Mapping of secondary metabolites to their synthase genes in Aspergillus species

Klejnstrup, Marie Louise; Larsen, Thomas Ostenfeld; Gotfredsen, Charlotte Held; Mortensen, Uffe Hasbro

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Mapping of secondary metabolites to their synthase genes in *Aspergillus* species

Marie Louise Klejnstrup
PhD Thesis
September 2012

Center for Microbial Biotechnology
Department of Systems Biology
Mapping of secondary metabolites to their synthase genes in *Aspergillus* species

*Marie Louise Klejnstrup*

PhD thesis

Center for Microbial Biotechnology

DTU Systems Biology
Preface

This thesis is submitted to the Technical University of Denmark in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Chemistry. The work has been carried out between June 2009 and June 2012, at the Center for Microbial Biotechnology at the Department of Systems Biology under the main supervision of associate professor Thomas Ostenfeld Larsen (CMB) with associate professors Charlotte Held Gottfredsen (Department of Chemistry, DTU) and Uffe Hasbro Mortensen (CMB) as co-supervisors. The project was funded by DTU, Novozymes A/S and FOBI (Research School for Biotechnology).

First I would like to thank my three supervisors for giving me the opportunity of working on this project and for their guidance and enthusiasm for the last three years. It has been a pleasure.

I would also like to thank the molecular biologists I have had the pleasure of working with for the past three years. Starting with Michael L. Nielsen who introduced me to the world of genetics, Morten T. Nielsen for picking up afterwards and last but not least for Jakob B. Nielsen for his patience with my endless stream of questions and for making evening and weekend work a pleasure.

A special thanks also goes to Christian Rank, Anita Iversen, Richard K. Phipps, Kristian F. Nielsen and Maria Månsson for taking time for discussions and helpful suggestions in the daily work.

Another thank goes to the technical staff at CMB, Hanne Jakobsen for running MS samples, Jesper Mogensen, Anne Hector and Ellen Kirstine Lyhne for helping when needed and especially Lisette Knoth-Nielsen for being an invaluable help in the lab.

During my PhD I had the privilege to spend three months at the University of Southern California with associate professor Clay C.C. Wang’s group. I would like to thank Clay for letting me come and work at his lab and enjoy the sun of California. I also own the rest of his group a big thanks but especially CJ, Mike, Ken and Jim for helping me get settled and guiding me in the world of making mutants. I would also like to thank the Oticon and Otto Månsted foundations financial support.

A big thanks also goes out to all my former and present officemates; Tanja, Richard, Anita, Olivera, Andreas, Christian, Lucas and Morten for creating a great atmosphere. A special thanks also goes out to the rest of CMB for the good times and to Paiman for all the trips to Mexichar and bringing strawberry cakes when needed.

Credit is also due to the Carlsberg laboratories for allowing me to use their 800 MHz NMR spectrometer.

Lastly I would like to thank my family and friends for their support over the past three years.
Summary

Secondary metabolites have for centuries been used as drugs and inspiring structural leads in the pursuit for new potent drugs. The main sources for secondary metabolites are marine and terrestrial plants, bacteria, and filamentous fungi. During the past decade the genomes of several filamentous fungi and bacteria have been sequenced opening new possibilities for further analysis of the secondary metabolite potential of the organisms, and directing the focus of natural product chemists towards the elucidation of biosynthetic pathways of the metabolites. Thereby, aiding in identifying the genes encoding for the enzymes responsible for constructing these complex molecules.

The aim of this Ph.D. study has been to explore the secondary metabolite potential of important filamentous fungi and link specific metabolites to their synthase genes and biosynthetic pathways. Throughout this study three important filamentous fungi have been examined; Aspergillus oryzae, A. niger and A. nidulans. A. oryzae and A. niger are used in the fermentation industry for production of various products such as sake, miso and soy sauce from A. oryzae and organic acids as well as enzymes for A. oryzae. At the same time A. niger is also an important contaminant of crops and feed due to the possibility of mycotoxins or other metabolites ending up in human and animal food. A. nidulans is a model organism for genetic studies and vast efforts have been used in the development of a molecular platform used for studying the genetic of the organism.

The work with A. oryzae and A. niger focused on discovery of the secondary metabolite potential of these two filamentous fungi. Several new metabolites were isolated and structure elucidated from A. oryzae including; 13-desoxypaxiline a tremorgenic intermediate of the aflatrem biosynthesis seen in A. flavus, diode- and 14-deacetyl parasiticicolide A, and the novel non-ribosomal peptides ditryptoleucine and oryzamide A_1, which are variants of known A. flavus metabolites. From a strain of A. niger a new isomer of the mycotoxin group of fumonisins, fumonisin B_6 was isolated and characterized through NMR-spectroscopy.

Due to A. nidulans' role as a platform for genetic studies emphasis was put on identifying and linking metabolites to their synthase genes. To identify novel metabolites of A. nidulans several attempts, both chemical and genetic, were made to activate the secondary metabolites production. The study was initiated with the creation of a deletion library of 32 putative polyketide synthase genes encoding for the enzymes responsible for catalyzing the production polyketides. To initiate secondary metabolite production these strains were cultivated on several complex media. The study linked austinol and dehydroaustinol to their synthase gene as well as providing new insights into two already partly identified biosynthetic pathways; monodictyphenone and orsellinic acid.

To enhance secondary metabolite production in A. nidulans overexpression of a transcription factor belonging to a polyketide gene cluster identified in the deletion library resulted in the isolation of several metabolites including cichorine, nidulol, 4-hydroxy-3,6-dimethyl-2-pyrone and the novel metabolites cichonidulol and demethylcichonidulol. A deletion study of the surrounding genes led to a proposed biosynthetic pathway towards these 3-methylorsellinic acid derived metabolites.
Summary

A study which included another strain of *A. nidulans* and the deletion of a global regulator gene challenges on complex media resulted in the isolation of asperugin A and B, which through a deletion study were linked to their synthase gene, and nidubenzal A and B. The study ends with a report on the production of the insect juvenile hormones in *A. nidulans* provoked by the heterologous expression of a regulatory gene of *A. niger*.

In conclusion this PhD study has added to the knowledge of the chemical diversity of three important filamentous fungi through activation of silent genes through the use of complex media and genome modifications.
Sammenfatning

Sekundære metabolitter er i løbet af de sidste mange århundrede blevet brugt som medikamenter og som inspiration til syntese af nye biologisk potente stoffer, der kan bruges til ny medicin. Hovedkilderne til sekundære metabolitter i naturen er planter, bakterier og skimmelsvampe både fra land og marint ophav. Genommerne af flere skimmelsvampe og bakterier er gennem det sidste årti blevet sekventeret, hvilket har åbnet op for forskningen i potentialet af sekundære metabolitter i disse organismer. Dette har for naturstofkemikere medført et øget fokus på udredning af biosyntesen for disse metabolitter, og dertil identificering af generne som koder for de enzymer, der katalyserer konstruktionen af disse komplekse metabolitter.


I forsøget på at øge produktionen af metabolitter i en interessant biosyntese gengruppe i A. nidulans, blev et regulatorisk gen overudtrykt. Dette medførte en øget produktion af flere metabolitter inklusiv cichorine, nidulol, 4-hydroxy-3,6-dimethyl-2-pyrene og de nye metabolitter cichonidulol og demethylcichonidulol. Ved hjælp af en stor serie gene deletioner, og efterfølgende kemisk analyse af mutanterne, kunne vi opstille et forslag til biosyntesen af de 3-methylorsellinsyre afledte metabolitter. Et andet studie af rapporteret i denne afhandling
Sammenfatning

inkluderede et vildtypeisolat af A. nidulans, og her blev de to nye metabolitter nidubenzal A og B isoleret og strukturbestemt. Derudover blev asperugin A og B gennem et nyt gendeletionssstudie, der påvirkede tilgængeligheden i kromatin, koblet til et syntasegen. Afhandlingen slutter med et studie, hvori produktionen af juvenile hormone III blev aktiveret ved heterolog ekspression af et regulatorisk gen fra A. niger kombineret med dyrkning på komplekse medier.

Dette PhD studie har bidraget til viden om den kemiske diversitet af tre, på hver sin måde vigtige, skimmelsvampe, og arbejdet demonstrerer hvorledes avanceret kemisk analyse og effektive gensplejsningsværktøj sammenspiller til at kortlægge den fascinerende sekundære metabolisme i skimmelsvampe.
List of papers and other publications

Paper 1  ML Kleijnstrup, RJN Frandsen, DK Holm, MT Nielsen, UH Mortensen, TO Larsen, JB Nielsen; “Genetics of polyketide metabolism in Aspergillus nidulans”, Metabolites 2, 100-133 (2012).

Paper 2  C Rank and ML Kleijnstrup (Joint 1st author), LM Petersen, S Kildgaard, JC Frisvad, CH Gottfredsen, TO Larsen; “Comparative chemistry of Aspergillus oryzae (RIB40) and A. flavus (NRRL 3357)”, Metabolites 2, 39-56 (2012).


Paper 7  MT Nielsen and ML Kleijnstrup (Joint 1st author), CH Gottfredsen, MR Andersen, BG Hansen, UH Mortensen, TO Larsen; “A regulatory protein from Aspergillus niger induces juvenile hormones upon heterologous expression in A. nidulans”, Submitted to PLoS ONE – under revision.

Conference proceeding:

ML Kleijnstrup, MT Nielsen, JC Frisvad, UH Mortensen, TO Larsen; A combined genetic and multi medium approach reveals new secondary metabolites in Aspergillus nidulans. Poster at Eight International Aspergillus meeting Asperfest 8 and 26th fungal genetics conference, Asilomar, California, USA. March 14th – 20th 2011.
List of papers and other publications
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP</td>
<td>Acyl Carrier Protein</td>
</tr>
<tr>
<td>AT</td>
<td>acyltransferase</td>
</tr>
<tr>
<td>BPC</td>
<td>Base Peak Chromatogram</td>
</tr>
<tr>
<td>CLC</td>
<td>Claisen Cyclase</td>
</tr>
<tr>
<td>CMB</td>
<td>Center for Microbial Biotechnology</td>
</tr>
<tr>
<td>CY20</td>
<td>Czapek Yeast sucrose</td>
</tr>
<tr>
<td>CYA</td>
<td>Czapek Yeast Agar</td>
</tr>
<tr>
<td>CYAs</td>
<td>Czapek Yeast Agar salt</td>
</tr>
<tr>
<td>DAD</td>
<td>Diode Array Detection</td>
</tr>
<tr>
<td>DH</td>
<td>Dehydratase</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DQF-COSY</td>
<td>Double-quantum filtered correlated spectroscopy</td>
</tr>
<tr>
<td>DTU</td>
<td>Technical University of Denmark</td>
</tr>
<tr>
<td>ESI</td>
<td>Extracted Ion Chromatogram</td>
</tr>
<tr>
<td>ER</td>
<td>Enoyl Reductase</td>
</tr>
<tr>
<td>H2BC</td>
<td>Heteronuclear two bond correlation</td>
</tr>
<tr>
<td>HMBC</td>
<td>Heteronuclear multiple bond correlation</td>
</tr>
<tr>
<td>HR</td>
<td>Highly Reduced</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear single quantum coherence</td>
</tr>
<tr>
<td>IPKS</td>
<td>iterative Polyketide synthase</td>
</tr>
<tr>
<td>KR</td>
<td>β-ketoreductase</td>
</tr>
<tr>
<td>KS</td>
<td>β-ketosynthase</td>
</tr>
<tr>
<td>MM</td>
<td>Minimal medium</td>
</tr>
<tr>
<td>MS</td>
<td>Mass Spectrometry</td>
</tr>
<tr>
<td>MS/MS</td>
<td>Tandem Mass Spectrometry</td>
</tr>
<tr>
<td>MT</td>
<td>Methyl transferase</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>NR</td>
<td>Non-Reduced</td>
</tr>
<tr>
<td>NRP</td>
<td>Non-Ribosomal Peptide</td>
</tr>
<tr>
<td>NRPS</td>
<td>Non-Ribosomal Peptide Synthetase</td>
</tr>
<tr>
<td>Oex</td>
<td>Overexpression</td>
</tr>
<tr>
<td>OSMAC</td>
<td>One Strain Many Compounds</td>
</tr>
<tr>
<td>PK</td>
<td>Polyketide</td>
</tr>
<tr>
<td>PKS</td>
<td>Polyketide Synthase</td>
</tr>
<tr>
<td>PR</td>
<td>Partially Reduced</td>
</tr>
<tr>
<td>PT</td>
<td>Product Template</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>quantitative Reverse Transcription Polymerase Chain Reaction</td>
</tr>
<tr>
<td>RT</td>
<td>Raulin-Thom</td>
</tr>
<tr>
<td>RTO</td>
<td>Raulin-Thom Oatmeal</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosylmethionine</td>
</tr>
<tr>
<td>SAT</td>
<td>Starter-unit ACP Transacylase</td>
</tr>
<tr>
<td>TE</td>
<td>Thioesterase</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>TOF</td>
<td>Time of Flight</td>
</tr>
<tr>
<td>UHPLC</td>
<td>Ultra High Pressure Liquid Chromatography</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra Violet</td>
</tr>
<tr>
<td>YE</td>
<td>Yeast Extract</td>
</tr>
<tr>
<td>YES</td>
<td>Yeast Extract Sucrose</td>
</tr>
</tbody>
</table>
# Table of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>I</td>
</tr>
<tr>
<td>Summary</td>
<td>III</td>
</tr>
<tr>
<td>Sammenfatning</td>
<td>IV</td>
</tr>
<tr>
<td>List of original papers and other publications</td>
<td>VII</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>IX</td>
</tr>
<tr>
<td>Table of contents</td>
<td>XI</td>
</tr>
<tr>
<td><strong>1 Background</strong></td>
<td></td>
</tr>
<tr>
<td>2 Production of secondary metabolites in filamentous fungi</td>
<td>3</td>
</tr>
<tr>
<td>2.1 Biosynthesis of secondary metabolites</td>
<td>3</td>
</tr>
<tr>
<td>2.1.1 Polyketides</td>
<td>3</td>
</tr>
<tr>
<td>2.1.2 Non-ribosomal peptides</td>
<td>5</td>
</tr>
<tr>
<td>2.1.3 Terpenoids</td>
<td>7</td>
</tr>
<tr>
<td>2.1.4 Hybrid metabolites</td>
<td>8</td>
</tr>
<tr>
<td>2.2 Genetics of secondary metabolite biosynthesis</td>
<td>9</td>
</tr>
<tr>
<td>2.2.1 Prediction of gene function</td>
<td>9</td>
</tr>
<tr>
<td>2.2.2 Activation of silent genes</td>
<td>10</td>
</tr>
<tr>
<td>2.2.2.1 One strain – many compounds</td>
<td>11</td>
</tr>
<tr>
<td>2.2.2.2 Discovery of secondary metabolites through genome mining</td>
<td>12</td>
</tr>
<tr>
<td>2.2.2.3 Other methods and combinations of activation strategies</td>
<td>14</td>
</tr>
<tr>
<td><strong>3 Overview over experimental work</strong></td>
<td>17</td>
</tr>
<tr>
<td>3.1 Linking of secondary metabolites to gene</td>
<td>17</td>
</tr>
<tr>
<td>3.1.1 Targeted genetic manipulations including deletions and</td>
<td>18</td>
</tr>
<tr>
<td>overexpressions</td>
<td></td>
</tr>
<tr>
<td>3.1.2 Metabolome analysis</td>
<td>18</td>
</tr>
<tr>
<td>3.1.3 Isolation and structure elucidation</td>
<td>18</td>
</tr>
</tbody>
</table>
Table of contents

4  Results and discussion 19
4.1  *Aspergillus oryzae* 19
4.1.1  Comparison of secondary metabolite profiles of *A. oryzae* and *A. flavus* 20
4.1.1.1  13-desoxypaxilline 22
4.1.1.2  Parasiticolides 23
4.1.1.3  Ditypoteleucine and oryzamide A$_{1,2}$ 24
4.1.2  Summary and part conclusion 26
4.2  *Aspergillus niger* 26
4.2.1  Fumonisin production in *A. niger* 27
4.2.1.1  Isolation and NMR characterization of fumonisin B$_2$ and B$_6$ from *A. niger* 27
4.2.3  Summary and part conclusion 29
4.3  *Aspergillus nidulans* 29
4.3.1  Discovery and linking of polyketides to their PKS through a deletion library 29
4.3.1.1  Identification of the austinol/dehydroaustinol synthase gene 32
4.3.1.2  Linking of arugosin to the monodictyphenone gene cluster 34
4.3.1.3  Linking of orsellinic acid derived metabolites to the orsellinic acid gene cluster and *AN7903* 36
4.3.2  Characterization of the 3-methylorsellinic acid gene cluster 37
4.3.3  Genetic activation with regulatory genes combined with complex media 51
4.3.3.1  Identification and linking of *A. nidulans* metabolites to genes through a combined approach of complex media and genetic modifications 52
4.3.3.2  Activation of juvenile hormone production in *A. nidulans* 60
4.3.4  Summary and part conclusion 63
5  Overall discussion and conclusion 65
6  References 71

Papers 1-7


1 Background

Natural products (or secondary metabolites) have for centuries been used as drugs or inspiring structural leads in the pursuit for new potent drugs. The main sources for secondary metabolites are marine and terrestrial plants, bacteria, and filamentous fungi. Secondary metabolites are classified depending on their structural scaffold. These classes include among others polyketides (PKs), non-ribosomal peptides (NRPs), terpenoids, alkaloids and hybrid metabolites of the different classes. The function of secondary metabolites in their hosts has not yet been exploited to its fullest; however, they seem to have several functions including defense, communication, and signaling (Calvo 2002, Fox 2008, Kempken 2010, Rohlf 2010). Secondary metabolites have, especially after the discovery of the penicillins in 1929 (Fleming 1929), been used as inspiration in the discovery of new pharmaceutical candidates. In fact, a study by Newman and Cragg showed that 34 % of all small-molecule new chemical entries between 1981 and 2010 were natural products or derived thereof and further 25 % were mimics of natural products (Newman 2012).

The filamentous fungi within the Aspergillus family are important due to both their vast uses as industrial workhorses (A. niger) (Schuster 2002) and as model organisms for genetic studies (A. nidulans) (Pontecorvo 1953, Chiang 2010a), but also due to their prevalence as food contaminants (A. niger) (Perrone 2007, Nielsen 2009). A. niger is both an industrial workhorse for the production of citric acid and various enzymes (Schuster 2002) as well as a contaminant of food and feeds on which it can produce the mycotoxins ochratoxin and the fumonisins (Perrone 2007, Nielsen 2009). The Aspergillus section includes other mycotoxin and drug producing species such as A. flavus (aflatoxin) and A. terreus (lovastatin (mevinolin)), respectively (Hoffmeister 2007).

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. aculeatus</td>
<td>ATCC 16872</td>
<td>DOE Joint Genome Institute 2012a</td>
</tr>
<tr>
<td>A. awamori</td>
<td>IFO 4308</td>
<td>NCBI 2012a</td>
</tr>
<tr>
<td>A. carbonarius</td>
<td>ITEM 5010</td>
<td>DOE Joint Genome Institute 2012b</td>
</tr>
<tr>
<td>A. clavatus</td>
<td>NRRL 1</td>
<td>Fedorova 2008</td>
</tr>
<tr>
<td>A. fischerianus</td>
<td>NRRL 181</td>
<td>Fedorova 2008</td>
</tr>
<tr>
<td>A. flavus</td>
<td>NRRL 3357</td>
<td>Center for Integrated Fungal Research 2012</td>
</tr>
<tr>
<td>A. fumigatus</td>
<td>A293</td>
<td>Nierman 2005</td>
</tr>
<tr>
<td>A. fumigatus</td>
<td>A1163</td>
<td>Fedorova 2008</td>
</tr>
<tr>
<td>A. nidulans</td>
<td>FGSC A4</td>
<td>Galagan 2005</td>
</tr>
<tr>
<td>A. niger</td>
<td>CBS 513.88</td>
<td>Pel 2007</td>
</tr>
<tr>
<td>A. niger</td>
<td>ATCC 1015</td>
<td>Baker 2006, Andersen 2011</td>
</tr>
<tr>
<td>A. oryzae</td>
<td>RIB40</td>
<td>Machida 2005</td>
</tr>
<tr>
<td>A. sojae</td>
<td>NBRC 4239</td>
<td>NCBI 2012b</td>
</tr>
<tr>
<td>A. terreus</td>
<td>NIH 2624</td>
<td>Broad Institute 2012</td>
</tr>
</tbody>
</table>

During the past decade, the genome of several Aspergillus species have been fully sequenced and become available for the public, table 1.1. The availability of genome sequences have led to a shift in natural product research from discovery of new drug-like metabolites to linking metabolites to the responsible genes as well as further elucidation of biosynthetic pathways (Walsh 2010). Detailed information on the genetic and enzymatic work in an organism enable the possibility of

Genome sequences have, through bioinformatics studies, revealed that the number of secondary metabolite synthase genes far exceeds the number of metabolites known. For example, A. nidulans appears to contain as many as 32 polyketide synthases (PKSs) of which, until recently, only nine gene clusters have been identified and linked to metabolites (paper 1). The information from the bioinformatic studies reveals that the secondary metabolite potential by far exceeds the known secondary metabolite pool. This huge, untapped potential for secondary metabolite production has led to collaborations between molecular biologists and natural product chemists to explore the full potential of filamentous fungi through genome mining, optimized chemical screening, and other advanced techniques towards activation of silent genes (Sanchez 2012).

The aim of this Ph.D. study has been to explore the secondary metabolite potential of important filamentous fungi, to link specific metabolites to their synthase genes and elucidate the biosynthetic pathways. The work presented in the following chapters includes exploring the chemistry of A. oryzae (paper 2) by comparison of the metabolic profiles to A. flavus and isolation and characterization of several new metabolites. This is followed by isolation and characterization of a new fumonisin analogue isolated from A. niger (paper 3). To discover the hidden secondary metabolic potential of A. nidulans attempts have been performed to activate the silent genes through both media variation and genetic modifications. The activation of genes has led to the identification of several new metabolites being identified, isolated and structure elucidated (paper 4-7). Several of these metabolites have been linked to their biosynthetic genes. An overview over the genes which at the time had been linked to polyketides in A. nidulans is given in paper 1.

Altogether this thesis consists of 6 chapters where chapter 1 gives background information for the Ph.D. study. Chapter 2 gives a general introduction to production of secondary metabolites in filamentous fungi and ways of activating silent genes. Chapter 3 gives an overview of the experimental work and a short introduction to experimental procedures used throughout the work. In chapter 4 the results are presented and discussed and chapter 5 is an overall discussion and conclusions of the work presented. The references are in chapter 6. The papers presented in this thesis are in the accompanying appendix (paper 1-7) where paper 5 and 6 are rough drafts.
2 Production of secondary metabolites in filamentous fungi

This section will focus on the production of secondary metabolites in filamentous fungi including biosynthesis, regulation and activation of genes.

2.1 Biosynthesis of secondary metabolites

Secondary metabolites are divided into classes according to their biosynthetic origin. The following sections will describe some of the general biosynthetic features of three main classes of secondary metabolites including examples of metabolites from *A. nidulans*. The focus will be on PKs, however, general characteristics on NRPs, terpenoids and examples of hybrid metabolites will be described.

2.1.1 Polyketides

PKs are one of the most important classes of secondary metabolites and are produced by several terrestrial and marine organisms including bacteria, plants and fungi. Though the structures of PKs are highly diverse they share a common biosynthetic origin of the carbon skeleton which is synthesized from small activated carboxylic acids such as acetate, propionate and butyrate. The carbon skeleton is biosynthesized by polyketide synthases (PKSs). PKSs can be divided into three types, type I, II, and III based on their catalytic organization; however, only the iterative type I PKS (iPKS) has been reported in *A. nidulans*. The iPKS repeat the use of a single module, containing several catalytic domains, until the growing chain has reached the desired length.

Fungal PKSs are generally type I iPKSs. The fungal iPKSs can, based on their catalytic domains, be divided into three classes: non-reducing (NR), highly reducing (HR) and partially reducing (PR) (Bingle 1999, Nicholson 2001, Kroken 2003). The difference between the three classes reflects their ability to reduce the β-keto carbon of the growing carbon chain. The reduction occurs in the catalytic domains of the iPKS-enzymes.

![Figure 2.1 Chain extension of the PK chain in iPKSs through Claisen-like condensations of the extender and starter unit. The synthesis is started by binding a starter unit, here acetyl-CoA, to the KS domain and loading of the extender unit, here malonyl-CoA, to the ACP domain, releasing the CoA. The KS catalyzes the claisen-like condensation extending the polyketide chain by an acetate unit (marked in red) whereby CO₂ is released. To start another extension the polyketide chain is transferred back to the KS domain and the ACP domain is loaded with a new extender unit. Abbreviations: KS: β-ketosynthase, ACP: Acyl carrier protein.](image)

Three fundamental domains are in general found in all filamentous fungal iPKSs; β-ketosynthase (KS), acyltransferase (AT) and the acyl carrier protein (ACP). The KS catalyzes the C-C bond formation through Claisen-type condensations. The ACP domain is responsible for transiently
2 | Production of secondary metabolites in filamentous fungi

holding the growing acyl chain, hereby allowing the loading of the extender units. AT recognizes
the acyl extender units which are transferred from CoA onto KS and transferred to the growing PK
chain through Claisen-like condensations catalyzed by KS, figure 2.1. The iterative use of the three
domains results in a NR-PK.

In partially reduced and highly reduced PKSs, reduction occurs through the β-ketoreductase (KR)
domain that reduces the β-ketone to a hydroxyl group. Loss of the hydroxyl group can occur by
the elimination of water through the dehydratase (DH) domain which can be followed by
hydrogenation catalyzed by the enoyl reductase (ER), figure 2.2a.

Figure 2.2 Functional domains which may be present in a PKS. The reaction schemes show examples of the
general reactions of the PKS domains a) reduction of ketone to saturation catalyzed by the KR, DH, ER
domains. b) methylation catalyzed by the MT domain. Abbreviations: ACP: Acyl carrier protein, KR:
ketoreductase, DH: dehydratase, ER: enoyl reductase, MT: methyltransferase, SAM: S-adenosylmethionine

Another domain which can be present in the reducing PKSs is the methyltransferase domain (MT)
which is responsible for C-methylation of the growing PK chain, using S-adenosylmethionine
(SAM) as a carbon-donor. The degree of eventual reductions and modifications, and their
positions in the PK product, is always the same for the individual PKSs; however, it is presently
unknown how deployment of the basic modifying domains are programmed into the PKS.

The NR-PKSs do not contain the reducing domains, however, they contain an N-terminal starter
unit-ACP transacylase (SAT) domain, which is responsible for selecting the starter unit (Crawford
2008a), and an internal product template (PT) domain, which is responsible for folding and
cyclization of the NR-PK backbone (Crawford 2008b, Crawford 2009, Crawford 2010). The number
of iterations within the PKS and thereby the display of functional groups and size of the final
product is likely determined by the size of the active site cavity in the iPKS (Yadav 2009). The
iterative nature of the iPKSs makes it difficult to predict the structure of the final product of the
iPKSs both with respect to the degree of saturation, methylation pattern and size; however,
domain swapping experiments by Fisch and co-workers (Fisch 2011) are starting to reveal the
programming of the iPKSs (Fisch 2011).

Once the final length of the PK backbone has been achieved, the PK chain is released from the
PKS, catalyzed by either a thioesterase (TE), a Claisen cyclase (CLC) or by other accessory enzymes
(Du 2010).
The product coming directly from the PKS rarely seems to be the only product in the biosynthesis, but usually undergoes further modifications by tailoring enzymes which can range from small changes (methylations, reductions, oxidations) to large modifications including ring formation and condensations with other metabolites. The PKS genes tend to reside in gene clusters where the genes responsible for further modifications, e.g. genes encoding monooxygenases, O-methyltransferase, dehydrogenases, are present.

For a more detailed description on polyketides than the one given here, reviews by Cox (Cox 2007), Hertweck (Hertweck 2009) and Crawford (Crawford, 2010) can be consulted.

An example of a complex PK biosynthesis found in several Aspergillus species including A. nidulans is sterigmatocystin (Rank 2011). Sterigmatocystin is an intermediate in the biosynthesis of the aflatoxins which are among the most carcinogenic mycotoxins. Research into the biosynthesis and biological function of aflatoxin and sterigmatocystin intensified with the Turkey X disease caused by aflatoxins in the middle of the last century (Blount 1961) and it is therefore one of the most extensively studied biosynthetic pathways and there are several reviews for further reading (Yu 2004, Yabe 2004). The proposed biosynthesis of sterigmatocystin in A. nidulans can be seen in figure 2.3. The gene cluster consists of 27 genes (Brown 1996) where stcA is a NR-PKS which codes for the enzyme that catalyzes the condensation of the starting unit hexanoyl-CoA and seven malonyl-CoA extender units into the PK backbone, which is followed by cyclization and release of norsolorinic acid anthrone (Yu 1995). Several enzymatic steps are required for converting norsolorinic acid anthrone to sterigmatocystin which include oxidations, reductions, O-methylations and Bayer-Villiger oxidation (Yabe 1993, Yabe 1991, Trail 1994, Keller 1995, Kelkar 1996, Silva 1996, Kelkar 1997, Chang 2000, Keller 2000, Watanabe 2002, Sakuno 2003, Chang 2004, Ehrlich 2005, Henry 2005, McDonald 2005, Cary 2006, Ehrlich 2010). For a more detailed description see paper 1.

Even though the biosynthesis of sterigmatocystin has been studied extensively all the biosynthetic steps are still not accounted for. This highlights that the biosynthesis of secondary metabolites can be very complex and difficult to decipher.

2.1.2 Non-ribosomal peptides

NRPs are another large group of secondary metabolites produced in microorganisms. They are made up of amino acids which are linked by a non-ribosomal peptide synthethase (NRPS) through condensations. The NRPSs are single multimodular enzymes where each module includes several catalytic domains. The modules contain the necessary domains; adenylation, peptidyl carrier protein and condensation. The adenylation domain is specific for each amino acid which is recognized, adenylated, and loaded onto the peptidyl carrier protein domain where the condensation domain catalyzes the formation of the peptide bond between the two amino acids. The number of modules corresponds to the number of amino acids in the final product and their order in the metabolite follows the order of the adenylation domains. Other catalytic functions that may be present in an NRPS includes epimerization, peptide cyclization and N-methylation domains. The amino acids are produced from two pathways. The aliphatic amino acids are produced from the central carbon metabolism whereas the aromatic amino acids are derived
from shikimic acid, figure 2.4. Shikimic acid can also be incorporated into secondary metabolites and may be recognized for the substitution pattern of the three hydroxyl-groups on the ring.

The emericellamides, figure 2.7 are examples of secondary metabolites of a combined PKS and NRPS origin produced by A. nidulans. The product of the PKS is loaded onto the NRPS which consist of five modules each responsible for addition of an amino acid (glycine, valine, isoleucine,
alanine and alanine) to the metabolite backbone followed by release and cyclization (Chiang 2008).

Beside NRPS genes there are also NRPS-like genes present in the genomes. These genes share the catalytic domains found in NRPSs but are missing the condensation domain necessary for peptide-bond formation (von Döhren 2009). So far products of two NRPS-like genes in A. nidulans have been identified, terrequinone A (Bok 2006, Balibar 2007, Schneider 2008), which does not have a peptide bond in its structure, and microperfuranone, without a nitrogen atom in its structure (Yeh 2012), figure 2.4.

Figure 2.4 Structures of shikimic acid, microperfuranone and terrequinone A

For reviews on the biosynthesis of NRPs and the shikimic pathway please see reviews by Fischbach (Fischbach 2006), Koglin (Koglin 2009), Du (Du 2010), and Matthias (Matthias 2010) and Dewick and Knaggs respectively (Dewick 1998, Knaggs 2003).

2.1.3 Terpenoids

Terpenoids is a group of natural products made up of activated C5 isoprene units; dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP), which are condensed in a head to tail fashion, figure 2.5, in the terpene synthases. These carbon-chains can then undergo several modifications like cyclization, reduction or oxidations to reach the final product. DMAPP and IPP are derived by the mevalonate (Miziorko 2011) and the methylerythritol phosphate pathways (Rohmer 2008).

Figure 2.5 Structure of an isoprene unit and example of linking of two isoprene units in a head to tail fashion. The two isoprene units in the monoterpane is highlighted in bold.

The biosynthetic potential of terpenoids in A. nidulans has not yet been exploited to its fullest. In fact, only one terpene, ent-pimara-8(14),15-diene, have, so far, been identified and linked to a gene cluster which includes a putative geranylgeranyl-diphosphate (GGPP) synthase, AN1593, and a pimaradiene synthase, AN1594 (figure 2.6) (Broman 2012). The DDPP-synthase is responsible for the formation of geranylgeranyl-PP formation and AN1594 catalyzes the cyclization reactions...
Production of secondary metabolites in filamentous fungi

into the final product (Broman 2012). In the figure the head to tail linking of the isoprene backbone can be seen.

Figure 2.6 Folding and cyclization of the terpene ent-pimara-8(14),15-diene as proposed by Broman and co-workers (Broman 2012). The four isoprene units in geranylgeranyl-PP is highlighted in bold.

For further reading of the details of the biosynthesis of terpenoids please consult review by Dewick (Dewick 2002) and Oldfield (Oldfield 2012).

2.1.4 Hybrid metabolites

Hybrid metabolites are metabolites where different parts of the molecule originate from different biosynthetic routes. In A. nidulans both the emericellamides and the aspyridones, figure 2.7 are hybrid metabolites which have a PK and NRP part they are; however, produced by different types of enzymes. Aspyridone is a PK-NRP which is assembled from the activity of a single fusion enzyme, whereas the biosynthesis of the emericellamides requires a PKS and a NRPS where the product from the PKS acts as the starter unit in the NRPS (Bergmann 2007, Chiang 2008). A further discussion of the biosynthesis of the emericellamides and aspyridones can be found in paper 1.

Figure 2.7 The structures of aspyridone A and B and emericellamide A, C, D, E and F produced by A. nidulans.

Meroterpenoids is another group of natural products of mixed biosynthetic origin where one or more parts are derived from terpenoids. The meroterpenoids can be divided into two main classes, PK-terpenoids and non-PK-terpenoids. As the name indicates this division is based on whether the non-terpenoid part of the molecule is PK derived or derived from other biosynthetic routes, e.g. from the shikimate route (Geris 2009).

In A. nidulans austinol and dehydroaustinol are examples of PK-meroterpenoids. The biosynthesis of 3,5-dimethylorsellinic acid, the PK, is catalyzed by AusA and then prenylated by the
prenyltransferase AusN (Nielsen 2011, Lo 2012). The meroterpenoid then undergoes several modifications to the products austinol and dehydroaustinol, figure 2.8 (Lo 2012).

**Figure 2.8** The structures of 3,5-dimethylorsellinic acid, austinol and dehydroaustinol. The carbon skeleton of 3,5-dimethylorsellinic acid incorporated into austinol and dehydroaustinol is presented in bold.

Isoprene units can be linked to PKs and NRPs by a prenyltransferase (dimethylallyl tryptophan synthases DMAT) through either O-C or C-C bonds. Emericellin, figure 2.9 is a mixed terpenoid and PK metabolite which is produced in the monodictyphenone pathway in *A. nidulans*. The origin of the two isoprene units arise from two separate prenyltransferases, XptA and XptB, and are coupled to the PK through a C-O and a C-C bond, respectively (Sanchez 2011).

**Figure 2.9** Structure of emericellin. The carbon skeleton of the two isoprene units is highlighted in bold.

### 2.2 Genetics of secondary metabolite biosynthesis

To understand the biosynthesis of secondary metabolites, one of the important tasks is to target genes encoding the enzymes that catalyze steps within the biosynthetic pathways. Due to the availability of fungal genome sequences, the search for genes coding for these enzymes can be intensified among the thousands of genes, e.g. the 11223 genes in the *A. nidulans* genome (Arnaud 2012). In this section the prediction of gene function, regulation and activation of genes related to secondary metabolite biosynthesis in filamentous fungi will be introduced with a focus on *A. nidulans*.

#### 2.2.1 Prediction of gene function

It is possible to identify putative secondary metabolite genes within genome sequences, through computational analysis. Automated annotations and manual validations of predictions are already present for several of the *Aspergillus* genomes. Therefore Blast analysis on amino acid sequences of predicted protein activities will offer a list of homologous targets and their respective loci, especially due to their functional domains which can be highly homologous. As the gene function prediction is based on available knowledge on genes and protein activity, it is important to stress
Production of secondary metabolites in filamentous fungi

that these predictions need to be confirmed through follow-up experiments, unless sequence homology and coverage of sequence is high. Still sequence analysis of the genomes gives a qualified putative overview of the gene functions. In table 2.1 the number of secondary metabolite synthase genes which have been predicted in the literature for the fully genome sequenced strains of *A. niger*, *A. oryzae*, *A. flavus* and *A. nidulans*, using a BlastP analysis approach. The total amount of synthase genes exceeds the number of known biosynthetic pathways. For example *A. nidulans* are predicted to contain 32 PKS genes (Nielsen 2011) of which 17 have been linked to metabolites (Brown 1996, Watanabe 1998, Watanabe 1999, Bergmann 2007, Chiang 2008, Szewczyk 2008, Bok 2009, Chiang 2009a, Schroechk 2009, Nielsen 2011, Ahuja 2012) indicating a large potential for new metabolites, even in one of the most studied fungi.

Table 2.1: Overview of the predicted number of synthase genes in the fully genome sequenced strains of *A. flavus*, *A. nidulans*, *A. oryzae*, and *A. niger*. *These numbers includes NRPS- or PKS-like genes.

<table>
<thead>
<tr>
<th>Species</th>
<th>PKS*</th>
<th>NRPS*</th>
<th>Hybrid</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. flavus</em></td>
<td>28</td>
<td>32</td>
<td>2</td>
<td>Cleveland 2009</td>
</tr>
<tr>
<td><em>A. nidulans</em></td>
<td>32</td>
<td>27</td>
<td>1</td>
<td>von Döhren 2009, Nielsen 2011</td>
</tr>
<tr>
<td><em>A. niger</em> (ATCC 1015)</td>
<td>33</td>
<td>15</td>
<td>9</td>
<td>Sanchez 2012</td>
</tr>
<tr>
<td><em>A. niger</em> (CBS 513.88)</td>
<td>39</td>
<td>33</td>
<td>7</td>
<td>Sanchez 2012</td>
</tr>
<tr>
<td><em>A. oryzae</em></td>
<td>27</td>
<td>32</td>
<td>2</td>
<td>Cleveland 2009</td>
</tr>
</tbody>
</table>

In filamentous fungi genes linked to biosynthesis of a given secondary metabolite tend to cluster, which enables the discovery of genes encoding tailoring enzymes both through Blast analysis but also the use of techniques like microarray and qRT-PCR (quantitative reverse transcription polymerase chain reaction) can assist in discovering connections between the expression of genes and gene clustering in the genome. Microarray is applicable for the overview of the expression level of genes in a genome under specific conditions. The data can be used to analyze whether genes are downregulated, unaltered, or upregulated at certain conditions or between different strains. Hereby, it is possible to establish connections between large groups of genes at the same time. In qRT-PCR, the strength is the ease of measuring for example relative expression of a gene in a mutant strain compared to the expression in a reference strain, or in one strain at different growth parameters. In the latter case, the overview is limited to genes chosen for the study.

Due to the fact that the total amount of synthase genes and thereby biosynthetic pathways far exceeds the number of known genes different terms to distinguish between unknown and not expressed genes and pathways has been developed. Orphan biosynthetic genes are defined as biosynthetic loci for which the corresponding metabolite(s) is/are unknown whereas silent pathways are pathways where the genes are not expressed under the given growth conditions (Gross 2007).

2.2.2 Activation of silent genes

With the recent recognition of the high percentage of orphan secondary metabolite genes several different techniques of activating biosynthetic genes have been explored. In this section some of the activation techniques, with relations to the methods used in this thesis, will be described with emphasis on *A. nidulans*. These approaches include the One Strain MAny Compounds (OSMAC) method, epigenome manipulations, gene regulation, and co-cultivation.
2.2.2.1 One strain – many compounds

One classic way to activate the production of secondary metabolites is to vary the growth conditions. The principle behind the method, named one strain – many compounds (OSMAC) by Zeeck and co-workers, is to expose the microorganism to other cultivating conditions than the standards used in laboratories (Fuchser 1997, Schiewe 1999, Höfs 2000, Bode 2002). There are several factors which can be altered for example temperature, pH, culture vessel, aeration, cultivation time, light intensity and media composition including carbon and nitrogen sources and salt concentration. All these factors can influence the biosynthesis of secondary metabolites in several places including transcription, translation and enzyme inhibition or activation (figure 2.10), however, the understanding of the exact mechanisms for the change in metabolic profile due to change in culture conditions is usually not completely understood and therefore difficult to predict (Bode 2002).

![Figure 2.10](image)

Figure 2.10. Storage of information in the cells. a) DNA is folded up on the histones into nucleosomes which coil up and are packed into the chromosome. b) The DNA is transcribed into mRNA which is translated into proteins which in the case of PKS and NRPSs are enzymes catalyzes the formation of a product.

![Figure 2.11](image)

Figure 2.11 Structure of orsellinic acid, F-9775A and F-9775B.

The use of the OSMAC approach which included 20 different culture conditions of an *A. nidulans* strain, led to the activation of the orsellinic acid cluster resulting in the production of orsellinic acid and compounds derived thereof including F-9775A and F-9775B, figure 2.11 (Sanchez 2010). In another study by Scherlach and co-workers an *A. nidulans* strain was cultivated under 45
Production of secondary metabolites in filamentous fungi

different culture conditions, which resulted in the isolation of two novel metabolites, aspernidine A and B, figure 2.12 (Scherlach 2010).

Figure 2.12 The structure of aspernidine A and B.

Interestingly these two examples of the use of the OSMAC approach have cultivated the strains used many conditions to activate the production of these metabolites. In 2011 Nielsen and co-workers showed the metabolic profile of an A. nidulans reference strain grown on seven complex media displayed large differences in the production of metabolites including the orsellinic acid derived compounds (Nielsen 2011).

2.2.2.2 Discovery of secondary metabolites through genome mining

The elucidation of the secondary metabolite production potential can, as stated above, be initiated through genome mining. This strategy is confined to genome sequenced species, and involves the search in the genome for genes which are predicted to have enzymatic activities that are associated with production of secondary metabolites. This enables subsequent systematic genetic manipulations and validation through chemical analysis. This includes deletion and/or overexpression of genes in connection to secondary metabolite gene clusters or regulatory genes. The regulatory genes can act globally within the fungus or within a single secondary metabolite gene cluster.

Epigenetic manipulations of global regulatory genes

The presence and possibilities of manipulation techniques of global regulators have been a vast challenge; however, several advances have increased the interest in epigenetic manipulations where examples of some of these manipulations will be described below (Keller 2005, Yu 2005a, Cichewicz 2009, Scherlach 2009).

The presence of secondary metabolite clusters in silent areas of the chromosomes, e.g. near telomeres and centromeres, suggests that chromatin remodeling factors can influence the expression of genes responsible for secondary metabolite production. Regulation of gene expression by chromatin regulation is directed by modifications of the histones through e.g. methylation, which closes, and acetylation, which opens the histones, since it is thought that histone modifications may control interactions of histones with transcriptional activators and repressors (Jenuwein 2001).

In 2004 Bok and Keller reported the LaeA protein, a putative methyl transferase with homology to histone methyltransferases, as a global regulator of secondary metabolism in Aspergillus species, however, the mechanism of gene activation remains elusive (Bok 2004, Bayram 2012). Deletion of laeA resulted in a significantly decreased production of secondary metabolites in A. nidulans including sterigmatocystin and penicillin (Bok 2004). The importance of LaeA was further
strengthened as it was shown to be part of the conserved velvet complex, which is a major factor for regulation of fungal development hereby coupling the triggering of secondary metabolism to central fungal biology (Bayram 2008).

Another example of chromatin regulated gene expression came from the deletion of a histone deacetylase, hdaA, which led to significant increase in production of penicillin and norsolorinic acid (intermediate in sterigmatocystin biosynthesis, figure 2.3) (Shwab 2007).

The SUMO protein is also a chromatin regulating gene which has been shown to have an effect on secondary metabolism. It is an ubiquitin-like protein which is added to many proteins including histones posttranslationally (reviewed in Verger 2003, Gill 2004, Johnson 2004). A. nidulans contains one sumO encoding gene (Wong 2008) which upon deletion was shown to give a decrease in the production of secondary metabolites such as austinol, dehydroaustinol and sterigmatocystin as well as an increase in asperthecin production (Szewczyk 2008). Ten NR-PKSs were identified by genome mining and a deletion series of these in the sumOΔ background was performed and the PKS for asperthecin biosynthesis was identified (Szewczyk 2008).

Investigation of other possible chromatin-level regulation of secondary metabolite biosynthesis in A. nidulans led Bok and co-workers to delete cclA, an ortholog to the BRE2 yeast gene, encoding an enzyme partner of the COMPASS transcriptional regulator complex conserved in eukaryotes, which both facilitate and represses chromatin-mediating processes through methylation of the histone tails (Bok 2009). Deletion of cclA showed an altered chemical profile compared to the reference strain where two secondary metabolite gene clusters were activated and afterwards identified through deletions studies of mdpG and orsA which catalyzed the biosynthesis of monodictyphenone, figure 2.13, and several emodin analogs, as well as F9775A and F9775B, respectively (Bok 2009).

**Transcription factor manipulations**

Secondary metabolite gene clusters share common regulation of its member genes that can be activated or repressed by transcriptional regulators. Therefore activation or repression of the transcriptional regulators for example transcription factor offers a convenient way to control the expression of whole cluster by one genetic operation. Thus in studies based on mining genomes, the presence of a transcription factor within the cluster has been the driver in characterizing the gene cluster. Asperfuranone and the aspyridones are examples of metabolites discovered through genome mining in A. nidulans followed by activation of transcription factors (Bergmann 2007, Chiang 2009a). Chiang and co-workers noticed that a NR-PKS encoding gene, afoE, and afoG a gene coding for a HR-PKS were located close to each other on chromosome VIII. Since no product had ever been detected from activity of this locus, and due to the rare constellation of two
neighbor PKSs, the authors speculated whether a novel metabolite could be identified. A putative transcription activator, \(afoA\), was found near the PKSs and the authors replaced the upstream sequence of \(afoA\), estimated to be the native promoter, with an inducible promoter, \(alcA\), which led to the production of asperfurane (Chiang 2009a).

Bergmann and co-workers identified a putative transcription factor in the genome and overexpressed it under the control of an inducible promoter, \(alcA\). They demonstrated by Northern blot analysis that six of the nearest neighbor genes were upregulated on inductive medium and detected the production of aspyridones and two intermediates (Bergmann 2007). In a recent study the promoters of the NR-PKS genes and other genes necessary for NR-PKS product formation or release were also replaced with \(alcA\), leading to the identification of several NR-PKSs (Ahuja 2012).

The above mentioned studies illustrated the action of local acting transcription factors; however, it has also been reported that a transcription factor associated with a cluster can act on genes outside the cluster. Bergmann and co-workers showed that overexpression of a transcription factor on chromosome II of \(A. nidulans\) led to the activation of the asperfurane gene cluster on chromosome VIII (Bergmann 2010). Another example is the transcription factor that regulates sterigmatocystin biosynthesis in \(A. nidulans\) and aflatoxin biosynthesis in \(A. flavus\) and \(A. parasiticus\), \(aflR\), also acts on genes outside the cluster (Price 2006).

For further reading on the regulation of secondary metabolite biosynthesis see reviews by Hoffmeister (Hoffmeister 2007), Palmer (Palmer 2010), Fox (Fox 2008), Strauss (Strauss 2011), Yin (Yin 2011).

2.2.2.3 Other methods and combinations of activation strategies

There are several other methods for activation of secondary metabolite production. One method is to mimic growth conditions resembling known natural environments. This is utilized in the OSMAC approach as described, and also in co-cultivation of microorganisms, which might induce the production of new secondary metabolites due to the influence of the different organisms to each other. The co-cultivation of \(A. nidulans\) with 58 soil-dwelling actinomycetes led to the activation of the orsellinic acid cluster resulting in the production of orsellinic acid and compounds derived thereof including F-9775A and F-9775B (Schroeckh 2009).

In section 2.2.2.2 molecular methods for altering the epigenome was presented. However, chemical methods which have similar effects are beginning to emerge. The research on chemical epigenetic modifiers focuses on the inhibition of DNA methyltransferases and histone deacetylase which modifies DNA and histones resulting in changes in the folding of DNA (Cichewicz 2010). A screening of nine DNA methyltransferase and histone deacetylase inhibitors against 12 fungal species, by Cichewicz and co-workers, led to production of new or enhanced levels of secondary metabolites in 11 species (Williams 2008). A study by the same group which tested inhibitors of a DNA methyl transferase (5-azacytidine) and a histone deacetylase (suberoylanilide hydroxamic acid) on \(A. niger\), figure 2.14, led to increased transcriptional rates of all but seven of the total 55 PKSs, NRPSs and hybrid PKS-NRPSs clusters (Fisch 2009). One of the new metabolites identified was the novel metabolite nygerone A, figure 2.14 (Henrikson 2009).
Several of the above mentioned strategies have an effect on the production of the same metabolites. As mentioned in the previous sections several orsellinic acid derived compounds were discovered in *A. nidulans* through both co-cultivation, histone-modification and the OSMAC approach (Bok 2009, Schroeckh 2009, Sanchez 2010, Nielsen 2011).

For further reading on activation of secondary metabolite gene clusters there are several reviews present in the literature (Gross 2007, Chiang 2009b, Scherlach 2009, Brakhage 2011, Chiang 2011, Winter 2011, Yin 2011).
3 Overview over experimental work

The figure below gives an overview of the experimental work performed over the past three years. The work in this thesis has been focused on *Aspergillus niger*, *A. oryzae* and *A. nidulans*. The results are presented and discussed in chapter 4 and the experimental sections can be found in the papers 2-7. The outcome of the work at the external stay at the University of Southern California was not as expected and is therefore not ready for publication yet. Since paper 1 is a review it is not listed in figure 3.1. Paper 5 and 6 are rough drafts; however, further experiments need to be conducted prior to submission.

Figure 3.1 Experimental overview over the work for duration of the past three years. Abbreviations: Oex: overexpression, PKS: polyketide synthase.

The aim of the work presented in this thesis has been to link secondary metabolites to genes; however, to do this several experiments and analysis needs to be performed. This includes bioinformatics, molecular biology, chemical analysis, isolation of metabolites and structure elucidation. My part of have consisted of the latter; chemical analysis, isolation and structure elucidation.

3.1 Linking of secondary metabolites to genes

This section will briefly outline the tools and instruments applied in the thesis. The experimental descriptions are presented in the paper 2-7; however, examples will be included in chapter 4.
Overview over experimental work

3.1.1 Targeted genetic manipulations including deletions and over-expressions

Most of the strains analyzed in this work were genetically modified. The host strains used for genetic transformation in *A. nidulans* had been disrupted in the non-homologous end-joining pathway enabling efficient gene targeting (Nielsen 2008), referred to as the reference strain in this thesis. This was supported by the development of versatile tools to construct DNA substrates, e.g. fusion PCR and USER cloning technology. The mutant strains analyzed were gene deletions, point mutations or overexpressions where the gene of interest has been moved to an insertion site under the control of a constitutive or inducible promoter (Nielsen 2006, Nielsen 2008, Nielsen 2011, Hansen 2011).

3.1.2 Metabolome analysis

The chemical analysis of all the strains has been performed on three instruments. A UHPLC (Dionex RSLC Ultimate 3000) equipped with a Kinetex C$_{18}$ column coupled with diode array detection (DAD), throughout the thesis called UHPLC-DAD. Two TOF-MSs; a LCT oTOF mass spectrometer (Micromass), equipped with a Phenomenex C$_{18}$ column and diode array detection, and a MaXis 3G QTOF (Bruker Daltronics) coupled to a Dionex Ultimate 3000 UHPLC system equipped with a Kinetex C$_{18}$ column, throughout the thesis called LC-TOF-MS (LC-DAD when DAD detection in reported) and UHPLC-TOF-MS respectively.

The approach in the metabolome analysis has been to compare mutant strains with the reference strain to detect metabolites either appearing or disappearing. Generally the strains have been analyzed in UHPLC-chromatograms (210 nm) and, in some cases, both positive and negative ionization mode, enabling the identification of metabolites which could only be detected in one of the techniques. When interesting peaks had been identified, the metabolite data was compared to the literature, and the metabolite database Antibase (Laatsch 2010) in order to identify already reported metabolites. The MaXis UHPLC-TOF-MS was acquired during the study and provided invaluable information on both the molecular formula and isotope pattern due to the high mass accuracy (1 ppm).

3.1.3 Isolation and structure elucidation

The approach described above was done in micro-scale where plugs from a colony was extracted and analyzed (Frisvad 1987, Smidsgaard 1997). When the metabolome analysis of the mutants revealed new interesting metabolites the cultivations were scaled up to allow isolation and structure elucidation. The biomass was extracted and metabolites were isolated using different LC techniques with various column material including normal phase, reverse phase and size exclusion chromatography. The structures of the isolated metabolites were solved using 1D and 2D NMR-spectroscopy. The NMR spectra were acquired using standard experiments on either a Varian Unity Inova 500 MHz spectrometer placed at DTU Chemistry and/or on a Bruker Avance 800 MHz spectrometer at the Danish Instrument Center for NMR Spectroscopy of Biological Macromolecules.
4 Results and Discussion

The experimental work performed for the past three years have focused on the industrially important filamentous fungi; *A. oryzae* and *A. niger* and the genetic model organism *A. nidulans*. The results of the work will be presented in this section starting with the exploration of *A. oryzae* wildtype chemistry including several new metabolites and comparison of these to *A. flavus* metabolites (paper 2). The next section (4.2) will focus on the work done with *A. niger* where a new fumonisin analog was isolated and characterized through NMR spectroscopy (paper 3). Section 4.3 describes the results of the *A. nidulans* work. This included the analysis of a deletion library of 32 putative PKSs encoding genes including activation of genes using several complex media and linking of austinol and dehydroaustinol to their synthase gene (section 4.3.1, paper 4). This was followed by identification of several new *A. nidulans* metabolites through activation of a transcription factor, and a proposed biosynthetic pathway through deletion studies (section 4.3.2, paper 5). The last part of this section (4.3.3) focus on the novel metabolites, nidubenzal A and B and the linking of asperugin A and B to their PKS encoding gene (section 4.3.3.1, paper 6) and the juvenile hormones (section 4.3.3.2, paper 7) which were isolated after activation through genetic modifications combined with the use of complex media. The biosynthetic origin of asperugin A and B was identified by reference to the deletion library followed by genetic modifications.

4.1 *Aspergillus oryzae*

*Aspergillus oryzae* is one of industry’s most used workhorses and has been used for centuries in food fermentation for the production of for example sake, miso, soy sauce and other traditional Asian foods (Machida 2008a). *A. oryzae* is also a widely used organism for production of amylase, lipases and proteases, and, more recently, also for heterologous expression of secondary metabolite genes and non-fungal proteins (Punt 2002, Meyer 2008, Fisch 2011). For many years, *A. oryzae*, has been suspected to be a domesticated form of *A. flavus*, a plant and mammalian pathogenic saprophyte, capable of producing some of the most carcinogenic compounds known, the aflatoxins, figure 4.1. Genetic work and subsequent genome sequencing of strains of both species have verified the tight link between the species (Geiser 1998, Geiser 2000, Machida 2005, Yu 2005b, Abe 2006, Kobayashi 2007, Machida 2008a).

![Figure 4.1 The structure of aflatoxin B1.](image)

The availability of the sequenced strains of *A. oryzae* (RI840 (ATCC 42149)) (Machida 2005) and *A. flavus* (NRRL 3357) (Yu 2005b) have led to comparison of the two genomes. The comparison shows high homology: 99.5% genome homology and 98% at the protein level (Rokas 2007). The
high homology could lead one to expect A. oryzae to produce most of the metabolites found in A. flavus (Machida 2008a, Machida 2008b, Machida 2005, Payne 2006, Yu 2008) and the preliminary bioinformatics studies performed along with the genome sequencing reported roughly the same number of predicted genes: 28 PKSs, 32 NRPSs and 2 hybrid PKS-NRPSs for A. flavus and 27 PKSs, 32 NRPSs and 2 hybrid PKS-NRPS for A. oryzae with two NRPSs apparently unique for each strain (Cleveland 2009). However, published metabolic data indicated a very low chemical correlation between the two strains (Laatsch 2010). Most of the predicted genes for secondary metabolites of A. oryzae (or A. flavus) have not been linked to specific metabolic products. Only genes of the most important toxins: aflatoxin (Yu 2008, Lee 2006, Tominaga 2006), cyclopiazonic acid (Tokuoka 2008, Chang 2009) and aflatrem (Nicholson 2009) have been annotated in both species which leaves much to be explored.

The aim of the work in this section was to perform an initial comparative investigation of the small molecule chemistry from the two genome sequenced strains of A. oryzae (RIB40) and A. flavus (NRRL 3357) along with isolation and structure elucidation of new metabolites from A. oryzae, in order to get further insights into possible homologies in secondary metabolite production for these two important and related species. The results are described and discussed below and further experimental details can be found in paper 2.

4.1.1 Comparison of secondary metabolite profiles of A. oryzae and A. flavus

For the analysis of A. oryzae RIB40 (and A. flavus NRRL 3357) chemistry, we investigated a series of solid media cultivated at 25°C in the dark for 7 and 14 days followed by micro-scale extractions (Frisvad 1987, Smedsgaard 1997) and subsequently analysis with HPLC-DAD-MS for investigation of optimal secondary metabolite production conditions. The screening indicated the greatest chemodiversity and metabolite production from CYA, YES and WATM agar.

The comparison of the secondary metabolite profiles of the two strains showed a high degree of chemical difference on all media as illustrated in figure 4.2 for the WATM medium. The major metabolite repetitions between the two genome sequenced strains were merely kojic acid and ergosterol as well as a number of minor metabolites (not analyzed here), which seemed to be shared between the two strains. Altogether, this is in sharp contrast to the high gene homology; however, it reflects the difference in metabolic published data (Laatsch 2010).
Several previously reported metabolites of *A. oryzae* were dereplicated and we found that the RIB40 strain did not produce detectable levels (LC-MS) of cyclopiazonic acid (Orth 1977) (as also noted by Tokuoka *et al.* (Tokuoka 2008)), asperfuran (Pfefferle 1990), sporogen AO1 (Tanaka 1984a, Tanaka 1984b), maltoryzine (Iizuka 1962) or aspergillomarasmine A (Robert 1962, Barbier 1963), under these growth conditions. The RIB40 strain did produce kojic acid (Manabe 1984, Bentley 2006) and aspirochlorin (Sakata 1982, Sakata 1983, Sakata 1987, Monti 1999, Klausmeyer 2005), figure 4.3 and a series of potentially new metabolites of which some were isolated and structure elucidated. These metabolites included the parasiticolites, ditryptoleucine, oryzamide A_{1,2} and 13-desoxy-paxilline.
4 | Results and discussion

4.1.1.1 13-desoxypaxilline

During fermentation of the chemically potent RIB40 strain, we have been interested in the tremorgenic compounds, allegedly coupled to fungal sclerotia (Wicklow 1982, Gloer 1988, Gloer 1989, Staub 1992, Staub 1993, Tepaske 1989a, Tepaske 1989b, Tepaske 1990, Tepaske 1992) and whether these could be found in *A. oryzae* as they have been in *A. flavus*. The RIB40 strain produces large and abundant sclerotia, especially on WATM agar, a fact not widely announced in literature although sclerotia have been observed in *A. oryzae* sporadically (Raper 1965, Wicklow 2007, Jin 2010). No sclerotia were observable after 14 days on YES agar, but although these metabolites are often characterized as sclerotial metabolites, there is not a strict correlation between the biosynthesis of these metabolites and the formation of sclerotia, as also noted by Wilson (Wilson 1966), and this extract was used for the described isolations.

We isolated the aflatoxin precursor 13-desoxypaxilline which was originally isolated from *Penicillium paxilli* (Springer 1975, Longland 2000, Bilmen 2002, Sabater-Vilar 2003, Sheehan 2009). Aflatoxin is a known metabolite of *A. flavus* and was discovered by Wilson and Wilson in 1964 (Wilson 1964) and structure elucidated by Gallagher et al. (Gallagher 1978, Gallagher 1980). 13-desoxypaxilline was present in several media micro-scale extracts including YES, CYA, OAT and WATM. It was isolated from the above mentioned YES extract and the structure was confirmed by LC-MS, LC-MS/MS and NMR analysis. It was investigated through HPLC-DAD and LC-MS/MS whether aflatoxin or other intermediates in the biosynthetic pathway were present in the extracts, however, none of the proposed intermediates towards aflatoxin or aflatoxin itself, figure 4.4 could be detected and only one sample (WATM, 7d) showed traces of paspaline, a precursor for 13-desoxypaxilline.

The discovery of 13-desoxypaxilline as the end-product of *A. oryzae* RIB40 is in agreement with the analysis of Nicholson et al., who showed that a frameshift mutation in the *atmQ* gene presumably accounts for the 13-desoxypaxilline not being converted into paspalicine and paspalinine (Nicholson 2009). This mutation is therefore likely responsible for terminating the
aflatrem biosynthesis in RIB40 prematurely. Contrary to our discovery, Nicholson et al. did not find the aflatrem gene cluster of RIB40 to be transcribed during their fermentations (Nicholson 2009).

Figure 4.4 The final steps in the proposed biosynthesis of aflatrem (in A. flavus). A. oryzae RIB40 biosynthesis stops at 13-desoxypaxilline (Nicholson 2009).

A second extract was made from 100 plates of a 14 day old A. oryzae culture grown on WATM agar with abundant sclerotia formation to validate the findings from the YES extract. The analysis of the WATM extract showed 13-desoxypaxilline as a major metabolite alongside other sclerotium-related metabolites, such as aflavinines (based on UV – data not shown) which were unfortunately not isolated in pure and large enough quantities for NMR analysis.

The isolated 13-desoxypaxilline is a member of the paspalitrem tremorgens, a widely distributed group of metabolites that have been isolated from several genera: Penicillium, Eupenicillium, Claviceps, Emericella, Aspergillus and Phomopsis (Cole 1981, Steyn 1985, Bills 1992). Besides the tremorgenic activity in animals, these metabolites have been shown to be insecticides (Steyn 1985), which is believed to be their ecological function together with aflatoxin and cyclopiazonic acid for protection of the sclerotia against fungivorous insects (Wicklow 1982, Gloer 1988).

