



**HAL**  
open science

**Deciphering the biosynthetic pathways of PKS-NRPS  
derived metabolites from the endophytic fungus  
Sarocladium zeae: Elucidation of pyrrocidines  
biosynthesis**

Youwei Chen

► **To cite this version:**

Youwei Chen. Deciphering the biosynthetic pathways of PKS-NRPS derived metabolites from the endophytic fungus *Sarocladium zeae*: Elucidation of pyrrocidines biosynthesis. Organic chemistry. Sorbonne Université, 2022. English. NNT: 2022SORUS133 . tel-04156320

**HAL Id: tel-04156320**

**<https://theses.hal.science/tel-04156320v1>**

Submitted on 8 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Sorbonne Université

École doctorale 406 – Chimie Moléculaire Paris Centre  
Molécules de Communication et Adaptation des Micro-organismes  
UMR 7245, CNRS-MNHN

# **Deciphering the biosynthetic pathways of PKS-NRPS derived metabolites from the endophytic fungus *Sarocladium zae*: Elucidation of pyrrocidines biosynthesis**

Par Youwei CHEN

Thèse de doctorat de Chimie Moléculaire

Dirigée par Dr. Didier BUISSON

Co-encadrée Dr. Stéphane MANN

Présentée et soutenue publiquement le 7 juillet 2022

Devant un jury composé de :

Pr. Dominique GUIANVARCH	Professeur	Rapporteur
Pr. Marie-Agnès SARI	Professeur	Rapporteur
Dr. Béatrice GOLINELLI-PIMPANEAU	Directrice de Recherche	Examineur
Dr. Hervé LALUCQUE	Maitre de conférences-HDR	Examineur
Dr. Stéphane MANN	Chargé de Recherche	Co-encadrant de Thèse
Dr. Didier BUISSON	Directeur de Recherche	Directeur de Thèse



## ACKNOWLEDGMENTS

This thesis work was carried out within the Muséum National d'Histoire Naturelle (MNHN) -Molécules de Communication et Adaptation des Micro-organismes (MCAM) UMR7245 - CPNFB Chimie des Produits Naturels Fongiques et Bactériens.

Looking backward, time was fleeting. Four years is like a sudden meteor that vanishes in just a second. Paris is an open and humanistic city. The four years I have spent here are the treasures of my life. Obviously, my doctoral career, it's a "bittersweet" and (best of all) hopeful film. I'd like to take this opportunity to thank you folks who surrounded, encouraged, and supported me during these years.

I would like to express my sincere thanks to **Dominique GUIANVARC'H**, **Marie-Agnès SARI** for accepting to be my thesis rapporteurs. I thank them for taking the time to read and judge my work. Thanks to **Hervé LALUCQUE** for judging my work. I am very honored to count you among the members of my thesis jury.

I warmly thank **Béatrice GOLINELLI-PIMPANEAU** for agreeing to be President of my thesis jury and for providing excellent advice and suggestions to help my work progress.

Thanks to **Philippe GRELLIER** for accepting me into MCAM. I am so grateful to him for giving me the opportunity to complete my thesis within his team.

I am extremely grateful to my supervisors **Stéphane MANN** and **Didier BUISSON** for offering me this opportunity to pursue my Ph.D in Paris and also for their kindly help. I thank **Didier BUISSON** for his trust and the freedom he has given me to move forward step by step during these four years. Sincere thanks to **Stéphane MANN** for dedicating himself to the dissemination of engineering knowledge, and, especially to the instruction of tedious data analysis. He teaches by precept and example, showing me always to be rigorous and with unique insights in part of scientific research. Thank you for the always interesting discussions and the advice, whether for the posters, the orals, the manuscript or the future career. I sincerely appreciate the help of **Stéphane MANN** and **Didier BUISSON** revising my thesis. I have to say, all that they taught will be my lifelong treasure.

Thanks to **Yanyan LI** for having contributed to my development both scientifically and humanly.

Thank **Zhilai HONG** for his help in life and experiments when I first came to the laboratory. I greatly appreciate **all lab members** for their kindly help, whether for our scientific discussions

or less, for the lunches and all those moments spent eating! Taking this opportunity to express my most sincere thanks to **Steffi SEWSURN**, **Benedict BOSTOCK**, **Coralie PAVESI**, **Emilie ADOUANE**, **Man ZHOU** and **Xiaohan LIU** in MCAM. Because of their company, the life of my Ph. D has become rich and colorful.

I would like to acknowledge **Kevin CALABRO** for the structure elucidation course, which helped me a lot. A big thank to **S verine AMAND** and **Delphine CHAMPEVAL** for always being there to help me out whenever I have a technical problem. I have greatly appreciated all the scientific advices of **Soizic PRADO** and **Caroline KUNZ**. I would like to acknowledge **Dr. Arul Marie**, **Alexandre Deville** and **Dr. Alain Blond** for providing convenience on NMR and LC-MS/MS analytical platforms.

Finally, my most sincere thanks go to **my family**. I really appreciate my wife for her meticulous care and support so that I could successfully complete my studies. Without her, none of this would have been possible. I would like to thank my parents for their support and encouragement. I will always be by your side in happiness and in difficult times.

I gratefully acknowledge financial support from **China council scholarship (CSC)**.

# Table of Contents

<b>ABBREVIATIONS</b> .....	<b>7</b>
<b>FRENCH SUMMARY</b> .....	<b>9</b>
<b>FOREWORD</b> .....	<b>39</b>
<b>CHAPTER I: GENERAL INTRODUCTION</b> .....	<b>43</b>
<b>1.1 Diversity of natural products and their biosynthetic pathways in fungi.</b> .....	<b>44</b>
1.1.1 Terpenoids .....	45
1.1.2 Alkaloids.....	48
1.1.3 Peptides and peptide-derived natural products .....	49
1.1.3.1 Non-ribosomal peptides (NRPs) .....	49
1.1.3.2 Diketopiperazines (DKPs) .....	52
1.1.3.3 Ribosomally synthesized and post-translationally modified peptides (RiPPs).....	54
1.1.4 Polyketides (PKs) .....	55
1.1.5 Metabolites from hybrid pathways .....	60
1.1.5.1 Metabolites requiring two synthases .....	60
1.1.5.2 Metabolites from hybrid synthases .....	60
1.1.6 Regulation of secondary metabolism .....	65
<b>1.2 Discovery of new natural products</b> .....	<b>66</b>
1.2.1 Traditional natural product discovery strategies .....	66
1.2.2 Genome mining and activation of silent gene clusters.....	68
1.2.3 Heterologous expression .....	72
1.2.4 Combinatorial biosynthesis and synthetic biology.....	75
<b>1.3 Pyrrocidines and the endophytic fungi <i>Sarocladium zeae</i></b> .....	<b>78</b>
1.3.1 Pyrrocidines and their analogs, an expanding compound family .....	78
1.3.2 Biosynthetic origin of pyrrocidines.....	80
<b>1.4 Objectives</b> .....	<b>87</b>
<b>CHAPTER II: BIOSYNTHESIS OF PYRROCIDINES IN <i>SAROCLADIUM ZEA</i></b> .....	<b>89</b>
<b>2.1 Pyrrocidines and analogues produced by <i>Sarocladium zeae</i></b> .....	<b>90</b>
2.1.1 Culture conditions for pyrrocidines production in <i>S. zeae</i> .....	90
2.1.2 Characterization of pyrrocidines and analogs in wild type <i>S. zeae</i> . .....	91
<b>2.2 BGCs involving PKS-NRPS in the genome of <i>S. zeae</i></b> .....	<b>104</b>
<b>2.3 Identification of the PKS-NRPS producing pyrrocidines by gene knock-out</b> .....	<b>109</b>
<b>2.4 Annotation of pyrrocidine BGC</b> .....	<b>114</b>
<b>2.5 Identification of the function of auxiliary enzymes by gene inactivation</b> .....	<b>116</b>
2.5.1 Construction of the deletion mutants of auxiliary enzymes and LC-MS analysis.....	116

2.5.2 Analysis of the inactivation of the gene coding PrcI.....	119
2.5.3 Analysis of the inactivation of the gene coding PrcX.....	120
2.5.4 Analysis of the inactivation of the gene coding PrcH .....	125
2.5.5 Inactivation of the genes <i>prcB</i> , <i>prcD</i> and <i>prcE</i> .....	129
<b>2.6 Proposed biosynthetic pathway of pyrrocidines.....</b>	<b>142</b>
<b>2.7 Evaluation of the effect of the pyrrocidine production during the interaction between <i>S. zeae</i> and <i>Fusarium verticillioides</i> .....</b>	<b>149</b>
<b>2.8 Pyrrocidine BGC in other fungi: Genome mining .....</b>	<b>153</b>
<b>CHAPTER III: RECONSTITUTION OF THE BIOSYNTHESIS PATHWAY OF PYRROCIDINES.....</b>	<b>157</b>
3.1 Introduction.....	158
3.2 Cloning of genes <i>prcC</i> coding the ER and <i>prcA</i> coding the PKS-NRPS .....	160
3.3 Co-expression of PKS-NRPS-67 and ER in <i>S. cerevisiae</i> BJ5464-NpgA.....	165
3.4 Characterization of the efficiency of promoter pRPL18B in BJ5464-NpgA .....	168
3.5 Verification of the production of PKS-NRPS and ER in yeast by Western blot.....	169
<b>CHAPTER IV: ANALYSIS OF OTHER GENE CLUSTERS .....</b>	<b>171</b>
4.1 Introduction.....	172
4.2 Bioinformatics analysis of BGC involving <i>prbA</i> from scaffold 47 .....	172
4.3 Analysis of cluster <i>pra</i> .....	178
<b>CHAPTER V: MATERIALS AND METHODS.....</b>	<b>181</b>
<b>CONCLUSIONS AND PERSPECTIVES.....</b>	<b>195</b>
<b>REFERENCE .....</b>	<b>197</b>
<b>ANNEX.....</b>	<b>218</b>

# Abbreviations

A domain	Adenylation domain
ACP domain	Acyl carrier protein
Amp	Ampicillin
AT domain	Acyl transferase
BGC	Biosynthetic gene cluster
C domain	Condensation domain
Cam	Chloramphenicol
<sup>13</sup> C NMR	Carbon nuclear magnetic resonance
cDNA	Complementary DNA
CDPS	Aminoacyl-tRNA-dependent CDP synthase
CoA	Coenzyme A
COSY	Correlation spectroscopy
1D/2D	One Dimension / Two Dimension
DH domain	Dehydratase
DKP	Diketopiperazine
DMSO	Dimethyl sulfoxide
EDTA	Ethylenediaminetetraacetic acid
EIC	Extract ion chromatogram
ER domain	Enoyl reductase
gDNA	Genomic DNA
GNPS	Global Natural Product Social Molecular Networking
HMBC	Heteronuclear Multiple bond correlation
<sup>1</sup> H NMR	Proton nuclear magnetic resonance
HPLC	High pressure liquid chromatography
HR-PKS	Highly-reducing polyketide synthase
HSQC	Heteronuclear Single Quantum Coherence spectroscopy
Hyg	Hygromycin
Kan	Kanamycin
KOAc	Potassium acetate
KR domain	Keto-reductase
KS domain	Ketosynthase
MS	Mass spectrometry



MS <sup>2</sup>	Tandem mass spectroscopy
MT domain	C-methyltransferase
nm	Nanometer
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser effect spectroscopy
NR-PKS	Non-reducing polyketide synthase
NRPS	Non ribosomal peptide synthetase
OD	Optical density
OSMAC	One strain many compounds
PCR	Polymerase chain reaction
PKS	Polyketide synthase
Ppm	Parts per million
PPTase	Phosphopantetheinyl transferase
PR-PKS	Partially reducing polyketide synthase
R domain	Reductive domain
Ripp	Ribosomally synthesized and post-translationally modifies peptode
RT-PCR	Reverse transcription polymerase chain reaction
SAM	S-adenosylmethonine
SDR	Short-chain dehydrogenase
T domain	Thiolation domain
UV	Ultraviolet

**Décryptage des voies biosynthétiques de métabolites dérivés de PKS-NRPS  
chez le champignon endophyte *Sarocladium zae*:**

**Élucidation de la biosynthèse des pyrrocidines**

Avant-propos

Les micro-organismes endophytes sont des bactéries ou des champignons qui se développent dans les plantes de façon asymptomatique. Ils peuvent participer à la défense de la plante contre les maladies, notamment en produisant des métabolites secondaires ayant une activité antagoniste contre des phytopathogènes. L'un des axes de recherche de l'équipe Chimie des Produits Naturels Fongiques et Bactériens est l'étude de la communication chimique entre les différents protagonistes dans la relation plante – endophytes- pathogènes depuis la biosynthèse des métabolites impliqués jusqu'à leur rôle biologique.

Dans ce contexte, la thèse concerne l'étude du champignons *S. Zae*, qui est un endophyte du maïs, et de ces métabolites qui ont une activité contre des pathogènes tels que des *Fusarium* (fig. 1).

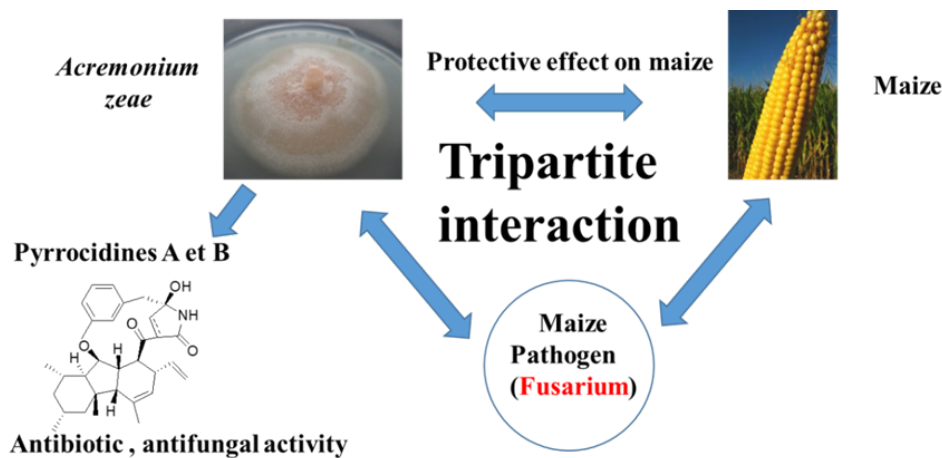


Fig. 1

Le chapitre 1 est consacré aux métabolites secondaires, avec en première partie l'objectif de montrer leur diversité et de discuter des réactions impliquées dans leur voie de biosynthèse. Ensuite, les approches pour découvrir de nouveaux produits naturels sont exposées. Enfin, l'état de l'art sur les pyrrocidines au début du travail de thèse est rapporté.

Le chapitre 2 décrit les résultats qui ont permis de proposer une voie de biosynthèse pour les pyrrocidines et leurs analogues. Celle-ci a été obtenue après avoir identifié le cluster de gènes impliqué, généré des mutants d'inactivation des différents gènes et étudié les métabolites alors produits.

Le chapitre 3 est consacré aux travaux menés pour reconstruire une voie de biosynthèse de la pyrrocidine chez la levure *Saccharomyces cerevisiae*.

Le chapitre 4 présente les travaux préliminaires en vue de l'identification des métabolites produits par les deux autres BGC impliquant un gène codant une PKS-NRPS chez *S. zeae*.

Le chapitre 5 est la partie expérimentale

## Chapitre 1

### Introduction générale

#### 1.1 Diversité des produits naturels et leurs voies biosynthétiques chez les champignons.

Les produits naturels sont utilisés comme médicaments pour prévenir et traiter les maladies depuis des milliers d'années.<sup>18</sup> De 1981 à 2019, il a été rapporté qu'environ 1 881 médicaments à base de petites molécules ont été approuvés par la FDA. Environ 50% des médicaments sur le marché provenaient de produits naturels et de leurs dérivés.

Les produits naturels proviennent essentiellement du métabolisme secondaire qui englobe tous les processus et molécules non nécessaires à la croissance du micro-organisme. Cela contraste avec le métabolisme primaire qui produit des molécules nécessaires au développement de la cellule, comme les lipides, les glucides, les acides nucléiques, les vitamines ou les hormones. Les métabolites secondaires peuvent être classés en fonction de leur structure dans les cinq classes suivantes : 1) terpénoïdes, 2) alcaloïdes, 3) peptides, 4) polycétides et 5) métabolites hybrides.

Dans la plupart des voies biosynthétiques produisant ces métabolites, le squelette de la molécule est d'abord assemblé par des synthases ou des synthétases, puis des fonctionnalisations chimiques sont installées par diverses enzymes auxiliaires. Ce mode de programmation permet d'atteindre une grande diversité de métabolites selon la nature et la spécificité de ces enzymes auxiliaires. Les gènes codant pour les enzymes impliquées dans une voie biosynthétique sont généralement regroupés sous forme de cluster de gènes biosynthétiques (BGC) et leur expression est souvent coréglée.

Les champignons produisent de nombreux produits naturels, mais plus de 90% des clusters de gènes sont silencieux dans les conditions de culture de laboratoire. Les champignons sont donc un réservoir potentiellement énorme pour la découverte de produits naturels. Dans ce chapitre nous présentons d'abord brièvement les différentes classes de métabolites secondaires et les réactions impliquées dans leur voie de biosynthèse, puis le cas des pyrrocidines de façon plus détaillée.

### 1.1.1 Terpénoïdes

Les terpénoïdes sont des terpènes, formés de plusieurs unités isoprènes et fonctionnalisés. La formule générale des terpènes est  $(C_5H_8)_n$  (n est le nombre d'unité d'isoprène). Selon le nombre de n, les terpènes peuvent être divisés en hémiterpènes (C5), monoterpènes (C10), sesquiterpènes (C15), diterpènes (C20), disesquiterpènes (C25), triterpènes (C30), tétraterpènes (C40) et polyterpènes (n > C40). Leur squelette provient de deux précurseurs à savoir l'isopentenyl pyrophosphate (IPP) et le pyrophosphate de diméthylallyle (DMAPP) (Fig.1.3). Il existe deux voies de formation des précurseurs DMAPP et IPP. Chez les champignons, ils sont formés à partir d'acetyl-CoA par la voie du mévalonate tandis que chez les plantes et bactéries, la voie du méthylerythritol phosphate prédomine.

### 1.1.2 Alcaloïdes

Les alcaloïdes sont définis comme des composés organiques contenant un atome d'azote inséré dans un cycle, avec une diversité de structure et une multiplicité d'activités. La biosynthèse des alcaloïdes implique généralement un acide aminé comme précurseur de l'azote. Par exemple, un acide aminé peut réagir avec d'autres blocs de construction du métabolisme primaire et être transformé en d'autres produits par oxydation et cyclisation. Les alcaloïdes peuvent être classés comme 1) la quinoléine et l'isoquinoléine comme pénicilotam, 2) les amines et les amides comme 7-amino-4-méthylcoumarine, 3) les dérivés d'indole comme 2-(3,3-diméthylprop-1-ène)-épiostaclavine, 4) les pyridines comme 7,8-diméthyl-isoalloxazine, 5) les quinazolines comme la chaétominine,<sup>38</sup> et 6) d'autres comme la trichoderme C.

### 1.1.3 Peptides

Les peptides et les dérivés peptidiques proviennent d'acides aminés, qui sont assemblés au cours de la biosynthèse du polymère par la formation de liaisons amides. On peut les classer en trois groupes 1) les peptides non ribosomiques (NRPs) 2) les diketopipérazines (DKPs) et 3) les peptides ribosomiques post-traductionnellement modifiés (RiPPs).

Les NRPs sont composées d'acides aminés protéinogènes et non protéinogènes, et plus de 300 résidus ont été identifiés. Ils sont produits par des méga-enzymes, les Peptides Synthétases Non Ribosomiques (NRPS). Les NRPS sont constitués de trois domaines essentiels : les domaines condensation (C), adénylation (A) et thiolation (T) (Fig.1.7). Le domaine d'adénylation (A) est chargé de sélectionner un acide aminé et de l'activer par l'intermédiaire d'un amino-acyl-AMP. Cet intermédiaire réagit avec le groupe thiol du

domaine T qui a été activé de manière préparatoire à son holoforme en ajoutant un bras phosphopanthéinyle. Il en résulte le transfert de l'acide aminé sur le domaine T. Le domaine C contrôle la stéréochimie peptidique en amont (donneur) et la sélectivité des acides aminés en aval (récepteur) avec des thioesters formés simultanément sous l'action du domaine de thiolation (T) (Fig. 1.1) et catalyse la formation de liaisons peptidiques. Le domaine C agit également comme un filtre de substrat. Chez les champignons, une sous-unité du domaine C (Ct) libère l'intermédiaire. De plus, toutes les chaînes peptidiques cycliques sont catalysées par le domaine Ct. Enfin, certaines NRPS itératives ont également été caractérisées.

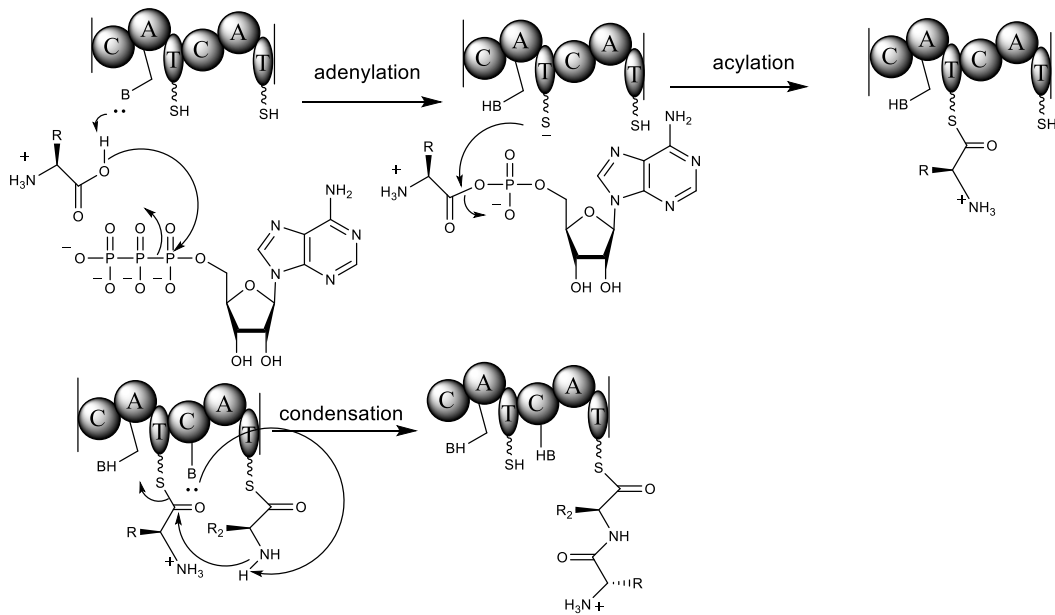


Fig. 1.1 Mécanisme des Peptide Synthétase Non-Ribosomiques (NRPS)

Les dicétopipérazines sont des dipeptides cycliques pouvant être synthétisé par deux voies différentes : soit *via* des NRPSs soit *via* des Aminoacyl (AA)-tRNA-Dependent Cyclo-Di-Peptide Synthases (CDPSs) (Fig. 1.2). D'autres enzymes peuvent modifier le squelette formé et donner des molécules complexes.

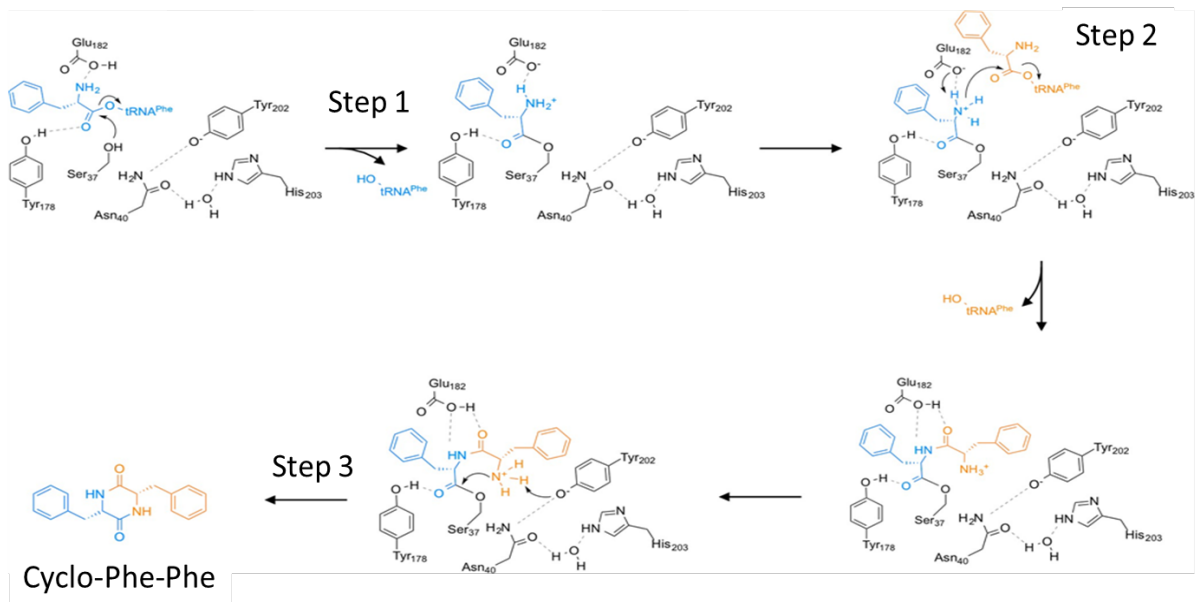


Fig. 1.2 Mécanisme Cyclo-Di-Peptide Synthases (CDPSs)

Les peptides ribosomiques post-traductionnellement modifiés (RiPPs) sont biosynthétisés *via* la transcription de l'ADN génomique puis la traduction de l'ARN. Le peptide est alors mûré par différents enzymes.

### 1.1.4 Polycétides

Les polycétides, les principaux métabolites secondaires des champignons, sont biosynthétisés par des polycétides synthases (PKS). Ces enzymes peuvent être classées en trois types en fonction de leur mécanisme catalytique, nommés type-I, II et III. Les PKS de type I peuvent être multimodulaires ou itératives, la plupart des PKS modulaires de type I se trouvant chez les bactéries et les PKS itératives de type I étant principalement décrites chez les champignons. Ces protéines sont composées de domaines aux fonctions catalytiques diverses. L'acyltransférase (AT), la kétosynthase (KS) et la protéine porteuse d'acyle (ACP) sont les domaines essentiels requis pour la biosynthèse du polycétide (PK). AT est en charge du recrutement et du transfert des unités d'extension du malonyl-CoA vers l'ACP.<sup>65</sup> KS catalyse une condensation de Claisen entre le groupe malonyle porté par l'ACP et une chaîne acyle se liant de manière covalente au KS lui-même.

Les PKS de type I comprennent d'autres domaines capables de modifier la chaîne en cours d'élongation : cétoréductase (KR), déshydratase (DH), méthyltransférase et enoyl réductase.

Les PKS itérative de type I peuvent avoir un domaine réducteur plus ou moins actif, cela se traduit par une chaîne polycétide plus ou moins oxydée. On distingue les PKS non réductrices (NR-PKS), les PKS partiellement réductrices (PR-PKS) et les hautement réductrices (HR-PKS)

Les PKS de type II sont en fait des complexes d'enzymes plus petites et multifonctionnelles ; Le mécanisme des PKS de type III est encore différent en ne faisant pas intervenir de domaine ACP.

## 1.1.5 Les métabolites issus de voies hybrides

Un certain nombre de molécules isolées des champignons ont des structures résultant d'une biosynthèse hybridant plusieurs des quatre voies précédemment présentées.

Chez les champignons, de nombreux métabolites sont issus de PKS et de NRPS. Un exemple correspond au produit d'une NRPS-PKS, où seul le domaine KS de la PKS intervient dans une réaction de Dieckmann.

Les PKS-NRPS sont principalement des hybrides de PKS itératives de type I, qui produisent le polycétide qui est ensuite transféré sur l'acide aminé chargé sur le domaine T de la partie NRPS. Le relargage du produit peut se faire suivant deux mécanismes suivant l'activité du domaine R, soit par réduction du thioester pour donner un aldéhyde, soit par cyclisation intramoléculaire pour donner une pyrrolidine dione (fig. 1.3)

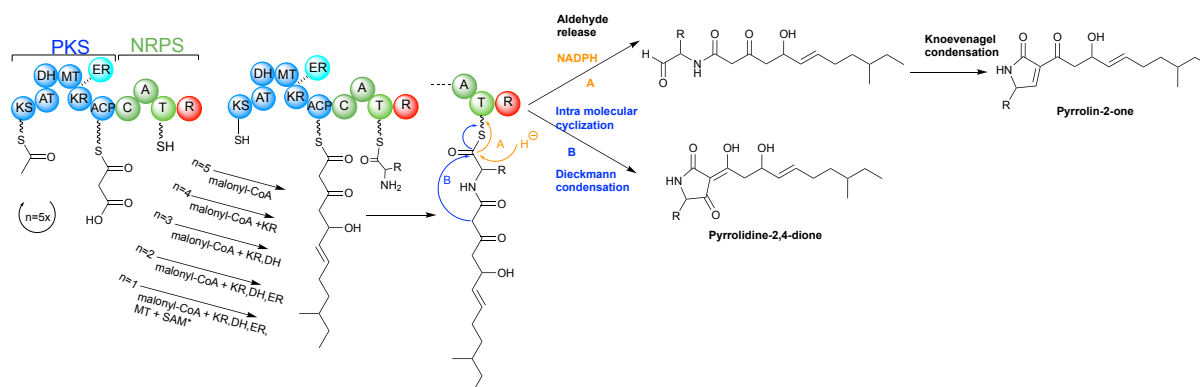


Fig. 1.3 Mécanisme des PKS-NRPS

## 1.1.6 Régulation du métabolisme secondaire

L'expression des gènes est sous la dépendance de nombreux facteurs ; ce peut être des facteurs environnementaux, pH, source de carbone, d'azote, lumière ou des facteurs épigénétiques.

## 1.2 Découverte de nouveaux produits naturels

### 1.2.1 Stratégies traditionnelles de découverte de produits naturels.

La modification des conditions de cultures d'un même micro-organisme est une approche permettant d'obtenir des métabolites dont la production est régulée, et est nommée OSMAC (one strain many compounds).

## **1.2.2 Exploration du génome et activation de clusters de gènes silencieux**

Ces approches permettent de produire des métabolites issus de cluster de gènes silencieux.

La comparaison de génomes avec les outils informatiques « genome mining » permet de mettre en évidence des homologies de séquences et donc de trouver des clusters de gènes proches conduisant à une voie de biosynthèse d'analogues à la molécule prise comme référence.

La dérégulation de l'expression des gènes peut se faire par des approches de modifications épigénétiques, par surexpression des facteurs de transcription ou échange de promoteurs.

## **1.2.3 Expression hétérologue des clusters de gènes**

De nombreux systèmes d'expression hétérologue ont été développés pour exprimer des clusters de gènes fongiques. *Escherichia coli* est un hôte idéal pour les systèmes d'expression hétérologue procaryote en raison de son utilisation simple et de son taux de croissance rapide. Cependant, les levures (*Saccharomyces cerevisiae*, *Pichia pastoris*) ou les champignons (*Aspergillus* sp.) sont plus adaptés pour l'expression de clusters de gènes fongiques.

## **1.2.4 Biosynthèse combinatoire et biologie synthétique**

Les analogues de produits naturels ont un grand potentiel pour la découverte de médicaments. En raison de la complexité de leur structure, leur synthèse chimique est souvent compliquée et le rendement insatisfaisant. Les progrès en biologie moléculaire et l'augmentation des séquences de gènes déposées dans les bases de données permettent d'une part de générer de nouvelles molécules par combinaison de gènes de clusters d'origines différentes afin d'obtenir des analogues de produits naturels mais aussi de construire de nouvelles voies de biosynthèse pour créer de nouvelles molécules.

## **1.3 Pyrrocidines et champignons endophytes *Sarocladium zeae***

### **1.3.1 Pyrrocidines et leurs analogues, une famille de composés en expansion**

On trouve dans la littérature de plus en plus d'exemples d'analogues des pyrrocidines, avec des activités variées. Les modifications sont le nombre de méthyl sur le polycétide (hirsutellone B), la stéréochimie (GKK 1032A<sub>2</sub>), la longueur de la chaîne du polycétide ou encore des cyclisations supplémentaires (trichobamide, pencipyrrodiether, pyrrospirone)

### **1.3.2 Origine biosynthétique des pyrrocidines**

La première étude sur la voie de biosynthèse de ce type de métabolite concerne le composé GKK 1032A<sub>2</sub>. Elle a montré l'implication d'une PKS-NRS et d'enzymes auxiliaires pour les cyclisations et oxydations (fig. 1.4).



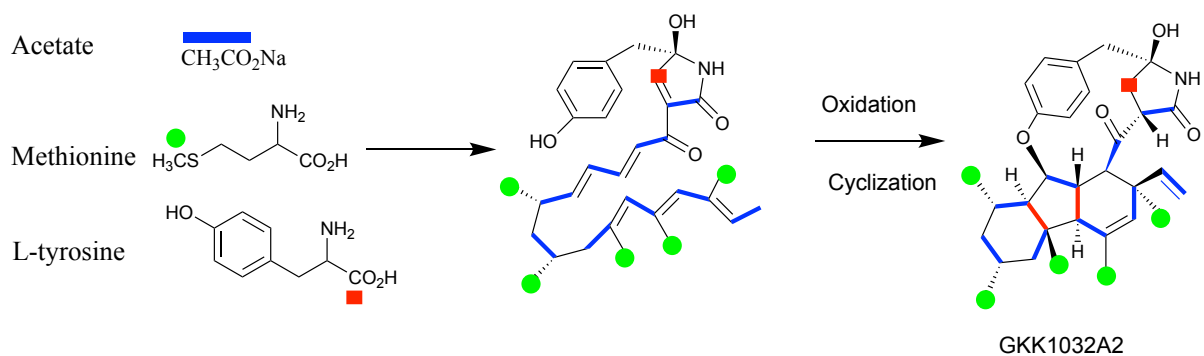


Fig.1.4 Analyse isotopique du GKK 1032A<sub>2</sub> produit par *Penicillium sp.* en présence de substrats marqués.

Il est supposé que la biosynthèse des pyrrocidines suit le même schéma avec la formation de l'aldéhyde par une PKS-NRPS et les six modifications structurales suivantes : (Fig. 1.5): 1) formation du lactame 2) formation de la liaison éther 3) formation de la liaison C7-C12 4) une réaction de Diels-Alder 5) l'hydroxylation en C20 6) réarrangement des doubles liaisons.

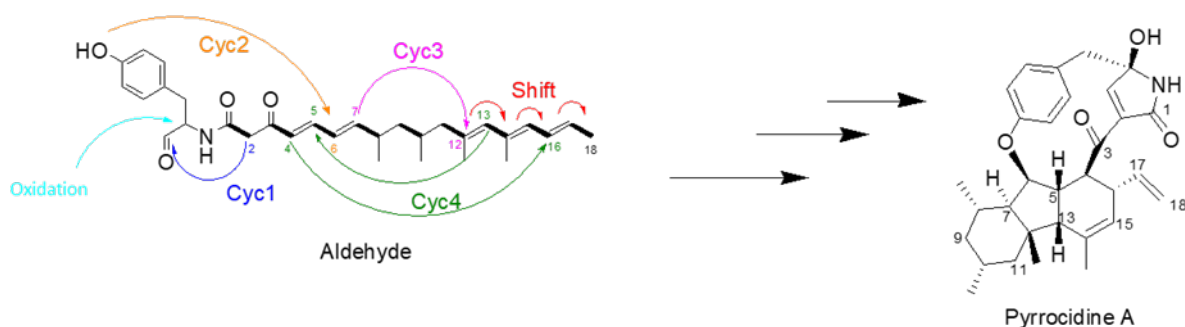


Fig. 1.5 Modifications nécessaires pour passer de l'aldéhyde linéaire à la pyrrocidine A

Une étude de la formation du paracyclophane a été menée au laboratoire en incubant le champignon en présence de tyrosine doublement marquée (<sup>18</sup>O et <sup>13</sup>C). Le double marquage des pyrrocidines obtenues montre que l'oxygène phénolique est incorporé dans les pyrrocidines.(Fig.1.6).

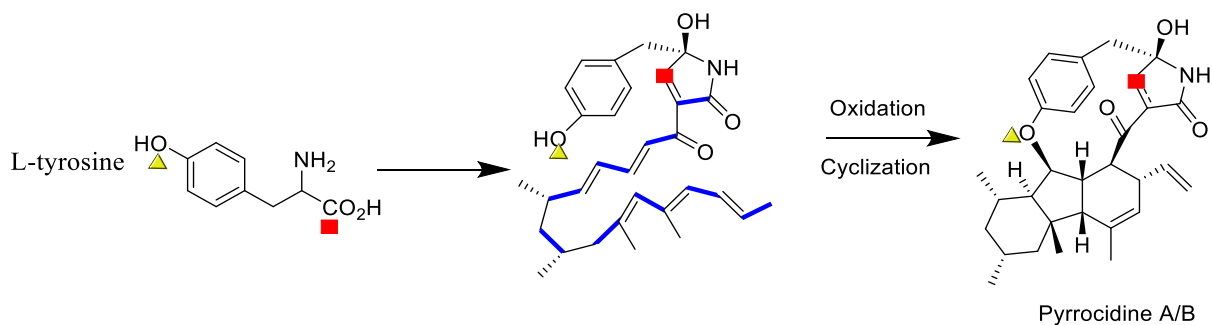


Fig.1.6 Pyrrocidines obtenues lors de culture de *S. zae* en présence de tyrosine marquée.

## 1.4 Objectifs

Ce projet se concentre sur l'élucidation de la voie de biosynthèse des pyrrocidines chez le champignon endophyte *Sarocladium zae*, afin de comprendre les rôles physiologiques des pyrrocidines mais aussi pour obtenir de nouveaux analogues des pyrrocidines. L'objectif principal est de déterminer le rôle biosynthétique des différentes protéines du cluster de gènes et d'obtenir leur expression hétérologue pour une étude pas à pas des mécanismes impliqués.

## Chapitre 2

### Biosynthèse des Pyrrocidines chez *Sarocladium zeae*

#### 2.1 Introduction

#### 2.2 Pyrrocidines and analogs produced by *S.zeae*

##### 2.2.1. Conditions de culture pour la production de pyrrocidines par *S.zeae*

La première étape de l'étude a été de déterminer les meilleures conditions de culture du champignon permettant une production des pyrrocidines pour des analyses par LC/MS et pour l'isolement des métabolites et leur identification par techniques spectroscopiques (RMN, SM). Alors que le champignon est capable de survivre plus de huit semaines sur un milieu composé de grains de maïs, il semble sénescer dans les autres milieux testés après des temps beaucoup plus courts comme le montre les données du tableau 2.1. Les milieux ont été extraits avec de l'acétate d'éthyle pour des analyses par LC-ESI-HRMS.

Tableau. 2.1 Masses des extraits bruts et temps de culture de *S. zeae* sur différents milieux

Medium	Mass of medium	Mass of extract	Time culture (weeks)
PDA	20 g	12 mg	2
Maize	50 g	2 g	8 or more
Rice	50 g	1.1 g	3
Oat	50 g	1.6 g	4
Wheat	50 g	2.4 g	4

Les chromatogrammes sont présentés dans la figure 2.1. Par comparaison des chromatogrammes d'ions extraits (ECI), il semble que la production de pyrrocidines est la plus importante dans le milieu maïs, ce qui est confirmé par la précipitation d'une quantité importante de pyrrocidine A durant la phase de concentration de l'extrait (100 mg à partir de 50 g de maïs). Ce milieu a été choisi pour l'isolement des métabolites. Cependant la production sur milieu PDA est suffisante pour des analyses par LC/MS, et ce milieu sera utilisé pour voir l'effet de l'inactivation des gènes impliqués dans la biosynthèse des pyrrocidines A et B.

##### 2.2.2. Caractérisation des pyrrocidines et analogues produits par la souche sauvage

La purification des métabolites produit par la souche sauvage *S. zeae* cultivée sur maïs a été réalisée par fractionnement sur colonne de silice (« flash » chromatographie) puis par HPLC préparative. Les spectres RMN (COSY, HSQC, HMBC and NOESY) ainsi que les résultats des analyses par

LC-HRMS sont donnés en annexe de la thèse. La figure 2.2 donne les chromatogrammes des différents composés isolés.

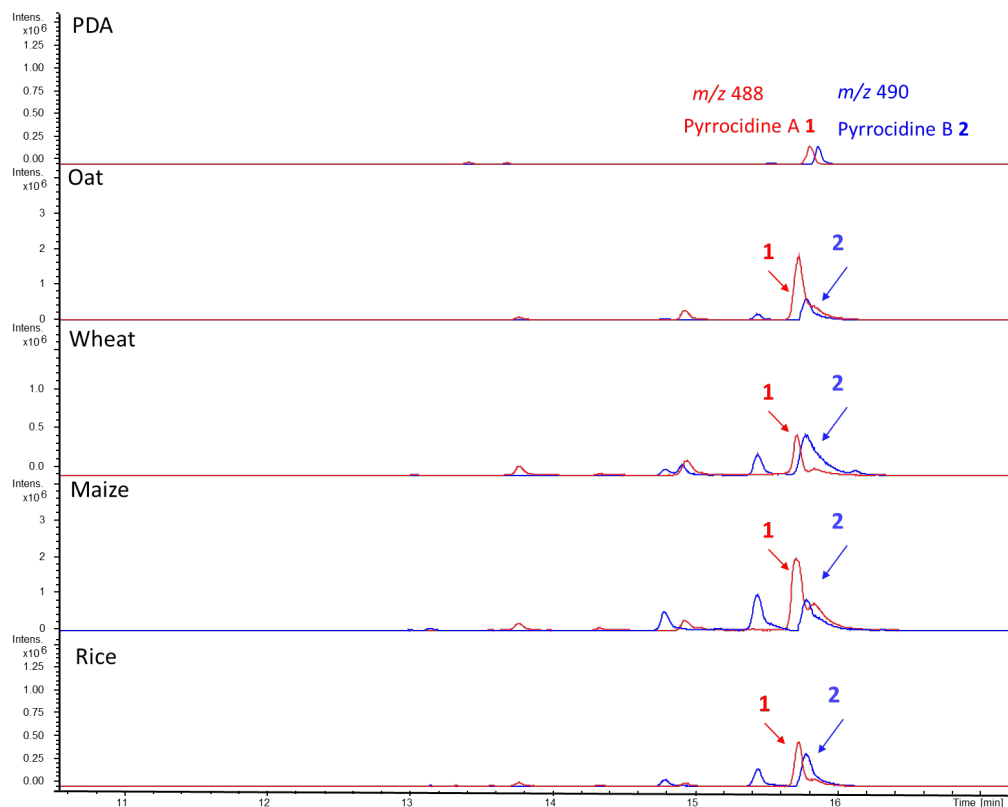


Fig. 2.1: Analyse par LC-MS des extraits bruts des cultures de *S. zeae* sur différents milieux. EIC of  $m/z$  488.2795 (rouge) and  $m/z$  490.2950 (bleu) correspondant à l'ion moléculaire  $[M+H]^+$  des pyrrocidines A and B, respectivement.

A côté de la pyrrocidine B **2**, deux composés de même masse  $m/z$  490.2952 ont été isolés et identifiés comme étant des stéréoisomères. L'un est la forme cétonique de l'isomère en C3 et C6 de **2** et l'autre est sa forme énol moins stable (Fig. 2.3).

Un composé **8** avec  $m/z$  556.3057 pour l'ion  $[M+H]^+$  et un temps de rétention de 16,1 min. a été identifié comme étant un isomère du trichobamide A **5** connu (Fig. 2.4). Le composé **11** dont l'ion  $[M+H]^+$  a une masse  $m/z$  488.276 a été détecté à 15.0 min. Ses spectres RMN sont

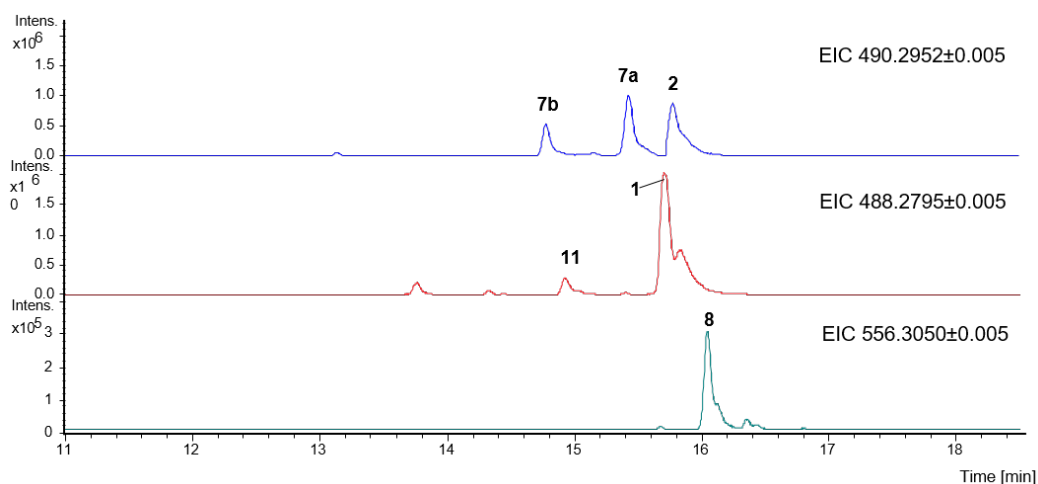


Fig. 2.2: Chromatogrammes d'ions extraits obtenus par analyse LC-MS des extraits métaboliques de la souche sauvage *S. zae* cultivée sur maïs pendant 30 jours.

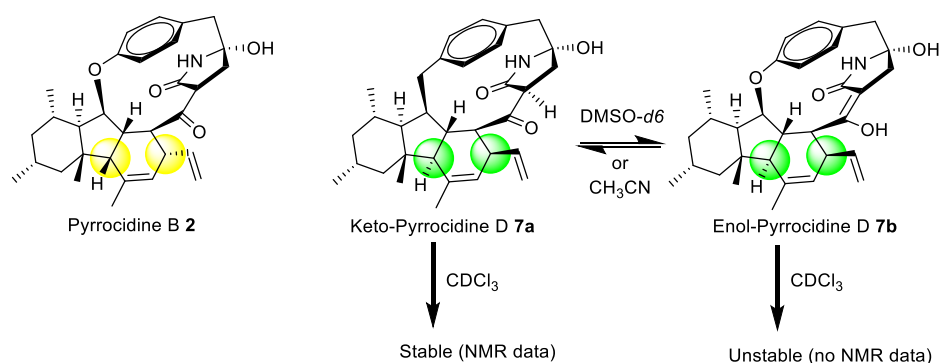


Fig. 2.3 Structures des pyrrocidine B, D, et l'équilibre céto-énolique de la pyrrocidine D

identiques à ceux décrits par Shiono *et al.* pour l'épi-pyrrocidine isolée du champignon *Neonectria ramulariae* In-2. Il est nommé pyrrocidine F et correspond à l'épimère en C20 de la pyrrocidine A (Fig. 2.5). Enfin, la pyrrocidine C 9 avec  $[M+H]^+$  de  $m/z$  504.2744 a aussi été isolée. (Fig. 2.6). Bien que ce composé ait été plusieurs fois mentionné dans la littérature décrivant sa formation spontanée à partir de la pyrrocidine A, sa caractérisation par RMN n'a pas encore été publiée.

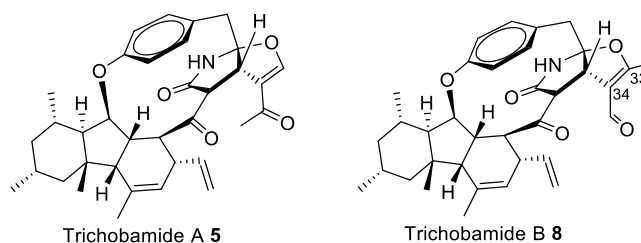


Fig. 2.4 Structures des trichobamides A (connu) and B (nouveau)

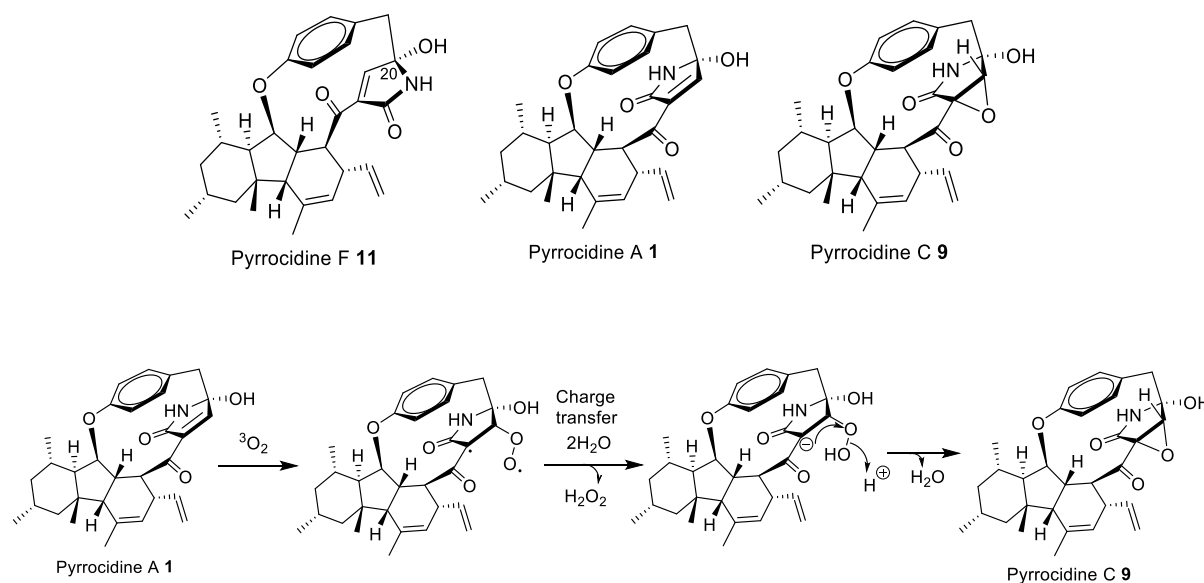


Fig. 2.5 Structures des pyrrocidines F, A et C, et proposition de mécanisme pour la formation spontanée de la pyrrocidine C 9.

### 2.3 BGCs avec PKS-NRPS dans le génome de *S. zeae*

Le génome de *S. zeae* a été séquencé par notre équipe. Trois gènes codant pour des PKS-NRPS putatives ont été identifiés par BLAST en utilisant la PKS-NRPS TenS comme référence. Ces trois gènes se trouvent sur trois *scaffolds* différents et nommés *praA*, *prbA* et *prcA* (fig. 2.6)

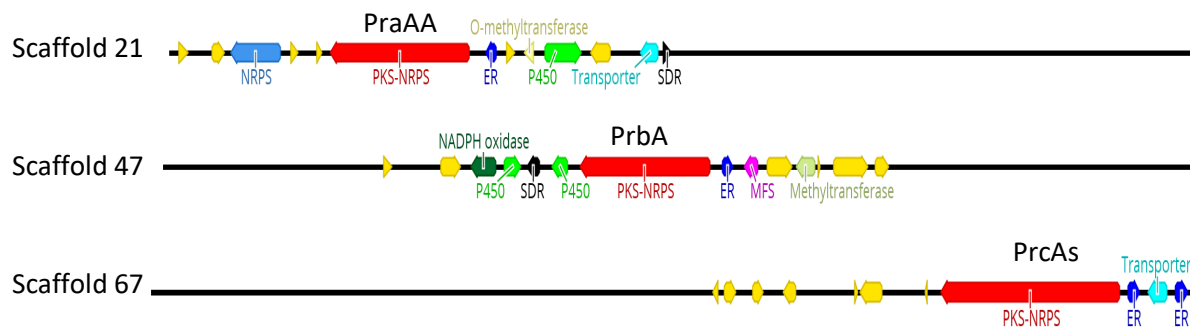


Fig. 2.6 Les trois gènes candidats pouvant coder la PKS-NRPS impliquée dans la biosynthèse des pyrrocidines chez *S. zeae*. Fonctions identifiées des protéines codées dans ces clusters de gènes : PKS-NRPS en rouge, NRPS en bleu, ER en bleu foncé, P450 en vert, Transporteur en bleu ciel, O-methyltransférase en couleur or, SDR en noir, MFS en rose, Methyltransférase en vert clair, NADPH oxydase en vert foncé.

Les trois gènes ont été analysés en utilisant des outils de prédiction (NCBI BLAST et PKS/NRPS analysis) afin d'identifier les différents domaines. Nous avons tenté de déterminer le gène impliqué par comparaison avec ce qui est connu sur les PKS-NRPS.

D'abord, nous avons recherché la possibilité d'obtenir des informations de prédiction sur le domaine A puisqu'il sélectionne l'acide-amino incorporé dans le métabolite. Nous n'avons pas obtenu de correspondance parfaite entre l'un des domaines A de *S. Zeae* et ceux connus.

Ensuite, nous avons examiné le domaine R qui est en charge du relargage de l'intermédiaire qui peut se faire suivant deux mécanismes : avec réduction pour libérer l'aldéhyde qui se cyclise suivant une réaction de Knoevenagel pour former la pyrrolidinone ou sans réduction (domaine R incompetent), le  $\beta$ -cétamide se cyclise *via* une réaction de Dieckmann en pyrrolidine-2,4-dione. La présence du cycle pyrrolidone dans la pyrrocidine A suggère que la mégasynthase a un domaine réducteur avec une triade fonctionnelle Ser-Tyr-Lys. Le domaine R de PrbA présente une mutation dans la triade catalytique (Tyr en Leu). Par conséquent, PrbA a été exclu de la participation à la biosynthèse des pyrrocidines. Les deux autres PraA et PrcA ont une triade intacte et pourraient potentiellement être impliqués.

Finalement, nous avons constaté qu'un fragment d'ADN de 2,5 kb sans prédiction de domaine était adjacent au gène *PrcA* (Fig. 2.8). Une analyse plus approfondie à l'aide du logiciel FGENESH a révélé la présence d'un gène dans cette séquence de 2,5 kb nommée *prcB* et codant une protéine à fonction Diels-Aldérase putative. Ce gène pourrait être nécessaire en raison de la présence d'un cyclohexène dans les pyrrocidines. Par conséquent, le cluster de gènes 67 serait le meilleur candidat pour la biosynthèse des pyrrocidines.

## 2.4 Identification de la PKS-NRPS produisant les pyrrocidines

Afin de valider notre hypothèse, nous avons inactivé par délétion indépendamment les trois gènes codant les PKS-NRPS chez *S. zeae* et analysé l'effet de ces délétions sur la production de pyrrocidines. Pour obtenir les mutants souhaités, nous avons mis au point une méthode de transformation du champignon *S. zeae* par recombinaison homologe d'une cassette d'inactivation *via* la transformation de protoplastes traités avec du PEG/CaCl<sub>2</sub>.

Une étude avait montré que l'hygromycine est active sur *S. zeae* et nous avons donc utilisé le gène de résistance *hph* pour sélectionner les mutants. Les cassettes, comprenant le gène de résistance à l'hygromycine flanqué de part et d'autre d'une séquence d'ADN (1kbase) correspondant aux séquences amont et aval du gène ciblé ont été assemblées chez *E. coli*, amplifiées et transférées dans le champignon. Cette stratégie a permis d'obtenir les trois mutants souhaités.

Ces mutants ont été cultivés et les extraits métaboliques à l'acétate d'éthyle ont été analysés par LC/MS (Fig. 2.7). Seul le mutant au niveau du cluster *prc* n'est plus producteur de pyrrocidines alors que le dihydroresorcylide est toujours produit dans les mêmes quantités. On peut donc conclure que la PKS-NRPS du cluster *prc* est impliquée dans la biosynthèse des pyrrocidines.

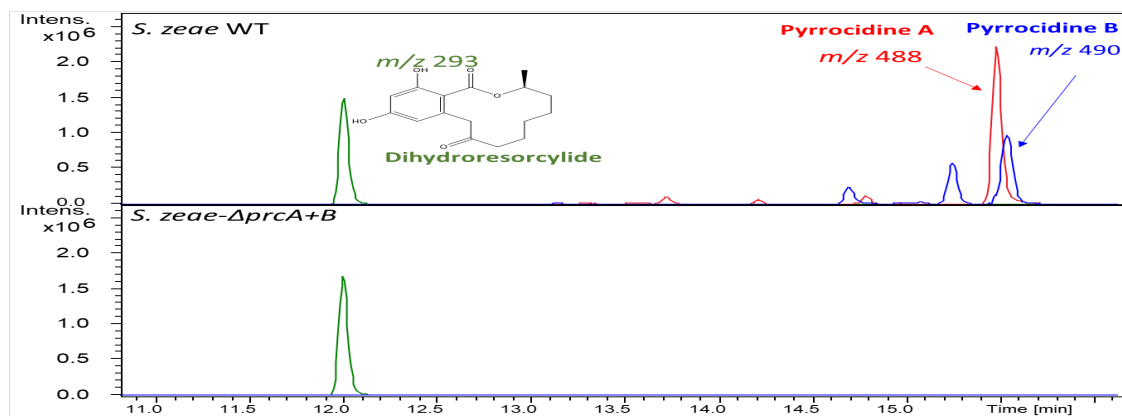
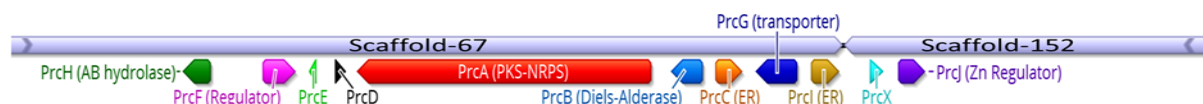


Fig. 2.7 Analyses par LC-MS des extraits bruts des cultures de la souche sauvage *S. zeae* et du mutant *S. zeae-ΔprcA+B*. Chromatogrammes d'ions extraits de  $m/z$  293.137  $[M+H]^+$  (dihydroresorcyllide) utilisé comme contrôle,  $m/z$  488.277  $[M+H]^+$  pyrrocidine A et  $m/z$  490.293  $[M+H]^+$  pyrrocidine B).

## 2.5 Annotation du BGC des pyrrocidines

La prédiction des séquences codantes (CDS) entourant le gène *prcA* avec le logiciel FgenesH a révélé neuf protéines putatives codées par des gènes annotés de *prcH* à *prcI* (Fig. 2.8). À partir de l'analyse des séquences, nous avons identifié PrcC comme une ényol réductase (ER) associée à l'hybride PKS-NRPS PrcA, PrcH comme une  $\alpha/\beta$ -hydrolase, PrcF comme facteur de transcription, PrcG comme transporteur, PrcI comme deuxième ER et PrcB homologue aux Diels-Aldérase, ainsi que PrcD et PrcE apparaissant comme des protéines membranaires intégrales à fonction inconnue (Fig 2.8).



Lane	Size (bp)	Predicted function	Percent identity/known function	NCBI accession/species
<i>prcA</i>	12063	PKS-NRPS	78%/FsdS	GJC90835.1/ Colletotrichum spaethianum
<i>prcB</i>	1299	Diels-Aldérase	77%/CcsF	GJC90834.1/Colletotrichum spaethianum
<i>prcC</i>	1188	Enoyl reductase	76%/Ccsc	GJC90833.1/Colletotrichum spaethianum
<i>prcD</i>	489	Unknown	69%/unknown	XP_018074793.1/[Phialocephala scopiformis]
<i>prcE</i>	375	Unknown	87%/unknown	KXH65045.1/[Colletotrichum salicis]
<i>prcF</i>	1491	Regulator	53%/Acell	XP_003069900.1/[Coccidioides posadasii ]
<i>prcG</i>	1846	Transporter	81%/Fus6	GJC90832.1/Colletotrichum spaethianum
<i>prcH</i>	1257	A/B hydrolyse	83%/CcsE	GJC90837.1/ Colletotrichum spaethianum
<i>prcI</i>	1098	Enoyl reductase	32%/PoxH	A0A1W5T4X6.1/Penicillium oxalicum
<i>prcJ</i>	1317	Zn regulator	21%/Fus10	XP_007826306.1/[Metarhizium robertsii ARSEF 23]
<i>prcX</i>	555	Unknown	83%/unknown	KXH56094.1/[Colletotrichum salicis]

Fig. 2.8 Annotation du cluster de gènes des pyrrocidines et fonctions putatives des protéines.

## 2.6 Identification des fonctions des enzymes auxiliaires par inactivation des gènes



### 2.6.1 Construction de mutants de délétion des enzymes auxiliaires

Les mutants des gènes d'enzymes auxiliaires ont été construits selon la même méthode que celle pour les PKS-NRPS. Après culture, les métabolites extraits de ces mutants ont été analysés par LC/MS (fig. 2.9).

### 2.6.2 Analyse de l'effet de l'inactivation du gène codant pour PrcI

L'analyse par LC/MS de l'extrait de la culture du mutant *S. zea*- $\Delta prcI$  montre que ce mutant ne produit plus les pyrrocidines B **2**, et D **7a** et **7b** alors que la pyrrocidine A **1** est toujours produite. On peut conclure que PrcI est responsable de la formation de la pyrrocidine B par réduction de la double liaison de la pyrrolidone de la pyrrocidine A (fig 2.9) .

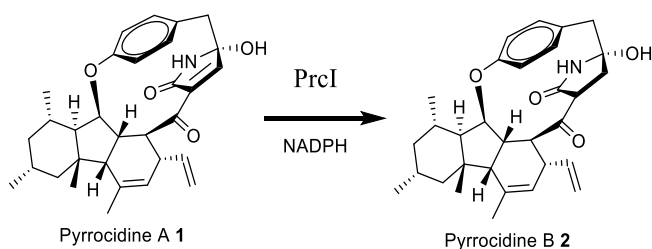


Fig. 2.23 Réduction de la pyrrocidine A en pyrrocidine B par PrcI.

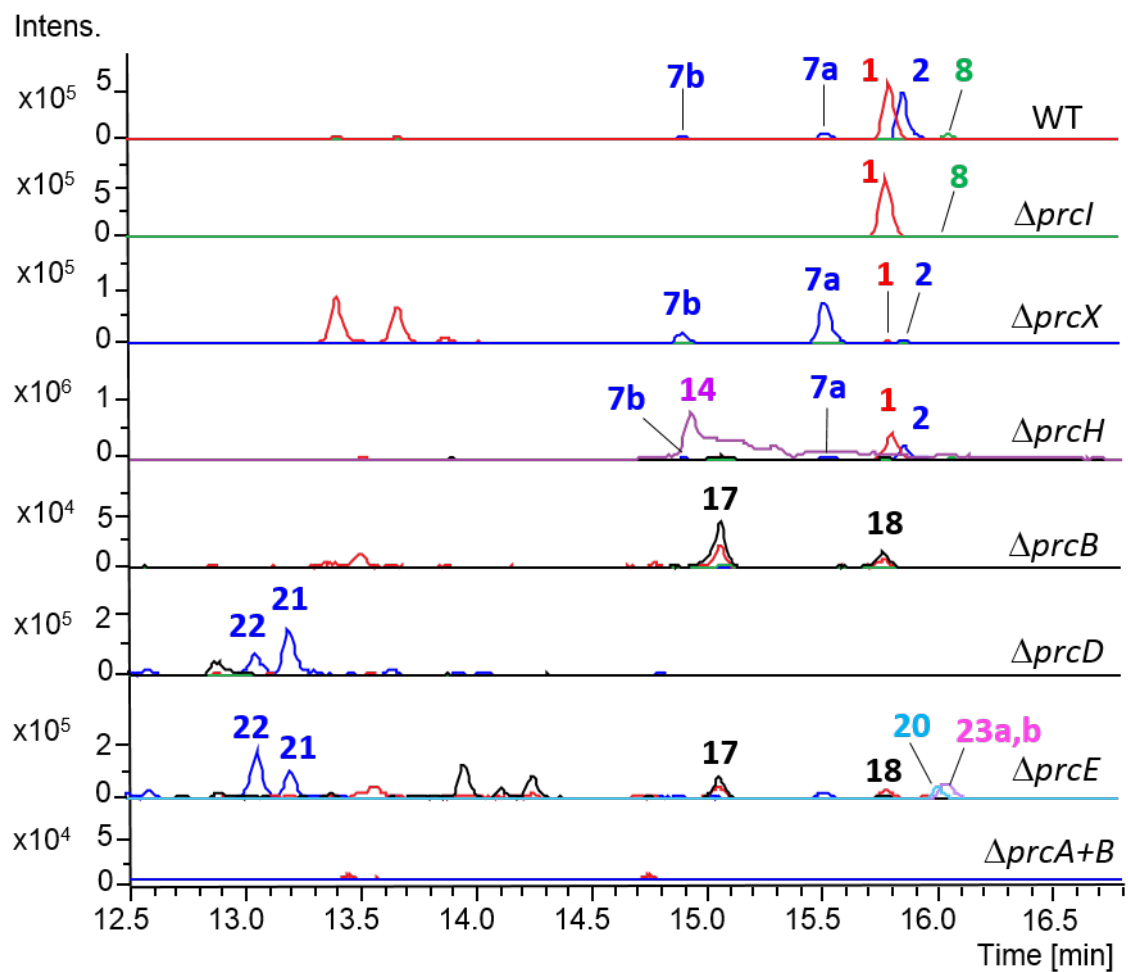


Fig. 2.9 Analyses par LC-MS des extraits bruts des cultures de la souche sauvage *S. zeae* et des différents mutants. Chromatogrammes d'ions extraits correspondants aux métabolites accumulés.

### 2.6.3 Analyse de l'effet de l'inactivation du gène codant pour PrcX

La suppression de *prcX* donne lieu à une diminution substantielle de la production de **1** et **2** sans impact sur la production des pyrrocidines D **7a,b** (Fig.2.10). Ce résultat montre que **1** et **2** sont naturellement produits sans cette lipocaline.

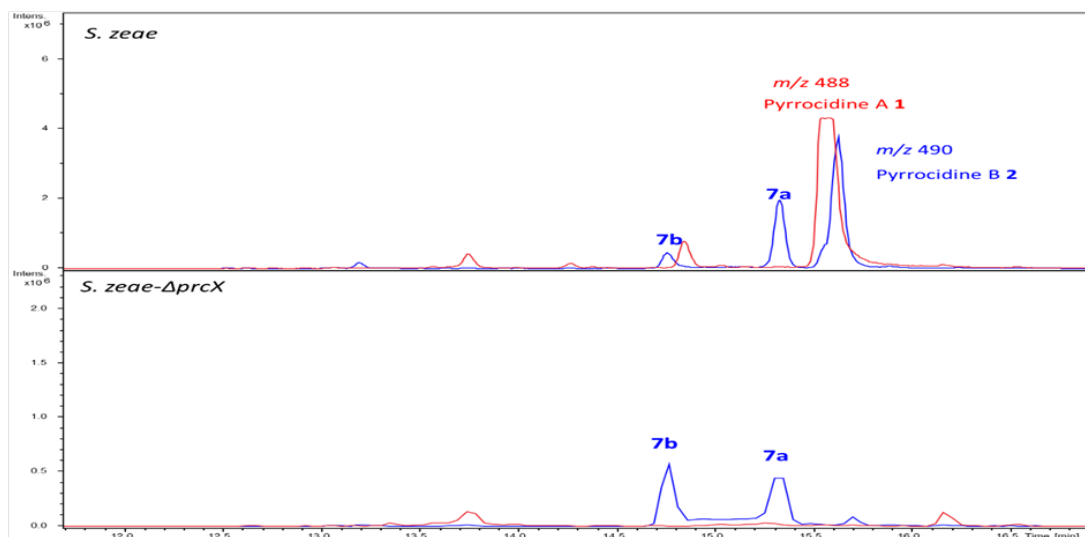


Fig. 2.10 Analyses par LC-MS des extraits bruts des cultures de de la souche sauvage *S. zeae* et du mutant *S. zeae-ΔprcX*. Chromatogrammes d'ions extraits correspondants aux pyrrocidines avec  $m/z$  488.277  $[M+H]^+$  pyrrocidine A et  $m/z$  490.293  $[M+H]^+$  pyrrocidine B.

La diastéréomérisation observée entre **2** et **7a/b** en C3/C6 peut s'expliquer par la réaction IMDA formant le cycle C à partir du précurseur **25** (Fig. 2.11). En effet, le diène et le diénophile de **25** peuvent réagir selon les approches *exo* ou *endo* pour donner respectivement **1** ou **27**. Nous en avons déduit que PrcX entraîne la cycloaddition de Diels-Alder en faveur du produit *exo*.

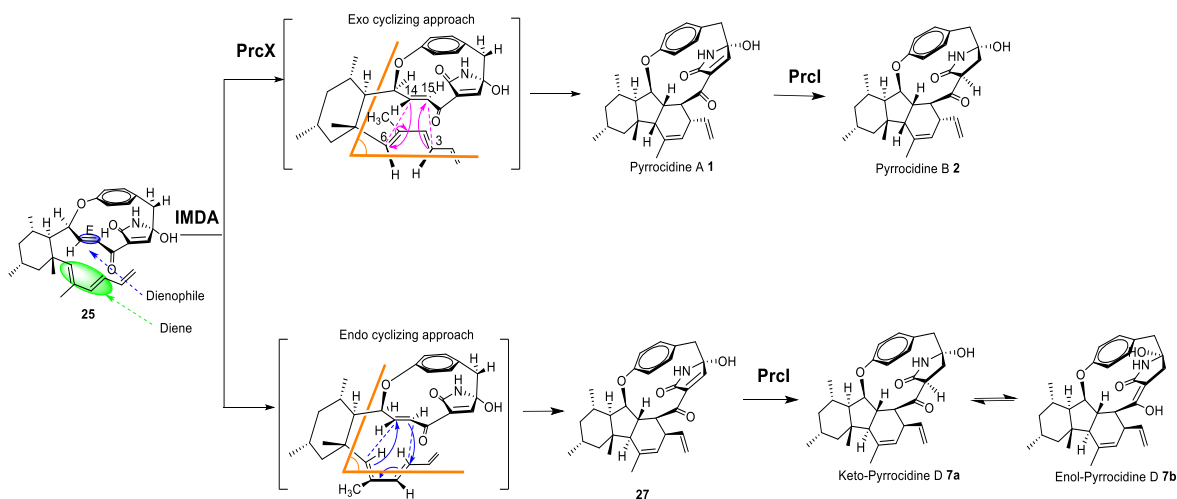


Fig. 2.11. Formation des pyrrocidines *via* IMDA à partir du précurseur **25**. La cycloaddition du diène et du diénophile suivant une approche *exo* conduit à la formation des pyrrocidines **A 1** et **B 2**. Une approche *endo* conduit à la formation d'un précurseur non isolé **27** qui est réduit en pyrrocidines **D 7a,b**.

## 2.6.4 Analyse de l'effet de l'inactivation du gène codant pour PrcH

Le gène codant pour PrcH (annotée  $\alpha/\beta$ -hydrolase) a été inactivée et le mutant *S. zeae- $\Delta$ prcH* a été obtenu. La comparaison des spectres UV (Fig. 2.12) montre qu'un nouveau métabolite est produit par ce mutant.

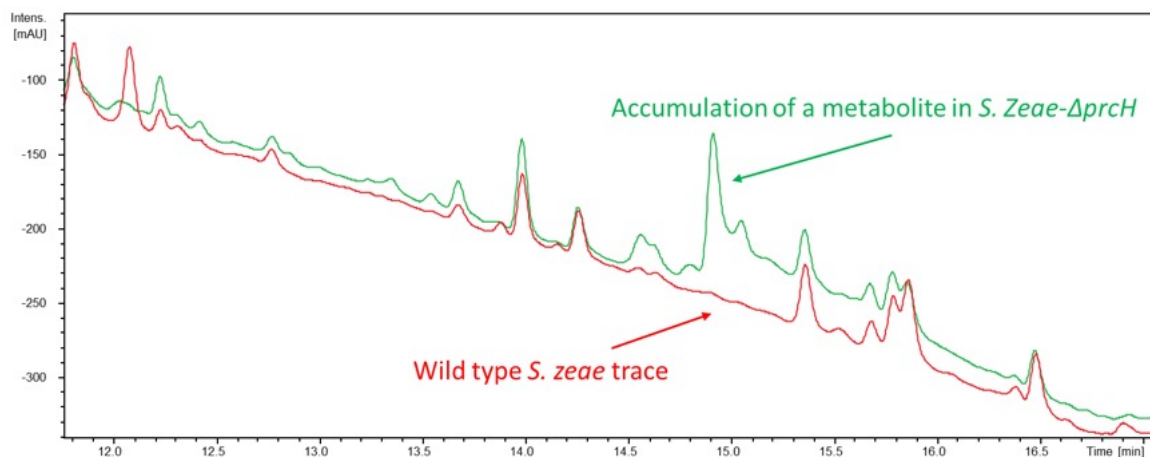


Fig. 2.12 Chromatogramme UV (200-500nm) des extraits bruts d'une culture de *S. zeae* (en rouge) et du mutant *S. zeae- $\Delta$ prcH* (en vert)

Ce métabolite avec une masse pour l'ion  $[M+H]^+$   $m/z$  494.3297 et de formule moléculaire  $C_{31}H_{43}NO_4$  a été identifié comme étant l'alcool linéaire **14** (Table 2.12). La structure linéaire est déduite grâce aux corrélations observées sur les spectres HMBC et COSY. Les constantes de couplage  $^3J$  pour les protons éthyléniques indiquent que les doubles liaisons carbone-carbone sont toutes de configuration *E* ( $^3J = 15.5$  Hz pour  $C_{15}=C_{14}$ ,  $^3J = 15.2$  Hz pour  $C_{13}=C_{12}$ ,  $^3J = 15.0$  Hz pour  $C_3=C_2$ ). On peut suggérer que ce composé linéaire dérive directement du produit relargué par la PKS-NRPS.

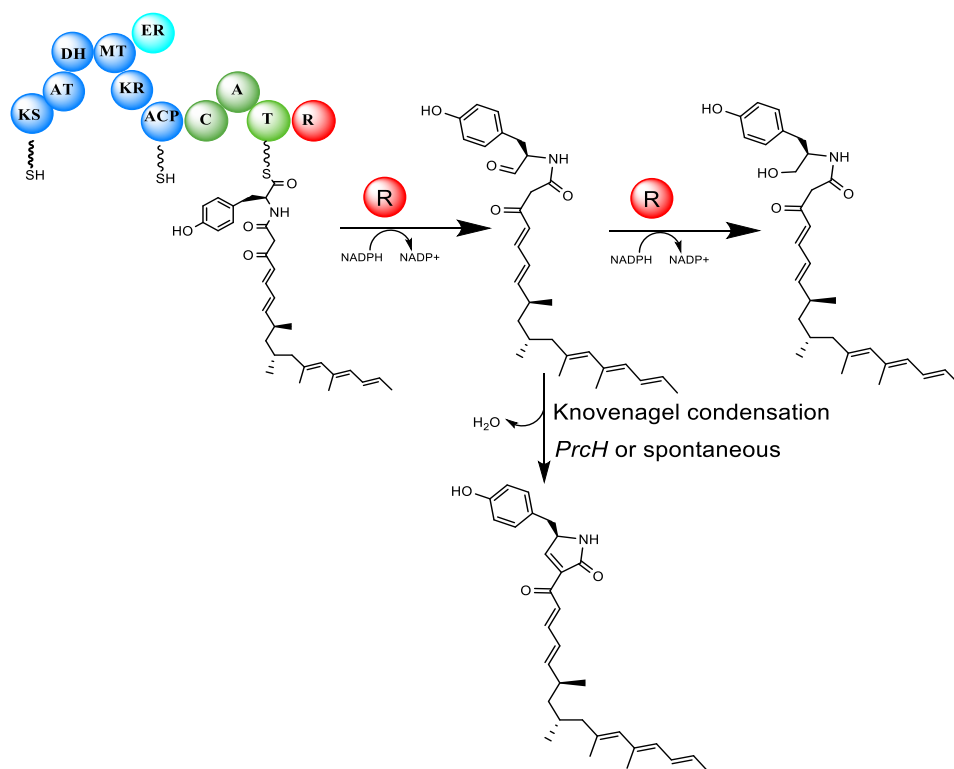


Fig. 2.18 Fonction proposée pour PrcH

PrcH est homologue à des  $\alpha/\beta$ -hydrolases associées à des PKS-NRPS et leur inactivation conduit comme ici à l'accumulation de l'alcool. Ces produits n'entrent pas dans la voie de biosynthèse et sont appelés produits de shunt (Fig. 2.18).

### 2.6.5 Analyse de l'effet de l'inactivation du gène *PrcB*, *PrcD* et *PrcE*.

Les trois gènes *prcB*, *D* and *E* ont été inactivés individuellement, et les mutants obtenus ne produisent plus de pyrrocodines mais de nouveaux métabolites sont formés, ce qui montre l'implication de ces gènes dans la biosynthèse.

La culture du mutant  $\Delta prcB$  a permis d'isoler les composés **17** et **18** avec pour l'ion  $[M+H]^+$  une  $m/z$  506.289 et élués à 14.9 et 15.8 min, respectivement (Figure 2.19). Le composé **17** est une 5-hydroxy-3,4-epoxy-pyrrolidone substituée par une chaîne polyoléfinique comme le composé **14**. Le composé **18** est lui une décaline avec une jonction *cis* et a été nommé pyrrocaline A. Sa formation peut être expliquée par une IMDA de **17** suivant une approche *exo* (fig.2.20).

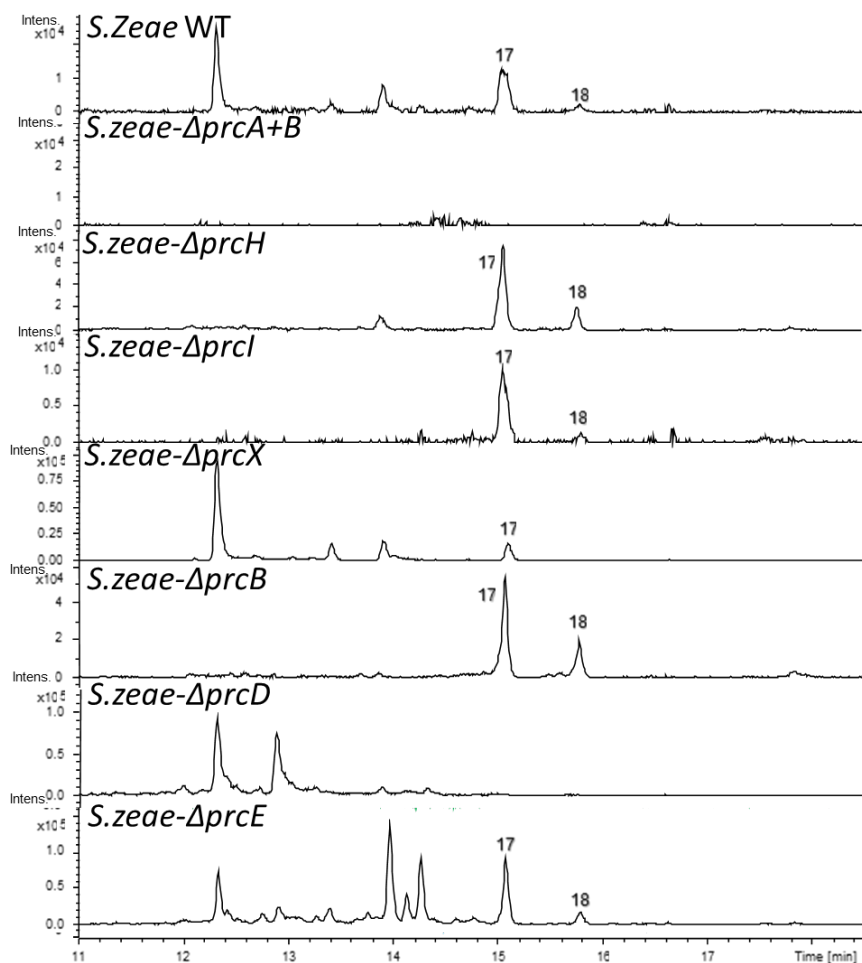


Fig. 2.19 Analyses par LC-MS des extraits bruts des cultures de la souche sauvage *S. zeae* et de différents mutants. Chromatogrammes d'ions extraits correspondants aux métabolites **17** et **18**.

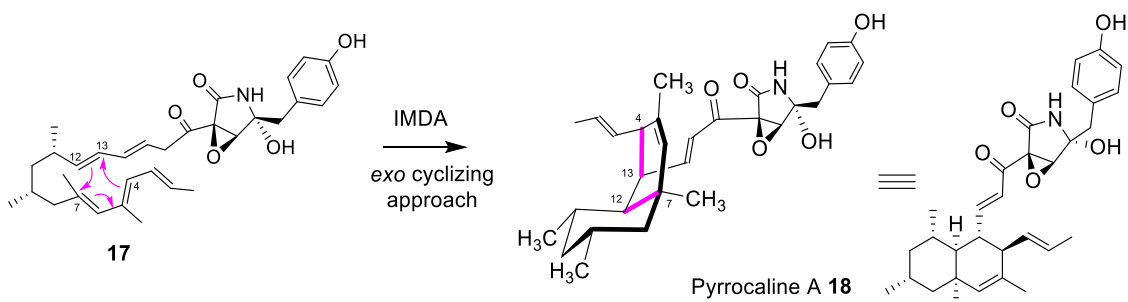


Fig. 2.20: Structure des composés **17** et **18** et mécanisme proposé pour la formation de la cis-décaldine **18** à partir du polycétide **17**.

L'analyse LC-MS du milieu de culture du mutant  $\Delta prcE$  a permis de mettre en évidence à 16.0 min un ion  $[M+H]^+$  de  $m/z$  476.3146 correspondant à la formule moléculaire  $C_{31}H_{41}NO_3$  (Fig. 2.20) tandis que l'analyse par RMN révèle la présence de composés **23a,b** dans un ratio 1:1. L'élucidation de leur structure montre qu'il s'agit du triméthyl-cyclohexane, correspondant au cycle A des pyrrocidines,

substitué par des oléfines en position C7 et C12. Contrairement aux composés **14** et **17**, **23a,b** montrent un déplacement des double liaisons sur la chaîne au niveau du carbone 6.

Ce mutant produit aussi le composé **20** détecté à 15,8 min. dont la formule est  $C_{31}H_{39}NO_3$  (déterminée d'après l'ion  $[M+H]^+$   $m/z$  474.2998). L'analyse RMN révèle une structure polycyclique unique 6/10/6/5 impliquant le résidu d'une tyrosine. Il a été nommé pyrrochalsine A et représente le premier membre nonacétide-tyrosine de cette famille.

Les chromatogrammes des analyses LC/MS des extraits des mutants  $\Delta prcE$  et  $\Delta prcD$  montrent la présence de deux autres composés à 13.2 et 13.0 min (fig. 2.9) et ayant la même formule moléculaire  $C_{31}H_{39}NO_4$  (attribuée à partir de  $m/z$  490.295 pour l'ion moléculaire  $[M+H]^+$ ). Ces composés nommés pyrrocyclines A **21** and B **22** présentent une structure unique avec un macrocycle à 15 carbones fusionné à des cycles à 6 et 5 chaînons (fig. 2.21).

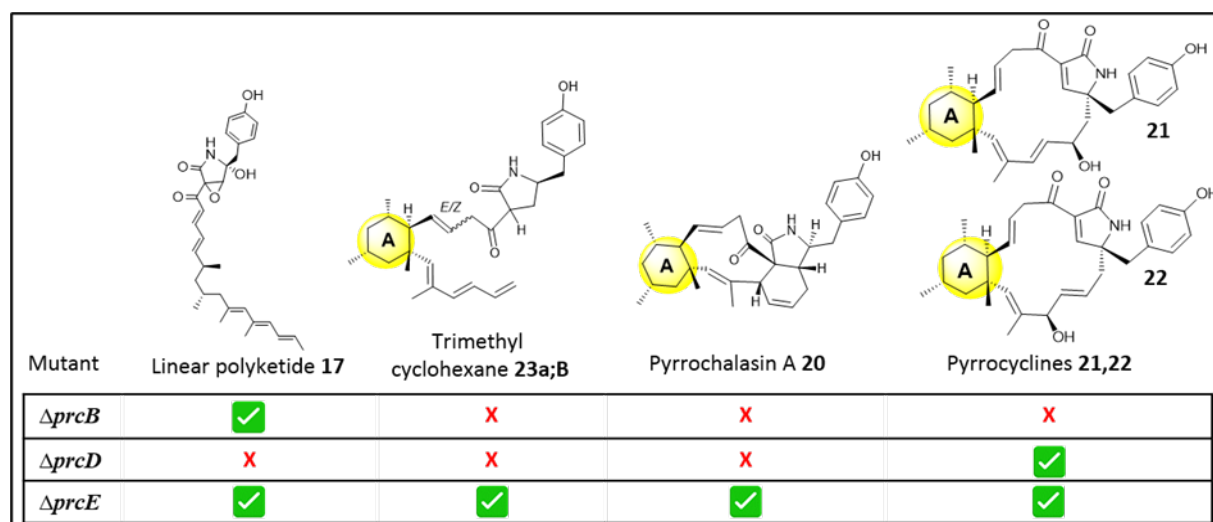


Fig. 2.21 Métabolites produits par les mutants  $\Delta prcB$ ,  $\Delta prcD$  and  $\Delta prcE$ .

Alors que des composés avec un cyclohexane triméthylé sont observés avec les mutants  $\Delta prcE$  et  $\Delta prcD$ , celui-ci n'est pas observé dans les composés isolés de la culture du mutant  $\Delta prcB$ , ce qui nous permet de conclure que PrcB est responsable de la formation du cycle A avec une étape d'isomérisation-cyclisation du produit linéaire **15** en intermédiaire clé **19**. Ce dernier peut donner les différents composés isolés **20**, **21** et **22** comme le montre la figure 2.22.

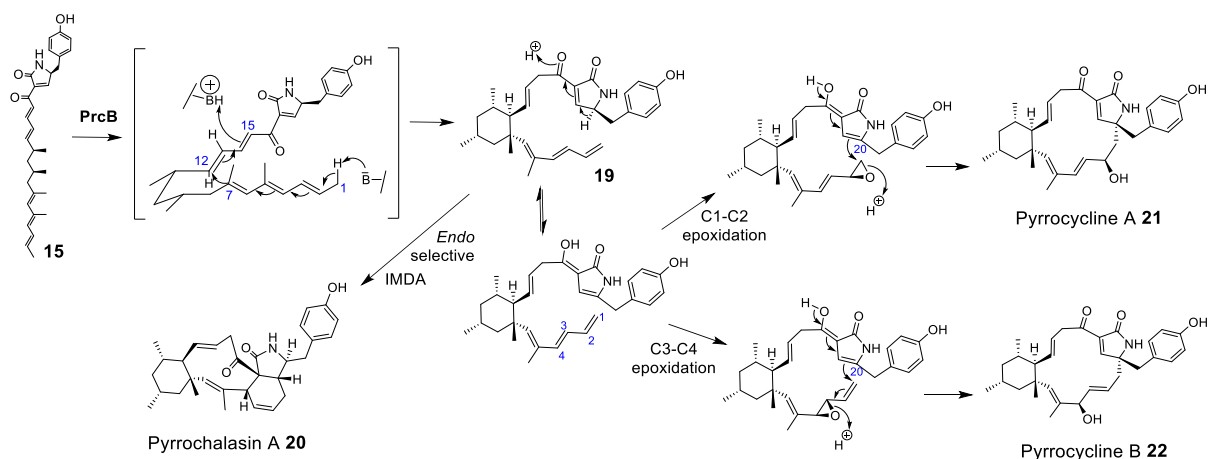


Fig. 2.22: Voie proposée pour la formation des composés polycycliques **20**, **21** et **22** à partir de l'intermédiaire commun **19**.

Enfin, les composés **17** et **18** ne sont pas observés dans la culture du mutant  $\Delta prcD$ , ce qui suggère que PrcD est impliquée dans l'oxydation de la pyrrolidone **15** en **16**. L'époxydation pouvant être une réaction spontanée donne **17**, puis une IMDA conduit à la formation de **18** (fig. 23).

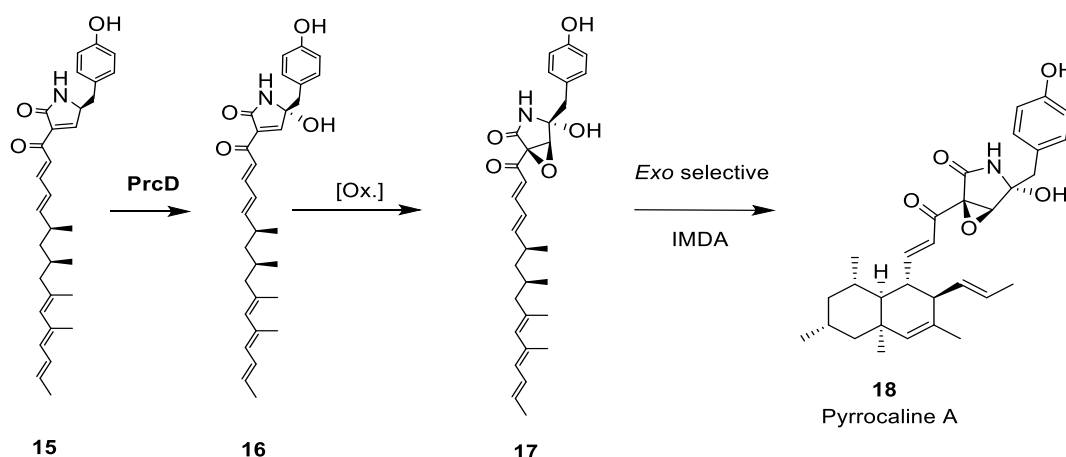


Fig. 2.23 Voie proposée pour la formation de **17** et **18**

## 2.7 Proposition d'une voie de biosynthèse des pyrrocidines

A partir de cette étude, nous proposons cette voie de biosynthèse pour les pyrrocidines :

- En utilisant comme substrat un acétyl-CoA, huit malonyl-CoA, quatre S-Adénosyl-méthionines et une tyrosine, la PKS-NRPS (PrcA) et son ényol-réductase (PrcC) produisent l'aldéhyde **13**.
- Celui-ci subit ensuite une condensation de Knoevenagel pour donner la pyrrolidone intermédiaire **15**, sous le contrôle de PrcH (protection de l'aldéhyde et/ou la catalyse).
- Le cycle A des pyrrocidines est formé par PrcB, et le mécanisme de cyclisation conduit au réarrangement des doubles liaisons pour donner **19**.
- L'hydroxylation du cycle pyrrolidone est catalysée par la protéine membranaire PrcD pour donner **24**.



- La deuxième protéine membranaire (PrcE) catalyse la formation du paracyclophane.
- Une réaction de Diels-Alder intramoléculaire avec une approche *exo* dirigée par PrcX forme les cycles B et C pour donner la pyrrocidine A **1**.
- Enfin la réduction de **1** en pyrrocidine B **2** est réalisée par PrcI.

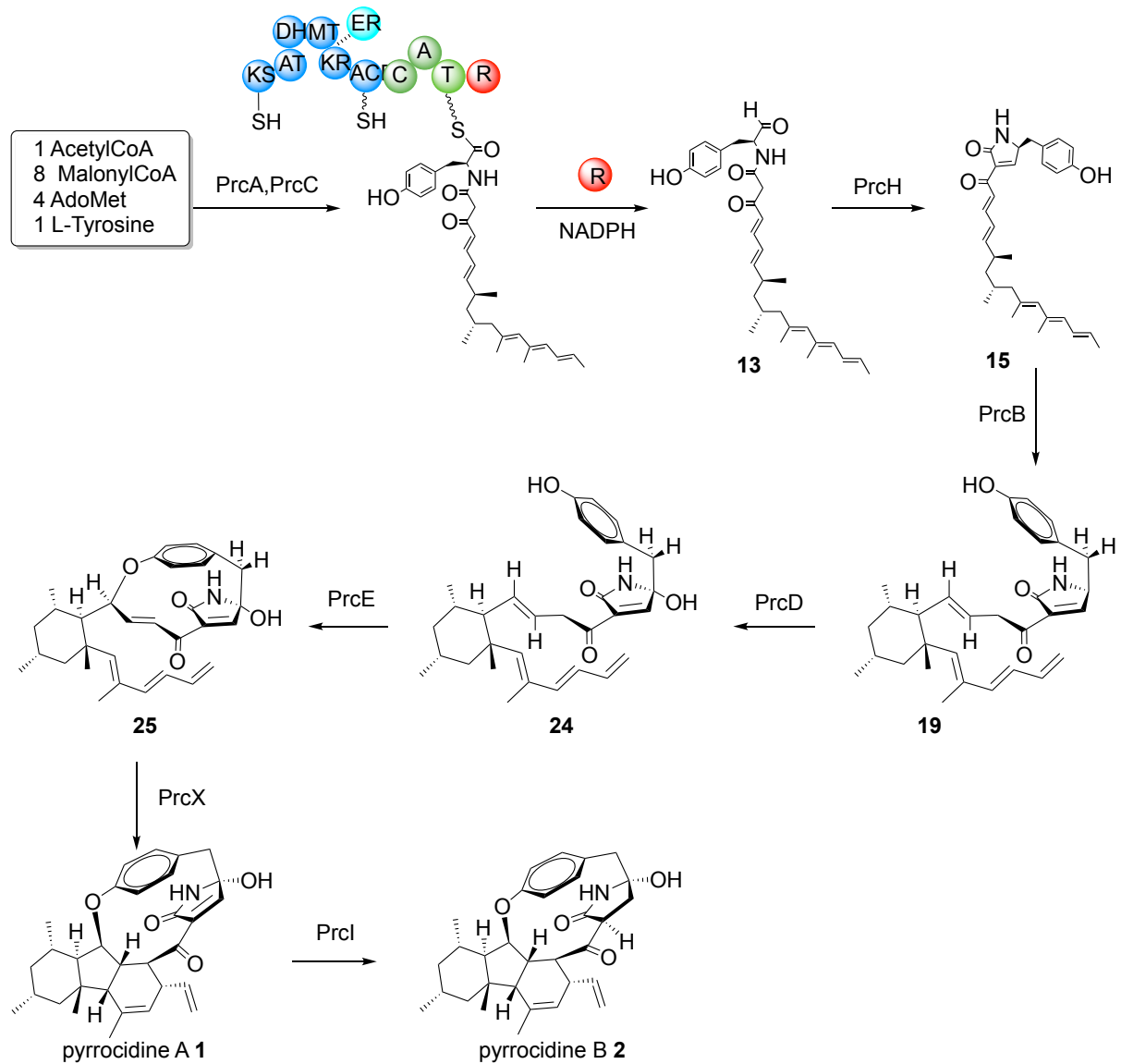


Fig. 2. Voie de biosynthèse proposée pour les pyrrocidines A et B

Pendant ce travail, deux études ont été publiées et proposent des voies de biosynthèse pour la pyrrocidine A et pour un analogue. Nos résultats montrent certaines incohérences de ces propositions.

## **2.8 Evaluation de l'effet de pyrrocidine durant l'interaction entre *S. zeae* and *Fusarium verticilloides***

Plusieurs études ont montré que *S. zeae* avait un effet antagoniste sur *F. verticilloides*, producteur de fumonisine. Le rôle des pyrrocidines a été mis en évidence dans cette interaction.

Récemment, il a été montré que les pyrrocidines pouvaient entraîner une diminution de la production de fumonisine même à des concentrations sub-fongitoxiques.

Nous avons confronté la souche sauvage *S. zeae* et le mutant  $\Delta prcA-B$ , qui ne produit pas de pyrrocidine, avec trois souches de *F. verticilloides*. Les résultats préliminaires montrent que les pyrrocidines semblent jouer un rôle dans ces interactions fongiques.

## **2.9 BGC des Pyrrocidines chez d'autres champignons : « Genome mining »**

Nous avons recherché dans les bases de données publiques qui comptent plus de 2 000 génomes fongiques, des BGC similaires à celui des pyrrocidines afin de trouver de nouveaux métabolites analogues avec des activités biologiques mais aussi de nouvelles enzymes auxiliaires.

En utilisant les enzymes auxiliaires de *S. zeae* comme références nous avons trouvé de nombreux candidats et après annotation, quatorze BGC ont été identifiés comme des clusters apparentés à celui des pyrrocidines. Huit ont été identifiés dans des souches disponibles commercialement et ont pu être étudiés suivant l'approche OSMAC.

L'analyse préliminaire par LC-MS et LC-MS/MS a montré qu'aucun d'entre eux ne produisait de pyrrocidines.

## Chapter III

# Reconstitution du chemin de biosynthèse des pyrrocidines

### 3.1 Introduction

Dans le chapitre précédent, nous avons montré la diversité des composés produits par le BGC des pyrrocidines et identifié les enzymes auxiliaires impliquées. Pour aller plus loin nous avons envisagé de reconstruire le schéma de biosynthèse étape par étape. Le choix de l'hôte est crucial pour obtenir une production. Parmi les différentes possibilités nous avons choisi de réaliser la reconstitution chez *Saccharomyces cerevisiae* qui a déjà été employée pour l'expression de PKS et PKS-NRPS et la production de métabolites.

### 3.2 Clonage des gènes *prcC* codant l'ER et *prcA* codant la PKS-NRPS

De nombreux outils génétiques sont disponibles pour cette levure et particulièrement la stratégie de clonage modulaire MoClo. Le principe de MoClo repose sur l'unicité de l'enzyme de restriction de type II (ici BsaI). Si le gène cible ne contient pas le site de restriction de BsaI, il suffit de le préparer par PCR. Lorsque le gène cible est excessivement long, il contient inévitablement des sites de restriction BsaI, ce qui était le cas pour *prcA* (12 kbases). En revanche, *prcC* n'héberge pas la séquence de restriction de BsaI. Par conséquent, nous avons construit le vecteur d'expression avec *prcC* par MoClo, puis remplacer *prcC* par *prcA* par assemblage *in vivo* (IVA) chez *E. coli*.

### 3.3 Co-expression de PrcA et PrcC chez *S. cerevisiae*

Différentes conditions de fermentation ont été testées pour la production de métabolites par la levure transformée. Fort de notre expérience sur le mutant *S. zeae-ΔprcH*, nous savions que le produit de PrcA-PrcC pouvant être modifié par d'autres enzymes, nous avons donc recherché la formation de l'aldéhyde mais aussi de la pyrrolidone et des produits de réduction (fig. 3.1).

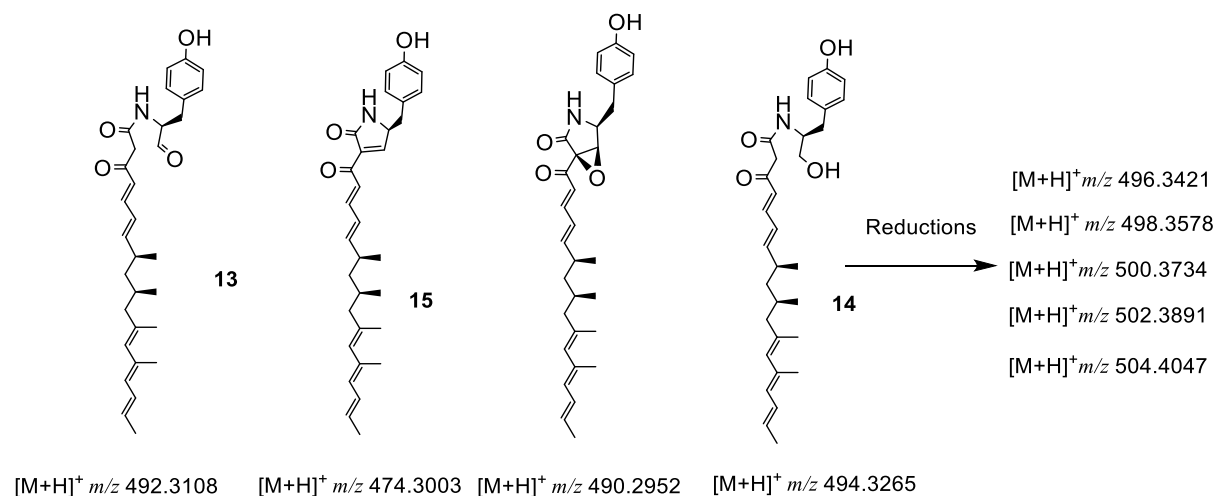


Fig. 3.1 Métabolites recherchés dans les cultures de la levure transformée

Aucuns de ces composés n'a été trouvé, peut-être en raison d'une mauvaise expression des protéines due à une efficacité limitée du promoteur.

### **3.4 Caractérisation de l'efficacité du promoteur pRPL18B**

Afin d'identifier le problème lié à l'absence de production métabolique dans la souche BJ5464-NpgA transformée nous avons vérifié la fonctionnalité du promoteur pRPL18B grâce à une construction de gène rapporteur. Celui-ci, construit à partir du kit Moclo de la levure nous a permis de confirmer que le promoteur était opérationnel.

### **3.5 Recherche de la PKS-NRPS et la ER chez la levure transformée**

Nous avons ensuite cherché à savoir si les protéines cibles PrcA et PrcC étaient produites lors de la fermentation. Pour cela les protéines ont été recherchées par Western blot, mais ces expériences sont restées sans succès.

# Chapitre IV

## Analyse des autres clusters de gènes

### 4.1 Introduction

Avec l'augmentation du nombre de génomes fongiques séquencés, des centaines de gènes de PKS-NRPS peuvent être identifiés mais sans être reliés au métabolite produit. Il apparaît donc que les BGC concernés sont des sources de nouveaux métabolites avec potentiellement des structures chimiques complexes et des activités biologiques intéressantes mais aussi de nouvelles fonctions enzymatiques.

Nous avons montré dans le chapitre 2 que des PKS-NRPS étaient localisés dans scaffolds 21 et 47. Nous avons tenté d'identifier les composés produits et analysé ces voies.

### 4.2 Analyses bioinformatiques du BGC impliquant *prbA* du scaffold 47

L'analyse du scaffold 47 a permis l'annotation de dix gènes dans le cluster *prb* et d'identifier des clusters de gènes similaires dans différents champignons (fig. 4.1).

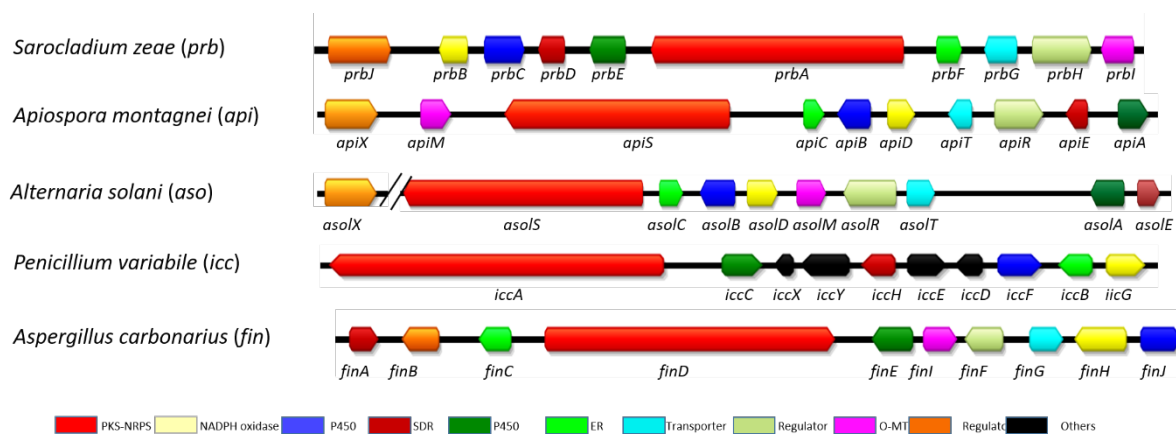


Fig 4.1. Le cluster de gènes *prb* et les homologues *api*, *aso*, *icc* et *fin* chez d'autres champignons.

Pour certains BGC, les métabolites sont connus, nous les avons cherchés sans succès chez *S. zeae*. De plus, la comparaison par LC/MS des extraits métaboliques de la culture de la souche sauvage de *S. zeae* avec celle du mutant  $\Delta prbA$  *S. zeae* n'a pas permis de mettre en évidence des différences résultant de la non production de métabolites chez le mutant.

Nous avons cherché à activer la voie de biosynthèse par surexpression des facteurs transcription et la substitution de promoteurs mais dans aucun cas nous n'avons obtenu de métabolite résultant de *prcb*.

### 4.3 Analyses du cluster de gènes *pra*

Quatorze gènes ont été identifiés dans le cluster *pra* localisé sur le scaffold 21.

