Genomics driven metabolomics novel strategies for the discovery and identification of secondary metabolites
Ries, M.

Citation

Version: Corrected Publisher’s Version
License: Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from: https://hdl.handle.net/1887/24303

Note: To cite this publication please use the final published version (if applicable).
The handle http://hdl.handle.net/1887/24303 holds various files of this Leiden University dissertation

**Author:** Ries, Marco  
**Title:** Genomics driven metabolomics: novel strategies for the discovery and identification of secondary metabolites  
**Issue Date:** 2014-02-25
Chapter 2

A branched biosynthetic pathway is involved in production of roquefortine and related compounds in *Penicillium chrysogenum*

*Based on*

*A branched biosynthetic pathway is involved in production of roquefortine and related compounds in Penicillium chrysogenum*

*PLos ONE, 2013, 8 (6), e65328*

*these authors contributed equally*
Abstract

Profiling and structural elucidation of secondary metabolites produced by the filamentous fungus *Penicillium chrysogenum* and derived deletion strains were used to identify the various metabolites and enzymatic steps belonging to the roquefortine/meleagrin pathway. Major abundant metabolites of this pathway were identified as histidyltryptophanyldiketopiperazine (HTD), dehydrohistidyltryptophanyldiketopiperazine (DHTD), roquefortine D, roquefortine C, glandicoline A, glandicoline B and meleagrin. Specific genes could be assigned to each enzymatic reaction step. The nonribosomal peptide synthetase RoqA accepts L-histidine and L-tryptophan as substrates leading to the production of the diketopiperazine HTD. DHTD, previously suggested to be a degradation product of roquefortine C, was found to be derived from HTD involving the cytochrome P450 oxidoreductase RoqR. The dimethylallyltryptophan synthetase RoqD prenylates both HTD and DHTD yielding directly the products roquefortine D and roquefortine C without the synthesis of a previously suggested intermediate and the involvement of RoqM. This leads to a branch in the otherwise linear pathway. Roquefortine C is subsequently converted into glandicoline B with glandicoline A as intermediates, involving two monooxygenases (RoqM and RoqO) which were mixed up in an earlier attempt to elucidate the biosynthetic pathway. Eventually, meleagrin is produced from glandicoline B involving a methyltransferase (RoqN). It is concluded that roquefortine C and meleagrin are derived from a branched biosynthetic pathway.
Introduction

Fungi produce a variety of secondary metabolites, which have diverse activities, ranging from natural antibiotics to simple toxins or immunosuppressants (Demain and Fang, 2000). Most of these metabolites are synthesized either by nonribosomal peptide synthetases (NRPS) or polyketide synthases (PKS) and can be further modified by a range of biosynthetic enzymes. The genes encoding the NRPS or PKS and the modifying enzymes are usually located in the same genetic region and are often co-expressed (Keller, et al., 2005; Shwab and Keller, 2008). Several methods have been developed to identify the secondary metabolite biosynthetic gene clusters and their respective products. Key methods are heterologous gene expression, bioinformatic (SMURF) identification of gene clusters (Khaldi, et al., 2010) and genome mining strategies in which core synthetase genes are deleted and changes in metabolite production are identified by comparative metabolite profiling (Challis, 2008). Secondary metabolites genes and/or their clusters were frequently transferred from one organism to another during evolution (Donadio, et al., 2005; Fischbach, et al., 2008). The distribution of these clusters in fungal genomes supports the hypothesis that different organisms produce similar metabolites (Ma, et al., 2010). Various fungal species were reported to produce diketopiperazines, a class of pharmaceutically important naturally produced secondary metabolites (Maristela and Carvalho, 2007; Scherlach and Hertweck, 2006). Roquefortine C, a diketopiperazine that was first isolated from *Penicillium roquefortine* (Scott and Kennedy, 1976) has now been reported from 25 different *Penicillium* species (Frisvad, et al., 2004). It displays bacteriostatic activity against Gram-positive bacteria (Kopp-Holtwiesche and Rehm, 1990). Although the exact mechanism of action is not known it appears to interact with cytochrome p450 and interferes with RNA synthesis (Aninat, et al., 2001; Kopp and Rehm, 1981). Roquefortine C also shows neurotoxic activity in mice and is considered as a contaminant in blue cheese (Polonsky, et al., 1977; Scott, et al., 1979). Meleagrin, a downstream product of roquefortine C, has been proposed to be the precursor of neoxaline, a compound with antimicrobial activity (Overy, et al., 2005).

The filamentous fungus *P. chrysogenum* has been explored for more than eighty years for its excellent fermentation capacity and penicillin production (Fleming, 2001; Weber, et al., 2012). In recent years, genome sequencing, as well as microarray analysis in combination with genetic modification of *P. chrysogenum*, has provided a strong base to elucidate the role of the secondary metabolite clusters found in this organism. The genome of *P. chrysogenum* encodes 20 putative PKS and 11 NRPS genes including the gene encoding δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase involved in penicillin biosynthesis (van den Berg, et al., 2008). This allowed us to identify the corresponding biosynthetic genes of the roquefortine/meleagrin pathway using a gene deletion strategy as well as quantitative metabolite profiling. Here, we revisit an earlier proposal for a linear pathway for the biosynthesis of roquefortine C that included several ambiguities as well speculations because of incomplete metabolite profiling and enzyme miss assignment. We have elucidated the entire biosynthetic pathway up to meleagrin production including a series of precursors. It is demonstrated that meleagrin is derived from a branched pathway with two biologically derived diketopiperazines as major intermediates.
Experimental procedures

**Chemicals**
HPLC-grade acetonitrile and methanol were purchased from Biosolve (the Netherlands). Dichloromethane, formic acid the internal standards reserpine, ranitidine and ampicillin were acquired from Sigma-Aldrich (St. Louis, MO). Meleagrin and neoxaline were purchased from Bio-Connect (the Netherlands). Roquefortine C was obtained from Bioaustalix (Australia). L-Tryptophan, L-Histidine and Mevalonic acid lactone were purchased from Sigma-Aldrich. All purchased compounds were of highest available purity.