4.1.1.2 Parasiticolides

In addition to 13-desoxypaxilline, two new analogues of parasiticolide A were also isolated. The metabolites showed, through LC-MS and NMR analysis, to be dide- and 14-deacetyl analogues and are most likely precursors to the sesquiterpene parasiticolide A (also called astellolide A), figure 4.5. The shift values of all the parasiticolide analogues correlated with the published data for parasiticolide A, except for the missing signals of the acetate units and their minor influence on the chemical shift values of adjacent protons and carbons (see paper 2 for NMR details) (Hamasaki 1975, Gould 1981). Several different extraction procedures were tested to verify the correctness of the compounds as genuine metabolites and not in vitro degraded parasiticolide A products, but all samples showed only dide- and 14-deacetyl parasiticolide A and no traceable (LC-MS) levels of parasiticolide A itself, even with different non-acidic extraction procedures. Parasiticolide A have been isolated from A. flavus var. columnaris once (Shiomi 2002) and was originally isolated and characterized from A. parasiticus (Fukuyama 1975, Ishikawa 1984) and later
4 Results and discussion

also from a mutant of *Emericella variecolor* (Gould 1981). Recently parasiticolides have been detected in the newly described species *A. arachidica* and *A. minisclerotigenes* (Pildain 2008). There have to our knowledge not been published any toxic studies on the parasiticolides, but the related peniopholides from the fungus *Peniophora polygonia* have been reported to have antifungal properties (Ayer 1992).

**Figure 4.5** Structures of Parasiticolide A, and 14-deacetyl parasiticolide A, dideacetyl parasiticolide A.

According to our observations parasiticolides are more often detectable metabolites of *A. oryzae* than of *A. flavus* under the same fermentation conditions, suggesting that the pathway is partly silenced for *A. flavus* and may need to be activated under otherwise normal growth conditions. It is interesting that parasiticolide A is scarcely observed in *A. flavus*, when it is an important product of *A. oryzae* and also of *A. parasiticus*. We also isolated and elucidated a formyl variant of parasiticolide A; however, it was not possible to exclude the possibility of in vitro chemistry due to the formic acid added during the ethyl acetate extraction, so the correctness of this metabolite remains tentative (Hamasaki et. al. used benzene to extract parasiticolide A (Hamasaki 1975)).

To further verify these observations, a MS/MS method was used to analyze several different microscale extracts of *A. oryzae* RIB40 for parasiticolide A. Trace amounts of parasiticolide A could be measured, indicating a complete transformation into the end product. The small amount of parasiticolide A in RIB40 (roughly 1:1000 ratio compared to 14-deacetyl parasiticolide A, presuming the same response factor) might be the result of spontaneous acetylation involving the first acetylating enzyme. When the gene cluster responsible for the biosynthesis of these metabolites is mapped, it is likely that the gene responsible for the last (specific) acetylation will be found to be mutated.

4.1.1.3 Ditryptoleucine and oryzamide A$_{1-2}$

Four new *A. flavus* related NRP metabolites were also isolated from *A. oryzae* RIB40. The metabolites showed, through structure elucidating by $^1$H, DQF-COSY, HSQC and HMBC NMR data, to be two isomeric metabolites named ditryptoleucine related to the ditryptophenalines (Springer 1977, Barrow 1993, Oleynek 1994, Barrow 1994), figure 4.6, and two indole-enamides named oryzamide A$_{1-2}$ (*cis*- and *trans*-isomers) structurally similar to the antibiotic miyakamide B$_{1-2}$ (Shiomi 2002), figure 4.7.
Figure 4.6 Structures of ditryptoleucine (A. oryzae metabolite) and ditryptophenaline (A. flavus metabolite).

The diketopeptide ditryptoleucine was isolated in two variants with $^1$H-NMR shifts varying around the C-C dimeric bond, but with the same base structure. This unspecific dimerization is in line with previously isolated compounds (Barrow 1994). A hybrid between the ditryptophenaline and ditryptoleucine was isolated from an Aspergillus sp. as WIN 64745, with both a phenylalanine and leucine moiety, but with no N-methylation (Barrow 1993). These compounds were tested and proved antagonistic against substance P at the NK1 receptor (Movassaghi 2008). Intriguingly, these new A. oryzae metabolites have apparently exchanged phenylalanine with leucine compared to the similar A. flavus metabolites, indicating either a common trait in the domestication process or coupling between the two pathways.

Figure 4.7 Structures of oryzamide A$_{1,2}$ and the miyakamides B$_{1,2}$.

It proved difficult to unambiguously isolate the cis- and trans-isomers of oryzamide A in a pure form due to isomerization around the double bond. Due to this difficulty the structure elucidation has been performed on mixtures of compounds. It is evident from the structure elucidation that oryzamides A$_{1,2}$ only differ in the configuration around the double bond of the enamine as evident from the size of the coupling constant. Marfey analysis of the amino acid derived compounds only established the presence of L-tyrosine in the oryzamides. The oryzamides A$_{1,2}$ are variants of the miyakamides B$_{1,2}$ (Shiomi 2002). The phenylalanine in the miyakamides has been substituted with leucine in the oryzamides. However, this exchange of phenylalanine and leucine has often been found for compounds produced within the same species. Penicillium polonicum produces both fructigenine A (also called puberulione and rugulosuvine A) containing phenylalanine and fructigenine B (also called verrucofortine) containing leucine (Arai 1989, Frisvad 2004a, Frisvad 2004b). We could also detect two slightly later eluting compounds in the A. oryzae RIB40 extracts likely to be further indole-enamides having phenylalanine incorporated instead of tyrosine as evident from LC-DAD-MS analysis (data not shown) as also seen for the miyakamides (Shiomi 2002).
4 Results and discussion

4.1.2 Summary and part conclusion

The tremorgenic 13-desoxypaxillin have been isolated from A. oryzae RIB40 and verified under several different growth conditions contrary to previous studies. We believe that 13-desoxypaxillin is the end-product of the aflatrem biosynthesis for the RIB40 strain which correlated well with the results of Nicholson et al. (Nicholson 2009) who described a mutation in the gene responsible for further modifications of 13-desoxypaxillin in A. oryzae RIB40.

The new metabolites dide- and 14-deacetyl parasiticolide A were also found as genuine products from the RIB40 strain where parasiticolide A was only detected in trace amounts using a LC-MS/MS method. This indicates that a defective acetylation of the 14-deacetyl parasiticolide A and the small amount of parasiticolide A in RIB40 could be the result of spontaneous acetylation in the cell cytosol.

The new NRPs ditryptoleucine and oryzamide A1-2 appear to be natural variants of known A. flavus metabolites. They share the exchange of a phenylalanine for a leucine, although they are believed to originate from two unrelated pathways. As for 13-desoxypaxillin, dide- and 14-deacetyl parasiticolide A are almost certainly products of a prematurely ended biosynthesis, here parasiticolide A. Further analysis of the genomes of A. oryzae and A. flavus could give possible candidates for the NRPS genes responsible for the production of ditryptoleucine and oryzamide A1-2. After identification of the synthetase genes it could be interesting to investigate whether the aminoacid exchange is due to genetic mutations.

Altogether the finding contribute to understanding why the overall chemical profiles of A. oryzae (RIB40) and A. flavus (NRRL 3357) appear quite different since some of the end-products usually seen in A. flavus are apparently not reached in A. oryzae or the metabolites differ in one incorporated amino acid. Whether the different chemical profiles are merely the result of different regulation or are a result of genuine mutations, as with 13-desoxypaxillin, remains to be settled. A. oryzae RIB40 is clearly a chemically potent strain, and as more of its chemistry is unfolded it is likely to be revealed that most of the biosynthetic pathways of A. flavus will be found more or less functional which correlates to the high degree of homology between the two strains.

The findings above shows that genome comparisons are an important tool to explore the chemical potential of different strains, however, a regulatory difference or mutations (small or large) in the genome can lead to large chemical differences between species who seem alike.

4.2 Aspergillus niger

Aspergillus niger is another industrially important filamentous fungus used in the fermentation industry to produce organic acids such as citric acid and gluconic acid as well as enzymes such as amylases and lipases (Schuster 2002). Besides the uses in the fermentation industry A. niger is also a contaminant of crops and feed distributed worldwide, including corn, peanuts, onions, apples, grapes, raisins, mangos and dried meat products (Perrone 2007, Nielsen 2009).

Due to the vast importance of A. niger, both as contaminant and its uses in fermentations, the genome of two strains of A. niger have been sequenced (Baker 2006, Pel 2007) giving the
possibility of searching the genomes for secondary metabolites gene clusters. The next section describes how the genome sequences revealed the presence of a mycotoxin (fumonisins) gene cluster known from other species making researchers look for and report the production of fumonisins in *A. niger*. We have identified, isolated, and structure elucidated a new fumonisin isomer (paper 3).

4.2.1 Fumonisin production in *A. niger*

The sequencing of the genome of *A. niger* revealed the presence of a fumonisin-like gene cluster in two different isolates, ATCC 1015 (Baker 2006) and CBS 513.88 (Pel 2007). The fumonisins are a group of HR-PKs produced by *Fusarium verticillioides* and other Fusaria (Bezuidenhout 1988, Gelderblom 1988, Marin 2004). Several analogues of fumonisin have been reported, however, fumonisin B₁, B₂ and B₃ of the B-series are the most abundant in nature and are all carcinogenic (Gelderblom 1993, Rheeder 2002, Bartók 2006). The B-series analogues contain a terminal 2-amino-3-hydroxy motif on an eicosane backbone and two hydroxyl groups esterified with tricarboxylic acids. There had not, at the time of the publications of the genome sequences, been any reports of fumonisin production in *A. niger*. Therefore, after the discovery of the putative fumonisin gene cluster, Frisvad *et al.* (Frisvad 2007) investigated the possibility of fumonisin production in *A. niger*. They reported detection, through LC-MS comparisons to standards, of fumonisin B₂ (FB₂) from four strains of *A. niger* including the two fully genome sequenced strains (ATCC 1015 and CBS 513.88) and ATCC 9029 and NRRL 326 (Frisvad 2007) Later it was reported that *A. niger* could also produce fumonisin B₄ (Nielsen 2009, Noonim 2009). The ability of *A. niger* to produce fumonisins gave speculations to whether this could be an overlooked health risk, since *A. niger* is used in industrial fermentations and is a common contaminant of feed and foods due to the possibility of fumonisins ending up in the final products (Frisvad 2007).

4.2.1.1 Isolation and NMR characterization of fumonisin B₂ and B₆ from *A. niger*

Further analysis of the *A. niger* strains revealed the ability of *A. niger* to produce another isomer besides fumonisin B₂ and B₄ with the same elemental composition as fumonisin B₁ and iso-fumonisin B₁ but with a different retention time. We isolated and characterized this isomer, named fumonisin B₆, from *A. niger* NRRL 326, along with fumonisin B₂ which were NMR characterized to validate that fumonisin B₂ from *A. niger* is identical to that reported in several *Fusarium* species.

![Figure 4.8 The structures of fumonisin B₁, iso-B₁, B₂, B₃, B₄ and B₆.](image-url)
The structure of fumonisin B₆ were solved based on DQF-COSY, HSQC, HMBC and H2BC NMR experiments along with comparison to literature values of fumonisin B₁ (Savard 1994a) and iso-B₁ (MacKenzie 1998). Unambiguous assignment of the proton and carbon resonances of fumonisin B₆ could be obtained for most of the resonances. However due to spectral overlap individual assignment of C₇-C₉ was not possible. The two tricarboxylic acid side chains were assigned as two separate spin systems with only minor chemical shift differences. The attachment of the C₂₉-C₃₄ side chain to the backbone structure were confirmed by the presence of an HMBC correlation between H₁₅ and C₂₉, whereas the position of the C₂₃-C₂₈ side chain is based on the chemical shifts of H₁₄/C₁₄ and comparison with the assignment obtained from fumonisin B₁ and iso-B₁. The backbone fragments C₁-C₆ and C₁₀-C₂₀ were unambiguously assigned and only showed minor deviations compared to fumonisin B₁ and iso-B₁. Fumonisin B₁ and iso-B₁ have OH-groups attached to C₃, C₅, and C₁₀ and to C₃, C₄, and C₁₀, respectively, whereas fumonisin B₆ has OH-groups at C₇, C₈, and C₁₇ (figure 4.8). Chemical shift values of 1.25 ppm for H₁₀ and 27.5 ppm for C₁₀, DQF-COSY connectivities from H₁₁ to H₁₀ and HMBC correlations from H₉ to C₁₀ and from H₁₀ to C₉ confirm that no OH-group is bound at this position due to the large upfield shift of the resonances when compared to fumonisin B₁, where H₁₀ and C₁₀ resonances at 3.62 and 69.8 ppm, respectively. DQF-COSY connectivities can be traced along the backbone from H₃ at 3.64 ppm, H₄ at 3.46 ppm, and H₅ at 3.71 ppm to H₆ at 1.53/1.59 ppm. In addition HMBC correlations can be seen from H₃ to C₄/C₅ and from H₅ to C₆, confirming the presence of the third OH-group at C₆ in the structure. The chemical shifts obtained for H₄/C₅ of 3.46 ppm/74.9 ppm compared to 3.65 ppm/72.2 ppm as seen in iso-fumonisin B₁ and equivalently for H₅/C₆ of 3.71 ppm/71.4 ppm compared to 3.84 ppm/68.4 ppm for fumonisin B₁ is in agreement with the proposed structure of fumonisin B₆.

The NMR data of the isolated fumonisin B₂ and the authentic standard from Fusarium were compared to establish whether they were identical. The chemical shifts were comparable except for the shifts of the two tricarboxylic acid side chains. The differences could be explained by a difference in the structural fold of the fumonisin B₂ between the two samples, likely to be caused by differences in pH and thereby overall level of protonation of the four carboxylic acid functionalities. The fact that fumonisin B₂ does have a unique globular folded structure has previously been addressed by Beier and Stanker (Beier 1997). The fumonisin B₂ from A. niger was purified on a cation exchange column and to ensure that the two samples had been subjected to the same purification protocol the standard fumonisin B₂ were also run over the column. The NMR spectra of the two fumonisins were now identical. The coupling constants for the protons in the amino terminal (C₂-C₅) and the attachment of the side chains (C₁₄ and C₁₅) in the fumonisins from A. niger and Fusarium showed that they are of the same size, proving that the relative stereochemistry for these parts of the backbone are the same. The stereochemistry at position C₁₂ and C₁₆ could not be determined from the data due to chemical shifts overlap. The optical rotation of fumonisin B₂ from A. niger was similar to that known from Fusarium; however, with several stereocenters in the molecule this is not a direct proof of identical stereochemistry. The data points toward identical stereochemistry of the two fumonisins; however, a genetic comparison of the fumonisin gene cluster in the two genera revealed several differences in the placement and orientation of the different involved biosynthetic genes (Baker 2006) which could lead to differences in the stereochemistry in the final product.
4.2.3 Summary and part conclusion

Two fumonisin isomers were isolated and structure elucidated from an extract of *A. niger* where one was a new analogue. The absolute stereochemistry of both fumonisin B₂ and B₆ still needs to be determined either through crystallization or derivation studies. The production of fumonisins in *A. niger* is an example of how the genome sequences followed by the prediction of gene function revealed the possible production of a mycotoxin in a species, which is present in both food and feeds and which is used in the industry. The reason for the fumonisins not having been reported in *A. niger* before may be that it was not expected for them to be produced and that the extraction and detection methods are different from other mycotoxins.

4.3 *Aspergillus nidulans*

*Aspergillus nidulans* is one of the most significant biological model systems in the fungal kingdom. This was pioneered by Pontecorvos’ (Pontecorvo 1953) work in the middle of the last century, which demonstrated that *A. nidulans*, in addition to the asexual state, also proliferate via sexual and parasexual life cycles giving an ideal platform for genetic studies. Due to the wide uses of *A. nidulans* as a model organism, efficient tools for targeted gene transformations have been developed (Nayak 2006, Nielsen 2006, Nielsen 2008) and the availability of the genome sequences has facilitated the possibilities of genome mining (Galagan 2005). Prior to the release of the genome sequence the research of secondary metabolism in *A. nidulans* was focused on a few biosynthetic pathways whereas the focus have now been shifted towards discovery and linking of metabolites to the silent genes.

The aim of the work in this section was to discover and link new secondary metabolites, especially PKs, of *A. nidulans* to their synthase genes. The strategies for activation of the secondary metabolite genes were the OSMAC approach, where the reference strain was grown on up to eight different media (section 4.3.1, paper 4), and through genetic modification (section 4.3.2, paper 5). These two approaches were combined and led in two studies to the discovery of metabolites not previously described from *A. nidulans* (section 4.3.3, paper 6 and 7). The results are described and further experimental details can be found in the papers.

4.3.1 Discovery and linking of polyketides to their PKS through a deletion library

As described in section 2.2.2 there are several ways to activate secondary metabolite genes in filamentous fungi. Due to previous experience at CMB with cultivating filamentous fungi at various media conditions we decided to use the OSMAC approach to possible enhance secondary metabolite production of *A. nidulans* (Frisvad 1981, Frisvad 1983, Filtenborg 1990, Frisvad 2007). Reports of novel metabolites from *A. nidulans* for example the aspoquinolones (Scherlach 2006) and aspernidine A and B (Scherlach 2010) indicated that this was a reasonable approach.

Three base media, Czapek yeast agar (CYA), Raulin-Thom (RT) and yeast extract (YE), were chosen and these were varied through the addition of salt (CYAs), sucrose (YES, CY20) and oatmeal (RTO). In addition a minimal media (MM) was used, giving eight different media in total. The media were selected based on the knowledge at CMB to ensure exploration of different growth conditions.
4 | Results and discussion

The reference strain, IBT29539 were cultivated on the eight media and inoculated for seven days at 37°C in the dark. Figure 4.9 show cultures of the strain on the different media.

![Figure 4.9](image)

Figure 4.9 The reference strain cultivated on eight different media: RTO, RT, CY20, CYAs, CYA, YES, YE and MM.

The cultures appear quite different at the different cultivation conditions. To analyze the metabolite production micro-extraction was performed (Frisvad 1987, Smedsgaard 1997) and the extracts were analyzed by both UHPLC-DAD and LC-TOF-MS. The UHPLC analysis revealed a diverse production of secondary metabolites on the eight different media, figure 4.10.

![Figure 4.10](image)

Figure 4.10 UHPLC-DAD chromatograms (210 nm) of extracts of the reference strain cultivated on eight different media compositions demonstrate the difference in secondary metabolite production. From bottom to top: MM, YE, YES, CYA, CYAs, CY20, RT and RTO. Note that the emericellamides are not visible in the UHPLC chromatogram (210 nm); however, they were detected in ESI+ and ESI- in all media except YE.
The secondary metabolite production was highly diverse and formed the basis for isolation and characterization of several metabolites which will be elaborated on the following sections. The metabolic profiles of MM and YE were considered to be redundant and these two media were therefore not used for further screening.

To investigate whether any of the metabolites in figure 4.10 were PKs we decided on a strategy to make a global deletion library where all 32 putative PKS encoding genes, as defined from the annotation of the genome databases at the Broad Institute of Harvard and MIT and the Aspergillus Genome Database at Stanford (Arnaud 2012), were deleted leaving us with 32 strains to be analyzed on the six chosen media. The deletion samples were prepared using the same micro-extraction procedure as for the reference strain. Since the deletion of a PKS gene should abolish the production of the metabolite, we hypothesized that comparison of a deletion strain to the reference would show missing peaks. To verify the approach we included deletions of genes which had already been linked, by others, to metabolites. Figure 4.11 show the UHPLC chromatogram of the reference strain compared to a deletion of stcA which is the PKS gene responsible for the biosynthesis of the sterigmatocystin backbone (Yu 1995). It is seen that the large peak at 4.6 minutes disappears. To confirm this result the same strains are analyzed with LC-TOF-MS where an extracted ion chromatogram of the [M+H]$^+$ of sterigmatocystin, 325 (±0.5 Da), disappeared in the deletion strain (data not shown).

![Figure 4.11 UHPLC chromatogram of the stcAΔ strain compared to the reference strain on CYA.](image)

This analysis was also performed on the remaining eight PKS encoding genes which, at the time, had been linked to metabolites. Metabolites representing three, beside sterigmatocystin, of the eight gene clusters could be detected: monodictyphenone, orsellinic acid and emericellamide. In agreement with previously published data the metabolites were absent in the corresponding single PKS gene deletion mutant strains; mdpG (Bok 2009), orsA (Schroeckh 2009) and easB (Chiang 2008), respectively (data not shown). Beside the secondary metabolite effect seen in these four strains, the strain carrying the wAΔ mutation formed white conidiospores, figure 4.12. Formation of the white spores occurred as it fails to produce the naphthopyrone, YWA1, the precursor of green conidial pigment (Watanabe 1998, Watanabe 1999).
Results and discussion

Figure 4.12 The wΔ strain on RTO, YES and CY20 medium.

The compounds resulting from the remaining three PKS gene clusters were originally identified through activation of the gene clusters by controlled expression of the transcription factor gene (Bergmann 2007, Chiang 2009) or by deletion of the sumO gene (Szewczyk 2008) and were not expressed under our growth conditions.

The metabolic profiles of the remaining 24 strains were compared to the reference strain on all media to uncover novel genetic links between PKS genes and PKs. The UHPLC-DAD chromatograms were first compared and the interesting peaks were afterwards identified on the LC-TOF-MS. The measured \( m/z \) was used to perform dereplication by use of the natural product database AntiBase (Laatsch 2010) which contains around 37000 metabolites. The hits were subsequently sorted based on the UV-data and reports in the literature. The novel and linked metabolites were isolated and structure elucidated through NMR analysis. UHPLC-DAD was chosen as the initial analysis based on the conjugated nature of PKs and better separation of metabolites than on the LC-TOF-MS. When a metabolite had been identified in the reference strain and corresponding peak was missing in a deletion strain, it was confirmed by HR-MS data through an extracted ion chromatogram. Through the comparison, several PKs which could be linked to PKS genes were identified; arugosin and shamixanthones \((\text{mdpG})\), orsellinic acid \((\text{AN7903 and orsA})\) as well as austinol and dehydroaustinol \((\text{AN8383})\) as described in detail below.

4.3.1.1. Identification of the austinol/dehydroaustinol synthase gene

The meroterpenoids austinol and dehydroaustinol and related metabolites were first isolated from \( A.\ ustus \) by Simpson and co-workers in 1982 (Simpson 1982) and reported from \( A.\ nidulans \) in 2007 (Márquez-Fernández 2007). Austinol and dehydroaustinol are examples of a group of meroterpenoids which Simpson and co-workers through labeling studies showed to be derived from the polyketide 3,5-dimethylorsellinic acid (Simpson 1981, Scott 1986, Geris 2009), however, the genes responsible for the production was not identified.

In this study production of austinol and dehydroaustinol was observed on all the six media used for screening of the deletion library. The structures of austinol and dehydroaustinol were confirmed through LC-TOF-MS and for dehydroaustinol through extraction, purification and NMR analysis. Interestingly we noticed that on RTO five of the single PKS deletion strains, AN0523Δ, \( pkbA\Delta \) \((\text{AN6448}\Delta)\), AN7022Δ, \( ausA\Delta \) \((\text{AN8383}\Delta)\) and \( adpA\Delta \) did not produce austinol and dehydroaustinol, figure 4.13 indicating a cross regulation of austinol and dehydroaustinol production involving these five genes. When the rest of the media was examined it was only \( ausA \) which failed to produce austinol and dehydroaustinol, see figure 4.13 for YES medium.
Figure 4.13 UHPLC-DAD chromatograms (210 nm) showing the presence of austinol and dehydroaustinol in AN0523Δ, pkbΔ (AN6448Δ), AN7022Δ, ausAΔ (AN8383Δ) and adpAΔ compared to the reference strain on RTO and YES medium.

To confirm that the loss of austinol and dehydroaustinol production was directly due to the ausA deletion rather than to silencing of another gene caused by chromatin changes provoked by the orsA deletion, a point mutation which has been shown to disrupt the enzymatic activity of the PKS production was created (Evans 2008). Like with the ausAΔ strain the production of austinol and dehydroaustinol was completely abolished on all six media (data not shown) showing that AusA is directly involved in the production of austinol and dehydroaustinol. To identify the product of AusA we overexpressed the gene by creating a strain which expressed ausA controlled by an inducible promoter, alcA, in integration site I (Hansen 2011). On inductive medium the LC-TOF-MS analysis showed a large peak eluting at around 6 minutes which was also present in the reference strain in small amounts but not produced in the deletion strain, figure 4.14.

Figure 4.14 Extracted ion chromatograms of 3,5-dimethylorsellinic acid (m/z 195 ±0.5 Da) [M-H]- of a) the reference strain and ausAΔ on RTO medium and b) the reference strain and the ausA overexpression strain (Oex) on induction medium (glycerol with threonine).
4 | Results and discussion

The metabolite was purified and the structure was elucidated by NMR to be 3,5-dimethylorsellinic acid which is therefore believed to be the first step in the biosynthesis of austinol and dehydroaustinol as shown in figure 2.8 in section 2. This result confirms the results of the labeling experiments performed by Simpson and co-workers 30 years ago (Simpson 1981, Scott 1986). After our publication Lo and co-workers have shown that the conversion of 3,5-dimethylorsellinic acid into austinol and dehydroaustinol is catalyzed by enzymes whose genes are present in two separately located gene clusters (Lo 2012).

4.3.1.2 Linking of arugosin to the monodictyphenone gene cluster

The deletion which gave the most prominent change in metabolite profile was mdpGΔ which, in an earlier study by Bok and co-workers, had been activated through epigenetic modulation of the chromatin landscape by the deletion of cclA resulting in the linking of eight emodin analogues to MdpG (Bok 2009, Chiang 2010b). Figure 4.15 show a segment of the UHPLC chromatogram of mdpGΔ and the reference strain on RTO medium where the largest differences are observed.

![Figure 4.15](image)

Figure 4.15 UHPLC-DAD chromatogram (210 nm) from four to eight minutes of the reference strain (bottom) and mdpGΔ (top) on RTO medium. The deletion eliminates the production of arugosin H, arugosin A, emericellin, shamixanthone and epi-shamixanthone as well as some unidentified peaks.

Several peaks disappeared, the most prominent being the broad peak which elutes at 7.2 minutes. This compound had a characteristic UV-spectrum, figure 4.16a, and through analysis of the adducts present in the MS-spectrum, figure 4.16b, the peak was identified to have an m/z 425.214 for [M+H]+. A search in the AntiBase database (Laatsch 2010) tentatively gave 32 possible hits based on the molecular ion. To narrow the possible number of candidates the fragment pattern was investigated and an ion of m/z 357.161 was also observed which corresponds to [M+H-C₅H₈]⁺ where C₅H₈ could correspond to a prenyl group. This information excluded metabolites of the 32 candidates to four, arugosin A, B, C and variecoxanthone C since they were the only compounds containing a labile prenyl group. The arugosins had the same UV-chromophore as reported in the literature (Ballantine 1970, Kawahara 1988a).
Figure 4.16 a) UV-spectrum of arugosin A b) MS-spectrum of arugosin A, top: ESI+ bottom: ESI-.

A large scale extraction of the reference strain on RTO medium was made for isolation of arugosin A. The NMR data were in agreement with the data reported in the literature (Kawahara 1988a) of the hemiacetal form of arugosin A except that in DMSO-\textit{d}_6 the equilibrium is shifted completely to the open form. In contrast the NMR data, in methanol-\textit{d}_4, figure 4.17, showed that arugosin A exists in an equilibrium between the open and closed form, explaining the broad peak observed in the UHPLC-chromatograms. Arugosin A had not previously been reported from \textit{A. nidulans} but from \textit{A. rugulosus} and \textit{A. silvaticus} (Ballantine 1970, Kawahara 1988a).

Figure 4.17 $^1$H-NMR spectrum of arugosin A in a) methanol-\textit{d}_4 and b) DMSO-\textit{d}_6.
Further investigations of the metabolites disappearing in the \textit{mdpGΔ} strain revealed a minor peak which could be assigned as a mono-prenylated arugosin with a [M+H]$^+$ of \textit{m/z} 357.131, arugosin H, figure 4.18. The MS data did not show fragments due to loss of a prenyl, as in the case of arugosin A, indicating that the peak is a precursor to arugosin A and that arugosin H is C-prenylated since this is less labile than an O-prenyl. Arugosin H have along with arugosin G been reported from \textit{Emericella nidulans var. acristata} (Krajl 2006) Emericellin and shamixanthone have previously been reported from \textit{A. nidulans} (Márquez-Fernández 2007) and labeling studies have shown that the arugosins are precursors for emericellin and the shamixanthones but have not established a genetic link (Ahmed 1992), as done in this study. A recent paper of Simpson (Simpson 2012) have elaborated on the information published, both through chemical and genetic studies and proposed a biosynthesis of the shamixanthones combining recent findings with the information from the biosynthesis established through labeling studies 20 years ago (Ahmed 1992).

The \textit{mdpG} gene cluster is responsible for the biosynthesis of a family of structurally similar compounds which is not unusual (Walsh 2002, Kroken 2003, Frisvad 2004b, Amoutzias 2008).

4.3.1.3. Linking of orsellinic acid derived metabolites to the orsellinic acid gene cluster to and AN7903

One of the most studied PK families in \textit{A. nidulans} is the products derived from the PK of the PKS encoding \textit{orsA}, orsellinic acid. The metabolites F-9775A and F-9775B, lecanonic acid, gerfelin, C10-deoxy-gerfelin, diocinol, orcinol cordyol C, violaceol I, violaceol II, sanghaspirodin A and sanghaspirodin B have all been shown to be linked to \textit{orsA} (Bok 2009, Schroechk 2009, Sanchez 2010, Nahlik 2010, Nützmann 2011, Scherlach 2011).
In this study orsellinic acid, lecanoric acid, F-9775A, F-9775B, violaceol I and violaceol II, figure 4.19, were present in the extracts from the reference strains on the complex media and the production was abolished in the orsAΔ strain. Interestingly the metabolites which disappear in the orsAΔ strain also disappear in the AN7903Δ strain. This result does not contradict the original assignment of orsA as the PKS gene responsible for production of orsellinic acid. AN7903 is larger than orsA and contains a methyl-transferase domain, which is not required for orsellinic acid production. Schroeckh et al. (Schroeckh 2009) observed that both AN7903 and orsA were upregulated when orsellinic acid was induced by co-cultivation with *Streptomyces hygroscopicus*, which could indicate either cross-talk between the two clusters or induction of expression of both clusters due to interactions. What appears to be trace amounts of orsellinic acid could be detected in both the AN7903Δ and the orsAΔ strains. The source of the orsellinic acid remains elusive but could be produced by other PKSs. Ahuja and co-workers reported the production of orsellinaldehyde from PkfA (Ahuja 2012). It is interesting to note that in most of the strategies for activating secondary metabolite production through epigenetic modifications, both chemical and molecular, in *A. nidulans* the products of the orsellinic acid cluster seems to be activated.

### 4.3.2 Characterization of the 3-methylorsellinic acid gene cluster

*Aspergillus nidulans* has been shown to produce at least four variants of orsellinic acid; orsellinic acid (OrsA (Schroeckh 2009)), 3-methylorsellinic acid (Hansen 2011), 3,5-dimethylorsellinic acid (AusA (Nielsen 2011)) and orsellinic acid aldehyde (PkfA (Ahuja 2012)). The detection of 3-methylorsellinic acid reported by Hansen et al. (Hansen 2011) led us to examine the extracts of the reference strain used for creation of the PKS deletion library described in section 4.3.1. The extracts of the deletion library was analyzed on the LC-TOF-MS; however, due to the better mass accuracy and sensitivity of the newly acquired UHPLC-TOF-MS this instrument was used for the extract analysis. We found 3-methylorsellinic acid to be produced in trace amounts on CYA, CY20 and RT medium and of the 32 deletions it was only in AN6448Δ (pkbΔ) that production failed. Figure 4.20 shows EICs of a 3-methylorsellinic acid [M+H]⁺ (m/z 183.06483±0.001) standard, and the reference and pkbΔ strains cultivated on CYA medium. This data corresponds to the recently published findings by Ahuja and co-workers that the PKS PkbA is required to produce 3-methylorsellinic acid along with cichorine (Ahuja 2012).
4 Results and discussion

![Figure 4.20](image)

Figure 4.20 EIC of 3-methylorsellinic acid [M+H]^+ (calc. m/z 183.06483±0.001) (UHPLC-TOF-MS)) of a standard, the reference and pkbAΔ strains on CYA medium cultivated for seven days at 37°C.

As production of 3-methylorsellinic acid was relatively low under our cultivation conditions we searched the genome sequence for neighboring genes predicted to encode for transcriptional regulation. AN6446 was predicted to have DNA binding activity and we speculated that this could be a cluster-specific transcription factor and constructed an AN6446 overexpression strain (AN6446-Oex) in IS1 (Hansen 2011). The overexpression of AN6446 clearly had an effect on *A. nidulans*, since the AN6446-Oex compared to the reference displayed a decrease in pigmentation of conidia to a dusty green appearance and a light brown center of the colony after seven days growth on MM, figure 4.21.

![Figure 4.21](image)

Figure 4.21 The reference strain compared to the AN6446-Oex strain cultivated on MM at 37°C for three, five days, and seven days.

To analyze the metabolite production micro-extraction was performed (Frisvad 1987, Smedsgaard 1997) and the extracts were analyzed by UHPLC-TOF-MS, figure 4.22. The strains were extracted
after three, five, and seven days to test the effect of growth time on the production of secondary metabolites.

**Figure 4.22** UHPLC-TOFMS ESI+ BPC of the micro-extracts of the reference and AN6446-Oex strains cultivated on MM media at 37°C for a) three, b) five, and c) seven days. The metabolites isolated in this study are marked. *This metabolite is the siderophore triacetylfusarine.*
4 | Results and discussion

The UHPLC-TOF-MS analysis revealed several new peaks appearing on the AN6446-Oex strain compared to the reference indicating that the overexpression had indeed turned on the likely 3-methylorsellinic acid biosynthetic pathway. 3-methylorsellinic acid was present in trace amounts in the AN6446-Oex strain. The metabolic profile seemed to be rather similar for the three different days although new peaks were apparently more dominant after seven days and the peak at 4.1 which was identified to be the siderophore triacetylfusarine was less dominant. Consequently the strains were grown for seven days throughout the study. Moreover this showed that MM medium was a suitable growth medium for isolation of candidate metabolites.

A large scale extract of 200 petri dishes carrying the AN6446-Oex strain was thus prepared for isolation and structure elucidation of several metabolites. In figure 4.23, the UHPLC-TOF-MS ESI+ BPC of the large extract compared to the micro-extracts is shown.

![Figure 4.23 UHPLC-TOF-MS ESI+ BPC of the micro-extract (bottom) and 200 plate extract (top) of the overexpression of AN6446-Oex strain cultivated on MM medium for seven days at 37°C.](image)

Sample preparations and cultivation conditions of a large number of plates could influence the loss of metabolites observed; however, further analysis revealed more metabolites than visible in the chromatogram as cichorine were present in very high concentrations. Isolation and structure elucidation revealed the presence of several metabolites; cichorine, demethylnidulol and 4-hydroxy-3,5-dimethyl-2-pyrone and two dimers of cichorine and nidulol/demethylnidulol, named cichonidulol and demethylcichonidulol figure 4.24. Throughout the study several deletion strains were constructed where some of these will not be mentioned further; however, nidulol and 3-methylorsellinic acid was isolated from two of these, AN6444Δ and AN6445Δ respectively. Cichorine and 3-methylorsellinic was recently reported to be the products of the overexpression of the AN6448 PKS (Ahuja 2012). Cichorine has previously been isolated from A. silvaticus (Kawahara 1988b) and Alternaria cichorii (Stierle 1993) whereas nidulol have been reported from A. nidulans (Aucamp 1968), A. silvaticus (Fujita 1984), and Emericella desertorum (Nozawa 1987).
Figure 4.24 Structures of metabolites isolated and elucidated throughout this study. 3-methylorsellinic acid, cichorine, demethylcichorine, nidulol, demethylnidulol, cichonidulol, demethylcichonidulol, 4-hydroxy-3,6-dimethyl-2-pyone.

The structures were elucidated using 2D NMR spectroscopy, where especially heteronuclear HSQC and HMBC spectra were extensively used due to the absence of proton-proton couplings and low proton density of the metabolites. For further details on NMR analysis and 1D spectra see paper 5. In addition the results were in accordance with chemical shift values reported in literature for cichorine (Monreau 2005), 3-methylorsellinic acid (Ahuja 2012), demethylnidulol (El-Feraly 1985), and 4-hydroxy-3,6-dimethyl-2-pyone (Savard 1994b).

The resemblance of the molecular formulas of cichorine (C_{10}H_{11}NO_{3}) and nidulol (C_{10}H_{10}O_{4}) and initial comparison of the $^1$H-spectra that suggested these metabolites could be related. The main issue in structure elucidation of these two metabolites was whether the carbonyl was situated on C_7 or C_8. The problem was solved by comparison of the carbon chemical shift of cichorine to the shifts reported in the literature obtained after total synthesis of cichorine (Moreau 2005). For all the metabolites key HMBC correlations from H_5 to C_7 and/or C_8 and from H_7/H_8 to the carbons present in the aromatic ring, figure 4.25, aided in the structure elucidation. A NOESY correlation between H_8 and H_10 in cichorine which were missing in nidulol confirmed the difference in the two structures.
Results and discussion

Figure 4.25 Key HMBC correlations of cichorine, nidulol, cichonidulol and 4-hydroxy-3,6-dimethyl-2-pyrone.

As these metabolites are biosynthetic derivatives of 3-methylorsellinic acid it was expected that the carbonyl would be present at C_8. These structures indicate that more enzymatic activities are involved in the biosynthesis due to the oxidations at C_7.

The UV spectrum of demethylnidulol was identical to the one of nidulol and the molecular formula C_{20}H_{21}NO_{6} showed that the difference between the two metabolites was a carbon and two protons. Through analysis of the NMR-spectra it was confirmed that the two metabolites were related and that the difference was the methylation of C_3 in nidulol which is not present in demethylnidulol.

Table 4.1 13C-chemical shifts (ppm) of 3-methylorsellinic acid, cichorine, nidulol, demethylnidulol, and cichonidulol. All data were obtained in DMSO-d_6 and referenced to the solvent peak at 39.5 ppm.

<table>
<thead>
<tr>
<th>ID</th>
<th>3-methylorsellinic acid</th>
<th>Cichorine</th>
<th>Nidulol</th>
<th>Demethylnidulol</th>
<th>Cichonidulol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1'</td>
<td>103.7/-</td>
<td>123.1/-</td>
<td>107.2/-</td>
<td>92.9/-</td>
<td>121.3/116.1</td>
</tr>
<tr>
<td>2/2'</td>
<td>163.0/-</td>
<td>153.6/-</td>
<td>157.5/-</td>
<td>159.6/-</td>
<td>153.6/153.8</td>
</tr>
<tr>
<td>3/3'</td>
<td>107.7/-</td>
<td>118.9/-</td>
<td>117.5/-</td>
<td>101.1/-</td>
<td>120.5/115.3</td>
</tr>
<tr>
<td>4/4'</td>
<td>159.7/-</td>
<td>156.3/-</td>
<td>162.7/-</td>
<td>168.8/-</td>
<td>156.9/158.1</td>
</tr>
<tr>
<td>5/5'</td>
<td>109.9/-</td>
<td>102.9/-</td>
<td>103.0/-</td>
<td>103.2/-</td>
<td>103.4/102.4</td>
</tr>
<tr>
<td>6/6'</td>
<td>139.4/-</td>
<td>131.9/-</td>
<td>148.6/-</td>
<td>151.4/-</td>
<td>131.0/139.5</td>
</tr>
<tr>
<td>7/7'</td>
<td>214.4/-</td>
<td>169.8/-</td>
<td>68.2/-</td>
<td>66.9/-</td>
<td>167.1/71.7</td>
</tr>
<tr>
<td>8/8'</td>
<td>173.8/-</td>
<td>43.1/-</td>
<td>168.5/-</td>
<td>178.1/-</td>
<td>43.2/83.5</td>
</tr>
<tr>
<td>9/9'</td>
<td>7.9/-</td>
<td>9.1/-</td>
<td>8.2/-</td>
<td>-/-</td>
<td>9.3/8.6</td>
</tr>
<tr>
<td>10/10'</td>
<td>-/-</td>
<td>58.9/-</td>
<td>53.9/-</td>
<td>53.9/-</td>
<td>59.0/59.2</td>
</tr>
</tbody>
</table>

Through analysis of the large scale AN6446-Oex extract we noticed two small peaks eluting around five minutes which had UV-spectra that resembled those of nidulol and cichorine. These metabolites were present in small amounts in the large extract and based on the molecular formula of these metabolites C_{20}H_{21}NO_{6} and C_{19}H_{19}NO_{6} we speculated whether they could be dimers of cichorine, 3-methylorsellinic acid, nidulol or demethylnidulol through condensation and formation of e.g. an ester as seen with the formation of lecanoric acid in the orsellinic acid pathway (Schroeckh 2009). 2.0 mg of C_{20}H_{21}NO_{6} and 0.6 mg of C_{19}H_{19}NO_{6} were isolated from the extract and the structure of C_{19}H_{19}NO_{6} was elucidated based on 2D NMR analysis to be a dimer of cichorine and nidulol, named cichonidulol. Analysis of the HMBC spectrum revealed that the link of the two molecules was not through an ester formation but interestingly through a C-N bond formed apparently with the loss of a carbonyl in nidulol. A search of the literature showed that cichonidulol and demethylnidulol were novel metabolites with a new base carbon skeleton. The metabolite which showed the closest resemblance was the plant metabolite, capaurine, which
contains the C-N condensation bond between two ringsystems; however, the nitrogen containing ring is a six membered ring and the substitution patterns of are quite different (Manske 1945).

Table 4.1 lists the $^{13}$C-chemical shifts of the above mentioned isolated metabolites where it is seen that the largest differences in shift values occur around at C$_1$ and C$_6$ which differs depending on the position of the carbonyl.

The final metabolite isolated was 4-hydroxy-3,6-dimethyl-2-pyrene whose chemical shifts were in accordance with shifts reported in the literature (Savard 1994b). This metabolite is quite different from the above mentioned having a completely different carbon skeleton. Consequently this metabolite was either a product of pkbA or we could have activated another PKS encoding gene.

To verify that AN6446 had an activator effect on the surrounding genes, relative gene expression measurement by qRT-PCR analysis was performed. A deletion of AN6446, which displayed the same phenotype, as the reference was added to the analysis with the hypothesis that this would have the inverse effect from the overexpression strain minimizing the risk of picking false cluster members. AN6446-Oex and AN6446Δ strains compared to the reference strain was performed for all predicted genes in the range AN6435-AN6457, figure 4.26.

![Figure 4.26](image)

**Figure 4.26** Results of the qRT-PCR analysis of AN6446-Oex and AN6446Δ on the genes from AN6435-AN6457 compared to the reference strain. An expression of 1 means unaltered expression, over 1 means higher expression and below 1 means lower expression. These data show that AN6444, AN6446, AN6447, AN6448, AN6449, AN6450 and AN6451 are upregulated in the overexpression strain and downregulated in the deletion strain indicating that the pkbA cluster contains these genes.

The average fold change of three runs of all 25 genes in the AN6446Δ and AN6446-Oex strains compared to the reference is shown in figure 4.26. A level of one means unaltered expression whereas a level below means lower expression and above means higher. For a gene to be a putative cluster member one criterion was to have fold change above one for the AN6446-Oex
and close to zero for the \( \Delta \) strains. Deviation from one or both of these criteria made us disregard the candidate as a cluster member based on the gene expression data.

Based on this analysis the values of the genes AN6444, AN6447, AN6448, AN6449, AN6450 and AN6451 along with AN6446 were considered as cluster candidates. To verify the results and to determine the biosynthetic role of the enzymes encoded by these genes we constructed deletions of all the 25 initial candidate genes in the AN6446-Oex background. The six most distant genes of AN6446 were deleted in triplets (AN6435-AN6437, AN6438-AN6440, AN6452-AN6454 and AN6455-AN6457) and the remaining as singular mutants.

To see whether AN6446 had regulatory effects on secondary metabolite production outside the cluster, a full cluster deletion (AN6435-AN6457) in the overexpression strain was constructed as well. Lastly to confirm the role of \( \text{PkbA} \) as the 3-methylorsellinic acid producing PKS an overexpression of this gene was constructed.

![Figure 4.27](image_url)

Figure 4.27 The appearance of the AN6446-Oex, \( \text{pkbA-Oex} \), AN6446-Oex AN6435-6437\( \Delta \), AN6446-Oex AN6438-6440\( \Delta \), AN6446-Oex AN6441\( \Delta \), AN6446-Oex AN11920\( \Delta \), AN6446-Oex AN11921\( \Delta \), AN6446-Oex AN6443\( \Delta \), AN6446-Oex AN6444\( \Delta \), AN6446-Oex AN6445\( \Delta \), AN6446-Oex AN6446\( \Delta \), AN6446-Oex AN6447\( \Delta \), AN6446-Oex AN6448\( \Delta \), AN6446-Oex AN6449\( \Delta \), AN6446-Oex AN6450\( \Delta \), AN6446-Oex AN6451\( \Delta \), AN6446-Oex AN6452-6454\( \Delta \), AN6446-Oex AN6455-6457\( \Delta \) and AN6446-Oex AN6435-6457\( \Delta \) strains cultivated on MM medium for seven days at 37°C.
The phenotypes of the constructed strains can be compared in figure 4.27. The deletion of the genes, predicted to form the gene cluster of PkbA, indeed stood out from the others in the phenotype display. The phenotypes varied greatly, indicating altered metabolism in the respective strains. To investigate which metabolites originated from this cluster and if correlated with the phenotype of the mutant strains micro-extraction was performed (Frisvad 1987, Smedsgaard 1997) and the extracts were analyzed by UHPLC-TOF-MS. In figure 4.28, 4.29 and 4.30 ESI+ BPC of all the extracts can be seen.

Figure 4.28 UHPLC-TOF-MS ESI+ BPC of the micro-extracts of the (from bottom to top) reference AN6446-Oex, AN6446-Oex AN6436-6457Δ, AN6446-Oex AN6436-6440Δ, AN6446-Oex AN6441Δ, AN6446-Oex AN11920Δ, AN6446-Oex AN11921Δ, AN6446-Oex AN6435-6457Δ strains cultivated on MM medium for seven days at 37°C. The metabolites identified in the AN6446-Oex strain are missing in all the deletions.
In figure 4.28 the strains where all the metabolites of the cluster had disappeared and the AN6446-Oex strain are depicted. In all the deletions from AN6435 to AN11921 all the key metabolites shown in the AN6446-Oex strain are missing. Since the qRT-PCR data did not support these belonging to the cluster we speculate that this links to local chromosomal aberration due to change in the strain. Inactivating one or some of these genes instead with point mutations would test this hypothesis.

Moreover, the phenotypes of these strains were very similar as the metabolic profiles were. The strains all showed the same differences from the reference strain, namely the massive production of the siderophore triacetylfulusarinine.

In the whole cluster deletion strain, AN6446-Oex AN6435-6457Δ, all metabolite production is eliminated including 4-hydroxy-3,6-dimethyl-2-pyrene strongly indicating that this metabolite is also a product of the cluster. However, eliminated production may reflect the deletions of AN6435 to AN11921. A cluster deletion from AN6444 to AN6451 would be an alternative. A thorough investigation of the PKS deletion library described in section 4.3.1 revealed trace amounts of 4-hydroxy-3,6-dimethyl-2-pyrene on CYA medium in the reference strain. The 32 deletions on CYA medium were analyzed for 4-hydroxy-3,6-dimethyl-2-pyrene and it was only in pkbA that production failed.

Figure 4.29 UHPLC-TOF-MS ESI+ BPC of the micro-extracts of the (from bottom to top) AN6446-Oex, AN6446-Oex AN6443Δ, AN6446-Oex AN6445Δ, AN6446-Oex AN11922Δ, AN6446-Oex AN6452-6454Δ, AN6446-Oex AN6454-6457Δ strains cultivated on MM medium for seven days at 37°C. The figure illustrates that these genes are not involved in the pkbA gene cluster in agreement with the qRT-PCR data.
The extracts depicted in figure 4.29 are the strains of which the chromatograms mimic the profile of AN6446-Oex. The colonies of these strains in figure 4.27 resembled the overexpression strain, and the results are in agreement with the qRT-PCR data, which excluded these genes to be part of the 3-methylorsellinic acid gene cluster.

*This metabolite is the siderophile triacetylfusarine. The figure illustrated that these genes indeed are part of the 3-methylorsellinic acid gene cluster.*
4 | Results and discussion

The BPC chromatograms of the micro-extracts of the deletion strains predicted through the qRT-PCR analysis to belong to the 3-methylorsellinic acid cluster are depicted in figure 4.30. Detection of the metabolites isolated in this study are listed in table 4.2 along with several additional peaks appearing in the chromatograms; however, the structures of these have not yet been identified.

**Table 4.2** Detection of metabolites (through EIC+) in the micro-extracts of the deletion and overexpression strains constructed and identified as belonging to the 3-methylorsellinic clusters in this study; (from bottom to top) AN6446-Oex, AN6448-Oex, AN6446-Oex AN6444Δ, AN6446-Oex AN6447Δ, AN6446-Oex AN6444Δ AN6446-Oex AN6450Δ, AN6446-Oex AN6451Δ. Parenthesis indicates that the metabolites are present in trace amounts.*Strain having AN6446-Oex allele. *A metabolite eluting close to cichorine with a small mass difference makes further analysis needed for clarification. **Eluting at 2.9 minutes. ***Eluting at 4.7 minutes.

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>AN6446-Oex</th>
<th>pkbA-AN6444Δ</th>
<th>AN6447Δ</th>
<th>pkbAΔ</th>
<th>AN6449Δ</th>
<th>AN6450Δ</th>
<th>AN6451Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-methyl-orsellinic acid</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
<td></td>
<td>+</td>
<td>(+)</td>
<td>+</td>
</tr>
<tr>
<td>Cichorine*</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Nidolol</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demethyl-nidolol</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>(+)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Cichonidulol</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Demethyl-cichonidulol</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>4-hydroxy-3,6-dimethyl-2-pyrone</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>(+)</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>C9H10O3</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C9H15O3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(+)</td>
<td>-</td>
<td>(+)</td>
</tr>
<tr>
<td>C12H21O4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C15H19NO5</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>C10H16NO3</td>
<td>+</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

In the deletion of pkbA in the AN6446-Oex strain the isolated metabolites are all, except demethylnidolul, missing confirming that these metabolites are part of the 3-methyllellinic acid gene cluster. Demethylnidolul was detected in trace amount and could be due to the conversion of orsellinic acid into the product. A double deletion strain of orsA and pkbA could verify this hypothesis.

AN6450 is predicted to encode an oxidoreductase, and showed in the qRT-PCR analysis to be regulated by AN6446; however, the metabolite profile of AN6446-Oex AN6450Δ was equal to that of AN6446-Oex indicating that either it does not take part in the biosynthesis, or the product of this enzymatic step was undetected. Interestingly, AN6451 is predicted to encode a transporter of the Major Facilitator Superfamily and this deletion mutant is heavily impaired in growth. This indicated that some compartmentalization of the biosynthesis may occur and failure to transport either 3-methylorsellinic acid or 4-hydroxy-3,6-dimethyl-2-pyrone, or even other products in the...
pathway may be toxic for the fungus, at least under the artificial conditions that overexpression of AN6446 offers.

Table 4.3 Predicted function of the genes identified as belonging to the 3-methylorsellinic acid cluster.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Predicted function</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN6444</td>
<td>NRPS-like</td>
</tr>
<tr>
<td>AN6446</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>AN6447</td>
<td>O-methyltransferase</td>
</tr>
<tr>
<td>pkbA</td>
<td>PKS</td>
</tr>
<tr>
<td>AN6449</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>AN6450</td>
<td>Oxidoreductase</td>
</tr>
<tr>
<td>AN6451</td>
<td>Transporter</td>
</tr>
</tbody>
</table>

Based on the metabolic profile of the deletion strains a biosynthetic pathway from 3-methylorsellinic acid to cichoridulol is proposed, figure 4.31. 4-hydroxy-3,6-dimethyl-2-pyrene disappears in the AN6446-Oex pkbAΔ strain and are only present in trace amounts in the AN6446-Oex AN6447Δ strain. As oppose to 3-methylorsellinic acid the polyketide backbone seems to be C₆ and not C₈ and may be a shunt of PkbA. If the methylation pattern in the assembly of the polyketide backbones of 3-methylorsellinic acid and 4-hydroxy-3,6-dimethyl-2-pyrene were the same one would expect the latter to be methylated at C₁ and not C₃ as is observed based on careful investigation of the 2D NMR-spectra. This shows that there are several differences between the assembly of the two metabolites. Other metabolites with the same UV-spectrum as 4-hydroxy-3,6-dimethyl-2-pyrene were detected in the deletion strains but have not yet been isolated. Ahuja and co-workers (Ahuja 2012) isolated two products of the PkgA enzyme, dehydrocitreoisocoumarin and 6,8-dihydroxy-3-(2-oxopropyl)-isocoumarin where the difference of the polyketide backbone also was the incorporation of one less Mal-CoA in the latter.

The first step of the biosynthesis is likely to be oxidation and reduction of 3-methylorsellinic acid. The proposed intermediates are the 3-methylorsellinic acid in which the carboxylic acid has been reduced to an aldehyde (C₉H₁₀O₃) and the 6-methyl oxidized (C₉H₁₀O₃) compound. Due to the appearance of metabolites identified by LC-MS to have these molecular formulas in the AN6447Δ strain the next step is proposed to be O-methylation of the 2-OH catalyzed by the enzyme encoded by AN6447. This metabolite could then undergo ring closure to form nidulol. The C₁-carbonyl of nidulol, silvaticol, has also been reported in the literature isolated from A. silvaticus (Fujita 1984) and Kawahara and co-workers (Kawahara 1988b) suggest quadrilineatin as the intermediate before ring closure thereby explaining that both metabolites nidulol and silvaticol are present in A. silvaticus. Neither silvaticol nor quadrilineatin have been detected in any of the examined extracts yet.
4 | Results and discussion

**Figure 4.31** Proposed biosynthetic route of the elucidated metabolites of the pkbA cluster. The metabolites in brackets are proposed intermediates which have been detected through MS-analysis except for the dialdehyde (C$_{10}$H$_{10}$O$_4$) which have not been detected.

The metabolic profile of the AN6446-Oex AN6449Δ and AN6446-Oex AN6451Δ appear to be similar. Several peaks appear in these two deletions where the molecular formula for one of these is equal to the proposed methylated product of AN6447 (C$_{10}$H$_{12}$O$_4$) indicating that these enzymes are involved in the next steps of the biosynthesis. The roles for these two enzymes in the biosynthesis are still unknown but identification of the intermediates could provide helpful answers. AN6449 encodes for a cytochrome P450 which could be responsible for the oxidation of the C$_{10}$H$_{12}$O$_4$ to quadrilineatin proposed in figure 4.31 which could then undergo further modifications to cichorine.

In the AN6444Δ strain the production of several metabolites have terminated including cichorine and other nitrogen containing metabolites for example a metabolite eluding at 4.2 minutes with the molecular formular C$_{15}$H$_{19}$NO$_5$ (identified through UHPLC-TOF-MS). The role of AN6444 strongly indicates contribution of the nitrogen to the molecule and ring closure. As no new intermediates were detected in the extracts the exact role of the enzyme remains elusive. The nitrogen incorporated in the structure could arise from either transamination as seen in the biosynthesis of amino acids or from an amino acid which reacts with quadrilineatin. This reaction could yield a lactam where the rest of the amino acid is still attached to the nitrogen followed by
decarboxylation as seen for erinacerin A, figure 4.32, where the residual part of decarboxylated phenyalanine is attached to the lactam-N (Yaoita 2005). The above mentioned metabolite, $C_{15}H_{19}NO_5$, which was present in the AN6446-Oex and AN6446-Oex AN6449Δ strains could indeed resemble the product, figure 4.32, of these reactions if the amino acid was 2-aminoadipic acid. Bioinformatic analysis of AN6444 supports this hypothesis since it predicts the adenylating domain of AN6444 to encode for 2-aminoadipic acid. This amino acid residue would then have to be cleaved from the molecule to reach cichorine. A way to test this hypothesis is to do feeding experiments with labeled 2-aminoadipic acid. Aspernidine A and B (not detected in these extracts) (figure 4.34) are two other metabolites reported from *A. nidulans* where a nitrogen is incorporated into aromatic part of the structure and is proposed to be derived from an orsellinic acid (Scherlach 2010). The nitrogen of these metabolites could also arise from AN6444 or other NRPS-like enzymes.

![Figure 4.32](image)

**Figure 4.32** The structure of erinacerin A, 2-aminoadipic acid and proposed structure of $C_{15}H_{19}NO_5$.

It remains unknown whether the dimers, cichonidulol and demethylcichonidulol, were biosynthesized by an enzyme or the dimerization occurred non-enzymatically. Due to the small production of the metabolites in the overexpression strain the data obtained is not sufficient for identifying the potential enzyme.

Altogether this study has through qRT-PCR and subsequent chemical analysis identified the genes belonging to the 3-methylorsellinic acid gene cluster. Several metabolites have been isolated, structure elucidated and linked to the gene cluster e.g. nidulol, cichonidulol, demethylcichonidulol and 4-hydroxy-3,6-dimethyl-2-pyrene. The function of AN6447 has been shown to be O-methylation and data strongly suggests that AN6444 is involved in the incorporation of a nitrogen atom into the structure which might occur through an aminoadipic intermediate. Several interesting intermediates have been identified through UHPLC-TOF-MS analysis and remain to be isolated and structure elucidated to shed more light over the biosynthesis. A shunt product of the PKS have been identified and through analysis of the PKS deletion library and data of the strains constructed in this work it has been linked to PkbA.

### 4.3.3 Genetic activation with regulatory genes combined with complex media

The work in the previous sections indicated that it is possible to activate secondary metabolite production through both genetic and OSMAC approaches. However, as there were still a lot of silent genes we decided to look into combining these approaches of activation. In this section two examples of metabolites not previously reported from *A. nidulans* are described.
4 Results and discussion

4.3.3.1 Identification and linking of A. nidulans metabolites to genes through a combined approach of complex media and genetic modifications

Due to the availability of the PKS deletion library, described in section 4.3.1 we wanted to identify possible PK candidates to isolate, structure elucidate and, if possible, link to their synthase genes. Linking of metabolites present in small quantities in the reference strain could be possible due to the possibility of comparison with an isolated standard. As seen in the previous study of the 3-methylorsellinic acid gene cluster the disappearing of a small peak helped identify a candidate for genetic studies, resulting in the discovery and linking of a family of both known and previously undescribed metabolites to their synthase genes. In the search of new secondary metabolites of A. nidulans we investigated the IBT culture collection at CMB.

The strain chosen for this study was IBT24909 which was cultivated on 200 plates of CYAs for 13 days at 25°C for optimal production of 1, 2, and 3. The LC-DAD chromatogram at 210 nm of the extract can be seen in figure 4.33.

![Figure 4.33 LC-DAD chromatogram (210 nm) of 200 plate extract of IBT24909 cultivated on CYAs for 13 days. The two peaks 1 and 2 were identified as potential PKs. The large peak at 10.5 minutes is sterigmatocystin.](image)

After inspection of the chromatogram two rather large peaks, 1 and 2, were identified as putative PKs based on their UV-spectra, figure 4.33, which had absorption maxima at wavelengths over 300 nm indicating that the metabolite or part of it contained a highly conjugated system.

![Figure 4.34 UV-spectra of a) 1 and b) 2.](image)
Analysis of the LC-TOF-MS data revealed an [M+H]+ of 343.231 and 387.219 of 1 and 2, respectively. A search of the Antibase database (Laatsch 2010) (tolerance ± 0.05 Da) gave 29 hits for 1 and 54 hits for 2 where several contained conjugated double bonds which could explain the absorption in the UV-spectra. So to be able to identify the metabolites they were isolated and structure elucidated by 2D NMR (DQF-COSY, HSQC, HMBC, H2BC and NOESY).

Figure 4.35 Structures of nidubenzal A and B (1), asperugin A, B (2), and C, and aspernidine A and B

NMR analysis of 1 revealed that two isomers, that we named nidubenzal A and B, figure 4.35, were present in the solution in a 5:1 ratio. The difference of the chemical shift values of the two isomers were small and only visible for H₁'/C₁' and in the aromatic moiety of the molecule, figure 4.36.

Figure 4.36 ¹H-NMR spectrum of nidubenzal A and B in DMSO-d₆. Which illustrates the two metabolites present in the solution.

NMR analysis of 2 revealed that two isomers, that we named nidubenzal A and B, figure 4.35, were present in the solution in a 5:1 ratio. The difference of the chemical shift values of the two isomers were small and only visible for H₁'/C₁' and in the aromatic moiety of the molecule, figure 4.36.
4 | Results and discussion

The structure of the farnesyl side chain was deduced based on DQF-COSY, HSQC, HMBC and H2BC (for detailed NMR data see paper 6) correlations and comparison with literature data from the related structures aspernidine A and B (Scherlach 2010). The attachment of the farnesyl side chain to the two isomers was based on HMBC correlations of H₁ into the aromatic ring, figure 4.37. It had not been possible to see the two different isomers in any of the HPLC/UHPLC analysis due to co-elution; however, NMR analysis revealed the presence of them. Nidubenzal A has been reported in a synthetic screening library (Aurora Fine Chemicals LLC); however, neither nidubenzal A nor B has previously been reported from nature.

The structure of 2 was also solved through 2D-NMR analysis (DQF-COSY, HSQC, and HMBC) which identified it to be asperugin B a metabolite previously isolated from the related fungus A. rugulosus (Ballantine 1967). Two other asperugins, A and C, have been isolated from A. rugulosus (Ballantine 1964, 1965, and 1971) and through LC-TOF-MS-analysis we tentatively identified asperugin A (3), figure 4.33, but not asperugin C in the A. nidulans IBT24909 extract. We have not yet been able to isolate asperugin A in quantities sufficient for NMR-analysis. In 2005 An et al. (An 2005) reported production of asperugin A and B (also with B as the major product), as well as violaceol, in a strain where cosmid-size DNA from slow growing fungi was cloned and introduced into A. nidulans. During the course of this study Scherlach et al. reported the characterization of aspernidine A and B isolated from A. nidulans (Scherlach 2010); however, we were not able to trace these metabolites through LC-TOF-MS analysis of our extract.

All the isolated metabolites, nidubenzal A and B, asperugin A-C and the related aspernidines, seem to be hybrid metabolites which consist of a farnesyl side chain and an aromatic moiety, which may be of PK origin. The aromatic nature of the metabolites indicated that the pathway could be initiate with formation of a NR-PK catalyzed by a NR-PKS. The products of the NR-PKS could be orsellinic acid or even more likely orsellinaldehyde, figure 4.38, which have, just recently, been shown to be a metabolite of A. nidulans from the AN3230 (PfkA) NR-PKS (Ahuja 2012). The only difference between the aromatic part of asperugin A and B, and aspernidine A and B, are an O-methylation indicating that the two metabolites in each group are derived from the same biosynthetic pathway. Scherlach et al. (Scherlach 2010) suggests that the aspernidines are formed by either orsellinic acid or the corresponding aldehyde which traps ammonia nitrogen in an intermediary aldehyde, where oxidoreduction and condensation steps leads to the isoidolinone moiety as suggested for cichorine in the previous section. The precursor of the aromatic moiety of asperugin C could also be orsellinaldehyde which, in a shunt pathway, will be methylated instead of oxidated. The two new metabolites nidubenzal A and B differs in the aromatic moiety which containing seven carbons instead of the eight of the backbone of the other metabolites. This
could indicate a different biosynthetic pathway or a decarboxylation from a C8-PK product. We speculated whether these metabolites might be a product of the shikimic acid pathway since the substitution pattern resembles protocatechuic acid, figure 4.38, a shunt product from the biosynthesis of shikimic acid, differing only by the aldehyde in the place of the carboxylic acid.

