

Functional Characterization of Lichen Fungal (*Cladonia uncialis*) Genes and Exploration of Its Secondary Metabolome.

By

**Navriti Mittal**

A Thesis submitted to the Faculty of Graduate Studies  
of The University of Manitoba in partial fulfillment  
of the requirements for the degree of

**DOCTOR OF PHILOSOPHY**

Department of Chemistry  
Faculty of Science  
University of Manitoba  
Winnipeg, Manitoba, Canada

November 2022

© Copyright 2022, Navriti Mittal

# Front Matter

## Abstract

Lichens are traditionally described as symbionts of fungi and algae and are renowned for their diverse secondary metabolites. Our group identified and annotated 48 secondary metabolite biosynthetic gene clusters of the fungal partner of the lichen *Cladonia uncialis* using de novo whole genome sequencing. Out of these, putative assignment of biosynthetic functions of ‘ten’ gene clusters in *C. uncialis* was done by Bertrand (2018) including the usnic acid (UA) gene cluster (the most extensively studied lichen secondary metabolite). This research work deduced biosynthetic pathways and proposed biosynthetic function of seven more gene clusters using a ‘homology mapping’ approach in combination with phylogenetics. The UA gene cluster contains two genes, one encodes for PKS enzyme and other for a post PKS tailoring enzyme, that have been respectively named as methylphloracetophenone synthase (MPAS) and methylphloroacetophenone oxidase (MPAO, a cytochrome p450). MPAO is believed to catalyze the oxidative dimerization of methylphloroacetophenone (MPA) to usnic acid (UA). Electron transfer reactions from NAD(P)H to a cytochrome p450 are supported by redox partner, cytochrome P450 oxidoreductase (CPR). We have identified a gene, *cu-cpr* in the genome of *C. uncialis* that encodes for a similar redox partner. We have successfully developed heterologous expression protocols for first lichen cytochrome p450 (*mpao*) and CPR (*cu-cpr*) in *E. coli* and, the purification protocols for these lichen proteins. We have also established a protocol for heterologous co-expression system for *mpao* and *cpr* in *E. coli*. This study demonstrates the first successful biosynthesis of UA (a lichen polyketide) with the development of bioconversion protocol of MPA to UA catalyzed by both MPAO and CPR. Biofilm-disruption and antibacterial assays were carried out to compare the bioactivity of MPA and UA, where UA was found to be more bioactive. That could suggest a rationale for why lichen fungus exerts the effort required to produce UA. Phloroglucinol derivatives (including MPA derivatives) were chemically synthesized and tested for their antibiotic properties against *S. aureus* (an important CF pathogen). Out of which, diacylated derivatives displayed the strongest bactericidal activity against MRSA clinical isolates.

## Acknowledgements

There are many people to whom I owe the success of this research program. I would not be here where I am standing today with successfully completed the entire PhD program without the indelible efforts of my Supervisor, Dr. John Sorensen. That brought me ample of opportunities to learn, grow and develop myself beyond my potential. I thank him immensely for his mentorship and patience throughout this roller coaster journey of ‘7’ years (It was not easy to keep up with a student like me). I thank my committee members, Teresa De Kievit (have always been so kind & motivating), Helene Perreault, and Rebecca Davis, for their encouragement and guidance, and for helping me pass through several critical junctures of my doctoral program. I also extend my thanks to our collaborators, late Tuhinadri Sen and Ratul Sarkar for their significant contribution in assessing antibacterial & biofilm disruption activities of UA & MPA. And my thanks to our research collaborator, Silvia Cardona and her graduate students, Andrew Hogan & Haben Tesfu for their contribution in characterizing antibiotic properties of ‘12’ chemically synthesized phloroglucinols against *S. aureus*, an important CF pathogen. A huge thanks to Hamid Karbalaieheidari & Evan Booy, of the Department of Chemistry, who are thoroughly knowledgeable in biochemistry and molecular biology techniques. Because of them I learnt many of the molecular biology techniques with having me coming from a pure chemistry, physics & mathematics background w/o any prior knowledge of molecular biology. Evan Booy service to the chemistry department by operating the ‘BioBar’ has also allowed me to conduct my research at a much faster pace. I also extend my thanks to Manu Singh, Tasia Lightly, Brijesh Kumar, Amit Kaul, Adis, Mona Abdel-Hameed & Robert Bertrand ex-PhD students of the Department of Microbiology & Department of Chemistry, who all helped me in many ways in my molecular biology project and their advice allowed me to overcome several frustrating obstacles during my research program. Manu Singh, who is well-versed in bioinformatics and genome assembly, and I am very thankful for his guidance this important phase of my project. I also wish to thank Sean Mckenna, also from the Department of Chemistry, for kindly providing me with access to his expensive lab equipment (for molecular biology work). A special thanks to Kirk Marat, who was NMR facility manager at Department of Chemistry when I started my PhD project, and I learned a lot about NMR from him. Andrew Laluk, of the Sorensen group, is an aspiring organic chemist who kindly prepared for me a chemical compound (MPA) when I was out of this molecule (what

I prepared) during usnic acid bioconversion experiments, and I thank him for this contribution. I am also very thankful to all my wonderful lab-mates (Harman, Randy, Alex, Susan, April, Edwin, Tanya, Mila, Scott, Sabina, Vienna, Harley, Patricia) in all these years who were the important of this PhD journey and had a significant contribution to it. I have been lucky to be able to lean upon Tannis Wills, Keith Travis & Tricia Lewis, the academic programs administrator in the Department of Chemistry, for helping me navigate through the many administrative concerns that inevitably arise within this line of work. I have had a good experience as a teaching assistant in the Department of Chemistry. I thank Horace Luong, senior instructor for the organic chemistry teaching laboratories, for his inexhaustible patience and his sincere desire to develop his teaching assistants. I have become a better instructor because of him. Many thanks to all my friends from India & friends that I have made here in Winnipeg who were/have been there for me through thick & thin, to cry with me for failing my experiments or to encourage me not to give up or to celebrate my very rare happy moments of getting results in the lab or to drive me to the lab at midnight to start a bacterial culture or to pick me up early in morning after spending a sleepless night in the lab, to listen to practice talks of all my research presentations (where they did not understand a word) & now they all know what a flash column chromatography is. I love you guys- Rajbir, Deepak, Muntahi, Preeti, Manu, Gulshan, Mitali, Sakshi, Ishwinder, Anshika, Rishikesh, Vikram, Param, Vimal, Divyani, Chandrika, Navrin, Karan, Ankita, Pavan and thank you so much for being my family when I was away from my own family in Canada. And at last, the biggest thanks to my loveliest family, my strength & the pillars of my life: Mamma, Papa (every girl should have a father like you), my darling sister (Guddu), my sweetheart (my brother Kaku), my aunt (Ritu Massi), my uncles (Arun & Babbal Mamaji)- my world! There is no word in the dictionary to express my love & gratitude for you all. Y'all have always motivated me no matter what, have shaped me to be a better version of myself & thank you so much for all your unconditional support & love. I feel God is with me in forms of my family, friends & my mentors!



## Dedication

To all the graduate students who come from a humble background, specially from the developing countries and leave their friends & families behind to pursue their careers in developed countries without any physical, mental & emotional support (all by themselves)!!

## Contribution of Authors

**Chapter 1:** *Non-applicable*

**Chapter 2:** *Non-applicable*

**Chapter 3:** Ratul Sarkar is the primary author of this published research work, who performed the biofilm-disruption and antibacterial assays to compare the bioactivity of MPA & UA. And I contributed towards the chemical synthesis of MPA as it is not commercially available. I am the primary contributor of other all presented work within this chapter.

**Chapter 4:** Haben Tesfu and Andrew Hogan tested ‘12’ phloroglucinol derivatives (including MPA derivatives) for their antibiotic properties against *S. aureus* (an important CF pathogen), where I carried out the chemical synthesis of these ‘12’ phloroglucinol derivatives. I am the primary contributor of all other presented work within this chapter.

**Chapter 5:** I am the primary contributor of the presented work within this chapter.

**Chapter 6:** I am the primary contributor of the presented work within this chapter.

**Chapter 7:** *Non-applicable*

# Contents

<b>Front Matter .....</b>	<b>II</b>
<b>Abstract.....</b>	<b>II</b>
<b>Acknowledgements .....</b>	<b>III</b>
<b>Dedication.....</b>	<b>V</b>
<b>Contribution of Authors .....</b>	<b>VI</b>
<b>List of Figures.....</b>	<b>X</b>
<b>List of Tables .....</b>	<b>XVI</b>
<b>List of Supporting Information Items .....</b>	<b>XVII</b>
<b>Non-IUPAC Abbreviations .....</b>	<b>XIX</b>
<b>Chapter 1 .....</b>	<b>1</b>
<b>Introduction to lichens and their secondary metabolites/natural products.....</b>	<b>1</b>
1.1. Natural Products .....	1
1.2. Primary versus secondary metabolism .....	3
1.3. Lichens Natural Products .....	5
1.4. Polyketides .....	15
1.5. Biosynthetic gene clusters (BGCs).....	22
1.6. Biosynthesis of Aromatic/Non-reducing Polyketides .....	23
1.6.3. Background on cytochrome p450s .....	27
1.6.3d. Reactions catalyzed by CYPs other than hydroxylation.....	32
1.7. Proposed mechanisms of cytochrome p450 .....	35
1.8. Biosynthesis of Reducing Polyketides .....	44
1.9. Terpenes .....	49
1.10. Genome sequencing, BGCs and mining of novel SMs .....	50
1.11. Problems associated with lichens and the production of their SMs .....	54
1.12. Summary of this research project objectives.....	56
References .....	58
<b>Chapter 2 .....</b>	<b>78</b>
<b>Materials and methods .....</b>	<b>78</b>
2.1. Materials and methods related to Chapter 1 .....	78
2.2. Materials and methods related to Chapter 2 .....	78
2.3. Materials and methods related to Chapter 3 .....	78
2.4. Materials and methods related to Chapter 4 .....	83
2.5. Materials and methods related to Chapter 5 .....	87
2.6. Materials and methods related to Chapter 6 .....	100
2.7. Materials and methods related to Chapter 7 .....	101
References .....	102
<b>Chapter 3 .....</b>	<b>107</b>

<b>Identification of the UA gene cluster, its biosynthetic pathway and chemical synthesis of MPA and bioassays of MPA v/s UA. ....</b>	<b>107</b>
3.1. Introduction .....	107
3.2. History of biosynthesis of UA .....	109
3.3. Steps towards genome annotation of <i>C. uncialis</i> .....	110
3.4. Identification of UA gene cluster .....	112
3.5. Biosynthesis of UA .....	114
3.6. Chemical synthesis of MPA .....	115
3.7. Bio-film disruption activity, MPA v/s UA .....	116
3.8. Summary .....	121
References .....	122
<b>Chapter 4 .....</b>	<b>129</b>
<b>Synthesis of novel MPA analogues, acylated phloroglucinols and their antibiotic activity against methicillin-resistant <i>Staphylococcus aureus</i> .....</b>	<b>129</b>
4.1. Introduction .....	129
4.2. Synthesis of phloroglucinol compounds .....	131
4.3. Evaluation of antibiotic properties .....	133
4.4. Summary .....	137
References .....	138
<b>Chapter 5 .....</b>	<b>140</b>
<b>Functional heterologous expression of lichen cytochrome p450 (CYP) and p450-reductase (CPR) in bacteria responsible for UA biosynthesis .....</b>	<b>140</b>
5.1. Introduction .....	140
5.2. An Outline of all experiments for UA biosynthesis described in this chapter .....	141
5.3. Importance and the goal of this study .....	143
5.4. Why biocatalysts over chemical synthesis .....	144
5.5. Advantages of using <i>E. coli</i> as a heterologous host (Figure 5.2) .....	144
5.6. Challenges of using a prokaryotic host to express eukaryotic proteins (Figure 5.2) .....	145
5.7. Strategies to improve and achieve a high expression of Eukaryotic P-450s in bacteria .....	146
5.8. Co-expression of redox partners (RPs) .....	151
5.9. Significance of cytochrome p450 reductase (the regulating effect of RPs) .....	153
5.10. Classification of CYPs based on their interaction with RPs .....	153
5.11. Heterologous expression of native MPAO (1554 bp) in <i>E. coli</i> .....	154
5.12. Heterologous expression of MPAO (w/o TMD &, codon-optimized for <i>E. coli</i> ) .....	159
5.12. Attempt towards bioconversion of MPA to UA catalyzed by MPAO .....	163
5.13. Cloning and heterologous expression of Redox proteins .....	166
5.14. Attempt towards bioconversion of MPA to UA catalyzed by both MPAO & CPR357 .....	170
5.15. Attempts toward the heterologous expression of MPAO using pET28b vector .....	173
5.16. Attempts toward co-expression of MPAO & CPR357 in <i>E. coli</i> .....	175
5.17. Attempts toward the heterologous expression of Fungal p450s .....	177
5.18. Proteins co-expression experiments in <i>E. coli</i> .....	182
5.19. Bioconversion experiments of MPA to UA .....	184
5.20. LC-MS analysis .....	189
5.21. Summary .....	193
References .....	195
<b>Chapter 6 .....</b>	<b>206</b>
<b>Assignment of Biosynthetic pathways to Lichen Gene Clusters .....</b>	<b>206</b>
6.1. Introduction .....	206
6.2. Terpene synthase gene clusters .....	209

6.3. Non-reducing polyketide synthase gene clusters.....	214
6.4. Reducing polyketide synthase gene clusters .....	214
6.5. Summary.....	214
References .....	214
<b><i>Chapter 7</i> .....</b>	<b>214</b>
<b>Conclusions and Future Prospects .....</b>	<b>214</b>
<b><i>Appendix</i>.....</b>	<b>214</b>
<b>Supporting information items.....</b>	<b>214</b>

## List of Schemes

Scheme 1.1: Hydroxylation reaction catalyzed by cytochrome p450 monooxygenase by activating the molecular oxygen and inactivated C-H bond.....	29
Scheme 1.2: Oxidative coupling of two MPA molecules to UA catalyzed by a CYP .....	34
Scheme 1.3: Post-PKS C-O coupling catalyzed by cytochrome p450 (CYP) and methylation of -OH catalyzed by O-methyltransferase (OMT) in grayanic acid biosynthesis .....	34
Scheme 1.4: Post-PKS reactions catalyzed by OMT and CYP in atranorin biosynthesis .....	35
Scheme 1.5: Cytochrome p450 (MPAO) in UA biosynthesis catalyzes the two consecutive H-abstractions of from two MPA substrate molecules .....	37
Scheme 1.6: Radical-radical C-C oxidative coupling of two MPA radicals to UA. ....	37
Scheme 1.7: Proposed cationic mechanism of cytochrome p450 that catalyzes the oxidative conversion of two molecules of MPA to UA.....	39
Scheme 1.8: Grayanic acid biosynthesis.....	40
Scheme 1.9: First proposed radical-2 C-O coupling in grayanic acid biosynthesis. ....	42
Scheme 1.10: Second proposed radical-2 C-O coupling in grayanic acid biosynthesis. ....	43
Scheme 1.11: Third proposed radical-2 C-O coupling in grayanic acid biosynthesis that involves an epoxide formation. ....	44
Scheme 1.12: Post-PKS reactions catalyzed PatG; 6-MSA decarboxylase, PatH; m-cresol hydroxylase, PatI; m-hydroxy benzyl alcohol hydroxylase, PatN; Isoepoxydon dehydrogenase, in patulin biosynthesis.....	48
Scheme 3.1: Chemical synthesis of MPA.....	116
Scheme 4.1: The synthesis of compounds 1 – 4 is shown. ....	132
Scheme 4.2: The synthesis of compounds 5 – 8 is shown. ....	132
Scheme 4.3: The synthesis of compounds 9 – 12 is shown. ....	133
Scheme 5.1: Schematic representation of Baeyer villager style oxidative cleavage of averufin catalyzed by StcB/CypX/afIV.....	180
Scheme 6.1: Cytochrome p450 part of terpene synthase gene cluster found in <i>C. uncialis</i> . ....	211
Scheme 6.2: Proposed biosynthetic pathway of a novel ‘Compounds 13-15’. ....	214
Scheme 6.3: Hydroxylation reaction catalyzed by <i>TRIII</i> ; trichothecene C-15 hydroxylase. ....	215

Scheme 6.4: Putative biosynthetic pathway of a novel molecule with a terpenoid core & a polyketide side chain.....	217
Scheme 6.5: Enzymatic demethylation of (+) pisatin catalyzed by pisatin demethylase into 3,6a-dihydroxy-8,9-methylenedioxypterocarpan (DMDP).....	221
Scheme 6.6: Proposed biosynthesis of 2-butyl-1,3,6,8-tetrahydroxy-naphthalene (BTHN), 2-acetyl-1,3,6,8-tetrahydroxy-naphthalene (ATHN) and, naphthopyrone. ....	225
Scheme 6.7: Proposed biosynthetic pathway of four novel polyketide molecules named as ‘Compounds f-i’.....	228
Scheme 6.8: Proposed oxidative C-C couplings of ‘Compounds f-i’ to six novel polyketide molecules named as ‘Compound 17-22’.....	229
Scheme 6.9: Flaviolin biosynthesis, where deacetylation of I-2 and chain chain shortening of naphthopyrone to 4THN. ....	229
Scheme 6.10: Proposed radical-2 coupling of two molecules of ‘Compound f’.....	231
Scheme 6.11: Hydroxylation & lactonization of 3,5-dimethylorsellinic acid catalyzed by p450 ( <i>andk</i> ). ....	236
Scheme 6.12: <i>S. alpinum</i> gene cluster responsible for the biosynthesis of atranorin. ....	237
Scheme 6.13: Experimentally supported pathway for solanapyrone biosynthesis in <i>A. solani</i> and in parallel a biosynthetic pathway for solanapyrone biosynthesis that is proposed to be encoded in <i>C. uncialis</i> . ....	243

## List of Figures

### Chapter 1

Figure 1.1: Natural products isolated from different species of bacteria, fungi and plants, and their biological properties.....	3
Figure 1.2: Abundance of various adaptive lichen species in diverse habitats.....	6
Figure 1.3: Classification of lichens based on its morphology and anatomy of mutualistic photobiont and a mycobiont.....	8
Figure 1.4: The anatomy of lichens .....	9
Figure 1.5: Pictorial representation of primary and SM biosynthesis in lichen symbionts .....	11
Figure 1.6: Examples of lichen SMs with their biological properties .....	14
Figure 1.7: The experiment resulted in the first polyketone structure, orcinol during the work of 6-MSA.....	16
Figure 1.8: Types of polyketide synthases (PKSs).....	17
Figure 1.9: Classification of PKS based on its domain architecture.....	18
Figure 1.10: Detailed schematic representation of the function of each catalytic domain of PKS results in the formation of aromatic polyketides, partially or highly reduced polyketides and, fatty acids .....	20
Figure 1.11: Examples of polyketides with their biological properties, divided in two categories based on the length of original polyketide chain .....	22
Figure 1.12: Pictorial representation of a biosynthetic gene cluster (BGC).....	23
Figure 1.13: Schematic outline of two modes of cyclization of a polyketide chain .....	26
Figure 1.14: Examples of post-PKS/accessory/tailoring genes .....	27
Figure 1.15: Prosthetic of cysteinato-heme enzymes .....	29
Figure 1.16: part A) Catalytic cycle of a cytochrome p450, part B) Hydroxylation reaction catalyzed by ‘Compound I’, one of the intermediates in the catalytic cycle of CYP .....	31
Figure 1.17: Post-PKS genes encode for different post-PKS enzymes .....	33
Figure 1.18: Patulin biosynthetic gene cluster with 15 different genes .....	46
Figure 1.19: Biosynthesis of 6-methylsalicylic acid (6-MSA) catalyzed by 6-methylsalicylic acid synthase (6-MSAS).....	48
Figure 1.20: Classification of lichen terpenoids .....	51



Figure 1.21: Experimental evidence of slow growth of a lichen fungus named *C. uncialis*. ..... 55

### Chapter 3

Figure 3.1: Chemical structure of Usnic acid and its biological properties. .... 108

Figure 3.2: Isotope-labelled (<sup>14</sup>C) feeding experiments of lichens Taguchi in 1969. .... 110

Figure 3.3: Schematic representation of *C. uncialis* genome sequencing, assembly, and annotation of biosynthetic genes. .... 112

Figure 3.4: Gene-knockout experiment for the identification of UA gene clusters. .... 114

Figure 3.5: A schematic outline of biosynthesis of UA catalyzed MPAS and MPAO ..... 115

Figure 3.6: Effect of UA and MPA on biofilm formation. .... 120

### Chapter 4

Figure 4.1: The biosynthesis of UA. .... 130

Figure 4.2: MPA analogs to UA analogs catalyzed by cytochrome p450. .... 132

Figure 4.3: Chemical structure of the synthesized compounds 1-12. .... 133

### Chapter 5

Figure 5.1: Summary of the experimental work of functional heterologous expression of MPAO, and CPR. .... 143

Figure 5.2: Advantages and Dis-advantages associated with a prokaryotic expression system. 145

Figure 5.3: Strategies/key approaches to improve the eukaryotic CYP production in bacteria. 147

Figure 5.4: Three different electron transfer (ET) mechanisms of redox partners (RPs) of eukaryotic CYPs ..... 153

Figure 5.5: Steps of cloning and heterologous expression of MPAO in *E. coli* ..... 156

Figure 5.6: Cloning and heterologous expression of MPAO codon optimized for *E. coli* without TMD. .... 161

Figure 5.7: Screening of soluble expression of MPAO and purified MPAO protein with SDS-PAGE gel and Western Blot ..... 162

Figure 5.8: Whole cell bio-conversion assays of MPA to UA catalyzed by MPAO. .... 164

Figure 5.9: Pictorial representation of sample analysis method, HPLC coupled UV detector for whole-cell bioassays ..... 165

Figure 5.10: HPLC traces along with UV spectra at 254 nm wavelength of bioconversion experiments .....	166
Figure 5.11: Schematic representation of identification of CPR, CB5 and CB5R (RPs).....	168
Figure 5.12: Pictorial description of primer design to remove introns manually for <i>cu-cpr357</i> gene .....	169
Figure 5.13: Purified MPAO and CPR357 were screened with 10% SDS-PAGE gel.....	170
Figure 5.14: Bio-conversion experiments of MPA to UA catalyzed by both MPAO and CPR357 (cell free extracts, in-vitro). .....	171
Figure 5.15: UV coupled HPLC analysis of bioconversion experiments of MPA to UA. ....	173
Figure 5.16: Schematic representation of cloning and heterologous expression of MPAO.....	174
Figure 5.17: Cloning of both <i>cu-mpao</i> and <i>cu-cpr357</i> . .....	175
Figure 5.18: Steps illustrating double transformation of pET28b harboring MPAO & chaperone plasmids .....	176
Figure 5.19: Phylogenetic tree reconstruction of MPAO protein shows two fungal CYP orthologs. ....	178
Figure 5.20: Multiple sequence alignment of <i>A. nidulans</i> CYP, <i>C. uncialis</i> MPAO and <i>T. marneffii</i> CYP.....	179
Figure 5.21: Results of heterologous expression of STCB, and SCYP in <i>E. coli</i> . ....	181
Figure 5.22: Double transformation experiments. ....	183
Figure 5.23: Results of Co-expression experiments screened with 10% SDS-PAGE gels. ....	184
Figure 5.24: Co-expression experiments .....	185
Figure 5.25: Sample preparation of all the bioconversion experiments for UV-HPLC and LC-MS analysis.....	186
Figure 5.26: Co-expression experiments. ....	187
Figure 5.27: Bioconversion of MPA to UA catalyzed by these proteins.....	187
Figure 5.28: Co-expression experiments.. ....	188
Figure 5.29: Bioconversion of MPA to UA catalyzed by these proteins.....	189
Figure 5.30: LC-MS results of MPA to UA bioconversion experiments. ....	193

## Chapter 6

Figure 6.1: Terpene synthase gene clusters in the <i>C. uncialis</i> mycobiont genome. ....	210
---	-----

Figure 6.2: Truncated phylogenetic trees of genes part of <i>Cu-terp-4</i> gene cluster.....	212
Figure 6.3: Truncated phylogenetic trees of genes part of <i>Cu-terp-6</i> gene cluster.....	213
Figure 6.4: Truncated phylogenetic trees of genes part of <i>Cu-terp-9</i> gene cluster.....	218
Figure 6.5: Type I nonreducing polyketide synthase (PKS) gene clusters in the <i>C. uncialis</i> mycobiont genome.....	219
Figure 6.6: Truncated phylogenetic tree of cytochrome p450; <i>Cu-cyp7-nrpks4</i> .....	220
Figure 6.7: Uncharacterized <i>C. uncialis</i> gene cluster, <i>Cu-nr-pks-6</i> .....	223
Figure 6.8: Truncated phylogenetic trees of genes part of <i>Cu-nr-pks-6</i> gene cluster.....	224
Figure 6.9: <i>Stereocaulon alipnum</i> gene cluster involved in the biosynthesis of atranorin .....	233
Figure 6.10: Truncated phylogenetic tree of <i>Cu-nr-pks-12</i> .....	236
Figure 6.11: Type I reducing polyketide synthase (PKS; backbone genes) gene clusters in the <i>C. uncialis</i> mycobiont genome .....	238
Figure 6.12: <i>Alternaria solani</i> gene cluster involved in the biosynthesis of solanapyrone .....	238
Figure 6.13: Truncated phylogenetic trees of genes part of <i>Cu-r-pks-18</i> gene cluster.....	243

Most of the figures are created with [BioRender.com](https://www.biorender.com)

## List of Tables

Table 1.1: Enlisting three biosynthetic pathways involved in the synthesis of different categories of lichen SMs. ....	11
Table 3.1: Assessment of minimum inhibitory concentration (MIC) of UA and MPA against different bacterial strains.....	118
Table 3.2: Pyoverdine and Pyocyanin production in <i>Pseudomonas aeruginosa</i> .....	119
Table 3.3: Swimming and twitching motility of <i>P. aeruginosa</i> .....	120
Table 4.1: MIC of the 12 novel phloroglucinol derivatives against <i>S. aureus</i> strains .....	135
Table 4.2: MBC and ratio of MBC to MIC for the four novel phloroglucinol derivatives against <i>S. aureus</i> strains. ....	135
Table 4.3: <i>In vitro</i> hemolysis of ovine red blood cells by novel phloroglucinol derivatives. ....	136
Table 4.4: Interaction of Compound 9 with several clinically relevant antibiotics against <i>S. aureus</i> . ....	136
Table 5.1: Bacterial chaperone plasmids that express different chaperone proteins .....	149
Table 5.2: BLAST statistics of <i>C. uncialis</i> MPAO protein sequence shows two fungal CYP homologs.....	178
Table 6.1: BLAST statistics of <i>C. uncialis</i> cytochrome p450 ( <i>Cu-cyp6-nrpks-4</i> ).....	221
Table 6.2: Blast statistics that show a genetic similarity of CYP named as pisatin demethylase. ....	221
Table 6.3: BLAST statistics of <i>C. uncialis</i> NR-PKS ( <i>Cu-nr-pks-6</i> ).....	223
Table 6.4: BLAST statistics of <i>C. uncialis</i> NR-PKS ( <i>Cu-nr-pks-12</i> ).....	232
Table 6.5: BLAST statistics of <i>C. uncialis</i> putative GNAT-family acetyltransferase.....	235
Table 6.6: BLAST statistics of <i>C. uncialis</i> putative CYP ( <i>Cu-cyp8-nrpks-12</i> ).....	236
Table 6.7: BLAST statistics of <i>C. uncialis</i> putative reducing polyketide synthase ( <i>Cu-r-pks-18</i> ). ....	240
Table 6.8: BLAST statistics of <i>C. uncialis</i> putative O-methyltransferase ( <i>Cu-r-pks18-OMT</i> )..	240
Table 6.9: BLAST statistics of <i>C. uncialis</i> putative cytochrome p450 ( <i>Cu-cyp9-rpks18</i> ).....	241
Table 6.10: BLAST statistics of <i>C. uncialis</i> putative FAD oxidase ( <i>Cu-r-pks18-FAD oxidase</i> )..	241
Table 6.11: BLAST statistics of <i>C. uncialis</i> putative short-chain dehydrogenase/reductase ( <i>Cu-r-pks18-SDR</i> ).....	242

## List of Supporting Information Items

Figure S1: Phylogenetic tree of a putative <i>Squalene cyclase</i> of <i>Cladonia uncialis</i> .....	305
Figure S2: Phylogenetic tree of a putative <i>Cytochrome p450</i> of <i>Cladonia uncialis</i> . ....	306
Figure S3: Phylogenetic tree of a putative <i>Cytochrome p450</i> of <i>Cladonia uncialis</i> . ....	307
Figure S4: Phylogenetic tree of a putative <i>Cytochrome p450</i> of <i>Cladonia uncialis</i> . ....	309
Figure S5: Phylogenetic tree of a putative <i>farnesyl-diphosphate farnesyl transferase</i> of <i>Cladonia uncialis</i> .....	309
Figure S6: Phylogenetic tree of a putative <i>aristolochene synthase</i> of <i>Cladonia uncialis</i> . ....	310
Figure S7: Phylogenetic tree of a putative <i>Cytochrome p450</i> of <i>Cladonia uncialis</i> . ....	311
Figure S8: Phylogenetic tree of a putative <i>Cytochrome p450</i> of <i>Cladonia uncialis</i> . ....	312
Figure S9: Phylogenetic relationship between a putative <i>acetyl transferase</i> of <i>Cladonia uncialis</i> . .....	313
Figure S10: Phylogenetic tree of a putative <i>short chain dehydrogenase</i> of <i>Cladonia uncialis</i> ..	314
Figure S11: Phylogenetic tree of a putative <i>type I PKS</i> of <i>Cladonia uncialis</i> .....	315
Figure S12: Phylogenetic tree of a putative <i>Cytochrome p450</i> of <i>Cladonia uncialis</i> . ....	316
Figure S13: Phylogenetic tree of a putative <i>Type I PKS</i> of <i>Cladonia uncialis</i> .....	317
Figure S14: Phylogenetic tree of a putative <i>short-chain dehydrogenase</i> of <i>Cladonia uncialis</i> .	319
Figure S15: Phylogenetic tree of a putative <i>monooxygenase</i> of <i>Cladonia uncialis</i> . ....	319
Figure S16: Phylogenetic tree of a putative <i>scytalone dehydratase</i> of <i>Cladonia uncialis</i> . ....	321
Figure S17: Phylogenetic tree of a putative <i>Cytochrome p450</i> of <i>Cladonia uncialis</i> . ....	322
Figure S18: Phylogenetic tree of a putative <i>type I PKS</i> of <i>Cladonia uncialis</i> .....	323
Figure S19: Phylogenetic tree of a putative <i>Cytochrome p450</i> of <i>Cladonia uncialis</i> . ....	324
Figure S20: Phylogenetic tree of a putative <i>Type I reducing PKS</i> of <i>Cladonia uncialis</i> .....	325
Figure S21: Phylogenetic tree of a putative <i>cytochrome p450</i> of <i>Cladonia uncialis</i> . ....	326
Figure S22: Phylogenetic tree of a putative <i>O-methyltransferase</i> of <i>Cladonia uncialis</i> . ....	327
Figure S23: Phylogenetic tree of a putative <i>FAD oxidase</i> of <i>Cladonia uncialis</i> .....	328
Figure S24: Phylogenetic tree of a putative <i>short-chain dehydrogenase/reductase</i> of <i>Cladonia uncialis</i> . .....	329
Figure S25: Phylogenetic tree of a putative <i>short-chain dehydrogenase/reductase</i> of <i>Cladonia uncialis</i> . .....	330

Figure S26: Rightslink® by Copyright Clearance of Figures 5, 6 and 7 from <i>Journal of Natural Products</i> 2018 81 (4), 723-731. ....	331
Table S1: Primers used throughout this work.....	262
Table S2: Plasmids used throughout this work.....	265
Table S3: Bacterial Strains.....	267
Table S4: Minimum biofilm eradication concentration (MBEC) of active phloroglucinol derivatives against <i>S. aureus</i> strains. ....	268
Table S5: BLAST statistics for Cluster 15 in Figure 31.....	298
Table S6: BLAST statistics for Cluster 17 in Figure 31.....	300
Table S7: BLAST statistics for Cluster 20 in Figure 31.....	301
Table S8: BLAST statistics for Cluster 24 in Figure 35.....	301
Table S9: BLAST statistics for Cluster 26 in Figure 35.....	303
Table S10: BLAST statistics for Cluster 32 in Figure 35.....	304
Table S11: BLAST statistics for Cluster 48 in Figure 41.....	304

## Non-IUPAC Abbreviations

6MSAS:	6-methylsalicylic acid synthase
ACP:	Acyl carrier protein
AT:	Acyltransferase
CB5R:	Cytochrome B5 reductase
CLC:	Claisen cyclase (An alternative abbreviation is “CYC”)
TE:	Thioesterase
TH:	Thioester hydrolase
THID:	Thioester hydrolase interdomain
CYP:	Cytochrome p450
CPR:	Cytochrome p450 reductase
GNAT:	General control non-responsible 5 (GCN5)-related N-acetyltransferase
SDR:	Short-chain dehydrogenase/reductase
SD:	Scytalone dehydratase
NHH:	Non-heme halogenase
SQTKS:	Squalene tetraketide synthase
3-MOA:	3-Methyl orsellinic acid
PA:	Phloroacetophenone
KRDH:	Ketoreductase-dehydratase-like protein
KR:	Ketosynthase
MPAO:	Methylphloroacetophenone oxidase
MPAS:	Methylphloroacetophenone synthase
MT:	C-methyltransferase (An alternative abbreviation is “cMeT”)
NRPS:	Non-ribosomal peptide synthetase
PKS:	Polyketide synthase
PT:	Product template
R:	Reductase
SAT:	Starter acyltransferase (May also be referred to in the literature as “starter unit-acyl carrier protein transacylase” (SAT))

# Chapter 1

## Introduction to lichens and their secondary metabolites/natural products

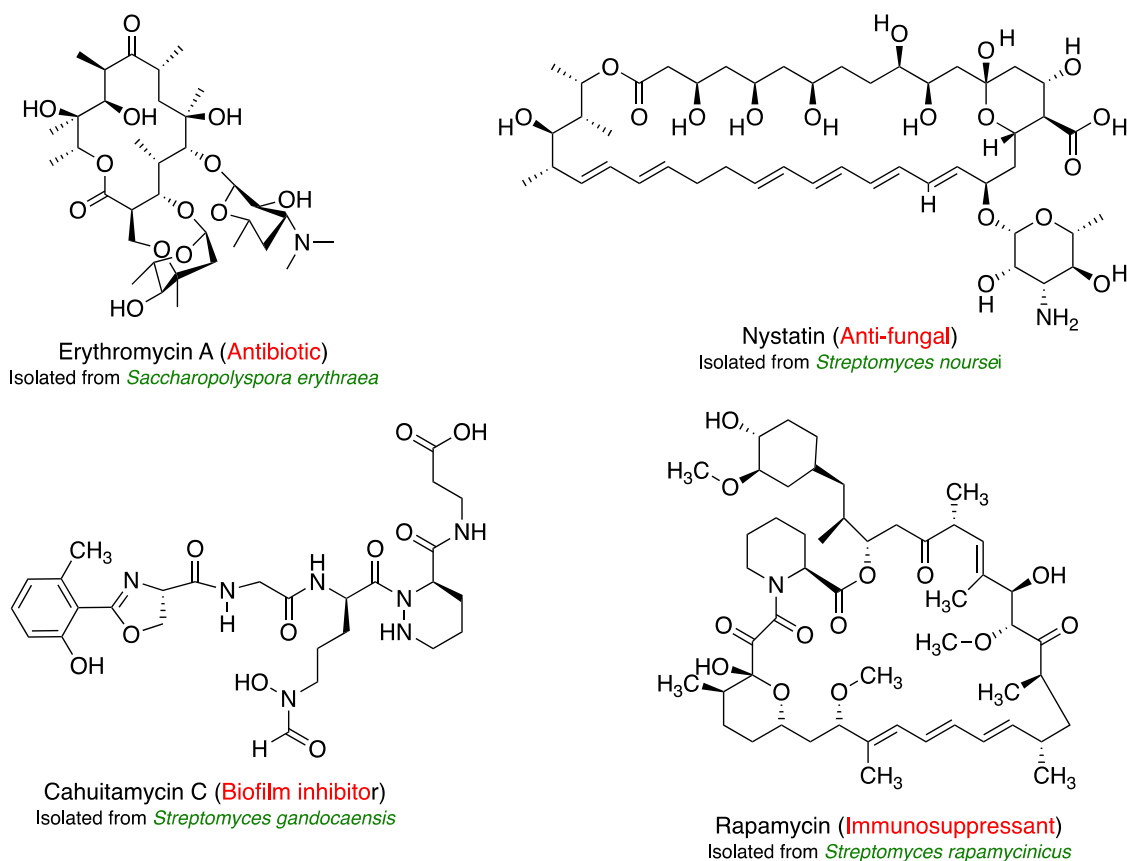
### 1.1. Natural Products

Natural products are a collection of small molecules derived from plants, animals, micro-organisms, terrestrial vertebrates, and invertebrates [Chin et al., 2006]. These molecules have been a rich source of structurally diverse and biologically active organic molecules with great pharmacological, pharmaceutical, and therapeutic importance. Many years of exploration of natural products have resulted in the discovery of almost 300,000 natural product-based drug molecules. The beginning of the scientific research that pertained to these natural products was from the isolation of morphine from opium poppy (*Papaver somniferum*) by Friedrich Wilhelm Sertürner in 1805 [Goerig and Esch, 1991]. Before that, Mesopotamia's medicinal documentation from 2900 to 2600 BCE [Borchardt, 2002; Siddiqui et al., 2014] witnessed the earliest application of natural products to treat human ailments. By the 1900s, most people were using natural products (80%) derived from plants as medicines [Sneader, 1997; Siddiqui et al., 2014]. The discovery of penicillin from *Penicillium notatum* by Alexander Fleming in 1928 [Fleming 1944] and its subsequent development into medicine by Florey and Chain in the 1940s miraculously laid a foundation for the development of microbial natural products as a cornerstone of new drug discovery in the 20<sup>th</sup> century [Chain et al., 1940/2005]. Natural products have a plethora of biological activities positively relevant to the human health, such as antibiotics: Erythromycin A, [McGuire et al., 1952; Zhang et al., 2010; Cobb et al., 2013], anti-fungal: Nystatin [Stanley and English, 1965], anti-tumour/cancer: Actinomycin D [Waksman and Woodruff, 1940 and 1941;



Hollstein, 1974], immunosuppressants: Rapamycin [Vezina, C. et al., 1975], biofilm inhibition: Cahuitamycins C [Park et al., 2016] etc. (**Figure 1.1**).

Vancomycin, artemisinin, taxol, aspirin and cephalosporins are some of the most widely used natural products in the modern pharmaceutical industry (**Figure 1.1**). Meanwhile, avermectins, pyrethrins and azadirachtins (*Azadirachta indica* plant) are the most widely used natural pesticides [Butterworth, J. H. et al. 1968; Kraus, W. 1995] (**Figure 1.1**). In this study, the focus is on the biosynthesis of lichen natural products (LNPs).



**Figure 1.1:** Natural products isolated from different species of bacteria, fungi and plants, and their biological properties [Butterworth, J. H. et al. 1968; Cobb et al., 2013; Hollstein, 1974; Kraus, W. 1995; McGuire et al., 1952; Park et al., 2016; Stanley and English, 1965; Vezina, C. et al., 1975; Waksman and Woodruff, 1940 and 1941; Zhang et al., 2010].

## 1.2. Primary versus secondary metabolism

Metabolism is the sum of all the biochemical reactions within a living cell, providing energy for vital processes and the metabolites are the intermediates of these chemical reactions. Despite the immense difference in characteristics of various living organisms, the pathways to modify and synthesize these metabolites are essentially the same, with minor variations in all the living organisms. There are two types of metabolisms: primary and secondary, and the combination of these two results in the synthesis of NPs. Therefore, there is a categorization of NPs as primary and secondary metabolites (SMs).

### 1.2.1. Primary metabolites

Primary metabolites are usually limited to small molecules, and living cells synthesize these molecules by themselves. These metabolites are directly involved in cell growth, development, and reproduction. Therefore, these molecules are essential for the survival of a living being, and their absence would lead to cell death. The cells assemble these macromolecules using different building blocks like nucleotides and amino acids. Primary metabolic processes that produce these building blocks are Krebs/citric acid/tricarboxylic acid cycle, glycolysis, beta-oxidation of fatty acids and oxidative phosphorylation (part of photosynthesis) in aerobic organisms. Carbohydrates, proteins, and fats are primary metabolites, which further catabolize into citric acid, acetic acid, ethanol, etc. Primary metabolites are also modified to other compounds such as vitamins and coenzymes (**Figure 1.5**).

### 1.2.2. Secondary metabolites

In contrast to primary metabolism, secondary metabolism, which is limited to certain organisms, is responsible for producing organic compounds or natural products that are not directly involved in cell growth, development, and reproduction. These organic molecules are known as

secondary metabolites (SMs) which have unusual structures and their production arises from primary metabolites (**Figure 1.5**). Simple structures of primary metabolites are condensed into more complex structures of SMs by defined biochemical pathways. A. Kossel introduced the term secondary in 1891, which means their absence does not curtail the life of a living organism as they do not take part in the synthesis of the basic molecular skeleton of an organism (vegetative growth of an organism). But the survival of an organism is significantly affected due to the lack of these metabolites. SMs are used for different purposes by different organisms, for example:

1. Protection from predation, from environmental stress
2. Competition and toxicity against other animals, plants, insects, and microorganisms [Fox and Howlett, 2008; Derntl et al., 2017]
3. Metal transporting agents
4. Agents for symbiotic relations with other organisms and communication between organisms
5. Reproductive agents and differentiation effectors

Therefore, each species is characterized by a specific profile of SMs [Mitrovic et al., 2011].

### **1.3. Lichens Natural Products**

#### *1.3.1 Lichen*

Lichens, by definition, are symbiotic organisms. Symbiotism is a mutualism (where all the partners gain benefits from the association) between a fungal partner (a mycobiont) and a photobiont; which could be green algae or cyanobacteria [Ahmadjian and Jacobs, 1981; Seaward, 1997; Brodo et al., 2001; Papazi et al., 2015]. This dual nature of lichen was established by Simon Schwedner (1844) [Honegger R. Great Discoveries in bryology and lichenology - Simon Schwendener (1829-1919)]. Studies have shown that the fungal partner may have a parasitic,

commensalistic, mutualistic, or saprophytic/saprobic relationship with the algal partner. Overall, the lichen symbiosis has been the most successfully established and adaptive entente (evolutionary divergent) on Earth so far. That is because of their ability to withstand all types of environmental conditions ranging from the tropics to polar regions [Hawksworth, 1991] (**Figure 1.2**). Lichens can be found on or within rocks, on soil, on tree trunks and shrubs, on the surface of living leaves, on animal carapaces, and on any stationary, undisturbed artificial surface such as wood, leather, bone, glass, metal, concrete, mortar, brick, rubber, and plastic [Brightman and Seaward, 1977; Seaward, 2008] (**Figure 1.2**). Lisick (2008) reported 18 lichen species on an acrylic-coated aluminum roof. Most lichens are terrestrial, but a few species occur in freshwater streams and others in marine intertidal zones [Nash, 2008].



Desert Sand



Arctic Tundra



Tree Trunk



Rock

**Figure 1.2:** Abundance of various adaptive lichen species in diverse habitats [Brightman and Seaward, 1977; Seaward, 2008],

[http://i1272.photobucket.com/albums/y387/TulaKrystal2/Moab%20Desert/LichenSteven1\\_zps9b3d7c9a.jpg](http://i1272.photobucket.com/albums/y387/TulaKrystal2/Moab%20Desert/LichenSteven1_zps9b3d7c9a.jpg),

[https://c1.staticflickr.com/7/6004/5981642447\\_dcea004c14\\_z.jpg](https://c1.staticflickr.com/7/6004/5981642447_dcea004c14_z.jpg),

<https://www.lamtrees.com/lichen-on-your-trees-good-or-bad/>,

<https://lagomera1.blogspot.com/2014/06/lichens-on-rock.html>

One of the main reasons for their abundance in diverse habitats is the unique mutualism between fungus and green algae, where both organisms separately would be rare or non-existent. The oldest fossil of the lichen on record is 600 million years old, and it came from China [Yuan, X. et al., 2005]. Therefore, lichens are known to be the first colonizers of terrestrial habitats (pioneers). Approximately 13,500 to 18,500 lichen species have been described worldwide. Nevertheless, many lichens are very sensitive to various air pollutants, mainly nitrogen because they receive all their nutrients and water from wet and dry atmospheric deposition (fall out). Nitrogen deposition can increase the load of nutrients. Too much nitrogen can harm and kill the algae's chlorophyll which is used to produce sugars that feed the fungi. Lichens are also sensitive to other atmospheric pollutants such as sulphur and heavy metal-based compounds. Therefore, lichens have been widely used as bioindicators [Fernandez-Salegui et al., 2007; Glavich and Geiser, 2008; Gries, 1996; Sheppard et al., 2007 – only a few of many studies].

Although most of the general textbooks and researchers refer to lichens as a classic example of mutualism, the potential relationship of a mycobiont and a photobiont may be complex. A rigorous classification of many types of relationships was developed by Rambold and Triebel (1992), where they proclaimed lichens as holo-bionts (many organisms living together), including yeast, or secondary fungi or another algal species and sometimes bacteria as a third or a fourth partner. Lichens are mainly classified as fungi (a dominant mycobiont), and the most significant number of lichen fungi belong to the phylum *Ascomycota*. There are also a few known lichen fungi that belong to *Basidiomycota* and *Deuteromycota* phyla. So far seventeen thousand fungal species as mycobionts of lichens are known.

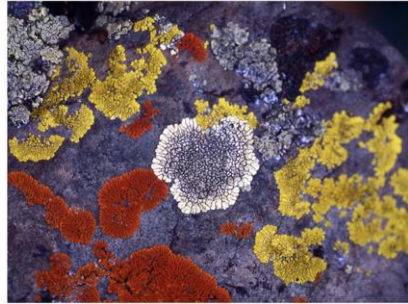
The photobionts of approximately 98% of lichens are not known at the species level [Honegger, 2001]. Only 40 genera of algae/blue-green algae are known as photosynthetic partners

in lichen formation, of which 25 are algae, and 15 are cyanobacteria/blue-green algae [Kirk et al., 2008].

### *1.3.2 Classification and anatomy of lichens*

Lichens grow in many forms, and these organisms exhibit a fantastic array of colours. The morphology of lichen thallus is mainly influenced by the type of photobiont and mycobiont associated with it. Based on the morphology and the anatomy of lichens, they are classified into four different categories such as (**Figures 1.3 and 1.4**) [Nash, T. H., 2008]; **1)** Crustose, adhering tightly to the substrate like a thick coat of the paint where a photobiont forms a separate layer from an upper mycobiont cortical layer without any cortex underneath; **2)** Foliose, a flat leafy structure where a photobiont is in a layer below an upper cortex with a discrete cortex below, separated from a substrate on which it grows; **4)** Fruticose, tiny, leafless hair-like branches could be either strap-shaped or shrubby, they stand out from the surface of the substrate and majority of them possess radial symmetric thalli; **3)** Squamulose, exist as numerous small rounded lobes that are tightly clustered and slightly flattened pebble-like units. They comprise some scaly patterns or squamules. In the *Cladonia* genus, the thallus is usually differentiated into fruticose thallus-verticalis and crustose-squamulose to folios thallus-horizontalis.





A) Crustose



B) Foliose



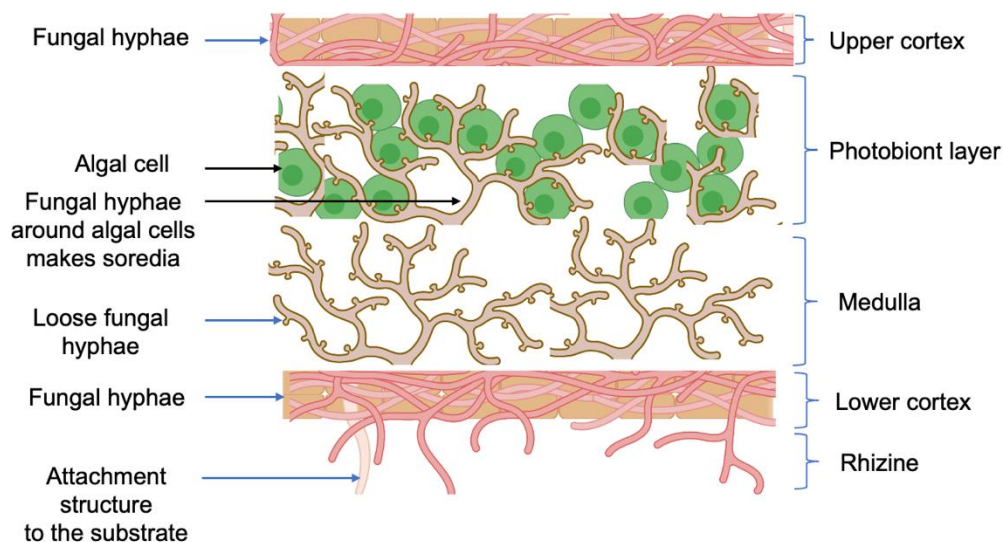
C) Fruticose



D) Squamulose

**Figure 1.3:** Classification of lichens based on its morphology and anatomy of mutualistic photobiont and a mycobiont  
[https://www.fs.usda.gov/wildflowers/beauty/lichens/images/biology/crustose\\_lichen.jpg](https://www.fs.usda.gov/wildflowers/beauty/lichens/images/biology/crustose_lichen.jpg)  
<https://www2.palomar.edu/users/warmstrong/pljan98f.htm>, <https://pixels.com/featured/fruticose-lichen-on-tree-branch-yvonne-johnstone.html>, <https://www.projectnoah.org/spottings/1576186027/fullscreen> [Nash, T. H., 2008].

The anatomy of the lichens varies based on these different forms of lichens but there is a universal anatomical lichen structure detailing different components of lichens are shown in **Figure 1.4**. The function of each component is as follows, a) the upper cortex (fungal hyphae) provides some small measure of protection, as well as provides color in some species, b) underneath upper cortex there is a layer of algal cells or cyanobacteria in some cases that are responsible for the production of nutrient molecules via photosynthesis and these nutrients are absorbed by fungal hyphae wrapped around the algal cells, c) the third layer is called the medulla and the majority of the lichen thallus is comprised of this medulla with the loosely packed fungal cells, d) the fungal hyphae from the medulla extend further to form the lower cortex and finally the rhizines. Rhizines have no vascular capabilities like the roots in plants. They do not move water or nutrients to the lichen; they simply attach the lichen to its substrate.



**Figure 1.4:** The anatomy of lichens consists of a compact fungal hypha (fungal partner in a symbiotic relationship) making the upper cortex layer. Below that the algal cells wrapped in fungal hyphae make an algal layer or a photobiont layer. Then, there is a layer of medulla in which fungal hyphae are loosely packed, below that there is a lower cortex layer consists of compact fungal hyphae and, finally the rhizine through which the lichen attaches to the substrate. Created with [BioRender.com](https://www.biorender.com)

### 1.3.2 Lichen SMs

The specific existence of a symbiotic relationship between an algal and a fungal partner and their ability to grow almost everywhere bestows lichens with unique bioactive natural products/SMs. More than 1,050 different lichen chemical structures have been reported from various species of lichens such as *Ramalina*, *Usnia*, *Cladonia*, *Evernia*, *Alectoria*, *Peltigera*, *Letharia*, *Bryoria* etc. and, apo-symbiotically cultured mycobionts [Molnar and Farkas, 2010].

### 1.3.3 Mycobiont is a significant contributor to lichen SMs

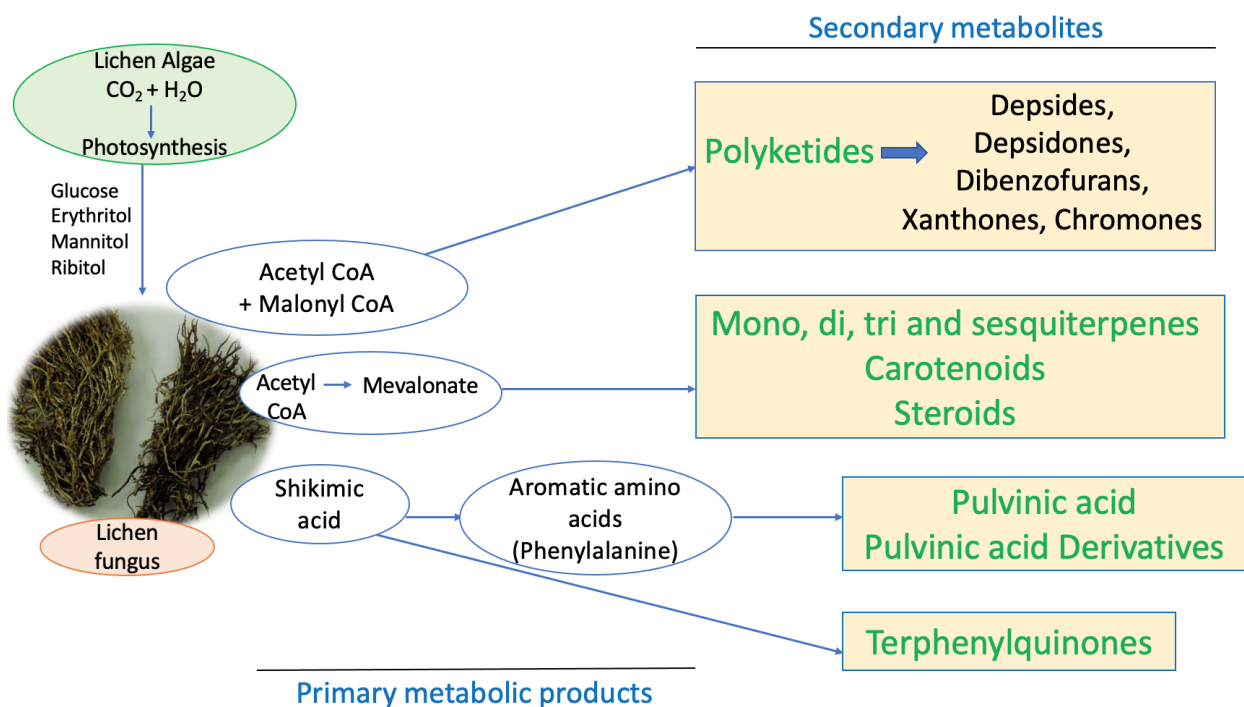
Most of these lichen SMs are produced by the fungal partner in a symbiotic association and are exported outside the fungal hyphae and deposited as crystals in different parts of the thallus, often in the upper cortex or specialized structures such as fruiting bodies (**Figure 1.4**) [Culbertson and Armaleo 1992; Fahselt 1994; Hamada et al. 1996; Kon et al. 1997; Stocker-Worgotter and Elix, 2002].



Although lichen SMs are exclusive in fungal origin, the metabolic interaction between the mycobiont and photobiont is essential to produce these secondary chemicals. Many articles show that a mycobiont grown without a photobiont does not synthesize the same metabolites as the intact lichen species and sometimes synthesizes an entirely different set of chemical products [Molina et al., 2003; Fazio et al., 2009]. Some researchers have also shown that the photobiont, mainly cyanobacteria, is responsible for synthesizing SMs [Cox et al., 2005; Yang et al., 1993]. Over 1,050 SMs have been reported for lichens, and most of these SMs are unique to lichens. Out of these, very few have been reported in non-lichenized fungi or higher plants [Elix and Stocker-Worgotter 2008]. One example is the anthraquinone parietin which is present in other fungi like *Aspergillus* and *Penicillium*, as well as in the vascular plant genera *Rheum*, *Rumex* and *Ventilago* [Romagni and Dayan 2002].

#### *1.3.4 Biosynthetic pathways and biological properties associated with lichen SMs*

The source of carbon in lichens is a photobiont (a cyanobacteria or algae) that produces carbohydrates via photosynthesis (**Figure 1.5**). These carbohydrates (primary metabolites) are transported to the mycobiont (the fungal partner) by the photobiont. Mycobiont converts primary metabolites to SMs (**Figure 1.5**), and the mechanism of metabolism of these primary metabolites is summarized by Mosbach in 1969. The term biosynthesis refers to the synthesis of structurally diverse organic molecules (SMs) catalyzed by different enzymes encoded by specific genes found in the genome of a microorganism. A pathway that leads to the biosynthesis of a SM is called a biosynthetic pathway.



**Figure 1.5:** Pictorial representation of primary and SM biosynthesis in lichen symbionts [Mosbach, K. 1969].

**Table 1.1:** Enlisting three biosynthetic pathways involved in the synthesis of different categories of lichen SMs.

1. *Acetate/Polymalonate pathway:* Most lichen SMs are aromatic polyketides derived from the polyketide pathway. These aromatic compounds are derived from either one or more than one molecule of acetyl CoA or a combination of acetyl CoA and malonyl CoA units. Therefore, this biosynthetic pathway is also known as the acetate–polymalonate pathway. These compounds are unique to lichens and formed from linking two aromatic units (orcinol and p-orcinol) through ester, ether, and C-C bonds. Examples are depsides, depsidones and dibenzofurans. Xanthenes, chromones and anthraquinones are some more interesting examples of polyketides (**Figure 1.6A**).
2. *Mevalonic acid pathway:* Terpenes and steroids are biosynthesized from MEP (methylerythritol phosphate) pathway. Most of these compounds are unique to lichens; only some are found in higher plants (**Figure 1.6B**).

---

3. *Shikimic acid/Shikimate pathway*: This pathway provides an alternate route to aromatic compounds and is mainly responsible for synthesizing two categories of pigmented molecules, terphenyl quinones and pulvinic acid derivatives. The pentose phosphate pathway leads to the formation of shikimic acid from erythritol (a sugar molecule synthesized by algae via photosynthesis and transported to the fungal partner in lichens), which ultimately gets converted to the aromatic amino acids such as L-phenylalanine, L-tyrosine, and L-tryptophan. And most of the pulvinic acid derivatives are derived from phenylalanine (**Figure 1.6C**).

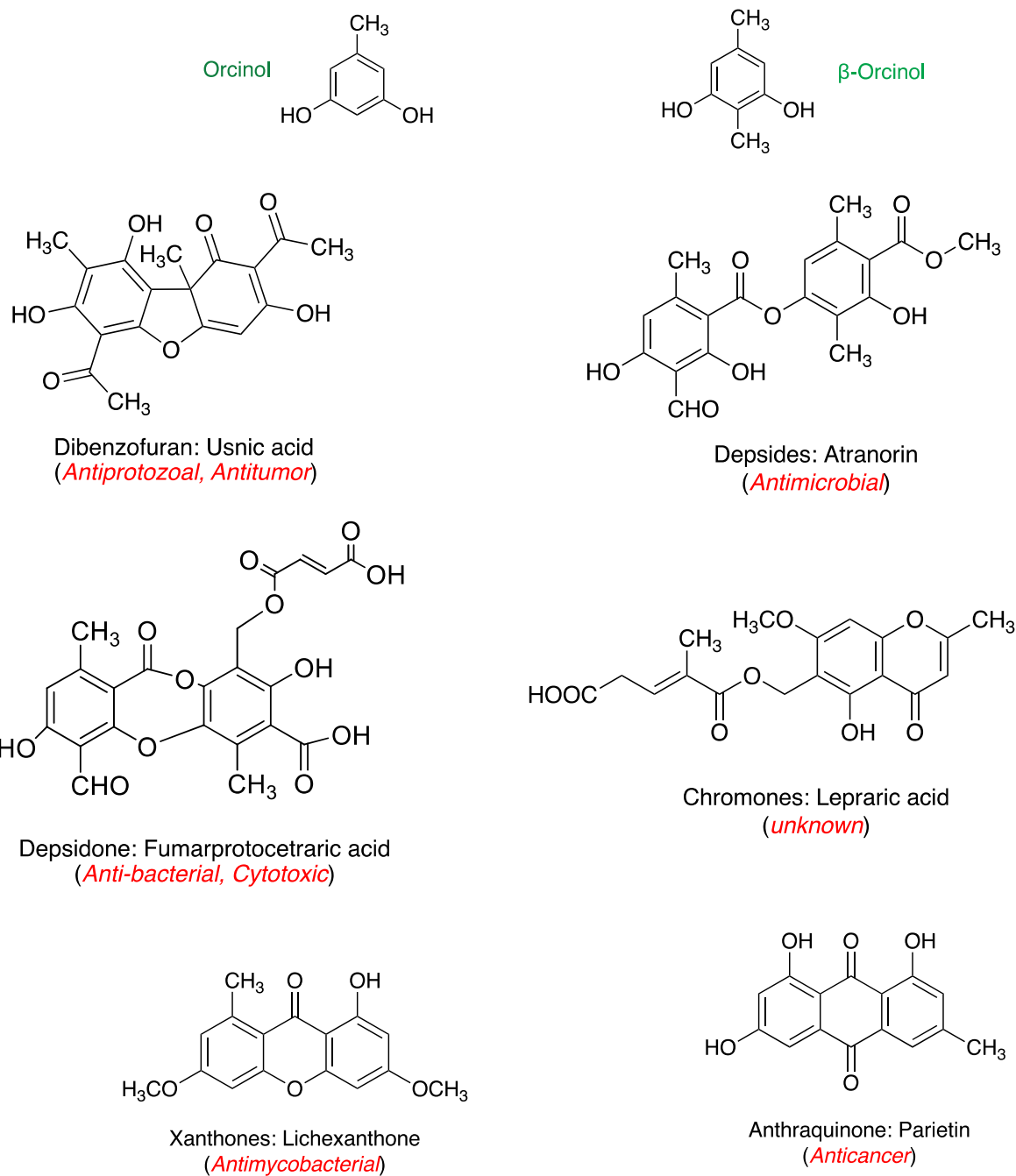
---

There are other categories of lichen SMs, such as non-ribosomal peptides (NRPS), and hybrid compounds include PKS-NRPS and meroterpenoids (a combination of polyketides and terpenoids).

Lichen SMs have been used in folk medicines for centuries to treat a broad array of common ailments such as blood and heart diseases, bronchitis, scabies, leprosy, asthma inflammations, stomach disorders, and many more, all around the globe [Dayan and Romagna 2001, Shukla et al. 2010] (**Figure 1.6**). Advances in medical research resulted in the exploration of various biological activities that have been demonstrated to be associated with these lichen SMs, including anti-tumour [Bucar et al. 2004; Burlando et al. 2009; Kumar and Muller 1999a-c], antifungal, anticancer [Bezivin et al. 2003; Bezivin et al. 2004; Mayer et al. 2005; O'Neill et al. 2010; Ren et al. 2009], antibacterial/antibiotic [Balaji et al. 2006; Burkholder et al. 1944; Paudel et al. 2010; Turk et al. 2003], antiviral, antiprotozoal [De Carvalho et al. 2005; Schmeda-Hirschmann et al. 2008], anti-inflammatory, analgesic, antipyretic, anti-HIV [Nakanishi et al. 1998; Neamati et al. 1997], antioxidant [Bhattarai et al. 2008; Gulluce et al. 2006; Hidalgo et al.

1994], and anti-herbivore etc. [Muller 2001; Molnár and Farkas, 2010; Shukla et al., 2010; Shrestha and St. Clair, 2013] (**Figure 1.6**).

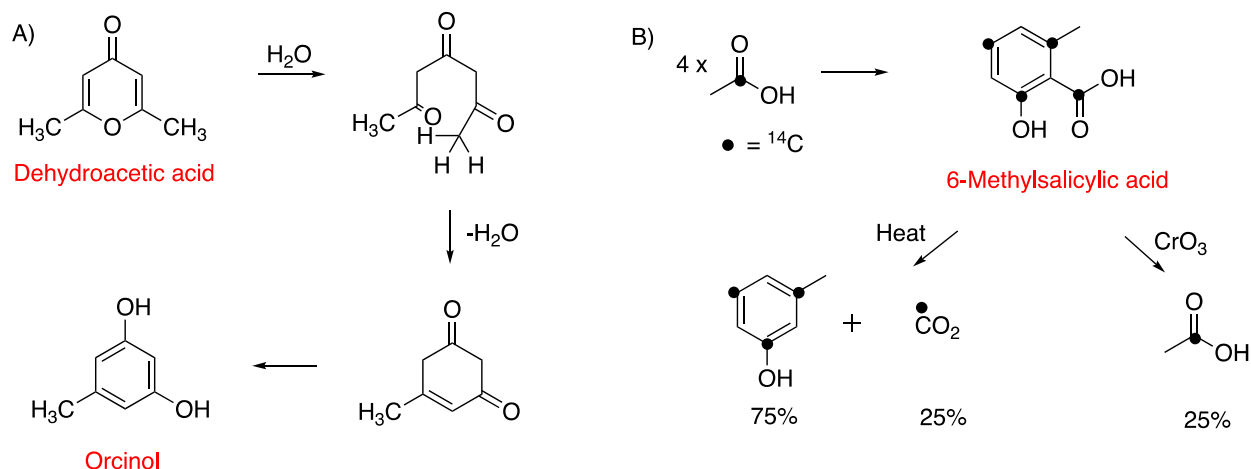
A) Lichen polyketides, xanthones, chromones and, anthraquinones



B) Lichen terpenoids and steroids



[Collie, J. N. et al., 1893] (**Figure 1.7**). He later also suggested the formation of a triketone intermediate during orcinol biosynthesis and proposed a mechanism of polyphenol biosynthesis. However, most of his work was pre-emptive due to the lack of vision and knowledge of natural products of organic chemists during the 1900s [Collie, J. N. 1907]. After many years of Collie's pioneering work, Arthur Birch 1950 established that the biosynthesis of patulin (a toxin produced by a fungus *Penicillium patulum*) involves the formation of an aromatic polyketide intermediate named 6-methylsalicylic acid (6-MSA) and later demonstrated the incorporation of the acetate (radiolabeled acetate was used) to 6-MSA using *Penicillium griseofulvin* in 1955 [Birch, A. J. et al., 1955]. He also showed the formation of m-cresol and acetic acid upon heating and oxidation of 6-MSA. In 1990 Schweizer's group successfully cloned 6-acetylsalicylic acid synthase (6-MSAS), the first fungal PKS gene that putatively catalyzes the biosynthesis of 6-MSA [Beck, J. et al., 1990].



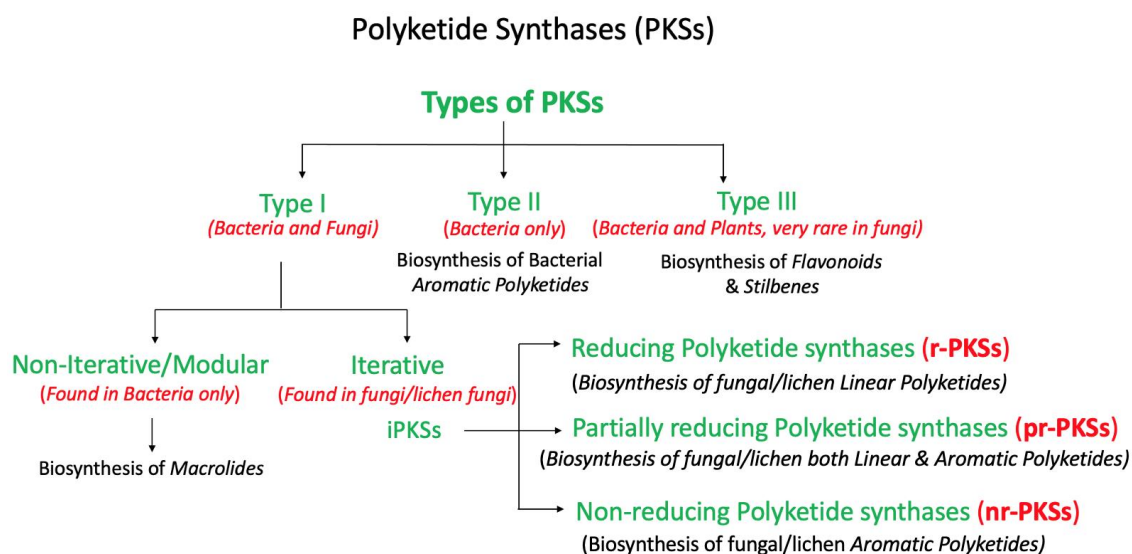
**Figure 1.7:** The experiment resulted in the first polyketone structure, orcinol during the work of 6-MSA [Collie, J. N. 1907; Birch, A. J. et al., 1955]

#### 1.4.2. Classification of PKSs

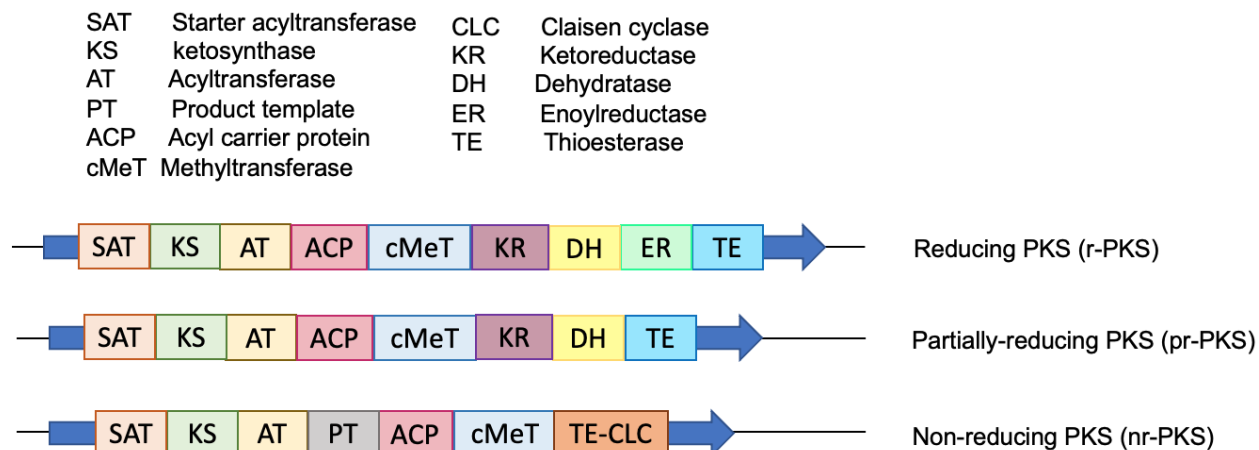
Polyketide synthases (PKSs) are multi-domain enzymes where each domain has a specific catalytic function [Staunton, J. and Weissman, K. J. 2001; Dewick, P. M. 2009]. The detailed study

of PKSs and their mechanistic nature led to the categorization of PKSs into three general types [Staunton, J. and Weissman, K. J. 2001; Dewick, P. M. 2009] (**Figure 1.8**): Type I PKSs are extensive multifunctional proteins with individual catalytic functional domains and are found in bacteria, fungi, and the lichen fungi. Type II PKSs are composed of a complex of monofunctional proteins that are restricted to bacteria only, and Type III are homo-dimeric proteins that utilize coenzyme A esters rather than ACPs to make polyketide intermediates. Type III PKSs employ a single active site to perform a series of decarboxylation, condensation, cyclization, and aromatization reactions. These are also known as chalcone-synthase like PKS and could be isolated from plants, bacteria, and fungi. In addition to acetyl-CoA, there are examples in the literature that show propionyl-CoA and hexanoyl-CoA as starter units and sometimes methyl malonyl-CoA as an extender unit in polyketide pathways (usually in Type I). Type III PKSs use cinnamoyl-CoA as starter units for the chain extension and result in the formation of flavonoids and stilbenes. Type I PKSs are subdivided into iterative PKSs, iPKSs (i.e., repeating) found in fungi and lichen fungi and non-iterative modular systems found in bacteria only. Iterative PKSs use its functional domains repeatedly, whereas non-iterative PKSs possess a distinctive site for every enzyme-catalyzed step to produce a polyketide chain. Iterative PKSs are further subcategorized into three classes: non-reducing polyketide synthase (nr-PKS), partially reducing synthase (pr-PKS) and reducing polyketide synthase (r-PKS). Non-reducing polyketide synthase catalyzes the biosynthesis of fungal/lichen fungal linear polyketides, pr-PKS catalyzes the biosynthesis of fungal/lichen fungal both linear and aromatic polyketides. Reducing PKS catalyzes the biosynthesis of fungal/lichen fungal aromatic polyketides (**Figure 1.8**). This classification has been done based on the domain architecture of polyketide synthase (**Figure 1.9**), and the function of each domain is elaborated in **Section 1.4.3. (Figure 1.10)** In non-reducing and partially reducing

PKS, the reduction step is either entirely or partially emitted because of the absence of reducing domains (DH, ER, KR: elaborated in **Section 1.4.3.**) [Staunton, J. and Weissman, K. J. 2001; Dewick, P. M. 2009] (**Figure 1.10**).



**Figure 1.8:** Types of polyketide synthases (PKSs) [Staunton, J. and Weissman, K. J. 2001; Dewick, P. M. 2009].



**Figure 1.9:** Classification of PKS based on its domain architecture [Staunton, J. and Weissman, K. J. 2001; Dewick, P. M. 2009].

### 1.4.3 Function of each catalytic domain of PKSs

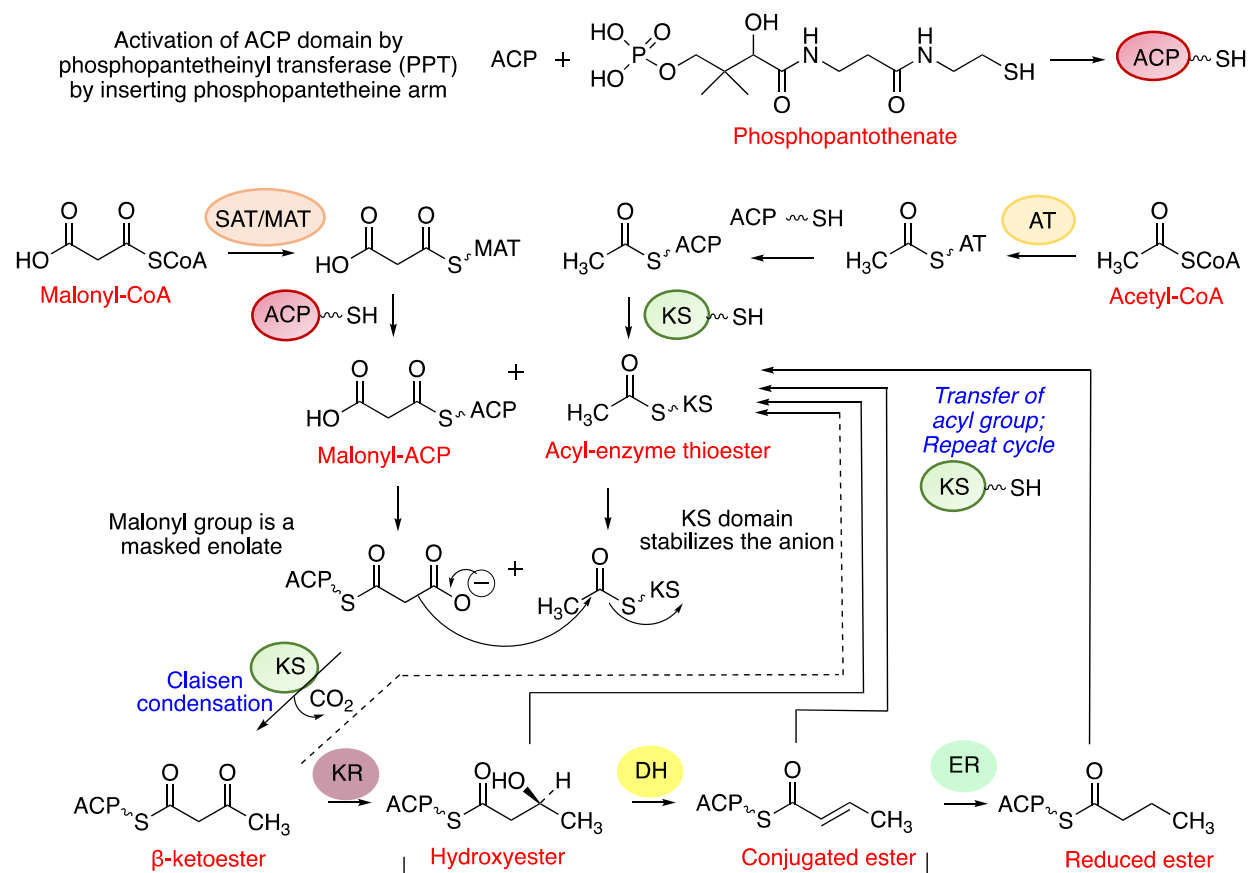
Like fatty acids, polyketides (both fatty acid and polyketide pathways have very high homology) are assembled from C<sub>2</sub> units by repeated head-to-tail linkages catalyzed by PKSs. All the chain extension chemical steps [Staunton, J. and Weissman, K. J. 2001; Dewick, P. M. 2009]

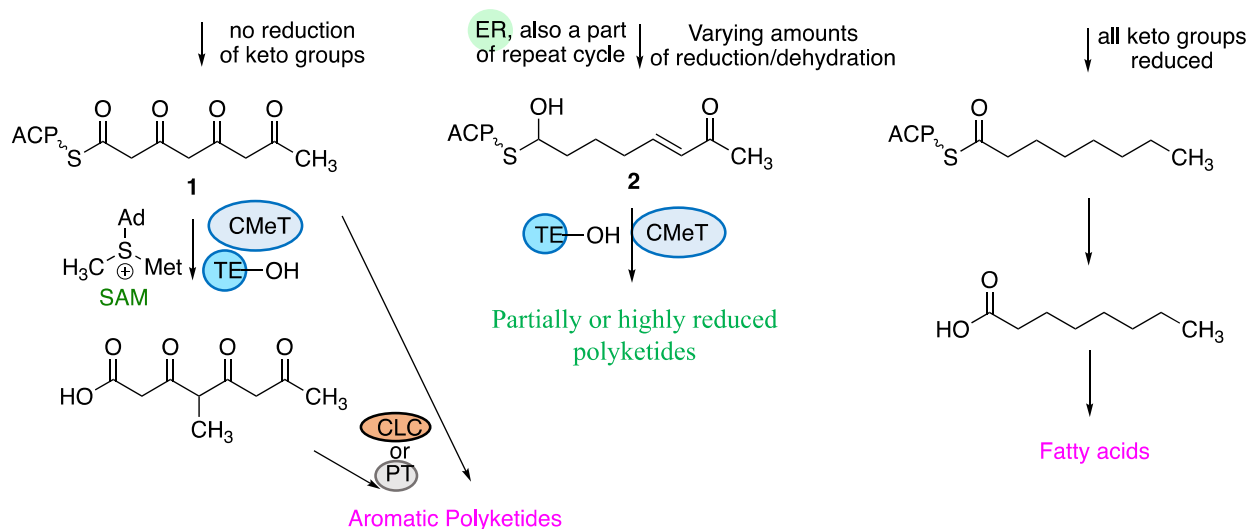


are shown in **Figure 1.10**, where a starter unit is an acyl/acetyl CoA and malonyl-CoA acts as an extender unit that undergo concerted decarboxylation to facilitate the electrons for the new C-C bond formation. Acetyl-CoA and malonyl-CoA do not get involved in the condensation step directly but rather get converted into enzyme-bound thioester and the malonyl ester using an acyl carrier protein (ACP). PKS enzyme contains an ACP binding site and an active-site cysteine residue in the ketosynthase (KS) domain. The ACP carries a phosphopantetheine group that provides a long flexible arm which enables a growing polyketide chain to reach the active site of each enzyme in the complex and, therefore, allows the different chemical reactions to be performed without releasing intermediates from the enzyme [Dewick, P. M. 2009].

Acetyl and malonyl groups are successively transferred from coenzyme A esters and attached to the thiol groups of Cys and ACP, respectively, by the same malonyl/acetyltransferase (MAT/AT; **Figure 1.10**). First, the growing polyketide chain is constructed by Claisen condensation catalyzed by the KS domain, which leads to the formation of  $\beta$ -ketoester (acetoacetyl-ACP). Next, the ketone group is reduced to alcohol by the action of the ketoreductase domain (KR), followed by dehydration to the conjugated ester catalyzed by dehydratase (DH), and then reduction of the double bond using the enoyl reductase (ER) domain. This 3-step reduction process occurs after each condensation step and before the next round of chain extension. As shown in **Figure 1.10**, after several rounds of chain extension of beta-ketoesters and the partially or highly (hydroxy esters/conjugated/reduced esters), reduced forms of beta-ketoesters lead to the formation of compounds **1** (poly beta-ketoesters) and **2**. Then both these compounds undergo the methylation step catalyzed by the cMeT domain (methyltransferase), and the TE domain (thioesterase) is often used to terminate the chain reaction and release the final product. Ultimately that leads to the formation of non-reducing aromatic polyketides from compound **1** and partially

or highly reduced complex polyketides from compound **2** (**Figure 1.10**). Product complexity arises from the number of cycles (chain length control), and the selective reductive processing carried out during product assembly. Both CLC (or CYC/Claisen cyclase) and PT (Product template) domains are responsible for the proper folding/cyclization of a polyketide chain [Tsai, S. C. et al., 2009], and it has also been proposed that the PT domain is required for chemical stabilization of reactive poly- $\beta$ -keto intermediates [Udwary, D. W. et al., 2002]. However, these domains are present in only NR-PKSs.





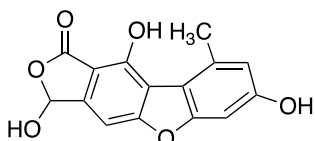
**Figure 1.10:** Detailed schematic representation of the function of each catalytic domain of PKS results in the formation of aromatic polyketides, partially or highly reduced polyketides and, fatty acids [Staunton, J. and Weissman, K. J. 2001; Dewick, P. M. 2009; Tsai, S. C. et al., 2009; Udvary, D. W. et al., 2002].

#### 1.4.4. Polyketides derived from either single or two and more polyketide chains

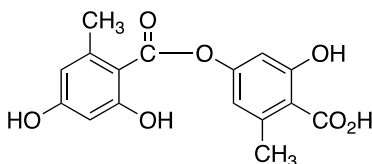
**Figure 1.11** presents some interesting examples of lichen polyketides with a plethora of biological activities ranging from antibacterial, antifungal, antitumor, photosystem II (PSII) inhibitors, cytotoxic agents etc. Biosynthesis of these molecules involves the addition of a specific number of 2-carbon units which leads to the formation of a desired linear polyketide chain. All these molecules are of non-reducing type I iPKS or NR-PKS (solely present in lichens) that have been classified into two sets: One set of molecules is derived from a single polyketide chain that folds and cyclizes in a particular fashion to facilitate the final product (**Figure 1.11**). For example, lichen aliphatic acids (protolichestrinic acid), monocyclic aromatic compounds (orcinol, beta-orcinol and, and divaric acid) cyclized from 4 basic units (a tetrakide), chromones (lepraric acid) are derived from the cyclization of 5 basic unit (pentaketides), and xanthenes (lichexanthone) from 6 basic unit (hexaketide) precursors. Larger polyketide-derived compounds, such as the anthraquinones parietin and haemoventosin, are derived from the cyclization of 8 basic units (octaketide) (**Figure 1.6 Part A** and **Figure 1.11**) [Dayan, F. E. et al., 2001]. The other set of

molecules are derived from two or more polyketide chains [Karunaratne, V. et al., 2005] and are known to be the products of oxidative coupling/linking between either two or more polyketide chains.

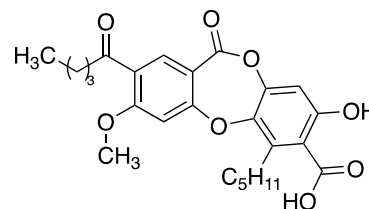
A) Polyketides derived from two or more polyketide chains



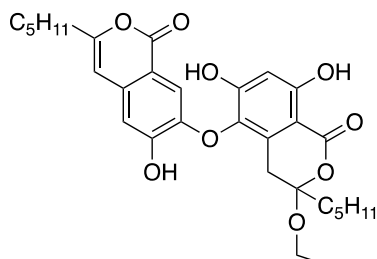
Dibenzofuran: Alectosarmentosin  
(*Antimicrobial*)



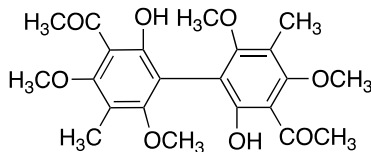
para-depside: Lecanoric acid  
(*Fungitoxin, PSII inhibitor*)



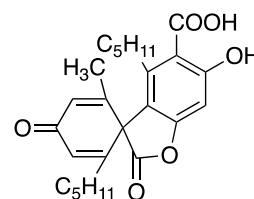
Depsidone: Lobaric acid  
(*Antitumor, Cytotoxic*)



Diphenyl ether:  $\beta$ -Alectoronic acid  
(*Antimicrobial*)

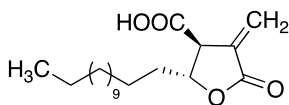


Diphenyl: Contortin  
(*Unknown*)

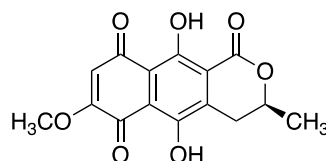


Depsone: Pichrolichenic acid  
(*Unknown*)

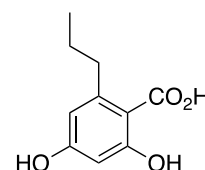
B) Polyketides derived from a single polyketide chains



Aliphatic acids: (+)-Protolichesterinic acid  
(*Antibacterial, Antitumor*)



Naphthoquinone: Haemoventosin  
(*Antibacterial, antifungal, cytotoxic*)

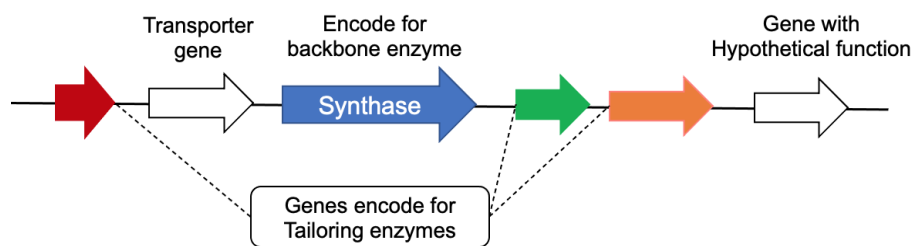


Monocyclic phenolics: Divaric acid  
(*Antibacterial*)

**Figure 1.11:** Examples of polyketides with their biological properties, divided in two categories based on the length of original polyketide chain [Dayan, F. E. et al., 2001; Karunaratne, V. et al., 2005].

## 1.5. Biosynthetic gene clusters (BGCs)

Genes involved in the biosynthesis of SMs in lichen fungi are co-localized in the genome of a micro-organism and organized as clusters of genes and known as biosynthetic gene clusters (BGCs). Typically, BGCs are minimally composed of one or more synthase or synthetase genes encoding backbone enzymes, which produce the core structure of the compound (**Figure 1.12**). Backbone enzymes determine the class of SM produced by a BGC. BGCs may also contain other genes such as those encoding tailoring enzymes, cluster-specific transcription factors, transporters, and genes with hypothetical functions (**Figure 1.12**). Tailoring enzymes are responsible for modifying the core compound to generate various SMs.



**Figure 1.12:** Pictorial representation of a biosynthetic gene cluster (BGC) with a synthase gene that encode for backbone enzyme and the other genes such as transporter gene, gene with a hypothetical function and genes encoding tailoring enzymes [Bertrand, R. L. et al., 2018].

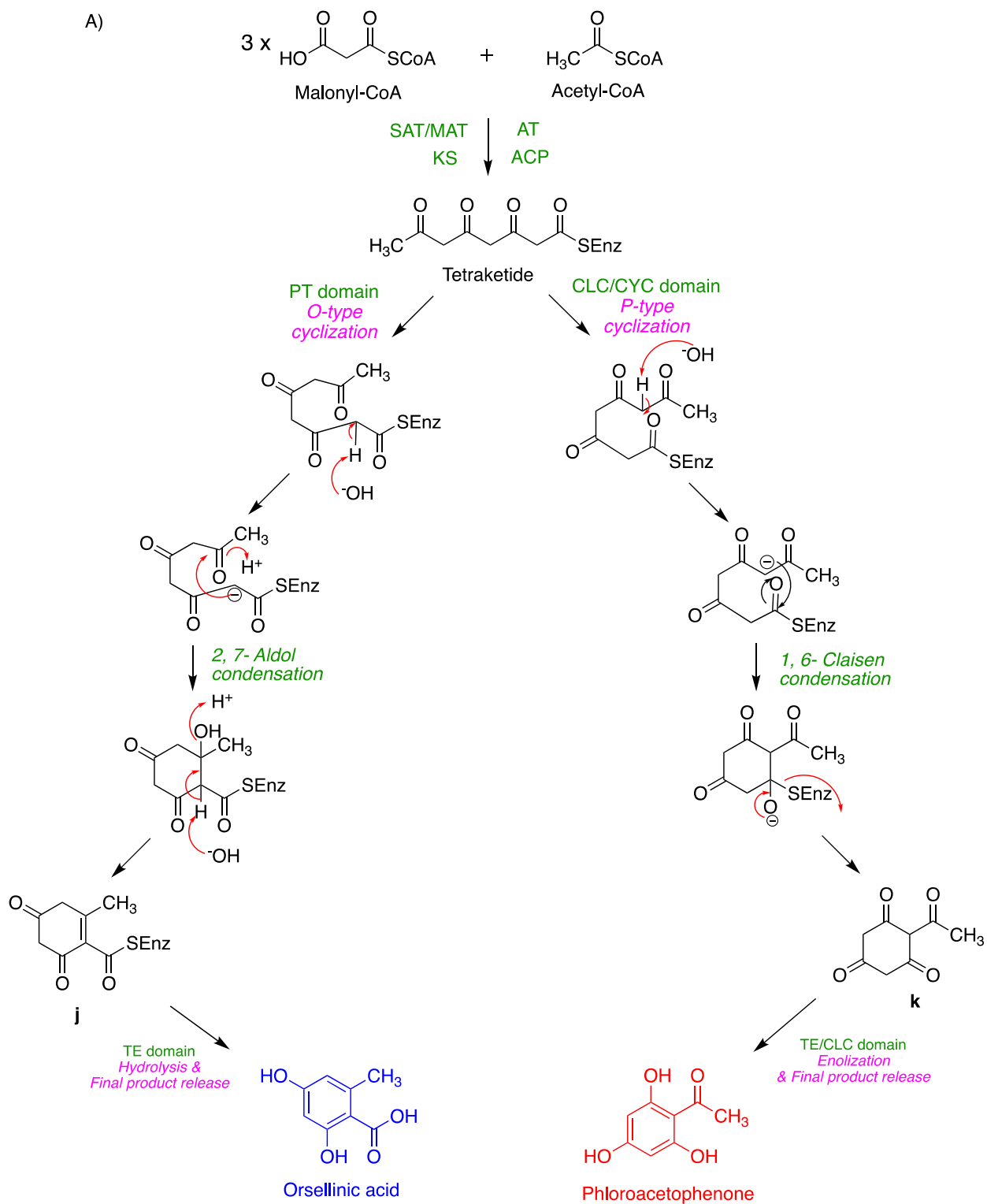
## 1.6. Biosynthesis of Aromatic/Non-reducing Polyketides

### 1.6.1 Function of PKS gene (the key enzyme) in a polyketide BGC

In polyketide BGCs, polyketide synthetase/synthase genes code for backbone enzymes/key enzymes. The most common example of a fungal non-reducing polyketide is aflatoxin biosynthesis catalyzed by NR-PKS. There are many examples of non-reducing lichen polyketides that have been reported in the literature, such as UA, atranorin, lecanoric acid, fumarprotocetraric acid and grayanic acid etc [Abdel-Hameed, M. et al., 2016; Calchera, A. et al., 2019; Wang, Y. et al., 2018;

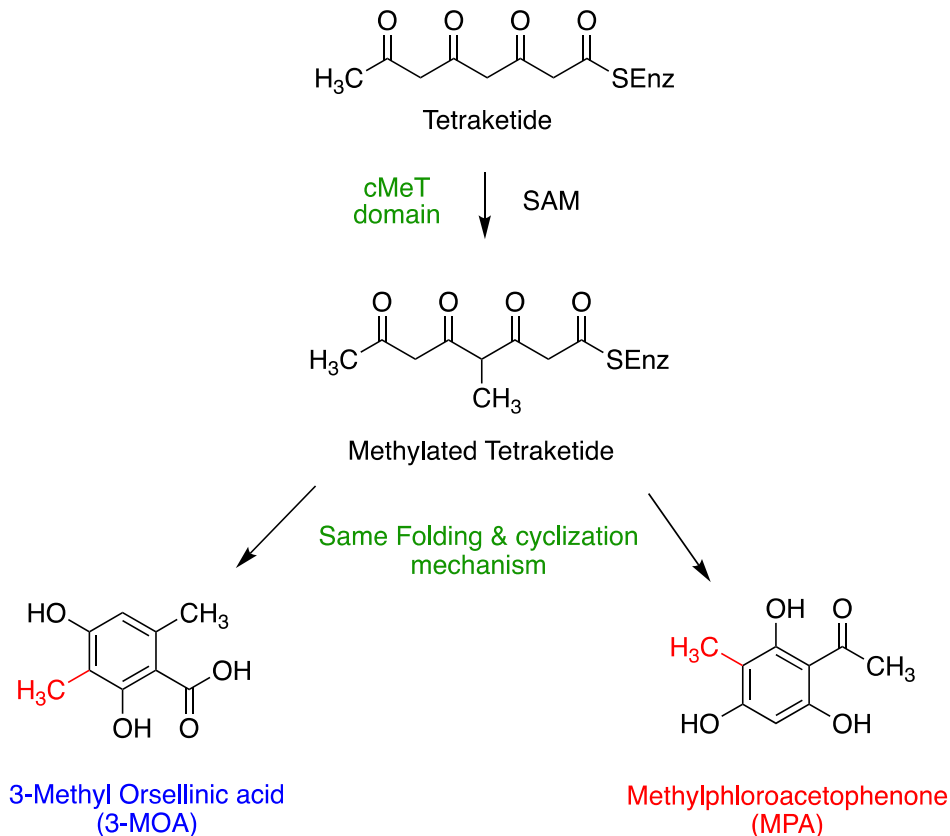
Kim, W. et al., 2021; Chooi, Y. H. et al., 2008; Armaleo, D. et al., 2011]. During the biosynthetic assembly of these lichen non-reducing polyketides catalyzed by NR-PKSs, the first step common to all polyketide biosynthetic pathways, is the formation of a polyketide chain (a tetraketide, **Figure 1.13, part A**) from a starter unit acetyl CoA and an extender unit, malonyl CoA following recursive chain extension steps catalyzed by SAT, ACP, AT and KS catalytic domains as shown in (**Figure 1.10**). The folding of a tetraketide chain in a particular fashion has been facilitated by a PT domain followed by a ring closure via regiospecific 2, 7-aldol condensation also known as O-type of cyclization [Korman, T. P. et al., 2010]. CLC (CYC) and in some cases TE/CLC domain, are deemed to direct the C-C cyclization (P-type) via 1, 6 -Claisen condensation. There has been a quite elaborative research work done to establish the exact role of TE and TE/CLC domains [Korman, T. P. et al., 2010; Zhou, H. et al., 2008; Yuquan Xu et al., 2013; Chooi, Y. H. et al., 2012; Crawford, J. M. et al., 2008, Liangcheng Du and Lili Lou, 2010]. There are some examples where the TE/CLC domain has shown to be involved in C-C bond formation through Dieckmann cyclization. Very few fungal iPKSs feature the TE domain catalyzing O-C cyclization for example, RAL or DAL; resorcylic acid lactone and the dihydroxyphenylacetic acid lactone biosynthesis [Zhou, H. et al., 2008; Wang, M. et al. 2009.] and in some cases it is coupled to product truncation by diacylation [Vagstad, A. L. et al., 2012]. The TE domain achieves the precise control over the mode of final product release by hydrolysis of **Compound j** to orsellinic acid (OA) and enolization of **Compound k** towards phloroacetophenone (PA) as shown in **Figure 1.13, part A**. The methylation of a tetraketide chain catalyzed by the cMeT domain followed by a similar set of folding and cyclization steps results in the formation of 3-methyl orsellinic (3-MOA, b-orsellinic acid) and methylphloroacetophenone (MPA) (**Figure 1.13, part B**) from a methylated tetraketide [Chooi, Y. H. et al., 2008; Abdel-Hameed, M. et al., 2016]. These are monocyclic

aromatic polyketides and are key intermediates in the biosynthesis of dibenzofurans, depsides and depsidones (UA, lecanoric acid, atranorin, fumarprotocetraric acid and grayanic acid) unique to lichens.





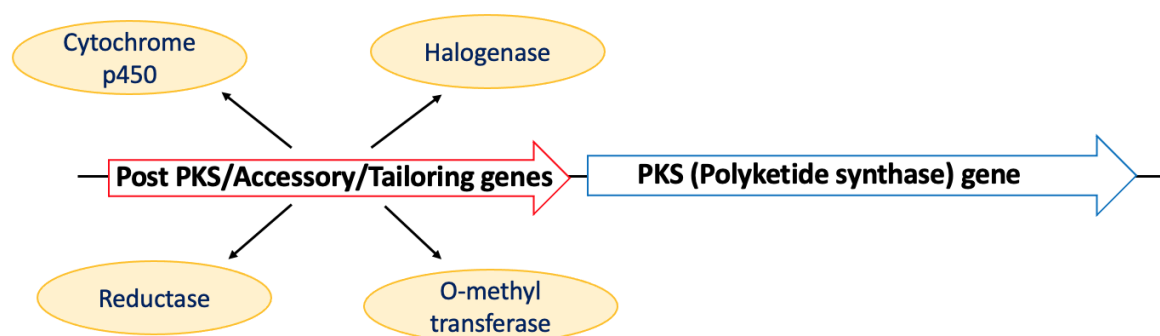
B)



**Figure 1.13:** Schematic outline of two modes of cyclization of a polyketide chain (tetraketide) with the representation of the difference in the function of PT v/s CYC/CLC domains and TE v/s TE/CLC domains. A) O-type (orsellinic acid-type) and P-type (phloroacetophenone-type) cyclization of a tetraketide leads to the formation of orsellinic acid and phloroacetophenone B) O-type and P-type cyclization of a tetraketide with an addition of the cMeT domain leads to the formation of 3-MOA and MPA [Abdel-Hameed, M. et al., 2016; Chooi, Y. H. et al., 2008; Dewick, P. M. 2009].

### 1.6.2. Post PKS/Accessory genes (tailoring enzymes)

Conversion of 3-MOA and MPA to diverse polyketide structures such as UA and atranorin is catalyzed by post-PKS/tailoring enzymes encoded by accessory/post-pks genes. These are present either upstream or downstream of a polyketide synthase gene in a polyketide producing gene cluster. Monooxygenase, cytochrome p450 (CYP), FAD oxidase/monooxygenase, O-methyltransferase, short-chain dehydrogenase/reductase (SDR), oxidoreductase, dehydratase, GNAT (general control non-repressible 5 (GCN5)-related N-acetyltransferases), hydrolase, non-heme halogenase (NHH) are some examples of post-PKS enzymes (**Figure 1.14**).



**Figure 1.14:** Examples of post-PKS/accessory/tailoring genes [Bertrand, R. L. et al., 2018].

### 1.6.3. Background on cytochrome p450s

The superfamily of cytochrome p450 (CYP) enzymes was first identified in 1958 as an absorption spectrum taken from rat liver microsomes with an unusual peak at 450 nm. Cytochrome p450 monooxygenases, also known as CYPs, are b type heme-containing enzymes that play key roles in nature, particularly in the evolution of organisms, including the dawn of multicellular life [Nelson, D. R. 2013; Silverman, R. B. 2002]. By far these oxidative enzymes represent one of the largest and most ancient gene superfamilies that encode for enzymes [Degtyarenko, K.N. et al., 1993]. Since their identification occurred five decades ago, more than 18,500 CYPs from hundreds of species have been discovered across all the biological kingdoms from bacteria to humans.

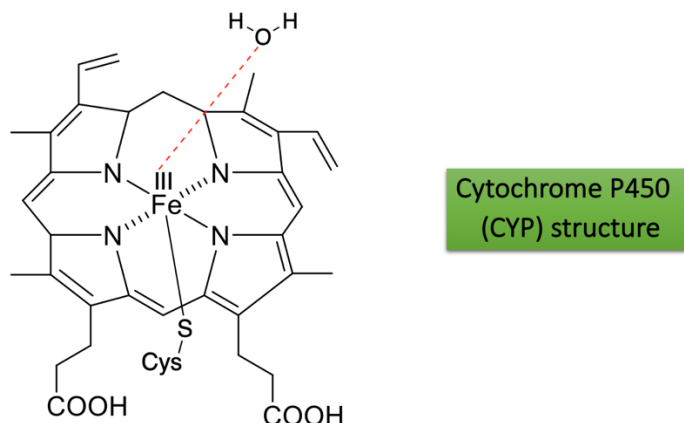
The “Fungal Cytochrome p450 Database” hosted in Korea lists more than 8700 fungal p450 sequences from 113 species (<http://P450.riceblast.snu.ac.kr>; Seoul National University: 2014/03/20), and the “SuperCYP” database contains 1170 drugs, 2785 cytochrome p450-drug interactions, and 1200 p450 alleles ([http:// bioinformatics.charite.de/supercyp](http://bioinformatics.charite.de/supercyp); Charité-University Medicine Berlin; 2014/03/20). The first fungal p450 sequence identified from *Saccharomyces cerevisiae* (yeast) is known as CYP51.

These versatile biocatalysts have been very useful to pharmaceutical industry, agriculture, biotechnology and even cosmetics [Urlacher, V.B. and Girhard, M. 2019]. During the biosynthetic assembly of natural products, at many instances the pathways utilize unique oxidation reactions to

generate diverse functions. Many reduced chemical scaffolds undergo a diverse range of redox reactions and ultimately generate highly complex bioactive molecules. These oxidative enzymes that catalyze redox reactions are categorized as monooxygenases and dioxygenases. The former category includes cytochrome p450s and FAD dependent monooxygenases and the latter includes non-heme iron alpha-ketoglutarate-dependent oxygenases [Silverman, R. B. 2002].

#### *1.6.3a. Prosthetic group of Cytochrome p450*

The active site of a cytochrome p450 contains a heme-iron center specifically heme 'b' (**Figure 1.15**). Heme b is the most abundant heme, also present in hemoglobin and myoglobin. Heme consists of a complex organic ring structure called protoporphyrin, where a single iron atom/Fe(III) is bound to the four nitrogen atoms of the tetrapyrrole ring system forming a planar active site. The heme is also chelated to the cysteine residue (fifth ligand) and connected to the heme iron through a sulphur atom of the thiol axially within the p450 active site [Silverman, R. B. 2002] (**Figure 1.15**). Therefore, cytochrome p450 enzymes are classified as heme thiolate proteins [NC-IUB, Nomenclature Committee of the International Union of Biochemistry [NC-IUB, 1991]. The nomenclature of these well-known heme-thiolate pigments is not based on their function (which is usual for enzymes) but their spectral properties which displays an adsorption band at 450 nm by their reduced carbon monoxide bound form [Omura, T. et al., 1964]. This unusual spectral feature is induced due to a cysteine thiolate group [Ichikawa, Y. et al., 1967] and a water ligand which is also bound to heme-iron in the resting state of p450.



**Figure 1.15:** Prosthetic of cysteinato-heme enzymes: an iron-(III) protoporphyrin-IX linked with a proximal cysteine ligand and an axial water ligand [Silverman, R. B. 2002].

### 1.6.3b. Hydroxylation reaction catalyzed by CYPs

Heme enables the activation of molecular oxygen and is responsible for the oxidation of nonactivated C–H bonds. The most common reaction that CYPs are known to catalyze (**Scheme 1.1, Figure 1.16; part A**) is the transfer of molecular oxygen to X-H bonds (X: -C, -N, -S) of a substrate with the reduction of the other oxygen atom to water [Ruckpaul, K., Rein, H. and, Blanck, J. 1989; Mansuy, D. 1998].

### Most catalyzed reaction



**Scheme 1.1:** Hydroxylation reaction catalyzed by cytochrome p450 monooxygenase by activating the molecular oxygen and inactivated C-H bond [Silverman, R. B. 2002].

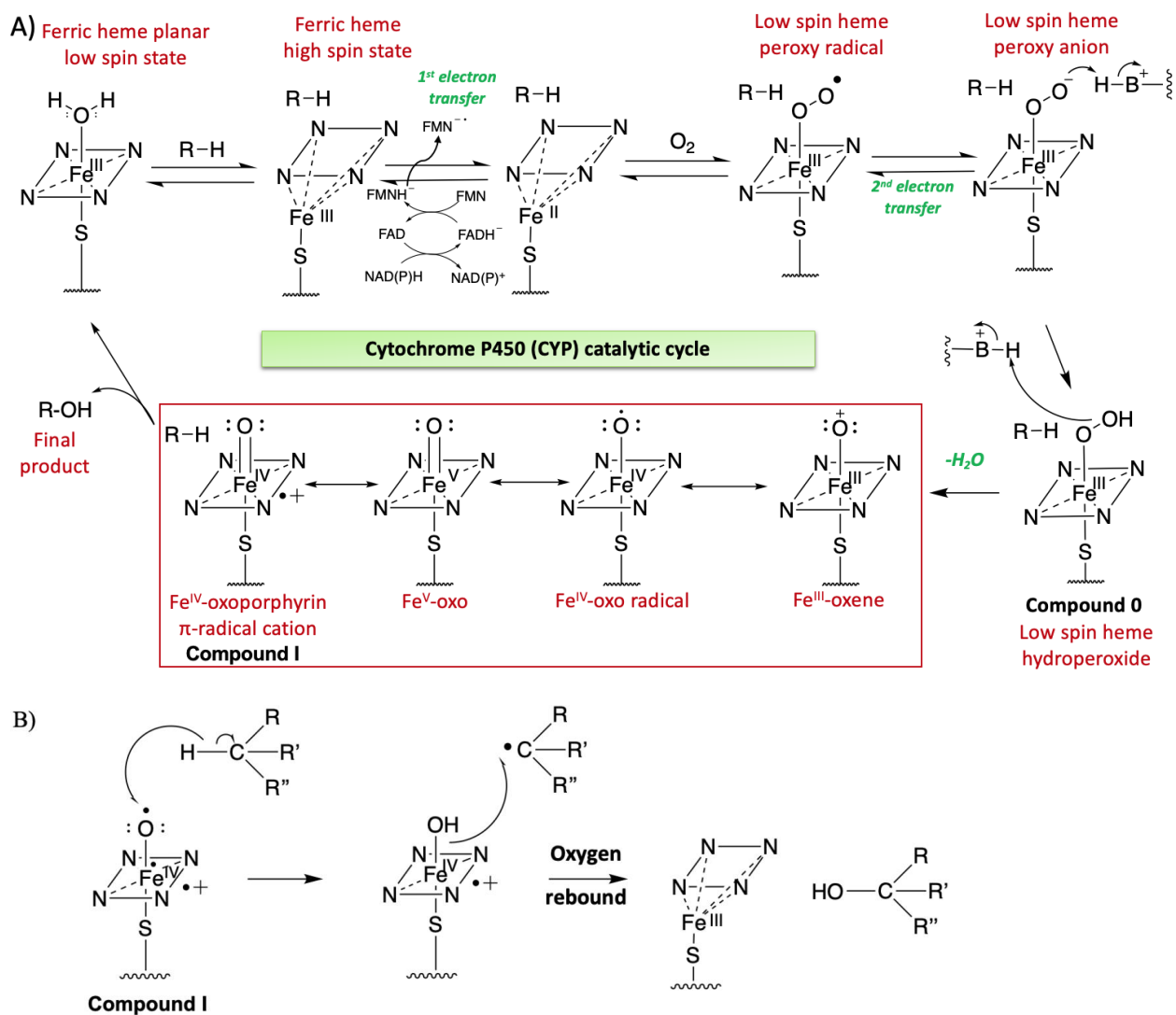
In catalysis, a direct reaction of a molecular oxygen ( $\text{O}_2$ ) which is in triplet state with a carbon molecule in a singlet state is spin-forbidden [Bugg, T.D. 2003; Guengerich, F. P. and Yoshimoto, F. K. 2018; Hernández, A. et al. 2015]. Cytochrome p450 metalloproteins overcome this barrier by complexing the oxygen to a metal (iron), so that the metal-oxygen complex can react with carbon substrates, as well as some heteroatoms. The first industrial list of bio-

transformations in 1950s undertaken by filamentous fungi in the production of corticosteroids by using their cytochrome p450 enzymes includes hydroxylation reactions at carbon 11 of the steroid skeleton. These are a class of oxidation reactions that are difficult to achieve via chemical synthesis.

### *1.6.3c. Catalytic cycle of CYPs*

The catalytic cycle of p450, is shown in **Part A** of **Figure 1.16**. The first step in the catalytic cycle of p450 is the binding of substrate molecule (RH) that replaces the water ligand, and the low spin ferric heme state gets converted to high-spin state. The sequential input of two electrons is required to complete the catalytic cycle of p450 enzymes and these two electrons are derived from the cellular cofactors such as NADH (nicotinamide adenine dinucleotide) or NAD(P)H (nicotinamide adenine dinucleotide phosphate). These electrons, required to reduce the active site to restart the catalytic cycle, are transported to CYPs by their redox protein partners known as flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN). The next step in the catalytic cycle is the reduction of enzyme-substrate complex that results in the increase in redox potential of heme prosthetic from ferric (+3) to the ferrous (+2) state by the associated reductases (FAD and FMN). The following step is the binding of molecular oxygen to the iron center where electrons are transferred from heme-iron to the oxygen which gives a low spin heme peroxy radical also known as ferric superoxide complex. The second electron transfer generates a low spin heme peroxy anion or iron-peroxo (**Compound 0**), which is then protonated to give a low spin heme-hydroperoxide/iron-hydroperoxy (**Figure 1.16, part A**). A subsequent protonation promotes heterolytic cleavage of the O–O bond with the release of water molecule to form Fe (V)-oxo radical (**Compound I**). There are other possible high valent resonance forms of Compound I such as Fe, Fe (V)-oxo species and Fe (IV)-oxoporphyrin  $\pi$ -radical cation. Most of the reactions have been

shown to be catalyzed by  $\pi$ -radical cationic form of Compound I (the catalytic intermediate of CYP) via hydrogen abstraction from a substrate R–H bond (formally a proton-coupled 1-electron oxidation of the substrate) yielding compound II and a substrate radical. Following a hydroxyl radical rebound from the heme to the substrate to release a hydroxylated product which is both thermodynamically and kinetically favored (**Figure 1.16, part B**). The enzyme returns to the default ferric resting state restarting the catalytic cycle (**Figure 1.16, part A**) [McIntosh, J. A. et al., 2014; Silverman, R. B. 2002].



**Figure 1.16:** part A) Catalytic cycle of a cytochrome p450, part B) Hydroxylation reaction catalyzed by ‘**Compound I**’, one of the intermediates in the catalytic cycle of CYP [Silverman, R. B. 2002].

### 1.6.3d. Reactions catalyzed by CYPs other than hydroxylation

One of the distinguishing features of cytochrome p450 enzymes are the capabilities for regio- and stereo-specific oxidation of substrates, which makes these enzymes essential in the primary and secondary metabolism of organisms. The reactions catalyzed by p450 are more diverse than simply hydroxylation reactions. Some examples of the other types of reactions catalyzed by p450 enzymes are 1) alkene, arene epoxidation, 2) dealkylation of amines, sulphides, and ethers, 3) oxygenation of amines and sulfides, 4) oxidative deamination that leads to the conversion of primary amines to aldehydes, ketones and  $\text{NH}_3$ , 5) oxidative dehalogenation-conversion of halides to aldehydes, ketones and HX, and 6) dehydrogenation, conversion of alcohols to aldehydes and aldehydes to acids [Guengerich, F. P. 2018; Guengerich, F. P. 2001; de Montellano, O. 2015; Isin, E. M. & Guengerich, F. P. 2007]. Most of these oxidations can be rationalized by mechanisms involving **Compound I (Figure 1.16, part A)**. **Compounds 0** (heme hydroperoxide) and other intermediates in the catalytic cycle of p450 (**Figure 1.16, part A**) have also been reported to catalyze some of the p450 transformations. For example, the iron-peroxo (or hydroperoxide) intermediate can mediate epoxidation and sulfoxidation under some circumstances; in others this species also carries out C–C bond cleavage [De Voss, J. J. et al., 2006; Jin, S. et al., 2003]. **Compound I** act as an electrophile contrary to **Compound 0** and the other intermediates, which are nucleophilic in nature.

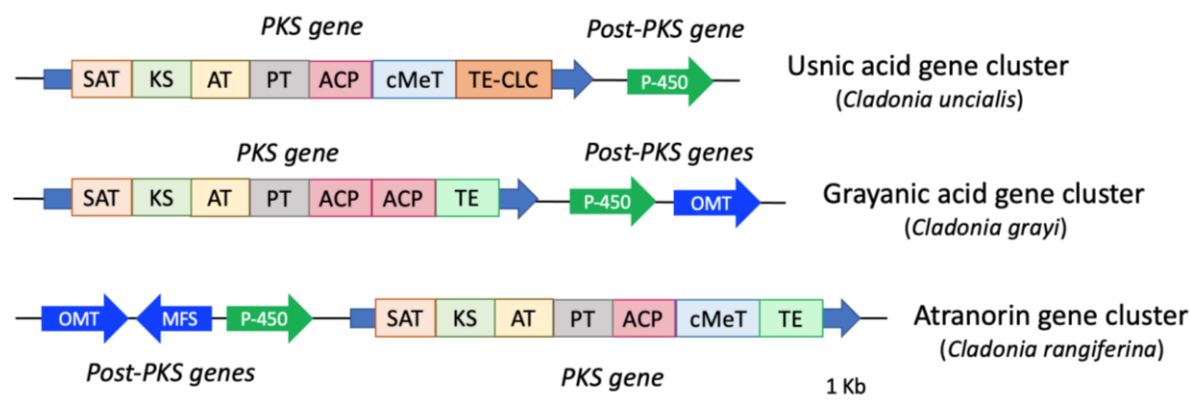
Though many potential oxidants occur during the cycle of P-450s, but often these enzymes are quite specific in the reactions that they catalyze. Specificity is directed by protein sequences molded by the force and filter of natural selection to favour certain intermediates while tuning their reactivity and selectivity.

