Nonribosomal peptide synthetases
Zwahlen, Reto Daniel

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

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

Publication date:
2018

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
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.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
CHAPTER VI

Prokaryotic MbtH like proteins stimulate secondary metabolism in the filamentous fungus *Penicillium chrysogenum*

Reto D. Zwahlen,¹ Carsten Pohl,¹ Roel A.L. Bovenberg,²,³ and Arnold J.M. Driessen¹,⁴

¹Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands
²Synthetic Biology and Cell Engineering, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands
³DSM Biotechnology Centre, Delft, The Netherlands
⁴Kluwer Centre for Genomics of Industrial Fermentations, Julianalaan 67, 2628BC Delft, The Netherlands

*manuscript in preparation*
Abstract

MbtH-like proteins or MLPs, are small (~10 kDa) proteins, which associate non-covalently with adenylation domains of nonribosomal peptide synthetases (NRPS). MLPs increase folding speed, stability, and improve kinetic properties of the interacting NRPS adenylation domain. MLPs are highly conserved amongst a wide range of prokaryotic species, however, they appear absent from all fungal species. Here we transform a selection of five prokaryotic MLPs into two industrially relevant P.chrysogenum strains in order to evaluate potential changes in nonribosomal peptide metabolite profiles. The over 60 obtained strains shown significantly changed dynamics of nonribosomal peptide formation, with the majority of the strains showing increased concentrations of intermediate and final products of the roquefortine, chrysogine and penicillin biosynthetic pathways. These findings demonstrate that the expression of bacterial MLPs can be highly instrumental to stimulate the secondary metabolism of various fungal NRPS biosynthetic gene clusters.
**Introduction**

*Penicillium chrysogenum* is a filamentous fungus, most prominently known as a producing host of β-lactam antibiotics [1–4]. In addition, *P. chrysogenum* generates numerous other secondary metabolites such as the mycotoxin roquefortine, the yellow pigment chrysogine, and others [5–10]. However, in high yielding penicillin producing strains, the production of undesired secondary metabolites was reduced by mutations introduced during several decades of classical strain improvement resulting in an increased flux of metabolic resources into penicillin production [11–12]. This re-distribution of metabolic fluxes, efficiently eliminated low abundant secondary metabolites under penicillin production conditions. As typical for secondary metabolism the core enzymes of biosynthetic gene clusters (BGC) are either nonribosomal peptide synthetases (NRPS), polyketide synthetases (PKS) or a fusion thereof, termed NRPS-PKS hybrid enzymes. These enzymes belong to a family of poly-catalytic molecular nano-machineries, composed of functionally distinct domains and modules [13–14]. The genome of *P. chrysogenum* encodes 10 NRPS, 20 PKS and 2 NRPS-PKS hybrid BGCs [15]. Enzymes and associated compounds have been identified for several of the BGCs [5–6;10;16–17]. The most prominent example is pcbAB, encoding the three modular NRPS L-δ-(α-aminoadipyl)-L-cysteinyl-D-valine synthetase (ACVS) which synthetizes L-δ-(α-aminoadipyl)-L-cysteinyl-D-valine (LLD-ACV), a tripeptide precursor of penicillin [18–20]. Another well described pathway includes RoqA, a NRPS synthesizing histidyltryptophanyl-diketopiperazine (HTD), which is further converted into a series of roquefortine and melagrin derivatives[5–6;21–22]. The two-modular NRPS ChyA, is yet another example of an NRPS, it is responsible for the formation of 2-aminobenzamide structures [23], the precursor for chrysogine biosynthesis. RoqA, ChyA and PcbAB are all multi-modular NRPS and, given their complex nature and typical low turn-over-rates, they may represent potential bottlenecks in their biosynthetic pathways [24–26]. NRPS have a high degree of functional conservation with a very strict modular setup, much like a molecular assembly line. Each module possesses a minimal domain requirement, comprising an adenylation domain (A), consisting of a N-terminal Acore part and a C-terminal ~110 residue Asub part, a thiolation domain (T) and a condensation domain (C), which selects, activates, transfers and condensates a specific substrate amino acid, respectively. In addition, optional domains could be part of the NRPS like domains for methylation and epimerization and thioesterification domains for final product release. To facilitate the complex biochemical process of nonribosomal peptide (NRP) formation, activating
and chaperoning factors are additionally required. Most notably are the 4'-phosphopanthetenyl-transferase (PPTase) [27], an essential thiolation domain activator and the MbtH-like proteins (MLPs), an A domain associating, catalytically inactive chaperone.

PPTases are essential for NRPS activation (Weber and Marahiel 2001)[28] and different classes of PPTase genes are conserved in one or more genomic copies across a majority of organisms. Sfp-type PPTases are essential for NRPS phosphopantetheination and show an outstanding degree of promiscuity in bacterial and fungal species, respectively [29–30]. MLPs however, are more elusive, as they are not essential to fungal NRPSs, though partially essential to prokaryotic NRPSs [31]. MLPs are relatively small, i.e., ~10kD, catalytically inactive and associate with the A domain of NRPS enzymes [32–33]. Initially discovered in *Mycobacterium tuberculosis*, it has become increasingly apparent that they play a key role in bacterial NRPS function as well as NRP formation [34–36]. MLPs often cluster among NRPS relevant genes, although one genomic copy is often sufficient to facilitate a generic NRPS activating function [37]. Upon binding, the MLP does not only enable or enhance the adenylation activity [38], but also increases soluble NRPS levels [39]. Despite the essential function it fulfills for some bacterial NRPSs, it appears completely absent from any fungal genome analyzed up to date. Despite a high level of MLP conservation [40], only two structures are available displaying an A-MLP interface in SlgNI (PDB:4GR4) and EntF (4ZXJ) [33;41]. Both structures highlight the same potential interaction region, differing only in the absence of an $A_{ub}$ domain in the structure of SlgNI as well as covalently linked versus un-linked MLP in SlgNI versus EntF, respectively. Thus, there seems to be a defined and rather conserved part of the $A_{core}$ surface which enables the dynamic association of different MLP variants.