Since the metabolites might be of PK origin we went back to the reference strain used in section 4.3.1 to see if whether trace amounts of the metabolites could be detected in EICs on any of the media.

Nidubenzal A and B were present in small quantities on the CYAs medium and analysis of the deletion library identified two PKS deletions, ausAΔ and AN11191Δ, where the production of nidubenzal A and B seemed to have been abolished.

Asperugin A could not be detected in any of the strains whereas asperugin B could be traced. The same procedure for identification of putative PKSs as for the nidubenzals was done and six PKS candidates were identified, AN0523 (pkdA), AN11191, AN3230 (pkfA), ausA, AN6791, and AN9005. Due to the small quantities of metabolites present in the extract it was, in some cases, difficult to determine the presence of the metabolites. It has recently been reported that AN3230 produces orsellinaldehyde and AN0523 2-ethyl-4,6-dihydroxy-3,5-dimethylbenzaldehyde, figure 4.38 (Ahuja 2012). ausA, pkfA and pkdA are NR-PKSs whereas AN11191, AN6791 and AN9005 are PR-PKSs (Sanchez 2012). These results gave several possible genes so we were not able to link the metabolites to their synthase genes.

At the same time as this study a strain where the cclA gene had been deleted activated the production of the asperugins and nidubenzals. It had already been shown that this deletion strain increased the production of secondary metabolites in A. nidulans (Bok 2009). We wanted to investigate whether challenging the cclAΔ strain on complex media could activate even more secondary metabolite production.

Figure 4.39 show the phenotypical appearance of the cclAΔ strain compared to the reference strain when grown on the same eight media as used initially in the deletion study. It is evident that the fungus was affected by the deletion, and as the reference strain it responded differently to the different medium conditions. Generally it seems that the cclAΔ strains grow slower than the reference. To analyze the metabolite production micro-extraction was performed (Frisvad 1987, Smedsgaard 1997) and the extracts were analyzed by both UHPLC-DAD and UHPLC-TOF-MS. The UHPLC analysis revealed a diverse and different production of secondary metabolites on the eight media compared to the reference strain, figure 4.40.
Figure 4.39 The reference strain compared to the cclAΔ strain cultivated for seven days at 37°C on RTO, RT, CY20, CYAs, CYA, YES, YE, and MM medium.

In all of the media except YE and CYAs there seemed to be an activation of secondary metabolite production. It was especially noticeable that many of the metabolites produced in the cclAΔ strain are derived from the orsellinic acid and monodictyphenone pathway, which was also the gene clusters activated by Bok et al. (Bok 2009). Arugosin A which was linked to the monodictyphenone pathway in section 4.3.1.2 seemed to disappear in the cclAΔ strain whereas intermediates in this pathway, for example emodin, seemed to accumulate. Most noticeable for our study, and what is focused on in this section, was the increased production of asperugin A and B as well as nidubenzal A and B on several of the media; YES, CY20, CYAs, CYA, RTO and MM, figure 4.41.
Figure 4.40 UHPLC-DAD chromatograms (210 nm) of micro-extracts of the cclΔ strain cultivated on eight different media compositions for seven days at 37°C compared to the reference strain. From top to bottom RTO, RT, CY20, CYAs, CYA, YES, YE and MM.*This peak corresponds to an internal standard of chloroamphenicol added to the extracts. The concentration of chloroamphenicol is the same in all the samples. Metabolites were identified through comparison with standards or based on exact mass and UV-data.
Results and discussion

Figure 4.41 UHPLC-DAD chromatograms (210 nm) from seven to eight minutes of micro-extracts of the cclAΔ on the eight different media shows the production of nidubenzal A and B, and asperugin A and B. From top to bottom: RTO, RT, CY20, CYAs, CYA, YES, YE and MM.

The increased production of asperugin A and B, and nidubenzal A and B in this strain allowed for creating double deletions of cclA and the candidate PKS genes identified previously. Therefore cclA was deleted in the respective PKS gene deletions identified previously. Both the NR- and PR-PKSs were included in the study since they might have an effect on asperugin production. It was originally thought that these products could be derived from orsellinic acid, and due to the activation of several orsellinic acid derived metabolites in the cclAΔ strain, a deletion of cclA in orsAΔ was also constructed. As described previously we wondered whether nidubenzal A and B could be shikimic derived so a deletion of the qutC gene which has been shown to encode for an enzyme, dehydroshikimic acid dehydratase that catalyzes the production of protocatechic acid from 3-dehydroshikimic acid in A. nidulans were included in the study (Lamb 1992).

Figure 4.42 UHPLC chromatograms (210 nm) of micro-extracts of the qutCΔcclAΔ, ausAΔcclAΔ, AN11191ΔcclAΔ and cclAΔ strains after cultivation for seven days at 37°C on CY20 medium. The other parts of the chromatogram were also identical.

The double deletion mutants made were analyzed on both UHPLC-DAD and UHPLC-TOF-MS on all eight media. As seen in figure 4.42 nidubenzal A and B were present in all the deletion strains and on the different media; however, spore-PCR analysis of the AN11191ΔcclAΔ strain showed presence of WT AN11191 nuclei indicating this strain was not suited for analysis. A new double deletion based on the original confirmed AN11191Δ strain is in progress of being constructed. As
the $qutC\Delta clA\Delta$ strain also produced nidubenzal A and B it indicates that they are derived from either AN11191 or an alternative route to protocatechuic acid exists in *A. nidulans*.

Figure 4.43 EIC of a) asperugin A (calc. [M-H]$^-$ 401.2325±0.002) and b) asperugin B (calc. [M+H]$^+$ 387.21660±0.002) of micro-extracts of the (from top to bottom) *pkdA\Delta clA\Delta, pkfA\Delta clA\Delta, ausA\Delta clA\Delta, AN6791\Delta clA\Delta, AN9005\Delta clA\Delta* and *clA\Delta* strains after cultivation on CY20 medium for seven days at 37°C.

The EICs of asperugin A and B in the double deletion mutants and the reference strain can be seen in figure 4.43. As the AN11191\Delta clA\Delta strain was not correct it have been removed from the study; however, a correct strain will be included in the analysis for the presence of the asperugins. As can be seen in the figure production of asperugin A and B are abolished in the *pkfA\Delta clA\Delta* strain and this was consistent on the eight analyzed media. Ahuja et al. (Ahuja 2012) recently showed that overexpression of this PKS with an inducible promoter led to the production of orsellinaldehyde. Further biosynthesis of orsellinaldehyde to the asperugins would require several tailoring enzymes catalyzing oxidation of a methyl to an aldehyde, oxidation of the aromatic ring, a terpene synthase and an O-methyltransferase. As suggested by Scherlach et al. (Scherlach 2010) further modifications to the aromatic moiety of the metabolites could lead to the biosynthesis of the aspernides. This biosynthetic pathway resembles in some ways the proposed biosynthesis of the cichorine gene cluster. Preliminary bioinformatics studies indicated the presence of O-methyltransferase, prenyltransferase, oxidase, cytochrome P450, and monooxidase enzymatic functions surrounding the *pkfA* gene. Construction of the double deletions of *clA* and the surrounding genes to determine the biosynthetic pathway of these metabolites including identifying intermediates are ongoing. In the AN6791\Delta clA\Delta strain it appears that the production of asperugin A and B have increased compared to the other strains. This is interesting due to the fact that the deletion was constructed based on analysis of the deletion library where the metabolites were missing in the single deletion strain, AN6791. This phenomena is seen in
4 | Results and discussion

several media (YES, CY20, CYAs and MM); however, the reason for this apparently increase in production is so far unknown.

![UHPLC chromatograms (210 nm) of micro-extracts of the orsAΔcclAΔ strain compared to the cclAΔ strain on CY20 medium.](image)

**Figure 4.44** UHPLC chromatograms (210 nm) of micro-extracts of the orsAΔcclAΔ strain compared to the cclAΔ strain on CY20 medium.

As mentioned a double deletion of orsA and cclA were included in the study. As seen from figure 4.44 the deletions did not abolish the production of the asperugin; however, the metabolites which have been shown to be derived from the orsellinic acid biosynthetic pathway were indeed absent in the double deletion strain (Bok 2009, Schroekh 2009, Sanchez 2010, Nahlik 2010).

In this study two novel metabolites nidubenzal A and B have been isolated. Asperugin A and B have been linked to the PKS encoding gene AN3230. Deletions of the AN3230 gene cluster is underway and deletions to link nidubenzal A and B to a gene is undergoing.

4.3.3.2 Activation of juvenile hormone production in *A. nidulans*

As all the genetic modifications described so far have been in the *A. nidulans* genome we decided to investigate whether heterologous expression of regulatory genes from other filamentous fungi could induce secondary metabolite production in *A. nidulans*, using *A. niger* as a test case. As we have seen a significant difference in secondary metabolite production on different complex cultivation media we tested these strains on five media; MM, CYAs, RTO, YES and OAT (oatmeal agar).

The genes selected for expression was identified based on a collection of microarray experiments from *A. niger* grown under diverse conditions to identify regulatory genes associated with predicted secondary metabolite gene clusters using a local co-expression algorithm. Seven genes associated with predicted gene clusters containing either PKSs or NRPSs were identified. All seven genes belonged to the binuclear zinc finger class which is often associated with secondary metabolism in filamentous fungi (MacPherson 2006). The seven genes were expressed individually in *A. nidulans* (Hansen 2011) and incubated on the five media for seven days at 37°C and micro-extraction was performed (Frisvad 1987, Smedsgaard 1997) and the extracts were analyzed by both UHPLC-DAD and UHPLC-TOF-MS. Of the 35 combinations of strains and growth media only one strain, grown on CYAs, had a significant impact on secondary metabolite
production, resulting in an increased production of several metabolites not previously identified from *A. nidulans*, figure 4.45. The gene which was expressed in the strain was renamed (from est_fge1_pg_C_150220, Broad annotation) Secondary Metabolism associated Regulatory protein A (SmrA). BLAST analysis revealed that *smrA* did not have a potential ortholog in *A. nidulans*.

**Figure 4.45** UHPLC-HRMS ESI+ BPC of the reference strain (bottom) and the strain where *smrA* (top) is expressed grown on CYAs at 37°C for seven days.

To purify the induced products we made a large extract and in this extract X2 was the dominant peak (data not shown). As X1 and X2 had a similar UV-chromatogram, a mass difference of 27.9948 (CO), and similar adduct patterns in their mass spectra, figure 4.46, it was speculated that these structures were related and both were purified.

**Figure 4.46** Mass spectra of X1, X2, JH-III and methyl-farnesoate.

The structures were elucidated using the following 2D NMR experiments; DQF-COSY, HSQC and HMBC. X1 was identified to be methyl (2E,6E)-10,11-epoxid-3,7,11-trimethyl-2,6-dodecadienoate (Kuhnz 1981) and X2 was indeed a formylated analogue of X1, figure 4.47. As formic acid was added to both the micro extraction and the large extract, a new micro extract was prepared.
without formic acid. X2 was not produced in this extract and must therefore result from chemical modification resulting from the presence of formic acid in the extraction process.

Figure 4.47 The structures of X1, X2, juvenile hormone III and methyl-farnesoate

The NMR data of X1 indicates that two diastereomers were present in the sample. The $^1$H- and $^{13}$C-chemical shifts differ most in the reduced end of X1 where a stereocenter (C$_{10}$) is present. The chemical shifts of H$_{1}$/C$_{1}$ to H$_{5}$/C$_{5}$ are identical. The difference of chemical shifts of the two methyl groups (H$_{12}$/C$_{12}$ and H$_{13}$/C$_{13}$) and the two CH$_2$ groups next to the stereocenter are due to the presence of the chiral center. The two diastereomers present in the X1 solution must be due to the presence of X1 in both the E- and Z-conformation at the C$_{6}$ and C$_{7}$ double bond. The presence of both an E and a Z double bond conformation could a result from chemical modifications due to the presence of formic acid in the extraction process.

As the structure of X1 resembled juvenile hormone III, a hormone thought to be unique for insects with essential developmental roles (Wilson 2004, Gilbert 2000) and the crustacean hormone methyl-farnesoate (Nagaraju 2007) we searched for these metabolites in our extract. The two metabolites were putatively identified through extracted ion chromatograms of the calculated [M+H]$^+$ (juvenile hormone III: 267.1955 and methyl-farnesoate: 251.2006) of the metabolites and putatively identified both metabolites (juvenile hormone III: 267.1957 and methyl-farnesoate: 251.2007). The mass spectra of these two metabolites resembled the spectra of X1 and X2, figure 4.46; however, to confirm the structure of the metabolites an authentic sample of juvenile hormone III was compared to the extract and it confirmed that the metabolite was present. An attempt was made to purify both juvenile hormone III and methyl-farnesoate; however, obtaining sufficient amounts of methyl-farnesoate for NMR analysis was not possible and the juvenile hormone was unstable and had degraded before an NMR analysis was possible. As methyl-farnesoate was thought to be volatile we analyzed the production of volatiles in the mutant strain by GC-MS and were able to confirm the production of methyl-farnesoate through comparison with the Xcalibur software package (Thermo Scientific). Further GC-MS analysis of volatiles collected by the mutant strain showed that methyl-farnesoate was a major metabolite of the volatile fraction in both the smrA expression strain, and in the reference strain whereas juvenile hormone III and X1 were undetectable. Indicating that juvenile hormone III is retained in the fungus. The strongest metabolite induction was observed when *A. nidulans* was grown on CYAs (high salt concentration); however, data from the UHPLC-TOF-MS indicated presence of juvenile hormone III and increased levels of X1 production on YES media (data not shown).

The biological function of juvenile hormone III and methyl-farnesoate in *A. nidulans* have not yet been explored; however, they may serve as part of a defense strategy against insects and/or other predators since fungal secondary metabolites are known to play an important role in
fungal/insects interactions (Kempken 2010, Rohfis 2011). As A. nidulans is generally believed to inhabit a variety of ecological niches, including dead plant materials, it is likely that competing insects will be part of the natural environment. However, it cannot be excluded that the metabolites can serve hormonal functions. It has been demonstrated that the juvenile hormone precursor farnesol (Cao 2009) in Candida albicans both regulates the transition from yeast to filamentous growth (Hornby 2001) and regulates induction of apoptosis in competing fungal species (Dinamarco 2011). To test the biological function of these metabolites it could be interesting to either feed the metabolites to the growing fungi or to test this new strain towards insects.

The biosynthetic origin of these metabolites is likely to be terpenoids with farnesol as the precursor where a methylation and oxidation occurs in one end of the molecule. Afterwards an epoxid, juvenile hormone III is produced, the ring is opened and the diol, X1, is formed. Bromann et al. have identified several putative terpene synthases in A. nidulans which could be involved in the production of these metabolites (Broman 2012). qRT-PCR of the putative terpene synthases and the SmrA strain could give an indication of possible biosynthetic origin which then could be followed up by deletion studies to determine the biosynthetic origin.

4.3.4 Summary and part conclusion

Cultivation of A. nidulans under different media compositions have led to the activation of several biosynthetic pathways, including austinol, monodictyphenone, and orsellinic acid. Through the creation of a deletion library austinol and dehydroaustinol was linked to the PKS ausA and several metabolites were contributed to the monodictyphenone (arugosin A and H) and orsellinic acid (violaceols) pathways.

Overexpression of a transcription factor led to the activation of a biosynthetic pathway where 3-methylorsellinic acid and several derivatives, nidulol, cichorine, cichonidulol, and demethylcichonidulol were produced. Through the qRT-PCR analysis and the creation of a deletion study a cluster of seven genes were identified. Based on metabolite screening a biosynthetic pathway of the metabolites have been proposed which includes a putative NRPS-like encoding gene that is suspected to be involved in the incorporation of a nitrogen atom to the structure.

Further analysis of A. nidulans cultivated on complex media led to the identification of asperugin A and B, and the novel metabolites; nidubenzal A and B. Through analysis of the PKS deletion library several PKSs were identified as putative synthase encoding genes. The combination deletion of the cclA gene and the putative PKS genes led to the linking of asperugin production to AN3230 where work is still in progress for the nidubenzals.

Heterologous expression of a transcription factor from A. niger led to the identification of the juvenile hormone III produced from A. nidulans. Clarification of the distribution of the juvenile hormones in filamentous fungi as well as their biological function in A. nidulans will improve our understanding of fungal/insect interactions. The finding that heterologous expression of transcription factors may influence secondary metabolism is of general relevance for activation of secondary metabolite production.
5 Overall discussion and conclusion

In summary the work performed during the last three years have contributed to the chemical knowledge of several important filamentous fungi including identifying several novel metabolites; oryzamid A1-2, ditryptoleucine, fumonisin B1, nidubenzal A and B, cichonidulol, and demethylcichonidulol as well as linking of several metabolites of A. nidulans to their synthase genes; austinol and dehydroaustinol, asperugin A and B, the arugosins, and finally elaboration on the biosynthetic pathway of cichorine and nidulol. The work have been an example of the advantages of the close collaboration of two fields; molecular biology and natural product chemistry, which have led to discoveries that neither could have achieved separately.

The study on the genetic important filamentous fungus A. nidulans has contributed to and elaborated on the understanding of secondary metabolism activation and biosynthesis. The availability of the genome of A. nidulans has changed the focus of research into the area of secondary metabolites from elucidation of the biosynthetic pathway of a few important metabolites to the search for new metabolites (Walsh 2009). Figure 5.1 shows the distribution of the 32 putative PKSs of A. nidulans on the eight chromosomes. The products of the PKSs marked in green were known at the beginning of this PhD study ((Brown 1996, Watanabe 1998, Watanabe 1999, Bergmann 2007, Chiang 2008, Szewczyk 2008, Bok 2009, Schroeckh 2009, Nielsen 2011), the ones in black have been reported during the course of this PhD (Nielsen 2011, Ahuja 2012, Lo 2012), and the results of this study have contributed to the knowledge of the biosynthetic pathways of the clusters marked by black circles (papers 4-6). The products of the genes in red have not yet been identified.

As the figure shows, progress in mapping of the PKS encoding genes in A. nidulans has intensified in the recent years and many academic groups have added to the knowledge, illustrating the competitive nature of the field. Even though intense efforts have been made, there is still a lot of work to be performed before the complete potential of secondary metabolite biosynthesis of A. nidulans is elucidated. The results, both in this PhD and in the literature, indicate that attempts to affect gene regulation including the OSMAC approach and genetic modifications, have a tendency to activate some of the same gene clusters especially the orsellinic acid, monodictyphenone, and austinol clusters (Bok 2009, Schroeckh 2009, Chiang 2010b, Sanchez 2010, Nahlik 2010, Nielsen 2011, Nützmann 2011, Sanchez 2011, Scherlach 2011, Lo 2012). This may indicate that these metabolites are important to the fungus and a change from normal conditions (both cultivation and genetic) causes the fungus to react by producing other metabolites.
Figure 5.1 Overview of the distribution of the 32 putative PKSs in *A. nidulans*. The products of the genes shown in green were known when this study started, the ones in black have been elucidated in the course of this study where results of this study have contributed to and/or discovered the ones in the black circles. The products of the PKSs in red have not yet been identified. Dark grey circles and ends symbolize centromeric regions, respectively, and should not be considered to scale.

In this thesis the focus have been on combining genetic modifications with the OSMAC approach and through these attempts the discovery of several new metabolites have added to the chemical knowledge of these fungi. One of the more interesting results is the detection of the juvenile hormone III in *A. nidulans* (paper 7) which is known to have a biological function against insects and/or other predators. Further studies of the biological function of these and other secondary metabolites against the fungus itself could add to the knowledge of secondary metabolism and aid in the discovery process through cultivations under more natural conditions than the lab media (for example co-cultivation). Other metabolites which have been discovered in *A. nidulans* through the use of OSMAC in this thesis is the arugosins (Nielsen 2011), asperugin A and B (paper 6) and the novel metabolites nidubenzal A and B (paper 6). These results confirm the potential of challenging the fungus through differing of the cultivation conditions. Several methods of activation of silent genes have been reported in the literature. The individual exchange of the PKS promoters with an inducible promoter by Ahuja et al. have, recently, shown promising results both with respect to PKS and NRPSs (Ahuja 2012, Yeh 2012).

The discovery of new secondary metabolites has been focused on activation of silent genes and thereby, hopefully, new products can be formed. Even if the optimal strategies are found and even more effort is put into secondary metabolite discovery, it may not be possible to discover the products of all secondary metabolite synthase genes and clusters. This may be due to mutations in the genes which can have an effect on the transcription or translation of DNA and
RNA, respectively, in turn providing a non-functional enzyme. The mutations can have an effect on several steps in the assembly for example they may not be able to release the product, catch the substrate, or the folding may be incorrect. This was the case for *A. oryzae* where we showed that 13-desoxypxaxilaine was the final product in the biosynthetic cluster of aflatrem (paper 2) which confirmed the reports of Nicholson *et al.* (Nicholson 2009), who observed a mutation in the *atmQ* gene responsible for the next steps of the biosynthesis of aflatrem in *A. flavus*.

Another area which could be relevant to explore in the search for silent secondary metabolites is the extraction of the metabolites. Methods for secondary metabolite extraction have been developed and implemented as standard procedures in several laboratories. However, some metabolites, for example the fumonisins in *A. niger* (paper 3), went undetected because the extraction procedures are different from most other metabolites.

Even though a biosynthetic pathway is expressed it does not mean that the metabolites are detected. The detection of different metabolites depends of the analytical methods used. Two methods of combined separation and detection, UHLPC-DAD and HPLC-TOF-MS, were used throughout this study. In both methods some metabolites may not be detected. The emericellamides and the fumonisins does not absorb UV-light so they are not detected in a DAD detector whereas metabolites, which do not ionize, will not be detectable through LC-MS. During the course of this study, a new UHPLC-TOF-MS instrument was implemented which enhanced the detection of metabolites and provided valuable information, for example better predictions of the molecular formula due to improved mass accuracy and isotope pattern, in the identification of metabolites.

The discovery of the production of 3-methylorsellinic acid in *A. nidulans* was more prominent on the new instrument and it was not until this new data was available that the metabolite could be confirmed to be present in the reference strain. The availability of a more sensitive instrument also reveals the difficulty of determining whether a metabolite is present or production have been eliminated. This raises the question of the influence of the detection limit and if the metabolites, which cannot be detected, are due to the concentration falling below the detection limit of the instruments, or the loss of production. As in the case of nidubenzal A and B, several PKS candidates were identified even on the most sensitive instrument and it was not until a method of enhancing the metabolite production was found that it was possible to confirm the PKS gene AN3230 to be responsible for PK biosynthesis. Cichonidulol and demethylcichonidulol were isolated from a 200 plate extract in quantities sufficient for NMR analysis; however, only cichonidulol could be detected in the micro-scale extracts of the strain indicating that either the metabolites were produced in too small amount for detection, the cultivation conditions might have been slightly different in one of the experiments, or variations in extraction methods.

The availability of more sensitive and precise instruments also opens up to new methods of data analysis. Multivariate data analyses, which can decipher between co-eluting peaks and identify small variations in large amounts of date are tools which can be useful in future analysis (Toh 2010, Gao 2012, Tian 2012). However, for these programs to be able to give as correct results as possible it is important to be able to perform controlled experiments, especially, if data obtained over longer periods of time are compared. This is both in terms of variations of cultivation (media, temperature fluctuations, time), extraction procedures, and data acquisition. The sensitive TOF-
Overall discussion and conclusion

MS instruments are useful for these methods of data analysis and identification of small differences (for example appearance or disappearances of peaks); nonetheless, to confirm the metabolite structures a substantial amount of material is needed. Bioinformatic analysis can aid in the structure elucidation especially in the case of NRPs due to the modular nature of NRPSs (Koglin 2009, Matthias 2010); however, in the case of PKs it is still difficult to predict structures based on the iPKS (Crawford 2010). It is even more difficult when several tailoring enzymes are present in the secondary metabolite clusters due to the possibility of the biosynthesis of several isomers. Fragmentation patterns of metabolites can assist in the structure elucidation as seen in the example of arugosin A though further structural studies, for example NMR-spectroscopy or X-ray crystallography, may be necessary.

The identification of new secondary metabolites and linking these to genes have, especially since the availability of the genome sequences focused on creating conditions for activation of silent gene clusters in the fully genome sequences strain. Another approach could be to identify metabolites in other high-producer strains of the same species and afterwards use advanced data analysis programs to discover these metabolites in the genome sequenced strain. Since the genome sequenced strain of for example A. nidulans is a strain, which have been used in the lab for decades, the strain may have lost the ability to produce more secondary metabolites than needed due to the lack of natural competition in the surroundings. An example of the use of another strain of A. nidulans in this PhD has been the discovery of nidubenzal A and B, and asperugin A and B which were isolated from a strain from the IBT culture collection and afterwards found to be produced in trace amounts in the lab strain. Another advantage from this approach is the possibility to predict the steps in the biosynthetic pathway and thereby predict which functionalities the tailoring enzymes should hold. Due to the clustering of the synthase and tailoring genes, genome mining could predict possible gene cluster candidates which could then be a target for genetic studies, thereby minimizing the amount of genetic work.

Discovery of the full biosynthetic potential of filamentous fungi can have several purposes. First, as many medical drugs are secondary metabolites or derivatives thereof (Newman 2012) the discovery of new structures enhances the possibilities of discovering new drugs. Second, the possibilities of synthetic biology where swapping of domains or genes can result in the production of designed metabolites. Third, it can help identify possible pathogenic species as the genes can reveal the possibility of mycotoxin production, as the example of fumonisins production of A. niger where a new analogue fumonisin B₆ which have not been reported from Fusarium was identified (paper 3). Fourth, for the filamentous fungi used for industrial purposes, it is easier to identify a new metabolite appearing in a process when the full potential is known. A lot of the genetic work performed on filamentous fungi has been focused on A. nidulans. The strategies and genetic tools developed in this fungus can be transferred into other organisms since the price of genome sequencing is decreasing and within the reach of individual laboratories (Nowrousian 2010).

Due to the vast interest in A. nidulans there are several groups working in this area. During this study several groups have published results of the same biosynthetic pathways, for example orsellinic acid (Bok 2009, Schroeck 2009, Sanchez 2010, Nahlik 2010, Nützmann 2011, Scherlach 2011), austinol (Nielsen 2011, Lo 2012) and monodictyphenone (Bok 2009, Chiang 2010b, Sanchez 2011). Due to the complexity of the secondary metabolism it has been possible for the groups to contribute with new information on already discovered biosynthesis. For example three papers on
the monodictyphenone pathway in A. nidulans were published (Bok 2009, Chiang 2010b, Sanchez 2011) before the report of the arugosin intermediates was linked to mpdG as intermediates in the biosynthesis of the shamixanthones in this study (paper 4). The results of these works along with labeling studies (Ahmed 1992) have led to a detailed biosynthetic pathway proposed by Simpson (Simpson 2012). The examples above, illustrated, that even when the product of a PKS has been identified, the full potential of a biosynthetic pathway may not have been fully elucidated. Ahuja et al. reported linked the production of 3-methylorsellinic acid and cichorine to PkbA; however, overexpression of a transcription factor led the isolation of several metabolites including the two novel metabolites cichonidulol and demethylnidulol (paper 5).

The work in this thesis illustrates the enormous potential of secondary metabolism in filamentous fungi which gives discoveries of new metabolites of previously silent pathways and where new discoveries adds to the pool of knowledge, and at the same time, challenges researchers to go in different directions in the search of novel metabolites.
Overall discussion and conclusion
6 References


Center for Integrated Fungal Research *Aspergillus flavus* genome sequencing project. 2012 Available at: http://aspergillusflavus.org/genomics/ (Accessed on June 17th 2012)


DOE Joint Genome Institute *Aspergillus aculeatus* ATCC16872 Genome portal v.1.1. **2012a** Available at: [http://genome.jgi-psf.org/Aspac1/Aspac1.home.html](http://genome.jgi-psf.org/Aspac1/Aspac1.home.html). (Accessed on June 17th 2012)

DOE Joint Genome Institute *Aspergillus carbonarius* ITEM 5010 Genome portal v.3. **2012b** Available at: [http://genome.jgi-psf.org/Aspca3/Aspca3.home.html](http://genome.jgi-psf.org/Aspca3/Aspca3.home.html). (Accessed on June 17th 2012)


References


References


Raper, K.B.; Fennell, D.I. *The genus Aspergillus*; Williams & Wilkins: Baltimore, MD, USA, **1965**.


6 | References


Wilson, T.G. The molecular site of action of juvenile hormone and juvenile hormone insecticides during metamorphosis: how these compounds kill insects. *J. Insect. Physiol.* 2004, 50, 111-121.


6 References


Paper 1

“Genetics of polyketide metabolism in Aspergillus nidulans”

ML Klejnstrup, RJN Frandsen, DK Holm, MT Nielsen, UH Mortensen, TO Larsen, and JB Nielsen

Accepted in Metabolites, 2, 100-133 (2012)
Review

Genetics of Polyketide Metabolism in *Aspergillus nidulans*

Marie L. Klejnstrup, Rasmus J. N. Frandsen, Dorte K. Holm, Morten T. Nielsen, Uffe H. Mortensen, Thomas O. Larsen and Jakob B. Nielsen

1 Department of Systems Biology, Center for Microbial Biotechnology, Technical University of Denmark, Søltofts Plads B221, DK-2800 Kgs. Lyngby, Denmark; E-Mails: marlk@bio.dtu.dk (M.L.K.); tol@bio.dtu.dk (T.O.L)

2 Department of Systems Biology, Center for Microbial Biotechnology, Technical University of Denmark, Søltofts Plads B223, DK-2800 Kgs. Lyngby, Denmark; E-Mails: rasf@bio.dtu.dk (R.J.N.F.); dmkp@bio.dtu.dk (D.K.H.); motni@bio.dtu.dk (M.T.N.); um@bio.dtu.dk (U.H.M.)

* Author to whom correspondence should be addressed; E-Mail: jbn@bio.dtu.dk; Tel.: +45-4525-2657; Fax: +45-4588-4148.

Received: 1 November 2011; in revised form: 23 December 2011 / Accepted: 17 January 2012 / Published: 30 January 2012

Abstract: Secondary metabolites are small molecules that show large structural diversity and a broad range of bioactivities. Some metabolites are attractive as drugs or pigments while others act as harmful mycotoxins. Filamentous fungi have the capacity to produce a wide array of secondary metabolites including polyketides. The majority of genes required for production of these metabolites are mostly organized in gene clusters, which often are silent or barely expressed under laboratory conditions, making discovery and analysis difficult. Fortunately, the genome sequences of several filamentous fungi are publicly available, greatly facilitating the establishment of links between genes and metabolites. This review covers the attempts being made to trigger the activation of polyketide metabolism in the fungal model organism *Aspergillus nidulans*. Moreover, it will provide an overview of the pathways where ten polyketide synthase genes have been coupled to polyketide products. Therefore, the proposed biosynthesis of the following metabolites will be presented; naphthopyrone, sterigmatocystin, aspyridones, emericellamides, asperthecin, asperfuraneone, monodictyphenone/ emodin, orsellinic acid, and the austinols.

**Keywords:** secondary metabolites; polyketides; polyketide synthases; gene clusters; biosynthesis; *Aspergillus nidulans*
1. Introduction

Aspergillus nidulans, teleomorph Emericella nidulans, is one of the most significant biological model systems in the fungal kingdom. This was pioneered by Pontecorvos’ [1] work in the middle of the last century, which demonstrated that A. nidulans, in addition to the asexual state, also proliferate via sexual and parasexual life cycles, hence, offering an ideal platform for genetic studies. Related species in the genus Aspergillus include important industrial cell factories, A. niger and A. oryzae, species that cause allergic diseases, A. clavatus, as well as opportunistic pathogens, such as A. fumigatus.

A common feature of aspergilli and filamentous fungi in general is their capacity to produce secondary metabolites (SMs). As opposed to the primary metabolites, SMs are not essential for cellular growth, but provide fungi, as well as bacteria and plants, with a competitive advantage in nature, e.g., by serving as agents for chemical warfare or as signal molecules. Hence, an impressive range of compounds with broad ranging bioactivities has evolved. SMs can be divided into four main chemical classes: Polyketides (PK), terpenoids, shikimic acid derived compounds, and non-ribosomal peptides (NRP). Moreover, hybrid metabolites composed of moieties from different classes are common, as in the meroterpenoids, which are fusions between PKs and terpenes. Hybrid molecules significantly add to the complexity and variety of the fungal metabolomes.

In addition to their likely important ecological roles in their natural biological niches, SMs also have a considerable impact on human life. For instance aflatoxins, ochratoxins, and fumonisins act as mycotoxins by having a detrimental effect on humans and livestock, whereas others are beneficial and serve as food additives, pigments, cholesterol-lowering drugs, immunosuppressants, antibiotics and anticancer agents. The different aspects of SM action and application have spurred a tremendous interest in fungal secondary metabolites, which is further underlined by the fact that around 63% of all small molecule drugs, which reached the market from 1981–2006 were inspired by natural products or derivatives thereof [2].

In filamentous fungi, the competitive race in SM development and the cost of producing and secreting complex compounds have resulted in the evolution of a multifaceted regulation of SM biosynthesis to avoid unnecessary use of resources. This hampers their discovery since production of most SMs is not induced under laboratory conditions. Analysis of full genome sequences of eight different aspergilli have demonstrated that for the majority of genes that putatively encode enzymes for SM production, the product is not known or detected. In this review, we will provide highlights of the use of genome mining, sophisticated molecular biological and chemical tools to trigger the production of SMs from cryptic gene clusters and discuss how these techniques have accelerated our understanding of PK production and regulation in A. nidulans.

1.1. Polyketide Biosynthesis in A. nidulans

PKs in fungi are synthesized by the use of acyl-CoA units. They act as the general substrates for large multi-domain enzymes named polyketide synthases (PKSs), which resemble eukaryotic fatty-acid synthases (FASs) in domain architecture. PKSs are divided into three types of PKSs based on their catalytic organization, however, only the iterative type I PKS (iPKS) has been reported in A. nidulans. The iPKS repeat the use of a single module containing several catalytic domains until the growing
chain of acyl-CoA units block further elongation. For descriptions of PKSs in general, excellent reviews by Crawford [3], Hertweck [4] and Cox [5] can be consulted. The most commonly encountered catalytic activities in fungal PKSs will be addressed as a general introduction to fungal PKSs in the following three paragraphs.

Three fundamental domains are found in all iPKSs in *A. nidulans* like in filamentous fungi in general; β-ketosynthase (KS), acyltransferase (AT), and the acyl carrier protein (ACP). The KS catalyzes the C–C bond formation via decarboxylation reactions through Claisen condensations between thioesters. The ACP domain is responsible for transiently holding the growing acyl chain, hereby allowing the loading of malonyl extender units. The acyl groups are transferred from CoA by AT onto KS and ACP. The iterative use of the three domains results in a non-reduced PK, a β-keto thioester. Additional domains can be present in the PKS allowing the introduction of further chemical complexity.

iPKSs in fungi can, based on their catalytic domains, be classified as non-reducing (NR-PKSs), partially reducing (PR-PKSs), or highly reducing (HR-PKSs) [6]. This is based on their ability to reduce the β-keto carbon. In PR- and HR-PKSs, reduction occurs through the β-ketoreductase (KR) domain that converts the β-ketone to a hydroxyl group. The resulting hydroxyl can go all the way to saturation by elimination of water through the dehydratase (DH) domain followed by hydrogenation by enoyl reductase (ER). In addition, reducing PKSs can also possess a methyltransferase domain (MT) responsible for C-methylation of the growing PK chain, using S-adenosylmethionine (SAM) as a carbon-donor. The degree of modifications and their position in the PK product is always the same for the individual PKSs. However, it is presently unknown how deployment of the various modifying domains is programmed into the PKS enzyme.

NR-PKSs differ in domain architecture from reducing PKS by not having any of the reducing domains and by having an N-terminal starter unit-ACP transacylase (SAT) domain and an internal product template (PT) domain. The SAT domain is responsible for selecting the starter unit to be extended by the enzyme [7], while the PT domain is responsible for folding and cyclization of the non-reduced PK backbone [3,8]. The number of iterations within the PKS and thereby the display of functional groups and the size of the final product is likely determined by the size of the active site cavity in the iPKS [9]. Once the length of the final product has been achieved, the PK chain is released from the PKS, catalyzed by either a thioesterase (TE), a Claisen cyclase (CLC) domain if present, or by accessory enzymes. A more detailed discussion on PKS release mechanisms is reviewed by Du and Lou [10].

It should be noted that it currently is impossible to reliably predict the product of iPKSs based on their amino acid sequences and domain architecture. This is in part due to the inability to predict the number of iterations performed by the iPKS and in part due the lack of understanding of how deployment of tailoring domains in individual iterations are programmed into the enzyme.

Interestingly, the PKS encoding gene tends to reside in clusters of genes coding for a broad range of enzymatic activities. The compound coming directly from the PKS rarely seem to be the final product in the biosynthesis, but usually undergoes further modifications by tailoring enzymes from small decorations to drastic and large intervention and couplings.

Through inspection of the genome sequence (genome mining), the latest estimate of genes encoding PKSs in *A. nidulans* is 32 open reading frames (ORFs) [11] (Figure 1), indicating that the number of PK containing end products in *A. nidulans* should count at least 32 plus stable intermediates. The compounds detected under a given condition do not necessarily reflect the final outcome of a PK
pathway, since the presence of intermediates and shunt products depends on other downstream enzymes and regulation.

**Figure 1.** An overview of the relative distribution of the 32 putative polyketide synthases (PKS) open reading frames (ORFs) on the eight chromosomes of *A. nidulans*. Green and red AN numbers represent assigned and unassigned PKS genes, respectively. Dark grey circles and ends symbolize centromeric and telomeric regions, respectively, and should not be considered to scale.

At present, a total of nine PKS genes have been coupled to the polyketome (collection of PKs and their synthesis) in *A. nidulans*, and numerous endeavors are currently attempting to unveil the mechanisms of PK metabolism in this model fungus. The pathways described in this review will follow in chronological order with respect to the discovery from PK to genes. For each of the PK gene clusters that have been linked to products so far we will focus on the PK compounds, their discovery, genetics as well as their biosynthetic pathway: Naphthopyrone, sterigmatocystin, aspyridones, emericellamides, asperthecin, asperfuranone, monodictyphenone (emodin), orsellinic acid, and the austinols.

2. Naphthopyrone

Spores from *A. nidulans* are characterized by a dark grey-green macroscopic appearance. This is due to deposition of pigments in the conidial wall as shown by ultrastructure studies using transmission electron microscopy (TEM) [12]. The responsible pigment is based on the PK-naphthopyrone YWA1 and the function of the pigment layer has been shown to include quenching of reactive oxygen species [13] and increased resistance to UV radiation [14]. The work on naphthopyrone synthesis in *A. nidulans* has paved the way for understanding iPKS domain structure.

The study of conidial pigmentation in *A. nidulans* has been extremely valuable for genetic screens. The first pigment mutant recorded in literature was the spontaneous white alba (*nua*) strain reported by...
Yuill in 1939 [15]. In addition to the white mutant class, \( yA^+ \) mutants producing yellow conidia were discovered. These available color variants served as easy recognizable markers (green, white, and yellow) that allowed the establishment of fundamental genetic tools in \( A. nidulans \) [1]. Interestingly, sexual crossing showed that the \( wA \) mutation masked the effect of the \( yA \) mutation (epistatic) [1,16]. Clutterbuck and co-workers [16] proposed that \( WA \) synthesized the yellow pigment that was observed in \( yA \) mutants and that the \( YA \) enzyme converted this compound into the green conidial pigment. In 1967, Agnihotri and co-workers [17] found that the wild type strains if grown under copper limiting conditions could mimic the yellow phenotype of the \( yA^+ \) strain. \( yA \) (AN6635) and \( wA \) (AN8209) were isolated and mapped to loci, chromosome I and II respectively, by complementation of a cosmid based library in 1989 and 1990, respectively [18,19]. Later, cross-feeding experiments performed by Clutterbuck [16] revealed that the \( yA \) phenotype was caused by the lack of a copper dependent extracellular laccase (p-diphenol oxidase). The \( wA \) functionality in pigment formation was confirmed by gene-deletion studies [19]. The lack of clustering of the two \( A. nidulans \) conidial pigment genes also became evident by their different expression patterns and the finding that they are controlled by different regulatory systems [20,21]. The \( yA \) gene is expressed in phialide cells and primary sterigmata (metulae) [18], and controlled by BrIA and AbaA [21], while \( wA \) is expressed only in phialides [22] and controlled by WetA [20]. Interestingly, none of the genes are expressed in the conidia. Characterization of the WA PKS was accomplished by Northern blotting, which revealed that \( wA \) encoded a 7.5 kb large transcript [19], and sequencing of the locus [22]. Re-sequencing of the 3' region in 1998 led to a revised gene model of the PKS with the following domain structure KS-AT-ACP-CLC. This novel CLC domain [23] catalyzed release of the product and cyclization of the second aromatic ring of YWA1 via a Claisen condensation reaction [24].

Heterologous expression of \( wA \) in \( A. oryzae \) resulted in the production of the yellow compound, as observed in \( yA^+ \) mutants, which was identified to be the heptaketide naphthopyrone named YWA1 [25]. In 2002, \( wA \) was used for constructing a collection of chimeric PKSs (cPKS) by mixing its domains with those of \( Colletotrichum lagenarium pks1 \), known to produce the tetraketide 1,3,6,8-tetrahydroxynaphthalene (T4HN). One of the resulting cPKSs, SW-B, produced several new compounds including both tetra- and pentaketides [26]. The results prompted a reanalysis of the two PKSs, which revealed the existence of two previously overlooked conserved domains; an N-terminal and a central domain. These domains were later identified as a SAT and PT domain, respectively, thus providing the full domain structure SAT-KS-AT-PT-ACP-CLC [3,7,8]. With the organization within WA in mind, the biosynthetic pathway can be envisioned as condensations of an acetyl-CoA with six malonyl-CoA units in six successive reactions resulting in the formation of YWA1 [25] (Figure 2).
Figure 2. Biosynthetic pathway for formation green conidial pigment in *A. nidulans*

YWA1 is then believed to be dimerized or polymerized by the YA laccase into the green conidial pigment(s) via phenolic oxidative coupling. However, to date no one has succeeded in characterizing the chemical structure of the green conidial pigmentation in detail. As the final product remains elusive, it is impossible to predict if other tailoring enzymes further modify the YWA1 backbone or reactions occur with other metabolites or cellular components, e.g., the cell wall.

3. Sterigmatocystin

Sterigmatocystin, a PK, was first partially purified from a *Sterigmatocytis* sp. in 1948 by Nekam and Polgar [27]. Hatsuda and co-workers [28,29] successfully isolated sterigmatocystin in 1954 from *A. versicolor*. The correct relative structure was determined in 1962 by Bullock *et al.* [30]. By performing degradative experiments it was shown that the stereochemistry of sterigmatocystin was the same as that of aflatoxin [31], which had the absolute stereochemistry determined in 1967 [32]. The absolute stereochemistry of sterigmatocystin was confirmed via crystallography [33,34].

The aflatoxins are among the most carcinogenic mycotoxins and the research in aflatoxin and sterigmatocystin intensified with the Turkey X disease caused by aflatoxins in the middle of the last century [35]. Aflatoxins are reported to be produced only by a few aspergilli. *A. nidulans* does not produce aflatoxins, as the biosynthesis stops at sterigmatocystin, a late, yet stable precursor of the pathway. Sterigmatocystin is a powerful mycotoxin, though it is estimated to be 150 times less carcinogenic than the most potent aflatoxin, B1 [36]. Fungi that are able to produce aflatoxins and/or sterigmatocystin are common contaminants of food, feed, and indoor environments and may be mammalian and plant pathogens [37,38]. Due to the high toxicity and prevalence of sterigmatocystin and aflatoxins, they are likely the most extensively studied examples of secondary metabolism in fungi both in terms of biosynthesis and biological function, and there are several excellent and comprehensive reviews for further reading on aflatoxin biosynthesis [39,40]. Studies on the biosynthesis of aflatoxin and sterigmatocystin have been carried out in several fungi (*A. flavus, A. nidulans* and *A. parasiticus*) and some of the assigned gene functions in *A. nidulans* are proposed based on gene homology to these two other species.

The biosynthetic cluster of sterigmatocystin in *A. nidulans* was first characterized by Brown and co-workers in 1996 [41]. They identified a 60 kb region in the *A. nidulans* genome responsible for the synthesis of sterigmatocystin. The cluster contains 27 genes named *stcA*-X (Figure 3), reflecting their order of appearance on the chromosome [41].
The PKS catalyzing the production of the PK backbone of sterigmatocystin was identified by Yu and co-workers in 1995 [42] and originally named pksST, but was later renamed to stcA by Brown et al. [41] to simplify nomenclature. Besides the PKS, the stc gene cluster is predicted to contain two transcription factors (aflR, aflJ), six monooxygenases (stcB, stcC, stcF, stcL, stcM, stcS, stcW), two dehydrogenases (stcG, stcF), an esterase (stcI), an O-methyltransferase (stcP), two ketoreductases (stcE, stcU), a VERB synthase (stcN), an oxidase (stcO), a monoxygenase/oxidase (stcQ), a Baeyer-Villiger oxidase (stcR), a fatty acid synthase composed by the two subunits HexA and HexB (encoded by stcJ and stcK, respectively), and five unassigned genes (stcD, stcH, stcR, stcT, stcX), which may also be part of the cluster [39–41,43–47].

The stc cluster is a relatively large gene cluster, and studying the gene regulation has led to important discoveries. The two TFs were found to be present within the cluster. The AfIR is a Zn$_2$Cys$_6$ TF that regulates transcription of the stc locus in A. nidulans [41,48], while AfJ (also named AflS) have been shown to have a role in the regulation of aflatoxin biosynthesis in A. flavus and is likely to have a similar function in A. nidulans [49]. Interestingly, Bok and Keller [50] discovered a novel regulator (LaeA) of secondary metabolism in A. nidulans in a mutant screen for loss of aflR expression. Deletion of laeA resulted in a significantly decreased production of different classes of SMs like sterigmatocystin and penicillin. LaeA, a putative methyl transferase was moreover acting in a feedback loop with AfIR since overexpression of aflR downregulates laeA expression, and overexpression of laeA could not increase production of sterigmatocystin [50]. LaeA was shown to be a part of the conserved Velvet complex, which is important for regulation of fungal development and secondary metabolism [51]. Another hint on chromatin regulated gene expression came from the deletion of a histone deacetylase, hdaAΔ, which led to significant increase in the expression of two stc cluster genes, stcU and aflR compared to the reference [52].

Applying this strategy of deleting and overexpressing genes encoding global epigenetic regulators has paved the way for novel discoveries in secondary metabolism. Moreover the alternative of utilizing a chemical epigenetic approach through epigenetic modifier molecules has proven successful in activating gene clusters in A. niger [53].

The first step in the biosynthesis of sterigmatocystin (Figure 4) is the production of hexanoate by the FAS units, StcJ and StcK [41]. Watanabe and Townsend [54] showed that the hexanoyl-CoA is not an intermediate freed from the complex, indicating that hexanoate is transferred directly to the SAT domain of the PKS. The PK backbone is assembled by StcA by condensation of the starter unit, hexanoyl-CoA and seven malonyl-CoA extender units followed by cyclization and release of norsolorinic acid anthrone [42]. The oxidation of norsolorinic acid anthrone to norsolorinic acid may be catalyzed by stcM, a monooxygenase ortholog to hypC that converts norsolorinic acid anthrone to norsolorinic acid in A. parasiticus [43]. Norsolorinic acid is the first stable intermediate in the biosynthesis of sterigmatocystin and is converted into averantin by StcE, reducing the hexanoate ketone to an alcohol [41,55,56].

**Figure 3.** The sterigmatocystin gene cluster. The black arrows are predicted stc ORFs, while light grey arrows are genes with unassigned functions.
Figure 4. Proposed biosynthesis of sterigmatocystin. StcA contains starter unit-ACP transacylase (SAT), β-ketosynthase (KS), acyltransferase (AT), product template (PT), acyl carrier protein (ACP) and thioesterase/claisen cyclase (TE/CLC) domains. *Indicates a proposed, but not confirmed, enzyme. Multiple arrows indicate that the number of enzymatic steps is unknown.
Yabe et al. [55] showed that 5'-hydroxyaverantin is a step towards aflatoxin in *A. parasiticus* and a later study showed that the oxidation of averantin into 5'-hydroxyaverantin is catalyzed by StcF [57]. The conversion of 5'-hydroxyaverantin to 5'-oxyaverantin is likely catalyzed by StcG [39,44]. In a study in *A. parasiticus* by Yabe and co-workers it was shown that both (1'S, 5'S)- and (1'S, 5'R)-hydroxyaveratin are formed in the conversion of averatin to 5'-oxyaveratin [58]. Disruption of *aflH* in *A. parasiticus* resulted in the accumulation of 5'-hydroxyaverantin, however, small amounts of *O*-methylsterigmatocystin present suggested that other enzymes may be involved in the reaction [44,59]. The gene(s) responsible for the conversion of 5'-oxyaverantin to averufin have not been identified [39,60].

Individual disruption of StcB and StcW resulted in elimination of sterigmatocystin and accumulation of averufin, indicating that both enzymes catalyze the conversion of averufin to 1-hydroxyversicolorone [57]. It was not possible for the authors to determine why two monooxygenases were required for this reaction step [57]. No gene products have been identified as being responsible for the conversion of 1-hydroxyversicolorone to versicolinal hemiacetal acetate. StcI is thought to catalyze the reaction from 1-hydroxyversicolorone to versicolorin based on studies of the ortholog AflJ in *A. parasiticus*, though other genes capable of this reaction may be present in *A. nidulans* [44,61].

Deletion of *stcN* did not result in the production of sterigmatocystin or other intermediates [44]. However, StcN show homology to AflK and the versicolorin B synthase, Vbs, in *A. parasiticus*, indicating that the biosynthetic step from versicolorin B to versicolorin A may be catalyzed by StcN [39,44,62]. StcL was shown by Kelkar et al. [63] to catalyze the conversion of versicolorin B to versicolorin A. Inactivation of *stcL* resulted in accumulation of dihydrosterigmatocystin, leading to a branching of the sterigmatocystin biosynthesis as seen in the aflatoxin biosynthesis. Addition of versicolorin A to the mutant gave production of sterigmatocystin, and that indicated that this enzyme functions before versicolorin A [63].

Keller and co-workers [64,65] showed that StcU and StcS are involved in the conversion of versicolorin A to demethylsterigmatocystin. Individual disruption of *stcU* and *stcS* led to the accumulation of versicolorin A and eliminated production of sterigmatocystin in *A. nidulans* [64,65]. Henry and Townsend [45] studied the same step in the aflatoxin biosynthesis in *A. parasiticus* and proposed an oxidation-reduction-oxidation mechanism, involving at least a ketoreductase AflM and a monooxygenase AflN, orthologs to *stcU* and *stcS*, respectively. Ehrlich et al. [46] and Cary et al. [47] identified two enzymes, AflY, a Baeyer-Villiger oxidase and AflX, an oxidoreductase, to be involved in the conversion of versicolorin A to demethylsterigmatocystin in *A. parasiticus* and *A. flavus*, respectively. *aflX* and *aflY* are homologous to *stcQ* and *stcR*, which suggests that these genes might be involved in the biosynthetic step from versicolorin A to demethylsterigmatocystin [46,47].

The final step in the biosynthesis of sterigmatocystin is the methylation of demethylsterigmatocystin catalyzed by StcP [66]. The conversion of sterigmatocystin to aflatoxin involves two additional biosynthetic steps; an *O*-methylation of sterigmatocystin by *aflP* followed by involvement of *aflQ* to produce aflatoxin [39]. Slot and Rokas have recently showed that the sterigmatocystin gene cluster in *Podospora anserina* was horizontally transferred from *Aspergillus*, which shows that transfer of large metabolite clusters between fungi are possible [67].
4. Aspyridone

Aspyridone is a PK-NRP hybrid and a fascinating example on how PK-NRP compounds in *A. nidulans* can be assembled from the activity of a single fusion enzyme. The aspyridones have shown to display moderate cytotoxicity [68]. The responsible gene cluster was discovered by Bergmann and co-workers [68] using a genome mining approach. Using the *Aspergillus* genome sequence, they identified a SM gene cluster, which contained a putative TF (AN8414/apdR) that the authors hypothesized could trigger activation of the genes in the cluster. Accordingly, the authors overexpressed the TF under the control of an inducible *alcA* promoter by integrating it randomly in the genome. In agreement with the hypothesis, it was demonstrated by Northern blot analysis that six of the nearest neighbor genes were upregulated in this strain on inductive medium, and that the aspyridones and two intermediates or shunt products could also be detected. Prediction of the catalytic potential for the six upregulated genes (*apdA, apdB, apdC, apdD, apdE*, and *apdG*) combined with the structure of the accumulating compounds allowed the authors to propose a model for the biosynthesis of aspyridones including an assignment of the involved enzymes (Figure 5).

**Figure 5.** Proposed biosynthesis of aspyridone A and B. (a) Based on proposed biosynthesis by [68,72,73]; (b) Based on results by Halo et al. [78] Multiple arrows indicate that the number of enzymatic steps is unknown.
Metabolites 2012, 2 110

apdA was deleted by Chiang et al. [69] and confirmed to be involved in aspyridone biosynthesis as reported by Bergmann et al. [68].

The structure of aspyridone A and B suggested that their synthesis involved both PKS and NRPS activity. Indeed, analysis of the AN8412 structure revealed domains characteristic for a HR-PKS as well as NRPS in one ORF spanning more than 11 kb. This is a special subclass of reducing PKSs, where the PKS has been directly fused with a single NRPS module at the C'-terminal end. This architecture allows for the incorporation of amino acids or carboxylic acids into the carboxylic end of the growing PK chain. Only one of these fusion enzymes has been found in A. nidulans, but has been reported in other fungi [70,71]. Since AN8412 is the first enzyme to act in the pathway, the gene was named apdA. ApdA catalyzes the assembly of the PK-amino acid backbone of the aspyridones by three Claisen condensations of malonyl-CoA, and KR-DH-ER-MT carries out full reduction of the β-keto and the methylations, which are required. However, as ApdA lacks a functional ER domain, the ER activity is most likely provided by ApdC, a homolog to an enoyl reductase (LovC) from the lovastatin biosynthetic gene cluster [68]. The resulting triketide is transferred to the NRPS module, where it is linked to tyrosine [68]. Bergmann and co-workers listed the domains through protein homology in the NRPS as condensation (C), adenylation (A), peptidyl carrier protein (PCP) and reductase domain (RED). The release of the PKS-NRPS hybrid product was proposed to be a NADPH-dependent reductive release followed by an intramolecular Knoevenagel condensation and enzymatic oxidation [68].

Biochemical studies of the role of ApdA and ApdC in the biosynthetic pathway of the aspyridones have been performed by Liu et al. [72] and Xu et al. [73]. Liu and co-workers [72] defined the NRPS module as C-A-T-R with the latter two being thiolation and reductase, which is an alternative to the more frequent C-A-T-TE found in these modules. However, this reductase domain (R*) in the NRPS module of ApdA is not the standard SDR superfamily dehydrogenase since tyrosine in the Ser-Tyr-Lys catalytic triad is mutated suggesting a redox-independent condensation reaction and the release of a tautomer of preaspyridone from ApdA by a Dieckmann cyclization, which was first shown by Halo and co-workers [74]. This result has been confirmed by Xu et al. who expressed the apdA and apdC genes in Saccharomyces cerevisiae and Escherichia coli, respectively [73]. The purified enzymes (ApdA and ApdC) were incubated in the presence of cofactors and building blocks and the predominant product was preaspyridone [73].

The formation of preaspyridone into aspyridone A and B was proposed by Bergmann et al. [68], (and outlined in Figure 5a) using the predicted functions of the remaining genes of the apd gene cluster. The proposed biosynthesis involved ApdB and ApdE which shows similarity to cytochrome P450 oxygenases and cytochrome P450 alkane hydroxylases, respectively, and were believed to catalyze the formation of hydroxypreaspyridone [68]. Based on the study of pyridone rearrangement in metabolites related to aspyridone it was suggested that ApdE or ApdB were involved in the pyridone rearrangement [68,75–77]. Moreover, aspyridone A was hypothesized to be converted into aspyridone B by ApdD, a putative FAD-dependent monooxygenase, which is related to other ring hydroxylases [68]. Aspyridone has a similar structure to other pyridines isolated from fungi, e.g., tenellin whose biosynthetic gene cluster also has been identified [75,76]. The proposed biosynthesis of aspyridone was, as described above, based on predicted gene functions and not isolated intermediates. However, a study on the biosynthesis of the related metabolite tenellin by Halo and co-workers [78] showed that the suggested biosynthesis may be incorrect and an alternative biosynthesis was suggested (as shown
in Figure 5b). In this biosynthesis preaspyridone is not converted into hydroxyaspyridone but ring expanded by ApdE to aspyridone A similar to the biosynthesis of tenellin [78]. Halo et al. also showed that the hydroxylated metabolite of pretenellin is a shunt metabolite as it could not be converted into tenellin.

5. Emericellamides

The emericellamides are other examples of hybrid compounds that are formed between PKs and NRPs. In this case the biosynthesis requires a PKS and a NRPS rather than a fusion PKS-NRPS as used in the production of the aspyridones. Emericellamides are cyclic depsipeptides and a total of five variants, A, C–F, of these metabolites have been found in A. nidulans. Initially, emericellamide A and B were isolated and described from an unidentified marine-derived Emericella strain in a screen due to their antibacterial activity against methicillin-resistant Staphylococcus aureus [79].

In order to discover novel natural products, Chiang and co-workers [69] searched the genome sequence of A. nidulans for NRPS gene candidates. Subsequently, six of these genes were randomly chosen and deleted by gene targeting. One of the resulting mutants, AN2545Δ, showed a metabolite profile where emericellamide A was missing. Furthermore, HPLC profiles and dereplication using mass spectrometry and database searches revealed four additional compounds, which disappeared in the mutant metabolite profile. Since these compounds had not previously been described in A. nidulans, they were purified and their structures solved by NMR analysis revealing that they were novel analogues of emericellamide A and B, thus named emericellamide C–F [69].

To investigate whether AN2545, now called easA, defines a gene cluster encoding all necessary enzymatic activities in the emericellamide biosynthetic pathway, the genes from AN2542 to AN10325, a total of ten, were deleted [69]. Most of these gene deletions did not affect emericellamide production as judged by LC-MS analysis, demonstrating that they do not participate in the biosynthesis. However, the emericellamides were absent in four of the deletion strains, now named easA-easD, indicating that these genes are involved in the pathway.

Bioinformatic analysis of the three additional genes suggested that they all encode activities that are relevant for emericellamide biosynthesis. Specifically, easB (AN2547), a PKS, easC (AN2548), an acyl transferase, and easD (AN2549), an acyl-CoA ligase. Based on these putative activities, the authors proposed a biosynthetic pathway for emericellamide production (Figure 6). In this model, the biosynthesis is initiated by EasB, a HR-PKS composed of the domains KS-AT-MT-DH-ER-KR-ACP. Since the PK component of the different emericellamide variants differ with respect to chain length and methylation pattern, it indicates that iterativity of this PKS is flexible [69].

Next, the PK carboxylic acid is converted to a CoA thioester by the acyl-CoA ligase, EasD, loaded onto the acyltransferase EasC, and then transferred to the thiolation (T) domain of EasA. This NRPS is a multi-modular enzymatic assembly containing 18 domains grouped into five modules. Among those, the authors propose that the first T domain is responsible for accepting the incoming PK from EasC (Figure 6). Moreover, the remaining domains fit well with the fact that five amino acids are incorporated into emericellamides. The authors note that this NRPS does not contain a TE domain at the end of module 5, indicating that this enzymatic activity is not necessary for cyclization of the emericellamides [69].
**Figure 6.** Proposed biosynthesis of the emericallamides. The order of the methyltransferase (MT) and dehydratase (DH) domain as suggested by Chiang et al. [69], however a BLASTp analysis suggests a swapping of the MT and DH domains. The NRPS, EasA, contains 18 (T, E (epimerization), C, A) domains grouped into five modules.
6. Asperthecin

Asperthecin is a PK compound that was first isolated from *A. quadrilineatus* by Howard and Raistrick in 1955 [80], however, the structure was not fully determined until six years later [81]. Initially, various chemical derivatizations and spectroscopic data determined a partial structure of asperthecin [80]. Neelakantan and co-workers [82] reduced the number of possible structures to two, and further derivatizations of asperthecin by Birkinshaw and Gourlay resulted in the final structure [81]. *A. quadrilineatus* is a member of the *A. nidulans* group, therefore Howard and Raistrick [80] extended the search of asperthecin to additional members of the *A. nidulans* group. No other was as rich in asperthecin production as *A. quadrilineatus*, yet small amounts of crystallized asperthecin could be obtained from cultures of *A. nidulans* and *A. rugulosus* indicating that production of asperthecin was possible in other aspergilli [80].

About 50 years later, Szewczyk and co-workers [83] used a molecular genetics approach to find the gene cluster responsible for the production of asperthecin in *A. nidulans*. Since many aspects of the regulation in the polyketome were unknown, the authors speculated whether sumoylation had an effect. SUMO is a small ubiquitin-like protein which is post-translationally added to proteins in the cell, as it plays a role in regulating transcription. *A. nidulans* contains one SUMO encoding gene, *sumO* [84], deletion of which led to a decrease in the production of SMs such as austinol, dehydroaustinol, and sterigmatocystin, and an increase in the production of a metabolite identified to be asperthecin, whereas the production of emericellamides were not affected [83]. Due to the aromatic structure of asperthecin, Szewczyk *et al.* [83] studied the domain prediction in 27 putative PKS-protein sequences using the *A. nidulans* genome sequence and available tools, in order to identify potential producers of non-reduced PKs. Ten NR-PKSs were identified and a deletion series of all NR-PKS genes was performed in the *sumO* background [83]. While nine of the PKS-deletion strains still produced asperthecin, the AN6000 (*aptA*) PKS-deletion strain failed to synthesize asperthecin [83]. With the notion that most end compounds in PK biosynthesis are made by a clustered gene collective, six candidate genes surrounding *aptA* were picked in an attempt to identify the *apt* biosynthetic cluster. Two of these genes, *aptB* (AN6001) and *aptC* (AN6002), were found to be required for asperthecin production [83]. One strain (AN5999Δ) had a significantly lower production of asperthecin compared to the reference strain, but asperthecin was still present in the metabolite extracts, and as the strain showed poor growth, it was not included in the *apt* gene cluster.

