Discovery and characterization of novel natural compounds and biosynthetic gene clusters from *Streptomyces*

Dissertation

zur Erlangung des Grades

des Doktors der Naturwissenschaften

der Naturwissenschaftlich-Technischen Fakultät

der Universität des Saarlandes

von

Patrick Oberhäuser

Saarbrücken

2025

Tag des Kolloquiums:	21.04.2026
Dekan:	Prof. Dr.-Ing. Dirk Bähre
Berichterstatter:	Prof. Dr. Andriy Luzhetskyy
	Prof. Dr. Christoph Wittmann
Vorsitz:	Prof. Dr. Uli Kazmaier
Akad. Mitglied:	Dr. Stefan Boettcher

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 August 2020 bis Dezember 2024.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in einem Verfahren zur Erlangung eines akademischen Grades vorgelegt.

Ort, Datum (Patrick Oberhäuser)

Acknowledgements

First and foremost, my deepest gratitude to Prof. Dr. Andriy Luzhetskyy for letting me join the AMEG family as a master student and giving me the opportunity to do my PhD in his group. I am grateful to him for his continuous support, always having an open ear, valuable guidance and constant optimism and positivity to tackle the biggest challenges.

I am thankful to Dr. Maksym Myronovskyi for leading my projects and helping me by sharing his excellent expertise with me. Also many thanks for his sense of details and patience, not only with the experimental part but also with the writing part.

I would like to thank Prof. Dr. Christoph Wittmann for the time invested to review this thesis, share his opinion and the time I spent here during my biotechnology master.

I was very lucky to work in a group full of great people and I am very thankful for that. There are countless stories and moments enough to fill a book. Special thanks goes to Marc, who supervised me during my master thesis, solving all the NMR problems and from whom I was able to learn a lot in many different fields. Thank you Anja for always being positive, joking and laughing around and helping me with HPLC problems. Thank you Niko for all the scientific help, boulder sessions, night outs in the bar or at Staden and discussions. Thank you Ruslan for funny night outs, discussions and all the time we spent doing all kinds of sports. Thank you Alina for all the discussions we had, regardless the topic, always having an ear and all the good night outs we spent together. Thank you Negin, Christopher and all the other colleagues for the time we spent outside of the lab.

Zum Abschluss will ich meiner Familie danken, vor allem meiner Mutter und meinem Bruder die immer an meine Seite standen, immer zu mir gehalten haben egal was passiert und egal wie schwer es war. Nicht nur im Zuge meiner Promotion, auch in all den Jahren davor. Auch will ich meinem Vater danken, der leider für mich zu früh gestorben ist und meine Zeit nach der Schule nie erleben durfte.

Zusammenfassung

Sekundärmetabolite stellen einen wichtigen Bestandteil in der Entwicklung neuer Wirkstoffe dar. Aktinobakterien, insbesondere die Gattung *Streptomyceten,* sind eine reiche Quelle für neue bioaktive Stoffe. Ein Großteil ihres biosynthetischen Potenzials ist unerforscht, da viele biosynthetische Gencluster nicht unter Laborbedingungen exprimiert werden. Das Ziel dieser Arbeit war die Entdeckung neuer Naturstoffe aus *Streptomyceten* Stämme und die Entwicklung Strategien zur Aktivierung bisher stiller Gencluster.

Im ersten Teil dieser Arbeit wurden die neuartigen trisubstituierten Pyrazinone Ichizinone A-C aus *Streptomyces* sp. LV45-129 entdeckt und charakterisiert. Genomanalysen und heterologe Expression ermöglichte die Identifizierung des zugrundeliegenden biosynthetischen Genclusters. Die Ergebnisse geben neue Einblicke in die Biosynthese seltener Pyrazinone und erweitern die bekannte strukturelle Vielfalt dieser Naturstoffklasse.

Der zweite Teil zeigt das Zusammenspiel zwischen dem nativen Mansouramycin BGC und verschiedenen heterologen BGC, was zur Formierung verschiedener Hybridstoffen führte. Teil dieser Stoffe ist Mansevorone, dass ein seltenes 7-azochromone Grundgerüst enthält. Diese Ergebnisse tragen zum Verständnis bei, wie regulatorische Interaktionen zwischen Genclustern zur Entstehung neuer Naturstoffe beitragen.

Der dritten Teil fokussierte sich auf das neue Depsipeptid Atrevomycin. Dessen Biosynthese erfolgt durch das koordinierte Zusammenspiel zweier getrennter NRPS/PKS Gencluster, die vermutlich aus einem gemeinsamen, großen Vorläufer-Cluster hervorgegangen sind.

Abstract

Secondary metabolites are important in the development of new pharmaceuticals. Actinobacteria, especially the genus *Streptomyces*, are prolific producers of bioactive compounds. Much of their biosynthetic potential remains unexploited, as many biosynthetic gene clusters (BGCs) are not expressed under standard laboratory conditions. This dissertation aimed to discover novel natural products from *Streptomyces* strains and to develop strategies for the activation of previously silent gene clusters.

In the first part of this work, the novel trisubstituted pyrazinones Ichizinones A-C were discovered and characterized from *Streptomyces* sp. LV45-129. Genome analysis and heterologous expression enabled the identification of the corresponding BGC. The results provide new insights into the biosynthesis of rare pyrazinone compounds and expand the known structural diversity of this class of natural product.

Demonstrated in the second part is the crosstalk between the native mansouramycin BGC with different heterologous BGC, which led to the formation of different novel hybrid compounds. Among these is mansevorone featuring a rare 7-azochromone scaffold. These findings could contribute to the understanding of how regulatory interactions between gene cluster can contribute to the generation of new natural products.

The third part focused on the new depsipeptide atrevomycin. Its biosynthesis is mediated by the coordinated action of two distinct NRPS/PKS gene clusters, likely evolved from one big ancestor BGC.

Publications

Oberhäuser, P., Myronovskyi, M., Stierhof, M., Gromyko, O., Luzetskyy, A., Identification and Heterologous Expression of a NRPS Biosynthetic Gene Cluster Responsible for the Production of the Pyrazinones Ichizinones A, B and C, Microb Cell Fact 24, 131 (2025). https://doi.org/10.1186/s12934-025-02753-6

Oberhäuser, P., Stierhof, M., Horbal, L., Zaiachkivska, A., Ullrich, G., Gromyko, O., Luzhetskyy, A., Atrevomycin, a non-ribosomally synthesized cyclopeptide from the non-categorized *Streptomyces* LV1-209GEK biosynthesized by two interfering NRPS gene cluster, *to be submitted*

Stierhof, M., Horbal, L., **Oberhäuser, P**., Paluszak, A., Cox, P., Lopatniuk, M., Ruf, C., Zapp, J., Luzhetskyy, A., Hybrid compounds obtained via interactions of the Native Mansouramycin Biosynthesis in the Host *Streptomyces albus* Del14 with Heterologously Expressed Clusters, *to be submitted*

Conference contributions

Oberhäuser, P., Myronovskyi, M., Stierhof, M., Luzhetskyy, A., *Streptomyces albus* as an advances chassis strain for the heterologous production of secondary metabolites, VAAM international Symposium Biology of Bacterial Natural Producers, Saarbrücken, Germany, **2023** (Poster)

Oberhäuser, P., Myronovskyi, M., Stierhof, M., Luzhetskyy, A., *Streptomyces albus* as an advances chassis strain for the heterologous production of secondary metabolites, LCS summer school of engineered living materials, Nonnweiler, Germany, **2023** (Poster)

Table of Contents

Table of Contents

1. Introduction

1.1 Historical perspective: From traditional medicine to modern pharmaceuticals

Nature has been a vital resource for humans since the beginning of its evolution. We depend on it to secure our survival and to meet our needs. Despite all challenges humans had to face, our immune system is vital to protect us from invisible threats like microbial infections such as the pest, tuberculosis and cholera. Early on, the human species experienced and learned by trial and error, that nature offers resources with certain biological effects on the body. These effects were recognized and systemized in many ancient cultures around the world. A tremendous amount of documents evidence this by describing the use of nature to treat diseases and to manage pain. One of the earliest records, dating back to 2600 BC, from old Mesopotamia already describe the use of oils from cedar trees and herbs like myrrh to treat colds, coughs and inflammations.[2-3]The "Ebers Papyrus" from Egypt from around 1500 BC, is the most comprehensive record listing over 700 drugs used to promote health of the human body, most of them derived from plants. Plants have been the main source used as medicines due to their accessibility and renewability. Some of these plants used in traditional medicine are still being used today. Ancient China already mentioned the use of ginseng roots in the book Shennong Bencao Jing (Shennong's herbal, 神農本草經), written 2000 years ago.[4] Ginseng roots are still used today to boost the immune system, improve our memory and as an ingredient for skin and health care products.[5]

Despite the wide use of medicinal plants, the components causing the biological effects responsible for the observed therapeutic effects remained a mystery. Abstract concepts like "the power of life" was long believed in traditional medicine to be the root of their efficacy.[6] This concept was soon replaced with the idea of active ingredients related to the observed medical effects. As scientific knowledge progressed and analytical tools advanced during the 18th and 19th century, researchers were able to isolate specific compounds from crude extracts.[6] Nowadays, those specific compounds are known as natural products (NP). They are produced by the organism's secondary metabolism and are not essential for the growth and survival unlike primary metabolites like nucleotides, amino acids or carbohydrates.[7-9] The secondary metabolism is important for the survival of the organism and provides it with a selective advantage by producing compounds such as antimicrobials, insecticides, antifungals, and poisons.[7-8, 10] A breakthrough was achieved in 1805 by the German pharmacist Friedrich Wilhelm Sertürner, who isolated the active component morphine from opium (Papaver somniferum).[2] Merck later commercialized morphine in 1826, which changed medicinal

applications from the use of crude plant extracts to purified compounds. The first commercially available painkillers now existed.[3] This event led to the isolation of other medically important NPs such as salicin (an analgesic from *Salix alba*)[9], codeine (an analgesic and sedative from *Papaver somniferum*)[11], and quinine (an anti-malarial from *Cinshona sp*.)[12]. A huge interest in natural products ignited leading to the isolation of a vast variety of complex structures with a broad spectrum of activities from all kinds of sources. By the early 1900s, the majority of medicines in use were mainly derived from plants and the basis of modern natural product discovery and drug development was layed.[13-14,10] The field would soon be revolutionized and see a shift from plants to microbes.[14]

1.2 Microbial natural products as a source of drug discovery

Natural products are organic compounds produced by all kingdoms of living things; animals, plants, fungi, protests and monera. The natural environment and evolutionary pressure from opponents led to the requirement for these metabolites. They enable mechanisms that for example support chemical defense against predation, protection against UV radiation and chemical signalling to ensure gene regulation. The molecular weight of NPs typically is less than 3000 Dalton and they depict a tremendous chemical and structural diversity.[15] According to the US Food and Drug Administration (FDA), more than two-third of all approved drugs are either NPs or based on NP compounds, indicating their crucial role in drug development.[14, 16]

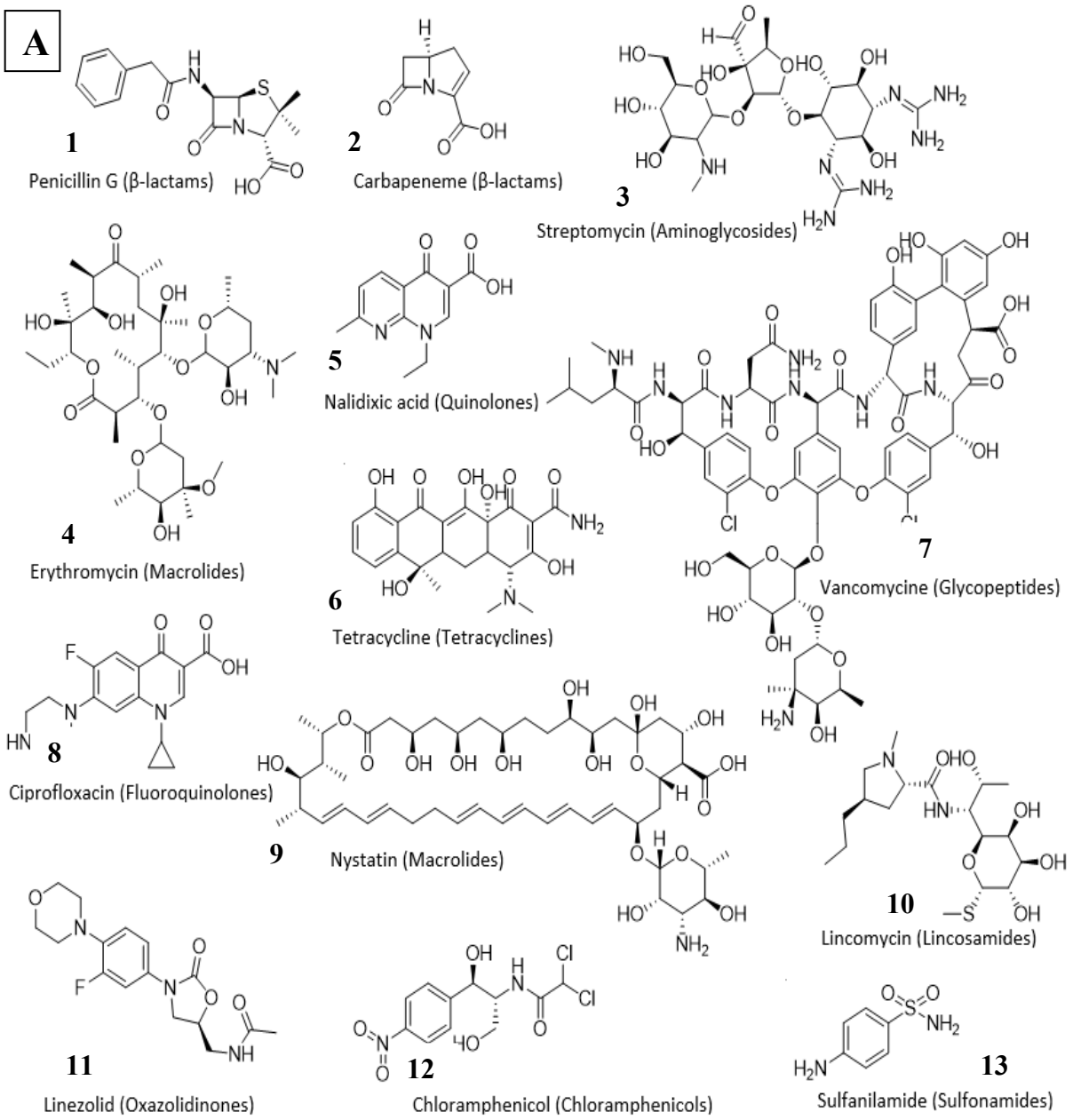
The discovery of penicillin from Penicillium notatum by Sir Alexander Fleming in 1929 marked a new era in drug development and shifted the focus in research of NPs to microbial sources.[14, 17] A period of systematic screening of microorganisms followed mainly driven by an activity guided approach to detect more NPs of clinical interest. Culture broths obtained from cultivated strains showing an activity were separated by chromatographic methods. The obtained fractions were then again screened for activity and this process repeated until the pure compound responsible for the observed activity was isolated.[18-19] Several methods were developed to improve the detection of new metabolites and to further enhance their production (OSMAC – one strain many compounds). This includes the altering of the cultivation conditions (e.g. media composition, aeriation, cultivation conditions like temperature and pH) and co-cultivation of two different microbial strains.[20] The classical activity guided Waksman approach in the early 20th century resulted in the “golden era” of antibiotics in which most of the antibiotic classes known today were characterized between the period of the 1940s and the 1970s.[21] An important example includes streptomycin (**3**), the first effective NP for the treatment of tuberculosis (Fig.

1). This NP is an aminoglycoside, a group of antibiotics active against Gram-negative bacteria by targeting the bacterial protein synthesis by binding to the 30S ribosomal subunit.[22-23] Other classes characterized during this time include the β-lactams such as penicillin (**1**) and carbapeneme (**2**) which inhibit bacterial cell wall synthesis[1, 24-25]; the macrolides like erythromycin (**4**), which inhibit bacterial protein synthesis by binding to the 50S ribosomal subunit[26-27]; the tetracyclines (**6**), that prevent the attachment of the aminoacyl-rRNA to the ribosome[28]; the glycopeptides like vancomycin (**7**), which are crucial against Gram-positive bacteria[29] and the chloramphenicols, which exhibit various mechanisms of action including DNA intercalation and transcription inhibition.[30] These findings made effective treatment of infectious diseases now possible and mortality rates dropped as a consequence.[1] Despite the high success rate with approximately 2500 identified metabolites by the year 1970, there was a significant gap of newly introduced antibiotics until the beginning of the 21st century.[31]The reasons for the huge decline of the NP discovery rate are widespread. Constant rediscovery of already known compounds like streptomycin and tetracycline revealed the limitations of the Waksman approach, an issue referred to as the "dereplication problem".[32] Additionally, the extensive exploitations of the terrestrial environment strongly reduced new potential NP producers and the production yields of potential new compounds were too small.[1]

Alternative strategies of combinatorial chemistry gained popularity to combat the decreasing discovery rate. The focus shifted towards high-throughput screening of huge synthetic compound libraries. These libraries were created by the generation of a vast pool of structurally diverse compounds through systematic assembly of various building blocks. Depending on the interest, the library was screened against a biological target and positive compounds could then be identified through individual interactions.[33] Different approaches with this idea were created like the OBOC (one-bead one-compound) combinatorial peptide library and the phage display peptide library.[34-35] However, this approach had limited success. Bacterial cell wall penetration represents a boundary for many synthetic compounds and they often possess the same basic skeletons which results in a lack of structural diversity and reduced bioavailability. Those drawbacks are further emphasized by the average hit rate of less than 0.001 % of synthetic libraries compared to the 0.3 % hit rate of NP screening.[36-37]

An early solution to avoid the abundant rediscovery issue was the creation of NP databases enabled by the constant progression of analytical tools. Advanced separation methods like high performance liquid chromatography (HPLC) coupled with detection methods like diode array detectors (DAD), mass spectrometry (MS) and nuclear magnetic resonance spectroscopy

(NMR) allowed for precise separation of crude extracts and recording of the eluting compounds.[32, 38] These techniques make it possible to analyse metabolic profiles of strains and to gather valuable chemical and physical information of compounds regarding their retention time, exact mass, UV/Vis absorption, MS/MS and isotopic fragmentation pattern and more. This data is systematically collected in NP databases and makes comparison with newly obtained datasets possible. This process known as "dereplication" is an effective strategy to distinguish between new and known compounds as early as possible and helps to avoid the characterization of already known NPs.[32, 38] Several commercially and publicly accessible databases have emerged to serve this purpose. Based on the Journal of Cheminformatics article titled "Review on natural products databases: where to find data in 2020", the world's largest collection of chemical information, in terms of the estimated number of NP entries, is PubChem with ~438,000 entries.[39-40] Other big databases frequently used in the dereplication process are CAS/SciFinder[41] with >300.000 and the Dictionary of Natural Products (DNP) with >230.000 entries, respectively[42]. They provide a comprehensive coverage of compounds with extensive metadata allowing for advanced searches, including substructures, and include links to relevant literature.[39] Dereplication has become a crucial strategy to avoid redundant characterization of known compounds and opened together with advances in genomics a modern approach in NPs discovery. These developments will be discussed in chapter 1.4.



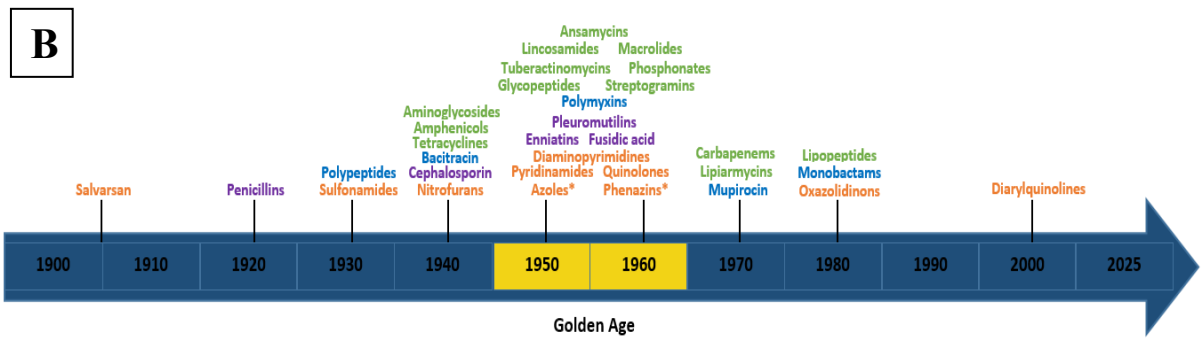


Figure 1: (**A**) Structures of antibiotics and their respective compound class. (**B**) Antibiotics discovery timeline. Green: Actinomycete natural products; Blue: Other bacterial natural products; Purple: Fungal natural products; Orange: Synthetic antibiotics, * indicates synthesis inspired by a natural product[1]

Despite its efficiency and utility, the process of dereplication has its limitations in accuracy and automation. It has the potential for misidentification due to spectral similarity between distinct compounds. The mass-to-charge ratio of adduct ions or fragments formed during mass spectrometry can mistakenly be considered as a new compound. Similarly, the mass-to-charge ratio can be identical with a known compound and a potential new compound will be overlooked. Many natural product databases lack complete spectral data and standardized metadata, such as source organism information, which limits search specificity. Moreover, dereplication workflows often rely on fragmented software solutions requiring manual interpretation, especially in ranking or interpretative searches. These limitations can lead to false positives or negatives, complicating the distinction between novel and known compounds. As such, while dereplication accelerates natural product research, it is not yet a fully error-free or automated process and requires continued improvement in database completeness, spectral resolution, and computational tools.[43-44]

1.3 Exploring Actinobacteria: Key natural product groups and their significance

The phylum Actinobacteria are among the most prolific producers of biologically active secondary metabolites. More than half of all clinically used antibiotic classes derive from them, underlying their huge contribution to the success of microbial NPs since the golden era of antibiotics.[45] Actinomycetes are Gram-positive, aerobic, filamentous bacteria with high G+C content in their genome and are typically found in terrestrial and aquatic ecosystems. Initially, they were regarded as bacteria with fungal-like features due to their characteristic formation of branched mycelium, which allows the bacteria to efficiently colonize. The mycelium of Actinobacteria typically includes both substrate (vegetative) and aerial hyphae, with the latter being able to differentiate into spores.[46-47]

Streptomyces are the biggest genus of Actinobacteria and one of the main sources of NPs. With about 50% of clinically used antibiotics of microbial origin, the majority is originating from *Streptomyces.*[48] Improvements in recombinant DNA methodology and bioinformatic tools allowed the sequencing of whole genomes revealing an enormous chemical and biosynthetic potential hidden in their genes. Together with comparative analyses of the growing database of described NPs biosynthetic pathways, it became apparent that genes involved in the biosynthesis of secondary metabolites in bacteria are clustered and not randomly distributed throughout the whole genome.[49] This finding revealed a discrepancy between biosynthetic gene

clusters (BGCs) present and characterized compounds, indicating that the metabolic potential of these bacteria is even greater than estimated.[50-51]

The genus *Streptomyces* produce NPs with diverse structures and biological activities and can generally be classified into classes according to their biosynthetic origin like macrolides (e.g. nystatin (**9**), erythromycin (**4**)), tetracyclines (**6**), aminoglycosides (e.g. streptomycin (**3**)) and terpenes (Fig. 1).[52] Central to their biosynthesis are several enzymatic platforms, amongst them the non-ribosomal peptide synthetases (NRPs) and the polyketide synthases (PKS). Both pathways utilize different building blocks to biosynthesis a variety of core structures. Polyketides, similar to the fatty acid pathway, are built from acyl-CoA esters like acetyl-CoA and malonyl-CoA, while peptides from NRPS are assembled from α- and β-amino acids. Post-translational modifications can further alternate the core of NPs extending the broad variation of them.[53-54] To understanding the biosynthesis of NPs formation it is essential to know the fundamental genetic and enzymatic principles underlying these core pathways.

1.3.1 Biosynthesis of peptides

Peptides produced by Actinobacteria fall into two major categories, the ribosomally synthesized and post-translationally modified peptides (RiPPs) and the NRPs. The biosynthesis of RiPPs starts with the ribosomal synthesis of a linear precursor that is divided into a N-terminal leader and a C-terminal core region. While the leader region is important for substrate recognition, the core region undergoes posttranslational modifications and is subsequently cleaved off by proteolysis to form the mature peptide. The biosynthesis of RiPPs can be highly specific due to the physical separation of both motifs. Their substrate tolerance alongside their diverse core structures have been used to characterize many RiPP classes, including lantipeptides, thiopeptides, and lasso peptides many of which exhibit strong antimicrobial activity.[55-57]

Unlike RiPPs, peptides from NRPS pathways are assembled by large, multi-modular enzyme complexes and are synthesized independently of the ribosome. Although the NRPS enzymes themselves are produced via the canonical transcription and translation machinery, the assembly of the peptide product does not involve mRNA templates or ribosomal machinery (Fig. 2). Instead, their products are biosynthesized via the coordinated action of specialized enzymatic domains with each NRPS module typically incorporating a single amino acid into the growing peptide chain, following a collinear logic. A module typically consists of three core domains. The adenylation (A) domain is responsible for substrate recognition and activates a specific amino acid that can be proteinogenic and non-proteinogenic. Once the amino acid is

activated it is passed on to the thiolation domain (T) or peptidyl carrier protein (PCP). There it becomes covalently linked through the flexible 4'-phosphopantetheine arm, which positions the substrate for the next catalytic step. The condensation domain (C) then joins the tethered intermediates from adjacent modules by forming a peptide bond, thereby elongating the chain in a sequential manner. The release of the mature peptide is typically mediated by a thioesterase (TE) domain by either catalyzing a hydrolytic cleavage or cyclization.[58-60] Additional modification domains can tailor the mature peptide by introducing diverse chemical modifications. Epimerases (E) can change the stereochemistry of individual residues by converting L-amino acids into their D-enantiomers. Methyltransferases (MT) catalyze the transfer of methyl groups to the peptide backbone and introduce small but functionally important modifications. Oxidative (Ox) enzymes can change the stereochemistry of individual residues and cyclization (Cy) domains catalyze oxidative cross-linking and ring structures that stabilize the scaffold and shape the biological activity.[54, 61-62] These modular features are also used for metabolic engineering as they provide opportunities to fine-tune yields and to create novel compounds.[59-60, 63]

The vast diversity of NRPS compounds is mainly achieved by the broad substrate pool of A-domains and a series of different C-domains types. While the A-domain can activate all 20 proteinogenic amino acids and hundreds of non-proteinogenic amino acids, the C-domain can catalyse reactions other than peptide bond formation.[64] Various functional subtypes of these C-domain exist. Common examples include $^{L}C_{L}$ and $^{D}C_{L}$, which catalyse peptide bond formation between two L-amino acids ($^{L}C_{L}$) and an L-amino acid and a growing peptide ending with a D-amino acid ($^{D}C_{L}$). A Cstart domain can acylate the first amino acid typically with a β-hydroxyl fatty acid while Cyc domains can catalyse heterocyclizations combining peptide bond formation and subsequent cyclization.[64-66]

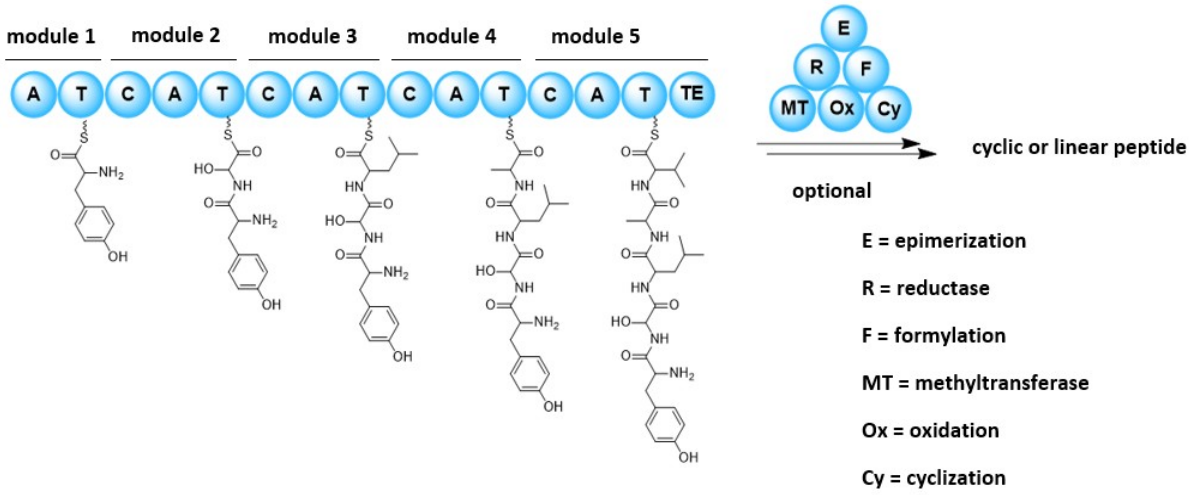


Figure 2: Schematic representation of a hypothetical NRPS pathway consisting of 5 modules with the minimal set of domains. Module 1 acts as the loading module and lacks the condensation domain. The hypothetical peptide consists of the amino acids tyrosin, serine, leucine, alanine and valine.

The biosynthetic flexibility allows NRPS system to incorporate non-proteinogenic amino acids, contributing to their unique structures and biological activities. Examples of NRPS derived compounds relevant to medicine and biotechnology involves for example the 'last resort' antibiotic vancomycin and daptomycin for treating drug-resistant infections.[63, 67] The potential for more NRPS metabolites become evident by advanced bioinformatic and mining techniques revealing many NRPS BGC in actinobacterial genomes. This capacity for novel compounds is supported by modern approaches like domain swapping and module engineering to optimize these systems and to create next-generation therapeutics.[59-60, 63]

1.3.2 Biosynthesis of polyketides and fatty acids

Polyketides represent a large and versatile family of NPs. Their biosynthesis proceeds in a modular, assembly line manner similar to the biosynthesis of NRPS and also shares similarities with fatty acid biosynthesis (FAS) in terms of both, chemical logic and architectural design.[62, 68]

The mega enzyme complexes of PKS systems can be categorized on the basis of their organization and functionality. Type I systems functions in a modular, non-iterative manner. Type II systems consists of dissociated enzymes that act iteratively and type III systems are minimal homodimers with an iterative mechanism. PKS systems build their products similar to FAS by sequential incorporation of acyl units. The biosynthesis usually starts with the loading of a short acyl starter, usually acetyl- or propionyl-CoA, onto a ketosynthase (KS) domain by an acyltransferase (AT). The chain gets then elongated with malonyl- or methylmalonyl extender units through the decarboxylative condensation catalysed by the KS. A specialized AT, the malonyl- acetyl transferase (MAT), delivers in bacterial systems these extender units. This

enzyme ensures, that the initiation and the elongation substrates do no compete for the same catalytic site. Once transferred to the phosphopantheine arm of the ACP, the intermediated are correctly positioned for condensation.[62, 69]

The enzymes in type I PKS are organized into modular chains, where each module typically consists of the minimal set of KS, AT and ACP domains. Optional domains like the ketoreductase (KR), dehydratase (DH) and enoylreductase (ER) can be present in a module. They can alter the oxidation level of the intermediate at each step. A thioesterase (TE) domain located at the end of the assembly line terminates the biosynthesis through hydrolytic cleavage and the mature product is released. The number of modules present in the assembly line largely determined the length of the final product.[70] A well-studied example of this biosynthetic strategy is the assembly of 6-deoxyerythronolide B, the macrolactone intermediate of erythromycin (Fig. 3).[71]

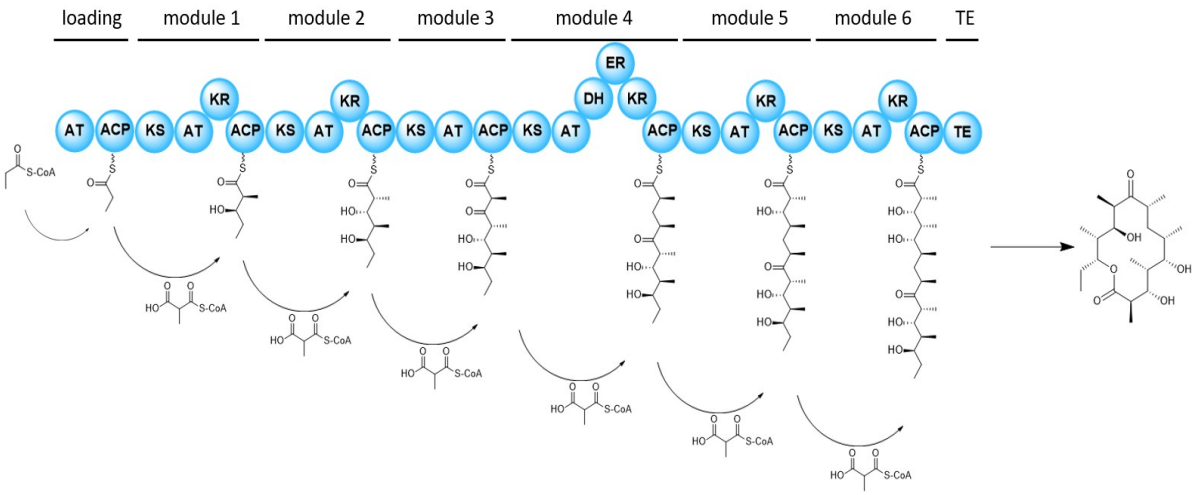


Figure 3: Schematic representation of a modular type I PKS biosynthesis using the 6-deoxyerythronolide B synthase required for erythromycin biosynthesis.

In contrast to the modular architecture of type I PKS, the type II PKS relies on a small set of discrete enzymes. The catalytic core consists of the monoenzymatic subunit KS_{α} and KS_{β} as well as an ACP. Together, these enzymes extend the polyketide chain in an iterative cycle of condensation reactions, with the ACP being responsible for holding and transferring the growing chain during biosynthesis. The number of iterations and the ultimate length of the final product is determined by the length of the channel formed by KS_{β}, which is often also referred to as the chain length factor (CLF) (Fig. 4).[72]

Type III PKS, often referred as chalcone/stilbene synthase superfamily, is built from a single KS domain forming a homodimer. It operates iteratively like type II, but without ACP and direct

utilization of coenzyme A to deliver the extender units and hold the growing chain during the assembly process. This minimal system performs all the functions of type I and II, carrying out priming, chain extension and often cyclization.[73] Based on the post-PKS processing enzymes, this type can be divided into five groups (RppA, PhlD, DpqA, ArsB, PKS18) (Fig. 4).[74]

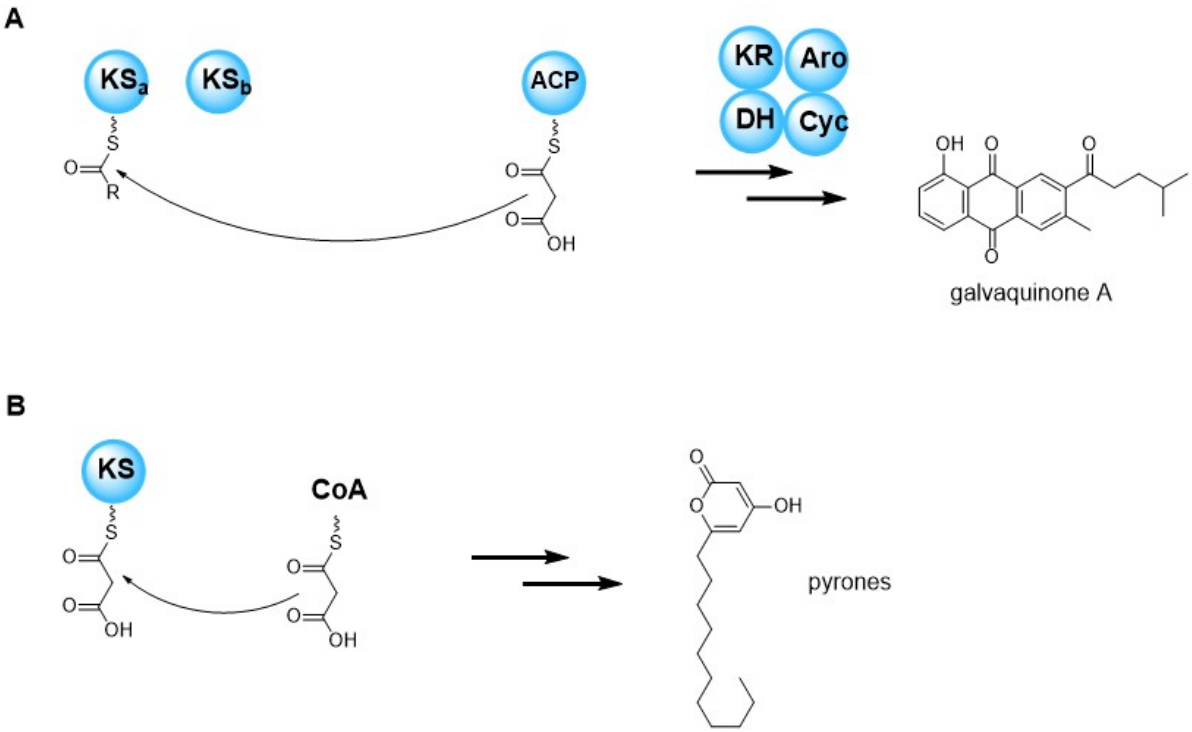


Figure 4: Schematic representation of a type II PKS biosynthesis using galvaquinone A as an example (**A**)[75] and schematic representation of a type III PKS biosynthesis using pyrones as an example (**B**)[76].

After the core polyketide chain is assembled, tailoring enzymes can modify these scaffolds to further diversify the final products by introducing structural complexity and biological functionality. Those enzymes are typically encoded in the same BGC and can include for example cyclases, oxidases, methyltransferases, glycosyltransferases and halogenases.[77-78] Also type III PKSs, despite their relatively simple homodimeric structure, can yield products with broad chemical diversity. Crystal and mutational studies of chalcone synthase–type III PKSs show that small changes in the active-site can alter product length, starter/ extender specificity, and cyclization patterns. This structural flexibility enables the generation of small hydroxylated metabolites, most often simple aromatic rings or lactones.[79] More broadly, the product classes tend to align with the PKS type. Modular type I systems often generate macrolides, polyethers or polyenes while type II systems typically produce aromatic polyketides whose scaffolds are defined by specific cyclization patterns. For example, this group includes compound families like the tetracyclines and anthracyclines.[54, 80-83] In addition to clusters encoding one of the three PKS types, nature also features hybrid systems that combine PKS modules with other

biosynthetic domains. Examples include NRPS-PKS assemblies that generate mixed peptide-polyketide frameworks.[84-85] Starter units from FAS pathways can likewise feed into NRPS or PKS routes and further enhance the structural diversity.[70]

Fatty acid synthases share the same iterative logic as PKS. Both systems rely on dedicated enzymes with specialized functions and on a stepwise chain extension process using different acyl-CoA building blocks.[86-87] The similar architectures led to the idea that PKS systems evolved from an ancestral FAS. However, the functional outcomes differ between both systems. In FAS pathways, the β-keto groups are fully reduced during each elongation cycle, resulting in saturated fatty acids of various chain length. The incomplete reduction in PKS generates greater structural diversity in contrast to FAS[70, 88]. Several secondary metabolites exemplify the integration of fatty acid moieties into NP scaffolds which shows the biochemical interplay between PKS/FAS and NRPS systems. An example is the cyclic lipopeptide surfactin, whose fatty acid is essential for membrane interaction and biological activity, thereby linking primary fatty acid metabolism with secondary metabolite function.[89]

1.4 Decoding bioactive potential: Discovery of new NPs and their BGC

High-throughput sequencing entered routine use in the early 2000s and revealed the true extent of biosynthetic capacity encoded in bacterial genomes. Sequencing efforts in *Streptomyces* and other actinobacteria revealed a striking mismatch between the identified BGCs encoded in their genomes and the number of NPs that had been isolated and assigned to them.[90-92] This discrepancy was shown by comparing elucidated biosynthetic pathways of already characterized NPs with BGCs identified in the genomes. The integration of biosynthetic knowledge into databases enabled systematic correlation of BGCs with their corresponding metabolites, allowing researchers to distinguish between characterized and cryptic clusters. This means, that actinobacteria still harbour an immense reservoir of potential new NPs with unexplored chemical scaffolds and biological activities.[93-94] With this milestone, the concept of genome mining was introduced. It involves using available genomic information to predict and screen biosynthetic gene clusters (BGCs). Once identified, these clusters can be activated to trigger the expression of specific genes leading to the production, isolation, and eventual characterization of the corresponding natural compounds. Multiple methods have been developed to facilitate each step of this process, from in silico prediction to experimental validation and compound analysis.[93, 95] The rising accessibility of genome sequence data in conjunction with deeper understanding of NP biosynthetic pathways opened the development of specialized software platforms.[96] Widely used examples incldue antiSMASH[97], BiG-SCAPE[98],

and PRISM[99], which allow automated annotation and preliminary functional prediction of BGCs across sequenced genomes These tools made it possible to catalogue BGCs encoding diverse systems like PKSs, NRPSs, RiPPs and even hybrid pathways, opening the possibility to directly characterize genomes *in silico*.[92, 100]

One of the most adopted strategy to verify predicted BGCs experimentally is heterologous expression, where BGCs of interest are cloned and expressed in genetically tractable surrogate hosts. It's utility in the functional characterization of cryptic or orphan gene clusters represents an advantage, as the resulting metabolites can be directly traced back to the introduced BGC.[101-102] The strains used for this approach are optimized for secondary metabolite production with reduced metabolic background. This enhances the availability of precursors and energy, resulting in more efficient flux of those resources towards the biosynthetic pathway of the heterologously expressed NP. The lower complexity of these chassis organisms not only increases the chance of detecting novel NPs, it also makes purification of new NPs more straightforward.[101] Metabolic profiles obtained from extracted fermentation broths of host strains with and without the target cluster can now easily be compared via appropriate methods like LC-MS. Produced compounds represented by peaks in the chromatogram of the metabolic profile that are present in the host strain but absent in the empty heterologous host are likely products of the introduced BGC.[103-105] In some cases, fusion or hybrid compounds can be observed as a result of metabolic crosstalk between the introduced BGC and endogenous biosynthetic pathways of the host.[106] Model host systems such as *Streptomyces albus*, *Escherichia coli* and *Bacillus subtilis* are highly genetically tractable, which allows targeted modification and refactoring of BGCs. Promotor swaps, insertion of strong constitutive or inducible elements, removal of repressors and overexpression of pathway specific activators are among the most common modifications applied to foreign BGCs. These changes cannot only boost the expression, they can also enable combinatorial biosynthesis. By mixing and matching pathway components, new structural analogs with altered or improved bioactivity can be generated with this approach.[107-108]

Heterologous systems also make it possible to work with metagenomic DNA from microbes that cannot yet be grown in laboratory conditions, expanding the pool of biosynthetic diversity available to study[109]. This is valuable for organisms found in soil, marine and host-associated microbiomes, where much of the chemical potential remains untapped because of those limitations.[110-111]

To capture this diversity, researchers often build genomic libraries which are collections of cloned DNA fragments that together represent an organism's entire genomic content.[101, 107] These libraries provide a route to examine specific genes or pathways including BGCs that can range in size from just a few kilobases to over 100 kb. Because of this size variation, vector choice matters. Cosmids can carry around 40 – 50 kb and are often used for mid-sized clusters.[112-113] Bacterial artificial chromosomes (BAC) on the other hand are more suited for capturing very large BGCs intact, allowing precise reconstruction of the biosynthetic pathway.[114]

Still, heterologous expression often fails to deliver efficient metabolite production. Yields often remain below those of the native producer. This can result from regulatory mismatches, limited precursors, compound toxicity or missing cofactors in the surrogate host. To improve output, researchers adjust both the host and the pathway. Common strategies include genomic deletions, regulatory tuning or co-expression of accessory genes.[101, 115]

1.4.1 Heterologous expression systems in natural product discovery

Heterologous expression systems play a central role in natural products research. They can bring silent BGCs to life and can also be used to create new derivatives and raise production levels of certain compounds. Early efforts often relied on *Escherichia coli* and *Bacillus subtilis*. These bacteria are familiar laboratory workhorses values for their quick growth and easy manipulation.[116-117] Yet, they fall short when asked to produce complex secondary metabolites. Many of the necessary precursors or cofactors are missing and some compounds are toxic to the host. The intricate enzyme assemblies involved in these pathways may also fail to fold or function properly. Another problem is the way these bacteria regulate gene expression, which differs from that of organisms naturally producing such compounds. As a result, attempts to express large and complex gene clusters in them often lead to poor or no yields.[118]

To overcome these limitations, researchers have turned more towards *Streptomyces* strains. They are known for their prolific production of secondary metabolites and possess the inherent metabolic strengths for the efficient expression of foreign BGCs, a characteristic trait for an ideal heterologous host. Several features must be combined in effective surrogate hosts. It should be genetically easy to manipulate and fast growing, while at the same time producing few interfering metabolites of its own. A steady supply of precursors and cofactors is equally important to channel resources more towards the introduced pathway and to achieve high production yields. The heterologous host also needs to accommodate large DNA constructs, which often requires vector systems such as cosmids and BACs.[101, 119-121]

Streptomyces lividans and *Streptomyces albus* have emerged as preferred host systems within the actinobacteria. Both strains have been extensively engineered with targeted genetic modifications to improve their performance in heterologous expression. *Streptomyces lividans* TK24 has long served as a favoured host due to its genetic stability and reliable transformation protocols.[108] This *Streptomyces lividans* 66 derived strain lacks the self-replicating plasmids SLP2 and SLP3 and contains the point mutation [K88E] in the *rpsL* gene, which was shown to enhance secondary metabolite production.[122-123] However, its genome still harbours several endogenous BGCs, which can complicate the detection of heterologous product.[104] This has been addressed by Ahmed et al. by systematically deleting clusters to generate a set of genome-minimized derivatives, removing 11 endogenous BGCs from the genome covered by more than 200 kb of chromosomal DNA in total.[104] The resulting ΔYA strains displayed a simplified metabolic background that allowed a clearer attribution of newly observed metabolites to the introduced pathways. Importantly, these modifications did not affect the growth of the strain negatively. Further improvements were achieved through introducing multiple φC31 *attB* integration sites. These sites allow for stable integration and multi-copy expression of BGCs and thereby boosting product titers significantlly.[104]

A similar approach has been applied to engineer the heterologous host *Streptomyces albus*. The *S. albus* J1074 strain, a derivative of *S. albus* G, was already a widely used host for heterologous expression.[124] This strain offers a fast growth rate, efficient DNA uptake, and a compact genome of 6.8 Mb with relatively few interfering secondary metabolites.[124-125] Myronovskyi et al. further advanced this strain by developing the *S. albus* Del14 host by deleting 15 native clusters that were covered by roughly 500 kb of DNA. This resulted in a streamlined metabolic background that significantly enhanced the deletion and purification of heterologously expressed compounds.[103] Additional *attB* integration sites were added to further engineer *S. albus* Del14, resulting in the derivatives B2P1, B3P1 and B4.[126] These *attB* sites are used by the phage φC31 based integrative vector commonly used for heterologous expression of BGCs. The vector can integrate simultaneously into several *attB* sites in the chromosome and therefor increase the copy number of the BGC. This modular integration strategy translated into higher yields across diverse NPs. This success demonstrated that *S. albus* Del14 is not only highly versatile but can also outperform host strains like *S. coelicolor* or even the native producer. The engineered host also demonstrated compatibility with foreign DNA from phylogenetically distant organisms such as *Frankia.*[103]

There is currently no host strain that supports all BGCs equally well. The outcome of expression depends on the origin of the cluster, the types of enzymes it encodes and the metabolic resources required for product formation. An example is *Streptomyces lividans*, which was shown to produce amino acid derived compounds like tunicamycin and deoxycinnamycin effectively. In contrast, *Streptomyces albus* performs better when handling polyketide synthase pathways.[127-129]The introduction of additional cluster copies through *attb* integration can sometimes raise production, but the effect is not consistent and often depends on how the pathway is regulated inside the host.[103-104] The size of many BGCs adds another layer of complexity. Some clusters exceed 100 kb, which makes them difficult to handle. Cosmid and BAC libraries can be constructed to capture and express these large sequences. Nevertheless, this approach does not eliminate the problem of incompatibility between the host and the introduced pathway. Such mismatches may occur at different levels: promoters may not be recognized, precursors may not be available, enzymes may not be folded or modified correctly. This needs to be considered and can be solved by different solutions provided by synthetic biology: Pathways can be refactored, promoters exchanged or additional regulatory and resistance genes supplied to restore activity.[113]

Recent advances also come from the design of genome reduced strains. Variants such as *S. albus* Del14 and *S. lividans* ΔYA9–11 have been stripped of competing pathways, which simplifies their metabolism and makes them easier to manipulate genetically. These strains create an environment that is more receptive for foreign BGCs, including those from organisms that cannot yet be cultured. As methods for genome mining and pathway engineering continue to improve, these optimized hosts are expected to remain central tools for natural product discovery.