In this study, we use five prokaryotic MLP proteins from various sources in order to analyze their potential for fungal NRPS activation upon introduction into the filamentous fungus *Penicillium chrysogenum*. Remarkably, the bacterial MLP proteins broadly impacts the activity of various NRPSs, showing changed kinetics and levels of secondary metabolites, including the primary NRPS products. This study shows for the first time that MLPs can be utilized to enhance secondary metabolism in a fungal host, providing a potentially new approach for host strain development and pathway engineering.
Materials and methods

Cloning, plasmids and culture conditions for E. coli

Cloning was performed using E. coli DH5α. Plasmid selection was conducted with 25 µg/ml zeocin for pIAT and pBAD plasmids and 15 µg/ml chloramphenicol for pACYCtac-MbtH (p15Aori; cat; pTac; lacIq), respectively. All cultures were grown using 2xPY (15 g/l bacto-tryptone, 10 g/l yeast extract, 10 g/l sodium chloride, pH 7.0) at 37 ºC and 200 rpm (New Brunswick Innova 44R, Eppendorf, Hamburg, Germany).

MLP were derived from 4 different sources as indicated in Table 1. Genes were codon optimized for P. chrysogenum, and ordered as synthetic DNA (gBlocks, IDT, Leuven, Belgium) Subcloning into the pIAT or pACYCtac vector was done using restriction with Ndel and Nsil. Assembled constructs were sequence-verified using Sanger sequencing (Macrogen, Amsterdam, The Netherlands). The N. lactamdurans ACVS gene was kindly provided by DSM-Sinochem Pharmaceuticals BV, (DSM biotechnology, Delft, The Netherlands). (Supplementary data — primers constructs)

Model creation and evaluation

To establish a basis for interaction probabilities of prokaryotic MLP and eu-karyotic NRPS, we utilized P.chrysogenum ACV synthetase (ACVS) as a model NRPS and focused on interactions concerning the adenylation domain (M2 A) of the second module. Adenylation domain boundaries were determined using a prediction by the conserved domain database (CDD) of the NCBI [42].

Table 1 — Overview of MLP variants in this study.
The displayed MLP variants were identified and in part characterized for their essential status to NRPS. A full list of all considered homologues can be found in chapter I.

<table>
<thead>
<tr>
<th>Name</th>
<th>Organism</th>
<th>Biosynthetic cluster</th>
<th>Accession</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>COM</td>
<td>Streptomyces lavendulae</td>
<td>Complestatine</td>
<td>WP_030951033</td>
<td></td>
</tr>
<tr>
<td>CDAI</td>
<td>Streptomyces coelicolor</td>
<td>Calcium dependent antibiotic</td>
<td>WP_003978376</td>
<td></td>
</tr>
<tr>
<td>Tcp13</td>
<td>Actinoplanes teichomyceticus</td>
<td>Teicoplanin</td>
<td>CAE53354.1</td>
<td></td>
</tr>
<tr>
<td>VEG8</td>
<td>uncultured</td>
<td>VEG</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>TEG</td>
<td>uncultured</td>
<td>TEG</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
The obtained sequences were subsequently used to generate models using RaptorX [43] and SWISS-MODEL [44]. Models were classified according to the model quality (global model quality estimation, GMQE) and were structurally aligned to SlgNI (PDB: 4GR4) [33] in a global and interface targeted fashion, respectively, using PyMOL (Schrödinger 2015). The resulting models, were then used for superpositioning and visualization in combination with 4GR4. Finally, a docking analysis was performed using 3DIANA [46].

Protein expression analysis

In order to evaluate interactions, *N. lactamdurans* ACVS was co-expressed with each MLP variant separately. Expression of the proteins was performed using the *E. coli* K12 derivative strain HM0079 [26]. Cells were grown at 37 °C and 200 rpm to an OD<sub>600</sub> = 0.6, transferred to 18 °C, 200 rpm for 1 h and subsequently induced using 0.2 % L-arabinose (pBAD — ACVS) and 0.3 mM IPTG (pACYCtac — MLP). Cells were collected 18 hours after induction by spinning at 3500 g for 15 minutes at 4 °C. After resuspension in lysis buffer (50 mM HEPES pH 7.0, 300 mM NaCl, 2 mM DTT, complete EDTA free protease inhibitor; Roche, Basel, Switzerland), cells were disrupted using sonication (6 s/15 s; on/off, 60 cycles, 10 μm amplitude) and cell-free lysate obtained by centrifugation for 15 minutes 13000 g and 4 °C. Enzymes were purified by means of his-tag affinity purification using Ni-NTA beads (Qiagen, Venlo, The Netherlands). Wash steps were performed using two column volumes wash buffer (50 mM HEPES pH 7.0, 300 mM NaCl, 20 mM imidazole) and a one-step elution using 5 bed volumes elution buffer (50 mM HEPES pH 7.0, 300 mM NaCl, 250 mM imidazole). Protein concentrations were determined using the BioRad DC assay kit (BioRad, Utrecht, The Netherlands). All fractions were analyzed on 5 % SDS-PAGE gels and stained using 0.025 % coomassie blue solution. Images were acquired using a FUJIFILM LAS-4000 scanner (Fujifilm, Tilburg, The Netherlands). Protein levels were additionally compared in-gel by 2-D densitometry using LAS-4000 AIDA software.