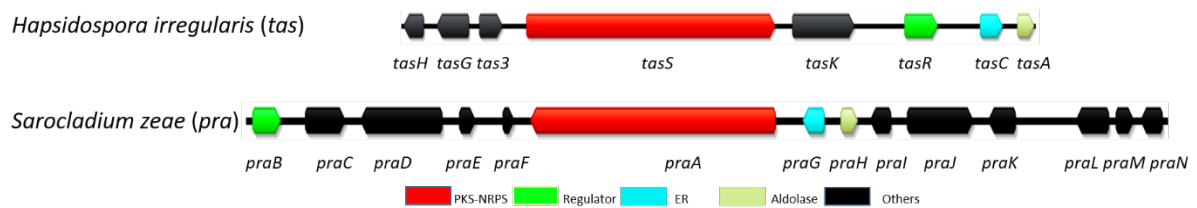


Fig. 4.2 Le cluster de gènes *pra* et un homologue *tas* chez *H. irregularis*

Très peu d'homologies avec d'autres clusters de gènes ont été mises en évidence par comparaison avec les bases de données. Une analyse comparative des profils métaboliques entre le mutant de délétion de la PKS-NRPS et la souche sauvage de *S. zeae* a été menée mais elle n'a pas permis d'identifier un métabolite associé au cluster. Celui-ci reste silencieux dans les conditions de cultures utilisées.



# **Foreword**



Microorganisms, prokaryotes (archaea, bacteria) and eukaryotes (fungi) can have different habitats and are present in all ecological niches. Some of them live in extreme conditions (temperature, pressure, high radiation) and others constitute the microbiota present in and on all multicellular organisms. In these different environments, communication and competition between microorganisms is constant and they have developed different strategies to defend themselves, among which there is antibiosis that corresponds to the production of secondary metabolites with antagonistic activities.

Endophytes are defined as bacteria or fungi that parasitic in plants tissue for the whole or part of the life cycle, causing no significant damage or infection to the host.<sup>1</sup> They co-evolve with the host and are mutually beneficial, meanwhile they may become pathogen when plants senesce or under environmental stress, and they have a dynamic equilibrium antagonistic relationship with plants. Endophytes adapts to the host environment and obtain nutrients by the production of extracellular hydrolases.<sup>2</sup> They are distributed in the roots, stems, leaves and seeds of plants and are spread by air, seeds, soil. So, they are generally obtained by screening on agar plates after light disinfection of plant tissue.

Endophytes play essential roles in nature: 1) by promoting host growth. It is well known that the rhizobia of legumes can play the role of nitrogen fixation to promote plant growth. Endophytic fungi produces gibberellins and indole acetic acid to promote host-plant growth.<sup>3</sup> 2) by protecting plants from insects, herbivores and pathogens. Peramine was characterized from the endophytic fungus *Acremonium lolii* in ryegrass. It is an antifeedant that protects ryegrass from weevils.<sup>4</sup> Several plant endophytic fungi produce ergot alkaloids as a repellent for plants.<sup>5</sup> 41 % of the isolated endophytic fungi from rice showed antagonism to rice pathogens.<sup>6</sup>

Plant pests and pathogens reduce global crop yields by 30-50%, resulting in poverty and hunger. Chemically synthesized pesticides are widely used to protect the normal growth of crops but in the meantime cause severe environmental pollution. The biocontrol agents are a feasible approach to replace pesticides but requires significant effort to develop.

Endophyte fungi also play an essential role in the exploitation of natural medicines. World-renowned drugs paclitaxel and lovastatin were characterized in a series of secondary metabolites of endophytic fungi.<sup>7,8</sup> Therefore, the exploration of endophytic fungi as the chemical and biological level is fundamental.<sup>5</sup>

In this context, the CPNFB team studies the chemical communication between endophytic microorganisms, the host plant and plant pathogens. The diversity of metabolites produced by the endophyte and their role in the defense of the host plant is explored. This also includes the study of their biosynthesis pathways and associated mechanisms of regulation. Our thesis work concerns the tripartite model maize - endophytic fungus *Sarocladium zae* — phytopathogen *Fusarium verticillioides* (Fig. 1.1).

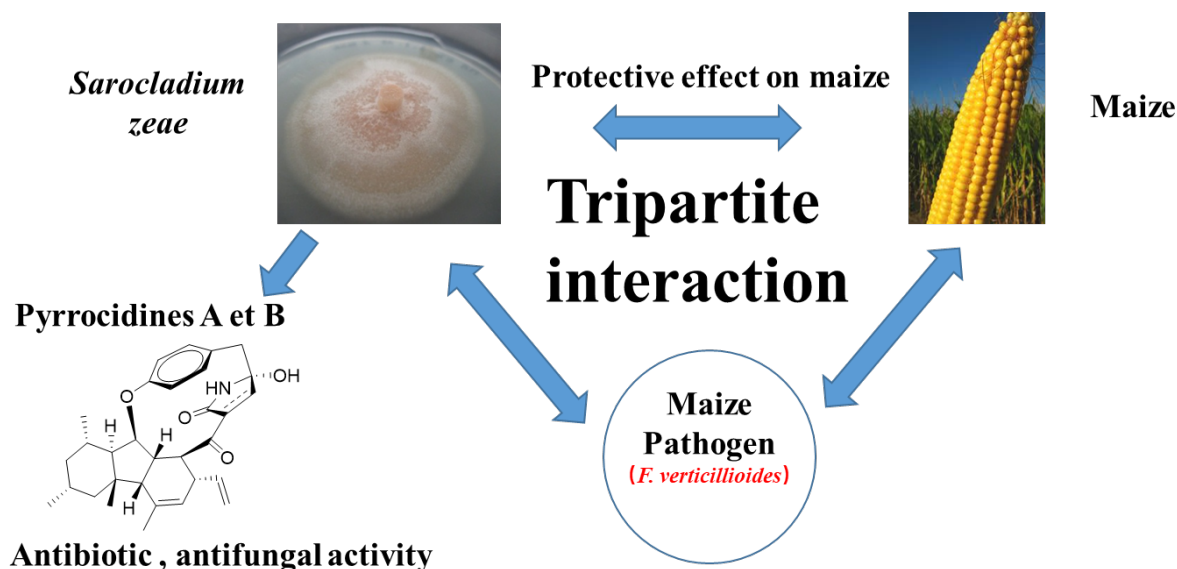


Fig. 1.1 The tripartite model of the chemical communication mediated by pyrrocidines.

*Sarocladium zae* has an antagonistic effect on maize pathogens (*Aspergillus flavus* and *Fusarium verticillioides*).<sup>9</sup> *F. verticillioides* has been defined as a notorious mycotoxigenic pathogen and posed a severe threat to food security since it survives in the pedicel and abscission layer of maize seed. However, *S. zae* survives in the embryo and endosperm of maize.<sup>10,11</sup> Furthermore, no report showed *S. zae* producing harmful compounds to maize.<sup>12</sup> Consequently, *S. zae* has been described as a protective endophyte of maize.<sup>9</sup>

Chemical studies showed that *S. zae* produces pyrrocidines A and B with antifungal activity against maize pathogens (*F. verticillioides*).<sup>13</sup> These molecules are representative of a growing family of fungal natural products sharing a decahydrofluorene core connected to a highly strained paracyclophane including a pyrrolidone moiety (Fig. 1.2). The mechanism of formation of these complex polycyclic structures is still an enigma and our thesis work consisted in identifying the reactions involved and the associated enzymes.

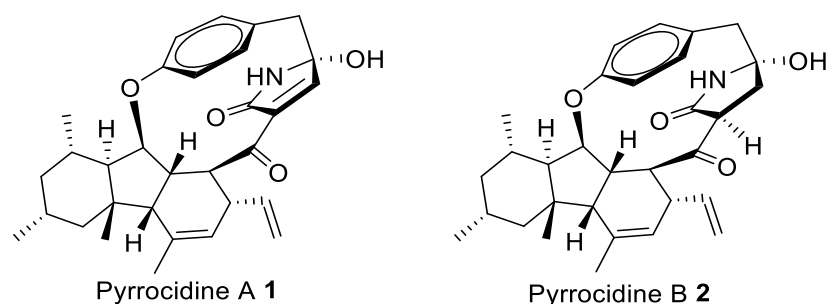


Fig. 1.2 The unique structures of pyrrocidine A and B.

The first part of Chapter I is devoted to secondary metabolites with the aim of showing their diversity and discussing the reactions involved in biosynthesis pathways. In the second part are exposed the approaches described to discover new natural products. Finally, the third part reports the knowledge on pyrrocides and analogues.

Chapter 2 describes the results that have made it possible to propose a pyrrocidine biosynthetic pathway. The study of the genome of *S. zeae* shows the presence of three biosynthetic gene clusters involving a PKS-NRPS and potentially at the origin of pyrrocidines. After identifying the BGC involved, we were able to generate inactivation mutants for these different enzymes. The identification of the new metabolites produced allowed us to determine the role of each of the enzymes.

Chapter 3 is devoted to the work carried out to reconstruct a pyrrocidine biosynthesis pathway in yeast *Saccharomyces cerevisiae*.

Chapter 4 presents the preliminary work towards the identification of metabolites produced by the two other BGCs involving a PKS-NRPS gene in *S. zeae*.

Chapter 5 is the experiment part.

# **Chapter I: General Introduction**

## 1.1 Diversity of natural products and their biosynthetic pathways in fungi.

Natural products have been utilized as medicines to prevent and treat diseases for thousands of years.<sup>14</sup> From 1981 to 2019, it has been reported that around 1,881 small-molecule drugs have been approved by the FDA. About 50% of the drugs on the market were originated from natural products and their derivatives.<sup>15</sup> Natural products mainly refer to molecules derived from animals, plants and microorganisms. Since animals and plants are slow-growing and extraction steps of compounds are either troublesome or unfriendly to the environment, microorganisms, with their fast proliferation and wide distribution, have become essential resources for developing clinical antibacterial, antitumor, immunosuppressive, and other drugs.<sup>16</sup> Natural compounds are also commonly utilized in agriculture, food and cosmetic industries. Currently, fungi produce about half of the bioactive microbial molecules.<sup>17</sup>

Natural products of pharmacological interest originate from the secondary metabolism which encompasses all the processes and molecules not required for the growth of the microorganisms. This is in contrast to the primary metabolism which produces molecules necessary for the development of the cell, like lipids, carbohydrates, nucleic acids, vitamins or hormones. Secondary metabolites can be categorized according to their structure into the five following classes: 1) terpenoids, 2) alkaloids, 3) peptides, 4) polyketides, and 5) hybrid metabolites. In most of the biosynthetic pathways producing these metabolites, first the backbone of the molecule is enzymatically assembled by synthases or synthetases and then chemical functionalizations are installed by various auxiliary (or tailoring) enzymes. This mode of programming allows reaching a vast diversity of metabolites according to the nature and the specificity of the auxiliary enzymes.

In a biosynthetic pathway, these enzymes are generally encoded in fungal genomes at one locus as a biosynthetic gene cluster (BGC). Sometimes more tricky configurations occur where the BGC displays a bipartite genomic structure like in the case of botrydial produced by *Botrytis cinerea*.<sup>18</sup> Hybrid metabolite biosynthesis can also require enzymes encoded on two different chromosomes as in the biosynthesis of nidulanin A produced by *Aspergillus nidulans*.<sup>19</sup> In even more complex cases, two BGCs can be intertwined as for fumagillin and pseurotin in *Aspergillus fumigatus*.<sup>20</sup>

Fungi usually encode more than 60 natural products, and more than 90% of gene clusters in fungi are silent in laboratory cultures.<sup>21</sup> So fungi are a potentially tremendous reservoir for

natural product discovery. In the last decade, technological advances in several fields were determinant steps for the studies of fungal biosynthetic pathways. Firstly, sequencing facilities based on different technologies, like Illumina and Nanopore, allowed the access to 20-100 Mb genomes corresponding to fungi with cheap prices. Thus, more than 1000 representative fungal genomes have been released in public databases.<sup>22</sup> Secondly, analytical tools developed in the fields of genome data mining (AntiSMASH...) and metabolomics (XCMS, GNPS...) as well as heterologous expression host system (in yeast, *E. coli* or *Aspergillus sp.*) and *de novo* gene synthesis at low-cost facilitated the study of biosynthetic pathways from the gene to the molecule. Moreover, the emergence of synthetic biology provides new possibilities to recombine genes from various pathways which is a promising field to generate synthetic metabolites with potential new bioactivities.

Fungi plays an essential role in the discovery of natural products. Here, we shed light on the variety of secondary metabolites of fungi.

### 1.1.1 Terpenoids

Terpenoids are defined as terpene or its derivatives that are formed by two or more isoprene units. The general formula of terpenes is  $(C_5H_8)_n$  ( $n$  is the number of isoprene units). According to the number of  $n$ , terpenoids can be divided into hemiterpenes (C5), monoterpenes (C10), sesquiterpenes (C15), diterpenes (C20), disesquiterpenes (C25), triterpenes (C30), tetraterpenes (C40), and polyterpenes ( $n > 8$ ).<sup>23</sup> The skeleton of terpenoids originated from two precursors namely isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) by undergoing a different number of repeats, rearrangement, and cyclization (Fig.1.3).<sup>24</sup> The DMAPP and IPP were accumulated using acetyl-CoA as the substrate through the mevalonate (MVA) pathway in fungi, which is different from bacteria and plants (mevalonate pathway or methylerythritol phosphate (MEP) pathway). Terpene synthases are mainly divided into two types: one is ionization-dependent terpene synthases that cause the departure of pyrophosphate groups catalyzed by metal ions such as  $Mg^{2+}$ ,  $Fe^{2+}$ , etc, including monoterpene, sesquiterpene, and diterpene synthases. Another is protonation-dependent terpene synthases that cause the removal of the pyrophosphate group by epoxidation or protonation of the carbon-carbon double bond, including partial diterpene synthases and triterpenoid synthases.<sup>25</sup> Both types build the carbon backbone of the molecule and then decorating enzymes introduce various modifications such as cyclization, oxidation, methoxylation, or glucosylation yielding the structure diversity of terpenes.

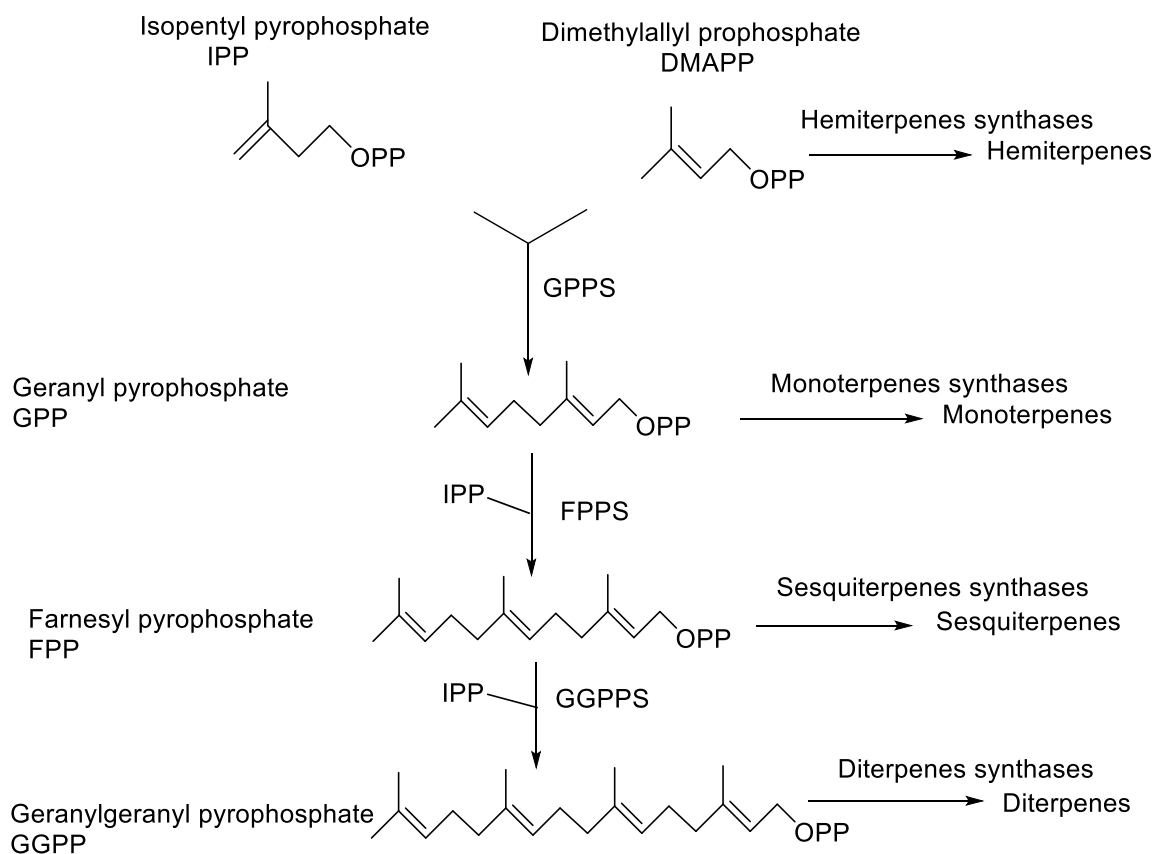


Fig 1.3 Biosynthetic pathway of terpenoids.<sup>24</sup>

As an example, geranyl-PP gives the  $\alpha$ -terpinyl cation, which constitutes the skeleton of monoterpenes. The carbocation then undergoes a series of rearrangement or quenching to form a diversity of structures (Fig. 1.4). The ten-carbon skeleton of monoterpenes can evolve to so many different structures that it gives an idea of how wide and complex could be the variety of structures produced by carbocation chemistry from larger precursors like sesquiterpenes, diterpenes, triterpenes.

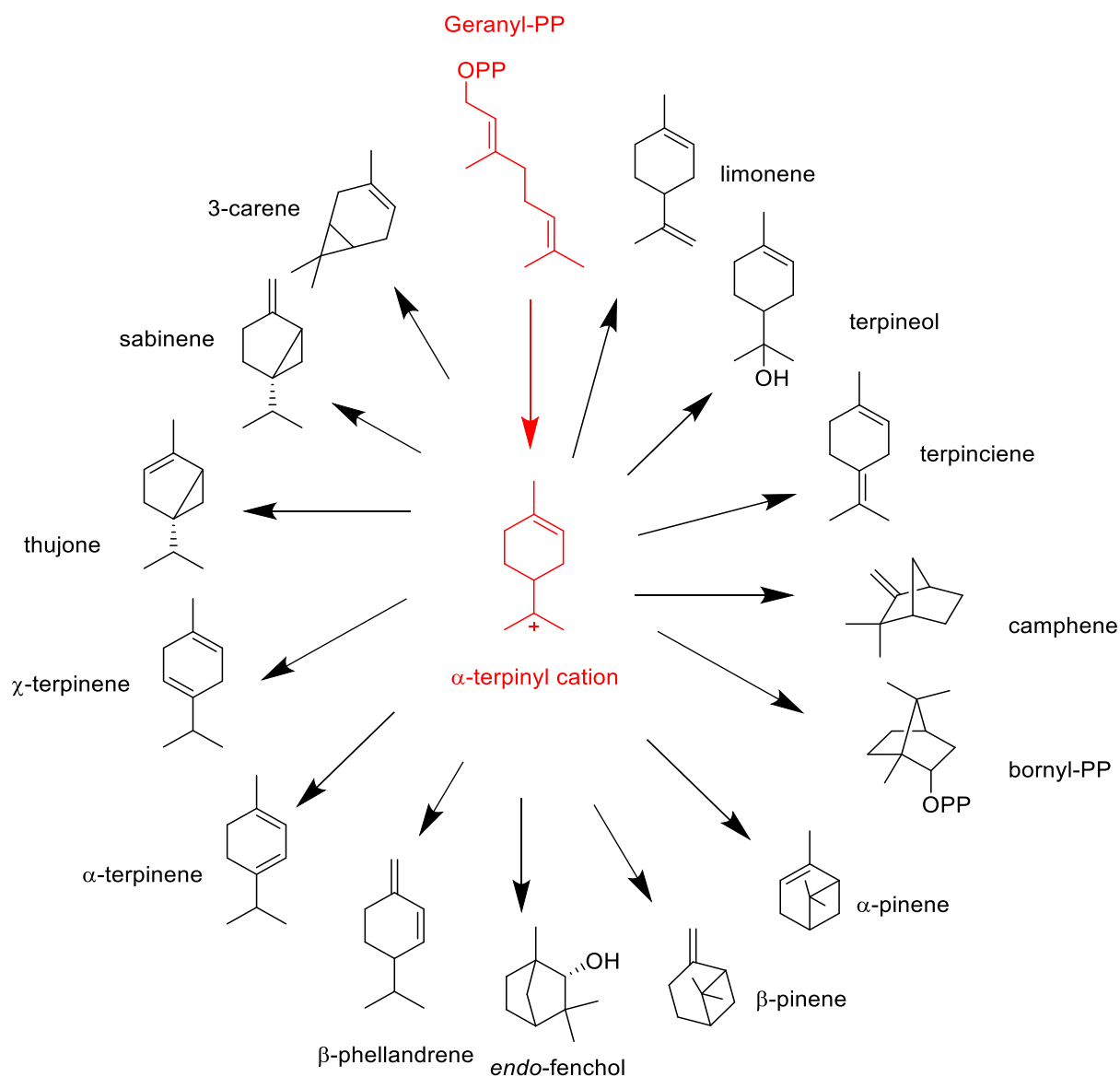


Fig. 1.4 Carbocation chemistry of monoterpenes. The  $\alpha$ -terpinyl cation, initially formed from geranyl-PP, can undergo different quenching reactions like water addition, proton removal and carbon bond migration. It can also undergo intramolecular carbon bond migration or carbon-carbon bond formation followed by quenching.

Artemisinin and paclitaxel are the most renowned terpenoids. Artemisinin has a tremendous anti-malarial impact, while paclitaxel has an excellent anti-cancer (breast, ovarian and lung cancer) effects. They were first discovered from plant extracts, and paclitaxel was also found in a series of secondary metabolites of endophytic fungi.<sup>7</sup> Some terpenes of fungal origin have diverse biological activities (Fig. 1.5), such as: ganoderic acid DM,<sup>26</sup> erinacine A,<sup>27</sup> eburicoic acid,<sup>28</sup> brassicene,<sup>29</sup> as well as sterenin E.<sup>30</sup>



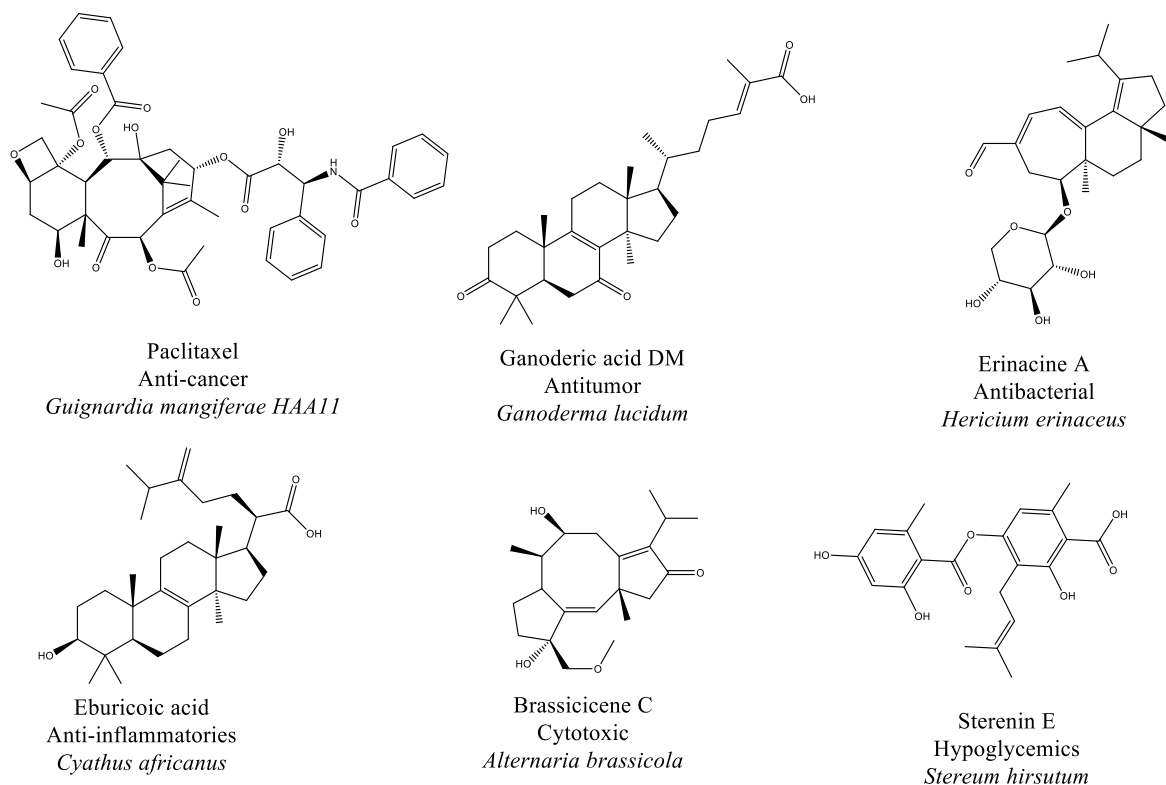


Fig.1.5 Structures of biologically active terpenes derived from fungi.

### 1.1.2 Alkaloids

Alkaloids are defined as organic compounds containing nitrogen, leading to a diversity of structure and a multiplicity of activities.<sup>31</sup> The biosynthesis of alkaloids typically involves the ligation of a nitrogen precursor and most of them contains at least one N-containing ring. For instance, an amino acid can react with other primary metabolism building blocks and further be transformed into other products by oxidation and cyclization. Alkaloids can be classified as 1) quinoline and isoquinoline as penicinetam,<sup>32</sup> 2) amines and amides as 7-amino-4-methylcoumarin,<sup>33</sup> 3) indole derivatives as 2-(3,3-dimethylprop-1-ene)-epicostaclavine,<sup>34</sup> 4) pyridines as 7,8-dimethyl-isoalloxazine,<sup>35</sup> 5) quinazolines as chaetominine,<sup>36</sup> and 6) others as trichoderme C<sup>37</sup> (Fig 1.6).

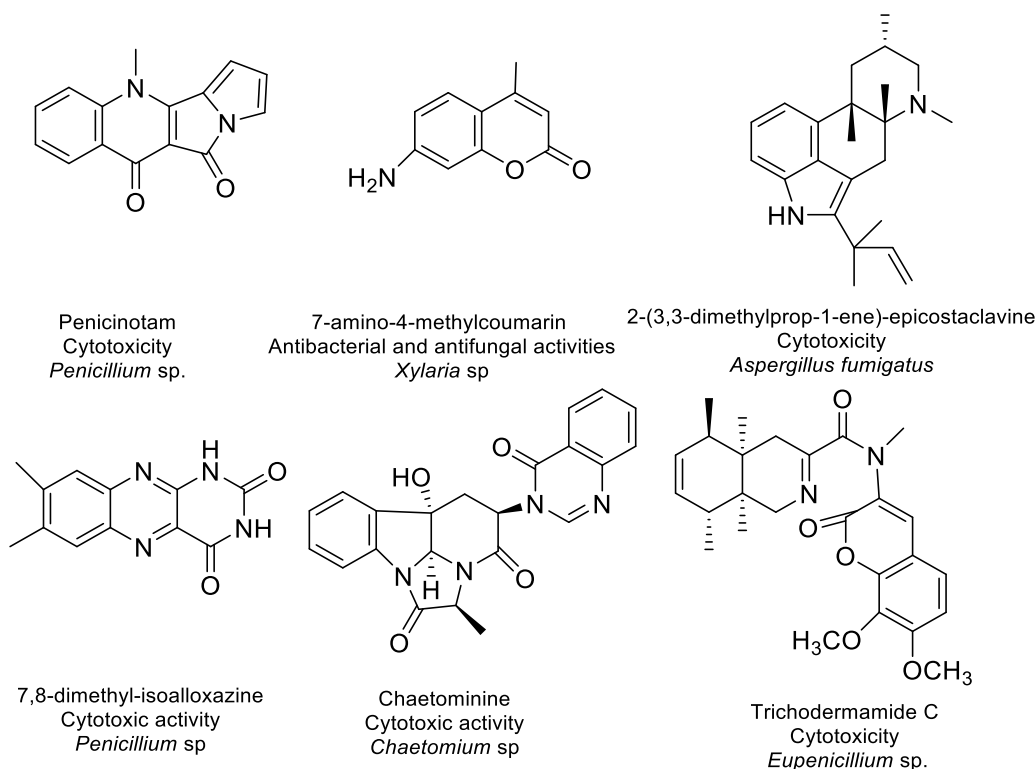


Fig. 1.6 Structures diversity of fungal alkaloids.

### 1.1.3 Peptides and peptide-derived natural products

Peptides and peptide derivatives originate from amino acids, which are assembled during the polymer biosynthesis through amide bond formation. In fungi, three different enzymatic systems have been reported to achieve peptide synthesis: 1) the mega-enzymes Non-Ribosomal Peptide Synthetases (NRPSs), 2) Aminoacyl (AA)-tRNA-Dependent Cyclo-Di-Peptide Synthases (CDPSs)<sup>38</sup> producing diketopiperazines (DKPs) and 3) Ribosomally synthesized and post-translationally modified peptides (RiPPs).

#### 1.1.3.1 Non-ribosomal peptides (NRPs)

NRPS catalyzes the condensation of different amino acids to produce peptides. The monomer units used by NRPS are composed of proteinogenic and non-proteinogenic amino acids. Over 300 dissimilar building blocks have been identified as substrate for NRPS, which is one of the foundations of their structural diversity.<sup>39</sup> NRPS are constituted of three essential domains: condensation (C), adenylation (A) and thiolation (T) domains (Fig. 1.7).<sup>40</sup> The adenylation domain (A) is in charge of selecting an amino-acid and activating it through an amino-acyl-AMP. This intermediate reacts with the thiol group of T domain which was preparatorily activated to its holo form by adding a phosphopantetheinyl arm. This results in

the loading of the amino-acid on the T domain. C domain controls the upstream peptide (donor) stereochemistry and the selectivity of downstream (receptor) amino acids with thioesters concomitantly formed under the action of thiolation domain (T) (Fig. 1.7) and catalyzes the peptide bond formation.<sup>41</sup> When A domain is not rigorous enough to select the substrate, C domain can play a gatekeeping role.<sup>42</sup> In bacteria, TE domain uses hydrolysis, ammonolysis, and intermolecular cyclization to release the peptide chain from the assembly line.<sup>42</sup> Fungi use this kind of domain, but another approach is also encountered to release the peptide from the assembly line: C<sub>T</sub> is the subunit of C domain that is recruited to replace the function of TE. All cyclically released peptide chains in fungi are catalyzed by the C<sub>T</sub> domain.<sup>43</sup> Some iterative NRPS have also been characterized.<sup>44</sup>

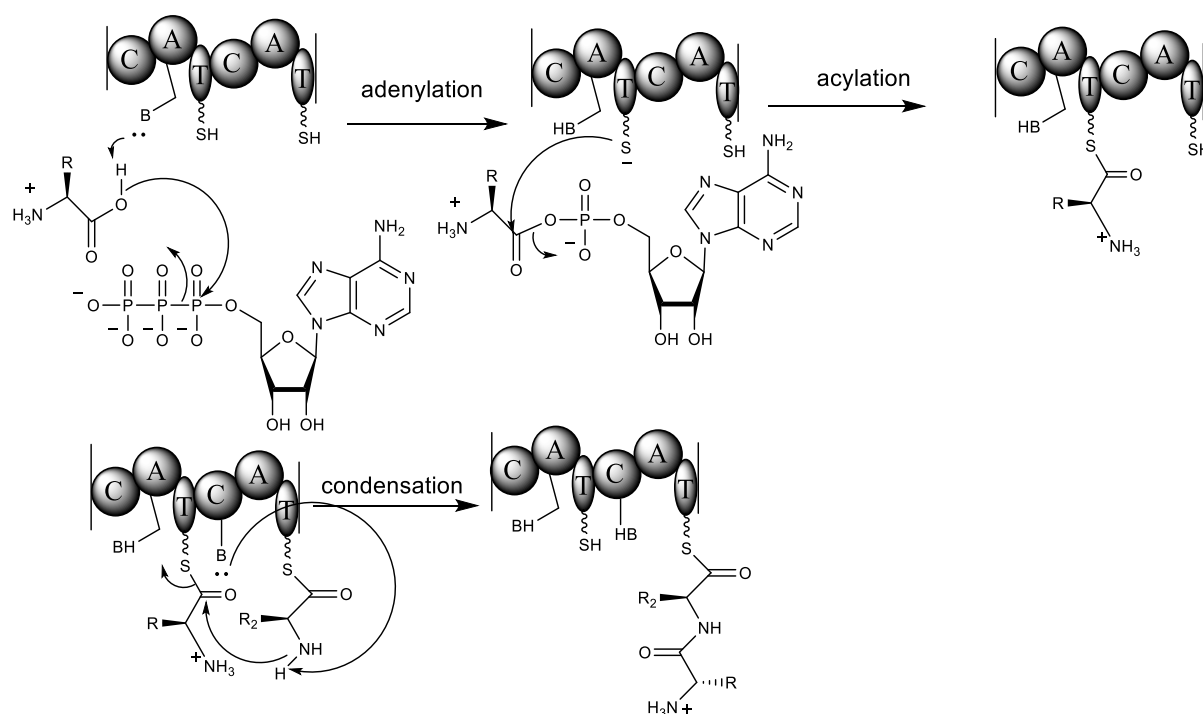



Fig. 1.7 The mechanism of non-ribosomal peptide synthetase (NRPS).

Double function C domain have also been found, like C/E or C<sub>y</sub> performing epimerization or heterocyclization, respectively. Besides, some NRPS have domains like: epimerase (E), oxidase (Ox), cyclase (CYC), methyltransferase (MT), reductase (Re), acyl-CoA ligase (AL).<sup>45</sup> For instance, in the biosynthesis of the fungal conidiophore pigment, the E domain of LvoA (NRPS) converted the natural L-tryptophan into D-tryptophan.<sup>46</sup> Through the isomerization, methylation, acylation and other functions, the domains mentioned above further modify the peptide chain, thereby deriving various structural products.

Ten critical amino acid residues were supposed to control the specific selection of the A domain, which can be apply to predict the bound substrate (Fig. 1.8).<sup>47</sup>



The diagram shows a purple rectangular box labeled "A domain" with dashed lines extending from its top and bottom edges, indicating its position within a larger protein structure.

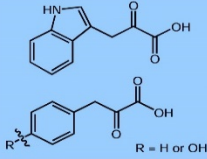
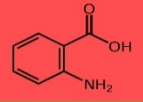
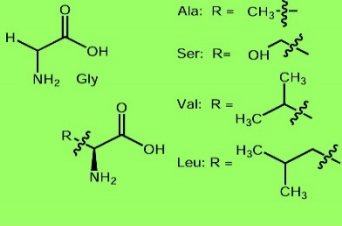
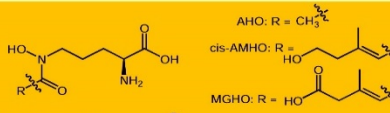
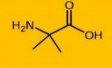
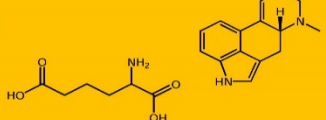
Enzyme	Nonribosomal code										Substrate
Position	1	2	3	4	5	6	7	8	9	10	
Residue	235	236	239	278	299	301	322	330	331	517	
TdiA	V	A	H	F	T	G	A	A	C	K	indolyl-3-pyruvic acid 4-hydroxyphenylpyruvic acid 4-hydroxyphenylpyruvic acid phenylpyruvic acid 
GreA	V	A	E	F	S	G	G	A	C	K	
AtrA	V	A	E	F	S	G	G	A	C	K	
MicA	V	A	T	F	I	G	G	A	G	K	
AnaPS (A1)	G	A	L	F	F	A	A	G	V	K	anthranilic acid 
ArdA (A1)	G	M	I	L	L	A	A	G	I	K	
TqaA (A1)	G	V	I	F	M	A	A	G	V	K	
GlIP (A2)	D	Y	N	T	Y	T	A	I	C	K	L-serine glycine L-valine L-valine L-leucine L-alanine L-alanine L-alanine 
EasA (A1)	D	I	Q	G	V	L	A	M	Q	K	
EasA (A2)	D	A	S	Q	I	G	G	I	Y	K	
LPS4 (A1)	D	A	I	F	C	G	G	P	L	K	
EasA (A3)	D	I	H	F	V	G	A	I	A	K	
EasA (A4)	D	L	L	V	V	A	G	I	L	K	
EasA (A5)	D	I	A	I	L	V	A	I	L	K	
LPS1 (A1)	D	L	F	F	C	G	G	P	L	K	
LPS3	D	I	F	L	A	G	I	I	G	K	
CpaA	D	M	A	L	C	G	S	A	C	K	
CheA	D	M	I	I	C	G	C	A	A	K	
ApdA	D	M	V	I	C	G	C	A	A	K	
TenS	D	M	V	I	C	G	C	A	A	K	
LnaA/LnbA	D	V	F	A	F	G	A	I	F	K	
PsoA	D	A	Y	T	M	A	A	I	C	K	
GlIP (A1)	D	A	G	T	L	G	A	L	M	K	
NPS2 (A4)	D	V	L	D	I	G	G	I	G	K	
SidN (A3)	D	V	G	G	G	G	V	I	G	K	AHO cis-AMHO MGHO 
Fso1 (A3)	D	I	I	T	I	T	A	T	L	R	
TqaB	D	L	F	M	M	C	G	C	I	K	
Tex1	D	L	G	Y/F	L	A	G	V	F/C	K	α-aminoisobutyric acid α-aminoisobutyric acid 
LPS2	D	V	F	S	V	G	L	V	M	K	
Lys2	D	P	R	H	F	V	L	R	A	K	D-lysergic acid α-amino-adipic acid 

Fig. 1.8 Ten key positions that are relevant for substrate specificity in fungi.<sup>47</sup>

A crowd of NRPs was extracted from fungi (Fig. 1.9). For instance, penicillin from *Aspergillus fumigatus* is a world-renowned broad-spectrum antibiotic that saved countless lives in World War II.<sup>48</sup> Cephalosporin is another eminent broad-spectrum antibiotic originated from *Acremonium chrysogenum*.<sup>48</sup> Nanangelenin A has cytotoxicity obtained from *Aspergillus nanangensis*.<sup>49</sup> Asperphenamate has antitumor effect extracted from *Aspergillus flavipes*.<sup>48</sup>

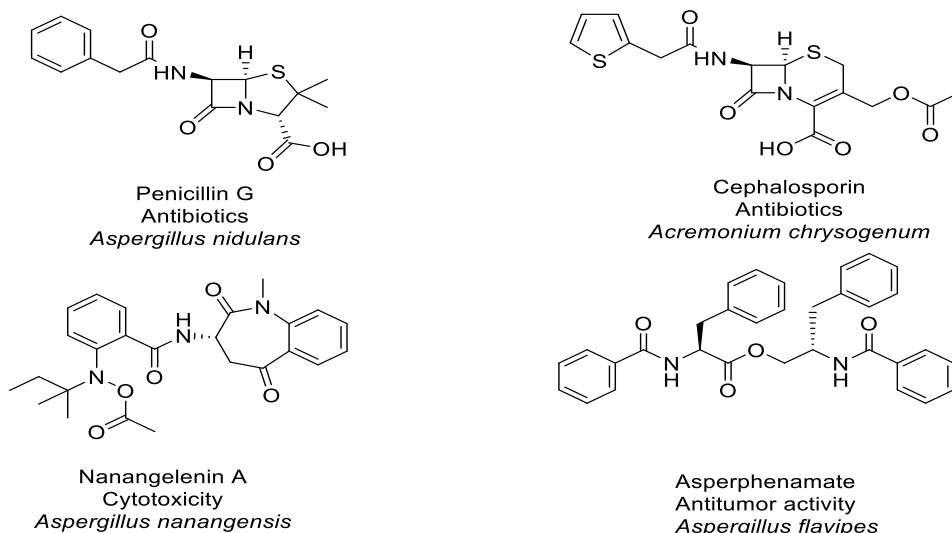


Fig. 1.9 The structure diversity of fungal non-ribosomal peptides.

### 1.1.3.2 Diketopiperazines (DKPs)

The basic structure of DPKs is a cyclic dipeptide formed by the condensation of two amino acids, and its backbone contains a stable six-membered ring. Many of them are produced by bi-modular NRPSs. For instance, the cytotoxic compound trichoderamide A isolated from *Trichoderma virens* fungus belongs to the diketopiperazine family. Its biosynthetic pathway has been determined by bioinformatics analysis and gene inactivation. Two phenylalanines are condensed in the presence of the NRPS TdaA to form cyclo-Phe-Phe. Then P450 cytochrome introduces the hydroxyl group, followed by substituting sulfhydryl moieties thanks to four enzymes (TdaL, TdaK, TdaJ, TdaF). Next, a FAD-dependent oxidase (TdaR) creates a disulfide bond. The biosynthesis of the trichoderamide A was finalized with the help of P450 and methyl-transferases (Fig. 1.10).<sup>50</sup>

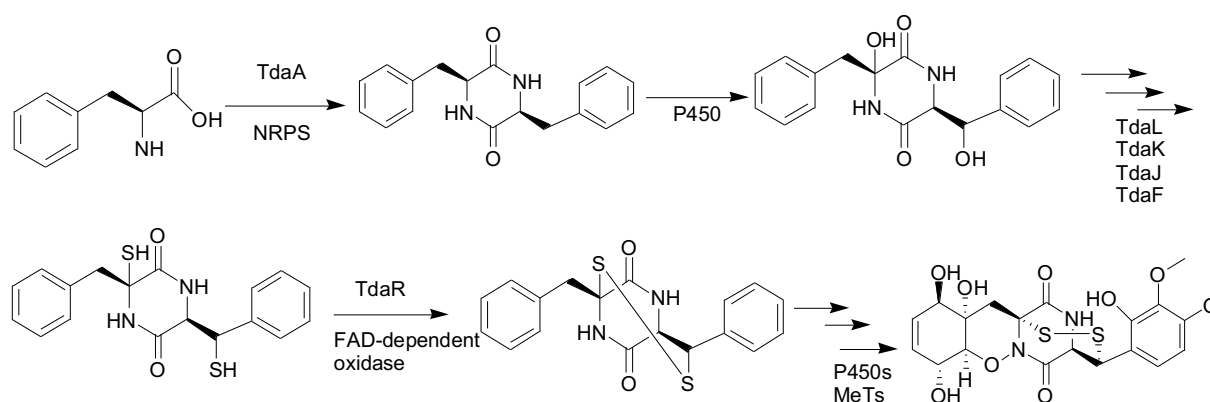


Fig.1.10 The biosynthetic pathway of trichoderamide A

Several other fungal DPKs originated from NRPS have been characterized, such as, roquefortine,<sup>51</sup> cristatumin A<sup>52</sup> and phomazine B<sup>53</sup> (Fig. 1.12).

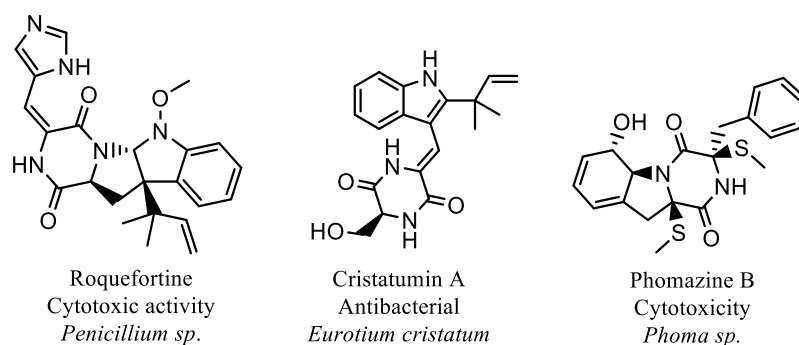


Fig. 1.12 The structure diversity of fungal DPKs originating from NRPS.

Another way to build the backbone involves a single aa-tRNA-dependent cyclodipeptide synthases (CDPS). Three steps were involved in synthesizing the backbone of DPKs. Firstly, the aminoacyl group of a first aa-tRNA was transferred to a conserved serine; secondly, aminoacyl serine intermediate combines with the second aa-tRNA by forming an amide bond; thirdly, the second amide bond was catalyzed by intramolecular cyclization, giving the formation of the piperazine-2,5-dione ring.<sup>54</sup> For instance, the CDPS pathway of cyclo-Phe-Phe was shown in Fig 1.11.<sup>54</sup>

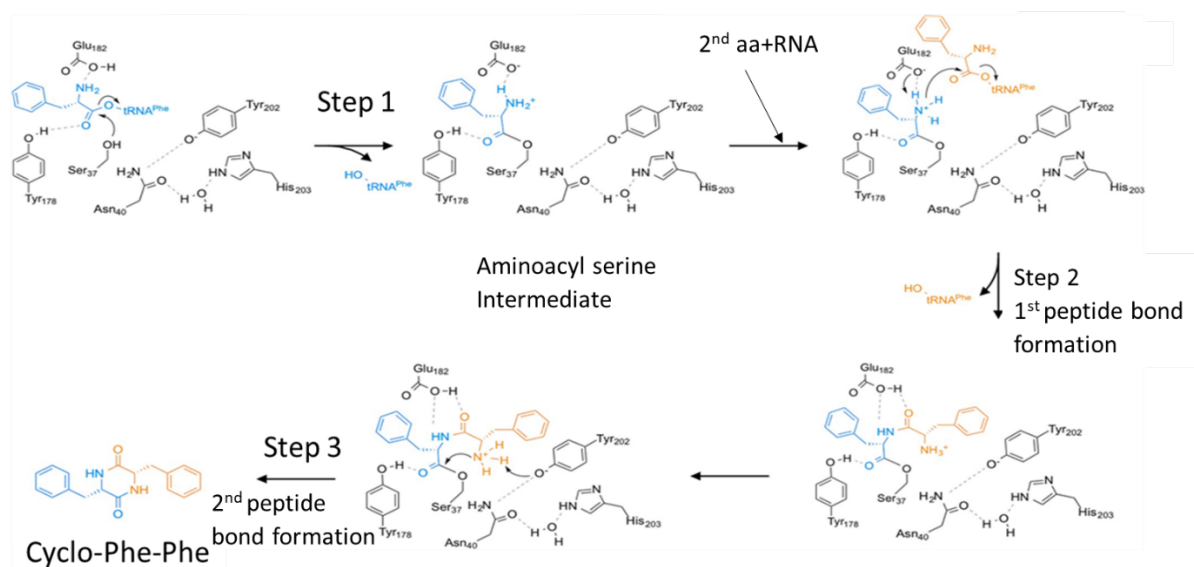


Fig. 1.11 The cyclodipeptide synthases pathway of cyclo-Phe-Phe.<sup>54,55</sup>

Some reviews reported that CDPS-dependent pathways have been identified in fungi.<sup>55</sup> However, after verification, we did not identify any references supporting these allegations. To the best of our knowledge, only one review proposed that the presence of protein

PH1(XP\_384791.1) from *Fusarium graminearum* is related to bacterial-derived CDPS (13% sequence identity),<sup>55</sup> but, this fungal sequence is now removed from the database. It seems that no CDPS pathway has been identified in fungal so far.

### 1.1.3.3 Ribosomally synthesized and post-translationally modified peptides (RiPPs)

In the biosynthesis of RiPPs, genomic DNA is transcribed and translated to form a precursor peptide composed of different parts with specific functions: the core peptide whose sequence will constitute the backbone of the final peptide is completed by a signal sequence and a leader peptide at its N-terminus and a recognition sequence at the C-terminus. The core peptide first undergoes post-translational modifications which are leader peptide-dependent or not.<sup>56</sup> The resulting modified precursor peptide is finally cleaved to generate a mature peptide (Fig. 1.13).

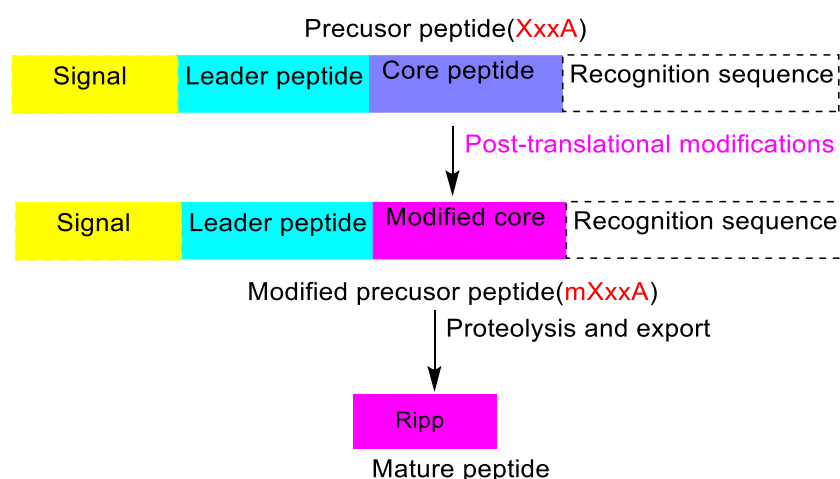


Fig 1.13 Biosynthesis of RiPPs.<sup>56</sup>

The overwhelming majority of RiPPs found in fungi belongs to the amatoxins/phallotoxins family. The remaining handful of fungi source RiPPS belongs to Borosins (containing heterocycles and methylation) and Dikaritins.<sup>57</sup> For example, Amanullic acid,<sup>58</sup> Phallacidin,<sup>59</sup> Antamanide,<sup>60</sup> belongs to Amatoxins/Phallotoxin family. Omphalotin A-C<sup>61</sup> are classified into Borosins family. Ustiloxin A,<sup>62</sup> Ustiloxin F,<sup>63</sup> Asperipin-2a<sup>64</sup> are part of Dikaritins (Fig. 1.14). Because gene manipulations in fungi are difficult to perform and the structure of the RiPPS compounds are complex, most fungi-source RiPPS synthesis routes are not completely analyzed. For instance, only two critical enzymes have been identified in the biosynthesis of omphalotin A: methyltransferase OphMA catalyzes the methylation of its C-terminus, and the protease OphP is responsible for release and cyclization.<sup>65</sup>

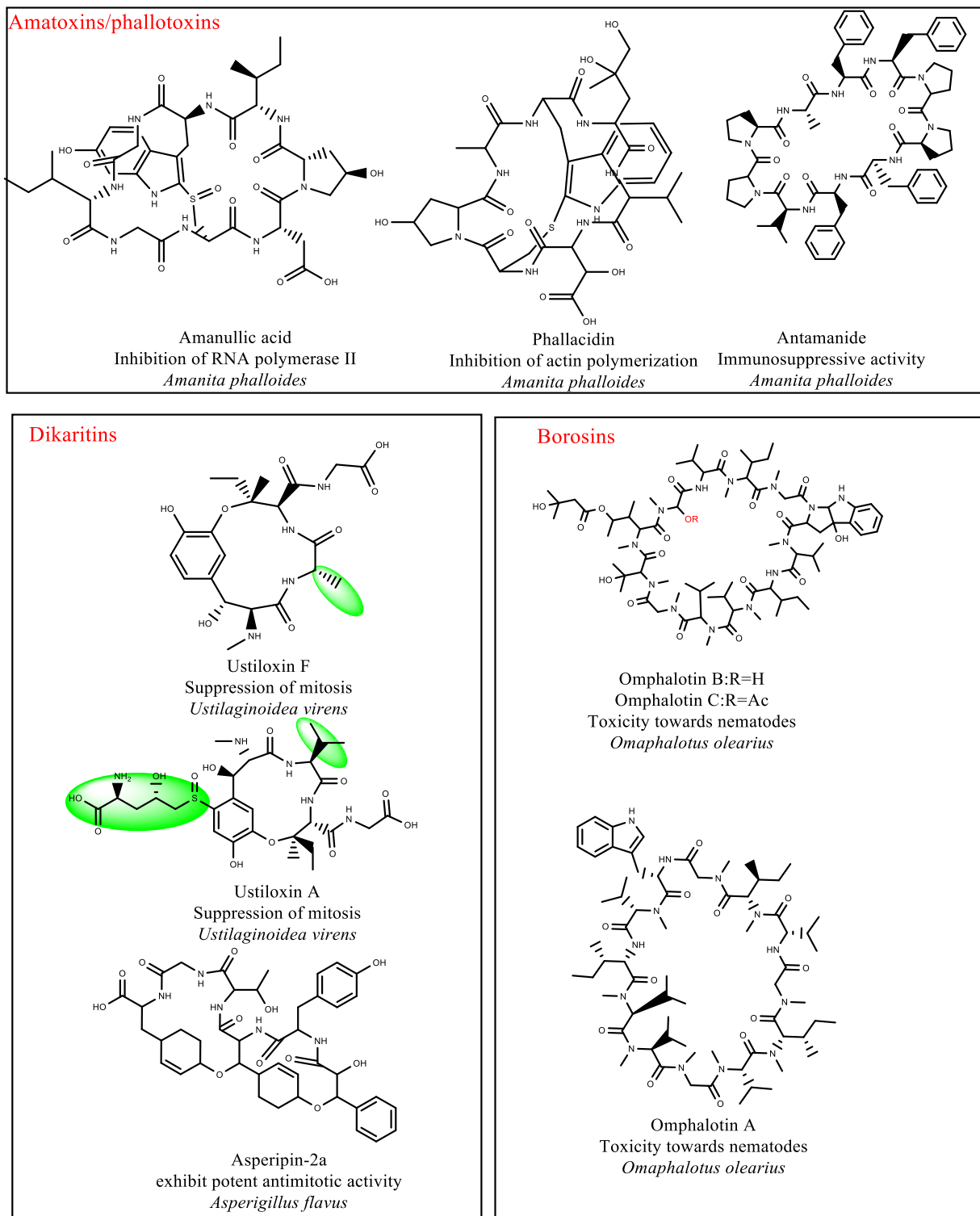


Fig. 1.14. The structure diversity of fungal RiPPs

### 1.1.4 Polyketides (PKs)

Polyketides, the major secondary metabolites of fungi, are biosynthesized by polyketide synthases (PKS). These enzymes can be categorized into three types according to their catalytic mechanism, named type-I, II and III.<sup>66</sup>



Type I PKS can be either multi-modular or iterative, with most of modular type I PKS being found in bacteria and iterative type I PKS mainly described in fungi. These proteins are composed of domains with diverse catalytic functions. Acyltransferase (AT), ketosynthase (KS) and acyl carrier protein (ACP) are the essential domains required for the polyketide (PK) biosynthesis. AT is in charge of recruiting and transferring extension units from malonyl-CoA to ACP.<sup>67</sup> KS catalyzes a Claisen condensation between the malonyl group carried by ACP and an acyl chain covalently bound to KS itself.

In addition, type I PKS has other domains that can modify the growing chain of the polyketide. These optional domains are ketoreductase (KR), in charge of reducing the  $\beta$ -keto group to  $\beta$ -hydroxy group; dehydratase (DH), catalyzing hydroxyl groups dehydration to form unsaturated olefins; methyltransferase (MT), methylating the carbon chain; enoyl reductase (ER), reducing the  $\alpha$ - $\beta$  double bond to a single bond. This domain can be found in the PKS sequence or recruited as a discrete protein encoded by a single gene (*in trans*) located in the BGC. Thioesterase (TE) and cyclase (CYC) catalyze the hydrolysis or cyclization reaction to release mature polyketide chains.<sup>68</sup>

Fungal type I iterative polyketide synthase is divided into non-reducing PKSs (NR-PKS), partially reducing PKSs (PR-PKS), and highly-reducing PKSs (HR-PKS), according to the degree of reduction of polyketide intermediates (Fig. 1.15). An NRPS-liked C domain was frequently found at the C-terminal of HR-PKS. It was proposed to be involved in the release of the polyketide chain.<sup>69</sup> Compared with HR-PKS, PR-PKS does not contain the MT and ER domains.<sup>70</sup> In NR-PKS, a starter unit-ACP transacylase (SAT domain) is located at the N-terminal and is responsible for selecting the start unit. None of the DH, MT or ER domain was found in NR-PKS. A product template domain (PT) was speculated to control the length, folding and stabilization of the polyketide chain. An optional TE domain is in charge of the Claisen condensation to release and cyclase the linear polyketide chain. Without TE domain, the polyketide chain automatically forms C-O cyclization to release the intermediate.

## HR-PKS



## PR-PKS



## NR-PKS



Fig. 1.15 The general domain constitution of the three types of fungal iterative PKS.

The mechanism of the fungal type I iterative HR-PKS biosynthesis pathway is showed in Fig. 1.16. Hydrolysis of acetyl-CoA thioester bond has a  $\Delta G$  around  $-40$  kJ/mol. For this reason, this thioester bond is called a “high energy bond”, making the S-CoA as an easy leaving group. The  $-OH$  of Ser residue in AT domain’s active site specifically select acetyl-CoA as start unit. The  $-SH$  group of a Cys residue in KS domain’s active site act as a point of attachment of the start unit. A functional ACP must be covalently linked with a phosphopantetheine arm (PPant arm).<sup>71</sup> The PPant arm is originated from a dispersed PPTase, which produces sulfhydryl functional groups critical for subsequent reactions. ‘Robot arm’ like conformational flexibility enables ACP domain to communicate with a variety of cis- and trans-acting components.<sup>42</sup> Malonyl-CoA was originated from acetyl-CoA catalyzed by acetyl-CoA carboxylase. AT domain selects malonyl-CoA as the extend unit, and attached to ACP domain. The KS domain performs Claisen condensation between the acetyl-CoA and malonyl-ACP to give  $\beta$ -ketoacyl-ACP, of which reaction is driven by the emitted energy of decarboxylation. MT adds one methyl to the polyketide chain originated from S-adenosyl-methionine. MT catalyzed methylation must precede  $\beta$ -carbonyl reduction by the KR. Intermediates that require methylation prefer to bind MT domain rather than KR domain and can only attach to KR when methylation is completed.<sup>70</sup> KR reduces the  $\beta$ -keto moiety to give  $\beta$ -hydroxybutyryl-ACP in the presence of NADPH. Work of Cox *et al.* showed by domain swapping that KR control the length of the PK chain.<sup>72</sup> DH catalyzed dehydration to form *trans*- $\Delta^2$ -butenoyl-ACP. ER reduced the double bond to provide the butyryl-ACP with the help of NADPH. These four steps are optional in each cycle. Then, the polyketide chain is translocated from ACP to KS. Next, ACP recharged a new malonyl-CoA for another cycle. After several cycles, the thioesterase (TE) domain releases the polyketide chain from ACP.

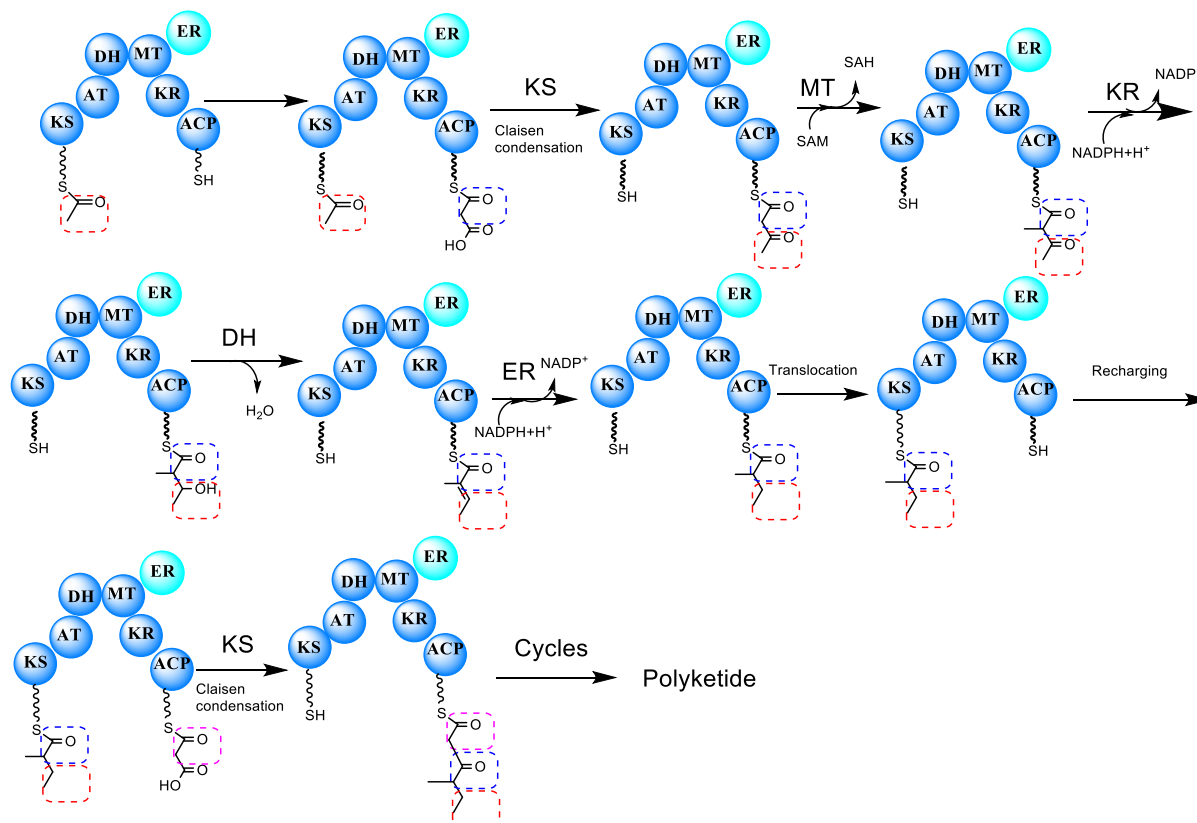


Fig.1.16 The mechanism of fungal type I iterative HR-PKS.

The biosynthetic pathway of type I multi-modular PKS was predominantly found in bacteria. Each module contains a different composition of domains and is responsible for a one-time extension of the carbon chain. These lead to bacterial PKs requiring multiple modules and even several PKS participation. Finally, the PKS of bacteria often has a more gigantic gene size than fungi.

The sophisticated iterative PKS gives rise to the high diversity of fungal polyketides. Some example of the diversity is represented in Fig. 1.17: fumonisins B<sub>1</sub><sup>73</sup> and lovastatin<sup>74</sup> (HR-PKS); 6-methylsalicylic acid<sup>75</sup> (PR-PKS) and aflatoxin B<sup>76</sup>, griseofulvin<sup>77</sup> and DHN<sup>78</sup> (NR-PKS).

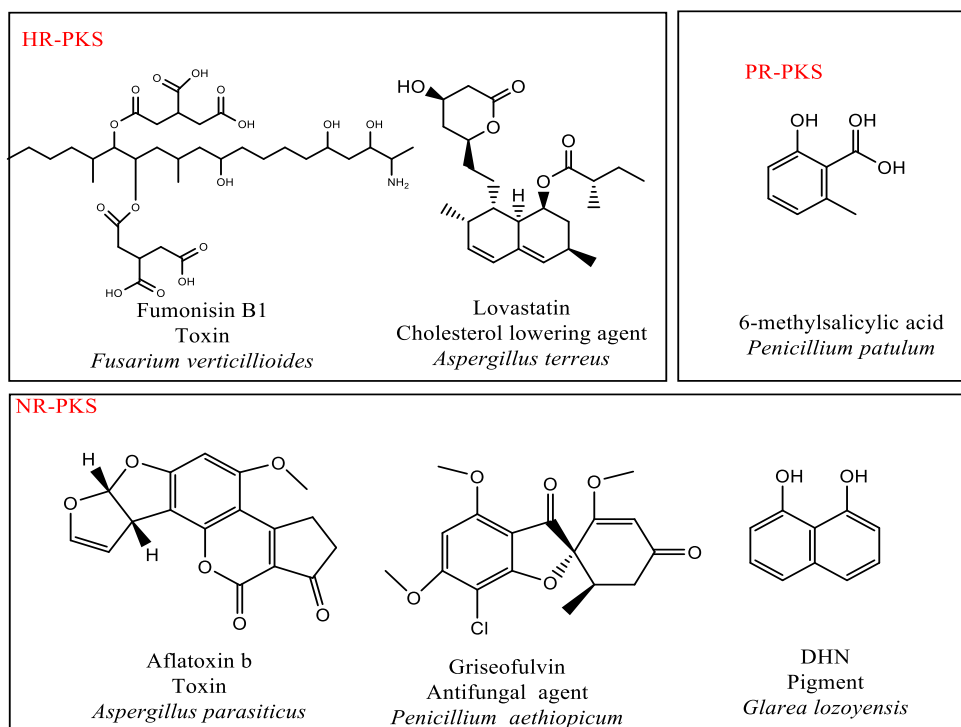


Fig. 1.17 The structure diversity of fungal polyketides.

In contrast to type I PKS, type II PKS was found only in bacteria and are dissociable multi-enzyme complexes. Each protein in the complex bears a single and independent catalytic domain. KS and AT domain, acting iteratively during polyketide formation (Fig.1.18).<sup>18</sup>

Type III PKS are homodimeric enzymes found in fungi but mainly studied in plants. These enzymes only contain KS domain and do not depend on the acyl carrier protein to catalyze the extension of polyketone chains, making the type III PKS directly use malonyl-CoA as the substrate (Fig. 1.18).<sup>79</sup> A small amount of fungal type III PKSs have also been described. Fungal iterative PKSs have a cryptic programming mechanism. It's impossible to predict the number of catalyzed cycles performed by the mega-synthase or the reduction/oxidation introduced on the PK.

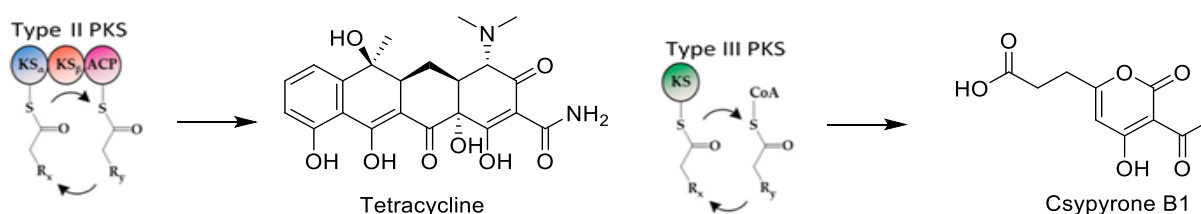


Fig.1.18 Workflow of type II and III of PKS.<sup>80</sup>

### 1.1.5 Metabolites from hybrid pathways

In fungi, several types of hybrid pathways have been described. Some pathways result from the association of two synthases while in other cases, the synthase itself is the product of the hybridization of two synthases. Illustrating examples are given here.

#### 1.1.5.1 Metabolites requiring two synthases

Hybrid pathways involving highly reducing PKS (HR-PKSs) and non-reducing PKS (NR-PKS) have been characterized in fungi. In this system, HR-PKS first forms an intermediate which is directly transferred to the AT domain of NR-PKS to perform the following reaction. For example, Desmethyl-lasiiodiplodin, *R*-Zearalane and Lasicicol were obtained by the collaboration of HR-PKS and NR-PKS (Fig. 1.19).<sup>81</sup> Their biosynthetic pathways have been extensively studied through combinatorial biology.

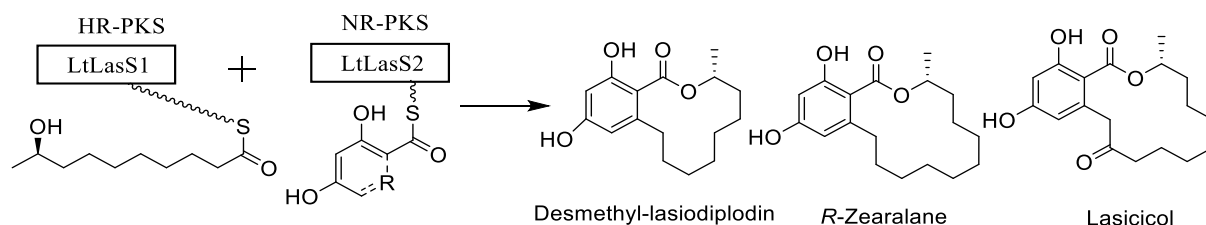


Fig. 1.19 Three PKSs obtained by collaboration of HR-PKS and NR-PKS. (R indicates the HR-PKS product that is elaborated by the NR-PKS).

The biosynthetic pathway of meroterpenoid is another example of hybrid pathway in which a NR-PKS generates an intermediate which is then integrated into a terpene pathway (Fig. 1.20).<sup>82</sup>

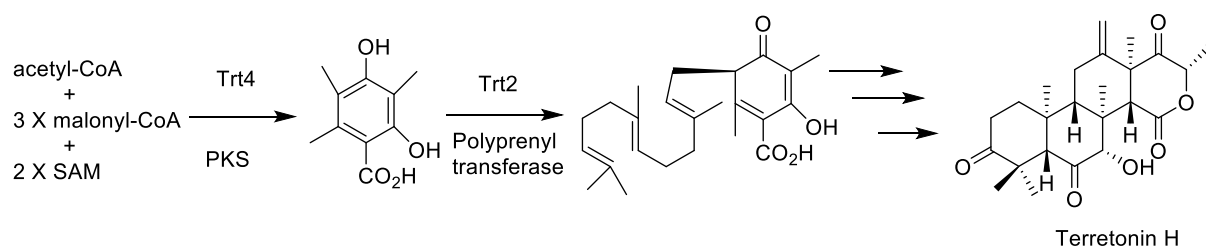


Fig. 1.20 Biosynthetic pathway of terretonin H

#### 1.1.5.2 Metabolites from hybrid synthases

##### Hybrid NRPS-PKS

The deciphering of the biosynthetic pathway of tenuazonic acid gave the first example of a hybrid NRPS-PKS in fungi.<sup>83</sup> This mega enzyme is composed of three domains (C-A-T) for

the NRPS part and a single KS domain for the PKS part. The latter plays the role of R domain and is responsible for Dieckmann cyclization (Fig. 1.21).

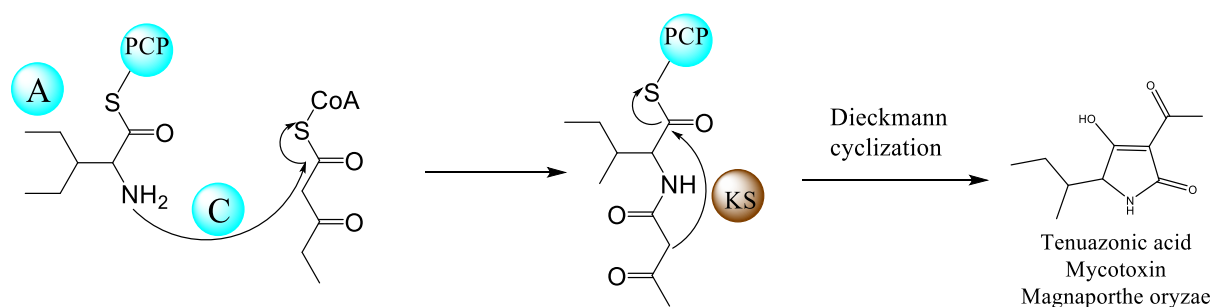


Fig. 1.21 Biosynthetic pathway of Tenuazonic acid

### Hybrid PKS-NRPS

Polyketide synthase and peptides synthetase (PKS-NRPS) is in charge of the biosynthesis of PK-NRPs. PKS-NRPS is composed of a hybrid of type-I iterative PKS and NRPS. The first fungus-derived hybrid PKS-NRPS *fusA* was linked to a metabolite fusarin C in 2004.<sup>84</sup> The hybrid PKS-NRPS is a mega-enzyme composed of about 4000 amino acids and ten catalytic domains associated with an extra enoyl reductase domain. PKS is devoted to the synthesis of polyketide. The NRPS machinery is in charge of selecting and activating the amino acid and binding it to the polyketide. (Fig. 1.22).<sup>67</sup> A reductase or Dieckman cyclization domain (R / DKC) is located at the end of the PKS-NRPS polymerase to perform the release of the intermediate. Thus, two possible approaches were identified to release the intermediate from the T domain. One way of the embedded terminal reductase domain of an NRPS catalyzes reduction and Knoevenagel condensation of the thioester intermediate to give a pyrrolin-2-one product via route A (Fig. 1.22), such as in the biosynthesis of fusarin C. Another way catalyzes a Dieckmann condensation of the intermediate to give a pyrrolidine-2,4-dione product via route B (Fig 1.22), like in the biosynthesis of equisetin.<sup>85</sup> Then other enzymes coded in the genome convert the intermediate to the final product by cyclase, oxidase, and reductase.

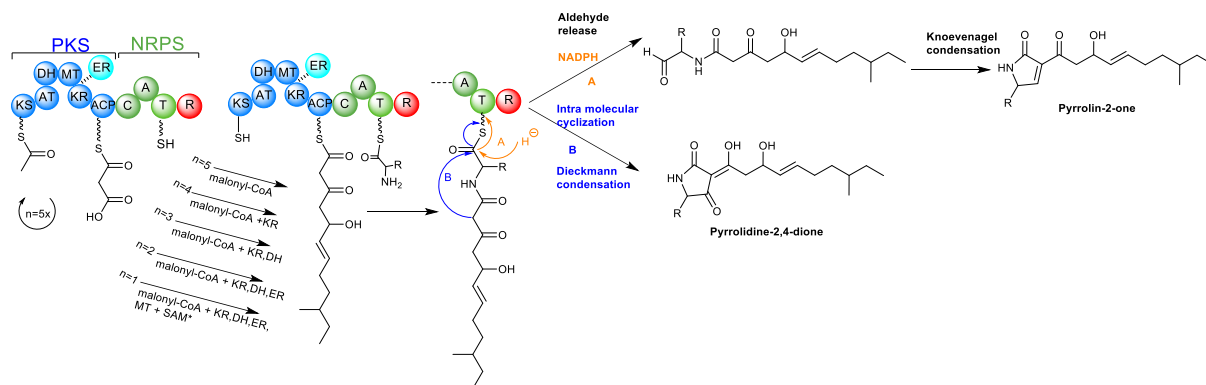


Fig. 1.22 Biosynthesis process of PKS-NRPS and release mechanism of R domain.<sup>67, 85</sup>

With the advancement of molecular biology and bioinformatics, so far, 25 fungal PKS-NRPS metabolic pathways have been elucidated (Fig. 1.23/1.24), namely fusarin C,<sup>84</sup> NG391,<sup>86</sup> equisetin,<sup>87</sup> tenellin,<sup>88</sup> bassianin,<sup>89</sup> aspidone,<sup>90</sup> chaetoglobosin,<sup>91</sup> cytochalasin,<sup>92</sup> pseurotin,<sup>93</sup> cyclopiazonic acid,<sup>94</sup> isoflavipucine,<sup>95</sup> pyranonigirin E,<sup>96</sup> pyrrolocin A,<sup>97</sup> leporin A,<sup>98</sup> sch210972,<sup>99</sup> myceliothermophin,<sup>100</sup> curvupallide A,<sup>101</sup> oxaleimide C,<sup>102</sup> UCS1025s,<sup>103</sup> burnettramnic acid A,<sup>104</sup> citridone B,<sup>105</sup> harzianic acid,<sup>106</sup> pyranterrone,<sup>107</sup> thermolide 1,<sup>108</sup> and harzianopyridone<sup>109</sup>.

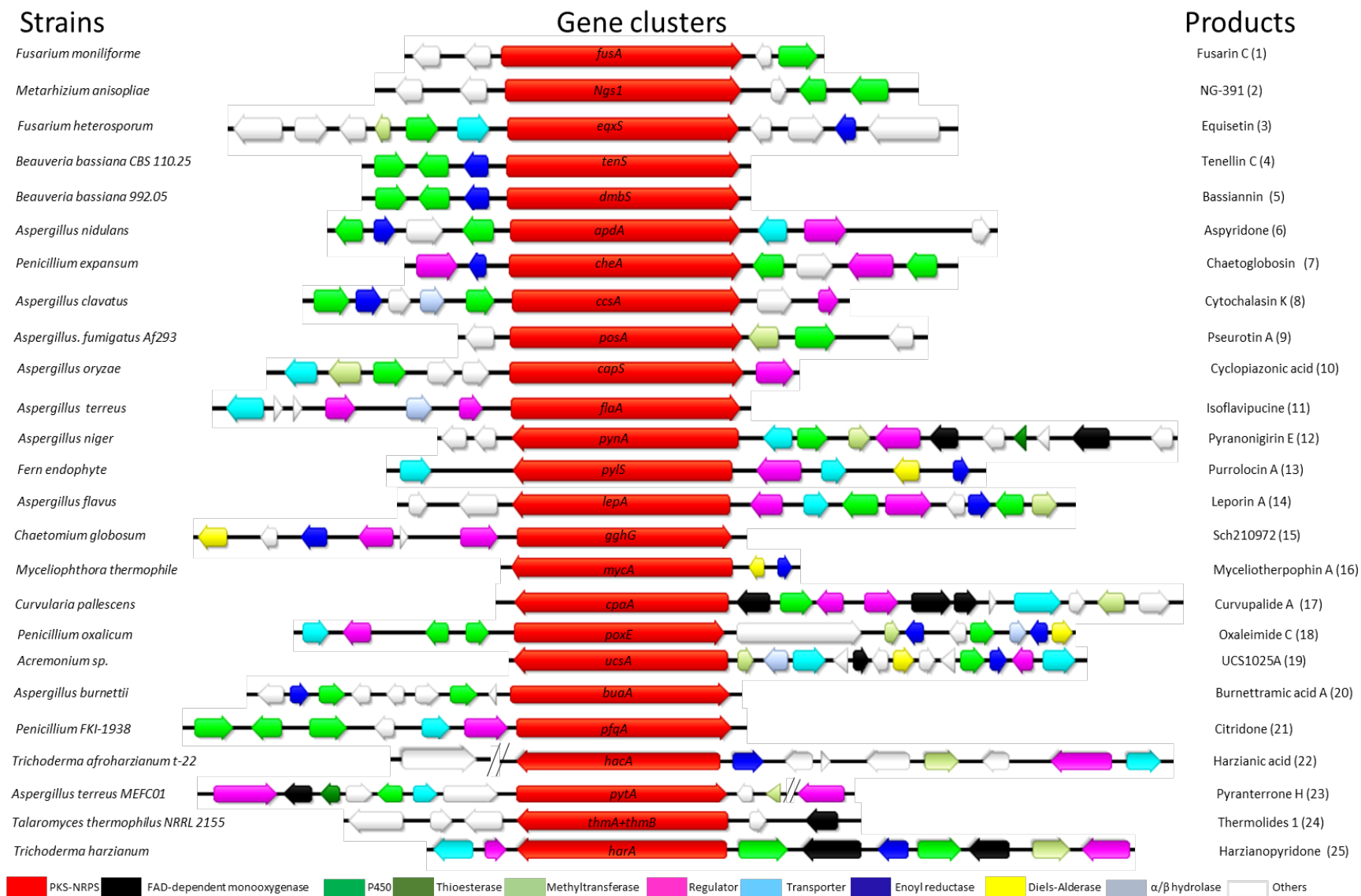


Fig. 1.23 The map of published fungal PKS-NRPS biosynthetic gene clusters corresponding to product.



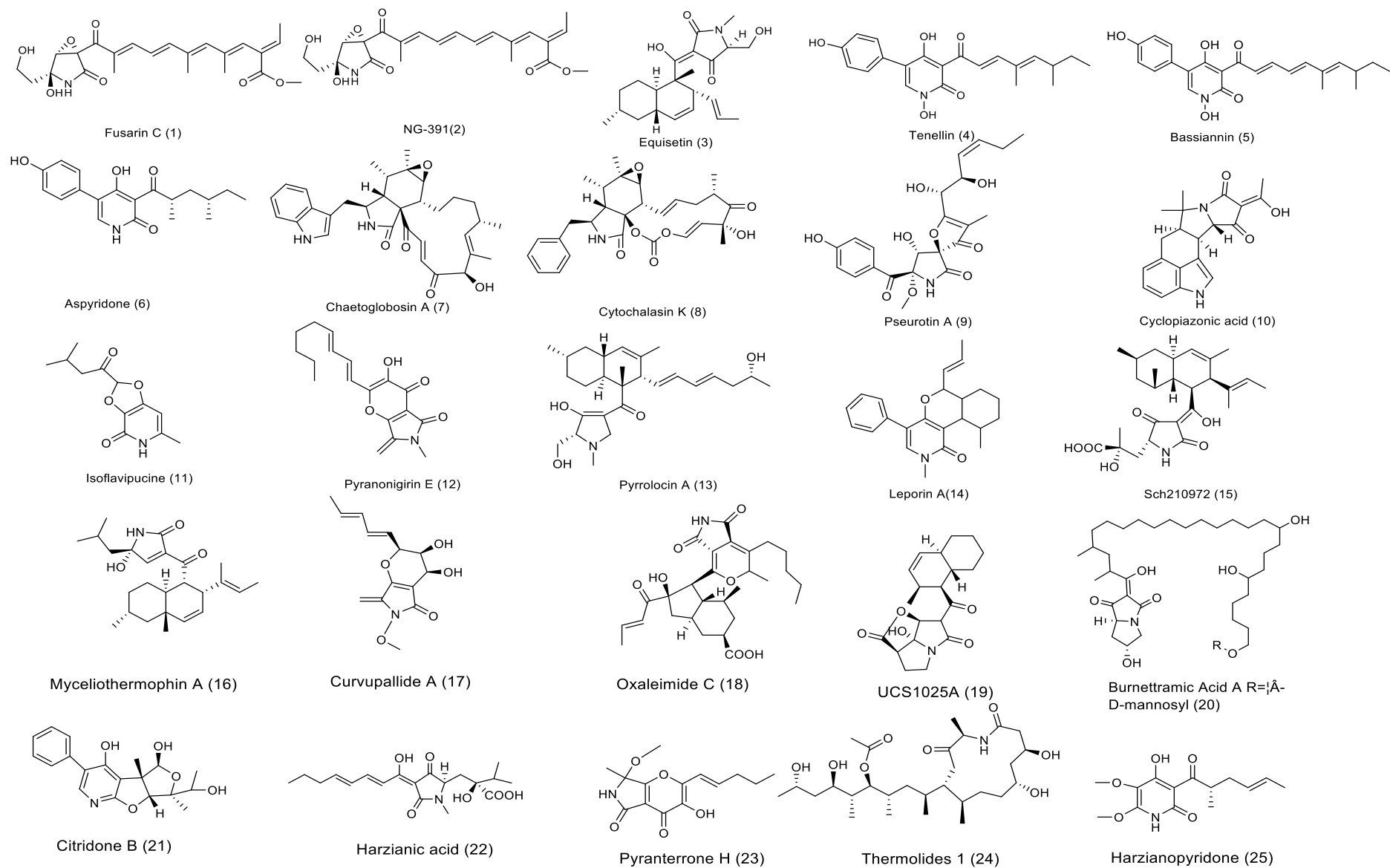


Fig. 1.24 The corresponding compounds of twenty-five fungal PKS-NRPS gene clusters.

### 1.1.6 Regulation of secondary metabolism

Epigenetics regulates the biosynthesis of secondary metabolites in various filamentous fungi. DNA methylation and histone modifications are the major epigenetic modifications in filamentous fungi which impact the structure of chromatin. Epigenetics refers to the heritable changes in genes under the condition that the DNA sequence in the chromosome remains unchanged, and this change can be inherited stably. These two types of modifications are closely related to various physiological states of filamentous fungi. DNA methylation refers to the methylation of a carbon atom of cytosine in the genome to form methyl-cytosine, which changes the molecular conformation of DNA and affects the binding of DNA and protein molecules, thereby affecting the expression of secondary metabolites. Histones are essential chromosome-related proteins that can provide binding sites for DNA and control transcription initiation. Histone acetylation is a switch that controls transcription through the affinity of lysine amino acids in histones to DNA, resulting in gene silencing and expression. Histone methylation represses or activates gene transcription through dynamic methylation state interactions with other functional proteins (Fig. 1.25).<sup>110</sup> Both histone and DNA methylation repress gene transcriptions. However, histone acetylation indicates that the gene is activated. Histone methylation and acetylation are reversible processes.

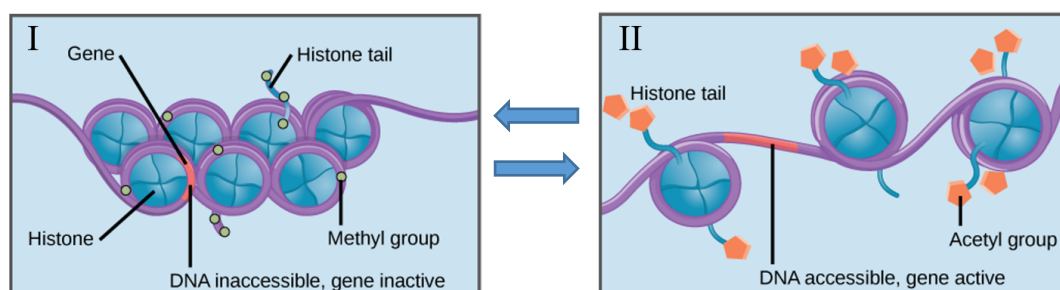


Fig. 1.25: Impact of histone modification on chromatin. I: Methylation of histones causes nucleosomes to pack tightly together corresponding to hetero-chromatin. Transcription factors cannot bind the DNA, and genes are not expressed. II: Histone acetylation results in loose packing of nucleosomes corresponding to euchromatin. Transcription factors can bind DNA and genes are expressed.

Fungi produce different secondary metabolites (gene clusters to be silent or activated) to response to external environment factors (pH, C source, N source and light).<sup>111</sup> This process is mainly achieved through regulatory factors (pathway-specific regulation or global regulation). Most of the regulatory factors have the characteristics of binding to DNA. By identifying and binding to the promoter sequence of the target gene, they can activate or silence the expression of target genes. Global regulation is achieved through the global regulator. Genes encoding

global regulators do not belong to any gene cluster but regulate the expression of a range of genes.<sup>112</sup> 60% of fungal secondary metabolite synthesis gene clusters contain pathway-specific regulatory genes.<sup>112</sup> Some pathway-specific regulators are indispensable for the activation of the entire gene cluster. A gene cluster can contain several regulatory genes. For example, three transcription factors ATEG\_03638, ATEG\_07666 and ATEG\_07667 regulate the expression of two gene clusters to synthesize Azasperpyranone A.<sup>113</sup> The regulation of biosynthesis of Azasperpyranone A also showed the interactive regulation mode in fungal secondary metabolism regulation. One regulatory gene can regulate two or more secondary metabolite synthesis of gene clusters. Pathway-specific regulator also has a positional effect on gene regulation with its location on the chromosome being crucial for the regulation. When transferring a gene out of the cluster, it loses the control under the specific regulators in the cluster. Moving specific regulators out of the cluster also lose the ability to control this gene cluster.

## **1.2 Discovery of new natural products**

In recent years, with the outbreak of viruses such as Influenza A, Ebola, SARS-CoV-2, researchers need to find suitable drugs for fighting against the new virus. With the development of AI technology, potential new drugs can be discovered through deep learning based on known compounds.<sup>114</sup> Based on the crystal structure of the critical hydrolase (mpro) in SARS-CoV-2, a combination of virtual screening and enzymatic testing can be used to obtain targeted drugs quickly.<sup>115</sup> The commercialization of a drug requires the screening of tens of thousands of compounds. So, the most critical and foundation things are to expand the diversity of natural products and to establish efficient biological or chemical synthetic pathways. Drug discovery is highly dependent on studies of natural products.

### **1.2.1 Traditional natural product discovery strategies**

The initial natural product discovery strategy is mainly to obtain natural products with various activities through chemical screening and activity tracking, followed by undergoing a series of chemical modifications to obtain more active compounds.<sup>116</sup> Diverse sample screening combined with activity tracking is an effective method for new natural product discovery (Fig 1.30). For five isolated fungal strains, only the crude extracts of *Emericella* sp. SMA01 exhibited strong antimicrobial activities. Rice medium was chosen as the best fermentation medium by comparing the yield and activities of the products from the fungus. After tracking the activity step by step during the separation process, phenazine-1-carboxylic acid with intense

antimicrobial activities was obtained.<sup>117</sup> Fifteen ocean fungi were isolated from *Onchidium struma*, and anti-inflammatory bioactivity studies were carried out on the crude extract of these fungi. The crude extract of *Aspergillus terreus* GZU-31-1 showed intense anti-inflammatory bioactivity. Aspermeroterpene A, B, and C were obtained by activity tracking to have strong anti-inflammatory bioactivity.<sup>118</sup>

Fungi accumulate different secondary metabolites in response of various physiological conditions (Fig. 1.26e). Culture condition screening known as “OSMAC”(one strain many compounds) is a feasible way to activate silent genes.<sup>119</sup> New compounds could be obtained by comparing metabolomic profiling from different culture mediums. Inactivation of the critical gene *rbtJ* in charge of the biosynthesis of the main product rubratoxins was performed in *Penicillium dangeardii*. Comparison of the different metabolomic profiling of the mutant and wild type extracts from different culture mediums leads to the identification of several new azaphilone with some of them have novel skeletons.<sup>120</sup> Using the OSMAC method, the yield of an unknown compound fusarielin (anticancer effects) was increased by 80 times, allowing it to be isolated and characterized.<sup>121</sup> Two new compounds fusarielin K and fusarielin J were obtained by culturing *Fusarium tricinctum* in rice medium mixed with fruit and vegetable juice.<sup>121</sup> Meanwhile, co-culture with other organisms is another efficient approach to activate the silent gene clusters under laboratory conditions.<sup>122</sup> Pestalone, a new chlorinated benzophenone antibiotic was isolated by co-culture of a marine fungus *pestalotia* with bacterium *SNJ-328*.<sup>123</sup> A new compound, namely wheldone that has extensive cytotoxicities was obtained by co-culture of *Aspergillus fischeri* and *Xylaria flabelliformic*.<sup>124</sup>

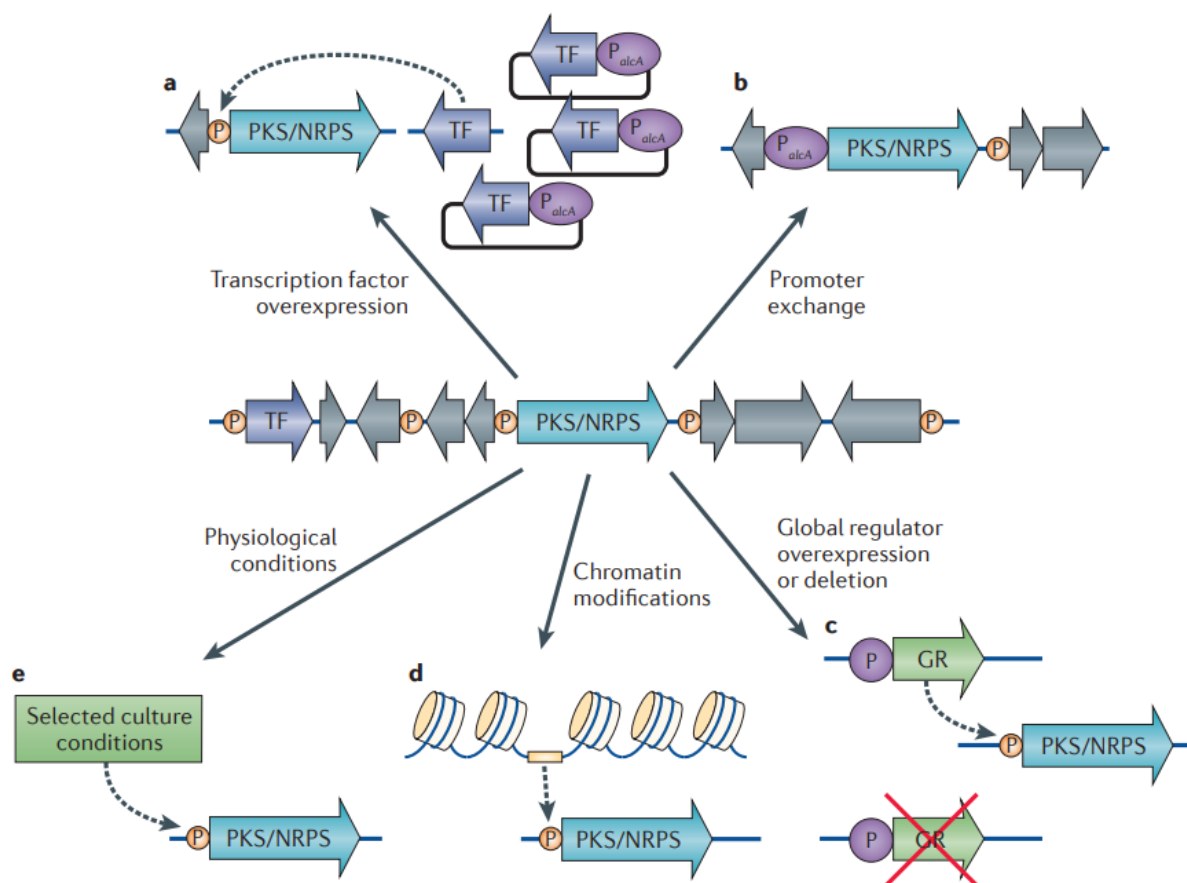


Fig 1.26 Strategies to discover novel natural product.<sup>125</sup>

### 1.2.2 Genome mining and activation of silent gene clusters

Traditional methods to obtain new compounds with good bioactivities are time-consuming and this approach often leads to reisolation of known compounds. Genes in fungi that encode biosynthetic pathways responsible to produce these secondary metabolites are often gathered into clusters, making it convenient to identify similar clusters through bioinformatics. Therefore the mined clusters can give access to new natural products.

The development of high-throughput sequencing technology, faster sequencing and affordable price allow more and more fungi to be sequenced and publicly available. Considerable software and websites have been developed to analyze gene clusters. For instance, HMMER is a website for gene homology analysis based on Hidden Markov Models, which is very useful for gene function analysis.<sup>126</sup> AntiSMASH is the software for predicting the gene cluster according to known gene clusters.<sup>126</sup> These all paved the way for genome mining.

For example, gene mining has been applied to discover natural herbicides in fungi. Dihydroacid dehydratase (DHAD) is an essential enzyme for branched-chain amino acids

synthesis in plants. Inhibiting the function of DHAD would negatively affect the growth of plants. So it's a good target as an herbicide. The gene clusters responsible for producing the metabolite that inhibit DHAD should contain a resistant gene whose sequence is similar to DHAD. Genome mining showed a resistance gene coding AstD with 60% similarity to DHAD in fungi, and a new natural herbicide aspterric acid was successfully discovered (Fig.1.27).<sup>127</sup>

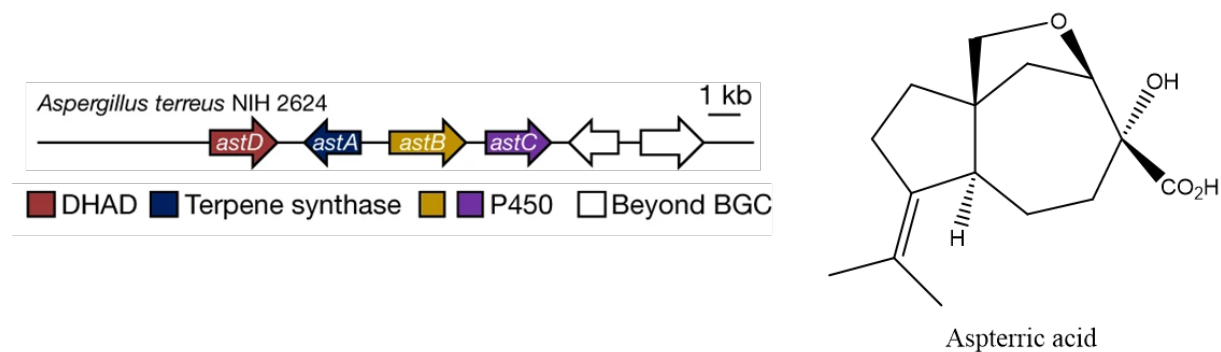


Fig. 1.27 The structure of aspterric acid and its related gene cluster.

Another example of genome mining is that the compound synthesis by hybrid NRPS and terpene synthase has not been reported previously. By genome mining, a gene cluster containing both enzymes were found in *Aspergillus flavus*. A new alkaloid heteropterpenoids flavunoidine was obtained by heterologous expression of the gene cluster in *Aspergillus nidulans*.<sup>128</sup>

Fungi usually encode more than 60 natural products. However more than 90% of gene clusters in fungi are silent in laboratory cultures.<sup>21</sup> It is necessary to harness this vast potential to obtain new metabolites that could be potential drugs. Herein several approaches have been established to activate the silent gene clusters.

Overexpression of the pathway-specific regulator has been successfully applied to activate the silent gene cluster (Fig. 1.26a). For example, the overexpression of specific transcription factor *BlnR* has activated the cryptic gene cluster *bln* to obtain balanol and its series of derivatives.<sup>129</sup> The over-expression of specific transcription factor *PabHLH* has accelerated the accumulation of bisbibenzyl.<sup>130</sup>

Another approach to activate the silent gene cluster is to swap the native promoter with other strong promoters (Fig. 1.26b). For instance, the exchange of the promoter of *Aspergillus.nidulans acvA* with *alcA* can result in a thirty-fold increase in penicillin production.<sup>131</sup> An inducible promoter combined with brevianamide gene was integrated into

the genome of *A. fumigatus* to discover several compounds that has not been produced before.<sup>132</sup>

The inactivation or overexpression of the global regulator can also result to inactivating the gene clusters (Fig. 1.26c). For instance, LaeA is a well-described global regulator which regulates the production of secondary metabolites through chromatin remodeling. Antitumor compound terrequinone A was obtained by knocking out the *laeA* in *Aspergillus nidulans*.<sup>133</sup> Another example is that transcriptomic analysis of the original strain of *Aspergillus fumigatus* and mutants with a knockout of *laeA* revealed that 12 of the 22 secondary gene clusters were regulated by *laeA*.<sup>134</sup> The overexpression of a global regulator *laeA* increased the production of penicillin in *Aspergillus. nidulans*.<sup>135</sup>

Epigenetic regulation has been applied to activate the gene clusters (Fig. 1.26d). Inactivation of histone deacetylase and histone methylase or overexpression of histone acetylase is another feasible way to perform epigenetic regulation. Inactivation of the histone deacetylase *hdaA* in *Aspergillus nidulans* resulted in activating the penicillin and sterigmatocystin synthetic gene clusters.<sup>136</sup> Overexpression of the histone acetylase GcnE in *Aspergillus nidulans* efficiently activates the expression of the aspergillus and penicillin gene clusters.<sup>137</sup> Inactivation of histone methyltransferase CclA in *Aspergillus nidulans* activates multiple gene clusters and obtains eight new compounds.<sup>138</sup>

For fungi that are difficult to genetically manipulate, adding inhibitors to the medium is a straightforward approach for epigenetic regulation (Fig. 1.28). DNA methyltransferase inhibitors can covalently bind to DNA methyltransferases to competitively inhibit the methylation of cytosine and uracil, thus reducing the methylation level of DNA molecules and increasing promoters and transcriptional regulation in genes specificity of factor recognition sites to activate gene transcription and expression. Two new PKs lunalin A and B were obtained after adding DNA methyltransferase inhibitor 5-azacytidine to the medium of *Cladosporium*.<sup>141</sup> Histone deacetylase inhibitors can bind to the active site of histone deacetylase to inhibit the activity of the enzyme, thus increasing the level of histone acetylation. The relaxation of nucleosome structure is conducive to transcription factors, and cooperative transcription factors, binding to DNA molecules increases the level of transcription to activate the expression of silent genes. The histone deacetyltransferase inhibitor SAHA can effectively activate *Cladosporium* to produce novel compounds cladochromes and calphostin.<sup>141</sup>

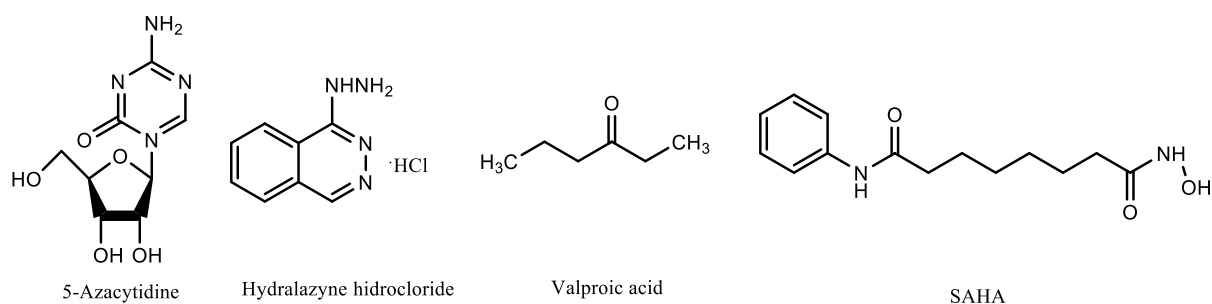


Fig.1.28 The structures of inhibitors (DNA methyltransferase inhibitors: 5-Azacytidine and hydralazine hydrochloride. Histone-deacetylase inhibitors: valproic acid and SAHA).

Most of the previous processes require genetic manipulation in fungi. So how to perform it in fungi? Compared to prokaryotes and yeast, fungal genetic manipulation has low targeting efficiency and troublesome steps. Several approaches have been established for gene manipulation in filamentous fungi as below:<sup>142</sup> 1) *Agrobacterium tumefaciens*-mediated transformation (ATMT): the plasmid constructed in *Escherichia coli* was transferred into *Agrobacterium* by electro-transformation. *Agrobacterium* was co-cultured with fungal spores and the virulence region (*vir* gene) on the plasmid that began to be expressed under the inducibility of acetosyringone, resulting in homologous recombination or ectopic insertion (random integration of foreign genes into the genome of the target strain) occurring when the cassette was transferred into fungi. No restriction apply to the transformation recipients of AMAT, but this method is cumbersome, time-consuming and the targeting efficiency is affected by many factors. 2) Electroschock conversion: this requires the preparation of protoplasts by removing the cellular wall through enzyme digestion or fresh spores under high voltage for a short period to increase the permeability of cells, thereby leading to the intake of foreign DNA. The electroschock method is to place fungal protoplasts, it's a cumbersome process that is fungal dependent and needs to be treated with distinct enzymes.<sup>143</sup> 3) Biolistic transformation: DNA is coated with tungsten or gold and accelerated into cells at high velocity. The requirement of specific machines limits the application of this method.<sup>144</sup> 4) Lithium acetate conversion: lithium acetate was used to prepare fungal spores or protoplasts competent cells, and DNA enters the fungal cells under heat shock with the help of PEG to obtain mutants.<sup>143</sup> 5) PEG/CaCl<sub>2</sub>-Mediated Transformation: the control of osmotic pressure is critical in protoplast preparation and regeneration. Fungal protoplasts were treated with CaCl<sub>2</sub> to open the channels in the cytomembrane, and exogenous genes entered the protoplasts with the help of PEG to obtain mutant strains. High copy number of inserted DNA was required by using this approach. 6) CRISPR/Cas9: The occurrence of double-strand breaks at specific DNA sites in the genome



will significantly improve the efficiency of homologous recombination.<sup>145</sup> Therefore, CRISPR/Cas9 was developed in fungi by inducing the repair and editing of the broken DNA to realize the specific editing of the genome. CRISPR/Cas 9 technology is an accessible, high-efficiency, and multi-function gene-editing technology.