1.5 Outline of this work

Natural products have served as a rich source of clinically useful antibiotics and remain central to drug discovery. The antibiotic resistance problem is a global threat that constantly askes for new scaffolds that can also serve as lead structures in the field of synthetic chemistry. This includes the identification of BGCs and the elucidation of biosynthetic pathways. The aim of the presented work was the investigation of *Streptomyces* strains with the goal to isolate and to characterize novel secondary metabolites and their BGC.

Chapter 2.1 and 2.2 focuses on the uncharacterized *Streptomyces* sp. LV45-129. A cosmid library was constructed in order to systematically express BGCs in heterologous hosts. This approach led to the production of ichizinones A-C representing novel member of the pyrazinone class of natural products. Their structures were elucidated which helped together with BGC organization and gene deletions to propose their biosynthesis.

The expression of another BGC led to biosynthetic crosstalk between the introduced genes and the native mansouramycin gene cluster in *S. albus* Del14. The novel hybrid compound mansevorone was produced featuring a rare 7-azochromone scaffold. So far, the biosynthesis of this scaffold is not described and only hardly achieved by synthetic chemistry. Furthermore, azochromones are associated with various biological and pharmaceutical properties.

Chapter 2.3 focuses on the isolation and characterization of the novel depsipeptides atrevomycin A and B from the uncharacterized *Streptomyces* sp. LV1-209GEK. This compound was found to be produced by the interaction of two separate NRPS/PKS hybrid gene clusters. Comparison of the gene cluster with BGC of known compounds in order to propose the biosynthesis showed, that the identified separate gene clusters likely evolved from one big ancestor BGC. The structurally related compound atrovimycin has previously demonstrated activity against *Physarum* and *Mycobacterium tuberculosis*, prompting current activity tests to evaluate whether atrevomycin A and B exhibit similar bioactivity.

1.6 References

1. Hutchings, M. I.; Truman, A. W.; Wilkinson, B., Antibiotics: past, present and future. *Current opinion in microbiology* **2019,** *51*, 72-80.
2. Ji, H. F.; Li, X. J.; Zhang, H. Y., Natural products and drug discovery: can thousands of years of ancient medical knowledge lead us to new and powerful drug combinations in the fight against cancer and dementia? *EMBO reports* **2009,** *10* (3), 194-200.
3. Newman, D. J.; Cragg, G. M.; Snader, K. M., The influence of natural products upon drug discovery. *Natural product reports* **2000,** *17* (3), 215-234.
4. Park, H. J.; Kim, D. H.; Park, S. J.; Kim, J. M.; Ryu, J. H., Ginseng in traditional herbal prescriptions. *Journal of ginseng research* **2012,** *36* (3), 225.
5. Newman, D. J.; Cragg, G. M., Natural products as sources of new drugs over the last 25 years. *Journal of natural products* **2007,** *70* (3), 461-477.
6. Lahlou, M., The success of natural products in drug discovery. **2013**.
7. Pang, Z.; Chen, J.; Wang, T.; Gao, C.; Li, Z.; Guo, L.; Xu, J.; Cheng, Y., Linking plant secondary metabolites and plant microbiomes: a review. *Frontiers in plant science* **2021,** *12*, 621276.
8. Abdel-Razek, A. S.; El-Naggar, M. E.; Allam, A.; Morsy, O. M.; Othman, S. I., Microbial natural products in drug discovery. *Processes* **2020,** *8* (4), 470.
9. Schmid, B.; Kötter, I.; Heide, L., Pharmacokinetics of salicin after oral administration of a standardised willow bark extract. *European journal of clinical pharmacology* **2001,** *57*, 387-391.
10. Jakubiec-Krzesniak, K.; Rajnisz-Mateusiak, A.; Guspiel, A.; Ziemska, J.; Solecka, J., Secondary metabolites of actinomycetes and their antibacterial, antifungal and antiviral properties. *Polish journal of microbiology* **2018,** *67* (3), 259.
11. Sindrup, S. H.; Brøsen, K., The pharmacogenetics of codeine hypoalgesia. *Pharmacogenetics and Genomics* **1995,** *5* (6), 335-346.
12. Achan, J.; Talisuna, A. O.; Erhart, A.; Yeka, A.; Tibenderana, J. K.; Baliraine, F. N.; Rosenthal, P. J.; D'Alessandro, U., Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. *Malaria journal* **2011,** *10*, 1-12.
13. Cragg, G. M.; Newman, D. J., Natural products: a continuing source of novel drug leads. *Biochimica et Biophysica Acta (BBA)-General Subjects* **2013,** *1830* (6), 3670-3695.
14. Pham, J. V.; Yilma, M. A.; Feliz, A.; Majid, M. T.; Maffetone, N.; Walker, J. R.; Kim, E.; Cho, H. J.; Reynolds, J. M.; Song, M. C., A review of the microbial production of bioactive natural products and biologics. *Frontiers in microbiology* **2019,** *10*, 1404.
15. Pilkington, L. I., A chemometric analysis of deep-sea natural products. *Molecules* **2019,** *24* (21), 3942.
16. Newman, D. J.; Cragg, G. M., Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *Journal of natural products* **2020,** *83* (3), 770-803.
17. Patridge, E.; Gareiss, P.; Kinch, M. S.; Hoyer, D., An analysis of FDA-approved drugs: natural products and their derivatives. *Drug discovery today* **2016,** *21* (2), 204-207.
18. Nothias, L.-F.; Nothias-Esposito, M.; Da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J., Bioactivity-based molecular networking for the discovery of drug leads in natural product bioassay-guided fractionation. *Journal of natural products* **2018,** *81* (4), 758-767.
19. Katz, L.; Baltz, R. H., Natural product discovery: past, present, and future. *Journal of Industrial Microbiology and Biotechnology* **2016,** *43* (2-3), 155-176.
20. Pan, R.; Bai, X.; Chen, J.; Zhang, H.; Wang, H., Exploring structural diversity of microbe secondary metabolites using OSMAC strategy: A literature review. *Frontiers in Microbiology* **2019,** *10*, 294.

21. Ribeiro da Cunha, B.; Fonseca, L. P.; Calado, C. R., Antibiotic discovery: where have we come from, where do we go? *Antibiotics* **2019,** *8* (2), 45.
22. Waksman, S. A.; Schatz, A., Streptomycin–origin, nature, and properties. *Journal of the American Pharmaceutical Association* **1945,** *34* (11), 273-291.
23. Garneau-Tsodikova, S.; Labby, K. J., Mechanisms of resistance to aminoglycoside antibiotics: overview and perspectives. *Medchemcomm* **2016,** *7* (1), 11-27.
24. Lederberg, J., Mechanism of action of penicillin. *Journal of bacteriology* **1957,** *73* (1), 144-144.
25. Aurilio, C.; Sansone, P.; Barbarisi, M.; Pota, V.; Giaccari, L. G.; Coppolino, F.; Barbarisi, A.; Passavanti, M. B.; Pace, M. C., Mechanisms of action of carbapenem resistance. *Antibiotics* **2022,** *11* (3), 421.
26. Haight, T. H.; Finland, M., The antibacterial action of erythromycin. *Proceedings of the Society for Experimental Biology and Medicine* **1952,** *81* (1), 175-183.
27. Vázquez-Laslop, N.; Mankin, A. S., How macrolide antibiotics work. *Trends in biochemical sciences* **2018,** *43* (9), 668-684.
28. Ramachanderan, R.; Schaefer, B., Tetracycline antibiotics. *ChemTexts* **2021,** *7* (3), 18.
29. Stogios, P. J.; Savchenko, A., Molecular mechanisms of vancomycin resistance. *Protein Science* **2020,** *29* (3), 654-669.
30. Xue, L.; Spahn, C. M.; Schacherl, M.; Mahamid, J., Structural insights into context-dependent inhibitory mechanisms of chloramphenicol in cells. *Nature Structural & Molecular Biology* **2025,** *32* (2), 257-267.
31. Newman, D. J.; Cragg, G. M., Natural products as sources of new drugs from 1981 to 2014. *Journal of natural products* **2016,** *79* (3), 629-661.
32. Gaudêncio, S. P.; Pereira, F., Dereplication: racing to speed up the natural products discovery process. *Natural product reports* **2015,** *32* (6), 779-810.
33. Liu, R.; Li, X.; Lam, K. S., Combinatorial chemistry in drug discovery. *Current opinion in chemical biology* **2017,** *38*, 117-126.
34. Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J., A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* **1991,** *354* (6348), 82-84.
35. Smith, G. P., Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **1985,** *228* (4705), 1315-1317.
36. Nisar, B.; Sultan, A.; Rubab, S., Comparison of medicinally important natural products versus synthetic drugs-a short commentary. *Nat. Prod. Chem. Res* **2018,** *6* (2), 308.
37. Ayon, N. J., High-throughput screening of natural product and synthetic molecule libraries for antibacterial drug discovery. *Metabolites* **2023,** *13* (5), 625.
38. Hubert, J.; Nuzillard, J.-M.; Renault, J.-H., Dereplication strategies in natural product research: How many tools and methodologies behind the same concept? *Phytochemistry Reviews* **2017,** *16*, 55-95.
39. Sorokina, M.; Steinbeck, C., Review on natural products databases: where to find data in 2020. *Journal of cheminformatics* **2020,** *12* (1), 20.
40. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B., PubChem 2023 update. *Nucleic acids research* **2023,** *51* (D1), D1373-D1380.
41. CAS SciFinder^{n}. Chemical Abstracts Service: Columbus, OH, 2025.
42. Buckingham, J. E. Dictionary of Natural Products. https://dnp.chemnetbase.com.
43. Mohamed, A.; Nguyen, C. H.; Mamitsuka, H., Current status and prospects of computational resources for natural product dereplication: a review. *Briefings in bioinformatics* **2016,** *17* (2), 309-321.

44. Zani, C. L.; Carroll, A. R., Database for rapid dereplication of known natural products using data from MS and fast NMR experiments. *Journal of natural products* **2017,** *80* (6), 1758-1766.
45. De Simeis, D.; Serra, S., Actinomycetes: A never-ending source of bioactive compounds—An overview on antibiotics production. *Antibiotics* **2021,** *10* (5), 483.
46. Barka, E. A.; Vatsa, P.; Sanchez, L.; Gaveau-Vaillant, N.; Jacquard, C.; Klenk, H.-P.; Clément, C.; Ouhdouch, Y.; van Wezel, G. P., Taxonomy, physiology, and natural products of Actinobacteria. *Microbiology and molecular biology reviews* **2016,** *80* (1), 1-43.
47. Van der Meij, A.; Worsley, S. F.; Hutchings, M. I.; van Wezel, G. P., Chemical ecology of antibiotic production by actinomycetes. *FEMS microbiology reviews* **2017,** *41* (3), 392-416.
48. Parra, J.; Beaton, A.; Seipke, R. F.; Wilkinson, B.; Hutchings, M. I.; Duncan, K. R., Antibiotics from rare actinomycetes, beyond the genus Streptomyces. *Current Opinion in Microbiology* **2023,** *76*, 102385.
49. Jensen, P. R., Natural products and the gene cluster revolution. *Trends in microbiology* **2016,** *24* (12), 968-977.
50. Rutledge, P. J.; Challis, G. L., Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nature reviews microbiology* **2015,** *13* (8), 509-523.
51. Doroghazi, J. R.; Albright, J. C.; Goering, A. W.; Ju, K.-S.; Haines, R. R.; Tchalukov, K. A.; Labeda, D. P.; Kelleher, N. L.; Metcalf, W. W., A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nature chemical biology* **2014,** *10* (11), 963-968.
52. Alam, K.; Mazumder, A.; Sikdar, S.; Zhao, Y.-M.; Hao, J.; Song, C.; Wang, Y.; Sarkar, R.; Islam, S.; Zhang, Y., Streptomyces: The biofactory of secondary metabolites. *Frontiers in Microbiology* **2022,** *13*, 968053.
53. Zhang, S.; Chen, Y.; Zhu, J.; Lu, Q.; Cryle, M. J.; Zhang, Y.; Yan, F., Structural diversity, biosynthesis, and biological functions of lipopeptides from Streptomyces. *Natural Product Reports* **2023,** *40* (3), 557-594.
54. Risdian, C.; Mozef, T.; Wink, J., Biosynthesis of polyketides in Streptomyces. *Microorganisms* **2019,** *7* (5), 124.
55. Le, T.; van der Donk, W. A., Mechanisms and evolution of diversity-generating RiPP biosynthesis. *Trends in Chemistry* **2021,** *3* (4), 266-278.
56. Hudson, G. A.; Mitchell, D. A., RiPP antibiotics: biosynthesis and engineering potential. *Current opinion in microbiology* **2018,** *45*, 61-69.
57. Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; Van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y., New developments in RiPP discovery, enzymology and engineering. *Natural product reports* **2021,** *38* (1), 130-239.
58. Caswell, B. T.; de Carvalho, C. C.; Nguyen, H.; Roy, M.; Nguyen, T.; Cantu, D. C., Thioesterase enzyme families: Functions, structures, and mechanisms. *Protein Science* **2022,** *31* (3), 652-676.
59. Bozhüyük, K. A.; Linck, A.; Tietze, A.; Kranz, J.; Wesche, F.; Nowak, S.; Fleischhacker, F.; Shi, Y.-N.; Grün, P.; Bode, H. B., Modification and de novo design of non-ribosomal peptide synthetases using specific assembly points within condensation domains. *Nature chemistry* **2019,** *11* (7), 653-661.
60. Brown, A. S.; Calcott, M. J.; Owen, J. G.; Ackerley, D. F., Structural, functional and evolutionary perspectives on effective re-engineering of non-ribosomal peptide synthetase assembly lines. *Natural product reports* **2018,** *35* (11), 1210-1228.
61. Süssmuth, R. D.; Mainz, A., Nonribosomal peptide synthesis—principles and prospects. *Angewandte Chemie International Edition* **2017,** *56* (14), 3770-3821.
62. Hertweck, C., The biosynthetic logic of polyketide diversity. *Angewandte Chemie International Edition* **2009,** *48* (26), 4688-4716.

63. Calcott, M. J.; Ackerley, D. F., Genetic manipulation of non-ribosomal peptide synthetases to generate novel bioactive peptide products. *Biotechnology letters* **2014,** *36*, 2407-2416.
64. Niu, W.; Liu, J.; Duan, Y.; Zhong, L.; Pang, L.; Zhong, G.; Zhang, Y.; Bian, X., Biosynthesis of Nonribosomal Peptides Chitinimides Reveal a Special Type of Thioesterase Domains. *Chemistry–A European Journal* **2024,** *30* (69), e202402763.
65. Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H., Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC evolutionary biology* **2007,** *7*, 1-15.
66. He, R.; Zhang, J.; Shao, Y.; Gu, S.; Song, C.; Qian, L.; Yin, W.-B.; Li, Z., Knowledge-guided data mining on the standardized architecture of NRPS: Subtypes, novel motifs, and sequence entanglements. *PLOS Computational Biology* **2023,** *19* (5), e1011100.
67. Dang, T.; Süssmuth, R. D., Bioactive peptide natural products as lead structures for medicinal use. *Accounts of chemical research* **2017,** *50* (7), 1566-1576.
68. Kohli, G. S.; John, U.; Van Dolah, F. M.; Murray, S. A., Evolutionary distinctiveness of fatty acid and polyketide synthesis in eukaryotes. *The ISME Journal* **2016,** *10* (8), 1877-1890.
69. Ray, L.; Moore, B. S., Recent advances in the biosynthesis of unusual polyketide synthase substrates. *Natural product reports* **2016,** *33* (2), 150-161.
70. Herbst, D. A.; Townsend, C. A.; Maier, T., The architectures of iterative type I PKS and FAS. *Natural product reports* **2018,** *35* (10), 1046-1069.
71. Khosla, C.; Tang, Y.; Chen, A. Y.; Schnarr, N. A.; Cane, D. E., Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu. Rev. Biochem.* **2007,** *76* (1), 195-221.
72. Chen, A.; Re, R. N.; Burkart, M. D., Type II fatty acid and polyketide synthases: deciphering protein–protein and protein–substrate interactions. *Natural product reports* **2018,** *35* (10), 1029-1045.
73. Shimizu, Y.; Ogata, H.; Goto, S., Type III polyketide synthases: functional classification and phylogenomics. *ChemBioChem* **2017,** *18* (1), 50-65.
74. Yu, D.; Xu, F.; Zeng, J.; Zhan, J., Type III polyketide synthases in natural product biosynthesis. *IUBMB life* **2012,** *64* (4), 285-295.
75. Sottorff, I.; Künzel, S.; Wiese, J.; Lipfert, M.; Preußke, N.; Sönnichsen, F. D.; Imhoff, J. F., Antitumor anthraquinones from an Easter Island Sea Anemone: animal or bacterial origin? *Marine drugs* **2019,** *17* (3), 154.
76. Bisht, R.; Bhattacharyya, A.; Shrivastava, A.; Saxena, P., An overview of the medicinally important plant type III PKS derived polyketides. *Frontiers in plant science* **2021,** *12*, 746908.
77. Gober, R.; Wheeler, R.; Rohr, J., Post-PKS enzyme complexes. *MedChemComm* **2019,** *10* (11), 1855-1866.
78. Walsh, C. T., Tailoring enzyme strategies and functional groups in biosynthetic pathways. *Natural product reports* **2023,** *40* (2), 326-386.
79. Pandith, S. A.; Ramazan, S.; Khan, M. I.; Reshi, Z. A.; Shah, M. A., Chalcone synthases (CHSs): the symbolic type III polyketide synthases. *Planta* **2020,** *251* (1), 15.
80. Wang, Q.; Liu, N.; Deng, Y.; Guan, Y.; Xiao, H.; Nitka, T. A.; Yang, H.; Yadav, A.; Vukovic, L.; Mathews, I. I., Triepoxide formation by a flavin-dependent monooxygenase in monensin biosynthesis. *Nature communications* **2023,** *14* (1), 6273.
81. Otsuka, R.; Sato, Y.; Okano, K.; Okamura, E.; Tomita, H.; Honda, K.; Kitani, S., Identification of a critical gene involved in the biosynthesis of the polyene macrolide lavencidin in Streptomyces lavendulae FRI-5 using the Target-AID (activation-induced cytidine deaminase) base editing technology. *Applied and Environmental Microbiology* **2025,** *91* (5), e00975-24.

82. Wang, J.; Zhang, R.; Chen, X.; Sun, X.; Yan, Y.; Shen, X.; Yuan, Q., Biosynthesis of aromatic polyketides in microorganisms using type II polyketide synthases. *Microbial cell factories* **2020,** *19* (1), 110.
83. Rivers, M. A.; Lowell, A. N., Expanding the biosynthetic toolbox: the potential and challenges of in vitro type II polyketide synthase research. *SynBio* **2024,** *2* (1), 85-111.
84. Shen, B.; Du, L.; Sanchez, C.; Edwards, D.; Chen, M.; Murrell, J., The biosynthetic gene cluster for the anticancer drug bleomycin from Streptomyces verticillus ATCC15003 as a model for hybrid peptide–polyketide natural product biosynthesis. *Journal of Industrial Microbiology and Biotechnology* **2001,** *27* (6), 378-385.
85. Paulus, C.; Myronovskyi, M.; Zapp, J.; Rodríguez Estévez, M.; Lopatniuk, M.; Rosenkränzer, B.; Palusczak, A.; Luzhetskyy, A., Miramides A–D: Identification of detoxin-like depsipeptides after heterologous expression of a hybrid NRPS-PKS gene cluster from Streptomyces mirabilis Lu17588. *Microorganisms* **2022,** *10* (9), 1752.
86. Wang, H.; Sivonen, K.; Fewer, D. P., Genomic insights into the distribution, genetic diversity and evolution of polyketide synthases and nonribosomal peptide synthetases. *Current opinion in genetics & development* **2015,** *35*, 79-85.
87. Jenke-Kodama, H.; Sandmann, A.; Müller, R.; Dittmann, E., Evolutionary implications of bacterial polyketide synthases. *Molecular biology and evolution* **2005,** *22* (10), 2027-2039.
88. Gago, G.; Diacovich, L.; Arabolaza, A.; Tsai, S.-C.; Gramajo, H., Fatty acid biosynthesis in actinomycetes. *FEMS microbiology reviews* **2011,** *35* (3), 475-497.
89. Hu, F.; Cai, W.; Lin, J.; Wang, W.; Li, S., Genetic engineering of the precursor supply pathway for the overproduction of the n C14-surfactin isoform with promising MEOR applications. *Microbial Cell Factories* **2021,** *20* (1), 96.
90. Bentley, S. D.; Chater, K. F.; Cerdeño-Tárraga, A.-M.; Challis, G. L.; Thomson, N.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D., Complete genome sequence of the model actinomycete Streptomyces coelicolor A3 (2). *nature* **2002,** *417* (6885), 141-147.
91. Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Ōmura, S., Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. *Nature biotechnology* **2003,** *21* (5), 526-531.
92. Mohite, O. S.; Jørgensen, T. S.; Booth, T. J.; Charusanti, P.; Phaneuf, P. V.; Weber, T.; Palsson, B. O., Pangenome mining of the Streptomyces genus redefines species' biosynthetic potential. *Genome Biology* **2025,** *26* (1), 9.
93. Ziemert, N.; Alanjary, M.; Weber, T., The evolution of genome mining in microbes–a review. *Natural product reports* **2016,** *33* (8), 988-1005.
94. Medema, M. H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J. B.; Blin, K.; De Bruijn, I.; Chooi, Y. H.; Claesen, J.; Coates, R. C., Minimum information about a biosynthetic gene cluster. *Nature chemical biology* **2015,** *11* (9), 625-631.
95. Bauman, K. D.; Butler, K. S.; Moore, B. S.; Chekan, J. R., Genome mining methods to discover bioactive natural products. *Natural product reports* **2021,** *38* (11), 2100-2129.
96. Lee, N.; Hwang, S.; Kim, J.; Cho, S.; Palsson, B.; Cho, B.-K., Mini review: Genome mining approaches for the identification of secondary metabolite biosynthetic gene clusters in Streptomyces. *Computational and Structural Biotechnology Journal* **2020,** *18*, 1548-1556.
97. Blin, K.; Shaw, S.; Augustijn, H. E.; Reitz, Z. L.; Biermann, F.; Alanjary, M.; Fetter, A.; Terlouw, B. R.; Metcalf, W. W.; Helfrich, E. J., antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic acids research* **2023,** *51* (W1), W46-W50.
98. Navarro-Muñoz, J. C.; Selem-Mojica, N.; Mullowney, M. W.; Kautsar, S. A.; Tryon, J. H.; Parkinson, E. I.; De Los Santos, E. L.; Yeong, M.; Cruz-Morales, P.; Abubucker, S., A

computational framework to explore large-scale biosynthetic diversity. *Nature chemical biology* **2020,** *16* (1), 60-68.
99. Skinnider, M. A.; Johnston, C. W.; Gunabalasingam, M.; Merwin, N. J.; Kieliszek, A. M.; MacLellan, R. J.; Li, H.; Ranieri, M. R.; Webster, A. L.; Cao, M. P., Comprehensive prediction of secondary metabolite structure and biological activity from microbial genome sequences. *Nature communications* **2020,** *11* (1), 6058.
100. Otani, H.; Udwary, D. W.; Mouncey, N. J., Comparative and pangenomic analysis of the genus Streptomyces. *Scientific reports* **2022,** *12* (1), 18909.
101. Huo, L.; Hug, J. J.; Fu, C.; Bian, X.; Zhang, Y.; Müller, R., Heterologous expression of bacterial natural product biosynthetic pathways. *Natural product reports* **2019,** *36* (10), 1412-1436.
102. Baltz, R. H., Synthetic biology, genome mining, and combinatorial biosynthesis of NRPS-derived antibiotics: a perspective. *Journal of Industrial Microbiology and Biotechnology* **2018,** *45* (7), 635-649.
103. 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. *Metabolic engineering* **2018,** *49*, 316-324.
104. 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. *Microbial cell factories* **2020,** *19*, 1-16.
105. Kim, H. U.; Charusanti, P.; Lee, S. Y.; Weber, T., Metabolic engineering with systems biology tools to optimize production of prokaryotic secondary metabolites. *Natural product reports* **2016,** *33* (8), 933-941.
106. Xu, M.; Wright, G. D., Heterologous expression-facilitated natural products' discovery in actinomycetes. *Journal of Industrial Microbiology and Biotechnology* **2019,** *46* (3-4), 415-431.
107. Yamanaka, K.; Reynolds, K. A.; Kersten, R. D.; Ryan, K. S.; Gonzalez, D. J.; Nizet, V.; Dorrestein, P. C.; Moore, B. S., Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proceedings of the National Academy of Sciences* **2014,** *111* (5), 1957-1962.
108. Kang, H.-S.; Kim, E.-S., Recent advances in heterologous expression of natural product biosynthetic gene clusters in Streptomyces hosts. *Current Opinion in Biotechnology* **2021,** *69*, 118-127.
109. Crits-Christoph, A.; Diamond, S.; Butterfield, C. N.; Thomas, B. C.; Banfield, J. F., Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. *Nature* **2018,** *558* (7710), 440-444.
110. Hemmerling, F.; Piel, J., Strategies to access biosynthetic novelty in bacterial genomes for drug discovery. *Nature Reviews Drug Discovery* **2022,** *21* (5), 359-378.
111. Li, L., Accessing hidden microbial biosynthetic potential from underexplored sources for novel drug discovery. *Biotechnology Advances* **2023,** *66*, 108176.
112. Zhang, J. J.; Tang, X.; Moore, B. S., Genetic platforms for heterologous expression of microbial natural products. *Natural product reports* **2019,** *36* (9), 1313-1332.
113. Nah, H.-J.; Pyeon, H.-R.; Kang, S.-H.; Choi, S.-S.; Kim, E.-S., Cloning and heterologous expression of a large-sized natural product biosynthetic gene cluster in Streptomyces species. *Frontiers in microbiology* **2017,** *8*, 394.
114. Zerikly, M.; Challis, G. L., Strategies for the discovery of new natural products by genome mining. *ChemBioChem* **2009,** *10* (4), 625-633.
115. Medema, M. H.; Fischbach, M. A., Computational approaches to natural product discovery. *Nature chemical biology* **2015,** *11* (9), 639-648.

116. Zhang, G.; Lin, M.; Qin, M.; Xie, Q.; Liang, M.; Jiang, J.; Dai, H.; Xu, S.; Feng, S.; Liao, M., Establishing heterologous production of microcins J25 and Y in Bacillus subtilis. *Journal of Agricultural and Food Chemistry* **2023,** *71* (14), 5600-5613.
117. Viel, J. H.; Jaarsma, A. H.; Kuipers, O. P., Heterologous expression of mersacidin in Escherichia coli elucidates the mode of leader processing. *ACS Synthetic Biology* **2021,** *10* (3), 600-608.
118. Yang, H.; Qu, J.; Zou, W.; Shen, W.; Chen, X., An overview and future prospects of recombinant protein production in Bacillus subtilis. *Applied Microbiology and Biotechnology* **2021,** *105* (18), 6607-6626.
119. Klumbys, E.; Xu, W.; Koduru, L.; Heng, E.; Wei, Y.; Wong, F. T.; Zhao, H.; Ang, E. L., Discovery, characterization, and engineering of an advantageous Streptomyces host for heterologous expression of natural product biosynthetic gene clusters. *Microbial cell factories* **2024,** *23* (1), 149.
120. Liu, J.; Wang, X.; Dai, G.; Zhang, Y.; Bian, X., Microbial chassis engineering drives heterologous production of complex secondary metabolites. *Biotechnology Advances* **2022,** *59*, 107966.
121. Hwang, S.; Lee, Y.; Kim, J. H.; Kim, G.; Kim, H.; Kim, W.; Cho, S.; Palsson, B. O.; Cho, B.-K., Streptomyces as microbial chassis for heterologous protein expression. *Frontiers in bioengineering and biotechnology* **2021,** *9*, 804295.
122. Okamoto-Hosoya, Y.; Okamoto, S.; Ochi, K., Development of antibiotic-overproducing strains by site-directed mutagenesis of the rpsL gene in Streptomyces lividans. *Applied and environmental microbiology* **2003,** *69* (7), 4256-4259.
123. Rückert, C.; Albersmeier, A.; Busche, T.; Jaenicke, S.; Winkler, A.; Friðjónsson, Ó. H.; Hreggviðsson, G. Ó.; Lambert, C.; Badcock, D.; Bernaerts, K., Complete genome sequence of Streptomyces lividans TK24. *Journal of biotechnology* **2015,** *199*, 21-22.
124. Ahmed, Y.; Rebets, Y.; Tokovenko, B.; Brötz, E.; Luzhetskyy, A., Identification of butenolide regulatory system controlling secondary metabolism in Streptomyces albus J1074. *Scientific Reports* **2017,** *7* (1), 9784.
125. Zaburannyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A., Insights into naturally minimised Streptomyces albus J1074 genome. *BMC genomics* **2014,** *15*, 1-11.
126. Wang, W.; Zheng, G.; Lu, Y., Recent advances in strategies for the cloning of natural product biosynthetic gene clusters. *Frontiers in bioengineering and biotechnology* **2021,** *9*, 692797.
127. Widdick, D.; Royer, S. F.; Wang, H.; Vior, N. M.; Gomez-Escribano, J. P.; Davis, B. G.; Bibb, M. J., Analysis of the tunicamycin biosynthetic gene cluster of Streptomyces chartreusis reveals new insights into tunicamycin production and immunity. *Antimicrobial agents and chemotherapy* **2018,** *62* (8), 10.1128/aac. 00130-18.
128. Lopatniuk, M.; Myronovskyi, M.; Luzhetskyy, A., Streptomyces albus: a new cell factory for non-canonical amino acids incorporation into ribosomally synthesized natural products. *ACS chemical biology* **2017,** *12* (9), 2362-2370.
129. Hwang, S.; Joung, C.; Kim, W.; Palsson, B.; Cho, B.-K., Recent advances in non-model bacterial chassis construction. *Current Opinion in Systems Biology* **2023,** *36*, 100471.

2. Results

2.1 Identification and Heterologous Expression of a NRPS Biosynthetic Gene Cluster Responsible for the Production of the Pyrazinones Ichizinones A, B and C

Patrick Oberhäuser, Maksym Myronovskyi, Marc Stierhof, Oleksandr Gromyko and Andriy Luzhetskyy

Microbial Cell Factories (2025) 24:131

DOI: https://doi.org/10.1186/s12934-025-02753-6

Published online: 07 June 2025

2.1.1 Abstract

Pyrazinones are a growing family of microbial NRPS-derived natural products showing interesting biological activities. These compounds are characterized by the presence of either a di- or trisubstituted heterocyclic, nonaromatic 2(1 H)-pyrazinone core in their structure. The most commonly occurring disubstituted pyrazinone natural products are synthesized through a dipeptide intermediate, which is further cyclized to yield the pyrazinone moiety. Trisubstituted pyrazinones are seldom found in natural products, with JBIR56 and JBIR57, isolated from marine *Streptomyces*, being notable examples. In contrast to the simply organized disubstituted pyrazinones, JBIR56 and JBIR57 are synthesized as tetrapeptides with unnatural beta-amino acid residue involved in the formation of the pyrazinone moiety. Despite interesting structural features, biosynthetic routes leading to the production of these compounds have not been reported yet. Here we report the discovery of new members of trisubstituted pyrazinone family–tetrapeptides ichizinones A-C in *Streptomyces* sp. LV45-129. Through sequence analysis and heterologous expression, a biosynthetic gene cluster encoding ichizinone production was identified. Based on gene annotation and sequence homology, a biosynthetic model was suggested. The presented results provide insights into the biosynthesis of rare trisubstituted pyrazinone natural products.

Keywords: Actinobacteria, *Streptomyces*, Pyrazinone, NRPS, Heterologous expression, Gene cluster, Biosynthesis

2.1.2 Introduction

Pyrazinones represent a small family of natural products structurally related to diketopiperazines[1, 2]. The hallmark of these compounds is the six-membered heterocyclic 2(1 H)-pyrazinone ring (**1**) in their structure (Fig. 1)[1]. Pyrazinones are widely distributed across various domains of life, including fungi, staphylococci, streptomycetes, myxobacteria, and several marine sponge species[3–12]. From a structural perspective, pyrazinones can be classified into two subgroups: the more commonly occurring disubstituted 2(1 H)-pyrazinones and the less frequently found trisubstituted pyrazinones[7, 8, 13]. The disubstituted compounds include deoxyaspergillic acid and flavacol isolated from *Aspergillus flavus*, leuvalin (**2**), phevalin, and tyrvalin (**3**) from *Staphylococcus* species, phileucin, arglecin, and argvalin (**4**) from *Streptomyces*, as well as dragmacidin D and ma'edamines A and B from deep-water marine sponge species (Fig. 1)[2–4, 9, 11, 12, 14, 15]. The trisubstituted pyrazinones include enhypyrazinones A

and B, coralinones A and B, and sorazinone B (**5**) from various myxobacteria, butrepyrazinone (**6**) from *Verrucosispora* sp. K51G, and JBIR-56 (**7**) and JBIR-57 (**8**) from *Streptomyces* sp. SpD081030SC-03 (Fig. 1)[5–8, 10].

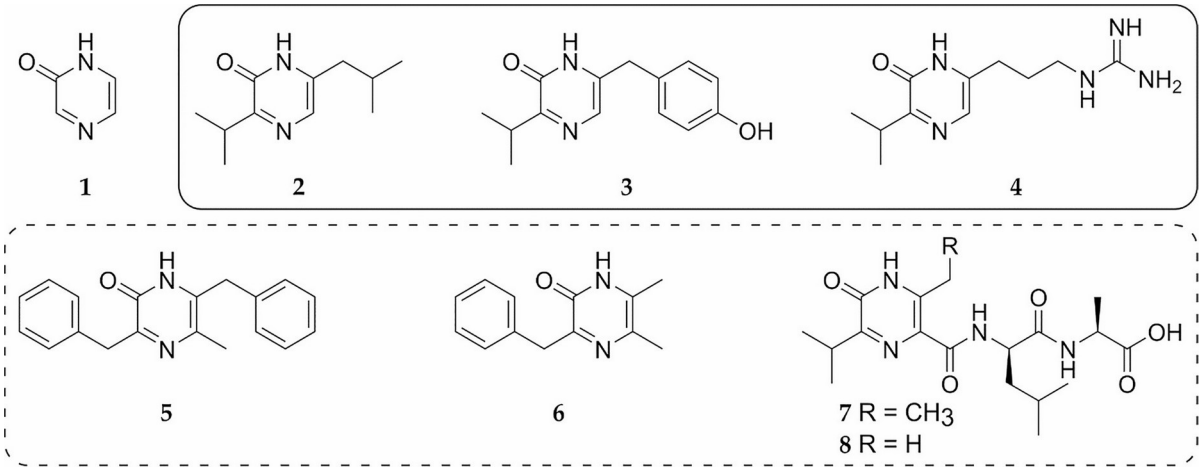


Figure 1: Representatives of the pyrazinone family of natural products: 1– 2(1 H)-pyrazinone core structure (**1**), 2– leuvalin (**2**), 3– tyrvalin (**3**), 4– argvalin (**4**), 5– sorazinone B (**5**), 6– butrepyrazinone (**6**), 7– JBIR-56 (**7**), 8– JBIR-57 (**8**). The disubstituted pyrazinone natural products are enclosed within a solid-line rectangle, while the trisubstituted pyrazinone natural products are enclosed within a dashed-line rectangle

Biosynthetically, the majority of pyrazinones are condensation products of two amino acids. Typically, a dimodular nonribosomal peptide synthetase (NRPS) is involved in synthesizing the dipeptide precursor, which cyclizes upon release to form the pyrazinone moiety[2, 6, 11]. However, in contrast to the widespread, simply organized dipeptide-derived compounds, the pyrazinone family also includes more complexly synthesized members. For example, JBIR-56 (**7**) and JBIR-57 (**8**) are assembled from four amino acid residues. An unnatural beta-amino acid appears to play a role in forming the pyrazinone moiety of these compounds, which is further modified by attachment to a dipeptide chain[8]. Unfortunately, despite their intriguing structural features, the biosynthesis of JBIR-56 (**7**) and JBIR-57 (**8**) has not yet been elucidated.

In this study, we report the discovery of three new members of the trisubstituted pyrazinone family: ichizinones A (**9**), B (**10**), and C (**11**). These compounds were identified in extracts of *Streptomyces* sp. LV45-129 and are structurally closely related to JBIR-56 (**7**) and JBIR-57 (**8**). The ichizinones were isolated, and their structures were elucidated using nuclear magnetic resonance spectroscopy, MS/MS analysis, and Marfey's analysis. To identify the biosynthetic gene cluster encoding the production of these compounds, we performed bioinformatics analysis of the genome sequence of the producing strain and confirmed the findings through heterologous expression in *Streptomyces albus* Del14. DNA deletion experiments and sequence analysis enabled us to propose a biosynthetic pathway leading to ichizinone production.

2.1.3 Results and Discussion

Identification and isolation of ichizinones A, B and C in the culture broth of *Streptomyces* sp. LV45-129

The strain *S.* sp. LV45-129 was thoroughly investigated as part of a broader screening for new bioactive metabolites. For this purpose, the strain was fermented in the production medium DNPM. The culture broth was then extracted with butanol, and the resulting extract analyzed using high-resolution liquid chromatography mass spectrometry (LC-MS). Besides the identification of known metabolites such as puromycin and pamamycin, three peaks were detected at the retention times of 8.8 min, 9.2 min and 10.6 min, corresponding to the compounds with $[M + H]^+$ 409.244 m/z, $[M + H]^+$ 423.258 m/z and $[M + H]^+$ 485.273 m/z respectively (Fig. 2, Fig. S1). A search in the natural product database for these high-resolution masses did not generate any matches, implying that the identified compounds might be new.

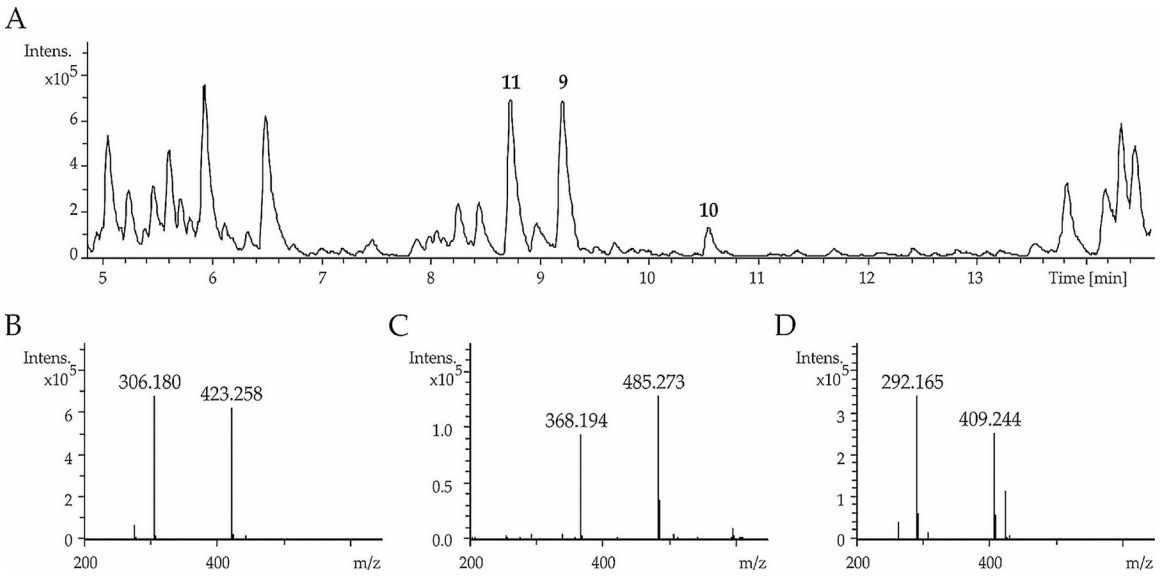


Figure 2: LC-MS detection of ichizinones. **A**– Base peak chromatogram of crude extract from *S.* sp. LV45-129. Peaks corresponding to ichizinones A (**9**), B (**10**), and C (**11**) are marked with the numbers 9, 10 and 11, respectively. **B**, **C**, and **D**– Mass spectra of the peaks corresponding to ichizinones A (**9**), B (**10**), and C (**11**), respectively

In order to isolate the identified compounds, we performed a large scale cultivation of the strain *S.* sp. LV45129. The strain was inoculated into 10 L of DNPM production medium, and the culture broth was subsequently extracted with butanol. The compounds were first separated from contaminants using size-exclusion chromatography and then further purified through preparative reverse-phase chromatography. With this procedure, we obtained 1.4 mg of the compound with $[M + H]^+$ 423.258 m/z, 1 mg of the compound with $[M + H]^+$ 485.273 m/z, and 4.4 mg of the compound with $[M + H]^+$ 409.244. The isolated compounds were named

ichizinone A (**9**), B (**10**) and C (**11**), respectively. To gain insights into the structure of the ichizinones, the purified compounds were used for the structure elucidation experiments. Additionally, the isolated compounds were subjected to biological activity studies to assess their potential antimicrobial properties. The ichizinones showed no activity against the tested Gram-positive and Gram-negative bacteria or the fungal strains.

Structure elucidation of ichizinone A, B and C

The molecular formula of ichizinone A (**9**) was determined to be $C_{21}H_{34}N_4O_5$ based on HRMS data ($[M + H]^+$, m/z 423.258), indicating that the compound is a peptide. The structure of ichizinone A (**9**) was elucidated using 1D and 2D NMR experiments and corroborated by MS fragmentation data (Table 1, Fig. 3, Fig. S2–S9, S26). For NMR analysis, the compound was thoroughly dried and measured in DMSO and DMSO containing traces of TFA to achieve full protonation of NH groups. The latter resulted in sharper signals due to reduced proton exchange. The data quality generally improved upon the addition of TFA, however, two carbon signals at δ_H 122.59 and δ_H 157.85 vanished and could only be observed in TFA free DMSO (Fig. S9). The amino signals at δ_H 8.23 (2'-NH) and δ_H 8.41 (2''-NH) showing COSY correlations to the methines H-2' (δ_H 4.69) and H-2'' (δ_H 4.17) indicated the presence of two amino acids. Analyzing the spin system starting from the α-methine proton at H-2' together and the aliphatic proton signals at δ_H 1.53 (CH_2), 1.57 (CH), 0.92 (CH_3) and 0.89 (CH_3) revealed the amino acid leucine. The spin system starting from the α-methine proton at H-2'', followed by δ_H 2.07 (CH), 0.86 (CH_3) and 0.87 (CH_3) was assigned to valine. An HMBC correlation of the carbonyl C-1' (δ_C 172.59) of leucine to 2''-NH of valine showed that the two amino acids are connected via a peptide bond, while the carbonyl group of valine did not show any correlation, revealing a carboxylic acid terminus. The remaining two spin systems were assigned based on COSY correlations of H-7, H-8 and H-9 revealing an isopropyl group and H-11 and H-12 revealing an ethyl group. HMBC correlations of H-11 to C-6 (δ_C 146.09), C-5 (δ_C 122.59) and N-1 (δ_N 175.38), and H-7(δ_H 1.16) to C-2 (δ_C 157.85), C-3 (δ_C 158.92) and N-4 (δ_N 317.96) suggests that the isopropyl and the ethyl group are attached to a pyrazinone heterocycle. A database search of the suggested final structure elements revealed strong similarities to compound JBIR-56 (**7**) [8]. Correlations in the ^{1}H-^{13}CHMBC from 1-NH to 2-C or 3-C and in the ^{15}N-HMBC to either 1-NH or 4 N could not observed. However, alignment of ^{1}H and ^{13}C shifts of ichizinone A (**9**) and JBIR-56 (**7**) and analysis of MS/ MS fragmentation patterns (Fig. S26) strongly indicated the structural moieties are arranged in a pyrazinone moiety (Fig. 3; Table 1). The elucidated compound ichizinone A (**9**) was identified as a novel natural compound.

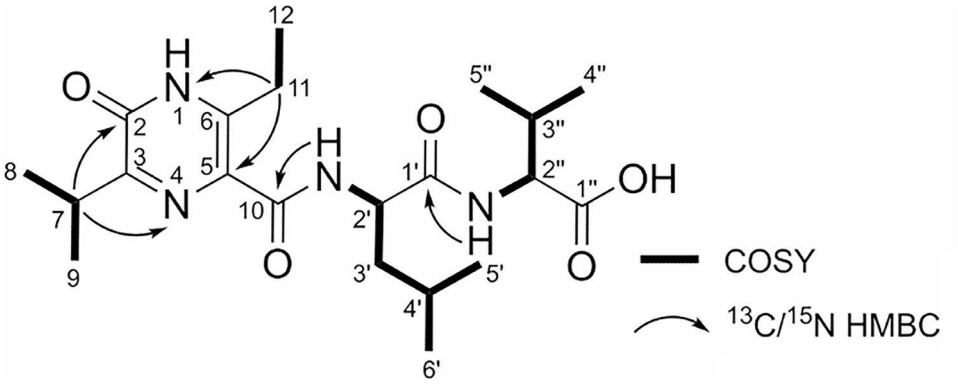


Figure 3: $^{13}C/^{15}N$ HMBC ( ) and COSY ( ) key correlations of ichizinone A.

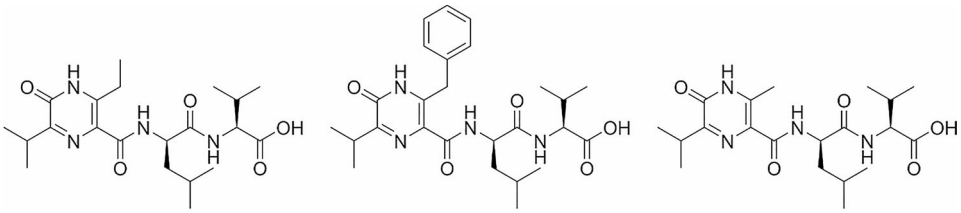
The molecular formula of ichizinone B (**10**) was determined as $C_{26}H_{36}N_4O_5$ based on the HRMS data ($[M + H]^+$, m/z 485.273). The structure of ichizinone B (**10**), as shown in Fig. 4, was determined by 1D and 2D NMR experiments (Fig. S10–S17, S27, Table S4). The ^{1}H-NMR spectrum revealed aromatic methines δH 7.34 (H-13/17), 7.25 (H-14/26) and 7.20 (H-15) which were assigned to a benzyl group. The benzyl group was found to be attached to the pyrazinone moiety at C-6, while the core structure of ichizinone B (**10**) is the same as that of ichizinone A (**9**). The structure of the compound was further confirmed through MS/MS fragmentation analysis. The observed fragmentation pattern of ichizinone B (**10**) closely corresponded to that of ichizinone A (**9**), with the only differences arising from the distinct substituents at position C-6 of the pyrazinone moiety (Fig. S26–S27). The elucidated compound ichizinone B (**10**) was identified as a novel natural compound.

The molecular formula of ichizinone C (**11**) was determined as $C_{20}H_{32}N_4O_5$ based on the HRMS data ($[M + H]^+$, m/z 409.244). The structure of ichizinone C (**11**) (Fig. 4) was determined by 1D and 2D NMR experiments (Fig. S18–S25, S28, Table S5). This compound is structurally similar to other ichizinones, differing only by the presence of a methyl group attached to the pyrazinone moiety at C-6, instead of an ethyl or benzyl group as found in ichizinones A (**9**) and B (**10**), respectively. The absolute stereochemistry of the amino acids not incorporated in the pyrazinone moiety was determined for ichizinone C (**11**) using Marfey's method[20], revealing L-valine and D-leucine as the constituent amino acids (Fig. S29).

Table 1*:* NMR data of ichizinone A (**9**) in DMSO-d_6 and JBIR-56 for comparison.

