

Technical University of Denmark



The cyclochlorotine mycotoxin is produced by the nonribosomal peptide synthetase CctN in *Talaromyces islandicus* (“*Penicillium islandicum*”)

Schafhauser, Thomas; Kirchner, Norbert; Kulik, Andreas; Huijbers, Mieke M.E.; Flor, Liane; Caradec, Thibault; Fewer, David P.; Gross, Harald; Jacques, Philippe; Jahn, Linda; Jokela, Jouni; Leclère, Valérie; Ludwig-Müller, Jutta; Sivonen, Kaarina; van Berkel, Willem J.H.; Weber, Tilmann; Wohlleben, Wolfgang; van Pée, Karl-Heinz

Published in:
Environmental Microbiology

Link to article, DOI:
[10.1111/1462-2920.13294](https://doi.org/10.1111/1462-2920.13294)

Publication date:
2016

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Schafhauser, T., Kirchner, N., Kulik, A., Huijbers, M. M. E., Flor, L., Caradec, T., ... van Pée, K-H. (2016). The cyclochlorotine mycotoxin is produced by the nonribosomal peptide synthetase CctN in *Talaromyces islandicus* (“*Penicillium islandicum*”). *Environmental Microbiology*, 18(11), 3728–3741. DOI: 10.1111/1462-2920.13294

DTU Library
Technical Information Center of Denmark

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.

The cyclochlorotine mycotoxin is produced by the nonribosomal peptide synthetase CctN in *Talaromyces islandicus* (“*Penicillium islandicum*”)

Thomas Schafhauser*¹, Norbert Kirchner^{2,3}, Andreas Kulik¹, Mieke M.E. Huijbers⁴, Liane Flor⁵, Thibault Caradec⁶, David P. Fewer⁷, Harald Gross^{2,3}, Philippe Jacques⁶, Linda Jahn⁸, Jouni Jokela⁷, Valérie Leclère⁶, Jutta Ludwig-Müller⁸, Kaarina Sivonen⁷, Willem J.H. van Berkel⁴, Tilmann Weber^{1,3,9}, Wolfgang Wohlleben*^{1,3} and Karl-Heinz van Pée⁵

¹Mikrobiologie und Biotechnologie, Interfakultäres Institut für Mikrobiologie und Infektionsmedizin, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

²Pharmaceutical Institute, Department of Pharmaceutical Biology, University of Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

³German Centre for Infection Research (DZIF), Partner site Tübingen, 72076 Tübingen, Germany

⁴Laboratory of Biochemistry, Wageningen University, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

⁵Allgemeine Biochemie, Technische Universität Dresden, 01069 Dresden, Germany

⁶Research Laboratory in Agro-Food and Biotechnology, Charles Viollette Institute, Team ProBioGEM, Polytech-Lille, Université Lille1- Sciences et Technologies, 59655 Villeneuve d'Ascq, France

⁷Microbiology and Biotechnology Division, Department of Food and Environmental Sciences, University of Helsinki, Viikinkaari 9, FIN-00014, Helsinki, Finland

⁸Institut für Botanik, Technische Universität Dresden, 01062 Dresden, Germany

⁹The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kogle Alle 6, 2970 Hørsholm, Denmark

*Corresponding authors: thomas.schafhauser@biotech.uni-tuebingen.de;
wolfgang.wohlleben@biotech.uni-tuebingen.de

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/1462-2920.13294

SUMMARY

Talaromyces islandicus (“*Penicillium islandicum*”) is a widespread foodborne mold that produces numerous secondary metabolites, among them potent mycotoxins belonging to different chemical classes. A notable metabolite is the hepatotoxic and carcinogenic pentapeptide cyclochlorotine that contains the unusual amino acids β -phenylalanine, 2-aminobutyric acid and 3,4-dichloroproline. Although the chemical structure has been known for over five decades, nothing is known about the biosynthetic pathway of cyclochlorotine. Bioinformatic analysis of the recently sequenced genome of *T. islandicus* identified a wealth of gene clusters potentially coding for the synthesis of secondary metabolites. Here we show by RNA interference-mediated gene silencing that a nonribosomal peptide synthetase, *CctN*, is responsible for the synthesis of cyclochlorotine. Moreover, we identified novel cyclochlorotine chemical variants, whose production also depended on *cctN* expression. Surprisingly, the halogenase required for cyclochlorotine biosynthesis is not encoded in the *cct* cluster. Nonetheless, our findings enabled us to propose a detailed model for cyclochlorotine biosynthesis. In addition, comparative genomics revealed *cct*-like clusters are present in all of the sequenced *Talaromyces* strains indicating a high prevalence of cyclochlorotine production ability.

INTRODUCTION

Filamentous fungi are prolific producers of secondary metabolites (SMs) with various bioactivities that range from toxic (e.g. aflatoxins, fumonisins) to beneficial and pharmaceutically important (e.g. penicillin, cyclosporine, lovastatin). The biosynthesis genes of fungal SMs are usually organized on a single genetic locus as a biosynthetic gene cluster (BGC) (Keller and Hohn, 1997; Hoffmeister and Keller, 2007). The genus *Talaromyces* is among the chemically most versatile producers of all fungi (Frisvad, 2014). To date, four genomes of the 88 known *Talaromyces* strains have been sequenced and annotated. Bioinformatic analyses based on biosynthetic key genes revealed between 48 and 62 BGCs in the investigated genomes (Nierman et al., 2015; Schafhauser et al., 2015), indicating that the genus has the ability to produce more SMs than currently described in the literature.

Talaromyces islandicus, formerly designated as *Penicillium islandicum* (Samson et al., 2011), is a commonly found grain decaying mold and an important foodborne species in tropical regions (Frisvad and Thrane, 2004). *T. islandicus* produces more than 20 mainly polyene-like or quinoid pigments (Yilmaz et al., 2014) including skyrin, erythrokyrin, and islandicin, some of which show pronounced hepatotoxicity (Cole et al., 1976; Liberman et al., 1980; Kawai et al., 1984). Furthermore, *T. islandicus* is known to produce the halogenated peptide cyclochlorotine which is toxic and carcinogenic to liver cells of mice (Uraguchi et al., 1972). Cyclochlorotine, also referred to as islanditoxin (Samson et al., 2011), was identified as one of the so-called yellow-rice syndrome causing agents in Japan after World War II (Uraguchi, 1961), when *T. islandicus* became prevalent on foodstuff deliveries from overseas. The molecular target of cyclochlorotine was presumed to be cytoskeletal filaments (Ohmi et al., 2001). However, the specific mechanisms of toxicity and carcinogenicity remain poorly understood. Structurally, cyclochlorotine is composed of the nonproteinogenic amino acids L-2-aminobutyrate (2Abu), L- β -phenylalanine (β Phe), the unique L-*cis*-3,4-dichloroproline [Pro(Cl₂)] and two L-serines (Ser), which results in the cyclic chlorinated pentapeptide cyclo[Ser- β Phe-Ser-Pro(Cl₂)-2Abu]. It thus shows

remarkable similarities to the plant-derived astins of *Aster tataricus* (Kosemura et al., 1993; Morita et al., 1993). The only difference between cyclochlorotine and astin C is the replacement of one of the serines by a second 2-aminobutyrate [Ser-βPhe-2Abu-Pro(Cl₂)-2Abu]. Nonetheless, nothing is known about the biosynthesis of cyclochlorotine and the same applies for the formation of the non-proteinogenic moieties βPhe, 2Abu and the unique Pro(Cl₂).

Modified cyclic pentapeptides are typically synthesized by non-ribosomal peptide synthetases (NRPSs), large multifunctional enzymes that assemble simple building blocks to complex molecules (Hori et al., 1989; Finking and Marahiel, 2004). About 500 different amino acids have been identified to date in nonribosomal peptides (reviewed in Walsh et al., 2013), thereby extending the range of selectable building blocks far beyond the 21 proteinogenic amino acids (a comprehensive documentation and examples are given in the Norine database, Caboche et al., 2008). However, modified cyclic peptides may as well derive from ribosomally produced natural products (RiPPs) that undergo posttranslational modification and cyclization. This has recently been exemplified by the characterization of the ustiloxin B pathway in *Aspergillus flavus* (Umemura et al., 2014).

Here, we used a comprehensive genome mining approach to identify and characterize the cyclochlorotine biosynthesis gene cluster from *Talaromyces islandicus*. Based on our results, we propose an NRPS pathway model which to our knowledge represents the first characterized NRPS pathway in the genus *Talaromyces*.

MATERIAL AND METHODS

General

Unless otherwise stated, all chemicals were purchased from Sigma-Aldrich, St. Louis, USA. Plasmids are listed in Table 1 and sequences of primers used in this study are shown in Table S1 in supplementary information. Purification of the putative halogenase Pisl3812_03494 and the halogenation assays as well as purification of the proline dehydrogenase Pisl3812_07821 and the dehydrogenation assay are described in the supplementary information.

As standard cultivation, *T. islandicus* WF-38-12 (ATCC 26535, obtained as *Penicillium islandicum* Sopp) was grown in PDB (potato dextrose broth, see Table S2) at 27°C in the dark. Cultures were incubated under shaking (150 rpm) for strain maintenance and under non-shaking for metabolite production purposes and for qPCR studies, unless otherwise stated.