Transformation of MLP into *P. chrysogenum*

To introduce the MLPs into *P. chrysogenum*, pIAT-MLP plasmids were cut with NotI and SmaI, recovered from gel, purified and subsequently concentrated by desalting. Protoplasts were prepared as described [47]. Co-transformation was performed using the obtained MLP expression cassette together with
linearized pDONR221AMDS marker vector in a ratio of 1:10. After 5 to 7 days, colonies were screened for integration of the MLP expression cassette with primers (Sigma Aldrich, Zwijndrecht, The Netherlands) (Supplementary table 2) targeting the IPNS promoter and the AT terminator of each MLP expression cassette using colony PCR (Phire Plant Direct PCR Kit, Thermo Fisher Scientific, Bleiswijk, The Netherlands). Positive colonies were purified using three sporulation-selection transfer cycles.

**Small scale fermentations of P. chrysogenum**

*Penicillium* strains carrying the MLP genes were subjected to shake-flask cultivation experiments for up to 5 days. For preculture, spores stored on rice grains were inoculated in 25 ml YGG medium. After 24 h, the sporulated preculture was 10-fold diluted in a total volume of 30 ml penicillin production medium [48]. All cultures were grown at 25 °C and 200 rpm under semi-dark conditions in standard laboratory shakers (Innova44, New Brunswick Scientific, Nijmegen, The Netherlands). Sampling was performed on day 2 and day 5 after pre-culture transfer by aseptically harvesting 2 ml of culture from the shake flask followed by centrifugation at 4 °C and 14000 g for 10 minutes to pellet mycelium. From 4 strains expressing the MLP COM, mycelium was mixed with trizol reagent (Thermo Fisher Scientific, Bleiswijk, The Netherlands), transferred to screw-cap tubes containing glass beads (0.75–1 mm, Sigma Aldrich, Zwijndrecht, The Netherlands) and stored at −80 °C until analysis. The supernatant was subsequently filtered using a 0.2 µM PTFE membrane syringe filter (VWR, Amsterdam, The Netherlands) and stored at −80 °C up until analysis. Samples were reduced using 10 mM TCEP prior to analysis. The remaining volume of the culture after 5 days was additionally used for dry weight determination. In all experiments two biological and two technical replicates were use.

**qPCR analysis of gene copy number and NRPS expression in MLP strains**

Mycelium stored in Trizol was disrupted with a FastPrep FP120 system (Qbiogene, France). Total RNA was isolated from aqueous phase by chloroform/isopropanol isolation and purified using the Ambion Turbo DNA-free kit (Thermo Fisher Scientific, Bleiswijk, The Netherlands). RNA degradation was analyzed by electrophoresis using a 2 % agarose gel and concentrations were determined with a nanodrop ND-1000 (ISOGEN, Utrecht, the Netherlands).
The iScript cDNA synthesis kit (Bio-Rad, Veenendaal, The Netherlands) was used for reverse transcription with 500 ng total RNA as input. The primers used for qPCR of Pc21g21390 (pcbAB), Pc21g15480 (roqA) and Pc21g12630 (chyA) as well as the expression level of com are listed in (Table 2). The γ-actin gene (Pc20g11630) was used as internal standard for data normalization. As master mix for qPCR, the SensiMix SYBR Hi-ROX (Bioline Reagents, London, England) was used. All runs were performed on a MiniOpticon system (Bio-Rad). The following conditions were employed for amplification: 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, following an acquisition step. Raw ct data were exported and analysis of relative gene expression was performed with the 2-ΔΔCT method [49]. The expression analysis was performed with three technical replicates per strain.

The copy number of pcbAB and the MLP homologue COM was determined from isolated genomic DNA of a selection of strains, following the same protocols of qPCR as described above. For normalization and reference point of single copy, γ-actin was used.

**LC/MS analysis and evaluation of metabolite profiles**

Clarified culture broth samples from fermentations were subjected to LC/MS analysis. Analysis was performed by injecting 5 µl sample on aC18 column

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’ to 3’)</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>pIPNS-FW</td>
<td>GTCTGCCATTGCAAGGTATATATGGC</td>
<td>Verification of MLP-expression cassette integration</td>
</tr>
<tr>
<td>tAT-R</td>
<td>TAGTGACGGTGCAAGGTAAAGC</td>
<td>Verification of MLP-expression cassette integration</td>
</tr>
<tr>
<td>BPS-CN-FW</td>
<td>CTTGTCTCTTTAAGCGAGGAGGCC</td>
<td>Calculation of BPS copy number</td>
</tr>
<tr>
<td>BPS-CN-R</td>
<td>GAAATGAAAGACGGAGTCCGGTGAAAGC</td>
<td>Calculation of BPS copy number</td>
</tr>
<tr>
<td>act-CN-FW</td>
<td>ATGGAGGGATATGTTATTCAGTTGGAAG</td>
<td>Actin (reference for copy number/ normalization of qPCR data)</td>
</tr>
<tr>
<td>act-CN-R</td>
<td>TGCGGTGAACGATGGAAGGACC</td>
<td>Actin (reference for copy number/ normalization of qPCR data)</td>
</tr>
<tr>
<td>pcbAB-FW</td>
<td>CGATATCAGTCGATCGACTG</td>
<td>Calculation of ACVS copy number/ expression</td>
</tr>
<tr>
<td>pcbAB-R</td>
<td>CGCTTCGATACCTCGACC</td>
<td>Calculation of ACVS copy number/ expression</td>
</tr>
<tr>
<td>roqA-FW</td>
<td>CTTGATGGATCGACGGAGG</td>
<td>roqA expression</td>
</tr>
<tr>
<td>roqA-R</td>
<td>CTGTGAGAGGGCTTGTGAGTA</td>
<td>roqA expression</td>
</tr>
<tr>
<td>chyA-FW</td>
<td>GAGCCAACTCTGTGTCTCAG</td>
<td>chyA expression</td>
</tr>
<tr>
<td>chyA-R</td>
<td>CAGGGCAATTGCTCATTCTG</td>
<td>chyA expression</td>
</tr>
</tbody>
</table>
Prokaryotic MbtH like proteins stimulate secondary metabolism in the filamentous fungus Penicillium chrysogenum