	Ichizinone A (**9**)				JBIR-56 (**7**)	
No, type	δ(13C, 15N) [ppm]	δ(1H) [ppm], mult(J)	COSY (H-)	HMBC (C-/N-)	δ(13C, 15N) [ppm]	δ(1H) [ppm], mult(J)
1-NH	175.38			11		
2-C*	157.85*			7*	155.9	
3-C	158.92			7, 8, 9	158.7	
4-N	317.96			7		
5-C*	122.59*			11*	120.3	
6-C	146.09			11, 12	145.9	
7-CH	29.94	3.23, m	8, 9	8, 9	29.5	3.22, q (6.6)
8-CH3	20.25	1.16, ovl**	7	7, 9	19.9	1.15, d (6.6)
9-CH3	20.33	1.16, ovl**	7	7, 8	19.8	1.15, d (6.6)
10-C	163.62			2'-NH	163.3	
11-CH2	23.35	3.02, m 2.93, m	12	12	23.2	2.99, dq (12.6, 7.2) 2.93, dq (12.6, 7.2)
12-CH3	14.27	1.13, t (7.5)	11	11	14.0	1.13, (7.2)
Leu					Leu	
1'-C	172.59			2', 3', 2''-NH	171.4	
2'-CH	51.18	4.69, dt (7.5, 8.4)	3', 2'-NH	3', 2'-NH	50.7	4.52, dd (13.8, 8.4)
3'-CH2	43.10	1.53, m	2', 4'	2', 5', 6'	42.6	1.52, dd (13.8, 6.0)
4'-CH	25.22	1.57, m	3', 5', 6'	2', 3', 5', 6'	24.7	1.56, d (6.0)
5'-CH3	22.72	0.92, d (5.8)	4'	3', 6'	23.1	0.89, d (6.0)
6'-CH3	23.66	0.89, d (5.8)	4'	3', 5'	22.4	0.88, d (6.0)
2'-NH	113.98	8.23, d (8.7)	2'	3'		8.18, d (8.4)
Val					Ala	
1''-C	173.43	-		2"	174.0	
2''-CH	57.52	4.17, dd (6.7, 8.7)	2''-NH, 3"	3", 4", 5", 2''-NH	48.1	4.12, dq (7.2, 6.6)
3''-CH	30.53	2.07, m	2", 4", 5"	2", 4", 5"	17.7	1.23, d (7.2)
4''-CH3	18.34	0.86, ovl**	3"	2", 3", 5"		
5''-CH3	19.90	0.87, ovl**	3"	2", 3", 4"		
2''-NH	114.71	8.41, d (8.6)	2"	3"		8.34, br s

*ovl. = overlap with other signals



9 **10** **11**

Figure 4: Structures of the isolated compounds: ichizinone A (**9**), ichizinone B (**10**), ichizinone C (**11**).

Identification of the ichizinone gene cluster through heterologous expression in *Streptomyces* albus Del14

The isolated ichizinones are pyrazinones, each consisting of four amino acid residues. The pyrazinone moiety of ichizinones A (**9**), B (**10**), and C (**11**) is derived from a valine residue and a beta-amino acid residue, which varies among the isolated compounds: 3-amino-pentanoic acid, 3-amino-4-phenylbutanoic acid, and 3-aminobutanoic acid, respectively. The free carboxyl group of the beta-amino acid is extended by the attachment of a dipeptide composed of leucine and valine residues. The pyrazinone moiety of natural products is typically assembled by a small dimodular nonribosomal peptide synthetase[2, 6, 11]. However, the tetrapeptide nature of ichizinones suggests the involvement of a larger NRPS system composed of four modules. Genome sequence analysis of the ichizinone producer *S.* sp. LV45-129 using the antiSMASH software revealed 38 putative gene clusters associated with secondary metabolite production, including eight involved in the biosynthesis of NRPStype products[21]. The detailed analysis of NRPS clusters identified one cluster as the most likely candidate for encoding ichizinone production. This cluster encodes a four-module NRPS with adenylation domains that show specificity for the amino acids valine, phenylalanine, and leucine, some of which are present in the ichizinone structures. To confirm that the identified cluster is responsible for ichizinone production, a genome library of the producer strain was constructed on the integrative cosmid vector cos15A_gus-AmInt and subsequently sequenced[23]. The cosmid E514, containing the entire NRPS cluster presumably involved in ichizinone production, was transferred into the heterologous host strain *Streptomyces albus* Del14 via conjugation[24]. The exconjugant strain *S. albus* E514 and the corresponding control strain, *S. albus* Del14, which lacks the cosmid, were fermented in DNPM production medium. The culture filtrates of both strains were extracted with butanol, and the resulting extracts were analyzed using LC-MS. This analysis identified ichizinones A (**9**), B (**10**), and C (**11**) in the extract of *S. albus* E514 (Fig. 5). The retention times and m/z values of the detected peaks were identical to those observed in the extract of the natural producer strain *S.* sp. LV45-129. In contrast, the extract of the control strain, *S. albus* Del14, showed no peaks corresponding to ichizinones. These results clearly demonstrate that the cosmid E514 contains the entire biosynthetic gene cluster for the production of ichizinones.

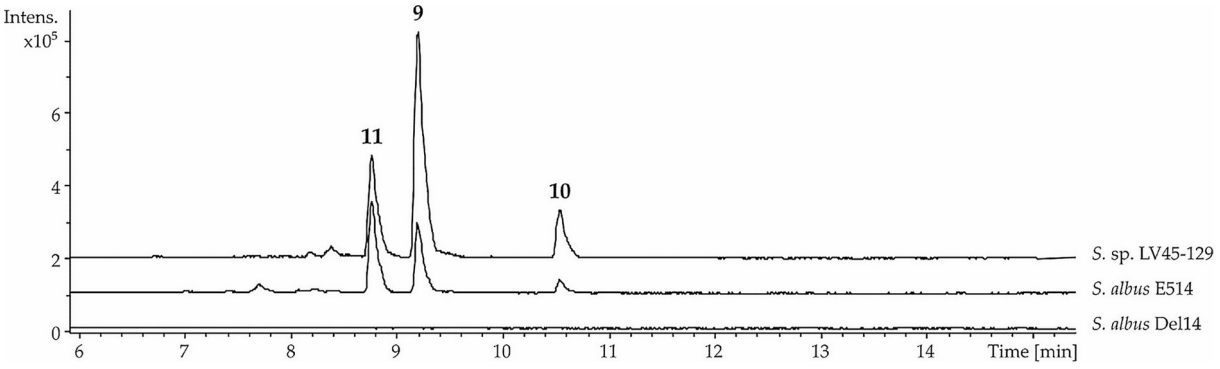


Figure 5: LC-MS analysis of ichizinone production by the heterologous host *S. albus* E514. Extracted ion chromatograms (409.3 ± 0.5 Da, 423.3 ± 0.5 Da, 485.3 ± 0.5 Da) of crude extracts from *S.* sp. LV45-129, *S. albus* E514, and *S. albus* Del14 are shown. Peaks corresponding to ichizinones A (**9**), B (**10**), and C (**11**) are marked with the numbers 9, 10, and 11, respectively

The chromosome fragment cloned in the cosmid E514 spans a 52.6 kb DNA region and contains 40 genes (accession number PQ885478). Sequence analysis of this DNA fragment was conducted to define the borders of the ichizinone cluster. A set of 16 genes, *ichA–ichP*, was predicted to be involved in the biosynthesis of the compound (Table 2, Fig. 6). These genes encode four nonribosomal peptide synthetases, three regulatory proteins, three transport proteins, an MbtH-like protein, a cyclase, a thioesterase, a type I polyketide synthase, a monooxygenase, and a hypothetical protein. ClusterBlast analysis indicates that these genes are clustered together across various *Streptomyces* species, suggesting they constitute the ichizinone biosynthetic cluster. The 5' upstream region of *ichA* contains genes encoding a Ku protein and two ATP-dependent DNA ligases– proteins involved in DNA repair mechanisms and not in ichizinone production[25,26]. Similarly, the genes located downstream of *ichP* encode an arginyl-tRNA and an arginyl-tRNA synthetase, which are also unrelated to ichizinone biosynthesis. These findings further support the hypothesis that the genes *ichA– ichP* comprise the ichizinone biosynthetic cluster. However, since the inactivation of the genes flanking the *ichA–ichP* region was not performed, the possibility that additional genes are involved in ichizinone biosynthesis cannot be excluded.

Figure 6: The chromosomal fragment of *S.* sp. LV45-129 containing the ichizinone biosynthetic gene cluster.

Table 2: Proposed functions of genes within the ichizinone gene cluster.

Gene	Proposed Function	GeneBank homologue
orf17	ATP-dependent DNA ligase	WP_055529559.1
orf18	ATP-dependent DNA ligase	WP_055529557.1
orf19	KU protein	WP_055529555.1
IchA	MbtH family NRPS accessory protein	WP_150477341.1
ichB	MFS Transporter	WP_150477342.1
ichC	NRPS	WP_055529551.1
ichD	FMN-dependent luciferase-like monooxygenase	WP_055529549.1
ichE	NRPS	WP_167532726.1
ichF	NRPS	WP_150477345.1
ichG	NRPS	WP_150477346.1
ichH	TypeI Polyketide synthase	WP_150477347.1
ichI	Thioesterase	WP_055529989.1
ichJ	Polyketide cyclase	WP_055529992.1

ichK	ABC transporter permease	WP_055529993.1
ichL	ABC transporter ATP-binding protein	WP_055529995.1
ichM	Sensor histidine kinase	WP_055529997.1
ichN	Response regulator	WP_055530004.1
ichO	LysR family transcriptional regulator	WP_055530006.1
ichP	Hypothetical protein	WP_055530008.1
orf36	Arg tRNA	WP_246201615.1
orf37	Response regulator	WP_234336420.1
orf38	Arginyl-tRNA synthetase	WP_030792088.1

To confirm the involvement of *ichA– ichP* genes in ichizinone production, the genes *ichH*, *ichI*, and *ichJ*, which encode a type I polyketide synthase, a thioesterase, and a cyclase, respectively, were deleted in cosmid E514. The resulting recombinant constructs were expressed in the heterologous host *S. albus* J1074. Inactivation of the polyketide synthase led to a complete cessation of ichizinone production, while inactivation of the other two genes resulted in severely impaired production, with only trace amounts of the compounds being detected (Fig. 7). These findings corroborate that the identified set of genes is involved in ichizinone biosynthesis

Insights into the biosynthesis of ichizinones

Ichizinones isolated in this study belong to the group of pyrazinone natural products. The most commonly occurring pyrazinones are small cyclic dipeptides composed of two proteinogenic amino acids[2–7, 11, 12, 15]. Their biosynthesis involves a dimodular NRPS that forms a dipeptide precursor, which is cyclized upon release to yield a disubstituted six-membered heterocyclic 2(1 H)-pyrazinone ring (**1**)[2, 11]. The substituents are positioned para to each other and correspond to the side chains of the amino acids involved in the ring formation (Fig. 1). In contrast to these dipeptide pyrazinones, the ichizinones produced by *S.* sp. LV45-129 are more complex (Fig. 4). These compounds are tetrapeptides composed of three proteinogenic amino acid residues and one nonproteinogenic beta-amino acid residue in the second position. The N-terminal amino acid residue and the second nonproteinogenic amino acid residue participate in forming the trisubstituted pyrazinone ring. The mechanism of pyrazinone ring formation of ichizinones appears to differ from that of disubstituted dipeptide pyrazinones. In the latter, ring closure is proposed to occur through the interaction of the free amino group of the N-terminal amino acid and the carboxyl group of the C-terminal amino acid[2, 11]. In ichizinones, however, the pyrazinone ring formation likely involves the interaction of the amino group of the N-terminal amino acid residue with the alpha carbon of the second nonproteinogenic beta-amino acid residue. Consequently, the carboxyl group of this beta-amino acid becomes the third substituent of the pyrazinone ring and serves as the attachment site for the remaining two amino acid residues

(Fig. 4). As with disubstituted pyrazinones, the side chains of the amino acid residues involved in pyrazinone ring formation form the first and second substituents of the trisubstituted pyrazinone ring in ichizinones. Structurally, the isolated ichizinones strongly resemble compounds JBIR-56 (**7**) and JBIR-57 (**8**) (Fig. 1), which were identified in 2011 from a marine sponge-derived *Streptomyces sp.* SpD081030SC-03[8]. However, the biosynthesis of these compounds has not been published. The results of sequence analysis and heterologous expression experiments demonstrated that, as in the case of disubstituted pyrazinones, an NRPS is involved in the biosynthesis of ichizinones. The ichizinone biosynthetic cluster was predicted to comprise 16 genes, *ichA– ichP*, of which only eight encode structural biosynthetic enzymes. The genes *ichC*, *ichE*, *ichF*, and *ichG* encode individual modules of a four-module NRPS (Fig. 8), while *ichH* encodes a single-module type I polyketide synthase.

The genes *ichD* and *ichI* encode a flavin-dependent oxidoreductase and a type II thioesterase, respectively. The product of the *ichJ* gene exhibits low sequence similarity to proteins of the SRPBCC family with uncharacterized functions and to members of the polyketide cyclase/ dehydrase family. The closest characterized homologs of IchJ are the aromatase/cyclase proteins ZhuI and TcmN from polyketide biosynthetic pathways[27, 28]. The *ichA* gene encodes an MbtH family protein, which is often associated with bacterial NRPS. MbtH proteins are reported to bind noncovalently to the adenylation domains of some NRPS, thereby promoting their folding, stability, and activity[29, 30].

The number of modules in the NRPS encoded by *ichC*, *ichE*, *ichF*, and *ichG* aligns with the length of the ichizinone tetrapeptide precursor. At the first, third, and fourth positions, all ichizinones contain valine, leucine, and valine residues, respectively. Marfey's analysis of the ichizinone C (**11**) hydrolysate demonstrated that the valine residue of its dipeptide tail is in the L-configuration, while the leucine residue is in the D-configuration. The isolated ichizinones differ in the nonproteinogenic amino acid residue at the second position: ichizinones A (**9**), B (**10**), and C (**11**) contain 3-aminopentanoic acid, 3-amino-4-phenylbutanoic acid, and 3-aminobutanoic acid, respectively. According to predictions made by antiSMASH and PKS/NRPS analysis bioinformatics tools, the adenylation domains of IchC, IchE, and IchF exhibit substrate specificity for valine, leucine, and valine, respectively– amino acids that are components of ichizinones[21, 31]. The NRPS module encoded by *ichE* contains an additional epimerization domain, which converts L-leucine to its D-stereoisomer. The substrate specificity of the adenylation domain of IchG could not be clearly predicted. We propose that the NRPS module encoded by *ichG* is responsible for the incorporation of various nonproteinogenic beta-

amino acids found in the structures of ichizinones. According to the domain organization of NRPS modules encoded by *ichC*, *ichE*, *ichF*, and *ichG*, we propose that they form an NRPS assembly line in the following order: *ichC* encodes an initiation module responsible for the incorporation of the N-terminal L-valine residue (Fig. 8). The module encoded by *ichG* acts as the first elongation module, incorporating the nonproteinogenic amino acid. The second elongation module, encoded by *ichE*, incorporates L-leucine and converts it to the D-configuration. Finally, the termination NRPS module, encoded by *ichF*, incorporates the C-terminal L-valine residue. The thioesterase domain of IchG releases the tetrapeptide precursor, which is subsequently processed to yield the ichizinone molecule.

The origin of nonproteinogenic beta-amino acids is one of the most intriguing aspects of ichizinone biosynthesis but remains incompletely understood. We propose that the genes *ichH* and *ichI*, whose deletions result in the cessation of ichizinone production, are involved in the biosynthesis of these beta-amino acids. The *ichH* gene encodes a single-module type I PKS comprising only ketosynthase (KS) and acyl carrier protein (ACP) domains, while *ichI* encodes a type II thioesterase. According to our hypothesis, the PKS utilizes free fatty acyl-CoA esters: acetyl-CoA, propionyl-CoA, and phenylacetyl-CoA as starter units for decarboxylative condensation with malonate (Fig. 8). Acetyl-CoA and propionyl-CoA are common acyl-CoA esters present in bacterial cells. The detection of phenylacetic acid and phenylacetate-CoA ligase activity has also been reported in *Streptomyces*[32–34].

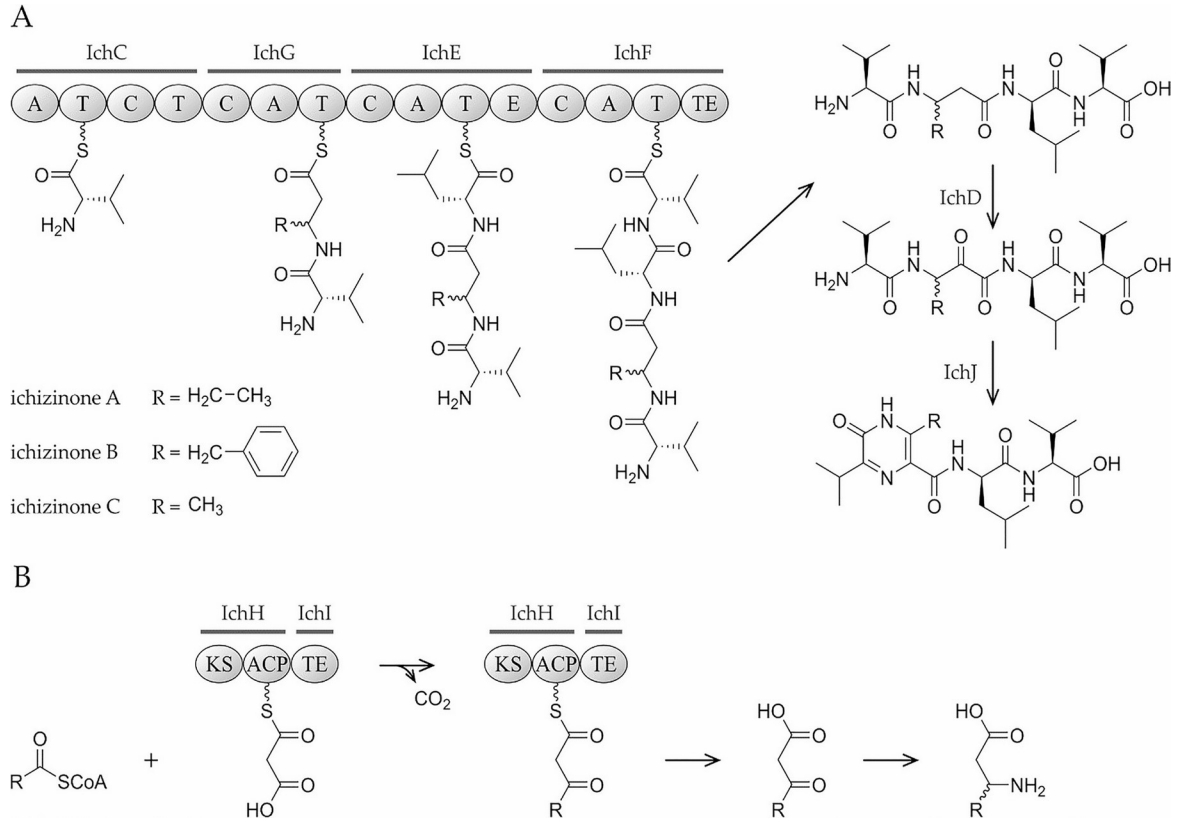


Figure 7: Proposed biosynthetic pathway of ichizinones. A – Assembly of ichizinones by linear NRPS. B – Biosynthesis of beta-amino acid precursors of ichizinones.

Since the PKS lacks a dedicated acyltransferase (AT) domain, we hypothesize that an unidentified trans-acyltransferase, encoded elsewhere in the host strain's genome, transfers the malonyl moiety from malonyl-CoA to the ACP domain of *IchH*. The resulting diketides 3-oxobutanoic acid, 3-oxopentanoic acid, and 3-oxo-4-phenylbutanoic acid are released from the PKS by the action of the thioesterase IchI. We further propose that the resulting beta-keto acids are aminated at the beta position by an unidentified aminotransferase encoded in the host strain's genome, yielding the betaamino acids 3-aminobutanoic acid, 3-aminopentanoic acid, and 3-amino-4-phenylbutanoic acid.

To confirm the involvement of the PKS *IchH* in the biosynthesis of beta-amino acid precursors of ichizinones, the *S. albus* E514 Δ*ichH* strain harboring the ichizinone cluster with the inactivated gene *ichH* was supplemented with beta-amino acids: 3-aminobutanoic acid, 3-aminopentanoic acid, and 3-amino-4-phenylbutanoic acid. However, LC-MS analysis of the culture extracts revealed that the supplementation did not restore ichizinone production. This could be due to inefficient transport of the non-natural amino acids into the cells or disruptions in protein-protein interactions between the PKS and the NRPS assembly line.

The NRPS assembly line encoded by the genes *ichC*, *ichE*, *ichF* and *ichG* produces a linear tetrapeptide precursor, which undergoes further modifications to yield ichizinones. We propose that during this maturation process, the alpha carbon of beta-amino acid residue is oxidized, and the pyrazinone ring is formed through a cyclization reaction. The oxidation reaction might be catalyzed by the flavin-dependent oxidoreductase encoded by the gene *ichD*, while the cyclization may be mediated by the putative cyclase encoded by *ichJ*.

In this paper, we report the isolation of three new members of the pyrazinone family of natural products– ichizinones A (**9**), B (**10**), and C (**11**). The isolated ichizinones are tetrapeptides featuring a rare trisubstituted pyrazinone ring. The biosynthetic gene cluster responsible for the production of these compounds was identified through heterologous expression and gene inactivation studies. The structures of the isolated ichizinones, as well as the sequence of their biosynthetic gene cluster, indicate that the biosynthesis of tetrapeptide, trisubstituted pyrazinones substantially differs from the biosynthetic pathways of smaller dipeptide compounds. The results presented in this paper provide insights into the possible mechanisms involved in the production of ichizinones.

2.1.4 Materials and Methods

General Experimental Procedures

All strains and plasmids used in this study are listed in Tables S1– S2. *Escherichia coli* strains were cultured in Luria-Bertani (LB) medium[16]. *Streptomyces* strains were cultivated on soya flour mannitol agar (MS agar)[17] for sporulation and conjugation 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) was used for cluster expression and secondary metabolite production. When required, the antibiotics kanamycin, apramycin, hygromycin and nalidixic acid were supplemented.

DNA Manipulations

Isolation of cosmids from the constructed genomic library of *S.* sp. LV45-129, DNA manipulations, transformation into *E. coli* and intergeneric conjugation between *E. coli* and *Streptomyces* were performed according to standard protocols[16–18]. Purification of cosmids was done by using the BACMAX ™ DNA purification kit (Lucigen, Middleton, WI, USA). All

restriction endonucleases were used according to manufacturer's recommendations (New England Biolabs, Ipswich, MA, USA).

To determine the involvement of certain genes in the ichizinone biosynthesis, cosmids with the deletion of these genes were constructed. In the cosmid E514_KOA the cyclase gene was inactivated. In the cosmid E5124_ KOB, the thioesterase gene was inactivated and in cosmid E514_KOC, the PKS gene was inactivated. For this purpose, the hygromycin cassette was amplified from the pACS-hyg plasmid with the pairs of primers E514_ KOA_F/E514_KOA_R, E514_KOB_F/E514_KOB_R, E514_KOC_F/E514_KOC_R (Table S3). The obtained PCR fragments were utilized for the construction of the abovementioned cosmids using Red-ET modification[19]. The construction was confirmed using restriction mapping, PCR with the pairs of primers E514_KOA_ chkF/E514_KOA_chkR, E514_KOB_chkF/E514_KOB_ chkR, E514_KOC_chkF/E514_KOC_chkR (Table S3) and sequencing.

Metabolite Extraction and Analysis

All *Streptomyces* strains were precultured in 25 mL of TSB for 24 h before 1 mL of each seed culture was used to inoculate 100 mL of DNPM production medium. Cultures were incubated at 28 °C for seven days. Metabolites were extracted with equal amounts of butanol from the supernatant of the culture broth, evaporated and dissolved in methanol. HPLC-MS analysis was performed by separating 1 µL of the extract using a Dionex UltiMate 3000 UPLC (Thermo Fisher Scientific, Waltham, MA, USA), a 10-cm ACQUITY UPLC BEH C18 column, 1.7 µm (Waters, Milford, MA, USA) and a linear gradient of 0.1% formic acid solution in acetonitrile against 0.1% formic acid solution in water from 5 to 95% in 18 min at a flow rate of 0.6 mL/min. Samples were analyzed using an amaZon speed mass spectrometer or maXis high resolution QTOF system (Bruker, USA). Data were collected and analyzed with the Bruker Compass Data Analysis software, version 4.1 (Bruker, Billerica, MA, USA). The monoisotopic mass was searched in the Dictionary of Natural Products database

Ichizinone Isolation

For preparative metabolite isolation *Streptomyces* strains were precultured in 6 Flasks each containing 25 mL of TSB for 24 h before 100 flasks each containing 100 mL of DNPM were inoculated with 1 mL of the seed culture. Cultures were incubated for 7 days at 28 °C. The mycelial part was separated by centrifugation and metabolites were extracted from the supernatant as described above. The ichizinones were purified using size-exclusion and

subsequent reverse phase chromatography. Size-exclusion chromatography was performed by using a Sephadex LH-20 column (GE Healthcare, USA) and methanol as a solvent. Fractions containing ichizinones were detected by LC-MS analysis, pooled together, evaporated and dissolved in methanol. Subsequently, preparative HPLC was performed on a Waters Autopurification system (Waters, Milford, MA, USA) equipped with a Waters 2545 Binary Gradient module using a Nucleodur C18 HTec column (5 µm, 250 × 21 mm, Macherey-Nagel, Düren, Germany). UV spectra were recorded with a photodiode array detector (Waters 2998, Waters, Milford, MA, USA). A composition of 0.1% formic acid containing water and methanol was used as a mobile phase. Individual peaks were collected and analyzed by LC-MS as described above. Finally, all collected peaks were purified by semipreparative HPLC (Dionex UltiMate 3000, Thermo Fisher Scientific, USA) using a C18 column (Synergi 10 µm, 250 × 10 mm; Phenomenex, Aschaffenburg, Germany).

Ichizinone A (**9**): White powder; 1.4 mg; [α]D20 -25 (c 0.10, MeOH); UV (ACN/ H2O + 0.1% FA) λmax 226, 266 and 316 nm; NMR data, see Table 1; ESI-TOF-MS m/z 423.2597 $[M + H]^+$ (calc. for $C_{21}H_{35}N_4O_5$ 423.2602), see Fig. 2.

Ichizinone B (**10**): White powder; 1.0 mg; [α]D20 -20 (c 0.02, MeOH); UV data not available; NMR data, see Table S4; ESI-TOF-MS m/z 423.2597 $[M + H]^+$ (calc. for $C_{21}H_{35}N_4O_5$ 423.2602), see Fig. 2.

Ichizinone C (**11**): White powder; 4.4 mg; [α]D20 -22 (c 0.44, MeOH); UV (ACN/ H2O + 0.1% FA) λmax 226, 266 and 316 nm; NMR data, see Table S5; ESI-TOF-MS m/z 485.2752 $[M + H]^+$ (calc. for $C_{26}H_{37}N_4O_5$ 485.2758), see Fig. 2.

NMR Data Acquisition and Optical Rotation

The chemical structures of all the compounds were determined via multidimensional NMR analysis. ^{1}HNMR, ^{13}C-NMR, and 2D spectra were recorded either at 700 MHz (1H)/175 MHz (^{13}C) ichizinone A (**9**) or at 500 MHz (1H)/126 MHz (^{13}C) for ichizinones B (**10**) and C (**11**). Experiments were conducted in the Ascend 700 spectrometer using a cryogenically cooled triple resonance probe (Bruker BioSpin, Rheinstetten, Germany) or in the Bruker Avance Neo 500 MHz, equipped with a Prodigy Cryo-probe. Samples were dissolved in methanol-d_3 or dimethyl sulfoxide-d_6. Chemical shifts are reported in ppm relative to tetramethylsilane; the solvent was used as the internal standard. Coupling constants are reported in Hertz (Hz). Multiplicity is

reported with the usual abbreviations (s: singlet, br s: broad singlet, d: doublet, dd: doublet of doublets, t: triplet, dt: doublet of triplets, q: quartet, dq: doublet of quartets, m: multiplet).

Optical rotations were measured using a JASCO P-2000 digital polarimeter (28600 Mary's Ct, Easton, MD, USA)

Marfey's Method

200 µg of ichizinone C (**11**) was hydrolyzed in 100 µL 6 N HCl at 100 °C for 45 min in a closed vial filled with nitrogen. The sample was then dried for 15 min and dissolved in 110 µL of water before 50 µL were transferred to a 1.5 mL Eppendorf tube. Subsequently, 20 µL of 1 N $NaHCO_3$ and 20 µL of 1% L-FDLA or D-FDLA in acetone were added to the hydrolysate[20]. The amino acid standards were prepared in the same way using only L-FDLA. All reaction mixtures were incubated at 40 °C for 2 h at 700 rpm and subsequently quenched with 10 µL 2 N HCL to stop the reaction. The samples were diluted with 300 µL ACN to a total volume of 400 µL, from which 1 µL of each sample was analyzed with a maXis high-resolution LC-QTOF system using aqueous ACN with 0.1% FA and an adjusted gradient of 5–10% for 2 min, 10–25% for 13 min, 25–50% for 7 min and 50–95% for 2 min. Sample detection was carried out at 340 nm.

Genome mining and Bioinformatics Analysis

The *S.* sp. LV45-129 genome was screened for secondary metabolite biosynthetic gene clusters using the antiSMASH online tool[21]. Genetic data was analyzed with the software Geneious prime 2022.2.2[22]. Selected genes were further analyzed using the BLAST online tool from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/) and the Universal Protein Resource (UniProt) (https://www.uniprot.org/blast).

Antimicrobial susceptibility test

Minimum inhibitory concentrations (MICs) were determined according to standard procedures. Single colonies of the tested strains were suspended in cation-adjusted Müller-Hinton broth to achieve a final inoculum of approximately 104 CFU mL^{-1}. Serial dilutions of 409 (0.03 to 64 µg mL^{-1}) were prepared in sterile 96-well plates before the strain suspension was added. Growth inhibition was assessed after overnight incubation (16– 18 h) at 30–37 °C. A panel consisting of the following strains was tested: *B. subtilis* DSM-10, *S. aureus* Newman, *Mycobacterium smegmatis* MC2155, *Citrobacter freundii* DSM-30,039, *E. coli* BW25113 (wt), *E. coli* JW0451-2 (ΔacrB), *Pseudomonas aeruginosa* PA14 DSM19,882, Acinetobacter baumanii DSM-30,008,

Mucor hiemalis DSM-2656, *Pichia anomala DSM-6766*, *Cryptococcus neoformans* DSM-11,959, *Candida albicans* DSM1665, CHO-K1 and HepG2.

Author contributions

P.O., M.M., A. L. designed the experiments. P.O. performed the experiments. M.S. performed and evaluated the NMR analysis. P.O. and M.S. wrote the manuscript. M.M. and A.L. aided in interpreting the results and also worked on the manuscript. All authors have read and agreed to the published version of the manuscript.

2.1.5 References

1. Riesco-Llach G, Planas M, Feliu L, Joule JA. 2(1H)-Pyrazinones from acyclic Building blocks: methods of synthesis and further derivatizations. RSC Adv. 2023;13:1162–84.
2. Zimmermann M, Fischbach MA. A family of Pyrazinone natural products from a conserved nonribosomal peptide synthetase in Staphylococcus aureus. Chem Biol. 2010;17:925–30.
3. Tatsuta K, Tsuchiya T, Someno T, Umezawa S, Umezawa H. Arglecin, a new microbial metabolite isolation and chemical structure. J Antibiot (Tokyo). 1971;24:735–46.
4. Tatsuta K, Fujimoto K, Yamashita M, Tsuchiya T, Umezawa S. Argvalin, a new microbial metabolite: isolation and structure. J Antibiot (Tokyo). 1973;26:606–8.
5. Kyeremeh K, Acquah KS, Camas M, Tabudravu J, Houssen W, Deng H, et al. Butrepyrazinone, a new Pyrazinone with an unusual methylation pattern from a Ghanaian Verrucosispora Sp. K51G. Mar Drugs. 2014;12:5197–208.
6. Zhu L-L, Yang Q, Wang D-G, Niu L, Pan Z, Li S, et al. Deciphering the biosynthesis and physiological function of 5-Methylated Pyrazinones produced by Myxobacteria. ACS Cent Sci. 2024;10:555–68.
7. Zhang F, Braun DR, Rajski SR, DeMaria D, Bugni TS. Enhypyrazinones A and B, Pyrazinone natural products from a Marine-Derived myxobacterium Enhygromyxa Sp. Mar Drugs. 2019;17:698.
8. Motohashi K, Inaba K, Fuse S, Doi T, Izumikawa M, Khan ST, et al. JBIR-56 and JBIR-57, 2(1H)-pyrazinones from a marine Sp.nge-derived *Streptomyces* Sp. SpD081030SC-03. J Nat Prod. 2011;74:1630–5.
9. Kurimoto S, Okamoto A, Seino S, Fromont J, Kobayashi J, Kubota T. Ma'edamines E and F, rare bromotyrosine alkaloids possessing a 1,2,3,5-tetrasubstituted pyridinium moiety from an Okinawan marine Sponge Suberea Sp. Tetrahedron Lett. 2022;103:153985.
10. Jansen R, Sood S, Mohr KI, Kunze B, Irschik H, Stadler M, et al. Nannozinones and Sorazinones, unprecedented Pyrazinones from Myxobacteria. J Nat Prod. 2014;77:2545–52.
11. Wyatt MA, Wang W, Roux CM, Beasley FC, Heinrichs DE, Dunman PM, et al. Staphylococcus aureus nonribosomal peptide secondary metabolites regulate virulence. Science. 2010;329:294–6.
12. Capon RJ, Rooney F, Murray LM, Collins E, Sim ATR, Rostas Ja. Dragmacidins: new protein phosphatase inhibitors from a Southern Australian deep-water marine Sponge, Spongosorites Sp. J Nat Prod. 1998;61:660–2.

13. Caldwell JJ, Veillard N, Collins I, Design. and synthesis of 2(1H)-pyrazinones as inhibitors of protein kinases. Tetrahedron. 2012;68:9713–28.
14. Lebar MD, Mack BM, Carter-Wientjes CH, Wei Q, Mattison CP, Cary JW. Small NRPS-like enzymes in Aspergillus sections Flavi and Circumdati selectively form substituted Pyrazinone metabolites. Front Fungal Biol. 2022;3:1029195.
15. Böhringer N, Gütschow M, König GM, Schäberle TF. Phileucin - A Cyclic dipeptide similar to Phevalin (Aureusimine B) from *Streptomyces coelicolor* M1146. Nat Prod Commun. 2017;12:107–9.
16. Green MR, Sambrook J. Molecular cloning: A laboratory manual (Fourth Edition). Cold Sping Harbor Laboratory; 2012.
17. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA, Practical. *Streptomyces* Genetics. John Innes Centre; 2000.
18. Rebets Y, Kormanec J, Lutzhetskyy A, Bernaerts K, Anné J. Cloning and expression of metagenomic DNA in *Streptomyces lividans* and its subsequent fermentation for optimized production. Methods Mol Biol. 2023;2555:213–60.
19. Muyrers JPP, Zhang Y, Benes V, Testa G, Rientjes JMJ, Stewart AF. ET recombination: DNA engineering using homologous recombination in E. coli. Methods Mol Biol. 2004;256:107–21.
20. Harada K, Fujii K, Hayashi K, Suzuki M, Ikai Y, Oka H. Application of d,l-FDLA derivatization to determination of absolute configuration of constituent amino acids in peptide by advanced Marfey's method. Tetrahedron Lett. 1996;37:3001–4.
21. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. AntiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Res. 2023;51:W46–50.
22. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28:1647–9.
23. Marques F, Luzhetskyy A, Mendes MV. Engineering Corynebacterium glutamicum with a comprehensive genomic library and phage-based vectors. Metab Eng. 2020;62:221–34.
24. 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;49:316–24.
25. Fell VL, Schild-Poulter C. The Ku heterodimer: function in DNA repair and beyond. Mutat Res Rev Mutat Res. 2015;763:15–29.
26. Martin IV, MacNeill SA. ATP-dependent DNA ligases. Genome Biol. 2002;3:REVIEWS3005.
27. Ames BD, Lee M-Y, Moody C, Zhang W, Tang Y, Tsai S-C. Structural and biochemical characterization of ZhuI aromatase/cyclase from the R1128 polyketide pathway. Biochemistry. 2011;50:8392–406.
28. Ames BD, Korman TP, Zhang W, Smith P, Vu T, Tang Y, et al. Crystal structure and functional analysis of tetracenomycin ARO/CYC: implications for cyclization specificity of aromatic polyketides. Proc Natl Acad Sci U S A. 2008;105:5349–54.
29. Zwahlen RD, Pohl C, Bovenberg RAL, Driessen AJM. Bacterial MbtH-like proteins stimulate nonribosomal peptide Synthetase-Derived secondary metabolism in filamentous Fungi. ACS Synth Biol. 2019;8:1776–87.
30. Baltz RH. Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery. J Ind Microbiol Biotechnol. 2011;38:1747–60.
31. Bachmann BO, Ravel J. Chapter 8. Methods for in Silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. Methods Enzymol. 2009;458:181–217.

32. Hwang BK, Lim SW, Kim BS, Lee JY, Moon SS. Isolation and in vivo and in vitro antifungal activity of phenylacetic acid and sodium phenylacetate from *Streptomyces* humidus. Appl Environ Microbiol. 2001;67:3739–45.
33. Burckhardt RM, VanDrisse CM, Tucker AC, Escalante-Semerena JC. New AMP-forming acid:coa ligases from *Streptomyces lividans*, some of which are posttranslationally regulated by reversible lysine acetylation. Mol Microbiol. 2020;113:253–69.
34. Zhang D, Feng Y, Chu M, Dai Y, Jiang L, Li H. Anti-vibriosis bioactive molecules from marine-derived variant *Streptomyces* Sp. ZZ741A. Nat Prod Res. 2024;1–12.

2.1.6 Supplementary Material

Strains and Plasmids

Table S1: Bacterial strains used in this work

Strain	Description	Reference or source
Streptomyces LV45-129	The wild type strain; the source of the ichizinone cluster	
Streptomyces albus Del14	The heterologous host strain; cluster free derivative of the *S. albus* J1074	[1]
S. albus E514	Derivative of *S. albus* Del14 harboring the E514 cosmid	This work
S. albus E514_KOA	Derivative of *S. albus* Del14 harboring the E514_KOA cosmid	This work
S. albus E514_KOB	Derivative of *S. albus* Del14 harboring the E514_KOB cosmid	This work
S. albus E514_KOC	Derivative of *S. albus* Del14 harboring the E514_KOC cosmid	This work
Escherichia coli ET12567 pUB307	Donor strain for intergeneric conjugation	[2]
Escherichia coli DH10β	General cloning strain	[3]

Table S2: Plasmids and cosmids used in this work

Strain	Description	Reference or source
E514	The cosmid containing 40 kb chromosomal fragment of *Streptomyces* LV45-129; contains ichizinone gene cluster	This work
E514_KOA	The derivative of E514 with the deletion of gene A	This work
E514_KOB	The derivative of E514 with the deletion of gene B	This work
E514_KOC	The derivative of E514 with the deletion of gene C	This work
pACS-hyg	The plasmid containing hygromycin resistance gene	[4]

Table S3: Primers used in this work

Primer	Sequence
E514_KOA_F	ATGGCCTCCTCGCTGCTCGAAGCCGTAGACATCAAGGCGCCGGTCGCCGTGTT TAAACAATACTTGACATATCACTGT
E514_KOA_R	CTAGTCCCCGGTGCCTTCGGAGACCGTTCGTTCCGCCATCTTCTTGAATTGTTT AAACTCAGGCGCCGGGGGCGGTGT
E514_KOB_F	GTGGGCGACTGGATACGTTGCTGCCACCCCGCTCCCGACGCCGGGGTCCGGTT TAAACAATACTTGACATATCACTGT
E514_KOB_R	TCATCGCCTCCCGATCGGATCGGGAAGAGAGGAGGCGATGACGTCGGCGAGT TTAAACTCAGGCGCCGGGGGCGGTGT
E514_KOC_F	ATGACCGAGCCAGCACCCCGGGCCTCCGACGCCGACGCCACCCCCGAGCCGT TTAAACAATACTTGACATATCACTGT
E514_KOC_R	CTACTCCTCCGCCTCGAACTCCGCCAGGAGCGCGTCCAGCGAGTCCTCCGGTT TAAACTCAGGCGCCGGGGGCGGTGT
E514_KOA_chkF	TGGCAGCAACGTATCCAGTC
E514_KOA_chkR	GTCCGCCAGCTCCTCAGG
E514_KOB_chkF	GCTCGACGTTCTCCCACAG
E514_KOB_chkR	CGATCACCCTGTTCCTGGAG
E514_KOC_chkF	CACGCTGGTCGACCTGTTC
E514_KOC_chkR	CTTCTCCTGCCACGTCCTG

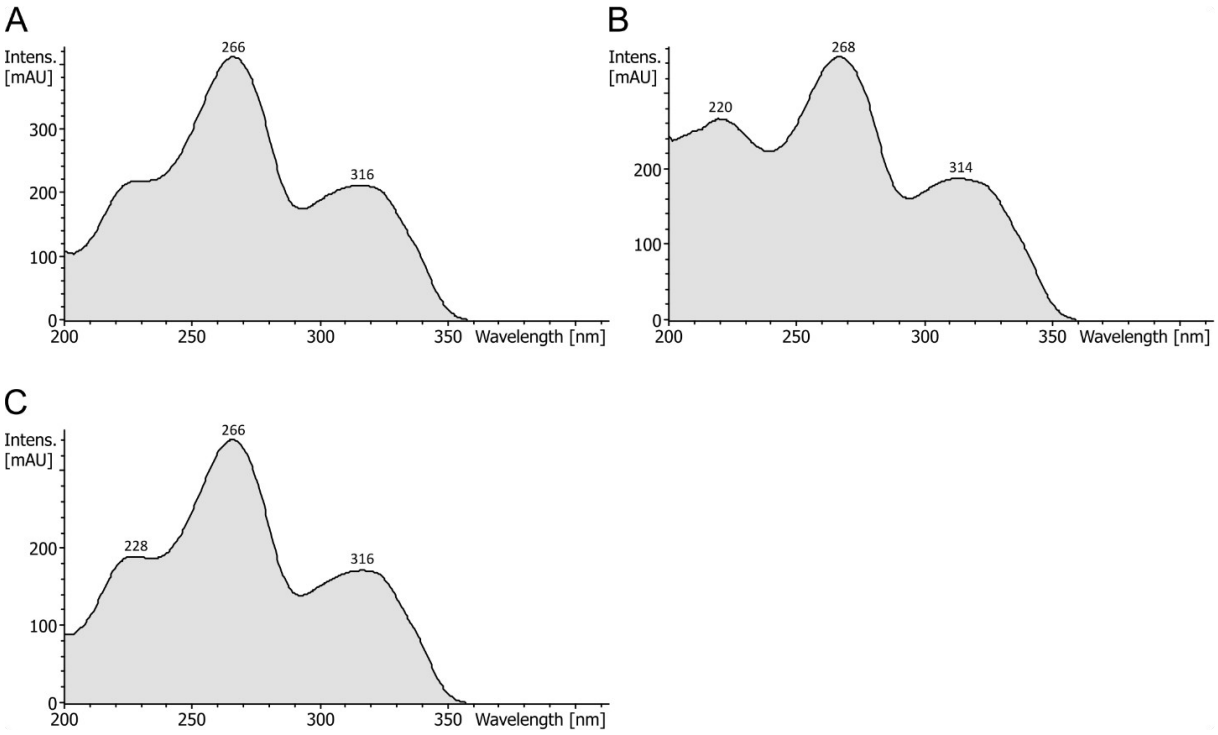


Figure S1: UV-Vis absorption spectra of isolated ichizinones. A – UV-Vis absorption spectrum of ichizinone A, B – UV-Vis absorption spectrum of ichizinone B, C – UV-Vis absorption spectrum of ichizinone C.

NMR spectroscopy

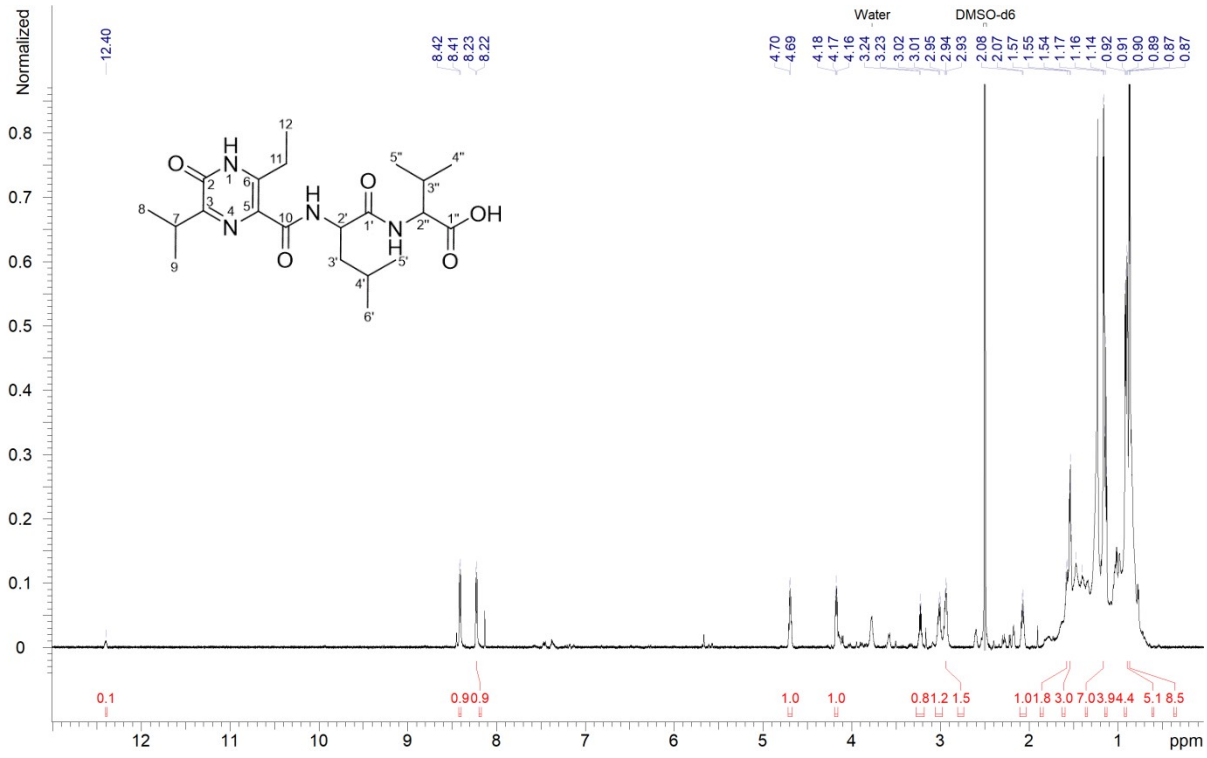


Figure S2: ^{1}H-NMR spectrum of ichizinone A in DMSO-d_6 and TFA.