Gene expression assay

Quantitative real-time PCR (qPCR) on an IQ5 iCycler (Bio-Rad, Munich, Germany) was used to determine the expression of the *cct* cluster genes, the putative flavin-dependent halogenase gene *pisl3812_03494* and the *cct*-like NRPS gene *pisl3812_04906*. Total RNA was extracted from liquid nitrogen ground material of 21-day-old standing cultures grown in PDB (in case of the halogenase gene expression assays 7-d-old shaking cultures were used) with the RNeasy Mini Kit (Qiagen, Hilden, Germany) and using the QIAshredder homogenizers. Isolated RNA was treated with DNase (Thermo Scientific) to remove genomic DNA contamination. Total cDNA was prepared from RNA (50 ng/μl) using the RevertAid reverse transcriptase (0.2 μl, Thermo Scientific) and 100 pmol anchored oligo(dT)₁₀ primers (sequence: 5'-NVT TTT TTT TTT TTT-3') in a 20 μl volume reaction. The qPCR was assessed with the iQ SYBR Green supermix (2x, Bio-Rad), 40 pmol primers and 1.5 μl cDNA in a 20 μl volume reaction. The cycle threshold (Ct) values

were determined automatically by the Bio-Rad IQ5 software (V2.1) after automated base-line subtraction. A two-step program on the thermocycler was used: 95°C for 3 min and subsequently 40 cycles at 95°C for 10 sec and 55°C for 60 sec. For each sample, RNA was used as a negative control to confirm absence of contaminating genomic DNA. For each gene, 3 biological replicates were performed (in case of the halogenase gene expression assays only 2 replicates) with primers amplifying a ~250 bp region near the 3' end of the gene. Melting curve analyses were performed at the end of the PCR program, to exclude possible nonspecific amplification products. PCR efficiency (E) for each pair of primers was calculated in advance with the help of the regression of a standard curve from dilution series of genomic DNA with the formula $E_{\text{gene}} = 10^{(-1/\text{regression})}$. Finally, all gene expression values were normalized to the expression of the reference genes *gpd* (glyceraldehyde-3-phosphate dehydrogenase coding) or *act-1* (β -actin coding) from the same sample with the formula

$$E_{\text{ref}}^{\Delta Ct_{\text{ref}}}/E_{\text{gene}}^{\Delta Ct_{\text{gene}}}$$

Silencing the NRPS gene *cctN*

RNA interference-mediated gene silencing was realized with the hairpin RNA-expressing plasmid pSilent-1. A ~500 bp fragment of gene *cctN* was PCR-amplified with the primer pair *cctN*_KD-F / *cctN*_KDHind-R (generating a 3' HindIII-site) and ligated into the pJET1.2/blunt vector resulting in a template construct that is flanked by primer- and vector-derived restriction sites (BglII-XhoI-*cctN* fragment-HindIII-BglII). *Sense* and *antisense* fragments were consecutively ligated into pSilent-1 using XhoI/HindIII-restriction and BglII-restriction and the T4 DNA ligase (Thermo Scientific), respectively, in consideration of the correct orientation of the latter fragment (*antisense*). The resulting plasmid pST-*cctN* was transferred into *T. islandicus* WT protoplasts. Gene expression of some hygromycin resistant transformants was analyzed and the strain with lowest *cctN*-expression was selected for further studies.

Overexpression of genes *cctM*, *cctT* and *pisI3812_03494*

The plasmid pSilent-1 was used as basis for gene overexpression taking advantage of the strong *Aspergillus nidulans* PtrpC promoter in front of the first MCS. The hygromycin resistance cassette was cut out with NotI and exchanged against the (terminator-free) nourseothricin resistance cassette of plasmid pD-NAT1 using the primer pair PtrpCNot-F/Nat1Not-R and T4 DNA ligase. The resulting plasmid pOE was linearized with XhoI/BglII, which eliminated the loop forming spacer sequence. The genes *cctM*, *cctT* and *pisI3812_03494* were PCR-amplified with the primer pairs cctM_OE-F / -R, cctT_OE-F / -R and halo3494_OE-F / -R, which generated overhangs complementary to the plasmid endings, and ligated into the linearized pOE by Gibson assembly (New England Biolabs, Ipswich, USA) according to the manufacturer protocol. The resulting plasmids pOE-cctM, pOE-cctT and pOE-halo3494 were transferred into *T. islandicus* WT protoplasts. For each gene, expression analysis of some nourseothricin resistant transformants was performed and strains with highest gene-expression were selected for further studies.

Protoplast transformation

Protocols for generation and transformation of protoplasts were adapted from Huang, et al.(1995). *T. islandicus* spores were incubated for 2-3 days in PDB under shaking. Cells from 15 mL culture were pelleted (5000 rpm, 10 min), washed twice with osmotic medium (OM, 1.2 M MgSO₄, 50 mM phosphate buffer, pH 5.8) and resuspended in 10 mL OM. Generation of protoplasts was achieved by adding 5 mg/mL Yatalase (Takara, Kyoto, Japan) and 5 mg/mL Glucanex (lysing enzymes from *T. harzianum*, Sigma-Aldrich) followed by incubation under shaking at 30°C for 0.5-1.5 h. After protoplast formation, the mixture was filtered using a glass frit (pore size 40-100 µm), carefully covered with 10 mL STC buffer

(1.2 M Sorbitol, 10 mM CaCl₂, 10 mM Tris-HCl, pH 7.5) and centrifuged at 1500 g for 15 min. Protoplasts at the interphase were generously removed, pelleted, washed twice with 5 mL STC and finally resuspended in 1 mL STC.

For transformation of protoplasts, 1 µg plasmid DNA and 40 µl PEG (polyethylene glycol) solution consisting of 40 % (w/v) PEG 6000 (Merck, Darmstadt, Germany), 50 mM CaCl₂ and 10 mM Tris-HCl (pH 7.5) were added to 1-5x10⁶ protoplasts (usually 200 µl). The mixture was incubated for 10 min on ice, another 1 mL PEG solution was added followed by 10 min incubation at room temperature and finally 5 mL STC buffer were added. Transformed protoplasts were pelleted and plated onto PDA (potato dextrose agar) containing 1.2 M sorbitol and either 200 µg/mL nourseothricin (cloNAT, Werner Bioagents, Jena, Germany) or 100 µg/mL hygromycin B (Sigma-Aldrich).

Cyclopeptide screening by HPLC-MS and HPLC-MS²

Cyclochlorotine and variants were extracted from 20 mL of culture filtrate with the same volume of butanol. The butanol phase was evaporated to dryness *in vacuo* and extracted material was dissolved in 0.5 mL MeOH/H₂O (50:50) and subsequently analyzed by mass spectrometry as described in Spohn et al. (2014).

HPLC-MS guided cyclopeptide isolation

Ten L of CYA medium (Czapek Yeast Autolysate, see Table S2) were inoculated with 0.2 L of a 3-day-old *T. islandicus* pre-culture (grown under shaking in PDB), divided over 24 Erlenmeyer flasks and incubated for 12 days under non-shaking conditions at 27°C. Likewise, 2.5 L of CYA medium were inoculated, divided over 6 Fernbach flasks and incubated for 31 days. Cultures from both approaches were merged, filtered and adsorbed on Amberlite XAD16 resin (40x8 cm column). MeOH eluents containing cyclopeptides were evaporated *in vacuo*, dissolved in H₂O and separated by size-exclusion

chromatography on a Bio-Gel P2 column (95x5.2 cm, Bio-Rad). Fractions containing compounds **1-2** were combined, purified by preparative RP-HPLC (Reprosil-Pur Basic-C18, 10 μm , 250x20 mm, Maisch, Ammerbuch, Germany) with a linear gradient from 30-70 % MeOH over 20 min. The flow rate for RP-HPLC was 24 mL min⁻¹. The isolation yielded 30 mg of cyclochlorotine (**1**). Fractions containing hydroxycyclochlorotine (**2**) were further purified by RP-HPLC (Reprosil Pur ODS-3, 10 μm , 250x20 mm, Maisch) and eluted with 30 % acetonitrile (ACN, isocratic) which yielded 10 mg of **2**. Bio-Gel fractions containing compound **3** as well as combined fractions with compounds **4-5** were evaporated *in vacuo*, dissolved in MeOH and each consecutively purified on Sephadex LH-20 (90x5 cm, Amersham, Freiburg, Germany) and Toyopearl HW-40-F (95x2.5 cm, Toyo Biosep, Stuttgart, Germany) columns. Fractions containing **3** were purified by two consecutive RP-HPLC steps on Reprosil Pur Basic-C18 (linear gradient from 30-60 % MeOH over 15 min) and Reprosil Pur ODS-3 (20 % ACN, isocratic), which yielded 3 mg of cyclotine (**3**). Compounds **4** and **5** from Toyopearl fractions were combined and purified on Reprosil Pur Basic-C18 (linear gradient from 30-60 % MeOH over 20 min) and finally separately purified on Reprosil Pur ODS-3. Elution with ACN (linear gradient of 20-26 % ACN over 1.5 min, followed by 26 % ACN isocratic) yielded ~2 mg of compound **4** and the elution with a linear gradient from 20-60 % ACN over 15 min yielded ~2 mg of compound **5**.

Structure elucidation (HR-MS and NMR)

Optical rotations were measured on a JASCO P-2000 polarimeter. Nuclear magnetic resonance (NMR) spectra of compound **1** were recorded on a Bruker AMX-600 spectrometer. NMR spectra of compounds **2, 3, 4,** and **5** were acquired on a Bruker Avance III 500 HD spectrometer, equipped with a BBFO cryo probe head. Spectra were referenced to residual protonated solvent signals with resonances at $\delta_{\text{H/C}}$ 2.50/39.5 (deuterated dimethyl sulfoxide [d_6 -DMSO]). High-resolution (HR) electrospray ionization-time

of flight mass spectrometry (ESI-TOF MS) data were recorded using a Bruker Daltonic maXis 4G instrument.

Accepted Article

RESULTS AND DISCUSSION

Identification of candidate NRPS genes for cyclochlorotine biosynthesis

The ability of *T. islandicus* to produce nonribosomal peptides was investigated by means of the SM analysis software antiSMASH (version 3.0, Weber et al., 2015). A total of 62 BGCs were identified from the genome, among them as many as 13 putative NRPS gene clusters. Four of the putative 13 NRPS genes encode a predicted penta-modular composition consistent with the pentapeptide structure of cyclochlorotine. Given that all amino acids in cyclochlorotine occur in L-configuration, no epimerization domains are expected. Furthermore, a terminal domain displaying a C_T domain specific signature (Caradec et al., 2014) is expected in the NRPS for cyclization and release of cyclochlorotine. These considerations allowed us to further narrow down the plausible NRPSs to two candidates with the deduced domain architecture A-T-(C-A-T)₄-C_T (see Fig. S1), in which A, T, C and C_T stand for adenylation domain, thiolation domain, condensation domain and terminal condensation-like domain, respectively. These remaining NRPSs share a remarkably high amino acid sequence identity of 57 % over the whole length of around 5,600 amino acids (see Fig. S1). They are encoded by the genes *cctN* (16926 bp) and *pis13812_04907* (16874 bp) which are located on different chromosomes, respectively.