(Shim pack XR-ODS 2.2; 3.0×75 mm, Shimadzu, Japan) coupled to a LC/MS Orbitrap device (Thermo Fisher Scientific, Bleiswijk, The Netherlands) operated in positive ionization mode. A gradient program with water (A), acetonitrile (B) and 2 % formic acid in water (D) was run; 0 min; A 90 %, B 5 %, C 5 %; 4 min, A 90 %, B 5 %, C 5 %; 13 min, A 0 %, B 95 %, C 5 %; 16 min A 0 %, B 95 %, C 5 %; 16 min, A 90 %, B 5 %, C 5 %; 20 min A 90 %, B 5 %, C 5 % at a flow rate of 0.3 ml min⁻¹. Two technical replicates were recorded for each sample. Available standards were used to identify peaks according to retention time and accurate mass. A complete list of all metabolites and intermediates which were identified is summarized in Supplementary table 1.

Results

MLP selection, model evaluation and in vitro analysis

In order to examine if a prokaryotic MLP has the potential to interact with eukaryotic NRPS, possible co-complexes were modeled based on the SlgNI structure in complex with various MLP. For this purpose, 5 MLPs were selected, which are derived from 4 sources. TEG and VEG8, were extracted from metagenomic data and Tcp13, CDAI and COM from various prokaryotic sources (Table 1). Moreover, 7 adenylation domains were selected, derived from the three fungal NRPS enzymes ACVS (3), RoqA (2) and ChyA (2), associated with the penicillin, roquefortine and chrysogenin pathways, respectively. The two platforms RaptorX and SWISS-MODEL served as a basis for the creation of over 100 in silico models, using 23 PDB templates of 18 proteins. General model quality estimates (GMQE 0–1) for adenylation domain models are in a range from 0.44 for ChyA up to 0.71 for ACVS and 0.59–0.82 for the MLPs (Supplementary table 2). The obtained models were structurally aligned to SlgNI (PDB: 4GR4) using PyMOL, in a global as well as interface directed fashion. Due to the absence of an Asub domain in 4GR4, only the ~340 Acore residues were aligned, covering the hypothetical A-MLP interface. The process allowed for a precise alignment of up to 305 (340) Cα carbons in ACVS models and 49 (61) in COM MLP models, at RMSD values of <1 (Supplementary table 2).

All sequences of the adenylation domains, MLP and the corresponding templates were subsequently aligned and the secondary structural data extracted and overlaid (Figure 1A/B). Hereby all hypothetically interface forming (IF) and interactive residues (IA) were determined, according to the SlgNI structure, which counts 29 IF and 6 IA on the MLP domain and 32 IF and
The alignment reveals that the selected MLP are conserved on 20–23 (29) IF positions as well as for all 6 (6) IA residues (Figure 1C). With respect to the adenylation domains, 8–13 (32) IF and 4–10 (12) IA are conserved (Figure 1D). The region resolving around α-helix 11 and β-sheet 22 (4GR4) on the adenylation domains, furthermore includes two core alanine residues at position 428 and 433, centered at the interface area. Although A433 seems to be completely absent from all investigated domains, ACVS M2 and RoqA M2 both contain A428 nonetheless. Especially ACVS M2 A appears to have a high similarity to SlgNI. With respect to the MLPs, not only the A428/A433 interacting partners Ser23 and Lys24 are conserved, but all MLPs contain the minimal MLP functionality motif (Figure 1D). Overall, the MLPs structures are highly conserved, exhibiting a ubiquitously seen structural pattern, composed of three anti-parallel β-sheets and one α-helix. The most significant differences can be found in the disordered N-terminus, which is however not thought to be involved in A-MLP association. A more diverse secondary structure topology is observed in the adenylation domains. Even though the upstream region of α-helix 11 is fairly conserved, there is a 15 amino acid stretch downstream of β-sheet 22, containing a short insertion, only observed in the P. chrysogenum adenylation domains. However, this does not involve any hypothetical interacting residues. Compared to the given templates, the local structure seems well preserved, which is further underlined by the comparison of ACVS M2 A and COM MLP with SlgNI (Figure 1A). In spite of the absent A_sub domain in SlgNI, the remaining analogous structures align well and the local interface alignment reveals a good positioning of α-helix 11 and β-sheet 22 (Figure 1B) at the core interface.

In order to test the presumed A-MLP interaction findings, in vitro co-purification experiments were conducted, employing A-domains of the Nocardia lactamdurans ACVS. The N. lactamdurans ACVS is a functionally conserved ACVS with a high homology to the P. chrysogenum ACVS. However, the latter is poorly expressed in E. coli (Unpublished data/Chapter II). Nonetheless, neither of these proteins has known interactions with MLPs. Modules were determined using the Conserved Domain Database (CCD) prediction tool or were split at the center of the inter-module linkers (Unpublished data/Chapter II). Subsequently, the modules were co-expressed along with one of the MLPs, purified using a C-terminal bound HIS-tag and visualized using SDS-PAGE. Despite the considerable high level of MLP conservation, a distinct co-purification pattern emerged (Figure 2). COM, TEG and Tcp13 co-purify significantly with their respective co-expressed ACVS module while CDAI nor VEG8 failed to co-purify with the ACVS module or full length protein (Supplementary Data).
Prokaryotic MbtH-like proteins stimulate secondary metabolism in the filamentous fungus Penicillium chrysogenum.
**MLP strain selection, analysis and evaluation**

To examine the impact of a set of bacterial MLPs in vivo, the genes of all five proteins were transformed into *P. chrysogenum* using a non-targeted genome insertion strategy. Two strains served as basis for this purpose, *P. chrysogenum* DS47274 and DS17690. Strain DS17690 emerged from the classical strain improvement program (CSI) to optimize penicillin production and contains eight copies of the penicillin biosynthetic gene clusters. Strain DS47274 was derived from strain DS17690 and differs only by containing a single biosynthetic gene cluster, but is further genetically identical. The genomic integration of the different MLP variants was determined using colony PCR (Supplementary figure 3 and 4).