### 1.2.3 Heterologous expression

Many heterologous expression systems have been developed to express fungal gene clusters. *E.coli* is an ideal host for prokaryotic heterologous expression systems due to its simple operation, fast growth rate, and transparent genetic background. Considering the codon preference, shortage of precursors required for fungal compound synthesis and the presence of introns in fungi, *E.coli* has significant limitations as a heterologous expression host for fungal gene clusters. Nevertheless, it is a suitable platform for purification and study the key enzymes from fungi.<sup>146</sup> A special example is that the prokaryotic *Bacillus subtilis* was successfully used as the expression of the NRPs gene cluster in the fungus *Fusarium oxysporum*; the non-ribosomal cyclic depsipeptide enniatin was obtained, and enniatin was directly excluded from the extracellular space, making it beneficial to the fermentative production of enniatin.<sup>147</sup>

As a eukaryotic organism, yeast grows fast and has mature genetic manipulation tools, without endotoxin. It can also correctly fold and post-translationally modify proteins. *Saccharomyces cerevisiae* has multiple auxotrophic selection markers, making it a good host for heterologous fungal gene clusters. Gene clusters from diverse fungal species were separately expressed in modified *Saccharomyces cerevisiae*. Related products could be detected within twenty-two of them. Some new compounds with unexpected biosynthetic origins have been obtained.<sup>148</sup> *Saccharomyces cerevisiae* expresses the *matB* and *npgA* genes through modification, *matB* encoded malonyl-CoA synthetase and converts malonate into malonyl-CoA to provided start units or elongation units for PKS synthesis. *NpgA* encoded PPTase to ensures the flexibility of the carrier protein. It was successfully used for the expression of five genes in *chaetomium globosum* individually, resulting in the accumulation of six polyketides.<sup>149</sup> The *apdA* gene cluster was silent in *Aspergillus nidulans*, and aspyridone A was obtained by overexpressing the specific transcription factor. The correlation between *apdA* and aspyridone was further confirmed by co-expression of *apdA* and *apdC* in *Saccharomyces cerevisiae*. The biosynthetic pathway was also proposed.<sup>90</sup> Yeast plays an essential role in the discovery of fungal natural products, but it also has shortcomings. It cannot recognize fungal introns and lacks an advanced mRNA splicing system. Moreover, it is challenging to operate on gene

clusters containing more than three genes. *Pichia pastoris* can also be used as a fungal heterologous expression host, however, due to low recombination efficiency and lack of neutral sites for genome integration, it is difficult to genetically modify *Pichia pastoris*, which significantly limits the application of *Pichia pastoris*

*Aspergillus oryzae* is a safe strain for brewing soy sauce and wine. It has a transparent genetic background which can discharge considerable products into the extracellular space, cut introns and post-translational modification functions. It is an excellent host for the heterologous expression of fungal gene clusters.<sup>150</sup> More than 10 HR-PKS genes were found in the genome of *Alternaria solani*, Four new compounds were obtained by heterologous expression of these gene clusters in *Aspergillus oryzae*.<sup>151</sup> Using this method, didymellamide was isolated by further gene mining based on the *Aspergillus solani* transcriptome.<sup>152</sup> Four terpene synthase genes with the equivalent cyclization mechanisms were found in *Neosartorya fischeri* and *Aspergillus oryzae*, but no terpenoids were extracted from *Neosartorya fischeri* and *Aspergillus oryzae*, which means these genes are silent in lab conditions. Four genes were integrated into two plasmids and transferred into *Aspergillus oryzae* to obtain a new sesterterpene sesterfisherol. The biosynthetic pathway of sesterfisherol was elucidated by isotope labeling combined with in vivo experiments.<sup>153</sup> Four novel meroterpenoids funiculolides A-D were isolated by heterologous expression of a silent gene cluster from *Aspergillus funiculosus* CBS 116.56. Their biosynthetic pathways were further elucidated by in vitro assay.<sup>154</sup> These results demonstrated that *A. oryzae* is a good heterologous expression system for mining fungal natural products and dissecting the metabolic pathways of fungal-derived natural products.

One of the original strains for the heterologous expression of penicillin is *Aspergillus niger*, which has significant advantages for producing non-ribosomal polypeptides. *Aspergillus niger* is a safe strain that does not contain endotoxin. The potent immunosuppressant FR901483 was isolated from *Cladobotrym* sp. Its biosynthetic pathway was characterized by heterologous expression in *Aspergillus niger*.<sup>155</sup> Terrein produced by *Aspergillus terreus* has antimicrobial, antiproliferative, and antioxidative effects. The biosynthesis gene cluster of this compound was unknown. When expressing a PKS gene *terA* in *Aspergillus niger* that may synthesize pigment, it occasionally discovered the analogs of terrein which means it successfully linked the *terA* to the terrein.<sup>156</sup>

*Aspergillus nidulans* is one of the earliest fungi with its genome sequencing to be completed. It has the most advanced and effective genetic system that can produce abundant

secondary metabolites with simple genetic manipulation in a mature transformation system as well as efficient targeting system.<sup>157</sup> An efficient heterologous expression system of *Aspergillus nidulans* was established by Berl Oakley, Nancy Keller, and Clay Wang.<sup>158-160</sup> First, a model strain with a simple background was obtained by knocking out the synthetic pathway genes of some primary metabolites of *Aspergillus nidullans*, which reduced the influence of its metabolites on the heterologous expression system. Then, A process for efficient conversion of secondary metabolites has also been established. Phusion PCR combined the target fragment with a unique promoter. *amyB*, *gpdA*, and *alcA* were used to promote the expression of the target gene. The vector was assembled by homologous recombination in yeast and transferred into *A. nidulans*. The side sequence in the vector was designed to be the homologous sequence of the *wA* gene related to the color change of *Aspergillus nidulans* spores. The heterologously expressed gene is integrated into the *ma* site, resulting in the incompleteness of the *wA* gene and changing the color from green to white in the successfully transformed spores. The screening efficiency is significantly improved. This strain has been widely used for heterologous expression of large gene clusters in fungi. Most of the PKS-NRPS gene clusters of fungi use this system for heterologous expression. The gene cluster for biosynthesis of asperfuranone in *Aspergillus terreus* was silent under different culture conditions. By the heterologous expression of the entire gene cluster in this system, asperfuranone was linked to the gene cluster. The biosynthetic pathway of asperfuranone was elucidated by combining different genes expressed in *Aspergillus nidulans*.<sup>159</sup> By genome mining of pathogenic fungi, a homologous gene cluster homologous to the gene cluster in *Aspergillus fumigatus* that produces an immunosuppressive polyketide is found in pathogenic fungi *Trichophyton* and *Arthroderma*. Neosartoricin B was obtained through heterologous expression of this gene cluster in *Aspergillus nidulans*, suggesting it play an important role in the pathogenesis.<sup>161</sup>

After the in-depth design of gene clusters, the heterologous expression can promote the discovery of natural products. For example, by inserting strong promoters in front of gene clusters, new compounds can be obtained by activating silenced gene clusters through heterologous expression.<sup>162</sup> More and more heterologous expression systems are being developed to express fungal gene clusters. The expression of fungal natural products in heterologous hosts will be a powerful tool for discovering new compounds. Heterologous expression is also suitable for producing uncultivable fungal natural product and analyzing of biosynthetic pathways.

#### 1.2.4 Combinatorial biosynthesis and synthetic biology

The analogs of natural products have great potential for drug discovery. Due to the complexity of the structure, the chemical synthesis process was often complicated, and the yield was unsatisfactory. On the contrary, synthetic biology is an excellent way to discover analogs of natural products.

Water solution and metabolic stability are the key factors in ensuring drug efficacy. The sugar-based modification can improve the water-soluble of drug molecules, thereby improving effectiveness. For example, erythromycin (medical) and tyladin (agricultural) showed antibiotic activity only after glycosylation modification. However, the compounds with glycosylation modification are easily hydrolyzed *in vivo* and have poor metabolic stability: Further modification of the glucose group into a sugar methyl group can increase the metabolic stability while maintaining the water solution and better exerting the efficacy of medicinal effects. Compared with chemical methods, biological methods for methyl glycosylation of drug precursors with significant application potential have the advantages of accurate modification and environmental friendliness. Synthetic biology was applied to solve this problem.

The research published in PNAS demonstrated the discovery of a glycosyltransferase-methyltransferase module belonging to a novel glycosyltransferase family in the fungus *Beauveria bassiana* and the glycosyltransferase has substrate promiscuity.<sup>163</sup> Using a combinatorial biology platform, the structural modification of 45 prodrugs such as flavonoids, anthraquinones, and naphthols was successfully achieved in yeast, which significantly improved the water solubility of these compounds. The metabolic stability of these compounds increased from an average of around 50% to more than 95% in various models. It is worth mentioning that this modification significantly improves the ability of two compounds, desmethyl-lasiodiplodin and monocillin II, to inhibit the spread of cancer cells while reducing the toxicity to normal cells, greatly increasing the potential of this compound in the development of anticancer drugs.

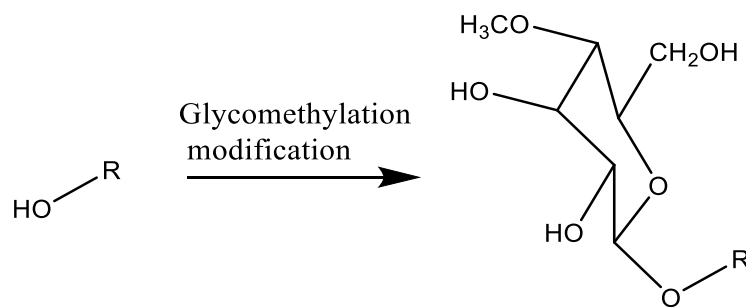


Fig.1.29 Glycomethylation modification by enzymes.

Benzenediol lactones are a class of fungal polyketide natural products with diverse chemical structures and biological activities. The assembly of the core backbone is catalyzed by a pair of iterative polyketide synthases (iPKSs). In these dual iPKSs systems, highly reducing PKS (hrPKS) synthesizes a linear polyketide intermediate. The SAT domain of non-reducing PKS (nrPKS) accepts this intermediate. Regioselective aromatic cyclization of extended polyketide intermediates was catalyzed by product template (PT) domains. The final product is released by the C-terminal thioesterase (TE) or reductase (R) domains of the nrPKS.

The combination of hrPKS and nrPKS from different sources or the combination of hrPKS, SAT domain from its related nrPKS and nrPKS lacking SAT domain from other organisms were expressed in the *S. cerevisiae* BJ5464-NpgA. This combinatorial biology work of Molnár et al. showed that combinatorial biology holds great potential for access to new compounds and SAT domains are polyspecific.<sup>81</sup>

Their more profound job is to combine the hrPKS, SAT domain, TE domain and the rest of nrPKS originated from different sources of benzenediol lactones synthase (Fig. 1.30),<sup>164</sup> showing that SAT<sub>AtCurS2</sub> has good promiscuity, and nrPKS carrying SAT<sub>AtCurS2</sub> can smoothly transport different starting units with varying chain lengths, redox degree and terminal hydroxyl configuration. However, SAT<sub>RrDalS2</sub> shows apparent selectivity to the hrPKS product structure. The SAT<sub>RrDalS2</sub> of nrPKS tends to transport unconventional, short-chain (preferably triketone) initiation units. This indicates that SAT can "actively" compete with the catalytic domain of hrPKS to block the polyketide intermediate (exogenous program) attached to ACP based on its preference for specific substrates and catalytic efficiency (endogenous program) to intercept hrPKS in advance unprogrammed product. The nrPKS of TE<sub>AtCurS2</sub> generally releases cyclic dihydroxyphenylacetic acid lactones through intramolecular esterification, while TE<sub>RrDalS2</sub> generally generates linear acyl dihydroxyphenylacetic acid esters through intermolecular esterification. TEs can determine the preference of hydroquinone lactone synthases for

intramolecular or intermolecular esterification reactions according to their endogenous programs. Meanwhile, TE can play a "molecular sieve" function, promoting or hindering the release of products derived from specific starting units according to its endogenous program, thereby regulating the metaprogram of upstream hrPKS and even the entire dual iPKSs system.

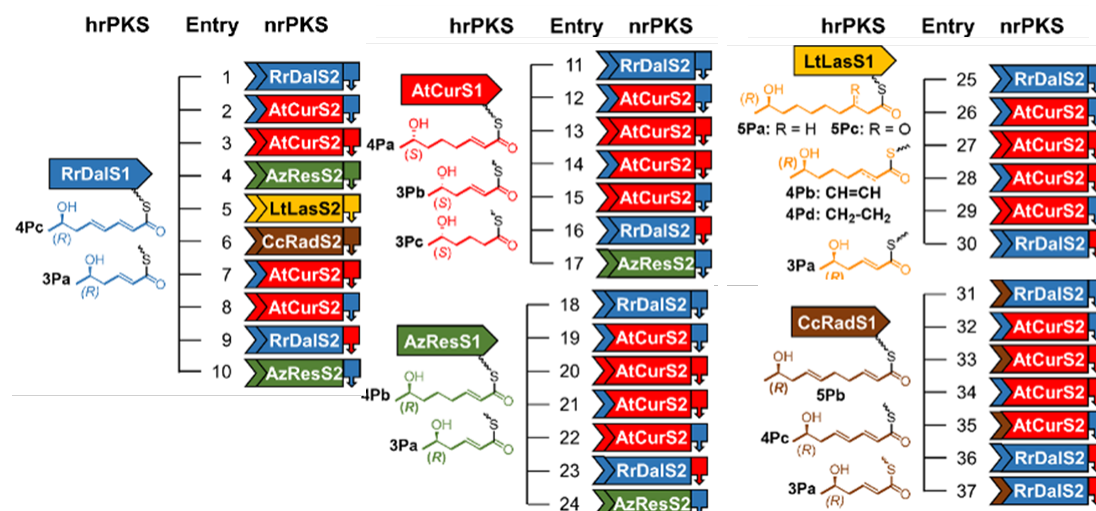


Fig. 1.30 Combinatorial biosynthesis with the benzenediol lactones synthase heterocombinations.

Synthetic biology is a rigorous engineering discipline to create, control and program cellular behavior. Here we introduce two platforms that are recently established based on synthetic biology for heterogeneous expression for fungal genes. One is termed as DIVERSIFY platform (Fig. 1.31).<sup>165</sup> First, they designed a common synthetic gene integration site (COSI) which is composed of 500 bp upstream homologous arm (A1), *gpdA* promoter (P), *uidA* reporter gene (*uidA*), *trpC* terminator (T) and 500 bp downstream homologous arm (B1)). Next, the COSI was intergreted into *A. nidulans*, *A. aculeatus*, *A. niger*, and *A. oryzae* separately. And then, a gene expression cassette (GEC) was designed which is composed of A1, P, gene of interest (GOI), T and B1. Finally, CRISPR-Cas9 mediated the homogous recombination to replace the reporter gene with the GOI. Correctly targeted transformants will be white, whereas incorrectly targeted transformants will be blue due to  $\beta$ -glucuronidase activity (Fig. 1.31). The disadvantage of this platform is that it can only be used for the expression of a single fungal gene.

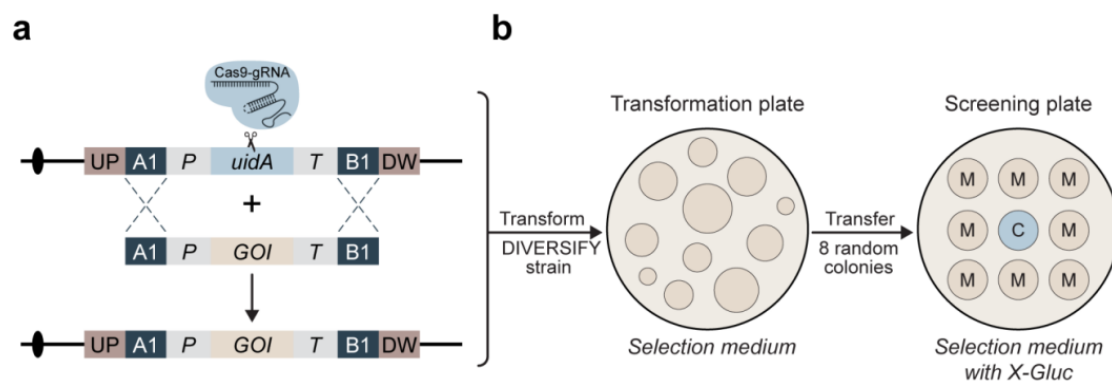


Fig. 1.31 Overview of the DIVERSIFY platform. (a) CRISPR mediated GEC insertion into *COSI-1*. (b) Screening for positive transformants.

Another platform used the same principle as the yeast toolkit we introduced in the introduction of chapter III. *Aspergillus nidulans* or *Penicillium rubens* is applied in this platform as a heterologous expression host for multiple fungal genes.<sup>166</sup> Successfully expressed genes can also be screened through the color of a single clone.

### 1.3 Pyrrocidines and the endophytic fungi *Sarocladium zae*

#### 1.3.1 Pyrrocidines and their analogs, an expanding compound family

Several fungal metabolite analogs of pyrrocidines have been described from various fungi (Fig. 1.32). Such as Penicipyrrodiether A,<sup>167</sup> GKK1032A2,<sup>168</sup> Hirsutellone B,<sup>169</sup> Pyrrospirone,<sup>170</sup> Embellicine A,<sup>171</sup> Trichobamide A,<sup>172</sup> Ascomylactam A,<sup>173</sup> Embellicine B,<sup>171</sup> Phomapyrrolidine B,<sup>174</sup> Hirsutellone F,<sup>175</sup> 19-epi-pyrrocidine A,<sup>176</sup> Phomapyrrolidone C,<sup>174</sup> Pyrrocidine C\*,<sup>177</sup> O-methyl-pyrrocidine B,<sup>178</sup> Penicipyrroether A.<sup>179</sup> They are attractive compounds because of their different bioactivities. If we decipher the biosynthetic pathway of pyrrocidines, we can envisage accessing new analogs by performing genome mining in sequenced fungal genome databases.

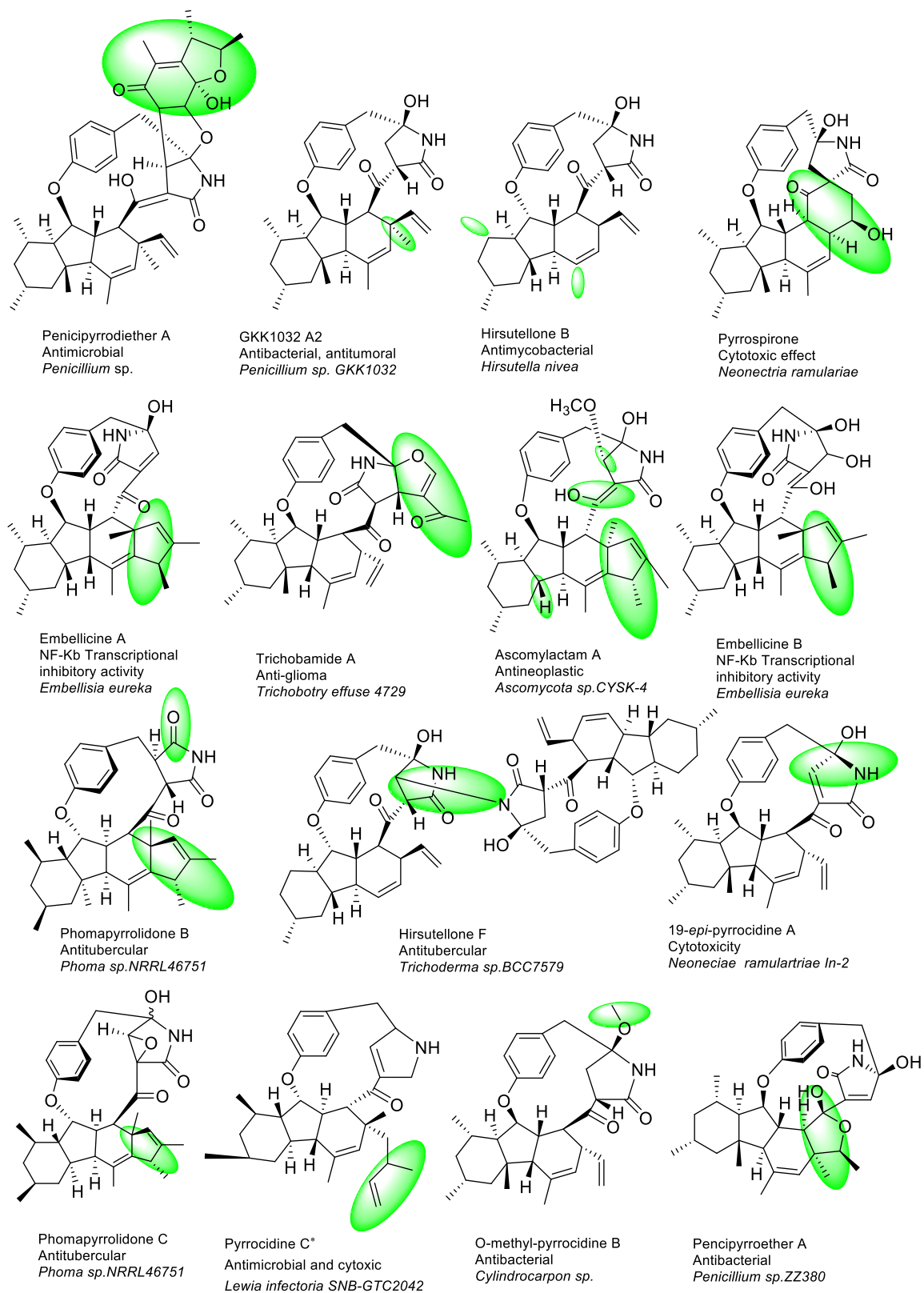


Fig. 1.32 Analogs of pyrrocidines from the fungal metabolites.



### 1.3.2 Biosynthetic origin of pyrrocidines

In order to determine the metabolic pathway leading to molecules in this family, some supplementation studies with labeled precursors have been conducted. First, feeding experiments with  $^{13}\text{C}$ -labeled compounds (tyrosine, acetate and S-methyl-methionine) as substrates for the fungus *Penicillium* sp. 1032GKK were performed. Isotopic analysis of the resulting products concluded that the precursor of GKK1032A is constructed from a tyrosine and a chain of nine acetate units flanked by five methyl groups from methionines probably by a hybrid PKS-NRPS. (Fig.1.33).<sup>168</sup>

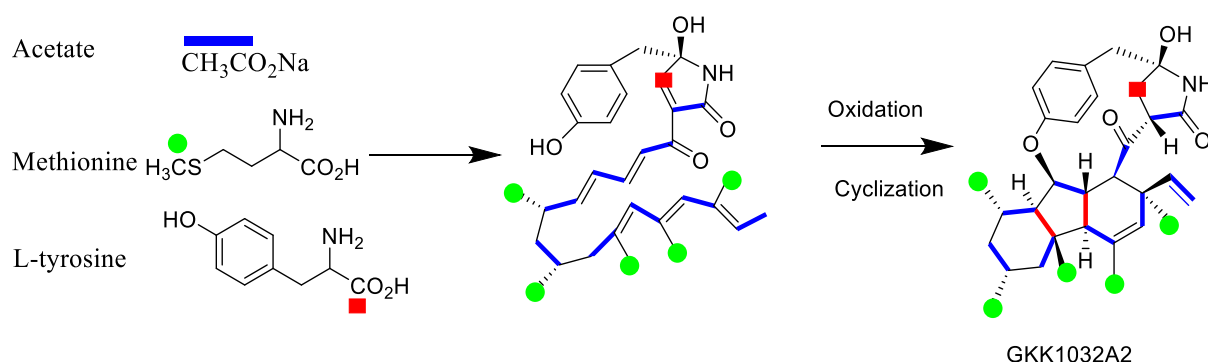


Fig.1.33 Isotopic labeling experiments performed with the fungus *Penicillium* sp. producing GKK1032A<sub>2</sub>.

Considering the structural analogy between pyrrocidines and GKK1032A<sub>2</sub>, it could be hypothesized that a hybrid PKS-NRPS is involved in the building of the pyrrocidine backbone. This PKS-NRPS should assemble nine acetates, one tyrosine and four S-methyl-methionines to give a linear intermediate aldehyde which is subsequently cyclized and oxidized by other enzymes (auxiliary enzymes) to form pyrrocidines (Fig. 1.34).

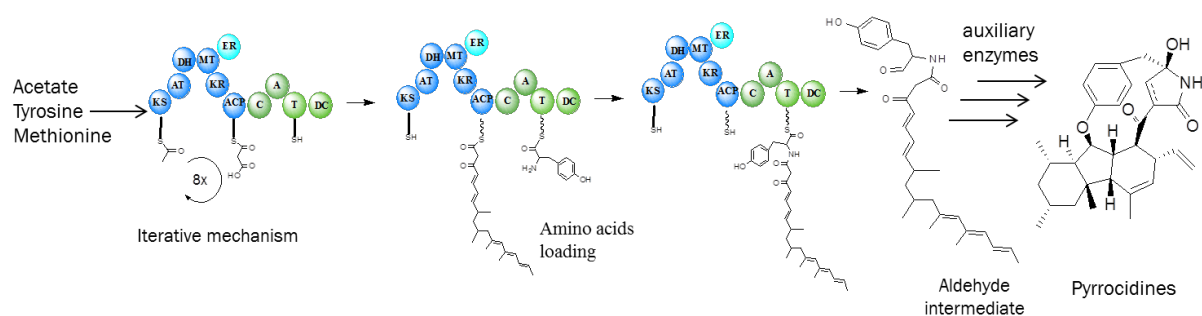


Fig.1.34 Proposed biosynthetic pathway for pyrrocidines.

Six steps are required to convert this aldehyde intermediate to pyrrocidine A (Fig. 1.35): 1) the formation of lactam ring; 2) the formation of the ether bond; 3) C7-C12 bond formation to give the cyclohexane; 4) the Diels-Alder reaction; 5) the insertion of the hydroxyl group on C20; 6) the rearrangement of the double bonds.

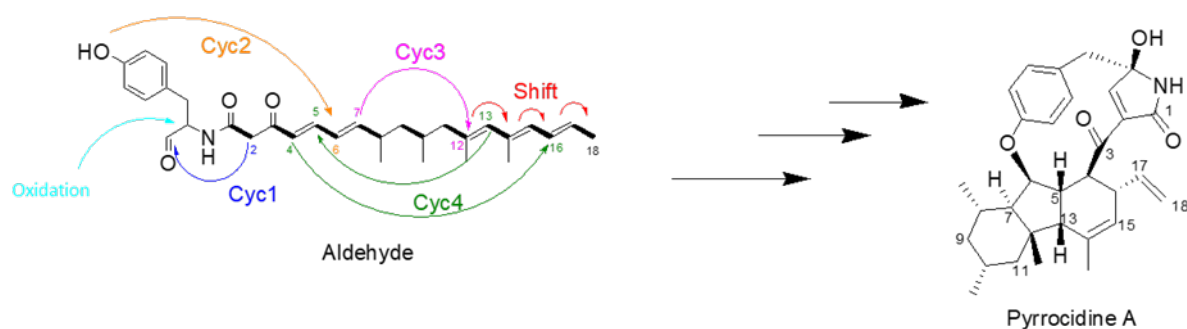


Fig. 1.35 Required steps from linear aldehyde to pyrrocidine A.

It can easily be admitted that the first step is the formation of lactam by a Knoevenagel reaction, indeed aldehyde is the most reactive function present in the molecule. An experiment was carried out in our laboratory, aiming to explore the mode of formation of the paracyclophane in pyrrocidine biosynthesis. For this purpose, doubly labeled tyrosine,  $^{13}\text{C}$  and  $^{18}\text{O}$ , was synthesized and used to supplement the culture medium of the fungus *S. zeae*. LC/MS analysis showed that ( $^{13}\text{C}$ ,  $^{18}\text{O}$ )-tyrosine had been incorporated into the pyrrocidines A and B produced during the experiment. This result showed that phenolic oxygen was retained during cyclization (Fig.1.36).<sup>180</sup>

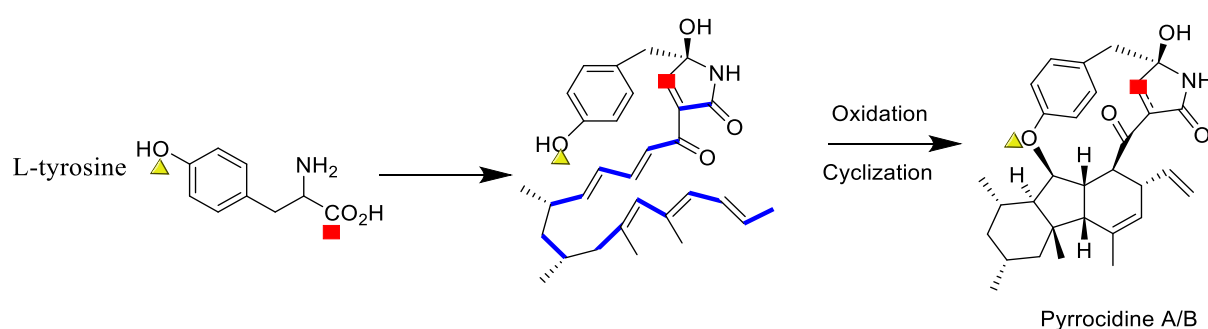


Fig.1.36 Isotopic labeling experiments performed with the fungus *S. zeae* showing the incorporation of the dual labeled tyrosine in pyrrocidines.

According to this results, three different mechanisms were proposed for the formation of cycles depending on the type of activation:

- The first possibility is the hydroxylation at C1, corresponding to allylic oxidation (Fig. 1.37a); Dehydration and rearrangement of the double bonds lead to the formation of the cyclohexane followed by the attack of phenolic oxygen on the formed carbocation to give paracyclophane;
- The second possibility is the epoxidation of the C12-C13 double bond (Fig. 1.37b); Removal of a C1 proton leads to the rearrangement of the double bonds which ends with the attack of carbon C12 for the formation of the cyclohexane. The subsequent dehydration step leads to the same carbocation as mentioned in the first possibility;
- The third proposal is a radical mechanism (Fig. 1.37c), where cyclisations are concerted.

Finally, cyclohexene ring is formed via an intra molecular Diels Alder reaction (IMDA).

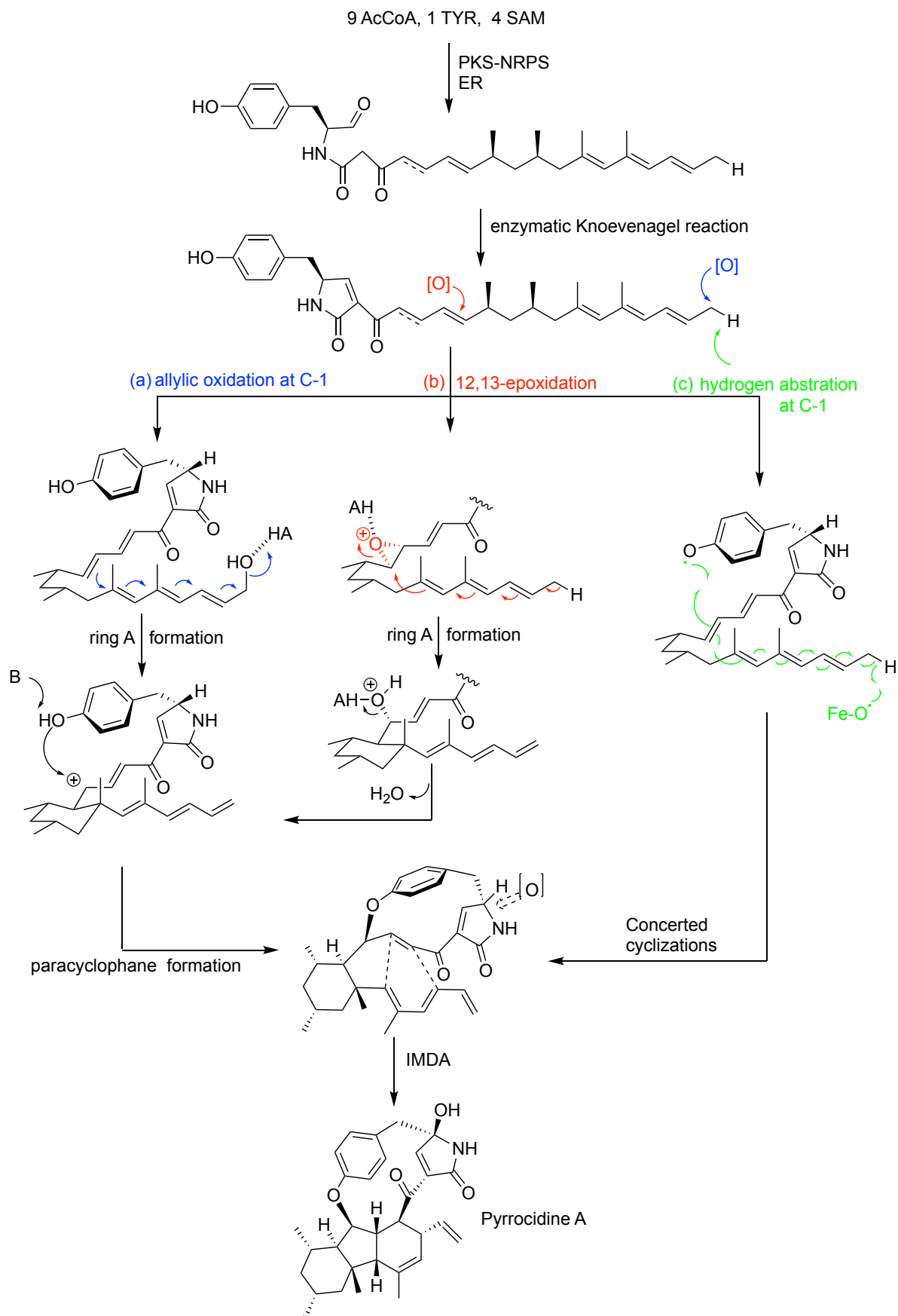


Fig. 1.37 The proposed biosynthetic pathway of pyrrocidines in the literature.

The cyclohexane moiety of pyrrocidines is likely formed by Diels-Alder reaction. Diels-Alderase catalyzed the formation of some cyclohexane group in natural products, such as Fsa2 (equisetin), SpnF (spinosyn aglycone), CghA (sch210972) (Fig.1.38). Based on that, it was proposed that a Diels-Alderase is involved in the biosynthesis of pyrrocidines.

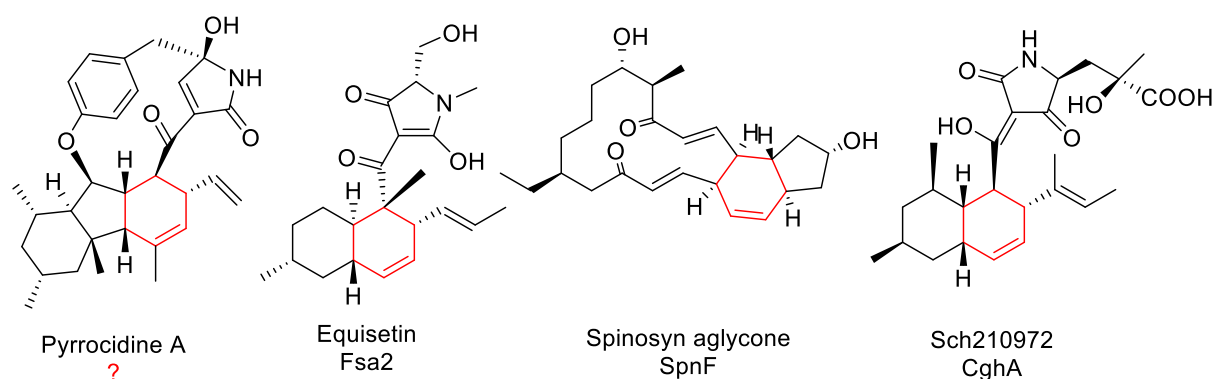


Fig. 1.38 Diels-Alder reaction catalyzed by enzymes in some natural products.

So, what is Diels-Alderase and how was Diels-Alderase discovered? Pericyclic reactions are often applied in organic synthesis to build carbon-carbon bonds efficiently. Diels-Alder reaction is the most famous pericyclic reaction owing to its versatility and remarkable stereo-selectivity.<sup>181</sup> Otto Paul Hermann Diels and Kurt Alder discovered the Diels-Alder reaction in 1928 and won the Nobel Prize in 1950. The Diels-Alder reaction is also called the 4 + 2 cycloaddition reaction, electron-rich conjugated dienes react with electron-deficient substituted olefins to form cyclohexene. There is no intermediate in the reaction process, but only a single peripheral transition state exists. The Diels-Alder reaction forms two stereoisomers though *endo* and *exo* approaches (Fig.1.39).<sup>182</sup> The organic synthesis products of the Diels-Alder reaction tend to *endo*-selectivity.

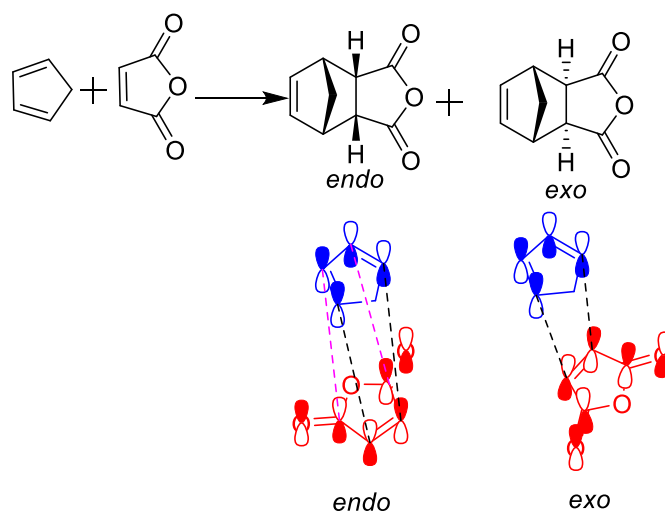


Fig. 1.39 Two approaches of Diels-Alder reaction.<sup>183</sup>

Since the Diels-Alder reaction occurs spontaneously, there has been controversy over whether there are enzymes that can catalyze the Diels-Alder reaction in nature for a long time. Some people hold the opinion that these enzymes are not necessarily authentic Diels-Alderase, but these enzymes can increase the [4 + 2] -cycloaddition reaction rate and strictly control the stereo configuration of the adduct.<sup>184</sup> The crude enzyme obtained from a cell-free extract of *Alternaria solani* can catalyze the Diels-Alder reaction from practically *endo* products to a majority of unfavourable *exo* products (Fig 1.40).<sup>185</sup> It confirmed the feasibility of the emergence of Diels-Alderase.

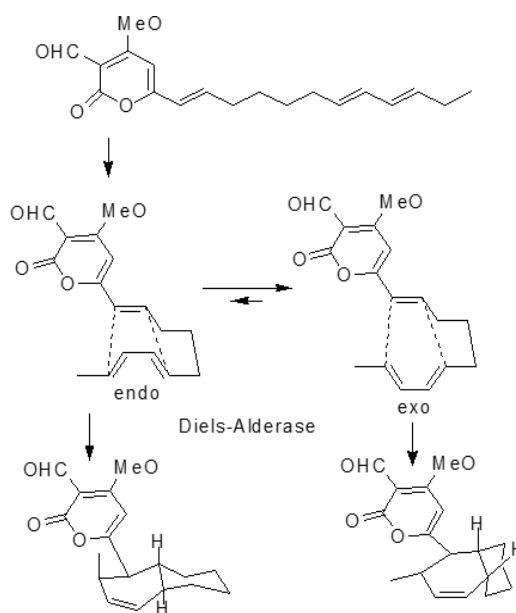


Fig. 1.40 Two different stereochemistry products originated from prosolanapyrone III.

LNKS is a polyketide synthase and is forecasted to catalyze the formation of decalin ring system in lovastatin by *in vitro* assay and heterologous expression.<sup>74,186</sup> MPS is a multifunctional enzyme and catalyzes oxidation, polyketide chain formation, decarboxylation, and Diels–Alder reaction. It was proposed to catalyze Diels-Alder reaction by the stereospecificity of macrophomate.<sup>187</sup> The crystal structure of MPS showed the reaction proceeds via a large-scale structural reorganization of the product.<sup>188</sup> These enzymes consistently demonstrate more than one catalytic activity, leaving their specific influence on the cycloaddition step uncertain. Until the advent of SpnF, it was a dedicated Diels-Alderase that accelerate the rate of Diels Alder reaction approximately 500 fold by *in vitro* assay(Fig. 1.41).<sup>189</sup> SpnF is recognized as the first proven Diels Alderase.

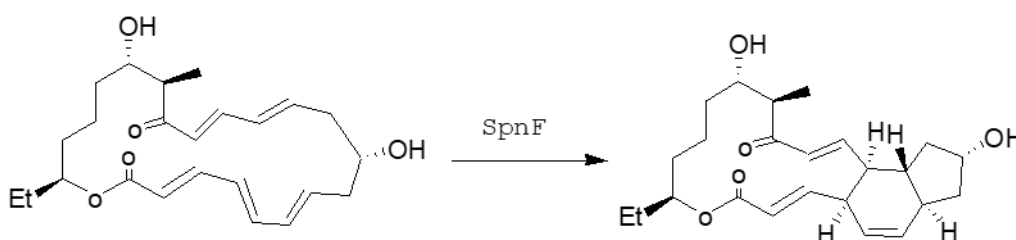


Fig. 1.41 SpnF catalyzed the Diels-Alder reaction.

The sequences of Diels-Alderase were unconservative, which resulted in other enzymes that could also catalyze the Diels-Alder reaction. O-methyltransferase LepI was a multifunctional enzyme (retro claisen rearrangement enzyme, dehydration) that was also identified to catalyze the hetero Diels-Alder reaction.<sup>190</sup> FAD-dependent Diels-Alderase MadA catalyzed the formation of chalconoracin has been characterized.<sup>191</sup> lipocalin-like protein Fsa2 functions as an *endo*-selective Diels-Alderase in the decalin formation, which was verified by gene inactivation.<sup>192</sup> In a similar way, CghA strictly controlled the stereoselectivity of decalin formation, which was identified by gene inactivation and complementation.<sup>193</sup>

The switch of *exo/endo* selectivity by changing the structure of the enzyme has been demonstrated. The crystal structure of CghA has been analyzed to explain the possible mechanism of Diels Alder reaction. The Site-directed mutation resulted in the product from *endo* stereoselectivity to major unfavoured *exo* stereoselectivity, which demonstrated a deep understanding of the Diels-Alderase.<sup>85</sup> Two homology Diels-Alderases MaDA and MaDA-3 utilized the same substrates to obtain stereoisomers with opposite *endo/exo* selectivity and high enantioselectivity. The stereoselectivity from *exo* to *endo* was achieved by point mutations after finding critical amino acids by protein crystallization.<sup>194</sup>

## 1.4 Objectives

This project focuses on the elucidation of the biosynthetic pathway of pyrrocidines from the endophytic fungus *Sarocladium zeae*, to understand the physiological roles of pyrrocidines, to characterize new analogs of pyrrocidines as well as to obtain other two PK-NRPs with bioactivities. The primary goal is to elucidate the biosynthetic pathway of pyrrocidines through gene inactivation and heterologous expression. Characterizing the novel analogs of pyrrocidines by performing genome mining and “OSMAC” strategy would lead to the discovery of potential new drugs. Understanding the physiological role of pyrrocidines by antagonistic experiments could potentially provide more possibilities for biocontrol agents. Since almost all PK-NRPs have good activities, our second target is to link the other two PKS-NRPS gene clusters to metabolites by gene inactivation. We also shed a light on activating the silent PKS-NRPS gene cluster by overexpression of the transcription factor which could lead to a structure diversity of PK-NRPs.





## **Chapter II: Biosynthesis of Pyrrocidines in *Sarocladium zae***

## 2.1 Pyrrocidines and analogues produced by *Sarocladium zeae*

### 2.1.1 Culture conditions for pyrrocidines production in *S. zeae*.

In a preliminary step, we tried to estimate the best culture conditions of the fungus *S. zeae* allowing enough production of pyrrocidines to perform LC-MS analysis and isolation experiments. The strain was cultivated in various media at 25°C and its growth aspect was examined over the time. We observed that *S. zeae* can survive for more than eight weeks on maize, while it seems dying after two weeks on Potatoe Dextrose Agar (PDA) medium, three weeks on rice medium and four weeks on oat or wheat media (Table. 2.1). After these times of culture, the media were extracted by ethyl acetate and analyzed by LC-ESI-HRMS. Extracted ion chromatograms (EIC) of  $[M+H]^+$  ions at  $m/z$  488.2795 ( $C_{31}H_{38}NO_4^+$ ) and  $m/z$  490.2952 ( $C_{31}H_{40}NO_4^+$ ) corresponding to pyrrocidines A and B allowed us to detect both metabolites at the retention times of 15.8 min and 15.9 min, respectively.

Table. 2.1 The yield and survival time of *S. zeae* on different media.

Medium	Mass of medium	Mass of extract	Time culture (weeks)
PDA	20 g	12 mg	2
Maize	50 g	2 g	8 or more
Rice	50 g	1.1 g	3
Oat	50 g	1.6 g	4
Wheat	50 g	2.4 g	4

By comparing the EIC, *S. zeae* seems to have the most prolific production of pyrrocidines on maize media (Fig. 2.1). This was confirmed by the large amount of pyrrocidine A precipitating during the concentration step of the ethyl acetate crude metabolic extract from maize (more than 100 mg of **1** is obtained from 50 g of medium) whereas this precipitation was not observed in the other tested media. NMR analysis of this precipitate confirmed the structure of pyrrocidine A. The NMR data are reported in Table 2.2 and are in agreement with the data published in the literature.<sup>195</sup>

It is worth noting that in the case of PDA medium, despite the low yield of metabolic extract (12 mg /20 g of PDA), a sufficient amount of pyrrocidines is accumulating. Hence, PDA seems to be a convenient medium to rapidly assess the effect of gene inactivation in the biosynthetic pathway of pyrrocidines by LC-MS analysis, while a long culture on maize medium is suitable for scale up fermentation to isolate and characterize pyrrocidines and related compounds or accumulating metabolites resulting from gene inactivation.

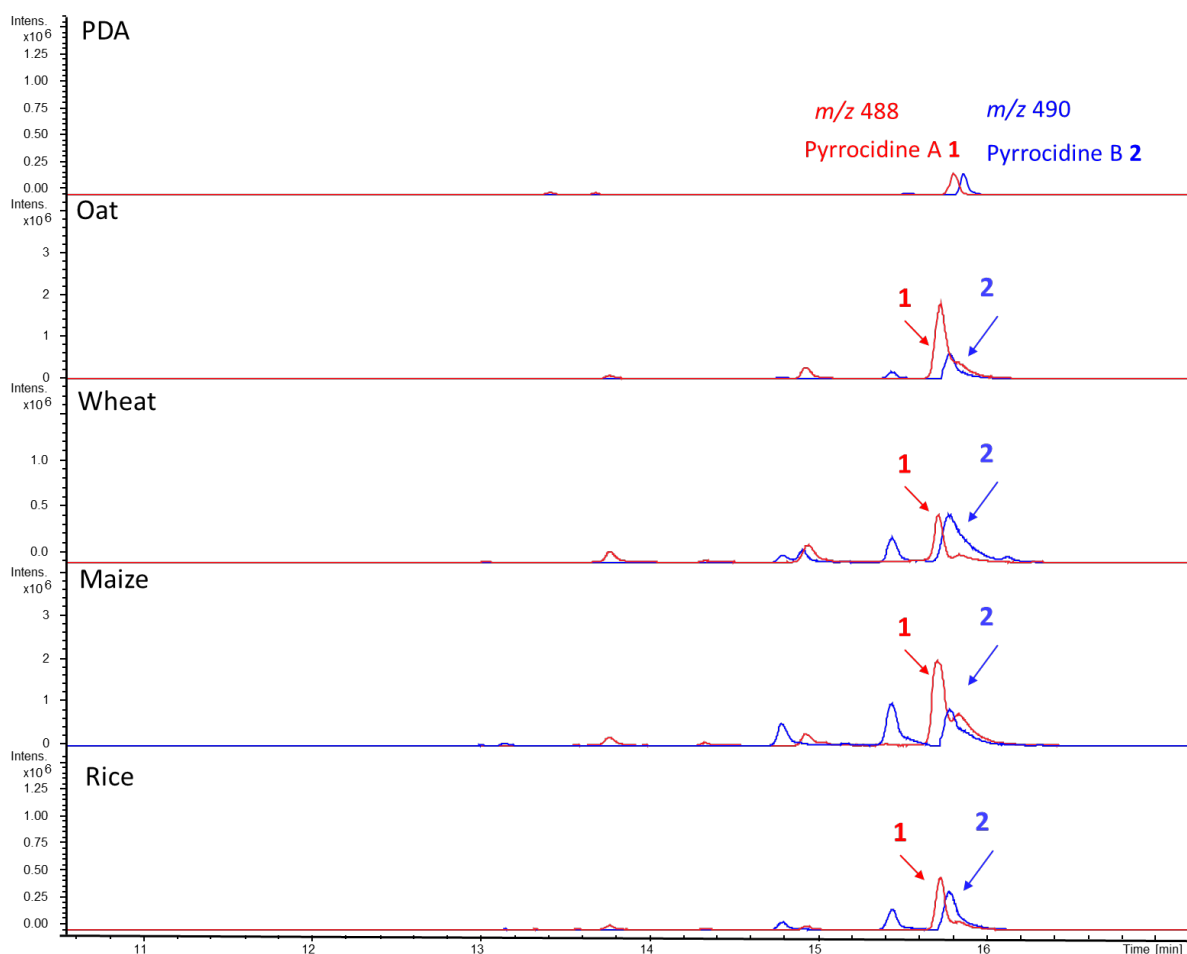


Fig. 2.1 LC-MS analysis of crude metabolite extracts of *S. zeae* cultured on various media. EIC of  $m/z$  488.2795 (red) and  $m/z$  490.2950 (blue) corresponding to the  $[M+H]^+$  molecular ions of pyrrocidines A and B, respectively.

### 2.1.2 Characterization of pyrrocidines and analogs in wild type *S. zeae*.

From the culture conditions established above, we investigated the diversity of pyrrocidines and analogues produced in WT *S. zeae*. In a first step, the crude extract was fractionated by flash chromatography on silica gel using a cyclohexane/ethyl acetate gradient. After analysis by LC-MS and LC-MS/MS of all fractions, a further purification of selected metabolites was performed on preparative HPLC with a water/acetonitrile gradient. The purified products were then characterized by NMR experiments. For each reported molecule  $^1\text{H}$  and  $^{13}\text{C}$  spectral data were recorded and detailed 2D analysis was performed with COSY, HSQC, HMBC and NOESY experiments. The purified compounds were also analyzed by high resolution mass spectrometry (LC-HRMS) and their rotatory powers measured (Annex 05 and 06).

Pyrrocidine B **2** was thus isolated and the NMR characterization (Table. 2.3) as well as  $[\alpha]_D$  (Annex 06) are in agreement with published data.<sup>195</sup> Careful analysis of the  $m/z$  490.2952 EIC from PDA medium shows two other peaks with retention time at 15.5 min and 14.9 min. The respective assigned compounds **7a** and **7b** have similar LC-MS<sup>2</sup> fragmentation patterns as pyrrocidine B (Fig. 2.2), suggesting stereoisomers of **2**.

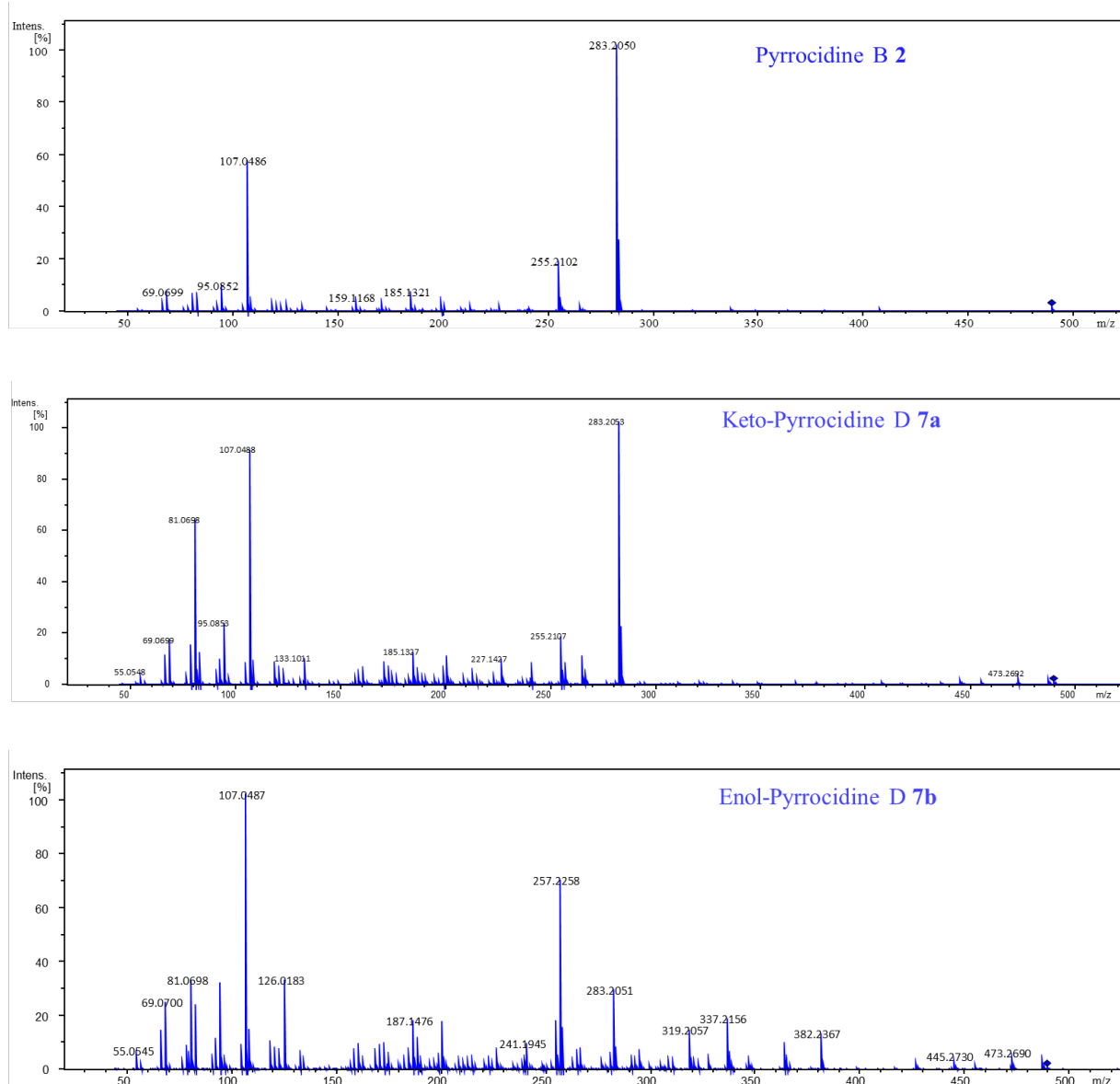


Fig. 2.2 The LC-MS<sup>2</sup> spectra of pyrrocidine B **2**, **7a** and **7b**.

Interestingly, the culture of the fungus on maize medium for 30 days leads to the accumulation of these compounds compared to **2** (Fig. 2.3).

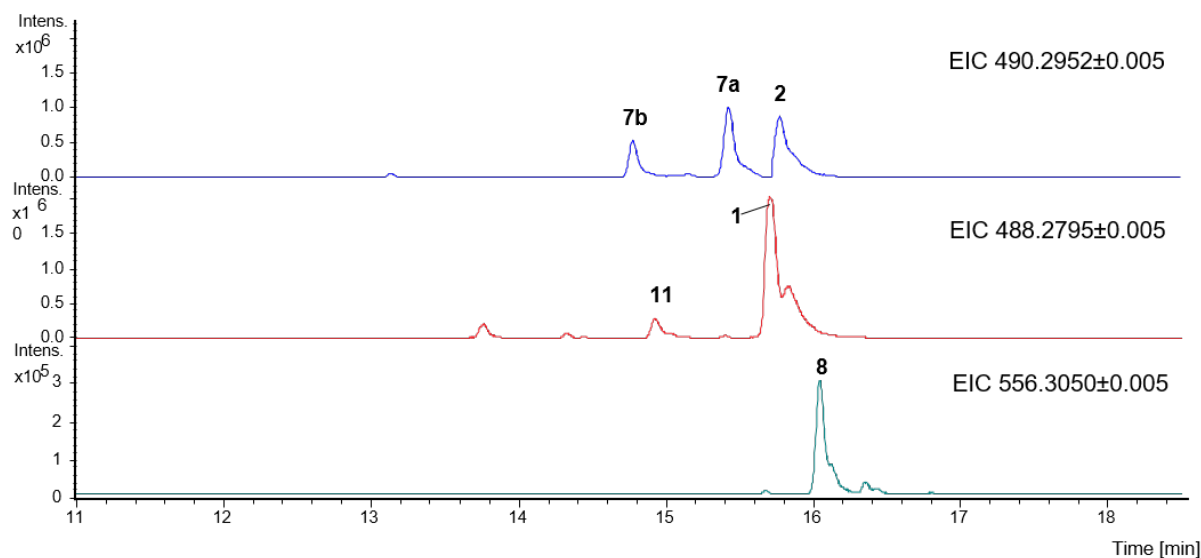


Fig. 2.3 Extracted ion chromatograms obtained from LC-MS analysis of the metabolite extract from wild type *S. zeae* cultured on maize medium for 30 days.

Compound **7a** was purified and we established its planar structure from 2D NMR experiments in  $\text{CDCl}_3$  and by comparison with NMR data of **2** (Table 2.4 and 2.5). The structure of the tetramethylated decahydrofluorene core was established by HMBC correlations from H14 to C4, C5, C6, C13, C15, C16, from H15 to C3, C13, C14, C16, from H10b to C9, C11, C30, C31, from H6 to C4, C5, C7, C14, C29, and COSY cross correlations from H1 to H12. Connection to the pyrrolidone moiety was highlighted by HMBC correlations from H19a/b to C16, C17, C18, C20, C21. The presence of the locked paracyclophane was deduced from the inequivalent  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts of aromatic positions 27/23 and 24/23 and the HMBC correlations from C25 to H13. The relative stereochemistry was elucidated from NOESY spectrum with the key cross-correlations H3-H6-H12-H8b, H15-H13-H31-H24 and H29-H14-H26, as well as  $^3J$  coupling values from H11 to H3, defining **7a** as the diastereomer of **2** at C3 and C6 position. This diastereomery is identifiable by  $^{13}\text{C}$  NMR with notable shielding of C7, C28 and C29 chemical shifts of ca. 6 ppm in compound **7a** with respect to **2**. These downfield shifts are thus a signature of a *trans*-fused junction of rings B and C in pyrrocidines.

Compound **7b** turned out to be unstable in  $\text{CDCl}_3$  but, when analyzed by LC-MS pure ketone **7a** in acetonitrile showed partial conversion to **7b** that we interpreted as the enol form of **7a**. In turn, purified enol **7b** interconverted with **7a**. This was also observed by NMR when ketone **7a** is solubilized in  $\text{DMSO-d}_6$  which equilibrates with its enol form in a respective 2:1 ratio (Fig. 2.4). Enol **7b** was identified by the key changings of chemical shifts at C16, C17 and C19.

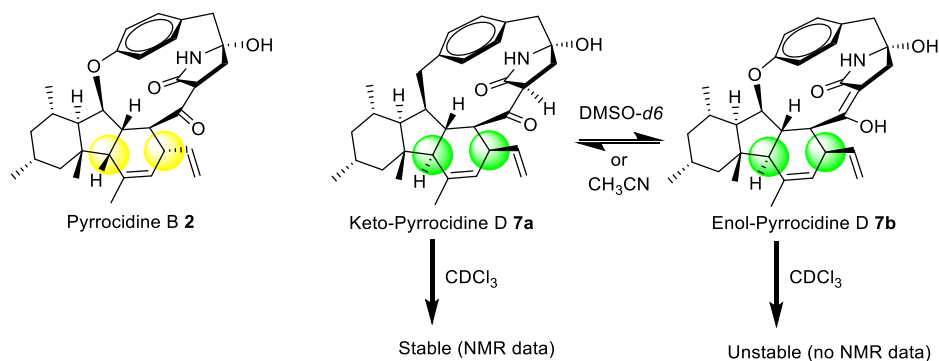


Fig. 2.4 Structures of pyrrocidine B, D, and keto-enol equilibrium of pyrrocidine D.

Besides these metabolites, we detected the compound **8** with  $[M+H]^+$  ion at  $m/z$  556.3057 (retention time of 16.1 min) (Fig. 2.3) corresponding to the molecular formula  $C_{35}H_{41}NO_5$  identical to the anti-glioma trichobamide A.<sup>172</sup> Structure elucidation by NMR analysis shows that both compounds are isomers differentiating in the substitution of the tetrahydro furo-pyrrolone moiety at C33 and C34 (Fig. 2.5). Relative stereochemistries were determined from NOESY spectrum. Indeed, the cross-correlations of H11-H29-H6-H14-H26 and H10b-H12-H8b-H13-H24-H15-H2 show that H12, H13, H15 and the vinyl group are cofacial by respect to the equatorial plane of decahydrofluorene and H6, H14, H29, H3 are positioned on the opposite face (Table 2.7).

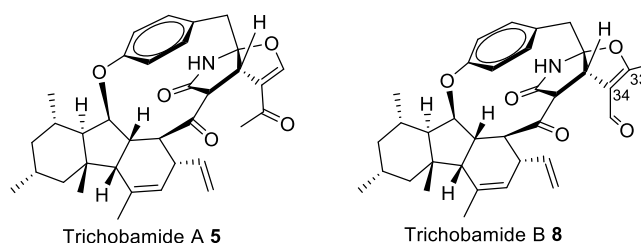


Fig. 2.5 The structure of trichobamides A and B.

Additionally, the presence of the compound **11** with  $[M+H]^+$  of  $m/z$  488.276 was detected at 15.0 min (Fig. 2.3). NMR spectra of purified **11** are identical to those reported by Shiono *et al.* for epi-pyrrocidine isolated from the insect derived fungus *Neonectria ramulariae* In-2.<sup>176</sup> In this metabolite, the pyrrolidone ring belonging to the strained paracyclophane is in reverse orientation by respect to the one in pyrrocidine A. For convenience, it was named here pyrrocidine F and it represents the C20-epimer of pyrrocidine A.

From the metabolite extract of *S. zeae*, we also isolated a compound with  $[M+H]^+$  of  $m/z$  504.2744 that we assign to pyrrocidine C **9** (Fig. 2.6). Despite this compound was already mentioned in several studies<sup>13,180</sup> describing its spontaneous formation from pyrrocidine A, to the best of our knowledge, NMR characterization was not published and we provide here its  $^1\text{H}$  and  $^{13}\text{C}$  spectral data (Table 2.8).

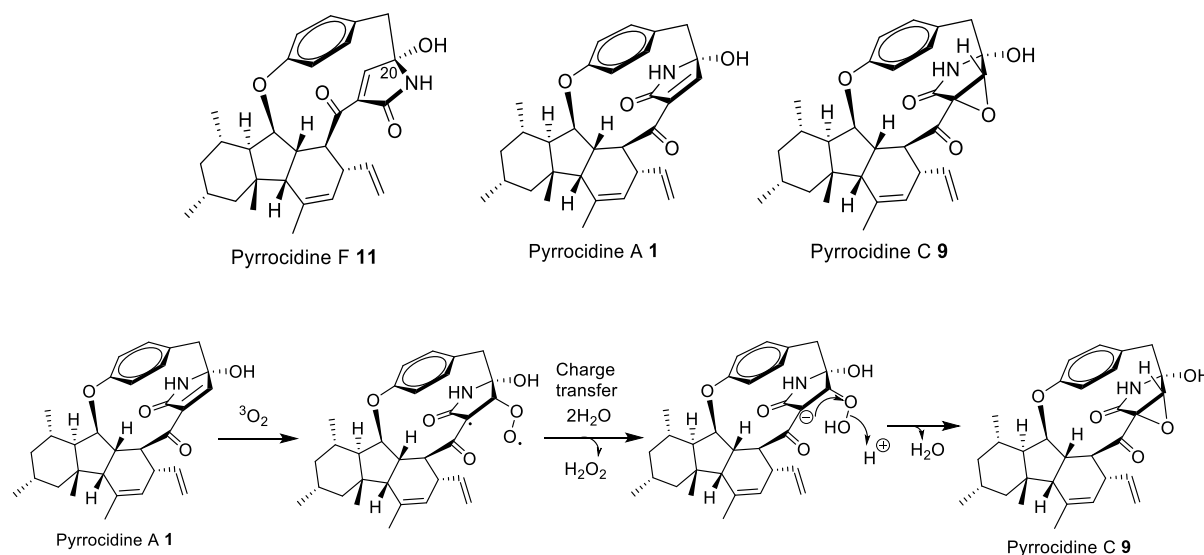
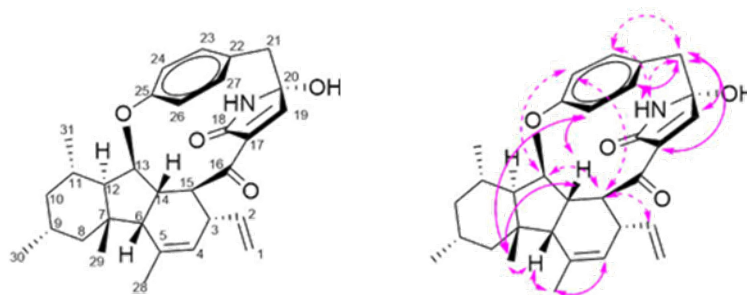


Fig. 2.6 The structure of pyrrocidines F, A and C and possible mechanism for the spontaneous formation of pyrrocidine C **9**.

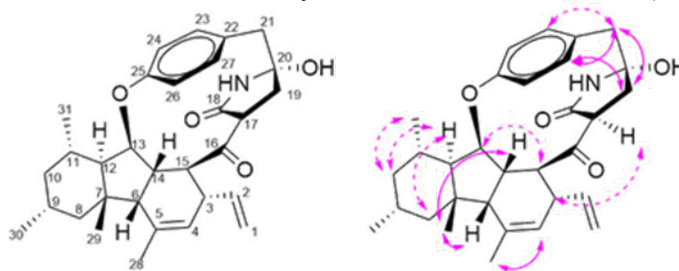
This chemical study of wild type *S. zeae* allowed us to identify three new members in the pyrrocidine family, namely trichobamide B, keto-pyrrocidine D and enol-pyrrocidine D and also the production of pyrrocidine F in this fungus. Since we set up the conditions of culture for pyrrocidines production, we initiated the deciphering of their biosynthetic pathway.



Table 2.2: NMR data of Pyrrocidine A **1** in CDCl<sub>3</sub> (400 Hz).

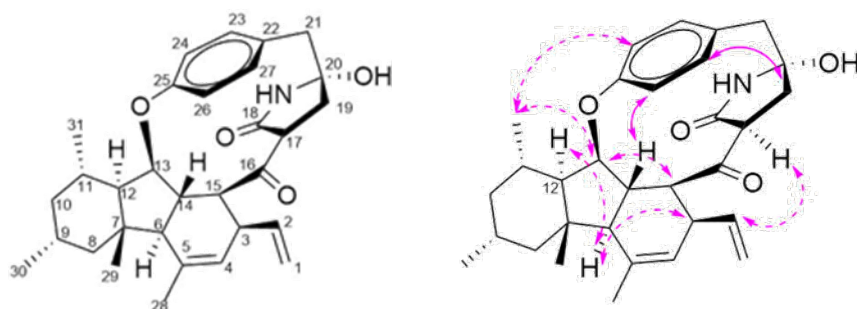
Key NOESY correlations

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. (J in Hz)	HMBC	COSY	NOESY
1a	115.8	4.93	d (17.1)		1b; 2; 12	
1b		4.79	d (10.1)		1a; 2; 3; 12	
2	140.7	5.89	ddd (17.0; 10.2; 4.3)	1a; 15	1a; 1b; 3	15
3	38.7	2.95	m	1a; 1b; 4; 14; 15	1b; 2; 4	
4	121.9	5.62	dd (6.4)	6; 15; 28	3; 6; 28	28
5	138.1	-	-	6; 14; 28	-	
6	52.2	2.55	d (9.2)	4; 15; 28; 29	4; 14	28; 29
7	47.7	-	-	6; 12; 13; 29	-	
8a	48.4	1.77	m	30	8b	
8b		0.94	m		8a	
9	28.4	1.73	m	30	30	
10a	44.8	1.72	m	30; 31	10b	
10b		0.47	ddd (11.9)		9; 10a; 11	
11	27.2	1.81	m	12; 31	10b; 12; 31	
12	53.9	1.21	m	6; 29; 31	1a; 1b; 11; 13	
13	92.1	4.34	d (5.4)	6; 14; 15	12; 14	15; 24
14	42.6	2.34	d (9.2)	15	6; 13	26; 29
15	51.5	2.92	s	13; 14	-	2; 13; 24
16	202.9	-	-	14; 15; 19	-	
17	139.8	-	-	19, NH	-	21b
18	172.0	-	-	19	-	
19	151.9	7.55	m	NH	NH	21b
20	88.8	-	-	19; 21a; 21b	-	
21a	43.9	3.19	d (11.7)	23; 27	21b	23; NH
21b		2.65	d (11.7)		21a	17; 19; 27
22	129.1	-	-	24; 26	-	
23	130.2	7.48	d (8.2)	27	24	21a; NH
24	122.2	7.18	dd (8.2; 1.7)	26	23	13; 15
25	157.9	-	-	23; 27; 24; 26	-	
26	125.4	6.75	dd (8.2; 1.7)	24	27	14; 29
27	132.2	7.15	d (8.2)	23	26	21b
28	25.3	1.81	s	4	4	4; 6
29	24.0	1.25	s	6; 12	-	6; 14; 26
30	23.1	0.90	d (5.8)		9	
31	19.7	1.01	d (6.0)		11	
NH		8.62	s		19	21a; 23
OH		6.56	broad			

Table 2.3 : NMR data of Pyrrocidine B **2** in CDCl<sub>3</sub> (600 Hz).

## Key NOESY correlations

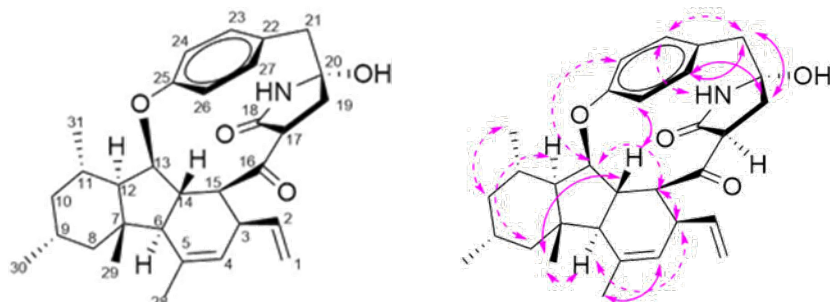
position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. (J in Hz)	HMBC	COSY	NOESY
1a	115.3	5.06	d (17.2)		2	
1b		4.98	d (10.1)			
2	141.0	5.01	ddd (17.2; 10.1; 4.4)	1a; 15, 28	1; 3	
3	37.5	3.16	m	1a; 1b; 4; 14; 15	2; 4; 28	17
4	121.4	5.27	dd (6.7; 1.1)	2; 6; 15; 28	3; 6; 28	28
5	137.4	-	-	6; 14; 28	-	
6	53.4	2.95	d (8.9)	4; 14; 15; 28; 29	4; 14; 28	29
7	47.6	-	-	6; 12; 13	-	
8a	48.1	1.70	dd (11.7; 2.1)	30	8b	
8b		0.92	dd (11.7; 11.7)		8a; 29	12
9	28.4	1.78	m	8b; 10b	28	
10a	44.9	1.72	m (eq.)	30; 31	10b	12
10b		0.46	ddd (12.2) (ax.)		10a	
11	27.3	1.86	m	10a; 12	14; 31	
12	53.1	1.07	dd (11.3; 6.8)	6	-	8b; 10b
13	92.2	4.26	dd (6.7; 2.9)	6; 14; 15	14	15
14	44.2	2.33	d (8.8)	6; 15	6; 11; 13	24; 29
15	51.2	1.87	m	4; 13; 14	-	13
16	208.8	-	-	15; 17; 19b	-	
17	53.4	3.29	dd (12.3; 6.5)	19b	19a; 19b	3
18	174.3	-	-	17; 19a	-	
19a	37.2	2.40	dd (15.2; 2.7)	17	17; 19b	
19b		2.12	dd (15.2; 6.3)		17; 19a	21; 27
20	89.0	-	-	19a, 19b; 21a; 21b	-	
21a	45.6	3.13	d (13.7)	19b; 21a; 23; 27	21b	23
21b		2.94	d (13.7)		21a	27
22	130.8	-	-	21a; 21b; 24; 27	-	
23	132.8	7.25	dd (8.4; 2.0)	21a; 21b; 27	27	
24	125.2	6.96	dd (8.1; 2.3)	26	26	14
25	158.8	-	-	23; 27; 24; 26	-	
26	122.8	7.16	dd (8.4; 2.2)	24	24	
27	131.7	7.16	dd (8.4; 2.2)	21a; 21b, 23	23	
28	25.1	1.76	s	6	3; 4; 6; 9	4
29	23.3	1.20	s	8b	6	6
30	23.1	0.88	d (6.4)	10b		
31	20.0	1.01	d (6.2)	10b	11	10b
NH		7.29	S broad			
OH		4.10	broad			

Table 2.4: NMR data of *keto*-Pyrrocidine D **7a** in CDCl<sub>3</sub> (600 MHz).

Key NOESY correlations

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1a	116.3	5.07	d (17.5)		2	
1b		5.00	d (10.5)			
2	138.3	5.78	ddd (17.5; 10.5; 6.7)	1a	1a; 1b; 3	17
3	43.0	2.85	m	1a; 1b; 2; 15	2; 4; 15; 28	6
4	121.6	5.20	m	2; 6; 14; 28	3; 6; 28	
5	141.6	-	-	6; 14; 28	-	
6	56.9	1.51	d (12.2)	14; 29	4; 14	3; 12
7	41.3	-	-	6; 10; 12; 13; 29	-	
8a	49.1	1.92	m	12; 29; 30	8b	
8b		0.78	dd (12.3; 12.3)		8a; 29	12
9	28.0	1.83	m	8a; 8b; 10b; 30	30	
10a	45.5	1.80	m	11; 30; 31	10b	
10b		0.59	ddd (12.5 x3)		10a	
11	27.1	1.98	m	10b; 12; 31	31	
12	61.1	1.04	dd (11.3; 7.4)	11; 29; 31	13	6
13	89.9	4.33	dd (7.1; 4.3)	14; 15	14; 12	15; 31
14	50.3	2.40	ddd (12.3; 7.1; 4.2)	6; 15	6; 13; 15	26
15	50.6	3.66	dd (7.1)	13; 14	13; 14	13
16	201.9	-	-	14; 15; 17; 19a; 19b	-	
17	56.5	3.10	dd (12.3; 4.1)	19a; 19b; 28	19a; 19b	2
18	172.7	-	-	17; 19a; 19b	-	
19a	34.7	2.87	m	17; 21a; 21b	17	27
19b		2.00	m		17	
20	88.8	-	-	19a; 21a; 21b	-	
21a	46.8	2.97	d (13.0)	19b; 27	21b	
21b		2.91	d (13.0)		21a	
22	128.8	-	-	21a; 21b; 24; 26	-	
23	136.6	7.03	d (8.5)	21a; 21b; 27	24	
24	120.6	7.00	dd (8.5; 2.1)	26	23	31
25	158.2	-	-	13; 26; 27	-	
26	124.4	6.75	dd (8.2; 2.1)	24	27	14
27	131.8	7.08	dd (8.2; 2.1)	21a; 21b, 23	26	
28	19.9	1.94	s		3; 4	
29	16.9	1.25	s	6; 8b; 12	8b	
30	22.9	0.90	d (6.2)	8b; 10b	9	
31	19.7	1.09	d (6.2)	10b	11	13; 24
NH		6.55	s (broad)			
OH		3.55	s (broad)			

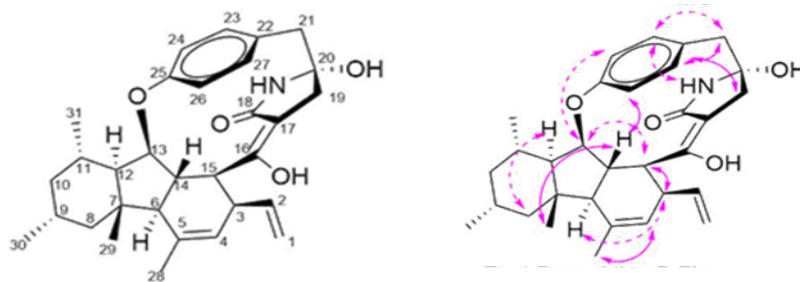
Table. 2.5: NMR data of *keto*-Pyrrocidine D form **7a** in DMSO (600 Hz).



Key NOESY correlations

positio	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. (J in Hz)	HMBC	COSY	NOESY
1a	115.1	5.01	d (17.3)		2	
1b		5.00	d (10.7)		2	
2	138.5	5.83	ddd (17.3; 10.7; 6.5)	1a	1a, 1b; 3	
3	42.0	2.83	m	1a; 1b; 2; 15	2; 15	6; 15
4	121.6	5.18	s	2; 15; 28	6	
5	139.7	-	-	6; 14; 28	-	
6	55.5	1.65	d (12.4)	8b; 14; 28; 29	4; 8a; 14; 28	29
7	40.5	-	-	6; 8b; 13; 29	-	
8a	48.1	1.86	m	29; 30	6; 8b	
8b		0.80	dd (12.0; 12.0)		8a	12
9	27.4	1.79	m	8b; 10b; 30	30	
10a	44.9	1.76	m	30; 31	10b	
10b		0.59	ddd (12.1 x3)		10a	
11	26.6	1.84	m	10b; 29; 31	12; 31	
12	59.7	1.13	m	11; 29; 31	11; 13	
13	88.7	4.26	dd (7.2; 4.3)	14; 15	12; 14	15; 24
14	49.35	2.28	ddd (12.1; 7.5; 4.3)	15	6; 13; 15	26; 29
15	49.3	3.68	dd (7.5; 7.5)	13; 14	3; 14	3; 13
16	202.4	-	-	3; 14; 16; 19a; 19b	-	
17	55.9	2.82	m	19a; NH	19a; 19b	
18	172.2	-	-	19b	-	
19a	33.7	2.45	dd (14.4; 4.3)	17	17; 19b	
19b		1.81	dd (14.4; 12.5)		17; 19a	21; 27
20	87.4	-	-	19a, 19b; 21a; 21b	-	
21a	45.9	2.79	d (12.6)	19b; 21a; 23; 27	21b	23
21b		2.63	d (16.6)		21a	27
22	130.0	-	-	21a; 21b; 24; 27	-	
23	133.7	7.11	dd (8.6; 2.0)	21a; 21b; 27	27	H21a; NH
24	119.4	6.85	dd (8.6; 2.4)	26	26	14
25	156.7	-	-	23; 27; 24; 26	-	
26	122.8	6.52	dd (8.1; 2.4)	24	24	
27	131.4	7.00	dd (8.1; 2.0)	21a; 21b, 23	23	H21b; H19a
28	19.5	1.87	s	6	3; 4; 6; 9	4
29	16.5	1.16	s	8b	6	6
30	22.7	0.88	d (6.2)	10b		
31	19.5	1.07	d (6.5)	10b	11	10b
NH		8.25				
OH		5.90				

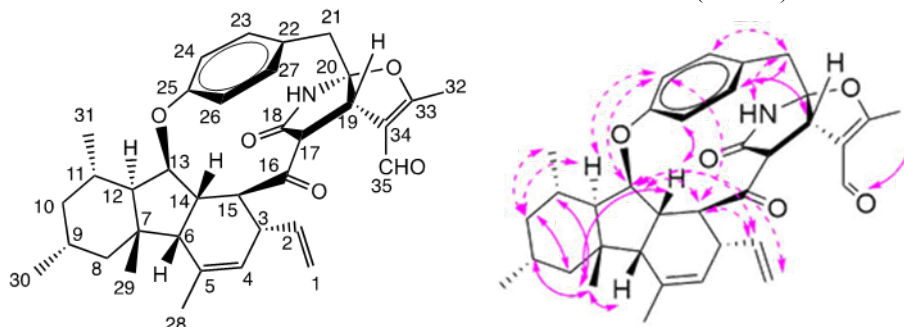
Table 2.6: NMR data of Pyrrocidine D-enol form **7b** in DMSO (600 Hz).



Key NOESY correlations

positio	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. (J in Hz)	HMBC	COSY	NOESY
1a	113.5	4.89	d (17.4)		2	
1b		4.77	d (10.5)		2	
2	138.9	5.50	ddd (17.4; 10.5; 6.8)	1a	1a, 1b; 3	
3	41.9	2.63	m		2; 15	6; 15
4	125.9	5.30	s	2; 15; 28	6	28
5	136.4	-	-	6	-	
6	55.6	1.62	d (12.2)	8b; 28; 29	4; 8a; 14	3
7	40.7	-	-	6; 8b; 13; 29	-	
8a	48.2	1.86	m	29; 30	6; 8b	
8b		0.81	dd (12.0; 12.0)		8a	12
9	27.4	1.79	m	8b; 10b; 30		
10a	44.9	1.76	m	30; 31	10b	
10b		0.59	ddd (12.0 x3)		10a	
11	26.6	1.84	m	10b; 29; 31		
12	58.9	1.13	m	11; 29; 31	13	
13	89.4	4.26	dd (7.5; 4.3)	14; 15	12; 14	15; 24
14	49.4	2.08	ddd (12.2; 7.5; 4.3)	15	6; 13; 15	26; 29
15	40.6	3.68	dd (7.5; 7.5)		3; 14	3; 13
16	157.7	-	-	14; 15; 19a; 19b	-	
17	103.5	-	-	19a; 19b; NH		
18	171.1	-	-	19a; NH	-	
19a	39.0	2.47	d (14.9)	21a; NH	19b	27
19b		2.05	d (14.9)		19a	
20	86.2	-	-	19a, 19b; 21a; 21b; NH; OH	-	
21a	46.2	2.72	d (12.6)	23; 27; OH	21b	
21b		2.63	d (16.6)		21a	27
22	131.0	-	-	21a; 21b; 26; 27	-	
23	132.3	7.06	dd (8.4; 2.0)	21a; 21b; 27	24	21a; NH
24	119.6	6.77	dd (8.4; 2.3)	26	23	13
25	156.9	-	-	23; 27; 24; 26	-	
26	121.4	6.56	dd (8.1; 2.3)	24	27	14
27	130.2	7.00	dd (8.1; 2.0)	21a; 21b, 23	26	19a; 21b
28	19.5	1.87	s			4
29	16.4	1.13	s	6; 8b; 12		14
30	22.7	0.88	d (6.2)			
31	19.5	1.09	d (6.4)	10b; 12		
NH		7.44				23
OH		5.72				
OHeno		7.45				

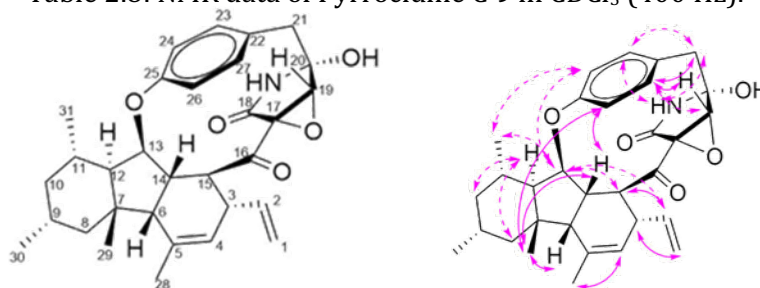
Table 2.7: NMR data of Trichobamide B **8** in CDCl<sub>3</sub> (600 Hz).



Key NOESY correlations

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1a	115.1	5.03	d (17.6)	H3	2	13
1b		4.96	d (10.1)			
2	141.2	5.87	ddd (14.7; 10.1; 4.6)	1a; 3	1a; 1b; 3	13; 15
3	37.2	3.22	m	1a; 1b; 2; 4; 13; 14; 15	2; 4	
4	121.7	5.22	d (6.7)	2; 28	3; 28	
5	136.9	-	-	6; 14; 28	-	
6	53.2	2.55	s	3; 4; 13; 14; 28; 29	-	29
7	47.7	-	-	12; 13; 29	-	
8a	48.1	1.70	m	30	8b	
8b		0.92	dd (12.0; 12.0)		8a	10b
9	28.4	1.78	m	10b; 30	30	29
10a	45.0	1.72	m	30; 31	10b	
10b		0.46	ddd (12.0)		10a	8b; 12; 31
11	27.3	1.87	m	10b; 31	31	29
12	53.5	1.06	dd (11.3; 6.7)	31	13	10b; 24
13	93.5	4.14	dd (6.6)	6; 12; 14; 15	12; 14	1a; 2; 15; 24
14	43.4	2.55	s	3; 6; 15	13	26; 29
15	52.3	1.63	s	4; 13	17	2; 13; 24
16	205.0	-	-	14; 15; 17; 19	-	
17	59.1	3.10	dd (10.0; 3.0)	19; 32; NH	19; 15	
18	173.6	-	-	17; 19; NH	-	
19	48.6	3.61	m	17; 21a; 21b; 32; 35; NH	17; 32	27
20	103.6	-	-	21a; 21b; NH	-	
21a	42.5	3.36	d (13.7)	19; 23; 27	21b	23; NH
21b		3.01	d (13.7)		21a	27
22	128.1	-	-	21a; 21b; 24; 27	-	
23	132.7	7.23	dd (8.3; 2.3)	21a; 21b; 27	24	
24	123.5	7.15	dd (8.3; 2.3)	26	23	
25	159.3	-	-	13; 23; 27; 26	-	
26	126.3	6.95	d (8.3; 2.3)	24	27	
27	131.0	7.13	d (8.3; 2.3)	21a; 21b; 23	26	19
28	25.1	1.73	s	6	4; 6	
29	23.4	1.21	s	8b; 12	-	9; 11; 14
30	23.1	0.89	d (6.4)	10b	9	
31	20.0	1.02	d (6.3)	10b; 12	11	10b
32	13.0	2.23	d (1.2)		19	35
33	117.1	-	-	17; 19; 32; 35	-	
34	171.5	-	-	19; 32	-	
35	184.3	9.73	s	-	-	32
NH		6.92	broad			

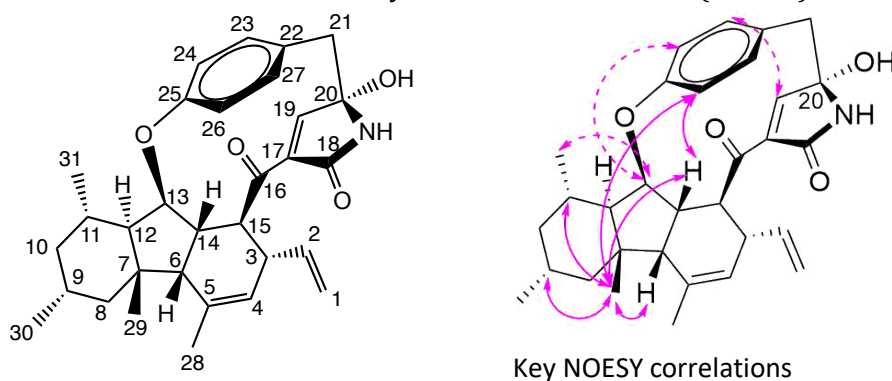
Table 2.8: NMR data of Pyrrocidine C 9 in CDCl<sub>3</sub> (400 Hz).



Key NOESY correlations

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1a	115.8	5.12	d (17.2)	3	2	3
1b		5.03	d (10.0)			
2	140.9	5.89	ddd (17.2; 10.0; 4.7)	1a; 3; 15; 28	1a; 1b; 3	13; 15
3	36.9	3.10	m	1a; 1b; 2; 4; 14; 15	2; 3; 14; 28	1a
4	121.8	5.37	d (6.6)	2; 6; 15; 28	3; 6; 28	28
5	137.3	-	-	6; 14; 28	-	
6	52.9	2.35	d (9.0)	4; 14; 15; 28; 29	4; 14; 28	29
7	47.8	-	-	6; 12; 8b 29	-	
8a	48.1	1.71	m	13; 30	8b	
8b		0.95	dd (11.9; 11.9)		8a	12
9	28.3	1.77	m	8b; 10b; 30	30	
10a	44.9	1.73	m	8a; 12; 30; 31	10b	
10b		0.48	ddd (12.0)		10a; 11	12
11	27.3	1.86	m	12; 31	10b; 12; 31	
12	53.2	1.14	dd (11.3; 6.8)	6; 31	11; 13	8b; 10b
13	91.4	4.39	dd (6.7; 2.8)	6; 14; 15	12; 14	2; 15; 24; 31
14	43.7	2.22	d (8.6)	6; 15	3; 6; 13	26; 29
15	50.2	2.37	s	4; 13; 14	-	2; 13
16	200.7	-	-	3; 14; 15	-	
17	63.3	-	-	15, NH	-	
18	167.9	-	-	19; NH	-	
19	65.2	3.60	d (2.4)	21a; 21b; NH		27
20	85.4	-	-	19; 21a; 21b; NH	-	
21a	43.2	3.25	d (13.9)	23; 27	-	23
21b		3.18	d (13.9)			27
22	128.3	-	-	21a; 21b; 24; 26	-	
23	131.8	7.28	dd (8.5; 2.1)	27	24	21a; NH
24	123.1	7.20	dd (8.3; 2.2)	26	23	13; 31
25	158.6	-	-	23; 27; 26	-	
26	125.3	6.93	dd (8.2; 2.1)	24	27	14; 29
27	131.3	7.19	dd (8.2; 2.1)	23	26	19; 21b
28	25.2	1.79	s	4; 6	3; 4; 6	4
29	23.3	1.18	s	6; 8b; 12	-	6; 14; 26
30	23.1	0.89	d (6.2)	10b	9	
31	20.0	1.01	d (6.2)	10b; 12	11	13; 24
NH		6.55	d (2.0)		19	23
OH		4.09	broad			

Table 2.9: NMR data of Pyrrocidine F **11** in CDCl<sub>3</sub> (400 Hz).



position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1a	116.0	5.09	ddd (17.1; 1.8; 1.8)		2	
1b		5.03	ddd (10.0; 1.8; 1.8)			
2	140.5	6.78	ddd (15.4; 10.0; 5.3)	15	1a; 1b	
3	38.0	2.87	m	4; 14; 15	4	
4	123.8	5.56	d (6.4)	6; 15; 28	3; 28	
5	134.4	-	-	6; 28	-	
6	52.2	2.13	d (9.0)	4; 28; 29	14	29
7	47.3	-	-	6; 13; 29	-	
8a	48.5	1.63	d (12.3)	30	-	
8b		0.88	m		-	
9	28.4	1.70	m	10b; 29; 30	-	29
10a	44.7	1.69	m	30; 31	-	
10b		0.47	ddd (12.3)		-	
11	27.3	1.87	m	12; 31	-	29
12	53.2	1.21	dd (5.6)	6; 29; 31	13	
13	90.8	4.51	dd (6.1; 2.0)	6; 14; 15	12; 14	24; 31
14	40.9	2.30	d (7.4)	15	6; 13	26; 29
15	55.0	2.40	s	4; 13; 14	-	
16	199.3	-	-	14; 15; 19	-	
17	135.8	-	-	19	-	
18	168.4	-	-	19	-	
19	148.8	6.43	d (1.4)	21	-	23
20	89.1	-	-	19; 21	-	
21a	44.0	3.26	d (13.1)	23	-	} 23/27
21b		3.30	d (13.1)		-	
22	126.6	-	-	24; 26	-	
23	132.7	7.19	dd (8.5; 2.2)	21; 27	24	19; 21
24	120.9	7.06	dd (8.4; 2.4)	26	23	13
25	158.3	-	-	23; 27	-	
26	124.4	6.94	dd (8.4; 2.4)	24	27	14; 29
27	129.4	7.10	dd (8.4; 2.2)	23	26	21
28	25.1	1.73	s	4	-	
29	23.8	1.13	s	8b; 12	-	6; 9; 11; 14; 26
30	23.1	0.87	d (6.2)		-	
31	19.9	0.99	d (6.2)		-	13
NH		7.01	broad			
OH			Not visible			



## 2.2 BGCs involving PKS-NRPS in the genome of *S. zeae*

In order to identify the biosynthesis gene cluster of pyrrocidines, the genome of *S. zeae* was sequenced and assembled by our team. Among the resulting 230 DNA scaffolds covering 31.2 Mb genome, three genes coding for a putative hybrid PKS-NRPS were identified in the genome of *S. zeae* by BLAST with the hybrid PKS-NRPS TenS as reference.<sup>196</sup> These three genes located in three distinct scaffolds Sc21, Sc47 and Sc67 were termed *praA*, *prbA*, and *prcA*, respectively.

In a first approach, we performed a gene annotation with the web server AntiSMASH in its fungal version.<sup>197</sup> This software predicts CDS and gives annotations of putative gene clusters and also comparisons with known BGC. The resulting annotations of the three scaffolds afforded three BGCs represented in Figure 2.7, which don't have any evident correspondence with known BGC.

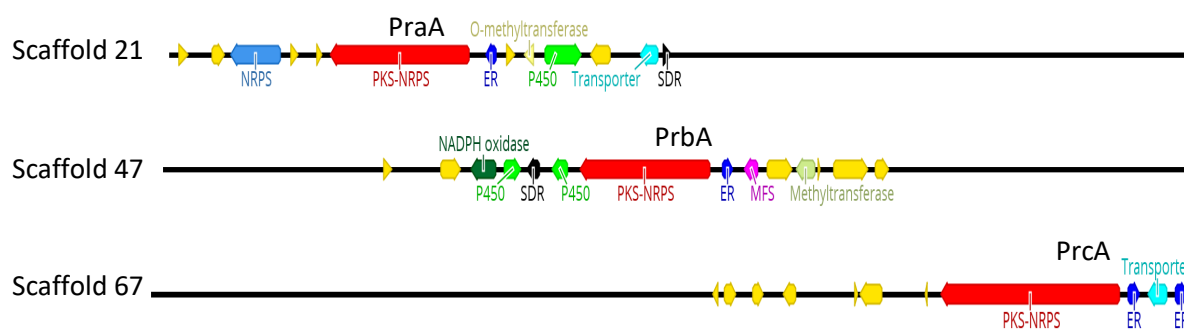


Fig. 2.7 Three genes coding for putative hybrid PKS-NRPSs identified in the genome and identified functions of proteins encoding in the gene clusters: PKS-NRPS is in red, NRPS is in blue, ER is in dark blue, P450 is in green, Transporter is in light blue, O-methyltransferase is in gold, SDR is in black, MFS is in light red, Methyltransferase is in light green, NADPH oxidase is in deep green.

Each putative PKS-NRPS was then analysed by using prediction tools like NCBI BLAST and PKS/NRPS analysis<sup>198</sup> to identify the different catalytic domains. As expected the three hybrid megasynthase harbored ten catalytic domains: the PKS part is composed of KS, AT, DH, MT, KR and ACP domains, and the NRPS part of C, A, T and R domains. (Fig. 2.8). However, we found that a 2.5 kb DNA fragment with no domain prediction was situated at the beginning of the PKS-NRPS-67 gene (Fig. 2.8). Further analysis using FGENESH webserver<sup>199</sup>, allowing open reading frames identification, revealed the presence of another gene in this 2.5 kb sequence named *prcB*.



Fig. 2.8 Architecture of the three putative PKS-NRPSs found in *S. zeae*. A. Ten catalytic domains were identified in the PKS-NRPS coding genes annotated by AntiSMASH: Ketosynthase (KS), Acyl-transferase (AT), Dehydratase (DH), Methyl-transferase (MT), Ketoreductase (KR), Acyl-carrier protein (ACP), Condensation (C), Adenylation (A), Thiolation (T), Reductase (R). B. PKS-NRPS-67 gene from AntiSMASH analyzed by FGENESH.

This indicates that AntiSMASH can give, at first glance, a useful annotation for the whole cluster, however, the results must be considered with caution and verified by manual annotations with a more robust program (e.g. FGENESH).

Prior to this thesis project, the team investigated the possibility to obtain predictive information from the A domain since it selects the amino acid incorporated into the metabolite as mentioned in the introduction. Isotopic incorporation experiments showed that tyrosine is integrated in pyrrocidines, thus the residues composing the active site of this domain were compared with those of already characterized PKS-NRPS A domains. (Tab. 2.10). No exact match has been found between *S. zeae* domains and known domains, only a similarity between residues from A domain 47 with DmbS, TenS or ApdA appeared, suggesting that Tyr is recruited by PKS-NRPS 47. Similarly, PKS-NRPS would likely incorporate Phe, whereas no prediction is possible for PKS-NRPS 67. At this stage the cluster 47 could be a candidate for pyrrocidine biosynthesis. However, due to low number of characterized enzymes and also the variability of the fungal non-ribosomal specificity code<sup>47</sup> (as seen in Tab. 2.10 for ApdA, TenS, DmbS, FsdS, AceI, all selecting Tyr) a high uncertainty remained in this kind of prediction.

Table. 2.10 Comparison of the residues defining the A domain active site of characterized PKS-NRPS with those extracted from the three *Sarocladium zeae* PKS-NRPS.

Organism	Metabolite	Protein	Loaded amino-acid	Residues of the active site in adenylation domain									
				173	174	177	216	239	241	272	280	281	395
<i>Giberella moniliformis</i>	fusarin	FUSS	Hse	D	M	T	F	V	W	G	I	I	K
<i>Metharhizium robertsii</i>	NG-391	NGS1	Hse	D	M	T	F	S	W	G	I	N	K
<i>Xylaria spp</i>	xyrrolin	Pks3	Ser	D	L	L	M	T	W	W	I	V	K
<i>Aspergillus terreus</i>	isoflavipucinin	ATEG_00325	Leu	D	A	S	L	Q	W	A	I	M	K
<i>Aspergillus nidulans</i>	aspyridone	ApdA	<b>Tyr</b>	<b>D</b>	<b>M</b>	<b>V</b>	<b>I</b>	<b>Y</b>	<b>W</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>K</b>
<i>Beauveria bassiana</i>	tenellin	TenS	<b>Tyr</b>	<b>D</b>	<b>M</b>	<b>V</b>	<b>I</b>	<b>T</b>	<b>W</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>K</b>
<i>Beauveria bassiana</i>	desmethylobassianin	DmbS	<b>Tyr</b>	<b>D</b>	<b>M</b>	<b>V</b>	<b>I</b>	<b>T</b>	<b>W</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>K</b>
<i>F. heterosporum</i>	fusaridione	FsdS	Tyr	D	F	E	S	H	W	N	I	A	K
<i>Magnaporthe grisea</i>	?	Ace1	Tyr	D	M	A	Q	S	W	G	I	N	K
<i>Penicillium expansum</i>	chaetoglobosin	CheA	Trp	D	M	I	I	T	W	C	A	A	K
<i>Aspergillus oryzae</i>	cyclopiazonic acid	CpaS	Trp	D	M	A	L	T	W	S	A	C	K
<i>Aspergillus flavus</i>	cyclopiazonic acid	CpaS	Trp	D	M	A	L	A	W	S	A	C	K
<i>Aspergillus fumigatus</i>	pseurotin	PsoA	<b>Phe</b>	<b>D</b>	<b>A</b>	<b>Y</b>	<b>T</b>	<b>S</b>	<b>W</b>	<b>A</b>	<b>I</b>	<b>C</b>	<b>K</b>
<i>Aspergillus clavatus</i>	cytochalasin	CesA	Phe	D	M	S	E	S	W	C	F	C	K
<i>Sarocladium zeae</i>	-	PKS-NRPS 47	<b>Tyr?</b>	<b>D</b>	<b>M</b>	<b>I</b>	<b>I</b>	<b>S</b>	<b>W</b>	<b>C</b>	<b>A</b>	<b>A</b>	<b>K</b>
	-	PKS-NRPS 67	?	D	M	L	Q	T	W	N	L	S	K
	-	PKS-NRPS 21	<b>Phe?</b>	<b>D</b>	<b>A</b>	<b>N</b>	<b>S</b>	<b>T</b>	<b>W</b>	<b>A</b>	<b>I</b>	<b>S</b>	<b>K</b>

We also examined another domain of PKS-NRPS, the R domain which is in charge of the release of the PK-amino acid intermediate from the megasynthase. Two mechanisms of release were proposed (Fig. 2.9): one is reductive release catalyzed by a genuine R domain which generates aldehyde via hydride transfer from NAD(P)H cofactor and a thioacetal intermediate. The aldehyde then undergoes a Knoevenagel condensation to form the pyrrolidinone moiety. The other is catalyzed by a redox-incompetent R\* domain via the formation of the carbanion of the  $\beta$ -ketoamide which undergoes a Dieckmann condensation to afford a pyrrolidine-2,4-dione known as tetramic acid. The work of Liu and Walsh<sup>200</sup> showed that a signature of R\* domain is a mutation in the catalytic triad Ser-Tyr-Lys found in R domains where Tyr is mutated into Leu (in Cyclopiazonic acid synthase, Cpas) or Phe (in Tenellin A synthase, TenS).

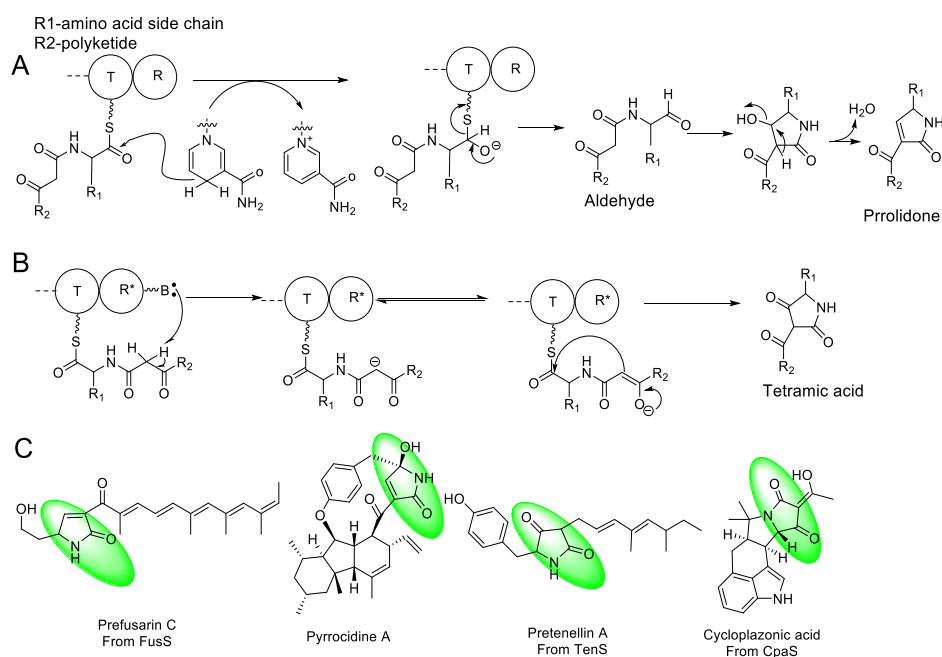


Fig. 2.9 The R domain-catalyzed release of PK-amino acid from the PKS-NRPS (A: reductive release followed by Knoevenagel condensation; B: Dieckmann condensation) and resulting structural differences observed in representative metabolites (C).

By comparing the catalytic triad of R domains, it can be determined which PKS-NRPS cluster is suitable for synthesizing pyrrocidines. The presence of the pyrrolidone moiety in pyrrocidine A suggested that the megasynthase has a reductive domain with a functional triad Ser-Tyr-Lys. The R domain of PKS-NRPS-47 has a mutation in the catalytic triad (Tyr into Leu) (Fig. 2.10). Consequently, PKS-NRPS-47 was excluded from participating in the biosynthesis of pyrrocidines. The other two PKS-NRPSs (21 and 67) have intact triad and could potentially be involved.

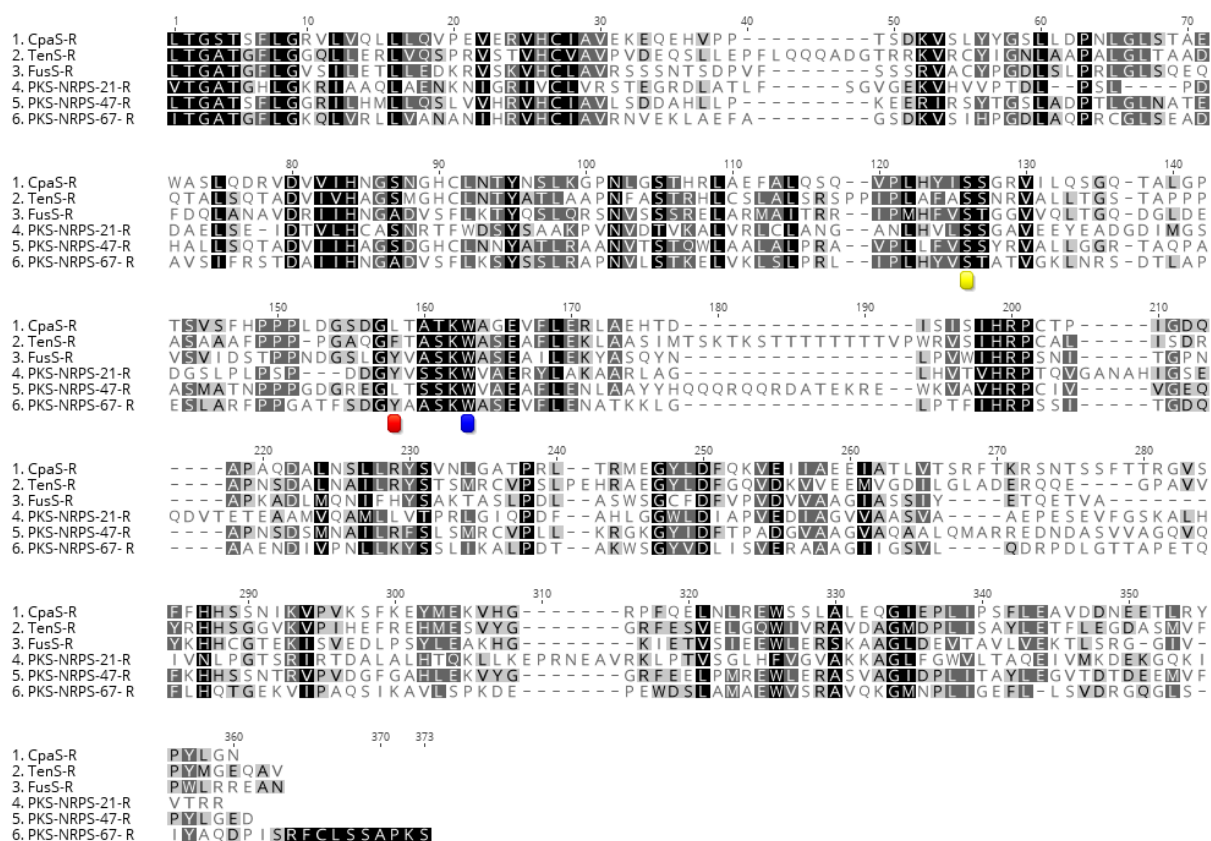


Fig. 2.10 The comparison of protein sequence of R domain from different PKS-NRPS. CapS: cyclopiazonic acid synthase, TenS: tenellin synthase, FusS: fusarin synthase.