T. islandicus produced cyclochlorotine under standard cultivation conditions. Metabolite production was analyzed by high performance liquid chromatography coupled to mass spectrometry (HPLC-MS) of culture filtrate butanol extracts. Cyclochlorotine was identified by the mass of its pseudo-molecular ion ($[M+H]^+$, m/z 572.4, Fig. 1A) and due to the characteristic isotope pattern caused by the two chlorine atoms of the Pro(Cl₂) residue (Fig. 1B). According to the initial genome sequence analysis two candidate NRPS genes might play a role in cyclochlorotine biosynthesis. To elucidate whether both candidate NRPS genes are transcribed, quantitative real-time PCR (qPCR) studies under standard cultivation conditions were performed in the *T. islandicus* control strain (WT + empty plasmid pSilent-1, see below). The expression analysis revealed that only *cctN* is transcribed in these cyclochlorotine producing cultures

(21.9 % \pm 5.9, normalized to reference gene, Fig. 1C), while *pis/3812_04907* is hardly expressed (0.05 % \pm 0.01, data not shown). Investigation of the WT without the empty plasmid pSilent-1 confirmed these results (not shown) suggesting that only the NRPS CctN is involved in the biosynthesis of cyclochlorotine.

The NRPS CctN is required for cyclochlorotine production

Due to extremely inefficient homologous recombination, RNA interference based gene silencing in combination with ectopic integration of plasmids was considered for further exploring the functions of the candidate genes. For this, an internal 500 bp sequence of *cctN* was ligated in both orientations around the loop forming spacer sequence of the plasmid pSilent-1. The resulting hairpin RNA-expressing plasmid (pST-cctN) was subsequently transferred into *T. islandicus* wild type protoplasts. Two of several transformants were analyzed by quantitative real-time PCR (qPCR) revealing a decrease of *cctN* expression to ~8 % and ~16 % (with regard to the *cctN* expression level of the WT control strain with an empty plasmid [WT+pSilent-1]). The transformant with lowest *cctN* expression was grown in triplicates and qPCR analysis at days 7, 14, and 21 confirmed a stable decrease of *cctN* expression at around 10 % (Fig. 1C). Remarkably, metabolite profile analysis showed the complete abolishment of cyclochlorotine production, whereas the control strain (WT+pSilent-1) was not affected in this respect (Fig. 1A). This clearly demonstrated the requirement of the NRPS CctN in the biosynthesis of cyclochlorotine. In addition, comparison of the metabolite profiles of the *cctN* silenced strain (WT+pST-cctN) with the control strain (WT+pSilent-1) allowed us to screen for further pentapeptide metabolites that potentially originate from the same pathway. Indeed, besides cyclochlorotine, an additional seven compounds with ion masses, retention times and isotope pattern suitable for variants of cyclochlorotine were detected in wild type culture extracts (Fig. 1A and 1B). These metabolites were completely absent in the silenced strain indicating that their productions also rely on the NRPS CctN.

Isolation of new pentapeptides and structure determination

Out of 15 different cultivation media tested, CYA medium was selected because most of the cyclochlorotine variants were produced in appropriate amounts (Fig. S3). Since production rates of the individual metabolites slightly altered over time, filtrates of cultures differing in age were collected and applied to HPLC-MS guided column chromatography and preparative HPLC. The isolated compounds **1** and **2** were verified as cyclochlorotine and the *allo*-threonine-containing hydroxy-cyclochlorotine (Mizutani et al., 2008), respectively, on the basis of their pseudo-molecular ion peaks in high-resolution electrospray-ionization time-of-fly MS (HR-ESI-TOF-MS; $[M+Na]^+$ at m/z 594.1493 for **1**; $[M+Na]^+$ at m/z 610.1442 for **2**) and by comparison of their recorded 1D-NMR data with the respective literature data. (A detailed discussion of the spectral and mass data is provided in the supplementary information.)

Likewise, for the structure elucidation of the new compounds **3**, **4**, and **5** HR-ESI-TOF-MS data as well as extensive 1D- and 2D-NMR data were recorded (see supplementary information). Accordingly, compound **3** ($[M+Na]^+$ at m/z 526.2270) represents the dehalogenated derivative of cyclochlorotine (**1**) for which the trivial name cyclotine is proposed. Compound **4** ($[M+Na]^+$ at m/z 560.1883) is the mono-chlorinated derivative of **1** with the proline residue being chlorinated only at its β -position. For this compound, the trivial name cyclochlorotine B is proposed. Compound **5** ($[M+Na]^+$ at m/z 578.1544) possesses an alanine residue instead of a serine residue between β Phe and Pro(Cl₂), with respect to cyclochlorotine (**1**), and therefore the trivial name deoxy-cyclochlorotine is proposed. The compounds **6**, **7**, and **8** were only obtained in minor amounts and no NMR studies could be performed. ESI-MS² experiments of these compounds revealed identical fragmentation patterns as for compounds **1**, **3**, and **4** respectively (Fig S36), so that they can be assumed as stereo- or regioisomers of their respective counterparts.

Of all variants from *T. islandicus*, cyclochlorotine (**1**) was produced in the highest quantities and production rates decreased from compound **2** to **8**. The new pentapeptide variants from *T. islandicus* will help to identify the biological target and exploring the structure-function relationship in future experiments. Especially the pentapeptides varying in chlorination will be investigated in this context in order to clarify whether chlorination is crucial for biological activity.

Characterization of the cyclochlorotine biosynthesis cluster

To determine the full extent of the cyclochlorotine gene cluster, the genetic environment of the *cctN* gene was investigated (see Fig. 2). The left border is marked by a presumably non-coding region of around 270 kbp with a GC content below 20 %. Notably, the whole genome of *T. islandicus* is sporadically interspersed by such A+T rich islands that span up to 300-400 kbp. These islands consist of repetitive as well as mobile elements and they encode, if any, only fragments of genes. They might represent areas of increased horizontal gene transfer activity which, however, has yet to be supported by experimental or further theoretical evidence. Genome comparisons of the right cluster region with MultiGeneBlast (Medema et al., 2013) revealed that the right border of the cyclochlorotine cluster is defined by *pisl3812_02626* representing the first gene of a conserved synteny block occurring in many ascomycete genomes, among them in *Penicillium roqueforti*. Hence, we concluded that the cyclochlorotine gene cluster encompasses the eight genes *pisl3812_02618-25*, designated as *cctM* to *cctT* (Fig. 3 and Table 2). Based on sequence comparisons with the NCBI Conserved Domain Database, *cctQ* and *cctS* encode proteins belonging to the major facilitator superfamily (MFS) and to ABC-transporters, respectively. Such transporters mediate the efflux of SMs from fungi. The gene *cctP* is assumed to encode a phenylalanine ammonia lyase (PAL). Genomic analyses revealed that PALs are found predominantly in plants and fungi aside from one report from prokaryotes (Wu et al., 2014). Members of this subfamily of

methylideneimidazole-5-one (MIO) prosthetic group enzymes catalyze the conversion of Phe into cinnamic acid. However, minor structural modifications could turn the lyase activity into a phenylalanine aminomutase (PAM) activity. The rigidity of an active-site inner loop is discussed to be crucial for forcing the MIO-amine adduct in the active site to promote re-addition of the amino group onto cinnamic acid to form β Phe as the final product (Heberling et al., 2015). In this model, the inner loop enables the chemically challenging 2,3-amine shift when closed and causes a lyase reaction when open. In order to predict the enzymatic function of CctP, the primary sequence was compared with those of the *Taxus chinensis* PAM and *Petroselinum crispum* PAL prototypes. Small and hydrophobic residues at characteristic inner loop positions were identified (Fig. S37) which, in accordance with the findings of Heberling et al. (2015), may prevent the exposure to solvent and hence promote rigidity to the loop. Thus we suggest that CctP rather acts as a PAM and provides the uncommon building block β Phe for cyclochlorotine biosynthesis. Interestingly, the C-terminus of CctP comprises a domain of unknown function (DUF3328) that is completely uncharacterized and only found in eukaryotes. The small genes *cctO* and *cctR* encode DUF3328 domains, too. In the ascomycete *Sordaria macrospora*, a model organism for the analysis of sexual development and cell differentiation, a cluster of also three DUF3328 encoding genes is upregulated in young fruiting bodies compared to sexual mycelium which might indicate an involvement in cell differentiation (Teichert et al., 2012). Notably though, the cluster in *S. macrospora* does not comprise any SM genes. The scarce knowledge about DUF3328 proteins does not allow any speculation of their role in cyclochlorotine biosynthesis. The remaining two genes *cctM* and *cctT* encode proteins with domains belonging to the short chain dehydrogenase/reductase (SDR) superfamily and although they have no obvious role in cyclochlorotine biosynthesis we provisionally assigned them as members of the cluster.

Quantitative real-time PCR studies on cDNA prepared from cyclochlorotine producing cultures were carried out showing that all genes in the *cct* gene cluster were strongly expressed (Fig. 4A). The

comparable expression levels suggest a co-regulated role during biosynthesis, although no transcriptional regulator is encoded within the cluster. More detailed expression studies will be necessary to address the exact regulation of the cyclochlorotine biosynthesis genes. The examination of gene expression did not help to assign the borders of the *cct* gene cluster, since the neighboring genes of *cctT* were transcribed at similar levels as the cluster genes (data not shown).