Interestingly, *AptA* was shown to have SAT-KS-AT-PT-ACP domains, but lack a TE/CLC domain [83,85–88]. Independent groups have used this case as a model system to study the mode of PK release, and two alternating mechanisms for the biosynthesis of asperthecin are shown in Figure 7 [86,87]. The first model suggests the formation of the PK backbone by condensations of one acetyl-CoA and seven malonyl-CoA units [86], with the β-lactamase *AptB* releasing the octaketide from *AptA* [83]. This assumption was based on a study by Awakawa and co-workers [85] in *A. terreus*, where there was a release of atrochrysone carboxylic acid from the atrochrysone carboxylic acid synthase (ACAS) lacking a TE/CLC domain, in the presence of the atrochrysone carboxylic ACP thioesterase (ACTE), a member of the β-lactamase superfamily [85]. The unstable atrochrysone carboxylic acid then undergoes a series of reactions; decarboxylation, dehydration, and various oxidations where the monoxygenase *AptC* is believed to be involved, and in the end yielding asperthecin [86].
In another approach, Li et al. [87] introduced aptA, aptB, and aptC into S. cerevisiae, which resulted in the production of a nonaketide (here called preasperthecin), and not the octaketide as proposed by Chiang et al. [86]. Expressing aptA and aptB without aptC resulted in a product identical to preasperthecin except for the lack of C2-OH, confirming that AptC is responsible for this step [87]. Expression of aptA and aptC alone did not lead to the production of preasperthecin or any other traceable compounds, confirming that AptB is needed for release of the PK from AptA [87]. These results were confirmed by an in vitro assay after expressing aptA, aptB and aptC in E. coli [87]. Further insight into AptA functionality came from expressing the AptA-PT domain in E. coli, and combining it with the Gibberella fujikuroi PKS4, which can produce nonaketide products in vitro [88]. The experiment revealed that AptA-PT can catalyze C6-C11 cyclization, and most likely also the C4-C13 cyclization. Further, a spontaneous C2-C15 cyclization was followed by a C1-C17
Metabolites 2012, 2 115

esterification [88]. The apt gene cluster appears likely to consist of additional genes, which are responsible for the conversion of preasperthecin into asperthecin.

7. Asperfuranone

Asperfuranone is an example of a novel PK metabolite discovered through genetic mining in A. nidulans, as this compound had not previously been reported in the literature before Chiang and co-workers [89] in 2009. Asperfuranone was later shown to possess bioactive properties as it inhibited proliferation of human non-small A549 cancer cells [90]. Investigating the loci containing putative PKS gene clusters, Chiang and co-workers [89] noticed that a NR-PKS gene (AN1034, afoE) and a HR-PKS gene (AN1036, afoG) were located close to each other on chromosome VIII. Since no products had ever been detected from activity of this locus, and due to the rare constellation of two neighbor PKSs, the authors speculated whether a novel metabolite could be revealed. A putative transcriptional activator (AN1029, afoA) was found near the PKS and the authors replaced the upstream sequence of afoA, estimated to be the native promoter, with the inducible alcA promoter [89]. This indeed turned on the expression of the cluster, since asperfuranone and a precursor metabolite were detected. The structure of asperfuranone was determined based on one- and two-dimensional NMR experiments and the absolute configuration by a modified Mosher’s method, whereas the precursor preasperfuranone had already been determined in the literature [89,91]. With these two compounds being identified, a gene-deletion strategy was performed to map the other genes assigned to the afo gene cluster, which involved twelve surrounding genes including the two PKSs [89]. Four deletion strains afoDΔ, afoEΔ, afoFΔ, and afoGΔ fully eliminated asperfuranone production whereas afoBΔ and afoCΔ strongly reduced production of asperfuranone [82]. The deletions confirmed that both afoE and afoG were responsible for the production of asperfuranone, and that the deletion of afoD, encoding a putative hydroxylase, resulted in the production of preasperfuranone [89]. Deletion of afoB reduced the production of asperfuranone and due to a high homology to efflux pumps, Chiang and co-workers [89] suggested that it was responsible for the transport of asperfuranone out of the cell.

With the gene cluster and predicted functionalities of the gene products defined, a biosynthetic pathway of asperfuranone was proposed (Figure 8) [89]. The assembly of the primary reduced tetraketide is synthesized by AfoG from one acetyl-CoA, three malonyl-CoA, and two SAM. The tetraketide is transferred to the SAT domain of AfoE and extended with four malonyl-CoA and one SAM [89]. The octaketide is released from AfoE after aldol condensation and reductive release from a C-terminal reductase (R) domain, which resembles a reductive release mechanism to generate the aldehydes described by Bailey et al. [92], forming the aldehyde preasperfuranone [89]. The biosynthetic steps from preasperfuranone to asperfuranone are uncharacterized and the suggestions are not based on identified metabolites [89]. Accumulation of preasperfuranone in the afoDΔ suggested AfoD to be the next enzyme in the biosynthesis of asperfuranone. The deletions of afoE, encoding a putative FAD/FMN-dependent oxygenase and afoC, initially believed to code for a homologue to citrinin biosynthesis oxidoreductase, did not reveal any intermediates, the order of reactions and the exact enzymatic functions for AfoF and AfoC have not been determined. In afoCΔ, the production of asperfuranone was not fully eliminated, which Chiang and co-workers [89] suggested could be due to other enzymes catalyzing the reaction, however less efficiently.
Figure 8. Proposed biosynthesis of asperfuranone. The highly reducing (HR)-PKS, AfoG, contains KS, AT, DH, CM, ER, KR, and ACP domains whereas the NR-PKS, AfoE, contains SAT, KS, AT, PT, ACP, CM, and R domains. The only intermediate isolated in the biosynthesis is preasperfuranone [89]. Multiple arrows indicate that the number of enzymatic steps and reaction order is unknown.

Other puzzling discoveries have been made in connection to the asperfuranone production. When trying to activate a cryptic NRPS gene cluster containing two NRPSs, *inpA* (AN3495) and *inpB* (AN3496), by overexpression of a regulatory gene, *scpR* (AN3492), Bergmann *et al.* [93] also activated asperfuranone. This is an interesting example of a regulatory gene located on chromosome II that activates the *afo* cluster located on chromosome VIII [93]. Lui *et al.* [94] have attempted to engineer the production of a new metabolite by swapping the SAT domain of AfoE with the StcA-SAT. This led to the production of a new metabolite though having the same length as the native AfoE product, asperfuranone [94].

8. Monodictyphenone/Emodin

The PK monodictyphenone was first reported in *A. nidulans* in 2005 [95] and the genes behind the production of monodictyphenone were mapped four years later [96]. This discovery not only enabled the establishment of a biosynthesis model for monodictyphenone in *A. nidulans*, it has subsequently revealed that more than ten different stable products among different classes of related polyketides can be linked to monodictyphenone biosynthesis [11,96–98]. These metabolites count monodictyphenone, emodin and the emodin derivatives 2-hydroxyemodin, 2-aminoemodin, α-hydroxy emodin, and emodic acid. Moreover, the arugosins and prenyl-xanthones are also coupled to the pathway [11,98]. The compound emodin has been studied for more than a century [99], and is an anthraquinone found in a wide array of both plants and fungi [100,101]. Emodin and several derivatives (e.g. emodic acid) have been shown to possess anti-bacterial and cancer preventive properties [102–106].
The presence of the SM clusters in silent areas of chromosomes, e.g. near telomeres and centromeres, suggests that chromatin remodeling factors can influence the expression of genes responsible for SMs. As rationalized by Bok and co-workers [96], removal of histone-tail methylation could open heterochromatic regions for transcription. The authors deleted an ortholog, cclA, to the yeast BRE2 gene, encoding an enzyme partner of the COMPASS transcriptional regulator complex conserved in eukaryotes, which rendered A. nidulans defective in di- and trimethylation of lysine 4 of the histone 3 tails (H3K4). The cclA deletion was established in a mutant strain, stcJΔ, to avoid interference of high amounts of sterigmatocystin in purification of other metabolites. The effect was striking as the loss of CclA in HPLC analysis showed an altered chemical landscape compared to the stcJΔ reference [96].

As the compounds appearing were UV-active suggesting high conjugation likely due to aromaticity, the ten NR-PKSs investigated in the asperthecin and asperfuranone studies were individually deleted in the cclAΔ stcJΔ double deletion background. This screen revealed six products that all were linked to one PKS, AN0150 (mdpG) [96]. Delineation of the cluster was achieved by inspecting the genome sequence for possible cluster candidates followed by Northern blotting for gene-expression analysis in the cclAΔ stcJΔ, where the products were detected. The cluster was found to span 12 putative ORFs (AN10021-AN10023 (mdp4-L)) [97] of which two genes AN0147 (mdpD) and AN10035 (mdpI), did not show altered expression from the reference [96,97]. The mdp-cluster candidates were also deleted in the cclAΔ stcJΔ mutant strain to confirm the expression analysis data and to draw the borders of the cluster [97]. The authors suggest that two transcriptional activators are present within the cluster; MdpE as a main activator (homologue to AflR) and that MdpA is a co-activator. The mdp locus is located near the telomere of chromosome VIII, and activation of the genes in the cclAΔ strain supports the hypothesis of epigenetic regulation in these areas through chromatin remodeling [96].

Two groups cultivated A. nidulans on complex growth media, which revealed six additional metabolites. First Sanchez et al. [98] discovered that emericillin, variecoxanthone A, shamixanthone, and epi-shamixanthone were also products of the mdp cluster, and subsequently Nielsen and co-workers [11] added arugosin A and H to the pathway. Since these PKs are prenylated, a BLAST search of the A. nidulans genome sequence was performed and pathway-candidate genes were deleted. Two prenyl transferases encoded by xptA and xptB, and one neighbor GMC oxidoreductase encoded by xptC, were found to be involved in the pathway, though they were located on other chromosomes than the mdp cluster (for cluster overview see Sanchez et al. [98]). This is an intriguing example of SM-cluster members located on more than one chromosome, however, prenyl transferases are known to have broad substrate specificity, and it is currently not known whether they are involved in other processes than prenyl-xanthone formation [98].

MdpG synthesizes the main PK backbone. Since MdpG lacks a CLC/TE domain, MdpF, a putative zinc dependent hydrolase, is believed to catalyze the release of the PK from MdpG [97]. The mechanism is believed to follow the case of ACAS and ACTE as introduced previously in the asperthecin section. Awakawa and co-workers [85] demonstrated that the direct product of the ACAS/ACTE is not emodin anthrone as proposed earlier [107,108], but more likely atrochrysone carboxylic acid (Figure 9). Atrochrysone carboxylic acid was not observed in vitro, instead the decarboxylated product atrochrysone was the major product in the assays and therefore proposed to be an intermediate to emodin, as suggested by Couch and Gautier [109]. Conversion of atrochrysone to
emodin requires dehydration (forming emodin anthrone) and a final oxidation (Figure 9), however, Awakawa and co-workers [85] observed small amounts of both emodin anthrone and emodin in vitro showing that these reactions may occur non-enzymatically. Based on these observations, Chiang and co-workers [97] proposed that mdpH encodes a decarboxylase, catalyzing the conversion of atrochrysone carboxylic acid to atrochrysone. The deletion of mdpH resulted in accumulation of a shunt product endocrocin produced via endocrocin anthrone. Enzymes responsible for dehydration of atrochrysone or modification of emodin into the observed derivatives have not yet been identified.

Figure 9. Proposed biosynthesis of emodin. MdpG contains SAT, KS, AT, PT, and ACP domains.

The first stable intermediate following emodin towards the prenyl-xanthones is monodictyphenone [97,98], and gene-deletion studies points to at least the five following enzymes are involved; a dehydratase (MdpB), a ketoreductase (MdpC), a glutathione S transferase (MdpJ), an oxidoreductase (MdpK), and a Baeyer-Villiger oxidase (MdpL) [97]. The mechanism has been proposed to be analogous to the conversion of versicolorin A to demethylsterigmatocystin which is known to proceed through oxidation-reduction-oxidation catalyzed by a cytochrome P450 monoxygenase (VerA) and a ketoreductase (StcU) [45,65,110]. However, none of the above mentioned Mdp enzymes appear to be homologous to VerA, and the role of the individual enzymes has not been investigated further [97].

The biosynthesis of the six monodictophenone derived metabolites is based on hydroxylation (MdpD), C-prenylation (XptA), O-prenylation (XptB), and carboxylic acid reduction (unidentified enzyme) [11,98]. Central in the pathway is the hydroxylation of C2 in monodictyphenone accompanied by reduction of the carboxylic acid. The carboxylic acid is suggested by Sanchez and co-workers [98] to be reduced to a hydroxy group, the B-ring is closed by dehydration and the
intermediate is O-prenylated at C2 to yield variecoxanthone A, which in turn is C-prenylated to emericillin (Figure 10). The final known step in prenyl-xanthone biosynthesis gives rise to the stereoisomers shamixanthone and epishamixanthone and is catalyzed by XptC \[98\]. Alternatively, Nielsen and co-workers include synthesis of arugosins by partially reducing the carboxylic acid to an aldehyde, followed by C-prenylation, yielding arugosin H and O-prenylation to give arugosin A. Subsequent reduction of the aldehyde to a hydroxyl group, and ring closure by dehydration then gives emericillin and shamixanthones \[11\].

**Figure 10.** Suggested biosynthesis of the shamixanthons from emodin. Multiple arrows indicate that the number of enzymatic steps are unknown.
9. Orsellinic Acid

In addition to the *mdp* cluster, the loss of CclA also led to the discovery of another gene cluster driven by an NR-PKS [96]. Two PK products, the cathepsin K inhibitors F-9775A and F-9775B, first isolated from *Paecilomyces carneus* [111], were detected and mapped to AN7909. Following this discovery, Schroeckh and co-workers [112] found the primary metabolite from AN7909 (*orsA*) to be orsellinic acid, an archetypal metabolite [113]. Moreover, the metabolite lecanoric acid typically found in lichens and produced by mycobiots such as *Umbilicaria antarctica* [114] was linked to OrsA [112]. Following the initial observations, the number of detected metabolites from the *orsA* gene cluster has expanded to gerfelin, a C10-deoxy-gerfelin, doricinol, orcinol, cordyol C, and violaceol I and II [115,116]. The biosynthetic activities of the *orsA* cluster are as yet not elucidated, and this illustrates the need for applying different eloquent strategies to trigger production of these metabolites.

The deficiency in methylation of H3K4 in the *cclA*Δ strain resulted in activation of both *mdp* and *ors* gene clusters. Expression analysis revealed that the annotated ORFs from AN7909-AN7915 were possible cluster members, hereby indicating candidates for a gene cluster [96]. The *ors* gene cluster, *orsB*-*orsE*, was identified by Schroeckh and co-workers [112] as four additional ORFs spanning AN7911-AN7914, which was confirmed by Sanchez *et al.* [115], who deleted all genes from AN7901 to AN7915. Interestingly, the neighbor PKS to the *ors* locus, AN7903, was deleted by Nielsen *et al.* [11] and the resulting strain failed to produce F-9775A and B like AN7909Δ under the conditions tested. Schroeckh and co-workers [112] defined *orsA*-*E* using gene-expression analysis through both an *Aspergillus* secondary metabolism array (ASMA) and relative expression analysis in quantitative reverse-transcriptase PCR (qRT-PCR). The induction of *orsA* was achieved by co-cultivating with a soil bacterium, *Streptomyces rapamycinicus* (initially named *S. hygroscopicus*) and extracting mRNA from the fungus [112]. This response on SM level was further investigated by Nützmann and co-workers [117]. Since the loss of H3K4 methylation induced gene expression in *ors* locus [96], the rationale was that the transcriptional activation of silent secondary-metabolism genes by acetylation of lysines on histone tails, especially H3K9, is equally important and the search for histone acetyl transferases (HATs) in the genome sequence was commenced [117]. Forty HATs were found and deleted, and only four proved to be essential. Of the 36 deletions of nonessential HATs in *A. nidulans*, the deletions of *gcnE* and *adaA*, both essential core parts of the multi-subunit Saga/Ada complex, an important complex for HAT activity in *A. nidulans*, significantly lowered the *ors* transcripts investigated [117]. Thus, Saga/Ada plays a role in the response to *S. rapamycinicus* and loss of this complex downregulated orsellinic acid metabolites, as well as sterigmatocystin, terrequinone, and penicillin [117].

Four additional orsellinic acid derived compounds were found in a defect COP9 signalosome (CSN) mutant strain [116]. The multiunit CSN complex is found in higher eukaryotes, albeit with different functional roles depending on the tissues. In *A. nidulans* the CSN is required for fruiting body formation and is not essential for asexual growth. By deleting *csnE*; orcinol, cordyol C, and violaceol I+II were produced, and the genes *orsA*-*orsE* were shown to be differentially expressed [116]. The link of the violaceol metabolites to *ors* was confirmed by Nielsen and co-workers [11] who applied an OSMAC strategy on their reference strain and compared this to their deletion library.
Very little is known about the biosynthesis of the metabolites of the ors locus. One acetyl-CoA and three malonyl-CoA units can yield a C8 aldol intermediate, and as proposed by Nielsen et al. [11], this can lead to the tetraketide orsellinic acid through loss of water and enolization and to the C7 compound orcinol by decarboxylation and enolization. Oxidation of orcinol in the para position then leads to 5-methyl-benzene-1,2,3-triol which is believed to either dimerize with the loss of water to give violaceol I and II, Figure 11, or to give F-9775A+B, Figure 12, in an unknown series of synthesis steps. Another outcome is the formation of lecanoric acid by dimerization of orsellinic acid. Though the steps in the pathway have been hypothesized, most steps are not accounted for. It has been reported that OrsA having the domains SAT-KS-AT-PT-ACP-TE is responsible solely to form orsellinic acid. OrsA, OrsB, and OrsC seem to be sole responsible for F-9775A+B formation. Moreover, it has been shown that gerfelin and a C10-deoxy derivative of gerfelin accumulate in orsBΔ, whereas diorcinol was found in high amounts in the orsCΔ strain [115]. Gerfelin, C10-deoxy gerfelin and cordyol C are all dimers built up of two of the three suggested monomer units, orsellinic acid, orcinol and 5-methylbenzene-1,2,3-triol.

Recently Scherlach and co-workers [118] continuously cultivated A. nidulans under nitrogen-limitation and carbon-limitation. At nitrogen limiting conditions in continuous cultivations two novel products, denoted as spiroanthrones, were found. They could not be detected at batch cultivation. The metabolites were based on anthraquinone and orsellinic acid derived phenols. The induced expression of both mdpG and orsA confirmed increased activity under the N-limiting continuous cultivation conditions.

10. Austinol and Dehydroaustinol

The meroterpenoids austinol and the related compound dehydroaustin were first isolated from A. ustus by Simpson and co-workers in 1982 [119], where the structure of austinol was elucidated by $^1$H and $^{13}$C NMR. Austinol and dehydroaustinol are just two examples out of many meroterpenoids that are derived from 3,5-dimethyl orsellinic acid as presented in the excellent review by Geris and Simpson [120].

The two austinols were detected for the first time in A. nidulans four years ago [121], where it was further substantiated that austinol was indeed of partly PK origin. Deletion of the phosphopantetheinyl transferase (PPTase) cfwA/npgA in A. nidulans resulted in a strain that among many other compounds did not produce austinol and dehydroaustinol [121]. The PPTase is responsible for attaching the phosphopantetheine moiety to the acyl carrier domain of the PKSs and NRPSs, thus it is an activator of the enzyme complexes. Hence, the abolition of austinol and dehydroaustinol production in the PPTase deficient strain strongly suggests a PK origin of these compounds.

In 2011, Nielsen and co-workers [11] discovered the PKS responsible for synthesis of the PK part of austinol and dehydroaustinol in A. nidulans. A deletion library of all 32 putative PKS genes in A. nidulans was created and screened using an OSMAC approach [122] to enable activation of different clusters on different media. One single strain deleted in AN8383 (ausA) failed to produce austinol and dehydroaustinol [11]. This discovery was supported by the introduction of a point mutation at the phosphopantetheine attachment site, thus abolishing activation of the enzyme by the PPTase to ensure the loss of austinols was not an indirect effect e.g., at chromatin level. The ausAΔ
strain was complemented by re-introducing the *ausA* gene into the deletion strain. Introducing the gene under control of the inducible *alcA* promoter revealed that 3,5-dimethyl orsellinic acid (3,5-MOA) was indeed the precursor for austinol and dehydroaustinol in *A. nidulans*, as shown experimentally via labeling studies by Simpson and co-workers 20 years earlier [11,123,124].

**Figure 11.** Proposed biosynthesis of orsellinic acid and its derivatives of orsellinic acid. The enzymes that catalyze the individual reactions in the biosynthesis of the metabolites are so far unknown and biosynthesis is proposed based on the observed metabolites.
A more detailed understanding of the biosynthesis of the austinols has not been established yet. However, it is well established that 3,5-dimethyl orsellinic acid is synthesized from condensation of an acetyl-CoA unit with 3 malonyl-CoA units to form the PK backbone, which is methylated twice, catalyzed by AusA. The PK part is then alkylated with farnesyl pyrophosphate to form a transient intermediate (Figure 13), which can then act as a precursor for several similar meroterpenoids, such as andibenins, austin, berkeleyones and andrastins [120]. Recently we tentatively identified neoaustin and austinolide in A. nidulans extracts by LC-MS analysis (unpublished data), which makes us hypothesize that the biosynthesis towards the austinols involves (i) oxidation and acyl shift in the D ring (ii) lactonization from the substituent groups of the D ring, a Baeyer-Villiger type oxidation and 1,2 alkyl shift in the A ring to give neoaustin. Neoaustin is subsequently oxidized in the D ring by another Baeyer-Villiger type oxidation to give austinolide that upon further oxidation and ring condensation leads to austinol and dehydroaustinol, Figure 13.

11. Concluding Remarks

Secondary metabolism represents chemical diversity and span in biological functionality to the extreme. As shown above, individual compound classes can even form hybrid molecules to other compound classes. There is a high commercial interest in discovery and utilization of SMs in general as drugs or additives, or to avoid mycotoxins in food and feeds. Mapping PK biosynthesis to genes in A. nidulans as presented in this review involves great complexity. One challenge is to find and activate the genes required to produce the compounds. As shown, it takes in-depth understanding of fungal biology, nutrient sensing, chromatin remodeling, as well as analysis on all levels from DNA to metabolites to unveil cryptic gene clusters and their products. Moreover, the majority of the pathways described in this review have been elucidated in a relatively short time span. This has been possible due to bioinformatics. The availability of the genome sequence, as well as resources and tools in, e.g., Aspergillus Comparative Database (ACD), Aspergillus Genome Database (AspGD), and the Central Aspergillus REsource (CADRE) have been key aids to perform the extensive genome mining.

The ability to predict enzymatic function based on gene sequences has proven fruitful in characterization of secondary metabolism, since this has revealed the location of e.g., PKS genes. Additionally, the presence of gene cluster specific TFs was utilized in activation of silent SM clusters both in the case of aspyridones and asperfuraneone. The asperfuraneone biosynthesis is moreover an example of a previously unknown compound, with a potential to be a novel drug, has been found in a
well-known filamentous fungus. Another case of a metabolite with attractive properties is emodin, which has been known for more than a century, but just recently had the biosynthetic machinery uncovered in A. nidulans. Both asperthecin and the emericellamides were firstly discovered and isolated from less well described aspergilli, however, after the compounds were observed in A. nidulans, the candidates for responsible genes in biosynthesis were found in a few months. The study of conidial pigment biosynthesis in A. nidulans has contributed to our basic understanding of fungal development and PKS organization, and provided researchers with an easy assayable marker system for genetic studies. Interestingly, the structure of the final pigment(s) still remains unknown after more than 70 years of research, underlining the difficulties in elucidating structures of highly polymerized PKs.

**Figure 13.** Proposed biosynthesis of austinol and dehydroaustinol. All genes in the biosynthesis of austinol and dehydroaustinol, except AusA, are unknown.
Epigenetic regulation through chromatin remodeling has shown to be involved in activation of several SM gene clusters. Conserved chromatin remodeling factors have influenced both local activation of some SM clusters and more global responses within the genome. Gene clusters producing sterigmatocystin, orsellinic acid, emodins, and austins have shown to respond to specific factors. The presence of SM producing genes outside gene clusters, e.g., in prenyl-xanthone production, is probably more common than observed so far. Moreover, cross-talk between pathways is frequently observed, as more pathways become known. This can open a discussion whether common pools of intermediates or enzymes can exist. In addition, the compartmentalization of SM production is an area to be explored. Furthermore, controlling compartmentalization of production as well as secretion will influence yields and downstream applications which are important factors for exploiting SM production commercially. Existing compounds can be modified genetically to add/remove chemical groups on existing drugs, mix moieties from different SM classes or species by for example domain swapping, and to considerably increase/abolish a specific production.

Altogether the recent uncovering of secondary metabolism in *A. nidulans* is an illustrative example of strong interdisciplinary efforts requiring strong expertise within chemistry, biology, microbiology, molecular genetics, protein chemistry, computer science, and engineering. Ultimately, the efforts described in this review can form the basis for uncovering of the specific biological roles of the chemical arsenal in the fungus.

Acknowledgments

This work was supported by the Danish Research Agency for Technology and Production grant # 09-064967, FOBI, DTU, and Novozymes. Additionally, we thank Maria Månsson for critical proof reading of the manuscript and for suggestions.

Conflict of Interest

The authors declare no conflict of interest.

References and Notes


   adhA in Aspergillus parasiticus is involved in conversion of 5'-hydroxyaverantin to averufin. 
60. Sakuno, E.; Yabe, K.; Nakajima, H. Involvement of two cytosolic enzymes and a novel 
   intermediate 5'-oxoaverantin in the pathway from 5'-hydroxyaverantin to averufin in aflatoxin 
   versiconal hemiacetal acetate to versiconal and versiconol acetate to versiconol in aflatoxin 
62. Silva, J.C.; Minto, R.E.; Barry, C.E., III; Holland, K.A.; Townsend, C.A. Isolation and 
   encodes a putative cytochrome P-450 monooxygenase required for bisfuran desaturation during 
64. Keller, N.P.; Kantz, N.J.; Adams, T.H. Aspergillus nidulans verA is required for production of the 
   required for the conversion of versicolorin A to sterigmatocystin in Aspergillus nidulans. 
67. Slot, J.C.; Rokas, A. Horizontal transfer of a large and highly toxic secondary metabolic gene 
68. Bergmann, S.; Schümann, J.; Scherlach, K.; Lange, C.; Brakhage, A.A.; Hertweck, C. Genomics- 
   Simiityan, H.; Kuo, E.; Praseuth, A.; et al. Molecular genetic mining of the Aspergillus secondary 
71. Song, Z.; Cox, R.J.; Lazarus, C.M.; Simpson, T.J. Fusarin C biosynthesis in Fusarium moniliforme and 
72. 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-
73. Xu, W.; Cai, X.; Jung, M.E.; Tang, Y. Analysis of intact and dissected fungal polyketide 


© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).
Paper 2

“Comparative chemistry of Aspergillus oryzae (RIB40) and A. flavus (NRRL 3357)”

C Rank and ML Klejnstrup (joint 1st author), LM Petersen, S Kildgaard, JC Frisvad, CH Gotfredsen, and TO Larsen

Accepted in Metabolites, 2, 39-56 (2012)
Comparative Chemistry of *Aspergillus oryzae* (RIB40) and *A. flavus* (NRRL 3357)

Christian Rank 1,†, Marie Louise Klejnstrup 1,†, Lene Maj Petersen 1, Sara Kildgaard 1, Jens Christian Frisvad 1, Charlotte Held Gottfredsen 2 and Thomas Ostenfeld Larsen 1,*

1 Department of Systems Biology, Center for Microbial Biotechnology, Technical University of Denmark, Soltofts Plads B221, DK-2800 Kgs. Lyngby, Denmark; E-Mails: christian.rank@gmail.com (C.R.); marlk@bio.dtu.dk (M.L.K.); lmape@bio.dtu.dk (L.M.P.); sarakildgaard@hotmail.com (S.K.); jcf@bio.dtu.dk (J.C.F.)

2 Department of Chemistry, Technical University of Denmark, Kemitorvet B201, DK-2800 Kgs. Lyngby, Denmark; E-Mail: chg@kemi.dtu.dk

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: tol@bio.dtu.dk; Tel.: +45-4525-2632; Fax: +45-4588-4148.

Received: 18 November 2011; in revised form: 14 December 2011 / Accepted: 22 December 2011 / Published: 5 January 2012

Abstract: *Aspergillus oryzae* and *A. flavus* are important species in industrial biotechnology and food safety and have been some of the first aspergilli to be fully genome sequenced. Bioinformatic analysis has revealed 99.5% gene homology between the two species pointing towards a large coherence in the secondary metabolite production. In this study we report on the first comparison of secondary metabolite production between the full genome sequenced strains of *A. oryzae* (RIB40) and *A. flavus* (NRRL 3357). Surprisingly, the overall chemical profiles of the two strains were mostly very different across 15 growth conditions. Contrary to previous studies we found the aflatrem precursor 13-desoxypaxilline to be a major metabolite from *A. oryzae* under certain growth conditions. For the first time, we additionally report *A. oryzae* to produce parasiticolide A and two new analogues hereof, along with four new alkaloids related to the *A. flavus* metabolites ditryptophenalines and miyakamides. Generally the secondary metabolite capability of *A. oryzae* presents several novel end products likely to result from the domestication process from *A. flavus*. 
Keywords: Aspergillus oryzae; (RIB40); Aspergillus flavus; (NRRL 3357); parasiticolide; ditryptoleucine; oryzamide

1. Introduction

Aspergillus oryzae is one of industry’s most used “workhorses” and has been used for centuries in food fermentation for the production of e.g., sake, soy sauce and other traditional Asian foods [1]. A. oryzae is also a widely used organism for production of amylase, lipases and proteases and more recently also for heterologous expression of secondary metabolite genes and non-fungal proteins [2–4]. For many years, A. oryzae has been suspected to be a domesticated form of A. flavus, a plant and mammalian pathogenic saprophyte, capable of producing some of the most carcinogenic compounds known: the aflatoxins. Genetic work and subsequent genome sequencing of strains of both species have verified the tight link between the species [1,5–8].

The relationship of the two species has resulted in extensive screening of the toxic potential of A. oryzae, but no genuine evidence of aflatoxin production in validated A. oryzae isolates has ever been shown. Other important toxins, known from A. flavus, have on the other hand been shown in A. oryzae: 3-Nitropropionic acid [9] and cyclopiazonic acid (CPA) [10] along with kojic acid [11,12] (Figure 1). Additional metabolites previously reported from A. oryzae are asperfurane [13], sporogen AO1 [14,15], maltoryzine [16], and aspergillomarasmine A [17,18]. Aspirochlorine has been found in A. flavus, A. oryzae and A. tamarii [19–23] (Figure 1). For reviews on the safety and taxonomy of A. oryzae, see [7,24–26].

Figure 1. Known compounds from Aspergillus flavus or A. oryzae.
The few predicted differences between the genomes of *A. oryzae* and *A. flavus* (ca. 99.5% genome homology and 98% at the protein level for RIB40/ATCC 42149 and NRRL 3357 [27]), could lead one to expect *A. oryzae* to produce most of the metabolites found in *A. flavus* [1,28-31], but published metabolic data indicates a very low chemical correlation [32]. It is with reference to the established genetic heritage of *A. oryzae* from *A. flavus* remarkable that maltoryzine, sporogen AO1, asperfuran and aspergillomarasmine A have never been unambiguously identified in *A. flavus*. Though research on *A. flavus* chemistry has focused primarily on toxic compounds, one would expect that these metabolites should be part of its chemical potential as they are for the domesticated *A. oryzae*. The preliminary bioinformatic studies in conjunction with the genome sequencing shows roughly the same number of predicted genes: 32 Polyketide synthases (PKSs) and 28 non-ribosomal synthases (NRPSs) for *A. flavus* and 32 PKSs and 27 NRPSs for *A. oryzae* with 2 NRPSs apparently unique for each strain [33]. The exclusiveness of these genes in terms of end product has yet to be verified chemically.

Most of the predicted genes for secondary metabolites of *A. oryzae* (or *A. flavus*) have not been mapped to specific metabolic products, despite the genome sequencing of RIB40 in 2005 [27]. Only genes of the most important toxins: Aflatoxin [31,34,35], CPA [36,37] and aflatrem [38] have been fully annotated in both species, which leaves much to be explored. The full chemical potential of either species is unknown and epigenetic modifiers [39-43] may be necessary, alongside with the use of different growth conditions to aid triggering the full potential of secondary metabolite expression in these two closely related species.

The aim of the current work has been to perform an initial comparative investigation of the chemistry from the two genome sequenced strains of *A. oryzae* (RIB40) and *A. flavus* (NRRL 3357), in order to get further insights into possible homologies and differences in secondary metabolite production for these two important species.

### 2. Results and Discussion

#### 2.1. De-Replication of *A. oryzae* RIB40

For the analysis of *A. oryzae* RIB40 chemistry, we investigated a series of solid media (YES, YESBEE, DRYES, CYA, CYAS, CY20, CY40, DUL, GAK, GMMS, MEA, OAT, PDA, TGY, WATM (see Methods and Materials for explanation) cultivations with micro-scale extractions [44,45] and subsequently analyzed with HPLC-DAD-MS for selection of optimal conditions. The different media were tested on a collection of *A. oryzae* (RIB40, IBT 28103) and *A. flavus* (NRRL 3357, IBT 23106, IBT 3642) and these strains were cultivated at 25 °C in the dark for 7 and 14 days. The media screening for *A. oryzae* and *A. flavus* indicated the greatest chemodiversity and metabolite production from CYA, YES and WATM agar for our purpose.

The comparison of the secondary metabolite profiles of the two strains, NRRL 3357 and RIB40, exposed a surprisingly high degree of chemical difference on all media as illustrated in Figure 2 and Table 1 for the WATM medium. The major metabolite repetitions between the two genome sequenced strains were merely kojic acid and ergosterol, like a number of minor metabolites (not analyzed here) seemed to be shared between the two strains. Altogether, this is in sharp contrast to the high gene homology, particularly for the secondary metabolite genes.
Figure 2. ESI+ BPC chromatogram of 7 day micro scale extract from WATM, bottom: *A. oryzae* RIB40, top: *A. flavus* NRRL 3357. Besides kojic acid (shown in box) and analogues in the beginning of chromatogram and ergosterol in the end (not shown), there are very few identical compounds between the genetically almost identical strains. Note that aspirochlorine is only detectable in negative ionization, and therefore not visible in this chromatogram.

Table 1. LC-MS de-replication of some of the important secondary metabolite pathways from the two full genome sequenced siblings, *A. flavus* (NRRL 3357) and *A. oryzae* (RIB40). Based on 7 day fermentation on solid WATM agar in the dark. (+) indicates the presence of these types of compounds in *A. flavus* based on UV spectroscopic analysis. * New compounds reported here for the first time.

<table>
<thead>
<tr>
<th>Metabolite</th>
<th><em>A. flavus</em> NRRL 3357</th>
<th><em>A. oryzae</em> RIB40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kojic acid</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Aflatoxin</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Aflavinines</td>
<td>(+)</td>
<td>+</td>
</tr>
<tr>
<td>Aflatrem</td>
<td>+</td>
<td>13-desoxypaxilline</td>
</tr>
<tr>
<td>Miyakamides</td>
<td>(+)</td>
<td>Oryzamides *</td>
</tr>
<tr>
<td>Aspirochlorine</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Cyclopiazonic acid</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Ditryptophenaline</td>
<td>+</td>
<td>Ditryptoleucine *</td>
</tr>
<tr>
<td>Parasiticolide A</td>
<td>-</td>
<td>14-deacetyl parasiticolide A *</td>
</tr>
</tbody>
</table>
Known metabolites to *A. oryzae* were de-replicated and we found that the RIB40 strain did not produce detectable levels (LC-MS) of CPA (as also noted by Tokuoka *et al.* [36]), asperfuran, sporogen AO1, maltoryzine or aspergillomarasmine under these growth conditions. It did, however, produce kojic acid and aspirochlorine and a series of potentially new metabolites of which some were isolated, structurally characterized and reported here.

### 2.2. New Metabolites to *A. oryzae* RIB40

During fermentation of the chemically potent RIB40 strain, we have been interested in the tremorgenic compounds, allegedly coupled to fungal sclerotia [46–54] and whether these could be found in *A. oryzae* as they have been in *A. flavus*. The RIB40 strain produces large and abundant sclerotia, especially on WATM agar, a fact not widely announced in literature although sclerotia have been observed in *A. oryzae* sporadically [55–57]. No sclerotia were observable after 14 days on YES agar, but although these metabolites are often characterized as sclerotial metabolites, there is not a strict correlation between the biosynthesis of these metabolites and the formation of sclerotia, as also noted by Wilson [58], and this extract was used for the described isolations.

Here, we report the discovery of the aflatrem precursor 13-desoxypaxilline (13-dehydroxypaxilline) in *A. oryzae* RIB40, originally isolated from *Penicillium paxilli* [59–63]. Aflatrem is known from *A. flavus* and was discovered by Wilson and Wilson in 1964 [64] and structure elucidated by Gallagher *et al.* in 1978 and 1980 [65,66]. 13-desoxypaxilline was present in YES, CYA, OAT and WATM agar 7 day old micro-scale extracts. From the 14 days old YES 200 plate extract used for isolation, 13-desoxypaxilline was recovered as an intermediate metabolite. LC-MS, LC-MS/MS and NMR data analysis (Supplementary Material) confirmed the structure. Naturally the prospect of finding aflatrem itself was investigated, though no apparent peak was visible in HPLC-DAD data files. The use of a LC-MS/MS method further confirmed 13-desoxypaxilline as an end-product of *A. oryzae* RIB40 for the above cultivation conditions, since none of the proposed intermediate steps towards aflatrem could be detected (LC-MS/MS) and only one sample (WATM, 7d) showed traces of paspaline, a precursor for 13-desoxypaxilline (Figure 3).

**Figure 3.** The final steps in the proposed biosynthesis of aflatrem (in *A. flavus*). *A. oryzae* RIB40 biosynthesis stops at 13-desoxypaxilline [38].
A second extract was made from 100 plates of a 14-day old A. oryzae RIB40 culture grown on WATM agar with abundant sclerotia formation to validate the findings from the YES extract. The analysis of the WATM extract showed 13-desoxypaxilline as a major metabolite alongside other sclerotium-related metabolites, such as aflavinines (based on UV-data not shown). The discovery of 13-desoxypaxilline as the apparent end-product of A. oryzae RIB40 is in agreement with the analysis of Nicholson et al. [38], who showed that a frameshift mutation in the atmQ gene presumably accounts for 13-desoxypaxilline not being converted into paspalicine and paspalinine. This mutation is therefore likely responsible for terminating the aflatem biosynthesis in RIB40 prematurely, short of the acetal ring closure, C-13 hydroxylation and isoprene attachment. Nicholson et al. [38] demonstrated that AtmQ is a multifunctional cytochrome P450 monoxygenase likely to catalyze the several oxidative steps needed for biosynthesis of the acetal ring present in the structures of paspalicine, paspalinine and aflatem, altogether pointing towards a more complex biosynthesis than shown here. Contrary to our discovery, Nicholson et al. did not find the aflatem gene cluster of RIB40 to be transcribed during their fermentations [38].

The isolated 13-desoxypaxilline is a member of the paspalitrem tremorgens, a widely distributed group of metabolites that have been isolated from several genera: Penicillium, Eupenicillium, Claviceps, Emericella, Aspergillus and Phomopsis [67–69]. Besides the tremorgenic activity in animals, these metabolites have been shown to be insecticides [68], which is believed to be their ecological function together with aflatoxin and CPA for protection of the sclerotia against fungivorous insects [46,47].

In addition to 13-desoxypaxilline, two new analogues of parasiticolide A were also isolated and are here reported for the first time. The metabolites showed to be dide- and 14-deacetoxy analogues, and are most likely precursors to the sesquiterpene parasiticolide A (= astellolide A) (see Figure 4). The metabolites were present in CYA, YES and WATM extracts and isolated from the same 14 day old YES extract as 13-desoxypaxilline, and the dide- and 14-deacetyl analogues were also found in the sclerotia enriched WATM extract. Again the metabolites were analyzed using LC-MS and NMR. Several different extraction procedures were tested to verify the correctness of the compounds as genuine metabolites and not as in vitro degraded parasiticolide A products, but all samples showed only dide- and 14-deacetyl parasiticolide A and no traceable (LC-MS) levels of parasiticolide A itself, even with different non-acidic extractions. Parasiticolide A have been isolated from A. flavus var. columnaris (FKI-0739) once [71] and was originally isolated and characterized from A. parasiticus (IFO 4082) [72,73] and later also from a mutant of Emericella variecolor (= A. stellatus Curzi) [74]. Recently parasiticolides have been detected in the newly described species A. arachidcola (CBS 117610) and A. minisclerotigenes (CBS 117635) [75]. There have to our knowledge not been published any toxic studies on the parasiticolides, but the related peniopholides from the fungus Peniophora polygonia have been reported to have antifungal properties [76].
In our observations parasiticolides are more often detectable metabolites of *A. oryzae* than of *A. flavus* under the same fermentation conditions, suggesting that the pathway is partly silenced for *A. flavus* and may need epigenetic modification to be expressed under otherwise normal growth conditions. It is interesting that parasiticolide A is scarcely observed in *A. flavus*, when it is an important product of *A. oryzae* and also of *A. parasiticus*. As for 13-desoxyxanthonine, dide- and 14-deacetyl parasiticolide A are almost certainly products of a prematurely ended biosynthesis, here parasiticolide A. We also isolated and elucidated a third parasiticolide A analogue; a formyl variant of parasiticolide A, but it was not possible to exclude the possibility of in vitro chemistry due to the formic acid added during the ethyl acetate extraction, so the correctness of this metabolite remains tentative. (Hamasaki *et al.* used benzene to extract parasiticolide A [77]). All parasiticolide analogues were structure elucidated by NMR and shift values correlated with the published data for parasiticolide A, except for the missing signals of the acetate units and their minor influence on the chemical shifts values of adjacent protons and carbons (See Supplementary Material for NMR data).

To further verify these observations, a MS/MS method was used to analyze several different microscale extracts of RIB40 for parasiticolide A itself. Trace amounts of parasiticolide A was found under these conditions and compared to an isolate of *A. parasiticus* (IBT 4387) capable of producing parasiticolide A. In the *A. parasiticus* isolate no dide- or 14-deacetyl parasiticolide A could be measured, indicating a complete transformation into the end-product (Figure 4). The small amount of parasiticolide A in RIB40 (roughly 1:1.000 ratio compared to 14-acetyl parasiticolide A, presuming the same response factor) might be the result of spontaneous acetylation involving the first acetylating enzyme. When the gene cluster of this metabolite is mapped, it is likely that the gene responsible for the last (specific) acetylation will be found to be mutated. Except for the section *Nidulantes* member *Emericella variecolor*, all other producers of these metabolites have been members of section *Flavi*. No indication points to these metabolites being part of the previously mentioned sclerotium metabolites since they are not found in selective sclerotium extracts, but are found for example in *A. arachidicola*, which is not known to produce sclerotia.

Four new *A. flavus* related NRPS compounds were also isolated from *A. oryzae* RIB40 and characterized based on standard interpretation of $^1H$, COSY, HMQC and HMBC NMR data (see Supplementary Material). The compounds showed to be two isomeric metabolites named ditryptoleucine related to the ditryptophenalines [78–81] and the two indole-enamides named oryzamides A$_{1,2}$ (cis- and trans-forms) structurally similar to the antibiotic miyakamide B$_{1,2}$ [71] (Figure 5).
**Figure 5.** Structure of the new *A. oryzae* metabolites: Ditryptoleucine and oryzamide A1-2.

It proved difficult to unambiguously isolate the cis- and trans-isomer of oryzamide A in a pure form due to isomerization around the double bond. Due to this the structure elucidation has been performed on mixtures of compounds. It is evident from the structure elucidation that oryzamides A1-2 only differ in the configuration around the double bond of the enamine as evident from the size of the coupling constant. Marfey analysis of the amino acid derived compounds only established the presence of L-tyrosine in the oryzamides. NMR structural data on the trans isomer (A1) is present in the supporting material.

Intriguingly, these *A. oryzae* metabolites have apparently exchanged phenylalanine with leucine compared to the similar *A. flavus* metabolites, indicating either a common trait in the domestication process or a coupling between the two pathways. The diketopiperazine ditryptoleucine was isolated in two variants with 1H-NMR shifts varying around the C-C dimeric bond, but with the same base structure. This unspecific dimerization is in line with previously isolated compounds. A hybrid between the ditryptophenaline and ditryptoleucine was isolated from an *Aspergillus* sp. as WIN 64745, with both a phenylalanine and leucine moiety, but with no N-methylation [79]. These compounds were tested and proved antagonistic against substance P at the NK1 receptor [82]. The oryzamides A1-2 are variants of the miyakamides B1-2 [71]. The exchanged phenylalanine has been substituted with leucine between a tryptophan and a tyrosine in the oryzamides. However this exchange of phenylalanine and leucine has often been found for compounds produced within the same species. *Penicillium polonicum* (= *P. fructigenum*) [83-85] produces both fructigenine A = puberuline = rugulosuvine A, containing phenylalanine and fructigenine B = verrucofortine, containing leucine. We could also detect two slightly later eluting compounds in the *A. oryzae* RIB40 extracts likely to be two further indole-enamides having phenylalanine incorporated instead of tyrosine as evident from LC-DAD-MS analysis (data not shown) as also seen for the miyakamides [71].

3. **Experimental**

General procedures and methods for analysis are described in [43,44,86].

Mass spectrometric analysis was performed using an Agilent HP 1100 liquid chromatograph with a DAD system (Waldbronn, Germany) on a LCT oaTOF mass spectrometer (Micromass, Manchester, UK) with a Z-spray ESI source and a LockSpray probe. For general procedures see [87]; method 1 for
Metabolites 2012, 2 47

LC-DAD-TOF was used in this study. All solvents used were HPLC grade from Sigma-Aldrich (St. Louis, MO, USA).

3.1. Fungal Material and Fermentation

A. oryzae RIB40 (IBT28103); A. flavus NRRL3357 (IBT3696), (IBT15934), NRRL 13462; A. parvisclerotigenus IBT16807 and A. minisclerotigenes IBT13353 were obtained from the IBT Culture Collection at DTU Systems Biology, Technical University of Denmark. The RIB40 isolate used for isolation of 13-dehydroxypaxilline was cultured for 14 days at 25 °C in the dark on 200 Petri dishes with Yeast Extract Sucrose agar (YES). All strains were grown for 1 week at 25 °C on YES, Czapek Yeast Autolysate (CYA), Wickerhams Antibiotic Test Medium (WATM) agar [88], YESBEE [89] (YES+50 g Bee pollen Type III, granulate, Sigma, P-8753, pr. 1 L medium), DRYES (Dichloran rose Bengal chloramphenicol agar), AFPA (Aspergillus flavus, A. parasiticus agar), CYAS (CYA + 50 g NaCl pr. 1 L medium), CY20 (CYA + 170 g sucrose pr. 1 L), CY40 (CYA with added 370 g sucrose pr. 1 L medium), DUL (Dulaney’s medium for Penicillin), GAK (Potato-carrot agar), GMMS (Glucose minimal media (GMM) + 2% sorbitol), MEA (Malt extract agar), OAT (Oat meal agar), PDA (Potato-dextrose agar). For medium formulations see Samson et al. [90].

3.2. Extraction and Isolation of Pure Compounds

13-Desoxypaxilline. The plates were homogenized using a Stomacher and 100 mL EtOAc with 1% HCO$_2$H pr. 10 plates. The extract was filtered and dried down on a freeze drier. The crude extract was separated on a KP-C18-HS 60 g SNAP column using a Biotage Isolera One (Biotage, Uppsala, Sweden), resulting in a 22 mg fraction. The fraction was segmented with a 10 g ISOL Diol column, using 12 steps of stepwise Heptane-dichloromethane-EtOAc-MeOH. 13-desoxypaxilline was predominant in a 100% EtOAc fraction (6 mg), and purified on a Waters HPLC W600/996PDA (Milford, MA, USA) and a RP column (Phenomenex Luna C18(2), 250 × 10 mm, 5 μm, Torrance, CA, USA) using a gradient of 80% MeCN (H$_2$O–Milli-Q (Millipore, MA, USA)) to 90% over 10 min. with 50 ppm TFA added to the solvents. The collection was concentrated on a rotary evaporator (Büchi V-855/R-215, Flawil, Switzerland) and dried under N$_2$ to yield 0.5 mg of white, amorphous 13-desoxypaxilline.

Dideacetyl-, 14-deacetyl-, and 18-formyl parasiticolide A. From the same fermentation described for 13-desoxypaxilline a more polar, 90 mg fraction was fractionated with a 10 g ISOL Diol column, using 12 steps of stepwise Heptane-dichloromethane-EtOAc-MeOH. The parasiticolide A-analogues were predominant in a 100% EtOAc fraction (10 mg), and purified on a Waters HPLC W600/996PDA (Milford, MA, USA) and a RP column (Phenomenex Luna C18(2), 250 × 10 mm, 5 μm, Torrance, CA, USA) using a gradient of 72% MeCN (H$_2$O–Milli-Q (Millipore, MA, USA)) to 87% over 15 min. with 50 ppm TFA. The collection was concentrated on a rotary evaporator (Büchi V-855/R-215) and dried under N$_2$ to yield 0.3, 1.0 and 0.8 mg of white, amorphous di-, 14-deactyl- and 18-formyl parasiticolide A, respectively. See Supplemental Material for NMR data.

Ditryptoleucine. 400 plates of A. oryzae RIB40 were cultured on YES medium. The plates were homogenized using a Stomacher and 100 mL EtOAc pr. 10 plates. The extract was filtered and dried
down on a freeze drier. The crude extract were separated into three phases by dissolving it in 9:1 MeOH:H₂O–Milli-Q and extracted into a heptan phase and afterwards a DCM phase. The DCM phase was separated on a KP-C18-HS SNAP column using a Biotage Isolera One (Biotage, Uppsala, Sweden) using a gradient of 10% MeOH (H₂O–Milli-Q (Millipore, MA, USA)) to 100% over 15 CVs (column volumes) resulting in a 402 mg fraction. The fraction was further separated on another KP-C18-HS SNAP column using a Biotage Isolera One (Biotage, Uppsala, Sweden) using a gradient of 10% MeOH (H₂O–Milli-Q (Millipore, MA, USA)) to 25% over 2 CVs, 25% MeOH to 50% over 8 CVs and 50% to 100% over 4 CVs resulting in a 60.1 mg fraction. The fraction was fractionated on a LH-20 column using 100% MeOH. The ditryptoleucines were present in the earlier fractions and were purified on a Waters HPLC W600/996PDA (Milford, MA, USA) using a RP column (Phenomenex Luna C18(2), 250 × 10 mm, 5 μm, Torrance, CA, USA) using a gradient of 50% MeCN (H₂O–Milli-Q (Millipore, MA, USA)) to 100% over 20 min. with 50 ppm TFA and a flow of 4 mL/min. The collections were concentrated on a rotary evaporator (Büchi V-855/R-215) and dried under N₂ to yield 4.0 mg of the two isomers. See Supplemental Data for NMR data.

**Oryzamides A₁,₂.** The remaining broth from the fermentation described for ditryptoleucine was extracted with EtOAc + 1% HCO₂H for 24 h. The extract was filtered and dried down on a freeze drier. The crude extract were separated into three phases by dissolving it in 9:1 MeOH:H₂O–Milli-Q and extracted into a heptan phase and afterwards a DCM phase. The DCM phase was separated on a KP-C18-HS SNAP column using a Biotage Isolera One (Biotage, Uppsala, Sweden) using a gradient of 10% MeOH (H₂O–Milli-Q (Millipore, MA, USA)) to 100% over 15 CVs (column volumes) resulting in a 654 mg fraction. The fraction was fractionated with a 10 g ISOL Diol column, using 13 steps of stepwise Hexane-dichloromethane-EtOAc-MeOH. The miyakamides were predominant in the 40:60 DCM:EtOAc and 20:80 DCM:EtOAc fractions (43.8 mg), and purified on a Waters HPLC W600/996PDA (Milford, MA, USA) and a RP column (Phenomenex Luna C18(2), 250 × 10 mm, 5 μm, Torrance, CA, USA) using a gradient of 45% MeCN (H₂O–Milli-Q (Millipore, MA, USA)) to 62% over 15 min. with 50 ppm TFA. The collection was concentrated on a rotavap (Büchi V-855/R-215) and dried down under N₂ to yield 4.0 and 2.0 mg of oryzamide A₁ and A₂, respectively. See Supplemental Data for NMR data.

**3.3. Marfeys Method**

100 μg of each peptide was hydrolysed with 200 μL 6 M HCl at 110 °C for 20 h. To the hydrolysis products (or 50 μL (2.5 μmol) solutions of standard D- and L-amino acids) was added 50 μL water, 20 μL 1 M NaHCO₃ solution and 100 μL 1% FDAA in acetone, followed by reaction at 40 °C for 1 h. The reaction mixture was removed from the heat, neutralized with 10 μL 2 M HCl and the solution was diluted with 820 μL MeOH to a total volume of 1 mL. The FDAA derivatives were analysed by UHPLC-DAD on a Dionex Ultimate 3000 RS DAD equipped with a Kinetex C18 column (2.6 μm, 150 × 2.10mm, Phenomenex). The analyses were run with a gradient elution of MeCN/H₂O–Milli-Q (Millipore, MA, USA) added 50 ppm TFA from 15 to 100% MeCN over 7 min (60 °C, 0.8 mL/min). The retention times of the FDAA derivatives were compared to retention times of the standard amino acid derivatives.
3.4. Selective Extraction of Sclerotium Metabolites

The selective extraction of sclerotia from IBT 15934, NRRL 13462, IBT 16807 and IBT 13353 was made from harvested sclerotia of a 7 day old cultivation on WATM and CYA agar (25 °C in dark). The sclerotia were washed several times with Milli-Q (Millipore, Millford, USA) 0.22 μm H₂O and dried. The sclerotia were transferred to a 2 mL Eppendorf tube together with three stainless steel balls (2 × 1 mm and 1 × 5 mm) and frozen with liquid N₂ before mechanical crushed. The pulverized sclerotia were suspended in 1mL methanol and transferred to a 2 mL vial with 1 mL of 1:2:3 methanol:dichloromethane:ethylacetate and left for evaporation overnight in a fume hood. The dried extract was resolved in 1 mL methanol in ultrasonicated for 10 min and then filtered with a 0.45 μm PTFE filter to a clean vial for analysis.

3.5. HPLC-DAD-TOF Method

Mass was measured using a Agilent HP 1100 liquid chromatograph with a DAD system (Waldbronn, Germany) on a LCT oaTOF mass spectrometer (Micromass, Manchester, UK) with a Z-spray ESI source and a LockSpray probe. For general procedures see [86], method 1.

3.6. NMR Instrumentation

NMR spectra were recorded on a Varian Unity Inova 500 MHz spectrometer equipped with a 5 mm probe using standard pulse sequences. The signals of the residual solvent protons and solvent carbons were used as internal references (δH 2.50 and δC 39.5 ppm for DMSO). In cases where higher field was needed the NMR spectra were recorded on a Bruker Avance 800-MHz spectrometer equipped with a 5-mm TCI Cryoprobe at the Danish Instrument Center for NMR Spectroscopy of Biological Macromolecules.

3.7. MS/MS Method Used for Aflatrem and Parasiticolide Screening

Liquid chromatography was performed on an Agilent (Torrence, CA, USA) 1100 HPLC system coupled to a triple-quadrupole mass spectrometer (Waters-Micromass, Manchester, UK) with a Z-spray ESI operated in positive mode source using a flow of 700 L/h nitrogen desolvated at 350 °C. The hexapole was held at 50 V. The system was controlled by MassLynx v4.1 (Waters-Micromass). Nitrogen was used as collision gas, and the MS operated in MRM mode (dwell time = 100 ms) with the parameters listed in Table 2. Extracts of 2 μL were injected and separated on a Phenomenex Gemini C18-phenyl, 3 μm, 2 × 50 mm column with a flow of 0.3 μL/min. Water contained 20 mM ammonium formate. Oven temperature 40 °C. Two different methods were applied to the aflatrem- and parasiticolide screen: Aflatrem inlet method: Linearly gradient from 50 to 100% MeCN over 5 min, increased to 0.5 μL/min over 1.5 min. The column was washed additionally 1.5 min with 100% MeCN at 0.5 μL/min, followed by a return to 50% MeCN over 2.5 min and kept at this level for another 1 min with a linearly decrease in flow to 0.3 μL/min, prior to the next sample. Standards used for analysis of this pathway were from Sigma-Aldrich Aldrich (St. Louis, MO, USA). Parasiticolides: linearly gradient from 20 to 90% MeCN over 15 min, increased to 0.5μL/min from 90 to 100% MeCN in additional 1 min. The column was washed from 2 min with 100% MeCN at 0.5 μL/min, followed by a return to
20% MeCN over 1.5 min and kept at this level for another 3.5 min with a linearly decrease in flow to 0.3 μL/min, before the next sample. Standards were internal standards from other extracts of the known parasiticolide A producer *A. parasiticus*: IBT 4387 (= CBS 260.67) and IBT 11863 (= CBS 115.37).

**Table 2.** MS/MS method including scan event, retention times, transition ions and the cone and collision energies used.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Scan event</th>
<th>RT (min)</th>
<th>Ion type</th>
<th>Transition (m/z)</th>
<th>Cone (V)</th>
<th>Collision energy (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paxilline</td>
<td>1</td>
<td>4.0</td>
<td>Quantifier</td>
<td>436 → 130</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>436 → 182</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.3</td>
<td>Quantifier</td>
<td>434 → 130</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>434 → 376</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>Paspalinine</td>
<td>3</td>
<td>4.8</td>
<td>Quantifier</td>
<td>420 → 182</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>13-desoxypaxilline</td>
<td>4</td>
<td>5.2</td>
<td>Quantifier</td>
<td>502 → 198</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>Aflatrem</td>
<td>5</td>
<td>5.5</td>
<td>Quantifier</td>
<td>422 → 130</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>422 → 275</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>Dideacetyl-parasiticolide A</td>
<td>1</td>
<td>7.0</td>
<td>Quantifier</td>
<td>387 → 217</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>387 → 189</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>14-deacetyl parasiticolide A</td>
<td>2</td>
<td>8.7</td>
<td>Quantifier</td>
<td>429 → 217</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>429 → 189</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>Parasiticolide A</td>
<td>3</td>
<td>10.4</td>
<td>Quantifier</td>
<td>488 → 229</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>488 → 247</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>

*a* All transitions were made from [M+H]+, except for parasiticolide A: [M+NH₄]+.

4. Conclusions

The tremorgenic 13-desoxypaxilline has been isolated from *A. oryzae* RIB40 and verified under several different growth conditions contrary to previous studies. We believe that 13-desoxypaxilline is the end-product of the aflatrem biosynthesis for the RIB40 strain since no aflatrem could be detected in any fermentation using LC-MS/MS as detection.

The new metabolites dide- and 14-deacetyl parasiticolide A were also found as genuine products from the RIB40 strain and the compounds were present in multiple fermentations, however parasiticolide A was only detected in trace amounts using a LC-MS/MS method. This indicates a defective acetylation of the 14-deacetyl parasiticolide A and the small amount of parasiticolide A in RIB40 could be the result of spontaneous acetylation in the cell cytosol. The mono-deacetylated analogue detected in both *A. flavus* and *A. oryzae* had same retention times, suggesting a selective acetylation.

The new NRPS compounds ditryptoleucine and oryzamides A₁-2 appear to be natural variants of known *A. flavus* metabolites. They share the exchange of a phenylalanine for a leucine, although they are believed to originate from two unrelated pathways.

Altogether, our findings contribute to understanding why the overall chemical profiles of *A. oryzae* (RIB40) and *A. flavus* (NRRL 3357) appear quite different since some of the end-products usually seen in *A. flavus* are apparently not reached in *A. oryzae*. Whether the different chemical profiles are merely
the result of different regulation that can be overcome by the use of epigenetic modifiers or are a result of genuine mutations remains to be settled. A. oryzae RIB40 is clearly a chemically potent strain, and as more of its chemistry is unfolded we believe that most of the biosynthetic pathways of A. flavus will be found to be more or less functional. Future research will reveal whether the many different A. oryzae strains that are used as industrial workhorses are as chemically potent as the RIB40 strain.

Supplementary Materials


Acknowledgments

This work was funded by the Danish Research Agency for Technology and Production (Grant: 09-064967). We thank the Danish Instrument Center for NMR Spectroscopy of Biological Macromolecules at the Carlsberg Laboratory for 800 MHz NMR time.

Conflict of Interest

The authors declare no conflict of interest.

References and Notes


© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).
Comparative chemistry of *Aspergillus oryzae* (RIB40) and *A. flavus* (NRRL 3357)

Christian Rank 1,†, Marie Louise Kleinstrup 1,†, Lene Maj Petersen 1, Sara Kildgaard 1, Jens Christian Frisvad 1, Charlotte Held Gotfredsen 2 and Thomas Ostenfeld Larsen 1,*

13-Desoxypaxilline

HRESIMS: m/z = 420.2551 [M+H]+, calculated for [C_{27}H_{33}NO_3+H]+: 420.2533.

NMR spectra were acquired in DMSO-d6 on a Bruker Avance 800 MHz spectrometer using standard pulse sequences. 1H-NMR (799.30 MHz, DMSO-d6, 25 °C, 2.50 ppm): 0.88 (3H, s, H-23), 1.00 (3H, s, H-25), 1.16 (3H, s, H-29), 1.20 (3H, s, H-28), 1.52 (1H, ddd, J = 25.4, 12.8, 4.4 Hz, H-14a), 1.60 (1H, m, H-14b), 1.65 (1H, m, H-15a), 1.74 (1H, d, J = 12.2 Hz, H-15b), 1.81 (1H, ddd, J = 17.9, 13.8, 4.2 Hz, H-6a), 1.98 (1H, ddd, J = 13.8, 13.6, 4.2 Hz, H-5a), 2.07 (1H, m, H-5b), 2.22 (1H, m, H-6b), 2.32 (1H, dd, J = 12.8, 11.0 Hz, H-17a), 2.53 (1H, m, H-13), 2.62 (1H, dd, J = 12.8, 6.3 Hz, H-17b), 2.71 (1H, m, H-16), 3.74 (1H, d, J = 1.6 Hz, H-9), 4.34 (1H, br. s, 27-OH), 4.41 (1H, m, H-7), 5.73 (1H, s, H-11), 6.91 (1H, dd, J = 7.6, 7.6 Hz, H-21), 6.95 (1H, dd, J = 7.6, 7.6 Hz, H-22), 7.27 (1H, d, J = 7.6 Hz, H-23), 7.28 (1H, d, J = 7.6 Hz, H-20), 10.76 (1H, s, N-H).

13C-NMR (201.00, DMSO-d6, 25 °C, 39.5 ppm): 14.4 (C-25), 15.4 (C-26), 23.5 (C-15), 24.8 (C-14), 25.5 (C-28), 25.7 (C-29), 26.7 (C-17), 29.5 (C-6), 30.7 (C-5), 41.6 (C-13), 48.5 (C-16), 49.2 (C-3), 49.8 (C-4), 70.7 (C-27), 74.0 (C-7), 82.4 (C-9), 111.6 (C-23), 115.8 (C-18), 117.5 (C-20), 118.1 (C-21), 119.2 (C-22), 120.7 (C-11), 124.4 (C-19), 140.2 (C-24), 150.4 (C-2), 168.6 (C-12), 196.1 (C-10).