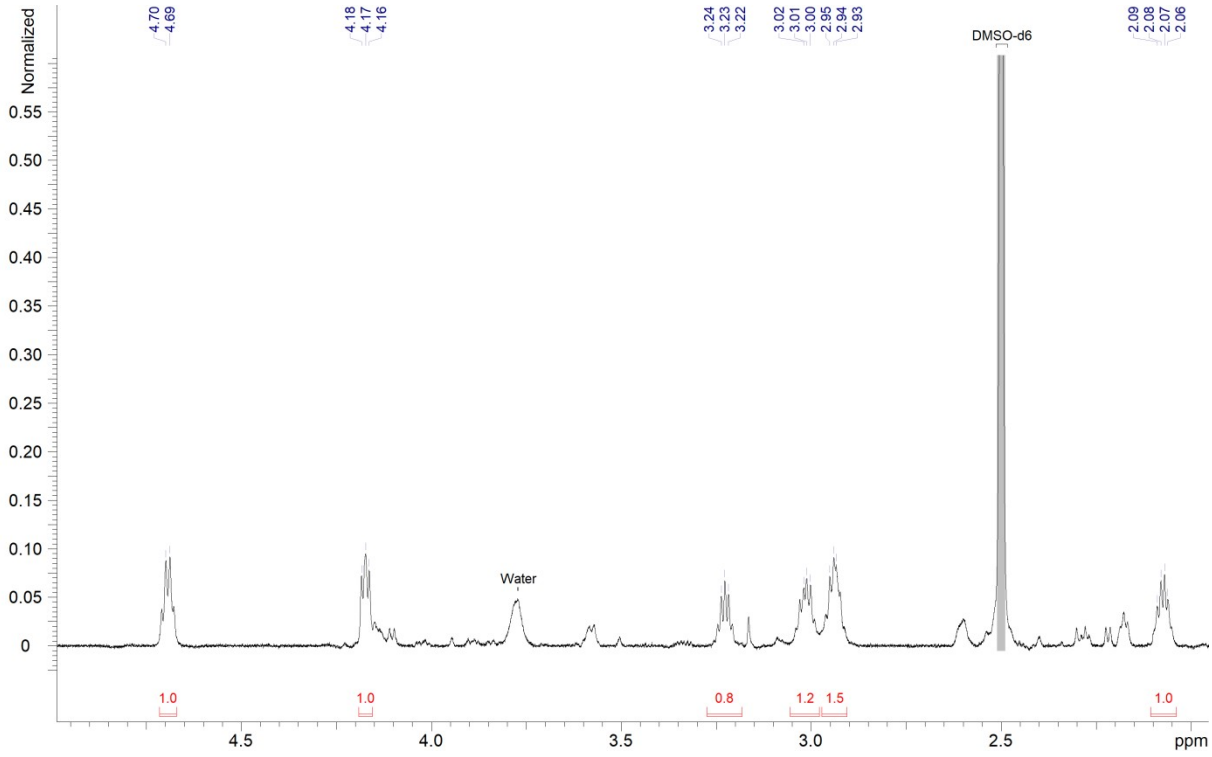


Figure S3: ^{1}H-NMR spectrum of ichizinone A in DMSO-d_6 and TFA

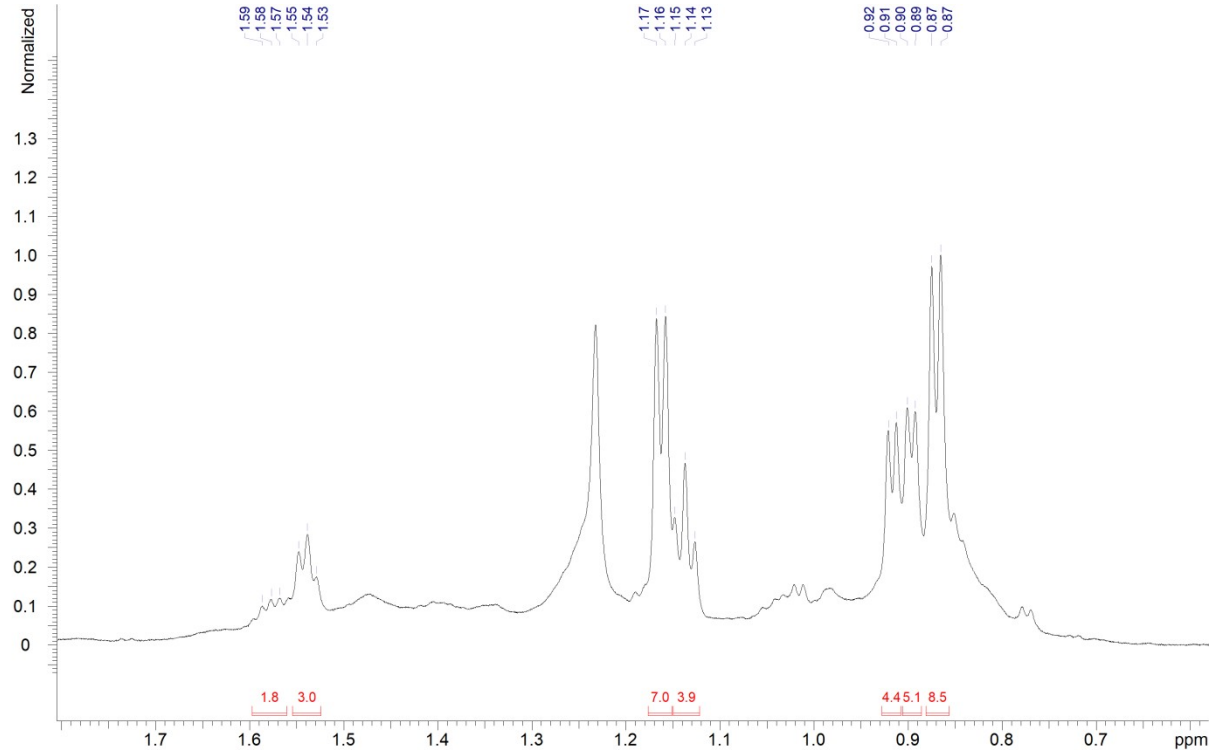


Figure S4: ^{1}H-NMR spectrum of ichizinone A in DMSO-d_6 and TFA.

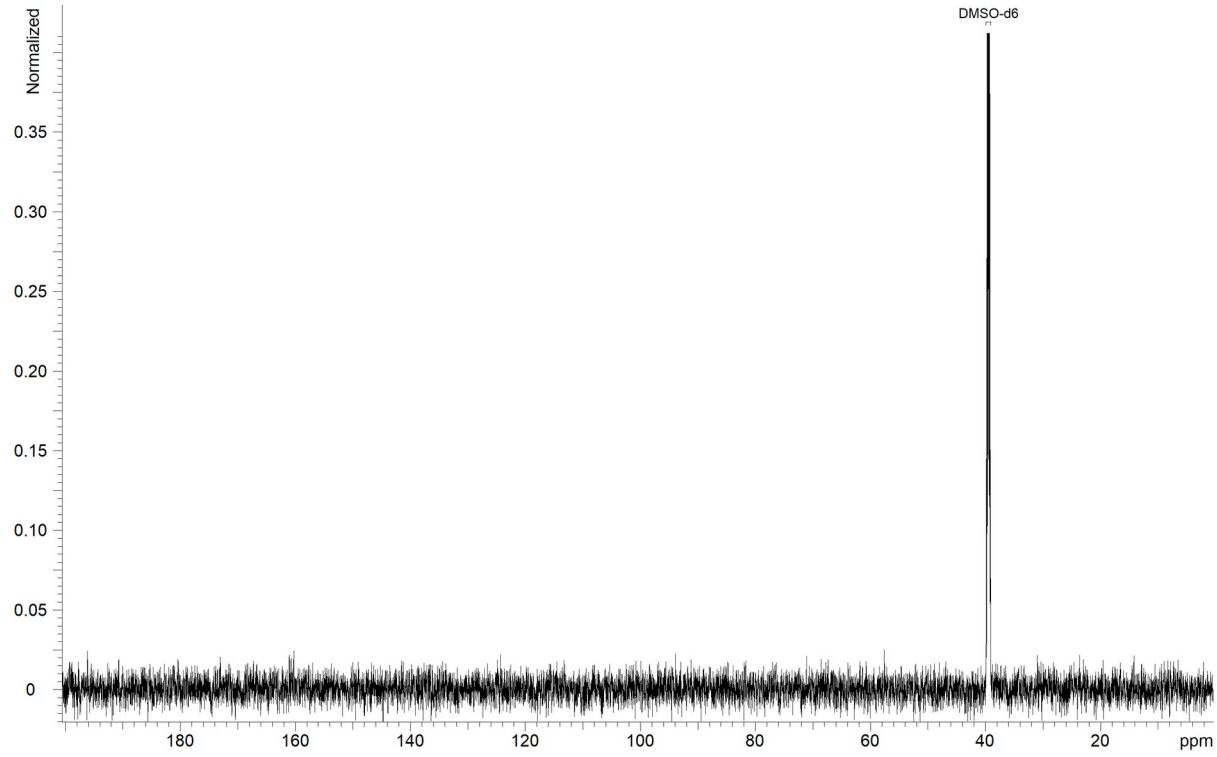


Figure S5: ^{13}C-NMR spectrum of ichizinone A in DMSO-d_6 and TFA

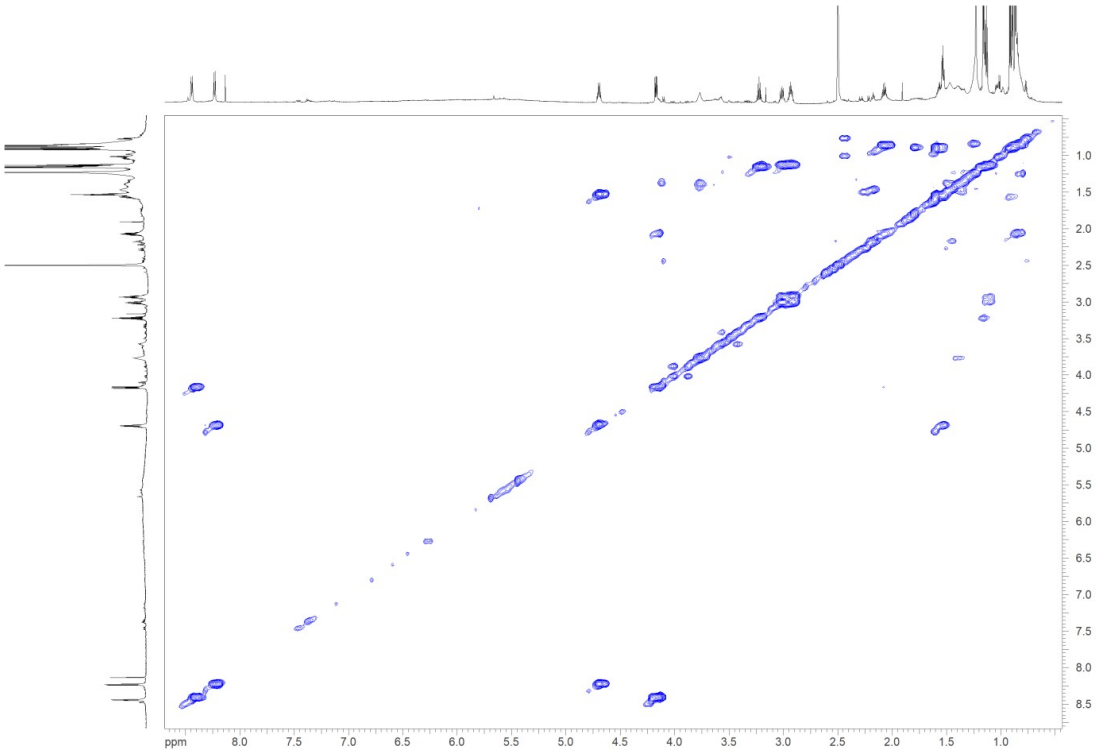


Figure S6: ^{1}H-^{1}H-COSY spectrum of ichizinone A in DMSO-d_6 and TFA.

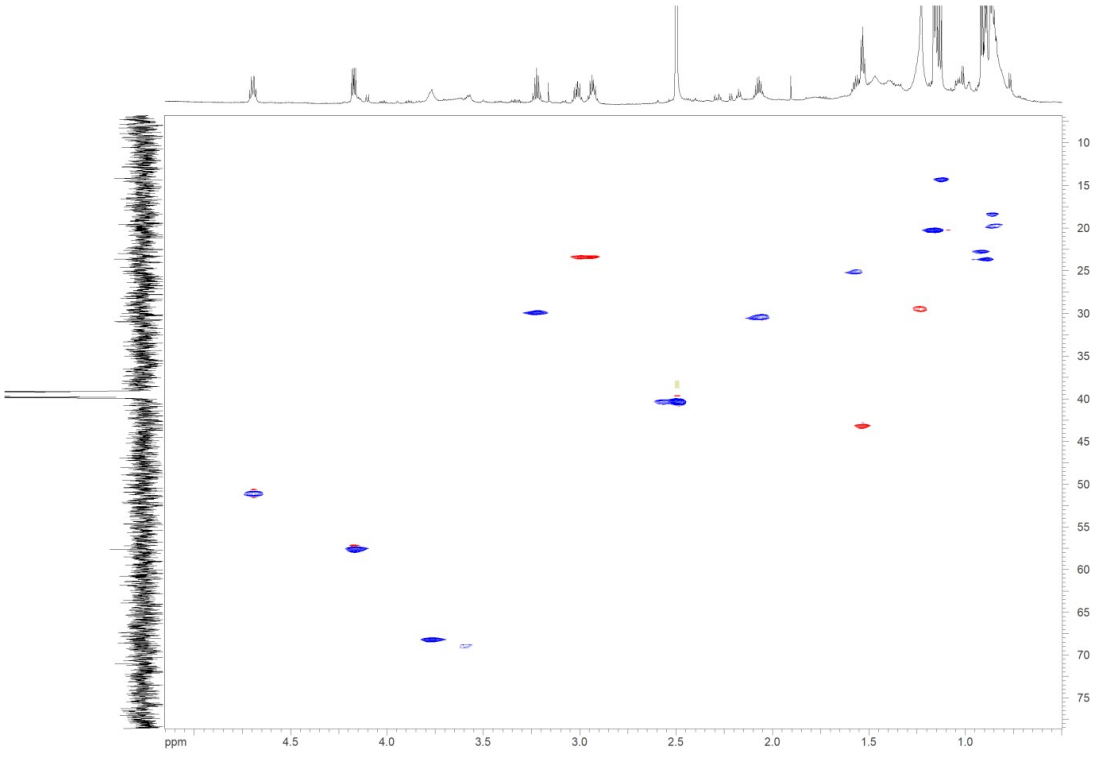


Figure S7: Edited HSQC spectrum of ichizinone A in DMSO-d_6 and TFA.

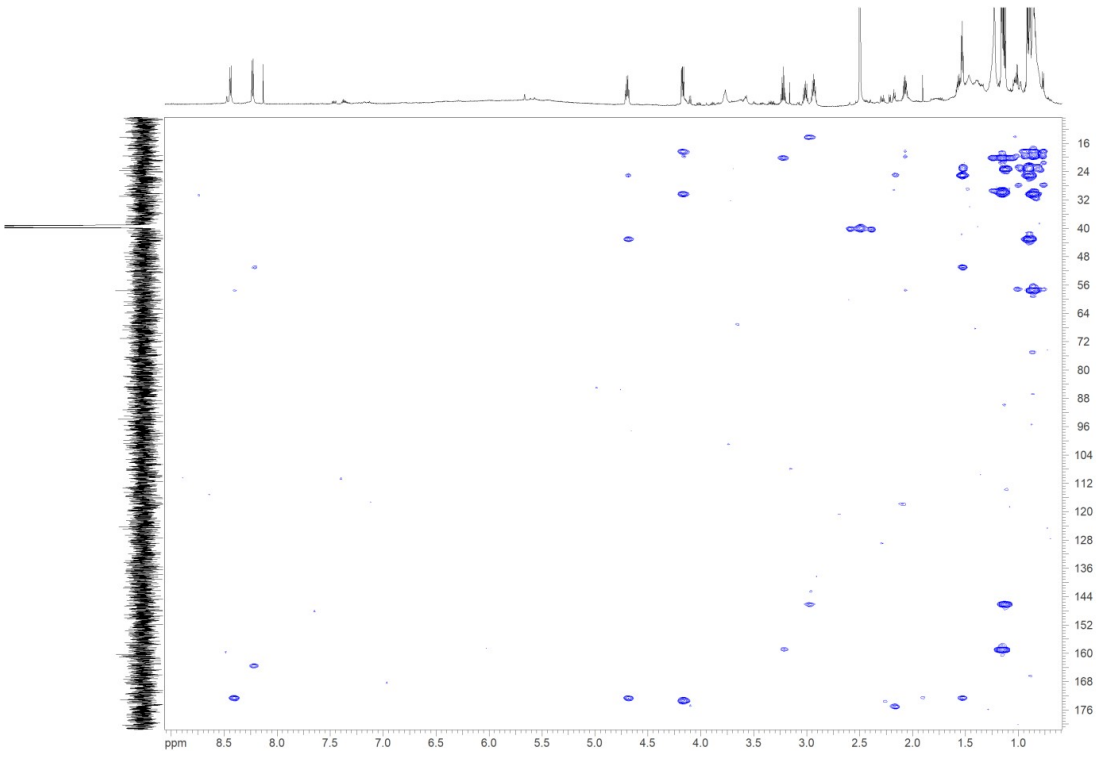


Figure S8: HMBC spectrum of ichizinone A in DMSO-d_6 and TFA.

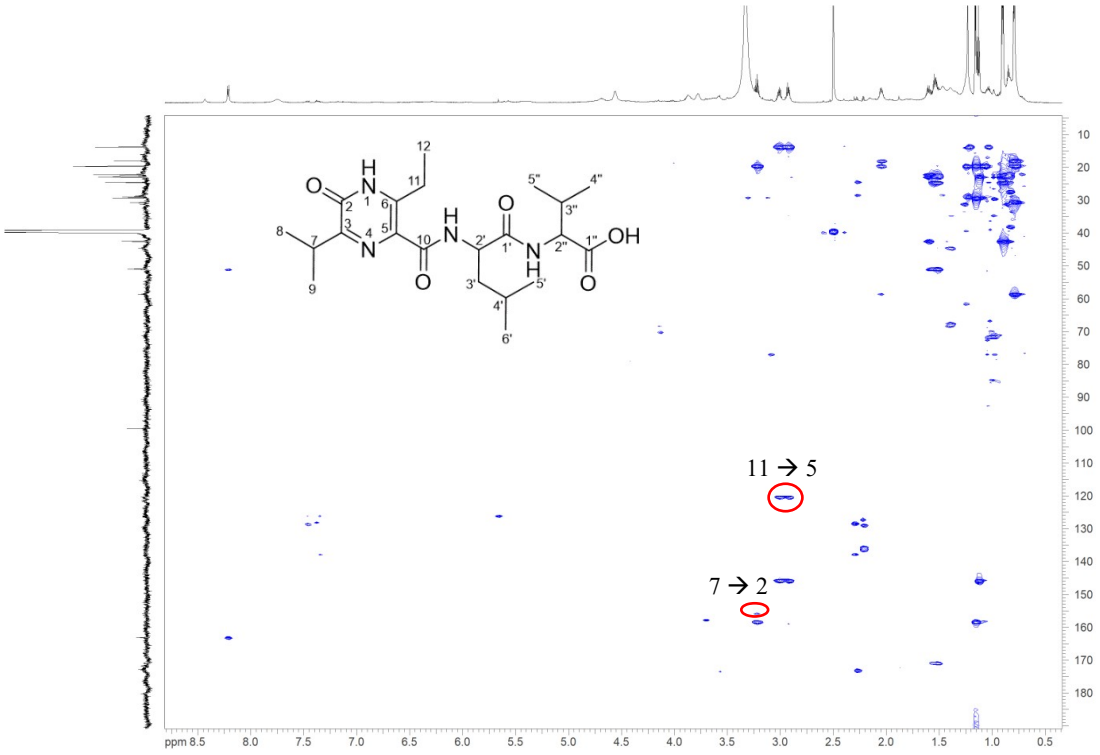


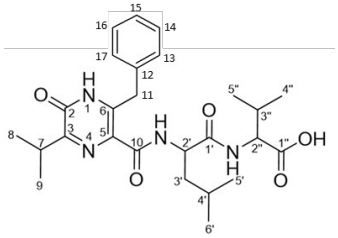
Figure S9: HMBC spectrum of key correlations of ichizinone A in DMSO-d_6; HMBC correlation H7 → C2 and H11 → C5 were not observed when TFA was added.

Table S4: NMR Table of ichizinone B measured in DMSO-d_6.

Ichizinone B

No, type	δ(13C, 15N) [ppm]	δ(1H) [ppm], mult (J)	COSY (H-)	HMBC (C-/N-)
1-NH				
2-C	n.a.*			n.a.*
3-C	158.9			7, 8, 9
4-N				
5-C	120.8			11
6-C	142.5			11
7-CH	29.4	3.22, dt (13.7, 6.9)	8, 9	8, 9
8-CH3	19.7	1.16, d (6.8)	7	7, 9
9-CH3	19.6	1.15, d (6.8)	7	7, 8
10-C	163.4	-		2’-NH
11-CH2	34.2	4.42, q (13.5)		13/17
12-C	138.2	-		12, 14, 16
13/17-CH	128.7	7.34, d (7.5)	14/16	11, 14/16, 15
14/16-CH	128.3	7.25, t (7.5)	13/17, 15	13/17, 15
15-CH	126.4	7.20, t (7.5)	14/16	13/17, 14/16
Leu				
1’-C	171.0			2’-NH , 2', 3'
2’-CH	51.0	4.62, m	2’-NH, 3'	2’-NH , 3', 4'
3’-CH2	42.7	1.54, m	2', 4'	2', 4', 5', 6'
4’-CH	24.6	1.59, m	3', 5', 6'	3', 5', 6'
5’-CH3	22.2	0.82, d (6.0)	4'	3', 4', 6'
6’-CH3	23.0	0.89, d (5.8)	4'	3', 4', 5'
2’-NH		8.30, d (8.6)	2'	-
Val				
1’’-C	172.4			3''
2’’-CH	58.7	3.90, ovl.**		4'', 5''
3’’-CH	30.6	2.05, m	4'', 5''	4'', 5''
4’’-CH3	19.6	0.80, d (6.9)	3''	3'', 5''
5’’-CH3	18.1	0.80, d (6.9)	3''	3'', 4''
2’’-NH				

*n.a. = no signal or correlation available **ovl. = overlap with other signals



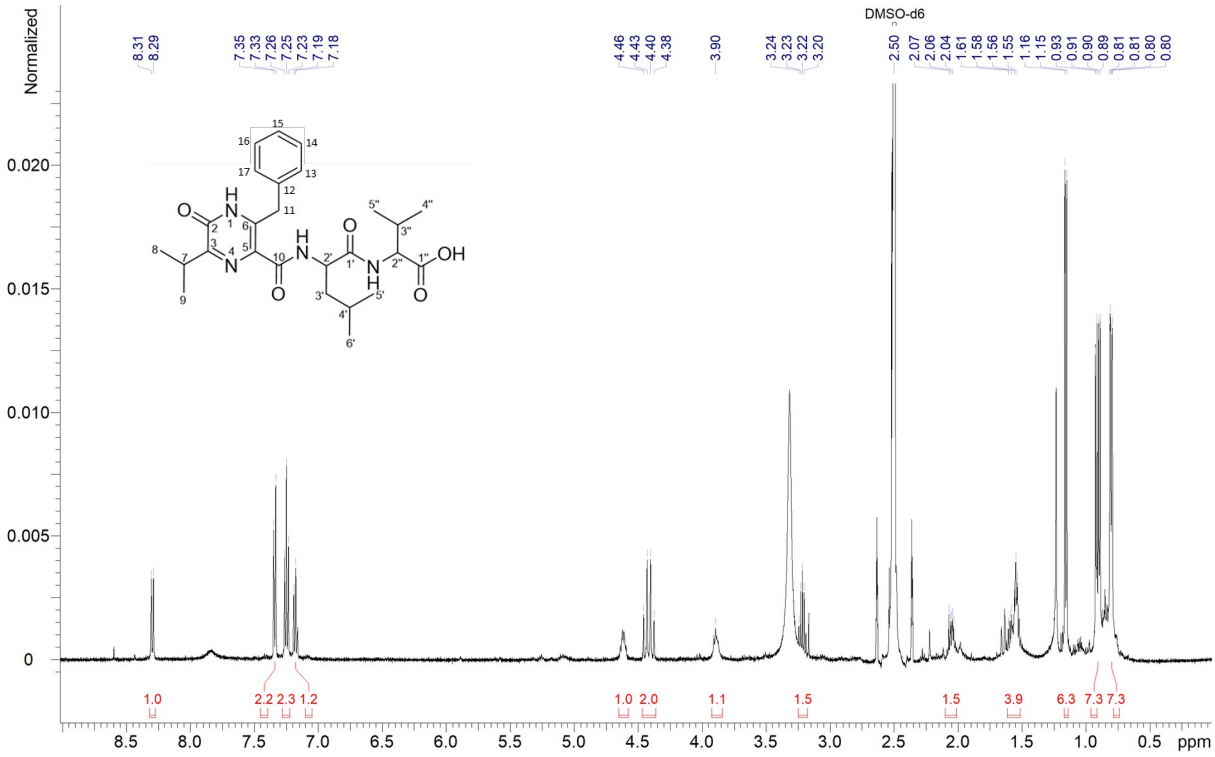


Figure S10: ^{1}H-NMR spectrum of ichizinone B in DMSO-d_6.

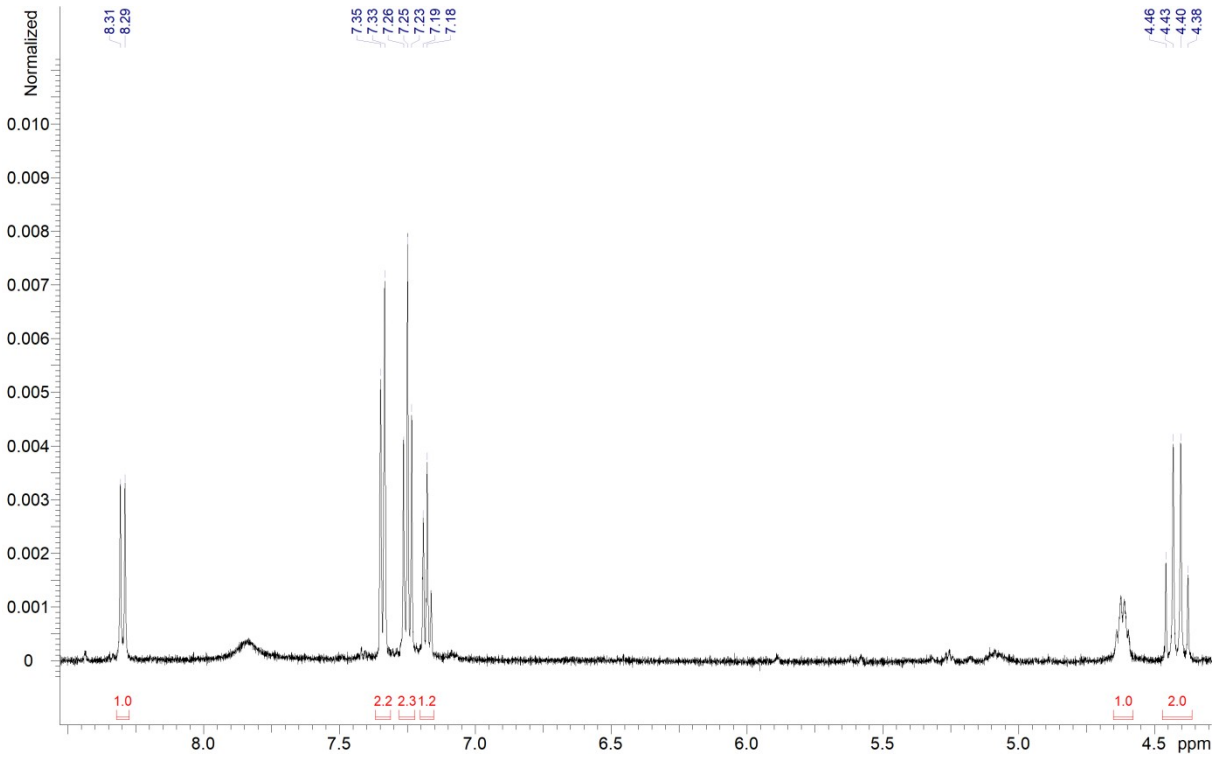


Figure S11: ^{1}H-NMR spectrum of ichizinone B in DMSO-d_6.

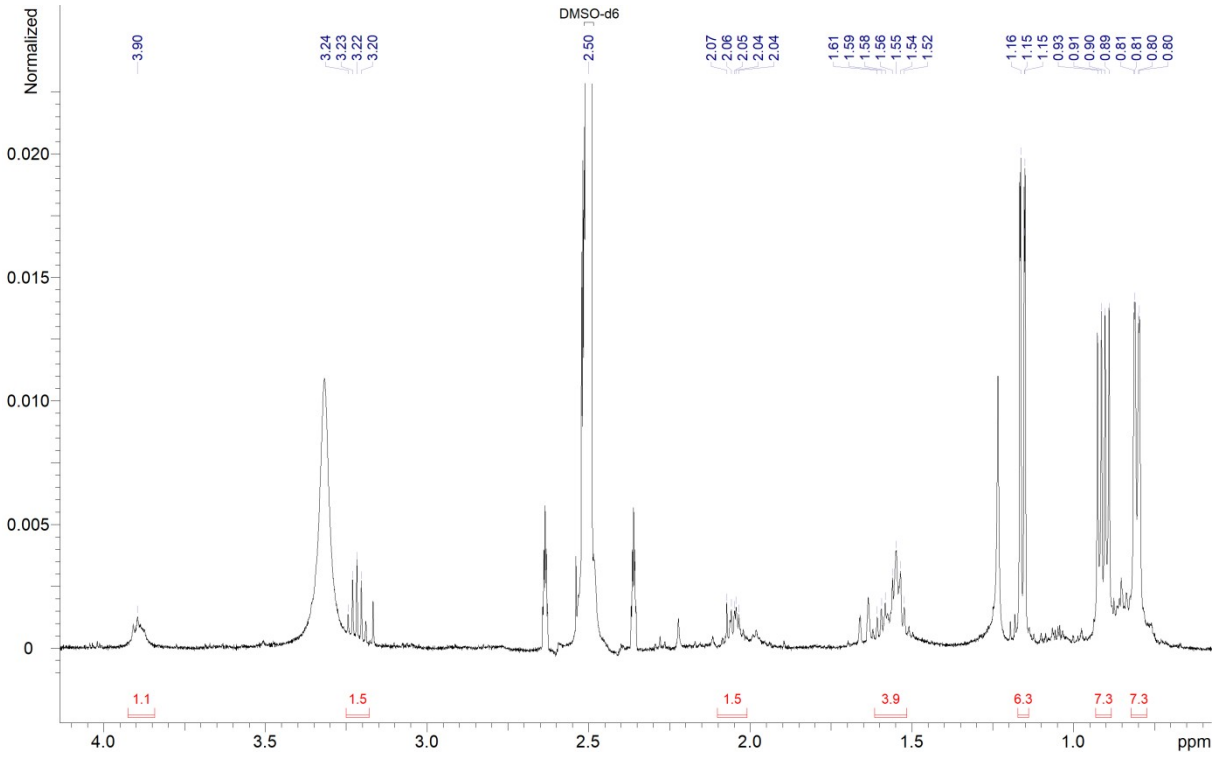


Figure S12: ^{1}H-NMR spectrum of ichizinone B in DMSO-d_6.

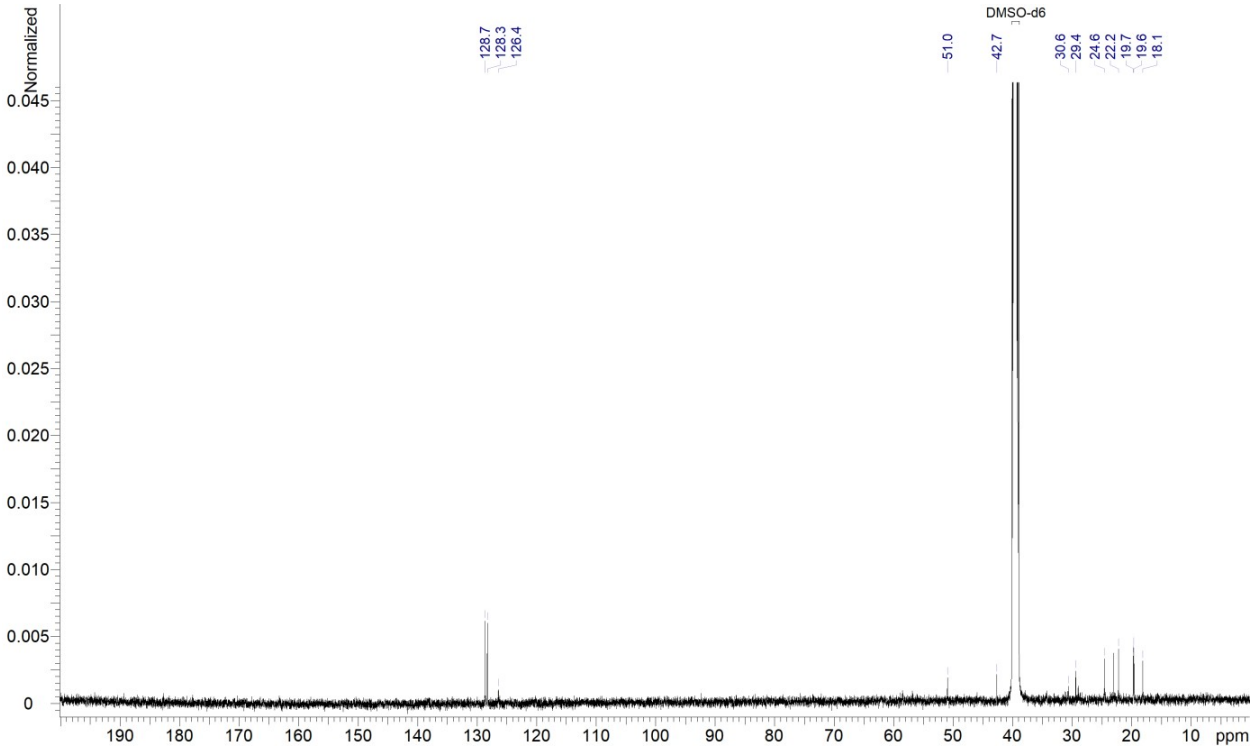


Figure S13: ^{13}C-NMR spectrum of ichizinone B in DMSO-d_6.

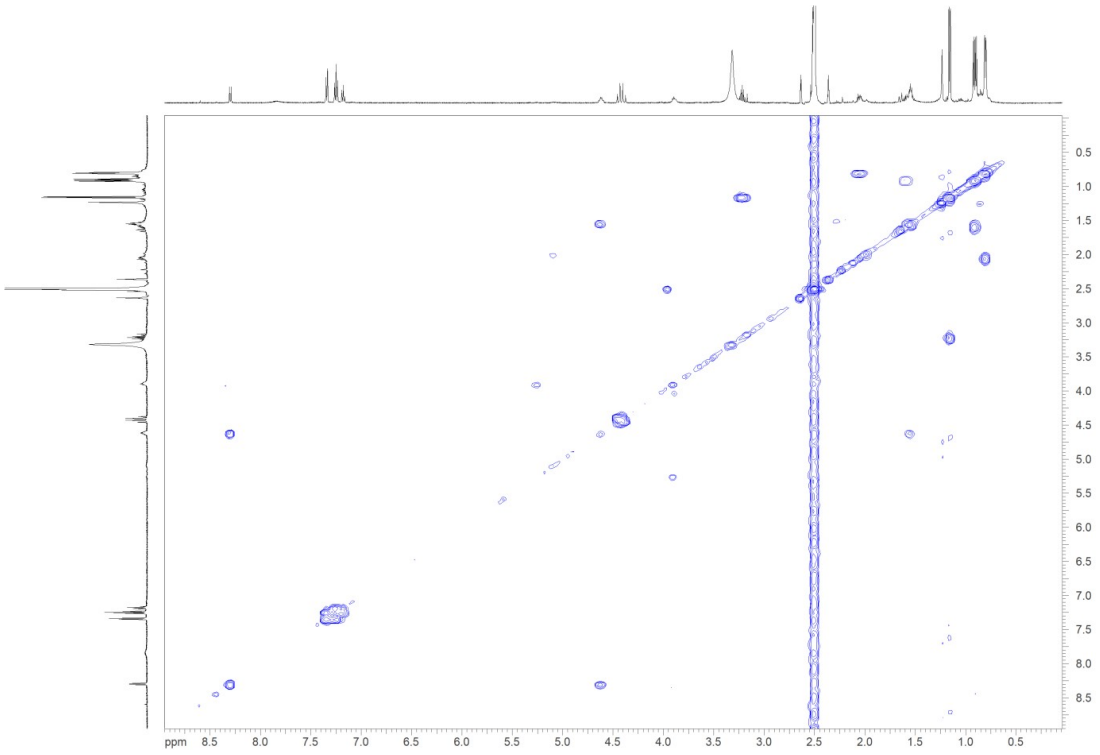


Figure S14: ^{1}H-^{1}H-COSY spectrum of ichizinone B in DMSO-d_6.

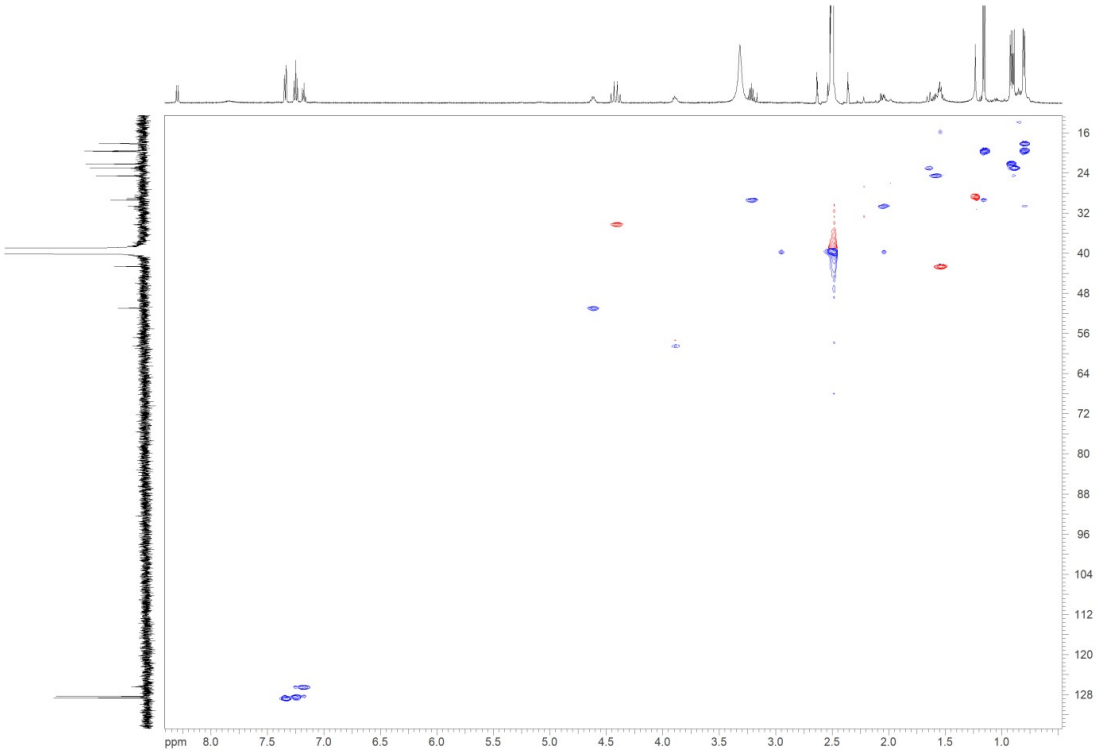


Figure S15: Edited HSQC spectrum of ichizinone B in DMSO-d_6.

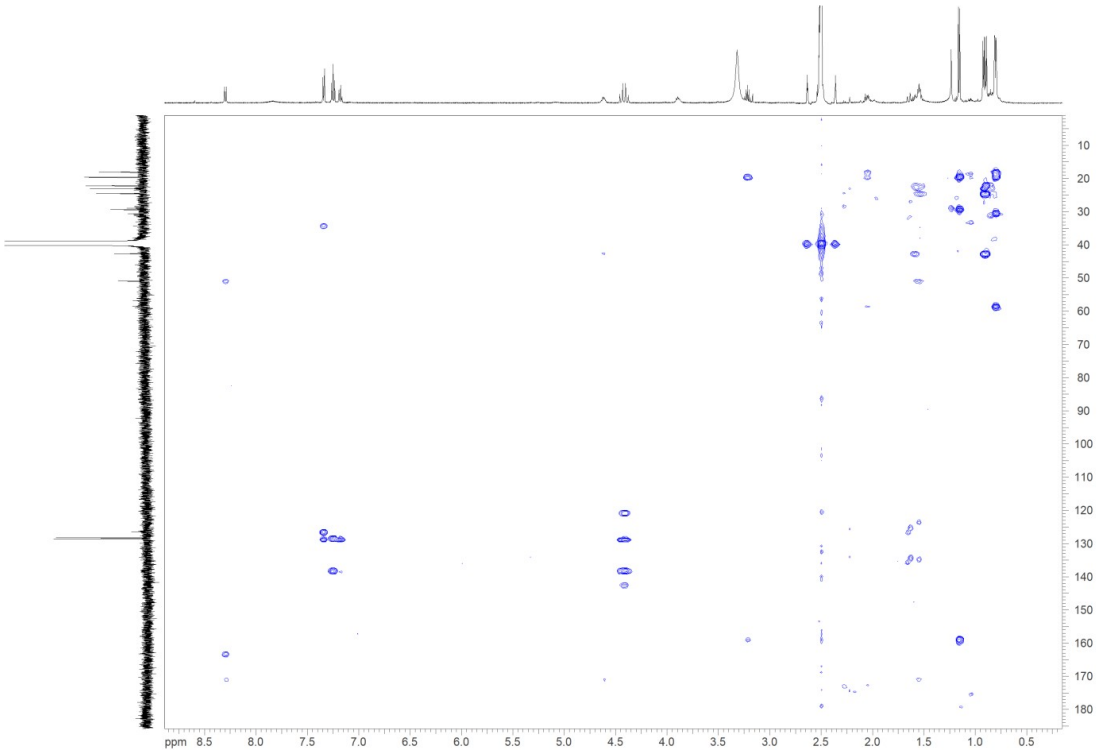


Figure S16: HMBC spectrum of ichizinone B in DMSO-d_6.

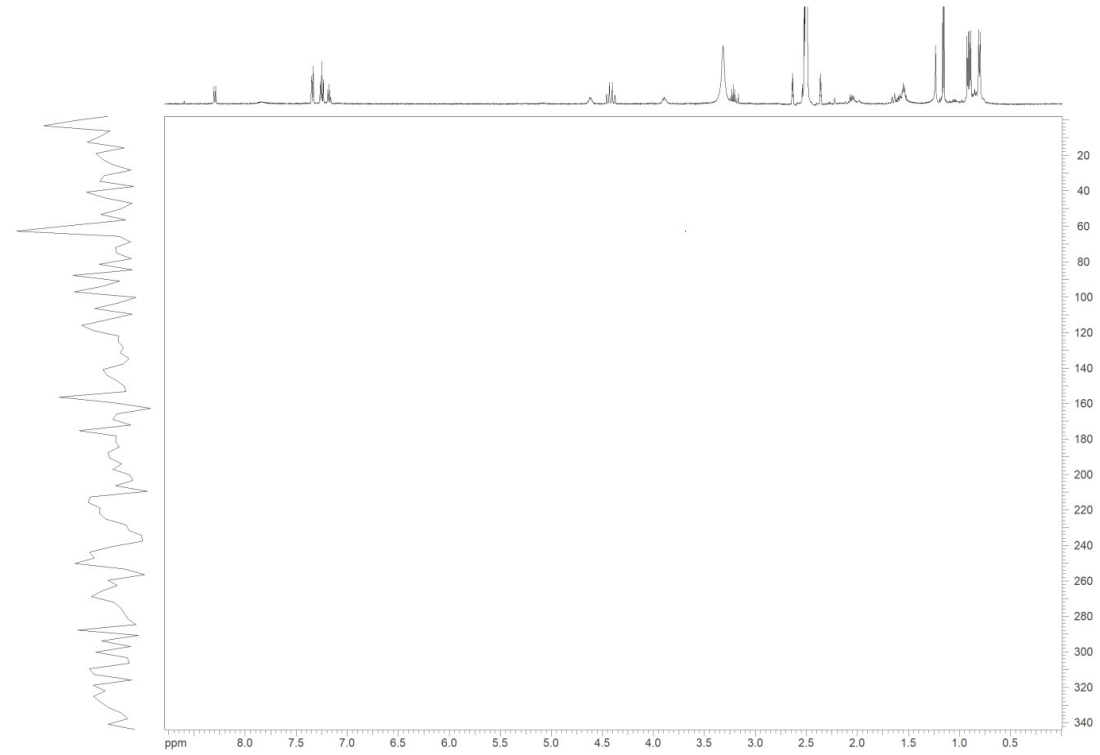
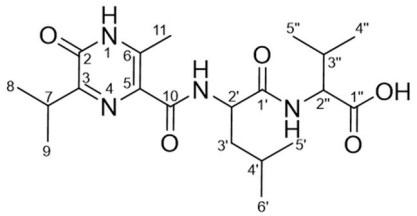


Figure S17: ^{15}N-HMBC spectrum of ichizinone B in DMSO-d_6

Table S5: NMR Table of ichizinone C measured in MeOD-d_3 and JBIR-57 for comparison.

Ichizinone C					**JBIR-57**	
No, type	**δ(13C, 15N) [ppm]**	**δ(1H) [ppm], mult(J)**	**COSY (H-)**	**HMBC (C-/N-)**	**δ(13C, 15N) [ppm]**	**δ(1H) [ppm], mult(J)**
1-NH	179.65			11		
2-C	157.93			7	155.8	
3-C	160.41			7, 8, 9	158.3	
4-N	320.15			7, 2’-NH, 11		
5-C	123.37			2’-NH, 11	120.9	
6-C	142.24			11	140.8	
7-CH	31.17	3.32, m	8, 9	8, 9	29.5	3.21, q (6.8)
8-CH3	20.20	1.24, d (6.9)	7	7, 9	19.9	1.15, d (6.8)
9-CH3	20.25	1.25, d (6.9)	7	7, 8	19.8	1.15, d (6.8)
10-C	166.43			2’-NH, 2'	163.6	
11-CH3	17.12	2.61, s			16.7	2.53, s
Leu					**Leu**	
1’-C	174.92			2’-NH, 2’’-NH, 2', 3'	171.6	
2’-CH	53.14	4.66, dt (7.6, 6.8)	2’-NH, 3'	2’-NH, 3' ,4'	50.6	4.53, dd (14.8, 8.4)
3’-CH2	43.06	1.70, ovl*	2', 4'	2’-NH, 2', 4', 5', 6'	42.6	1.52, dd (14.8, 6.2)
4’-CH	26.35	1.71, ovl*	3', 5', 6'	2', 3', 5', 6'	24.7	1.56, q (6.2)
5’-CH3	22.51	0.98, d (6.2)	4'	3', 4', 6'	23.3	0.89, d (6.2)
6’-CH3	23.52	0.99, d (6.4)	4'	3', 4', 5'	22.4	0.88, d (6.2)
2’-NH	117.54	8.34, d (8.2)	2'	2', 3', 4'		8.16, d (8.4)
Val					**Ala**	
1’’-C	174.85			2'', 3''	174.1	
2’’-CH	59.29	4.33, dd (8.7, 5.7)	2’’-NH, 3''	2’’-NH, 3'', 4'', 5''	47.8	4.14, dq (7.0, 6.2)
3’’-CH	31.86	2.19, m	2'', 4'', 5''	2’’-NH, 2'', 4'', 5''	17.7	1.24, d (7.0)
4’’-CH3	18.37	0.95, d (6.8)	3''	2'', 3'', 5''		
5’’-CH3	19.73	0.97, d (7.0)	3''	2'', 3'', 4''		
2’’-NH	117.54	8.16, d (8.5)	2''	2'', 3''		8.41, br s

*ovl. = overlap with other signals



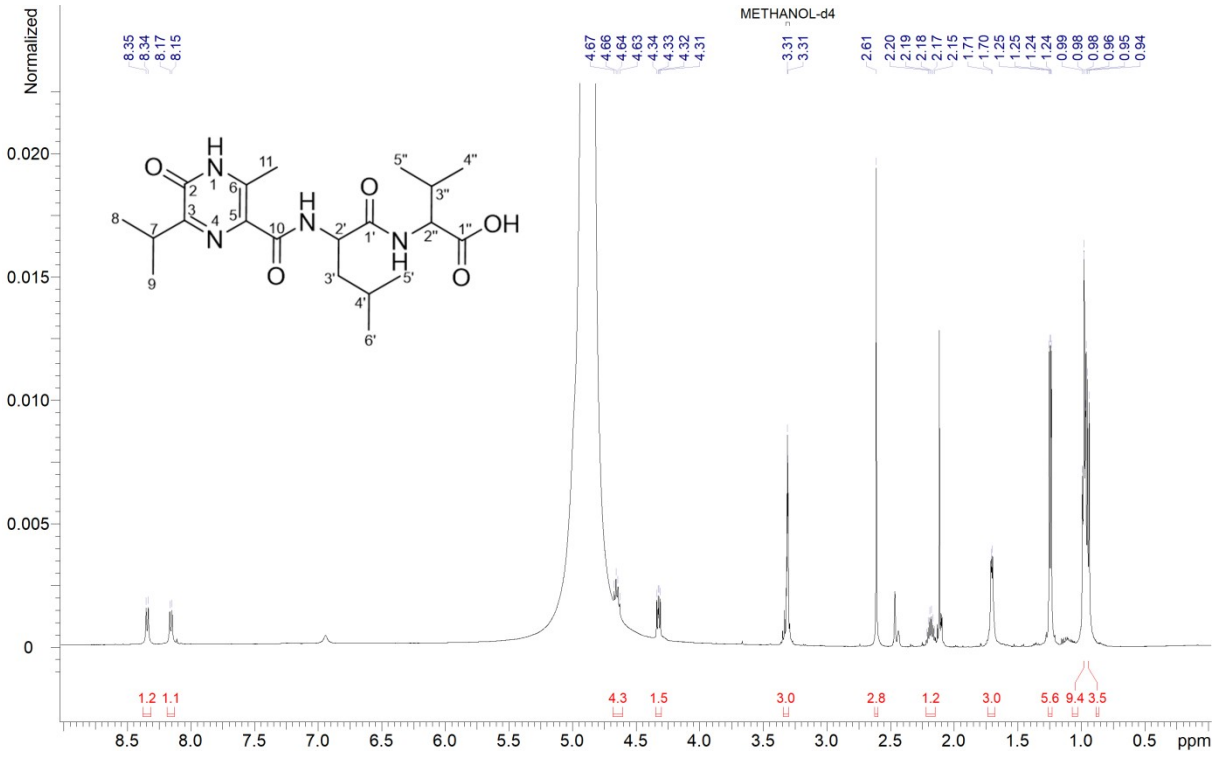


Figure S18: ^{1}H-NMR spectrum of ichizinone C in MeOD-d_3.