Cct-like gene clusters are prevalent in *Talaromyces* strains

Many *Talaromyces* species are of economical relevance. For instance, *T. marneffeii* is a medically important pathogen causing mycosis in immunocompromised patients (Zheng et al., 2015) and *T. stipitatus* has an impact on food industry as it causes spoilage of fruit based products (Yilmaz et al., 2014). Therefore, the occurrence of *cct*-like genes in these and other fungi was investigated in consideration of the toxicity of cyclochlorotine pentapeptides that may derive from such a genetic disposition. Indeed, comparative genomics using MultiGeneBlast (Medema et al., 2013) revealed that the cyclochlorotine synthetase coding gene *cctN* and five adjacent genes, *cctO-S*, in the *cct* gene cluster of *T. islandicus* show remarkable similarities (yet low synteny) to genes of the uncharacterized and orphan NRPS gene clusters from *T. stipitatus* ATCC 18224 (genes *tsta_038070-130*), *T. marneffeii* ATCC 18224 (genes *pmaa_075440-540*) and *T. cellulolyticus* Y-94 (genes *tce0_017r03506-24*) suggesting conserved function in the respective pathways (Fig. 3). Thus, all of the *Talaromyces* genomes sequenced and annotated to date harbor a *cct*-like NRPS cluster. Moreover, the NRPSs of all four species share very high amino acid identities (53-81 %), identical deduced domain architectures (A-T-(C-A-T)₄-C_T) as well as extremely similar A domain binding pockets (see Fig. S1). A further common feature of the *cct*-like gene clusters is the presence of DUF3328 coding genes including their general bias to be repeatedly encoded up to four times. These findings indicate that many *Talaromyces* species have the ability to produce a cyclic pentapeptide of the cyclochlorotine group. No *cct*-like gene clusters were found in genomes of

other organisms, neither eukaryotes nor bacteria which could indicate that cyclochlorotines are specific to the genus *Talaromyces*. The only structurally similar pentapeptides described in literature are the astins of the plant *A. tataricus* (Kosemura et al., 1993; Morita et al., 1993). However, no genome information is available for the plant. It is conceivable that astins are actually produced by a fungal endophyte potentially related to the genus *Talaromyces*.

Cyclochlorotine biosynthesis model

Supply of the non-proteinogenic amino acids

Cyclochlorotine and variants consist of five amino acids including modified and quite uncommon building blocks that have to be supplied before or during peptide assembly. The cluster-encoded PAM CctP most likely provides β Phe (see above). Furthermore, 2Abu can be synthesized from threonine which in *E. coli* has been shown to depend on the concerted actions of a dehydratase and a transaminase (Fotheringham et al., 1999). In the first step of 2Abu synthesis, the dehydration of threonine is accompanied by a hydrolytic release of ammonia resulting in an oxo-acid, which requires the transfer of an amino group in the second step (see Fig. 5). Genes encoding putative threonine dehydratases (e.g. *pis13812_05069*, *pis13812_05443*) and several transaminases are scattered over the genome and the respective enzymes should be capable of 2Abu biosynthesis.

The building block *allo*-threonine for the side-product hydroxy-cyclochlorotine (**2**) is likely formed from acetaldehyde and glycine via an aldose reaction (such reactions are reviewed in Duckers et al., 2010) and as many as five enzymes that contain threonine aldolase domains are encoded in the genome, none of them however is located near the cyclochlorotine cluster. These enzymes, together with the putative threonine dehydratases, constitute a remarkable accumulation of threonine metabolizing enzymes in

T. islandicus which, besides being used for catabolic purposes, are apparently also used for biosyntheses of SMs.

Intriguingly, no halogenase for the synthesis of the characteristic Pro(Cl₂) moiety is encoded in the *cct* gene cluster and therefore the search for an appropriate gene was expanded to the complete genome. Formation of Pro(Cl₂) requires regio- and stereospecific chlorination which is basically attributed to two groups of halogenating enzymes, the flavin-dependent halogenases and the non-heme iron and α -ketoglutarate-dependent [2OG-Fe(II)] halogenases (e. g. reviewed by Neumann et al., 2008; van Pée, 2012). Candidate 2OG-Fe(II) halogenase genes were identified by searching for conserved signatures (InterPro domain IPR006620) of prolyl oxidizing enzymes in deduced amino acid sequences. However, alignments of all candidate sequences did not reveal the halogenase typical HxA motif described by Blasiak et al. (2006) but showed the oxygenase characteristic HxD dyad (Fig. S37). We therefore assume that no 2OG-Fe(II) halogenase is involved in Pro(Cl₂) synthesis. The alternative Pro(Cl₂) formation via the action of flavin-dependent halogenases requires proline to be initially oxidized to pyrrole carboxylic acid in order to provide a double bond for the subsequent halogenase reaction (Thomas et al., 2002; Dong et al., 2005; van Pée and Patallo, 2006). One of three putative proline dehydrogenases identified by a BlastP search was heterologously produced (Fig. S38) and showed a specific proline dehydrogenating activity of 0.6 U/mg (comparable to other ProDHs, see Huijbers and van Berkel, 2015) which demonstrates that the initial proline oxidation step can be accomplished. Likewise, a genome wide BlastP search revealed one putative flavin-dependent halogenase (Pisl3812_03494) featuring characteristic halogenase motifs. To investigate chlorination efficiency, the putative halogenase Pisl3812_03494 was heterologously produced and its activity was assayed *in vitro*. However, no chlorination could be observed in the assay by using different substrates such as the unchlorinated pentapeptide **3**, pyrrole-2-carboxylic acid and 1H-pyrrole-2-carbonyl-SNAC. The latter substrate was supposed to imitate the native substrate being tethered to peptidyl carrier proteins (see also Fig. 5). The missing halogenation activity strongly

challenges the hypothesis that Pisl3812_03494 is involved in cyclochlorotine biosynthesis. In line with this, qPCR experiments demonstrate that the halogenase encoding gene *pisl3812_03494* is virtually not transcribed during metabolite production (Fig. S39A). Moreover, the insertion of an extra gene copy under the promoter *P_{trpC}* resulted in a strong overexpression of the putative halogenase gene but did not influence the ratio of produced chlorinated and non-chlorinated pentapeptides, as represented by constant amounts of the compounds cyclochlorotine (**1**) and cyclotine (**3**) (Fig. S39B). In summary, the biosynthesis of the unusual Pro(Cl₂) moiety remains elusive. Even though the genome was thoroughly examined for candidate genes, the actual halogenase still awaits to be detected which includes the possibility that proline halogenation is accomplished by a so far unknown type of halogenase.

Cyclochlorotine assembly, release and export

The deduced domain organization A-T-(C-A-T)₄-C_T of the cyclochlorotine synthetase CctN (see Fig. S1) implies a linear assembly of five building blocks in which the selection of the substrate amino acids is based on the specificity conferring signatures of the A domains. This signature, also known as the nonribosomal code, is composed of 10 key amino acids that form the A domain active site (Stachelhaus et al., 1999; Challis et al., 2000). In order to assign the A domains of CctN to their substrate specificities, nonribosomal codes were extracted (Fig. 6 and Fig. S1) with the software NRPSpredictor2 (Röttig et al., 2011) and manually compared with characterized fungal A domains that incorporate the same building blocks as CctN presumably does. Indeed, the nonribosomal code of CctN-A3 shares conserved residues with that of the SimA-A6 domain from cyclosporine biosynthesis (Fig. 6). From *in vitro* experiments it was concluded that SimA-A6 accepts 2Abu (Dittmann et al., 1994). Likewise, the CctN-A2 code resembles that of the proline activating Tex1-A13 domain in the paracelsin pathway, whose specificity to proline in turn could be concluded *a priori* (Wiest et al., 2002). Finally, both signatures of CctN-A4 and CctN-A1 are reminiscent of that of the biochemically characterized GliP-A2 domain which incorporates serine into

gliotoxin (Balibar and Walsh, 2006). Accordingly, the assembly order in CctN is very likely Ser-Pro-2Abu-Ser- β Phe which leads (after variable chlorination) to the compounds **1**, **3**, and **4**. The variants **2** and **5** might result from incorporation of the alternative substrates *allo*-threonine or alanine by the domains A3 or A1, respectively. Considering the lower production yields of compounds **5** and **2**, the alternative substrates though are much less favored.

On the other hand, the proposed A domain promiscuity regarding Thr/2Abu and Ser/Ala is questionable since in both cases the substrate pairs differ by one hydrogen donor/acceptor which rather would expect differences in the shape of the binding pockets. Thus, it is also conceivable that compound **1** derives from **2** by α,β -dehydration of the *allo*-threonine moiety which would yield an intermediate containing dehydrobutyrine (similar to the biosynthesis of the lanthipeptide nisin, Karakas Sen et al., 1999). In contrast to the reaction on free *allo*-threonine, the dehydration of a mature peptide would not be accompanied by a hydrolysis of ammonia but result in formation of a C=C bond which can easily be reduced to form the 2Abu residue. Likewise, compound **5** might result from **1** in a similar manner with an intermediate peptide harboring a dehydroalanine residue instead of serine, which subsequently is dehydrogenated to alanine. To test whether the *cct* cluster-encoded putative dehydrogenases CctM or CctT are involved in these dehydration reactions, their genes were overexpressed in *T. islandicus* (Fig. 4A), which should change the ratio of compounds **2**, **1**, and **5** in favor of the latter ones. However, compound analysis of the overproduction strains did not show any changes in this production ratio (Fig. 4B), indicating that neither CctM nor CctT play a role in this hypothetical scenario.

The assembled linear pentapeptide, tethered to domain T₅ of module five, is then assumed to be released through cyclization by the terminal condensation-like (C_T) domain which shows the characteristic variations in the first histidine residue of the highly conserved HHxxxDGxS motif of C domains (Bergendahl et al., 2002) (⁵³⁶⁰SHMQNDGWC in CctN, Fig. S1), as it has been demonstrated for other fungal NRPSs that produce macrocyclic peptides (Gao et al., 2012). Finally, the cluster encoded

transporters CctQ and CctS are most likely responsible for cyclochlorotine secretion and thereby may contribute to intrinsic resistance.

CONCLUSION

This study represents the first discovery of an NRPS gene cluster responsible for the biosynthesis of the cyclic chlorinated pentapeptide cyclochlorotine and newly described variants by taking advantage of the recently sequenced genome of *T. islandicus* (Schafhauser et al., 2015). In addition, the strain possesses a transcriptionally inactive and cryptic NRPS gene which is apparently paralogous to *cctN*. The high similarity of both genes is a strong argument for an intragenomical duplication event of an ancestral gene. The incidence of homologous NRPS biosynthetic genes of which one is silent has been found earlier in work on *Aspergillus fumigatus* (Steinchen et al., 2013). Genetic engineering and metabolite analysis revealed that both NRPS are functional and produce a similar set of small metabolites (Kalb et al., 2015). Apparently, the effort of harboring redundant NRPS pathways is prevalent in filamentous fungi and might emphasize the significance of the corresponding SMs.

