New natural products from Actinobacteria:

Potential and constraints of heterologous expression in *Streptomyces* hosts

Dissertation

zur Erlangung des Grades

des Doktors der Naturwissenschaften

der Naturwissenschaftlich-Technischen Fakultät

der Universität des Saarlandes

von

Constanze Lasch

Saarbrücken

2022

Tag des Kolloquiums:	9. Dezember 2022
Dekan:	Prof. Dr. Ludger Santen
Berichterstatter:	Prof. Dr. Andriy Luzhetskyy
	Prof. Dr. Rolf Müller
Akad. Mitglied:	Priv.-Doz. Dr. Martin Frotscher
Vorsitz:	Prof. Dr. Volkhard Helms

Diese Arbeit entstand unter der Anleitung von Prof. Dr. Andriy Luzhetskyy in der Fachrichtung 8.2, Pharmazeutische Biotechnologie der Naturwissenschaftlich-Technischen Fakultät der Universität des Saarlandes von Oktober 2017 bis August 2022.

"What speaks to the soul, escapes our measurements."

Alexander von Humboldt

Acknowledgments

Words cannot express my gratitude to my supervisor Prof. Dr. Andriy Luzhetskyy for his endless encouragement. I would like to thank him for accepting me as member of AMEG group, for giving me the freedom to undertake even unorthodox experiments and to make new discoveries. Thanks for being patient when results where rather deceiving and for allowing me to grow personally with my tasks.

I am extremely grateful to my second supervisor Prof. Dr. Rolf Müller who contributed with useful advices during the committee meetings.

I could not have undertaken this journey without the generous support of my mentor Dr. Maksym Myronovskyi. He shared his knowledge, both practically in the lab as well as in the writing of articles. Thanks for your thorough work.

Many thanks go to AMEG group for an extraordinary experience of teamwork. I feel lucky to be member of this marvellous work environment. Birgit Rosenkränzer introduced me to the group, the lab techniques and was always around as a friend. Dr. Marc Stierhof and Dr. Nils Gummerlich not only contributed their NMR expertise, but also kept me laughing with their unstoppable production of silly jokes. Dr. Hilda Sucipto and Nikolas Eckert shared my best days and my largest personal crisis. Thanks for standing by my side while never asking for anything in return.

I would like to recognize my good friend Dr. Rebecca Hahn and her family who took me out to nature whenever I needed recreation. I would be remiss in not mentioning my family, especially my parents Karin Gahmig and Friedemann Lasch and my sister Ulrike Himpel. Thanks for accepting my decision to resign from my career in the pharmaceutical industry to return to university. Further thanks go to Roswitha and Egbert Mündörfer for being my second family. I would also like to thank my small guinea pig family for all the entertainment and emotional support.

I am lacking words to describe the gratitude I have for Dr. Marco Mündörfer. I feel lucky that you came into my life, opened my view for the beauty of love and stood by me for twelve unbelievably amazing and rich years. It is incredibly hard to accept that you are gone forever. I miss you.

Zusammenfassung

Naturstoffe finden breite Anwendung als Antiinfektiva oder Krebstherapeutika. Eine reiche Quelle für solche Naturstoffe sind Bakterien aus der Ordnung Actinomycetales. Weit verbreitete Arzneistoffe, wie zum Beispiel Erythromycin oder Tetracyclin, wurden aus Actinomyceten isoliert.

Die vorliegende Arbeit beschreibt die Entdeckung von drei neuen Naturstoffen aus Actinomyceten. Die neuen Moleküle Loseolamycin, Dudomycin und Bonsecamin wurden mittels Expression von Genen in einem Heterologen Wirt entdeckt. Die Beispiele in dieser Arbeit zeigen deutlich die Vorteile wie auch die Schranken dieser Technik für die Entdeckung neuer Moleküle auf.

Neben der Darlegung der Entdeckung und Aufreinigung der Moleküle wird in dieser Arbeit auch die Anwendung von Methoden wie Massenspektrometrie zur Molekülcharakterisierung beschrieben. Gendeletionen sowie die Nutzung von individuellen Kulturmedien halfen, die Biosynthese von Loseolamycin, Dudomycin und Bonsecamin in den jeweiligen Actinomyceten aufzuklären. Des Weiteren wurden Bioaktivitätstests durchgeführt, um mögliche Anwendungen der neuen Moleküle in Medizin oder der Landwirtschaft zu erforschen. Diese Experimente lieferten vielversprechende Ergebnisse für eine der neuen Substanzen.

Abstract

Natural products (NPs) are widely used as anti-infective or anticancer drugs. A prolific source for new natural products are bacteria of the biological order of actinomycetales. As a matter of fact, numerous well-established drugs such as erythromycin or tetracycline are actinomycete products.

The present thesis describes the discovery of three new natural products from actinomycetes. The novel molecules loseolamycin, dudomycin and bonsecamin were discovered using the approach of gene cluster expression in a heterologous host. The examples in this work clearly show the advantages as well as the boundaries of heterologous expression in the discovery of new natural products.

Besides the molecules' discovery and isolation, this work describes the use of methods such as high resolution mass spectrometry to characterize the new NPs. Gene deletion experiments as well as the cultivation in customized media helped to elucidate the biosynthesis of the new molecules in the actinomycetes. Further, bioactivity tests were also performed to explore potential applications of loseolamycin, dudomycin and bonsecamin in medicine or agriculture and led to interesting results in one case.

Publications

C. Lasch, N. Gummerlich, M. Myronovskyi, A. Palusczak, J. Zapp, A. Luzhetskyy, Loseolamycins: A Group of New Bioactive Alkylresorcinols Produced after Heterologous Expression of a Type III PKS from Micromonospora endolithica, Molecules. 25 (2020) 4594. https://doi.org/10.3390/molecules25204594.

C. Lasch, M. Stierhof, M. R. Estévez, M. Myronovskyi, J. Zapp, A. Luzhetskyy, Dudomycins: New Secondary Metabolites Produced after Heterologous Expression of an Nrps Cluster from Streptomyces albus ssp. chlorinus NRRL B-24108, Microorganisms. 8 (2020) 1800. https://doi.org/10.3390/microorganisms8111800.

C. Lasch, M. Stierhof, M. R. Estévez, M. Myronovskyi, J. Zapp, A. Luzhetskyy, Bonsecamin: A New Cyclic Pentapeptide Discovered through Heterologous Expression of a Cryptic Gene Cluster, Microorganisms. 9 (2021) 1640. https://doi.org/10.3390/microorganisms9081640.

Conference contributions

C. Lasch, M. R. Estévez, M. Myronovskyi, M. Stierhof, N. Gummerlich, A. Luzhetskyy, Characterization of novel secondary metabolites originating from Actinobacteria, Doktorandentag Saarland University 2019 (poster presentation)

C. Lasch, M. R. Estévez, M. Myronovskyi, A. Luzhetskyy, Characterization of novel secondary metabolites originating from Streptomyces albus ssp. chlorinus, 32. Irseer Naturstofftage 2020 (poster presentation)

Table of contents

1   Introduction

1.1   Microbial natural products as anti-infectives

Disease has been an inevitable part of life as long as humankind can remember. Only 100 years back at the beginning of the 20th century infectious diseases such as pneumonia, tuberculosis or cholera had high level mortality rates and the average life expectancy at birth in industrial nations rarely exceeded 50 years [1,2]. The chance observation of growth inhibiting activity of Penicillium moulds against *Staphylococci* made by Alexander Fleming advanced the treatment of severe infections tremendously and paved the way to the isolation and therapeutic use of the first antibiotic substance known today as penicillin [3]. This finding heralded a "golden era of antibiotics" with discovery of numerous compounds of biological origin which were successfully used for treatment of infectious diseases in human. As a result of this achievement a drastic decrease in infant mortality and an increase of life expectancy to about 80 years were observed at the beginning of the 21st century [1,2].

The relevance of natural organisms as source of new antibiotics is uncontested [4]. In the case of Fleming the antibacterial compound was produced by a mould. Besides fungi, the clade of actinobacteria are considered to be the most prolific source of natural products (NPs) [5]. Especially the actinobacterial genus *Streptomyces* encompasses outstanding producers of antibiotics and contributes in a large extend to the pool of biologically active substances [6,7]. For instance, streptomycin, tetracycline and fosfomycin belong to currently used antibacterial drugs that originate from *Streptomyces* [8–10].

The problem. Despite the successful isolation of numerous antibacterial compounds during the "golden era", today's clinical need for new antibiotic substances remains unbroken as resistance to known antibacterial compounds is on the rise. In the 1980s an overall decline in the discovery rates of new NP scaffolds gave serious cause for concern whether the demand for new antibiotics can still be met by the established sources such as bacteria and fungi [4,11]. Fortunately, few encouraging exceptions recently demonstrated that *Streptomyces* still harbour the potential to produce novel chemical compounds [12]. The compounds platensimycin and daptomycin with convincing antibiotic activities have been successfully isolated from *S. platensis* and *S. roseosporus* in the 21st century [13,14]. However the question still remains open whether this discovery was something more than a pure serendipity.

The cause. This question may be answered by looking closer at the techniques used for NP discovery. The approach of NP discovery used during the "golden era" was based on the analysis of the metabolites extracted from a culture broth of a strain [11]. Compound isolation was either driven by the peak signals detected during chromatographic runs or – as in the case of Fleming's penicillin – by testing antibiotic activity of the strains. All metabolites produced by the investigated strain seemed to be captured. But this conservative approach was holistic on the first glance only.

Due to the rapid advances in sequencing and bioinformatics technologies hundreds of *Streptomyces* genomes have been fully characterized in the recent decade [15] and the biosynthetic gene clusters (BGCs) governing the production of NPs have been identified [16]. The obtained data revealed a huge discrepancy between the number of the bioinformatically identified BGCs and the number of experimentally detected NPs implying that the bacteria do not make use of their whole potential to produce NPs when cultivated under axenic conditions in the laboratory [5]. Further analysis revealed that many of those dormant BGCs are indeed silent, meaning that their encoded products are either not produced under standard laboratory conditions or that low production levels prevent their detection [15,17]. Must those clusters remain untapped or is there a way to capture this biosynthetic wealth?

The way out. One possibility to access those hidden compounds lies in the optimization of the cultivation conditions of the strain – a laborious, undirected and often frustrating approach as the success is hard to predict. But what if it was possible to turn on those clusters on demand? Today's molecular biological understanding and available tools for genetic manipulation allow the targeted activation of silent clusters. For instance, BGC activation can be performed in the native strains using e.g. CRISPR/Cas9 facilitated promoter knock-in [18,19]. However, many native *Streptomyces* strains are refractory to genetic manipulations and most of them are rather slow and complicated in cultivation. If production can be achieved it often remains on a barely detectable level as rival enzymes compete for the limited amount of precursors available in the cell [15]. A more efficient method to get access to NPs encoded in the silent clusters seems to be the outsourcing of the metabolite production to other characterized, well-established surrogate strains – an approach referred to as heterologous expression and production. The purpose of this work is to shed light on the potentials of this method in drug discovery, but as well to critically discuss its constraints.

1.2   Potential and constraints of heterologous expression

Streptomyces harbour in average dozens of biosynthetic pathways for NPs, but control the production of secondary metabolites according to their needs [20]. While the pattern of expressed proteins and produced NPs can vary depending on the cultivation conditions, the genetic information on the bacterial DNA is a constant, ready to be used as a template for synthesis of molecules upon bacterial request. In bacterial genomes the genes encoding the biosynthesis of NPs are not randomly distributed within the chromosome but are rather placed together in so called biosynthetic gene clusters [11]. Such spatial arrangement of the biosynthetic genes into the clusters is not only of great advantage for their identification, but also facilitates their heterologous expression. To date a number of bioinformatic programs are available for identification of biosynthetic gene clusters within the bacterial genomes, assignment of their functions by comparison to already known genes or even prognostication of probable metabolic products. The most widely used is the web-based, non-commercial antiSMASH tool [21] which allows for identification of a broad range of (even putative) cluster types [22]. Other programs use algorithms designed for the identification of pathway specific types of clusters such as non-ribosomal peptide synthases (NRPS), polyketide synthases (PKS) [23,24] or ribosomally synthesized peptide (RiPPs) clusters [25]. A comprehensive overview of the currently available bioinformatic tools is provided by different authors [16]. Based on the prediction one can select a unique cluster candidate for heterologous expression. In order to increase the probability to discover a novel natural product during heterologous expression a preference is often given to cryptic clusters [26], clusters coding for poorly investigated compound classes [27,28], hybrid clusters [29] or BGCs containing genes that code for interesting features such as halogens [30] or rare sugars [31]. Further steps in the process of heterologous production are the cloning of the selected BGC from the native strain, the selection of an appropriate heterologous host organism, transfer of the cluster, analysis of the secondary metabolite profile of the resulting strain and finally the purification of the compound. In case when no product is detected in the course of heterologous expression, genetic engineering of the studied pathway is performed to ensure successful expression of the biosynthetic genes in the heterologous host (Figure 1). These steps are explained in more detail in the following sections.
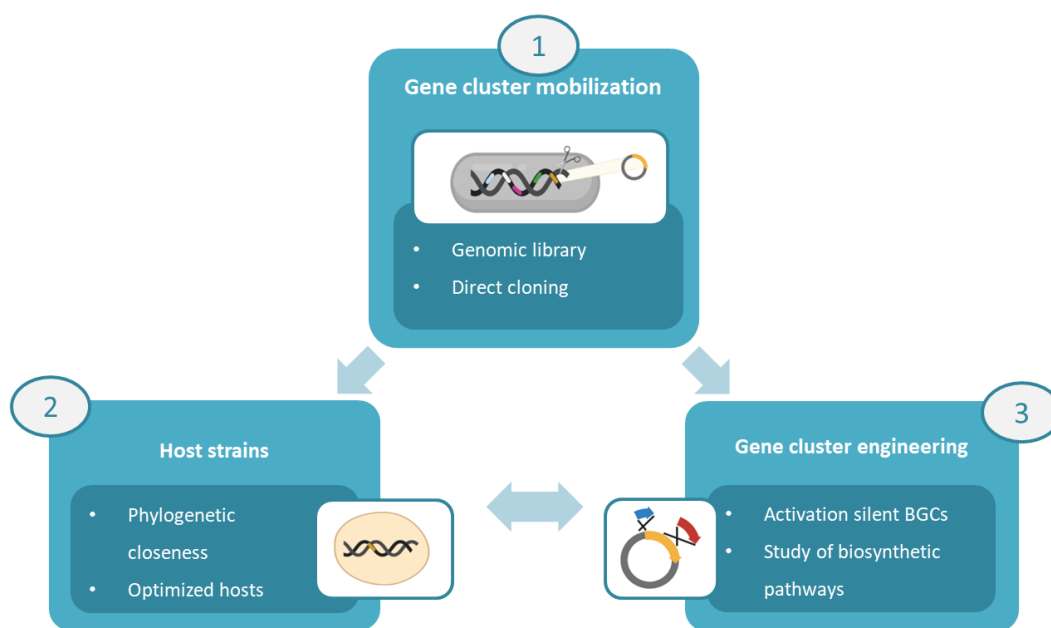
Figure 1. The three core elements of heterologous expression.

1.2.1 Gene cluster mobilization

Once the choice for a pathway of interest on the bacterial genome is made, the corresponding gene cluster needs to be cloned from the chromosome into a specific vector which could be maintained in the heterologous host.

Genomic libraries. If available, one may rely on a genomic library covering fragments of the entire bacterial genome. Those libraries are based on fragmentation of the genomic DNA which can be achieved by restriction endonuclease digestions or mechanical shear. DNA fragments are ligated *in vitro* with the backbone of the expression vector and transformed to *Escherichia coli* for propagation.

The vector serves as an artificial carrier of the DNA insert and can contain elements relevant for selection, transfer, maintenance, replication and expression of the DNA. Depending on the mechanism of propagation in the cells the vectors can be subdivided into two groups: replicative (often high copy number) and integrative vectors. The phage-based integrative vectors (mainly *phi*C31-based) have proven themselves as most suitable for the expression of natural product biosynthetic clusters in *Streptomyces* hosts, since the introduced

DNA is stably maintained as a part of the host chromosome and is not lost over time even without antibiotic selection [32]. Vector classification according insert size plays an eminent role with capacity increasing from cosmids/fosmids (40-50 kb) to artificial chromosomes (> 100 kb).

As the generation of genomic libraries is cluster unspecific, screening for the correct DNA fragment is mandatory. Libraries are valuable if the genome harbours a multitude of BGCs interesting for expression. The respective DNA fragment covering the BGC can easily be picked from the library for further transfer to the host strain and heterologous expression. Straightforward work might be hindered if the desired BGC is not represented as a single continuous fragment in the genomic library. This can occur either due to gaps in the library or when the huge genetic pathways of megasynthetases are scattered on several DNA fragments. The second can be resolved either by parallel expression of all relevant vectors [33,34] or – more elegant and more common – by reconstitution of the cluster on a single vector prior transfer to the host. Different tools are available to stich fragments to complete clusters, e.g. homologous recombination in *E. coli* [35] or *Saccharomyces cerevisiae* [36,37].

Direct cloning. In those cases where the BGC of interest cannot easily be picked from a constructed library, it is possible to directly capture and clone the specific BGC of interest from the bacterial genome. Several *in vivo* and *in vitro* techniques were established and are reviewed in detail by different authors [38–40]. Established methods to obtain specific DNA fragments are PCR amplification, CRISPR/Cas9 mediated cut-out or custom synthesis of the desired DNA fragment. The latter bears the advantage that codons may be designed individually to suit the host's preferred tRNAs. To join the thus captured DNA fragments to the vector, either ligation, Gibson assembly or *in vivo* recombination can be used. Of the whole panel of complete direct cloning strategies the most appealing ones are probably transformation associated recombination (TAR), linear plus linear homologous recombination (LLHR) and integrase mediated DNA excision (IR). These techniques are discussed below.

TAR. Transformation associated recombination (TAR) cloning technology was increasingly employed in the recent decade and has become a routine method for assembly or cloning of bacterial DNA [41]. For direct cloning the method requires a linear DNA fragment containing the cluster of interest and a linearized cloning vector harbouring hooks on both ends. The hooks share homology to the 5' and the 3' ends of the desired DNA fragment. Both, the linear DNA fragment and the vector, are co-transformed into the yeast *S. cerevisiae*. The

yeast's inherent ability for homologous recombination facilitates the construction of the intact circular artificial chromosome (yeast artificial chromosome YAC) [42]. This shuttle vector carries elements from the yeast itself, but as well from *E. coli* where it is maintained and from *Streptomyces* as the desired expression host [38,43] and provides capacity for uptake of up to 250 kb DNA fragments [42]. For instance, the TAR system was successfully used to clone a biosynthetic cluster from the actinomycete *Saccharomonospora* sp. CNQ-490. Further heterologous expression of this BGC in the host *Streptomyces coelicolor* led to the discovery of the potent antibiotics taromycin A and B [44,45]. Meanwhile, several bottlenecks of the TAR methodology were addressed: efforts were reported for application of TAR for cloning of GC-rich bacterial DNA, such as from actinobacteria [41] and also attempts to increase the overall cloning efficiency were undertaken [46,47]. Still, the method necessarily requires a step in yeast (recombination) making its use arduous.

LLHR. The principle of linear plus linear homologous recombination (LLHR) is comparable to TAR, but not requiring the step in yeast. In contrast to TAR, the recombination is carried out in *E. coli* directly and is mediated by the RecE/RecT proteins of Rac prophage [48]. The LLHR system was initially limited to BGCs up to 52 kb, but recently combination with other technologies was reported to overcome the size issue. Thus, the optimized LLHR allowed for direct capture of the intact 106 kb salinomycin BGC from *S. albus* [49].

IR. Integrase mediated recombination (IR) is another noteworthy method for the direct cloning of biosynthetic gene clusters. The method is based on utilisation of the integrase of the phage *phi*BT1 and its cognate attachment sites *att*P and *att*B. *att*P, *att*B and the vector elements required for replication in *E. coli* are introduced into the chromosome of the strain in the way that the attachment sites flank the desired cluster from both the 5' and the 3' ends. Upon expression of the *phi*BT1 integrase recombination of *att*B with *att*P is triggered and the cluster of interest is excised from the chromosome of the strain as a circular DNA molecule containing a vector backbone for replication in *E. coli*. IR was successfully applied by the authors, e.g for cloning of the actinorhodin gene cluster [50]. The method convinces with its high cloning efficiency. However, IR is reserved to BGC isolation from genetically tractable organisms and so far the method did not prevail as routine in the labs.

1.2.2 Host strains

Once the gene cluster is cloned in a suitable vector it can be transferred to an appropriate host organism, either through direct transformation or in case of *Streptomyces* hosts using intergeneric bacterial conjugation. In each case the recombinant strains are selected on a medium supplemented with an appropriate antibiotic. The transferred BGC enables the exconjugant strain to produce the encoded NP upon cultivation. Further extraction and chromatographic analysis will reveal the success of the experiment. This success is strongly dependent on the choice of the optimal host used for the purpose. This section provides an overview on the host options and will discuss their advantages and disadvantages.

Classic industrial hosts. Since the beginning of systematic heterologous expression in the 1980s bacterial and yeast cells have been extensively used as hosts. The long-standing surrogate strains include the gram-negative bacterium *Escherichia coli* as well as the eukaryotic yeast *Saccharomyces cerevisiae*. Both strains provide comfortable and economic lab handling, short cell doubling times, they are genetically well-characterized and come along with highly developed tools for strain engineering. The gram-positive bacterium *Bacillus subtilis* is a further favoured host convincing with a functional endogenous system for efficient excretion of metabolites. The latter two organisms are in addition classified as GRAS (generally recognized as safe) organisms by the FDA. Attempts to express the complex erythromycin pathway originating from the actinomycete *Saccharopolyspora erythraea* were made for the three above-mentioned strains. The successful expression of erythromycin A and C in *E. coli* was achieved after intense strain engineering [51,52]. The non-glycosylated erythromycin core was produced by *B. subtilis* [53] and a partial reconstruction of the pathway was reported for *S. cerevisiae* [54]. This raises the question whether those strains are indeed the best choice for expression of actinobacterial NPs.

Despite the seemingly advantageous properties of *E. coli*, *S. cerevisiae* and *B. subtilis* surrogate strains, there is one huge drawback when using them in combination with actinobacterial pathways – their phylogenetic distance. Interspecies differences can hamper the successful heterologous expression on different levels. Functional promoters and complex regulatory circuits are prerequisites for successful transcription [53]. The high prevalence of GC bases in *Streptomyces* may mismatch the preferred codons used in classic hosts and lead to decreased translation levels. Problems in folding of the nascent peptide chain to form the protein complex as well as absence of crucial posttranslational modifications required for the

activity of heterologously expressed enzymes were reported. For instance, the lack of phosphopantheteinyl transferases (PPTases) which are required for activation of nonribosomal peptide synthetases and polyketide sythases is one of the commonly encountered issues [55]. The pool of precursors of the heterologous host may also hinder successful production of the compound if rare precursors such as methylmalonyl CoA, propionyl CoA or non-canonical amino acids are required for the biosynthesis of the product [53–57]. Furthermore, toxicity issues are common during heterologous production. Biologically active compounds accumulating in the host cell will cause the producer's premature death if no resistance or efficient excretion system is inherent in the host [51]. The reported cases of successful heterologous expression in these hosts demonstrate that above issues can be optimistically regarded as challenges. Production of the desired compound can still be achieved if the requirements for its biosynthesis are systematically addressed. However, the immense overall workload and time investment required for the construction of compound-tailored hosts cannot be underestimated. Furthermore, once constructed the strains will likely exhibit low versatility required for the expression of other, structurally different compounds (Figure 2).
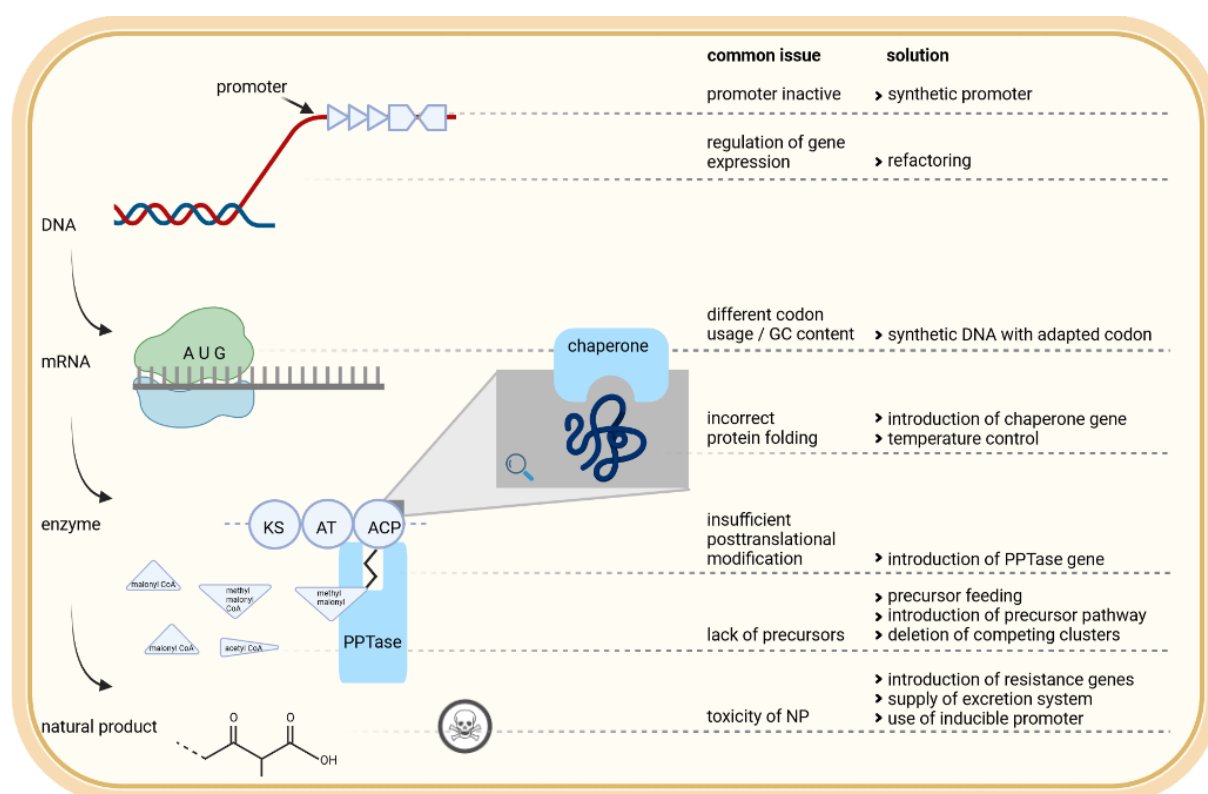
Figure 2. Common issues encountered during heterologous expression and ways to address them.

***Streptomyces* hosts.** The phylogenetic distance between the source organism of the genetic pathway and the host organism needs to be taken into consideration during heterologous expression. Although actinobacteria are generally quite complicated to handle in the laboratory, the clade is huge enough to bear some cultivable and well-studied strains that are susceptible to genetic manipulations. The *Streptomyces* strains *S. coelicolor* and its close relative *S. lividans* belong to the established model strains and were predominantly used as hosts in recent years. The advantage of a further laboratory strain *S. albus* is its small, well characterized genome and its captivating fast growth [58]. To date, successful heterologous production of NPs in above hosts has been demonstrated and examples were nicely reviewed and tabulated by several groups [43,59,60]. Remarkably, those examples are not limited to heterologous expression of already characterized BGCs, but also include cases of NP discovery after expression of silent clusters [61]. The successful use of *Streptomyces* strains as surrogate strains encouraged scientists to improve their properties. A variety of alterations were introduced into the host's genomes in order to facilitate the heterologous production of compounds. Currently a whole platform of various chassis hosts is available including engineered *S. coelicolor* [62–64], *S. lividans* [65,66], *S. albus* [67,68], *S. avermitilis* [69] , *S. chattanoogensis* [70], *S. venezuelae* [71], *S. roseosporus* [72], and *S.* FR-008 [73]. The undertaken strain improvement efforts addressed the availability of precursors, the maintenance of the BGC in the host, the upregulation of transcription and translation levels as well as the secretion of the metabolites. These aspects are discussed in the following section.

Precursor availability. Complex natural products are built up by a number of small precursor molecules. Depending on the product type these can include acyl-units, amino acids, fatty acids, sugars and many more. An ideal super-host provides a wide variety of those building blocks to suit the demand of different compound classes. Successful production of the desired metabolite largely depends on the availability of precursor substances and also on their concentration. In contrast to *E. coli* or *S. cerevisiae*, *Streptomyces* host strains convince with a broad range of biosynthetic precursors required for production of all kinds of metabolites. The *Streptomyces* hosts are, however, not without their disadvantages – an extraordinary abundance of endogenous biosynthetic gene clusters is encoded in their genomes. Endogenic clusters can act as competitors for the cell's precursor units and thus limit their availability for the production of metabolites encoded by the newly introduced pathway. A strategy to reduce this competition is the deletion of dispensable endogenous gene clusters. This streamlining approach was applied to several established *Streptomyces* hosts

leading to optimized strains with reduced number of indigenous clusters. The further pursuit of this cluster deletion approach led to the generation of the first complete cluster-free chassis hosts. Cluster-free derivatives of *S. avermitilis* (SUKA17/SUKA22) [69], *S. albus* (*S. albus* Del14) [67] and *S. lividans* (*S. lividans* DelYA9) [65] were successfully validated as host systems and facilitated the discovery of e.g. pentangumycin or huimycin in our group [26,27]. Another considerable advantage of cluster-free chassis strains is a substantially simplified metabolite background of their extracts. This facilitates greatly the detection of new peak signals which otherwise could easily be overseen. Furthermore, the clean metabolic background is of great benefit during substance isolation. The lower level of contaminants in the crude extract reduces the number of purification steps required to obtain the pure new substance (Figure 3).

Cluster maintenance. Bacterial conjugation is commonly used to transfer the vector from *E. coli* to the *Streptomyces* host. Subsequently, the vector is integrated into the *Streptomyces* chromosome to stably maintain the new genes. This step is empowered by the *phi*C31 integrase which is encoded on common vectors. This enzyme allows for the site specific integration of the cluster in the chromosome at so called *att*B sites. These *phi*C31 attachment sites are naturally present in chromosomes of most actinobacteria. As the yield of heterologously expressed products depends on the location and the copy number of the BGC [74], improved versions of the cluster free chassis strains *S. albus* Del14 and *S. lividans* DelYA9 were constructed. Two additional *att*B sites were introduced in the host genomes enabling simultaneous integration of several copies of the exogenous cluster. Several known metabolic pathways were expressed in the engineered strains to proof the concept. The constructed *att*B-mutant strains demonstrated increased production levels of e.g. griseorhodin, but generally the magnitude of the production increase was pathway dependent [65,67].

Transcriptional upregulation. The production of secondary metabolites is strictly controlled by the bacterium involving complex regulatory mechanisms on both transcriptional and translational level. Most cell regulation processes appear to be pathway specific. However, some mechanisms were found to pleiotropically enhance the expression of secondary metabolite BGCs. To date, a number of regulatory systems are known to govern globally the transcription of secondary metabolite BGCs in an organism [20,75]. Different groups used this knowledge to remodel those global regulatory circuits in order to increase the expression of heterologous BGCs. Various pleiotropic positive regulators were introduced in

the genomes of *S. lividans* and *S. albus* or alternatively, the existing endogenous negative regulators were knocked-out [66,68]. A different strategy was followed for the upregulation of cluster transcription in *S. coelicolor*. It is based on the finding that some spontaneous mutations in the RNA polymerase genes are accompanied by enhanced secondary metabolite production [76,77]. The scientists exchanged the respective RNA polymerase gene by its variants with the mutations. [62]. In all three studies, the proof of principle study with exogenous model clusters was carried out and an increased metabolite production by the mutant strains was demonstrated.

Translational upregulation. An approach similar to the above described utilisation of mutated versions of RNA polymerase was used to pleiotropically increase the translation of secondary metabolite pathways. The empirical discovery that certain point mutations in the genes encoding ribosome proteins enhance the overall translation of secondary metabolite clusters is the basis of ribosome engineering [78]. Mutations in the gene coding for a ribosomal protein S12 were generated. The respective gene in an already streamlined genome of *S. coelicolor* was substituted by the mutated gene. The mutant displayed increased production of heterologously expressed compounds [62]. Similarly, ribosome mutants of *S. albus* were constructed for the expression of silent gene clusters by our group. The subsequent mutant analysis indicated that ribosome engineering affects not only the translational level, but also has a feedback on transcriptional regulation [79].

Metabolite secretion. New NPs exhibiting good bioactivity are highly desired molecules. However, if this activity is antibacterial it may be harmful for the producing heterologous host itself if it does not inhere a native resistance mechanism. This problem can be addressed by the use of inducible promoters (see next section) or the transfer of genes conferring resistance. A less specific way to manage this obstacle is the use of efflux pumps. Thus, the genes encoding for two multi-drug transporters were transferred to *S. lividans* and integrated into its genome. The resulting mutant showed slightly increased production of compounds encoded by the introduced model pathway [66].

Many examples can be found where chassis strains with deleted endogenous clusters were successfully used to identify new compounds from cryptic actinobacterial clusters. However, engineered host strains that were rationally optimized with regard to gene maintenance, regulation or metabolite secretion are not yet used in labs in a routine manner.

Their high potential in heterologous expression was demonstrated. Still, their role in discovery of new NPs will be shown in the forthcoming years.
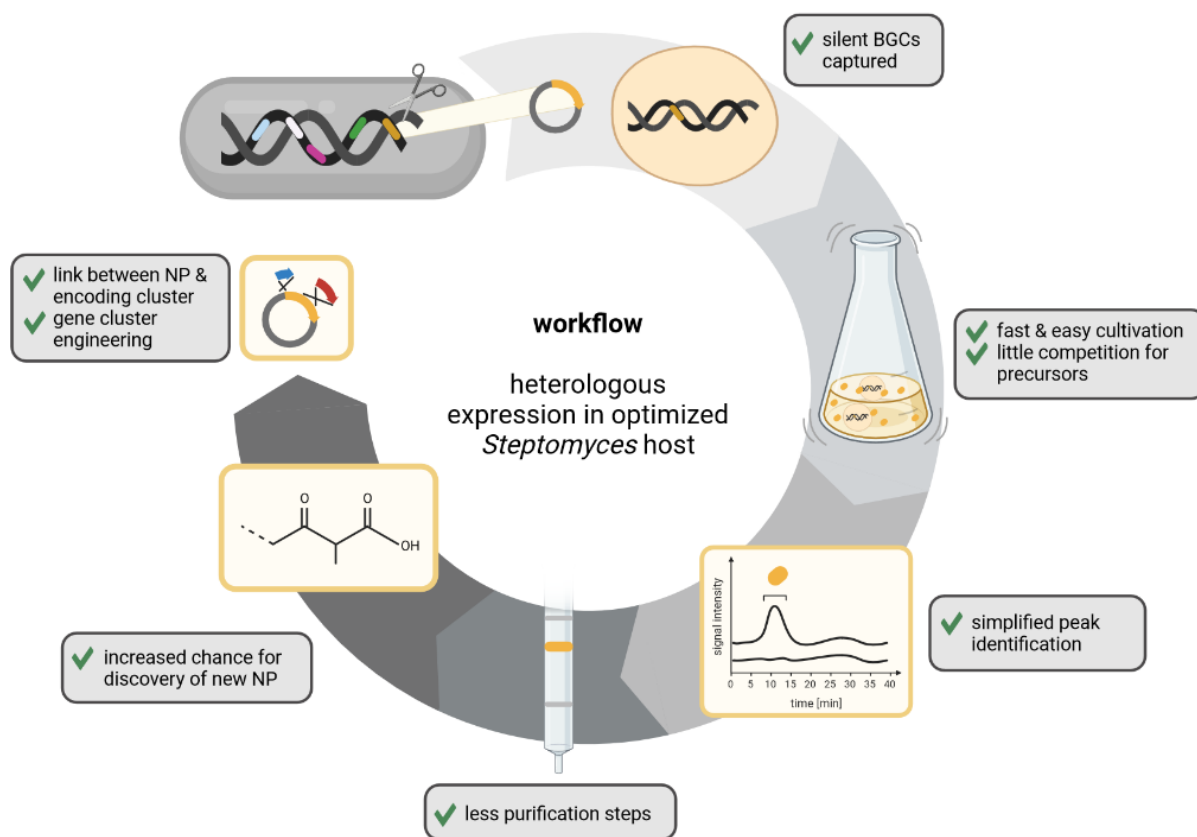

Figure 3. Workflow & benefits of heterologous expression in optimized Streptomyces chasses.

Other actinobacteria hosts. The plenty of available genetic tools and optimized *Streptomyces* hosts make heterologous expression of silent BGCs a reasonable and straight strategy in NP discovery. However, the true value of this approach can be measured by its outcome only. We systematically expressed all relevant silent gene clusters of two actinobacteria strains and determined the success rates by measurement of metabolite production. After transfer of 17 gene clusters of *S. albus* ssp. *chlorinus* to the hosts *S. albus* Del14 and *S. lividans* DelYA9 we detected new peaks in around 44 % of the cases [unpublished data]. In contrast, the expression of 17 gene clusters of the genetically more distant *Saccharothrix espanaensis* yielded production in 11 % of the cases only [26]. This observation highlights one more time the role of a specie's phylogenetic closeness for successful heterologous expression.

The discovery of the new compounds huimycin from *Kutzneria albida* and pentangumycin from *Saccharothrix espanaensis* are two examples from our lab demonstrating that in principle the concept of heterologous expression in *Streptomyces* hosts also works for BGCs from distantly related, rare actinobacteria, too [26,27]. Compared to *Streptomyces* those rare actinobacteria are still regarded as under-investigated and thus a promising source for new NP structures. All the more it is desirable to increase the expression success rates from those rare actinobacteria to fully use their biosynthetic potential. Two methods are suggested to increase the compound output.

The first option is to utilize heterologous host strains with little phylogenetic distance to the organisms which were used as a source of BGCs. So far this approach is little popular in the research of new antibiotics, probably because rare actinobacteria which could be used as potential hosts are often challenging in cultivation and genetically intractable. Despite such poor preconditions the following attempts give reason for optimism: the brasilicardin terpene core was successfully produced after heterologous expression of the respective *Nocardia terpenica* cluster in an *Amycolatopsis japonicum* host [80]. An engineered *Amycolatopsis* sp. ATCC 39116 host with additional *att*B site was used in our lab to heterologously produce cinnamycin from *Streptomyces cinnamoneus* and flaviolin, a compound encoded by a BGC of *Saccharopolyspora erythraea* [unpublished data]. *S. erythraea* is a well-known industrial producer of erythromycin. An engineered *S. erythraea* strain where the erythromycin genes were replaced by an *att*B site was constructed for the purpose of heterologous production of polyketides [81]. Similarly, the marine actinomycete *Salinispora tropica* was optimized for heterologous cluster expression. A part of endogenous BGCs was deleted and a functional *att*B site was introduced in the chromosome of *S. tropica*. The constructed strain is a promising host for the expression of BGCs of various *Salinispora* species [82]. *Corynebacterium glutamicum* is another more distant actinobacterium which recently came into focus as potential host for production of diverse classes of natural products. *C. glutamicum* is a gifted producer of amino acids. Recently the strain was engineered to facilitate the production of polyketides [83].

The development of rare actinobacteria as host strains for heterologous expression of BGCs seriously lags behind the advances made for *Streptomyces* hosts. A second option to activate expression of silent clusters from unconventional actinobacterial strains is their extensive refactoring for the expression in the established *Streptomyces* hosts.

1.2.3 Gene cluster engineering

One of the biggest advantages of heterologous expression of BGCs is the possibility to link the detected product to its encoding genes. This knowledge about the biosynthetic pathways serves as key to activate and optimize the NP production. The insertion of promoters or the refactoring of entire pathways have facilitated the production of many new compounds (figures 1 and 3).

Promoter insertion. Efficient transcription as the first step of gene expression is extremely important for the success of heterologous expression and NP production. The promoter is a short DNA sequence upstream of the genes and operons which is recognized by the RNA polymerase and is responsible for initiation of transcription. Native promoters are often controlled by bacterial regulatory networks. A widespread and most-effective approach to activate the transcription of sleeping genes is the insertion of defined active promoter elements upstream of the cluster [84]. The promoter insertions are directly carried out on the cloned vector using either homologous recombination or CRISPR/Cas9 technique. Besides natural promoters, a number of (semi)synthetic promoter elements suitable for *Streptomyces* were constructed and are available in respective libraries [85,86]. Promoters can either be constitutive or inducible, meaning that a stimulus is needed to trigger their activation. Inducible promoters are helpful in case the production should be started on demand and are advantageous during the production of toxic metabolites when an unwanted feedback regulation during biomass accumulation should be avoided [40]. The insertion of single promoters is applied for the activation of BGCs on single operons. In case of complex cluster architecture with multiple operons refactoring of the cluster is the method of choice.

Refactoring. Gene refactoring refers to the complete reorganisation of metabolic pathways including the replacement of promoters, the tuning of ribosomal binding sites, the insertion of terminators and if required inactivation of repressors. This approach is beneficial mainly when dealing with complex clusters where biosynthetic genes are scattered on several operons or in cases where the genes of the pathway are not entirely clustered in one genomic locus. The huge economic and time investments required during cluster refactoring did not make this strategy an established method for drug discovery. The shortened version of this approach which is often referred to as "promoter refactoring" focuses solely on the replacement of regulatory elements and is used more widely. For instance the removal of a repressor gene led to identification of the new promising antibiotics taromycin A and B

through heterologous expression of their biosynthetic cluster from *Saccharomonospora* sp. CNQ-490 in an *S. coelicolor* host [44,45]. Still, it must be considered that the process of refactoring is not trivial and redesign of the cluster needs to be performed thoroughly. Previous studies pointed out the importance of balancing the strength of individual promoters used for refactoring in order to obtain well-equilibrated transcription of biosynthetic genes and optimal production of the target product [84,87]. In contrast, the intentionally unbalanced transcription of a biosynthetic pathway was suggested as a tool to increase the production level of otherwise low abundant biosynthetic intermediates or shunt products. This strategy has successfully been applied in our lab and led to the discovery of novel rabelomycin and tetrangulol derivatives from the landomycin gene cluster [87] and drastically altered the spectrum of produced pamamycins [unpublished data].

Gene inactivation. Besides the use of genetic engineering to access new metabolites and optimize their production, it is also of great importance to contribute to the general understanding of the bacterial biosynthetic machinery. Targeted gene inactivation experiments in order to determine the borders of the gene cluster as well as inactivation of pathway genes to produce biosynthetic intermediates are crucial for the reconstruction of the genuine metabolite pathways. These experiments add value in all cases where the biosynthesis does not follow regular, well-investigated mechanisms. In return, those valuable insights will allow us to systematically generate new metabolites in further applications such as combinatorial biosynthesis.

1.3 Prospects

Let's return to the initial question whether the traditional family of actinobacteria has been disused or is still fruitful enough to provide new anti-infective compounds. The above described examples illustrate vividly that actinobacteria still bear a prolific reservoir of new metabolites that are accessible by means of modern biotechnology. Heterologous expression is a powerful tool to reach those hidden compounds. The chances inhered in heterologous expression seem to be even more convincing when the source of biosynthetic pathways is expanded to environmental samples. Considering the fact that over 99 % of the natural strains are not cultivable at all, the expression of environmental DNA (eDNA) in surrogate strains is a tricky, yet very interesting concept to be followed. This metagenomics approach facilitated the discovery of the potent MRSA active drug tetarimycin A produced by an *S. albus* host

[88]. Beyond the capture of native compounds, heterologous expression is a playground for creation of "unnatural" natural compounds. This includes the targeted shuffling of known scaffolds by combinatorial biosynthesis. The probably most popular examples for molecule hybrids are the mithramycin variants with diversified glycosylation or the rebeccamycins with varied halogenation pattern [89,90]. Further, the production of new substances arising through cross-talk of the heterologously expressed clusters with the endogenous host genes deserves more attention in the future. Several examples from our lab demonstrate that some advanced chassis strains such as *S. albus* Del14 preferably invest their precursor resources in the creation of those products, namely fridamycins F and G, benzanthric acid and SEK90 [26,87,91]. However, biological activity for those molecules is not guaranteed and needs to be tested meticulously.

Three more examples of successful production of new molecules after heterologous expression are presented in this PhD thesis work. Loseolamycin, dudomycin and bonsecamin were discovered after expression of exogenous pathways in *Streptomyces* hosts. Although the general applied workflow was the same for all three molecules, the details such as source strain, product types and tools for gene cluster engineering varied. Problems encountered on the way towards the new molecules and solutions to overcome those hindrances are summarized in the conclusion section. Hence, this work reinforces the capabilities as well as the limits of the heterologous expression approach in the discovery of new natural products from actinobacteria.

