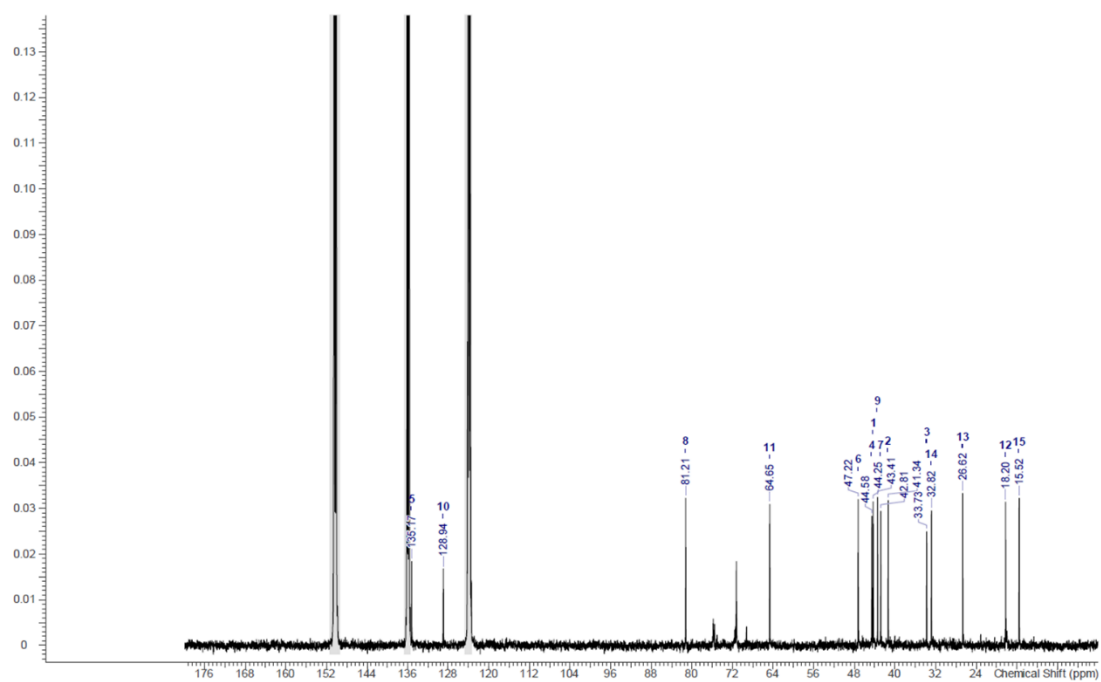


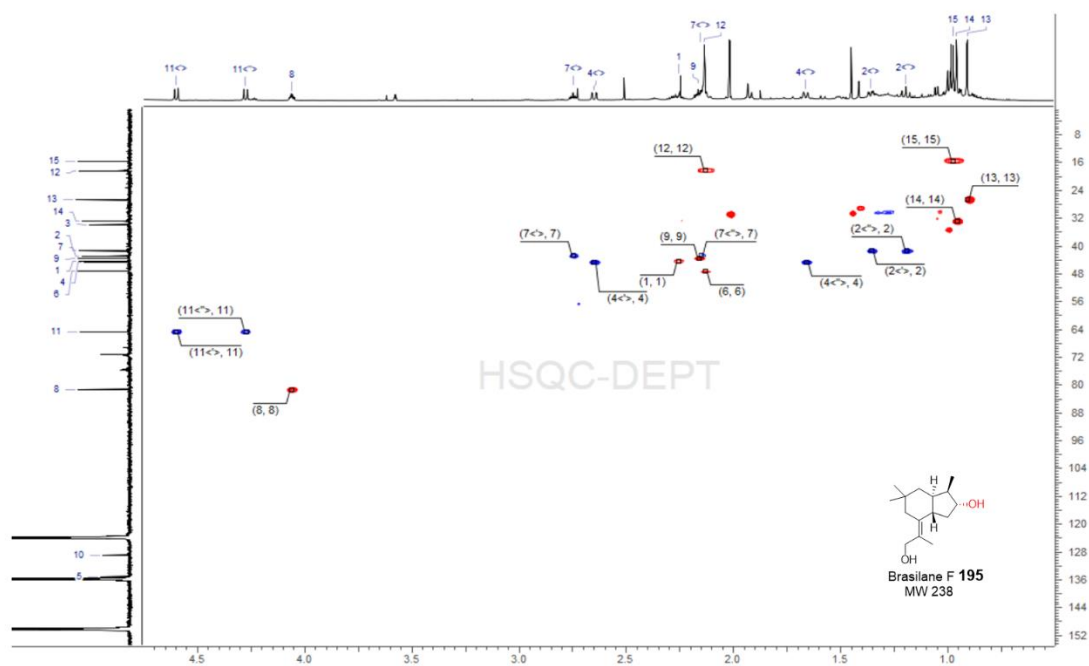
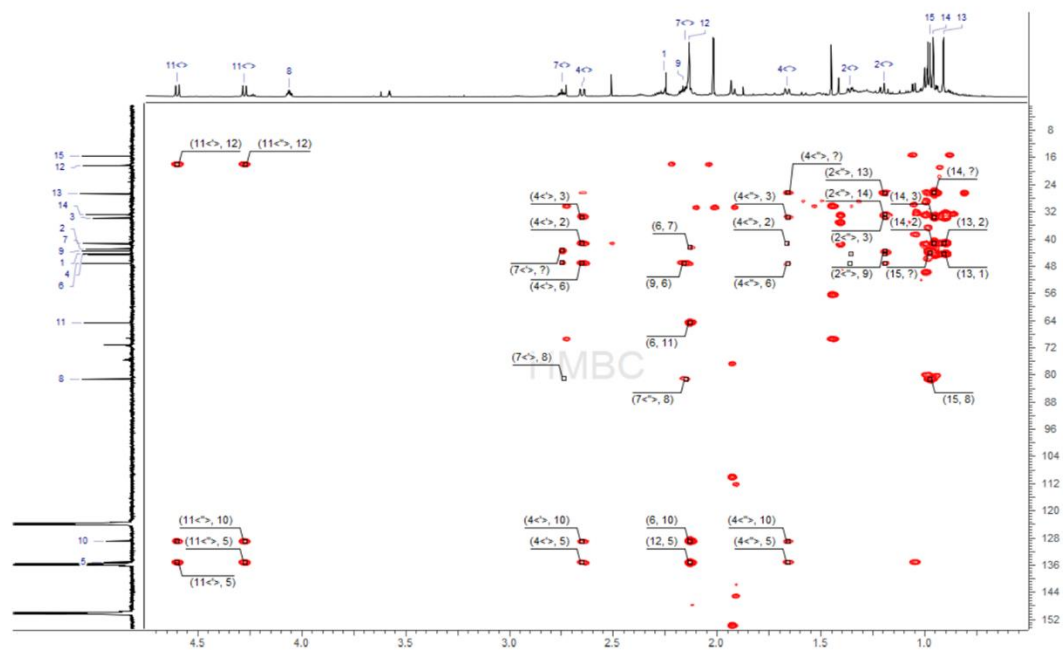
HMBC NMR spectrum (700 MHz, pyridine-*d*<sub>5</sub>) of brasilane E 4.