14-Deacetyl parasiticolide A

All NMR spectra of parasiticolide A analogues were acquired in DMSO-d6 on a Varian Unity Inova 500 MHz spectrometer with 4 mm gHX Nano probe and with a spin rate of 2 kHz for all samples, using standard pulse sequences. The spectra were referenced to this solvent with resonances δ_H = 2.50 and δ_C = 39.5.

HRESIMS: m/z = 429.1901 [M+H]+, calculated for [C_{26}H_{38}O_5+H]+: 429.1908.
$^1$H-NMR (499.87 MHz, DMSO-$d_6$, 25 °C, 2.50 ppm): 0.96 (1H, td, $J = 13.6, 3.7$ Hz, H-3a), 1.01 (3H, s, H-13), 1.36 (1H, td, $J = 13.3, 3.3$ Hz, H-1a), 1.45 (1H, m, H-2a), 1.63 (1H, m, H-2b), 1.91 (1H, s, H-5), 1.97 (1H, m, H-3b), 2.11 (1H, d, $J = 13.2, 1.2$ Hz, H-1b), 2.36 (1H, d, $J = 19.0$ Hz, H-7a), 2.76 (1H, ddt, $J = 19.0, 6.0, 3.1$ Hz, H-7b), 3.14 (1H, d, $J = 10.6$ Hz, H-14a), 3.61 (1H, d, $J = 10.6$ Hz, H-14b), 4.67 (1H, d, $J = 11.0$ Hz, H-15a), 4.80 (1H, d, $J = 11.0$ Hz, H-15b), 4.96 (2H, m, H-11), 5.83 (1H, d, $J = 6.0$ Hz, H-6), 7.52 (2H, m, H-4′/6′), 7.67 (1H, tt, $J = 7.3, 1.2$ Hz, H-5′), 7.97 (2H, m, H-3′/7′).

$^{13}$C-NMR (125.70 MHz, DMSO-$d_6$, 25 °C, 39.5 ppm): 17.3 (C-2), 20.6 (C-17), 28.2 (C-7), 30.7 (C-1), 34.6 (C-3), 39.4 (C-4), 39.8 (C-10), 52.5 (C-5), 62.9 (C-14), 65.4 (C-15), 66.5 (C-6), 71.1 (C-11), 121.2 (C-8), 128.7 (4′/6′), 129.3 (3′/7′), 129.8 (C-2′), 133.5 (5′), 165.6 (C-1′), 166.8 (C-9), 170.5 (C-16), 173.2 (C-12).

Dideacetyl parasiticolide A

HRESIMS: $m/z = 387.1817$ [M+H]$^+$, calculated for [C$_{22}$H$_{26}$O$_6$+H]$^+$: 387.1802.

$^1$H-NMR (499.87 MHz, DMSO-$d_6$, 25 °C, 2.50 ppm): 0.91 (1H, ddd, $J = 13.4, 13.4, 3.7$ Hz, H-3a), 0.99 (3H, s, H-13), 1.17 (1H, ddd, $J = 12.9, 3.5$ Hz, H-1a), 1.42 (1H, m, H-2a), 1.65 (1H, m, H-2b), 1.81 (1H, s, H-5), 1.97 (1H, d, $J = 13.4$ Hz, H-3b), 2.28 (1H, d, $J = 12.9$ Hz, H-1b), 2.31 (1H, d, $J = 19.3$ Hz, H-7a), 2.73 (1H, m, H-7b), 3.09 (1H, dd, $J = 10.5, 5.3$ Hz), 3.60 (1H, dd, $J = 10.5, 5.3$ Hz, H-14b), 4.07 (1H, m, H-15a), 4.12 (1H, m, H-15b), 4.42 (1H, dd, $J = 5.3$, 5.3 Hz, 14-OH), 4.86 (1H, dd, $J = 17.5$, 2.0 Hz, H-11a), 5.06 (1H, dt, $J = 17.5, 2.5$ Hz, H-11b), 5.10 (1H, dd, $J = 5.2$, 5.2 Hz, 15-OH), 5.80 (1H, d, $J = 5.8$ Hz, H-6), 7.55 (2H, m, H-4′/6′), 7.67 (1H, m, H-5′), 7.90 (2H, m, H-3′/7′).

$^{13}$C-NMR (125.70 MHz, DMSO-$d_6$, 25 °C, 39.5 ppm): 17.6 (C-2), 27.3 (C-13), 28.5 (C-7), 30.5 (C-1), 34.9 (C-3), 39.4 (C-4), 42.9 (C-10), 52.7 (C-5), 62.4 (C-15), 63.1 (C-14), 66.9 (C-6), 72.0 (C-11), 119.5 (C-8), 128.9 (4′/6′), 129.0 (C-3′/7′), 129.8 (C-2′), 133.5 (5′), 165.7 (C-1′), 169.4 (C-9), 173.6 (C-12).
18-Formyl parasiticolide A

HRESIMS: m/z = 457.1889 [M+H]+, calculated for [C_{25}H_{28}O_{8}+H]^+: 457.1857.

$^1$H-NMR (499.87 MHz, DMSO-$d_6$, 25 °C, 2.50 ppm): 1.10 (3H, s, H-13), 1.18 (1H, ddd, $J = 13.7, 13.7, 3.3$ Hz, H-3a), 1.43 (1H, ddd, $J = 13.2, 13.2, 2.9$ Hz, H-1a), 1.51 (1H, m, H-2a), 1.71 (1H, m, H-2b), 1.82 (1H, d, $J = 13.7$ Hz, H-3a), 2.06 (1H, s, H-5), 2.09 (3H, s, H-17), 2.16 (1H, d, $J = 13.2$ Hz, H-1b), 2.43 (1H, d, $J = 19.3$ Hz, H-7a), 2.80 (1H, m, H-7b), 3.99 (1H, d, $J = 11.0$ Hz, H-14a), 4.36 (1H, d, $J = 11.0$ Hz, H-14b), 4.71 (1H, d, $J = 10.8$ Hz, H-15a), 4.78 (1H, d, $J = 10.8$ Hz, H-15b), 4.98 (2H, m, H-11), 5.83 (1H, d, $J = 5.6$ Hz, H-6), 7.52 (2H, dd, $J = 7.6, 7.6$ Hz, H-3'/7'), 7.68 (1H, dd, $J = 7.6, 7.6$ Hz, H-4''), 8.06 (1H, s, H-18).

$^{13}$C-NMR (125.70 MHz, DMSO-$d_6$, 25 °C, 39.5 ppm): 16.9 (C-2), 20.5 (C-17), 26.6 (C-13), 27.9 (C-7), 30.5 (C-1), 35.4 (C-3), 36.9 (C-4), 39.6 (C-10), 52.1 (C-5), 65.0 (C-14), 65.1 (C-15), 66.5 (C-6), 71.1 (C-11), 121.0 (C-8), 128.6 (C-4'/6'), 129.0 (C-3'/7'), 129.2 (C-2'), 133.4 (C-5'), 161.6 (C-18), 165.1 (C-1'), 165.9 (C-9), 170.1 (C-16), 172.8 (C-12).
**Metabolites 2012, 2**

**Ditryptoleucine A**

HRESIMS: *m/z* = 625.3472 [M+H]^+, calculated for [C_{36}H_{44}N_6O_4+H]^+: 625.3497.

$^1$H-NMR (499.87 MHz, DMSO-d$_6$, 25 °C, 2.50 ppm): 0.81 (6H, d, *J* = 6.4 Hz, H-21), 0.83 (6H, d, *J* = 6.4 Hz, H-20), 1.42 (H2, m, H-18a), 1.48 (H2, m, H-18b), 1.58 (H2, m, H-19), 2.44 (H4, m, H-11), 2.80 (H6, s, H-17), 3.76 (H2, dd, *J* = 10.4, 3.8 Hz, H-12), 3.81 (H2, dd, *J* = 7.9, 6.0 Hz, H-15), 5.09 (H2, m, H-2), 6.61 (H2, d, *J* = 7.8, H-5), 6.66 (H2, t, *J* = 7.5, H-7), 6.68 (H2, s, H-3), 7.05 (H2, t, 7.8, H-6), 7.23 (H2, d, *J* = 7.8, H-8).

$^{13}$C-NMR (125.70 MHz, DMSO-d$_6$, 25 °C, 39.5 ppm): 21.9 (C-20), 22.6 (C-21), 23.7 (C-19), 31.8 (C-17), 36.1 (C-11), 39.4 (C-18), 56.7 (C-12), 57.9 (C-10), 61.6 (C-15), 76.7 (C-2), 108.7 (C-5), 117.2 (C-7), 124.6 (C-8), 126.7 (C-9), 129.0 (C-6), 150.9 (C-4), 165.0 (C-16), 166.1 (C-13).

**Ditryptoleucine B**

$^1$H-NMR (499.87 MHz, DMSO-d$_6$, 25 °C, 2.50 ppm): 0.81 (6H, d, *J* = 6.1 Hz, H-21), 0.83 (6H, d, *J* = 6.2 Hz, H-20), 1.42 (H2, m, H-18), 1.49 (H2, m, H-19), 2.45 (H2, dd, *J* = 13.5, 4.7, H-11a), 2.75 (H6, s, H-17), 3.13 (H2, m, H-11b), 3.33 (H2, s, H-2), 3.82 (H2, t, *J* = 6.9 Hz, H-15), 4.23 (H2, m, H-12), 6.58 (H2, d, *J* = 7.5, H-5), 6.65 (H2, t, *J* = 7.1, H-7), 6.70 (H2, s, H-3), 7.01 (H2, t, 7.2, H-6), 7.37 (H2, d, *J* = 7.1, H-8).

$^{13}$C-NMR (125.70 MHz, DMSO-d$_6$, 25 °C, 39.5 ppm): 21.5 (C-21), 22.6 (C-20), 23.6 (C-19), 31.8 (C-17), 37.0 (C-11), 38.0 (C-18), 55.3 (C-12), 60.8 (C-10), 61.7 (C-15), 76.3 (C-2), 108.5 (C-5), 117.4 (C-7), 124.0 (C-8), 130.2 (C-9), 128.3 (C-6), 149.6 (C-4) 166.9 (C-16), 166.6 (C-13).

**Oryzamide A**

HRESIMS: *m/z* = 491.26526 [M+H]^+, calculated for [C_{28}H_{34}N_4O_4+H]^+: 491.26525.

$^1$H-NMR (499.87 MHz, DMSO-d$_6$, 25 °C, 2.50 ppm): 0.83 (3H, d, *J* = 6.5 Hz, H-5'), 0.88 (3H, d, *J* = 6.6 J, H-5'), 1.36 (1H, m, H-4'), 1.59 (2H, m, H-3'), 1.77 (3H, s, H-2''), 2.67 (1H, m, H-3''), 2.83 (1H, m, H-3''), 2.87 (3H, s, N-CH), 4.76 (1H, m, H-2''), 5.02 (1H, m, H-2'), 6.47 (1H, d, *J* = 14.8, H-2), 6.61 (2H, d, *J* = 8.4 Hz, H-6''/8''), 7.01 (2H, d, *J* = 8.4 Hz, H-5''/9''), 7.05 (1H, m, H-8), 7.10
(1H, m, H-7), 7.26 (1H, m, H-1), 7.36 (1H, m, H-6), 7.41 (1H, s, H-4), 7.62 (1H, d, J = 7.9 Hz, H-9), 8.24 (1H, d, J = 8.0 Hz, 2″-NH), 9.16 (1H, s, 7″-OH), 9.84 (1H, d, J = 9.8 Hz, 1-NH), 11.11 (1H, s, 4-NH).

$^{13}$C-NMR (125.70 MHz, DMSO-$d_6$, 25 °C, 39.5 ppm): 21.6 (C-5′), 22.0 (C-2″′), 22.7 (C-5′′), 24.1 (C-4′), 30.9 (C-N), 36.2 (C-3″), 36.9 (C-3′), 50.5 (C-2″), 54.2 (C-2′), 106.4 (C-2), 111.4 (C-3), 111.5 (C-6), 114.6 (C-6″/8″), 118.5 (C-9), 118.8 (C-8), 119.2 (C-1), 121.1 (C-7), 123.1 (C-4), 125.0 (C-10), 127.4 (C-4″), 129.7 (C-5″/9″), 136.9 (C-5), 156.0 (C-7″), 167.7 (C-1′″), 168.8 (C-1″′), 172.3 (C-1″).
Paper 3

“Isolation and NMR characterization of fumonisin B$_2$ and a new fumonisin B$_6$ from *Aspergillus niger*”

M Månsson, ML Klejnstrup, RK Phipps, KF Nielsen, JC Frisvad, CH Gotfredsen, and TO Larsen

Accepted in *Journal of Agricultural and Food Chemistry*, 58, 949-953 (2010)
Isolation and NMR Characterization of Fumonisin B₂ and a New Fumonisin B₆ from Aspergillus niger

MARIO MANSSON,† MARIE LOUISE KLEINSTRUP,† RICHARD K. PIPLE,† KRISTIAN F. NIELSEN,‡ JENS C. FRESVAD,‡ CHARLOTTE H. GOTFREDSEN,‡ and THOMAS O. LARSEN*†‡

†Department of Systems Biology, Center for Microbial Biotechnology, Technical University of Denmark, Building 221, DK-2800 Kgs. Lyngby, Denmark and ‡Department of Chemistry, Kemiløvervej, Technical University of Denmark, Building 201, DK 2800 Kgs. Lyngby, Denmark

A new fumonisin, fumonisin B₆ (1), has been isolated by cation-exchange and reverse-phase chromatography, together with fumonisin B₂ (2), from stationary cultures of the fungus Aspergillus niger NRRL 326. Analysis of mass spectrometric and NMR data determined that FB₆ is a positional isomer of FB₁ and iso-FB₁, having hydroxyl functions at C3, C4, and C5. Analysis of the NMR data for FB₆ showed very similar chemical shift values when compared to an authentic Fusarium FB₂ standard, strongly indicating identical molecules despite that an absolute stereochemical assignment of FB₆ from A. niger was not possible.

KEYWORDS: Fumonisins; Aspergillus niger; mycotoxins; food safety; NMR

INTRODUCTION

Recently, our group was the first to report the production of the mycotoxin fumonisin B₁ (FB₁) from Aspergillus niger NRRL 3122, using HPLC–MS (1). The finding did not come as a great surprise, since others (2, 3) had already reported a fumonisin-like gene cluster in the A. niger genome of two different strains of the species. Later, we also reported FB₁ from A. niger, generally produced in amounts around 10–25% of the FB₂ (4, 5). Until these findings, fumonisin production has only been previously reported in certain species within the Gibberella fujikuroi and Fusarium oxysporum species complexes, which are common fungal contaminants of maize-based foods and feeds (6–9).

The first fumonisins were described by Gelderblom et al. (9) from cultures of F. verticilloides. Since then, more than 25 fumonisins have been identified by NMR (6) and a further 25 have been putatively identified by LC–MS² (10). The most abundant and most toxic fumonisins are the B-series analogues (11–13), which contain a terminal 2-amino-3-hydroxy motif on an eicosane backbone and two hydroxy groups esterified with tricarboxylic acids (TCA).

In most Fusaria, FB₁ is predominant, usually representing some 70% of the total fumonisin content. FB₂ and FB₆ usually account for up to 15–25% and 3–8%, respectively, while FB₃ is normally present in insignificant amounts (4–16).

The recent documentation of fumonisin B-series production by A. niger has raised the question of whether fumonisins might be an overlooked health risk due to the ubiquitous presence of A. niger on a wide range of food stuffs not usually contaminated with Fusarium species (17). This includes grapes and thereby raisins and wine, peanuts, coffee, tea, and several other products (1). Moreover, one could be concerned that industrial products such as enzymes and fine chemicals could be contaminated with traces of fumonisins because of the wide application of A. niger as “workhorse” in industrial fermentations (18–21).

Our recent LC–HRMS based investigations of fumonisin production by A. niger and other black Aspergilli suggested that besides FB₂ and FB₁, A. niger can also produce an additional third fumonisin with the same elementel composition as FB₁ and iso-FB₁, but with a different retention time than the two latter compounds. The present study reports the isolation and characterization of this novel compound, which we named fumonisin B₆. In addition, we isolated and characterized FB₂ from A. niger by NMR in order to further validate that FB₂ from A. niger is indeed identical to that of FB₂ previously reported from several Fusarium species.

MATERIALS AND METHODS

Fungal Incubation. In a 10 L glass vessel, 2 L of still culture of 55% rice meal and 45% corn steep liquor (RC) medium was prepared according to Bullermann (22) but with a reduced amount of agar (1.5 g/L). It was inoculated with a spore suspension from 7 d old ex-type cultures of A. niger NRRL 326, ex-tannin-gallic acid fermentation (=IBT 27876 = ATCC 16888 = CBS 554.63) from CVA. After incubation for 14 d in darkness at 25 °C, the mycelia and media were homogenized with 2 L of MeCN/H₂O (50:50, v/v) and shaken overnight. After filtration the extract was concentrated on a rotary evaporator and freeze-dried.

Preparative Chromatography. The raw extract, 9.3 g, was initially subjected to flash chromatography on a 155 g, 50 μm C₂₂ material (Phenomenex, Torrance, CA) which was eluted with H₂O–MeOH mixtures in steps of 10%, starting from 10% MeOH. The 70% MeOH fraction (163 mg) was further enriched for fumonisins on a 1000 mg, 33 μm Strata X-C mixed-mode RP–cation-exchange material (Phenomenex), which was washed with MeCN and eluted with MeCN–H₂O (containing 4% NH₄OH) in steps of 20% from 4% NH₄OH. The fumonisin-containing fraction (20% MeCN, 19 mg), was finally subjected...
to RP separation on a 250 mm x 10 mm i.d., 5 μm Luna C8 (II) column (Phenomenex) using an MeCN–H2O gradient (starting at 30% MeCN + 50 ppm TFA increasing to 70% over 15 min). Fractions were collected automatically every 20 s and pooled for each chromatographic run after analysis of the content of each fraction by LC–MS had shown which fractions contained fumonisin B1 (1.8 mg) and fumonisin B2 (4.2 mg).

Analytical LC–MS. LC–DAD-HRMS was performed as described by Frisvad et al. (1) or by LC–tandem MS (MS/MS) on an Agilent 1100 liquid chromatograph (Waldbronn, Germany) coupled to a Quattro Ultima triple mass spectrometer (Micromass, Manchester, U.K.) with an ESI source. The separation was performed on a 50 mm x 2 mm i.d., 3 μm Gemini C18 (Phenomenex) column fitted with a security guard system using a linear gradient starting from 20% MeCN in H2O (both with 20 mM formic acid) to 55% MeCN for 6 min at a flow rate of 300 μL min−1. MS/MS was performed using nitrogen as collision gas and making daughter ion scans of the protonated molecular ions (m/z 722 and 796), using fragmentation potentials from 15 to 50 V. Solvents were HPLC grade, and all other chemicals were analytical grade unless otherwise stated. Fumonisin standards (FB1 and FB2, mixture 50 μg mL−1) used for MS analysis were acquired from Biopure (Tulln, Austria). FB2 was from Chiron (Trondheim, Norway). FB1 used for NMR analysis was from Alexis Biochemicals (Lausen, Switzerland).

**NMR Spectroscopy.** NMR spectra were recorded on a Bruker Avance 800 MHz spectrometer with a 5 mm TCI cryoprobe at the Danish Instrument Center for NMR Spectroscopy of Biological Macromolecules. The 2D DQF-COSY, NOESY, gHSQC, gH2BC, and gHMBC spectra were acquired using standard pulse sequences. The NMR data used for the structural assignment of the fumonisins were acquired in CDCl3 (δH 3.31 ppm and δC 49.15 ppm).

**RESULTS AND DISCUSSION**

In the current study, the ex-type culture of A. niger was used, as it has been found to be a strong producer strain of fumonisins (1). With no published protocols for purification of fumonisins from A. niger, we chose to use mixed-mode RP cation-exchange for purification. This was done because A. niger produces high amounts of acidic naphtho-γ-pyrones and organic acids, which would likely also bind to a strong anion-exchanger (SAX), which is the usual way to purify fumonisins (23).

Analytical LC–MS was used to compare the retention times and daughter-ion spectra of the purified A. niger fumonisin analogues to those of authentic standards of FB1, FB2, and FB3. The daughter-ion spectra of the FB2 standard and FB1 from A. niger, eluting at the same time, were identical and consistent with previously reported ESI-MS/MS data (24–26), with losses of H2O and TCA groups from the alkyl backbone dominating the spectra. This confirmed the presence of FB2 or an isomer thereof in the A. niger extract. LC–HRMS analyses also showed a peak with the same elemental composition as FB2, however, eluting slightly later (Figure 1). LC–MS/MS also confirmed this and showed that it eluted 0.6 min after the authentic standard of FB2 and 0.15 min before FB2 on the C60 phenyl column. Significant differences in the daughter-ion spectra clearly showed that A. niger produces an FB3 analogue and not FB1 (Figure 1). The same was observed in the crude A. niger extract; thus, the FB2 analogue was not an artifact of the purification process but proved to be the novel analogue fumonisin B3 (FB3). FB3 displayed significantly less abundant losses of H2O than FB2, which indicates the presence of adjacent hydroxyl groups blocking the hydrogens from detaching, thereby allowing the H2O loss. This is in agreement with our previous findings during our studies of the fragmentation patterns of 474 microbial metabolites using LC–MS ESI+ and ESI− MS (27), where we have never observed H2O losses unless a β-hydrogen could be lost due to double bond formation.

**NMR Based Structural Investigations.** To further validate that the FB2 purified from A. niger is identical to that produced by Fusarium, 1D and 2D homo- and heteronuclear NMR spectra of the purified A. niger fumonisin FB2 were carefully inspected and compared to those of an authentic standard of FB2. The NMR data obtained in methanol-d4 displayed comparable chemical shifts for the 34 expected 13C NMR shift values except for the resonances originating from the two TCA side chains and their sites of attachment to the fumonisin backbone. In this region, deviations in the 13C NMR chemical shifts of between −1.8 to +0.3 ppm were observed. These results could be explained by a difference in the structural fold of the FB2 between the two samples, likely to be caused by differences in pH and thereby overall level of protonation of the four carboxylic acid functionalities. The fact that FB2 does have a unique globular
Table 1. NMR Data for Fumonisin $B_2$, and Iso-$B_2$

<table>
<thead>
<tr>
<th></th>
<th>fumonisin $B_2$</th>
<th>fumonisin $B_1$</th>
<th>fumonisin iso-$B_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>H (ppm)</td>
<td>C (ppm)</td>
<td>HMBCC</td>
</tr>
<tr>
<td>1</td>
<td>1.34</td>
<td>16.3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3.67</td>
<td>50.0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3.64</td>
<td>71.8</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3.46</td>
<td>74.9</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>3.71</td>
<td>71.4</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>1.53</td>
<td>34.8</td>
<td>1</td>
</tr>
<tr>
<td>7 &amp; 8</td>
<td>1.26</td>
<td>30.6</td>
<td>4.8</td>
</tr>
<tr>
<td>9 &amp; 10</td>
<td>1.26</td>
<td>30.6</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>1.04</td>
<td>3.61</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>1.50</td>
<td>29.9</td>
<td>15</td>
</tr>
<tr>
<td>13</td>
<td>1.43</td>
<td>36.1</td>
<td>14, 16, 21</td>
</tr>
<tr>
<td>14</td>
<td>1.62</td>
<td>1.70</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>5.18</td>
<td>1.56</td>
<td>72.6</td>
</tr>
<tr>
<td>16</td>
<td>4.92</td>
<td>78.7</td>
<td>13, 14, 16, 17, 22</td>
</tr>
<tr>
<td>17</td>
<td>1.70</td>
<td>34.7</td>
<td>15</td>
</tr>
<tr>
<td>18</td>
<td>1.09</td>
<td>32.9</td>
<td>16, 22</td>
</tr>
<tr>
<td>19</td>
<td>1.44</td>
<td>1.44</td>
<td>1.43</td>
</tr>
<tr>
<td>20</td>
<td>1.31</td>
<td>29.7</td>
<td>17, 20</td>
</tr>
<tr>
<td>21</td>
<td>1.26</td>
<td>1.40</td>
<td>1.38</td>
</tr>
<tr>
<td>22</td>
<td>0.89</td>
<td>14.2</td>
<td>18, 19</td>
</tr>
<tr>
<td>23</td>
<td>0.92</td>
<td>20.8</td>
<td>12, 13</td>
</tr>
<tr>
<td>24</td>
<td>0.94</td>
<td>15.7</td>
<td>14, 16</td>
</tr>
<tr>
<td>25</td>
<td>172.9</td>
<td>173.5</td>
<td>173.5</td>
</tr>
<tr>
<td>26</td>
<td>2.51</td>
<td>24.5</td>
<td>37.1</td>
</tr>
<tr>
<td>27</td>
<td>2.72</td>
<td>2.71</td>
<td>2.71</td>
</tr>
<tr>
<td>29</td>
<td>2.68</td>
<td>2.50</td>
<td>3.78</td>
</tr>
<tr>
<td>30</td>
<td>1754</td>
<td>1774</td>
<td>1772</td>
</tr>
<tr>
<td>31</td>
<td>1729</td>
<td>1785</td>
<td>1787</td>
</tr>
<tr>
<td>32</td>
<td>1729</td>
<td>1730</td>
<td>1732</td>
</tr>
<tr>
<td>33</td>
<td>2.84</td>
<td>2.60</td>
<td>3.69</td>
</tr>
<tr>
<td>34</td>
<td>2.80</td>
<td>2.79</td>
<td>2.79</td>
</tr>
<tr>
<td>36</td>
<td>3.58</td>
<td>3.52</td>
<td>3.75</td>
</tr>
<tr>
<td>37</td>
<td>2.74</td>
<td>2.75</td>
<td>2.75</td>
</tr>
<tr>
<td>38</td>
<td>175</td>
<td>177.0</td>
<td>176.7</td>
</tr>
<tr>
<td>39</td>
<td>176.8</td>
<td>178.0</td>
<td>178.2</td>
</tr>
</tbody>
</table>

*Only approximate J coupling constants (in Hz) due to overlap in the region. Ambiguous assignment due to total overlap in the region. HMBCC and HBCB connectivities are from the assigned proton to the indicated carbon(s).*

folded structure has previously been addressed by Beier and Stanker (29). In order to ensure that the two samples had been subjected to the same purification protocol, the standard FB$_2$ sample was run over a cation exchange column, in the same way as the last purification step for the isolated A. niger FB$_2$ sample. This proved to be very important, since the spectra obtained for the final stereochemistry of fumonisins produced by A. niger and Fusarium, showed that they are of the same size, proving that only minor changes in the stereochemistry at positions C12 and C16 could not be elucidated from the assigned proton to the indicated carbon(s).
which is why we conclude that further structural studies such as X-ray analysis need to be performed to establish the absolute stereochemistry of FB$_6$ from *A. niger*.

NMR studies of FB$_1$, substantiated the MS analyses, indicating a new fumonisin of the B type. Unambiguous assignments of the proton and carbon resonances of FB$_1$ could be obtained for most of the resonances. However, spectral overlap hampered some specific assignments in the alkane-like part of the molecule. The assignment was compared to literature values for fumonisins (16, 29). A series of 2D DQF-COSY, HSQC, HMBC, and H2BC experiments established partial assignment of the backbone as well as the TCA side chains present in the structure. Thirty-four carbon resonances could be identified comprising four CH$_3$ groups, 14 CH$_2$ groups, 10 CH groups, and 6 carboxyl carbon resonances (Table 1). The two TCA side chains bound to the backbone structure could be assigned as two separate spin systems C23–C28 and C29–C34 (Figure 2) with only minor differences in their chemical shifts. The attachment of the C29–C34 TCA side chain to the backbone structure could be confirmed by the presence of an HMBC correlation between H15 and C29, whereas the position of the C23–C28 TCA side chain is based on the chemical shift of H14/C14 and comparison with the assignment obtained for FB$_1$ and iso-FB$_1$. Unambiguous assignments of the fragments C1–C6 and C10–C20 were obtained. In Table 1, a comparison of the chemical shifts of FB$_1$, iso-FB$_1$, and FB$_2$ can be seen. Comparing the resonance assignment of FB$_1$ to FB$_1$ and iso-FB$_1$, showed only minor deviations between the structures except at the positions of OH attachment in the backbone structure. Where FB$_1$ and iso-FB$_1$ have OH groups attached to C3, C5, and C10 and to C3, C4, and C10, respectively, the new fumonisin FB$_6$ has OH groups at C3, C4, and C5 (Figure 2). The chemical shift values of 1.25 ppm for H10 and 27.5 ppm for C10, DQF-COSY connectivities from H11 to H10 and HMBC correlations from H9 to C10 and from H10 to C9, confirm that no OH group is bound at this position due to the large upfield shift of the resonances when compared to FB$_1$, where H10 and C10 resonate at 3.62 and 69.8 ppm, respectively. In the amino terminal portion of the backbone structure, DQF-COSY connectivities can be traced along the chain from H3 to 3.64 ppm, H4 at 3.46 ppm, and H5 at 3.71 ppm to H6 at 1.53/1.59 ppm. In addition, HMBC correlations can be seen from H4 to C2; C5 and from H3 to C4, confirming the presence of the third OH group at C4 in the structure. The chemical shift obtained for H4/C4 of 3.46 ppm/74.9 ppm compared to 3.65 ppm/72.2 ppm as seen in iso-FB$_1$ and equivalently for H5/C5 of 3.71 ppm/71.4 ppm compared to 3.84 ppm/68.4 ppm for FB$_6$, is in agreement with the proposed structure for FB$_6$.

Altogether, the analysis of these data proves FB$_6$, with hydroxyl functionalities at C3, C4, and C5, to be a structural isomer of FB$_1$ and iso-FB$_1$. As can be seen in Figure 2 FB$_1$, and iso-FB$_1$ have a hydroxyl at C4, which is rare for the B-series fumonisins and only previously seen for iso-FB$_{29}$. As is the case for FB$_2$, the absolute stereochemistry of FB$_6$ still needs to be resolved, which is work that is ongoing in our laboratory. Analysis of several *A. niger* strains in our laboratory (results not shown) has shown that the frequencies of production of fumonisins in *A. niger* on average are 100% (FB$_1$), 10–25% (FB$_2$), and 5–10% (FB$_6$). Testing of the toxicological properties of not only FB$_1$ and FB$_2$, but also FB$_6$, including investigations of the different conformations that we have seen for these *A. niger* fumonisins, is also in progress in our laboratory.

**ACKNOWLEDGMENT**

We thank the Danish Instrument Center for NMR Spectroscopy of Biological Macromolecules for NMR time.

**LITERATURE CITED**


(22) Bullermann, L. B. A screening medium and method to detect several mycotoxins in mold cultures. **J. Milk Food Technol.** 1974, 37, 1–3.


Paper 4

“A genome-wide polyketide synthase deletion library uncovers novel genetic links to polyketides and meroterpenoids in _Aspergillus nidulans_”

ML Nielsen, JB Nielsen, C Rank, ML Klejnstrup, DK Holm, KH Brogaard, BG Hansen, JC Frisvad, TO Larsen, and UH Mortensen

Accepted in Journal of FEMS Microbiology Letters, 321, 157-166 (2011)
RESEARCH LETTER

A genome-wide polyketide synthase deletion library uncovers novel genetic links to polyketides and meroterpenoids in Aspergillus nidulans


Department of Systems Biology, Center for Microbial Biotechnology, Technical University of Denmark, Lyngby, Denmark

Correspondence: Uffe H. Mortensen, Department of Systems Biology, Center for Microbial Biotechnology, Technical University of Denmark, Building 223, 2800 Kgs. Lyngby, Denmark. Tel.: +45 4525 2701; fax: +45 4588 4148; e-mail: um@bio.dtu.dk

Received 14 April 2011; revised 26 May 2011; accepted 27 May 2011. Final version published online 27 June 2011.

DOI:10.1111/j.1574-6968.2011.02327.x

Editor: Richard Staples

Abstract

Fungi possess an advanced secondary metabolism that is regulated and coordinated in a complex manner depending on environmental challenges. To understand this complexity, a holistic approach is necessary. We initiated such an analysis in the important model fungus Aspergillus nidulans by systematically deleting all 32 individual genes encoding polyketide synthases. Wild-type and all mutant strains were challenged on different complex media to provoke induction of the secondary metabolism. Screening of the mutant library revealed direct genetic links to two austinol meroterpenoids and expanded the current understanding of the biosynthetic pathways leading to arugosins and violaceols. We expect that the library will be an important resource towards a systemic understanding of polyketide production in A. nidulans.

Introduction

It is well known that filamentous fungi produce a large number of bioactive secondary metabolites (Berdy, 2005; Cox, 2007; Newman & Cragg, 2007). Polyketide (PK) compounds constitute a major part of these metabolites and have long been recognized as a valuable source of diverse natural compounds of medical importance, for example lovastatin (cholesterol lowering) (Lai et al., 2005), griseofulvin (antibiotic) (Cho et al., 2010) and mycophenolic acid (immunosuppressant) (Bentley, 2000). However, polyketides also include many toxic compounds that pose a serious threat to human health, for example patulin, ochratoxins, fumonisins and aflatoxin (Frisvad et al., 2004; Månsson et al., 2010). Polyketides are biosynthesized by large multidomain polyketide synthases (PKSs), which besides acyl transferase, β-ketoacyl synthase and acyl carrier domains may also contain keto reductase, dehydratase, cyclization and methyltransferase domains (Cox, 2007; Smith & Tsai, 2007; Hertweck, 2009). In fungi, the different catalytic activities often work in an iterative manner (fungal type I) and it is generally difficult to predict the exact product formed by a given PKS from its sequence alone (Keller et al., 2005). Product prediction is further complicated by the fact that the resulting polyketide structure may be decorated by tailoring enzymes. Such genes are often physically associated with the PKS gene in a gene cluster allowing for coordinated regulation (Schumann & Hertweck, 2006). The fact that natural products may be of mixed biosynthetic origin, combining elements such as polyketides with terpenes (meroterpenoids) and/or nonribosomal peptide units, adds to the complexity (Chang et al., 2009; Geris & Simpson, 2009; Hertweck, 2009; Scherlach et al., 2010).

As a consequence of their bioactivity, societal importance and also the prospect of reprogramming the biosynthetic machinery for drug development (Cox, 2007), there is tremendous interest in the discovery and understanding of fungal polyketide biosynthesis. The availability of full genome sequences of a number of filamentous fungi has provided a means to address the discovery of polyketides because the PKS genes are large and contain several conserved protein domains. Importantly, analysis of the...
genomic sequences from filamentous fungi (including *Aspergillus nidulans*, teleomorph, *Emergeria nidulans*) predict numbers of individual PKS genes that exceeds significantly the number of polyketides that these fungi are known to produce (Galagan et al., 2005). In fact, the genome of *A. nidulans* (Galagan et al., 2005) appears to contain as many as 32 individual PKS genes (Nierman et al., 2005; Szweczyk et al., 2008; von Döhren, 2009), but until now only nine genes have been linked to eight polyketides (Yamazaki & Maebayas, 1982; Bergmann et al., 2007; Chiang et al., 2008; Szweczyk et al., 2008; Bok et al., 2009; Chiang et al., 2009; Schroebb et al., 2009) (see Supporting Information, Fig. S1). This may reflect that many polyketides are either produced in small amounts, under special conditions, or in developmental stages that are rarely observed under laboratory conditions.

Towards a more complete genetic mapping of the secondary metabolism in *A. nidulans*, we first cultivated a reference strain on an array of different growth media to uncover polyketides that were not previously linked to a gene cluster. This analysis revealed several compounds, including austinols, violacols, arugosins and prenylated xanthones. Next, genetic links to these compounds were established by constructing and screening an *A. nidulans* mutant library containing individual deletions of 32 putative PKS genes.

**Materials and methods**

**Strains**

The *A. nidulans* strain IBT29539 (argB2, pyrG89, veA1) (Nielsen et al., 2008) was used as the reference strain and for deletion-strain constructions. *Escherichia coli* strain DH5α was used for cloning.

**Media**

Fungal minimal medium (MM) was as described in Cove (1966), but with 1% glucose, 10 mM NaNO₃, and 2% agar. Medium for *aA* promoter induction consisted of MM supplemented with 100 mM L-threonine and 100 mM glycerol as carbon source instead of 1% glucose. Polyketide screening media variants CYA, CYAs, YES and RT were prepared as per Frisvad & Samson (2004). CY20 medium consisted of CYA with 170 g sucrose; RTO contained RT with 30 g organic oat meal; and YE was made as YES but without sucrose. All media variants were supplemented with 10 mM uridine, 10 mM uracil and/or 4 mM L-arginine where appropriate.

**Construction of *A. nidulans* gene deletion library**

Individual PKS gene deletions were carried out as described previously (Nielsen et al., 2006), except that the targeting fragments were assembled using Gateway technology (Hartley et al., 2000) (see Table S1 for PCR oligonucleotide and Fig. S2 for an overview of the procedure). The *A. nidulans* transformants were streak purified and rigorously screened through three complementing diagnostic PCRs. Subsequently, the *Aspergillus fumigatus* pyrG marker was eliminated from all strains by selecting on 5-fluoroorotic acid medium before final verification by two additional complementing diagnostic PCRs (see Fig. S3 and Table S2). All strains have been deposited in the IBT strain collection, DTU, (http://www.fbd.dtu.dk/straincollection/).

**Construction of the ausA-S1660A strain**

The amino acid substitution of serine to alanine, S1660A, in *ausA* (AN8383) was created by USER fusion (Geu-Flores et al., 2007) according to the method described by Nielsen et al. (2011). The allele was transferred to IBT29539 and the pyrG marker was eliminated by direct repeat recombination, creating strain IBT31032 containing only the desired point mutation. The strain was verified to contain the *ausA-S1660A* allele by sequencing (StarSEQ, Germany). See Table S3 for primer details.

**Ectopic integration of *ausA* into IS1**

The gene, *ausA*, was PCR amplified by USER fusion (Geu-Flores et al., 2007) and inserted into both pUT1111-1 and pU1211-1 (Hansen et al., 2011) creating pDH23 (gpdA promoter) and pDH24 (aA promoter), respectively. For both plasmids, the gene-targeting substrate was excised by NotI digestion and transformed into IBT28738 using *A. nidulans* argB as a selectable marker. Transformants were streak purified and verified for correct integration into the IS1 site (Hansen et al., 2011) by two complementing diagnostic PCRs.

**Chemical analysis of the mutants**

Strains were inoculated as three point stabs on solid media and incubated for 7 days at 37°C in the dark. Metabolite extraction was performed according to the micro extraction procedure (Smedsgaard, 1997). Extracts were analyzed by two methods: (1) Ultra-high performance liquid chromatography-diode array detection (UHPLC-DAD) analyses using a Dionex RSLC Ultimate 3000 (Dionex, Sunnyvale, CA) equipped with a diode-array detector. Separation of 1 μL extract was obtained on a Kinetex C₁₈ column (150 × 2.1 mm, 2.6 μm; Phenomenex, Torrence, CA) at 60°C using a linear water–acetonitrile gradient starting from 15% CH₃CN to 100% (50 ppm trifluoroacetic acid) over 7 min at a flow rate of 0.8 mL min⁻¹. (2) Exact mass, HPLC-DAD-HRMS, was performed on a 5 cm × 3 μm, Luna C₁₈(2) column (Phenomenex) using a water–acetonitrile gradient from 15% CH₃CN to 100% to 20% in 20 min (20 mM formic acid). LC-DAD-MS analysis was performed on a LCT oaTOF mass spectrometer (Micromass,
Isolation and structure elucidation of secondary metabolites

3,5-Dimethylorsellinic acid and dehydroaustinol were purified from large-scale ethyl acetate extracts prepared from 100 MM agar plates after 4 days’ cultivation in darkness at 37 °C. The compounds were purified using a 10 × 250 mm Phenomenex pentafluorophenyl column (5 μm particles) with a water–acetonitrile gradient from 15% to 100% CH₃CN in 20 min using a flow of 5 mL min⁻¹. Arugosin A was isolated from an ethyl acetate extract of the reference strain grown on 200 CYAs agar plates using a Waters 19 × 300 mm C18 Delta Pak column (15 μm particles), gradient from 80% to 90% CH₃CN in 10 min, and a flow of 30 mL min⁻¹. The NMR spectra were acquired on a Varian Unity Inova 500 MHz spectrometer using standard pulse sequences. Additional details about the compound identification can be found in the supporting information.

Results and discussion

Selecting media that support secondary metabolite production in *A. nidulans*

The principle of using different media and/or incubation conditions for fungal secondary metabolite production has often been promoted (Oxford et al., 1935; Davis et al., 1966; Pitt et al., 1983; Bode et al., 2002; Scherlach & Hertweck, 2006). Based on our previous experiences (Frisvad, 1981; Frisvad & Filtenborg, 1983; Filtenborg et al., 1990; Frisvad et al., 2007), eight different media, CYA, CYAs, CY20, MM, RT, RTO, YE and YES, were initially selected for the analysis (Fig. 1a). HPLC analyses revealed a large number of different secondary metabolites produced by the *A. nidulans* reference strain on CYA, CYAs, CY20, RT, RTO, YE and YES (Fig. 1b) and these metabolites served as a source for further investigation.

Fig. 1. (a) The appearance of the Aspergillus nidulans reference strain, IBT29539, cultivated on different media. (b) UV-chromatograms (210 nm) of extracts of the reference strain provides an excellent demonstration of the profound effect different media compositions have on production of secondary metabolites. From the bottom: minimal media (MM), yeast extract (YE), Raulin-Thom oatmeal (RTO), yeast extract sucrose (YES), Czapek yeast salt (CY20), Raulin-Thom (RT), Czapek yeast agar salt (CYAs), and Czapek yeast agar (CYA). The metabolite profiles on MM and YE media were considered to be redundant and these two media were therefore not used for further screening. Numbers above peaks correspond to compound identity as follows: monodictyphenone (1), 3,5-dimethylorsellinic acid (2), sterigmatocystin (3), arugosin A (4), arugosin H (5), 2-α-dihydroxyemodin (7), α-hydroxyemodin (8), emodin (9), emericellin (10), shamixanthone (11), epi-shamixanthone (12), violaceol I (13), violaceol II (14), lecanoric acid (15), F9775A (16), F9775B (17), austinol (18) and dehydroaustinol (19).
Construction and initial characterization of a genome-wide PKS gene deletion library

To investigate whether any of the compounds observed in Fig. 1 could be genetically linked to a PKS gene, we decided to take a global approach and individually deleted all 32 (putative and known) PKS genes in the A. nidulans genome (see Fig. S4), as defined from the annotation of the genome databases at the Broad Institute of Harvard and MIT and the Aspergillus Genome Database at Stanford (Arnaud et al., 2010). All genes were individually deleted by replacing the entire ORFs using gene-targeting substrates based on the pyrG marker from A. fumigatus for selection. Before analyzing the deletion mutant strains, the pyrG marker was excised by direct repeat recombination (Nielsen et al., 2006) in each case. This was carried out to ensure that the analyses of individual mutant strains were comparable to and not influenced by differences in the primary metabolism due to gene cluster-specific expression levels of the pyrG marker. All 32 deletion mutant strains (see Table S4) were viable and able to sporulate, showing that none of the 32 genes are essential for growth and that no polyketide product is essential for conidiation. As expected, the one strain carrying the wA D mutation formed white conidiospores as it fails to produce the naphthopyrone, YWA1, the precursor of green conidial pigment (Watanabe, 1998; Watanabe et al., 1999).

In addition to wA, eight additional PKS genes have previously been linked to metabolites. In our analysis, key compounds representing four of these gene clusters could be detected: monodictyphenone (1) (observed on RTO, YES and CY20), orsellinic acid (2) (observed on YES, CY20, RT, CYAs and CYA), emericellamide (A) (3) (observed on all media) and sterigmatocystin (4) (observed on RTO, CYAs and CYA). To verify the previously published gene links to these compounds, we individually compared the metabolic profiles of the reference strain to the corresponding profiles obtained with the single PKS gene deletion mutant strains. In agreement with previous analyses, these four compounds

Fig. 2. The late portion of the UV-chromatogram (210 nm) of the Aspergillus nidulans reference strain and mdpG D on RTO medium. The deletion abolishes formation of arugosin A (5), arugosin H (6), emericellin (10), shamixanthone (11) and epi-shamixanthone (12) and a few unidentified peaks.

Fig. 3. Proposed biosynthetic route of the arugosins and shamixanthones. As seen in Fig. 2, arugosin A is produced in large amounts relative to the expected end products: emericellin (10) and shamixanthones (11 and 12). DMAPP, dimethylallyl pyrophosphate. MdpG and XptA are enzymatic activities as proposed by Sanchez et al. (2011).

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

disappeared in mdpG (Bok et al., 2009), orsA (Schroeckh et al., 2009), osAB (Chiang et al., 2008) and stcA (Yu & Leonard, 1995) deletion strains of our library (Fig. S5). Compounds resulting from the remaining four PKS genes were identified by activating the gene clusters by controlled expression of the transcription factor gene in the cluster (Bergmann et al., 2007; Chiang et al., 2009) or by deleting sumO (Szewczyk et al., 2008). Expression from these clusters is apparently not triggered by growth on any of our media, and natural conditions provoking their activation remain to be discovered.

Next, we performed a comparison of the metabolite profiles from the 32 deletion mutants with those obtained with the reference strain with the aim of uncovering novel genetic links between PKS genes and polyketides. The most significant changes are described below.

**mdpG is involved in the biosynthesis of arugosins**

First we focused our attention on the most prominent compound produced on RTO, YES, CY20 and RT media, which eluted as a broad peak around 7.2 min. This compound completely disappeared in the mdpGΔ strain (Fig. 2 and Fig. S6). Moreover, it had a characteristic UV spectrum in the UHPLC analysis indicating an arugosin-like metabolite. MS analysis indicated the compound to be arugosin A (m/z 425 a.m.u. for [M + H]+), which to our knowledge has not been reported before from *A. nidulans*. We therefore decided to confirm the structure of this compound (5). A large-scale extraction was performed and the metabolite was purified. The NMR data in dimethyl sulfoxide are in agreement with the literature (Kawahara et al., 1988) for the hemiacetal form of arugosin A except that the equilibrium was shifted completely to the open form (Fig. 3). In methanol, the NMR data showed that the compound exists in equilibrium between the closed and open ring form (data not shown), explaining the broad peak observed in Fig. 2. A minor peak could be assigned as a mono-prenylated arugosin as [M + H]+ at m/z 357 a.m.u. The MS data of this compound did not indicate loss of a prenyl moiety, suggesting that it is arugosin H (6), a likely immediate precursor of arugosin A (Fig. 3). Hence, our data show that mdpG, which is known for its role in formation of monodictyphenone, is also involved in formation of arugosins.

It is not unusual that one PKS gene cluster is responsible for the biosynthesis of a family of structurally similar compounds (Walsch, 2002; Kroken et al., 2003; Frisvad et al., 2004; Arnaoutzias et al., 2008). In the original analysis of the mdpG gene cluster, it was activated due to remodeling of the chromatin landscape, which occurs in a cell deletion strain (Chiang et al., 2010). That study genetically linked the mdpG gene cluster to eight emodin analogues, including several aminated species, which were detected and tentatively identified. In our analyses, we also detected several emodins including 2-oxo-dihydroxyemodin (7), 6-hydroxyemodin (8) and emodin (9), as well as the more apolar compounds emericellin (10), shamanxanthone (11) and epishamanxanthone (12) (Fig. 1 and Fig. S7). Like in the original study, all emodins disappear in our mdpGΔ strain.

Recently, it was demonstrated that the polyketide part of prenylated xanthones also could be coupled to mdpG (Sanchez et al., 2011). Our finding that mdpG is involved in formation of arugosins indicates that these compounds serve as intermediates in the conversion of monodictyphenone into xanthones, Fig. 3. In agreement with this, previous studies have reported arugosins to be precursors for emericellin (10) and shamanxanthones (11) and (12) (Ahmed et al., 1992; Kralj et al., 2006; Mársquez-Fernández et al., 2007), but have not established a genetic link to mdpG.

**AN7903 and orsA are involved in violaceol production**

Our reference strain produces the antibiotic violaceol I (13) and II (14), in significant amounts (Fig. 4 and Fig. S8). These two diphenyl ethers have been identified in *Emericella*...
The violaceols are formed by dimerization of two C7 monomers of 5-methylbenzene-1,2,3-triol, a compound that we could tentatively detect as [M-H]⁻ at m/z 139 in cultivation extracts. The C7 backbone of 5-methylbenzene-1,2,3-triol, may conceivably be formed by decarboxylation of a C8 aldol intermediate as suggested by Turner 40 years ago (Turner, 1971) (Fig. 5). This C8 intermediate also serves as a branch point towards orsellinic acid.

Interestingly, the same compounds that disappear in the orsAΔ strain also disappear in AN7903Δ, a strain missing a PKS gene separated from orsA by only ~20 kb (Fig. 4). This result does not contradict the original assignment of orsA as the PKS gene responsible for production of orsellinic acid. Although the enzymes encoded by the two genes are predicted to share many of the same functional domains, AN7903 is larger by around 500 amino acid residues and contains a methyl-transferase domain, which is not required for orsellinic acid production. Moreover, we note that Schroeckh et al. (2009) observed that both AN7903 and orsA were upregulated when orsellinic acid was induced by co-cultivation with Streptomyces hygroscopicus, indicating cross-talk between the two clusters. Surprisingly, what appear to be trace amounts of orsellinic acid can be detected as m/z 167 [M-H]⁻ in both the AN7903Δ and the orsAΔ strains (Fig. 4). The source of this residual orsellinic acid remains elusive, but it could possibly stem from unmethylated byproducts from the PKS, AN8383, that produces 3,5-dimethylorsellinic acid, see below.

**AN8383 is responsible for austinol and dehydroaustinol biosynthesis**

Interestingly, production of austinol (18) and dehydroaustinol (19) was observed in the reference strain on several media (Fig. 1). Despite the fact that the production of these compounds is known from *A. nidulans* (Szewczyk et al., 2008), they have not yet been assigned to a specific gene.
Only the AN8383 strain failed to produce the two austinols on all the media, which triggered austinol production in the reference strain (Fig. 6a). This phenotype could be rescued by inserting the structural gene of AN8383 under the control of the constitutive gpdA promoter from IS1 (Hansen et al., 2011) (Fig. 6a). Moreover, a point mutant strain AN8383–S1660A also failed to produce austinols on these six media (Fig. 6a). In this strain, the DSL motif of the AN8383 PKS has been mutated to DAL, preventing the phosphopantetheine moiety of coenzyme A to attach to the acyl carrier protein domain of the PKS, thus disrupting polyketide synthesis (Evans et al., 2008). The lack of austinols can thus be linked directly to an AN8383-encoded function rather than to silencing of another gene caused by chromatin changes provoked by the AN8383 deletion.

To confirm the role of AN8383 in austinol production, we constructed a new strain that expresses the AN8383 ORF controlled by the inducible alcA promoter from the ectopic locus, IS1 (Hansen et al., 2011). On inductive medium, the subsequent LC-MS analysis showed a large novel peak eluting at ca. 6 min (see Fig. S9). The corresponding compound was purified and the structure was elucidated by NMR (Fig. S10), identifying 3,5-dimethylorsellinic acid (20), which is in good agreement with the route of synthesis previously proposed for austinol (Fig. 6b; Geris & Simpson, 2008). The lack of austinols can thus be linked directly to an AN8383-encoded function rather than to silencing of another gene caused by chromatin changes provoked by the AN8383 deletion.

Together, the results strongly indicate that AN8383 encodes a PKS producing 3,5-dimethylorsellinic acid and that this compound serves as the first intermediate in the complex biosynthesis of austinol and dehydroaustinol that also involves a yet unidentified prenyl transferase(s). Based on these results, we have named the locus AN8383, ausA.

Concluding remarks

In the present study, we constructed a genome-wide PKS deletion library, which we screened for effects on polyketide production on a variety of media. This analysis has provided novel links between PKS genes and polyketide products demonstrating the strength of this approach. Many PKS genes still remain to be functionally connected to products, as many gene clusters have not yet been activated. As the repertoire of tools and methods to induce gene expression is rapidly increasing, new polyketide compounds will likely soon be uncovered in A. nidulans. To this end, the genome-wide PKS gene deletion library presented here will undoubtedly serve as a useful resource.

Acknowledgements

This work was supported by the Danish Research Agency for Technology and Production, grant # 09-064967. We thank Lisette Knuth-Nielsen for her dedicated and valuable technical assistance in the laboratory. In addition, we thank Rasmus John Normand Frandsen for suggestions and critical reading of the manuscript.

References


Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Eight known metabolites that have been linked to specific PKS genes in Aspergillus nidulans.

Fig. S2. A graphical illustration of the procedure used to make the gene targeting fragments for the PKS deletions.

Fig. S3. Procedure for diagnostic PCR.

Fig. S4. Chromosome map showing the position of the 32 PKS genes.

Fig. S5. Verification that the polyketide is absent in selected deletion mutants.
Fig. S6. Positive mode extracted ion chromatograms for the mdpGΔ strain cultivated on RTO.

Fig. S7. Additional polyketides that were detected in metabolite extracts in this study.

Fig. S8. The mutants ortsΔ and AN7903Δ lack production of violaceols.

Fig. S9. Extracted ion chromatograms of metabolic extracts from the reference and ausΔ strains.

Fig. S10. $^1$H NMR spectrum of 3,5-dimethylorsellinic acid in dimethyl sulfoxide-d$_6$.

Fig. S11. $^1$H NMR spectrum of dehydroaustinol in CDCl$_3$.

Fig. S12. $^1$H NMR spectrum of arugosin A open form in dimethyl sulfoxide-d$_6$.

Table S1. PCR primers for Gateway assembly.

Table S2. Oligonucleotide primers for diagnostic PCR.

Table S3. Additional oligonucleotides used in the study.

Table S4. The constructed Aspergillus nidulans strains have been deposited into the IBT Culture Collection.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.
A Genome-wide Polyketide Synthase Deletion Library uncovers Novel Genetic Links to Polyketides and Meroterpenoids in *Aspergillus nidulans*


**Supporting Information**

Additional Supporting Information may be found in the online version of this article:

**Figures** (page 3)

Fig. S1. Eight known metabolites that have been linked to specific PKS genes in *A. nidulans*.

Fig. S2. A graphical illustration of the procedure employed to make the gene targeting fragments for the PKS deletions.

Fig. S3. Procedure for diagnostic PCR.

Fig. S4. Chromosome map showing the position of the 32 PKS genes.

Fig. S5. Verification that the polyketide is absent in selected deletion mutants.

Fig. S6. Positive mode extracted ion chromatograms for the *mdpGΔ* strain cultivated on RTO.

Fig. S7. Additional polyketides that were detected in metabolite extracts in this study.

Fig. S8. The mutants *orsAΔ* and AN7903Δ lack production of violaceols.

Fig. S9. Extracted ion chromatograms of metabolic extracts from the reference and *ausAΔ* strains.

Fig. S10. $^1$H NMR spectrum of 3,5-dimethyl orsellinic acid in DMSO-$_d_6$.

Fig. S11. $^1$H NMR spectrum of dehydroaustinol in CDCl$_3$.

Fig. S12. $^1$H NMR spectrum of arugosin A open form in DMSO-$_d_6$. 
Supporting Information

**Text** (page 15)

Text S1. Details about compound identification.

**Tables** (page 17)

Table S1. PCR primers for Gateway assembly.

Table S2. Oligonucleotide primers for diagnostic PCR.

Table S3. Additional oligonucleotides used in the study.

Table S4. The constructed *A. nidulans* strains have been deposited into the IBT Culture Collection.
Fig. S1. Eight known metabolites that have been linked to specific PKS genes in *A. nidulans*.

Aspyridone (*adpA*), asperfuraneone (*afoE/G*), asperthecin (*aptA*), YWA1 (*wA*), (1)

monodictyphenone (*mdpG*), (2) orsellinic acid (*orsA*), (3) emericellamide (A) (*easB*) and (4)

sterigmatocystin (*stcA*).
Fig. S2. A graphical illustration of the procedure employed to make the gene targeting fragments for the PKS deletions. Genomic DNA serves as template for two separate PCR reactions (A) that amplify the UP-stream and DOWN-stream regions of the PKS gene targeted for deletion. The oligonucleotide pairs contain approx. 20 nucleotide attB sequences that allow BP-recombination according to the Gateway cloning system (Invitrogen). (B) The PCR products with the specific attB sequences are incorporated into donor vectors (C) that contain A. fumigatus pyrG flanked by a sequence (DR) that serves as a direct repeat in the final mutant. The resulting plasmids (D) are purified from E. coli, and used as template to amplify gene-targeting substrates by PCR. Universal oligonucleotides U1 & U2 are used to amplify the UP-stream targeting fragments; and D1 and D2 are used to amplify DOWN-stream targeting fragments. The final targeting substrates (E) are co-transformed as bipartite gene-targeting substrates.
Fig. S3. Illustration of the diagnostic PCR procedure. A standard repertoire of PCR reactions were used in the analysis of PKS deletion mutant strains. (A) shows PCR check of a targeting event for a given PKS locus. Up and Dw depicts the up- and downstream flanks of the PKS, respectively. The forward primer, P1, anneals outside the targeting sequence and the reverse primer, P2, is placed within the A. fumigatus pyrG (AFpyrG) marker gene. (B) the marker gene has been excised from the locus and the primer pair, P1 or P3 with P4, binds on either side of the remaining direct repeat sequence. The reaction can also occur on reference strain genomic DNA template, but in this case to produce a significantly larger fragment as illustrated in (C). In addition to this, a PCR reaction (C) using P1 or P3 with P5 positioned inside the PKS gene, identifies false positives. The PKS gene is shown as a blue arrow, the marker gene as a light grey arrow and the direct repeat as black squares. Drawing is not to scale. A control PCR reaction with reference strain genomic DNA as template is always included. (D) shows a typical agarose gel result for the PCR verification of a deletion strain (AN0150 (mdpG) in this case). Lanes are indicated by numbers 1-8 below the wells. Lane 1 & 8, 1 kb ladder from New England Biolabs. Lane 2 shows the test for correct replacement of mdpG with the AFpyrG marker by P1 & P2 (test A) in the deletion strain candidate giving the 3.5 kb band, and the expected absence of a band in the WT control reaction in lane 3. Lanes 4 and 5, confirms marker loss after counter-selecting on 5-FOA via primers P1 and P4 (test B), the pop-out event is verified by the 2.9 kb band in lane 4 whereas WT gave the expected larger 7.9 kb band. Lane 6 and 7 with primers P1 & P5 confirm that the deletion strain is pure and does not contain WT nuclei (intact mdpG sequence). All spore PCR reactions were run in triplicates.
Fig. S4. Chromosome map showing the position of the 32 PKS genes on the eight chromosomes of *A. nidulans*. The map was generated based mainly on information from the *Aspergillus* Genome Database at Stanford. The approximate location of the centromere is shown for each of the eight chromosomes. Telomeres are indicated as black tips and are not to scale. PKS genes that have previously been linked to polyketide compounds have been marked in red with the corresponding gene names. The scale bar on the left indicates size in mega bases (Mb).  

a In the present study, we name the PKS gene AN8383, *ausA*, (marked in green) due to its role in the biosynthesis of austinols.  
b These ORFs have been removed from the genome annotation of the *Aspergillus* Genome Database at Stanford during preparation of Version 5. We have made the deletions according to Version 4 of annotation.  
c The annotation of this gene was taken from the genome databases at the Broad Institute of Harvard and MIT.
Fig. S5. Verification that the polyketide is absent in selected deletion mutants by either positive or negative mode extracted ion chromatograms. The mdpGΔ, orsAΔ, easBΔ and stcAΔ mutants are compared to the reference strain for the mass corresponding to the metabolites these genes are known to be involved in. (A) Monodictyphenone m/z 287 [M-H]⁻, (B) orsellinic acid m/z 167 [M-H]⁻, (C) sterigmatocystin m/z 325 [M+H]⁺ and (D) emericellamide A m/z 610 [M+H]⁺, emericellamide C/D m/z 596 [M+H]⁺, emericellamide E/F m/z 624 [M+H]⁺. * Chiang et al, (Chiang et al. 2008) also observed a minor compound with the same mass.
Fig. S6. Positive mode extracted ion chromatograms for the \textit{mdpG\textDelta} strain cultivated on RTO, identifying arugosin A \textit{m/z} 425 [M+H]$^+$.
Fig. S7. Additional polyketides that were detected in metabolite extracts in this study.
Supporting Information  Figures

Fig. S8. (A) UHPLC profiles (UV 210 nm) of orsAΔ, AN7903Δ compared to the reference strain. The peaks are numbered as follows: orsellinic acid (2), ω-hydroxyemodin (8), violaceol I (13), violaceol II (14), lecanoric acid (15), F9775A (16), F9775B (17), austinol (18), dehydroaustinol (19). (B) UV-spectra of violaceol I (left panel) and II (right panel) standard are shown in black, with an overlay shown in red of the UV-spectra of the sample representing peak 14 (left panel) and peak 13 (right panel). There is agreement between standards and samples.
Fig. S9. Extracted ion chromatograms (detection of 3,5-dimethyl orsellinic acid as [M-H]⁻ at m/z 195) of metabolic extracts from (A) the reference strain and the ausAΔ (AN8383Δ) strain on RTO medium and (B) the reference and the ausA overexpression strain on induction medium (glycerol with threonine).
Fig. S10. $^1$H NMR spectrum of 3,5-dimethyl orsellinic acid in DMSO-$d_6$. 
Fig. S11. $^1$H NMR spectrum of dehydroaustinol in CDCl$_3$. 
Fig. S12. $^1$H NMR spectrum of arugosin A open form in DMSO-$d_6$. 
Supporting Information  

Text S1. Details about compound identification.

**Compound Identification**

3,5-dimethyl orsellinic acid. Yellow/brownish powder; High resolution electron spray ionization mass spectra, HRESIMS, m/z 197.0818 [M+H]+, calculated for [C10H12O4+H]+: 197.0814. 1H NMR (DMSO-d6): δ (ppm) = 2.01 (3H, s), 2.05 (3H, s), 2.36 (3H, s), 8.89 (1H, s), 11.85 (1H, s), 13.43 (1H, s); 13C (DMSO-d6): δ (ppm) = 8.7, 12.0, 18.1, 106.0, 108.2, 115.7, 136.4, 157.5, 158.9, 173.9. These 1H and 13C NMR data are in good agreement with published data (Hirota et al., 1997; Andres et al., 1967) (see supplemental Fig. S10).