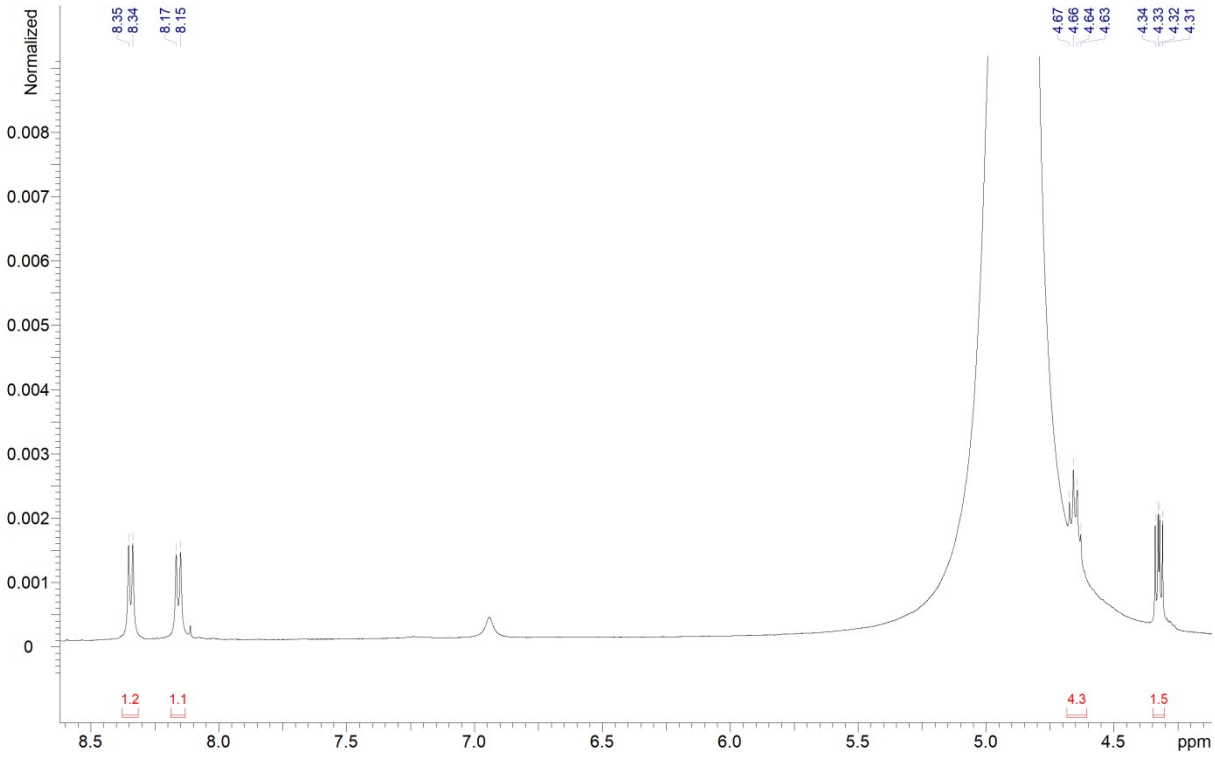


Figure S19: ^{1}H-NMR spectrum of ichizinone C in MeOD-d_3.

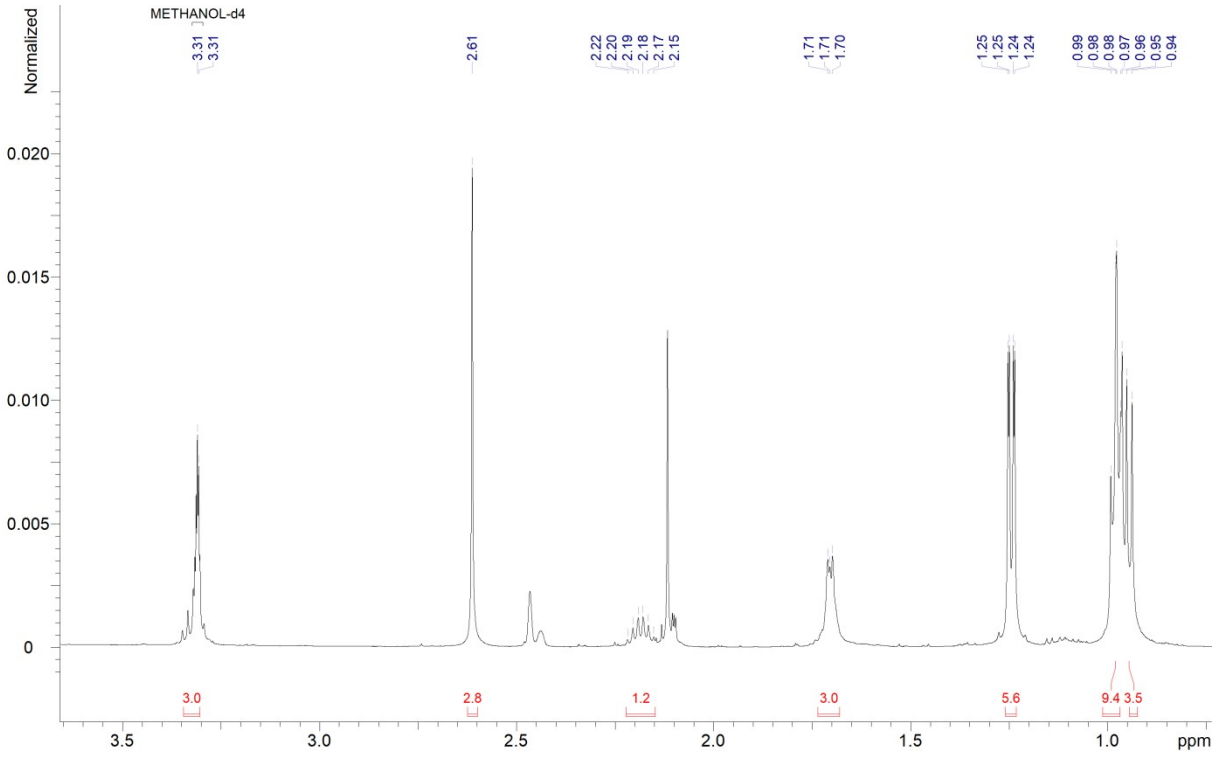


Figure S20: ^{1}H-NMR spectrum of ichizinone C in MeOD-d_3.

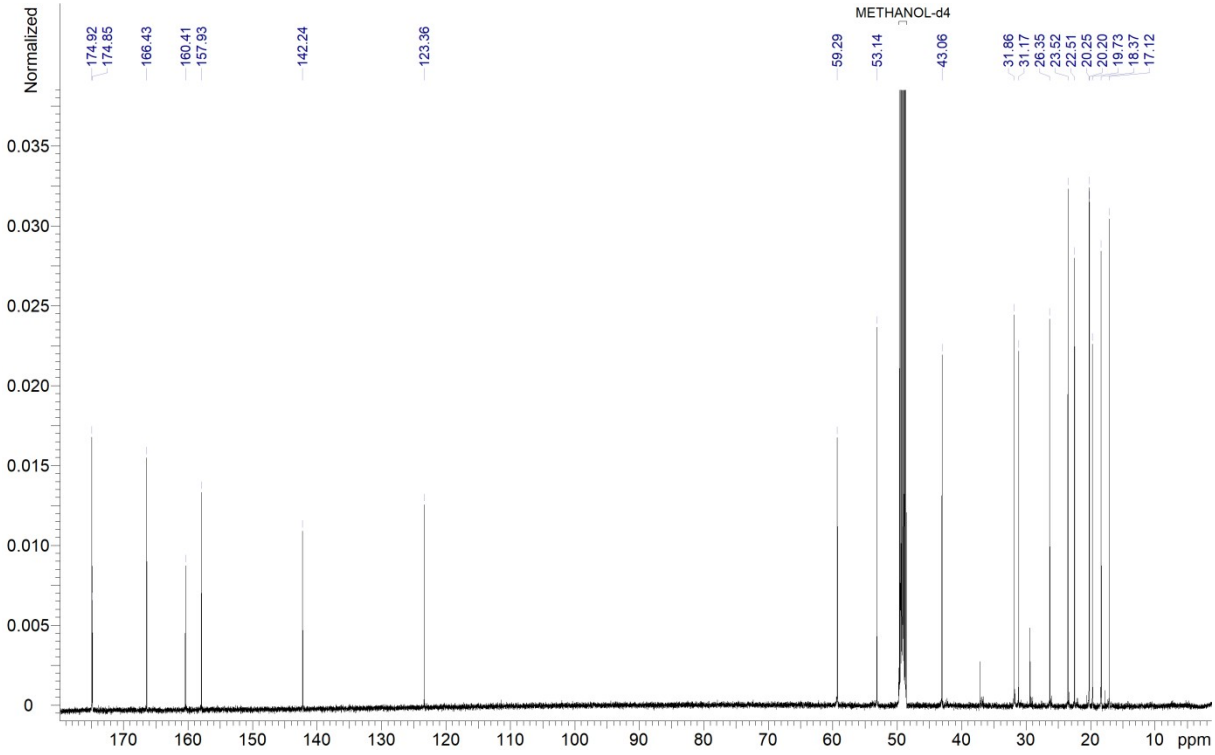


Figure S21: ^{13}C-MR spectrum of ichizinone C in MeOD-d_3.

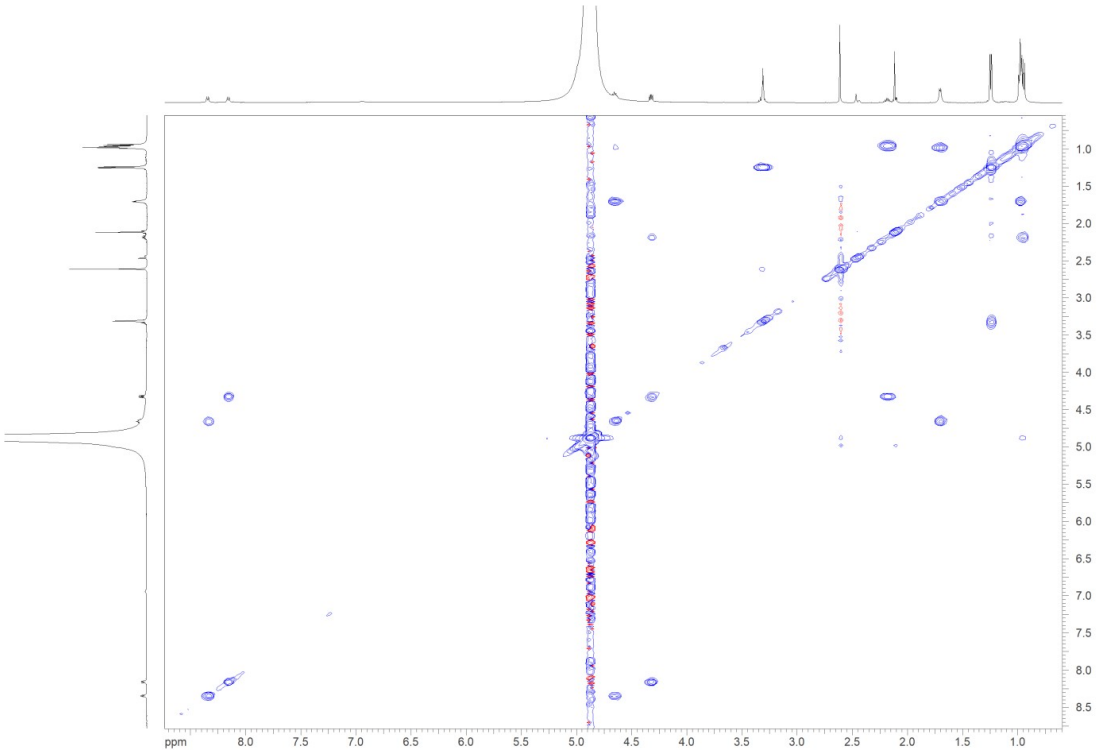


Figure S22: ^{1}H-^{1}H-COSY spectrum of ichizinone C in MeOD-d_3.

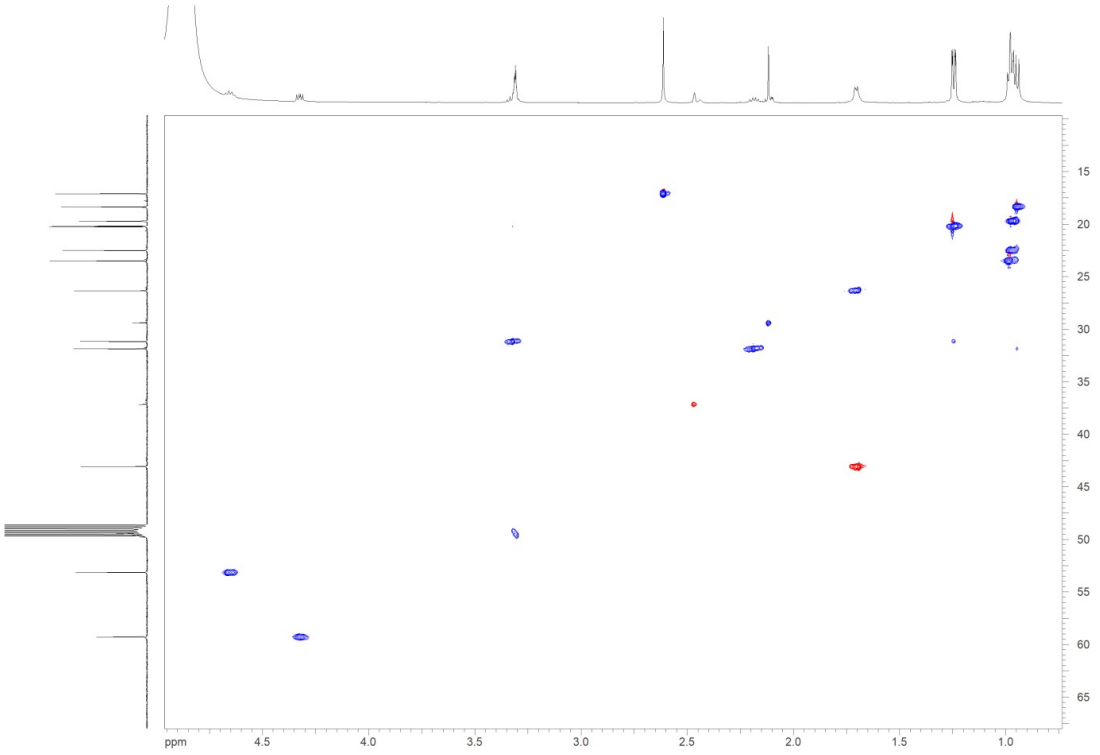


Figure S23: Edited HSQC spectrum of ichizinone C in MeOD-d_3.

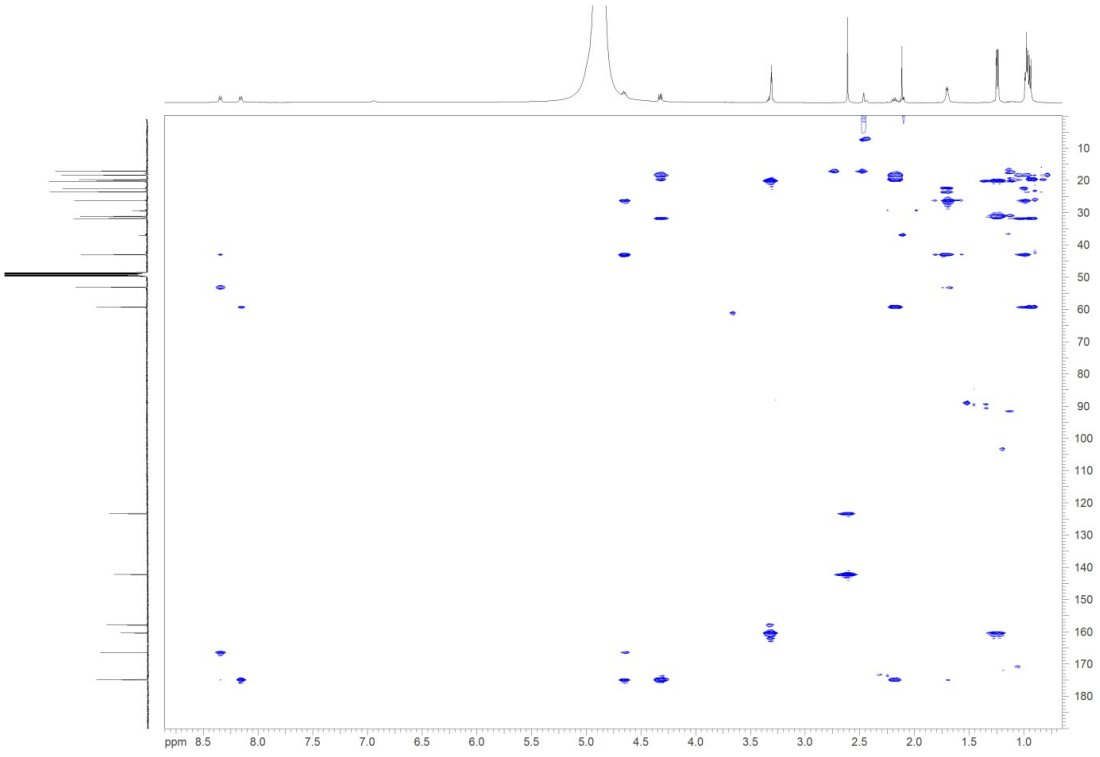


Figure S24: HMBC spectrum of ichizinone C in MeOD-d_3.

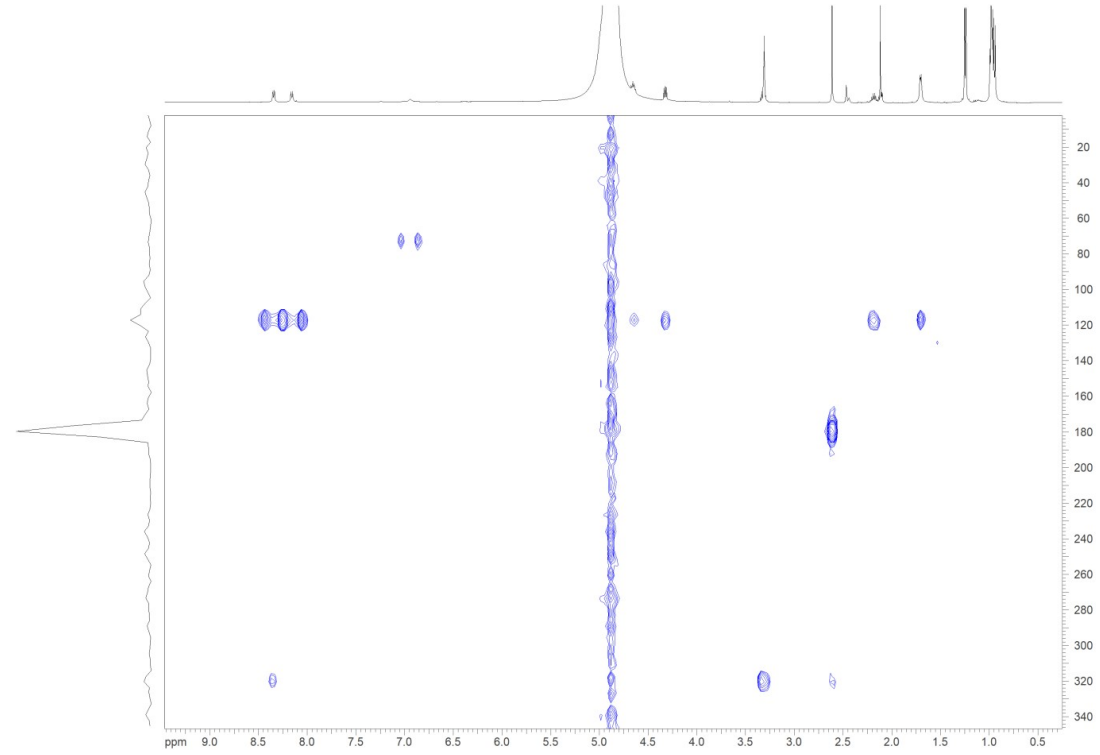


Figure S25: ^{15}N-HMBC spectrum of ichizinone C in MeOD-d_3.

MS/MS Fragmentation

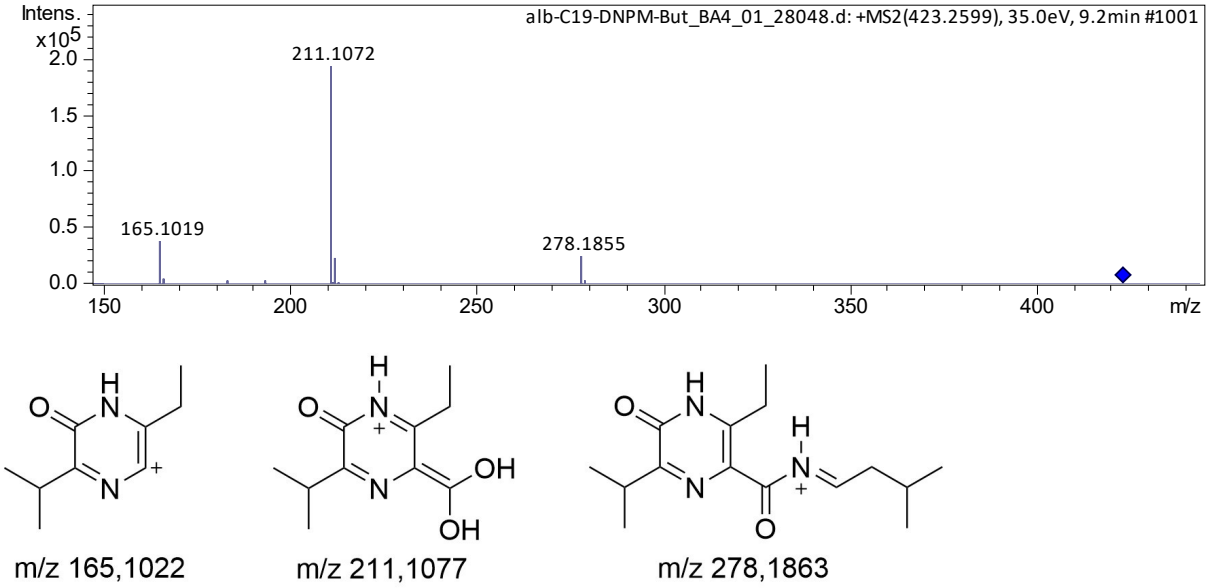


Figure S26: MS-fragmentation spectrum of ichizinone A and the suggested fragments including the calculated monoisotopic mass.

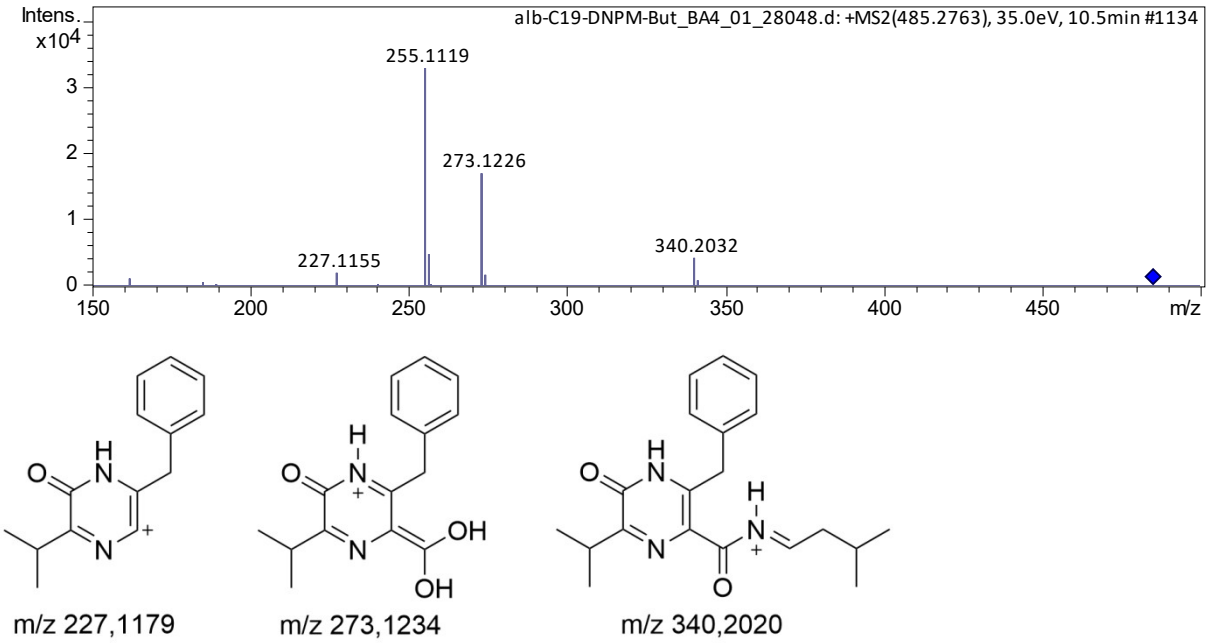


Figure S27: MS-fragmentation spectrum of ichizinone B and the suggested fragments including the calculated monoisotopic mass.

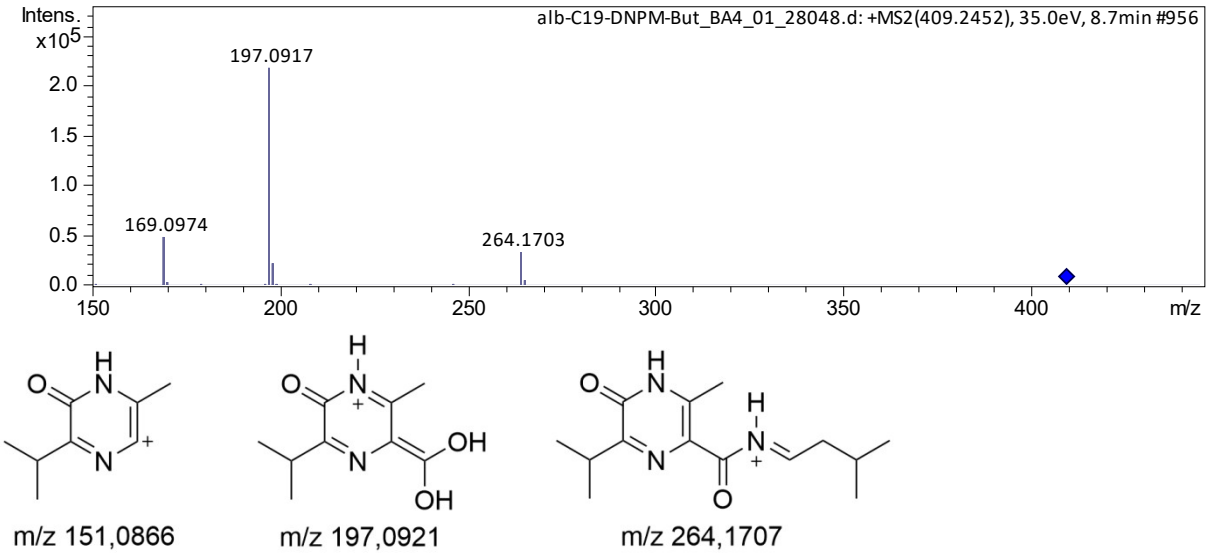


Figure S28: MS-fragmentation spectrum of ichizinone C and the suggested fragments including the calculated monoisotopic mass.

Marfey's analysis

Standards in D- and L-configuration were derivatized with L-FDLA and compared with the amino acids derived from hydrolysis of ichizinone C (fig. S 30)

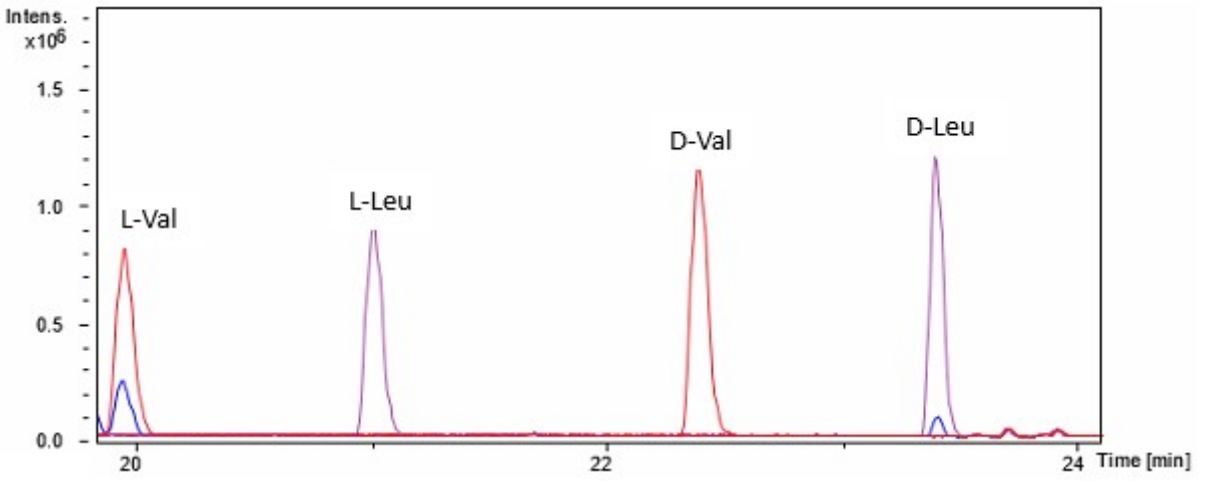


Figure S29: Marfey's chromatograms of the ichizinone hydrolysate (blue) and the amino acids standards (red: val; purple: leu) derivatized with L-FDLA.

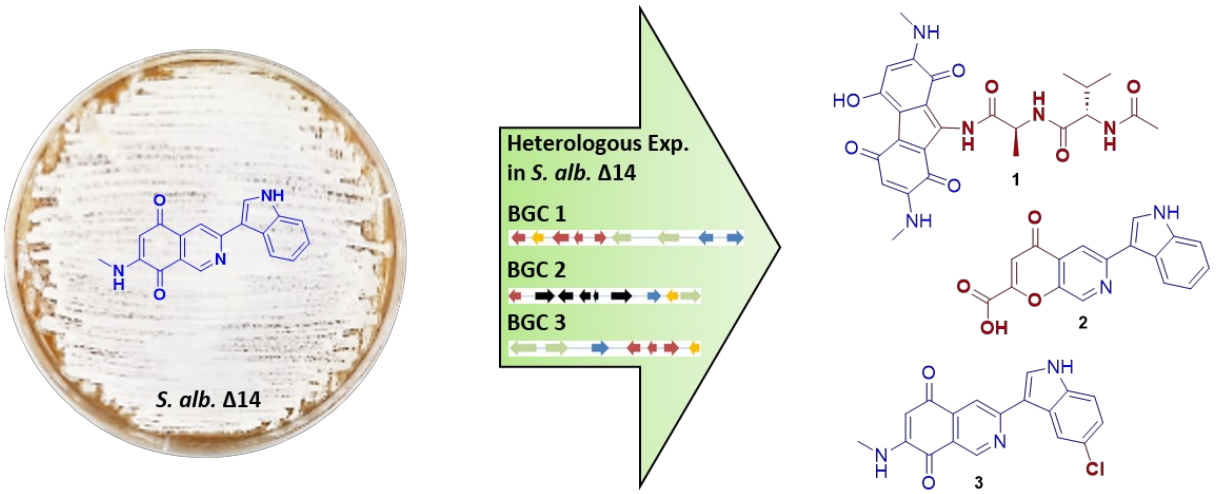
References

1. Myronovskyi, M., et al., Generation of a cluster-free Streptomyces albus chassis strains for improved heterologous expression of secondary metabolite clusters. Metabolic engineering, 2018. 49: p. 316-324.
2. Flett, F., V. Mersinias, and C.P. Smith, High efficiency intergeneric conjugal transfer of plasmid DNA from Escherichia coli to methyl DNA-restricting streptomycetes. FEMS microbiology letters, 1997. 155(2): p. 223-229.
3. Grant, S.G., et al., Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. Proceedings of the National Academy of Sciences, 1990. 87(12): p. 4645-4649.
4. Myronovskyi, M., et al., Generation of new compounds through unbalanced transcription of landomycin A cluster. Applied microbiology and biotechnology, 2016. 100: p. 9175-9186.

2.2 Hybrid compounds obtained *via* interaction of The Native Mansouramycin Biosynthesis in the Host *Streptomyces albus* Del14 with Heterologously Expressed Clusters

Marc Stierhof, Liliya Horbal, Patrick Oberhäuser, Anja Paluszak, Peyton Cox, Maria Lopatniuk, Christopfer Ruf , Josef Zapp and Andriy Luzhetskyy

To be submitted



2.2.1 Abstract

Streptomyces albus J1074 (now *S. albidoflavus* J1074) is a widely used heterologous host for natural product discovery due to its capacity to express biosynthetic gene clusters (BGCs) from diverse organisms. A derivative of this strain, *S. albus* Del14, enhances heterologous expression by reducing background metabolite production enabling the identification of the previously hidden BGC responsible for producing mansouramycins. In this study, we demonstrate the biosynthetic crosstalk between the native mansouramycin BGC in *S. albus* Del14 and introduced BGCs from three different organisms results in the production of novel hybrid compounds, some featuring rare and complex chemical scaffolds. These include malevonin, which combines NRPS- and mansouramycin-derived building blocks forming a fluorene scaffold, as well as 5'-chloromansouramycin D, a halogenated derivative of mansouramycin D. Additionally, we identified mansevorone, a compound structurally similar to mansouramycin D but utilizing a different tryptophan-derived C7 precursor. This precursor likely arises from the activation of native genes in the host *S. albus* Del14, triggered by SARP regulators present on the introduced BGC. These findings highlight the evolutionary significance of BGC interactions and underscore their potential as a powerful tool for discovering novel natural products, providing insights that could inform innovative strategies in biosynthetic engineering and the guided evolution of new bioactive compounds.

2.2.2 Introduction

Over the last two decades *Streptomyces albus* J1047 has successfully been used as a heterologous expression host to study and discover a large variety of natural products[1-6]. To facilitate access to these metabolites and optimize expression conditions, a chassis strain was developed based on *S. albus* J1047, namely *S. albus* Del14. Deletion of 15 biosynthetic gene clusters (BGCs) in *S. albus* Del14 led to a reduced metabolic background and higher production

titers, which enabled the successful expression of previously inaccessible natural products[7]. The simplified production profile of *S. albus* Del14 resulted in the appearance of new peaks which were later identified as mansouramycin A and D. Due to their unknown biosynthesis, the corresponding biosynthetic gene cluster was not annotated by genomic search tools like antiSMASH and only showed production when the rest of the metabolic background was cleared. In a study by Hui et al., it was discovered that the cluster utilized tryptophan (Trp) as a building block, which draw the focus on Trp-related genes. This finding subsequently enabled the identification of the BGC responsible for mansouramycin biosynthesis and the elucidation of its biosynthetic pathway until the C7 precursor, while the final biosynthetic mechanism leading to mansouramycins A-G remains unknown (Fig. 1)[8-10]. In the course of this study, the mansouramycin BGC-deficient strain *Streptomyces albus* Del15 was generated. However, it exhibited reduced fitness compared to *S. albus* Del14. Therefore, *S. albus* Del14 was selected as the primary host for our heterologous expression pipeline[11].

Most of the BGCs successfully expressed in *S. albus* Del14 produced compounds that were predicted using in silico tools. However, in some cases, the structure of the newly identified metabolites showed significant differences to the predicted one and seemed to incorporate structural elements from mansouramycin D, resulting in a great variety of hybrid compounds. We identified three of such compounds, including malevonin (**1**) with a substituted fluorene scaffold, mansevorone (**2**) with an azachromone scaffold and 5'-Cl-Mansourmaycin D (**3**), a chlorinated variant of mansouramycin D. By conducting feeding studies, knockout experiments, and heterologous production in the mansouramycin BGC deficient strain *S. albus* Del15, we demonstrated that mansouramycin biosynthesis interacts with the foreign biosynthetic genes, leading to a variety of different new compounds. This provides a unique insight into cluster evolution in *Streptomyces* by mimicking the natural occurring horizontal gene transfer through heterologous expression. Understanding the mechanism behind these cluster interactions could be utilized to provide new strategies for biosynthetic engineering and direct evolution.

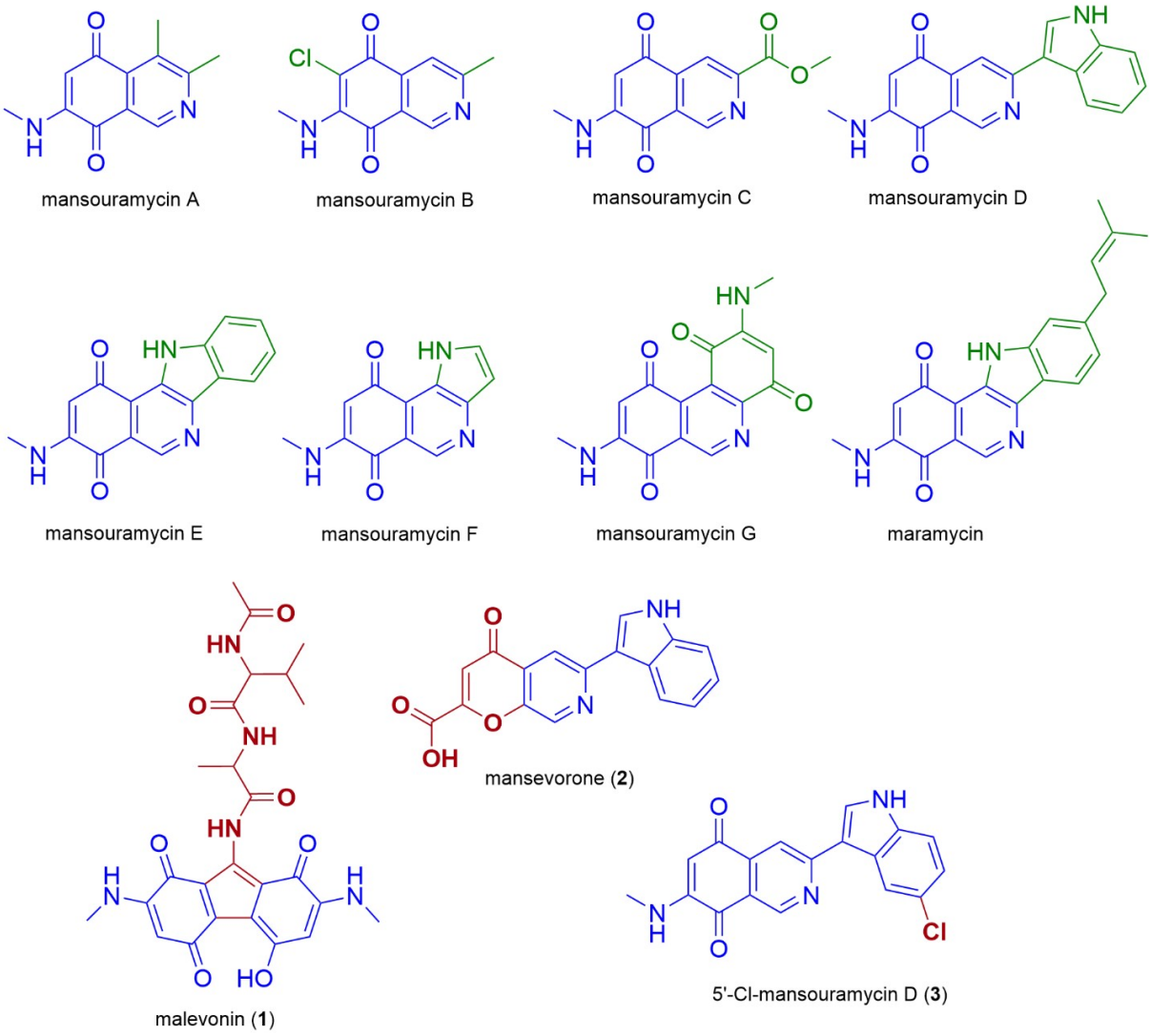


Figure 1: Structures of all known mansouramycins, maramycin and the newly discovered malevonin (**1**), mansevorone (**2**) and 5'-Cl-mansouramycin (**3**).

2.2.3 Results and Discussion

Heterologous Production of Malevonin:

In our heterologous expression pipeline we constantly express biosynthetic gene clusters from various organisms in our in house chassis strains *S. lividans* Del8 and *S. albus* Del14[12-18]. Mutant strains are cultivated in SG and/or DNPM medium, metabolites are extracted by butanol and/or ethylacetat and analysed by LC-HRMS. New peaks are identified by dereplication using the dictionary of natural products and natural product atlas. One of the clusters from our pipeline, cluster 3, originated from cosmid library created for *S. kitasatoensis* (NCBI acc. nr. PV759347.1). Cluster 3 was selected based on 54% similarity to vazabitide A biosynthetic genes by antiSMASH[19]. The cluster showed differences in the NRPS core genes and some

additional biosynthetic genes that suggested a new compound different from vazabitide A. Heterologous expression of the cluster 3 in our second chassis strain *S. lividans* Del8 did not lead to any new peaks. However, in *S. albus* Del14 we observed new masses including a major compound at $[M+H]^+$ = 512.214 accompanied by dark coloration of the bacterial culture (Fig. S1, S2; Fig. 2A-C). The compound was rather insoluble, therefore, it was possible to purify it by washing the bacterial butanol extract with hexane and methanol resulting in 6 mg of the compound.

The molecular formula of the new compound was established as $C_{25}H_{29}N_5O_7$ with 14 degrees of unsaturation based on m/z = 512.2145 $[M+H]^+$. Due to its low solubility, the compound was dissolved in DMSO-d_6 to saturation (ca. 2 mg in 300 µL) and the structure was determined by NMR (Table S6; Fig S3-S9). Analysis of ^1H, ^{13}C and edited-HSQC spectra revealed five CH, six CH_3 and nine quaternary carbons. Four proton signals above 7 ppm did not show correlations to carbon signals and were determined as NH groups. COSY spectra revealed two spin systems, which were assigned to alanine (Ala) and valine (Val). Long-range HMBC correlations showed that the two amino acids form a dipeptide with an N-terminal acetyl group (Ac) resulting in Ac-Val-Ala. The calculated remaining molecular formula $C_{15}H_{12}N_3O_4$ did not match the observed signals from the 1D experiment that indicated a molecular formula of $C_8H_6N_2O_2$, suggesting a symmetrical moiety. This was confirmed by comparing the integral values in the ^1H-NMR, where 9-NH showed a value of 1, while 3/6-CH, 2/7-NH, and 2/7-NMe displayed twofold integral values. Assignment of all peaks with twofold integral values led to 2-(Methylamino)-1,4-d benzoquinone (MAB), a structural feature of mansouramycin The HMBC correlation from 9-NH to 1a/8a-C, along with a weak correlation to 9-C, suggested that 9-C is adjacent to both 9-NH and 1a/8a-C, thereby linking the two MAB moieties (Fig. 2F). Based on the established structure, the final connection needed to achieve the required 14 degrees of unsaturation is between the quaternary carbons 4a-C and 5a-C resulting in a 2,7-N-methyl-5-Hydroxy-1H-fluorene-1,4,8-trione. The symmetry is achieved by the enol hydroxyl proton that forms hydrogen bonds between the oxygen atoms at position 4 and 5, resulting in a mesomeric diketonate (Fig. 2F). In this configuration, the proton is significantly deshielded, resulting in a downfield-shifted signal at 18.47 ppm (Fig. S4). A review of the literature revealed a structurally similar compound, namely hipposudoric acid, which exhibited a comparable chemica shift of of 16.05 ppm for the deshielded proton, supporting our observations[20]. The absolute configuration of the amino acid residues was determined by Marfey's method, which resulted in L-Ala and L-Val (Fig. S10), and the compound was named malevonin (**1**)[21].

Cluster 3 Enzymes Utilizes the Tryptophan-Derived C7 Precursor.

Malevonin differs significantly from the expected cluster product, which was anticipated to be a vazabitide A derivative. Instead, it features a substituted fluorene scaffold incorporating two MAB moieties, structural elements derived from mansouramycin D, suggesting an interaction between the native mansouramycin biosynthetic cluster and the introduced cluster 3. To verify this, cluster 3 was expressed in the mansouramycin BGC deficient strain *S. albus* Del15. Subsequent analysis of the metabolic profile of the *S. albus* Del15_cluster3 mutant did not show production of malevonin or other new masses (Fig. 2D). This indicates that cluster 3 requires structural elements generated by the native mansouramycin BGC to form malevonin.

As demonstrated by Hui et al., the mansouramycin BGC uses tryptophan (Trp) as a building block for the Trp-derived C7 precursor 5-(methylamino)-3,6-dioxocyclohexa-1,4-diene-1-carboxamide (**5**)[11]. To determine if the fluorene core in malevonin is constructed from the C7 fragment, *S. albus* Del14_cluster3 was fed with $^{13}C_{11}$ labelled Trp (**4**) and the culture extracts were analysed by LC-MS. The feeding was done in rich medium, providing both labelled and unlabelled Trp as building blocks. Subsequent analysis revealed the unlabelled malevonin mass peak and masses differing by +6, +7 and +13 Da, indicating an incorporation of a C7 and a C6 fragment, both originating from **5** (Fig. S11; Fig. 3). Concluding from that, the fluorene core is likely synthesized from two moieties of the Trp-derived C7 precursor (**5**).

The peptide moiety of malevonin was hypothesized to originate from cluster 3, which was analysed through *in silico* gene analysis. The chromosomal fragment of cluster 3 consists of 53 genes, including *orf6-orf36* which contain the genes *vzb0*, *vzb3* and *vzb5-28* from vazabitide biosynthesis (Table S5)[19]. The presence of three NRPS genes (*orf22*, *orf23* and *orf32*) and an N-acetyltransferase (*orf29*) indicates that the Ac-Val-Ala (6) moiety is synthesized by cluster 3. The adenylation (A)-domain specificity prediction by antiSMASH and other prediction tools revealed that *orf32* is incorporating Val, while *orf23* revealed unclear results from including Ala, Val, Trp, methionine (Met) or lysine (Lys). *Orf22* carries two A-domains, one with a prediction for threonine (Thr), while the other revealed the same unclear specificity as the one in *orf23*. Alignment of the two undefined A-domains from *orf22* and *orf23* revealed a 90% identity (Fig. S12), suggesting they have similar function. To reveal their role in the malevonin biosynthesis, *orf22* it was deleted from cluster 3. As a result, production of malevonin in the deletion mutant *S. albus* Del14_cluster3_delOrf22 was completely abolished and no new compound was observed, indicating that *orf22* is essential in the biosynthesis of malevonin (Fig. 2E).