MLP transformation yielded 62 strains, whereof 32 originated from DS17690 and 30 from DS47274 as parental strain, respectively. Shake flask cultivations were subsequently performed and the supernatant was subjected to secondary metabolites profiling using LC-MS (Supplementary table 1). The metabolites of three pathways were analyzed in more detail, i.e., penicillin (ACVS — Pc21g21390), chrysogine (ChyA — Pc21g12630) and roquefortine (RoqA — Pc21g15480), respectively. A total of at least 26 metabolites is known to be associated with these pathways. Metabolite levels represent peak area values, normalized for dry weight, expressed in a relative ratio to the parental strain and were determined after two and five days of the fermentation.

**Figure 1 — Structural model comparison and alignment of SlgNI with ACVS and COM MLP.**

Alignment of SlgNI (PDB: 4GR4 — red) structure with the most refined A domain model retrieved from SWISS Model (GMQE = 0.80), showing ACVS M2 A (rose-pink) (A). Indicated are MLP domain (green), MLP-A linker region (blue) as well as the two adenylation domain parts (Asub and Acore). The structural alignment of SlgNI — ACVS M2 A indicates a structural conserved interface, shown in an enhanced representation (B). The core binding structure α-11 and β-22 as well as the important interactive residues S23, L24 in addition to A428 and A433 are labelled and furthermore also indicated in the A domain and MLP alignments. The alignments cover all relevant residues for interface formation IF (*) and interaction IA (X) as compared to SlgNI (PDB:4GR4). Conservation of IA and IF is indicated in the utmost right column. The list includes the five MLPs in this study (D), the seven A domains of 3 NRPS (C) alongside all relevant templates used for the generation of the models, including common gene names and PDB entry. In addition, the essential MLP motif is shown in A [37]. Sequences in C are further annotated according to fungal origin (*) as well as MLP dependency (**) according to STRING [51]. A second layer containing the extracted secondary structures of all proteins was added, showing α-helices (red), β-sheets (green) and 3_10-helices (yellow).
Results

LC-MS analysis of total ion chromatograms revealed large scale changes in the metabolic profile of *P. chrysogenum* DS17690 and DS47274 (Figure 3). These concerned in particular chrysogine and the associated metabolite chrysogine 13 as well as the penicillin precursor LLD-ACV and further derivatives of the penicillin pathway. In addition to these NRPS products, a series of lower abundant products appear, including all key metabolites of the three investigated clusters, penicillin, chrysogine and roquefortine and a series of unidentified compounds. With respect to the data of all 62 strains in both genetic backgrounds, product levels associated with all three pathways

Figure 2 — Co-purification of ACVS M2 with various MLP proteins.
ACVS M2 (CAT domain composition) and MLPs were co-expressed and purified using the ACVS M2 C-terminal his-tag. The NRPS module as well as the MLP are indicated with red arrows. The distinct co-purification pattern favoring COM, Tcp13 and TEG emerges clearly. M = marker; CFE = cell free extract; E = elution (2-fold (II) or 5-fold (V) diluted).
Prokaryotic MbtH-like proteins stimulate secondary metabolism in the filamentous fungus Penicillium chrysogenum.

Figure 3 — Example of total ion current (TIC) of DS47274 (A) and DS17690 (B) with and without the COM MLP.

The region displayed spans from 4–12 minutes of the 20 minutes total run, representing the majority of all investigated metabolites, as well as predicted, but non-assigned metabolites. Indicated in numerical order are (1) LLD-ACV, (2) Chrysogine 13, (3) Chrysogine, (4) Monocerin*, (5) Penicilloic G acid (Penicillin G degradation product), (6) Penicillin G.
Results

increased (Figure 4). This trend is supported by median levels of >1 for a variety of pathway associated products, underlining the consistency of the effect across a majority of the MLP expressing strains. However, multiple strains show varying effects, including lower product levels or even the abolishment.
of any secondary metabolic activity in one strain (Veg8 — Figure 4). This is likely a side effect of the non-targeted genomic insertion strategy of the MLPs. Moreover, a marked difference between the two strain backgrounds DS47274 and DS17690 arises, showing overall stronger effects in lower penicillin producing DS47274 strains as compared to DS17690 strains. DS47274 strains expressing MLPs showed positive effects on the penicillin, roquefortine and chrysogine biosynthesis, while DS17690 borne strains experience the strongest effects on chrysogine derived metabolites. The NRPS targeting effect is highly evident with the core NRPS derived metabolites LLD-ACV and HTD, which are significantly increased in many DS47274 strains and are also stimulated but to a lesser degree in DS17690 strains. In addition, a series of low abundance metabolites of the roquefortine biosynthesis (Roq12, Roq13, Roq15, Roq16) seem to experience very high relative gains (>1000-fold), but it should be stressed that these are minor compounds in the WT strains. Increased levels of penicillin and roquefortine-related compounds were already evident after two days, while the chrysogine-related compound that already is detected early in fermentation does not show an altered time-dependent trend (Figure 4).