Dehydroaustinol. Pale reddish compound; HRESIMS m/z 457.1867 [M+H]+, calculated for [C25H28O8+H]+: 457.1857. 1H NMR (CDCl3): δ (ppm) = 6.89 (1H, d, J = 9.8 Hz*), 6.32 (1H, s), 5.89 ((1H, d, J = 9.8 Hz), 5.83 (1H, s), 5.75 (1H, s), 5.65 (1H, s), 5.26 (1H, q, J = 6.8), 4.35 (s, 1H), 2.30 (1H, br. s), 2.10 (1H, ddd, J = 13.8, 13.8, 3.8 Hz*), 1.77 (1H, ddd, J = 13.8, 13.8, 3.8 Hz*), 1.72 (1H, m), 1.71 (3H, s), 1.63 (3H, d, J = 6.8 Hz), 1.51 (3H, s), 1.49 (3H, s), 1.32 (1H, ddd, J = 3.8, 13.8, 13.8 Hz), 1.24 (3H, s). These 1H NMR data are in good agreement with published data (Márquez-Fernández et al., 2007) except for the coupling constants reported for protons resonating at 6.89 ppm, 1.71 ppm and 2.10 ppm that were 2, 27.7 and 27.7 Hz, respectively (the discrepancy is indicated by asterisk above). A coupling constant of 27.7 Hz is not possible due to Karplus equation and dihedral angles, and upon closer inspection, we observed that the signals for protons at 1.71 and 2.10 ppm should have been listed as a ddd, not a dt, thus making the third coupling constant equivalent to the second one. The coupling constant listed in (Márquez-Fernández et al., 2007) for 6.89 ppm does not match up with that given for its partner at 5.89 ppm, both of which should be around 10 Hz, due to cis coupling (see supplemental Fig. S11), indicating that our data is correct.

Arugosin A open form. Yellow compound; HRESIMS m/z 425.1936 [M+H]+, calculated for [C25H28O6+H]+: 425.1959. 1H NMR (DMSO-d6): δ (ppm) = 10.08 (1H, s), 7.07 (1H, d, J = 8.3 Hz), 6.98 (1H, s), 6.10 (1H, d, J = 8.3 Hz), 5.49 (1H, t, J = 7.0 Hz), 5.29 (1H, t, J = 7.1 Hz), 4.39 (2H, d, J = 7.0 Hz), 3.17 (2H, d, J = 7.0 Hz), 2.26 (3H, s), 1.71 (3H, s), 1.70 (3H, s), 1.68 (3H, s), 1.56 (3H, s); 13C (DMSO-d6): δ (ppm) =
15.3, 17.4, 17.6, 25.3, 25.3, 26.7, 71.6, 105.4, 110.3, 118.4, 119.1, 122.5, 124.0, 126.9, 128.5, 131.1, 132.5, 135.7, 139.3, 148.6, 152.3, 158.4, 160.5, 190.1, 201.4 (see supplemental Fig. S12).

**Violaceol I and II.** De-replicated and validated against in-house standards of both compounds. HRESIMS m/z 263.0886 [M+H]^+, calculated for [C_{14}H_{14}O_{5} +H]^+: 263.0914.
### Table S1. PCR primers for Gateway assembly.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Oligonucleotide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN0150-U-1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN0150-U-2</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTctgtatggtggggacgaaaa</td>
</tr>
<tr>
<td>AN0150-D-3</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN0150-D-4</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN0523-U-1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN0523-U-2</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN0523-D-3</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN0523-D-4</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1034-U-1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1034-U-2</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1034-D-3</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1034-D-4</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1036-U-1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1036-U-2</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1036-D-3</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1036-D-4</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN11191-U-1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN11191-U-2</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN11191-D-3</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN11191-D-4</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1784-U-1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1784-U-2</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1784-D-3</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN1784-D-4</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2032-U-1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2032-U-2</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2032-D-3</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2032-D-4</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2035-U-1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2035-U-2</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2035-D-3</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2035-D-4</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2547-U-1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2547-U-2</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2547-D-3</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN2547-D-4</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN3230-U-1</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN3230-U-2</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN3230-D-3</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
<tr>
<td>AN3230-D-4</td>
<td>GGGGACAAGTTTGTACAAAAAAGCAGGCTggactgaagtagactgtaatctgcg</td>
</tr>
</tbody>
</table>
AN3230-U-2  GGGGACAAGTTTGTACAAAAAAGCAGGCTtaacttgtggcagtggagggctc
AN3230-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTccgtagcatttcgctgtgac
AN3230-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTccgctgcatttcgctgtgac
AN3273-U-1  GGGGACAAGTTTGTACAAAAAAGCAGGCTtaactgtggcgactggagggctc
AN3273-U-2  GGGGACAACCTTTATATACAAAAAGCAGGCTccgtagcatttcgctgtgac
AN3273-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTccgtagcatttcgctgtgac
AN3273-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTccgtagcatttcgctgtgac
AN3386-U-1  GGGGACAAGTTTGTACAAAAAAGCAGGCTccatgttgatgttcgctcgt
AN3386-U-2  GGGGACAACCTTTATATACAAAAAGCAGGCTccatgttgatgttcgctcgt
AN3386-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTccatgttgatgttcgctcgt
AN3386-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTccatgttgatgttcgctcgt
AN3612-U-1  GGGGACAAGTTTGTACAAAAAAGCAGGCTGCTCCTGTTGGAGAAGCTGG
AN3612-U-2  GGGGACAACCTTTATATACAAAAAGCAGGCTGCTCCTGTTGGAGAAGCTGG
AN3612-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTGCTCCTGTTGGAGAAGCTGG
AN3612-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTGCTCCTGTTGGAGAAGCTGG
AN5475-U-1  GGGGACAACCTTTATATACAAAAAGCAGGCTacagtacttgaagaatcgaagtcctgccacc
AN5475-U-2  GGGGACAACCTTTATATACAAAAAGCAGGCTacagtacttgaagaatcgaagtcctgccacc
AN5475-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTacagtacttgaagaatcgaagtcctgccacc
AN5475-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTacagtacttgaagaatcgaagtcctgccacc
AN6000-U-1  GGGGACAAGTTTGTACAAAAAAGCAGGCTtcaatgcgaaaaagctgggagaacc
AN6000-U-2  GGGGACAACCTTTATATACAAAAAGCAGGCTtcaatgcgaaaaagctgggagaacc
AN6000-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTtcaatgcgaaaaagctgggagaacc
AN6000-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTtcaatgcgaaaaagctgggagaacc
AN6431-U-1  GGGGACAACCTTTATATACAAAAAGCAGGCTgaaacccgttcagccaacca
AN6431-U-2  GGGGACAACCTTTATATACAAAAAGCAGGCTgaaacccgttcagccaacca
AN6431-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTgaaacccgttcagccaacca
AN6431-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTgaaacccgttcagccaacca
AN6791-U-1  GGGGACAACCTTTATATACAAAAAGCAGGCTcacgccatgagacggatggcac
AN6791-U-2  GGGGACAACCTTTATATACAAAAAGCAGGCTcacgccatgagacggatggcac
AN6791-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTcacgccatgagacggatggcac
AN6791-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTcacgccatgagacggatggcac
AN7022-U-1  GGGGACAACCTTTATATACAAAAAGCAGGCTtcgatccagacgttcattctgttcggaggagg
AN7022-U-2  GGGGACAACCTTTATATACAAAAAGCAGGCTtcgatccagacgttcattctgttcggaggagg
AN7022-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTtcgatccagacgttcattctgttcggaggagg
AN7022-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTtcgatccagacgttcattctgttcggaggagg
AN7071-U-1  GGGGACAACCTTTATATACAAAAAGCAGGCTttactcgtcgcccttctgggctg
AN7071-U-2  GGGGACAACCTTTATATACAAAAAGCAGGCTttactcgtcgcccttctgggctg
AN7071-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTttactcgtcgcccttctgggctg
AN7071-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTttactcgtcgcccttctgggctg
AN7084-U-1  GGGGACAACCTTTATATACAAAAAGCAGGCTttactcgtcgcccttctgggctg
AN7084-U-2  GGGGACAACCTTTATATACAAAAAGCAGGCTttactcgtcgcccttctgggctg
AN7084-D-3  GGGGACAACCTTTATATACAAAAAGCAGGCTttactcgtcgcccttctgggctg
AN7084-D-4  GGGGACAACCTTTATATACAAAAAGCAGGCTttactcgtcgcccttctgggctg
Supporting Information

Table S1.

<table>
<thead>
<tr>
<th>Sequence ID</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN7084-U-2</td>
<td>GGGGACAAGTTTTGTACAAAAAAACGAGGCTccgtctctgagcgtgctggaa</td>
</tr>
<tr>
<td>AN7084-D-3</td>
<td>GGGGACAACCTTTTATATACAAAGTTGtagttacctgtgcgcaagagac</td>
</tr>
<tr>
<td>AN7084-D-4</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGccctcgtaaccaaatcacaccactccc</td>
</tr>
<tr>
<td>AN7825-U-1</td>
<td>GGGGACAAGTTTTGTACAAAAAAGCAGGCTtcatcccagatactcgagctggctgaa</td>
</tr>
<tr>
<td>AN7825-U-2</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTgtggtgtaaggaggaacctcgatgctgga</td>
</tr>
<tr>
<td>AN7825-D-3</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTtcgtttacggtatctgctggtagctgca</td>
</tr>
<tr>
<td>AN7825-D-4</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGTccctcgtaaccaaatcacaccactccc</td>
</tr>
<tr>
<td>AN7838-U-1</td>
<td>GGGGACAAGTTTTGTACAAAAAAGCAGGCTgacactgaaagatgggagcgtgctggctgaa</td>
</tr>
<tr>
<td>AN7838-U-2</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTgtggtgaaggaggacctgagtgctgctgga</td>
</tr>
<tr>
<td>AN7838-D-3</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTggtctcagtattctgtgtctgtagggtgagattcagattgcct</td>
</tr>
<tr>
<td>AN7838-D-4</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTggtctcagtattctgtgtctgtagggtgagattcagattgcct</td>
</tr>
<tr>
<td>AN7903-U-1</td>
<td>GGGGACAAGTTTTGTACAAAAAAGCAGGCTggctctactcgagctgacgtgctgctgaa</td>
</tr>
<tr>
<td>AN7903-U-2</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTgtggtgaaggaggacctgagtgctgctgga</td>
</tr>
<tr>
<td>AN7903-D-3</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTggtctcagtattctgtgtctgtagggtgagattcagattgcct</td>
</tr>
<tr>
<td>AN7903-D-4</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTggtctcagtattctgtgtctgtagggtgagattcagattgcct</td>
</tr>
<tr>
<td>AN7909-U-1</td>
<td>GGGGACAAGTTTTGTACAAAAAAGCAGGCTgagatgcttccctatatgctcgaacgagcca</td>
</tr>
<tr>
<td>AN7909-U-2</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTgtggtgaaggaggacctgagtgctgctgga</td>
</tr>
<tr>
<td>AN7909-D-3</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTgtcggcttcatcattctcggggtcatgtgtaagtaagtagtgctgcaacgagcca</td>
</tr>
<tr>
<td>AN7909-D-4</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTgtcggcttcatcattctcggggtcatgtgtaagtaagtagtgctgcaacgagcca</td>
</tr>
<tr>
<td>AN8209-U-1</td>
<td>GGGGACAAGTTTTGTACAAAAAAGCAGGCTacccaaggttgacaacacgatatctgc</td>
</tr>
<tr>
<td>AN8209-U-2</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTgtggtgaaggaggacctgagtgctgctgga</td>
</tr>
<tr>
<td>AN8209-D-3</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTatgctcgccgcgatcttctatgagcag</td>
</tr>
<tr>
<td>AN8209-D-4</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTggtctcagtattctgtgtctgtagggtgagattcagattgcct</td>
</tr>
<tr>
<td>AN8383-U-1</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTgtgccgtgaagagtcaagctccgtg</td>
</tr>
<tr>
<td>AN8383-U-2</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTgtgccgtgaagagtcaagctccgtg</td>
</tr>
<tr>
<td>AN8383-D-3</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTTCCTGCGAAGACCTGTCTCGGCAAG</td>
</tr>
<tr>
<td>AN8383-D-4</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGGTGGGTGTTGACTGAGGAGC</td>
</tr>
<tr>
<td>AN8412-U-1</td>
<td>GGGGACAAGTTTTGTACAAAAAAGCAGGCTaagacggtgcgctttcatagatgtcagc</td>
</tr>
<tr>
<td>AN8412-U-2</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTtgctggtacctatactgcatcaccaagg</td>
</tr>
<tr>
<td>AN8412-D-3</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTatgctcgccgcgatcttctatgagcag</td>
</tr>
<tr>
<td>AN8412-D-4</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTggtctcagtattctgtgtctgtagggtgagattcagattgcct</td>
</tr>
<tr>
<td>AN8910-U-1</td>
<td>GGGGACAAGTTTTGTACAAAAAAGCAGGCTcccaagacggactacgcagc</td>
</tr>
<tr>
<td>AN8910-U-2</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTggtgcaggatacgcggaggac</td>
</tr>
<tr>
<td>AN8910-D-3</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTgtcggcttcatcattctcggggtcatgtgtaagtaagtagtgctgcaacgagcca</td>
</tr>
<tr>
<td>AN8910-D-4</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTgtcggcttcatcattctcggggtcatgtgtaagtaagtagtgctgcaacgagcca</td>
</tr>
<tr>
<td>AN9005-U-1</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTggctctactcgagctgacgt</td>
</tr>
<tr>
<td>AN9005-U-2</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGGTggaccaaacacttcctttgcgggggtacaa</td>
</tr>
<tr>
<td>AN9005-D-3</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGGTggctctactcgagctgacgt</td>
</tr>
<tr>
<td>AN9005-D-4</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGGTggctctactcgagctgacgt</td>
</tr>
<tr>
<td>anid_12168-U-1</td>
<td>GGGGACAAGTTTTGTACAAAAAAGCAGGCTggaggacgctttggtgaagg</td>
</tr>
<tr>
<td>anid_12168-U-2</td>
<td>GGGGACAACCTTTTATATACAAAGTTGGGTggtcggcttcatcattctcggggtcatgtgtaagtaagtagtgctgcaacgagcca</td>
</tr>
<tr>
<td>anid_12168-D-3</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGTggtcggcttcatcattctcggggtcatgtgtaagtaagtagtgctgcaacgagcca</td>
</tr>
<tr>
<td>anid_12168-D-4</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGGTggtcggcttcatcattctcggggtcatgtgtaagtaagtagtgctgcaacgagcca</td>
</tr>
<tr>
<td>Uni-U1</td>
<td>GGGGACAACCTTTTCTATACAAAGTTGGGTggtcggcttcatcattctcggggtcatgtgtaagtaagtagtgctgcaacgagcca</td>
</tr>
</tbody>
</table>
For each oligonucleotide, the *attB* sequence required for Gateway recombination is capitalized.

* Oligonucleotide for amplifying the final upstream gene targeting substrate from the Gateway upstream plasmid. These oligonucleotides are labeled U1 and U2 in Fig. S2, respectively.

* Oligonucleotide for amplifying the final downstream gene targeting substrate from the Gateway downstream plasmid. These oligonucleotides are labeled D1 and D2 in Fig. S2, respectively.

| Uni-U2<sup>b</sup> | tgctgtccagcttacctcc |
|-------------------|--|---|
| Uni-D1<sup>c</sup> | tacatcgccgtcatcaag |
| Uni-D2<sup>c</sup> | gattttgacacgagccag |

<sup>a</sup> For each oligonucleotide, the *attB* sequence required for Gateway recombination is capitalized.

<sup>b</sup> Oligonucleotide for amplifying the final upstream gene targeting substrate from the Gateway upstream plasmid. These oligonucleotides are labeled U1 and U2 in Fig. S2, respectively.

<sup>c</sup> Oligonucleotide for amplifying the final downstream gene targeting substrate from the Gateway downstream plasmid. These oligonucleotides are labeled D1 and D2 in Fig. S2, respectively.
Supporting Information

Table S2. Oligonucleotide primers for diagnostic PCR.

<table>
<thead>
<tr>
<th>Gene</th>
<th>External Forward</th>
<th>Gap Check Forward</th>
<th>Gap Check Reverse</th>
<th>Internal Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN0150</td>
<td>GAGTTCGCGGGTGGTATAAG</td>
<td>CGACGGGGATCTAGTTAGT</td>
<td>GTGCCGGGGATTTGGGAGT</td>
<td>ACTACCTGCAACGCACTAC</td>
</tr>
<tr>
<td>AN0523</td>
<td>GAGAGACGCATCGGGAAGAA</td>
<td>CCGATGCTGGATCTTTAG</td>
<td>GTGGCCATGTTTCTTTTCT</td>
<td>TGCTTTAAGTTCGCCAG</td>
</tr>
<tr>
<td>AN1034</td>
<td>CATTCTATCCGCTCTCCTCC</td>
<td>CTTAATCTCCGGGCTTCAAAT</td>
<td>GAGGACTGTAAGGCGAATAG</td>
<td>TCTGGTGCTGCCTCTGGT</td>
</tr>
<tr>
<td>AN1036</td>
<td>CATCAGCGAGAGCGGTCAC</td>
<td>GCATCAGCGGAGCGGTCAC</td>
<td>CACTGAGCCGAGGAGGCAA</td>
<td>TAGACAGAGTGTGTTGTC</td>
</tr>
<tr>
<td>AN10430</td>
<td>GAGGACTGCAAGCCACTG</td>
<td>CTTTTCTCGGTTTGAAGTTG</td>
<td>GTCAAGACGCCAAGAAC</td>
<td>TGCCGCCGGTCTGATGCG</td>
</tr>
<tr>
<td>AN11191</td>
<td>CGGCTTCTCGCTCTCCCTC</td>
<td>GCCAATCTTACATGCTCCAG</td>
<td>CTTTCTCAATGTTTTGTC</td>
<td>CCAAGCGAGTATGATAG</td>
</tr>
<tr>
<td>AN1784</td>
<td>CGGGTTCTGCGCCTCCCTCCCT</td>
<td>CGGTTCTGCGCCTCCCTCCCT</td>
<td>GAGGACTGGAAGGCTGAGT</td>
<td>GGCAAGTCGTCGAGGAC</td>
</tr>
<tr>
<td>AN2032</td>
<td>CAACTGCGCGCGGAGCGG</td>
<td>CGCTGCGCGGAGCGGAGGAATT</td>
<td>GTAGGGATGCTGACTAGT</td>
<td>AGACGATTACCAGTGAAG</td>
</tr>
<tr>
<td>AN2035</td>
<td>GAGATGGTGGATGCTGCGT</td>
<td>GTGGTGGGATTGGTGGTGGT</td>
<td>GTGACGAGGAGGAGGAGG</td>
<td>GAGGAGGATGCTGACTAGT</td>
</tr>
<tr>
<td>AN2547</td>
<td>GGCTGCTGCGCTGGAAGTT</td>
<td>GCCTGCGCTGCGCTGGAAGTT</td>
<td>GTGACGAGGAGGAGGAGG</td>
<td>GAGGAGGATGCTGACTAGT</td>
</tr>
<tr>
<td>AN3230</td>
<td>GCCTGCTGAGGAACCCAT</td>
<td>GCTGGATGCTGAGGAACCCAT</td>
<td>GTGACGAGGAGGAGGAGG</td>
<td>GAGGAGGATGCTGACTAGT</td>
</tr>
<tr>
<td>AN3273</td>
<td>GGCGAGCAGAGGAAACAT</td>
<td>GGCACTGCGCTGCGCTGCG</td>
<td>GAGGAGGATGCTGACTAGT</td>
<td>AGACGATTACCAGTGAAG</td>
</tr>
<tr>
<td>AN3386</td>
<td>GCTGCCTGCGCTGAGGCCAT</td>
<td>CGTGGCGCGCTGCGCTGCG</td>
<td>GAGGAGGATGCTGACTAGT</td>
<td>AGACGATTACCAGTGAAG</td>
</tr>
<tr>
<td>AN3612</td>
<td>CTGCGCTGCGCTGAGGAT</td>
<td>GCCTGCGCTGCGCTGAGGG</td>
<td>GTGACGAGGAGGAGGAGG</td>
<td>GAGGAGGATGCTGACTAGT</td>
</tr>
<tr>
<td>AN5475</td>
<td>CAACGACAGTGCACATG</td>
<td>GCAGGAGAGGAACACTAA</td>
<td>CCTCTTTTGCTGCAATGCTA</td>
<td>CCAGCAATCCCTTCCGAAAG</td>
</tr>
<tr>
<td>AN6000</td>
<td>CTGAGACGACGACGACCTAC</td>
<td>GTCGTCGAAGTGCTGCAAATAT</td>
<td>GGAGACCAACGAAGAAACT</td>
<td>CCGATACTCAAGCCAGCAAG</td>
</tr>
<tr>
<td>AN6431</td>
<td>CTGATCGTGCTCGACGG</td>
<td>GAGGTTTATGTGCGCCGCT</td>
<td>GTCGGTATGCTGAGTGGT</td>
<td>AGCAACCCCTCTTCTGCC</td>
</tr>
<tr>
<td>AN6448</td>
<td>GTCTGCGAGAAGCTGACA</td>
<td>GCTTTTTGCTGCTTATTGTT</td>
<td>CATTGGAAAGCTGCTGGA</td>
<td>TCTTACACCCAGGAAATTG</td>
</tr>
<tr>
<td>AN6791</td>
<td>GGAAGATAGCTCCTGGTAC</td>
<td>GATGAGTCTCAAAAGGGGTG</td>
<td>CCCCTTTGCTGCTGGTGT</td>
<td>GTTCTGGAAGGCTAATTGG</td>
</tr>
<tr>
<td>AN7022</td>
<td>GTGGCCACATCTTCCATTAG</td>
<td>CTCTGGTCTGTGAGTAC</td>
<td>GCAGTGGTGCTGCTGAGT</td>
<td>TCGCTCAGCTTATAGAGT</td>
</tr>
<tr>
<td>AN7071</td>
<td>CCTCACTGACTCTTTTTACCTC</td>
<td>GGTCTGGATGTGCTTGGG</td>
<td>CTTGGAAGCTGCTGAGT</td>
<td>CATTTGCAGCCAGGCTTCT</td>
</tr>
<tr>
<td>AN7084</td>
<td>GCAGCCCGTGTCGTTACG</td>
<td>GTGGTTTGTGAAAGGGGTG</td>
<td>GCCCTTTGCTGCTGAGT</td>
<td>CTCCTACAAGCAGCTCAAC</td>
</tr>
<tr>
<td>AN7825</td>
<td>TCTTGCTTGTCGTAGTATTC</td>
<td>GCAATCAGAGGAGCAATCT</td>
<td>CTGGGATTTGCTGCTGAGT</td>
<td>GCTTTTGGCTTGGGTTTC</td>
</tr>
<tr>
<td>AN7838</td>
<td>TCTCTCTGACAGAGACTTG</td>
<td>GCTGAGTCTAACCCTCTGGA</td>
<td>GATGAGTCTGCTGCAATG</td>
<td>TGGGATGAGTTGATG</td>
</tr>
<tr>
<td>AN7903</td>
<td>GTGTGTATTCCTCTCTCCC</td>
<td>GGGAGCGGATTACCTGGCA</td>
<td>CAGTTTCAGGCGGATCCA</td>
<td>AGAATTCCAGAGCCACTA</td>
</tr>
<tr>
<td>AN7909</td>
<td>GGCTTTCCAGAATATTACG</td>
<td>GGGCTCTATCAGAGTGTG</td>
<td>CTCTGGTATGCGCTGTGGA</td>
<td>CAGCGAGGTTAGGCTTAC</td>
</tr>
<tr>
<td>AN8209</td>
<td>CAACCATTTTTGCGGAGCAC</td>
<td>GAATCGCACAAATGAGGCC</td>
<td>GGTCTTGGTGCAACCCTT</td>
<td>CGAGACTCAAATTTGCGG</td>
</tr>
<tr>
<td>AN8383</td>
<td>GCTCTTGCAATCTGCGC</td>
<td>GTGCAAGTACCCCTCTTGA</td>
<td>GTCGGACAAACCCGCTTC</td>
<td>GGAAGAACCAGACACTG</td>
</tr>
<tr>
<td>AN8412</td>
<td>GACGACATCCCTCCCTACGGA</td>
<td>CTTGCTGGATGGTGCTTGGGA</td>
<td>GAGACTGAGGGCTGGGCT</td>
<td>ATTCTTTGCCACATACCTG</td>
</tr>
<tr>
<td>AN8910</td>
<td>AGACTCCTACTAAAGGCTAAG</td>
<td>CACTATTGATCAGAGCGCA</td>
<td>GGGTACGCTGGTAATGAAG</td>
<td>GTGCTGTTGCTGCTGAGT</td>
</tr>
<tr>
<td>AN9005</td>
<td>GGAAGTCGCGAGCGAAAG</td>
<td>GGCAACTGACGACATAGAC</td>
<td>CGGGGAGTTATGAGTACG</td>
<td>ACCTACAACACACATTTG</td>
</tr>
<tr>
<td>Anid_12168</td>
<td>CTTTTAAAGACTCGGCTGCT</td>
<td>GAGCTAGAGTGGGGTTGGG</td>
<td>CCGAGCACATCTGAACTCAA</td>
<td>CACTAGCATAACGCCCTTC</td>
</tr>
</tbody>
</table>
Supporting Information

Table S2.

| AFpyrG²       | - | - | - | - | GGAAGAGAGGTTCACACC |

² These oligonucleotides anneal upstream of the “upstream targeting sequence” for each PKS gene. Labeled “P1” in Fig. S3.

³ These oligonucleotide pairs, “Gap Check” forward and reverse anneal on either side of the targeted gene in the upstream and downstream targeting sequences, respectively. Labeled “P3” and “P4” in Fig. S3.

⁴ These oligonucleotides anneal to the open reading frame of the targeted gene in reverse direction. Labeled “P5” in Fig. S3 (except AFpyrG).

⁵ This oligonucleotide anneals to the A. fumigatus pyrG, priming in reverse direction. Labeled “P2” in Fig. S3.
### Supporting Information

#### Table S3.

Additional oligonucleotides used in the study.

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Name</th>
<th>Oligonucleotide sequence*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction of pU4002</td>
<td>BGHA195</td>
<td>ggttttaaUtaagtcctagcgtctctagacgcggagtt</td>
</tr>
<tr>
<td></td>
<td>BGHA196</td>
<td>ggttttaaUtaagtgctctacgcgaaggaactcttgacga</td>
</tr>
<tr>
<td>Construction of ausA-S1660A</td>
<td>BGHA600</td>
<td>ggttttaaUgaatgatcggtcatagctttgg</td>
</tr>
<tr>
<td></td>
<td>BGHA601</td>
<td>ataaagcgUcgacaccgatatiecgg</td>
</tr>
<tr>
<td></td>
<td>BGHA602</td>
<td>acaagcUgacactgaagctcttaacgaaat</td>
</tr>
<tr>
<td></td>
<td>BGHA603</td>
<td>ggttttaaUggattacctgtctggatgtcagat</td>
</tr>
<tr>
<td></td>
<td>BGHA604</td>
<td>ggttttaaUgaatgatcggtcatagctttgg</td>
</tr>
<tr>
<td></td>
<td>BGHA605</td>
<td>ggttttaaUggattacctgtctggatgtcagat</td>
</tr>
<tr>
<td>Integration of ausA into IS1</td>
<td>FW-AN8383-1</td>
<td>agagcgaUatggggtcccttgatgataata</td>
</tr>
<tr>
<td></td>
<td>RV-AN8383-1</td>
<td>aatgtgUatgcggatatgctgacgg</td>
</tr>
<tr>
<td></td>
<td>FW-AN8383-2</td>
<td>acaattUggcacggaagcgaacgagatatg</td>
</tr>
<tr>
<td></td>
<td>RV-AN8383-2</td>
<td>atttgtgaUttcgtctcgacgcgccc</td>
</tr>
<tr>
<td></td>
<td>FW-AN8383-3</td>
<td>atccaaUccagcagcagagccaa</td>
</tr>
<tr>
<td></td>
<td>RV-AN8383-3</td>
<td>tctgcaUttataaagccacgtctctcg</td>
</tr>
<tr>
<td>Sequencing of ausA</td>
<td>BGHA576</td>
<td>ttgtgcaacactcagcaag</td>
</tr>
<tr>
<td></td>
<td>BGHA577</td>
<td>tcaacagcgaacgaaccaaca</td>
</tr>
<tr>
<td></td>
<td>BGHA578</td>
<td>ctggtatatcaactcagcc</td>
</tr>
<tr>
<td></td>
<td>BGHA579</td>
<td>gatacatcaccagcagcattt</td>
</tr>
<tr>
<td></td>
<td>BGHA580</td>
<td>gcacaggagggacccaaat</td>
</tr>
<tr>
<td></td>
<td>BGHA581</td>
<td>ctctcaacacagggagccgat</td>
</tr>
</tbody>
</table>

*The uracil base necessary for USER cloning is indicated by a “U” in the oligonucleotide sequence.
Table S4. The constructed *A. nidulans* strains have been deposited into the IBT Culture Collection.

<table>
<thead>
<tr>
<th>IBT Strain</th>
<th>Genotype*</th>
</tr>
</thead>
<tbody>
<tr>
<td>29539</td>
<td>argB2, pyrG89, nkuAΔ</td>
</tr>
<tr>
<td>28738</td>
<td>argB2, pyrG89, nkuA-trS::AFpyrG</td>
</tr>
<tr>
<td>30995</td>
<td>argB2, pyrG89, nkuAΔ, mpdΔ</td>
</tr>
<tr>
<td>30996</td>
<td>argB2, pyrG89, nkuAΔ, AN0523Δ</td>
</tr>
<tr>
<td>30997</td>
<td>argB2, pyrG89, nkuAΔ, afoEA</td>
</tr>
<tr>
<td>30998</td>
<td>argB2, pyrG89, nkuAΔ, afoGA</td>
</tr>
<tr>
<td>30999</td>
<td>argB2, pyrG89, nkuAΔ, AN10430Δ</td>
</tr>
<tr>
<td>31000</td>
<td>argB2, pyrG89, nkuAΔ, AN11191Δ</td>
</tr>
<tr>
<td>31001</td>
<td>argB2, pyrG89, nkuAΔ, AN1784Δ</td>
</tr>
<tr>
<td>31002</td>
<td>argB2, pyrG89, nkuAΔ, AN2032Δ</td>
</tr>
<tr>
<td>31003</td>
<td>argB2, pyrG89, nkuAΔ, AN2035Δ</td>
</tr>
<tr>
<td>31004</td>
<td>argB2, pyrG89, nkuAΔ, casBΔ</td>
</tr>
<tr>
<td>31005</td>
<td>argB2, pyrG89, nkuAΔ, AN3230Δ</td>
</tr>
<tr>
<td>31006</td>
<td>argB2, pyrG89, nkuAΔ, AN3273Δ</td>
</tr>
<tr>
<td>31007</td>
<td>argB2, pyrG89, nkuAΔ, AN3386Δ</td>
</tr>
<tr>
<td>31008</td>
<td>argB2, pyrG89, nkuAΔ, AN3612Δ</td>
</tr>
<tr>
<td>31009</td>
<td>argB2, pyrG89, nkuAΔ, AN5475Δ</td>
</tr>
<tr>
<td>31010</td>
<td>argB2, pyrG89, nkuAΔ, aptΔ</td>
</tr>
<tr>
<td>31011</td>
<td>argB2, pyrG89, nkuAΔ, AN6431Δ</td>
</tr>
<tr>
<td>31012</td>
<td>argB2, pyrG89, nkuAΔ, AN6448Δ</td>
</tr>
<tr>
<td>31013</td>
<td>argB2, pyrG89, nkuAΔ, AN6791Δ</td>
</tr>
<tr>
<td>31014</td>
<td>argB2, pyrG89, nkuAΔ, AN7022Δ</td>
</tr>
<tr>
<td>31015</td>
<td>argB2, pyrG89, nkuAΔ, AN7071Δ</td>
</tr>
<tr>
<td>31016</td>
<td>argB2, pyrG89, nkuAΔ, AN7084Δ</td>
</tr>
<tr>
<td>31017</td>
<td>argB2, pyrG89, nkuAΔ, stcAΔ</td>
</tr>
<tr>
<td>31018</td>
<td>argB2, pyrG89, nkuAΔ, AN7838Δ</td>
</tr>
<tr>
<td>31019</td>
<td>argB2, pyrG89, nkuAΔ, AN7903Δ</td>
</tr>
<tr>
<td>31020</td>
<td>argB2, pyrG89, nkuAΔ, orsAΔ</td>
</tr>
<tr>
<td>31021</td>
<td>argB2, pyrG89, nkuAΔ, w4Δ</td>
</tr>
<tr>
<td>31022&lt;sup&gt;a&lt;/sup&gt;</td>
<td>argB2, pyrG89, nkuAΔ, ausAΔ</td>
</tr>
<tr>
<td>31023</td>
<td>argB2, pyrG89, nkuAΔ, adpAΔ</td>
</tr>
<tr>
<td>31024</td>
<td>argB2, pyrG89, nkuAΔ, AN8910Δ</td>
</tr>
<tr>
<td>31025</td>
<td>argB2, pyrG89, nkuAΔ, AN9005Δ</td>
</tr>
<tr>
<td>31026&lt;sup&gt;f&lt;/sup&gt;</td>
<td>argB2, pyrG89, nkuAΔ, anid12168Δ</td>
</tr>
<tr>
<td>31030&lt;sup&gt;f&lt;/sup&gt;</td>
<td>argB2, pyrG89, IS1::P&lt;sub&gt;alcA&lt;/sub&gt;:ausA::T&lt;sub&gt;trpC&lt;/sub&gt;:argB</td>
</tr>
<tr>
<td>31031&lt;sup&gt;f&lt;/sup&gt;</td>
<td>argB2, pyrG89, IS1::P&lt;sub&gt;gpdA&lt;/sub&gt;:ausA::T&lt;sub&gt;trpC&lt;/sub&gt;:argB</td>
</tr>
<tr>
<td>31032&lt;sup&gt;a&lt;/sup&gt;</td>
<td>argB2, pyrG89, nkuAΔ, ausA-S1660A</td>
</tr>
<tr>
<td>31281&lt;sup&gt;i&lt;/sup&gt;</td>
<td>argB2, pyrG89, nkuAΔ, ausAΔ, IS1::P&lt;sub&gt;gpdA&lt;/sub&gt;:ausA::T&lt;sub&gt;trpC&lt;/sub&gt;:argB</td>
</tr>
</tbody>
</table>

<sup>a</sup> all strains carry the veA1 mutation  
<sup>b</sup> nkuAΔ deletion strain (Nielsen et al., 2008)  
<sup>c</sup> transient small repeat in nkuA (Nielsen et al., 2008)  
<sup>d</sup> AN8383 was designated ausA in this study  
<sup>e</sup> The annotation of this gene was from the *Aspergillus* Comparative Sequencing Project, Broad Institute of Harvard and MIT  
<sup>f</sup> Strain contains inducible expression (*alcA* promoter) of ausA in IS1  
<sup>g</sup> Strain contains constitutive expression (*gpdA* promoter) of ausA in IS1  
<sup>h</sup> Strain contains a point mutation in ausA with S166A (serine to alanine)  
<sup>i</sup> The ausA deletion (strain, IBT31022) was complemented by the expression of ausA from IS1
Supporting References


Paper 5

“Characterization of the Aspergillus nidulans 3-methylorsellinic acid gene cluster”

JB Nielsen and ML Klejnstrup (joint 1st author), P Khorsand-Jamal, DK Holm, A Kabat, ML Nielsen, JC Frisvad, CH Gotfredsen, TO Larsen, and UH Mortensen

Manuscript in preparation
**Introduction.** Filamentous fungi are reservoirs of bioactive natural products, i.e. secondary metabolites that are classified into groups such as polyketides, nonribosomal peptides, terpenoids, and hybrid molecules mixing moieties of these classes. The selection of secondary metabolites produced by one fungus inevitably pinpoints to the environmental niche and the challenges it faces. As these products display high chemical diversity, the bioactivities are similarly broad ranged, in extremes ranging from harmful mycotoxins to beneficial drugs. This fuels the interest to link secondary metabolites to their corresponding genes often residing in gene clusters to ultimately map the full biosynthetic pathway. This task has however shown to be cumbersome due to the expression of most gene clusters responsible for secondary metabolite production has been challenging to stimulate under lab conditions. Even in a well-studied fungus as *Aspergillus nidulans*, the biosynthetic pathways of many clusters are still unidentified regarding if they are active at all. Moreover, in several cases only a few metabolites have been coupled to a specific PKS and the remaining products remains to be found (Ahuja 2012).

In *A. nidulans* at least four PKSs have been shown to produce variants based on an orsellinic acid backbone, namely orsellinic acid (OrsA; Schroeckh 2009), 3,5-dimethylorsellinic acid (3,5-DMOA, AusA; Nielsen 2011), 3-methylorsellinic acid (3-MOA; Hansen 2011; Ahuja 2012) and orsellinic acid aldehyde (PkfA; Ahuja 2012). Moreover, Hansen and co-workers showed that 5-MOA could be heterologously produced by transferring the PKS, *mpaC*, the first step in mycophenolic acid production from *Penicillium brevicompactum* (Hansen 2011). Furthermore, Nielsen et al. reported that small amounts of orsellinic acid were still present in extracts of *orsAΔ* (Nielsen 2011). This could indicate that synthesis of orsellinic acid variants as potentially leaky, which could complicate characterization of these pathways. Recently, 3-MOA and cichorine were reported to be produced by AN6448, PkbA (Ahuja 2012). We set out to map and characterize the biosynthetic activities present in the *pkbA* gene cluster. In addition to determining several of the biochemical steps, we found that nitrogen was incorporated by a to our knowledge novel mechanism by the NRPS-like enzyme AN6444. Moreover a novel carbon skeleton was discovered in the metabolites cichonidulol/demethylcichonidulol.
Results and Discussion

Ectopic overexpression of the putative transcription factor AN6446 results in production of 3-MOA and associated metabolites. 32 putative PKS ORFs were deleted and challenged in an OSMAC approach by Nielsen et al., 2011. Through analysis of the data we also found loss of 3-methylorsellinic acid production in the \( pkbA \Delta \) mutant strain grown on CYA, CY20 and RT compared to the reference strain by UHPLC-MS analysis employing a 3-MOA standard, see Fig 1 for comparison on CYA. All of the remaining 31 PKS deletion strains produced 3-MOA under these conditions (data not shown). Our observation corresponds to the finding recently published by Ahuja et al (Ahuja 2012) that the PKS encoding \( pkbA \) is required to produce 3-MOA. The complete loss of 3-MOA production also showed that 3-MOA cannot be derived from the three other orsellinic acid producing gene clusters in \( A. nidulans \), \( ors \), \( pkf \) and \( aus \) under the tested conditions.

![Figure 1](image-url)  

**Figure 1** Extracted ion chromatograms of 3-methylorsellinic acid \([M+H]^+\) (calc. \( m/z \ 183.06483 \pm 0.001 \)) of a standard, the reference strain and \( pkbA \Delta \) strains on CYA medium cultivated for seven days at 37°C.
As 3-MOA production was relatively low in the reference strain extracts and no other metabolites were found to be significantly differentiated in the *pkbAΔ* extracts compared to the reference, we decided to induce production of potential metabolites from PkbA, and searched the genome sequence for neighbor ORFs predicted to encompass transcriptional regulator activity. AN6446 was predicted to have a DNA binding domain of the Myb/SANT family both associated with transcriptional regulation, e.g. the SANT domain is present in numerous chromatin remodeling enzymes. We speculated that AN6446 could be a cluster-specific activator or repressor, and thus constructed an AN6446 overexpression strain (AN6446-Oex) based on ectopic expression from the IS1 site (Hansen 2011). The overexpression of AN6446 clearly had an effect on *A. nidulans*, since the AN6446-Oex compared to the reference displayed a decrease in pigmentation of conidia to a dusty green appearance and a light brown center of the colony after seven days growth on MM, see Fig 2.

![Figure 2](image.png)

**Figure 2.** The reference strain compared to the AN6446-Oex strain cultivated on MM at 37°C for three, five and seven days.

UHPLC-TOF-MS analysis revealed several new peaks appearing in the AN6446-Oex strain compared to the reference indicating that the overexpression had indeed turned on the likely 3-MOA biosynthetic pathway. 3-MOA was, as in the reference strain, present in trace amounts in the AN6446-Oex strain.
Figure 3 UHPLC-TOFMS ESI+ BPC of the micro-extracts of the reference (black) and AN6446-Oex (blue) strains cultivated on MM media at 37°C for a) three, b) five, and c) seven days.
Extracts of the AN6446-Oex strain cultivated for three different periods, i.e. 3, 5 and 7 days, displayed highly similar metabolite profiles, albeit some metabolite were not visible before 5 days. Consequently, the strains were grown for seven days throughout the study. Moreover this showed that MM medium was a suited growth medium for isolation of candidate metabolites under the artificial conditions encountered in the overexpression strain.

**Overexpression of AN6446 resulted in production of two dimers having a novel carbonskeleton**

A large-scale extract of 200 MM dishes cultivating the AN6446-Oex strain was prepared for isolation and structure elucidation of several metabolites. In figure 3, the UHPLC-TOF-MS ESI⁺ BPC of the large-scale extract versus the micro-extracts showed that several of the peaks were not present in the large extracts. Sample preparations and cultivation conditions of a large number of plates could influence this loss of metabolites observed, however further analysis revealed more compounds than visible in the chromatogram as cichorine eluting at 3.2 min was in high concentrations.

![UHPLC-TOF-MS ESI⁺ BPC of the micro-extract (bottom) and 200 plate extract (top) of the overexpression of AN6446-Oex strain cultivated on MM medium for seven days at 37°C.](image-url)
Isolation and structure elucidation by NMR revealed the presence of several metabolites; cichorine, demethylnidulol and 4-hydroxy-3,5-dimethyl-2-pyrone and two dimers of cichorine and nidulol/demethylnidulol, named cichonidulol and demethylcichonidulol, see figure 5. Although present in the extract of AN6446-Oex, nidulol was isolated from the singular deletion mutant strain, AN6444Δ. Both cichorine and nidulol have been reported previously in literature. Cichorine have been isolated from *A. silvaticus* (Kawahara 1988), *Alternaria cichorii* (Stierle 1993) and *A. nidulans* (Ahuja 2012) whereas nidulol have been reported from *A. nidulans* (Aucamp 1968), *A. silvaticus* (Fujita 1984), and *Emericella desertorum* (Nozawa 1987).

![Structures of metabolites](image)

**Figure 5** Structures of metabolites isolated and elucidated throughout this study. 3-methylorsellinic acid, cichorine, demethylcichorine, nidulol, demethylnidulol, cichonidulol, demethylcichonidulol, 4-hydroxy-3,6-dimethyl-2-pyrone.

The structures were elucidated using 2D NMR-spectra, where especially heteronuclear HSQC and HMBC spectra were extensively used due to the absence of proton-proton couplings and low proton density of the metabolites. In addition the results were in accordance with chemical shift values reported in literature for
cichorine (Monreau 2005), 3-methylorsellinic acid (Ahuja 2012), demethylnidulol (El-Feraly 1985), and 4-hydroxy-3,6-dimethyl-2-pyrone (Savard 1994).

The resemblance of the molecular formulas of cichorine ($\text{C}_{10}\text{H}_{11}\text{NO}_4$) and nidulol ($\text{C}_{10}\text{H}_{10}\text{O}_4$) and initial comparison of the $^1\text{H}$-spectra suggested that these metabolites could be related. The main issue in structure elucidation of these two metabolites was whether the carbonyl was situated on $\text{C}_7$ or $\text{C}_6$. This problem was solved by comparison of the carbon chemical shift of cichorine to the shifts reported in the obtained after total synthesis of cichorine (Moreau 2005). For all the metabolites key HMBC correlations from $\text{H}_5$ the $\text{C}_7$ and/or $\text{C}_6$ and from $\text{H}_7/\text{H}_8$ to the carbons present in the aromatic ring, figure 6, aided in the structure elucidation. A NOESY correlation between $\text{H}_8$ and $\text{H}_{10}$ in cichorine which were missing in nidulol confirmed the difference in the two structures.

The UV spectrum of demethylnidulol was identical to the one of nidulol and the molecular formula $\text{C}_9\text{H}_8\text{O}_4$ showed that the difference between the two metabolites was a carbon and two protons. Through analysis of the NMR-spectra it was confirmed that the two metabolites were related and that the difference was a demethylation of $\text{C}_3$ in nidulol which gave demethylnidulol.

Through the analysis of the large-scale AN6446-Oex extract we noticed two small peaks eluting around five minutes which had UV-spectra that resembled the ones of nidulol and cichorine. These metabolites were present in small amounts in the large extract and based on the molecular formula of these metabolites $\text{C}_{20}\text{H}_{21}\text{NO}_6$ and $\text{C}_{19}\text{H}_{19}\text{NO}_6$ we speculated whether they could be dimers of cichorine, 3-MOA, nidulol or demethylnidulol through condensation and the formation of e.g. an ester as seen with the formation of lecanoric acid in the orsellinic acid pathway (Schroeckh 2009). 2.0 mg of $\text{C}_{20}\text{H}_{21}\text{NO}_6$ and 0.6 mg of
$C_{19}H_{19}NO_6$ were isolated from the extract and the structure of $C_{19}H_{19}NO_6$ were elucidated based on 2D NMR analysis to be a dimer of cichorine and nidulol, named cichonidulol. Analysis of the HMBC spectrum, key HMBC couplings of cichonidulol shown in figure 6, revealed that the link of the two molecules was not through an ester formation but interestingly through a C-N bond formed at the carbonyl position in nidulol. A search of the literature showed that cichonidulol and demethylnidulol were novel metabolites with a new carbon skeleton. The metabolite which showed the closest resemblance was the plant metabolite, capaurine, which contains the C-N condensation bond between two ringsystems; however, the nitrogen containing ring is a six membered ring and the substitution patterns of are quite different (Manske 1945).

Table 1 $^{13}$C-chemical shifts (ppm) of 3-methylorsellinic acid, cichorine, nidulol, demethylnidulol, cichonidulol and demethylcichonidulol. All data obtained in DMSO-$d_6$ (referenced to 39.5 ppm)

<table>
<thead>
<tr>
<th>ID</th>
<th>3-MOA</th>
<th>Cichorine</th>
<th>Nidulol</th>
<th>Demethyl-nidulol</th>
<th>Cichonidulol</th>
<th>Demethylcichonidulol</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1'</td>
<td>103.7/-</td>
<td>123.1/-</td>
<td>107.2/-</td>
<td>92.9/-</td>
<td>121.3/116.1</td>
<td></td>
</tr>
<tr>
<td>2/2'</td>
<td>163.0/-</td>
<td>153.6/-</td>
<td>157.5/-</td>
<td>159.6/-</td>
<td>153.6/153.8</td>
<td></td>
</tr>
<tr>
<td>3/3'</td>
<td>107.7/-</td>
<td>118.9/-</td>
<td>117.5/-</td>
<td>101.1/-</td>
<td>120.5/115.3</td>
<td></td>
</tr>
<tr>
<td>4/4'</td>
<td>159.7/-</td>
<td>156.3/-</td>
<td>162.7/-</td>
<td>168.8/-</td>
<td>156.9/158.1</td>
<td></td>
</tr>
<tr>
<td>5/5'</td>
<td>109.9/-</td>
<td>102.9/-</td>
<td>103.0/-</td>
<td>103.2/-</td>
<td>103.4/102.4</td>
<td></td>
</tr>
<tr>
<td>6/6'</td>
<td>139.4/-</td>
<td>131.9/-</td>
<td>148.6/-</td>
<td>151.4/-</td>
<td>131.0/139.5</td>
<td></td>
</tr>
<tr>
<td>7/7'</td>
<td>23.4/-</td>
<td>169.8/-</td>
<td>68.2/-</td>
<td>66.9/-</td>
<td>167.1/71.7</td>
<td></td>
</tr>
<tr>
<td>8/8'</td>
<td>173.8/-</td>
<td>43.1/-</td>
<td>168.5/-</td>
<td>178.1/-</td>
<td>43.2/83.5</td>
<td></td>
</tr>
<tr>
<td>9/9'</td>
<td>7.9/-</td>
<td>9.1/-</td>
<td>8.2/-</td>
<td>-/-</td>
<td>9.3/8.6</td>
<td></td>
</tr>
<tr>
<td>10/10'</td>
<td>-/-</td>
<td>58.9/-</td>
<td>53.9/-</td>
<td>53.9/-</td>
<td>59.0/59.2</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.1 lists the $^1$C-chemical shifts of the above mentioned isolated metabolites where it is seen that the largest differences in shift values occur around at C$_1$ and C$_6$ differs depending on the position of the carbonyl.

The amount of C$_{19}$H$_{19}$NO$_6$ isolated was not sufficient for full 2D analysis; however, comparison of the $^1$H-spectrum to the one of cichonidulol revealed the only difference between the two were a demethylation in nidulol on C$_3$ thereby generating a dimer of cichorine and demethylnidulol, named demethylcichonidulol.

The final metabolite isolated was 4-hydroxy-3,6-dimethyl-2-pyrone whose chemical shifts were in close resemblance to those in the literature (Savard 1994). This metabolite is different from the above mentioned compounds having a completely different carbon skeleton. We therefore speculated if this metabolite was a product of pkbA or we could have activated another PKS encoding gene. We overexpressed pkbA and extracts of pkbA-Oex grown on MM showed high production of both 3-MOA and 4-hydroxy-3,6-dimethyl-2-pyrone (data not shown). Moreover, screening the data from all the PKS deletion strain grown on CYA medium revealed that only the pkbA$\Delta$ strain lost the ability to produce 4-hydroxy-3,6-dimethyl-2-pyrone.

**Delineation of the cluster by gene expression analysis in overexpression and deletion strains of AN6446.**

To verify that AN6446p had an activator effect on the surrounding genes, relative gene expression measurement by qRT-PCR analysis was performed. A deletion of AN6446, which displayed the same phenotype, as the reference was added to the analysis with the hypothesis that this would have the inverse effects from the overexpression strain minimizing the risk of picking false cluster members. Moreover, none of the peaks observed as an effect of the overexpression strains were present in the extracts of AN6446$\Delta$ (data not shown).

RNA was extracted from the reference strain as well as the deletion and overexpression strain of AN6446 grown on MM for four days. Relative gene expression levels were monitored by qRT-PCR on all 25 ORFs ranging from AN6435 to AN6457 both included. In Figure 7 the average fold change (fc) of the three runs
for all genes in the AN6446 overexpression and deletion strains compared to the reference are shown. A level of 1 in fc denotes unaltered expression in respect to the reference strain. One criterion for genes to be a putative cluster member was for AN6446-Oex to have fc above 1 coupled to fc close to 0 for AN6446Δ. As shown in Figure 7, the two levels of AN6446 expression had dramatic effects on the expression of numerous genes in the study. However, based on fc criteria as stated above the number of candidates was narrowed down to six genes besides AN6446: AN6444, AN6447, AN6448, AN6449, AN6450, AN6451. AN6445 and AN6452 both have fc’s close to 1 for AN6446-Oex considering the standard deviation, and will thus not be considered cluster members, though they both have fc’s close to 0 for the AN6446Δ, which could reflect local organization in the chromatin. In both mutant strains, the expression of distant neighbor genes were influenced, since only AN6435, AN6440, AN6441 and AN11921 were unchanged from the wild type with fc=1 in both mutant strains. AN6453 and AN6454 were significantly downregulated in respect to the reference strain. AN6438 to AN6439 were increased 4-6 fold in fc AN6446-Oex, but did not respond to the deletion of the AN6446-Oex. These examples indicate that the genetic alterations did affect the expression in the locus, and therefore some changes could be artifacts and not cluster specific reactions. Taken together the qRT-PCR data suggests the action of seven gene products.
**Figure 7** Results of the qRT-PCR analysis of AN6446-Oex and AN6446Δ on the genes from AN6435-AN6457 compared to the reference strain. An expression of 1 means unaltered expression, over 1 means higher expression and below 1 means lower expression. These data show that AN6444, AN6446, AN6447, AN6448, AN6449, AN6450 and AN6451 are upregulated in the overexpression strain and downregulated in the deletion strain indicating that the pkbA cluster contains these genes.

**Deletion of all genes spanning AN6435-AN6457 to confirm results from gene expression analysis**

Gene deletions of all the 25 initial candidate genes were performed in the AN6446-Oex background to confirm the conclusions obtained from qRT-PCR and to elucidate on the biosynthetic role of the gene products. The six most distant genes in each side of the cluster were deleted in triplets (AN6435-AN6437; AN6438-AN6440; AN6452-AN6454; AN6455-AN6457) and the gene in between as singular deletion mutants.
Figure 8 The phenotypes of the AN6446-Oex, pkbA-Oex, AN6446-Oex AN6435-6437Δ, AN6446-Oex AN6438-6440Δ, AN6446-Oex AN6441Δ, AN6446-Oex AN11920Δ, AN6446-Oex AN11921Δ, AN6446-Oex AN6443Δ, AN6446-Oex AN6444Δ, AN6446-Oex AN6445Δ, AN6446-Oex AN6446Δ, AN6446-Oex AN6447Δ, AN6446-Oex AN6448Δ, AN6446-Oex AN11922Δ, AN6446-Oex AN6449Δ, AN6446-Oex AN6450Δ, AN6446-Oex AN6451Δ, AN6446-Oex AN6452-6454Δ, AN6446-Oex AN6455-6457Δ and AN6446-Oex AN6435-6457Δ strains cultivated on MM medium for seven days at 37°C.

To see whether AN6446 had regulatory effects on secondary metabolite production outside the cluster, a full cluster deletion (AN6435-AN6457) in the overexpression strain was constructed as well. The phenotypes can be compared in figure 8, and the loss of genes, predicted to form the gene cluster of PkbA, indeed stood out from the others in the phenotype display. The phenotypes varied greatly, indicating altered metabolism in the respective strains. To investigate which metabolites originated from this cluster and if correlated with the phenotype of the mutant strains plug micro-extraction was performed (Frisvad 1987, Smedsgaard 1997) and the extracts were analyzed by UHPLC-TOF-MS. In figure 9, 10 and 11 ESI+ BPC of all the extracts can be seen, and they are grouped based on their metabolic patterns.

Interestingly, in figure 9, a number of deletion strains, from AN6435 to AN11921, are displayed where all the key metabolites appearing in the AN6446-Oex strain had disappeared. Since the qRT-PCR data did not support that these genes belonged to the cluster, we speculated that this effect links to local chromosomal aberration due to change in the strain. Inactivating one or some of the genes instead with point mutations would test this hypothesis. In the whole cluster deletion strain, AN6446-Oex AN6435-AN6457Δ, all metabolite production is eliminated including 4-hydroxy-3,6-dimethyl-2-pyrone strongly indicating that this metabolite is a product of the cluster. However, the eliminated production may reflect the deletions of AN6435 to AN11921. A cluster deletion from AN6444 to AN6451 would be an alternative. Moreover, the phenotypes of these strains were very similar as the metabolic profiles were. The strains all showed the same differences from the reference strain, namely the massive production of the siderophore triacetylfusarinine.
Figure 9 UHPLC-TOF-MS ESI+ BPC of the micro-extracts of the (from bottom to top) reference, AN6446-Oex, AN6446-Oex AN6435-6437Δ, AN6446-Oex AN6438-6440Δ, AN6446-Oex AN6441Δ, AN6446-Oex AN11920Δ, AN6446-Oex AN11921Δ, AN6446-Oex AN6435-6457Δ strains cultivated on MM medium for seven days at 37°C.
Figure 10 UHPLC-TOF-MS ESI$^+$ BPC of the micro-extracts of the (from bottom to top) AN6446-Oex, AN6446-Oex AN6443Δ, AN6446-Oex AN6445Δ, AN6446-Oex AN11922Δ, AN6446-Oex AN6452-6454Δ, AN6446-Oex AN6454-6457Δ strains cultivated on MM medium for seven days at 37°C. The figure illustrates that these genes are not involved in the pkbA gene cluster in agreement with the qRT-PCR data.

The ESI$^+$ BPC chromatograms of the micro-extracts of the second group of strains are collected in figure 10. These mutant strains all have chromatograms mimicking the profile of AN6446-Oex. In this group, the colonies of the strains in figure 8 resembled the overexpression strain, and these results are in agreement with the qRT-PCR data, which excluded these genes to be part of 3-MOA gene cluster.

The last group of mutant strains collected in figure 11 as BPC chromatograms of the micro-extracts, actually constitute the strains pointing out from the gene expression analysis as belonging to the cluster as depicted in figure 12. The metabolites isolated in this study are listed in table 1 if present in the extracts of mutant strains. There are also other peaks appearing in the chromatograms, however no metabolite have been assigned to these at the present time.
Figure 11 UHPLC-TOF-MS ESI’ RPC of the micro-extracts of the AN6446-Oex, AN6448-Oex, AN6446-Oex AN6444Δ, AN6446-Oex AN6446Δ, AN6446-Oex AN6447Δ, AN6446-Oex pkbAΔ, AN6446-Oex AN6449Δ, AN6446-Oex AN6450Δ and AN6446-Oex AN6451Δ strains cultivated on MM medium for seven days at 37°C.
Table 2  Detection of metabolites (through EIC\textsuperscript{+}) in the micro-extracts of the deletion and overexpression strains constructed and identified as belonging to the 3-methylorsellinic clusters in this study; (from bottom to top) AN6446-Oex, AN6448-Oex, AN6446-Oex AN6444\textDelta, AN6446-Oex AN6447\textDelta, AN6446-Oex pkbA\textDelta, AN6446-Oex AN6449\textDelta, AN6446-Oex AN6450\textDelta, AN6446-Oex AN6451\textDelta. Parenthesis indicates the metabolites are present in trace amounts.  
*A metabolite eluting close to cichorine with a small mass difference makes further analysis needed for clarification.  
*A metabolite eluting close to cichorine with a small mass difference makes further analysis needed for clarification.**Eluting at 2.9 minutes.***Eluting at 4.7 minutes

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>AN6446-Oex</th>
<th>pkbA -Oex</th>
<th>AN6444\textDelta</th>
<th>AN6447\textDelta</th>
<th>pkbA\textDelta</th>
<th>AN6449\textDelta</th>
<th>AN6450\textDelta</th>
<th>AN6451\textDelta</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-methyl-orsellinic acid</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>(+)</td>
<td>+</td>
</tr>
<tr>
<td>Cichorine*</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Nidulol</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Demethyl-nidulol</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Cichonidulol</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Demethyl-cichonidulol</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4-hydroxy-3,6-dimethyl-2-pyrene</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>(+)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>C\textsubscript{9}H\textsubscript{12}O\textsubscript{2}</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C\textsubscript{10}H\textsubscript{15}O\textsubscript{3}</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>(+)</td>
<td>-</td>
<td>(+)</td>
</tr>
<tr>
<td>C\textsubscript{10}H\textsubscript{15}O\textsubscript{4}**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C\textsubscript{10}H\textsubscript{15}O\textsubscript{4}***</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>C\textsubscript{10}H\textsubscript{15}NO\textsubscript{3}</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

\* strain having AN6446-Oex allele

In the deletion of \textit{pkbA} in the AN6446-Oex strain the isolated metabolites are all, except demethylnidulol, missing confirming that these metabolites are part of the 3-MOA gene cluster. Demethylnidulol were detected in trace amount and could be due to the conversion of orsellinic acid into the product, however this metabolite cannot be detected in several of the mutants.
AN6450 is predicted to encode an oxidoreductase, and showed in the qRT-PCR analysis to be regulated by AN6446, however the metabolite profile of AN6446-Oex AN6450Δ was equal to that of AN6446-Oex indicating that either it does not take part in the biosynthesis, or the product this enzymatic step was not detected.

Interestingly, AN6451 is predicted to be a transporter of the Major Facilitator Superfamily (MFS) having close to full sequence coverage and more than 50 % sequence identity to transporters from other fungi, e.g. the CtnC transporter of the citrinin biosynthesis in Monascus purpureus, and this deletion mutant is heavily impaired in growth. This indicates that some compartmentalization of the biosynthesis may occur and failure to transport either 3-MOA or 4-hydroxy-3,6-dimethyl-2-pyrone, or even other products in the pathway can be toxic, at least under the artificial conditions that overexpression of AN6446 offers.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Predicted function</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN6444</td>
<td>NRPS-like</td>
</tr>
<tr>
<td>AN6446</td>
<td>Transcription factor</td>
</tr>
<tr>
<td>AN6447</td>
<td>O-methyltransferase</td>
</tr>
<tr>
<td>pkbA</td>
<td>PKS</td>
</tr>
<tr>
<td>AN6449</td>
<td>Cytochrome P450</td>
</tr>
<tr>
<td>AN6450</td>
<td>Oxidoreductase</td>
</tr>
<tr>
<td>AN6451</td>
<td>Transporter</td>
</tr>
</tbody>
</table>

Based on the metabolic profile of the deletion strains a biosynthetic pathway from 3-methylorsellinic acid to cichonidulol is proposed, figure 13. 4-hydroxy-3,6-dimethyl-2-pyrone disappears in the AN6446-Oex pkbAΔ strain and are only present in trace amounts in the AN6446-Oex AN6447Δ strain. As opposed to 3-methylorsellinic acid the polyketide backbone seems to be C₆ and not C₈ and may be a shunt product of PkbA. If the methylation pattern in the assembly of the polyketide backbones of 3-methylorsellinic acid and
4-hydroxy-3,6-dimethyl-2-pyrone were the same one would expect the latter to be methylated at C₃ and not C₁ as is observed. This shows that there are several differences between the assembly of the two metabolites. Other metabolites with the same UV-spectrum as 4-hydroxy-3,6-dimethyl-2-pyrone were detected in the deletion strains but have not yet been isolated. Ahuja and co-workers (Ahuja 2012) isolated two products of the PkgA enzyme, dehydrocitreoisocoumarin and 6,8-dihydroxy-3-(2-oxopropyl)-isocoumarin where the difference of the polyketide backbone also was the incorporation of one malonyl-CoA less in the latter.

The first step of the biosynthesis is likely to be oxidation and reduction of 3-MOA. The proposed intermediates are the 3-MOA in which the carboxylic acid has been reduced to an aldehyde (C₉H₁₀O₃) and the 6-methyl oxidated (C₉H₁₀O₃). Due to the appearance of metabolites identified to have these molecular formulas in the AN6447Δ strain the next step is proposed to be O-methylation of the 2-OH catalyzed by the enzyme encoded by AN6447. This metabolite could then undergo ring closure to form nidulol. The C₇-carbonyl of nidulol, silvaticol, has also been reported in the literature isolated from A. silvaticus (Fujita 1984) and Kawahara and co-workers (Kawahara 1988) suggest quadrilineatin as the intermediate before ring closure thereby explaining that both metabolites nidulol and silvaticol are present in A. silvaticus. Neither silvaticol nor quadrilineatin have been detected in any of the examined extracts yet.

The metabolic profile of the AN6446-Oex AN6449Δ and AN6446-Oex AN6451Δ appear to be similar. Several peaks appear in these two deletions where the molecular formula for one of these is equal to the proposed methylated product of AN6447 (C₁₀H₁₂O₄) indicating that these enzymes are involved in the next steps of the biosynthesis. The roles for these two enzymes in the biosynthesis are still unknown but identification of the intermediates could provide helpful answers. AN6449 encodes for a cytochrome P450 which could be responsible for the oxidation of the C₁₀H₁₂O₄ to a quadrilineatin proposed in figure 13 which could then undergo further modifications to cichorine.
Figure 13 Proposed biosynthetic route of the elucidated metabolites of the pkbA cluster. The metabolites in boxes have all been identified through NMR-analyses, whereas the metabolites in brackets are proposed intermediates which have been detected through MS-analysis and the rest have not been detected.