Interestingly, in cluster 67, sequence analysis of the protein encoded by the gene *prcB* adjacent to *prcA* predicts a putative Diels-Alderase function. Indeed, the protein shares similarities with decalin synthases CghA<sup>193</sup>, Fsa2 and Phm<sup>785,201</sup> with 20 %, 22 % and 25 % sequence identities, respectively (Fig. 2.11), but with differences like longer sequences at N and C terminus with N terminal predicted to be a transmembrane segment. Diels Alderase seems to be required in pyrrocidine biosynthesis due to the presence of a cyclohexene. Since no other gene encoding a putative Diels-Alderase was detected in cluster 21, the best candidate for pyrrocidine biosynthesis would be the *prc* gene cluster found in scaffold 67.

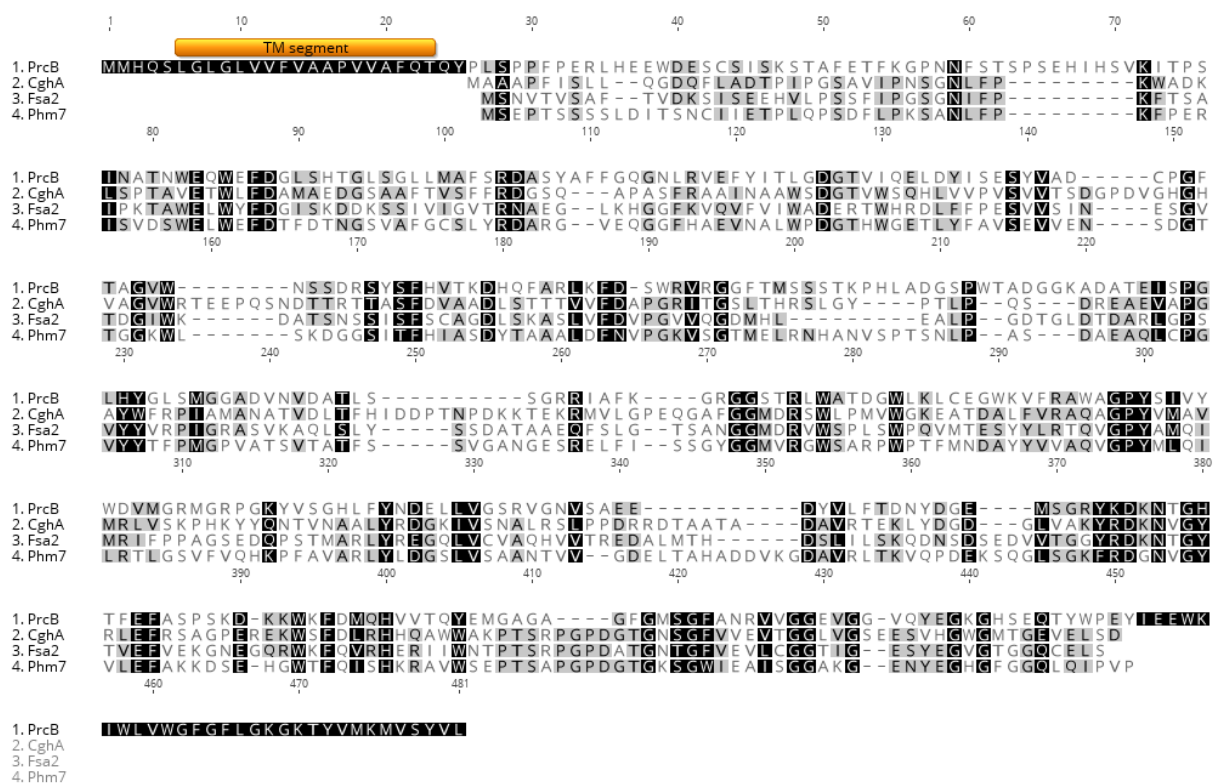


Fig. 2.11 Protein alignment of PrcB and the decalinesynthases CghA, Fsa2 and Phm7 (Diels-Alderases).

### 2.3 Identification of the PKS-NRPS producing pyrrocidines by gene knock-out

In order to validate our hypothesis, we intended to independently knock-out the three PKS-NRPS genes in *S. zae* and analyze the effect of these deletions on the production of pyrrocidines.

Preliminary resistance screening in the laboratory showed that hygromycin has an inhibitory effect on the growth of *S. zae* at a concentration of 50 µg/mL. So *hph* gene (hygromycin resistance) was selected as a screening marker in selecting mutants.

Several strategies were established to transfer DNA into a filamentous fungal genome, as described in the introduction. Homologous recombination of an inactivation cassette via PEG mediated transformation of protoplasts is the easiest method to test and it has been successfully implemented for *S. zae* in the laboratory (Nouha Dallel, M2 2017)

In our case, the inactivation cassettes are composed of three DNA fragments, up and downstream of targeted gene sequences (1 kb homologous arms) and *hph* gene. We choose to use the *in vivo* assembly (IVA) cloning technique to build and clone this sequence in the pUC18

vector directly. IVA cloning is an effective, robust and cheap technology to assemble the DNA fragments using the intrinsic bacterial recombination pathway, the *recA*-independent recombination pathway. This method is known from the early nineties<sup>202</sup> and has been used and developed by some laboratories until a more recent publication that standardizes the process.<sup>203</sup> The method requires that the DNA fragments to be assembled should have homologous sequences around 30 bp at their termini (Fig. 2.12). Before undergoing the IVA cloning, all required DNA fragments and primers should be amplified by PCR. Therefore, in the IVA cloning, the primers should consist of two regions: the 3' end binds template DNA for PCR amplification, while modifications and homologous sequences are encoded at the 5' end. The *hph* gene was amplified from pSN44 vector by using the primers RH-hphF (197) and RH-hphR (198). The linear backbone of pUC18 was obtained by PCR using the primers RhpUC18F(73) and RhpUC18R(74) on the native vector. Both one kb homologous arms were obtained by using two pairs of primers FL1F(xxx) and FL1R(xxx), FL2F(xxx) and FL2R(xxx) (Fig. 2.12). PCR and sequencing identified the correct construction of the plasmids.

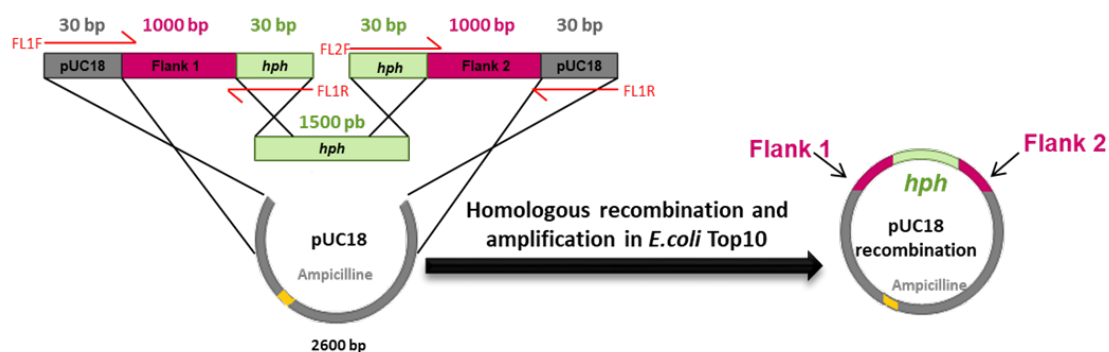


Fig. 2.12. The construction of plasmids with inactivation cassette.

To inactivate our targeted genes, we used the split-marker method.<sup>204</sup> For that, the gene inactivation cassettes were then amplified from the plasmids with two couple of primers, ConF(xxx), RhHphSplt1R(201) and ConR(xxx), RhHphSplt1F(202), yielding two DNA fragments for each gene inactivation. The first one contains the upstream homology to target genes fused to the first portion of the hygromycin resistance gene. Another include the second portion of the hygromycin resistance gene fused to the downstream homology of target genes (Fig 2.13). These sequences will be the place of three crossing-over events catalyzed by the fungal machinery in the presence of an inactivation cassette having the same flanking sequence and a selection marker (*hph* resistance gene) (Fig. 2.13). As a result, the target gene will be replaced by the *hph* expression cassette, giving rise to the inactivation of the target gene. Two

pairs of primers Ext F(xxx) and Ext R(xxx), Int F(xxx) and Int R(xxx) were both used to ensure accurate knockout of the target gene (Fig 2.13).

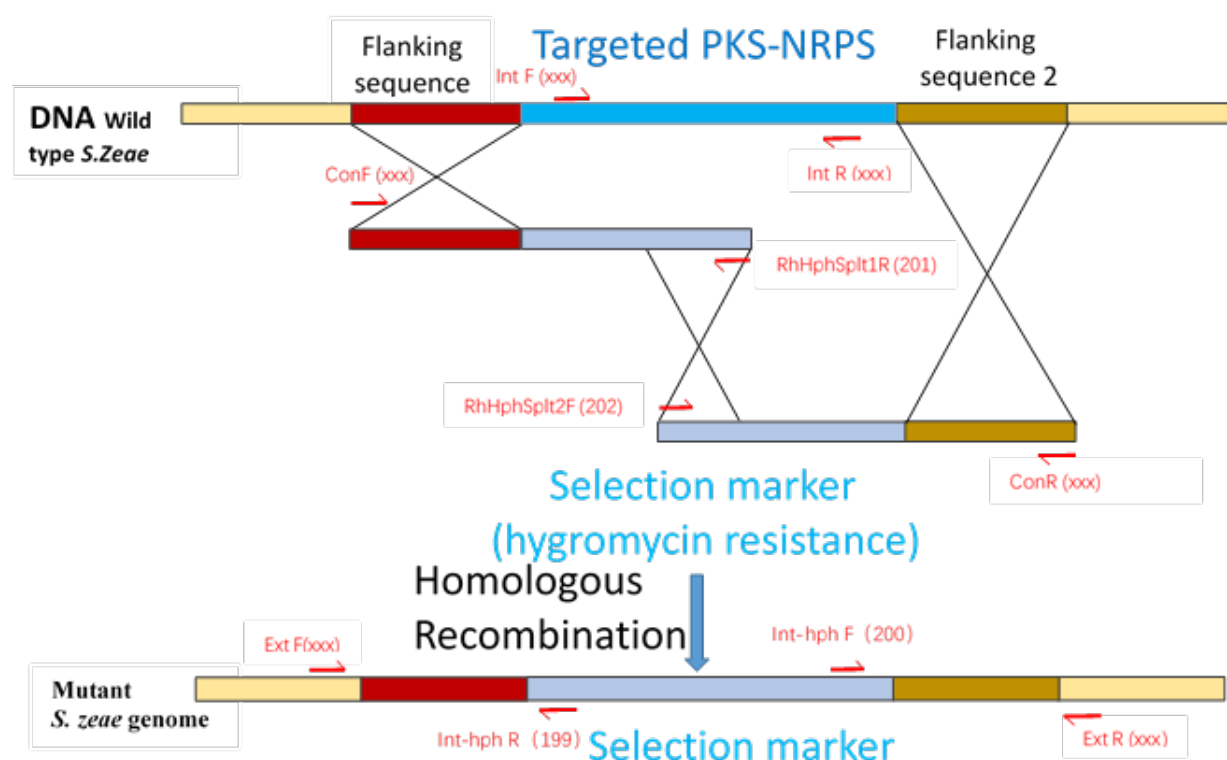


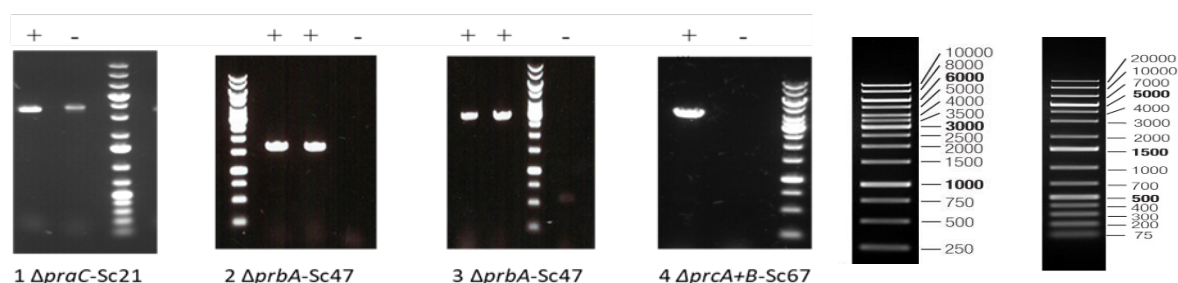
Fig. 2.13 The schemes of split approaches for gene inactivation.

To verify our hypothesis that PKS-NRPS-67 gene cluster is linked to pyrrocidines, first the gene inactivation was performed using a cassette targeting the DNA sequence comprising *prcA-prcB* genes. After generation of transformants cultivated for five days on a hygromycin selecting minimum medium, the colonies were separated and grown in individual plates for one week. Mycelium of each transformants were produced during one more week in a liquid medium culture. A PCR screening on the extracted DNA of the transformants was performed to verify the presence of the inactivation cassette at the desired locus and the deletion of the targeted DNA. After screening of about 50 single transformants, one positive mutant strain *S. zeae-ΔprcA+B* with accurate inactivation of *prcA* and *prcB* was obtained. Fig. 2.14-4 showed the *prcA+prcB* was replaced by *hph* in *ΔprcA+B*. Fig. 2.15-3 demonstrated the *prcA+prcB* was deleted in *ΔprcA+B*.

Our first trials to knocked out the whole gene sequence *praA* and *prbA* coding either PKS-NRPS-21 or PKS-NRPS-47 separately were unsuccessful. So, we tried to delete shorter DNA sequence which could also impact the metabolite production of the BGC. Finally, we could



isolate two transformants  $\Delta prbA$  with the partial deletion of PKS-NRPS-47 gene on the first 2.5 kbases and one transformant  $\Delta praC$  with the deletion of gene coding the enoyl-reductase (ER) associated to the putative PKS-NRPS-21. PCR characterizations of these mutants are presented on Fig. 2.14 and 2.15.



Lane	Deletion mutants expected size (bp)	Gel band size (bp)	Lane	Wild type expected size (bp)	Gel band size (bp)
1+	3506	~3500	1-	3767	~3800
2+	1661	~1500	2-	0	0
3+	2862	~2800	3-	0	0
4+	3994	~4000	4-	15246	X

Fig. 2.14 Verification of the insertion of the cassette in the mutants, 1 and 4 used the external primers (Ext XXX), 2 and 3 used the external primers (Ext XXX) and hph internal primers (Int-hph XXX).

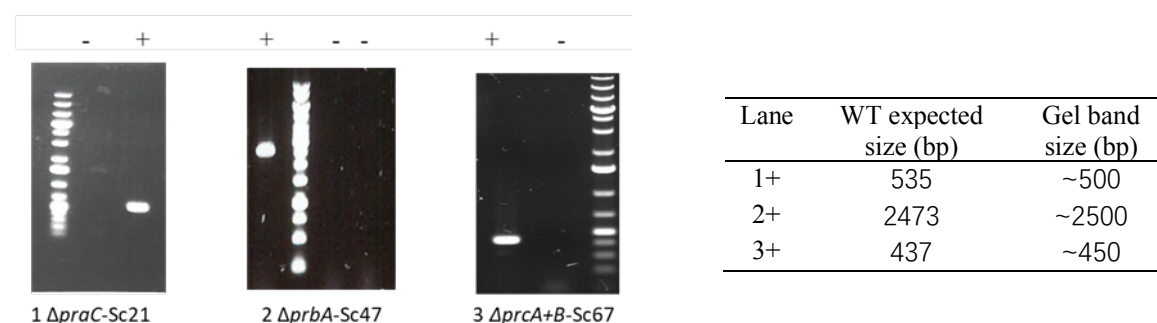


Fig. 2.15 Verification of the deletion of the targeted gene of mutants, using the internal primers (Int XXX).

The three deletion mutants were cultivated and their ethyl acetate metabolic extracts were analyzed by LC-MS. EIC of  $m/z$  488 and 490 corresponding to  $[M+H]^+$  molecular ion of pyrrocidines A and B revealed that the production of this metabolites is abolished in the  $\Delta praC + prcB$  mutant but not in other mutants. Since dihydroresorcylyde is produced by *S. zeae*,<sup>205</sup> we used it as a positive control in LC-MS experiments ( $[M+H]^+$  ion at  $m/z$  293). Its intensity did not change in the mutants. So, we can conclude that the PKS-NRPS encoded by *prcA* gene in cluster 67 is responsible for synthesizing pyrrocidines.

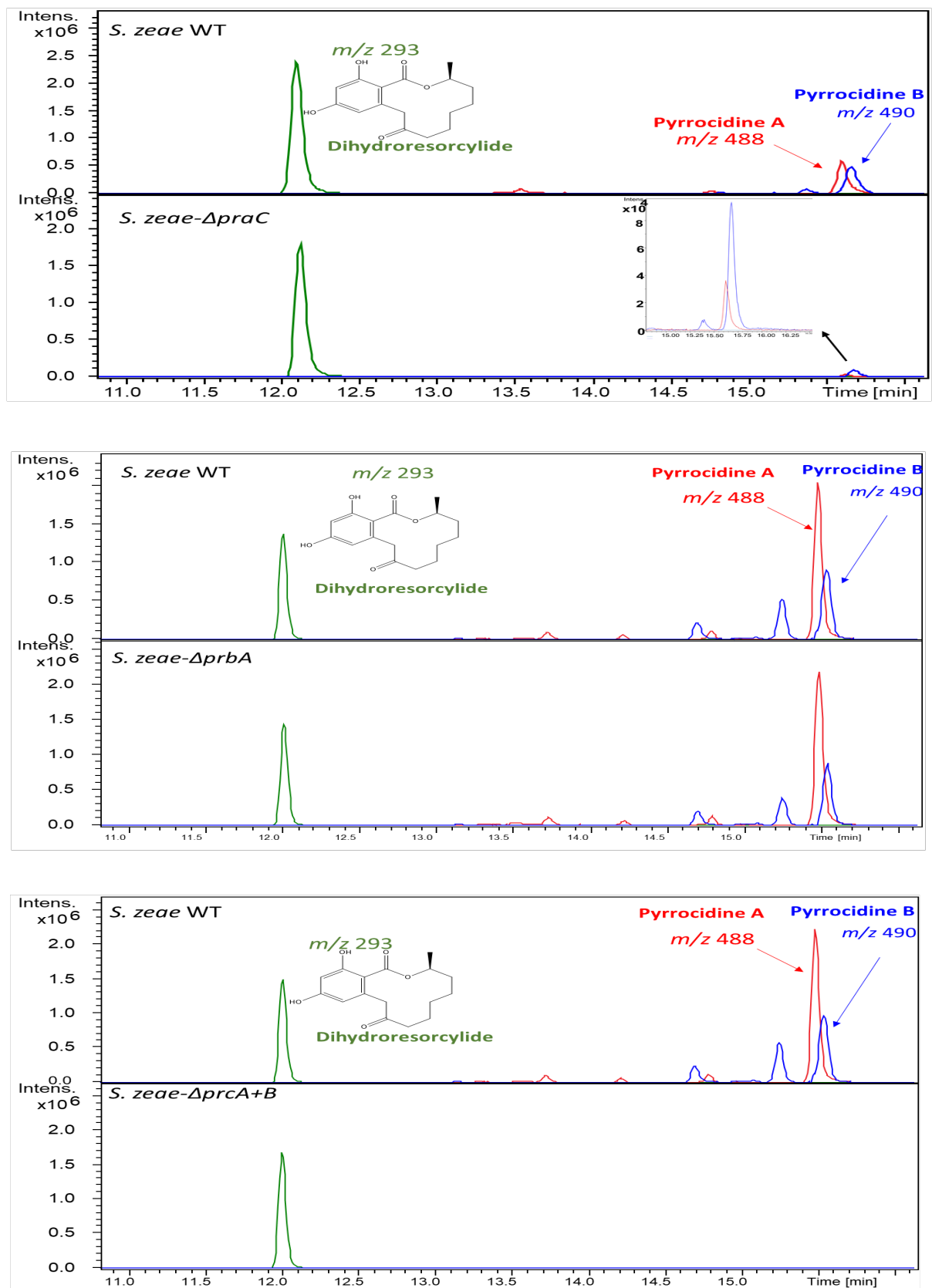


Fig. 2.16 The Extraction Ion Chromatograms of the wild type *S. zeae* and the mutants ( $m/z$  293.137 [M+H]<sup>+</sup> (dihydroresorcylic acid) control,  $m/z$  488.277 [M+H]<sup>+</sup> pyrrocyclin A,  $m/z$  490.293 [M+H]<sup>+</sup> pyrrocyclin B).

## 2.4 Annotation of pyrrocidine BGC

Prediction of coding sequences (CDS) surrounding *prcA* with FGENESH software revealed nine putative proteins encoded by genes annotated from *prcH* to *prcI* (Fig. 2.18). From sequence analysis we identified PrcC as the *trans*-acting enoyl reductase (ER) associated to the hybrid PKS-NRPS PrcA, PrcH as an  $\alpha/\beta$ -hydrolase, PrcF as a transcription factor, PrcG as a transporter, PrcI as a second ER and PrcB homologous to Diels-Alderases, as well as PrcD and PrcE appearing as integral membrane proteins with unknown function (Fig. 2.18).

Because *prcI* is located at the extremity of scaffold 67, we investigated the possibility that the BGC is split between two scaffolds. At the origin of this project, a 5,000 cloned fosmid library built from *S. zeae* genomic DNA was screened by dot blot hybridization using the *prcI* DNA sequence as a probe (Alexandre Delporte M2 2016). This allowed the team to identify a fosmid containing the BGC and its right-contiguous scaffold. Thus, scaffold 152 was supposed to be adjacent to the tail of scaffold 67. To further confirm the relationship between scaffold 67 and scaffold 152, one pair of primers (PbSc67->Sc152(156) and PbSc152->Sc67(157)) were designed to amplify the junction between both (Fig. 2.17). The size of the PCR amplicon is around 450 bp (Fig. 2.17) whereas the minimum length expected for this amplification is 432 bp, based on the scaffold sequences meaning that there is no gap between scaffold 67 and 152. Sequencing of the PCR band confirmed the contiguity of scaffold 67 with scaffold 152.

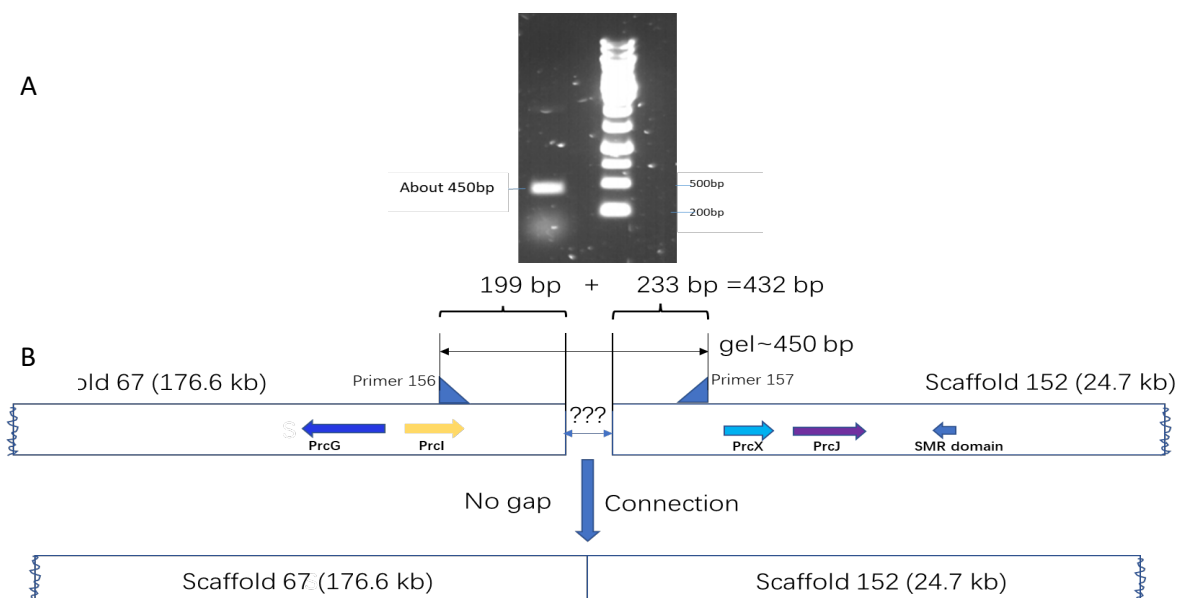
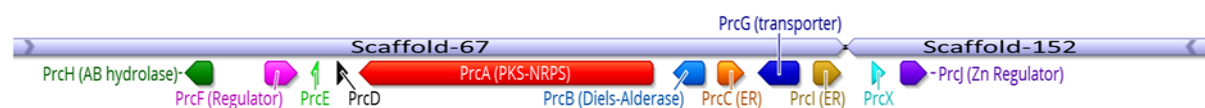


Fig. 2.17 Analysis of the junction between Scaffolds 67 and 152. A: Agarose gel electrophoresis of the PCR product obtained from primers 156 and 157. B: View of the junction scaffold 67-scaffold 152.

Sequence analysis of the beginning of Scaffold 152 indicated the presence of two genes coding for a protein belonging to the lipocalin family with unknown function and a transcription factor (Fig. 2.18). These two genes possibly belonging to the BGC were named *prcX* and *prcJ*, respectively. Then, a sequence coding for a Smr domain (Small MutS related protein) possibly involved in nicking endonuclease, marks very likely the limit of the BGC.



Protein	Size (aa)	Predicted function	Percent identity/known function	NCBI accession/Species (with predicted function)
PrcA	4021	PKS-NRPS	54%/FsdS	GJC90835.1/ <i>Colletotrichum spaethianum</i>
PrcB	433	Diels-Alderase	77%/CcsF	GJC90834.1/ <i>Colletotrichum spaethianum</i>
PrcC	396	Enoyl reductase	76%/Ccsc	GJC90833.1/ <i>Colletotrichum spaethianum</i>
PrcD	163	Unknown	69%/unknown	XP_018074793.1/ <i>Phialocephala scopiformis</i>
PrcE	125	Unknown	87%/unknown	KXH65045.1/ <i>Colletotrichum salicis</i>
PrcF	497	Zn regulator	53%/AceII	XP_003069900.1/ <i>Coccidioides posadasii</i>
PrcG	615	Transporter	81%/Fus6	GJC90832.1/ <i>Colletotrichum spaethianum</i>
PrcH	419	$\alpha/\beta$ hydrolyse	83%/CcsE	GJC90837.1/ <i>Colletotrichum spaethianum</i>
PrcI	366	Enoyl reductase	32%/PoxH	A0A1W5T4X6.1/ <i>Penicillium oxalicum</i>
PrcJ	439	Zn regulator	21%/Fus10	XP_007826306.1/ <i>Metarhizium robertsii</i> ARSEF 23
PrcX	185	Unknown	83%/unknown	KXH56094.1/ <i>Colletotrichum salicis</i>

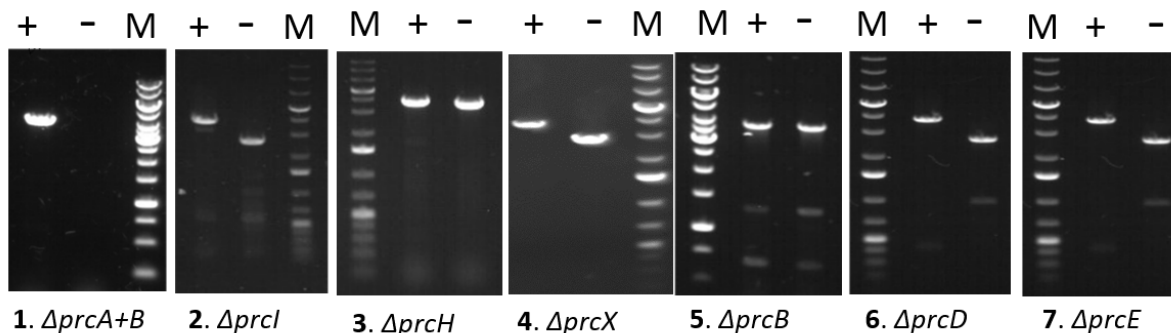
Fig. 2.18 The annotation of the related genes in PKS-NRPS-67 gene clusters.

## 2.5 Identification of the function of auxiliary enzymes by gene inactivation

### 2.5.1 Construction of the deletion mutants of auxiliary enzymes and LC-MS analysis

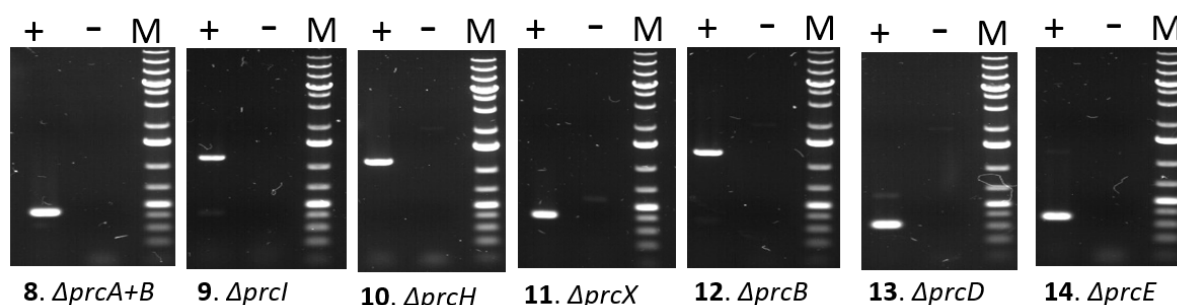
While the PKS-NRPS assisted by its *in trans* enoyl-reductase PrcC is in charge of the construction of the linear nonaketide-tyrosine intermediate, corresponding to the backbone of pyrrocidines, the auxiliary enzymes organize the different cyclization or redox steps to give pyrrocidines. To assign a precise function to the latter proteins, we performed gene inactivation in *S. zeae* with the same procedure as described above and analyzed their effects on the metabolite production. Thus, we generated individual deletion mutants for each gene coding putative auxiliary enzymes (i.e.  $\Delta prcB$ ,  $\Delta prcD$ ,  $\Delta prcE$ ,  $\Delta prcH$ ,  $\Delta prcI$ ,  $\Delta prcX$ ). These mutants were verified by PCR amplification (Fig. 2.19 and 2.20) and sequencing of the amplicons. Attempts to inactivate both gene *prcF* and *prcJ* coding transcription factors remained unsuccessful.

Each mutant was cultivated on PDA medium and the metabolic extracts were analyzed by LC-MS. A selection of extracted ion chromatograms corresponding to accumulating metabolites are represented on the Fig. 2.21 and detailed EIC are presented in annex 7-14. These accumulating products were then isolated and purified from maize medium culture their structure elucidated by NMR analysis. These results are presented below.



Lane	Deletion mutants Expected size (bp)	Gel band Size (bp)	Lane	Wild type Expected size (bp)	Gel band Size (bp)
1+	3 994	~ 4 000	1-	15 246	x
2+	3 589	~ 3 600	2-	2 578	~ 2 500
3+	3 702	~ 3 700	3-	3 555	~ 3 500
4+	3 583	~ 3 500	4-	2 781	~ 2 700
5+	3 951	~ 4 000	5-	3 812	~ 3 800
6+	3 858	~ 3 900	6-	2 919	~ 2 900
7+	3 611	~ 3 600	7-	2 558	~ 2 500

Fig. 2.19 Verification of the insertion of the inactivation cassette in deletion mutants by PCR. Amplifications were performed with pairs of primers external to the cassette (ExtF(xxx) / ExtR(xxx)) using genomic DNA from the mutants (Lane +) and the wild type *S. zeae* (Lane -). Expected and estimated band sizes obtained by 1 % agarose gel electrophoresis are reported in the table.



Lane	Wild type Expected size (bp)	Gel band Size (bp)
8+	437	~ 400
9+	1 188	~ 1 200
10+	1 172	~ 1 200
11+	404	~ 400
12+	1 224	~ 1 200
13+	332	~ 300
14+	349	~ 400

Fig. 2.20: Verification of the deletion of the targeted genes in *S. zeae* mutants by PCR. Amplifications were performed with pairs of primers internal to the cassette (IntF(xxx) / IntR(xxx)) using genomic DNA from the wild type *S. zeae* (Lane +) and the mutants (Lane -). No amplification was detected in mutants compared to the wild type showing the deletion of the targeted genes. Expected and measured band sizes obtained by 1 % agarose gel electrophoresis are reported in the table.

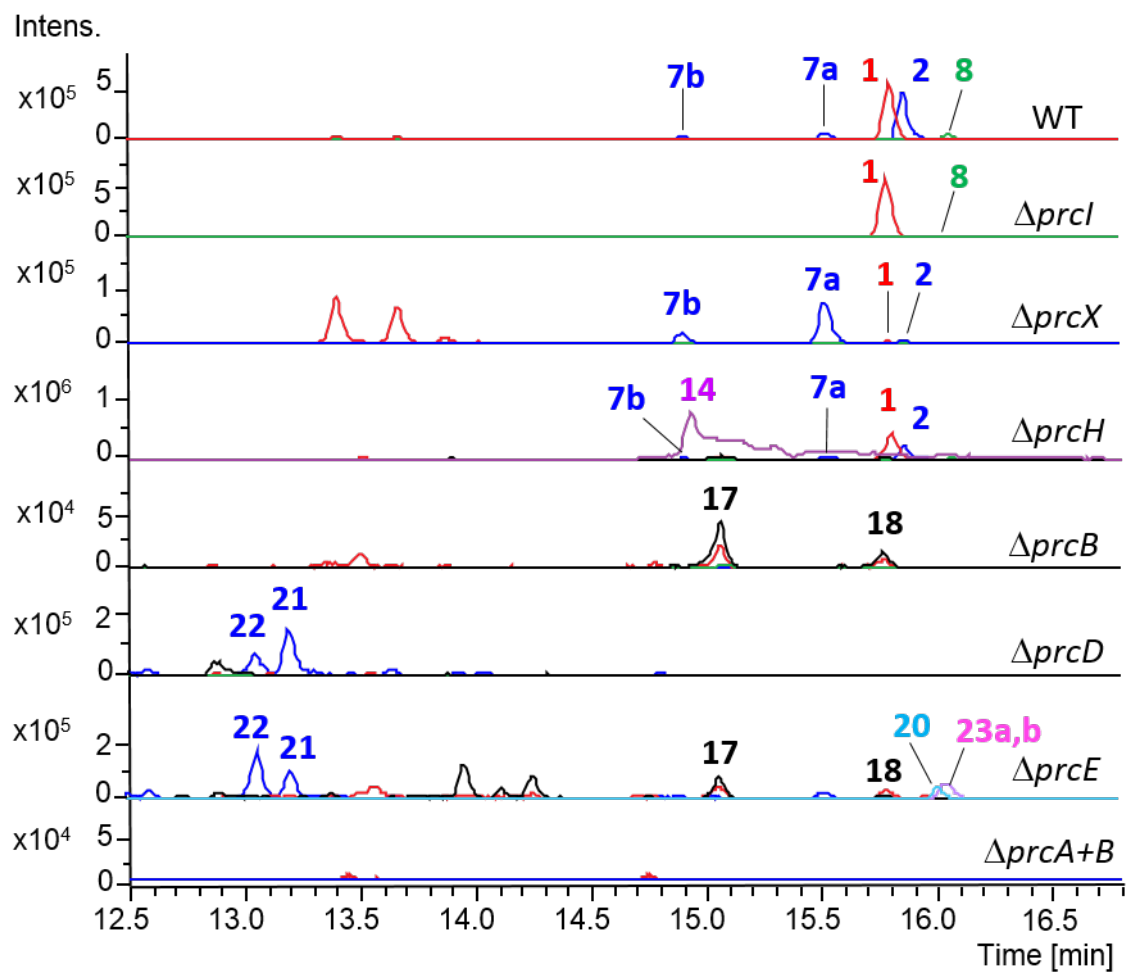


Fig. 2.21 A selection of extracted ion chromatograms corresponding to accumulating metabolites in each mutant and wild type.

## 2.5.2 Analysis of the inactivation of the gene coding PrcI

*S. zeae-ΔprcI* mutant was obtained by gene inactivation of *prcI*. The production of pyrrocidines B **2**, D **7a** and **7b** were abolished in *S.zeae-ΔprcI* mutant. The identical concentration of pyrrocidine A was still produced in this mutant compared to *S.zeae* (Fig. 2.22). PrcI converts pyrrocidine A to pyrrocidine B by reducing the double bond of the pyrrolidone moiety (Fig 2.23).

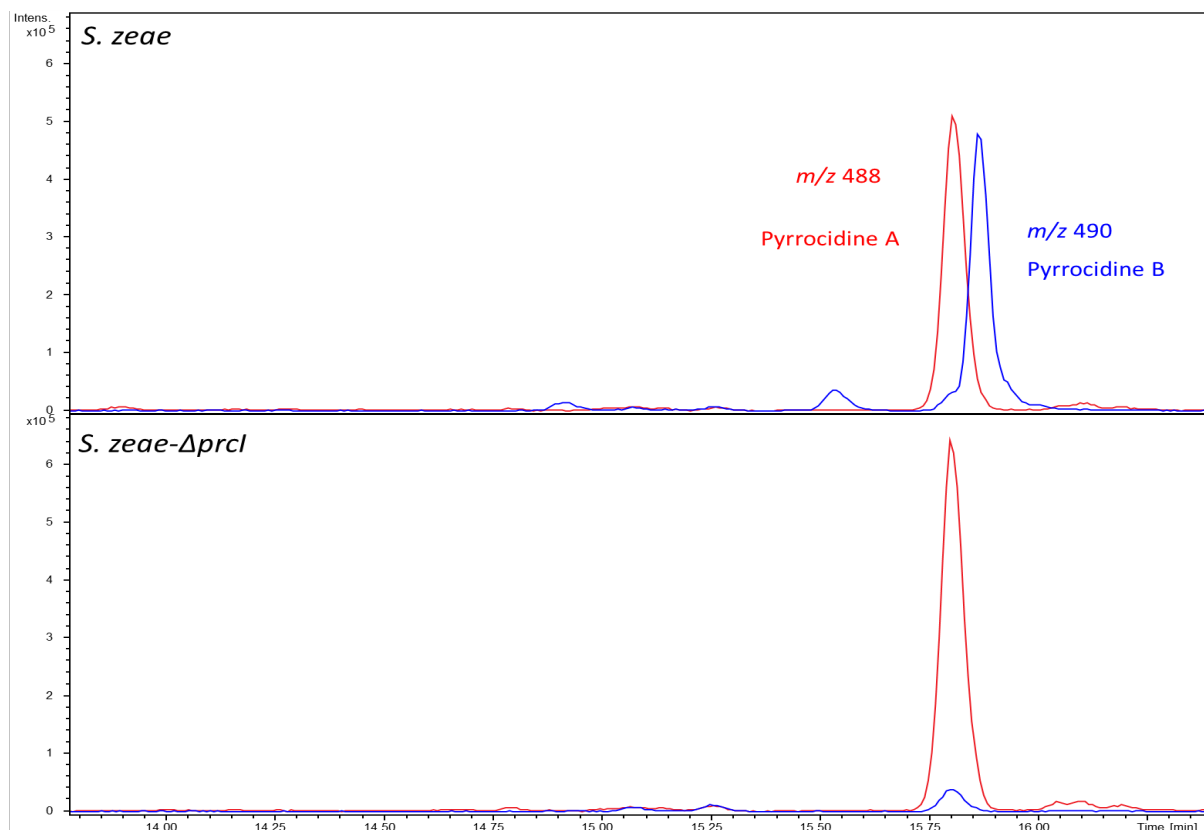


Fig. 2.22 The different extraction ion chromatogram between WT and *S. zeae- ΔprcI*.

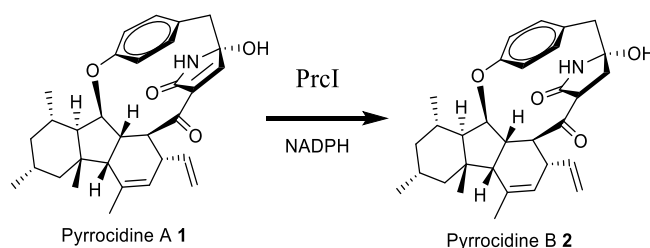


Fig. 2.23 Conversion of pyrrocidine A into pyrrocidine B catalyze by PrcI.



### 2.5.3 Analysis of the inactivation of the gene coding PrcX

The deletion of *prcX* gives rise to a substantial production decrease of **1** and **2** without impacting the production of pyrrocidines D **7a,b** (Fig. 2.24 and annex 11). This result shows that **1** and **2** are naturally produced without this lipocalin protein.

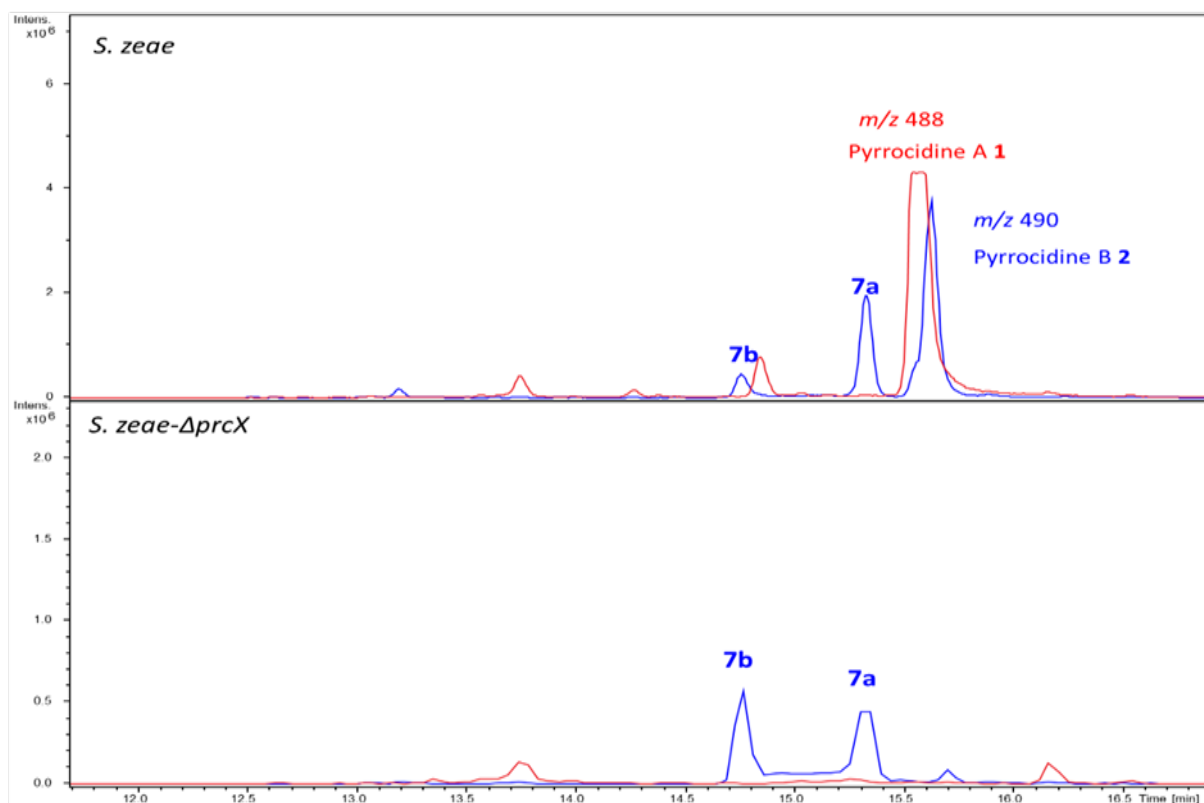


Fig. 2.24 The different extraction ion chromatogram between WT and *S. zeae-ΔprcX*.

The diastereomery observed between **2** and **7a/b** at C3/C6 can be explained by IMDA reaction forming ring C from the precursor **25**. Indeed, diene and dienophile of **25** can react according to *exo* or *endo* approaches to give **1** or **27**, respectively. We deduced that PrcX drives the Diels-Alder cycloaddition in favor of the *exo* product (Fig. 2.25).

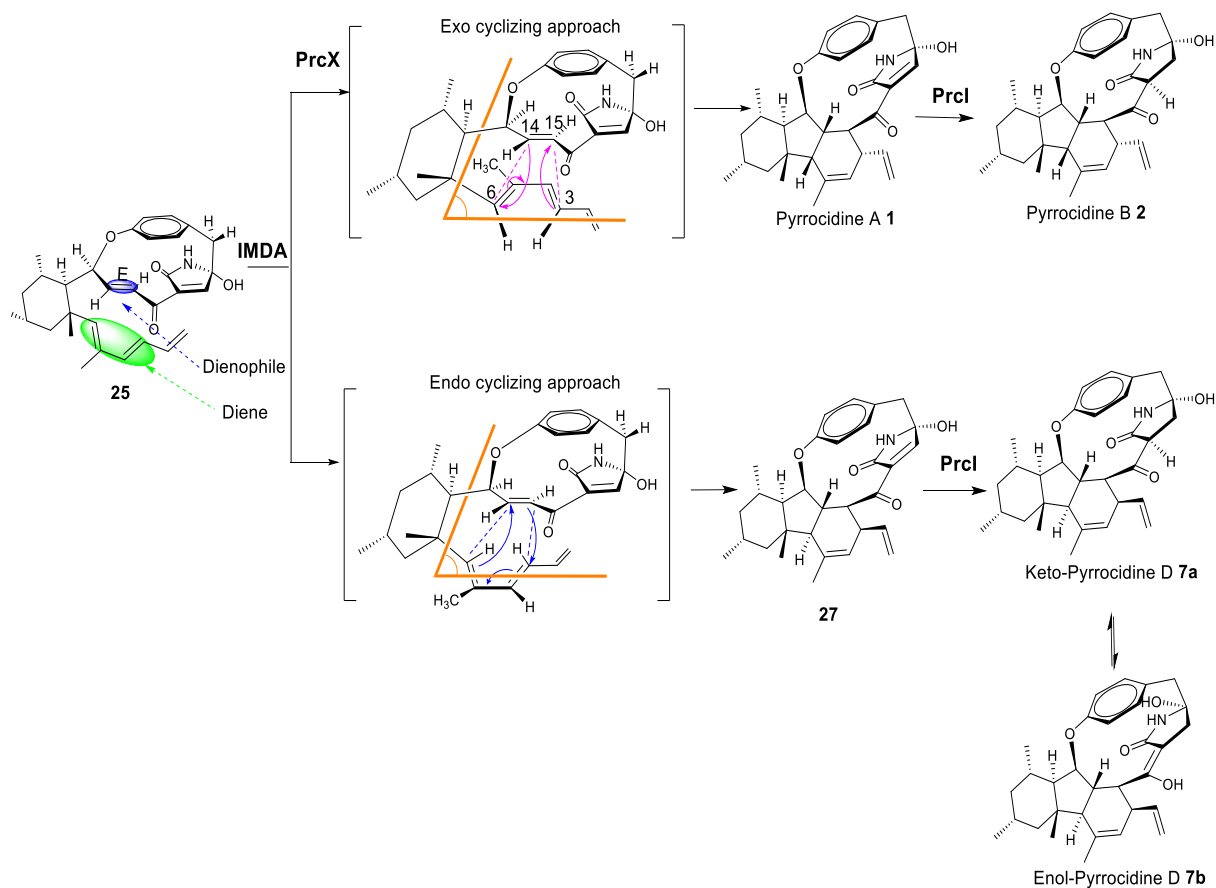


Fig. 2.25. Formation of pyrrocidines *via* IMDA from precursor **25**. Cycloaddition of the diene and dienophile according to an *exo*-cyclizing approach affords pyrrocidine **A 1** and the **B 2**. *Endo* cyclization gives rise to the non-isolated intermediate **27** which is reduced into pyrrocidines **D 7a,b**. PrcX favors the *exo*-pathway.

Moreover, in this mutant due to the dramatic decrease of pyrrocidine B, we could detect and purified from maize medium two new metabolites named pyrrocidine E **10** and sarocladiene A **12**. The structure of the latter, also produced by wild type *S. zea*e, with a molecular formula  $C_{32}H_{39}NO_5$  ( $[M+H]^+$  ion at  $m/z$  518.2901) was assigned by NMR analysis to the C3-epimer of GKK1032C.<sup>206</sup> Indeed, NOESY spectrum through the cross-correlations H31-H24-H13-H6-H12-H15, H13-H15-H1a-H6 and H14-H26-H29 shows that H12, H13, H15, H6, methyl 31 and vinyl group are cofacial while H14 and methyl 29 are on the opposite face of the decahydrofluorene (Table 2.11). The formation of **12** indicates that PrcA is able to perform four and five methylations on the polyketide which is to our knowledge the first reported example for PKS-NRPS.

Compound **10** has the molecular formula  $C_{31}H_{37}NO_5$  ( $[M+H]^+$  at 15.4 min,  $m/z$  504.2736) whose structure was assigned to the C3, C6 diastereomer of pyrrocidine C **9** (Fig. 2.26). Like pyrrocidine C **9** comes from **1**, spontaneous formation of pyrrocidine E **10** is formed from

intermediate **27** (not isolated). The *trans*-fused junction of rings B and C in **10** and also in **12** was corroborated by the shielding of C7, C28 and C29 chemical shifts as mentioned in **7a,b**.

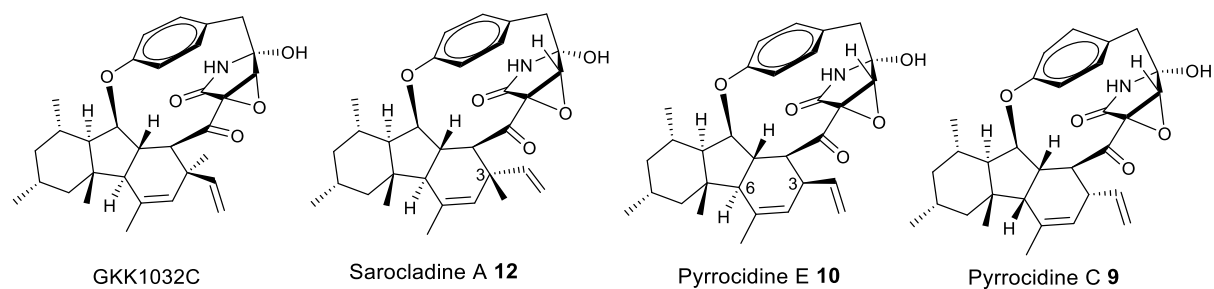
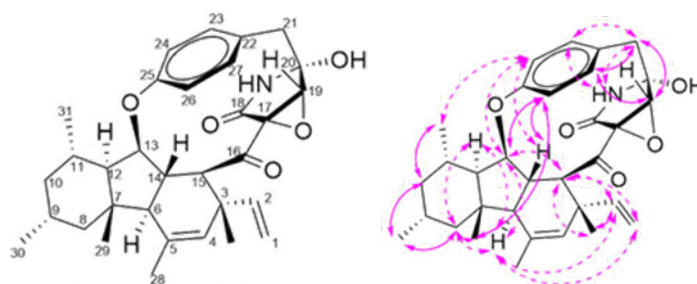
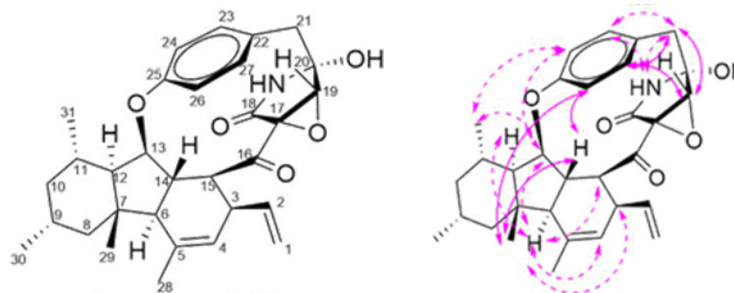


Fig. 2.26: Structure of GKK1032C, Sarocladine, pyrrocidine E and pyrrocidine C.

Table 2.11: NMR data of Sarocladine A **12** in CDCl<sub>3</sub> (600 Hz).

Key NOESY correlations

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1a	111.6	5.02	dd (17.2; 1.0)		2	6;15
1b		4.93	dd (10.1; 1.0)		2	
2	145.7	5.01	dd (17.2; 10.1)	1a; 15; 32	1a; 1b	6; 15; 32
3	45.7	-	-	1a; 1b; 2; 15; 32	-	
4	128.9	4.98	ddd (4.0; 3.2; 1.5)	2; 15; 28; 32	3; 28	
5	140.5	-	-	28	-	
6	53.2	1.98	d (12.8)	4; 14; 28; 29	4; 14	1a; 2; 8b; 12; 13
7	41.3	-	-	8b; 12; 13; 29	-	
8a	48.6	1.88	dd (12.1; 4.4)	29; 30	8b	29; 30
8b		0.74	dd(12.1; 12.1)		8a	6; 12, 30
9	28.0	1.83	m	8b; 10b; 30	30	
10a	45.6	1.79	m ( <i>eq.</i> )	30; 31	10b	30
10b		0.59	ddd (12.1 x3) ( <i>ax.</i> )		10a	30; 31
11	27.2	1.87	m	10a; 31	31	
12	60.7	1.04	dd (11.3; 7.9)	29; 31	13; 14	8b; 6
13	88.3	4.61	dd (7.9; 5.2)	14; 15	12; 14	6; 15; 24
14	51.7	2.41	ddd (12.8; 6.5; 5.3)	15	6; 12; 13	26
15	54.9	3.95	d (6.6)	2; 13; 14; 32	-	1a; 2; 13; 32
16	203.5	-	-	14; 15	-	
17	58.6	-	-		-	
18	168.1	-	-		-	
19	66.1	3.79	d (2.8)	21b	NH	21a; 27
20	84.3	-	-	19; 21a; 21b	-	
21a	43.8	3.18	d (13.4)	23; 27	21b	19; 27
21b		3.12	d (13.4)		21a	NH; 23
22	127.0	-	-	21a; 21b; 24; 26	-	
23	132.1	7.07	dd (8.5; 2.5)	21a; 21b; 27	24	NH
24	119.7	7.15	dd (8.5; 2.5)		23	14; 31
25	158.6	-	-	23; 27; 24; 26	-	
26	122.9	6.74	dd (8.5; 2.5)		27	13
27	131.7	7.07	dd (8.5; 2.5)	21a; 21b, 23	26	19; 21a
28	20.1	1.91	t (1.3)	4	4	
29	16.3	1.50	s	8b; 12		8a; 26
30	22.9	0.89	d (6.3)		9	8b; 10b
31	19.7	1.07	d (6.2)		11	10b; 24
32	26.9	0.99	s	4		2; 15
NH		5.63	broad		19	21b; 23
OH			broad			

Table 2.12: NMR data of Pyrrocidine E **10** in CDCl<sub>3</sub> (400 Hz).


Key NOESY correlations

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1a	117.2	5.08	d (17.5)		2	
1b		5.06	d (10.5)			
2	136.0	5.82	ddd (17.5; 10.2; 7.2)	1b	1a; 1b	
3	43.7	2.91	m	1a; 1b; 2; 15	4; 15; 28	6
4	121.8	5.30	s (broad)	2; 15; 28	3; 28	
5	140.0	-	-	6; 14; 28	-	
6	56.1	1.61	d (11.0)	14; 28; 29	14	3; 12; 13; 15
7	41.5	-	-	13	-	
8a	48.6	1.94	m	29; 30	8b	
8b		0.81	dd (11.8; 11.8)		8a	
9	28.0	1.85	m	8b; 10b; 30	-	
10a	45.6	1.83	m	8a; 30; 31	10b	
10b		0.62	ddd (13.3)		10a	
11	27.2	1.92	m	31	12	
12	61.0	1.15	m	13; 29; 31	11; 13	6
13	89.2	4.58	dd (7.8; 4.8)	14; 15	12; 14	6; 24; 31
14	51.7	2.18	ddd (12.1; 6.8; 4.8)	6; 15	6; 13; 15	26; 29
15	52.6	3.98	dd (7.4)	6; 13; 14	3; 14	6
16	203.8	-	-	14; 15	-	
17	58.1	-	-	NH	-	
18	167.9	-	-		-	
19	66.1	3.72	d (2.7)	21a; 21b; NH	-	21a; 27
20	84.2	-	-	19; 21a; 21b	-	
21a	43.8	3.17	d (13.4)	27	-	19
21b		3.09	d (13.4)			
22	127.5	-	-	21a; 21b; 24; 26	-	
23	132.3	7.06	dd (8.2; 2.1)	26; 27	24	21b; NH
24	120.5	7.11	dd (8.2; 2.1)	26	23	13; 31
25	158.6	-	-	13; 23; 27; 26	-	
26	123.0	6.72	dd (8.2; 2.1)	24	27	14; 29
27	132.0	7.04	dd (8.2; 2.1)	23	26	19; 21a
28	19.9	1.94	s		3; 4	
29	16.6	1.19	s	6; 8b; 12	-	26
30	22.9	0.91	d (6.2)		-	
31	19.7	1.01	d (6.2)		-	13; 24
NH		5.70	s broad		-	21b; 23
OH		3.17	s broad			

## 2.5.4 Analysis of the inactivation of the gene coding PrcH

The gene coding PrcH (annotated as  $\alpha/\beta$ -hydrolase) was inactivated to obtain the mutant *S. zeae- $\Delta$ prcH*. By comparing the UV chromatogram between wild type *S. zeae* and *S. zeae- $\Delta$ prcH*, a new metabolite with a prominent peak was found in *S. zeae- $\Delta$ prcH* (Fig. 2.27).

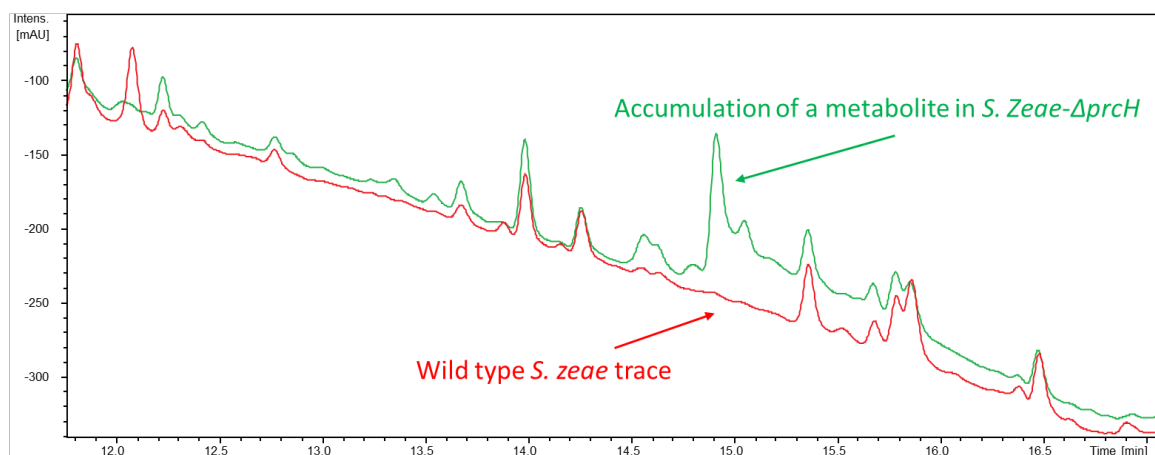


Fig. 2.27 The UV chromatogram of 200-500nm comparison between *S. zeae* and *S. zeae- $\Delta$ prcH*.

This metabolites with  $[M+H]^+$   $m/z$  494.3297 has the molecular formula  $C_{31}H_{43}NO_4$ . Its structure was assigned by NMR to the alcohol **14** (Table 2.13). The linear structure was established thanks to HMBC and COSY correlations. The all-E double bonds were deduced from the  $^3J$  coupling constants measured for ethylenic protons ( $^3J = 15.5$  Hz for C15=C14,  $^3J = 15.2$  Hz for C13=C12,  $^3J = 15.0$  Hz for C3=C2) and the NOESY spectrum showing the correlations H3-H28 and H4-H6 for C4=C5 alkene and H8-H6 for C5=C6 bond. The linear PK structure of this compound suggests that it directly derives from the release of the PKS-NRPS product. This gives us valuable elements on the structure of the polyketide generated by the cryptic PKS part of the megasynthase, which is essential to establish the biosynthetic route of pyrrocidines.

PrcH is homologous to other  $\alpha/\beta$ -hydrolase associated with PKS-NRPS. When the genes coding  $\alpha/\beta$ -hydrolase were knocked out in these systems, accumulations of alcohol were observed.<sup>207–209</sup> These products are not expected to be produced in the normal pathway and they are called shunt products (Fig. 2.28).

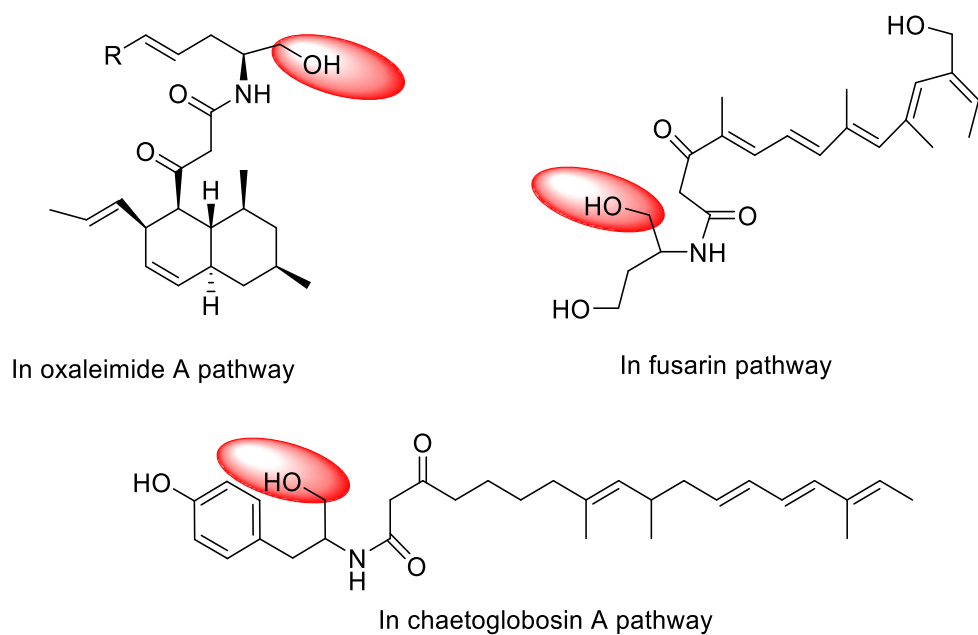
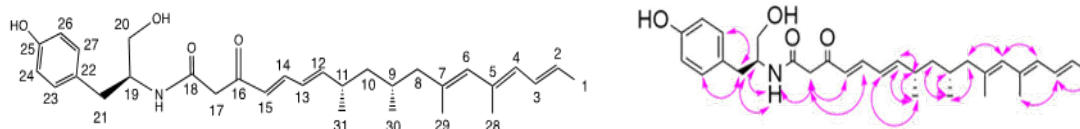


Fig. 2.28 Examples of alcohol shunt products produced when gene coding  $\alpha/\beta$ -hydrolase (PoxO, Fus2, OrfZ) is inactivated in PKS-NRPS BGCs or when PKS-NRPS and ER are heterologously co-expressed without  $\alpha/\beta$  hydrolase.

Table 2.13: NMR data of compound **14** in CDCl<sub>3</sub> (600 Hz).

## Key NOESY correlations

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC from C to	COSY	NOESY
1	18.7	1.80	d (6.6)	2	2	-
2	128.5	5.67	dq (15.0; 7.3)	-	1; 3	-
3	128.3	6.33	ddd (15.0; 11.0; 1.6)	-	2; 4	28; 1
4	128.8	5.87	d (10.6)	2; 6; 28	3	29; 6
5	132.7	-	-	28	-	-
6	130.9	5.63	s	4; 8a; 8b; 28; 29	-	8a; 8b; 4
7	135.2	-	-	8a; 8b; 29	-	-
8a	49.7	1.98	dd (13.2; 6.7)	6; 29; 30	9	6
8b		1.82	m	-	9	6; 30
9	28.9	1.59	m	8a; 8b; 10b; 30	8a; 8b; 10a; 10b; 30	30
10a	43.9	1.35	ddd (14.0; 11.6; 9.7)	8a; 8b; 12; 30; 31	9; 10b; 11	30
10b		1.09	ddd (14.0; 9.3; 5.0)	-	9; 10a; 11	30
11	35.5	2.42	m	14; 15; 31	10a; 10b; 12; 31	12; 30; 31
12	154.0	6.10	dd (15.2; 8.1)	13; 31	11; 13	11; 31
13	127.2	6.17	dd (15.2; 10.5)	-	12; 14	31
14	146.4	7.22	dd (15.5; 10.5)	13	13; 15	17
15	127.3	6.09	d (15.5)	-	14	-
16	196.0	-	-	13; 15; 17	-	-
17a	46.5	3.53	d (17.0)	-	-	14; 15; NH
17b		3.48	d (17.0)	-	-	14; 15; NH
18	166.9	-	-	17	-	-
19a	64.6	3.70	dd (11.2; 3.2)	21a; 21b	-	NH
19b		3.57	dd (11.2; 5.6)	-	-	NH
20	53.7	4.14	m	21a; 21b	19a; 19b; 21a; 21b; NH	-
21a	36.3	2.83	dd (14.0; 7.0)	23; 27	21b; 23; 27	23; 27
21b		2.76	dd (14.0; 7.5)	-	21a; 23; 27	23; 27; NH
22	129.6	-	-	21a; 21b; 23; 27	-	-
23	130.5	7.07	d (8.4)	21a; 21b; 27	24	-
24	115.6	6.75	d (8.4)	26	23	-
25	154.6	-	-	23; 27; 24; 26	-	-
26	115.6	6.75	d (8.4)	24	27	-
27	130.5	7.07	d (8.4)	21a; 21b; 23	26	-
28	17.6	1.85	s	4; 6; 29	-	3
29	18.4	1.73	d (1.1)	6; 8a; 8b	-	4
30	19.6	0.82	d (6.6)	8a; 8b	9	9; 10ab; 11
31	21.1	1.04	d (6.6)	19	11	12
NH	-	7.33	d (7.3)	-	-	19; 17; 21b
OH	-	5.20	broad	-	-	-
		2.65	broad	-	-	-



As explained in part 2.3 of this chapter, the reductive R domain located at the end of PKS-NRPS operates the reduction of thioester bound to release an aldehyde, which then undergo a Knoevenagel condensation to produce the pyrrolidone moiety. The formation of alcohol **14** likely finds its origin in the over-reduction by the R domain of the intermediate aldehyde **13**, as already exemplified in the myxochelin biosynthesis.<sup>210</sup>

Beside the production of the shunt product **14**, we observed that pyrrocidine production is not impaired in  $\Delta prcH$  mutant (Fig. 2.21). PCR analysis of the mutant and BLAST search in the genome showed the absence of another copy of *prcH* gene or a similar protein encoded in the genome able to substitute PrcH. Recently, Zhang *et al.* showed from synthetic aldehyde that Knoevenagel condensation can proceed spontaneously to afford pyrrolone moiety.<sup>211</sup> This suggests that aldehyde **13** is spontaneously converted to **15** and next to pyrrocidines by other auxiliary enzymes of the pathway and the fact that **13** does not accumulate and is not detected in the extract. Therefore, together these results indicate that PrcH is not required for the completion of the biosynthesis of pyrrocidines and its role is here conceivably to prevent the reduction of the PKS-NRPS off-loaded reactive aldehyde **13** in order to promote its ring closing into pyrrolidone **15** through a Knoevenagel condensation (Fig. 2.29).

The role of PrcH in the ring closing of the pyrrolidone need to be further investigated more particularly at the molecular level to identify the catalytic contribution of the protein in the Knoevenagel condensation. The first step has been initiated by cloning and expressing the protein in *E. coli*. Conditions for soluble protein production were defined, which will allow us to perform enzymatic assays with substrate surrogates and protein crystallization.

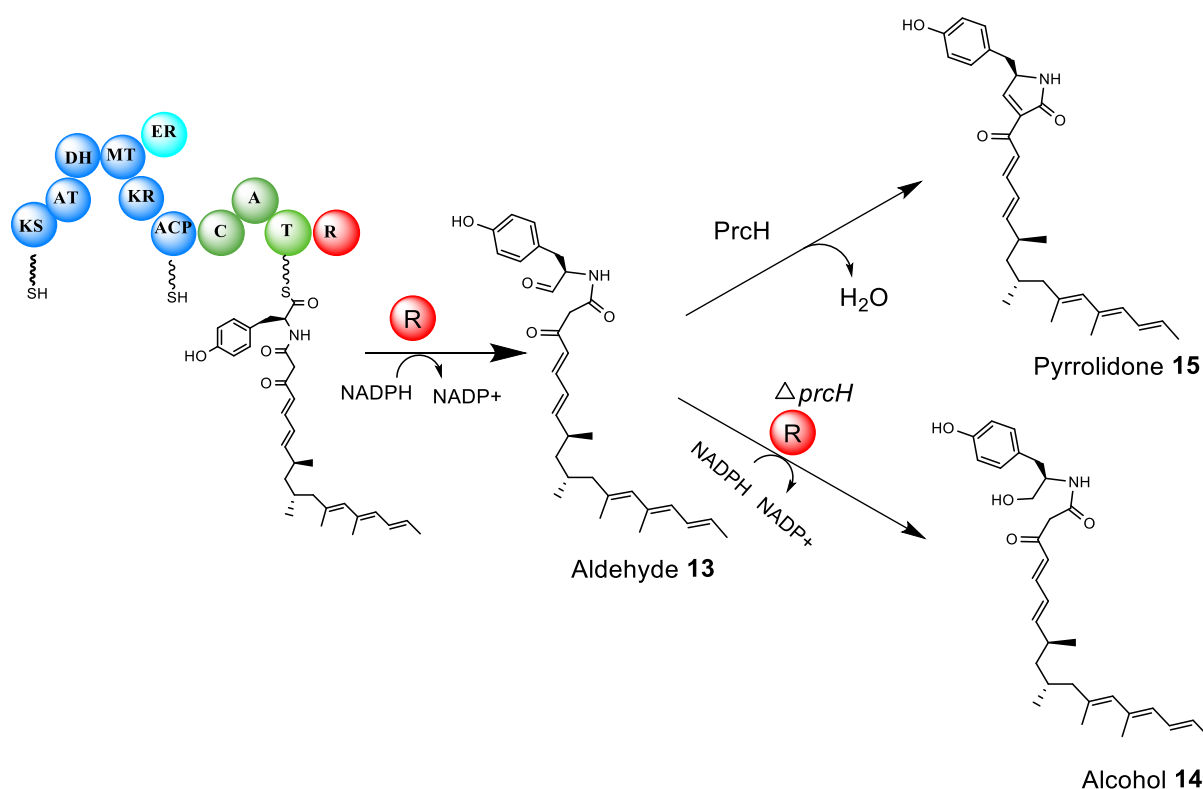


Fig. 2.29 Proposed function of PrcH: in  $\Delta prcH$  aldehyde is reduced to the alcohol very likely by R domain. In the presence of PrcH, the aldehyde is connected into pyrrolidone via a Knoevenagel condensation.

### 2.5.5 Inactivation of the genes *prcB*, *prcD* and *prcE*

Individual disruptions of the *prcB*, *D* and *E* genes all lead to annihilation of the production of pyrrocidines and to accumulations of new metabolites showing that these genes are directly involved in the biosynthesis.

In  $\Delta prcB$  mutant, we observed the formation of products **17** and **18** according to  $[M+H]^+$  ion of  $m/z$  506.289 eluted at 14.9 and 15.8 min, respectively (Fig. 2.30). Structure elucidation showed that **17** is a 5-hydroxy-3,4-epoxy-pyrrolidone substituted with a linear polyolefinic chain as in **14**. The *E*-configuration of the double bonds were deduced from the  $^3J$  coupling constants measured for ethylenic protons ( $^3J = 15.2$  Hz for C15=C14,  $^3J = 14.8$  Hz for C3=C2) and the NOESY spectrum showing the correlations H3-H28 for C4=C5 alkene, H8-H6 for C6=C7 bond and H13-H31 for C12=C13 bond.

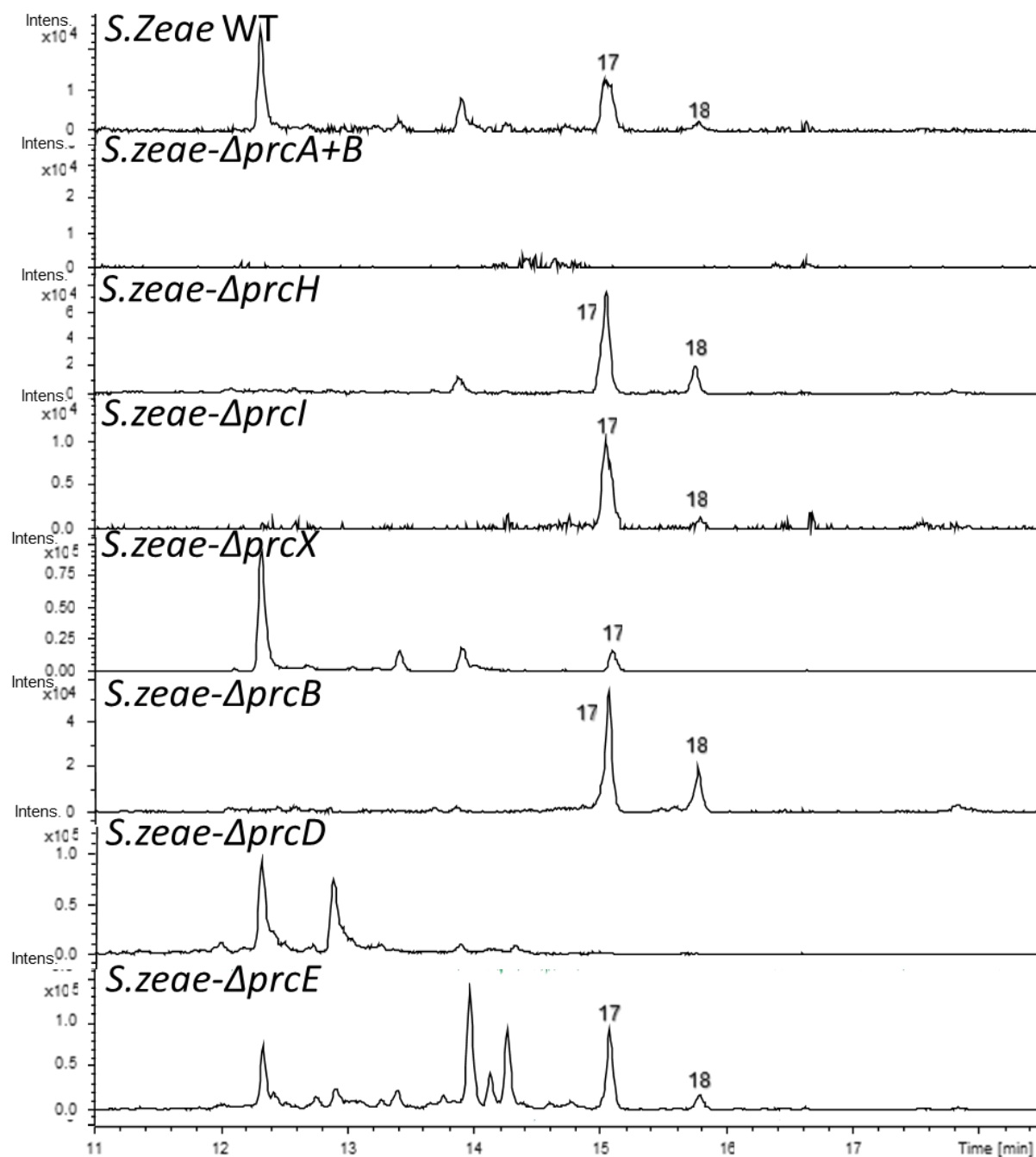


Fig. 2.30 The comparison of the EIC of  $[M+H]^+$  at  $m/z$  506.2901 showing the presence of the **17** and **18** in different mutants and wild type of *S.zeae*, noteworthy **17,18** are not present in *S.zeae-ΔprcE*.

Compound **18** was established to be a decaline from HMBC and COSY spectra with a *cis* junction and named pyrrocaline A. The relative stereochemistries of the bicycle substituents were obtained from NOESY data thanks to cross-correlations H12-H29-H10b-H14-H15-H4 showing that these protons are located on the same face of the decaline. The formation of the *cis*-decaline can be explained by an IMDA of the linear polyketide **17** through an *exo*-selective approach (Fig. 2.31).

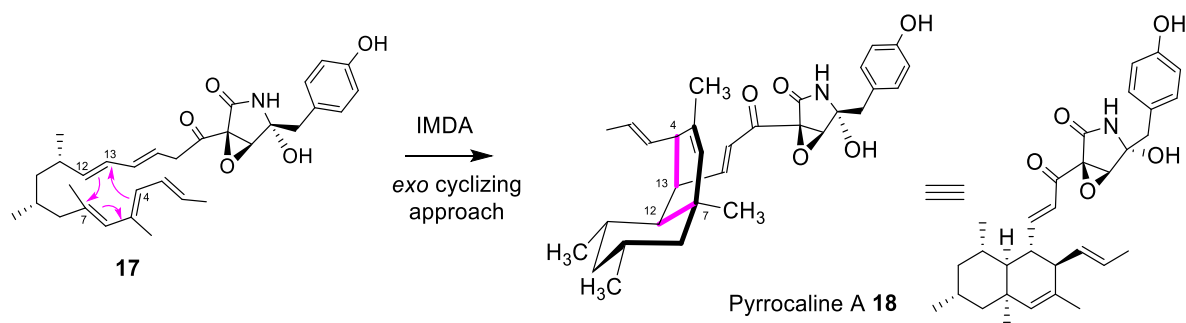


Fig. 2.31: Proposed mechanism for the formation of *cis*-decalin **18** from linear polyketide **17**.

In  $\Delta prcE$  mutant, we detected a peak by LC-MS at 16.0 min, with  $m/z$  476.3146 for  $[M+H]^+$  corresponding to the molecular formula  $C_{31}H_{41}NO_3$  (Fig. 2.21 and annex 14). NMR analysis revealed a mixture of two compounds **23a,b** in a 1:1 ratio. The structure elucidation from HMBC and COSY spectra shows the presence of a trimethyl-cyclohexane corresponding to ring A of pyrrocidines substituted in C7 and C12 by olefins. Unlike to the polyketides **14** and **17** having a trienic system with a methyl at C1, **23a,b** show three isomerized double bounds from C6 until C1 positions (Fig. 2.33). The other substituent at C12 is C13-C14 olefin with either E or Z configuration (measured coupling constants for C13=C14: **23a**  $^3J_E = 15.3$  Hz and **23b**  $^3J_Z = 9.3$  Hz). The *trans* relationship of C7 and C12 substituents were deduced from the key cross-correlations H29-H13-H11-H9 and H12-H6-H14 in NOESY spectrum.

The same mutant also produces compound **20** detected at 15.8 min (Fig. 2.21) with a molecular formula of  $C_{31}H_{39}NO_3$  (assigned from  $[M+H]^+$   $m/z$  474.2998). NMR analysis revealed a unique 6/10/6/5 fused polycyclic structure substituted by a tyrosine residue. This macrocycle fused to an amino acid-derived octahydro-isoindole belongs to cytochalasans.<sup>212</sup> It was named pyrrochalin A and represents the first nonaketide-tyrosine member of this family. Relative stereochemistry was assigned by NOESY experiment, on the base of two identified groups of cross-correlations H30-H10b-H31-H12-H14-H28 and H29-H6-H4-H19-H23 indicating that these groups of protons are respectively on each side of equatorial plan of the molecule (Table 2.20).

From LC-MS chromatograms of  $\Delta prcE$  and  $\Delta prcD$  mutants, we identified two other accumulating compounds named pyrrocyclines A **21** and B **22** detected at 13.2 and 13.0 min (Fig. 2.32), respectively, with the same molecular formula  $C_{31}H_{39}NO_4$  (assigned from  $[M+H]^+$  molecular ion  $m/z$  490.295).

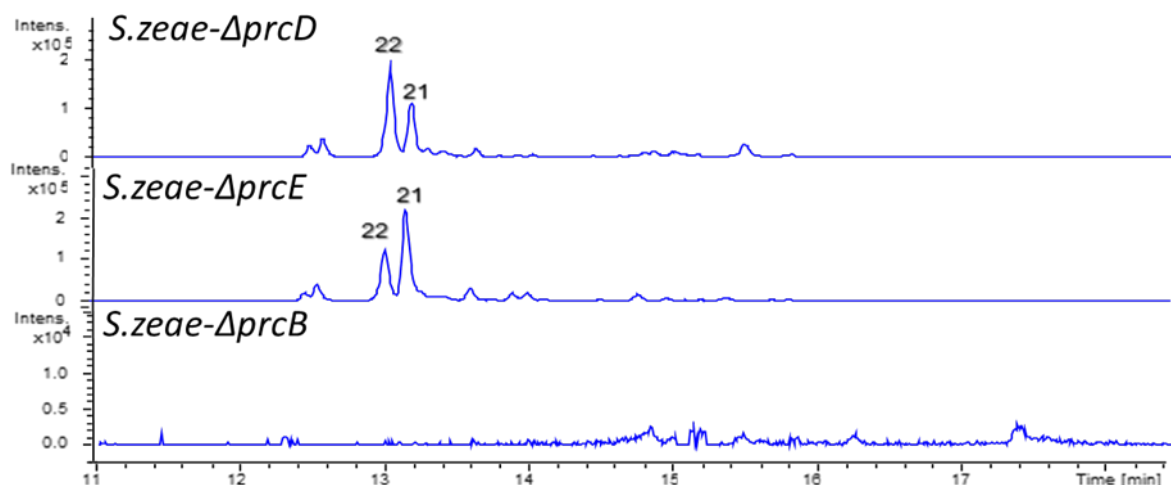


Fig. 2.32 The comparison of the content of compounds **21** and **22** in different mutants.

Their planar structures were established showing a unique backbone consisting of a central 15-membered macrocycle fused to 6- and 5-membered rings. NOESY experiments performed on **21** and **22** indicates that they adopt a conformation in which protons H14, H12, H6, H8, H4, H2 and NH are oriented on the same face of the equatorial plan of the molecule while H23/27, H21, H19, H1, H3, H28, H29, H13, H11 are located on the opposite face. (Table 2.17 and 2.18)

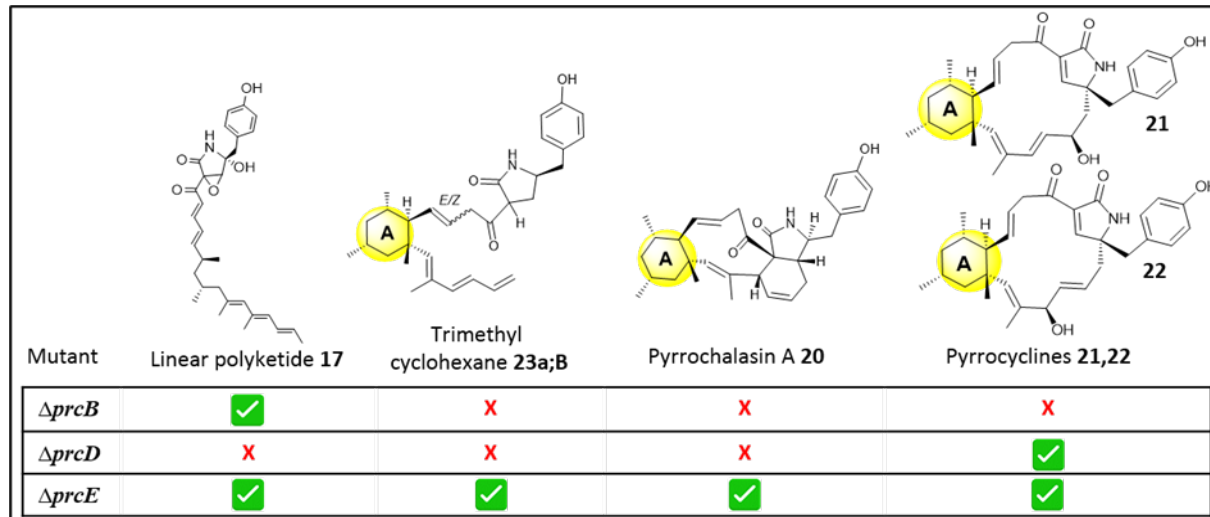


Fig. 2.33 Metabolites isolated from mutants  $\Delta prcB$ ,  $\Delta prcD$  and  $\Delta prcE$ .

In both mutant  $\Delta prcE$  or  $\Delta prcD$  accumulating metabolites with the trimethyl cyclohexane motif corresponding to the ring A of pyrrocidines were isolated (Fig. 2.33). By contrast, this cyclohexane is not observed among the accumulating metabolites in the  $\Delta prcB$  mutant. We deduced that PrcB installs ring A with an isomerization-cyclization step from pyrrolidone **15** to the key intermediate **19**. The existence of this intermediate can explain the formation of the

isolated metabolites in  $\Delta prcE$  or  $\Delta prcD$ . Indeed **19** gives rise to **20** through an *endo*-selective IMDA as in the biosynthesis of cytochalasans. Such a stereocontrolled cyclization required a pericyclase as demonstrated by Hankte *et al.* in pyrrochalin H biosynthesis<sup>212</sup> suggesting that the formation of **20** is not spontaneous. This reaction may be governed by a protein like PrcX which seems to have the capacity to bind a polyenic chain. Intermediate **19** can also provide the macrocycles **21** and **22** by attack of enolised C20 on activated C1 position via epoxidation of C1-C2 or C3-C4 olefins (Fig. 2. 34).

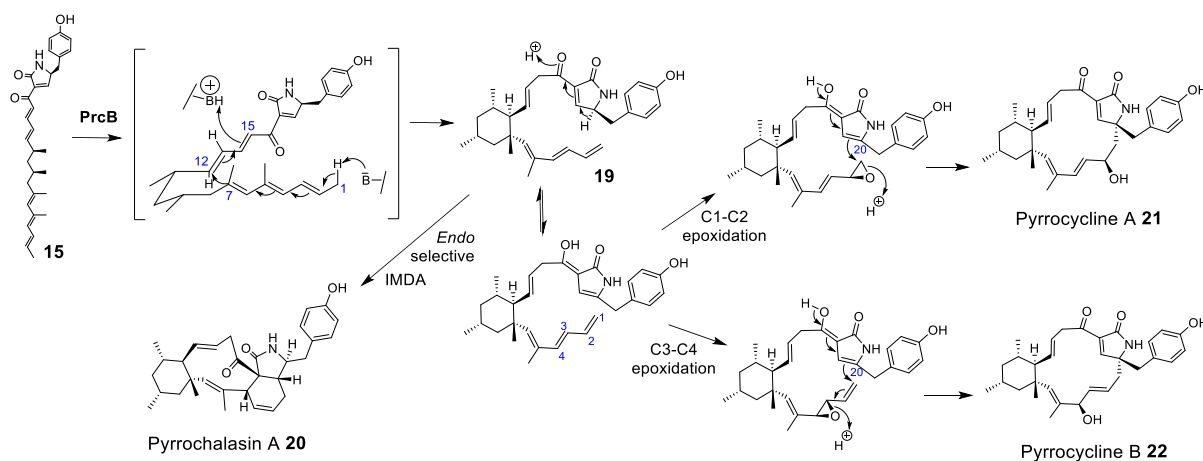


Fig. 2.34 Proposed pathway for the formation of polycyclic compounds **20**, **21** and **22** from the shared intermediate **19**.

Comparison of LC-MS analysis of all the mutants shows that **17** and **18** are not produced when *prcD* is inactivated (Fig. 2.30). This indicates that PrcD is involved in the oxidation of pyrrolidone moiety of **15**. Autooxidation of strained olefins has been reported to give epoxides<sup>213</sup> and this was also exemplified by the known conversion of **1** to **9**<sup>214,215</sup>. It can be deduced that PrcD catalyses the hydroxylation of **15** on C20 to give **16** which is then converted into **17** by spontaneous olefin epoxidation (Fig. 2.35).

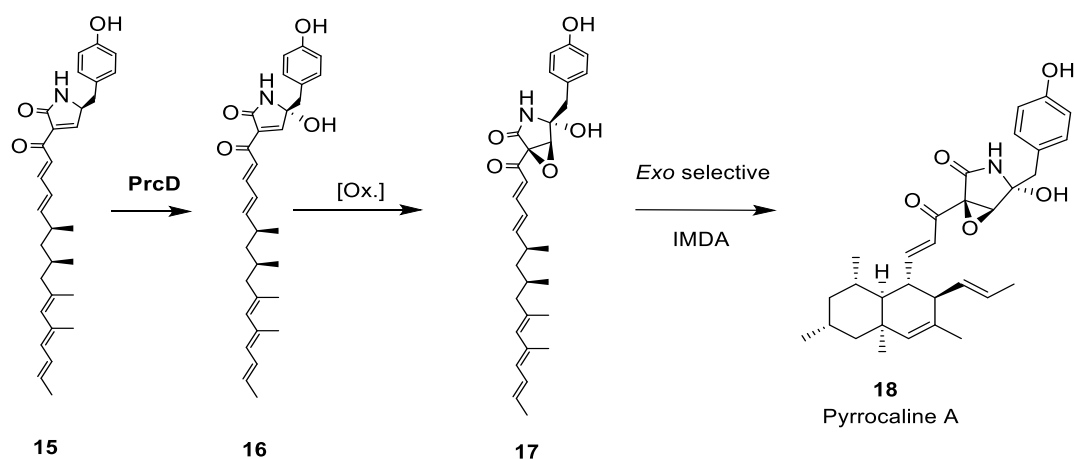


Fig. 2.35 Proposed pathway for the formation of **17** and **18**

The isolation of cyclohexanes **20**, **21** and **22** implies that **15** is substrate of PrcB to afford cyclohexane **19** and then PrcD-catalysed C20 oxidation occurs to give **24**. It cannot be excluded as well that hydroxylated **16** is substrate of PrcB to give **24**. The latter in  $\Delta prcE$  mutant can give **23a,b** after a reduction-dehydration sequence (Fig. 2.36). Combining all these results, we deduced that PrcE build the paracyclophane **25** from **24**, which next undergoes an *exo* IMDA assisted by PrcX to afford **1**.

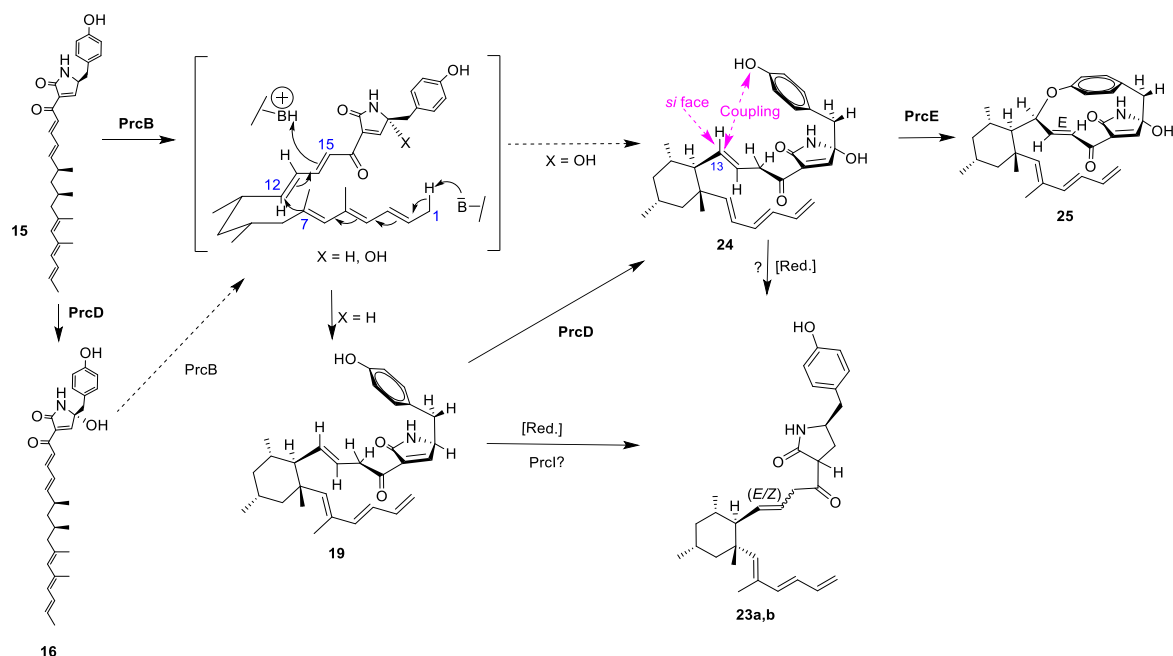
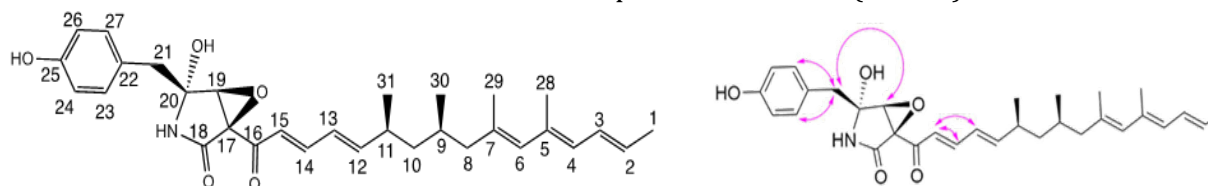


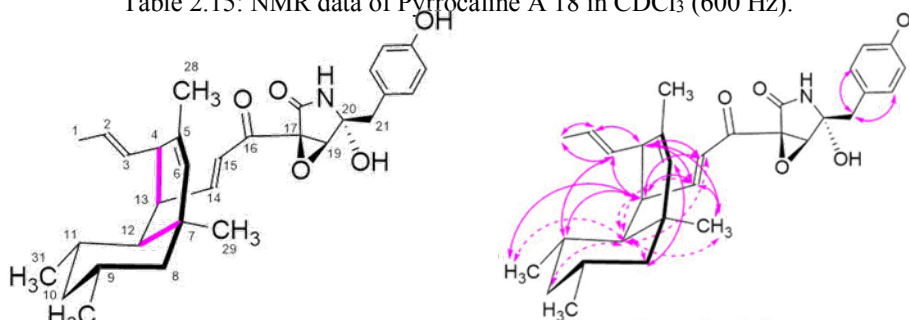
Fig. 2.36 Proposed pathway for the formation of **23 a,b**

Table 2.14: NMR data of compound **17** in CDCl<sub>3</sub> (600 Hz).

## Key NOESY correlations

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1	18.7	1.80	d (6.7)	2	2; 3	
2	128.5	5.66	dq (14.8; 7.0)	1; 4	1; 3	
3	128.3	6.32	ddd (14.8; 10.0; 1.6)	-	1; 2; 4	28
4	128.8	5.87	d (11.1)	2; 28	3; 6; 28	
5	132.7	-	-	28	-	
6	130.9	5.64	s	4; 8a; 8b; 28; 29	4; 28; 29	8a,b
7	135.2	-	-	8a; 8b; 29	-	
8a	49.7	1.97	dd (13.0; 6.3)	6; 29; 30	9	6
8b		1.83	dd (13.0; 7.9)		9	6
9	29.0	1.59	m	8a; 8b	8a; 8b; 10a; 10b; 30	
10a	44.0	1.34	ddd (14.0; 9.4; 4.6)	8a; 8b; 30; 31	9; 10b; 11	
10b		1.09	ddd (14.0; 9.4; 5.1)		9; 10a; 11	
11	35.5	2.42	m	12; 31	10a; 10b; 12; 31	
12	155.3	6.15	m		11; 13	
13	127.4	6.15	m	15	14; 15	31
14	147.1	7.36	dd (15.2; 10.0)	12; 13	12; 13	15
15	122.1	6.12	d (15.2)	-	14	13
16	187.4	-	-	14	-	
17	63.2	-	-	-	-	
18	166.8	-	-	-	-	
19	63.4	3.99	d (2.3)	21a; 21b	-	21a,b
20	84.6	-	-	19; 21a; 21b	19a; 19b; 21a; 21b; NH	
21a	42.6	3.11	d (14.1)	-	21b; 23; 27	19; 23; 27
21b		3.00	d (14.1)		21a; 23; 27	19; 23; 27
22	124.8	-	-	24; 26	-	
23	132.0	7.12	d (8.3)	21a; 21b; 27	24	21a,b
24	115.8	6.77	d (8.3)	26	23	
25	155.4	-	-	23; 27; 24; 26	-	
26	115.8	6.77	d (8.3)	24	27	
27	132.0	7.12	d (8.3)	21a; 21b, 23	26	21a,b
28	17.6	1.85	s	4; 6	-	3
29	18.4	1.73	d (1.0)	6; 8a; 8b	-	
30	19.6	0.83	d (6.6)	8a; 8b	9	
31	21.0	1.03	d (6.7)	-	11	13
NH		6.77	broad			
OH		3.19	broad			

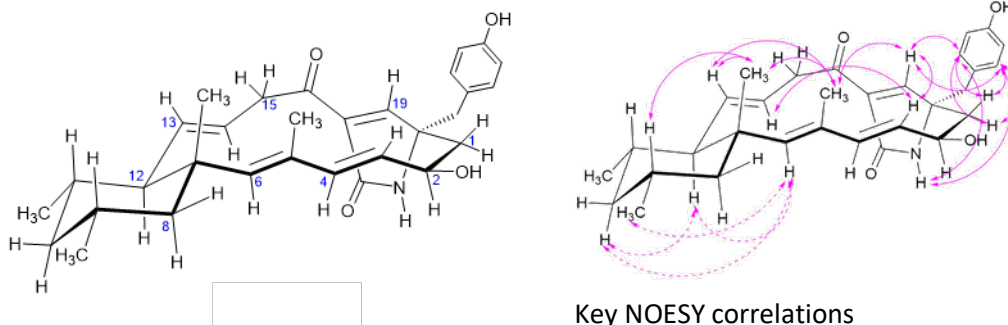


Table 2.15: NMR data of Pyrrocaline A 18 in CDCl<sub>3</sub> (600 Hz).

Key NOESY correlations

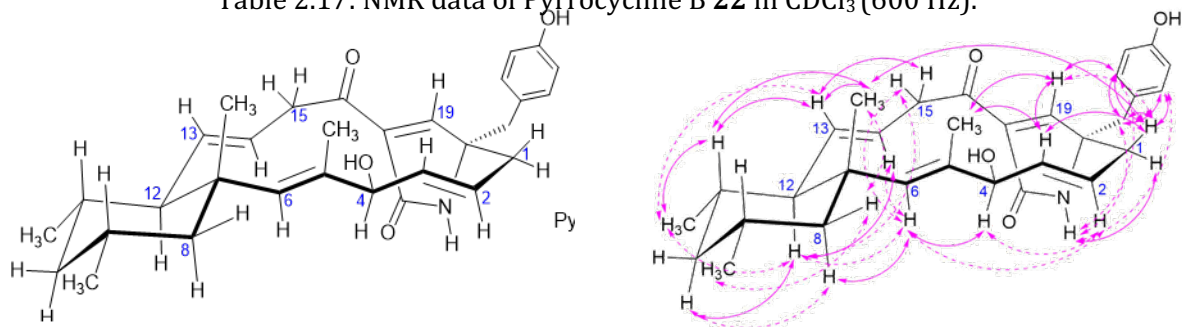
position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1	17.9	1.66	dd (6.4; 1.5)	2; 3	2; 3	
2	127.0	5.45	dq (14.9; 6.4)	1; 3; 4; 13	1; 3	1; 4
3	134.1	5.30	ddd (14.9; 9.3; 3.2; 1.6)	1; 2; 4; 13	1; 2; 4	1; 11; 13
4	45.3	2.63	d (9.2)	2; 3; 6; 12; 13; 14; 28	3; 6; 13; 28	2; 14; 15; 29
5	132.0	-	-	4	-	
6	132.4	5.20	s	8b; 12; 29	4; 12; 28	8a; 29
7	36.9	-	-	6; 12; 13	-	
8a	51.1	1.44	dd (12.7; 2.7) eq.	6; 29; 30	8b; 10a	6
8b		0.88	d (12.7) ax.		8a	
9	28.9	1.40	m	8b; 10b; 30	30	
10a	44.8	1.56	ddd (12.7; 6.2; 3.1) eq.	8a; 30; 31	8a; 10b	
10b		0.62	ddd (12.7) ax.		10a	
11	33.4	1.50	m	10b; 12; 13; 31	12; 31	3
12	53.9	1.12	d (10.7)	6; 8a; 13; 14; 29; 31	6; 11; 13	8b; 10b; 15; 29; 31
13	45.0	2.76	d (6.5)	4; 15	4; 12; 15	3; 11; 12; 15; 31
14	161.7	7.37	dd (15.8; 6.8)	4; 12; 13	15	4; 12; 13; 29
15	121.5	6.36	dd (15.8; 1.6)	13	13; 14	4; 12; 13
16	189.0	-	-	14; 15	-	
17	61.6	-	-	NH	-	
18	168.0	-	-	19	-	
19	64.2	3.95	d (2.4)	21b	NH	
20	84.9	-	-	19; 21a; 21b	-	
21a	40.5	3.11	d (13.8)	23; 27	21b	23; 27
21b		2.93	d (13.8)		21a	23; 27
22	126.0	-	-	21a; 21b	-	
23	131.7	7.18	d (8.3)	21a; 21b; 27	24	21a,b
24	115.9	6.81	d (8.3)	26	23	
25	155.4	-	-	23; 27; 24; 26	-	
26	115.9	6.81	d (8.3)	24	27	
27	131.7	7.18	d (8.3)	21a; 21b, 23	26	21a,b
28	22.7	1.62	m	6	4; 6	
29	32.1	0.85	s	8b; 28	-	4; 6; 12; 14
30	22.6	0.82	d (6.4)	8b	9	
31	21.1	0.82	d (6.3)	10b; 12	11	10b; 12
NH		5.97	broad		19	
OH		5.18	Broad			
OH		3.75	broad			

Table 2.16: NMR data of Pyrrocycline A **21** in CDCl<sub>3</sub> (600 Hz).