### 1.6.3e. C-C & C-O coupling reactions catalyzed by CYPs

Organic molecules known as biaryls with direct carbon-carbon bonds between aromatic rings, are notoriously difficult to chemically synthesize in the laboratory. However, CYPs have been known to catalyze these biaryl formations efficiently. There are many examples that have been reported of p450 catalyzing phenolic coupling reactions such as lignin, melanin, teicoplanin, vancomycin, flaviolin dimers and trimers, bicoumarin *P*-kotanin [Stadler, R., and Zenk, M. H. 1993]. The mechanism by which **Compound I** can mediate oxidation reactions in these cases is through sequential one-electron oxidations which is distinct from the above-mentioned hydrogen atom abstraction [Guengerich, F. P. 2018].

### 1.6.4. Function of Post PKS/Accessory genes (tailoring enzymes) in a polyketide BGC

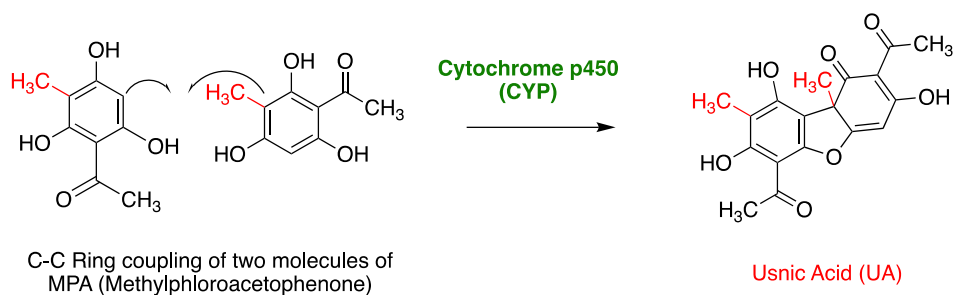
Investigation of fungal and a few lichen-fungal gene clusters reveals the function of many post-PKS/tailoring enzymes. UA BGC consists of two genes: a PKS gene and a post-PKS gene encoding a cytochrome p450 enzyme. Examples of lichen gene clusters involved in the biosynthesis of non-reducing polyketides such as UA, grayanic acid and atranorin are found in three different species of lichens named *C. uncialis*, *C. grayi*, and *C. rangiferina* where post-PKS genes are either upstream or downstream of a PKS gene (**Figure 1.17**).



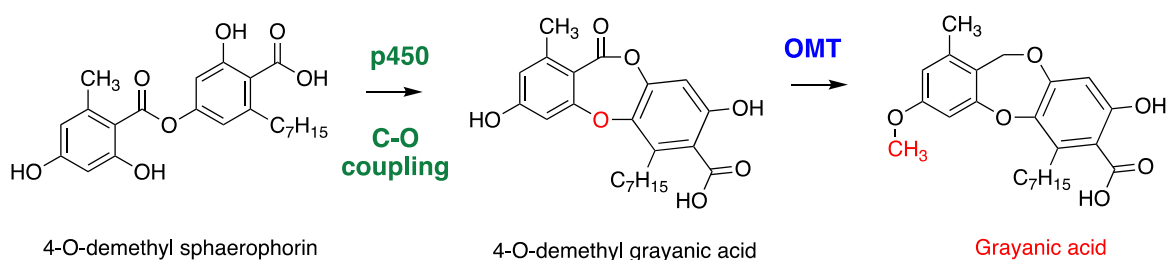


**Figure 1.17:** Post-PKS genes encode for different post-PKS enzymes such as cytochrome p-450 (P-450), O-methyltransferase (OMT). In three gene clusters responsible for grayanic acid, UA and, atranorin biosynthesis. MFS is a major facilitator superfamily (a general substrate transporter) [Bertrand, R. L. et al., 2018].

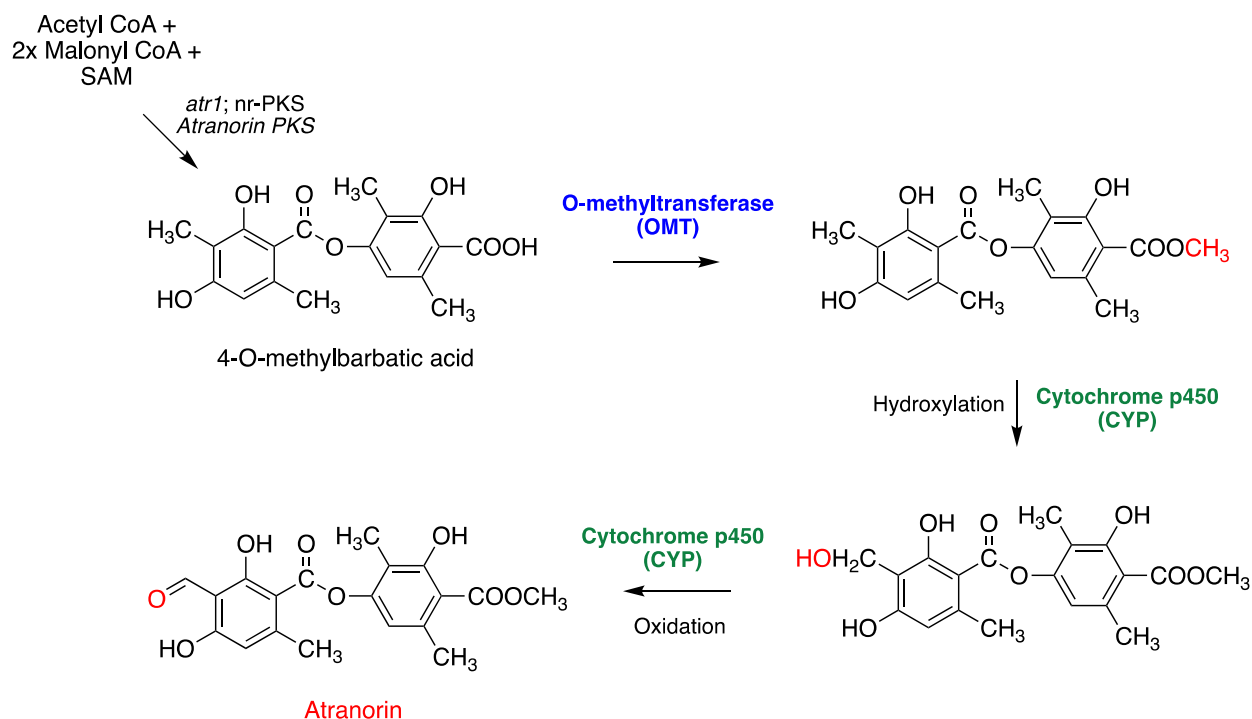
The function of CYP (a post-PKS enzyme) in UA (dibenzofuran) biosynthesis is the catalysis of iC-C coupling (a post-PKS reaction) of two molecules of MPA (**Scheme 1.2**) [Abdel-Hameed, M. et al., 2016]. In grayanic acid biosynthesis, cytochrome p450 catalyzes the C-O coupling between two aromatics to convert a depside (4-O-demethyl sphaerosporin) to a depsidone (4-O-demethyl grayanic acid) and O-methyltransferase (**OMT**) catalyzes the O-methylation of hydroxyl functional group of 4-O-demethyl grayanic acid leads to the formation of grayanic acid (**Scheme 1.3**) [Armaleo, D. et al., 2011; Bertrand, R. L. et al., 2018]. Post-PKS reactions in atranorin biosynthesis involves, 1) Oxidation of methyl group to hydroxy and subsequently to an aldehyde functional group on phenolic ring structure is catalyzed by p450 tailoring enzyme and, 2) O-methylation of hydroxyl and ester functional groups are catalyzed by OMT tailoring enzyme [Kealey, J. T. et al., 2021] (**Scheme 1.4**).



**Scheme 1.2:** Oxidative coupling of two MPA molecules to UA catalyzed by a CYP [Abdel-Hameed, M. et al., 2016].



**Scheme 1.3:** Post-PKS C-O coupling catalyzed by cytochrome p450 (CYP) and methylation of -OH catalyzed by O-methyltransferase (OMT) in grayanic acid biosynthesis [Armaleo, D. et al., 2011; Bertrand, R. L. et al., 2018].



**Scheme 1.4:** Post-PKS reactions catalyzed by OMT and CYP in atranorin biosynthesis [Kealey, J. T. et al., 2021].

## 1.7. Proposed mechanisms of cytochrome p450

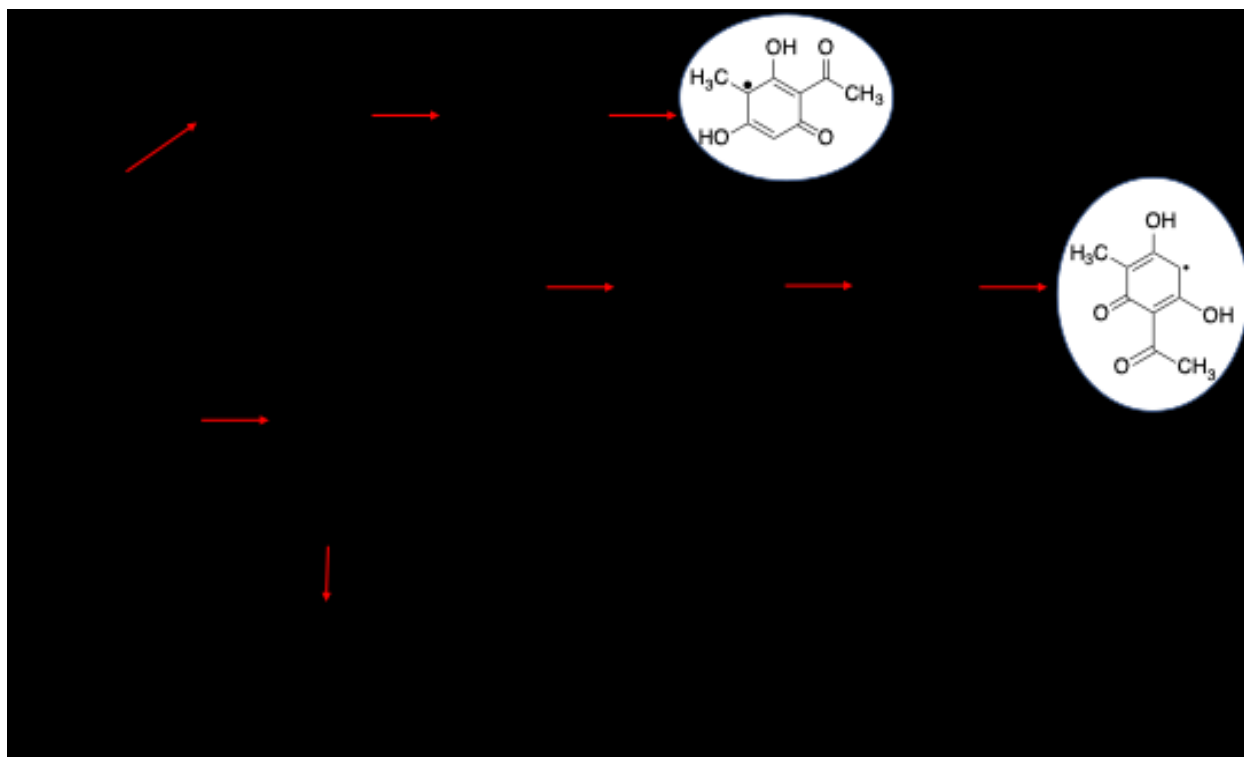
In this study we propose the two different catalytic mechanisms: radical and cationic, of cytochrome p450 mediated oxidation reactions during the biosynthesis of UA and grayanic acid.

### 1.7.1. Proposed radical-2 coupling mechanism catalyzed by MPAO in UA biosynthesis

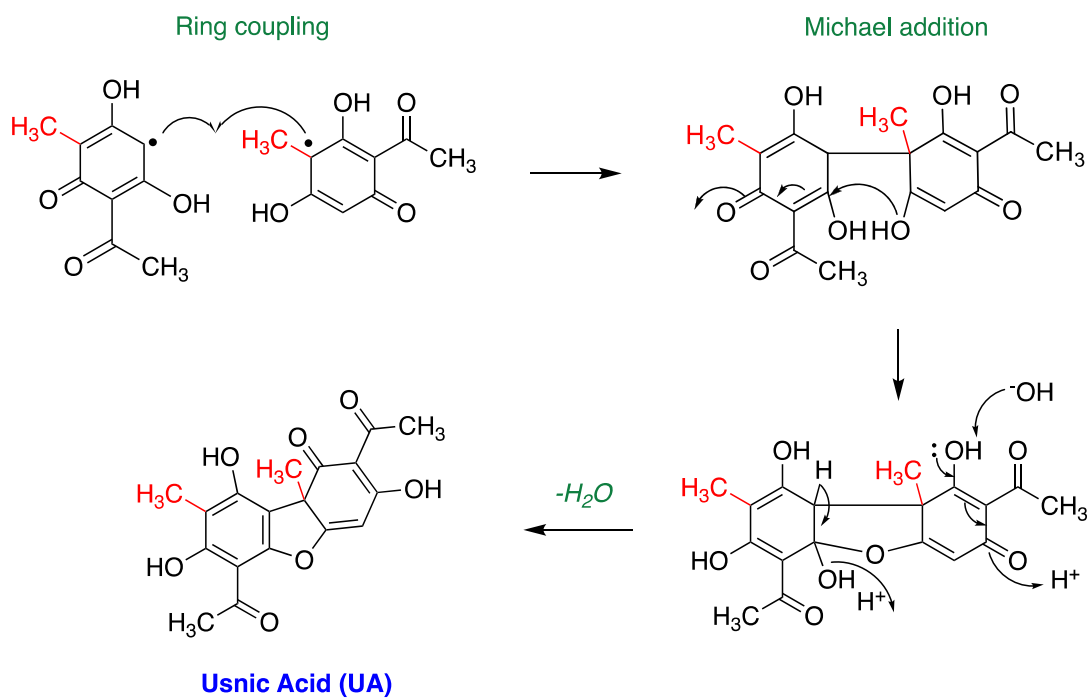
The identification of the UA gene cluster will be discussed in detail in **Chapter 3** based on genome sequencing work of *C. uncialis* [Abdel-Hameed, M. et al., 2016]. The oxidation reaction in UA biosynthesis is catalyzed by a cytochrome p450 (CYP) enzyme, also named as methylphloroacetophenone oxidase (MPAO) encoded by an accessory gene in UA gene cluster. MPAO catalyzes the oxidative dimerization of two molecules of MPA to UA (**Scheme 1.2**). Here, we propose the two different catalytic mechanisms for the oxidative enzyme (MPAO) that are plausible and consistent with the known chemistry of CYPs.

The first mechanistic step in the coupling of two radical forms of MPA to UA is the conventional hydrogen abstraction by the Fe<sup>IV</sup>-oxoporphyrin  $\pi$ -radical cation intermediate (**Compound I**) of CYP catalytic cycle. The initial proton abstraction has been presumed to occur at the phenolic -OH that in turn leads to the formation of an MPA radical with the electron density concentrated on the oxygen atom. The Fe<sup>IV</sup>-oxoporphyrin  $\pi$ -radical cation gets converted to Fe (IV)-hydroxo species (**Scheme 1.5**). It is expected that the MPA radical has the most significant resonance contributor as a tertiary radical (most stable radical) with the electron density highest on the carbon adjacent to the methyl group (**A**). The hydrogen abstraction from the second molecule of MPA by Fe (IV)-hydroxo species leads to formation electron density localized at O-atom which again gives the resonance stabilized MPA radical (**B**) (**Scheme 1.5**).

Further, we propose a radical coupling of two resonance-stabilized radicals; **A** and **B** followed by the Michael addition and condensation steps that could lead to the most widely occurring form of the UA molecule (**Scheme 1.6**). A similar CYP mechanism has been proposed in case of the synthesis of flaviolin dimers, extracted from a bacterial species *Streptomyces coelicolor* A3(2) [Zhao, B. et al., 2005].



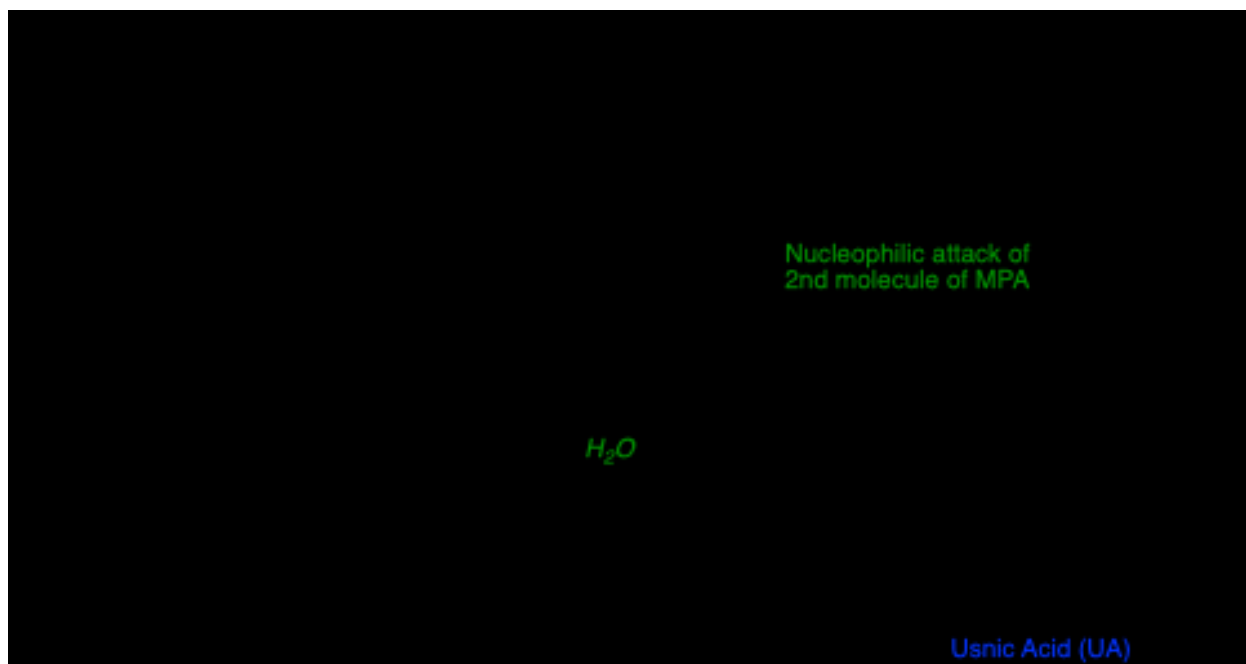
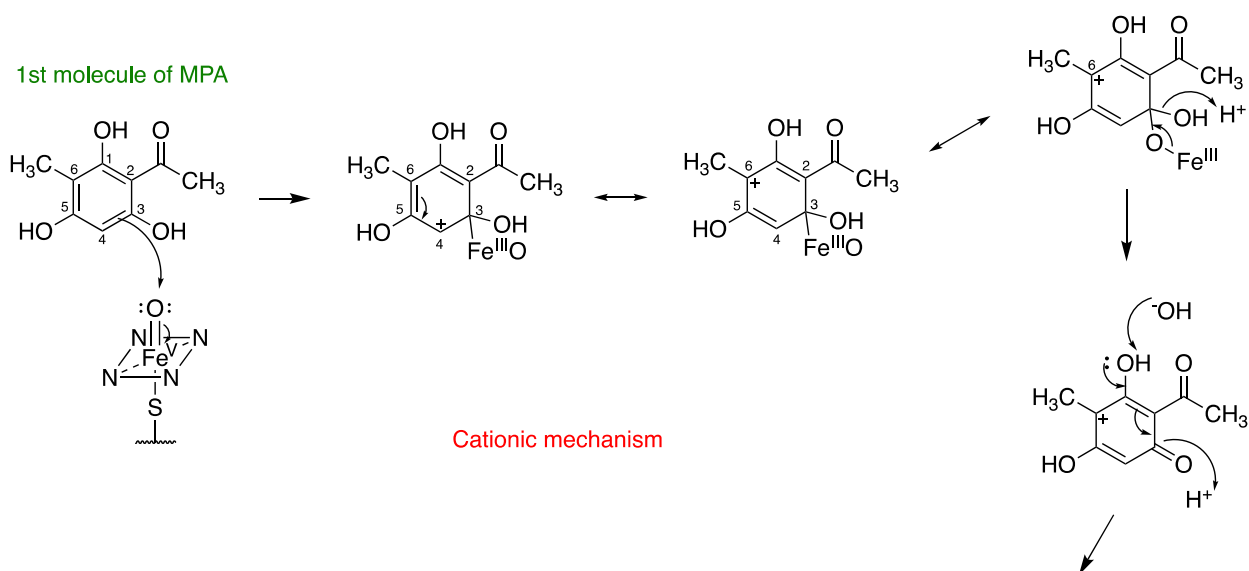
**Scheme 1.5:** Cytochrome p450 (MPAO) in UA biosynthesis catalyzes the two consecutive H-abstractions of from two MPA substrate molecules to two resonance-stabilized MPA radicals; **A & B**.



**Scheme 1.6:** Radical-radical C-C oxidative coupling of two MPA radicals to UA.

### 1.7.2. Cationic mechanism of MPAO

The **cationic mechanism** that has been proposed in case of MPA oxidation catalyzed by MPAO is also via cationic intermediate, where the nucleophilic attack on Fe<sup>V</sup>-oxo (one of the intermediates in the catalytic cycle of CYP) from C3-C4 double bond of benzene ring of first MPA molecule makes a FeO-C-3 complex (**Scheme 1.7**). This covalent intermediate then further undergoes the rearrangement to achieve the localization of positive charge at C-6 that is the most stable tertiary carbocation followed by the scission of Fe-O bond. This results in the oxidation of the benzene ring along with the dissociation of cationic MPA intermediate from the enzyme (**Scheme 1.7**) and a concerted removal of water molecules takes place. Rather than attack of the iron-heme center of the p450 on the second molecule of MPA, there is a nucleophilic attack of this second molecule of MPA on the cationic MPA. These steps then lead directly to the formation of a C-C bond (**Scheme 1.7**). The next two steps are the formation of a furan ring via Michael addition and the release of a water molecule giving the final product (UA) are identical to the earlier proposed radical mechanism (**Scheme 1.7**). Yet there is no experimental evidence to support these proposed mechanisms. However, we have some very preliminary data of DFT calculations, and this work has been in collaboration with Dr. Rebecca Davis that offers some support for the radical mechanism proposed in **Scheme 1.5** (not reported in this study). Specifically, the calculations suggest that the second oxidation potential to the cation proposed in **Scheme 1.7** is too high to be achievable by CYP enzyme. However, this work is still in progress and will require further computational or experimental results to confirm this hypothesis.

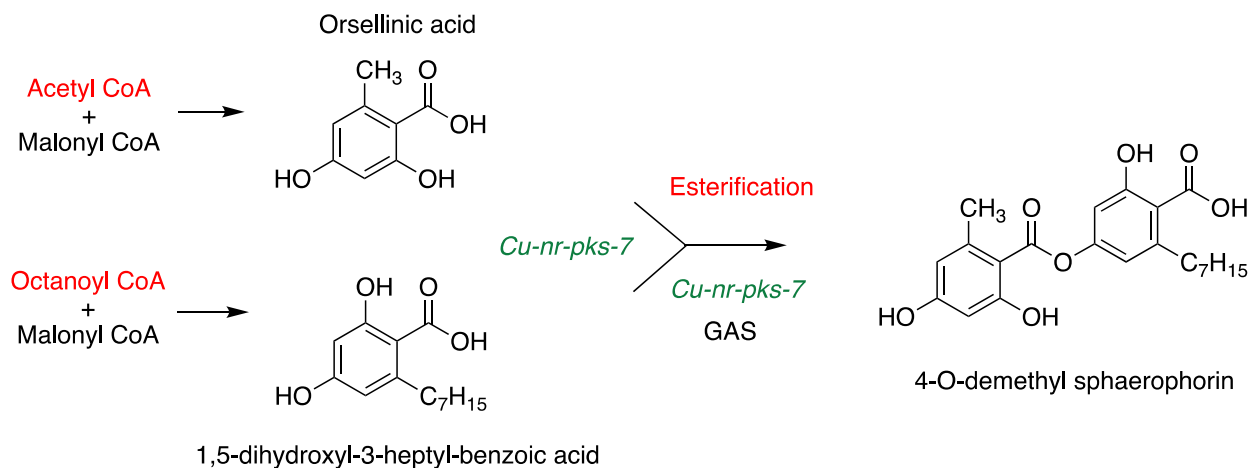


**Scheme 1.7:** Proposed cationic mechanism '2' of cytochrome p450 that catalyzes the oxidative conversion of two molecules of MPA to UA.

### 1.7.3. Grayanic acid gene cluster and its biosynthetic scheme

The lichen natural product grayanic acid has been putatively linked to a gene cluster found in *C. uncialis* genome which consisted of a core non reducing polyketide synthase *Cu-nr-pks-7* and was flanked by a series of accessory genes [Bertrand, R. L. et al., 2018]. Based on homology

mapping and AntiSMASH results, this *C. uncialis* gene cluster is found to be a possible functional homologue [Bertrand, R. L. et al., 2018] of grayanic acid cluster found in the genome of *C. grayi* [Armaleo, D. 2011]. The gene cluster from *C. uncialis* has three genes: one encodes for a PKS enzyme that has been proposed to catalyze the synthesis of orsellinic acid from acetyl CoA and malonyl CoA and the synthesis of 1,5-dihydroxyl-3-heptyl benzoic acid from octanoyl CoA and malonyl CoA (**Scheme 1.8**). The functions of two accessory genes encode for cytochrome p450 and O-methyltransferase (OMT) enzymes have already been discussed in **Section 1.6.4 (Scheme 1.3)**.



**Scheme 1.8:** Grayanic acid biosynthesis [Armaleo, D. 2011; Bertrand, R. L. et al., 2018].

#### 1.7.4. Proposed radical-2 C-O coupling mechanism catalyzed by CYP

The first p450 catalytic mechanism in grayanic acid biosynthesis is rationalized with ‘**Compound I**’ (**Scheme 1.9**), which is an Fe<sup>IV</sup>-oxoporphyrin  $\pi$ -radical cation. The initial oxidation of 4-O-demethyl sphaerophorin is postulated to occur from hydrogen atom abstraction from the (phenolic) OH (**Scheme 1.9**). This oxidation leads to the formation of a radical intermediate of 4-O-demethyl sphaerophorin and further rearrangement of this radical is proposed to localize the electron density at C-4 atom of 4-O-demethyl sphaerophorin (**Scheme 1.9**). When

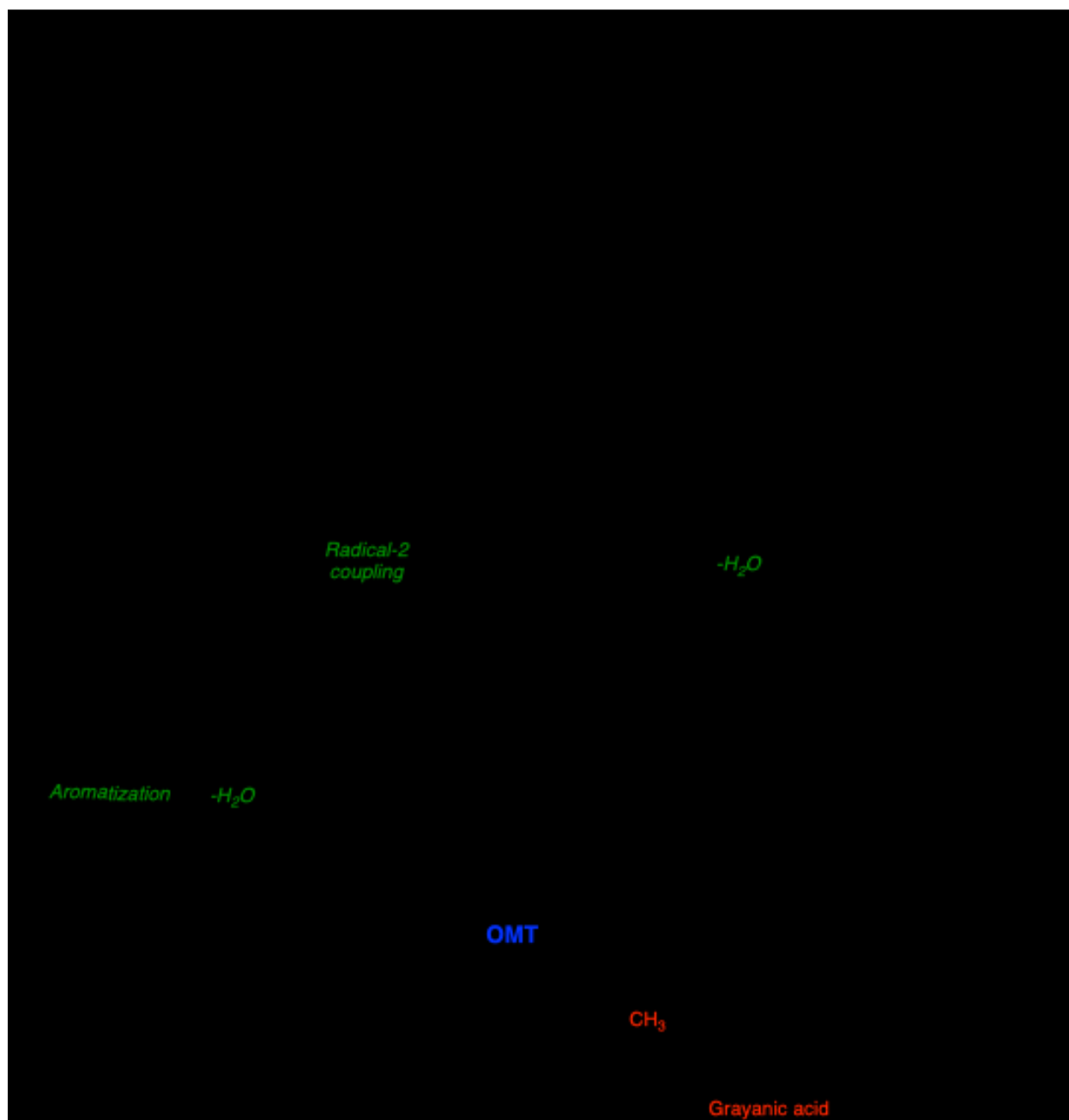
the substrate is still in the active pocket of the p450 enzyme, the second H-abstraction can be proposed to proceed via a Fe<sup>IV</sup>-hydroxo species that in turn results in the formation a biradical species (**Scheme 1.9**). This biradical species undergoes an intramolecular radical C-O coupling followed by a release of water molecule and the re-aromatization of this ring drives the formation of 4-O-demethyl grayanic acid (**Scheme 1.9**). The final step has been proposed to be catalyzed by OMT (O-methyltransferase) to convert 4-O-demethyl grayanic acid to grayanic acid (**Scheme 1.9**) [Armaleo, D. et al., 2011].

There is a second plausible radical C-O coupling mechanism (**Scheme 1.10**) which follows steps nearly identical to those in **Scheme 1.9**. The second mechanism differs only in the conformation of an active pocket of CYP enzyme that could impact the angle at which the hydrogen atom abstraction occurs. This speculation is just theoretical, there is no experimental data to back this theory. This theory is being postulated based on the previously reported crystal structures of cytochrome p450s that catalyze C-O and C-C radical-2 couplings [Zhao B. et al., 2005; Guengerich, F. P. et al., 2016].

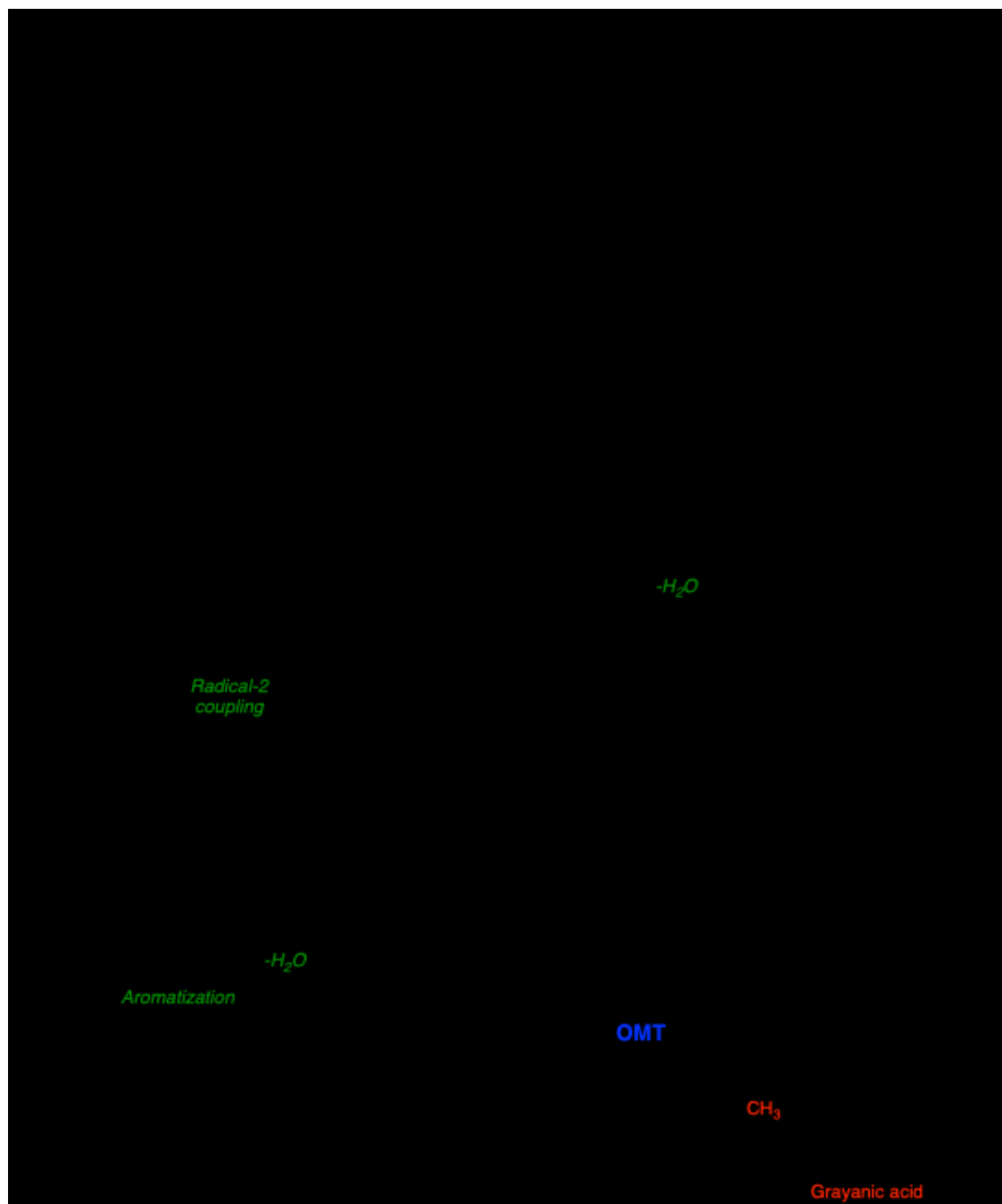
The third proposed radical-radical C-O coupling mechanism (**Scheme 1.11**) involves the formation of an arene oxide, which is a well-established precedent for cytochrome p450 enzymes [Guengerich, F. P. 2001]. In this study, we propose that the **Compound I** make a complex with a substrate molecule via one-electron oxidation followed by arene oxide/epoxide formation. Rather than a usual rearrangement of the arene oxide to an aromatic system, a nucleophilic attack by the -OH of the second phenol ring of the substrate results in the formation of ether linkage between two phenol rings. A subsequent elimination of water would lead to the formation of 4-O-demethyl grayanic acid (**Scheme 1.11**). The last step is catalyzed by OMT that converts 4-O-demethyl



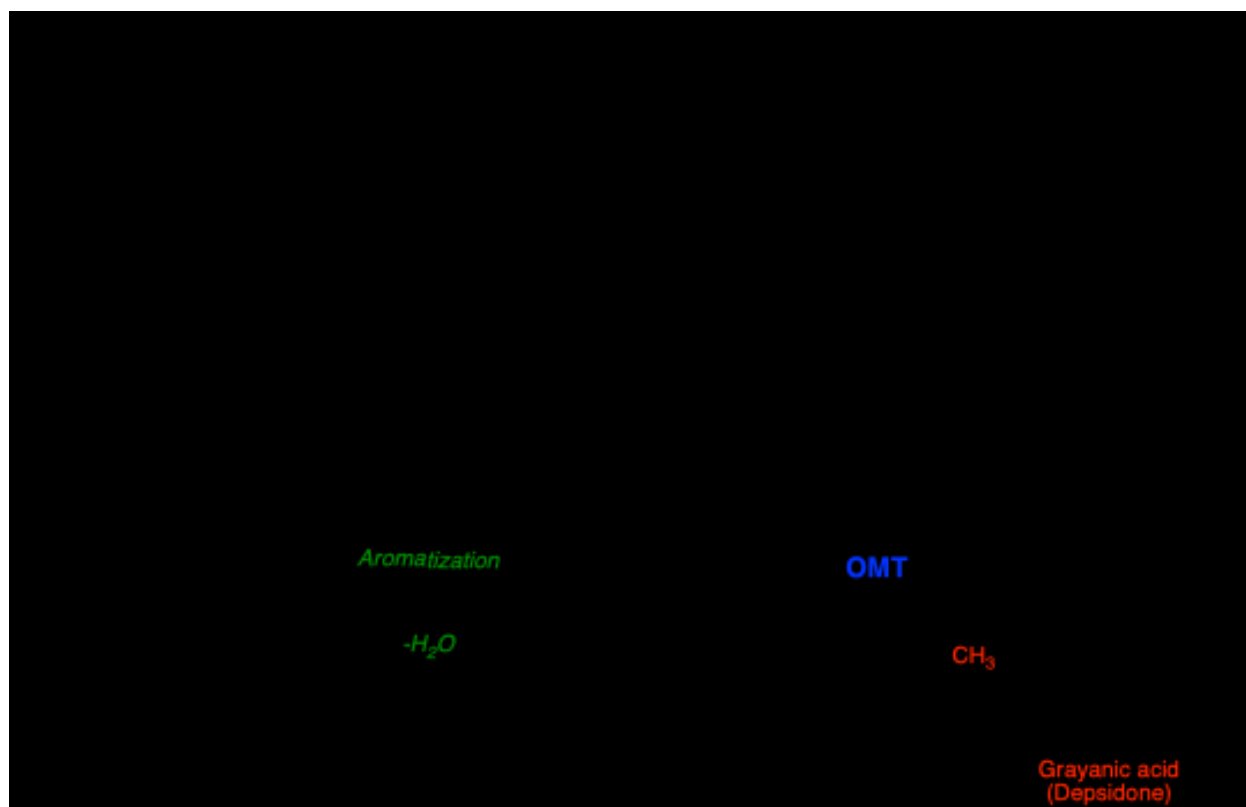
grayanic acid (a depside) to the desired product grayanic acid (a depsidone) identical with the other proposed biosynthetic schemes [Armaleo, D. et al., 2011].



**Scheme 1.9:** First proposed radical-2 C-O coupling in grayanic acid biosynthesis.



**Scheme 1.10:** Second proposed radical-2 C-O coupling in grayanic acid biosynthesis.



**Scheme 1.11:** Third proposed radical-2 C-O coupling in grayanic acid biosynthesis that involves an epoxide formation.

## 1.8. Biosynthesis of Reducing Polyketides

Very few examples of reduced polyketides (r-PKS) have been published so far that are found in lichens. Out of those very few examples, the two research articles that have been published our research group in 2018 demonstrate 16 fully annotated putative reducing polyketide synthase gene clusters based on the genomic studies of the lichen species *C. uncialis*. Where we proposed [Bertrand, R. L. et al. 2018, 723-731 and 732-748]. Some of these have been linked to the biosynthesis of already known polyketide molecules, for example patulin (mycotoxin) and betaenone. There is one more article published where the author reported the identification of 6-MSAS like PKS in a lichen species [Schmitt, I. et al., 2008]. The **Chapter 6** of this study, embarks upon more work on proposed biosynthetic schemes of reduced polyketides catalyzed by r-PKSs gene clusters found in *C. uncialis*. There are some popular model examples of fungal r-PKS gene

clusters that are involved in the biosynthesis of reduced fungal polyketides such as lovastatin (cholesterol-lowering drug), compactin, equisetin, solanapyrones A-D, alternapyrone etc [Kasahara, K. et al., 2010; Fujii I. et al., 2005].

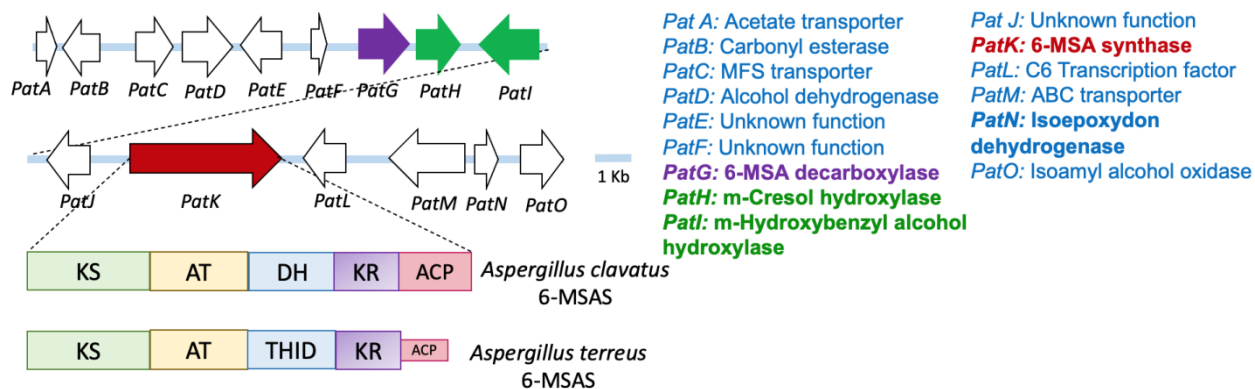
### 1.8.1. Highly reducing v/s partially reducing PKSs

As mentioned earlier in **Section 1.4.3**, the unique feature of r-PKSs (reducing polyketide synthases) is the presence of KR-DH domains, where a keto-reductase (KR) domain reduces the beta-keto group to hydroxyl, succeeding dehydration of the hydroxyl group to an alkene catalyzed by a dehydratase (DH) (**Figure 1.10**). On some occasions, the participation of the ER domain (reducing catalytic domain) has also been shown to be involved in the reduction of alkene which has been generated by the DH catalytic step (**Figure 1.10**). LovB (lovastatin nonaketide) is an example of highly reducing PKS (HR-PKS/R-PKS), where all three reductive domains (KR, DH, and, ER) are present [Chantel D. Campbell and John C. Vederas, 2010; Ma S. M. et al., 2009]. 6-methyl salicylic acid synthase (6-MSAS) is a partially reducing PKS (PR-PKS) where only KR-DH domains are involved in the biosynthesis of 6-methyl salicylic acid (6-MSA) and the ER domain is absent [Peul, O. et al., 2010] (**Figure 1.19**).

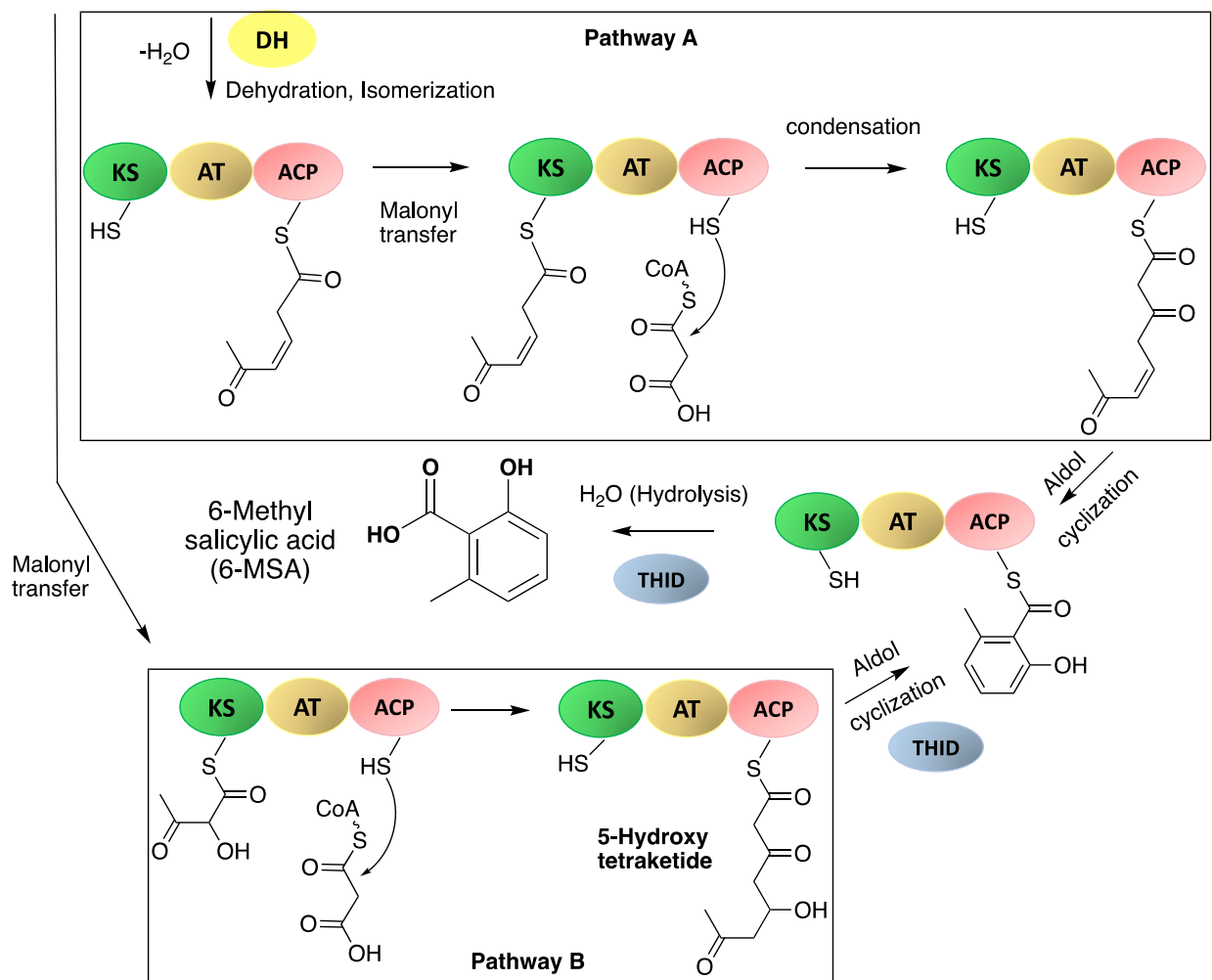
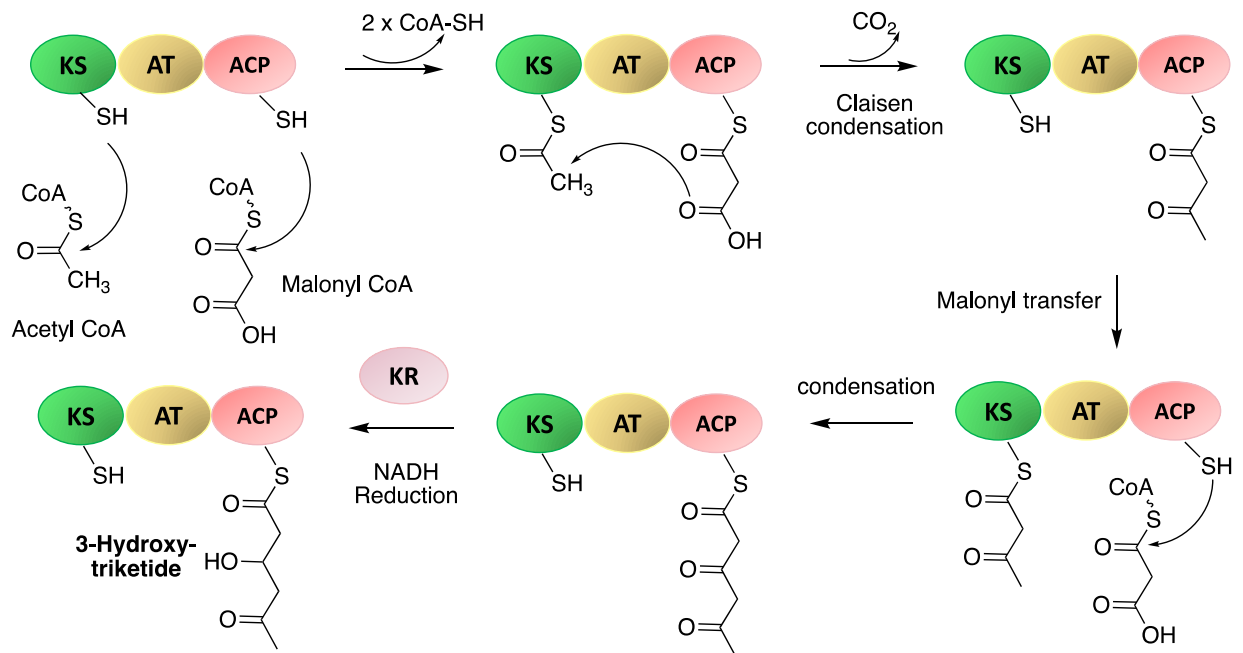
### 1.8.2. Biosynthesis of first polyketide, 6-MSA

6-MSA was the first polyketide to be biosynthetically investigated [Birch, A. J. et al., 1955; Lynen, F. et al., 1961] and isolated from various species of fungi belonging to *Penicillium*, *Aspergillus* etc. 6-MSA is a catalytic intermediate in patulin biosynthesis and it is assembled by the programmed chain extension cycles of 6-MSAS from acetyl CoA and malonyl CoA (**Figure 1.18**). A few more additional post-PKS reaction steps ultimately convert 6-MSA to patulin (**Scheme 1.12**) [Tannous, J. et al., 2014]. Patulin biosynthetic gene clusters found in different species of *Aspergillus* fungus that have been reported to have a total of 15 genes out of

which one is the PKS gene (6-MSAS), few are regulatory genes, and others are the post-PKS genes. The function of many these post-PKS genes is unknown except '4' genes are with the known function, 6-MSA decarboxylase (PatG), Isoepoxydon dehydrogenase (PatN), m-cresol hydroxylase (PatH) and m-hydroxybenzyl alcohol hydroxylase (PatI) (**Figure 1.18**) [Peul, O. et al., 2010]. Two separate pathways have been reported that lead to the formation of 6-MSA, 1) pathway A, which involves a 3-hydroxytriketide intermediate tethered to an acyl carrier protein that undergoes dehydration and isomerization and, 2) pathway B involves the conversion of 3-hydroxytriketide to 5-hydroxytetraketide which further cyclizes, dehydrates, and aromatizes prior to the TH (thioester hydrolase)-mediated product release (hydrolysis) (**Figure 1.19**). TH domain has been recently renamed as THID; thioester hydrolase with interdomain linker [Parascandolo, J. S. et al., 2016] which is a DH domain with an adjacent region of interdomain (ID) linker (**Figure 1.19**). These two different pathways exist because of the presence of the THID domain of 6-MSAS found in *Aspergillus terreus* in place of the DH domain of *Aspergillus clavatus* 6-MSAS (**Figure 1.19**) [Moriguchi, T. et al., 2010, Parascandolo, J. S. et al., 2016].

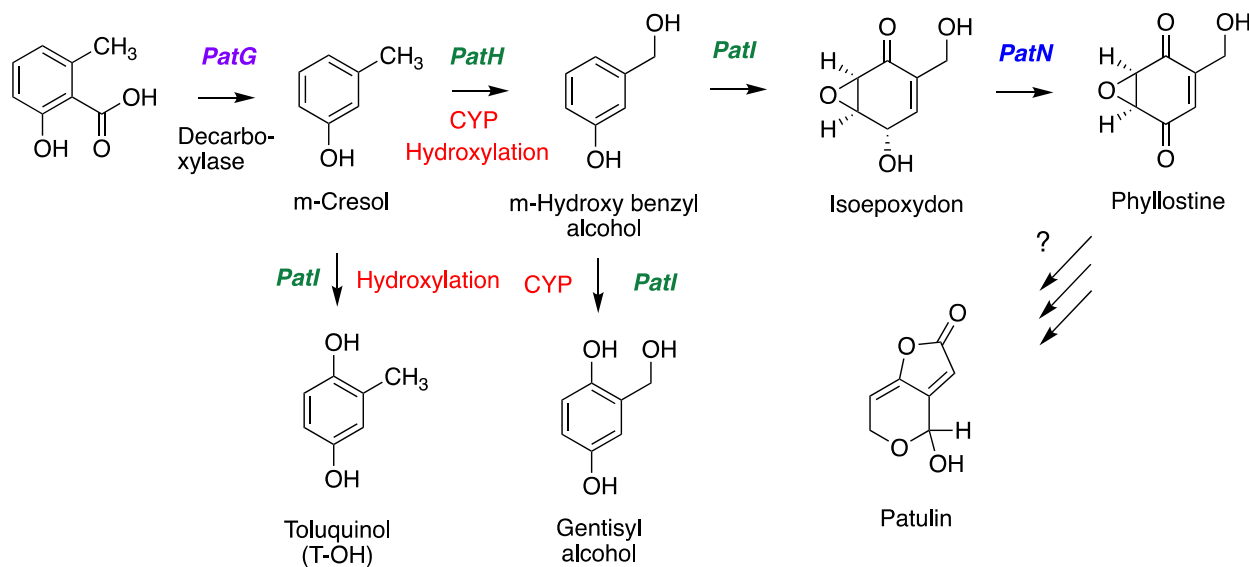


**Figure 1.18:** Patulin biosynthetic gene cluster with 15 different genes named *PatA-O*. *Aspergillus clavatus* 6-methylsalicylic acid synthase (6-MSAS)/PKS/*PatK* domain architecture: ketosynthase (KR) domain, acyltransferase (AT) domain, dehydratase (DH) domain, ketoreductase (KR) domain and acyl carrier protein (ACP) and domain architecture of 6-MSAS found in *A. terreus* with THID domain in place of DH domain [Bertrand, R. L. et al. 2018, 723-731 and 732-748].



**Figure 1.19:** Biosynthesis of 6-methylsalicylic acid (6-MSA) catalyzed by 6-methylsalicylic acid synthase (6-MSAS). Two different catalytic pathways: Pathways A and B with two different intermediates, 3-hydroxytriketide in former and 5-hydroxtetraketide in the latter. Enzymatic activity of two different domains; dehydratase (DH) in pathway A and thioester hydrolase interdomain linker (THID) in pathway B, ultimately leads to the formation of 6-methylsalicylic acid [Moriguchi, T. et al.,2010, Parascandolo, J. S. et al., 2016].

Post-PKS reactions in patulin biosynthesis are catalyzed by post-PKS enzymes named as PatG, PatH, PatI and PatN that catalyze the modification of 6-MSA (**Scheme 12**). PatG is a decarboxylase that catalyzes the decarboxylation of 6-MSA leads to the formation of m-cresol. PatH and PatI are the two cytochrome p450s, PatH catalyzes the hydroxylation of a methyl group on m-cresol leading to the formation of m-hydroxy benzyl alcohol. Whereas PatI catalyzes the hydroxylation of the benzene ring in m-cresol to toluquinol and of m-hydroxy benzyl alcohol to gentisyl alcohol. PatI has also been known to catalyze the epoxidation of m-hydroxy benzyl alcohol to isoepoxydon. PatN, named isoepoxydon dehydrogenase catalyzes the conversion of isoepoxydon to phyllostine. The remaining genes with the unknown function in this gene cluster are predicted to catalyze the further steps in the biosynthesis final molecule, patulin.



**Scheme 1.12:** Post-PKS reactions catalyzed PatG; 6-MSA decarboxylase, PatH; m-cresol hydroxylase, PatI; m-hydroxy benzyl alcohol hydroxylase, PatN; Isoepoxydon dehydrogenase, in patulin biosynthesis [Moriguchi, T. et al.,2010, Parascandolo, J. S. et al., 2016].

## 1.9. Terpenes

Terpenoids constitute a large group of natural products [Christianson, D. W., 2008; Quin, M. B. et al., 2014]. This structurally diverse class of terpenes is derived from the simple five-carbon precursor (C5 isoprene units) molecules dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP) (**Figure 1.20**). In lichen fungi, these two isomers are produced from acetyl-CoA via mevalonate pathway (**Figure 1.5, Table 1.1, and Figure 1.20**) [Miziorko, H. M., 2011]. Condensation of IPP and DMAPP monomers results in linear prenyl diphosphate chains of varying length such as C10 geranyl pyrophosphate (GPP), C15 (2E,6E)-farnesyl pyrophosphate ((2E,6E)-FPP, or FPP), and C20 geranylgeranyl pyrophosphate (GGPP). These linear prenyl diphosphate chains undergo a dephosphorylation and cyclization cascade to produce terpenes. These highly complex reactions are catalyzed by enzymes known as terpene synthases [Christianson, D. W., 2006]. Depending on the length of the precursor molecule, fungal terpene synthases are known to produce monoterpenes (C10), sesquiterpenes (C15), diterpenes (C20), triterpenes and steroids (C30) (**Figure 1.20**). Further tailoring enzymes such as cytochrome p450s, different group transferases and oxidoreductases modify the initial terpene scaffold and produce the final terpenoid SM.

The combination of two isoprene units, IPP and DMAPP in a head-to-tail fashion leads to the formation of geranyl pyrophosphate (GPP) and GPP further undergoes some catalytic modifications to produce a monoterpene, limonene (**Figure 1.20**). Limonene shows antimicrobial and antifungal activities and is extracted by hydro-distillation from *Evernia prunastri* [Kahriman, N. et al., 2011]. GPP reacts with the first isoprene unit to give farnesyl PP (3-isoprene units) and all the sesquiterpenes are derived from FPP. Graphilane has been recently isolated from a lichen fungal species and found to be cytotoxic against K592 cancer cell lines upon evaluation. This

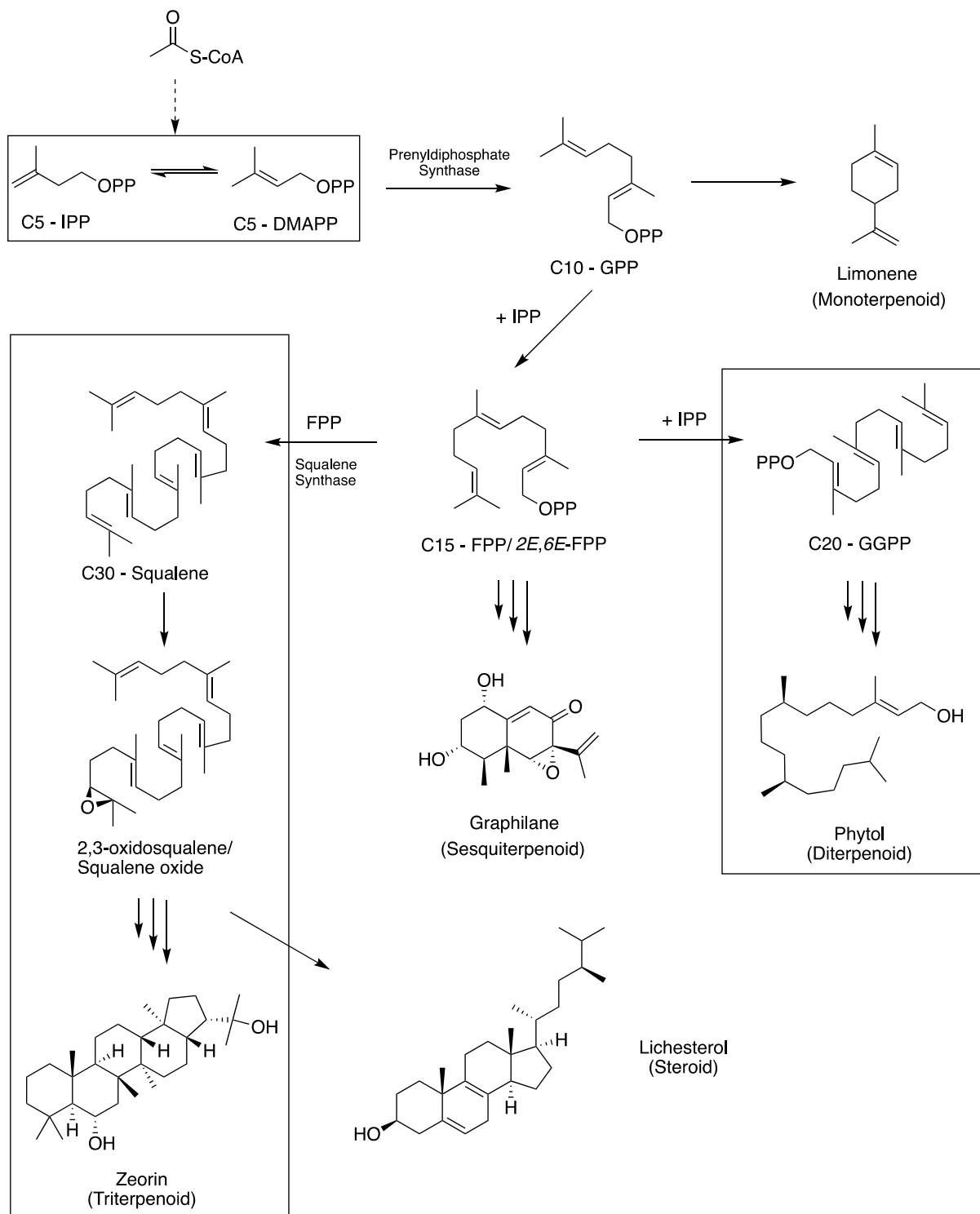


lichen mycobiont is isolated from native lichen *Graphis* sp. in Vietnam and cultivated in test tubes [Vo, V. G. et al., 2022]. Farnesyl PP reacts with the second isoprene to give geranylgeranyl PP (GGPP, 4-isoprene units). GGPP ultimately gets converted to phytol, one of the best-known diterpenes (**Figure 1.20**) that shows antimycobacterial activity against *Mycobacterium tuberculosis* [Rajab, M. S. et al., 1998]. Triterpenes are derived from squalene, six isoprene units containing a compound where one molecule of farnesyl pyrophosphate added to another molecule of FPP leads to the formation of C<sub>30</sub> – Squalene. Squalene further oxidizes to 2,3-oxidosqualene/squalene oxide and this squalene oxide cyclizes to a lichen triterpenoid, zeorin [Shukla, V. et al., 2010] and a lichen steroid, lichesterol (**Figure 1.20**) [Goga, M. et al., 2018]. Triterpenes and sterols are the cyclized products of squalene. Zeorin, isolated from lichen *Lecanora frustulosa* exhibits antibacterial and antifungal activities [Kosanić, M. et al., 2015]. Previous studies have demonstrated that sterols may play a role in the membrane permeability in the lichen thallus and lichesterol has been isolated and characterized in many lichen species such as *Usnea longissima*, and *Lobaria pulmonaria*, *Lobaria scrobiculate* and *Ramalina Africana* [Shukla, V. et al., 2004].

### **1.10. Genome sequencing, BGCs and mining of novel SMs**

The discovery of novel natural products or SMs has ramped up with the advent of whole genome sequencing technology. The declining cost and expanding availability of this technology over the years have increased the input of information on biosynthetic gene clusters [Goodwin et al., 2016]. Examples include large-scale microbial genome sequencing using Illumina and low-throughput short read Sanger sequencing. In addition to genome sequencing, the development and improvement of programs designed to detect genes associated with secondary metabolism, such

as, AntiSMASH [Blin et al., 2017], SMURF [Khaldi et al., 2010], and PRISM [Skinnider et al., 2017] have also revolutionized the quantum of the secondary metabolome.



**Figure 1.20:** Classification of lichen terpenoids [Kahrman, N. et al., 2011; Goga, M. et al., 2018].

Although genome sequencing projects in bacteria, plants, and fungi have resulted in the discovery of a plethora of biosynthetic gene clusters encoded within organisms, this number vastly exceeds the number of SMs known to be produced by these organisms [Schorn et al., 2016; Huang et al., 2017; Nielsen et al., 2017]. Most gene clusters are ‘cryptic’ meaning they cannot be associated with any SM or are speculated to encode undiscovered SMs. Investigating these ‘cryptic’ clusters using biotechnological and bioinformatic tools is a process known as ‘genome mining’. This process has led to the identification of new natural products [Cacho et al., 2015; Ziemert et al., 2016].

There are several means of investigating the biosynthetic function of gene clusters. In the case of ‘cryptic’ gene clusters, one method is to activate their expression and look for any metabolic changes. If silence is due to a regulatory ‘switch’ that is only ‘turned on’ given a specific stimulus (e.g., UV radiation, co-culturing with other organisms, etc.), applying that stimulus could result in gene expression and metabolite biosynthesis [Pettit, 2009; Bode et al., 2002; Tanaka et al. 2010]. Alternatively, the regulatory ‘switches’ can be engineered within the organism to either induce or increase the rate of gene expression [Aigle and Corre, 2012; Suroto et al., 2017; Saha et al., 2017]. This strategy is becoming increasingly widespread due to the advent of CRISPR Cas9 gene-editing technology [Zhang et al., 2017]. In cases where the metabolite is known but the associated gene cluster is speculative, the second method of investigation is to knock out candidate genes and observe whether the studied metabolite is no longer being produced. This technique is particularly useful for dissecting the function of encoded enzymes and how they individually contribute to the formation of a complete biosynthetic pathway [Alhawatema et al., 2017; Fuller et al., 2015]. The third method of investigation is to insert biosynthetic genes into a host and observe *de novo* metabolite biosynthesis within the host, a process known as a heterologous

expression [Zhang et al., 2015; Alberti et al., 2017]. A variety of expression hosts have been developed specifically for the functional heterologous expression of biosynthetic gene clusters from prokaryotes and eukaryotes alike [Anyagou and Mortensen, 2015; Yaegashi et al., 2014; Billingsley et al., 2016; Li and Neubauer, 2014; Loeschcke and Thies, 2015; Gomez-Escribano and Bibb, 2014; Baltz, 2010; Ikram et al., 2015].

The motivation to find and functionally characterize biosynthetic gene clusters rests in the biotechnological applications arising from their cloning and heterologous expression. One common barrier to the commercialization of natural products concerns their mass production. The most obvious way to produce a biogenic compound is to cultivate the producing organism at a commercial scale, then extract the compound from it. However, this approach is often untenable because cultivation is difficult, and expensive, and the native organism produces the compound in minuscule quantities. Although one may initially suggest that this problem could be overcome through chemical syntheses, such a task is greatly complicated by the numerous stereocenters and functional groups that are often present on complex SMs. For instance, the major barrier to the commercialization of Taxol, an FDA-approved anti-cancer drug produced by Pacific yew trees, was that it required cutting down three trees to produce enough drugs to treat one patient. This was an ecologically devastating process that was met with adverse public opinion. The development of a protocol for the complete chemical synthesis of Taxol was justifiably heralded as a breakthrough, nonetheless, requiring 17 (!) chemical steps [Nicolaou et al., 1994].

These problems could be overcome by harnessing the molecular machinery responsible for producing these complex molecules. The enzymes that are responsible for producing SMs have been ‘programmed’ by evolution to perform precise chemical modifications. If the genes encoding these biosynthetic enzymes were expressed in a heterologous host, and the enzymes purified, a

chemist could use these enzymes to resolve difficult steps via a semi-synthetic approach. Alternatively, if the complete biosynthetic gene cluster were to be expressed in a fast-growing host such as yeast or bacteria, commercially relevant amounts of the molecule might be produced in large bioreactors. For example, the mass production of the potent anti-malarial drug artemisinin by heterologous expression of its biosynthetic gene cluster in yeast is arguably one of the greatest success stories within the field of biotechnology [Paddon et al., 2013].

Another fascinating field of natural products chemistry, known as ‘combinatorial biosynthesis’ emerged during the early 2000s, where various components of these molecular factories could be swapped to produce ‘unnatural’ derivatives of natural products. This aims to develop enzymes that can produce libraries of derivatives of naturally occurring molecules with improved drug-like properties [Sun et al., 2015; Bayly and Yadav, 2017; Winn et al., 2016].

### **1.11. Problems associated with lichens and the production of their SMs**

Although lichen SMs have a great deal of industrial importance and applications because of their extensive biological properties, yet their potential to be fully explored and utilized. One of the major reasons behind this is the very slow growth of lichens, less than 1cm/year. Lichenologists have developed a biometric dating method known as ‘lichenometry’ that can be used to estimate the age of ancient ruins or monitor climate changes, with a timescale on the order of several thousand years [Benedict, 2009; Armstrong, 2004]. Despite their wide range of ecological adaptation (mentioned earlier in **Section 1.3.1**), lichens prefer to grow in a familiar environment. As a result, it is very difficult to grow them in their non-native habitat. **Figure 1.21** shows one of the examples of the slow growth of lichens, where an experiment was conducted by Dr. Abdel-Hameed (an ex-PhD student in our lab) to grow a lichen fungus named *C. uncialis* in the laboratory conditions. Only a tiny circle of lichen fungal growth was observed after 18 months (**Figure 1.21**).



**Figure 1.21:** Experimental evidence of slow growth of a lichen fungus named *Cladonia uncialis*.

It is also very challenging to replicate the natural mutualism of lichen fungus and green alga in the lab. Many articles talk about the change in the secondary metabolome of a lichen species when comparing the samples from axenic cultures of mycobionts to the samples of cultures where mycobionts are grown in combination with photobionts [Calcott, M. J. et al., 2018]. All these challenges make the typical extraction method to discover novel lichen SMs/biomolecules untenable. Also, the elucidation of the function of gene clusters via gene knockout experiments in lichens is limited because of their extended lifespans.

Therefore, heterologous expression of a single gene, a cassette of genes, or an entire biosynthetic gene cluster within a genetically tractable host is a more practical approach to identifying and engineering the lichen natural product. In the post-genomic era, besides the transcriptional studies and heterologous expression of lichen BGCs, lichenologists have started going back to more traditional and non-definitive approaches to linking genes to metabolites. This includes the study of an evolutionary relationship between organisms by comparing their DNA/protein sequences, known as phylogenetics. A core premise of phylogenetics is that genes with common ancestry that evolved within a recent timeframe are more likely to encode proteins with similar functions compared to related genes within a distant timeframe. In SM biosynthetic studies, observation of common and recent ancestry between a subject and reference gene is used

as supporting evidence that the encoded proteins possess similar roles in metabolite biosynthesis [Ziemert and Jensen, 2012].

### **1.12. Summary of this research project objectives**

The lichen species that our research group works with is *Cladonia uncialis*. To achieve the successful functional expression of lichen biosynthetic genes within a heterologous host, the first rational step was the *de novo* whole-genome sequencing of *Cladonia uncialis*. The following step was the functional annotation of lichen biosynthetic gene clusters. The whole-genome sequencing of *C. uncialis* and functional annotation of its biosynthetic gene clusters was carried out between 2010-2018 by Dr. Abdel-Hameed and Dr. Bertrand (ex-PhD students). This was the time when whole-genome sequencing was quite expensive, and limited lichen genetic information was available. My Ph.D. research focus was to develop a method of heterologous expression and functional characterization of these annotated lichen genes (*C. uncialis*). This would in turn provide the methodological groundwork for the mass production and commercialization of interesting SMs from lichens. I have summarized the goals of my doctoral research into four objectives:

1. Chemical syntheses of small polyketide molecules that have been found to act as catalytic intermediates (**Chapter 3**) during the biosynthesis of lichen polyketides. Examples are given in **Section 1.6.1** and their biosynthesis is summarized in **Figure 1.13**. Chemical synthesis of analogs of these catalytic intermediates (**Chapter 4**) and exploration of the bioactivities of all these molecules by performing biofilm disruption and anti-microbial assays.
2. To develop a reliable protocol for the heterologous expression of lichen biosynthetic genes within a fast-growing host.

3. Functional characterization of lichen biosynthetic post-PKS genes to catalyze the conversion of chemically synthesized polyketide intermediates to final natural products (more complex polyketide molecules, **Schemes 2, 3 and, 4**).
4. To generate putative assignments of the function of the annotated *C. uncialis* gene clusters and propose whole biosynthetic pathways catalyzed by these gene clusters based on 'homology mapping' and phylogenetics.



## References

- Abdel-Hameed, M., Bertrand, R. L., Piercey-Normore, M. D., Sorensen, J. L. Putative identification of the usnic acid biosynthetic gene cluster by de novo whole-genome sequencing of a lichen-forming fungus. *Fungal Biol.* **2016**, *120*, 306-16.
- Ahmadjian, V. and Jacobs, J. B. Relationship between fungus and alga in the lichen *Cladonia cristatella* tuck. *Nature.* **1981**, *289*, 169–172.
- Aigle, B. and Corre, C. Waking up *Streptomyces* secondary metabolism by constitutive expression of activators or genetic disruption of repressors. *Methods Enzymol.* **2012**, *517*: 343-366.
- Aitor Hernández-et. al. *Journal of the American Chemical Society* **2015** *137* (23), 7474-7487.
- Alberti, F., Foster, G. D., Bailey, A. M., Natural products from filamentous fungi and production by heterologous expression. *Appl. Microbiol. Biotechnol.* **2017**, *101*: 493-500.
- Alhawatema, M. S., Gebril, S., Cook, D., Creamer, R. RNAi-mediated down-regulation of a melanin polyketide synthase (pks1) gene in the fungus *Slafractonia leguminicola*. *World J. Microbiol. Biotechnol.* **2017**, *33*: 179.
- Anyago, D. C., Mortensen, U. H. Heterologous production of fungal secondary metabolites in *Aspergilli*. *Front. Microbiol.* **2015**, *6*: 77.
- Armaleo, D., Sun, X., & Culberson, C. Insights from the first putative biosynthetic gene cluster for a lichen depside and depsidone. *Mycologia*, **2011**, *103*(4), 741–754.
- Armaleo, D., Sun, X., Culberson, C. Insights from the first putative biosynthetic gene cluster for a lichen depside and depsidone. *Mycologia.* **2011** Jul-Aug;103 (4): 741-54.
- Armstrong, R. Lichens, lichenometry and global warming. *Microbiologist.* **2004**: 32-35.
- Balaji, P., Bharath, P., Satyan, R. S., Hariharan, G. N. In vitro antimicrobial activity of *Rocella montagnei* thallus extracts. *J Trop Med Plants.* **2006**, *7*:169–173.

- Baltz, R. H. *Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters. *J. Ind. Microbiol. Biotechnol.* **2010**, 37: 759-772.
- Bayly, C. L., Yadav, V. G. Towards precision engineering of canonical polyketide synthase domains: Recent advances and future prospects. *Molecules.* **2017**, 22: e235.
- Beck, J., Ripka, S., Signer, A., Schiltz, E., & Schweizer, E. The multifunctional 6-methylsalicylic acid synthase gene of *Penicillium patulum*. Its gene structure relative to that of other polyketide synthases. *Eur. J. Biochem.* **1990**, 192, 487–498.
- Benedict, J. B. A review of lichenometric dating and its applications to archaeology. *Am. Antiquity.* **2009**, 74: 143-172.
- Bertrand, R. L., Abdel-Hameed, M., and Sorensen, J. L. Lichen Biosynthetic Gene Clusters Part I: Homology Mapping Suggests a Functional Diversity. *Journal of Natural Products* **2018**, 81 (4), 723-731.
- Bertrand, R. L., Abdel-Hameed, M., and Sorensen, J. L. Lichen Biosynthetic Gene Clusters Part II: Homology Mapping Suggests a Functional Diversity. *Journal of Natural Products*, **2018**, 81 (4), 732-748.
- Bézivin, C., Tomasi, S., Lohézic-Le Dévéhat, F., and Boustie, J. Cytotoxic activity of some lichen extracts on murine and human cancer cell lines. *Phytomedicine.* **2003**, 10, 499–503.
- Bézivin, C., Tomasi, S., Rouaud, I., Delcros, J., and Boustie J. Cytotoxic activity of compounds from the lichen: *Cladonia convoluta*. *Planta Med.* **2004**, 70, 874–877.
- Billingsley, J. M., DeNichola, A. B., Tang, Y. Technology development for natural product biosynthesis in *Saccharomyces cerevisiae*. *Curr. Opin. Biotechnol.* **2016**, 42, 74-83.

- Birch, A. J., Massy-Westrop R.A., Moye, C. J. Studies in relation to biosynthesis. VII. 2-Hydroxy-6-methylbenzoic acid in *Penicillium griseofulvum* Dierckx. *Aust. J. Chem.* **1955**, 8, 539–544.12.
- Birch, A. J., Massy-Westropp, R. A. & Moye, C. J. Biosynthesis (VII) 2-hydroxy-6-methyl benzoic acid in *Penicillium griseofulvum*. *Aust. J. Chem.* **1955**, 8, 539–544.
- Blin, K., Wolf, T., Chevrette, M. G., Lu, X., Schwalen, C. J., Kautsar, S. A., Suarez-Duran, H. G., De los Santos, E. L., Kim, H. U., Nave, M., Dickschat, J. S., Mitchell, D. A., Shelest, E., Breitling, R., Takano, E., Lee, S. Y., Weber, T., Medema, M. H. AntiSMASH 4.0 – Improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **2017**, 45: W36-W41.
- Bode, H. B., Bethe B, Höfs R, Zeeck A. Big effects from small changes: possible ways to explore nature's chemical diversity. *ChemBioChem* **2002**, 3: 619–627.
- Borchardt, J. K. The Beginnings of Drug Therapy: Ancient Mesopotamian Medicine. *Drug News Perspect.* **2002** Apr;15(3):187-192
- Brodo, I. M., Sharnoff, S. D., and Sharnoff, S. *Lichens of North America*. New Haven; London: Yale University Press, **2001**.
- Bucar, F., Schneider, I., Ogmundsdottir, H., Ingolfssdottir K. Antiproliferative lichen compounds with inhibitory activity on 12(S)-HETE production in human platelets. *Phytomedicine.* **2004**, 11, 602-6.
- Bugg, T.D. Dioxygenase enzymes: catalytic mechanisms and chemical models. *Tetrahedron*, 59, **2003**, 7075-7101.
- Burkholder, P. R., Evans, A. W., McVeigh, I., Thornton, H. K. Antibiotic activity of lichens. *Proc Natl Acad Sci USA*, **1944**, 30, 250–255.

- Burlando, B., Ranzato, E., Volante, A., Appendino, G., Pollastro, F., Verotta, L. Anti-proliferative effects on tumour cells and promotion of keratinocyte wound healing by different lichen compounds. *Planta Med.* **2009**, 75, 607-13.
- Cacho, R. A., Tang, Y., Chooi, Y. H. Next-generation sequencing approach for connecting secondary metabolites to biosynthetic gene clusters in fungi. *Front. Microbiol.* **2015**, 5, 774.
- Calchera, A., Dal Grande, F., Bode, H.B., Schmitt, I. Biosynthetic Gene Content of the ‘Perfume Lichens’ *Evernia prunastri* and *Pseudevernia furfuracea*. *Molecules* **2019**, 24, 203
- Calcott, M. J., Ackerley, D. F., Knight, A., Keyzers, R. A., Owen, J. G. Secondary metabolism in the lichen symbiosis. *Chem Soc Rev.* **2018**, 47, 1730-1760.
- Campbell, C. D., Vederas, J. C. Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers.* **2010**, 93 (9), 755-63.
- Chain, E., Florey, H. W., Gardner, A. D., Heatley, N. G., Jennings, M. A., Orr-Ewing, J., Sanders, A. G. The classic: penicillin as a chemotherapeutic agent. 1940. *Clin. Orthop. Relat. Res.* **2005**, 439, 23–2610.
- Chin, Young-Won et al. “Drug discovery from natural sources.” *The AAPS journal* **2006**, 8 (2), E239-53. doi:10.1007/BF02854894.
- Chooi, Y. H., Stalker D. M., Davis M. A., Fujii, I., Elix J. A., Louwhoff, S. H., et al., Cloning and sequence characterization of a non-reducing polyketide synthase gene from the lichen *Xanthoparmelia semiviridis*. *Mycological Research*, **2008**, 112:147–161.
- Chooi, Y. H.; Tang, Y. *J. Org. Chem.* **2012**, 77, 9933.
- Christianson, D. W. Structural Biology and Chemistry of the Terpenoid Cyclases. *Chemical Reviews* **2006**, 106 (8), 3412-3442

- Christianson, D. W. Unearthing the roots of the terpenome. *Curr Opin Chem Biol.* **2008**, 12(2), 141-50.
- Cobb, R. E., Luo, Y., Freestone, T., Zhao, H. Chapter 10- Drug Discovery and Development via Synthetic Biology, **2013**, pp. 183-206.
- Cox, P. A., Banack, S. A., Murch, S. J., Rasmussen, U. Tien, G., Bidigare, R. R., Metcalf, J. S., Morrison, L. F., Codd, G. A., and Bergman, B. Diverse taxa of cyanobacteria produce-N-methylamino-L-alanine, a neurotoxic amino acid. *PNAS*, **2005**, Vol. 102, No. 14, 5074-5078.
- Crawford, J. M., Thomas, P. M., Scheerer, J. R., Vagstad, A. L., Kelleher, N. L., Townsend, C. A. *Science* **2008**, 320, 243.
- Culberson, C. F. and Armaleo, D. Induction of a complete secondary-product pathway in a cultured lichen fungus. *Exp. Mycol.* **1992**, 16, 52–63.
- Mansuy, D. The great diversity of reactions catalyzed by cytochromes P450, *Comp. Biochem. Physiol., Part C: Pharmacol., Toxicol. Endocrinol.* **1998**, 121, 5–14.
- Dayan, F. E. and Romagni, J. G. Lichens as a potential source of pesticides. *Pestic. Outlook*, **2001**, 12, 229–232.
- De Carvalho, E. A. B., Andrade, P. P., Silva, N. H., Pereira, E. C., Figueiredo, R. C. B. Q. Effect of usnic acid from the lichen *Cladonia substellata* on *Trypanosoma cruzi* in vitro: an ultrastructural study. *Micron* **2005**, 36, 155–161.
- De Voss, J. J., Cryle, M. J. Is the ferric hydroperoxy species responsible for sulfur oxidation in cytochrome P450s? *Angew. Chem. Int. Ed. Engl.* **2006**, 45, 8221-8223.

- Derntl, C., Kluger, B., Bueschl, C., Schuhmacher, R., Mach, R. L., Mach-Aigner, A. R.,  
Transcription factor Xpp1 is a switch between primary and secondary fungal metabolism.  
*Proceedings of the National Academy of Sciences*, **2017**, 114, E560–E569.
- Dewick, P. M. Medicinal Natural Products: A Biosynthetic Approach. *John Wiley & Sons*, **2009**,  
3rd Edition.
- Elix, J. A. and Stocker-Wörgötter, E. Biochemistry and secondary metabolites. In: *Lichen Biology*,  
*2nd ed. (Nash T. H. III, ed.)*. Cambridge University Press, Cambridge, **2008**, pp. 104 –  
133.
- Guengerich, F. P. and Yoshimoto, F. K. *Chemical Reviews* **2018**, 118 (14), 6573-6655.
- Brightman, F. H., Seaward, M.R.D. Lichens of man-made substrates. *M. R.*  
*D. Seaward (Ed.), Lichen Ecology*, Academic Press, London, **1977**, pp. 253-293.
- Fahselt, D. Secondary biochemistry of lichens. *Symbiosis*. **1994**, 16, 117–165.
- Fazio, A., Bertoni, M., Adler, M., Ruiz, L., Rosso, M., Muggia, L., Hager, A., Stocker, E. &  
Maier, M. Culture studies on the mycobiont isolated from Parmotrema reticulatum Choisy:  
Metabolite production under different conditions. *Mycological Progress*. **2009**, 8, 359-365.
- Fernández-Salegui, A. B., Terrón, A., Barreno, E., and Nimis, P. L. Biomonitoring with  
cryptogams near the power station of La Robla. *Bryologist*. **2007**, 110, 723 – 737.
- Fleming, A. The discovery of Penicillin. *British Medical Bulletin*, **1944**, 2(1), 4-5.
- Fox, E. M., Howlett, B. J. Secondary metabolism: regulation and role in fungal biology. *Curr Opin*  
*Microbiol*. **2008**, 11(6), 481-7.
- Dayan, F. E. and Romagni, J. G. *Pestic. Outlook*, **2001**, 12, 229-232.
- Fuller, K. K., Chen, S., Loros, J. J., Dunlap, J. C. Development of the CRISPR/Cas9 system for  
targeted gene disruption in *Aspergillus fumigatus*. *Eukaryot. Cell*. **2015**, 14, 1073-1080.

- Glavich, D. A. and Geiser, L. H. Potential approaches to developing lichen-based critical loads and levels for nitrogen, sulphur and metal-containing atmospheric pollutants in North America. *Bryologist*. **2008**, 111, 638 – 649.
- Goerig, M., Schulte am Esch, J. Friedrich Wilhelm Adam Sertürner—the discoverer of morphine. *Anästhesiol Intensivmed Notfallmed Schmerzther*. **1991**, 26(8), 492-498.
- Michal, G., Elečko, J., Marcinčinová, M., Dajana, R., Bačkorová, M., Martin, B. Lichen Metabolites: An Overview of Some Secondary Metabolites and Their Biological Potential, **2018**.
- Gomez-Escribano, J. P., Bibb, M. J. Heterologous expression of natural product biosynthetic gene clusters in *Streptomyces coelicolor*: From genome mining to manipulation of biosynthetic pathways. *J. Ind. Microbiol. Biotechnol*. **2014**, 41, 425-431.
- Goodwin, S., McPherson, J. D., McCombie, R. Coming of age: Ten years of next-generation sequencing technologies. *Nat. Rev. Genet*. **2016**, 17, 333-351.
- Gries, C. Lichens as indicators of air pollution. In: *Lichen Biology, 1st ed. (Nash T. H. III, ed.)*. Cambridge University Press, Cambridge, **1996**, 240 –254.
- Guengerich, F. P. Common and Uncommon Cytochrome P450 Reactions Related to Metabolism and Chemical Toxicity. *Chemical Research in Toxicology* **2001**, 14 (6), 611-650.
- Guengerich, F. P. Mechanisms of Cytochrome P450-Catalyzed Oxidations. *ACS catalysis*, **2018**, 8(12), 10964–10976.
- Guengerich, F. P., Waterman, M. R., & Egli, M. Recent Structural Insights into Cytochrome P450 Function. *Trends in pharmacological sciences*, **2016**, 37(8), 625–640.

- Zang, H., Wang, Y., Wu, J., Shakila, K., Pfeifer, B. A. Complete biosynthesis of erythromycin A and designed analogs using *E. coli* as a heterologous host. *Chem. Biol.*, **2010**, 17, 1232-1240.
- Hamada, N. The effect of various culture conditions on depside production by an isolated lichen mycobiont. *Bryologist*. **1989**, 92, 310–313.
- Hawksworth, D. L. The fungal dimension of biodiversity: magnitude, significance, and conservation. *Mycol. Res.* **1991**, 95, 641–655
- Honegger R. The symbiotic phenotype of lichen-forming Ascomycetes. In: *The Mycota IX (Hock B., ed.)*. Springer-Verlag, Berlin, Heidelberg, **2001**, 165–188.
- Hollstein, U. Actinomycin. Chemistry and mechanism of action. *Chem. Rev.* **1974**, 74 (6), 625-652.
- Honegger, R. Simon Schwendener (1829–1919) and the Dual Hypothesis of Lichens. *The Bryologist*. **2009**, 103, 307-313.
- Huang, A. C., Kautsar, S. A., Hong, Y. J., Medema, M. H., Bond, A. D., Tantillo, D. J., Osbourn, A. Unearthing a sesterterpene biosynthetic repertoire in the *Brassicaceae* through genome mining reveals convergent evolution. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, 114, E6005-E6014.
- Ichikawa, Y., Yamano, T., Murakami, K., Mason, H. S. *J. Biol. Chem.* **1967**, 242,1102–1110.
- Ikram, N. K., Zhan, X., Pan, X. W., King, B. C., Simonsen, H. T. Stable heterologous expression of biologically active terpenoids in green plant cells. *Front. Plant Sci.* **2015**, 6, 129.
- Fujii, I., Yoshida, N., Shimomaki, S., Oikawa, H., Ebizuka, Y. An Iterative Type I Polyketide Synthase PKSN Catalyzes Synthesis of the Decaketide Alternapyrone with Regio-Specific Octa-Methylation. *Chemistry & Biology*, **2005**, 12(12),1301-1309.



- Isin, E. M., Guengerich, F. P. Complex reactions catalyzed by cytochrome P450 enzymes. *Biochim Biophys Acta*. **2007**, 1770(3), 314-29.
- Butterworth, J. H. & Morgan, E. D. Isolation of a substance that suppresses feeding in locusts, *Chem. Commun. (London)*, **1968**, 23-24.
- Collie, J. N. and Myers, W. S. VII. — The formation of orcinol and other condensation products from dehydracetic acid. *J. Chem. Soc.*, **1893**, 63, 122.
- Collie, J. N. CLXXI. — Derivatives of the multiple keten group *J. Chem. Soc.*, **1907**, 91, 1806.
- Staunton, J. and Weissman, K. J. Polyketide Biosynthesis: A millennium review. *Nat. Prod. Rep.*, **2001**, 18, 380–416.
- Kealey, J. T., Craig, J. P., Barr, P. J. Identification of a lichen depside polyketide synthase gene by heterologous expression in *Saccharomyces cerevisiae*. *Metabolic Engineering Communications*, **2021**, 13.
- Jin, S., Markris, T. M., Bryson, T. A., Sligar, S. G., Dawson, J. H. Epoxidation of olefins by hydroperoxo-ferric cytochrome P450. *J. Am. Chem. Soc.* **2003**, 125, 3406-3407.
- Ruckpaul, K., Rein, H., Blanck, J. Regulation mechanism of the activity of the hepatic endoplasmic cytochrome P-450, in: K. Ruckpaul, H. Rein (Eds.), *Basis and Mechanism of Regulation of Cytochrome P450*, Akademie-Verlag, Berlin, **1989**, pp. 1–55.
- Degtyarenko, K. N. Archakov, A. I. Molecular evolution of P450 superfamily and P450-containing monooxygenase systems, *FEBS Lett.* **1993**, 332, 1–8.
- Kahriman, N., Yazici, K., Arslan, T., Aslan, A., Karaoğlu, S. A., Yayli, N. Chemical Composition and Antimicrobial Activity of the Essential Oils from *Evernia prunastri* (L.) Ach. and *Evernia divaricata* (L.) Ach. *Asian Journal of Chemistry*, **2011**, 23, 1937-1939.

- Karunaratne, V., Bombuwela, K., Kathirgamanathar, S., and Thadani, V. M., Lichens: A Chemical Important Biota. *Journal of the National Science Foundation of Sri Lanka*, **2005**, 33 (3), pp. 169-186.
- Kasahara, K., Miyamoto, T., Fujimoto, T., Oguri, H., Tokiwano, T., Oikawa, H., Ebizuka, Y., Fujii, I. Solanapyrone synthase, a possible Diels-Alderase and iterative type I polyketide synthase encoded in a biosynthetic gene cluster from *Alternaria solani*. *Chembiochem*. **2010**, 11(9), 1245-52.
- Khalidi, N., Seifuddin, F. T., Turner, G., Haft, D., Nierman, W. C., Wolfe, K. H., Fedorova, N. D. SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genet. Biol.* **2010**, 47, 736-741.
- Kim, W., Liu, R., Woo, S., Kang, K. B., Park, H., Yu, Y. H., Ha, H. H., Oh, S. Y., Yang, J. H., Kim, H., Yun, S. H., Hur, J. S. Linking a Gene Cluster to Atranorin, a Major Cortical Substance of Lichens, through Genetic Dereplication and Heterologous Expression. *mBio*. **2021**, 12(3).
- Kirk P. M., Cannon P. F., Minter D. W., and Stalpers J. A. (eds.) Dictionary of the Fungi, 10th ed. CAB International, Wallingford, Oxon, UK, **2008**.
- Kon, Y., Iwashina, T., Kashiwadani, H., Wardlaw, J. D., and Elix, J. A. A new dibenzofuran, isostrepsilic acid, produced by cultured mycobiont of the lichenized ascomycete *Usnea orientalis*. *Journal of Japanese Botany* **1997**, 72, 67-71.
- Korman, T. P., Crawford, J. M., Labonte, J. W., Newman, A. G., Wong, J., Townsend, C. A., Tsai, S. C. Structure and function of an iterative polyketide synthase thioesterase domain catalyzing Claisen cyclization in aflatoxin biosynthesis. *Proc. Natl. Acad. Sci. USA*. **2010**, 107(14), 6246-51.

- Kosanić, M. & Ranković, B. Studies on Antioxidant Properties of Lichen Secondary Metabolites. In: Ranković B. Lichen Secondary Metabolites Bioactive Properties and Pharmaceutical Potential. *Springer International Publishing*. **2015**. p. 105-125.
- Kraus, W. Biologically active ingredients in: Schmitterer H editor. The neem tree *Azadirachta indica* A. Juss and other meliaceous plants. Sources of unique natural products for pest management, medicine, industry, and other purposes. Weinheim, Germany: *VCH Verlagsgesellschaft mbH*. **1995**, Vol. 3, pp. 35-88.
- Kumar K. C. S., Muller, K. Depsides as non-redox inhibitors of leukotriene B4 biosynthesis and HaCaT cell growth. 1. Novel analogs of barbatic and diffractaic acid. *Eur J Med Chem*. **1991c**, 34: 1035–1042.
- Kumar, K. C. S., Muller, K. Lichen metabolites. 1. Inhibitory action against leukotriene B4 biosynthesis by a nonredox mechanism. *J Nat Pro*. **1999a**, 62: 817–820.
- Kumar, K. C. S., Muller, K. Lichen metabolites. 2. Antiproliferative and cytotoxic activity of gyrophoric, usnic, and diffractaic acid on human keratinocyte growth. *J Nat Pro*. **1999b**, 62: 821–823.
- Li, J., Neubauer, P. *Escherichia coli* as a cell factory for heterologous production of nonribosomal peptides and polyketides. *New Biotechnol*. **2014**, 31: 579-585.
- Liangcheng Du and Lili Lou. PKS and NRPS release mechanism. *Nat. Prod. Rep*. **2010**, 27, 255.
- Loeschcke A, Thies S. *Pseudomonas putida*: A versatile host for the production of natural products. *Appl. Microbiol. Biotechnol*. **2015**, 99: 6197-6214.
- Lynen, F.; Tada, M. Die. biochemischen grundlagen der polyacetat-regel. *Angew. Chem*. **1961**, 73, 513–519.

- Ma S. M., Li, J. W., Choi, J. W., Zhou, H., Lee, K. K., Moorthie, V. A., Xie, X., Kealey, J. T., Da Silva, N. A., Vederas, J. C., Tang, Y. Complete reconstitution of a highly reducing iterative polyketide synthase. *Science*. **2009**, 326 (5952): 589-92.
- Mayer, M., O'Neill, M. A., Murry, K. E., Santos-Magal-hães, N. S., Carneiro-Leão, A. M. A., Thompson, A. M., and Appleyard, V. C. L. Usnic acid: a non-genotoxic compound with anti-cancer properties. *Anti-Cancer Drugs*, **2005**, 16, 805–809.
- McGuire, J. M.; Bunch R. L., Anderson R. C., Boaz H. E., Flynn E. H., Powell H. M. & Smith J. W. "Ilotycin," a new antibiotic. *Antibiot. Chemother.* **1952**, 2: 281 -283.
- McIntosh, J. A., Farwell, C. C., & Arnold, F. H. Expanding P450 catalytic reaction space through evolution and engineering. *Current opinion in chemical biology*, **2014**, 19, 126–134.
- Mitrović, T., Stamenković, S., Cvetković, V., Nikolić, M., Tošić, S., Stojičić, D.: Lichens as source of versatile bioactive compounds, *Biologica Nyssana*, **2011**, 2 (1), 1-6.
- Miziorko, H. M. Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch Biochem Biophys.* **2011**, 505(2) 131-43.
- Molina, M. C., Crespo, A., Vicente, C., Elix, J. A. Differences in the composition of phenolics and fatty acids of cultured mycobiont and thallus of *Physconia distorta*, *Plant Physiology and Biochemistry*, **2003**, Volume 41, Issue 2.
- Molnár, K., Farkas, E. Current Results on Biological Activities of Lichen Secondary Metabolites: A Review. *Zeitschrift für Naturforschung*, **2010**, 65, 157–173.
- Moriguchi, T., Kezuka, Y., Nonaka, T., Ebizuka, Y., and Fujii, I. Hidden Function of Catalytic Domain in 6-Methylsalicylic Acid Synthase for Product Release. *The Journal of Biological Chemistry*, Vol. 285, No. 20, 15637–15643.

- Mosbach, K. Biosynthesis of lichen substances, products of a symbiotic association. *Angewandte Chemie*, International Edition 8, **1969**, 240-250.
- Müller, K. Pharmaceutically relevant metabolites from lichens. *Appl Microbiol Biotechnol.* **2001** Jul;56(1-2):9-16.
- Nakanishi, T., Murata, H., Inatomi, Y., Inada, A., Murata, J., Lang, F. A., Yamasaki, K., Nakano, M., Kawahata, T., Mori, H., Otake, T. Screening of anti-HIV-1 activity of North American plants. Anti-HIV-1 activities of plant extracts, and active components of *Letharia vulpina* (L.) *Hue. Nat Med.* **1998**, 52521–526.
- NC-IUB, Nomenclature Committee of the International Union of Biochemistry (NC-IUB). Nomenclature of electron-transfer proteins Recommendations 1989, *Eur. J. Biochem.* **1991**, 200, 599–611.
- Neamati, N., Hong, H., Mazumder, A., Wang, S., Sunder, S., Nicklaus, M. C., Milne, G. W., Proksa, B., Pommier, Y. Depsides and depsidones as inhibitors of HIV-1 integrase: discovery of novel inhibitors through 3D database searching. *J. Med. Chem.* **1997**, 40, 942–951.
- Nelson, D. R. A world of cytochrome P450s. *Phil. Trans. R. Soc. B. Biol. Sci.* **2013**, 368.
- Nicolaou, K. C., Yang Z., Liu, J. J., Ueno, H., Nantermet, P. G., Guy, R. K., Claiborne, C. F., Renaud, J., Couladouros, E. A., Paulvannan, K., Sorensen, E. J. Total synthesis of taxol. *Nature* **1994**, 367, 630-634.
- Nielsen, J. C., Grijseels, S., Prigent, S., Ji, B., Dainat, J., Nielsen, K. F., Frisvad, J. C., Workman, M., Nielsen, J. Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium species*. *Nat. Microbiol.* **2017**, 2, 17044.

- O'Neill, M. A., Mayer, M., Murray, K. E., Rolim-Santos, H. M., Santos-Magalhães, N. S., Thompson, A. M., Appleyard, V. C. Does usnic acid affect microtubules in human cancer cells? *Braz J Biol.* **2010**, 70(3), 659-64.
- Ortiz de Montellano, P. R. Substrate Oxidation. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 4th ed; *Ortiz de Montellano PR*, Ed.; Springer: New York, **2015**, pp 111–176.
- Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M. D., Tai, A., Main, A., Eng, D., Polichuk, D. R., Teoh, K. H., Reed, D. W., Treynor, T., Lenihan, J., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K. W., Fickes, S., Galazzo, J., Gaucher, S. P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Jiang, H., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secret, S., Tsuruta, H., Vazquez, R., Westblade, L. F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievens, J., Covello, P. S., Keasling, J. D., Reiling, K. K., Renninger, N. S., Newman, J. D. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **2013**, 496, 528-532.
- Papazi, A., Kastanaki, E., Pirintsos, S., and Kotzabasis, K. Lichen symbiosis: nature's high yielding machines for induced hydrogen production. *PLoS ONE* **2015**, 10: e0121325. doi: 10.1371/journal.pone.0121325
- Parascandolo, J. S., Havemann, J., Potter, H. K. et al. Insights into 6-Methylsalicylic Acid Bio-assembly by Using Chemical Probes. *Angew Chem Weinheim Bergstr Ger.* **2016**, 128(10), 3524-3528.
- Park, S., Tripathi, A., Wu, J. *et al.* Discovery of cahuitamycins as biofilm inhibitors derived from a convergent biosynthetic pathway. *Nat. Commun.* **2016**, 7, 10710.

- Paudel, B., Bhattarai, H. D., Lee, H. K., Oh, H., Shin, H. W., Yim, J. H. Antibacterial activities of ramalin, usnic acid and its three derivatives isolated from the antarctic lichen ramalina terebrata. *Z. Naturforsch C*. **2010**, 65, 34–38.
- Pettit, R. K. Mixed fermentation for natural product drug discovery. *Appl. Microbiol. Biotechnol.* **2009**, 83, 19–25.
- Puel, O., Galtier, P., Oswald, I. P. Biosynthesis and toxicological effects of patulin. *Toxins (Basel)*. **2010**, 2(4), 613-631.
- Quin, M. B., Flynn, C. M. and Schmidt-Dannert, C. Traversing the fungal terpenome. *Nat. Prod. Rep.*, **2014**, 31, 1449-1473.
- Rajab, M. S., Cantrell, C. L., Franzblau, S. G. and Fischer, N. H. *Planta Med.*, **1998**, 64, 2.
- Ren, M. R., Hur, J. S., Kim, J. Y., Park, K. W., Park, S. C., Seong, C. N., Jeong, I. Y., Byun, M. W., Lee, M. K., Seo, K. I. Anti-proliferative effects of *Lethariella zahlbruckneri* extracts in human HT-29 human colon cancer cells. *Food Chem. Toxicol.* **2009**, 47, 2157–2162
- Romagni J. G. and Dayan F. E. Structural diversity of lichen metabolites and their potential use. In: *Advances in Microbial Toxin Research and its Biotechnological Exploitation* (Upadhyay R. K., ed.). *Kluwer Academic/Plenum Publishers*, New York, **2002**, pp. 151–169.
- Silverman, R. B. *The Organic Chemistry of Enzyme-Catalyzed Reactions*; Elsevier Science: Texas, **2002**.
- Saha, S., Zhang, W., Zhang, G., Zhu, Y., Chen, Y., Liu, W., Yuan, C., Zhang, Q., Zhang, H., Zhang, L., Zhang, W., Zhang, C. Activation and characterization of a cryptic gene cluster reveals a cyclization cascade for polycyclic tetramate macrolactams. *Chem. Sci.* **2017**, 8, 1607-1612.

- Schmeda-Hirschmann, G., Tapia, A., Lima, B., Pertino, M., Sortino, M., Zacchino, S., de Arias, A. R., Feresin, G. E. A new antifungal and antiprotozoal depside from the andean lichen *Protousnea poeppigii*. *Phytother. Res.* **2008**, 22, 349–355.
- Schmitt, I., Kautz, S., Lumbsch, H. Thorsten. 6-MSAS-like polyketide synthase genes occur in lichenized ascomycetes. *Mycological Research* **2008**, 112(Pt 2), 289-296.
- Schorn, M. A., Alanjary, M. M., Aquinaldo, K., Korobeynikov, A., Podell, S., Patin, N., Linecum, T., Jensen, P. R., Ziemert, N., Moore, B. S. Sequencing rare marine actinomycete genomes reveals a high density of unique natural product biosynthetic gene clusters. *Microbiology* **2016**, 162: 2075-2086.
- Seaward M.R.D. Environmental role of lichens T.H. Nash (Ed.), *Lichen Biology*, Cambridge University Press, Cambridge. **2008**, pp. 274-295.
- Seaward, M. R. D. Major impacts made by lichens in biodeterioration processes. *Int. Biodeterior. Biodegr.* **1997**, 40, 269–273.
- Sheppard P. R., Speakman R. J., Ridenour G., and Witten M. L. Using lichen chemistry to assess airborne tungsten and cobalt in Fallon, Nevada. *Environ. Monit. Assess.* **2007**, 130, 511–518.
- Shrestha, Gajendra and Larry L. St. Clair. “Lichens: a promising source of antibiotic and anticancer drugs.” *Phytochemistry Reviews* **2013**, 12, 229-244
- Shukla V., Joshi, G. P., Rawat, M. S. M. Lichens as a potential natural source of bioactive compounds: a review. *Phytochem Rev.* **2010**, 9: 303–314
- Shukla, V., Negi, S., Rawat, M. S. M., Pant, G., Nagatsu, A. Chemical Study of *Ramalina africana* (Ramaliniaceae) from Garhwal Himalayas. *Biochem Systemat Ecol.* **2004**, 32, 449–453.



- Shukla, V., Joshi, G.P. & Rawat, M.S.M. Lichens as a potential natural source of bioactive compounds: a review. *Phytochem Rev.* **2010**, 9, 303–314.
- Siddiqui, Anees & Iram, Farah & Siddiqui, Seemi & Sahu, Kapendra. Role of Natural Products in Drug Discovery Process. *International Journal of Drug Development and Research.* **2014**, 6. 172-204.
- Skininder, M. A., Merwin, N. J., Johnston, C. W., Magarvey, N. A. PRISM 3: Expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res.* **2017**, 45, W49-W54.
- Sneider, W. Drug Prototypes and Their Exploitation. *J. Am. Chem. Soc.* **1997**, 119, 6, 1500
- Stadler, R., and Zenk, M. H. *J. Biol. Chem.* **1993**, 268, 823–831.
- Stanley, V. C., English, M. P. Some effects of nystatin on the growth of four *Aspergillus* species. *J Gen Microbiol.* **1965**, 40(1),107-18.
- Stocker-Wörgötter E. and Elix J. A. Secondary chemistry of cultured mycobionts: formation of a complete chemosyndrome by the lichen fungus of *Lobaria spathulata*. *Lichenologist*, **2002**, 34, 351–359.
- Sun, H., Liu, Z., Zhao, H., Ang, E. L. Recent advances in combinatorial biosynthesis for drug discovery. *Drug. Des. Devel. Ther.* **2015**, 9, 823-833.
- Suroto, D. A., Kitani, S., Miyamoto, K. T., Sakihama, Y., Arai, M., Ikeda, H., Nihira, T. Activation of cryptic phthoxazolin A production in *Streptomyces avermitilis* by the disruption of autoregulator-receptor homologue AvaR3. *J. Biosci. Bioeng.* **2017**, 124: 611-617.
- Nash, T. H., Frontmatter: Lichen Biology Edition III. Cambridge University Press, **2008**.
- T. Omura, R. Sato *J. Biol. Chem.* **1964**, 239, 2379–2385.

- Tanaka, Y., Hosaka, T., Ochi, K. Rare earth elements activate the secondary metabolite-biosynthetic gene clusters in *Streptomyces coelicolor* A3(2). *J. Antibiot.* **2010**, 63, 477-481.
- Tannous, J.; El Khoury, R.; Snini, S.P.; Lippi, Y.; El Khoury, A.; Atoui, A.; Lteif, R.; Oswald, I.P.; Puel, O. Sequencing, physical organization and kinetic expression of the patulin biosynthetic gene cluster from *Penicillium expansum*. *Int. J. Food Microbiol.* **2014**, 189.
- Tsai, S. C., Ames, B. D. Structural enzymology of polyketide synthases. *Methods Enzymol.* **2009**, 459, 17-47.
- Turk, A. O., Yilmaz, M., Kivanc, M., Turk, H. The antimicrobial activity of extracts of the lichen *Cetraria aculeata* and its protolichesterinic acid constituent. *Z Naturforsch C*, **2003**, 58:850–854.
- Udwaray, D. W. & Townsend, C. A. *J. Mol. Biol.* **2002**, 323, 585–598
- Urlacher, V. B. and Girhard, M. Cytochrome P450 monooxygenases in biotechnology and synthetic biology. *Trends Biotechnol.* **2019**, 37, 882–897.
- Vagstad, A. L., Hill, E. A., Labonte, J. W., Townsend, C. A. Characterization of a fungal thioesterase having Claisen cyclase and deacetylase activities in melanin biosynthesis. *Chem Biol.* **2012**, 19(12), 1525-34.
- Vézina, C., Kudelski, A., Sehgal, S. N. Rapamycin (AY-22,989), a new antifungal antibiotic. I. Taxonomy of the producing streptomycete and isolation of the active principle. *J Antibiot (Tokyo)*. **1975**, 28(10), 721-6.
- Vo, V. G., Le, H. D., Tran, T. N., Nguyen, N. H., Vo, T. G., Sichaem, J., Nguyen, V. K., Duong, T. H. A new eremophilane-sesquiterpene from the cultured lichen mycobiont of *Graphis* sp. *Nat Prod Res.* **2022**, 36(1), 319-325.

- Waksman, S. A., and Woodruff, H. B. The soil as a source of microorganisms antagonistic to diseaseproducing bacteria. *J. Bacteriol.* **1940**, 40, 600-605.
- Waksman, S. A., and Woodruff, H. B. Actinomyces antibioticus, a new soil organism antagonistic to pathogenic and non-pathogenic bacteria. *J. Bacteriol.* **1941**, 42, 231-249.
- Wang, M., Zhou, H., Wirz, M., Tang, Y., and Boddy, C. N. A Thioesterase from an Iterative Fungal Polyketide Synthase Shows Macrocyclization and Cross Coupling Activity and May Play a Role in Controlling Iterative Cycling through Product Offloading. *Biochemistry*, **2009**, 48 (27), 6288-6290.
- Wang, Y., Geng, C., Yuan, X., Hua, M., Tian, F., Li, C. Identification of a putative polyketide synthase gene involved in usnic acid biosynthesis in the lichen *Nephromopsis pallescens*. *PLoS ONE* **2018**, 13(7).
- Winn, M., Fyans, J. K., Zhuo, Y., Micklefield, J. Recent advances in engineering nonribosomal peptide assembly lines. *Nat. Prod. Rep.* **2016**, 33: 317-347.
- Yaegashi J, Oakley BR, Wang CCC. Recent advances in genome mining of secondary metabolite biosynthetic gene clusters and the development of heterologous expression systems in *Aspergillus nidulans*. *J. Ind. Microbiol. Biotechnol.* **2014**, 41: 433-442.
- Yang, Z. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Molecular Biology and Evolution*, **1993**, 10: 1396–1401.
- Yuan, X., Xiao, S., & Taylor, T. N. Lichen-like symbiosis 600 million years ago, *Science* **2005**, 308, 1017-1020.
- Xu, Y., Zhou, T., Zhang, S., Xuan, L. J., Zhan, J., and Molnár, I. Thioesterase Domains of Fungal Nonreducing Polyketide Synthases Act as Decision Gates during Combinatorial Biosynthesis. *J. Am. Chem. Soc.* **2013** 135 (29), 10783-10791.

- Zhang, M. M., Wang, Y., Ang, E. L., Zhao, H. Engineering microbial hosts for production of bacterial natural products. *Nat. Prod. Rep.* **2015**, 33: 963-987.
- Zhang, M. M., Wong, F. T., Wang, Y., Luo, S., Lim, Y. H., Heng, E., Yeo, W. L., Cobb, R. E., Enghiad, B., Ang, E. L., Zhao, H. CRISPR-Cas9 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat. Chem. Biol.* **2017**, 13: 607-607.
- Zhao, B., Guengerich, F. P., Bellamine, A., Lamb, D. C., Izumikawa, M., Lei, L.; Podust, L. M., Sundaramoorthy, M., Kalaitzis, J. A., Reddy, L. M., Kelly, S. L., Moore, B. S., Stec, D., Voehler, M., Falck, J. R., Shimada, T., Waterman, M. R. Binding of Two Flaviolin Substrate Molecules, Oxidative Coupling, and Crystal Structure of *Streptomyces coelicolor* A3(2) Cytochrome P450 158A2\*, *J. of Biol. Chemistry.* **2005**, 280 (12),11599-11607.
- Zhao, B., Guengerich, F. P., Bellamine, A., Lamb, D. C., Izumikawa, M., Lei, L., Podust, L. M., Sundaramoorthy, M., Kalaitzis, J. A., Reddy, L. M., Kelly, S. L., Moore, B. S., Stec, D., Voehler, M., Falck, J. R., Shimada, T., and Waterman, M. R. *The J. of Biol. Chemistry.* **2005**, Vol. 280, No. 12, 11599-11607.
- Zhou, H., Zhan, J., Watanabe, K., Xie, X., Tang, Y. A polyketide macrolactone synthase from the filamentous fungus *Gibberella zeae*. *Proc Natl Acad Sci USA.* **2008**,105 (17), 6249-54.
- Ziemert, N., Alanjary, M., Weber, T. The evolution of genome mining in microbes – a review. *Nat. Prod. Rep.* **2018**, 33: 988-1005.
- Ziemert, N., Jensen PR. Phylogenetic approaches to natural product structure prediction. *Methods Enzymol.* **2012**, 517: 161-182.

# Chapter 2

## Materials and methods

### 2.1. Materials and methods related to Chapter 1

*Not applicable.*

### 2.2. Materials and methods related to Chapter 2

For a list of primers used in all chapters, see **Table S1** in the Appendix. For a list of plasmids constructed in all chapters, see **Table S2** in the Appendix. For a list of bacterial strains used in **Chapter 4**, see **Table S3** in the Appendix.

### 2.3. Materials and methods related to Chapter 3

#### 2.3.1. *Collection and taxonomic identification of C. uncialis*

This section is described in detail by Abdel-Hameed [2015]. The internal transcribed spacer (ITS) sequence that was acquired by Abdel-Hameed [2015] matched entries for *C. uncialis* with >95 % sequence identity [Bertrand, R. L. 2019]. Throughout this thesis, BLAST alignment statistics will be reported in terms of “Identity % / Coverage %”. The “identity” is the percentage of nucleotides or amino acids within a queried sequence that are identical to those of the reference gene or protein. The “coverage” is the percentage of overlap between the analyzed portions of the queried gene and the reference gene. These statistics are reported as XX / YY throughout this thesis.