**Host strains, media, grown condition and plasmid construction**
P. chrysogenum strain DS54555, which lacks penicillin cluster genes and Ku70 gene was used as a host strain for deletion analysis and was kindly supplied by DSM Anti-infective. All the strains were grown on YGG-medium (Kovalchuk, et al., 2012) for protoplasts formation and transformation. For analysis, cells were grown on SMP medium (glucose, 5.0 g/L; lactose, 75 g/L; urea, 4.0 g/L; Na₂SO₄, 4.0 g/L; CH₃COONH₄, 5.0 g/L; K₂HPO₄, 2.12 g/L; KH₂PO₄, 5.1 g/L) for secondary metabolites production using a shaking incubator at 200 rpm for 168 hours at 25°C. Deletion plasmids were constructed by amplifying the flanking regions of the targeted gene with the Multisite Gateway Three-Fragment Vector Construction Kit (Invitrogen). Escherichia coli DH5α was used as host strain for high frequency transformation and plasmid DNA amplification (Sambrook, et al., 1989). Tryptophan, histidine and mevalonic acid lactone were dissolve in the same phosphate buffer as used in the culture medium.

**Transformation procedure**
Deletion plasmids were transformed to the protoplasts of P. chrysogenum strain DS54555 (Alvarez, et al., 1987). The phleomycin resistance gene was used as selection marker for the deletion of roqA, roqT and roqD while acetamidase gene (amDS) was used as selection marker for the deletion of roqM, roqO, roqN and roqR using acetamide as the only nitrogen source for selection (Kolar, et al., 1988; Kovalchuk, et al., 2012).

**Genomic DNA extraction, total RNA extraction and cDNA amplification**
Genomic DNA (gDNA) was isolated after 96 hours of growth on SMP medium using the modified yeast gDNA isolation protocol (Harju, et al., 2004) in which the fungal mycelium is broken in a FastPrep FP120 system (Qbiogene). Isolated gDNA was measured using a NanoDrop ND-1000 and 5 µg was used for southern hybridization. Total RNA of the host strain was isolated after 168 hours of growth in SMP medium using Trizol (Invitrogen), with additional DNase treatment using the Turbo DNA-free kit (Ambion). Total RNA was measured with the NanoDrop ND-1000 and a concentration of 500 ng per cDNA reaction was used. cDNA was synthesized using the iScript cDNA synthesis kit (Bio-Rad) in a 10-µl end volume.
Southern blot confirmation

gDNA of the host and various deletion strains was isolated using the E.Z.N.A. Fungal DNA kit (Omega Bio-tek). Southern blotting was carried out by digesting gDNA (5 μg) with the indicated restriction enzymes. Digested DNA fragments were separated on a 0.8 % agarose gel and blotted onto a Zeta-Probe membrane (Biorad) described earlier (Nijland, et al., 2008), and hybridized with the indicated probes that were DIG labeled.

qPCR analysis

The primers used to analyze the expression of all the genes in the Roquefortine/Meleagrin biosynthetic gene cluster i.e. Pc21g15480 (roqA), Pc21g15420 (roqT), Pc21g15430 (roqD), Pc21g15440 (roqT), Pc21g15450 (roqN), Pc21g15460 (roqO) and Pc21g15470 (roqR) were designed around an intron to avoid amplification on gDNA (Table S3). For expression analyses, the y-actin gene was used as a control for normalization. A negative reverse transcriptase (RT) control was used to determine the gDNA contamination in isolated total RNA. The expression levels were determined in triplicate with a MiniOpticon system (Bio-Rad) using the Bio-Rad CFX manager software, with in which the threshold cycle (CT) values were determined automatically by regression. The SensiMix SYBR mix (Bioline) was used as a master mix for qPCR with 0.4 μM primers. The following thermocycler conditions were used: 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, 60°C for 30 s, and 72°C for 30 s. Subsequently, a melting curve was generated to determine the specificity of the qPCRs.

Microarray methods

Triplicate shake flask cultivations were performed with the P. chrysogenum strain DS17690 to prepare a proprietary DNA microarray, using the Affymetrix Custom GeneChip program (Affymetrix): After 90 hours of cultivation, samples from shake flask cultures were filtered within seconds and quenched in liquid nitrogen. A standard protocol using Trizol reagent (Invitrogen) and acid phenol-chloroform was used to extract the total RNA followed by cDNA synthesis and cRNA synthesis. Hybridized arrays were scanned and analyzed using the Affymetrix GeneChip Operating Software (GCOS, Affymetrix) as described earlier by van der Berg et al. (van den Berg, et al., 2008).

HPLC-MS validation and analysis - Sample preparation

All strains used for gene assignments were grown in quintuplicate according to the procedure described above. Samples for determination of growth curves were grown in eight replicates whereas feeding experiment samples were grown in quadruplicate. Up to 50 μL of a thawed fermentation broth, 8 μL internal standard mixture containing 855 nmol/mL ranitidine, 657 nmol/mL reserpine and 1144 nmol/mL ampicillin was added. Subsequently, 230 μL of methanol was added for protein precipitation and vortexed for 10 minutes. The sample was then centrifuged at 14,000 g for 10 minutes at 10°C. 100 μL supernatant was transferred to an Eppendorf vial and evaporated for 30 minutes in a Thermo-Speedvac (Thermo Scientific, San Jose, CA). The dried sample was redissolved in 100 μL water containing 2 % acetonitrile, vortexed for 10 minutes and transferred to an autosampler vial.
Reversed-Phase LC-MS
For separation, an Agilent 1200 Capillary pump (Agilent, Santa Clara, CA) coupled to a Surveyor PDA detector (Thermo Scientific, San Jose, CA) and LTQ-FT Ultra mass spectrometer (Thermo Scientific, San Jose, CA) were used. A sample of 5 μL was injected onto a Waters Atlantis T3 column (2.1 x 100 mm, 3 µm) (Waters, Milford, MA). The elution was performed with a linear gradient starting with 98 % of solvent A (1 % acetonitrile and 0.1 % formic acid in water) and 2 % solvent B (1 % water and 0.1 % formic acid in acetonitrile) for 1.5 minutes at a flow rate of 300 μL/min. The first linear gradient reached 40 % B at 22 minutes, the second 100 % B at 25 minutes. The column was flushed for 10 minutes at 100% B followed by equilibration for 8 minutes at 100 % A. The column effluent was directed to the ESI-LTQ-FT Ultra MS, operated in full scan \((\text{m/z } 100-2000)\) in pos/neg switching mode with following settings: Positive ion mode (4kV source voltage, 14V capillary voltage, 65V tube lens), negative ion mode (3kV source voltage, -18V capillary voltage, -85V tube lens) with capillary temperature 275°C, sheath gas flow 50 and auxiliary gas flow 2.