Considering the presence of homologous cyclochlorotine NRPS gene clusters in all of the *Talaromyces* strains from which annotated genomes are available and the fact that species within the genera *Aspergillus*, *Penicillium* and *Talaromyces* often have SMs in common (Frisvad, 2014), production of cyclochlorotine or similar pentapeptides might be more frequent than previously assumed. This, together with the fact that cyclic pentapeptides are easily overlooked by HPLC-coupled diode array detectors (due to only weak chromophore groups) could make a thorough screening for such SMs worthwhile in the genus *Talaromyces*. Furthermore, determining the cyclochlorotine gene cluster opens up the possibility of developing diagnostic PCR based techniques which are of relevance in environmental microbiology in terms of food control.

Investigations of the cyclochlorotine biosynthesis genes and rational interpretation of the biosynthetic capacity deduced from genetic analysis allowed us to draw a detailed model for cyclochlorotine biosynthesis. Although several biosynthetic steps have still to be verified in future biochemical experiments, this biosynthesis model provides the entry for the manipulation of the biosynthetic machinery towards higher yields and for rational design of cyclochlorotine variants. For instance, introduction of strong fungal promoters in front of biosynthetic key genes should increase formation of selected cyclopeptides.

One of the main hallmarks of fungal SMs, the clustering of biosynthesis genes (Keller and Hohn, 1997), does not fully apply to the cyclochlorotine pathway, since important genes for precursor supply and tailoring reactions are not localized in the vicinity of the cyclochlorotine synthetase gene *cctN*. Especially, the formation of Pro(Cl₂) remains mysterious, since no halogenase gene is part of the gene cluster and the only putative flavin-dependent halogenase encoded in the genome is not involved. This could imply that *T. islandicus* harbors a hitherto unknown kind of halogenating enzyme that awaits detection in future experiments. Remarkably, the unexpected issue of halogenases not being co-localized with the other biosynthesis genes was recently also demonstrated for the melleolide pathway in *Armillaria mellea* (Wick et al., 2015). It appears that for some SM pathways the biosynthesis genes and the halogenase gene have independent phylogenetic histories, which is in disagreement with the common assumption that SM biosynthesis genes usually are inherited as a joint entity. In these examples, halogenation might be a type of modification event that evolved independently and at a later evolutionary time point.

ACKNOWLEDGMENTS

This work was supported by the ERA-IB project MESIAB (EIB 10.004). K.H.v.P., L.F., J.L.-M., and L.J. were funded by the Sächsische Staatsministerium für Wissenschaft und Kunst (SMWK) and the Sächsische Aufbaubank (SAB). W.W. and T.S. were supported by the Bundesministerium für Bildung und Forschung

(BMBF, grant 0315934). T.W. is supported by a grant of the Novo Nordisk Foundation. M.M.E.H. and W.J.H.v.B. are supported by grants from the Netherlands Organization for Scientific Research (NWO) and The Graduate School VLAG (Wageningen, The Netherlands). Research at the University of Helsinki was supported by the Finnish Funding Agency for Technology and Innovation TEKES 400077/11 (to K.S.) and the Academy of Finland grants 1258827 (to K.S.) and 1259505 (to D.P.F.). We thank M. Brock, Hans Knoell Institute Jena, for valuable advice on protoplast transformation. The plasmids pD-NAT1 and pSilent-1 were obtained from the Fungal Genetics Stock Center (FGSC, Manhattan, Kansas, USA). We are very grateful to A. Steck, Analytical Services & Applied NMR Development, at Bruker Biospin GmbH and M. Kramer, NMR Spectroscopy Division, Institute for Organic Chemistry, University of Tübingen for support and NMR measurements. We also thank D. Wistuba and her team of the Mass Spectrometry Division of the Institute for Organic Chemistry, University of Tübingen for HR-MS measurements.

REFERENCES

- Balibar, C.J., and Walsh, C.T. (2006) GliP, a multimodular nonribosomal peptide synthetase in *Aspergillus fumigatus*, makes the diketopiperazine scaffold of gliotoxin. *Biochemistry* **45**: 15029-15038.
- Bergendahl, V., Linne, U., and Marahiel, M.A. (2002) Mutational analysis of the C-domain in nonribosomal peptide synthesis. *Eur J Biochem* **269**: 620-629.
- Blasiak, L.C., Vaillancourt, F.H., Walsh, C.T., and Drennan, C.L. (2006) Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **440**: 368-371.
- Caboche, S., Pupin, M., Leclère, V., Fontaine, A., Jacques, P., and Kucherov, G. (2008) NORINE: a database of nonribosomal peptides. *Nucleic Acids Res* **36**: D326-331.
- Caradec, T., Pupin, M., Vanvlassenbroeck, A., Devignes, M.D., Smail-Tabbone, M., Jacques, P., and Leclère, V. (2014) Prediction of monomer isomery in Florine: a workflow dedicated to nonribosomal peptide discovery. *Plos One* **9**: e85667.
- Challis, G.L., Ravel, J., and Townsend, C.A. (2000) Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. *Chem Biol* **7**: 211-224.
- Cole, R.J., Kirksey, J.W., Cutler, H.G., Wilson, D.M., and Morganjones, G. (1976) Two toxic indole alkaloids from *Penicillium islandicum*. *Can J Microbiol* **22**: 741-744.
- Dittmann, J., Wenger, R.M., Kleinkauf, H., and Lawen, A. (1994) Mechanism of cyclosporin A biosynthesis. Evidence for synthesis via a single linear undecapeptide precursor. *J Biol Chem* **269**: 2841-2846.
- Dong, C., Flecks, S., Unversucht, S., Haupt, C., van Pée, K.H., and Naismith, J.H. (2005) Tryptophan 7-halogenase (PrnA) structure suggests a mechanism for regioselective chlorination. *Science* **309**: 2216-2219.
- Duckers, N., Baer, K., Simon, S., Groger, H., and Hummel, W. (2010) Threonine aldolases-screening, properties and applications in the synthesis of non-proteinogenic beta-hydroxy-alpha-amino acids. *Appl Microbiol Biotechnol* **88**: 409-424.
- Finking, R., and Marahiel, M.A. (2004) Biosynthesis of nonribosomal peptides. *Annu Rev Microbiol* **58**: 453-488.
- Fotheringham, I.G., Grinter, N., Pantaleone, D.P., Senkpeil, R.F., and Taylor, P.P. (1999) Engineering of a novel biochemical pathway for the biosynthesis of L-2-aminobutyric acid in *Escherichia coli* K12. *Bioorg Med Chem* **7**: 2209-2213.
- Frisvad, J.C. (2014) Taxonomy, chemodiversity, and chemoconsistency of *Aspergillus*, *Penicillium*, and *Talaromyces* species. *Front Microbiol* **5**: 773-780.
- Frisvad, J.C., and Thrane, U. (2004) Mycotoxin production by common filamentous fungi. In *Introduction to food- and airborne fungi*. Samson, R.A., Hoekstra, E.S., and Frisvad, J.C. (eds). Utrecht: Centraalbureau voor Schimmelcultures, pp. 321-331.
- Gao, X., Haynes, S.W., Ames, B.D., Wang, P., Vien, L.P., Walsh, C.T., and Tang, Y. (2012) Cyclization of fungal nonribosomal peptides by a terminal condensation-like domain. *Nat Chem Biol* **8**: 823-830.
- Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**: 4121-4130.
- Heberling, M.M., Masman, M.F., Bartsch, S., Wybenga, G.G., Dijkstra, B.W., Marrink, S.J., and Janssen, D.B. (2015) Ironing out their differences: dissecting the structural determinants of a phenylalanine aminomutase and ammonia lyase. *Acs Chem Biol* **10**: 989-997.
- Hoffmeister, D., and Keller, N.P. (2007) Natural products of filamentous fungi: enzymes, genes, and their regulation. *Nat Prod Rep* **24**: 393-416.
- Hori, K., Yamamoto, Y., Minetoki, T., Kurotsu, T., Kanda, M., Miura, S. et al. (1989) Molecular cloning and nucleotide sequence of the gramicidin S synthetase 1 gene. *J Biochem* **106**: 639-645.