#### 2.3.2. *Sub-culturing of C. uncialis from the algal partner*

The ITS sequence that was acquired by Abdel-Hameed [2015] matched entries for *C. uncialis* with >95 % sequence identity [Bertrand, R. L. 2019].

### 2.3.3. Genome sequencing and assembly

As described by [Bertrand, R. L. 2019], the genome of the fungal partner was *de novo* sequenced by MICB DNA sequencing services (Manitoba Institute of Cell Biology, CancerCare Manitoba, University of Manitoba), using an Illumina MiSeq sequencer with a MiSeq Micro V2 sequencing kit. The genomic DNA sample from the fungal culture was used to generate a paired-end DNA library for 150 bp paired-end sequencing reads. The genome was estimated to be 30 MB in length, an estimate based on studies of genome length of *Cladonia grayi* [Armaleo, D. & May, S. 2009]. The average length of the raw reads was 147 nucleotides with a standard deviation of 16 nucleotides. A total of 515.7 million nucleotides of information was generated. A Phred quality score of 30 or greater was achieved for 441.1 million nucleotides, signifying that 85.54 % of the nucleotides generated were correctly identified as A, T, G, or C with confidence of 99.9 % or greater [Ewing et al., 1998; Ewing & Green, 1998]. Raw sequence reads were deposited in GenBank and are available under accession number SRR4418292. The raw data obtained from Illumina MiSeq sequencing were assembled into contigs using four DNA assembly programs: DNASTAR, Geneious, SPAdes, and Velvet [Kearse et al., 2012; Bankevich et al., 2012; Zerbino & Birney, 2008]. The results generated by SPAdes was chosen for subsequent analysis. This assembly generated a total of 2109 contigs ( $\geq 1$  KB) with a combined total of 32.9 million nucleotides. This result was consistent with our prior estimation of 30 million nucleotides based on the estimated length of the *C. grayi* genome [Armaleo, D. & May, S. 2009]. The GC/AT ratio of the assembled DNA was 46.38 %. The contig N50 (defined as the value at which half of the genomic DNA is contained within contigs of X length or greater) was 34.7 KB. The longest contig was 143.1 KB. The 2109 contigs were deposited in GenBank under accession number NAP000000000.

#### 2.3.4. Annotation of SM gene clusters

The rapid *in silico* identification of putative polyketide synthase (PKS) genes was performed using Antibiotics and Secondary Metabolite Analysis Shell (AntiSMASH Version 2.0) [Blin et al., 2013], freely available for academic use at <http://www.secondarymetabolites.org/>

#### 2.3.5. Synthesis of MPA

The synthesis of MPA was described by [Hawranik, D. J. *et al.* 2009]. In this study the reaction protocol and product purification methodology were modified to increase the yield of MPA. Biotage automated flash chromatography was used to get the final purified product with a solvent gradient of EtOAc/Hexane.

#### 2.3.6. Test Compounds for bioassays

UA was purchased from ChromaDex USA (>98% purity by HPLC) and used as received. MPA was synthesized using a previously reported method [D. J. Hawranik et al., 2009] and the spectroscopic characterization data were identical to the previously reported data.

#### 2.3.7. Bacterial Strains

Reference strains (*S. aureus* MTCC 96, *B. subtilis* MTCC 441, and *Pseudomonas aeruginosa* MTCC 2453) were used for assay. The strains were maintained on Nutrient Agar (NA) plates and stored at 4 °C. A single colony was transferred to Mueller Hinton broth (MHB) and, incubated at 37 °C. Density of the broth (containing the suspended organisms) was adjusted to 0.5 McFarland standard [Sarkar et al., 2014].

#### 2.3.8. Bacterial susceptibility assay

The minimum inhibitory concentration (MIC) was determined using the broth micro-dilution method. Briefly, a standardized test inoculum (10 µL of a  $1-5 \times 10^5$  CFU/mL suspension)

was added to the wells of 96 well microtiter plate, containing 100  $\mu$ L of two-fold serially diluted sample (UA and MPA) in MHB (final concentrations ranging from 0.015 to 2 mg/mL). Plates were then incubated (100 rpm, 37 °C) for 18 h. DMSO (0.1%) was used as the negative control. The MIC value is defined as the lowest concentration where no viability was observed after 18 h [Sarkar et al., 2014].

### 2.3.9. Effect on biofilm formation

The effect of UA and MPA on biofilm formation was performed in 96-well polystyrene plates. A standardized inoculum (5  $\mu$ L of a  $1-5 \times 10^5$  CFU/mL suspension) was inoculated with 100  $\mu$ L of fresh MHB in presence or absence (non-treated control) of UA and MPA. Following incubation (24 h), non-adherent bacteria were removed by washing with sterile phosphate buffer saline (PBS; pH 7.2). Biofilms were stained with 1% crystal violet solution. The absorbance of the crystal violet solution (stain bound to biofilm was removed from each well employing 33% glacial acetic acid) was measured at 492 nm (Spectramax M5, Molecular Device). Wells containing medium and extract were used as blanks [Sarkar et al., 2015]. The percentage of inhibition of biofilm formation was calculated with the following formula:

$$\left( \frac{\text{Optical density at 492 nm of the test sample}}{\text{Optical density at 492 nm of non - treated control}} \right) \times 100 \quad \text{Equation 2.1}$$

### 2.3.10. Pyoverdine bioassay

Samples of UA and MPA over a range of concentrations were incubated for 24 h with a 10% culture of *P. aeruginosa*. The cells were removed by centrifugation, and the cell-free supernatant was used for the pyoverdine assay. The pyoverdine concentration in the supernatant was measured with fluorescence spectroscopy with excitation @ 405 nm and emission @ 465 nm in a multimode microplate reader (Spectramax M5; Molecular Device). The activity was recorded



in relative fluorescence units [Sarkar et al., 2014] and reported as the mean with the error expressed as the standard error of the mean (SEM).

#### 2.3.11. *Pyocyanin bioassay*

The test for the effect of UA and MPA on pyocyanin production was performed in 15 mL falcon tubes. A standardized inoculum (250  $\mu$ L of a  $1-5 \times 10^5$  CFU/mL suspension) of *P. aeruginosa* was inoculated with 5 mL of fresh MHB in absence (non-treated control) and presence (treated) of UA and MPA. The cell culture (5 mL) was extracted with chloroform (3 mL) and 1 mL of 0.2 M hydrochloric acid was added. The organic layer was collected by centrifugation and the absorbance was recorded at 520 nm [Chang, C. Y. et al., 2014] and reported as the mean with the error expressed as the standard error of the mean (SEM).

#### 2.3.12. *Swimming motility bioassay*

Tryptone swim plates (1% tryptone, 0.5% NaCl, 0.3% agar) were inoculated with a sterile toothpick and incubated for 16 h at 25 °C. Motility was assessed qualitatively by examining the circular turbid zone formed by the bacterial cells migrating away from the point of inoculation [Dézziel, E. et al., 2002].

#### 2.3.13. *Swarming motility bioassay*

Swarm plates were composed of 0.5% Bacto Agar and 8 g of nutrient broth/liter (Difco, Detroit, Mich.), supplemented with 5 g of dextrose/liter, and dried overnight at room temperature. Cells were point inoculated with a sterile toothpick, and the plates were incubated at 30 °C for 24 h [Dézziel, E. et al., 2002].

#### 2.3.14. Twitching motility bioassay

Cells were stab inoculated with a toothpick through a thin (approximately 3-mm) LB agar layer (1% agar) to the bottom of the petri dish. After incubation for 24 to 48 h at 30 °C, a hazy zone of growth at the interface between the agar and the polystyrene surface was observed [Déziel, E. et al., 2002].

### 2.4. Materials and methods related to Chapter 4

#### 2.4.1. Bacterial strains and growth conditions

Strains (Appendix **Table S3**) were grown in tryptic soy broth (TSB) (Difco) at 37 °C and 230 rpm shaking for routine overnight culturing unless otherwise stated.

#### 2.4.2. Minimum inhibitory concentration (MIC) and Minimum bactericidal concentration (MBC) assays

The CLSI protocol for determining the MIC and MBC of antibiotic activities in *S. aureus* was followed [CLSI, 2012]. Briefly, an overnight culture was diluted equivalent to a 0.5 McFarland standard and then further diluted to inoculate wells of a 96-well plate, containing an antibiotic dilution gradient, with  $5 \times 10^5$  cells each. The plate was grown statically at 37 °C for 24 h. To determine the MIC the plates were visually inspected for growth and by reading the optical density (OD<sub>600</sub>) with a BioTek plate reader. To determine the MBC,  $10^{-1}$ - $10^{-3}$  dilutions of cell suspensions exposed to the antibiotic dilution gradients were plated on LB plates and incubated at 37 °C for 24 h. The number of colony forming units (CFU) per dilution was recorded. As a positive control, cells suspensions treated similarly but without exposure to the compounds were used. The MBC was calculated using the formula:

$$MBC = \frac{\text{no. of CFU in a specific concentration}}{\text{no. of CFU in the positive control}} \times 100 \quad \text{Equation 2.2}$$

Where no. of CFU in a specific concentration is the average CFU of each dilution at each concentration tested. A concentration was considered bactericidal if it killed 99.9% of the bacterial population [Pankey, G. A. & Sabath, L. D., 2004].

#### 2.4.3. Minimum biofilm inhibitory concentration (MBIC) assay

Assays were performed as previously described [Zhang, E. et al., 2018]. Briefly, an overnight culture was diluted 100-fold and incubated at 37 °C for 3 h to reach mid-log phase. The subculture of *S. aureus* was prepared by diluting the overnight culture in 1:1000 (~10<sup>5</sup> CFU ml<sup>-1</sup>) into TSB. Wells of a 96-well plate were used to serially dilute the desired compound in TSB and then inoculated with 5x10<sup>6</sup> cells each and grown statically at 37 °C for 24 h. The wells were then rinsed twice with PBS (pH 7.2, 0.8% NaCl, 0.02% KCl, 0.17% Na<sub>2</sub>HPO<sub>4</sub>, 0.8% KH<sub>2</sub>PO<sub>4</sub>) and this was followed with the resazurin cell variability assay below.

#### 2.4.4. Minimum biofilm eradication concentration (MBEC) assay

Assays (Appendix **Table S4**) were performed as previously described [Zhang, E. et al., 2018]. Briefly, an overnight culture was diluted 100-fold and incubated at 37 °C for 3 h to reach mid-log phase. The subculture of *S. aureus* was prepared by diluting the overnight culture in 1:1000 (~10<sup>5</sup> CFU ml<sup>-1</sup>) into TSB. Wells of a 96-well plate were inoculated with 5x10<sup>6</sup> cells each and grown statically at 37 °C for 24 h. The media was then removed from each well. Then serial dilutions of the desired compound in TSB were added to the wells and grown statically at 37°C for 24 h. The wells were then rinsed twice with PBS (pH 7.2, 0.8% NaCl, 0.02% KCl, 0.17% Na<sub>2</sub>HPO<sub>4</sub>, 0.8% KH<sub>2</sub>PO<sub>4</sub>) and this was followed with the resazurin cell viability assay.

#### 2.4.5. Resazurin cell viability assay

Viable cells with active metabolism reduce resazurin into fluorescent resorufin [Evie, I. M. et al., 2017]. To the washed wells, 20  $\mu$ L of Cell Titre-Blue (Promega) was added with 100  $\mu$ L of PBS and cells were incubated at 37°C for 1 h. Fluorescence was measured (530/25 nm excitation and 590/35 nm emission) using a BioTek Synergy plate reader. Wells that had the same fluorescence as a resazurin-only control were considered to have no cell viability.

#### 2.4.6. Checkerboard Assay

Checkerboard assays in 96-well format were performed as previously described [Garcia, L. S., 2016]. Briefly, the inoculum was prepared as for the MIC testing above and added to a two-dimensional gradient of the compound of interest. Plates were grown statically at 37 °C for 24 h then visually inspected for growth. The fractional inhibition concentration (FIC) index is calculated using the formula.

$$FIC\ Index = \frac{C_1}{MIC_1} + \frac{C_2}{MIC_2} \quad \text{Equation 2.3}$$

where  $MIC_1$  and  $MIC_2$  are the MICs of antibiotics 1 and 2 alone, respectively, and  $C_1$  and  $C_2$  are the concentrations of antibiotics 1 and 2, respectively, when at the combined MIC. When the FIC index was  $\leq 0.5$  the antibiotic combination was synergistic; when the FIC index was  $> 0.5$  and  $\leq 4.0$  the effect of the two antibiotics was considered to be additive. Finally, when the FIC index was  $\geq 4.0$ , the antibiotic combination was defined as antagonistic [Belkum, V. A. et al., 2015].

#### 2.4.7. Hemolysis assay

Hemolytic activity on ovine erythrocytes was assessed as previously described [Hogan, A. M. et al., 2018 and Selin, C. et al., 2015]. Briefly, ovine erythrocytes were gently pelleted by

centrifugation and washed thrice with PBS then resuspended in PBS at a 1:5 dilution. In 96-well format, 100  $\mu$ l erythrocyte suspension was added to a dilution gradient of the desired compound and incubated statically at 37 °C for 1 hour. Triton X-100 at 0.1% was included as a positive control and a dimethyl sulfoxide (DMSO) concentration gradient as a negative control. Erythrocytes were pelleted and the absorbance of the supernatant was measured at 540 nm. The percentage (%) of lysis was calculated as follows.

$$\% \text{ Lysis} = \frac{X - BT - B}{X - BT} \times 100\% \quad \text{Equation 2.4}$$

where B is the  $A_{540\text{nm}}$  of the negative DMSO control, T is the  $A_{540\text{nm}}$  of the positive Triton X-100 control, and X is the  $A_{540\text{nm}}$  of the analyzed sample.  $A_{540\text{nm}} > 40\%$  hemolysis was considered as high percent hemolysis and 5–10% hemolysis was considered as low percent hemolysis for a given compound [Sperandio, D. et al., 2010].

#### 2.4.8. *Synthesis and characterization of compounds 1 – 12*

Reagents and solvents were used as purchased without further purification. Thin Layer Chromatography (TLC) was conducted with 0.25 mm F<sub>254</sub> silica gel on aluminum plates and visualization was conducted with 254 nm UV light. The purification of all compounds was carried out over flash grade-silica gel and fractions combined according to TLC. Solvent was removed under reduced pressure. The <sup>1</sup>H (300 MHz) and <sup>13</sup>C NMR (125 MHz) spectra were collected on a Bruker Avance-300 console and were referenced to solvent (CDCl<sub>3</sub> = 7.24 ppm). The complete synthetic procedures and characterization data for compounds **1 – 12** are included in the **Chapter 4 Section** of the Appendix.

## 2.5. Materials and methods related to Chapter 5

### 2.5.1. RNA extraction from *C. uncialis* and cDNA synthesis

The lichen sample of *C. uncialis* (lichen species), collection number Normore 8774 (Michelle Piercey Normore lichen collection), was collected in Manitoba, Canada, on a Jack Pine dominated south-facing granite ridge. The lichen material was air dried and stored in the University of Manitoba Herbarium cryptogam division until needed. The lichen sample was crushed in liquid nitrogen and 150 mg of lichen fungal cells were used to extract & purify RNA using a MasterPure™ Yeast RNA Purification Kit from Lucigen. But this kit did not work for lichen fungus RNA extraction and purification. Subsequently total mRNA was extracted using the E.Z.N.A. Fungal RNA Mini Kit (*Omega*), in accordance with the manufacturer's protocol to successfully get an average quality RNA (209.4 ng/ µl). An aliquot of mRNA extract was set aside for analysis, and the remainder was treated with DNase (*ThermoScientific*). An aliquot of DNase-treated mRNA was set aside for analysis, and reverse transcription was applied to the remainder using the RevertAid H Minus First Strand cDNA Synthesis Kit (*ThermoScientific*) using Oligo (dt)<sub>18</sub> primers, in accordance with the manufacturer's protocol. Aliquots of the DNA-contaminated mRNA and DNase-treated mRNA were cleaned using a PCR cleanup kit (*ThermoScientific*). The efficacy of DNase was evaluated by using the cleaned aliquots as templates for PCR reactions.

### 2.5.2. Polymerase chain reaction

All PCR reactions were performed using Phusion polymerase and reagents (*ThermoScientific*). Per 20 µL: 4 µL Phusion HF Buffer (containing 7.5 mM MgCl<sub>2</sub>), 0.8 µL 10 mM dNTP mix, 0.6 µL of 50 mM EDTA, 0.8 µL of 10 µM forward primer, 0.8 µL of 10 µM reverse primer, template DNA (variable concentration), 0.2 µL of 2 U/ µL Phusion polymerase, 11.8 µL of water. Primers are provided in **Table S1** in the Appendix. PCR was performed with a

SimpliAmp thermocycler (*Applied Biosystems*): Initial denaturation was at 94°C for 60 sec., 25 rounds of denaturation at 94°C for 10 sec., 25 rounds of annealing at 3-5°C below primer T<sub>m</sub> for 15 sec., 25 rounds of extension at 72 °C for 30 sec. per KB, and final extension at 72 °C for 30 sec. per KB. Amplicons were electrophoresed on a 1 % agarose gel containing 1.25 µL of CYBR Safe DNA Gel Stain (*Invitrogen*) and electrophoresed for 45-60 min. using a gel box (*BioRad*) and Powerpack (*VWR*) programmed at 230 mA and 200 V. Amplicons were visualized using blue light, excised from the gel, and the DNA purified using the GeneJet Gel Purification Kit (*ThermoScientific*) in accordance with the manufacturer's protocols. The amplified DNA was prepared for plasmid construction by concentrating the template using a SpeedVac. DNA concentration and quality were evaluated using a model 2000c NanoDrop (*ThermoScientific*).

### 2.5.3. Preparation of plasmids

#### 2.5.3a. InFusion method

Transformation plasmids (pETite N-His SUMO & pET21a) were linearized by restriction digestion with FastDigest NotI (*ThermoScientific*) in accordance with the manufacturer's instructions. The linearized plasmid was electrophoresed on agarose gel, purified, concentrated, and evaluated for concentration and quality (see 'PCR', above). Genes amplified from *C. uncialis* were bound to linearized plasmid using InFusion (*Takara*) in accordance with the manufacturer's instructions. Plasmid constructs were screened by transformation into Stellar *E. coli* (*Takara*) using the manufacturer's recommended heat-shock method, then plated on LB media (Per 1 L of water: 10 g tryptone, 5 g yeast extract, 5 g NaCl, 0.03 g kanamycin; pH unadjusted for pETite N-His SUMO) & (Per 1 L of water: 10 g tryptone, 5 g yeast extract, 5 g NaCl, 0.1 g ampicillin; pH unadjusted for pET21a). Transformants were inoculated by sterile toothpick into liquid LB media containing appropriate selection marker/antibiotic and grown overnight. Plasmids were extracted

from *E. coli* suspension using the GenJet Plasmid Miniprep Kit (*ThermoScientific*), in accordance with the manufacturer's instructions. The extracted plasmids were evaluated by DNA sequencing using a commercial service provider (Eurofins MWG Operon).

#### 2.5.3b. Conventional ligation method

Transformation plasmids (pET28b, pET21c & pETDuet-1) were linearized by restriction digestion with FastDigest restriction enzymes (*ThermoScientific*); ECoRI/NdeI for pET28b, HindIII/NotI for pET21c and, ECoRI/NotI to insert *cu-cpr357* gene in pETDuet-1 and FseI/KpnI to insert MPAO in pETDuet-1, in accordance with the manufacturer's instructions. Linearized plasmids were electrophoresed on agarose gel, purified, concentrated, and evaluated for concentration and quality (see 'PCR', above). Genes from *C. uncialis* were amplified (see 'PCR', above) and digested with appropriate FastDigest restriction enzymes. After restriction digestion, genes with their sticky ends were bound to linearized plasmid using T4 DNA Ligase (*ThermoScientific*) in accordance with the manufacturer's instructions. Plasmid constructs were screened by transformation into Stellar *E. coli* (*Takara*) using the manufacturer's recommended heat-shock method, then plated on LB media (Per 1 L of water: 10 g tryptone, 5 g yeast extract, 5 g NaCl, 0.1 g ampicillin; pH unadjusted for). Transformants were inoculated by sterile toothpick into liquid LB media containing ampicillin and grown overnight. Plasmids were extracted from *E. coli* suspension using the GenJet Plasmid Miniprep Kit (*ThermoScientific*), in accordance with the manufacturer's instructions. The extracted plasmids were evaluated by DNA sequencing using a commercial service provider (Eurofins MWG Operon). A list of all plasmids prepared for this work is provided in **Table S2** in the Appendix.



#### 2.5.4. cDNA amplification & cloning of *cu-mpao*

Further, cDNA was amplified with MPAO specific primers. The primers were designed for In-Fusion seamless directional cloning in Stellar cells using pETite N-His SUMO vector system. The plasmid was linearized as mentioned above (**Section 2.5.3a**). The PCR product (see ‘PCR’, above), and the linear plasmid was purified from agarose gel using a gel extraction kit (*ThermoScientific*). The gel purified PCR product and the pETite vector were then incubated using InFusion master mix for 15 min. (In-Fusion HD Cloning Kit by Takara Bio) and the reaction product was transformed into Stellar cells (*Takara*). The extracted plasmid was evaluated by DNA sequencing using a commercial service provider (Eurofins MWG Operon).

#### 2.5.5. Transformation and expression of Truncated MPAO, codon optimized for *E. coli* in *BL21(DE3)pLySs*

Custom synthesized plasmid pET28a carrying MPAO without a transmembrane domain and codon optimized for *E. coli* was obtained from ABE Ikuro’s laboratory at University of Tokyo, Japan. Primers were designed to amplify MPAO out of pET28a and clone into pETite N-His SUMO vector (see ‘PCR’, above). In-Fusion cloning strategy was used to clone MPAO codon-optimized for *E. coli* w/o TMD gene (named as MPAO(Eco)-pETite) into stellar competent cells. Further, pETite vector with MPAO was gel purified from the selected and tested stellar cell colonies and then transformed into *E. coli* BL21(DE3)pLySs (*ThermoScientific*) using the manufacturer’s recommended heat-shock method. Colonies were selected on the agar plate with a specific antibiotic (kanamycin in this case). A single colony was resuspended in 5 mL of LB media with kanamycin (at a conc. of 30 mg/mL) and was grown overnight at 37 °C. The overnight grown bacterial culture was divided into multiple microcentrifuge tubes (each with 500 µL of sample) and 500 µL of 40% glucose in LB solution was added to each of the tubes to make the final 1 mL

of 20% stock solution and stored at -80 °C. 10 mL of LB containing 30 mg/mL kanamycin was inoculated with 1 mL of stock solution (thawed on ice) and the culture was grown overnight at 37 °C. The following morning, overnight growth with  $OD_{600} = 1.0$  was measured using a UV spectrophotometer and the cultures were diluted to  $OD_{600} = 0.02$  in 250 mL LB media plus kanamycin. When cultures reached an optical density at 600 nm ( $OD_{600}$ ) of 0.4-0.6, a 1-mL aliquot of uninduced cells was collected by pelleting in a microcentrifuge tube (12,000 x g for 1 min). The cell pellet was resuspended in 50  $\mu$ L SDS-PAGE loading buffer (*BioRad*) and stored at -20 °C until SDS-PAGE analysis. To induce the protein expression, IPTG was added (*ThermoScientific*) to a final concentration of 0.5 mM and the cultures were shaken continuously at 37 °C for 2 h. The  $OD_{600}$  of the induced culture was recorded and a 500  $\mu$ L culture was harvested by microcentrifugation. The cell pellet was resuspended in 50  $\mu$ L SDS-PAGE loading buffer (*BioRad*) and stored at -20 °C until SDS-PAGE analysis to evaluate the protein expression. The remaining culture was harvested by centrifugation at 4000 x g for 15 min. The supernatant was discarded, and pellet was re-suspended in 5 mL (1/20 to 1/50 lysis buffer of the volume of culture used) of lysis buffer (50 mM  $NaH_2PO_4$  or 20 mM Tris HCl, 300 mM NaCl or 150 mM NaCl, pH 8.0) on ice. Protease inhibitor cocktail (10  $\mu$ L/100X) (*ThermoScientific*) was added to 1 mL of the lysis buffer, which protects an intact, active cellular protein from degradation by endogenous proteases. Cells were lysed by sonication on ice. Six-ten pulses were used for 10 sec each with a microtip and the samples were allowed to cool for 30 sec between pulses to avoid frothing. Avoid frothing. The lysate was centrifuged at 10000 x g for 20 min at 4 °C. The supernatant was collected in a microcentrifuge tube on ice, which contained the soluble protein and protein concentration was measured using nanodrop. The sample were saved at -20 °C for the further purification step. The pellet was resuspended, containing insoluble proteins and un-lysed cells in 5 mL of urea lysis

buffer (7.5-8M Urea, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, pH 12.5). Samples after treatment with SDS-loading buffer were analyzed with SDS-PAGE gel. For minigels, 0.05 OD<sub>600</sub> equivalent per lane usually contains the appropriate amount of protein for Coomassie blue staining.

#### 2.5.6. Expression of CPR357, MPAO w/o SUMO tag, STCB (*A. nidulans* Sterigmatocystin B) and SCYP (*Penicillium marneffii* 1 Sterigmatocystin cytochrome p450)

A similar protein expression protocol was followed as detailed above in **Section 2.5.5** except with some modifications in protein induction temperature, time conditions and IPTG concentration.

#### 2.5.7. Double transformation of two genes in a bacterial cell.

A plasmid harboring one gene (G<sub>1</sub>) was transformed into *E. coli* BL21(DE3)pLySs (*ThermoScientific*) using the manufacturer's recommended heat-shock method. Colonies were selected on the agar plate with a specific antibiotic. A single colony was resuspended in 5 mL of LB media with the respective antibiotic and was grown overnight at 37 °C. The overnight growth with OD<sub>600</sub> = 1.0 was measured using a UV spectrophotometer. The following morning, the overnight growth with OD<sub>600</sub> = 1.0 was measured using a UV spectrophotometer and the competent bacterial cells were synthesized in accordance with the manufacturer's instructions ([http://mcb.berkeley.edu/labs/krantz/protocols/calcium\\_comp\\_cells.pdf](http://mcb.berkeley.edu/labs/krantz/protocols/calcium_comp_cells.pdf)). The competent bacterial cells harboring G<sub>1</sub> clone were thawed on ice and transformed with the second plasmid with another gene (G<sub>2</sub>) using the manufacturer's recommended heat-shock method. Colonies were selected on the agar plate with two antibiotics (each for two plasmids). A single colony with both genes (G<sub>1</sub> & G<sub>2</sub>) was resuspended in 5 mL of LB media plus antibiotic and was grown overnight at 37 °C. 500 μL of 40% LB with glucose was inoculated with 500 μL of the overnight grown bacterial culture to make the final 20% of the stock solution (1-mL) and stored at -80 °C.

#### 2.5.8. Co-expression of two proteins

Colonies were selected on the agar plate with a specific antibiotic. 10 mL of LB containing specific antibiotic was inoculated with 1 mL of stock solution (thawed on ice) and the culture was grown overnight at 37 °C. The following morning, the overnight growth with  $OD_{600} = 1.0$  was measured using a UV spectrophotometer and the cultures were diluted to  $OD_{600} = 0.02$  in 250 mL TB (Terrific Broth) media with a specific antibiotic. When cultures reached an optical density at 600 nm ( $OD_{600}$ ) of 0.4-0.6, a 1 mL aliquot of uninduced cells was collected by pelleting in a microcentrifuge tube (12,000 x g for 1 min). The cell pellet was resuspended in 50  $\mu$ L SDS-PAGE loading buffer (*BioRad*) and stored at -20 °C until SDS-PAGE analysis. IPTG (1 mM) (*ThermoScientific*) was added to induce the expression of recombinant proteins and the cultures were shaken continuously at 22 °C for 16 h. The  $OD_{600}$  of the induced cultures was recorded and a 500  $\mu$ L of each culture was harvested by microcentrifugation. Resuspended the cell pellet in 50  $\mu$ L SDS-PAGE loading buffer (*BioRad*) and stored at -20 °C until SDS-PAGE analysis to evaluate the protein expression. The protocol used for the co-expression of two proteins was same in both cases, either two genes were cloned on a single plasmid (pETDuet-1) and transformed in BL21(DE3)pLySs cells or two genes were cloned into two separate plasmids and doubly transformed in BL21(DE3)pLySs cells using the manufacturer's recommended heat-shock method.

#### 2.5.9. Protein Purification

Purification of the soluble proteins tagged with six N-terminal histidine residues from *E. coli* was performed utilizing a  $Ni^{2+}$ -NTA (*ThermoScientific*) resin under nondenaturing conditions, in accordance with manufacturer's instructions. Concentration of all eluted fractions of a pure protein was measured using UV spectrophotometrically at 450 nm.

### 2.5.10. SDS-PAGE

A stacking gel was prepared within Mini-Protean System Glass Plates (*BioRad*) to a final acrylamide concentration of 7.5 % or 10.0 % (Resolving layer, per 20 mL: 4.68 mL or 6.24 mL of 40% 29/1 acrylamide/bisacrylamide, 7.50 mL of 1 M Tris (pH 8.8), 0.20 mL of 10 % ammonium persulfate, 0.02 mL of TEMED, 7.86 mL or 9.42 mL of water; stacking layer, per 10 mL: 0.93 mL or 1.24 mL of 40% 29/1 acrylamide/bisacrylamide, 3.76 mL of 1M Tris (pH 6.8), 0.10 mL of 10 % ammonium persulfate, 0.01 mL of TEMED, 5.20 mL or 5.51 mL of water). The concentration of protein samples was measured using the Bradford technique (Bradford, 1976). Protein samples were mixed in a 3:1 ratio with 4x loading buffer (2 mL of 1 M Tris (pH 6.8), 0.8 g of sodium dodecyl sulfate, 4 mL of glycerol, 0.4 mL of 14.7 M  $\beta$ -mercaptoethanol, 1 mL of 0.5 M EDTA, 0.008 g of bromophenol blue, water to 10 mL). Proteins were denatured by incubation at 95 °C for 5 min., and 15 mg of total protein extract was loaded into each well. Molecular weight standards used were either 5  $\mu$ L of PageRuler Plus Pre-Stained Protein Ladder (*ThermoScientific*) or 10  $\mu$ L of HiMark Pre-stained Protein Standard (*Invitrogen*). Electrophoresis was performed using a Mini-Protean Tetra System (*BioRad*) and a Powerpack (*VWR*) set to 230 mA and 200 V. Samples were electrophoresed under constant current for 1-3 h.

### 2.5.11. Western Blot

Two Trans-Blot Turbo Mini-size Transfer Stacks (*BioRad*) and the SDS-PAGE gel were washed/wetted with Trans-Blot Turbo Transfer Buffer (*BioRad*) and assembled into a tower along with a methanol-activated Immun-Blot PVDF membrane (*BioRad*), in accordance with the manufacturer's instructions. Blotting was performed using a Trans-Blot Turbo system (*Biorad*) set to 1.3 A, 25 V, and 7 min. The tower was disassembled, and the membrane was washed for five min. on an end-to-end rotator in 10 mL of TBS solution (Per 1 L of water: 3 g Tris, 0.2 g KCl, 8 g

NaCl, pH adjusted to 7.4). Blocking was performed by incubating the membrane in a 5 % milk powder-TBS solution for one hour on rotator. Primary antibody application was provided by applying for one hour on rotator 10 ml of 5 % milk powder-TBS solution containing 1  $\mu$ L of Goat Anti-Mouse IgG (*EMD Millipore*). The membrane was washed three times with TBS solution in ten min applications. Chemiluminescence was performed by applying 1.5 mL of Immobilon Forte Western HRP Substrate (*EMD Millipore*) for five min. on rotator. Membranes were imaged using FluorChem Q system (*Cell Biosciences*) in accordance with manufacturer's instructions.

#### *2.6.12. Bioassays of converting MPA to UA catalyzed by MPAO (whole-cell catalyst; in vivo)*

MPA (5 mg, the substrate) was dissolved in 400  $\mu$ L water + 100  $\mu$ L acetone, as MPA is not water soluble but sparingly soluble in acetone and the high percentage of acetone solvent causes bacterial cell death [Dyrda, G. et al., 2019]. Water plus acetone solution containing MPA (5 mL) was added to 250 mL of bacterial culture (LB media) harboring MPAO clone which was grown to OD<sub>600</sub> 0.569 at 37 °C. Protein was expressed following the protocol discussed in **Section 2.5.5**. After 1 h of protein induction at 37 °C and the substrate incubation, 25 mL of culture was collected from each flask, acidified using 1M HCl to quench the reaction and the product was extracted using 25 mL of ethyl acetate. Combined ethyl acetate fractions were dried over sodium sulphate (Na<sub>2</sub>SO<sub>4</sub>) and solvent was evaporated using rotovap and the sample was stored at 4 °C. The same procedure was repeated after 2 h of protein induction period (25 mL collection from the remaining bacterial culture, followed by extraction & sample preparation). The rest of the 200 mL of culture was harvested after 3 h of protein induction and extracted in portions, each of 25 mL following the same steps as for earlier extraction. Experiments were done in triplicates and all the samples were dissolved in 1mL of HPLC-grade methanol to screen for UA using UV coupled HPLC detector.

### 2.5.13. Incubation conditions of MPA with both MPAO and CPR357 enzymes (*in vitro*)

Both MPAO and CPR357 proteins were expressed separately in *E. coli* (protein expression protocol was already discussed in **Section 2.5.5**). Protein expression was induced for 2-3 h using 0.5 mM IPTG. Both bacterial cultures (250 mL each) were harvested by centrifugation at 4000 x g for 15 min where one was harboring the MPAO clone and the other was harboring the CPR357 clone. The supernatants were discarded, and the cell pellets were re-suspended in 5 mL of lysis buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub> or 20 mM Tris HCl, 300 mM NaCl or 150 mM NaCl, pH 8.0) on ice. Protease inhibitor cocktail (50 µL, 100X) (*ThermoScientific*) was added to the 5 mL of the lysis buffer solutions, which protects intact, active cellular proteins from degradation by endogenous proteases. Cells were lysed by sonication on ice. Six-ten pulses were used for 10 sec each with a microtip and the samples were allowed to cool for 30 sec between pulses to avoid frothing. The lysate was centrifuged at 10000 x g for 20 min at 4 °C. The supernatants were collected in a microcentrifuge tube on ice, which contained the soluble MPAO and CPR357 proteins and their concentration was measured using nanodrop. 4 mg of MPA (substrate) & 24 mg of NADH (electron donor) were dissolved in 1 mL of acetone. The reaction mixture of 150 µL of MPAO protein sample (in lysis buffer), 150 µL of CPR357 protein sample (in lysis buffer) and 75 µL of 20% glycerol (prepared in water) was prepared in a microcentrifuge tube where 125 µL of acetone (dissolved MPA & NADH) was added to the mixture on ice. Both proteins were incubated with the substrate and the electron donor for 2 h at 4 °C in the cold room. Reaction was quenched with 500 µL mixture of 1:1 MeOH & acetone. Samples were vortexed for a few seconds and then centrifuged at 16000 x g for 2 min. Filtered the supernatant using 0.45-micron filter and 250 µL was transferred to an HPLC vial and further diluted by adding 750 µL of HPLC-grade MeOH to

be analysed for UA on HPLC. Another 250  $\mu$ L was diluted with 750  $\mu$ L acetonitrile to be analysed for UA on LC-MS.

*2.5.14. Incubation experiments of MPA with co-expressed MPAO+CPR357 proteins and co-expressed STCB+CPR357 proteins (in vivo)*

The protein expression protocol was already discussed in **Section 2.5.7**. Protein expression induction period was 48 h using 1 mM IPTG. Bacterial cultures (125 mL each) were harvested by centrifugation at 4000 x g for 15 min and the supernatant was discarded after centrifugation. Cells were resuspended in 1 mL of 100 mM potassium phosphate buffer with 12.5 mM glucose (prepared in sterile water) in a microcentrifuge tube. MPA (20 mg) was dissolved in 1 mL of DMSO and 100 mg of NADH was dissolved separately in 1 mL of water. The reaction mixture was incubated on the rocker for 4 h at room temperature in a shaker. MPA (50  $\mu$ L, final concentration is 0.2 mM) and 70  $\mu$ L of NADH (final concentration is 0.7 mM) were added to 1 mL of the potassium phosphate buffer (*SigmaAldrich*) reaction mixture. Then the reaction mixture was transferred in a glass vial and the reaction was quenched with 1 mL of MeOH and a drop of conc. HCl. Samples were vortexed for a few seconds and then centrifuged at 16000 x g for 2 min. The supernatant was filtered using 0.20-micron filter and 250  $\mu$ L was transferred to an HPLC vial. Further the supernatant was diluted by adding 750  $\mu$ L of HPLC-grade MeOH to be analysed for UA on HPLC. Another 250  $\mu$ L was diluted with 750  $\mu$ L acetonitrile to be analysed for UA on LC-MS.

*2.5.15. Incubation experiments of MPA with co-expressed MPAO+CPR357 proteins and co-expressed STCB+CPR357 proteins (in vitro)*

The protein expression protocol was already discussed in **Section 2.5.7**. The protein expression induction period was 48 h using 1 mM IPTG. Bacterial cultures (125 mL each) were harvested by centrifugation at 4000 x g for 15 min and the supernatant was discarded after



centrifugation. The cell pellet was re-suspended carefully in 1mL of lysis buffer (composition is provided in **Section 2.5.5**) and added 20  $\mu$ L of protease inhibitor. Cells were frozen in liquid nitrogen and thawed on ice and then sonicated at 50% amplitude for 15 sec on and 15 sec off with 6 pulses on ice. The cells were centrifuged at 10,000xg for 20 min at 4 °C and supernatant (soluble fraction) was collected in a 1 mL microcentrifuge and pellet was redissolved in 1 mL of 7.5M urea lysis buffer (lipid fraction) in a cold room at 4 °C. In each, soluble and lipid fractions, 50  $\mu$ L of MPA (substrate, 20 mg was dissolved in 1 mL of DMSO) at a concentration of 0.2 mM/ mL. And 70  $\mu$ L of NADH was also added to each of the reaction mixture (electron donor, 100 mg dissolved in 1mL of water) to a concentration of 0.7 mM/ mL. Incubated these samples on the rocker for 3 hrs at 4 °C and added 1 mL of methanol after incubation to quench the reactions. Sample preparation procedure is mentioned above in **Section 2.5.12** for analyses on HPLC and LC-MS.

#### 2.5.16. HPLC

Analysis by HPLC of extracts of *E. coli* media was performed using a Waters HPLC Separations Module 2695 combined with a PDA Detector Model 2996. The column was a  $\mu$ Bondapak® Waters C18 (3.9 X 300 mm) column, with a particle diameter of 15-20  $\mu$ m with 125 Å pores. The flow rate was 1 mL/min, and the eluent was monitored continuously at 210-600 nm. The elution was done at 20 % methanol and 80 % water containing 0.075 % aqueous trifluoroacetic acid for 10 min, following which a linear gradient was applied to 80 % methanol and held at that composition for 20 min, followed by application of a linear gradient back to 20 % methanol for 10 min and held there for 10 min. The total run time was 60 min.

#### 2.5.17. LC-MS/MS

LC-MS/MS analyses were performed on a Bruker system coupled to a Bruker compact qTOF mass spectrometer equipped with an electrospray ionization interface (ESI).

Chromatographic separations were performed on an Intensity Solo 2 C18 (100 × 2.1 mm 1.7 μm) column. The mobile phase comprised H<sub>2</sub>O (A) and MeCN (B), both of which were acidified with 0.1% formic acid. The column temperature and sample organizer were maintained at 35 °C and 15 °C, respectively. A stepwise gradient method at constant flow rate of 0.4 mL/min was used to elute the column with the following conditions: 40-80% B (0–25 min), followed by a 30 min run of MeCN for washing and 5 min of reconditioning. Analyses of the samples (5.0 μL injected) were performed in negative ion mode consisting of a full MS survey scan in the m/z 100–1500 Da range (scan time: 100 ms) followed by MS/MS scans for the three most abundant ions. The collision energy was applied at 10 V. The negative ESI conditions were set as follows: capillary voltage 3.5 kV, source temperature 200 °C, gas flow 4 L/h. Samples were prepared by extracting 10 mg of sample into 1 mL of acetonitrile, centrifuging at 15.0 rpm at 20 min twice so that the samples were ready for HRMS. All data analysis was through the Bruker Data Analysis program.

#### 2.5.18. BLAST search, multiple sequence alignment and phylogenetic

Amino acid sequences of *C. uncialis* CYP (also known as MPAO, accession no. AUW31051.1) were selected for pBLAST [Atschul, S. F. et al. 1990] and multiple sequence alignment analysis. Two out-groups were selected, a human and a fungal p450 (Accession no. 1OG2\_A, XP\_749134.1). The 30 most genetically similar entries in GenBank were compiled for phylogenetic analysis with more than 90% query cover and 35% amino acid identity Multiple sequence alignment of MPAO, PM1 or SCYP (Accession no. XP\_002145795.1) & STCB (Accession no. XP\_081093.1) was performed using MView from EMBL-EBI [Brown N. P. et al., 1998]. Phylogenetic tree was constructed using MEGAX [Kumar et al., 2018]. Evolutionary history was inferred using the maximum-likelihood (ML) method [Whelan, S. and Godman, N., 2001]. The assembled tree was evaluated with 1000 replicates of the bootstrap test [Felsenstein, J.

1985]. The evolutionary distance was estimated using the poisson model [Zuckerkindl, E. & Pauling, L. 1965] and are in the units of the number of amino acid substitutions per site. The rate variation across sites was uniform, the gaps/missing data treatment selected as partial deletion and the site coverage cut-off% was 95. Under the system resource usage, number of threads were chosen to be '4'. The resulting tree was saved as a PDF file and is provided in Appendix.

## **2.6. Materials and methods related to Chapter 6**

### *2.6.1. Annotation of lichen gene clusters*

As described in detail by [Bertrand, R. L. 2019], the annotation of the gene clusters was performed using Antibiotics and Secondary Metabolite Analysis Shell (AntiSMASH version 3.0) [Weber et al., 2015], freely available for academic use at <http://www.secondarymetabolites.org/>.

### *2.6.2. BLAST search of genetically similar gene clusters*

Data was modified and added (Appendix **Table S5-S11**) to those already described by [Bertrand, R. L. 2019] based on the new entries in GenBank. pBLAST analysis was performed on each genetically similar gene to determine percent similarity score [Altschul, S. F. et al., 1990]. Lichen genes not identified as genetically similar to genes of the reference clusters were manually analyzed using comparative pBLAST to confirm absence of similarity. The reporting of 'no significant similarity found' was interpreted as a negative result.

### *2.6.3 Phylogenetics*

Genes were analyzed via pBLAST [Altschul, S. F. et al. 1990] and the 38 most genetically similar entries in GenBank were compiled for phylogenetic analysis with more than 90% query cover and 50% amino acid identity. Two out-group proteins were selected from organisms known to be more distantly related than organisms of the in-group genes that also share the general

function of in-group genes. Two out-group proteins were included for a total of 40 entries per phylogenetic tree. Amino acid sequences (Multiple sequence alignment) were aligned by *ClustalW* [Higgins, D. et al., 1994] package of MEGAX [Kumar, S. et al., 2018] and exported as a mega file, freely available for download at <http://www.megasoftware.net/>. This mega file was processed in Neighbour-Joining (NJ) statistical method, which was used while reconstructing phylogenetic trees. In some cases, the maximum-likelihood (ML) [Whelan, S. and Godman, N., 2001] method was used to perform evolutionary analyses. One-thousand replicates of pseudo samples were used of bootstrap support [Felsenstein, J. 1985] for analyses. The poisson model [Zuckerandl, E. & Pauling, L. 1965] was used as a substitutional model to evaluate the evolutionary distances and data are expressed in the units of the number of amino acid substitutions per site. The rate variation across sites was uniform, the gaps/missing data treatment selected as partial deletion and the site coverage cut-off% was 95. Under the system resource usage, the number of threads was chosen to be '4'. The resulting trees were saved as PDF files (**Figures S1-S25**) and are provided in the **Chapter 6 Section** of the Appendix.

## **2.7. Materials and methods related to Chapter 7**

*Not applicable*

## References

- Abdel-Hameed, M. Polyketide Biosynthesis in Lichen Fungi *Cladonia Uncialis* (Doctoral dissertation). University of Manitoba, Winnipeg, Canada, **2015**, Retrieved from [mspace.lib.umanitoba.ca › bitstream › Abde-Hameed\\_Mona](https://mspace.lib.umanitoba.ca/bitstream/Abde-Hameed_Mona)
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, 215: 403-410.
- Armaleo, D., May, S. Sizing the fungal and algal genomes of the lichen *Cladonia grayi* through quantitative PCR. *Symbiosis* **2009**, 49: 43-51.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., Pevzner, P. A. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, 19: 455-477.
- Bertrand, R. L. Characterization of Secondary Metabolome of a Lichenizing Fungus. University of Manitoba, Winnipeg, Canada, **2019**, Retrieved from <https://mspace.lib.umanitoba.ca/handle/1993/33943>
- Blin, K., Medema, M. H., Kazempour, D., Fischbach, M. A., Breitling, R., Takano, E., Weber, T. AntiSMASH 2.0 – a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **2013**, 41, 204-212.
- Chang, C. Y., Krishnan, T., Wang, H., Chen, Y., Yin, W. F., Chong, Y. M., Tan, L. Y., Chong, T. M., Chan, K. G. Non-antibiotic quorum sensing inhibitors acting against N-acyl homoserine lactone synthase as druggable target. *Scientific Reports* **2014**, 4, 7245.

- Clinical and Laboratory Standards Institute CLSI. *Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically; approved standard—ninth edition*. CLSI, **2012**.
- Déziel, E., Comeau, Y., Villemur, R. Initiation of biofilm formation by *Pseudomonas aeruginosa* 57RP correlates with emergence of hyperpiliated and highly adherent phenotypic variants deficient in swimming, swarming, and twitching motilities. *Journal of Bacteriology*. **2002**, 283, 1195-1204.
- Dyrda, G., Boniewska-Bernacka, E., Man, D. *et al.* The effect of organic solvents on selected microorganisms and model liposome membrane. *Mol Biol Rep* **2019**, 46, 3225–323.
- Evie, I. M., Dickson, A. J. & Elvin, M. Metabolite profiling of mammalian cell culture processes to evaluate cellular viability. in *Cell Viability Assays: Methods and Protocols* (eds. Gilbert, D. F. & Friedrich, O. Springer New York, **2017**, 137–152.
- Ewing, B., Green, P. Base-calling of automated sequencer tracers using phred. II. Error probabilities. *Genome Res.* **1998**, 8: 186-194.
- Ewing, B., Hillier, L., Wendl, M. C., Green, P. Base-calling of automated sequencing traces using phred. I. Accuracy assessment. *Genome Res.* **1998**, 8: 175-185.
- Felsenstein, J. Phylogenies and the comparative method. *Am. Nat.* **1985a**, 125: 1-15.
- Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap, *Evolution*. **1985b**, 39: 783-791.
- Garcia, Lynne S. *Clinical Microbiology Procedures Handbook, Volumes 1-3 (3rd Edition)*. American Society for Microbiology (ASM). **2016**, 3.
- Hawranik, D. J., Anderson, K. S., Simmonds, R., Sorensen, J. L. The chemoenzymatic synthesis of usnic acid. *Bioorganic & Medicinal Chemistry Letters*. 2009 May;19(9):2383-2385.

- Higgins, D., Thompson, J., Gibson, T. Thompson, J. D., Higgins, D. G., Gibson, T. J. *Nucleic Acids Res.* **1994**, 22:4673-4680.
- Hogan, A. M. *et al.* Competitive fitness of essential gene knockdowns reveals a broad-spectrum antibacterial inhibitor of the cell division protein FtsZ. *Antimicrobial Agents and Chemotherapy.* **2018**, 62, e01231-18.
- Kearse, M., Moir. R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., Drummond, A. Geneious basic: An integrated and extendable desktop software program for the organization and analysis of sequence data. *Bioinformatics*, **2012**, 28 1647-1649.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms, *Molecular Biology and Evolution*, **2018**, 35(6), 1547–1549.
- Pankey, G. A. & Sabath, L. D. Clinical relevance of bacteriostatic versus bactericidal mechanisms of action in the treatment of Gram-positive bacterial infections. *Clin. Infect. Dis.* **2004**, 38, 864–870.
- Sarkar, R., Chaudhary, S. K., Sharma, A., Yadav, K. K., Nema, N. K., Sekhoacha, M., Karmakar, S., Braga, F. C., Matsabisa, M. G., Mukherjee, P. K., Sen T. Anti-biofilm activity of Marula - a study with the standardized bark extract. *Journal of Ethnopharmacology*, **2014**,154, 170-175.
- Sarkar, R., Mondal, C., Bera, R., Chakraborty, S., Barik, R., Roy, P., Kumar, A., Yadav, K. K., Choudhury, J., Chaudhary, S. K., Samanta, S. K., Karmakar, S., Das, S., Mukherjee, P. K., Mukherjee J, Sen T. Antimicrobial properties of *Kalanchoe blossfeldiana*: a focus on drug resistance with particular reference to quorum sensing-mediated bacterial biofilm

- formation. *Journal of Pharmacy and Pharmacology*, **2015**, 67, 951-962.
- Selin, C. *et al.* A pipeline for screening small molecules with growth inhibitory activity against *Burkholderia cenocepacia*. *PloS one*, **2015**, 10, e0128587.
- Singh, A., Upadhyay, V., Upadhyay, A. K. *et al.* Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microb Cell Fact*, **2015**, 14, 41.
- Sperandio, D. *et al.* Cell-associated hemolysis activity in the clinical strain of *Pseudomonas fluorescens* MFN1032. *BMC Microbiol.* **10**, 124 (2010).
- Van Belkum, A. *et al.* Meropenem/colistin synergy testing for multidrug-resistant *Acinetobacter baumannii* strains by a two-dimensional gradient technique applicable in routine microbiology. *J Antimicrob Chemother.* **2015**, 70, 167–172.
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., Lee, S. Y., Fischbach, M. A., Müller, R., Wohlleben, W., Breitling, R., Takano, E., Medema, M. H. AntiSMASH 3.0 – a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **2015**, 43: W237-W243.
- Whelan, S., & Goldman, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular biology and evolution*, **2001**, 18(5), 691–699.
- Zerbino, D. R., Birney, E. Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 2008, 18, 821-829.
- Zhang, E. *et al.* Synthesis and bioactivities study of new antibacterial peptide mimics: The dialkyl cationic amphiphiles. *European Journal of Medicinal Chemistry* **143**, 1489–1509 (2018).



Zuckerlandl, E., Pauling, L. Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (Eds.) *Evolving Genes and Proteins*, Academic Press, New York, **1965**, p. 97-166.

## Chapter 3

### Identification of the UA gene cluster, its biosynthetic pathway and chemical synthesis of MPA and bioassays of MPA v/s UA.

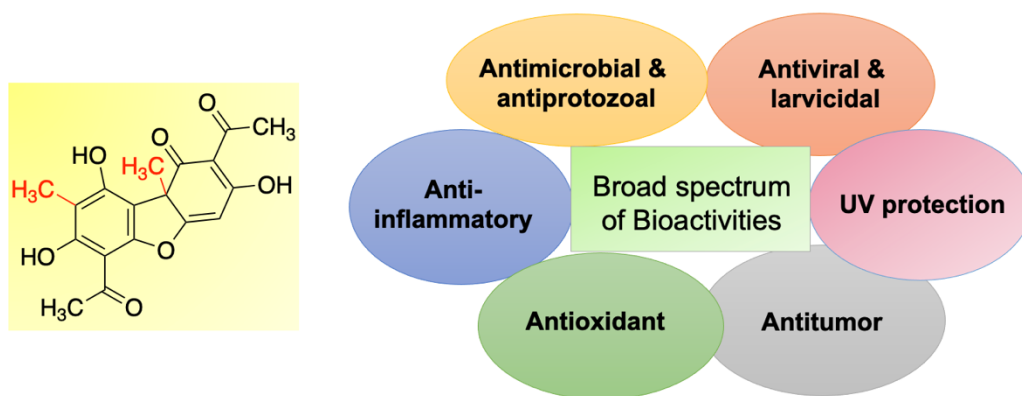
This chapter is already published. Sarkar R, Mittal N, Sorensen J, Sen T. A Comparison of the Bioactivity of UA versus Methylphloroacetophenone. *Natural Product Communications*. December 2018. doi:[10.1177/1934578X1801301224](https://doi.org/10.1177/1934578X1801301224). My contribution was the chemical synthesis of Methylphloroacetophenone (MPA).

#### 3.1. Introduction

UA is a yellow-coloured dibenzofuran, and it is one of the most well-studied lichen natural products so far. It is the first polyketide isolated from lichens in 1844 (**Figure 3.1**). UA is widely distributed amongst many species of lichens such as *Cladonia*, *Usnea*, *Lecanora*, *Evernia* and *Ramalina*, etc. (**Figure 3.1**). Many biological activities are associated with this molecule such as anti-bacterial, antitumor, antiviral, anti-inflammatory, antiprotozoal, larvicidal and treating tuberculosis (**Figure 3.1**) [Araújo et al. 2015]. The existence of such rich literature highlighting its pharmacological importance makes it an important candidate in the world of research on lichen NPs. Commercial availability of the chemical standard of UA at *Sigma-Aldrich* increases the tenability of UA biosynthesis research. The test organism that we chose to understand the science behind the biosynthesis of this unique lichen SM is *C. uncialis*. This lichen species can be found locally in Manitoba and UA is one of the main chemical constituents of this lichen species. Freshly field-collected *C. uncialis* was provided to our research group by Dr. Michelle Piercey-Normore, a lichenologist at the

Department of Biology, the University of Manitoba (Presently, Memorial University). Morphological identification was done by Dr. Michelle Piercey-Normore and genetic identification of this species was done by Dr. Abdel-Hameed (ex-PhD student in Sorensen's lab) by PCR amplification of the internal transcribed spacer (ITS) region and sequenced the amplicon using a conventional Sanger platform [Abdel-Hameed, M. et al., 2016]. In addition, previous *in situ* imaging of UA within the thallus of *C. uncialis* performed in collaboration with Dr. Michelle Piercey-Normore research group provided us with in-house expertise on stimulating UA biosynthesis in *C. uncialis* [Liao et al., 2010].

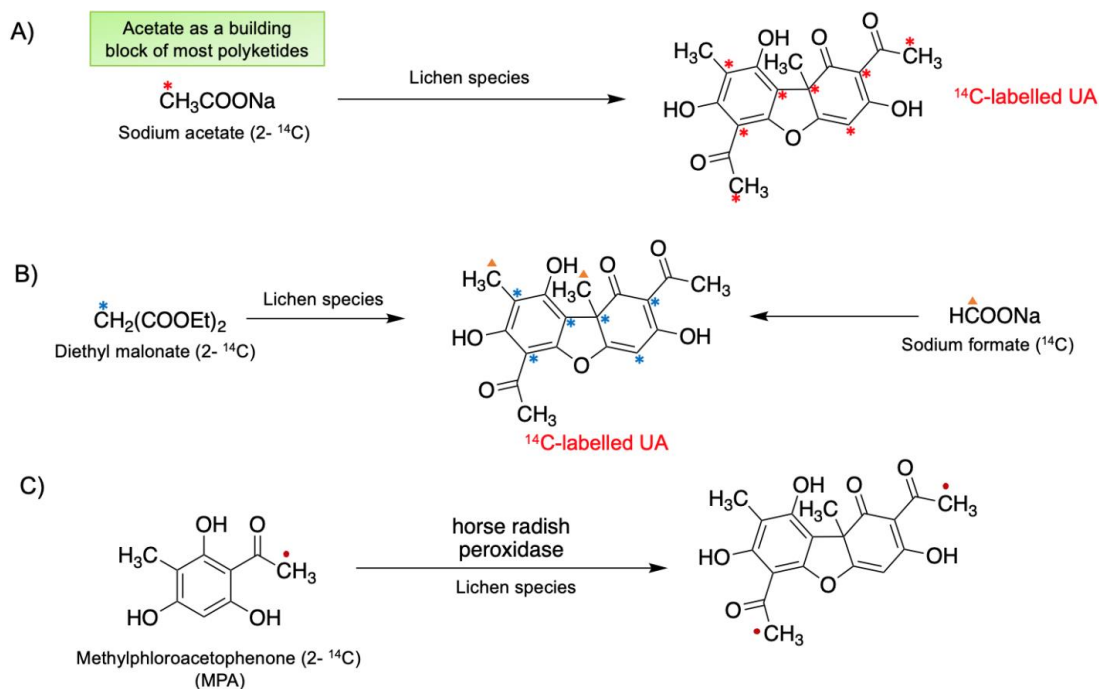
*C. uncialis* belongs to the order Lecanorales and the family Cladoniaceae. The thallus of this lichen is yellow-green or greenish, often brownish towards the pointed apices. *C. uncialis* frequently grows on soil among mosses, in well-lit and typically, dry places. It is common in coniferous forests, on heathland and on sand dunes in Europe, North America, and Asia [Purvis et al., 1992]. *C. uncialis* produces dibenzofurans such as UA and depsides: squamatic acid and thamnolic acid [Culberson 1969; Purvis et al., 1992; BeGora & Fahlset 2001].



**Figure 3.1:** Chemical structure of UA and its biological properties.

### 3.2. History of biosynthesis of UA

A biosynthetic scheme was proposed by Taguchi et al. in 1969, where feeding experiments of isotope-labelled sodium acetate and diethyl malonate were performed to four different lichen species, the resulting product enrichment pattern was consistent with UA biosynthesis (**Figure 3.2**). A  $^{14}\text{C}$ -labelled sodium formate feeding experiment confirms the incorporation of methyl groups into UA to be derived from a C1-unit (**Figure 3.2**).  $^{14}\text{C}$ -labelled MPA was also fed to lichen species and enriched UA was re-isolated, which ultimately demonstrated that UA is a dimer of two molecules of MPA (**Figure 3.2**). An *in vitro* oxidation experiment of MPA to UA using horseradish peroxidase also demonstrated the oxidative coupling of MPA to UA (**Figure 3.2**). On the contrary, when  $^{14}\text{C}$ -labelled phloroacetophenone was fed to lichen species, enriched UA was not isolated which demonstrated the addition of methyl group during polyketide chain elongation (**Figure 3.2**). The isotope-labelling experiments by Taguchi et al. (1969) also demonstrate that the PKS must also have a C-methyltransferase (cMeT) domain. We now know that S-adenosyl methionine substrate is required as a methyl donor for cMeT to catalyze the addition of methyl group to a polyketide chain. This work laid a foundation for the discovery of the categorical classification of UA-associated PKS as a non-reducing polyketide synthase (NR-PKS) [Taguchi et al., 1969].

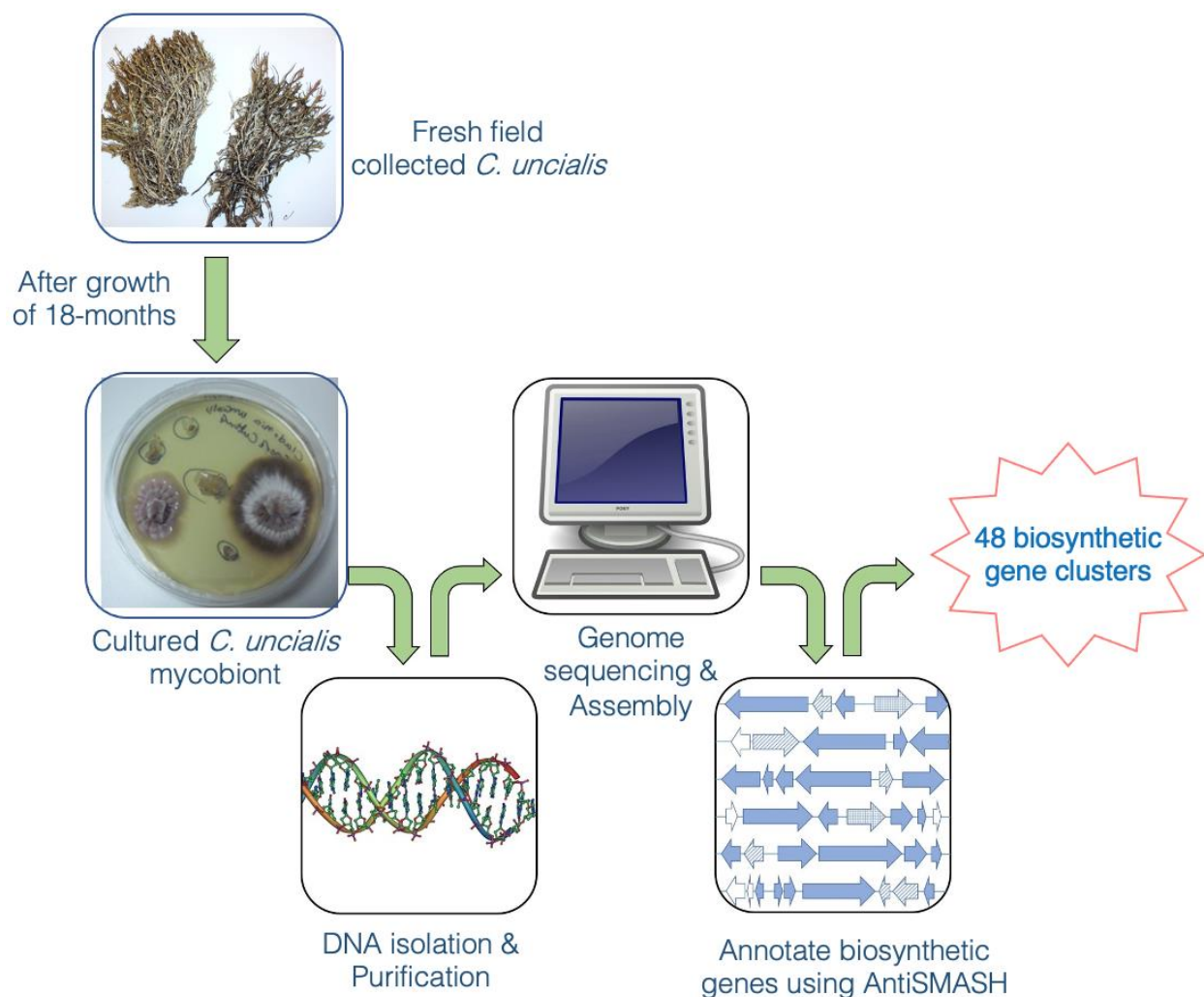


**Figure 3.2:** Isotope-labelled ( $^{14}C$ ) feeding experiments of lichens (Taguchi in 1969) & proposed UA biosynthetic pathway. A) Feeding experiment of  $^{14}C$ -labelled sodium acetate to different lichen species and enriched UA was isolated (\* =  $^{14}C$ ) B) Feeding experiment of  $^{14}C$ -labelled diethyl malonate and sodium formate to different lichen species and enriched UA was re-isolated (\* =  $^{14}C$  from diethyl malonate,  $\Delta$  =  $^{14}C$  from sodium formate) C) Both feeding experiment of  $^{14}C$ -labelled MPA to different lichen species and its oxidation to  $^{14}C$ -labelled UA using horse-radish peroxidase.

### 3.3. Steps towards genome annotation of *C. uncialis*

The methodology that was adopted to annotate biosynthetic genes found in the genome of *C. uncialis* [Abdel-Hameed et al. 2016] is as follows: **1)** Spores from *C. uncialis* were collected onto a water-agar media, then transferred to nutrient-rich agar and to generate enough biomass for *de novo* whole-genome sequencing, this fungus was grown on agar for one year. As mentioned earlier in **Chapter 1**, lichen is a symbiont of a fungus and algae or cyanobacterium. Although the photosynthetic partner is known to produce a few of the SMs arising from the lichen partnership [Kassalainen et al. 2012, Kampa et al. 2013], the fungal partner is responsible for most of the lichen secondary metabolome. Therefore, it was imperative to grow the fungal partner separately (an axenic culture) from the photosynthetic partner in the lab. **2)** The genomic DNA of *C. uncialis* fungus was isolated and sequenced using an Illumina platform, then assembled using

SPAdes [Kearse et al., 2012]. **3)** Contigs generated from SPAdes were submitted to the Antibiotic and Secondary Metabolite Analysis Shell (AntiSMASH V. 2.0) [Blin et al., 2013]. This program annotates biosynthetic genes and when motifs indicative of a biosynthetic gene (e.g., a polyketide synthase) are discovered, the program then expands its analysis of the nucleotides flanking the biosynthetic gene to search for accessory genes (e.g., p450, *O*-methyltransferase, halogenase) [Medema et al., 2011; Weber et al., 2015; Blin et al., 2017] (**Figure 3.3**). As described earlier in **Chapter 1**, a combination of a biosynthetic gene that encodes for a backbone enzyme (e.g., a polyketide synthase or a terpene cyclase) and one or more accessory genes encoding for accessory/tailoring enzymes makes a biosynthetic gene cluster (BGC). And a total of 48 biosynthetic gene clusters were found in the genome *C. uncialis* (**Figure 3.3**).



**Figure 3.3:** Schematic representation of *C. uncialis* genome sequencing, assembly, and annotation of biosynthetic genes [Abdel-Hameed, M. et al. 2016].

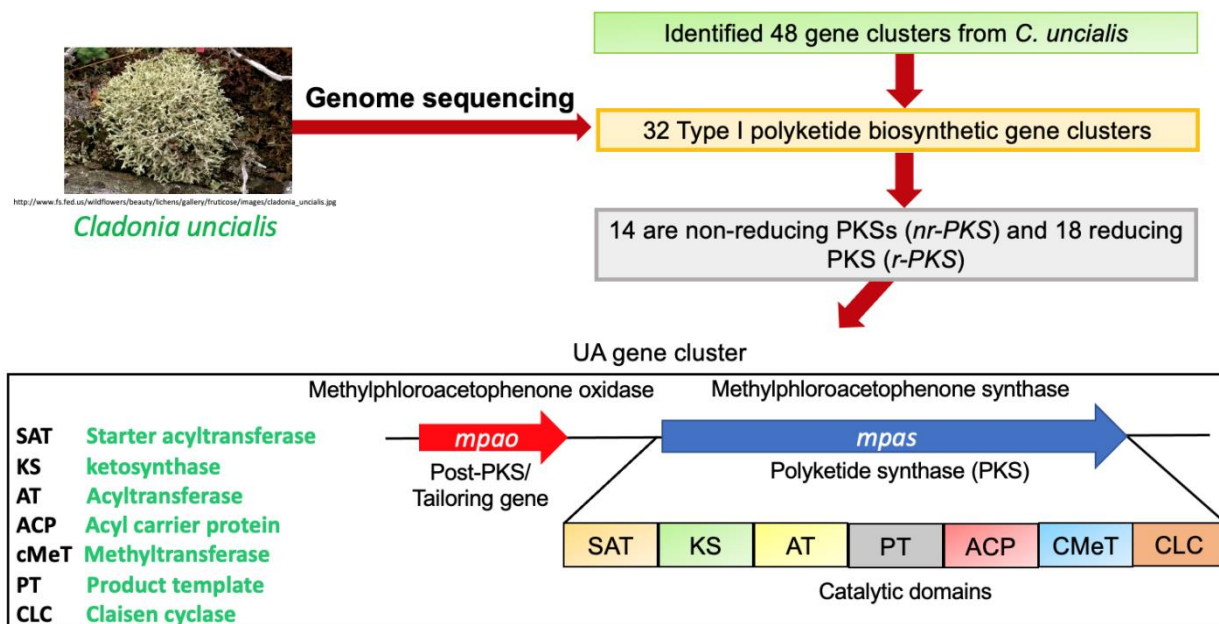
### 3.4. Identification of UA gene cluster

The first successful gene architecture of UA-associated *nr-pks* and UA biosynthetic gene clusters was proposed by our research group by carrying out an extensive *C. uncialis* genome sequencing and gene knockout experiments [Abdel-Hameed, M. et al., 2016]. A total of 32 types I PKS genes (**Figure 3.4**) were identified within the assembled contigs of 48 annotated gene clusters using AntiSMASH V. 2.0. (Described above in **Section 3.4**).

Each identified gene was also submitted to the Basic Local Alignment Search Tool (BLAST) [Altschul et al., 1990] and the probable function of each gene was noted based on

consensus similarity to characterized genes deposited in GenBank. Earlier mentioned isotope-labelling experiments [Taguchi et al., 1969] also demonstrated that UA must be produced by a non-reducing PKS possessing methyltransferase (cMeT) and methyl group introduces before cyclization of polyketide chain. This cyclization is catalyzed by the Claisen cyclase (CLC) domain via Claisen condensation instead of conventional aldol cyclization catalyzed by the TE domain (explained in detail in **Chapter 1**). A single oxidative enzyme, most probably a cytochrome p450, should be found near the PKS and known as a post-PKS/tailoring enzyme (**Chapter 1**) [Bernhardt 2006; Guengerich & Munro, 2013; Isin and Guengerich 2007]. Taking into the account presence/absence of reducing domains (**Chapter 1**), out of 32 PKS genes, 14 genes were found to be non-reducing and 18 were reducing. Of these 14 remaining non-reducing genes, only five had the requisite cMeT domain. Out of these 5 genes, only one was consistent with UA biosynthesis with regards to the presence of the CLC domain and one tailoring gene, a cytochrome p450 (**Figure 3.4**). Domain architecture of UA *nr-pks* gene (**Figure 3.4**): ACP trans-acylase domain (SAT), this advanced priming unit starts the polyketide chain extension, acetyl transferase domain (AT), acyl-carrier protein (ACP), keto-synthase domain (KS), methyltransferase (cMeT), product template domain (PT) and Claisen cyclase (CLC) domain and function of each domain has been elaborated in **Chapter 1**. This *nr-pks* gene was named by Bertrand & Abdel-Hameed [Abdel-Hameed et al. 2016] as *mpas* (methylphloroacetophenone synthase) encodes for MPAS enzyme and the cytochrome p450 as *mpao* (methylphloroacetophenone oxidase) encodes MPAO enzyme (**Figure 3.4**).

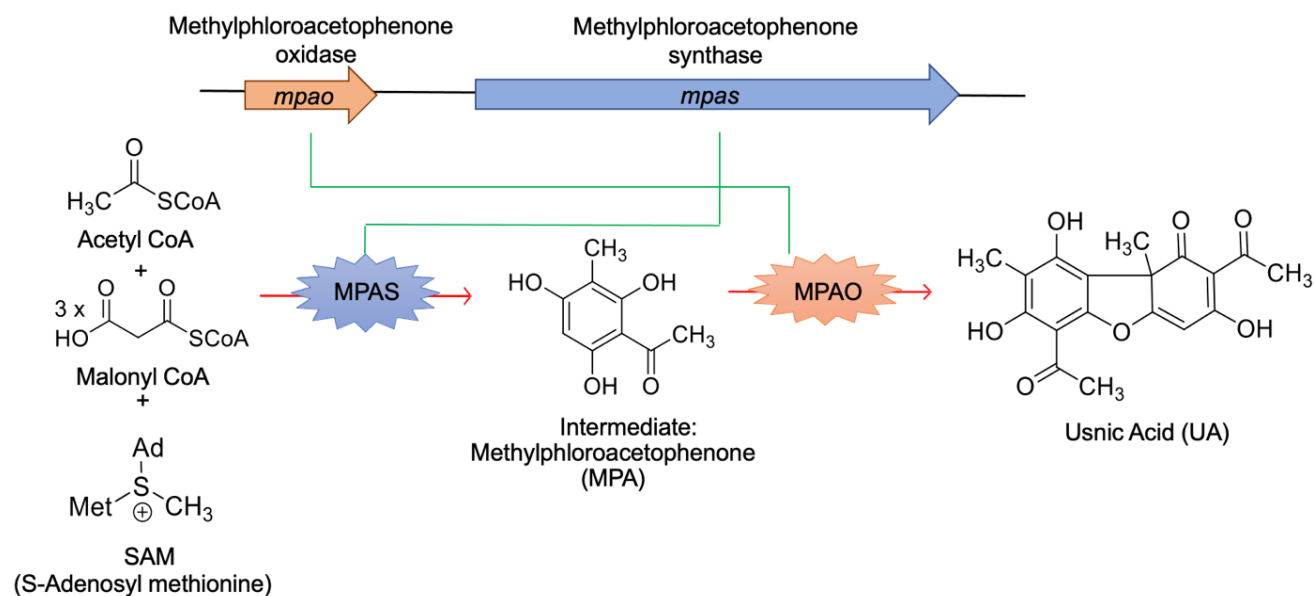




**Figure 3.4:** Gene-knockout experiment for the identification of UA gene clusters after *de-novo* genome sequencing of *C. uncialis*, gene architecture of UA gene cluster and domain architecture of *mpas* (Abdel-Hameed, M. *et al.* 2016).

### 3.5. Biosynthesis of UA

The biosynthesis of UA has been demonstrated [Taguchi *et al.*, 1969; Abdel-Hameed *et al.*, 2016] to proceed via the intermediate MPA (**Figure 3.5**) which is produced from acetyl-CoA, malonyl-CoA and *S*-adenosylmethionine (SAM) by an MPAS. The intermediate MPA is in turn dimerized by an MPAO enzyme to produce UA (**Figure 3.5**).



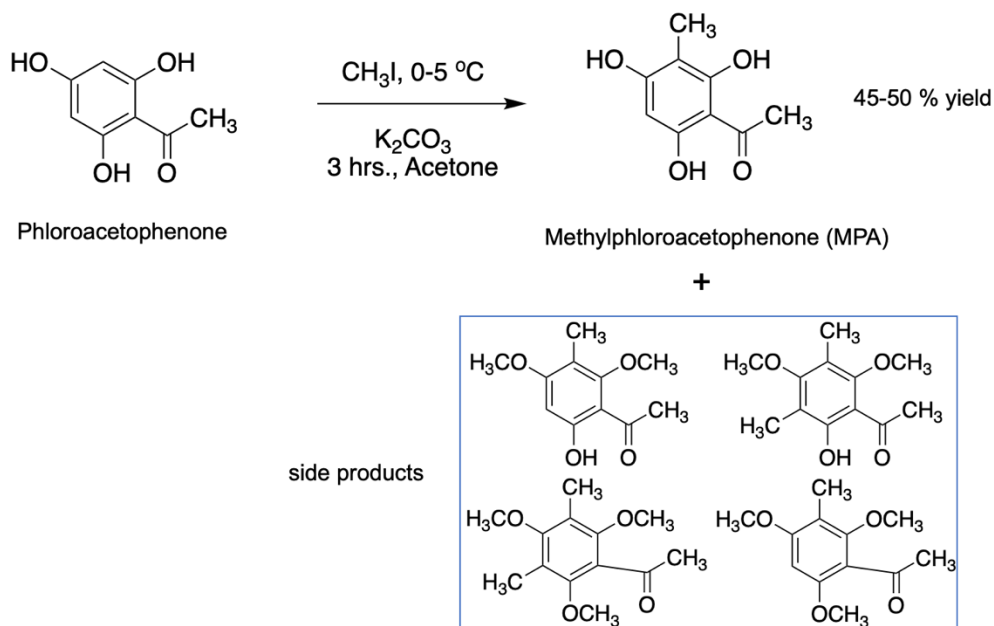
**Figure 3.5:** A schematic outline of biosynthesis of UA catalyzed MPAS and MPAO.

Experiments of functional heterologous expression of MPAS in both fungal and bacterial hosts, were performed by Dr. Robert Bertrand (ex-PhD student). Unfortunately no functional heterologous expression of MPAS could be detected.

### 3.6. Chemical synthesis of MPA

The first key experiment was the chemical synthesis of MPA. Both phloroacetophenone (substrate) and potassium carbonate ( $K_2CO_3$ , a Lewis base) were dissolved in acetone at r. t. This reaction mixture was kept in an ice bath to bring the temperature between 0 to 5 °C. As a methylated reagent, 3 equivalents of iodomethane were added dropwise for 3 h. 1 equivalent at the start, 2<sup>nd</sup> equivalent after 1 h and the 3<sup>rd</sup> equivalent after 2 h of reaction time. The reaction mixture was stirred for 1 hour and then quenched with 1 M HCl. The product was extracted using ethyl acetate and was stored at 4 °C. The reaction resulted in many O-methylated side products along with the synthesis of MPA (**Scheme 3.1**). The crude product was purified with flash column chromatography and the final yield of MPA was 39% and the yield was improved up to 48% when purified using automated flash chromatography (*Biotage*). As already mentioned in **Chapter**

2, that this reaction was already published by [Hawranik, D. J. *et al.* in 2009], and we only did some modifications to increase the % yield and reduce the number of O-methylated side products.



**Scheme 3.1:** Chemical synthesis of MPA. K<sub>2</sub>CO<sub>3</sub>: Potassium carbonate and CH<sub>3</sub>I: Iodomethane.

### 3.7. Bio-film disruption activity, MPA v/s UA

This work was done in collaboration with the late Dr. Tuhinadri Sen and his research group at the Department of Pharmaceutical Technology, Jadavpur University, Kolkata, India. UA has been shown to possess antimicrobial activity against several planktonic gram-positive bacteria, including *Staphylococcus aureus*, *Enterococcus faecalis*, and *Enterococcus faecium*. Although the exact mechanism of action is still unknown [Shibata et al., 1948; Lauterwein, M. et al., 1995] there is some evidence that the antibacterial activity of UA is due to the inhibition of RNA transcription [Campanella, L. et al., 2002]. In addition, it appears that UA can also affect biofilm formation in organisms such as *Staphylococcus aureus* and *Pseudomonas aeruginosa* [Francolini, I. et al., 2004]. Opportunistic human pathogens such as *Pseudomonas aeruginosa* can sense their population density by using an intercellular signalling system known as quorum-sensing (QS),

which in turn regulates a diverse array of physiological activities such as biofilm formation and the production of virulence factors [Müh, U. et al., 2006; Fuqua, C., 2002]. Opportunistic pathogens embed themselves in an extracellular polymeric matrix that forms the basis of the biofilm, which can contribute to bacterial resistance and pathogenicity [O'Brien, K. T. et al., 2016]. Targeting the quorum-sensing circuit in *P. aeruginosa* could be a useful approach to preventing bacterial pathogenesis. This approach could render pathogenic bacteria non-virulent without affecting their viability, which will result in less pressure to develop resistance as is observed in antibiotic therapy [Sarabhai, S. et al., 2013; Welsh, M. A. et al., 2015]. Recently natural products have become of interest as a source of QS inhibitors. For example, ajoene from *Allium sativum* [Jakobsen, T. H. et al., 2012], curcumin, from *Curcuma longa* [Rudrappa, T., et al., 2008], iberin from *Armoracia rusticana* [Jakobsen, T. H. et al., 2012] have all been shown to interrupt the bacterial QS signalling system. In addition, methanolic extracts from plants used in traditional medicine applications have been also shown to disrupt the formation of biofilms [Sarkar, R. et al., 2014; Sen, T. et al., 2015]. There has been some recent interest in exploring lichen natural products and their ability to disrupt biofilm formation [Singh, N. et al., 2015]. There has been a demonstration of the potential application of UA in thin-film coatings for this purpose [Grumezescu, V. et al., 2014]. However, little is known about the potential of the biosynthetic intermediate MPA for disruption of biofilm formation. One possible explanation for the conversion of MPA into UA in a lichen is that the latter may have increased biological activity, providing greater protection from bacterial infection for the slow-growing lichen. Therefore, we set out to compare the bioactivity, specifically concerning biofilm formation, of UA and the biosynthetic precursor MPA. It is interesting to note that there have been no reports of the isolation of MPA as a natural product, suggesting that MPA is rather efficiently converted into UA in lichen species.

We demonstrate here, for the first time, that in our experiments UA displays more biological activity than MPA.

The presence of UA in some lichen species may offer some protective effect from colonization by bacteria through the mechanism of disrupting biofilm formation and the quorum sensing pathways. UA is known to inhibit the growth of numerous Gram-positive bacteria, while it has no effect on Gram-negative microorganisms [Maciąg-Dorszyńska, M. et al., 2014; Lauterwein, M. et al., 1995], but interestingly in our present investigation UA was found to exhibit strain-specific antibacterial activity against both Gram-positive and Gram-negative standard reference strains at concentrations in the 7.8–500 µg/mL range (**Table 3.1**).

**Table 3.1:** Assessment of minimum inhibitory concentration (MIC) of UA and MPA against different bacterial strains.

Micro-organisms	Agents	MIC (µg/mL)
<i>Escherichia coli</i> MTCC 2939	UA	7.8
	MPA	250
<i>Pseudomonas aeruginosa</i> MTCC 2453	UA	500
	MPA	>500
<i>Staphylococcus aureus</i> MTCC 96	UA	150
	MPA	500
<i>Bacillus subtilis</i> MTCC 441	UA	25
	MPA	500

UA has been found to be more active as compared to MPA against all tested reference strains. Pyoverdine and Pyocyanin are virulence factors regulated by QS and play an important role in bacterial invasion of the host cell. Pyoverdine promotes microbial growth by competing for iron with mammalian transferrin, causing iron deficiency in the host cells with subsequent improvement of iron availability in the bacterial cells [Sarkar, R. et al., 2014; Adonizio, A. et al., 2008]. Pyocyanin is functioning as a redox-active toxin and promoting extracellular DNA (eDNA)

release for biofilm formation [Das, T. et al., 2012; Chang, C. Y. et al., 2014]. UA and MPA have significantly inhibited the production of these virulence factors in a dose-dependent manner (**Table 3.2**). UA was found to be more active than its intermediate at any given concentration. Bacterial motility is one of the major QS regulated factors responsible for biofilm formation [Sarkar, R. et al., 2014]. Flagellum-dependent swimming, swarming motility, and type IV pilus-based twitching motility are required for the initial attachment, development of a biofilm by *P. aeruginosa* and for the overproduction of virulence factors [de la Fuente-Núñez, C. et al., 2012; Déziel, E. et al., 2002].

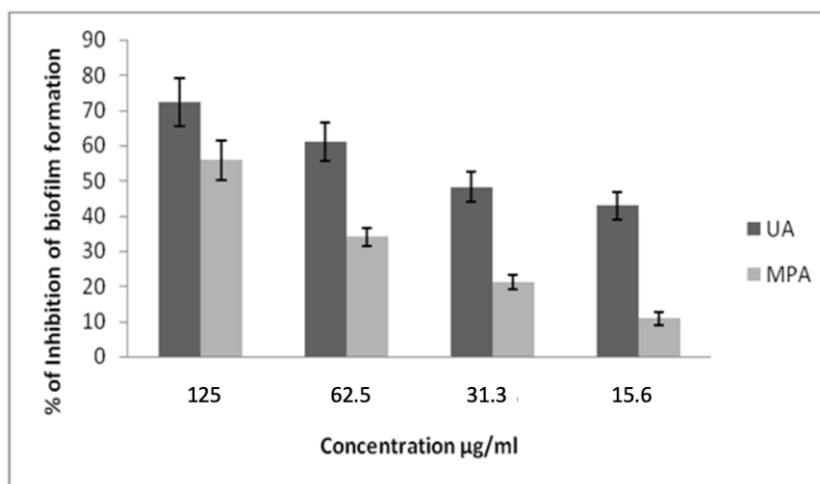
**Table 2.2:** Pyoverdine and Pyocyanin production in *Pseudomonas aeruginosa*, grown in presence or absence of UA and MPA. The fluorescence emission spectrum was recorded by exciting at 405 nm. The activity was expressed in relative fluorescence units for Pyoverdine production. Salicylic acid was used as a positive control. Values are expressed as mean  $\pm$  SEM (standard error of the mean).; (n = 3); \*P < 0.05 (vs control), variable were compared using the *t* test.

Treatment	Concentration ( $\mu\text{g/mL}$ )	Pyoverdine production		Pyocyanin production	
		RFU (Excitation 405 nm &, Emission 465 nm)	% Inhibition	Absorbance at 520 nm	% Inhibition
Control		9600 $\pm$ 100	0	0.408 $\pm$ 0.006	0
UA	31.3	3400 $\pm$ 200*	63.9	0.230 $\pm$ 0.01*	43.9
	15.6	5600 $\pm$ 100*	41.4	0.267 $\pm$ 0.017*	34.6
	7.8	6900 $\pm$ 300*	27.8	0.316 $\pm$ 0.021*	22.5
MPA	31.2	4200 $\pm$ 200*	55.9	0.259 $\pm$ 0.011*	36.5
	15.6	6900 $\pm$ 200*	28.0	0.312 $\pm$ 0.013*	23.5
	7.8	8100 $\pm$ 300*	15.2	0.352 $\pm$ 0.023*	13.7
Salicylic acid	50	4400 $\pm$ 100*	54.4	0.265 $\pm$ 0.007*	35.1

UA and MPA inhibited swimming, swarming and twitching motility in a dose-dependent manner (**Table 3.3**). The effect of UA and MPA on bacterial biofilm formation was assessed by crystal violet assay. Both UA and MPA were found to inhibit bacterial biofilm formation in a dose-dependent manner (**Figure 3.6**). UA (31.25  $\mu\text{g/mL}$ ; 48.2% inhibition) was found to be more active as compared to MPA (31.25  $\mu\text{g/mL}$ : 21.1% inhibition).

**Table 3.3:** Swimming and twitching motility of *P. aeruginosa* on semi-solid agar in presence and absence of UA and MPA. Values are expressed as mean  $\pm$  SEM (standard error of the mean).; (n = 3); \*P < 0.05 (vs control).

Treatment	Concentration ( $\mu\text{g/mL}$ )	Distance of migration (mm)		
		Swimming motility	Swarming motility	Twitching motility
Control		31.2 $\pm$ 2.6	16 $\pm$ 2.1	23 $\pm$ 1.9
UA	15.6	14 $\pm$ 1.2*	6 $\pm$ 0.4*	5 $\pm$ 0.5*
	7.8	19 $\pm$ 1.4*	10 $\pm$ 0.8*	7 $\pm$ 0.9*
MPA	15.6	20 $\pm$ 1.3*	9 $\pm$ 0.5*	15 $\pm$ 1.1*
	7.8	23 $\pm$ 1.8*	12 $\pm$ 0.7*	20 $\pm$ 1.2*



**Figure 3.6:** Effect of UA and MPA on biofilm formation. Biofilm formation was assessed by a colorimetric technique using crystal violet. The results are expressed as a percentage of biofilm inhibition for untreated control. Values are expressed as mean  $\pm$  SEM (standard error of the mean).; n = 6.

Our present investigation suggests that both UA and its intermediate MPA inhibit bacterial biofilm formation and virulence factor production and may be by interfere with the bacterial QS signalling system.

### **3.8. Summary**

In this study, we have outlined the details of *the C. uncialis* genome sequencing project resulted in the identification of 48 gene clusters [Abdel-Hameed, M. 2014 & Bertrand, R. L. 2019]. Description of the methodology of identification of the UA gene cluster & the proposed biosynthesis of UA. We have also discussed the chemical synthesis of MPA (an intermediate during UA biosynthesis). Bioactivity of MPA & UA was compared by doing the biofilm-disruption & antibacterial assays. In conclusion, we have demonstrated that UA is more biologically active than MPA in the whole-cell and biofilm disruption assays. This suggests a rationale for why lichen fungi exert the effort required to produce UA. This knowledge may help guide the discovery of other natural products that disrupt biofilm production.





- Blin, K., Wolf, T., Chevrette, M. G., Lu, X., Schwalen, C. J., Kautsar, S. A., Suarez-Duran, H. G., de los Santos, E. L., Kim, H. U., Nave, M., Dickschat, J. S., Mitchell, D. A., Shelest, E., Breitling, R., Takano, E., Lee, S. Y., Weber, T., Medema, M. H. AntiSMASH 4.0 – Improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **2017**, 45: W36-W41.
- Campanella, L., Delfini, M., Ercole, P., Iacoangeli, A., Risuleo, G. Molecular characterization and action of usnic acid: a drug that inhibits proliferation of mouse polyomavirus in vitro and whose main target is RNA transcription, *Biochimie*, **2002**, 84, 329-334.
- Chang CY, Krishnan T, Wang H, Chen Y, Yin WF, Chong YM, Tan LY, Chong TM, Chan KG. Non-antibiotic quorum sensing inhibitors acting against N-acyl homoserine lactone synthase as druggable target. *Scientific Reports*, **2014**, 4, 7245.
- Culberson, L. W. The use of chemistry in the systematics of the lichens. *Taxon*, **1969**, Vol. 18 (2), 152-166.
- Das T, Manefield M. Pyocyanin promotes extracellular DNA release in *Pseudomonas aeruginosa*. *PLoS ONE*, **2012**, 7, e46718.
- de la Fuente-Núñez C, Korolik V, Bains M, Nguyen U, Breidenstein EB, Horsman S, Lewenza S, Burrows L, Hancock RE. Inhibition of bacterial biofilm formation and swarming motility by a small synthetic cationic peptide. *Antimicrobial Agents and Chemotherapy*, **2012**, 56, 2696-2704.
- Déziel E, Comeau Y, Villemur R. (2002) Initiation of biofilm formation by *Pseudomonas aeruginosa* 57RP correlates with emergence of hyperpiliated and highly adherent phenotypic variants deficient in swimming, swarming, and twitching motilities. *Journal of Bacteriology*, **283**, 1195-1204.

- Francolini I, Norris P, Piozzi A, Donelli G, Stoodley P. Usnic acid, a natural antimicrobial agent able to inhibit bacterial biofilm formation on polymer surfaces. *Antimicrobial Agents and Chemotherapy*, **2004**, 48, 4360-4365.
- Fuqua C, Greenberg EP. Listening in on bacteria: acyl-homoserine lactone signalling. *Nature Reviews Molecular Cell Biology*, **2002**, 3, 685-695.
- Guengerich FP, Munro AW. Unusual cytochrome p450 enzymes and reactions. *J. Biol. Chem.* **2013**, 288, 17065-17073.
- Hawranik DJ, Anderson KS, Simmonds R, Sorensen JL. The chemoenzymatic synthesis of usnic acid. *Bioorganic & Medicinal Chemistry Letters*. **2009**, 9(9), 2383-2385.
- Isin EM, Guengerich FP. Complex reactions catalyzed by cytochrome p450 enzymes. *Biochim. Biophys. Acta* **2007**, 1770, 314-329.
- Jakobsen TH, Bragason SK, Phipps RK, Christensen LD, van Gennip M, Alhede M, Skindersoe M, Larsen TO, Høiby N, Bjarnsholt T, Givskov M. Food as a source for quorum sensing inhibitors: iberin from horseradish revealed as a quorum sensing inhibitor of *Pseudomonas aeruginosa*. *Applied and Environmental Microbiology*, **2012**, 78, 2410-2421.
- Jakobsen TH, van Gennip M, Phipps RK, Shanmugham MS, Christensen LD, Alhede M, Skindersoe ME, Rasmussen TB, Friedrich K, Uthe F, Jensen PØ, Moser C, Nielsen KF, Eberl L, Larsen TO, Tanner D, Høiby N, Bjarnsholt T, Givskov M. Ajoene, a sulfur-rich molecule from garlic, inhibits genes controlled by quorum sensing. *Antimicrobial Agents and Chemotherapy*, **2012**, 56, 2314-2325.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentjies P, Drummond A. Geneious basic:

An integrated and extendable desktop software program for the organization and analysis of sequence data. *Bioinformatics* **2012**, 28, 1647-1649.

Lauterwein M, Oethinger M, Belsner K, Peters T, Marre R. *In vitro* activities of the lichen secondary metabolites vulpinic acid, (+)-usnic acid, and (-)-usnic acid against aerobic and anaerobic microorganisms. *Antimicrobial Agents and Chemotherapy*, **1995**, 39, 2541-2543.

Liao C, Piercey-Normore MD, Sorensen JL, Gough K. *In situ* imaging of usnic acid in selected *Cladonia* spp. by vibrational spectroscopy. *Analyst* **2010**, 135, 3242-3248.

Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, de Bruijn I, Chooi YH, Claesen J, Coates RC, Cruz-Morales P, Duddela S, Düsterhus S, Edwards DJ, Fewer DP, Garg N, Geiger C, Gomez-Escribano JP, Greule A, Hadjithomas M, Haines AS, Helfrich EJ, Hillwig ML, Ishida K, Jones AC, Jones CS, Jungmann K, Kegler C, Kim HU, Kötter P, Krug D, Masschelein J, Melnik AV, Mantovani SM, Monroe EA, Moore M, Moss N, Nützmann HW, Pan G, Pati A, Petras D, Reen FJ, Rosconi F, Rui Z, Tian Z, Tobias NJ, Tsunematsu Y, Wiemann P, Wyckoff E, Yan X, Yim G, Yu F, Xie Y, Aigle B, Apel AK, Balibar CJ, Balskus EP, Barona-Gómez F, Bechthold A, Bode HB, Borriss R, Brady SF, Brakhage AA, Caffrey P, Cheng YQ, Clardy J, Cox RJ, De Mot R, Donadio S, Donia MS, van der Donk WA, Dorrestein PC, Doyle S, Driessen AJ, Ehling-Schulz M, Entian KD, Fischbach MA, Gerwick L, Gerwick WH, Gross H, Gust B, Hertweck C, Höfte M, Jensen SE, Ju J, Katz L, Kaysser L, Klassen JL, Keller NP, Kormanec J, Kuipers OP, Kuzuyama T, Kyrpides NC, Kwon HJ, Lautru S, Lavigne R, Lee CY, Linquan B, Liu X, Liu W, Luzhetskyy A, Mahmud T, Mast Y, Méndez C, Metsä-Ketelä M, Micklefield J, Mitchell DA, Moore BS, Moreira LM, Müller R, Neilan BA, Nett M, Nielsen J, O'Gara F, Oikawa

- H, Osbourn A, Osburne MS, Ostash B, Payne SM, Pernodet JL, Petricek M, Piel J, Ploux O, Raaijmakers JM, Salas JA, Schmitt EK, Scott B, Seipke RF, Shen B, Sherman DH, Sivonen K, Smanski MJ, Sosio M, Stegmann E, Süssmuth RD, Tahlan K, Thomas CM, Tang Y, Truman AW, Viaud M, Walton JD, Walsh CT, Weber T, van Wezel GP, Wilkinson B, Willey JM, Wohlleben W, Wright GD, Ziemert N, Zhang C, Zotchev SB, Breitling R, Takano E, Glöckner FO. Minimum information about a biosynthetic gene cluster. *Nat. Chem. Biol.* **2015**, 11, 625-631.
- Mona Abdel-Hameed, Robert L. Bertrand, Michele D. Piercey-Normore, John L. Sorensen, Putative identification of the usnic acid biosynthetic gene cluster by de novo whole-genome sequencing of a lichen-forming fungus, *Fungal Biology*, Volume 120, Issue 3, **2016**, 306-316.
- Müh U, Schuster M, Heim R, Singh A, Olson ER, Greenberg EP. Novel *Pseudomonas aeruginosa* quorum-sensing inhibitors identified in an ultra-high-throughput screen. *Antimicrobial Agents and Chemotherapy*, **2006**, 50, 3674-3679.
- O'Brien KT, Noto JG, Nichols-O'Neill L, Perez LJ. Potent irreversible inhibitors of LasR Quorum sensing in *Pseudomonas aeruginosa*. *ACS Medicinal Chemistry Letters*, **2016**, 6, 162–167.
- Purvis OW, COPPINS BJ, Hawksworth DL, JAMES PW, Moore DM, eds. The Lichen Flora of Great Britain and Ireland. London: *The Natural History Museum*, **1992**.
- Rudrappa T, Bais HP. Curcumin, a known phenolic from *Curcuma longa*, attenuates the virulence of *Pseudomonas aeruginosa* PAO1 in whole plant and animal pathogenicity models. *Journal of Agricultural and Food Chemistry*, **2008**, 56, 1955-1962.

- Sarabhai S, Sharma P, Capalash N. Ellagic acid derivatives from *Terminalia chebula* Retz. downregulate the expression of quorum sensing genes to attenuate *Pseudomonas aeruginosa* PAO1 virulence. *PLoS One*, **2013**, 8, e53441.
- Sarkar R, Chaudhary SK, Sharma A, Yadav KK, Nema NK, Sekhoacha M, Karmakar S, Braga FC, Matsabisa MG, Mukherjee PK, Sen T. Anti-biofilm activity of Marula - a study with the standardized bark extract. *Journal of Ethnopharmacology*, **2014**, 154, 170-175.
- Sen T, Karmakar S, Sarkar R. Evaluation of natural products against biofilm-mediated bacterial resistance. [ed.] PK Mukherjee. Evidence- Based Validation of Herbal Medicine. Amsterdam: *Elsevier*, **2015**, 321-338.
- Shibata S, Ukita T, Tamura T, Miura Y. Relation between chemical constitutions and antibacterial effects of usnic acid and derivatives. *The Japanese Medical Journal*, **1948**, 1, 152-155
- Singh N, Prateeksha, Pandey G, Jadaun V, Singh S, Bajpai R, Nayaka S, Naqvai A, Rawat A, Upreti D, Singh B. Development and characterization of a novel Swarna-based herbo-metallic colloidal nanoformulation – inhibitor of *Streptococcus mutans* quorum sensing *RSC Advances* **2015**, 5, 5809–5822.
- Grumezescu V, Socol G, Grumezescu, A, Holban A, Ficai A, Trusca R, Bleotu C, Balaure P, Cristescu R, Chifiriuc M. Functionalized antibiofilm thin coatings based on PLA–PVAmicrospheres loaded with usnic acid natural compounds fabricated by MAPLE *Applied Surface Science* **2014**, 302, 262-267.
- Taguchi H, Sankawa U, Shibata S. Biosynthesis of natural products. VI. Biosynthesis of usnic acid in lichens. A general scheme of biosynthesis of usnic acid. *Chem. Pharm. Bull.* **1969**, 17, 2054-2060.

- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Brucocoleri, R., Lee, S. Y., Fischbach, M. A., Müller, R., Wohlleben, W., Breitling, R., Takano, E., Medema, M. H. AntiSMASH 3.0 – a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **2015**, 43, W237-W243.
- Welsh, M. A., Eibergen, N. R., Moore, J. D., Blackwell, H. E. Small molecule disruption of Quorum sensing cross-regulation in *Pseudomonas aeruginosa* causes major and unexpected alterations to virulence phenotypes. *Journal of the American Chemical Society*, **2015**, 137, 1510–1519.

## Chapter 4

### Synthesis of novel MPA analogues, acylated phloroglucinols and their antibiotic activity against methicillin-resistant *Staphylococcus aureus*

This work was done and published in collaboration with Dr. Silvia Cardona research group at the Department of Microbiology, University of Manitoba, Winnipeg, Canada. Where I (Navriti Mittal<sup>1</sup>) contributed towards the chemical synthesis of Compounds 1-12, and bioassays were carried out by two students, Haben H. Tesfu<sup>2</sup> and Andrew M. Hogan<sup>3</sup>. Mittal, N., Tesfu, H.H., Hogan, A.M. *et al.* Synthesis and antibiotic activity of novel acylated phloroglucinol compounds against methicillin-resistant *Staphylococcus aureus*. *J Antibiot.* **72**, 253-259 (2019).

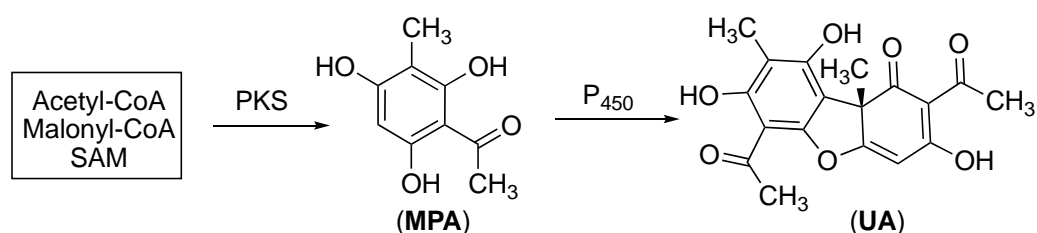
<https://doi-org.uml.idm.oclc.org/10.1038/s41429-019-0153-4>

#### **4.1. Introduction**

The emergence of multidrug resistant bacterial infections is a major health threat [Munita, J. M. et al., 2017]. Agencies like the World Health Organization and the Public Health Agency of Canada [2017 Report] have highlighted the need for investment in research and development of new antibiotics. A remarkable example is the Gram-positive bacterium *Staphylococcus aureus*, a ubiquitous commensal of the human skin. Naturally susceptible to most antibiotics, *S. aureus* can cause infections in predisposed individuals and can acquire formidable antibiotic resistance [Chambers, H. F. et al., 2009]. One of the most notorious antibiotic-resistant members of the species is methicillin-resistant *S. aureus* (MRSA) [Stryjewski, M. E. & Corey, G. R. 2014], which are grouped into two categories referred as hospital-associated (HA-MRSA) and community-associated (CA-MRSA) [Miller, L. G. & Kaplan, S. L. 2009]. MRSA is also a concerning pathogen



for people with the genetic disease cystic fibrosis (CF) [Akil, N. & Muhlebach, M. S. 2016]. Studies show that lower airway inflammation in children with CF is associated with early colonization with *S. aureus* [Gangell, C. 2011] and that the prevalence of MRSA in CF patients is high [Pena Amaya, P. et al., 2017; Muhlebach, M. S. 2017]. Microorganisms such as bacteria and fungi produce chemical compounds that have been successfully used as therapeutic antibiotics [Harvey, A. L. et al., 2015]. Lichenized fungi are a rich source of mostly under-explored natural products, many of which have chemical structures that suggest an origin from acetate via the polyketide pathway. As already discussed in **Chapter 3**, UA is one of the most widely occurring lichen natural products, and the biological activity of this polyketide has been extensively studied. **Chapter 3** also discusses the biosynthesis of UA, which has been demonstrated to proceed via the intermediate MPA, which is produced from acetyl-CoA, malonyl-CoA, and S-adenosylmethionine (SAM) by a polyketide synthase (PKS) [Heihachiro, T. et al., 1969] (**Figure 4.1**). The intermediate MPA is in turn dimerized by an oxidative p450-type enzyme to produce UA (**Figure 4.1**). It has been concluded in **Chapter 3** that UA is more bioactive than MPA based on the results of series of antibacterial and biofilm disruption assays.



**Figure 4.1:** The biosynthesis of UA from acetyl-CoA, malonyl-CoA and S-adenosylmethionine (SAM) has been demonstrated to proceed via MPA.

In this study, we systematically examined and explored other related phloracetophenone analogs including MPA analogs as potential source of novel bioactivity. Although MPA is a known natural product none of the compounds reported here appear to be of natural origin. We synthesized

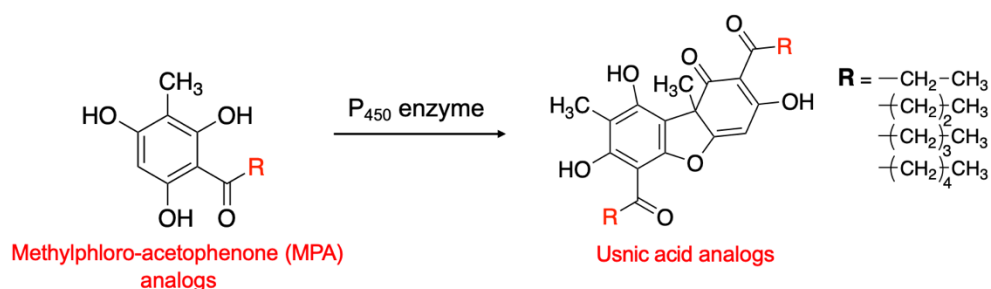
12 novel phloroacetophenone analogs (**Compounds 1–12, Schemes 4.2-4.4, Figure 4.3**) and explored their biological activity against *S. aureus* ATCC 29213 and MRSA strains isolated from CF patients. Our results show that 4 of the 12 novel compounds displayed promising activity as a monotherapy and one strongly synergized with doxycycline. These results suggest that phloroacetophenone analogs may serve as promising lead compounds for the design of novel antibiotics.

#### **4.2. Synthesis of phloroglucinol compounds**

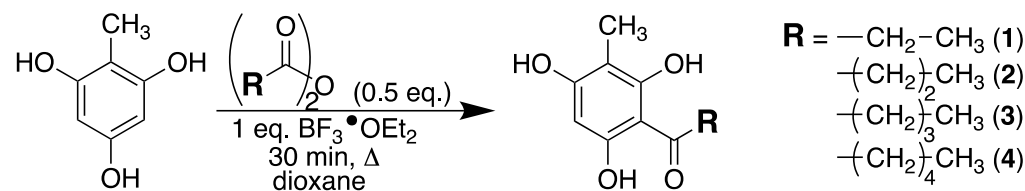
We first set out to synthesize a series of analogs of MPA where the methyl ketone had been replaced with a sequentially longer hydrocarbon chain of 3–6 carbon atoms in length. The rationale behind the synthesis of analogs of MPA was to ultimately biosynthesize the analogs of UA catalyzed by p450 type-enzyme (**Figure 4.2**). UA has been reported to show some side effects such as hepatotoxicity [García-Cortés, M. et al., 2016] when it was taken as a dietary supplement in higher concentrations and allergic contact dermatitis, sometimes accompanied by conjunctivitis [Ingólfssdóttir, K., 2002]. Therefore, synthesizing the analogs of UA with long hydrocarbon acyl side chains ultimately increase their lipophilicity and hence, the membrane permeability, could result in prevention of hepatotoxicity (**Figure 4.2**). However, other studies [Sahu et al., 2012] indicate that low nontoxic concentrations of UA do not cause damage to the liver.

The synthesis of compounds 1–4 is summarized in (**Scheme 4.1**). Commercially available trihydroxy-toluene was reacted with 0.5 equivalent of the appropriate anhydride in the presence of 1 equivalent of Lewis acid ( $\text{BF}_3 \cdot \text{OEt}_2$ ) using dioxane as the solvent. Using the anhydride as the limiting reagent resulted in a correspondingly low yield (18–40%) but prevented over acylation of the substrate. We also undertook the synthesis of a series of acylated phloroglucinols, compounds 5–8, as described in (**Scheme 4.2**). Phloroglucinol was dissolved in carbon disulfide and

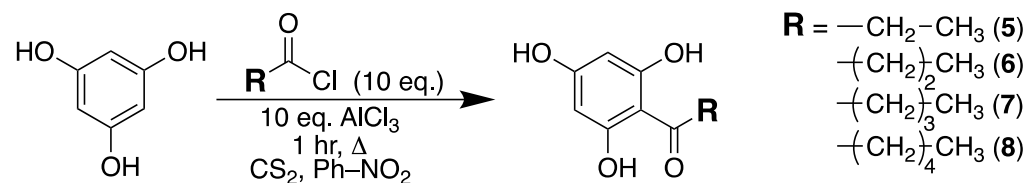
nitrobenzene and treated with 10 equivalents of the appropriate acid chloride using  $\text{AlCl}_3$  (10 equivalents) as catalyst. Heating for one hour followed by work-up afforded compounds 5–8 in a range of 41–53% yield. A series of diacylated phloroglucinol derivatives were synthesized as described in **Scheme 4.3**. Phloroglucinol was dissolved in dioxane and reacted with an excess of the appropriate anhydride (10 equivalents) and an excess (10 equivalents) of Lewis acid. Heating for 30 min followed by work up produced di-acylphloroglucinols 9–12 in a range of 17 to 55% yield. The complete characterization data for each compound are provided in the appendix.



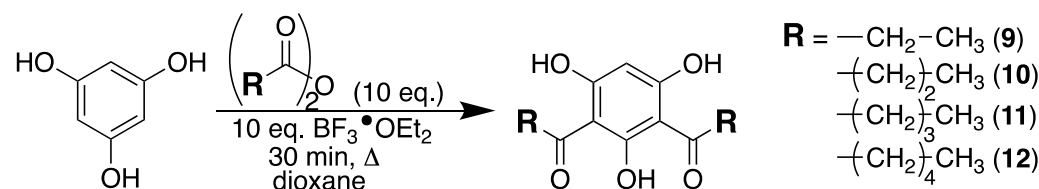
**Figure 4.2:** MPA analogs, to be used in a proposed conversion to UA analogs catalyzed by cytochrome p450.



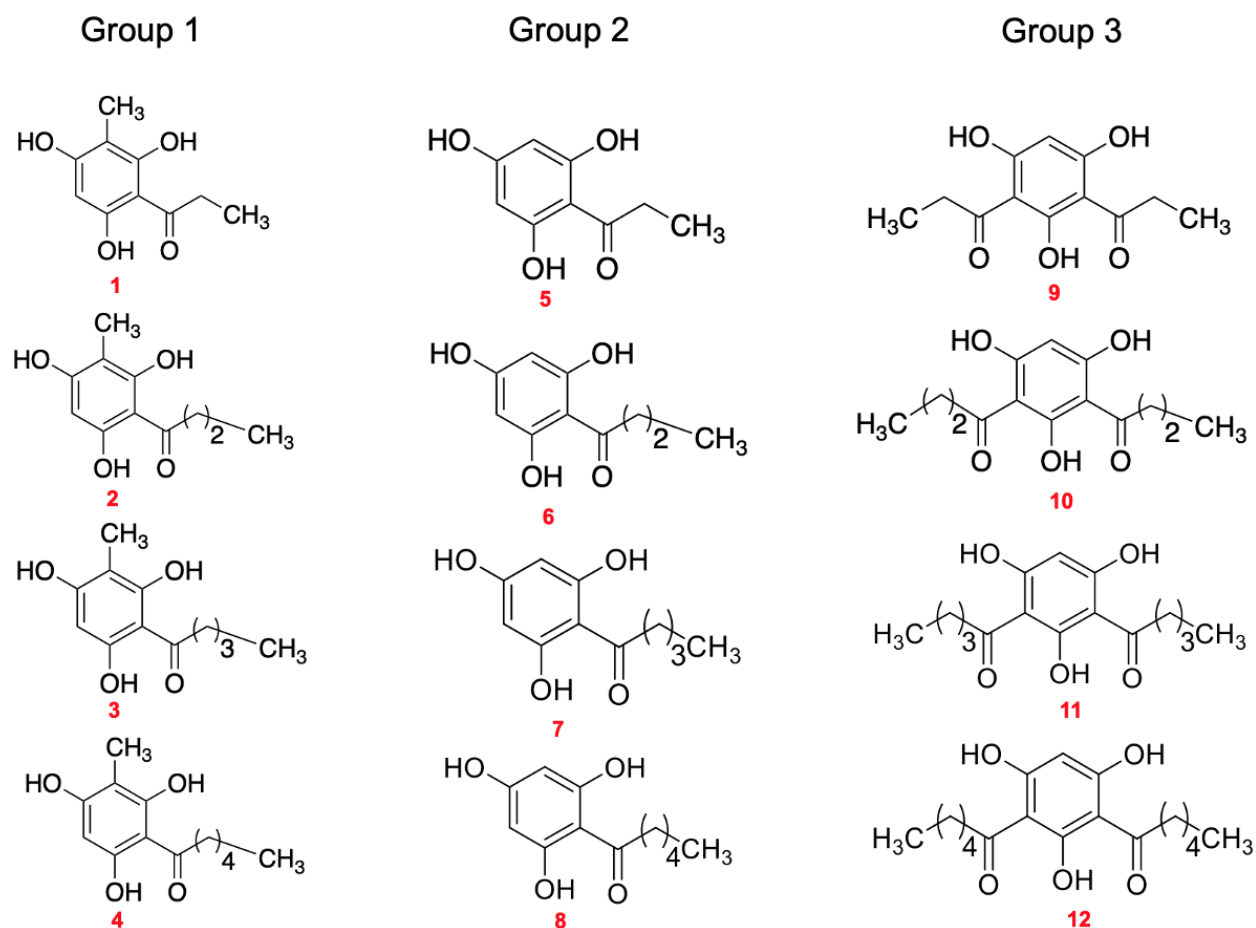
**Scheme 4.1:** The synthesis of compounds 1 – 4 is shown.



**Scheme 4.2:** The synthesis of compounds 5 – 8 is shown.



**Scheme 4.3:** The synthesis of compounds 9 – 12 is shown.



**Figure 4.3:** Chemical structure of the synthesized compounds. Compounds were grouped according to the synthesis scheme. Where Group 1 compounds are the analogs of MPA, Group 2 are mono-acylated phloroglucinols and Group 3 contains the most bioactive compounds (di-acylated phloroglucinols).

### 4.3. Evaluation of antibiotic properties

The minimum inhibitory concentration (MIC) of UA against *S. aureus* ranges between 6 and 32  $\mu\text{g ml}^{-1}$  depending on the strain [Sarkar, R. et al., 2018]. To evaluate the antibiotic activity of the synthesized compounds, we performed MIC assays using a laboratory strain of *S. aureus* (ATCC 29213) and MRSA strains isolated from the sputum of CF patients [Pena Amaya, P. et al., 2018]. **Table 4.1** shows MICs of the synthesized compounds. Group 3 had the lowest MICs, ranging between 0.125 and 8  $\mu\text{g ml}^{-1}$ . Among them, compound 11 had the lowest MIC against all the strains

of *S. aureus* tested ranging from 0.125 to 0.5  $\mu\text{g ml}^{-1}$ , while compounds 9 and 10 had MIC values ranging from 2 to 8  $\mu\text{g ml}^{-1}$ . The common structural feature for the four most active compounds (9–12) is a di-keto moiety with a mutual *ortho* phenolic group. The compounds differ only in the chain length of the alkyl group on the anhydride that was used in the acylation step. We then tested the ability of compounds 9–12 to effectively kill *S. aureus* by performing minimum bactericidal concentration (MBC) assays. All compounds displayed MBC to MIC ratio lower than 4 for all the strains tested with exception of compound 11 and strain CF188 (**Table 4.2**), indicating a bactericidal action [Pankey, G. A. & Sabath, L. D., 2004]. We also evaluated the ability of the synthesized phloroglucinol compounds to eradicate biofilms by performing the minimum biofilm eradication concentration (MBEC) assay (**Table S4**). None of the four compounds from group 3 that had antibacterial activity possessed biofilm eradication activity (**Table S4**).