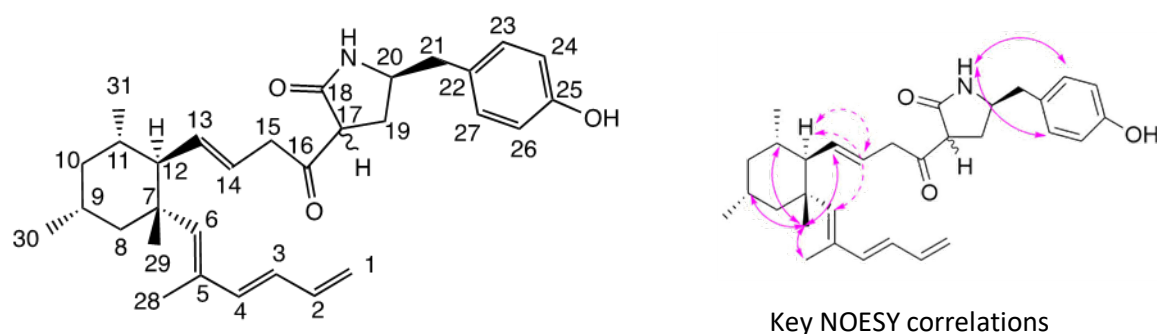
position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1a	41.9	2.33	dd (13.5; 3.2)	3; 21a; 21b	1b; 2	23; 27
1b		2.00	dd (13.5; 10.7)		1a; 2	23; 27
2	71.0	4.07	m	1a; 1b; 4	1a; 1b; 3	
3	125.8	5.01	dd (15.3; 9.2)	-	2; 4	14; 19
4	143.1	5.85	d (15.3)	6; 28	3	
5	129.7	-	-	3; 4; 6; 28	-	
6	148.4	4.93	s	4; 12; 28; 29	11	12; 30
7	43.5	-	-	12; 29	-	
8a	48	1.49	d (12.3)	6; 29; 30	8b; 10a	
8b		0.86	m		8a; 29	
9	26.6	1.63	m	10a; 30	30	29
10a	44.2	1.67	m	8a; 30; 31	8a	
10b		0.53	d(14.2)			12
11	30.6	1.57	m	10b; 12; 13; 31	6; 31	
12	59.7	1.30	dd (10.0; 10.0)	6; 8a; 10a; 10b; 13; 14; 29; 31	13	6; 10b
13	135.0	5.22	dd (15.3; 10.0)	12; 15a; 15b	12; 14	28; 29
14	123.6	5.49	ddd (15.4; 9.7; 5.9)	12; 15a; 15b	13; 15a; 15b	3
15a	46.8	3.84	dd (10.3; 7.7)	13; 14	14; 15b	
15b		2.89	dd (10.1; 10.1)		14; 15a	
16	191.5	-	-	15a; 15b; 19	-	
17	134.2	-	-	15b	-	
18	169.6	-	-	19, 21a; 21b	-	
19	164.1	7.78	d (1.7)	-	-	3; 23; 27; 28
20	63.4	-	-	1a, 1b; 19; 21a; 21b	-	
21a	46.6	2.81	d (13.4)	23; 27	21b	
21b		2.70	d (13.4)		21a	
22	126.1	-	-	1a, 1b; 21a; 21b 24; 26	-	
23	131.9	6.95	d (8.0)	21a; 21b; 27	24	1a; 1b; 19; NH
24	115.8	6.73	d (8.0)	23; 26	23	
25	155.6	-	-	23; 27; 24; 26	-	
26	115.8	6.73	d (8.0)	24; 27	27	
27	131.9	6.95	d (8.0)	21a; 21b, 23	26	1a; 1b; 19; NH
28	12.9	1.54	s	4; 6	-	13; 19; 29
29	18.9	1.11	s	6; 8a; 12	8b	9; 13; 28
30	22.8	0.81	d (6.4)	8b; 10b	9	6
31	21.8	0.79	d (6.4)	12	11	
NH						23; 27
OH						

Table 2.17: NMR data of Pyrrocycline B **22** in CDCl<sub>3</sub> (600 Hz).

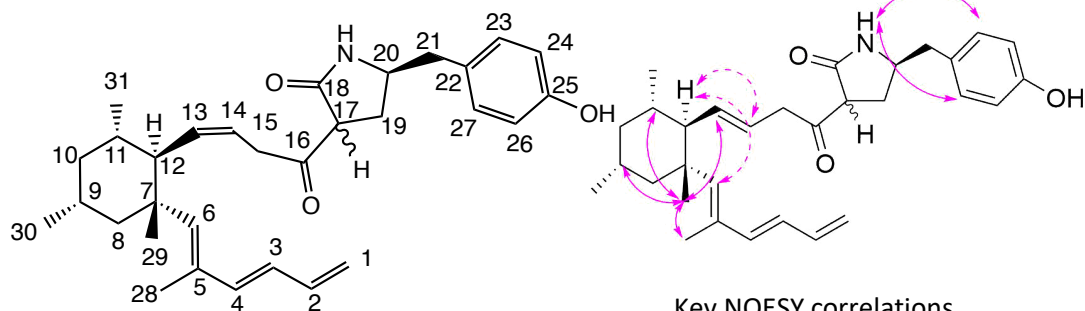


Key NOESY correlations

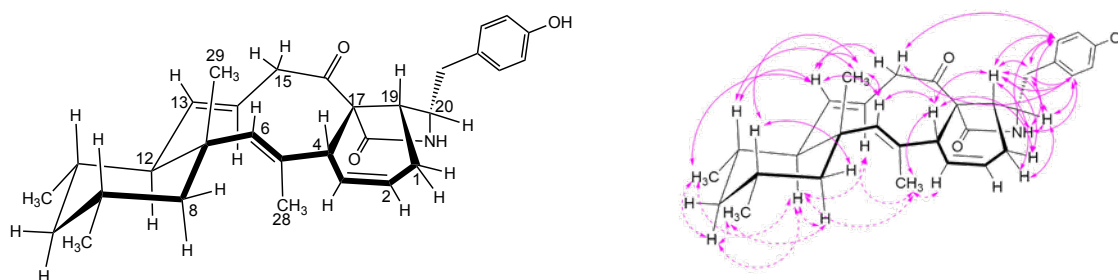
position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1	38.4	2.43	m	2; 3; 21a; 21b	2	3; 21ab; 23; 27; 29; NH
2	124.3	5.25	m	1	1; 3	4
3	136.8	5.54	dd (15.4; 7.2)	1; 4	2; 4	1; 19; 28
4	82.0	4.24	d (7.2)	2; 6; 31	3	2; 6
5	133.5	-	-	3; 4; 6; 31	-	
6	140.9	5.09	s	4; 8b; 12; 28; 29	-	2; 4; 8b; 12; 14
7	41.0	-	-	6; 8b; 12; 28; 29	-	
8a	48.9	1.53	m	9; 10b; 29; 30	8b	29; 30
8b		0.84	d (13.0)		8a	6; 10b
9	27.1	1.56	m	-	30	
10a	44.0	1.64	m	8a; 30; 31	10b	
10b		0.53	ddd (12.4)		10a	8b; 12
11	31.5	1.45	m	10b; 13; 31	12; 31	13; 29; 31
12	57.6	1.32	dd (10.2; 10.2)	6; 8a; 14; 29; 31	11; 13	6; 10b; 14; 31
13	137.9	5.27	m	6; 12; 14; 15a; 15b	12; 14	11; 15b; 29; 31
14	121.6	5.46	ddd (15.3; 7.7; 6.8)	12; 15a; 15b	13; 15a; 15b	6, 12
15a	45.4	4.02	dd (10.6; 6.6)		14; 15b	
15b		2.98	dd (11.7; 7.9)		14; 15a	
16	192.7	-	-	-	-	
17	135.7	-	-	15b; 19; NH	-	
18	169.3	-	-	19	-	
19	161.5	7.69	d (2.0)	1; 14; 15a; 15b; 19; 21a; 21b	-	1; 3; 23; 28
20	65.3	-	-	1; 19; 21a; 21b; NH	-	
21a	43.6	2.89	d (13.6)	23; 27	-	1; 23; 27; NH
21b		2.84	d (13.6)			
22	126.4	-	-	21a; 21b 24; 26	-	
23	131.5	6.97	d (8.4)	27	24	1; 21; NH
24	115.7	6.74	d (8.4)	23; 26	23	
25	155.4	-	-	23; 27; 24; 26	-	
26	115.7	6.74	d (8.4)	24; 27	27	
27	131.5	6.97	d (8.4)	23	26	1; 21; NH
28	11.6	1.57	s	4; 6	-	3; 19
29	19.6	1.10	s	6; 12		8a; 11; 13
30	22.8	0.81	d (6.3)	8b	9	8a
31	21.6	0.76	d (6.4)	10b; 12	11	12; 13
NH		6.18	s (broad)		19	1; 21; 23; 27
OH		5.87	s (broad)			

Table 2.18: NMR data of compound **23a** in CDCl<sub>3</sub> (600 Hz).

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1a	115.7	5.16	d (17.3)	3	1b; 2	
1b		5.10	d (10.0)		1a; 2	
2	137.8	6.35	ddd (16.9; 10.0; 10.0)	1a; 3; 4	1a; 1b; 3	
3	126.3	6.15	m	1a; 1b; 2; 4	2	
4	140.7	6.15	dd (15.5; 4.0)	6; 22	-	
5	132.3	-	-	4; 6	-	
6	146.2	5.47	s	4; 8b; 12; 28; 29	28	8b; 12
7	40.6	-	-	6; 12; 29	-	
8a	44.4	1.69	m	6; 29; 30	-	
8b		1.04	dd (13.3)		-	6
9	27.4	1.62	m	8b; 30	30	29
10a	44.1	1.68	m	8a; 30; 31	10b	
10b		0.59	ddd (12.3)		10a	
11	31.0	1.54	m	10b; 12; 31	31	29
12	58.0	1.60	dd (9.9)	6; 14; 29; 31	13	6; 14
13	136.9	5.30	dd (15.3; 9.5)	12; 15a; 15b	12; 14	29
14	123.9	5.57	ddd (15.3; 7.2; 7.2)	12; 15a; 15b	13; 15a; 15b	12
15a	46.9	3.48	m	13; 14	-	
15b		3.40	m			
16	204.3	-	-	15a; 15b, 19a; 20	-	
17	53.1	3.70	dd (9.1; 8.2)	19a; 19b; NH	19a; 19b	
18	172.8	-	-	19a; 19b; 20; 21a	-	
19a	28.2	2.30	ddd (13.4; 8.1; 6.6)	17; 21a; 21b; NH	17; 19b	
19b		2.14	ddd (13.4; 9.3; 7.5)		17; 19a	
20	54.2	3.68	m	17; NH	21a; 21b	
21a	42.4	2.80	dd (13.6; 7.9)	19b; 23; 27	20; 21b	
21b		2.58	dd (13.6; 9.0)		20; 21a	
22	129.3	-	-	21a; 21b; 24; 26	-	
23	130.2	6.99	d (8.4)	21a; 21b; 27	27	NH
24	115.9	6.75	d (8.4)	23	26	
25	155.2	-	-	23; 27; 24; 26	-	
26	115.9	6.75	d (8.4)	27	24	
27	130.2	6.99	d (8.4)	21a; 21b	23	NH
28	14.2	1.82	d (1.1)	4; 6	6	29
29	20.8	1.15	s	6; 8b; 12	-	9; 11; 13; 28
30	22.8	0.84	d (6.2)	8b	9	
31	21.6	0.80	d (6.2)	10b	11	
NH		7.75	broad		-	23; 27

Table 2.19: NMR data of compound **23b** in CDCl<sub>3</sub> (600 Hz).

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	Key NOESY correlations		
				HMBC	COSY	NOESY
1a	115.7	5.13	d (17.3)	3	1b; 2	
1b		4.97	d (10.1)		1a; 2	
2	137.7	6.35	ddd (16.9; 10.0; 10.0)	1a; 3; 4	1a; 1b; 3	
3	126.2	6.15	m	1a; 1b; 2; 4	2	
4	140.7	6.15	dd (15.5; 4.0)	2; 6; 22	-	
5	132.2	-	-	4; 6	-	
6	146.1	5.49	s	4; 8b; 12; 28; 29	28	8b; 12
7	40.3	-	-	6	-	
8a	47.3	1.69	m	6; 29; 30	8b	
8b		1.01	dd (13.3)		8a	6
9	27.5	1.62	m	8b; 30	30	29
10a	44.1	1.62	m	8a; 30; 31	10b	
10b		0.59	ddd (12.3)		10a	
11	30.9	1.54	m	10b; 31	31	29
12	58.3	1.54	dd (9.9)	6; 14; 29; 31	13	6; 14
13	137.3	5.39	m	12; 15a; 15b	12	29
14	123.5	5.39	m	12; 15a; 15b	15a; 15b	12
15a	46.9	3.45	m	13; 14	14	
15b		3.38	m			
16	204.1	-	-	15a; 15b, 19a; 19b; 20	-	
17	53.2	3.65	dd (9.1; 5.0)	19a; 19b; NH	19b	
18	172.8	-	-	19b; 20; 21a	-	
19a	29.0	2.68	ddd (12.8; 7.6; 5.0)	21a; 21b; NH	17; 19b	
19b		1.76	ddd (13.3; 7.2; 5.7)		17; 19a; 20	
20	54.7	3.85	m	17; 21a; NH	19a; 21a; 21b	
21a	42.4	2.78	dd (13.6; 7.3)	19a; 19b; 23; 27	20; 21b	
21b		2.47	dd (13.6; 9.0)		20; 21a	
22	128.9	-	-	21a; 21b; 24; 26	-	
23	130.2	6.97	d (8.4)	21a; 21b; 24; 27	27	NH
24	115.8	6.75	d (8.4)	23; 26	-	
25	155.1	-	-	23; 27; 24; 26	-	
26	115.8	6.75	d (8.4)	24; 27	-	
27	130.2	6.97	d (8.4)	21a; 21b; 23	23	NH
28	14.1	1.84	d (1.1)	4; 6	6	29
29	20.8	1.19	s	6; 8b	-	9; 11; 13; 28
30	22.8	0.84	d (6.2)	8b	9	
31	21.5	0.80	d (6.2)	10b	11	
NH		5.77	broad		-	

Table 2.20: NMR data of Pyrrochalin **20** in CDCl<sub>3</sub> (600 Hz).

## Key NOESY correlations

position	$\delta$ C (ppm)	$\delta$ H (ppm)	Mult. ( <i>J</i> in Hz)	HMBC	COSY	NOESY
1a	30.5	2.29	m	2; 3; 19; 20	16; 19	4; 19
1b		2.17	dd (6.4; 14.2)	-	2	20
2	127.4	5.97	ddd (12.4; 6.4; 3.0)	1a; 1b; 19	1b; 3	
3	131.1	6.10	ddd (9.15; 3.7; 3.7)	1a; 1b	2; 4	28
4	55.2	3.11	m	3	3	1a; 19; 28
5	132.3	-	-	6; 28; 29	-	
6	143.6	4.70	s	28; 29	28	4; 13; 29
7	44.8	-	-	8a; 8b; 12; 13; 14; 28; 30	-	
8a	44.5	1.64	dd (12.3)	6; 29	8b	30; 12
8b		1.46	dd (12.3; 2.7)	-	8a	9
9	27.8	1.76	m	10b; 30	30	29; 9
10a	44.9	1.73	d (12.3)	31	10b	
10b		0.60	ddd (12.3)	-	10a	12
11	29.7	1.71	m	10b; 12; 13; 31	31	29; 13
12	57.7	2.16	dd (10.7)	8a; 8b; 10a; 13; 14; 28; 29; 31	13	8a; 10b; 14; 31; 28
13	137.5	5.34	dd (15.7; 10.6)	12; 15a; 15b	12; 14	6; 11; 15b; 29; 31
14	124.5	5.92	ddd (16.0; 10.4; 6.1)	12; 15a; 15b	13; 15a; 15b	28; 12
15a	47.9	3.22	dd (9.7; 6.0)	14	14; 15b	23; 27
15b		2.91	dd (10.2)	-	14; 15a	13
16	207.2	-	-	1; 13; 15a; 15b; 19	-	
17	71.6	-	-	1b; 4; 15b; 16; 19; 28; NH	-	
18	175.9	-	-	4; 19; 20	-	
19	47.7	2.85	d (7.4)	20; NH	1a	1a,b; 4; 20; 21a,b; 23; 27
20	61.5	3.15	m	1a; 1b; 19; 21a; 21b; NH	21a; 21b	1b; 19
21a	42.6	2.70	dd (13.8; 5.2)	19; 23; 27	20; 21b	19
21b		2.42	dd (13.8; 7.9)	-	20; 21a	19
22	128.3	-	-	20; 21a; 21b	-	
23	130.7	6.94	d (8.4)	21a; 21b; 27	24	15a; 19; 20; 21a, 21b
24	115.8	6.75	d (8.4)	23; 26	23	
25	155.1	-	-	23; 27; 24; 26	-	
26	115.8	6.75	d (8.4)	24; 27	27	
27	130.7	6.94	d (8.4)	21a; 21b, 23	26	15a; 19; 20; 21a,b
28	17.4	2.01	d (1.0)	4; 6	6	4; 14
29	24.4	0.93	s	6	-	6; 9; 11; 13
30	23.1	0.85	d (6.3)	8a; 10b	9	8a; 10b
31	21.8	0.80	d (6.2)	10b; 12	11	12; 13; 10b
NH		5.81	s	-	-	20; 21a,b; 23; 27
OH		2.17	s	-	-	

## 2.6 Proposed biosynthetic pathway of pyrrocidines

Based on the gene inactivation study, the following biosynthetic pathway of pyrrocidines was proposed (Fig. 2.37). Using one acetyl-CoA, eight malonyl-CoA, four S-adenosyl-methionines and one tyrosine as substrate, PKS-NRPS (PrcA) and its associated ER (PrcC) assemble the linear aldehyde **13**. The latter then undergoes a Knoevenagel condensation to give the pyrrolidone intermediate **15** under the protection or catalysis of PrcH. Ring A of pyrrocidines is installed by PrcB with a concomitant rearrangement of the double bonds affording **19**. Hydroxylation of pyrrolidone ring is then performed by the integral membrane protein PrcD to give **24**. An alternative pathway is also conceivable by reversing the order between PrcB and PrcD. In this case **15** is hydroxylated into **16** on which ring A is introduced to give **24**. The second integral membrane protein (PrcE) catalyzes the formation of the paracyclophane. Finally, intramolecular Diels-Alder reaction with *exo* approach in the presence of PrcX forms rings B and C to give pyrrocidine A **1** which is then reduced into **2** by PrcI.

After we established this pathway, two concurrent studies on the biosynthesis of pyrrocidine family were published in the literature.

In the first one, the authors sequenced the genome of *Sarocladium zae* and reconstituted the biosynthesis of pyrrocidine B **2** by heterologous expression in the fungus *Aspergillus nidulans*.<sup>216</sup> For that, the genes of the cluster were introduced in this host and expressed. Diverse combinations of the auxiliary enzymes were tested. By comparing the metabolite production profiles in the reconstitution experiments with or without PrcX, the authors showed that this protein controls the stereochemistry of IMDA in favor of the *exo*-approach to give **2**. This is in full agreement with our results obtained by gene inactivation in the naturally producing strain. However, for other auxiliary enzymes, the group could not identify any function for PrcI, PrcB, PrcD and PrcE, despite the analysis of diverse genetic constructions expressing these proteins. Nevertheless, a mechanism was proposed in which the four proteins function synergistically to form the paracyclophane intermediate I from the all-*E* polyene **15** (Fig. 2.38). In this mechanism, the latter is configured in an inverse S-shape positioning the C13 closed to phenol oxygen while the hydrocarbon chain from C7 to C12 is conformed as a pre-chair motif. The cyclisation cascade is triggered by deprotonation of phenate which attacks olefinic C13 and subsequently generates the attack of C12 on C7 to form ring A and initiate the displacement of three conjugated double bonds up to C1. Consequently, a hydride acceptor is required at C1 position which was proposed to be the NAD(P)<sup>+</sup> cofactor of PrcI. However, this hypothesis is ruled out

considering our results demonstrating that *prcI* deletion does not affect the production of pyrrocidine A **1** and it is involved in the conversion of **1** to **2**.

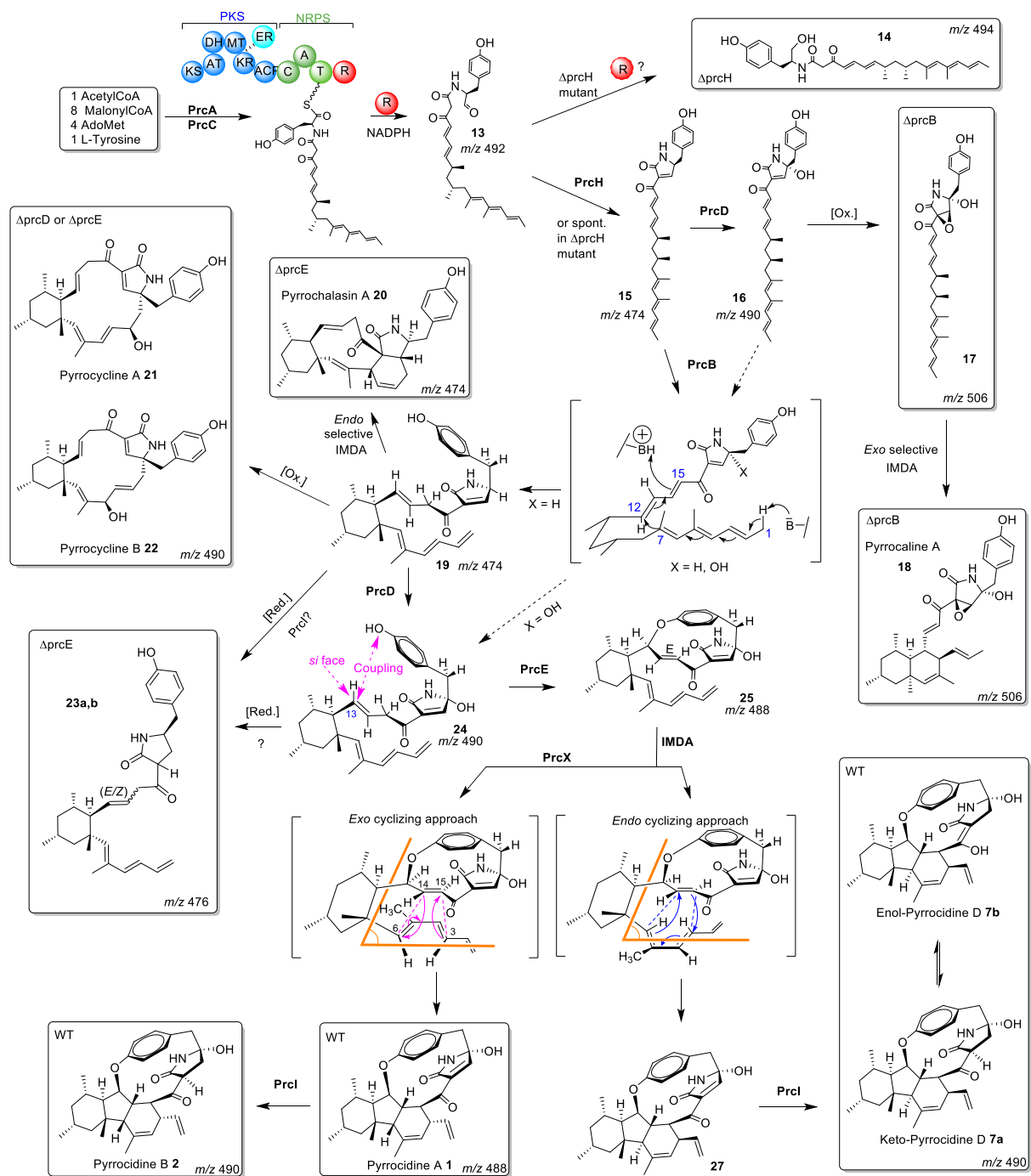


Fig. 2.37 Proposed biosynthesis pathway of pyrrocidines in *Sarocladium zeae*. Structures of the metabolites isolated from the wild type and mutant strains are represented in the boxes.



More importantly, from a stereochemical point of view, the proposed mechanism cannot produce pyrrocidines. Indeed, the attack of the phenate on C13 of *E*-C12=C13 alkene is expected to be antiperiplanar to the attack of C12 towards C7 (in other words the phenate attacks the olefin at C13 on its *si* face allowing C12 attack from its *si* face towards *si* face of C7). This results in the (*R*)-C13 configuration instead of the expected (*S*)-C13 in pyrrocidines, disqualifying the proposed mechanism for the formation of pyrrocidines.

Similarly to pyrrocidine B, the reconstitution of the biosynthesis of GKK1032A<sub>2</sub>, an analogue of pyrrocidine D 7 with one more methyl in C3, was attempted from a gene cluster identified in *Penicillium citrinum* and lacking both genes coding PrcX and PrcI homologues. The GKK1032A<sub>2</sub> were successfully produced with the presence of the gene PrcI and without PrcX. Here PrcX is not required given that the ring C stereochemistry of GKK1032A<sub>2</sub> is the same as in pyrrocidine D 7 and originates from spontaneous IMDA according to an endo cyclization (Fig. 2.38).

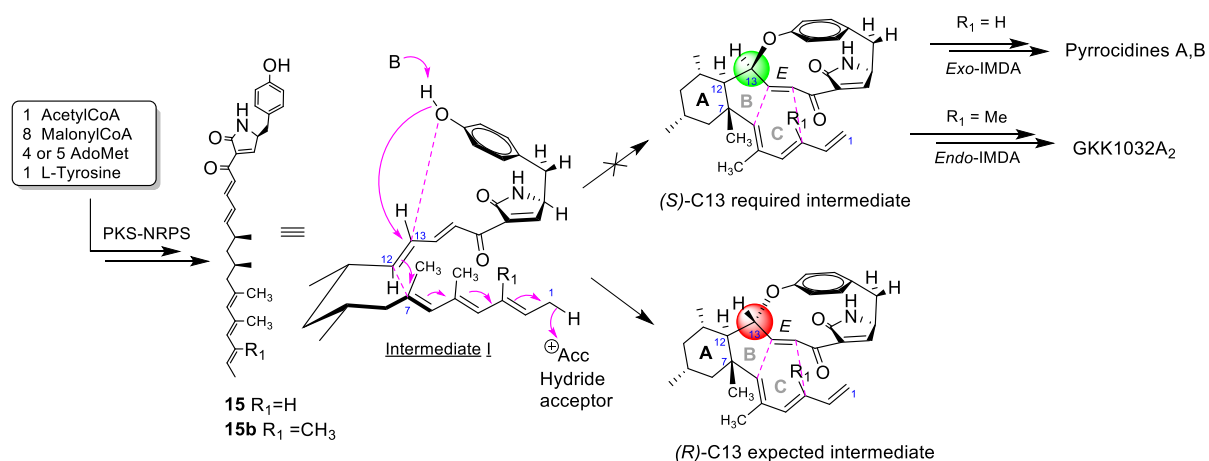


Fig. 2.38 Biosynthetic pathway for pyrrocidines and GKK1032A<sub>2</sub> proposed by Ohashi *et al.* (ref). Upon deprotonation, the phenate attacks the C13 which trigger the conjugated addition of C12 onto the triene at C7 to form ring A. The (*R*)-C13 intermediate would be expected to be formed in this approach instead of the required *S*-C13 found in pyrrocidines and GKK1032A<sub>2</sub> and proposed by the authors.

The second publication reported the isolation of the three xenoacremones A to C and the characterization of their *xen* biosynthetic cluster from the fungal strain *Xenoacremonium sinensis* ML-31. This cluster is composed of seven genes coding for proteins homologous to those found in pyrrocidine cluster (Fig. 2.39).

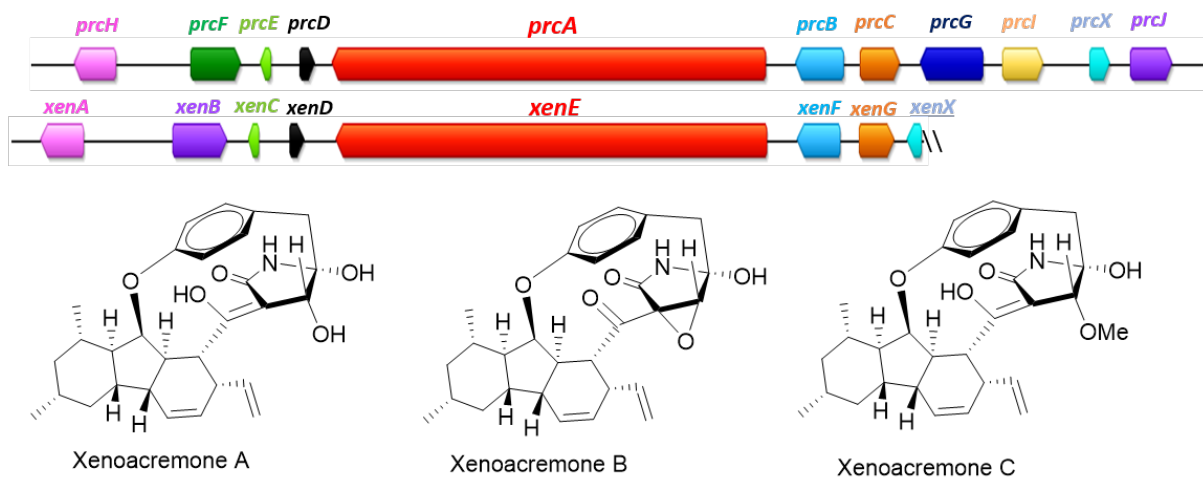


Fig. 2.39 *prc* and *xen* biosynthetic gene clusters producing pyrrocidines and xenoacremones, respectively. Structures of xenoacremones isolated by Liu *et al.*

Using gene inactivation in the producing strain and heterologous expression of the *xen* BGC in *A. nidulans*, the authors isolated several accumulating metabolites and give a function to the auxiliary enzymes encoded in the BGC. An enzymatic cascade forming the paracyclophane-decahydrofluorene structure was proposed in three steps from the all-*E* polyene intermediate I: 1) first, XenF (PrcB homologue) catalyzes the formation of ring A; 2) then XenC (PrcE homologue) build the cyclophane after XenD (PrcD homologue) oxidized the pyrrolidone ring; 3) finally, ring B and C of xenoacremones are formed by a spontaneous IMDA.

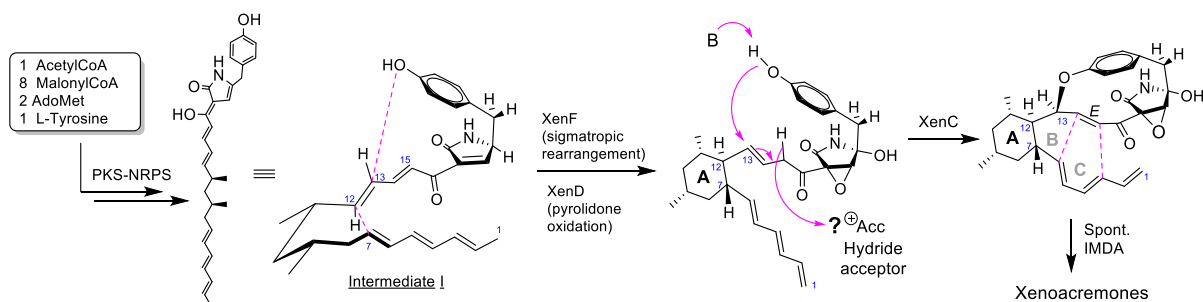


Fig. 2.40 Proposed biosynthesis for xenoacremones by Liu *et al.* The bonds to be formed are represented with pink dashed lines. XenF would catalyzed a [1,11]-sigmatropic rearrangement and the pyrrolidone ring would be hydroxylated and epoxidized by XenD, Nucleophilic attack of the phenate to C13 would be catalyzed by XenC to form the cyclophane.

The proposed functions of the auxiliary enzymes are in good agreement with the one we found in pyrrocidine cluster. However, we identified several questionable points in the proposed biosynthetic pathway:

- The authors speculated that a [1,11]-sigmatropic rearrangement is catalyzed by XenF. However, they confusingly illustrated it according to a base catalyzed mechanism with

abstraction of the C1 proton and protonation at C15 (Annex 15). Sigmatropic rearrangement is a kind of pericyclic reaction which, in the present case, would be the transfer of one hydrogen atom from C1 to C15 with a concerted displacement of the double bonds without any chemical catalysis.

- XenE was proposed to be the catalyzer of the paracyclophane formation. After deprotonation of the phenol, C13 undergoes the attack of the phenate which triggers the shift of the double bond to C14=C15 position. The formation of the paracyclophane corresponds to an oxidation step and would require a hydride acceptor (as already mentioned in the first concurrent publication) positioned at C15 in this case. The presence of such acceptor is not evocated in this publication and leads to erroneous mechanistic representation (Annex 15).
- The authors proposed that xenoacremone A is formed from xenoacremone B through a spontaneous hydration (Annex 15). However, this conversion implies a reduction process which is very likely not spontaneous. Hence, xenoacremone B is probably not the precursor of A.

Beside these inaccuracies, we also analysed the DNA sequence deposited in genbank corresponding to the xenoacremone BGC. Our analysis revealed the presence of a partial gene at one BGC extremity coding for a N-terminal truncated homolog of PrcX. The amino acid sequence corresponds to 79 % of the length of PrcX with 61% sequence identity. This suggests that xenoacremone BGC is not completely identified and that a homologue of PrcX named XenX is involved in the biosynthesis to control the IMDA. This is supported by the supplementary data provided by the authors. Indeed, the LC-MS data of the reconstitution pathway of xenoacremones A and B (without XenX) shows that these two metabolites are minorly produced beside other isomers, whereas in the wild type these compounds correspond to single peaks (see EIC  $m/z$  478.26 and  $m/z$  476.24 of Fig. 2.41). This is coherent with our finding in *S. zeae*  $\Delta prcX$  mutant and wild type where the ratio of pyrrocidines B **2** and D **7** is modulated depending on whether PrcX is present or not.

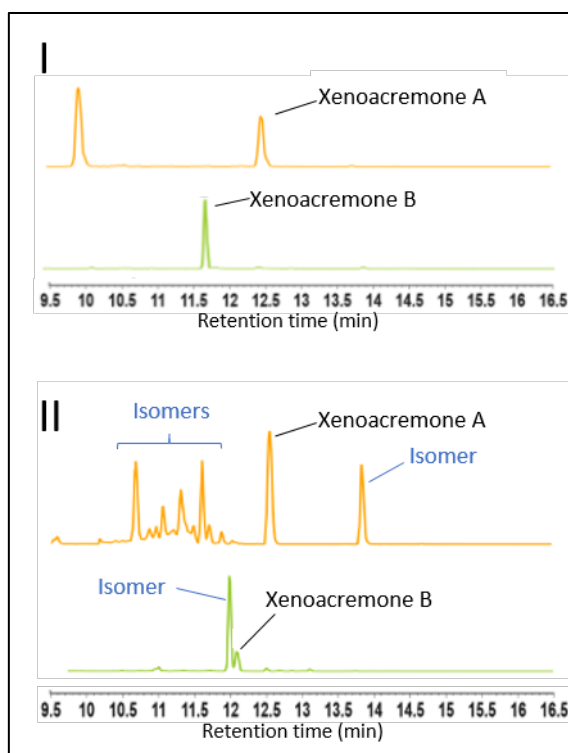


Fig. 2.41 LC-MS data published by Liu et al.<sup>217</sup> on xenoacremone biosynthesis showing the EIC of  $m/z$  478.26 (yellow) and  $m/z$  476.24 (green), corresponding to xenoacremones A and B, respectively. I. EIC from wild type *Xenoacremonium sinensis* ML-31 producing xenoacremones A and B. II. EIC from reconstitution of the pathway in *A. nidulans* (without XenX) showing the presence of xenoacremones A and B and other isomers.

Globally, these points and analysis highlight that the proposed mechanisms for each step of the polycyclic system construction in xenoacremones is not convincing and should be clarified. Moreover, despite these two publications, the mechanism by which the paracyclophane is formed still remains an enigma. Indeed, crystal structures of pyrrocidines and analogues all showed an aromatic ring bending in the paracyclophane moiety (more than  $12^\circ$  from planarity). The hypothesized ring closing by the attack of the phenol oxygen to C13 resulting in a strained paracyclophane does not consider the necessary highly energetic distortion of the aromatic ring.

In order to decipher the molecular mechanisms leading to the decahydrofluorene-paracyclophane moiety, the study of the auxiliary enzymes was initiated in our group. This is one part of the ongoing PhD project of Steffi Sewsum that we won't describe here in detail. Her results combined with the present biosynthetic study gave rise to the submitted publication joined to this manuscript. Briefly, Steffi Sewsum performed *in silico* study on PrcB and PrcE in collaboration with Dr. Nicolas Pietrancosta (LBM, Sorbonne University). The modelisation of PrcB and docking experiments with the substrate **15** indicate that the enzyme builds ring A

through a base catalyzed mechanism and not via a [1,1]-sigmatropic rearrangement as claimed by Liu *et al.* Her work also showed that the substrate adopts several conformations in the active site that can explain the formation of E/Z-alkene configurations, more particularly in compounds **23a,b**, and also the stereochemistry observed in saroclidine A.

Based on a structural analysis of the four-transmembrane protein PrcE, our group proposed an enzymatic mechanism for the formation of the paracyclophane via transient loss of aromaticity of tyrosine residue accounting for the presence of distorted aromatic ring (Fig. 2.42).

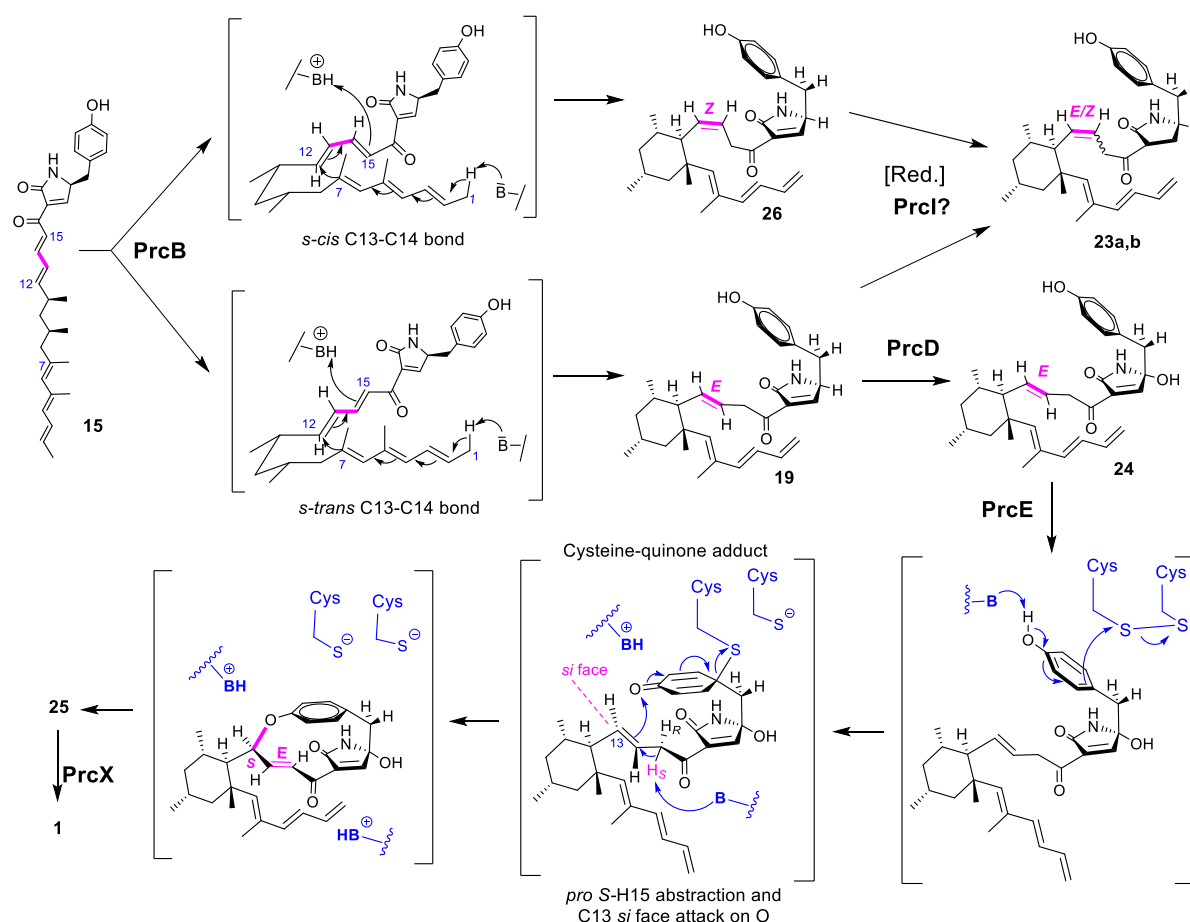


Fig. 2.42 Proposed stepwise enzymatic mechanism for the formation of paracyclophane-decahydrofluorene motif in pyrrocidine pathway. PrcB catalyzes the formation of ring A and PrcD performs hydroxylation on C20. PrcE builds the paracyclophane with the transient formation of a cysteine-quinone adduct which undergoes a nucleophilic attack on its oxygen by C13.

## 2.7 Evaluation of the effect of the pyrrocidine production during the interaction between *S. zeae* and *Fusarium verticillioides*

As explained in the introduction, a mutual antagonism has been reported between the fumonisin producer *Fusarium verticillioides* and the pyrrocidine producer *S. zeae* by Wicklow *et al.*<sup>218</sup> leading these authors to characterize *S. zeae* as “protective endophyte” of maize. Significant antifungal activities of pyrrocidines A and B have also been reported against *F. verticillioides*. Moreover, a recent study showed that sub-fungitoxic levels of pyrrocidine B in *F. verticillioides* cultures nearly abolish the production of the fumonisins. Therefore, since we had in hands the  $\Delta prcA-B$  mutant strain which does not produce pyrrocidines, we assess the possibility for *S. zeae* to employ pyrrocidines as chemical weapon to control *F. verticillioides* in the reported antagonism and we attempted to measure the impact of the presence of pyrrocidines on the production of fumonisins during the fungal interaction.

For that, we selected three different strains of *F. verticillioides* from the fungal collection of CIRM-CF and known to produce fumonisins. These strains are referenced as BRFM 2250, BRFM 2251 and BRFM 2252 in the collection. Each of them was cultivated in the presence of wild type *S. zeae* and  $\Delta prcA-B$  *S. zeae* on PDA and the culture was monitored for seven days. First of all, we observed that wild type and mutant *S. zeae* grow at the same rate and 2 to 3 times slower than *Fusariums*. We observed the mutual antagonism between wild type *S. zeae* and the three *Fusarium* strains as described by Wicklow *et al.*<sup>218</sup> *i.e.* “mutual inhibition on contact, the space between the two colonies is small, but clearly marked”. In the three cultures wild type and mutant *S. zeae* are equally impacted by *Fusariums*. By contrast, a slight difference is observed in the progression of *Fusarium* towards the mutant compared to the wild type *S. zeae*. Indeed, the three *fusarium* strains progress farther when facing the mutant than the wild type strain and they tend to grow around their pyrrocidine non-producing competitor (Fig. 2.43). This preliminary experiment indicates that pyrrocidines play a role in the fungal interaction.

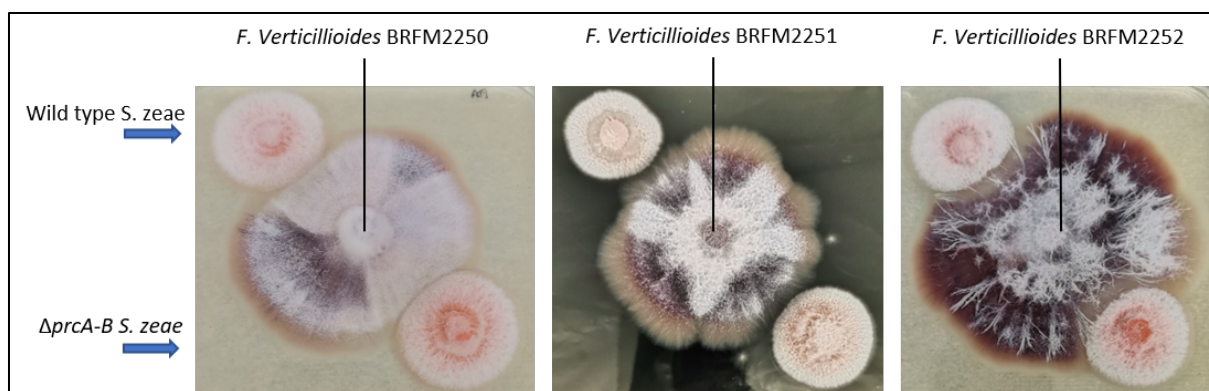


Fig. 2.43 Evolution of cultural antagonism between *F. verticillioides* BRFM 2250, BRFM 2251 or BRFM 2252 with wild type and  $\Delta prcA-B$  *S. zeae*. The culture was maintained at 25 °C for seven days on PDA.

During our biosynthetic study, we generated several other mutants which produce different kind of metabolites. One of them, *S. zeae-ΔprcI*, is particularly interesting because it generates pyrrocidine A but not pyrrocidine B and by similar experiments we can evaluate the impact of pyrrocidine B in the antagonism. Bioactivities of the metabolites that we identified from other *S. zeae* mutants will also be evaluated according to same assay.

In order to estimate the impact of pyrrocidines on the biosynthesis of fumonisins during the fungal interaction, we tried to inventory the fumonisins produced by the three strains of *F. verticillioides*. A large diversity of these mycotoxins is produced by *Fusarium* as reported in Table X. The linear fumonisin backbone can be functionalized with tricarballylate (TCA), 3-hydroxypyridinium (3HP), hydroxy or amino groups at various positions.

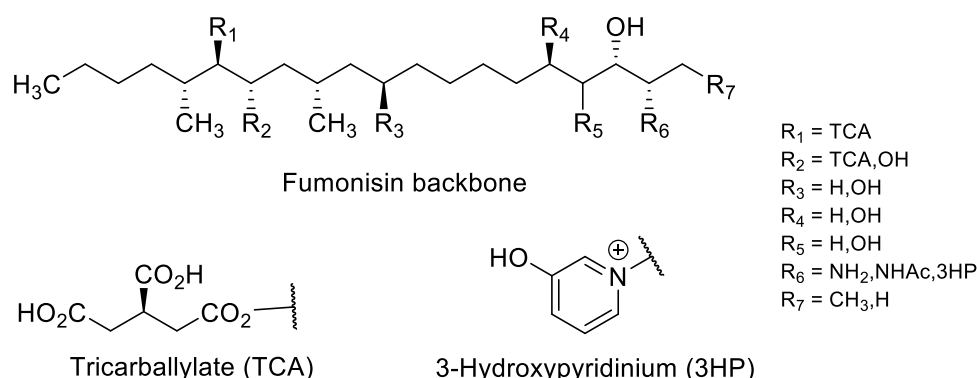


Fig. 2.44 Backbone of fumonisins and substituents encountered in these mycotoxins.

Name	[M+H] <sup>+</sup>	Side chains of fumonisin backbone						
		R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	R <sub>6</sub>	R <sub>7</sub>
Fumonisin B1	722.3963	TCA	TCA	OH	OH	H	NH <sub>2</sub>	CH <sub>3</sub>
Fumonisin B2	706.4014	TCA	TCA	H	OH	H	NH <sub>2</sub>	CH <sub>3</sub>
Fumonisin B3	706.4014	TCA	TCA	OH	H	H	NH <sub>2</sub>	CH <sub>3</sub>
Fumonisin B4	690.4065	TCA	TCA	H	H	H	NH <sub>2</sub>	CH <sub>3</sub>
Fumonisin A1	764.4069	TCA	TCA	OH	OH	H	NHCOCH <sub>3</sub>	CH <sub>3</sub>
Fumonisin A2	748.4119	TCA	TCA	H	OH	H	NHCOCH <sub>3</sub>	CH <sub>3</sub>
Fumonisin A3	748.4119	TCA	TCA	OH	H	H	NHCOCH <sub>3</sub>	CH <sub>3</sub>
Fumonisin C1	708.3806	TCA	TCA	OH	OH	H	NH <sub>2</sub>	H
Fumonisin C3	692.3857	TCA	TCA	OH	H	H	NH <sub>2</sub>	H
Fumonisin C4	676.3908	TCA	TCA	H	H	H	NH <sub>2</sub>	H
Fumonisin P1	801.4147	TCA	TCA	OH	OH	H	3HP	CH <sub>3</sub>
Fumonisin P2	795.4198	TCA	TCA	H	OH	H	3HP	CH <sub>3</sub>
Fumonisin P3	795.4198	TCA	TCA	OH	H	H	3HP	CH <sub>3</sub>
Hydrolyzed fumonisins	590.3904	TCA	OH	OH	H	H	NHCOCH <sub>3</sub>	CH <sub>3</sub>
N-acetyl-fumonisin C1	750.3912	TCA	TCA	OH	OH	H	NHCOCH <sub>3</sub>	H
N-acetyl-OH-fumonisin C1	766.3861	TCA	TCA	OH	OH	OH	NHCOCH <sub>3</sub>	H

Fig. 2.45 Fumonisin B and its analogs characterized in the literature.<sup>73</sup>

The three strains were independently cultivated on PDA for seven days in duplicates and the metabolites after extraction with ethyl acetate were analysed by LC-MS. However, neither fumonisin B1 nor its analogs were detected in the extracts. We envisaged that these strains does not produce these metabolites on PDA, therefore we performed the analysis from another medium, GYAM with agar, used for fumonisin production assay,<sup>219</sup> but still the detection of the mycotoxins remained unsuccessful (Fig. 2.46).

We envisaged that the absence of production of fumonisins in these strains could be due to an inappropriate choice of the solid medium. Consequently, the production should be tested in liquid culture media. However, this liquid condition is unsuitable for the planned antagonism experiments. Another possibility to explain the absence of mycotoxin is that the biosynthesis of this metabolites became silent after storage and repeated cultivations of the fungi. The capacity of these fusarium strains to produce the mycotoxins will be further investigated and, alternatively, other strains will be tested to carry out this study.



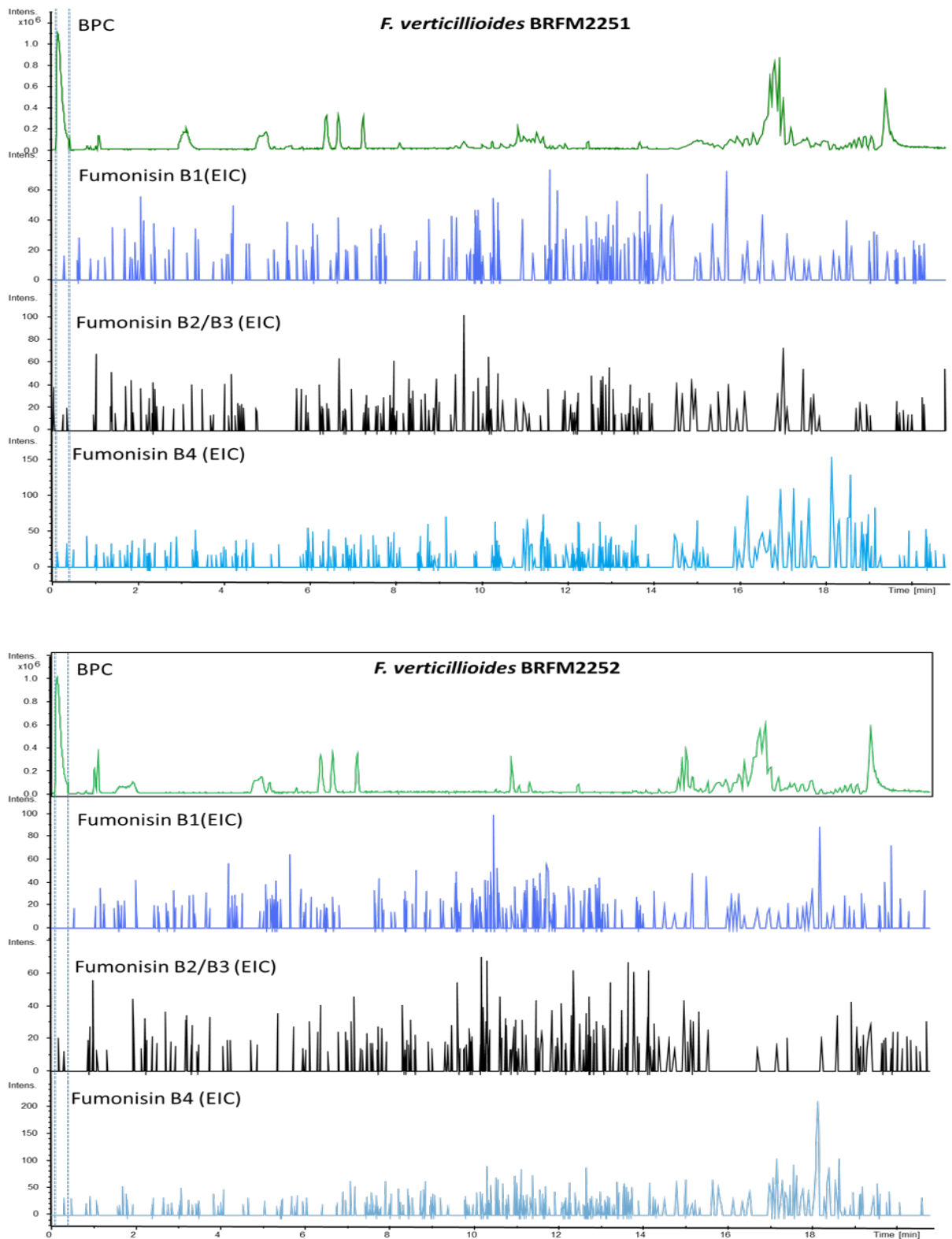


Fig. 2.46 BPC and EIC used to analysis the fumonisin B in PDA medium with different *F. Verticilloides* strains.

## 2.8 Pyrrocidine BGC in other fungi: Genome mining

The development of sequencing technologies and DNA assembling tools rendered the access to fungal genomes easier so that the public databases have grown considerably over the past five years and now number more than 2 000 fungal genomes. This still represents an untapped source of metabolic pathways in which we can expect to find similar pyrrocidine BGC producing new analogous metabolites with potential bioactivities and even new auxiliary enzymes. Therefore, throughout a genome mining approach we explored these databases. Using the auxiliary enzymes as reference sequences with BLAST, we could retrieve many BGC candidates that we manually annotated. From those, fifteen pyrrocidine-like biosynthetic gene clusters were identified (Fig. 2.47) and eight of them were available from commercial collections (2, 5, 7, 9, 10, 11, 14 and 15).

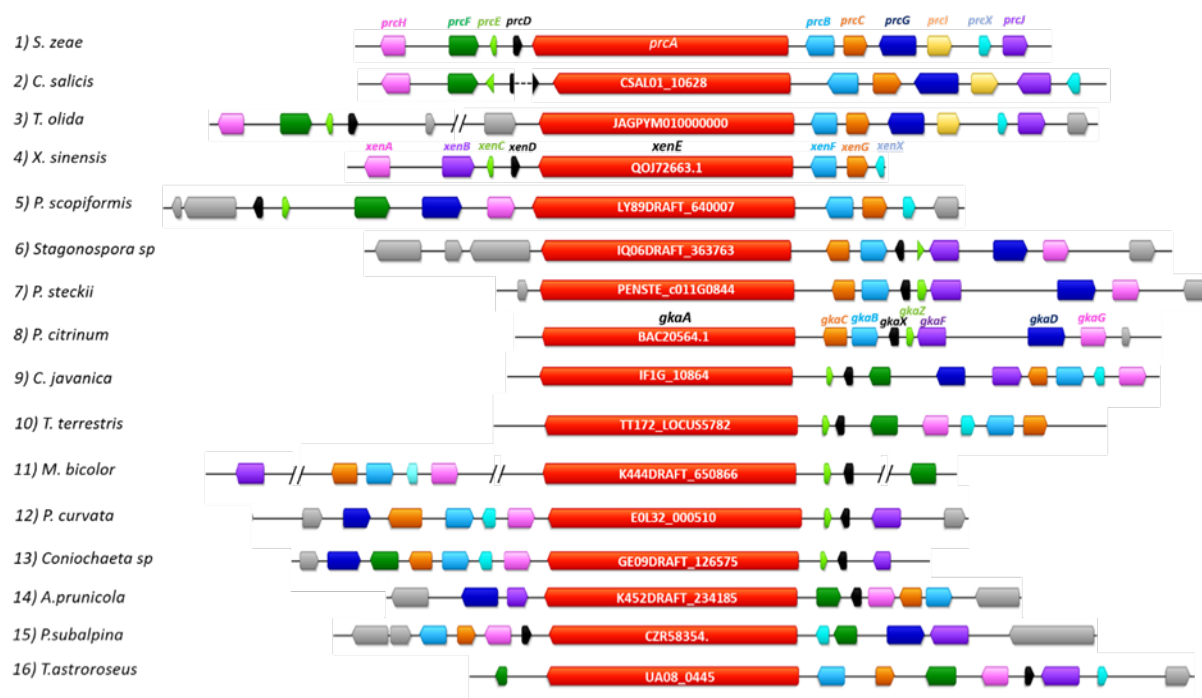


Fig. 2.47 Pyrrocidine-like BGCs identified in published fungal databases. Identified strains: 2) *Colletotrichum salicis*. 3) *Thelonectria olida*. 4) *Xenoacremonium sinensis*. 5) *Phialocephala scopiformis*. 6) *Stagonospora* sp. 7) *Penicillium steckii*. 8) *Penicillium citrinum*. 9) *Cordyceps javanica*. 10) *Thermothielavioides terrestris*. 11) *Meliniomyces bicolor*. 12) *Phialemoniopsis curvata*. 13) *Coninchaeta* sp. 14) *Aplosporella prunicola*. 15) *Phialocephala subalpine*. 16) *Talaromyces atroseus*.

We acquired these strains and we cultivated them in triplicates on various solid media (maize, rice, wheat, oat) in an “OSMAC” approach. After extraction with ethyl acetate, the metabolite extracts were analysed by LC-MS and LC-MS/MS. A rapid search of pyrrocidines molecular ions showed that none of them were produced by these fungi in the selected culture

conditions. Therefore, we more systematically search for pyrrocidine analogues from the molecular ion fragmentations generated in the LC-MS/MS analysis of the metabolite extracts. The web-server GNPS is dedicated to this kind of approach and provides different open-source tools. More specifically Molecular Networking tool allows the display of the chemical space present in tandem mass spectrometry data (MS/MS) by gathering sets of spectra from related molecules. The software builds molecular networks in which each spectrum corresponding to one molecule appears as a node and node connections indicate the molecules sharing similar fragmentations (i.e. similar structure).

The LC-MS/MS data of each fungal extracts were supplemented with those of an extract containing pyrrocidines as reference (wild type *S. zea* extract) and processed with GNPS.

We obtained a molecular network involving pyrrocidines only from *Cordyceps javanica* extract in which four compounds were presumed to be analogs of pyrrocidines (Fig. 2.48) and characterized by their molecular ions  $[M+H]^+$   $m/z$  528.309 ( $C_{34}H_{42}NO_4^+$ ),  $m/z$  530.307 ( $C_{34}H_{44}NO_4^+$ ),  $m/z$  544.304 ( $C_{34}H_{42}NO_5^+$ ),  $m/z$  546.319 ( $C_{34}H_{44}NO_5^+$ ).

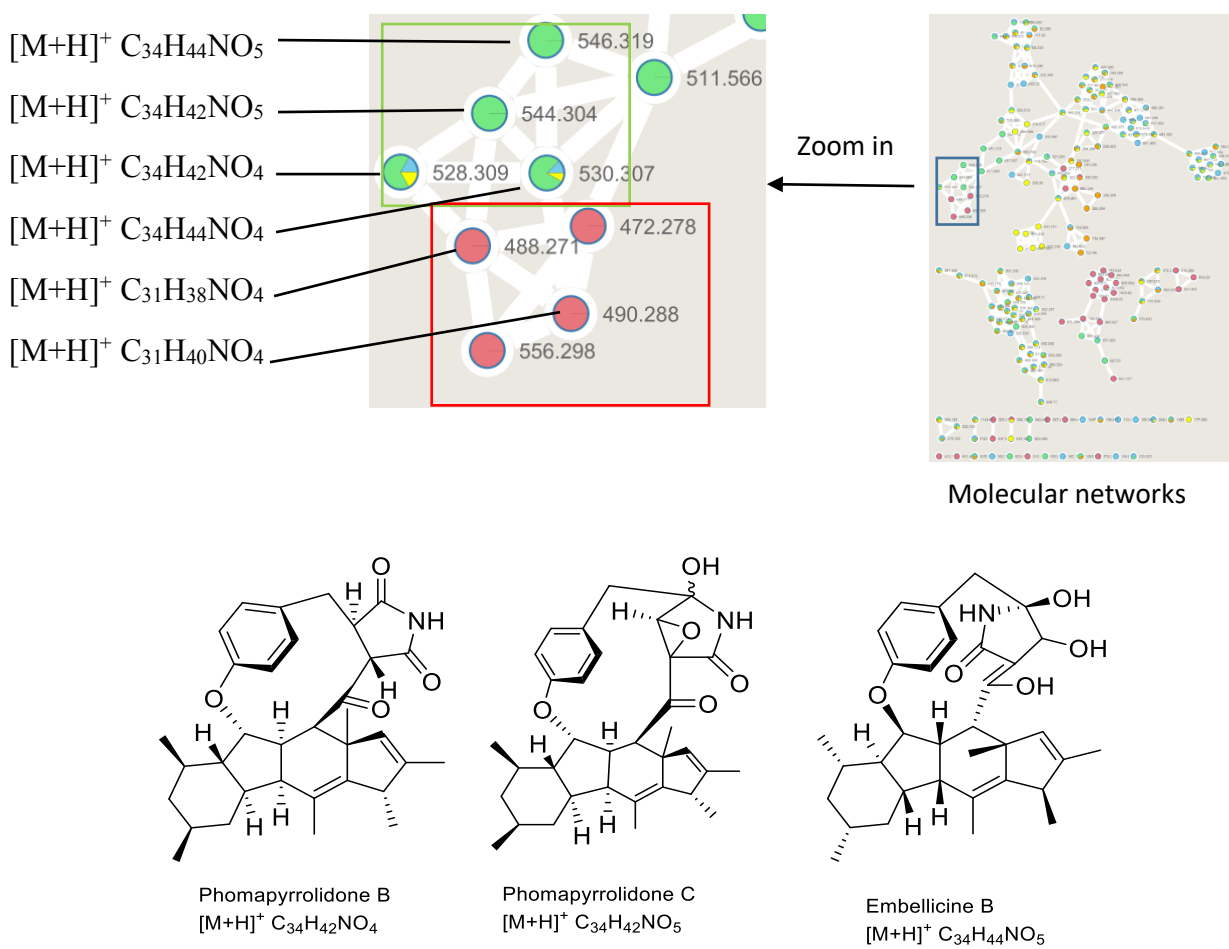


Fig. 2.48 Molecular Network obtained by GNPS analysis, and structures of three known metabolites possibly produce by *Cordyceps javanica*.

These molecules share some fragments with pyrrocidines A and B (Fig. 2.49). From the calculated formula given by the [M+H]<sup>+</sup> molecular ion, these compounds could belong to the families of embellicines or phomapyrrolidones (Fig. 2.48). After scaled up fermentation the extract was purified but the amount produced was not sufficient to characterize the metabolites by NMR. Scale up culture and optimized purification steps will be investigated to further perform structure elucidation.

M/Z \ Fragments	488	490	530	528	544	546
81.0698	✓	✓	✓	✓	✓	✓
95.0855	✓	✓	✓	✓	✓	✓
107.0492	✓	✓	✓	✓	✓	✓
126.0182	✓	✓	-	✓	-	-
185.1324	✓	✓	✓	-	✓	✓
	Pyrrocidine A	Pyrrocidine B	C <sub>34</sub> H <sub>43</sub> NO <sub>4</sub>	C <sub>34</sub> H <sub>41</sub> NO <sub>4</sub>	C <sub>34</sub> H <sub>41</sub> NO <sub>5</sub>	C <sub>34</sub> H <sub>43</sub> NO <sub>5</sub>

Fig 2.49 Four analogs shared the same fragments with pyrrocidine A and B.



## **Chapter III: Reconstitution of the biosynthesis pathway of pyrrocidines**

### 3.1 Introduction

In the previous chapter, we showed the diversity of compounds produced by pyrrocidine BGC which demonstrates the metabolic potential of the pyrrocidines pathway. We also identified the function of auxiliary enzymes encoded in the gene cluster. To further explore the chemical space of this pathway and investigate the mechanism of the involved proteins, it is crucial to reconstruct or remodel the pathway step by step. A suitable heterologous expression host is critical in this reconstitution. Several hosts are available: *Escherichia coli* is very commonly used for heterologous expression, however, by belonging to the prokaryotes it can neither properly splice introns nor post-translationally modify eukaryotic-derived proteins. More recently, fungal expression systems which manages intron splicing were developed for reconstitution of fungal BGC, more particularly in *Aspergillus sp.* As mentioned in chapter 2, pyrrocidine pathway was reconstituted in *Aspergillus nidulans*, however, the metabolic yield remained very low.<sup>216</sup> The use of these fungal strains is developing and the range of genetic tools still need to be further expanded. By contrast, *Saccharomyces cerevisiae* is a powerful expression host. Several strains have been developed for the heterologous expression of fungal gene clusters, despite their inability to splice introns.<sup>220</sup> Particularly, the specific host *S. cerevisiae* BJ5464-NpgA was used for PKS and hybrid PKS-NRPS expressions and metabolite productions. This strain was genetically modified with the integration of *npgA* gene from *A. nidulans* encoding the phosphopantetheine transferase (PPTase) required to activate the apo-form of ACP and T domains into their holo-form functionalized with a phosphopantetheine arm (Fig. 3.1).

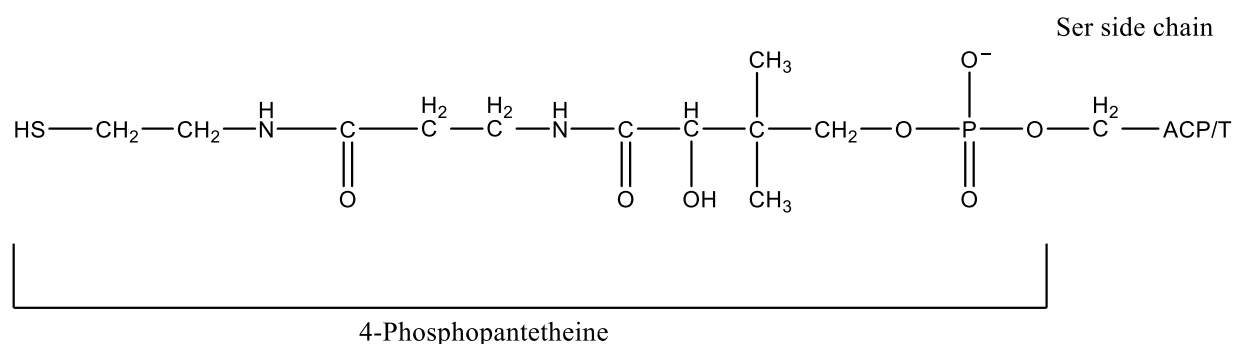


Fig. 3.1 A functional ACP/T domain (holo ACP/T) covalently links with a phosphopantetheine arm (PPant arm).

Because of its long history in industrial fermentation and as model organism, numerous genetic tools are available for *S. cerevisiae*. With the advent of synthetic biology, an effort of standardization of these tools has been made which leads to the development of versatile

engineering platforms. This is the case of the highly characterized yeast tool kit designed by Dueber and co-workers which provides well characterized elements (categorized in types like promoters, terminators, tags, selection markers, replication origins...) to build expression vectors among other things.<sup>221</sup> All the required types can be assembled by the modular cloning strategy. The latter uses type II restriction enzymes which cut outside their recognition sequence and provide unique cohesive ends enabling directional cloning. For that, the available types in the kit were preliminary and separately cloned in an identical entry vector by the designers, so that the types can be excised from the vector by the type II restriction enzyme BsaI and directionally ligated with a ligase on demand. The assembly is performed *in vitro* on a thermocycler through repeated digestion-ligation cycles.

The flexibility of this approach allows, for instance, to quickly build expression vectors with promoters having diverse strengths, consequently we undertook the reconstitution of the pyrrocidine pathway in *S. cerevisiae* with this tool kit. Our starting point was the cloning on two different plasmids of the PKS-NRPS gene *prcA* and the ER coding *prcC* (each harboring a specific selection marker) whose co-expression should lead to the formation of linear alcohol **14**.

The plasmid construction requires the participation of nine parts each belonging to one specific type. When the nine type-containing plasmids are treated with BsaI, the parts are released giving fragments with four-bases cohesive ends at each extremity. Both cohesive ends of a specific part (n) are designed to complement differently the cohesive ends of both next (n+1) and previous (n-1) parts, so that the subsequent ligation lead to directional assembly. For example, in MoClo, after treating with BsaI restriction enzyme, the cohesive front end (3' end) of type 1 parts is TTGC and the cohesive back end (in 5') of type 2 parts is AACG. Followed by treating with T4 ligase, the front of type 1 part is connected to the back end of type 2 parts. The cohesive back end of type 2 parts is ATAC and the cohesive front end of type 3 parts is TATG, so they are also attached end to end after treating with T4 ligase (Fig. 3.2). Based on this principle, the nine parts can finally be linked together *in vitro* to form a shuttle vector for *E. coli* (for replication) and for Yeast as an expression plasmid (Fig. 3.2).



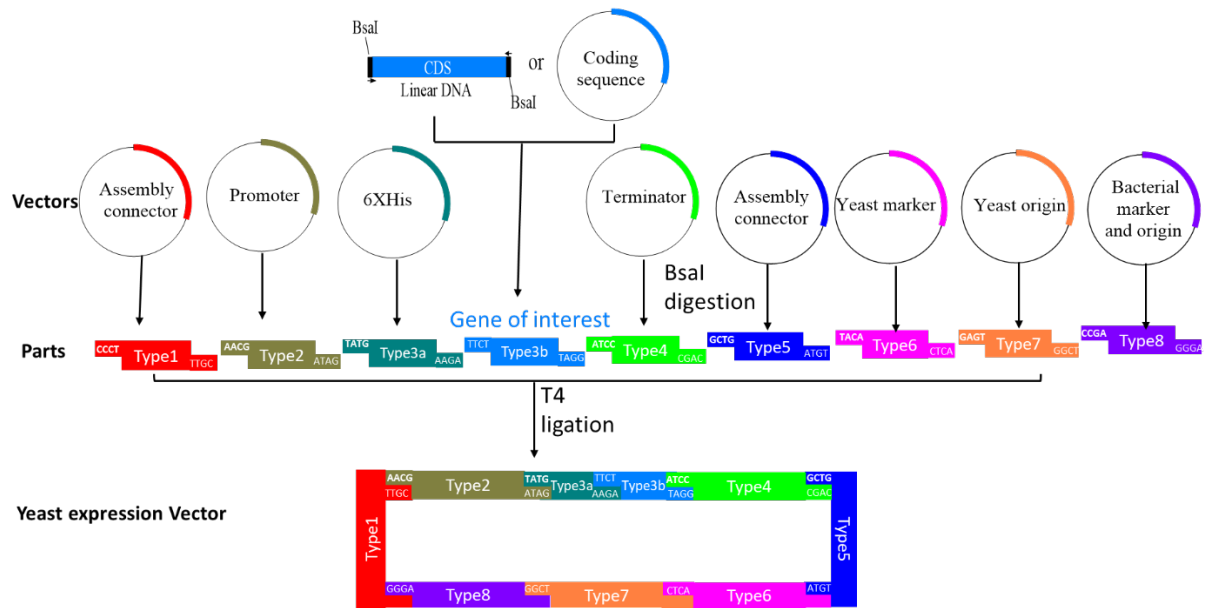


Fig 3.2 Principle of the assembly with modular cloning (MoClo). The plasmid construction requires the participation of nine parts each belonging to one specific type (Type 1: connectors. Type 2: promoters. Type 3a: 6XHis. Type 3b: coding sequences. Types 4: terminators. Type 5: assembly connectors. Type 6: *S. cerevisiae* selection markers. Type 7: *S. cerevisiae* origins. Type 8: *E. coli* markers and origins). Both sides of each part link to a BsaI restriction and recognition site, and then integrate into a Part plasmid for storage.

### 3.2 Cloning of genes *prcC* coding the ER and *prcA* coding the PKS-NRPS

The principle of MoClo relies on the uniqueness of type II restriction enzyme (here BsaI). If the target gene does not contain the restriction site of BsaI, we only need to prepare it by PCR including tails of a BsaI restriction site by primer design. Otherwise, the target gene without BsaI recognition site could be obtained by mutation or synthetic ways. When the target gene is excessively long, it inevitably contains BsaI restriction sites, which was the case for *prcA* (12 kbases). In contrast, *prcC* does not harbor the restriction sequence of BsaI. Therefore, we envisage to build the expression vector with *prcC* by MoClo and then substitute *prcC* with *prcA* by *in vivo* assembly (IVA) cloning in *E. coli*. All the plasmids (in red) used in this chapter and their critical composition see table 3.1.

Table. 3.1 Key building parts of plasmids used in this chapter.

Plasmids	PKS-NRPS gene			
	Promoter	Terminator	Selection marker	Yeast origin
pGG-P67-RPL18B	pRBL18B	tADH1	Leu	CEN6/ARS4
pGG-P67-PGK1	pPGK1	tADH1	Leu	2micron
pGG-P69-TEF1	pTEF1	tADH1	Leu	2micron
Plasmids	ER gene			
	Promoter	Terminator	Selection marker	Yeast origin
pGG-ER-RPL18B	pRBL18B	tADH1	Ura	CEN6/ARS4
PGG-ER-PGK1	pPGK1	tADH1	Ura	2micron
PGG-ER-TEF1	pTEF1	tADH1	Ura	2micron
PGG-EP-RPL18B	pRBL18B	tADH1	Leu	CEN6/ARS4
PGG-EP-PGK1	pPGK1	tADH1	Leu	2micron
PGG-EP-TEF1	pTEF1	tADH1	Leu	2micron
Plasmid	mRuby gene			
	Promoter	Terminator	Selection marker	Yeast origin
pGG-mRuby	pRBL18B	tADH1	Ura	CEN6/ARS4
Plasmid	Venus gene			
	Promoter	Terminator	Selection marker	Yeast origin
pGG-Venus	pRBL18B	tADH1	Leu	CEN6/ARS4

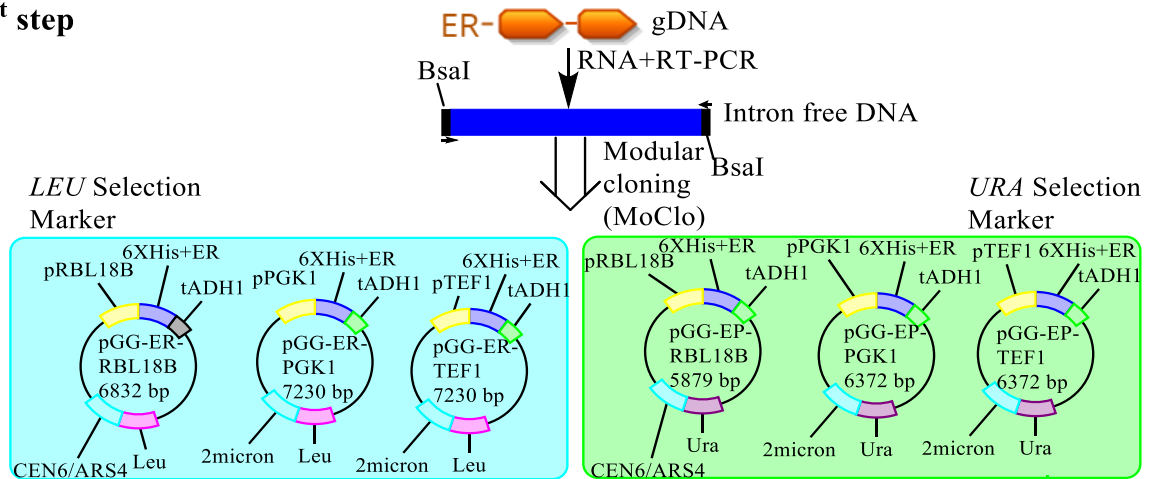
Since *S. cerevisiae* is not able to splice fungal introns, the latter must be removed before the cloning. Analysis with FGENESH web-server predicted six introns in *prcA* and one in *prcC* (Fig. 3.3).



Fig. 3.3 Prediction of the position of introns by FGENESH.

Intron-free sequence of *prcC* was obtained by RT-PCR amplification from extracted total RNAs of *S. zeae*. BsaI restriction site were introduced at both extremities of the gene sequence with adapted primers (P384 and P385) during the PCR. The amplicons were then directly integrated as type 3b part (coding sequence) in MoClo assembly experiments. Once in vitro assembly was performed, *E. coli* was transformed with the mixture and the transformants were screened to obtain the desired plasmids. Three vectors were prepared, all including a leucine auxotrophy selection marker (LEU2) and tADH1 terminator but with different promoters controlling the ER gene, the medium strength promoter pRPL18B and the strong promoters pPGK1 and pTEF1 (Table 3.1). These constructions were respectively named pGG-ER-RBL18B, pGG-ER-PGK1 and pGG-ER-TEF1 (Fig.3.4).

1<sup>st</sup> step



2<sup>st</sup> step

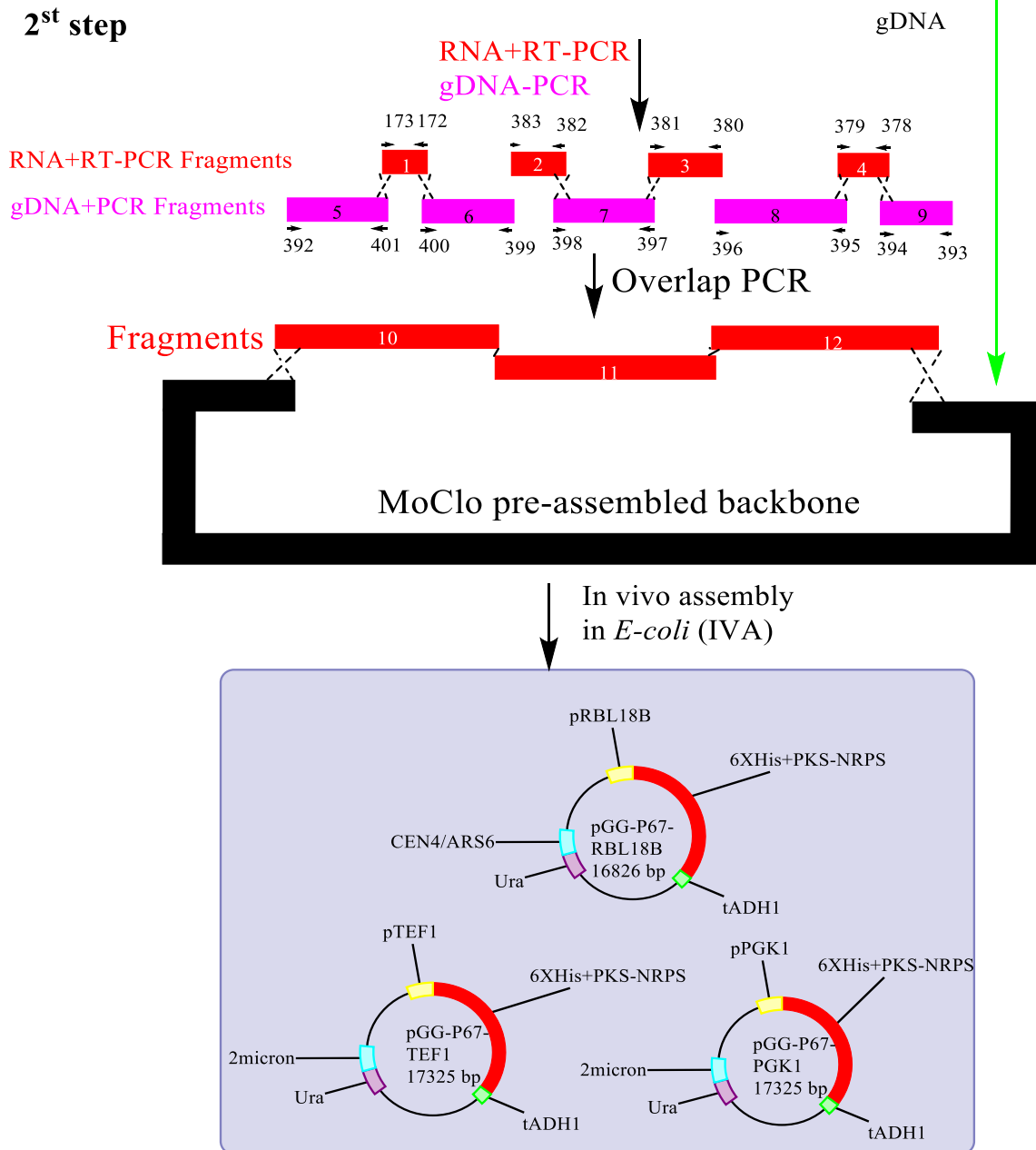


Fig. 3.4 The strategy of cloning of PKS-NRPS and ER.

The right construction of the plasmids were validated by PCR amplification of the transcription unit promoter-ER-terminator with primers P388 and P389 (Fig. 3.5) and sequencing.

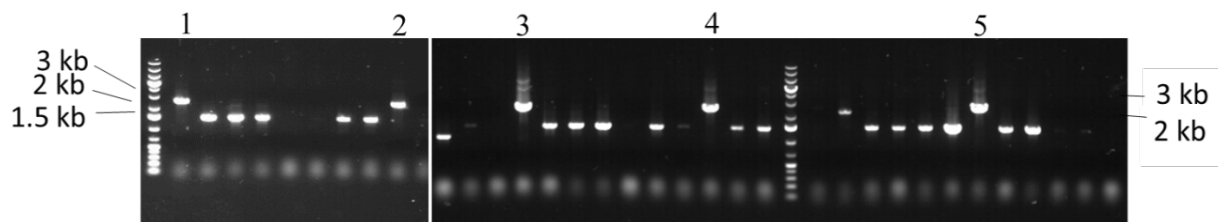
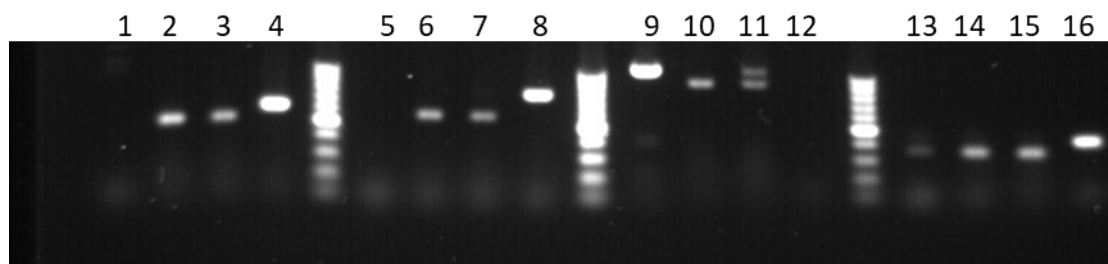


Fig. 3.5 PCR identification of the positive colonies. lanes 1 and 2 were the positive plasmids of pGG-ER-RBL18B (2530 bp). lanes 3 and 4 were the positive plasmids of pGG-ER-PGK1 (2530 bp). Lane 5 was the positive plasmid of pGG-ER-TEF1 (2530 bp).

In order to co-express the PKS-NRPS with ER, another selection marker for *prcA* vector was required. Consequently, we built the same three vectors with uracil auxotrophy selection marker (*URA3*) named pGG-EP-RBL18B, pGG-EP-PGK1 and pGG-EP-TEF1 in which we could substitute ER *prcC* gene by the spliced *prcA* gene (Tab.3.1 and Fig. 3.4).

Since RNA is prone to degradation, the quality of cDNA obtained from reverse transcription of RNA can be challenging to support PCR amplification of large fragment genes. Therefore, the whole *prcA* gene (12 kb) was divided into nine fragments (1-9) with 30 bp overlap between adjacent fragments (Fig. 3.4). Fragments 5 to 9 of PKS-NRPS gene could be obtained by using gDNA as a PCR template with the five pairs of primers P393/P394, P395/P396, P397/P398, P399/P400, P401/P392, because no intron was identified in these sequences. However, shorter fragments 1 to 4 of PKS-NRPS gene containing introns were amplified by using cDNA as template with the four pairs of primers P172/P173, P382/P383, P380/P381, P378/P379 (Fig. 3.4). PCR amplicons of fragments 1 to 4 obtained by using cDNA are shorter than those obtained with gDNA in the PCR positive controls. This validates the existence of predicted introns (Fig. 3.6). The precise location and the size of the six introns were obtained by sequencing. The following sizes of introns were found: Int1: 80 bp; Int2: 70 bp; Int3: 67 bp; Int4: 99 bp; Int5: 61 bp; Int6: 91 bp.



Primers	Lanes	Expect size (bp)	Gel band size (bp)	Lanes	Expect size (bp)	Gel band size (bp)	Lane	
P172,P173	2,3	357	350	4	437	450	1	Negative control
P382,P383	6,7	861	850	8	998	1000	5	Negative control
P381,P380	10,11	567	550	9	757	750	12	Negative control
P379,P378	14,15	514	500	16	594	600	13	Negative control

Fig. 3.6 The PCR results used for identifying the existence of intron (1, 5, 12 and 13 were negative control, H<sub>2</sub>O as the PCR template. 4, 8, 9 and 16 were positive control, gDNA as the PCR template. 2/3, 6/7, 10/11, 14/15 were obtained by using two batches of cDNA as the PCR template).

Considering that the size of the final plasmid used for heterologous expression of PKS-NRPS is more than 17 kb, it is perilous to assemble in *E. coli* too many fragments to form a gigantic plasmid. Therefore, we chose to preassemble fragments 1, 5, 6 by overlap PCR with primers P392/P399 to give fragment 10, as well as the fragments 2, 3, 7, and 8, 9, 4 to give fragments 11 and 12 with primers P380/P383, P396/P393, respectively (Fig. 3.4). The steps of overlap PCR were as follows. To build fragment 10, fragments 1, 5, 6, were used as megaprimers and matrix in the Phusion polymerase standard PCR mix for ten cycles, and then PCR was run with the primers (P392/P399) for 35 cycles as usual. The construction of fragments 11 and 12 used the same approach. These three resulting fragments were then directionally cloned with IVA cloning method into the linearized backbone of vectors pGG-EP-RBL18B, pGG-EP-PKG1 and pGG-EP-TEF1. The linearization of these plasmids was obtained by PCR with primers P390/P391 hybridizing at the end of the promoter and at the beginning of the terminator, and thus discarding the ER gene of the sequence. Both primers also harbored 30 bp homologous arms at their 5' ends allowing the introduction of recombination sites homologous to the extremities of fragments 10, 11 and 12. The three expected constructions including *prcA* were termed as pGG-P67-RBL18B, pGG-P67-PKG1 and pGG-P67-TEF1 (Tab.3.1 and Fig. 3.4).

The expected size of pGG-P67-RBL18B is 16,826 bp. From the cloning, twenty transformants were screened by plasmid extraction and only one retained a size between 10 kb

and 20 kb (Fig. 3.7). The sequence of cloned *prcA* was then verified by Illumina sequencing at the Molecular Systematic Service (SSM) of the MNHN.

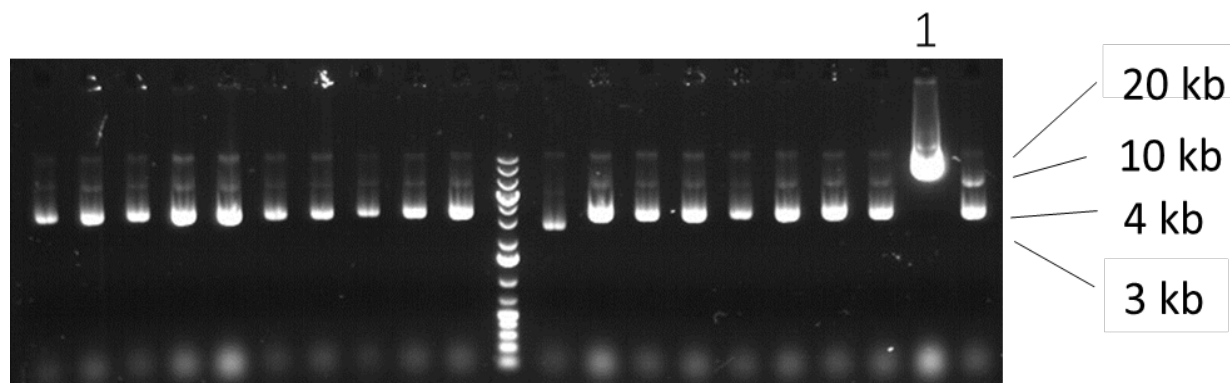


Fig 3.7 Transformant screening by plasmid isolation of the cloning. Electrophoresis on agarose gel of the isolated plasmids. Lane 1 shows the plasmid with the expected size of pGG-67-RBL18B.

Similarly, we obtained the construction of pGG-67-PGK1 and pGG-67-TEF1 after scale screening (Fig 3.8).

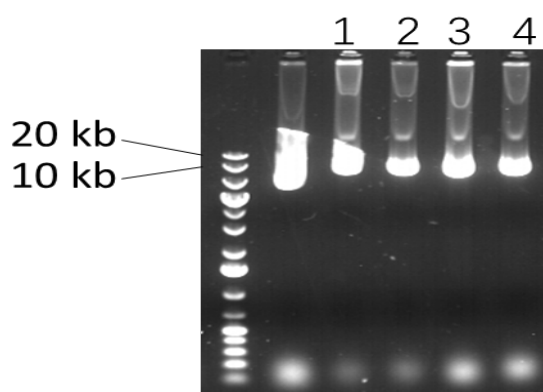


Fig. 3.8 Transformant screening by plasmid isolation of the cloning. Electrophoresis on agarose gel of the isolated plasmids. Lanes 1 and 2 shows the plasmids with the expected size of pGG-67-PGK1. Lanes 3 and 4 shows the plasmids with the expected size of pGG-67-TEF1.

### 3.3 Co-expression of PKS-NRPS-67 and ER in *S. cerevisiae* BJ5464-NpgA

Two plasmids transferred into *S. cerevisiae* BJ5464-NpgA simultaneously is inefficient. Meanwhile, the smaller plasmid replicates more stably in yeast. So, the plasmid pGG-ER-RBL18B was first introduced into BJ5464-NpgA to obtain BJ-ER-RPL18B, and then the plasmid pGG-P67-RBL18B was transferred into BJ-ER-RPL18B to obtain BJ-EP67-RPL18B. The same procedure was performed to obtain BJ-EP67-PGK1 and BJ-EP67-TEF1 by using two pairs of plasmids pGG-ER-PGK1 and pGG-P67-PGK1, pGG-ER-TEF1 and pGG-P67-TEF1.

Two fermentation conditions were applied for *S. cerevisiae* to accumulate metabolites: 1) *S. cerevisiae* BJ-EP67-RPL18B/PGK1/TEF1 were grown in YSM medium that lack of leucine and uracil at 25°C for three days, followed by 0.5 mL seed culture as incubated into 50 mL YPD medium for five days.<sup>90</sup> 2) BJ-EP67-RPL18B/PGK1/TEF1 were grown in 50 mL YSM medium that lack of leucine and uracil at 30 °C until the optical density at 600 nm of the culture reaches 0.6, then 50 mL of YP medium (1% yeast extract, 2% peptone) were added to the flask and the fermentation extended for three days.<sup>44</sup>

*S. cerevisiae* BJ5464-NpgA was used as the control cultured in the same conditions except for the YSM medium containing all the amino acids. Then, all the fermentation was centrifuged at 4 °C, 4000 g for 10 min and extracted with ethyl acetate twice, followed by evaporation and analyzed by LC-MS.

According to our experience of inactivation of *prcH*, the linear alcohol **14** should accumulate by co-expression of PKS-NRPS and ER in yeast. But, we also observed the production of pyrrocidines in *S. zeae-ΔprcH* which means that the released aldehyde **13** undergoes spontaneous Knoevenagel condensation to afford **15** as mentioned in chapter 2. Consequently, from the co-expression of PrcA-PrcC in yeast we could expect the formation of aldehyde **13**, pyrrolidone **15**, or alcohol **14**. Some biotransformation of these metabolites can also occur in yeast, it may be possible to observe reduction of double bonds (Fig. 3.9).

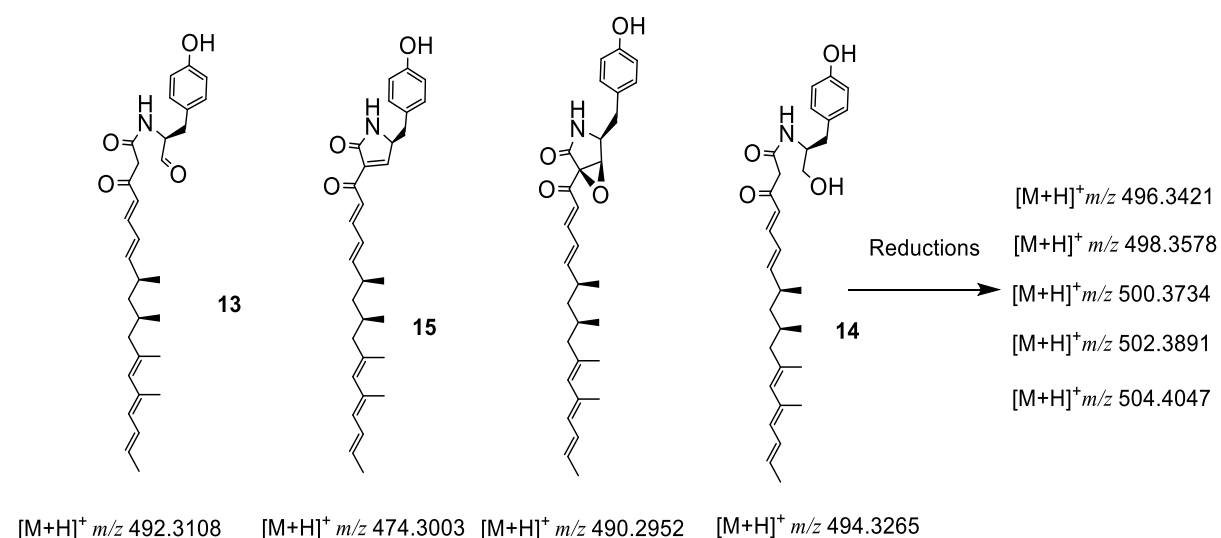


Fig. 3.9 Expected metabolites produced in yeast with the reconstitution of PrcA-PrcC steps. Alcohol **14** for example can undergoes several reductions giving various metabolites whose  $[M+H]^+$  ion mass are indicated.

After LC-MS analysis, we didn't observed the formation of the expected products among the metabolites of *S. cerevisiae* BJ-EP67-RPL18B/PGK1/TEF1 compared to *S. cerevisiae* BJ5464-NpgA as shown in Fig. 3.10.

Since *S. cerevisiae* BY4741 was chosen as the heterologous expression host in the yeast toolkit, it might be these promoters were not efficient in *S. cerevisiae* BJ5464-NpgA leading to the non-expression of the genes, which could explain why we did not detect any expected products.

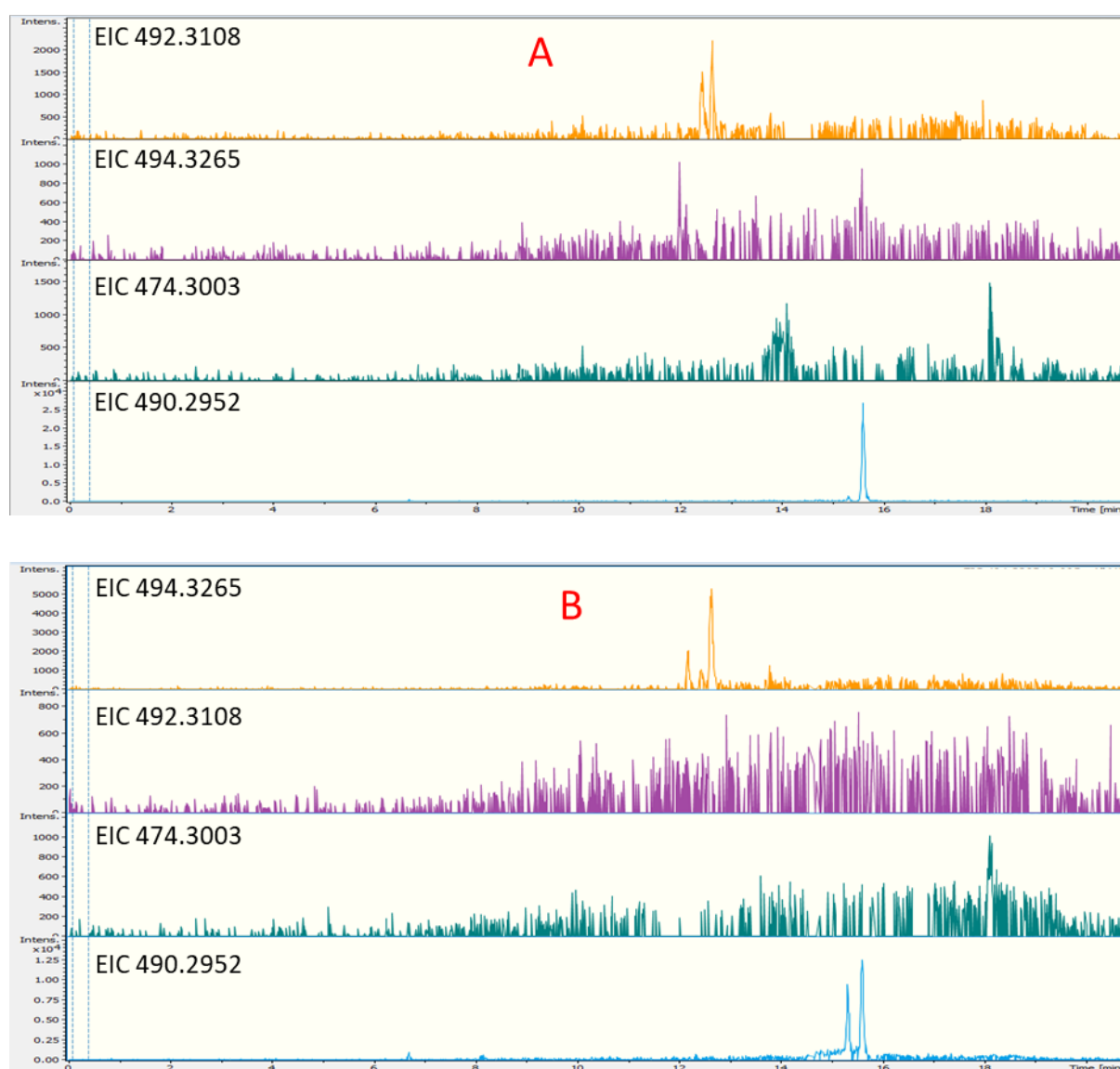


Fig. 3.10 EIC of expected compounds in BJ-EP67-RPL18B(A) compare to the wild type *S. cerevisiae* BJ5464-NpgA (B).



### 3.4 Characterization of the efficiency of promoter pRPL18B in BJ5464-NpgA

To characterize the efficiency of the promoter, we used the fluorescence signal of proteins mRuby and Venus as reporter. For that, by using the Yeast tool kit, we built two plasmids pGG-Venus and pGG-mRuby in which the genes coding mRuby or Venus are under the control of the promoter pRPL18B (Tab. 3.1 and Fig. 3.11).

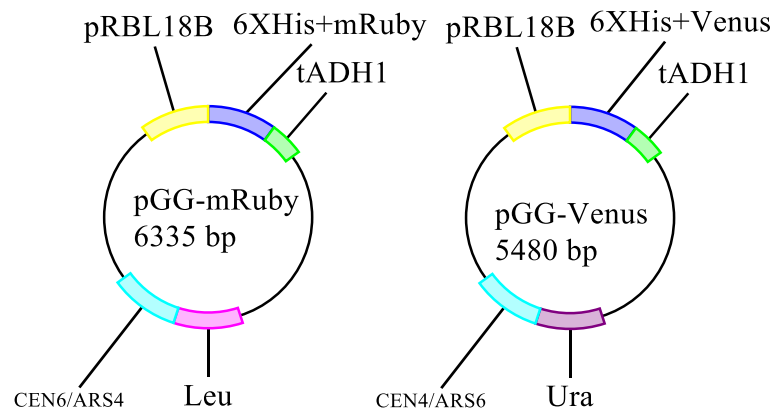


Fig. 3.11 The cloning of fluorescent genes.

The screening of twenty transformants was carried out by PCR amplification of the transcription unit (with primers P388/P389). One transformant for each construction were identified (Fig. 3.12).

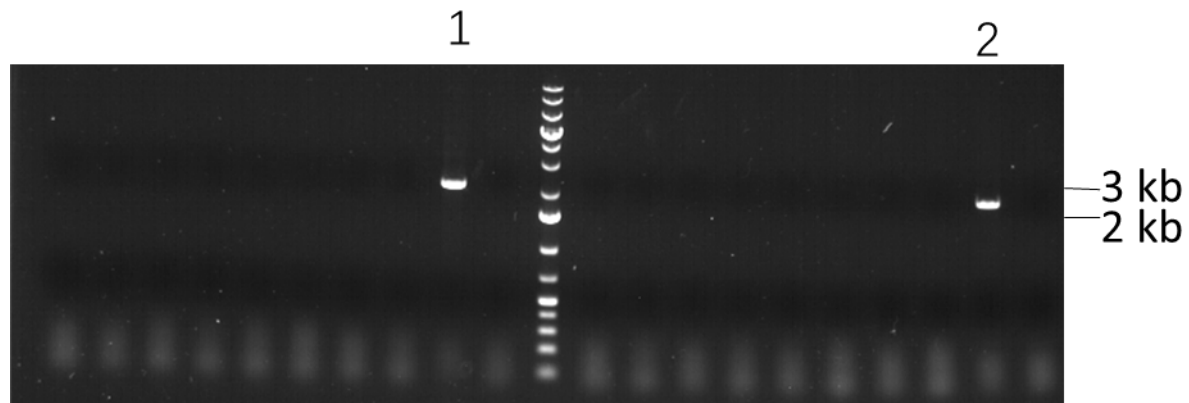


Fig. 3.12 PCR identification of the positive colonies. Lane 1: amplicon of the transcription unit in plasmid of pGG-Venus (expected size 2111bp), Lane 2: amplicon of the transcription unit in plasmid pGG-mRuby (expected size 2108bp).

pGG-Venus and pGG-mRuby were transferred into *S. cerevisiae* BJ5464-NpgA using the same procedure as before to obtain *BJ-Venus* and *BJ-mRuby*, respectively. *BJ-Venus* and *BJ-mRuby* were grown in of YSM medium lacking leucine or uracil up to saturation and then 50 mL of fresh medium were inoculated 0.5 mL of this preculture in triplicates. The fermentation

was extended for 12 hours. A control experiment was carried out with the culture of *S. cerevisiae* BJ5464-NpgA in the same conditions. The cultures were then diluted three times in microplates with fresh medium for fluorescence detection on a plate reader. The 530 nm excitation and 590 nm emission wavelengths were used to detect the mRuby fluorescent protein and 520 nm and 530 nm wavelengths to detect Venus fluorescent protein.

*BJ-mRuby* showed about twice level of fluorescence as *S. cerevisiae* BJ5464-NpgA. *BJ-Venus* showed about 2.5 level of fluorescence as *S. cerevisiae* BJ5464-NpgA (Fig. 3.13). These results demonstrated that the moderate promoter *RPL18B* is operational in *S. cerevisiae* BJ5464-NpgA. Because of this result with moderate promoter, other used strong promoters pTEF1 and pPGK1 in our constructions were not tested (we therefore presumed they should also work well in *S. cerevisiae* BJ5464-NpgA). This suggests that the absence of production of expected metabolites is not related to our plasmid construction.

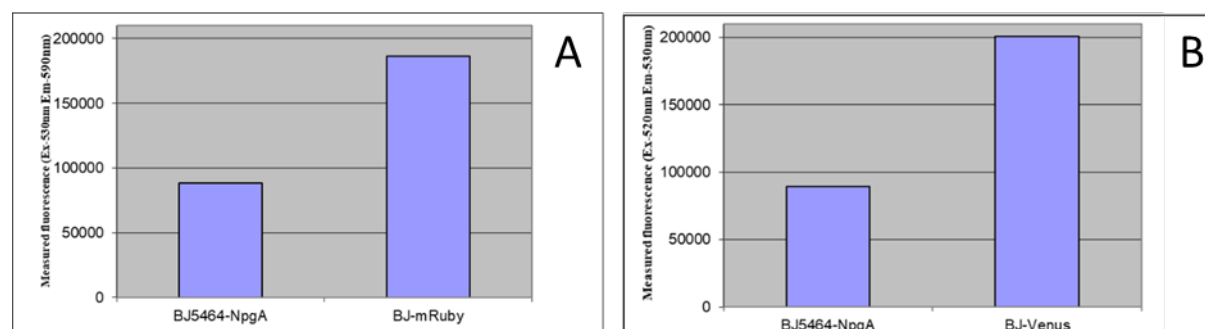


Fig 3.13 The measurement of fluorescence intensity of different strains: A- mRuby fluorescence intensity detection, B- Venus fluorescence intensity detection.

### 3.5 Verification of the production of PKS-NRPS and ER in yeast by Western blot

Type 3a (6XHis) was integrated into all the constructed plasmids, making it convenient to verify the production of PKS-NRPS and ER in yeast by Western blot. So Western blot was performed by using enzyme with His-tag as a positive control and *S. cerevisiae* BJ5464-NpgA as a negative control. After the protein was transferred to the nylon membrane, the membrane was dyed with Ponceau Red. Protein bands and markers can be clearly observed on the nylon membrane (Fig. 3.14-A), proving that the protein was successfully transferred to the nylon membrane. Anti-His tag was used as the primary antibody, Anti-mouse was used as the secondary antibody and Immun-Star™ AP Substrate was chosen as the chemiluminescent agent. Chemiluminescence showed that positive control with expected size 49.3 KDa (Fig. 3.14-B). The protein bands corresponding to PKS-NRPS and ER in the *BJ-EP67-RBL18B*, *BJ-EP67-*

*PGK1*, *BJ-EP67-TEF1* and *S. cerevisiae* BJ5464-NpgA were not observed (Fig. 3.13-B). Western blot results indicated that proteins was not expressed in *S. cerevisiae* BJ5464-NpgA in spite of promoter efficiency.