1.4   References

[1]   J.R. Wilmoth, Demography of longevity: past, present, and future trends, Exp. Gerontol. 35 (2000) 1111–1129. https://doi.org/10.1016/S0531-5565(00)00194-7.

[2]   Office for National Statistics, Statistical bulletin: English Life Tables No.17: 2010 to 2012, 2015.

[3]   Fleming Alexander, Nobel Lecture: Penicillin, 1945.

[4]   D.J. Newman, G.M. Cragg, Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019, J. Nat. Prod. 83 (2020) 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285.

[5]   R. Baltz, Renaissance in antibacterial discovery from actinomycetes, Curr. Opin. Pharmacol. 8 (2008) 557–563. https://doi.org/10.1016/j.coph.2008.04.008.

[6]   M. Watve, R. Tickoo, M. Jog, B. Bhole, How many antibiotics are produced by the genus Streptomyces ?, Arch. Microbiol. 176 (2001) 386–390. https://doi.org/10.1007/s002030100345.

[7]   J. Bérdy, Bioactive Microbial Metabolites, J. Antibiot. (Tokyo). 58 (2005) 1–26. https://doi.org/10.1038/ja.2005.1.

[8]   D. Hendlin, E.O. Stapley, M. Jackson, H. Wallick, A.K. Miller, F.J. Wolf, T.W. Miller, L. Chaiet, F.M. Kahan, E.L. Foltz, H.B. Woodruff, J.M. Mata, S. Hernandez, S. Mochales, Phosphonomycin, a New Antibiotic Produced by Strains of Streptomyces, Science (80-.). 166 (1969) 122–123. https://doi.org/10.1126/science.166.3901.122.

[9]   B.M. Duggar, Aureomycin: A product of the continuing search for new antibiotics, Ann. N. Y. Acad. Sci. 51 (1948) 177–181. https://doi.org/10.1111/j.1749-6632.1948.tb27262.x.

[10]   A. Schatz, E. Bugle, S.A. Waksman, Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria.*, Exp. Biol. Med. 55 (1944) 66–69. https://doi.org/10.3181/00379727-55-14461.

[11]   R.H. Baltz, Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration?, J. Ind. Microbiol. Biotechnol. 33

(2006) 507–513. https://doi.org/10.1007/s10295-005-0077-9.

[12]   H.A. Kirst, Developing new antibacterials through natural product research, Expert Opin. Drug Discov. 8 (2013) 479–493. https://doi.org/10.1517/17460441.2013.779666.

[13]   J.N. Steenbergen, J. Alder, G.M. Thorne, F.P. Tally, Daptomycin: a lipopeptide antibiotic for the treatment of serious Gram-positive infections, J. Antimicrob. Chemother. 55 (2005) 283–288. https://doi.org/10.1093/jac/dkh546.

[14]   D. Häbich, F. von Nussbaum, Platensimycin, a New Antibiotic and "Superbug Challenger" from Nature, ChemMedChem. 1 (2006) 951–954. https://doi.org/10.1002/cmdc.200600145.

[15]   N. Lee, S. Hwang, J. Kim, S. Cho, B. Palsson, B.-K. Cho, Mini review: Genome mining approaches for the identification of secondary metabolite biosynthetic gene clusters in Streptomyces, Comput. Struct. Biotechnol. J. 18 (2020) 1548–1556. https://doi.org/10.1016/j.csbj.2020.06.024.

[16]   L. Sukmarini, Recent Advances in Discovery of Lead Structures from Microbial Natural Products: Genomics- and Metabolomics-Guided Acceleration, Molecules. 26 (2021) 2542. https://doi.org/10.3390/molecules26092542.

[17]   Y. Zhao, G. Li, Y. Chen, Y. Lu, Challenges and Advances in Genome Editing Technologies in Streptomyces, Biomolecules. 10 (2020) 734. https://doi.org/10.3390/biom10050734.

[18]   M.M. Zhang, F.T. Wong, Y. Wang, S. Luo, Y.H. Lim, E. Heng, W.L. Yeo, R.E. Cobb, B. Enghiad, E.L. Ang, H. Zhao, CRISPR–Cas9 strategy for activation of silent Streptomyces biosynthetic gene clusters, Nat. Chem. Biol. 13 (2017) 607–609. https://doi.org/10.1038/nchembio.2341.

[19]   H. Ren, B. Wang, H. Zhao, Breaking the silence: new strategies for discovering novel natural products, Curr. Opin. Biotechnol. 48 (2017) 21–27. https://doi.org/10.1016/j.copbio.2017.02.008.

[20]   G. Liu, K.F. Chater, G. Chandra, G. Niu, H. Tan, Molecular Regulation of Antibiotic Biosynthesis in Streptomyces, Microbiol. Mol. Biol. Rev. 77 (2013) 112–143. https://doi.org/10.1128/mmbr.00054-12.

[21] M.H. Medema, K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M.A. Fischbach, T. Weber, E. Takano, R. Breitling, antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences, Nucleic Acids Res. 39 (2011) W339–W346. https://doi.org/10.1093/nar/gkr466.

[22] T. Weber, K. Blin, S. Duddela, D. Krug, H.U. Kim, R. Bruccoleri, S.Y. Lee, M.A. Fischbach, R. Müller, W. Wohlleben, R. Breitling, E. Takano, M.H. Medema, antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters, Nucleic Acids Res. 43 (2015) W237–W243. https://doi.org/10.1093/nar/gkv437.

[23] M.H. Li, P.M. Ung, J. Zajkowski, S. Garneau-Tsodikova, D.H. Sherman, Automated genome mining for natural products, BMC Bioinformatics. 10 (2009) 185. https://doi.org/10.1186/1471-2105-10-185.

[24] A. Starcevic, J. Zucko, J. Simunkovic, P.F. Long, J. Cullum, D. Hranueli, ClustScan : an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures, Nucleic Acids Res. 36 (2008) 6882–6892. https://doi.org/10.1093/nar/gkn685.

[25] A. de Jong, S.A.F.T. van Hijum, J.J.E. Bijlsma, J. Kok, O.P. Kuipers, BAGEL: a web-based bacteriocin genome mining tool, Nucleic Acids Res. 34 (2006) W273–W279. https://doi.org/10.1093/nar/gkl237.

[26] N. Gummerlich, Y. Rebets, C. Paulus, J. Zapp, A. Luzhetskyy, Targeted Genome Mining—From Compound Discovery to Biosynthetic Pathway Elucidation, Microorganisms. 8 (2020) 2034. https://doi.org/10.3390/microorganisms8122034.

[27] H. Shuai, M. Myronovskyi, S. Nadmid, A. Luzhetskyy, Identification of a Biosynthetic Gene Cluster Responsible for the Production of a New Pyrrolopyrimidine Natural Product—Huimycin, Biomolecules. 10 (2020) 1074. https://doi.org/10.3390/biom10071074.

[28] M. Rodríguez Estévez, M. Myronovskyi, N. Gummerlich, S. Nadmid, A. Luzhetskyy, Heterologous Expression of the Nybomycin Gene Cluster from the Marine Strain Streptomyces albus subsp. chlorinus NRRL B-24108, Mar. Drugs. 16 (2018) 435.

https://doi.org/10.3390/md16110435.

[29] T. Kawasaki, Y. Hayashi, T. Kuzuyama, K. Furihata, N. Itoh, H. Seto, T. Dairi, Biosynthesis of a Natural Polyketide-Isoprenoid Hybrid Compound, Furaquinocin A: Identification and Heterologous Expression of the Gene Cluster, J. Bacteriol. 188 (2006) 1236–1244. https://doi.org/10.1128/JB.188.4.1236-1244.2006.

[30] J. Shi, Y.J. Zeng, B. Zhang, F.L. Shao, Y.C. Chen, X. Xu, Y. Sun, Q. Xu, R.X. Tan, H.M. Ge, Comparative genome mining and heterologous expression of an orphan NRPS gene cluster direct the production of ashimides, Chem. Sci. 10 (2019) 3042–3048. https://doi.org/10.1039/C8SC05670F.

[31] M. Yu, J. Luo, D. Luo, Q. He, Y. Yan, X. Ji, S.-X. Huang, Discovery and heterologous production of sarubicins and quinazolinone C -glycosides with protecting activity for cardiomyocytes, Org. Chem. Front. (2021). https://doi.org/10.1039/D1QO00470K.

[32] A.C. Groth, M.P. Calos, Phage Integrases: Biology and Applications, J. Mol. Biol. 335 (2004) 667–678. https://doi.org/10.1016/j.jmb.2003.09.082.

[33] C. Kao, L. Katz, C. Khosla, Engineered biosynthesis of a complete macrolactone in a heterologous host, Science (80-.). 265 (1994) 509–512. https://doi.org/10.1126/science.8036492.

[34] J. Jaitzig, J. Li, R.D. Süssmuth, P. Neubauer, Reconstituted Biosynthesis of the Nonribosomal Macrolactone Antibiotic Valinomycin in Escherichia coli, ACS Synth. Biol. 3 (2014) 432–438. https://doi.org/10.1021/sb400082j.

[35] S.C. Wenzel, F. Gross, Y. Zhang, J. Fu, A.F. Stewart, R. Müller, Heterologous Expression of a Myxobacterial Natural Products Assembly Line in Pseudomonads via Red/ET Recombineering, Chem. Biol. 12 (2005) 349–356. https://doi.org/10.1016/j.chembiol.2004.12.012.

[36] Z. Shao, H. Zhao, H. Zhao, DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways, Nucleic Acids Res. 37 (2009) e16–e16. https://doi.org/10.1093/nar/gkn991.

[37] Z. Shao, Y. Luo, H. Zhao, Rapid characterization and engineering of natural product biosynthetic pathways via DNA assembler, Mol. Biosyst. 7 (2011) 1056.

https://doi.org/10.1039/c0mb00338g.

[38] V.T.T. Pham, C.T. Nguyen, D. Dhakal, H.T. Nguyen, T.-S. Kim, J.K. Sohng, Recent Advances in the Heterologous Biosynthesis of Natural Products from Streptomyces, Appl. Sci. 11 (2021) 1851. https://doi.org/10.3390/app11041851.

[39] S.E. Ongley, X. Bian, B.A. Neilan, R. Müller, Recent advances in the heterologous expression of microbial natural product biosynthetic pathways, Nat. Prod. Rep. 30 (2013) 1121. https://doi.org/10.1039/c3np70034h.

[40] J.J. Zhang, X. Tang, B.S. Moore, Genetic platforms for heterologous expression of microbial natural products, Nat. Prod. Rep. 36 (2019) 1313–1332. https://doi.org/10.1039/C9NP00025A.

[41] V.N. Noskov, B.J. Karas, L. Young, R.-Y. Chuang, D.G. Gibson, Y.-C. Lin, J. Stam, I.T. Yonemoto, Y. Suzuki, C. Andrews-Pfannkoch, J.I. Glass, H.O. Smith, C.A. Hutchison, J.C. Venter, P.D. Weyman, Assembly of Large, High G+C Bacterial DNA Fragments in Yeast, ACS Synth. Biol. 1 (2012) 267–273. https://doi.org/10.1021/sb3000194.

[42] N. Kouprina, V. Larionov, Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology, Chromosoma. 125 (2016) 621–632. https://doi.org/10.1007/s00412-016-0588-3.

[43] H.-J. Nah, H.-R. Pyeon, S.-H. Kang, S.-S. Choi, E.-S. Kim, Cloning and Heterologous Expression of a Large-sized Natural Product Biosynthetic Gene Cluster in Streptomyces Species, Front. Microbiol. 8 (2017). https://doi.org/10.3389/fmicb.2017.00394.

[44] K. Yamanaka, K.A. Reynolds, R.D. Kersten, K.S. Ryan, D.J. Gonzalez, V. Nizet, P.C. Dorrestein, B.S. Moore, Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A, Proc. Natl. Acad. Sci. 111 (2014) 1957–1962. https://doi.org/10.1073/pnas.1319584111.

[45] K.A. Reynolds, H. Luhavaya, J. Li, S. Dahesh, V. Nizet, K. Yamanaka, B.S. Moore, Isolation and structure elucidation of lipopeptide antibiotic taromycin B from the activated taromycin biosynthetic gene cluster, J. Antibiot. (Tokyo). 71 (2018) 333–338. https://doi.org/10.1038/ja.2017.146.

[46] M. Xu, G.D. Wright, Heterologous expression-facilitated natural products' discovery in actinomycetes, J. Ind. Microbiol. Biotechnol. 46 (2019) 415–431. https://doi.org/10.1007/s10295-018-2097-2.

[47] N.C.O. Lee, V. Larionov, N. Kouprina, Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast, Nucleic Acids Res. 43 (2015) e55–e55. https://doi.org/10.1093/nar/gkv112.

[48] J. Fu, X. Bian, S. Hu, H. Wang, F. Huang, P.M. Seibert, A. Plaza, L. Xia, R. Müller, A.F. Stewart, Y. Zhang, Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting, Nat. Biotechnol. 30 (2012) 440–446. https://doi.org/10.1038/nbt.2183.

[49] H. Wang, Z. Li, R. Jia, J. Yin, A. Li, L. Xia, Y. Yin, R. Müller, J. Fu, A.F. Stewart, Y. Zhang, ExoCET: exonuclease in vitro assembly combined with RecET recombination for highly efficient direct DNA cloning from complex genomes, Nucleic Acids Res. 46 (2018) e28–e28. https://doi.org/10.1093/nar/gkx1249.

[50] D. Du, L. Wang, Y. Tian, H. Liu, H. Tan, G. Niu, Genome engineering and direct cloning of antibiotic gene clusters via phage φBT1 integrase-mediated site-specific recombination in Streptomyces, Sci. Rep. 5 (2015) 8740. https://doi.org/10.1038/srep08740.

[51] S. Peirú, H.G. Menzella, E. Rodríguez, J. Carney, H. Gramajo, Production of the Potent Antibacterial Polyketide Erythromycin C in Escherichia coli, Appl. Environ. Microbiol. 71 (2005) 2539–2547. https://doi.org/10.1128/AEM.71.5.2539-2547.2005.

[52] H. Zhang, Y. Wang, J. Wu, K. Skalina, B.A. Pfeifer, Complete Biosynthesis of Erythromycin A and Designed Analogs Using E. coli as a Heterologous Host, Chem. Biol. 17 (2010) 1232–1240. https://doi.org/10.1016/j.chembiol.2010.09.013.

[53] J. Kumpfmüller, K. Methling, L. Fang, B.A. Pfeifer, M. Lalk, T. Schweder, Production of the polyketide 6-deoxyerythronolide B in the heterologous host Bacillus subtilis, Appl. Microbiol. Biotechnol. 100 (2016) 1209–1220. https://doi.org/10.1007/s00253-015-6990-6.

[54] S.C. Mutka, S.M. Bondi, J.R. Carney, N.A. Da Silva, J.T. Kealey, Metabolic pathway engineering for complex polyketide biosynthesis in Saccharomyces cerevisiae, FEMS

Yeast Res. 6 (2006) 40–47. https://doi.org/10.1111/j.1567-1356.2005.00001.x.

[55]   B.A. Pfeifer, Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of E. coli, Science (80-.). 291 (2001) 1790–1792. https://doi.org/10.1126/science.1058092.

[56]   S. Peirú, E. Rodríguez, H.G. Menzella, J.R. Carney, H. Gramajo, Metabolically engineered Escherichia coli for efficient production of glycosylated natural products, Microb. Biotechnol. 1 (2008) 476–486. https://doi.org/10.1111/j.1751-7915.2008.00046.x.

[57]   C.-G. Hyun, T. Bililign, J. Liao, J.S. Thorson, The Biosynthesis of Indolocarbazoles in a Heterologous E. coli Host, ChemBioChem. 4 (2003) 114–117. https://doi.org/10.1002/cbic.200390004.

[58]   N. Zaburannyi, M. Rabyk, B. Ostash, V. Fedorenko, A. Luzhetskyy, Insights into naturally minimised Streptomyces albus J1074 genome, BMC Genomics. 15 (2014) 97. https://doi.org/10.1186/1471-2164-15-97.

[59]   U. Galm, B. Shen, Expression of biosynthetic gene clusters in heterologous hosts for natural product production and combinatorial biosynthesis, Expert Opin. Drug Discov. 1 (2006) 409–437. https://doi.org/10.1517/17460441.1.5.409.

[60]   R.H. Baltz, Streptomyces and Saccharopolyspora hosts for heterologous expression of secondary metabolite gene clusters, J. Ind. Microbiol. Biotechnol. 37 (2010) 759–772. https://doi.org/10.1007/s10295-010-0730-9.

[61]   F. Alberti, D.J. Leng, I. Wilkening, L. Song, M. Tosin, C. Corre, Triggering the expression of a silent gene cluster from genetically intractable bacteria results in scleric acid discovery, Chem. Sci. 10 (2019) 453–463. https://doi.org/10.1039/C8SC03814G.

[62]   J.P. Gomez-Escribano, M.J. Bibb, Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters, Microb. Biotechnol. 4 (2011) 207–215. https://doi.org/10.1111/j.1751-7915.2010.00219.x.

[63]   A. Thanapipatsiri, J. Claesen, J.-P. Gomez-Escribano, M. Bibb, A. Thamchaipenet, A Streptomyces coelicolor host for the heterologous expression of Type III polyketide synthase genes, Microb. Cell Fact. 14 (2015) 145. https://doi.org/10.1186/s12934-015-

0335-0.

[64] M. Zhou, X. Jing, P. Xie, W. Chen, T. Wang, H. Xia, Z. Qin, Sequential deletion of all the polyketide synthase and nonribosomal peptide synthetase biosynthetic gene clusters and a 900-kb subtelomeric sequence of the linear chromosome of Streptomyces coelicolor, FEMS Microbiol. Lett. 333 (2012) 169–179. https://doi.org/10.1111/j.1574-6968.2012.02609.x.

[65] Y. Ahmed, Y. Rebets, M.R. Estévez, J. Zapp, M. Myronovskyi, A. Luzhetskyy, Engineering of Streptomyces lividans for heterologous expression of secondary metabolite gene clusters, Microb. Cell Fact. 19 (2020) 5. https://doi.org/10.1186/s12934-020-1277-8.

[66] Q. Peng, G. Gao, J. Lü, Q. Long, X. Chen, F. Zhang, M. Xu, K. Liu, Y. Wang, Z. Deng, Z. Li, M. Tao, Engineered Streptomyces lividans Strains for Optimal Identification and Expression of Cryptic Biosynthetic Gene Clusters, Front. Microbiol. 9 (2018). https://doi.org/10.3389/fmicb.2018.03042.

[67] M. Myronovskyi, B. Rosenkränzer, S. Nadmid, P. Pujic, P. Normand, A. Luzhetskyy, Generation of a cluster-free Streptomyces albus chassis strains for improved heterologous expression of secondary metabolite clusters, Metab. Eng. 49 (2018) 316–324. https://doi.org/10.1016/j.ymben.2018.09.004.

[68] D. Kallifidas, G. Jiang, Y. Ding, H. Luesch, Rational engineering of Streptomyces albus J1074 for the overexpression of secondary metabolite gene clusters, Microb. Cell Fact. 17 (2018) 25. https://doi.org/10.1186/s12934-018-0874-2.

[69] M. Komatsu, T. Uchiyama, S. Omura, D.E. Cane, H. Ikeda, Genome-minimized Streptomyces host for the heterologous expression of secondary metabolism, Proc. Natl. Acad. Sci. 107 (2010) 2646–2651. https://doi.org/10.1073/pnas.0914833107.

[70] Q.-T. Bu, P. Yu, J. Wang, Z.-Y. Li, X.-A. Chen, X.-M. Mao, Y.-Q. Li, Rational construction of genome-reduced and high-efficient industrial Streptomyces chassis based on multiple comparative genomic approaches, Microb. Cell Fact. 18 (2019) 16. https://doi.org/10.1186/s12934-019-1055-7.

[71] W.S. Jung, S.K. Lee, J.S.J. Hong, S.R. Park, S.J. Jeong, A.R. Han, J.K. Sohng, B.G. Kim, C.Y. Choi, D.H. Sherman, Y.J. Yoon, Heterologous expression of tylosin

polyketide synthase and production of a hybrid bioactive macrolide in Streptomyces venezuelae, Appl. Microbiol. Biotechnol. 72 (2006) 763–769. https://doi.org/10.1007/s00253-006-0318-5.

[72] V. Miao, M.-F. Coëffet-Le Gal, K. Nguyen, P. Brian, J. Penn, A. Whiting, J. Steele, D. Kau, S. Martin, R. Ford, T. Gibson, M. Bouchard, S.K. Wrigley, R.H. Baltz, Genetic Engineering in Streptomyces roseosporus to Produce Hybrid Lipopeptide Antibiotics, Chem. Biol. 13 (2006) 269–276. https://doi.org/10.1016/j.chembiol.2005.12.012.

[73] Q. Liu, L. Xiao, Y. Zhou, K. Deng, G. Tan, Y. Han, X. Liu, Z. Deng, T. Liu, Development of Streptomyces sp. FR-008 as an emerging chassis, Synth. Syst. Biotechnol. 1 (2016) 207–214. https://doi.org/10.1016/j.synbio.2016.07.002.

[74] E.M. Widenbrant, H.-H. Tsai, C.W. Chen, C.M. Kao, Spontaneous Amplification of the Actinorhodin Gene Cluster in Streptomyces coelicolor Involving Native Insertion Sequence IS 466, J. Bacteriol. 190 (2008) 4754–4758. https://doi.org/10.1128/JB.00131-08.

[75] S. Horinouchi, T. Beppu, Regulation of secondary metabolism and cell differentiation in Streptomyces: A-factor as a microbial hormone and the AfsR protein as a component of a two-component regulatory system, Gene. 115 (1992) 167–172. https://doi.org/10.1016/0378-1119(92)90555-4.

[76] H. Hu, Q. Zhang, K. Ochi, Activation of Antibiotic Biosynthesis by Specified Mutations in the rpoB Gene (Encoding the RNA Polymerase β Subunit) of Streptomyces lividans, J. Bacteriol. 184 (2002) 3984–3991. https://doi.org/10.1128/JB.184.14.3984-3991.2002.

[77] B. Baral, A. Akhgari, M. Metsä-Ketelä, Activation of microbial secondary metabolic pathways: Avenues and challenges, Synth. Syst. Biotechnol. 3 (2018) 163–178. https://doi.org/10.1016/j.synbio.2018.09.001.

[78] J. Shima, A. Hesketh, S. Okamoto, S. Kawamoto, K. Ochi, Induction of actinorhodin production by rpsL (encoding ribosomal protein S12) mutations that confer streptomycin resistance in Streptomyces lividans and Streptomyces coelicolor A3(2), J. Bacteriol. 178 (1996) 7276–7284. https://doi.org/10.1128/jb.178.24.7276-7284.1996.

[79] M. Lopatniuk, M. Myronovskyi, A. Nottebrock, T. Busche, J. Kalinowski, B. Ostash,

V. Fedorenko, A. Luzhetskyy, Effect of "ribosome engineering" on the transcription level and production of S. albus indigenous secondary metabolites, Appl. Microbiol. Biotechnol. 103 (2019) 7097–7110. https://doi.org/10.1007/s00253-019-10005-y.

[80] P.N. Schwarz, A. Buchmann, L. Roller, A. Kulik, H. Gross, W. Wohlleben, E. Stegmann, The Immunosuppressant Brasilicardin: Determination of the Biosynthetic Gene Cluster in the Heterologous Host Amycolatopsis japonicum, Biotechnol. J. 13 (2018) 1700527. https://doi.org/10.1002/biot.201700527.

[81] E. Rodriguez, Z. Hu, S. Ou, Y. Volchegursky, C.R. Hutchinson, R. McDaniel, Rapid engineering of polyketide overproduction by gene transfer to industrially optimized strains, J. Ind. Microbiol. Biotechnol. 30 (2003) 480–488. https://doi.org/10.1007/s10295-003-0045-1.

[82] J.J. Zhang, B.S. Moore, X. Tang, Engineering Salinispora tropica for heterologous expression of natural product biosynthetic gene clusters, Appl. Microbiol. Biotechnol. 102 (2018) 8437–8446. https://doi.org/10.1007/s00253-018-9283-z.

[83] N. Kallscheuer, H. Kage, L. Milke, M. Nett, J. Marienhagen, Competence of Corynebacterium glutamicum as a host for the production of type I polyketides, (2019). https://doi.org/10.1101/622399.

[84] M. Myronovskyi, A. Luzhetskyy, Heterologous production of small molecules in the optimized: Streptomyces hosts, Nat. Prod. Rep. 36 (2019) 1281–1294. https://doi.org/10.1039/c9np00023b.

[85] T. Siegl, B. Tokovenko, M. Myronovskyi, A. Luzhetskyy, Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes, Metab. Eng. 19 (2013) 98–106. https://doi.org/10.1016/j.ymben.2013.07.006.

[86] N. Seghezzi, P. Amar, B. Koebmann, P.R. Jensen, M.-J. Virolle, The construction of a library of synthetic promoters revealed some specific features of strong Streptomyces promoters, Appl. Microbiol. Biotechnol. 90 (2011) 615–623. https://doi.org/10.1007/s00253-010-3018-0.

[87] M. Myronovskyi, E. Brötz, B. Rosenkränzer, N. Manderscheid, B. Tokovenko, Y. Rebets, A. Luzhetskyy, Generation of new compounds through unbalanced

transcription of landomycin A cluster, Appl. Microbiol. Biotechnol. 100 (2016) 9175–9186. https://doi.org/10.1007/s00253-016-7721-3.

[88]   D. Kallifidas, H.-S. Kang, S.F. Brady, Tetarimycin A, an MRSA-Active Antibiotic Identified through Induced Expression of Environmental DNA Gene Clusters, J. Am. Chem. Soc. 134 (2012) 19552–19555. https://doi.org/10.1021/ja3093828.

[89]   A. Trefzer, G. Blanco, L. Remsing, E. Künzel, U. Rix, F. Lipata, A.F. Braña, C. Méndez, J. Rohr, A. Bechthold, J.A. Salas, Rationally Designed Glycosylated Premithramycins: Hybrid Aromatic Polyketides Using Genes from Three Different Biosynthetic Pathways, J. Am. Chem. Soc. 124 (2002) 6056–6062. https://doi.org/10.1021/ja017385l.

[90]   C. Sanchez, L. Zhu, A.F. Brana, A.P. Salas, J. Rohr, C. Mendez, J.A. Salas, From The Cover: Combinatorial biosynthesis of antitumor indolocarbazole compounds, Proc. Natl. Acad. Sci. 102 (2005) 461–466. https://doi.org/10.1073/pnas.0407809102.

[91]   M. Rodríguez Estévez, N. Gummerlich, M. Myronovskyi, J. Zapp, A. Luzhetskyy, Benzanthric Acid, a Novel Metabolite From Streptomyces albus Del14 Expressing the Nybomycin Gene Cluster, Front. Chem. 7 (2020). https://doi.org/10.3389/fchem.2019.00896.

2   Publications

I.

Loseolamycins:

A Group of New Bioactive Alkylresorcinols Produced after Heterologous Expression of a Type III PKS from Micromonospora endolithica

Constanze Lasch, Nils Gummerlich, Maksym Myronovskyi, Anja Palusczak,

Josef Zapp and Andriy Luzhetskyy

Loseolamycins: A Group of New Bioactive Alkylresorcinols Produced after Heterologous Expression of a Type III PKS from *Micromonospora endolithica*

Constanze Lasch [1], **Nils Gummerlich** [1], **Maksym Myronovskyi** [1], **Anja Palusczak** [1], **Josef Zapp** [2] **and Andriy Luzhetskyy** [1,3,*]

1   Pharmaceutical Biotechnology, Saarland University, 66123 Saarbruecken, Germany;
    constanze.lasch@uni-saarland.de (C.L.); nils.gummerlich@gmx.de (N.G.);
    maksym.myronovskyi@uni-saarland.de (M.M.); anja.palusczak@uni-saarland.de (A.P.)
2   Pharmaceutical Biology, Saarland University, 66123 Saarbruecken, Germany; j.zapp@mx.uni-saarland.de
3   Helmholtz Institute for Pharmaceutical Research Saarland, 66123 Saarbruecken, Germany
*   Correspondence: a.luzhetskyy@mx.uni-saarland.de; Tel.: +49-681-302-70200

Abstract: Natural products are a valuable source of biologically active compounds with potential applications in medicine and agriculture. Unprecedented scaffold diversity of natural products and biocatalysts from their biosynthetic pathways are of fundamental importance. Heterologous expression and refactoring of natural product biosynthetic pathways are generally regarded as a promising approach to discover new secondary metabolites of microbial origin. Here, we present the identification of a new group of alkylresorcinols after transcriptional activation and heterologous expression of the type III polyketide synthase of *Micromonospora endolithica*. The most abundant compounds loseolamycins A1 and A2 have been purified and their structures were elucidated by NMR. Loseolamycins contain an unusual branched hydroxylated aliphatic chain which is provided by the host metabolism and is incorporated as a starter fatty acid unit. The isolated loseolamycins show activity against gram-positive bacteria and inhibit the growth of the monocot weed *Agrostis stolonifera* in a germination assay. The biosynthetic pathway leading to the production of loseolamycins is proposed in this paper.

1. Introduction

In the last century, actinobacteria have proven to be a reliable source of natural products with various biological activities. Many of these compounds have found commercial application in medicine and agriculture as antibiotics, anticancer drugs, immunosuppressives, anthelmintics, insecticides, etc. [1,2]. Besides their interesting biological activities, natural products are of great fundamental interest. Unprecedented structural entities of natural products and biosynthetic routes leading to their production are the object of study of numerous research groups [3–5]. For the above mentioned reasons, the discovery of new natural products still remains an important research topic.

The discovery of structurally new secondary metabolites of microbial origin has become a tough task. The classical approach for metabolite discovery which is based on the analysis of crude extracts of naturally occurring strains has lost its efficiency due to, often, the rediscovery of already known compounds. High rediscovery rates have led to the assumption that the potential of actinobacteria as a

source of new natural products is exhausted [6]. However, the constantly growing amount of genomic data implies that the potential of actinobacteria to produce secondary metabolites greatly exceeds the number of compounds isolated from the strains. Obviously, only few pathways from several dozens encoded in the genome of each single actinobacterial strain are efficiently expressed under standard laboratory conditions. A great amount of natural product biosynthetic pathways remains silent or is expressed on a very low level in laboratory, preventing the encoded compounds from being discovered. This situation enforced the need for new strategies to uncover the biosynthetic potential inhered in actinobacteria [7,8]. The availability of genome sequencing data as well as of genome-optimized actinobacterial chassis strains led to the development of a new heterologous expression approach for the discovery of new natural products [9]. The genes encoding biosynthetic pathways potentially leading to new compounds are first identified in the genomes with the help of established secondary metabolite genome mining tools. The identified pathways are then cloned and expressed in a panel of optimized heterologous host strains. This technique of heterologous gene expression allows the targeted expression of promising natural products encoding gene clusters and almost completely eliminates the problem of assignment of the identified natural products to the biosynthetic clusters responsible for their production [10]. Optimized levels of biosynthetic precursors and simplified metabolic background of heterologous expression hosts positively affect the production yields of the expressed compounds and simplify their purification for structure elucidation and biological activity studies [11,12]. Due to the mentioned advantages, the heterologous expression approach is regarded as superior to the classical one and is believed to solve the rediscovery problem [10].

In this work, we present the discovery and characterization of structurally new and biologically active compounds. The isolated compounds belong to the alkylresorcinol family of natural products and were given the name loseolamycins. The production of loseolamycins was achieved through the transcriptional activation of the gene encoding a type III polyketide synthase from *Micromonospora endolithica* and its heterologous expression in the cluster-free *Streptomyces albus* Del14 strain. The most abundant products, loseolamycins A1 and A2, were successfully isolated and their structures were solved by NMR. Loseolamycins possess an unusual branched aliphatic chain which is hydroxylated in the one but last position. Individual loseolamycins differ in the length of the chain and its saturation grade. The bioactivity studies revealed antibacterial and herbicidal activities of loseolamycins. The biosynthetic route leading to the production of loseolamycins is discussed in this paper.

2. Results and Discussion

2.1. Identification and Expression of the Type III PKS Gene

Strains of the class of actinobacteria not belonging to the ubiquitous genus of *Streptomyces* are usually referred to as rare. They are frequently underexplored and represent a remarkable source for new metabolites. In the frame of the project aiming to discover new natural products through heterologous expression, the rare actinomycete *Micromonospora endolithica* has been chosen as a source of secondary metabolite clusters. The putative clusters encoding the production of natural products were identified and annotated in the chromosome of the strain using the genome mining software antiSMASH [13]. To facilitate the heterologous expression of the identified clusters, a genomic library of *M. endolithica* was constructed on an integrative BAC vector. In this study, we focused on gene *LU17765_001730* encoding a putative type III polyketide synthase [14]. From this point on, the gene *LU17765_001730* is designated in the text as *losA*. The BAC I7 containing the *losA* gene was selected from the constructed genomic library. A 24-kb fragment of the *M. endolithica* chromosome is cloned in BAC I7. Besides the gene *losA*, 24 open reading frames were annotated within the cloned fragment (Figure 1, Table 1). In order to express the *losA* gene heterologously, the BAC I7 was transferred into the genome-minimized host *Streptomyces albus* Del14 by conjugation [15]. The obtained exconjugant strain *S. albus* I7 was analyzed for the production of new compounds. For this purpose, the constructed strain *S. albus* I7 harboring the *losA* gene was cultivated in the production medium DNPM. The *S. albus* Del14

strain was used as negative control. LC-MS analysis of the culture extracts did not reveal any new peaks in the culture of the recombinant strain. One of the possible reasons for the failed heterologous expression might be poor expression of the *losA* gene in heterologous environment.
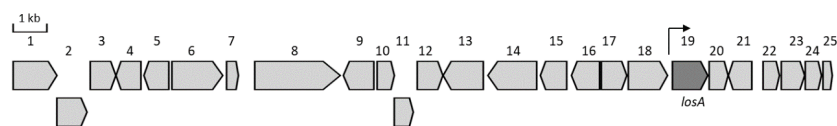
Figure 1. Gene organization of the *M. endolithica* chromosomal fragment cloned in BAC I7. The gene *losA* encoding the putative type III PKS is highlighted in dark grey. The arrow indicates the insertion position of the promoter cassette.

Table 1. Proposed functions of the genes in the chromosomal fragment cloned in BAC I7.

Gene	Identifier	Putative Function
1	*LU17765_001550*	L-aspartate oxidase
2	*LU17765_001560*	carboxylating nicotinate-nucleotide diphosphorylase
3	*LU17765_001570*	Type III pantothenate kinase
4	*LU17765_001580*	putative methyltransferase
5	*LU17765_001590*	putative methyltransferase
6	*LU17765_001600*	Lysine-tRNA ligase
7	*LU17765_001610*	Nucleoid-associated protein Lsr2
8	*LU17765_001620*	ATP-dependent Clp protease ATP-binding subunit
9	*LU17765_001630*	A/G-specific adenine glycosylase
10	*LU17765_001640*	ACT domain-containing protein
11	*LU17765_001650*	Peptide deformylase
12	*LU17765_001660*	Hypothetical protein
13	*LU17765_001670*	DNA integrity scanning protein DisA
14	*LU17765_001680*	DNA repair protein RadA
15	*LU17765_001690*	Hypothetical protein
16	*LU17765_001700*	UbiA family prenyltransferase
17	*LU17765_001710*	putative methyltransferase
18	*LU17765_001720*	acyl-CoA dehydrogenase
19 [*losA*]	*LU17765_001730*	type III polyketide synthase
20	*LU17765_001740*	Hypothetical protein
21	*LU17765_001750*	Hypothetical protein
22	*LU17765_001760*	CarD family transcriptional regulator
23	*LU17765_001770*	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase
24	*LU17765_001780*	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
25	*LU17765_001790*	tetratricopeptide repeat protein

To obtain the product of the type III PKS encoded by the *losA* gene, the latter was transcriptionally activated. The promoter cassette was PCR amplified and inserted in front of the *losA* gene by Red/ET [12,16]. In the obtained BAC I7act, the *losA* gene was set under the control of the strong synthetic promoter TS61 [17]. The constructed BAC was transferred into the *S. albus* Del14 strain by conjugation and the secondary metabolite profile of the obtained *S. albus* I7act strain was analyzed using LC-MS. This analysis revealed a number of new compounds in the extract of *S. albus* I7act which could not be observed in the extract of the control strain *S. albus* Del14 (Figure 2). Six most abundant compounds were the object of this study. For the sake of simplicity these compounds will be designated in the text by the numbers **1** to **6**. Besides UV/Vis detection, the compounds were analyzed with high-resolution mass spectrometry. Dissolved in methanol the compounds show a λ_{max} of 272 nm (Figure S1) and their experimental monoisotopic masses were calculated as follows: 336.2656 Da (compound **1**), 336.2657 Da (compound **2**), 334.2502 Da (compound **3**), 334.2497 Da (compound **4**), 348.2651 Da (compound **5**), and 362.2813 Da (compound **6**). The following molecular formulas were determined for the compounds **1** to **6**: $C_{21}H_{36}O_3$ (compound **1**), $C_{21}H_{36}O_3$ (compound **2**), $C_{21}H_{34}O_3$

(compound **3**), $C_{21}H_{34}O_3$ (compound **4**), $C_{22}H_{36}O_3$ (compound **5**), and $C_{23}H_{38}O_3$ (compound **6**) (Figure S2). The differences in the identified monoisotopic masses as well as the retention times of the compounds **1** to **6** indicate that they are probably structurally related. The pairs of the compounds **1**, **2** and **3**, **4** have similar monoisotopic masses, meaning that they might be constitutional isomers. The mass difference of 2 Da between the groups of compounds **1**, **2** and **3**, **4** implies that the corresponding compounds differ from each other in saturation grade of the alkyl chain. The mass difference of 14 Da between compound **5** and the group of compounds **3** and **4** suggests that the former contains one additional methylene group compared to the latter compounds. Similarly, compound **6** likely contains an additional methylene group compared to **5**.
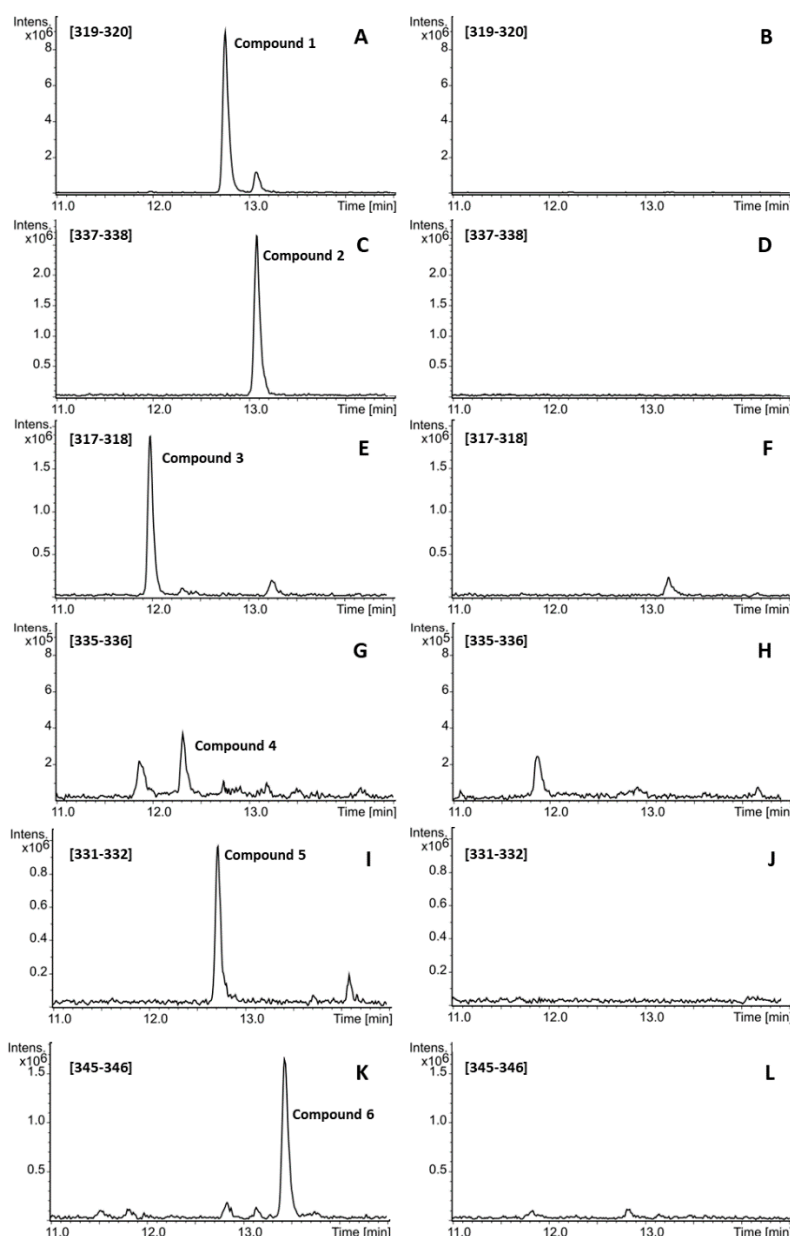
Figure 2. LC-MS analysis of loseolamycin production (compounds **1** to **6**) by the strain *S. albus* I7act. (**A,C,E,G,I,K**)—extracted ion chromatograms of crude extract of *S. albus* I7act showing the production of the compounds **1**, **2**, **3**, **4**, **5**, and **6**, respectively. (**B,D,F,H,J,L**)—extracted ion chromatograms of crude extract of the negative control strain *S. albus* Del14. Compounds **1** to **6** cannot be detected in the extract of the control strain. The extracted mass ranges are shown on each single chromatogram.

The identified high-resolution monoisotopic masses of the compounds **1** to **6** were used for a search in the Dictionary of Natural Products (DNP) database of natural products. This search did not generate any matches. Further, the metabolic profiling of the wild type strain *M. endolithica* did not reveal the mass signals of compounds **1** to **6**. We therefore assumed that the identified compounds **1** to **6** might be new and heterologous expression facilitates their production under laboratory conditions.

2.2. Purification and Structure Elucidation

The results of the high-resolution LC-MS analysis and database surveys indicate that the compounds identified after activation of the *losA* gene might be new. In order to obtain structural information about the detected compounds, we set out to purify them. For this purpose, the *S. albus* J7act strain was inoculated into 10 L of the production medium. The biosynthetic products were extracted from the culture broth with ethyl acetate and the organic solvent was evaporated under vacuum. The compounds **1** to **6** were purified using size-exclusion and reverse phase chromatography (Figure S3). Only the compounds **1** and **2** with the determined monoisotopic masses of 336.2656 Da and 336.2657 Da were isolated in quantities sufficient for structure elucidation by NMR spectroscopy (approximately 1 mg). The other four compounds (**3**, **4**, **5**, and **6**) were purified in submilligram quantities which were not sufficient to perform structure elucidation by NMR. All of the isolated compounds showed yellowish color and oily consistency.

The structure elucidation of compounds **1** and **2** was performed using nuclear magnet resonance spectroscopy (^{1}H-NMR, HSQCed, HMBC, ^{1}H-^{1}H-COSY, and ^{13}C-NMR) (Figure 3). All observed chemical shifts and correlations are presented in Table 2 (Figures S4–S11).
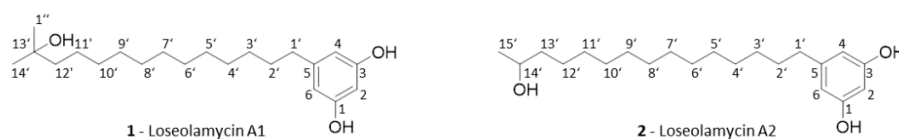
Figure 3. Structures of the loseolamycins A1 and A2 with numbered atoms according to the NMR data.

Table 2. NMR Spectroscopic data (500 MHz, d_6-DMSO) for loseolamycin A1 and loseolamycin A2.