**Table 4.1:** MIC of the 12 novel phloroglucinol derivatives against *S. aureus* strains. Values presented are the median of three biological replicates. *SCV* = Small colony variant.

Compound no.	MIC ( $\mu\text{g/mL}$ )							
	CF66 (MSSA)	ATTC 29213	CF225 (MRSA)	CF188 (MRSA)	CF224 (MRSA)	CF4 (MRSA)	CF250 ( <i>SCV</i> MRSA)	CF3 (MRSA)
1	>32	>32	>32	>32	>32	>32	>32	>32
2	>32	>32	>32	>32	>32	>32	>32	>32
3	>32	>32	>32	>32	>32	>32	>32	>32
4	>32	>32	>32	>32	>32	>32	>32	>32
5	>32	>32	>32	>32	>32	>32	>32	>32
6	>32	>32	>32	>32	>32	>32	>32	>32
7	>32	>32	>32	>32	>32	>32	>32	>32
8	>32	>32	>32	>32	>32	>32	>32	>32

<b>9</b>	8	2	2	2	2	4	4	8
<b>10</b>	8	2	4	8	8	8	4	8
<b>11</b>	0.25	0.25	0.5	0.125	0.125	0.25	0.5	0.125
<b>12</b>	0.5	0.25	0.5	0.5	0.25	0.5	1	0.5

**Table 4.2:** MBC and ratio of MBC to MIC for the four novel phloroglucinol derivatives against *S. aureus* strains.

Compound no.	<i>S. aureus</i> CF3		<i>S. aureus</i> CF4		<i>S. aureus</i> CF66		<i>S. aureus</i> CF188	
	MBC ( $\mu\text{g/mL}$ )	MBC/ MIC ratio	MBC ( $\mu\text{g/mL}$ )	MBC/ MIC ratio	MBC ( $\mu\text{g/mL}$ )	MBC/ MIC ratio	MBC ( $\mu\text{g/mL}$ )	MBC/ MIC ratio
<b>9</b>	16	2	16	4	16	2	8	4
<b>10</b>	32	4	32	4	32	4	32	4
<b>11</b>	0.5	4	0.5	2	0.5	2	1	8
<b>12</b>	1	2	1	2	1	2	1	2

To further investigate the therapeutic potential of the four active compounds, we evaluated their hemolytic properties against red blood cells (RBC) (**Table 4.3**). Only 11 displayed minor hemolytic activity of 8% at the highest concentration tested ( $32 \mu\text{g mL}^{-1}$ ), suggesting it would be unsuitable for therapeutic purposes.

**Table 4.3:** *In vitro* hemolysis of ovine red blood cells by novel phloroglucinol derivatives. Values presented are the mean of three biological replicates  $\pm$  SD.

Compound no.	Percentage hemolysis <sup>a</sup>		
	<b>32 <math>\mu\text{g/mL}</math></b>	<b>16 <math>\mu\text{g/mL}</math></b>	<b>8 <math>\mu\text{g/mL}</math></b>
<b>9</b>	$0.60 \pm 1.14$	$1.07 \pm 1.48$	$0.74 \pm 1.18$
<b>10</b>	$0.17 \pm 1.08$	$0.07 \pm 0.56$	$0.28 \pm 1.0$

<b>11</b>	8.10 ± 0.89	0.49 ± 0.55	-0.52 ± 0.93
<b>12</b>	2.86 ± 1.88	2.27 ± 2.14	1.44 ± 1.49

<sup>a</sup>Values presented are the mean of three biological replicates ± SD

**Table 4.4:** Interaction of Compound 9 with several clinically relevant antibiotics against *S. aureus*. Values presented are the median of three biological replicates.

Antibiotic	<i>S. aureus</i> ATCC 29213		<i>S. aureus</i> CF225	
	FIC Index	Interpretation	FIC Index	Interpretation
Oxacillin	0.625	Additive	0.75	Additive
Cephalexin	1	Additive	1	Additive
Rifampicin	0.75	Additive	1	Additive
Doxycycline	0.325	Synergistic	0.325	Synergistic
Trimethoprim	1.5	Additive	1.5	Additive

Furthermore, the lack of hemolytic activity for compounds 9, 10, and 12 suggests that this class of compounds may have promising selectivity against bacterial cells. Hence, more research on understanding the mechanism of action and the apparent selectivity is merited. We also investigated the interaction of the compounds with clinically relevant antibiotics that have various mechanisms of action. We selected compound 9 as representative for checkerboard assays as it displayed strong activity, but without hemolysis. As shown in **Table 4.4**, Compound 9 interacted additively with oxacillin, cephalexin, rifampicin, and trimethoprim, and acted synergistically with doxycycline. This trend was seen in ATCC 29213 and the MRSA strains. It has been frequently observed that synergies often occur between drugs that target the same cellular process [Brochado, A. R. et al., 2018] and also can happen due to modulation of intracellular drug concentrations. Doxycycline targets the small ribosome subunit (30S subunit) interrupting protein synthesis. In a

study conducted by Sahuquillo-Arce et al. it was shown that linezolid interacted synergistically with doxycycline [Sahuquillo Arce, J. M. et al., 2006]. Linezolid binds to the 50S ribosomal subunit inhibiting protein synthesis. Therefore, the synergistic interaction between doxycycline and linezolid occurs because the two compounds act on interacting targets [Sahuquillo Arce, J. M. et al., 2006]. It is tantalizing to speculate that the mechanism of action of compound 9 is related to the inhibition of protein synthesis, a hypothesis worth to further be tested.

#### **4.4. Summary**

In the present study, we have synthesized 12 phloroglucinol derivatives and characterized their antibiotic properties against *S. aureus*, an important CF pathogen. Diacylated derivatives displayed the strongest bactericidal activity against MRSA clinical isolates. Furthermore, these compounds displayed apparent selectivity in their action, as these were not hemolytic. Of interest, compound 9 interacted synergistically with doxycycline. Hence, it could potentially be used as a drug combination to treat highly resistant MRSA strains.



## References

- Akil, N., Muhlebach, M. S. Biology and management of methicillin resistant *Staphylococcus aureus* in cystic fibrosis. *Pediatr Pulmonol.* **2018**.
- Brochado, A. R. *et al.* Species-specific activity of antibacterial drug combinations. *Nature* **2018**, 559, 259–263.
- Chambers, H. F. & DeLeo, F. R. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat Rev Microbiol* **2009**, 7, 629–641.
- Chen, S., Zhang, Z., Qing, T., Ren, Z., Yu, D., Couch, L., Ning, B., Mei, N., Shi, L., Tolleson, W. H., & Guo, L. Activation of the Nrf2 signaling pathway in usnic acid-induced toxicity in HepG2 cells. *Archives of toxicology*, **2017**, 91(3), 1293–1307.
- Gangell, C. *et al.* Inflammatory responses to individual microorganisms in the lungs of children with cystic fibrosis. *Clin. Infect. Dis.* **2011**, 53, 425–432.
- García-Cortés, M., Robles-Díaz, M., Ortega-Alonso, A., Medina-Caliz, I., & Andrade, R. J. Hepatotoxicity by Dietary Supplements: A Tabular Listing and Clinical Characteristics. *International journal of molecular sciences*, **2016**, 17(4), 537.
- Harvey, A. L., Edrada-Ebel, R. & Quinn, R. J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* **2015**, 14, 111–129.
- Heihachiro, T., Ushio, S. & Shoji, S. Biosynthesis of natural products.VI) Biosynthesis of usnic acid in lichen's (1). A general scheme of biosynthesis of usnic Acid. *Chem Pharma* **1969**, 17, 2054–2060.
- Ingólfssdóttir, K. Usnic acid, *Phytochemistry*, **2002**, 61(7), 729-736.
- Miller, L. G. & Kaplan, S. L. *Staphylococcus aureus*: A community pathogen. *Infectious Disease Clinics of North America* **2009**, 23, 35–52.

- Muhlebach, M. S. Methicillin-resistant *Staphylococcus aureus* in cystic fibrosis: how should it be managed? *Curr Opin Pulm Med* **2017**, *23*, 544–550.
- Munita, J. M., Shelburne, S., Greenberg, D. E. & Arias, C. A. The Growing threat of antimicrobial resistance. *Tex Med* **2017**, *113*, 48–52.
- Pankey, G. A. & Sabath, L. D. Clinical relevance of bacteriostatic versus bactericidal mechanisms of action in the treatment of Gram-positive bacterial infections. *Clin. Infect. Dis.* **2004**, *38*, 864–870.
- Pena Amaya, P. *et al.* Molecular Epidemiology of Methicillin-Resistant *Staphylococcus aureus* in Cystic Fibrosis Patients from Argentina. *Microbial. Drug Resistance* **2017**, doi:10.1089/mdr.2017.0162
- Pena Amaya, P. *et al.* Molecular epidemiology of Methicillin-Resistant *Staphylococcus aureus* in Cystic Fibrosis patients from Argentina. *Microb. Drug Resist.* **2018**, *24*, 613–620.
- Public Health Agency of Canada. Canadian Antimicrobial Resistance Surveillance System **2017** Report.
- Sahuquillo Arce, J. M., Colombo Gainza, E., Gil Brusola, A., Ortiz Estévez, R., Cantón, E., & Gobernado, M. In vitro activity of linezolid in combination with doxycycline, fosfomicin, levofloxacin, rifampicin and vancomycin against methicillin-susceptible *Staphylococcus aureus*. *Revista española de quimioterapia: publicación oficial de la Sociedad Española de Quimioterapia*, **2006**, *19* (3), 252–257.
- Sarkar, R., Mittal, N., Sorensen, J., Sen T. A comparison of the bioactivity of usnic acid versus methylphloracetophenone *Nat. Prod. Commun.* **2018**, *13*, 1673–1676.
- Stryjewski, M. E. & Corey, G. R. Methicillin-Resistant *Staphylococcus aureus*: An evolving pathogen. *Clin Infect Dis* **2014**, *58*, S10–S19

# Chapter 5

## Functional heterologous expression of lichen cytochrome p450 (CYP) and p450-reductase (CPR) in bacteria responsible for UA biosynthesis

This chapter has been edited in the form of a manuscript and, planned to submit to *Journal of Nature Communications* in December 2022.

### 5.1. Introduction

**Chapter 3** described in some detail that the genome sequencing, assembly, and annotation of *C. uncialis* resulted in the discovery of 48 putative gene clusters out of which 32 are polyketide synthase (PKS) gene clusters. Further, these 32 PKS gene clusters are subcategorized into 16 non-reducing & 16 reducing PKSs [Abdel-Hameed, M. et al., 2016]. As discussed in **Chapter 1**, non-reducing PKSs do not have reducing domains contrary to the reducing PKSs that have reducing domains such as enoyl & keto-reductases. In polyketide synthase gene clusters, there are PKS genes encode for backbone enzymes and genes, up or downstream of PKSs encode for tailoring/accessory/post-pks enzymes. PKSs are known to assemble different poly-ketonic chains using different catalytic domains. Which further catalyzes the cyclization of poly-ketonic chains in a particular fashion following different cyclization mechanisms leading to the formation of unique polyphenol molecules. These polyphenols are usually the key intermediates in the biosynthesis of different polyketides. Tailoring/accessory/post-PKS enzymes catalyze the conversion of these key intermediates to more complex polyketide molecules. Polyketides are one main class of natural products produced by lichens and are typically aromatic polyphenols with varying degrees of oxidations and other modifications.

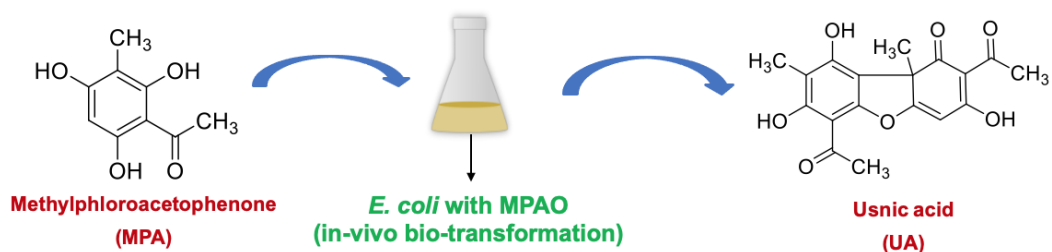
The first NR-PKS gene cluster of *C. uncialis* that was identified and completely annotated by Dr. Mona Abdel-Hameed and Dr. Robert Bertrand (ex-PhD candidates in our lab) is involved in the biosynthesis of a natural product named UA (see **Chapter 3**) [Abdel-Hameed, M. et al., 2016]. Earlier mentioned in **Chapter 3**, the details of various lichen sources of UA, its biological importance, and its potential of being a versatile drug molecule in the pharmaceutical industry. The UA gene cluster (see **Chapter 3**) consists of two genes, a non-reducing polyketide synthase gene, encodes for NR-PKS enzyme named MPAS and a post-PKS gene encodes for a cytochrome p450 enzyme named MPAO. The proposed biosynthetic scheme of UA (see **Chapter 3**) involves a catalytic intermediate MPA and MPAO catalyzes the oxidation of two molecules of MPA to UA. Cytochrome p450 (CYP) is an oxidative enzyme and **Chapter 1** has already given a detailed insight into cytochrome p450s (CYPs) in general and eukaryotic CYPs, their catalytic reaction types, mechanisms, and their redox partners/proteins (RPs). Eukaryotic RPs are cytochrome p450 reductase (CPR), cytochrome b5 (CB5) and cytochrome b5 reductase (CB5R).

## **5.2. An Outline of all experiments for UA biosynthesis described in this chapter**

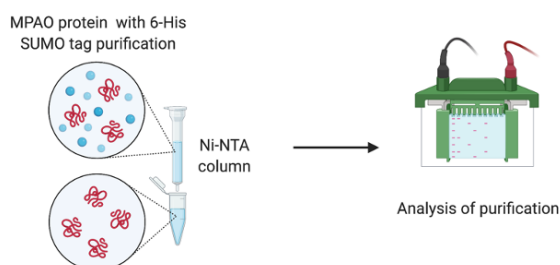
This chapter illustrates:

- 1) Attempts toward a functional heterologous expression of MPAO in *E. coli* (**Figure 5.1A**).
- 2) The soluble expression of MPAO and its redox partner CPR proteins and their purification (**Figure 5.1B, C & D**).
- 3) Attempt towards bioconversion of MPA to UA catalyzed by pure MPAO protein,
- 4) Attempt towards bioconversion of MPA to UA catalyzed by both purified; MPAO & CPR proteins (**Figure 5.1E**).
- 5) Successful functional heterologous co-expression of *C. uncialis* CPR (cytochrome p450 reductase) with MPAO in *E. coli* (**Figure 5.1F**).

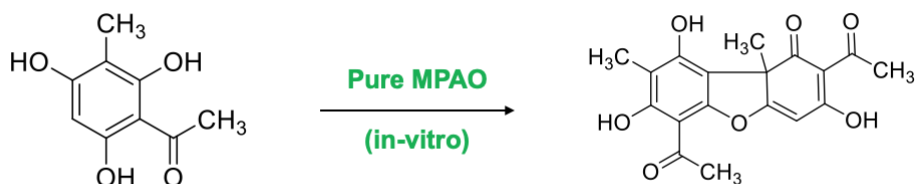
A) Functional heterologous expression of MPAO in *E. coli*



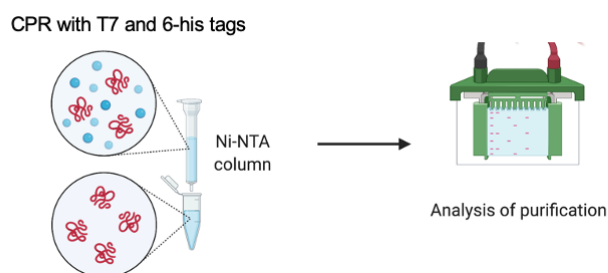
B) Large scale soluble MPAO expression in *E. coli* and its purification

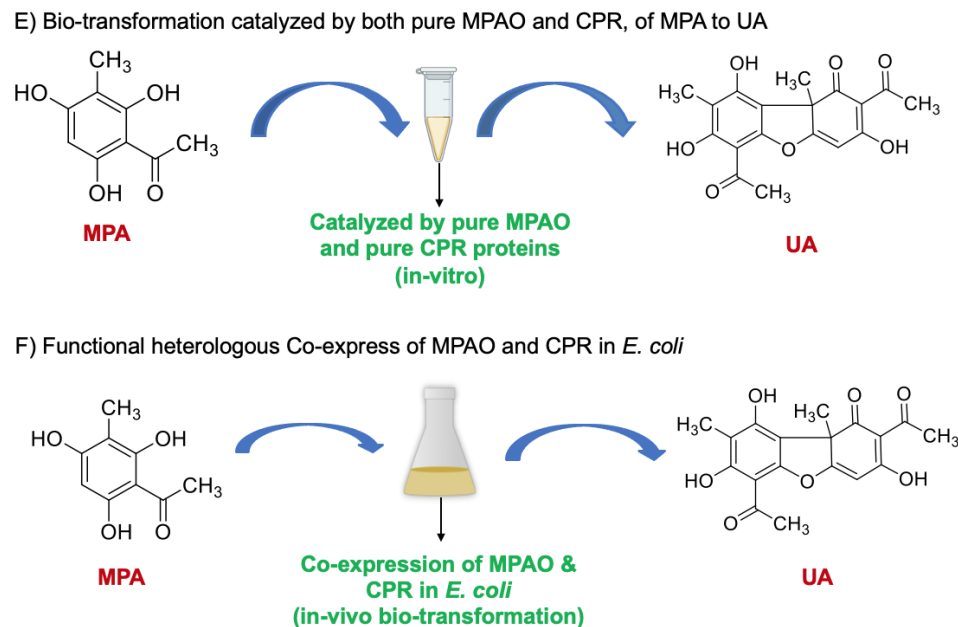


C) Bio-transformation of MPA to UA catalyzed by pure MPAO



D) Large scale soluble CPR expression in *E. coli* and its purification





**Figure 5.1:** Summary of the experimental work of functional heterologous expression of MPAO, and CPR; **A.** Schematic outline of a bio-conversion experiment of MPA to UA catalyzed by MPAO by incubating MPA with *E. coli* transformants (*in vivo*), **B.** Large scale production of soluble novel lichen CYP (MPAO with N-His SUMO tag), its purification aided by Ni-NTA (Ni-nitrilo acetate) affinity column chromatography and screened with SDS-PAGE (sodium dodecyl-sulphate polyacrylamide gel electrophoresis), **C.** Schematic outline of a bio-conversion experiment of MPA to UA by incubating MPA with purified MPAO (*in vitro* experiments), **D.** Large scale production of soluble novel lichen CPR with T7 and His<sub>6</sub> tags, its purification aided by Ni-NTA affinity column chromatography, **E.** Schematic outline of a bio-conversion experiment of MPA to UA by incubating MPA with both purified MPAO and CPR (*in vitro* experiments), **F.** Schematic outline of a bio-conversion experiment of MPA to UA by incubated MPA with *E. coli* transformants (*in vivo*) co-expressing MPAO and CPR.

### 5.3. Importance and the goal of this study

It is of great importance to characterize the biochemistry of these lichen proteins to better understand lichen biology. The exploitation of the biochemistry of this lichen p450s is not only necessary for lichen NPs biosynthesis but also for their applications in biotechnology and the pharmaceutical industry. Isolation of p450s from native tissues, however, often results in modest yields. Therefore, recombinant protein production is an attractive alternative. The development of a heterologous expression system is often a principal step in biochemical and biotechnological studies of eukaryotic proteins. This study aims to establish a methodology to achieve a functional heterologous expression of these lichen fungal proteins, that are membrane-bound CYPs and their

RPs such as cytochrome p450 oxidoreductase/reductase (CPR) in bacteria. To better understand how to improve experimental procedures to institute purified lichen proteins for their formal *in vitro* characterization.

#### **5.4. Why biocatalysts over chemical synthesis**

1) Biocatalysts can help reduce the number of tedious steps such as the use of protecting groups and deprotecting those that are commonly required in conventional chemical synthesis, 2) biocatalysts produce a single product with no by-product formation, which reduces the cost of downstream purification steps, and 3) biocatalysts produce enantiomeric mixtures (rather than racemic mixtures) that are strictly required in some fields, such as the pharmaceutical industry.

#### **5.5. Advantages of using *E. coli* as a heterologous host (Figure 5.2)**

It is important to produce large quantities of purified, properly folded and catalytically active CYP proteins for the downstream biochemical and structural studies. Therefore, a fast-growing bacterial system has always been a dominant host for the recombinant expression of eukaryotic proteins. These prokaryotic cell factories provide a cost-effective avenue using inexpensive media for large-scale protein production. *E. coli* is one of the most powerful tools for heterologous protein production because of the well-established methods for their genetic manipulation, the multitude of molecular tools and protocols at hand, the vast catalogue of expression plasmids and many cultivation strategies [Demain, A. L. & Vaishnav, P. 2009].

Pros	Cons
1. Fast growth kinetics	1. Requirement of redox proteins
2. High cell densities	2. Requirement of NAD(P)H cofactor
3. Inexpensive media	3. Lack of membrane binding systems
4. Established methods for genetic manipulation	4. Lack of post-translational modifications
5. Vast catalog of expression plasmids	5. Protein folding and solubility issues
6. Various cloning and expression protocols at hand	6. Enzyme stability issues
7. Many cultivation strategies	7. Requirement of heme precursor

**Figure 5.2:** Advantages and Disadvantages associated with a prokaryotic expression system [Durairaj, P. et al., 2019].

### 5.6. Challenges of using a prokaryotic host to express eukaryotic proteins (Figure 5.2)

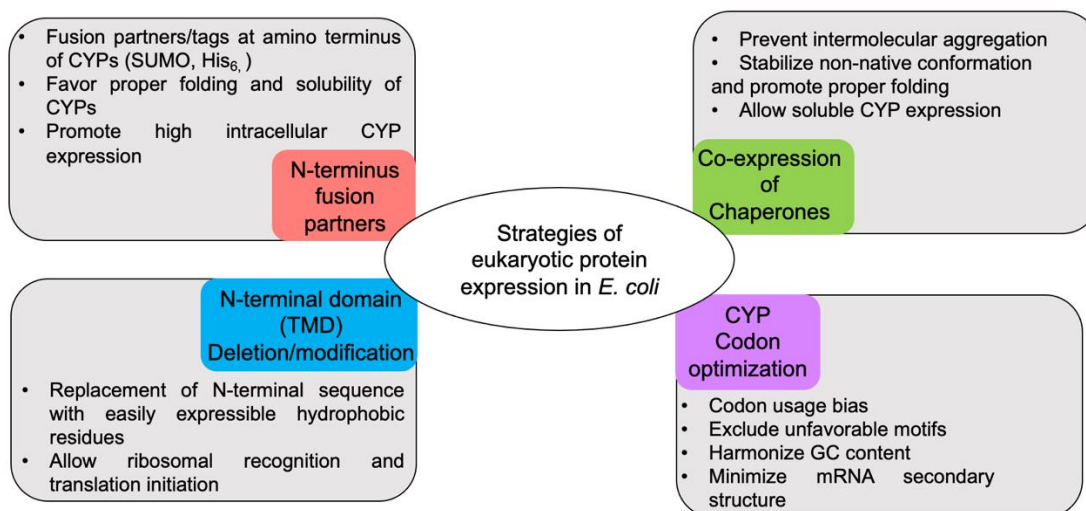
Nevertheless, the recombinant production of soluble eukaryotic proteins in *E. coli* always remains challenging and a laborious task owing to their membrane-bound nature [Durairaj, P. et al., 2019]. The latest model has determined these eukaryotic CYPs as membrane-associated proteins where they are anchored to the endoplasmic reticulum (ER) through N-terminal transmembrane alpha helix on the luminal side and the catalytic domain is sprawling in the cytosol as shown in (Figure 5.4A) [Šrejber, M. et al., 2018; Urban, P. et al., 2018]. The lack of membrane-bound organelles in prokaryotes results in either non-expression of eukaryotic CYPs or their misfolding or incomplete folding into insoluble aggregates or inclusion bodies. Which eventually hinders their functional and structural studies [Denisov, I. G. et al., 2012; Zelasko, S. et al., 2013; Durairaj, P. et al., 2016]. Other factors impeding expression include the incorporation of a heme group, the requirement of redox partners (CPR, CB5 & CB5R) for the electron transfer (ET) to complete the catalytic cycle of CYP (Figure 5.4) and the co-expression of signal proteins to achieve a soluble expression of these membrane proteins. Also, bacterial systems are not adapted to perform post-translational modifications (PTMs) that are often necessary for a eukaryotic protein to be functional. PTMs are chemical modifications of proteins after their translation and



enable cells to respond quickly to any environmental changes. This includes the addition of chemical groups such as phosphates and acetates or more complex molecules, e.g., carbohydrates or lipids to proteins, the covalent linkage of small proteins (like ubiquitin and ubiquitin-like proteins (UBLs)) or the modification of side chain residues of specific amino acids (like deamidation or elimination) and all these activities are catalyzed by tightly regulated enzymes [Chen, J. & Tsai, Y. H. 2021]. However, literature studies have shown that there are a few post-translational modifications that occur in some species of bacteria [Macek, B. et al., 2019; Chapot-Chartier, M. P. and Buddelmeijer, N. 2021] but still, many complex PTMs are unique to eukaryotes. The absence of spliceosome machinery to automatically remove introns (the non-coding regions in a DNA sequence) in prokaryotes results in some extra tedious steps to be able to express a eukaryotic gene with introns, in *E. coli*. The codon mismatch between the host cell (prokaryotic) and protein of interest (eukaryotic) slows down the rate of translation, which ultimately results in low or faulty eukaryotic protein expression in *E. coli*.

### **5.7. Strategies to improve and achieve a high expression of Eukaryotic P-450s in bacteria**

Several efforts have been put and many changes have been undertaken to achieve maximal expression of eukaryotic p450s in bacteria (**Figure 5.3**) [Cheng, D. et al., 2004].



**Figure 5.3:** Strategies/key approaches to improve the eukaryotic CYP production in bacteria [Cheng, D. et al., 2004].

### 5.7.1. Codon optimization

Modern biotechnological techniques of gene synthesis facilitate the appropriate codon usage based on the choice of host expression system by harmonizing the favourable GC content and eradicating the mRNA secondary structure at the 5' untranslated region. Codon optimization of eukaryotic p450 exons for *E. coli* leads to high protein expression in lesser time as shown in studies (**Figure 5.3**) [Claassens, N. J. et al., 2017].

### 5.7.2. N-terminal domain (TMD) modification or deletion

There is a lot of work that has been done so far related to the modification of the N-terminal domain/transmembrane domain (NTD/TMD) [Pan, Y. et al., 2011; Hatakeyama, M. et al. 2016], or the complete elimination of NTD [Budriang, C. et al., 2011; Ichinose, H. & Wariishi, H. 2012]. NTD is almost always cleaved or deleted to get solubilized p450 by redistributing it in the cytosol upon expression in bacteria. The only exceptions are the two full-length microsomal CYPs, *viz.*, CYP51 lanosterol 14 alpha-demethylase of *Saccharomyces cerevisiae* (PDB ID: 4LXJ and 5EQB) [Monk, B. C. et al., 2014]) and CYP19 aromatase of *Homo sapiens* that were expressed successfully in bacteria (PDB ID: 3EQM) [Ghosh, S et al., 2009]. The function of the NTD is

mostly the protein trafficking to ER membrane upon recognition by a signal recognition particle (SRP). Modifications have been carried out either by changing the native NTD or inserting a previously established sequence (easily recognized by bacteria) to facilitate the full-length p450 expression in *E. coli*. The modification of this 20-30 amino acid hydrophobic NTD is preferred over its deletion in the cause of retaining protein's membrane-binding characteristics. NTD correctly directs the protein towards the bacterial plasma membrane. Studies have shown that NTD/TMD plays an important role in the electron transfer mechanism of CYPs by affecting the interaction of CPR and CYP [Šrejber, M. et al., 2018; Gideon, D. A. et al., 2012]. Hence, TMD truncation might impair the catalytic activity of CYPs and affect the phospholipid bilayer composition leading to folding and stability issues (**Figure 5.3**).

### 5.7.3. Co-expression of Chaperone proteins

Co-expression of ER-residing chaperone proteins with CYPs improves the yield of properly folded, active, and soluble eukaryotic p450s expressed in bacteria (**Figure 5.3**) [Schlapschy, M., & Skerra, A. 2011]. Takara Bio's Chaperone Plasmid Set consists of five different types of chaperone plasmids developed by HSP Research Institute Inc as shown in (**Table 5.1**). The plasmids are designed to enable efficient expression of multiple molecular chaperones and have been reported to increase the recovery of target CYPs in the soluble fractions upon co-expression in the BL21 (DE3) strain of bacteria. Specifically, GroEL-GroES molecular chaperone has been used widely in co-expressing eukaryotic CYPs [Sørensen, H. P. and Mortensen, K. K. 2005; Ichinose, H. and Wariishi, H. 2013; Ichinose, H. et al., 2015; Wu, Z. L. et al., 2006; Wu, Z. L. et al., 2009; Hatakeyama, M. et al., 2016].

**Table 5.1:** Bacterial chaperone plasmids that express different chaperone proteins [Sørensen, H. P. and Mortensen, K. K. 2005; Ichinose, H. and Wariishi, H. 2013; Ichinose, H. et al., 2015; Wu, Z. L. et al., 2006; Wu, Z. L. et al., 2009; Hatakeyama, M. et al., 2016].

No.	Plasmids	Chaperone
1.	pG-KJE8	<i>dnaK-dnaJ-grpE, groES-groEL</i>
2.	pGro7	<i>groES-groEL</i>
3.	pKJE7	<i>dnaK-dnaJ-grpE</i>
4.	pG-Tf2	<i>groES-groEL-tig</i>
5.	pTf16	tig

#### 5.7.4. Expression plasmids with Fusion proteins/tags

Another tailor-made strategy that has emerged greatly in the past few years to combat the recombinant protein proper folding, solubility and purification issues in the bacterial system is the gene fusion technology (**Figure 5.3**) [Costa, S. et al., 2014]. Where a bacterial expression plasmid is used to incorporate fusion tags such as SUMO, 6-histidine (His<sub>6</sub>), Fh8, Flag etc. that assist in proper protein folding and increase the production of soluble and purified proteins. These fusion proteins/tags are usually located at the N-terminal region of the expression plasmid. The commonly used N-terminal fusion tag as a protein purification handle is the His(6)tag, where 6-histidine residues bind with the nickel of Ni-NTA (Nickel-nitrilo triacetate) resin used in affinity column chromatography and pure protein is later eluted by adding imidazole which displaces the poly-histidine residues [Lee, C. D. et al., 2008; Gao, X. et al., 2010; Wang, H. et al., 2010; Satakarni, M. et al., 2011]. Other fusion proteins are used in conjugation with the His<sub>6</sub> tag to assist both proper folding and protein purification, examples are maltose-binding protein (MBP); MBP-

His<sub>6</sub> and small ubiquitin-related modifier (SUMO); SUMO-His<sub>6</sub> [Wang, Z. et al., 2012]. SUMO is a small protein (100 amino acids; approx. 12 kDa), which promotes the proper folding and solubility of CYP protein possibly by exerting chaperoning effects in a similar mechanism to the described for its structural homolog Ubiquitin (Ub; Khorasanizadehetal, S. 1996). In addition, Ub and Ub-like proteins (Ulp) have a highly hydrophobic inner core. The hydrophilic surface of Ub, together with a tight & rapidly folding soluble structure, may explain SUMO's behaviour as a nucleation site for the proper folding of target proteins [Malakhovetal, M. P. et al., 2004; Marblestoneetal, J. G. et al., 2006].

#### *5.7.5. Fusion partners mimic eukaryotic PTMs catalyzing enzymes*

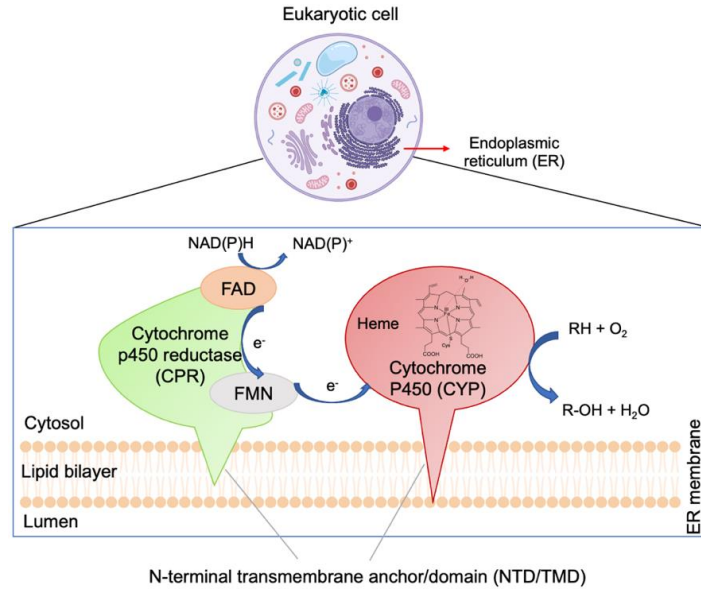
Ubiquitination is one of the many complex post-translational modifications (PTMs) reported in eukaryotes catalyzed by Ubiquitin enzyme [Katarzyna, K. et al., 2020]. It includes coordinating cellular localization of proteins, activating, and inactivating proteins, regulating protein degradation, and modulating protein-protein interactions. Fusion partners are also used for translocating proteins to a different cellular location with less concentration of proteases present to prevent protein degradation [Butt, T. R. et al., 2005]. SUMO is the structural homolog of Ub [Ribet, D., & Cossart, P. 2010] and MBP fusion tags and is known to exhibit the same feature [Costa, S. et al., 2014]. In the case of these fusion tags, a protease recognition sequence is present between the fusion partner coding gene and the target protein-coding gene that allows the tags removal after the protein purification step for downstream applications of proteins such as protein therapies, vaccine development and structural analyses. The unique features present by fusion partners have been constantly explored, and many novel strategies for facile and rapid protein production are now available [Costa, S. et al., 2014].

Despite all the above-mentioned strategies, there are still no well-established methods and universal guidelines to enable successful eukaryotic p450 overexpression in bacteria. Instead, all these tremendous efforts toward optimizing CYP protein yields, diverse expression systems and the purification strategies are tedious trial and error based and all these strategies are case-specific [Durairaj, P. et al., 2016; Maroutsos, D. et al., 2019; Zelasko, S. et al., 2013; Yun, C. H. et al., 2006; Ichinose, H. et al., 2015; Hausjell, J. et al., 2018]

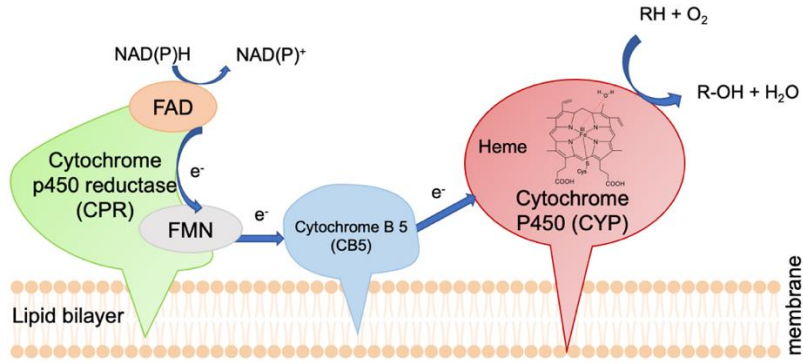
### **5.8. Co-expression of redox partners (RPs)**

CYPs have a requirement of sequential input of two electrons to complete their catalytic cycle. To conduct this two-electron transfer (ET) process, CYPs employ highly diversified redox chains comprised of varied redox partners (RPs) such as ferredoxins (Fdx; an iron Sulphur protein), ferredoxin reductase (FdR), cytochrome p450 reductase (CPR), cytochrome b5 (CB5) and cytochrome b5 reductase (CB5R) etc [Durairaj, P. et al., 2016]. Where Fdx and FdR, RPs are commonly found in prokaryotes and CPR, CB5, and CB5R are unique to eukaryotic systems. Like CYP, CPR is also a membrane-bound multidomain protein containing the prosthetic cofactors FAD and FMN, which transfer two electrons from NAD(P)H to the heme moiety of eukaryotic endoplasmic reticulum-associated CYPs (**Figure 5.4A**). Alternatively, CPR may also transfer one electron directly to CYP and the other electron to a third protein component, Cyt b5, which then transfers it to an oxyferrous CYP (**Figure 5.4B**). Rarely, certain CYPs (CYP5150A2) can be directly activated by Cyt b5 and NADH-dependant Cyt b5 reductase (CB5R) in the absence of CPR [Ichinose, H. & Wariishi, H. 2012; Hatakeyama, M. et al., 2016] as shown in **Figure 5.4C**. CPR, being an imperative redox partner of CYP and considering the requirement of CB5 and CB5R in some cases, the co-expression of these RPs with CYPs is essential for the optimal catalytic activity of CYPs.

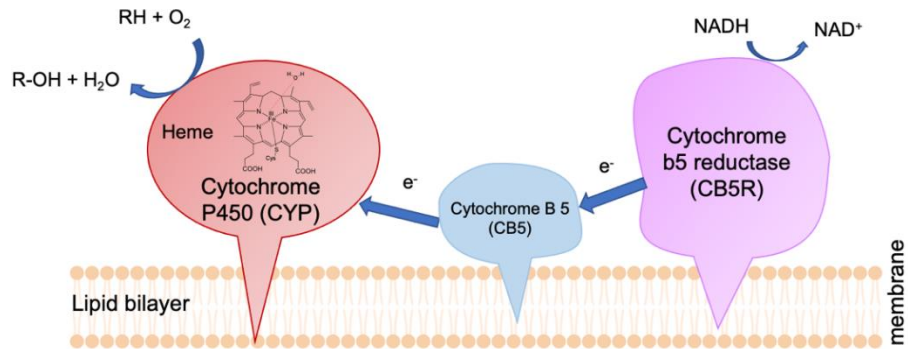
A)



B)



C)



**Figure 5.4:** Three different electron transfer (ET) mechanisms of redox partners (RPs) of eukaryotic CYPs; CPR, CB5, and CB5R. A. Microsomal Class II: NAD(P)H-[CPR]-[CYP]; B. NAD(P)H-[CPR]-[Cyt B5]-[CYP]; C. NADH-[CB5]-[Cyt B5]-[CYP] [Durairaj, P. et al., 2016].

### **5.9. Significance of cytochrome p450 reductase (the regulating effect of RPs)**

There is experimental evidence of electrostatic interactions between negatively charged amino acids of the FMN domain of CPR protein and positively charged amino acids of CYP protein [Campelo et al., 2018; Esteves et al., 2020]. With some recent breakthroughs, it has been found that the CPR enzymes not only function as auxiliary electron transfer proteins but also have a significant impact on the product formation catalyzed by CYP. The efficacy of CYP in terms of substrate selectivity and reactivity is dependent on the abundance of CPR and its ET compatibility with CYP [Braun et al., 2012; Durairaj et al., 2015; Durairaj et al., 2016; Zhang et al., 2018; Neunzig et al., 2013; Sagadin et al., 2019; Sagadin et al., 2018; Lah et al., 2011; Lv et al., 2019; Novak et al., 2015; Li et al., 2020]. As a result of the CYP functioning-regulating role of RPs, it has become extremely important to choose the right RP for a specific CYP and it is mainly dependent on the aim of the respective study.

### **5.10. Classification of CYPs based on their interaction with RPs**

Even with the low amino acid sequence identity of up to 16% within the superfamily of diversified p450s, the 3D structural fold where cysteine is axially linked to the heme prosthetic group is well conserved across all kingdoms of life. Depending on the topology of the protein components involved in electron transfer, CYP systems are classified into 10 different classes, and the fungal CYPs fall into classes II, VIII, and IX systems [Hannemann, F. et al., 2007; Crešnar, B., & Petrič, S. 2011]. Class II usually comprises two integral membrane proteins (microsomal): CYP and CPR, but in some cases, ET is catalyzed by either a combination of CPR and CB5 or CB5 & CB5R (**Figure 5.4**). Most of the fungal CYPs belong to class II and perform extremely

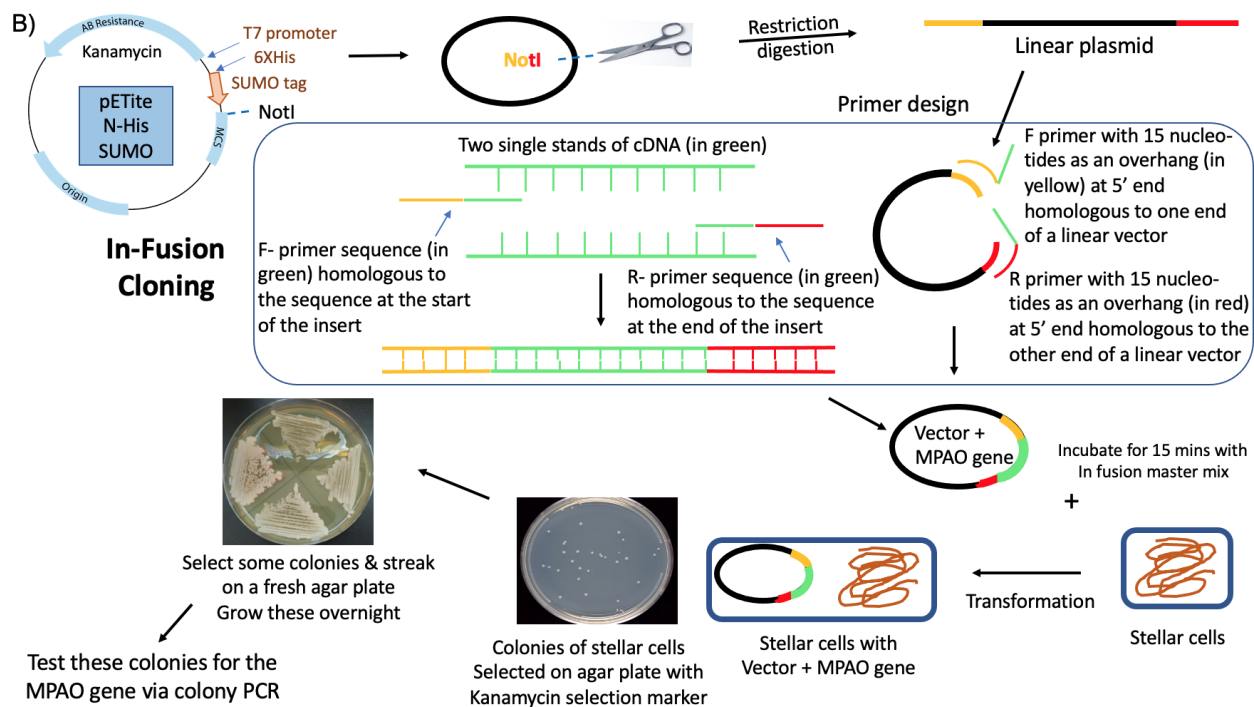
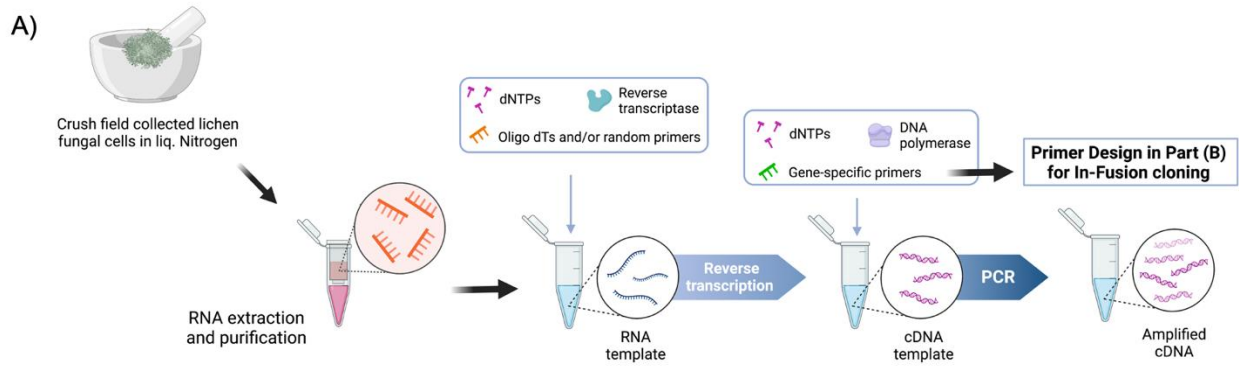


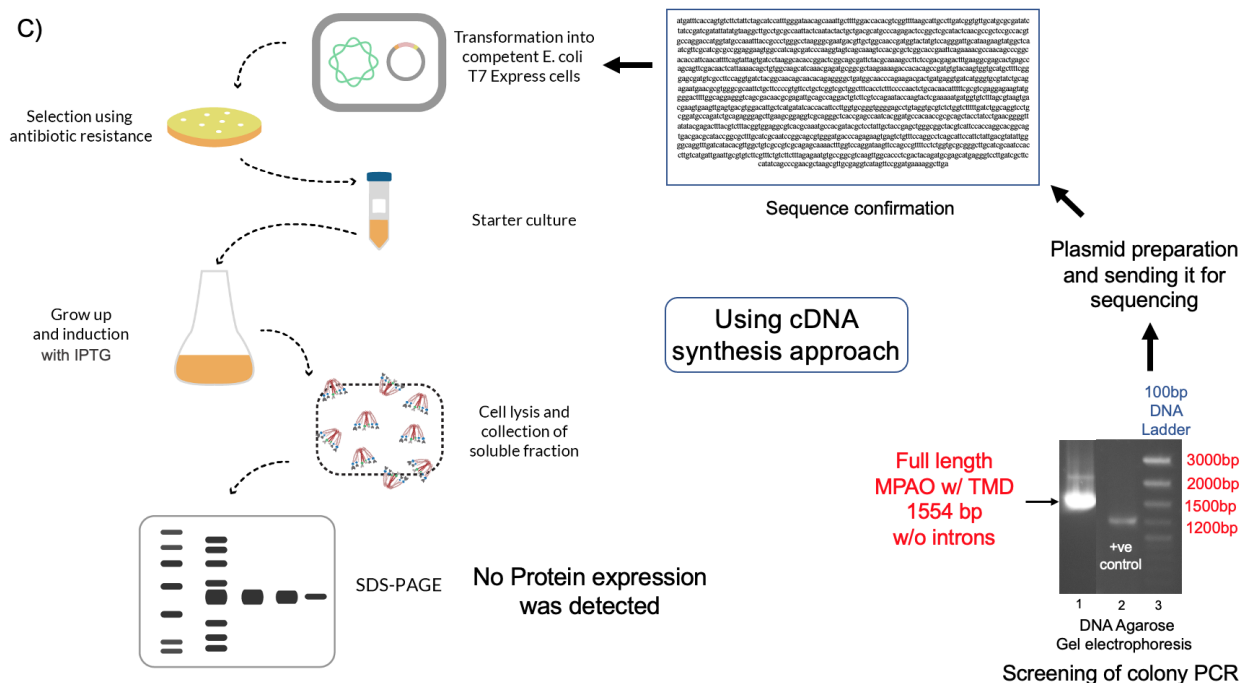
diverse catalytic reactions. Class VIII p450s are known to catalyze the subterminal hydroxylation in fatty acid synthesis, high homology with bacterial p450s CYP (P450BM3) and Class IX is exceptionally unique and so far, only one CYP is known to belong to this class (P-450nor), which is known to catalyze denitrification reactions. These CYPs do not need any redox partners to be functional and are the only known soluble fungal CYPs. [Shoun, H. & Takaya, N. 2002; Durairaj, P. et al., 2016]. Based on the above criteria of classification, MPAO posits to fall in the category of class II CYPs.

### **5.11. Heterologous expression of native MPAO (1554 bp) in *E. coli***

At the start of this research project when very little was known about lichen p450s & nothing was published regarding the expression of lichen genes within a heterologous host. We attempted the first experiment of functional heterologous expression of MPAO (a lichen p450) within bacteria. RNA was extracted from freshly field-collected *C. uncialis* fungus and cDNA was synthesized using reverse transcriptase (see **Section 2.5.1**) (**Figure 5.5A**).

MPAO cDNA (intact transmembrane domain) was amplified using gene-specific primers with the overhangs homologous to two opposite ends of a linear plasmid (**Figure 5.5B**) (see **Section 2.5.2**). The highly efficient In-fusion cloning technique (**Figure 5.5B**) for homologous recombination of MPAO & pETite N-His SUMO plasmid was used in contrast to the traditional cloning technique (see **Section 2.5.3a**). After the successful transformation in stellar cells, colonies were tested for MPAO clone (1554 bp) using colony PCR with 1% DNA agarose gel (see **Section 2.5.5**) (**Figure 5.5C**). After the sequence was confirmed, the pETite plasmid with the MPAO clone was transformed into BL21 (DE3)pLySs cells. MPAO protein was induced with IPTG (**Figure 5.5C**) and its expression was screened with 10% SDS-PAGE gel (see **Section 2.5.5**). No protein was detected after many expression trials.





**Figure 5.5:** Steps of cloning and heterologous expression of MPAO in *E. coli*. A) RNA extraction of lichen fungus and reverse transcription for cDNA synthesis & its amplification; B) plasmid map of pETite N-His SUMO vector, detailed steps of primer design, restriction digestion to linearize pETite vector, PCR amplification of cDNA for In-fusion cloning of MPAO in stellar cells and colonies of stellar cells were selected on agar plate with kanamycin selection marker and finally tested for MPAO clone by colony PCR; C) Colony PCR, analyzed by running DNA agarose gel in which lane 1 shows the correct size band of full length MPAO with an intact TMD at about 1554 bp when compared with 1 kb DNA ladder in lane 3 and lane 2 shows a band for +ve control, the same colony was prepared for sequence confirmation, after the right sequence was confirmed, pETite vector with MPAO transformed into BL21 (DE3)pLySs (T7 promoter) cells and MPAO expression was induced using IPTG inducer, then cells were lysed and both soluble and lipid fractions were tested for MPAO protein expression with 10% SDS-PAGE Gel.

### 5.11.1. Importance of choosing the appropriate bacterial strain

The choices of different expression vectors and bacterial strains are also the key variables for the successful lichen CYP expression in bacteria. BL21(DE3)pLySs Competent Cells allow high-efficiency protein expression of any gene that is under the control of a T7 promoter and has a ribosome binding site. BL21(DE3)pLySs is lysogenic for  $\lambda$ -DE3, which contains the T7 bacteriophage gene I, encoding T7 RNA polymerase under the control of the *lac* UV5 promoter. BL21(DE3)pLySs also contains a plasmid, pLySs, which carries the gene encoding T7 lysozyme. T7 lysozyme is a natural inhibitor of T7 RNA polymerase and serves to suppress basal expression

of T7 RNA polymerase prior to induction of target protein with IPTG, under the control of the T7 promoter. Thus, stabilizing recombinants encoding target proteins that may also affect cell growth and viability. In addition, empirical selection yields *E. coli* strains that are superior to the traditional BL21(DE3) host strain by overcoming the toxic effects associated with the overproduction of membrane and globular proteins under T7 transcriptional control [Terpe, K. 20006].

#### *5.11.2. Importance of choosing the right expression vector*

The selection of the expression vector is also crucial to confer specific features to eukaryotic proteins such as solubility and/or affinity tags and to direct the protein synthesis in either cytoplasm or the periplasm of bacteria. Other features of an expression vector such as plasmid copy number, transcriptional promoters and antibiotic-selection markers also play important role in recombinant eukaryotic protein expression. pET vectors are amongst the most frequently used expression plasmids for bacterial hosts. Like a pET vector system, the pETite N-His SUMO vector has also a T7 promoter where the transcription of a given CYP is initiated by T7 RNA polymerase. This polymerase is not native to the bacterial host and must be introduced into the host chromosome under a separate inducible *lac* operator [Studier, F. W. & Moffatt, B. A. 1986]. A tight production of T7 polymerase results in virtually no basal expression and once the *lacI* repressor is deactivated by IPTG, the pET vector is over-expressed to produce very high levels of protein [Domanski, T. L. et al., 2001]. T7 polymerase is also very specific towards the replication of only vector plasmid. The pETite system also uses an engineered form of the SUMO protein as a fusion partner and a His<sub>6</sub> motif at the amino terminus of the SUMO tag. SUMO is derived from the yeast *SMT3* gene product, and as mentioned earlier in **Section 5.7.5** it enhances the expression and solubility of proteins that are otherwise poorly expressed or insoluble. The His<sub>6</sub>-SUMO tag (moves around 15-18 kDa on SDS-PAGE gel) can later be removed precisely and

efficiently by SUMO Express Protease. The protease recognizes the tertiary structure of SUMO and cleaves precisely at its carboxyl terminus, allowing recovery of the intact protein. [Malakhov, M.P. et al., 2004; Marblestone, J. G. et al., 2006; Mossessova, E. & Lima, C.D. 2000].

No modification or deletion of the N-terminal sequence (the hydrophobic transmembrane domain; NTD) and no codon optimization of MPAO for *E. coli* was undertaken at this stage. Because we found some examples in the literature where a eukaryotic p450 was successfully expressed in *E. coli* without excision of the transmembrane domain (TMD), although the expression levels were quite low [Ichinose, H. & Wariishi, H. 2013]. People also demonstrated a loss of functionality of CYPs by removing NTD/TMD in some cases. Because the TMD truncation does not facilitate proper folding of the protein and results in the inactive form of CYP [Durairaj, P. et al. 2015]. We also found examples where *E. coli* strains were modified by adding some rare codons recognized by tRNAs (also known as rare tRNAs) to achieve high levels of protein expression via proper folding of a recombinant protein in codon bias-adjusted bacterial strains [Rosano, G. L. & Ceccarelli, E. A. 2009]. Because sometimes incorporating the codon-biased strategy to achieve a faster rate of protein translation confuses the bacteria and, eventually leads to misfolding and protein aggregation into inactive insoluble inclusion bodies.

We did not co-express a redox protein; either a homologous CPR (from the same source as CYP) or a heterologous CPR (from yeast or other fungi) with CYP. Bacteria also have redox proteins (RPs) which act in conjunction with bacterial p450s for their proper catalytic activity [Benarroch, J. M. & Asally, M. 2020]. Although *E. coli* lacks p450 enzymes, but they do produce pyruvate dehydrogenase and ferredoxin oxidoreductase redox proteins [Blaschkowski, H. P. et al., 1982; Noth, J. et al., 2013]. At the beginning of this project when my knowledge regarding these proteins was very little, I presumed that the bacterial RPs could interact with a recombinant

eukaryotic CYP and facilitate electron transfer from NAD(P)H or pyruvate to a recombinant CYP. However, there is no experimental evidence so far that shows the transfer of electrons catalyzed by bacterial RPs from NAD(P)H to recombinant eukaryotic CYP proteins in bacteria.

### **5.12. Heterologous expression of MPAO (w/o TMD &, codon-optimized for *E. coli*)**

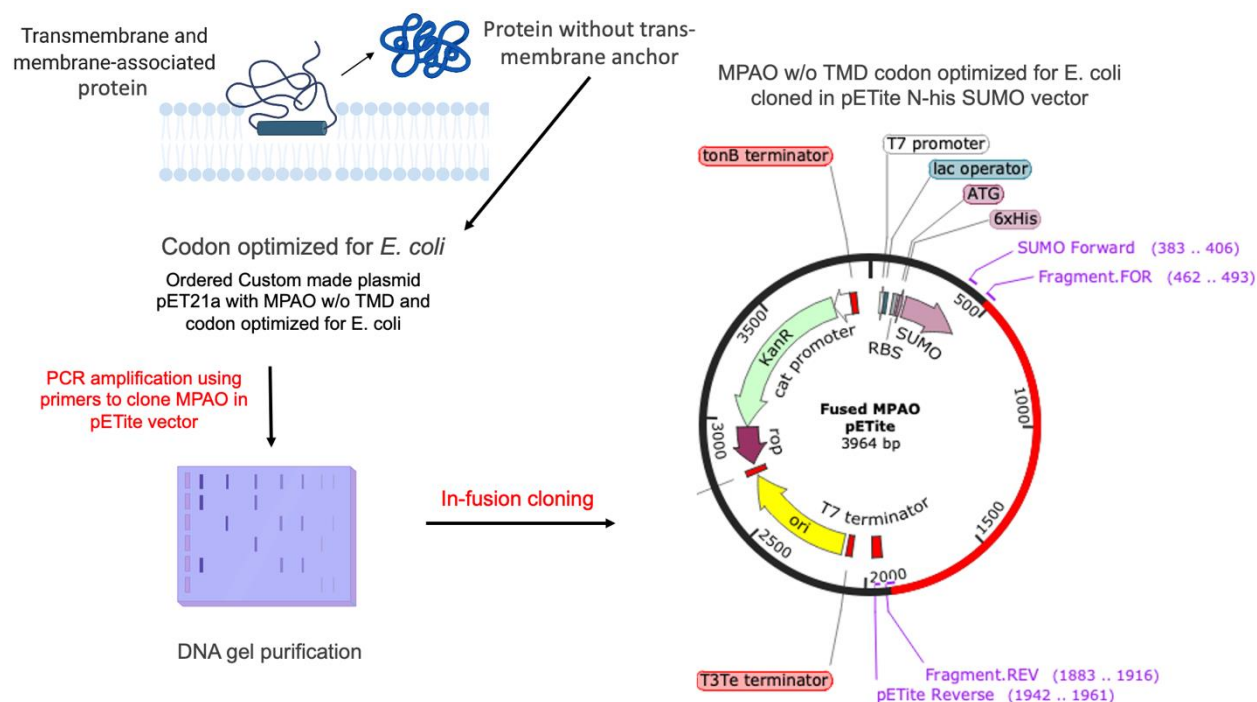
After getting no success in working with full-length MPAO w/ transmembrane domain, the next approach was to use a custom synthesized MPAO gene. Which was codon-optimized for *E. coli* and transmembrane-domain (TMD) was truncated. This gene was named CO(Eco)\_MPAO (C: codon O: optimized Eco: *E. coli*). Many articles reported the functional sustainability of cytochrome p450s even after the removal of TMD [Hausjell, J. et al., 2018].

The vector pET21a with CO(Eco)\_MPAO clone was custom ordered from [Twist Bioscience] (South San Francisco, CA, USA) and was later amplified using gene-specific primers to clone into pETite N-His SUMO vector (In-fusion cloning) (**Figure 5.6A**). Stellar cells harbouring CO(Eco)\_MPAO clone (1425 bp) were confirmed by colony PCR with 1% DNA agarose gel 1 (**Figure 5.6B**). Successful MPAO expression in BL21 (DE3)pLySs strain was detected with SDS-PAGE gel 2 (**Figure 5.6B**). An important observation was made after many protein expression trials at different temperatures and induction periods that there was a soluble expression of MPAO protein before lysing the cells with sonication. But, as soon as the cells were lysed via sonication, the protein was aggregating into insoluble fractions (**SDS-PAGE Gel 3, Figure 5.6B**), which was a result of localized heating during sonication. Upon troubleshooting, 1% Triton X-100 in lysis buffer was used; a non-ionic detergent that is widely used for cell lysis to extract proteins. Especially in those cases where it is important to maintain the protein function as well. Additionally, during sonication number of pulses was reduced from 8 to 4 and the cooling time was increased from 30 secs to 1 min in between the pulses to avoid excessive heating. Finally,

the protein was recovered successfully in the soluble fraction after cell lysis, but in low concentration (**Gel 1, Figure 5.7**). The successful soluble protein expression was also confirmed with Western blot using an anti-His<sub>6</sub> antibody (**Figure 5.7**). MPAO was purified using a Ni-NTA column (**Gel 2, Figure 5.7**).

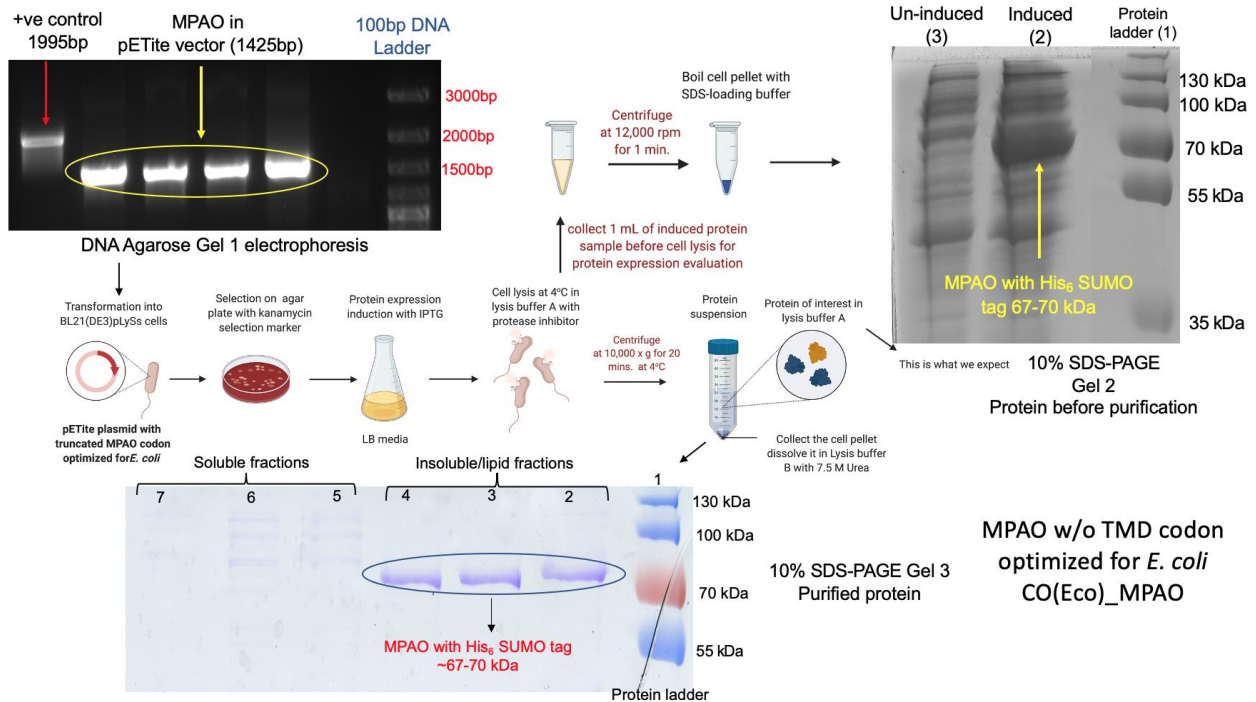
The urea-based lysis buffer was used to resuspend the cell pellet in the cue of solubilizing the insoluble protein fractions at higher pH [Singh, A. et al., 2015]. In all these steps the quantity of protein recovered was utilized to carry out the *in vitro* bioconversion assays of MPA to UA catalyzed by pure MPAO protein.

A)





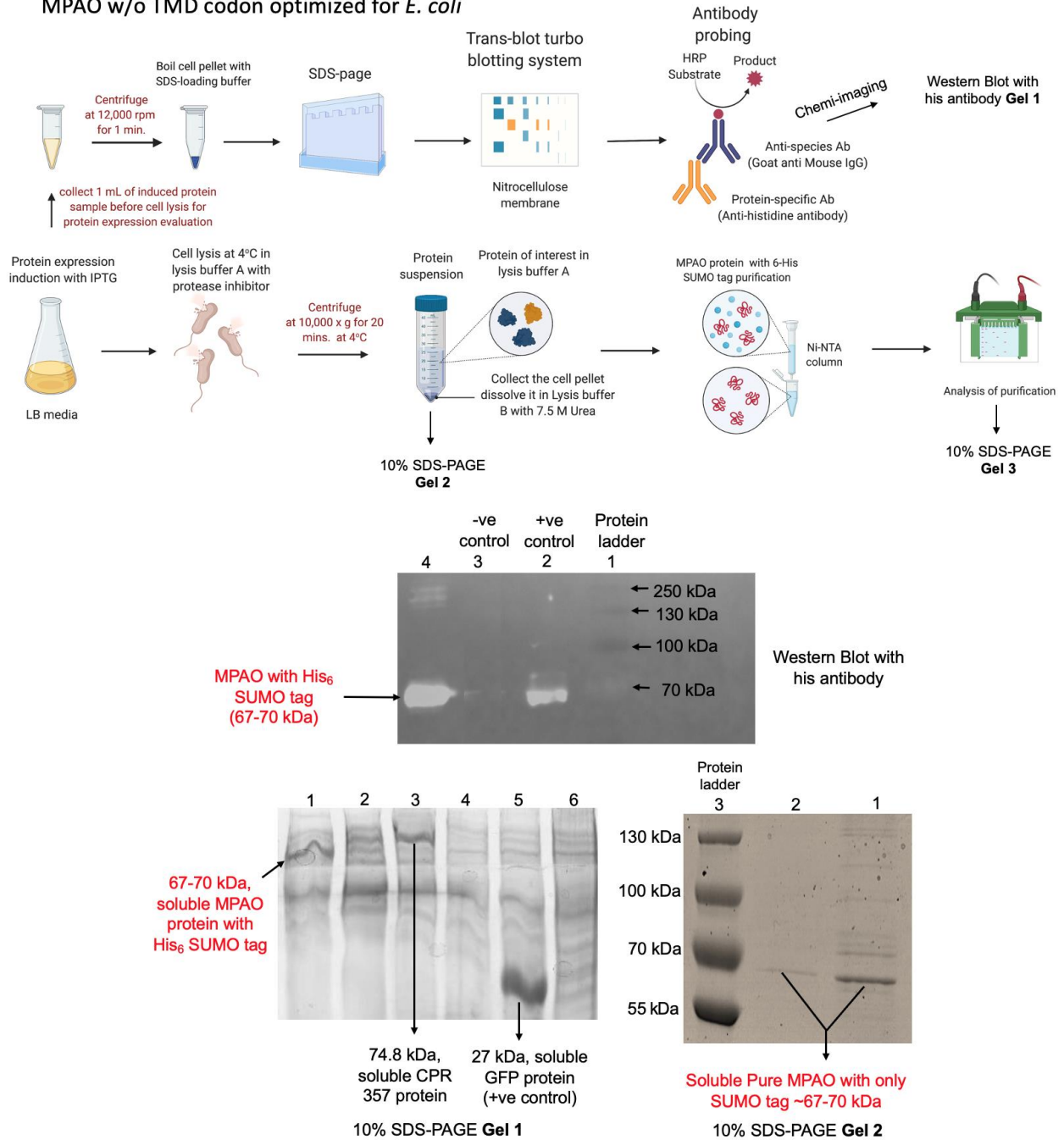
B)



**Figure 5.6:** Cloning and heterologous expression of MPAO codon-optimized for *E. coli* without transmembrane-domain (TMD) using pETite N-His SUMO vector. **A)** Pictorial representation of a membrane CYP with TMD anchored in lipid bilayer and CYP after TMD truncation (w/o TMD) in blue named as CO(Eco)\_MPAO], PCR amplification of CO(Eco)\_MPAO for In-fusion cloning in pETite vector; snapgene created pETite vector map with MPAO; **B)** Colony PCR for CO(Eco)\_MPAO clones with right size band around 1425 bp and band for +ve control at 1995 bp (**1% DNA agarose Gel 1**); Protein expression was induced with IPTG and screened with **10% SDS-PAGE Gel 2**, lane 1: protein ladder, lane 2: a thick right size band at about 67-70 kDa of MPAO protein and lane 3: un-induced sample (-ve control); Samples ran on **Gel 3 (10% SDS-PAGE)** were collected after bacterial cell lysis with sonication, lane 1: protein ladder, lanes 2-4: protein (MPAO ~67-70 kDa) was detected in lipid fractions, lanes 5-7: no protein was detected in soluble fractions.



MPAO w/o TMD codon optimized for *E. coli*



**Figure 5.7:** Screening of soluble expression of MPAO (Gel 1), purified MPAO protein with SDS-PAGE Gel 2 and Western Blot. Western blot using his antibody to detect His<sub>6</sub> SUMO tagged MPAO, lane 1: protein ladder, lane 2: -ve control and lane 3: +ve control with 69 kDa protein band, lane 4: CO(Eco)\_MPAO with 67-70 kDa band; Gel 1 is the SDS-PAGE gel where lane 1: the soluble fraction of MPAO protein (~67-70 kDa) after cell lysis, lane 2: un-induced sample of MPAO protein (-ve control), lane 3 and 5 are +ve control soluble fractions with bands at ~74.8 kDa (CPR357 protein) and 27 kDa (GFP protein), lanes 4 & 6 are un-induced samples of CPR357 and GFP proteins (-ve controls); Gel 2 is the SDS-PAGE gel where lanes 1 & 2: fractions of pure MPAO protein (~67-70 kDa), lane 3 is a protein ladder.

## 5.12. Attempt towards bioconversion of MPA to UA catalyzed by MPAO

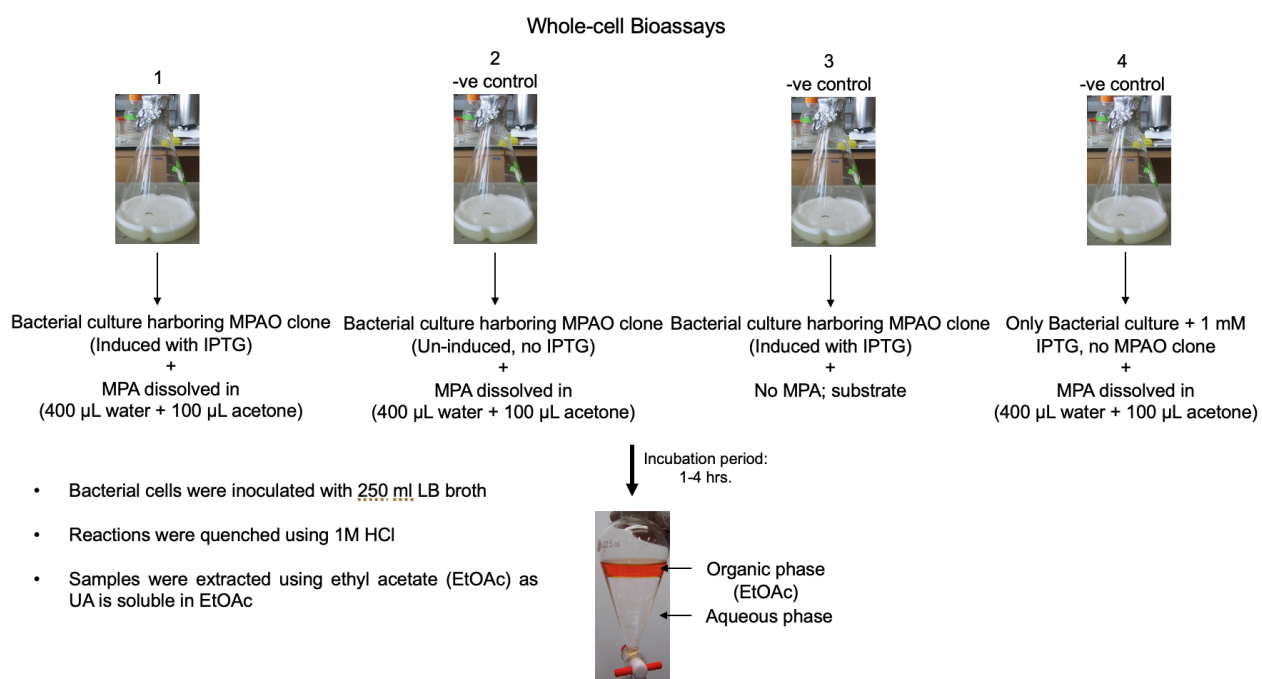
Whole-cell bioconversion experiment (*in vivo*) of MPA to UA was performed by incubating MPA (substrate) with *E. coli* transformants expressing MPAO protein. Protein was induced with 1 mM IPTG and three -ve control experiments (2-4) were set in parallel to experiment 1 (**Figure 5.8**). After 1 h of protein induction at 37 °C, 25 mL of the culture was collected from each flask, acidified using 1 M concentrated HCl to quench the reaction and the product was extracted using 25 mL of ethyl acetate solvent. Combined ethyl acetate fractions were evaporated using rotovap and dried over sodium sulphate (Na<sub>2</sub>SO<sub>4</sub>) and stored at 4 °C. The same procedure was repeated after 2 h of protein induction period (25 mL culture collected from the remaining bacterial culture, followed by extraction & sample preparation). The rest of the 200 mL culture was harvested after 3 h of protein induction and extracted in portions with 25 mL each, following the same steps as for earlier extractions (**Figure 5.8**).

### 5.12.1. HPLC analysis

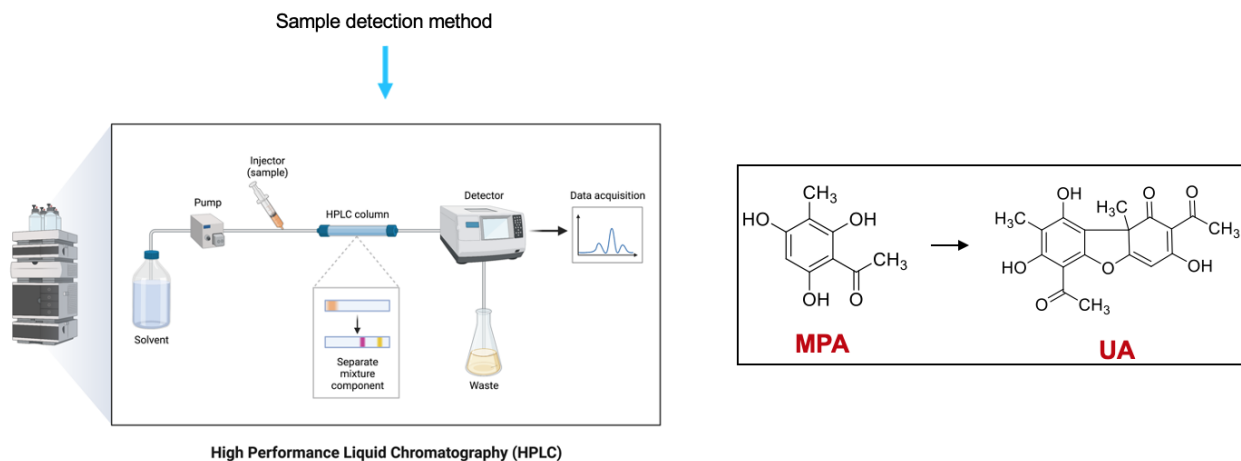
Experiments were done in triplicates and all the samples were dissolved in 1 mL of HPLC-grade methanol to screen for UA using UV coupled HPLC detector (**Figure 5.9**) (HPLC method details are discussed in **Section 2.5.17 of Chapter 2**).

Comparing the HPLC chromatogram of the extract from experiment 1 with that of standard UA (with a retention time 20-22 min and a UV spectrum with distinct bands at 231-232 and 282-284 nm), there was no UA in the experimental extract (**Figure 5.10**). Protein expression was tested with 10% SDS-PAGE gel and the MPAO protein was observed around ~67-70 kDa. The substrate, MPA peak was absent in the chromatogram of experiment 1 when compared to the standard MPA, has a retention time ( $t_R$ ) of 7.3 min & sometimes it elutes at 9.3 min and UV absorption band at wavelength 293 nm. Three possible conclusions could be made based on the HPLC results. The

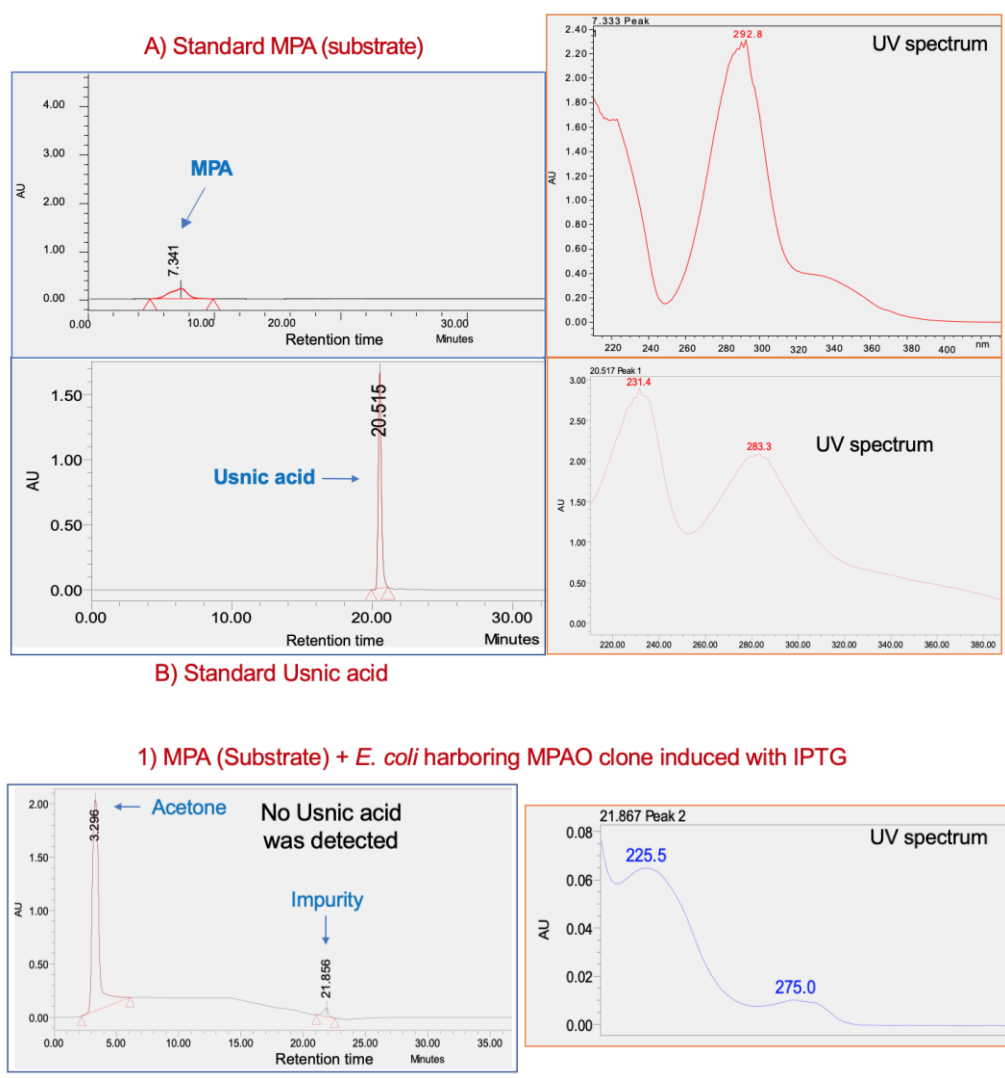
first conclusion was that the MPA degrades during the experiment, the second was that the MPA gets converted to some unknown compound catalyzed by a native bacterial protein or MPAO & the third was that the MPAO catalyzes MPA to something else but not UA. The first two possibilities were ruled out because of the presence of the MPA peak in the control experiments 2 & 4 (**Figure 5.10**). So, the only possible conclusion from these results was that the MPA is permeable to bacterial membrane and gets converted to some unknown compound but not UA, catalyzed by MPAO



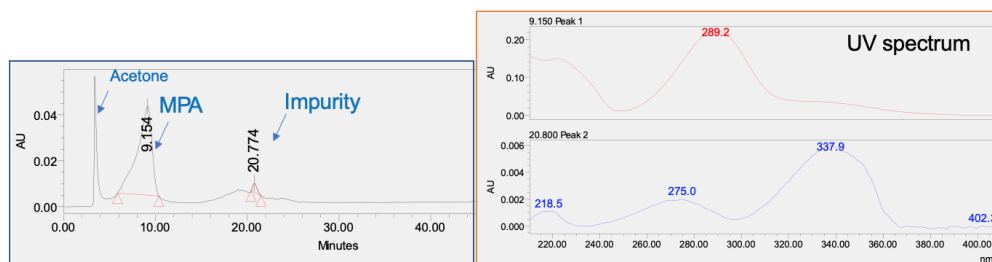
**Figure 5.8:** Whole-cell bio-conversion assays of MPA to UA catalyzed by MPAO with His<sub>6</sub> SUMO tag. **1.** MPA (substrate) dissolved in water + acetone, was incubated with *E. coli* transformants expressing MPAO; **2.** MPA was incubated with *E. coli* transformants but did not induce MPAO expression with IPTG; **3.** *E. coli* transformants expressing MPAO but did not add MPA; **4.** MPA was incubated with Only BL21 (DE3)pLySs cells + IPTG. The incubation period was 1-4 h and samples were collected at different time intervals, quenched with 1M HCl and extracted with EtOAc.



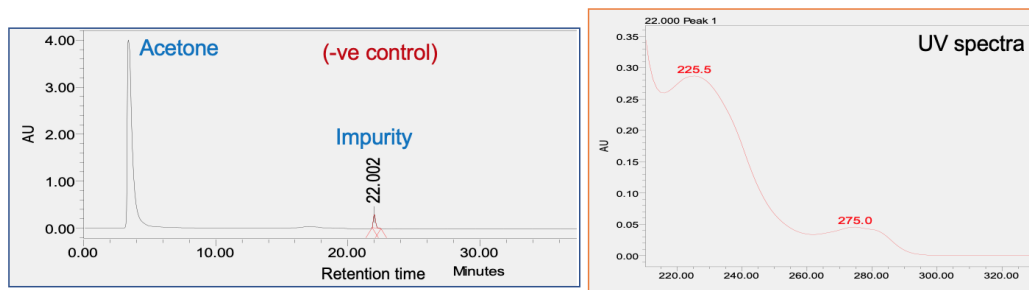
**Figure 5.9:** Pictorial representation of sample analysis method, HPLC coupled UV detector for whole-cell bioassays.



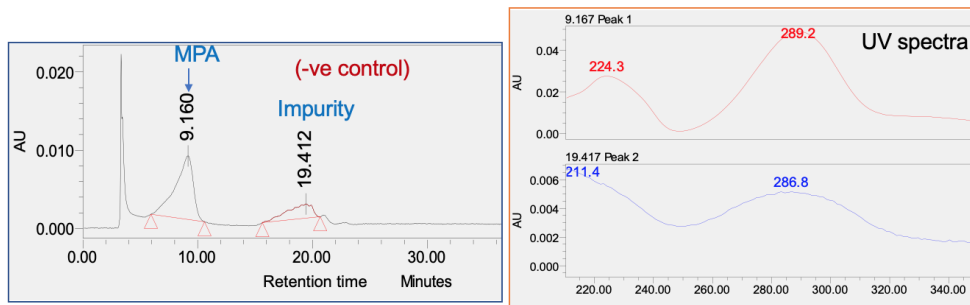
### 2) MPA (Substrate) + *E. coli* harbouring MPAO clone un-induced (No IPTG)



### 3) No MPA + *E. coli* with induced MPAO



### 4) MPA + Only *E. coli* in LB with kanamycin + 1 mM IPTG (No MPAO protein)



**Figure 5.10:** HPLC traces along with UV spectra at 254 nm wavelength of bioconversion experiment of MPA to UA catalyzed by whole-cell *E. coli* catalysts. **A)** MPA in 1 mL MeOH; **B)** UA in 1 mL MeOH; **1)** MPA incubated with *E. coli* transformant expressing MPAO; **2)** MPA incubated with *E. coli* transformant, did not induce MPAO expression (no IPTG); **3)** Only *E. coli* transformant expressing MPAO, no substrate; **4)** MPA + Only *E. coli* cells (no transformation with MPAO).

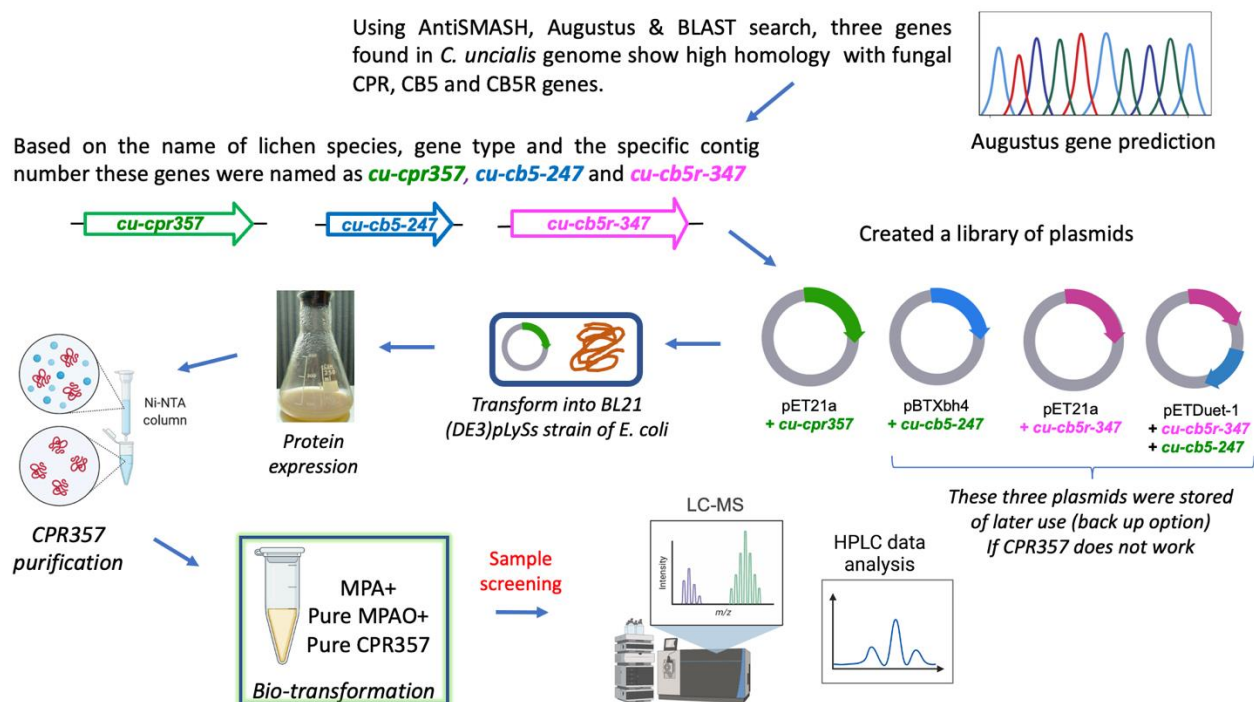
## 5.13. Cloning and heterologous expression of Redox proteins

No functional catalytic activity of truncated MPAO with His<sub>6</sub> SUMO tag (codon-optimized for *E. coli*) was observed in whole-cell biotransformation experiments. The next strategy was to clone and express CYP redox partners CPR, CB5 and CB5R in *E. coli*, and purify those proteins, and then to perform a biotransformation with the entire MPAO/reductase system *in vitro*.

As mentioned earlier in **Section 5.8**, the efficacy of CYP catalyzed reactions also relies on the abundance of their RPs and their electron transfer compatibility with CYPs [Braun, A. et al., 2012; Sono, M. et al., 1996]. Compared to the large and diverse CYPomes, most fungi possess only one or two CPRs, except four putative CPRs in found *F. oxysporum* [Durairaj, P. et al., 2016; Durairaj, P. et al., 2015; Lah, L. et al., 2011; Lah, L. et al., 2008] The reconstitution of recombinant eukaryotic CYP with the RP belonging to three major sources of origin: endogenous CPR (eukaryotic expression host's native CPR), homologous (the same source of recombinant CYP origin) and heterologous CPR (from yeast or other fungal strains). There are many examples of fungal p450s that are catalytically active when co-expressed with heterologous CPR, where expression profiles of CPRs are well established [Ide, M. et al., 2012; 194; Nazir, K. H. et al., 2011; Hirose, S. et al., 2011]. But research has shown that the electron transfer compatibility and coupling efficiency of homologous CYP-CPR interactions are relatively higher [Durairaj, P. et al., 2015, Braun, A. et al., 2012; Jennewein, S. et al., 2005]. Also, the substrate specificity of fungal CYPs was significantly influenced by the source of reductase. CYP539A7 and CYP655C2 proteins produced in higher yields with their homologous redox partners compared to the heterologous CPRs [Durairaj, P. et al., 2015]. We found three genes in the genome of *C. uncialis* using Augustus (gene prediction program) that are highly homologous to the known fungal CPRs, CB5s and CB5Rs. These genes were named, *cu-cpr357*, *cu-cb5-247* &, and *cu-cb5r-347* based on the name of the lichen species, the function of the gene itself and the contig number in *the C. uncialis* genome (**Figure 5.11**). Genes *cu-cpr357*, *cu-cb5-247* &, *cu-cb5r-347* encode for proteins CPR357, CB5247 &, CB5R347.

A library of plasmids was generated (**Figure 5.11**), where *cu-cpr357* was cloned in pET21a, *cu-cb5-247* in pET21c, *cu-cb5r-347* in pBTXbh4, both *cu-cb5-247* & *cu-cb5r-347* in a

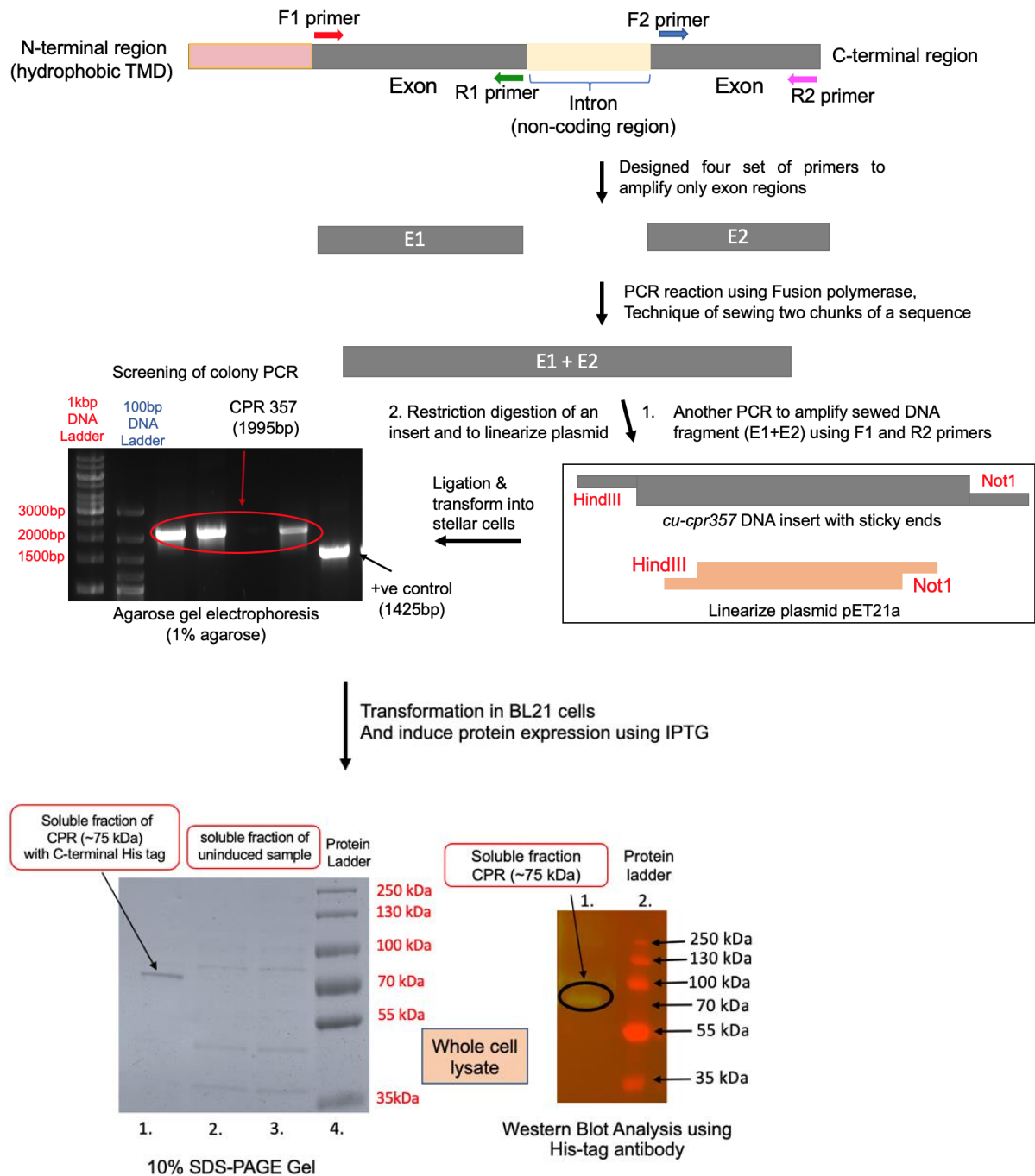
single plasmid pETDuet-1. Earlier mentioned in **Section 5.8** that in most cases only CPR carries both electrons from NADPH/NADH to CYP. So, the first strategy was to only express CPR357 in *E. coli* and to perform bioconversion experiments of MPA to UA catalyzed by both CPR357 & MPAO (**Figure 5.11**). Plasmids with *cu-cb5-24* & *cu-cb5r-347* were kept as backup options if CPR alone did not transfer both electrons to CYP from NADPH/NADH.



**Figure 5.11:** Schematic representation of identification of CPR, CB5 and CB5R (redox proteins) homologs in *C. uncialis*, creation of a library of plasmids with these redox proteins and the set of experiments performed with only CPR357 to get UA from MPA.

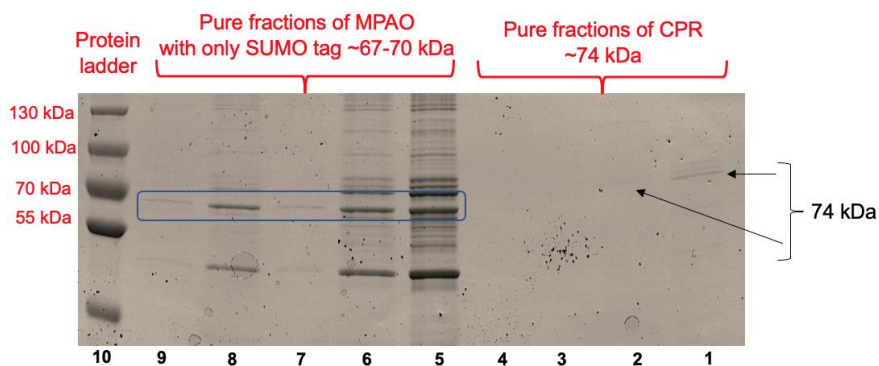
Successful cloning of *cu-cpr357* was confirmed with colony PCR results (1995 bp, 1% DNA agarose gel) and the soluble expression of CPR357 protein in *E. coli* was confirmed with both, 10% SDS-PAGE gel and Western blot (**Figure 5.12**). Purified proteins (both CPR357 & MPAO) were confirmed with 10% SDS-PAGE gel (**Figure 5.13**).





**Figure 5.12:** Pictorial description of primer design to remove introns manually for *cu-cpr357* gene and of sewing technique to conjoin exon fragments using PCR, amplification of *cu-cpr357* w/o (PCR), cloning into pET21a and transformation into stellar cells. Results of colony PCR (DNA agarose gel) by testing three colonies of stellar cells show the right size band at 1995 bp for *cu-cpr357*, the first two lanes are 1 kb and 100 bp DNA ladders and the last lane is the +ve control with a band at 1425 bp. Protein expression results screened with SDS-PAGE gel, lane 1: soluble fraction of induced CPR357 protein with a right size band around 75 kDa, lane 2, 3: soluble fractions of uninduced protein (CPR357) samples, lane 4: protein ladder and Western blot (his antibody), lane 1: band lights up around 75 kDa (C-terminus his tag) for CPR357 and lane 2: protein ladder.



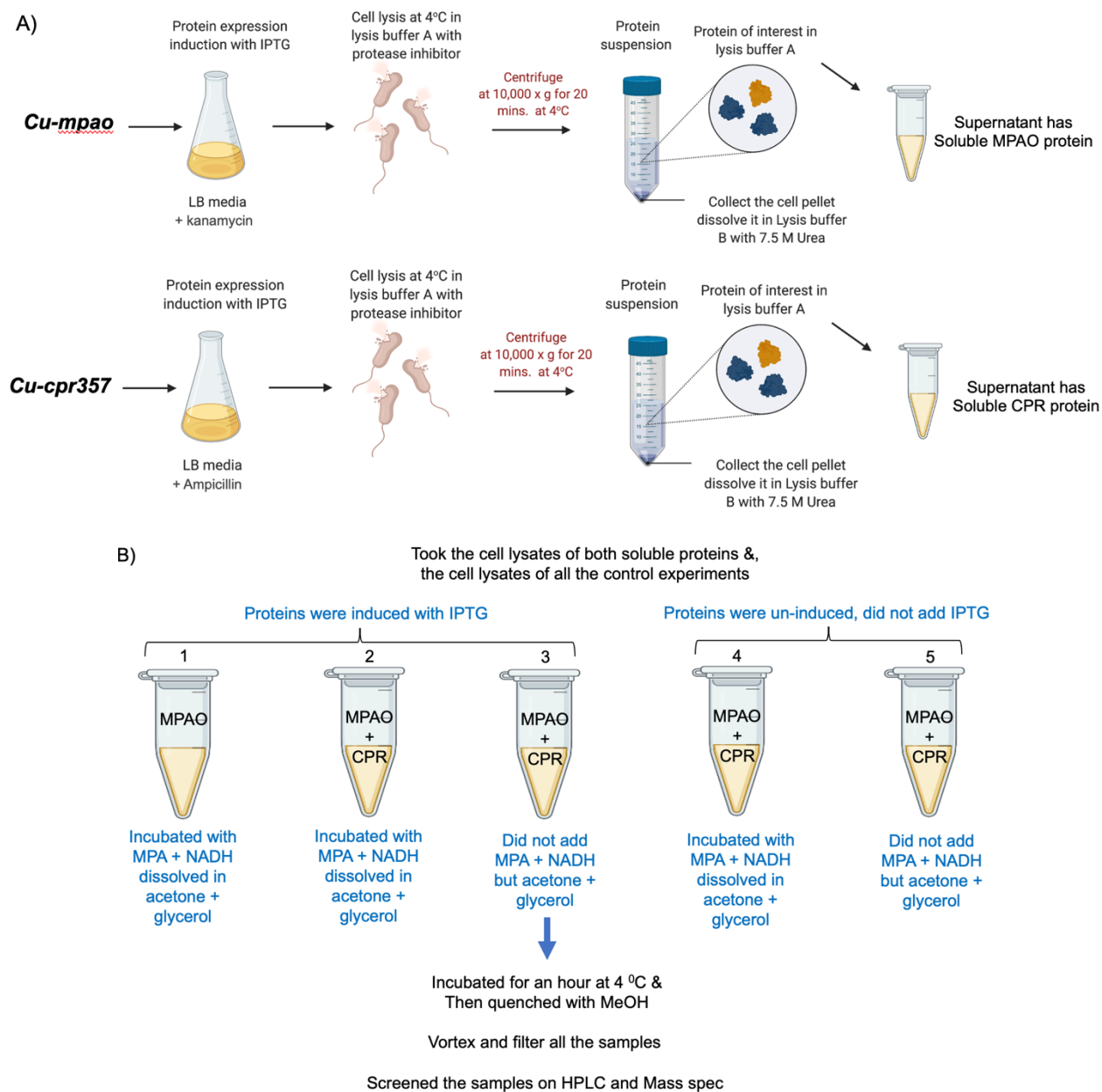


Purified proteins were separated on a 10% SDS-PAGE gel and stained with Coomassie blue stain

**Figure 5.13:** Purified MPAO and CPR357 were screened with 10% SDS-PAGE gel. Lanes 1-4: pure fractions of CPR357 proteins (1<sup>st</sup> two fractions show a band around 75 kDa); Lanes 5-9: pure fractions of MPAO protein with a right size band around 67-70 kDa; Lane 10: protein ladder.

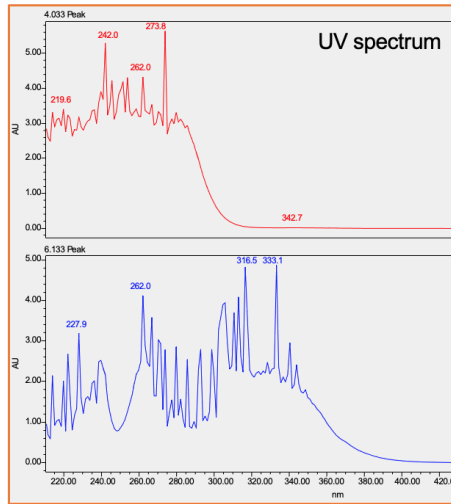
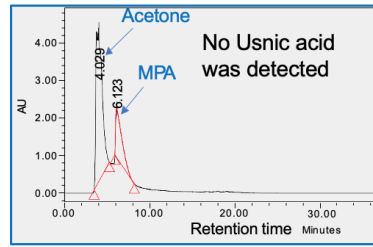
#### 5.14. Attempt towards bioconversion of MPA to UA catalyzed by both MPAO & CPR357

The concentration of pure CPR357 protein was very low, so, the bioconversion experiments were carried out using cell lysates of *E. coli* expressing CPR357 & MPAO instead of using pure proteins (**Figure 5.14A**). Control experiments (2-5) were set up in parallel to avoid any false positives (**Figure 5.14B**). Incubated MPA (substrate) and NADH (electron donor) with CPR357 & MPAO in lysis buffer at 4 °C for an hour. Reactions were quenched with a drop of conc. HCl, and samples were filtered using 0.45-micron filters. Samples were, dissolved in HPLC-grade MeOH to run on HPLC analysis and dissolved in acetonitrile to run on LC-MS (**Figure 5.14B**). Based on HPLC and LC-MS data, there was no trace of UA in HPLC data of experiments 1 & 2 (**Figure 5.15**), when compared with the standard of UA (**Figure 5.10B**). But the presence of MPA peak in both these experiments suggests that neither MPAO itself nor a combination of MPAO & CPR357 catalyzes MPA to UA. It also contradicts the hypothesis proposed in **Section 5.12** that MPA is converted to some unknown compound *in vivo*.

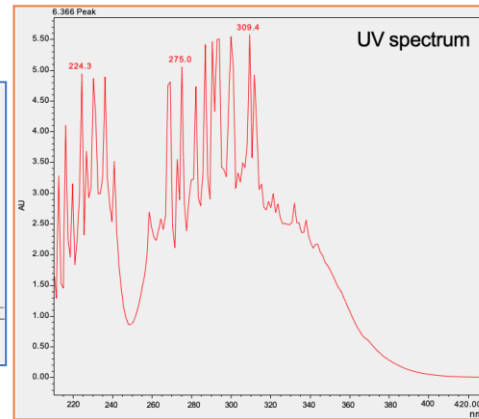
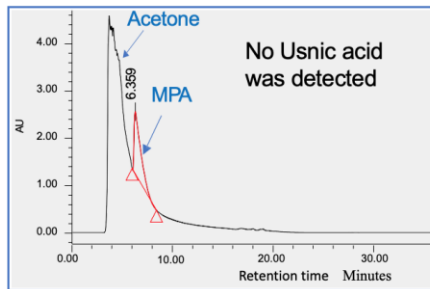


**Figure 5.14:** Bio-conversion experiments of MPA to UA catalyzed by both MPAO and CPR357 (cell free extracts, *in vitro*). **A)** Both MPAO and CPR357 proteins were expressed in parallel (In separate *E. coli* cultures) and soluble fractions of both proteins were collected after cell lysis; **B)** Bioassays (*in vitro*) by taking cell-free extracts **1.** MPAO (in lysis buffer A) + MPA + NADH (electron donor), both were dissolved in acetone & glycerol, **2.** MPAO + CPR (in lysis buffer A) + MPA + NADH; **3.** MPAO + CPR (in lysis buffer A) but did not add MPA & NADH (-ve control) just acetone + glycerol, **4.** Did not induce MPAO & CPR proteins (no IPTG), added MPA & NADH (-ve control), **5.** Just *E. coli* cells harboring MPAO & CPR clones (no protein induction, no substrate & electron donor) + acetone & glycerol (-ve control).

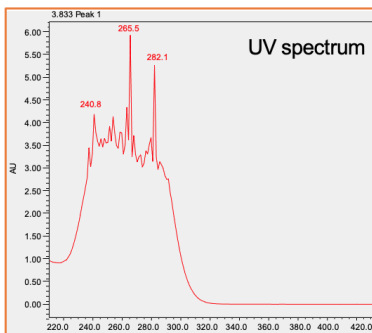
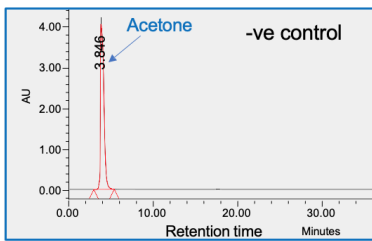
1) MPAO protein (in lysis buffer A) incubated with MPA + NADH in acetone + glycerol



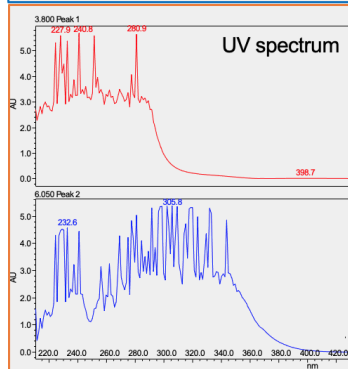
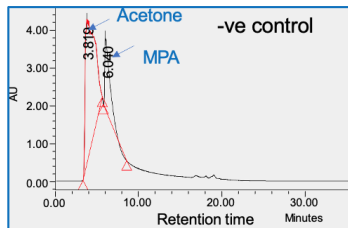
2) MPAO protein + CPR protein (in lysis buffer A) + MPA (substrate) + NADH in acetone + glycerol



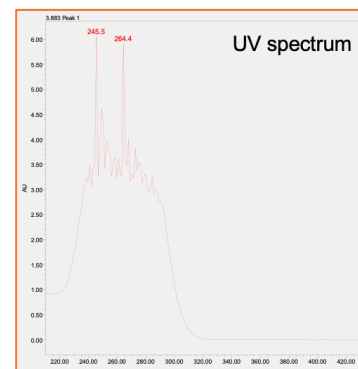
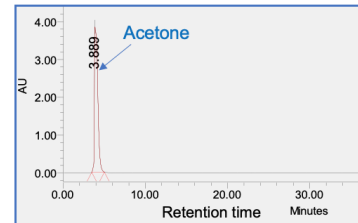
3) MPAO + CPR (induced proteins)  
No MPA + no NADH



4) MPAO (uninduced) + CPR (uninduced)  
+ MPA + NADH



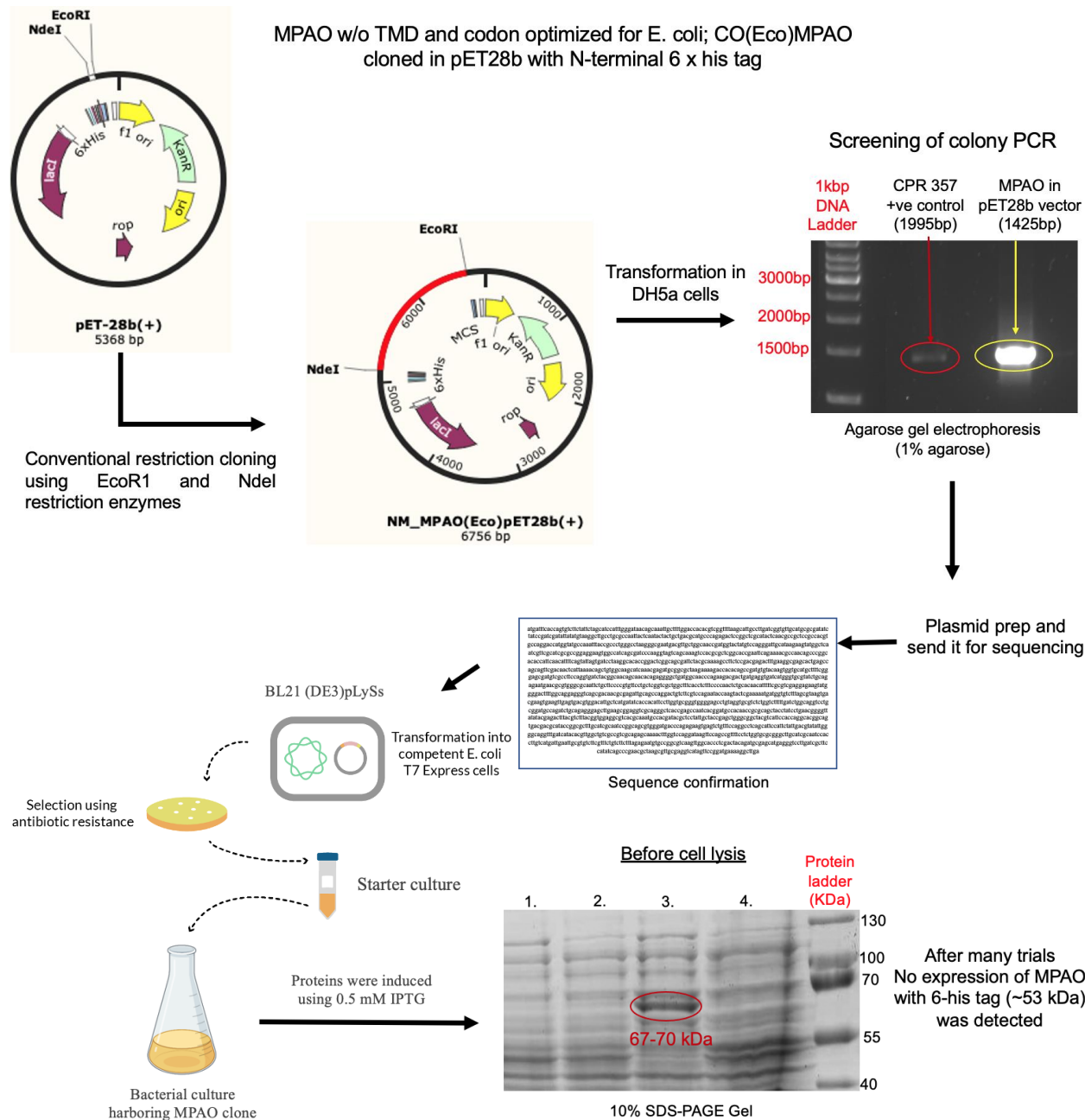
5) MPAO (uninduced)+CPR (uninduced)+  
NO MPA+ NO NADH



**Figure 5.15:** UV coupled HPLC analysis of bioconversion experiments of MPA to UA catalyzed MPAO only, MPAO+CPR both. **1)** MPAO (in lysis buffer A) + MPA + NADH (electron donor), both were dissolved in acetone & glycerol; **2)** MPAO + CPR (in lysis buffer A) + MPA + NADH; **3)** MPAO + CPR (in lysis buffer A) but did not add MPA & NADH (-ve control) just acetone + glycerol; **4)** Did not induce MPAO & CPR proteins in lysis buffer A (no IPTG), added MPA & NADH (-ve control); **5)** Just *E. coli* cells harboring MPAO & CPR clones in lysis buffer A (no protein induction, no substrate & electron donor) + acetone & glycerol (-ve control).

### **5.15. Attempts toward the heterologous expression of MPAO using pET28b vector**

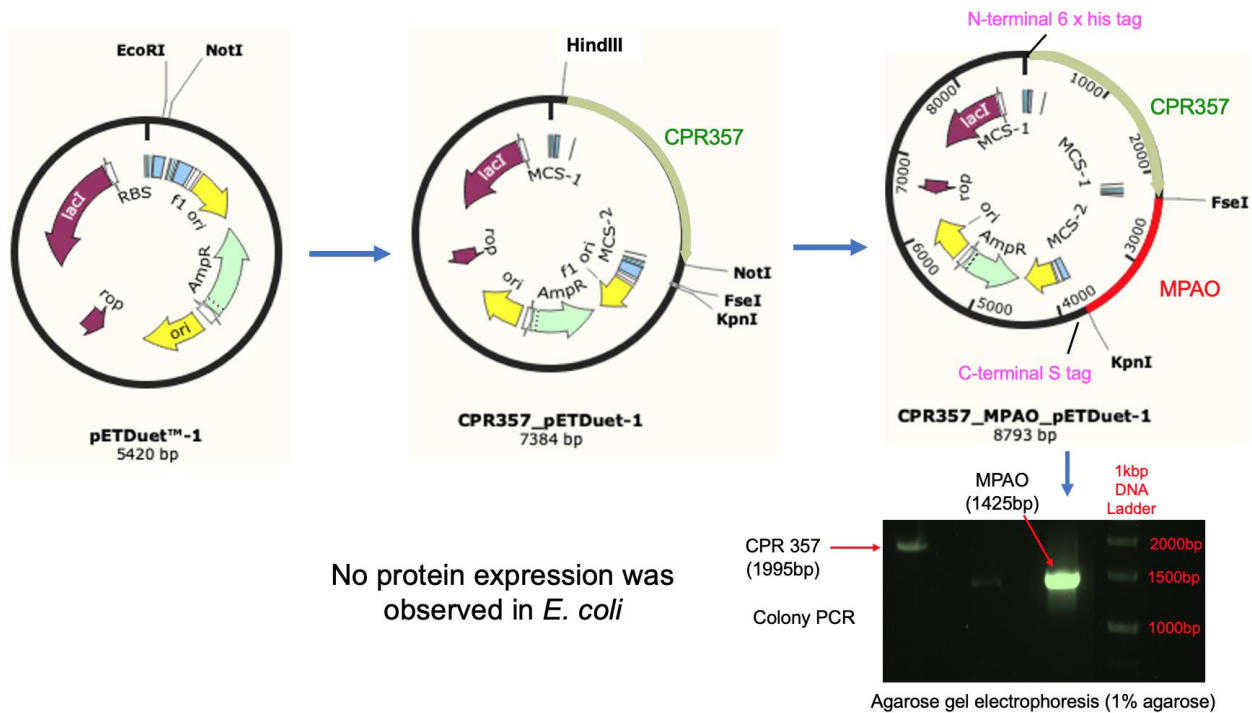
After 3-4 unsuccessful attempts of quantifying purified CPR357 proteins, and bioconversion experiments of MPA to UA using both MPAO and CPR357 proteins, the next strategy was to clone MPAO (w/o TMD &, codon-optimized for *E. coli*) in pET28b vector (N-terminal His<sub>6</sub> tag). We decided to clone the MPAO gene into a pET28b vector which has only an N-terminal His<sub>6</sub> tag contrary to a 12.7 kDa His<sub>6</sub> SUMO tag in pETite vector systems. The cloning of pET28b with MPAO in DH5a cells; a band at 1425 bp resulting from a colony PCR was successful and the sequence was confirmed (**Figure 5.16**). Expression of MPAO protein, induced with 0.5 mM IPTG in *E. coli* was tested with SDS-PAGE gel and nothing was detected (**Figure 5.16**). Many expression trials were performed using different temperatures, media conditions, IPTG concentrations and the varied number of induction periods. All attempts failed to give any visible protein expression.