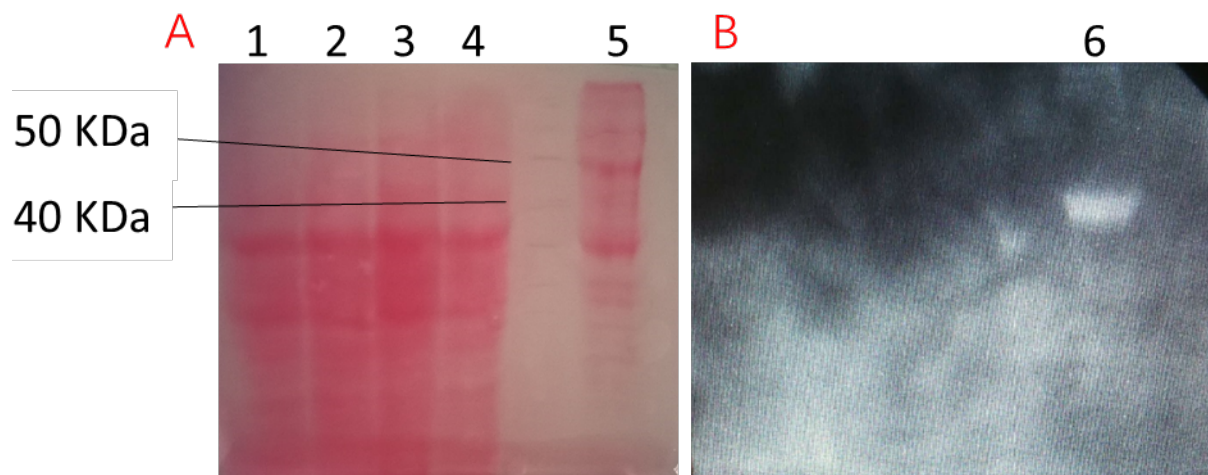


Fig.3.14 Nylon membrane stained with Ponceau Red (A-1: BJ5464-NpgA, A-2: BJ-EP67-RBL18B, A-3: BJ-EP67-PGK1, A-4: BJ-EP67-TEF1, A-5: PrCH), nylon membrane under chemiluminescent (B-6: PrCH).

We also extracted the plasmids in *BJ-EP67-RBL18B*, *BJ-EP67-PGK1*, *BJ-EP67-TEF*. Plasmids containing the genes coding PKS-NRPS and ER can replicate in *S. cerevisiae* BJ5464-NpgA was identified by PCR. The reasonable explanation is that *S. cerevisiae* BJ5464-NpgA is not a suitable host for the expression of gene originating from *S. zeae* due to codon preference. RT-PCR will be used to see if the gene is transcribed, synthetic gene to change the codons to make it ideal for yeast expression is a feasible way to solve this problem.

## **Chapter IV: Analysis of other gene clusters**

## 4.1 Introduction

Among the always increasing number of sequenced fungal genomes, hundreds of encoding PKS-NRPS genes can be identified. However, only few of them are linked to the production of a metabolite. Therefore, BGCs involving PKS-NRPS are an untapped and attractive source to discover new metabolites with potential high chemical complexity and promising bioactivities, as well as new enzymatic functions.

Beside the biosynthetic pathway of pyrrocidines described in chapter 2, we identified two other putative BGCs *pra* and *prb* involving PKS-NRPS in *S. zeae* and located in scaffold 21 and 47, respectively. Since we already inactivated genes belonging to these BGC that could impair the production of associated metabolites, it was interesting to further investigate these pathways and identify the compounds produced.

## 4.2 Bioinformatics analysis of BGC involving *prbA* from scaffold 47

Ten genes were described in the *prb* gene clusters by annotation of the scaffold 47. We identified PrbF as the *trans*-acting enoyl reductase (ER) associated to the hybrid PKS-NRPS PrbA, PrbB as a NADPH dehydrogenase, PrbC and PrbF homologous to cytochrome P450. PrbD appears as a short chain dehydrogenase, PrbI as a methyltransferase, PrbG as a transporter. PrbH and PrbJ are two putative transcription factors (Fig. 4.1).

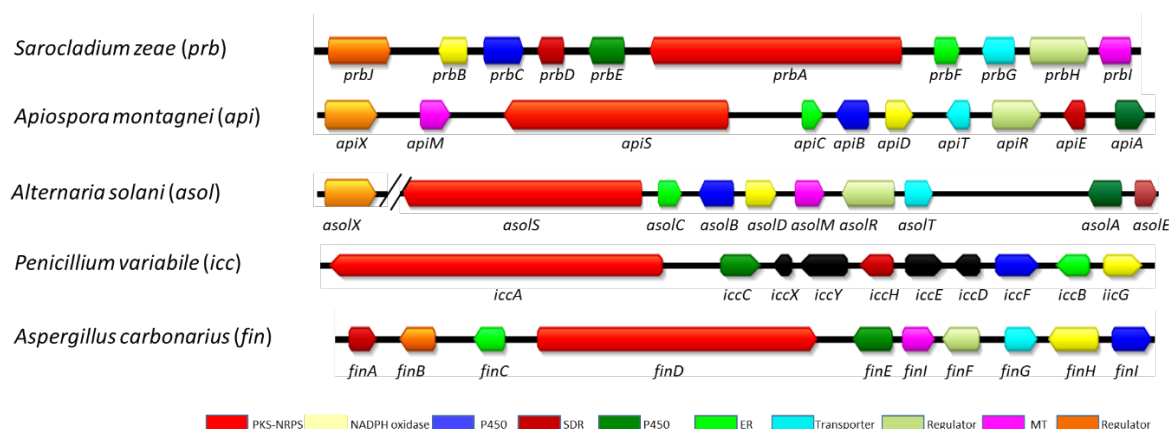


Fig. 4.1 The *prb* gene cluster and its homologues: *api*, *asol*, *icc* and *fin* gene clusters.

Thanks to bioinformatics comparisons, the gene cluster *asol* characterized in *Alternaria solani* was found to be similar to *prb* gene cluster (Fig 4.1). The metabolites produced by *asol* gene cluster were identified by Oikawa and co-workers with reconstitution of the pathway in *A. oryzae*.<sup>152</sup> Indeed, despite the apparent expression of the genes of the cluster detected by RT-

PCR in *Alternaria solani*, no metabolite was identified under various culture conditions and thus the pathway remained cryptic. Heterologous co-expression of PKS-NRPS (*asolS*), ER (*asolC*) and cytochrome P450 (*asolB*) afforded the *trans*-decaline didymellamide B<sup>222</sup>. However in this study, the authors only partially reconstituted the pathway since four other genes coding tailoring enzymes *AsolA*, *D*, *E*, *M* (for putative oxidoreductases and methyltransferase) are present in the BGC. Didymellamide B was originally isolated from the fungus *Stagonosporopsis cucurbitacearum*, with more complex analogues like didymellamide A, C and D (Fig. 4.2).<sup>223</sup> This suggests that the didymellamide B is probably modified in the producing fungus into other more complex metabolites by *AsolA*, *D*, *E*, *M* whose role is very likely to break the aromaticity of the phenol moiety and functionalize it.

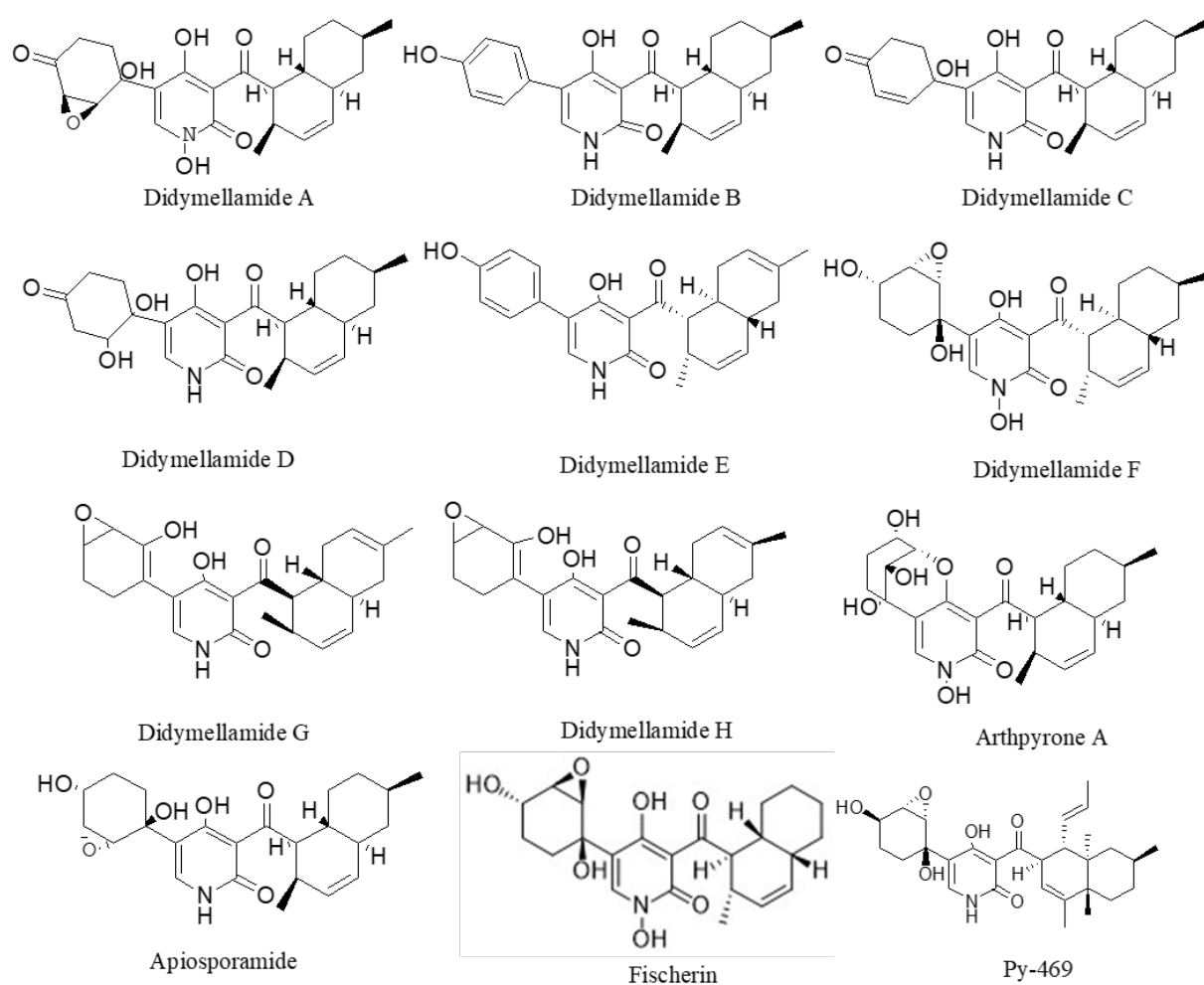


Fig. 4.2 Didymellamide compounds and their analogues

Genome mining of the *asol* gene clusters revealed the similar *api* cluster in the fungus *Apiospora montagnei* NRRL25634 which was known to produce the antifungal agent apiosporamide. Both BGCs contain ten genes with identical functions but different arrangements.<sup>222</sup> Because of the structure similarity between apiosporamide and didymellamide B, it has been proposed that apiosporamide was biosynthesized by the *api* gene cluster.<sup>222</sup>

The genes in *prb* cluster share at least 50 % sequence identity with those of *asol* (Fig. 4.3) excepted for *prbJ* which has no equivalent in the published *asol* cluster annotation. However, we identify in *A. solani* an homologous gene, named *asolX*, in another genomic scaffold probably contiguous to the cluster. These similarities suggest that *prb* cluster produce didymellamides or analogues.

Protein	Size(aa)	Identity percent compared to Asol	NCBI accession	Identity percent compared to Api	Predicted function
PrbA	3676	57%/AsolS	BBJ34510.1	60%/ApiS	PKS-NRPS
PrbB	432	62%/AsolD	BBJ34513.1	71%/ApiD	NADPH dehydrogenase
PrbC	582	68%/AsolB	BBJ34512.1	75%/ApiB	P450
PrbD	349	55%/AsolE	BBJ34518.1	52%/ApiE	SDR
PrbE	511	72%/AsolA	BBJ34517.1	70%/ApiA	P450
PrbF	372	71%/AsolC	BBJ34511.1	72%/ApiC	ER
PrbG	453	66%/AsolT	BBJ34516.1	66%/ApiT	Transporter
PrbH	826	52%/AsolR	BBJ34515.1	58%/ApiR	Transcription regulator
PrbI	384	54%/AsolM	BBJ34514.1	57%/ApiM	O-MT
PrbJ	933	46%/AsolX		60%/ApiX	Transcription regulator

Fig. 4.3 Proposed function of the proteins encoded by *prb* gene cluster.

To establish the link between this cluster and such metabolites, we compared the LC-MS metabolic profiles from  $\Delta prbA$  *S. zeae* mutant and wild type *S. zeae* obtained from a thirty days culture on maize medium (Fig. 4.4). The chromatograms do not show any differences in the production of metabolites.

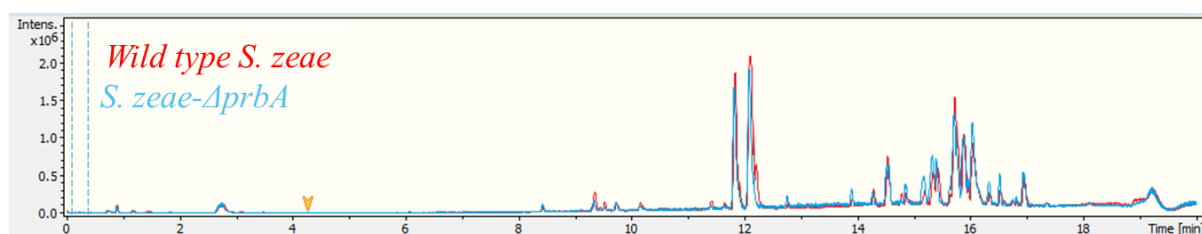


Fig 4.4 Superimposed Base Peak Chromatograms of the metabolic extract LC-MS analysis from  $\Delta prbA$  mutant and WT *S. zeae*.

We performed a more precise analysis of the data with the webserver *XCMS-online*. This tool allows accurate comparisons of sets of LC-MS data obtained from replicates of two culture conditions and lists the metabolites up- and down-regulated between these conditions. We more

precisely looked at the metabolites which were strongly down-regulated in the mutant compared to the wild type *S. zeae*. However, none of the compounds of the didymellamide family (Tab. 4.1) could be identified among the down-regulated metabolites. This suggests that the *prb* cluster remains silent under laboratory culture conditions.

Known analogs of didymellamide compounds	Formula	[M+H] <sup>+</sup> ( <i>m/z</i> )
didymellamide A	C <sub>24</sub> H <sub>29</sub> NO <sub>7</sub>	444.2017
didymellamide B	C <sub>24</sub> H <sub>27</sub> NO <sub>4</sub>	394.2013
didymellamide C	C <sub>24</sub> H <sub>29</sub> NO <sub>5</sub>	412.2119
didymellamide D	C <sub>24</sub> H <sub>31</sub> NO <sub>6</sub>	430.2224
didymellamide E	C <sub>24</sub> H <sub>25</sub> NO <sub>4</sub>	392.1856
didymellamide F	C <sub>24</sub> H <sub>31</sub> NO <sub>7</sub>	446.2173
didymellamide G	C <sub>24</sub> H <sub>27</sub> NO <sub>5</sub>	410.1962
didymellamide H	C <sub>24</sub> H <sub>27</sub> NO <sub>5</sub>	410.1962
Arthpyrone A	C <sub>24</sub> H <sub>31</sub> NO <sub>7</sub>	446.2173
Apiosporamide	C <sub>24</sub> H <sub>31</sub> NO <sub>6</sub>	430.2224
Fischerin	C <sub>23</sub> H <sub>29</sub> NO <sub>7</sub>	432.2022
Py-469	C <sub>27</sub> H <sub>35</sub> NO <sub>6</sub>	470.2543

Table. 4.1 List of the didymellamides compounds and analogues with their corresponding [M+H]<sup>+</sup> molecular ion mass.

In order to identify the metabolites produced by *prb* cluster we tried to activate the pathway by expressing the transcription factors. Two transcriptional regulators were identified in the cluster. Using BLAST analysis, we found that PrbH and PrbJ are predicted to be specific transcription factors and belong to the zinc finger family (Zn(II)<sub>2</sub>Cys<sub>6</sub>).<sup>224</sup> However, PrbJ has a lower similarity with known transcriptional factors compared to PrbH. So, we selected *prbH* as a first target to activate the *prb* gene cluster. To express this transcription factor in *S. zeae*, we envisaged to substitute its promoter by those of both transcription factors found in pyrrocidine BGC, *PprcF* and *PprcJ*, since pyrrocidines is easily produced by *S. zeae* in various media (Fig. 4.5). We also selected the known constitutive promoter *PgdaA* of the gene coding glyceraldehyde-3-phosphate dehydrogenase (GpdA), which is commonly employed in fungal heterologous expression.





Fig 4.5 Promoters *PprcF* and *PprcJ* in pyrrocidine BGC.

These promoters were obtained by PCR amplification of the sequences preceding the genes (1 kb for *PprcF*, 565 bp for *PprcJ* and 1053 bp for *PgpdA*) from genomic DNA of *S. zeae*. Then, amplicons were integrated by IVA cloning into a recombination cassette including the hygromycin resistance transcription unit and flanked by two recombination arms (Fig. 4.6).

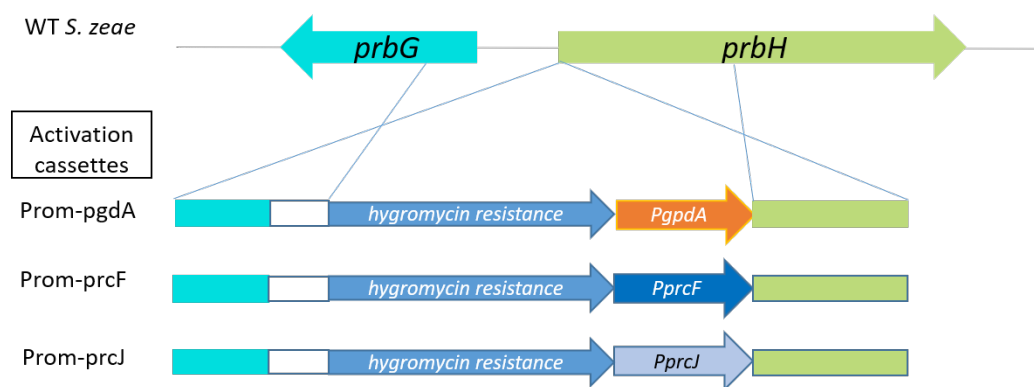


Fig. 4.6. Activation cassettes design to overexpress the transcriptional factor *prbH*. Three constructions were performed by integrating the three promoters *PgpdA*, *PprcF* and *PprcJ* downstream the transcription unit of the hygromycin resistance.

The three linear activation cassettes cloned into pUC18 were amplified by PCR and integrated into *S. zeae*, separately. Two of them, Prom-*pgdA* and Prom-*prcF*, were successfully integrated at the targeted locus thus substituting native *prbH* promoter by *PgpdA* or *PprcF*. The two strains were cultivated on maize medium and analyzed by LC-MS/MS as previously described. However, none of the  $[M+H]^+$  ions corresponding to the didymellamide family was identified in metabolite extracts of both mutants.

In order to activate this pathway in *S. zeae* further investigations will be conducted. Several options are possible. For example, an activation cassette harboring a strong promoter can be tested. The other transcriptional factor encoded in the cluster by *prbJ* could also trigger the production of the metabolites. Therefore, a similar approach could be tested on this gene.

Alternatively, we envisaged to heterologously produce the metabolites by reconstitution of the pathway in an *Aspergillus* host.

Very recently, two similar gene clusters *icc* and *fin* were found to produce the apiosporamide analogues Py-469 and fischerin, respectively (Fig. 4.1).<sup>225</sup> The biosynthetic pathway of Py-469 was characterized by heterologous expression of the *icc* gene cluster in *Aspergillus nidulans*.<sup>225</sup> This study assigned a functional role to auxiliary enzymes in the pathway. More specifically, the three proteins IccF (P450 cytochrome homologous to AsolB), IccG (NADPH dehydrogenase homologous to AsolD) and IccH (short chain dehydrogenase/reductase homologous to AsolE) convert a phenol moiety to the final epoxy-cyclohexanediol. Interestingly, these three proteins are also conserved in the *api* and *fin* cluster (Fig. 4.1). However, the epoxy-cyclohexanediols of the resulting metabolites PY-469, apiosporamide and fischerin harbor various stereochemistries, suggesting that these three enzymes have their own selectivity depending on the cluster, and that the predictions of final products of *asol* or *prb* clusters remain uncertain.

Nevertheless, a general biosynthetic pathway from *asol* BGC can be proposed as followed: (Fig. 4.7): PKS-NRPS (AsolS) and ER (AsolC) gives the tetramic acids intermediate as we deduced in chapter 2. P450 cytochrome (AsolA) expands the ring of tetramic acids to form dydimellamide B. The O-methyltransferase AsolM catalyzes the Diels-Alder reaction like the O-methyltransferase LepI<sup>226</sup> (24.2% identity to AsolM) and then P450 cytochrome AsolB controls the oxidative dearomatization and epoxidation of the phenol moiety. Finally, two reduction steps occurred on the cyclohexenone: the NADPH dehydrogenase AsolD catalyzes ene-reduction and short chain dehydrogenase AsolE reduces the ketone. This pathway is very likely similar in the case of *prb* gene cluster.

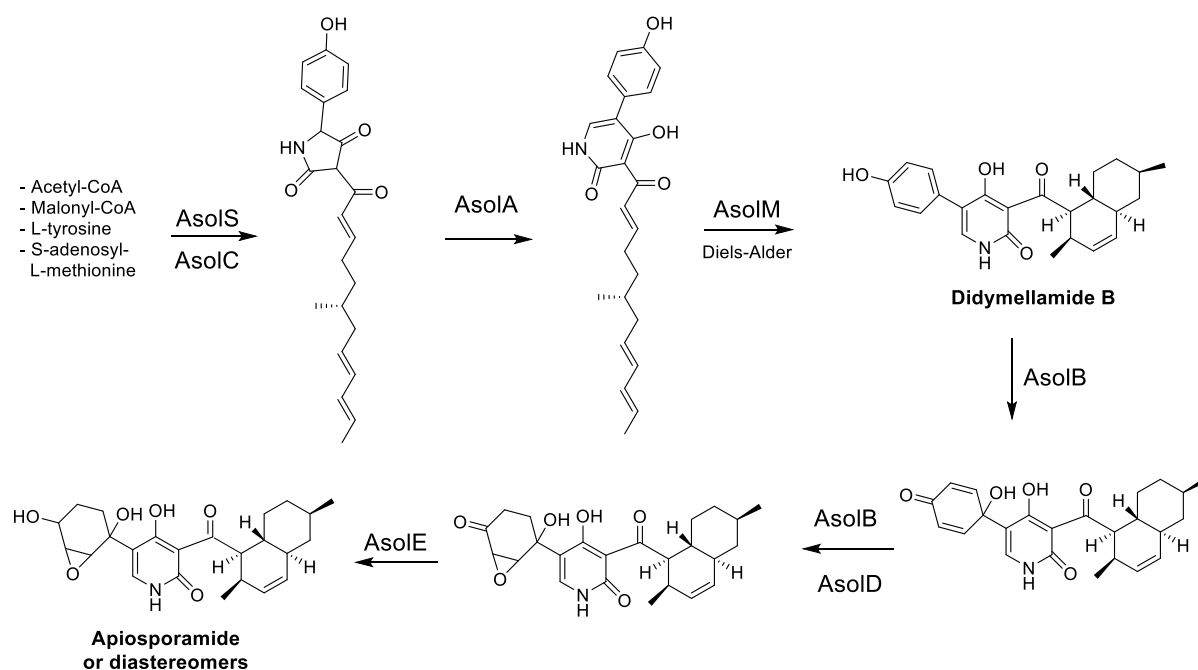


Fig.4.7 Proposed biosynthetic pathway of Apiosporamide or analogues from *asol* BGC.

### 4.3 Analysis of cluster *pra*

Fourteen genes were described in the *pra* gene cluster by annotation of the scaffold 21. We identified PraG as the *trans*-acting enoyl reductase (ER) associated to the hybrid PKS-NRPS PraA, PraB and PraC as two transcription factors, PraD as a NRPS, PraE with unknown function, PraF as a calycin, PraH is an aldolase. PraI as an O-methyltransferase. PraJ homologous to cytochrome P450, PraK appeared as a sterol-reductase. PraL is a transporter. PraM has a putative function of a short chain dehydrogenase and PraN as a  $\alpha/\beta$  hydrolase (Fig. 4.8).

Protein	Size (aa)	Predicted function	Percent identity with Known function or predicted function	NCBI accession (protein)/Species
PraA	4047	PKS-NRPS	59%/TasS	AOA0F7GFS4.1/ <i>Hapsidospora irregularis</i>
PraB	506	Transcription factor	32%/TasR	AOA0F7GH49.1/ <i>Hapsidospora irregularis</i>
PraC	635	Transcription factor	56%	CRL26838.1/ <i>Penicillium camemberti</i>
PraD	1444	NRPS	40%	KAF3402511.1/ <i>Talaromyces pinophilus</i>
PraE	271	Unknown	65%	XP_040776382.1/ <i>Cryphonectria parasitica</i>
PraF	169	Calycin	67%	KGO74914.1/ <i>Penicillium italicum</i>
PraG	388	ER	66%/TasC	AOA0F7GF15.1/ <i>Hapsidospora irregularis</i>
PraH	284	Aldolase	71%/TasA	AOA0F7GG11.1/ <i>Hapsidospora irregularis</i>
PraI	366	O-MT	43%/p hm5	AOA2Z5XAK6.1/ <i>Pyrenochaetopsis</i> sp.
PraJ	1077	P450	57%	KGO74909.1/ <i>Penicillium italicum</i>
PraK	394	Reductase	79%	MBE3049343.1/ <i>Candidatus Bathyarchaeota</i>
PraL	563	Transporter	87%	XP_006667144.1/ <i>Cordyceps militaris</i> CM01
PraM	265	SDR	99%	XP_018699599.1/ <i>Cordyceps fumosorosea</i>
PraN	445	$\alpha/\beta$ hydrolase	43%	XP_028481814.1/ <i>Paecilomyces variotii</i>

Fig. 4.8 Gene annotation of the putative *pra* biosynthetic gene cluster.

*Tas* gene clusters contain eight genes, four of which were proposed to be indispensable for the biosynthesis of Sch210971 and Sch210972 by heterologous expression in *F. heterosorum* and labeling experiments. *Tas* gene cluster shares three tailoring genes with *pra* gene cluster.

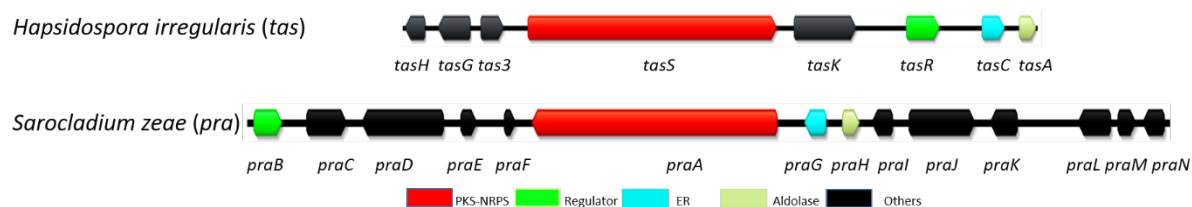


Fig. 4.9 Comparison of the *tas* and *pra* gene clusters from *H. irregularis* and *S. zeae*, respectively.

Several attempts were performed to knock out the gene *praA* but this remained unsuccessful. Because PraA requires the presence of its partnering ER (PraG) to be properly functional, we targeted the deletion of *praG*, which was successful.

The metabolite profiles analysis of LC-MS experiments of the extracts from  $\Delta$ *praG* - *S. zeae* mutant and wild type was performed as previously exposed for *prb* BGC. The data did not allow us to discriminate a metabolite linked to the *pra* gene cluster suggesting that the latter remains silent in the tested culture conditions.

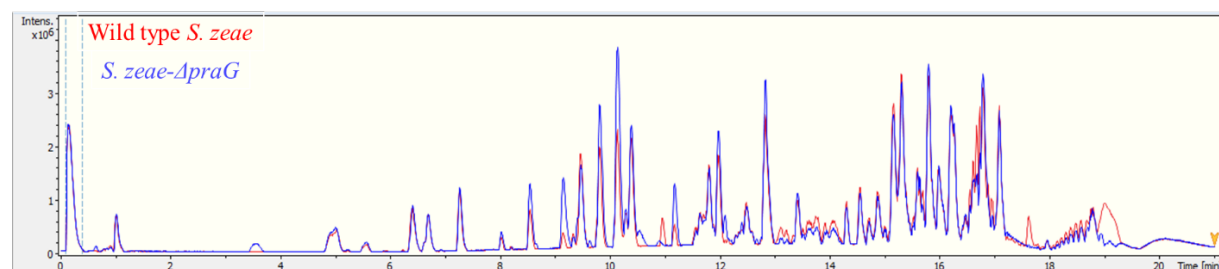


Fig. 4.10 LC-MS analysis of the metabolite extracts of  $\Delta$ *praG* mutant and wild type *S. zeae*. Superimposed base peak chromatograms (BPC).

To further investigate this *pra* encoded biosynthetic pathway, we plan to activate the transcription factor as proposed for the *prb* cluster.



## **Chapter V: Materials and Methods**

Table 5.1 Strains and plasmids used in chapter II.

Strain or plasmid	Relevant Characteristics	Source or reference
<b>Plasmids</b>		
pUC18	Cloning vector	Invitrogen
pInact- <i>prcA+B</i>	Inactivation cassette( $\Delta prcA+B$ ) cloned into pUC18	This study
pInact18- <i>praG</i>	Inactivation cassette ( $\Delta praG$ ) cloned into pUC18	This study
pInact- <i>prcB</i>	Inactivation cassette ( $\Delta prcB$ ) cloned into pUC1	This study
pInact- <i>prcD</i>	Inactivation cassette( $\Delta prcD$ ) cloned into pUC18	This study
pInact- <i>prcE</i>	Inactivation cassette( $\Delta prcE$ ) cloned into pUC18	This study
pInact- <i>prcH</i>	Inactivation cassette( $\Delta prcH$ ) cloned into pUC18	This study
pInact- <i>prcI</i>	Inactivation cassette ( $\Delta prcI$ ) cloned into pUC18	This study
pInact- <i>prcX</i>	Inactivation cassette ( $\Delta prcX$ ) cloned into pUC18	This study
<b>Strains</b>		
<i>Sarocladium zeae</i> NRPL13540		
<i>Sarocladium zeae</i> - $\Delta prcA+B$	<i>Sarocladium zeae</i> with deletion of the gene of <i>prcA</i> and <i>prcB</i>	This study
<i>Sarocladium zeae</i> - $\Delta praG$	<i>Sarocladium zeae</i> -with deletion of the gene of <i>praG</i>	This study
<i>Sarocladium zeae</i> - $\Delta prcB$	<i>Sarocladium zeae</i> with deletion of the gene of <i>prcB</i>	This study
<i>Sarocladium zeae</i> - $\Delta prcD$	<i>Sarocladium zeae</i> with deletion of the gene of <i>prcD</i>	This study
<i>Sarocladium zeae</i> - $\Delta prcE$	<i>Sarocladium zeae</i> with deletion of the gene of <i>prcE</i>	This study
<i>Sarocladium zeae</i> - $\Delta prcH$	<i>Sarocladium zeae</i> with deletion of the gene of <i>prcH</i>	This study
<i>Sarocladium zeae</i> - $\Delta prcI$	<i>Sarocladium zeae</i> with deletion of the gene of <i>prcI</i>	This study
<i>Sarocladium zeae</i> - $\Delta prcX$	<i>Sarocladium zeae</i> with deletion of the gene of <i>prcX</i>	This study
<i>Sarocladium zeae</i> - <i>prcE</i> -G418	<i>Sarocladium zeae</i> - $\Delta prcE$ with restore the function of <i>prcE</i>	This study
<i>E. coli</i> , TOP10	Genotype: <i>F</i> - <i>mcrA</i> $\Delta$ ( <i>mrr</i> - <i>hsdRMS</i> - <i>mcrBC</i> ) $\phi$ 80 <i>lacZ</i> <i>AM15</i> $\Delta$ <i>lacX74</i> <i>nupG</i> <i>recA1</i> <i>araD139</i> $\Delta$ ( <i>ara-leu</i> )7697 <i>galE15</i> <i>galK16</i> <i>rpsL</i> ( <i>StrR</i> ) <i>endA1</i> $\lambda$ -	Invitrogen
<i>E. coli</i> , codon plus (DE3)	<i>F</i> - <i>ompT</i> <i>gal</i> <i>dcm</i> <i>lon</i> <i>hsdSB</i> ( <i>rB</i> - <i>mB</i> -) $\lambda$ (DE3 [ <i>lacI</i> <i>lacUV5</i> -T7p07 <i>ind1</i> <i>sam7</i> <i>nin5</i> ]) [ <i>malB</i> +] <i>JK</i> -12( $\lambda$ S)	Invitrogen

Table 5.2 Strains and plasmids used in chapter III.

Strain or plasmid	Relevant Characteristics	Source or reference
<b>Plasmids</b>		
pGG-Venus	Expression plasmid (RPL18B-Venus) assembled by Golden Gate	This study
pGG-mRuby	Expression plasmid (RPL18B-mRuby) assembled by Golden Gate	This study
pGG-ER-RPL18B	Expression plasmid (RPL18B-ER) assembled by Golden Gate	This study
pGG-EP-RPL18B	Expression plasmid (RPL18B-ER) assembled by Golden Gate	This study
pGG-P67-RPL18B	Expression plasmid (RPL18B-PKS-NRPS) assembled by homologous recombination	This study
pGG-ER-PGK1	Expression plasmid (PGK1-ER) assembled by Golden Gate	This study
pGG-EP-PGK1	Backbone plasmid (PGK1-ER) assembled by Golden Gate	This study
pGG-P67-PGK1	Expression plasmid (PGK1-PKS-NRPS) assembled by homologous recombination	This study
pGG-ER-TEF1	Expression plasmid (TEF1-ER) assembled by Golden Gate	This study
pGG-EP-TEF1	Backbone plasmid (TEF1-ER) assembled by Golden Gate	This study
pGG-P67-TEF1	Expression plasmid (TEF1-PKS-NRPS) assembled by homologous recombination	This study
<b>Strains</b>		
<i>BJ5464-NpgA</i>	<i>Saccharomyces cerevisiae</i> used for heterologous expression host	
<i>BJ-Venus</i>	Expression of Venus in <i>BJ5464-NpgA</i>	This study
<i>BJ-mRuby</i>	Expression of mRuby in <i>BJ5464-NpgA</i>	This study
<i>BJ-ER-PRL18B</i>	Expression of pGG-ER-PRL18B in <i>BJ5464-NpgA</i>	This study
<i>BJ-EP67-RPL18B</i>	Expression of pGG-P67-RPL18B in <i>BJ-ER-PRL18B</i>	This study
<i>BJ-ER-PGK1</i>	Expression of pGG-ER-PGK1 in <i>BJ5464-NpgA</i>	This study
<i>BJ-ER-TEF1</i>	Expression of pGG-P67-PGK1 in <i>BJ-ER-PGK1</i>	This study
<i>BJ-P67-PGK1</i>	Expression of pGG-ER-TEF1 in <i>BJ5464-NpgA</i>	This study
<i>BJ-P67-TEF1</i>	Expression of pGG-P67-TEF1 in <i>BJ-ER-TEF1</i>	This study



Table 5.3 Strains and plasmids used in chapter IV.

Strain or plasmid	Relevant Characteristics	Source or reference
Plasmids		
pUC18-gg	Ectopic insertion cassette (pgpdA-GFP) cloned into pUC18	This study
pUC18-ag	Ectopic insertion cassette (pamyB-GFP) cloned into pUC18	This study
pUC18-gm	Ectopic insertion cassette (pgpdA-mRuby) cloned into pUC18	This study
pUC18-am	Ectopic insertion cassette (pamyB-mRuby) cloned into pUC18	This study
pUC18-av	Ectopic insertion cassette (pamyB-Venus) cloned into pUC18	This study
pUC18-gv	Ectopic insertion cassette (pgpdA-Venus) cloned into pUC18	This study
pUC18-tfg	Ectopic insertion cassette (pgpdA- <i>prbH</i> ) cloned into pUC18	This study
pUC18-gf	Recombination cassette (pgpdA) cloned into pUC18	This study
pUC18-tf1f	Recombination cassette (pTF1) cloned into pUC18	This study
pUC18-tf2f	Recombination cassette (pTF2) cloned into pUC18	This study
Strains		
<i>S. zeae</i> -gg	<i>S. zeae</i> with ectopic insertion of pgpdA-GFP	This study
<i>S. zeae</i> -ag	<i>S. zeae</i> with ectopic insertion of pamyB-GFP	This study
<i>S. zeae</i> -gm	<i>S. zeae</i> with ectopic insertion of pgpdA-mRuby	This study
<i>S. zeae</i> -am	<i>S. zeae</i> with ectopic insertion of pamyB-mRuby	This study
<i>S. zeae</i> -av	<i>S. zeae</i> with ectopic insertion of pamyB-Venus	This study
<i>S. zeae</i> -etfg	<i>S. zeae</i> with ectopic insertion of pgpdA-PrbH	This study
<i>S. zeae</i> -gf	<i>S. zeae</i> with the gene of gpdA inserted before <i>prbH</i>	This study
<i>S. zeae</i> -tf1f	<i>S. zeae</i> with the gene of tf1f inserted before <i>prbH</i>	This study

## 01. General growth conditions and reagents

*Sarocladium zeae* and its mutants were cultivated and maintained on 1% Malt agar medium at room temperature. *E. coli* was cultured on LB medium (10 g tryptone, 5 g yeast extract, 10 g NaCl in 1 L MilliQ water) for further being used for plasmid construction and amplification. Each bottle containing of 50 g corn, oat, rice or wheat mixed with 50 mL water was sterilized for OSMAC experiment. *S. cerevisiae* BJ5464-NpgA were cultivated and maintained on YPD medium (1% yeast extract, 2% peptone, 2% glucose, 2% agar for solid medium) for further used for heterologous expression of PKS-NRPS-67. The plasmid pSN44 harboring hygromycin resistance gene was a kind gift of Muriel VIAUD. Hygromycin (Hyg<sup>+</sup>) 50 µg/mL was used to select the mutants of *Sarocladium zeae*, verified by PCR. Ampicillin (Amp<sup>+</sup>) 100 µg/mL, chloramphenicol (Cam<sup>+</sup>) 35 µg/mL, kanamycin (Kan<sup>+</sup>) 50 µg/mL were used to select *E. coli* Top 10 harboring the right construction of plasmids. PCR were applied using Q5 High-Fidelity DNA polymerase (NEB) and Phusion High-Fidelity DNA polymerase

(Thermo Fischer Scientific™) according to the manufacturer's recommendations. Restriction enzymes and T4 ligase were purchased from Thermo Fischer Scientific™. RNA extraction and DNA removal kits were obtained from Invitrogen (Thermo Fischer Scientific™). RT-PCR kits was achieved from Takara Bio.

## 02. Preparation of competent cells

Top10 was cultured on LB liquid medium at 37 °C overnight. 150 µL of the pre-culture solution was inoculated to 15 mL LB. When the OD reached 0.6 after about three at 37 °C, the cells were collected by centrifugation at 4 °C, 2000 g for 10 min. The pellet was resuspended in TSS solution (PEG 8000 (5 g), 1 M MgCl<sub>2</sub> (1.5 mL) and DMSO (2.5 mL) mixed with LB (final volume 50 mL).) and kept on ice. The cell suspension was aliquoted into 90 µL each tube and frozen with liquid nitrogen to be stored at -80 °C.

## 03. General PCR Steps

Phusion DNA polymerase was used for most of the PCR processes. When the amplification carried by Phusion was not efficient, Phusion should be replaced by Q5 DNA polymerase. For PCR, each 20 µL reaction mixture contains 0.4 Unit Phusion polymerase/Q 5 polymerase, 4 µL HF buffer/Q 5 reaction buffer, 200 µM dNTP, 0.5 µM forward and reverse primers and 1 ng DNA template. The annealing temperature were calculated by the online website Phusion TM calculator/Q 5 TM calculator. PCR were programmed according to the following protocol in Table 5.4.

Table.5.4 PCR system of Phusion/Q 5.

Initial Denaturation	98 °C	30 s
Denaturation	98 °C	10 s
Annealing	50-72 °C	20 s
Extension	72 °C	25 s/kb
Final Extension	72 °C	10 min
Final Hold	10 °C	∞

## 04. Restriction enzyme digestion

For EcoR I and Hind III digestion, 1 µL enzyme was used for 1 µg plasmid digestion at 37 °C for 1 hour. For Dpn I digestion, 1 µL enzyme was used for 0.2 µg PCR products at 37 °C for 30 minutes.

## 05. Agarose electrophoresis

5xTAE buffer (Tris 242 g, CH<sub>3</sub>COOH 57.1 mL, Na<sub>2</sub>EDTA.2H<sub>2</sub>O) was used as an electrophoresis buffer. The gel was made of 1xTAE buffer and 0.8% agarose. Agarose electrophoresis was executed at 110 V for 25 min.

## 06. Vector Assembly cloning by in vivo assembly (IVA) in *E-coli*

IVA cloning technique was exploited to construct the plasmid *in vivo* of the *E. coli* Top10. The competent cells were treated with 10 µL of KCM solution and incubated on ice for 10 min. After that, interested DNA was added to the cells, and the mixture was incubated on ice for 20 minutes followed by heat shock at 42 °C for 90 seconds and maintained for 15 minutes at 4 °C. Finally, 100 µL of SOC medium were added to the tube and the bacteria were incubated for 1 hour in 37 °C with shaking and then plated on LB medium with antibiotics.

KCM solution consists of 0.1 M KCl, 0.03 M CaCl<sub>2</sub>, 0.05 M MgCl<sub>2</sub>. SOC medium contains 0.5% (w/v) yeast extract, 2% (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, 20 mM MgSO<sub>4</sub> and 0.4% glucose.

## 07. Plasmid purification

1.5 mL of *E. coli* bacteria were centrifuged for 30 s at 11,000 g to harvest the cells. The pellet was re-suspended in 250 µL of solution P1 containing 3.5 µL RNase (10 mg/mL), followed by 250 µL solution P2 and immediately up- down inverted several times. 300 µL solution P3 was added to the suspension and inverted several times, centrifugation at 16000 g for 15 min and transferred the supernatant to a 1.5 mL Eppendorf tube. 560 µL isopropanol was added to the tube and inverted several times, then centrifuged at 16000 g for 10 min. The supernatant was discarded and the pellet was washed with 70% acetic acid. Centrifugation was operated at 16000 g, followed by discarding the supernatant, the pellet was dried at 37 °C for 10-15 min, 30 µL sterile H<sub>2</sub>O were used to re-suspend the pellet and the plasmids were stored at -20 °C. Solution P1 contains 50 mM Tris-HCl (pH=8.0) and 10 mM EDTA with adjusting pH to 8.0. Solution P2 contains 20 mM NaOH and 1% SDS. Solution P3 consists of 3 M KOAc with adjusting pH to 5.5 with acetic acid. All the solutions were autoclaved and stored at room temperature.

## 08. PCR products clean-up

NT1 buffer was added with twice the volume of PCR products. The mixture was transferred to a column. The PCR product was combined to the column after centrifugation at 11,000 g for 30 s. The column was treated twice with 700  $\mu$ L NT3 and then centrifugation at 11,000 g for 30 s. The column was dried by idling at 11,000 g for 1 min. 30  $\mu$ L Elution Buffer was added to the column and incubated for 5 minutes at 70 °C by dry-bath. Clean PCR products were obtained by centrifugation at 11, 000 g for 2 min.

#### **09. Isolation of the genomic DNA of *S. zeae*<sup>227</sup>**

100 mg mycelium of *S. zeae* from 7 days liquid culture in 1% Malt medium were collected in 1.5 mL microtubes for centrifugation (22100 g, 10 min). The pellet was homogenized on a cell homogenizer sample preparation system with a speed of 6 m/s in 400  $\mu$ L of buffer (0.4 M NaCl, 10 mM Tris-HCl pH 8.0 and 2 mM EDTA pH 8.0). 40  $\mu$ L SDS (w/v 20%) and 8  $\mu$ L proteinase K (20 mg/mL) were added to the tubes. The samples were incubated at 55 °C overnight and 300  $\mu$ L of 6 M NaCl were added after vortexing for 30 s. The genomic DNA was precipitated by adding an equal volume of isopropanol and stored at -20 °C for at least 1 h. Samples were centrifuged for 20 min at 6200 g, 4 °C. The pellet was rinsed with 500  $\mu$ L of 70% ethanol. DNA was dried at 37 °C for 1 h and dissolved in 300  $\mu$ L sterile Milli-Q H<sub>2</sub>O.

#### **10. Preparation of the protoplasts of *S. zeae***

Fresh spores were obtained by scraping the surface of the plate culture with 5 mL of Milli-Q H<sub>2</sub>O. The spores were filtered through 10  $\mu$ M Nylon cloth and numbered on Malassez cell. Malt liquid medium 50 mL was inoculated with 10<sup>7</sup> spores and germinated at 25 °C, 110 rpm overnight. Mycelium was harvested by centrifugation at 1789 g for 15 min. and washed with 25 mL of autoclaved Milli-Q H<sub>2</sub>O. The suspension was centrifuged for 15 min at 2000 g, 4 °C and the supernatant was discarded. The pellet was recovered and mixed with 1 mL of autoclaved Milli-Q H<sub>2</sub>O and transferred to a 2 mL microtube to be centrifuged for 5 min at 6200 g, 4 °C. The pellet was resuspended in 600  $\mu$ L of sterilized lysing solution composed of 60 mg/mL lysing enzyme in KK solution (0.6 M KCl, 25 mM KH<sub>2</sub>PO<sub>4</sub>, adjusting pH to 5.8 by 1 M KOH). The mycelium was incubated for 2 h at 25 °C under mild agitation (Fig. 5.1). The protoplast suspension was diluted with 10 mL of KK solution filtered through 1  $\mu$ M bolting cloth, and the solution was centrifuged for 15 min at 1000 g, 4 °C. The pellet was resuspended with 2 mL of KCa solution (0.6 M KCl, 50 mM CaCl<sub>2</sub>) and transferred to a 2 mL microtube to be centrifuged

for 5 min at 1000 g, 4 °C. The protoplasts were resuspended in a final volume of 200  $\mu$ L of KCa solution and kept on ice.

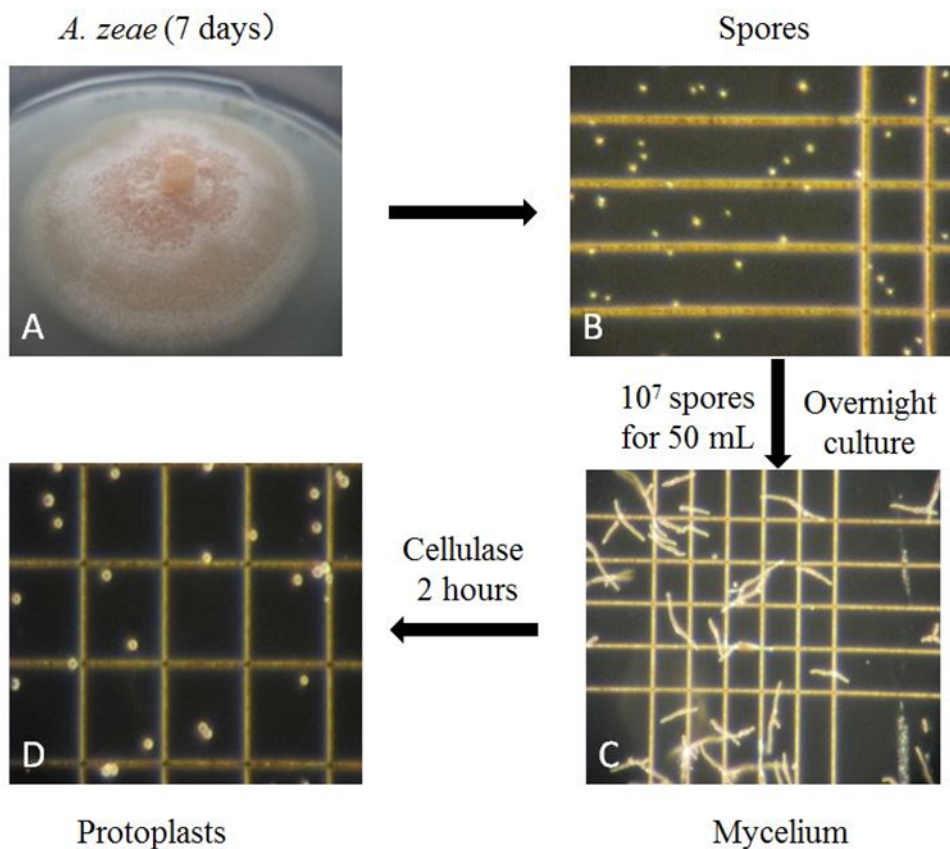


Fig. 5.1 Preparation of the protoplasts.

## 11. Transformation of *S. zeae*

The protoplasts were diluted to 100-fold suspension and counted using Malassez cell. Six plates were prepared with each plate including 12.5 mL MSHA (0.6% pure agar, 5 mM HEPES, 0.6 M sucrose and 400 mL Milli-Q water by adjusting pH to 7) and 12.5  $\mu$ L 2.5% yeast extract. Protoplasts, TE CaCl<sub>2</sub> (1 mM EDTA, 10 mM CaCl<sub>2</sub> and 10 mM Tris-HCl), DNA (Two split inactivation cassette) and PEG 4000 (PEG PH 7.5 25% PEG, 50 mM CaCl<sub>2</sub>, 10 mM Tris-HCl) were mixed in different proportions and kept on ice for thirty minutes. 170  $\mu$ L PEG4000 was added to each tube and kept on room temperature for 10 min. KCa = 0.6 M KCl 50 mM CaCl<sub>2</sub> were added to each tube to a final volume of 1 mL and mixed them well with 11 mL MSHA, each 4 mL mixture were added to each plate before solidification. The plates were kept in 25 °C for 16 h for protoplast regeneration, then 4 mL MSHA with 20  $\mu$ L hygromycin were added to each plate except for one plate used as the control.

After five days, the mutants were transferred to small plates with 3 mL malt agar containing hygromycin. The small plates were cultured for another 5 days, then the spores were harvested by scratching the plate with a loop and 1 mL of autoclaved Milli-Q water. 10  $\mu$ L suspension were transferred to 3 mL Malt liquid with hygromycin for DNA extraction and PCR verification, the rest were used for preservation. The mutants were cultured for about five days. After centrifugation, DNA extraction were performed to validate the mutants.

## **12. The analysis of the metabolites**

Spores were prepared as before; About  $10^7$  spores were inoculated on PDA plate (20 mL) and cultured at 25 °C for 7 days or inoculated on 50g rice medium at 25 °C for one month. The culture was extracted with ethyl acetate, agitating on a magnetic stirrer for three hours at room temperature. The extracts were concentrated under a vacuum. Samples were solubilized in acetonitrile to a final concentration of 1mg/mL for LC-MS/MS-MS analyses, which were performed on an ultra-high performance LC system (ULTimate 3000 RSLC, Thermo Scientific) coupled to a high-resolution electrospray ionization-quadrupole-time of flight (ESI-Q-TOF) mass spectrometer (MaXis II ETD, Bruker Daltonics). An Acclaim RSLC Polar Advantage II column (2.2  $\mu$ M, 2.1  $\times$  100 mm, Thermo Scientific) was used for LC separation with a flow rate of 0.3 mL/min. Chromatographic separation was carried out with a linear gradient of MeCH-H<sub>2</sub>O within 20 min, 0.1% formic acid was added to both solvents. Mass dates were exported as xzXML Files and uploaded to XCMS online for pairwise comparison to discover subtle metabolic profile changes between the mutant and wild type.

## **13. Isolation of pyrrocidines in WT and accumulation metabolites in mutants**

For the isolation of the compounds, about 12 g of crude extracts were dissolved by CH<sub>2</sub>Cl<sub>2</sub>. Initial purification was performed by flash chromatography on silica gel (The amount of silica gel was 40-60 times the mass of the crude extract. Eluted fractions started with 100% cyclohexane and 0% ethyl acetate. The cyclohexane content was decremented by a ratio of 10%, until eluting with 100% ethyl acetate and 0% cyclohexane. 90% CH<sub>2</sub>Cl<sub>2</sub>/10% EtOH, 80% CH<sub>2</sub>Cl<sub>2</sub>/20% EtOH, 70% CH<sub>2</sub>Cl<sub>2</sub>/30% EtOH, 100% acetone/0% MeOH and 60% acetone/40% MeOH were added to the column respectively in sequence for elution according to the TLC results after each step. Each fraction was analyzed by LC-MS-MS with an Acclaim RSLC Polar Advantage II column (2.2  $\mu$ M, 2.1  $\times$  100 mm, Thermo Scientific). The further purification was performed by preparative HPLC with a Kinetex<sup>®</sup> 5 $\mu$ M C18 LC Column (Elution with different

gradients of water and acetonitrile). Thin layer chromatography was adopted for final purification if necessary. The chemical structures were elucidated by extensive NMR experiments. 1D and 2D NMR spectra were acquired on a Bruker Avance 400 MHz spectrometer/Bruker AVANCE III HD 600 MHz spectrometer equipped with a triple resonance TCI cryoprobe at 298 K.

#### **14. Link the scaffold 67 to scaffold 152**

The reason for obtaining scaffolds of different sizes is that the GC content often leads to the sequencing to stop. Therefore, to obtain high-quality PCR products for sequencing, a high GC buffer was used and the extension time of each cycle was increased to 3min compared to the normal PCR.

#### **15. RNA extraction of *S. zeae***

10  $\mu$ M Nylon cloth was cut to the size of the plate; after autoclave, the Nylon cloth was tiled on a PDA plate surface, about  $10^7$  spores were inoculated upon the Nylon cloth for one week. The mycelium was scraped and quickly frozen by liquid nitrogen, then kept at  $-80^{\circ}\text{C}$  for RNA extraction. The mortar and pestle were kept at  $200^{\circ}\text{C}$  for at least 12 h after autoclaving to remove the RNase. Then after cooling down, they were kept at  $-80^{\circ}\text{C}$ . The desktop was washed with DEPC water, 100 mg samples were mixed with 1 mL RLT buffer harboring 10  $\mu$ L  $\beta$ -ME. The mixture was thoroughly ground in the environment of liquid nitrogen. Then transferred to an RNase-free Eppendorf tube and solubilized with 450  $\mu$ L RLT buffer. After vortexing, the lysate was transferred to the purple column; the tube was centrifuged at maximum speed for 2 min. The supernatant was transferred to a new Eppendorf tube and half volume of ethanol was added to the tube, mixing well by pipetting. The mixture was transferred to a pink column immediately and centrifugation at 9000 g for 15 s. 700  $\mu$ L RW1 buffer was added to the column and centrifuged at 9000 g for 15 s. 500  $\mu$ L RPE buffer was added to the tube, followed by centrifugation at 9000 g for 15 s. Another 500  $\mu$ L RPE buffer was repeated at the same condition except for 2 min. The pink column was transferred to a clean Eppendorf tube after tube idling for 1 min at maximum speed. 50  $\mu$ L RNase was used to solubilize the RNA by adding to the column with centrifugation at 10000 g for 1 min, repeated as the previous operation.

#### **16. Removal of DNA from RNA and RT-PCR**

44  $\mu\text{L}$  of RNA was obtained from the last procedure mixed with 5  $\mu\text{L}$  of 10x DNase Max buffer and 1  $\mu\text{L}$  DNase Max enzyme. The mixture was treated at 37  $^{\circ}\text{C}$  for 25 min. 5  $\mu\text{L}$  DNase Max Removal Resin was added to the tubes and kept at room temperature for 10 minutes, blending every two minutes. After centrifugation at 13000 g for 1 min, the supernatant was transferred to a new tube to obtain the clean RNA.

1  $\mu\text{L}$  of 50  $\mu\text{M}$  dTT, 1  $\mu\text{L}$  of 10 mM dNTP, 9  $\mu\text{L}$  RNA (about 1  $\mu\text{g}$ ), and 2  $\mu\text{L}$  DEPC-treated water were well mixed and then incubated at 65  $^{\circ}\text{C}$  for 5 min, followed by kept on the ice at least for 1 min. 4  $\mu\text{L}$  5x SSIV buffer, 1  $\mu\text{L}$  100 mM DTT, 1  $\mu\text{L}$  RNase inhibitor were added to the mixture and kept at 55  $^{\circ}\text{C}$  for 10 min followed by treated at 80  $^{\circ}\text{C}$  for 10 min.

### **17. Extraction of plasmids provided in the yeast toolkit**

Macherey-Nagel™ NucleoSpin Plasmid QuickPure™ Kit was used for the plasmid extraction provided in the yeast toolkit. All required plasmids were recovered on plates with appropriate corresponding antibiotics and then inoculated into 4 mL of liquid medium containing relevant antibiotics for overnight culture. 2 mL culture was centrifuged at 5900 g for 5 min. The pellet was re-suspended in 250  $\mu\text{L}$  Buffer A1. 250  $\mu\text{L}$  Buffer A2 was then added to the mixture and kept at room temperature for 5 min. Then 300  $\mu\text{L}$  blue Buffer A3 was added to the solution mixed well until the solution became clear. The mixture was centrifuged for 10 min at 11,000 g at room temperature. The supernatant was transferred to the column for DNA binding and then centrifuged for 1 min at 11,000 g at room temperature. The column was washed with 500  $\mu\text{L}$  Buffer AW at 11,000 g for 1 min. Next, the column was rewashed with 600  $\mu\text{L}$  Buffer A4 at 11,000 g for 1 min. Finally, the column was dried by centrifugation at 11,000 g for 2 min and eluted with a 60  $\mu\text{L}$  elution buffer followed by centrifugation at 11,000 g for 1 min. The concentration of the plasmids determined by  $\text{OD}_{260} \times 5 \mu\text{g/ml}$ .

### **18. The steps required for Golden gate**

1  $\mu\text{L}$  T4 ligase buffer, 0.5  $\mu\text{L}$  T4 ligase, 0.5  $\mu\text{L}$  bsaI, 20 fmol of each PCR product and plasmid backbone were mixed and adjust to the final volume of 10  $\mu\text{L}$  with water. The PCR program was 35 cycles of digestion and ligation (42  $^{\circ}\text{C}$  for 2 min, 16  $^{\circ}\text{C}$  for 5 min) followed by a final digestion step (60  $^{\circ}\text{C}$  for 10 min), and a heat inactivation step (80  $^{\circ}\text{C}$  for 10 min).<sup>221</sup>

### **19. Expression of the two target genes in *S. cerevisiae***



*S. cerevisiae* BJ5464-NpgA was grown on YPD plate for three days. Single clones were picked and grown overnight in a liquid YPD medium. Cells were collected by centrifugation at 1500 g, 4 °C for 10 min, the pellet was washed once with water and twice with 100 mM Lithium Acetate. Cells were resuspended in 2.4 mL consisting of 50% PEG-3350, 360 µL of 1 M Lithium Acetate, 250 µL of salmon sperm DNA and 500 µL of water. 100 ng of plasmid was added to 200 µL of transformation mixture and incubated at 42 °C for 25 min. The mixture was centrifuged at 1500 g for 5 min. The precipitate was dissolved in water and spread on YSM plate (0.68% yeast nitrogen base without amino acids, 0.5% glucose, 0.14% yeast synthetic drop-out medium supplements without histidine, leucine, tryptophan and uracil, these four amino acids are selectively added to the medium based on auxotrophy). After about 3 to 4 days of culture at 30 °C, the *S. cerevisiae* mutants were grown on the plate. If *S. cerevisiae* harboring double plasmid needed to be constructed, using the same protocol and changing the amino acid composition of the selection medium according to the auxotrophy of the plasmid.

## **20. Extraction of *S. cerevisiae* genome**

The well-cultured yeast were harvested from the YPD plate and suspended in 50 µL of 0.02 M NaOH. The suspension was incubated at 95 °C for 15 min. After centrifugation at 5900 g for 5 min, the genome of *S. cerevisiae* was acquired.

## **21. SDS-PAGE**

BJ-EP67-RPL18B, BJ-EP67-PGK1 and BJ-EP67-TEF1 were grown in YSM medium lacking uracil and leucine at 25 °C for three days respectively. BJ5464-NpgA was cultured in the same conditions except for the YSM medium containing all the amino acids and used as the control. Then, 1% seed culture was inoculated into YPD media with 1% glucose and continued to grow at 25 °C for three days. The cells were harvested by centrifugation at 2500 g, 4 °C for 20 min. 30 mg of each sample was mixed with 1mL lysis buffer (Tris pH 7.5 20 mM, NaCl 150 mM). The cell lysates were obtained by sonication and then centrifuged to obtain supernatant.

Running buffer (Tris 7.57 g, Glycine 46.90 g, SDS 0.1%, 500 mL milli-Q H<sub>2</sub>O) and 2x loading buffer were prepared for SDS-PAGE gel analysis.

Table 5.5 Preparation of the samples for SDS-PAGE.

	Marker	supernatant	pellet
Ladder	7 $\mu$ l	0	0
Sample	0	70 $\mu$ l	70 $\mu$ l
2x protein loading buffer	7 $\mu$ l	70 $\mu$ l	70 $\mu$ l
DTT	1 $\mu$ l	1 $\mu$ l	1 $\mu$ l

The samples were well mixed, as shown in the table 5.5 and heated at 95°C for 5min. Then 15  $\mu$ l of each sample was loaded on 10% protein gel. The gel ran at a constant current of 25 mA for about 2 hours and then stained with Coomassie Brilliant Blue for at least 45 min, followed by unstained for 2 hours with 10% acetic acid.

## 22. Western blot

TBST and transfer buffer were prepared as following: transfer buffer consists of 3 g tris base, 11.25 g glycine, 200 mL ethanol and MilliQ H<sub>2</sub>O to 1 L. 0.1% of tween 20 was added on transfer buffer to prepare TBST. PVDF membrane was activated in 100% ethanol for 20 seconds. The gel was rinsed with the transfer buffer and the filter paper was wet with transfer buffer. The sandwich was well built and ran for 1.5 hours at 100 V. The membrane was stained with ponceau red (ponceau red 0.1% in 1% acetic acid) for 1 min and the ladder position was marked. Then the membrane was washed with water to eliminate the stain. The membrane was incubated with 5% milk in TBST) for 5 min, then combined with anti-His antibody (1:2000 in TBST) by incubating at 4 °C for overnight. The membrane was washed with a consecutive bath of TBST for 5 min and repeated four times. The membrane was covered with 1ml of AP substrate solution and waited for 5 min to make them fully integrated. The bolt of chemiluminescence was detected by the OMEGA.



## Conclusions and Perspectives

In the present study, we investigated the biosynthetic pathway of pyrrocidines and their analogues in the endophytic fungus *Sarocladium zae*. First of all, by analyzing the metabolic diversity of the wild type strain, we identified several new metabolites in the pyrrocidine family namely pyrrocidines D, trichobamide B, sarocladine A as well as the already known pyrrocidine F from another fungus.

Thanks to the sequenced genome of *S. zae* and established conditions enabling its genetic manipulation, we characterized the encoding PKS-NRPS BGC of pyrrocidines encompassing genes coding for six auxiliary enzymes, two transcription factors and one transporter. Gene knock-out of the tailoring enzymes generated a diversity of metabolites ranging from linear polyketides to complex polycyclic compounds. Some of them harbors new backbones having the rare *cis*-decalin motif or unique 6/10/6/5 fused-ring structure or 15-membered macrocycles fused to 5- and 6-numbered rings. However, none of the isolated metabolites correspond to biosynthetic intermediates which, due to their high reactivity, remained elusive. Nevertheless, comparisons of the chemical modifications between the isolated metabolites allowed us to give a function to each of the tailoring enzymes. After the synthesis of the linear nonaketide-tyrosine by PKS-NRPS the formation of pyrrolidone ring is controlled by PrcH. Then, the building of the fused-decahydrofluorene-paracyclophane structure of pyrrocidines A and B implies three proteins: the enzyme PrcB, originally predicted as Diels-Alderase, is an isomerizing cyclase installing ring A of decahydrofluorene; the integral transmembrane protein PrcE catalyzes the paracyclophane formation; the lipocaline PrcX controls the stereochemistry of a Diels-Alder cycloaddition which finalizes the formation rings B and C of the decahydrofluorene. Another integral transmembrane protein, PrcD which catalyzes hydroxylation, is also required in this enzymatic sequence very likely by making a transmembrane complex with PrcE. Finally, the last enzyme PrcI converts pyrrocidine A into B by reduction of carbon-carbon double bond.

Based on our functional analysis of the auxiliary enzymes, the study of the pathway is in progress in the team as part of the thesis project of Steffi Sewsum. Combining sequence analysis of the tailoring enzymes as well as protein modelisation and docking experiments with the results of the present work, we propose a detailed biosynthetic pathway for pyrrocidines via transient loss of aromaticity of tyrosine residue accounting for the formation of the distorted aromatic ring of paracyclophane moiety. This analysis also gives molecular bases to understand

the stereochemical diversity obtained in this family of natural products leading to a unified biosynthetic pathway for these compounds. A manuscript gathering these results is about to be submitted for publication (see annex).

Gene inactivation approach developed in *S. zeae*, allowed us to reveal the plasticity of pyrrocidine pathway which is promising to extend the chemical diversity by synthetic biology. This requires a heterologous expression system of the BGC that was investigated. The preliminary work of the cloning of PKS-NRPS and ER genes was achieved. However, the expression in *Saccharomyces cerevisiae* was not successful and this step is under optimization.

The exploration of the sequenced fungal genome with pyrrocidine BGC as target, allowed us to identify twelve similar clusters. The capacity to produce pyrrocidine related compounds of the eight commercial strains was investigated by OSMAC strategy. Only the fungus *Cordiceps javanica* seems to produce analogue traces which will need to be characterized after culture optimization. The silent clusters require their heterologous expression in a suitable host in order to characterize the linked metabolites.

The role of pyrrocidine in the antagonistic interaction between *S. zeae* and *Fusarium verticillioides* has been assessed. Our antagonistic experiments demonstrated that pyrrocidines reduce the progression of the pathogen and is in agreement with the protective role of *S. zeae* in maize.

The metabolites isolated from wild type and mutant *S. zeae* in this study are a potential source of bioactive compounds to be evaluated. Therefore, in one hand the power of growth inhibition of purified molecules will be tested on *F. verticillioides* and other pathogens. In the other hand the antagonistic effect of the *S. zeae* mutant producing these metabolites will be studied.

Two other encoding PKS-NRPS gene clusters identified in the *S. zeae* were also studied but no metabolites could be linked neither by gene inactivation study neither by overexpression of transcription factors of the pathways. These BGCs will be further investigated by heterologous expression.

# Reference

- (1) Petrini, O. Fungal Endophytes of Tree Leaves. In *Microbial Ecology of Leaves*; Andrews, J. H., Hirano, S. S., Eds.; Brock, T. D., Series Ed.; Brock/Springer Series in Contemporary Bioscience; Springer New York: New York, NY, 1991; pp 179–197. [https://doi.org/10.1007/978-1-4612-3168-4\\_9](https://doi.org/10.1007/978-1-4612-3168-4_9).
- (2) *Ecology, metabolite production, and substrate utilization in endophytic fungi* - PubMed. <https://pubmed.ncbi.nlm.nih.gov/1344919/> (accessed 2022-05-03).
- (3) Waqas, M.; Khan, A. L.; Kamran, M.; Hamayun, M.; Kang, S.-M.; Kim, Y.-H.; Lee, I.-J. Endophytic Fungi Produce Gibberellins and Indoleacetic Acid and Promotes Host-Plant Growth during Stress. *Molecules* **2012**, *17* (9), 10754–10773. <https://doi.org/10.3390/molecules170910754>.
- (4) Rowan, D. D. Lolitrems, Peramine and Paxilline: Mycotoxins of the Ryegrass/Endophyte Interaction. *Agriculture, Ecosystems & Environment* **1993**, *44* (1), 103–122. [https://doi.org/10.1016/0167-8809\(93\)90041-M](https://doi.org/10.1016/0167-8809(93)90041-M).
- (5) Mousa, W. K.; Raizada, M. N. The Diversity of Anti-Microbial Secondary Metabolites Produced by Fungal Endophytes: An Interdisciplinary Perspective. *Front. Microbiol.* **2013**, *4*. <https://doi.org/10.3389/fmicb.2013.00065>.
- (6) Tian, X. L.; Cao, L.; Tan, H. M.; Zeng, Q. G.; Jia, Y. Y.; Han, W.; Zhou, S. N. Study on the Communities of Endophytic Fungi and Endophytic Actinomycetes from Rice and Their Antipathogenic Activities in Vitro. *World Journal of Microbiology and Biotechnology* **2004**, *20*, 303–309. <https://doi.org/10.1023/B:WIBI.0000023843.83692.3f>.
- (7) Stierle, A.; Strobel, G.; Stierle, D.; Grothaus, P.; Bignami, G. The Search for a Taxol-Producing Microorganism Among the Endophytic Fungi of the Pacific Yew, *Taxus Brevifolia*. *J. Nat. Prod.* **1995**, *58* (9), 1315–1324. <https://doi.org/10.1021/np50123a002>.
- (8) Parthasarathy, R.; Sathiyabama, M. Lovastatin-Producing Endophytic Fungus Isolated from a Medicinal Plant *Solanum Xanthocarpum*. *Nat Prod Res* **2015**, *29* (24), 2282–2286. <https://doi.org/10.1080/14786419.2015.1016938>.
- (9) Wicklow, D. T.; Roth, S.; Deyrup, S. T.; Gloer, J. B. A Protective Endophyte of Maize: *Acremonium Zeae* Antibiotics Inhibitory to *Aspergillus Flavus* and *Fusarium Verticillioides*. *Mycological Research* **2005**, *109* (5), 610–618. <https://doi.org/10.1017/S0953756205002820>.
- (10) *ECOLOGY OF CORN STALK ROT IN NEBRASKA* - ProQuest. <https://www.proquest.com/openview/b23b7b2cfe1e290e3c575e34c2d3ddfb/1?pq-origsite=gscholar&cbl=18750&diss=y> (accessed 2022-02-21).
- (11) *Fusarium verticillioides: Advancements in Understanding the Toxicity, Virulence, and Niche Adaptations of a Model Mycotoxigenic Pathogen of Maize | Phytopathology®*. <https://apsjournals.apsnet.org/doi/10.1094/PHYTO-06-17-0203-RVW> (accessed 2022-02-21).
- (12) Munkvold, G. P.; White, D. G. *Compendium of Corn Diseases, Fourth Edition*; Diseases and Pests Compendium Series; The American Phytopathological Society, 2016. <https://doi.org/10.1094/9780890544945>.
- (13) Wicklow, D. T.; Poling, S. M. Antimicrobial Activity of Pyrrocidines from *Acremonium Zeae* Against Endophytes and Pathogens of Maize. *Phytopathology* **2009**, *99* (1), 109–115. <https://doi.org/10.1094/PHYTO-99-1-0109>.

- (14) Yuan, H.; Ma, Q.; Ye, L.; Piao, G. The Traditional Medicine and Modern Medicine from Natural Products. *Molecules* **2016**, *21* (5), 559. <https://doi.org/10.3390/molecules21050559>.
- (15) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83* (3), 770–803. <https://doi.org/10.1021/acs.jnatprod.9b01285>.
- (16) Pham, J. V.; Yilma, M. A.; Feliz, A.; Majid, M. T.; Maffetone, N.; Walker, J. R.; Kim, E.; Cho, H. J.; Reynolds, J. M.; Song, M. C.; Park, S. R.; Yoon, Y. J. A Review of the Microbial Production of Bioactive Natural Products and Biologics. *Frontiers in Microbiology* **2019**, *10*.
- (17) Bérdy, J. Bioactive Microbial Metabolites. *J Antibiot (Tokyo)* **2005**, *58* (1), 1–26. <https://doi.org/10.1038/ja.2005.1>.
- (18) Porquier, A.; Morgant, G.; Moraga, J.; Dalmais, B.; Luyten, I.; Simon, A.; Pradier, J.-M.; Amselem, J.; Collado, I. G.; Viaud, M. The Botrydial Biosynthetic Gene Cluster of *Botrytis Cinerea* Displays a Bipartite Genomic Structure and Is Positively Regulated by the Putative Zn(II)2Cys6 Transcription Factor BcBot6. *Fungal Genet Biol* **2016**, *96*, 33–46. <https://doi.org/10.1016/j.fgb.2016.10.003>.
- (19) Andersen, M. R.; Nielsen, J. B.; Klitgaard, A.; Petersen, L. M.; Zachariassen, M.; Hansen, T. J.; Blicher, L. H.; Gotfredsen, C. H.; Larsen, T. O.; Nielsen, K. F.; Mortensen, U. H. Accurate Prediction of Secondary Metabolite Gene Clusters in Filamentous Fungi. *Proc Natl Acad Sci U S A* **2013**, *110* (1), E99–107. <https://doi.org/10.1073/pnas.1205532110>.
- (20) Wiemann, P.; Guo, C.-J.; Palmer, J. M.; Sekonyela, R.; Wang, C. C. C.; Keller, N. P. Prototype of an Intertwined Secondary-Metabolite Supercluster. *Proceedings of the National Academy of Sciences* **2013**, *110* (42), 17065–17070. <https://doi.org/10.1073/pnas.1313258110>.
- (21) Lyu, H.-N.; Liu, H.-W.; Keller, N. P.; Yin, W.-B. Harnessing Diverse Transcriptional Regulators for Natural Product Discovery in Fungi. *Nat. Prod. Rep.* **2019**, *10.1039/C8NP00027A*. <https://doi.org/10.1039/C8NP00027A>.
- (22) Grigoriev, I. V.; Nikitin, R.; Haridas, S.; Kuo, A.; Ohm, R.; Otilar, R.; Riley, R.; Salamov, A.; Zhao, X.; Korzeniewski, F.; Smirnova, T.; Nordberg, H.; Dubchak, I.; Shabalov, I. MycoCosm Portal: Gearing up for 1000 Fungal Genomes. *Nucleic Acids Res* **2014**, *42* (Database issue), D699–704. <https://doi.org/10.1093/nar/gkt1183>.
- (23) *Biochemistry of Terpenoids: Monoterpenes, Sesquiterpenes and Diterpenes - Ashour - - Major Reference Works - Wiley Online Library*. <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119312994.apr0427> (accessed 2020-03-20).
- (24) Abdallah, I.; Quax, W. A Glimpse into the Biosynthesis of Terpenoids. *KnE Life Sciences* **2017**, *2017*, 81–98. <https://doi.org/10.18502/kls.v3i5.981>.
- (25) Mb, Q.; Cm, F.; C, S.-D. Traversing the Fungal Terpenome. *Natural product reports* **2014**, *31* (10). <https://doi.org/10.1039/c4np00075g>.
- (26) Lee, M. K.; Hung, T. M.; Cuong, T. D.; Na, M.; Kim, J. C.; Kim, E.-J.; Park, H.-S.; Choi, J. S.; Lee, I.; Bae, K.; Hattori, M.; Min, B. S. Ergosta-7,22-Diene-2 $\beta$ ,3 $\alpha$ ,9 $\alpha$ -Triol from the Fruit Bodies of *Ganoderma Lucidum* Induces Apoptosis in Human Myelocytic HL-60 Cells. *Phytother Res* **2011**, *25* (11), 1579–1585. <https://doi.org/10.1002/ptr.3447>.
- (27) Duru, M. E.; Çayan, G. T. Biologically Active Terpenoids from Mushroom Origin: A Review. *Records of Natural Products* **2015**, *9* (4), 456–483.

- (28) Han, J.; Chen, Y.; Bao, L.; Yang, X.; Liu, D.; Li, S.; Zhao, F.; Liu, H. Anti-Inflammatory and Cytotoxic Cyathane Diterpenoids from the Medicinal Fungus *Cyathus Africanus*. *Fitoterapia* **2013**, *84*, 22–31. <https://doi.org/10.1016/j.fitote.2012.10.001>.
- (29) Minami, A.; Tajima, N.; Higuchi, Y.; Toyomasu, T.; Sassa, T.; Kato, N.; Dairi, T. Identification and Functional Analysis of Brassicicene C Biosynthetic Gene Cluster in *Alternaria Brassicicola*. *Bioorg Med Chem Lett* **2009**, *19* (3), 870–874. <https://doi.org/10.1016/j.bmcl.2008.11.108>.
- (30) Wang, B. T.; Qi, Q. Y.; Ma, K.; Pei, Y. F.; Han, J. J.; Xu, W.; Li, E. W.; Liu, H. W. Depside  $\alpha$ -Glucosidase Inhibitors from a Culture of the Mushroom *Stereum Hirsutum*. *Planta Med.* **2014**, *80* (11), 918–924. <https://doi.org/10.1055/s-0034-1382828>.
- (31) Dostál, J. *Two Faces of Alkaloids*. ACS Publications. <https://pubs.acs.org/doi/pdf/10.1021/ed077p993> (accessed 2022-01-30). <https://doi.org/10.1021/ed077p993>.
- (32) Shao, C.-L.; Wang, C.-Y.; Gu, Y.-C.; Wei, M.-Y.; Pan, J.-H.; Deng, D.-S.; She, Z.-G.; Lin, Y.-C. Penicinoline, a New Pyrrolyl 4-Quinolinone Alkaloid with an Unprecedented Ring System from an Endophytic Fungus *Penicillium* Sp. *Bioorg Med Chem Lett* **2010**, *20* (11), 3284–3286. <https://doi.org/10.1016/j.bmcl.2010.04.043>.
- (33) Liu, X.; Dong, M.; Chen, X.; Jiang, M.; Lv, X.; Zhou, J. Antimicrobial Activity of an Endophytic *Xylaria* Sp. YX-28 and Identification of Its Antimicrobial Compound 7-Amino-4-Methylcoumarin. *Appl Microbiol Biotechnol* **2008**, *78* (2), 241–247. <https://doi.org/10.1007/s00253-007-1305-1>.
- (34) Zhang, D.; Satake, M.; Fukuzawa, S.; Sugahara, K.; Niitsu, A.; Shirai, T.; Tachibana, K. Two New Indole Alkaloids, 2-(3,3-Dimethylprop-1-Ene)-Costaclavine and 2-(3,3-Dimethylprop-1-Ene)-Epicostaclavine, from the Marine-Derived Fungus *Aspergillus Fumigatus*. *Journal of natural medicines* **2011**, *66*, 222–226. <https://doi.org/10.1007/s11418-011-0565-3>.
- (35) Devi, N. Bioactive Metabolites from an Endophytic Fungus *Penicillium* Sp. Isolated from *Centella Asiatica*. *CREAM* **2014**, *4* (1), 34–43. <https://doi.org/10.5943/cream/4/1/3>.
- (36) Jiao, R. H.; Xu, S.; Liu, J. Y.; Ge, H. M.; Ding, H.; Xu, C.; Zhu, H. L.; Tan, R. X. Chaetominine, a Cytotoxic Alkaloid Produced by Endophytic *Chaetomium* Sp. IFB-E015. *Org. Lett.* **2006**, *8* (25), 5709–5712. <https://doi.org/10.1021/ol062257t>.
- (37) Davis, R. A.; Longden, J.; Avery, V. M.; Healy, P. C. The Isolation, Structure Determination and Cytotoxicity of the New Fungal Metabolite, Trichodermamide C. *Bioorg Med Chem Lett* **2008**, *18* (9), 2836–2839. <https://doi.org/10.1016/j.bmcl.2008.03.090>.
- (38) Hong, Z. Peptide Natural Products from *Pseudomonas*: Discovery, Biosynthesis and Ecological Roles. **2018**.
- (39) Döhren, H. von. Chapter 15. Compilation of Peptide Structures - A Biogenetic Approach. In *Chapter 15. Compilation of Peptide Structures - A Biogenetic Approach*; De Gruyter, 2019; pp 411–508. <https://doi.org/10.1515/9783110886139-016>.
- (40) Du, L.; Lou, L. PKS and NRPS Release Mechanisms. *Nat. Prod. Rep.* **2010**, *27* (2), 255–278. <https://doi.org/10.1039/B912037H>.
- (41) Kaniusaite, M.; Tailhades, J.; A. Marschall, E.; A. Goode, R. J.; B. Schittenhelm, R.; J. Cryle, M. A Proof-Reading Mechanism for Non-Proteinogenic Amino Acid Incorporation into Glycopeptide Antibiotics. *Chemical Science* **2019**, *10* (41), 9466–9482. <https://doi.org/10.1039/C9SC03678D>.



- (42) Oide, S.; Turgeon, B. G. Natural Roles of Nonribosomal Peptide Metabolites in Fungi. *Mycoscience* **2020**, *61* (3), 101–110. <https://doi.org/10.1016/j.myc.2020.03.001>.
- (43) Gao, X.; Haynes, S. W.; Ames, B. D.; Wang, P.; Vien, L. P.; Walsh, C. T.; Tang, Y. Cyclization of Fungal Nonribosomal Peptides by a Terminal Condensation-like Domain. *Nat Chem Biol* **2012**, *8* (10), 823–830. <https://doi.org/10.1038/nchembio.1047>.
- (44) Yu, D.; Xu, F.; Zhang, S.; Zhan, J. Decoding and Reprogramming Fungal Iterative Nonribosomal Peptide Synthetases. *Nat Commun* **2017**, *8* (1), 15349. <https://doi.org/10.1038/ncomms15349>.
- (45) Fischbach, M. A.; Walsh, C. T. Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *Chem. Rev.* **2006**, *106* (8), 3468–3496. <https://doi.org/10.1021/cr0503097>.
- (46) Hai, Y.; Jenner, M.; Tang, Y. Complete Stereoinversion of L-Tryptophan by a Fungal Single-Module Nonribosomal Peptide Synthetase. *J. Am. Chem. Soc.* **2019**, *141* (41), 16222–16226. <https://doi.org/10.1021/jacs.9b08898>.
- (47) Kalb, D.; Lackner, G.; Hoffmeister, D. Fungal Peptide Synthetases: An Update on Functions and Specificity Signatures. *Fungal Biology Reviews* **2013**, *27* (2), 43–50. <https://doi.org/10.1016/j.fbr.2013.05.002>.
- (48) Keller, N. P.; Turner, G.; Bennett, J. W. Fungal Secondary Metabolism — from Biochemistry to Genomics. *Nat Rev Microbiol* **2005**, *3* (12), 937–947. <https://doi.org/10.1038/nrmicro1286>.
- (49) Li, H.; Gilchrist, C. L. M.; Phan, C.-S.; Lacey, H. J.; Vuong, D.; Moggach, S. A.; Lacey, E.; Piggott, A. M.; Chooi, Y.-H. Biosynthesis of a New Benzazepine Alkaloid Nanangelenin A from *Aspergillus Nanangensis* Involves an Unusual L-Kynurenine-Incorporating NRPS Catalyzing Regioselective Lactamization. *J. Am. Chem. Soc.* **2020**, *142* (15), 7145–7152. <https://doi.org/10.1021/jacs.0c01605>.
- (50) Liu, H.; Fan, J.; Zhang, P.; Hu, Y.; Liu, X.; Li, S.-M.; Yin, W.-B. New Insights into the Disulfide Bond Formation Enzymes in Epidithiodiketopiperazine Alkaloids. *Chem. Sci.* **2021**, *12* (11), 4132–4138. <https://doi.org/10.1039/D0SC06647H>.
- (51) García-Estrada, C.; Ullán, R. V.; Albillos, S. M.; Fernández-Bodega, M. Á.; Durek, P.; von Döhren, H.; Martín, J. F. A Single Cluster of Coregulated Genes Encodes the Biosynthesis of the Mycotoxins Roquefortine C and Meleagrins in *Penicillium Chrysogenum*. *Chem Biol* **2011**, *18* (11), 1499–1512. <https://doi.org/10.1016/j.chembiol.2011.08.012>.
- (52) Huang, R.-M.; Yi, X.-X.; Zhou, Y.; Su, X.; Peng, Y.; Gao, C.-H. An Update on 2,5-Diketopiperazines from Marine Organisms. *Marine Drugs* **2014**, *12* (12), 6213–6235. <https://doi.org/10.3390/md12126213>.
- (53) Le Govic, Y.; Papon, N.; Le Gal, S.; Bouchara, J.-P.; Vandeputte, P. Non-Ribosomal Peptide Synthetase Gene Clusters in the Human Pathogenic Fungus *Scedosporium Apiospermum*. *Frontiers in Microbiology* **2019**, *10*.
- (54) Moutiez, M.; Schmitt, E.; Seguin, J.; Thai, R.; Favry, E.; Belin, P.; Mechulam, Y.; Gondry, M. Unravelling the Mechanism of Non-Ribosomal Peptide Synthesis by Cyclodipeptide Synthases. *Nat Commun* **2014**, *5* (1), 5141. <https://doi.org/10.1038/ncomms6141>.
- (55) *Aminoacyl-tRNA-Utilizing Enzymes in Natural Product Biosynthesis | Chemical Reviews*. <https://pubs.acs.org/doi/abs/10.1021/acs.chemrev.6b00523> (accessed 2022-04-29).

- (56) G. Arnison, P.; J. Bibb, M.; Bierbaum, G.; A. Bowers, A.; S. Bugni, T.; Bulaj, G.; A. Camarero, J.; J. Campopiano, D.; L. Challis, G.; Clardy, J.; D. Cotter, P.; J. Craik, D.; Dawson, M.; Dittmann, E.; Donadio, S.; C. Dorrestein, P.; Entian, K.-D.; A. Fischbach, M.; S. Garavelli, J.; Göransson, U.; W. Gruber, C.; H. Haft, D.; K. Hemscheidt, T.; Hertweck, C.; Hill, C.; R. Horswill, A.; Jaspars, M.; L. Kelly, W.; P. Klinman, J.; P. Kuipers, O.; James Link, A.; Liu, W.; A. Marahiel, M.; A. Mitchell, D.; N. Moll, G.; S. Moore, B.; Müller, R.; K. Nair, S.; F. Nes, I.; E. Norris, G.; M. Olivera, B.; Onaka, H.; L. Patchett, M.; Piel, J.; T. Reaney, M. J.; Rebuffat, S.; Paul Ross, R.; Sahl, H.-G.; W. Schmidt, E.; E. Selsted, M.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; D. Süßmuth, R.; R. Tagg, J.; Tang, G.-L.; W. Truman, A.; C. Vederas, J.; T. Walsh, C.; D. Walton, J.; C. Wenzel, S.; M. Willey, J.; Donk, W. A. van der. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Natural Product Reports* **2013**, 30 (1), 108–160. <https://doi.org/10.1039/C2NP20085F>.
- (57) *New developments in RiPP discovery, enzymology and engineering - Natural Product Reports (RSC Publishing)* DOI:10.1039/D0NP00027B. <https://pubs.rsc.org/en/content/articlehtml/2021/np/d0np00027b> (accessed 2022-02-10).
- (58) Baumann, K.; Zanotti, G.; Faulstich, H. A  $\beta$ -Turn in  $\alpha$ -Amanitin Is the Most Important Structural Feature for Binding to RNA Polymerase II and Three Monoclonal Antibodies. *Protein Science* **1994**, 3 (5), 750–756. <https://doi.org/10.1002/pro.5560030504>.
- (59) Vries, J. D.; Wieland, T. *Influence of phallotoxins and metal ions on the rate of proteolysis of actin*. ACS Publications. <https://pubs.acs.org/doi/pdf/10.1021/bi00603a025> (accessed 2022-04-16). <https://doi.org/10.1021/bi00603a025>.
- (60) Wieland, Th.; Lüben, G.; Ottenheym, H.; Faesel, J.; De Vries, J. X.; Prox, A.; Schmid, J. The Discovery, Isolation, Elucidation of Structure, and Synthesis of Antamanide. *Angewandte Chemie International Edition in English* **1968**, 7 (3), 204–208. <https://doi.org/10.1002/anie.196802041>.
- (61) Büchel, E.; Martini, U.; Mayer, A.; Anke, H.; Sterner, O. Omphalotins B, C and D, Nematicidal Cyclopeptides from *Omphalotus Olearius*. Absolute Configuration of Omphalotin A. *Tetrahedron* **1998**, 54 (20), 5345–5352. [https://doi.org/10.1016/S0040-4020\(98\)00209-9](https://doi.org/10.1016/S0040-4020(98)00209-9).
- (62) *USTILOXINS, ANTIMITOTIC CYCLIC PEPTIDES FROM FALSE SMUT BALLS ON RICE PANICLES CAUSED BY Ustilaginoidea virens*. [https://www.jstage.jst.go.jp/article/antibiotics1968/47/7/47\\_7\\_765/\\_article/-char/ja/](https://www.jstage.jst.go.jp/article/antibiotics1968/47/7/47_7_765/_article/-char/ja/) (accessed 2022-04-16).
- (63) Koiso, Y.; Morisaki, N.; Yamashita, Y.; Mitsui, Y.; Shirai, R.; Hashimoto, Y.; Iwasaki, S. Isolation and Structure of an Antimitotic Cyclic Peptide, Ustiloxin F: Chemical Interrelation with a Homologous Peptide, Ustiloxin B. *J. Antibiot.* **1998**, 51 (4), 418–422. <https://doi.org/10.7164/antibiotics.51.418>.
- (64) Ye, Y.; Ozaki, T.; Umemura, M.; Liu, C.; Minami, A.; Oikawa, H. Heterologous Production of Asperipin-2a: Proposal for Sequential Oxidative Macrocyclization by a Fungi-Specific DUF3328 Oxidase. *Org. Biomol. Chem.* **2018**, 17 (1), 39–43. <https://doi.org/10.1039/C8OB02824A>.
- (65) *A Self-Sacrificing N-Methyltransferase Is the Precursor of the Fungal Natural Product Omphalotin - Ramm - 2017 - Angewandte Chemie International Edition - Wiley Online Library*. <https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201703488> (accessed 2022-02-10).

- (66) Shen, B. Polyketide Biosynthesis beyond the Type I, II and III Polyketide Synthase Paradigms. *Current Opinion in Chemical Biology* **2003**, 7 (2), 285–295. [https://doi.org/10.1016/S1367-5931\(03\)00020-6](https://doi.org/10.1016/S1367-5931(03)00020-6).
- (67) Boettger, D.; Hertweck, C. Molecular Diversity Sculpted by Fungal PKS/NRPS Hybrids. **2013**, 15.
- (68) Nguyen, T.; Ishida, K.; Jenke-Kodama, H.; Dittmann, E.; Gurgui, C.; Hochmuth, T.; Taudien, S.; Platzer, M.; Hertweck, C.; Piel, J. Exploiting the Mosaic Structure of Trans-Acyltransferase Polyketide Synthases for Natural Product Discovery and Pathway Dissection. *Nat. Biotechnol.* **2008**, 26 (2), 225–233. <https://doi.org/10.1038/nbt1379>.
- (69) Cox, R. J.; Simpson, T. J. Fungal Type I Polyketide Synthases. *Methods Enzymol* **2009**, 459, 49–78. [https://doi.org/10.1016/S0076-6879\(09\)04603-5](https://doi.org/10.1016/S0076-6879(09)04603-5).
- (70) Herbst, D. A.; Townsend, C. A.; Maier, T. The Architectures of Iterative Type I PKS and FAS. *Nat. Prod. Rep.* **2018**, 35 (10), 1046–1069. <https://doi.org/10.1039/C8NP00039E>.
- (71) Cane, D. E.; Walsh, C. T. The Parallel and Convergent Universes of Polyketide Synthases and Nonribosomal Peptide Synthetases. *Chemistry & Biology* **1999**, 6 (12), R319–R325. [https://doi.org/10.1016/S1074-5521\(00\)80001-0](https://doi.org/10.1016/S1074-5521(00)80001-0).
- (72) Fisch, K. M.; Bakeer, W.; Yakasai, A. A.; Song, Z.; Pedrick, J.; Wasil, Z.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. Rational Domain Swaps Decipher Programming in Fungal Highly Reducing Polyketide Synthases and Resurrect an Extinct Metabolite. *J. Am. Chem. Soc.* **2011**, 133 (41), 16635–16641. <https://doi.org/10.1021/ja206914q>.
- (73) *Production of Fumonisin Analogs by Fusarium Species | Applied and Environmental Microbiology.* <https://journals.asm.org/doi/10.1128/AEM.68.5.2101-2105.2002> (accessed 2022-04-05).
- (74) *Lovastatin Nonaketide Synthase Catalyzes an Intramolecular Diels–Alder Reaction of a Substrate Analogue | Journal of the American Chemical Society.* <https://pubs.acs.org/doi/10.1021/ja003216%2B> (accessed 2021-04-13).
- (75) Shao, L.; Qu, X.-D.; Jia, X.-Y.; Zhao, Q.-F.; Tian, Z.-H.; Wang, M.; Tang, G.-L.; Liu, W. Cloning and Characterization of a Bacterial Iterative Type I Polyketide Synthase Gene Encoding the 6-Methylsalicylic Acid Synthase. *Biochem Biophys Res Commun* **2006**, 345 (1), 133–139. <https://doi.org/10.1016/j.bbrc.2006.04.069>.
- (76) Yu, J. Current Understanding on Aflatoxin Biosynthesis and Future Perspective in Reducing Aflatoxin Contamination. *Toxins (Basel)* **2012**, 4 (11), 1024–1057. <https://doi.org/10.3390/toxins4111024>.
- (77) Chooi, Y.-H.; Cacho, R.; Tang, Y. Identification of the Viridicatumtoxin and Griseofulvin Gene Clusters from *Penicillium Aethiopicum*. *Chem Biol* **2010**, 17 (5), 483–494. <https://doi.org/10.1016/j.chembiol.2010.03.015>.
- (78) Zhang, A.; Lu, P.; Dahl-Roshak, A. M.; Paress, P. S.; Kennedy, S.; Tkacz, J. S.; An, Z. Efficient Disruption of a Polyketide Synthase Gene ( Pks1) Required for Melanin Synthesis through Agrobacterium-Mediated Transformation of *Glarea Lozoyensis*. *Mol Genet Genomics* **2003**, 268 (5), 645–655. <https://doi.org/10.1007/s00438-002-0780-4>.

- (79) Stewart, C.; Vickery, C. R.; Burkart, M. D.; Noel, J. P. Confluence of Structural and Chemical Biology: Plant Polyketide Synthases as Biocatalysts for a Bio-Based Future. *Current Opinion in Plant Biology* **2013**, *16* (3), 365–372. <https://doi.org/10.1016/j.pbi.2013.02.004>.
- (80) Lim, Y. P.; Go, M. K.; Yew, W. S. Exploiting the Biosynthetic Potential of Type III Polyketide Synthases. *Molecules* **2016**, *21* (6), 806. <https://doi.org/10.3390/molecules21060806>.
- (81) Xu, Y.; Zhou, T.; Zhang, S.; Espinosa-Artiles, P.; Wang, L.; Zhang, W.; Lin, M.; Gunatilaka, A. A. L.; Zhan, J.; Molnár, I. Diversity-Oriented Combinatorial Biosynthesis of Benzenediol Lactone Scaffolds by Subunit Shuffling of Fungal Polyketide Synthases. *Proceedings of the National Academy of Sciences* **2014**, *111* (34), 12354–12359. <https://doi.org/10.1073/pnas.1406999111>.
- (82) Matsuda, Y.; Awakawa, T.; Mori, T.; Abe, I. Unusual Chemistries in Fungal Meroterpenoid Biosynthesis. *Curr Opin Chem Biol* **2016**, *31*, 1–7. <https://doi.org/10.1016/j.cbpa.2015.11.001>.
- (83) Yun, C.-S.; Motoyama, T.; Osada, H. Biosynthesis of the Mycotoxin Tenuazonic Acid by a Fungal NRPS–PKS Hybrid Enzyme. *Nat Commun* **2015**, *6* (1), 8758. <https://doi.org/10.1038/ncomms9758>.
- (84) *Fusarin C Biosynthesis in Fusarium moniliforme and Fusarium venenatum* - Song - 2004 - *ChemBioChem* - Wiley Online Library. <https://onlinelibrary.wiley.com/doi/abs/10.1002/cbic.200400138> (accessed 2020-03-21).
- (85) Sato, M.; Kishimoto, S.; Yokoyama, M.; Jamieson, C. S.; Narita, K.; Maeda, N.; Hara, K.; Hashimoto, H.; Tsunematsu, Y.; Houk, K. N.; Tang, Y.; Watanabe, K. Catalytic Mechanism and Endo-to-Exo Selectivity Reversion of an Octalin-Forming Natural Diels–Alderase. *Nature Catalysis* **2021**, *4* (3), 223–232. <https://doi.org/10.1038/s41929-021-00577-2>.
- (86) Donzelli, B. G. G.; Krasnoff, S. B.; Churchill, A. C. L.; Vandenberg, J. D.; Gibson, D. M. Identification of a Hybrid PKS–NRPS Required for the Biosynthesis of NG-391 in *Metarhizium robertsii*. *Curr Genet* **2010**, *56* (2), 151–162. <https://doi.org/10.1007/s00294-010-0288-0>.
- (87) Kakule, T. B.; Sardar, D.; Lin, Z.; Schmidt, E. W. Two Related Pyrrolidinedione Synthetase Loci in *Fusarium heterosporum* ATCC 74349 Produce Divergent Metabolites. *ACS Chem. Biol.* **2013**, *8* (7), 1549–1557. <https://doi.org/10.1021/cb400159f>.
- (88) Eley, K. L.; Halo, L. M.; Song, Z.; Powles, H.; Cox, R. J.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J. Biosynthesis of the 2-Pyridone Tenellin in the Insect Pathogenic Fungus *Beauveria bassiana*. *ChemBioChem* **2007**, *8* (3), 289–297. <https://doi.org/10.1002/cbic.200600398>.
- (89) Heneghan, M. N.; Yakasai, A. A.; Williams, K.; Kadir, K. A.; Wasil, Z.; Bakeer, W.; Fisch, K. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J.; Lazarus, C. M. The Programming Role of Trans-Acting Enoyl Reductases during the Biosynthesis of Highly Reduced Fungal Polyketides. *Chem. Sci.* **2011**, *2* (5), 972–979. <https://doi.org/10.1039/C1SC00023C>.
- (90) Xu, W.; Cai, X.; Jung, M. E.; Tang, Y. Analysis of Intact and Dissected Fungal Polyketide Synthase–Nonribosomal Peptide Synthetase in Vitro and in *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* **2010**, *132* (39), 13604–13607. <https://doi.org/10.1021/ja107084d>.
- (91) Schümann, J.; Hertweck, C. Molecular Basis of Cytochalasin Biosynthesis in Fungi: Gene Cluster Analysis and Evidence for the Involvement of a PKS–NRPS Hybrid Synthase by RNA Silencing. *J. Am. Chem. Soc.* **2007**, *129* (31), 9564–9565. <https://doi.org/10.1021/ja072884t>.
- (92) Fujii, R.; Minami, A.; Gomi, K.; Oikawa, H. Biosynthetic Assembly of Cytochalasin Backbone. *Tetrahedron Letters* **2013**, *54* (23), 2999–3002. <https://doi.org/10.1016/j.tetlet.2013.03.120>.

- (93) Maiya, S.; Grundmann, A.; Li, X.; Li, S.-M.; Turner, G. Identification of a Hybrid PKS/NRPS Required for Pseurotin A Biosynthesis in the Human Pathogen *Aspergillus Fumigatus*. *Chembiochem* **2007**, *8* (14), 1736–1743. <https://doi.org/10.1002/cbic.200700202>.
- (94) Liu, X.; Walsh, C. T. Characterization of Cyclo-Acetoacetyl-L-Tryptophan Dimethylallyltransferase in Cyclopiiazonic Acid Biosynthesis: Substrate Promiscuity and Site Directed Mutagenesis Studies. *Biochemistry* **2009**, *48* (46), 11032–11044. <https://doi.org/10.1021/bi901597j>.
- (95) *Multifactorial induction of an orphan PKS-NRPS gene cluster in Aspergillus terreus. - Leibniz Institute for Natural Product Research and Infection Biology - Hans Knöll Institute (HKI)*. <https://www.leibniz-hki.de/en/publication.html?publication=104> (accessed 2022-02-06).
- (96) Yamamoto, T.; Tsunematsu, Y.; Noguchi, H.; Hotta, K.; Watanabe, K. Elucidation of Pyranonigrin Biosynthetic Pathway Reveals a Mode of Tetramic Acid, Fused  $\gamma$ -Pyrone, and Exo-Methylene Formation. *Org. Lett.* **2015**, *17* (20), 4992–4995. <https://doi.org/10.1021/acs.orglett.5b02435>.
- (97) Kakule, T. B.; Jadulco, R. C.; Koch, M.; Janso, J. E.; Barrows, L. R.; Schmidt, E. W. Native Promoter Strategy for High-Yielding Synthesis and Engineering of Fungal Secondary Metabolites. *ACS Synth Biol* **2015**, *4* (5), 625–633. <https://doi.org/10.1021/sb500296p>.
- (98) Cary, J. W.; Uka, V.; Han, Z.; Buyst, D.; Harris-Coward, P. Y.; Ehrlich, K. C.; Wei, Q.; Bhatnagar, D.; Dowd, P. F.; Martens, S. L.; Calvo, A. M.; Martins, J. C.; Vanhaecke, L.; Coenye, T.; De Saeger, S.; Di Mavungu, J. D. An *Aspergillus Flavus* Secondary Metabolic Gene Cluster Containing a Hybrid PKS-NRPS Is Necessary for Synthesis of the 2-Pyridones, Leporins. *Fungal Genet. Biol.* **2015**, *81*, 88–97. <https://doi.org/10.1016/j.fgb.2015.05.010>.
- (99) Sato, D. M.; Yagishita, D. F.; Mino, P. D. T.; Uchiyama, D. N.; Patel, D. A.; Chooi, D. Y.-H.; Goda, D. Y.; Xu, D. W.; Noguchi, P. D. H.; Yamamoto, T.; Hotta, P. D. K.; Houk, P. D. K. N.; Tang, P. D. Y.; Watanabe, P. D. K. Involvement of Lipocalin-like CghA in Decalin-Forming Stereoselective Intramolecular [4+2] Cycloaddition. *Chembiochem : a European journal of chemical biology* **2015**, *16* (16), 2294. <https://doi.org/10.1002/cbic.201500386>.
- (100) Li, L.; Yu, P.; Tang, M.-C.; Zou, Y.; Gao, S.-S.; Hung, Y.-S.; Zhao, M.; Watanabe, K.; Houk, K. N.; Tang, Y. Biochemical Characterization of a Eukaryotic Decalin-Forming Diels–Alderase. *J. Am. Chem. Soc.* **2016**, *138* (49), 15837–15840. <https://doi.org/10.1021/jacs.6b10452>.
- (101) Yokoyama, M.; Hirayama, Y.; Yamamoto, T.; Kishimoto, S.; Tsunematsu, Y.; Watanabe, K. Integration of Chemical, Genetic, and Bioinformatic Approaches Delineates Fungal Polyketide–Peptide Hybrid Biosynthesis. *Org. Lett.* **2017**, *19* (8), 2002–2005. <https://doi.org/10.1021/acs.orglett.7b00559>.
- (102) Sato, M.; Dander, J. E.; Sato, C.; Hung, Y.-S.; Gao, S.-S.; Tang, M.-C.; Hang, L.; Winter, J. M.; Garg, N. K.; Watanabe, K.; Tang, Y. Collaborative Biosynthesis of Maleimide- and Succinimide-Containing Natural Products by Fungal Polyketide Megasyntases. *J. Am. Chem. Soc.* **2017**, *139* (15), 5317–5320. <https://doi.org/10.1021/jacs.7b02432>.
- (103) Li, L.; Tang, M.-C.; Tang, S.; Gao, S.; Soliman, S.; Hang, L.; Xu, W.; Ye, T.; Watanabe, K.; Tang, Y. Genome Mining and Assembly-Line Biosynthesis of the UCS1025A Pyrrolizidinone Family of Fungal Alkaloids. *J. Am. Chem. Soc.* **2018**, *140* (6), 2067–2071. <https://doi.org/10.1021/jacs.8b00056>.
- (104) Wright, A. D.; Osterhage, C.; König, G. M. Epicoccamide, a Novel Secondary Metabolite from a Jellyfish-Derived Culture of *Epicoccum Purpurascens*. *Org. Biomol. Chem.* **2003**, *1* (3), 507–510. <https://doi.org/10.1039/B208588G>.