In the AN6444Δ strain the production of several metabolites have terminated including cichorine and other nitrogen containing metabolites for example a metabolite eluding at 4.2 minutes with the molecular formula C_{15}H_{12}NO_{5} (identified through LC-TOF-MS). The role of AN6444 strongly indicates contribution of the nitrogen to the molecule and ring closure. As no new intermediates were detected in the extracts the exact role of the enzyme remains elusive. The nitrogen incorporated in the structure could arise from either transamination as seen in the biosynthesis of amino acids or from an amino acid which reacts with the dialdehyde. This reaction could yield a lactam where the rest of the amino acid is still attached to the
nitrogen followed by decarboxylation as seen for erinacerin A, figure 14, where the residual part of decarboxylated phenyalanine is attached to the lactam-N (Yaoita 2005). The above mentioned metabolite, C$_{15}$H$_{19}$NO$_{5}$, which was present in the AN6446-Oex and AN6446-Oex AN6449Δ strains could indeed resemble the product, figure 14, of these reactions if the amino acid was 2-aminoadipic acid. Bioinformatic analysis of AN6444 supports this hypothesis since predicts the adenylation domain of AN6444 to encode for this amino acid. This amino acid residue would then have to be cleaved from the molecule to reach cichorine. A way to test this hypothesis is to do feeding experiments with labeled 2-aminoadipic acid. Aspernidine A and B (not detected in these extracts) are two other metabolites reported from *A. nidulans* by Scherlach and co-workers (Scherlach 2010) where nitrogen is incorporated into aromatic part of the structure and is proposed to be derived from an orsellinic acid. The nitrogen of these metabolites could also arise from AN6444 or other NRPS-like enzymes.

![Figure 4.32](image_url) The structure of erinacerin A, 2-aminoadipic acid and proposed structure of C$_{15}$H$_{19}$NO$_{5}$.

It remains unknown whether the dimers, cichonidulol and demethylcichonidulol, were catalyzing by an enzyme or the dimerization occurred non-enzymatic. Due to the small production of the metabolites in the overexpression strain the data obtained is not sufficient for identifying the enzyme.

This study has through qRT-PCR and subsequent chemical analysis identified the genes belonging to the 3-methylorsellinic acid gene cluster. Several metabolites have been isolated, structure elucidated and linked to the gene cluster e.g. nidulol, cichonidulol, demethylcichonidulol and 4-hydroxy-3,6-dimethyl-2-pyrone. The function of AN6447 has been shown to be O-methylation and data strongly suggests that AN6444 is involved in the incorporation of a nitrogen atom into the structure which might occur through an aminoadipic intermediate. Several interesting intermediates have been identified through QTOF-MS analysis and remain
to be isolated and structure elucidated to shed more light over the biosynthesis. A shunt product of the PKS has been identified and through analysis of the PKS deletion library and data of the strains constructed in this work it have been linked to PkbA.

**Materials and methods**

**Strains and media.** The *A. nidulans* strain IBT29539, *nkuΔ, (argB2, pyrG89, veA1, nkuΔ)* was used for overexpressing AN6446, and the IBT28738 (*nkuΔ-trS*) was used as reference in qRT-PCR experiments as well as a control in metabolite analysis (*argB2, pyrG89, veA1, nkuΔ-trS::APyrG*) (Nielsen 2008). In table 1, all strains available and constructed in this study are listed. *Escherichia coli* strain DH5α was used for cloning.

Fungal minimal medium (MM) was as described in (Cove 1966), but with 1% glucose, 10 mM NaNO₃ and 2% agar (Sorbigel, Hendaye, France), when necessary supplemented with 10 mM uridine (Uri), 10 mM uracil (Ura) and 4 mM L-arginine (Arg). Elimination of the *pyrG* marker was done on 5-FOA (1.3 mg/ml) supplemented MM, as previously described in Nielsen (2006). All chemicals were from Sigma-Aldrich (St. Louis, Mo, USA) unless otherwise stated.

**Table 4. List of strains used in this study**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>nkuΔ</em></td>
<td><em>nkuΔ</em></td>
</tr>
<tr>
<td><em>nkuΔ-trS</em></td>
<td><em>nkuΔ-trS::APyrG</em></td>
</tr>
<tr>
<td>AN6444Δ</td>
<td><em>nkuΔ</em>, AN6444Δ</td>
</tr>
<tr>
<td>AN6445Δ</td>
<td><em>nkuΔ</em>, AN6445Δ</td>
</tr>
<tr>
<td>AN6446Δ</td>
<td><em>nkuΔ</em>, AN6446Δ</td>
</tr>
<tr>
<td>AN6446-Oex</td>
<td><em>nkuΔ</em>, IS1::PgpdA-AN6446-TrpC::APyrG</td>
</tr>
<tr>
<td>AN6446-Oex AN6451Δ</td>
<td><em>nkuΔ</em>, IS1::PgpdA-AN6446-TrpC, AN6451Δ::APyrG</td>
</tr>
<tr>
<td>AN6446-Oex AN6450Δ</td>
<td><em>nkuΔ</em>, IS1::PgpdA-AN6446-TrpC, AN6450Δ::APyrG</td>
</tr>
<tr>
<td>AN6446-Oex AN6449Δ</td>
<td><em>nkuΔ</em>, IS1::PgpdA-AN6446-TrpC, AN6449Δ::APyrG</td>
</tr>
</tbody>
</table>
AN6446-Oex AN11922Δ nkuΔ, IS1::PgpdA-AN6446-TtrpC, AN11922Δ::AFpyrG
AN6446-Oex AN6448Δ nkuΔ, IS1::PgpdA-AN6446-TtrpC, AN6448Δ::AFpyrG
AN6446-Oex AN6447Δ nkuΔ, IS1::PgpdA-AN6446-TtrpC, AN6447Δ::AFpyrG
AN6446-Oex AN6446Δ nkuΔ, IS1::PgpdA-AN6446-TtrpC, AN6446Δ::AFpyrG
AN6446-Oex AN6455Δ nkuΔ, IS1::PgpdA-AN6446-TtrpC, AN6455Δ::AFpyrG
AN6446-Oex AN6454-AN6452Δ nkuΔ, IS1::PgpdA-AN6446-TtrpC::AFpyrG, AN6454-AN6452Δ::AFpyrG
AN6446-Oex AN6440-AN6438Δ nkuΔ, IS1::PgpdA-AN6446-TtrpC::AFpyrG, AN6440-AN6438Δ::AFpyrG
AN6446-Oex AN6437-AN6457Δ nkuΔ, IS1::PgpdA-AN6446-TtrpC::AFpyrG, AN6437-AN6457Δ::AFpyrG
pkbA-Oex nkuΔ, IS1::PgpdA-pkbA-TtrpC::argB

* Strain from Nielsen et al., 2008

### Primers used in this study

<table>
<thead>
<tr>
<th>USER Deletion Primers</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN6451-DL-Up-FU</td>
<td>GGGTTTAAUTGTCAGAGAAGCGAGGAAC</td>
<td></td>
</tr>
<tr>
<td>AN6451-DL-Up-RU</td>
<td>GGAATTAAUGGCCGCGCAACAATAGATA</td>
<td></td>
</tr>
<tr>
<td>AN6451-DL-Dw-FU</td>
<td>GCCATTAAUCCCTGTTTCAATGCTTTC</td>
<td></td>
</tr>
<tr>
<td>AN6451-DL-Dw-RU</td>
<td>GGGTTTAATAGGATTATCAGTCAATGCTTTC</td>
<td></td>
</tr>
<tr>
<td>AN6450-DL-Up-FU</td>
<td>GGGTTAAUGGAGATGACAGTCAATGCTTTC</td>
<td></td>
</tr>
<tr>
<td>AN6450-DL-Up-RU</td>
<td>GGACTTAAUGAGTACGAAGATGAACAGCC</td>
<td></td>
</tr>
<tr>
<td>AN6450-DL-Dw-FU</td>
<td>GGGTTAAUAGGCTGTACAGCTTACGCTTTC</td>
<td></td>
</tr>
<tr>
<td>AN6450-DL-Dw-RU</td>
<td>GGGTTAAUAGGCTGTACAGCTTACGCTTTC</td>
<td></td>
</tr>
<tr>
<td>AN11922-DL-Up-FU</td>
<td>GGGTTAAUAGGCTGTACAGCTTACGCTTTC</td>
<td></td>
</tr>
<tr>
<td>AN11922-DL-Up-RU</td>
<td>GGACTTAAUGGAGTACGAAGATGAACAGCC</td>
<td></td>
</tr>
<tr>
<td>AN11922-DL-Dw-FU</td>
<td>GGACTTAAUAGGCTGTACAGCTTACGCTTTC</td>
<td></td>
</tr>
<tr>
<td>AN11922-DL-Dw-RU</td>
<td>GGACTTAAUAGGCTGTACAGCTTACGCTTTC</td>
<td></td>
</tr>
<tr>
<td>Primer Name</td>
<td>Sequence 1</td>
<td>Sequence 2</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>AN11922-DL-Dw-FU</td>
<td>GGCATTAAUGCCAAATCATCCCATACAGTG</td>
<td></td>
</tr>
<tr>
<td>AN11922-DL-Dw-RU</td>
<td>GGTCTTAAUTGCGCTGACACACCTATGAA</td>
<td></td>
</tr>
<tr>
<td>AN6448-DL-Up-FU</td>
<td>GGGTTAAUCGGCTAGACTCCACTTCTG</td>
<td></td>
</tr>
<tr>
<td>AN6448-DL-Up-RU</td>
<td>GGGTTAAUCGGCTAGACTCCACTTCTG</td>
<td></td>
</tr>
<tr>
<td>AN6447-DL-Up-FU</td>
<td>GGGTTAAUTGAGATGGTAATGGTGGAGTG</td>
<td></td>
</tr>
<tr>
<td>AN6447-DL-Up-RU</td>
<td>GGGTTAAUCGGCTAGACTCCACTTCTG</td>
<td></td>
</tr>
<tr>
<td>AN6446-DL-Up-FU</td>
<td>GGGTTAAUCGGCTGACACTATGTTG</td>
<td></td>
</tr>
<tr>
<td>AN6446-DL-Up-RU</td>
<td>GGGTTAAUCGGCTGACACTATGTTG</td>
<td></td>
</tr>
<tr>
<td>AN6445-DL-Dw-FU</td>
<td>GGCATTAAUGAGATGGTAATGGTGGAGTG</td>
<td></td>
</tr>
<tr>
<td>AN6445-DL-Dw-RU</td>
<td>GGCATTAAUGAGATGGTAATGGTGGAGTG</td>
<td></td>
</tr>
<tr>
<td>AN6444-DL-Dw-FU</td>
<td>GGCATTAAUGAGATGGTAATGGTGGAGTG</td>
<td></td>
</tr>
<tr>
<td>AN6444-DL-Dw-RU</td>
<td>GGCATTAAUGAGATGGTAATGGTGGAGTG</td>
<td></td>
</tr>
<tr>
<td>AN6443-DL-Dw-FU</td>
<td>GGCATTAAUGAGATGGTAATGGTGGAGTG</td>
<td></td>
</tr>
<tr>
<td>AN6443-DL-Dw-RU</td>
<td>GGCATTAAUGAGATGGTAATGGTGGAGTG</td>
<td></td>
</tr>
<tr>
<td>AN6441-DL-Dw-FU</td>
<td>GGCATTAAUGCCAAATCATCCCATACAGTG</td>
<td></td>
</tr>
<tr>
<td>AN6441-DL-Dw-RU</td>
<td>GGCATTAAUGCCAAATCATCCCATACAGTG</td>
<td></td>
</tr>
<tr>
<td>AN11921-DL-Dw-FU</td>
<td>GGCATTAAUGCCAAATCATCCCATACAGTG</td>
<td></td>
</tr>
<tr>
<td>AN11921-DL-Dw-RU</td>
<td>GGCATTAAUGCCAAATCATCCCATACAGTG</td>
<td></td>
</tr>
<tr>
<td>AN6457-DL-Dw-Fad</td>
<td>aattccagctgaacaccatgGCTGACACCGCTGCAGCACG</td>
<td></td>
</tr>
<tr>
<td>AN6457-DL-Dw-R</td>
<td>GCTGACACGCGCGGTATCTCC</td>
<td></td>
</tr>
<tr>
<td>AN6454-DL-Dw-Fad</td>
<td>aattccagctgaacaccatgGCTGACACCGCTGCAGCACG</td>
<td></td>
</tr>
<tr>
<td>AN6454-DL-Dw-Rad</td>
<td>gatccctggtggttttgctgcagctgcagctgcaccccct</td>
<td></td>
</tr>
<tr>
<td>AN6452-DL-Up-F</td>
<td>GCTGACACGCGCGGTATCTCC</td>
<td></td>
</tr>
<tr>
<td>AN6452-DL-Dw-Rad</td>
<td>gatccctggtggttttgctgcagctgcagctgcaccccct</td>
<td></td>
</tr>
<tr>
<td>AN6440-DL-Dw-Fad</td>
<td>aattccagctgaacaccatgGCTGACACCGCTGCAGCACG</td>
<td></td>
</tr>
<tr>
<td>AN6440-DL-Dw-R</td>
<td>GACAGAAGCCTCACCACATG</td>
<td></td>
</tr>
<tr>
<td>AN6438-DL-Up-F</td>
<td>GACAGAAGCCTCACCACATG</td>
<td></td>
</tr>
<tr>
<td>AN6438-DL-Up-Rad</td>
<td>gatccctggtggttttgctgcagctgcagctgcaccccct</td>
<td></td>
</tr>
<tr>
<td>AN6437-DL-Dw-Fad</td>
<td>aattccagctgaacaccatgGCTGACACCGCTGCAGCACG</td>
<td></td>
</tr>
<tr>
<td>AN6437-DL-Dw-R</td>
<td>GCTGACACGCGCGGTATCTCC</td>
<td></td>
</tr>
<tr>
<td>AN6435-DL-Up-F</td>
<td>GCTGACACGCGCGGTATCTCC</td>
<td></td>
</tr>
<tr>
<td>AN6435-DL-Up-Rad</td>
<td>gatccctggtggttttgctgcagctgcagctgcaccccct</td>
<td></td>
</tr>
<tr>
<td>Check Primers</td>
<td>AN64TT-CHK-F</td>
<td>GGCATTCAAGCAGAAGCTG</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------</td>
<td>---------------------</td>
</tr>
<tr>
<td></td>
<td>AN64TT-CHK-R</td>
<td>GCGTCTGCTTCAGTITG</td>
</tr>
<tr>
<td>AN6457-CHK-Int-R</td>
<td>GCAACGGAAGGAACTGCG</td>
<td></td>
</tr>
<tr>
<td>AN6455-CHK-Dw-F</td>
<td>GCTGCCATGTTCAGTITG</td>
<td></td>
</tr>
<tr>
<td>AN6455-CHK-Dw-R</td>
<td>CCTGCCAGAGGAGCAGTIT</td>
<td></td>
</tr>
<tr>
<td>AN6454-CHK-F</td>
<td>GGGCTAFAACGACGAGAACG</td>
<td></td>
</tr>
<tr>
<td>AN6454-CHK-Int-R</td>
<td>GCTCCCAGGGACGAGGAC</td>
<td></td>
</tr>
<tr>
<td>AN6453-CHK-F</td>
<td>GGGAGGTTCATACGACAGT</td>
<td></td>
</tr>
<tr>
<td>AN6452-CHK-F</td>
<td>GCACAGTGTAGAGAGCAG</td>
<td></td>
</tr>
<tr>
<td>AN6451-CHK-F</td>
<td>TGTTAATCGTGTGGGGCTCA</td>
<td></td>
</tr>
<tr>
<td>AN6451-CHK-R</td>
<td>TGCGATACGAGCCCTTG</td>
<td></td>
</tr>
<tr>
<td>AN6450-CHK-F</td>
<td>GGGTGGTGGTAAAGTAGAAA</td>
<td></td>
</tr>
<tr>
<td>AN6450-CHK-R</td>
<td>GGCGTAGCAGTCTCCCAGT</td>
<td></td>
</tr>
<tr>
<td>AN6449-CHK-F</td>
<td>TGCCGCACTTAAGAAGCTAG</td>
<td></td>
</tr>
<tr>
<td>AN6449-CHK-R</td>
<td>GAGGAACACTGCAAGGCTAT</td>
<td></td>
</tr>
<tr>
<td>AN6448-CHK-F</td>
<td>GCAGAGACGGAACCTCGG</td>
<td></td>
</tr>
<tr>
<td>AN6448-CHK-R</td>
<td>GCCGACCATATGCAGAG</td>
<td></td>
</tr>
<tr>
<td>AN6448-CHK-F</td>
<td>TAATCGACATCAGCAGAG</td>
<td></td>
</tr>
<tr>
<td>AN6448-CHK-R</td>
<td>GAATTCGCGCAATGGCTAGT</td>
<td></td>
</tr>
<tr>
<td>AN6447-CHK-F</td>
<td>TGGGCTCTAGAAGCTCTGTAT</td>
<td></td>
</tr>
<tr>
<td>AN6447-CHK-R</td>
<td>GACACGACACAGAAACTT</td>
<td></td>
</tr>
<tr>
<td>AN6446-CHK-F</td>
<td>GCACCTGACAACTGACATC</td>
<td></td>
</tr>
<tr>
<td>AN6446-CHK-R</td>
<td>TAATGTAGACGAGAGCTGAAAT</td>
<td></td>
</tr>
<tr>
<td>AN6445-CHK-F</td>
<td>GCACACCAATCTCAAAACTG</td>
<td></td>
</tr>
<tr>
<td>AN6445-CHK-R</td>
<td>TACACCTTGCGGAGACATCA</td>
<td></td>
</tr>
<tr>
<td>AN6445-CHK-F</td>
<td>GCCGTCGAGAGGAGGAG</td>
<td></td>
</tr>
<tr>
<td>AN6445-CHK-R</td>
<td>GCACAGAGACGACGGAG</td>
<td></td>
</tr>
<tr>
<td>AN6444-CHK-F</td>
<td>TTGTATTTGCCTTTCGCGCTC</td>
<td></td>
</tr>
<tr>
<td>AN6444-CHK-R</td>
<td>CACGGCATAAAGCTCAAGAAAG</td>
<td></td>
</tr>
<tr>
<td>AN6443-CHK-F</td>
<td>CTCCTACGGTCTCTCTCT</td>
<td></td>
</tr>
<tr>
<td>AN6443-CHK-R</td>
<td>TCGGCCACGTTTACTAC</td>
<td></td>
</tr>
<tr>
<td>AN11922-CHK-F</td>
<td>ATGGGAAGGATCTTACAAAG</td>
<td></td>
</tr>
<tr>
<td>AN11922-CHK-R</td>
<td>TGTCAAAACCTGCTTTCT</td>
<td></td>
</tr>
<tr>
<td>AN11920-CHK-F</td>
<td>GCTCAACGGATACGACAGAG</td>
<td></td>
</tr>
<tr>
<td>AN11920-CHK-R</td>
<td>GCCGACGGAGTCTTCTCTG</td>
<td></td>
</tr>
<tr>
<td>AN6441-CHK-F</td>
<td>GCAGGGAGAGCTGAGTGTG</td>
<td></td>
</tr>
<tr>
<td>AN6441-CHK-R</td>
<td>GCACACGACCGCGCGAG</td>
<td></td>
</tr>
<tr>
<td>AN6439-CHK-Int-F</td>
<td>GCCATACGATCTCCTTTAGC</td>
<td></td>
</tr>
<tr>
<td>AN6438-CHK-F</td>
<td>GCACCTTGAACAGAGGACTG</td>
<td></td>
</tr>
<tr>
<td>AN6436-CHK-R</td>
<td>GCCGCTGGAGAGAGACT</td>
<td></td>
</tr>
<tr>
<td>AN6435-CHK-F</td>
<td>GCCAGGCTGACCAACACTC</td>
<td></td>
</tr>
<tr>
<td>qRT-PCR Primers</td>
<td>AN6457-3-qRT-F</td>
<td>CCTCCGAGAGACGAAACCTG</td>
</tr>
<tr>
<td>AN6457-3-qRT-R</td>
<td>CCTGATCCGCTCGAGAGCC</td>
<td></td>
</tr>
<tr>
<td>AN6456-3-qRT-F</td>
<td>GCCGACATGAGAGCAAGAG</td>
<td></td>
</tr>
<tr>
<td>AN6456-3-qRT-R</td>
<td>GAATCTGATGCTGGGGCTT</td>
<td></td>
</tr>
<tr>
<td>AN6453-3-qRT-F</td>
<td>CCTATCGCTACGCTGAGT</td>
<td></td>
</tr>
<tr>
<td>AN6453-3-qRT-R</td>
<td>GCCAAGGTGGGAGGAGGCG</td>
<td></td>
</tr>
<tr>
<td>Primer ID</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------</td>
<td></td>
</tr>
<tr>
<td>AN6453-3-qRT-R</td>
<td>GGACAAAAGGCTTGAGGGGG</td>
<td></td>
</tr>
<tr>
<td>AN6452-3-qRT-F</td>
<td>CACGACAATCATACTCTCGCG</td>
<td></td>
</tr>
<tr>
<td>AN6452-3-qRT-R</td>
<td>CTTTTCAGTCCCCACTAGCCC</td>
<td></td>
</tr>
<tr>
<td>AN6451-3-qRT-F</td>
<td>CGCCCATGTACCAATCGAC</td>
<td></td>
</tr>
<tr>
<td>AN6451-3-qRT-R</td>
<td>CACACTGGGAGACCCTCTCG</td>
<td></td>
</tr>
<tr>
<td>AN6450-3-qRT-F</td>
<td>CTCTCTGACCGCCAGCATCAGAC</td>
<td></td>
</tr>
<tr>
<td>AN6450-3-qRT-R</td>
<td>GTCCGGAACCCAGAAGTCG</td>
<td></td>
</tr>
<tr>
<td>AN6449-3-qRT-F</td>
<td>CAAAGCGCTGTAACACTCGTG</td>
<td></td>
</tr>
<tr>
<td>AN6449-3-qRT-R</td>
<td>CCCACCGTCTTCAAGAATGC</td>
<td></td>
</tr>
<tr>
<td>AN11922-3-qRT-F</td>
<td>CACCTGTGCTACGAGTAACATATCC</td>
<td></td>
</tr>
<tr>
<td>AN11922-3-qRT-R</td>
<td>GTCCCCTACGAGCAGATCC</td>
<td></td>
</tr>
<tr>
<td>AN6448-3-qRT-F</td>
<td>CTCCCAAGGACACCTCTCCCTTC</td>
<td></td>
</tr>
<tr>
<td>AN6447-3-qRT-F</td>
<td>TATATGCGGATGACTGGGAC</td>
<td></td>
</tr>
<tr>
<td>AN6447-3-qRT-R</td>
<td>GCAATCTCCAAATCCCAACC</td>
<td></td>
</tr>
<tr>
<td>AN6446-3-qRT-F</td>
<td>GGAGATACACGAATGACG</td>
<td></td>
</tr>
<tr>
<td>AN6446-3-qRT-R</td>
<td>GGTACTGGTAAATGTTGGAGGTG</td>
<td></td>
</tr>
<tr>
<td>AN6445-3-qRT-F</td>
<td>GTATGAGATGCCTGATGGCTG</td>
<td></td>
</tr>
<tr>
<td>AN6445-3-qRT-R</td>
<td>GTGGGAGGAAGAAGGAGCATG</td>
<td></td>
</tr>
<tr>
<td>AN6444-3-qRT-F</td>
<td>GACAAACCAGGAAGAAACTCC</td>
<td></td>
</tr>
<tr>
<td>AN6444-3-qRT-R</td>
<td>CTCCAAACCACCCAAATCTCCTCTCCA</td>
<td></td>
</tr>
<tr>
<td>AN6443-3-qRT-F</td>
<td>GTATTGCGATCGACTGAGAC</td>
<td></td>
</tr>
<tr>
<td>AN6443-3-qRT-R</td>
<td>CTCCACCTTCGCACGAAATATCC</td>
<td></td>
</tr>
<tr>
<td>AN11921-3-qRT-F</td>
<td>CTGCCCTACCTTTCTCCTCCCTCGC</td>
<td></td>
</tr>
<tr>
<td>AN11921-3-qRT-R</td>
<td>CCCAGAGTGTACAGACAC</td>
<td></td>
</tr>
<tr>
<td>AN11920-3-qRT-F</td>
<td>CCGTGTACCTCCTCCACCATGC</td>
<td></td>
</tr>
<tr>
<td>AN11920-3-qRT-R</td>
<td>CACACCCGTTTCATTCCAC</td>
<td></td>
</tr>
<tr>
<td>AN6441-3-qRT-F</td>
<td>GGGCAATGGAATACATCCCTGGGG</td>
<td></td>
</tr>
<tr>
<td>AN6441-3-qRT-R</td>
<td>GAGGGAGCCGCTAGGCTTGAG</td>
<td></td>
</tr>
<tr>
<td>AN6440-3-qRT-F</td>
<td>CCGACACAATCAGCTGCC</td>
<td></td>
</tr>
<tr>
<td>AN6440-3-qRT-R</td>
<td>CCGCCCTTGATATCTCTATCCAC</td>
<td></td>
</tr>
<tr>
<td>AN6439-3-qRT-F</td>
<td>CGAGATACCTTAATCTTGCGGC</td>
<td></td>
</tr>
<tr>
<td>AN6439-3-qRT-R</td>
<td>GTCACTCCATCGCAGGCAC</td>
<td></td>
</tr>
<tr>
<td>AN6438-3-qRT-F</td>
<td>CGAGATAGGATACAGATCC</td>
<td></td>
</tr>
<tr>
<td>AN6438-3-qRT-R</td>
<td>CGAGATCCTGCTGCTCTCTTG</td>
<td></td>
</tr>
<tr>
<td>AN6437-3-qRT-F</td>
<td>CACCCGTACCTCGACAC</td>
<td></td>
</tr>
<tr>
<td>AN6437-3-qRT-R</td>
<td>CCAGAGACTGCGAGCCCTC</td>
<td></td>
</tr>
<tr>
<td>AN6436-3-qRT-F</td>
<td>GAGAGCCGACGAAATGCGTGC</td>
<td></td>
</tr>
<tr>
<td>AN6436-3-qRT-R</td>
<td>GGTTCACCGACGAGAGATGAGTGTG</td>
<td></td>
</tr>
<tr>
<td>AN6435-qRT-3-F</td>
<td>GCCCTCAACCTTGATTCCCCCTTG</td>
<td></td>
</tr>
<tr>
<td>AN6435-qRT-3-R</td>
<td>GAGCCCCCTCTTGTGGCGG</td>
<td></td>
</tr>
</tbody>
</table>

Overexpression Primers

<table>
<thead>
<tr>
<th>Primer ID</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN6446-Oex-FU</td>
<td>AGAGGCCGAAUTGAGAACGACACTTTCTGAGTCGAC</td>
</tr>
<tr>
<td>AN6446-Oex-RU</td>
<td>TCCTGCAGAUTCAAGGACCCCTGTGTGAC</td>
</tr>
<tr>
<td>AN6448-Oex-5-FU</td>
<td>AGAGGCCGAAUTGAGACCCTGCTGGCACCATCC</td>
</tr>
<tr>
<td>AN6448-Oex-5-RU</td>
<td>AGGCCGAGAACCAAUGGCGAGCCGCTTTATGAGTGC</td>
</tr>
<tr>
<td>AN6448-Oex-3-FU</td>
<td>AGTTCCTCCCCGGTTGCGGCGCTTTGTAAGACA</td>
</tr>
<tr>
<td>AN6448-Oex-3-RU</td>
<td>TCTGCAGAUTTACCTCCACGACACCTCTCCTC</td>
</tr>
</tbody>
</table>

PCR and DNA substrate construction.
The annotation of AspGD was used for retrieving gene sequence information. Oligonucleotides were produced by IDT (Integrated DNA Technologies Inc, Coralville, IA, USA) or Sigma-Aldrich, see table x for list. The PfuX7 polymerase (Nørholm, 2010) was used in all PCR reactions as described in Hansen et al. (Hansen 2011). All DNA fragments for gene deletion and overexpression were assembled into transformation ready cassettes via the USER cloning technology as presented in Hansen et al., 2011, except for the triplet gene deletions, which were constructed by fusion-PCR as described in Nielsen et al., 2008. Overexpression construct for AN6446 were propagated in pU2111 having AsiSI/Nb.BtsI USER cassettes and AFpyrG flanked by direct repeats (Hansen 2011). The pU2002A vector designed to incorporate gene-targeting sequences flanking DR-AFpyrG-DR was digested with Pacl/Nt.BbvCI prior to mixing with appropriate PCR fragments and USER cloning. All gene-targeting substrates were liberated from their vector backbone with SwaI (1U/μl). Overexpression of pkbA was established in a modified pU1111 vector employing A. nidulans argB as genetic marker (Hansen et al., 2011), pDH57. The pDH57 vector carries the ccdB suicide gene from the Gateway™ cloning system (Invitrogen, Carlsbad, CA, USA), allowing elimination of background from uncut plasmid from the USER™ cloning reaction. The ccdB::camR cassette was amplified from pDONR (Invitrogen) using primers containing USER tails that were designed to restore the AsiSI/Nb.BtsI cassette. Furthermore, a silent point mutation was introduced in order to eliminate an Nb.BtsI nicking site in ampR. Primers for pkbA-Oex construct, split in two parts for USER fusion, were designed using the PHUSER program (Olsen 2011). The transformation substrate was excised from the vector by NotI digestion after manufacturer’s protocol.

USER mix, restriction and nicking enzymes were purchased from New England Biolabs, Ipswich, MA, USA. Protoplasting and gene-targeting procedures were performed as described previously (Johnstone 2005, Nielsen 2006). Gene targeting events were verified by diagnostic spore-PCR in 40 μL reaction volume applying 20 min of initial denaturation.

Relative gene expression by qRT-PCR. The AN6446-Oex, AN6446Δ and the reference strain were inoculated as three-points on MM+Arg for 5 days in the dark at 37 °C. Cultivations for plug extraction were done in parallel to confirm that the levels of gene expression also were resulting in appearance or absence of
metabolites as expected. Total RNA from the three strains was isolated with Qiagen Plant RNAeasy kit (Qiagen, Hilden, Germany). The samples were disrupted by a TissueLyser LT (Qiagen) using 45 Hz for 1 min. 10 μg of RNA was DNase I (Qiagen) treated after manufacturer’s protocol, with addition of 10U of RNAguard RNAsse inhibitor (Amersham Biosciences, Nassau, NY, USA). 1 μg of DNase I treated RNA samples were used in cDNA synthesis by the Phusion RT-PCR Kit (Finnzymes, Vantaa, Finland) according to manufacturer’s protocol. The subsequent qRT-PCR was performed in a Chromo 4TM Detector/PTC-200 (MJ Research, St. Bruno, Canada) using the SYBR® Green JumpStart Taq ReadyMix (Sigma). *hhtA* (AN0733) encoding histone protein H3 was chosen as internal standard for normalization of expression levels. Only one of them, *hhtA*, was used for the fold change calculations. Primer combinations for the qPCR and sequences are listed in Supplementary table1. Moreover, two types of control samples were initially included for the qPCR; the DNAse treated RNA sample used for the reverse transcriptase reaction, and a template-free reaction to test for primer-dimer influence on the overall fluorescence. The final individual cDNA samples were added to the reactions as 5 times diluted samples. Samples were run in triplicates. The program was 94 °C for 2 min and cycling conditions 40 times; 94 °C for 10 s, 60 °C for 15 s, 72 °C for 30 s. A melting curve from 65 °C to 95 °C with reads every 0.2 min was ending the program to evaluate the purity of the reaction products. The fluorescence threshold values, C(t), was determined through the OpticonMonitor 3.1 software (MJ Research). The relative expression levels was approximated by $2^{\Delta\Delta C(t)}$ as

$$\Delta\Delta C(t) = \Delta C(t)_{\text{normalized}} - \Delta C(t)_{\text{calibrator}}.$$ 

The $\Delta C(t)_{\text{normalized}} = \Delta C(t)_{\text{target gene}} - \Delta C(t)_{\text{internal std}}$ and the calibrator $C(t)$ values were the corresponding values from the reference strain.
**Chemical characterization of mutant strains by LC-DAD-MS:** All strains were grown as three point inoculations for 7 days at 37 degrees in the dark on solid media. Extraction of metabolites were performed by the agar plug extraction method (Smedsgaard 1997) using three 6mm agar plugs/extract. Extracts were analyzed by LC-HRMS. LC-HRMS analysis were made on a MaXis 3G QTOF (Bruker Daltronics) coupled to a Dionex Ultimate 3000 UHPLC system equipped with a 100 x 2.0 mm, 2.6 μm, Kinetex C-18 column. The separation column was held at a temperature of 40°C and a gradient system composed of A: 20 mM formic acid in water, and B: 20 mM formic acid in acetonitrile was used. The flow was 0.4 ml/min, 85% A graduating to 100% B in 0-10 min, 100% B 10-13 min, 85% A 13.1-15 minutes. For calibration, a mass spectrum of sodium formate was recorded at the beginning of each chromatogram using the divert valve (0.3-0.4 min). Samples were analyzed both in positive and negative ionization mode. De-replication of induced compounds were performed by comparison of accurate mass to the metabolite database Antibase2009 (Laatsch 2010) as well as comparison of UV spectra to published data.

**Isolation of 3-methylorsellinic acid**

AN6444Δ was cultured on 100 MM plates for 7 days at 37°C in the dark. The plates were homogenized using a Stomacher and 100 mL ethyl acetate (EtOAc) + 1% formic acid (FA) pr. 10 plates. The extract was filtered after 1 hour and the remaining broth was extracted with EtOAc + 1% FA for 24 hours. The extract was filtered and the two fractions pooled and dried down on a freeze drier. The crude extract were separated into three phases by dissolving it in 9:1 MeOH:H_2O – Milli-Q and extracted into a heptane phase followed by addition of H_2O-Milli-Q to 1:1 methanol(MeOH):H_2O – Milli-Q and extracted into a dichloromethane (DCM) phase and a aqueous phase. The aqueous phase was separated on a KP-C18-HS SNAP column using a Biotage Isolera One (Biotage, Uppsala, Sweden) using a gradient of 10% MeOH (H_2O-Milli-Q (Millipore, MA, USA)) to 100% over 15 CVs (column volumes) resulting in 22 fractions. Fraction 17 contained 9.2 mg of pure 3-methylorsellinic acid and were concentrated on a rotavap (Büchi V-855/R-215) and dried down under N_2(g).

**Isolation of cichorine (X1)**

AN6446-Oex was cultured on 200 MM plates for 7 days at 37°C in the dark. The plates were homogenized using a Stomacher and 100 mL EtOAc + 1% FA pr. 10 plates. The extract was filtered after 1 hour and the remaining broth was extracted with EtOAc + 1% FA for 24 hours. The extract was filtered and the two fractions pooled and dried down on a freeze drier. The crude extract were separated into three phases by dissolving it in 9:1 MeOH:H_2O – Milli-Q and extracted into a heptane phase followed by addition of H_2O-Milli-Q to 1:1 MeOH:H_2O – Milli-Q and extracted into a DCM phase and a aqueous phase. The DCM phase
was separated on a KP-C18-HS SNAP column using a Biotage Isolera One (Biotage, Uppsala, Sweden) using a gradient of 10% MeOH (H$_2$O-Milli-Q (Millipore, MA, USA)) to 100% over 15 CVs (column volumes) resulting in 22 fractions. Cichorine was present in fractions 12 and 13 and was purified on a Waters HPLC W600/996PDA (Milford, MA, USA) using a RP column (Phenomenex Luna C18(2), 250 x 21.20 mm, 5 μm, Torrance, CA, USA) using an isocratic gradient of 20% MeCN (H$_2$O – Milli-Q (Millipore, MA, USA)) over 20 minutes with a flow of 15 mL/min. The collections were concentrated on a rotavap (Büchi V-855/R-215) and dried down under N$_2$(g) to yield 9.3 mg of cichorine.

**Isolation of nidulol**

AN6445A was cultured on 100 MM plates for 7 days at 37°C in the dark. The plates were homogenized using a Stomacher and 100 mL EtOAc + 1% FA pr. 10 plates. The extract was filtered after 1 hour and the remaining broth was extracted with EtOAc + 1% FA for 24 hours. The extract was filtered and the two fractions pooled and dried down on a freeze drier. The crude extract were separated into three phases by dissolving it in 9:1 MeOH:H$_2$O – Milli-Q and extracted into a heptane phase followed by addition of H$_2$O-Milli-Q to 1:1 MeOH:H$_2$O – Milli-Q and extracted into a DCM phase and a aqueous phase. The DCM and MeOH:H$_2$O phases were separated seperately on a KP-C18-HS SNAP column using a Biotage Isolera One (Biotage, Uppsala, Sweden) using a gradient of 10% MeOH (H$_2$O-Milli-Q (Millipore, MA, USA)) to 100% over 15 CVs (column volumes) resulting in 22 fractions each. Nidulol was present in fraction 17 of both the phases and was purified on a Waters HPLC W600/996PDA (Milford, MA, USA) using a RP column (Phenomenex Luna C18(2), 250 x 10 mm, 5 μm, Torrance, CA, USA) using a gradient of 15% MeCN (H$_2$O – Milli-Q (Millipore, MA, USA)) to 30% MeCN over 5 min followed by 15 minutes of isocratic flow of 30% MeCN with 50 ppm TFA added to the solvents and a flow of 5 mL/min. The collections were concentrated on a rotavap (Büchi V-855/R-215) and dried down under N$_2$(g) to yield 1.2 mg of nidulol.

**Isolation of demethylnidulol**

The MeOH:H$_2$O phase of the extract of the AN6446-Oex described for the isolation of cichorine was separated on a KP-C18-HS SNAP column using a Biotage Isolera One (Biotage, Uppsala, Sweden) using a gradient of 10% MeOH (H$_2$O-Milli-Q (Millipore, MA, USA)) to 100% over 15 CVs (column volumes) resulting in 22 fractions. Demethylnidulol was present in fractions 12 and was purified on a Gilson 322 liquid chromatograph with a 215 liquid handler/injector (BioLab, Risskov, Denmark) using a RP column (Phenomenex Luna C18(2), 250 x 10 mm, 5 μm, Torrance, CA, USA) using a flow of 15% MeCN (H$_2$O – Milli-Q (Millipore, MA, USA)) over 10 minutes followed by a gradient from 15% to 95% MeCN with a flow of 5
mL/min. The collections were concentrated on a rotarvap (Büchi V-855/R-215) and dried down under N\(_2\) (g) to yield 1.0 mg of cichorine.

**Isolation of cichonidulol**

Cichonidulol was present in fraction 18 of the separated DCM phase described in the isolation of cichorine and was purified on a Gilson 322 liquid chromatograph with a 215 liquid handler/injector (BioLab, Risskov, Denmark) using a RP column (Phenomenex Luna C18(2), 250 x 10 mm, 5 \( \mu \text{m} \), Torrance, CA, USA) using a flow of 15% MeCN (H\(_2\)O – Milli-Q (Millipore, MA, USA)) over 10 minutes followed by a gradient from 15% to 95% MeCN over 30 minutes with a flow of 5 mL/min. The collections were concentrated on a rotarvap (Büchi V-855/R-215) and dried down under N\(_2\) (g) to yield 2.0 mg of cichonidulol.

**Isolation of demethylcichonidulol**

One of the fractions from the last purification step of cichonidulol contained demethylcichonidulol and was purified on a Gilson 322 liquid chromatograph with a 215 liquid handler/injector (BioLab, Risskov, Denmark) using a RP column (Phenomenex Luna C18(2), 250 x 10 mm, 5 \( \mu \text{m} \), Torrance, CA, USA) using a flow of 15% MeCN (H\(_2\)O – Milli-Q (Millipore, MA, USA)) over 10 minutes followed by a gradient from 15% MeCN to 95% MeCN over 20 minutes with a flow of 5 mL/min. The collections were concentrated on a rotarvap (Büchi V-855/R-215) and dried down under N\(_2\) (g) to yield 0.6 mg of demethylcichonidulol.

**Isolation of 4-hydroxy-3,6-dimethyl-2-pyrone**

4-hydroxy-3,6-dimethyl-2-pyrone was present in an earlier fraction than demethylcichonidulol in the purification procedure described previously. The collections were concentrated on a rotarvap (Büchi V-855/R-215) and dried down under N\(_2\) (g) to yield 1.0 mg of cichorine.

**NMR studies and structure elucidation:** NMR spectra were acquired in DMSO-d\(_6\) on a Varian Unity Inova 500 MHz spectrometer for cichorine, 3-methylorsellinic acid, nidulol and 4-hydroxy-3,6-dimethyl-2-pyrone and on a Bruker Avance 800 MHz spectrometer at the Danish Instrument Center for NMR Spectroscopy of Biological Macromolecules for demethylcichonidulol, cichonidulol and demethylcichonidulol using standard pulse sequences. The spectra were referenced to this solvent with resonances \( \delta_H = 2.49 \) and \( \delta_C = 39.5 \).

**Characterization data of 3-methylorsellinic acid.** \(^1\)H NMR (500 MHz, DMSO-d\(_6\)): \( \delta \) 13.06 (1H, br. s, 8-OH), 9.98 (1H, s, 4-OH), 6.23 (1H, s, H-5), 2.38 (3H, s, H-7), 1.91 (3H, s, H-9); \(^{13}\)C NMR (125 MHz): \( \delta \) 173.8 (C-8),
Characterization data of cichorine. $^1$H NMR (500 MHz, DMSO-$d_6$): $\delta$ 9.71 (1H, s, 4-OH), 8.41 (1H, s, N-H), 6.81 (1H, s, H-5), 4.40 (2H, s, H-8), 3.84 (3H, s, H-10), 2.05 (3H, s, H-9); $^{13}$C NMR (125 MHz): $\delta$ 169.8 (C-7), 156.3 (C-4), 153.6 (C-2), 131.9 (C-6), 123.1 (C-1), 118.9 (C-3), 102.9 (C-5), 58.9 (C-10), 9.1 (C-9). HRMS (m/z): [M+H]$^+$ calcd. for C$_{10}$H$_{12}$NO$_3$, 194.08117; found, 194.08123.

NMR-table:

<table>
<thead>
<tr>
<th>Number</th>
<th>$^1$H (ppm)</th>
<th>$^{13}$C (ppm)</th>
<th>Multiplicity (J/Hz)</th>
<th>Integral</th>
<th>HMBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>103.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>163.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>107.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>159.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4-OH</td>
<td>9.98</td>
<td>-</td>
<td>s</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.23</td>
<td>109.9</td>
<td>s</td>
<td>1</td>
<td>1, 3, 4, 7, 8, 9</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>139.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>2.38</td>
<td>23.4</td>
<td>s</td>
<td>3</td>
<td>1, 2, 5, 6, 8</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>173.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>1.91</td>
<td>7.9</td>
<td>s</td>
<td>3</td>
<td>2, 3, 4, 6</td>
</tr>
<tr>
<td>8-OH</td>
<td>13.06</td>
<td>-</td>
<td>br. s</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>
Characterization data of nidulol. $^1$H NMR (500 MHz, DMSO-$d_6$): $\delta$ 10.72 (1H, s, 4-OH), 6.73 (1H, s, H-5), 5.17 (2H, s, H-7), 3.87 (3H, s, H-10), 2.05 (3H, s, H-9); $^{13}$C NMR (125 MHz): $\delta$ 168.5 (C-8), 162.7 (C-4), 157.5 (C-2), 148.6 (C-6), 117.5 (C-3), 107.2 (C-1), 103.0 (C-5), 68.2 (C-7), 61.1 (C-10), 8.2 (C-9). HRMS (m/z): [M+H]$^+$ calcd. for C$_{10}$H$_{11}$O$_4$, 195.0652; found, 195.0652.
**Characterization data of demethylnidulol:** $^1$H NMR (800 MHz, DMSO-$d_6$): $\delta$ 5.67 (1H, s, H-5), 5.64 (1H, s, H-3), 4.82 (2H, s, H-7), 3.61 (3H, s, H-10); $^{13}$C NMR (200 MHz): $\delta$ 178.1 (C-8), 168.8 (C-4), 159.6 (C-2), 151.4 (C-6), 103.2 (C-5), 101.1 (C-3), 92.9 (C-1), 66.9 (C-7), 53.9 (C-10). HRMS (m/z): [M+H]$^+$ calcd. for C$_9$H$_{10}$O$_4$, 181.0495; found, 181.0497.

<table>
<thead>
<tr>
<th>Number</th>
<th>$^1$H (ppm)</th>
<th>$^{13}$C (ppm)</th>
<th>Multiplicity (J/Hz)</th>
<th>Integral</th>
<th>HMBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>107.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>157.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>117.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>162.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4-OH</td>
<td>10.72</td>
<td>-</td>
<td>s</td>
<td>1</td>
<td>3, 4</td>
</tr>
<tr>
<td>5</td>
<td>6.73</td>
<td>103.0</td>
<td>s</td>
<td>1</td>
<td>1, 3, 7</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>148.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>5.17</td>
<td>68.2</td>
<td>s</td>
<td>2</td>
<td>1, 6, 8</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>168.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>2.02</td>
<td>8.2</td>
<td>s</td>
<td>3</td>
<td>2, 3, 4</td>
</tr>
<tr>
<td>10</td>
<td>3.87</td>
<td>61.1</td>
<td>s</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>
NMR-table:

<table>
<thead>
<tr>
<th>Number</th>
<th>$^1$H (ppm)</th>
<th>$^{13}$C (ppm)</th>
<th>Multiplicity (J/Hz)</th>
<th>Integral</th>
<th>HMBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>92.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>159.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>5.64</td>
<td>101.1</td>
<td>s</td>
<td>1</td>
<td>5, 6</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>168.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>5.67</td>
<td>103.2</td>
<td>s</td>
<td>1</td>
<td>3, 6</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>151.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>4.82</td>
<td>66.9</td>
<td>s</td>
<td>2</td>
<td>1, 2, 4, 5, 6, 8</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>178.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>3.61</td>
<td>53.9</td>
<td>s</td>
<td>3</td>
<td>2, 3</td>
</tr>
</tbody>
</table>

Characterization data of cichonidulol. $^1$H NMR (800 MHz, DMSO-$d_6$): $\delta$ 9.74-10.11 (1H, br. s), 7.21 (1H, d, $J$ = 1.8 Hz, H-8'), 6.91 (1H, s, H-5), 6.59 (3H, s, H-5'), 5.17 (1H, d, $J$ = 12.3 Hz, H-7'b), 4.88 (1H, d, $J$ = 12.3 Hz, 7'a), 4.28 (1H, d, $J$ = 16.3 Hz, H-8b), 3.75 (1H, d, $J$ = 16.3 Hz, H-8a), 3.70 (3H, s, H-10), 3.56 (3H, s, H-10'), 2.05 (3H, s, H-9), 1.98 (3H, s, H-9'); $^{13}$C NMR (200 MHz): $\delta$ 167.1 (C-7), 158.1 (C-4'), 156.9 (C-4), 153.8 (C-2'), 153.6 (C-2), 139.5 (C-6'), 131.0 (C-6), 121.3 (C-1), 120.5 (C-3), 116.1 (C-1'), 115.3 (C-3'), 103.4 (C-5), 102.4 (C-5'), 83.5 (C-8'), 71.7 (C-7'), 95.2 (C-10'), 59.0 (C-10), 43.2 (C-8), 9.3 (C-9), 8.6 (C-9'). HRMS (m/z): [M+H]$^+$ calcd. for C$_{20}$H$_{22}$NO$_6$, 372.1442; found, 372.1441.
<table>
<thead>
<tr>
<th>Number</th>
<th>H (ppm)</th>
<th>C (ppm)</th>
<th>Multiplicity (J/Hz)</th>
<th>Integral</th>
<th>HMBC</th>
<th>DQF-COSY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>121.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>153.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>120.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>156.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>6.91</td>
<td>103.4</td>
<td>s</td>
<td>1</td>
<td>1, 4, 7</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>131.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>167.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8a</td>
<td>3.75</td>
<td>43.2</td>
<td>d (16.3)</td>
<td>1</td>
<td>1, 6</td>
<td>8b</td>
</tr>
<tr>
<td>8b</td>
<td>4.28</td>
<td>43.2</td>
<td>d (16.3)</td>
<td>1</td>
<td>1, 6</td>
<td>8a</td>
</tr>
<tr>
<td>9</td>
<td>2.05</td>
<td>9.3</td>
<td>s</td>
<td>3</td>
<td>2, 3, 4</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>3.70</td>
<td>59.0</td>
<td>s</td>
<td>3</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>1'</td>
<td>-</td>
<td>116.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2'</td>
<td>-</td>
<td>153.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3'</td>
<td>-</td>
<td>115.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4'</td>
<td>-</td>
<td>158.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5'</td>
<td>6.59</td>
<td>102.4</td>
<td>s</td>
<td>1</td>
<td>1', 4', 7'</td>
<td>-</td>
</tr>
<tr>
<td>6'</td>
<td>-</td>
<td>139.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7'a</td>
<td>4.88</td>
<td>71.7</td>
<td>d (12.3)</td>
<td>1</td>
<td>1', 6', 8'</td>
<td>7'b</td>
</tr>
<tr>
<td>7'b</td>
<td>5.17</td>
<td>71.7</td>
<td>d (12.3)</td>
<td>1</td>
<td>6'</td>
<td>7'a, 8'</td>
</tr>
<tr>
<td>8'</td>
<td>7.21</td>
<td>83.5</td>
<td>d (1.8)</td>
<td>1</td>
<td>1', 7', 8</td>
<td>7'b</td>
</tr>
</tbody>
</table>
Characterization data of demethylcichonidulol. $^1$H NMR (500 MHz, DMSO-$d_6$): $\delta$ HRMS ($m/z$): 6.97 (1H, d, J = 2.2, H-8'), 6.90 (1H, s, H-5), 6.35 (1H, s, H-3' or H-5'), 6.34 (1H, s, H-3' or H-5'), 5.16 (1H, dd, J = 12.8, 2.2 Hz, H-7'a), 4.86 (1H, d, J = 12.8 Hz, H-7'b), 4.25 (1H, d, J = 16.0 Hz, H-8a), 3.78 (1H, d, J = 16 Hz, H-8b), 3.71 (3H, s, H-10), 3.63 (3H, s, H-10'), 2.05 (3H, s, H-9). HRMS ($m/z$): $\left[M+H\right]^+$ calcd. for $C_{19}H_{20}O_{6}$, 358.1285; found 358.1283.

Characterization data of 4-hydroxy-3,6-dimethyl-2-pyrone. $^1$H NMR (500 MHz, DMSO-$d_6$): $\delta$ 5.46 (1H, s, H-5), 1.93 (3H, s, 3-CH$_3$), 1.56 (3H, s, 6-CH$_3$); $^{13}$C NMR (125 MHz): $\delta$ 174.9 (C-4), 166.0 (C-2), 156.0 (C-6), 106.6 (C-5), 91.0 (C-3), 19.0 (6-CH$_3$), 9.1 (3-CH$_3$). HRMS ($m/z$): $\left[M+H\right]^+$ calcd. for $C_{7}H_{9}O_{3}$, 141.0546; found, 141.0543.

NMR-table:

<table>
<thead>
<tr>
<th>Number</th>
<th>$^1$H (ppm)</th>
<th>$^{13}$C (ppm)</th>
<th>Multiplicity (J/Hz)</th>
<th>Integral</th>
<th>HMBC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-</td>
<td>166.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>91.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>174.9</td>
</tr>
<tr>
<td>5</td>
<td>5.46</td>
<td>106.0</td>
<td>s</td>
<td>1</td>
<td>3, 6, 6-CH₃</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>156.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-CH₃</td>
<td>1.93</td>
<td>19.1</td>
<td>s</td>
<td>3</td>
<td>5, 6</td>
</tr>
<tr>
<td>3-CH₃</td>
<td>1.56</td>
<td>9.1</td>
<td>s</td>
<td>3</td>
<td>3, 4, 5</td>
</tr>
</tbody>
</table>

$^1$H NMR spectrum 3-methylorsellinic acid in DMSO-$d_6$
$^1$H NMR spectrum cichorine in DMSO-$d_6$

$^1$H NMR spectrum nidulol in DMSO-$d_6$
$^1$H NMR spectrum demethylnidulol in DMSO-$d_6$

$^1$H NMR spectrum cichonidulol in DMSO-$d_6$
$^1$H NMR spectrum demethylcichonidulol in DMSO-$d_6$

$^2$H NMR spectrum 4-hydroxy-3,6-dimethyl-2-pyrone in DMSO-$d_6$
References


Nørholm 2010


Paper 6

“A combined genetic and multi medium approach reveals new secondary metabolites of *Aspergillus nidulans*”

ML Klejnstrup, MT Nielsen, JB Nielsen, JC Frisvad, CH Gotfredsen, UH Mortensen and TO Larsen

Rough draft
A combined genetic and multi medium approach reveals new secondary metabolites of
*Aspergillus nidulans*

Marie Louise Klejstrup¹, Morten Thrane Nielsen¹,², Jakob Blæsbjerg Nielsen¹, Jens Christian
Frisvad¹, Charlotte Held Gotfredsen³, Uffe Hasbro Mortensen¹, Thomas Ostenfeld Larsen¹

**Affiliation**

¹ Center for Microbial Biotechnology, Department of Systems Biology, Technical University of Denmark,
2800 Kgs Lyngby, Denmark.

² Current address: Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark,
Fremtidsvej 3, 2970 Hoersholm, Denmark.

³ Department of Chemistry, Technical University of Denmark, 2800 Kgs Lyngby, Denmark.

**Introduction**

The construction of a deletion library of 32 putative PKS encoding genes constructed in our group by
Nielsen *et al.* (Nielsen 2011) made us look into the production of polyketides in other strains of *A. nidulans*
to identify candidates to isolate, structure elucidate and, if possible, link to their synthase genes. Linking of
metabolites present in small quantities in the reference strain could be possible due to the possibility of
comparison with an isolated standard. In the search of new secondary metabolites of *A. nidulans* we
investigated the IBT culture collection at CMB. The strain chosen for the study IBT24909 produced, among
others, four metabolites of which three, asperugin B, nidubenzal A, and B were isolated and structure
elucidated and the last, asperugin A were putatively identified through LC-MS analysis. These metabolites
were detected in trace amounts in the reference strains of the above mentioned deletion study and
candidates were chosen for a deletion study in a cclAΔ strain which also produced these metabolites.
Asperugin A and B were linked to the PKS encoding gene AN3230, pkfA, which have been shown by Ahuja
and co-workers (Ahuja 2012) to catalyze the production of orsellinaldehyd. The biosynthetic origin of the
novel metabolites nidubenzal A and B remains to be clarified.

**Materials and methods**

**Strains**

Nine strains of *A. nidulans* strains were used in a cclA deletion series, see supplementary table 1.
Protoplasting and genetic transformation was as described in Johnstone *et al.* (Johnstone 1985) and Nielsen
*et al.* (Nielsen 2006). IBT28738 was used as reference strain for chemical analysis (argB2, pyrG89, veA1,
nkuA-tr5:AFpyrG) (Nielsen 2008). *Escherichia coli* strain DH5α was used for cloning. The IBT24909 was used
for isolation of aspuridone B and nidubenzal A and B.
Media

Fungal minimal medium (MM) was as described in (Cove 1966), but with 1% glucose, 10 mM NaNO₃ and 2% agar (Sorbigel, Hendaye, France), when necessary supplemented with 10 mM uridine (Uri), 10 mM uracil (Ura) and 4 mM L-arginine (Arg). For transformation medium, glucose was replaced with 1 M sucrose. Elimination of the Aspergillus fumigatus pyrG marker was done on 5-FOA (1.3 mg/ml) supplemented MM, as previously described by Nielsen et al. 2006 (Nielsen 2006). All chemicals were from Sigma-Aldrich (St. Louis, Mo, USA) unless otherwise stated. CYA, CYAs, YES, RT, CY20, RTO, and YE were prepared as described in Nielsen et al. (Nielsen 2011).

PCR and DNA substrate construction

The Aspergillus genome database (AspGD) was used for retrieving gene sequence information. Oligonucleotides were produced by IDT (Integrated DNA Technologies Inc, Coralville, IA, USA) or Sigma-Aldrich, see supplementary table 2 for list. The PfuX7 polymerase (Nørholm, 2010) was used in all PCR reactions as described in Hansen et al., 2011. Deletion of qutC was based on bipartite gene targeting substrates by fusion-PCR and the recyclable AFpyrG marker with direct repeats (Nielsen et al., 2006; 2008), whereas the cclΔ DNA substrate for the deletion series was based on the USER technology. The pU2002A vector designed to incorporate gene-targeting sequences flanking the AFpyrG marker with direct repeats was digested with PacI/Nt.BbvCI prior to mixing with appropriate PCR fragments and USER cloning. The cclΔ deletion substrate was linearized prior to transformation with Swal (1U/μl) according to manufacturer’s protocol. USER mix, restriction and nicking enzymes were from New England Biolabs, Ipswich, MA, USA.

Chemical analysis

All cclΔ strains were grown as three point inoculations for 7 days at 37 degrees in the dark on solid media. Extraction of metabolites were performed by the agar plug extraction method (Smedsgaard 1997) using three 6mm agar plugs/extract. Extracts were analyzed by UHPLC-DAD and LC-HRMS. UHPLC-DAD analysis was performed on a Dionex RSLC Ultimate 3000 (Dionex, Sunnyvale, CA) equipped with a diode-array detector. Separation was performed at 60°C on a 150 mm × 2.1 mm ID, 2.6 μm Kinetex C₁₈ column (Phenomenex, Torrence, CA) using a linear water/MeCN (both buffered with 50 ppm tri-fluoroacetic acid (TFA)) gradient starting from 15% MeCN to 100% over 7 min at a flow rate of 0.8 mL min⁻¹. LC-HRMS analysis were made on a MaXis 3G QTof (Bruker Daltonics) coupled to a Dionex Ultimate 3000 UHPLC system equipped with a 100 x 2.0 mm, 2.6 μm, Kinetex C-18 column. The separation column was held at a temperature of 40°C and a gradient system composed of A: 20 mM formic acid in water, and B: 20 mM formic acid in acetonitrile was used. The flow was 0.4 ml/min, 85% A graduating to 100% B in 0-10 min, 100% B 10-13 min, 85% A 13.1-15 min. For calibration, a mass spectrum of sodium formate was recorded at the beginning of each chromatogram using the divert valve (0.3-0.4 min). Samples were analyzed both in positive and negative ionization mode. De-replication of induced compounds were performed by comparison of accurate mass to the metabolite database Antibase2010 (Laatsch 2010) as well as comparison of UV spectra to published data.

The IBT24909 extract was analyzed by a HPLC-DAD-HRMS on a 5 cm x 3 μm, Luna C₁₈ (2) column (Phenomenex) using a water-acetonitrile gradient from 15 % CH₃CN to 100 % over 20 minutes (20 mM
formic acid). LC-DAD-MS analysis was performed on a LCT oTOF mass spectrometer (Micromass, Manchester, UK) as in Nielsen et al. (Nielsen 2003, Nielsen 2009)

**Isolation of nidubenzal A and B**

IBT24909 was cultured on 200 CYAs plates for 13 days at 25°C in the dark. The plates were homogenized using a Stomacher and 100 mL ethyl acetate (EtOAc) + 1% formic acid (FA) pr. 10 plates. The extract was filtered after 1 hour and the remaining broth was extracted with EtOAc + 1% FA for 24 hours. The extract was filtered and the two fractions pooled and dried down on a freeze drier. The extract was separated on a KP-C18-HS SNAP column using a Biotage Isolera One (Biotage, Uppsala, Sweden) using a gradient of 10% MeOH (H₂O-Milli-Q (Millipore, MA, USA)) to 100% over 15 CVs (column volumes) resulting in 27 fractions. Asperugin B was present in fractions 20 to 23 which were pooled and purified on a Waters HPLC W600/996PDA (Milford, MA, USA) using a RP column (Phenomenex Luna C18(2), 250 x 10 mm, 5 μm, Torrance, CA, USA) using a gradient of 80% MeCN (H₂O – Milli-Q, Millipore, MA, USA)) to 100 % over 20 minutes with a flow of 4 mL/min. The collections were concentrated on a rotavarp (Büchi V-855/R-215) and dried down under N₂(g) to yield 3.7 mg of nidubenzal A/B.

**Isolation of asperugin B**

Asperugin was collected in a later fraction than nidubenzal A/B in the last purification step described above. The collections were concentrated on a rotavarp (Büchi V-855/R-215) and dried down under N₂(g) to yield 4.3 mg of asperugin.

**NMR studies and structure elucidation**

NMR spectra were acquired in DMSO-d₆ on a Varian Unity Inova 500 MHz spectrometer for asperugin B and on a Bruker Avance 800 MHz spectrometer at the Danish Instrument Center for NMR Spectroscopy of Biological Macromolecules for nidubenzal A/B using standard pulse sequences. The spectra were referenced to this solvent with resonances δH = 2.49 and δC = 39.5.

**Characterization data of nidubenzal A**

1H NMR (800 MHz, DMSO-d₆): δ 9.72 (1H, s, H-7), 7.38 (1H, m, H-6), 7.36 (1H, m, H-2), 6.92 (1H, d, J = 7.9 Hz, H-5), 5.42 (1H, t, J = 6.2 Hz, H-2'), 5.03 (1H, m, H-6), 5.03 (1H, m, H-10'), 4.61 (2H, d, J = 6.2 Hz, H-1'), 2.06 (2H, m, H-5'), 2.03 (2H, m, H-4'), 1.98 (2H, m, H-9'), 1.89 (2H, t, J = 7.3 Hz, H-8'), 1.71 (3H, s, H-15'), 1.61 (3H, s, H-12'), 1.54 (3H, s, H-14'), 1.53 (3H, s, H-13'); 13C NMR (200 MHz): δ 190.9 (C-7), 153.7 (C-4), 147.1 (C-3), 140.0 (C-3'), 130.4 (C-11'), 128.2 (C-1), 125.8 (C-6), 123.8 (C-10'), 119.5 (C-2'), 115.4 (C-5), 112.3 (C-2), 64.9 (C-1'), 38.7 (C-4'), 25.4 (C-5'), 25.3 (C-12'), 17.3 (C-13'), 16.1 (15'), 15.5 (C-14'). HRMS (m/z): [M+H]+ calcd. for C₂₄H₃₂O₅, 343.2268; found, 343.2267.