**Figure 5.16:** Schematic representation of cloning and heterologous expression of MPAO using pET28b vector system in *E. coli*. Traditional restriction digestion cloning of codon-optimized MPAO w/o TMD using EcoRI & NdeI restriction enzymes in pET28b with N-His tag, colony PCR result (1% DNA agarose gel) shows right size band at 1425 bp of MPAO clone, lane 1: 1 kb DNA ladder, lane 2: +ve control (1995 bp); Sequence confirmation of MPAO, transformation in BL21 (DE3)pLySs cells and protein induction with 0.5 mM IPTG, protein expression was tested with 10% SDS-PAGE gel, lane 1: induced MPAO protein in pET28b, lane 2: uninduced MPAO protein in pET28b, lane 3: induced MPAO protein in pETite vector (+ve control) with a right size band around 67-70 kDa, lane 4: uninduced MPAO protein in pETite vector, compared with a protein ladder in lane 5.

### 5.16. Attempts toward co-expression of MPAO & CPR357 in *E. coli*.

The theory of protein-protein interaction in the case of eukaryotic p450s with their RPs such as CPR & CB5 led us to clone both *cu-mpao* and *cu-cpr357* in a single vector. pETDuet-1 vector with two T7 promoters and two MCSs was used to co-express MPAO and CPR357 proteins in *E. coli*. Cloning was successful, results of colony PCR of DH5a cells harbouring pETDuet-1 with both *cu-mpao* (1425 bp) and *cu-cpr357* (1995 bp) are shown in **Figure 5.17**.

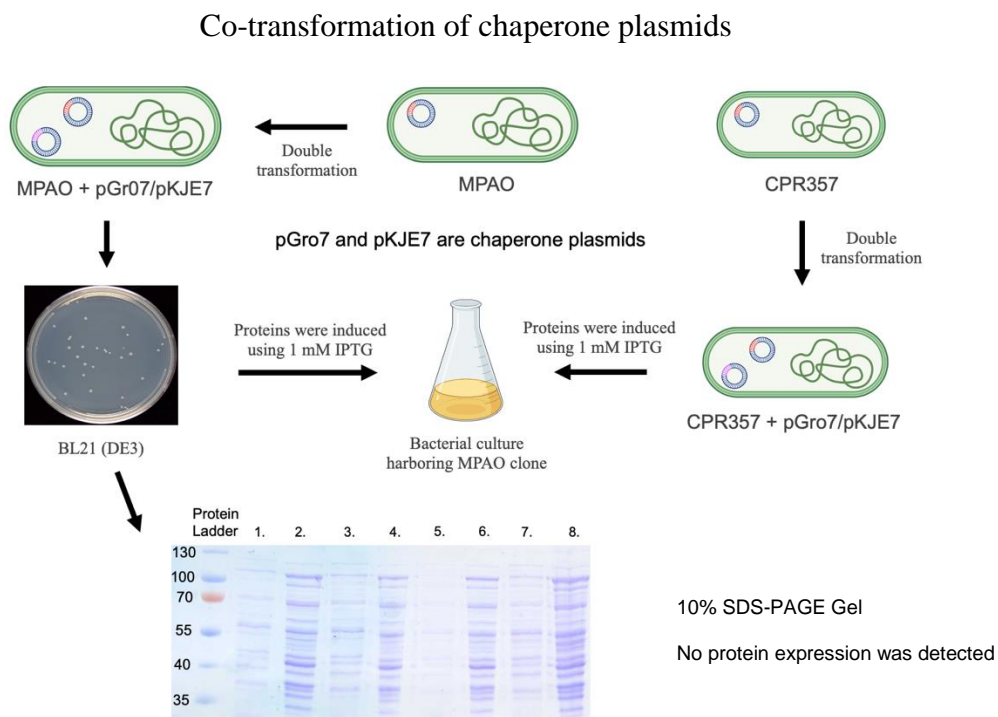


**Figure 5.17:** Cloning of both *cu-mpao* and *cu-cpr357* on a plasmid pETDuet-1 with N-terminus His<sub>6</sub> tag and C-terminus S tag. Colony PCR shows the right size band for *cu-mpao* (1425 bp) and *cu-cpr357* (1995 bp).

The Co-expression of MPAO and CPR357 using the pETDuet-1 vector system was unsuccessful. The next strategy was to co-express a chaperone protein with MPAO with His<sub>6</sub> tag (pET28b vector) in *E. coli* and this was to get a successful soluble expression of MPAO with His<sub>6</sub> tag. Chaperone proteins have been known to be used for the soluble expression of eukaryotic membrane proteins. As their name suggests a helper/caretaker, these proteins assist the



conformational folding of a membrane protein. Co-expression experiments of CPR357 with different chaperone proteins in *E. coli* were also performed to achieve high levels of expression of CPR357 protein (**Figure 5.18**). Several expression plasmids have been used in molecular biology to express these chaperone proteins as mentioned in **Section 5.7.3 (Table 5.1)**. Co-transformed strains of BL21 (DE3) were generated as: pGro7 (**Table 9**) + pET28b (*mpao*), pKJE7 (**Table 5.1**) + pET28b (*mpao*), pGro7+ pET21a (*cu-cpr357*), pKJE7 + pET21a (*cu-cpr357*). Co-expression of these proteins was tested with 10% SDS-PAGE gel (**Figure 5.18**). There was no successful protein expression was observed.



**Figure 5.18:** Steps illustrating double transformation of pET28b harboring MPAO & pGro7 (chaperone plasmid with GroEL/GroES chaperone proteins listed in Table 9) in *E. coli*, MPAO (pET28b) + pKJE7 (chaperone plasmid with DnaK/DnaJ/GrpE chaperone proteins listed in Table 9) in *E. coli*, CPR357 (pET21a) & pGro7 (chaperone plasmid with GroEL/GroES chaperone proteins listed in Table 9) in *E. coli*, CPR357 (pET21a) + pKJE7 (chaperone plasmid with DnaK/DnaJ/GrpE chaperone proteins listed in Table 9) in *E. coli* and co-expression of these proteins were tested with 10% SDS-PAGE gel, lane 1: Un-Induced samples of MPAO with 6-his tag + DnaK, DnaJ, GrpE, lane 2: Induced samples of MPAO with 6-his tag (53 kDa) + DnaK (70 kDa), DnaJ (40 kDa), GrpE (22 kDa), lane 3: Un-Induced samples of MPAO with 6-his tag + GroEL, GroES, lane 4: Induced samples of MPAO with 6-his tag (53 kDa) + GroEL (60 kDa), GroES (10 kDa), lane 5: Un-Induced samples of CPR357 with 6-his tag + DnaK, DnaJ, GrpE, lane 6: Induced samples of CPR357 with 6-his tag (75 kDa) + DnaK (70 kDa), DnaJ (40 kDa), GrpE (22 kDa), lane 7: Un-Induced samples of CPR357 with 6-his tag + GroEL, GroES, lane 8: Induced samples of CPR357 with 6-his tag (75 kDa) + GroEL (60 kDa), GroES (10 kDa).

## 5.17. Attempts toward the heterologous expression of Fungal p450s

When we were unsuccessful in demonstrating the function of MPAO protein, we decided to find the functional homolog of this lichen protein in fungus. The experimental plan was to take the known one or two fungal p450s that are homologous to MPAO and express those fungal p450s in *E. coli*. Furthermore, see if they could catalyze the conversion of MPA to UA when used in combination with the lichen CPR357 redox protein.

### 5.17.1 BLAST and phylogenetic analysis

Based on an extensive BLAST search (**Table 5.2**) and phylogenetic tree reconstruction (Maximum likelihood method; ML) (**Figure 5.19**), two already known fungal p450 proteins were found to be highly homologous to the lichen p450, MPAO. pBLAST results infer more than 95% score of A.A. query (MPAO protein) cover of fungal p450s found in *Penicillium/Talaromyces marneffi* and *Aspergillus nidulans* (**Table 5.2**). Using Clustal Omega, the multiple sequence alignment (MSA) data was computed (**Figure 5.20**) and used to construct a phylogenetic tree. The resultant phylogenetic tree showed that *T. marneffi* p450 and MPAO shared a close evolutionary relationship based on the high bootstrap support of 94% (**Figure 5.19**). Also, both proteins were clustered in one clade of the phylogenetic reconstruction (**Figure 5.19**). The evolutionary proximity of *A. nidulans* p450 and MPAO was supported with a more than average bootstrap % value (71). The topology of the phylogenetic tree also suggested that *A. nidulans* p450 & MPAO had evolved in a recent time frame based on the short horizontal evolutionary distance between them through the node. As **Chapter 6** discusses in detail, the premise of phylogenetics is that genes with close evolutionary relationships are more likely to share similar functions than ones with a distant common origin. Therefore, based on this theory, we predicted that the two fungal genes could be orthologous to MPAO. These fungal genes were named *an-stcb* (*A.*



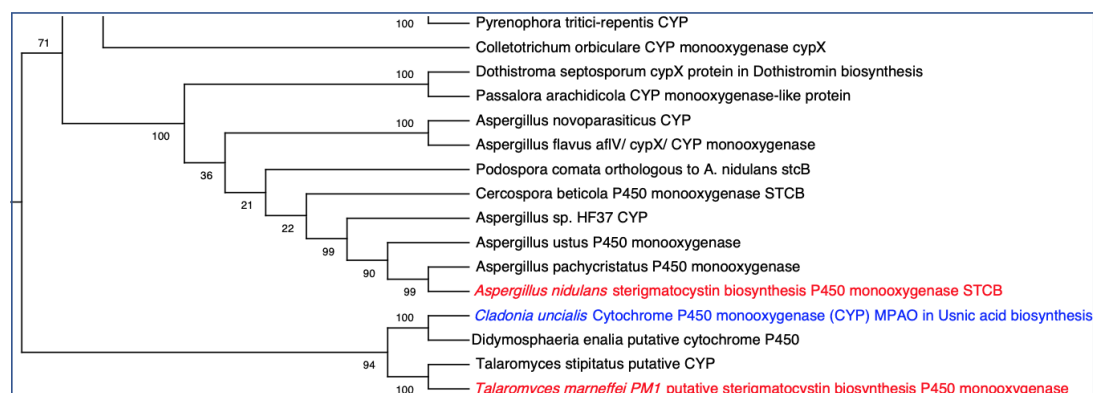
*nidulans* Sterigmatocystin B) and *pm1-scyp* (*Penicillium marneffi* 1 Sterigmatocystin cytochrome p450).

**Table 5.2:** BLAST statistics of *C. uncialis* MPAO protein sequence shows two fungal CYP homologs.

Blast results of MPAO

Description	Total Score	Query Cover	E value	Per. Identity	Accession no.
Methylphloracetophenone oxidase (MPAO) [ <i>Cladonia uncialis</i> ]	922	100%	0	100%	AUW31051.1
Hypothetical protein [ <i>Alectoria fallacina</i> ]	797	99%	0	84.91%	CAF9912111.1
Putative Sterigmatocystin biosynthesis p450 monooxygenase [ <i>Talaromyces marneffi</i> PM1]	386	95	3e-125	41.85%	XP_002145795.1
Probable Sterigmatocystin biosynthesis p450 monooxygenase STCB [ <i>Aspergillus nidulans</i> ]	327	90	3e-102	38.91%	XP_681093.1

Phylogenetic analysis of to find homologs of MPAO



**Figure 5.19:** Phylogenetic tree reconstruction of MPAO protein shows two fungal CYP orthologs; one found in *T. marneffi* with high bootstrap support of 94% present in the same group as MPAO (monophyletic) and the other found in *A.nidulans* with an avg. bootstrap support of 71% in a group other than MPAO (non-monophyletic) (PDF version of complete phylogenetic trees are in Appendix Figure S25).

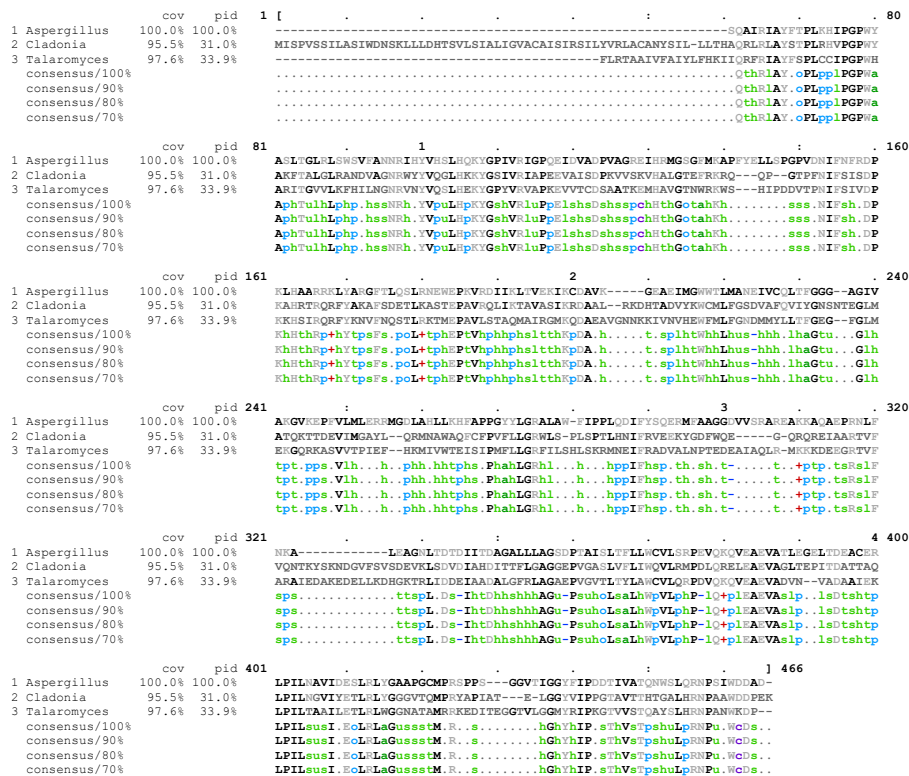
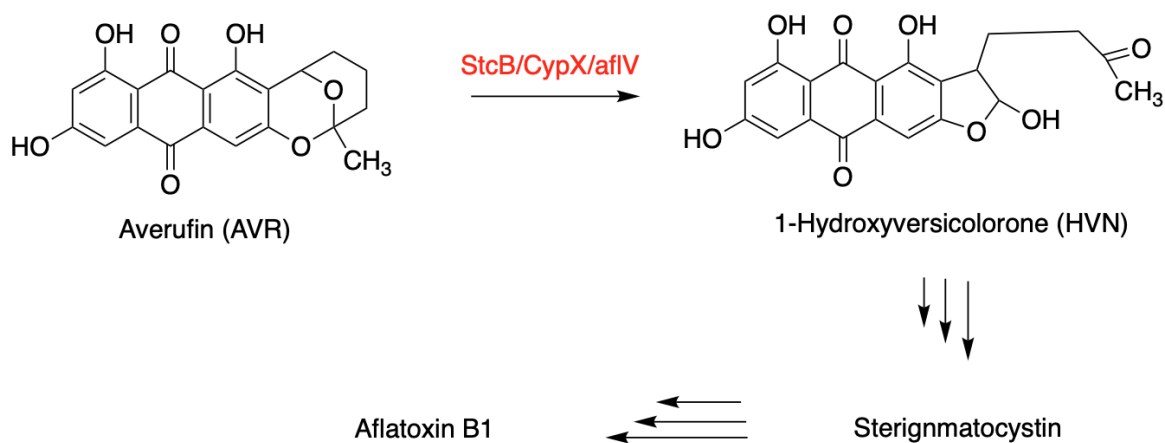


Figure 5.20: Multiple sequence alignment (Clustal Omega) of *A. nidulans* CYP, *C. uncialis* MPAO and *T. marneffii* CYP.

### 5.17.2. Oxidation reactions catalyzed by STCB

The fungal homolog of MPAO; StcB/CypX/afIV (*A. nidulans* cytochrome p450) is involved in aflatoxin biosynthesis [Keller, N. P. et al., 2000; Yu J. et al., 2004]. Where StcB catalyzes Baeyer-villager style oxidation reaction and converts averufin (AVR) to 1-hydroxyversicolorone (HVN) in aflatoxin biosynthesis (Scheme 5.1) [Henry, K. M. & Townsend, C. A. 2005]. However, no expression work of StcB has been reported so far. Gene disruption experiments of *cypx* (encodes CypX enzyme) were performed where a mutant *A. parasiticus* NIAH-26 strain was created, and accumulation of AVR was confirmed with TLC [Wen, Y. et al., 2005]. Feeding experiment of SM; AVR to *A. parasiticus* to analyze the enzymatic activity of StcB by measuring the formation of HVN with HPLC analysis [Yabe K. et al., 2003].

No experiments have been done so far regarding *pm1-scyp*. pBLAST search, MSA, and phylogenetic analysis show that *pm1-scyp* is highly homologous to *an-stcb*.

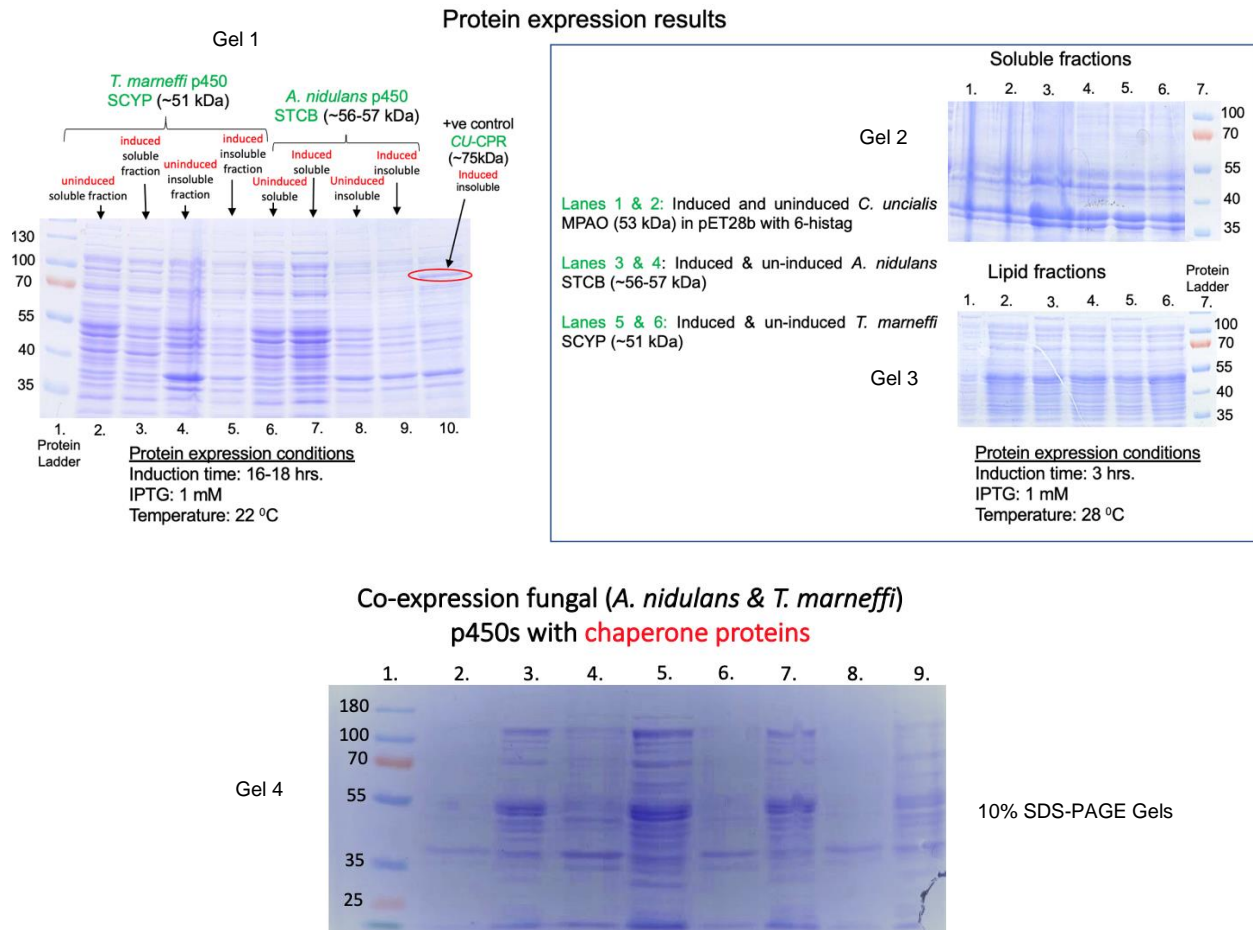


Wen Y. et al. *Applied and environmental microbiology*, 2005, 71(6), 3192–3198

**Scheme 5.1:** Schematic representation of Baeyer villager style oxidative cleavage of averufin catalyzed by StcB/CypX/aflV (*A. nidulans* cytochrome p450) during aflatoxin biosynthesis [Wen, Y. et al., 2005].

### 5.17.3 Heterologous expression of *an-stcb* & *pm1-scyp* in *E. coli*

Two custom synthesized plasmids, pET28a with *an-stcb* and *pm1-scyp* w/o introns, were ordered from Twist Bioscience (Biotech. Co. located in the Bay area, California). Both genes were not codon-optimized for *E. coli*, and in the case of *an-stcb*, TMD was truncated, but for *pm1-scyp*, TMD was intact. Both plasmids were transformed into BL21 (DE3)pLySs cells separately, and protein expression was induced by 1 mM IPTG. After many experiments and trials analyzed with SDS-PAGE gels, no protein expression was observed for STCB & SCYP proteins (**Gel 1, 2 & 3 Figure 5.21**). Co-expression of chaperone proteins was also attempted with STCB and SCYP, and all these attempts were unsuccessful (**Gel 4 Figure 5.21**).



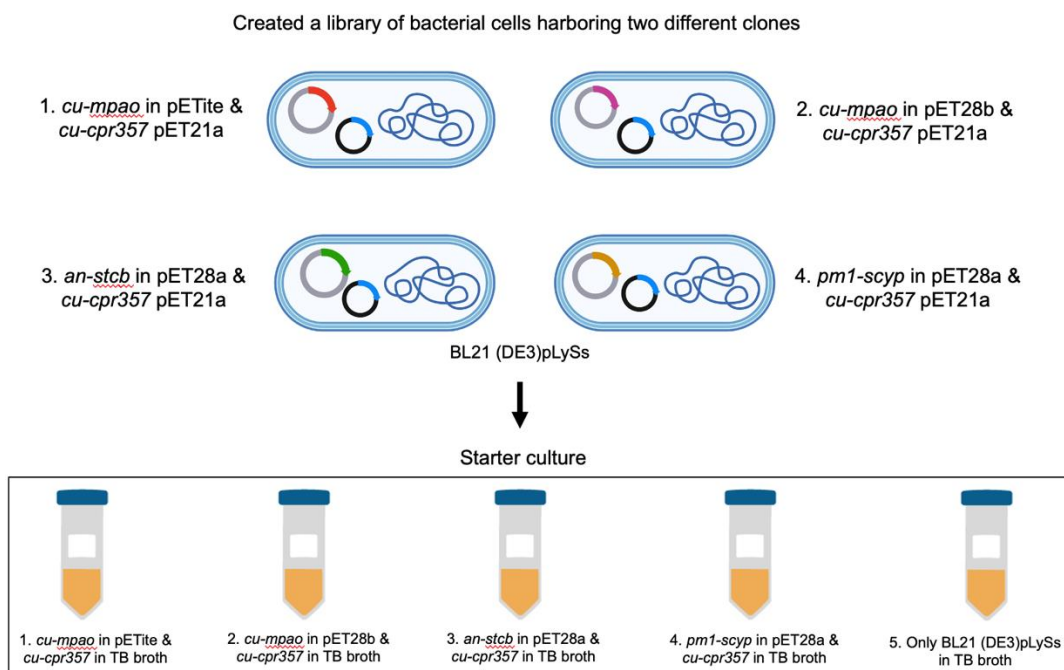
**Figure 5.21:** Results of heterologous expression of STCB and SCYP in *E. coli* were screened with 10% SDS-PAGE gels. **Gel 1**, lane 1: Protein ladder, lane 2: Soluble fraction of un-induced sample of SCYP protein, lane 3: Soluble fraction of induced sample of SCYP protein (expected band ~51 kDa), lane 4: lipid/insoluble fraction of un-induced sample of SCYP protein, lane 5: lipid/insoluble fraction of induced sample of SCYP protein, lane 6: Soluble fraction of un-induced sample of STCB protein, lane 7: Soluble fraction of induced sample of STCB protein (expected band ~56-57 kDa), lane 8: lipid/insoluble fraction of un-induced sample of STCB protein, lane 9: lipid/insoluble fraction of induced sample of STCB protein, lane 10: lipid/insoluble fraction of induced sample of CPR357 protein with a right size band around 75 kDa (+ve control). **Gel 2** has all soluble fractions and **Gel 3** has all lipid fractions, **Gel 4** lane 1: Induced MPAO with 6-his tag (expected band ~53 kDa), lane 2: Un-induced MPAO with 6-his tag, lane 3: Un-Induced samples of SCYP with 6-his tag + DnaK, DnaJ, GrpE, lane 4: Induced samples of SCYP with 6-his tag (51 kDa) + DnaK (70 kDa), DnaJ (40 kDa), GrpE (22 kDa), lane 5: Un-Induced samples of SCYP with 6-his tag + GroEL, GroES, lane 6: Induced samples of SCYP with 6-his tag (51 kDa) + GroEL (60 kDa), GroES (10 kDa), lane 7: Un-Induced samples of STCB with 6-his tag + DnaK, DnaJ, GrpE, lane 8: Induced samples of STCB with 6-his tag (~56-57 kDa) + DnaK (70 kDa), DnaJ (40 kDa), GrpE (22 kDa), lane 9: Un-Induced samples of STCB with 6-his tag + GroEL, GroES, lane 10: Induced samples of STCB with 6-his tag (~56-57 kDa) + GroEL (60 kDa), GroES (10 kDa).

### 5.18 Proteins co-expression experiments in *E. coli*

Several lines of evidence demonstrate the successful co-transformation and the persistence of two plasmids with the same origin of replication (ori) in the single cell of *E. coli* [Velappan, N. et al., 2007]. In molecular biology, plasmids with the same origin of replication are usually considered incompatible for co-transformation [Novick, R. P. 1987; Nordstrom and Austin, 1989; Austin and Nordstrom, 1990]. The only downside of using this technique is the time of persistence of these kinds of plasmids in a bacteria cell. Therefore, we did some experiments to test the persistence time of two plasmids with the same origin of replication in a bacterial cell. We found that such two plasmids cannot co-exist stably in a bacterial cell for more than two months, and one plasmid was lost after that period. The role of antibiotic resistance markers in the persistence of plasmids is not very clear yet, but it has been proved that a high copy plasmid dominates the cell after many growth cycles.

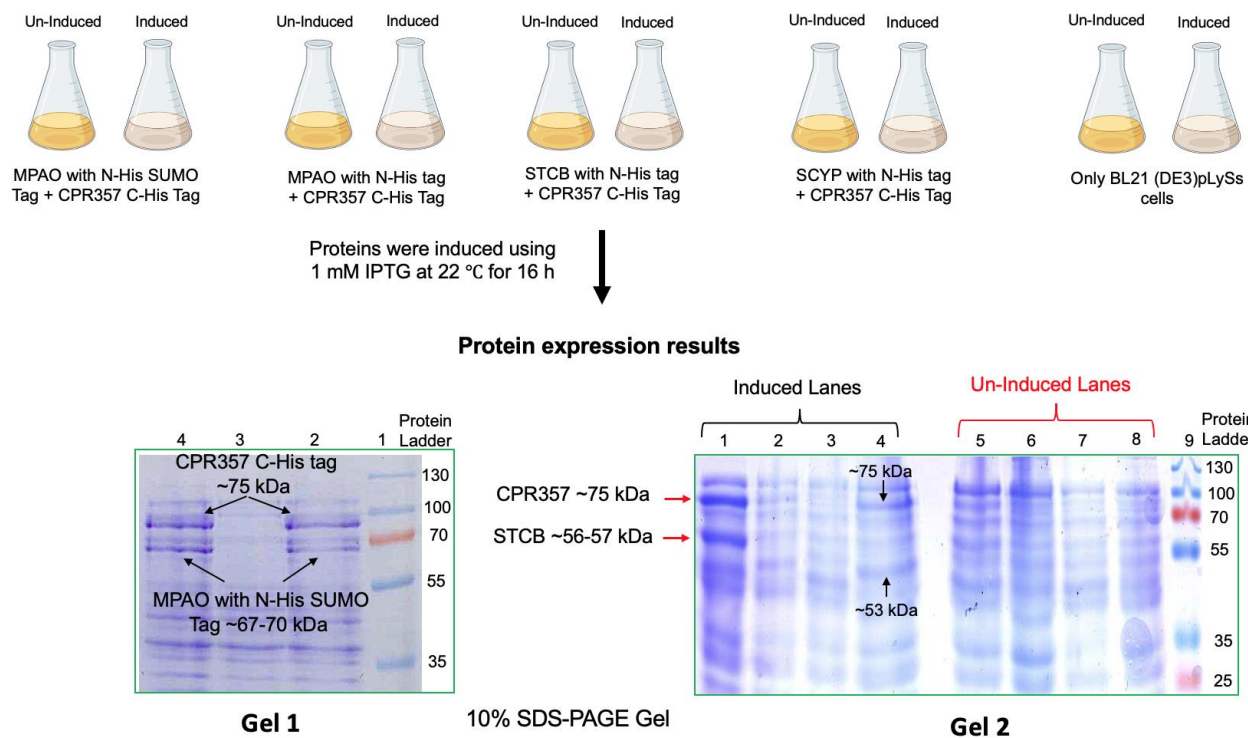
In the case of MPAO and CPR357, both pETite N-His SUMO & pET21a vectors have low copy pBR based origin of replication. So, the time of persistence of both plasmids in a single bacterial cell was quite significant and independent of the antibiotic choice. For the protocol of double transformation see **Section 2.5.7** of **Chapter 2**. The following experiments were performed simultaneously, pictorially represented in **Figure 5.22**. 1) Double transformation of *E. coli* with pETite N-His SUMO vector harboring *cu-mpao* clone followed by pET21a harboring *cu-cpr357* clone. 2) Double transformation of *E. coli* with pET28a (with only His<sub>6</sub> tag) vector harboring *cu-mpao* clone followed by pET21a harboring *cu-cpr357* clone. 3) Double transformation of *E. coli* with pET28a vector harboring *an-stcb* clone followed by pET21a harboring *cu-cpr357* clone. 4) Double transformation of *E. coli* with pET28a vector harboring *pm1-scyp* clone followed by pET21a harboring *cu-cpr357* clone. After successful transformation experiments and colonies

were screened for all these combinations of genes via colony PCR, the co-expression experiments started by setting up the starter cultures in BL21 (DE3) pLySs cells inoculated with LB media + selection marker (**Figure 5.22**). Proteins Co-expression was induced using 1 mM IPTG, and the induction period was quite long compared to previous experiments (**Figure 5.23**). The successful reproducible co-expression of MPAO with SUMO tag and CPR357 and the co-expression of STCB (fungal p450 from *A. nidulans*) and lichen fungal CPR357 were confirmed with SDS-PAGE gels (**Figure 5.23**). In the first experiment, co-expression of MPAO with His<sub>6</sub> tag and CPR357 was observed (**SDS-PAGE gel on the right, Figure 5.23**), but a similar result could not be reproduced in the second and third attempt. There was no co-expression detected in the case of SCYP (fungal p450 from *T. marneffi*) and lichen fungal CPR357 (**SDS-PAGE gel on the right, Figure 5.23**).



**Figure 5.22:** Double transformation of *cu-mpao* (N-His SUMO pETite vector) & *cu-cpr357* (C-His pET21a), *cu-mpao* (N-His pET28b vector) & *cu-cpr357* (C-His pET21a), *an-stcb* (N-His pET28a vector) & *cu-cpr357* (C-His pET21a), *an-scyp* (N-His pET28a vector) & *cu-cpr357* (C-His pET21a) in *E. coli*; BL21 (DE3)pLySs cells. Preparation of starter cultures with respective antibiotic-resistance markers for co-expression of all these combination of proteins in bacteria.





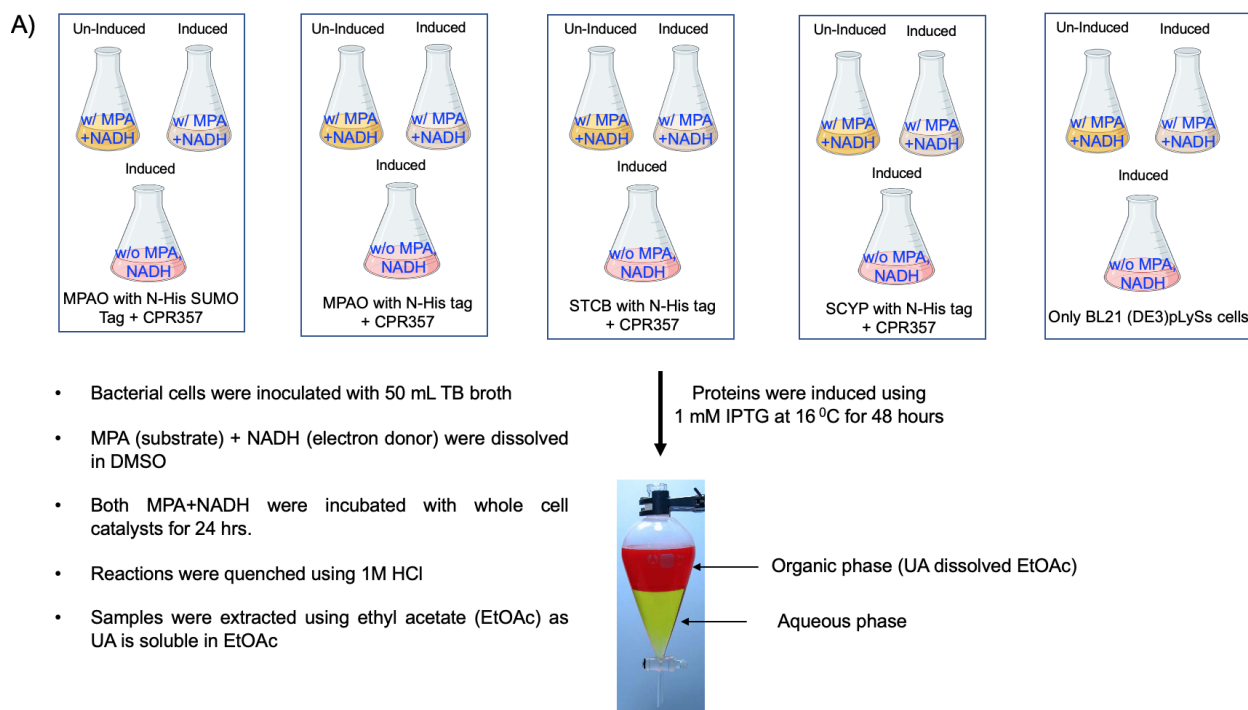
**Figure 5.23:** Results of Co-expression experiments screened with 10% SDS-PAGE gels. Baffled flasks with LB media + selection markers for the Co-expression of MPAO (N-His SUMO tag) & CPR357 (C-His tag), MPAO (N-His tag) & CPR357 (C-His tag), STCB (N-His tag) & CPR357 (C-His tag), SCYP (N-His tag) & CPR357 (C-His tag) proteins in bacteria were induced with IPTG and -ve controls (un-induced, no IPTG). **SDS-PAGE Gel 1** with lane 1 is a protein ladder, lanes 2 & 4: Induced samples of co-expressed MPAO (N-His SUMO tag ~67-70 kDa) & CPR357 (C-His tag ~75 kDa) proteins, lane 3: Un-induced samples. **SDS-PAGE Gel 2** lane 1: Induced co-expression of STCB (N-His tag ~56-57 kDa) & CPR357 (C-His tag ~75 kDa), lane 2: Induced co-expression of SCYP (N-His tag ~51 kDa) & CPR357 (C-His tag ~75 kDa), lane 3: Just BL21 cells with IPTG (-ve control), lane 4: Induced co-expression of MPAO (N-His tag ~53 kDa) & CPR357 (C-His tag ~75 kDa), lanes 5-8 are un-induced samples (-ve controls) in the same order as of lanes 1-4, lane 9: Protein ladder.

### 5.19. Bioconversion experiments of MPA to UA

Bioconversion experiments of MPA to UA were performed as follows: 1. MPA + NADH + MPAO (with His<sub>6</sub> SUMO tag) + CPR357. 2) MPA + NADH + MPAO with His<sub>6</sub> tag + CPR357. 3) MPA + NADH + STCB (fungal *A. nidulans* p450) + lichen CPR357. 4) MPA + NADH + SCYP (fungal *T. marneffi* p450) + lichen CPR357. Whole-cell bioconversion (*in vivo*) experiments were performed twice and the experiments using cell free extracts were performed only once (*in vitro*). Each time a different strategy was used.

### 5.19.1. Strategy I of whole-cell bioconversion experiments (*in vivo*)

In the first strategy of whole-cell bioconversion experiments, 15 cultures were set up, including 11 controls, as shown in **Figure 5.24**. BL21 (DE3)pLySs cells were inoculated with 50 mL of TB broth containing nutrients required to grow a bacterium for a longer period, and proteins were expressed for 48 h at 22 °C using 1 mM IPTG. The substrate MPA and the electron donor NADH were dissolved in DMSO and added directly to 10 out of 15 growing bacterial cultures when proteins were expressed for 24 h. The substrate incubation period was 24 h, and then all the samples were acidified to quench the reaction using a few drops of conc. HCl. Samples were extracted with ethyl acetate, dried over sodium sulfate, and the solvent was evaporated using rotovap. Samples were dissolved in MeOH and filtered using a 0.45-micron filter to run on HPLC, and samples were prepared in acetonitrile for LC-MS analysis (**Figure 5.25**).

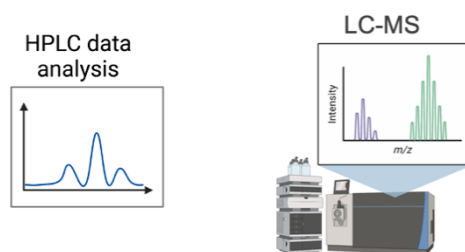


**Figure 5.24:** Co-expression of MPAO (N-His SUMO tag) & CPR357, MPAO (N-His tag) & CPR357, STCB (N-His tag) & CPR357, SCYP (N-His tag) & CPR357, and bioconversion of MPA to UA catalyzed by these proteins. MPA (substrate) & NADH (electron donor) were dissolved in DMSO and were directly added to bacterial cultures co-expressing earlier mentioned protein



combinations inoculated with TB broth. There are 4-types of co-transformants, and three experiments were set up for each co-transformant: Flasks with yellow and pink color cultures: Induced proteins with & w/o (-ve controls) MPA & NADH and flasks with brown color cultures are un-induced proteins with MPA & NADH (-ve controls). Bioconversion reactions were quenched with 1M HCl and extracted with ethyl acetate (EtOAc) because of high solubility of UA in EtOAc.

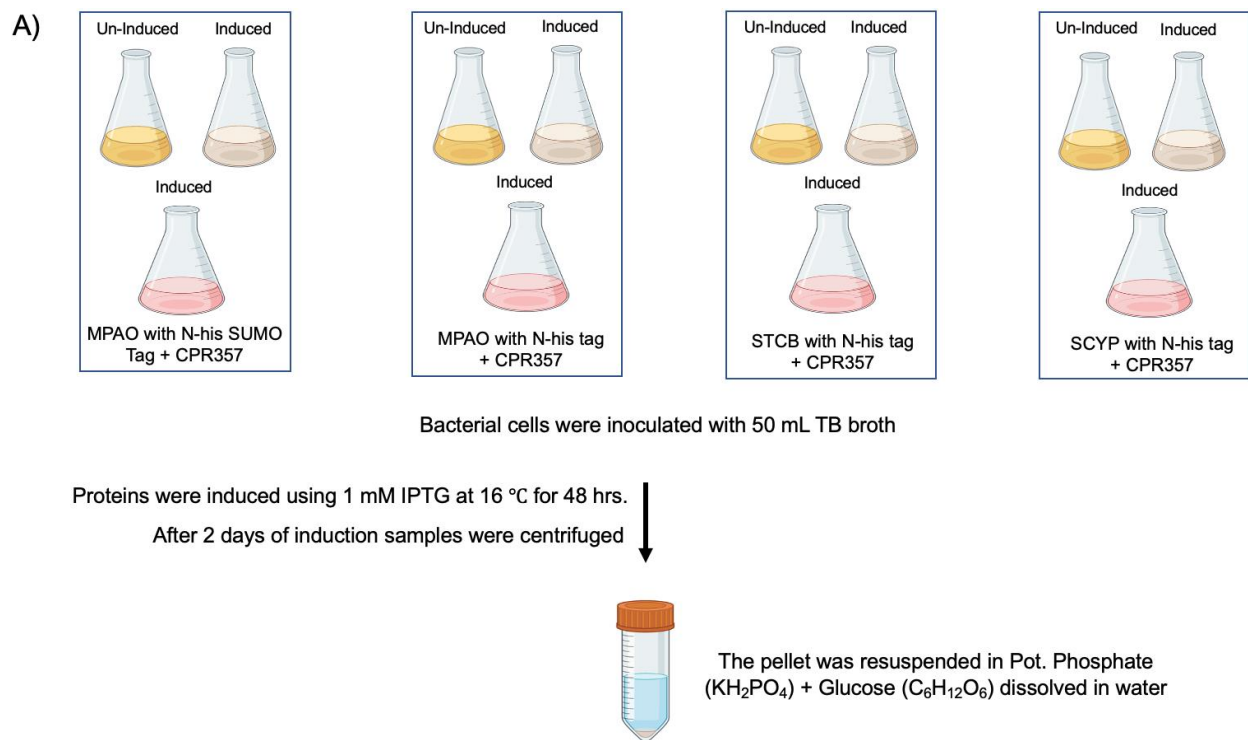
- 1) Samples were dried over  $\text{Na}_2\text{SO}_4$ , and solvent was evaporated using rotovap
- 2) Samples were dissolved in 1 mL of MeOH and filtered using 0.45-micron filter
- 3) Samples were analyzed using HPLC &, LC-MS



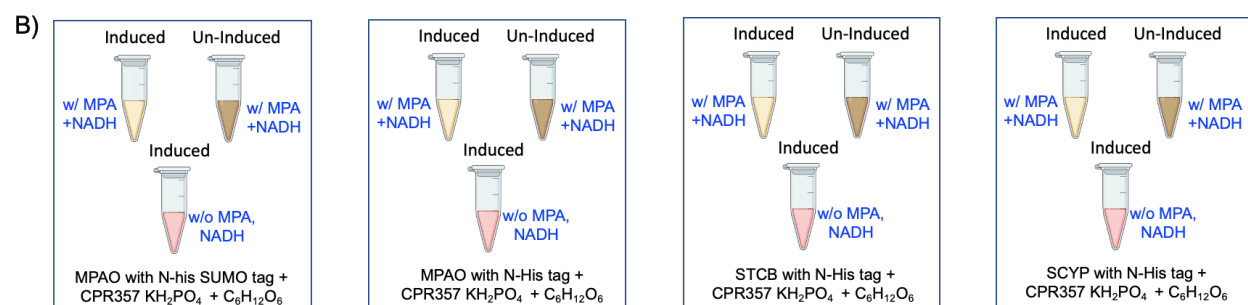
**Figure 5.25:** Sample preparation of all the bioconversion experiments by evaporating EtOAc, drying over sodium sulfate ( $\text{Na}_2\text{SO}_4$ ), dissolved in HPLC-grade methanol (MeOH) and filtered for UV-HPLC analysis. Samples were further diluted with acetonitrile for LC-MS analysis.

### 5.19.2. Strategy II of whole-cell bioconversion experiments (*in vivo*)

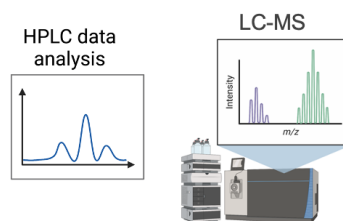
In strategy II, 15 whole-cell bioconversion experiments, including 11 controls, were set, as shown in **Figure 5.26**. BL21 (DE3)pLySs cells were inoculated with 50 mL of TB broth, and proteins were expressed for 48 h at 16 °C using 1 mM IPTG. Then cells were harvested, after centrifugation, the supernatant was discarded, and the cell pellet was resuspended in potassium phosphate and glucose buffer. MPA + NADH were dissolved in DMSO and incubated with 10 out of 15 samples at r. t. for 3 h All the samples were acidified to quench the reaction with two drops of conc. HCl and extracted with ethyl acetate. Samples were dissolved in HPLC-grade MeOH and filtered to run on HPLC. Also, samples were prepared in acetonitrile for LC-MS analysis (**Figure 5.27**).



**Figure 5.26:** Bacterial cultures co-expressing MPAO (N-His SUMO tag) & CPR357, MPAO (N-His tag) & CPR357, STCB (N-His tag) & CPR357, SCYP (N-His tag) & CPR357 inoculated with TB broth. There are 4-types of co-transformants, and three experiments were set up for each co-transformant: Flasks with yellow and pink color cultures: Induced proteins, and brown color cultures are un-induced proteins. Proteins were induced for 2 days with 1 mM IPTG at 16 °C. Then cells were harvested, centrifuged, and pellet was resuspended in potassium phosphate ( $\text{KH}_2\text{PO}_4$ ) + glucose buffer.



- 1) MPA (substrate) + NADH (electron donor) were dissolved in DMSO and were added to  $\text{KH}_2\text{PO}_4 + \text{C}_6\text{H}_{12}\text{O}_6$  dissolved cell pellets.
- 2) Incubated for 3 h at r. t.
- 3) Reactions were quenched by adding a drop of conc. HCl
- 4) Then samples were dissolved in MeOH as UA is sparingly soluble in MeOH
- 5) Samples were filtered using a 0.45-micron filter and screened for HPLC and LC-MS

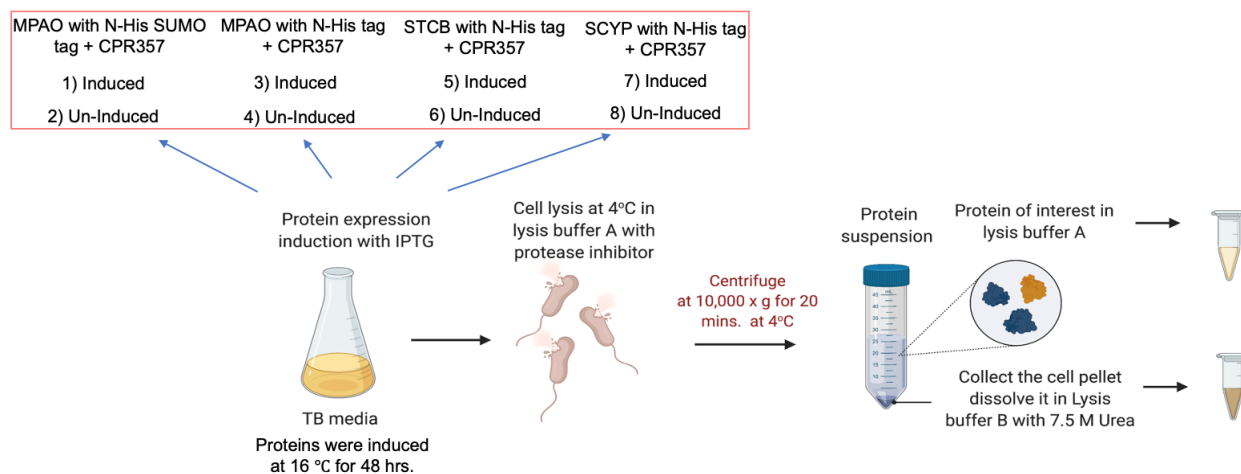


**Figure 5.27:** Bioconversion of MPA to UA catalyzed by these proteins. MPA (substrate) & NADH (electron donor) were dissolved in DMSO and were incubated for 3 h at r. t. with induced and un-induced samples (yellow & brown color coded). Bioconversion

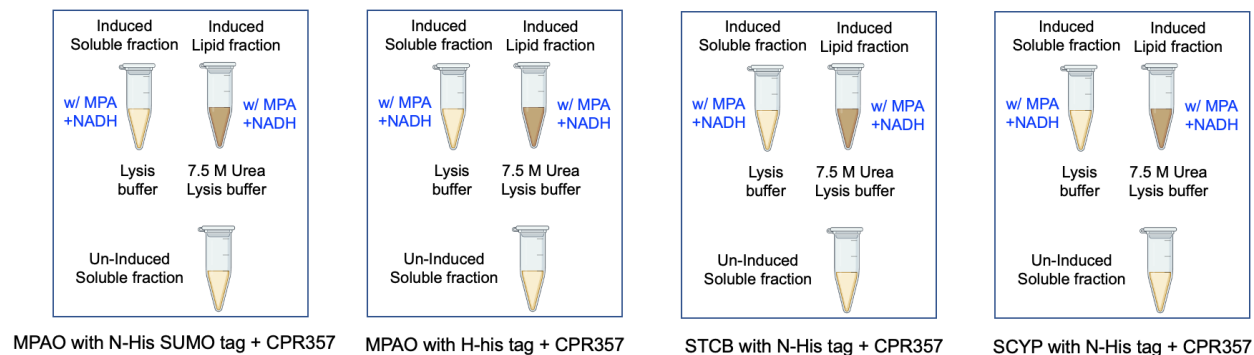
reactions were quenched with a drop of HCl, diluted in 500  $\mu$ L HPLC-grade methanol (MeOH), and filtered for UV-HPLC analysis. Samples were also prepared in acetonitrile for LC-MS analysis.

### 5.19.3. Strategy III of whole-cell bioconversion experiments (*in vivo*)

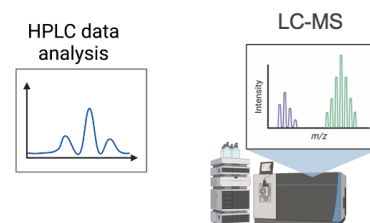
In strategy III, eight co-expression experiments were set up, of which four samples were induced with IPTG & four were not induced (**Section 5.18**). BL21 (DE3)pLySs cells were inoculated with 50 mL of TB broth, and proteins were expressed for 48 h at 16 °C using 1 mM IPTG. Then cells were harvested and lysed on ice with sonication (**Figure 5.28**). Both soluble protein fractions in lysis buffer A and lipid fractions in lysis buffer B were incubated with DMSO dissolved MPA + NADH. Also, the soluble fractions of un-induced samples were incubated with MPA & NADH. The incubation was done at 4 °C for 2 h and after that, all the reactions were quenched with two drops of conc. HCl. Samples were extracted using ethyl acetate, dissolved in HPLC-grade MeOH, and filtered to run on HPLC. Samples were also prepared in acetonitrile for LC-MS analysis (**Figure 5.29**).



**Figure 5.28:** Bacterial cultures co-expressing MPAO (N-His SUMO tag) & CPR357, MPAO (N-His tag) & CPR357, STCB (N-His tag) & CPR357, SCYP (N-His tag) & CPR357 inoculated with TB broth. Proteins were induced for 2 days with 1 mM IPTG at 16 °C. Then cells were harvested, centrifuged, lysed using lysis buffer A and the pellet was resuspended in lysis buffer B (7.5 M Urea).



- 1) MPA (substrate) + NADH (electron donor) were dissolved in DMSO and were added to **soluble fractions in lysis buffer** and **lipid fractions in 7.5M urea buffer**.
- 2) Incubated for 1.5 h at 4 °C.
- 3) Reactions were quenched by adding a drop of conc. HCl
- 4) Then samples were dissolved in MeOH as UA is sparingly soluble in MeOH
- 5) Samples were filtered using a 0.45-micron filter and screened for HPLC and LC-MS

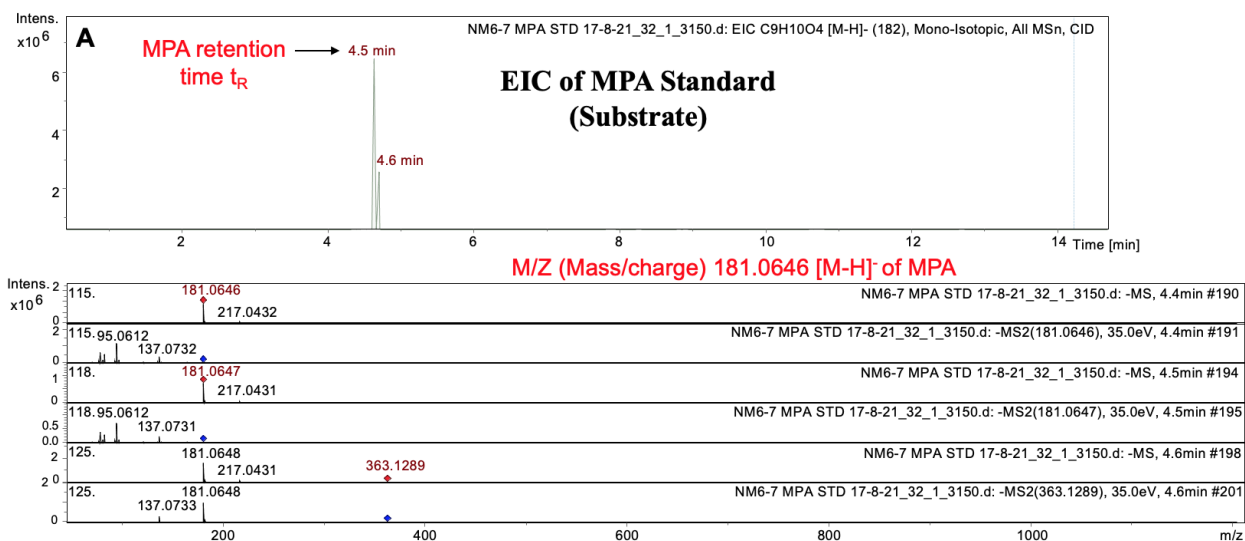


**Figure 5.29:** Bioconversion of MPA to UA. There are 4-types of co-expressed protein samples [MPAO (N-His SUMO tag) & CPR357, MPAO (N-His tag) & CPR357, STCB (N-His tag) & CPR357, SCYP (N-His tag) & CPR357]. Two experiments were set up for each co-expressed protein sample. Eppendorf tubes with yellow samples are induced soluble protein fractions, and tubes with brown samples are induced lipid protein fractions. MPA (substrate) & NADH (electron donor) were dissolved in DMSO and were incubated for 1.5 h at 4 °C with both soluble and lipid fractions in lysis buffer A & B. Bioconversion reactions were quenched with a drop of HCl, diluted in 500 mL HPLC-grade methanol (MeOH), and filtered for UV-HPLC analysis. And samples were prepared in acetonitrile for LC-MS analysis.

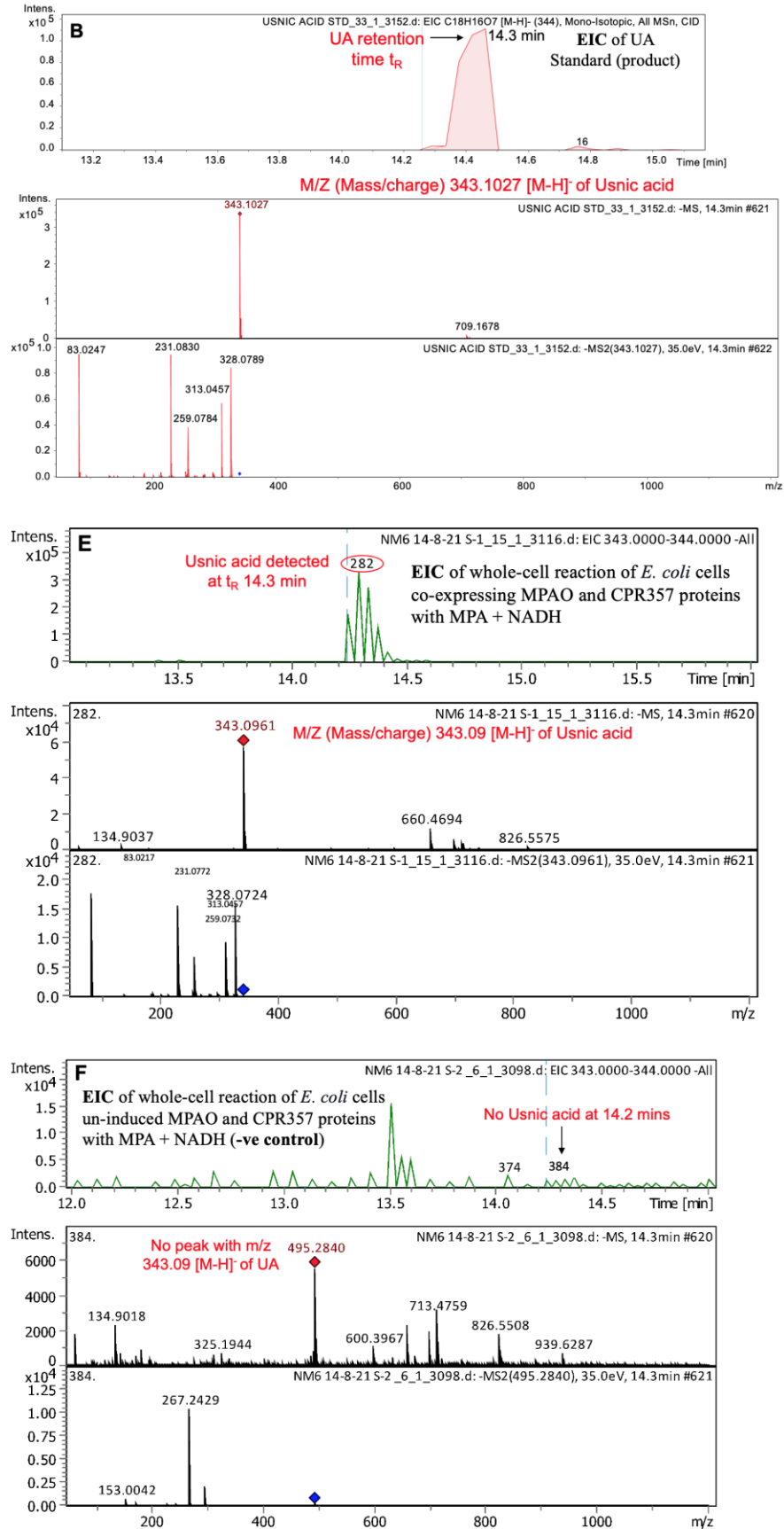
## 5.20. LC-MS analysis

Nothing was detected on HPLC using our standard lab-created UA detection method. In **Figure 5.30**, sample **A** is an extracted ion chromatogram (EIC) and mass spectrum (MS<sup>2</sup>) of standard MPA (substrate) with a peak at retention time ( $t_R$ ) 4.5 min & a base ion peak at  $m/z=181.06$  (-ve ion mode). Sample **B** of **Figure 5.30** is an extracted ion chromatogram (EIC) and mass spectrum (MS<sup>2</sup>) of standard UA (product) with a peak at retention time ( $t_R$ ) 14.2 min & a base ion peak at  $m/z=343.10$  (-ve ion mode). UA (product) was detected in LC-MS analysis of a sample **C** (**Figure 5.30**) of *in vitro* reaction of, STCB (*A. nidulans* p450) and CPR357 proteins lysis buffer A with MPA + NADH dissolved in DMSO (*strategy III, in vitro Section 5.19.3*). EIC of sample **C** has a peak at 14.3 min retention time ( $t_R$ ) and MS<sup>2</sup> has a parent ion peak at  $m/z=343$  in the -ve

ion mode with fragmentation ions 83.02, 231.07, 259.07, 313.04 & 348.07 in MS2 similar to standard UA (**Figure 5.30**). Sample **D** is a negative control experiment of sample **C**, where STCB & CPR357 proteins were not induced (No IPTG) (**Figure 5.30**). The peak of the substrate, MPA, at  $t_R$  4.5 min with a base ion peak at  $m/z$ = 181.06 with fragmentation ions 95.06, 137.07 was observed in the Sample **D** (-ve control) (**Figure 5.30**) after comparing it to the standard MPA. UA was also detected in the sample **E** (**Figure 5.30**) of whole-cell reaction of *E. coli* cells in potassium phosphate and glucose buffer, co-expressing MPAO (with His<sub>6</sub> SUMO tag) and CPR357 proteins with MPA + NADH dissolved in DMSO (*strategy II in vivo* **Section 5.19.2**). The peak at 14.3 min  $t_R$  in EIC with a mass of about 343.14 in MS2 was consistent with the standard UA spectra (Sample **E**, **Figure 5.30**). This peak was missing in the negative control (sample **F**) where MPAO (with His<sub>6</sub> SUMO tag) and CPR357 proteins were not induced (**Figure 5.30**).







**Figure 5.30:** LC-MS results of MPA to UA bioconversion experiments, liquid chromatogram (on the top of each chromatogram) with a retention time on the x-axis and intensity on the y-axis. Mass spectra (on the bottom of each chromatogram) with a base peak and (MS/MS) & the fragmentation pattern/ product ion (MS2). All spectra were acquired in the negative ion mode. A. MPA in acetonitrile; B. UA in acetonitrile; C. MPA + NADH + *E. coli* transformants co-expressed STCB & CPR357 (Sample C) (*in vitro*, strategy 3 bioconversion experiments); D. MPA + NADH + No induction of STCB & CPR357 proteins (-ve control, Sample D); E. MPA + NADH + *E. coli* transformants co-expressed MPAO (N-His SUMO tag) & CPR357 (Sample E) (*in vivo*, strategy 2 bioconversion experiments); F. MPA + NADH + No induction of MPAO (N-His SUMO tag) & CPR357 proteins (-ve control, Sample F).

## 5.21. Summary

This study has experimentally demonstrated the first successful heterologous expression of a lichen fungal cytochrome p450 (MPAO), lichen fungal cytochrome p450 reductase (CPR) in bacteria. First successful purification of lichen fungal CYP and CPR proteins. First successful functional characterization of lichen fungal CYP and CPR proteins in bacteria. First successful UA (lichen SM) biosynthesis catalyzed by MPAO & CPR357 proteins both *in vivo* & *in vitro*. This study establishes a successful protocol for soluble expression of a lichen CYP & its redox partner in bacteria and, also a protocol for the biosynthesis of a lichen SM (UA) both *in vivo* & *in vitro* reproducibly.

Here, is the summary of both successful & failed experiments (all proteins were expressed within a bacterial host; *E. coli*):

1. Full length-native MPAO-pETite N-His SUMO expression plasmid- No protein expression was observed.
2. Truncated MPAO (w/o TMD) codon-optimized for *E. coli*-pETite N-His SUMO expression plasmid- Soluble protein expression was observed.
3. Truncated MPAO (w/o TMD) codon-optimized for *E. coli*-pET28b N-His<sub>6</sub> expression plasmid- No protein expression was observed.
4. Truncated CPR357 (w/o TMD) non-codon-optimized for *E. coli*-pET21a N-His<sub>6</sub> expression plasmid- Low soluble protein expression was observed.



5. Fungal STCB (*A. nidulans* p450), w/o TMD non-codon-optimized for *E. coli*-pET28b N-His<sub>6</sub> expression plasmid- No protein expression was observed.
6. Fungal SCYP (*T. marneffi* p450), with TMD non-codon-optimized for *E. coli*-pET28b N-His<sub>6</sub> expression plasmid- No protein expression was observed.
7. Co-expression of the chaperone with fungal STCB (*A. nidulans* p450), w/o TMD non-codon-optimized for *E. coli*-pET28b N-His<sub>6</sub> expression plasmid- No protein expression was observed.
8. Co-expression of the chaperone with fungal SCYP (*T. marneffi* p450), with TMD non-codon-optimized for *E. coli*-pET28b N-His<sub>6</sub> expression plasmid- No protein expression was observed.
9. Co-expression of the chaperone with Truncated MPAO (w/o TMD) codon-optimized for *E. coli*-pET28b N-His<sub>6</sub> expression plasmid- No protein expression was observed.
10. Co-expression of fungal STCB (*A. nidulans* p450), w/o TMD non-codon-optimized for *E. coli*-pET28b N-His<sub>6</sub> expression plasmid + Truncated CPR357 (w/o TMD) non-codon-optimized for *E. coli*-pET21a N-His<sub>6</sub> expression plasmid- functional expression of proteins was observed.
11. Co-expression of fungal SCYP (*T. marneffi* p450), w/o TMD non-codon-optimized for *E. coli*-pET28b N-His<sub>6</sub> expression plasmid + Truncated CPR357 (w/o TMD) non-codon-optimized for *E. coli*-pET21a N-His<sub>6</sub> expression plasmid- No expression of SCYP was observed
12. Co-expression of truncated MPAO (w/o TMD) codon-optimized for *E. coli*-pETite N-His SUMO expression plasmid + Truncated CPR357 (w/o TMD) non-codon-optimized for *E. coli*-pET21a N-His<sub>6</sub> expression plasmid- functional expression of proteins was observed.
13. Co-expression of truncated MPAO (w/o TMD) codon-optimized for *E. coli*-pET28b N-His<sub>6</sub> expression plasmid + Truncated CPR357 (w/o TMD) non-codon-optimized for *E. coli*-pET21a N-His<sub>6</sub> expression plasmid- No expression of MPAO was observed.

## References

- Abdel-Hameed, M., Bertrand, R. L., Piercey-Normore, M. D., Sorensen, J. L. Putative identification of the usnic acid biosynthetic gene cluster by de novo whole-genome sequencing of a lichen-forming fungus. *Fungal Biol.* **2016**, 120 (3), 306-16.
- Austin, S. & Nordström, K. Partition-mediated incompatibility of bacterial plasmids. *Cell.* **1990** 60 (3), 351-4.
- Benarroch, J. M. & Asally, M. The Microbiologist's Guide to Membrane Potential Dynamics, *Trends in Microbiology*, **2020**, Vol. 28, Issue 4, Pages 304-314.
- Blaschkowski, H. P., Neuer, G., Ludwig-Festl, M., & Knappe, J. Routes of flavodoxin and ferredoxin reduction in Escherichia coli. CoA-acylating pyruvate: flavodoxin and NADPH: flavodoxin oxidoreductases participating in the activation of pyruvate formate-lyase. *European journal of biochemistry*, **1982**, 123 (3), 563–569.
- Braun, A., Geier, M., Bühler, B. *et al.* Steroid biotransformations in biphasic systems with *Yarrowia lipolytica* expressing human liver cytochrome P450 genes. *Microb. Cell Fact.* **2012**, 11, 106.
- Braun, A., Geier, M., Bühler, B., *et al.* Steroid biotransformations in biphasic systems with *Yarrowia lipolytica* expressing human liver cytochrome P450 genes. *Microb. Cell Fact.* **2012**, 11 (1), 106.
- Budriang, C., Rongnoparut, P., & Yuvaniyama, J. An expression of an insect membrane-bound cytochrome P450 CYP6AA3 in the Escherichia coli in relation to insecticide resistance in a malarial vector. *Pakistan journal of biological sciences: PJBS*, **2011**, 14 (8), 466–475.
- Butt, T. R., Edavettal, S. C., Hall, J. P., & Mattern, M. R. SUMO fusion technology for difficult-to-express proteins. *Protein expression and purification*, **2005**, 43 (1), 1–9.

- Campelo, D., Esteves, F., Brito Palma, B., Costa Gomes, B., Rueff, J., Lautier, T., Urban, P., Truan, G., & Kranendonk, M. Probing the Role of the Hinge Segment of Cytochrome P450 Oxidoreductase in the Interaction with Cytochrome P450. *International journal of molecular sciences*, **2018**, 19 (12), 3914. <https://doi.org/10.3390/ijms19123914>
- Chapot-Chartier, M. P. and Buddelmeijer, N. Post-translational modifications in bacteria – *The dynamics of bacterial physiology*, *Research in Microbiology*, **2021**, Volume 172, Issues 7–8, 103887.
- Chen, J. & Tsai, Y. H. Applications of Genetic Code Expansion in Studying Protein Post-Translational Modification, *Journal of Molecular Biology*, **2021**, Vol. 434, Issue 8, 167424.
- Cheng D., Kelley R. W., Cawley G. F. and Backes W. L. High-level expression of recombinant rabbit cytochrome P450 2E1 in Escherichia coli C41 and its purification. *Protein Expr. Purif.* **2004**, 33, 66–71.
- Claassens, N. J., Siliakus, M. F., Spaans, S. K., Creutzburg, S., Nijse, B., Schaap, P. J., Quax, T., & van der Oost, J. Improving heterologous membrane protein production in Escherichia coli by combining transcriptional tuning and codon usage algorithms. *PloS one*, **2017**, 12 (9), e0184355.
- Costa, S., Almeida, A., Castro, A., & Domingues, L. Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: the novel Fh8 system. *Frontiers in microbiology*, **2014**, 5, 63.
- Crešnar, B., & Petrič, S. Cytochrome P450 enzymes in the fungal kingdom. *Biochimica et biophysica acta*, **2011**, 1814 (1), 29–35.

- Demain, A. L., & Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnology advances*, **2009**, 27 (3), 297–306.
- Denisov, I. G., Shih, A. Y., & Sligar, S. G. Structural differences between soluble and membrane bound cytochrome P450s. *Journal of inorganic biochemistry*, **2012**, 108, 150–158.
- Domanski, T. L.; Finta, C.; Halpert, J. R.; Zaphiropoulos, P. G. CDNA cloning and initial characterization of CYP3A43, a novel human cytochrome P450, *Mol. Pharmacol.* **2001**, 59.
- Durairaj, P., Fan, L., Machalz, D., Wolber, G., & Bureik, M. Functional characterization and mechanistic modeling of the human cytochrome P450 enzyme CYP4A22. *FEBS letters*, **2019**, 593 (16), 2214–2225.
- Durairaj, P., Hur, JS. & Yun, H. Versatile biocatalysis of fungal cytochrome P450 monooxygenases. *Microb. Cell Fact.* **2016**, 15, 125.
- Durairaj, P., Jung, E., Park, H. H., et al. Comparative functional characterization of a novel benzoate hydroxylase cytochrome P450 of *Fusarium oxysporum*. *Enzyme Microb. Technol.* **2015**, 70, 58–65.
- Durairaj, P., Malla, S., Nadarajan, S. P., et al. Fungal cytochrome P450 monooxygenases of *Fusarium oxysporum* for the synthesis of  $\omega$ -hydroxy fatty acids in engineered *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **2015**, 14 (1), 1–16.
- Esteves, F., Campelo, D., Gomes, B. C., et al. The Role of the FMN-domain of human cytochrome P450 oxidoreductase in its promiscuous interactions with structurally diverse redox partners. *Front. Pharmacol.* **2020**, 11, 299.
- Gao, X., Chen, W., Guo, C. et al. Soluble cytoplasmic expression, rapid purification, and characterization of cyanovirin-N as a His-SUMO fusion. *Applied Microbiology and*

- Biotechnology. 2010 Jan;85(4):1051-1060. DOI: 10.1007/s00253-009-2078-5. PMID: 19547966; PMCID: PMC7080120.
- Gideon, D. A., Kumari, R., Lynn, A. M., et al. What is the functional role of N-terminal transmembrane helices in the metabolism mediated by liver microsomal cytochrome P450 and its reductase? *Cell Biochem. Biophys.* **2012**, 63 (1), 35–45.
- Hannemann, F., Bichet, A., Ewen, K. M., & Bernhardt, R. Cytochrome P450 systems--biological variations of electron transport chains. *Biochimica et biophysica acta*, **2007**, 1770 (3), 330–344.
- Hatakeyama, M., Kitaoka, T., Ichinose, H. Heterologous expression of fungal cytochromes P450 (CYP5136A1 and CYP5136A3) from the white-rot basidiomycete *Phanerochaete chrysosporium*: Functionalization with cytochrome b5 in *Escherichia coli*, *Enzyme and Microbial Technology*, **2016**, Volume 89, Pages 7-14.
- Hausjell, J., Halbwirth, H., & Spadiut, O. Recombinant production of eukaryotic cytochrome P450s in microbial cell factories. *Bioscience reports*, **2018**, 38 (2)
- Hausjell, J., Halbwirth, H., & Spadiut, O. Recombinant production of eukaryotic cytochrome P450s in microbial cell factories. *Bioscience reports*, **2018**, 38 (2), BSR20171290.
- Henry, K. M., and Townsend, C. A. Ordering the Reductive and Cytochrome P450 Oxidative Steps in Demethylsterigmatocystin Formation Yields General Insights into the Biosynthesis of Aflatoxin and Related Fungal Metabolites. *Journal of the American Chemical Society*, **2005**, 127 (11), 3724-3733.
- Hirosue, S., Tazaki, M., Hiratsuka, N., Yanai, S., Kabumoto, H., Shinkyō, R., Arisawa, A., Sakaki, T., Tsunekawa, H., Johdo, O., Ichinose, H., Wariishi, H. Insight into functional diversity of cytochrome P450 in the white-rot basidiomycete *Phanerochaete chrysosporium*:

- Involvement of versatile monooxygenase. *Biochemical and Biophysical Research Communications*, **2011**, Volume 407, Issue 1, Pages 118-123.
- Ichinose, H. & Wariishi, H. Heterologous expression and mechanistic investigation of a fungal cytochrome P450 (CYP5150A2): Involvement of alternative redox partners, *Archives of Biochemistry and Biophysics*, **2012**, Volume 518, Issue 1, Pages 8-15.
- Ichinose, H., Hatakeyama, M., Yamauchi, Y. Sequence modifications and heterologous expression of eukaryotic cytochromes P450 in Escherichia coli. *J. Biosci. Bioeng.* **2015**, 120 (3), 268–274.
- Ichinose, H., Wariishi, H. High-level heterologous expression of fungal cytochrome P450s in Escherichia coli. *Biochem. Biophys. Res. Commun.* **2013**, 438 (2), 289–294.
- Ide, M., Ichinose, H., & Wariishi, H. Molecular identification and functional characterization of cytochrome P450 monooxygenases from the brown-rot basidiomycete *Postia placenta*. *Archives of microbiology*, **2012**, 194 (4), 243–253.
- Jennewein, S., Park, H., DeJong, J. M., Long, R. M., Bollon, A. P., & Croteau, R. B. Coexpression in yeast of *Taxus* cytochrome P450 reductase with cytochrome P450 oxygenases involved in Taxol biosynthesis. *Biotechnology and bioengineering*, **2005**, 89 (5), 588–598.
- Katarzyna, K. and Koraljka, H. Resolving the Complexity of Ubiquitin Networks. *Frontiers in Molecular Biosciences*, **2020**, Vol. 7.
- Keller, N. P., Watanabe, C., Kelkar, H. S., Adams, T. H. and Townsend, C. A. Requirement of Monooxygenase-Mediated Steps for Sterigmatocystin Biosynthesis by *Aspergillus nidulans*. *Applied and Environmental Microbiology*, **2000**, 359–362.

- Khorasanizadeh, S., Peters, I. D., & Roder, H. Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nature structural biology*, **1996**, 3 (2), 193–205.
- Lah, L., Kraševc, N., Trontelj, P., et al. High diversity and complex evolution of fungal cytochrome P450 reductase: cytochrome P450 systems. *Fungal Genet. Biol.* **2008**, 45 (4), 446–458.
- Lah, L., Podobnik, B., Novak, M., et al. The versatility of the fungal cytochrome P450 monooxygenase system is instrumental in xenobiotic detoxification. *Mol. Microbiol.* **2011**, 81 (5), 1374–1389.
- Lee, C. D., Sun, H. C., Hu, S. M., Chiu, C. F., Homhuan, A., Liang, S. M., et al. An improved SUMO fusion protein system for effective production of native roteins. *Protein Sci.* **2008**, 17, 1241–1248.
- Li, S., Du, L., Bernhardt, R. Redox partners: function modulators of bacterial P450 enzymes. *Trends Microbiol.* **2020**, 28 (6), 445–454.
- Lv, Y., Marsafari, M., Koffas, M., et al. Optimizing oleaginous yeast cell factories for flavonoids and hydroxylated flavonoids biosynthesis. *ACS Synth. Biol.* **2019**, 8 (11), 2514–2523.
- Macek, B., Forchhammer, K., Hardouin, J., Weber-Ban, E., Grangeasse, C., & Mijakovic, I. Protein post-translational modifications in bacteria. *Nature reviews. Microbiology*, **2019**, 17 (11), 651–664.
- Malakhov, M. P., Mattern, M. R., Malakhova, O. A., Drinker, M., Weeks, S. D., & Butt, T. R. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *Journal of structural and functional genomics*, **2004**, 5 (1-2), 75–86.

- Marblestone, J. G., Edavettal, S. C., Lim, Y., Lim, P., Zuo, X., & Butt, T. R. Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. *Protein science: a publication of the Protein Society*, **2006**, 15 (1), 182–189.
- Maroutsos, D., Huff, H., & Das, A. Bacterial Expression of Membrane-Associated Cytochrome P450s and Their Activity Assay in Nanodiscs. *Methods in molecular biology (Clifton, N.J.)*, **2019**, 1927, 47–72.
- Monk, B. C., Tomasiak, T. M., Keniya, M. V., et al. Architecture of a single membrane spanning cytochrome P450 suggests constraints that orient the catalytic domain relative to a bilayer. *Proc. Natl. Acad. Sci.* **2014**, 111 (10), 3865–3870.
- Mossesso, E. and Lima, C.D. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell*. **2000**, 5, 865.
- Nazir, K. H., Ichinose, H., & Wariishi, H. Construction and application of a functional library of cytochrome P450 monooxygenases from the filamentous fungus *Aspergillus oryzae*. *Applied and environmental microbiology*, **2011**, 77 (9), 3147–3150.
- Neunzig, I., Widjaja, M., Peters, F. T., et al. Coexpression of CPR from various origins enhances biotransformation activity of human CYPs in *S. pombe*. *Appl. Biochem. Biotechnol.* **2013**, 170 (7), 1751–1766.
- Nordström, K. and Austin, S. J. Mechanisms that contribute to the stable segregation of plasmids. *Annu Rev Genet.* **1989**; 23: 37-69.
- Noth, J., Krawietz, D., Hemschemeier, A. and Happe, T. Pyruvate: Ferredoxin Oxidoreductase Is Coupled to Light-independent Hydrogen Production in *Chlamydomonas reinhardtii*. *The Journal of Biological Chemistry*, **2013**, Vol. 288, pp. 4368–4377.



- Novak, M., Lah, L., Šala, M., et al. Oleic acid metabolism via a conserved cytochrome p450 system-mediated  $\omega$ -hydroxylation in the bark beetle-associated fungus *Grosmannia clavigera*. *PLoS One* **2015**, 10 (3), e0120119.
- Novick, R. P. Plasmid incompatibility. *Microbiol Rev.* **1987**, 51(4), 381-95.
- Pan, Y., Abd-Rashid, B. A., Ismail, Z., Ismail, R., Mak, J. W., & Ong, C. E. Heterologous expression of human cytochromes P450 2D6 and CYP3A4 in *Escherichia coli* and their functional characterization. *The protein journal*, **2011**, 30 (8), 581–591.
- Ribet, D., & Cossart, P. Post-translational modifications in host cells during bacterial infection. *FEBS letters*, **2010**, 584 (13), 2748–2758.
- Rosano, G. L. & Ceccarelli, E. A. Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted *Escherichia coli* strain. *Microb. Cell Fact.* **2009**, 8, 41.
- Sagadin, T., Riehm, J. L., Milhim, M., et al. Binding modes of CYP106A2 redox partners determine differences in progesterone hydroxylation product patterns. *Commun. Biol.* **2018**, 1 (1), 1–9.
- Sagadin, T., Riehm, J., Putkaradze, N., et al. Novel approach to improve progesterone hydroxylation selectivity by CYP 106A2 via rational design of adrenodoxin binding. *FEBS J.* **2019**, 286 (6), 1240–1249.
- Sagar Ghosh, Ahsan Choudary, Sangeeta Ghosh, Nicolas Musi, Yanfen Hu, Rong Li, IKK $\beta$  Mediates Cell Shape-Induced Aromatase Expression and Estrogen Biosynthesis in Adipose Stromal Cells, *Molecular Endocrinology*, **2009**, Volume 23, Issue 5, Pages 662–670
- Satakarni, M., & Curtis, R. Production of recombinant peptides as fusions with SUMO. *Protein expression and purification*, **2011**, 78 (2), 113–119.

- Schlapschy, M., & Skerra, A. Periplasmic chaperones used to enhance functional secretion of proteins in *E. coli*. *Methods in molecular biology (Clifton, N.J.)*, **2011**, 705, 211–224.
- Shoun, H., Takaya, N. Cytochromes P450nor and P450foxy of the fungus *Fusarium oxysporum*. International Congress Series editors. Elsevier, **2002**.
- Singh, A., Upadhyay, V., Upadhyay, A.K. *et al.* Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microb Cell Fact* **14**, 41 (2015).
- Sono, M., Roach, M. P., Coulter, E. D., & Dawson, J. H. Heme-Containing Oxygenases. *Chemical reviews*, **1996**, 96 (7), 2841–2888.
- Sørensen, H. P., & Mortensen, K. K. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial cell factories*, **2005**, 4 (1), 1.
- Šrejber, M., Navrátilová, V., Paloncýová, M., Bazgier, V., Berka, K., Anzenbacher, P., Otyepka, M. Membrane-attached mammalian cytochromes P450: An overview of the membrane's effects on structure, drug binding, and interactions with redox partners, *Journal of Inorganic Biochemistry*, **2018**, 183, 117.
- Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, *J. Biol. Chem.* **1986**, 189, 113–130.
- Terpe K. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied microbiology and biotechnology*, **2006**, 72 (2), 211–222.
- Urban, P.; Lautier, T.; Pompon, D., et al., Ligand access channels in cytochrome P450 enzymes: *A review. Int. J. Mol. Sci.* **2018**, 19 (6), 1617.

- Velappan, N., Sblattero, D., Chasteen, L., Pavlik, P., Bradbury, A. R. M. Plasmid incompatibility: more compatible than previously thought? *Protein Engineering, Design and Selection*, **2007**, Volume 20, Issue 7, Pages 309–313.
- Wang, H., Xiao, Y., Fu, L. *et al.* High-level expression and purification of soluble recombinant FGF21 protein by SUMO fusion in *Escherichia coli*. *BMC Biotechnol.* **2010**, 10, 14.
- Wang, Z., Li, N., Wang, Y., Wu, Y., Mu, T., Zheng, Y., Huang, L., & Fang, X. Ubiquitin-intein and SUMO2-intein fusion systems for enhanced protein production and purification. *Protein expression and purification*, **2012**, 82 (1), 174–178.
- Wen, Y., Hatabayashi, H., Arai, H., Kitamoto, H. K., and Yabe, K. Function of the *cypX* and *moxY* Genes in Aflatoxin Biosynthesis in *Aspergillus parasiticus*. *Applied and Environmental Microbiology*, **2005**, Volume 71, Issue 6, Pages 3192-3198.
- Wu, Z. L., Qiao, J., Zhang, Z. G., et al. Enhanced bacterial expression of several mammalian cytochrome P450s by codon optimization and chaperone coexpression. *Biotechnol. Lett.* **2009**, 31 (10), 1589–1593.
- Wu, Z. L., Sohl, C. D., Shimada, T., et al. Recombinant enzymes overexpressed in bacteria show broad catalytic specificity of human cytochrome P450 2W1 and limited activity of human cytochrome P450 2S1. *Mol. Pharmacol.* **2006**, 69 (6), 2007–2014.
- Yabe, K., Chihaya, N., Hamamatsu, S., Sakuno, E., Hamasaki, T., Nakajima, H., and Bennett, J. W. Enzymatic Conversion of Averufin to Hydroxyversicolorone and Elucidation of a Novel Metabolic Grid Involved in Aflatoxin Biosynthesis. *Applied and Environmental Microbiology*, **2003**, Volume 69, Issue 1, Pages 66-73.
- Yu, J., Chang, P. K., Ehrlich, K. C., Cary, J. W., Bhatnagar, D., Cleveland, T. E., Payne, G. A., Linz, J. E., Woloshuk, C. P. and Bennett, J. W. Clustered Pathway Genes in Aflatoxin

- Biosynthesis. *Applied and Environmental Microbiology*, **2004**, Volume 70, Issue 3, , Pages 1253-1262.
- Yun, C. H., Yim, S. K., Kim, D. H., & Ahn, T. Functional expression of human cytochrome P450 enzymes in Escherichia coli. *Current drug metabolism*, **2006**, 7 (4), 411–429.
- Zelasko, S., Palaria, A., & Das, A. Optimizations to achieve high-level expression of cytochrome P450 proteins using Escherichia coli expression systems. *Protein expression and purification*, **2013**, 92 (1), 77–87.
- Zhang, X., Li, S. Expansion of chemical space for natural products by uncommon P450 reactions. *Nat. Prod. Rep.* **2017**, 34 (9), 1061–1089.

# Chapter 6

## Assignment of Biosynthetic pathways to Lichen Gene Clusters

### 6.1. Introduction

In **Chapters 1** and **3** the emergence of next-generation genome sequencing and its impact on the field of natural product chemistry was described. Along with the development of next-generation genome sequencing technology, there was a constant improvement and maturation of other techniques in genomics such as AntiSMASH, fungiSMASH (SM gene cluster annotation software) that have been developed to facilitate the annotation of assembled genomes. The release of third version of AntiSMASH (V. 3. 0) [Weber et al. 2015], introduced an additional feature called ‘KnownClusterBLAST’ which is an algorithm that compares the gene clusters identified within an imputed sequence to a compilation of well-characterized gene clusters [Medema et al., 2015]. In 2015 Abdel-Hameed used AntiSMASH V. 2.0 [Blin et al. 2013] to identify only the PKS genes within the draft assembly of their *C. uncialis* genome. A total of 48 SM biosynthetic gene clusters were identified in *C. uncialis* genome, out of which *Cu-nr-pks-2* cluster was identified as UA gene cluster using a gene annotation approach [Abdel-Hameed, M. et al., 2016]. The remaining 47 lichen gene clusters was submitted to AntiSMASH (V. 3.0) and using a ‘homology mapping’ approach in combination with AntiSMASH V. 3.0, Dr. Bertrand proposed the biosynthetic pathways of ‘9’ gene clusters out of these 47 clusters [Bertrand, R. L. et al., 2018 (part I & II)]. Whereas some genes were still of unknown function [Bertrand, R. L. et al., 2018 (part I & II)]. Homology mapping included BLAST consensus similarity and alignment to characterized genes deposited in GenBank and, genetic homology supported by phylogenetic analysis approach was used.

Originally our research group was only interested in identifying the UA biosynthetic gene cluster as this chemotype of *C. uncialis* is known to produce only UA. Profiling work in nonlichenized fungi and bacteria demonstrates that more biosynthetic clusters are present in these organisms than metabolites known to be produced by them [Nielsen, J. C. et al., 2017; Schorn, M. A. et al., 2016; Guo, C. & Wang, C. C. C. 2014; Aigle, B. et al., 2014]. And a similar observation was made in *C. uncialis* [Bertrand, R. L. et al., 2018 (part I)].

In continuation with the work published by Dr. Bertrand in 2018, I decided to analyze the rest of 38 biosynthetic clusters in the year 2020, because of the more fungal & lichen genes and gene clusters were discovered, reported and deposited in GenBank. I narrowed down my research to only those gene clusters (16 out of 38) which had cytochrome p450 as one of the accessory genes after having a better understanding of the classification, chemistry, and the functioning of lichen p450s. Bertrand already predicted and reported the function of all the genes found in these gene clusters [Bertrand, R. L. et al., 2018 (part I)]. Similar to Bertrand approach, I also used a ‘homology mapping’ approach to determine probable pathways of these gene clusters. The result of this work, described in this chapter, is the assignment of biosynthetic pathways of ‘7’ gene clusters out of 16 with cytochrome p450 as one of the accessory genes.

The procedure for identifying a plausible biosynthetic function via the ‘homology mapping’ approach is summarized by addressing the following three questions:

1. Are the genes genetically similar to those of a known biosynthetic gene cluster?
2. Do the identified genes form a coherent biosynthetic pathway?
3. Is genetic homology supported by phylogenetic analysis?

Predictions that contain many paired genes with a high degree of evolutionary homology and form a coherent series of chemical steps comprise a more useful prediction of function. Out

of seven putative biosynthetic pathways, there are only two gene clusters in this study that could be associated to already reported metabolites in the literature. The seven pathways proposed are described below, in order of terpene synthase gene clusters followed by non-reducing polyketide synthase (NR-PKS) gene clusters and then reducing polyketide synthase (R-PKS) gene clusters. Note that it is not obligatory for all genes in a pathway to be assigned with a specific biosynthetic function.

The strength of these predictions rest in the genetic similarity of already characterized genes deposited in GenBank, their genetic homology as supported by multiple sequence alignment, phylogenetics, and their ability to form a coherent chemical pathway. Throughout this work, phylogenetic analyses will be conducted to ascertain the degree of evolutionary homology between the *C. uncialis* and identified reference genes. Complete phylogenetic trees, including trees that do not support a homologous relationship, are shown as **Figure S1 to S24** in the Appendix.

### 6.1.1. Phylogenetics

Phylogenetics is the study of evolutionary relationships among biological entities, often species, individuals or genes. In this era of mega-scale genome sequencing, novel phylogenetic methodology has provided us a platform to depict the functionality of some of the known genes and identify the unknown genes present in these lichen gene clusters. We can predict the biological relatedness of two genes based on the same root, branch length, distance from one node to the other, grouping of the genes/proteins. Protein-based trees, taking amino acid sequences under consideration and their multiple sequence alignment (MSA) files for the comparative analysis of the function of two different proteins have been constructed. Bootstrap support values describe the statistical value for the nodes in a phylogenetic tree. Horizontal line distances or also be called scale bars, that represent number of substitutions per site, give you information about sequence

divergence between the tips through the nodes. If there are no horizontal lines between any sequences in a phylogenetic tree that means those sequences have no difference and are known as haplotypes. Descendants from a node share a common ancestor and those descendants are put together in one clade. An outgroup is a more distantly related sequence in an evolution.

## 6.2. Terpene synthase gene clusters

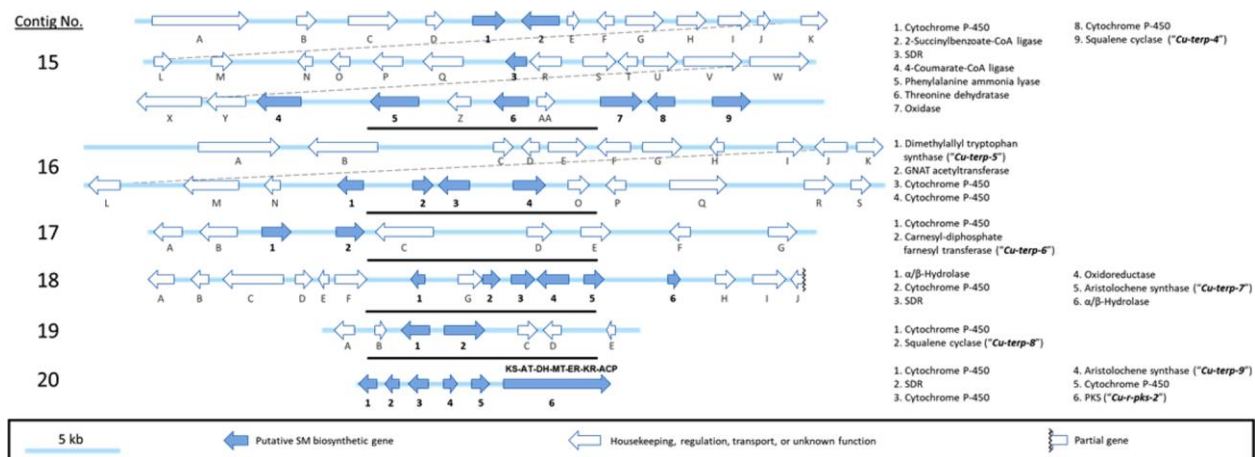
Bertrand in 2018 reported nine terpene synthase gene clusters, named *Cu-terp-1* to *Cu-terp-9* found in the *C. uncilias* mycobiont, where the gene that encodes for terpene synthase is to be considered a backbone enzyme and the other genes, both upstream and downstream to terpene synthase encode for tailoring enzymes. These tailoring enzymes are also known as accessory enzymes. Ultimately, a whole cluster (backbone and tailoring enzymes) was identified to be involved in the biosynthesis of a lichen SM.

Terpenes are the largest and most diverse class of natural compounds, structurally composed of isoprene units with the molecular formula  $(C_5H_8)_n$  [Schmidt-Dannert, C. 2015]. Although most terpenes are typically found in plants. In the last 10-15 years, many genes, and their biosynthetic pathways responsible for the biosynthesis of these SMs/terpenes have also been discovered and characterized in both, fungi [Quin, M. B. et al., 2014] and lichen fungi (mycobiont).

### 6.2.1 Terpene synthase gene cluster, *Cu-terp-4*

Bertrand in 2018 also reported that a gene cluster containing a terpene synthase *Cu-terp-4* is basically a 2,3-oxidosqualene cyclase (OSQCY). These are class II terpene cyclases (**Figure 6.1**) and usually have an alpha 6-barrel fold. 2,3-oxidosqualene cyclase (OSQCY) are integral membrane proteins that catalyze a cationic cyclization cascade converting linear triterpenes to fused ring compounds. Eukaryotic OSQCY transforms the 2,3-epoxide of squalene to compounds such as, lanosterol in fungi and in some mammals or, cycloartenol in plants.

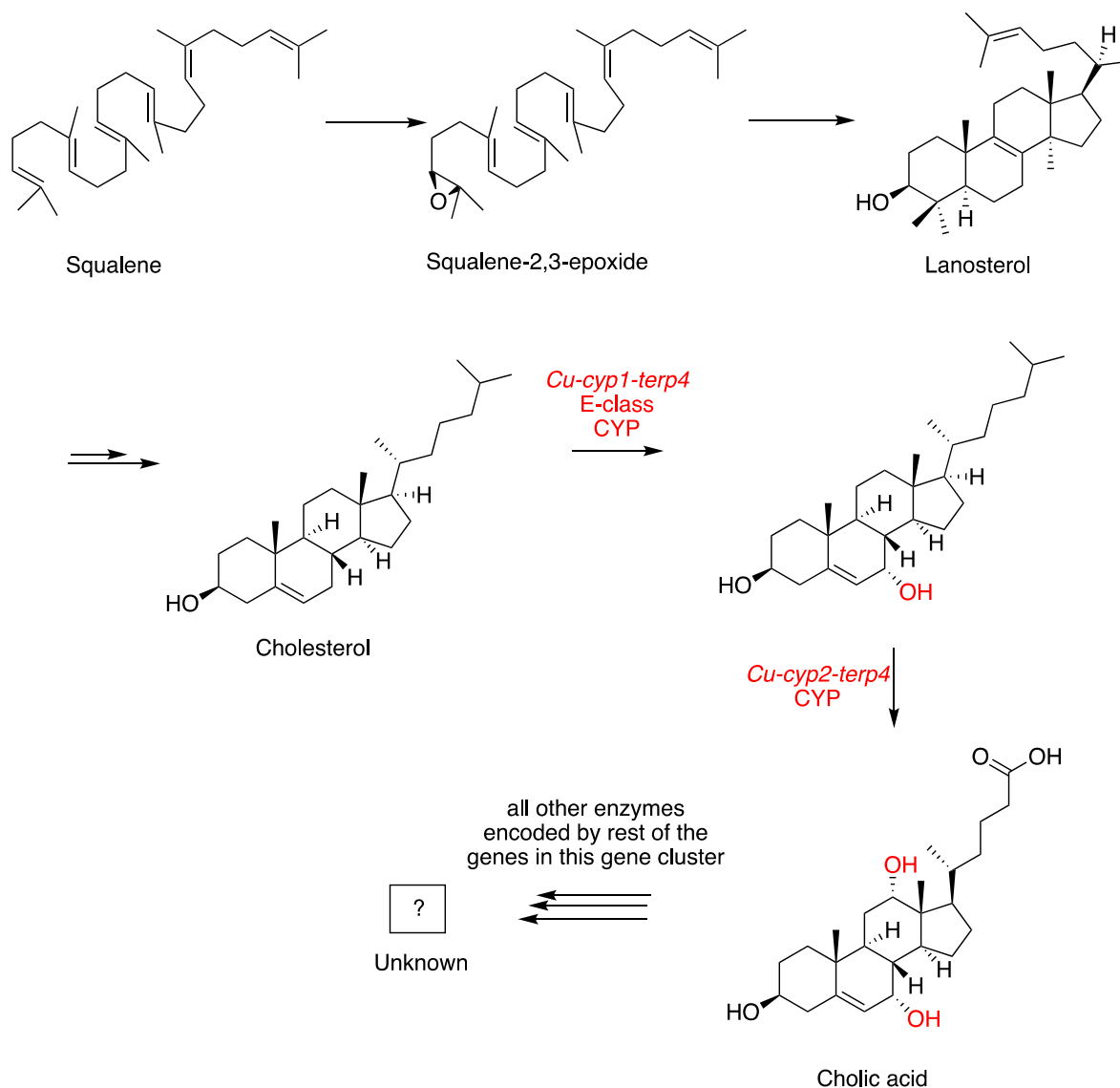




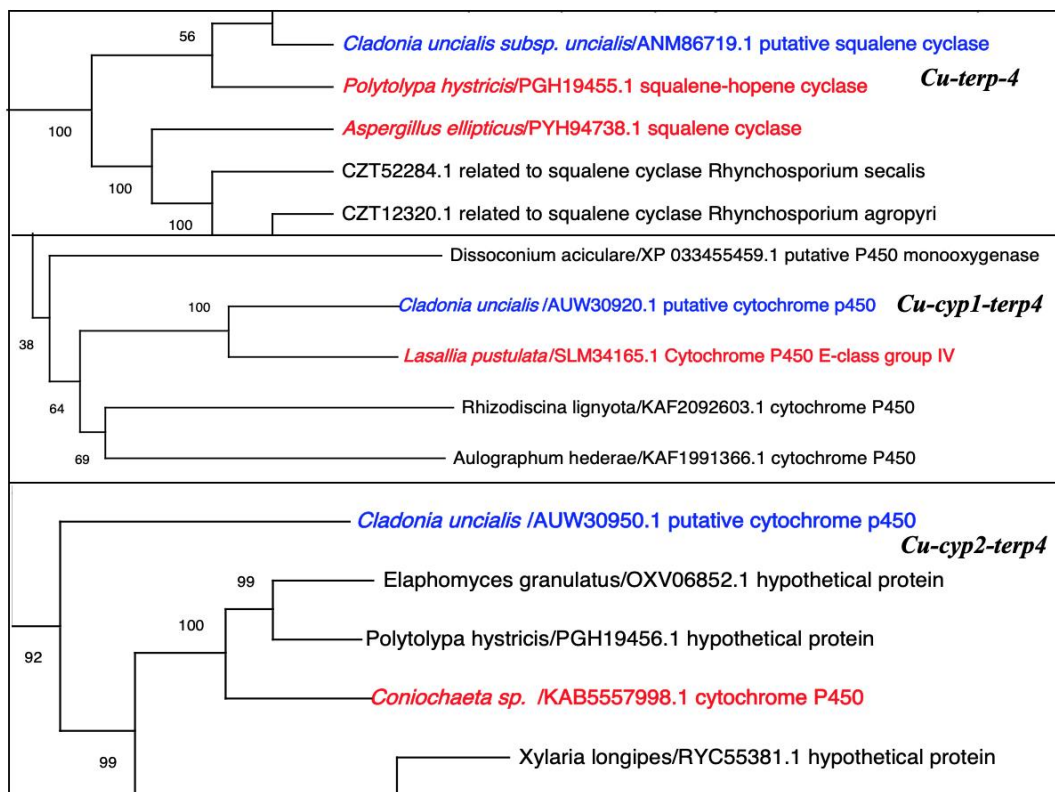
**Figure 6.1:** Terpene synthase gene clusters in the *C. uncialis* mycobiont genome. Abbreviations: KS, ketosynthase; AT, acyltransferase; DH, dehydratase; MT, C-methyltransferase; ER, enoylreductase; KR, ketoreductase; ACP, acetyl carrier protein; GNAT, Gcn5-related N-acetyltransferase and the representation of cytochrome p450s (tailoring genes); part of these lichen *cu-terp* gene clusters (Reprinted with the permission from *Journal of Natural Products* 2018 81 (4), 723-731. Copyright 2018 American Chemical Society, see Figure S26 in the Appendix for copyright clearance).

The work described in this chapter predicts the function of a putative cytochrome p450, an accessory gene in *Cu-terp-4* cluster, which I have named as *Cu-cyp1-terp4*, (**Figure 6.1**) and numbered as gene 1. Through an extensive sequence database search (BLASTp) we identified the closest homolog of *Cu-cyp1-terp4*; an already known group of CYPs present in a non-lichen fungal species named *Lasallia pustulata* (**Figure 6.2**). A phylogenetic tree was constructed (using Neighbour joining; NJ method) and in this evolutionary analysis, it was discovered that both lichen and fungal CYPs are orthologs and makes a monophyletic group with 100% bootstrap support. This group of CYPs represents class E cytochrome p450 proteins (oxidoreductases) that fall into sequence cluster group IV. Group IV comprises of CYP7A1 (cholesterol 7-alpha-hydroxylase/monooxygenase) type enzymes, which catalyze a rate-limiting step in cholesterol catabolism and bile acid biosynthesis by introducing a hydrophilic moiety at position 7 of cholesterol, ultimately converts cholesterol to 7-alpha-hydroxycholesterol. This enzyme is important for cholesterol homeostasis and inhibition of cholesterol 7-alpha-hydroxylase (CYP7A1) represses bile acid biosynthesis [Li, T. et al., 2013]. It is also proposed that the second

CYP in *Cu-terp-4* cluster (numbered as gene 8) named as *Cu-cyp2-terp4*, catalyzes the second hydroxylation step of 7- $\alpha$ -hydroxycholesterol, that leads to the formation of cholic acid (a primary bile acid) as shown in **Scheme 6.1**. The function of other genes could not be predicted.



**Scheme 6.1:** Cytochrome p450 part of terpene synthase gene cluster found in *C. uncialis* appears to be genetically similar to E-class group IV CYPs that catalyze the hydroxylation of cholesterol to 7- $\alpha$  hydroxycholesterol.



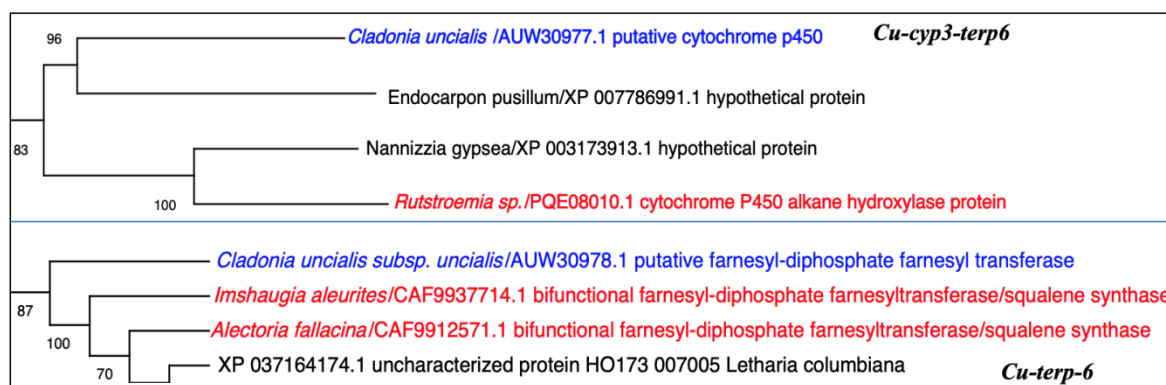
**Figure 6.2:** Truncated phylogenetic trees of *Cu-terp-4*, *Cu-cyp1-terp4* & *Cu-cyp2-terp4* illustrates the degree of relationship between *C. uncialis* genes and non-lichen fungal genes (PDF version of complete phylogenetic trees are in Appendix Figure S1-S3).

### 6.2.2. Terpene synthase gene cluster, *Cu-terp-6*

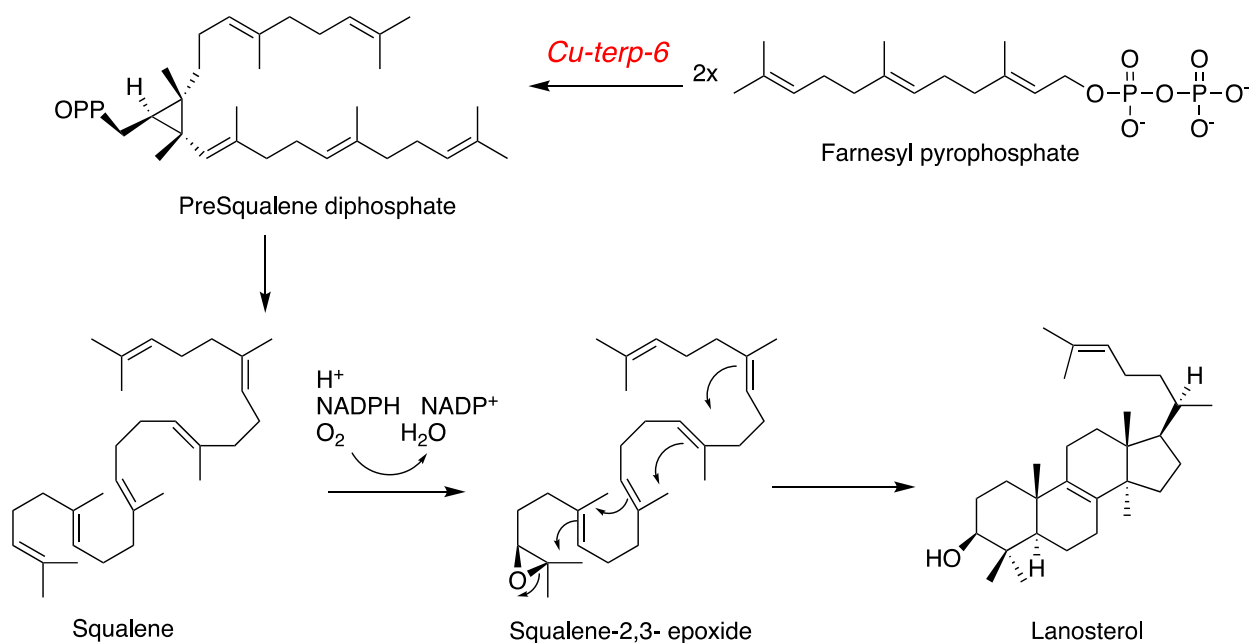
Bertrand in 2018 reported a farnesyl diphosphate farnesyl transferase, also known as squalene synthase, labelled *Cu-terp-6*, that catalyzes the synthesis of squalene via condensation of two molecules of farnesyl pyrophosphate (**Figure 6.1**) [Ha, N. T., & Lee, C. H. 2020].

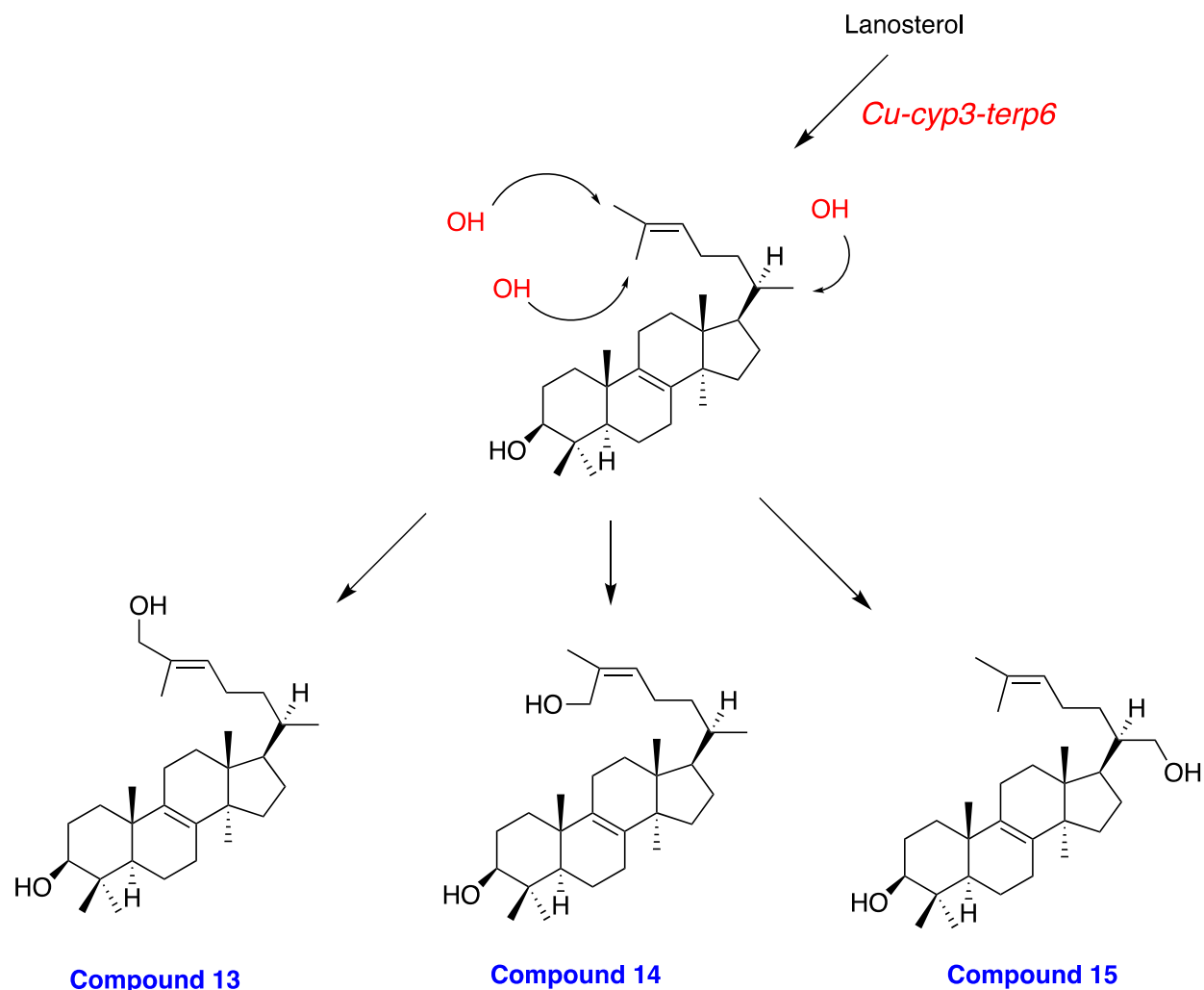
A cytochrome p450 upstream of this squalene synthase in this gene cluster (**Figure 6.1**) is named *Cu-cyp3-terp6*. The function of this cytochrome p450 is predicted to be involved in the catalysis of terminal hydroxylation of the alkyl side chain in lanosterol (which is synthesized from squalene-2,3-epoxide). The phylogenetic tree shows a high homology of *Cu-cyp3-terp6* with a fungal (*Rutstroemia* sp.) cytochrome p450 (**Figure 6.3**). Both CYPs are the part of one clade with strong bootstrap support (100% for NJ method). Also, the short evolutionary distance computed

between these two genes in the units of the number of amino acid substitutions per site conclude their biological and functional relatedness. The fungal CYP belongs to a CYP52A13 which is an alkane hydroxylase 2, together with an NADH cytochrome p450 enzyme system catalyzes the terminal hydroxylation as the first step in the assimilation of alkanes and fatty acids. So, based on the phylogenetic reconstruction we predict the function of *Cu-cyp5-terp6* that could catalyze the hydroxylation of any of terminal methyl groups present on alkyl side chain of lanosterol (**Scheme 6.2**) could result in the formation of either of the ‘3’ novel molecules named ‘**Compound 13-15**’.



**Figure 6.3:** Truncated phylogenetic trees of *Cu-terp-6* & *Cu-cyp3-terp6* illustrates the degree of relationship between *C. uncialis* genes and non-lichen fungal genes (PDF version of complete phylogenetic trees are in Appendix Figure S4 & S5).



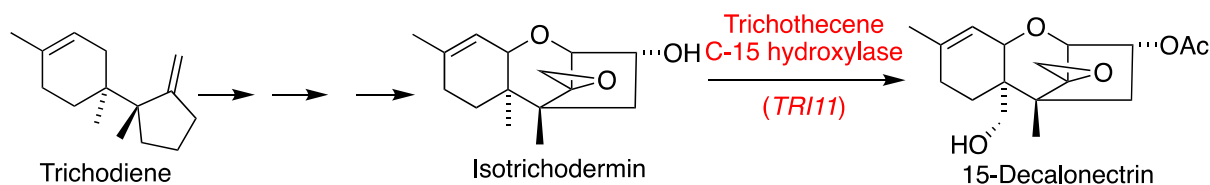


**Scheme 6.2:** Proposed biosynthetic pathway of a novel ‘**Compounds 13-15**’ catalyzed by terpene synthases and an accessory enzyme cytochrome p450 encoded by two genes in *Cu-terp-6* gene cluster (**Figure 31**).