- Huang, K.X., Iwakami, N., Fujii, I., Ebizuka, Y., and Sankawa, U. (1995) Transformations of *Penicillium islandicum* and *Penicillium frequentans* that produce anthraquinone-related compounds. *Curr Genet* **28**: 580-584.
- Huijbers, M.M., and van Berkel, W.J. (2015) High yields of active *Thermus thermophilus* proline dehydrogenase are obtained using maltose-binding protein as a solubility tag. *Biotechnol J* **10**: 395-403.
- Kalb, D., Heinekamp, T., Lackner, G., Scharf, D.H., Dahse, H.M., Brakhage, A.A., and Hoffmeister, D. (2015) Genetic engineering activates biosynthesis of aromatic fumaric acid amides in the human pathogen *Aspergillus fumigatus*. *Appl Environ Microb* **81**: 1594-1600.
- Karakas Sen, A., Narbad, A., Horn, N., Dodd, H.M., Parr, A.J., Colquhoun, I., and Gasson, M.J. (1999) Post-translational modification of nisin. The involvement of NisB in the dehydration process. *Eur J Biochem* **261**: 524-532.
- Kawai, K., Kato, T., Mori, H., Kitamura, J., and Nozawa, Y. (1984) A comparative study on cytotoxicities and biochemical properties of anthraquinone mycotoxins emodin and skyrin from *Penicillium islandicum* Sopp. *Toxicol Lett* **20**: 155-160.
- Keller, N.P., and Hohn, T.M. (1997) Metabolic pathway gene clusters in filamentous fungi. *Fungal Genet Biol* **21**: 17-29.
- Kosemura, S., Ogawa, T., and Totsuka, K. (1993) Isolation and structure of asterin, a new halogenated cyclic pentapeptide from *Aster tataricus*. *Tetrahedron Lett* **34**: 1291-1294.
- Kück, U., and Hoff, B. (2006) Application of the nourseothricin acetyltransferase gene (*nat1*) as dominant marker for the transformation of filamentous fungi. *Fungal Genet News*: 9-11.
- Liberman, D.F., Schaefer, F.L., Fink, R.C., Ramgopal, M., Ghosh, A.C., and Mulcahy, R. (1980) Mutagenicity of islandicin and chrysophanol in the *Salmonella* microsome system. *Appl Environ Microb* **40**: 476-479.
- Medema, M.H., Takano, E., and Breitling, R. (2013) Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol Biol Evol* **30**: 1218-1223.
- Mejean, A., Mann, S., Vassiliadis, G., Lombard, B., Loew, D., and Ploux, O. (2010) In vitro reconstitution of the first steps of anatoxin-a biosynthesis in *Oscillatoria* PCC 6506: from free L-proline to acyl carrier protein bound dehydroproline. *Biochemistry* **49**: 103-113.
- Mizutani, K., Hirasawa, Y., Sugita-Konishi, Y., Mochizuki, N., and Morita, H. (2008) Structural and conformational analysis of hydroxycyclochlorotine and cyclochlorotine, chlorinated cyclic peptides from *Penicillium islandicum*. *J Nat Prod* **71**: 1297-1300.
- Morita, H., Nagashima, S., Takeya, K., and Itokawa, H. (1993) Astins A and B, antitumor cyclic pentapeptides from *Aster tataricus*. *Chem Pharm Bull (Tokyo)* **41**: 992-993.
- Nakayashiki, H., Hanada, S., Quoc, N.B., Kadotani, N., Tosa, Y., and Mayama, S. (2005) RNA silencing as a tool for exploring gene function in ascomycete fungi. *Fungal Genet Biol* **42**: 275-283.
- Neumann, C.S., Fujimori, D.G., and Walsh, C.T. (2008) Halogenation strategies in natural product biosynthesis. *Chem Biol* **15**: 99-109.
- Nierman, W.C., Fedorova-Abrams, N.D., and Andrianopoulos, A. (2015) Genome sequence of the AIDS-associated pathogen *Penicillium marneffe* (ATCC18224) and its near taxonomic relative *Talaromyces stipitatus* (ATCC10500). *Genome Announc* **3**: e01559-01514.
- Ohmi, K., Enosawa, S., Nonomura, Y., Tatsuno, T., and Ueno, Y. (2001) Acceleration of actin polymerization and rapid microfilament reorganization in cultured hepatocytes by cyclochlorotin, a hepatotoxic cyclic peptide. *Toxicon* **39**: 303-308.
- Röttig, M., Medema, M.H., Blin, K., Weber, T., Rausch, C., and Kohlbacher, O. (2011) NRPSpredictor2 - a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res* **39**: W362-W367.

- Samson, R.A., Yilmaz, N., Houbraken, J., Spierenburg, H., Seifert, K.A., Peterson, S.W. et al. (2011) Phylogeny and nomenclature of the genus *Talaromyces* and taxa accommodated in *Penicillium* subgenus *Biverticillium*. *Stud Mycol* **70**: 159-183.
- Schafhauser, T., Wibberg, D., Rückert, C., Winkler, A., Flor, L., van Pée, K.H. et al. (2015) Draft genome sequence of *Talaromyces islandicus* ("*Penicillium islandicum*") WF-38-12, a neglected mold with significant biotechnological potential. *J Biotechnol* **211**: 101-102.
- Spohn, M., Kirchner, N., Kulik, A., Jochim, A., Wolf, F., Muenzer, P. et al. (2014) Overproduction of ristomycin A by activation of a silent gene cluster in *Amycolatopsis japonicum* MG417-CF17. *Antimicrob Agents Chemother* **58**: 6185-6196.
- Stachelhaus, T., Mootz, H.D., and Marahiel, M.A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem Biol* **6**: 493-505.
- Steinchen, W., Lackner, G., Yasmin, S., Schrettl, M., Dahse, H.M., Haas, H., and Hoffmeister, D. (2013) Bimodular peptide synthetase SidE produces fumarylalanine in the human pathogen *Aspergillus fumigatus*. *Appl Environ Microb* **79**: 6670-6676.
- Teichert, I., Wolff, G., Kück, U., and Nowrousian, M. (2012) Combining laser microdissection and RNA-seq to chart the transcriptional landscape of fungal development. *BMC Genomics* **13**: 511.
- Thomas, M.G., Burkart, M.D., and Walsh, C.T. (2002) Conversion of L-proline to pyrrolyl-2-carboxyl-S-PCP during undecylprodigiosin and pyoluteorin biosynthesis. *Chem Biol* **9**: 171-184.
- Umemura, M., Nagano, N., Koike, H., Kawano, J., Ishii, T., Miyamura, Y. et al. (2014) Characterization of the biosynthetic gene cluster for the ribosomally synthesized cyclic peptide ustiloxin B in *Aspergillus flavus*. *Fungal Genet Biol* **68**: 23-30.
- Uraguchi, K. (1961) Isolation of two toxic agents, luteoskyrin and chlorine-containing peptide, from the metabolites of *Penicillium islandicum* Sopp, with some properties thereof. *Jpn J Exp Med* **31**: 19-46.
- Uraguchi, K., Tatsuno, T., Takahashi, K., Saito, M., Enomoto, M., and Noguchi, Y. (1972) Chronic toxicity and carcinogenicity in mice of purified mycotoxins, luteoskyrin and cyclochlorotine. *Food Cosmet Toxicol* **10**: 193-207.
- van Pée, K.H. (2012) Enzymatic chlorination and bromination. *Method Enzymol* **516**: 237-257.
- van Pée, K.H., and Patallo, E.P. (2006) Flavin-dependent halogenases involved in secondary metabolism in bacteria. *Appl Microbiol Biotechnol* **70**: 631-641.
- Walsh, C.T., O'Brien, R.V., and Khosla, C. (2013) Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds. *Angew Chem Int Ed Engl* **52**: 7098-7124.
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H.U., Brucoleri, R. et al. (2015) antiSMASH 3.0 - a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* **43**: W237-W243.
- Wick, J., Heine, D., Lackner, G., Misiek, M., Tauber, J., Jagusch, H. et al. (2015) A fivefold parallelized biosynthetic process secures chlorination of *Armillaria mellea* (honey mushroom) toxins. *Appl Environ Microb* **82**: 1196-1204.
- Wiest, A., Grzegorski, D., Xu, B.W., Goulard, C., Rebuffat, S., Ebbole, D.J. et al. (2002) Identification of peptaibols from *Trichoderma virens* and cloning of a peptaibol synthetase. *J Biol Chem* **277**: 20862-20868.
- Wu, Z.H., Gui, S.T., Wang, S.Z., and Ding, Y. (2014) Molecular evolution and functional characterisation of an ancient phenylalanine ammonia-lyase gene (NnPAL1) from *Nelumbo nucifera*: novel insight into the evolution of the PAL family in angiosperms. *Bmc Evol Biol* **14**: 1-14.
- Yilmaz, N., Visagie, C.M., Houbraken, J., Frisvad, J.C., and Samson, R.A. (2014) Polyphasic taxonomy of the genus *Talaromyces*. *Stud Mycol* **78**: 175-341.
- Zheng, J., Gui, X., Cao, Q., Yang, R., Yan, Y., Deng, L., and Lio, J. (2015) A clinical study of acquired immunodeficiency syndrome associated *Penicillium marneffei* infection from a non-endemic area in China. *Plos One* **10**: e0130376.

TABLES

Table 1. Plasmids used in this study

Plasmid	Features and application	Reference (origin)
pJET1.2/blunt	Cloning of PCR-amplified DNA fragments	Thermo Scientific, Waltham, USA
pSilent-1	Contains a hygromycin B resistance cassette; Carries a transcriptional unit for expression of hairpin RNA: a loop forming spacer sequence is flanked by two multiple cloning sites (MCS) for <i>sense</i> and <i>antisense</i> cloning of target gene fragment; Served as template for the expression plasmid pOE	Nakayahiki et al., 2005 (obtained from FGSC)
pD-NAT1	Source of the nourseothricin resistance cassette	Kück and Hoff, 2006 (obtained from FGSC)
pOE	pSilent-1-based expression plasmid for fungi with the nourseothricin resistance cassette	This study
pBAD	For protein synthesis in <i>E. coli</i> with the N-terminal affinity tag maltose binding protein (MBP)	Guzman et al., 1995(56)

Table 2. Features of *cct* cluster genes

Locus tag	Gene name	No. of introns	No. of amino acids	deduced domain/proposed function	Putative biosynthetic role
<i>pisl3812_02618</i>	<i>cctM</i>	5	395	Short-chain dehydrogenase	
<i>pisl3812_02619</i>	<i>cctN</i>	1	5641	NRPS	Cyclochlorotine synthetase
<i>pisl3812_02620</i>	<i>cctO</i>	4	300	DUF3328	
<i>pisl3812_02621</i>	<i>cctP</i>	6	997	Phenylalanine aminomutase and DUF3328	β Phe-formation
<i>pisl3812_02622</i>	<i>cctQ</i>	2	252	MFS transporter	Efflux/resistance
<i>pisl3812_02623</i>	<i>cctR</i>	1	261	DUF3328	
<i>pisl3812_02624</i>	<i>cctS</i>	0	1553	ABC transporter	Efflux/resistance
<i>pisl3812_02625</i>	<i>cctT</i>	1	283	Short-chain dehydrogenase	