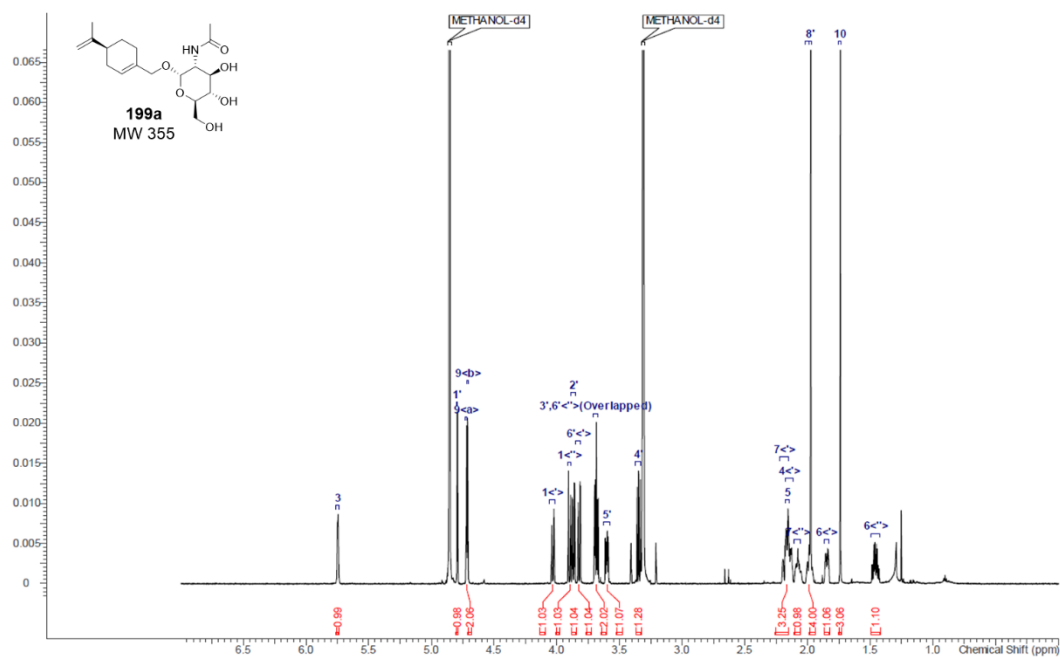
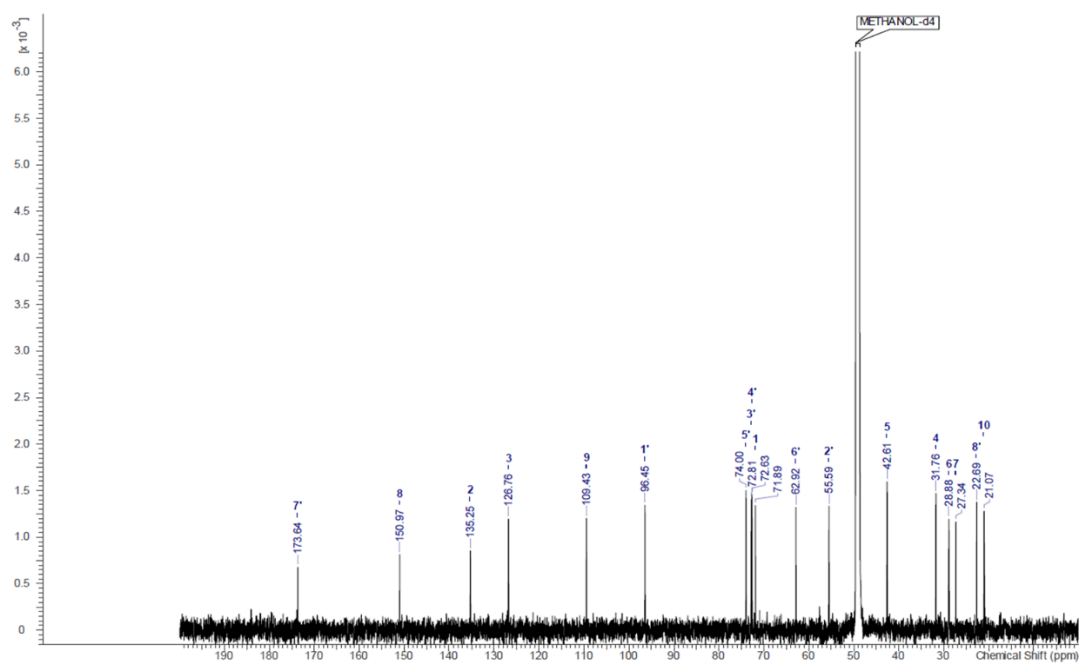


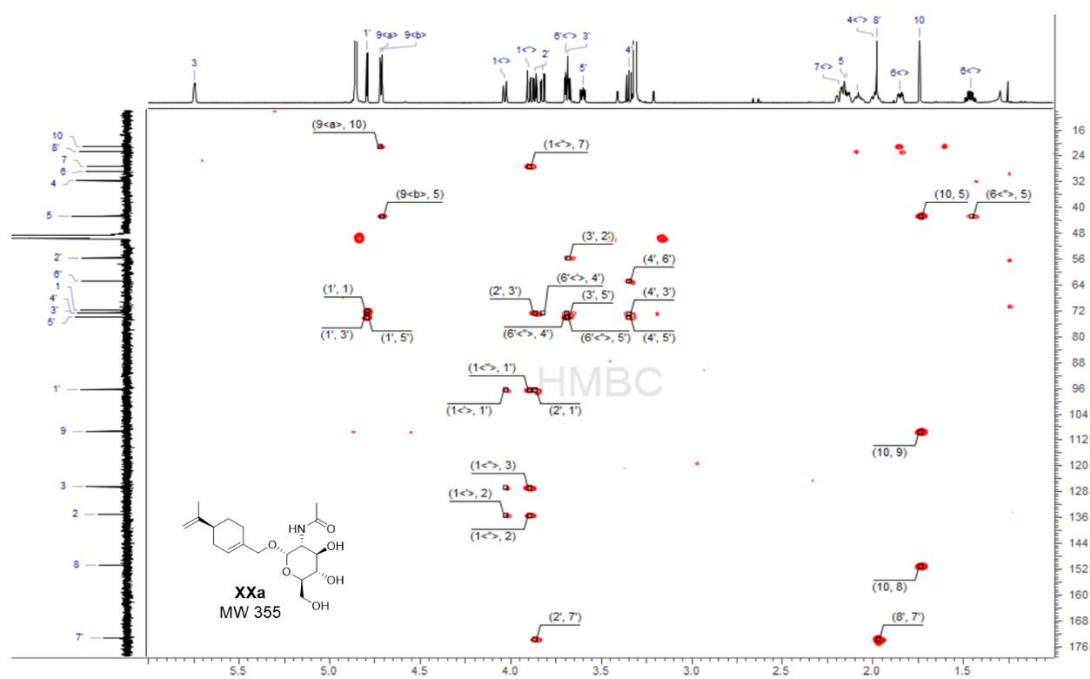
<sup>1</sup>H NMR spectrum (700 MHz, pyridine-*d*<sub>5</sub>) of xylarenic acid B **194**.<sup>13</sup>C NMR spectrum (125 MHz, pyridine-*d*<sub>5</sub>) of xylarenic acid B **194**.