	1-Loseolamycin A1			2-Loseolamycin A2	
Position	δ_C, Type	δ_H, Type (*J* [Hz])	HMBC [1]	Position	δ_H, Type
1	158.14, C	-	-	1	-
2	99.91, CH	5.98, s	3, 4, 6	2	5.98, m
3	158.14, C	-	-	3	-
4	106.24, CH	5.99, s	3, 6, 1′	4	5.98, m
5	144.15, C	-	-	5	-
6	106.24, CH	5.99, s	3, 4, 1′	6	5.98, m
1′	35.24, CH2	2.33, dd, 7.5 Hz	5, 4, 6, 2′, 8′	1′	2.32, m
2′	30.67, CH2	1.45, m	-	2′	1.44, m
3′	29.77, CH2	1.15–1.24, m	-	3′	1.15–1.34, m
4′	29.10, CH2	1.15–1.24, m	-	4′	1.15–1.34, m
5′	29.02, CH2	1.15–1.24, m	-	5′	1.15–1.34, m
6′	29.02, CH2	1.15–1.24, m	-	6′	1.15–1.34, m
7′	29.01, CH2	1.15–1.24, m	-	7′	1.15–1.34, m
8′	28.99, CH2	1.15–1.24, m	-	8′	1.15–1.34, m
9′	28.87, CH2	1.15–1.24, m	-	9′	1.15–1.34, m
10′	28.64, CH2	1.15–1.24, m	-	10′	1.15–1.34, m
11′	23.83, CH2	1.26, m	9′, 14′, 1″, 8′, 12′, 13′	11′	1.15–1.34, m
12′	43.67, CH2	1.29, m	13′, 14′, 1″, 10′	12′	1.44, m
13′	68.70, C	-	-	13′	1.52, m
14′	29.25, CH3	1.03, s	13′, 12′, 1″, 11′	14′	4.35, m
1″	29.25, CH3	1.03, s	13′, 12′, 14′, 11′	15′	1.49, d
1-OH	-	9.01, s br	-	1-OH	9.06, s br
3-OH	-	9.01, s br	-	3-OH	9.06, s br
13′-OH	-	4.01, s br	14′, 1″	14′-OH	5.30, s

[1] HMBC correlations, optimized for 6 Hz, are from protons stated to the indicated carbon.

The ^{1}H-NMR spectrum of the most abundant compound **1** in DMSO-d_6 showed signals for a 5-substituted *m*-dihydroxyphenol (δ_H 5.98, s, H-2; δ_H 5.99, s, 2H, H-4 and H-6; δ_H 9.01, s, 2H, 1-OH, and 3-OH) as well as for saturated long alkyl chain (δ_H 1.15–1.30, H-2′–H-12′). The singlet methyl at 1.03 ppm integrating to six protons together with the signal for a tertiary alcohol (δ_C 68.70, C-13′) in the ^{13}C-NMR spectrum gave hint to a terminal isopropyl alcohol moiety indicating a branched fatty alcohol. HMBC correlations (Table 2) supported these findings and led to the structure of the alkylresorcinol 5-(13-hydroxy-13-methyltetradecyl)benzene-1,3-diol, which was named loseolamycin A1 (compound **1**) (Figure 3).

Due to the low amount available, the structure elucidation of compound **2** with the monoisotopic mass of 336.2657 Da is solely based on the ^{1}H-NMR and ^{1}H-^{1}H-COSY experiments and on the calculated molecular formula (Table 2). The obtained signals of the *m*-dihydroxyphenol and the corresponding signals of the alkyl chain are identical to those of **1**. Therefore, the sole difference between loseolamycin A1 and the compound **2** is the terminal group of the fatty alcohol. The chemical shift of the methyl group of compound **2** increased slightly when compared to the dimethyl moiety of loseolamycin A1. Therefore, its signal was then located within the broad fatty acid proton peak between 1.15 and 1.34 ppm, when measured in d_6-DMSO. To circumvent this problem, a ^{1}H-NMR and a ^{1}H-^{1}H-COSY was recorded in CDCl$_3$. In the proton spectrum, a clear doublet becomes visible at 1.49 ppm, which shows a correlation in the COSY experiment to the signal of H-14′ at 4.35 ppm. The results of the NMR analysis and the observed monoisotopic mass allow for elucidation of compound **2** (5-(14-hydroxypentadecyl)benzene-1,3-diol). The compound **2** was named loseolamycin A2 (Figure 3). The determination of the optical rotation was not possible, because the substance had previously decomposed. The stereochemistry of the secondary alcohol C-14′ of loseolamycin A2 remains unsolved.

The compounds **3**, **4**, **5**, and **6** were purified in submilligram quantities which were not sufficient to perform NMR analysis. The determined monoisotopic masses and the calculated molecular formulas indicate that the compounds **3**, **4**, **5**, and **6** differ from loseolamycins A1 and A2 by saturation grade and the length of the alkyl chain (Figure S12).

2.3. Biosynthetic Scheme

Alkylresorcinols as a product of bacterial type III PKS were mechanistically investigated and described before [18]. We propose that the activated C16 to C18 hydroxy fatty acids from the precursor pool of *S. albus* Del14 serve as starter units for the biosynthesis of loseolamycins. The type III PKS encoded by the gene *losA* catalyzes the decarboxylative condensation of three malonyl-CoA units with the activated starter unit. The decarboxylative aldol cyclization of the nascent polyketide chain as well as its tautomerism lead to the favorable resorcinol form of the final product (Figure 4) [19].
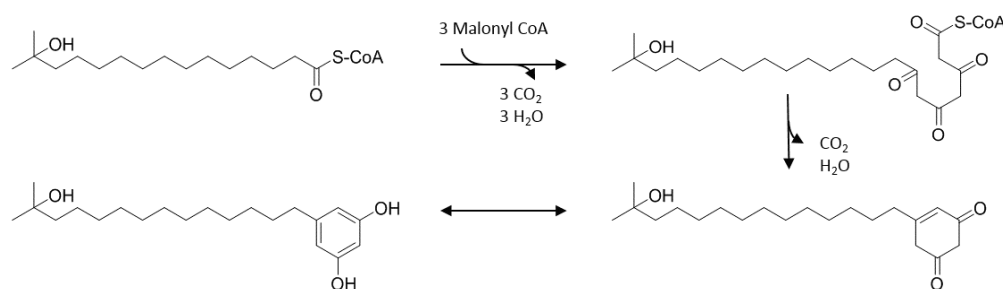
Figure 4. Proposed biosynthesis exemplarily for loseolamycin A1.

Bacterial type III PKS are often characterized by the lack of substrate specificity for the fatty acid starter unit [20–22]. Such promiscuity of the *losA* product might explain the production of numerous loseolamycin derivatives which differ in the length of their alkyl chain or its saturation state.

2.4. Biological Activity

Biological activity studies have been performed for the isolated loseolamycins. Agar disc diffusion tests with the test cultures *Escherichia coli*, *Pseudomonas putida*, and *Bacillus subtilis* were carried out to analyze if the compounds possess any antibacterial activity. No growth inhibition of *E. coli* and *P. putida* could be observed, indicating that loseolamycins A1 and A2 are not active against gram-negative bacteria. The compounds successfully inhibited the growth of the gram-positive bacterium *B. subtilis* (Figure S13). Additionally, loseolamycins were tested for herbicidal activity against the weed *Agrostis stolonifera*. Due to insufficient amounts of isolated individual loseolamycins, the prepurified fraction containing a mixture of loseolamycins with loseolamycin A1 as a major compound was used for the bioassay. Loseolamycins reproducibly demonstrated herbicidal activity by suppressing the germination of the *A. stolonifera* seeds (Figure S13). Further experiments are required to assess the activity of single derivatives and to propose a mode of action.

3. Materials and Methods

3.1. General Experimental Procedures

The strains, BACs and plasmids used in this work are listed in Table S1. The *Streptomyces* strains *S. albus* Del14, *S. albus* I7, and *S. albus* I7act were grown on soy flour mannitol (MS) agar [23] and in liquid tryptic soy broth (TSB; Sigma-Aldrich, St. Louis, MO, USA). Liquid DNPM medium (40 g/L dextrin, 7.5 g/L soytone, 5 g/L baking yeast, and 21 g/L MOPS, pH 6.8 as aqueous solution) was used for metabolite expression. DNPM as solid and liquid medium was used for cultivation of *Micromonospora endolithica*. *Bacillus subtilis*, *Pseudomonas putida*, and *Escherichia coli* strains were cultured in lysogeny broth (LB) medium [24]. The antibiotics kanamycin, apramycin, ampicillin, and nalidixic acid were added when required. Plant medium was prepared as aqueous solution containing 2.2 g/L Murashige and Skoog and 1.6 g/L Gamborg's B5 plant salts.

3.2. Isolation and Manipulation of DNA

DNA manipulation, transformation into *E. coli* as well as intergeneric conjugation between *E. coli* and *Streptomyces* were performed according to standard protocols [23–25]. BAC DNA from the constructed genomic library IG652BAC1-2 of *Micromonospora endolithica* was isolated with the BACMAX™ DNA purification kit (Lucigen, Middleton, WI, USA). Promoter TS61 was inserted on the BAC I7 using Red/ET [26]. A respective gene cassette from pUC19 plasmid harboring an ampicillin resistance was amplified by PCR. PCR primers 20180710_02_fw and 20180710_02_rev were constructed with overhang regions complementary to the DNA bases in front of *losA* gene for site specific introduction of the cassette. For control of the successful recombination, restriction mapping as well as sequencing were performed. The restriction enzymes were provided by ThermoFisher Scientific (Waltham, MA, USA) or New England BioLabs NEB (Ipswich, MA, USA) and used as outlined in the instruction manuals.

3.3. Metabolite Extraction

For metabolite extraction, *Streptomyces* strains were grown in 15 mL of TSB in a 100 mL baffled flask for 1 day, and 1 mL of seed culture was used to inoculate 100 mL of DNPM production medium in a 500 mL baffled flask. Cultures were grown for 7 to 8 days at 28 °C and 180 rpm in an Infors multitron shaker (Infors AG, Basel, Switzerland). Metabolites were extracted from the culture supernatant with an equal amount of ethyl acetate, evaporated at 40 °C and kept at storage condition at 4 °C. The strain *S. albus* I7act was used for 10-L scale and the above described procedures were carried out accordingly using 100 single flasks for main culture.

3.4. Mass Spectrometry (MS) Metabolite Analysis

Dry sample extracts were dissolved in methanol prior to MS analysis. Sample compounds were separated on a Dionex Ultimate 3000 UPLC system (ThermoFisher Scientific, Waltham, MA, USA) equipped with an ACQUITY BEH C18 column 1.7 μm, 2.1 mm × 30 or 50 or 100 mm (Waters Corporation, Milford, MA, USA). A linear gradient from hydrophilic to lipophilic mobile phase was run as follows: [A] water + 0.1% formic acid/[B] acetonitrile + 0.1% formic acid, 5% to 95% [B] at flow rate of 0.6 mL/min. Behind the PDA detector, an amaZon speed or LTQ Orbitrap XL mass spectrometer (Bruker, Billerica, MA, USA) were coupled to the system. LTQ Orbitrap XL provides a mass accuracy of 5 ppm. Mass analysis was performed using positive or negative ionization and a mass range selection of m/z 200 to 2000. In positive mode, loss of water occurs during ionization of most loseolamycin derivatives. Mass data was analyzed using the softwares Compass Data Analysis 4.1 and Xcalibur 3.

3.5. Purification

The crude extract from the 10-L culture of recombinant strain *S. albus* I7act was dissolved in methanol. Concentrated extract solution was loaded onto a LC column for a first purification step using size exclusion (SEC) as separation mode. The column was packed with a stationary phase of Sephadex-LH20 and run with an isocrate of pure methanol. Eluates were checked on mass spectrometer for content and purity of loseolamycin derivatives and pooled to prepurified loseolamycin fractions. Pooled fractions were dried at 40 °C and stored in the refrigerator at 4 °C. For separation of loseolamycin derivatives, a second chromatographic step was performed. The fraction pool was redissolved in methanol and applied on a reversed phase (RP). An Agilent Infinity 1200 series HPLC system equipped with Synergi^{TM} 4 μm Fusion-RP 80 Å 250 × 10 column (phenomenex, Torrance, CA, USA) was used and elution was carried out as follows: linear gradient of [A] water + 0.1% formic acid/[B] acetonitrile + 0.1% formic acid, 50% to 95% [B] in 20 min at flow rate of 4 mL/min and column oven temperature 45 °C. Fractions were controlled with HPLC-MS before final pooling to the six pure isolates **1** to **6**.

3.6. Nuclear Magnetic Resonance (NMR) Spectroscopy

The isolated compounds were dissolved in 300 μL deuterated solvent (DMSO-d_6, MEOD-d_4, CDCl$_3$) and measured in a corresponding 5-mm Shigemi-tube (DEUTERO GmbH, Kastellaun, Germany). NMR data (^{1}H, HH-COSY; TOCSY; HMBC, HSQC, ^{13}C) were acquired with standard pulse programs on either a Bruker Ascend 700 spectrometer equipped with a 5-mm TXA Cryoprobe at 300K or a Bruker Avance 500 spectrometer equipped with a 5-mm BBO Probe (Bruker BioSpin GmbH, Rheinstetten, Germany). Bruker TopSpin 3.5a software was used for the evaluation of NMR results.

3.7. Activity Testing

For testing of antibacterial activity, filter discs were impregnated with methanolic solutions of pure loseolamycins A1 and A2 and dried for 20 min. Meanwhile a concentrated cell suspension of each *E. coli*, *B. subtilis* and *P. putida*, was mixed with 1 mL of LB soft agar and poured onto a thin LB agar plate. The dried filter discs were placed onto the prepared plates and incubated at 29 °C for 15 h. Visual evaluation of the inhibition zone was performed against a methanol negative control which was processed as described before for the sample disc.

The herbicidal activity assay against the weed *A. stolonifera* was performed in 96-well plates. Serial dilutions of the sample containing loseolamycins were performed in the plant medium reducing the concentration of the compound by a factor of 0.5 each. 10 seeds of the test plant were put in each well containing 195 μL of medium and cultivated in a greenhouse with fluorescent light for 6 days at room temperature. As a control, dilutions of methanol in plant medium were used. The number of germinated seedlings was determined by visual analysis.

3.8. Genome Mining and Bioinformatic Analysis

The genome of *M. endolithica* was screened for secondary metabolite biosynthetic gene clusters using the antiSMASH online tool (https://antismash.secondarymetabolites.org/#!/start) [13]. Analysis of genetic data was performed using the Geneious software, version 11.0.3 [27]. The genomic sequence of BAC I7 was deposited in GenBank under accession number MT904273. For dereplication, the Dictionary of Natural Products (DNP) 28.1 was used as database of known natural products.

4. Conclusions

The production of secondary metabolites from actinobacteria strongly depends on the bacterial needs and many gene clusters remain silent under laboratory working conditions. Heterologous expression of a promoter activated type III PKS gene of *Micromonospora endolithica* facilitated the production of structurally new bacterial alkylresorcinols that we called loseolamycins. Unusual hydroxylated C16 to C18 fatty acids from the host's primary metabolism are utilized as starter units. The here isolated loseolamycins add further structural diversity to the class of alkylresorcinols. Furthermore, the new compounds exhibit herbicidal and antibacterial activity that might drive future research efforts.

References

1. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770–803. [CrossRef] [PubMed]

2. Shen, B. A New Golden Age of Natural Products Drug Discovery. *Cell* **2015**, *163*, 1297–1300. [CrossRef] [PubMed]

3. Myronovskyi, M.; Rosenkränzer, B.; Stierhof, M.; Petzke, L.; Seiser, T.; Luzhetskyy, A. Identification and Heterologous Expression of the Albucidin Gene Cluster from the Marine Strain Streptomyces Albus Subsp. Chlorinus NRRL B-24108. *Microorganisms* **2020**, *8*, 237. [CrossRef] [PubMed]

4. Keller, L.; Plaza, A.; Dubiella, C.; Groll, M.; Kaiser, M.; Müller, R. Macyranones: Structure, Biosynthesis, and Binding Mode of an Unprecedented Epoxyketone that Targets the 20S Proteasome. *J. Am. Chem. Soc.* **2015**, *137*, 8121–8130. [CrossRef] [PubMed]

5. Daum, M.; Peintner, I.; Linnenbrink, A.; Frerich, A.; Weber, M.; Paululat, T.; Bechthold, A. Organisation of the Biosynthetic Gene Cluster and Tailoring Enzymes in the Biosynthesis of the Tetracyclic Quinone Glycoside Antibiotic Polyketomycin. *ChemBioChem* **2009**, *10*, 1073–1083. [CrossRef]

6. Baltz, R.H. Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *J. Ind. Microbiol. Biotechnol.* **2006**, *33*, 507–513. [CrossRef]

7. Harvey, A.L.; Edrada-Ebel, R.; Quinn, R.J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* **2015**, *14*, 111–129. [CrossRef]

8. Newman, D. Screening and identification of novel biologically active natural compounds. *F1000Research* **2017**, *6*, 783. [CrossRef]

9. Luo, Y.; Cobb, R.E.; Zhao, H. Recent advances in natural product discovery. *Curr. Opin. Biotechnol.* **2014**, *30*, 230–237. [CrossRef]

10. Zarins-Tutt, J.S.; Barberi, T.T.; Gao, H.; Mearns-Spragg, A.; Zhang, L.; Newman, D.J.; Goss, R.J.M. Prospecting for new bacterial metabolites: A glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Nat. Prod. Rep.* **2016**, *33*, 54–72. [CrossRef]

11. Baltz, R.H. Renaissance in antibacterial discovery from actinomycetes. *Curr. Opin. Pharmacol.* **2008**, *8*, 557–563. [CrossRef]

12. Myronovskyi, M.; Luzhetskyy, A. Heterologous production of small molecules in the optimized Streptomyces hosts. *Nat. Prod. Rep.* **2019**, *36*, 1281–1294. [CrossRef] [PubMed]

13. Medema, M.H.; Blin, K.; Cimermancic, P.; De Jager, V.; Zakrzewski, P.; Fischbach, M.A.; Weber, T.; Takano, E.; Breitling, R. antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **2011**, *39*, W339–W346. [CrossRef] [PubMed]

14. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]

15. Myronovskyi, M.; Rosenkränzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A. Generation of a cluster-free Streptomyces albus chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab. Eng.* **2018**, *49*, 316–324. [CrossRef]

16. Olano, C.; García, I.; González, A.; Rodriguez, M.; Rozas, D.; Rubio, J.; Sánchez-Hidalgo, M.; Braña, A.F.; Méndez, C.; Salas, J.A. Activation and identification of five clusters for secondary metabolites in Streptomyces albus J1074. *Microb. Biotechnol.* **2014**, *7*, 242–256. [CrossRef] [PubMed]

17. Siegl, T.; Tokovenko, B.; Myronovskyi, M.; Luzhetskyy, A. Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes. *Metab. Eng.* **2013**, *19*, 98–106. [CrossRef] [PubMed]

18. Yu, D.; Xu, F.; Zeng, J.; Zhan, J. Type III polyketide synthases in natural product biosynthesis. *IUBMB Life* **2012**, *64*, 285–295. [CrossRef] [PubMed]

19. Funa, N.; Awakawa, T.; Horinouchi, S. Pentaketide Resorcylic Acid Synthesis by Type III Polyketide Synthase from *Neurospora crassa*. *J. Biol. Chem.* **2007**, *282*, 14476–14481. [CrossRef]

20. Awakawa, T.; Fujita, N.; Hayakawa, M.; Ohnishi, Y.; Horinouchi, S. Characterization of the Biosynthesis Gene Cluster for Alkyl-O-Dihydrogeranyl-Methoxyhydroquinones in Actinoplanes missouriensis. *ChemBioChem* **2011**, *12*, 439–448. [CrossRef]

21. Funabashi, M.; Funa, N.; Horinouchi, S. Phenolic Lipids Synthesized by Type III Polyketide Synthase Confer Penicillin Resistance on *Streptomyces griseus*. *J. Biol. Chem.* **2008**, *283*, 13983–13991. [CrossRef] [PubMed]

22. Funa, N.; Ozawa, H.; Hirata, A.; Horinouchi, S. Phenolic lipid synthesis by type III polyketide synthases is essential for cyst formation in *Azotobacter vinelandii*. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 6356–6361. [CrossRef] [PubMed]

23. Kieser, T.; Bibb, M.J.; Buttner, M.J.; Chater, K.F.; Hopwood, D.A. *Practical Streptomyces Genetics*; John Innes Foundation: Norwich, UK, 2000.

24. Green, M.R.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*, 4th ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2012.

25. Rebets, Y.; Kormanec, J.; Luzhetskyy, A.; Bernaerts, K.; Anné, J. Cloning and Expression of Metagenomic DNA in Streptomyces lividans and Subsequent Fermentation for Optimized Production. *Methods Mol. Biol.* **2016**, *1539*, 99–144. [CrossRef]

26. Court, D.L.; Sawitzke, J.A.; Thomason, L.C. Genetic Engineering Using Homologous Recombination. *Annu. Rev. Genet.* **2002**, *36*, 361–388. [CrossRef] [PubMed]

27. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef] [PubMed]

1 　*Article*

2 　# Loseolamycins: a group of new bioactive
3 　# alkylresorcinols produced after heterologous
4 　# expression of a type III PKS from *Micromonospora*
5 　# *endolithica*

6 　**Constanze Lasch [1], Nils Gummerlich [1], Maksym Myronovskyi [1], Anja Palusczak [1], Josef Zapp [2]**
7 　**and Andriy Luzhetskyy [1,3,*]**

8 　[1]　Pharmazeutische Biotechnologie, Universität des Saarlandes, Saarbrücken, Germany;
9 　　　constanze.lasch@uni-saarland.de (C.L.), nils.gummerlich@gmx.de (N.G.);
10 　　　maksym.myronovskyi@uni-saarland.de (M.M.); anja.palusczak@uni-saarland.de (A.P.);
11 　　　a.luzhetskyy@mx.uni-saarland.de (A.L.)
12 　[2]　Pharmazeutische Biologie, Universität des Saarlandes, Saarbrücken, Germany; j.zapp@mx.uni-saarland.de
13 　　　(J.Z.)
14 　[3]　Helmholtz-Institut für Pharmazeutische Forschung Saarland, Saarbrücken, Germany
15 　*　Correspondence: a.luzhetskyy@mx.uni-saarland.de; +49 681 302 70200 (A.L.)

17
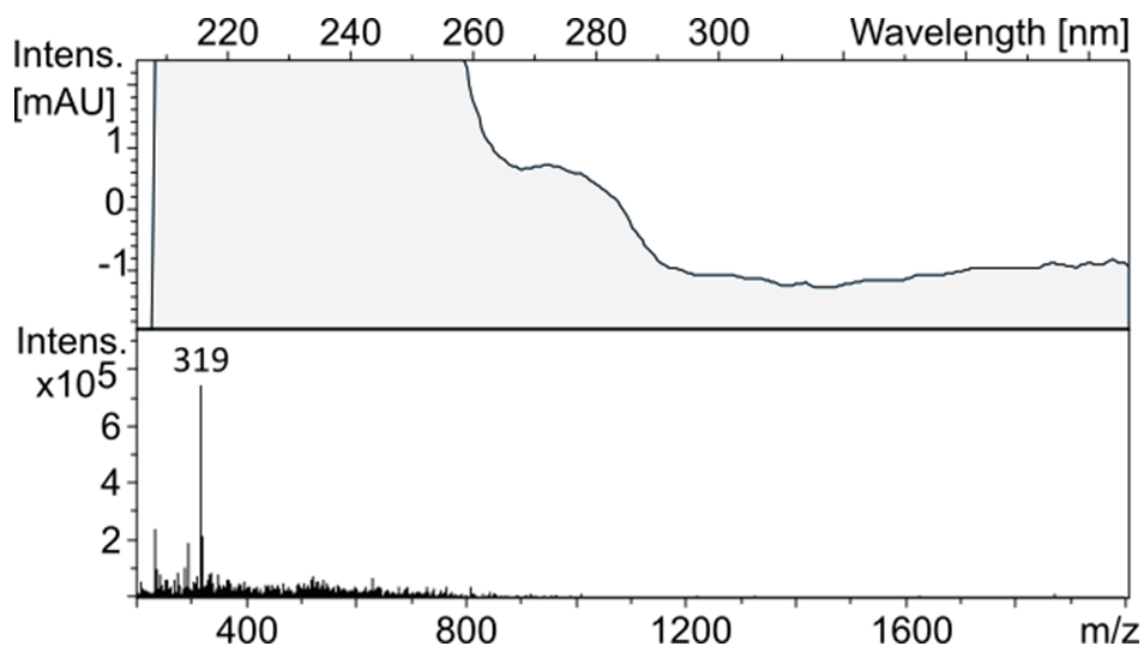
18

19   ***Figure S1.*** *UV/Vis spectrum of loseolamycin A1.*

20
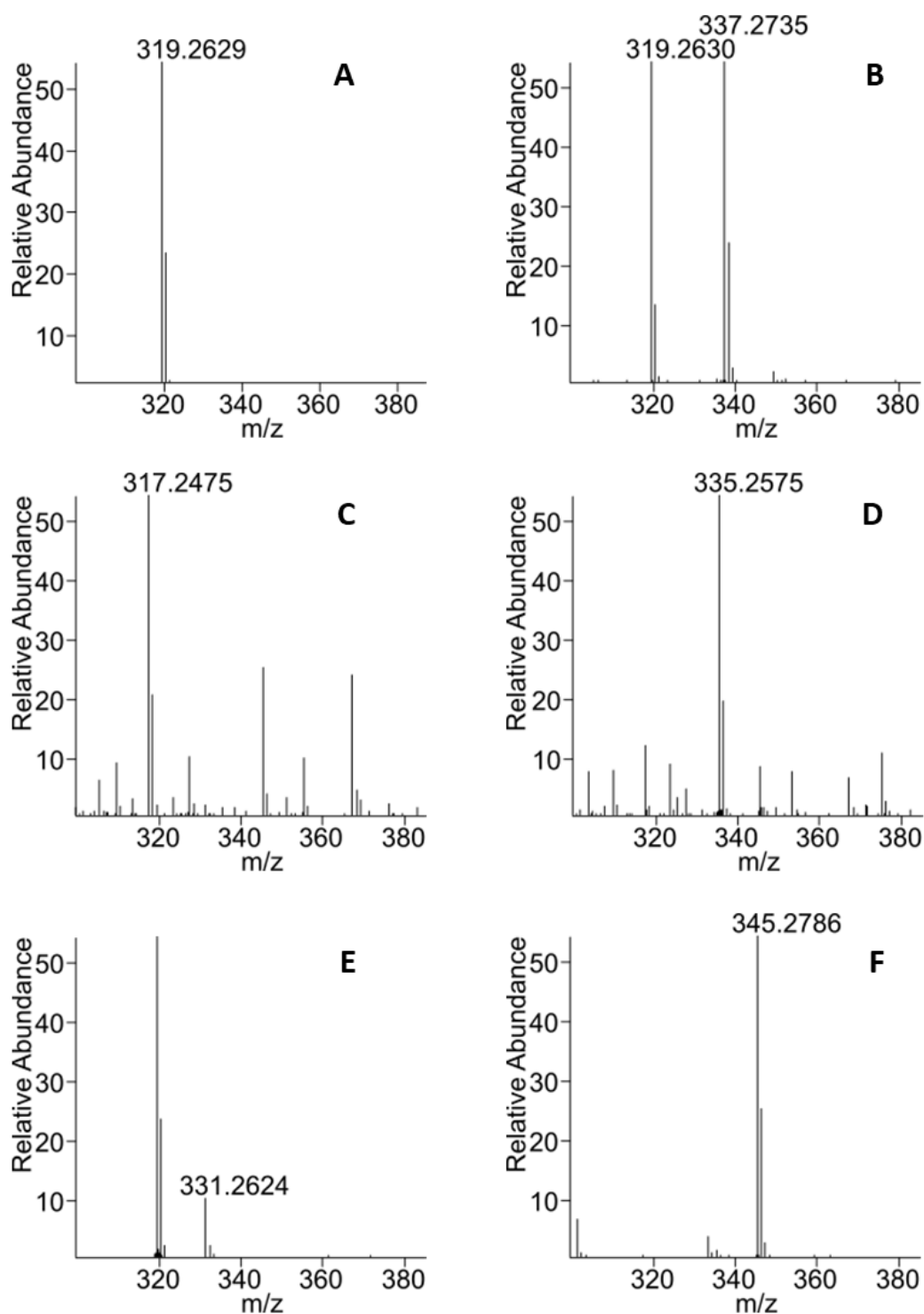
*Figure S2. Mass spectra of loseolamycins using positive ionization mode. A - compound **1**, B – compound **2**, C – compound*
*3, D – compound **4**, E – compound **5**, F – compound **6**.*
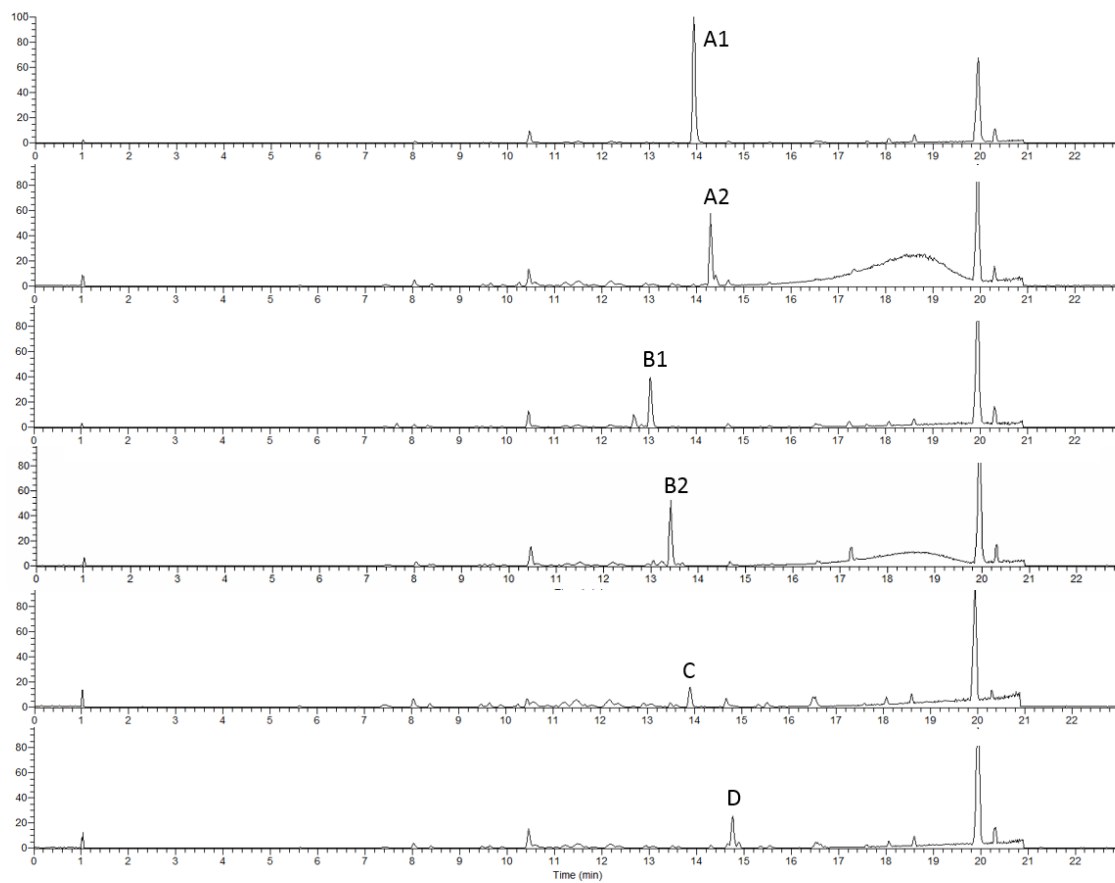
23

24 ***Figure S3.*** *Purity of the isolated loseolamycin derivatives A1 (compound **1**), A2 (compound **2**), B1 (compound **3**), B2*
25 *(compound **4**), C (compound **5**) and D (compound **6**) after two chromatographic steps. Base Peak Chromatograms (BPC)*
26 *are shown.*

27
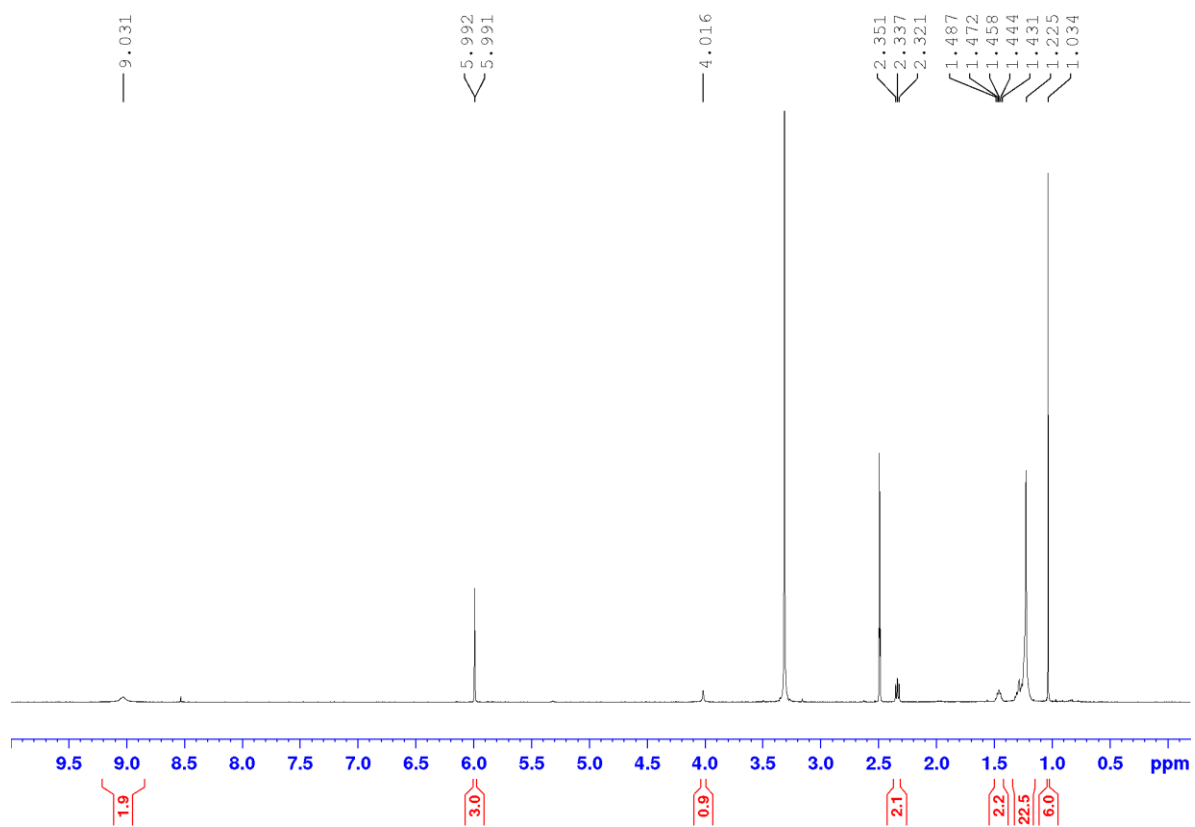
29 ***Figure S4.*** *¹H-NMR spectrum (500 MHz, DMSO-d6) of loseolamycin A1, complete spectrum.*

31
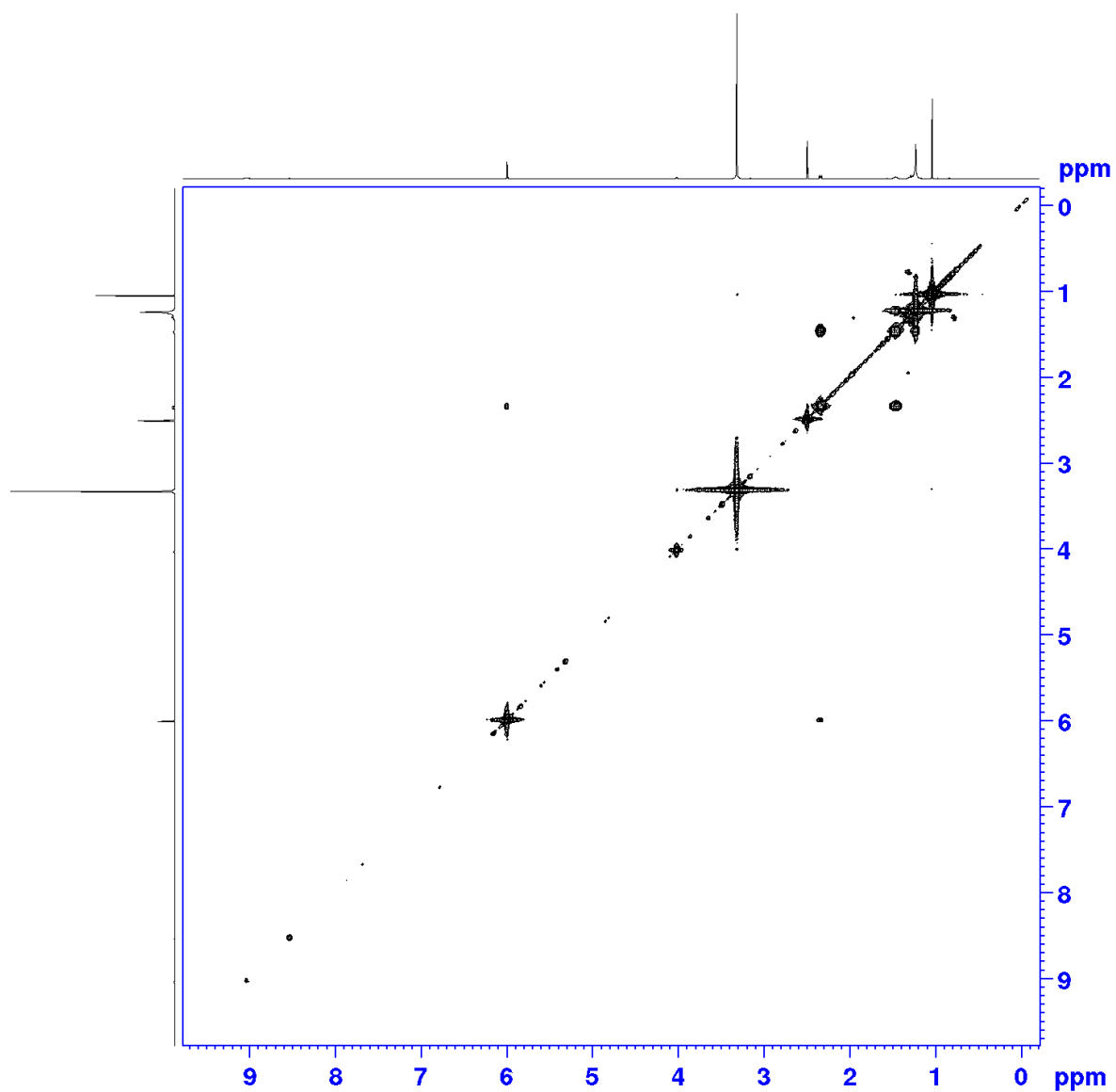
Figure S5. *¹H-¹H-COSY spectrum (500 MHz, DMSO-d6) of loseolamycin A1, complete spectrum.*

33
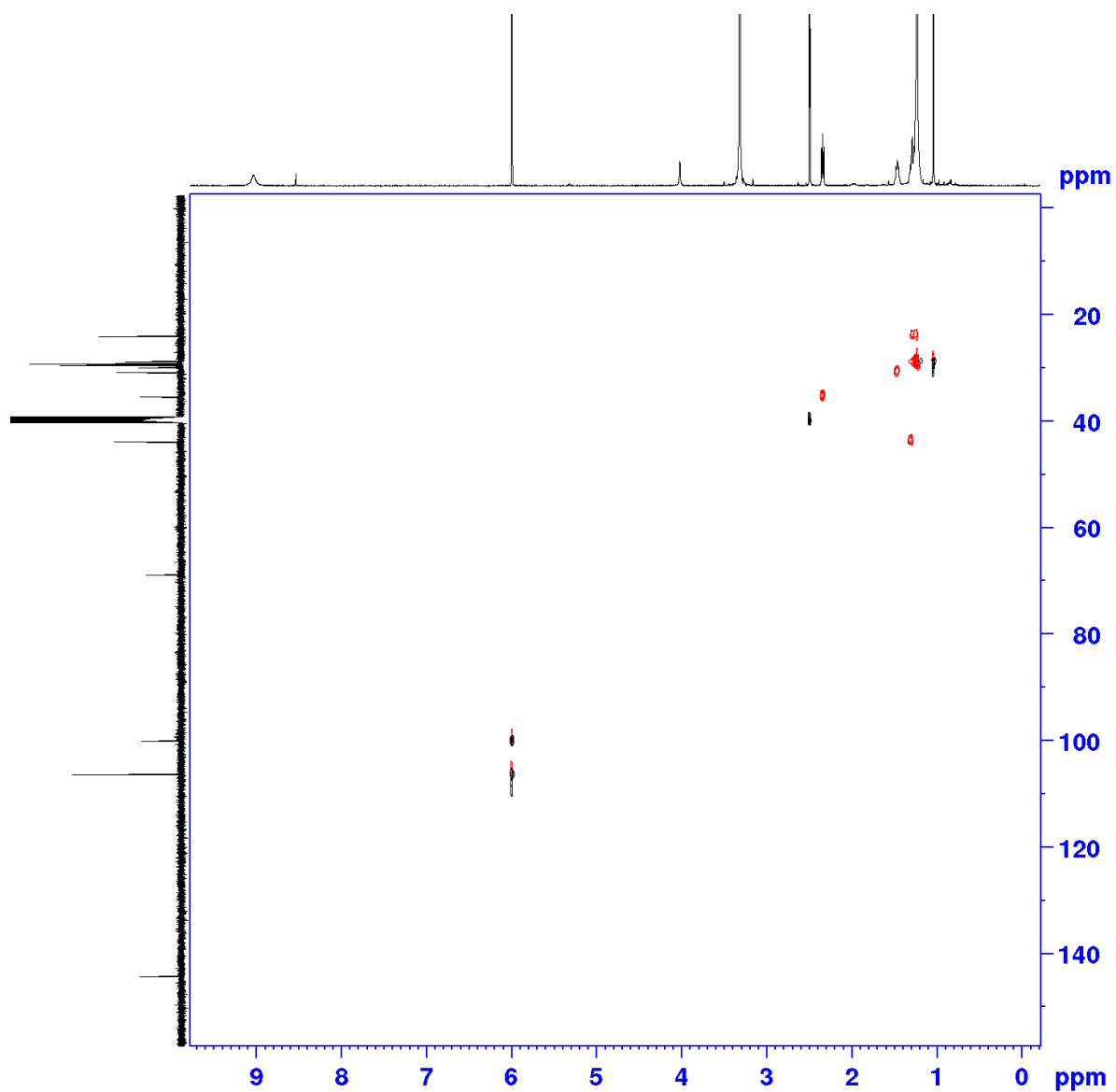
34

35 ***Figure S6.*** *HSQC-spectrum (500 MHz; 125 MHz, DMSO-d6) of loseolamycin A1, complete spectrum.*