Data Processing
Raw files were sliced into UV-trace, positive and negative mass trace and subsequently converted into NetCDF, using an in-house tool programmed in MATLAB (MathWorks, Natick, MA). NetCDF files with same polarity were subsequently deconvoluted in DataAnalysis 4.0 (Bruker Daltonik, Bremen, Germany) using the dissect function which was controlled through a macro. The resulting peak tables for each sample were combined and repeating features removed. The combined feature table was used as target list in which each feature was integrated in every individual sample. Samples were grouped regarding their biological origin and statistical tests were performed for determination of significant differences. Discovered features were selected and transferred to LCquan v2.6 (Thermo Scientific, San Jose, CA) for more accurate integration. Peaks were auto-integrated using base peak trace in a mass range of 10 ppm and retention time window of 60 seconds and manually corrected if necessary.

Secondary metabolite identification
The identity of 7 and 4 was confirmed by comparing retention time and HR-MS spectra of samples to commercially available standards. The structure of 1 was determined by Mass Spectrometry (MS² fragmentation and fragmentation tree comparison) as will be detailed elsewhere which was adapted from the metabolite identification pipeline developed with the Netherlands Metabolomics Centre (Kasper, et al., 2012; Rojas-Cherto, et al., 2012). Compound 2, 3 and 6 were identified by NMR. 3 and 6 were extracted based on a modified method of Ohmomo (Ohmomo, et al., 1980). A P. chrysogenum culture filtrate was made alkaline with 25 % ammonium hydroxide (pH 10) and extracted with dichloromethane. The alkaline dichloromethane layer was evaporated to dryness, redissolved in water containing 50 % acetonitrile, vortexed, centrifuged and transferred to an autosampler vial for fraction collection via preparative reversed phase LC on an Atlantis T3 column (10 x 100mm, 5 µm ) (Waters Milford, MA). Compound 2 was extracted following the isolation procedure above except using ethylacetate as extraction solvent instead of dichloromethane.
NMR spectra were recorded on a Bruker Avance III 700 MHz NMR spectrometer, equipped with a 5 mm TCI probe. Samples were dissolved in 0.6 mL DMSO/CDCl₃ 50/50 and acquired at 280K (2 and 3) and 320K (6).

Results

Identification of roquefortine biosynthetic gene cluster
In order to identify the secondary metabolite biosynthetic genes under shake flask culture conditions, DNA microarray analysis was performed on the high penicillin yielding *P. chrysogenum* strain DS17690 grown in the presence and absence of the precursor phenylacetic acid (PAA). Most of the secondary metabolite gene clusters were not expressed or only at low levels (van den Berg, et al., 2008). However, high levels of expression were observed in the absence of PAA for a gene cluster including the nonribosomal peptide synthetase PC21g15480 (*roq*A), and associated genes Pc21g15420 (*roq*T), Pc21g15430 (*roq*D), Pc21g15440 (*roq*N), Pc21g15450 (*roq*O), Pc21g15460 (*roq*M) and Pc21g15470 (*roq*R) (Figure 1A). These genes were all down regulated when the cells were grown in the presence of PAA and hence are considered as co-regulated in the genome (Figure 1B). For the remainder of the study these genes were abbreviated as *roq* genes because of their involvement in roquefortine production.

Untargeted secondary metabolite profiling by HPLC-UV-MS
A robust and sensitive quantitative platform for profiling of secondary fungal metabolites from culture broth was developed, validated and applied. As the majority of these compounds are unknown and often present at low concentrations only, the method should be sensitive and provide a high degree of versatility. Herein, we used a high resolution HPLC-UV-MS method with positive/negative ionization switching in combination with UV to allow for the detection of several thousand features. To extract these features from the acquired data, a combination of commercial software packages and in-house scripts were used for untargeted peak discovery and automatic peak integration. By applying a deconvolution algorithm, fragments, adducts and cluster ions as well as their isotopes were grouped and altogether represented mainly by the most abundant ion in the resulting target list which facilitated further data analysis.

The method was validated in which several analytical performance characteristics were determined, but here merely the most important outcomes are reported. As secondary metabolites from the roquefortine/meleagrin pathway showed much higher ionization efficiency in the positive ion mode, only this polarity is described here. For the determination of linear dynamic range, sensitivity and reproducibility, retention times and signal intensities were evaluated for the used (internal) standards at multiple concentrations. These standards comprised of commercially available compounds from the meleagrin/neoxaline pathway (neoxaline, meleagrin and roquefortine C) and a non-related non-endogenous compound (ranitidine), which was used as internal standard. Retention time variations for the standards and several “unknown” endogenous compounds, which are spread over the entire chromatogram of 35 minutes, were limited to maximal 7 seconds (500 injections over 2 weeks of time). The method proved to be linear for each of the standards...
Chapter 2

Figure 1. Organization of the roquefortine/meleagrin biosynthetic gene cluster and transcriptomic analysis.
A: Roquefortine/meleagrin biosynthetic gene cluster and their orthologs in phylogenetically relative species. Homologous proteins are indicated with the same color.
B: Microarray analysis of the roquefortine biosynthetic genes in *P. chrysogenum* DS54555 using shake flask culture conditions in the absence (-) or presence (+) phenylacetic acid (PAA).
C: Correlation between the expression level of roqA and the concentration of the product HTD (1) present in the growth media. The concentration of 1 was determined by HPLC-UV-MS.
A branched biosynthetic pathway is involved in production of roquefortine and related compounds in *Penicillium chrysogenum*

over at least 3 orders of magnitude ($R^2$ ranged from 0.998 – 0.999 in both, academic and spiked matrix samples) in the appropriate concentration range. Associated Limit of Detections (LOD’s) were determined from internal standard corrected calibration lines (8 levels) and ranged from 3-248 nM, depending on the compound of interest. Due to the absence of endogenous material in the roqA deletion strain used for validation, absolute quantitation at the reported low levels is realistic. The recovery of the method was determined primarily by the extraction of the spiked compounds from the matrix, i.e., fermentation broth. Recoveries ranged from 88 % (rsvd 7.8 %) to 55 % (rsvd 2.9 %) depending on the compound. However, reproducibility measurements showed that both within day and between day reproducibility are well below 15 %. In summary, the analytical profiling platform developed here is characterized by good reproducibility and linearity, high coverage and high sensitivity reaching nanomolar levels.