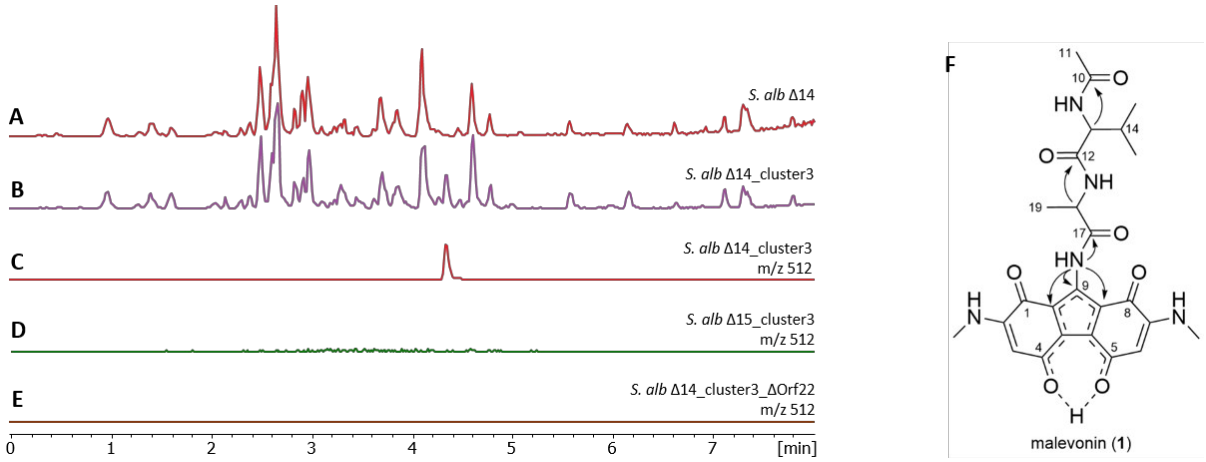


Figure 2*:* Base peak chromatogram of the *S. albus* Del14 chassis strain (**A**), heterologous expression of the cluster 3 in *S. albus* Del14 (**B**) showing production of malevonin (**C**), abolished malevonin production in the mansouramycin deficient strain *S. albus* Del15_cluster3 and (**D**) abolished malevonin production in the NRPS deletion mutant *S. albus* Del14_cluster3_delOrf22 (**E**). The mesomeric structure of malevonin and key HMBC correlations ( ) are shown in **F**.

The performed experiments indicate that malevonin biosynthesis is equally dependent on the Trp-derived C7 precursor from mansouramycin biosynthesis and the introduced cluster 3. A biosynthetic pathway was proposed starting with the formation of the acetylated dipeptide by the NRPS domains and the N-acetyltransferase in cluster 3, followed by the attachment of the substituted fluorine, which is derived from two moieties of the C7 precursor (**5**) from mansouramycin biosynthesis (Fig. 3). The assembly of the substituted fluorene scaffold and attachment to the dipeptide is yet unknown. The structural fluorene core of malevonin is significantly distinct from that of other mansouramycin derivatives. Whether its formation is regulated by enzymatic control or driven by spontaneous chemical processes remains unclear and will be investigated in future studies. Biological activity testing of malevonin was attempted but failed due to poor solubility, even in DMSO.

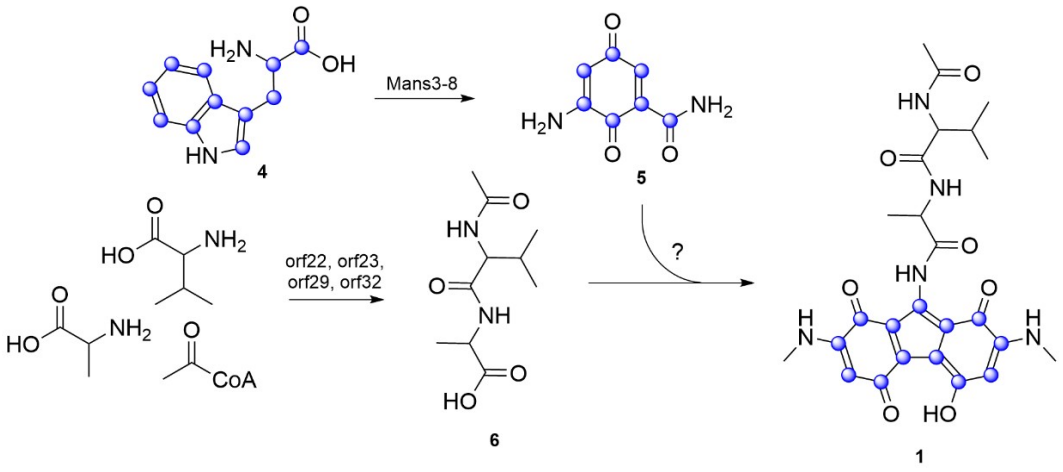


Figure 3: Proposed biosynthetic pathway of malevonin also showing the labeling pattern when $^{13}C_{11}$ Trp (**4**) is fed (blue dots)

Heterologous Production of Mansevorone.

By working on the expression of a terpenoid BGC, designated as cluster 16, from a cosmid library constructed for the uncharacterized *Streptomyces* strain LV45-129[22], we discovered another hybrid compound containing elements from mansouramycin D (CP023695.1, 7542688 – 7583435). Expression of the cluster 16 in *S. albus* Del14 led to the production of a new peak with m/z = 307.072 $[M+H]^+$ and an orange coloration of the culture (Fig. S13, S23; Fig. 4A-B). The production titers were insufficient for isolating the compound, requiring a refactoring of the cluster expression. Revisiting cluster 16, we identified a 2844-nucleotide gene encoding a 948-amino acid protein AfsR16 (locus tag) that was classified as a member of the SARP (Streptomyces Antibiotic Regulatory Protein) family through a BLAST search (Table S7). SARP regulators are transcriptional activators that directly influence the transcription of secondary metabolite biosynthesis genes; hence we aimed to overexpression of the gene to increase production using the pRT801 plasmid, which contains a strong synthetic promotor TS81[23-24]. The resulting expression vector pRT801_ampery_AfsR16 was introduced into *S. albus* Del14 harbouring cluster 16 via bacterial conjugation, generating the recombinant strain S. albus Del14_16h7. Subsequent determination of the production titer of the target mass m/z = 307.072 [M+H] resulted in a fivefold increase in target compound, indicating that the SARP regulator encoded by pRT801_ampery_AfsR16 plays a critical role in promoting the biosynthetic pathway of mansevorone in *S. albus* Del14_C16 (Fig. 4C).

The increase of production titer enabled the successful isolation of mansevorone after large-scale cultivation of *S. albus* Del14_C16h7 under the described conditions yielding 1 mg of the target compound, which was sufficient for subsequent NMR analysis (Table S8; Fig. S14-S22). The molecular formula was calculated as $C_{17}H_{10}N_2O_4$ based on the molecular ion $[M+H]^+$ = 307.072 Da. The ^{1}H-Proton NMR and edited-HSQC revealed eight methine signals in the downfield indicating a highly conjugated structure. In combination with COSY, ^{13}C-HMBC and ^{15}N-HMBC, 7 of these signals were assigned to an indole and a pyridine moiety. NOESY correlation between 1'-CH and 4-CH and HMBC correlations from 1'-CH to 3-C indicated that both moieties are connected between 2'-C and 3-C, revealing a structure similar to mansouramycin D lacking the MAB moiety (Fig. 4D). The new structural feature was calculated as $C_4H_2O_4$ and assigned to 2-hydroxy-4-oxo-2-butenoic acid using HMBC correlations. The calculated number of 14 ring double bond equivalents and the chemical shift of 1a-C at δC = 150.5 ppm indicate a cyclic structure forming a 4-pyrone-2-carboxylic acid unit. Comparison of the calculated shift with the experimental data confirmed cyclization via C7-OH rather than cyclization via C7a-COOH. The structure was confirmed through MS/MS

fragmentation, with experimental results compared to predicted outcomes generated by CFM-ID 4.0. (Fig. S24)[25-28]. The resulting structure is forming a rare 7-azachromone core and was named mansevorone (Fig. 4D).

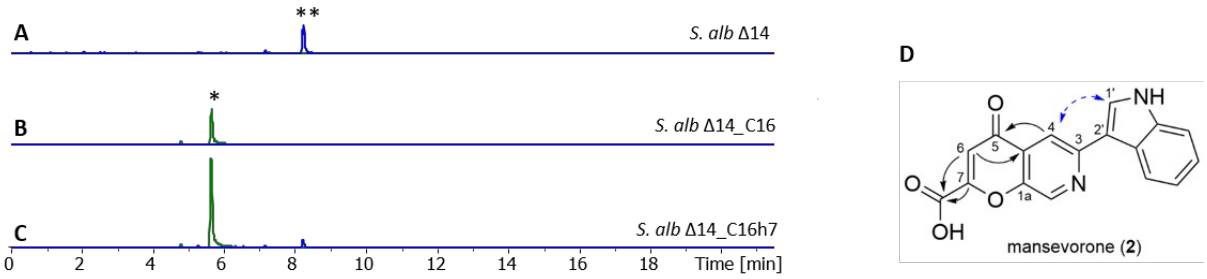


Figure 4: Extracted ion chromatograms showing mansevorone (**2**) and mansouramycin D (*) production in *S. albus* Del14 (**A**), *S. albus* C16 harboring cluster 16 (**B**) and in *S. albus* C16h7 harboring cluster16 and pRT801_ampery_AfsR16 for SARP overexpression (**C**). The mansevorone structure and key HMBC correlations ( ) and NOESY correlation ( ) are shown in **D**

Mansevorone Biosynthesis is Encoded by *S. albus* Del14

Mansevorone shares structural similarities with mansouramycin D, but features a 4-pyrone-2-carboxylic acid unit in place of the MAB moiety. This suggests that mansevorone (**2**) may result from an interaction between cluster 16 and the native mansouramycin BGC from *S. albus* Del14. To determine whether the mansouramycin BGC is involved in mansevorone synthesis, cluster 16 was transferred into the mansouramycin BGC-deficient strain *S. albus* Del15. No production of mansevorone or any other new peaks was observed in *S. albus* Del15_C16, indicating that mansevorone production is dependent on the presence of the mansouramycin BGC.

To further investigate the role of Trp in mansevorone biosynthesis, *S. albus* del14_C16h7 was fed with $^{13}C_{11}$-labeled tryptophan over the course of seven days. HPLC-MS analysis of the extracts revealed, in addition to the mansevorone peak (m/z = 307.072), additional ions with m/z values ranging from 313 to 324 Da, corresponding to mass shifts of +6 to +17 Da (Fig. S23). The most intense mass peaks were observed at m/z = 317.1049 and 324.1283 Da, indicating mass shifts of +10 and +17 Da, respectively, suggesting the involvement of two main biosynthetic precursor fragments, C7 and C10, both derived from Trp. The structural similarities between mansevorone and mansouramycin D suggest that the C10 fragment corresponds to the indole moiety, along with C-3 and C-4, which is likely attached in a similar manner to that proposed in the final biosynthetic steps of mansouramycin D (Fig. 5A). The 4-pyrone-2-carboxylic acid moiety, along with C-1, is derived from a different C7 precursor than the one used in mansouramycin D biosynthesis, suggesting a novel biosynthetic pathway.

The analysis of cluster 16 to identify potential genes involved in the biosynthesis of precursor fragments was unsuccessful. A possible explanation is that the SARP regulator *AfsR* from

cluster 16 may function as a global regulator, influencing native biosynthetic genes in *S. albus* Del14 which activated mansevorone biosynthesis. Supporting evidence was found when we revisited previous heterologous expression studies of various gene clusters containing SARP regulators to examine whether they also produced mansevorone. Among these studies, cluster C1.8, which encodes a PKS-I compound from a cosmid library constructed for *Streptomyces kitasatoensis* (unpublished data), produced a compound with the same mass as mansevorone in *S. albus* Del14_C1.8_pRT-DnrI (Fig. 5B-D). The identity of mansevorone was confirmed through matching retention time, UV/VIS spectra, and MS/MS fragmentation patterns (Fig. S13, S24/25).

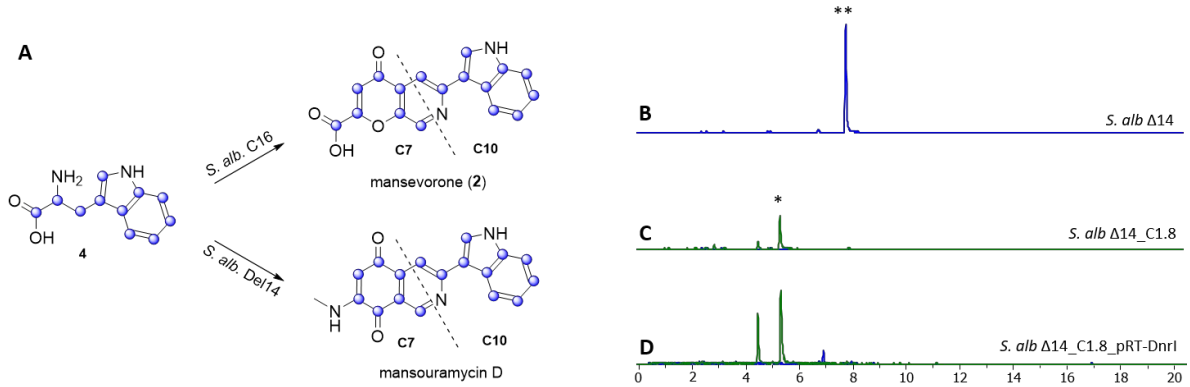


Figure 5: $^{13}C_{11}$-labeled tryptophan (blue dots) is incorporated by *S. albus* Del14_C16 into mansevorone and by *S. albus* Del14 into mansouramycin D; both compounds contain the similar C10 fragment but differ in the C7 fragment, while all fragments are synthesized from tryptophan (**A**). Extracted ion chromatograms showing mansevorone (*) and mansouramycin D (**) production in *S. albus* Del14 (**B**), *S. albus* Del14_C1.8 harboring cluster C1.8 (**C**) and in *S. albus* Del14_C1.8_rRT-DnrI harbouring cluster16 and pRT801_ampery_DnrI for SARP overexpression (**D**).

Similar to the previous case, overexpression of the SARP regulator *DnrI* on cluster C1.8 led to increased production of mansevorone (Fig. 5D). Cluster 16 and cluster 1.8 and their SARP regulators originate from two unrelated strains and do not share any homology (Fig. S26-S28). However, *AsfR* and *DnrI* are both from the same class of *AfsR*/*DnrI*/*RedD* of global transcriptional regulators and seem to activate mansevorone biosynthesis in *S. albus* Del14. Expression the SARP regulator containing plasmid pRT801_cat_ampery_Sk1.8_DnrI in *S. albus* Del14 without the respective clusters C1.8, however, did not lead to mansevorone production. This suggests that there are other genes on cluster C1.8 with regulatory function that rely on the SARP overexpression. However, the production of mansevorone after the heterologous transfer of two different gene clusters containing distinct SARP regulators clearly suggests that the biosynthetic genes for mansevorone are located in the genome of their heterologous host *S. albus* Del14.

Since biosynthesis of mansevorone is likely encoded in the native host *S. albus* Del14 and only expressed in the right conditions, a biosynthetic proposal cannot be made at this stage. A close examination of the biosynthetic mechanism will be conducted in future studies.

Heterologous Production of 5'Cl-Mansouramycin.

In the course of investigating additional heterologous expressed clusters containing SARP regulators in *S. albus* Del 14, no additional strains were identified producing mansevorone (**5**). However, heterologous expression of an NRPS cluster from *S. libani* (cluster 1.7, unpublished data) in *S. albus* Del14 led to the production of a new compound with m/z = 338.0697 showing a chlorinated isotopic pattern and a similar MS/MS fragmentation showing fragments with a difference of 34 Da compared to mansouramycin D (Fig. S33).

The molecular formula was calculated as $C_{18}H_{13}ClN_3O_2$ based on the mass peak m/z = 338.0697 $[M+H]^+$ confirming one chlorine atom attached to mansouramycin D. To determine the position of chlorination, the compound was produced and isolated from *S. albus* Del14_cluster1.7 which yielded 0.3 mg (Fig. 6). Due to insufficient amounts of compound HMBC and ^{13}C NMR spectra could not be obtained. However, using ^{1}H-NMR, COSY and edited-HSQC spectra (Fig. S29-S31), all proton signals except one showed close similarities to mansouramycin D (Table S9). The missing proton signal was located in the phenyl ring of the indole moiety that is part of the mansouramycin D structure. Analysis of ^{1}H-NMR revealed three signals including a singlet (4'-C), a doubled (7'-C) and a doublet of doublets (6'-C) while 7'-C and 6'-C appeared next to each other due to a COSY correlation. This would lead to two different derivatives with chlorination at position 5' or 6'. Comparison of the predicted shifts by ACD labs determined the chlorination at 5'-C resulting in 5'-Cl-mansouramycin.

A chlorinated version of mansouramycin D has not been observed in *S. albus* Del14, indicating that the chlorination is likely mediated by genes encoded within the NRPS cluster. This was confirmed by the absence of 5'-chloromansouramycin and mansouramycin D production after transferring cluster 1.7 into the mansouramycin BGC deficient strain *S. albus* Del15.

A potential gene responsible for chlorination, encoding the tryptophan halogenase enzyme *RebH*, was identified through BLAST analysis of the genes located on cluster 1.7. Deletion of *RebH* and heterologous expression of the resulting construct, *S. albus* Del14_cluster1.7-delRebH, resulted in the depletion of 5'-chloromansouramycin, with only mansouramycin D being produced (Fig. 6D). This indicates that *RebH* is involved in the indole chlorination of mansouramycin D. Since *RebH* is a tryptophan halogenase and the last steps in the mansouramycin biosynthesis are rather flexible, we propose that chlorination occurs on tryptophan prior to its conversion to the C10 precursor, which is subsequently used in the final biosynthetic step of 5'-chloromansouramycin.

The biosynthesis of maramycin and 5'-Cl-mansouramycin demonstrated that the mansouramycin pathway is flexible in its final steps. This flexibility presents an opportunity to

modify the residues at positions C3 and C4, potentially leading to the development of new mansouramycin derivatives with enhanced antitumor activity.

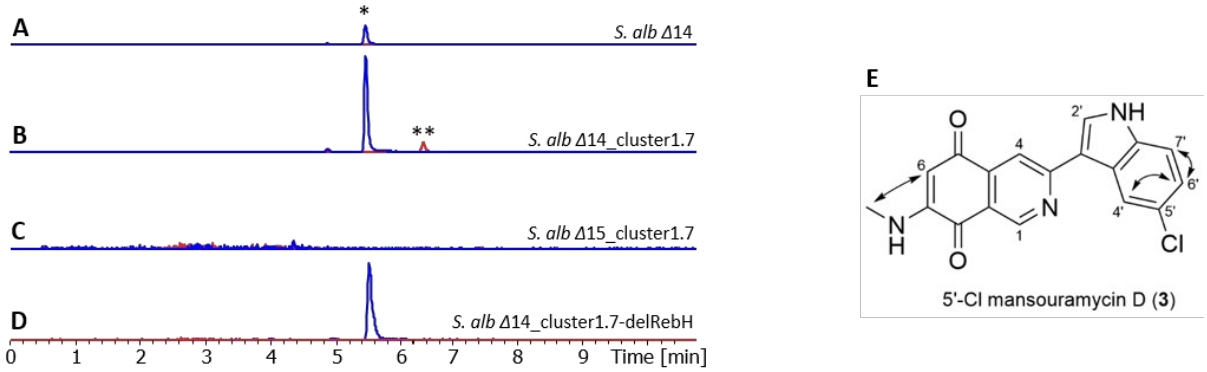


Figure 6: Production of mansouramycin D (*) and 5'-Cl-mansouramycin D (**) in *S. albus* Del14 chassis strain (**A**), the halogenase gene containing cluster 1.7 in *S. albus* Del14 (**B**), the mansouramycin cluster deficient strain *S. albus* Del15 (**C**) and in the *RebH* deletion mutant *S. albus* Del14_cluster1.7-delRebH (**D**). The structure of 5'-Cl-mansouramycin D and key COSY correlations (⌒) are shown in **E**.

Antimicrobial susceptibility test

Poor solubility of malevonin prevented adequate testing of the substance. 5'Cl-mansouramycin was not tested due to insufficient amounts. Mansevorone was tested against a panel of microorganisms, however no activity could be observed.

2.2.4 Conclusions

Three hybrid compounds with structure elements from mansouramycins were discovered after heterologous expression of three distinct BGCs, indicating a biosynthetic interaction between the native mansouramycin genes of *S. albus* Del14 and the introduced genes. This resulted in the rare fluorene and azachromone scaffolds of malevonin (**1**) and mansevorone (**2**) and a 5'-chlorinated version of mansouramycin D (**3**). The dynamic interaction of the mansouramycin biosynthesis with heterologously expressed clusters is intriguing and raises many questing regarding the final biosynthetic steps of mansouramycin D, the interaction of mansouramycin precursors and the regulation of gene expression. The mansouramycin pathway in *S. albus* Del14 is very responsive to heterologously introduced BGCs, which might influence the expression of compounds. This has to be taken in consideration when using *S. albus* Del14 as heterologous expression host. The interaction of biosynthetic gene cluster is the foundation of the cluster evolution and the huge variety of natural products from *Streptomyces*. Understanding the biosynthetic mechanisms underlying these interactions could enable new strategies for targeted production of novel compounds.

2.2.5 Experimental

General Experimental Procedures

All strains, plasmids, BACs and cosmids utilized in this study are listed in table S1-2. *Escherichia coli* strains were cultured in Luria Bertani (LB) medium[29]. Soya flour mannitol agar (MS agar) was used for the cultivation of *Streptomyces* strains, supporting both sporulation and conjugation, alongside tryptic soy broth (TSB; Sigma-Aldrich, St. Louis, MO, USA)[30]. 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) was utilized for expression and secondary metabolite production. Antibiotics such as kanamycin, apramycin, hygromycin, and nalidixic acid were added when necessary.

Genome mining and bioinformatic analysis

The S. sp. LV45-129 genome was screened for secondary metabolite biosynthetic gene clusters using the antiSMASH online tool[31]. Genetic data was analyzed with the Geneious prime 2022.2.2 software[32]. Selected genes were further analyzed using the BLAST online tool from the National Center of Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/) and the Universal Protein Resource (UniProt) (https://www.uniprot.org/blast).

Isolation and Manipulation of DNA

BAC and cosmid isolation from the constructed genomic library of *Streptomyces* LV45-129 and of *S. kitasatoensis*, DNA manipulation, transformation into *E. coli* and intergeneric conjugation between *E. coli* and *Streptomyces* were performed according to standard protocols[29-30, 33]. Purification of cosmids was carried out using the BACMAX™ DNA purification kit (Lucigen, Middleton, WI, USA). All restriction endonucleases were used according to manufacturer's recommendations (New England Biolabs, Ipswich, MA, USA). All of the Primers used in this study are listed in table S3-4.

Construction of the BAC vector containing cluster 3

Two cosmids 01A06 and 12F02 from the *S. kitasatoensis* library, containing left and right parts of the cluster 3, were used in the homologous recombination experiment[7, 34], using yeast[35] (Table S1) and pCLY10 vector[3] (Table S2) containing left and right homologous arms to the both cosmids, integrase gene and origin of replication for *Streptomyces* strains, in order to obtain full cluster 3 as one piece. Correct transformants were verified by the restriction analysis and

sequencing of the isolated BAC DNA (Table S2). As a result, the pBACcluster-3 expression vector was obtained.

The deletion of the orf22 gene, coding for NRPS-1, in cluster 3

Modification of the BAC containing cluster 3 by the Red/ET recombination.

Red/ET recombination technology was used in order to delete the *orf22* gene. For this purpose, linear DNA fragments flanked by suitable homology arms were generated by PCR using Cluster3ReddelNRPS_F and Cluster3ReddelNRPS_R primers (Table S3) and fragment containing hygromycin resistance gene. The PCR reactions were carried out with Phusion DNA polymerase (Thermo Fisher Scientific) according to the manufacturer's protocol. PCR products were concentrated by ethanol precipitation prior to further use. In general, 300 µl overnight culture of *E. coli* GB05-red (Table XX) cells harboring the cluster 3 BAC vector (Table XX) to be modified were inoculated into 15 ml of LB, and cultivated on a shaker at 37 °C and 200 rpm for 2 hours. Thereafter, 400 µl of 10% L-rhamnose was added to the culture to induce the expression of the recombinases. Cultivation was then continued for 45 min. The cells were subsequently harvested by centrifugation, washed twice with ice-cold distilled H_2O, and resuspended in 600 µl of 10% ice-cold glycerol. The PCR product was mixed with the electrocompetent cells, which were then transferred to an ice-cold electroporation cuvette (1 mm). The mixture was subsequently electroporated at 1800 V (Eppendorf electroporator), followed by the addition of 750 µl LB. The cells were cultivated at 37 °C for 90 min before the culture was transferred to LB agar plates with hygromycin. The plates were incubated at 37 °C overnight. Correct transformants were verified by the restriction analysis and sequencing of the isolated BAC DNA using the Cluster3checkF and Cluster3checkR primers (Table S3). As a result, the pBACcluster-3_Delorf22 vector was obtained.

Construction of the *S. albus* Del14_cluster3_delOrf22 Mutant.

The pBACcluster-3_Delorf22 (Table S2) gene disruption vector was transferred from *E. coli* ET12567/pUB307 cells into *S. albus* Del14 cells by means of conjugation. Transconjugants were selected for resistance to hygromycin (50 µg ml−1) and apramycin (50 µg ml−1). As a result, the *S. albus* Del14_cluster3_delOrf22 was obtained (Table S1).

Construction of the SARP expression vector

The SARP expression vector pRT801_ampery_AfsR16 and pRT801_ampery_DnrI were constructed based on the phage BT1 integrative vector pRT801, incorporating the SARP gene under the control of the strong, constitutive promoter TS_81. Amplification of the SARP gene was performed using the primers shown in table S4, which included additional *NdeI* and *EcoRV* restriction sites to facilitate subsequent cloning. AsfR16_Terp_For and C1.8_DnrI_act-ML-f also contained a sequence a 22bp sequence derived from the upstream region of the SARP gene, encompassing the ribosomal binding site. Successful integration of the SARP gene into the vector was confirmed by sequencing using the sequencing primers shown in table S4.

Construction of the BAC1.7-delRebH-delAp vector.

The halogenase *rebH* gene was functionally inactivated by an in-frame deletion of its main coding region. First, the ampicillin resistance gene was amplified using primers delRebH-f and delRebH-r (Table S3). The resulting PCR product was introduced into the BAC1.7 vector via Red/ET-mediated homologous recombination to replace the central portion of the *rebH* gene, yielding the intermediate construct BAC1.7-delRebH-Ap. The ampicillin resistance cassette was subsequently removed by *MssI* restriction digestion followed by self-ligation. The final plasmid, designated BAC1.7-delRebH-delAp, was introduced into *S. albus* Del14 strain via conjugation to assess the effect of the *rebH* gene deletion.

Metabolic Extraction and Analysis.

All *Streptomyces* strains were precultured in 25 mL of TSB medium for 24 h before 1 mL was taken from this seed culture to inoculate 100 mL of the DNPM production medium. These cultures were incubated for seven days at 28 °C. Metabolites were extracted with equal amounts of butanol or ethyl acetate from the supernatant of the culture broth, evaporated and dissolved in methanol. HPLC-MS analysis was performed by separating 1 µL of the extract using a Dionex Ultimate 3000 UPLC (Thermo Fisher Scientific, Waltham, MA, USA), a 10-cm ACQUITY UPLC BEH C18 column, 1.7 µm (Waters, Milford, MA, USA) and a linear gradient of 0.1 % formic acid solution in acetonitrile against 0.1 % formic acid solution in water from 5 % to 95 % in 18 min at a flow rate of 0.6 mL/min. Samples were analyzed using an amaZon speed mass spectrometer or maXis high resolution LC-QTOF system (Bruker, USA). Data were collected and analyzed with the Bruker Compass Data Analysis software, version 4.1 (Bruker, Billerica, MA, USA). Monoisotopic mass was searched in the natural products database DNP (Dictionary of Natural Products)[36].

Malevonin, Mansevorone and 5'-Cl-Mansouramycin Isolation.

S. albus Del14 cluster3, *S. albus* C16h7 and *S. albus* Del14_cluster1.7 each were precultured in 6 flasks, each containing 25 mL of TSB, for 24 h. Subsequently, 100 flasks, each containing 100 mL of DNPM (*S. albus* Del14_C16h7 and Del14_cluster1.7) or SG production medium (*S. albus* Del14_cluster3), were inoculated with 1 mL of the seed culture. The cultures were incubated for 7 days at 28 °C. The mycelial portion was separated by centrifugation, and metabolites were extracted from the supernatant with equal amount of butanol.

Malevonin Purification.

Malevonin precipitated from the crude extract dissolved in methanol. The precipitant was collected, washed with hexane and dried, resuling in the pure comound.

Mansevorone Purification.

In a first step, mansevorone was isolated from the crude mixture by flash chromatography using a Biotage Isolera system (Biotage, Uppsala, Sweden) equipped with CHROMABOND Flash RS 330 C18 ec cartrige (MACHEREY-NAGEL, Düren, Germany) and gradient elution using a mixture of water and methanol as solvents. The enriched fractions were pooled and subjected to size exclusion cloumn chromatography using Sephadex LH-20 (GE Healthcare, Chicago, USA) resin and methanol as eluent. Semi-preparative reversed phase HPLC purification was perfomred on a Agilent 1260 infinity (Agilent Technologies, Kalifornien, USA) equipped with a Synergi C18 column (Phenomenex, Aschaffenburg, Germany) and gradient elution using a 0.1% formic acid solution in acetonitrile and water as the mobile phase. Fractions containing mansevorone were detected by LC-MS analysis as described above, pooled together and evaporated for further experiments.

5'-Cl-Mansouramycin Purification.

5'-Cl-Mansouramycin was purified by size exclusion chromatography using a Sephadex LH-20 column (GE Healthcare, Chicago, USA) and methanol as a solvent. Preparative HPLC of the enriched fractions was carried out on a Waters Autopurification System equipped with a Nucleodur C18 Htec 250/21-C18 5 µm column (Macherey-Nagel, Düren, Germany) and

gradient elution using a 0.1% formic acid solution in acetonitrile and water as the mobile phase. Fractions containing 5'-Cl-Mansouramycin were detected by LC-MS analysis as described above, pooled together and evaporated for further experiment.

Feeding studies.

L-Tryptophan $^{13}C_{11}$ was purchased from Euroisotope. For feeding experiments, L-Tryptophan was prepared as a stock solution at a concentration of 5 mg/mL. From this stock solution, 125 µL was added to 25 mL cultures at inoculation and again after 24, 48, and 72 hours of cultivation. After 96 h, the culture supernatant was extracted with an equal volume of butanol. The solvent was evaporated under a nitrogen stream, and the residue was dissolved in 1 mL of methanol. The resulting extract was subsequently analyzed using LC-MS analysis as described above.

Antimicrobial susceptibility test.

Minimum inhibitory concentrations (MICs) were determined according to standard procedures. Single colonies of the tested strains were suspended in cation adjusted Müller-Hinton broth to achieve a final inoculum of approximately 104 CFU ml-1. Serial dilutions of mansevorone (0.03 to 64 µg mL-1) were prepared in sterile 96-well plates before the strain suspension was added. Growth inhibition was assessed after overnight incubation (16-18 h) at 30-37°C. A panel consisting of the following strains was tested: *B. subtilis* DSM-10, *S. aureus* Newman, *Mycobacterium smegmatis* MC2155, *Citrobacter freundii* DSM-30039, *E. coli* BW25113 (wt), *E. coli* JW0451-2 (ΔacrB), *Pseudomonas aeruginosa* PA14 DSM-19882, *Acinetobacter baumanii* DSM-30008, *Mucor hiemalis* DSM-2656, *Pichia anomala* DSM-6766, *Cryptococcus neoformans* DSM-11959, *Candida albicans* DSM-1665, CHO-K1 and HepG2.

Nuclear Magnetic Resonance Spectroscopy (NMR).

The chemical structures of all the compounds were determined via multidimensional NMR analysis. ^{1}H-NMR, ^{13}C-NMR, and 2D spectra were recorded at 500 MHz (^{1}H)/ 126 MHz (^{13}C), conducted in the Bruker Avance Neo 500 MHz, equipped with a Prodigy Cryo-probe. Samples were dissolved in methanol-d_4 or dimethyl sulfoxide-d_6. Chemical shifts are reported in ppm

relative to tetramethylsilane; the solvent was used as the internal standard. Coupling constants are reported in Hertz (Hz). Multiplicity is reported with the usual abbreviations (s: singlet, br s: broad singlet, d: doublet, dd: doublet of doublets, ddd: doublet of doublet of doublets, t: triplet, dt: doublet of triplets, q: quartet, p: pentet, dp: doublet of pentets, m: multiplet.

Author contributions

M.S., L.H., P.O. and M.L. designed experiments. M.S., L.H., P.O., M.L., A.P., P.C. and C.R. performed experiments. M.S. solved NMR structures, M.S. and P.O. wrote the manuscript. All authors edited the manuscript. A.L. and J.Z. supervised the work.

2.2.6 References

1. Gullón, S.; Olano, C.; Abdelfattah, M. S.; Braña, A. F.; Rohr, J.; Méndez, C.; Salas, J. A., Applied and environmental microbiology 2006, 72, 4172-4183.
2. Liu, X.; Liu, D.; Xu, M.; Tao, M.; Bai, L.; Deng, Z.; Pfeifer, B. A.; Jiang, M., Journal of natural products 2018, 81, 72-77.
3. Bilyk, O.; Sekurova, O. N.; Zotchev, S. B.; Luzhetskyy, A., PloS one 2016, 11, e0158682.
4. Lombó, F.; Velasco, A.; Castro, A.; De la Calle, F.; Braña, A. F.; Sánchez-Puelles, J. M.; Méndez, C.; Salas, J. A., ChemBioChem 2006, 7, 366-376.
5. Myronovskyi, M.; Brötz, E.; Rosenkränzer, B.; Manderscheid, N.; Tokovenko, B.; Rebets, Y.; Luzhetskyy, A., Applied microbiology and biotechnology 2016, 100, 9175-9186.
6. Wendt-Pienkowski, E.; Huang, Y.; Zhang, J.; Li, B.; Jiang, H.; Kwon, H.; Hutchinson, C. R.; Shen, B., Journal of the American Chemical Society 2005, 127, 16442-16452.
7. Myronovskyi, M.; Rosenkränzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A., Metabolic engineering 2018, 49, 316-324.
8. Maleckis, M.; Wibowo, M.; Williams, S. E.; Gotfredsen, C. H.; Sigrist, R.; Souza, L. D.; Cowled, M. S.; Charusanti, P.; Gren, T.; Saha, S., ACS chemical biology 2024.
9. Hawas, U. W.; Shaaban, M.; Shaaban, K. A.; Speitling, M.; Maier, A.; Kelter, G.; Fiebig, H. H.; Meiners, M.; Helmke, E.; Laatsch, H., Journal of natural products 2009, 72, 2120-2124.
10. Shaaban, M.; Shaaban, K. A.; Kelter, G.; Fiebig, H. H.; Laatsch, H., Marine Drugs 2021, 19, 715.
11. Shuai, H.; Myronovskyi, M.; Rosenkränzer, B.; Paulus, C.; Nadmid, S.; Stierhof, M.; Kolling, D.; Luzhetskyy, A., ACS Chemical Biology 2022, 17, 598-608.
12. Sikandar, A.; Lopatniuk, M.; Luzhetskyy, A.; Koehnke, J., ACS Chemical Biology 2020, 15, 2815-2819.
13. Dahlem, C.; Siow, W. X.; Lopatniuk, M.; Tse, W. K.; Kessler, S. M.; Kirsch, S. H.; Hoppstädter, J.; Vollmar, A. M.; Müller, R.; Luzhetskyy, A., Cancers 2020, 12, 1288.
14. Lopatniuk, M.; Riedel, F.; Wildfeuer, J.; Stierhof, M.; Dahlem, C.; Kiemer, A. K.; Luzhetskyy, A., Metabolic Engineering 2023, 78, 48-60.

15. Lasch, C.; Stierhof, M.; Estévez, M. R.; Myronovskyi, M.; Zapp, J.; Luzhetskyy, A., Microorganisms 2020, 8, 1800.
16. Lasch, C.; Stierhof, M.; Estévez, M. R.; Myronovskyi, M.; Zapp, J.; Luzhetskyy, A., Microorganisms 2021, 9, 1640.
17. Stierhof, M.; Myronovskyi, M.; Zapp, J.; Luzhetskyy, A., Journal of Natural Products 2023, 86, 2258-2269.
18. Hug, J. J.; Dastbaz, J.; Adam, S.; Revermann, O.; Koehnke, J.; Krug, D.; Müller, R., ACS chemical biology 2020, 15, 2221-2231.
19. Hasebe, F.; Matsuda, K.; Shiraishi, T.; Futamura, Y.; Nakano, T.; Tomita, T.; Ishigami, K.; Taka, H.; Mineki, R.; Fujimura, T., Nature Chemical Biology 2016, 12, 967-972.
20. Saikawa, Y.; Hashimoto, K.; Nakata, M.; Yoshihara, M.; Nagai, K.; Ida, M.; Komiya, T., Nature 2004, 429, 363-363.
21. Harada, K.-i.; Fujii, K.; Hayashi, K.; Suzuki, M.; Ikai, Y.; Oka, H., Tetrahedron letters 1996, 37, 3001-3004.
22. Oberhäuser, P.; Myronovskyi, M.; Stierhof, M.; Gromyko, O.; Luzhetskyy, A., *Microbial Cell Factories* 2025, **24**, 1-11
23. Liu, G.; Chater, K. F.; Chandra, G.; Niu, G.; Tan, H., Microbiology and molecular biology reviews 2013, 77, 112-143.
24. Gregory, M. A.; Till, R.; Smith, M. C., Journal of bacteriology 2003, 185, 5320-5323.
25. Wang, F.; Allen, D.; Tian, S.; Oler, E.; Gautam, V.; Greiner, R.; Metz, T. O.; Wishart, D. S., Nucleic acids research 2022, 50, W165-W174.
26. Wang, F.; Liigand, J.; Tian, S.; Arndt, D.; Greiner, R.; Wishart, D. S., Analytical chemistry 2021, 93, 11692-11700.
27. Djoumbou-Feunang, Y.; Pon, A.; Karu, N.; Zheng, J.; Li, C.; Arndt, D.; Gautam, M.; Allen, F.; Wishart, D. S., Metabolites 2019, 9, 72.
28. Allen, F.; Pon, A.; Wilson, M.; Greiner, R.; Wishart, D., Nucleic acids research 2014, 42, W94-W99.
29. Green, M. R.; Sambrook, J., A Laboratory Manual 4th 2012, 448.
30. Kieser, T.; Bibb, M.; Buttner, M.; Chater, K.; Hopwood, D., Norwich Research Park, Colney 2000, 44-61.
31. Blin, K.; Shaw, S.; Augustijn, H. E.; Reitz, Z. L.; Biermann, F.; Alanjary, M.; Fetter, A.; Terlouw, B. R.; Metcalf, W. W.; Helfrich, E. J., Nucleic acids research 2023, 51, W46-W50.
32. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C., Bioinformatics 2012, 28, 1647-1649.
33. Rebets, Y.; Kormanec, J.; Luzhetskyy, A.; Bernaerts, K.; Anné, J., Metagenomics: Methods and Protocols 2017, 99-144.
34. Myronovskyi, M.; Welle, E.; Fedorenko, V.; Luzhetskyy, A., Applied and environmental microbiology 2011, 77, 5370-5383.
35. Baker Brachmann, C.; Davies, A.; Cost, G. J.; Caputo, E.; Li, J.; Hieter, P.; Boeke, J. D., Yeast 1998, 14, 115-132.
36. Buckingham, J., Dictionary of natural products on CD-ROM. Bocakaton. Chapman & Hall/CRC Press: 2005.

2.2.7 Supplementary information

Table S1: Bacterial strains used in this study

Strain	Description	Source
S. lividans Del8	Cluster free derivative of *S. lividans* TK24	Yousra A., et al.[1]

S. albus Del14	Cluster free derivative of *S. albus* J1074	Myronovskyi, M., et al.[2]
S. albus Del15	Derivative of *S. albus* Del14 with the deletion of mansouramycin biosynthetic pathway	Shuai, H., et al.[3]
Escherichia coli ET12567 pUB307	Donor strain for intergeneric conjugation	Flett, F. et al.[4]
Escherichia coli DH10β	General cloning strain	Grant, S.G., et al[5]
Escherichia coli GB05-red	*E. coli* strain used for the Red/ET recombination	Zhang et al.[6]
Saccharomyces cerevisiae BY4742	Auxotrophic strain used for homologous recombination in yeast	Baker Brachmann, C., et al[7]
S. albus Del14_cluster3	Derivative of *S. albus* Del14 harboring the cluster 3	this study
S. albus Del15_cluster3	Derivative of *S. albus* Del15 harboring cluster 3	this study
S. albus Del14_cluster3_delOrf22	Derivative of *S. albus* Del14 harboring the cluster 3 with deletion of orf22 gene	this study
S. albus Del15_C16	Derivative of *S. albus* Del15 harboring cluster 16	this study
S. albus Del14_C16	Derivative of *S. albus* Del14 harboring cluster 16	this study
S. albus Del14_C16h7	Derivative of *S. albus* Del14 harboring the cluster 16 and pRT801_ampery-AfsR16 plasmid	this study
Streptomyces LV45-129	The wild-type strain; the source of cluster 16	this study
S. albus Del14_C1.8	Derivative of *S. albus* Del14 harboring BAC_C1.8	this study
S. albus Del14_C1.8_pRT-DnrI	Derivative of *S. albus* Del14 harboring BAC_C1.8 and pRT801_cat_ampery_Sk1.8_DnrI	this study
S. albus Del14_cluster1.7	Derivative of *S. albus* Del14 harboring BAC1.7	this study
S. albus Del14_cluster1.7-delRebH	Derivative of *S. albus* Del14 harboring BAC1.7 with RebH deletion	this study
S. albus Del15_cluster1.7	Derivative of *S. albus* Del15 harboring BAC1.7	this study

Table S2: Plasmids and BACs used in this study

Name	Despriction	Source
cluster 3	cluster originating from a COSMID library created for *Streptomyces kitasatoensis*	this study
cluster 3_ Delorf22	cluster 3 with deletion of NRPS gene *orf22*	this study
pCLY10	Shuttle vector for *E.coli*-yeast-actinomycetes	Bilyk O., et al[8]
cluster 16	cluster originating from a COSMID library created for *Streptomyces* strain LV45-129	Oberhäuser et al.[9]
pRT801	Plasmid; BT1 integrative vector for *Streptomyces* strains; apramycin resistance	Gregory, M.A.[10]
pRT801_cat_ampery	Derivative of pRT801 containing ampicilin-erythromycin resistance cassette and strong synthetic promoter TS81	Shuai, H., et al.[3]
pRT801_ampery_AfsR 16	Derivative of pRT801_cat_ampery, catalase gene was replaced with the SARP gene from cluster 16	this study
BAC_C1.8	Cluster originating from a COSMID library created for *Streptomyces kitasatoensis*	this study
pRT801_cat_ampery _Sk1.8_DnrI	Derivative of BT1 integrative vector pRT801_cat_ampery, catalase gene was replaced with the SARP gene from cluster 1.8	this study
BAC1.7	cluster 1.7	this study
BAC1.7-delRebH-delAp	cluster 1.7 with in-frame deletion of *RebH* halogenase gene	this study

Table S3: Primers used for deletion of genes

Primer	Sequence	Description
Cluster3ReddelNRPS_F	TTCCAGATCCGCCGCGGGTGCTTCGGTGCCCCGGAGGCCGGC GAGGTCGTGGACGAAGGCGGCCAGCAGCTCAACGGGAATCC TGCTCTGCGAG	Forward primer for the deletion of the NRPS-1 gene in the cluster 3.
Cluster3ReddelNRPS_R	CTGATGACCGACGCGACGCCCCGGCAGACGACGATCGAGGA GAAGCGGCGCGCGCTGCTGGCCCTGCGGCTGTCTTGTAGGCT GGAGCTGCTTC	Reverse primer for the deletion of the NRPS-1 gene in the cluster 3.
Cluster3checkF	ATCAGACCGGTGCGCTCCAGG	Forward primer for checking the deletion of the NRPS-1 gene.
Cluster3checkR	ACCGACTGCTGAGTGAGGAAG	Reverse primer for checking the deletion of the NRPS-1 gene.
delRebH-f	CTCAAGGCGGCCTACGGTGACCGCATCAACGTCACTCTGGTG GAATCCGAGTTTAAACAGCTGTTCCGGGGATCCGTC	Amplification of ampicilin cassette for in frame deletion of RebH, forward.
delRebH-r	CGCCCGGAGCCGCGCGAGCTCGGCGCGCGCGTTCGTCGGGTC GATGTGCGGTTTAAACTGTAGGCTGGAGCTGCTTCG	Amplification of ampicilin cassette for in frame deletion of *RebH*, reverse.
RebH-chk-F	ATGGCGCGAAGAATCTGCCG	Sequencing of deletion construct for verification, forward.
RebH-chk-R	GAAGGGTGGCCATGGTGAGG	Sequencing of deletion construct for verification, reverse.

Table S4: Primers used for SARP overexpression

Primer	Sequence	Description
AsfR16_Terp_For	CATATGGTCGAGCAACGGAGGTACGGACGTGGAGTTCCGGCTGCTCGG	*AsfR* amplification primer including rbs and NdelI restriction side
AsfR16_Terp_Rev	GATATCTCAGGTCCGCGCCATGAGCCGGACCAT	*AsfR* amplification primer including EcoRV restriction side
pRT_chk_For	TAGTTCCTTCGTCACCACAG	Primers for sequencing from pRT801 region
pRT_chk_Rev	TCACTCATTAGGCACCCCAG	Primers for sequencing from pRT801 region
AfsR16_chk1	CTCGACTGGTTCCACCGC	Primers for sequencing from *AsfR* Region
AfsR16_chk2	GTTTCAGCGACACCGGAGA	Primers for sequencing from *AsfR* Region
AfsR16_chk3	CGAACTGTCCGACGAGTCAC	Primers for sequencing from *AsfR* Region
AfsR16_chk4	CAGGACCTGGCGGATCATC	Primers for sequencing from *AsfR* Region
C1.8_DnrI_act-ML-f	AAAACATATGGTCGAGCAACGGAGGTACGGACATGCACGCACTCCGGACG	*DnrI* amplification primer including rbs and NdelI restriction side
C1.8_DnrI_act-r	TTTTGATATCTCAGCACCCCCTCAGCAC	*DnrI* amplification primer including EcoRV restriction side
prt801_cat-exchange_chk_f	TGAAGGAGGAAGACGAAGCG	Primers for sequencing from pRT801 region
prt801_cat-exchange_chk_r	AGTCAGTGAGCGAGGAAGC	Primers for sequencing from pRT801 region
Sk1.8_DnrI_seq-1	TGGAGTCCTACGCCGAGG	Primers for sequencing from *DnrI* Region
Sk1.8_DnrI_seq-2	TTACTGATGGAGCTCGCC	Primers for sequencing from *DnrI* Region
Sk1.8_DnrI_seq-3	TACTTCGTCATGCAGCTCCT	Primers for sequencing from *DnrI* Region
Sk1.8_DnrI_seq-4	TGCACATCCAGCTGGGGC	Primers for sequencing from *DnrI* Region

Table S5: Proposed function of the genes from cluster 3 and similarity comparison to the *vzb* gene cluster

	Cluster 3	Vzb cluster	
ORF	**Proposed function**	**Corresponding gene**	**% similarity to cluster 3**
1	ATP-dependent protease HslVU (ClpYQ), peptidase subunit		
2	Putative protein		
3	autoinducer-binding transcriptional regulator		
4	SARP		
5	SARP		
6	acyl-CoA_dehydrogenase	Vzb5	49
7	Cyclohexane-1-carbonyl-CoA dehydrogenase	vzb6	66
8	Thioesterase	vzb14	58
9	sulfate adenyl transferase	vzb18	77
10	Sulfate adenylyltransferase subunit 2	vzb19	86
11	Adenylyl-sulfate kinase	vzb20	73
12	Sulfotransferase family protein	vzb21	65
13	Alpha-aminoadipate carrier protein LysW, vzb22	vzb22	67
14	acetyl-CoA_carboxylase_biotin_carboxylase, vzb23	vzb23	70
15	N-acetyl-gamma-glutamyl-phosphate_reductase, vzb24	vzb24	75
16	acetylglutamate kinase-like protein LysZ, vzb25	vzb25	72
17	acetylornithine_deacetylase, vzb26	vzb26	69
18	transketolase, vzb27	vzb27	62
19	1-deoxy-D-xylulose-5-phosphate_synthase	vzb28	70
20	Glyoxalase/bleomycin resistance protein/dioxygenase		
21	major_facilitator_transporter		
22	NRPS		
23	NRPS		
24	Argininosuccinate lyase/adenylosuccinate lyase	vzb2	57
25	PCP	vzb8	50
26	Aminotransferase, vzb9	vzb9	70
27	Butirosin biosynthesis protein H, N-terminal, Azi29	vzb10	69
28	Azi28	vzb11	69
29	Acetyltransferase (GNAT) family protein (there is an acetyl group in the product)	vzb12	49
30	Succinyl-diaminopimelate desuccinylase	vzb13	69
31	cytochrome_P450		
32	NRPS	vzb15	59
33	mbtH-like protein	vzb16	49
34	Putative protein		
35	LuxR family DNA-binding response regulator	vzb17	50
36	Putative protein	vzb0	14
37	Helix-turn-helix domain protein		
38	Hypothetical protein		
39	LysR family transcriptional regulator		
40	2-amino-3-ketobutyrate coenzyme A ligase		
41	Alcohol dehydrogenase		
42	AsnC family transcriptional regulator		
43	L-asparaginase		
44	Aspartate ammonia-lyase		
45	NAD-dependent malic enzyme		
46	TetR_family_transcriptional_regulator		
47	GCN5-related_N-acetyltransferase		
48	Putative membrane protein		
49	Protease HtpX		
50	Acyl-CoA dehydrogenase		

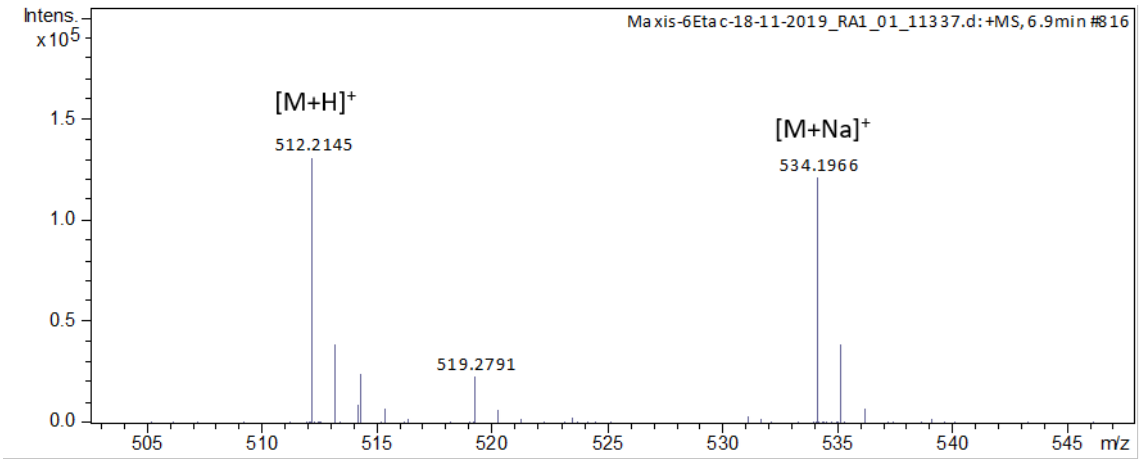


Figure S1: HRMS spectrum of malevonin.