### 6.2.3. Terpene synthase gene cluster, *Cu-terp-9*

The function of two p450 genes, numbered as genes 1 & 3 in terpene synthase the gene cluster labelled *Cu-terp-9* (**Figure 6.1**) could be predicted based on BLAST consensus similarity and alignment to be similar to functionally characterized genes deposited in GenBank (genetic homology). Homologs and orthologs of these two p450 matches found in non-lichen fungal species using iterative phylogenetic reconstruction (NJ method). The first p450 in *Cu-terp-9* gene cluster named as *Cu-cyp4-terp9* shows high homology with two genes coding for fungal p450. One gene found in *Elsinoe australis* encodes for cytochrome p450 monooxygenase like protein and the other

gene found in *Sphaerospora brunnea* encode for cytochrome p450 oxidoreductase (bootstrap support is 70%) (**Figure 6.4**). All three share the common ancestor and have evolved in a very recent time frame which means the evolutionary distance between all these three p450 genes is less and, they all are the part of one clade (**Figure 6.4**). Based on the phylogenetic results we can propose that the function of the proteins encode by these CYP genes is more likely to be similar. Again, based on phylogenetic reconstruction results, the second p450 in *Cu-terp-9* gene cluster named as *Cu-cyp5-terp9* has been found orthologous to a known fungal p450 gene encodes trichothecene C-15 hydroxylase protein (TRI11) found in *Elsinoe australis* as shown in **Figure 6.4**. The tree shows that these two genes form part of one clade and are likely strongly related in their functional properties. **Scheme 6.3** shows the hydroxylation step catalyzed by TRI11; an enzyme encoded by *TR111* cytochrome p450 monooxygenase which is a part of trichothecene gene cluster [Schmidt-Dannert C. 2015].



**Scheme 6.3:** Hydroxylation reaction catalyzed by *TR111*; trichothecene C-15 hydroxylase that converts isotrichodermin to 15-decalonectrin [Schmidt-Dannert C. 2015].

It is therefore possible to propose the function of these two CYP enzymes (**Scheme 6.4**), where one (*Cu-cyp5-terp9*) catalyzes the hydroxylation of methyl group of aristolochene molecule and leads to the formation of ‘**Compound c**’ (hydroxylated aristolochene), while the other (*Cu-cyp4-terp9*) acts as an oxidoreductase that catalyzes the oxidation of squalstatin tetraketide chain and leads to the formation of ‘**Compound d**’.

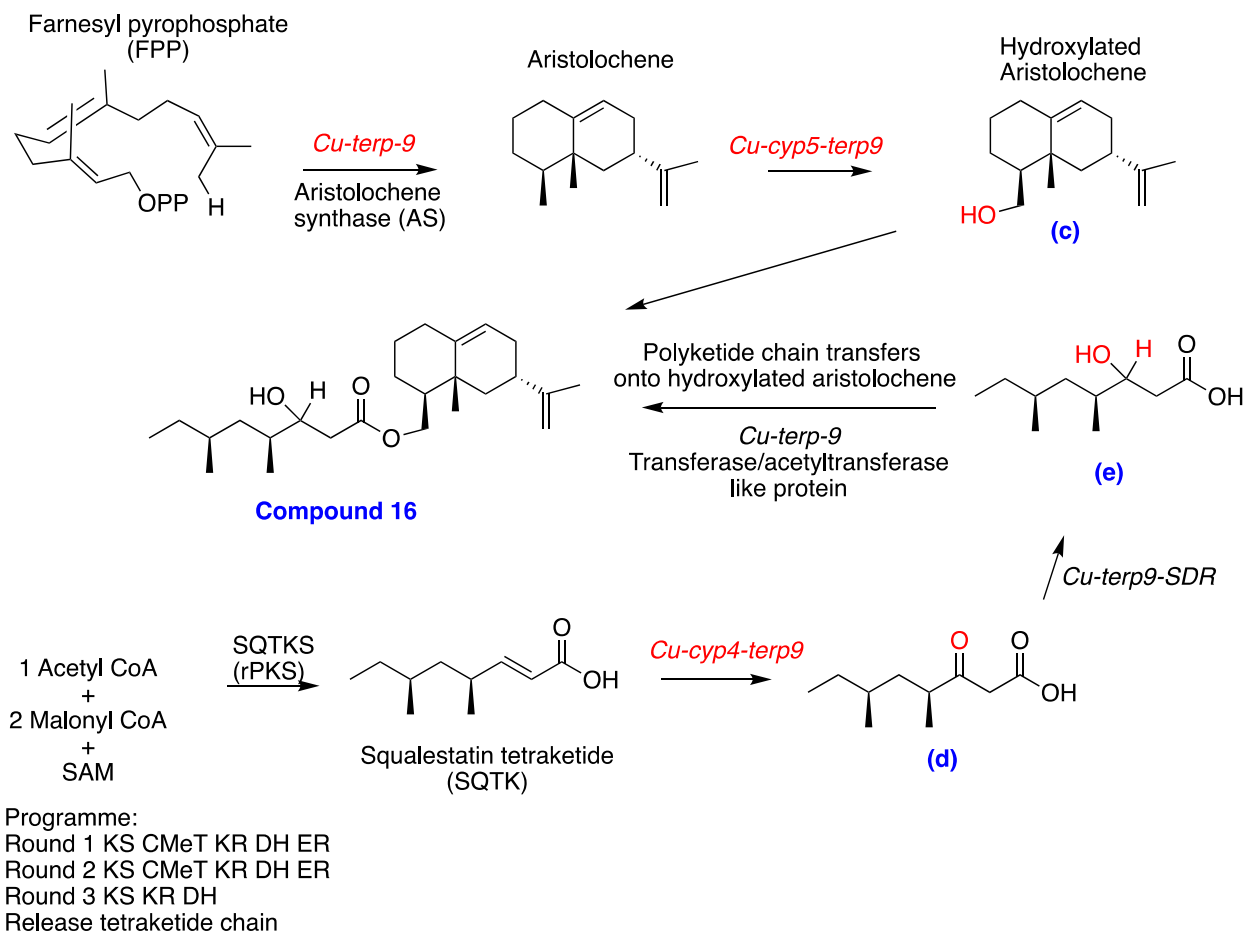
After analyzing the evolutionary history and the functional aspect of p450 genes, we deciphered through the sequences of other predicted genes in this gene cluster assembly. The aim

was to predict their function based on their genetic similarity, BLAST consensus alignment and homology comparison using phylogenetics with already characterized non lichen fungal genes deposited in GenBank. There are four more putative genes other than CYPs, aristolochene synthase (AS) reported earlier [Bertrand, R. L. et al., 2018], reducing polyketide synthase (*Cu-rPKS-2*) shows high homology with squalestatin synthase found in a fungus named *Phoma sp.*, short chain dehydrogenase/reductase (SDR) named as *Cu-terp9-SDR* has been found to be highly homologous to short-chain dehydrogenase like protein 44 found in *Elsinoe fawcetti* (fungus) with high bootstrap support of 87% and cytochrome p450 transferase/acetyltransferase shows high homology with transferase like protein found in *Elsinoe fawcetti* with high bootstrap support of 98%.

Aristolochene is a biosynthetic precursor of several toxins [Deligeorgopoulou, A. and Allemann, R. K. 2003]. Here, we propose the biosynthetic pathway of a putative SM which might show similar properties as aristolochene (a toxin molecule) with terpenoid core and a polyketide side chain (**Compound 16**) (**Scheme 6.4**). In this biosynthetic route, the first reaction is catalyzed by aristolochene synthase (AS) (*Cu-terp-9*) that converts farnesyl pyrophosphate (FPP, alicyclic sesquiterpene precursor) to aristolochene (a sesquiterpene) [Agger, S. et al., 2009] (**Scheme 6.4**). Aristolochene gets converted to '**Compound c**' by the action of p450 enzyme (*Cu-cyp4-terp9*) (as discussed above) (**Scheme 6.4**).

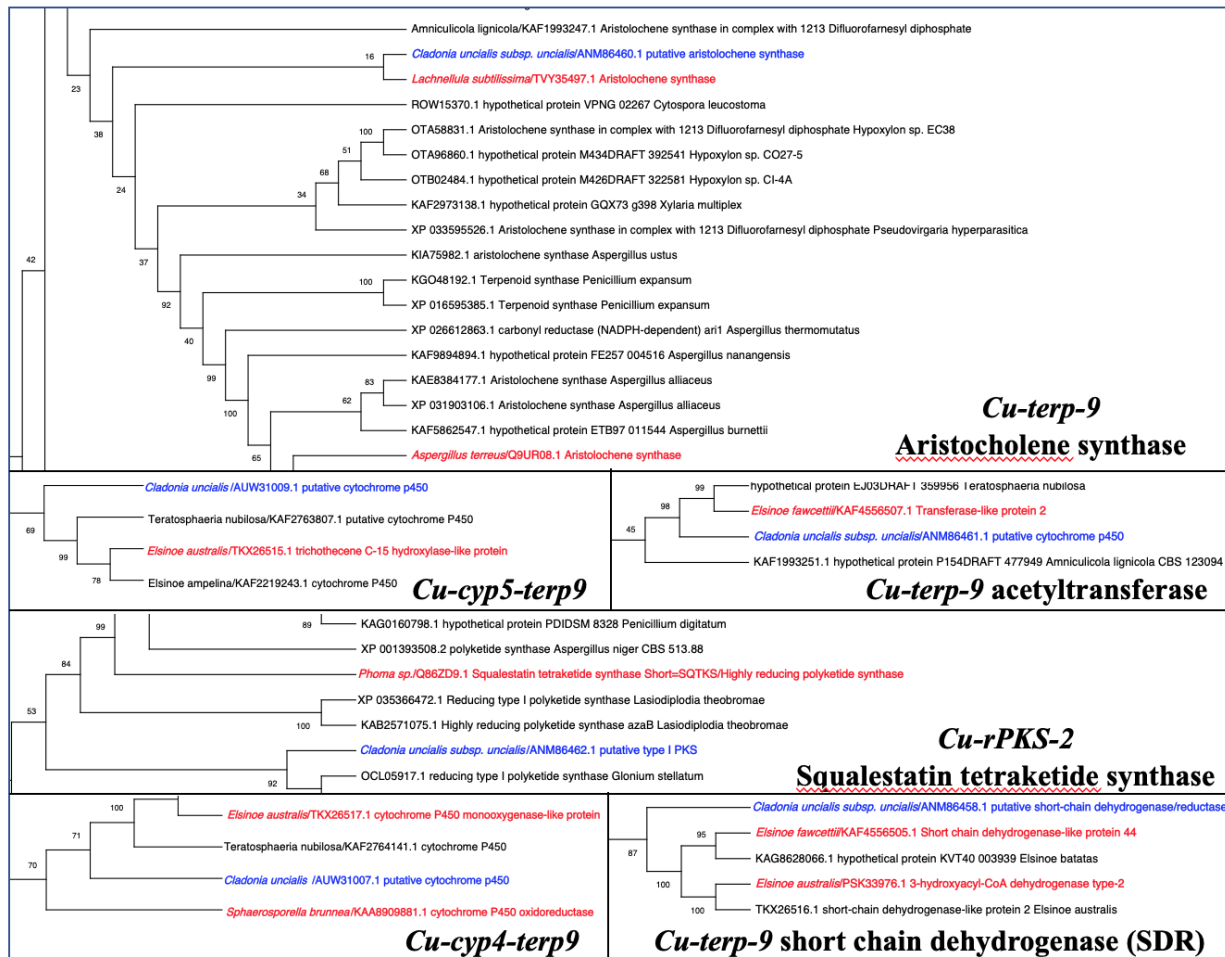
The second part of this biosynthetic route involves the synthesis of squalestatin tetraketide (SQTK), inhibits the formation of cholesterol and a key intermediate in the biosynthesis of squalestatin S1 [Lebe, K. E. and Cox, R. J. 2019]. Squalestatin tetraketide synthase (SQTKS), a reducing polyketide synthase (rPKS) catalyzes the synthesis of SQTK which followed by the action of p450 enzyme (*Cu-cyp4-terp9*) results in '**Compound d**' (as discussed above) (**Scheme 6.4**). *Cu-terp9-SDR* catalyzes the reduction of ketone to -OH moiety of '**Compound d**' could lead

to the formation of ‘**Compound e**’. The final step is proposed to be catalyzed by acetyl transferase which transfers the polyketide chain (Compound c) onto hydroxylated aristolochene (**Compound c**). Which could result in a novel molecule named as **Compound 16**.



**Scheme 6.4:** Putative biosynthetic pathway of a novel molecule with a terpenoid core & a polyketide side chain; **Compound 16**.

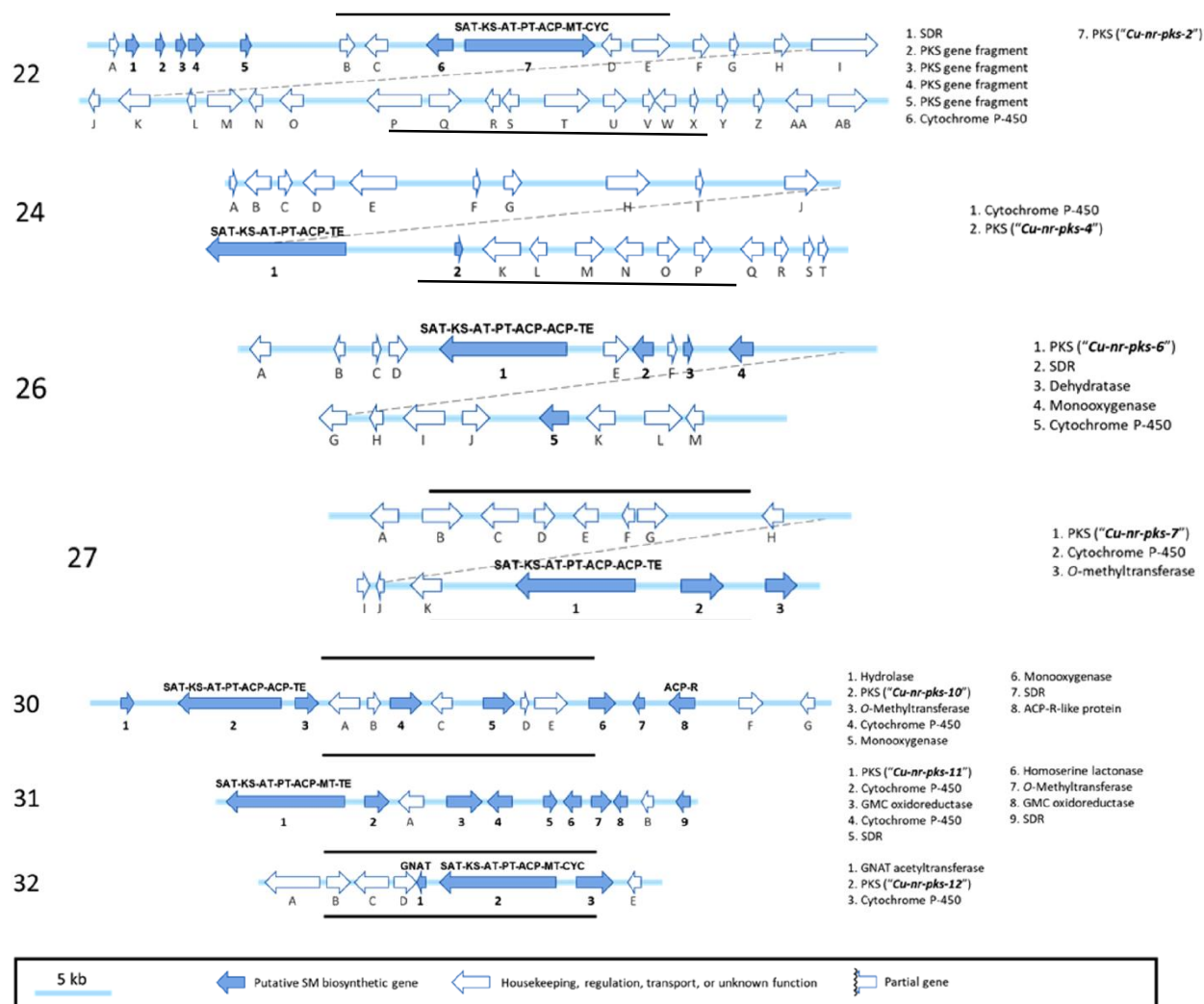




**Figure 6.4:** Truncated phylogenetic trees of *Cu-terp-9*, *Cu-cyp5-terp9*, *Cu-terp9 acetyltransferase*, *Cu-rPKS-2*, *Cu-cyp4-terp9* & *Cu-terp6-SDR* illustrates the degree of relationship between *C. uncialis* genes and non-lichen fungal genes (PDF version of complete phylogenetic trees are in Appendix Figure S6-S11).

### 6.3. Non-reducing polyketide synthase gene clusters

As discussed in **Chapter 3**, out of 32 putative type I PKS genes found in *C. uncialis*, 14 genes are non-reducing named from *Cu-nr-pks-1* to *Cu-nr-pks-14* [Abdel-Hameed, M. et al., 2016; Bertrand, R. L., 2018] (**Figure 6.5**). These fungal NR-PKSs are a family of multidomain enzymes that generate substituted aromatic natural products. Structural diversity is achieved through “programmed” selection of starter unit, polyketide chain length, regioselectivity of backbone cyclization(s), and mode of product release.

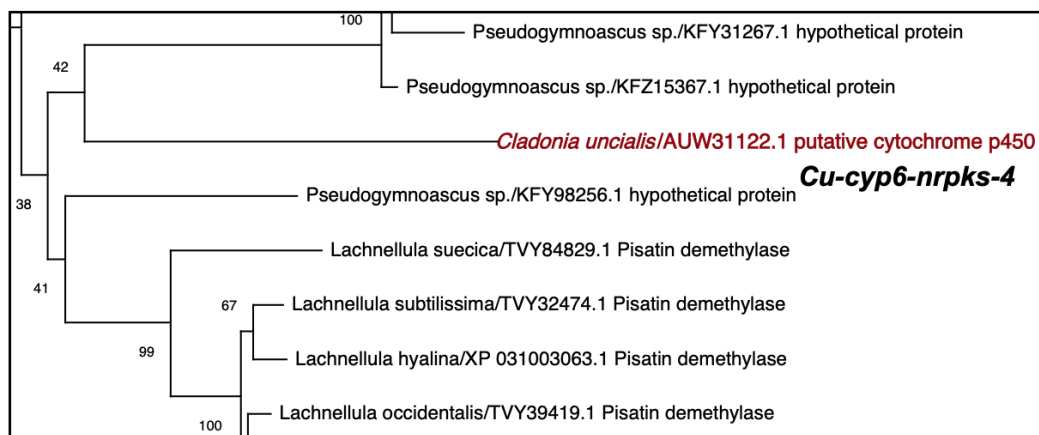


**Figure 6.5:** Type I nonreducing polyketide synthase (PKS) gene clusters in the *C. uncialis* mycobiont genome. Abbreviations: SAT, starter acyltransferase; KS, ketosynthase; AT, acyltransferase; PT, product template domain; MT, C-methyltransferase; ACP, acetyl carrier protein; KR, ketoreductase; DH, dehydratase; ER, enoylreductase; CYC, Claisen cyclase; TE, thioesterase; R, reductase; GNAT, Gcn5-related N-acetyltransferase and representation of cytochrome p450s as post-PKS/tailoring genes present in these gene clusters (Reprinted with the permission from *Journal of Natural Products* 2018 81 (4), 723-731. Copyright 2018 American Chemical Society, see Figure S26 in the Appendix for copyright clearance).

### 6.3.1. Non-reducing polyketide synthase gene cluster, *Cu-nr-pks-4*

A functionally characterized homolog of a non-reducing polyketide synthase could not be observed in this gene cluster (*Cu-nr-pks-4*) (Figure 6.5). However, the functionality of p450 named as *Cu-cyp6-nrpk-4* gene in this gene cluster (Figure 6.5) could be predicted. Based on protein BLAST search (Table 6.1); 99% query cover proposes a genetic proximity of *C. uncialis*

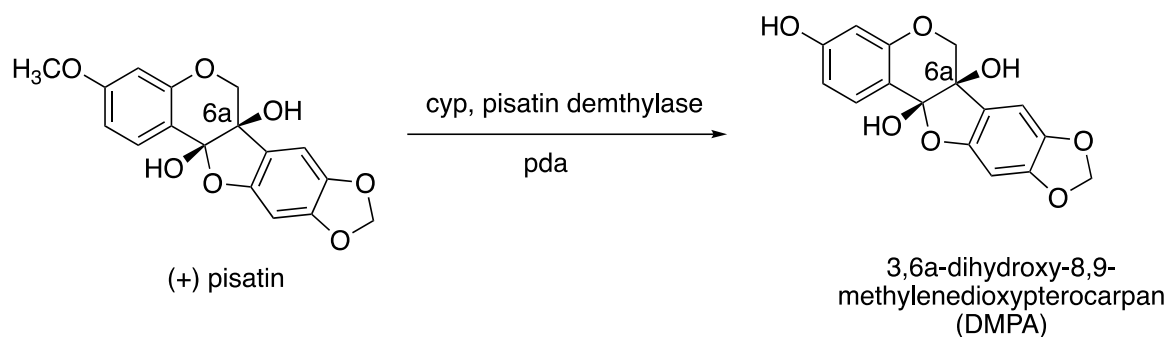
CYP to a fungal CYP found in many species of *Lachnellula* fungus. In spite of the low bootstrap support value in the phylogenetic tree reconstruction (**Figure 6.6**), we could still predict the putative function of CYP protein encoded by *Cu-cyp6-nrpks-4*, that has been found to be homologous to a fungal pisatin demethylases (PDAs). This PDA gene was initially found in a pea pathogen *Nectria haematococca* mating population VI (MPVI); which is a filamentous ascomycete, amongst few fungal species which were tolerant to an antimicrobial named phytoalexin pisatin produced by pea plants [Delserone, L.M. et al., 1999; George, H. L. & VanEtten, H. D., 2001]. These fungal species circumvent the host defense mechanism by detoxification of pisatin (a phytoalexin) and studies have shown that first step of degradation of pisatin molecule is catalyzed by a cytochrome p450 (Pda) via demethylation (**Scheme 6.5**) [Delserone, L.M. et al., 1999; George, H. L. & VanEtten, H. D., 2001]. Later, by performing some gene disruption experiments in other fungal species such as *Fusarium oxysporum*, *Fusarium polyferatum* the function of Pda was determined [Milani, N. A. et al., 2012] and BLASTp results show a strong genetic similarity of this fusarium Pda to cytochrome p450s found in *Lachnellula* species (**Table 6.2**).



**Figure 6.6:** Truncated phylogenetic tree of cytochrome p450; *Cu-cyp7-nrpks4* shows an evolutionary relationship between lichen CYP/AUW31122.1 to a known fungal CYP (pisatin demethylase; Pda) found in many *Lachnellula* species (PDF version of complete phylogenetic trees are in Appendix Figure S12).

**Table 3:** BLAST statistics of *C. uncialis* cytochrome p450 (*Cu-cyp6-nrpks-4*) which has been found genetically similar to pisatin demethylase, an enzyme found in many species of *Lachnellula* fungus.

Description	Max. Score	Total Score	Query Cover	E value	Per. Identity	Accession no.
<b>putative cytochrome p450 [<i>Cladnoia uncialis</i>]</b>	284	284	100%	8e-97	100%	AUW31122.1
<b>Pisatin demethylase [<i>Lachnellula willkommii</i>]</b>	104	104	99%	4e-25	40.88%	TVY87722.1



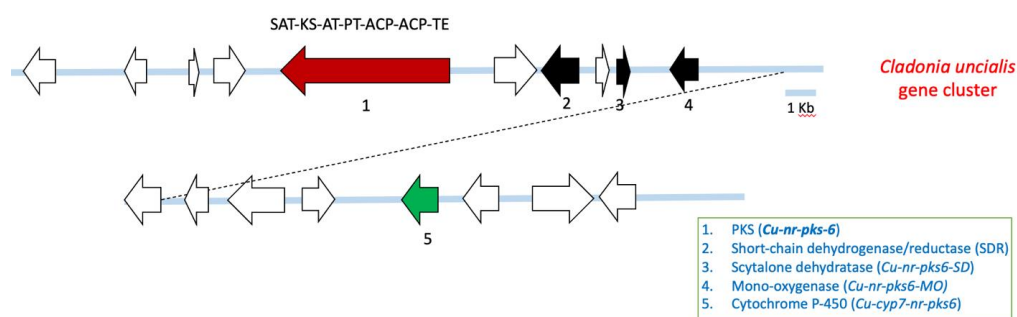
**Scheme 6.5:** Enzymatic demethylation of (+) pisatin catalyzed by pisatin demethylase into 3,6a-dihydroxy-8,9-methylenedioxypterocarpan (DMDP).

**Table 6.2:** Blast statistics that show a genetic similarity of CYP named as pisatin demethylase in *Lachenulla* fungus to a known CYP found in *F. proliferatum* whose function is also described.

Description	Max. Score	Query Cover	E value	Per. Identity	Accession no.
<b>Pisatin demethylase [<i>Lachnellula willkommii</i>]</b>	281	100%	7e-95	100%	TVY87722.1
Pisatin demethylase [ <i>Lachnellula cervina</i> ]	272	100%	2e-91	96.27%	TVY52984.1
Hypothetical protein [ <i>Exophiala xenobiotica</i> ]	184	99%	2e-54	62.41%	XP_013320560.1
Pisatin demethylase cytochrome p450 [ <i>Fusarium proliferatum ET1</i> ]	182	97%	2e-52	58.46%	XP_0131090391.1

### 6.3.2. Non-reducing polyketide synthase gene cluster, *Cu-nr-pks-6*

The *Cu-nr-pks-6* gene cluster as shown in **Figure 6.7**, has been reported [Bertrand, R. L. et al., 2018] to have one *nr-pks* gene encodes for non-reducing polyketide synthase (backbone enzyme) and, downstream of *nr-pks* there are four accessory genes encode for four post-PKS/tailoring enzymes: short chain dehydrogenase/reductase (*Cu-nr-pks6-SDR*), scytalone dehydratase (*Cu-nr-pks6-SD*), monooxygenase (*Cu-nr-pks6-MO*) &, a cytochrome P-450 (*Cu-cyp7-nr-pks6*).



**Figure 6.7:** Uncharacterized *C. uncialis* gene cluster, *Cu-nr-pks-6* (accession no. AUW31152.1). Gene in red color encodes for non-reducing polyketide synthase, gene in green color encodes for cytochrome p450 (*Cu-cyp6-nr-pks-6*), an accessory enzyme. And genes in black color encode for rest of the accessory enzymes: short-chain dehydrogenase/reductase (SDR), scytalone dehydratase (SD) & monooxygenase (MO). Domain abbreviations: Starter acyltransferase (SAT), ketosynthase (KS), acyltransferase (AT), product template (PT), acetyl carrier protein (ACP), thioesterase (TE).

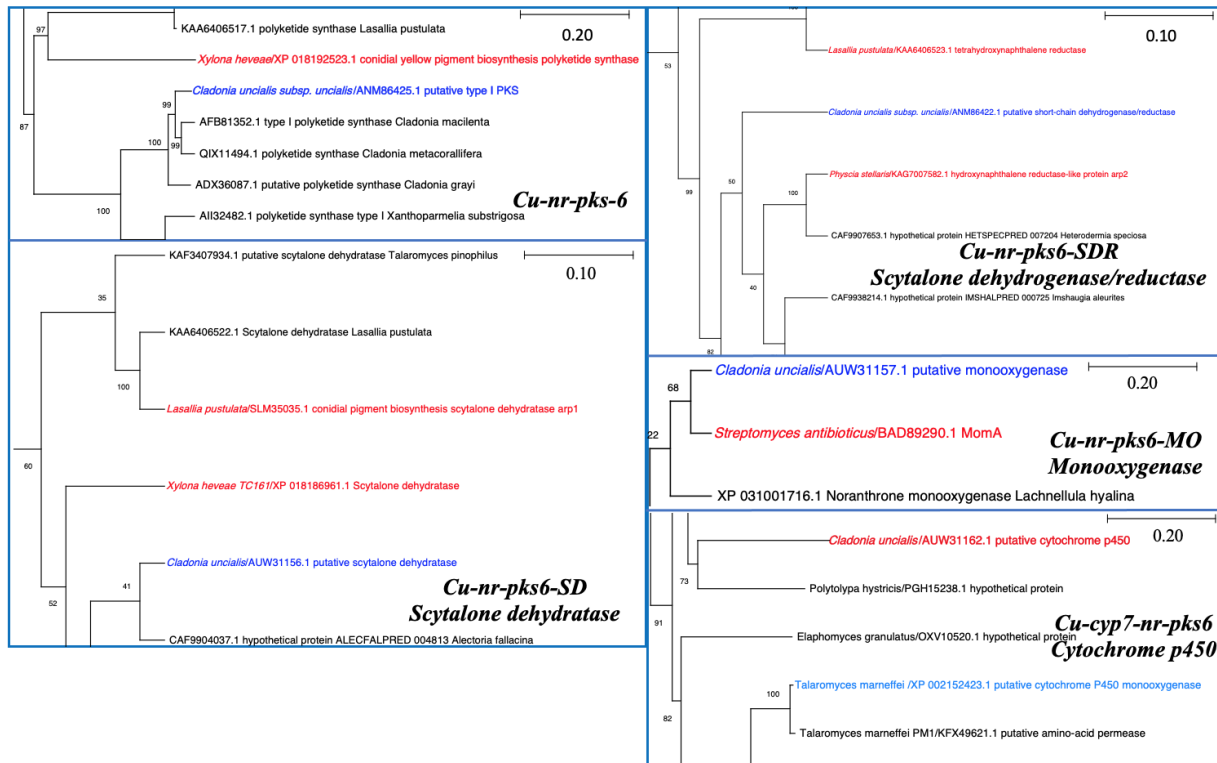
A putative biosynthetic pathway is proposed for *Cu-nr-pks-6* gene cluster, which suggests the biosynthesis of six novel polyketide molecules (**Compounds 17-22, Scheme 6.8**) that are unique to *C. uncialis*.

As shown in **Table 6.3**, two homologs of non-reducing PKS gene (part of *Cu-nr-pks-6* gene cluster) have been found after an extensive protein BLAST search that shows 99% query cover and approximately 53-54% amino acid (A.A.) identity to already known non-lichen fungal type I NR-PKSs (found in *Hyaloscypha variabilis* F & *Xylona heveae* species). The topology of the phylogenetic tree (using NJ method) (**Figure 6.8**) suggests that *Cu-nr-pks* and NR-PKS (*Xylona heveae*) genes have evolved quite recently based on short horizontal distance between

them through the node. Both genes are the part of two separate in-groups descending from one node (share a common ancestor) with 87 bootstrap support that ultimately makes a stronger case of their functional similarity.

**Table 6.3:** BLAST statistics of *C. uncialis* NR-PKS (*Cu-nr-pks-6*) which has been found genetically homologous to functionally characterized NR-PKS found in a non-lichen fungal species (*Xylona hevae*) and has been reported to be involved in the biosynthesis of a yellow pigment (Melanin).

Description	Max. Score	Query Cover	E value	Per. Identity	Accession no.
<b>Type I non-reducing pks</b> [ <i>Cladonia uncialis</i> subsp. <i>Uncialis</i> ]	<b>4185</b>	<b>100%</b>	<b>0</b>	<b>100%</b>	<b>AUW31152.1</b>
<b>BcPKS12, polyketide synthase</b> [ <i>Hyaloscypha variabilis</i> F]	<b>2214</b>	<b>99%</b>	<b>0</b>	<b>54%</b>	<b>PMD39418.1</b>
<b>conidial yellow pigment biosynthesis pks</b> [ <i>Xylona hevae</i> TC161]	<b>2155</b>	<b>99%</b>	<b>0</b>	<b>52.86%</b>	<b>XP_018192523.1</b>

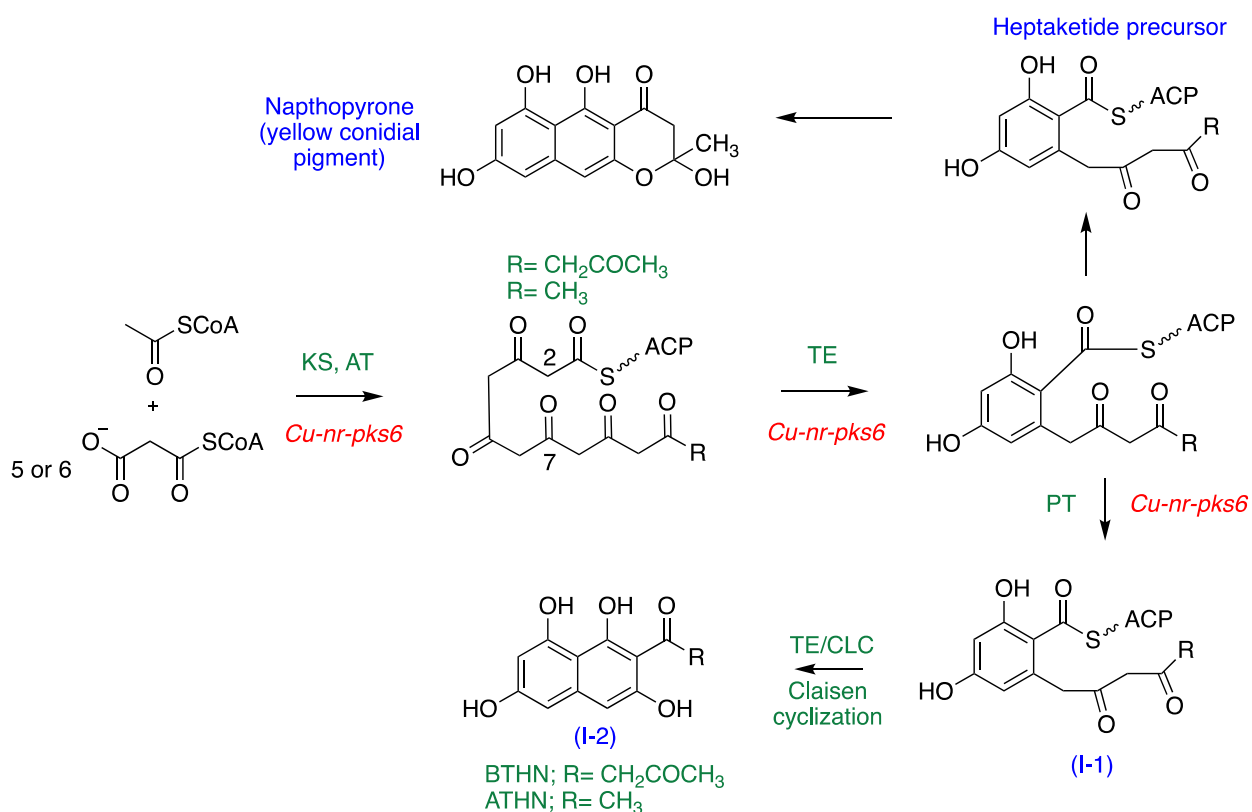


**Figure 6.8:** Truncated phylogenetic trees of *Cu-nr-pks-6*, *Cu-nr-pks6-SD*, *Cu-nr-pks6-SDR*, *Cu-nr-pks6-MO* and *Cu-cyp7-nrpk-6* illustrate the degree of relationship between *C. uncialis* genes and non-lichen fungal genes (PDF version of complete phylogenetic trees are in Appendix Figure S13-S17).

On the contrary NR-PKS gene found in *Hyaloscypha variabilis* *F* is very far from *Cu-nr-pks* in the phylogenetic tree (review **Figure S13** in the Appendix), which suggests a low functional similarity between these two genes. Therefore, we propose that *Cu-nr-pks-6* (NR-PKS) is orthologous to type I fungal NR-PKS (*Xylona heveae*) based on both the BLASTp, and phylogenetic reconstruction results. According to the BLAST search Type I NR-PKS found in *Xylona heveae* catalyzes the biosynthesis of a conidial yellow pigment molecule. This type I NR-PKS is homologous to wA naphthopyrone synthase found in *A. nidulans* that catalyzes the synthesis of a heptaketide naphthopyrone (a conidial yellow pigment molecule) (**Scheme 6.6**) [Fujii, I et al., 2001]. *BcPKS12* is involved in the biosynthesis of 1,8-dihydroxynaphthalene (DHN)-melanin pigment found in *Botrytis cinerea* [Zhu, P. et al., 2017]. The heptaketide naphthopyrone is a catalytic intermediate which is produced during melanin biosynthesis by inactivating or deletion of thioesterase/Claisen cyclase (TE/CLC) domain of NR-PKS [Watanabe, A. et al., 2000, Fujii, I et al., 2001]. The absence of TE/CLC domain results in the spontaneous O–C cyclization to release full-length pyrone shunt product, a naphthopyrone [Crawford, J. M. et al., 2010].

However, if the TE/CLC domain is present or fully functional in NR-PKS, it catalyzes C–C bond formation via Claisen cyclization. Hence, we propose the conversion of **intermediate I-1** to **intermediate I-1** catalyzed by TE/CLC domain of CU-NR-PKS (part of *Cu-nr-pks-6* gene cluster) (**Scheme 6.6**). Many of the recently published articles on melanin biosynthesis have discussed the difference in the architecture and the substrate specificity of these NR-PKSs results in the formation of three possible polyketide intermediate such as pentaketide, hexaketide and a heptaketide based on the varying number of malonyl CoA molecules [Chang, P. K. et al., 2020, Griffith, S. A. et al., 2018, Fujii, I et al., 2001, Tsai, H. F. et al., 2001, Zhu, P. et al., 2017]. Usually

melanin synthase (NR-PKS) produces a pentaketide intermediate that directly cyclizes to 1, 3, 6, 8-tetrahydroxy-naphthalene (4-THN) [Griffith, S. A. et al., 2018, Thompson, J. E. et al., 2000]. But there are many other fungal NR-PKSs that have been reported to produce hexaketide or heptaketide intermediates in melanin biosynthesis, such as 2-acetyl-1,3,6,8-tetrahydroxy-naphthalene (ATHN), 2-butyl-1,3,6,8-tetrahydroxy-naphthalene (BTHN) and heptaketide naphopyrone (**Scheme 6.6**). *Alb1* gene found in *Aspergillus fumigatus* is one of the examples of these kinds of NR-PKSs [Watanabe, A. et al., 2000].



**Scheme 6.6:** Proposed biosynthesis of 2-butyl-1,3,6,8-tetrahydroxy-naphthalene (BTHN), 2-acetyl-1,3,6,8-tetrahydroxy-naphthalene (ATHN) and, naphopyrone catalyzed by *Cu-nr-pks6*

The usual next step in melanin biosynthesis is the conversion of **intermediate I-2** or heptaketide naphopyrone to 1, 3, 6, 8 -tetrahydroxynaphthalene (4-THN) by the removal of acetyl/butyl side chain in case of **intermediate I-2** & chain-length shortening of heptaketide



naphthopyrone, catalyzed by a separate (other than thioesterases; TE) enzyme known as serine hydrolases [Tsai, H. F. et al., 2001] (**Scheme 6.9**). In the *Cu-nr-pks-6* gene cluster, there are no serine hydrolases and the formation of either ATHN or BTHN (**intermediate I-2**) has been predicted but not 4-THN. The proposed next step is the formation of **intermediate I-3**, catalyzed by *Cu-nr-pks6-SDR* (*C. uncialis* short-chain reductase/dehydrogenase) followed by a dehydration step catalyzed by *Cu-nr-pks6-SD* (scytalone dehydratase) could lead to the formation of **intermediate I-4** (**Scheme 6.7**). The prediction of the function of these two genes is again based on the BLAST consensus alignment and the constructed phylogenetic trees (using NJ method). Evolutionary history shows that *Cu-nr-pks6-SDR* has high homology with a fungal tetrahydroxynaphthalene reductase *arp2* (*Lasallia pustulata*) with a very high 99% bootstrap support, and the evolutionary distance between two genes is very less and similarly *Cu-nr-pks6-SDR* has high homology with a fungal scytalone dehydratase *arp1* (*Lasallia pustulata*) with a considerable bootstrap support value of 60% (**Figure 6.8**).

There are two different possible routes that can be considered here that could lead to the formation of six novel molecules named as **Compounds 17-22** from **Compounds f-i** (**Scheme 6.6**). **Compounds f-g** are basically hydroxy-naphthalenedione derivatives with/without acyl side chain (**Scheme 6.7**). The first route uses the SDR and SD enzymes again in a cyclic fashion and, results in the conversion of **intermediate I-4** to **I-5**. Followed by the oxidation of **intermediates I-5** which is being proposed to be catalyzed by *Cu-nr-pks6-MO* (Monooxygenase) could result in two different molecules '**Compounds h & i**' (**Scheme 6.7**). The function of this monooxygenase has been proposed based on the phylogenetic relationship (using NJ method) between *Cu-nr-pks6-MO* and, a known bacterial monooxygenase MomA with the 68% bootstrap support as shown (**Figure 6.8**). MomA monooxygenase is known to oxidize 4-THN to flaviolin [Funa, N. et al.,

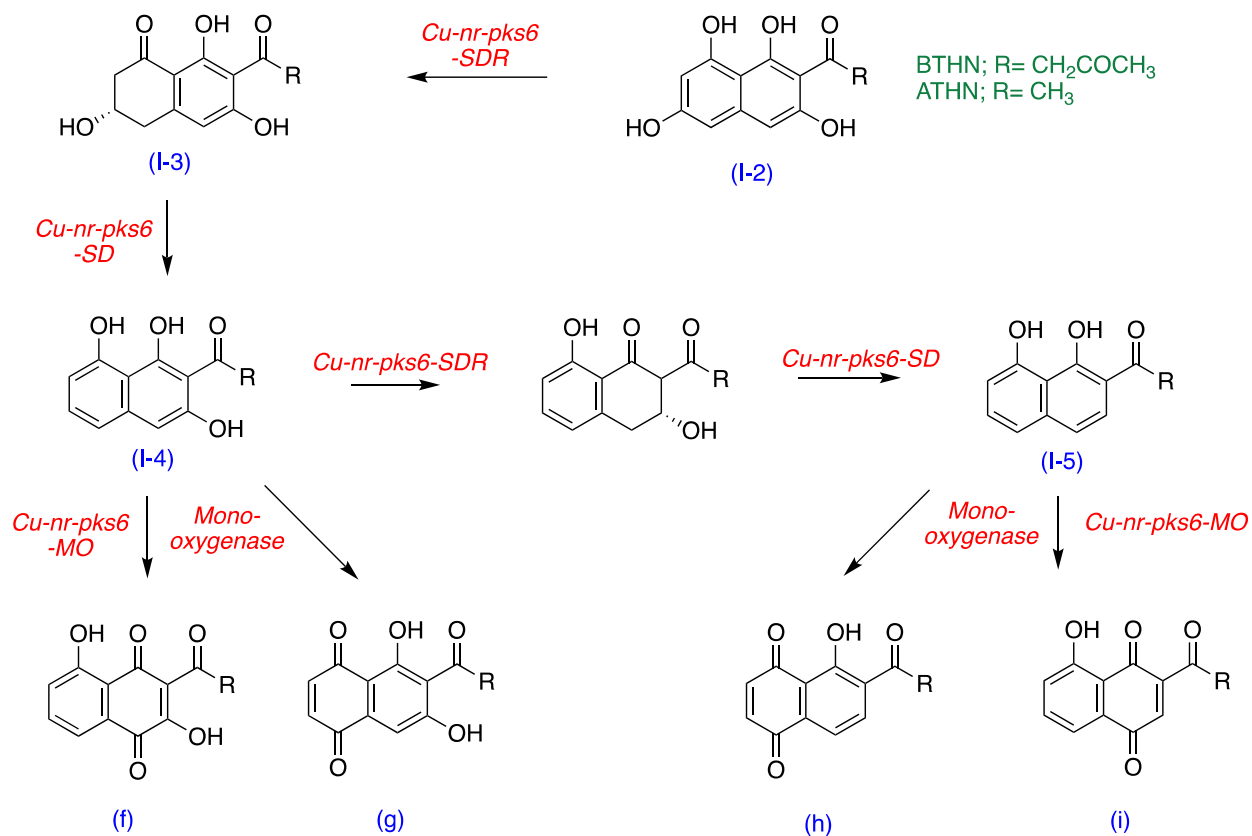
2005]. There was no fungal monooxygenase in the protein BLAST that could be related to *Cu-nr-pks6-MO*.

The final step that is postulated in this biosynthetic pathway is the oxidative dimerization of two molecules of ‘**Compound h**’ catalyzed by *Cu-cyp7-nr-pks6* (*C. uncialis* CYP) via radical coupling that could lead to formation of a novel polyketide, ‘**Compound 20**’ (**Scheme 6.8**). Similarly, two molecules of ‘**Compound i**’ could oxidize to either ‘**Compounds 21 or 22**’ via a radical or cationic mechanisms (**Scheme 6.8**). The functional assignment of this cytochrome p450 is purely theoretical. The closest homolog of *Cu-cyp7-nr-pks6* that have been found in the BLASTp search and through the evolutionary analysis (bootstrap support 91%) is a fungal p450 (*Talaromyces marneffi*). However, this gene has not been functionally characterized yet. The function of *C. uncialis* p450 can be predicted by taking into consideration of catalytic behavior of CYPs found in other NR-PKS gene clusters of *C. uncialis* which is either C-C or C-O oxidative coupling reactions in most of the cases so far.

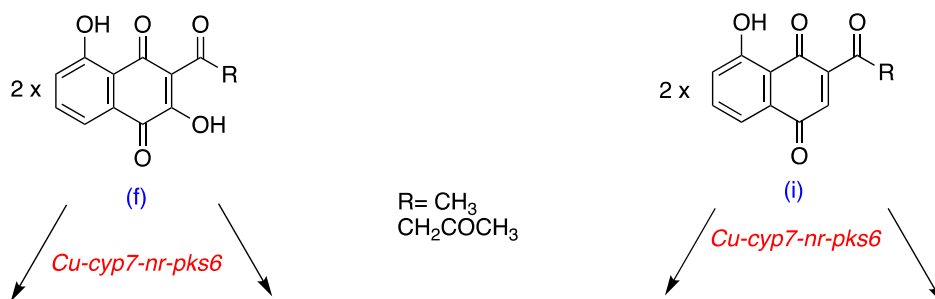
Cytochrome p450 158A2 (CYP158A2) has been found in the soil bacterium *Streptomyces coelicolor* A3(2) [Zhao, B. et al., 2005] and has been reported to catalyze the dimerization of two molecules of flaviolin (flaviolin dimers) via a radical mechanism. In this lichen gene cluster, we do not predict the biosynthesis of flaviolin (oxidation of 4THN catalyzed by MomA, **Scheme 6.9**), because the genes that encode for SDR, and SD enzymes are absent. However, the function of Mom A and CYP158A2 (part of flaviolin gene cluster) appear to be homologous the genes to *Cu-nr-pks6-MO* and, *Cu-cyp8-nr-pks6*.

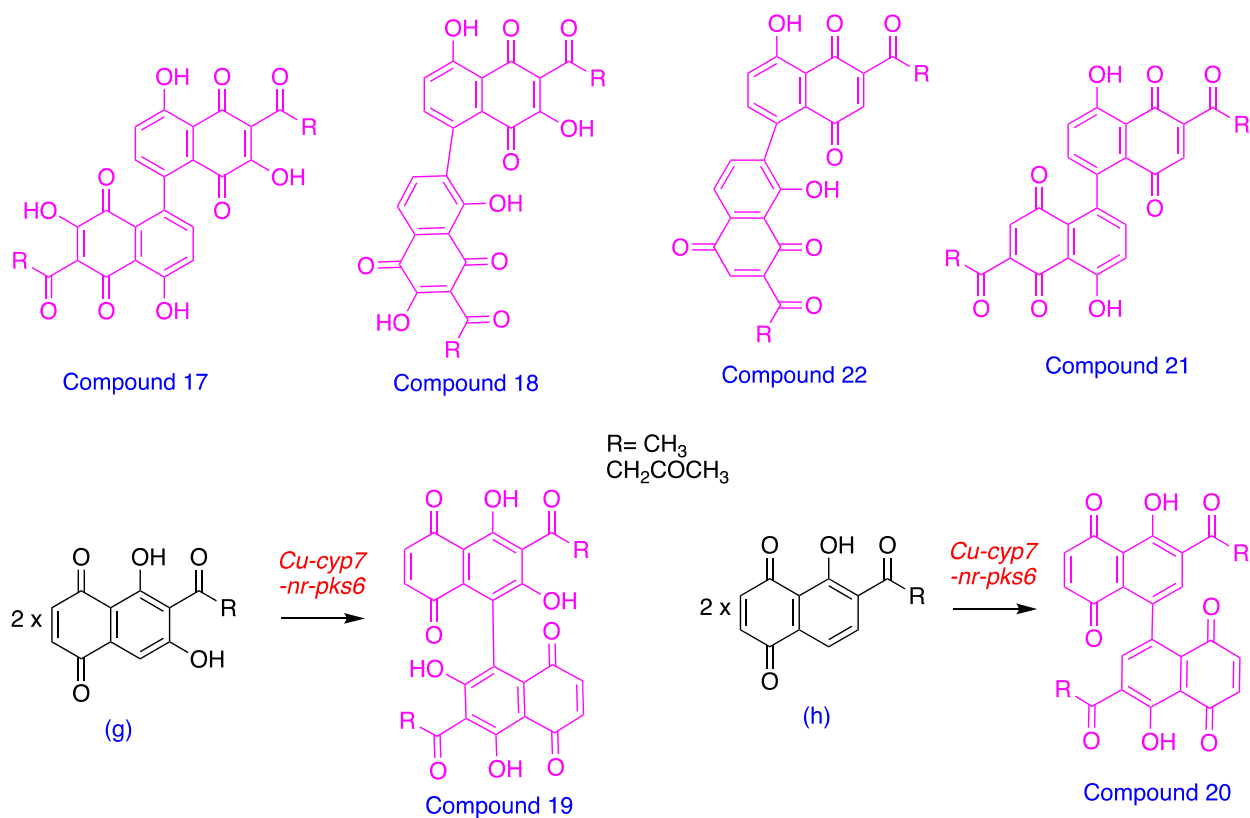
The second proposed route in **Scheme 6.7** is where both SDR and SD enzymes have not been repeated in the catalytic cycle and **intermediate I-4** could be oxidized to ‘**Compounds f & g**’ catalyzed by monooxygenase (MO). Two molecules ‘**Compounds g**’ could be oxidatively

dimerizes to ‘**Compound 19**’ catalyzed CYP enzyme (**Scheme 6.8**). Whereas two molecules of ‘**Compounds f**’ could either dimerizes to ‘**Compound 17**’ or ‘**Compound 18**’ (**Scheme 6.8**).

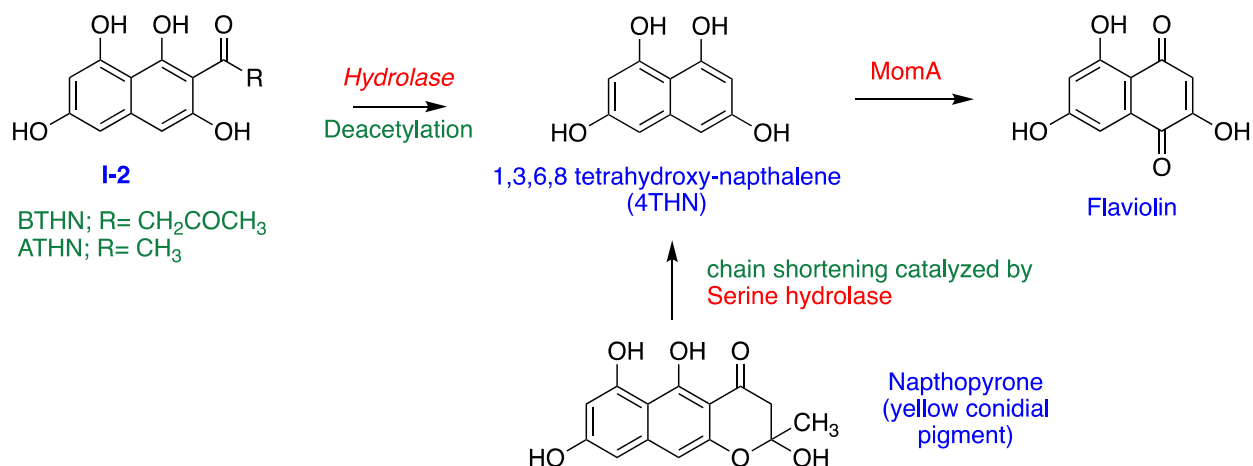


**Scheme 6.7:** Proposed biosynthetic pathway of four novel polyketide molecules named as ‘**Compounds f-i**’.





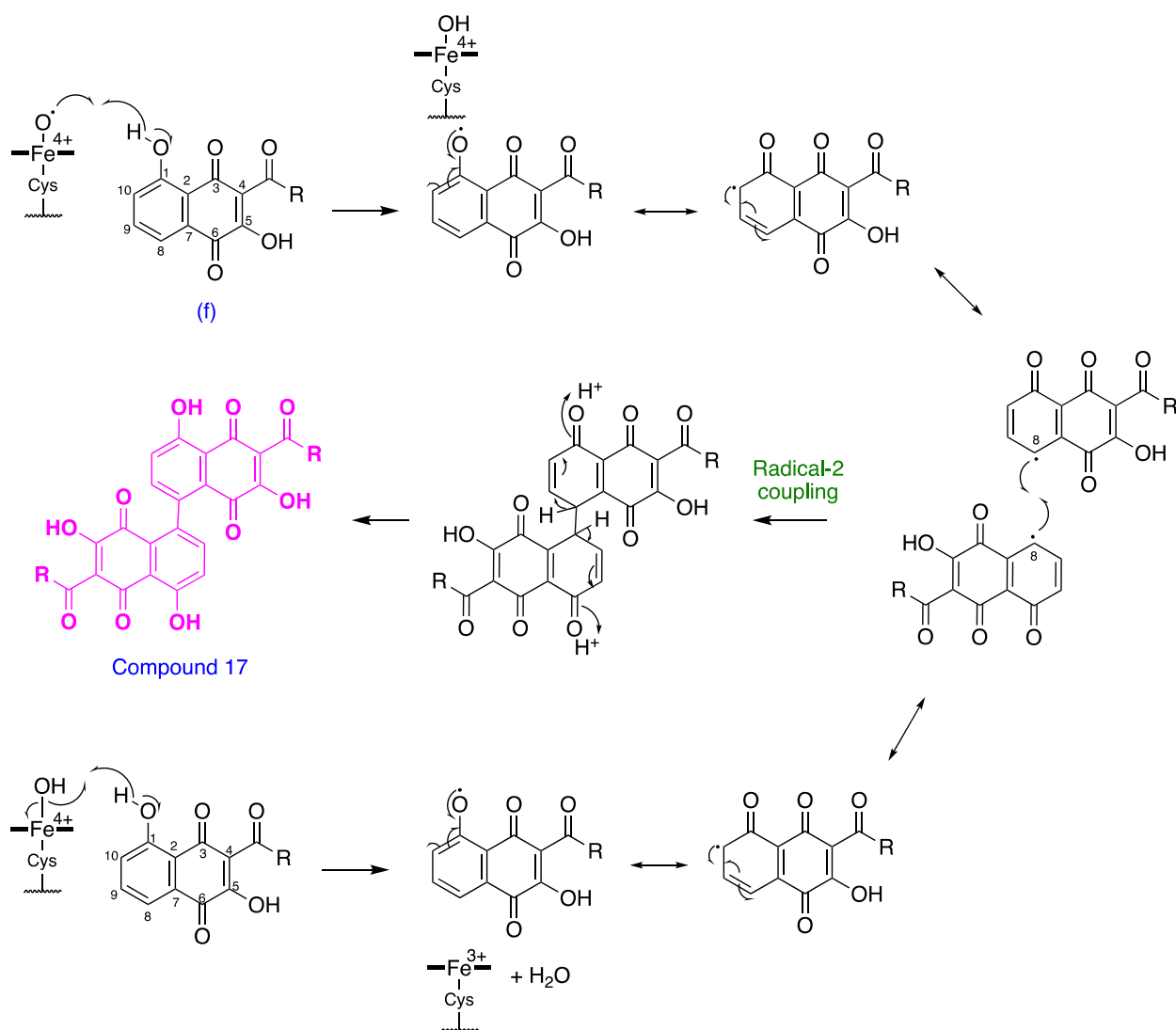
**Scheme 6.8:** Proposed oxidative C-C couplings of ‘Compounds f-1’ to six novel polyketide molecules named ‘Compound 17-22’ catalyzed by *Cu-cyp7-nr-pks6*, where R= CH<sub>3</sub> or CH<sub>2</sub>COCH<sub>3</sub>.



**Scheme 6.9:** Flaviolin biosynthesis, where deacetylation of I-2 and chain chain shortening of naphopyrone to 4THN (1,3,6,8 tetrahydroxy-naphthalene) catalyzed by serine hydrolases and MomA catalyzes the oxidation of 4THN to flaviolin.

### 6.3.2a. Proposed radical-2 coupling catalyzed by Cu-cyp7-nr-pks6

The first proposed catalytic mechanism of oxidation of ‘**Compound f**’ (**Scheme 6.10**) is rationalized with the conventional H-atom abstraction from (phenolic) OH at C-1 atom by ‘**Compound I**’ or Fe<sup>IV</sup>-oxoporphyrin  $\pi$ -radical cation of CYP (**Scheme 6.10**). The following proposed step is a second H-atom abstraction from the second molecule of ‘**Compound f**’ by Fe<sup>IV</sup>-hydroxo radical cation species. The radical forms of two molecules of ‘**Compound f**’ are predicted to have the maximum electron density on the carbon atoms *para* to the oxygen atom. These two radical molecules of ‘**Compound f**’ undergo radical phenolic coupling followed by re-aromatization lead to the formation of a novel polyketide molecule named as ‘**Compound 17**’ (**Scheme 6.10**). This mechanism is proposed based on an example in the literature of oxidative coupling of two molecules of flaviolin substrate catalyzed by cytochrome p450 158A2\* encoded by a gene found in bacterial species named *Streptomyces coelicolor* A3(2) [Zhao, B. et al. 2005].



**Scheme 6.10:** Proposed radical-2 coupling of two molecules of ‘**Compound f**’ catalyzed by CYP in the putative biosynthesis of a novel polyketide named as ‘**Compound 17**’.

Similarly, it is proposed that the radical oxidative coupling mechanism of two molecules of ‘**Compound g**’, ‘**Compound h**’ & ‘**Compound i**’ to predicted novel polyketide molecules be named as ‘**Compound 19**’, ‘**Compound 20**’ &, ‘**Compound 21**’.

### 6.3.3 Non-reducing polyketide synthase gene cluster, *Cu-nr-pks-12*

The earlier reported gene cluster *Cu-nr-pks-12* in the *C. uncialis* genome [Bertrand, R. L. et al., 2018] is annotated in contig 32 (**Figure 6.5**). It consists of a non-reducing polyketide

synthase (backbone enzyme) and, two accessory genes encoding a GNAT-family (General control non-repressible 5 (GCN5)-related N-acetyltransferases) enzyme and a cytochrome p450 (*Cucyp10-nr-pks12*).

Based on the BLASTp search *C. uncialis nr-pks-12* gene has found to be homologous to three lichen non-reducing PKSs found in *C. grayi*, *C. metacorallifera* and *S. alpinium* (Antarctic lichen) (**Table 6.4**). Out of these three genes, only one found in Antarctic lichen (*Stereocaulon alpinium*) has been functionally characterized [Kim, W. et al. in 2021]. In this article it has been reported that a non-reducing polyketide synthase named as *atr1* is associated with the biosynthesis of a very well-known lichen natural product, atranorin [Studzinska-Sroka, E. et al., 2017; Singh, G. et al., 2021].

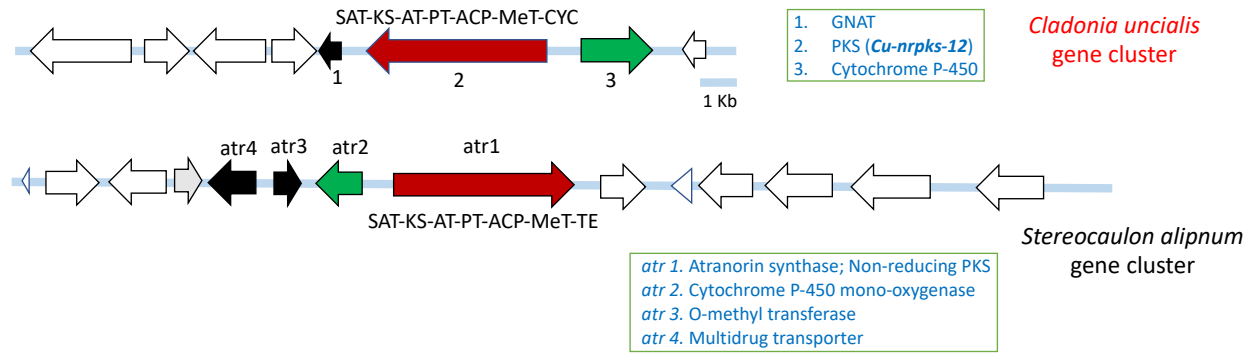
**Table 6.4:** BLAST statistics of *C. uncialis* NR-PKS (*Cu-nr-pks-12*) which has been found genetically homologous to functionally characterized NR-PKS found in a lichen species (*Stereocaulon alpinium*), which has been reported to be involved in the biosynthesis of atranorin.

Description	Max. Score	Query Cover	E value	Per. Identity	Accession no.
<b>Putative type 1 PKS [<i>Cladonia uncialis</i> subsp. <i>Uncialis</i>]</b>	<b>5188</b>	<b>100%</b>	<b>0.0</b>	<b>100%</b>	AUW31240.1
Putative non-reducing PKS [ <i>Cladonia grayi</i> ]	4137	99%	0.0	87.71%	ADM79462.1
Polyketide synthase [ <i>Cladonia metacorallifera</i> ]	3781	87%	0.0	84.08%	QIX11487.1
<i>Atr1</i> [ <i>Stereocaulon alpinium</i> ]	2122	98%	0.0	44.59%	QXF68953.1
Non-reducing polyketide synthase ascC [ <i>Phycia stellaris</i> ]	1599	95%	0.0	43.43%	KAG7007893.1

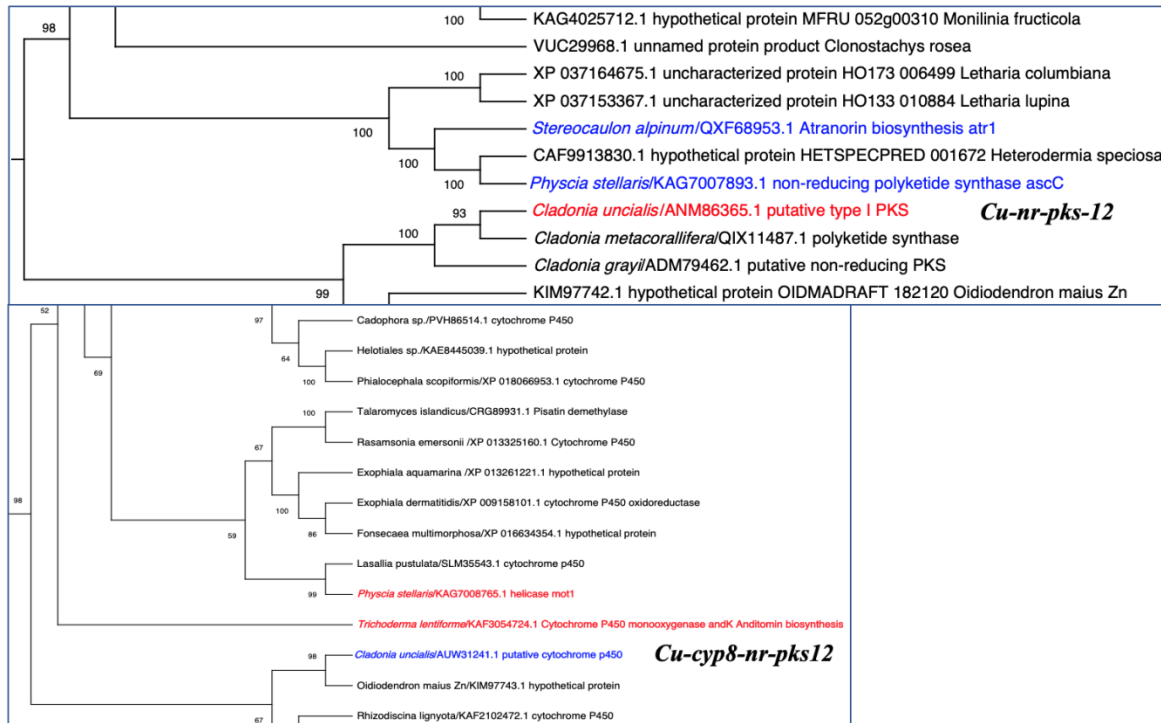
The percentage identity value is quite low as shown by BLAST statistics (**Table 6.4**). However, if we construct a phylogenetic tree as shown in **Figure 6.10**, high (98%) bootstrap support value and descendance of both proteins from one node with a very less horizontal distance between them suggest that these two proteins could be predicted to be orthologous. After

comparing both gene clusters as shown in **Figure 6.9**, *Cu-nr-pks-12* and *atr1* (*S. alpinum*) it was observed that other than PKS genes, all the accessory genes are different in both gene clusters.

Therefore, *Cu-nr-pks-12* gene cluster cannot be associated with the biosynthesis of atranorin.



**Figure 6.9:** *Stereocaulon alpinum* gene cluster involved in the biosynthesis of atranorin (accession no. QXF68953.1) & uncharacterized *C. uncialis* gene cluster, *Cu-nr-pks-6* (accession no. AUW31240.1). Genes in red color encode for non-reducing polyketide synthases, genes in green color encode for cytochrome p450 enzymes. And rest of the accessory genes in black color: GNAT in *C. uncialis*, O-methyltransferase (OMT) & Multidrug transporter *S. alpinum*. Domain abbreviations: Starter acyltransferase (SAT), ketosynthase (KS), acyltransferase (AT), acetyl carrier protein (ACP), C-methyltransferase (CMeT), claisen cyclase (CYC), thioesterase (TE). Gene abbreviations: Polyketide synthase (PKS), short-chain dehydrogenase/reductase (SDR), general control non-repressible 5 (GCN5)-related N-acetyltransferases (GNAT).





**Figure 6.10:** Truncated phylogenetic tree of *Cu-nr-pks-12* illustrates the degree of relationship between *C. uncialis* genes and a lichen *S. alpinum* NR-PKS gene (*atr1*). *Cu-cyp8-nrpk5-6* phylogenetic tree (truncated) shows a high homology b/w *C. uncialis* p450 & a non-lichen fungal p450 (*andk*) found in *T. lentiforme* (PDF version of complete phylogenetic trees are in Appendix Figure S18 & S19).

Next, the homologs of the accessory enzyme, GNAT (General control non-repressible 5 (GCN5)-related N-acetyltransferases) in *Cu-nr-pks-12* gene cluster was examined in detail. A fungal (*Lasallia pustulata*) GNAT-family acetyltransferase protein also known as Nat4 have been found to be homologous to *C. uncialis* GNAT based on the BLASTp consensus alignment (88% query cover, ~60% A.A. identity) (**Table 6.5**) 74% positive and less than 2% gap in MSA. GNAT-family enzyme is known to catalyze the transfer of an acyl moiety from acyl coenzyme A (acyl-CoA) to a diverse group of substrates. Most of the literature shows the function of GNAT is the transfer of acyl group from coenzyme A to acyl carrier protein (polyketide chain initiation) [Gu, L. et al., 2007; Salah Ud-Din, A. I. et al., 2016]. While in other cases it has been shown that GNAT catalyzes the decarboxylation of malonyl CoA during the claisen condensation step of acetyl CoA and malonyl CoA [Skiba, M. A. et al., 2020]. There are some examples in the literature that illustrate the role of evolutionarily conserved N-acetyltransferase, is the acetylation of the N-terminal residues of histones H4 and H2A in all species from yeast to mammals [Song, O. K. et al., 2003]. These articles also talk about the N-terminal acetylation of CYP proteins, which is basically a post-translation modification of a protein that eventually increases the protein stability (resistance to degradation). It also helps in the proper functioning of p450s. In another recently published article, it has been shown that how the acetylation of N-terminal residue of histone H3 attenuates the expression of cytochrome p450 proteins [Pande, P. et al., 2020].

**Table 6.5:** BLAST statistics of *C. uncialis* putative GNAT-family acetyltransferase which has been found genetically homologous to a GNAT family enzyme (*Nat4*) found in a non-lichen fungal species (*Lasallia pustulata*)

Description	Max. Score	Query Cover	E value	Per. Identity	Accession no.
<b>Putative GNAT-family acetyltransferase</b> [ <i>Cladonia uncialis</i> subsp. <i>Uncialis</i> ]	<b>1503</b>	<b>100%</b>	<b>0.0</b>	<b>100%</b>	AUW31241
GNAT family acetyltransferase <i>Nat4</i> [ <i>Lasallia pustulata</i> ]	259	88%	5e-84	59.62%	KAA6410295.1

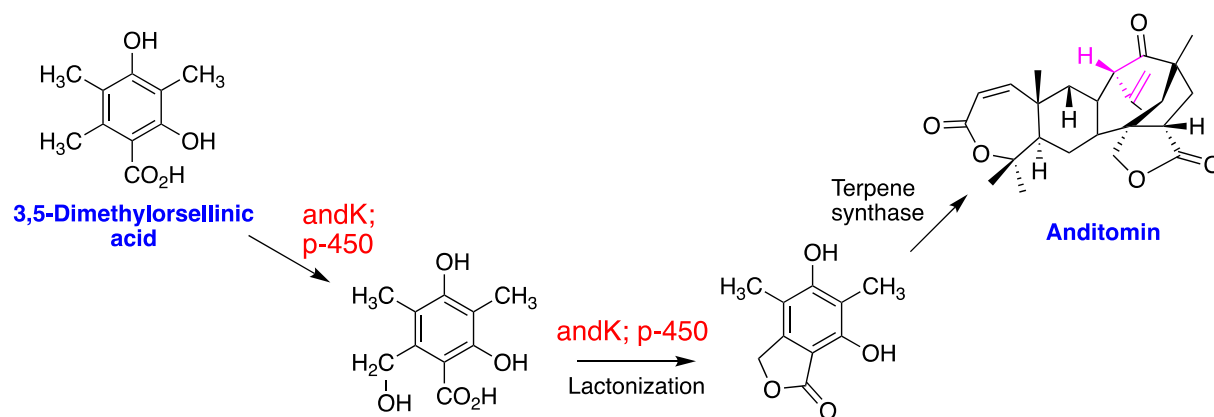
Similarly, the functional assignment of *C. uncialis* p450 (*Cu-cyp10-nr-pks12*) has been proposed based on the BLASTp (**Table 6.6**) search results and the phylogenetic reconstruction that shows *Cu-cyp10-nr-pks12* is homologous to a known fungal CYP also known as anditomin K (andK). The topology of the tree suggests high bootstrap support (98%), both share a common ancestor, and no other proteins are present between these two p450s, this laid a strong foundation to posit the functional similarity of *Cu-cyp8-nr-pks12* to anditomin K (**Figure 6.10**). Anditomin is a fungal meroterpenoid isolated from *Aspergillus varicolor* [Matsuda, Y. et al., 2014] and this article demonstrates that andK catalyzes the hydroxylation of the methyl group of 3,5-dimethylorsellinic acid ring in a terpenoid structure followed by the lactonization (**Scheme 6.11**).

Based on above predictions, we have proposed a biosynthetic pathway of a putative novel depside; named as ‘**Compound 23**’ and have been shown to be catalyzed by *Cu-nr-pks-12* gene cluster (**Scheme 6.15**). *Cu-nr-pks-12* gene shares functional similarity with *atr1*(*S. alpinium*). The first three proposed steps involve the formation of 4-O-methylbarbatic acid through two catalytic intermediates, a methylated tetraketide and 3-methyl orsellinic acid (3-MOA) (**Scheme 6.12**). Further the catalytic oxidation of 4-O-methylbarbatic to ‘**Compound 23**’ is proposed to be

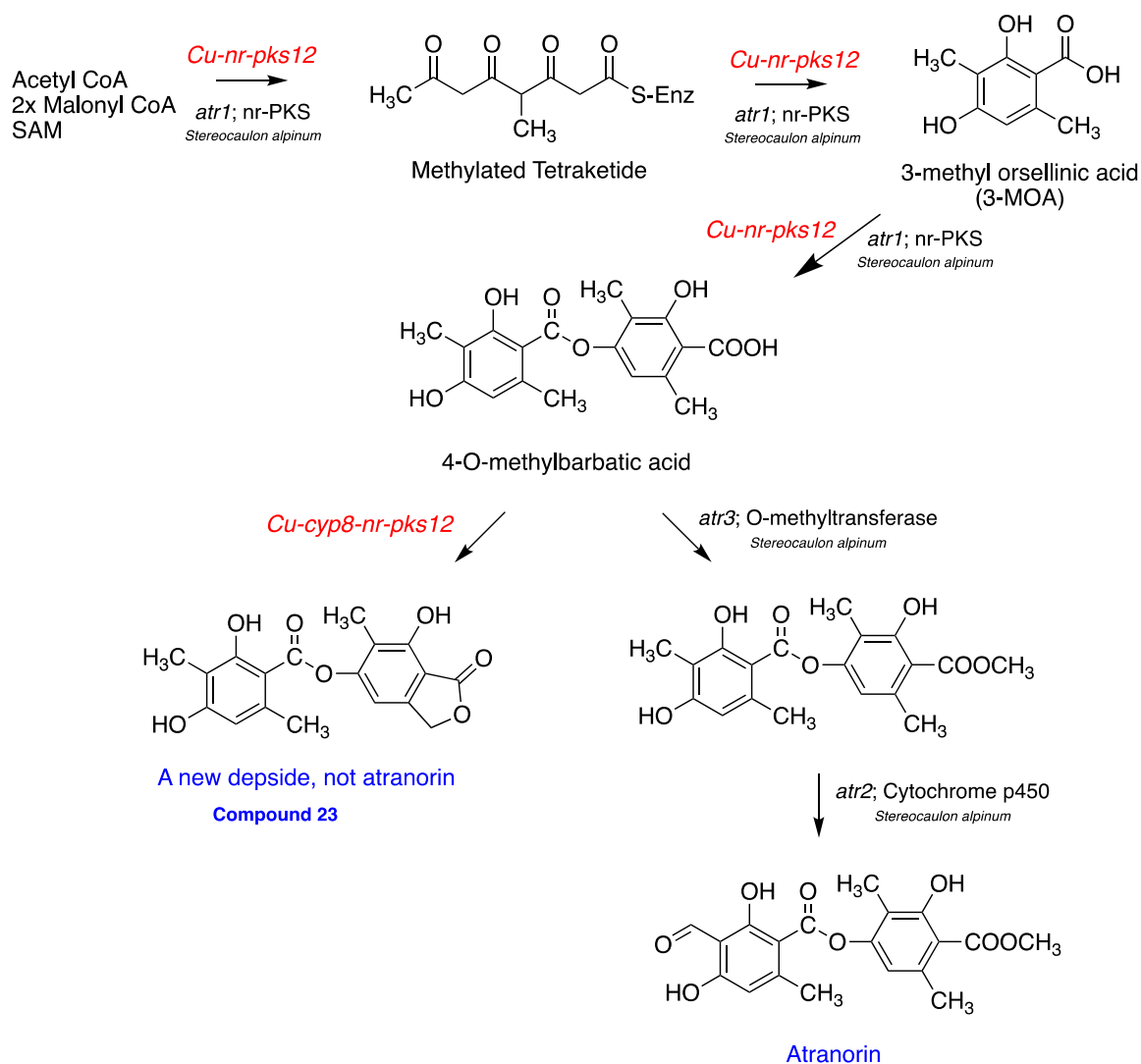
catalyzed by cytochrome p450 (*Cu-cyp8-nrpks-12*) (**Scheme 6.12**). Which is completely different from atranorin biosynthesis, where *atr2* (O-methyl transferase) catalyzes the methylation of carboxylic acid functional group followed by two oxidation steps catalyzed by *atr3* (CYP) leads to the formation of atranorin (**Scheme 6.12**).

**Table 6.6:** BLAST statistics of *C. uncialis* putative CYP (*Cu-cyp8-nrpks-12*) which has been found genetically homologous to a functionally characterized non-lichen fungal (*T. lentiforme*) CYP (*andk*).

Description	Max. Score	Query Cover	E value	Per. Identity	Accession no.
<b>Putative cytochrome p450</b> [ <i>Cladonia uncialis</i> subsp. <i>Uncialis</i> ]	<b>1503</b>	<b>100%</b>	<b>0.0</b>	<b>100%</b>	AUW31241
Cytochrome p450 monooxygenase <i>andk</i> [ <i>Trichoderma lentiforme</i> ]	488	71%	5e-162	47.46%	KAF3054724.1
Helicase mot 1 [ <i>Physcia stellaris</i> ]	431	69%	7e-140	44.01%	KAG7008765.1



**Scheme 6.11:** Hydroxylation & lactonization of 3,5-dimethylorsellinic acid catalyzed by p450 (*andk*), *anditomin K* during the biosynthesis of anditomin.

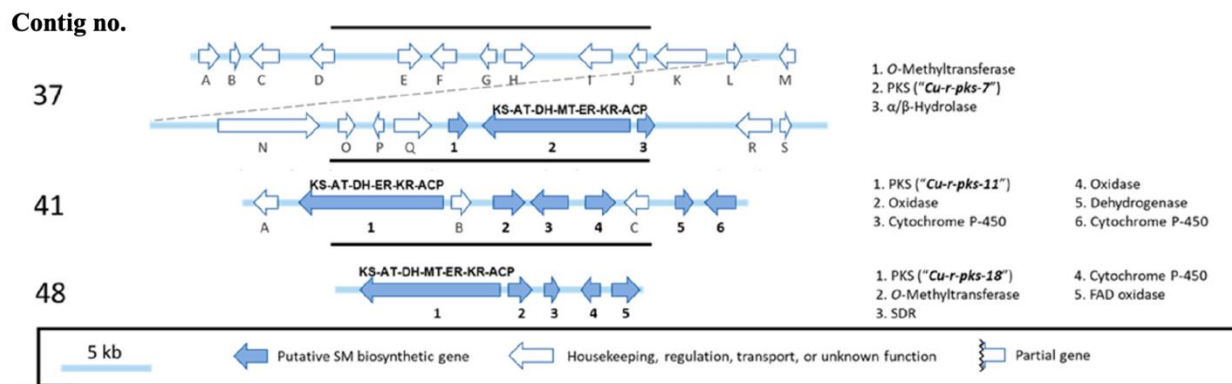


**Scheme 6.12:** *S. alpinum* gene cluster responsible for the biosynthesis of atranorin and juxtaposed, biosynthetic pathway of a novel lichen depside molecule named as ‘**Compound 23**’ that is proposed to be encoded in *C. uncialis*.

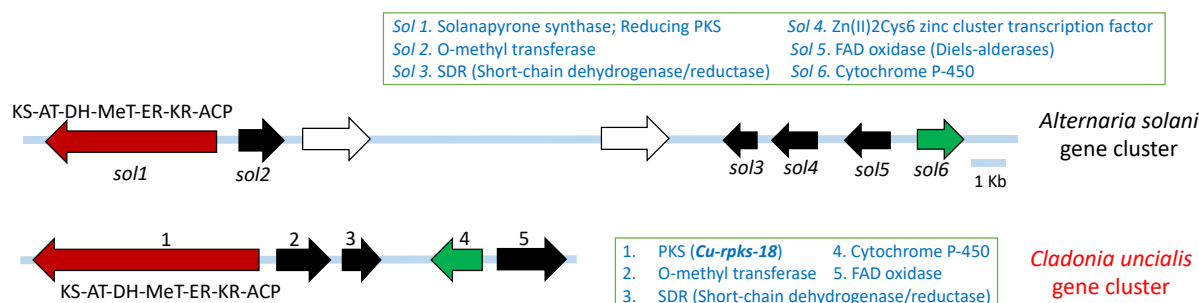
## 6.4. Reducing polyketide synthase gene clusters

### 6.4.1. *C. uncialis* reducing polyketide synthase gene clusters of *Cu-r-pks-18*

The earlier reported gene cluster [Bertrand, R. L. et al., 2018] that has been assigned as contig no. 48 (**Figure 6.11**), which belongs to a class of type I iterative reducing polyketide synthases (R-PKS). This gene cluster consists of one reducing PKS gene and four accessory genes such as, O-methyltransferase, cytochrome p450, FAD oxidase, short-chain dehydrogenase/reductase (**Figure 6.12**).



**Figure 6.11:** Type I reducing polyketide synthase (PKS; backbone genes) gene clusters in the *C. uncialis* mycobiont genome. Abbreviations: KS, ketosynthase; AT, acyltransferase; DH, dehydratase; MT, C-methyltransferase; ER, enoylreductase; KR, ketoreductase; ACP, acetyl carrier protein; R, reductase, and the representation of cytochrome p450s (tailoring genes); part of these lichen *cu-rpks* gene clusters. (Reprinted with the permission from *Journal of Natural Products* 2018 81 (4), 723-731. Copyright 2018 American Chemical Society, see Figure S26 in the Appendix for copyright clearance).



**Figure 22:** *Alternaria solani* gene cluster involved in the biosynthesis of solanapyrone (accession no. BAJ09789.1) & uncharacterized *C. uncialis* gene cluster, *Cu-r-pks-18* (accession no. AUW31423.1). Genes in red color encode for reducing polyketide synthases, genes in green color encode for cytochrome p450 enzymes. And rest of the accessory genes are in black color: O-methyltransferase (OMT), short-chain dehydrogenase/reductase (SDR), Zn (II)<sub>2</sub>Cys<sub>6</sub> zinc cluster transcription factor, flavin adenine dinucleotide (FAD) oxidase. Domain abbreviations: Ketosynthase (KS), acyltransferase (AT), dehydratase (DH), C-methyltransferase (CMeT), enoyl reductase (ER), keto-reductase (KR), acetyl carrier protein (ACP). Gene abbreviations: Polyketide synthase (PKS).

After analysing its evolutionary history and protein database search along with multiple sequence alignment results, it is predicted that this gene cluster could be linked to the biosynthesis of solanapyrone (**Scheme 6.13**). The entire gene cluster is very similar to the gene cluster that has been shown to produce solanapyrone found in a fungus *Alternaria solani* (sol1-sol6) (**Figure 6.12**). The only difference between gene clusters appears to be the absence in *C. uncialis* of

Zn<sub>2</sub>Cys<sub>6</sub> zinc cluster transcription factor. This gene is basically a GAL4-type transcriptional regulator (sol4) required for the expression of galactose-induced genes.

Solanapyrones were first isolated as phytotoxins in 1983 from a pathogenic fungus named *Alternaria solani* [Ichihara, A. et al., 1983]. The structure was deduced by spectroscopic analysis [Ichihara, A. et al., 1983] and chemical degradation has been confirmed by the total synthesis [Ichihara, A. et al., 1987].

Solanapyrones are biosynthesized from an acetate-derived octaketide with two C1 units from methionine. The isolation of the solanapyrones in optically active forms **solanapyrone B** and, **solanapyrone E (Scheme 6.13)**. Kasahara *et al.* showed that the decalin system is formed through a [4+2] cycloaddition from the achiral linear triene precursor prosolanapyrone III and this is an enzyme-catalyzed Diels–Alder reaction [Kasahara, K. et al., 2010]. Sol5 (FAD oxidase/first proven natural Diels-alderase) was the first example of an enzyme catalyzing a biological [4+2] cycloaddition reaction, this single enzyme catalyzes both the oxidation and stereoselective cycloaddition steps leads to formation the forms of two optically active forms of solanapyrone. Similarly, all the four accessory genes have been found to be closely related to the genes found in the same fungal species (*Colletotrichum gloeosporioides*) as we found in case of *Cu-r-pks-18* and all these genes are the part of one gene cluster found in *C. gloeosporioides* (solanapyrone biosynthetic gene cluster).

*Cu-rpks-18* has high homology (BLASTp, **Table 6.7**) to a fungal reducing PKS (*Colletotrichum gloeosporioides*) prosolanapyrone synthase (99% query cover with approx. 68% A.A. identity). The phylogenetic reconstruction (**Figure 6.13**) shows that the two proteins are present very close to each other (low evolutionary distance), although both are the part of two separate clades but both clades are related by 90 bootstrap value.

**Table 6.7:** BLAST statistics of *C. uncialis* putative reducing polyketide synthase (*Cu-r-pks18*) which is genetically similar to a non-lichen fungal (*C. gloeosporioides*) R-PKS (*prosolanapyrone synthase/sol1*).

Description	Max. Score	Query Cover	E value	Per. Identity	Accession no.
<b>Putative type I PKS [<i>Cladonia uncialis</i> subsp. <i>Uncialis</i>]</b>	<b>5365</b>	<b>100%</b>	<b>0.0</b>	<b>100%</b>	AUW31423.1
<b>Prosolanapyrone synthase sol1 [<i>Colletotrichum gloeosporioides</i>]</b>	<b>3592</b>	<b>99%</b>	<b>0.0</b>	<b>67.79%</b>	<b>KAF3806818.1</b>

*Cu-r-pks18-O-methyltransferase* homologous to the *sol2* (O-methyltransferase) found in *C. gloeosporioides* (99% query cover with 82% A.A. identity, **Table 6.8**), topology of the phylogenetic tree shows that both genes are in separate clades but present quite nearby (65 bootstrap support, **Figure 6.13**).

**Table 6.8:** BLAST statistics of *C. uncialis* putative O-methyltransferase (*Cu-r-pks18-OMT*) which is genetically similar to a non-lichen fungal (*C. gloeosporioides*) O-methyltransferase (*sol2*).

Description	Max. Score	Query Cover	E value	Per. Identity	Accession no.
<b>Putative O-methyltransferase [<i>Cladonia uncialis</i> subsp. <i>Uncialis</i>]</b>	<b>895</b>	<b>100%</b>	<b>0.0</b>	<b>100%</b>	AUW31424.1
<b>O-methyl transferase sol2 [<i>Colletotrichum gloeosporioides</i>]</b>	<b>742</b>	<b>99%</b>	<b>0.0</b>	<b>82.05%</b>	<b>XP_045265976.1/ KAF3806817.1</b>

*Cu-cyp9-r-pks18* homologous to *sol6* (cytochrome p450 monooxygenase, *C. gloeosporioides*) with 99% query cover with approx. 70% A.A. identity (**Table 6.9**). Phylogenetic tree shows that two proteins are orthologs with 100% bootstrap support (**Figure 6.13**).

**Table 6.9:** BLAST statistics of *C. uncialis* putative cytochrome p450 (*Cu-cyp9-rpks18*) which is genetically similar to a non-lichen fungal (*C. gloeosporioides*) CYP 4F5 (*sol6*).

Description	Max. Score	Query Cover	E value	Per. Identity	Accession no.
<b>Putative cytochrome p450</b> [ <i>Cladonia uncialis</i> subsp. <i>Uncialis</i> ]	<b>1213</b>	<b>100%</b>	<b>0.0</b>	<b>100%</b>	AUW31426.1
Cytochrome p450 4F5 sol 6 [ <i>Colletotrichum gloeosporioides</i> ]	887	99%	0.0	69.83%	KAF3806815.1

Both genes, *Cu-r-pks18-FAD oxidase* and *Cu-r-pks18-SDR* are found to be homologous to *sol5* and *sol3* with approximately 98% query cover with 70% A.A. identity (BLASTp, **Tables 6.10 & 6.11**). Although the topology of their phylogenetic trees shows that these genes are related to *sol5* and *sol3* of *C. gloeosporioides* with the low bootstrap value, but still have been shown to evolve in the recent time frame (low horizontal distances between the tips (proteins) through the nodes) (**Figure 6.13**).

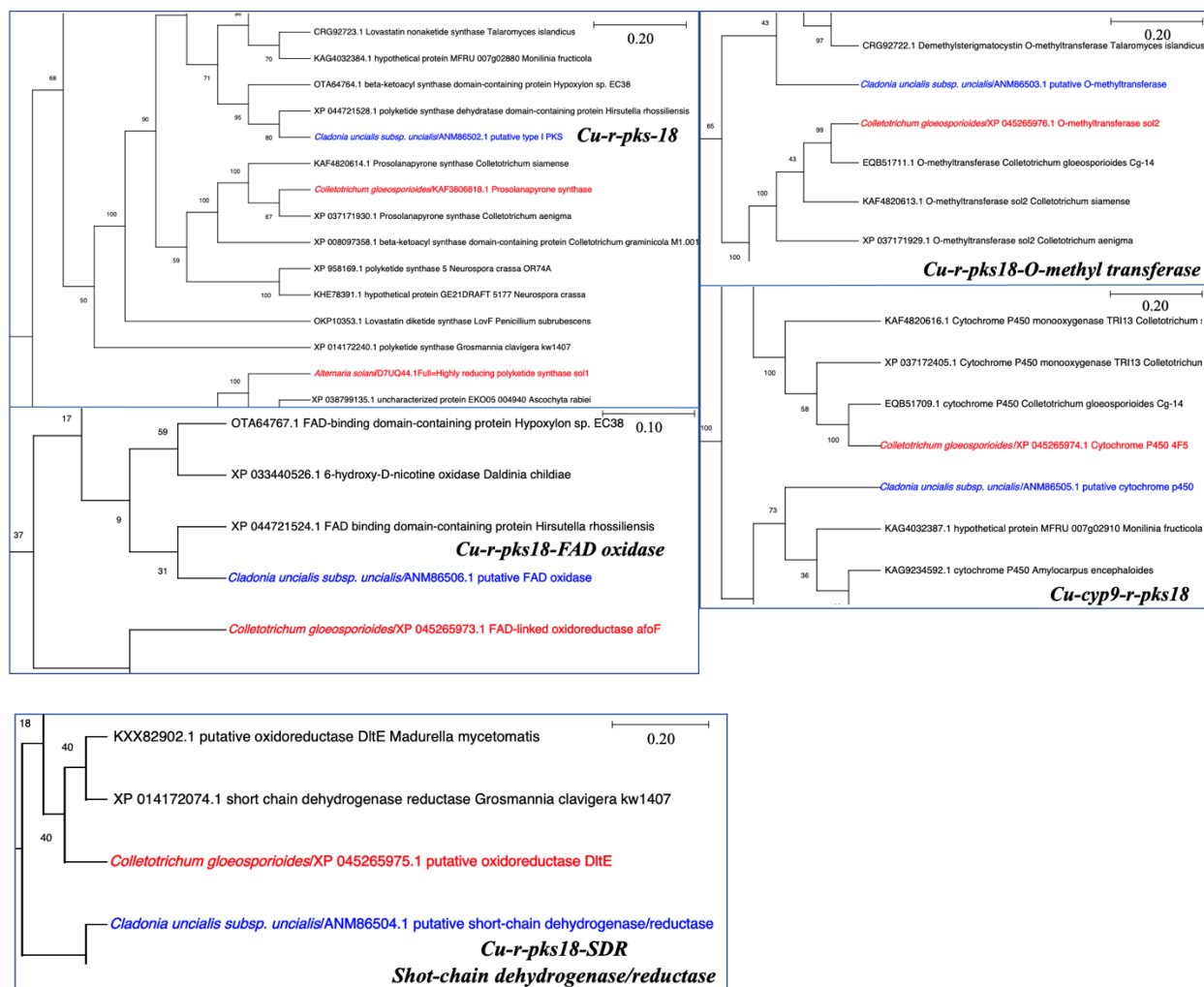
**Table 6.10:** BLAST statistics of *C. uncialis* putative FAD oxidase (*Cu-r-pks18-FAD oxidase*) which is genetically similar to a non-lichen fungal (*C. gloeosporioides*) FAD-linked oxidoreductase *afoF* (*sol5*).

Description	Max. Score	Query Cover	E value	Per. Identity	Accession no.
<b>Putative FAD oxidase</b> [ <i>Cladonia uncialis</i> subsp. <i>Uncialis</i> ]	<b>1077</b>	<b>100%</b>	<b>0.0</b>	<b>100%</b>	AUW31427.1
FAD-linked oxidoreductase <i>afoF</i> sol 5 [ <i>Colletotrichum gloeosporioides</i> ]	717	98%	0.0	69.34%	XP_045265973.1/ KAF3806814.1



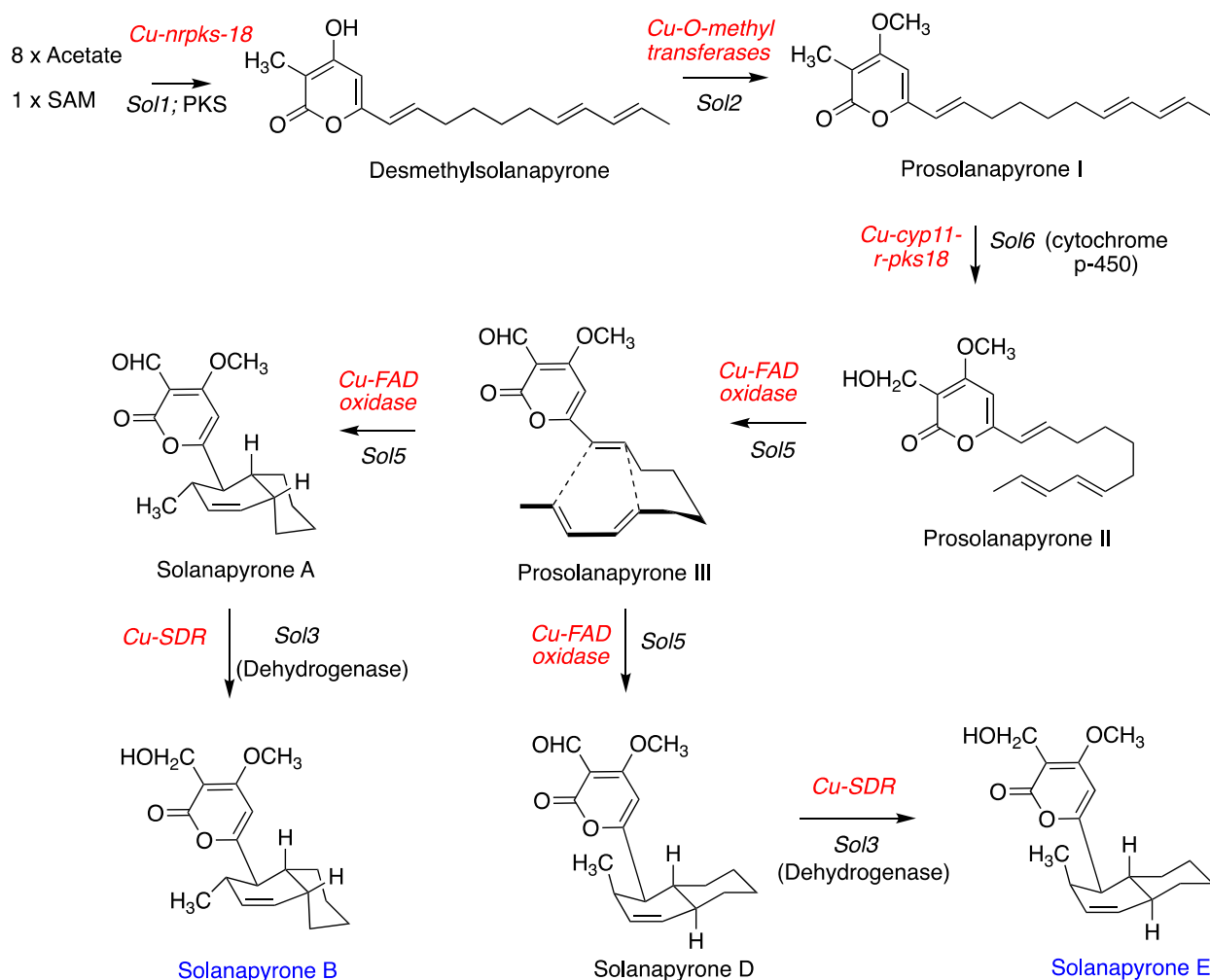
**Table 6.11:** BLAST statistics of *C. uncialis* putative short-chain dehydrogenase/reductase (*Cu-r-pks18-SDR*) which is genetically similar to a non-lichen fungal (*C. gloeosporioides*) FAD-linked oxidoreductase DtE (*sol3*).

Description	Max. Score	Query Cover	E value	Per. Identity	Accession no.
<b>Putative short-chain dehydrogenase/reductase</b> [ <i>Cladonia uncialis</i> subsp. <i>Uncialis</i> ]	<b>567</b>	<b>100%</b>	<b>0.0</b>	<b>100%</b>	AUW31425.1
<b>Putative oxidoreductase DtE sol3</b> [ <i>Colletotrichum gloeosporioides</i> ]	<b>407</b>	<b>98%</b>	<b>0.0</b>	<b>69.49%</b>	XP_045265975.1/ KAF3806816.1



**Figure 6.13:** Truncated phylogenetic trees of *Cu-r-pks-18*, *Cu-r-pks18-O methyl transferase*, *Cu-r-pks18-FAD oxidase*, *Cu-cyp9-r-pks18*, *Cu-r-pks18-SDR* illustrate the degree of relationship between *C. uncialis* genes and non-lichen fungal genes (PDF version of complete phylogenetic trees are in Appendix Figure S20-S24).

Taking all the genomic database search and the evolutionary studies into account we propose the biosynthesis of solanapyrone (**Scheme 34**) catalyzed by lichen reducing PKS gene cluster (*Cu-r-pks-18*).



**Scheme 6.13:** Experimentally supported pathway for solanapyrone biosynthesis in *A. solani* and in parallel a biosynthetic pathway for solanapyrone biosynthesis that is proposed to be encoded in *C. uncialis*.

## 6.5. Summary

By applying the ‘homology mapping’ approach described in this chapter, proposals of function for seven gene clusters were generated. These proposals include what appear to be complete

biosynthetic pathways for **Compounds 16, 23** and solanapyrone B & E. **Compounds 16 & 23** are completely novel molecules that have not been extracted naturally so far and their proposed biosynthetic pathways are entirely based on the functionality prediction model of individual genes in their respective gene clusters. The functional prediction of individual genes is based on the BLAST calculated genetic similarity of the proteins encoded by the individual genes and their evolutionary studies. Evolutionary relationships between different genes have been established by reconstructing a phylogenetic tree which predicts either a homolog or an ortholog of a particular gene (protein sequences were used to reconstruct all the trees). A similar concept was applied to predict the biosynthetic pathway of novel '**Compounds 17-22**' & '**Compounds 13-15**'. Any one of the **Compounds 17, 18, 19, 20, 21** or **22** could be biosynthesized by *Cu-nr-pks-6* gene cluster depending upon the substrate specificity of all the enzymes encoded by different genes in this gene cluster. A similar rationale could also be applied in case of biosynthesis of **Compounds 13, 14 & 15** catalyzed by *Cu-terp-6* gene cluster. The order of different catalytic steps in a biosynthetic pathway of proposed novels **Compounds 17-22** was rationalized by finding the closest NR-PKS gene cluster which was associated with a known SM (SM). Together with advances in bioinformatics and metabolomics, these approaches have started to begin the facilitation of high-throughput discovery of natural products pathways from genomic sequence information alone independent of lichen fungal cultivation. As methods evolve and computational prediction pipelines become more powerful, and these tools could not only be used to identify the gene clusters but also to predict the function of each individual gene in the respective gene clusters. But to confirm these predictions, it will be necessary to express these genes and search for *de novo* production of a metabolite in a heterologous host. Proposed catalytic mechanisms of a putative lichen CYP (*Cu-cyp7-nrpks6*) converting '**Compound f-I**' to '**Compound 17-22**' give the basic understanding of the catalytic mechanism of p450 in the context of electronic changes leading to oxygen activation and the oxidation of a

substrate. In particular, the chemistry of lichen CYPs have been described in detail. Building upon the already known different catalytic mechanisms of these novel lichen cytochrome p450s we have been proposed that these enzymes catalyze C-C and C-O oxidative coupling reactions. In this chapter, we also propose the two primary p450 catalytic mechanisms, either a radical mechanism or a possible cationic mechanism. These proposals are still purely theoretical and will need further confirmation by a combination of both computational and experimental results.

## References

- Abdel-Hameed, M., Bertrand R. L., Piercey-Normore, M. D., Sorensen, J. L. Putative identification of the usnic acid biosynthetic gene cluster by de novo whole-genome sequencing of a lichen-forming fungus. *Fungal Biol.* **2016**, 120(3), 306-16.
- Agger, S., Lopez-Gallego, F., & Schmidt-Dannert, C. Diversity of sesquiterpene synthases in the basidiomycete *Coprinus cinereus*. *Molecular microbiology*, **2009**, 72 (5), 1181–1195.
- Aigle, B., Lautru, S., Spitteller, D., Dickschat, J. S., Challis, G. L., Leblond, P., Pernodet, J. L. J. *Ind. Microbiol. Biotechnol.* 2014, 41, 251–263.
- Bertrand, R. L., Abdel-Hameed, M. and Sorensen, J. L. Lichen Biosynthetic Gene Clusters Part I: Homology Mapping Suggests a Functional Diversity. *Journal of Natural Products* **2018**, 81 (4), 723-731.
- Bertrand, R. L., Abdel-Hameed, M., and Sorensen, J. L. Lichen Biosynthetic Gene Clusters Part II: Homology Mapping Suggests a Functional Diversity. *Journal of Natural Products* **2018**, 81 (4), 732-748.
- Bertrand, R. L. Characterization of Secondary Metabolome of a Lichenizing Fungus. University of Manitoba, Winnipeg, Canada, **2019**, Retrieved from <https://mspace.lib.umanitoba.ca/handle/1993/33943>
- Blin, K., Medema, M. H., Kazempour, D., Fischbach, M. A., Breitling, R., Takano, E., Weber, T. AntiSMASH 2.0 – a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **2013**, 41, 204-212.
- Crawford, J. M., Townsend, C. A. New insights into the formation of fungal aromatic polyketides. *Nat Rev Microbiol.* **2010**; 8, 879–889.

- Chang, P-K., Carry, J. W., Lebar, M. D. Biosynthesis of conidial and sclerotial pigments in *Aspergillus* species. *Applied Microbiology and Biotechnology*, **2020**, 104, 2277-2286.
- Deligeorgopoulou, A. and Allemann, R. K. *Biochemistry* **2003** 42 (25), 7741-7747.
- Delserone, L. M., McCluskey, K. D., Matthews, E., VanEtten, H. D. Pisatin demethylation by fungal pathogens and nonpathogens of pea: association with pisatin tolerance and virulence. *Physiol. Mol. Plant Pathol.* **1999**, 55, 317–326.
- George, H. L., VanEtten, H. D. Characterization of pisatin-inducible cytochromeP450s in fungal pathogens of pea that detoxify the pea phytoalexin pisatin. *Fungal Genet. Biol.* **2001**, 33, 37–48
- Griffiths, S. A., Cox, R. J., Overdijk E. J. R., Mesarich, C. H., Saccomanno, B., Lazarus, C. M., et al. Assignment of a dubious gene cluster to melanin biosynthesis in the tomato fungal pathogen *Cladosporium fulvum*. *PLoS ONE*. **2018**,13(12): e0209600
- Gu, L., Geders, T. W., Wang, B., Gerwick, W. H., Håkansson, K., Smith, J. L., & Sherman, D. H. GNAT-like strategy for polyketide chain initiation. *Science (New York, N.Y.)*, **2007**, 318 (5852), 970–974.
- Guo, C., Wang, C. C. C. *Front. Microbiol.* **2014**, 5, 717.
- Ha, N. T., & Lee, C. H. Roles of Farnesyl-Diphosphate Farnesyltransferase 1 in Tumour and Tumour Microenvironments. *Cells*, **2020**, 9 (11), 2352.
- Ichihara, A., Miki, M., Tazaki, H., Sakamura, S. Synthesis of (±)-solanapyrone A, *Tetrahedron Letters*, **1987**, Volume 28, Issue 11, Pages 1175-1178.
- Ichihara, A., Tazaki, H., Sakamura, S. Solanapyrones A, B and C, phytotoxic metabolites from the fungus *Alternaria solani*, *Tetrahedron Letters*, **1983**, Volume 24, Issue 48, Pages 5373-5376.

- Fujii, I., Watanabe, A., Sankawa, U., Ebizuka, Y. Identification of Claisen cyclase domain in fungal polyketide synthase WA, a naphthopyrone synthase of *Aspergillus nidulans*, *Chemistry & Biology*, **2001**, Volume 8, Issue 2, Pages 189-197.
- Kasahara, K., Miyamoto, T., Fujimoto, T., Oguri, H., Tokiwano, T., Oikawa, H., Ebizuka, Y., and Fuji, I. Solanapyrone Synthase, a Possible Diels–Alderase and Iterative Type I Polyketide Synthase Encoded in a Biosynthetic Gene Cluster from *Alternaria solani*. *ChemBioChem* **2010**, 9, 1154.
- Kim, W., Liu, R., Woo, S., Kang, K. B., Park, H., Yu, Y. H., Ha, H. H., Oh, S. Y., Yang, J. H., Kim, H., Yun, S. H., & Hur, J. S. Linking a Gene Cluster to Atranorin, a Major Cortical Substance of Lichens, through Genetic Dereplication and Heterologous Expression. *mBio*, **2021**, 12 (3), e0111121.
- Lebe, K. E. and Cox, R. J. *Chem. Sci.*, **2019**, 10, 1227-1231.
- Li, T., Francl, J. M., Boehme, S., & Chiang, J. Y. Regulation of cholesterol and bile acid homeostasis by the cholesterol 7 $\alpha$ -hydroxylase/steroid response element-binding protein 2/microRNA-33a axis in mice. *Hepatology (Baltimore, Md.)*, **2013**, 58 (3), 1111–1121.
- Matsuda, Y., Wakimoto, T., Mori, T., Awakawa, T., and Abe, I. Complete Biosynthetic Pathway of Anditomin: Nature’s Sophisticated Synthetic Route to a Complex Fungal Meroterpenoid. *Journal of the American Chemical Society*, **2014**, 136 (43), 15326-15336
- Medema, M. H., Kottmann, R., Yilmaz, P., Cummings, M, Biggins, J. B., Blin, K., de Bruijn, I., Chooi, Y. H., Claesen, J., Coates, R. C., Cruz-Morales, P., Duddela, S., Düsterhus, S., Edwards, D. J., Fewer, D. P., Garg, N., Geiger C, Gomez-Escribano JP, Greule A, Hadjithomas M, Haines AS, Helfrich EJ, Hillwig ML, Ishida K, Jones AC, Jones CS, Jungmann K, Kegler C, Kim HU, Kötter P, Krug D, Masschelein J, Melnik AV, Mantovani

SM, Monroe EA, Moore M, Moss N, Nützmann HW, Pan G, Pati A, Petras D, Reen FJ, Rosconi F, Rui Z, Tian Z, Tobias NJ, Tsunematsu Y, Wiemann P, Wyckoff E, Yan X, Yim G, Yu F, Xie Y, Aigle B, Apel AK, Balibar CJ, Balskus EP, Barona-Gómez F, Bechthold A, Bode HB, Borriss R, Brady SF, Brakhage AA, Caffrey P, Cheng YQ, Clardy J, Cox RJ, De Mot R, Donadio S, Donia MS, van der Donk WA, Dorrestein PC, Doyle S, Driessen AJ, Ehling-Schulz M, Entian KD, Fischbach MA, Gerwick L, Gerwick WH, Gross H, Gust B, Hertweck C, Höfte M, Jensen SE, Ju J, Katz L, Kaysser L, Klassen JL, Keller NP, Kormanec J, Kuipers OP, Kuzuyama T, Kyrpides NC, Kwon HJ, Lautru S, Lavigne R, Lee CY, Linquan B, Liu X, Liu W, Luzhetskyy A, Mahmud T, Mast Y, Méndez C, Metsä-Ketelä, M., Micklefield, J., Mitchell, D. A., Moore, B. S., Moreira, L. M., Müller, R., Neilan, B. A., Nett, M., Nielsen, J., O'Gara, F., Oikawa, H., Osbourn, A., Osburne, M. S., Ostash, B., Payne, S. M., Pernodet, J. L., Petricek, M., Piel, J., Ploux, O., Raaijmakers, J. M., Salas, J. A., Schmitt, E. K., Scott, B., Seipke, R. F., Shen, B., Sherman, D. H., Sivonen, K., Smanski, M. J., Sosio, M., Stegmann, E., Süßmuth, R. D., Tahlan, K., Thomas, C. M., Tang, Y., Truman, A. W., Viaud, M., Walton, J. D., Walsh, C. T., Weber, T., Van Wezel, G. P., Wilkinson, B., Willey, J. M., Wohlleben, W., Wright, G. D., Ziemert, N., Zhang, C., Zotchev, S. B., Breitling, R., Takano, E., Glöckner, F. O. Minimum information about a biosynthetic gene cluster. *Nat. Chem. Biol.* **2015**, 11, 625-631

Milani, N. A., Lawrence, D. P., Arnold, A. E., & VanEtten, H. D. Origin of pisatin demethylase (PDA) in the genus *Fusarium*. *Fungal genetics and biology: FG & B*, **2012**, 49 (11), 933–942.

Nielsen, J. C., Grijseels, S., Prigent, S., Ji, B., Dainat, J., Nielsen, K. F., Frisvad, J. C., Workman, M., Nielsen, J. *Nat. Microbiol.* **2017**, 2, 17044.



- Pande, P., Zhong, X. B., & Ku, W. W. Histone Methyltransferase G9a Regulates Expression of Nuclear Receptors and Cytochrome P450 Enzymes in HepaRG Cells at Basal Level and in Fatty Acid Induced Steatosis. *Drug metabolism and disposition: the biological fate of chemicals*, **2020**, 48 (12), 1321–1329.
- Quin, M. B., Flynn, C. M. and Schmidt-Dannert, C. *Nat. Prod. Rep.*, 2014, 31, 1449-1473.
- Salah Ud-Din, A. I., Tikhomirova, A., & Roujeinikova, A. Structure and Functional Diversity of GCN5-Related N-Acetyltransferases (GNAT). *International journal of molecular sciences*, **2016**, 17 (7), 1018.
- Schmidt-Dannert, C. Biosynthesis of terpenoid natural products in fungi. *Advances in biochemical engineering/biotechnology* **2015**, 148, 19–61.
- Schorn, M. A.; Alanjary, M. M.; Aguinaldo, K.; Korobeynikov, A.; Podell, S.; Patin, N.; Lincecum, T.; Jensen, P. R.; Ziemert, N.; Moore, B. S. *Microbiology* **2016**, 162, 2075–2086.
- Singh, G.; Armaleo, D.; Dal Grande, F.; Schmitt, I. Depside and Depsidone Synthesis in Lichenized Fungi Comes into Focus through a Genome-Wide Comparison of the Olivetoric Acid and Physodic Acid Chemotypes of *Pseudevernia furfuracea*. *Biomolecules* **2021**, 11, 1445.
- Skiba, M. A., Tran, C. L., Dan, Q., Sikkema, A. P., Klaver, Z., Gerwick, W. H., Sherman, D. H., & Smith, J. L. (2020). Repurposing the GNAT Fold in the Initiation of Polyketide Biosynthesis. *Structure London, England* **1993**, 28(1), 63–74. e4.
- Song, O. K., Wang, X., Waterborg, J. H., & Sternglanz, R. An N-alpha-acetyltransferase responsible for acetylation of the N-terminal residues of histones H4 and H2A. *The Journal of biological chemistry*, **2003**, 278 (40), 38109–38112.

Studzinska-Sroka, E.; Galanty, A.; and Bylka W. Atranorin – An interesting Lichen Grayanic acid.

*Mini-Reviews in Medicinal Chemistry* **2017**, 17, 1633.

Thompson, J. E.; Fahnestock, S.; Farralli, L.; Liao, D-I.; Valenti, B.; and Jordan, D. B. The

Second Naphthol Reductase of Fungal Melanin Biosynthesis in *Magnaporthe grisea*.

*Journal of Biol. Chem.* **2000**, 275 (45), 34867-34872.

Tsai, H. F., Fujii, I., Watanabe, A., Wheeler, M. H., Chang, Y. C., Yasuoka, Y., Ebizuka, Y., &

Kwon-Chung, K. J. Pentaketide melanin biosynthesis in *Aspergillus fumigatus* requires

chain-length shortening of a heptaketide precursor. *The Journal of biological*

*chemistry*, **2001**, 276 (31), 29292–29298.

Watanabe, A., Fujii, I., Tsai, H.-F., Chang, Y. C., Kwon-Chung, K. J., and Ebizuka, Y. *FEMS*

*Microbiol. Lett.* **2000**, 192, 39–44.

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., Lee, S. Y., Fischbach, M.

A., Müller, R., Wohlleben, W., Breitling, R., Takano, E., Medema, M. H. AntiSMASH 3.0

– a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic*

*Acids Res.* **2015**, 43, W237-W243.

Zhao, B., Guengerich, F. P., Bellamine, A., Lamb, D. C., Izumikawa, M., Lei, L., Podust, L. M.,

Sundaramoorthy, M., Kalaitzis, J. A., Reddy, L. M., Kelly, S. L., Moore, B. S., Stec, D.,

Voehler, M., Falck, J. R., Shimada, T., and Waterman, M. R. *The Journal of Biological*

*Chemistry.* **2005**, Vol. 280, No. 12, 11599-11607.

Zhu, P., Li, Q., Zhang, C., Na, Y., Xu, L. Bcpks12 gene inactivation substantiates biological

functions of sclerotium melanization in *Botrytis cinerea*. *Physiological and Molecular*

*Plant Pathology*, **2017**, Vol. 98, 80-84.

# Chapter 7

## Conclusions and Future Prospects

There has been a lot of recent interest in the biosynthesis of lichen natural products. For instance, within the last few months two papers demonstrated the functional heterologous expression of a lichen PKS in *Saccharomyces cerevisiae* and a complete lichen gene cluster in the fungal expression system. The functional heterologous expression of lichen PKS associated with lecanoric acid biosynthesis was successfully observed in *Saccharomyces cerevisiae*. In addition, the whole lichen gene cluster that has been reported to be functionally expressed within a fungal host is associated with atranorin biosynthesis.

This study is the first experimental demonstration of functional heterologous expression lichen post-PKS genes that encode for post-PKS/accessory/tailoring enzymes in bacteria. These tailoring enzymes are cytochrome p450 (CYP) named MPAO & its redox partner (RP), cytochrome p450 oxidoreductase (CPR) named CPR357. We have also established a protocol to purify these lichen proteins. This research is the first conclusive experimental demonstration linking the putative lichen p450 gene encodes for an enzyme, MPAO, to its function. MPAO has been proposed to catalyze the biosynthesis of UA from MPA. Hence, this research is the first experimental demonstration of biosynthesis of UA in *E. coli* catalyzed by MPAO.