FIGURES

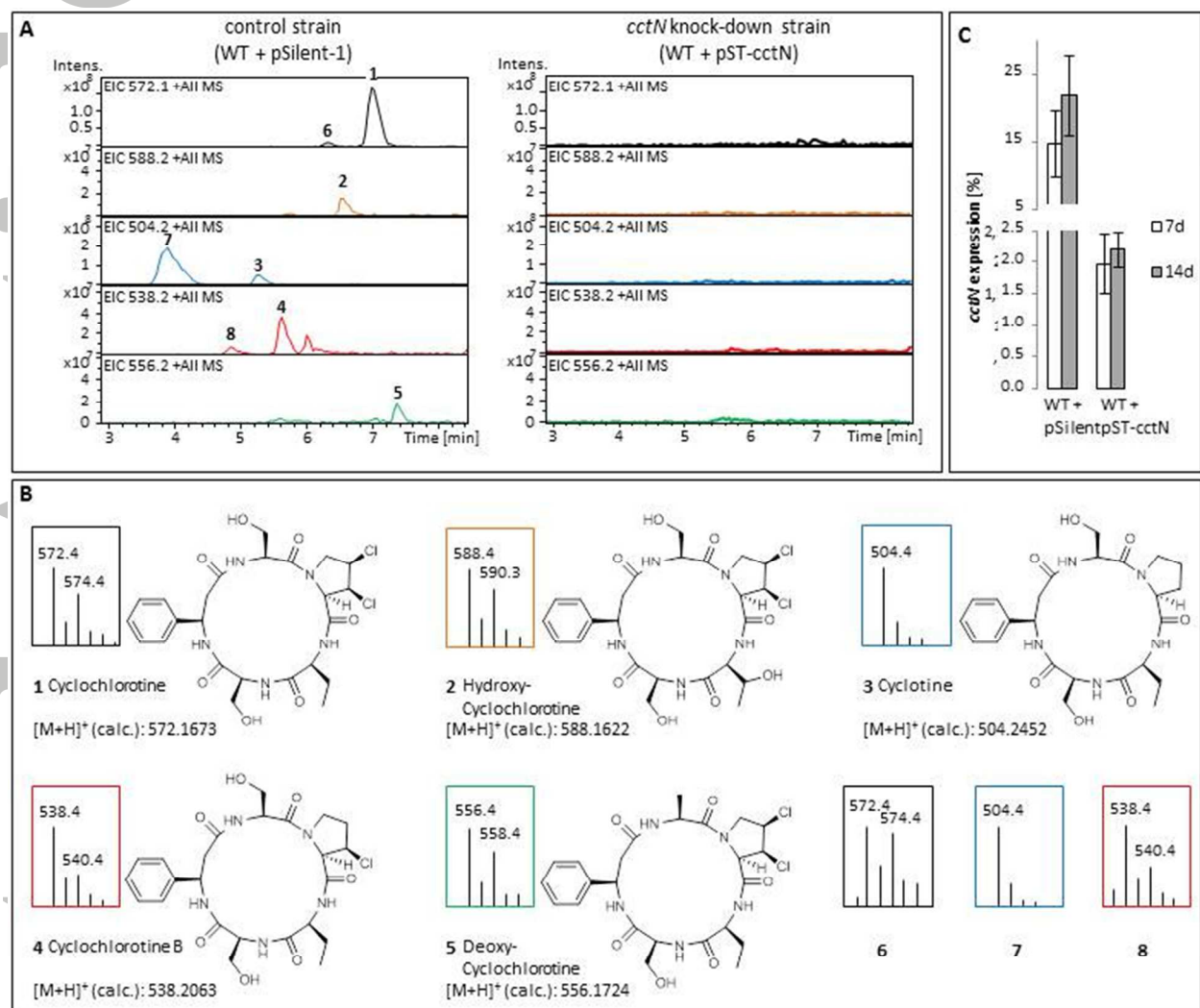


Figure 1. Production of cyclochlorotine and variants. A) Extracted ion chromatograms (EICs) of extracts from representative 3-week-old cultures analyzed by HPLC-MS show eight peaks that result from cyclochlorotine (**1**) and variants (**2-8**) in the *T. islandicus* control strain (WT + pSilent-1, left) but no compound peaks in the *cctN* knock-down strain (WT + pST-*cctN*, right). **B)** Mass spectra of compounds **1-8** from the control strain (WT + pSilent-1). The intensity ratios of the isotope peaks indicate the presence of two chlorine atoms (compounds **1**, **2**, **5**, and **6**), one chlorine atom (compounds **4** and **8**) and no chlorine (compounds **3** and **7**) in the molecules. Molecular structures of compounds **1-5** are indicated. **C)** Mean expression rate (± 1 SD) of the NRPS gene *cctN* (relative to the β -actin coding reference gene *act1*) measured by qPCR in a control strain (WT + pSilent-1) and a *cctN* knock-down strain (WT + pST-*cctN*) (n = 3).

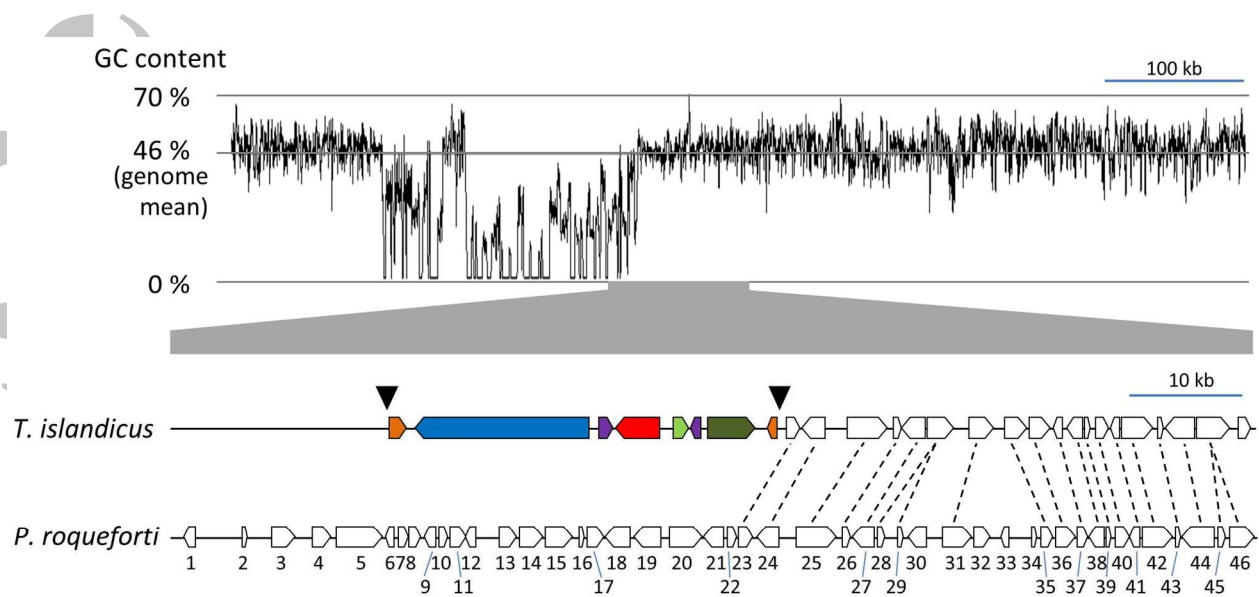


Figure 2. Borders and genetic environment of the *cct* cluster. GC content of the *T. islandicus* genomic region around the *cct* cluster is outlined above illustrating a dramatic drop of GC content left of the *cct* cluster compared to the genome mean of 46% (drops down to 0% are caused by missing sequence information). No genes are encoded in the low-GC region. Cluster borders are indicated with black arrow heads. The right environment is formed by a synteny block of highly conserved genes belonging to primary metabolism that occur in many ascomycetes including *Penicillium roqueforti*. Putative functions of the gene products are outlined as follows: 1 = exocyst complex component Sec10; 2, 3, and 4 = hyp. prot.; 5 = vacuolar protein sorting/targeting prot. 10; 6 = cytidine deaminase; 7 = peptidase C12; 8 = NmrA-like prot.; 9 = amino acid transporter arg-13; 10 = Cgr1; 11 = hyp. prot.; 12 = translation machinery-associated prot. 22; 13 = DUF1752; 14 = cell motility prot. ELMO; 15 = checkpoint prot. Rad24; 16 = hyp. prot.; 17 = UAA transporter; 18 = dynamin-related prot.; 19 = neutrophil cytosol factor 2 p67phox; 20 = AMP deaminase; 21 = hyp. prot.; 22 = DUF3752; 23 = porphobilinogen synthase; 24 = Zn(2)-C6 DNA-binding prot. ; 25 = exportin-1; 26 = phosphatase inhibitor; 27 = shugoshin; 28 = DASH complex; 29 = hyp. prot.; 30 = cytochrome P450; 31 = mitochondrial GPDH; 32 = amino acid transporter; 33 and 34 = hyp. prot.; 35 = coiled-coil domain containing prot. 25; 36 = GRIP; 37 = guanylate kinase; 38 = septin homolog; 39 = ribosomal prot. L32p; 40 = meiotic recombination prot.; 41 = kinetochore prot. spc25; 42 = RNA pol initiation factor; 43 = metal homeostasis factor ATX1; 44 = DUF2293; 45 = complex 1 LYR prot.; 46 = rubredoxin-type fold prot.

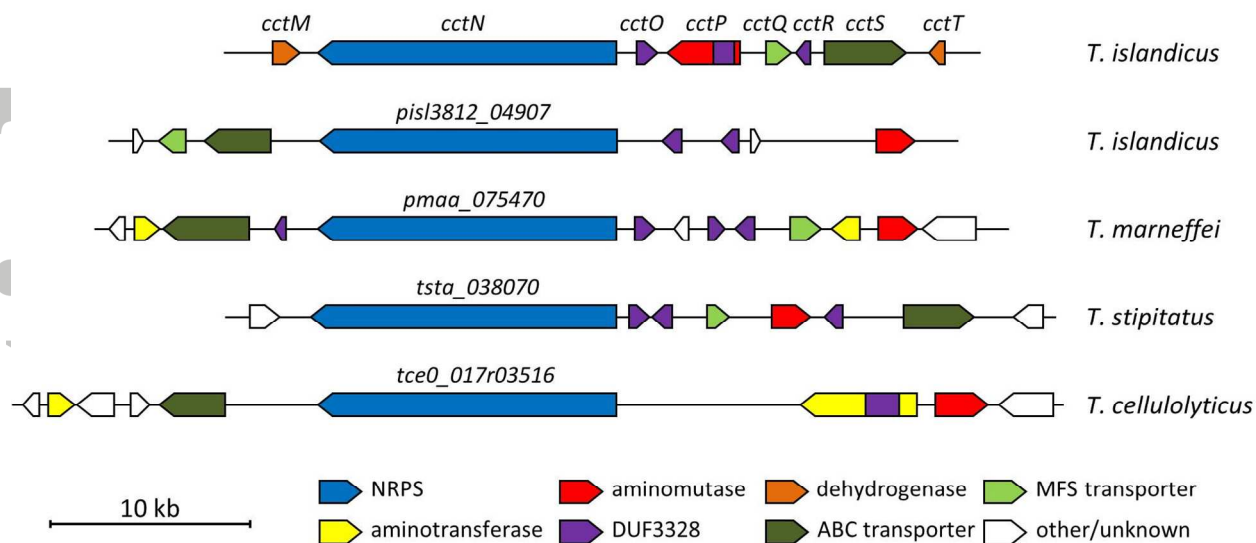


Figure 3. Genetic organization of the *cct* gene cluster and *cct*-like gene clusters. The *cct* gene cluster from *T. islandicus* encompasses eight genes *cctM* to *cctT*. A homologous *cct*-like gene cluster that is located on a different chromosome includes the NRPS gene *pisl3812-04907*. Further *cct*-like gene clusters were identified in the related species *T. marneffeii* ATCC 18224, *T. stipitatus* ATCC 10500 and *T. cellulolyticus* Y-94. Deduced functions of the genes are indicated.