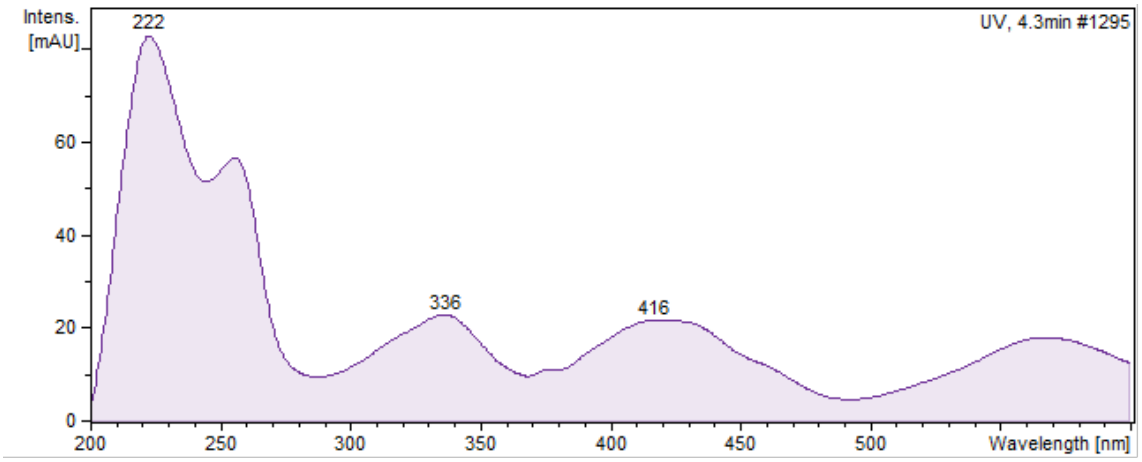
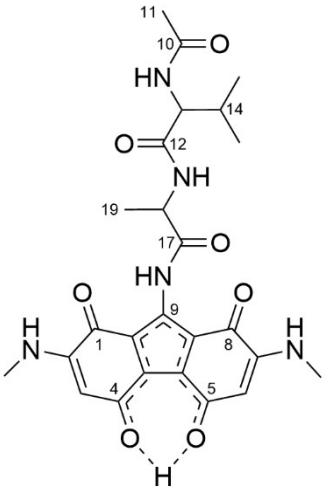


Figure S2: UV/VIS spectrum of malevonin.

Table S6: NMR data (500 MHz, DMSO-d_6) of malevonin.

residue	no	δ(^{13}C) [ppm], type	δ(^{1}H) [ppm], mult (J)	COSY (H-)	HMBC (C-)
AHMFT*	1/8	172.3, 2xC			3,6
	1a/8a	115.1, 2xC			9-NH, 3/6
	2/7	157.7, 2xC			3/6, 2/7-NMe
	2/7-NH		8.94, 2	2/7-NMe	
	2/7-NMe	30.8 2xCH_3	2.93, s	2/7-NH	
	3/6	94.5, 2xCH	5.53. s	2/7-NMe	
	4/5	178.2, 2xC			3/6
	4a/5a	124.8, 2xC			3/6
	9	138.5, C			9-NH
	9-NH		9.89, s		
Ac	10	170.1, C			13-NH, 13, 11
	11	23.3, CH_3	1.87, s		-
Val	12	171.7, C			18-NH, 13.14.18
	13	58.2, CH	4.22. t (7.9)	13-NH,14	13-NH, 14, 15, 16
	13-NH		7.91, d (8.9)	13	
	14	31.3, CH	1.94, m (6.7)	13, 15, 16	13-NH, 13, 15, 16
	15	20.1, CH_3	0.84, d (6.7)	14	12, 13, 16
	16	19.0, CH_3	0.81, d (6.7)	14	12 ,13 ,15
Ala	17	170.2, C			9-NH, 18, 18-NH, 19
	18	49.6, CH	4.53, p (6.9)	18-NH, 19	18-NH, 19
	18-NH		8.16, d (7.0)	18	
	19	18.6, CH_3	1.38, d (7.0)	18	18-NH, 18

*9-amino-5-hydroxy-2,7-bis(methylamino)-1H-fluorene-1,4,8-trione (AHMFT)



malevonin (**1**)

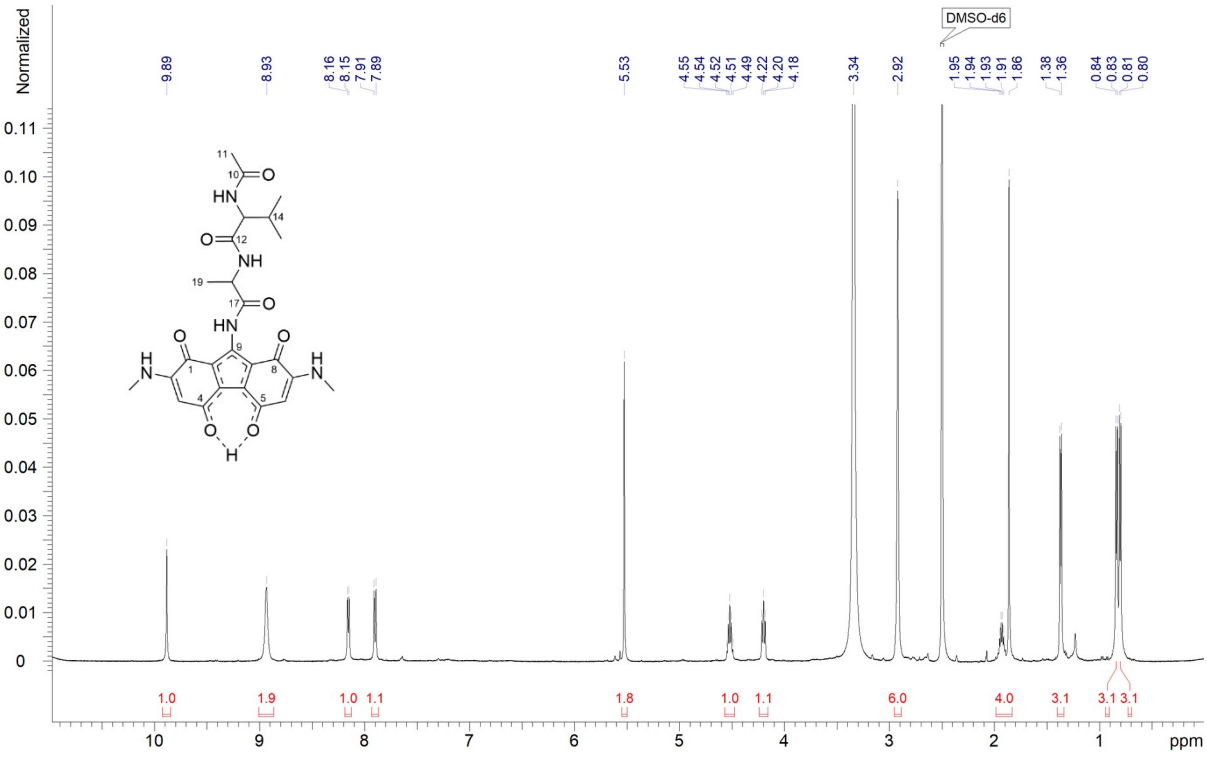


Figure S3: ^{1}H NMR spectrum (500 MHz, DMSO-d_6) of malevonin.

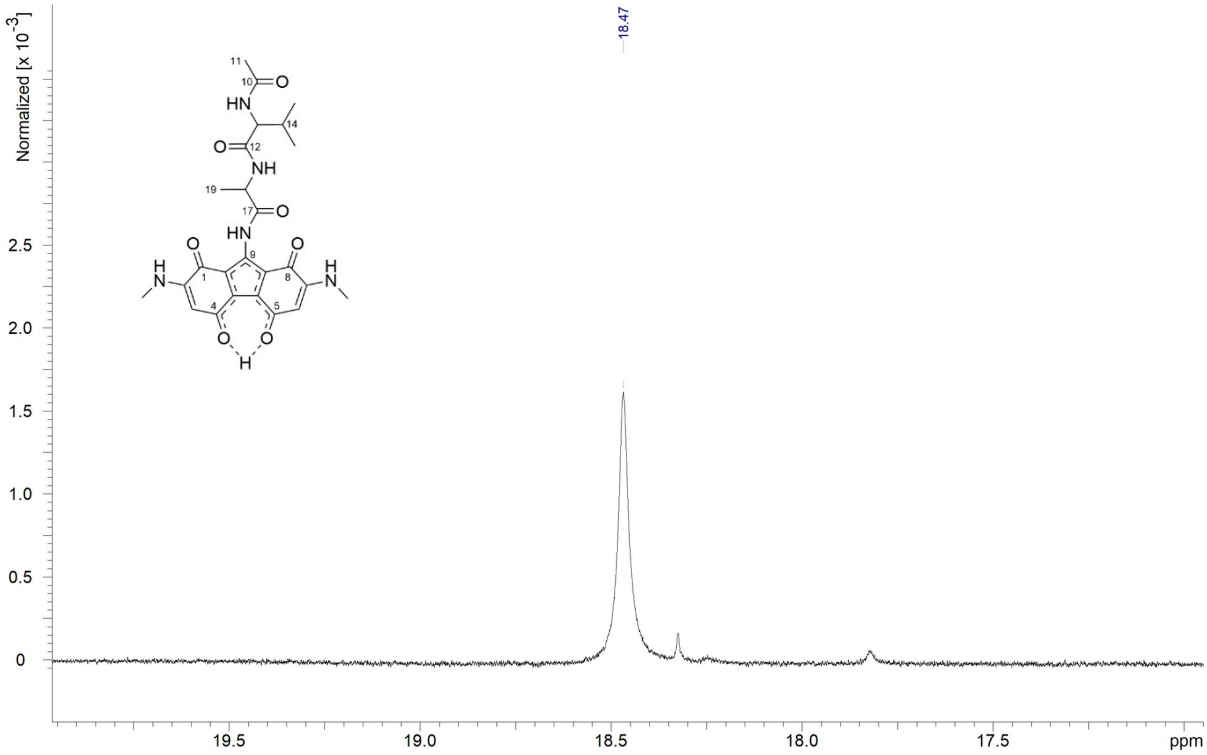


Figure S4: ^{1}H NMR spectrum (500 MHz, DMSO-d_6) of malevonin showing the hydrogen bridge stabilized proton signal in the downfield.

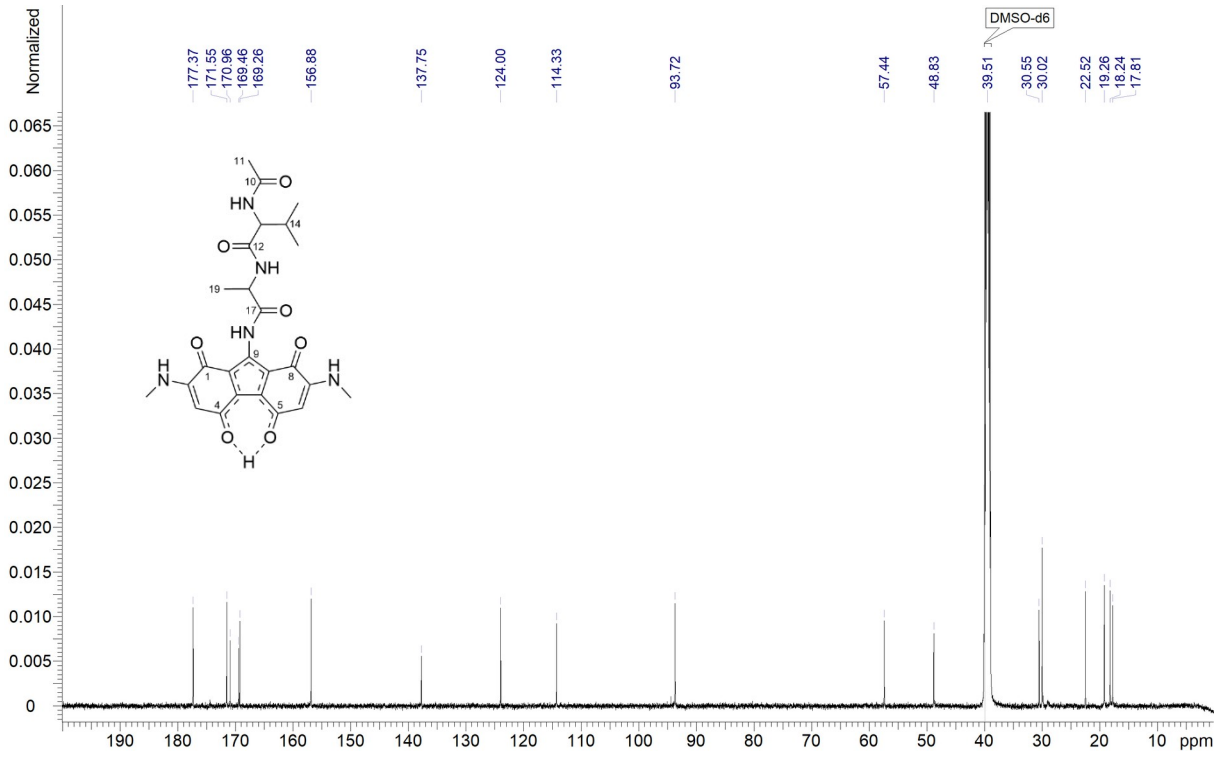


Figure S5: ^{13}H NMR spectrum (125 MHz, DMSO-d_6) of malevonin.

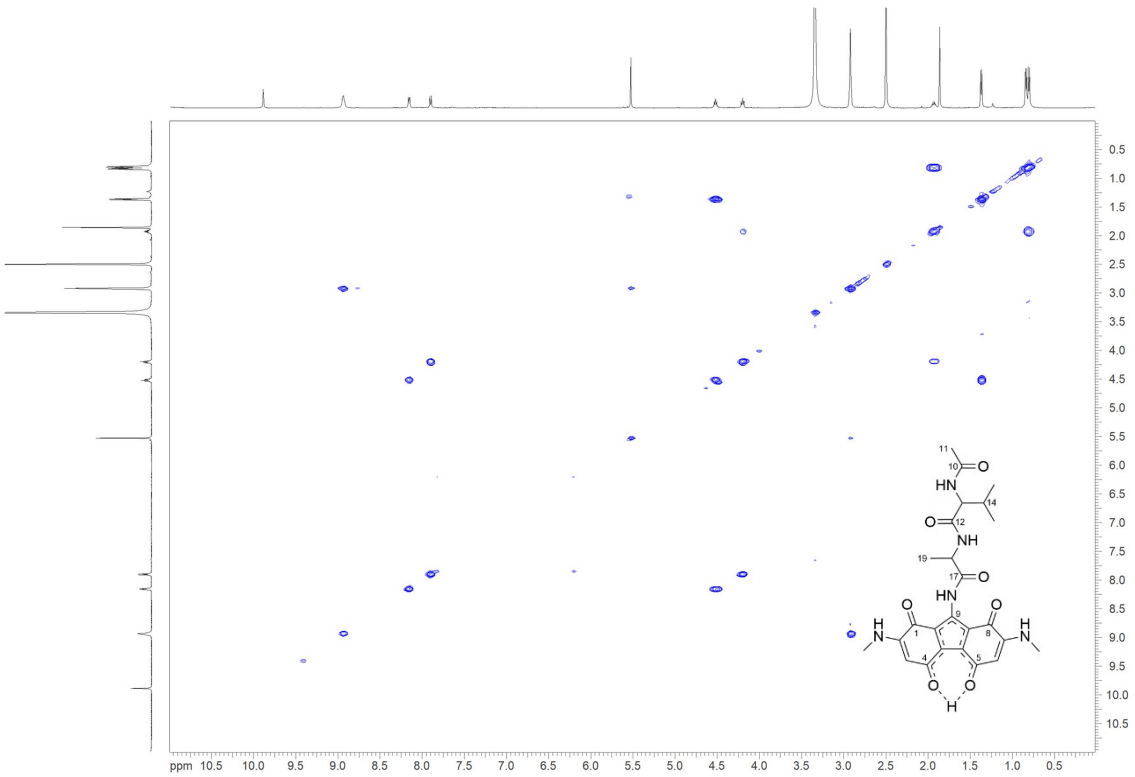


Figure S6: COSY spectrum (DMSO-d_6) of malevonin.

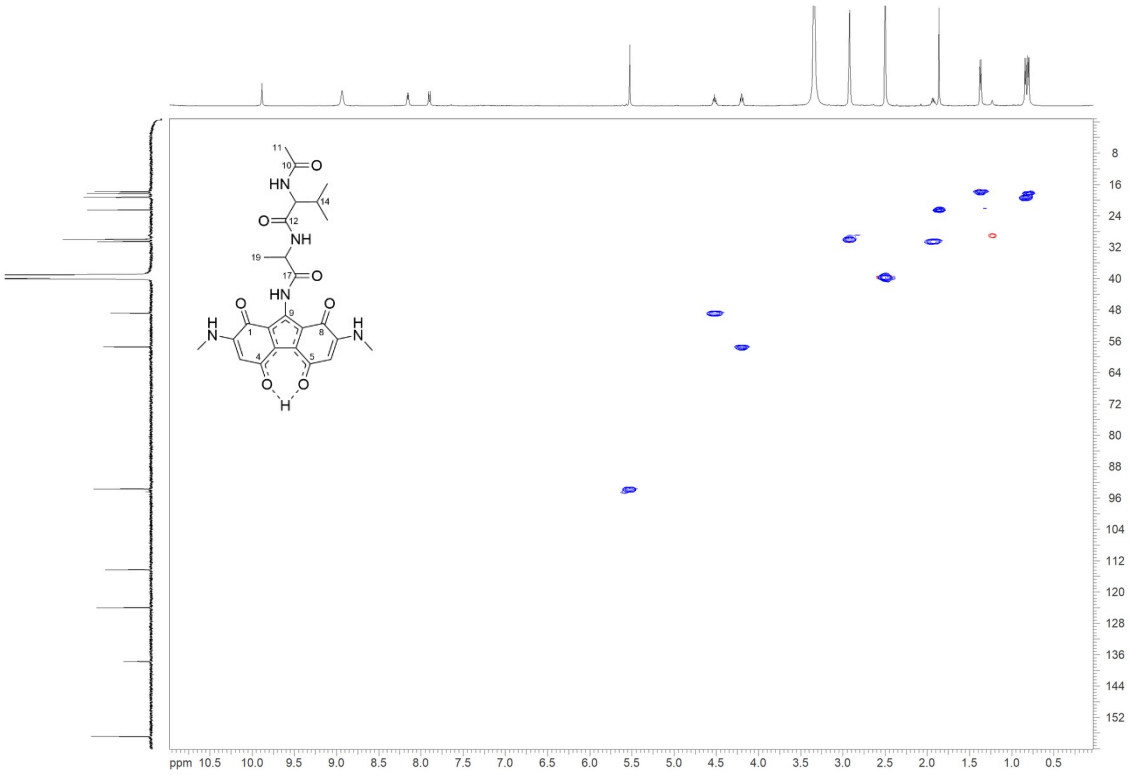


Figure S7: Edited HSQC spectrum (DMSO-d_6) of malevonin.

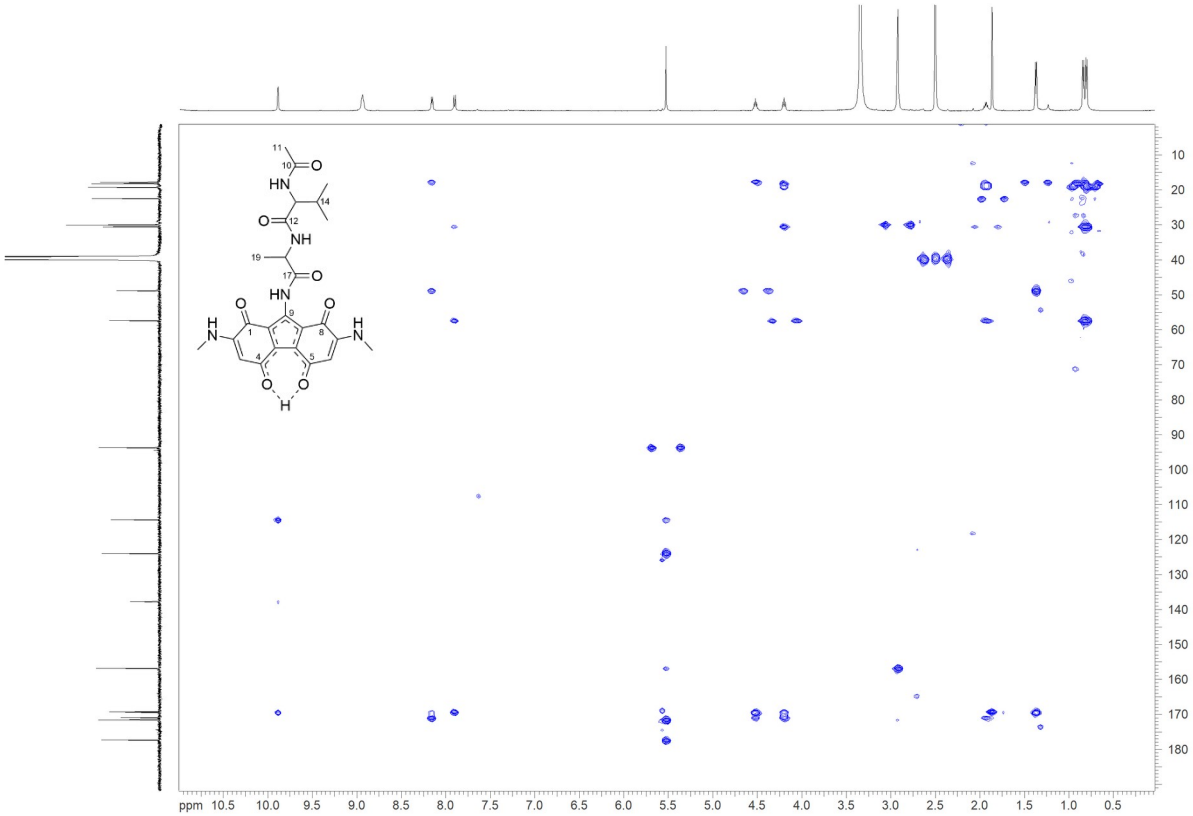


Figure S8: HMBC spectrum (DMSO-d_6) of malevonin.

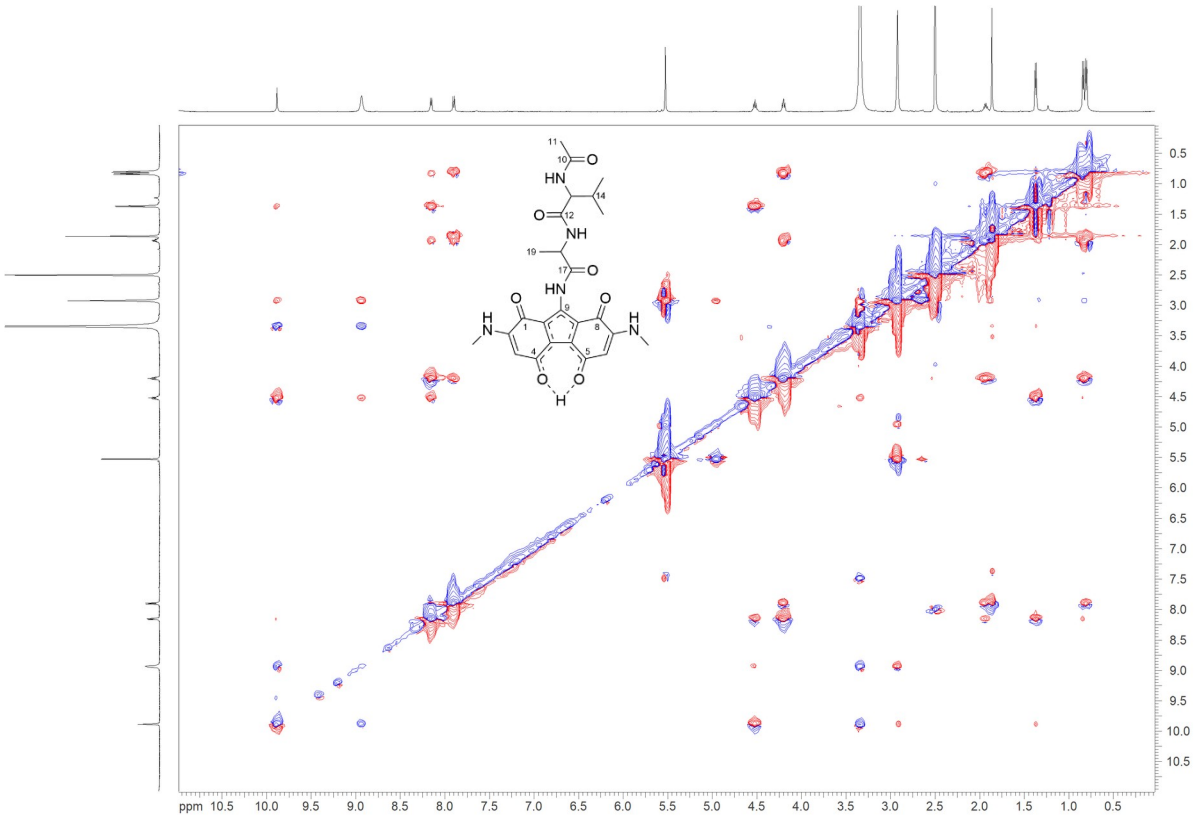


Figure S9: ROESY spectrum (DMSO-*d*6) of malevonin.

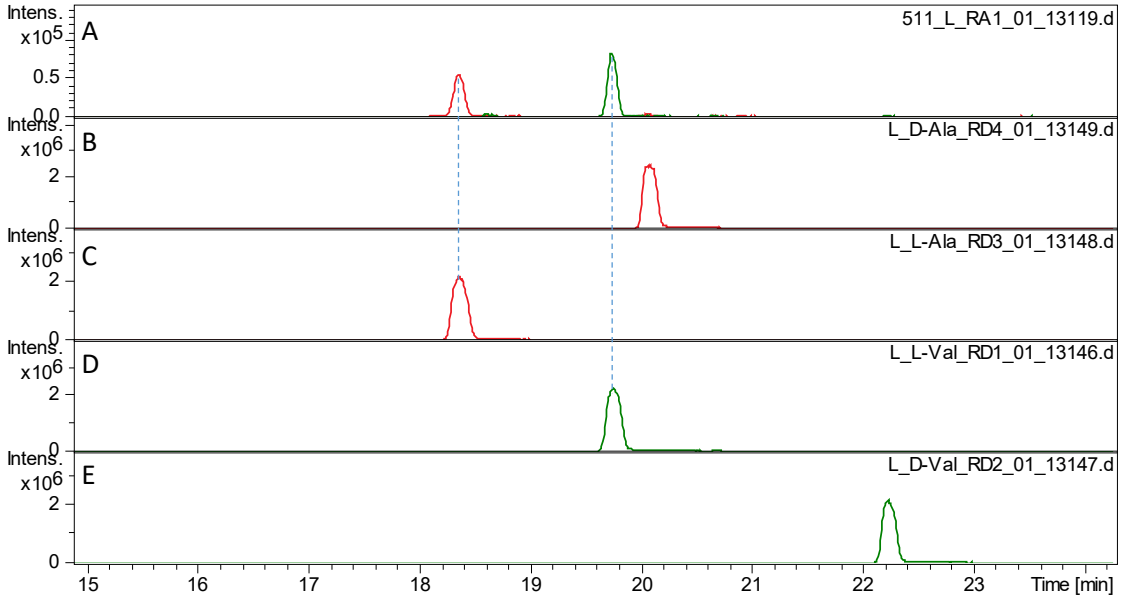


Figure S10: Determination of the absolute stereochemistry by Marfey's method using L-FDLA Derivatization agent with (**A**) malevonin hydrolyzed with 6M HCl and the standards (**B**, **C**) D/L-alanine and (**D**, **E**) L/D-valine.

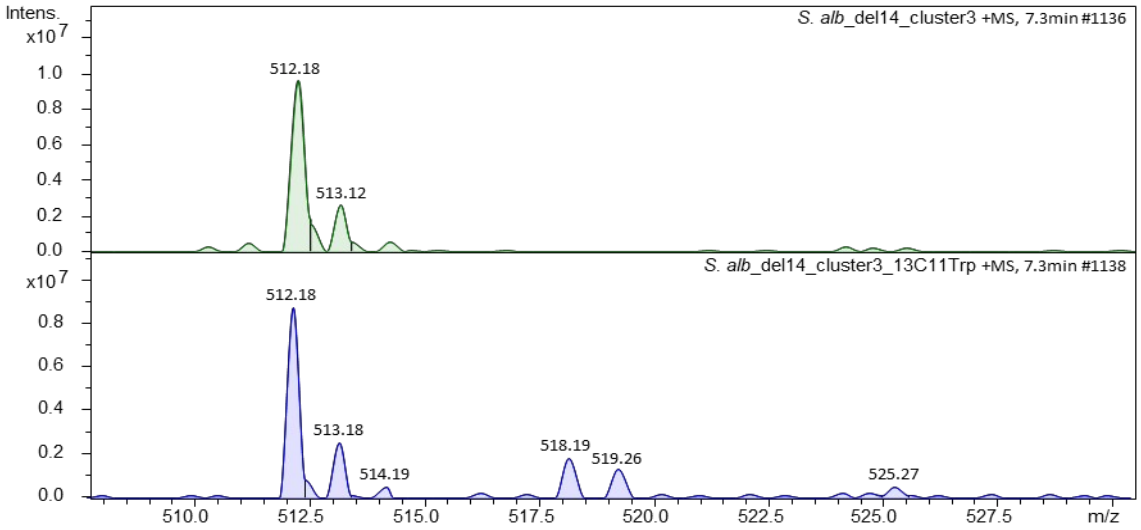


Figure S11: Feeding of *S. albus* Del14 with $^{13}C_{11}$ Trp (**2**) revealed (**B**) a mass of 518.18, 519.26 and 525.27 corresponding to +6, +7 and +13 Da compared to malevonin without labeling (**A**).

```
Consensus                       SYGELDXTXTRLAXVLRXXFGVGPDAVXGVLLXRGLDLPVAQLGVLKAGGAWLPLDPQHP  60
LU18909_1 - NRPS A-domain Orf24 SYGELDTTSTRLAYVLRTAFGVGPDAVGGVLLERGLDLPVAQLGVLKAGGAWLPLDPQHP  60
LU18909_1 - NRPS A-domain Orf25 SYGELDATATRLAHVLRALFGVGPDAVWGVLLDRGLDLPVAQLGVLKAGGAWLPLDPQHP  60

Consensus                       XERLXXXLXDAXAXXVVTTRALAXALPADXPRLXLDDPXXXXXLXRTPXTPLXXXXTXPX 120
LU18909_1 - NRPS A-domain Orf24 AERLTGQLADAGARVVVTTRALAHALPADVPRLCLDDPTTTHPLDRTPDTPLPHSGTKPE 120
LU18909_1 - NRPS A-domain Orf25 TERLASYLSDAGAAAVVTTRALAQALPADVPRLVLDDPAPRERLARTPETPLT-GNTDPD 119

Consensus                       XLAYJIYTSGSTGAPKGVMVSHGAAVNFVLNARELFGIRPGDRLLQFANPAFDVSVFDFY 180
LU18909_1 - NRPS A-domain Orf24 NLAYLIYTSGSTGAPKGVMVSHGAAVNFVLNARELFGIRPGDRLLQFANPAFDVSVFDFY 180
LU18909_1 - NRPS A-domain Orf25 HLAYLIYTSGSTGAPKGVMVSHGAAVNFVLNARELFGIRPGDRLLQFANPAFDVSVFDFY 179

Consensus                       GALGSGAAVVGASREXLLDPDALQELLVRERVSVADVPPAVLRLLDPGSLPDLRALFVGL 240
LU18909_1 - NRPS A-domain Orf24 GALGSGAAVVGASREVLLDPDALQELLVRERVSVADVPPAVLRLLDPGSLPDLRALFVGL 240
LU18909_1 - NRPS A-domain Orf25 GALGSGAAVVGASRELLDPDALQELLVRERVSVADVPPAVLRLLDPGSLPDLRALFVGL 239

Consensus                       EAFPAELVNRWSSEKREFHNGYGPTEATVACVDYLCPPGGLSASPPIGRAMANHRAYVLN 300
LU18909_1 - NRPS A-domain Orf24 EAFPAELVNRWSSEKREFHNGYGPTEATVACVDYLCPPGGLSASPPIGRAMANHRAYVLN 300
LU18909_1 - NRPS A-domain Orf25 EAFPAELVNRWSSEKREFHNGYGPTEATVACVDYLCPPGGLSASPPIGRAMANHRAYVLN 299

Consensus                       AETFEPVPVGVPGELFVAGAGLARGYLNRPDLTAERFVPDPFSGSGERMYRTGDVVRWRE 360
LU18909_1 - NRPS A-domain Orf24 AETFEPVPVGVPGELFVAGAGLARGYLNRPDLTAERFVPDPFSGSGERMYRTGDVVRWRE 360
LU18909_1 - NRPS A-domain Orf25 AETFEPVPVGVPGELFVAGAGLARGYLNRPDLTAERFVPDPFSGSGERMYRTGDVVRWRE 359

Consensus                       DGNLEFLGRADRQXKIRGLRIEPGEIEHALXXXXGVXQXXV                  401
LU18909_1 - NRPS A-domain Orf24 DGNLEFLGRADRQVKIRGLRIEPGEIEHALAGCEGVRQGTV                  401
LU18909_1 - NRPS A-domain Orf25 DGNLEFLGRADRQIKIRGLRIEPGEIEHALTTSPGVAQAVV                  400
```

Figure S12: Alignment of *Orf24* and *Orf25* from cluster 3 shows 90% similarity.

Table S7: Proposed functions of genes within cluster 16.

ORF	Proposed Function	GeneBank homologue
1	VOC family protein	WP_150477657.1
2	NAD(P)/FAD-dependent oxidoreductase	WP_150477658.1
3	CPBP family glutamic-type intramembrane protease	WP_150477659.1
4	helix-turn-helix domain-containing protein	WP_150478040.1
5	M64 family metallopeptidase	WP_055536098.1
6	LacI family DNA-binding transcriptional regulator	WP_150477660.1
7	nucleoside hydrolase	WP_055536099.1
8	MFS transporter	WP_055536100.1
9	terpene synthase family protein	WP_055536101.1
10	family 2B encapsulin nanocompartment shell protein	WP_055536102.1
11	hypothetical protein	WP_055536103.1
12	MBL fold metallo-hydrolase	WP_055536123.1
13	DUF4236 domain-containing protein	WP_030775935.1
14	hypothetical protein	WP_167532771.1
15	hypothetical protein	WP_055536104.1
16	AfsR/SARP family transcriptional regulator	WP_055536106.1
17	cytochrome P450	WP_055536124.1
18	HAD family hydrolase	WP_055536107.1
19	SigB/SigF/SigG family RNA polymerase sigma factor	WP_055536108.1
20	aminotransferase class I/II-fold pyridoxal phosphate-dependent enzyme	WP_055536109.1
21	hypothetical protein	WP_246201713.1
22	PHP domain-containing protein	WP_055536110.1
23	xanthine dehydrogenase family protein molybdopterin-binding subunit	WP_055536111.1
24	FAD binding domain-containing protein	WP_055536112.1
25	2Fe-2S iron-sulfur cluster-binding protein	WP_055536113.1
26	hypothetical protein	WP_055536114.1
27	alpha/beta hydrolase	WP_055536115.1
28	ATP-binding protein	WP_150477663.1
29	ribonuclease H family protein	WP_055536116.1
30	MarR family winged helix-turn-helix transcriptional regulator	WP_234335914.1
31	MFS transporter	WP_055536118.1
32	iron-containing redox enzyme family protein	WP_055536119.1
33	hypothetical protein CP975_33215	QEV21733.1
34	DUF5133 domain-containing protein	WP_055536120.1
35	catalase	WP_055536121.1

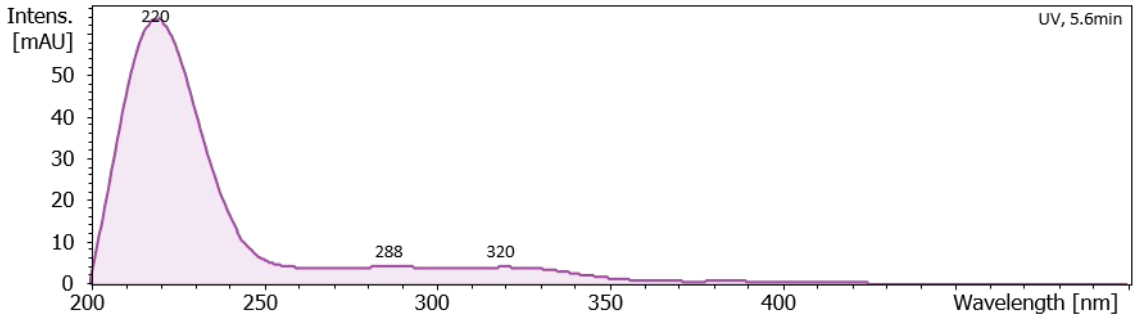
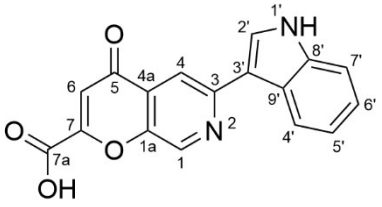


Figure S13: UV/VIS spectrum of mansevorone.

Table S8: NMR data (500 MHz, $MeOD_4$) of mansevorone.

no	δ(^{13}C, ^{15}N) [ppm], type	δ(^{1}H) [ppm], mult (J)	COSY (H-)	HMBC (H-)	NOESY (H-)
1	143.8, CH	9.14, d (0.5)	4 (^{4}J)	1a, 3, 4, 4a, 5	
1a	150.5, C	-	-	-	
2-N	319.0, N	-	-	-	
3	153.8, C	-	-	-	
4	112.8, CH	8.23, d (0.5)	1 (^{4}J)	1a, 2-N, 3, 3', 5	2'
4a	130.5, C	-	-	-	
5	181.3, C	-	-	-	
6	112.8, CH	7.04, s	-	4a, 7, 7a	
7	165.7, C	-	-	-	
7a	162.2, C	-	-	-	
1'-NH	133.6, NH	-	-	-	
2'	127.1, CH	7.97, s	-	1'-NH, 3, 3', 8', 9'	4
3'	116.5, C	-	-	-	
4'	121.7, CH	8.30, m	5', 7' (^{4}J)	5', 8', 9'	
5'	121.7, CH	7.19, m	4', 6'	6', 7', 9'	
6'	123.4, CH	7.20, m	5', 7'	4', 8'	
7'	112.9, CH	7.46, m	6', 4' (^{4}J)	5', 9'	
8'	139.0, C	-	-	-	
9'	126.6, C	-	-	-	



mansevorone (**5**)

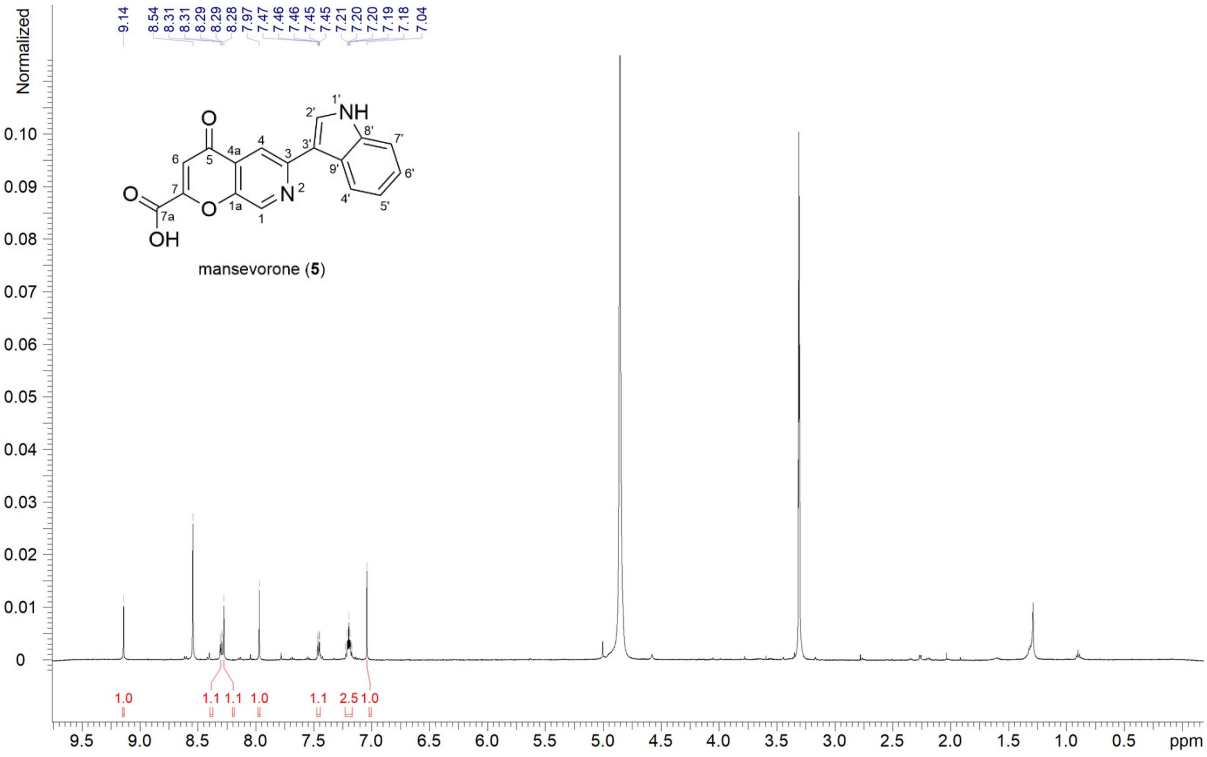


Figure S14: ^{1}H NMR spectrum (500 MHz, $MeOD_4$) of mansevorone.

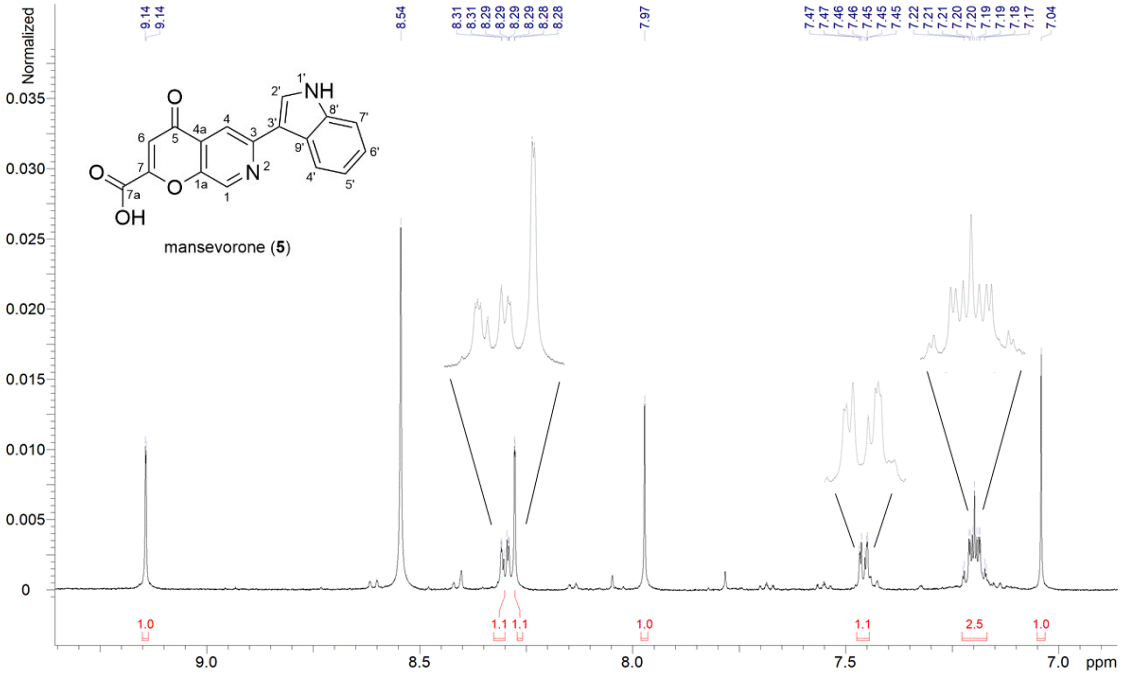


Figure S15: ^{1}H NMR spectrum (500 MHz, $MeOD_4$) of mansevorone zoomed in on the relevant peaks.