36
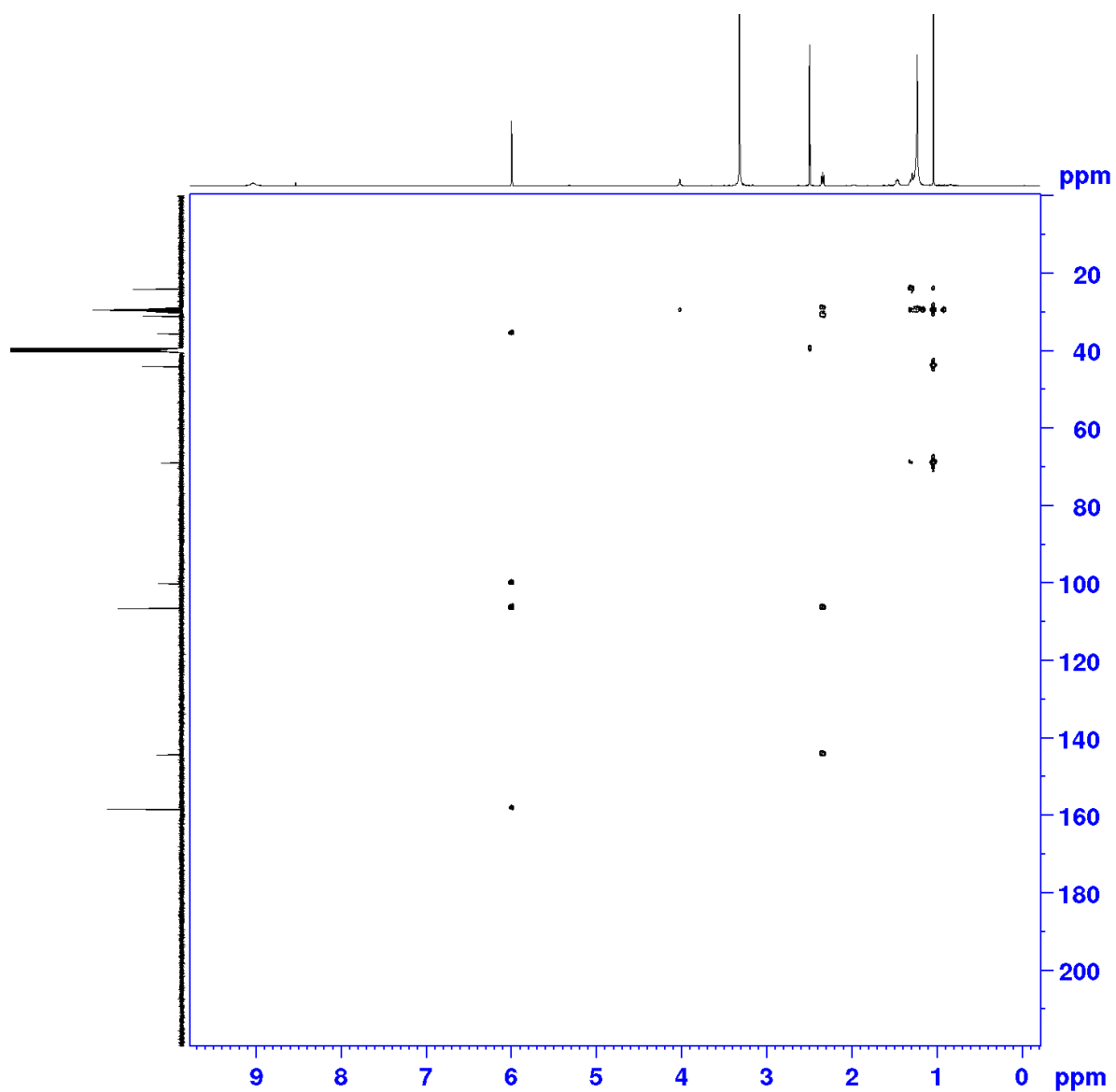
37
38   ***Figure S7.*** *HMBC-spectrum (500 MHz; 125 MHz, DMSO-d6) of loseolamycin A1, complete spectrum.*
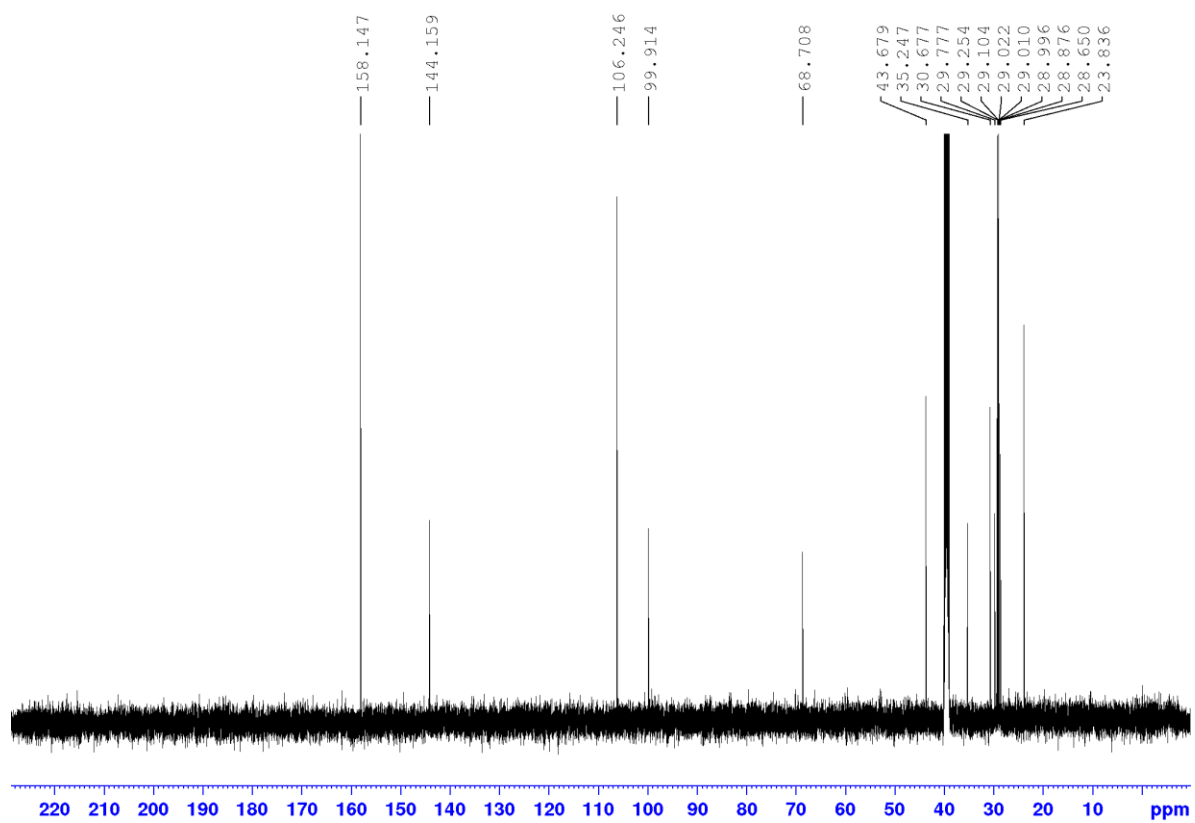
Figure S8. ^{13}C-spectrum (125 MHz, DMSO-d6) of loseolamycin A1, complete spectrum.
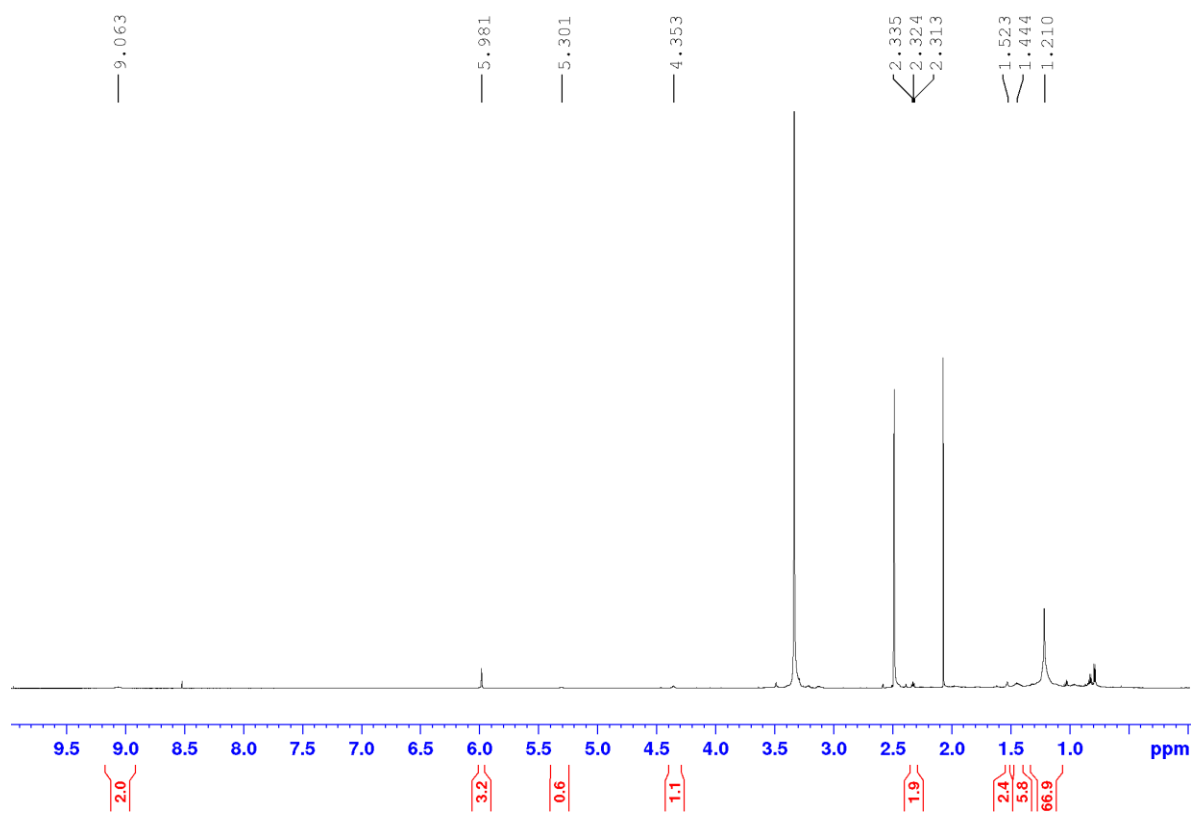
Figure S9. *¹H-NMR spectrum (500 MHz, DMSO-d6) of loseolamycin A2, complete spectrum.*
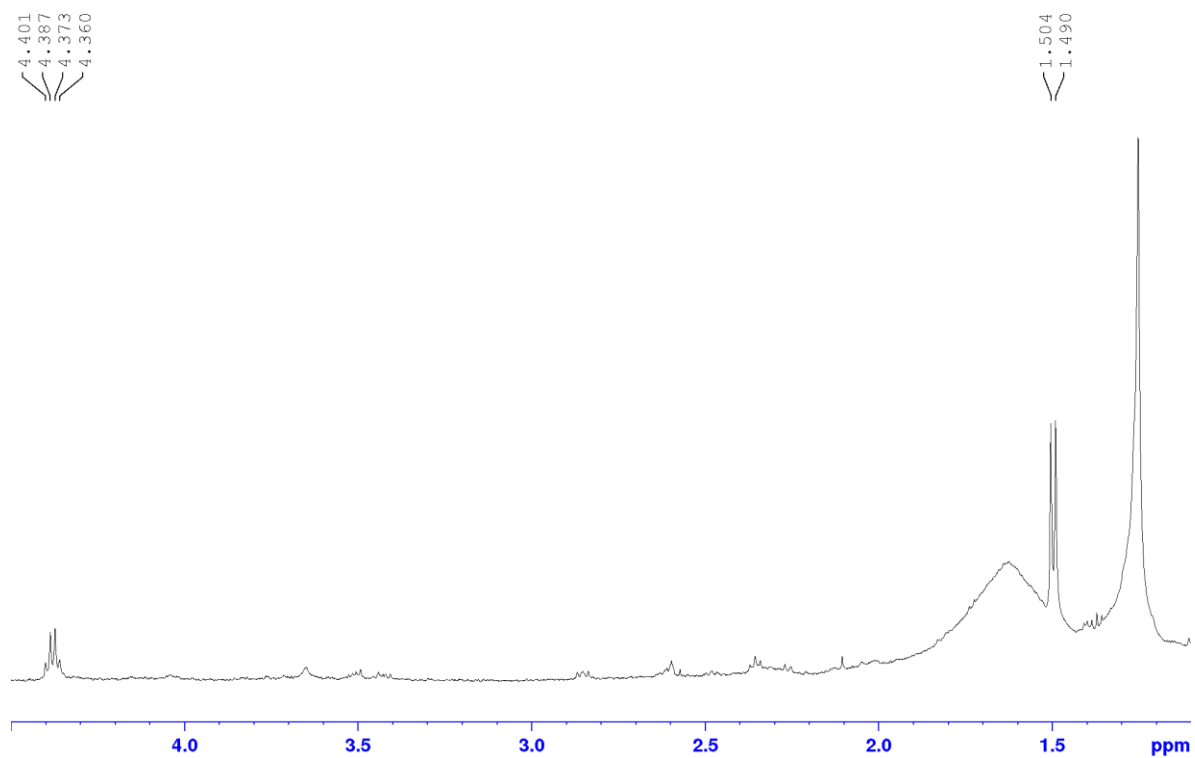
45

46   ***Figure S10.*** *^{1}H-NMR spectrum (500 MHz, CDCl₃) of loseolamycin A2, zoomed from 1.0 – 4.5 ppm.*
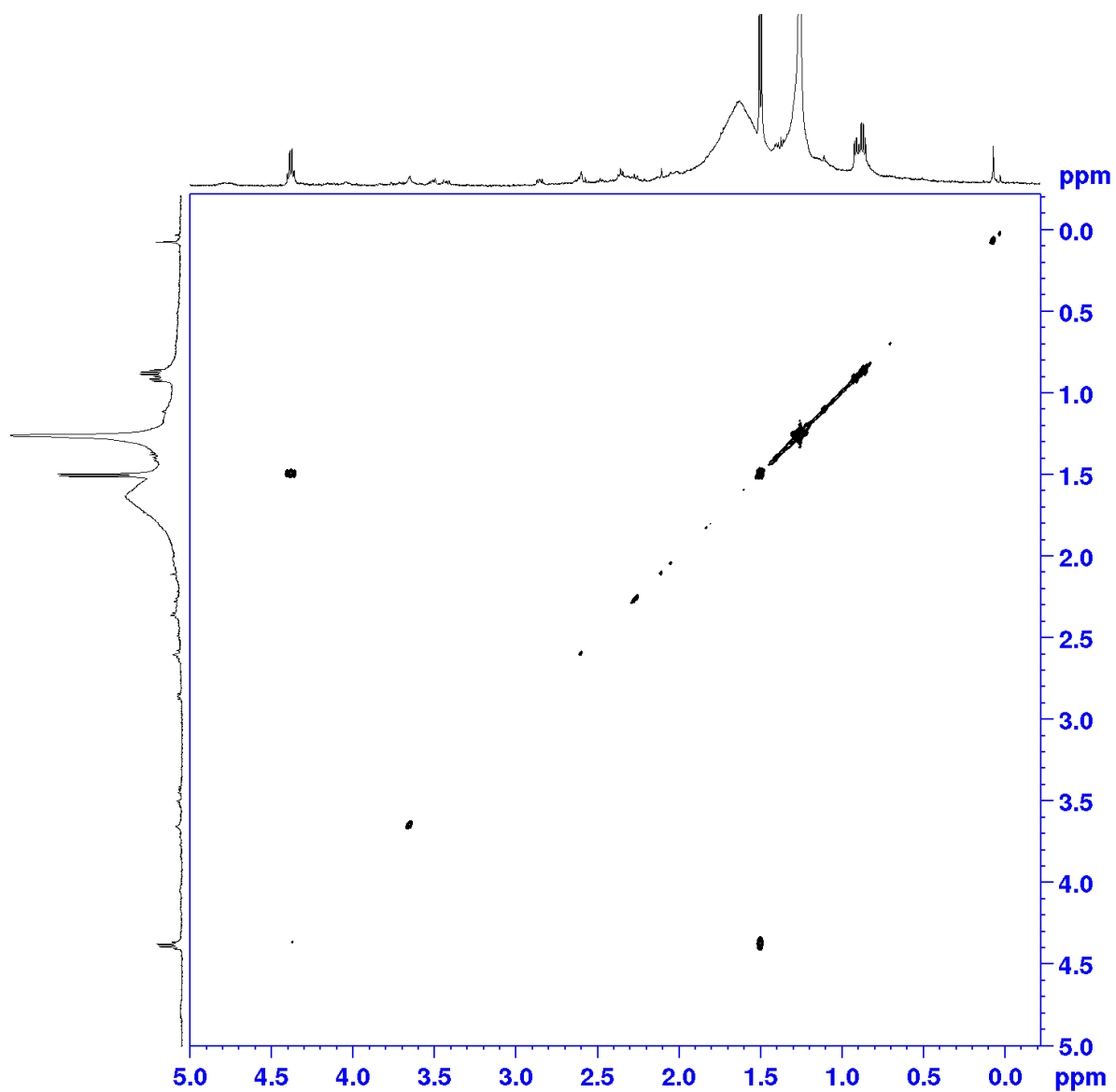
47

48 *Figure S11.* ^1H-^1H -COSY (500 MHz, CDCl$_3$) of loseolamycin A2, zoomed from 1.0 – 4.5 ppm.

49
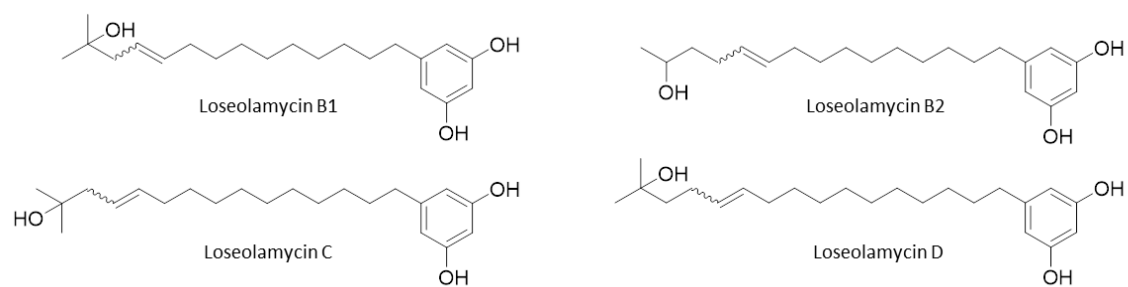
50
51   ***Figure S12.*** *Proposed structures of loseolamycins B1 (compound **3**), B2 (compound **4**), C (compound **5**) and D (compound*
52   ***6****) based on MS/MS experiments.*

53

(a)

(b)
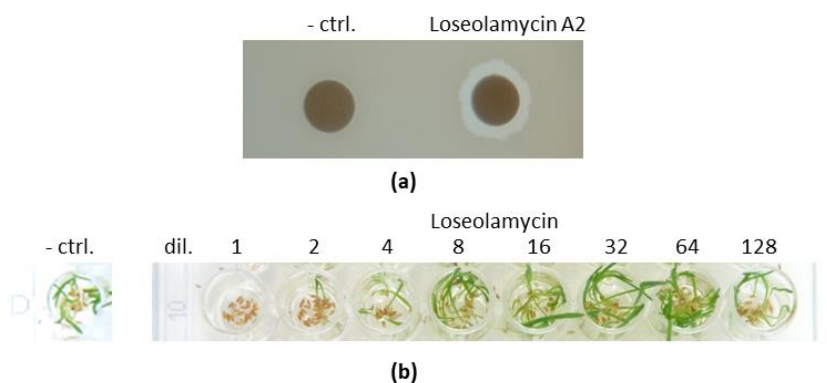
54
55 ***Figure S13.*** *Antibacterial and herbicidal activity. (a) Loseolamycin A2 on a filter disc inhibits the growth of B. subtilis. The*
56 *sample's solvent methanol was used as negative control (- ctrl) and shows no inhibition zone. (b) A mixture of loseolamycin*
57 *derivatives inhibits germination of the weed Agrostis stolonifera. The phytotoxic effect was reproduced once and is*
58 *concentration dependent. Methanol in plant medium was used as negative control (- ctrl) and shows no inhibition of seed*
59 *germination.*

60
61

62

63   *Table S1. Organisms, BACs, plasmids and primer used in this work.*

Material	Purpose
A. Organisms	
Micromonospora endolithica LU17765	originating strain of type III PKS [BASF]
Streptomyces albus Del14	optimized heterologous host [1]
Escherichia coli GB05 RedCC	cloning host [Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS)]
Escherichia coli ET12567 pUB307	alternate host intergeneric conjugation [2]
Escherichia coli GB 2005	bioactivity test
Pseudomonas putida KT2440	bioactivity test
Bacillus subtilis ATCC 6633	bioactivity test
Agrostis stolonifera	bioactivity test
B. BACs	
I7 [IG652BAC1-2]	heterologous expression of type III PKS [GenBank: MT904273]
I7act	heterologous expression of promoter activated type III PKS
C. Plasmids	
pUC19	promoter TS61 / ampicillin resistance marker
D. PCR primer	
20180710_02_fw [I7act]	TGAATCAGATTTGCGAGTCCCGCAGTCGCGAACGGACCGGACTCGTTGGTCGTCAGGTGGCACTTTTCG
20180710_02_rev [I7act]	CACCGGCACGCCCATGTCCCCCACCTCTCGTCCCCGATCCCCCACGCTTCGCGGATATCCTACTATGCCGAGGTATAATGTAGCCAGCGTGTTACCAATGCTTAATCAGTG

64

65

REFERENCES

[1]   M. Myronovskyi, B. Rosenkränzer, S. Nadmid, P. Pujic, P. Normand, A. Luzhetskyy, Generation of a cluster-free Streptomyces albus chassis strains for improved heterologous expression of secondary metabolite clusters, Metab. Eng. 49 (2018) 316–324. https://doi.org/10.1016/j.ymben.2018.09.004.

[2]   F. Flett, V. Mersinias, C.P. Smith, High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl DNA-restricting streptomycetes, FEMS Microbiol. Lett. 155 (2006) 223–229. https://doi.org/10.1111/j.1574-6968.1997.tb13882.x.

74

II.

Dudomycins:

New Secondary Metabolites Produced after Heterologous Expression of an NRPS Cluster from Streptomyces albus ssp. chlorinus NRRL B-24108

Constanze Lasch, Marc Stierhof, Marta Rodríguez Estévez, Maksym Myronovskyi,

Josef Zapp and Andriy Luzhetskyy

Dudomycins: New Secondary Metabolites Produced after Heterologous Expression of an Nrps Cluster from *Streptomyces albus* ssp. *Chlorinus* Nrrl B-24108

Constanze Lasch [1], Marc Stierhof [1] , Marta Rodríguez Estévez [1], Maksym Myronovskyi [1], Josef Zapp [2] and Andriy Luzhetskyy [1,3,*]

[1] Department of Pharmaceutical Biotechnology, Saarland University, 66123 Saarbruecken, Germany;
    constanze.lasch@uni-saarland.de (C.L.); m.stierhof@t-online.de (M.S.);
    marta.rodriguezestevez@uni-saarland.de (M.R.E.); maksym.myronovskyi@uni-saarland.de (M.M.)
[2] Department of Pharmaceutical Biology, Saarland University, 66123 Saarbruecken, Germany;
    j.zapp@mx.uni-saarland.de
[3] AMEG Department, Helmholtz Institute for Pharmaceutical Research Saarland,
    66123 Saarbruecken, Germany
[*] Correspondence: a.luzhetskyy@mx.uni-saarland.de; Tel.: +49-681-302-70200

Abstract: Since the 1950s, natural products of bacterial origin were systematically developed to be used as drugs with a wide range of medical applications. The available treatment options for many diseases are still not satisfying, wherefore, the discovery of new structures has not lost any of its importance. Beyond the great variety of already isolated and characterized metabolites, Streptomycetes still harbor uninvestigated gene clusters whose products can be accessed using heterologous expression in host organisms. This works presents the discovery of a set of structurally novel secondary metabolites, dudomycins A to D, through the expression of a cryptic NRPS cluster from *Streptomyces albus* ssp. *Chlorinus* NRRL B-24108 in the heterologous host strain *Streptomyces albus* Del14. A minimal set of genes, required for the production of dudomycins, was defined through gene inactivation experiments. This paper also proposes a model for dudomycin biosynthesis.

Keywords: Streptomyces; NRPS; heterologous expression

1. Introduction

Biologically active natural products of microbial origin are the result of natural design and evolutionary optimization to target essential biological processes, and are, therefore, a valuable source of potential drug leads [1–3]. Since the 1950s, the bacterial genus of Streptomyces largely contributed to the pool of diverse structural scaffolds, some of which were developed as successful drugs, e.g., vancomycin (antibacterial) [4], avermectin (antiparasitic) [5,6] and actinomycin D (anticancer) [7,8], etc. As a result of extensive screening, the discovery of structurally novel compounds with new biological targets has become a challenging task nowadays. However, only a small part of the biosynthetic gene clusters encoded in microbial genomes is readily expressed in the laboratory, while the rest remain silent. It seems that under standard conditions, only a confined number of clusters leads to the production of natural products, which are then often rediscovered. Identification of unique biosynthetic pathways within genome sequence data and their targeted expression in optimized chassis strains is regarded as a most promising approach to access new biologically active scaffolds [9,10].

Recently we have reported the first studies on the genome mining of the strain *Streptomyces albus* ssp. *Chlorinus* NRRL B-24108. Heterologous expression of *S. albus* ssp. *Chlorinus* genes enabled identifying the biosynthetic gene clusters of the antibiotic nybomycin and the herbicide albucidin, as well as the

isolation and characterization of the novel compounds benzanthric acid and fredericamycin C2 [11–14]. This study reports on a biosynthetic gene cluster of *Streptomyces albus* ssp. *Chlorinus* encoding an uncharacterized nonribosomal peptide synthetase. The bioinformatic analysis did not reveal any characterized homologs of the studied cluster, implying that it might encode a novel natural product. Heterologous expression of the NRPS cluster in our optimized hosts *Streptomyces albus* Del14 [15] and *Streptomyces lividans* Del8 [16] led to identifying a set of four new compounds we named dudomycins. The compounds were purified, and their structures were elucidated in ^{1}H and ^{13}C nuclear magnetic resonance (NMR) experiments. The identified compounds are structurally related and consist of a core lysine and three branched-chain hydroxy fatty acid residues. A hypothesis on dudomycin biosynthesis is proposed based on the results of gene deletion experiments.

2. Materials and Methods

2.1. General Experimental Procedures

The strains, bacterial artificial chromosomes (BACs), and plasmids used in this work are listed in Table S3. *Escherichia coli* strains were cultured in lysogeny broth (LB) medium [17]. Streptomyces strains were grown on soy flour mannitol (MS) agar [18] and in liquid tryptic soy broth (TSB; Sigma-Aldrich, St. Louis, MO, USA). Liquid DNPM medium (40 g/L dextrin, 7.5 g/L soytone, 5 g/L baking yeast, and 21 g/L MOPS, pH 6.8 as aqueous solution) and defined medium DM (mannitol 5 g/L, amino acid 0.5 g/L, K2PO4 0.5 g/L, MgSO4 × 7 H2O 0.2 g/L, FeSO4 × 7 H2O 0.01 g/L) were used for metabolite expression. When DM was used for production, the cells of the preculture were washed three times using amino acid-free defined medium prior to inoculation. Amino acids L-val, L-ile, L-lys, and D-lys were supplied to a defined medium as needed. The antibiotics kanamycin, apramycin, hygromycin, ampicillin, and nalidixic acid were added when required.

2.2. Isolation and Manipulation of DNA

DNA manipulation, transformation into *E. coli*, as well as intergeneric conjugation between *E. coli* and Streptomyces, were performed according to standard protocols [17–19]. BAC DNA from a constructed genomic library of *Streptomyces albus* ssp. *Chlorinus* NRRL B-24108 was isolated with the BACMAX™ DNA purification kit (Lucigen, Middleton, WI, USA). Deletion of several genes was performed on the BAC 4E8 itself using a two-step approach. Step one focused on deleting all genes 1 to 21, leading to BAC 4E8_del1, step two addressed further deletion of genes 25 and 26 on BAC 4E8_del1, resulting in BAC 4E8_del2. The genes were replaced by resistance markers ampicillin and hygromycin through homologous recombination using the Red/ET system [20]. PCR was performed for amplification of the respective gene cassettes from plasmids pUC19 and pXCM hygformax. PCR primers 20190429_1_fw, 20190429_1_rev, 20190429_2_fw, and 20190429_2_rev were constructed with overhang regions for site-specific introduction of the cassettes left or right from the dudomycin cluster and simultaneous removal of the genes 1 to 21 or 25 and 26. Restriction mapping and sequencing were used to control the success of the recombination. Restriction enzymes from ThermoFisher Scientific (Waltham, MA, USA) or New England BioLabs NEB (Ipswich, MA, USA) were used according to the manufacturer's instruction.

2.3. Metabolite Extraction

For metabolite extraction, Streptomyces strains were grown in 15 mL of TSB in a 100 mL baffled flask for 1 to 2 days, and 1 mL of seed culture was used to inoculate 100 mL of production medium in a 500 mL baffled flask. Cultures were grown for 7 days at 28 °C and 180 rpm in an Infors multitron shaker (Infors AG, Basel, Switzerland). Metabolites were extracted from the culture supernatant with an equal amount of either *n*-butanol or ethyl acetate, evaporated at 40 °C and kept at storage condition 4 °C.

2.4. Mass Spectrometry (MS) Metabolite Analysis

Dried extracts were dissolved in methanol prior to the mass spectrometry (MS) analysis. MS experiments were carried out on a Dionex Ultimate 3000 UPLC system (ThermoFisher Scientific, Waltham, MA, USA) coupled to PDA detector (stationary phase 30 or 100 mm ACQUITY UPLC BEH C18 1.7 µm column (Waters Corporation, Milford, MA, USA), mobile phase: Linear gradient of [A] ddH_2O + 0.1% formic acid/[B] acetonitrile + 0.1% formic acid, 5% to 95% at flow rate of 0.6 mL/min). Further mass detection was performed coupling either an amaZon speed (Bruker, Billerica, MA, USA) or LTQ Orbitrap XL mass spectrometer (ThermoFisher Scientific, Waltham, MA, USA) using positive ionization mode and mass range detection of m/z 200 to 2000. Data analysis was performed using software Compass Data Analysis v. 4.1 (Bruker) and Xcalibur v. 3.0 (ThermoFisher Scientific).

2.5. Purification

The extract from the 10 L culture was dissolved in methanol. A first purification step was carried out using Size Exclusion Chromatography (SEC; stationary phase: Sephadex-LH20; mobile phase: isocratic elution using 100% methanol). Fractions containing dudomycins were pooled, dried, redissolved in methanol and undergone a second purification step: Reversed Phase (RP) HPLC (Agilent Infinity 1200 series HPLC system; stationary phase: Synergi™ 4 µm Fusion-RP 80 Å 250 × 10 (Phenomenex, Torrance, CA, USA); mobile phase: Linear gradient of [A] H_2O + 0.1% formic acid/[B] acetonitrile + 0.1% formic acid, 30% to 95% [B] in 14.5 min at flow rate of 4 mL/min, column oven temperature 45 °C, detection UV 210 nm followed by fraction control on HPLC-MS). Fractions were pooled to obtain the four pure dudomycin isolates A to D.

2.6. Nuclear Magnetic Resonance (NMR) Spectroscopy

The ^{1}H-NMR spectrum in $CDCl_3$ (Deutero, Kastellaun, Germany) was recorded on a Bruker Avance 500 spectrometer (Bruker, BioSpin GmbH, Rheinstetten, Germany) equipped with a 5 mm BBO probe at 298 K. The chemical shifts were reported in parts per million (ppm) relative to TMS. The spectra were recorded with the standard ^{1}H pulse program using 64 scans. All other NMR spectra were acquired on a Bruker Ascend 700 MHz NMR spectrometer at 298 K equipped with a 5 mm TXI cryoprobe. As a solvent, deuterated CD_3OD was used. HSQC, HMBC, 1H-1H COSY spectra were recorded using the standard pulse programs from the TOPSPIN v. 3.6 software. Selective 1D TOCSY experiments were performed using mixing times of 120 ms.

2.7. Genome Mining and Bioinformatic Analysis

The genome of *S. albus* ssp. *Chlorinus* was screened for secondary metabolite biosynthetic gene clusters using the antiSMASH online tool (https://antismash.secondarymetabolites.org/#!/start) [21]. Analysis of genetic data was performed using Geneious software, v. 11.0.3 [22]. The genomic sequence of *Streptomyces albus* ssp. *Chlorinus* was deposited in GenBank under accession number VJOK00000000. For dereplication, the Dictionary of Natural Products (DNP) 28.1 was used as a database of known natural products.

3. Results and Discussion

3.1. Identification and Expression of the NRPS Gene Cluster

Genome mining of *Streptomyces albus* ssp. *Chlorinus* NRRL B-24108 using AntiSMASH software revealed several cryptic biosynthetic gene clusters (BGCs) within its chromosome [22]. A BGC encoding a putative nonribosomal peptide synthetase (NRPS) caught our attention as the software did not detect any homology to already characterized BGCs [23]. A BAC 4E8 covering the entire NRPS cluster was identified in the previously constructed genomic library of *S. albus* ssp. *Chlorinus* (GenBank accession number VJOK00000000). The BAC 4E8 was transferred into the optimized heterologous host strains

Streptomyces albus Del14 and *Streptomyces lividans* Del8, leading to the respective exconjugant strains *Streptomyces albus* 4E8 and *Streptomyces lividans* 4E8. The obtained exconjugant strains were cultivated in the production medium DNPM, and the metabolites were extracted from the culture supernatant with ethyl acetate or *N*-butanol. High-resolution HPLC-MS analysis revealed the presence of four new peaks in the ethyl acetate and *N*-butanol extracts of both *S. albus* and *S. lividans* harboring the BAC 4E8 (Figure 1; Figures S1 and S2). The identified peaks could not be observed in the extracts of the control strains without the BAC. Analysis of the mass spectra of the identified peaks revealed the molecular ions [M + H$^+$] with the masses 573.411, 587.426, 601.442, and 615.457 Da. A search in the DNP database of natural products for the identified high-resolution masses did not generate any matches, implying that the identified compounds might be new. The mass differences of 14 Da between the individual compounds imply that they might differ by the presence of additional CH$_2$ groups. The identified compounds were named dudomycins A, B, C, and D.
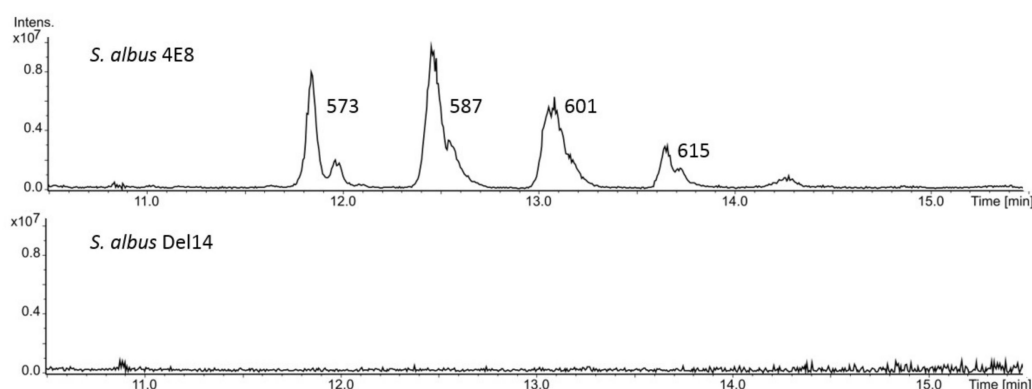
Figure 1. HPLC-MS analysis of dudomycin production by *S. albus* Del14 strain harboring the BAC (bacterial artificial chromosomes) 4E8. The strain *S. albus* Del14 without the BAC was used as a control. The extracted base peak chromatograms 573-574 Da, 587-588 Da, 601-602 Da, 615-616 Da are shown.

3.2. Purification and Structure Elucidation

To get insights into the structures of the produced compounds, the producer strain *S. albus* 4E8 was cultivated in 10 L of the production medium DNPM, and the metabolites were extracted from the culture supernatant with ethyl acetate. The strain *S. albus* 4E8 was preferred to *S. lividans* 4E8, due to its higher production level. The dudomycins A to D corresponding to the identified molecular ions [M + H$^+$] with the masses 573.411, 587.426, 601.442, and 615.457 Da were successfully purified from the extract: 1.2 mg of dudomycin A, 1.1 mg of dudomycin B, 0.7 mg of dudomycin C and 0.5 mg of dudomycin D were obtained. The structure elucidation of the isolated dudomycins was performed using 1D and 2D NMR with a special focus on 1D TOCSY experiments.

Due to its small molecular mass, dudomycin A was the first compound used for the structure elucidation. The molecular formula was calculated as C$_{30}$H$_{56}$O$_8$N$_2$ based on the calculated high-resolution mass of 572.403 Da, indicating four degrees of unsaturation. Analysis of ^{1}H, ^{13}C NMR, and edited HSQC led to 6 methyls, 13 methylenes, 7 methines, and 4 quaternary carbons. Two of the remaining five protons belonged to NH-groups, as the ^{1}H NMR measurement in CDCl$_3$ revealed the presence of two signals at δH 6.81 and δH 7.12 ppm (Figure S3), which disappeared in protic solvents. Due to broad signals and the compound's poor solubility in CDCl$_3$, all other spectra were recorded in CD$_3$OD.

Interpretation of HHCOSY revealed four discrete spin systems in the molecule. The first sequence starting from the methine signal at δH 4.26 followed by four methylenes at δH 1.83, 1.38, 1.51, and 3.17 was assigned to lysine, which was supported by HMBC data. The remaining three spin systems were mostly close to each other, leading to many overlaps of methylene and methyl resonances in the area of δH 0.8–1.7 (Figure S4). However, well-separated methine proton signals at δH 5.20,

3.95, 3.92 enabled selective 1D TOCSY experiments using these resonances as irradiation points. Their careful interpretation enabled the full assignment of all spin systems leading to three different 3-hydroxy-6-methyl heptanoic acid moieties (HMH1 to HMH3).

Long-range HMBC correlations of Lys-H-2 (δH 4.26, δC 56.2) to HMH1-C-1' (δC 171.9) and Lys-H-6 (δH 3.17, δC 40.4) to C-1' (δC 174.3) proved connections of HMH1 and HMH2 via amide bonds. The carbonyl group of lysine (δC 178.9) did not show any external connections, hence it was assigned as a free carboxyl group. The remaining long-range HMBC correlations between H-3' (δH 5.20, δC 73.3) and C-1''' (δC 173.2) indicated an ester bond between HMH1 and HMH3 leading to the final structure of dudomycin A (Figure 2; Figures S3–S10; Table S1).
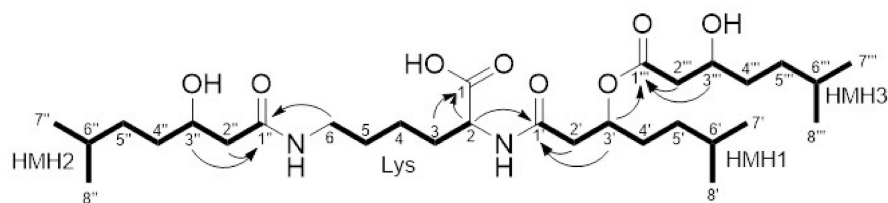
Figure 2. Observed HMBC (arrows) and selective 1D TOCSY (bold lines) key correlations of dudomycin A.

The results of NMR analysis indicate that the dudomycins B ($C_{31}H_{58}O_8N_2$), C ($C_{32}H_{60}O_8N_2$), and D ($C_{33}H_{62}O_8N_2$) are structurally related to dudomycin A. From these compounds, only dudomycin D (Table S2) was a pure substance, while dudomycins B and C were the mixtures of several isomers. Similar to dudomycin A, the structure of dudomycin D contains a lysine core, with three hydroxy fatty acids attached. However, in contrast to dudomycin, A three residues of 3-hydroxy-6-methyl octanoic acid (HMO1 to HMO3) can be found in the structure of dudomycin D instead of three residues of 3-hydroxy-6-methyl heptanoic acid (HMH1 to HMH3). The alkyl chain of the 3-hydroxy-6-methyl octanoic acid is extended by one additional CH_2 group compared to 3-hydroxy-6-methyl heptanoic acid what explains the overall mass difference of 42 Da between dudomycins A and D.

NMR spectra of dudomycins B and C revealed that both samples consist of three isomers each. Similar to dudomycins A and D, dudomycins B and C also contain a lysine core in their structures. In contrast to dudomycins A and D, which contain either HMH or HMO residues bound to the core, dudomycins B and C contain a mixture of HMH and HMO residues in their structures. The results of ^{1}H NMR and a multitude of 1D TOCSY measurements indicate that dudomycin B contains two HMH and one HMO residues, while dudomycin C contains one HMH and two HMO residues. These results are in accordance with the observed mass difference of 14 Da between Dudomycin A and B and of 28 Da between dudomycins A and C. Three isomeric forms are possible for both dudomycin B and C (Figure 3, Figures S11–S24).
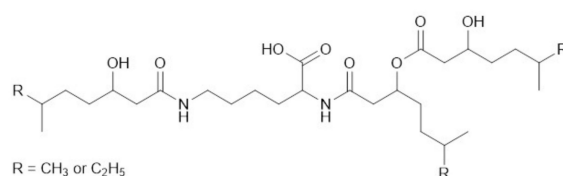
Figure 3. The structures of isolated dudomycins. In the case of dudomycin A, all three R groups correspond to the CH_3 group. Dudomycin B is a mixture of 3 constitutional isomers with two R groups corresponding to CH_3 and one R group—to C_2H_5. Dudomycin C is a mixture of 3 constitutional isomers with one R group corresponding to CH_3 and two R groups—to C_2H_5. In dudomycin D, all three R groups correspond to C_2H_5.

Experiments to determine the absolute configuration of dudomycins have not been performed, due to low amounts of the isolated compounds. The results of a feeding experiment imply that L-lysine

is used as a precursor for dudomycin biosynthesis (Figure S25). A higher dudomycin production level was observed upon the cultivation of *S. albus* 4E8 in a defined medium with ʟ-lysine as a nitrogen source than in the medium with ᴅ-lysine.

The occurrence of HMH and HMO in natural products is very rare, and to the best of our knowledge, the structures of dudomycin A to D have not been reported before.

3.3. Determination of the Minimal Dudomycin Gene Cluster

The 30 kb DNA fragment cloned in the BAC 4E8 contains 26 genes (Figure 4; Table 1). Gene 22 encodes a putative NRPS comprising condensation (C), adenylation (A), peptidyl carrier protein (PCP), and thioesterase (TE) domains. The A domain was predicted to have a weak preference for recognition of the amino acid ornithine. Due to the structural similarity of ornithine and lysine gene 22 encoding, an NRPS enzyme was regarded to be involved in dudomycin biosynthesis. A sequence analysis of the regions upstream and downstream of gene 22 was performed to identify the genes possibly involved in the production of dudomycins. This analysis did not reveal any gene whose product could be enzymatically involved in dudomycin biosynthesis. Sequence homology analysis revealed that the homologs of genes 22, 23, and 24 are clustered together in the genomes of seven different strains implying that those genes might be involved in the same pathway. Genes 23 and 24 encode a putative transport protein and a putative transcriptional regulator, respectively. To confirm that the genes upstream, gene 22 are not involved in dudomycin biosynthesis, the genes 1 to 21 were substituted in the BAC 4E8 with an ampicillin cassette using Red/ET. The constructed BAC 4E8_del1 was transferred into *S. albus* Del14 strain yielding *S. albus* 4E8_del1. HPLC-MS analysis of the metabolite production by the obtained strains did not detect any differences in dudomycin production compared to the *S. albus* 4E8 strain. This clearly indicated that the genes 1 to 21 encoded in the BAC 4E8 are not involved in the biosynthesis of dudomycins.

To find out if genes 25 and 26 are involved in dudomycin production, they were replaced in the BAC 4E8_del1 with a hygromycin resistance cassette using Red/ET. The constructed BAC 4E8_del2 contains genes 22, 23, and 24 only. The BAC was transferred into *S. albus* Del14, and the obtained strain *S. albus* 4E8_del2 was checked for dudomycin production. HPLC-MS analysis revealed that deleting genes 25 and 26 did not affect the production of dudomycins (Figure S26), indicating that these genes do not take part in the biosynthesis of the compounds.

The results of gene deletion experiments demonstrate that genes 22, 23, and 24 (designated as *dudA*, *dudB*, and *dudC*), encoding the putative NRPS, transport protein, and transcriptional regulator (Table 1, Figure 4), suffice for the production of dudomycins. Since the products of *dudB* and *dudC* do not have an enzymatic function, we suppose that only the product of *dudA* is responsible for the biosynthesis of dudomycins. The products of *dudB* and *dudC* are likely involved in the transport of the biosynthetic products and in the regulation of dudomycin production, respectively. Since the inactivation of the genes *dudB* and *dudC* was not performed, the possibility that the genes are not essential for dudomycin production cannot be completely excluded.
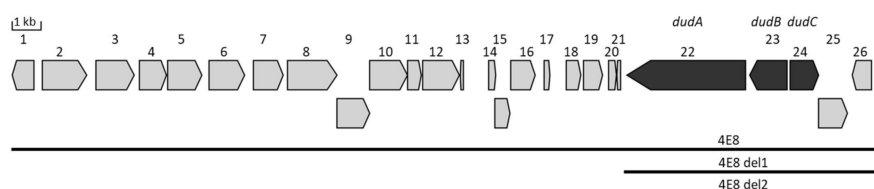
Figure 4. The chromosomal fragment of *Streptomyces albus* ssp. *Chlorinus* NRRL B-24108 containing the dudomycin gene cluster. The genes putatively involved in dudomycin biosynthesis are highlighted in dark grey. The chromosomal fragments cloned in BACs 4E8, 4E8_del1 and 4E8_del2 are shown with black bars.

Table 1. Proposed functions of the genes in the DNA fragment containing the dudomycin gene cluster.