Therefore, this is the first experimental evidence of both *in vivo* (a bacterial cell) and *in vitro* reproducible production of UA from MPA (substrate) catalyzed by MPAO & its redox partner CPR357 (see **Chapter 5**).

This study is a classic example of resolving untenable chemical synthesis of structurally complex lichen SMs via a semi-synthetic approach. This approach includes 1) The heterologous

expression of lichen accessory genes encoding accessory enzymes instead of the entire gene cluster within the fast-growing host. 2) Chemical synthesis of simple polyketide molecules in the lab. One of the examples is the chemical synthesis of MPA which is rather a simple polyketide molecule and a catalytic intermediate during UA biosynthesis. However, in nature, these biosynthetic intermediates are assembled by polyketide synthesis. 3) Either use the heterologously expressed accessory genes in bacteria or the pure accessory enzymes to catalyze the biosynthesis of complex polyketides/lichen SMs from simple polyketides molecules. **In the future**, the same approach could also be followed to biosynthesize the analogs of UA from MPA analogs and, SMs other than polyketides. The chemical synthesis of MPA analogs is discussed in **Chapter 4**.

In addition to MPA analogs, we also synthesized (see **Chapter 4**) many other (trivially modified) polyphenols (acyl phloroglucinols). Out of which four compounds (Di-acyl phloroglucinols) showed promising antibiotic properties, and one compound interacted synergistically with doxycycline against MRSA clinical isolates. An article also shows that various diacyl phloroglucinol analogs display antifungal activities. Similarly, more libraries of such chemical analogs of lichen biosynthetic intermediates or the analogs of various polyphenols could be generated **in the future**. Furthermore, these chemical analogs could exhibit multiple bioactivities and potentially be used as drug combinations to treat highly resistant strains.

**In the future**, the substrate scope of MPAO & CPR enzymes can be examined by catalyzing the bioconversion of MPA analogs to UA analogs. If these enzymes would stage a broad substrate spectrum, more avenues could be opened to make artificial SMs or analogs of natural products with new or improved biological activities. More accessory enzymes from different biosynthetic gene clusters could be expressed within bacteria & tested for their substrate scope. Although, most of the research work on CYPs shows that these enzymes are specific in their

catalytic activity. However, none of that work included lichen p450s, as we are the first group to get a purified lichen CYP. Therefore, at this stage, we cannot make any speculations about the substrate specificity of MPAO.

For the **future experiments**, the yield of MPAO protein can be improved by adding 5-aminolevulinic acid (a heme precursor) [Durairaj, P. 2022] and the functionality CPR357 protein could be enhanced by adding riboflavin (used to form the flavin nucleotide coenzymes) [Pinto, J. T. 2014], see **Section 5.8 of Chapter 5**. For **the future work**, pure MPAO and CPR proteins can be the ideal candidates for determining the X-ray crystal structures of first lichen CYP and CPR proteins. Enzyme-substrate (ES) complexes can also be studied using protein modeling in structural biology.

This study has also proposed some of the unique biosynthetic pathways of putative lichen SMs. All these predictions were made using bioinformatics tools. Phylogenetics combined with the ‘homology mapping’ approach was used to assign the putative function of both biosynthetic genes (encoding polyketide synthases & terpene cyclases) & their accessory genes. Based on the putative functions of the biosynthetic genes of a gene cluster, we predicted rational biosynthetic pathways of one known and a few unknown (novel) SMs. The functional annotation of biosynthetic genes and the proposed biosynthetic pathways results in the identification of many biosynthetic polyketide intermediates. Some of these molecules are structurally simple enough to be chemically synthesized in the lab. Using the protocols established for MPAO & CPR proteins expression, we could express more lichen accessory genes that have been functionally annotated so far (Bertrand 2019). Like UA synthesis, accessory enzymes could catalyze the conversion of chemically synthesized molecules to either novel or known SMs. However, all this data based on genetic similarity & phylogenetics is theoretical & still needs to be experientially demonstrated.

Still, the identification of biosynthetic gene clusters (BGCs) involved in the biosynthesis of lichen SMs can be a costly and complex task, especially due to the genomic diversity of lichen fungal BGCs in general. The shortage of lichen fungal genomic resources makes it challenging to investigate key questions about lichen ecology, evolution, genetics, and physiology. But, at the same time, it also makes them ideal targets for cost-effective genomics & metagenomic projects.

In **Section 5.7.** of **Chapter 5**, the strategies to achieve a soluble expression of eukaryotic membrane-bound proteins in bacteria have been enlisted (**Figure 3**). After many experimental trial-and-error, we got the successful soluble expression of eukaryotic membrane-bound proteins in bacteria. Based on all the experiments that have been discussed in **Chapter 5**, some conclusions have been drawn about the strategies that worked well for our protein functional expression work.

1. High-level soluble expression of membrane-bound lichen fungal proteins was observed after removing the trans-membrane domain (TMD, membrane anchor). On the contrary, we did not observe any expression for both lichen & non-lichen fungal p450s with an intact TMD (full-length genes). As discussed in **Section 5.7.**, there are reports showing a loss of p450 catalytic activity after removing the TMD, but this was not the case with lichen p450 & CPR proteins.
2. High-level soluble expression was observed when expression plasmid with fusion tag/protein such as SUMO (in case of MPAO) was used compared to a smaller fusion tag-His<sub>6</sub> (in case of CPR357 & fungal protein STCB). Therefore, we conclude that SUMO aids proper folding and improves protein solubility.
3. We also tried the co-expression of chaperone proteins with both lichen-fungal & fungal proteins to get the soluble expression of proteins. This strategy did not work in our experiments.

4. Codon-optimized genes for *E. coli* showed a high level of soluble expression compared to non-codon-optimized genes. But we did observe the protein expression in the case of non-codon-optimized genes, although it was low.
5. Another interesting observation we made was that if express fungal p450 (STCB from *A. nidulans*) using pET28b with his (6)-tag, there was no protein expression. When fungal STCB using pET28b with his(6)-tag & lichen CPR357 using pET21a with his(6)-tag co-expressed in a bacterial cell, we observed expression of both proteins. Co-expression experiments worked well at lower temperatures with an induction time of 16-24 h
6. Bioconversion experiments (*in vivo*) of MPA to UA were successful when MPAO & CPR357 co-expressed in a bacterial cell. A similar observation was made in the case of co-expressed STCB (fungal p450) & CPR357 (lichen fungal CPR). We did not observe any UA in the case of *in vitro* experiments when MPA was incubated with either cell lysates with expressed proteins or pure enzymes.

There are a few shortcomings of this research, **1)** To express the eukaryotic membrane-bound protein in bacteria, we still have to go through a laborious task of genetic manipulation of eukaryotic proteins. Including removal or modification of the N-terminal domain, codon-optimization for *E. coli*, and still, it is not easy to get high levels of soluble expression of lichen proteins. As we have seen in the CPR357 lichen redox protein case, the concentration of soluble protein was so low even after optimizing the expression conditions. Because of that, we could not perform bioconversion experiments of MPA to UA with pure proteins, **2)** Low percentage yield of chemically synthesized substrates and hence their limited availability to investigate the enzymatic activity of expressed proteins. Therefore, we cannot produce SMs in large amounts without improving the yield of substrate molecules.

In future, more experiments can be performed to quantify the purified proteins and perform more trials of *in vitro* biotransformation experiments of MPA to UA catalyzed by pure MPAO & CPR357 proteins.



## References

- Durairaj, P., Li, S. Functional expression and regulation of eukaryotic cytochrome P450 enzymes in surrogate microbial cell factories, *Engineering Microbiology*, **2022**, 2(1), 2667-3703, <https://doi.org/10.1016/j.engmic.2022.100011>.
- Pinto, J. T., Cooper, A. J. L. From Cholesterogenesis to Steroidogenesis: Role of Riboflavin and Flavoenzymes in the Biosynthesis of Vitamin D, *Advances in Nutrition*, **2014**, 5(2), 144–163, <https://doi-org.uml.idm.oclc.org/10.3945/an.113.005181>

# Appendix

## Supporting information items

### Chapter 1

*Not applicable*

### Chapter 2

**Table S1:** Primers used throughout this work

Name	Sequence (5' to 3')	Purpose (s)
CU-MPAO-F	CGCGAACAGATTGGAGGTGGA- ATGATTTACCCAGTGTCT	Amplifies <i>mpao</i> ( <i>C. uncialis</i> ); contains InFusion element for plasmid construction.
CU-MPAO-R	GTGGCGCCGCTATATTAAG- CCTTTTCATCCGGA ACTAT	Amplifies <i>mpao</i> ( <i>C. uncialis</i> ); contains 21nt to insert gene into plasmid
SUMO-His <sub>6</sub> - Eco-MPAO-F	CGCGAACAGATTGGAGGTGGC- ATGCGCTTACGCCTGGCG	Amplifies <i>mpao</i> ; without TMD (codon optimized for translation in <i>E. coli</i> ); attaches N-terminal His(6)-SUMO-tag; contains InFusion element for plasmid construction.

SUMO-His <sub>6</sub> - Eco-MPAO-R	GTGGCGGCCGCTATATTA- TGCCTTTTCGTCCGGCACAAT	Amplifies <i>mpao</i> ; without TMD (codon-optimized for translation in <i>E. coli</i> ); contains 21nt to insert gene into plasmid.
His <sub>6</sub> -Eco- MPAO-F	AGCCATATG- CGCTTACGCCTGGCGTATAG	Amplifies <i>mpao</i> (codon-optimized for translation in <i>E. coli</i> ); attaches N-terminal His(6)-tag; contains restriction enzyme element for plasmid construction
His <sub>6</sub> -Eco- MPAO-R	ATCGAATTC- TCATGCCTTTTCGTCCG	Amplifies <i>mpao</i> ; without TMD (codon optimized for translation in <i>E. coli</i> ); contains restriction enzyme element for plasmid construction
MPAO- pDuet-F	ATGGCCGGCC- ATGCGCTTACGCCTGGCGTATAGC	Amplifies <i>mpao</i> ; without TMD (codon optimized for translation in <i>E. coli</i> ); contains restriction enzyme element for plasmid construction
MPAO- pDuet-R	ATGGTACC- TCATGCCTTTTCGTCCGGCAC	Amplifies <i>mpao</i> ; without TMD (codon optimized for translation in <i>E. coli</i> ); contains restriction enzyme element for plasmid construction
CU-CPR 357-F	AAGGAGATATACATATG- GACCCCTTATGCAGGCGCTTACT	Amplifies <i>cpr357</i> ; without TMD; attaches N-terminal S-tag; contains restriction enzyme element for plasmid construction
CU-CPR 357-R	GACGGAGCTCGAATTCGG- AGACCAAACATCTTCTTGGTACTGATTACTC	Amplifies <i>cpr357</i> ; without TMD; contains restriction enzyme element for plasmid construction

CPR357- pDuet-F	ATGAAGCTT- ATGGACCCTTATGCAGGCG	Amplifies <i>mpao</i> (codon-optimized for translation in <i>E. coli</i> ); attaches N-terminal His(6)-tag; contains restriction enzyme element for plasmid construction
CPR357- pDuet-R	ATGCGGCCGC- CTAAGACCAAACATCTTCTTGG	Amplifies <i>mpao</i> (codon-optimized for translation in <i>E. coli</i> ); contains restriction enzyme element for plasmid construction

**Table S2:** Plasmids used throughout this work

<b>Plasmid name</b>	<b>Marker</b>	<b>Description</b>
MPAO-pETite	Kanamycin	Contains <i>mpao</i> from <i>C. uncialis</i>
Eco-MPAO-pETite	Kanamycin	Contains <i>mpao</i> codon-optimized for <i>E. coli</i> ; N-terminal His(6)-SUMO-tagged variant of entry above
Eco-MPAO-pET28b	Kanamycin	Contains <i>mpao</i> codon-optimized for <i>E. coli</i> ; N-terminal His(6)-tagged variant of entry above
CPR357-pET21a	Ampicillin	Contains <i>cpr357</i> from <i>C. uncialis</i> ; N-terminal S-tagged
Eco-MPAO-CPR357- pETDuet-1	Ampicillin	Contains both, <i>mpao</i> codon-optimized for <i>E. coli</i> & <i>cpr357</i> from <i>C. uncialis</i> ; N-terminal His(6)-tagged
STCB-pET28a	Kanamycin	Contains Cytochrome p450 ( <i>stcb</i> ) from <i>A. nidulans</i> ; N-terminal His(6)-tagged
SCYP-pET28b	Kanamycin	Contains Cytochrome p450 ( <i>scyp</i> ) from <i>T. marneffi</i> ; N-terminal His(6)-tagged
pG-KJE8	Chloramphenicol	Contains chaperones <i>dnaK-dnaJ-grpE</i> , <i>groES-groEL</i>
pGro7	Chloramphenicol	Contains chaperones <i>groES-groEL</i>
pKJE7	Chloramphenicol	Contains chaperones <i>dnaK-dnaJ-grpE</i>
pG-Tf2	Chloramphenicol	Contains chaperones <i>groES-groEL-tig</i>
pTf16	Chloramphenicol	Contains chaperone <i>tig</i>

# Chapter 3

*Not Applicable*

## Chapter 4

**Table S3:** Bacterial Strains

<i>S. aureus</i> strain	Genotype	Resistance	Reference
ATCC 29213	Lab isolate	N/A	28
CF 3	Clinical isolate MRSA ST 5, SCCmec Iv	ERY, CLI, CIP	13
CF 4	Clinical isolate MRSA ST 5, SCCmec Iv	ERY, CLI, GEN	13
CF 66	Clinical isolate MSSA		13
CF 188	Clinical isolate MRSA	GEN, ERY, TET, CIP, LZD	13
CF 224	Clinical isolate MRSA ST 30, SCCmec IV	N/A	13
CF 225	Clinical isolate MRSA ST 30, SCCmec IV	CIP, LZD	13
CF 250	Clinical isolate MRSA (SCV)	GEN, ERY, CLI, CIP	13

ERY- erythromycin, CLI-clindamycin, CIP-ciprofloxacin, GEN-gentamicin, TET-tetracycline, LZD- linezolid

**Table S4:** Minimum biofilm eradication concentration (MBEC) of active phloroglucinol derivatives against *S. aureus* strains.

	<b>MBEC (<math>\mu\text{g/ml}</math>)</b>						
<b>Compound no.</b>	<i>S. aureus</i> CF 3	<i>S. aureus</i> CF 4	<i>S. aureus</i> CF 66	<i>S. aureus</i> CF 188	<i>S. aureus</i> CF 224	<i>S. aureus</i> CF 225	<i>S. aureus</i> CF 250
9	>256	>256	>256	>256	>256	>256	>256
10	>256	>256	>256	>256	>256	>256	>256
11	>64	>64	>64	>64	>64	>64	>64
12	>256	>256	>256	>256	>256	>256	>256



## **Synthesis and characterization data for Compounds 1 – 12**

### **Synthesis of 1-(2,4,6-trihydroxy-3-methylphenyl)propanone (1)**

Trihydroxytoluene (0.154 g, 1.0 mmol) was dissolved in 6.0 ml of dioxane under argon, to which  $\text{BF}_3\text{OEt}_2$  (0.123 ml, 1.0 mmol, 1.0 equiv.) was added. The mixture was heated to 40-45°C and then propionic anhydride (0.065 ml, 0.5 mmol, 0.5 equiv.) was added slowly to the reaction mixture over 1.5 hr with each addition of 0.016 ml in every 30 min. After the complete addition of propionic anhydride, the reaction mixture was heated for another 30 min. Then the mixture was cooled down to room temperature (r. t.) and poured into 10.0 ml of 1M HCl and extracted with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as an eluent to provide **1** (0.038g, 19 % yield).

$^1\text{H}$  NMR (300 MHz, Acetone- $d_6$ ):  $\delta$  = 1.11 (3H, t,  $J$  = 7.2 Hz,  $\text{COCH}_2\text{CH}_3$ ), 1.95 (3H, s,  $\text{PhCH}_3$ ) 3.08 (2H, q,  $J$  = 7.2 Hz,  $\text{COCH}_2\text{CH}_3$ ), 6.05 (1H, s, aromatic  $H$ ) 8.03 (1H, br s,  $\text{OH}_a$ ) 9.48 (1H, br s,  $\text{OH}_b$ ) 13.85 (1H, s,  $\text{OH}_c$ ).

$^{13}\text{C}$  NMR (300 MHz Acetone- $d_6$ ):  $\delta$  = 7.5, 9.1, 37.6, 94.9, 103.4, 104.9, 160.4, 162.7, 165.0, 207.0.

MS (ESI, 80 eV):  $m/z$  (%) = 195.3 [ $\text{M} - \text{H}^+$ ] (100).

### **Synthesis of 1-(2, 4, 6-trihydroxy-3-methylphenyl)butanone (2)**

Trihydroxytoluene (0.154 g, 1.0 mmol) was dissolved in 6.0 ml of dioxane under argon, to which  $\text{BF}_3\text{OEt}_2$  (0.123 ml, 1.0 mmol, 1.0 equiv.) was added. The mixture was heated to 40-45°C and then

butanoic anhydride (0.082 ml, 0.5 mmol, 0.5 equiv.) was added to the reaction mixture over 1.5 hr with each addition of 0.020 ml in every 30 min. After the complete addition of propionic anhydride, the reaction mixture was stirred for another 30 min. Then the mixture was cooled down to r. t. and poured into 10.0 ml of 1M HCl and extracted with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as an eluent to provide **2** (0.030g, 14 % yield).

<sup>1</sup>H NMR (300 MHz, Acetone-*d*<sub>6</sub>): δ = 0.95 (3H, t, *J* = 7.4 Hz, COCH<sub>2</sub>CH<sub>3</sub>), 1.62-1.74 (2H, m, COCH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>) 1.95 (3H, s, PhCH<sub>3</sub>) 3.04 (2H, t, *J* = 7.4 Hz, COCH<sub>2</sub>CH<sub>3</sub>), 6.05 (1H, s, aromatic *H*) 9.05 (1H, br s, OH<sub>a</sub>) 9.39 (1H, br s, OH<sub>b</sub>) 13.87 (1H, s, OH<sub>c</sub>).

<sup>13</sup>C NMR (300 MHz Acetone-*d*<sub>6</sub>): δ = 7.5, 14.3, 18.9, 46.4, 94.9, 103.4, 105.0, 160.4, 162.7, 165.4, 206.5.

MS (ESI, 80 eV): *m/z* (%) = 209.3 [M - H<sup>+</sup>] (100).

### **Synthesis of 1-(2, 4, 6-trihydroxy-3-methylphenyl)pentanone (3)**

Trihydroxytoluene (0.154 g, 1.0 mmol) was dissolved in 6.0 ml of dioxane under argon, to which BF<sub>3</sub>OEt<sub>2</sub> (0.123 ml, 1.0 mmol, 1.0 equiv.) was added. The mixture was heated to 60°C and then pentanoic anhydride (0.098 ml, 0.5 mmol, 0.5 equiv.) was added slowly to the reaction mixture over 1.5 hr with each addition of 0.024 ml in every 30 min. After the complete addition of propionic anhydride, the reaction mixture was stirred for another 30 min. Then the mixture was cooled down to r. t. and poured into 10.0 ml of 1M HCl and extracted with EtOAc (3 x 10.0 ml).

The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as an eluent to provide **3** (0.020g, 9 % yield).

<sup>1</sup>H NMR (300 MHz, Acetone-*d*<sub>6</sub>): δ = 0.91 (3H, t, *J* = 7.4 Hz, COCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 1.31-1.43 (2H, m, COCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>) 1.59-1.69 (2H, m, COCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 1.95 (3H, s, PhCH<sub>3</sub>), 3.06 (2H, t, *J* = 7.2 Hz, COCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 6.05 (1H, s, aromatic *H*) 8.99 (1H, br s, OH<sub>a</sub>) 9.39 (1H, br s, OH<sub>b</sub>) 13.87 (1H, s, OH<sub>c</sub>).

<sup>13</sup>C NMR (300 MHz Acetone-*d*<sub>6</sub>): δ = 7.5, 14.3, 23.3, 27.8, 42.2, 94.9, 103.4, 105.1, 160.3, 162.72, 165.1, 206.7.

MS (ESI, 80 eV): *m/z* (%) = 223.4 [M - H<sup>+</sup>] (100), 224.2 [M<sup>+</sup>] (16).

#### **Synthesis of 1-(2, 4, 6-trihydroxy-3-methylphenyl)hexanone (4)**

Trihydroxytoluene (0.154 g, 1.0 mmol) is dissolved in 6.0 ml of dioxane under argon, to which BF<sub>3</sub>OEt<sub>2</sub> (0.123 ml, 1.0 mmol, 1.0 equiv.) was added. The mixture was heated to 55-60 °C and then hexanoic anhydride (0.115 ml, 0.5 mmol, 0.5 equiv.) was added to the reaction mixture all at once. The mixture was stirred for another 40 min. After completion, mixture was cooled down to r. t. and poured into 10.0 ml of 1M HCl and extracted with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as a eluent to provide **4** (0.020g, 8% yield).

$^1\text{H}$  NMR (300 MHz, Acetone- $d_6$ ):  $\delta$  = 0.89 (3H, t,  $J$  = 6.9 Hz,  $\text{COCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_3$ ), 1.31-1.36 (4H, m,  $\text{COCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_3$ ), 1.62-1.71 (2H, m,  $\text{COCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_3$ ), 1.95 (3H, s,  $\text{PhCH}_3$ ), 3.06 (2H, t,  $J$  = 7.2 Hz,  $\text{COCH}_2\text{CH}_2\text{CH}_2\text{CH}_3$ ), 6.05 (1H, s, aromatic  $H$ ) 8.95 (1H, br s,  $\text{OH}_a$ ) 9.50 (1H, br s,  $\text{OH}_b$ ) 13.88 (1H, s,  $\text{OH}_c$ ).

$^{13}\text{C}$  NMR (300 MHz Acetone- $d_6$ ):  $\delta$  = 7.5, 14.3, 23.3, 25.4, 32.5, 44.5, 94.9, 103.4, 105.0, 160.3, 162.7, 165.1, 206.6.

MS (ESI, 80 eV):  $m/z$  (%) = 237.4 [ $\text{M} - \text{H}^+$ ] (100) 238.3 [ $\text{M}^+$ ] (12).

### **Synthesis of 1-(2,4,6-trihydroxyphenyl)propanone (5)**

Phloroglucinol (0.630 g, 5.0 mmol) was dissolved in 5.0 ml of  $\text{CS}_2$  and 10.0 ml of  $\text{PhNO}_2$ , to which anhydrous  $\text{AlCl}_3$  was added (1.995g, 15.0 mmol, 3.0 equiv.). The mixture was refluxed at 45 °C for 10 min. Propionyl chloride (0.262 ml, 3.0 mmol, 0.6 equiv.) dissolved in 5 ml of nitrobenzene was added slowly to the reaction mixture for over 10 min before heating the mixture for another 1 h. The reaction mixture was allowed to cool down and poured over ice-water bath and then add 3-4 drops of conc. HCl. The mixture was stirred with a glass rod and brought to r. t. before extracting it with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as a eluent to provide **5** (0.288 g, 32% yield) (63–200 mesh; Merck).

MP: 101-103 °C.

Rf: 0.4 ( $\text{CHCl}_3$ -MeOH, 5:1).

IR (KBr): 3254, 3110, 1710, 1680, 1535, 1460, 970  $\text{cm}^{-1}$ .

UV/Vis  $\lambda_{\text{max}}$  (MeOH) nm ( $\log \epsilon$ ): 234 (3.80), 280 (4.52), 324 (3.45).

$^1\text{H}$  NMR (300 MHz, Acetone- $d_6$ ):  $\delta$  = 1.12 (3H, t,  $J$  = 7.3 Hz,  $\text{COCH}_2\text{CH}_3$ ), 3.10 (2H, q,  $J$  = 7.3 Hz,  $\text{COCH}_2\text{CH}_3$ ), 5.93 (2H, s, aromatic H), 9.13 (1H, br s, OH), 11.69 (2H, s, OH).

$^{13}\text{C}$  NMR (100 MHz DMSO- $d_6$ ):  $\delta$  = 8.9, 29.8, 37.4, 95.645, 105.289, 165.489, 166.502, 206.248.

MS (ESI, 80 eV):  $m/z$  (%) = 290.2 [ $\text{M} - \text{H}^+$ ] (100), 265.9 (90).

### **Synthesis of 1-(2,4,6-trihydroxyphenyl)butanone (6)**

Phloroglucinol (1.260 g, 10.0 mmol) was dissolved in 10.0 ml of  $\text{CS}_2$  and 15.0 ml of  $\text{PhNO}_2$ , to which anhydrous  $\text{AlCl}_3$  (1.333 g, 10.0 mmol, 1.0 equiv.) was added. The mixture was heated to 35-40°C and then butyryl chloride (1.03 ml, 10.0 mmol, 1.0 equiv.) was added slowly to the reaction mixture for over 10-12 min before heating the mixture for another 1 h. The reaction mixture was allowed to cool down to r. t. and poured over an ice-water bath and then add 3-4 drops of conc. HCl. The mixture was stirred with a glass rod and brought to r. t. before extracting it with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as an eluent to provide **6** (1.090 g, 55.6% yield). (63–200 mesh; Merck)

$^1\text{H}$  NMR (300 MHz, Acetone- $d_6$ ):  $\delta$  = 0.96 (3H, t,  $J$  = 7.3 Hz,  $\text{COCH}_2\text{CH}_2\text{CH}_3$ ), 1.69 (2H, m,  $\text{COCH}_2\text{CH}_2\text{CH}_3$ ), 3.05 (2H, t,  $J$  = 7.5 Hz,  $\text{COCH}_2\text{CH}_2\text{CH}_3$ ) 5.93 (2H, s, aromatic H), 9.14 (1H, br s, OH), 11.71 (2H, s, OH).

$^{13}\text{C}$  NMR (300 MHz Acetone- $d_6$ ):  $\delta$  = 14.3, 18.8, 46.4, 95.8, 105.2, 165.1, 165.4, 206.4.

MS (ESI, 80 eV):  $m/z$  (%) = 195.3 [ $\text{M} - \text{H}^+$ ] (100).

### **Synthesis of 1-(2,4,6-trihydroxyphenyl)pentanone (7)**

Phloroglucinol (0.630 g, 5.0 mmol) was dissolved in 10.0 ml of  $\text{CS}_2$  and 15.0 ml of  $\text{PhNO}_2$ , to which anhydrous  $\text{AlCl}_3$  (1.333 g, 10.0 mmol, 2.0 equiv.) was added. The mixture was heated to  $50^\circ\text{C}$  and then valeroyl chloride (1.42 ml, 12.0 mmol, 2.4 equiv.) was added slowly to the reaction mixture for over 10-12 min before heating the mixture for another 1 h. The reaction mixture was allowed to cool down to r. t. and poured over an ice-water bath and then add 3-4 drops of conc.  $\text{HCl}$ . The mixture was stirred with a glass rod and brought to r. t. before extracting it with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as an eluent to provide **7** (0.480 mg, 46 % yield).

$^1\text{H}$  NMR (300 MHz, Acetone- $d_6$ ):  $\delta$  = 0.93 (3H, t,  $J$  = 7.2 Hz,  $\text{COCH}_2\text{CH}_2\text{CH}_3$ ), 1.39 (2H, m,  $\text{COCH}_2\text{CH}_2\text{CH}_2\text{CH}_3$ ), 1.65 (2H, m,  $\text{COCH}_2\text{CH}_2\text{CH}_2\text{CH}_3$ ) 3.07 (2H, t,  $J$ =7.7 Hz,  $\text{COCH}_2\text{CH}_2\text{CH}_2\text{CH}_3$ ) 5.93 (2H, s, aromatic H), 9.13 (1H, br s,  $\text{OH}_a$ ), 11.70 (2H, s,  $\text{OH}_b$ ).

$^{13}\text{C}$  NMR (300 MHz Acetone- $d_6$ ):  $\delta$ = 14.3, 23.2, 27.7, 44.2, 95.8, 105.2, 165.1, 165.4, 206.6.

MS (ESI, 80 eV):  $m/z$  (%) = 209.4 [M – H<sup>+</sup>] (100).

### **Synthesis of 1-(2,4,6-trihydroxyphenyl)hexanone (8)**

Phloroglucinol (0.630 g, 5.0 mmol) was dissolved in 5.0 ml of CS<sub>2</sub> and 15.0 ml of PhNO<sub>2</sub>, to which anhydrous AlCl<sub>3</sub> (1.999 g, 15.0 mmol, 3.0 equiv.) was added. The mixture was heated to 55-60°C and then hexanoyl chloride (0.295 ml, 3.0 mmol, 0.6 equiv.) was added slowly to the reaction mixture for over 10-12 min before heating the mixture for another 30 min. The reaction mixture was allowed to cool down to r. t. and poured over an ice-water bath and then add 3-4 drops of conc. HCl. The mixture was stirred with a glass rod and brought to r. t. before extracting it with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as an eluent to provide **8** (0.275 g, 24 % yield)

<sup>1</sup>H NMR (300 MHz, Acetone-*d*<sub>6</sub>): δ = 0.90 (3H, t, *J* = 7.1 Hz, COCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 1.34 (4H, m, COCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 1.68 (2H, m, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 3.07 (2H, t, *J* = 7.4 Hz, COCH<sub>2</sub>), 5.94 (2H, s, aromatic H), 9.12 (1H, s, PhOH), 11.71 (2H, s, PhOH).

<sup>13</sup>C NMR (300 MHz Acetone-*d*<sub>6</sub>): δ = 14.3, 23.3, 25.3, 32.5, 44.4, 95.8, 105.2, 165.1, 165.4, 206.6.

MS (ESI, 80 eV):  $m/z$  (%) = 223.4 [M – H<sup>+</sup>] (100)

### **1,3-(2,4,6-trihydroxyphenyl)dipropanone (9)**

Phloroglucinol (0.630 g, 5.0 mmol) was dissolved in 2.0 ml of dioxane under argon, to which BF<sub>3</sub>OEt<sub>2</sub> (1.500 ml, 12.0 mmol, 2.4 equiv.) was added. The mixture was heated to 40-45°C and

then propanoic anhydride (1.548 ml, 12.0 mmol, 2.4 equiv.) was added slowly to the reaction mixture for over 12-15 min before heating the mixture for another 30 min. The reaction mixture was allowed to cool down to r. t. and poured over an ice-water bath and then add 3-4 drops of conc. HCl. The mixture was stirred with a glass rod and brought to r. t. before extracting it with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as an eluent to provide **9** (0.210 g, 18 % yield).

<sup>1</sup>H NMR (300 MHz, Acetone-*d*<sub>6</sub>): δ = 1.1 (6H, t, *J*=7.1 Hz, CH<sub>2</sub>CH<sub>3</sub>), 3.1 (4H, q, *J*=7.1 Hz, COCH<sub>2</sub>) 5.94 (1H, s, aromatic H), 12.6 (2H, s, PhOH) 16.4 (1H, s, PhOH).

<sup>13</sup>C NMR (300 MHz Acetone-*d*<sub>6</sub>): δ = 7.76, 36.9, 94.9, 103.5, 168.5, 171.7, 206.9.

MS (ESI, 80 eV): *m/z* (%) = 237.4 [M – H<sup>+</sup>] (100).



### **1,3-(2,4,6-trihydroxyphenyl)dibutanone (10)**

Phloroglucinol (0.630 g, 5.0 mmol) was dissolved in 4.0 ml of dioxane under argon, to which  $\text{BF}_3\text{OEt}_2$  (0.741 ml, 6.0 mmol, 1.2 equiv.) was added. The mixture was heated to  $50^\circ\text{C}$  and then butanoic anhydride (1.963 ml, 12.0 mmol, and 2.4 equiv.) was added slowly to the reaction mixture for over 10-12 min before heating the mixture for another 1 h. The reaction mixture was allowed to cool down to r. t. and poured over an ice-water bath and then add 3-4 drops of conc. HCl. The mixture was stirred with a glass rod and brought to r. t. before extracting it with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as an eluent to provide **10** (0.737 g, 55 % yield).

$^1\text{H}$  NMR (300 MHz, Acetone- $d_6$ ):  $\delta$  = 0.98 (6H, t,  $J=7.5$  Hz,  $\text{COCH}_2\text{CH}_2\text{CH}_3$ ), 1.71 (4H, m,  $\text{COCH}_2\text{CH}_2\text{CH}_3$ ), 3.09 (4H, t,  $J=7.3$  Hz,  $\text{COCH}_2$ ) 5.9 (1H, s, Ph-H), 12.6 (2H, s, PhOH) 16.4 (1H, s, PhO-H).

$^{13}\text{C}$  NMR (300 MHz Acetone- $d_6$ ):  $\delta$  = 14.2, 18.5, 46.5, 96.8, 104.6, 165.5, 172.8, 207.4.

MS (ESI, 80 eV):  $m/z$  (%) = 265.4 [ $\text{M} - \text{H}^+$ ] (100), 266.1 [ $\text{M}^+$ ] (14).

### **1,3-(2,4,6-trihydroxyphenyl)dipentanone (11)**

Phloroglucinol (0.630 g, 5.0 mmol) is dissolved in 2.0 ml of dioxane under argon, to which  $\text{BF}_3\text{OEt}_2$  (1.420 ml, 10.0 mmol, 2.0 equiv.) was added. The mixture was heated to  $40\text{-}45^\circ\text{C}$  and then pentanoic anhydride (1.973 ml, 10.0 mmol, 2.0 equiv.) was added slowly to the reaction mixture for over 10-12 min before heating the mixture for another 30 min. The reaction mixture

was allowed to cool down to r. t. and poured over an ice-water bath and then add 3-4 drops of conc. HCl. The mixture was stirred with a glass rod and brought to r. t. before extracting it with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as an eluent to provide **11** (0.089 g, 6 % yield).

<sup>1</sup>H NMR (300 MHz, Acetone-*d*<sub>6</sub>): δ = 0.98 (6H, t, *J* = 7.5 Hz, COCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 1.70 (4H, m, COCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>), 3.09 (4H, t, *J* = 7.3, COCH<sub>2</sub>) 5.92 (1H, s, aromatic H), 12.61 (2H, s, PhOH) 16.44 (1H, s, PhOH).

<sup>13</sup>C NMR (300 MHz Acetone-*d*<sub>6</sub>): δ = 14.2, 18.5, 46.5, 96.8, 104.6, 165.5, 172.8, 207.4.

MS (ESI, 80 eV): *m/z* (%) = 293.5 [M – H<sup>+</sup>] (100).

### **1,3-(2,4,6-trihydroxyphenyl)dihexanone (12)**

Phloroglucinol (0.630 g, 5.0 mmol) is dissolved 5.0 ml of CS<sub>2</sub> and 15.0 ml of PhNO<sub>2</sub>, to which anhydrous AlCl<sub>3</sub> (1.999 g, 15.0 mmol, 3.0 equiv.) was added. The mixture was heated to 55-60°C and then hexanoic anhydride (0.693 ml, 3.0 mmol, 0.6 equiv.) was added slowly to the reaction mixture for over 10-12 min before heating the mixture for another 30 min. The reaction mixture was allowed to cool down to r. t. and poured over an ice-water bath and then add 3-4 drops of conc. HCl. The mixture was stirred with a glass rod and brought to r. t. before extracting it with EtOAc (3 x 10.0 ml). The combined organic extracts were washed with water (3 x 10.0 ml), brine (2 x 10.0 ml), brine (2 x 10.0 ml), dried with anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The crude product was purified using flash column chromatography on silica

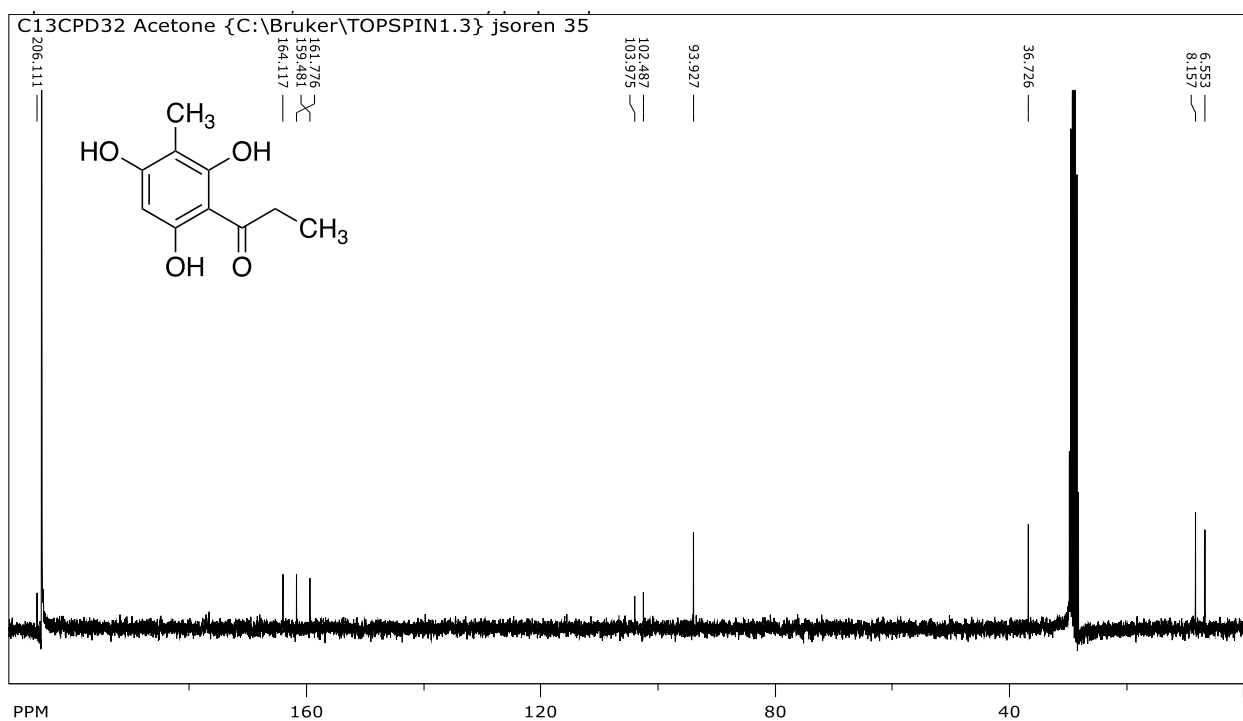
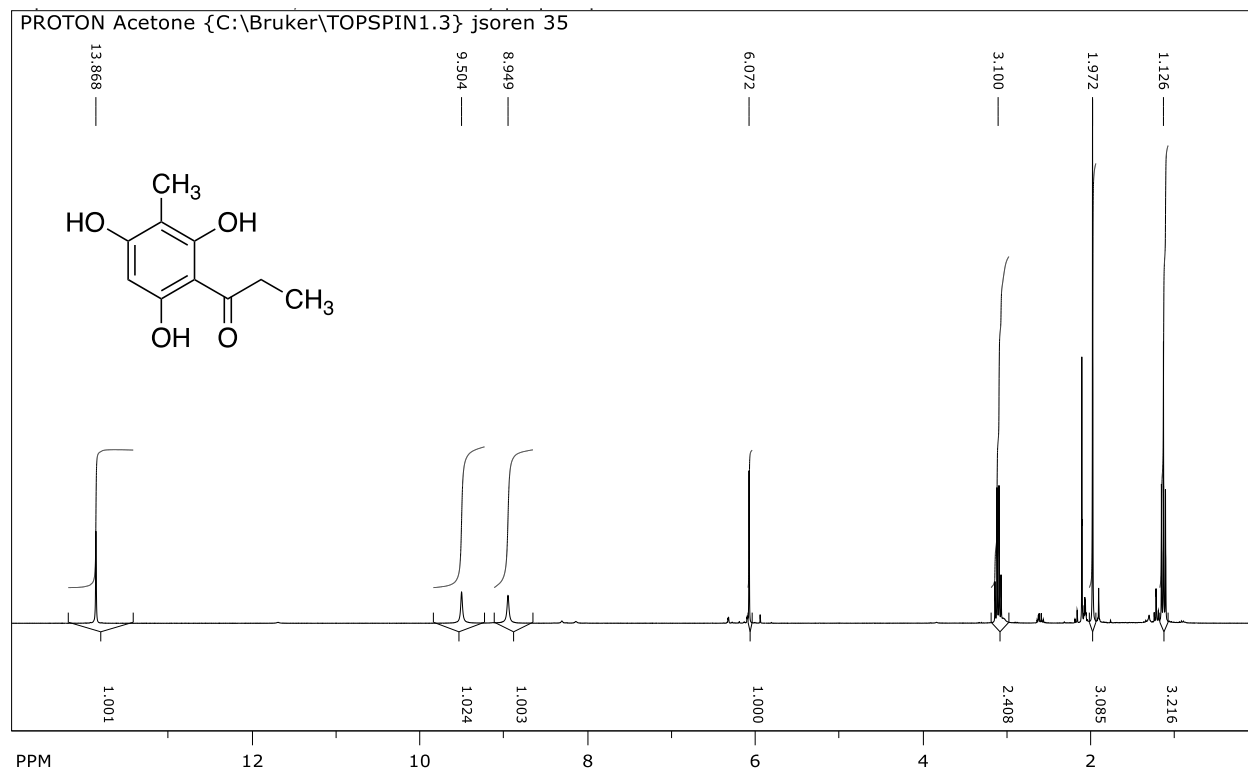
gel with a mixture of Hexane/EtOAc (9:1 to 6:4) as an eluent to provide **12** as a minor product (0.198 g, 12 % yield).

$^1\text{H}$  NMR (300 MHz, Acetone- $d_6$ ):  $\delta$  = 0.91 (6H, t,  $J$  = 7.5 Hz,  $\text{COCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_3$ ), 1.37 (8H, m,  $\text{COCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_3$ ), 1.68 (4H, m,  $\text{COCH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_3$ ), 3.12 (4H, t,  $J$  = 7.325,  $\text{COCH}_2$ ) 5.93 (1H, s, aromatic H), 12.69 (2H, bs, PhOH).

$^{13}\text{C}$  NMR (300 MHz Acetone- $d_6$ )  $\delta$  = 14.3, 23.3, 25.0, 32.3, 44.6, 95.8, 104.6, 172.8, 207.6.

HRMS (ESI, 70 eV):  $m/z$  (%) = 321.1936 [ $\text{M} - \text{H}^+$ ] (100) (321.1707 calcd. for  $\text{C}_{18}\text{H}_{25}\text{O}_5^-$ ).

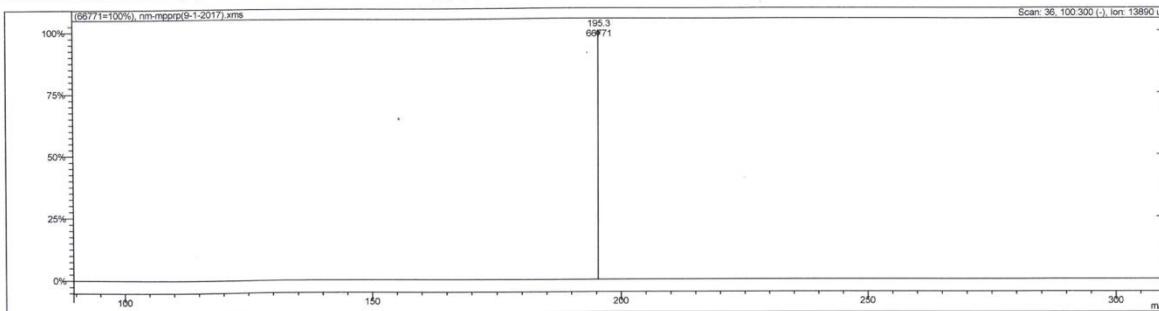
# NMR of compound 1



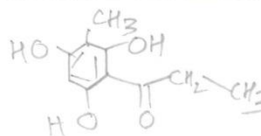
# Mass Spectra of Compound 1:

Print Date: 01 Sep 2017 12:54:47

Scan 36 from ...hemistry personalsorensen\navrit\m-mpprp(9-1-2017).xms



Spectrum from ...ersonalsorensen\navrit\m-mpprp(9-1-2017).xms  
 Scan No: 36, Time: 0.173 minutes  
 No averaging. Not background corrected.  
 Comment: 0.173 min. Scan: 36 100.300 (-) Ion: 13890 us RIC: 88155  
 Pair Count: 1 MW: 0 Formula: None  
 CAS No: None Acquired Range: 99.5 - 300.5 m/z

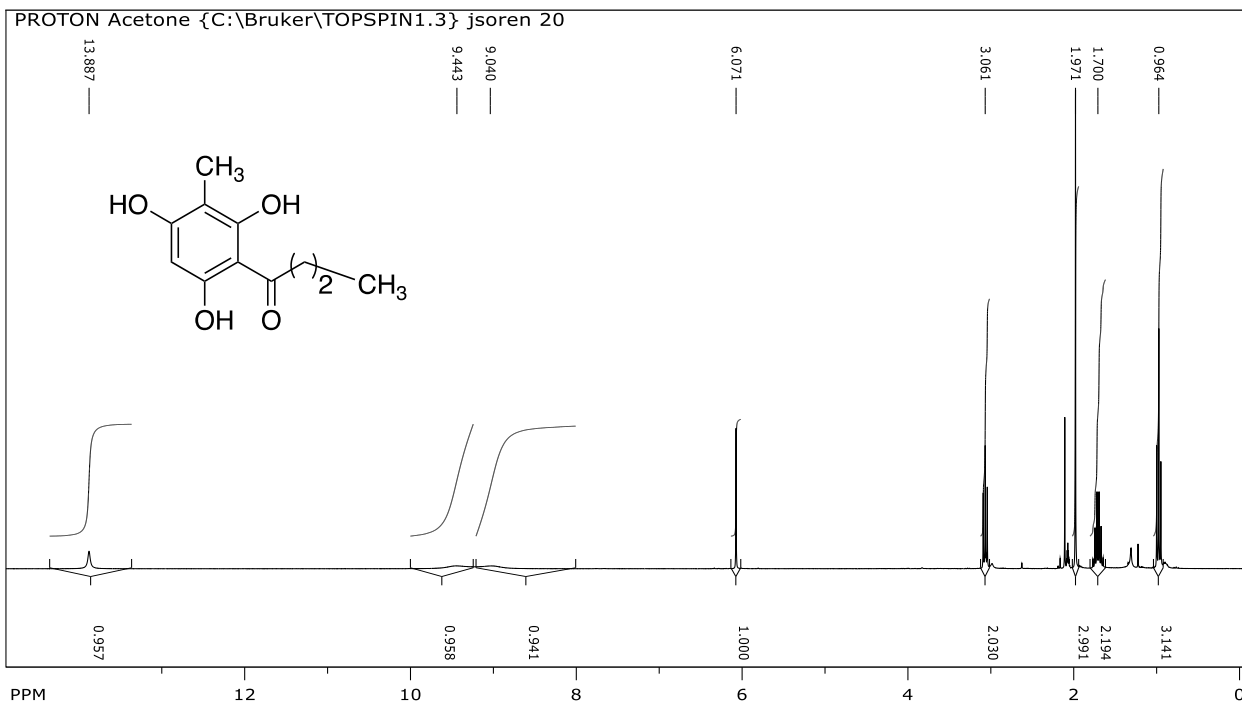


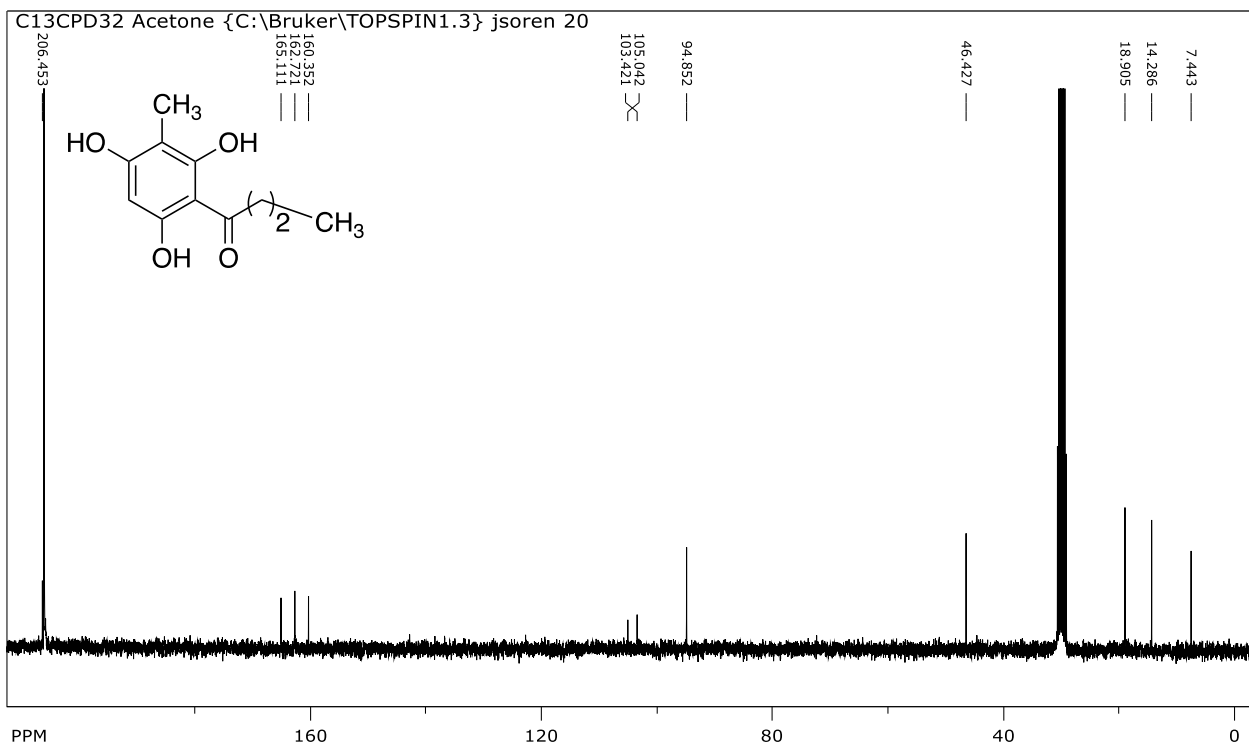
Methylpropano-  
phenone  
MW = 196g.

Method Time: 0.00-0.32, Centroid, Electrospray  
 Seg 1, Time: 0.00-0.32, Scan Functions: 1  
 1. 100:300 (-) 100.300 (-) ESI Standard 80.0[V] Full  
 Product Mass Range: 99.5 - 300.5 m/z

Ion	Int Norm	Ion	Int Norm	Ion	Int Norm	Ion	Int Norm
195.3	66771	999					

# NMR of compound 2

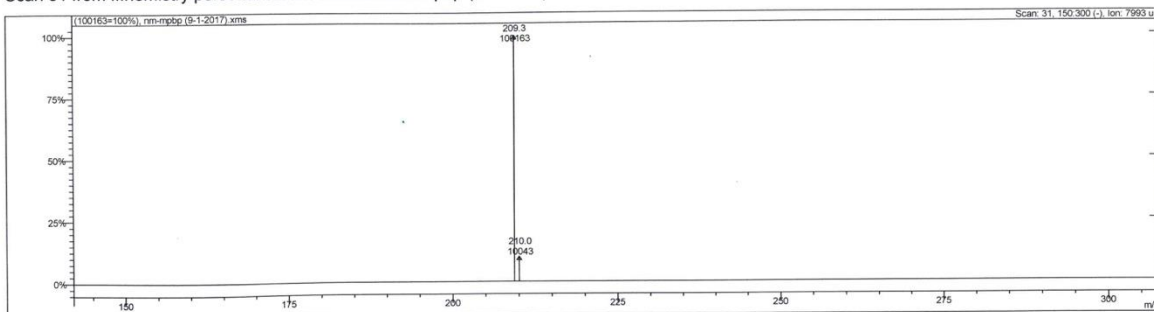




## Mass Spectra of Compound 2:

Print Date: 01 Sep 2017 13:10:38

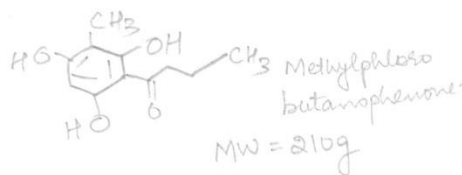
Scan 31 from ...hemistry personal\sorensen\navriti\nm-mpbp (9-1-2017).xms



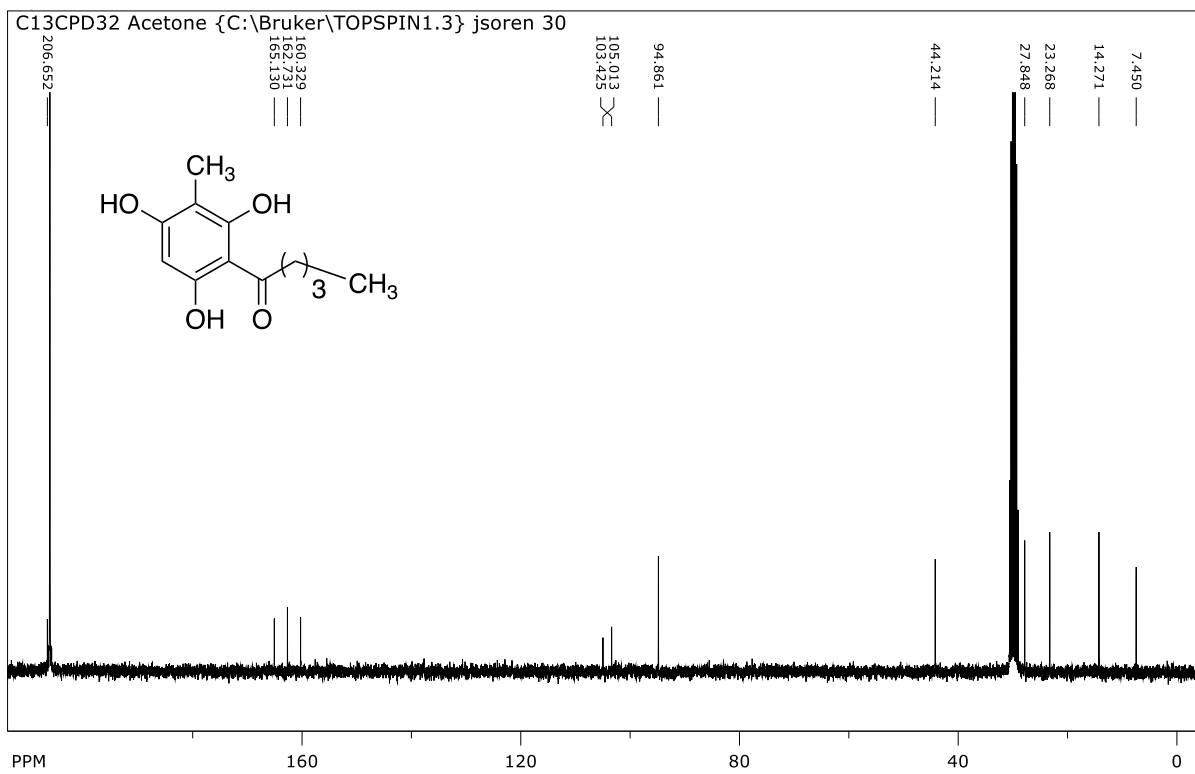
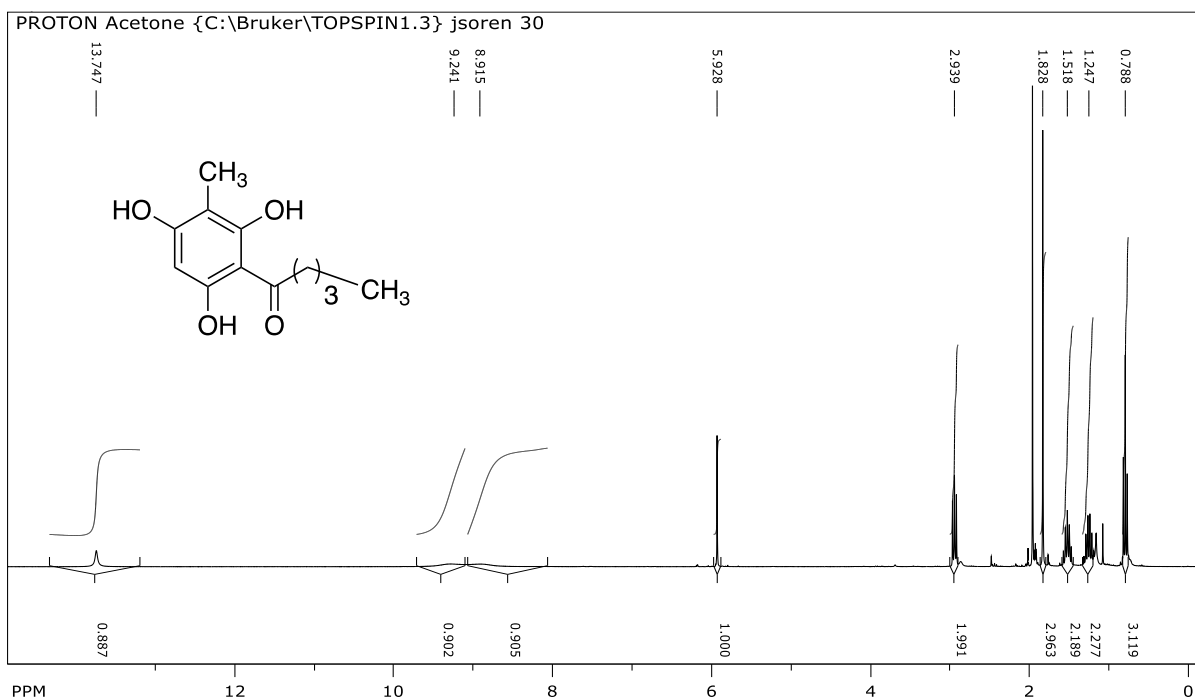
Spectrum from ...ersonal\sorensen\navriti\nm-mpbp (9-1-2017).xms  
 Scan No: 31, Time: 0.133 minutes  
 No averaging, Not background corrected.  
 Comment: 0.133 min. Scan: 31 150:300 (-) Ion: 7993 us RIC: 134145  
 Pair Count: 2 MW: 0 Formula: None  
 CAS No: None Acquired Range: 149.5 - 300.5 m/z

Method Time: 0.00-0.31, Centroid, Electrospray  
 Seg 1, Time: 0.00-0.31, Scan Functions: 1  
 1. 150:300 (-) 150:300 (-) ESI Standard 80.0[V] Full  
 Product Mass Range: 149.5 - 300.5 m/z

Ion	Int	Norm	Ion	Int	Norm	Ion	Int	Norm
209.3	100163	999	210.0	10043	100			



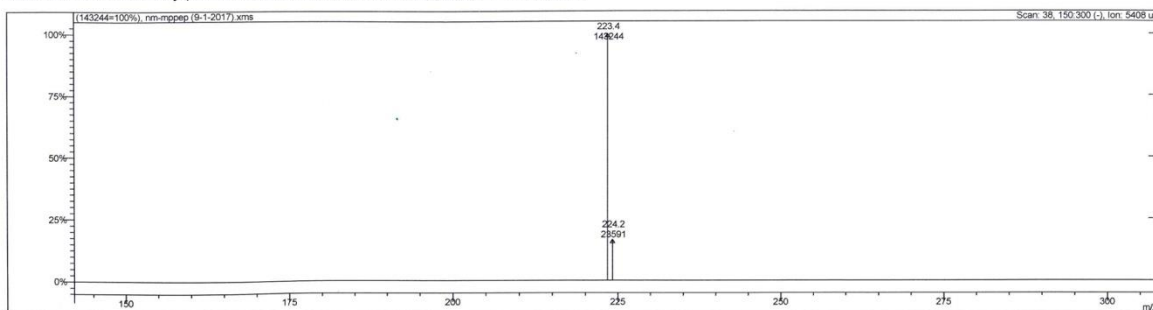
# NMR of compound 3



## Mass Spectra of Compound 3:

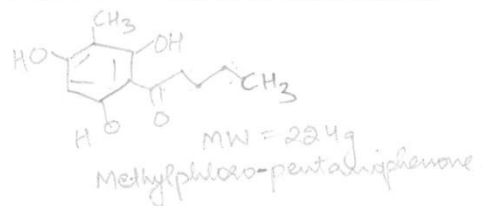
Print Date: 01 Sep 2017 13:20:02

Scan 38 from ...emistry personal\sorensen\navrit\m-mppep (9-1-2017).xms



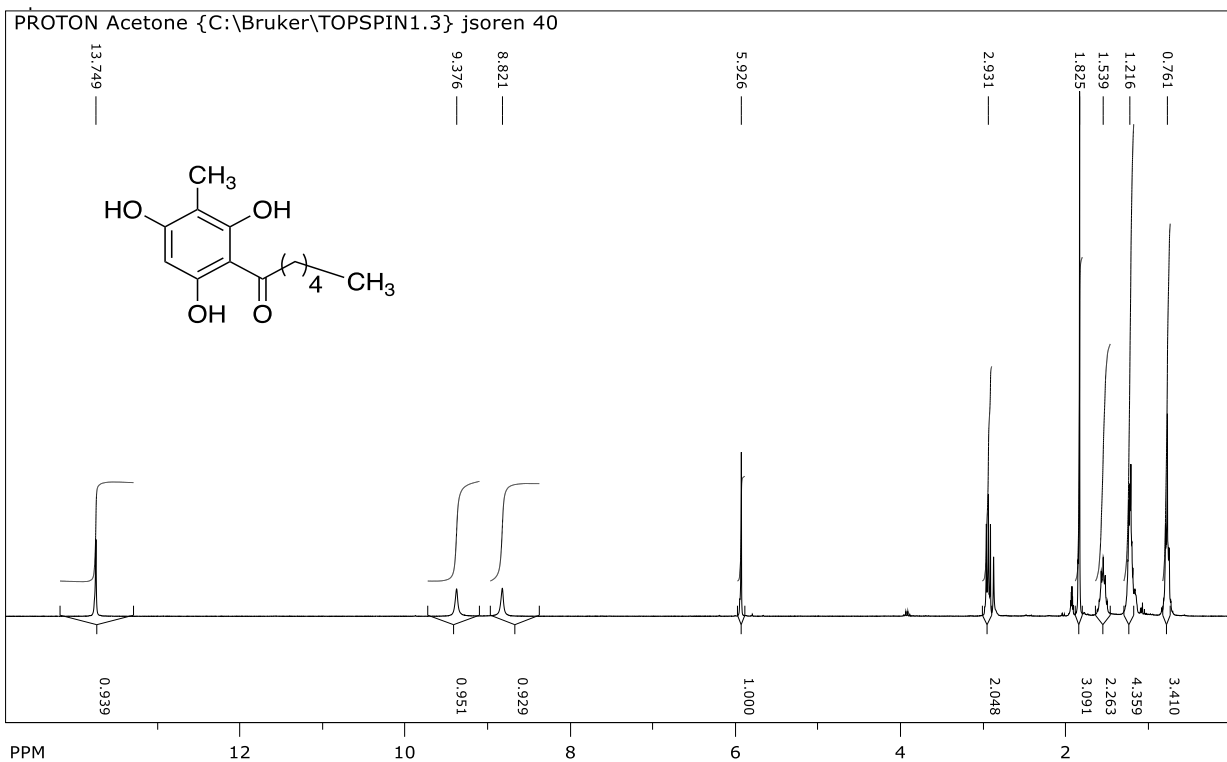
Spectrum from ...sonal\sorensen\navrit\m-mppep (9-1-2017).xms  
 Scan No: 38, Time: 0.157 minutes  
 No averaging, Not background corrected.  
 Comment: 0.157 min. Scan: 38 150.300 (-) Ion: 5408 us RIC: 199665  
 Pair Count: 2 MW: 0 Formula: None  
 CAS No: None Acquired Range: 149.5 - 300.5 m/z

Method Time: 0.00-0.25, Centroid, Electrospray  
 Seg 1, Time: 0.00-0.25, Scan Functions: 1  
 1. 150.300 (-) 150.300 (-) ESI Standard 80.0[V] Full  
 Product Mass Range: 149.5 - 300.5 m/z

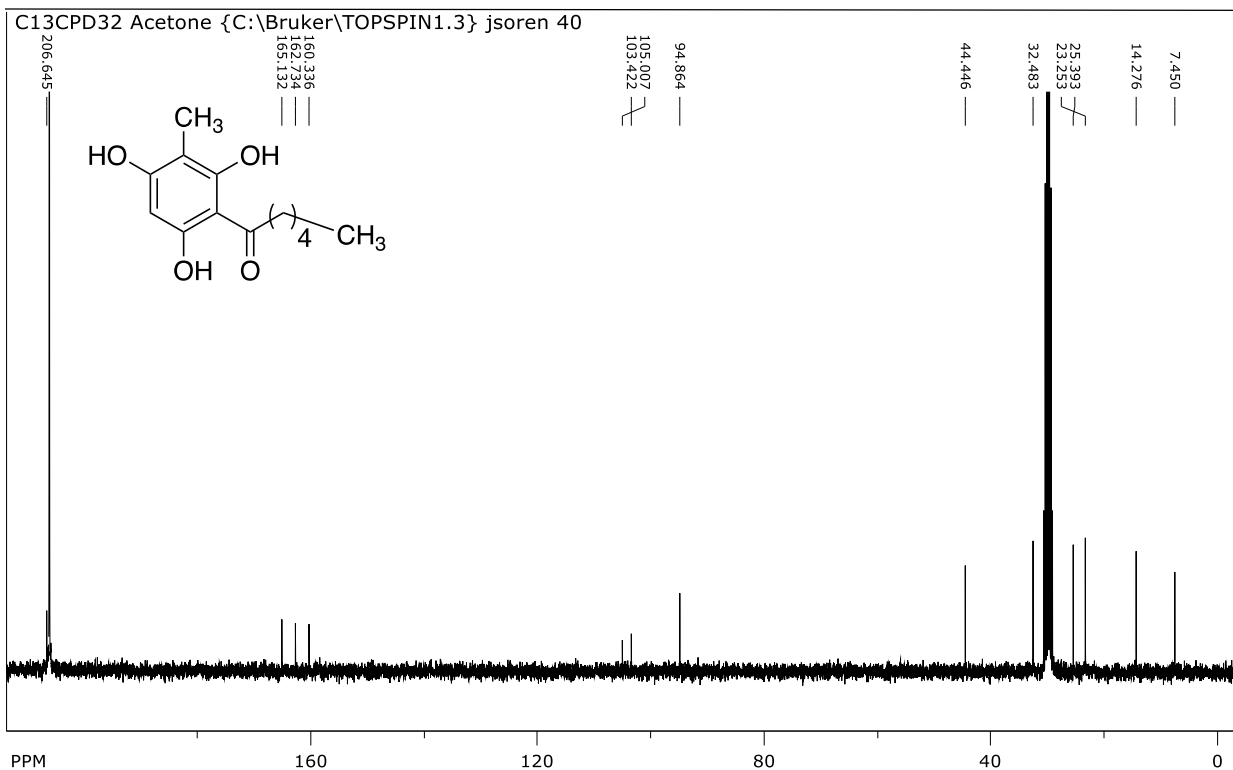


Ion	Int Norm	Ion	Int Norm	Ion	Int Norm	Ion	Int Norm
223.4	143244 999	224.2	23591 165				

## NMR of compound 4



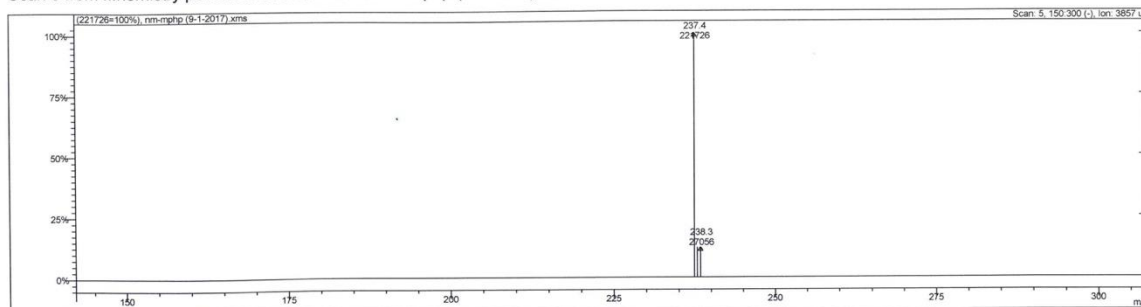




## Mass Spectra of Compound 4:

Print Date: 01 Sep 2017 13:35:29

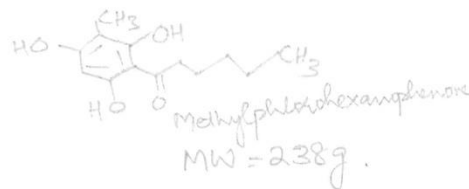
Scan 5 from ...hemistry personal\sorensen\navriti\m-mph (9-1-2017).xms



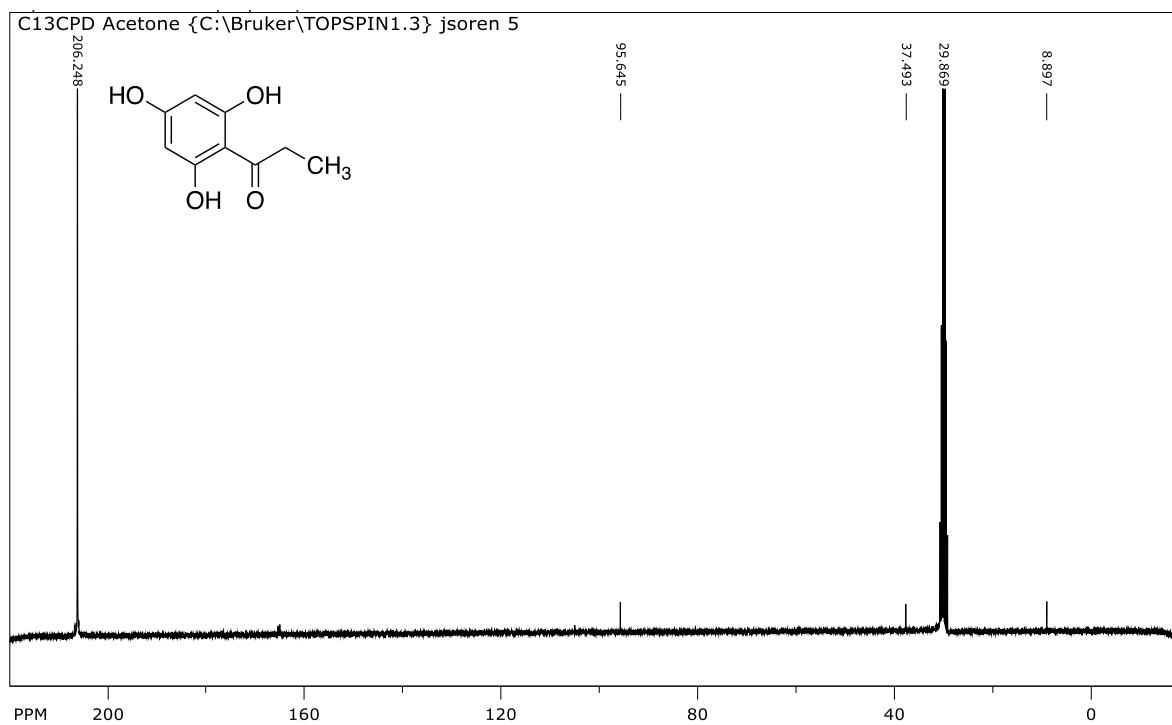
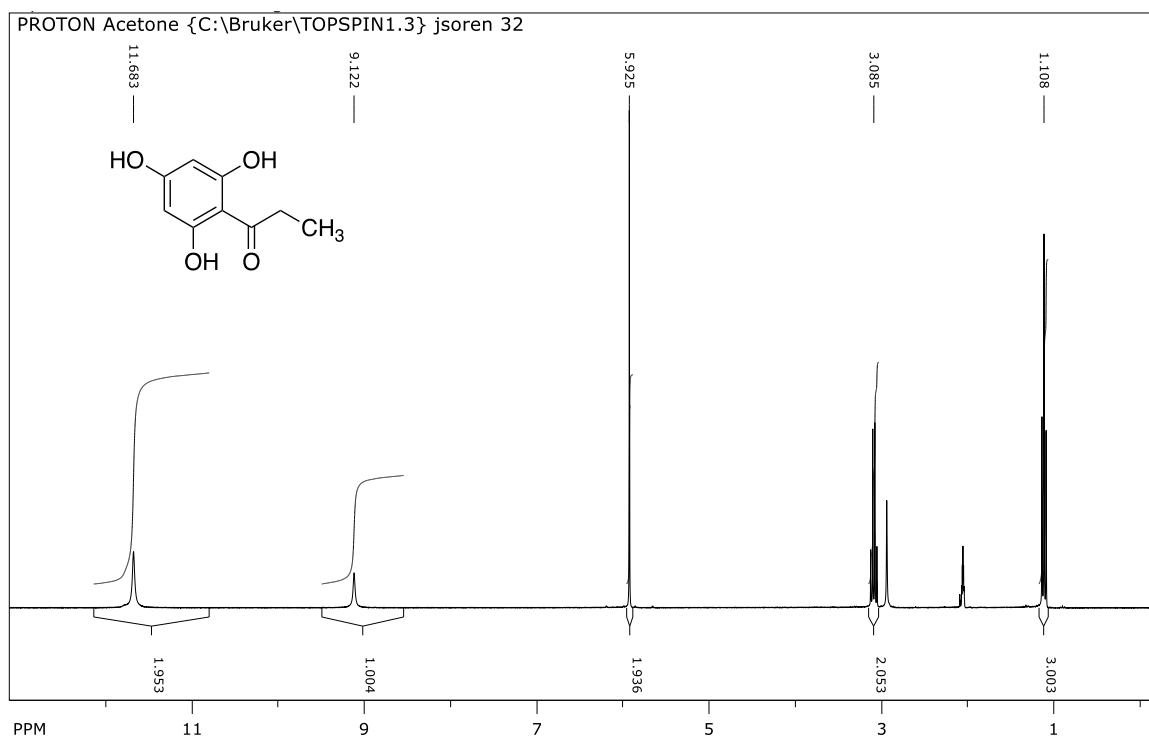
Spectrum from ...ersonal\sorensen\navriti\m-mph (9-1-2017).xms  
 Scan No: 5, Time: 0.017 minutes  
 No averaging. Not background corrected.  
 Comment: 0.017 min. Scan: 5 150.300 (-) Ion: 3857 us RIC: 309525  
 Pair Count: 3 MW: 0 Formula: None  
 CAS No: None Acquired Range: 149.5 - 300.5 m/z

Method Time: 0.00-0.08, Centroid, Electrospray  
 Sep 1, Time: 0.00-0.08, Scan Functions: 1  
 1, 150.300 (-) 150.300 (-)ESI Standard 80.0[V] Full  
 Product Mass Range: 149.5 - 300.5 m/z

Ion	Int Norm	Ion	Int Norm	Ion	Int Norm	Ion	Int Norm
237.4	221726 999	237.8	26806 121	238.3	27056 122		



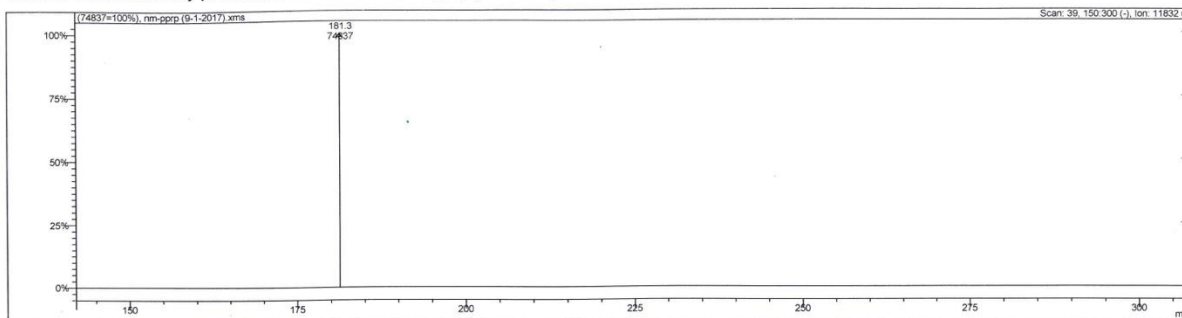
# NMR of Compound 5



## Mass Spectra of Compound 5:

Print Date: 01 Sep 2017 13:41:03

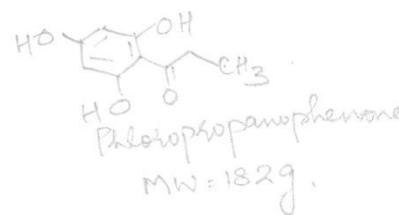
Scan 39 from ...hemistry personal\sorensen\navrit\m-prp (9-1-2017).xms



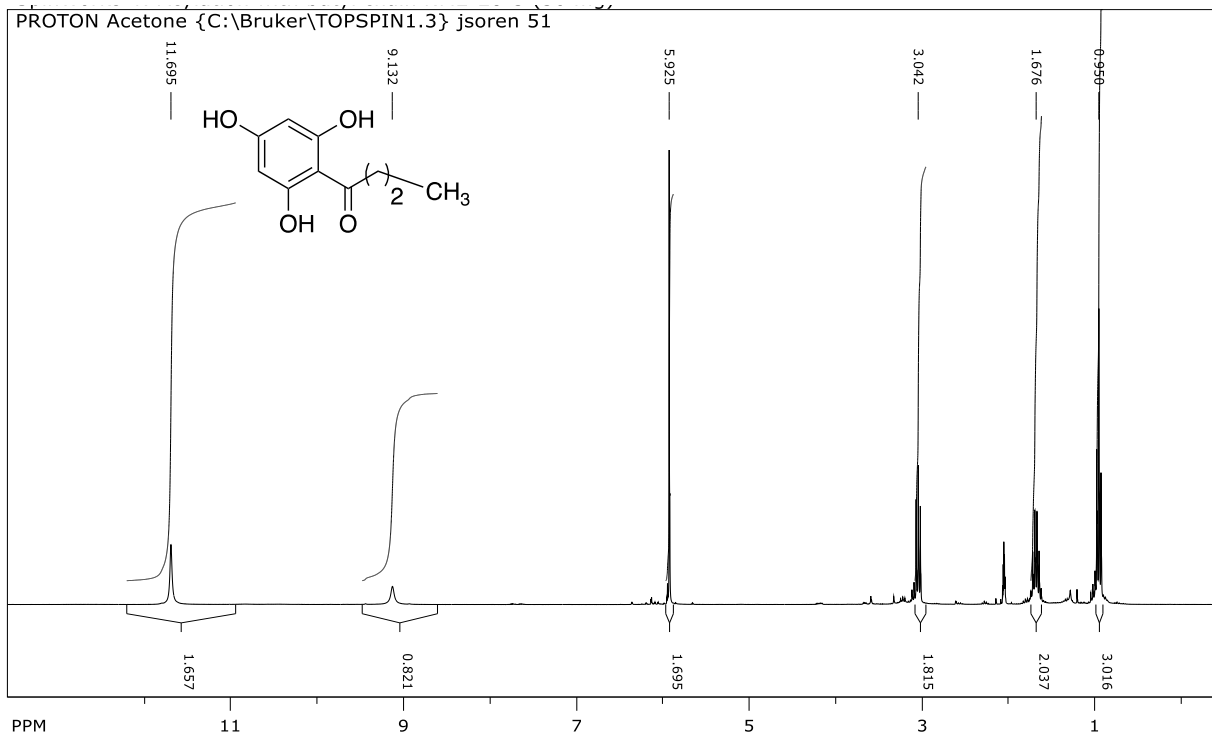
Spectrum from ...ersonal\sorensen\navrit\m-prp (9-1-2017).xms  
Scan No: 39, Time: 0.177 minutes  
No averaging. Not background corrected.  
Comment: 0.177 min. Scan: 39 150.300 (-) Ion: 11832 us RIC: 91428  
Pair Count: 1 MW: 0 Formula: None  
CAS No: None Acquired Range: 149.5 - 300.5 m/z

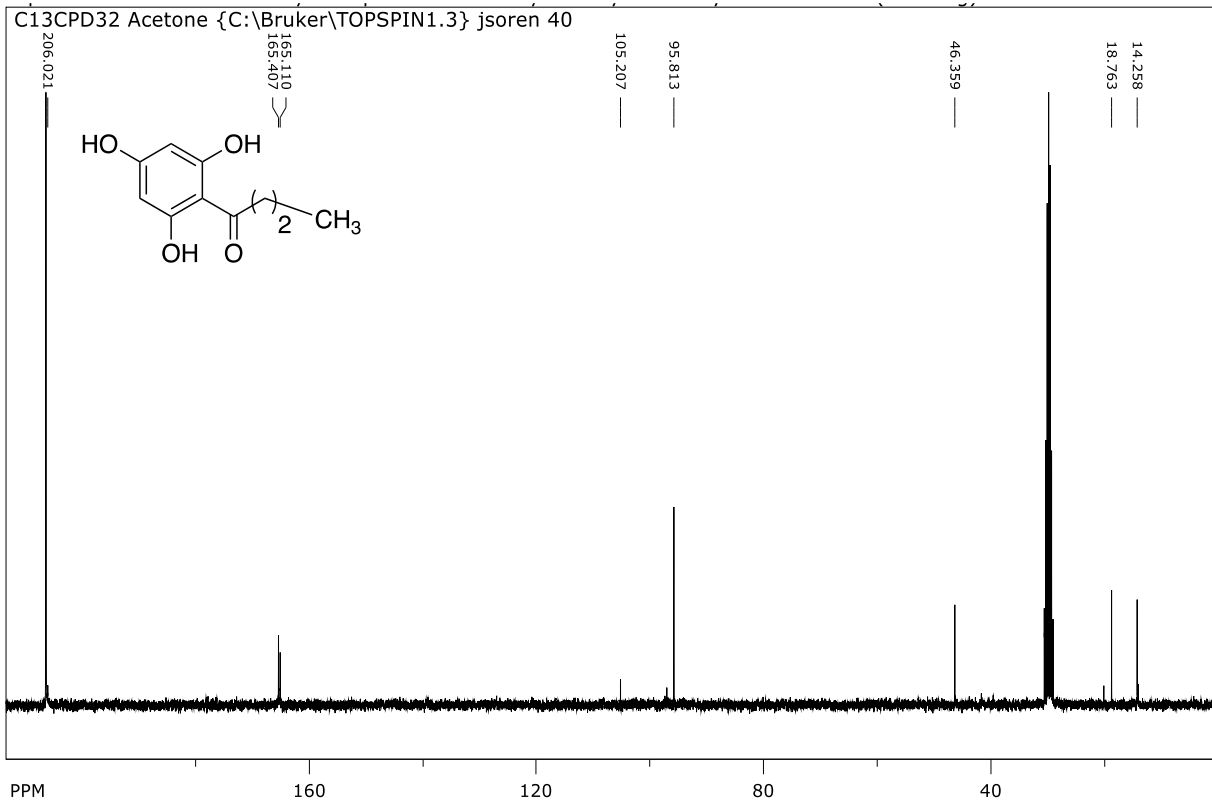
Method Time: 0.00-0.25, Centroid, Electrospray  
Seg 1, Time: 0.00-0.25. Scan Functions: 1  
1. 150.300 (-) 150.300 (-) ESI Standard 80.0[V] Full  
Product Mass Range: 149.5 - 300.5 m/z

Ion	Int Norm	Ion	Int Norm	Ion	Int Norm	Ion	Int Norm
181.3	74837.999						



## NMR of compound 6

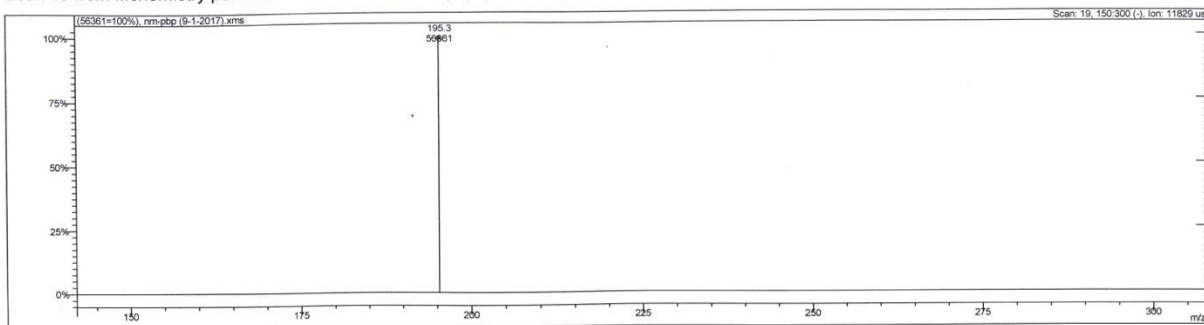




## Mass Spectra of Compound 6:

Print Date: 01 Sep 2017 13:49:17

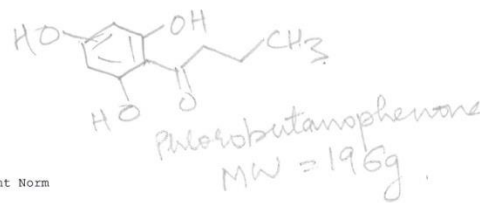
Scan 19 from ...chemistry personal\sorensen\navriti\m-pbp (9-1-2017).xms



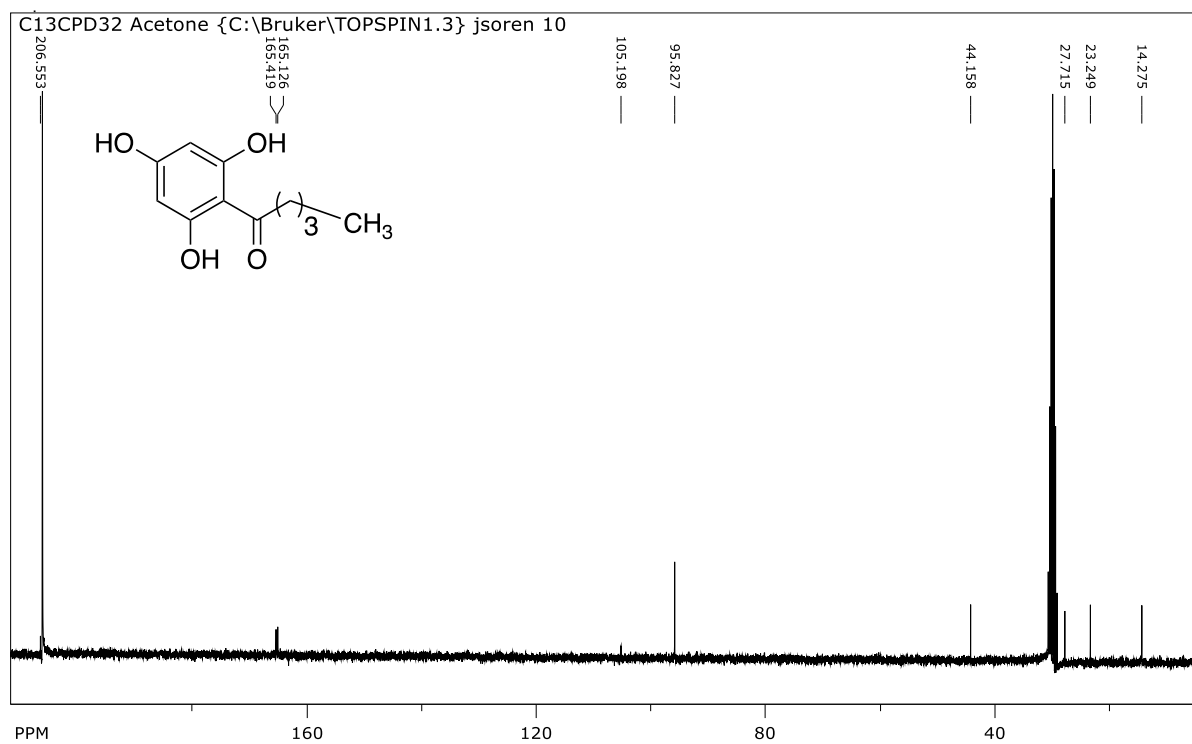
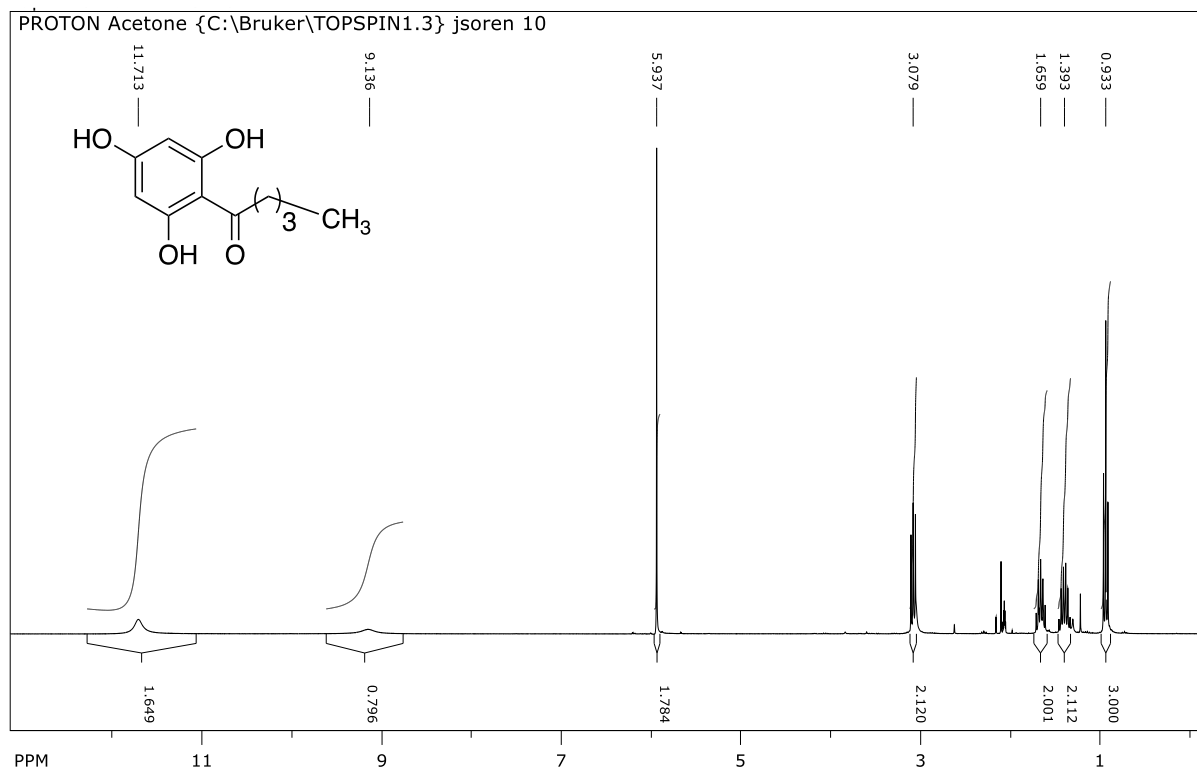
Spectrum from ...personal\sorensen\navriti\m-pbp (9-1-2017).xms  
 Scan No: 19, Time: 0.083 minutes  
 No averaging. Not background corrected.  
 Comment: 0.083 min. Scan: 19 150:300 (-) Ion: 11829 us RIC: 73893  
 Pair Count: 1 MW: 0 Formula: None  
 CAS No: None Acquired Range: 149.5 - 300.5 m/z

Method Time: 0.00-0.22, Centroid, Electrospray  
 Seg 1, Time: 0.00-0.22, Scan Functions: 1  
 1. 150:300 (-) 150:300 (-) ESI Standard 80.0[V] Full  
 Product Mass Range: 149.5 - 300.5 m/z

Ion	Int Norm	Ion	Int Norm	Ion	Int Norm
195.3	56361 999				

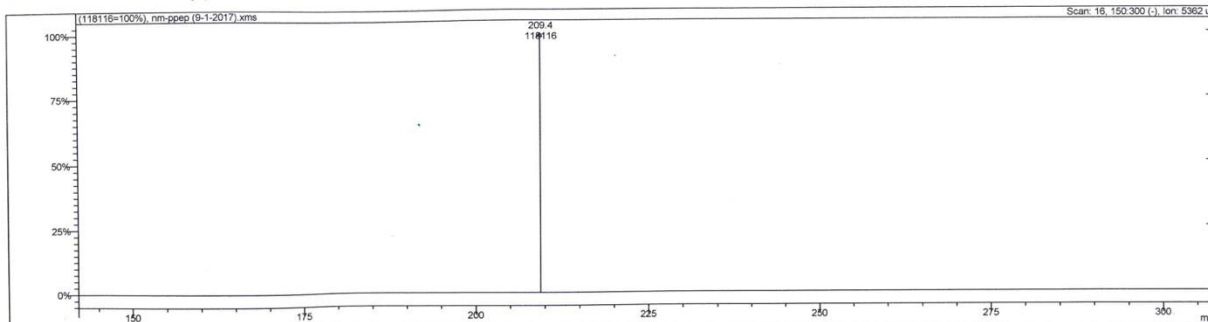


# NMR of Compound 7



## Mass Spectra of Compound 7:

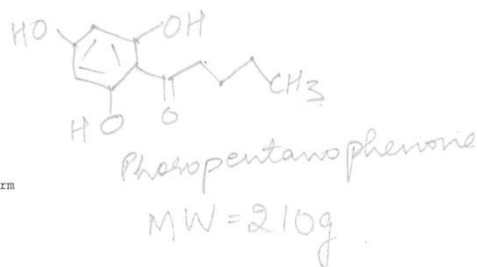
Scan 16 from ...hemistry personalsorensen\navriti\m-ppep (9-1-2017).xms



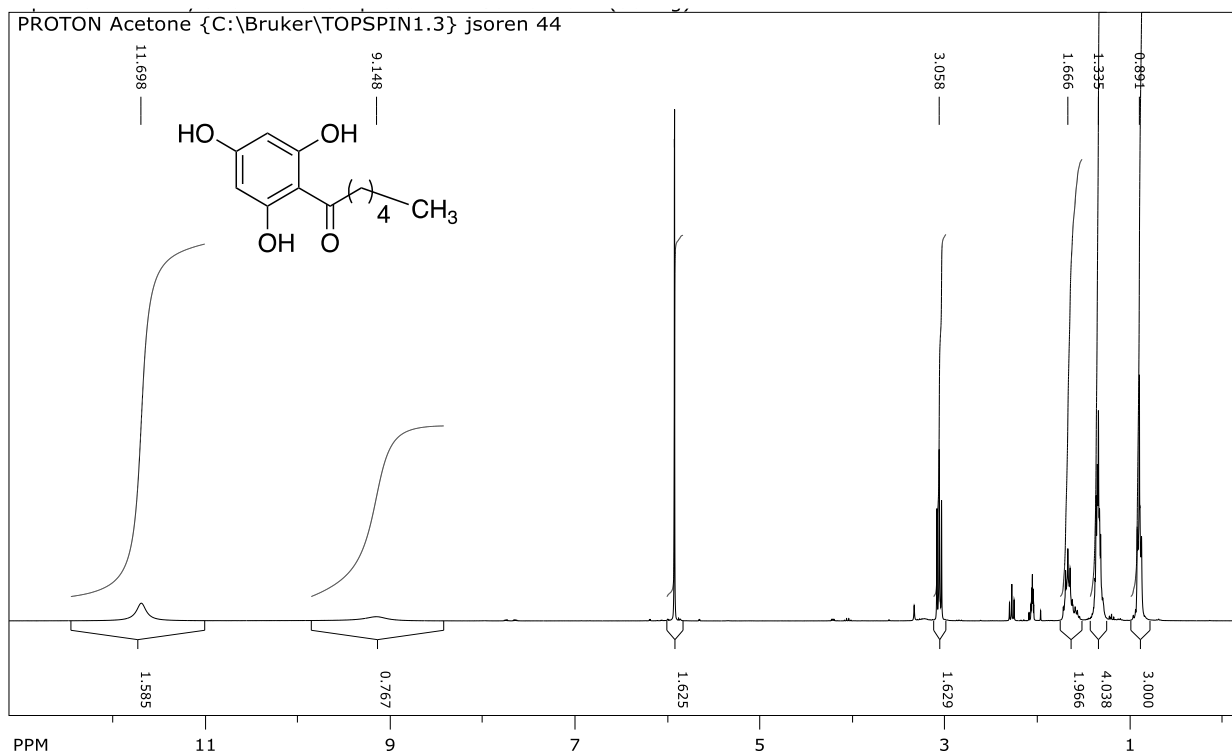
Spectrum from ...ersonalsorensen\navriti\m-ppep (9-1-2017).xms  
 Scan No: 16, Time: 0.064 minutes  
 No averaging, Not background corrected.  
 Comment: 0.064 min. Scan: 16 150:300 (-) Ion: 5362 us RIC: 176491  
 Pair Count: 1 MW: 0 Formula: None  
 CAS No: None Acquired Range: 149.5 - 300.5 m/z

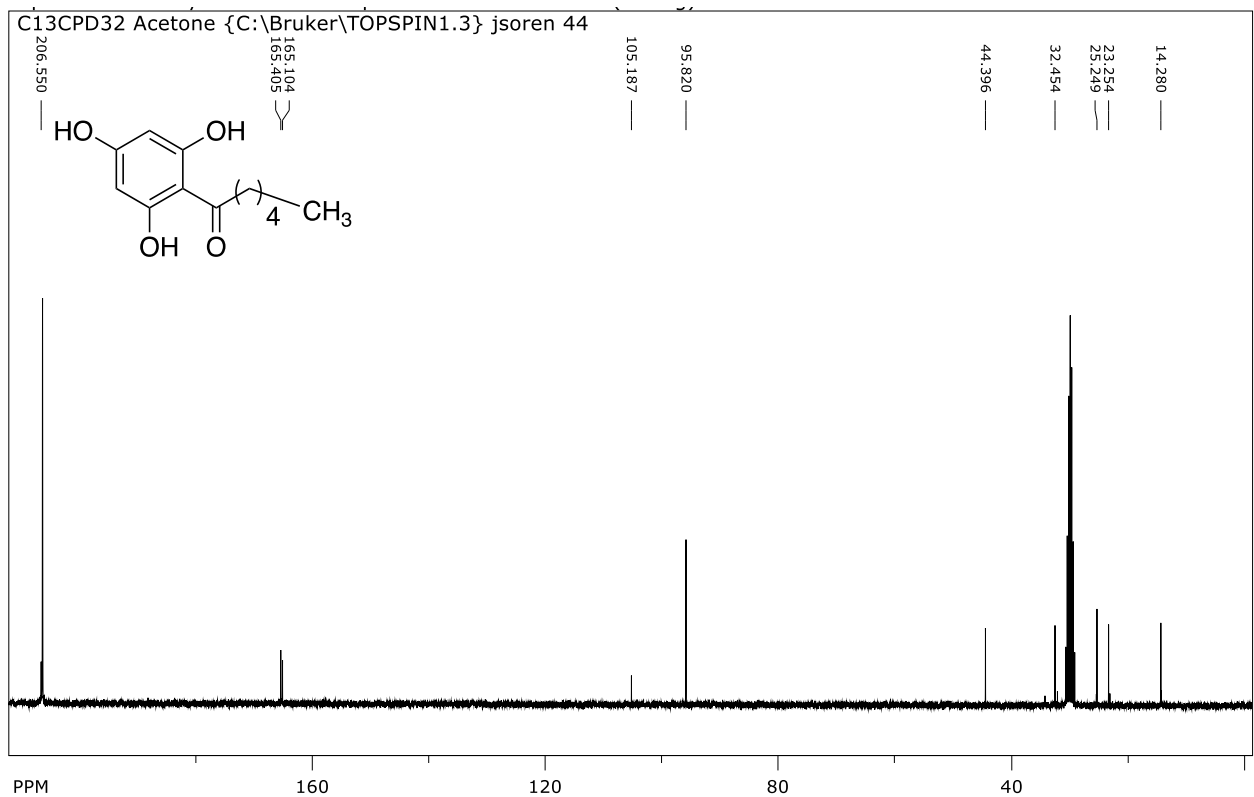
Method Time: 0.00-0.28, Centroid, Electrospray  
 Seg 1, Time: 0.00-0.28, Scan Functions: 1  
 1. 150:300 (-) 150:300 (-) ESI Standard 80.0[V] Full  
 Product Mass Range: 149.5 - 300.5 m/z

Ion	Int	Norm	Ion	Int	Norm	Ion	Int	Norm
209.4	118116	999						



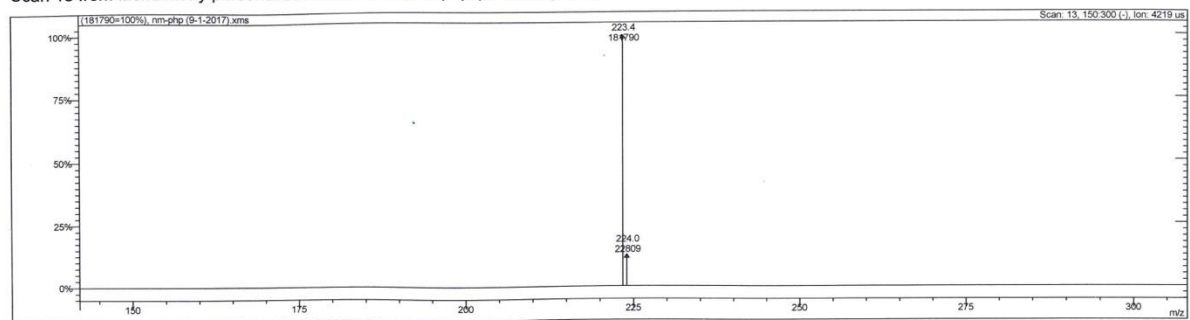
## NMR of Compound 8





### Mass Spectra of Compound 8:

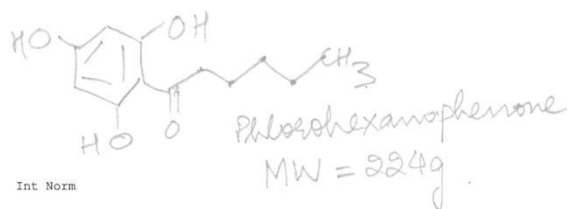
Scan 13 from ...chemistry personal\sorensen\navrit\m-pnp (9-1-2017).xms



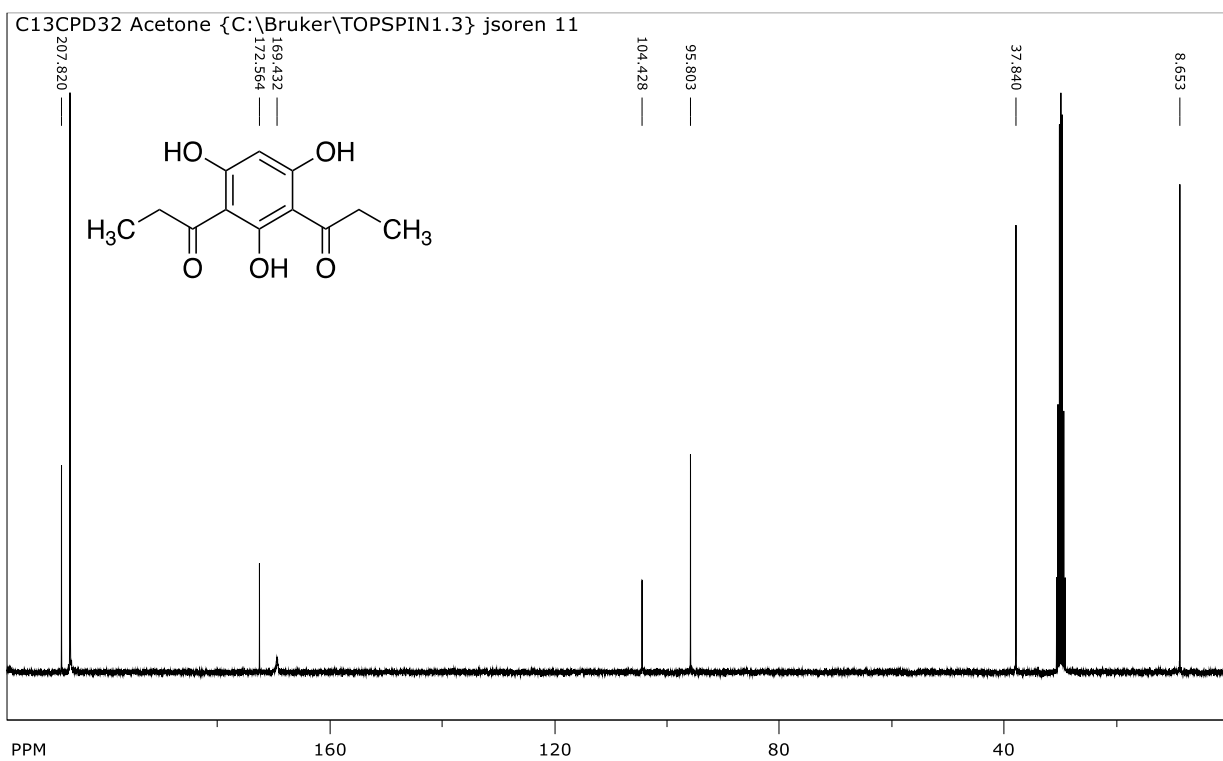
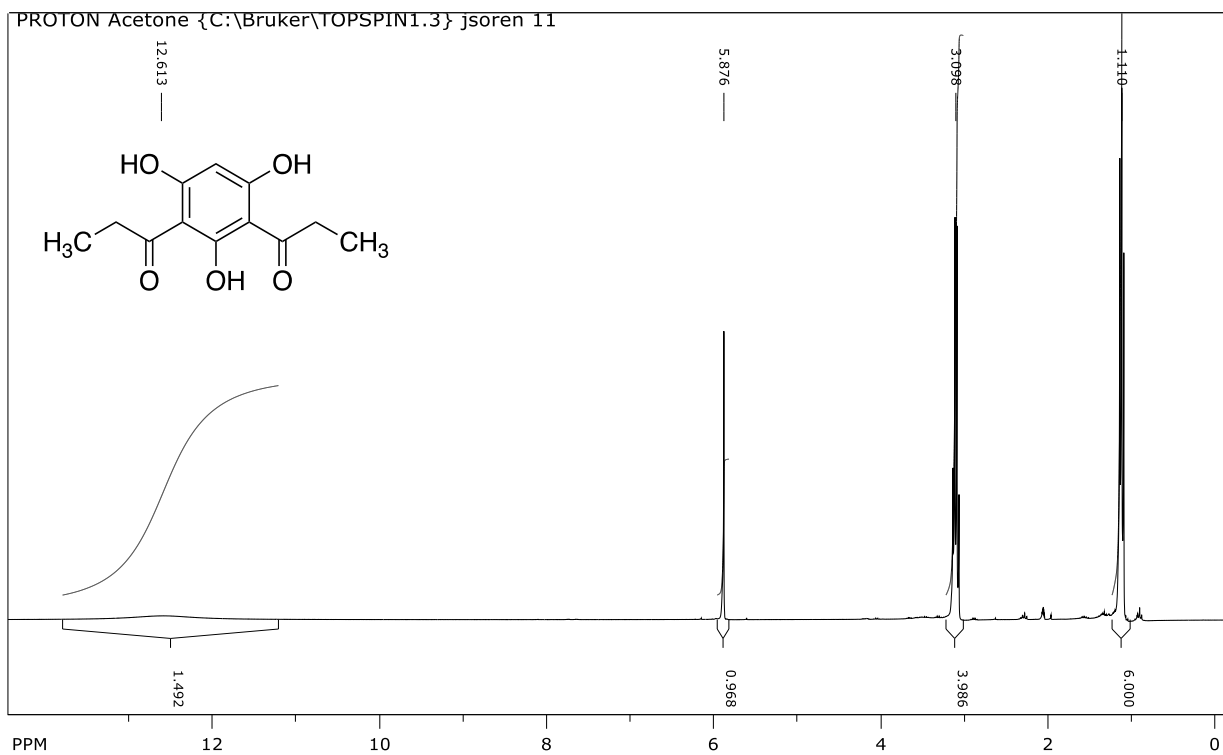
Spectrum from ...personal\sorensen\navrit\m-pnp (9-1-2017).xms  
Scan No: 13, Time: 0.050 minutes  
No averaging. Not background corrected.  
Comment: 0.050 min. Scan: 13 150:300 (-) Ion: 4219 us RIC: 248810  
Pair Count: 2 MW: 0 Formula: None  
CAS No: None Acquired Range: 149.5 - 300.5 m/z

Method Time: 0.00-0.13, Centroid, Electrospray  
Seg 1, Time: 0.00-0.13, Scan Functions: 1  
1. 150:300 (-) 150:300 (-) ESI Standard 80.0[V] Full  
Product Mass Range: 149.5 - 300.5 m/z

Ion	Int Norm	Ion	Int Norm	Ion	Int Norm	Ion	Int Norm
223.4	181790	999	224.0	22809	125		



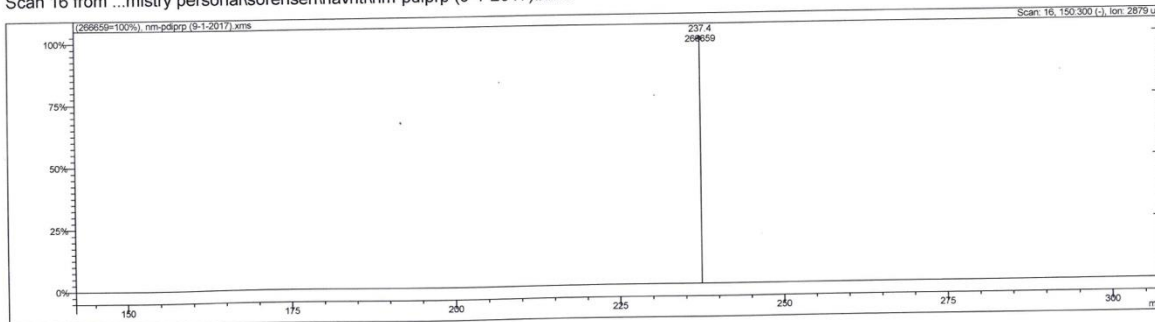
# NMR of Compound 9





## Mass Spectra of Compound 9:

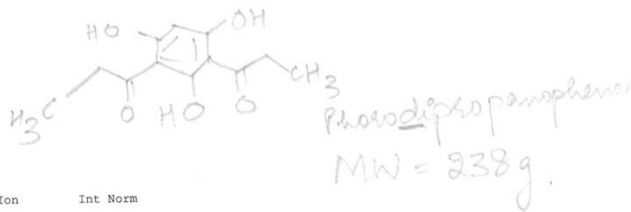
Scan 16 from ...mistry personalsorensen\navrit\m-pdiprp (9-1-2017).xms



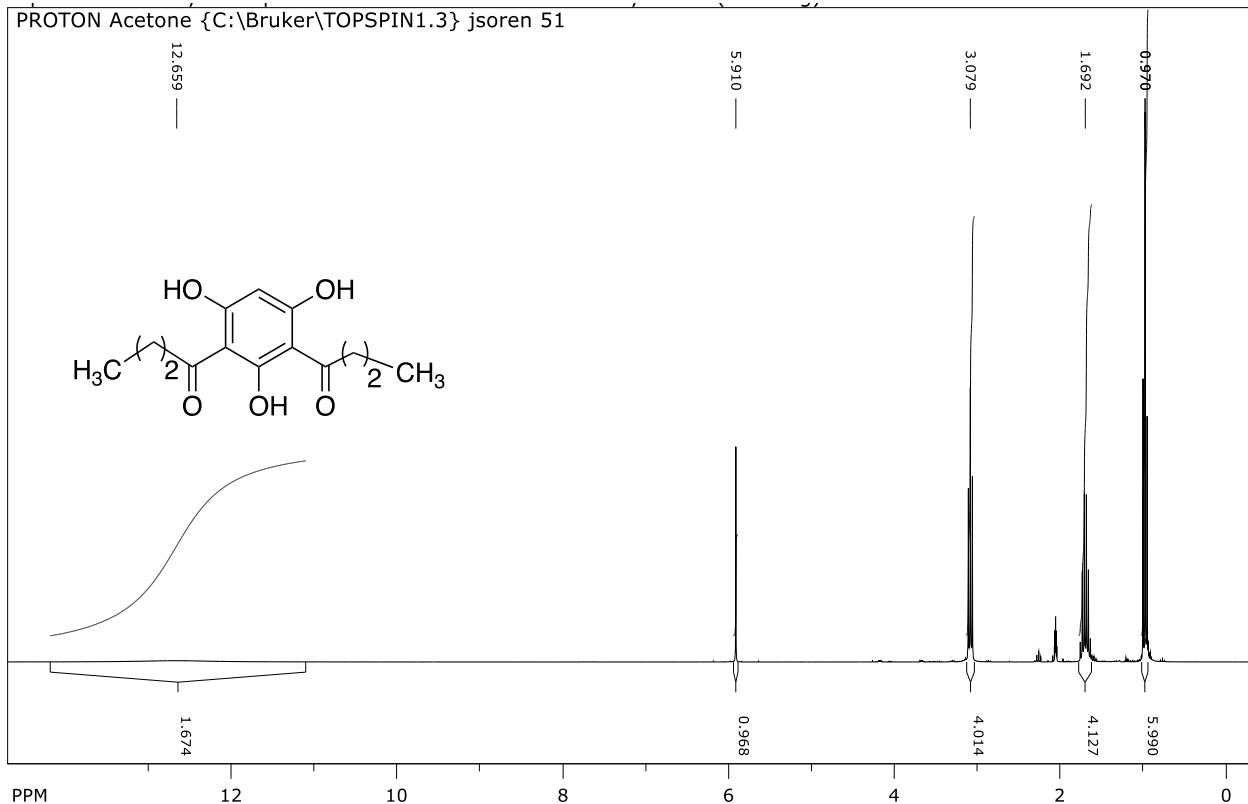
Spectrum from ...sonalsorensen\navrit\m-pdiprp (9-1-2017).xms  
 Scan No: 16, Time: 0.062 minutes  
 No averaging, Not background corrected.  
 Comment: 0.062 min. Scan: 16 150:300 (-) Ion: 2879 us RIC: 337411  
 Pair Count: 1 MW: 0 Formula: None  
 CAS No: None Acquired Range: 149.5 - 300.5 m/z