The results were further analyzed according to the specific transformed MLP (Supplementary figure 1). The 62 MLP expressing strains, represented as — (DS17690/DS47274) the following can be grouped into 22 strains containing COM (10/12), 10 containing CDAI (4/6), 12 Tcp13 (7/5), 10 TEG (6/4) and 8 VEG8 (5/3) — (DS17690/DS47274). Due to the experimental setup that involves random genomic integration of the MLP genes, it is not possible to clearly define differences in the abilities of the various MLPs to stimulate the different pathways. However, strains containing Tcp13, TEG and COM seem to be more stimulated than the CDAI and VEG8 bearing strains. An exceptional and coherent pattern is only observed with the COM MLP transformed strains, which experience a consistent increase in the chrysogine metabolites chrysogine B and N-pyrovoylantranilamid after 2 days of fermentation.

**COM MLP elevates the production of LLD-ACV and HTD in *P. chrysogenum***

Another transformation using the COM MLP was conducted and the obtained strains, 3 DS17690 and 6 DS47274, were further characterized for *pcbAB* and *com* copy numbers as well as *com* expression at 2 and 5 days of the fermentation (Figure 5). All strains contained one copy of the *com* gene, except for DS47274 7.1 and 8.1, which contain 5 and 8 copies of *com*,
Results

Figure 5 — Overview and characterization of obtained strains with the COM MLP. Determination of pcbAB (A) and MLP (B) copy number as well as com expression levels in selected strains (C). Strains are displayed in a separated manner, showing DS17690 (8xpcbAB — left) and DS47274 (1xpcbAB — right) strain backgrounds apart. Expression levels are additionally divided into two days (light grey) and five days (dark grey) after the start of the fermentation. Expression levels were normalized using α-actin and compared in a relative manner, setting the lowest value (7–3) to 1.
Prokaryotic MbtH-like proteins stimulate secondary metabolism in the filamentous fungus Penicillium chrysogenum

respectively (Figure 5B). Furthermore, DS47274 strains consistently contain one copy of pcbAB, while DS17690 strains show 6 to 8 copies (Figure 5A). The expression profile of the strains 1.2; 2.4 (8x pcbAB) and 7.3; 8.1 (1x pcbAB), indicates that the gene is expressed, after 2 and 5 days and that the expression increases with the copy number (Figure 5C). The strains were subsequently subjected to small scale fermentations and metabolite profiling (Figure 6). Again, various effects in all three secondary metabolite clusters can be observed. In particular, the penicillin and roquefortine cluster appear to be well transferable.

Figure 6 — Heat map of relative fold changes of the COM MLP strain set.

The data of the nine strains is separated by strain background, day and observed cluster. Data is derived of two independent experiments using two biological and two technical replicates. The effects targeting the penicillin and roquefortine cluster appear to be well transferable.
(Supplementary figure 2), the COM-dependent effect on penicillin production levels mostly concerns an earlier onset of production, i.e., after 2 days of fermentation, whereupon it stabilizes to the WT levels after 5 days. A similar behavior is observed with the chrysogine related metabolites, whereas the production of roquefortine-related compounds is stimulated in particular in the later stages of fermentation, i.e., after 5 days. The RoqA and ACVS derived products HTD and LLD-ACV increased up to 4.5- and 4.7-fold, and 2.3 and 2.4-fold in strain DS47274 and DS17690, respectively (Figure 7B). ACV levels were in particular impacted during the earlier stages (2 days) of fermentation, whereas HTD levels increased more dramatically during the later stages (5 days). The levels of ACV were additionally quantified and corrected for the cell dry weight (CDW) (Figure 7A). Extracellular levels increased from 0.144 µmol/ g(CDW) in DS47274 WT up to 0.700 in the DS47274 COM MLP strain 2-4_12 at day 2 and 0.608 to 1.476, in DS17690 WT and DS17690 COM MLP strain 2-4, respectively. All these observed effects are not related to

Figure 7 — LLD-ACV and HTD levels in COM MLP strains.
Levels of the secreted NRPS derived compounds LLD-ACV (ACVS) and HTD (RoqA) in R.chrysogenum strains DS 47274 and DS 17690. Figures are clustered based on metabolite and strain background and day 2 (◦) and day 5 (□) values shown in separate bars. Numbers are derived from two biological and two technical replicates, error bars indicate the standard deviation.
changes in expression of the NRPSs, since the changes measured by qPCR are neglectable. (Supplementary figure 4)

Discussion

MbtH like proteins (MLP) improve the solubility and catalytic properties of NRPS complexes or NRPS adenylation domains, respectively. Even though MLPs are not intrinsic to fungal species, fungi contain a multitude of secondary metabolite gene clusters harboring NRPS. Due to a significant degree of structural conservation among MLPs as well as NRPS adenylation domains, these NRPS may serve as a potential interaction targets for prokaryotic MLP variants. Here, 5 MLPs from various sources (Table 1) were used alongside the adenylation domains of the P.chrysogenum ACVS (PcbAB) to create a series of structural models, allowing for an evaluation of potential NRPS — MLP interactions. Herein, the models were derived from SlgNI (PDB: 4GR4), the first resolved structure showing an adenylation domain — MLP interaction. Despite a high functional conservation of all known adenylation domains, only prokaryotic domains partially rely, or even strictly require the presence of a MLP. Despite significant changes in the hypothetical interface composition, many directly interacting residues are conserved in the analyzed fungal adenylation domains (Figure 1). In conjunction with the strict MLP conservation and the very concise NRPS A — SlgNI — MLP superposition shown in Figure 1A, this data hints, that also fungal NRPS are potentially capable of interacting with bacterial MLPs. To address this experimentally, we used the N. lactamdurans ACVS, a homologue to the P. chrysogenum ACVS, though with significantly improved heterologous expression properties in E. coli. Each of the three PcbAB modules were found to interact with the MLPs COM, TEG and Tcp13, while CDAI and VEG8 appeared less effective. Surprisingly, the chosen MLPs do not show a high divergence in their amino acids sequences nor any specific motifs for the two classes of interacting and non-interacting MLPs [41]. Only detailed structural studies of A-MLP complexes may unveil precise mechanisms as to why CDAI and VEG8 performed poorly.