Gene #	Locus Tag	Putative Function
1	SACHL_42600	glycosyltransferase
2	SACHL_42610	hypothetical protein
3	SACHL_42620	phosphatase
4	SACHL_42630	*N,N'*-diacetyllegionaminic acid synthase
5	SACHL_42640	hypothetical protein
6	SACHL_42650	hydrolase
7	SACHL_42660	membrane lipoprotein precursor
8	SACHL_42670	galactose/methyl galactoside import ATP-binding protein
9	SACHL_42680	ribose transport system permease protein
10	SACHL_42690	branched-chain amino acid transport system/permease component
11	SACHL_42700	cytidine deaminase
12	SACHL_42710	pyrimidine-nucleoside phosphorylase
13	SACHL_42720	hypothetical protein
14	SACHL_42730	hypothetical protein
15	SACHL_42740	hypothetical protein
16	SACHL_42750	hypothetical protein
17	SACHL_42760	hypothetical protein
18	SACHL_42770	hypothetical protein
19	SACHL_42780	ubiquinone biosynthesis *O*-methyltransferase
20	SACHL_42790	zinc metallo-peptidase
21	SACHL_42800	hypothetical protein
22 [*dudA*]	SACHL_42810	dimodular nonribosomal peptide synthase
23 [*dudB*]	SACHL_42820	inner membrane transport protein
24 [*dudC*]	SACHL_42830	transcriptional regulator
25	SACHL_42840	demethylrebeccamycin-D-glucose *O*-methyltransferase
26	SACHL_42850	hypothetical protein

3.4. Biosynthesis of Dudomycins

Structurally dudomycins consist of a lysine core and three hydroxy fatty acid residues, two of which are attached directly to the lysine moiety through amide bonds, while the third residue is esterified with one of the lysine bound hydroxy fatty acids (Figure 3). Two types of hydroxy fatty acids are used for the biosynthesis of dudomycins: The shorter HMH and the longer HMO. Three HMH residues are found in the smallest compound—dudomycin A, while three HMO residues are used to form the largest compound—dudomycin D. Dudomycins B and C contain both types of hydroxy fatty acids in their structures. The results of gene inactivation studies and sequence analysis demonstrated that only the product of the *dudA* gene is responsible for biosynthesis of dudomycins. The gene encodes an NRPS comprising four domains only: C, A, PCP, and TE. We propose that the A domain activates the L-lysine and loads it on the PCP module. The C domain then attaches two hydroxy fatty acid moieties to the two available amino groups of PCP-bound lysine. The product is released through the hydrolytic activity of the TE domain. We propose that the third hydroxy fatty acid is spontaneously esterified with HMH1 or HMO1, respectively (Figures 2 and 5). It remains unclear if the attachment of the third hydroxy fatty acid occurs before or after the release of the NRPS product. No peaks corresponding to the dudomycin derivatives without HMH3 or HMO3 residue could be identified by HPLC-MS analysis in the extracts of Streptomyces strains harboring the dudomycin biosynthetic cluster.
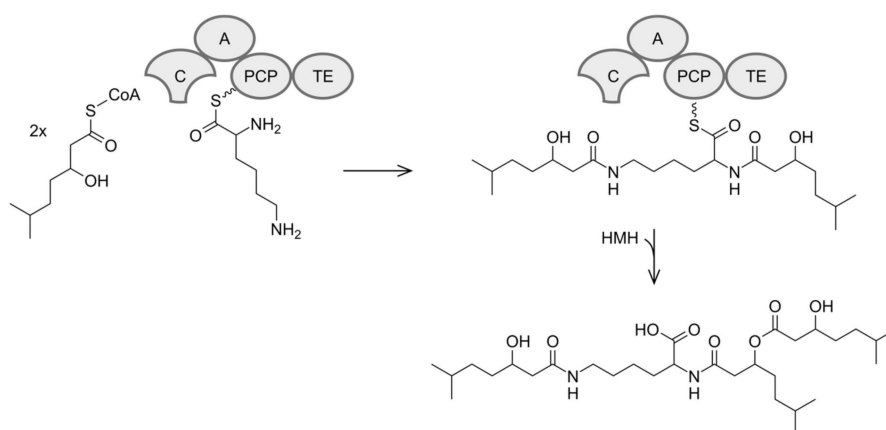
Figure 5. Proposed biosynthesis exemplarily for dudomycin A.

Attachment of two hydroxy fatty acid residues to a lysine, which is catalyzed by the single-module NRPS encoded by *dudA* is not typical. At least another two cases are known when a single NRPS module performs two rounds of condensation. During the biosynthesis of myxochelins a single-module NRPS encoded by the *mxcG* gene creates the amide linkage between two 2,3-dihydroxybenzoic acid residues and the two amino groups of lysine [24]. During the biosynthesis of vibriobactin the single-module NRPS encoded by the *vibF* gene also catalyzes two condensation events. However, in the case of vibriobactin, two condensation domains are present within *VibF* each of which is likely responsible for one condensation reaction [25].

Three genes, *dudA* to *dudC*, are required for the biosynthesis of dudomycins in *S. albus* Del14 and *S. lividans* Del8. The *dudA* gene encodes the above mentioned NRPS, and genes, *dudB* and *dudC*, encode the putative membrane transporter and transcriptional regulator. No genes involved in precursor supply could be identified in the DNA regions flanking genes, *dudA* to *dudC*, implying that all precursors required for the dudomycin biosynthesis, including the hydroxy fatty acids HMH and HMO are provided by the host metabolism.

Streptomyces are known for their extraordinary prevalence of branched-chain fatty acids, which are synthesized by type II fatty acid synthases [26]. Branched-chain amino acids often serve as precursors for biosynthesis of iso and anteiso carboxylic acids. Through oxidation of the amine group, the amino acids are converted into α-keto acids, which are then used as starter units in bacterial fatty acid biosynthesis [27]. Valine and isoleucine with high probability serve as main biosynthetic precursors for HMH and HMO, respectively. These amino acids are converted to 3-methyl-2-oxobutanoic acid and 3-methyl-2-oxopentanoic acid through the action of valine dehydrogenase or homologous enzymes. The formed 2-keto carboxylic acids are decarboxylated and used as starter units by a type II fatty acid synthase which converts them to HMH and HMO after decarboxylative condensation of two malonyl units (Figure 6). The precursor role of L-valine and L-isoleucine in biosynthesis of HMH and HMO is indirectly confirmed by feeding studies [28]. During cultivation in a defined medium with L-valine as nitrogen source, the strain *S. albus* 4E8_del2 produced mostly the smaller dudomycins (dudomycins A and B), which contain mainly HMH, which is derived from L-valine. With L-isoleucine, as a single nitrogen source, the strain produced mostly the bigger dudomycins C and D. These compounds contain mainly HMO, which is derived from L-isoleucine (Figure S25). Furthermore, heavily impaired production of dudomycins was observed when BAC 4E8 was expressed in an *S. albus* delVDH strain [29] with inactivated valine dehydrogenase gene *vdh* (Figure S27). This result suggests that the valine dehydrogenese is involved in the supply of HMH and HMO by oxidation of the branched-chain amino acids L-valine and L-isoleucine. Low dudomycin production by *S. albus* delVDH 4E8_del2 also indicates that other homologs of the valine dehydrogenase with broad substrate specificity are encoded in the genome of *S. albus*, in accordance with previously published data [30].
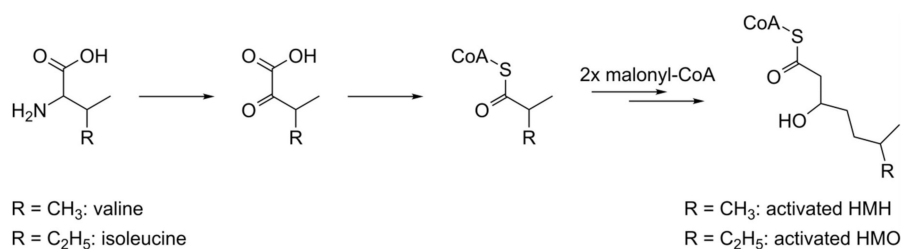
Figure 6. Proposed biosynthesis of HMH and HMO.

4. Conclusions

In this paper, we report the identification and successful heterologous expression of a new NRPS cluster leading to the production of a group of new compounds called dudomycins A to D. The single-module NRPS which is responsible for dudomycin production activates ʟ-lysine and catalyzes its condensation with two hydroxy fatty acid CoA precursors. Such examples where a single condensation domain catalyzes two condensation steps are very rare. The branched hydroxy fatty acids used for the dudomycin biosynthesis are provided by the host's metabolism. The amino acids ʟ-valine and ʟ-isoleucine serve as main precursors for their biosynthesis.

References

1. Beutler, J.A. Natural Products as a Foundation for Drug Discovery. *Curr. Protoc. Pharmacol.* **2009**, *46*. [CrossRef] [PubMed]
2. Shen, B. A New Golden Age of Natural Products Drug Discovery. *Cell* **2015**, *163*, 1297–1300. [CrossRef] [PubMed]
3. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770–803. [CrossRef] [PubMed]
4. Griffith, R.S. Vancomycin use—An historical review. *J. Antimicrob. Chemother.* **1984**, *14*, 1–5. [CrossRef]
5. Hotson, I.K. The avermectins: A new family of antiparasitic agents. *J. S. Afr. Vet. Assoc.* **1982**, *53*, 87–90.
6. Ottesen, E.A.; Campbell, W. Ivermectin in human medicine. *J. Antimicrob. Chemother.* **1994**, *34*, 195–203. [CrossRef]
7. D'Angio, G.J.; Evans, A.; Breslow, N.; Beckwith, B.; Bishop, H.; Farewell, V.; Goodwin, W.; Leape, L.; Palmer, N.; Sinks, L.; et al. The treatment of Wilms' Tumor: Results of the second national Wilms' Tumor study. *Cancer* **1981**, *47*, 2302–2311. [CrossRef]
8. Turan, T.; Karacay, O.; Tulunay, G.; Boran, N.; Koc, S.; Bozok, S.; Kose, M.F. Results with EMA/CO (etoposide, methotrexate, actinomycin D, cyclophosphamide, vincristine) chemotherapy in gestational trophoblastic neoplasia. *Int. J. Gynecol. Cancer.* **2006**, *16*, 1432–1438. [CrossRef]
9. Myronovskyi, M.; Luzhetskyy, A. Heterologous production of small molecules in the optimized Streptomyces hosts. *Nat. Prod. Rep.* **2019**, *36*, 1281–1294. [CrossRef]
10. Baltz, R. Renaissance in antibacterial discovery from actinomycetes. *Curr. Opin. Pharmacol.* **2008**, *8*, 557–563. [CrossRef]
11. Estévez, M.R.; Myronovskyi, M.; Gummerlich, N.; Nadmid, S.; Luzhetskyy, A. Heterologous Expression of the Nybomycin Gene Cluster from the Marine Strain *Streptomyces albus* subsp. *chlorinus* NRRL B-24108. *Mar. Drugs* **2018**, *16*, 435. [CrossRef] [PubMed]
12. Myronovskyi, M.; Rosenkränzer, B.; Stierhof, M.; Petzke, L.; Seiser, T.; Luzhetskyy, A. Identification and Heterologous Expression of the Albucidin Gene Cluster from the Marine Strain *Streptomyces albus* Subsp. *chlorinus* NRRL B-24108. *Microorganisms* **2020**, *8*, 237. [CrossRef] [PubMed]
13. Estévez, M.R.; Gummerlich, N.; Myronovskyi, M.; Zapp, J.; Luzhetskyy, A. Benzanthric Acid, a Novel Metabolite From *Streptomyces albus* Del14 Expressing the Nybomycin Gene Cluster. *Front. Chem.* **2020**, *7*. [CrossRef]
14. Estévez, M.R.; Myronovskyi, M.; Rosenkränzer, B.; Paululat, T.; Petzke, L.; Ristau, J.; Luzhetskyy, A. Novel fredericamycin variant overproduced by a streptomycin-resistant *Streptomyces albus* subsp. *chlorinus* Strain. *Mar. Drugs* **2020**, *18*, 284. [CrossRef] [PubMed]
15. Myronovskyi, M.; Rosenkränzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A. Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab. Eng.* **2018**, *49*, 316–324. [CrossRef] [PubMed]
16. Ahmed, Y.; Rebets, Y.; Estévez, M.R.; Zapp, J.; Myronovskyi, M.; Luzhetskyy, A. Engineering of *Streptomyces lividans* for heterologous expression of secondary metabolite gene clusters. *Microb. Cell Fact.* **2020**, *19*, 5. [CrossRef]
17. Green, J.M.R.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*, 4th ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2012.
18. Kieser, T.; Bibb, M.J.; Buttner, M.J.; Chater, K.F.; Hopwood, D.A. *Practical Streptomyces Genetics. A Laboratory Manual*; John Innes Foundation: Norwich, UK, 2000.
19. Rebets, Y.; Kormanec, J.; Lutzhetskyy, A.; Bernaerts, K.; Anné, J. Cloning and expression of metagenomic dna in *Streptomyces lividans* and subsequent fermentation for optimized production. *Methods Mol. Biol.* **2017**, *1539*, 99–144. [CrossRef]
20. Court, D.L.; Sawitzke, J.A.; Thomason, L.C. Genetic Engineering Using Homologous Recombination. *Annu. Rev. Genet.* **2002**, *36*, 361–388. [CrossRef]
21. Medema, M.H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M.A.; Weber, T.; Takano, E.; Breitling, R. antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **2011**, *39*, W339–W346. [CrossRef]

22. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef]

23. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]

24. Gaitatzis, N.; Kunze, B.; Muller, R. In vitro reconstitution of the myxochelin biosynthetic machinery of *Stigmatella aurantiaca* Sg a15: Biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 11136–11141. [CrossRef] [PubMed]

25. Keating, T.A.; Marshall, C.G.; Walsh, C.T. Reconstitution and Characterization of the *Vibrio cholerae* Vibriobactin Synthetase from VibB, VibE, VibF, and VibH. *Biochemistry* **2000**, *39*, 15522–15530. [CrossRef] [PubMed]

26. Gago, G.; Diacovich, L.; Arabolaza, A.; Tsai, S.-C.; Gramajo, H. Fatty acid biosynthesis in actinomycetes. *FEMS Microbiol. Rev.* **2011**, *35*, 475–497. [CrossRef] [PubMed]

27. Kaneda, T. Iso- and anteiso-fatty acids in bacteria: Biosynthesis, function, and taxonomic significance. *Microbiol. Rev.* **1991**, *55*, 288–302. [CrossRef] [PubMed]

28. Tang, L.; Hutchinson, C.R. Regulation of expression of the valine (branched-chain amino acid) dehydrogenase-encoding gene from *Streptomyces coelicolor*. *Gene* **1995**, *162*, 69–74. [CrossRef]

29. Manderscheid, N. Strain Development for Heterologous Expression of Secondary Metabolite Clusters in Actinobacteria. Ph.D. Thesis, Universität des Saarlandes, Saarbrücken, Germany, 2015.

30. Vancurová, I.; Vancura, A.; Volc, J.; Neuzil, J.; Flieger, M.; Basarová, G.; Běhal, V. Isolation and characterization of valine dehydrogenase from *Streptomyces aureofaciens*. *J. Bacteriol.* **1988**, *170*, 5192–5196. [CrossRef]

Dudomycins: new secondary metabolites produced after heterologous expression of an NRPS cluster from *Streptomyces albus* ssp. *chlorinus* NRRL B-24108

Constanze Lasch [1], Marc Stierhof [1], Marta Rodríguez Estévez [1], Maksym Myronovskyi [1], Josef Zapp [2] and Andriy Luzhetskyy [1,3,*]

[1] Pharmaceutical Biotechnology, Saarland University, Saarbruecken, Germany;
constanze.lasch@uni-saarland.de (C.L.), m.stierhof@t-online.de (M.S.),
marta.rodriguezestevez@uni-saarland.de (M.R.), maksym.myronovskyi@uni-saarland.de (M.M.),
a.luzhetskyy@mx.uni-saarland.de (A.L.)

[2] Pharmaceutical Biology, Saarland University, Saarbruecken, Germany; j.zapp@mx.uni-saarland.de (J.Z.)

[3] Helmholtz Institute for Pharmaceutical Research Saarland, Saarbruecken, Germany

* Correspondence: a.luzhetskyy@mx.uni-saarland.de; +49 681 302 70200 (A.L.)

Supplementary.

Figure S1. Production of dudomycins by different hosts using different organic solvents for extraction. Base Peak Chromatogram (BPC) extracted for masses M+H+ [573-574], [587-588], [601-602], [615-616]. Butanolic extractions led to more background peaks in crude extracts. *S. albus* 4E8 showed higher production levels compared to *S. lividans* 4E8.
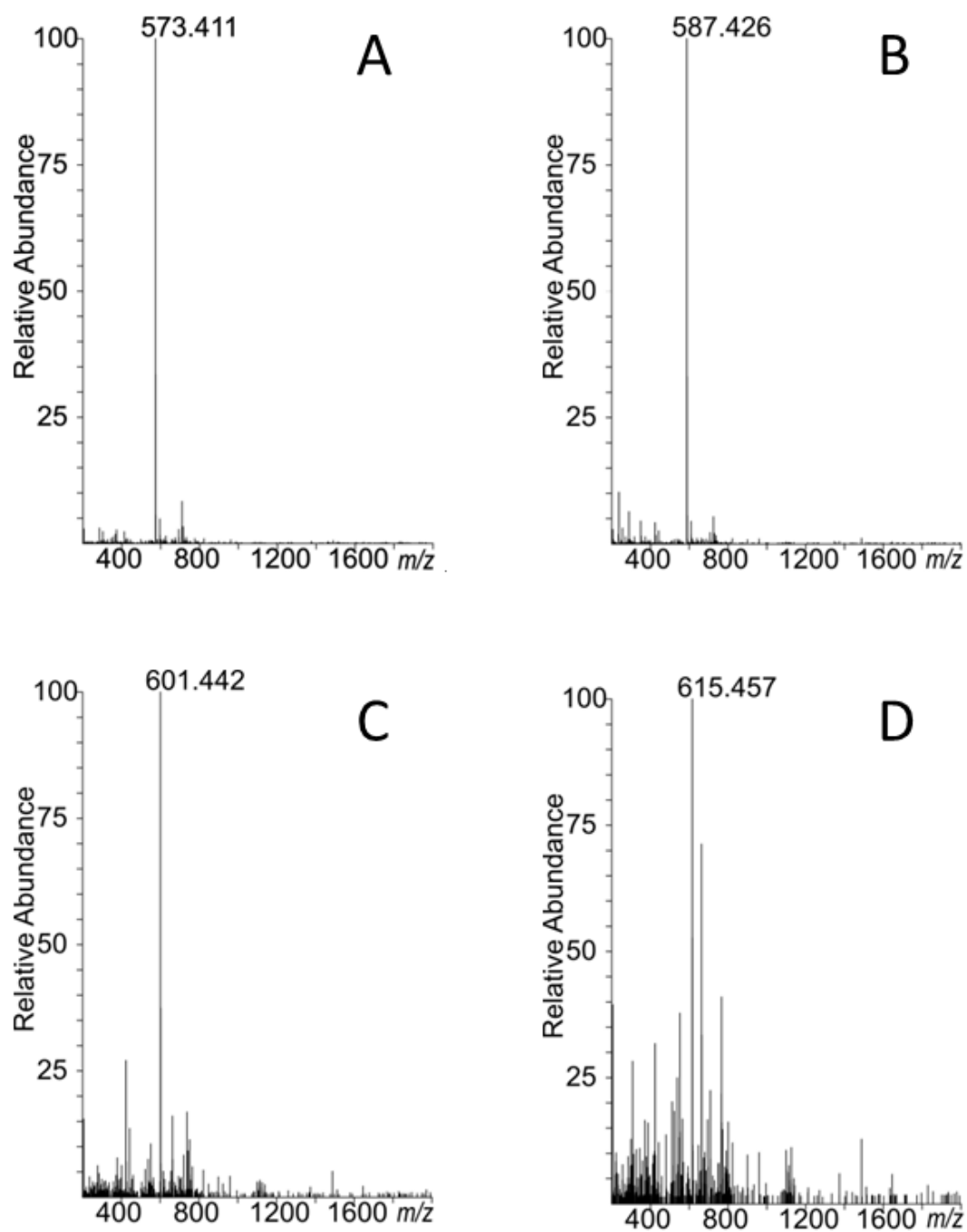
Figure S2. A to D: Mass spectra of dudomycins A to D.
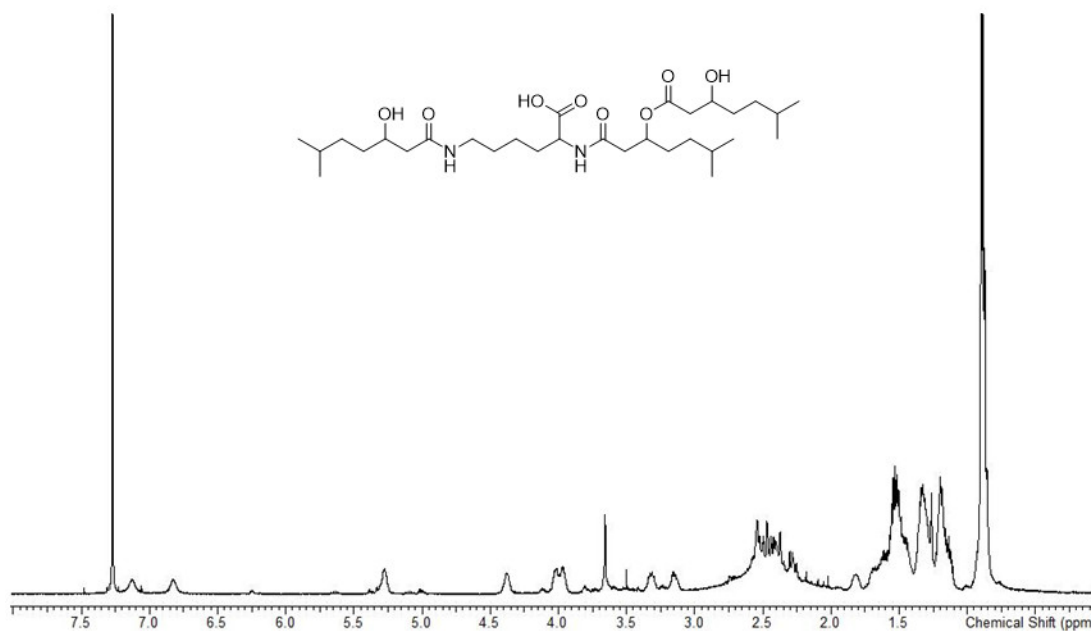
Figure S3. ¹H NMR (500 MHz, CDCl₃) spectrum of dudomycin A.
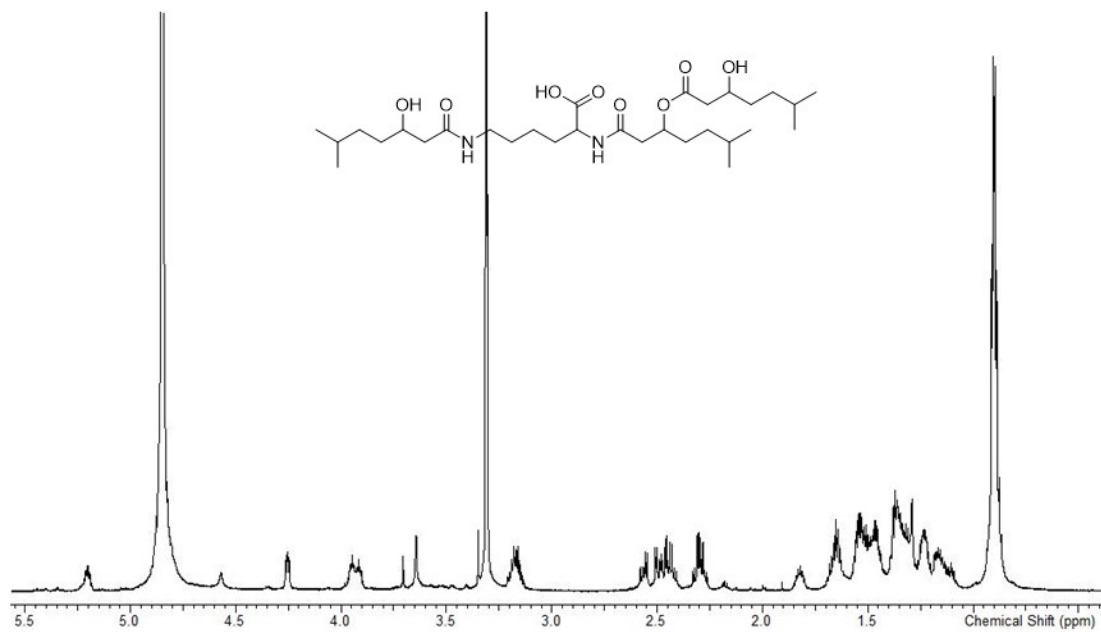
Figure S4. ¹H NMR (500 MHz, CD₃OD) spectrum of dudomycin A.
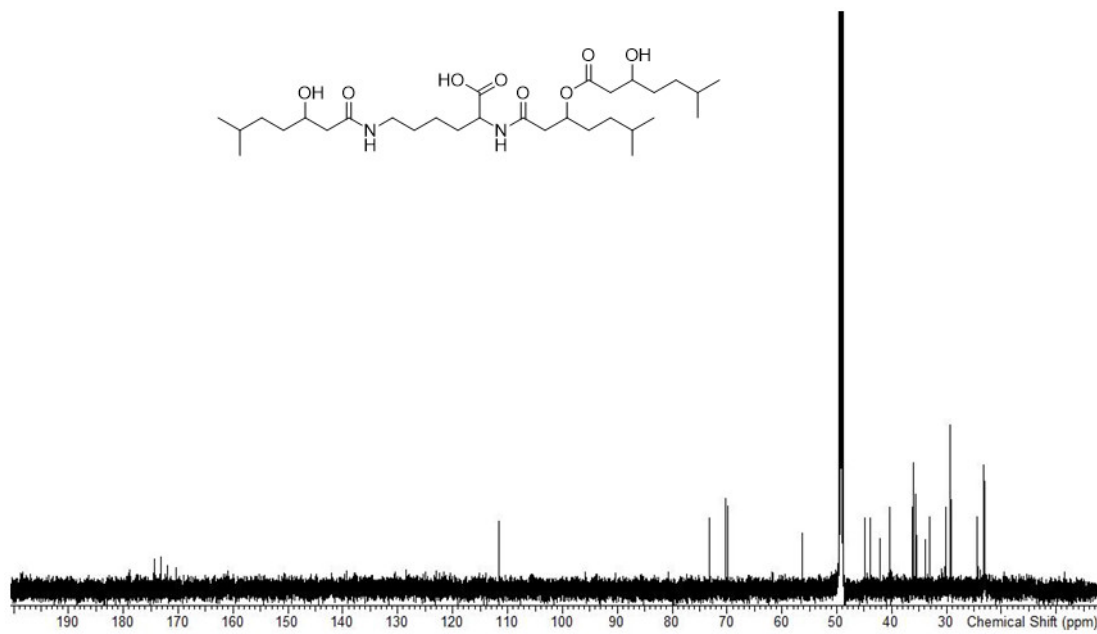
Figure S5. ^{13}C NMR (176 MHz, CD₃OD) spectrum of dudomycin A.
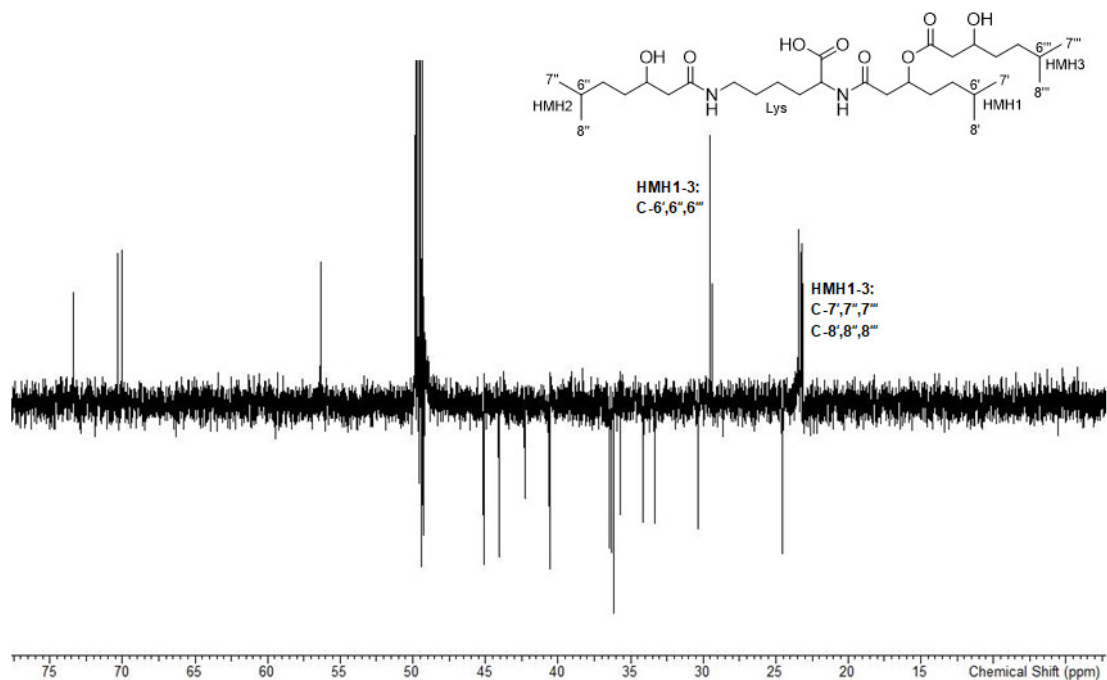
Figure S6. DEPT-135 (176 MHz, CD₃OD) spectrum of dudomycin A.
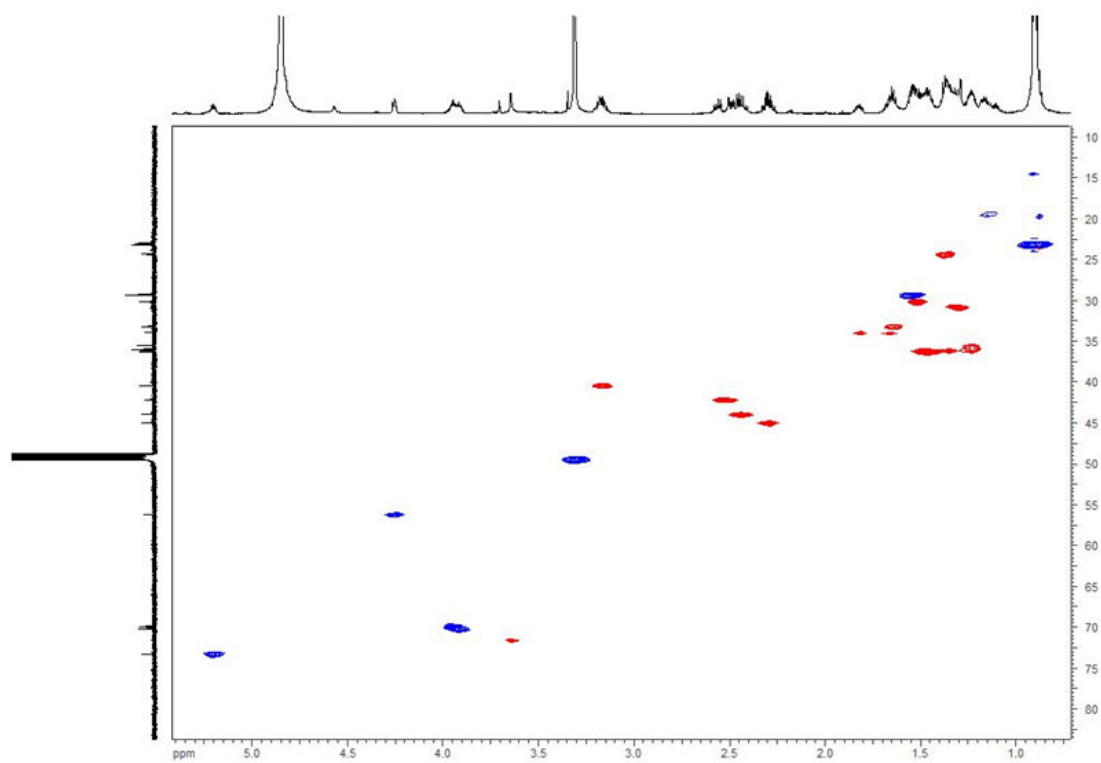
Figure S7. HSQC (700 MHz, CD₃OD) spectrum of dudomycin A.
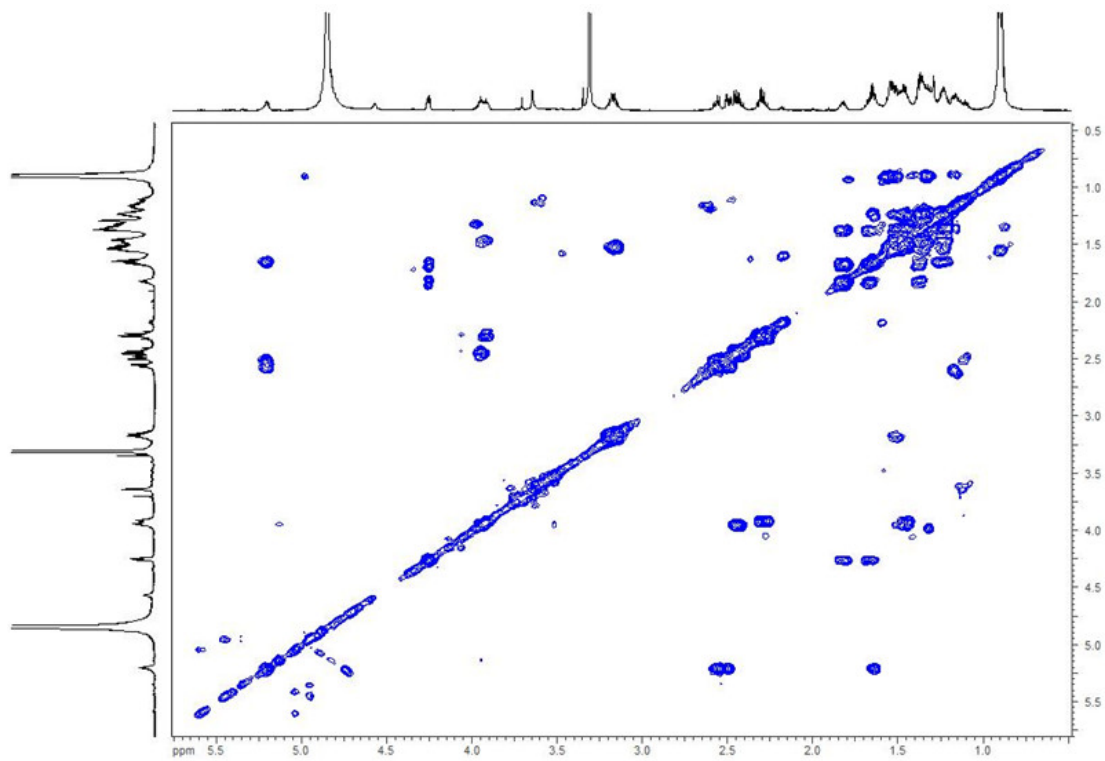
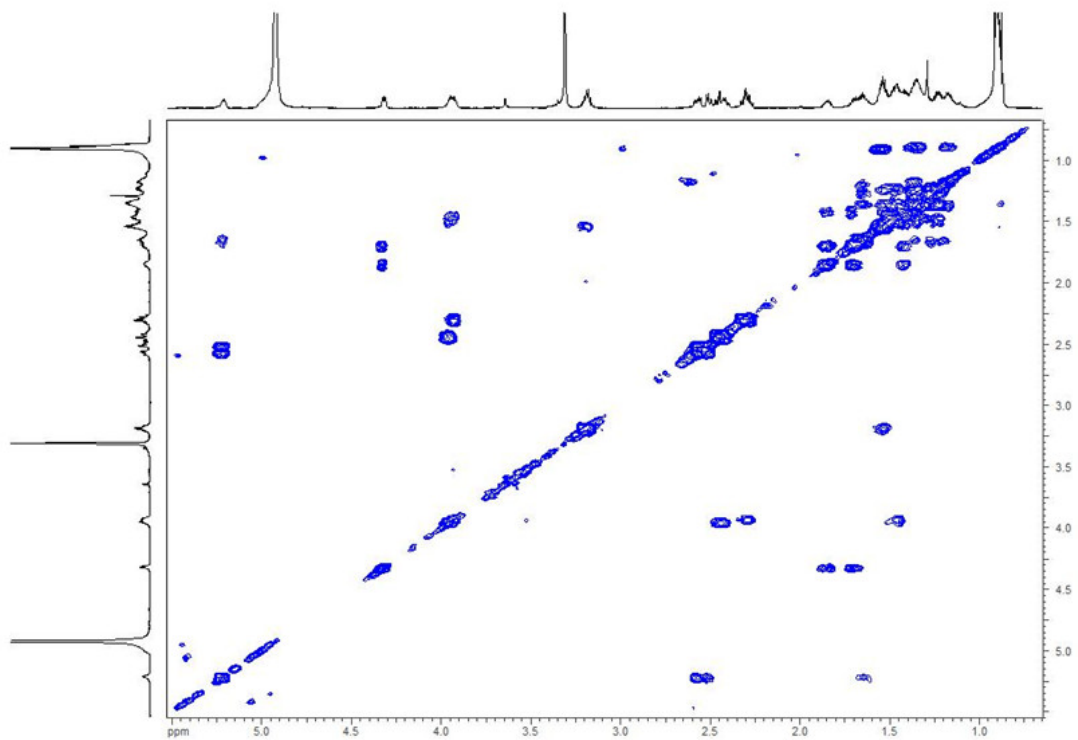**Figure S8.** COSY (700 MHz, CD₃OD) spectrum of dudomycin A.
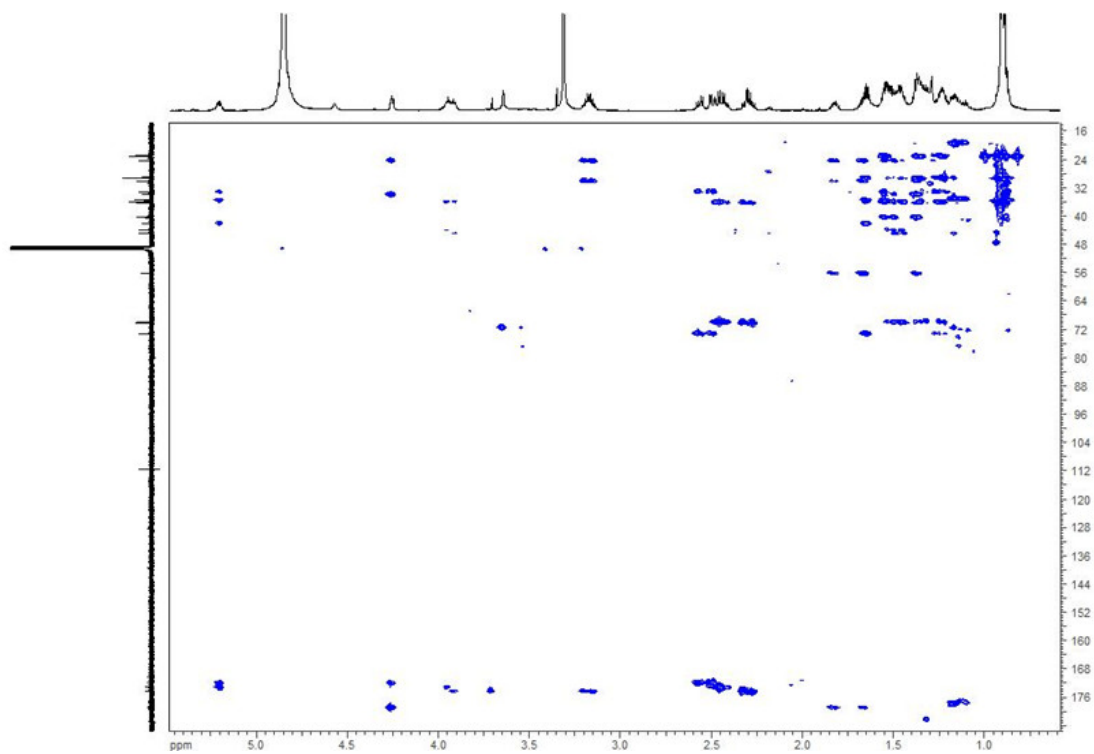
Figure S9. HMBC (700 MHz, CD₃OD) spectrum of dudomycin A.
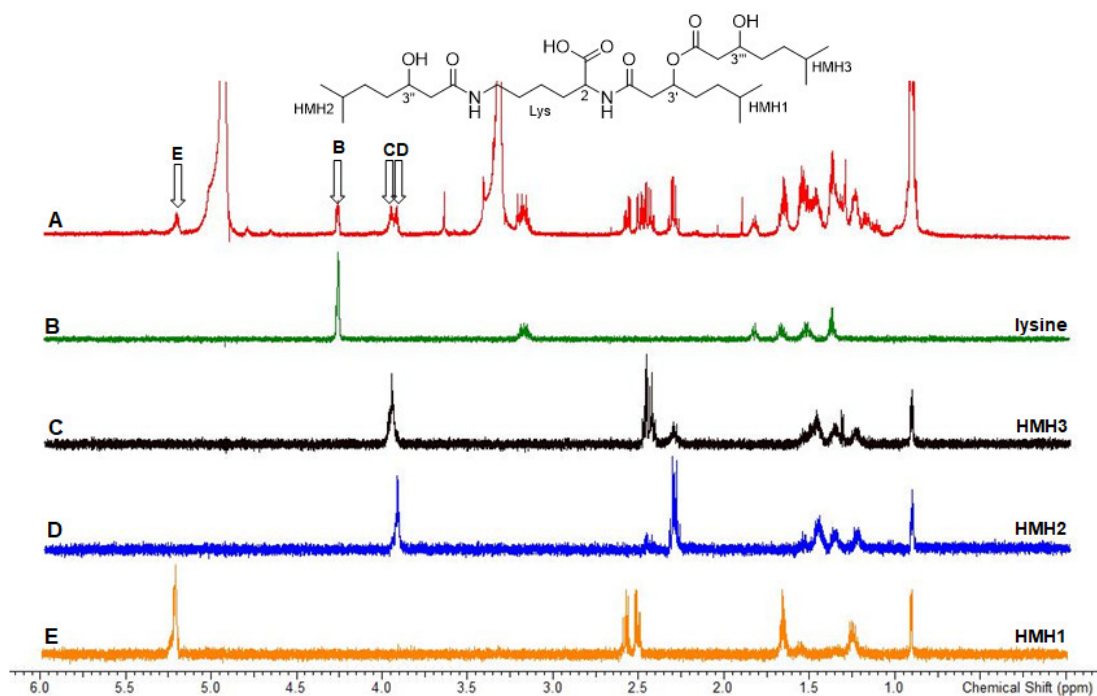
Figure S10. A. ¹H NMR spectrum of dudomycin A showing the irradiation points B (Lys, H-2), C (HMH3, H-3′′′), D (HMH2, H-3′′) and E (HMH1, H 3′). B-E. Selective 1D TOCSY (700 MHz, CD₃OD) spectra of the corresponding irradiation points.
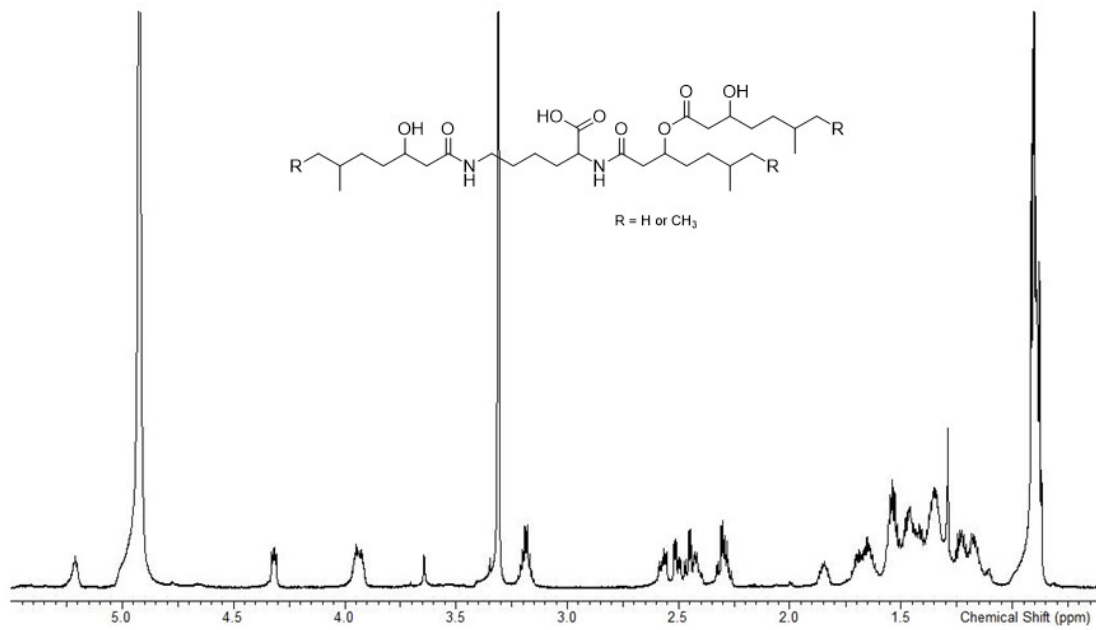
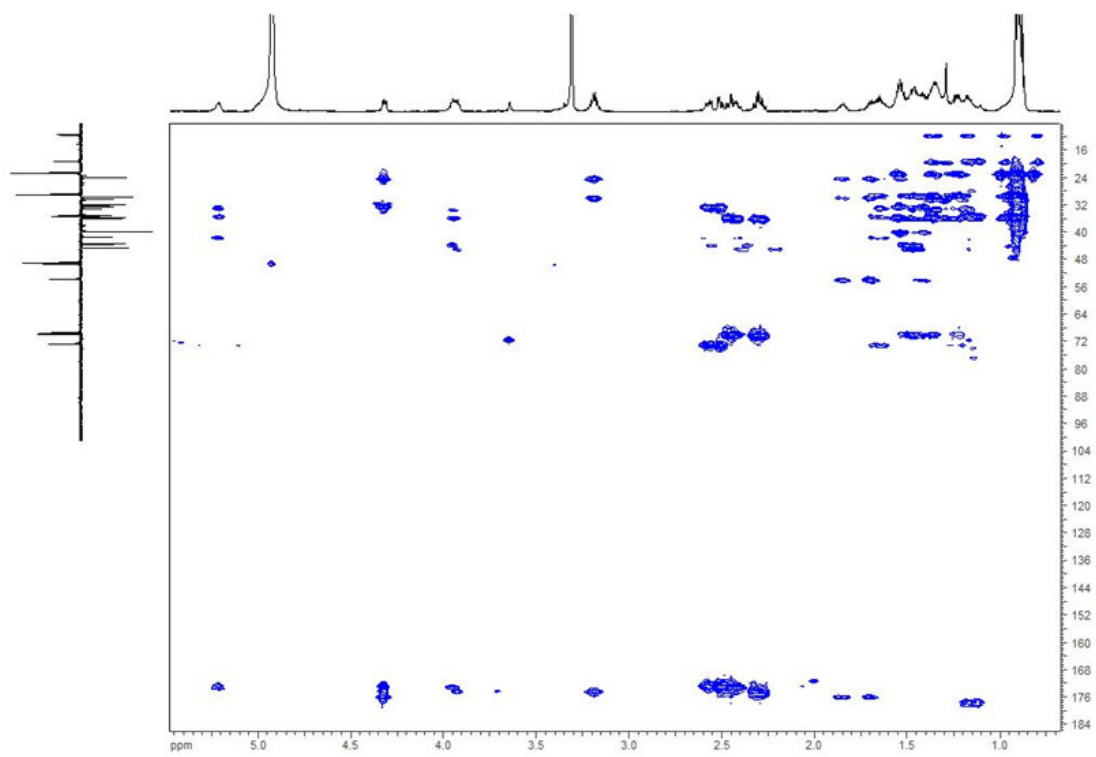**Figure S11.** ¹H NMR (700 MHz, CD₃OD) spectrum of dudomycin B.
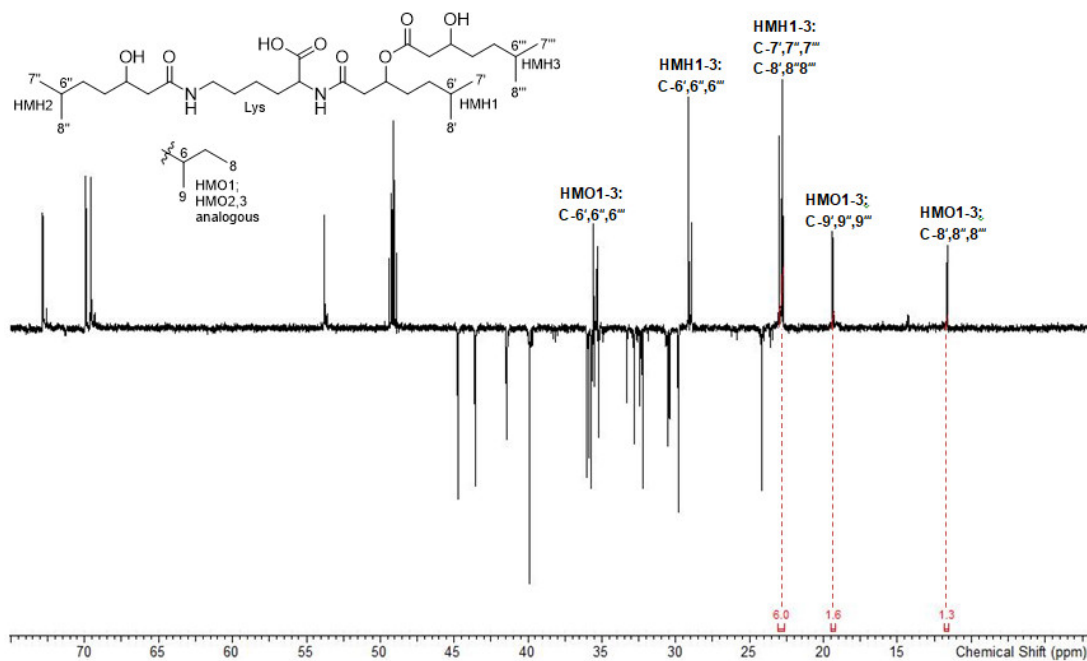
Figure S12. DEPT-135 (176 MHz, CD₃OD) spectrum of dudomycin B showing a 2/1 ratio of the methyl signals of HMH and HMO.
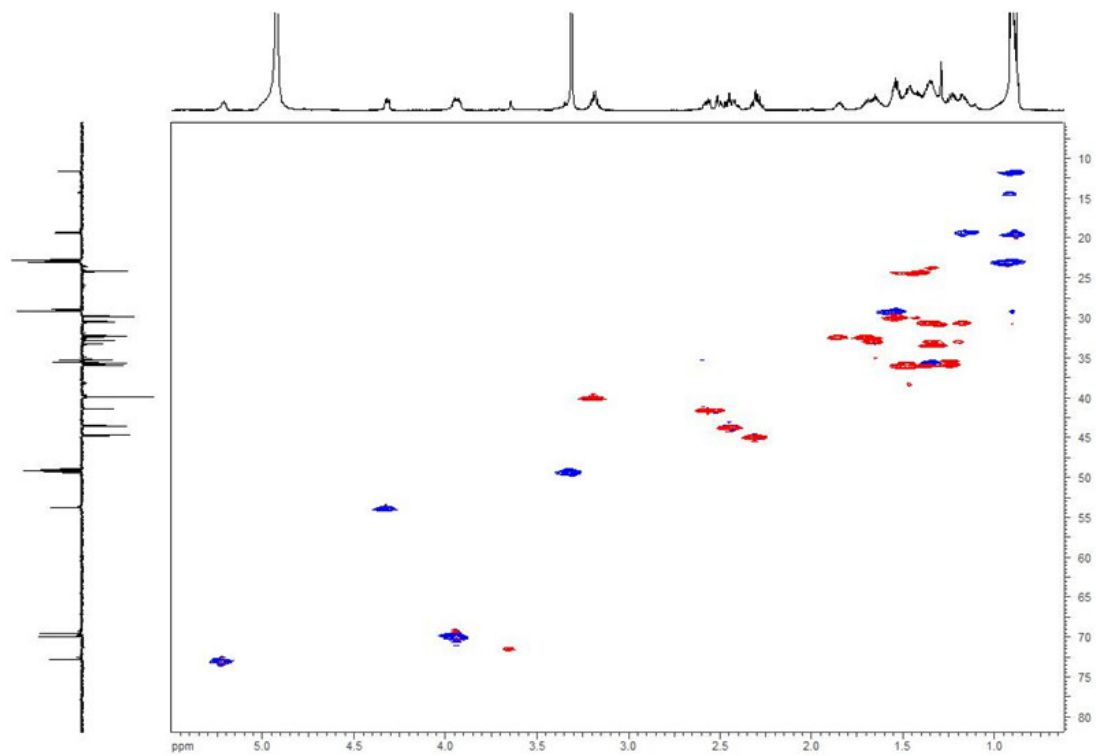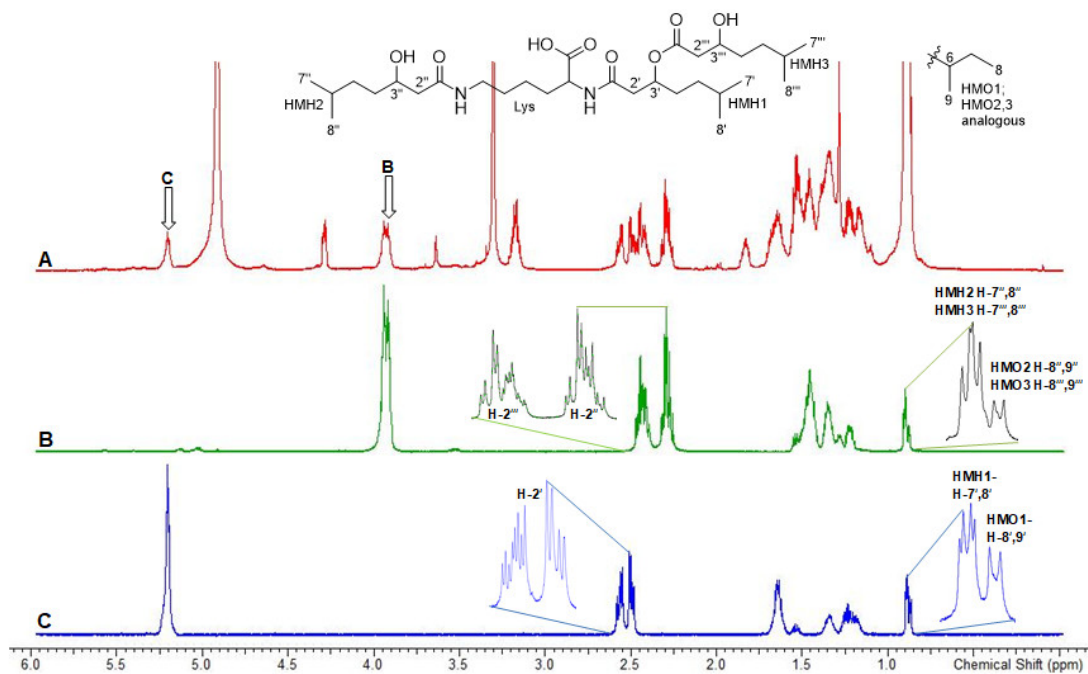
Figure S13. HSQC (700 MHz, CD₃OD) spectrum of dudomycin B.