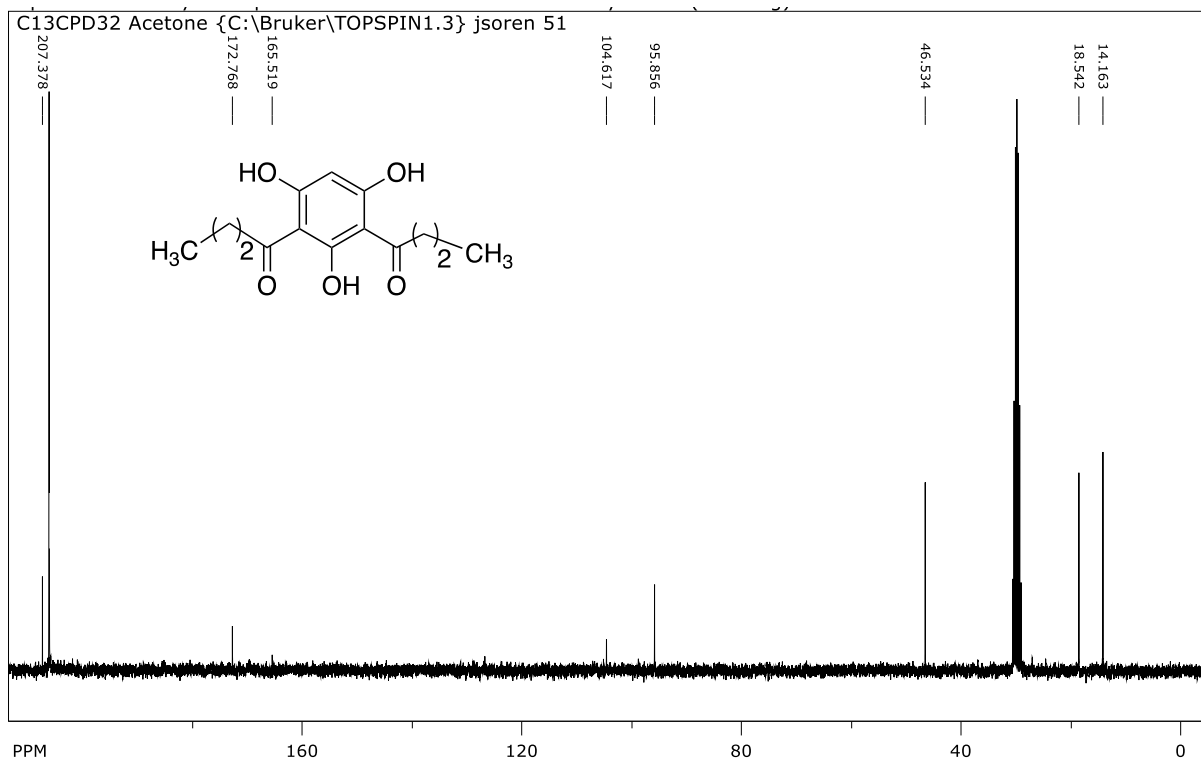
Method Time: 0.00-0.16, Centroid, Electropray  
 Seg 1, Time: 0.00-0.16, Scan Functions: 1  
 1. 150:300 (-) 150:300 (-) ESI Standard 80.0[V] Full  
 Product Mass Range: 149.5 - 300.5 m/z

Ion	Int	Norm	Ion	Int	Norm	Ion	Int	Norm
237.4	266659	999						



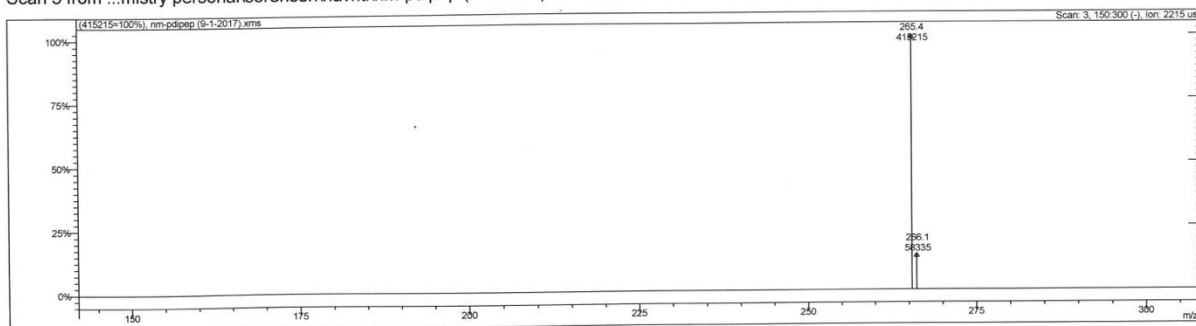
## NMR of Compound 10





### Mass Spectra of Compound 10:

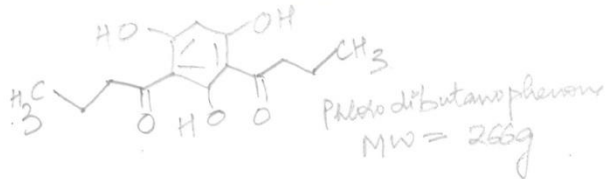
Scan 3 from ...mistry persona\sorensen\navrit\m-pdipep (9-1-2017).xms



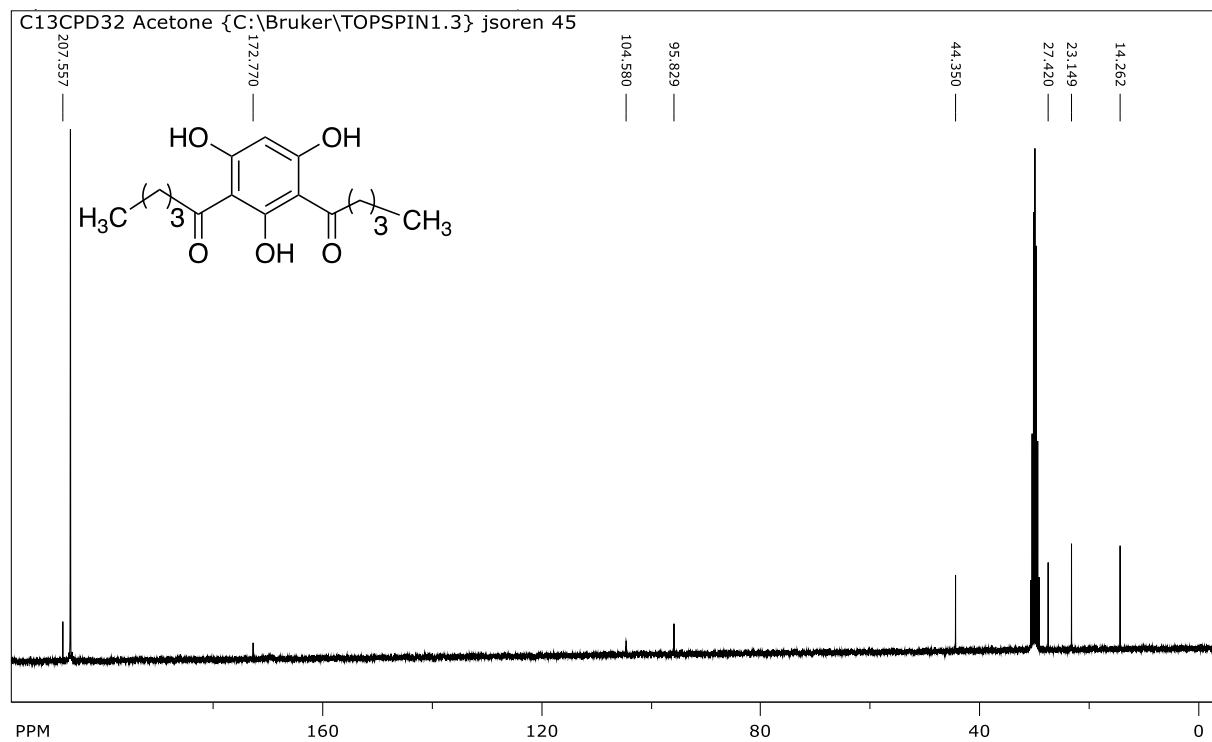
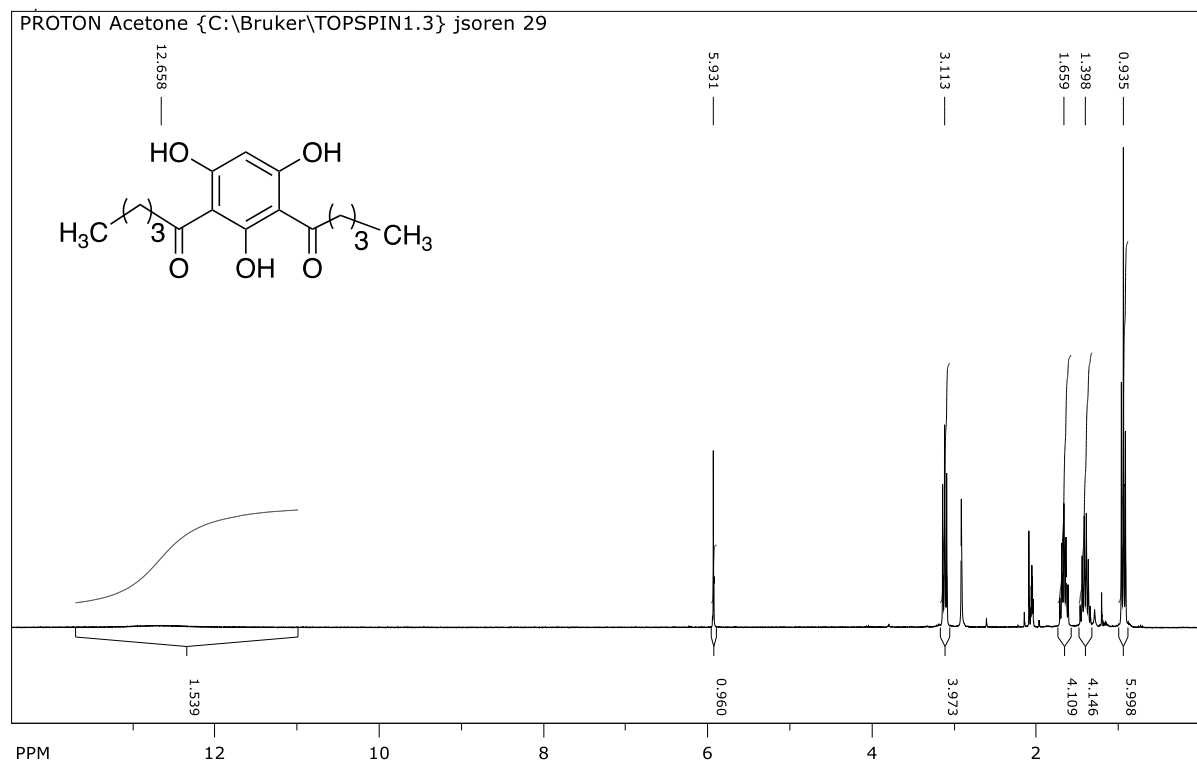
Spectrum from ...sorensen\navrit\m-pdipep (9-1-2017).xms  
Scan No: 3, Time: 0.008 minutes  
No averaging. Not background corrected.  
Comment: 0.008 min. Scan: 3 150:300 (-) Ion: 2215 us RIC: 507661  
Pair Count: 2 MW: 0 Formula: None  
CAS No: None Acquired Range: 149.5 - 300.5 m/z

Method Time: 0.00-0.15, Centroid, Electrospray  
Seg 1, Time: 0.00-0.15, Scan Functions: 1  
1, 150:300 (-) 150:300 (-) ESI Standard 80.0[V] Full  
Product Mass Range: 149.5 - 300.5 m/z

Ion	Int Norm	Ion	Int Norm	Ion	Int Norm	Ion	Int Norm
265.4	415215 999	266.1	58335 140				

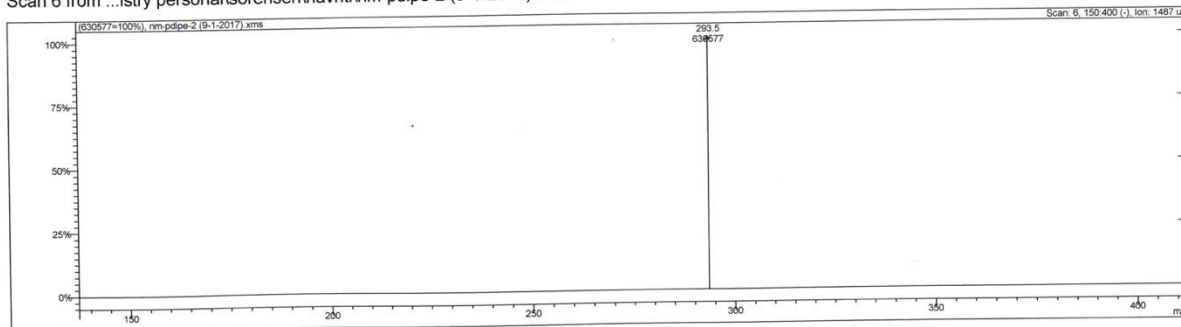


# NMR of compound **11**



## Mass Spectra of Compound 11:

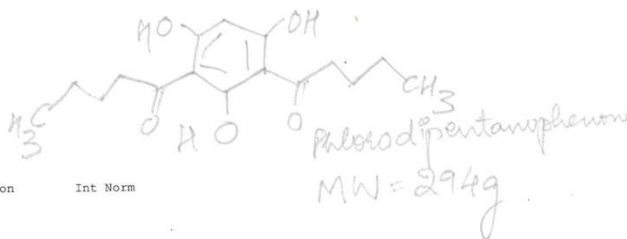
Scan 6 from ...istry personalsoreseninnavritim-pdipe-2 (9-1-2017).xms



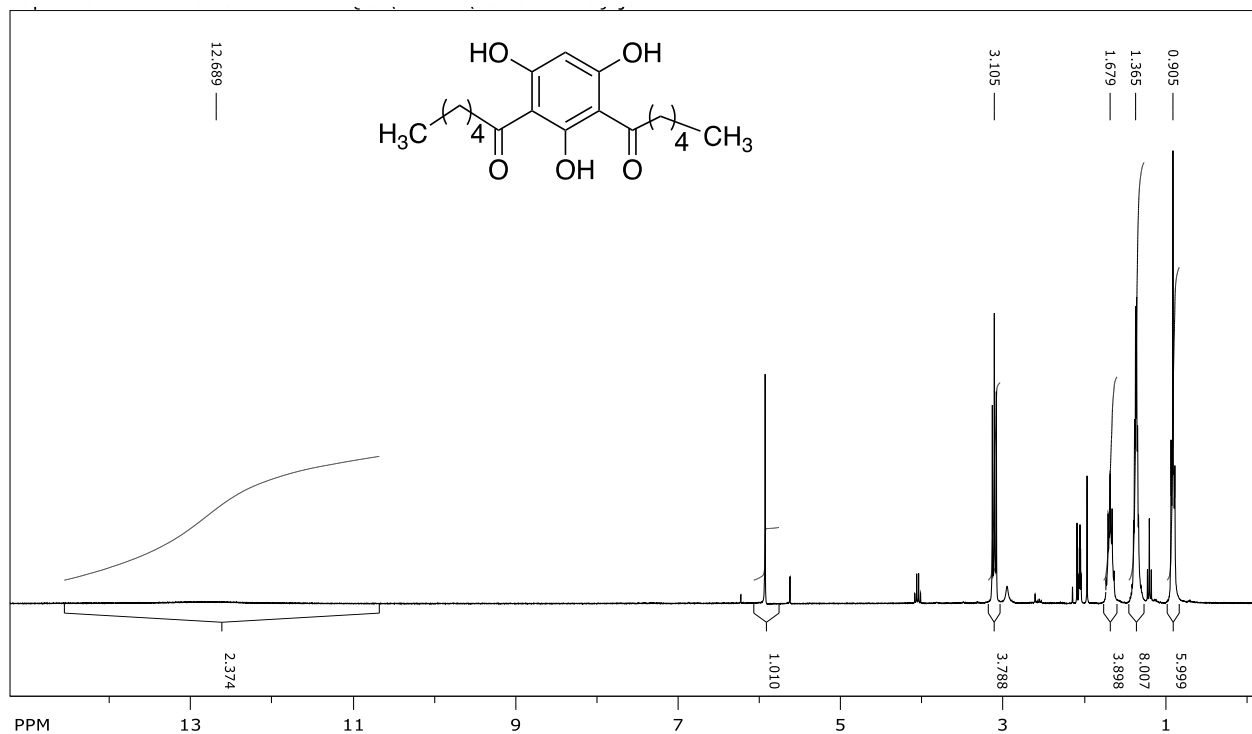
Spectrum from ...onalsoreseninnavritim-pdipe-2 (9-1-2017).xms  
 Scan No: 6, Time: 0.022 minutes  
 No averaging. Not background corrected.  
 Comment: 0.022 min. Scan: 6 150:400 (-) Ion: 1487 us RIC: 675475  
 Pair Count: 1 MW: 0 Formula: None  
 CAS No: None Acquired Range: 149.5 - 400.5 m/z

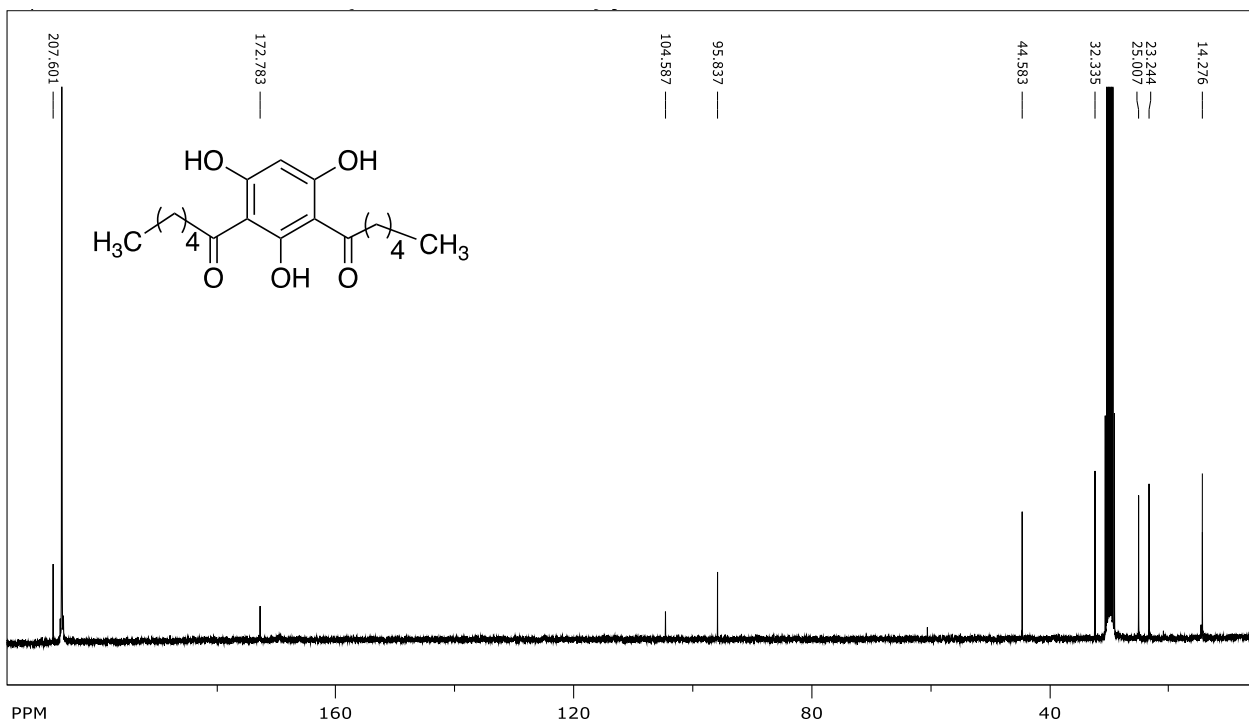
Method Time: 0.00-0.11, Centroid, Electrospray  
 Seg 1, Time: 0.00-0.11, Scan Functions: 1  
 1. 150:400 (-) 150:400 (-) ESI Standard 80.0[V] Full  
 Product Mass Range: 149.5 - 400.5 m/z

Ion	Int Norm	Ion	Int Norm	Ion	Int Norm	Ion	Int Norm
293.5	630577	999					

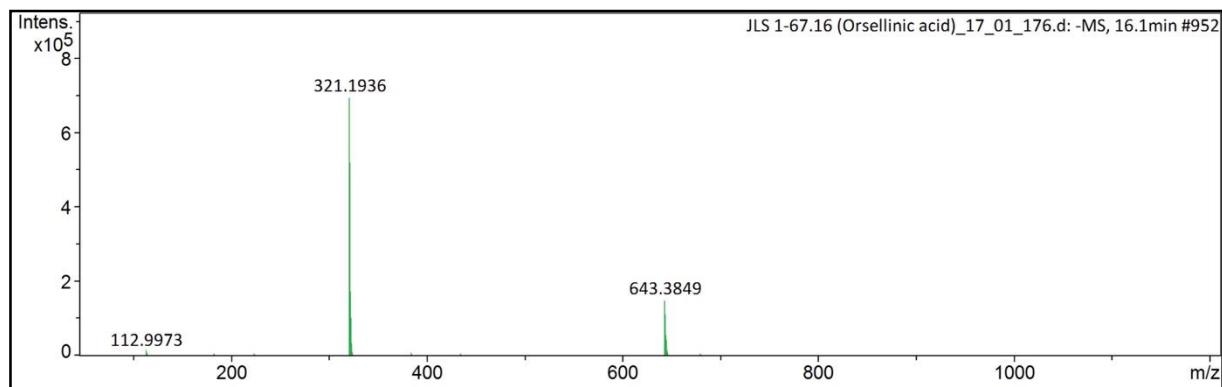


## NMR of compound 12





Mass Spectra of Compound 12:



## Chapter 6

**Table S5:** BLAST statistics for Cluster 15 in Figure 31. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retention of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30916	Hypothetical protein	<i>Pseudogymnoascus</i>	KFY65896	52/90
B	AUW30917	Thiosulfate sulfurtransferase	<i>U. pustulata</i>	SLM34737	71/91
C	AUW30918	No significant similarity	-	-	-
D	AUW30919	PLC-like phosphodiesterase	<i>P. scopiformis</i>	XP_018066270	66/92
1	AUW30920	Cytochrome P-450	<i>U. pustulata</i>	SLM34165	59/98
2	AUW30921	2-succinylbenzoate-CoA ligase	<i>S. lycopersici</i>	KNG47290	67/96
E	AUW30922	BolA-like protein	<i>U. pustulata</i>	SLM34753	80/52
F	AUW30923	GTP-binding protein Ypt5	<i>U. pustulata</i>	SLM38500	90/72
G	AUW30924	Phosphoinositide phosphatase	<i>U. pustulata</i>	SLM38501	75/97
H	AUW30925	Hypothetical protein	<i>A. oryzae</i>	XP_001821260	32/69
I	AUW30926	Nicotinate phosphoribosyltransferase	<i>U. pustulata</i>	SLM34755	61/94
J	AUW30927	No significant similarity	-	-	-
K	AUW30928	Hypothetical protein	<i>Pseudogymnoascus</i>	KFX93959	63/61
L	AUW30929	Hypothetical protein	<i>Fungal sp. 11243</i>	GAM88438	57/74
M	AUW30930	Translation initiation factor	<i>X. heveae</i>	XP_018189790	60/78
N	AUW30931	No significant similarity	-	-	-
O	AUW30932	Hypothetical protein	<i>M. phaseolina</i>	EKG13067	30/76

P	AUW30933	Hypothetical protein	<i>C. gloeosporioides</i>	EQB59441	34/77
Q	AUW30934	Hypothetical protein	<i>P. chrysogenum</i>	KZN85741	37/81
3	AUW30935	Short-chain dehydrogenase / reductase	<i>U. pustulata</i>	SLM34748	73/98
R	AUW30936	No significant similarity	-	-	-
S	AUW30937	No significant similarity	-	-	-
T	AUW30938	40S ribosomal protein	<i>P. brasilianum</i>	CEJ55024	91/83
U	AUW30939	UPF0183 domain protein	<i>P. brasilianum</i>	CEJ55020	49/98
V	AUW30940	Isoleucyl-tRNA synthetase	<i>T. stipitatus</i>	XP_002479535	51/98
W	AUW30941	No significant similarity	-	-	-
X	AUW30942	No significant similarity	-	-	-
Y	AUW30943	Hypothetical protein	<i>B. cinerea</i>	CCD51123	58/99
4	AUW30944	4-coumarate-CoA ligase	<i>E. dermatitidis</i>	XP_009156000	47/79
		AMP-dependent synthetase / ligase (not NRPS-like)	<i>P. camemberti</i>	CRL23284	47/80
5	AUW30945	Phenylalanine ammonia lyase	<i>L. vulpine</i>	BAN29055	64/94
Z	AUW30946	Opsin-1	<i>P. attae</i>	KPI45543	70/98
6	AUW30947	Threonine dehydratase	<i>B. spectabilis</i>	GAD96042	74/97
AA	AUW30948	Hypothetical protein	<i>M. brunnea</i>	XP_007288818	47/67
7	AUW30949	Multicopper oxidase	<i>P. chlamydsospora</i>	KKY23553	53/88
		Laccase-2	<i>Fusarium fujikuroi</i>	KLO79025	51/90
8	AUW30950	Cytochrome P-450	<i>E. lata</i>	XP_007795124	38/98
9	AUW30951	Squalene cyclase <i>(Cu-terp-4)</i>	<i>A. oryzae</i>	EIT72300	53/98

---

**Table S6:** BLAST statistics for Cluster 17 in Figure 31. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW30975	Molecular chaperone, heat shock protein Hsp40, DnaJ	<i>P. camemberti</i>	CRL27427	74/94
B	AUW30976	Hypothetical protein	<i>T. marneffeii</i>	XP_002143165	33/81
1	AUW30977	Cytochrome P-450	<i>A. flavus</i>	XP_002378813	40/92
2	AUW30978	Farnesyl transferase <i>(Cu-terp-6)</i>	<i>S. borealis</i>	ESZ91924	70/100
C	AUW30979	Small nucleolar ribonucleo- protein complex subunit	<i>A. fumigatus</i>	KMK57755	45/99
D	AUW30980	G1/S-specific cyclin Cln1	<i>A. oryzae</i>	XP_001818433	62/89
E	AUW30981	Cysteine protease	<i>C. higginsianum</i>	CCF41055	27/82
F	AUW30982	GPR/FUN34 family protein	<i>T. marneffeii</i>	XP_002150578	72/95
G	AUW30983	Png1p protein	<i>U. reesii</i>	XP_002542827	67/92



**Table S7:** BLAST statistics for Cluster 20 in Figure 31. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW31007	Cytochrome P-450	<i>P. rogueforti</i>	CDM31319	44/96
2	AUW31008	Short-chain dehydrogenase / reductase	<i>C. gloeosporioides</i>	XP_007286745	64/100
3	AUW31009	Cytochrome P-450	<i>M. phaseolina</i>	EKG12700	62/93
4	AUW31010	Aristolochene synthase ( <i>Cu-terp-9</i> )	<i>A. ustus</i>	KIA75982	59/99
5	AUW31011	Cytochrome P-450  Transferase	<i>C. fioriniae</i>  <i>P. expansum</i>	XP_007596781  KGO70752	35/98  41/98
6	AUW31012	Type I PKS	<i>P. membranacea</i>	AEE65373	53/99

**Table S8:** BLAST statistics for Cluster 24 in Figure 35. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31111	No significant similarity	-	-	-
B	AUW31112	Serine  hydroxymethyltransferase	<i>U. pustulata</i>	SLM37171	86/93
C	AUW31113	Hypothetical protein	<i>A. candida</i>	CCI43743	29/73
D	AUW31114	Hypothetical protein	<i>B. spectabilis</i>	GAD95252	31/62
E	AUW31115	Hypothetical protein	<i>U. pustulata</i>	SLM33372	41/65

F	AUW31116	Hypothetical protein	<i>G. lozoyensis</i>	XP_008078674	46/70
G	AUW31117	No significant similarity	-	-	-
H	AUW31118	Six-hairpin glycosidase	<i>G. lozoyensis</i>	XP_008083366	51/95
I	AUW31119	No significant similarity	-	-	-
J	AUW31120	No significant similarity	-	-	-
1	AUW31121	PKS ( <i>Cu-nr-pks-4</i> )	<i>S. borealis</i>	ESZ91464	58/99
2	AUW31122	Cytochrome P-450	<i>V. dahliae</i>	XP_009654521	38/97
K	AUW31123	Sugar transporter	<i>T. reesei</i>	ETR99312	63/91
L	AUW31124	No significant similarity	-	-	-
M	AUW31125	Ribonuclease	<i>N. parvum</i>	XP_007580708	62/83
N	AUW31126	Caspase-like protein	<i>G. lozoyensis</i>	XP_008086178	84/71
O	AUW31127	Isopenicillin N synthetase	<i>B. spectabilis</i>	GAD93339	61/91
		Clayaminate synthase-like protein	<i>G. lozoyensis</i>	XP_008077427	61/88
		1-aminocyclopropane-1-carboxylate oxidase	<i>P. attae</i>	XP_018002059	63/88
		Oxoglutarate/iron-dependent dioxygenase	<i>P. accitanis</i>	PCH03280	60/87
		Gibberellin 20-oxidase	<i>R. emersonii</i>	XP_013327414	58/90
		Thymine dioxygenase	<i>D. corticola</i>	XP_020130585	58/88
P	AUW31128	Glycoside hydrolase	<i>L. palustris</i>	OCK80932	58/92
Q	AUW31129	Hypothetical protein	<i>E. spinifera</i>	XP_016237774	41/66
R	AUW31130	No significant similarity	-	-	-
S	AUW31131	Hypothetical protein	<i>P. sporulosa</i>	XP_018031724	61/60

T	AUW31132	Hypothetical protein	<i>P. arizonense</i>	XP_022483471	40/82
---	----------	----------------------	----------------------	--------------	-------

**Table S9:** BLAST statistics for Cluster 26 in Figure 35. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31148	Ribosome biogenesis protein	<i>E. pusillum</i>	XP_007802122	87/100
B	AUW31149	No significant similarity	-	-	-
C	AUW31150	Hypothetical protein	<i>C. europaea</i>	XP_008718010	50/68
D	AUW31151	Hypothetical protein	<i>G. lozoyensis</i>	XP_008078151	38/95
1	AUW31152	PKS ( <i>Cu-nr-pks-6</i> )	<i>C. macilenta</i>	AFB81352	94/99
E	AUW31153	Transcription factor	<i>C. lagenaria</i>	BAE98094	32/97
2	AUW31154	Short-chain dehydrogenase / reductase	<i>M. phaseolina</i>	EKG20376	68/80
F	AUW31155	No significant similarity	-	-	-
3	AUW31156	Scytalone dehydratase	<i>E. aquamarina</i>	XP_013257419	70/92
4	AUW31157	Monoxygenase	<i>R. secalis</i>	CZT49571	31/87
G	AUW31158	Chromate transport protein	<i>P. secalis</i>	CZT47223	55/98
H	AUW31159	No significant similarity	-	-	-
I	AUW31160	Hypothetical protein	<i>G. stellatum</i>	OCL07820	37/69
J	AUW31161	MFS transporter	<i>P. roqueforti</i>	CDM36445	52/73
5	AUW31162	Cytochrome P-450	<i>T. stipitatus</i>	XP_002341512	53/95
K	AUW31163	Hypothetical protein	<i>A. fumigatus</i>	KMK54740	51/66
L	AUW31164	GABA permease	<i>A. fumigatus</i>	EDP50382	65/91

M	AUW31165	No significant similarity	-	-	-
---	----------	---------------------------	---	---	---

**Table S10:** BLAST statistics for Cluster 32 in Figure 35. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

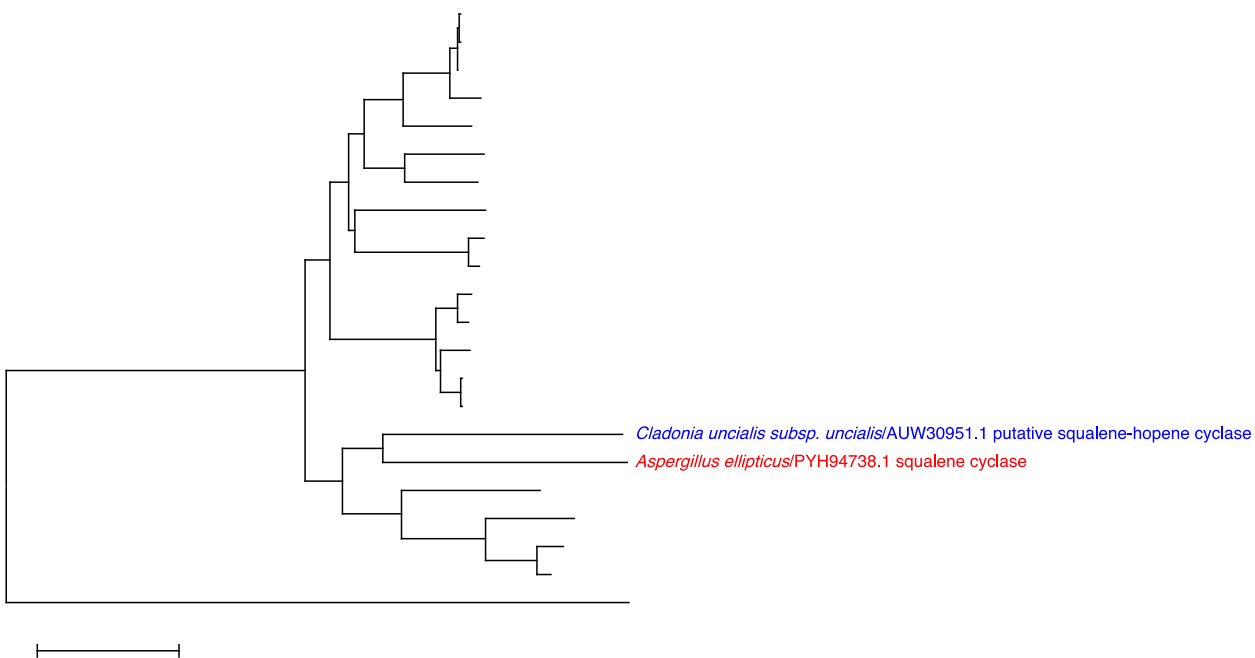
Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
A	AUW31235	Kinesin	<i>S. borealis</i>	ESZ93150	65/97
B	AUW31236	Tetrahydrofolylpolyglutamate synthase	<i>B. spectabilis</i>	GAD96053	62/98
		Tetrahydrofolate synthase	<i>R. emersonii</i>	XP_013329221	63/97
C	AUW31237	Ribosome biogenesis protein	<i>C. posadasii</i>	EFW17846	62/98
D	AUW31238	GPI transamidase component	<i>A. otae</i>	XP_002845857	64/100
1	AUW31239	GNAT acetyltransferase	<i>B. spectabilis</i>	GAD94037	49/88
2	AUW31240	PKS ( <i>Cu-nr-pks-12</i> )	<i>C. grayi</i>	ADM79462	82/99
3	AUW31241	Cytochrome P-450	<i>T. islandicus</i>	CRG89931	45/68
E	AUW31242	No significant similarity	-	-	-

**Table S11:** BLAST statistics for Cluster 48 in Figure 41. Reproduced from Bertrand et al. (2018a), Supporting Information file, in accordance with authors' retainment of privileges.

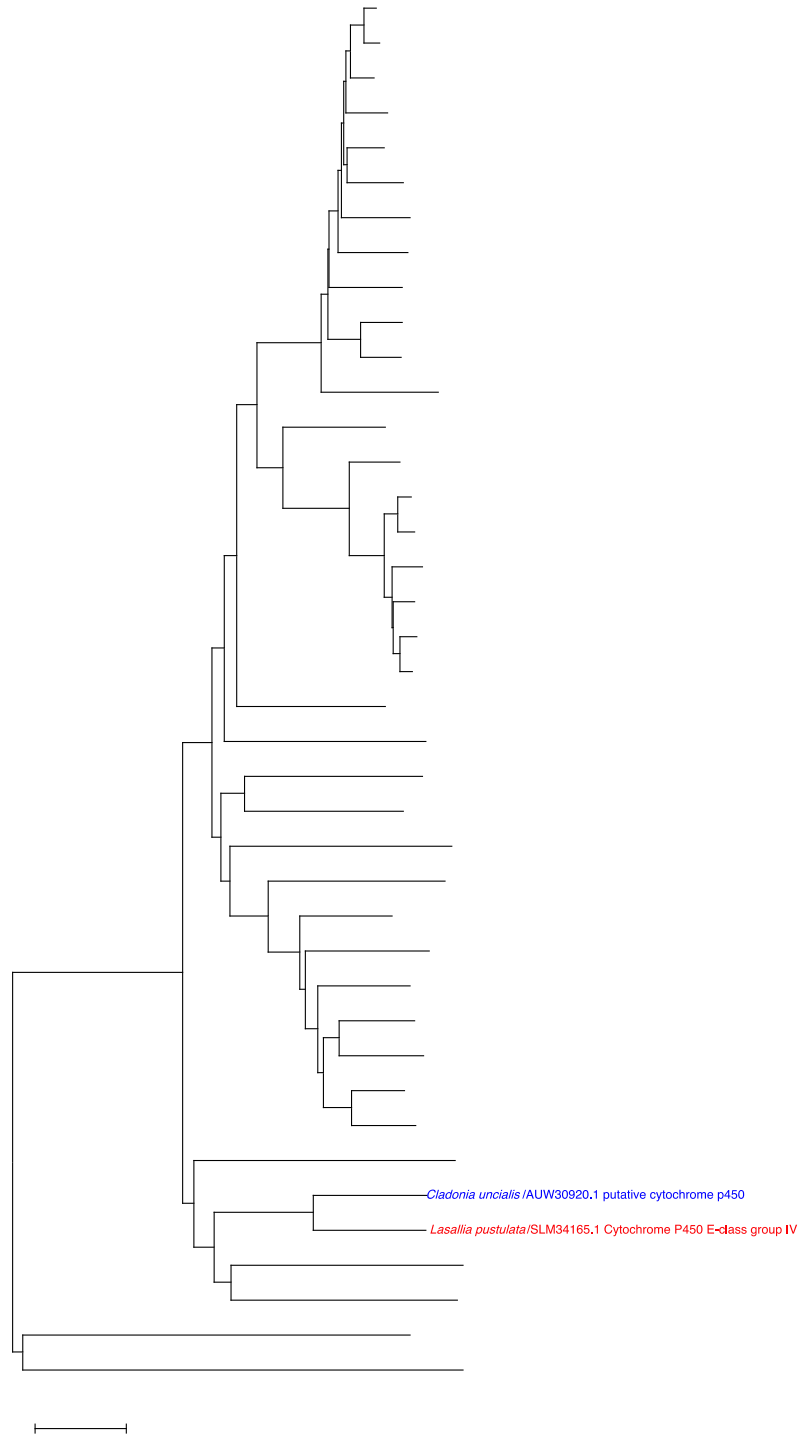
Gene (No.)	Protein Accession No.	Protein Putative function(s)	Species Closest homolog	Homolog Accession No.	Identity (%) / Coverage (%)
1	AUW31423	PKS ( <i>Cu-r-pks-18</i> )	<i>T. islandicus</i>	CRG92723	75/100
2	AUW31424	O-methyltransferase	<i>M. mycetomatis</i>	KOP50440	85/99

3	AUW31425	Short-chain dehydrogenase / reductase	<i>G. clavigera</i>	XP_014172074	68/99
4	AUW31426	Cytochrome P-450	<i>T. islandicus</i>	CRG92717	79/97
5	AUW31427	FAD oxidase	<i>M. mycetomatis</i>	KOP50437	71/98

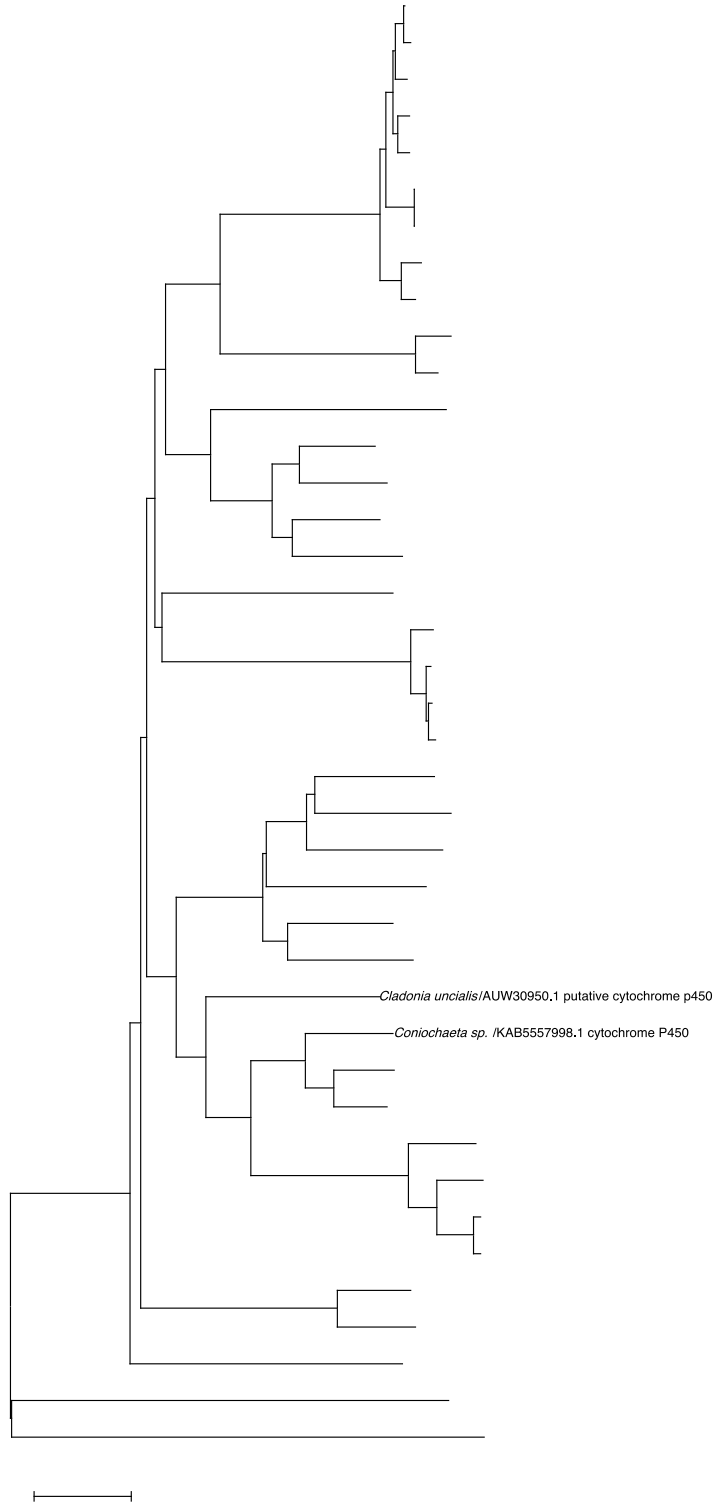
---



**Figure S1:** Phylogenetic relationship between a putative *Squalene cyclase* of *Cladonia uncialis* and a genetically similar gene encoding a *squalene cyclase* of *Aspergillus ellipticus*. The *Squalene cyclase* of *Alkalihalobacillus pseudofirmus* (ADC50961.1) was chosen as an out-group.



**Figure S2:** Phylogenetic relationship between a putative *Cytochrome p450* of *Cladonia uncialis* and a genetically similar gene encoding a *cytochrome p450 E-class group IV* of *Rhizodiscina lignyota*. The *cytochrome p450 (14-alpha demethylase)* of *Aspergillus fumigatus* (XP\_749134.1) and *cytochrome p450* of *Homo sapiens* were chosen as out-groups.

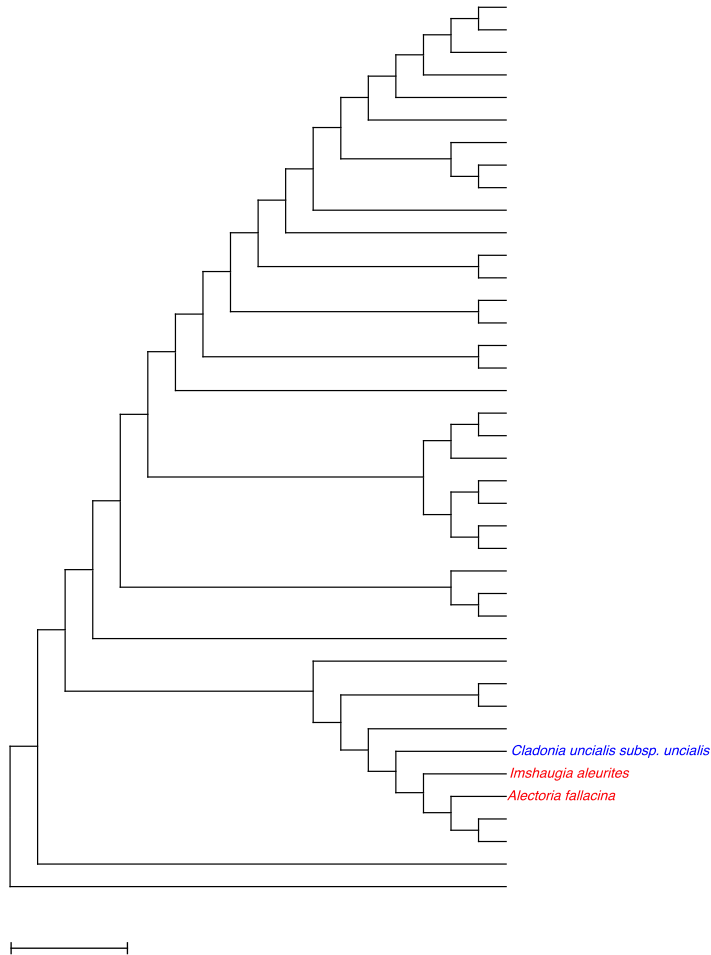


**Figure S3:** Phylogenetic relationship between a putative *Cytochrome p450* of *Cladonia uncialis* (*Cu-cyp2-terp-4*) and a genetically similar gene encoding a *CYP* of *Coniochaeta sp.* The *cytochrome p450* (*14-alpha demethylase*) of *Aspergillus fumigatus* (XP\_749134.1) and *cytochrome p450* of *Homo sapiens* were chosen as out-groups.

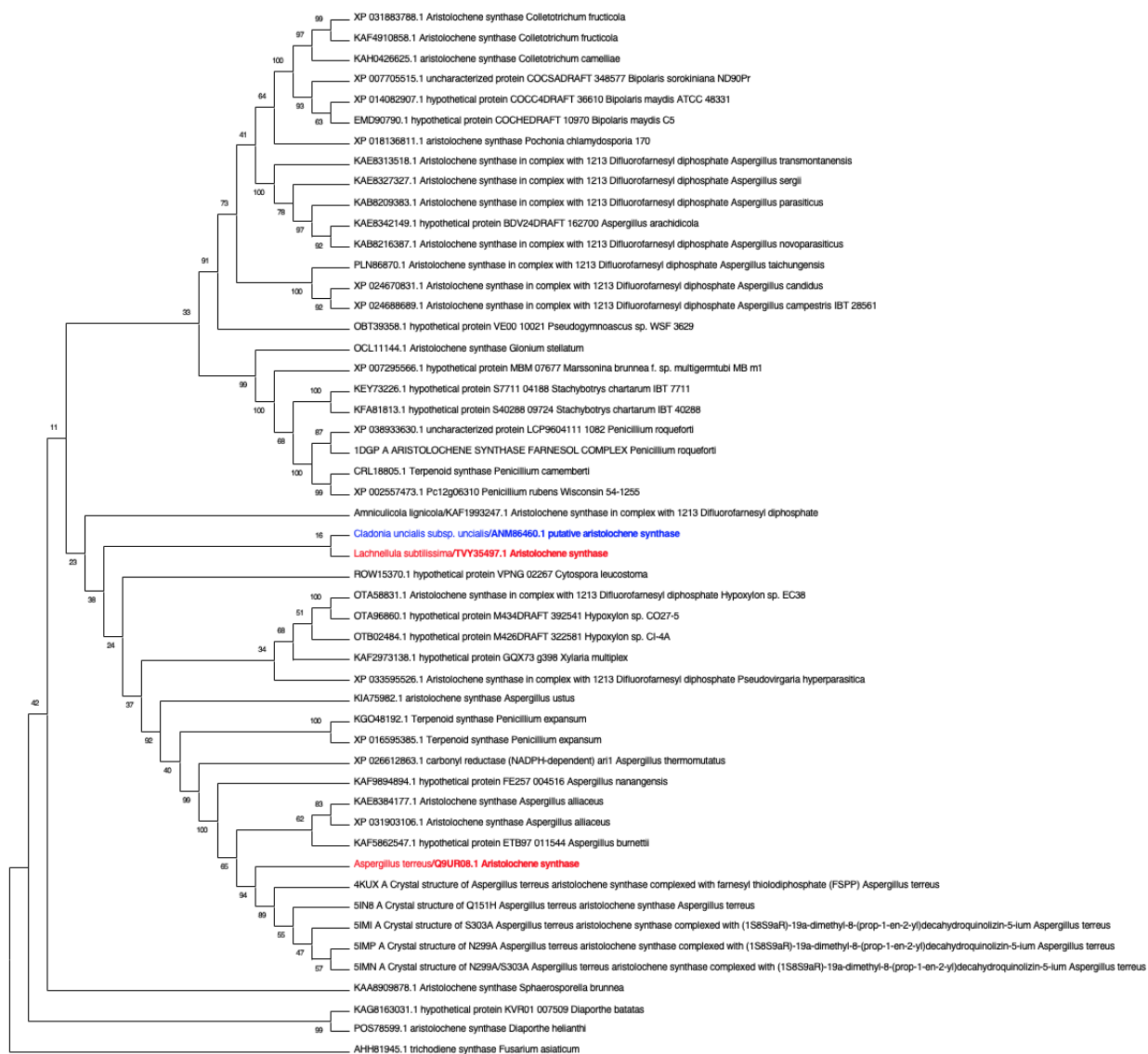




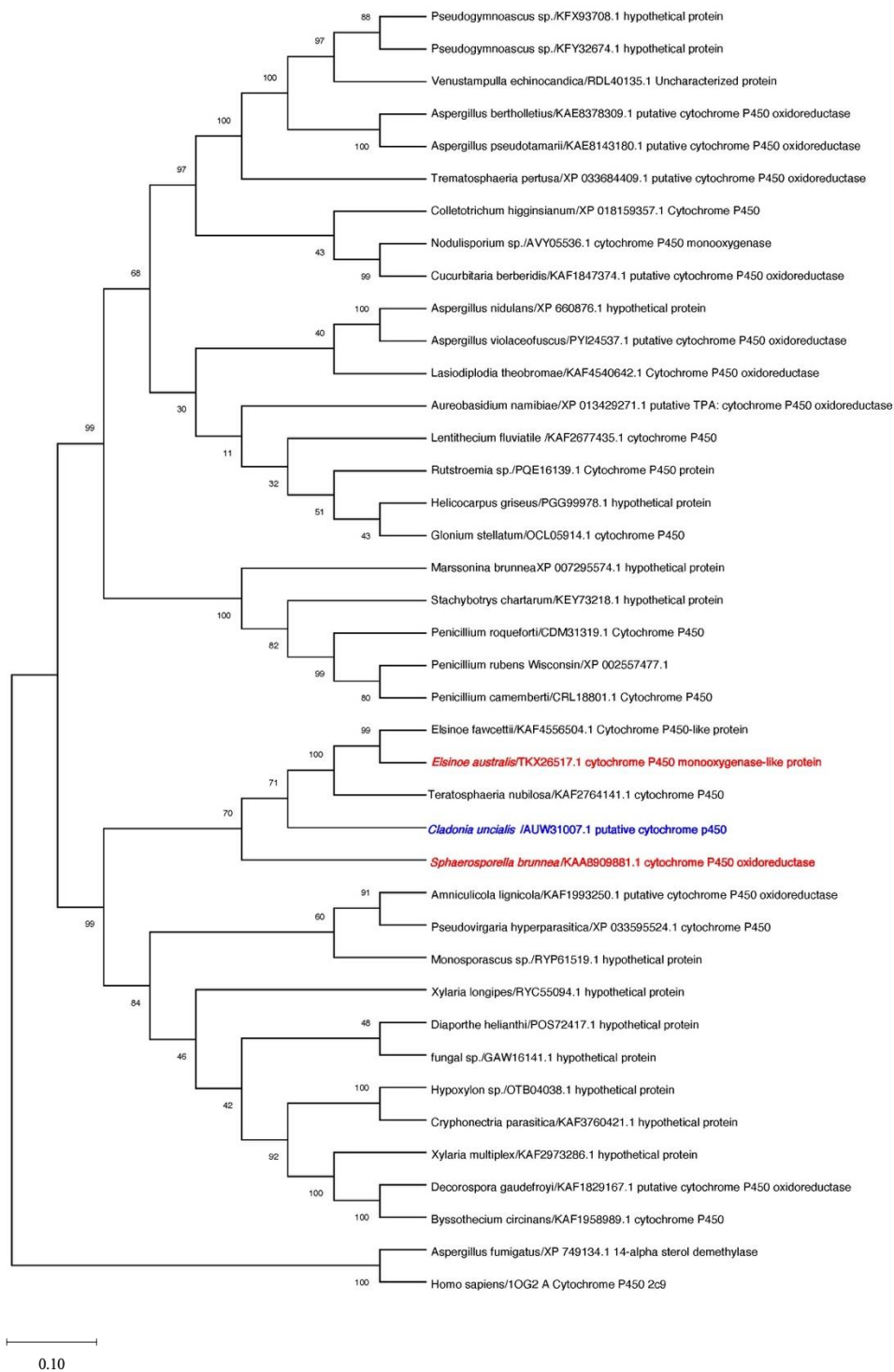
**Figure S4:** Phylogenetic relationship between a putative **Cytochrome p450** of *Cladonia uncialis* (*Cu-cyp3-terp-6*) and a genetically similar gene encoding a **CYP** of *Endocarpon gypsea*. The cytochrome p450 (14-alpha demethylase) of *Aspergillus fumigatus* (XP\_749134.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.



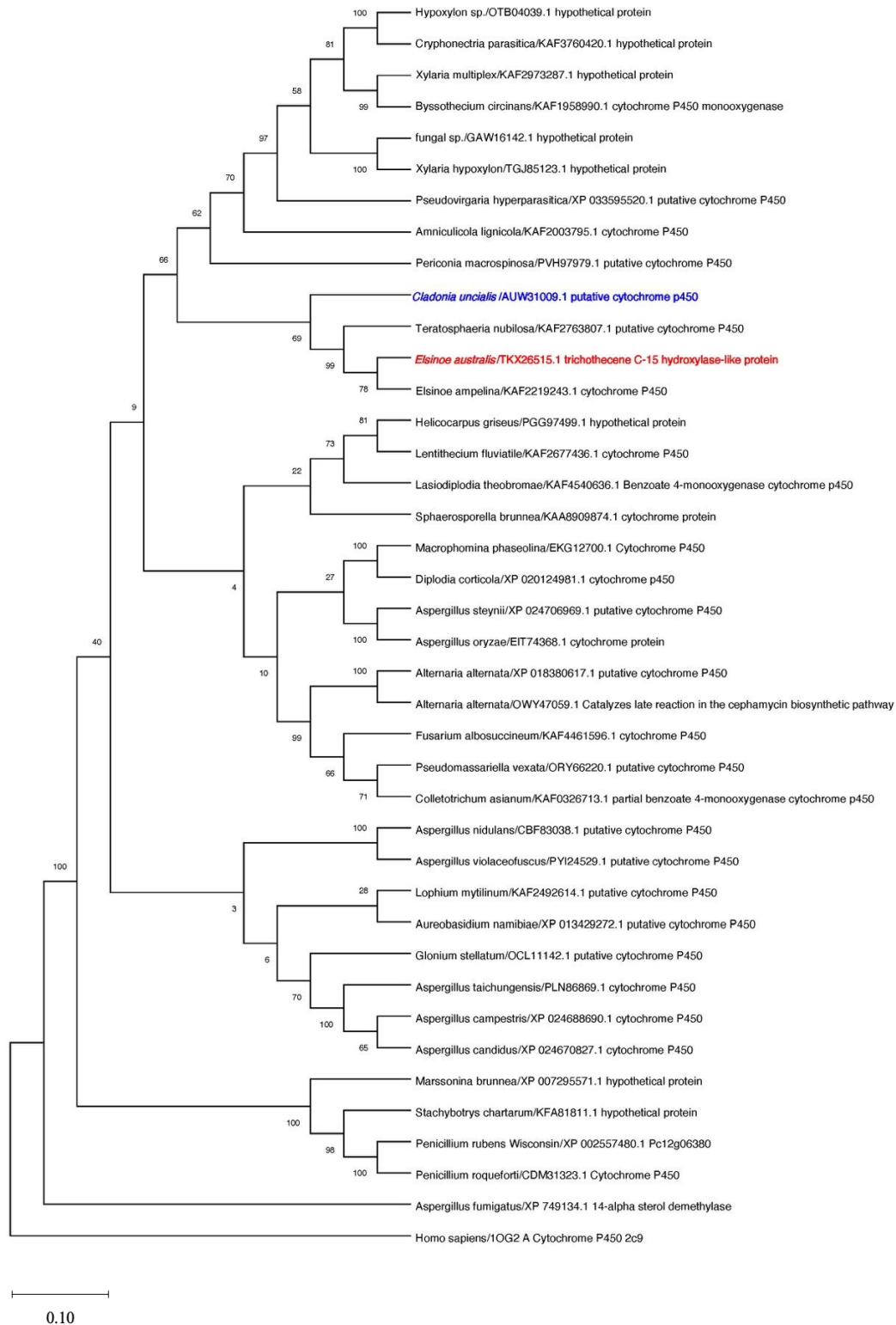
**Figure S5:** Phylogenetic relationship between a putative *farnesyl-diphosphate farnesyl transferase* of *Cladonia uncialis* (*Cu-terp-6*) and a genetically similar gene encoding a *bifunctional farnesyl-diphosphate farnesyl transferase* of *Imshaugia Aleurites*. The *farnesyl-diphosphate farnesyl transferase* of *Cryomyces minteri* (XP\_037155202.1) and *farnesyl-diphosphate farnesyl transferase* of *Saccharomyces cerevisiae* were chosen as out-groups.



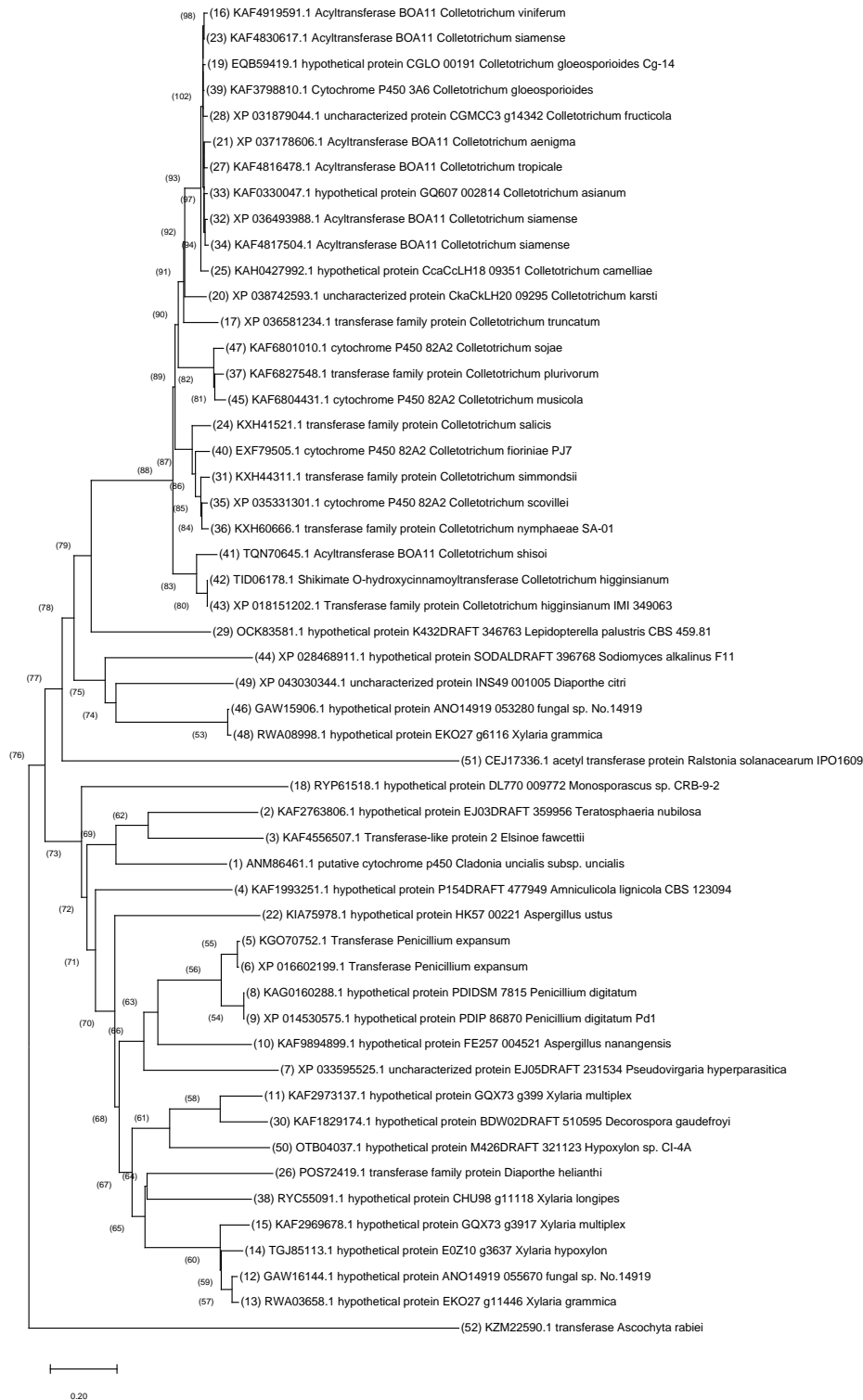
**Figure S6:** Phylogenetic relationship between a putative *aristolochene synthase* of *Cladonia uncialis* (*Cu-terp-9*) and a genetically similar gene encoding *aristolochene synthase* of *Lachnellula subtilissima*. The *trichodiene synthase* of *Fusarium asiaticum* (AHH81945.1) and was chosen as an out-group.



**Figure S7:** Phylogenetic relationship between a putative *Cytochrome p450* of *Cladonia uncialis* (*Cu-cyp4-terp-9*) and a genetically similar gene encoding a *CYP oxidoreductase* of *Sphaerosporella brunnea*. The cytochrome p450 (14-alpha demethylase) of *Aspergillus fumigatus* (XP\_749134.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.



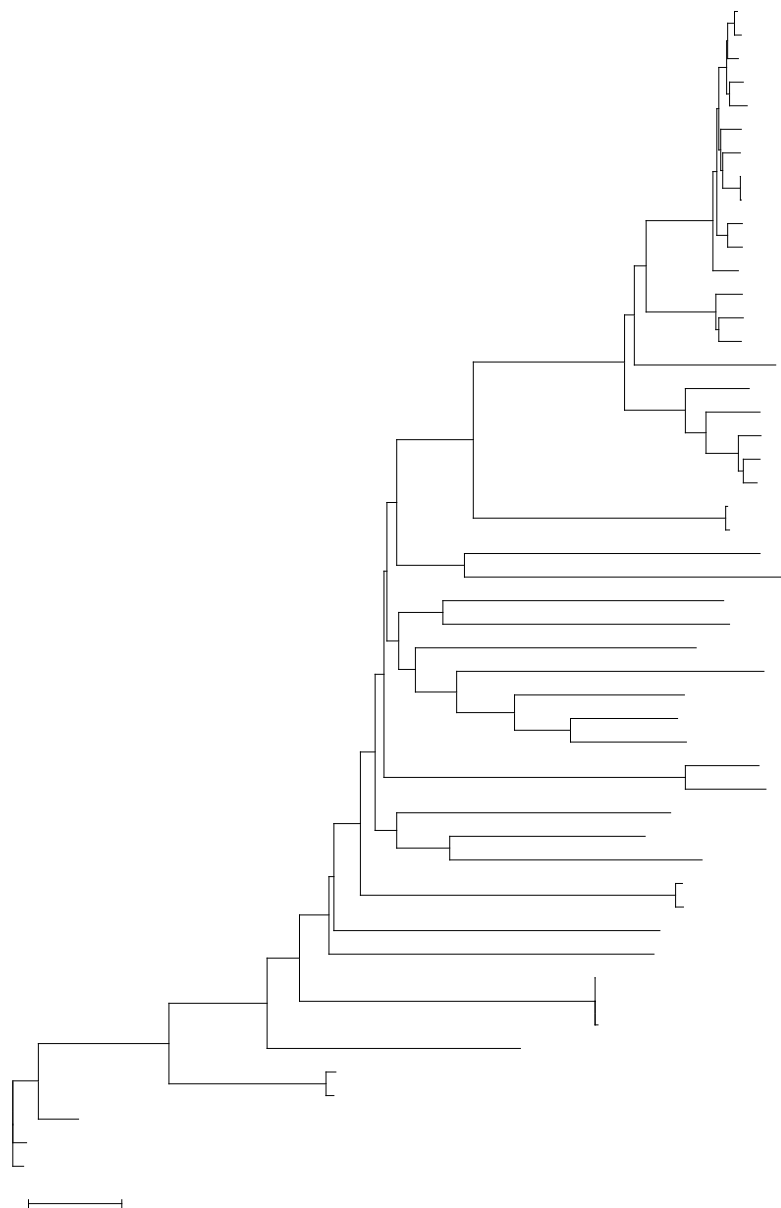
**Figure S8:** Phylogenetic relationship between a putative *Cytochrome p450* of *Cladonia uncialis* (*Cu-cyp5-terp-9*) and a genetically similar gene encoding a *Trichothecene C-15 hydroxylase* of *Elsinoe australis*. The cytochrome p450 (14-alpha demethylase) of *Aspergillus fumigatus* (XP\_749134.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.



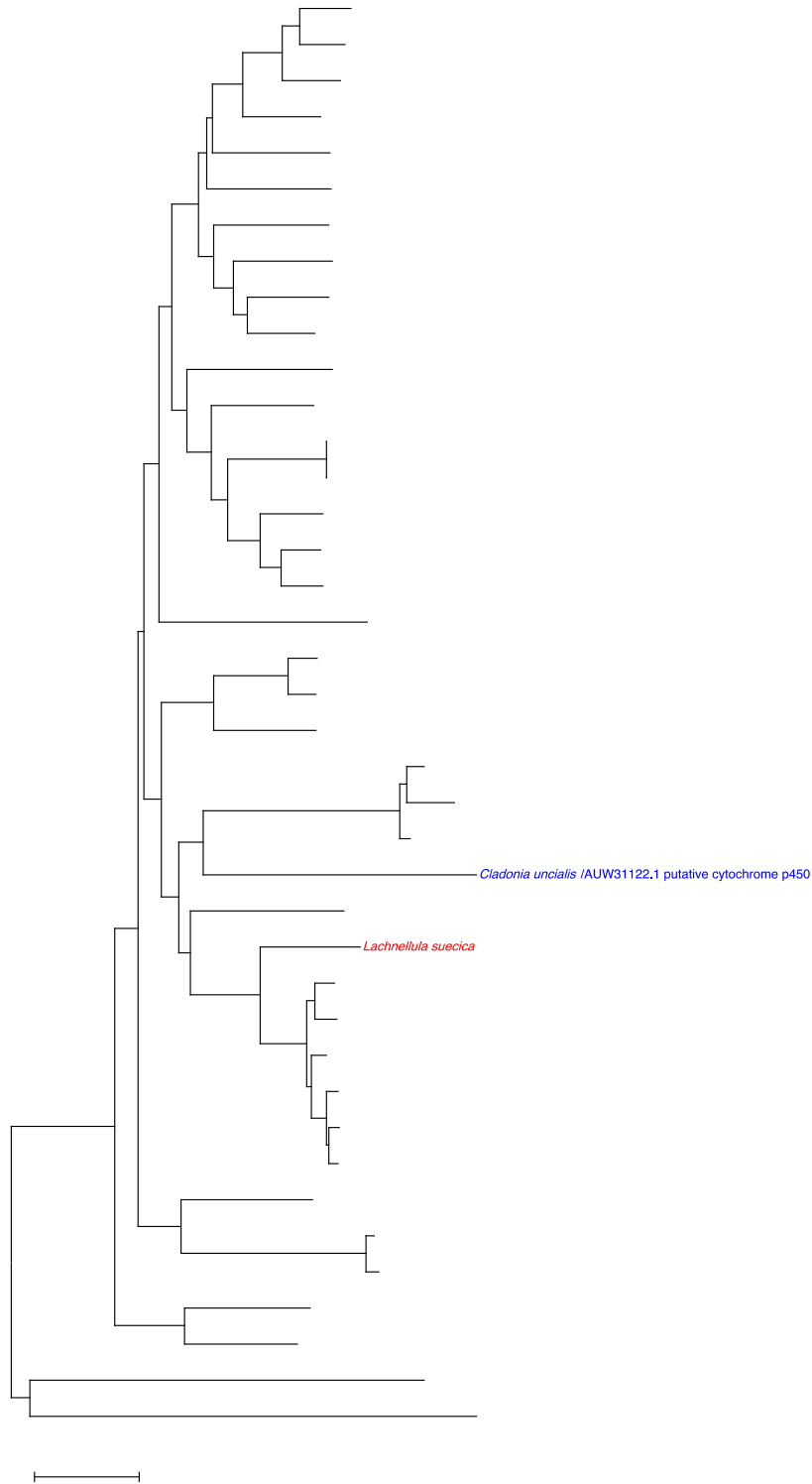
**Figure S9:** Phylogenetic relationship between a putative *acetyl transferase* of *Cladonia uncialis* (*Cu-terp9-acetyltransferase*) and a genetically similar gene encoding a *transferase-like protein 2* of *Elsinoe fawcettii*. The *transferase* of *Ascochyta rabiei* (KZM22590.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.



**Figure S10:** Phylogenetic relationship between a putative *short chain dehydrogenase* of *Cladonia uncialis* (*Cu-terp9-SDR*) and a genetically similar gene encoding a *short chain dehydrogenase-like protein 44* of *Elsinoe fawcettii*. The *short chain dehydrogenase* of *Dinoroseobacter shibae* (KZM22590.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.

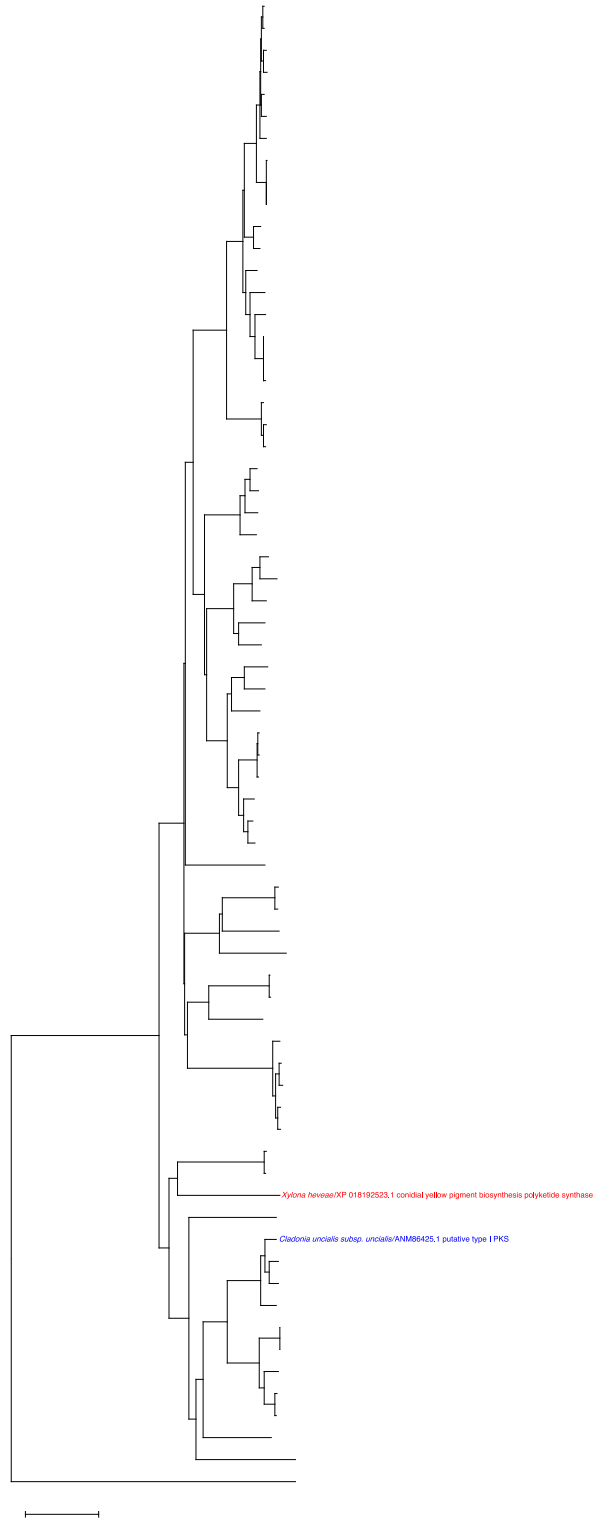


**Figure S11:** Phylogenetic relationship between a putative *type I PKS* of *Cladonia uncialis* (*Cu-terp9-SQTKS*) and a genetically similar gene encoding a *squalestatin tetraketide synthase* of *Phoma* sp. The *type I PKS* (*synthase 3*) of *Neurospora crassa* (XP\_959122.2.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.

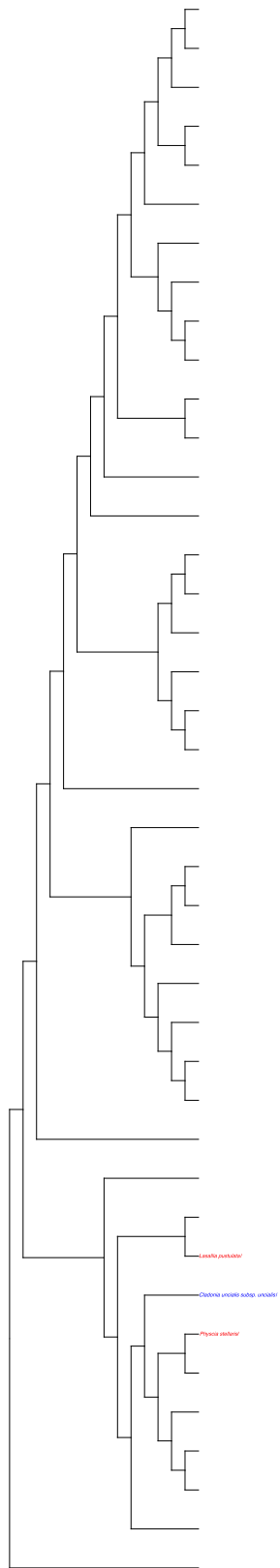


**Figure S12:** Phylogenetic relationship between a putative *Cytochrome p450* of *Cladonia uncialis* (*Cu-cyp6-nrpks-4*) and a genetically similar gene encoding a *Pisatin demethylase* of *Lachnellula suecica*. The cytochrome p450 (14- $\alpha$  demethylase) of *Aspergillus fumigatus* (XP\_749134.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.

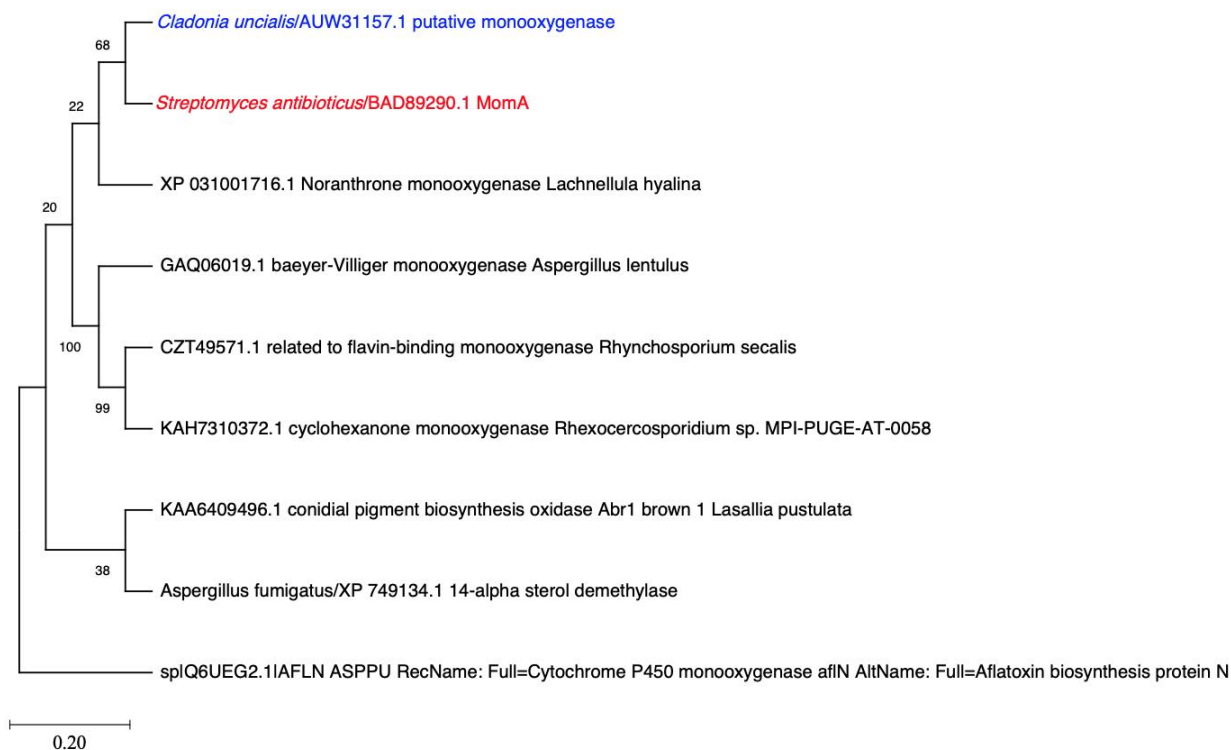




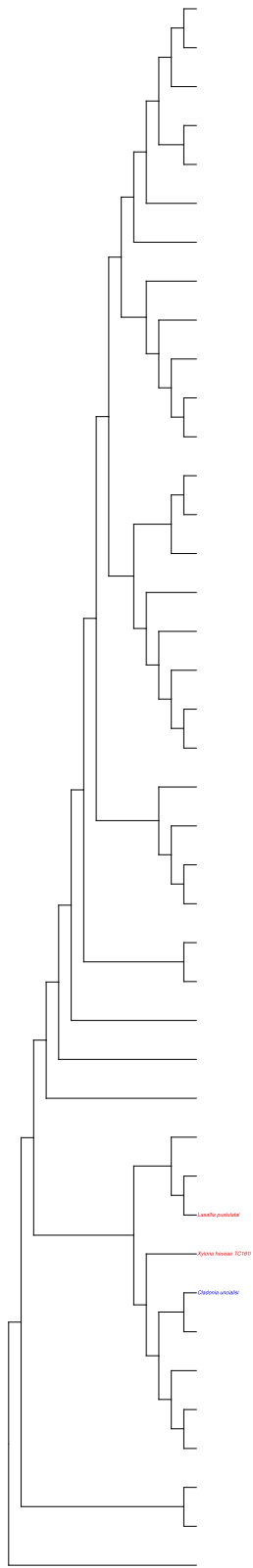
**Figure S13:** Phylogenetic relationship between a putative *Type I PKS* of *Cladonia uncialis* (*Cu-nr-pks-6*) and a genetically similar gene encoding a *conidial yellow pigment: melanin synthase* of *Xylona heveae*. The 6-MSAS of *Aspergillus terreus* (BAA20202.2) was chosen as an out-group.



**Figure S14:** Phylogenetic relationship between a putative *short-chain dehydrogenase* of *Cladonia uncialis* (*Cu-nrpk6-6-SDR*) and a genetically similar gene encoding a *tetrahydroxynaphthalene reductase* of *Lasallia pustulata*. The 6-MSAS of *Aspergillus terreus* (BAA20202.2) was chosen as an out-group.



**Figure S15:** Phylogenetic relationship between a putative *monooxygenase* of *Cladonia uncialis* (*Cu-nrpk6-6-MO*) and a genetically similar gene encoding a **MomA** of *Streptomyces antibioticus*. The cytochrome p450 monooxygenase of *Aspergillus parasiticus* (Q6UEG2.1) was chosen as an out-group.

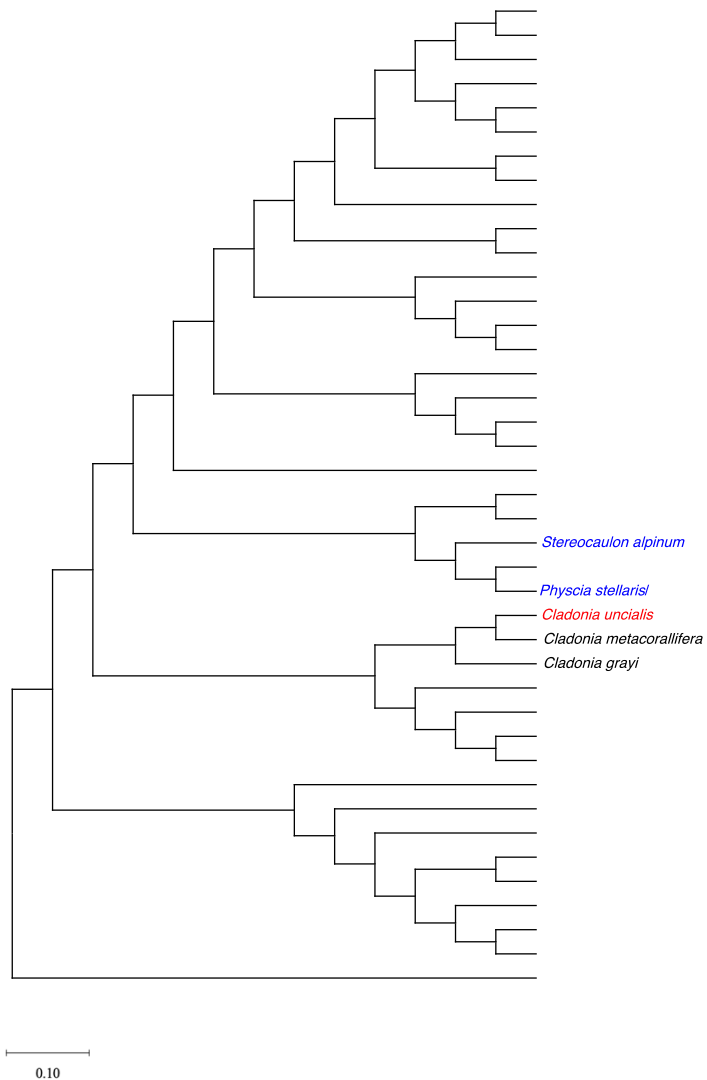


0.10

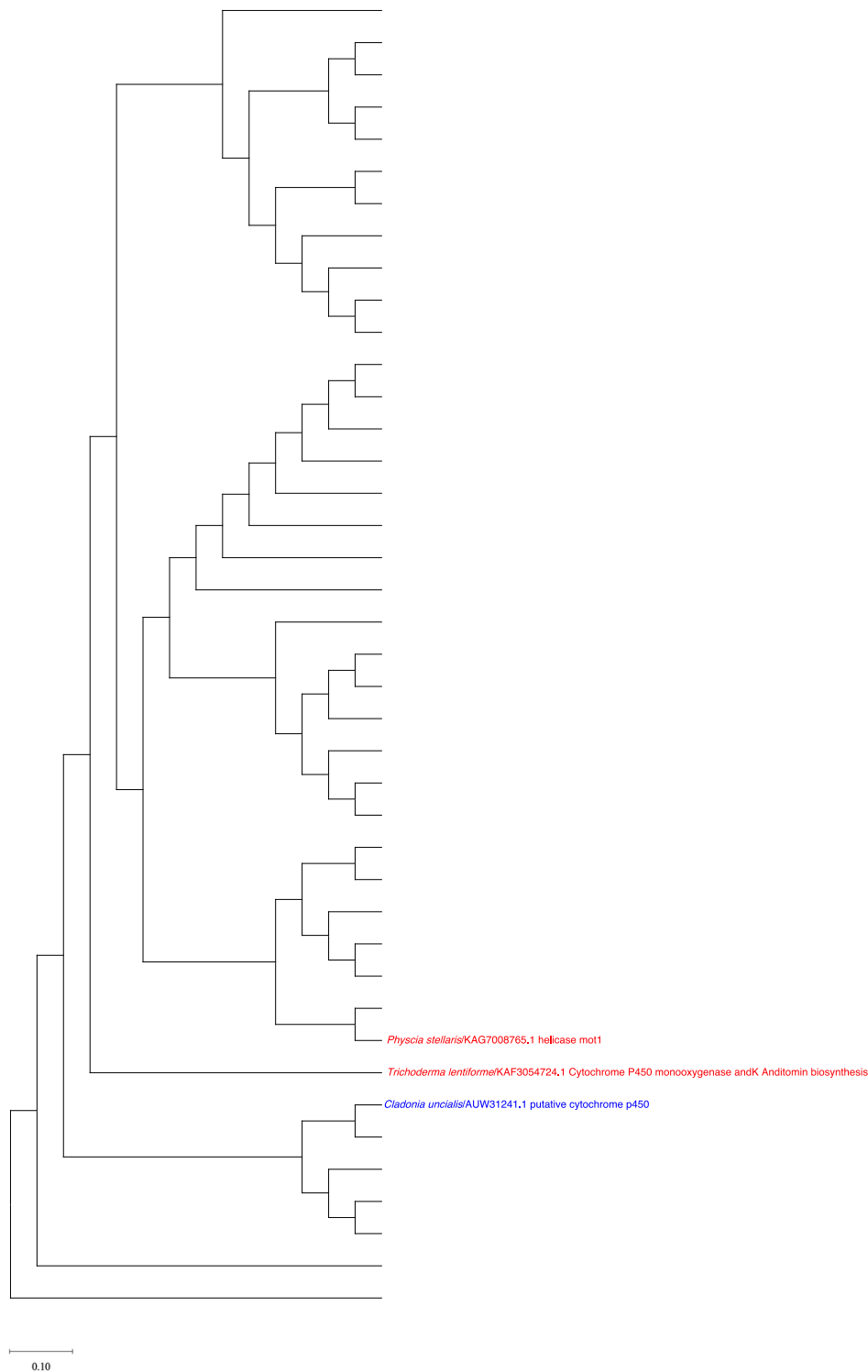
**Figure S16:** Phylogenetic relationship between a putative *scytalone dehydratase* of *Cladonia uncialis* (*Cu-nrpks-6-SD*) and a genetically similar gene encoding *scytalone dehydratase arp1* of *Lasallia pustulata*. The 6-MSAS of *Aspergillus terreus* (BAA20202.2) was chosen as an out-group.



**Figure S17:** Phylogenetic relationship between a putative *Cytochrome p450* of *Cladonia uncialis* (*Cu-cyp7-nrpk5-6*) and a genetically similar gene encoding a *CYP monooxygenase* of *Talaromyces marneffi*. The cytochrome p450 (14- $\alpha$  demethylase) of *Aspergillus fumigatus* (XP\_749134.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.

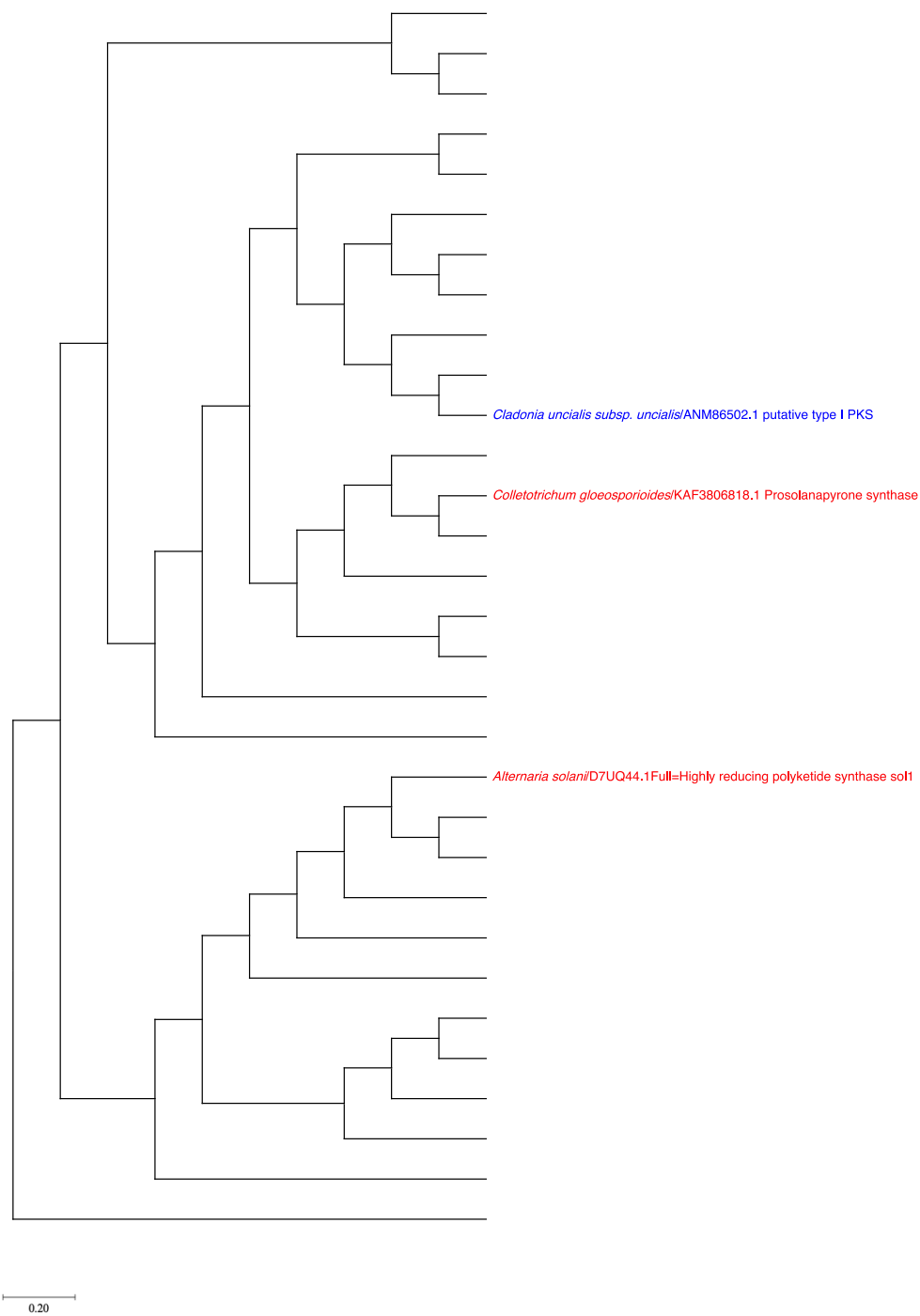


**Figure S18:** Phylogenetic relationship between a putative *type I PKS* of *Cladonia uncialis* (*Cu-nr-pks-12*) and a genetically similar gene encoding a *non-reducing PKS* (*atr1*, *atranorin biosynthesis*) of *Stereocaulon alpinum*. The cytochrome p450 (14-alpha demethylase) of *Xylaria sp.* (AAM93545.1) was chosen as an out-group.

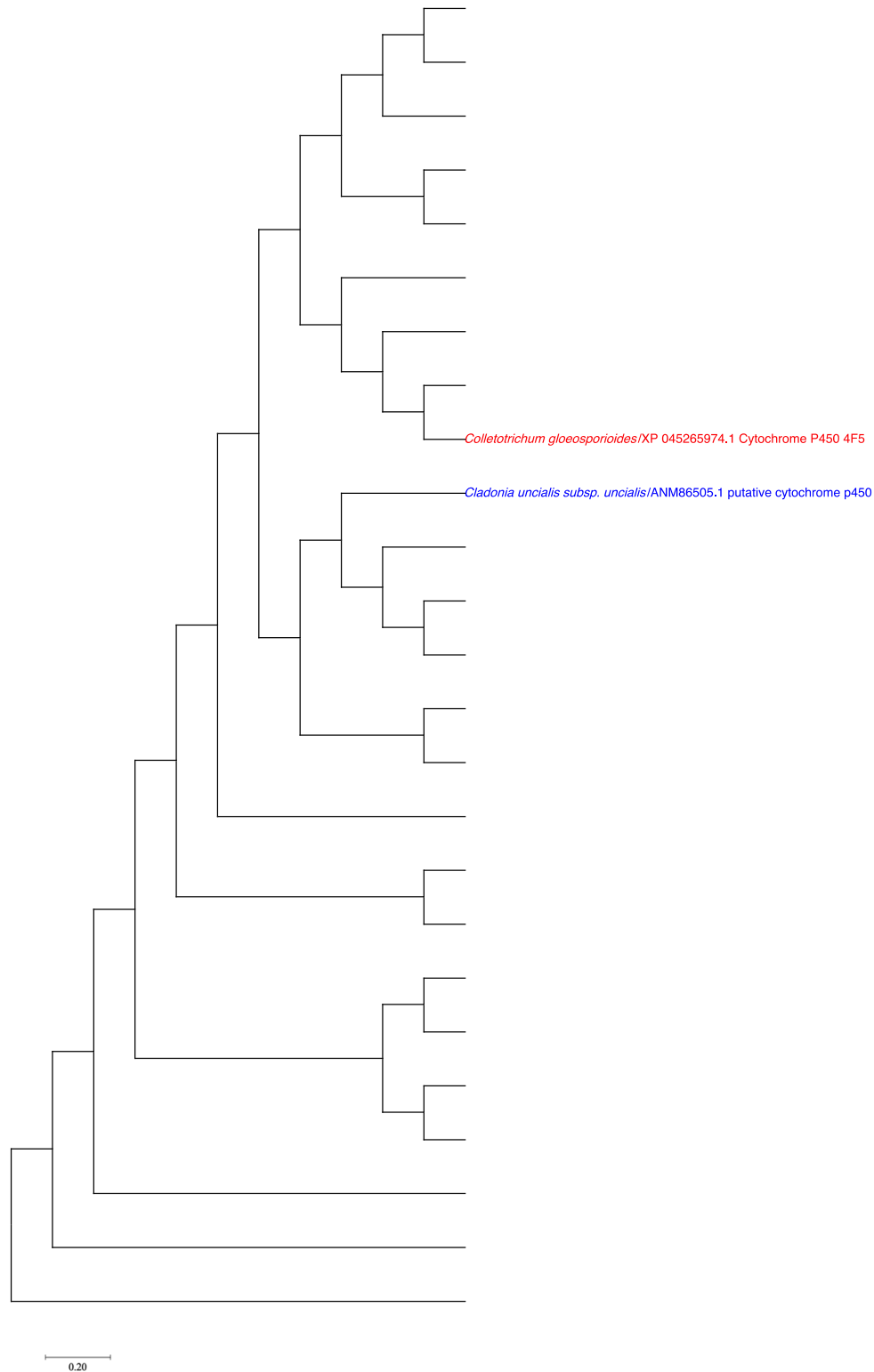


**Figure S19:** Phylogenetic relationship between a putative *Cytochrome p450* of *Cladonia uncialis* (*Cu-cyp8-nrpk5-12*) and a genetically similar gene encoding a *CYP monooxygenase (andK, part of Anditomin biosynthetic gene cluster)* of *Trichoderma lentiforme*. The cytochrome p450 (14-alpha demethylase) of *Aspergillus fumigatus* (XP\_749134.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.

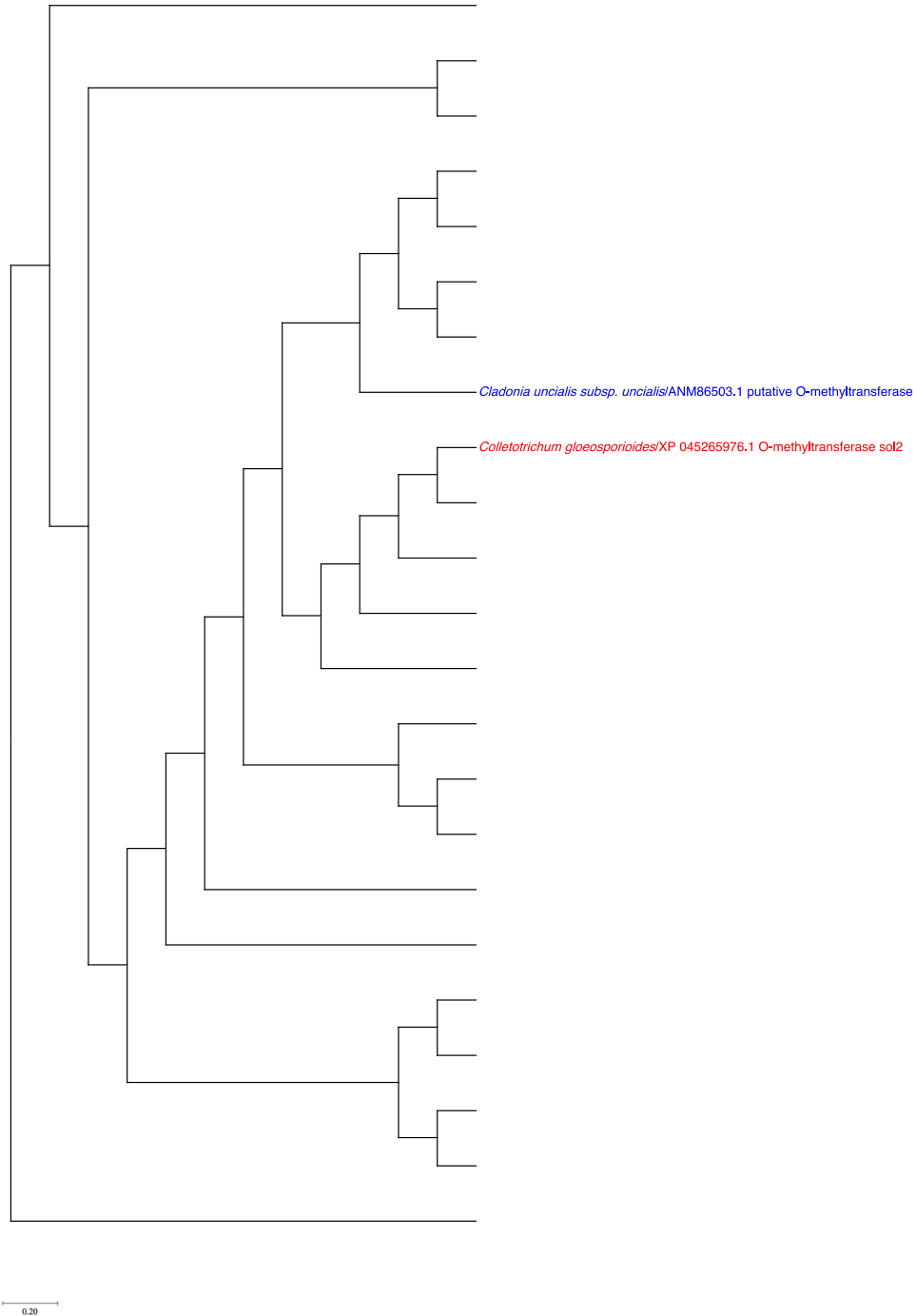




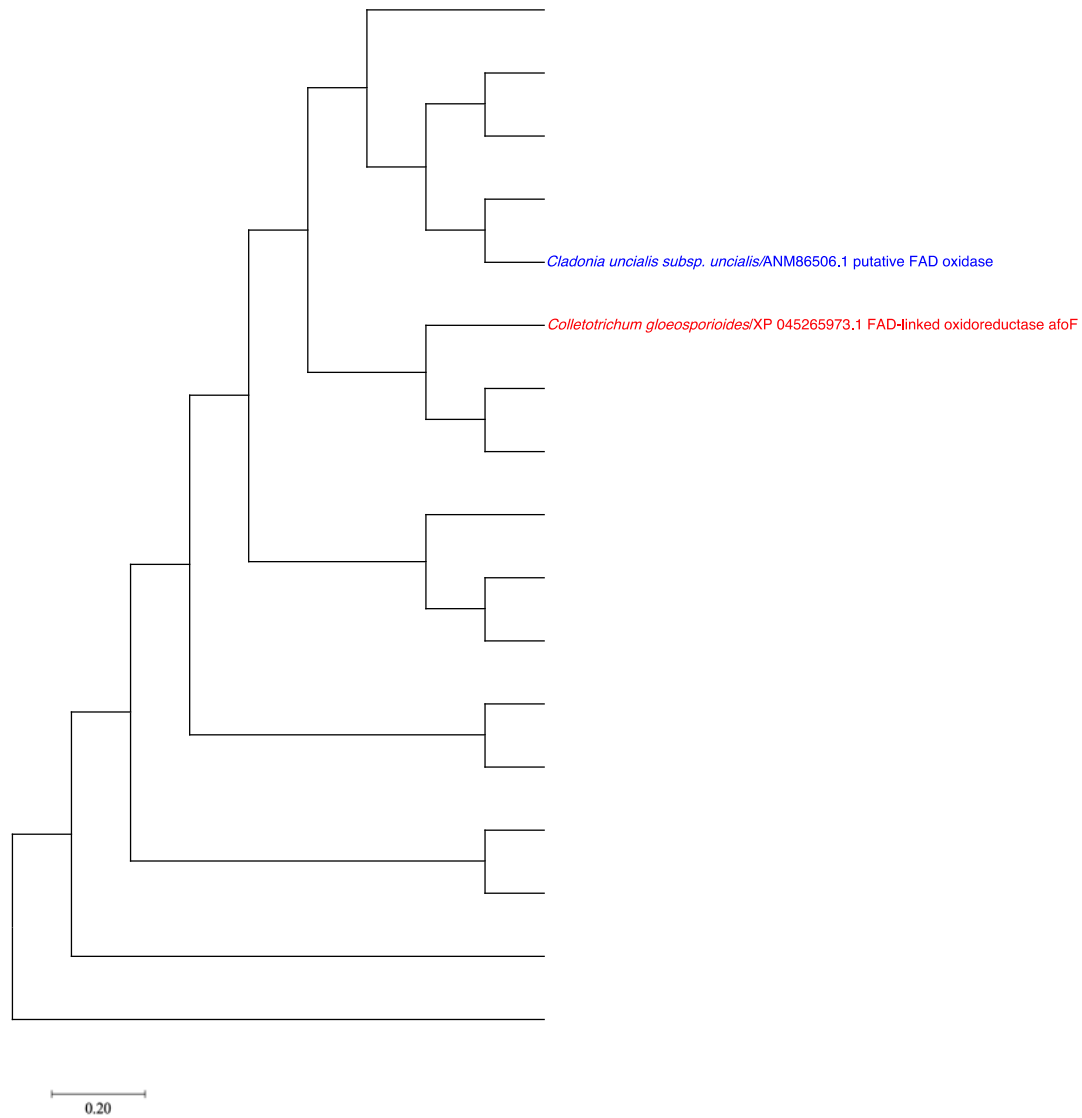
**Figure S20:** Phylogenetic relationship between a putative *Type I reducing PKS* of *Cladonia uncialis* (*C-r-pks-18*) and a genetically similar gene encoding a *Prosolanapyrone synthase* of *Colletotrichum gloeosporioides*. The melanin synthase of *Xylaria sp.* was chosen as an outgroup.



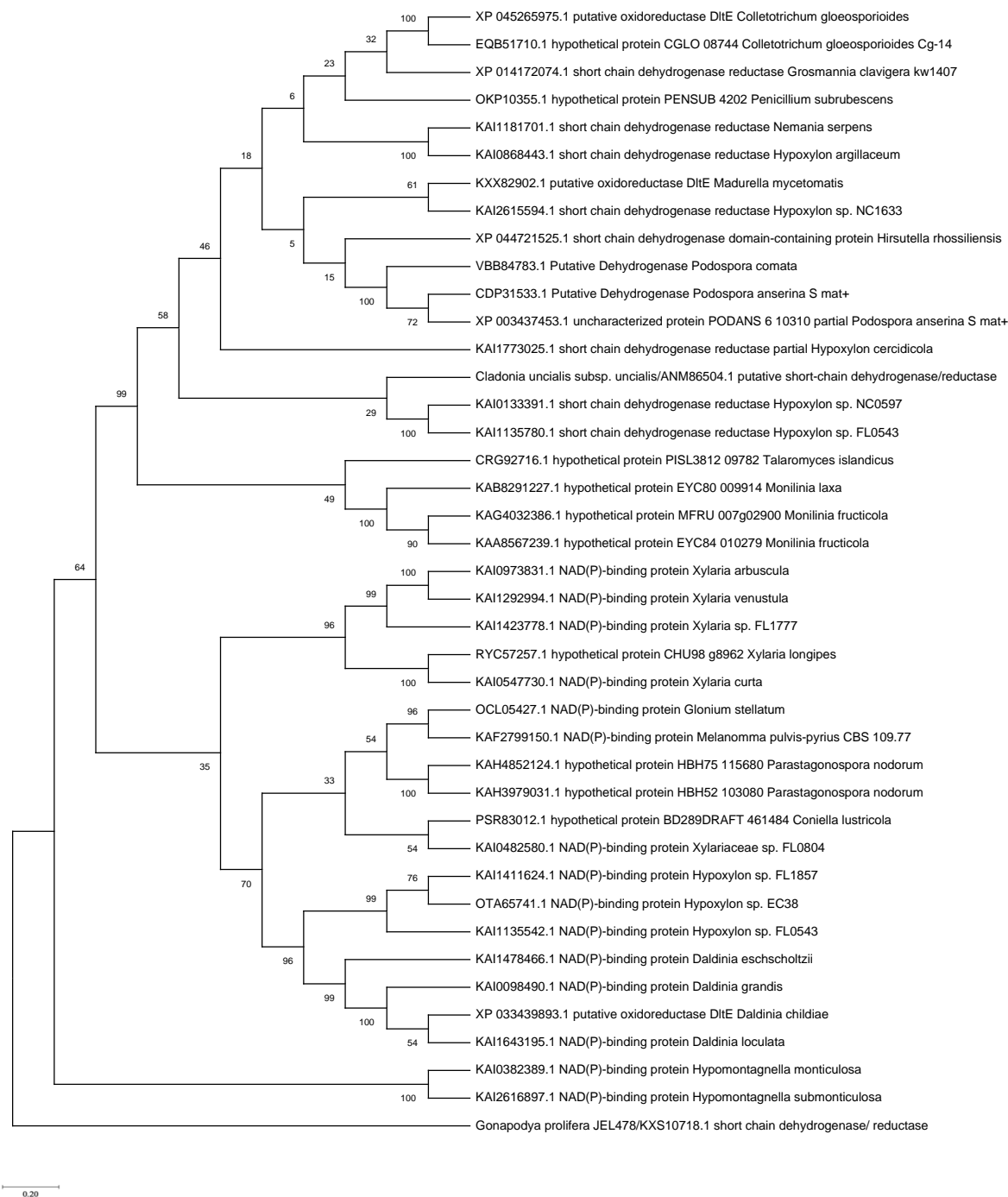
**Figure S21:** Phylogenetic relationship between a putative *cytochrome p450* of *Cladonia uncialis* (*C-cyp9-rpks-18*) and a genetically similar gene encoding a *cytochrome p450 4F5* of *Colletotrichum gloeosporioides*. The cytochrome p450 (14-alpha demethylase) of *Aspergillus fumigatus* (XP\_749134.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.



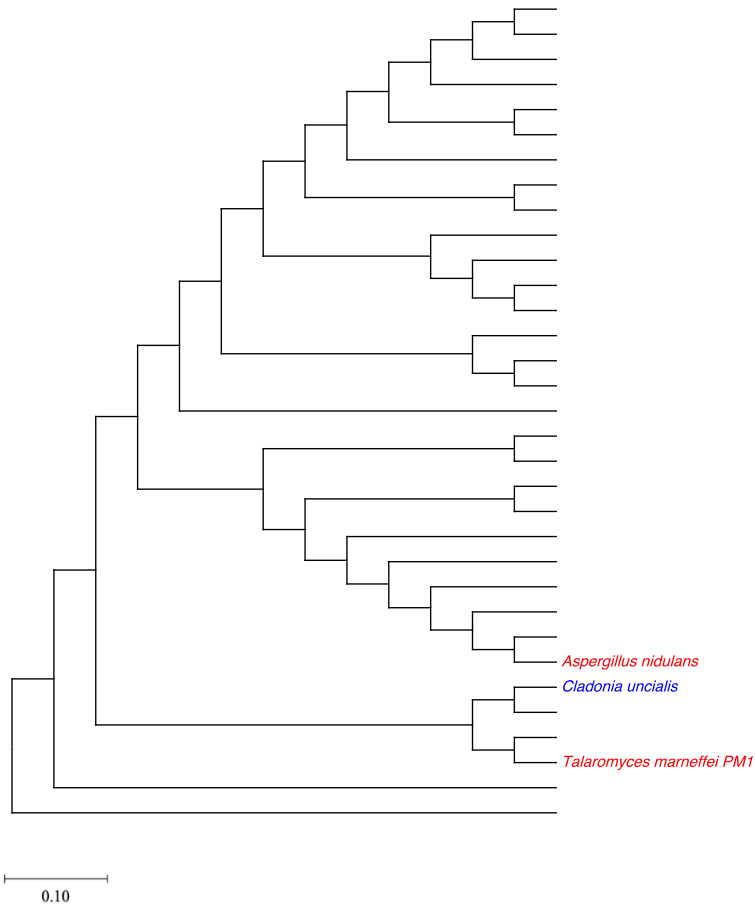
**Figure S22:** Phylogenetic relationship between a putative *O-methyltransferase* of *Cladonia uncialis* (*C-rpks-18-OMT*) and a genetically similar gene encoding a *O-methyltransferase* (*sol2*) of *Colletotrichum gloeosporioides*. The *O-methyltransferase* of *Pseudomonas aeruginosa* (AXL70985.1) was chosen as an out-group.




**Figure S23:** Phylogenetic relationship between a putative *FAD oxidase* of *Cladonia uncialis* (*C-rpks-18-FAD oxidase*) and a genetically similar gene encoding a *FAD-linked oxidoreductase afoF* of *Colletotrichum gloeosporioides*. The *FAD-dependant oxidase* of *Bacillus thuringiensis* (EEM38188.1) was chosen as an out-group.



**Figure S24:** Phylogenetic relationship between a putative *short-chain dehydrogenase/reductase* of *Cladonia uncialis* (*C-rpks-18-SDR*) and a genetically similar gene encoding a *putative oxidoreductase DltE* of *Colletotrichum gloeosporioides*. The of *short-chain dehydrogenase/reductase* of *Pseudomonas aeruginosa* (KSG75976.1) was chosen as an out-group.



**Figure S25:** Phylogenetic relationship between a putative *cytochrome p450* of *Cladonia uncialis* (*MPAO*) and genetically similar genes encoding *cytochrome p450* of *T. marneffi* & *p450 monooxygenase STCB* of *A. nidulans*. The cytochrome p450 (14- $\alpha$  demethylase) of *Aspergillus fumigatus* (XP\_749134.1) and cytochrome p450 of *Homo sapiens* were chosen as out-groups.



ACS Publications  
WILEY-INTERSCIENCE | WILEY | WILEY-VCH | WILEY-BLANKENHORN

**Lichen Biosynthetic Gene Clusters. Part I. Genome Sequencing Reveals a Rich Biosynthetic Potential**

Author: Robert L. Bertrand, Mona Abdel-Hameed, John L. Sorensen  
 Publication: Journal of Natural Products  
 Publisher: American Chemical Society  
 Date: Apr 1, 2018

Copyright © 2018, American Chemical Society

**Quick Price Estimate**

This service provides permission for reuse only. If you do not have a copy of the portion you are using, you may copy and paste the content and reuse according to the terms of your agreement. Please be advised that obtaining the content you license is a separate transaction not involving RightsLink.

If credit is given to another source for the material you requested from RightsLink, permission must be obtained from that source.

Note: Individual Scheme and Structure reuse is free of charge and does not require a license. If the scheme or structure is identified as a Figure in the article, permission is required.

Permission for this particular request is granted for print and electronic formats, and translations, at no charge. Figures and tables may be modified. Appropriate credit should be given. Please print this page for your records and provide a copy to your publisher. Requests for up to 4 figures require only this record. Five or more figures will generate a printout of additional terms and conditions. Appropriate credit should read: "Reprinted with permission from (COMPLETE REFERENCE CITATION). Copyright {YEAR} American Chemical Society." Insert appropriate information in place of the capitalized words.

I would like to... ⓘ reuse in a Thesis/Dissertat ▼

Requestor Type ⓘ Author (original work) ▼

Portion ⓘ Table/Figure/Micrograph ▼

Number of Table/Figure/Microgra  ⓘ

Format ⓘ Print and Electronic ▼

Select your currency CAD - \$ ▼

Quick Price ▶ Click Quick Price

QUICK PRICE CONTINUE

To request permission for a type of use not listed, please contact the publisher directly.

© 2022 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Data Security and Privacy  
 | For California Residents | Terms and Conditions Comments? We would like to hear from you. E-mail us at  
[customer-care@copyright.com](mailto:customer-care@copyright.com)

<https://s100.copyright.com/AppDispatchServlet#formTop>

**Figure S26:** Rightslink® by Copyright Clearance of Figures 5, 6 and 7 from *Journal of Natural Products* 2018 81 (4), 723-731.