**Metabolite profiling of the culture broth of a *P. chrysogenum* strain with a deletion of the nonribosomal peptide synthetase gene roqA**

In order to identify secondary metabolites synthesized by RoqA, its gene was deleted and comparative metabolite profiling was performed on the culture supernatant of the host and deletion strain. RoqA is specified by a 7.45 kbp long gene that translates into synthetase (2372 amino acids) with the typical domain motifs of nonribosomal peptide synthetases (NRPS). RoqA comprises two adenylation (A), thiolation (T) and condensation (C) domains arranged as ATCATC (Figure 1A) (van den Berg, et al., 2008). A host strain *P. chrysogenum* DS54555 was used which derived from a DS17690 strain lacking the *Ku70* gene and thus competent for homologous recombination. The DS54555 strain lacks all penicillin biosynthetic genes clusters, which facilitates the identification of unknown secondary metabolites in the culture broth as the profile is no longer dominated by β-lactam compounds. *RoqA* was deleted by homologous recombination using the deletion plasmid pDEST R4-R3a (Figure S1A) containing the phleomycin resistance gene surrounded by flanking regions of roqA. The deletion of roqA was confirmed by southern blot hybridization (Figure 2A). The host and ∆roqA strain were grown for 168 hours in secondary metabolite production medium (SMP Medium) and comparative metabolite analysis of these strains was carried out by HPLC-UV-MS (Figure S2). This revealed the loss of several secondary metabolites in the deletion strain as compared to the host strain. These metabolites were confirmed by LC-MS$^2$ analysis and subsequent fragmentation spectra, by retention time comparison with their extracted standards as well as by NMR (Figure 3, Figures S3, S4, S5 and S6, Table S1). The $m/z$ value for the compound 1 identified as HTD was observed at 324.144 dalton representing the protonated molecule $M^+ \text{H}^+$ with formula $C_{17}H_{18}N_5O_2$. Its chemical structure was elucidated using LC-MS$^2$ and high-resolution multi-dimensional fragmentation tree comparison. In addition, a protonated molecule $M^+ \text{H}^+$ at $m/z$ 404.171 and formula $C_{22}H_{22}N_5O_3$, which coincides with the protonated form of glandicoline A (5), was observed to be present in the host strain but absent in the deletion strain. Although, a full structure elucidation could not be performed, parts of the structure were identified by LC-
Southern blot hybridization was performed with total DNA extracted from P. chrysogenum D55455 strains with a deletion of the following genes: roqA (A), roqB (E), roqC (C), roqD (C), roqE (E), roqF (F), and roqT (G). The DNA was digested with the restriction enzymes as indicated in the schemes.

Figure 2. Southern blot analysis for deletion of the genes in the roquefortine/meleagrin pathway.
MS\textsuperscript{2} which indicated consistency with the chemical structure of glandicoline A (e.g. dimethylallyl group) (Figure S3E). Hence, this ion was assumed to correspond to protonated glandicoline A. In conclusion, the loss of the identified secondary metabolites in the deletion strain proves unequivocally that RoqA is responsible for the initial reaction in the roquefortine/meleagrin pathway by combining L-tryptophan and L-histidine to 1.

**Biochemical analysis of P. chrysogenum harboring a deletion of the roqR gene**

In order to elaborate the putative associated genes of the roquefortine/meleagrin biosynthetic pathway, the roqR gene encoding a putative cytochrome P450 oxidoreductase was deleted. Herein, a linearized deletion plasmid pDEST R4-R3r (Figure S1B) was transformed into the protoplast of the host strain using the *Aspergillus nidulans* acetamidase gene (*amdS*) for the positive selection of transformants growing on media with acetamide as a sole nitrogen source. The homologous recombination resulted in the complete deletion of the roqR gene as demonstrated by Southern blot hybridization (Figure 2B). Both the host and ΔroqR strains were grown as described above. Deletion of roqR resulted in an accumulation of compound 1 and 3 while 2, 4, 5, 6 and 7 were completely absent in the HPLC-MS profile of the ΔroqR strain (Figure 4A). These data suggest that 1 is a precursor of 2 resulting in a branch of the initially proposed linear pathway.

**Biochemical analysis of P. chrysogenum with a deletion of the roqD gene**

RoqD shares high sequence homology with dimethylallyltryptophan synthetase (DMATS). Transcriptional analysis of roqD showed high expression in the host strain grown in the absence of PAA in shake flasks (Figure 1B). To examine the role of RoqD in the biosynthesis of roquefortine and related metabolites, its gene was deleted with plasmid pDEST R4-R3d (Figure S1C) using phleomycin as a selection marker (Figure 2C). The deletion strain accumulated high levels of 1 (Figure 4B) in the culture broth as it was unable to add the dimethylallyl group needed for the conversion from 1 to 3. Moreover, the biosynthesis of the other (downstream) metabolites 3
Figure 4. Internal standard corrected concentrations (RR = response ratio) of secondary metabolites from roquefortine/meleagrin pathway. The metabolite concentrations in the culture broth of the ΔroqR (A), ΔroqD (B), ΔroqM (C), ΔroqO (D), ΔroqN (E) and ΔroqT (F) strains was compared to the host strain P. chrysogenum DS54555.
A branched biosynthetic pathway is involved in production of roquefortine and related compounds in Penicillium chrysogenum

Biochemical analysis of P. chrysogenum strains with a deletion of roqM, roqN and roqO gene
RoqM shows homology to MAK1-like monooxygenases. Its gene was deleted with the deletion plasmid pDEST R4-R3m (Figure S1D) using amdS as selection marker (Figure 2D). The deletion of roqM led to substantial accumulation of 3 and 4 compared to the host strain whereas (downstream) metabolites like 5 till 7 were not detected (Figure 4C). This demonstrates that roqM is involved in the synthesis of 5 using 4.