Subsequently, all five MLP variants were tested in vivo for their impact on secondary metabolite formation. For this purpose, two P.chrysogenum strain backgrounds were used DS47274 and DS17690, containing one and eight penicillin biosynthetic clusters, respectively. The resulting strains were screened for genomic MLP integration and subjected to small scale fermentations and metabolite profiling. The obtained profiles revealed a truly diverse effect. Not only could we observe changes in the levels of cluster
related compounds in penicillin, chrysogine and roquefortine biosynthetic pathways, but also other unrelated metabolites appear to be affected by the MLP presence (Figure 3). Even though a majority of strains seem to experience beneficial effects, some strains underperform relative to the respective wildtype strain. Due to the random genomic integration transformation approach, however, this seems to be an effect largely caused by disruptive insertions of the MLPs which can be omitted by a targeted integration in the future. Nonetheless, with the NRPS derived products LLD-ACV (PcbAB) and HTD (RoqA), a general increase in metabolite levels is observed. The effect, especially on LLD-ACV, strongly depends on the strain background and appears lower in the high penicillin yielding DS17690 strains as compared to single copy DS47274 strains suggesting that the extent of stimulation is limited by the secondary metabolic capacity of the respective strain.

In an attempt to further determine more specific effects of the MLPs, the strains were separated by MLP variant and strain background. Even though some specific effects were seen for N-Pyr and ChrB at Day 2 in the DS17690 COM MLP strains, overall the effect appears not to be limited to a specific MLP. Because of the coherent effects observed with a variety of NRPS derived products and related metabolites, it is evident that prokaryotic MLPs interact in a generic fashion with fungal NRPS. Novel strains bearing the ambivalent COM MLP were created resulting in 9 independent COM MLP containing strains. Further characterization of com and pcbAB copy number as well as com expression levels (Figure 5), revealed a single copy of com MLP with the exception of two DS47274 strains (Figure 5B). Despite showing significantly higher expression levels after 5 days, there was no further additional effect linked to copy number nor expression levels. MLP proteins are highly conserved and have a high intrinsic promiscuity for different NRPS [50], thus it seems that no additional benefits result from higher intracellular levels. The expression of the NRPS

Conclusion and perspectives

Using five selected bacterial MLPs expressed in two strains of the filamentous fungi Penicillium chrysogenum, we have successfully shown that fungal secondary metabolism can be effectively stimulated by prokaryotic MLPs. The presence of the MLPs not only increased metabolite levels in three individual NRP pathways, but consistent effects were observed across multiple strain backgrounds, as well as with all MLP variants used in this study. This suggests a transferable effect, applicable to a wide range of other NRPS
Prokaryotic MbtH-like proteins stimulate secondary metabolism in the filamentous fungus Penicillium chrysogenum.

Acknowledgments

We would like to thank the BE-Basic consortium for the financial support of this project. Furthermore, we thank DSM Sinochem Pharmaceuticals for their technical and intellectual support in this study. Carsten Pohl was supported by the European Commission Marie Curie Initial Training Network Quantfung (FP7-People-2013-ITN, grant no. 607332).

References


Prokaryotic MbtH-like proteins stimulate secondary metabolism in the filamentous fungus Penicillium chrysogenum.


35. Tatham E, Sundaram Chavadi S, Mohandas P, Edupuganti UR, Angala SK, Chatterjee D, et al. Production of mycobacterial cell wall glycopeptidolipids requires a member of the MbtH-like
References


45. Schrödinger, LLC. The {PyMOL} Molecular graphics system, Version-1.8. 2015 Nov.


49. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the
Prokaryotic MbtH like proteins stimulate secondary metabolism in the filamentous fungus Penicillium chrysogenum.


Supplementary data

Day 2
n=10

Day 5
n=4

TCP13

n=7

TGF

n=6

VEGF

n=5

Penicillin Chrysogenin Roquefortine
Supplementary figure 1 — First experiment separated by MLP. Separated view, displaying all strains of one MLP, divided by strain background and day of fermentation. The number n indicates the amount of different strains which were analyzed. Each point represents the average of data derived from 3 experiments using each, two biological and two technical replicates.
Supplementary data

Day 2

Day 5

Day 2 & 5

DS 17690

Relative fold change

DS 47274

Penicillin Chrysogine Roquefortine

A

B

C

n=32

n=30

A B C

0 1 2 3 4 5 6 7 8 9 10

egnahc dl of evit al eR Day 5

Day 2 & 5

n= 32

n= 30
Prokaryotic MbtH like proteins stimulate secondary metabolism in the filamentous fungus Penicillium chrysogenum.

Supplementary figure 3 — Colony PCR results confirm integration of MLP expression cassettes.

Colony PCR was performed on various colonies that were transformed with MLP expression cassettes using pIPNS-FW and AT-R. The integration of the helper protein expression cassette was confirmed in most of the colonies. The results suggest that only few colonies were not co-transformed with the helper protein construct. Expected product sizes: BPS = 1353 bp; COM = 1353 bp; CDAI = 1344 bp; Tcp13 = 1338 bp; TEG = 1341 bp; VEG8 = 341 bp. Marker: 1kb ladder (Fermentas). As positive control (+), pIAT-COM was used with primers used for colony PCR.