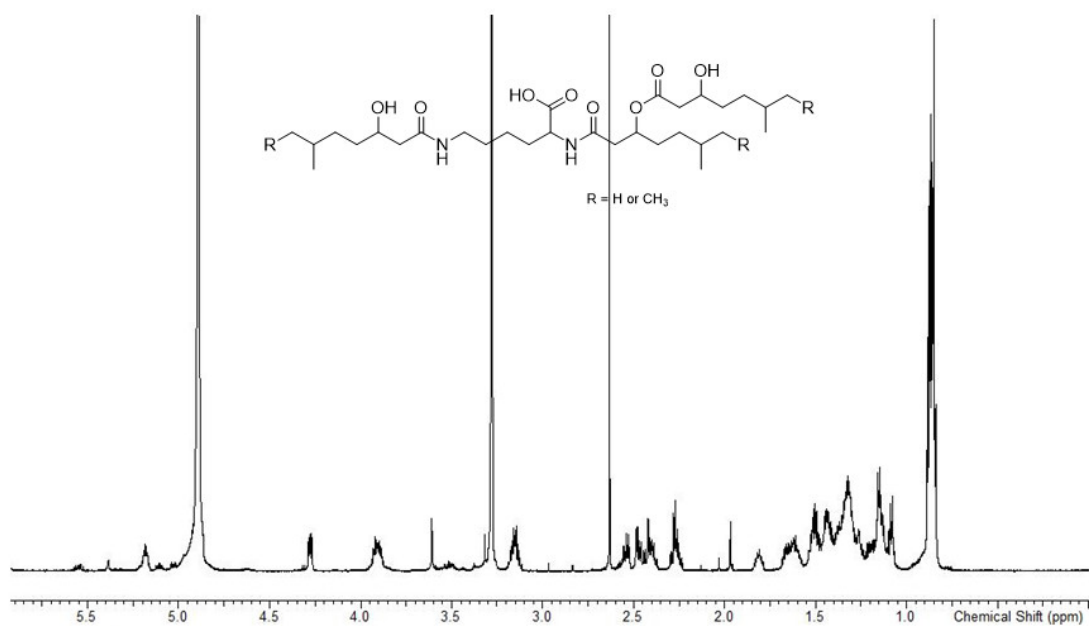**Figure S14.** COSY (700 MHz, CD₃OD) spectrum of dudomycin B.

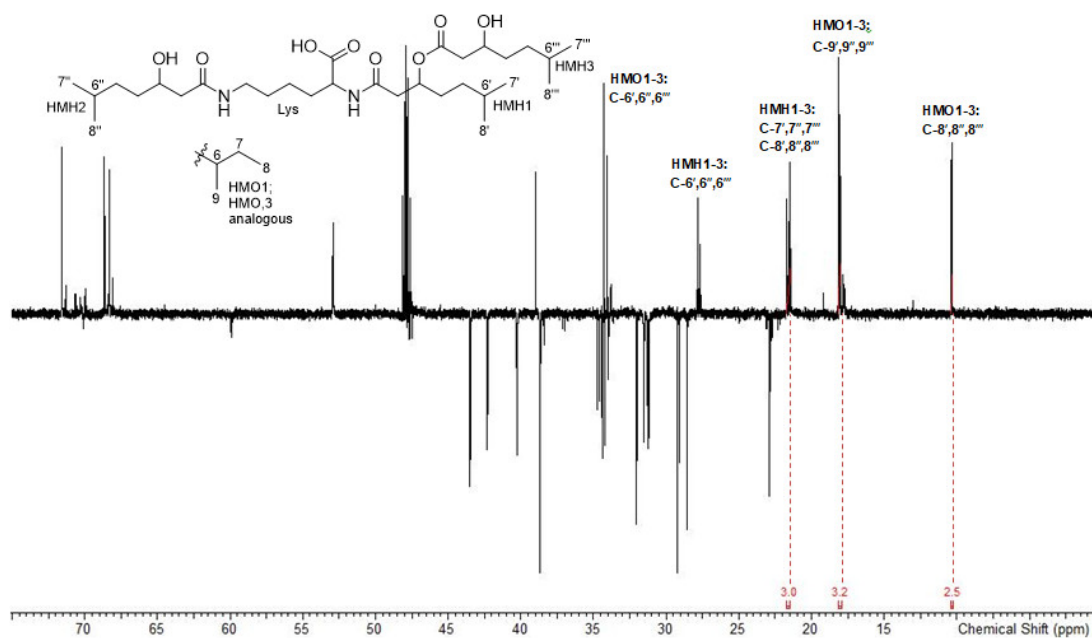**Figure S15.** HMBC (700 MHz, CD₃OD) spectrum of dudomycin B.

Figure S16. A. ¹H NMR spectrum of dudomycin B showing the irradiation points B (H-3″/ H-3‴) and C (H-3′). B-D. Selective 1D TOCSY (700 MHz, CD₃OD) spectra of the corresponding irradiation points.
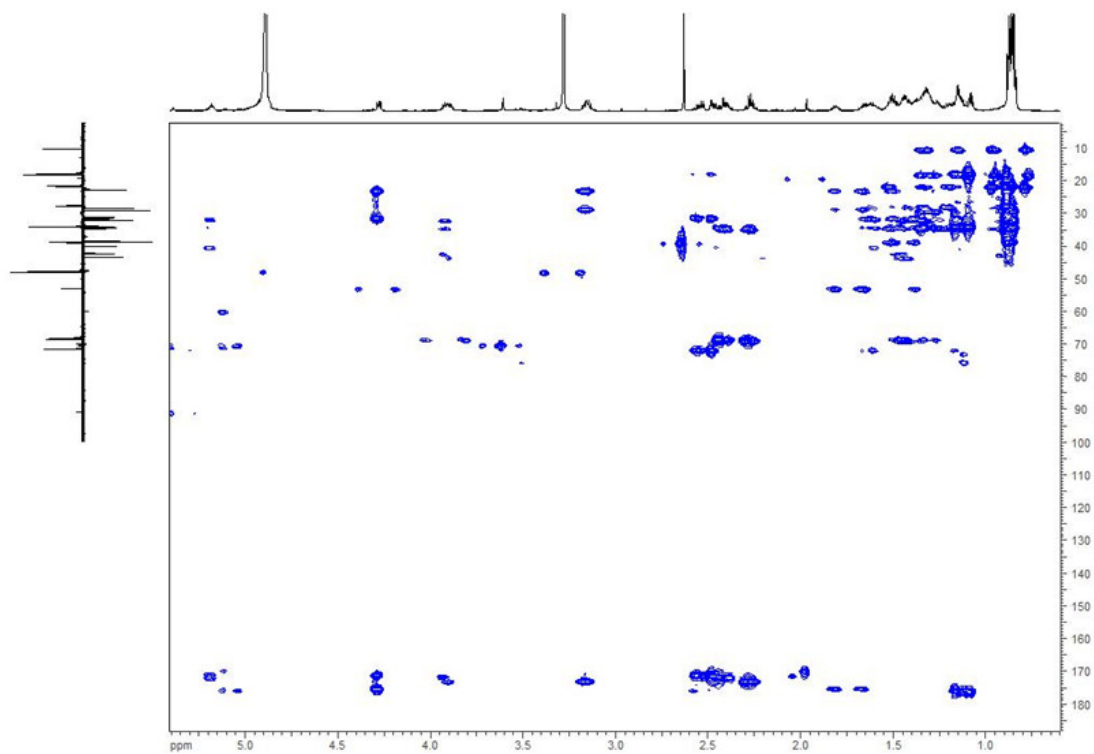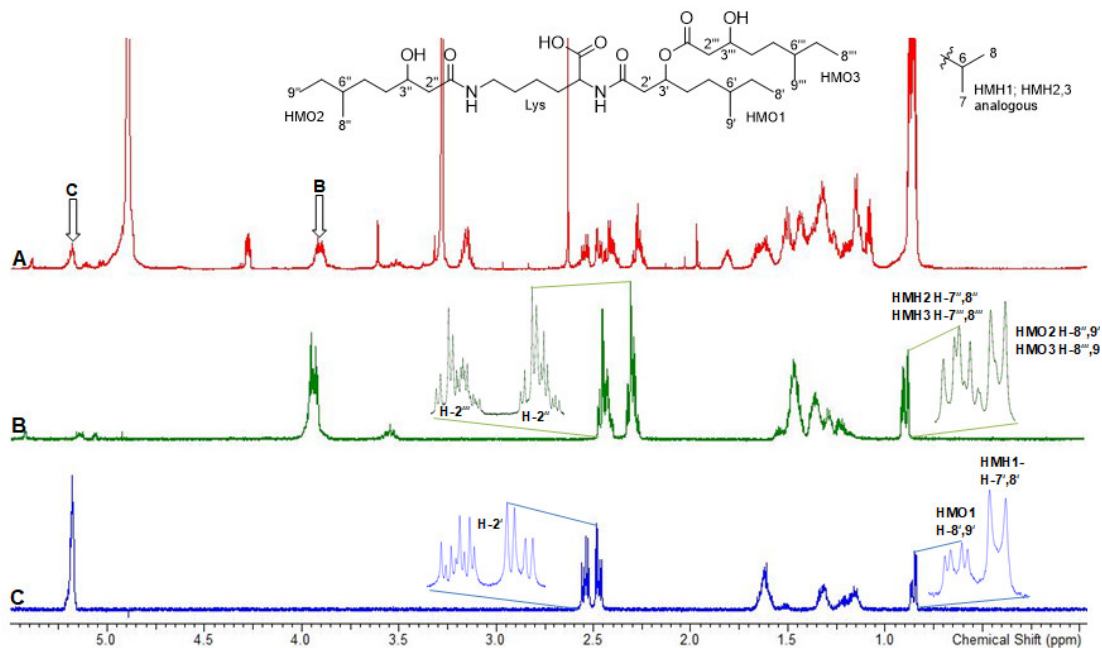
Figure S17. ^{1}H NMR (700 MHz, CD$_3$OD) spectrum of dudomycin C.

Figure S18. DEPT-135 (176 MHz, CD₃OD) spectrum of dudomycin C showing an approximate 1/2 ratio of the methyl signals of HMH and HMO.

Figure S19. HMBC (700 MHz, CD₃OD) spectrum of dudomycin C.

Figure S20. A. ¹H NMR spectrum of dudomycin C showing the irradiation points B (H-3''/ H-3''') and C (H-3'). B-D. Selective 1D TOCSY (700 MHz, CD₃OD) spectra of the corresponding irradiation points.

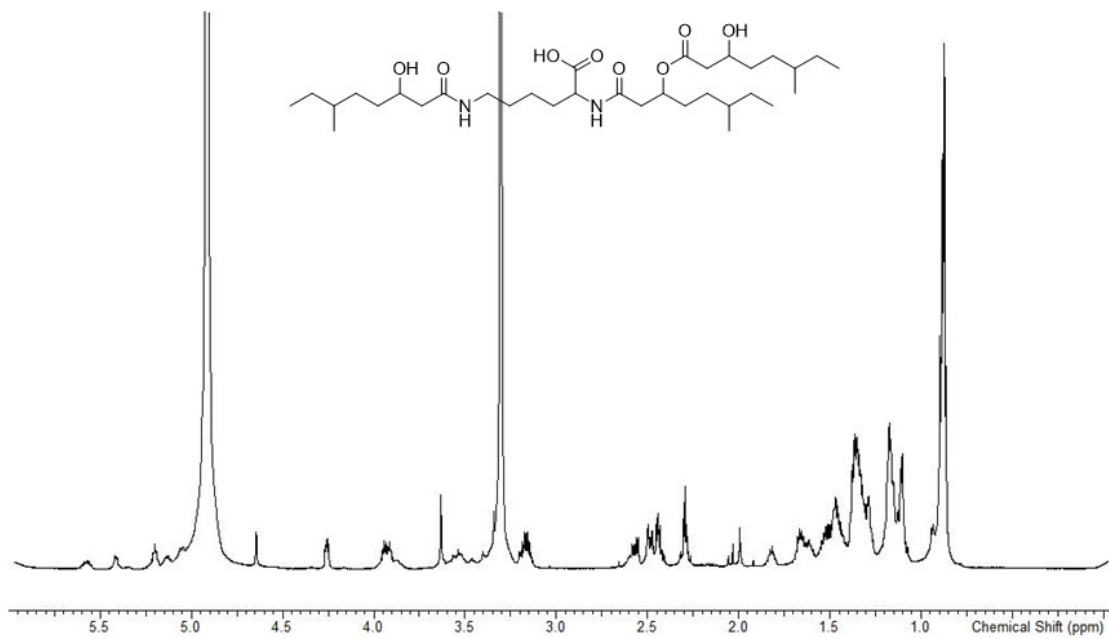**Figure S21.** ¹H NMR (700 MHz, CD₃OD) spectrum of dudomycin D.
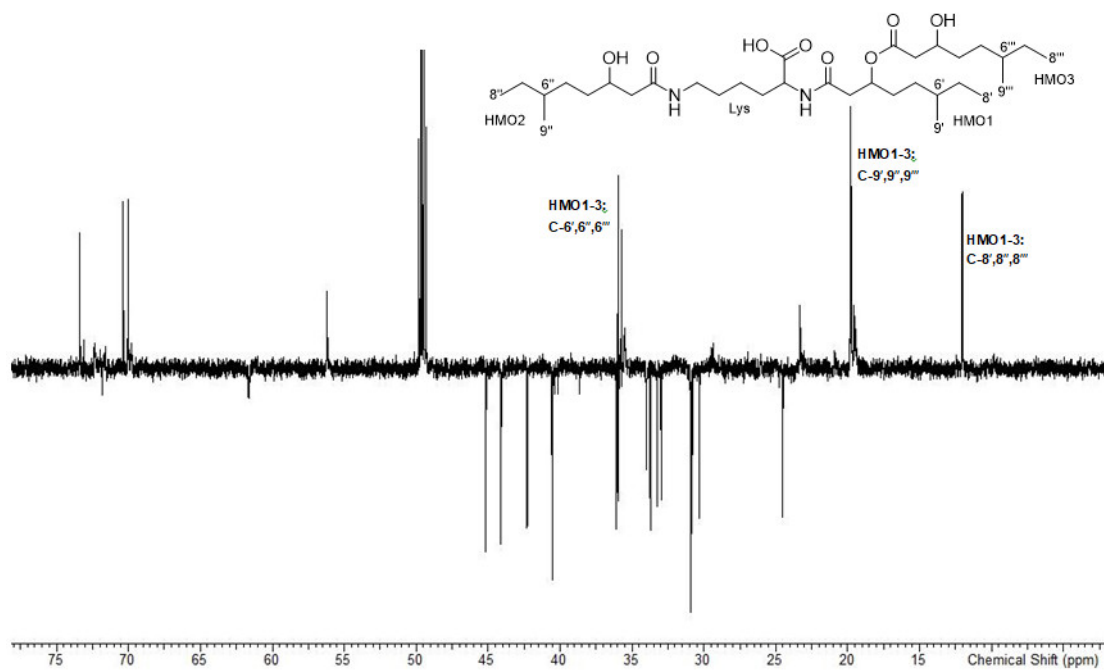
Figure S22. DEPT-135 (176 MHz, CD₃OD) spectrum of dudomycin D.
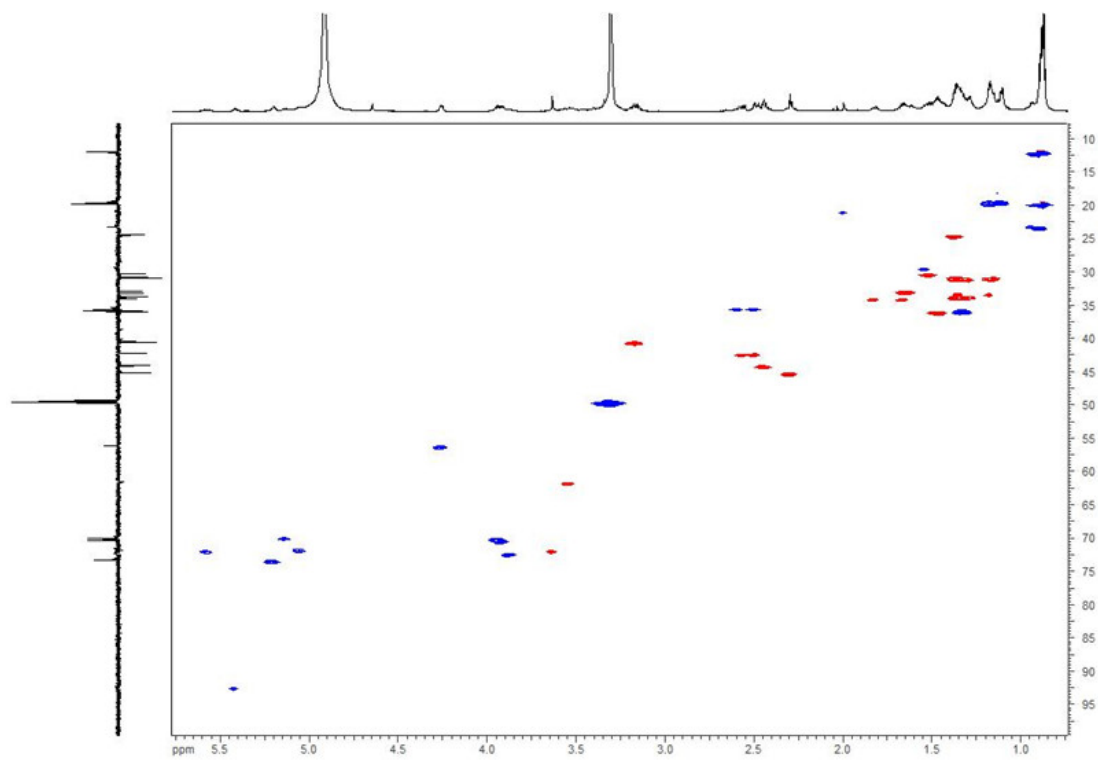
Figure S23. HSQC (700 MHz, CD₃OD) spectrum of dudomycin D.
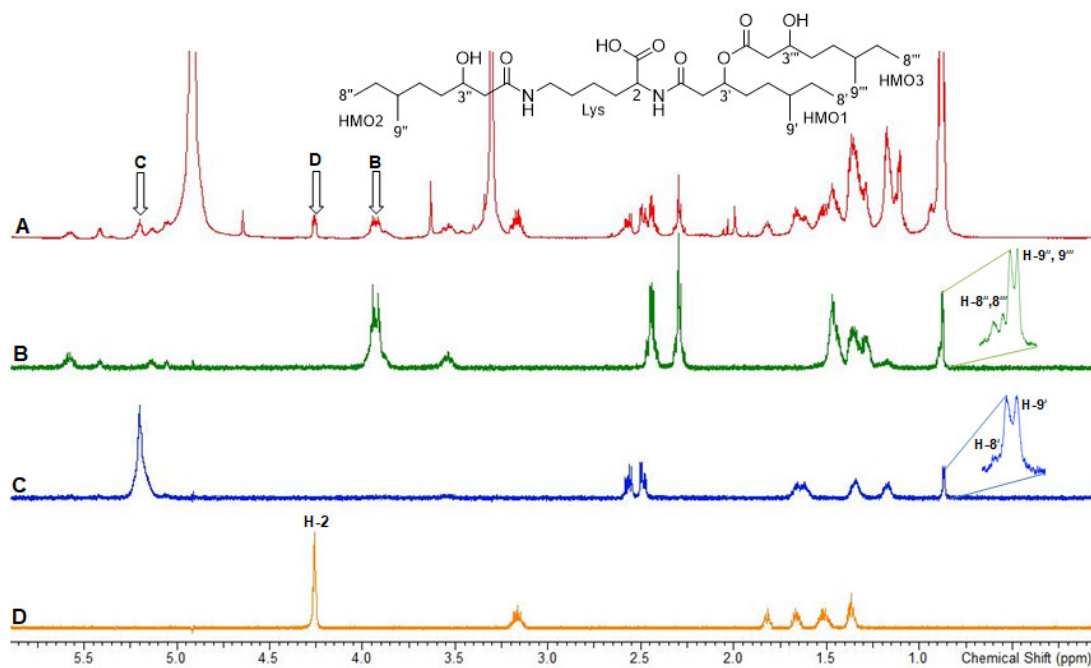
Figure S24. A. ¹H NMR spectrum of dudomycin D showing the irradiation points B (HMO2/3, H-3''/ H-3'''), C (HMO1, H-3') and D (Lys, H-2). B D. Selective 1D TOCSY (700 MHz, CD₃OD) spectra of the corresponding irradiation points.
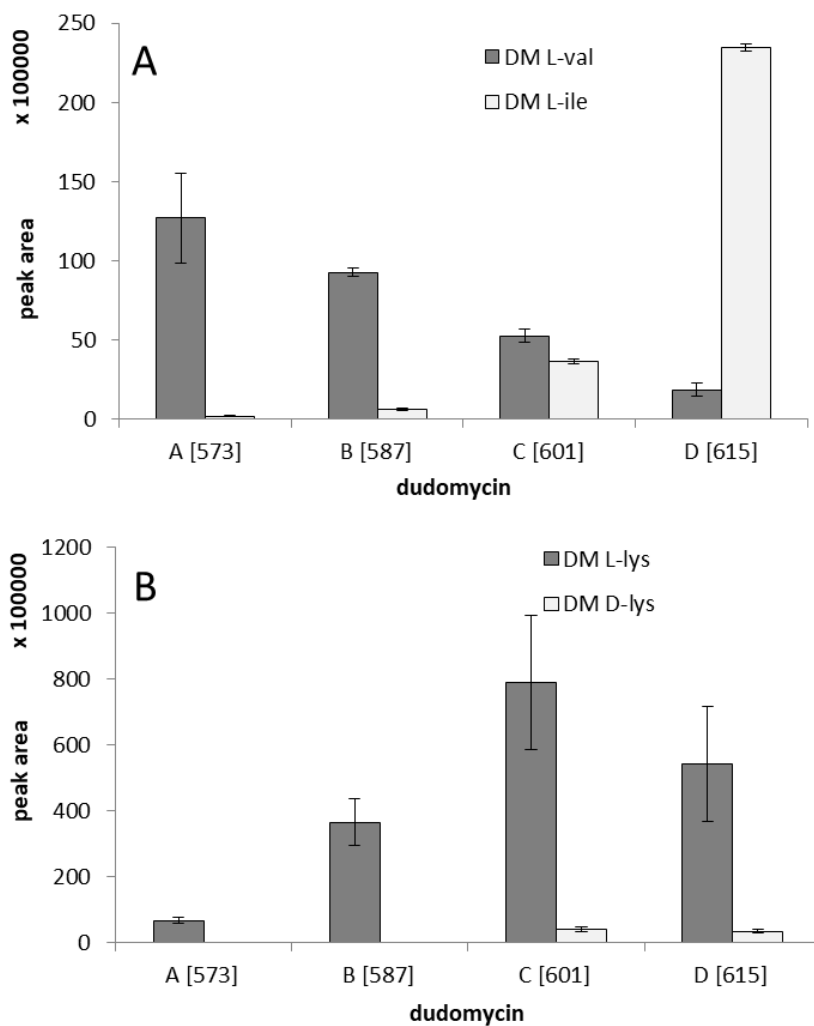
Figure S25. Production of dudomycin derivatives in defined medium (DM) with single sources of nitrogen.
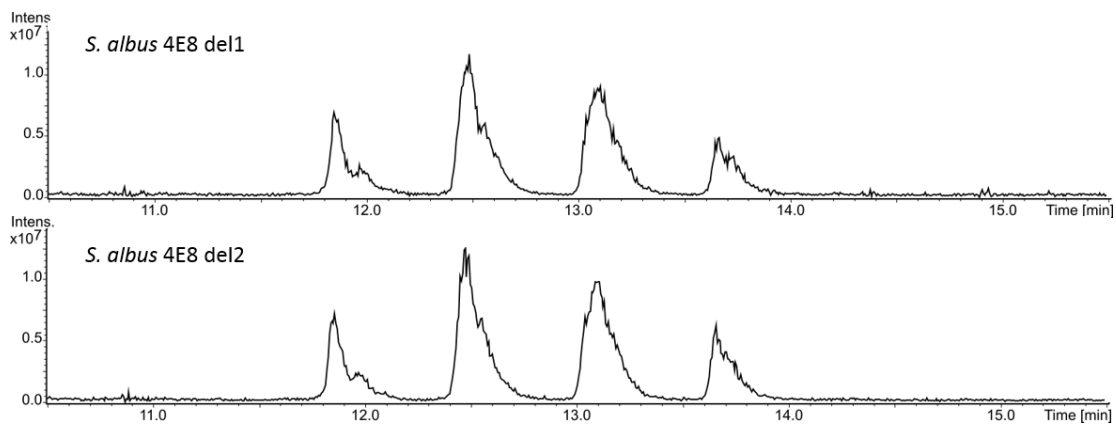
Figure S26. Production of dudomycins after gene deletion experiments. BPC extracted for masses [573-574], [587-588], [601-602], [615-616].
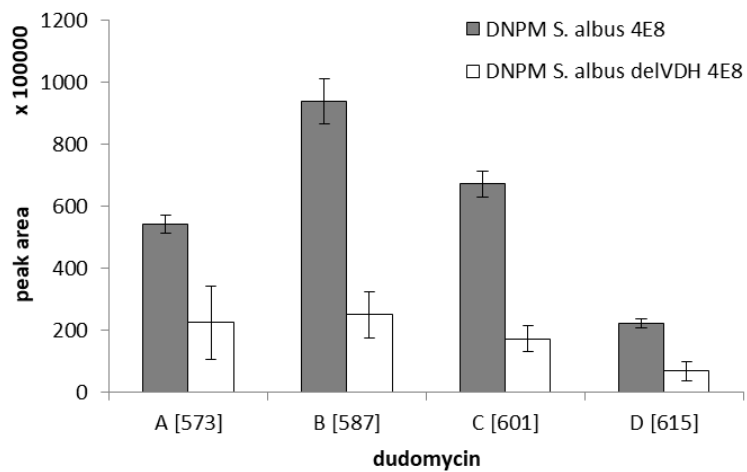
Figure S27. Decreased production of dudomycin derivatives by *S. albus* delVDH 4E8.

Table S1. NMR data of dudomycin A in CD$_3$OD.

unit	no.	δ_c, type	δH, (J in Hz)
Lys	1	178.9, C	-
	2	56.2, CH	4.26, dd (4.8, 7.5)
	3	34.0, CH$_2$	1.83, m; 1.677, m
	4	24.3, CH$_2$	1.38, m
	5	30.2, CH$_2$	1.51, m
	6	40.4, CH$_2$	3.17, m
	NH-1	-	7.12, bs[1]
	NH-2	-	6.81, bs[1]
HMH1	1'	171.9, C	-
	2'	42.2, CH$_2$	2.56, dd (7.7, 14.1); 2.50, dd (5.1, 14.4)
	3'	73.3, CH	5.20, dd (4.94, 7.5)
	4'	33.2, CH$_2$	1.64, m
	5'	35.6, CH$_2$	1.23, m
	6'	29.2, CH	1.54, m
	7'	23.1, CH$_3$	0.88, d (6.5)
	8'	23.0, CH$_3$	0.89, d (6.5)
HMH2	1"	174.3	-
	2"	44.9, CH$_2$	2.32, dd (4.5, 13.9); 2.28, dd (8.0, 13.9)
	3"	70.2, CH	3.92, m
	4"	36.3, CH$_2$	1.47, m
	5"	36.0, CH$_2$	1.23, m
	6"	29.4, CH	1.56, m
	7"	23.1, CH$_3$	0.90, d (4.4)
	8"	23.1, CH$_3$	0.89, d (4.4)
HMH3	1'''	173.2	-
	2'''	43.9, CH$_2$	2.47, dd (5.4, 14.8); 2.42, d (7.6, 14.9)
	3'''	69.9, CH	3.95, m
	4'''	36.2, CH$_2$	1.49, m; 1.348, m
	5'''	36.0, CH$_2$	1.23, m
	6'''	29.4, CH	1.56, m
	7'''	23.3, CH$_3$	0.91, d (4.8)
	8'''	23.3, CH$_3$	0.90, d (4.8)

[1] NH Signals measured in CDCl$_3$

Table S2. NMR data of dudomycin D in CD₃OD.

unit	no.	δC‡, type	δH, (J in Hz)
Lys	2	55.8, CH	4.26, dd (4.5, 6.8)
	3	33.6, CH₂	1.82, m; 1.66, m
	4	24.1, CH₂	1.38, m
	5	29.9, CH₂	1.51, m
	6	40.1, CH₂	3.17, m
HMO1	2'	41.9, CH₂	2.57, dd (8.5, 14.4); 2.49, dd (4.9, 14.4)
	3'	72.9, CH	5.20, m
	4'	32.8, CH₂	1.64, m
	5'	32.5, CH₂	1.36, m; 1.18, m
	6'	35.3, CH	1.34, m
	7'	30.4, CH₂	1.33, m; 1,17, m
	8'	11.6, CH₃	0.88, ovl
	9'	19.3, CH₃	0.87, d (6.0)
HMO2	2"	44.7, CH₂	2.31, dd (4.9, 13.5); 2.28, dd (8.2, 13.6)
	3"	69.9, CH	3.92, m
	4"	35.6, CH₂	1.47, m
	5"	33.3, CH₂	1.30, m; 1.37, m
	6"	35.5, CH	1.34, m
	7"	30.5, CH₂	1.35, m; 1.37, m
	8"	11.6, CH₃	0.89, ovl
	9"	19.4, CH₃	0.88, d (5.4)
HMO3	2‴	43.7, CH₂	2.46, dd (5.5, 15.1); 2.424, d (8.2, 15.1)
	3‴	69.6, CH	3.94, m
	4‴	35.5, CH₂	1.47, m
	5‴	33.3, CH₂	1.30, m; 1.37, m
	6‴	35.5, CH	1.34, m
	7‴	30.5, CH₂	1.35, m; 1.37, m
	8‴	11.6 CH₃	0.89, ovl
	9‴	19.4, CH₃	0.88, d (5.4)

‡from DEPT-135 measurement

Table S3. Strains, BACs, plasmids and primers used in this work.

Material	Purpose
A. Bacterial strains	
Streptomyces albus Del14	optimized heterologous host [1]
Streptomyces lividans Del8	optimized heterologous host [2]
Streptomyces albus J1074 delVDH	heterologous host [3]
Escherichia coli GB05 RedCC	cloning host [Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS)]
Escherichia coli ET12567 pUB307	alternate host intergeneric conjugation [4]
B. BACs	
4E8	heterologous expression of NRPS cluster
4E8 del1	determination left border of NRPS cluster
4E8 del2	determination right border of NRPS cluster
C. Plasmids	
pUC19	ampicillin resistance marker
pXCM hygformax	hygromycin resistance marker
D. PCR primer	
20190429_1_fw [4E8 del1]	CAGGGAGGAGGCCACCCCGCGCCGGTAGAAGACGCCCAGCAGGCCGACCGCGTCAGGTGGCACTTTTCG
20190429_1_rev [4E8 del1]	GATCGCGTGCATGCCCAGATGGTCGCGCGCATACCGCCGTCGGAACGGGCGCGGATATCCTACTATGCCGAGGTATAATGTAGCCAGCGTGTTACCAATGCTTAATCAGTG
20190429_2_fw [4E8 del2]	GGCGACGGCTCCCCGGGCCCGCGAGCCCACGACTCCCGTCCCCACGGCCATCAGGCGCCGGGGGCGGTGT
20190429_2_rev [4E8 del2]	GCGATGAGTATCCGTACTCATGTCCGGGCCGCGGGTGCCGCCTAGCGTGAAATACTTGACATATCACTGT

Table S4. Proposed functions of the genes in the DNA fragment containing the dudomycin gene cluster

gene #	locus tag	putative function
1	SACHL_42600	glycosyltransferase
2	SACHL_42610	hypothetical protein
3	SACHL_42620	phosphatase
4	SACHL_42630	N,N'-diacetyllegionaminic acid synthase
5	SACHL_42640	hypothetical protein
6	SACHL_42650	hydrolase
7	SACHL_42660	membrane lipoprotein precursor
8	SACHL_42670	galactose/methyl galactoside import ATP-binding protein
9	SACHL_42680	ribose transport system permease protein
10	SACHL_42690	branched-chain amino acid transport system / permease component
11	SACHL_42700	cytidine deaminase
12	SACHL_42710	pyrimidine-nucleoside phosphorylase
13	SACHL_42720	hypothetical protein
14	SACHL_42730	hypothetical protein
15	SACHL_42740	hypothetical protein
16	SACHL_42750	hypothetical protein
17	SACHL_42760	hypothetical protein
18	SACHL_42770	hypothetical protein
19	SACHL_42780	ubiquinone biosynthesis O-methyltransferase
20	SACHL_42790	zinc metallo-peptidase
21	SACHL_42800	hypothetical protein
22 [*dudA*]	SACHL_42810	dimodular nonribosomal peptide synthase
23 [*dudB*]	SACHL_42820	inner membrane transport protein
24 [*dudC*]	SACHL_42830	transcriptional regulator
25	SACHL_42840	demethylrebeccamycin-D-glucose O-methyltransferase
26	SACHL_42850	hypothetical protein

REFERENCES

[1] M. Myronovskyi, B. Rosenkränzer, S. Nadmid, P. Pujic, P. Normand, and A. Luzhetskyy, "Generation of a cluster-free Streptomyces albus chassis strains for improved heterologous expression of secondary metabolite clusters," *Metab. Eng.*, vol. 49, pp. 316–324, Sep. 2018, doi: 10.1016/j.ymben.2018.09.004.

[2] Y. Ahmed, Y. Rebets, M. R. Estévez, J. Zapp, M. Myronovskyi, and A. Luzhetskyy, "Engineering of Streptomyces lividans for heterologous expression of secondary metabolite gene clusters," *Microb. Cell Fact.*, vol. 19, no. 1, p. 5, Dec. 2020, doi: 10.1186/s12934-020-1277-8.

[3] N. Manderscheid, "Strain development for heterologous expression of secondary metabolite clusters in actinobacteria," Universität des Saarlandes, 2015.

[4] F. Flett, V. Mersinias, and C. P. Smith, "High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl DNA-restricting streptomycetes," *FEMS Microbiol. Lett.*, vol. 155, no. 2, pp. 223–229, Jan. 2006, doi: 10.1111/j.1574-6968.1997.tb13882.x.

III.

Bonsecamin:

A New Cyclic Pentapeptide Discovered through Heterologous Expression of a Cryptic Gene Cluster

Constanze Lasch, Marc Stierhof, Marta Rodríguez Estévez, Maksym Myronovskyi,

Josef Zapp and Andriy Luzhetskyy

Article

Bonsecamin: A New Cyclic Pentapeptide Discovered through Heterologous Expression of a Cryptic Gene Cluster

Constanze Lasch [1], **Marc Stierhof** [1], **Marta Rodríguez Estévez** [1], **Maksym Myronovskyi** [1], **Josef Zapp** [2] **and Andriy Luzhetskyy** [1,3,*]

[1] Department of Pharmaceutical Biotechnology, Saarland University, 66123 Saarbruecken, Germany; constanze.lasch@uni-saarland.de (C.L.); m.stierhof@t-online.de (M.S.); marta.rodriguezestevez@uni-saarland.de (M.R.E.); maksym.myronovskyi@uni-saarland.de (M.M.)