RoqO shows homology to cytochrome P450 monooxygenases, and its gene was deleted using plasmid pDEST R4-R3o (Figure 2E, Figure S1E) using amdS as selection marker. Metabolite profile comparison of the host and deletion strain showed similar concentrations of 5 and its upstream metabolites in both strains (Figure 4D). The absence of 6 and 7 in the deletion strain suggests 5 as final product and therefore indicates a role of roqO in the biosynthesis of 5 to 6.

RoqN specifies a putative methyltransferase. pDEST R4-R3n (Figure S1F) was used for the deletion of roqN using amdS for the positive selection of transformants (Figure 2F). The deletion of roqN resulted in the loss of 7 in fermentation broth (Figure 4E). In addition, the concentration of 6 was increased sixty times compared to the host strain. This increase confirms that RoqN is responsible for the conversion of 6 to 7 (Garcia-Estrada, et al., 2011).

Biochemical analysis of P. chrysogenum harboring a deletion of the roqT gene
RoqT shows high sequence homology with transporters of the major facilitator superfamily. The gene was removed from the genome (Figure 2G) using deletion plasmid pDEST R4-R3t (Figure S1G) and phleomycin as selection marker. Remarkably, the strain with a deletion of roqT did not show any marked changes in the metabolite profile as compared to host strain (Figure 4F), except that the production of 3 was increased five times whereas production of 4 was decreased by 60 %. This indicates that roqT is not essential for the biosynthesis of 1 till 7.

Gene expression and secondary metabolite production
To relate the extracellular metabolites of the roquefortine/meleagrin pathway to the expression levels of the biosynthetic genes, the host strain was grown for 216 hours in secondary SMP medium. Total mRNA extraction and extracellular metabolites analysis were carried with samples collected during growth. The metabolite concentrations were determined by HPLC-UV-MS, while transcript levels were determined by quantitative PCR using γ-actin as reference gene (Figure 1C, Figures S7 and S8). The expression levels of the various biosynthetic genes increased linear in time, except for roqA that was found to be highly up-regulated after 96 hours when the cells are in the late log phase. The concentration of metabolite 1 synthesized by RoqA increased almost equally with the expression levels of roqA (Figure 1C), while after 192 hours the concentrations of metabolites 5 till 7 in the media
decreased (Figure S7 and S8). The concentration of 4 dropped significantly already after 168 hours. The production of all metabolites was particularly high after 96 hours of growth. It is concluded that the production of these metabolites is a late event during growth and that it correlates with the expression of the respective biosynthetic genes.

**Precursor feed stimulates metabolite production**

To evaluate the role of the predicted precursors in roquefortine/meleagrin pathway and related metabolites production, feeding experiments were performed. L-Tryptophan, L-histidine and mevalonic acid lactone were added individually and in combination (at high and low concentration, i.e., 30 and 10 mM respectively) to the cultures. Since the production of 7 and most of its derivatives is most significant around 96 hours of growth, precursors were added at that time point. Filtered supernatants were analyzed on HPLC-UV-MS after 168 hours of growth. The metabolites concentrations were dry weight corrected and compared via t-test to samples without precursors addition. The concentration of metabolites 1 till 7, except 4 increased with the addition of 10 mM tryptophan (Figure 5). Addition of 30 mM tryptophan increased the production of 1, 2, 3, 5 and 7 while addition of histidine, mevalonic acid lactone and combinations of tryptophan and histidine did not reveal any significant increase. These data suggest that the roquefortine biosynthetic pathway is limited by the availability of tryptophan.

![Figure 5. Change in production of roquefortine/meleagrin metabolites after precursor addition compared to production in cultures without addition. Colored cells show mean levels that are significant (P < 0.05) different.](image)

**Discussion**

Here we have elucidated the biosynthetic pathway of *P. chrysogenum* responsible for the production of roquefortine, meleagrin and related compounds. Through a combined metabolic profiling, MS- and NMR based structure elucidation and gene inactivation analysis (7 genes in total, including a transporter-encoding gene); individual genes could be assigned to the various biosynthetic steps. The architecturally complex roquefortine and meleagrin scaffolds are synthesized from simple building blocks, i.e. histidine and tryptophan. RoqA is a di-modulated NRPS, containing two adenylation domains (van den Berg, et al., 2008) responsible for the condensation of tryptophan and histidine to pro-
A branched biosynthetic pathway is involved in production of roquefortine and related compounds in *Penicillium chrysogenum*