Supplementary figure 2 — Overview second round of COM MLP strain analysis.

The values indicated represent relative fold changes in peak area of a dedicated metabolite. Peaks were identified prior according to accurate mass and retention time. Numbers are further normalized for dry weight of the cultures and unchanged levels (-1) marked (- - - -). Strains are separated according to DS17690 (top) and DS47274 (bottom) strain background, day of fermentation, 2 (A) and 5 (B) and finally overlaid using day 2 (●) and day 5 (●) values (C). Individual strains are represented as DS17690 WT (●), 1-2 (X), 1-5 (+), 2-4(Δ) and DS47274 WT (●), 2-4_12 (●), 2-4_14 ( ), 7-1 (○), 7-3 (□), 8-1 (X) and 8-4 (Δ). Each point represents the average of data derived from 3 experiments using each, two biological and two technical replicates.
Supplementary figure 4 — expression changes in NRPS genes measured by qPCR.

A qPCR was performed to measure the expression of *pcbAB*, *chyA* and *roqA* in 4 strains that were transformed with the COM expression cassette. No strong change in gene expression was observed in either strain, suggesting that the effect is entirely kinetic and does not stimulate expression changes.
Supplementary table 1 — Measured metabolites and their corresponding biosynthetic pathways.

Compound names, Monoisotopic Ionized masses (M/Z [H]+), Molecular composition and retention times for the applied LC program are indicated.

<table>
<thead>
<tr>
<th>Compound</th>
<th>M/Z [H]+ (ppm)</th>
<th>Formula</th>
<th>RT (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Penicillin</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-Aminopenicillanic acid</td>
<td>217.06</td>
<td>C8H12N2O3S</td>
<td>1.84</td>
</tr>
<tr>
<td>6-(L-α-Aminoadipyl)-L-cysteinyl-D-valine</td>
<td>364.15</td>
<td>C14H25N3O6S</td>
<td>7.80</td>
</tr>
<tr>
<td>Isopenicillin N</td>
<td>360.12</td>
<td>C14H21N3O6S</td>
<td>2.44</td>
</tr>
<tr>
<td>Penicillin G</td>
<td>335.11</td>
<td>C16H18N2O4S</td>
<td>10.47</td>
</tr>
<tr>
<td>N-pyrooylanthranilamid</td>
<td>207.08</td>
<td>C10H10N2O3</td>
<td>7.48</td>
</tr>
<tr>
<td>Chrysogine VII</td>
<td>277.08</td>
<td>C13H12N2O5</td>
<td>7.94</td>
</tr>
<tr>
<td>Chrysogine XI</td>
<td>338.13</td>
<td>C15H19N3O6</td>
<td>8.17</td>
</tr>
<tr>
<td>Chrysogine XII</td>
<td>337.15</td>
<td>C15H20N4O5</td>
<td>7.07</td>
</tr>
<tr>
<td>Chrysogine XIII</td>
<td>295.11</td>
<td>C13H14N2O6</td>
<td>7.95</td>
</tr>
<tr>
<td>Chrysogine XIV</td>
<td>276.10</td>
<td>C13H13N3O4</td>
<td>8.28</td>
</tr>
<tr>
<td>Chrysogine XV</td>
<td>336.11</td>
<td>C15H18N3O6</td>
<td>9.43</td>
</tr>
<tr>
<td>Chrysogine XVI</td>
<td>413.15</td>
<td>C20H20N4O6</td>
<td>8.90</td>
</tr>
<tr>
<td>Chrysogine</td>
<td>191.08</td>
<td>C10H10N2O2</td>
<td>8.24</td>
</tr>
<tr>
<td>Chrysogine B</td>
<td>250.12</td>
<td>C13H15N3O3</td>
<td>7.99</td>
</tr>
<tr>
<td>Chrysogine C</td>
<td>294.11</td>
<td>C13H15N3O5</td>
<td>7.95</td>
</tr>
<tr>
<td><strong>Chrysogine</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HTD</td>
<td>324.15</td>
<td>C17H17N5O2</td>
<td>5.51</td>
</tr>
<tr>
<td>DHTD</td>
<td>322.13</td>
<td>C17H15N5O2</td>
<td>6.62</td>
</tr>
<tr>
<td>Roquefortine C</td>
<td>390.19</td>
<td>C22H23N5O2</td>
<td>9.49</td>
</tr>
<tr>
<td>Roquefortine D</td>
<td>392.21</td>
<td>C22H25N5O2</td>
<td>8.94</td>
</tr>
<tr>
<td>Roquefortine F</td>
<td>420.20</td>
<td>C23H25N5O3</td>
<td>9.75</td>
</tr>
<tr>
<td>Roquefortine M</td>
<td>422.18</td>
<td>C22H23N5O4</td>
<td>8.92</td>
</tr>
<tr>
<td>Roquefortine N</td>
<td>440.19</td>
<td>C22H25N5O5</td>
<td>8.19</td>
</tr>
<tr>
<td>Glandicoline A</td>
<td>404.17</td>
<td>C22H21N5O3</td>
<td>9.72</td>
</tr>
<tr>
<td>Glandicoline B</td>
<td>420.17</td>
<td>C22H21N5O4</td>
<td>8.93</td>
</tr>
<tr>
<td>Meleagrine</td>
<td>434.18</td>
<td>C23H23N5O4</td>
<td>9.17</td>
</tr>
<tr>
<td>Neoxaline</td>
<td>436.19</td>
<td>C23H26N5O4</td>
<td>9.17</td>
</tr>
</tbody>
</table>

Prokaryotic MbtH-like proteins stimulate secondary metabolism in the filamentous fungus Penicillium chrysogenum.