[2] Department of Pharmaceutical Biology, Saarland University, 66123 Saarbruecken, Germany; j.zapp@mx.uni-saarland.de

[3] Helmholtz Institute for Pharmaceutical Research Saarland, 66123 Saarbruecken, Germany

[*] Correspondence: a.luzhetskyy@mx.uni-saarland.de; Tel.: +49-681-302-70200

Abstract: The intriguing structural complexity of molecules produced by natural organisms is uncontested. Natural scaffolds serve as an important basis for the development of molecules with broad applications, e.g., therapeutics or agrochemicals. Research in recent decades has demonstrated that by means of classic metabolite extraction from microbes only a small portion of natural products can be accessed. The use of genome mining and heterologous expression approaches represents a promising way to discover new natural compounds. In this paper we report the discovery of a novel cyclic pentapeptide called bonsecamin through the heterologous expression of a cryptic NRPS gene cluster from *Streptomyces albus* ssp. *chlorinus* NRRL B-24108 in *Streptomyces albus* Del14. The new compound was successfully isolated and structurally characterized using NMR. The minimal set of genes required for bonsecamin production was determined through bioinformatic analysis and gene deletion experiments. A biosynthetic route leading to the production of bonsecamin is proposed in this paper.

Keywords: *Streptomyces*; NRPS; heterologous expression; cyclic peptide

1. Introduction

In the last two decades a considerable number of new small cyclic natural peptides produced by the bacterial genus of *Streptomyces* were discovered and published [1–8]. In nature, those molecules often have either toxic functions or serve the producing organism to coordinate the uptake of metal ions acting as a chelator [9,10]. Several naturally occurring cyclic peptides have proven their potential for use in pharmacotherapy. Remarkable antitumor and antibacterial activity in the recently discovered cyclic peptides chloptosin, hytramycins and mannopeptimycins, was reported by researchers [5,7,11]. Other compounds such as the antitumor agents actinomycin D and romidepsin, or immunosuppressant cyclosporine A, are already established as marketed drugs [12–15].

Cyclic peptides are often the product of enzymes belonging to the class of nonribosomal peptide synthetases (NRPS). NRPS are large modular enzymes or enzymatic complexes with each module responsible for the incorporation of a single amino acid residue into a nascent peptide chain. The individual modules can be split into separate domains. The typical elongation module contains a minimal set of condensation (C), adenylation (A) and peptidyl carrier (PCP) domains, where the A domain catalyzes the selective activation of an amino acid, the PCP domain holds the activated amino acid or the growing peptide chain and the C domain catalyzes the formation of the amide bond between two PCP-bound amino acid substrates. NRPS-derived molecules can display uncommon structural features, e.g., by incorporation of non-proteinogenic amino acids or when the amino acids

are structurally modified by tailoring domains. The chemical diversity of NRPS products is further expanded by the existence of hybrid NRPS clusters [16].

Recently we reported the successful expression of several cryptic secondary metabolite gene clusters of *Streptomyces albus* ssp. *chlorinus* in the engineered heterologous host strain *Streptomyces albus* Del14. This led to the identification of the biosynthetic genes of the characterized bioactive compounds, albucidin and nybomycin, as well as to the discovery of the new natural products, benzanthric acid, fredericamycin C_2 and dudomycins [17–21]. In this paper we report the discovery of a new cyclic pentapeptide, bonsecamin, through the heterologous expression of a cryptic NRPS gene cluster of *S. albus* ssp. *chlorinus* in *S. albus* Del14 (Figure 1). Bonsecamin was purified and its structure was elucidated by NMR. The bioinformatic analysis of the putative bonsecamin biosynthetic cluster as well as targeted gene inactivation experiments allowed the determination of the minimal set of genes responsible for the biosynthesis of the compound. A biosynthetic scheme leading to the production of bonsecamin is proposed in this paper.
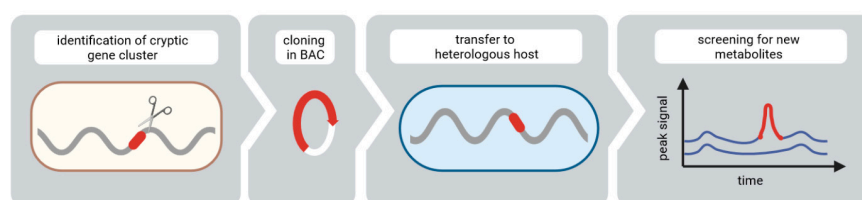
Figure 1. General workflow for compound discovery using heterologous expression.

2. Materials and Methods

2.1. General Procedures

Table S1 provides an overview of all bacterial strains, BACs and plasmids used in this work. Lysogeny broth (LB) medium was used for cultivation of *Escherichia coli* strains [22]. Soy flour mannitol (MS) agar [23] and tryptic soy broth (TSB; Sigma-Aldrich, St. Louis, MO, USA) were used to grow *Streptomyces* strains. Metabolites were extracted from either liquid DNPM medium (40 g/L dextrin, 7.5 g/L soytone, 5 g/L baking yeast, and 21 g/L MOPS, pH 6.8 as aqueous solution) or defined medium DM (mannitol 5 g/L, amino acid 0.5 g/L, K2PO$_4$ 0.5 g/L, MgSO$_4$ x 7 H$_2$O 0.2 g/L, FeSO$_4$ x 7 H$_2$O 0.01 g/L). Before inoculation of DM, the cells of the preculture were washed three times with amino-acid-free DM. Amino acids L-lys, L-val, L-ala, D-thr and D-ser were supplied to the defined medium as required by the experimental design. The antibiotics kanamycin, apramycin, ampicillin and nalidixic acid were added when needed.

2.2. DNA Isolation and Manipulation

Standard protocols were used to carry out DNA manipulation, transformation into *E. coli* and intergeneric conjugation between *E. coli* and *Streptomyces* [22–24]. The BACMAX™ DNA purification kit (Lucigen, Middleton, WI, USA) was used to isolate BAC DNA from a constructed genomic library of *Streptomyces albus* ssp. *chlorinus* NRRL B-24108. Cluster borders were determined by deletion of several genes downstream or upstream from the expected gene cluster with deletion of genes 15 to 28 leading to minimal BAC 2O18_del1 and deletion of genes 1 to 7 leading to BAC 2O18_del2. The genes were replaced by the resistance marker ampicillin through homologous recombination using the RedET system [25]. For amplification of the respective gene cassettes from the plasmid pUC19 PCR reactions were performed. The PCR primer pairs 20200815_1_fw / 20200815_1_rev and 20200815_2_fw / 20200815_2_rev were constructed with overhang regions allowing the site-specific introduction of the cassettes at both sides of the expected bonsecamin gene cluster. The success of the recombination was assessed by restriction mapping and DNA sequencing. The restriction enzymes purchased from ThermoFisher Scientific (Waltham, MA, USA) or New England BioLabs NEB (Ipswich, MA, USA) were used according to the manual. The deletion of three single genes was carried out as described above. To avoid

polar effects, additional *Pme*I recognition sites in primers 20201217_1_fw / 20201217_1_rev, 20201217_2_fw / 20201217_2_rev and 20201217_4_fw / 20201217_4_rev allowed the precise removal of the ampicillin resistance gene and religation to BACs 2O18_delKR_delbla, 2O18_delPCP_delbla, and 2O18_delTE2_delbla.

2.3. Metabolite Extraction

Streptomyces strains were grown in 15 mL of TSB medium in a 100 mL baffled flask for 1 to 2 days. 1 mL of this preculture was used to inoculate 100 mL of production medium in a 500 mL baffled flask. Cultures were incubated for 6 to 7 days at 28 °C and 180 rpm in an Infors multitron shaker (Infors AG, Basel, Switzerland) for metabolite production. The metabolites were extracted from the culture supernatant with an equal amount of *n*-butanol, dried at up to 50°C and stored at 4°C.

2.4. Mass Spectrometry (MS) Analysis of Metabolites

Prior to MS analysis the extracts were dissolved in methanol. MS experiments were performed on a Dionex Ultimate 3000 UPLC system (ThermoFisher Scientific, Waltham, MA, USA) with PDA detector (stationary phase 100 mm ACQUITY UPLC BEH C18 1.7 μm column (Waters Corporation, Milford, MA, USA), mobile phase: linear gradient of [A] ddH2O + 0.1% formic acid / [B] acetonitrile + 0.1% formic acid, 5% to 95% at a flow rate of 0.6 mL/min). For mass detection the system was further coupled to either an amaZon speed (Bruker, Billerica, MA, USA) or LTQ Orbitrap XL mass spectrometer (ThermoFisher Scientific, Waltham, MA, USA) applying standard settings of positive ionization and a mass range detection of m/z 200 to 2000. The data were analyzed by the softwares Compass Data Analysis 4.1 (Bruker) or Xcalibur 3.0 (ThermoFisher Scientific).

2.5. Extract Purification

The crude extracts from the 10 L cultures were dissolved in methanol. Four purification steps were carried out when using DNPM medium for cultivation: (1) Normal Phase (NP) Flash chromatography on an Isolera One system (Biotage, Uppsala, Sweden); stationary phase: SNAP Ultra 50 g (Biotage, Uppsala, Sweden), mobile phase: [A] n-hexane / [B] chloroform / [C] ethyl acetate / [D] methanol in linear gradients of [A]/[B] 10 column volumes (CV), [B]/[C] 15 CV, [C]/[D] 15 CV at a flow of 100 mL/min) was followed by (2) Size Exclusion Chromatography (SEC; stationary phase: Sephadex-LH20; mobile phase: isocratic elution using 100% methanol). Two Reversed Phase (RP) chromatography steps followed on a (3) Waters HPLC system (2545 Binary Gradient module, Waters, Milford, MA, USA); stationary phase: Nucleodur C18 HTec 250/21 5 μm (Macherey-Nagel, Düren, Germany); mobile phase: linear gradient of [A] H_2O + 0.1% formic acid / [B] methanol + 0.1% formic acid, 5% to 95% [B] for 17 min at flow rate of 20 mL/min, mass detection m/z 430 using software MassLynx (Waters)) and an (4) Agilent Infinity 1200 series HPLC system (stationary phase: Synergi 4 μm Fusion-RP 80 Å 250 × 10 (phenomenex, Torrance, CA, USA); mobile phase: linear gradient of [A] H_2O + 0.1% formic acid / [B] acetonitrile + 0.1% formic acid, 10% to 50% [B] for 15.5 min at flow rate of 4 mL/min, detection UV 201 nm followed by fraction control on HPLC-MS) yielding 0.7 mg of pure substance. In between all purification steps the fractions containing bonsecamin were pooled, dried (at up to 50°C) and redissolved in methanol.

2.6. Structure Elucidation by NMR Spectroscopy, Optical Rotation and Marfey's Method

NMR spectra were acquired on a Bruker Avance III 700 MHz spectrometer at 298 K equipped with a 5 mm TCI cryoprobe. The chemical shifts (δ) were reported in parts per million (ppm) relative to TMS. As solvents, deuterated DMSO-*d6* (δH 2.50 ppm, δC 39.51 ppm) from Deutero (Deutero, Kastellaun, Germany) were used. Edited-HSQC, HMBC, ^{1}H-^{1}H COSY, and N-HSQC spectra were recorded using the standard pulse programs from the TOPSPIN v.3.6 software. Optical rotations were measured using a Perkin Elmer Polarimeter Model 241 (Perkin Elmer, Ueberlingen, Germany).

For Marfey's method, bonsecamin was hydrolyzed in 100 μL 6 N HCl at 110 °C for 1 h. While cooling down, the sample was dried for 15 min under nitrogen, dissolved in 110 mL water and 50 μL each were transferred into 1.5 mL Eppendorf tubes. To the hydrolysate, 20 μL of 1N $NaHCO_3$ and 20 μL of 1% L-FDLA or D-FDLA in acetone were added, respectively. The amino acid standards were prepared the same way using L-FDLA only. The reaction mixtures were incubated at 40°C for 90 min at 700 rpm and subsequently quenched with 2N HCl to stop the reaction. The samples were diluted with 300 μL ACN and 1 μL was analyzed by maXis high-resolution LC-QTOF system using aqueous ACN with 0.1 vol% formic acid and an adjusted gradient of 5–10 vol% for 2 min, 10–25 vol% for 13 min, 25–50 vol% for 7 min and 50–95 vol% for 2 min. Detection was carried out at 340 nm.

2.7. Genome Mining and Bioinformatic Analysis

The antiSMASH online tool was used to screen the genome of *S. albus* ssp. *chlorinus* (https://antismash.secondarymetabolites.org/#!/start, accessed on 10 July 2021) [26]. Analysis of the genetic data was performed by Geneious software, version 11.0.3 [27]. The genomic sequence of *Streptomyces albus* ssp. *chlorinus* can be accessed in GenBank under VJOK00000000. The Dictionary of Natural Products (DNP) 28.1 was used as reference database of so far characterized metabolites.

3. Results and Discussion

3.1. Identification and Expression of the NRPS Gene Cluster

Recently, the potential of the strain *Streptomyces albus* ssp. *chlorinus*, as a source of new and undiscovered biomolecules, was demonstrated [17–19]. Genome mining of this strain (GenBank accession number VJOK00000000) using antiSMASH software revealed a further uncharacterized biosynthetic gene cluster encoding a putative nonribosomal peptide synthetase (NRPS) [26]. A search in the previously constructed genomic BAC library of *S. albus* ssp. *chlorinus* uncovered a BAC clone 2O18 with the cloned DNA fragment covering the entire NRPS gene cluster. The BAC 2O18 was transferred into the heterologous host *Streptomyces albus* Del14 by conjugation [28]. The obtained exconjugant strain *S. albus* 2O18 as well as the control strain *S. albus* Del14 were cultivated in the production medium DNPM and secondary metabolites were extracted with *n*-butanol. The extracts were analyzed using high resolution LC-MS. The analysis of the peak profile of *S. albus* 2O18 revealed the presence of a new mass peak with an m/z of 430.229 eluting at hydrophilic conditions at the very front of the chromatogram (Figure 2). The identified peak was not detected in the extract of the reference strain without the BAC. The calculated molecular mass of 429.221 [+/−5 ppm] of the identified compound was used for a search in the DNP database of natural products. This survey did not lead to any match with already registered metabolites from bacteria, implying that the structure of the molecule might be new. In order to gain insights into the structure of the identified compound, its purification for NMR analysis was carried out.
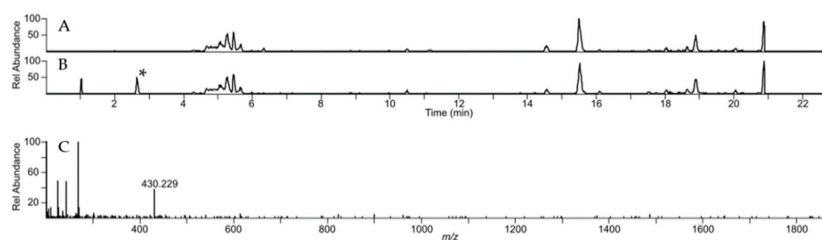
Figure 2. HPLC-MS analysis of bonsecamin production. A- and B-extracted ion chromatograms (430.5 ± 0.5 Da) of crude extracts of *S. albus* 2O18 and *S. albus* Del14, respectively. The new peak observed in the extract of *S. albus* 2O18 is marked with an asterisk (*). C–Mass spectrum of the new peak observed in the extract of *S. albus* 2O18.

3.2. Purification and Structure Elucidation

To obtain the identified compound in an amount sufficient for structure elucidation, *S. albus* 2O18 was cultivated in 10L of DNPM medium. The metabolites were extracted with *n*-butanol. The compound was purified from the crude extract using normal-phase, size-exclusion and reversed-phase chromatography. Only a low submilligram amount of pure compound was obtained as a white powder and physically characterized. The optical rotation was determined as $[\alpha]_D^{20}$ −13.5 (c 0.12, MeOH) and the measured λ_{max} (log ε) of the compound was 198 nm (2.11) in 11% ACN/H$_2$O + 0.1% formic acid. The dried isolate proved stable for > 1 year when stored at −20 °C and did not show signs of degradation when dissolved in organic solvents (methanol/DMSO). It was further used for structure elucidation by NMR experiments, MS/MS fragmentation and FDLA derivatization (Table 1, Figure S1—S8).

Table 1. NMR data of bonsecamin in DMSO-*d*6.

Pos.		δC/δN Type	δH, mult. (J in Hz)	COSY	Key HMBC (H-)
2,3-DABA	1	170.5 *, C			
	2	58.53, CH	3.96, dd (7.5, 2.5)	3, 5	1
	3	56.23, CH	3.23, m	2, 4	23
	4	13.41, CH3	0.52, d (6.3)	3	2
	5	122.4, NH	7.06, d (7.5)	2	1, 6
Val	6	169.6 *, C			
	7	58.40, CH	4.01, t (9.1)	8, 11	6, 12
	8	30.24, CH	1.96, m	7, 9, 10	
	9	18.39, CH3	0.83, d (7.0)	8	7
	10	19.44, CH3	0.84, d (7.0)	8	7
	11	114.5, NH	7.84, d (9.1)	7	12
Ala	12	172.2 *, C			
	13	50.67, CH	4.21, dq (8.6, 7.5)	14, 15	12, 16
	14	17.46, CH3	1.32, d (7.5)	13	12
	15	120.5, NH	8.14, d (8.6)	13	16
Ala	16	169.8 *, C			
	17	50.56, CH	4.08, quin (7.2)	18, 19	16, 20
	18	17.23, CH3	1.24, d (7.2)	17	16
	19	121.7, NH	7.87, d (7.2)	17	16
Ser	20	172.9 *, C			
	21	62.38, CH	2.80, m	22, 24	3, 20
	22	61.04, CH2	3.44, m	21, 23	
			3.57, m		
	23	—-, OH	5.01, m	22	
	24	48.3, NH	2.31, d (5.5)	21	2, 4, 20, 21, 22

* Not visible in the 13C NMR spectrum. The value was taken from the HMBC.

The molecular formula was calculated as C$_{18}$H$_{31}$N$_5$O$_7$ with 6 degrees of unsaturation corresponding to the monoisotopic mass of 429.222 Da. The analysis of ^{1}H NMR in DMSO-*d*6 revealed four doublet NH signals at δH 7.06, 7.84, 7.87 and 8.14, indicating four peptide bonds. The measurement of ^{15}N-HSQC confirmed this assumption by showing correlations to δN 114.5, 120.5, 121.7 and 122.4, and revealed an additional NH group at δH 2.31 and δN 48.3 suggesting a secondary amine (Figure S6). Analysis of ^{1}H and ^{13}C NMR, ^{1}H-^{1}H COSY and edited-HSQC revealed five amino acids corresponding to valine, two alanines, serine and threonine. The peptide was assigned by long-range HMBC correlations leading to the sequence Ser-Ala-Ala-Val-Thr. The remaining degree of unsaturation indicated a cyclic structure with a ring closure between serine and threonine via the aforementioned amine. This could be concluded from the altered chemical shifts of serine CH-α (δC 62.4, δH 2.80) and threonine CH-β (δC 56.2, δH 3.23). A final proof was provided by key correlations in the ^{1}H-^{1}H COSY and HMBC spectra. The amine proton at δH 2.31 was determined by a COSY correlation adjacent to CH-α of serine (δH 2.80) and showed strong HMBC correlations to CH-α and CH$_3$ of threonine, but not to the carboxyl group. In addition, a

COSY correlation of the hydroxyl group of serine (δH 5.01) to CH_2 β (δH 3.44 and 3.57) was observed, ruling out ring closure by an ester or ether.

In summary, ring closure between serine and CH-α of threonine was established. This will henceforth be referred to as 2, 3-diaminobutanoic acid (DABA). The missing COSY correlation between CH-β of DABA and the amine could be due to a dihedral angle of 90° between δH 2.31 (NH, Ser) and δH 3.23 (CH β, DABA) resulting in a very small 3J coupling constant. The structure was confirmed by MS/MS fragmentation by showing a- and x-ions patterns commonly observed for cyclic structures (Figure S8). The spectral data of bonsecamin are shown in Table 1.

The absolute configuration was determined by Marfey's method. The peptide was treated with 6N hydrochloric acid at 110°C for 1 h. The hydrolysate and the amino acid standards were derivatized with D- and L-FDLA and analyzed by LC-MS (Figure S7). Alanine and valine were determined in an L-configuration and showed the expected ratio of 2:1.

Serine was probably converted to 2,3-hydroxybutanoic acid and did not react with L-FDLA. Thus, its configuration could not be determined. The configuration of DABA was elucidated by the relative method using derivatization with D- and L-FDLA. When derivatized with D-FDLA, DABA showed a shorter retention time compared to L-FDLA, thus it had D-configuration. The absolute structure of bonsecamin is shown in Figure 3.
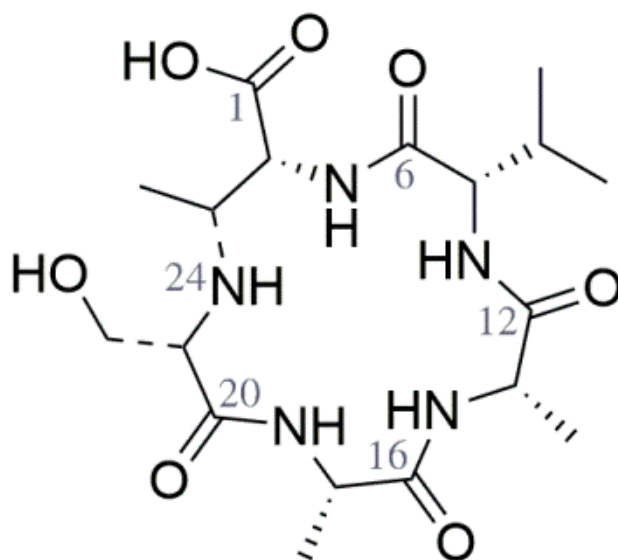

Figure 3. The structure of isolated bonsecamin.

3.3. Determination of the Minimal Gene Cluster

A sequence analysis of the 35 kb DNA fragment cloned in the BAC 2O18 revealed 28 putative genes (Figure 4). AntiSMASH analysis did not predict any significant homology between the expressed cluster and any other already characterized gene cluster. However, this analysis revealed that a DNA sequence highly similar to genes 8 to 14 is present in a number of *Streptomyces* strains. This implies indirectly that genes 8 to 14 cloned in the BAC 2O18 might correspond to the bonsecamin biosynthetic cluster. Genes 8 to 14 are further assigned as *bonA* to *bonG*. The genes *bonB*, *bon*C and *bonF* encode elements of an NRPS (Table S2). The gene *bonG* shows similarity to an amino acid ligase. The genes *bonD*, *bonE* and *bonA* encode two putative oxidoreductases and a transporter protein, respectively (Table S2).
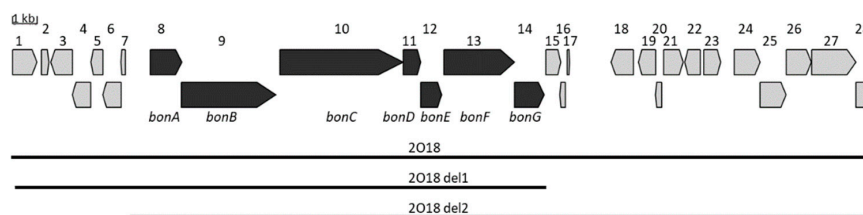
Figure 4. Fragment of the *Streptomyces albus* ssp. *chlorinus* NRRL B-24108 chromosome cloned in BAC 2O18. The genes putatively involved in bonsecamin biosynthesis are highlighted in dark grey. The black bars indicate the chromosomal fragments cloned in BACs 2O18, 2O18_del1 and 2O18_del2 [29].

To prove the minimal set of genes responsible for bonsecamin synthesis, a set of gene deletion experiments were performed. The genes 15 to 28 which were predicted not to be involved in bonsecamin biosynthesis were deleted in the BAC 2O18 using RedET. The constructed recombinant BAC 2O18_del1 was transferred into the host strain *S. albus* Del14 and the obtained strain *S. albus* 2O18_del1 was tested for bonsecamin production. The results of the HPLC-MS analysis demonstrated that the production of bonsecamin was not affected in the *S. albus* 2O18_del1 strain (Figure S9). This indicates that the genes 15–28 downstream of the *bonG* gene were not involved in bonsecamin production. Since *bonG* encoded a putative amino acid ligase and built an operon with the NRPS gene *bonF*, *bonG* was regarded as the last gene of the biosynthetic cluster (Figure 4). In order to determine the left border of the cluster, genes 1–7 upstream of the *bonA* gene were deleted in BAC 2O18. The constructed BAC 2O18_del2 was transferred into the heterologous host *S. albus* Del14 and the obtained exconjugant strain was analyzed for bonsecamin production. No difference in bonsecamin production between the strains *S. albus* 2O18 and *S. albus* 2O18_del2 was detected (Figure S9). This indicated that the deleted genes 1 to 7 were not involved in the production of bonsecamin. The gene *bonA* encoding a putative transporter was regarded as the first gene involved in bonsecamin production and *bonA* belongs to the same operon as the NRPS-encoding *bonB,* which further supports the assumption that the *bonA* gene constitutes the left border of the cluster. In general, our data indicated that the bonsecamin biosynthetic cluster encompassed the genes *bonA* to *bonG.*

3.4. Biosynthesis of Bonsecamin

Bonsecamin is a novel cyclic pentapeptide. The structure of bonsecamin implies that the compound might be synthesized through a linear peptide precursor–Ser-Ala-Ala-Val-Thr. The amino acid residues within the precursor are linked via conventional amide bonds. To generate the mature bonsecamin the linear precursor peptide likely undergoes intramolecular dehydrative cyclization. During this modification step the side chain of the threonine was joined with the free amino group of the serine residue.

The predicted minimal gene cluster for bonsecamin production consists of seven open reading frames from *bonA* to *bonG*. The *bonA* gene encodes a putative transporter. The genes *bonD* and *bonE* encode putative dehydrogenases and are presumably involved in the tailoring steps of the bonsecamin biosynthesis. The remaining four genes *bonB*, *bonC*, *bonF* and *bonG*, which encode putative elements of NRPS and an alanine ligase, might be involved in the biosynthesis of the linear bonsecamin precursor. The analysis of the NRPS genes revealed that *bonF* encodes A and PCP domains; *bonB* encodes A, PCP and C domains; and *bonC* encodes A, PCP, C and TE domains (Figure 5). This domain organization indicates that only three of five amino acid residues of bonsecamin were incorporated by the encoded NRPS. The A domains encoded by *bonC*, *bonB* and *bonF* were predicted to have substrate specificity towards the amino acids serine, valine and threonine, respectively. This prediction is in accordance with the amino acid composition of bonsecamin: Ser-Ala-Ala-Val-Thr. Two alanine residues are therefore expected to be incorporated into the bonsecamin precursor in an NRPS-independent manner. The putative alanine ligase encoded by the *bonG* gene might be responsible for the alanine incorporation.
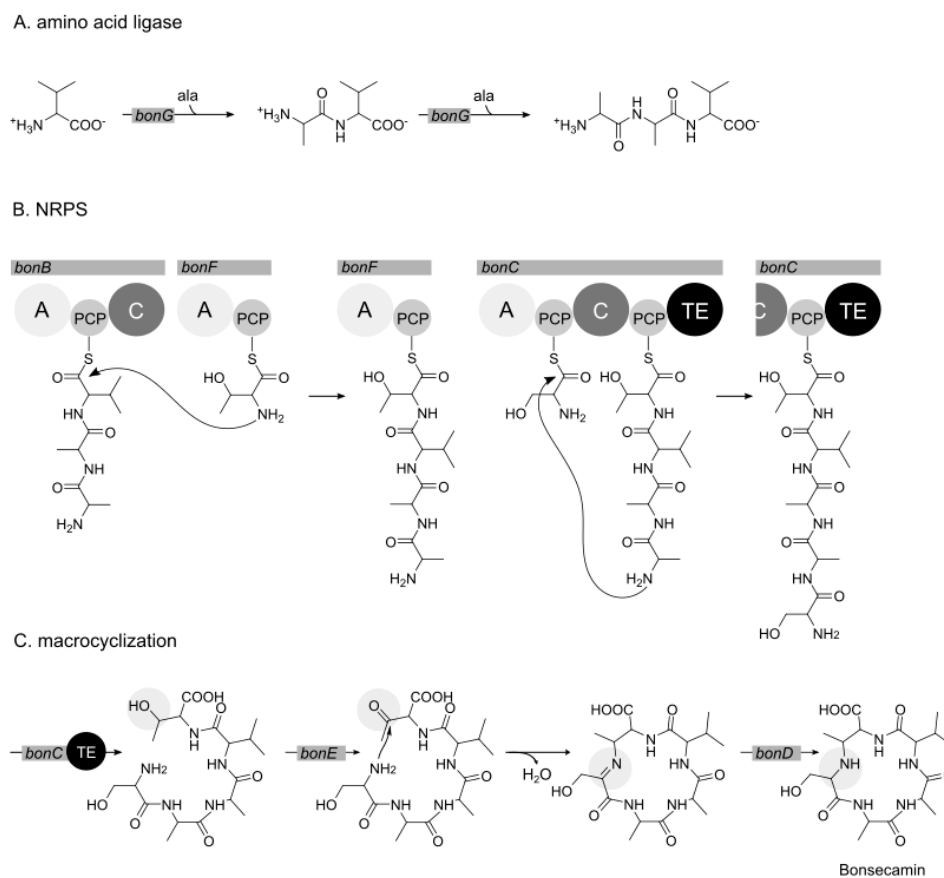
Figure 5. Proposed biosynthetic scheme for bonsecamin production. A. Formation of the Ala-Ala-Val tripeptide precursor catalyzed by the putative alanine ligase encoded by *bonG*. B. Conversion of the tripeptide intermediate into the linear pentapeptide precursor catalyzed by the NRPS encoded by *bonB*, *bonF* and *bonC*. C. Cyclization of the linear bonsecamin precursor catalyzed by the products of *bonE* and *bonD*.

Taking into consideration the structure of the compound, domain organization and the predicted substrate specificity of the NRPS, the following bonsecamin biosynthetic scheme is proposed. The first step in the biosynthesis is the assembly of the Ala-Ala-Val tripeptide catalyzed by a putative alanine ligase encoded by *bonG*. The enzyme uses valine as a starter substrate and carries out two rounds of alanine ligation. The participation of amino acid ligases in the biosynthesis of natural products was reported before [30]. Amino acid ligases usually attach small, non-polar amino acids to the amino group of the substrate by forming an amide bond. The enzymes are characterized as being quite specific for their extension unit (here alanine), but exhibit only little substrate specificity for the starter substrate, which is extended [31]. The relaxed substrate specificity might explain the attachment of the alanine residue to the valine in the first elongation step and then to the Ala-Val dipeptide in the second elongation step.

The conversion of the synthesized Ala-Ala-Val tripeptide into the linear bonsecamin precursor is proposed to be catalyzed by the NRPS genes encoded within the cluster. The attachment of the threonine residue to the tripeptide is supposedly catalyzed by the products of *bonB* and *bonF*, which together encode the starter module and the first elongation module of the bonsecamin NRPS. The tripeptide is probably activated by the A domain of *bonB*, presumably specific for valine, and loaded on the corresponding PCP domain (Figure 5). The activated tripeptide is then elongated with threonine which is activated by the A domain of *bonF* and bound to its PCP domain, leading to the Ala-Ala-Val-Thr tetrapeptide (Figure 5).

To finalize the assembly of the linear pentapeptide precursor, an amide bond between the amino group of the serine and the carboxyl group of the alanine within the tetrapeptide product needs to be formed. We propose that the product of *bonC* catalyzes this reaction. *BonC* contains a canonical starter module with A and PCP domains and a shortened termination module with C, PCP and TE domains. The termination module of *bonC* lacks an A domain. The A domain of the starter module is predicted to activate serine and to load it on the corresponding PCP domain. To elongate the serine residue with the tetrapeptide the latter needs to be loaded on the PCP domain of the termination module of *bonC*. Since the termination module of *bonC* does not contain any A domain, we propose that the tetrapeptide is transferred from the PCP domain of *bonF* on the PCP domain of the termination module of *bonC*. The mechanism of this assumed transesterification step remains elusive. The condensation of the serine residue with the tetrapeptide to yield the linear pentapeptide Ser-Ala-Ala-Val-Thr is catalyzed by the C domain of *bonC*. The linear pentapeptide precursor is proposed to be released through the action of a dedicated TE domain within the termination module of *bonC* (Figure 5).

The released linear pentapeptide precursor needs to be cyclized to yield mature bonsecamin. The cylization of bonsecamin occurs not through conventional amide bonds but through the secondary amine, which is formed between the side chain of the threonine and the amino group of the serine. We suppose that the conversion of the linear precursor into bonsecamin is catalyzed by the standalone enzymes encoded by *bonD* and *bonE*. Both enzymes were annotated to possess redox function. We assume that the first step towards cyclization is the oxidation of the hydroxyl group at the β carbon of the threonine, leading to a keto group with reactivity superior to the carboxylic acid carbon. This step might be catalyzed by the product of *bonE* which shows sequence similarity to 3-hydroxybutyrate dehydrogenase and 3-ketoacyl-ACP reductase–enzymes which catalyze similar reactions. Then the amino group of the serine residue reacts with the generated keto group under release of a water molecule to give an imine intermediate. Peptide macrocyclizations leading to imines are unusual, but have been described before [32–34]. The process of cyclization seems to happen spontaneously [32]. The generated imine group within the cyclized precursor is further reduced to a secondary amine leading to mature bonsecamin. We suppose that the final reduction step is catalyzed by the product of *bonD* which shows homology to the product of *lgrE*, involved in the reduction of linear gramicidin precursors [35].

The proposed scheme for bonsecamin biosynthesis includes a number of unusual enzymatic steps. In order to validate the proposed biosynthetic pathway, the inactivation of the *bonE* gene as well as the deletion of the second PCP domain or TE domain within *bonC* were undertaken and led to complete cessation of bonsecamin production (Figure S9). No putative derivatives or precursors of the compound could be detected in the culture broth of the engineered strains.

4. Conclusions

In this article we reported the identification of the new compound bonsecamin after successful heterologous expression of a cryptic NRPS cluster of *S. albus* ssp. *chlorinus* NRRL B-24108. Bonsecamin is a cyclic pentapeptide with a secondary amine moiety formed between the side chain of a threonine and the amino group of a serine. The identified minimal set of biosynthetic genes indicated that bonsecamin was a result of the interplay between an amino acid ligase and a nonlinear NRPS.

with D- or L-FDLA and the amino acid (aa) references derivatized with L-FDLA. Figure S8: MS/MS fragmentation of bonsecamin. Figure S9: Production of bonsecamin in S. albus Del14 mutant after gene deletion experiments.

References

1. Mohammadipanah, F.; Matasyoh, J.; Hamedi, J.; Klenk, H.-P.; Laatsch, H. Persipeptides A and B, two cyclic peptides from Streptomyces sp. UTMC 1154. *Bioorg. Med. Chem.* **2012**, *20*, 335–339. [CrossRef] [PubMed]
2. Wang, X.; Shaaban, K.A.; Elshahawi, S.I.; Ponomareva, L.V.; Sunkara, M.; Copley, G.C.; Hower, J.C.; Morris, A.J.; Kharel, M.K.; Thorson, J.S. Mullinamides A and B, new cyclopeptides produced by the Ruth Mullins coal mine fire isolate Streptomyces sp. RM-27-46. *J. Antibiot. (Tokyo)* **2014**, *67*, 571–575. [CrossRef] [PubMed]
3. Xiang, W.-S.; Wang, J.-D.; Wang, X.-J.; Zhang, J. Bingchamides A and B, two novel cyclic pentapeptides from the Streptomyces bingchenggensis: Fermentation, isolation, structure elucidation and biological properties. *J. Antibiot. (Tokyo)* **2009**, *62*, 501–505. [CrossRef]
4. Ji, Z.; Wei, S.; Fan, L.; Wu, W. Three novel cyclic hexapeptides from Streptomyces alboflavus 313 and their antibacterial activity. *Eur. J. Med. Chem.* **2012**, *50*, 296–303. [CrossRef]
5. Umezawa, K.; Ikeda, Y.; Uchihata, Y.; Naganawa, H.; Kondo, S. Chloptosin, an apoptosis-inducing dimeric cyclohexapeptide produced by Streptomyces. *J. Org. Chem.* **2000**, *65*, 459–463. [CrossRef]
6. Song, Y.; Li, Q.; Liu, X.; Chen, Y.; Zhang, Y.; Sun, A.; Zhang, W.; Zhang, J.; Ju, J. Cyclic Hexapeptides from the Deep South China Sea-Derived Streptomyces scopuliridis SCSIO ZJ46 Active Against Pathogenic Gram-Positive Bacteria. *J. Nat. Prod.* **2014**, *77*, 1937–1941. [CrossRef]
7. Cai, G.; Napolitano, J.G.; McAlpine, J.B.; Wang, Y.; Jaki, B.U.; Suh, J.-W.; Yang, S.H.; Lee, I.-A.; Franzblau, S.G.; Pauli, G.F.; et al. Hytramycins V and I, Anti-*Mycobacterium tuberculosis* Hexapeptides from a *Streptomyces hygroscopicus* Strain. *J. Nat. Prod.* **2013**, *76*, 2009–2018. [CrossRef]
8. He, H.; Williamson, R.T.; Shen, B.; Graziani, E.I.; Yang, H.Y.; Sakya, S.M.; Petersen, P.J.; Carter, G.T. Mannopeptimycins, Novel Antibacterial Glycopeptides from Streptomyces h ygroscopicus, LL-AC98. *J. Am. Chem. Soc.* **2002**, *124*, 9729–9736. [CrossRef]
9. Dawson, R. The toxicology of microcystins. *Toxicon* **1998**, *36*, 953–962. [CrossRef]
10. Raymond, K.N.; Dertz, E.A.; Kim, S.S. Enterobactin: An archetype for microbial iron transport. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3584–3588. [CrossRef] [PubMed]
11. Singh, M.P.; Petersen, P.J.; Weiss, W.J.; Janso, J.E.; Luckman, S.W.; Lenoy, E.B.; Bradford, P.A.; Testa, R.T.; Greenstein, M. Mannopeptimycins, new cyclic glycopeptide antibiotics produced by *Streptomyces hygroscopicus* LL-AC98: Antibacterial and mechanistic activities, Antimicrob. *Agents Chemother.* **2003**, *47*, 62–69. [CrossRef] [PubMed]
12. Turan, T.; Karacay, O.; Tulunay, G.; Boran, N.; Koc, S.; Bozok, S.; Kose, M.F. Results with EMA/CO (etoposide, methotrexate, actinomycin D, cyclophosphamide, vincristine) chemotherapy in gestational trophoblastic neoplasia. *Int. J. Gynecol. Cancer* **2006**, *16*, 1432–1438. [CrossRef] [PubMed]
13. Ueda, H.; Manda, T.; Matsumoto, S.; Mukumoto, S.; Nishigaki, F.; Kawamura, I.; Shimomura, K. FR901228, a novel antitumor bicyclic depsipeptide produced by chromobacterium violaceum no. 968. III. Antitumor activities on experimental tumors in mice. *J. Antibiot. (Tokyo)* **1994**, *47*, 315–323. [CrossRef]
14. Borel, J.F.; Feurer, C.; Gubler, H.U.; Stähelin, H. Biological effects of cyclosporin A: A new antilymphocytic agent. *Agents Actions* **1976**, *6*, 468–475. [CrossRef]
15. Kolata, G. FDA Speeds Approval of Cyclosporin. *Science* **1983**, *221*, 1273. [CrossRef]
16. Sieber, S.A.; Marahiel, M.A. Learning from Nature's Drug Factories: Nonribosomal Synthesisof MacrocyclicPeptides. *J. Bacteriol.* **2003**, *185*, 7036–7043. [CrossRef]

17. Rodríguez Estévez, M.; Gummerlich, N.; Myronovskyi, M.; Zapp, J.; Luzhetskyy, A. Benzanthric Acid, a Novel Metabolite From Streptomyces albus Del14 Expressing the Nybomycin Gene Cluster. *Front. Chem.* **2020**, *7*, 896. [CrossRef] [PubMed]

18. Estévez, M.R.; Myronovskyi, M.; Rosenkränzer, B.; Paululat, T.; Petzke, L.; Ristau, J.; Luzhetskyy, A. Novel fredericamycin variant overproduced by a streptomycin-resistant streptomyces albus subsp. *Chlorinus Strain. Mar. Drugs* **2020**, *18*, 284. [CrossRef]

19. Lasch, C.; Stierhof, M.; Estévez, M.R.; Myronovskyi, M.; Zapp, J.; Luzhetskyy, A. Dudomycins: New Secondary Metabolites Produced after Heterologous Expression of an Nrps Cluster from Streptomyces albus ssp. Chlorinus Nrrl B-24108. *Microorganisms* **2020**, *8*, 1800. [CrossRef]

20. Myronovskyi, M.; Rosenkränzer, B.; Stierhof, M.; Petzke, L.; Seiser, T.; Luzhetskyy, A. Identification and Heterologous Expression of the Albucidin Gene Cluster from the Marine Strain Streptomyces Albus Subsp. Chlorinus NRRL B-24108. *Microorganisms* **2020**, *8*, 237. [CrossRef] [PubMed]

21. Rodríguez Estévez, M.; Myronovskyi, M.; Gummerlich, N.; Nadmid, S.; Luzhetskyy, A. Heterologous Expression of the Nybomycin Gene Cluster from the Marine Strain Streptomyces albus subsp. chlorinus NRRL B-24108. *Mar. Drugs* **2018**, *16*, 435. [CrossRef]

22. Green, M.R.; Sambrook, J. *Molecular Cloning: A Laboratory Manual*, 4th ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2012.

23. Kieser, T.; Bibb, M.J.; Buttner, M.J.; Chater, K.F.; Hopwood, D.A. *Practical Streptomyces Genetics. A Laboratory Manual.*; John Innes Foundation: Norwich, UK, 2000.

24. Rebets, Y.; Kormanec, J.; Lutzhetskyy, A.; Bernaerts, K.; Anné, J. Cloning and expression of metagenomic dna in streptomyces lividans and subsequent fermentation for optimized production. *Methods Mol. Biol.* **2017**, *1539*, 99–144.

25. Court, D.L.; Sawitzke, J.A.; Thomason, L.C. Genetic Engineering Using Homologous Recombination. *Annu. Rev. Genet.* **2002**, *36*, 361–388. [CrossRef]

26. Medema, M.H.; Blin, K.; Cimermancic, P.; De Jager, V.; Zakrzewski, P.; Fischbach, M.A.; Weber, T.; Takano, E.; Breitling, R. antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **2011**, *39*, W339–W346. [CrossRef] [PubMed]

27. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef] [PubMed]

28. Myronovskyi, M.; Rosenkränzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A. Generation of a cluster-free Streptomyces albus chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab. Eng.* **2018**, *49*, 316–324. [CrossRef]

29. Harrison, K.J.; Crécy-Lagard, V. de; Zallot, R. Gene Graphics: A genomic neighborhood data visualization web application. *Bioinformatics* **2018**, *34*, 1406–1408. [CrossRef] [PubMed]

30. Parker, J.B.; Walsh, C.T. Action and Timing of BacC and BacD in the Late Stages of Biosynthesis of the Dipeptide Antibiotic Bacilysin. *Biochemistry* **2013**, *52*, 889–901. [CrossRef] [PubMed]

31. Antle, V.D.; Liu, D.; McKellar, B.R.; Caperelli, C.A.; Hua, M.; Vince, R. Substrate Specificity of Glycinamide Ribonucleotide Synthetase from Chicken Liver. *J. Biol. Chem.* **1996**, *271*, 8192–8195. [CrossRef] [PubMed]

32. Kopp, F.; Mahlert, C.; Grünewald, J.; Marahiel, M.A. Peptide Macrocyclization: The Reductase of the Nostocyclopeptide Synthetase Triggers the Self-Assembly of a Macrocyclic Imine. *J. Am. Chem. Soc.* **2006**, *128*, 16478–16479. [CrossRef]

33. Evans, B.S.; Ntai, I.; Chen, Y.; Robinson, S.J.; Kelleher, N.L. Proteomics-Based Discovery of Koranimine, a Cyclic Imine Natural Product. *J. Am. Chem. Soc.* **2011**, *133*, 7316–7319. [CrossRef] [PubMed]

34. Zipperer, A.; Konnerth, M.C.; Laux, C.; Berscheid, A.; Janek, D.; Weidenmaier, C.; Burian, M.; Schilling, N.A.; Slavetinsky, C.; Marschal, M.; et al. Human commensals producing a novel antibiotic impair pathogen colonization. *Nature* **2016**, *535*, 511–516. [CrossRef] [PubMed]

35. Schracke, N.; Linne, U.; Mahlert, C.; Marahiel, M.A. Synthesis of Linear Gramicidin Requires the Cooperation of Two Independent Reductases†. *Biochemistry* **2005**, *44*, 8507–8513. [CrossRef] [PubMed]

Supplementary

Bonsecamin: A New Cyclic Pentapeptide Discovered through Heterologous Expression of a Cryptic Gene Cluster

Constanze Lasch [1], Marc Stierhof [1], Marta Rodríguez Estévez [1], Maksym Myronovskyi [1], Josef Zapp [2] and Andriy Luzhetskyy [1,3,*]

[1] Pharmaceutical Biotechnology, Saarland University, Saarbruecken, Germany; constanze.lasch@uni-saarland.de (C.L.), m.stierhof@t-online.de (M.S.), marta.rodriguezestevez@uni-saarland.de (M.R.), maksym.myronovskyi@uni-saarland.de (M.M.), a.luzhetskyy@mx.uni-saarland.de (A.L.)