![Nidubenzal A](image-url)
<table>
<thead>
<tr>
<th>Number</th>
<th>$^1$H (ppm)</th>
<th>$^{13}$C (ppm)</th>
<th>Multiplicity</th>
<th>Integral</th>
<th>HMBC</th>
<th>DQF-COSY</th>
<th>H2BC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>128.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7.36</td>
<td>112.3</td>
<td>m</td>
<td>1</td>
<td>3, 4, 6, 7</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>147.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>153.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.92</td>
<td>115.4</td>
<td>d (J = 7.9 Hz)</td>
<td>1</td>
<td>1, 3, 4</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7.38</td>
<td>125.8</td>
<td>m</td>
<td>1</td>
<td>2, 4, 7</td>
<td>2, 5</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>9.72</td>
<td>190.9</td>
<td>s</td>
<td>1</td>
<td>1, 2, 3, 6</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>1'</td>
<td>4.61</td>
<td>64.9</td>
<td>d (J = 6.2 Hz)</td>
<td>2</td>
<td>3, 2', 3', 4', 5', 15'</td>
<td>2'</td>
<td></td>
</tr>
<tr>
<td>2'</td>
<td>5.42</td>
<td>119.5</td>
<td>t (J = 6.2 Hz)</td>
<td>1</td>
<td>1', 4', 5', 15'</td>
<td>1'</td>
<td></td>
</tr>
<tr>
<td>3'</td>
<td>-</td>
<td>140.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4'</td>
<td>2.03</td>
<td>38.7</td>
<td>m</td>
<td>2</td>
<td>2', 3', 5', 15'</td>
<td>5'</td>
<td></td>
</tr>
<tr>
<td>5'</td>
<td>2.06</td>
<td>25.4</td>
<td>m</td>
<td>2</td>
<td>3', 4', 6', 7'</td>
<td>4', 6'</td>
<td></td>
</tr>
<tr>
<td>6'</td>
<td>5.06</td>
<td>123.3</td>
<td>m</td>
<td>1</td>
<td>5', 8', 14'</td>
<td>5'</td>
<td></td>
</tr>
<tr>
<td>7'</td>
<td>-</td>
<td>134.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>8'</td>
<td>1.89</td>
<td>39.0</td>
<td>t (J = 7.3 Hz)</td>
<td>2</td>
<td>6', 7', 9', 14'</td>
<td>9'</td>
<td></td>
</tr>
<tr>
<td>9'</td>
<td>1.98</td>
<td>26.0</td>
<td>m</td>
<td>2</td>
<td>7', 8', 10', 11', 8', 10'</td>
<td>8', 10'</td>
<td></td>
</tr>
<tr>
<td>10'</td>
<td>5.03</td>
<td>128.8</td>
<td>m</td>
<td>1</td>
<td>12', 13'</td>
<td>9'</td>
<td></td>
</tr>
<tr>
<td>11'</td>
<td>-</td>
<td>130.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>12'</td>
<td>1.61</td>
<td>25.3</td>
<td>s</td>
<td>3</td>
<td>8', 9', 10', 11', 13'</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>13'</td>
<td>1.53</td>
<td>17.3</td>
<td>s</td>
<td>3</td>
<td>10', 11', 12'</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>14'</td>
<td>1.54</td>
<td>15.5</td>
<td>s</td>
<td>3</td>
<td>6', 7', 8'</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>15'</td>
<td>1.71</td>
<td>16.1</td>
<td>s</td>
<td>3</td>
<td>1', 2', 3', 4'</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

**Characterization data of nidubenzal B**

$^1$H NMR (800 MHz, DMSO-d$_6$): δ 9.72 (1H, s, H-7), 7.35 (1H, d, J = 1.7 Hz, H-6), 7.24 (1H, d, J = 1.8 Hz, H-2), 7.09 (1H, d, J = 8.2 Hz, H-5), 5.42 (1H, t, J = 6.2 Hz, H-2'), 5.06 (1H, m, H-6'), 5.03 (1H, m, H-10'), 4.66 (2H, d, J = 6.2 Hz, H-1'), 2.06 (2H, m, H-5'), 2.03 (2H, m, H-4'), 1.98 (2H, m, H-9'), 1.89 (2H, t, J = 7.3 Hz, H-8'), 1.71 (3H, s, H-15'), 1.61 (3H, s, H-12'), 1.54 (3H, s, H-14'), 1.53 (3H, s, H-13'); $^{13}$C NMR (200 MHz): δ 191.4 (C-7), 152.4 (C-4), 147.0 (C-3), 140.0 (C-3'), 130.4 (C-11'), 129.6 (C-1), 124.0 (C-6), 123.8 (C-10'), 119.5 (C-2'), 115.4
(C-5), 112.3 (C-2), 64.9 (C-1'), 38.7 (C-4'), 25.4 (C-5'), 25.3 (C-12'), 17.3 (C-13'), 16.1 (15'), 15.5 (C-14'). HRMS (m/z): [M+H]^+ calcd. for C_{24}H_{32}O_{5}, 343.2268; found, 343.2267.

<table>
<thead>
<tr>
<th>Number</th>
<th>ℎ (ppm)</th>
<th>ℜ (ppm)</th>
<th>Multiplicity (J/Hz)</th>
<th>Integral</th>
<th>HMBC</th>
<th>DQF-COSY</th>
<th>H2BC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>129.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>7.24</td>
<td>113.4</td>
<td>d (J = 1.8 Hz)</td>
<td>1</td>
<td>3, 4, 6, 7</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>147.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>152.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>7.09</td>
<td>112.6</td>
<td>d (J = 8.2 Hz)</td>
<td>1</td>
<td>1, 3, 4</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>7.35</td>
<td>124.0</td>
<td>m</td>
<td>1</td>
<td>2, 4, 7</td>
<td>2, 5</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>9.74</td>
<td>191.4</td>
<td>s</td>
<td>1</td>
<td>1, 2, 3, 6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1'</td>
<td>4.66</td>
<td>65.0</td>
<td>d (J = 6.2 Hz)</td>
<td>2</td>
<td>4, 2', 3', 4'</td>
<td>2'</td>
<td>2'</td>
</tr>
<tr>
<td>2'</td>
<td>5.42</td>
<td>119.5</td>
<td>t (J = 6.2 Hz)</td>
<td>1</td>
<td>1', 4', 5', 15'</td>
<td>1'</td>
<td>1'</td>
</tr>
<tr>
<td>3'</td>
<td>-</td>
<td>140.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4'</td>
<td>2.03</td>
<td>38.7</td>
<td>m</td>
<td>2</td>
<td>2', 3', 5', 15'</td>
<td>5'</td>
<td>5'</td>
</tr>
<tr>
<td>5'</td>
<td>2.06</td>
<td>25.4</td>
<td>m</td>
<td>2</td>
<td>3', 4', 6', 7'</td>
<td>4', 6'</td>
<td>4', 6'</td>
</tr>
<tr>
<td>6'</td>
<td>5.06</td>
<td>123.3</td>
<td>m</td>
<td>1</td>
<td>5', 8', 14'</td>
<td>5'</td>
<td>5'</td>
</tr>
<tr>
<td>7'</td>
<td>-</td>
<td>134.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8'</td>
<td>1.89</td>
<td>39.0</td>
<td>t (J = 7.3 Hz)</td>
<td>2</td>
<td>6', 7', 9', 14'</td>
<td>9'</td>
<td>9'</td>
</tr>
<tr>
<td>9'</td>
<td>1.98</td>
<td>26.0</td>
<td>m</td>
<td>2</td>
<td>7', 8', 10', 11'</td>
<td>8', 10'</td>
<td>8', 10'</td>
</tr>
<tr>
<td>10'</td>
<td>5.03</td>
<td>128.8</td>
<td>m</td>
<td>1</td>
<td>12', 13'</td>
<td>9'</td>
<td>9'</td>
</tr>
<tr>
<td>11'</td>
<td>-</td>
<td>130.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12'</td>
<td>1.61</td>
<td>25.3</td>
<td>s</td>
<td>3</td>
<td>8', 9', 10', 11', 13'</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13'</td>
<td>1.53</td>
<td>17.3</td>
<td>s</td>
<td>3</td>
<td>10', 11', 12'</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14'</td>
<td>1.54</td>
<td>15.5</td>
<td>s</td>
<td>3</td>
<td>6', 7', 8'</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15'</td>
<td>1.71</td>
<td>16.1</td>
<td>s</td>
<td>3</td>
<td>1', 2', 3', 4'</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Characterization data of asperugin A

HRMS (m/z) [M+H]^+ calcd. for C_{24}H_{32}O_{5}, 401.2322; found, 401.2314.

Characterization data of asperugin B

$^1$H NMR (500 MHz, DMSO-d$_6$): 10.49 (1H, s, H-7), 10.16 (1H, s, H-8), 6.99 (1H, s, H-3), 5.46 (1H, t, J = 7.1 Hz, H-2'), 5.03 (1H, m, H-10'), 5.02 (1H, m, H-6'), 4.62 (2H, d, J = 7.1, H-1'), 1.99 (2H, m, H-5'), 1.98 (2H, m, H-9'), 1.95 (2H, m, H-4'), 1.89 (2H, t, J = 7.6 Hz, H-8'), 1.62 (3H, s, H-13'), 1.58 (3H, s, H-15'), 1.54 (3H, s, H-12'), 1.52 (3H, s, H-14'); $^{13}$C NMR (125 MHz): δ 193.1 (C-7), 192.5 (C-8), 157.6 (C-6), 157.0 (C-4), 141.4 (C-3'), 136.9 (C-5), 134.6 (C-7'), 133.3 (C-2), 124.1 (C-10'), 123.6 (C-6'), 119.9 (C-2'), 113.7 (C-1), 112.6 (C-3), 68.0 (C-1'), 39.2 (C-8'), 39.1 (C-4'), 26.2 (C-9'), 25.8 (C-5'), 25.5 (C-13'), 17.6 (C-12'), 16.1 (C-15'), 15.7 (C-14'). HRMS (m/z) [M+H]^+ calcd. for C_{23}H_{30}O_{5}, 387.2166; found, 387.2154.

<table>
<thead>
<tr>
<th>Number</th>
<th>$^1$H (ppm)</th>
<th>$^{13}$C (ppm)</th>
<th>Multiplicity (J/Hz)</th>
<th>Integral</th>
<th>HMBC</th>
<th>DQF-COSY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>113.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>133.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>6.99</td>
<td>112.6</td>
<td>s</td>
<td>1</td>
<td>1, 2, 4, 5, 8</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>157.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>136.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>157.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>10.49</td>
<td>193.1</td>
<td>s</td>
<td>1</td>
<td>1, 2, 5, 6</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>10.16</td>
<td>195.5</td>
<td>s</td>
<td>1</td>
<td>1, 2, 5, 6</td>
<td>-</td>
</tr>
<tr>
<td>1'</td>
<td>4.62</td>
<td>68.0</td>
<td>d (J = 7.1 Hz)</td>
<td>2</td>
<td>5, 2', 3', 4', 5'</td>
<td>2'</td>
</tr>
<tr>
<td>2'</td>
<td>5.46</td>
<td>119.9</td>
<td>m</td>
<td>1</td>
<td>1', 4', 15'</td>
<td>1'</td>
</tr>
<tr>
<td>3'</td>
<td>-</td>
<td>141.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4'</td>
<td>1.95</td>
<td>39.1</td>
<td>m</td>
<td>2</td>
<td>2', 3', 5', 6', 15'</td>
<td>5'</td>
</tr>
<tr>
<td>5'</td>
<td>1.99</td>
<td>25.8</td>
<td>m</td>
<td>2</td>
<td>4', 6'</td>
<td>-</td>
</tr>
<tr>
<td>6'</td>
<td>5.02</td>
<td>123.6</td>
<td>m</td>
<td>1</td>
<td>5', 8', 14'</td>
<td>5'</td>
</tr>
<tr>
<td>7'</td>
<td>-</td>
<td>134.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Results and discussion

The strain chosen for this study was IBT24909 which was cultivated on 200 plates of CYAs for 13 days at 25°C for optimal production of 1, 2, and 3. The LC-DAD chromatogram at 210 nm of the extract can be seen in figure 1.

![Figure 1](image)

Figure 1 LC-DAD chromatogram (210 nm) of 200 plate extract of IBT24909 cultivated on CYAs for 13 days. The two peaks 1 and 2 were identified as potential PKs. The large peak at 10.5 minutes is sterigmatocystin.

After inspection of the chromatogram two rather large peaks, 1 and 2, were identified as putative PKs based on their UV-spectra which had absorption maxima at wavelengths over 300 nm indicating that the metabolite or part of it contained a highly conjugated system.
The metabolites were isolated and structure elucidated by 2D NMR (DQF-COSY, HSQC, HMBC, H2BC and NOESY).

![Chemical structures of nidubenzal A and B, asperugin A, B, C, and aspernidine A and B](image)

Figure 2: Structures of nidubenzal A and B (1), asperugin A, B (2), and C, and aspernidine A and B.

NMR analysis of A revealed that two isomers, that we named nidubenzal A and B, figure 2, were present in the solution in a 5:1 ratio. The difference of the chemical shift values of the two isomers were small and only visible for H1'/C1' and in the aromatic moiety of the molecule.

The structure of the farnesyl side chain was deduced based on DQF-COSY, HSQC, HMBC and H2BC correlations and comparison with literature data from the related structures aspernidine A and B (Scherlach 2010). The attachment of the farnesyl side chain to the two isomers was based on HMBC correlations of H1' to C3 and C4 for nidubenzal A and B, respectively. Nidubenzal A has been reported in a synthetic screening library (Aurora Fine Chemicals LLC); however, neither nidubenzal A nor B has previously been reported from nature.

The structure of 2 was also solved through 2D-NMR analysis (DQF-COSY, HSQC, and HMBC) which showed that it was asperugin B a metabolite previously isolated from the related fungus A. rugulosus (Ballantine 1967). Two other asperugin, A and C, have been isolated from A. rugulosus (Ballantine 1964, 1965, and 1971) and through LC-TOF-MS-analysis we tentatively identified asperugin A (3) but not asperugin C in the A. nidulans IBT24909 extract. We have not yet been able to isolate asperugin A in quantities sufficient for NMR-analysis. In 2005 An et al. (An 2005) reported production of asperugin A and B (also with B as the major product), as well as violaceol, in a strain where cosmid-size DNA from slow growing fungi was cloned and introduced into A. nidulans. During the course of this study Scherlach et al. reported the characterization of aspernidine A and B isolated from A. nidulans (Scherlach 2010); however, we were not able to trace these metabolites through LC-TOF-MS analysis of our extract.
All the isolated metabolites, nidubenzal A and B, asperugin A-C and the related aspernidines, seem to be hybrid metabolites which consist of a farnesyl side chain and an aromatic moiety, which may be of PK origin. The aromatic nature of the metabolites indicated that the pathway could be initiated with formation of a NR-PK catalyzed by a NR-PKS. The products of the NR-PKS could be orsellinic acid or even more likely orsellinaldehyde, figure 3, which have, just recently, been shown to be a metabolite of A. nidulans from the AN3230 (PkfA) NR-PKS (Ahuja 2012). The only difference between the aromatic part of asperugin A and B, and aspernidine A and B, are an O-methylation indicating that the two metabolites in each group are derived from the same biosynthetic pathway. Scherlach et al. (Scherlach 2010) suggests that the aspernidines are formed by either orsellinic acid or the corresponding aldehyde which traps ammonia nitrogen in an intermediary aldehyde, where oxidoreduction and condensation steps leads to the isoindolone moiety. The precursor of the aromatic moiety of asperugin C could also be orsellinaldehyde which, in a shunt pathway, will be methylated instead of oxidated. The new metabolites nidubenzal A and B differs in the aromatic moiety which containing seven carbons instead of the eight of the backbone of the other metabolites. This could indicate a different biosynthetic pathway or a decarboxylation from a C8-PK product. We speculated whether these metabolites might be a product of the shikimic acid pathway since the substitution pattern resembles protocatechuic acid, figure 3, a shunt product from the biosynthesis of shikimic acid, differing only by the aldehyde in the place of the carboxylic acid.

**Figure 3** Structures of orsellinaldehyde, protocatechuic acid and 2-ethyl-4,6-dihydroxy-3,5-dimethylbenzaldehyde

Since the metabolites might be of PK origin we went back to the reference strain of the PKS deletion library (Nielsen 2011) to see if we could trace the metabolites through extracted ion chromatograms (EICs) on any of the media used in the study.

Nidubenzal A and B were present in small quantities on the CYAs medium so we went back to the deletion library to see whether we could identify a possible PKS to link to these metabolites. In two of the PKS depletions, ausAΔ and AN11191Δ, the production seemed to have been abolished; however, due to the small quantities of metabolites present in the extract it was, in some cases, difficult to determine the presence of the metabolites.

Asperugin A could not be detected in any of the strains whereas asperugin B could be traced. The same procedure for identification of putative PKSs as for the nidubenzals was done and six PKS candidates were identified, AN0523 (pkdA), AN11191, AN3230 (pkfA), ausA, AN6791, and AN9005. It has recently been reported that AN3230 produces orsellinaldehyde and AN0523 2-ethyl-4,6-dihydroxy-3,5-dimethylbenzaldehyde, figure 4.37 (Ahuja 2012). ausA, pkfA and pkdA are NR-PKSs whereas AN11191, AN6791 and AN9005 are PR-PKSs (Sanchez 2012). These results gave several possible genes so we were not able to link the metabolites to their synthase genes.
At the same time as this study we were also working on a strain where the \textit{cclA} gene had been deleted which activated the production of the asperugins and nidubenzals. It had already been shown that this deletion strain increased the production of secondary metabolites in \textit{A. nidulans} (Bok 2009). We wanted to investigate whether challenging the \textit{cclA}Δ strain on complex media could activate even more secondary metabolite production.

\textbf{Figure 4} The appearance of the reference strain compared to the \textit{cclA}Δ strain cultivated for seven days at 37°C on RTO, RT, CY20, CYA, YES, YE, and MM medium.

Figure 4 show the phenotypical appearance of the \textit{cclA}Δ strain compared to the reference strain when grown on the same eight media as used initially in the deletion study. It is evident that the fungus was affected by the deletion and as the reference strain it responded differently to the different medium conditions. Generally it seems that the \textit{cclA}Δ strains grow slower than the reference. To analyze the metabolite production micro-extraction was performed (Frisvad 1987, Smedsgaard 1997) and the extracts were analyzed by both UHPLC-DAD and UHPLC-TOF-MS. The UHPLC analysis revealed a diverse and different production of secondary metabolites on the eight media compared to the reference strain, figure 5.
Figure 5 UHPLC-DAD chromatograms (210 nm) of micro-extracts of the cclΔ strain cultivated on eight different media compositions for seven days at 37°C compared to the reference strain. From top to bottom RTO, RT, CY20, CYAs, CYA, YES, YE and MM.*This peak corresponds to an internal standard of chloroamphenicol added to the extracts. The
concentration of chloroamphenicol is the same in all the samples. Metabolites were identified through comparison with standards or based on exact mass and UV-data.

In all of the media except YE and CYAs there seemed to be an activation of secondary metabolite production. It was especially noticeable that many of the metabolites produced in the cclAΔ strain are derived from the orsellinic acid and monodictyphenone pathway, which was also the gene clusters activated by Bok et al. (Bok 2009). Arugosin A which was linked to the monodictyphenone pathway in section 4.3.1.2 seemed to disappear in the cclAΔ strain whereas intermediates in this pathway, for example emodin, seemed to accumulate. Most noticeable for our study, and what is focused on in this section, was the increased production of asperugin A and B as well as nidubenzal A and B on several of the media; YES, CY20, CYAs, CYA, RTO and MM, figure 6.

The increased production of asperugin A and B, and nidubenzal A and B in this strain allowed for creating double deletions of cclA and the putative candidate PKS genes in order to understand the changes in production of the metabolites. Therefore cclA was deleted in the respective PKS gene deletions identified previously. We included both the NR- and PR-PKSs in the study since it seemed like they might have an effect on asperugin production. Since it was originally thought that these products could be derived from orsellinic acid, and due to the activation of several orsellinic acid derived metabolites in the cclAΔ strain, a deletion of cclA in orsAΔ was also created to see if there were an effect on the production of the nidubenzals and asperugins. As described previously we wondered whether the nidubenzal A and B could be shikimic derived so we also made a deletion mutant of the qutC gene which has been shown to encode for an enzyme, dehydroshikimic acid dehydratase that catalyzes the production of protocatechius acid from 3-dehydroshikimic acid in A. nidulans (Lamb 1992).
The double deletion mutants made were analyzed on both UHPLC-DAD and UHPLC-TOF-MS on all eight media. As seen in figure 7 nidubenzal A and B were present in all the deletion strains and on the different media; however, spore-PCRs analysis of the AN11191ΔcclΔ strain showed presence WT AN11191 nuclei indicating this strain was not suited for analysis. A new double deletion based on the original confirmed AN11191Δ strain is in progress of being constructed. As the qutCΔcclΔ strain also produced nidubenzal A and B it indicates that they are derived from either AN11191 or an alternative route to protocatechuic acid exists in A. nidulans.

![Figure 7](image7.png)

**Figure 7** UHPLC chromatograms (210 nm) of micro-extracts of the qutCΔcclΔ, ausΔcclΔ, AN11191ΔcclΔ and cclΔ strains after cultivation for seven days at 37°C on CY20 medium. The other parts of the chromatogram were also identical.

![Figure 8](image8.png)

**Figure 8** EIC of a) asperugin A (calc. [M-H]⁻ 401.23225±0.002] and b) asperugin B (calc. [M+H]⁺ 387.21660±0.002]) of micro-extracts of the (from top to bottom) pkdAΔcclΔ, pkfAΔcclΔ, ausAΔcclΔ, AN6791ΔcclΔ, AN9005ΔcclΔ and cclΔ strains after cultivation on CY20 medium for seven days at 37°C.
The extracted ion chromatograms (EIC) of asperugin A and B in the double deletion mutants and the reference strain can be seen in figure 8. As the AN11191ΔcclAΔ strain was not correct it have been removed from the study, however a correct strain will be included in the analysis for the presence of the asperugins. As can be seen on the figure production of asperugin A and B are abolished in the pkfAΔcclAΔ strain and this was consistent on the eight analyzed media. Ahuja et al. (Ahuja 2012) recently showed that overexpression of this PKS with an inducible promoter led to the production of orsellinaldehyde. Further biosynthesis of orsellinaldehyde to the asperugins would require several tailoring enzymes catalyzing oxidation of a methyl to an aldehyde, oxidation of the aromatic ring, a terpene synthase and an O-methyltransferase. As suggested by Scherlach et al. (Scherlach 2010) further modifications to the aromatic moiety of the metabolites could lead to the biosynthesis of the aspernides. This biosynthetic pathway resembles in some ways the proposed biosynthesis of the chicorine gene cluster. Preliminary bioinformatics studies indicated the presence of O-methyltransferase, prenyltransferase, oxidase, cytochrome P450, and monooxidase enzymatic functions surrounding the pkfA gene. Construction of the double deletions of cclA and the surrounding genes to determine the biosynthetic pathway of these metabolites including identifying intermediates are ongoing. In the AN6791ΔcclAΔ strain it appears that the production of asperugin A and B have increased compared to the other strains. This is interesting due to the fact that the deletion was constructed based on analysis of the deletion library where the metabolites were missing in the single deletion strain, AN6791Δ. This phenomena is seen in several media (YES, CY20, CYAs and MM); however, the reason for this apparently increase in production is so far unknown.

![Figure 9](image_url)

**Figure 9** UHPLC chromatograms (210 nm) of micro-extracts of the orsAΔcclAΔ strain compared to the cclAΔ strain on CY20 medium.

Due to the orsellinic acid like nature of asperugins we included a double deletion of orsA and cclA in the study. As seen from figure 9 the deletions did not abolish the production of the asperugins; however, the metabolites which have been shown to be derived from the orsellinic acid biosynthetic pathway were indeed absent in the double deletion strain (Bok 2009, Schroekh 2009, Sanchez 2010, Nahlik 2010).
References:


## Supplementary material

### Supplementary table 1: Strains used for deletions of cclA

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBT29539</td>
<td>argB2, pyrG89, veA1, nkuAΔ</td>
<td>Nielsen et al., 2008</td>
</tr>
<tr>
<td>IBT30996</td>
<td>argB2, pyrG89, veA1, nkuAΔ, AN0523Δ</td>
<td>Nielsen et al., 2011</td>
</tr>
<tr>
<td>-</td>
<td>argB2, pyrG89, veA1, nkuAΔ, qutCΔ</td>
<td>This study</td>
</tr>
<tr>
<td>IBT31005</td>
<td>argB2, pyrG89, veA1, nkuAΔ, pkfAΔ</td>
<td>Nielsen et al., 2011</td>
</tr>
<tr>
<td>IBT31013</td>
<td>argB2, pyrG89, veA1, nkuAΔ, AN6791Δ</td>
<td>Nielsen et al., 2011</td>
</tr>
<tr>
<td>IBT31020</td>
<td>argB2, pyrG89, veA1, nkuAΔ, orsAΔ</td>
<td>Nielsen et al., 2011</td>
</tr>
<tr>
<td>IBT31022</td>
<td>argB2, pyrG89, veA1, nkuAΔ, ausAΔ</td>
<td>Nielsen et al., 2011</td>
</tr>
<tr>
<td>IBT31025</td>
<td>argB2, pyrG89, veA1, nkuAΔ, AN9005Δ</td>
<td>Nielsen et al., 2011</td>
</tr>
<tr>
<td>IBT31000</td>
<td>argB2, pyrG89, veA1, nkuAΔ, AN11191Δ</td>
<td>Nielsen et al., 2011</td>
</tr>
</tbody>
</table>

### Supplementary table 2: Primers used in this study

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANqutC-DL-Up-F</td>
<td>CTTACAGCCTAGTCATTC</td>
<td>Deletion fragment Up</td>
</tr>
<tr>
<td>ANqutC-DL-Up-Rad</td>
<td>gatcccggaattgaccatgGA6CTGATTGAGTTGAAGAC</td>
<td>Deletion fragment Up</td>
</tr>
<tr>
<td>ANqutC-DL-Dw-Fad</td>
<td>aatccagctgaccacagtGGTAATGAGCTGCATGCACG</td>
<td>Deletion fragment Dw</td>
</tr>
<tr>
<td>ANqutC-DL-Dw-R</td>
<td>GATGCTGGGGAAATTTTGGGTG</td>
<td>Deletion fragment Dw</td>
</tr>
<tr>
<td>ANqutC-CHK-Up-F</td>
<td>CACCACTAGAACAACTACATACGG</td>
<td>Strain verification</td>
</tr>
<tr>
<td>ANqutC-CHK-Dw-R</td>
<td>GTGCTGGGGAAATTTTGGGTG</td>
<td>Strain verification</td>
</tr>
<tr>
<td>ANcclA-DL-Up-FU</td>
<td>GGGTTTAUGTGATCAGCTTCGAGGTAG</td>
<td>Deletion fragment Up</td>
</tr>
<tr>
<td>ANcclA-DL-Up-RU</td>
<td>GGACTTAAUCCTTGGCGAGATGGATTTCAC</td>
<td>Deletion fragment Up</td>
</tr>
<tr>
<td>ANcclA-DL-Dw-FU</td>
<td>GCCATTAUCAGCTGGTCATGGATTGG</td>
<td>Deletion fragment Dw</td>
</tr>
<tr>
<td>ANcclA-DL-Dw-RU</td>
<td>GGTCTTAUGACGGAGCCTGAGCTTACG</td>
<td>Deletion fragment Dw</td>
</tr>
<tr>
<td>Primer Name</td>
<td>Sequence</td>
<td>Comment</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>ANclIA-CHK-Up-F</td>
<td>CGTTTCTCTGCTGTCC</td>
<td>Strain verification</td>
</tr>
<tr>
<td>ANclIA-CHK-Dw-R</td>
<td>CGCCTACATCGACTAC</td>
<td>Strain verification</td>
</tr>
</tbody>
</table>

$^1$H NMR of nidubenzal A and B in DMSO-$d_6$
$^1$H NMR of asperugin B in DMSO-$d_6$
Paper 7

“A regulatory protein from Aspergillus niger induces juvenile hormones upon heterologous expression in A. nidulans”

MT Nielsen and ML Klejnstrup (joint 1st author), CH Gotfredsen, MR Andersen, BG Hansen, UH Mortensen, and TO Larsen

Submitted to PLoS ONE – under revision
A regulatory protein from *Aspergillus niger* induces juvenile hormones upon heterologous expression in *A. nidulans*.

**Authors**

Nielsen, M.T.\(^1,2,3\), Klejnstrup, M.L.\(^1,2\), Gotfredsen, C.H. \(^4\), Andersen, M.R.\(^1\), Hansen, B.G.\(^1,5\), Mortensen, U.H.\(^1\)* and Larsen, T. O.\(^1\)*

**Affiliation**

\(^1\) Center for Microbial Biotechnology, Department of Systems Biology, Technical University of Denmark, 2800 Kgs Lyngby, Denmark.

\(^2\) These authors contributed equally to the work.

\(^3\) Current address: Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Fremtidsvej 3, 2970 Hoersholm, Denmark.

\(^4\) Department of Chemistry, Technical University of Denmark, 2800 Kgs Lyngby, Denmark.

\(^5\) Current address: Novozymes A/S, Krogshoejvej 36, 2880 Bagsvaerd, Denmark.

* Authors to whom correspondence should be addressed ([um@bio.dtu.dk](mailto:um@bio.dtu.dk); tol@bio.dtu.dk)

**Abstract**
Fungal natural products are considered a rich, underexplored resource of bioactive compounds. A combined heterologous expression and growth condition variation approach using *Aspergillus nidulans* as model induced insect juvenile hormone III and methyl farnesoate, essential developmental hormones previously unobserved in fungi. The approach is of general relevance in endeavors to uncover metabolites produced by silent gene clusters while the finding of juvenile hormones is of ecological relevance.

**Introduction**

Filamentous fungi are capable of synthesizing a wide range of bioactive molecules important for growth in a complex and competitive ecological niche. These compounds have been found to have diverse pharmaceutical applications. With the release of full genome sequences of several filamentous fungi it became apparent that the number of predicted secondary metabolite synthases by far exceeds the number of known metabolites. These observations hints that specific environmental stimuli are required for induction of the majority of secondary metabolites, thus coining the terms silent and orphan synthases. Despite attempts to identify or mimic these stimuli in order to unravel the secondary metabolism of the model organism *Aspergillus nidulans*, the majority of predicted synthases are still considered silent.

Genetic approaches have been somewhat successful through manipulation of histone methylation or specific regulators. The latter approach is particularly attractive as it is a targeted approach and therefore promises rapid linking of induced metabolites to genes. Moreover, as biosynthetic pathways towards secondary metabolites tend to be clustered in the genome, specific regulators may be identified by genomic co-localization. However, the number of successful applications of this approach is limited, possibly because far from all predicted gene clusters contain regulatory proteins. We decided to investigate whether induction of secondary metabolites could be achieved through heterologous expression of regulatory genes from other filamentous fungi using *A. niger* as test case. A selection of putative pathway specific regulators was constitutively expressed individually from a defined locus. This genetic approach was combined with a screen of several complex media recently demonstrated to influence *A. nidulans*...
secondary metabolism. The combinatorial approach resulted in the identification of one regulatory protein that strongly induced metabolites not previously reported from A. nidulans. Among the induced metabolites were insect juvenile hormone III and methyl farnesoate, sesquiterpene hormones crucial for correct development of insects and crustaceans respectively. To the best of our knowledge, this is the first observation of juvenile hormones in fungi as well as the first example of transcriptional activation of secondary metabolism between distantly related species.

**Materials & Methods:**

**Strains and media.** *Escherichia coli* strain DH5α was used to propagate all plasmids. All *A. niger* genes were amplified from strain ATCC1015. The *A. nidulans* strain NID74 (argBΔ, veA1, pyrG89, nkuΔ) were used as background strain for all transformations as it allows gene targeting with the argB marker due to a complete deletion of the *A. nidulans argB*-open reading frame. NID74 was generated from NID1 using the fusion PCR technique essentially as described previously. NID545 (argBΔ, pyrG89, veA1, nkuΔ) IS1:pgpdA-LacZ-TrpC:argB) were used as reference strain for metabolite analysis. Genotypes of all strains are summarized in supplementary table 2. All *A. nidulans* strains were propagated on solid glucose minimal medium (MM) (1 % glucose, 10 mM NaNO3, 1x salt solution, 2 % agar (Bie & Berntsen, BBB 10030,SO-BI-GEL,Agar)), supplemented with 10 mM uridine (Uri), 10 mM uracil (Ura). Complex media used for chemical analysis were prepared as described by Frisvad and Samson and supplemented with 10 mM Uri and 10 mM Ura:  

- **CYAs** (Yeast extract 5,0 g/l (Biokar, A1202HA), Czapek dox broth 35,0 g/l (Difco, 233810), NaCl 50,0 g/l (J.T Baker, 0277), 1x trace metals, 1,5 % agar(Bie & Berntsen, BBB 10030,SO-BI-GEL,Agar)).  
- **OAT** (Oatmeal 30g/l, 1x trace metals, 1,5 % agar(Bie & Berntsen, BBB 10030,SO-BI-GEL,Agar)).  
- **RTO** (Oatmeal 30g/l, (NH₄)₂HPO₄ 0,4 g/l (Merck, 1126), K₂CO₃ 0,4 g (Merck, 4928), ZnSO₄·7H₂O 0,06 g/l (Merck, 8883), FeSO₄·7H₂O 0,06 g (Merck, 3965), (NH₄)₂SO₄ 0,16 g/l (Merck, 1217), MgCO₃ 0,25 g (Riedel-de Haén, 13118), Tartaric acid (C₄H₆O₆) 2,6 g (Merck, 804), di-Ammonium tartrate (C₄H₆N₂O₆) 2,6 g (Merck,
1222), Glucose (D+) 50.0 g (BHD, 10117), 1x trace metals, 2 % agar (Bie & Berntsen, BBB 10030,SO-BI-GEL,Agar)),

- YES (Yeast extract 20 g/l (Biokar, A1202HA), Sucrose 150 g/l (Fluka, 84100), MgSO₄·7H₂O 0.5 g/l (Merck, 5886), 1x trace metals, 2 % agar (Bie & Berntsen, BBB 10030,SO-BI-GEL,Agar))

**PCR, USER cloning and A. nidulans strain construction.** USER cloning compatible PCR products were amplified with 30 PCR cycles in 50 μl reaction mixtures using proof-reading PfuX7 polymerase. USER vectors are denoted according to the nomenclature introduced by Hansen et al. Putative *A. niger* genes were USER cloned into pU1111-IS1 and transformed into *A. nidulans* as described previously. In order to generate the NIDS45 reference strain the *E. coli lacZ* was cloned into a pU1014-IS1 vector generating pU1011-IS1:*LacZ* which was transformed to an pU1110-IS1-*LacZ* vector by insertion of *A. nidulans gpdA* promoter in the AsiSI/Nb.BtsI cassette. DNA sequences of *LacZ* and all inserted *A. niger* ORFs were verified by sequencing of the corresponding expression plasmids. Gene targeting events were verified in all *A. nidulans* transformants by analytical PCR as described previously. Supplementary table 3 summarizes the PCR primers used in this study.

**Chemical characterization of mutant strains by LC-DAD-MS:** All strains were grown as three point inoculations for 7 days at 37 degrees in the dark on solid media. Extraction of metabolites were performed by the agar plug extraction method using three 6mm agar plugs/extract. Extracts were analyzed by UHPLC-DAD and LC-HRMS. UHPLC-DAD analysis was performed on a Dionex RSLC Ultimate 3000 (Dionex, Sunnyvale, CA) equipped with a diode-array detector. Separation was performed at 60°C on a 150 mm × 2.1 mm ID, 2.6 μm Kinetex C₁₈ column (Phenomenex, Torrence, CA) using a linear water/MeCN (both buffered with 50 ppm tri-fluoroacetic acid (TFA)) gradient starting from 15% MeCN to 100% over 7 min at a flow rate of 0.8 mL min⁻¹. LC-HRMS analysis were made on a MaXis 3G QTOF (Bruker Daltronics) coupled to a Dionex Ultimate 3000 UHPLC system equipped with a 100 x 2.0 mm, 2.6 μm, Kinetex C-18 column. The separation
column was held at a temperature of 40°C and a gradient system composed of A: 20 mM formic acid in
water, and B: 20 mM formic acid in acetonitrile was used. The flow was 0.4 ml/min, 85% A graduating to
100% B in 0-10 min, 100% B 10-13 min, 85% A 13.1-15 min. For calibration, a mass spectrum of sodium
formate was recorded at the beginning of each chromatogram using the divert valve (0.3-0.4 min). Samples
were analyzed both in positive and negative ionization mode. De-replication of induced compounds were
performed by comparison of accurate mass to the metabolite database Antibase2009 as well as
column was held at a temperature of 40°C and a gradient system composed of A: 20 mM formic acid in
water, and B: 20 mM formic acid in acetonitrile was used. The flow was 0.4 ml/min, 85% A graduating to
100% B in 0-10 min, 100% B 10-13 min, 85% A 13.1-15 min. For calibration, a mass spectrum of sodium
formate was recorded at the beginning of each chromatogram using the divert valve (0.3-0.4 min). Samples
were analyzed both in positive and negative ionization mode. De-replication of induced compounds were
performed by comparison of accurate mass to the metabolite database Antibase2009 as well as
column was held at a temperature of 40°C and a gradient system composed of A: 20 mM formic acid in
water, and B: 20 mM formic acid in acetonitrile was used. The flow was 0.4 ml/min, 85% A graduating to
100% B in 0-10 min, 100% B 10-13 min, 85% A 13.1-15 min. For calibration, a mass spectrum of sodium
formate was recorded at the beginning of each chromatogram using the divert valve (0.3-0.4 min). Samples
were analyzed both in positive and negative ionization mode. De-replication of induced compounds were
performed by comparison of accurate mass to the metabolite database Antibase2009 as well as
column was held at a temperature of 40°C and a gradient system composed of A: 20 mM formic acid in
water, and B: 20 mM formic acid in acetonitrile was used. The flow was 0.4 ml/min, 85% A graduating to
100% B in 0-10 min, 100% B 10-13 min, 85% A 13.1-15 min. For calibration, a mass spectrum of sodium
formate was recorded at the beginning of each chromatogram using the divert valve (0.3-0.4 min). Samples
were analyzed both in positive and negative ionization mode. De-replication of induced compounds were
performed by comparison of accurate mass to the metabolite database Antibase2009 as well as
column was held at a temperature of 40°C and a gradient system composed of A: 20 mM formic acid in
water, and B: 20 mM formic acid in acetonitrile was used. The flow was 0.4 ml/min, 85% A graduating to
100% B in 0-10 min, 100% B 10-13 min, 85% A 13.1-15 min. For calibration, a mass spectrum of sodium
formate was recorded at the beginning of each chromatogram using the divert valve (0.3-0.4 min). Samples
were analyzed both in positive and negative ionization mode. De-replication of induced compounds were
performed by comparison of accurate mass to the metabolite database Antibase2009 as well as

Chemical characterization of mutant strains by GC-MS: Volatile metabolites were collected during days 5-7
for the strains inoculated in CYAs. To collect the volatiles, a stainless steel Petri dish lid with a standard 1/4
Swagelock™ replaced the usual lid. This lid possessed a standard 1/4 Swagelok fitting with PTFE insert in
the centre that is used to hold a charcoal tube (SKC, 226-01). The collected volatiles were extracted from
the charcoal tube with 0.3 mL of ether (Sigma Aldrich). The samples were concentrated to approximately
0.1 mL using a nitrogen flow in a GC vial and analysed using a Finnigan Focus GC coupled to a Finnigan
Focus DSQ mass selective detector. The separation of the volatiles was done on a Supelco SLB™-5 MS
capillary column, using He as carrier gas, at 1.2 mL/min. The injection and detection temperature was set to
220 °C. One microlitre of each sample was injected into the GC–MS system. Chromatographic conditions
were set to an initial temperature of 35 °C for 1 min, raised at 6 °C/min to 220 °C and then 20 °C/min to 260
°C for 1 min. The separated compounds were characterized by their mass spectra generated by electron
ionization (EI) at 70 eV at a scan range from m/z 35–300.

Isolation of methyl (2E,6E)-10,11-dihydroxy-3,7,11-trimethyl-2,6-dodecadienoate (X1).
NID477 was cultured on 100 CYAs plates for 7 days at 37°C in the dark. The plates were homogenized using
a Stomacher and 100 mL ethyl acetate (EtOAc) + 1% formic acid (FA) pr. 10 plates. The extract was filtered
after 1 hour and the remaining broth was extracted with EtOAc + 1% FA for 24 hours. The extract was filtered and the two fractions pooled and dried down on a freeze drier. The crude extract were separated into three phases by dissolving it in 9:1 MeOH:H₂O – Milli-Q and extracted into a heptane phase followed by a dicyromethane (DCM) phase. The DCM phase was fractionated with a 10g ISOL Diol column, using 13 steps of stepwise Hexane-dichloromethane-EtOAc-MeOH. X1 was present in the DCM fraction (9.5 mg) and was purified on a Waters HPLC W600/996PDA (Milford, MA, USA) using a RP column (Phenomenex Luna C18(2), 250 x 10 mm, 5 μm, Torrance, CA, USA) using a gradient of 40% MeCN (H₂O – Milli-Q (Millipore, MA, USA)) to 100% over 20 min. with 50 ppm TFA and a flow of 4 mL/min. The collections were concentrated on a rotavap (Büchi V-855/R-215) and dried down under N₂(g) to yield 2.0 mg of X1. 1 and 2D NMR characterization (1H, DQF-COSY, H2BC, HMBC and HSQC) of the compound showed that the compound present is a 2:3 mixture of X1a:X1b. The NMR data of X1 indicated that two diastereomers were present in the sample. The 1H- and 13C-chemical shifts differ most in the reduced end of X1 where a stereocenter (C10) is present. The chemical shifts of H1/C1 to H5/C5 are identical. The difference of chemical shifts of the two methyl groups (H12/C12 and H13/C13) and the two CH2 groups next to the stereocenter are due to the presence of the chiral center. The two diastereomers present in the X1 solution must be due to the presence of X1 in both the E- and Z-conformation at the C6 and C7 double bond. The presence of both an E and a Z double bond conformation could a result from chemical modifications due to the presence of formic acid in the extraction process. The carbon shifts are in good agreement with published data.

Isolation of methyl (2E,6E)-10-hydroxy-11-formyl-3,7,11-trimethyl-2,6-dodecadienoate (X2).

X2 was present in the 60:40 DCM:EtOAc fraction (13.1 mg) of the Diol fractionation as described above and was purified on a Waters HPLC W600/996PDA (Milford, MA, USA) using a RP column (Phenomenex Luna C18(2), 250 x 10 mm, 5 μm, Torrance, CA, USA) using a gradient of 40% MeCN (H₂O – Milli-Q (Millipore, MA, USA)) to 100% over 20 min. with 50 ppm TFA and a flow of 4 mL/min. The collections were concentrated on a rotavap (Büchi V-855/R-215) and dried down under N₂(g) to yield 2.6 mg of X2.
Isolation of methyl (2E,6E)-10,11-epoxid-3,7,11-trimethyl-2,6-dodecadienoate (JH III). JH-III was present in the 46:60 DCM:EtOAc fraction (26.2 mg) of the Diol fractionation as described for X1 and was purified on a Waters HPLC W600/996PDA (Milford, MA, USA) using a RP column (Phenomenex Luna C18(2), 250 x 10 mm, 5 μm, Torrance, CA, USA) using a gradient of 55% MeCN (H2O – Milli-Q (Millipore, MA, USA)) to 65% over 20 min. with 50 ppm TFA and a flow of 4 mL/min. The collections were concentrated on a rotavap (Büchi V-855/R-215) and dried down under N2 (g) to yield 1.4 mg of JH-III.

NMR studies and structure elucidation: NMR spectra were acquired in DMSO-d6 on a Varian Unity Inova 500 MHz spectrometer for X1 and JH-III and on a Bruker Avance 800 MHz spectrometer at the Danish Instrument Center for NMR Spectroscopy of Biological Macromolecules for X2 using standard pulse sequences. The spectra were referenced to this solvent with resonances δH = 2.49 and δC = 39.5.

Characterization data of methyl (2E,6E)-10,11-dihydroxy-3,7,11-trimethyl-2,6-dodecadienoate (X1). NMR data for X1a: 1H NMR (500 MHz, DMSO-d6): δ 5.65 (s, 1 H), 5.07 (m, 1 H), 4.25 (d, J = 5.6 Hz, 1 H), 4.01 (s, 1 H), 3.57 (s, 3 H), 2.15 (m, 2 H), 2.15-2.11 (m, 2 H), 2.09 (s, 3 H), 1.87 (m, 2 H), 1.60 (m, 1 H), 1.56 (s, 3 H), 1.15 (m, 1 H), 1.02 (s, 3 H), 0.97 (s, 3 H); 13C NMR (125 MHz): δ 166.1, 159.7, 135.9, 122.2, 114.7, 76.6, 71.5, 50.4, 39.7, 36.2, 29.4, 26.1, 25.1, 24.2, 18.2, 15.7. NMR data for X1b: 1H NMR (500 MHz, DMSO-d6): δ 5.65 (s, 1 H), 5.08 (m, 1 H), 3.74 (dd, J = 10.0, 3.0 Hz, 1 H), 2.15 (m, 2 H), 2.15-2.11 (m, 2 H), 2.10 (m, 3 H), 1.96 (m, 1 H), 1.57 (s, 3 H), 1.48 (m, 1 H), 1.41 (m, 1 H), 1.21 (s, 3 H), 1.08 (s, 3 H); 13C NMR (125 MHz): 166.1, 159.7, 134.9, 123.0, 114.7, 82.6, 79.4, 50.4, 39.7, 35.7, 29.2, 27.7, 25.1, 22.8, 18.2, 15.6. HRMS (m/z): [M+H]+ calcd. For C16H29O4, 285.2060; found, 285.2055; [M+Na]+ calcd. For C16H28O4Na, 307.1885; found, 307.1887; [α]D = 0.0 (MeOH).

Characterization data of methyl (2E,6E)-10-hydroxy-11-formyl-3,7,11-trimethyl-2,6-dodecadienoate (X2).
\[^{1}\text{H} \text{NMR (800 MHz, DMSO-}d_6\text{):} \delta 8.23 (s, 1 H), 5.65 (q, J = 1.0 Hz, 1 H), 5.05 (t, J = 6.9, 1 H), 4.62 (s, 1 H), 4.52 (d, J = 10.2 Hz, 1 H), 3.57 (s, 3 H), 2.15 (m, 2 H), 2.11 (m, 2 H), 2.09 (d, J = 1 Hz, 3 H), 1.93 (m, 1 H), 1.84 (m, 1 H), 1.75 (m, 1 H), 1.55 (s, 3 H), 1.46 (m, 1 H), 1.04 (s, 3 H), 1.03 (s, 3 H); \[^{13}\text{C} \text{NMR (200 MHz):} \delta 166.2, 162.3, 159.9, 134.6, 123.4, 114.7, 79.7, 70.2, 50.4, 39.6, 35.7, 26.9, 25.2, 25.1, 25.1, 18.3, 15.6; \text{HRMS (m/z): [M+H]}^+ \text{calcd. for C}_{17}\text{H}_{29}\text{O}_5, 313.2010; \text{found, 313.2010. [M+Na]}^+ \text{calcd. for C}_{17}\text{H}_{28}\text{O}_5\text{Na, 335.1828; found, 335.1831.}

\textbf{Characterization data of methyl (2E,6E)-10,11-epoxid-3,7,11-trimethyl-2,6-dodecadienoate (JH III).} \[^{1}\text{H} \text{NMR data of purified compound in DMSO-}d_6\text{ showed that the compound had degraded.}

\text{HRMS (m/z): [M+H]}^+ \text{calcd. for C}_{16}\text{H}_{27}\text{O}_3, 267.1955; \text{found, 267.1957. [M+Na]}^+ \text{calcd. for C}_{16}\text{H}_{26}\text{O}_3\text{Na, 289.1780; found, 289.1774.}

\textbf{Characterization data of methyl (2, 6, 10)-3,7,11-trimethyl-2,6-dodecadienoate (MF).}

\text{HRMS (m/z): [M+H]}^+ \text{calcd. for C}_{16}\text{H}_{27}\text{O}_3, 251.2006; \text{found, 251.2007.}

\textbf{Authentic standard of JH-III was purchased from Sigma Aldrich (J2000-10MG)}

\textbf{Results and discussion}

Selection of candidate pathway specific regulators was based on genomic co-localization of gene clusters.

We utilized a collection of microarray experiments from \textit{A. niger} grown under diverse conditions to identify regulatory genes associated with predicted secondary metabolite gene clusters using a local co-expression algorithm. Seven candidate genes associated with predicted gene clusters containing either polyketide synthases or non-ribosomal peptide synthases, were identified (supplementary table 1). All seven putative
transcription factors belonged to the binuclear zinc finger class, a class often associated with secondary metabolism in fungi. BLAST analysis using the predicted protein sequences against the annotated *A. nidulans* genome (Aspergillus Comparative Database, BROAD Institute) revealed that only one candidate (fge1_pg_C_4000037) had a potential ortholog (ANID_06396, 62% amino acid identity, supplementary table 1). All seven putative regulators were expressed individually in *A. nidulans* under control of the strong constitutive *pgpdA*-promoter from the defined loci Integration Site-1. The resulting mutant strains were grown on minimal glucose media as well as four complex media representing diverse physiological conditions. Metabolites profiles were analyzed with LC-HRMS as well as UHPLC-DAD and compared to a reference strain constitutively expressing the *Escherichia coli* β-galactosidase from the same locus. Of all combinations of candidate genes and growth media, only est_fge1_pg_C_150220 (annotation from Aspergillus Comparative Database, BROAD institute) grown under high salt conditions had a significant impact on secondary metabolism resulting in increased accumulation of several metabolites not previously reported to be produced by *A. nidulans* (figure 1A). Hence we renamed est_fge1_pg_C_150220 Secondary Metabolism associated Regulatory protein A (SmrA). Two of the induced metabolites, X1 and X2 had similar UV spectra indicating a biosynthetic relationship. The two compounds were isolated and identified by NMR analysis as a methyl (2E,6E)-10,11-epoxid-3,7,11-trimethyl-2,6-dodecadienoate (X1) and its formylated analogue (X2) (figure 1B). The formylation, however, was subsequently demonstrated to occur during the extraction procedure thus X2 is a derivative of X1. The NMR data of X1 indicated that two diastereomers were present in the sample. The ¹H- and ¹³C-chemical shifts differ most in the reduced end of X1 where a stereocenter (C₁₀) is present. The chemical shifts of H₁/C₁ to H₅/C₅ are identical. The difference of chemical shifts of the two methyl groups (H₁₂/C₁₂ and H₁₃/C₁₃) and the two CH₂ groups next to the stereocenter are due to the presence of the chiral center. The two diastereomers present in the X1 solution must be due to the presence of X1 in both the E- and Z-conformation at the C₆ and C₇ double bond. The presence of both an E and a Z double bond conformation could a result from chemical modifications due to the presence of formic acid in the extraction process. The sesquiterpene X1 closely resemble the Insect Juvenile Hormone
(JH-III), a well-established hormone thought to be unique for insects with essential developmental roles and hence an important target of insecticides\textsuperscript{12, 13}. The finding of X1 prompted us to search for related induced metabolites. The accurate mass of JH-III and the related crustacean hormone methyl-farnesoate (MF)\textsuperscript{11} were used as input for targeted LC-HRMS analysis. Indeed, both searches revealed the presence of a single induced metabolite with accurate mass practically identical to the predicted mass of JH-III and MF, respectively (JH-III HRMS (m/z): [M+H]\textsuperscript{+} calculated. 267.1955; found, 267.1957, MF HRMS (m/z): [M+H]\textsuperscript{+} calculated 251.2006; found, 251.2007 supplementary figure 1). Identification of these metabolites as JH-III and MF (figure 1B) was established by comparison of retention time and mass spectra with an authentic standard (supplementary figure 1) or the metabolite database of the Xcalibur software package (Thermo Scientific), respectively. The strongest metabolite induction was observed when \textit{A. nidulans} was grown under high salt conditions, however, close inspection of the LC-HRMS data of the remaining media revealed the presence of JH-III as well as increased levels of X1 in the \textit{SmrA} expressing strain on rich yeast extract based media (supplementary figure 2). In order to establish whether the newly discovered terpenes were retained intracellularly or released to the environment we analyzed the composition of collected volatiles as well as the growth media. In the volatile fraction, MF constituted a major metabolite in the \textit{SmrA} expressing strain as well as the reference (supplementary figure 3), whereas JH-III and X1 undetectable. None of the three terpenes were detectable in the growth media. The discovery of JH-III and MF represent to our knowledge the first report of production of juvenile hormones in fungi. It is tempting to speculate that accumulation of JH-III and release of MF may serve as a defense strategy against insect competitors or predators. Fungal secondary metabolites are known to play an important role in fungal/insect interactions\textsuperscript{22, 23} and JHs constitute attractive targets due to their critical role in insect development and the fact that both absence and overexposure are detrimental\textsuperscript{12}. Previous examples of both JH biosynthesis inhibition and overexposure as defense strategies are known as \textit{Penicillium brevicompactum} synthesize the JH biosynthesis inhibitor brevioxime\textsuperscript{24}, whereas the plant \textit{Cyperus iria} synthesize JH-III and MF\textsuperscript{25}. As \textit{A. nidulans} is generally believed to inhabit a great variety of ecological niches, among those dead plant
materials, it is most likely that competing or predating insects will be part of the natural environment. However, it cannot be excluded that X1, JH-III and/or MF could also serve hormonal functions in *A. nidulans*. A similar dual purpose has been demonstrated for the JH-III precursor farnesol in *C. albicans*, as it regulate the transition from yeast to filamentous growth as well as induce apoptosis in competing fungal species. Clarification of the distribution of JHs in fungi and their biological function in *A. nidulans* will improve our understanding of fungal/insect interactions and perchance fungal morphology regulation. Moreover, the finding that heterologous expression of transcription factors may influence secondary metabolism is of general relevance for characterization of secondary metabolites. As the genomes of a large number of fungal species are currently being sequenced (www.genomesonline.org) a similar approach may be undertaken for these species using well-characterized regulatory proteins to induce previously uncharacterized metabolites.

**Acknowledgements:**

This work was supported by the Danish Research Agency for Technology and Production, grant # 09-064967. The authors would like to acknowledge the Danish Instrument Center for NMR Spectroscopy of Biological Macromolecules for the use of NMR facilities, Malene Møller Jensen, Susanne Hannibal, Hanne Jakobsen and Jesper Mogensen for technical assistance, and the Research School for Biotechnology at the faculty of Life Sciences, University of Copenhagen for funding.

**Competing interests**

The authors declare no competing interests.


Figure legends:

Figure 1: Induction of metabolites by SmrA

Top panel: UHPLC-QTOFMS extracted ion chromatogram of m/z 251 (MF, [M+H]^+), 289 (JH-III, [M+Na]^+), 307 (X1, [M+H]^+) and 335 (X2, [M+H]^+) recorded in positive mode of extracts from the strain constitutively expressing SmrA (top) and reference (bottom) grown under high salt conditions. Chromatograms are normalized by intensity. Bottom panel: Structures of X1+2, JH-III and MF.
A regulatory protein from Aspergillus niger induces juvenile hormones upon heterologous expression in A. nidulans.

Authors

Nielsen, M.T.1,2,3, Klejnstrup, M.L.1,2, Gottfredsen, C.H.4, Andersen, M.R.1, Hansen, B.G.1,5, Mortensen, U.H.1 and Larsen, T. O.1*

Affiliation

1 Center for Microbial Biotechnology, Department of Systems Biology, Technical University of Denmark, 2800 Kgs Lyngby, Denmark.
2 These authors contributed equally to the work.
3 Current address: Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Fremtidsvej 3, 2970 Hoersholm, Denmark.
4 Department of Chemistry, Technical University of Denmark, 2800 Kgs Lyngby, Denmark.
5 Current address: Novozymes A/S, Kroghshoejvej 36, 2880 Bagsvaerd, Denmark.
6 * Author to whom correspondence should be addressed (tol@bio.dtu.dk)
Supplementary figure legends:

**Supplementary figure 1** Targeted LC-MS analysis of induced metabolites.
Extracted ion chromatogram traces of the accurate mass and corresponding mass spectra of X1, X2, JH-III and MF in the mutant strain constitutively expressing SmrA as well as the authentic JH-III standard (65% pure) purchased from Sigma Aldrich. Note that the standard contains several impurities.

**Supplementary figure 2** Detection of induced metabolites on Yeast based media
Extracted ion chromatogram of X1 or JH-III of SmrA (top) and reference (bottom) grown on YES media. MS spectra of the 8.3 area are embedded. Characteristic ions from JH-III are m/z [M+H]+ 267.1957 & [M+Na]+ 289.1774 Da.

**Supplementary figure 3** Excretion of MF by A. nidulans.
(top) Total MS chromatogram of the collected volatiles from the SmrA expressing strain and the reference. (bottom) mass spectrum of the compound eluting at 26.19 minutes. The compound was identified as MF by comparison to the metabolite library of the Xcalibur software package (Thermo Scientific).

**Supplementary table 1:** Candidate genes from A. niger.

<table>
<thead>
<tr>
<th>Referred name</th>
<th>Strain #</th>
<th>Broad annotation</th>
<th>Transcript ID</th>
<th>NidA-like enzymes</th>
<th>Identity percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Candidate #1</td>
<td>ND357</td>
<td>fgl1_pq_C_4000037</td>
<td>38716</td>
<td>AN0D_06096, AN0D_03269</td>
<td>02% 27%</td>
</tr>
<tr>
<td>Candidate #2</td>
<td>ND358</td>
<td>e_gw1_101945</td>
<td>178503</td>
<td>AN0D_07346</td>
<td>28%</td>
</tr>
<tr>
<td>Candidate #3</td>
<td>ND366</td>
<td>gwr1_16247</td>
<td>188333</td>
<td>AN0D_08894</td>
<td>25%</td>
</tr>
<tr>
<td>Candidate #4</td>
<td>ND476</td>
<td>gwr1_101945</td>
<td>123782</td>
<td>None</td>
<td>-</td>
</tr>
<tr>
<td>Candidate #5</td>
<td>ND367</td>
<td>fgl1_pq_C_19000192</td>
<td>49823</td>
<td>AN0D_11683, AN0D_07921</td>
<td>43% 22%</td>
</tr>
<tr>
<td>Candidate #6</td>
<td>ND368</td>
<td>e_gw1_101945</td>
<td>184813</td>
<td>AN0D_04465</td>
<td>30%</td>
</tr>
<tr>
<td>SmrA</td>
<td>ND477</td>
<td>est_fgl1_pq_C_160220</td>
<td>54836</td>
<td>None</td>
<td>-</td>
</tr>
</tbody>
</table>

**Supplementary table 2:** Name and description of fungal strains used in this work

<table>
<thead>
<tr>
<th>Strain #</th>
<th>Genotype</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NID74</td>
<td>argB, pyrG98, vasA, nisA</td>
<td>Parental strain with permanent deletion of nisA to facilitate gene targeting</td>
<td>The study</td>
</tr>
<tr>
<td>NID945</td>
<td>argB, pyrG98, vasA, nisA, IS1_pgdA, LeuZ-TfpC-argB</td>
<td>Reference strain with E.coli LacZ integrated in IS1.</td>
<td>This study</td>
</tr>
<tr>
<td>NID357</td>
<td>argB, pyrG98, vasA, nisA, IS1_pgdA, fgl1_pq_C_4000037-argB</td>
<td>Constitutive expression of putative biocin zinc finger transcription factor fgl1_pq_C_4000037 integrated in IS1.</td>
<td>This study</td>
</tr>
<tr>
<td>NID358</td>
<td>argB, pyrG98, vasA, nisA, IS1_pgdA, e_gw1_101945-argB</td>
<td>Constitutive expression of putative biocin zinc finger transcription factor e_gw1_101945 integrated in IS1.</td>
<td>This study</td>
</tr>
<tr>
<td>NID360</td>
<td>argB, pyrG98, vasA, nisA, IS1_pgdA, e_gw1_11945-argB</td>
<td>Constitutive expression of putative biocin zinc finger transcription factor e_gw1_11945 integrated in IS1.</td>
<td>This study</td>
</tr>
<tr>
<td>NID368</td>
<td>argB, pyrG98, vasA, nisA, IS1_pgdA, gwr1_10247-argB</td>
<td>Constitutive expression of putative biocin zinc finger transcription factor gwr1_10247 integrated in IS1.</td>
<td>This study</td>
</tr>
<tr>
<td>NID376</td>
<td>argB, pyrG98, vasA, nisA, IS1_pgdA, fgl1_pq_C_19000192-argB</td>
<td>Constitutive expression of putative biocin zinc finger transcription factor fgl1_pq_C_19000192 integrated in IS1.</td>
<td>This study</td>
</tr>
<tr>
<td>NID476</td>
<td>argB, pyrG98, vasA, nisA, IS1_pgdA, e_gw1_8296-argB</td>
<td>Constitutive expression of putative biocin zinc finger transcription factor e_gw1_8296 integrated in IS1.</td>
<td>This study</td>
</tr>
<tr>
<td>NID477</td>
<td>argB, pyrG98, vasA, nisA, IS1_pgdA, SmrA-argB</td>
<td>Constitutive expression of SmrA</td>
<td>This study</td>
</tr>
</tbody>
</table>
### Supplementary table 3: PCR primers used in this study

Upper case letters indicate annealing nucleotides, lower case indicate tails for user cloning.

<table>
<thead>
<tr>
<th>PCR product</th>
<th>Primer pair</th>
<th>Forward primer</th>
<th>Reverse Primer</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JRN2D/GOG</td>
<td>GCCACGTTGTTGATATGDS</td>
<td>ATACCTGCC</td>
<td>Amplifies 3'6k downstream region from Aa-Agpl (AR_04090) for fusion PCR</td>
</tr>
<tr>
<td>2</td>
<td>JRN4G1/GOG</td>
<td>CGGACTCATGATGATGATGDS</td>
<td>GCCGCGCTGTTGATATGDS</td>
<td>Amplifies 3'6k downstream region from Aa-Agpl (AR_04090) for fusion PCR</td>
</tr>
<tr>
<td>3</td>
<td>JRN4G1/AV3</td>
<td>CGGACTCATGATGATGATGDS</td>
<td>GCCGCGCTGTTGATATGDS</td>
<td>Amplifies 5' Fpr1-3 sequence including 300bp direct repeat and native promoter for fusion PCR</td>
</tr>
<tr>
<td>4</td>
<td>JRN4G1/AV2</td>
<td>CGGACTCATGATGATGATGDS</td>
<td>GCCGCGCTGTTGATATGDS</td>
<td>Amplifies 3' Fpr1-2 sequence including 500bp direct repeat and native terminator for fusion PCR</td>
</tr>
<tr>
<td>5</td>
<td>JRN4G1/AV1</td>
<td>CGGACTCATGATGATGATGDS</td>
<td>GCCGCGCTGTTGATATGDS</td>
<td>Amplifies 3' Fpr1-1 sequence including 500bp direct repeat and native terminator for fusion PCR</td>
</tr>
<tr>
<td>6</td>
<td>JRN4G1/AV0</td>
<td>CGGACTCATGATGATGATGDS</td>
<td>GCCGCGCTGTTGATATGDS</td>
<td>Amplifies 3' Fpr1-0 sequence including 500bp direct repeat and native terminator for fusion PCR</td>
</tr>
</tbody>
</table>

### Supplementary figure 1:
Supplementary figure 2:
Supplementary figure 3:
CMBI is an Engineering Center of Excellence funded by the Danish Research Agency. It is a collaboration between an acknowledged research manager, his/her institute and university, and the Research Agency. An Engineering Center of Excellence is a research institute of first-class quality with tradition for cooperation with industry.