**Figure 6. Proposed roquefortine/meleagrin biosynthetic pathway. See text for details.**

duce 1, a diketopiperazine. The *roqA* deletion strain no longer produced any of the roquefortine related metabolites (1 till 7 absent from the broth) which was also recognized by García-Estrada et al. (Garcia-Estrada, et al., 2011). Therefore, *roqA* encodes the core enzyme initiating the biosynthesis of 1 till 7 and it is present in all species reported to produce these compounds (Frisvad, et al., 2004; Nielsen, et al., 2006). *RoqD* encodes a dimethylallyltryptophan synthase catalyzing the reversed prenylation of 1 at the C3 position in its indole moiety utilizing dimethylallyl diphosphate derived from mevalonic acid lactone of the mevalonate pathway to form 3 (Figure 6). In addition, it also catalyzes the ring closure between C2 and N14 of
the diketopiperazine moiety (Ohmomo, et al., 1975). Both conversions seem to occur simultaneously, similar to the formation of aszonalenin from benzodiazepinedione, as no intermediate could be observed (Yin, et al., 2009). No prenylated cyclohistidyltryptophanyl-diketopiperazine was detected in both the host and deletion strain as previously suggested (Garcia-Estrada, et al., 2011). The latter we attribute to a lack of convincing analytical evidence in that particular study that indeed this compound is formed. Importantly, this resulted in a assignment of RoqD and RoqM in the biosynthetic pathway. Oxidation of 1 at position C12 – C15 of the histidinyl moiety to 2 is carried out by the cytochrome p450 oxidoreductase encoded by roqR leading to a previously unknown branch in the pathway. In other studies, 2 was reported as a degradation product of 4 under acidic conditions and in-vivo (Scott, et al., 1979; Vinokurova, et al., 2001). However, the accumulation of 2 in the ΔroqD strain, despite the absence of 4 unequivocally indicates a different origin for the production of 2. Two possible reactions of 1 lead to a branch of the roquefortine/meleagrin pathway, one via oxidation by RoqR to 2 and further to 4 by dimethylallyl addition by RoqD, and the other via dimethylallyl addition by RoqD to 3 and further to 4 via oxidation carried out by RoqR (Figure 6). 2 is therefore not just a degradation product of 4 but a true biological product of the roquefortine/meleagrin pathway. The absence of 3 and 4 in the roqD deletion strain and the absence of 2 and 4 in the roqR deletion strain strongly support our proposal that 2 is the precursor of 4 following an isoprene addition by RoqD at 2 to form 4, similar to the known reaction of 1 to 3 (Figure 6). These observations are very intriguing as they question the precursor role of 3 to 4 in the proposed roquefortine/meleagrin pathway (Ohmomo, et al., 1975). 3 could be the end of the branch leaving 2 as single precursor for 4 if RoqD is unable of isoprene addition on 2 to yield 3. Therefore, 2 or 3 or even both can be possible precursors of 4.

Based on the highly significant accumulation of 4 in the roqM deletion strain and the absence of all downstream metabolites 5 till 7, we concluded that roqM, encoding a MAK 1-monoxygenase like protein, is involved in the conversion of 4 to 5. This leads to the conclusion that only one enzyme is involved in this prominent conversion which combines an oxidation, ring opening and ring closure in one reaction cascade (Steyn and Vleggaar, 1983). A possible mechanism involves an initial hydroxylation at C9 with a subsequent cleavage of the bond between C9 and N14, yielding a nine-membered ring with a keto-group at C9. Further oxidation at N1 of the indole moiety could lead to the loss of water to generate an imine which is followed by an attack of N11, ultimately yielding glandicoline A. Maackiain monoxygenases are known to be responsible for the minor conversion of Maackiain to 1-Hydroxymackian by the addition of an oxygen, which differs from the more prominent reaction from 4 to 5 found here (Covert, et al., 1996). It should however be emphasized that the enzyme assignment was initially based on sequence alignment only. Further definition of the mechanism requires an enzymological characterization.

Our results also demonstrate an involvement of RoqO in the conversion of 5 to 6. roqO encodes a p450 monoxygenase that oxidizes 5 on its tryptophan moiety, resulting in formation of 6. Remarkably, García-Estrada et al. (Garcia-Estrada, et al., 2011) mixed up the reactions catalyzed by RoqO and RoqM most likely because the lower concentrations of 6 were not recognized, whereas 5 could not have been
A branched biosynthetic pathway is involved in production of roquefortine and related compounds in Penicillium chrysogenum

detected at all with the used HPLC-UV approach. Therefore, these authors wrongly assigned the biosynthetic pathway leading to the production of 5 by RoqO and 6 by RoqM. RoqN shares 99 % identity with the UbiE/COQ5 family methyltransferase in A. origami. Its activity was confirmed by inactivation of the roqN gene resulting in a loss of 7 concomitantly with a 60 times increase of 6. Thus 7 is the methylated form of 6, consistent with an involvement of methyltransferase as earlier recognized (Garcia-Estrada, et al., 2011).

Finally, roqT encodes a highly expressed member of the Multifacilitator Super Family of transporters that might be involved in active transport of some of these metabolites. However, its deletion had little effect on the metabolic profile suggesting that passive transport, diffusion or another transporter might be involved in secretion. It is remarkable that all of the described intermediates are found in the extracellular broth which suggests that these compounds effectively diffuse from the cell. Alternatively, secretion might be mediated by a relative unspecific export system that has not been identified so far. RoqT might fulfill a role in the retention of these metabolites (see also below) but this needs to be investigated further in feeding experiments.

In order to explore the evolutionary context of this gene cluster among filamentous fungi the roquefortine/meleagrin gene clusters in various fungi were compared using the antiSMASH algorithm (Medema, et al., 2011). Sorting the gene clusters by an empirical similarity score based on the number of BlastP hits between the predicted proteins of the gene clusters, gene order conservation and the percentage identity of the Blast hits resulted in three hits of gene clusters carrying more than two homologs of genes from the roquefortine/meleagrin biosynthesis gene cluster. One gene cluster, encoded in the genome of Neosartorya fisheria, contained three genes closely similar to genes of the P. chrysogenum roquefortine/meleagrin biosynthetic pathway (59, 67 and 60 % identity with roqA, roqD and roqO, respectively) (Figure 1A, Table S2). This suggests that N. fisheria may be capable of synthesizing compounds similar to 1, 2, 3 and 4, using similar biosynthetic mechanisms. A. fumigatus and A. terreus also have three genes in common with the P. chrysogenum roquefortine gene cluster (roqA, roqD and roqO). Metabolites produced by these organisms are likely similar to 1 and 3. However, the final structures may be different as for instance A. fumigatus contains at least four additional genes (annotated as phytanoyl-CoA dioxygenase, cytochrome P450, O-methyltransferase and cytochrome P450 monoxygenase) that might be involved in further modifications of these secondary metabolites (Table S2).