HSQC NMR spectrum (700 MHz, pyridine- $d_5$ ) of xylarenic acid B **194**.COSY NMR spectrum (700 MHz, pyridine- $d_5$ ) of xylarenic acid B **194**.

<sup>1</sup>H NMR spectrum (500 MHz, pyridine-*d*<sub>5</sub>) of brasilane F **195**.<sup>13</sup>C NMR spectrum (125 MHz, pyridine-*d*<sub>5</sub>) of brasilane F **195**.

HSQC NMR spectrum (500 MHz, pyridine- $d_5$ ) of brasilane F **195**.HMBC NMR spectrum (500 MHz, pyridine- $d_5$ ) of brasilane F **195**.

$^1\text{H}$  NMR spectrum (500 MHz, methanol- $d_4$ ) of **199a**. $^{13}\text{C}$  NMR spectrum (175 MHz, methanol- $d_4$ ) of **199a**.

HMBC NMR spectrum (500 MHz, methanol- $d_4$ ) of **199a**.

## Curriculum Vitae

Oct 2017 – Present	<b>PhD student</b> , Centre for Biomolecular Drug Research (BMWZ), Institute of Organic Chemistry, Gottfried Wilhelm Leibniz Universität Hannover, Germany
Aug 2016 – Jul 2017	<b>Research Assistant</b> , State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, China
Sep 2012 – Jul 2016	<b>Master of Science</b> , Department of Natural Medicines, School of Pharmaceutical Sciences, Peking University, China
Sep 2007 – Jun 2011	<b>Bachelor of Science</b> , School of Traditional Chinese Pharmacy, China Pharmaceutical University, China

## List of Publications

**Jin Feng**, Frank Surup, Maurice Hauser, Anna Miller, Jan-Peer Wennrich, Marc Stadler, Russell J. Cox, Eric Kuhnert. *Chem. Commun.* 2020, 56(82), 12419-12422.

**Jin Feng**, Peng Zhang, Ying-Lu Cui, Kai Li, Xue Qiao, Ying-Tao Zhang, Shu-Ming Li, Russell J. Cox, Bian Wu, Min Ye and Wen-Bing Yin. *Adv. Synth. Catal.* 2017, 359(6), 995-1006.

**Jin Feng**, Wen-Fei Liang, Shuai Ji, Xue Qiao, Ying-Tao Zhang, Si-Wang Yu and Min Ye. *J. Chin. Pharm. Sci.* 2015, 24, 285-291.

Xinran Xu, **Jin Feng**, Peng Zhang, Jie Fan, Wen-Bing Yin. *J. Microbiol. Biotechnol.* 2021, 31(1), 8-15.

Kai Li, **Jin Feng**, Yi Kuang, Wei Song, Meng Zhang, Shuai Ji, Xue Qiao, Min Ye. *Adv. Synth. Catal.* 2017, 359(21), 3765-3772.

Shuai Ji, Wen-Fei Liang, Zi-Wei Li, **Jin Feng**, Qi Wang, Xue Qiao and Min Ye. *RSC Adv.* 2016, 6, 20791–20799.

Wei Song, Xue Qiao, Kuan Chen, Ying Wang, Shuai Ji, **Jin Feng**, Kai Li, Yan Lin and Min Ye. *Anal. Chem.* 2017, 89, 3146–3153.