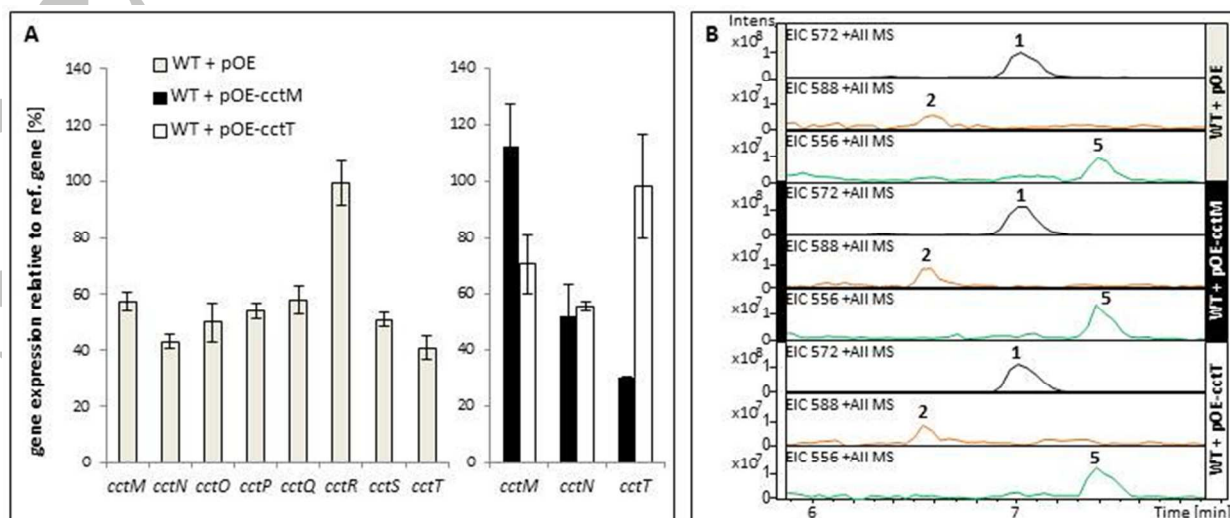


Figure 4. Expression of the *cct* cluster genes and production of selected cyclochlorotine variants. A) Mean gene expression rate (± 1 SD) relative to the β -actin coding reference gene *act1* measured by qPCR in non-shaking cultures ($n = 3$). In the control strain (WT + pOE) expression of all *cct* cluster genes is shown. In the strains overexpressing the putative dehydrogenase genes (WT + pOE-cctM and WT + pOE-cctT) expression of only three selected genes is shown. **B)** Extracted ion chromatograms (EICs) of extracts from one representative culture, each (WT + pOE, WT + pOE-cctM and WT + pOE-cctT) showing no differences in mass peaks of cyclochlorotine (1), hydroxy-cyclochlorotine (2) and deoxy-cyclochlorotine (5), respectively.

Accepted

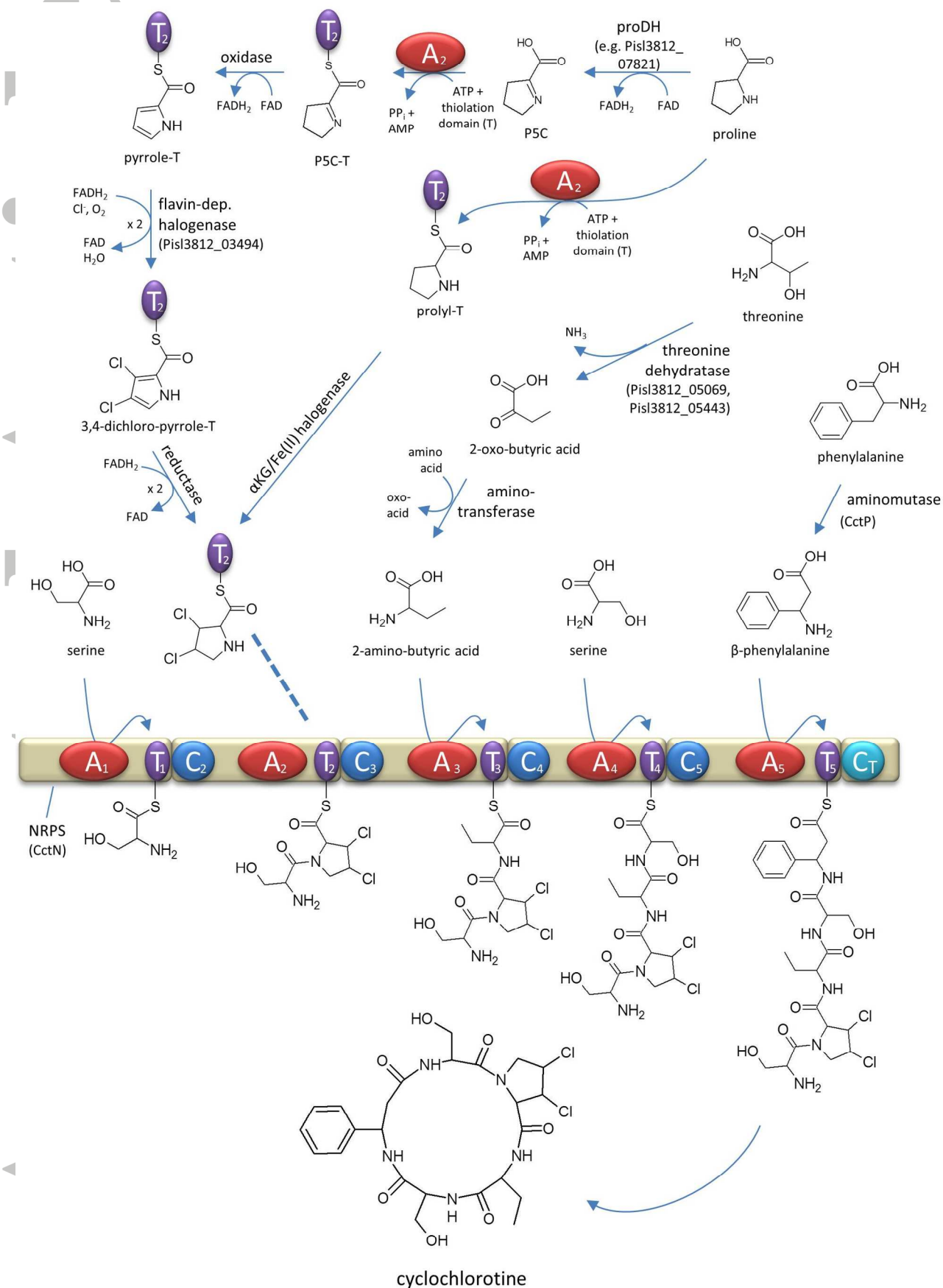


Figure 5. Model for cyclochlorotine biosynthesis. The upper part proposes pathways for the non-proteinogenic amino acid building blocks. Where applicable, candidate enzymes are indicated. Synthesis of 3,4-dichloroproline most likely is achieved on a carrier protein (e.g. T domain of module 2), and two possible pathways are shown. Halogenation occurs either directly by a 2KG/Fe(II)-dependent halogenase or by a flavin-dependent halogenase via a pyrrole carboxylic acid intermediate. The proline dehydrogenase (proDH) Pisl3812_07821 is considered to initiate the latter pathway by oxidation of free proline to form 1-pyrroline-5-carboxylic acid (P5C). Subsequent steps likely involve a P5C-carrier protein oxidase (according to Thomas et al., 2002; Mejean et al., 2010). In the lower part, NRPS domains are abbreviated as A for adenylation, T for thiolation, C for condensation and C_T for terminal condensation-like domains. The T domains are post-translationally modified by phosphopantetheinylation. Release and cyclization of cyclochlorotine is mediated by the C_T domain.

Accepted Article

Fungus	Nonribosomal peptide	Synthetase	Accession-Nr.	Domain	Specificity	Nonribosomal code (PheA numbering)									
						223	236	239	278	299	301	322	330	331	517
<i>Tolypocladium inflatum</i>	Cyclosporin	SimA	Q09164	A6	2Abu	D	A	W	F	H	A	V	A	Y	K
<i>T. islandicus</i>	Cyclochlorotine	CctN	CRG85572.1	A3		D	A	W	F	Y	G	A	P	L	K
<i>Trichoderma virens</i>	Paracelsin	Tex1	AAM78457.1	A13	Pro	D	V	L	F	C	G	L	I	C	K
<i>T. islandicus</i>	Cyclochlorotin	CctN	CRG85572.1	A2		D	I	H	D	V	S	A	I	C	K
<i>Aspergillus fumigatus</i>	Gliotoxin	GliP	Q5MBT9	A2	Ser	D	Y	N	T	Y	T	A	I	C	K
<i>T. islandicus</i>	Cyclochlorotine	CctN	CRG85572.1	A4		D	V	N	E	I	S	V	I	E	K
<i>T. islandicus</i>	Cyclochlorotine	CctN	CRG85572.1	A1		D	V	G	Q	E	S	T	I	T	K

Figure 6. Features of CctN A domains. Comparison of the nonribosomal codes of CctN A1, A2, A3, and A4 with those of substrate-assigned A domains of characterized fungal NRPS. The amino acid numbering corresponds to the PheA domain of the gramicidin S synthetase GrsA (Stachelhaus et al., 1999). Black: identical amino acids; gray: similar amino acids. The nonribosomal codes were extracted with the software NRPSpredictor2.