A time scale study for the production of roquefortine/meleagrin biosynthesis pathway intermediates and products showed an accumulation of 7 in the fermentation broth over time. After 192 hours of growth the metabolite concentrations reached their maximum before it declined due to a possible uptake from the media or degradation. Uptake of intermediates back into mycelia was previously reported for roquefortine (Kulakovskaya, et al., 1997) and it was proposed that these compounds serve as exogenous nitrogen source for colonial expansion (Overy, et al., 2005). As tryptophan, histidine and mevalonic acid lactone are the building blocks of the roquefortine/meleagrin pathway their effect on the production of 1-7 was determined in feeding experiments. Although labeling experiments showed the uptake and incorporation of these precursors (Barrow, et al., 1979) only the presence
of tryptophan stimulated secondary metabolite formation by this pathway (Figure 5). This implies that production is limited by the availability of tryptophan providing a lead for the optimization of roquefortine/meleagrin pathway by metabolic engineering of the shikimic acid or anthranilate route for tryptophan biosynthesis (Poulsen, et al., 1993). On the other hand, histidine reversed the effect of tryptophan. This might relate to nitrogen-dependent regulation of gene expression as roquefortine and related compounds have been implicated as exogenous nitrogen sources for colonial expansion (Overy, et al., 2005), and histidine is an excellent nitrogen source for P. chrysogenum. Finally, we are currently analyzing additional intermediates for structural elucidation. It is most likely that some of the enzymes of this pathway lack a certain degree of substrate specificity, which may lead to further branching and thus a wider palette of roquefortine/meleagrin related compounds.

**Conclusion**

Genome sequencing and microarray data analysis revealed seven genes clustered in the genome of *Penicillium chrysogenum* which are highly up regulated in the absence of the penicillin G precursor phenylacetic acid. These genes are involved in the biosynthesis of roquefortine/meleagrin metabolites. Complete deletion of all seven genes and detailed biochemical analysis of respective mutants strains via HPLC, MS and NMR revealed that dehydrohistidyltryptophanyldiketopiperazine (DHTD), which was previously identified as degradation product, is biologically synthesized and a precursor to Roquefortine C. Two enzymes of this pathway catalyze more than one reaction i.e. RoqD converts histidyltryptophanyldiketopiperazine (HTD) to roquefortine D and DHTD to roquefortine C while RoqR converts HTD to DHTD and roquefortine D to roquefortine C. Thus meleagrin is synthesized via a branched pathway rather than a linear biosynthesis.

**Acknowledgements**

This work was supported by the Perspective Genbiotics program subsidized by STW, and by the NWO-ACTS ibos program and (co)financed by the Netherlands Metabolomics Centre (NMC) which is a part of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research. HA was supported by HEC and STW. MR was supported by STW. We would like to thank Drs. H. Menke, W. Heijne and H. Roubos from the DSM Biotechnology Centre for making available the DNA microarray data.

**References**


Chapter 2


A branched biosynthetic pathway is involved in production of roquefortine and related compounds in Penicillium chrysogenum

Supplemental Data

Figure S1. Map of the deletion constructs for roqA, roqR, roqD, roqM, roqO, roqN and roqT which were used for deletion. Features of the vectors: Amp, Ampicillin resistance gene for the selection in E. coli; ori, pUC origin of replication; attP3, and attP4, Gateway att recombination sites; Pipns, promoter of P. chrysogenum pcbC gene; Phleomycin, resistance gene for selection in fungi. amdS, A. nidulans acetamidase gene.

Figure S2. Analytical approach in schematic view. Proteins, present in fermentation broth, were removed during sample preparation. Samples were analyzed by HPLC-UV-MS and comparative metabolite profiling performed. Statistical significant features were extracted using liquid-liquid extraction (LLE) and semi-preparative HPLC-UV-MS and their structure elucidated by NMR.
Figure S3. HPLC-MS² fragmentation spectra including chemical formula and calculated exact mass of the protonated HTD (A), DHTD (B), roquefortine D (C), roquefortine C (D), glandicoline A (E), glandicoline B (F) and meleagrin (G) acquired at LTQ-FT-MS Ultra at 35% normalized collision energy in positive mode.
A branched biosynthetic pathway is involved in production of roquefortine and related compounds in *Penicillium chrysogenum*.

Figure S4. $^1$H NMR spectrum of DHTD (2).

Figure S5. $^1$H NMR spectrum of roquefortine D (3). Small additional peaks are not due to impurities but to a second conformation of roquefortine D.
Figure S6. $^1$H NMR spectrum of glandicoline B (6).
A branched biosynthetic pathway is involved in production of roquefortine and related compounds in *Penicillium chrysogenum*

Figure S7. Internal standard corrected metabolite concentration in fermentation broth of *P. chrysogenum* AFF393 sampled at multiple time points and determined by HPLC-UV-MS.
Figure S8. Temporal expression of roquefortine/meleagrin biosynthetic gene cluster in *P. chrysogenum* AFF393 grown in shaking flask culture.
A branched biosynthetic pathway is involved in production of roquefortine and related compounds in Penicillium chrysogenum

Table S1. $^1$H and $^{13}$C NMR chemical shifts of DHTD (2), roquefortine D (3) and glandicoline B (6) in DMSO/CDCl$_3$ at 320K and 280K, respectively (δ in ppm).