[2] Pharmaceutical Biology, Saarland University, Saarbruecken, Germany; j.zapp@mx.uni-saarland.de (J.Z.)

[3] Helmholtz Institute for Pharmaceutical Research Saarland, Saarbruecken, Germany

[*] Correspondence: a.luzhetskyy@mx.uni-saarland.de; +49 681 302 70200 (A.L.)

Supplementary.

Table S1. Strains, BACs, plasmids and primers used in this work.

Material	Purpose
A. Bacterial strains	
Streptomyces albus Del14	heterologous host [1]
Escherichia coli GB05 RedCC	cloning host [Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS)]
Escherichia coli ET12567 pUB307	alternate host intergeneric conjugation [2]
B. BACs	
2O18	heterologous expression of NRPS cluster
2O18_del1	determination downstream border of NRPS cluster
2O18_del2	determination upstream border of NRPS cluster
2O18_delKR_delbla	single gene inactivation
2O18_delPCP_delbla	single gene inactivation
2O18_delTE2_delbla	single gene inactivation
C. Plasmids	
pUC19	ampicillin resistance marker
D. PCR primer Red/ET	
20200815_1_fw [2O18_del1]	TAGTCCAGCGTCATCAGCGGGGCGTCCGAGGCACTGCGGACCACGAGGCGCGTCAGGTGGCACTTTTCG
20200815_1_rev [2O18_del1]	TCCGACGGCGGGCGGCCCCGCACTAGGCTCGCCGCCATGACGGACGTCGACTTTTCTACGGGGTCTGAC
20210315_1_fw [2O18_del2]	CTATCGTCGCCACGCCTTGGTGCACGGGAAATCCGGTGTGATGCCGGTGCCGTCAGGTGGCACTTTCG
20200815_2_rev [2O18_del2]	CACTGGATGCCCAGGCAGGGGGTACGCAGCATGACCGAGGAGGACGCGGCCTTTTCTACGGGGTCTGAC
20201217_1_fw [2O18_delKR_delbla]	GCTGGTGAACCCGCCGTCGACGGTGACCGTGGAGCCGGTCACCTGGCGGGAGTTTAAACCGTCAGGTGGCACTTTTCG

|	
20201217_1_rev [2O18_delKR_delbla]	GGAGTGCTCACCGCGGCGCCGCTCGCGGGCAAGGCCGCCGTCATCACGGGTTTAAACGACTTTTCTACGGGGTCTGAC
20201217_2_fw [2O18_delPCP_delbla]	GGCCAGGGCGGCCAGTTCGCCCAGCCGCGGGATGCGGGTGAGGTCGGTGAAGTTTAAACCGTCAGGTGGCACTTTTCG
20201217_2_rev [2O18_delPCP_delbla]	CGCGCGGTCTGGCAGCAGATCCTGGGGCTGACGGCGGAGGAGATCGGTGGTTTAAACGACTTTTCTACGGGGTCTGAC
20201217_4_fw [2O18_delTE2_delbla]	CAGTTCCGCGGTCGCGGCCGGATTGCCGCGCACGAAGTAGTGGCCGCCCGAGTTTAAACCGTCAGGTGGCACTTTTCG
20201217_4_rev [2O18_delTE2_delbla]	TGCGTGCCGTATCCGTGCGGGCACCCGGTCAACTTCAAACCGCTGGCCGGTTTAAACGACTTTTCTACGGGGTCTGAC
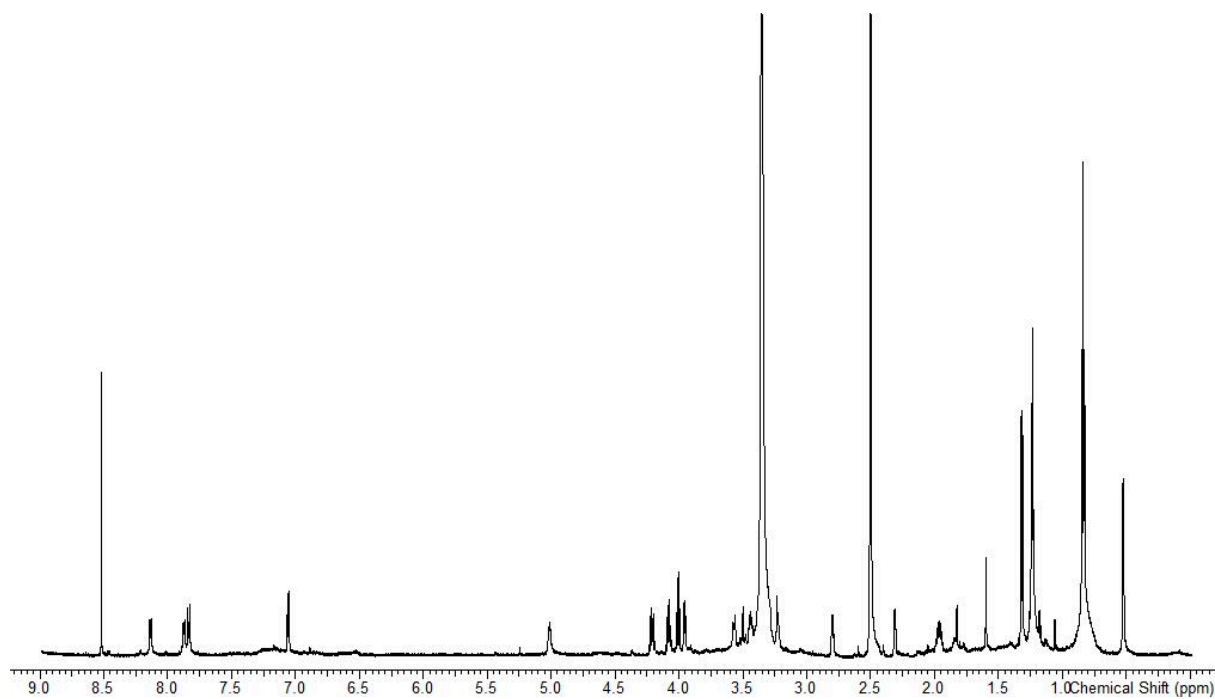
Figure S1: ^{1}H NMR spectrum (700 MHz, DMSO-*d*6) of bonsecamin.

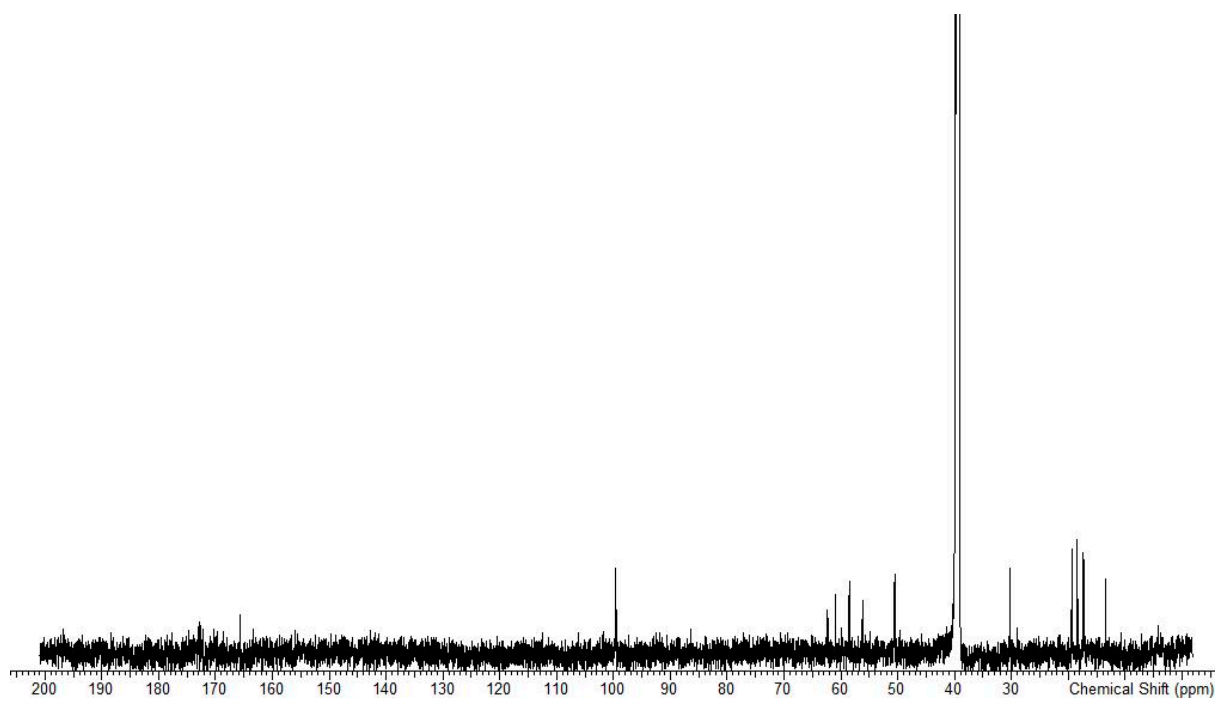**Figure S2:** ^{13}C NMR spectrum (700 MHz, DMSO-*d*6) of bonsecamin.
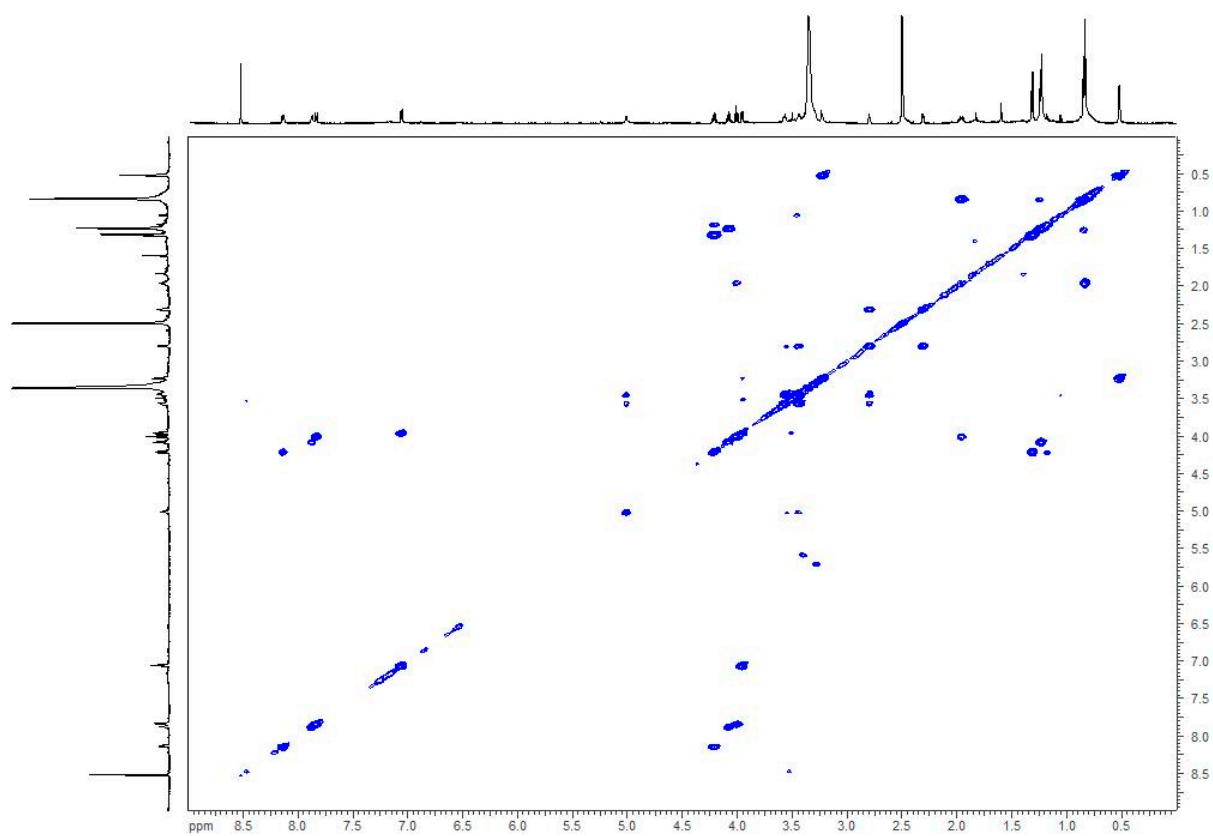
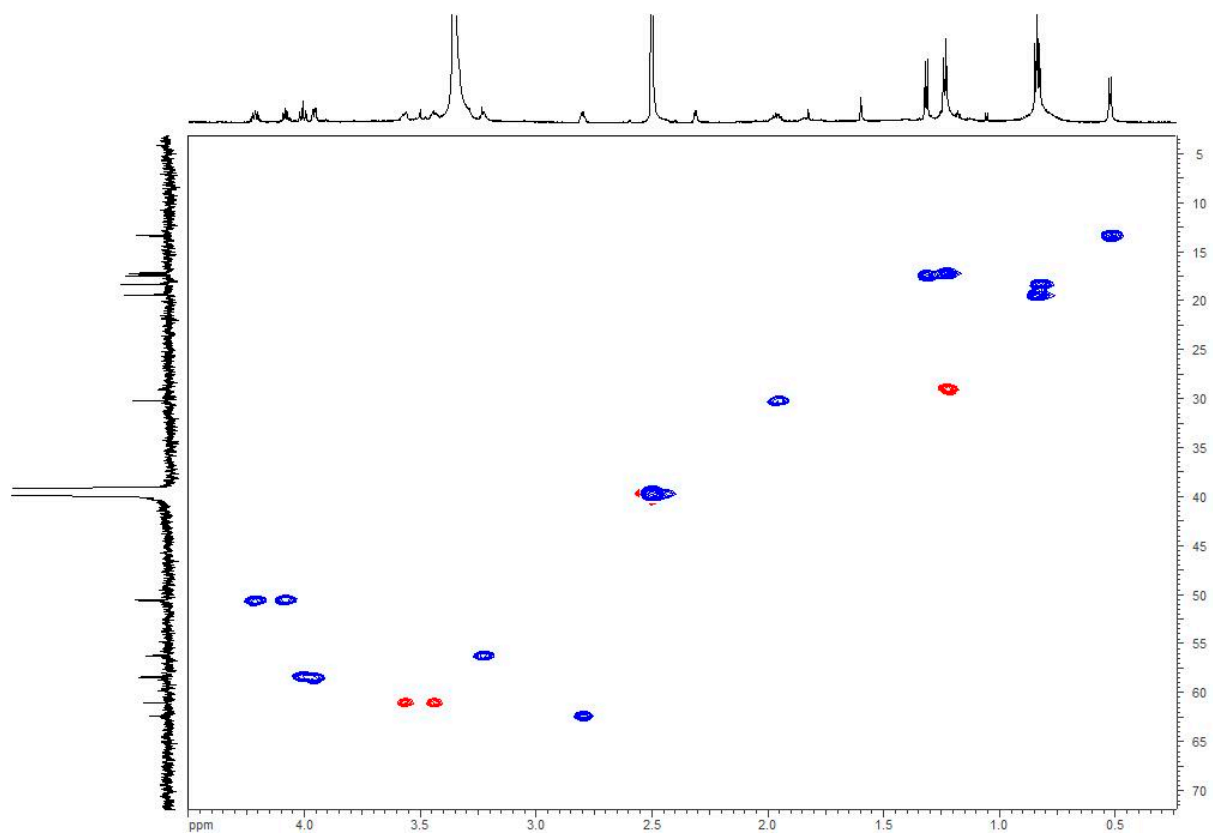**Figure S3:** ¹H-¹H COSY spectrum (700 MHz, DMSO-*d*6) of bonsecamin.

Figure S4: Edited-HSQC spectrum (700 MHz, DMSO-*d*6) of bonsecamin.
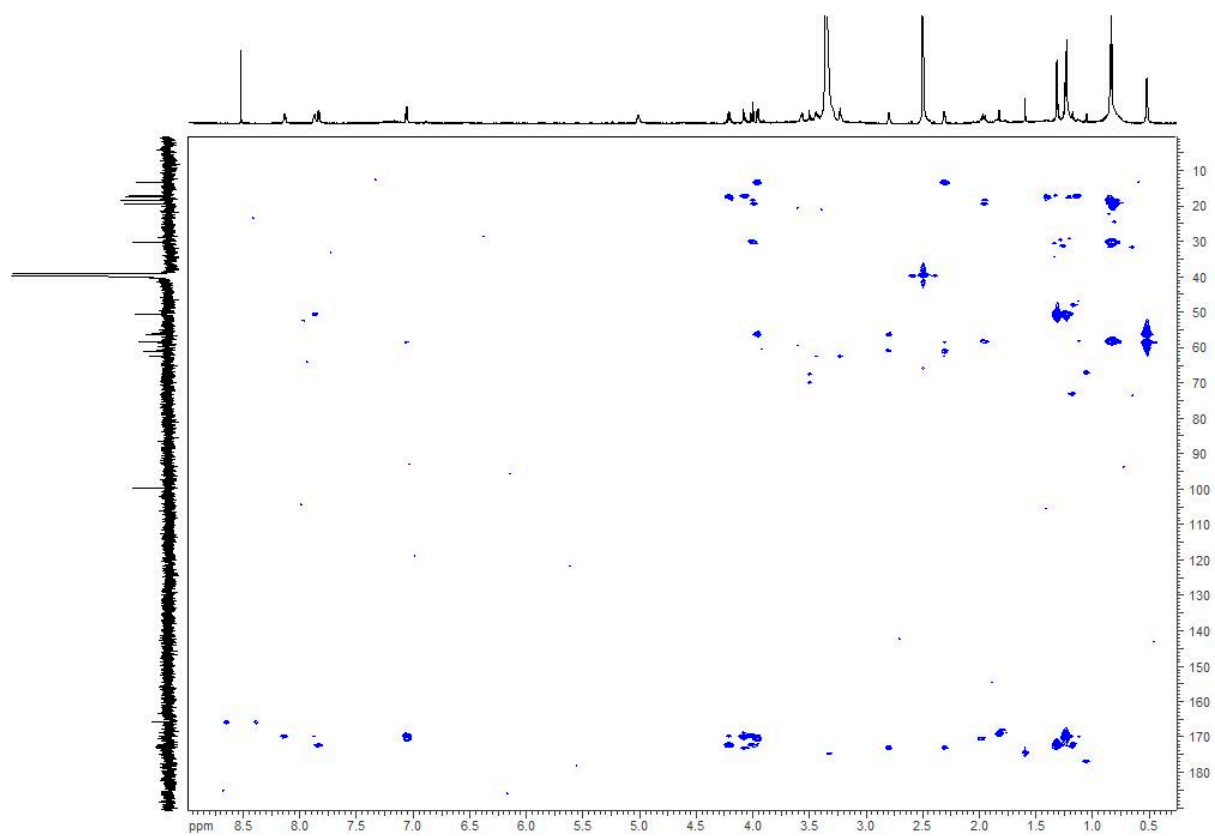
Figure S5: HMBC spectrum (700 MHz, DMSO-*d6*) of bonsecamin.
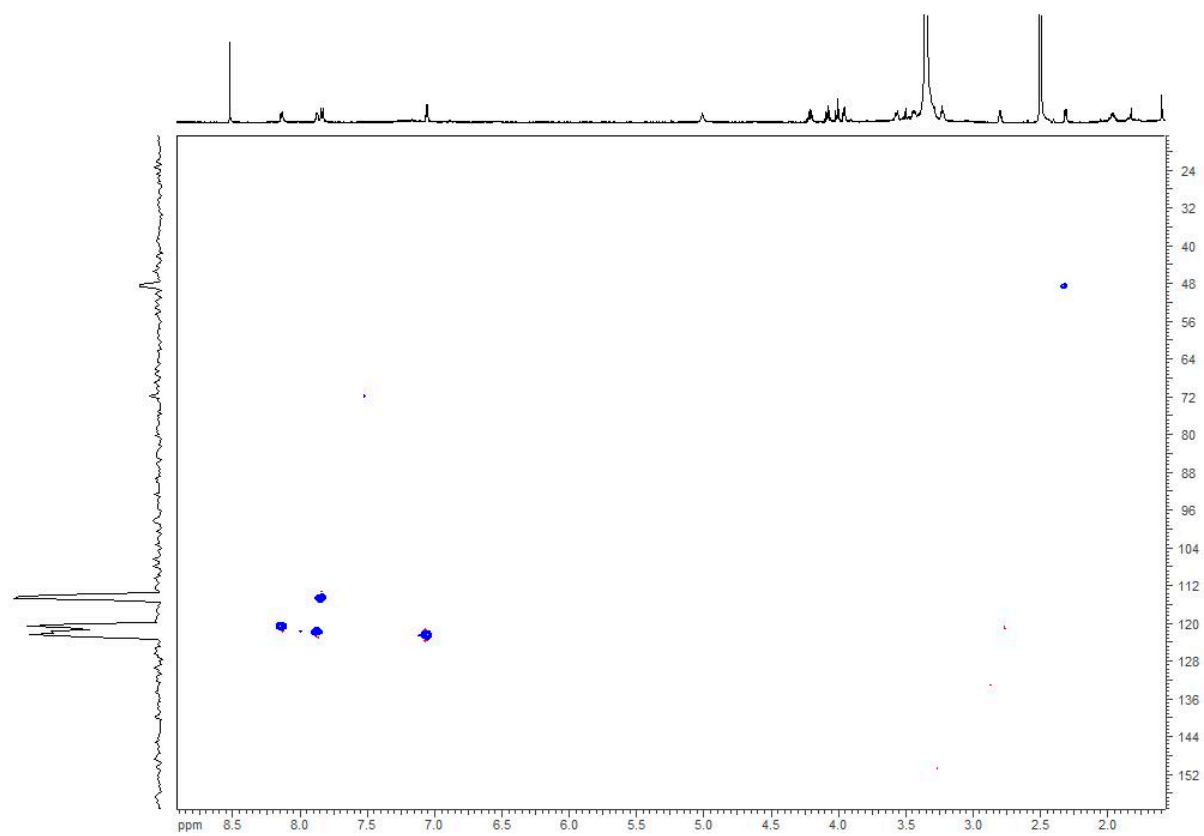
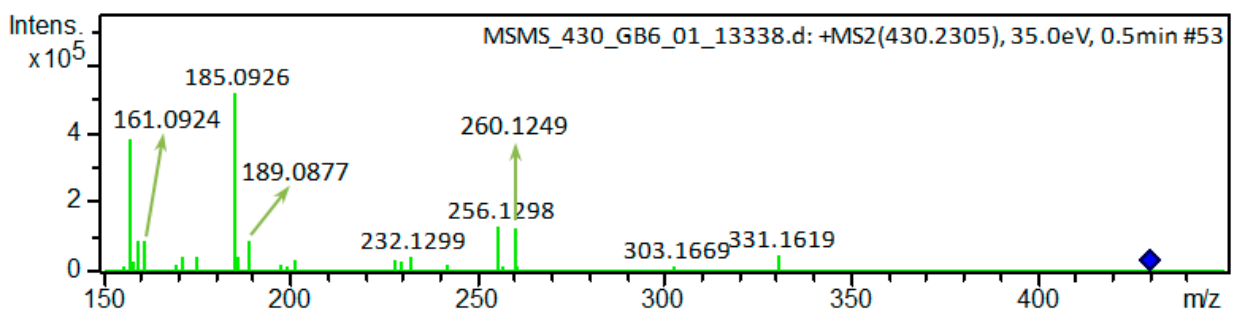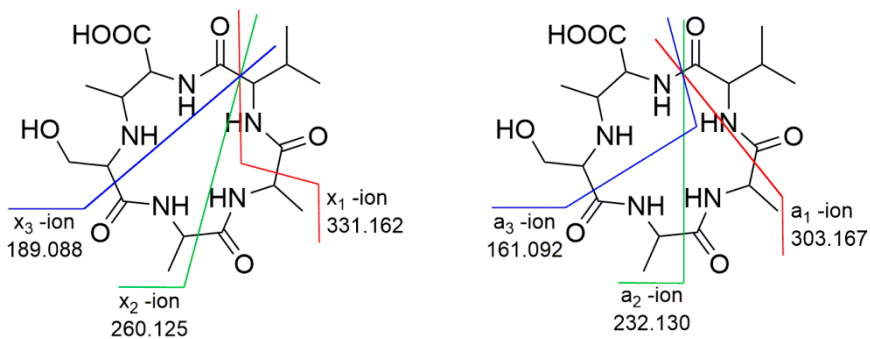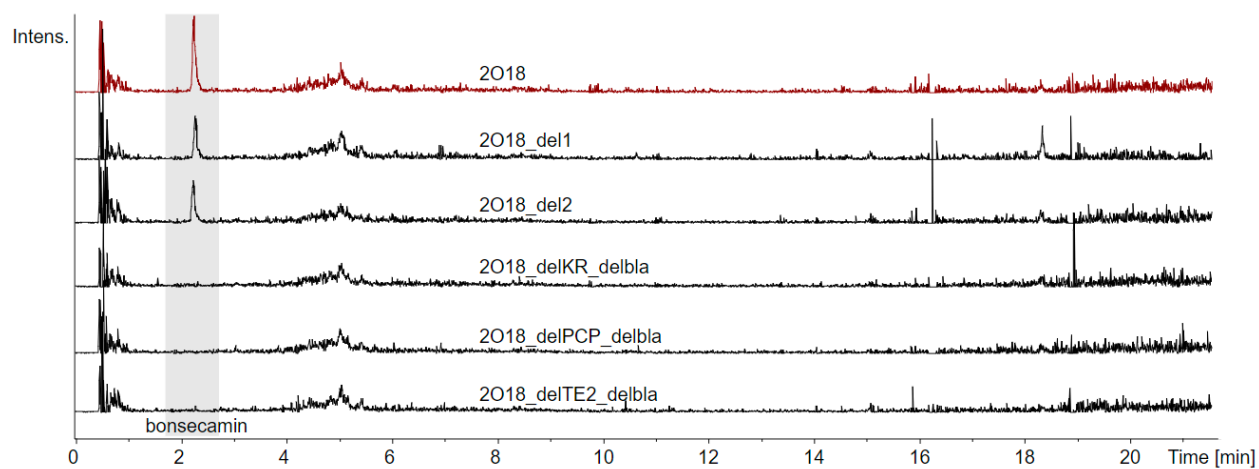**Figure S6:** ¹⁵N-HSQC spectrum (700 MHz, DMSO-*d6*) of bonsecamin.
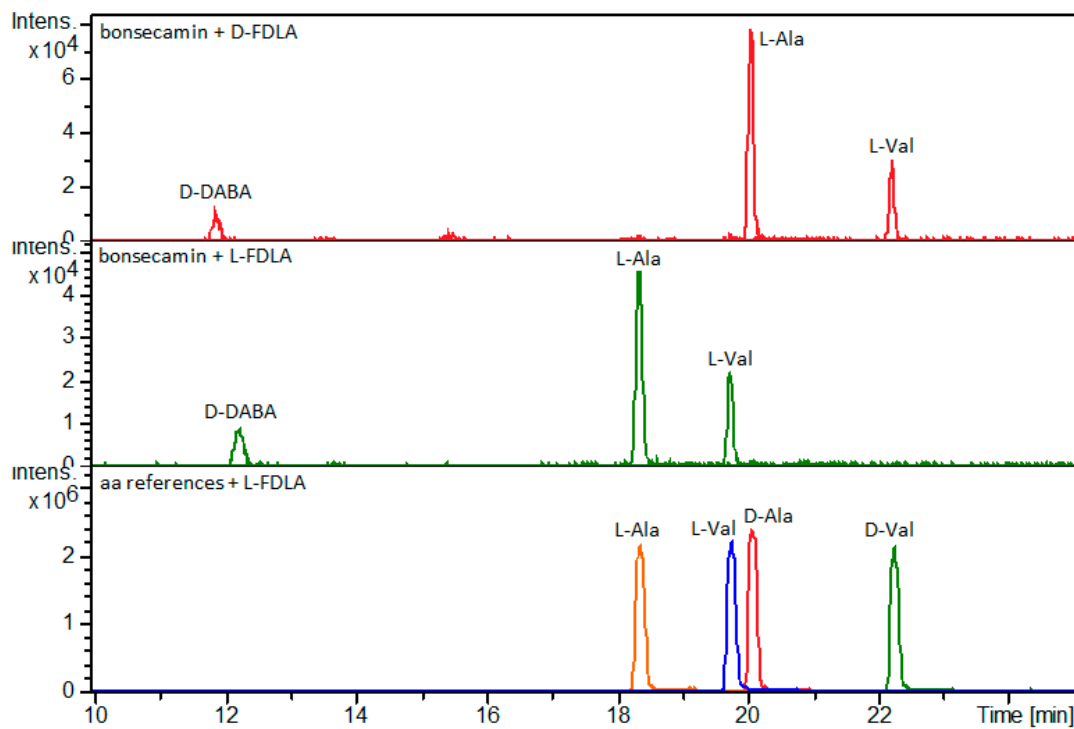
Figure S7: LC-MS chromatograms of hydrolyzed bonsecamin derivatized with D- or L-FDLA and the amino acid (aa) references derivatized with L-FDLA.

Figure S8: MS/MS fragmentation of bonsecamin.

Figure S9. Production of bonsecamin in *S. albus* Del14 mutant after gene deletion experiments. EIC extracted for masses [430-431].

Table S2. Putative products of the genes in the DNA fragment encoding bonsecamin production.

gene #	locus tag	putative product
1	SACHL_05130	Catalase
2	SACHL_05120	hypothetical protein
3	SACHL_05110	Cobalt import ATP-binding protein CbiO
4	SACHL_05100	Cobalt transport protein CbiQ
5	SACHL_05090	Cobalt transport protein CbiN
6	SACHL_05080	Cobalt transport protein CbiM
7	SACHL_05070	-
8 [*bonA*]	SACHL_05060	enterobactin exporter EntS
9 [*bonB*]	SACHL_05050	Tyrocidine synthase 3 - val
10 [*bonC*]	SACHL_05040	Tyrocidine synthase 3 - ser
11 [*bonD*]	SACHL_05030	Linear gramicidin dehydrogenase LgrE
12 [*bonE*]	SACHL_05020	(-)-trans-carveol dehydrogenase
13 [*bonF*]	SACHL_05010	Dimodular nonribosomal peptide synthase - thr
14 [*bonG*]	SACHL_05000	Alanine-anticapsin ligase BacD
15	SACHL_04990	hypothetical protein
16	SACHL_04980	hypothetical protein
17	SACHL_04970	-
18	SACHL_04960	P-aminobenzoate N-oxygenase AurF
19	SACHL_04950	hypothetical protein
20	SACHL_04940	hypothetical protein
21	SACHL_04930	hypothetical protein
22	SACHL_04920	CGNR zinc finger
23	SACHL_04910	(S)-2-haloacid dehalogenase
24	SACHL_04900	Putative phenylalanine aminotransferase
25	SACHL_04890	All-trans-nonaprenyl-diphosphate synthase (geranyl-diphosphate specific)
26	SACHL_04880	prenyltransferase
27	SACHL_04870	Squalene-hopene cyclase
28	SACHL_04860	2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase

[1] M. Myronovskyi, B. Rosenkränzer, S. Nadmid, P. Pujic, P. Normand, A. Luzhetskyy, Generation of a cluster-free Streptomyces albus chassis strains for improved heterologous expression of secondary metabolite clusters, Metab. Eng. 49 (2018) 316–324. https://doi.org/10.1016/j.ymben.2018.09.004.

[2] F. Flett, V. Mersinias, C.P. Smith, High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl DNA-restricting streptomycetes, FEMS Microbiol. Lett. 155 (2006) 223–229. https://doi.org/10.1111/j.1574-6968.1997.tb13882.x.

3   Conclusions

3.1   Dudomycin, bonsecamin and loseolamycin: Three novel molecules produced by optimized *Streptomyces* hosts

Silent actinobacterial gene clusters are a hidden, yet powerful source for the discovery of new natural products (NPs). In this work, three biosynthetic gene clusters (BGCs) of the two actinobacteria species *Streptomyces albus* ssp. *chlorinus* and *Micromonospora endolithica* were used as a source for BGCs for heterologous expression. The genome-minimized *Streptomyces* strains *S. albus* Del14 and *S. lividans* DelYA8 were chosen as chasses. For heterologous NP production a set of predefined conditions was used: the complex cultivation medium DNPM and minimal media had empirically proven to suit best for the production of numerous new molecules of different types and were selected for cultivation. Further, ethyl acetate and *n*-butanol served as extraction solvents to capture lipophilic as well as hydrophilic compounds. This approach enabled the successful heterologous production of the three new NPs dudomycin, bonsecamin and loseolamycin.
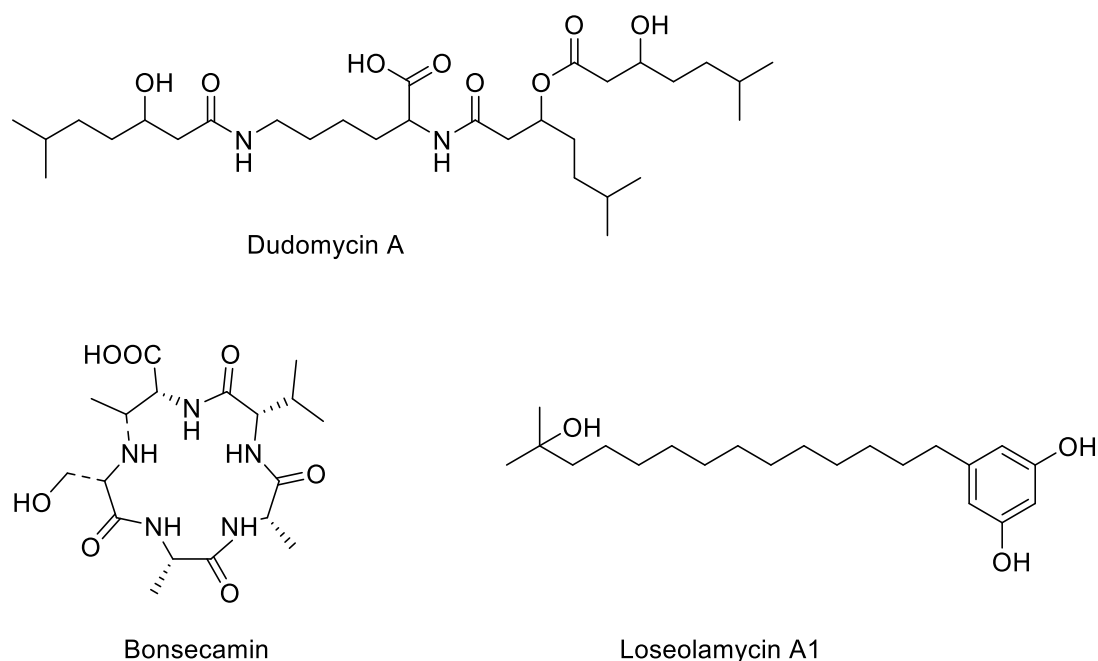

Figure 1. The structures of dudomycin, bonsecamin and loseolamycin.

Dudomycin [1]. In the past the majority of today's antibiotics were isolated from terrestrial microorganisms. Nowadays, the strain origin expanded to more unusual environments such as marine to avoid rediscoveries. Recently, heterologous expression of clusters from the marine strain *S. albus* ssp. *chlorinus* led to the identification of a remarkable number of new chemical structures and gene clusters in our lab [2–5]. In this work a dimodular non-ribosomal peptide synthase (NRPS) gene cluster was picked from the strain and transferred to our hosts which facilitated the production of the new compound dudomycin. Dudomycin has a core amino acid lysine and three medium chain fatty acids attached to it. A total number of eight dudomycin derivatives were identified differing in the length of the fatty acid residues. So far, no biological activity for those new structures was detected.

Bonsecamin [6]. Bonsecamin further illustrates the genomic wealth of the above described strain *S. albus* ssp. *chlorinus*. A gene cluster encoding a presumably unknown NRP was chosen for the heterologous expression. The optimization of the cultivation and extraction conditions was a prerequisite for the isolation of the new compound bonsecamin in amounts sufficient for structure elucidation. This cyclic pentapeptide harbours a structurally rare secondary amine function and is highly interesting from biosynthetic perspective as it involves a scarce amino acid ligase for its production. Due to the low production levels of bonsecamin no extended bioactivity testing could be performed up to now.

Loseolamycin [7]. *Micromonospora endolithica* is a representative of the non-*Streptomyces* actinobacteria strains and was expected to be a fruitful source for new NPs. A type III polyketide synthase (PKS) cluster was chosen for the expression in a heterologous host. However, production of a new molecule was only achieved after activation of the core biosynthetic gene through promoter insertion. Structurally, the discovered product loseolamycin is a resorcinol with hydroxylated alkyl chain. A total number of six loseolamycin derivatives were detected with variation in the length and saturation grade of the side chain and being of branched or linear type. Although of little structural complexity, loseolamycin indicated potency against gram positive bacteria and as herbicide in bioactivity tests.

3.2   Enabling the heterologous production of dudomycin, bonsecamin and loseolamycin

In the course of this thesis work three novel NPs were isolated from *Streptomyces* exconjugant strains. The following section discusses the role of heterologous expression in the discovery process of dudomycin, bonsecamin and loseolamycin.

Gene cluster mobilization. The availability of the sequenced genome data of the two strains *S. albus* ssp*. chlorinus* and *M. endolithica* was an important prerequisite for the implementation of this work. After thorough *in-silico* evaluation of all potential gene clusters of the two strains, the BGCs were prioritized according to their novelty and unusual features. Three clusters were selected for heterologous expression in the course of this project. Genomic libraries of both strains, *S. albus* ssp. *chlorinus* and *M. endolithica*, had been constructed in our lab previously and the BACs covering the entire selected clusters have been identified. Thus, no further cloning experiments were needed to capture our clusters of interest.

Surrogate hosts. The three clusters were transferred to two surrogate strains, *S. albus* Del14 and *S. lividans* DelYA8, both of which contain deletions of endogenous clusters and were previously constructed in our lab [8,9]. All three new compounds dudomycin, bonsecamin and loseolamycin were discovered in the extracts of the *S. albus* Del14 exconjugants. The low secondary metabolite background of the optimized host in combination with the use of minimal medium simplified enormously the identification and purification of the NPs, especially of the low abundant bonsecamin. None of the detected compounds could be identified in the extracts of the strains which were used as a source of the clusters. This indicates that the BGCs were silent in those strains and activated by heterologous expression.

Gene cluster engineering. A huge advantage of the heterologous expression approach for drug discovery is the possibility to link the newly identified NP to its encoding genes. While heterologous expression alone was sufficient for the activation of dudomycin and bonsecamin clusters and the production of the respective compounds, this approach did not lead to expression of the loseolamycin gene cluster from *M. endolithica*. Here, the heterologous expression in combination with targeted promoter insertion was successfully applied to induce the production of loseolamycin [10]. The promoter insertion step was required as wildtype and host belong to different genera of actinobacteria. In case of dudomycin and bonsecamin the minimal set of genes required for the production of

compounds was identified by gene inactivation experiments, where all genes up- and downstream from the presumed minimal cluster were deleted. In both cases this information helped to make a proposal on the molecules' biosynthesis. The data suggests that dudomycin is produced by an uncommon NRPS module which works in an iterative manner. The biosynthesis of bonsecamin is more complex. Besides the NRPS modules, an amino acid ligase and two standalone domains are involved to form the final pentapeptide which is reduced to an unusual secondary amine after ring closure. No further insights of the bonsecamin biosynthesis by the knock-out of single genes could be provided as none of the expected biosynthetic precursor could be identified in the extracts of the engineered strains.

3.3   Summary

Nature's creativity remains uncontested. By the discovery of the three new compounds dudomycin, bonsecamin and loseolamycin the relevance of heterologous expression in the discovery of new NPs was demonstrated. Heterologous expression not only shows the great undiscovered wealth inhered in actinobacterial genomes, but is also an effective tool for activation of the gene clusters. This method allowed switching on the production of dudomycin and bonsecamin which were not produced by the natural host strain under laboratory conditions. Drawbacks of heterologous expression such as low production yields of the new compounds cannot be neglected. But in combination with additional gene cluster engineering and the optimization of the fermentation process or extraction conditions, heterologous expression of gene clusters in *Streptomyces* hosts is a strong method in drug discovery.

3.4   References

[1]   C. Lasch, M. Stierhof, M.R. Estévez, M. Myronovskyi, J. Zapp, A. Luzhetskyy, Dudomycins: New Secondary Metabolites Produced after Heterologous Expression of an Nrps Cluster from Streptomyces albus ssp. Chlorinus Nrrl B-24108, Microorganisms. 8 (2020) 1800. https://doi.org/10.3390/microorganisms8111800.

[2]   M. Rodríguez Estévez, M. Myronovskyi, N. Gummerlich, S. Nadmid, A. Luzhetskyy, Heterologous Expression of the Nybomycin Gene Cluster from the Marine Strain Streptomyces albus subsp. chlorinus NRRL B-24108, Mar. Drugs. 16 (2018) 435. https://doi.org/10.3390/md16110435.

[3]   M. Rodríguez Estévez, N. Gummerlich, M. Myronovskyi, J. Zapp, A. Luzhetskyy, Benzanthric Acid, a Novel Metabolite From Streptomyces albus Del14 Expressing the Nybomycin Gene Cluster, Front. Chem. 7 (2020). https://doi.org/10.3389/fchem.2019.00896.

[4]   M.R. Estévez, M. Myronovskyi, B. Rosenkränzer, T. Paululat, L. Petzke, J. Ristau, A. Luzhetskyy, Novel fredericamycin variant overproduced by a streptomycin-resistant streptomyces albus subsp. Chlorinus Strain, Mar. Drugs. 18 (2020). https://doi.org/10.3390/md18060284.

[5]   M. Myronovskyi, B. Rosenkränzer, M. Stierhof, L. Petzke, T. Seiser, A. Luzhetskyy, Identification and Heterologous Expression of the Albucidin Gene Cluster from the Marine Strain Streptomyces Albus Subsp. Chlorinus NRRL B-24108, Microorganisms. 8 (2020) 237. https://doi.org/10.3390/microorganisms8020237.

[6]   C. Lasch, M. Stierhof, M.R. Estévez, M. Myronovskyi, J. Zapp, A. Luzhetskyy, Bonsecamin: A New Cyclic Pentapeptide Discovered through Heterologous Expression of a Cryptic Gene Cluster, Microorganisms. 9 (2021) 1640. https://doi.org/10.3390/microorganisms9081640.

[7]   C. Lasch, N. Gummerlich, M. Myronovskyi, A. Palusczak, J. Zapp, A. Luzhetskyy, Loseolamycins: A Group of New Bioactive Alkylresorcinols Produced after Heterologous Expression of a Type III PKS from Micromonospora endolithica, Molecules. 25 (2020) 4594. https://doi.org/10.3390/molecules25204594.

[8]   Y. Ahmed, Y. Rebets, M.R. Estévez, J. Zapp, M. Myronovskyi, A. Luzhetskyy,

Engineering of Streptomyces lividans for heterologous expression of secondary metabolite gene clusters, Microb. Cell Fact. 19 (2020) 5. https://doi.org/10.1186/s12934-020-1277-8.

[9]   M. Myronovskyi, B. Rosenkränzer, S. Nadmid, P. Pujic, P. Normand, A. Luzhetskyy, Generation of a cluster-free Streptomyces albus chassis strains for improved heterologous expression of secondary metabolite clusters, Metab. Eng. 49 (2018) 316–324. https://doi.org/10.1016/j.ymben.2018.09.004.

[10]   T. Siegl, B. Tokovenko, M. Myronovskyi, A. Luzhetskyy, Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes, Metab. Eng. 19 (2013) 98–106. https://doi.org/10.1016/j.ymben.2013.07.006.