- (105) Zhang, Z.; Qiao, T.; Watanabe, K.; Tang, Y. Concise Biosynthesis of Phenylfuropyridones in Fungi. *Angew. Chem. Int. Ed.* **2020**, *59* (45), 19889–19893. <https://doi.org/10.1002/anie.202008321>.
- (106) Xie, L.; Zang, X.; Cheng, W.; Zhang, Z.; Zhou, J.; Chen, M.; Tang, Y. Harzianic Acid from *Trichoderma Afroharzianum* Is a Natural Product Inhibitor of Acetohydroxyacid Synthase. *J. Am. Chem. Soc.* **2021**, *143* (25), 9575–9584. <https://doi.org/10.1021/jacs.1c03988>.
- (107) Tang, S.; Zhang, W.; Li, Z.; Li, H.; Geng, C.; Huang, X.; Lu, X. Discovery and Characterization of a PKS–NRPS Hybrid in *Aspergillus Terreus* by Genome Mining. *J. Nat. Prod.* **2020**, *83* (2), 473–480. <https://doi.org/10.1021/acs.jnatprod.9b01140>.
- (108) Zhang, J.-M.; Wang, H.-H.; Liu, X.; Hu, C.-H.; Zou, Y. Heterologous and Engineered Biosynthesis of Nematocidal Polyketide–Nonribosomal Peptide Hybrid Macrolactone from Extreme Thermophilic Fungi. *J. Am. Chem. Soc.* **2020**, *142* (4), 1957–1965. <https://doi.org/10.1021/jacs.9b11410>.
- (109) Bat-Erdene, U.; Kanayama, D.; Tan, D.; Turner, W. C.; Houk, K. N.; Ohashi, M.; Tang, Y. Iterative Catalysis in the Biosynthesis of Mitochondrial Complex II Inhibitors Harzianopyridone and Atpenin B. *J. Am. Chem. Soc.* **2020**, *142* (19), 8550–8554. <https://doi.org/10.1021/jacs.0c03438>.
- (110) Zhou, R.; Liao, G.; Hu, C. [Epigenetic regulation of secondary metabolite biosynthesis in filamentous fungi: a review]. *Sheng Wu Gong Cheng Xue Bao* **2011**, *27* (8), 1142–1148.
- (111) Keller, N. P. Fungal Secondary Metabolism: Regulation, Function and Drug Discovery. *Nat. Rev. Microbiol.* **2019**, *17* (3), 167–180. <https://doi.org/10.1038/s41579-018-0121-1>.
- (112) Brakhage, A. A. Regulation of Fungal Secondary Metabolism. *Nat Rev Microbiol* **2013**, *11* (1), 21–32. <https://doi.org/10.1038/nrmicro2916>.
- (113) Huang, X.; Zhang, W.; Tang, S.; Wei, S.; Lu, X. Collaborative Biosynthesis of a Class of Bioactive Azaphilones by Two Separate Gene Clusters Containing Four PKS/NRPSs with Transcriptional Crosstalk in Fungi. *Angewandte Chemie International Edition n/a* (n/a). <https://doi.org/10.1002/anie.201915514>.
- (114) Chen, Y. W.; Yiu, C.-P. B.; Wong, K.-Y. Prediction of the SARS-CoV-2 (2019-NCoV) 3C-like Protease (3CLpro) Structure: Virtual Screening Reveals Velpatasvir, Ledipasvir, and Other Drug Repurposing Candidates. *F1000Res* **2020**, *9*, 129. <https://doi.org/10.12688/f1000research.22457.1>.
- (115) Stokes, J. M.; Yang, K.; Swanson, K.; Jin, W.; Cubillos-Ruiz, A.; Donghia, N. M.; MacNair, C. R.; French, S.; Carfrae, L. A.; Bloom-Ackerman, Z.; Tran, V. M.; Chiappino-Pepe, A.; Badran, A. H.; Andrews, I. W.; Chory, E. J.; Church, G. M.; Brown, E. D.; Jaakkola, T. S.; Barzilay, R.; Collins, J. J. A Deep Learning Approach to Antibiotic Discovery. *Cell* **2020**, *180* (4), 688–702.e13. <https://doi.org/10.1016/j.cell.2020.01.021>.
- (116) Selvameenal<sup>1</sup>, L.; Radhakrishnan<sup>1</sup>, M.; Balagurunathan<sup>2</sup>, R. Antibiotic Pigment from Desert Soil Actinomycetes; Biological Activity, Purification and Chemical Screening. *Indian Journal of Pharmaceutical Sciences* **2009**, *71* (5), 499. <https://doi.org/10.4103/0250-474X.58174>.
- (117) Yue, Y.; Yu, H.; Li, R.; Hu, L.; Liu, S.; Xing, R.; Li, P. Isolation and Identification of Antimicrobial Metabolites from Sea Anemone-Derived Fungus *Emericella* Sp. SMA01. *J. Ocean. Limnol.* **2021**, *39* (3), 1010–1019. <https://doi.org/10.1007/s00343-020-0203-6>.
- (118) Tang, Y.; Liu, Y.; Ruan, Q.; Zhao, M.; Zhao, Z.; Cui, H. Aspermeroterpenes A–C: Three Meroterpenoids from the Marine-Derived Fungus *Aspergillus Terreus* GZU-31-1. *Org. Lett.* **2020**, *22* (4), 1336–1339. <https://doi.org/10.1021/acs.orglett.9b04648>.

- (119) Brakhage, A. A.; Schroeckh, V. Fungal Secondary Metabolites – Strategies to Activate Silent Gene Clusters. *Fungal Genetics and Biology* **2011**, *48* (1), 15–22. <https://doi.org/10.1016/j.fgb.2010.04.004>.
- (120) Wei, Q.; Bai, J.; Yan, D.; Bao, X.; Li, W.; Liu, B.; Zhang, D.; Qi, X.; Yu, D.; Hu, Y. Genome Mining Combined Metabolic Shunting and OSMAC Strategy of an Endophytic Fungus Leads to the Production of Diverse Natural Products. *Acta Pharmaceutica Sinica B* **2021**, *11* (2), 572–587. <https://doi.org/10.1016/j.apsb.2020.07.020>.
- (121) Hemphill, C. F. P.; Sureechachaiyan, P.; Kassack, M. U.; Orfali, R. S.; Lin, W.; Daletos, G.; Proksch, P. OSMAC Approach Leads to New Fusarielin Metabolites from *Fusarium Tricinatum*. *J Antibiot* **2017**, *70* (6), 726–732. <https://doi.org/10.1038/ja.2017.21>.
- (122) Netzker, T.; Fischer, J.; Weber, J.; Mattern, D. J.; König, C. C.; Valiante, V.; Schroeckh, V.; Brakhage, A. A. Microbial Communication Leading to the Activation of Silent Fungal Secondary Metabolite Gene Clusters. *Frontiers in Microbiology* **2015**, *6*.
- (123) Cueto, M.; Jensen, P. R.; Kauffman, C.; Fenical, W.; Lobkovsky, E.; Clardy, J. Pestalone, a New Antibiotic Produced by a Marine Fungus in Response to Bacterial Challenge. *J. Nat. Prod.* **2001**, *64* (11), 1444–1446. <https://doi.org/10.1021/np0102713>.
- (124) *Wheldone: Characterization of a Unique Scaffold from the Coculture of Aspergillus fischeri and Xylaria flabelliformis* | *Organic Letters*. <https://pubs.acs.org/doi/10.1021/acs.orglett.0c00219> (accessed 2022-02-13).
- (125) Yu, J.-H.; Keller, N. Regulation of Secondary Metabolism in Filamentous Fungi. *Annu. Rev. Phytopathol.* **2005**, *43* (1), 437–458. <https://doi.org/10.1146/annurev.phyto.43.040204.140214>.
- (126) I. Tietz, J.; A Mitchell, D. *Using Genomics for Natural Product Structure Elucidation*. <https://www.ingentaconnect.com/content/ben/ctmc/2016/00000016/00000015/art00003> (accessed 2019-11-14).
- (127) Yan, Y.; Liu, Q.; Zang, X.; Yuan, S.; Bat-Erdene, U.; Nguyen, C.; Gan, J.; Zhou, J.; Jacobsen, S. E.; Tang, Y. Resistance-Gene-Directed Discovery of a Natural-Product Herbicide with a New Mode of Action. *Nature* **2018**, *559* (7714), 415–418. <https://doi.org/10.1038/s41586-018-0319-4>.
- (128) Yee, D. A.; Kakule, T. B.; Cheng, W.; Chen, M.; Chong, C. T. Y.; Hai, Y.; Hang, L. F.; Hung, Y.-S.; Liu, N.; Ohashi, M.; Okorafor, I. C.; Song, Y.; Tang, M.; Zhang, Z.; Tang, Y. Genome Mining of Alkaloidal Terpenoids from a Hybrid Terpene and Nonribosomal Peptide Biosynthetic Pathway. *J. Am. Chem. Soc.* **2020**, *142* (2), 710–714. <https://doi.org/10.1021/jacs.9b13046>.
- (129) He, X.; Zhang, M.; Guo, Y.-Y.; Mao, X.-M.; Chen, X.-A.; Li, Y.-Q. Revelation of the Balanol Biosynthetic Pathway in *Tolypocladium Ophioglossoides*. *Org. Lett.* **2018**, *20* (20), 6323–6326. <https://doi.org/10.1021/acs.orglett.8b01543>.
- (130) Zhao, Y.; Zhang, Y.-Y.; Liu, H.; Zhang, X.-S.; Ni, R.; Wang, P.-Y.; Gao, S.; Lou, H.-X.; Cheng, A.-X. Functional Characterization of a Liverworts BHLH Transcription Factor Involved in the Regulation of Bisbibenzyls and Flavonoids Biosynthesis. *BMC Plant Biology* **2019**, *19* (1), 497. <https://doi.org/10.1186/s12870-019-2109-z>.
- (131) Kennedy, J.; Turner, G. Delta-(L-Alpha-Aminoadipyl)-L-Cysteinyl-D-Valine Synthetase Is a Rate Limiting Enzyme for Penicillin Production in *Aspergillus Nidulans*. *Mol Gen Genet* **1996**, *253* (1–2), 189–197. <https://doi.org/10.1007/s004380050312>.

- (132) Maiya, S.; Grundmann, A.; Li, S.-M.; Turner, G. The Fumitremorgin Gene Cluster of *Aspergillus Fumigatus*: Identification of a Gene Encoding Brevianamide F Synthetase. *Chembiochem* **2006**, *7* (7), 1062–1069. <https://doi.org/10.1002/cbic.200600003>.
- (133) Bouhired, S.; Weber, M.; Kempf-Sontag, A.; Keller, N. P.; Hoffmeister, D. Accurate Prediction of the *Aspergillus Nidulans* Terrequinone Gene Cluster Boundaries Using the Transcriptional Regulator LaeA. *Fungal Genet Biol* **2007**, *44* (11), 1134–1145. <https://doi.org/10.1016/j.fgb.2006.12.010>.
- (134) Bok, J. W.; Hoffmeister, D.; Maggio-Hall, L. A.; Murillo, R.; Glasner, J. D.; Keller, N. P. Genomic Mining for *Aspergillus* Natural Products. *Chemistry & Biology* **2006**, *13* (1), 31–37. <https://doi.org/10.1016/j.chembiol.2005.10.008>.
- (135) Bok, J. W.; Keller, N. P. LaeA, a Regulator of Secondary Metabolism in *Aspergillus* Spp. *Eukaryot Cell* **2004**, *3* (2), 527–535. <https://doi.org/10.1128/EC.3.2.527-535.2004>.
- (136) Shwab, E. K.; Bok, J. W.; Tribus, M.; Galehr, J.; Graessle, S.; Keller, N. P. Histone Deacetylase Activity Regulates Chemical Diversity in *Aspergillus*. *Eukaryot Cell* **2007**, *6* (9), 1656–1664. <https://doi.org/10.1128/EC.00186-07>.
- (137) Soukup, A. A.; Chiang, Y.-M.; Bok, J. W.; Reyes-Dominguez, Y.; Oakley, B. R.; Wang, C. C. C.; Strauss, J.; Keller, N. P. Overexpression of the *Aspergillus Nidulans* Histone 4 Acetyltransferase EsaA Increases Activation of Secondary Metabolite Production. *Mol Microbiol* **2012**, *86* (2), 314–330. <https://doi.org/10.1111/j.1365-2958.2012.08195.x>.
- (138) Bok, J. W.; Chiang, Y.-M.; Szewczyk, E.; Reyes-Dominguez, Y.; Davidson, A. D.; Sanchez, J. F.; Lo, H.-C.; Watanabe, K.; Strauss, J.; Oakley, B. R.; Wang, C. C. C.; Keller, N. P. Chromatin-Level Regulation of Biosynthetic Gene Clusters. *Nat Chem Biol* **2009**, *5* (7), 462–464. <https://doi.org/10.1038/nchembio.177>.
- (139) Wu, G.; Sun, X.; Yu, G.; Wang, W.; Zhu, T.; Gu, Q.; Li, D. Cladosins A–E, Hybrid Polyketides from a Deep-Sea-Derived Fungus, *Cladosporium Sphaerospermum*. *J. Nat. Prod.* **2014**, *77* (2), 270–275. <https://doi.org/10.1021/np400833x>.
- (140) Zhang, Z.; He, X.; Wu, G.; Liu, C.; Lu, C.; Gu, Q.; Che, Q.; Zhu, T.; Zhang, G.; Li, D. Aniline-Tetramic Acids from the Deep-Sea-Derived Fungus *Cladosporium Sphaerospermum* L3P3 Cultured with the HDAC Inhibitor SAHA. *J. Nat. Prod.* **2018**, *81* (7), 1651–1657. <https://doi.org/10.1021/acs.jnatprod.8b00289>.
- (141) Williams, R. B.; Henrikson, J. C.; Hoover, A. R.; Lee, A. E.; Cichewicz, R. H. Epigenetic Remodeling of the Fungal Secondary Metabolome. *Org Biomol Chem* **2008**, *6* (11), 1895–1897. <https://doi.org/10.1039/b804701d>.
- (142) Li, D.; Tang, Y.; Lin, J.; Cai, W. Methods for Genetic Transformation of Filamentous Fungi. *Microb Cell Fact* **2017**, *16* (1), 168. <https://doi.org/10.1186/s12934-017-0785-7>.
- (143) Ruiz-Díez, B. Strategies for the Transformation of Filamentous Fungi. *J Appl Microbiol* **2002**, *92* (2), 189–195. <https://doi.org/10.1046/j.1365-2672.2002.01516.x>.
- (144) Wang, S.; Chen, H.; Tang, X.; Zhang, H.; Chen, W.; Chen, Y. Q. Molecular Tools for Gene Manipulation in Filamentous Fungi. *Appl Microbiol Biotechnol* **2017**, *101* (22), 8063–8075. <https://doi.org/10.1007/s00253-017-8486-z>.
- (145) Porteus, M. H.; Baltimore, D. Chimeric Nucleases Stimulate Gene Targeting in Human Cells. *Science* **2003**, *300* (5620), 763–763. <https://doi.org/10.1126/science.1078395>.



- (146) Frommer, W. B.; Ninnemann, O. Heterologous Expression of Genes in Bacterial, Fungal, Animal, and Plant Cells. *Annual Review of Plant Physiology and Plant Molecular Biology* **1995**, *46* (1), 419–444. <https://doi.org/10.1146/annurev.pp.46.060195.002223>.
- (147) Zobel, S.; Kumpfmüller, J.; Süßmuth, R. D.; Schweder, T. Bacillus Subtilis as Heterologous Host for the Secretory Production of the Non-Ribosomal Cyclodepsipeptide Enniatin. *Appl Microbiol Biotechnol* **2015**, *99* (2), 681–691. <https://doi.org/10.1007/s00253-014-6199-0>.
- (148) *HEx: A heterologous expression platform for the discovery of fungal natural products*. <https://www.science.org/doi/full/10.1126/sciadv.aar5459> (accessed 2022-02-16).
- (149) Tsunematsu, Y.; Ishiuchi, K.; Hotta, K.; Watanabe, K. Yeast-Based Genome Mining, Production and Mechanistic Studies of the Biosynthesis of Fungal Polyketide and Peptide Natural Products. *Nat Prod Rep* **2013**, *30* (8), 1139–1149. <https://doi.org/10.1039/c3np70037b>.
- (150) Punya, J.; Tachaleat, A.; Wattanachaisaereekul, S.; Haritakun, R.; Boonlarppradab, C.; Cheevadhanarak, S. Functional Expression of a Foreign Gene in Aspergillus Oryzae Producing New Pyrone Compounds. *Fungal Genet Biol* **2013**, *50*, 55–62. <https://doi.org/10.1016/j.fgb.2012.10.005>.
- (151) Oikawa, H. Reconstitution of Biosynthetic Machinery of Fungal Natural Products in Heterologous Hosts. *Bioscience, Biotechnology, and Biochemistry* **2020**, *84* (3), 433–444. <https://doi.org/10.1080/09168451.2019.1690976>.
- (152) Ugai, T.; Minami, A.; Gomi, K.; Oikawa, H. Genome Mining Approach for Harnessing the Cryptic Gene Cluster in Alternaria Solani: Production of PKS–NRPS Hybrid Metabolite, Didymellamide B. *Tetrahedron Letters* **2016**, *57* (25), 2793–2796. <https://doi.org/10.1016/j.tetlet.2016.05.043>.
- (153) Ye, Y.; Minami, A.; Mandi, A.; Liu, C.; Taniguchi, T.; Kuzuyama, T.; Monde, K.; Gomi, K.; Oikawa, H. Genome Mining for Sesterterpenes Using Bifunctional Terpene Synthases Reveals a Unified Intermediate of Di/Sesterterpenes. *J. Am. Chem. Soc.* **2015**, *137* (36), 11846–11853. <https://doi.org/10.1021/jacs.5b08319>.
- (154) Yan, D.; Matsuda, Y. Genome Mining-Driven Discovery of 5-Methylorsellinate-Derived Meroterpenoids from Aspergillus Funiculosus. *Org. Lett.* **2021**, *23* (8), 3211–3215. <https://doi.org/10.1021/acs.orglett.1c00951>.
- (155) Zhang, Z.; Tamura, Y.; Tang, M.; Qiao, T.; Sato, M.; Otsu, Y.; Sasamura, S.; Taniguchi, M.; Watanabe, K.; Tang, Y. Biosynthesis of the Immunosuppressant (–)-FR901483. *J. Am. Chem. Soc.* **2021**, *143* (1), 132–136. <https://doi.org/10.1021/jacs.0c12352>.
- (156) Zaehle, C.; Gressler, M.; Shelest, E.; Geib, E.; Hertweck, C.; Brock, M. Terrein Biosynthesis in Aspergillus Terreus and Its Impact on Phytotoxicity. *Chem Biol* **2014**, *21* (6), 719–731. <https://doi.org/10.1016/j.chembiol.2014.03.010>.
- (157) Yaegashi, J.; Oakley, B. R.; Wang, C. C. Recent Advances in Genome Mining of Secondary Metabolite Biosynthetic Gene Clusters and the Development of Heterologous Expression Systems in Aspergillus Nidulans. *Journal of Industrial Microbiology and Biotechnology* **2014**, *41* (2), 433–442. <https://doi.org/10.1007/s10295-013-1386-z>.
- (158) *Frontiers | Recent advances in genome mining of secondary metabolites in Aspergillus terreus | Microbiology*. <https://www.frontiersin.org/articles/10.3389/fmicb.2014.00717/full> (accessed 2022-02-09).

- (159) Chiang, Y.-M.; Oakley, C. E.; Ahuja, M.; Entwistle, R.; Schultz, A.; Chang, S.-L.; Sung, C. T.; Wang, C. C. C.; Oakley, B. R. An Efficient System for Heterologous Expression of Secondary Metabolite Genes in *Aspergillus Nidulans*. *J Am Chem Soc* **2013**, *135* (20), 7720–7731. <https://doi.org/10.1021/ja401945a>.
- (160) Chiang, Y.-M.; Ahuja, M.; Oakley, C. E.; Entwistle, R.; Zutz, C.; Wang, C. C. C.; Oakley, B. R. Development of Genetic Dereplication Strains in *Aspergillus Nidulans* Results in the Discovery of Aspercryptin. *Angew Chem Int Ed Engl* **2016**, *55* (5), 1662–1665. <https://doi.org/10.1002/anie.201507097>.
- (161) Yin, W.-B.; Chooi, Y. H.; Smith, A. R.; Cacho, R. A.; Hu, Y.; White, T. C.; Tang, Y. Discovery of Cryptic Polyketide Metabolites from Dermatophytes Using Heterologous Expression in *Aspergillus Nidulans*. *ACS Synth. Biol.* **2013**, *2* (11), 629–634. <https://doi.org/10.1021/sb400048b>.
- (162) Cobb, R. E.; Zhao, H. Direct Cloning of Large Genomic Sequences. *Nat Biotechnol* **2012**, *30* (5), 405–406. <https://doi.org/10.1038/nbt.2207>.
- (163) Xie, L.; Zhang, L.; Wang, C.; Wang, X.; Xu, Y.; Yu, H.; Wu, P.; Li, S.; Han, L.; Gunatilaka, A. A. L.; Wei, X.; Lin, M.; Molnár, I.; Xu, Y. Methylglucosylation of Aromatic Amino and Phenolic Moieties of Drug-like Biosynthons by Combinatorial Biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115* (22). <https://doi.org/10.1073/pnas.1716046115>.
- (164) Wang, C.; Wang, X.; Zhang, L.; Yue, Q.; Liu, Q.; Xu, Y.; Gunatilaka, A. A. L.; Wei, X.; Xu, Y.; Molnár, I. Intrinsic and Extrinsic Programming of Product Chain Length and Release Mode in Fungal Collaborating Iterative Polyketide Synthases. *J. Am. Chem. Soc.* **2020**, *142* (40), 17093–17104. <https://doi.org/10.1021/jacs.0c07050>.
- (165) Jarczynska, Z. D.; Rendsvig, J. K. H.; Pagels, N.; Viana, V. R.; Nødvig, C. S.; Kirchner, F. H.; Strucko, T.; Nielsen, M. L.; Mortensen, U. H. DIVERSIFY: A Fungal Multispecies Gene Expression Platform. *ACS Synth. Biol.* **2021**, *10* (3), 579–588. <https://doi.org/10.1021/acssynbio.0c00587>.
- (166) Mózsik, L.; Pohl, C.; Meyer, V.; Bovenberg, R. A. L.; Nygård, Y.; Driessen, A. J. M. Modular Synthetic Biology Toolkit for Filamentous Fungi. *ACS Synth. Biol.* **2021**, *10* (11), 2850–2861. <https://doi.org/10.1021/acssynbio.1c00260>.
- (167) Song, T.; Chen, M.; Ge, Z.-W.; Chai, W.; Li, X.-C.; Zhang, Z.; Lian, X.-Y. Bioactive Penicypyrrodiether A, an Adduct of GKK1032 Analogue and Phenol A Derivative, from a Marine-Sourced Fungus *Penicillium* Sp. ZZ380. *J. Org. Chem.* **2018**, *83* (21), 13395–13401. <https://doi.org/10.1021/acs.joc.8b02172>.
- (168) Oikawa, H. Biosynthesis of Structurally Unique Fungal Metabolite GKK1032A<sub>2</sub>: Indication of Novel Carbocyclic Formation Mechanism in Polyketide Biosynthesis. *J. Org. Chem.* **2003**, *68* (9), 3552–3557. <https://doi.org/10.1021/jo0267596>.
- (169) Isaka, M.; Rugseree, N.; Maithip, P.; Kongsaree, P.; Prabpai, S.; Thebtaranonth, Y. Hirsutellones A–E, Antimycobacterial Alkaloids from the Insect Pathogenic Fungus *Hirsutella Nivea* BCC 2594. *Tetrahedron* **2005**, *61* (23), 5577–5583. <https://doi.org/10.1016/j.tet.2005.03.099>.
- (170) Shiono, Y.; Shimanuki, K.; Hiramatsu, F.; Koseki, T.; Tetsuya, M.; Fujisawa, N.; Kimura, K. Pyrrospirones A and B, Apoptosis Inducers in HL-60 Cells, from an Endophytic Fungus, *Neonectria Ramulariae* Wollenw KS-246. *Bioorganic & Medicinal Chemistry Letters* **2008**, *18* (23), 6050–6053. <https://doi.org/10.1016/j.bmcl.2008.10.032>.

- (171) Ebrahim, W.; Aly, A. H.; Wray, V.; Mándi, A.; Teiten, M.-H.; Gaascht, F.; Orlikova, B.; Kassack, M. U.; Lin, W.; Diederich, M.; Kurtán, T.; Debbab, A.; Proksch, P. Embellicines A and B: Absolute Configuration and NF-KB Transcriptional Inhibitory Activity. *J. Med. Chem.* **2013**, *56* (7), 2991–2999. <https://doi.org/10.1021/jm400034b>.
- (172) Chen, S.; Shen, H.; Zhang, P.; Cheng, H.; Dai, X.; Liu, L. Anti-Glioma Trichobamide A with an Unprecedented Tetrahydro-5 H -Furo[2,3- b ]Pyrrol-5-One Functionality from Ascidian-Derived Fungus *Trichobotrys Effuse* 4729. *Chem. Commun.* **2019**, *55* (10), 1438–1441. <https://doi.org/10.1039/C8CC08970A>.
- (173) Opatz, T.; Kolshorn, H.; Thines, E.; Anke, H. Ascomycones A–C, Heptaketide Metabolites from an Unidentified Ascomycete. *J. Nat. Prod.* **2008**, *71* (11), 1973–1976. <https://doi.org/10.1021/np800570w>.
- (174) Wijeratne, E. M. K.; He, H.; Franzblau, S. G.; Hoffman, A. M.; Gunatilaka, A. A. L. Phomapyrrolidones A–C, Antitubercular Alkaloids from the Endophytic Fungus *Phoma* Sp. NRRL 46751. *J. Nat. Prod.* **2013**, *76* (10), 1860–1865. <https://doi.org/10.1021/np400391p>.
- (175) Isaka, M.; Prathumpai, W.; Wongsas, P.; Tanticharoen, M. Hirsutellone F, a Dimer of Antitubercular Alkaloids from the Seed Fungus *Trichoderma* Species BCC 7579. *ChemInform* **2006**, *37* (43). <https://doi.org/10.1002/chin.200643203>.
- (176) Shiono, Y.; Furukawa, M.; Koseki, T.; Kwon, E.; Kurniawan, A. H.; Sato, S.; Harneti, D.; Maharani, R.; Supratman, U.; Uesugi, S.; Kimura, K. A Pyrrocidine Derivative Produced by Fungus *Neonectria Ramulariae* In-2 Isolated from a Beetle *Holotrichia Picea*. *Phytochemistry Letters* **2018**, *26*, 120–124. <https://doi.org/10.1016/j.phytol.2018.05.030>.
- (177) Casella, T. M.; Eparvier, V.; Mandavid, H.; Bendelac, A.; Odonne, G.; Dayan, L.; Duplais, C.; Espindola, L. S.; Stien, D. Antimicrobial and Cytotoxic Secondary Metabolites from Tropical Leaf Endophytes: Isolation of Antibacterial Agent Pyrrocidine C from *Lewia Infectoria* SNB-GTC2402. *Phytochemistry* **2013**, *96*, 370–377. <https://doi.org/10.1016/j.phytochem.2013.10.004>.
- (178) Shiono, Y.; Kosukegawa, A.; Koseki, T.; Murayama, T.; Kwon, E.; Uesugi, S.; Kimura, K. A Dimeric Pyrrocidine from *Neonectria Ramulariae* Is an Inhibitor of Prolyl Oligopeptidase. *Phytochemistry Letters* **2012**, *5* (1), 91–95. <https://doi.org/10.1016/j.phytol.2011.10.008>.
- (179) Song, T.; Tang, M.; Ge, H.; Chen, M.; Lian, X.; Zhang, Z. Novel Bioactive Penicypyrroether A and Pyrrospirone J from the Marine-Derived *Penicillium* Sp. ZZ380. *Marine Drugs* **2019**, *17* (5), 292. <https://doi.org/10.3390/md17050292>.
- (180) Ear, A.; Amand, S.; Blanchard, F.; Blond, A.; Dubost, L.; Buisson, D.; Nay, B. Direct Biosynthetic Cyclization of a Distorted Paracyclophane Highlighted by Double Isotopic Labelling of L -Tyrosine. *Org. Biomol. Chem.* **2015**, *13* (12), 3662–3666. <https://doi.org/10.1039/C5OB00114E>.
- (181) Oikawa, H.; Katayama, K.; Suzuki, Y.; Ichihara, A. Enzymatic Activity Catalysing Exo-Selective Diels–Alder Reaction in Solanapyrone Biosynthesis. *J. Chem. Soc., Chem. Commun.* **1995**, No. 13, 1321–1322. <https://doi.org/10.1039/C39950001321>.
- (182) Exo vs Endo Products In The Diels Alder: How To Tell Them Apart. *Master Organic Chemistry*, 2018.

- (183) Jamieson, C. S.; Ohashi, M.; Liu, F.; Tang, Y.; Houk, K. N. The Expanding World of Biosynthetic Pericyclases: Cooperation of Experiment and Theory for Discovery. *Nat. Prod. Rep.* **2019**, *36* (5), 698–713. <https://doi.org/10.1039/C8NP00075A>.
- (184) Zheng, Q.; Tian, Z.; Liu, W. Recent Advances in Understanding the Enzymatic Reactions of [4+2] Cycloaddition and Spiroketalization. *Current Opinion in Chemical Biology* **2016**, *31*, 95–102. <https://doi.org/10.1016/j.cbpa.2016.01.020>.
- (185) Oikawa, H.; Katayama, K.; Suzuki, Y.; Ichihara, A. Enzymatic Activity Catalysing Exo-Selective Diels–Alder Reaction in Solanapyrone Biosynthesis. *J. Chem. Soc., Chem. Commun.* **1995**, No. 13, 1321–1322. <https://doi.org/10.1039/C39950001321>.
- (186) Ma, S. M.; Li, J. W.-H.; Choi, J. W.; Zhou, H.; Lee, K. K. M.; 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), 589–592. <https://doi.org/10.1126/science.1175602>.
- (187) Watanabe, K.; Mie, T.; Ichihara, A.; Oikawa, H.; Honma, M. Detailed Reaction Mechanism of Macrophomate Synthase: EXTRAORDINARY ENZYME CATALYZING FIVE-STEP TRANSFORMATION FROM 2-PYRONES TO BENZOATES \*. *Journal of Biological Chemistry* **2000**, *275* (49), 38393–38401. <https://doi.org/10.1074/jbc.M003119200>.
- (188) Ose, T.; Watanabe, K.; Mie, T.; Honma, M.; Watanabe, H.; Yao, M.; Oikawa, H.; Tanaka, I. Insight into a Natural Diels–Alder Reaction from the Structure of Macrophomate Synthase. *Nature* **2003**, *422* (6928), 185–189. <https://doi.org/10.1038/nature01454>.
- (189) Kim, H. J.; Ruzsyczky, M. W.; Choi, S.; Liu, Y.; Liu, H. Enzyme-Catalysed [4+2] Cycloaddition Is a Key Step in the Biosynthesis of Spinosyn A. *Nature* **2011**, *473* (7345), 109–112. <https://doi.org/10.1038/nature09981>.
- (190) Cai, Y.; Hai, Y.; Ohashi, M.; Jamieson, C. S.; Garcia-Borras, M.; Houk, K. N.; Zhou, J.; Tang, Y. Structural Basis for Stereoselective Dehydration and Hydrogen-Bonding Catalysis by the SAM-Dependent Pericyclase LepI. *Nat. Chem.* **2019**, *11* (9), 812–820. <https://doi.org/10.1038/s41557-019-0294-x>.
- (191) *FAD-dependent enzyme-catalysed intermolecular [4+2] cycloaddition in natural product biosynthesis | Nature Chemistry*. <https://www.nature.com/articles/s41557-020-0467-7> (accessed 2021-04-20).
- (192) Kato, N.; Nogawa, T.; Hirota, H.; Jang, J.-H.; Takahashi, S.; Ahn, J.; Osada, H. A New Enzyme Involved in the Control of the Stereochemistry in the Decalin Formation during Equisetin Biosynthesis. *Biochemical and biophysical research communications* **2015**, *460*. <https://doi.org/10.1016/j.bbrc.2015.03.011>.
- (193) Sato, M.; Yagishita, F.; Mino, T.; Uchiyama, N.; Patel, A.; Chooi, Y.-H.; Goda, Y.; Xu, W.; Noguchi, H.; Yamamoto, T.; Hotta, K.; Houk, K. N.; Tang, Y.; Watanabe, K. Involvement of Lipocalin-like CghA in Decalin-Forming Stereoselective Intramolecular [4+2] Cycloaddition. *Chembiochem* **2015**, *16* (16), 2294–2298. <https://doi.org/10.1002/cbic.201500386>.
- (194) Gao, L.; Zou, Y.; Liu, X.; Yang, J.; Du, X.; Wang, J.; Yu, X.; Fan, J.; Jiang, M.; Li, Y.; Houk, K. N.; Lei, X. Enzymatic Control of Endo- and Exo-Stereoselective Diels–Alder Reactions with Broad Substrate Scope. *Nat Catal* **2021**, *4* (12), 1059–1069. <https://doi.org/10.1038/s41929-021-00717-8>.

- (195) Arai, N.; Ui, H.; Omura, S.; Kuwajima, I. Studies toward the Total Synthesis of GKK1032A2, a Structurally Unique Anti-tumor Compound: Stereoselective Construction of the Tricarboyclic System. *Synlett* **2005**, 2005 (11), 1691–1694. <https://doi.org/10.1055/s-2005-871549>.
- (196) Halo, L. M.; Marshall, J. W.; Yakasai, A. A.; Song, Z.; Butts, C. P.; Crump, M. P.; Heneghan, M. N.; Bailey, A. M.; Simpson, T. J.; Lazarus, C. M.; Cox, R. J. Authentic Heterologous Expression of the Tenellin Iterative Polyketide Synthase Nonribosomal Peptide Synthetase Requires Coexpression with an Enoyl Reductase. *ChemBioChem* **2008**, 9 (4), 585–594. <https://doi.org/10.1002/cbic.200700390>.
- (197) Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. AntiSMASH: Rapid Identification, Annotation and Analysis of Secondary Metabolite Biosynthesis Gene Clusters in Bacterial and Fungal Genome Sequences. *Nucleic Acids Research* **2011**, 39 (suppl\_2), W339–W346. <https://doi.org/10.1093/nar/gkr466>.
- (198) Ansari, Mohd. Z.; Yadav, G.; Gokhale, R. S.; Mohanty, D. NRPS-PKS: A Knowledge-Based Resource for Analysis of NRPS/PKS Megasyntases. *Nucleic Acids Res* **2004**, 32 (Web Server issue), W405–W413. <https://doi.org/10.1093/nar/gkh359>.
- (199) Besemer, J.; Borodovsky, M. GeneMark: Web Software for Gene Finding in Prokaryotes, Eukaryotes and Viruses. *Nucleic Acids Res* **2005**, 33 (Web Server issue), W451–454. <https://doi.org/10.1093/nar/gki487>.
- (200) Liu, X.; Walsh, C. T. Cyclopiazonic Acid Biosynthesis in *Aspergillus* Sp.: Characterization of a Reductase-like R\* Domain in Cyclopiazonate Synthetase That Forms and Releases Cyclo-Acetoacetyl-L-Tryptophan. *Biochemistry* **2009**, 48 (36), 8746–8757. <https://doi.org/10.1021/bi901123r>.
- (201) *Crystal Structures of Fsa2 and Phm7 Catalyzing [4 + 2] Cycloaddition Reactions with Reverse Stereoselectivities in Equisetin and Phomasetin Biosynthesis | ACS Omega*. <https://pubs.acs.org/doi/10.1021/acsomega.1c01593> (accessed 2022-04-20).
- (202) Oliner, J. D.; Kinzler, K. W.; Vogelstein, B. In Vivo Cloning of PCR Products in *E. Coli*. *Nucleic Acids Res* **1993**, 21 (22), 5192–5197.
- (203) García-Nafría, J.; Watson, J. F.; Greger, I. H. IVA Cloning: A Single-Tube Universal Cloning System Exploiting Bacterial In Vivo Assembly. *Sci Rep* **2016**, 6 (1), 27459. <https://doi.org/10.1038/srep27459>.
- (204) Gravelat, F. N.; Askew, D. S.; Sheppard, D. C. Targeted Gene Deletion in *Aspergillus Fumigatus* Using the Hygromycin-Resistance Split-Marker Approach. In *Host-Fungus Interactions: Methods and Protocols*; Brand, A. C., MacCallum, D. M., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2012; pp 119–130. [https://doi.org/10.1007/978-1-61779-539-8\\_8](https://doi.org/10.1007/978-1-61779-539-8_8).
- (205) *Occurrence of pyrrocidine and dihydroresorcyliide production among Acremonium zeae populations from maize grown in different regions: Canadian Journal of Plant Pathology: Vol 30, No 3*. <https://www.tandfonline.com/doi/abs/10.1080/07060660809507540> (accessed 2022-03-10).
- (206) Qi, X.; Li, X.; Zhao, J.; He, N.; Li, Y.; Zhang, T.; Wang, S.; Yu, L.; Xie, Y. GKK1032C, a New Alkaloid Compound from the Endophytic Fungus *Penicillium* Sp. CCCC 400817 with Activity against Methicillin-Resistant *S. Aureus*. *J Antibiot* **2019**, 72 (4), 237–240. <https://doi.org/10.1038/s41429-019-0144-5>.
- (207) Sato, M.; Dander, J. E.; Sato, C.; Hung, Y.-S.; Gao, S.-S.; Tang, M.-C.; Hang, L.; Winter, J. M.; Garg, N. K.; Watanabe, K.; Tang, Y. Collaborative Biosynthesis of Maleimide- and Succinimide-Containing

Natural Products by Fungal Polyketide Megasyntases. *J. Am. Chem. Soc.* **2017**, *139* (15), 5317–5320. <https://doi.org/10.1021/jacs.7b02432>.

(208) Song, Z.; Bakeer, W.; Marshall, J. W.; Yakasai, A. A.; Khalid, R. M.; Collemare, J.; Skellam, E.; Tharreau, D.; Lebrun, M.-H.; Lazarus, C. M.; Bailey, A. M.; Simpson, T. J.; Cox, R. J. Heterologous Expression of the Avirulence Gene ACE1 from the Fungal Rice Pathogen Magnaporthe Oryzae. *Chem. Sci.* **2015**, *6* (8), 4837–4845. <https://doi.org/10.1039/C4SC03707C>.

(209) Niehaus, E.-M.; Kleigrew, K.; Wiemann, P.; Studt, L.; Sieber, C. M. K.; Connolly, L. R.; Freitag, M.; Güldener, U.; Tudzynski, B.; Humpf, H.-U. Genetic Manipulation of the Fusarium Fujikuroi Fusarin Gene Cluster Yields Insight into the Complex Regulation and Fusarin Biosynthetic Pathway. *Chem Biol* **2013**, *20* (8), 1055–1066. <https://doi.org/10.1016/j.chembiol.2013.07.004>.

(210) Li, Y.; Weissman, K. J.; Müller, R. Myxochelin Biosynthesis: Direct Evidence for Two- and Four-Electron Reduction of a Carrier Protein-Bound Thioester. *J. Am. Chem. Soc.* **2008**, *130* (24), 7554–7555. <https://doi.org/10.1021/ja8025278>.

(211) Zhang, H.; Hantke, V.; Bruhnke, P.; Skellam, E. J.; Cox, R. J. Chemical and Genetic Studies on the Formation of Pyrrolones During the Biosynthesis of Cytochalasans. *Chem. Eur. J.* **2021**, *27* (9), 3106–3113. <https://doi.org/10.1002/chem.202004444>.

(212) Hantke, V.; Skellam, E. J.; Cox, R. J. Evidence for Enzyme Catalysed Intramolecular [4+2] Diels–Alder Cyclization during the Biosynthesis of Pyrichalasin H. *Chem. Commun.* **2020**, *56* (19), 2925–2928. <https://doi.org/10.1039/C9CC09590J>.

(213) Bartlett, P. D.; Banavali, R. Spontaneous Oxygenation of Cyclic Olefins. Effects of Strain. *J. Org. Chem.* **1991**, *56* (21), 6043–6050. <https://doi.org/10.1021/jo00021a016>.

(214) Wicklow, D. T.; Poling, S. M. Antimicrobial Activity of Pyrrocidines from *Acremonium Zeae* Against Endophytes and Pathogens of Maize. *Phytopathology* **2009**, *99* (1), 109–115. <https://doi.org/10.1094/PHYTO-99-1-0109>.

(215) Ear, A.; Amand, S.; Blanchard, F.; Blond, A.; Dubost, L.; Buisson, D.; Nay, B. Direct Biosynthetic Cyclization of a Distorted Paracyclophane Highlighted by Double Isotopic Labelling of L-Tyrosine. *Org. Biomol. Chem.* **2015**, *13* (12), 3662–3666. <https://doi.org/10.1039/C5OB00114E>.

(216) Ohashi, M.; Kakule, T. B.; Tang, M.-C.; Jamieson, C. S.; Liu, M.; Zhao, Y.-L.; Houk, K. N.; Tang, Y. Biosynthesis of Para-Cyclophane-Containing Hirsutellone Family of Fungal Natural Products. *J. Am. Chem. Soc.* **2021**, *143* (15), 5605–5609. <https://doi.org/10.1021/jacs.1c00098>.

(217) Liu, Z.; Li, W.; Zhang, P.; Fan, J.; Zhang, F.; Wang, C.; Li, S.; Sun, Y.; Chen, S.; Yin, W. Tricarbocyclic Core Formation of Tyrosine-Decahydrofluorenes Implies a Three-Enzyme Cascade with XenF-Mediated Sigmatropic Rearrangement as a Prerequisite. *Acta Pharmaceutica Sinica B* **2021**, *11* (11), 3655–3664. <https://doi.org/10.1016/j.apsb.2021.03.034>.

(218) Wicklow, D. T.; Roth, S.; Deyrup, S. T.; Gloer, J. B. A Protective Endophyte of Maize: *Acremonium Zeae* Antibiotics Inhibitory to *Aspergillus Flavus* and *Fusarium Verticillioides*. *Mycological Research* **2005**, *109* (5), 610–618. <https://doi.org/10.1017/S0953756205002820>.

(219) Gao, M.; Glenn, A. E.; Gu, X.; Mitchell, T. R.; Satterlee, T.; Duke, M. V.; Scheffler, B. E.; Gold, S. E. Pyrrocidine, a Molecular off Switch for Fumonisin Biosynthesis. *PLOS Pathogens* **2020**, *16* (7), e1008595. <https://doi.org/10.1371/journal.ppat.1008595>.

- (220) Billingsley, J. M.; DeNicola, A. B.; Tang, Y. Technology Development for Natural Product Biosynthesis in *Saccharomyces Cerevisiae*. *Curr Opin Biotechnol* **2016**, *42*, 74–83. <https://doi.org/10.1016/j.copbio.2016.02.033>.
- (221) *A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly* | *ACS Synthetic Biology*. <https://pubs.acs.org/doi/10.1021/sb500366v> (accessed 2022-02-25).
- (222) Ugai, T.; Minami, A.; Gomi, K.; Oikawa, H. Genome Mining Approach for Harnessing the Cryptic Gene Cluster in *Alternaria Solani*: Production of PKS–NRPS Hybrid Metabolite, Didymellamide B. *Tetrahedron Letters* **2016**, *57* (25), 2793–2796. <https://doi.org/10.1016/j.tetlet.2016.05.043>.
- (223) Haga, A.; Tamoto, H.; Ishino, M.; Kimura, E.; Sugita, T.; Kinoshita, K.; Takahashi, K.; Shiro, M.; Koyama, K. Pyridone Alkaloids from a Marine-Derived Fungus, *Stagonosporopsis Cucurbitacearum*, and Their Activities against Azole-Resistant *Candida Albicans*. *J. Nat. Prod.* **2013**, *76* (4), 750–754. <https://doi.org/10.1021/np300876t>.
- (224) *Sequencing of Aspergillus nidulans and comparative analysis with A. fumigatus and A. oryzae* | *Nature*. <https://www.nature.com/articles/nature04341> (accessed 2022-03-20).
- (225) Kim, L. J.; Ohashi, M.; Zhang, Z.; Tan, D.; Asay, M.; Cascio, D.; Rodriguez, J. A.; Tang, Y.; Nelson, H. M. Prospecting for Natural Product Structural Complexity Using Genome Mining and Microcrystal Electron Diffraction. *Nat Chem Biol* **2021**, *17* (8), 872–877. <https://doi.org/10.1038/s41589-021-00834-2>.
- (226) Ohashi, M.; Liu, F.; Hai, Y.; Chen, M.; Tang, M.; Yang, Z.; Sato, M.; Watanabe, K.; Houk, K. N.; Tang, Y. SAM-Dependent Enzyme-Catalysed Pericyclic Reactions in Natural Product Biosynthesis. *Nature* **2017**, *549* (7673), 502–506. <https://doi.org/10.1038/nature23882>.
- (227) Aljanabi, S. M.; Martinez, I. Universal and Rapid Salt-Extraction of High Quality Genomic DNA for PCR-Based Techniques. *Nucleic Acids Research* **1997**, *25* (22), 4692–4693. <https://doi.org/10.1093/nar/25.22.4692>.
- (228) He, H.; Yang, H. Y.; Bigelis, R.; Solum, E. H.; Greenstein, M.; Carter, G. T. Pyrrocidines A and B, New Antibiotics Produced by a Filamentous Fungus. *Tetrahedron Letters* **2002**, *43* (9), 1633–1636. [https://doi.org/10.1016/S0040-4039\(02\)00099-0](https://doi.org/10.1016/S0040-4039(02)00099-0).
- (229) Wicklow, D. T.; Poling, S. M.; Summerbell, R. C. Occurrence of Pyrrocidine and Dihydroresorcylic Acid Production among *Acremonium Zeae* Populations from Maize Grown in Different Regions. *Canadian Journal of Plant Pathology* **2008**, *30* (3), 425–433. <https://doi.org/10.1080/07060660809507540>.
- (230) Shiono, Y.; Furukawa, M.; Koseki, T.; Kwon, E.; Kurniawan, A. H.; Sato, S.; Harneti, D.; Maharani, R.; Supratman, U.; Uesugi, S.; Kimura, K. A Pyrrocidine Derivative Produced by Fungus *Neonectria Ramulariae* In-2 Isolated from a Beetle *Holotrichia Picea*. *Phytochemistry Letters* **2018**, *26*, 120–124. <https://doi.org/10.1016/j.phytol.2018.05.030>.
- (231) Błaszczuk, L.; Waśkiewicz, A.; Gromadzka, K.; Mikołajczak, K.; Chełkowski, J. *Sarocladium* and *Lecanicillium* Associated with Maize Seeds and Their Potential to Form Selected Secondary Metabolites. **2021**, 11.
- (232) Uesugi, S.; Fujisawa, N.; Yoshida, J.; Watanabe, M.; Dan, S.; Yamori, T.; Shiono, Y.; Kimura, K. Pyrrocidine A, a Metabolite of Endophytic Fungi, Has a Potent Apoptosis-Inducing Activity against HL60

Cells through Caspase Activation via the Michael Addition. *J Antibiot* **2016**, *69* (3), 133–140. <https://doi.org/10.1038/ja.2015.103>.

(233) Ebrahimi, K. S.; Ansari, M.; Hosseyni Moghaddam, M. S.; Ebrahimi, Z.; Salehi, Z.; Shahlaei, M.; Moradi, S. In Silico Investigation on the Inhibitory Effect of Fungal Secondary Metabolites on RNA Dependent RNA Polymerase of SARS-CoV-II: A Docking and Molecular Dynamic Simulation Study. *Computers in Biology and Medicine* **2021**, *135*, 104613. <https://doi.org/10.1016/j.combiomed.2021.104613>.

(234) Wicklow, D. T.; Roth, S.; Deyrup, S. T.; Gloer, J. B. A Protective Endophyte of Maize: *Acremonium Zeae* Antibiotics Inhibitory to *Aspergillus Flavus* and *Fusarium Verticillioides* 1 Dedicated to John Webster on the Occasion of His 80th Birthday. *Mycological Research* **2005**, *109* (5), 610–618. <https://doi.org/10.1017/S0953756205002820>.

(235) Oikawa, H. Biosynthesis of Structurally Unique Fungal Metabolite GKK1032A2: Indication of Novel Carbocyclic Formation Mechanism in Polyketide Biosynthesis. *J. Org. Chem.* **2003**, *68* (9), 3552–3557. <https://doi.org/10.1021/jo0267596>.

(236) Chen, S.; Shen, H.; Zhang, P.; Cheng, H.; Dai, X.; Liu, L. Anti-Glioma Trichobamide A with an Unprecedented Tetrahydro-5 *H*-Furo[2,3-*b*]Pyrrol-5-One Functionality from Ascidian-Derived Fungus *Trichobotrys Effuse* 4729. *Chemical Communications* **2019**, *55* (10), 1438–1441. <https://doi.org/10.1039/C8CC08970A>.

(237) Chen, Y.; Liu, Z.; Huang, Y.; Liu, L.; He, J.; Wang, L.; Yuan, J.; She, Z. Ascomylactams A–C, Cytotoxic 12- or 13-Membered-Ring Macrocyclic Alkaloids Isolated from the Mangrove Endophytic Fungus *Didymella* Sp. CYSK-4, and Structure Revisions of Phomapyrrolidones A and C. *J. Nat. Prod.* **2019**, *82* (7), 1752–1758. <https://doi.org/10.1021/acs.jnatprod.8b00918>.

(238) Wang, L.; Huang, Y.; Huang, C.; Yu, J.; Zheng, Y.; Chen, Y.; She, Z.; Yuan, J. A Marine Alkaloid, Ascomylactam A, Suppresses Lung Tumorigenesis via Inducing Cell Cycle G1/S Arrest through ROS/Akt/Rb Pathway. *Marine Drugs* **2020**, *18* (10), 494. <https://doi.org/10.3390/md18100494>.

(239) Song, T.; Chen, M.; Chai, W.; Zhang, Z.; Lian, X.-Y. New Bioactive Pyrrospirones C–I from a Marine-Derived Fungus *Penicillium* Sp. ZZ380. *Tetrahedron* **2018**, *74* (8), 884–891. <https://doi.org/10.1016/j.tet.2018.01.015>.

(240) Ebrahim, W.; Aly, A. H.; Wray, V.; Mándi, A.; Teiten, M.-H.; Gaascht, F.; Orlikova, B.; Kassack, M. U.; Lin, W.; Diederich, M.; Kurtán, T.; Debbab, A.; Proksch, P. Embellicines A and B: Absolute Configuration and NF- $\kappa$ B Transcriptional Inhibitory Activity. *Journal of Medicinal Chemistry* **2013**, *56* (7), 2991–2999. <https://doi.org/10.1021/jm400034b>.

(241) Ohashi, M.; Kakule, T. B.; Tang, M.-C.; Jamieson, C. S.; Liu, M.; Houk, K. N.; Tang, Y. Biosynthesis of Para-Cyclophane Containing Hirsutellone-Family of Fungal Natural Products. **71**.

(242) Liu, Z.; Li, W.; Zhang, P.; Fan, J.; Zhang, F.; Wang, C.; Li, S.; Sun, Y.; Chen, S.; Yin, W. Tricarbocyclic Core Formation of Tyrosine-Decahydrofluorenes Implies a Three-Enzyme Cascade with XenF-Mediated Sigmatropic Rearrangement as a Prerequisite. *Acta Pharmaceutica Sinica B* **2021**, S2211383521001064. <https://doi.org/10.1016/j.apsb.2021.03.034>.

(243) Fisch, K. M.; Bakeer, W.; Yakasai, A. A.; Song, Z.; Pedrick, J.; Wasil, Z.; Bailey, A. M.; Lazarus, C. M.; Simpson, T. J.; Cox, R. J. Rational Domain Swaps Decipher Programming in Fungal Highly Reducing Polyketide Synthases and Resurrect an Extinct Metabolite. *J. Am. Chem. Soc.* **2011**, *133* (41), 16635–16641. <https://doi.org/10.1021/ja206914q>.



- (244) Hamada, W.; Reignault, P.; Bompeix, G.; Boccara, M. Transformation of *Botrytis Cinerea* with the Hygromycin B Resistance Gene, *Hph*. *Curr Genet* **1994**, *26* (3), 251–255. <https://doi.org/10.1007/BF00309556>.
- (245) Solovyev, V.; Kosarev, P.; Seledsov, I.; Vorobyev, D. Automatic Annotation of Eukaryotic Genes, Pseudogenes and Promoters. *Genome Biol* **2006**, *7 Suppl 1*, S10.1-12. <https://doi.org/10.1186/gb-2006-7-s1-s10>.
- (246) Qi, X.; Li, X.; Zhao, J.; He, N.; Li, Y.; Zhang, T.; Wang, S.; Yu, L.; Xie, Y. GKK1032C, a New Alkaloid Compound from the Endophytic Fungus *Penicillium* Sp. CCCC 400817 with Activity against Methicillin-Resistant *S. Aureus*. *J Antibiot* **2019**, *72* (4), 237–240. <https://doi.org/10.1038/s41429-019-0144-5>.
- (247) Genetic Manipulation of the *Fusarium Fujikuroi* Fusarin Gene Cluster Yields Insight into the Complex Regulation and Fusarin Biosynthetic Pathway. *Chemistry & Biology* **2013**, *20* (8), 1055–1066. <https://doi.org/10.1016/j.chembiol.2013.07.004>.
- (248) Sato, M.; Dander, J. E.; Sato, C.; Hung, Y.-S.; Gao, S.-S.; Tang, M.-C.; Hang, L.; Winter, J. M.; Garg, N. K.; Watanabe, K.; Tang, Y. Collaborative Biosynthesis of Maleimide- and Succinimide-Containing Natural Products by Fungal Polyketide Megasyntases. *J. Am. Chem. Soc.* **2017**, *139* (15), 5317–5320. <https://doi.org/10.1021/jacs.7b02432>.
- (249) Li, Y.; Weissman, K. J.; Müller, R. Myxochelin Biosynthesis: Direct Evidence for Two- and Four-Electron Reduction of a Carrier Protein-Bound Thioester. *Journal of the American Chemical Society* **2008**, *130* (24), 7554–7555. <https://doi.org/10.1021/ja8025278>.
- (250) Fujiyama, K.; Kato, N.; Re, S.; Kinugasa, K.; Watanabe, K.; Takita, R.; Nogawa, T.; Hino, T.; Osada, H.; Sugita, Y.; Takahashi, S.; Nagano, S. Molecular Basis for Two Stereoselective Diels–Alderases That Produce Decalin Skeletons\*\*. *Angewandte Chemie International Edition* **2021**, *60* (41), 22401–22410. <https://doi.org/10.1002/anie.202106186>.
- (251) Sato, M.; Kishimoto, S.; Yokoyama, M.; Jamieson, C. S.; Narita, K.; Maeda, N.; Hara, K.; Hashimoto, H.; Tsunematsu, Y.; Houk, K. N.; Tang, Y.; Watanabe, K. Catalytic Mechanism and Endo-to-Exo Selectivity Reversion of an Octalin-Forming Natural Diels–Alderase. *Nat Catal* **2021**. <https://doi.org/10.1038/s41929-021-00577-2>.
- (252) Chi, C.; Wang, Z.; Liu, T.; Zhang, Z.; Zhou, H.; Li, A.; Jin, H.; Jia, H.; Yin, F.; Yang, D.; Ma, M. Crystal Structures of Fsa2 and Phm7 Catalyzing [4 + 2] Cycloaddition Reactions with Reverse Stereoselectivities in Equisetin and Phomasetin Biosynthesis. *ACS Omega* **2021**, *10*.
- (253) Guillén Schlippe, Y. V.; Hedstrom, L. A Twisted Base? The Role of Arginine in Enzyme-Catalyzed Proton Abstractions. *Archives of Biochemistry and Biophysics* **2005**, *433* (1), 266–278. <https://doi.org/10.1016/j.abb.2004.09.018>.
- (254) Friedt, J.; Leavens, F. M. V.; Mercier, E.; Wieden, H.-J.; Kothe, U. An Arginine-Aspartate Network in the Active Site of Bacterial TruB Is Critical for Catalyzing Pseudouridine Formation. *Nucleic Acids Research* **2014**, *42* (6), 3857–3870. <https://doi.org/10.1093/nar/gkt1331>.
- (255) Hirokawa, T.; Boon-Chieng, S.; Mitaku, S. SOSUI: Classification and Secondary Structure Prediction System for Membrane Proteins. *Bioinformatics* **1998**, *14* (4), 378–379. <https://doi.org/10.1093/bioinformatics/14.4.378>.
- (256) Bernhofer, M.; Dallago, C.; Karl, T.; Satagopam, V.; Heinzinger, M.; Littmann, M.; Olenyi, T.; Qiu, J.; Schütze, K.; Yachdav, G.; Ashkenazy, H.; Ben-Tal, N.; Bromberg, Y.; Goldberg, T.; Kajan, L.;

O'Donoghue, S.; Sander, C.; Schafferhans, A.; Schlessinger, A.; Vriend, G.; Mirdita, M.; Gawron, P.; Gu, W.; Jarosz, Y.; Trefois, C.; Steinegger, M.; Schneider, R.; Rost, B. PredictProtein - Predicting Protein Structure and Function for 29 Years. *Nucleic Acids Res* **2021**, *49* (W1), W535–W540. <https://doi.org/10.1093/nar/gkab354>.

(257) Wang, S.; Li, W.; Liu, S.; Xu, J. RaptorX-Property: A Web Server for Protein Structure Property Prediction. *Nucleic Acids Res* **2016**, *44* (W1), W430–435. <https://doi.org/10.1093/nar/gkw306>.

(258) Attwood, M. M.; Krishnan, A.; Pivotti, V.; Yazdi, S.; Almén, M. S.; Schiöth, H. B. Topology Based Identification and Comprehensive Classification of Four-Transmembrane Helix Containing Proteins (4TMs) in the Human Genome. *BMC Genomics* **2016**, *17*, 268. <https://doi.org/10.1186/s12864-016-2592-7>.

(259) Shen, G. The Catalytic Mechanism of Vitamin K Epoxide Reduction in a Cellular Environment. *10*.

(260) Liu, S.; Li, S.; Shen, G.; Sukumar, N.; Krezel, A. M.; Li, W. Structural Basis of Antagonizing the Vitamin K Catalytic Cycle for Anticoagulation. *Science* **2021**, *371* (6524), eabc5667. <https://doi.org/10.1126/science.abc5667>.

(261) Inaba, K.; Ito, K. Structure and Mechanisms of the DsbB-DsbA Disulfide Bond Generation Machine. *Biochim Biophys Acta* **2008**, *1783* (4), 520–529. <https://doi.org/10.1016/j.bbamcr.2007.11.006>.

(262) Malojčić, G.; Owen, R. L.; Grimshaw, J. P. A.; Glockshuber, R. Preparation and Structure of the Charge-Transfer Intermediate of the Transmembrane Redox Catalyst DsbB. *FEBS Letters* **2008**, *582* (23–24), 3301–3307. <https://doi.org/10.1016/j.febslet.2008.07.063>.

(263) Guo, X.; Mayr, H. Quantification of the Ambident Electrophilicities of Halogen-Substituted Quinones. *J. Am. Chem. Soc.* **2014**, *136* (32), 11499–11512. <https://doi.org/10.1021/ja505613b>.

(264) Gasper, R.; Effenberger, I.; Kolesinski, P.; Terlecka, B.; Hofmann, E.; Schaller, A. Dirigent Protein Mode of Action Revealed by the Crystal Structure of AtDIR6. *Plant Physiol.* **2016**, *172* (4), 2165–2175. <https://doi.org/10.1104/pp.16.01281>.

(265) Kim, K.-W.; Smith, C. A.; Daily, M. D.; Cort, J. R.; Davin, L. B.; Lewis, N. G. Trimeric Structure of (+)-Pinoresinol-Forming Dirigent Protein at 1.95 Å Resolution with Three Isolated Active Sites. *J. Biol. Chem.* **2015**, *290* (3), 1308–1318. <https://doi.org/10.1074/jbc.M114.611780>.

# Annex

## Annex 01: Purification and in vitro assay of PrcH

### Heterologous expression of *prcH*

To verify the function of PrcH. We tried to express and purify A in *E-coli*, and determine the function of PrcH through the in vitro experiments after purification

Since no intron was identified in *prcH* by FGENESH, the gene coding  $\alpha/\beta$  hydrolase was obtained by PCR from the gDNA of *Sarocladium zeae* using the primers RhAZABhydSc67F(297) and RhAZABhydSc67R(298). The plasmid pET28a was treated with restriction enzymes EcoRI and Hind III to obtain the backbone of the expression plasmid. The two fragments were then transferred into the *E.coil* (top 10) to obtain the integral plasmid pET28a-*prcH*. The recombinant plasmid was purified and sequenced to ensure that its ligation was correct. pET28a-*prcH* was extracted and transferred to the *E. coli* (codon plus) for protein expression.

pET28a-*prcH* expressed in *E.coli* (codon plus) was cultured in LB with antibiotics at 37 °C overnight. Next, 1 mL cultivation was added to 100 mL LB in a 500 mL flask to continue the cultivation. When the OD up to 0.6, kept the flask at 0 °C for 10 min, followed by 200  $\mu$ M IPTG was added to the flask to induce the protein expression. Then, the flask was cultured at 19 °C, 200 rpm for 19 h, centrifugation at 4 °C, 5900 g for 10 min to collect the pellet.

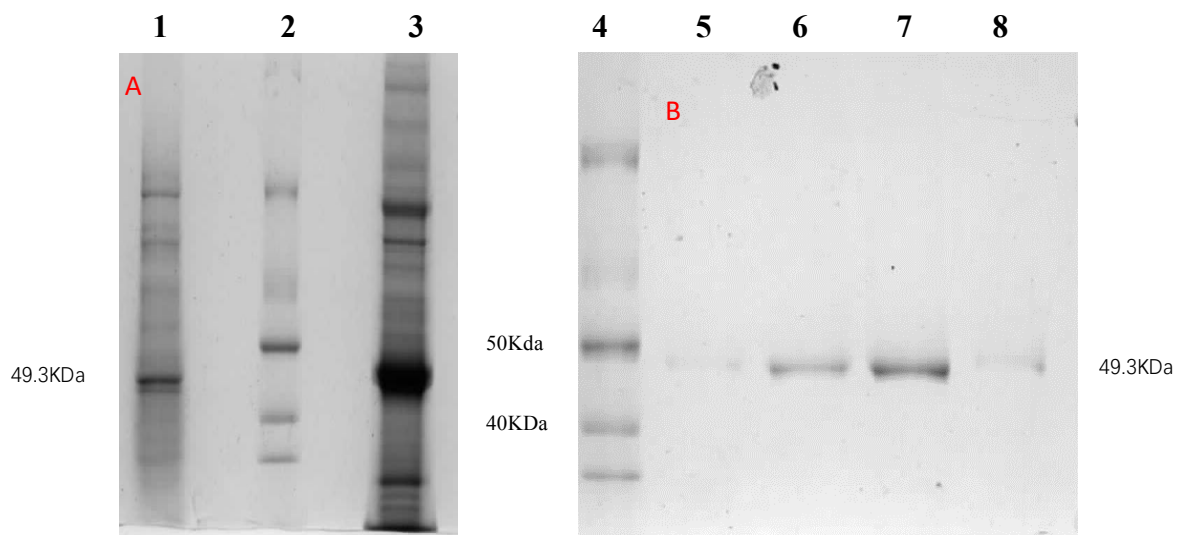
BugBuster toolkit was used for the rapid extraction of small amounts of protein. 1 mL BugBuster was added to the pellet (the pellet harvested from 5 mL fermentation liquid). After centrifugation, the upper layer was discarded. 1mg/ml lysozyme, 5% DNase I and 200  $\mu$ L bugbuster was added to the pellet, and the mixture was incubated at room temperature for 30 min. Then freezing it in liquid nitrogen followed by melting at room temperature. This operation was repeated three times. Centrifugation at 4°C, 14000 g for 15 min to get soluble fraction and inclusion body divided. SDS-PAGE was performed as before. The size of the *prcH* is 49.3KDa and soluble *prcH* was obtained see Annex 01-1-A.

### Protein purification

Affinity chromatography column(Ni) was used to purify the protein from the intracellular protein. Buffer A (NaH<sub>2</sub>PO<sub>4</sub> 20 mM, NaCl 500 mM, imidazole 10 mM, adjust pH to 7.6 ) and

buffer B (NaH<sub>2</sub>PO<sub>4</sub> 20 mM, NaCl 500 mM, imidazole 500 mM, adjust pH to 7.6) were prepared. The sample was mixed with buffer A and cell lysis was achieved by sonication. the supernatant was obtained by centrifugation at 20000 g, 4 °C for 15min. 5 mL beads were added to the column and the volume was concentrated to 3 mL by the pump. 30 mL buffer A was added to the column gently to wash the beads,

The sample was added to the upper layer of beads, followed by eluting with different concentrations of imidazole (10 mM, 50 mM, 200 mM, 400 mM). The nanoquant was used to identify the existence of protein depending on the change of color. The nanoquant was diluted five times and 100 µL nanoquant was mixed with 20 µL eluent to monitor protein elution. Finally, the target protein was isolated at 200mM imidazole. Four tubes of 1ml of protein were got (Annex: 01-1-B). The purified protein stored at -80 °C with 25% glycerol for long-term preservation.



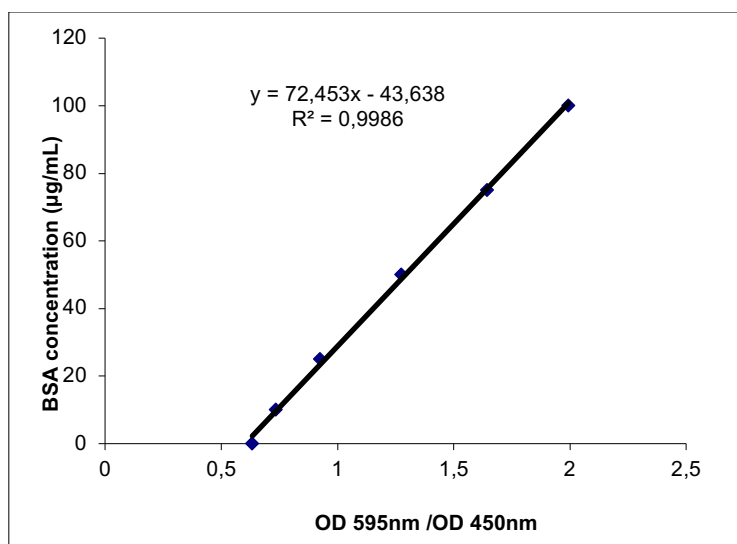
Annex 01-1: Sonicated proteins (A) and purified proteins (B) on SDS-PAGE gels stained with Coomassie blue. . lanes 2 and 4 are the markers, 1 is the supernatant. 3 is the precipitation.5-8 are the purified proteins.

### Desalting of protein

500 mM 10 × PEPPS-EPPS buffer was mixed with 10% glycerol, adjusting the pH to 8.5. The buffer was added to the column with protein inside to a final volume of 15 mL. Centrifugation at 4°C, 4000 g for 35 min to obtain the concentrate volume of 0.5 mL. 15 mL buffer was added to the column, centrifugation at 4 °C, 4000 g for 40 min to a concentrated volume 0.5 mL. Repeat this step several times until the final salt concentration of protein is less than 500 µM. The nanoquant was used to test whether the desalted protein was successfully obtained.

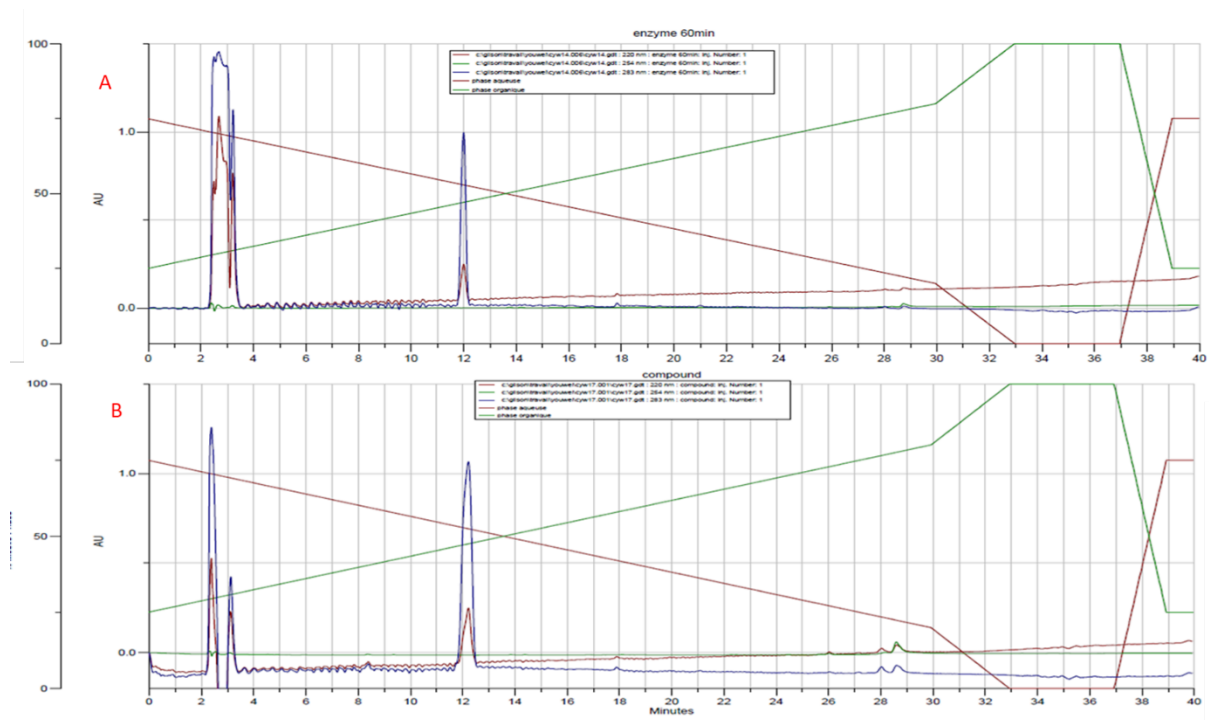
## Determination of protein content by the method of Bradford

To detect the concentration of the protein, 1 mg/mL bovine serum albumin was prepared and diluted to 100 µg/ml, 75 µg/ml, 50 µg/ml, 25 µg/ml, 10 µg/ml, 0 µg/ml separately. The nanoquant was diluted five times with water and participated in 96-well plates mixed with different concentrations of bovine serum albumin. The microplate reader was used to detect the protein concentration (OD 450 and OD 590). The standard curve for the determination of protein content was calculated to be  $y=72.453x-43.638$  (Annex 01-2), and then the content of the four tubes of protein were calculated to be 21, 46, 64 and 19 µg/ml respectively. Four tubes of protein were mixed and desalted to obtain 1mL desalted protein with a concentration of 110 µg/mL.



Annex 01-2: The standard curve for determination of protein content ( $y=72.453x-43.638$ )

First of all, we wanted to know if *prcH* can use linear alcohol as the substrate. In vitro assay was performed by using *prcH* and linear alcohol. Annex 01-2-A showed the alcohol compound has a 283 nm UV absorption peak at 12.2 min. After the alcohol compound reacted with *PrcH*, the substrate peak slightly reduced, but without a new peak formed. This result demonstrated the *prcH* can't use linear alcohol as the substrate. So we tried to use Dess martin reagent to convert the linear alcohol to the linear aldehyde to obtain the envisaged substrate but failed to convert.



Annex 01-3: The different HPLC chromatogram before (A) and after (B) *in vitro* reaction with alcohol compound as substrate.

## Annex 02 : Primers used in Chapter 2

Primer name	Sequence (5'→3')	Description
K7Sc67F11F (181)	GTAGCCAGATCTAATGGCGACAACAC	Construction of the inactivation cassette for $\Delta prcA+prcB$
RhSc67 Flank1F (127)	AAACAGCTATGACCATGATTACGAATTCGAGGTAGCCAGATCTAATGGCGACAACAC	Construction of plasmid for inactivation of $\Delta prcA+prcB$
RhSc67 Flank1R (128)	CCAAAAATGCTCCTTCAATATCATCTTCTGTCCATTGCTAGGACGTTGGCTTAG	Construction of plasmid for inactivation of $\Delta prcA+prcB$
K7Sc67F12R (182)	CACAACGCTGCACAGAATC	Construction of the inactivation cassette for $\Delta prcA+prcB$
RhSc67 Flank2R (130)	GCCAGTGCCAAGCTTGCATGCCTGCAGGTCGCACAACGCTGCACAGAATC	Construction of plasmid for inactivation of $\Delta prcA+prcB$
RhSc67 Flank2F (129)	GCCGACCGGATCCACTTAACGTTACTGAAATCGATACAAACTTGCATATATTGACTC	Construction of plasmid for inactivation of $\Delta prcA+prcB$
PbDIPK67R (266)	GTCCAATGGCAGCATAGCACTG	Validation of the $\Delta prcA+prcB$
PbDIPK67F12F (265)	GTTGGCGACCCAGTAGGGAC	Validation of the $\Delta prcA+prcB$
PbHph2F (183)	GTGTCACGTTGCAAGACCTG	Verification of target gene knockout( $\Delta prcA+prcB$ )
PbHph2R (184)	CGCTCGTCTGGCTAAGATC	Verification of target gene knockout( $\Delta prcA+prcB$ )
K7Sc67F11F (181)	GTAGCCAGATCTAATGGCGACAACAC	Construction of the inactivation cassette for $\Delta prcB$
RhSc67 Flank1F (127)	AAACAGCTATGACCATGATTACGAATTCGAGGTAGCCAGATCTAATGGCGACAACAC	Construction of plasmid for inactivation of $\Delta prcB$
RhSc67 Flank1R (128)	CCAAAAATGCTCCTTCAATATCATCTTCTGTCCATTGCTAGGACGTTGGCTTAG	Construction of plasmid for inactivation of $\Delta prcB$
K7DAF12BF (282)	CTGCACATCCGTTGTGCCATG	Construction of the inactivation cassette for $\Delta prcB$
RhSC67DAF2BR (281)	GCCAGTGCCAAGCTTGCATGCCTGCAGGTCGCTGCACATCCGTTGTGCCATG	Construction of plasmid for inactivation of $\Delta prcB$
RhSC67DAF12BF (280)	GCCGACCGGATCCACTTAACGTTACTGAAATCGATCTTGGGCTGGTAGCCAAG	Construction of plasmid for inactivation of $\Delta prcB$
PbSC67DA12F (315)	CTGCGGGTCGATGGCGTG	Validation of the $\Delta prcB$
PbDIPK67R (266)	GTCCAATGGCAGCATAGCACTG	Validation of the $\Delta prcB$
PbSC67DA11F (333)	GAGATCGCCCGAGCACACTC	Verification of target gene knockout( $\Delta prcB$ )
PbSC67DA11R (334)	GAGCGGCTGCACGAGGAATG	Verification of target gene knockout( $\Delta prcB$ )
K7Sc67ER2F (308)	GTGGGATGACCTGCGGCAC	Construction of the inactivation cassette for $\Delta prcI$

RhSc67ER2F11F (287)	AAACAGCTATGACCATGATTACGAATTCGAGGTGGGATGACCTGCGGCAC	Construction of plasmid for inactivation of <i>AprcI</i>
RhSc67ER2F11R (288)	CCAAAAATGCTCCTTCAATATCATCTTCTGTCCACTCTACAGCCAGCCAC	Construction of plasmid for inactivation of <i>AprcI</i>
K7Sc67ER2R (307)	CATGGGCCAGCGGCGAC	Construction of the inactivation cassette for <i>AprcI</i>
RhSc67ER2F12F (289)	GCCGACCGGATCCACTTAACGTTACTGAAATCGGTGGGGTGGTATGAGTGAAG	Construction of plasmid for inactivation of <i>AprcI</i>
RhSc67ER2F12R (290)	GCCAGTGCCAAGCTTGCATGCCTGCAGGTCGCATGGCCAGCGGCGAC	Construction of plasmid for inactivation of <i>AprcI</i>
PbMut67 ER2F (310)	CAGAAGGAGGATCTTTGACCTATAG	Validation of the <i>AprcI</i>
PbMut67 ER2R (309)	CTCCCTCAGGGCACCAC	Validation of the <i>AprcI</i>
PbSc67CDS29F (79)	GTCATCCACAATGAGAGGCTGGGTC	Verification of target gene knockout( <i>AprcI</i> )
PbSc67CDS29R (80)	CTTGCCACGTGCTCTCTTGCTC	Verification of target gene knockout( <i>AprcI</i> )
KOSc67PreDf11F (350)	CTGCTACGGTGGGAAGCTC	Construction of the inactivation cassette for <i>AprcD</i>
RhSc67PreDf11F (303)	AAACAGCTATGACCATGATTACGAATTCGAGCTGCTACGGTGGGCAAGCTC	Construction of plasmid for inactivation of <i>AprcD</i>
RhSc67PreDf11R (304)	CCAAAAATGCTCCTTCAATATCATCTTCTGTCTGCTGAAGTTCTGAAGTTAGTAG	Construction of plasmid for inactivation of <i>AprcD</i>
K7Sc67F12R (182)	CACAACGCTGCACAGAATC	Construction of the inactivation cassette for <i>AprcD</i>
RhSc67PreDf12F (305)	GCCGACCGGATCCACTTAACGTTACTGAAATCGTGATGGAGAACCAGATGACTGAG	Construction of plasmid for inactivation of <i>AprcD</i>
RhSc67PreDf12R (306)	GCCAGTGCCAAGCTTGCATGCCTGCAGGTCGCACAACGCTGCACAGAATC	Construction of plasmid for inactivation of <i>AprcD</i>
PbDIPK67F12F (265)	GCATTACACAGTAACGTCATACCCAATTG	Validation of the <i>AprcD</i>
PbSc67PreDf11F (351)	CATCATCCACAACGGGGCTG	Validation of the <i>AprcD</i>
PBsc67PreDf1R (352)	GTCCCTCATCGATTGCCAC	Verification of target gene knockout( <i>AprcD</i> )
PbSc67PreEfl1F (319)	GTACCCCATGTGCGCGAGAC	Verification of target gene knockout( <i>AprcD</i> )
KOSc67PreEfl1F (318)	GACGACGCCGAGAGCGTAG	Construction of the inactivation cassette for <i>AprcE</i>
RhSc67PreEfl1F (299)	AAACAGCTATGACCATGATTACGAATTCGAGGACGACGCCGAGAGCGTAG	Construction of plasmid for inactivation of <i>AprcE</i>
RhSc67PreEfl1R (300)	CCAAAAATGCTCCTTCAATATCATCTTCTGTGCGGAGTGCGTGTGTTGTTG	Construction of plasmid for inactivation of <i>AprcE</i>
KOSc67PreEfl2R (317)	CTCGGATTGCGACGCGGTATC	Construction of the inactivation cassette for <i>AprcE</i>
RhSc67PreEfl2F (301)	GCCGACCGGATCCACTTAACGTTACTGAAATCGGTGAGGTGAGGTGAGGCAAAG	Construction of plasmid for inactivation of <i>AprcE</i>
RhSc67PreEfl2R (302)	GCCAGTGCCAAGCTTGCATGCCTGCAGGTCGCTCGGATTGCGACGCGGTATC	Construction of plasmid for inactivation of <i>AprcE</i>
PbSc67PreEfl2R (316)	CAGGACGCTTTCGCGATG	Validation of the <i>AprcE</i>
PbSc67PreEfl1F (319)	GTACCCCATGTGCGCGAGAC	Validation of the <i>AprcE</i>
K7Sc67F12R (182)	CACAACGCTGCACAGAATC	Verification of target gene knockout( <i>AprcE</i> )
PbSc67PreEfl1F (335)	GAGTTCCGCCACGCTCTG	Verification of target gene knockout( <i>AprcE</i> )
K7Sc67hyd1F (313)	GTCGTCCCGTGGCGAGAAC	Construction of the inactivation cassette for <i>AprcH</i>
RhSc67BhF11F (291)	AAACAGCTATGACCATGATTACGAATTCGAGGTCGTCGCCGTGGCGAGAAC	Construction of plasmid for inactivation of <i>AprcH</i>
RhSc67BhF11R (292)	CCAAAAATGCTCCTTCAATATCATCTTCTGTGATGGCCTTGGTTGGTGGCTC	Construction of plasmid for inactivation of <i>AprcH</i>
K7Sc67hyd1R (312)	CACACCCCTGCTTGAAGGTG	Construction of the inactivation cassette for <i>AprcH</i>
RhSc67BhF12F (293)	GCCGACCGGATCCACTTAACGTTACTGAAATCTTTTGGATATGATAGAGACTAGACCCGCTTG	Construction of plasmid for inactivation of <i>AprcH</i>
RhSc67BhF12R (294)	GCCAGTGCCAAGCTTGCATGCCTGCAGGTCGCACACCCCTGCTCCTGAAGGTG	Construction of plasmid for inactivation of <i>AprcH</i>
PbMut67hyd1R (311)	GTGCTCGGGCTCTCTCGTAAC	Validation of the <i>AprcH</i>
PbMut67hyd1F (314)	GTCCGCGGGCTAAGTCCGTC	Validation of the <i>AprcH</i>
PbSc67AB1R (320)	CAACGCCACGACCAATC	Verification of target gene knockout( <i>AprcH</i> )
PbSc67AB1F (321)	GACTTCTCGGACGCGATC	Verification of target gene knockout( <i>AprcH</i> )
K7Sc67preXF11F (365)	GACTGTTGCTAAACCGGAGAG	Construction of the inactivation cassette for <i>AprcX</i>
RhSc67preXF11F (361)	AAACAGCTATGACCATGATTACGAATTCGAGGACTGTTGCTAAACCGGAGAG	Construction of plasmid for inactivation of <i>AprcX</i>
RhSc67preXF11R (362)	CCAAAAATGCTCCTTCAATATCATCTTCTGTGCTGCCATCGTACAGGAGATC	Construction of plasmid for inactivation of <i>AprcX</i>
K7Sc67preXF12R (366)	CACCAGGCACTCCGTCGTC	Construction of the inactivation cassette for <i>AprcX</i>
RhSc67preXF12F (363)	GCCGACCGGATCCACTTAACGTTACTGAAATCGACGCTGTTGTCGAGGATGAAG	Construction of plasmid for inactivation of <i>AprcX</i>
RhSc67preXF12R (364)	GCCAGTGCCAAGCTTGCATGCCTGCAGGTCGCACAGGCACTCCGTCGTC	Construction of plasmid for inactivation of <i>AprcX</i>

PbMutPreXF (367)	CTGGCTTCTGCCTGCCCATC	Validation of the $\Delta$ <i>prcX</i>
PbMutPreXR (368)	CATGGTCTCGGCCGTTGACAAGATC	Validation of the $\Delta$ <i>prcX</i>
PbPreXIntR (404)	CTGATCAAGCTCCTGGCTGTG	Verification of target gene knockout( $\Delta$ <i>prcX</i> )
PbPreXIntF (405)	GCGACGCCCGAAATGTC	Verification of target gene knockout( $\Delta$ <i>prcX</i> )
KOSc67PreEfl1F (318)	GACGACGCCGAGAGCGTAG	Construction of the complementation cassette for <i>prcE</i>
RhSc67PreEfl2R (302)	GCCAGTGCCAAGCTTGCATGCCTGCAGGTCGCTCGGATTGCGACGCGGTATC	Construction of plasmid for restore of <i>prcE</i>
PconG418-FL2E (509)	GCCGACCGGGATCCACTTAACGTTACTGAAATCGGTGAGGTGAGGTGAGGCAAAGGCACGAC	Construction of plasmid for restore of <i>prcE</i>
KOSc67PreEfl2R (317)	CTCGGATTGCGACGCGGTATC	Construction of the complementation cassette for <i>prcE</i>
RhSc67PreEfl1F (299)	AAACAGCTATGACCATGATTACGAATTCGAGGACGACGCCGAGAGCGTAG	Construction of plasmid for restore of <i>prcE</i>
PconG418-PRCE (510)	CCAAAAAGTGCTCCTTAATATCATCTTCTGTCTCACCTCACCAAGTATCGCGGCCCAAC	Construction of plasmid for restore of <i>prcE</i>
PbSc67PreEfl2R (316)	CAGGCAGCCTTTCGCGATG	Validation of complementation of <i>prcE</i>
PbSc67PreEfl1F (319)	GTACCCATGTGCGCGAGAC	Validation of the complementation of <i>prcE</i>
K7Sc67F12R (182)	CACAACGCTGCACAGAATC	Validation of the complementation of <i>prcE</i>
PbSc67PreEfl1F (335)	GAGTTCCGCCACGTCCTG	Verification of target gene complementation( <i>prcE</i> )
RhpUC18R (74)	CTCGAATTCGTAATCATGGTCATAGCTG	Verification of target gene complementation( <i>prcE</i> )
RH-hphF	GACAGAAGATGATATTGAAGGAGC	For screening of mutants
RH-hphR	GATTTCAGTAACGTTAAGTGGAT	For screening of mutants
RhHphSpl1R	CTTCTGCGGGCGATTGTG	Construction of the inactivation cassette
RhHphSpl2F	CGAGAAGTTTCTGATCGAAAAG	Construction of the inactivation cassette
RhpUC18F	CGACCTGCAGGCATGCAAG	Construction of plasmid for inactivation
RhpUC18R	CTCGAATTCGTAATCATGGTCATAGCTG	Construction of plasmid for inactivation
PbSc67->Sc152	GACATACCACGACACAGCCCTG	Identification of the correlation between <i>ssc67-sc152</i>
PbSc152->Sc67	GCCCGACTTAACTGTCTCAC	Identification of the correlation between <i>ssc67-sc152</i>

### Annex 03 : Primers used in Chapter 3

Primer name	Sequence (5'→3')	Description
PbSc67I2F (56)	CACGCGTCTCGCCTTCAT	Determination of the location of the intron
PbSc67I2R (57)	CCTGCTGTGACGTTCACTTGAAG	Determination of the location of the intron
Pid intron PN67-F1(378)	GAAGGAGTCGCCGCTGTG	Determination of the location of the intron
Pid intron PN67 R1(379)	GTGGGCTCTGTTGCGATC	Determination of the location of the intron
Pid intron PN67 R2(380)	GACAACCTGGGGCTGGTCAAC	Determination of the location of the intron
Pid intron PN67 F2(381)	GTTGGCCACGCTTTGCTTG	Determination of the location of the intron
Pid intron PN67 R3(382)	GTCGCCATCCGAATATCGATG	Determination of the location of the intron
Pid intron PN67 F3(383)	GTCCCGGAAGCCAGCAC	Determination of the location of the intron
Pcopy ER1 GG R1(384)	TTCGTCTCATCGGGTCTCATTTCTAGTCCCCAGCACTGCCTC	Determination of the location of the intron
Pcopy ER1 GG F1(385)	AACGCTCTGGTCGGTCTCAGGATGACTACTCCCTTGCTGGAGTTGTTG	Determination of the location of the intron
DI20(166)	TTTTTTTTTTTTTTTTTTT	Random primer for revers transcription PCR
Pid P-C-T F (388)	GTAAGTGCAGTGCAGTAC	Identification of the right construction of plasmids
Pid P-C-T R (389)	CCTGAATTCGCATCTAGA	Identification of the right construction of plasmids
Pcon pla-pksnrps67F (390)	ATCCTAACTCGAGGCGAATTC	construction of plasmids pGG-P67
Pcon pla-pksnrps67R (391)	GAACCGTGATGATGATGATGATGAC	construction of plasmids pGG-P67
Phom pks-nrps67R (393)	GTCATCATCATCATCATCACGGTTCTAAGGCCGAGAGACACCGAAAG	construction of plasmids pGG-P67
Poverlap PKSNRPS67F1(394)	CAGGACCACAGCGGCGAC	construction of plasmids pGG-P67
Poverlap pksnrps67R2(395)	CAACCTGAAGATCGCAACAGAG	construction of plasmids pGG-P67
Poverlap pnsnrps67F2(396)	GTTTTCCGTTGACCAGCCCCAG	construction of plasmids pGG-P67
Poverlap pksnrps67R2(397)	CAAGAATCTCAAGCAAACGCTG	construction of plasmids pGG-P67



Poverlap pksnrps67R3(398)	CACCTTGCATCGATATTCGGATG	construction of plasmids pGG-P67
Poverlap pksnrps67F3(399)	GTCTCAAGTCTGGCTTC	construction of plasmids pGG-P67
Poverlap pksnrps67F4(400)	CTGAAGCATGAAGGCGAGAC	construction of plasmids pGG-P67
Poverlap pksnrps67R4(401)	CTTGCTCAAGTGAACGTCACAG	construction of plasmids pGG-P67
Phom pks-nrps67f(392)	GAAATTGCGCTCGAGTTAGGATCATGCTTTTTTTCCTGGTCCGTC	construction of plasmids pGG-P67

## Annex 04 : Primers used in Chapter 4

Primer name	Sequence (5'→3')	Description
Conpuc-hphF (448)	AAACAGCTATGACCATGATTACGAATTCGAGGACAGAAGATGATATTGAAGGAGCATTTTTGG C	Construction of pUC18-gg/ag
RH-hphR (198)	GATTCAGTAACGTTAAGTGGAT	Construction of pUC18-gg /ga,/av/gv/am/g
Conhph-gfpF (449)	GCCGACCGGATCCACTTAACGTTACTGAAATCTCATTGTACAGTTCATCCATACCATG	Construction of pUC18-gfp-gpdA/amyB
Congfp-gpdAR (450)	CACATTGACCAATCTTAGATTATACATCATGCGTAAAGGCGAAGAGCTGTTCACTG	Construction of pUC18-gg
Rh-gpdAf (451)	GATGTATAATCTAAGATTGGTCAATGTG	Construction of pUC18-gg
CongpdA-pucR (452)	GCCAGTGCCAAGCTTGATGCTGCTGAGGTCGGTGGCTAAGAAGGATACCTGACCTATG	Construction of pUC18-gg
ConamyB-gfpR (455)	CTTACTCTTACAACGACCGTCTCGTGATGCGTAAAGGCGAAGAGCTGTTCACTG	Construction of pUC18-ag
Rh-amyBF (456)	CACGAAGCACGGTCGTTGTAAG	Construction of pUC18-ag
ConamyB-pucR (457)	GCCAGTGCCAAGCTTGATGCTGCTGAGGTCGGAATCAACATTGAGGGCGATACATAGTATTTG	Construction of pUC18-ag
ConTrpc-pucR (484)	GCCAGTGCCAAGCTTGATGCTGCTGAGGTCGGACAGCAAGATGATATTGAAGGAGC	Construction of pUC18-tg
PTrpcF (485)	ATCGATGCTTGGGTAGAAATAGGTAAGTC	Construction of pUC18-tg
Congfp-TrpcR (486)	GACTTACCTATTCTACCCAAGCATCGATATGCGTAAAGGCGAAGAGCTGTTCACTG	Construction of pUC18-tg
RhpUC18F (73)	CGACCTGCAGGCATGCAAG	Construction of pUC18-tg/tv/ta
ConamyB-mR ubyR (476)	CTTACTCTTACAACGACCGTCTCGTGGTGTCCAAAGGAGAGGAGTTAATCAAG	Construction of pUC18-am
Conhph-mRub yF (477)	CCGACCGGATCCACTTAACGTTACTGAAATCTTATACAATTCATCCATACCACCGCCTAAG	Construction of pUC18-am/tm/gm
Rh-amyBF (456)	CACGAAGCACGGTCGTTGTAAG	Construction of pUC18-am
Rh-gpdAf (451)	GATGTATAATCTAAGATTGGTCAATGTG	Construction of pUC18-gm
ConmRuby-gpdAR (478)	CACATTGACCAATCTTAGATTATACATCGTGTCCAAAGGAGAGGAGTTAATCAAG	Construction of pUC18-gm
ConamyB-venusR (480)	GTTACTCTTACAACGACCGTCTCGTGTCTAAAGGTGAAGAATTACTACTGGTG	Construction of pUC18-av
Conhph-venusF (481)	GCCGACCGGATCCACTTAACGTTACTGAAATCTTGTACAATTCATCCATACCATGGGTAATAC	Construction of pUC18-av/gv/tv
Congfp-venusR (482)	ACATTGACCAATCTTAGATTATACATCTCTAAAGGTGAAGAATTACTACTGGTG	Construction of pUC18-gv
ConamyB-venusR (479)	GACTTACCTATTCTACCCAAGCATCGATCTCTAAAGGTGAAGAATTACTACTGGTG	Construction of pUC18-av
ConTrpc-mRubyR (483)	GACTTACCTATTCTACCCAAGCATCGATGTGTCCAAAGGAGAGGAGTTAATCAAG	Construction of pUC18-cv
M13F (167)	CCCAGTCACGACGTTGTAAACG	Construction of the ectopic insertion cassette
M13R (168)	AGCGGATAACAATTCACACAGG	Construction of the ectopic insertion cassette

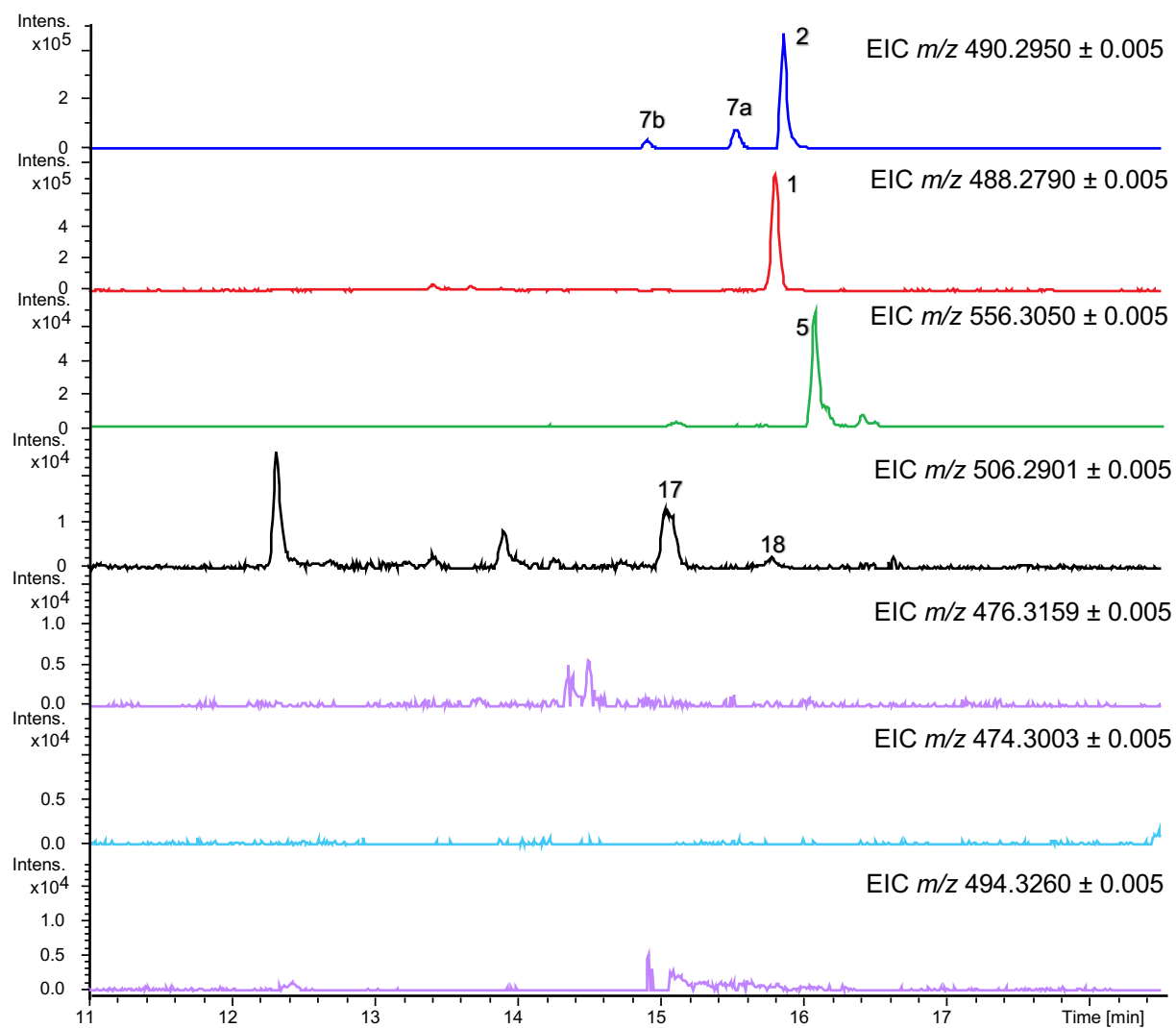
RH-hphF (197)	GACAGAAGATGATATTGAAGGAGC	Validation of the ectopic insertion of the mutants
PendAmyB (461)	GAATCAACATTGAGGGCGATACATAGTATTTG	Validation of the ectopic insertion of the mutants
PendGpaA (463)	GTGGCTAAGAAGGATACCTGACCTATG	Validation of the ectopic insertion of the mutants
Rh-gpdAf (451)	GATGTATAATCTAAGATTGGTCAATGTG	Construction of the pUC18-tfg
RH-hphR (198)	GATTCAGTAACTAAGTGGAT	Construction of the pUC18-tfg
Rhhph-prfh (507)	GCCGACCGGATCCACTTAACGTTACTGAAATCTTAGCGACCATCTAGATGCATTGGAC	Construction of the pUC18-tfg
RHPRFHGPDAR (508)	CACATTGACCAATCTTAGATTATACATCATGACCGAGCCTCCAGTCCCGATG	Construction of the pUC18-tfg
Rhsc47t1f (497)	AAACAGCTATGACCATGATTACGAATTCGAGGATATCTTGAAGCCTGCTCGTTGTG	Construction of the pUC18-gf/t1f/t2f
Rhsc47TFr (498)	CCAAAAATGCTCCTCAATATCATCTTCTGTCGGCTGCTGGTGGTGGCCACTTGTC	Construction of the pUC18-gf/t1f/t2f
Rhsc47t2f (499)	GCCGACCGGATCCACTTAACGTTACTGAAATCGTGGCTAAGAAGGATACCTGACCTAT	Construction of the pUC18-gf
Rh-gpdAf (451)	GATGTATAATCTAAGATTGGTCAATGTG	Construction of the pUC18-gf
RHsc47TF3F (500)	CACATTGACCAATCTTAGATTATACATCATGACCGAGCCTCCAGTCCCGATGGC	Construction of the pUC18-gf
Rhsc47t2f (501)	GCCAGTGCCAAGCTTGCATGCTGCAGGTCGCTTCTCCGAAGACTGCGATATTCAAG	Construction of the pUC18-gf
Rhsc47hphf1 (502)	GCCGACCGGATCCACTTAACGTTACTGAAATCCAATCTCACGACATGCATTGCTAAAG	Construction of the pUC18-tf1f
Rhsc47t2r (503)	CATCGGGACTGGAGGCTCGGTGATGAAGAGAGCGTTCGAACAAGCCCTTC	Construction of the pUC18-tf1f
Rhsc47hphf2f (504)	GCCGACCGGATCCACTTAACGTTACTGAAATCTGACGGCTGTTGTCGAGGATGAAGTG	Construction of the pUC18-t2f
Rhsc47t2f2r (505)	CATCGGGACTGGAGGCTCGGTGATGATGATGTGACTTGAGGGATGGAG	Construction of the pUC18-t2f
Pco ex pro tfr (496)	CTTCTCCGAAGACTGCGATATTC	Construction of the pUC18-tf1/t2f
pCopy back (506)	ATGACCGAGCCTCCAGTCCCGATG	Construction of the pUC18-tf1/t2f
pid ex Pro TF (493)	CATGAATGCGACTCTGGGAACATG	Validation of the insert of the target promoter
Pid ex pro TFr (494)	GAGAAGCGCTGATCGTACAGAC	Validation of the insert of the target promoter
RhSc21ERf1 F (426)	AAACAGCTATGACCATGATTACGAATTCGAGGAACAAAGGGTGCAAGCCGTTG	Construction of plasmid for inactivation of <i>ApraG</i>
RhSc21ERf1 R (427)	GCTCCTCAATATCATCTTCTGTCTGTTGGTATTCTTAGATACGGAGATG	Construction of plasmid for inactivation of <i>ApraG</i>
RhSc21ERf2 F (428)	ATCCACTTAACGTTACTGAAATCTTTCATAAAAACATTGGCCCTGATAG	Construction of plasmid for inactivation of <i>ApraG</i>
RhSc21ERf2 R (429)	GCCAGTGCCAAGCTTGCATGCCTGCAGGTCGGATTGCATTCAGCATCCGGAGTG	Construction of plasmid for inactivation of <i>ApraG</i>
KoSC21ERf1 (420)	GAACAAAGGGTGCAAGCCGTTG	Construction of the inactivation cassette for <i>ApraG</i>
KoSC21ERf2 (421)	GATTGCATTCAGCATCCGGAGTG	Construction of the inactivation cassette for <i>ApraG</i>
Pb-Mu-SC21ERf (422)	GATGTTCACTACACGATGGTAC	Validation of the <i>ApraG</i>
Pb-Mu-SC21ERr (423)	GTGCATGGCGGCTGTAAC	Validation of the <i>ApraG</i>
Pid-SC21ERf (424)	CATCTAGACCCTCCAACGCAC	Verification of target gene knockout( <i>ApraG</i> )
Pid-SC21ERr (425)	GTATTGGTCAACGGCGGTG	Verification of target gene knockout( <i>ApraG</i> )

**Annex 05: Rotatory powers of purified compounds.**

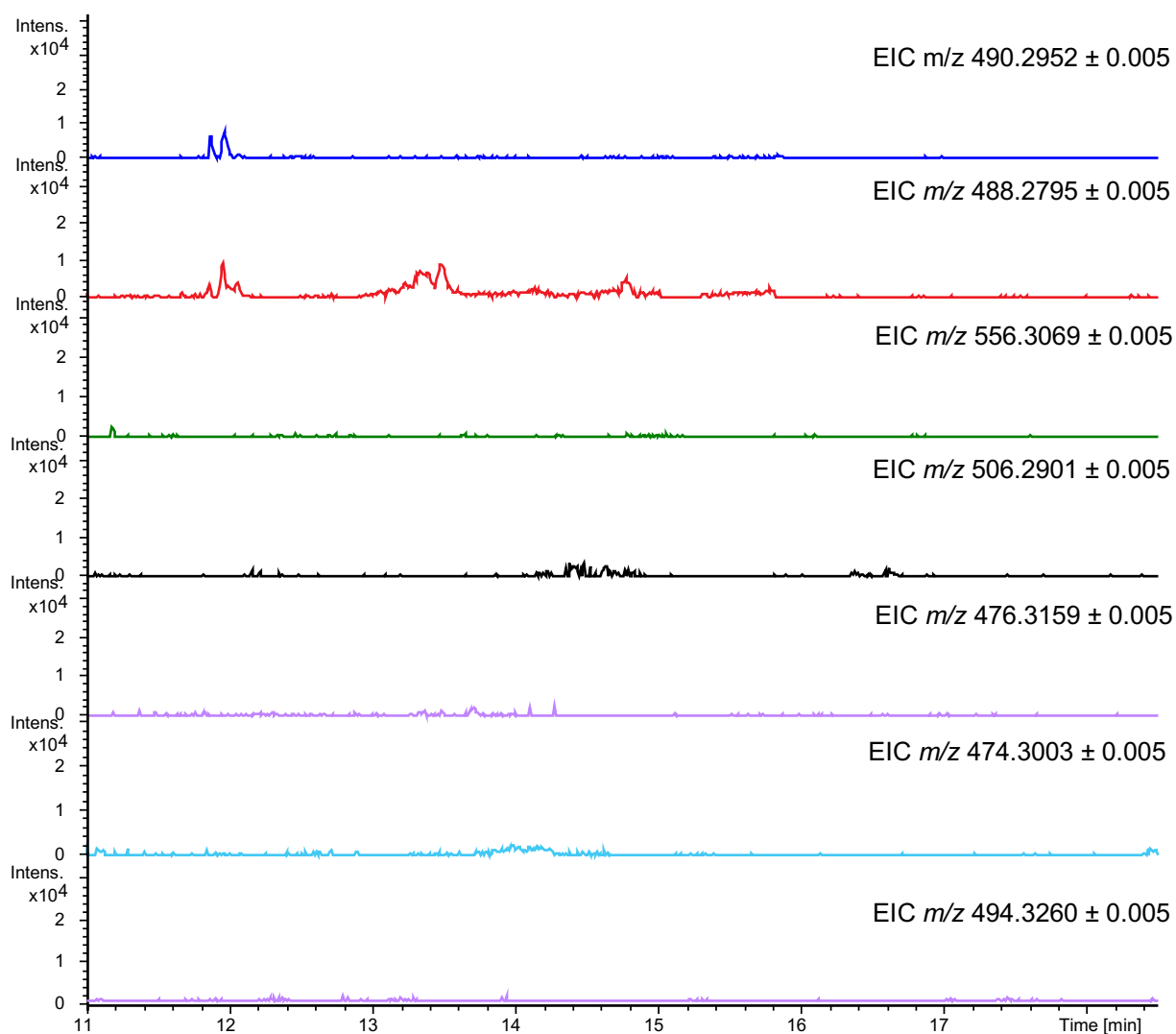
		Concentration (g/100mL)	$[\alpha]_D^{20}$	Solvent
<b>1</b>	Pyrrocidine A	1	+175.5	CHCl <sub>3</sub>
<b>2</b>	Pyrrocidine B	1	+83.6	CHCl <sub>3</sub>
<b>7a</b>	keto-Pyrrocidine D	1	+75.9	CHCl <sub>3</sub>
<b>8</b>	Trichobamide B	1	-120.9	CHCl <sub>3</sub>
<b>9</b>	Pyrrocidine C	0.7	+36.14	CHCl <sub>3</sub>
<b>10</b>	Pyrrocidine E	0.5	+77,2	CHCl <sub>3</sub>
<b>11</b>	Pyrrocidine F	0.3	-1,67	MeOH
<b>12</b>	Sarocladine A	0.55	+10.7	CHCl <sub>3</sub>
<b>14</b>		0.3	-5.33	CHCl <sub>3</sub>
<b>17</b>		0.4	-11.25	CHCl <sub>3</sub>
<b>18</b>	Pyrrocaline B	0.8	+21.75	CHCl <sub>3</sub>
<b>20</b>	Pyrrochalsin A	1	-4.4	CHCl <sub>3</sub>
<b>21</b>	Pyrrocycline A	0.3	-79.33	CHCl <sub>3</sub>
<b>22</b>	Pyrrocycline B	0.3	-37.7	CHCl <sub>3</sub>