<table>
<thead>
<tr>
<th></th>
<th>$^1$H (280K)</th>
<th>$^{13}$C (280K)</th>
<th>$^1$H (320K)</th>
<th>$^{13}$C (320K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH1</td>
<td>10.77</td>
<td>n.a.</td>
<td>6.27</td>
<td>n.a.</td>
</tr>
<tr>
<td>C2</td>
<td>7.03</td>
<td>124.4</td>
<td>5.48</td>
<td>76.7</td>
</tr>
<tr>
<td>C3</td>
<td>n.a.</td>
<td>106.9</td>
<td>n.a.</td>
<td>60.7</td>
</tr>
<tr>
<td>C4</td>
<td>7.58</td>
<td>118.6</td>
<td>7.12</td>
<td>124.5</td>
</tr>
<tr>
<td>C4a</td>
<td>n.a.</td>
<td>127.4</td>
<td>n.a.</td>
<td>128.6</td>
</tr>
<tr>
<td>C5</td>
<td>6.96</td>
<td>118.4</td>
<td>6.64</td>
<td>117.4</td>
</tr>
<tr>
<td>C6</td>
<td>7.00</td>
<td>120.7</td>
<td>7.00</td>
<td>128.3</td>
</tr>
<tr>
<td>C7</td>
<td>7.22</td>
<td>111.0</td>
<td>6.56</td>
<td>108.5</td>
</tr>
<tr>
<td>C7a</td>
<td>n.a.</td>
<td>136.1</td>
<td>n.a.</td>
<td>150.8</td>
</tr>
<tr>
<td>C8</td>
<td>3.38 3.12</td>
<td>29.6</td>
<td>2.39 2.28</td>
<td>36.5</td>
</tr>
<tr>
<td>C9</td>
<td>4.32</td>
<td>56.0</td>
<td>3.87</td>
<td>58.2</td>
</tr>
<tr>
<td>OH9</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>8.86</td>
</tr>
<tr>
<td>C10</td>
<td>n.a.</td>
<td>165.8</td>
<td>n.a.</td>
<td>168.2</td>
</tr>
<tr>
<td>NH11</td>
<td>8.81</td>
<td>n.a.</td>
<td>8.24</td>
<td>n.a.</td>
</tr>
<tr>
<td>C12</td>
<td>n.a.</td>
<td>121.6</td>
<td>4.14</td>
<td>55.2</td>
</tr>
<tr>
<td>C13</td>
<td>n.a.</td>
<td>160.9</td>
<td>n.a.</td>
<td>164.9</td>
</tr>
<tr>
<td>NH14</td>
<td>10.17</td>
<td>n.a.</td>
<td>n.a.</td>
<td>9.90</td>
</tr>
<tr>
<td>C15</td>
<td>5.90</td>
<td>107.4</td>
<td>2.77 3.24</td>
<td>28.0</td>
</tr>
<tr>
<td>C16</td>
<td>n.a.</td>
<td>125.5</td>
<td>n.a.</td>
<td>136.3</td>
</tr>
<tr>
<td>NH17</td>
<td>12.91</td>
<td>n.a.</td>
<td>11.93</td>
<td>n.a.</td>
</tr>
<tr>
<td>C18</td>
<td>7.59</td>
<td>135.8</td>
<td>7.54</td>
<td>134.8</td>
</tr>
<tr>
<td>C20</td>
<td>6.94</td>
<td>133.0</td>
<td>6.87</td>
<td>113.3</td>
</tr>
<tr>
<td>C21</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>40.6</td>
</tr>
<tr>
<td>C22</td>
<td>n.a.</td>
<td>n.a.</td>
<td>6.00</td>
<td>143.7</td>
</tr>
<tr>
<td>C23</td>
<td>n.a.</td>
<td>n.a.</td>
<td>5.13 5.06</td>
<td>113.9</td>
</tr>
<tr>
<td>C24</td>
<td>n.a.</td>
<td>n.a.</td>
<td>1.09</td>
<td>22.1</td>
</tr>
<tr>
<td>C25</td>
<td>n.a.</td>
<td>n.a.</td>
<td>0.95</td>
<td>22.6</td>
</tr>
</tbody>
</table>

Chapter 2
### Neosartorya fischeri NRRL 181

<table>
<thead>
<tr>
<th>Query gene</th>
<th>Subject gene</th>
<th>% Identity</th>
<th>Blast score</th>
<th>%Coverage</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>roq D</td>
<td>NFIA_074280</td>
<td>67</td>
<td>602</td>
<td>100</td>
<td>0.0</td>
</tr>
<tr>
<td>roq R</td>
<td>NFIA_074290</td>
<td>60</td>
<td>598</td>
<td>100</td>
<td>0.0</td>
</tr>
<tr>
<td>roq A</td>
<td>NFIA_074230</td>
<td>59</td>
<td>2644</td>
<td>92</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### Aspergillus fumigatus Af293

<table>
<thead>
<tr>
<th>Query gene</th>
<th>Subject gene</th>
<th>% Identity</th>
<th>Blast score</th>
<th>%Coverage</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>roq D</td>
<td>AFUA_8G00210</td>
<td>31</td>
<td>206</td>
<td>100</td>
<td>1e-62</td>
</tr>
<tr>
<td>roq O</td>
<td>AFUA_8G00240</td>
<td>65</td>
<td>646</td>
<td>94</td>
<td>0.0</td>
</tr>
<tr>
<td>roq A</td>
<td>AFUA_8G00170</td>
<td>32</td>
<td>955</td>
<td>89</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### Aspergillus terreus NIH2624

<table>
<thead>
<tr>
<th>Query gene</th>
<th>Subject gene</th>
<th>% Identity</th>
<th>Blast score</th>
<th>%Coverage</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>roq D</td>
<td>ATEG_10306</td>
<td>31</td>
<td>226</td>
<td>100</td>
<td>7e-71</td>
</tr>
<tr>
<td>roq O</td>
<td>ATEG_10307</td>
<td>47</td>
<td>333</td>
<td>70</td>
<td>5e-112</td>
</tr>
<tr>
<td>roq A</td>
<td>ATEG_10305</td>
<td>32</td>
<td>820</td>
<td>81</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table S2: BlastP analysis of roquefortine/meleagrine biosynthetic pathway genes.

### Primers designed for gene expression analysis of roquefortine/meleagrine biosynthetic gene

<table>
<thead>
<tr>
<th>Target</th>
<th>Primer sequence (5’- 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Forward</td>
</tr>
<tr>
<td>roqA</td>
<td>AATTAGTGCGTCTCATCTCC</td>
</tr>
<tr>
<td>roqD</td>
<td>CTTGGTGCGCGATTCCCGAGC</td>
</tr>
<tr>
<td>roqR</td>
<td>CCTGCGAATACACTGGGGG</td>
</tr>
<tr>
<td>roqM</td>
<td>CTGCAATCTCTGACTATAATCGC</td>
</tr>
<tr>
<td>roqO</td>
<td>GACGACGATTGCTGACACC</td>
</tr>
<tr>
<td>roqN</td>
<td>CGTCCACTGTGCTGGGCACC</td>
</tr>
<tr>
<td>roqT</td>
<td>AACTGATCCTCTACCGCAGG</td>
</tr>
</tbody>
</table>

Table S3. Primers designed for gene expression analysis of roquefortine/meleagrine biosynthetic gene.
A branched biosynthetic pathway is involved in production of roquefortine and related compounds in Penicillium chrysogenum.