## Annes 06: HRMS for isolated compounds

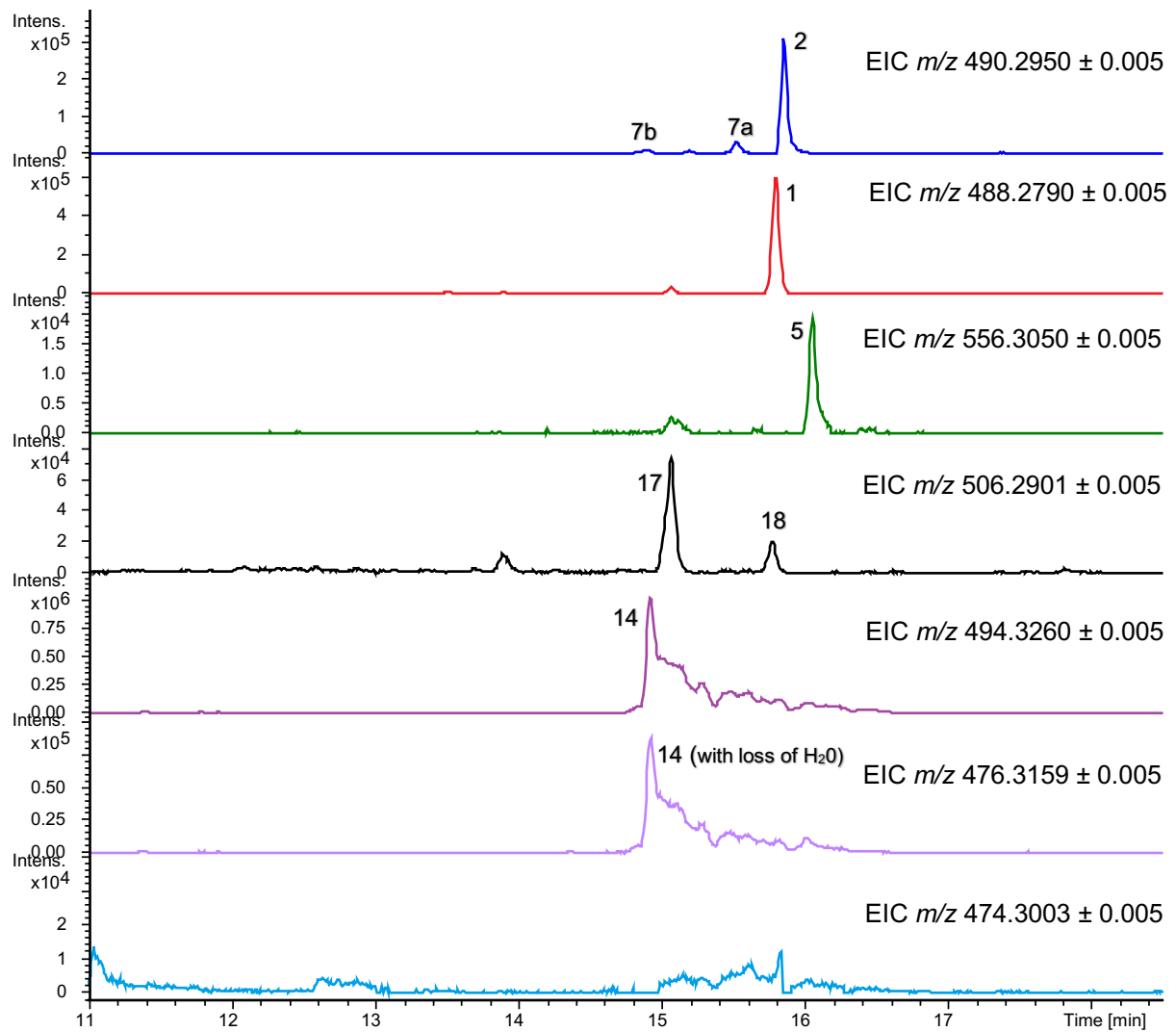
			[M + H] <sup>+</sup>		
			Found	Calculated	formula
<b>1</b>	Pyrrocidine A	C <sub>31</sub> H <sub>37</sub> NO <sub>4</sub>	488,2797	488,2795	C <sub>31</sub> H <sub>38</sub> NO <sub>4</sub> <sup>+</sup>
<b>2</b>	Pyrrocidine B	C <sub>31</sub> H <sub>39</sub> NO <sub>4</sub>	490,2945	490,2952	C <sub>31</sub> H <sub>40</sub> NO <sub>4</sub> <sup>+</sup>
<b>7a</b>	keto-Pyrrocidine D	C <sub>31</sub> H <sub>39</sub> NO <sub>4</sub>	490,2949	490,2952	C <sub>31</sub> H <sub>40</sub> NO <sub>4</sub> <sup>+</sup>
<b>7b</b>	enol-Pyrrocidine D	C <sub>31</sub> H <sub>39</sub> NO <sub>4</sub>	490,2947	490,2952	C <sub>31</sub> H <sub>40</sub> NO <sub>4</sub> <sup>+</sup>
<b>8</b>	Trichobamide B	C <sub>35</sub> H <sub>41</sub> NO <sub>5</sub>	556,3061	556,3057	C <sub>35</sub> H <sub>42</sub> NO <sub>5</sub> <sup>+</sup>
<b>9</b>	Pyrrocidine C	C <sub>31</sub> H <sub>37</sub> NO <sub>5</sub>	504,2745	504,2744	C <sub>31</sub> H <sub>38</sub> NO <sub>5</sub> <sup>+</sup>
<b>10</b>	Pyrrocidine E	C <sub>31</sub> H <sub>37</sub> NO <sub>5</sub>	504,2736	504,2744	C <sub>31</sub> H <sub>38</sub> NO <sub>5</sub> <sup>+</sup>
<b>11</b>	Pyrrocidine F	C <sub>31</sub> H <sub>37</sub> NO <sub>4</sub>	488,2827	488,2795	C <sub>31</sub> H <sub>38</sub> NO <sub>4</sub> <sup>+</sup>
<b>12</b>	Sarocladine	C <sub>32</sub> H <sub>39</sub> NO <sub>5</sub>	518,2884	518,2901	C <sub>32</sub> H <sub>40</sub> NO <sub>5</sub> <sup>+</sup>
<b>14</b>		C <sub>31</sub> H <sub>43</sub> NO <sub>4</sub>	494,3297	494,3265	C <sub>31</sub> H <sub>44</sub> NO <sub>4</sub> <sup>+</sup>
<b>17</b>		C <sub>31</sub> H <sub>39</sub> NO <sub>5</sub>	506,2890	506,2901	C <sub>31</sub> H <sub>40</sub> NO <sub>5</sub> <sup>+</sup>
<b>18</b>	Pyrrocaline A	C <sub>31</sub> H <sub>39</sub> NO <sub>5</sub>	506,2899	506,2901	C <sub>31</sub> H <sub>40</sub> NO <sub>5</sub>
<b>20</b>	Pyrrochalsin	C <sub>31</sub> H <sub>39</sub> NO <sub>3</sub>	474,2998	474,3003	C <sub>31</sub> H <sub>40</sub> NO <sub>3</sub> <sup>+</sup>
<b>21</b>	Pyrrocicline A	C <sub>31</sub> H <sub>39</sub> NO <sub>4</sub>	490,2929	490,2952	C <sub>31</sub> H <sub>40</sub> NO <sub>4</sub> <sup>+</sup>
<b>22</b>	Pyrrocicline B	C <sub>31</sub> H <sub>39</sub> NO <sub>4</sub>	490,2933	490,2952	C <sub>31</sub> H <sub>40</sub> NO <sub>4</sub> <sup>+</sup>
<b>23</b>		C <sub>31</sub> H <sub>41</sub> NO <sub>3</sub>	476,3146	476,3159	C <sub>31</sub> H <sub>42</sub> NO <sub>3</sub> <sup>+</sup>



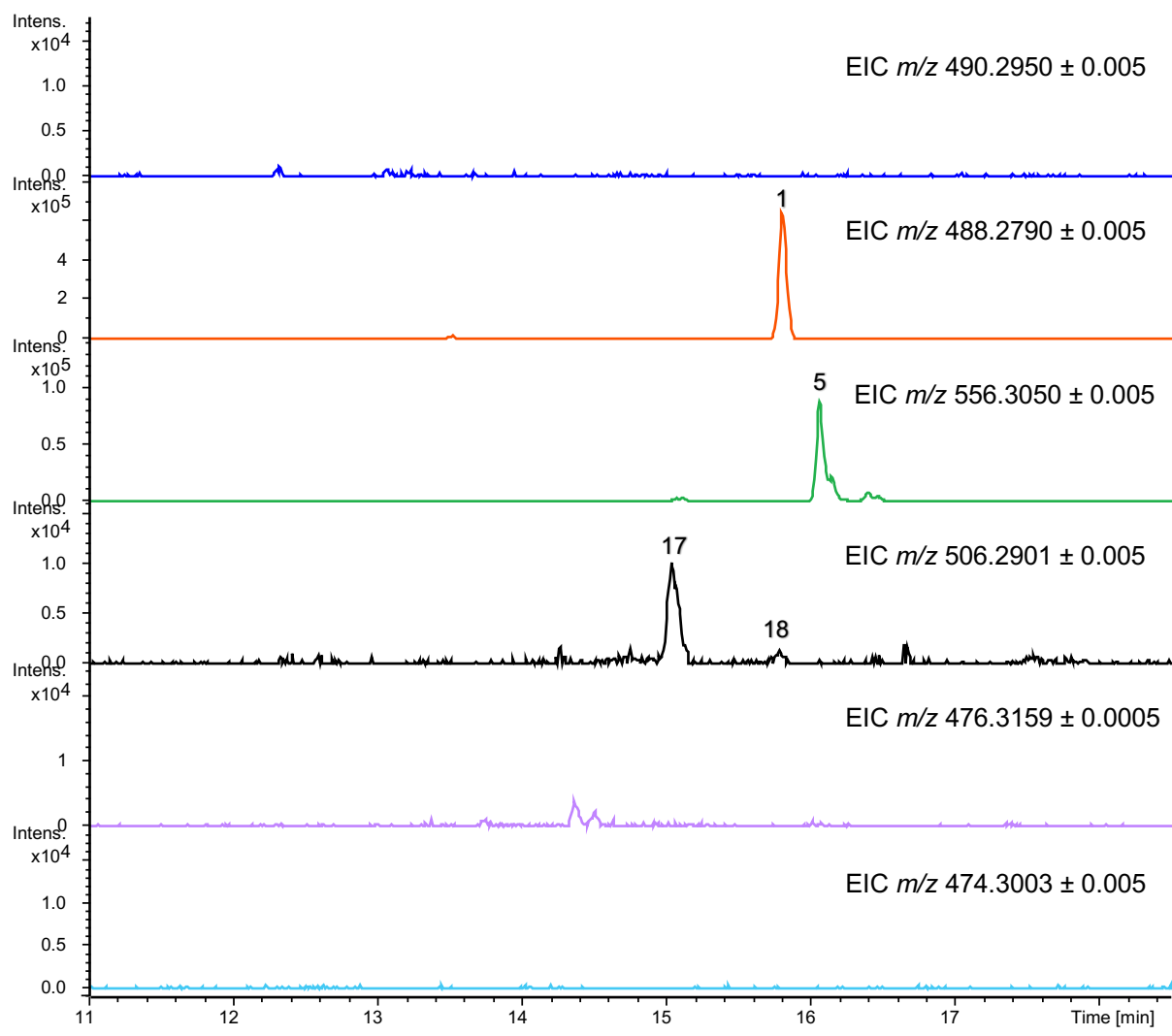
**Annex 07: Extract ion chromatograms obtained from LC-MS analysis of the metabolite extract from wild type *S. zeae* cultured on PDA.**



**Annex 08: Extract ion chromatograms obtained from LC-MS analysis of the metabolite extract from  $\Delta prcA+prcB$  *S. zeae* mutant cultured on PDA.**

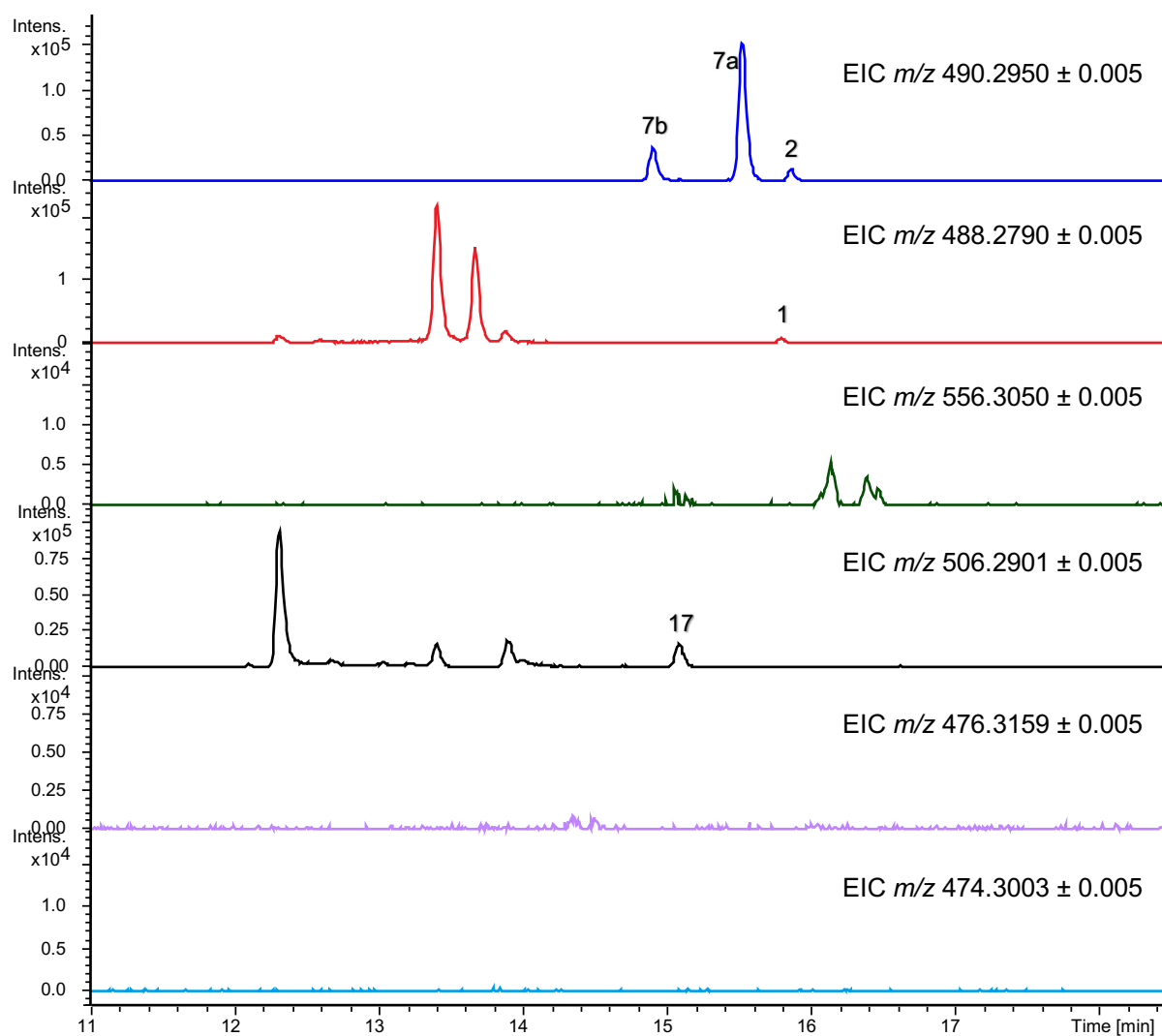


**Annex 09: Extract ion chromatograms obtained from LC-MS analysis of the metabolite extract from  $\Delta prcH$  *S. zeae* mutant cultured on PDA.**

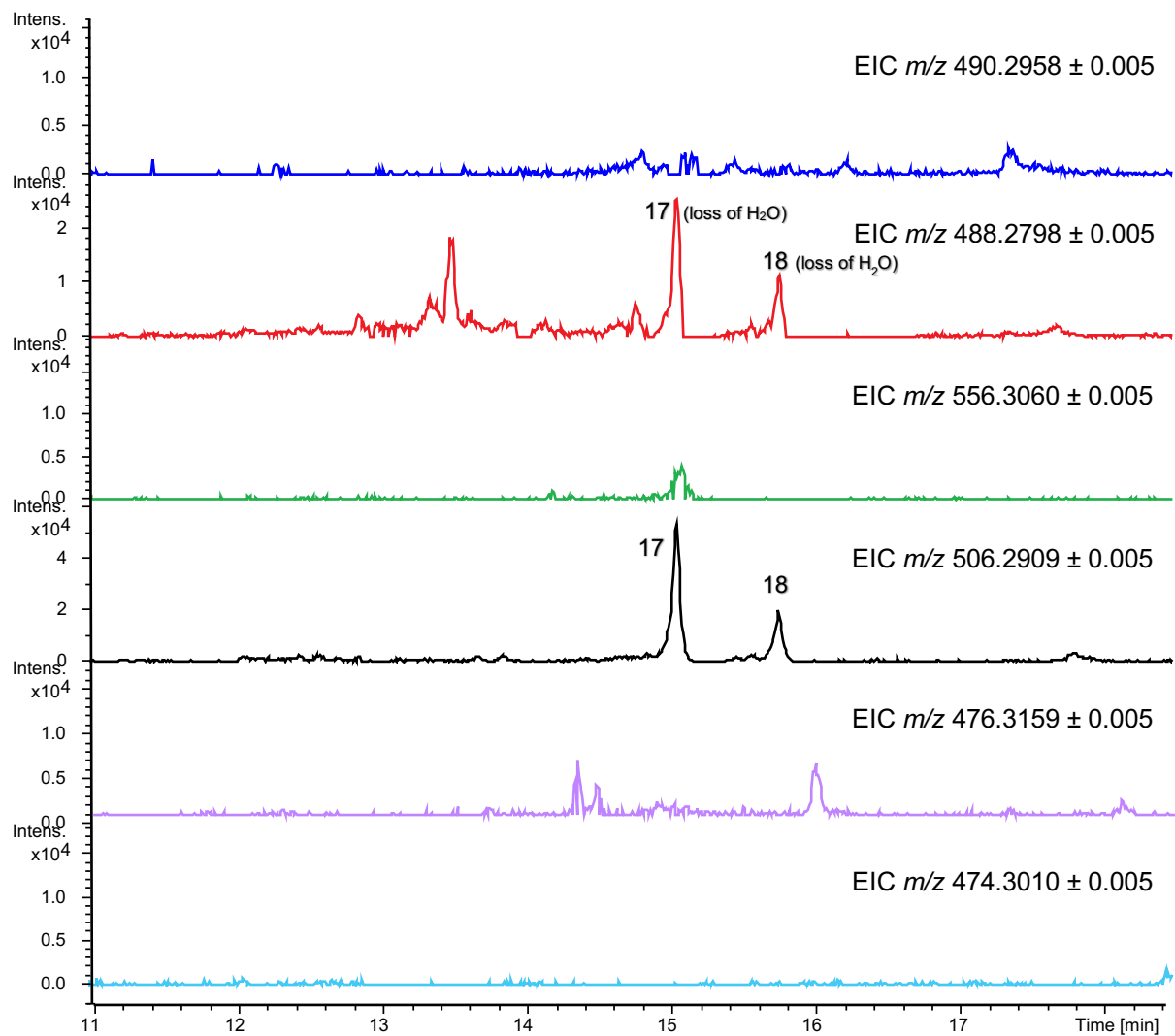


**Annex 10: Extract ion chromatograms obtained from LC-MS analysis of the metabolite extract from  $\Delta prcI$  *S. zeae* mutant cultured on PDA.**

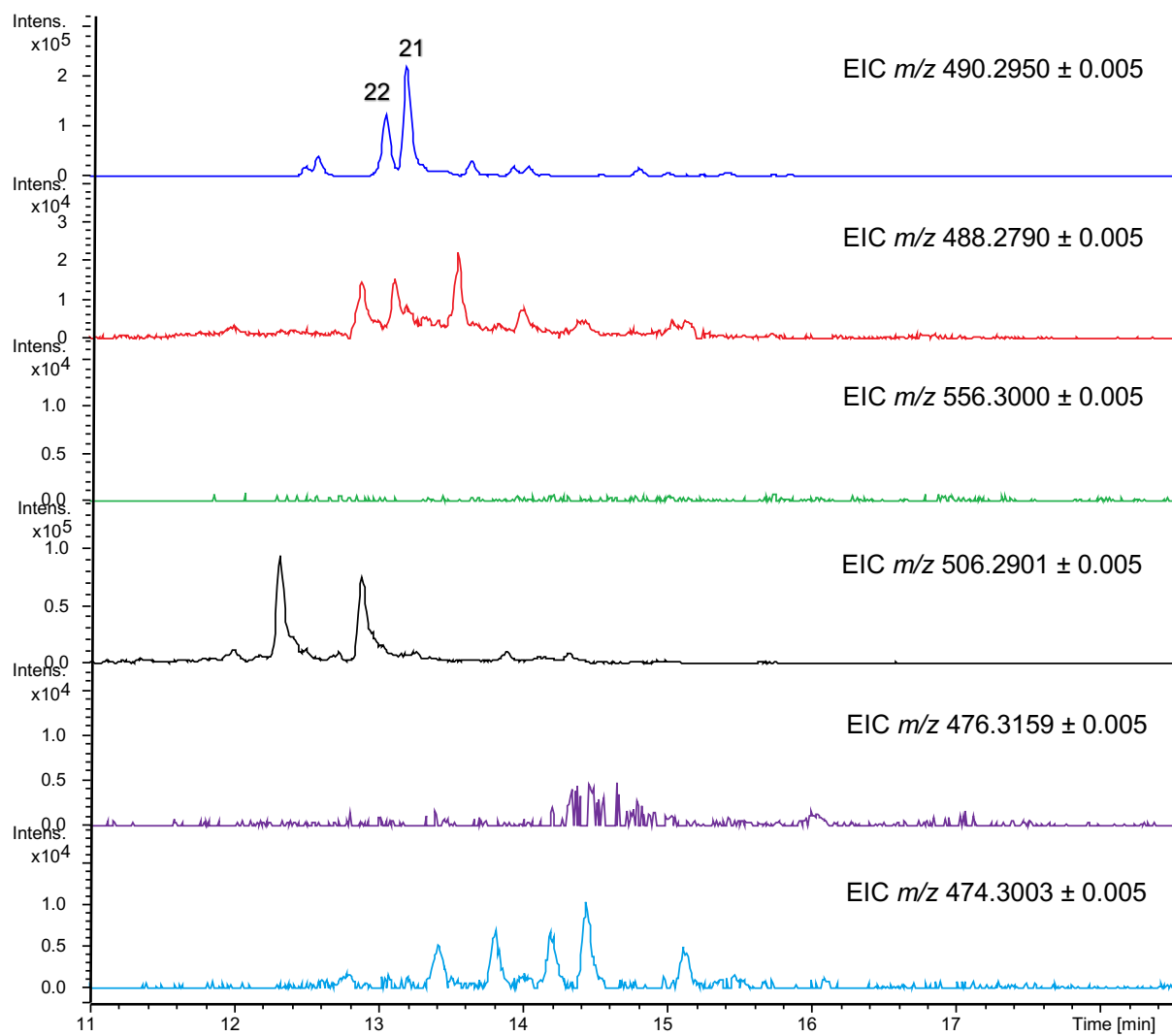




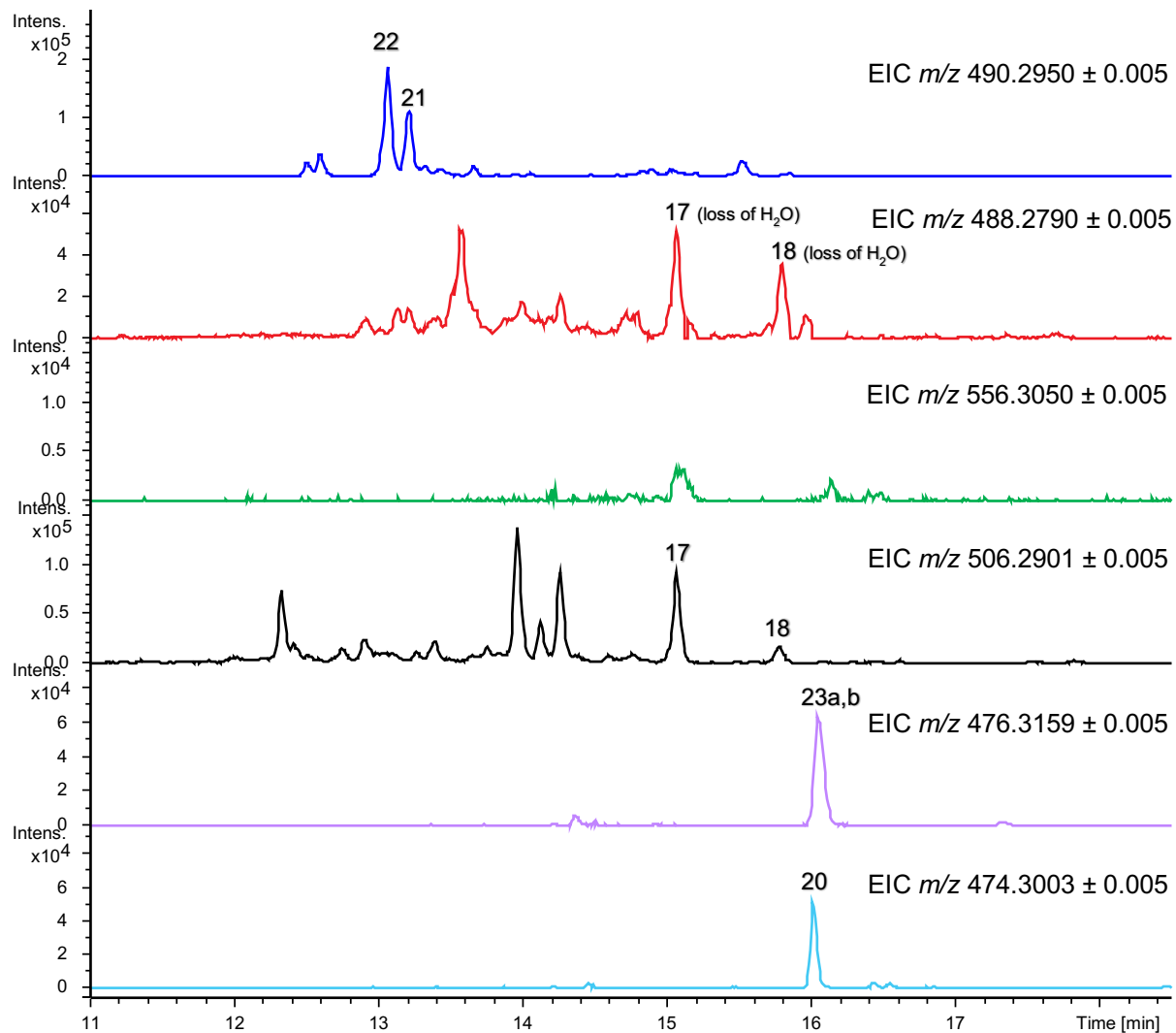
**Annex 011: Extract ion chromatograms obtained from LC-MS analysis of the metabolite extract from  $\Delta prcX$  *S. zeae* mutant cultured on PDA.**



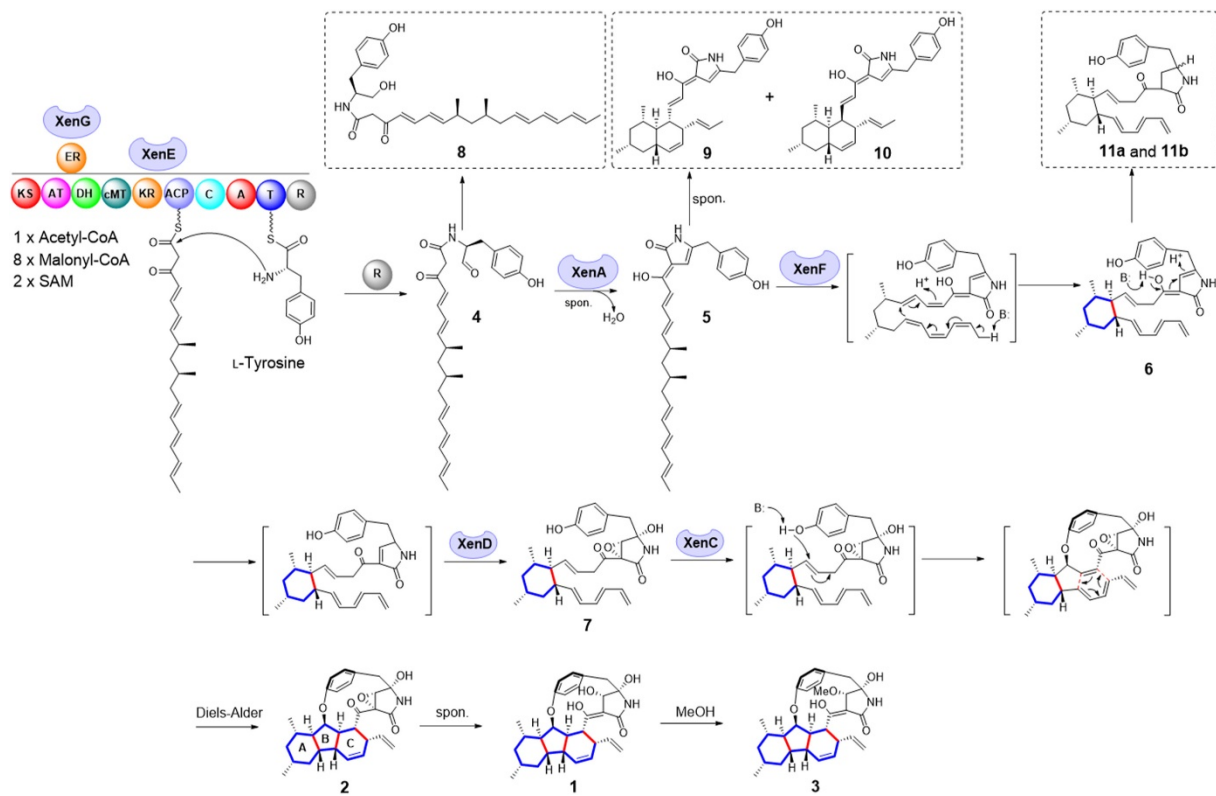
**Annex 12: Extract ion chromatograms obtained from LC-MS analysis of the metabolite extract from  $\Delta prcB$  *S. zeae* mutant cultured on PDA.**



**Annex 13: Extract ion chromatograms obtained from LC-MS analysis of the metabolite extract from  $\Delta prcD$  *S. zeae* mutant cultured on PDA.**



**Annex 14: Extract ion chromatograms obtained from LC-MS analysis of the metabolite extract from  $\Delta prcE$  *S. zeae* mutant cultured on PDA.**



**Annex 15: Proposed biosynthetic pathway of xenoacremones in *X. sinensis* ML-31**

## Annex 16: A manuscript is about to be submitted for publication

### Metabolic investigation of pyrrocidine pathway allows stereocontrol rationalization of paracyclophane-decahydrofluorene formation

Youwei Chen,<sup>a</sup> Steffi Sewurn,<sup>a</sup> Séverine Amand,<sup>a</sup> Caroline Kunz,<sup>ab</sup> Nicolas Pietrancosta,<sup>cd</sup> Kevin Calabro,<sup>a</sup> Didier Buisson<sup>a</sup> and Stéphane Mann<sup>a\*</sup>

<sup>a</sup> Laboratoire Molécules de Communication et Adaptation des Micro-organismes (MCAM), Muséum national d'Histoire naturelle, CNRS ; CP54, 57 rue Cuvier, 75005 Paris, France.

<sup>b</sup> Sorbonne Université, Faculté des Sciences et Ingénierie, UFR 927, F-75005 Paris, France.

<sup>c</sup> Laboratoire des Biomolécules, LBM, 75005 Paris, France. <sup>c</sup> Laboratoire des Biomolécules, LBM, Sorbonne Université, École Normale Supérieure, PSL University, CNRS, F-75005 Paris, France.

<sup>d</sup> Neurosciences Paris Seine - Institut de Biologie Paris Seine (NPS - IBPS), Sorbonne Université, INSERM, CNRS, F-75005 Paris, France.

#### Introduction

Pyrrocidines **1** and **2** (Fig. 1) are secondary metabolites isolated from diverse fungi [1–4] and characterized by antimicrobial [5] and apoptosis-inducing [6] activities, as well as potential inhibitory effects of SARS-CoV2 RNA-dependent RNA polymerase [7]. More importantly, pyrrocidines produced by the maize endophyte *Sarocladium zeae* (formerly classified as *Acremonium zeae*) are suggested to have a beneficial impact for the host plant, as sub-inhibitory concentrations of pyrrocidines can inhibit the fumonisin toxin production in the phytopathogen *Fusarium verticillioides* [8]. Pyrrocidines could thus be responsible, at least in part, for the protective effect of the producer in maize against phytopathogens. Pyrrocidines and analogues are therefore promising targets for the development of biocontrol agents. Pyrrocidines belong to a growing family of fungal natural products sharing a decahydrofluorene core connected to a highly strained paracyclophane moiety, which encompasses GKK1032A<sub>2</sub> **3** [9], hirsutellone B **4** [10], trichobamide A **5** [11], ascomylactams [12], [13], pyrrospirones [14] and embellicines [15]. These compounds attract much interest recently because of their diverse bioactivities (e.g. antimycobacterial, anti-glioma) and their complex structure. Each member possesses different combination of stereo-configuration at chiral centers present in the paracyclophane-decahydrofluorene motif, which represents a great challenge for chemical synthesis. This study is focused on a detailed understanding of pyrrocidine biosynthesis in *S. zeae*, in particular, the steps leading to polycyclic formation. Such knowledge will enable future pathway engineering efforts to generate pyrrocidine analogs with controlled stereochemistry.

The biosynthetic origin of this fungal scaffold was first studied by Oikawa [9] with isotopic incorporation of labelled acetate, L-tyrosine and methionine into GKK1032 biosynthesis. This suggested that a hybrid polyketide synthase – non-ribosomal peptide synthetase (PKS-NRPS) is involved. Later, feeding experiments on the maize endophyte *S. zeae* using dual labelling with <sup>18</sup>O,<sup>13</sup>C

L-tyrosine demonstrated the integration of the amino acid in pyrrocidines with conservation of the phenolic oxygen atom [16]. Consequently, radical or electrophilic cascades were suggested to form the paracyclophane and decahydrofluorene motifs. The proposed reaction path involved a pyrrolidone key intermediate I (Fig. 2) that is resulted from the PKS-NRPS nonaketide-tyrosine product by Knoevenagel condensation, and an intermediate II containing the paracyclophane and ring A (Figure 2). Given the presence of a cyclohexene motif (ring C), an intramolecular Diels-Alder reaction (IMDA) would occur from the intermediate II in an *exo*- or *endo*-cyclizing way to form pyrrocidines or GKK1032A<sub>2</sub>, respectively (Fig. 2). Very recently, Ohashi *et al.* heterologously expressed the biosynthesis pathways of pyrrocidine B **2** from *S. zeae* and GKK1032A<sub>2</sub> **3** from *Penicillium citrinum* in *Aspergillus nidulans* [17]. This work suggested that four proteins function together, likely as a complex, to transform intermediate I to II. However, the proposed mechanism of the cascade reaction, where the phenol oxygen needs to attack olefinic C13 at the *si* face to trigger conjugated addition of C12 to the triene at C7 to form a linkage *in trans*, would result in a reverse configuration of C13 in pyrrocidines and GKK1032A<sub>2</sub>. To note, this mechanism can instead account for the stereochemistry of C13 in hirsutellones. Concurrent to this publication, Liu *et al.* characterized xenoacremones and their biosynthetic pathway in *X. sinensis*, and proposed that one enzyme (XenF) catalyses [1,11]-sigmatropic rearrangement to install ring A in intermediate I (Fig. 2) [18]. However, this step was confusingly illustrated by a base catalyzed mechanism with abstraction of the C1 proton. Despite these proposed pathways for decahydrofluorene formation, the mechanism by which the paracyclophane is generated still remains an enigma. Indeed, crystal structures of pyrrocidines and analogues all showed an aromatic ring bending of the paracyclophane moiety (more than 12° from planarity). The hypothesized ring closing by the attack of the phenol oxygen to C13 resulting in a strained paracyclophane does not consider the highly energetic distortion of the aromatic ring. Moreover, this oxidizing process would require an unidentified hydride acceptor close to C1.

In the present study, we investigated pyrrocidine biosynthesis in *S. zeae* by gene knock-out and thorough metabolic analysis. This reveals an intrinsic plasticity of the pyrrocidine pathway and leads to the generation of complex metabolites with new cyclic backbones in the mutants. Moreover, structural analysis of key auxiliary enzymes by modelling and of the products accumulated in respective gene knock-out strains allowed us to propose a plausible mechanism of paracyclophane formation, via transient loss of aromaticity of tyrosine residue. Our analysis also gives molecular bases and enables a unified biosynthetic scheme that can reconcile the stereochemical diversity obtained in this family of natural products.

## **Results and Discussion**

## Identification and characterization of pyrrocidine metabolites produced by wild type *S. zeae*.

Firstly, we identified and characterized the pyrrocidines compounds produced by *S. zeae*. The fungi were cultured on PDA medium for seven days and the ethyl acetate metabolite extracts were analyzed by LC-MS. Pyrrocidines A **1** and B **2** were detected at 15.8 min and 15.9 min ( $[M+H]^+$  ions at  $m/z$  488.2795 and 490.2952), respectively, and characterized by NMR (Table S5,S6). Careful analysis of the  $m/z$  490.295 extracted ion chromatogram (EIC) (Fig. S5) shows two other small peaks for compounds **7a** at 15.5 min and **7b** at 14.9 min with similar LC-MS<sup>2</sup> fragmentation patterns as pyrrocidine B, suggesting stereoisomers of **2**. Interestingly, the culture of the fungi on maize medium for 30 days leads to the accumulation of these compounds compared to **1** and **2**, enabling access to sufficient amounts for isolation. Compound **7a** was purified and we established its planar structure from 2D NMR experiments (Fig. 3, Tables S7, S8) in CDCl<sub>3</sub> and by comparison with NMR data of **2**. The relative stereochemistry was elucidated from NOESY spectrum defining **7a** as the diastereomer of **2** at C3 and C6 position. This diastereomery is identifiable by <sup>13</sup>C NMR with notable shielding of C7, C28 and C29 chemical shifts of ca. 6 ppm in compound **7a** with respect to **2**. These downfield shifts are thus a signature of a *trans*-fused junction of rings B and C in pyrrocidines.

Compound **7b** turned out to be unstable in CDCl<sub>3</sub> but, when analyzed by LC-MS pure ketone **7a** in acetonitrile showed partial conversion to **7b** that we interpreted as the enol form of **7a** (Fig. 3). In turn, purified enol **7b** interconverted with **7a**. This was also observed by NMR when ketone **7a** is solubilized in DMSO-*d*<sub>6</sub> which equilibrates with its enol form (Table S9) in a respective 2:1 ratio. Enol **7b** was identified by the key changings of chemical shifts at C16, C17 and C19. The structures of **7a,b** are in agreement with those published by Ohashi *et al.* [17].

Beside these metabolites, we isolated trichobamide B **8** with the molecular formula C<sub>35</sub>H<sub>41</sub>NO<sub>5</sub> (assigned by HR-ESIMS,  $m/z$  556.3057  $[M+H]^+$ ) identical to trichobamide A **5** [11]. NMR analysis shows that both compounds are isomers differentiating in the substitution of the tetrahydro furo-pyrrolone moiety at C33 and C34 (Table S10). Additionally, pyrrocidine C **9** [5],[16] and epi-pyrrocidine A [3] named here pyrrocidine F **11** (Fig. 3), were detected by LC-MS at 15.5 min with  $m/z$  504.2744 ( $[M+H]^+$ ) and at 15.0 min with  $m/z$  488.2795 ( $[M+H]^+$ ) (Fig. S13), respectively, and characterized by NMR (Tables S11, S12).

## Identification of the biosynthetic gene cluster

In order to identify the biosynthetic gene cluster of pyrrocidines, we sequenced and assembled the genome of *Sarocladium zeae*. Among the resulting 230 DNA scaffolds covering the 31.2 Mb genome, we identified three putative hybrid PKS-NRPS genes by BLAST with the hybrid PKS-NRPS TenS as reference [19]. To establish the biosynthetic link between pyrrocidines and one of these genes, we



determined experimental procedures suitable for gene inactivation in *S. zeae* based on homologous recombination of a hygromycin resistant cassette *via* protoplastisation and PEG-mediated transformation of the fungi [20]. This procedure, independently applied to the three putative megasynthase genes, lead to the extinction of pyrrocidine production only in the case of the PKS-NRPS gene named *prcA*. Prediction of coding sequences (CDS) surrounding *prcA* with *FgenesH* software [21] revealed nine putative proteins encoded by genes annotated from *prcH* to *prcI* (Fig. 4A). From sequence analysis we identified PrcC as the *trans*-acting enoyl reductase (ER) associated to the hybrid PKS-NRPS PrcA, PrcH as an  $\alpha,\beta$ -hydrolase, PrcF as a transcription factor, PrcG as a transporter, PrcI as a second ER and PrcB homologous to Diels-Alderases. In contrast, PrcD and PrcE appeared as integral membrane proteins with unknown function.

Because *prcI* was located near the edge of the DNA scaffold, we investigated the possibility that the BGC was truncated. A 5,000 cloned fosmid library built from *S. zeae* genomic DNA was screened by dot blot hybridization using the *prcI* DNA sequence as a probe. The junction between BGC and its right-contiguous scaffold was established showing that the extended DNA sequence harbors two genes coding for PrcX, a protein belonging to the lipocalin family with unknown function and PrcJ as a transcription factor. Genome mining approach allowed us to identify twelve similar gene clusters in the sequenced fungi databases (Fig. S21).

### **Function assignments to the auxiliary enzymes of the pyrrocidine BGC**

To assign a precise function to the proteins encoded by the BGC, we performed gene inactivations in *S. zeae*. The metabolic extracts of the mutants were analyzed by LC-MS (Fig. 4B) and the accumulating products were purified and their structure elucidated by NMR analysis.

#### **Inactivation of *prcI* gene**

Firstly, the deletion of *prcI* resulted in the abolishment of the production of pyrrocidines B **2** and D **7a,b** but not pyrrocidine A **1** (Fig. S8) showing that the enoyl-reductase PrcI reduces the double bond of pyrrolone moiety.

#### **Inactivation of *prcX* gene**

The deletion of *prcX* gives rise to a strong decrease of production of **1** and **2**, without impacting the production of pyrrocidines D **7a,b**. This result shows that **1** and **2** are naturally produced without this lipocalin protein (Fig. S9). PrcX drives the Diels-Alder cycloaddition in favor of the *exo* product as mentioned by Ohashi *et al.* [17]. Moreover, in this mutant due to the dramatic decrease of pyrrocidine B, we could detect and purified two new metabolites named pyrrocidine E **10** and saroclidine A **12** (Fig. 3). The structure of the latter, also produced by wild type *S. zeae*, with a MF  $C_{32}H_{39}NO_5$  ( $[M+H]^+$  ion at  $m/z$  518.2901), was assigned by NMR analysis to the C3-epimer of GKK1032C [22]. Indeed,

NOESY spectrum through the cross-correlations H31-H24-H13-H6-H12-H15, H13-H15-H1a-H6 and H14-H26-H29 shows that H12, H13, H15, H6, methyl 31 and vinyl group are cofacial while H14 and methyl 29 are on the opposite face of the decahydrofluorene (Tables S12, S14). The formation of **12** thus indicates that PrcA is able to perform four and five methylations on the polyketide which is to our knowledge the first reported example for a PKS-NRPS. Compound **10** has the MF  $C_{31}H_{37}NO_5$  ( $[M+H]^+$  at 15.4 min,  $m/z$  504.2736) whose structure was assigned to the C3, C6 diastereomer of pyrrocidine C **9**. The *trans*-fused junction of rings B and C in **10** and also in **12** was corroborated by the shielding of C7, C28 and C29 chemical shifts as mentioned in **7a,b**.

#### Inactivation of *prcH* gene.

The inactivation of the encoding  $\alpha,\beta$ -hydrolase gene *prcH* generates accumulation of a metabolite with  $[M+H]^+$   $m/z$  494.3297 (MF:  $C_{31}H_{43}NO_4$ ) (Fig. S7) whose structure was assigned by NMR to the all-*E* polyene alcohol **14** (Table S15). Disruption of  $\alpha,\beta$ -hydrolases involved in PKS-NRPS BGCs was reported by others [23], [24] and also lead to the formation of alcohols as shunt products. Formation of the alcohol likely finds its origin in the over-reduction by the R domain [25] of an elusive aldehyde intermediate **13** released by the PKS-NRPS (Fig. 5). While PCR analysis of the mutant and BLAST search in the genome showed the absence of another copy of *prcH* gene or a similar protein encoded in the genome able to substitute PrcH, the production of pyrrocidines is not impaired in this mutant. Recently, Zhang *et al.* showed from synthetic aldehyde that Knoevenagel condensation can proceed spontaneously to afford pyrrolone moiety [26], which can explain the conversion of **13** to **15** and next to pyrrocidines and the fact that **13** was not detected in the extract. Therefore, together these results indicate that PrcH is not required for the completion of the biosynthesis of pyrrocidines and its role is here conceivably to prevent the reduction of the PKS-NRPS off-loaded reactive aldehyde **13** in order to promote its ring closing into pyrrolone **15** through a Knoevenagel condensation.

#### Inactivation of *prcB*, *prcD* and *prcE* genes

Individual disruptions of the *prcB*, *D* and *E* genes all lead to annihilation of the production of pyrrocidines and to accumulations of new metabolites (Fig. 4B). In  $\Delta prcB$  mutant, we observed the formation of products **17** and **18** with  $[M+H]^+$  of  $m/z$  506.289 eluted at 14.9 and 15.8 min, respectively (Fig. S10). Structure elucidation showed that **17** is a 5-hydroxy-3,4-epoxy-pyrrolidone substituted with a linear polyolefinic chain (Table S16) as in **14**. Compound **18** is a *cis*-decaline with a relative stereochemistry deduced from NOESY data thanks to cross correlations H12-H29-H10b-H14-H15-H4 (Table S17). An *exo*-selective IMDA of **17** can explain the formation of this decaline (Fig. S16).

In  $\Delta prcE$  mutant, we detected a peak by LC-MS at 16.0 min, with  $m/z$  476.3146 for  $[M+H]^+$  corresponding to the MF  $C_{31}H_{41}NO_3$  (Fig. S12). NMR analysis revealed a mixture of two trimethyl-cyclohexanes **23a,b** in a 1:1 ratio which are *E/Z* isomers of C13-C14 olefin (measured coupling

constants: **23a**  $^3J_E = 15.3$  Hz and **23b**  $^3J_Z = 9.3$  Hz) (Tables S21, S22). The trimethyl-cyclohexane corresponds to the ring A of pyrrocidines substituted with a conjugated trienic chain which was isomerized to the C1 terminal position unlike to the polyketides **14** and **17**.

The same mutant also produces compound **20** detected at 15.8 min with a MF of  $C_{31}H_{39}NO_3$  (assigned from  $[M+H]^+$   $m/z$  474.2998) (Fig. S12). NMR analysis revealed a unique 6/10/6/5 fused polycyclic structure substituted by a tyrosine residue. This macrocycle fused to an amino acid-derived octahydro-isoindeole belongs to cytochalasans [27]. It was named pyrrochalsin A and represents the first nonaketide-tyrosine member of this family. Relative stereochemistry was assigned by NOESY experiment, on the base of two identified groups of cross correlations H30-H10b-H31-H12-H14-H28 and H29-H6-H4-H19-H23 indicating that these groups of protons are respectively on each side of equatorial plan of the molecule (Table S18).

From LC-MS chromatograms of  $\Delta prcE$  and  $\Delta prcD$  mutants (Fig. S12, S11), we identified two other accumulating compounds named pyrrocyclines A **21** and B **22** detected at 13.2 and 13.0 min, respectively, with for both a MF  $C_{31}H_{39}NO_4$  (assigned from  $[M+H]^+$   $m/z$  490.293). Their planar structures were established showing a unique backbone consisting of a 15-membered macrocycle fused to 6- and 5-membered rings. NOESY experiments performed on **21** and **22** indicates that they adopt a conformation in which protons H14, H12, H6, H8, H4, H2 and NH are oriented on the same face of the equatorial plan of the molecule while H23/27, H21, H19, H1 H3, H28, H29, H13, H11 are located on the opposite face (Tables S19, S20).

All the accumulating metabolites from  $\Delta prcE$  or  $\Delta prcD$  harbored the trimethyl cyclohexane motif corresponding to the ring A of pyrrocidines. By contrast, this cyclohexane is not observed among the accumulating metabolites in the  $\Delta prcB$  mutant. We deduced that PrcB installs ring A with an isomerization-cyclization step from pyrrolidone **15** to the key intermediate **19** (Fig. 5). The latter then gives rise to **20** in  $\Delta prcE$  through an endo-selective IMDA as in the biosynthesis of cytochalasans. Such a stereocontrolled cyclization required a pericyclase as demonstrated by Hankte *et al.* in pyrrochalsin H biosynthesis [27] suggesting that the formation of **20** is not spontaneous and may be governed by a protein like PrcX. Macrocyclization of **19** after epoxidation at C1-C2 or C3-C4 olefins can provide the compounds **21** and **22** (Fig. S18).

Comparaison of LC-MS analysis of all the mutants shows that **17** and **18** are not produced when  $prcD$  is inactivated (Fig. S6-S12). This indicates that PrcD is involved in the oxidation of pyrrolidone moiety of **15**. Autooxidation of strained olefins has been reported to give epoxides [28] and this was also exemplified by the known conversion of **1** to **9** [5], [16]. It can be deduced that PrcD catalyses the hydroxylation of **15** on C20 to give **16** which is then converted into **17** by spontaneous olefin

epoxidation. The isolation of cyclohexanes **20**, **21** and **22** implies that **15** is substrate of PrcB to afford cyclohexane **19** and then PrcD-catalysed C20 oxidation occurs to give **24**. It cannot be excluded as well that hydroxylated **16** is substrate of PrcB to give **24**. The latter in  $\Delta prcE$  mutant can give **23ab** after a reduction-dehydration sequence (Fig. S19). Combining all these results, we deduced that from **24** PrcE build the paracyclophane **25** which next undergoes an *exo* IMDA assisted by PrcX to afford **1** (Fig. 5).

In order to validate our hypothesis and decipher the catalytic mechanism of the key auxiliary enzymes PrcB, D and E, the corresponding genes were cloned to carry out *in vitro* study. However, the production of heterologously expressed recombinant proteins was unsuccessful. Therefore, we turned towards *in silico* analysis of the proteins.

### **Tridimensional model of PrcB and docking calculations with biosynthetic intermediates.**

Sequenced alignments of PrcB show that its closest homologs with identified functions belongs to the fungal decalin synthases (DS) family like CghA, Fsa2 and Phm7 (21%, 22% and 24% sequence identities, respectively; Fig. S23). These three proteins adopt the same fold made of a N-terminus  $\beta$ -sandwich domain connected by a short  $\alpha$ -helix to a C-terminus  $\beta$ -barrel domain as revealed by recent crystal structure resolutions. [29–31] The DS active sites are located in the cavity at the contact of these two domains. Noteworthy, compared to the DS, PrcB has extended sequences predicted to structure as  $\alpha$ -helices at both of its extremities and conserved in all PrcB homologs found in pyrrocidine-like BGCs (Fig. S22). The N-terminus helix from residues 5 to 22 which has the signature of a transmembrane segment could thus anchor the protein to the membrane and direct it toward the integral transmembrane proteins PrcD and PrcE, forming a complex able to channel reactive intermediates up to the end product.

In order to investigate the catalytic mechanism of PrcB, we modeled this protein according to the crystal structure of CghA. We thus delineate a pocket at the interface of the two domains and to determine whether this could be the active site we performed docking calculations with substrate **15**. We obtained several similar poses fitting the cavity in which the molecule adopts an elbow shape with the phenoxy group anchored in a calix made of four tryptophan residues (Trp81, Trp254, Trp261, Trp276) as well as Leu248 (Fig. 7 and S27). The hydrocarbon chain from C7 to C12 is preorganized in a chair-like conformation with the C30 and C31 methyls in equatorial position and C29 in axial position as in the ring A. This part of the molecule interacts with residues Tyr120, Tyr360, Met362, Met370 and Leu132. In the vicinity of the C12-C15 diene we also identify the presence of the residue Glu118. Interestingly, the terminal methyl C1 points towards the guanidinium of Arg265 which interacts with both carboxylates of Asp87 and Glu85 in a salt bridge and charge-charge interaction, respectively. All these residues involved in the binding pocket are conserved in PrcB homologs found in BGC (Fig. S22).

In xenoacremone biosynthesis the formation of ring A by the protein XenF (80.2 % sequence identity with PrcB including conservation of all binding site residues of substrate **15**) was speculated to proceed via a [1,11] sigmatropic rearrangement [18]. However, our results on PrcB indicates that the conformation adopted by the substrate in the cavity strongly distances C1 from C15 which does not support a [1,11] sigmatropic rearrangement where the required cyclic transition state should bring closer C1 and C15 for H migration. By contrast arginine residues are known to act as general base catalysts in various enzymes in which the guanidinium is activated through an interaction network with carboxylate groups from other residues [32], [33]. Consequently, we proposed that Arg265 activated by Glu85 and Asp87 deprotonates C1 and initiates the isomerization of double bonds along the chain. This triggers the attack of C7 from its *si* face onto the *si* face of C12 to afford ring A after protonation of C15 by Glu118. Noteworthy, this tetrad of residues is specific to PrcB homologs found in similar pyrrocidine BGCs but not to DS: either Arg265 is conserved (CghA, Fsa2) either Glu85 and Glu118 are kept (Phm7).

Among the different docking poses of substrate **15**, the C13-C14 bond adopts either *s-trans* (most of the cases) or *s-cis* conformation. By the cycloisomerization mechanism we proposed, these conformers generate the (13-14)-*E*-alkene **19** and its (13-14)-*Z* isomer **26**, respectively (Fig. 6). Further reduction of pyrrolidone ring of these intermediates, likely by PrcI, thus accounts for the formation of isolated compounds **23a,b**.

Docking of the hydroxylated pyrrolidone **16** was also carried out and similar binding modes as **15** were obtained which places this molecule as a plausible biosynthetic intermediate.

### **Topology of the transmembrane proteins PrcD and PrcE and formation of the paracyclophane.**

Since PrcD and PrcE are predicted as integral membrane proteins, we investigated their topology with various prediction tools [34–36] and we identify four transmembrane segments (TM1-4) in both proteins. Four-transmembrane proteins present a wide diversity of topology [37]. Despite the lack of sequence homology of PrcD and E with reported structures in this protein family, comparison of the organization of TM domain in both proteins was carried out. We found that PrcD (TM1-4 located at residues 6-28, 61-83, 89-111 and 142-162, respectively) has an architecture resembling to connexin monomer (Fig. S28-S30). However, no enzymatic function was identified in this TM family allowing to propose a mechanism for PrcD.

In contrast, the architecture of PrcE with transmembrane segments (TM1-4 located at residues 10-29, 60-78, 84-104 and 107-123, respectively) and a 30 residues loop between TM1 and TM2 structured by a short  $\alpha$ -helix and a strand is similar to the vitamin K epoxide reductase (VKOR) [38], [39] and also to the disulfide bond formation protein DsbB [40], [41]. These enzymes use cysteine pairs

as active site residues which switch between reduced and disulfide-bridge states to reduce a quinone (vitamin K in VKOR and ubiquinone or menaquinone in DsbB) to its hydroquinone form. In this process one of the cysteines covalently binds to the substrate to form a transient cysteine-quinone adduct. Sequence analysis of PrcE reveals the presence of a unique pair of cysteines positioned almost in the middle of TM1 (Cys22) and TM4 (Cys114), facing each other in a way allowing a disulfide bond formation inside the four-helix bundle scaffold.

Similarly to the mechanism of VKOR and DsbB, we reasoned that a transient cysteine-quinone adduct can form in PrcE between C22 of the intermediate **24** and sulfur of one of the cysteines forming a disulfide bridge (Fig. 6). The resulting introduction of sp<sup>3</sup> hybridization on C22 releases the chain constrain allowing C13 and O to come closer and bind. We propose unlike to other publications that the ring closing occur by the attack of C13 on the oxygen of the transient quinone. This reactivity is supported by the work of Guo and Mayr reporting the O-electrophilicity of activated quinones towards Pi-nucleophiles [42]. This attack is triggered by the abstraction of the *pro-S* proton of C15 to generate the *E*-C14-C15 double bond of **25** required for the subsequent IMDA step. The conserved Lys87 residue is a good candidate to remove the H15 proton since it is located at the beginning of TM3 and can be positioned on the opposite side of the substrate compared to the cysteines (Fig. S31). Two conserved histidines are also well positioned to protonate/stabilize the cysteines after breaking of the disulfide bond. Indeed, these basic residues are located in TM1 (His26) and TM4 (His110) at one turn beyond the cysteines and aligned on the same side with them (see  $\alpha$ -helix wheel representation in Figure SX). To complete the catalytic cycle of PrcE, the disulfide bond can be then regenerated by ubiquinone cofactor as in DsbB.

After ring closing, the strained paracyclophane precludes any flip of the functional groups like the C14-C15 double bond or the pyrrolidone moiety. Thus, the formation of pyrrocidine F **11**, in which the pyrrolidone is oriented in the opposite way of the one found in pyrrocidine **1**, reveals that firstly epimerization at C20 occurs before cyclophane formation likely via enolization of **19** and secondly PrcE can tolerate both orientations of this motif.

### **Proposed biosynthetic pathway of xenoacremones**

The absolute stereochemistries of C14 and C15 in pyrrocidine A **1** implies that precursor **25** exposes the C14 and C15 *si* faces of the *E*-dienophile locked in the paracyclophane towards the diene to perform IMDA. By contrast, in xenoacremone C **6** the absolute configurations of the same carbons are reversed meaning that the *E*-dienophile precursor exposes its *re*-faces to the diene. In this case the ring closing of the paracyclophane occurs in a slightly different way compared to pyrrocidines. Indeed, since the C13 stereochemistry is the same in xenoacremone C **6** and pyrrocidine A **1**, the C13-C14

double bond in the substrate **30** of XenC has to be in *Z* configuration (Fig. S35). Thus, from **28** with *s-cis* C13-C14 bond the isomerizing-cyclase XenF forms the *Z*-C13-C14 alkene **29** which is then hydroxylated by XenD into **30**. Once the cysteine-quinone adduct is formed in XenC, abstraction of the *pro-R* H15 triggers the paracyclophane ring closing and locks the *E*-C14-C15 dienophile exposing its *re*-faces to the diene. We observed that the cysteine pair is conserved in XenC, Cys21 is in TM1 while Cys72 is in TM2 instead of TM4 in PrcE. This difference in cysteine positioning modify the orientation of the adduct by respect to the base, allowing the abstraction of *pro-S* H15 rather than *pro-R* in PrcE and consequently locking the dienophile in **31** in the opposite orientation of the one of **25**.

The next step finalizes the formation of decahydrofluorene moiety from **31** through an endo-cyclizing IMDA and was proposed to be spontaneous by *Liu et al.* However, our analysis of the BGC revealed at its extremity the presence of a partial gene coding for a N-terminal truncated homolog of the lipocalin PrcX (61 % sequence identity over 79 % of *prcX* length) that we called XenX (Fig. S26). This protein very likely controls the stereoselectivity of the IMDA in favor of endo-product as PrcX controls the *exo*-product formation in pyrrocidine pathway. The LC-MS data of the reconstitution of xenoacremones A and B (without XenX) [18] support this hypothesis since, whereas in the wild type these compounds correspond to unique peaks, in the reconstituted pathway these two metabolites are minorly produced beside other isomers (see EIC *m/z* 478.26 and *m/z* 476.24 of figures S5 and S19 in [18]). However, rather than to see PrcX and XenX selectively drive the IMDA in a respectively *exo* and *endo* manner, as it would suggest, these proteins should be considered from their common point which is their ability to orient the faces of dienes in a same way (namely C3-*re* and C6-*re* faces) towards differently constrained dienophiles. In this way they could be considered as dirigent proteins by analogy with these stereocontrolling phenoxy radical coupling proteins sharing the same lipocalin fold [43], [44].

### **A unified biosynthetic pathway for paracyclophane-decahydrofluorene in natural products**

The enzymatic mechanisms we proposed allow to explain how each stereogenic carbon center is controlled in the biosynthesis of paracyclophane-decahydrofluorene entity found in diverse reported natural products (Fig. 8). This control operates at three levels by proteins which are very closely linked. Indeed, at a first level PrcB-like isomerizing cyclase forms the ring A by establishing the always *trans* (*S*)-C7-(*R*)-C12 junction. However, this enzyme has a dual function given that it also sets the *E*- or *Z*-configuration of C13-C14 alkene precursor after double bond displacements. At a second level the PrcE-like cyclophane synthase controls the formation of *E/Z* stereochemistry of C14-C15 double bond by abstracting *pro-R* or *pro-S* proton on C15. Accordingly, this triggers the attack of C13 from its *si* or *re* face towards oxygen of cysteine-quinone adduct. We hypothesized that the observed stereoselectivity on C13, C14 and C15 is governed by the orientation of the cysteine-quinone adduct

in the enzyme which is guided by the relative location of both cysteines among the four TM domains. This is supported by the sequence analysis of the PrcE homologs identified in pyrrocidine-like BGCs (Figure SX) which shows several combinations for the repartition of cysteines. In most of the cases cysteine pair is positioned in TM1 and TM4 (e.g. PrcE in *S. zeae*) and also in TM1/TM3 (e.g. GkaZ *P. citrinum*) which both give the same stereochemistry at C13, C14 and C15. The couple could also be in TM1/TM2 (e.g. XenC in *Xenoacremonium sp.*), TM2/TM3 (*T. terrestris*), TM3/TM4 (*C. javanica*) or both cysteins in TM3 (*P. curvata*). The three latter combinations with any assigned metabolites can potentially generate stereo-related compounds to embellicines, ascomylactams or hirsutellones. According to this mechanism, the *E*- or *Z*- C14-C15 double bond formed in the paracyclophane, can be locked in a way selecting the exposed face of the dienophile towards the diene in the subsequent IMDA. In this last stage, the C3 and C6 stereochemistries are induced by spontaneous cycloaddition or controlled by the presence of a PrcX-like lipocalin.

**Acknowledgement:** Dr Yanyan Li and Dr Bastien Nay are thanked for initiating the project of genome sequencing, and Pr Rolf Müller and his group for assistance in the construction of genomic DNA library. We thank Alexandre Delporte, Nouha Dallel and Daria Barbash who contributed to the project during their internship. We acknowledge Dr Alain Blond and Alexandre Deville from the NMR service of MNHN and Dr Arul Marie from the mass spectrometry service of MNHN.

**Fundings** Emergence Project from Sorbonne University, ATM from MNHN, CNRS, China Scholarship Council, Doctoral School ED406 Sorbonne University.

## References

- [1] H. He, H.Y. Yang, R. Bigelis, E.H. Solum, M. Greenstein, G.T. Carter, Pyrrocidines A and B, new antibiotics produced by a filamentous fungus, *Tetrahedron Lett.* 43 (2002) 1633-1636.
- [2] D.T. Wicklow, S.M. Poling, R.C. Summerbell, Occurrence of pyrrocidine and dihydroresorcylyde production among *Acremonium zeae* populations from maize grown in different regions, *Can. J. Plant Pathol.* 30 (2008) 425-433.
- [3] Y. Shiono, M. Furukawa, T. Koseki, E. Kwon, A.H. Kurniawan, S. Sato, D. Harneti, R. Maharani, U. Supratman, S. Uesugi, K. Kimura, A pyrrocidine derivative produced by fungus *Neonectria ramulariae* In-2 isolated from a Beetle *Holotrichia picea*, *Phytochem. Lett.* 26 (2018) 120-124.
- [4] L. Błaszczyk, A. Waśkiewicz, K. Gromadzka, K. Mikołajczak, J. Chełkowski, *Sarocladium* and *Lecanicillium* Associated with Maize Seeds and Their Potential to Form Selected Secondary Metabolites, (2021) 11.
- [5] D.T. Wicklow, S.M. Poling, Antimicrobial Activity of Pyrrocidines from *Acremonium zeae* Against Endophytes and Pathogens of Maize, *Phytopathology* 99 (2009) 109-115.
- [6] S. Uesugi, N. Fujisawa, J. Yoshida, M. Watanabe, S. Dan, T. Yamori, Y. Shiono, K. Kimura, Pyrrocidine A, a metabolite of endophytic fungi, has a potent apoptosis-inducing activity against HL60 cells through caspase activation via the Michael addition, *J. Antibiot. (Tokyo)* 69 (2016) 133-140.



- [7] K.S. Ebrahimi, M. Ansari, M.S. Hosseyni Moghaddam, Z. Ebrahimi, Z. Salehi, M. Shahlaei, S. Moradi, In silico investigation on the inhibitory effect of fungal secondary metabolites on RNA dependent RNA polymerase of SARS-CoV-II: A docking and molecular dynamic simulation study, *Comput. Biol. Med.* 135 (2021) 104613.
- [8] D.T. Wicklow, S. Roth, S.T. Deyrup, J.B. Gloer, A protective endophyte of maize: *Acremonium zeae* antibiotics inhibitory to *Aspergillus flavus* and *Fusarium verticillioides* 1 Dedicated to John Webster on the occasion of his 80th birthday, *Mycol. Res.* 109 (2005) 610-618.
- [9] H. Oikawa, Biosynthesis of Structurally Unique Fungal Metabolite GKK1032A2: Indication of Novel Carbocyclic Formation Mechanism in Polyketide Biosynthesis, *J. Org. Chem.* 68 (2003) 3552-3557.
- [10] M. Isaka, N. Rugseree, P. Maithip, P. Kongsaree, S. Prabpai, Y. Thebtaranonth, Hirsutellones A–E, antimycobacterial alkaloids from the insect pathogenic fungus *Hirsutella nivea* BCC 2594, *Tetrahedron* 61 (2005) 5577-5583.
- [11] S. Chen, H. Shen, P. Zhang, H. Cheng, X. Dai, L. Liu, Anti-glioma trichobamide A with an unprecedented tetrahydro-5 *H*-furo[2,3-*b*]pyrrol-5-one functionality from ascidian-derived fungus *Trichobotrys effuse* 4729, *Chem. Commun.* 55 (2019) 1438-1441.
- [12] Y. Chen, Z. Liu, Y. Huang, L. Liu, J. He, L. Wang, J. Yuan, Z. She, Ascomylactams A–C, Cytotoxic 12- or 13-Membered-Ring Macrocyclic Alkaloids Isolated from the Mangrove Endophytic Fungus *Didymella* sp. CYSK-4, and Structure Revisions of Phomapyrrolidones A and C, *J. Nat. Prod.* 82 (2019) 1752-1758.
- [13] L. Wang, Y. Huang, C. Huang, J. Yu, Y. Zheng, Y. Chen, Z. She, J. Yuan, A Marine Alkaloid, Ascomylactam A, Suppresses Lung Tumorigenesis via Inducing Cell Cycle G1/S Arrest through ROS/Akt/Rb Pathway, *Mar. Drugs* 18 (2020) 494.
- [14] T. Song, M. Chen, W. Chai, Z. Zhang, X.-Y. Lian, New bioactive pyrrospirones C–I from a marine-derived fungus *Penicillium* sp. ZZ380, *Tetrahedron* 74 (2018) 884-891.
- [15] W. Ebrahim, A.H. Aly, V. Wray, A. Mándi, M.-H. Teiten, F. Gaascht, B. Orlikova, M.U. Kassack, W. Lin, M. Diederich, T. Kurtán, A. Debbab, P. Proksch, Embellicines A and B: Absolute Configuration and NF- $\kappa$ B Transcriptional Inhibitory Activity, *J. Med. Chem.* 56 (2013) 2991-2999.
- [16] A. Ear, S. Amand, F. Blanchard, A. Blond, L. Dubost, D. Buisson, B. Nay, Direct biosynthetic cyclization of a distorted paracyclophane highlighted by double isotopic labelling of L-tyrosine, *Org. Biomol. Chem.* 13 (2015) 3662-3666.
- [17] M. Ohashi, T.B. Kakule, M.-C. Tang, C.S. Jamieson, M. Liu, K.N. Houk, Y. Tang, Biosynthesis of para-Cyclophane Containing Hirsutellone-Family of Fungal Natural Products, 71.
- [18] Z. Liu, W. Li, P. Zhang, J. Fan, F. Zhang, C. Wang, S. Li, Y. Sun, S. Chen, W. Yin, Tricarbocyclic core formation of tyrosine-decahydrofluorenes implies a three-enzyme cascade with XenF-mediated sigmatropic rearrangement as a prerequisite, *Acta Pharm. Sin. B* (2021) S2211383521001064.
- [19] K.M. Fisch, W. Bakeer, A.A. Yakasai, Z. Song, J. Pedrick, Z. Wasil, A.M. Bailey, C.M. Lazarus, T.J. Simpson, R.J. Cox, Rational Domain Swaps Decipher Programming in Fungal Highly Reducing Polyketide Synthases and Resurrect an Extinct Metabolite, *J. Am. Chem. Soc.* 133 (2011) 16635-16641.
- [20] W. Hamada, P. Reignault, G. Bompeix, M. Boccara, Transformation of *Botrytis cinerea* with the hygromycin B resistance gene, *hph*, *Curr. Genet.* 26 (1994) 251-255.
- [21] V. Solovyev, P. Kosarev, I. Seledsov, D. Vorobyev, Automatic annotation of eukaryotic genes, pseudogenes and promoters, *Genome Biol.* 7 Suppl 1 (2006) S10.1-12.
- [22] X. Qi, X. Li, J. Zhao, N. He, Y. Li, T. Zhang, S. Wang, L. Yu, Y. Xie, GKK1032C, a new alkaloid compound from the endophytic fungus *Penicillium* sp. CPCC 400817 with activity against methicillin-resistant *S. aureus*, *J. Antibiot. (Tokyo)* 72 (2019) 237-240.
- [23] Genetic Manipulation of the *Fusarium fujikuroi* Fusarin Gene Cluster Yields Insight into the Complex Regulation and Fusarin Biosynthetic Pathway, *Chem. Biol.* 20 (2013) 1055-1066.
- [24] M. Sato, J.E. Dander, C. Sato, Y.-S. Hung, S.-S. Gao, M.-C. Tang, L. Hang, J.M. Winter, N.K. Garg, K. Watanabe, Y. Tang, Collaborative Biosynthesis of Maleimide- and Succinimide-Containing Natural Products by Fungal Polyketide Megasyntases, *J. Am. Chem. Soc.* 139 (2017) 5317-5320.

- [25] Y. Li, K.J. Weissman, R. Müller, Myxochelin Biosynthesis: Direct Evidence for Two- and Four-Electron Reduction of a Carrier Protein-Bound Thioester, *J. Am. Chem. Soc.* 130 (2008) 7554-7555.
- [26] H. Zhang, V. Hantke, P. Bruhnke, E.J. Skellam, R.J. Cox, Chemical and Genetic Studies on the Formation of Pyrrolones During the Biosynthesis of Cytochalasans, *Chem. – Eur. J.* 27 (2021) 3106-3113.
- [27] V. Hantke, E.J. Skellam, R.J. Cox, Evidence for enzyme catalysed intramolecular [4+2] Diels–Alder cyclization during the biosynthesis of pyrivalasin H, *Chem. Commun.* 56 (2020) 2925-2928.
- [28] P.D. Bartlett, R. Banavali, Spontaneous oxygenation of cyclic olefins. Effects of strain, *J. Org. Chem.* 56 (1991) 6043-6050.
- [29] K. Fujiyama, N. Kato, S. Re, K. Kinugasa, K. Watanabe, R. Takita, T. Nogawa, T. Hino, H. Osada, Y. Sugita, S. Takahashi, S. Nagano, Molecular Basis for Two Stereoselective Diels–Alderase that Produce Decalin Skeletons\*\*, *Angew. Chem. Int. Ed.* 60 (2021) 22401-22410.
- [30] M. Sato, S. Kishimoto, M. Yokoyama, C.S. Jamieson, K. Narita, N. Maeda, K. Hara, H. Hashimoto, Y. Tsunematsu, K.N. Houk, Y. Tang, K. Watanabe, Catalytic mechanism and endo-to-exo selectivity reversion of an octalin-forming natural Diels–Alderase, *Nat. Catal.* (2021).
- [31] C. Chi, Z. Wang, T. Liu, Z. Zhang, H. Zhou, A. Li, H. Jin, H. Jia, F. Yin, D. Yang, M. Ma, Crystal Structures of Fsa2 and Phm7 Catalyzing [4 + 2] Cycloaddition Reactions with Reverse Stereoselectivities in Equisetin and Phomasetin Biosynthesis, *ACS Omega* (2021) 10.
- [32] Y.V. Guillén Schlippe, L. Hedstrom, A twisted base? The role of arginine in enzyme-catalyzed proton abstractions, *Arch. Biochem. Biophys.* 433 (2005) 266-278.
- [33] J. Friedt, F.M.V. Leavens, E. Mercier, H.-J. Wieden, U. Kothe, An arginine-aspartate network in the active site of bacterial TruB is critical for catalyzing pseudouridine formation, *Nucleic Acids Res.* 42 (2014) 3857-3870.
- [34] T. Hirokawa, S. Boon-Chieng, S. Mitaku, SOSUI: classification and secondary structure prediction system for membrane proteins., *Bioinformatics* 14 (1998) 378-379.
- [35] M. Bernhofer, C. Dallago, T. Karl, V. Satagopam, M. Heinzinger, M. Littmann, T. Olenyi, J. Qiu, K. Schütze, G. Yachdav, H. Ashkenazy, N. Ben-Tal, Y. Bromberg, T. Goldberg, L. Kajan, S. O'Donoghue, C. Sander, A. Schafferhans, A. Schlessinger, G. Vriend, M. Mirdita, P. Gawron, W. Gu, Y. Jarosz, C. Trefois, M. Steinegger, R. Schneider, B. Rost, PredictProtein - Predicting Protein Structure and Function for 29 Years, *Nucleic Acids Res.* 49 (2021) W535-W540.
- [36] S. Wang, W. Li, S. Liu, J. Xu, RaptorX-Property: a web server for protein structure property prediction, *Nucleic Acids Res.* 44 (2016) W430-435.
- [37] M.M. Attwood, A. Krishnan, V. Pivotti, S. Yazdi, M.S. Almén, H.B. Schiöth, Topology based identification and comprehensive classification of four-transmembrane helix containing proteins (4TMs) in the human genome, *BMC Genomics* 17 (2016) 268.
- [38] G. Shen, The catalytic mechanism of vitamin K epoxide reduction in a cellular environment, 10.
- [39] S. Liu, S. Li, G. Shen, N. Sukumar, A.M. Krezel, W. Li, Structural basis of antagonizing the vitamin K catalytic cycle for anticoagulation, *Science* 371 (2021) eabc5667.
- [40] K. Inaba, K. Ito, Structure and mechanisms of the DsbB-DsbA disulfide bond generation machine, *Biochim. Biophys. Acta* 1783 (2008) 520-529.
- [41] G. Malojčić, R.L. Owen, J.P.A. Grimshaw, R. Glockshuber, Preparation and structure of the charge-transfer intermediate of the transmembrane redox catalyst DsbB, *FEBS Lett.* 582 (2008) 3301-3307.
- [42] X. Guo, H. Mayr, Quantification of the Ambident Electrophilicities of Halogen-Substituted Quinones, *J. Am. Chem. Soc.* 136 (2014) 11499-11512.
- [43] R. Gasper, I. Effenberger, P. Kolesinski, B. Terlecka, E. Hofmann, A. Schaller, Dirigent Protein Mode of Action Revealed by the Crystal Structure of AtDIR6, *Plant Physiol.* 172 (2016) 2165-2175.
- [44] K.-W. Kim, C.A. Smith, M.D. Daily, J.R. Cort, L.B. Davin, N.G. Lewis, Trimeric Structure of (+)-Pinoresinol-forming Dirigent Protein at 1.95 Å Resolution with Three Isolated Active Sites, *J. Biol. Chem.* 290 (2015) 1308-1318.

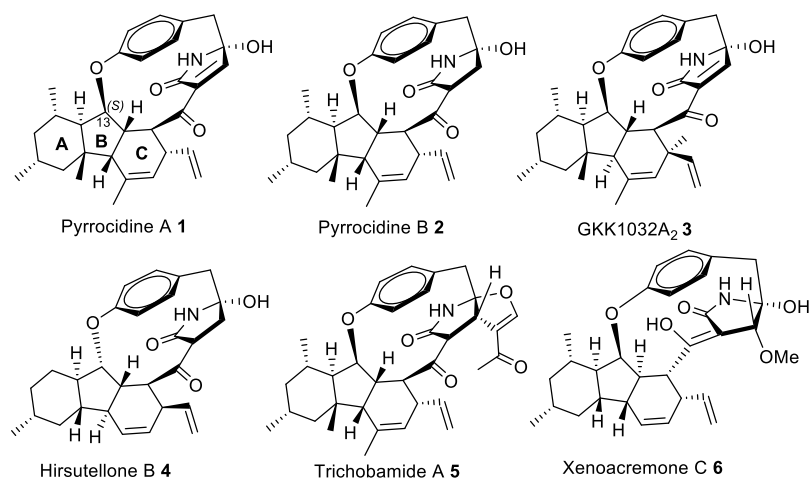


Figure 1: Pyrrocidine structures and related fungal *para*-cyclophane-decahydrofluorene metabolites

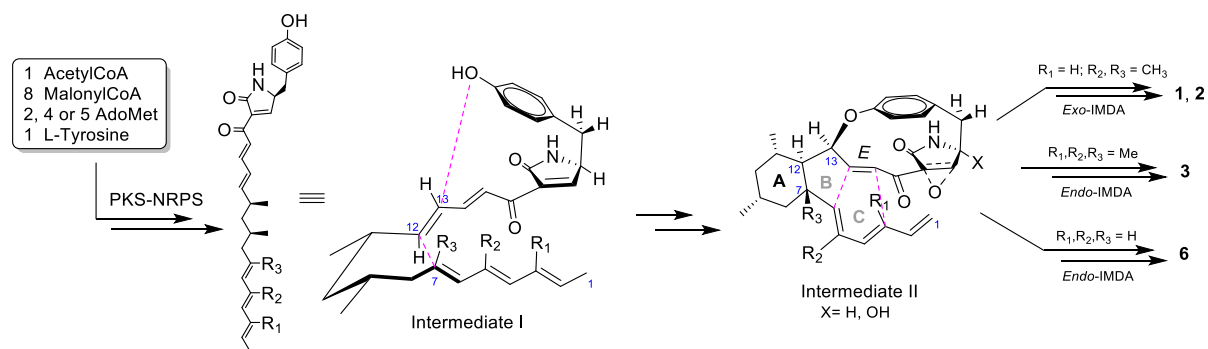


Figure 2: Proposed intermediates from previous publications on the biosynthesis of pyrrocidines, GKK1032A<sub>2</sub> or xenoacremones. The bonds to be formed to access to the final polycyclic metabolites are represented as dashed lines.

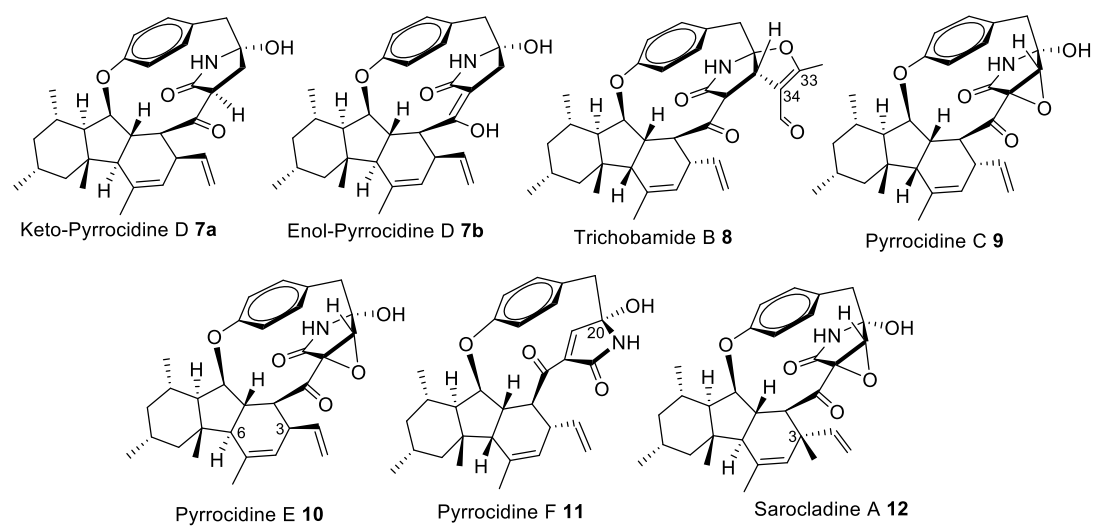
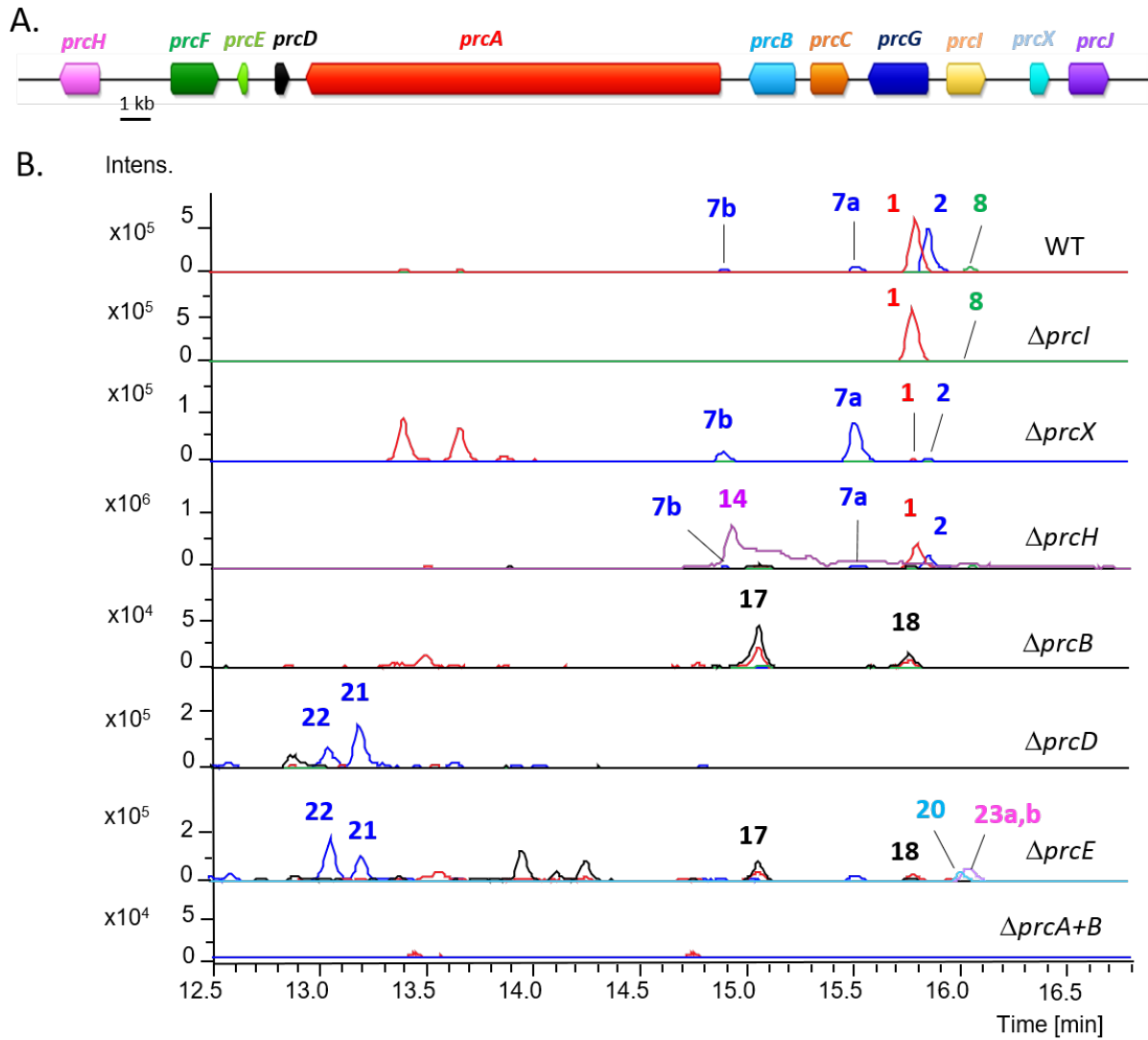


Figure 3: Analogues of pyrrocidines isolated and characterized from *Sarocladium zae* in this study.



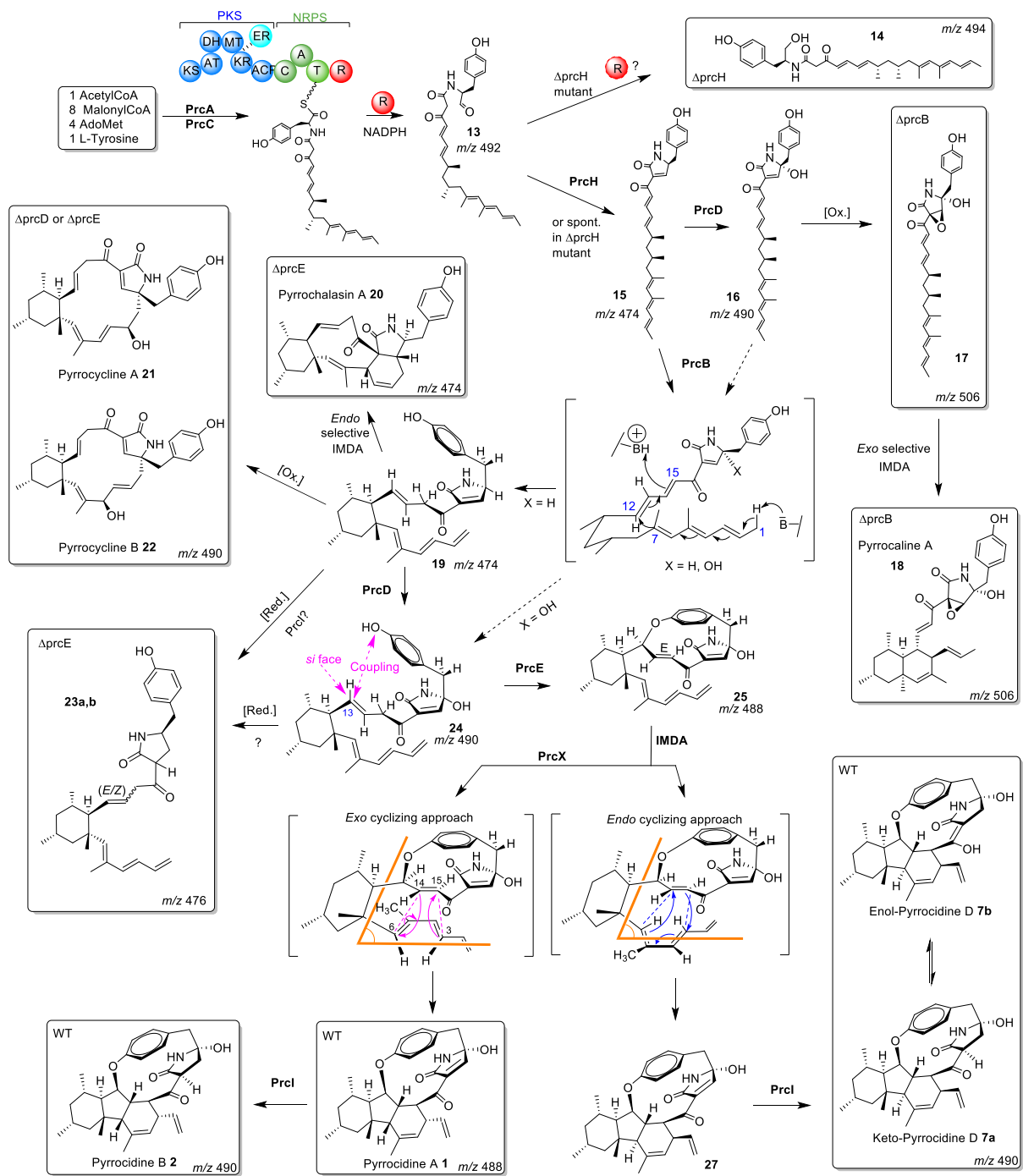


Figure 5: Proposed biosynthesis pathway of pyrrocidines in *Sarocladium zae*. Structures of the metabolites isolated from the wild type and mutant strains are represented in the boxes.

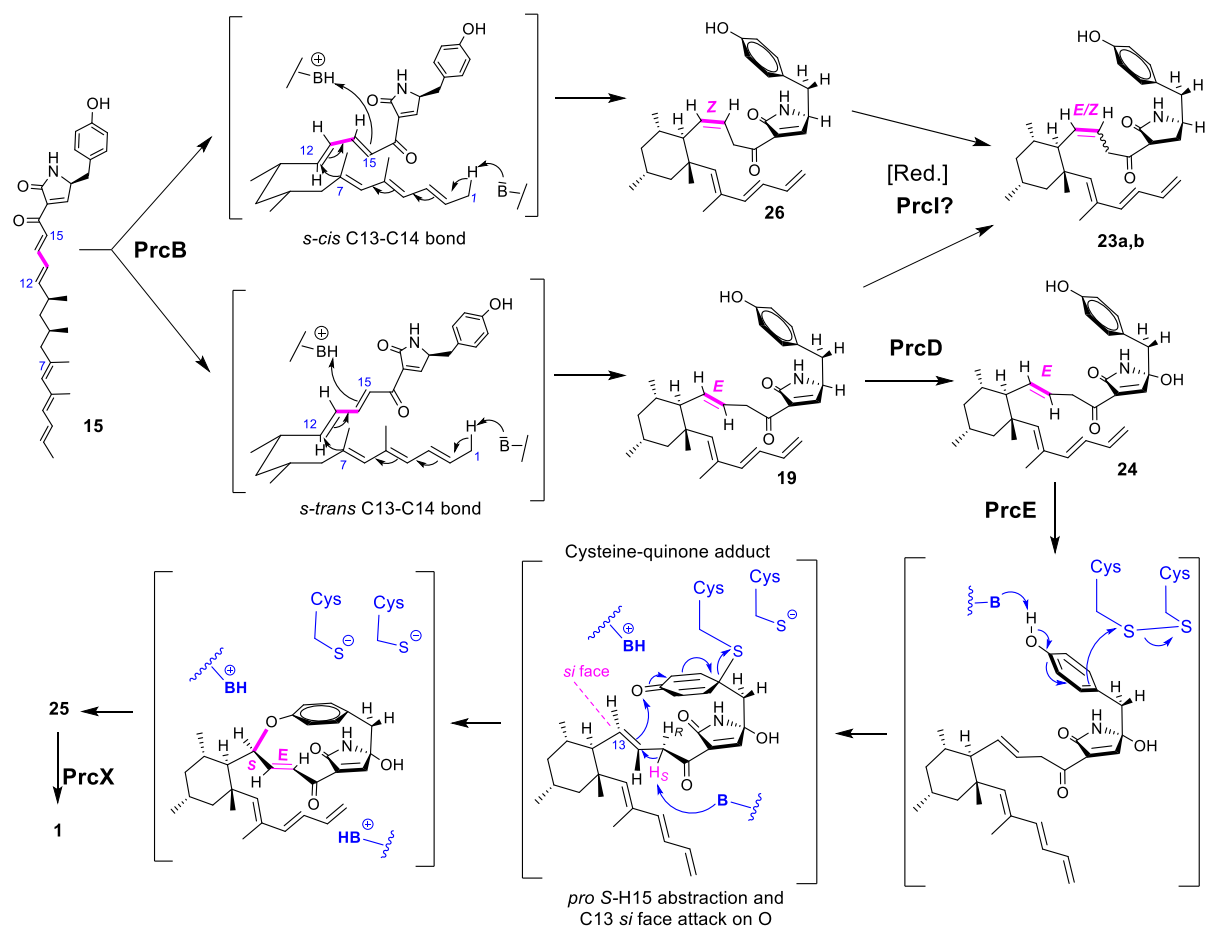


Figure 6: Proposed mechanism for the formation of the paracyclophane-decahydrofluorene moiety of pyrrocidine A **1** involving PrcB,D,E and X. C13-C14  $\sigma$ -bond of intermediate **15** adopts a *s-cis* or *s-trans* conformation in PrcB giving rise to *E*-**19** or *Z*-**26** intermediates after double bond displacement and then to accumulation of **23a,b** in  $\Delta prcE$  mutant after reduction of the double bond in pyrrolidone moiety. Binding of intermediate **24** to PrcE gives rise to a cysteine-quinone adduct. Abstraction of *pro S* H15 triggers the attack of C13 from the *si* face to quinone oxygen.

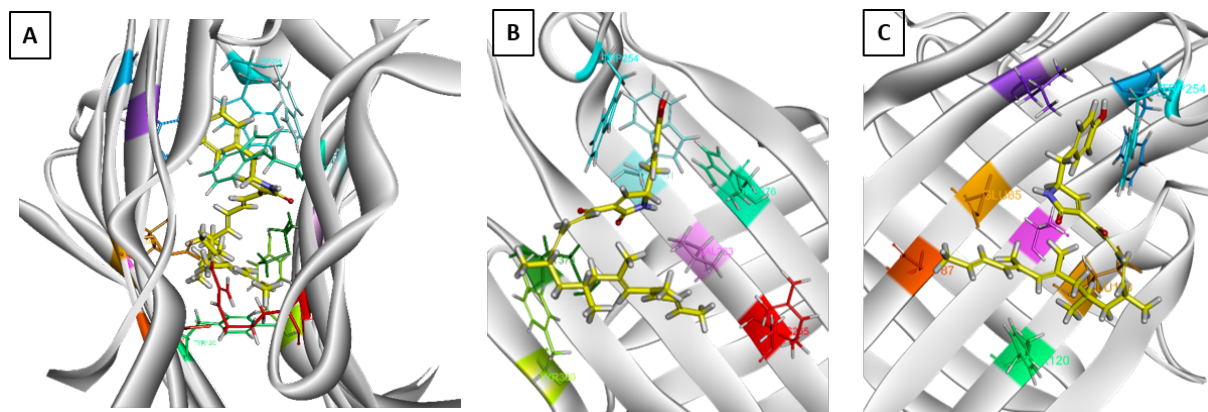


Figure 7: Docking of the intermediate **15** adopting a C14-C15 *s-trans* conformation in modeled PrcB: A. View of **15** (in yellow) in the pocket localized at the interface of both N- and C-terminal domains; B. View of the interactions of **15** with C-terminal domain residues (Arg265 red, Val263 pink, Tyr254 and Tyr261 cyan, Tyr276 spring green, Tyr360 green, Met370 dark green); C. View of the interactions of **15** with the N-terminal domain residues (Trp81 blue, Trp254 cyan, Leu248 violet, Tyr120 green, Asp87 orange, Glu85 ochre yellow, Glu118 brown).



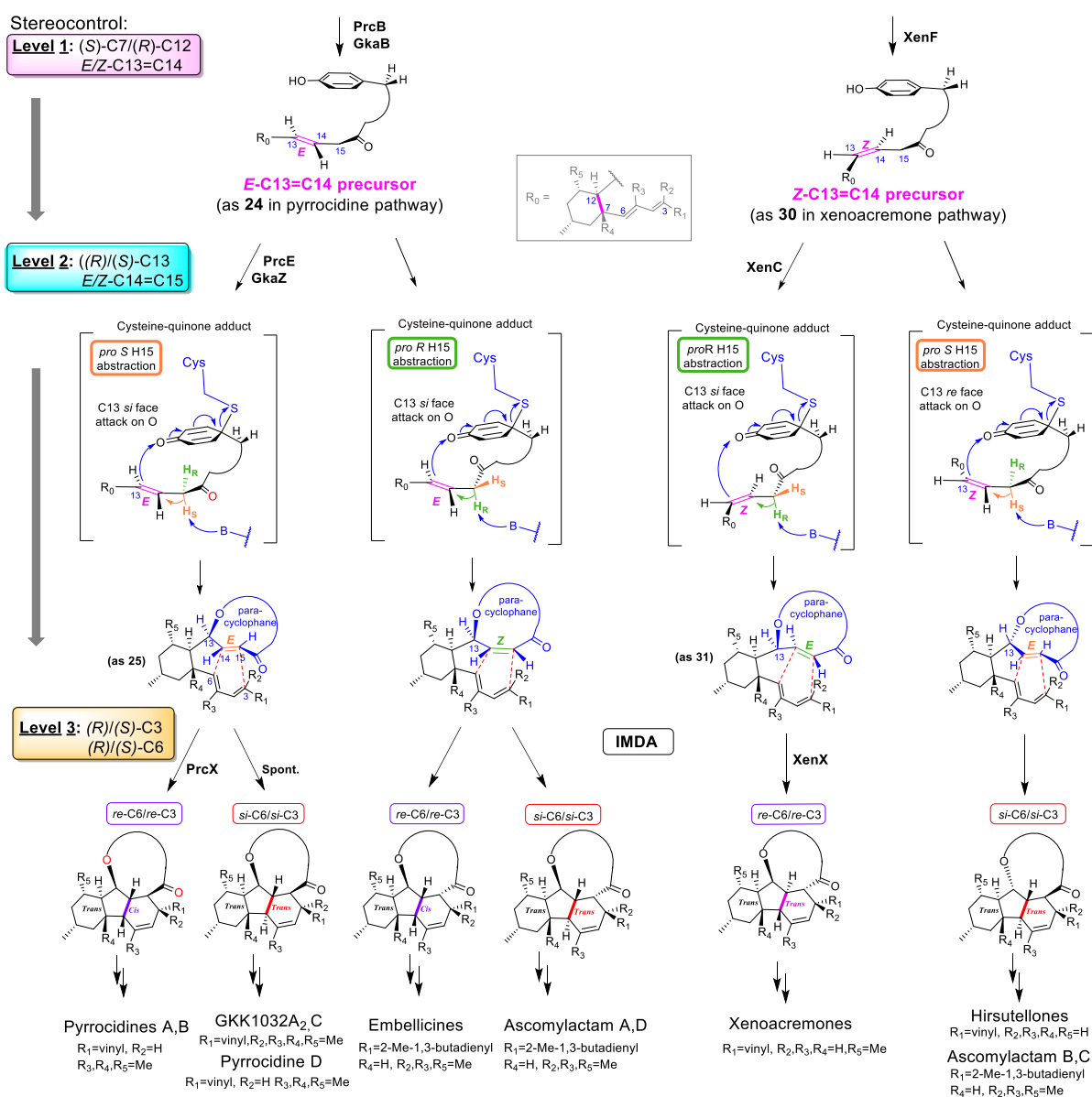


Figure 8: Proposed mechanism for the formation of the paracyclophane-decahydrofluorene motif in the representative natural products showing the three levels of control of the seven stereogenic centers C7, C12, C13, C14, C15, C3 and C6. Known enzymes involved in these reactions are mentioned.



**Titre :** Décryptage des voies biosynthétiques des métabolites dérivés de PKS-NRPS du champignon endophyte *Sarocladium zae*: élucidation de la biosynthèse des pyrrocidines

**Résumé :** Les micro-organismes endophytes ont souvent un rôle de défense de la plante hôte en produisant des métabolites à activité antibactérienne ou antifongique. C'est le cas de *Sarocladium zae*, champignon endophyte du maïs, qui produit des molécules complexes, les pyrrocidines. Ce sont des molécules polycycliques dont le précurseur linéaire, issu d'une mégasynthase PKS-NRPS, est transformé par des enzymes faisant intervenir différents types de cyclisations et d'oxydoréductions. Le génome ayant été séquencé, nous avons pu construire des mutants et déterminer le cluster de gènes de biosynthèse (BGC) impliqué qui comprend, entre autres, six gènes codant pour des enzymes auxiliaires. La culture des différents mutants a permis d'isoler de nouveaux métabolites et bien qu'ils ne soient pas des intermédiaires, nous avons pu proposer une fonction pour chaque enzyme auxiliaire et élaborer un schéma de biosynthèse. Le rôle des pyrrocidines dans l'interaction antagoniste entre *S. zae* et *Fusarium verticillioides* a été évalué et nous avons montré que la production des pyrrocidines ralentit la progression de l'agent pathogène, en accord avec le rôle protecteur de *S. zae* dans le maïs. Des essais de reconstitutions de la voie de biosynthèse chez la levure ont été initiés. Les constructions génétiques ont été réalisées mais la production des métabolites escomptés est restée sans succès. Deux autres clusters comprenant chacun un gène codant une PKS-NRPS ont été identifiés dans le génome de *S. zae*, mais aucun métabolite n'a pu être lié, ni par une étude d'inactivation des gènes ni par surexpression des facteurs de transcription. Ces BGC sont silencieux dans les conditions de culture de laboratoire.

**Mots clés :** Métabolisme secondaire, paracyclophane, decahydrofluorène, enzyme, Diels-Alderase, cyclase.

---

**Title:** Deciphering the biosynthetic pathways of PKS-NRPS derived metabolites from the endophytic fungus *Sarocladium zae*: Elucidation of pyrrocidines biosynthesis

**Abstract:** endophytic microorganisms often play a role in the defense of their host plant by producing metabolites with antifungal or antibiotic activities. This is the case of the maize endophytic fungus *Sarocladium zae* which produces complex molecules named pyrrocidines. A linear precursor of these polycyclic compounds is biosynthesized by a PKS-NRPS megaenzyme and transformed by auxiliary enzymes catalyzing various types of cyclizations and redox reactions. From the sequenced fungal genome, we could generate gene knock-out mutants and identify the biosynthetic gene cluster (BGC) of pyrrocidines including six genes coding for auxiliary enzymes. The culture of the different mutants allowed us to isolate new metabolites and, although these molecules are not biosynthetic intermediates, we could propose a function to each auxiliary enzymes and establish the biosynthetic pathway. The role of pyrrocidines in the antagonistic interaction between *S. zae* and *Fusarium verticillioides* have been assessed and we showed that pyrrocidine production slow down the growth of the pathogen, which is in good agreement with the reported protective effect of *S. zae* on maize. Reconstitution of the first steps of the biosynthesis were initiated in yeast. The genetic constructions were performed but the production of the expected metabolites remained unsuccessful. Two other BGC involving a PKS-NRPS were identified in the genome of *S. zae* and studied. However, neither gene inactivation approach nor transcription factors overexpression strategy allowed us to link a metabolite. This BGC keep silent in laboratory culture conditions.

**Key words:** Secondary metabolism, paracyclophane, decahydrofluoren, enzyme, Diels-Alderase, cyclase.