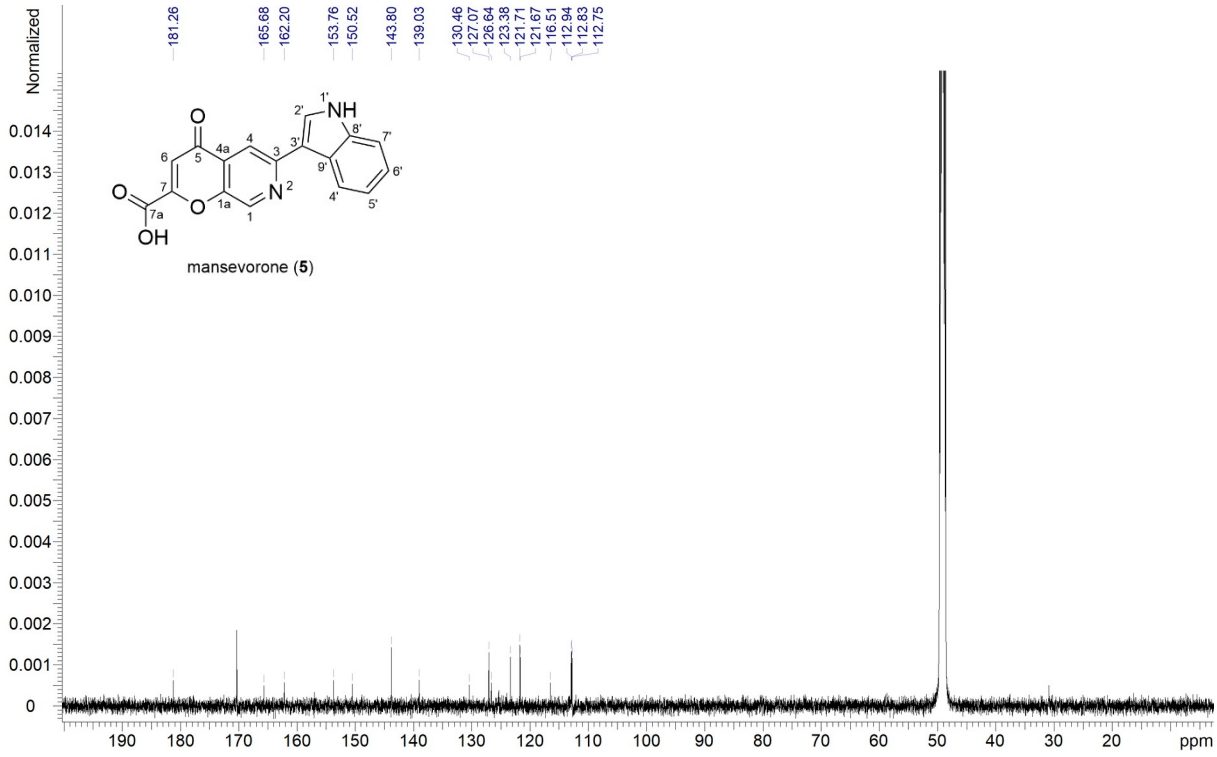


Figure S16: ^{13}C NMR spectrum (125MHz, $MeOD_4$) of mansevorone

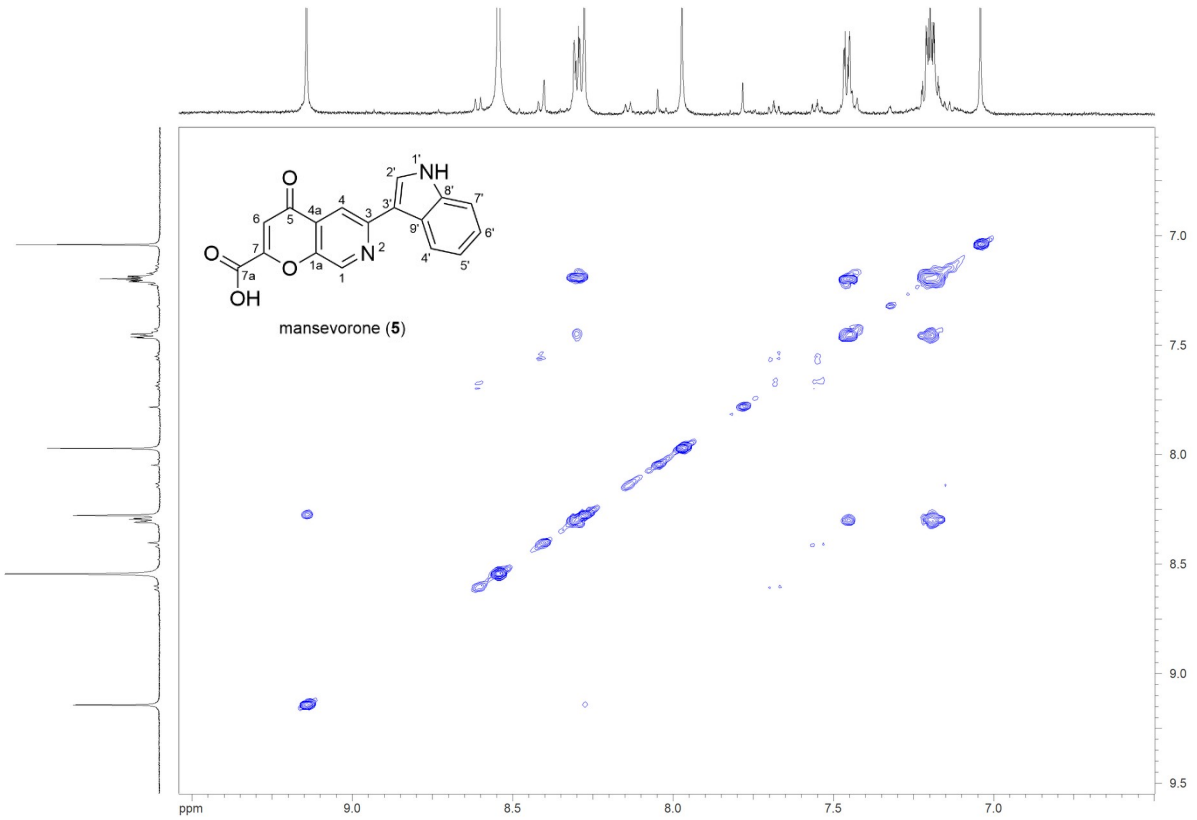


Figure S17: COSY spectrum ($MeOD_4$) of mansevorone.

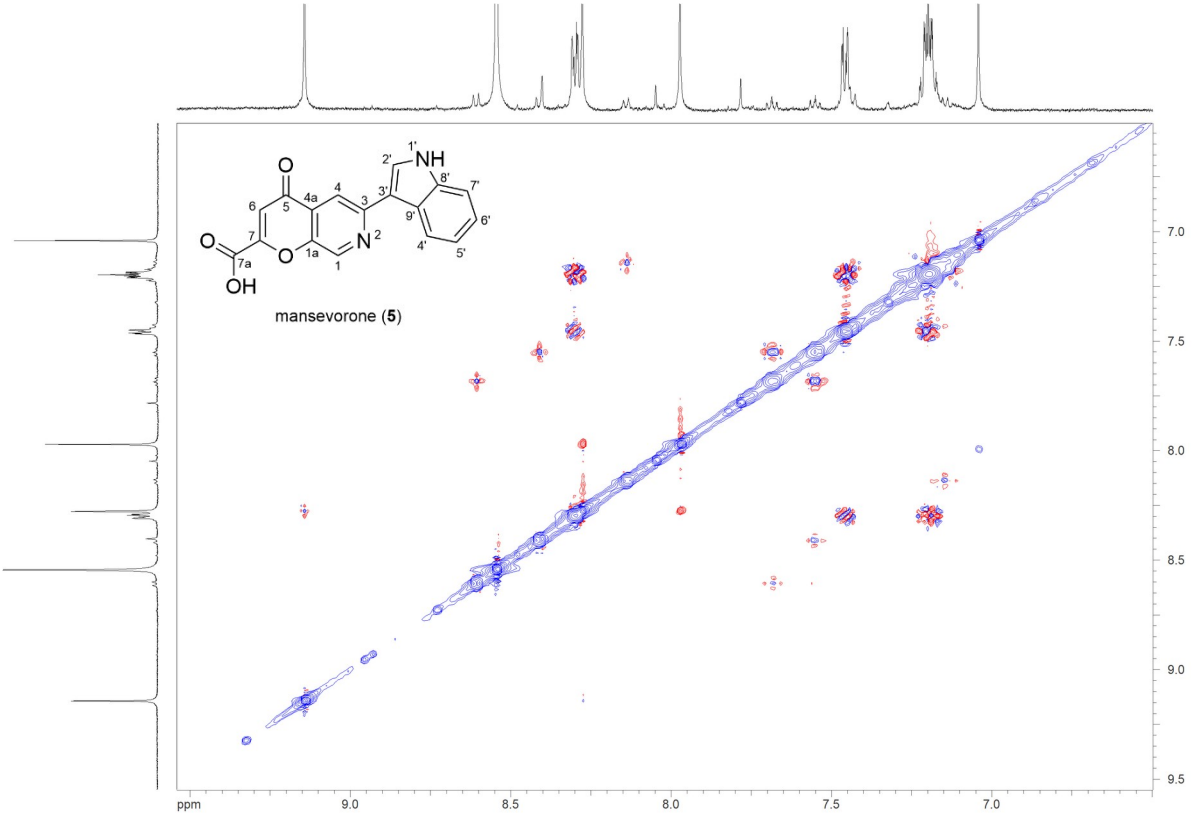


Figure S18: NOESY spectrum ($MeOD_4$) of mansevorone.

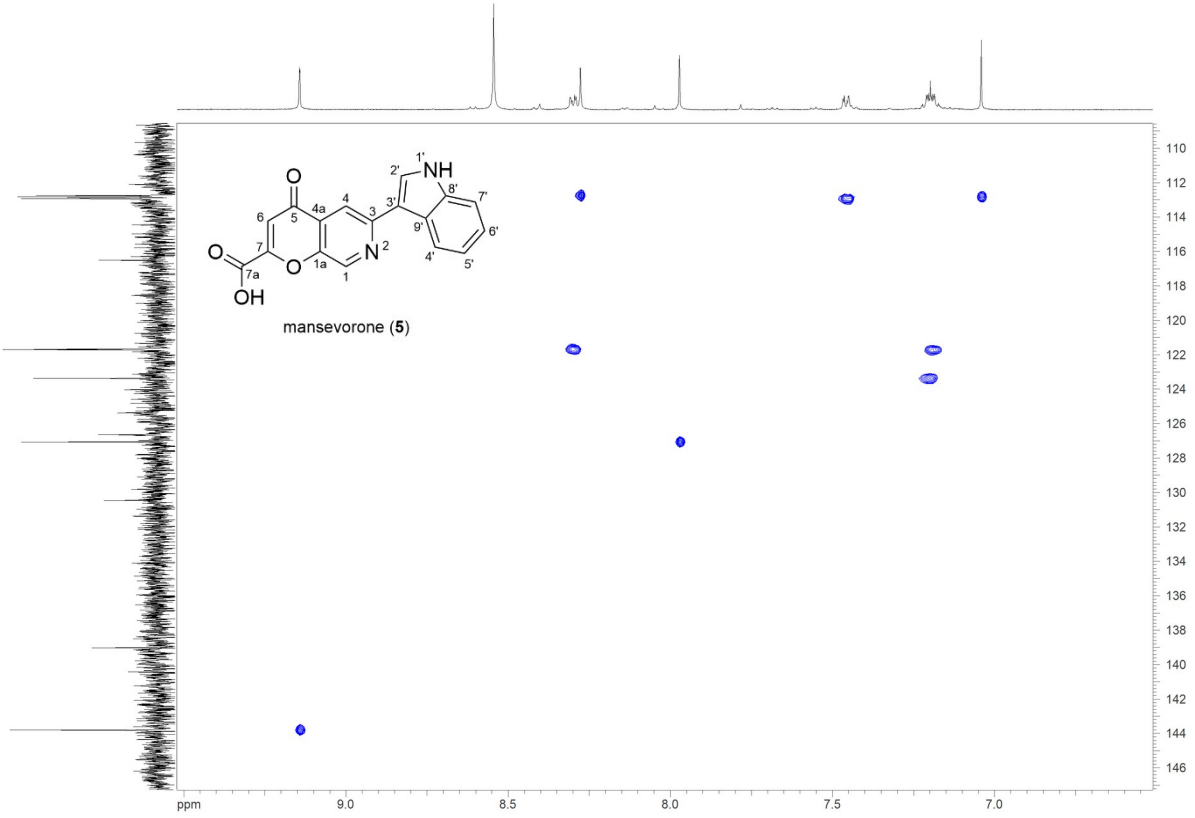


Figure S19: Edited HSQC spectrum ($MeOD_4$) of mansevorone.

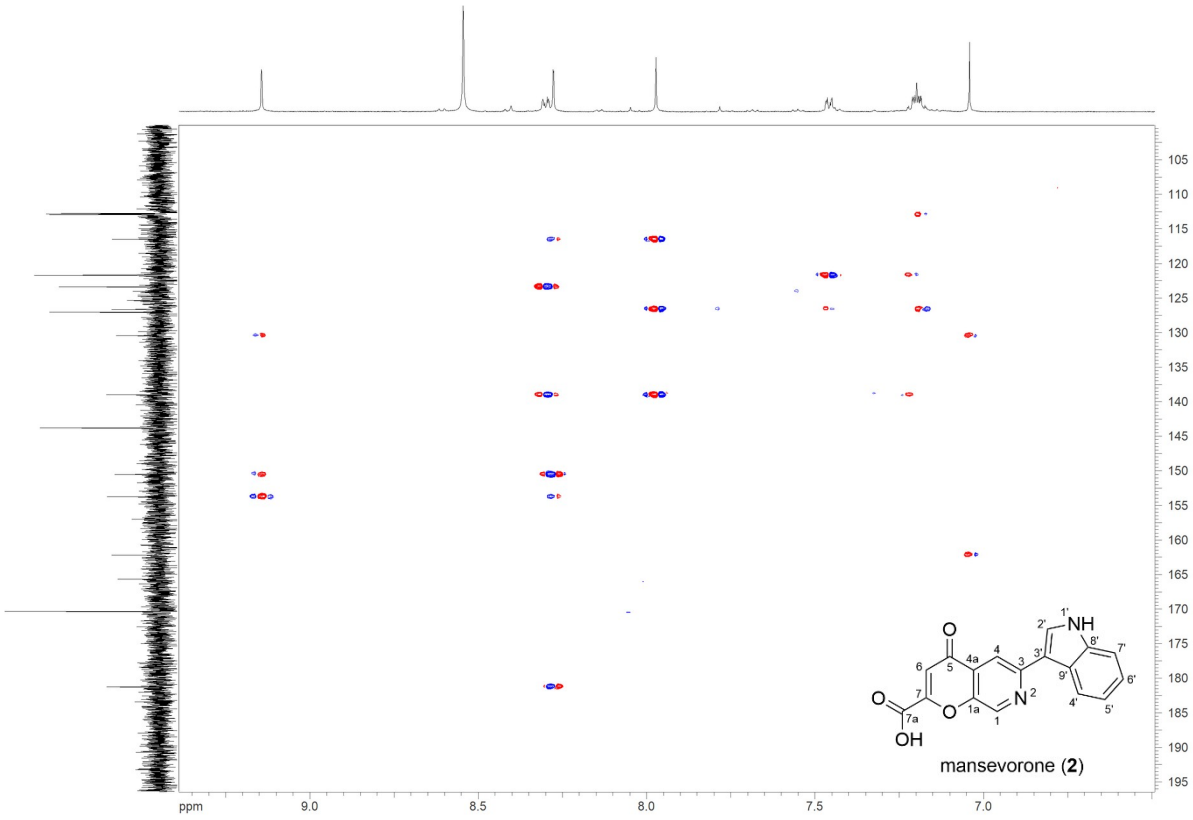


Figure S20: HMBC spectrum ($MeOD_4$) of mansevorone

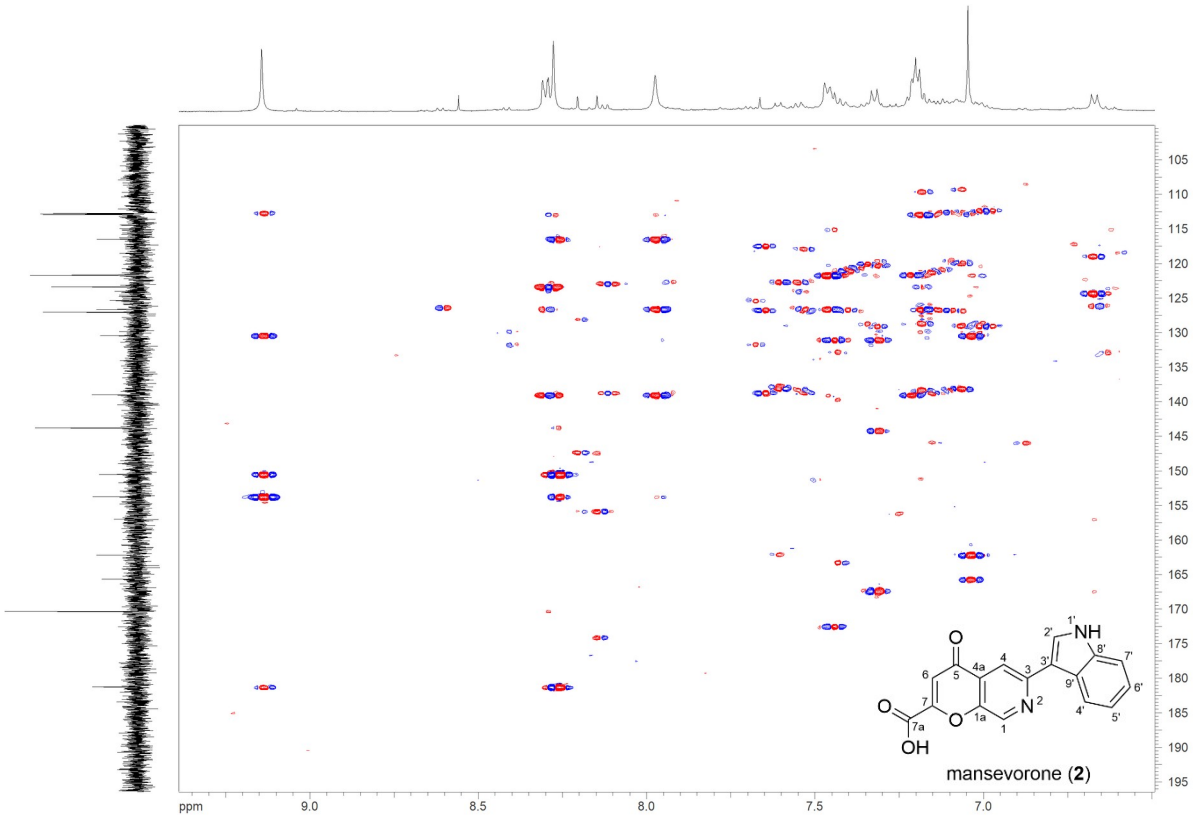


Figure S21: HMBC spectrum ($MeOD_4$) of mansevorone before final HPLC purification.

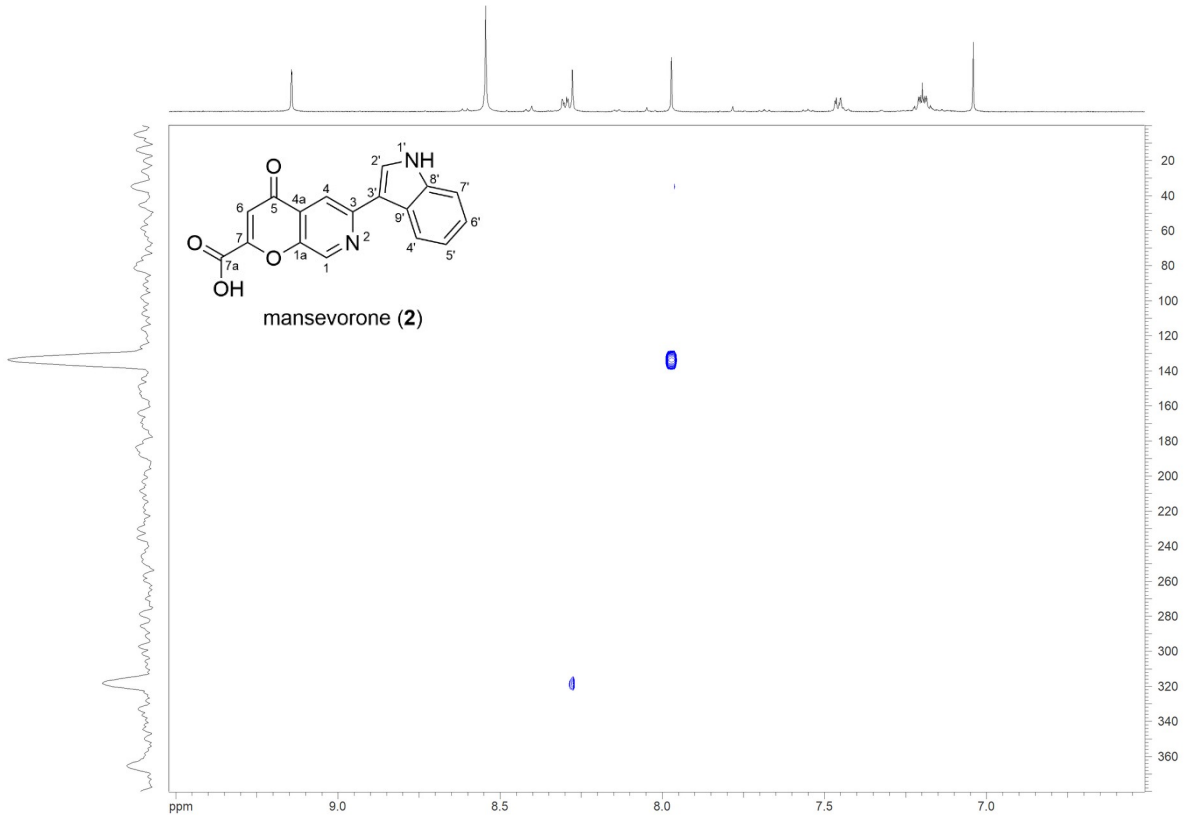


Figure S22: N-HMBC spectrum ($MeOD_4$) of mansevorone.

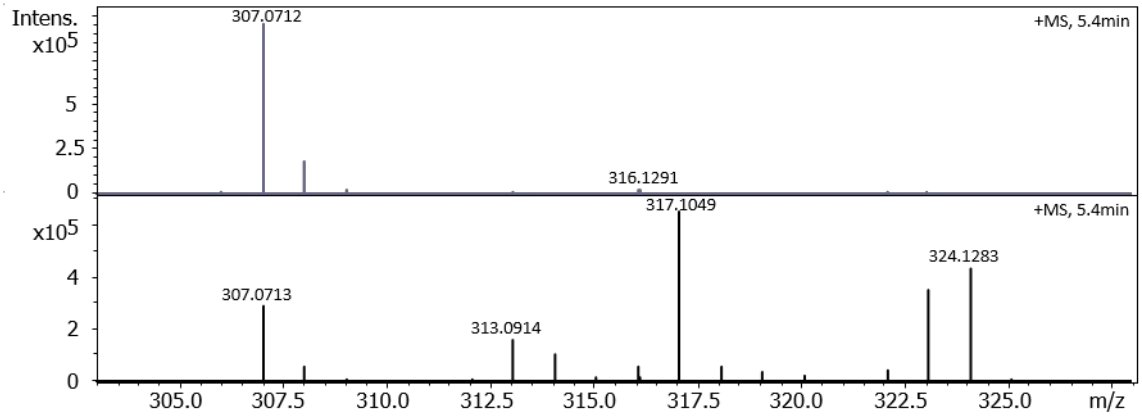


Figure S23: Comparison of MS spectra for mansevoron between (A) a culture of *S. albus* C16h7 and (B) a culture of *S. albus* C16h7 supplemented with $^{13}C_{11}$-labeled tryptophan.

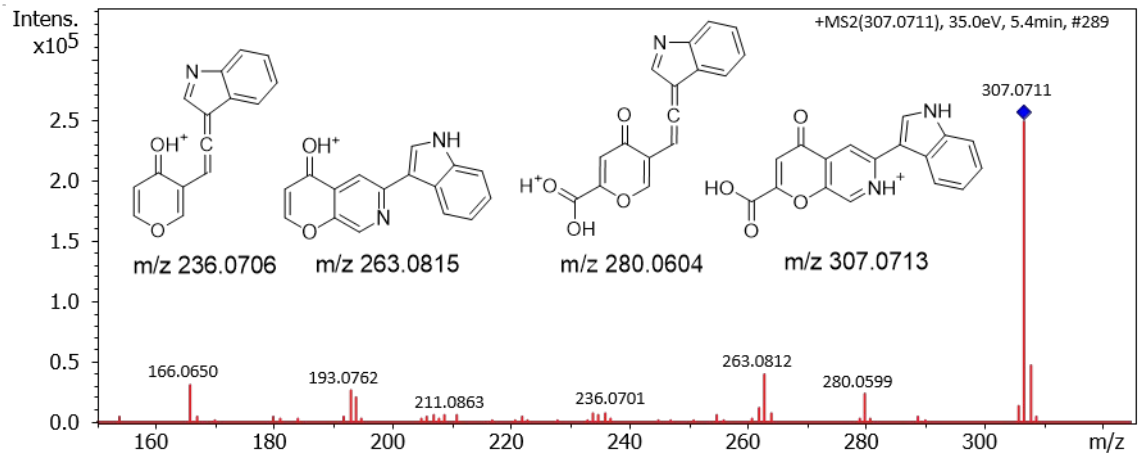


Figure S24: MS/MS fragmentation spectrum of mansevorone from cluster 16.

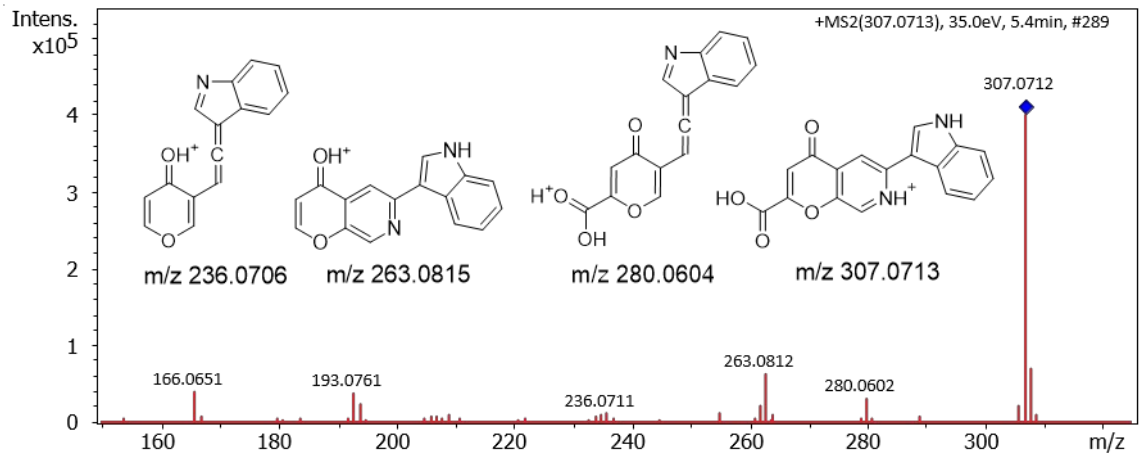


Figure S25: MS/MS fragmentation spectrum of mansevorone from cluster C1.8.

>AsfR sequence

MEFRLLGAVSVATEVGVLPLGPAKRRSLLAALLLRPNHPVPVDRLTAALWDQEPPARARGVIQGHVSQLRVLLTEAEAG
MFGVELVTQGTAYVLRMPESLLDAHRFEELVTLARGQRAPADAVAMYQEALSLWQGPALTGAYPSQPLQAAAQALEE
LRLASVESLAGAYTRMGEHARAAAVLRAEAGAHPLRESLSAELMRALQRAGRRTEALDWFHRTRRVLADELGVDPGRE
LADAYAAALRGAGDEADGAGGGASVADVPPSGGPVALAAPPASAPASLPAAPAAPGGSVVPAAPVASPASGPVLPPA
AIDLLPRMPRGFHGRAGELTALSRAAAGEAPVCLVTGPAGVGKTALVTHWAHRNREHFPGGLLYADLRGFSDTGEPAL
LEVLREFLLALGVAPRRIPESAGAASALFRSLSADRQLLVVLDNARASDQVRELLPGGARCVTVVTSRYRLRGLIASDAARP
VPVDVLEPEDSTALLAAVLGTDRVFAEEVAARRLAELCGGLPLALRVAAARLADQPESSLSTMCAELSDESRRLGLLDVED
TGVRAALRLTVRRLPADAARQFAHLGRHPGMYIDLYAAAALADTDPATAEAALDKLTAAHLVMRAGPDRWTLHDLVR
LYARDLDAEADALLRVLDLTVATALAAADAAEPGDESCFTLPADFRPPWAAREFADREQAMAWYAAERDDLTLAATA
ADSAGLHDRTWRIILGMWPQIVWRVQDGWAPLLKTALEATRADGDARAESRVLALLGWVLTEEGRIDEALGHLEAAP
LLAARAGDTRGEATALMNLSLAQAALGSPDEAAEGCALAAELAHSVGDRHTERLALCHLARHRLDAREWQAAHETAIA
ALDLVGPSEASAAARVLLLTAIGEALLGMGDETEGIRRLENAALEAEACGYDDGAVRALGALLRVSADAGLQARYDTAM
VRLMART*

Figure S26: Sequence of SARP regulator *AsfR16* from cluster C16.

>DnrI sequence

MHALRTPGPSPERTDPAGPLLRVLGPMSARFDGQDLPLGPPRRRALLALLLIRLGRVVPTELIVDELWGDEPPRQPVATL
QSHVSHLRRVLDPTAGPGAASVLRYRAPGYVLRLAPEQVDFCRFEDLVSTGRRSLEQRDPLAARDRLAEALELWHGSPY
TEFSAHPPLADETARLEQVRLAALESYAEARLTLGAAAEVAADLDREVRNHPMRERLVGHLMTALTRLGRQAEALEVYE
RTRSHLVEEFGVDTAVELQRVRTAILRQELGDGGPGGATAAPAVSAPSAPLTPPARPTRSPASAVPAPYGGTGRPARTP
VPAPAAAPARAEDPGTDVPDPSAPWPFTGRDQELHRLAAAAAGALTGHGHVVSVLGPAGVGKTRLLMELAPRLETAD
EHLEVVWSHCFPGEGVPPYWLWTQVLRRLSATRPDAFRAATAPFGALLGPLMPERSAGPGGAPGPEVDWAQARFLT
HDAVCEVLLALAAESPLVLLLEDLHWADTASLDLLRLLGVRRLGHPLSIVLTARDFEIESDATMRRLLAEVVRGPRSETLRL
DGLPRQAVATLVEAQAGPGVGAEVVEALHRRSKGNPYFVMQLLSLLGDVRHLHDPGASAVLLAQVPTGVREALRQRF
SALPEPVLRVLRLCAVIGTEVDTDLLHRTADGDEPVAEALESAIRAGLLGEDPHHPGRLHFAHALVQETLAEEPAREERHR
LHARVAEALCARGRGQMGDEEIERVAHHSWHAKDALAPGETLPRLLRAAERAEHHLAYEQVETWLRRAVHLAGFLPA
DDPSAPGLEQRLHIQLGQVLATIRGYGDAEAEAALGRGRALSAVTHSPEDPSVLWALCAALLVTGRYDDSRQFSGLLRD
LAGRTRQPVAVLGAAYGEGIVLHVRGRLPEALAELEHGVDMADRFAREGHSLARTFQHDPRVSCRSYDTFTHWLLGDR
RTAAERRDQLLRLTDYESRPSDRAFALYVDAVVAAWEGDVGTALASGGEGARVAGEHGLLYWKAMLSLPAGWALTHS
GREEEGLALMRTSLDELCPSRTHLRLPLHLGLLGQAQFHAGRREEATDTLRRMLSVVERRREYVYLDPALPATRLLHELV
GREAAEAVLRGC*

Figure S27: Sequence of SARP regulator *DnrI* from cluster C1.8.

Score	Expect	Method	Identities	Positives	Gaps
144 bits(364)	4e-39	Compositional matrix adjust.	144/372(39%)	178/372(47%)	13/372(3%)

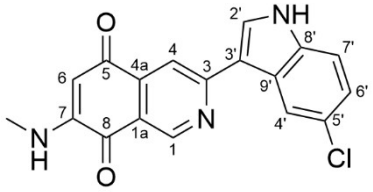
```
DnrI   21   LRVLGPMSARFDGQDLPLGPPRRRALLALLLIRLGRVVPTELIVDELWGDEPPRQPVATL  80
            R+LG +S   +   LPLGP +RR+LLA LL+R    VP + +   LW  EPP +    +
AsfR   3    FRLLGAVSVATEVGVLPLGPAKRRSLLAALLLRPNHPVPVDRLTAALWDQEPPARARGVI  62

DnrI   81   QSHVSHLRRVLDPTAGPGAASVLRYRAPGYVLRLAPEQVDFCRFEDLVSTGRRSLEQRDP  140
            Q HVS LR +L           L  +   YVLR+    +D  RFE+LV+  R    QR P
AsfR   63   QGHVSQLRVLLTEAEAGMFGVELVTQGTAYVLRMPESLLDAHRFEELVTLAR---GQRAP  119

DnrI   141  LAARDRLAEALELWHGSPYTEFSAHPPLADETARLEQVRLAALESYAEARLTLGAAAEVA  200
              A     EAL LW G   T      PL      LE++RLA++ES A A   +G  A  A
AsfR   120  ADAVAMYQEALSLWQGPALTGAYPSQPLQAAAQALEELRLASVESLAGAYTRMGEHARAA  179

DnrI   201  ADLDREVRNHPMRERLVGHLMTALTRLGRQAEALEVYERTRSHLVEEFGVDTAVELQRVR  260
            A L  E   HP+RE L   LM AL R GR+ EAL+ + RTR  L +E GVD   EL
AsfR   180  AVLRAEAGAHPLRESLSAELMRALQRAGRRTEALDWFHRTRRVLADELGVDPGRELADAY  239

DnrI   261  TAILRQELGDGGPGGATAAPAVSAPSA---PLTPPARPTRSPASAVPAPYGGTGRPARTP  317
             A LR    +    G  A+ A   PS     L  P     +   A PA  GG+  PA  P
AsfR   240  AAALRGAGDEADGAGGGASVADVPPSGGPVALAAPPASAPASLPAAPAAPGGSVVPA-AP  298

DnrI   318  VPAPAAAPARAEDPGTDVPDPSAPWPFTGRDQELHRLAAAAAGALTGHGHVVSVLGPAGV  377
            V +PA+ P     P      P  P  F GR  EL  L+ AAA    G   V  V GPAGV
AsfR   299  VASPASGPVL--PPAAIDLLPRMPRGFHGRAGELTALSRAAA----GEAPVCLVTGPAGV  352

DnrI   378  GKTRLLMELAPR  389
            GKT L+   A R
AsfR   353  GKTALVTHWAHR  364
```

Figure S28: Alignment of SARP regulators *DnrI* from cluster 1.8 and *AsfR* from cluster 16.

Table S9: NMR data of 5'-Cl-Mansouramycin D (500MHz, $MeOD_4$) and lit. data of mansouramycin D.

	5'-Cl-Mansouramycin D ($MeOD_4$)				Mansouramycin D (DMSO-*d*6)**	
No	**δ(^{13}C) [ppm], type**	**δ(^{13}C) [ppm], prediction***	**δ(^{1}H) [ppm], mult(J)**	**COSY (H-)**	**δ(^{13}C) [ppm]**	**δ(^{1}H) [ppm], mult(J)**
1	147.9, CH	149.0	9.20, s		147.6	9.11, d (0.7)
3						
4	113.6, CH	114.6	8.22, s		113.0	8.18, d (0.7)
5						
6	98.9, CH	99.6	5.79, s	N-Me	99.4	5.71, d (0.4)
7						
7-NH						
N-Me	27.8, CH	29.4	2.95, s	6	28.6	2.82 d (5.0)
8						
9						
10						
1'-NH						
2'	129.6, CH	131.6	8.25, s		129.8	8.53, br s
3'						
4'	121.1, CH	119.07	8.55, s	6'	121.7	8.52, dd (1.9, 6.0)
5'					122.2	7.22, m
6'	122.7, CH	123.85	7.22, dd (2.0, 8.6)	7', 4'	120.8	7.20, m
7'	112.7, CH	113.13	7.45, d (8.5)	6'	111.9	7.48, m
8'						
9'						

* predicted by ACD Labs ** ACS Chem. Biol. 2022, 17, 3, 598–608



5'-Cl mansouramycin D

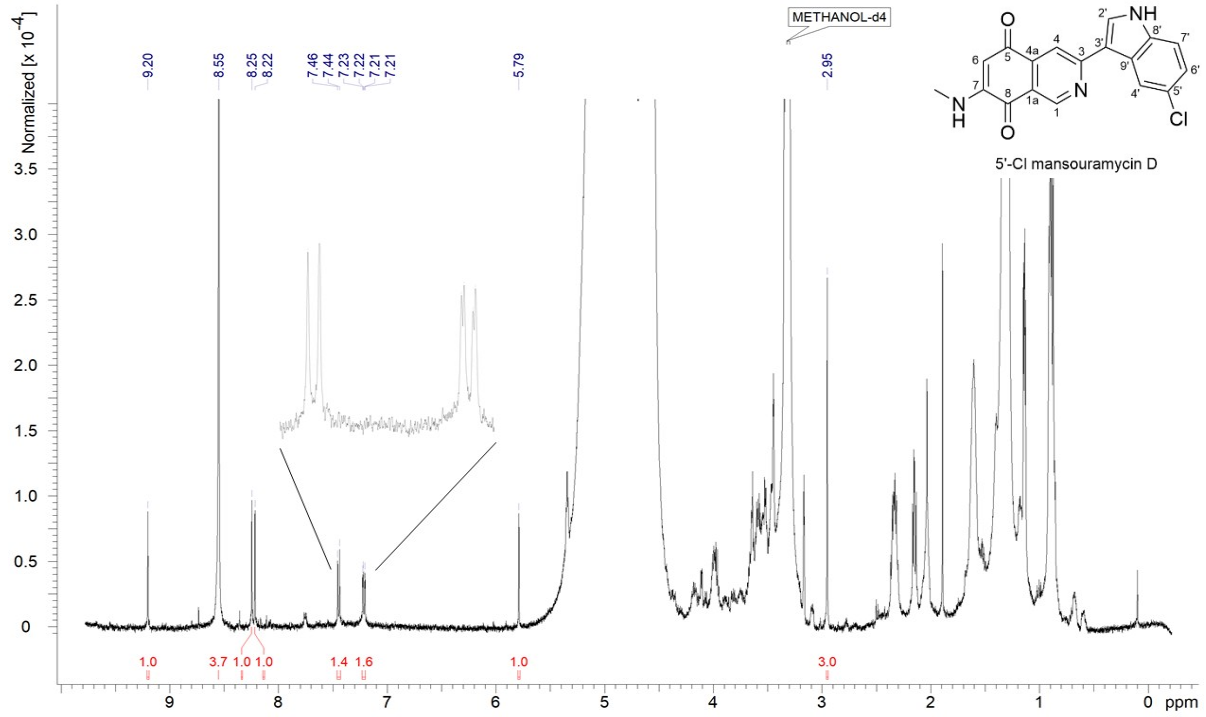


Figure S29: ^{1}H NMR spectrum (500 MHz, $MeOD_4$) of 5′-Cl-mansouramycin D.

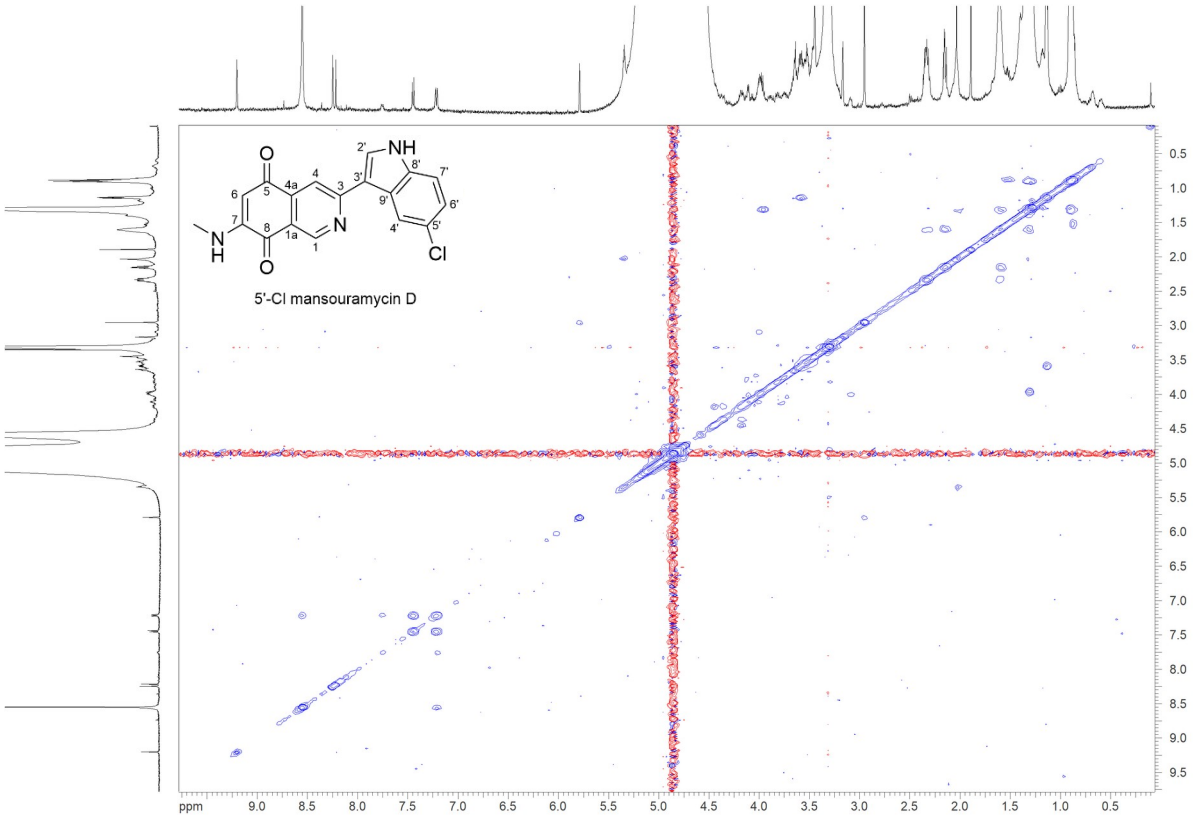


Figure S30: COSY spectrum ($MeOD_4$) of 5′-Cl-mansouramycin D.

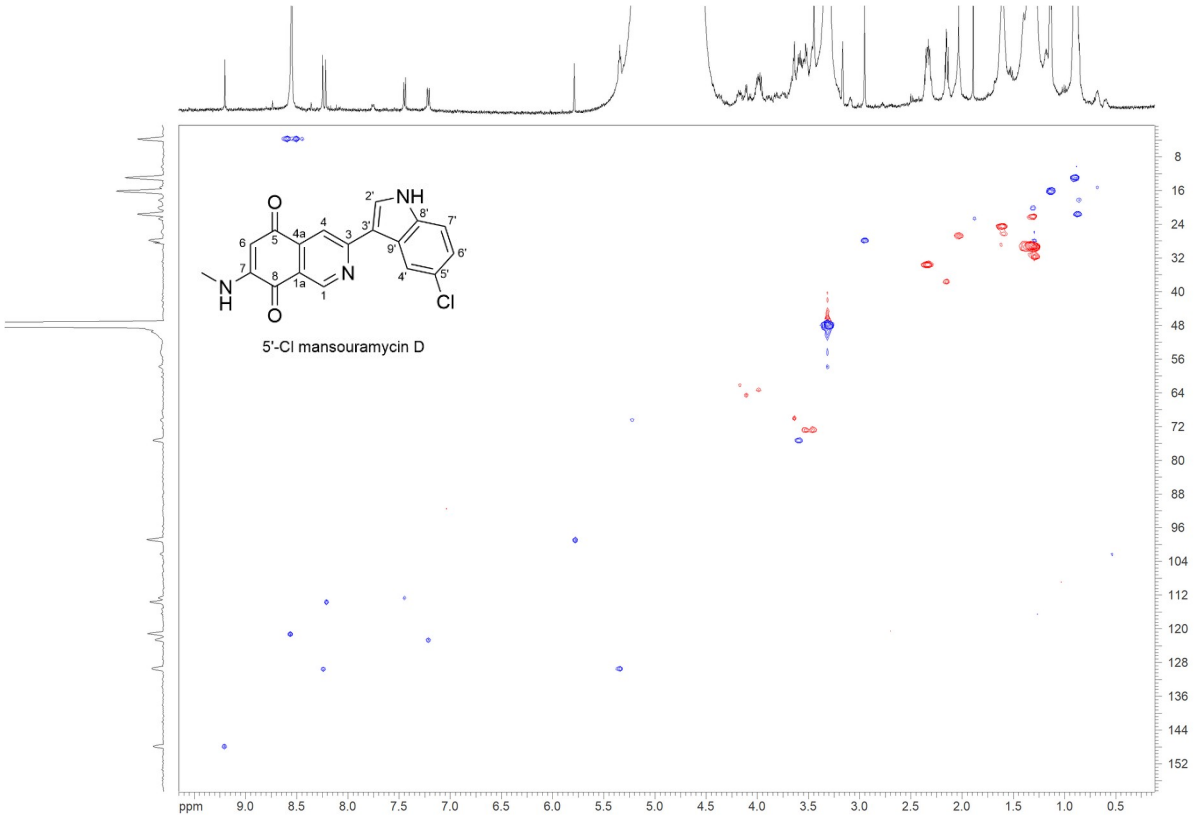


Figure S31: Edited HSQC spectrum ($MeOD_4$) of 5′-Cl-mansouramycin D.

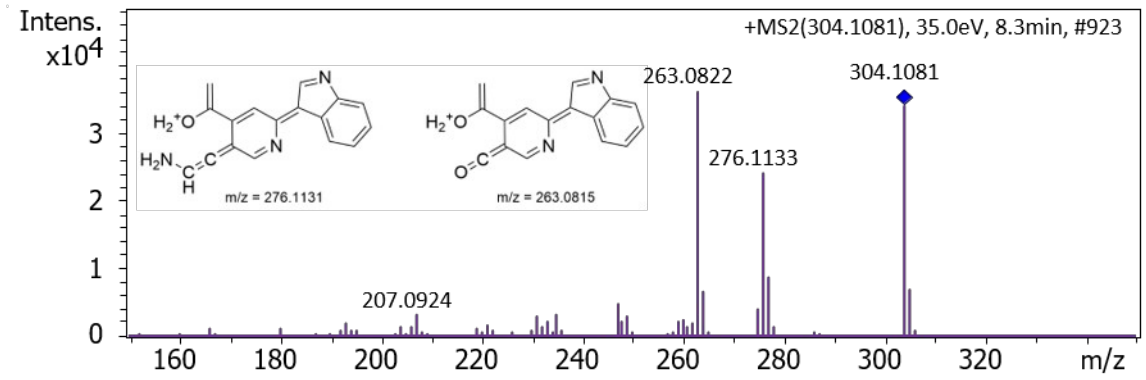


Figure S32: MS/MS fragmentation spectrum of mansouramycin D.

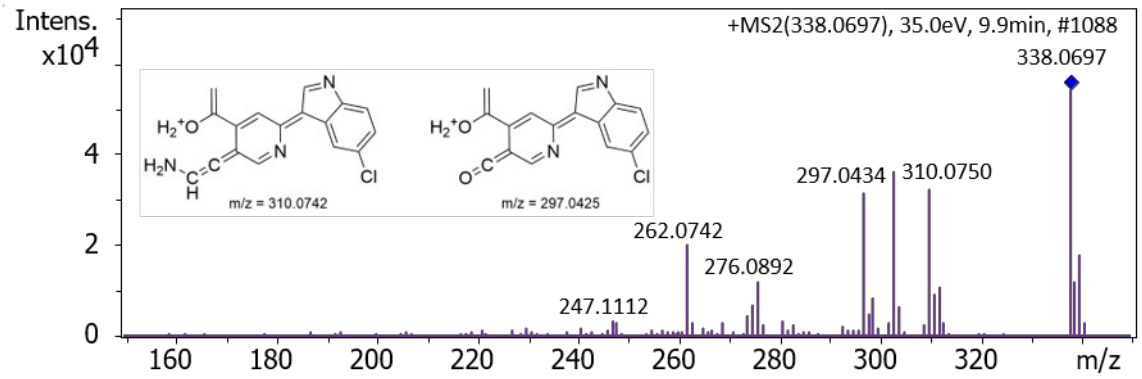


Figure S33: MS/MS fragmentation spectrum of 5'-Cl-mansouramycin D.

>RebH sequence

GTGATTAATTCAGTGGTCATTGTCGGCGGCGGCACTGCCGGATGGATGAGCGCCTCCTACCTCAAGGCGGCCTAC
GGTGACCGCATCAACGTCACTCTGGTGGAATCCGACCGCGTCGCGACCATCGGCGTCGGTGAAGCGACCTTCAGC
ACGGTCCGGCATTTCTTCGACTATCTCGGCCTGGACGAGCGCGAATGGATGCCGGAGTGCTCCGGTTCGTACAAGC
TGGGCATCCGGTTCGAGAACTGGCGGGAGCCGGGGCAGCACTTCTACCACCCCTTCGAGCGGCTGCGGACCTCCG
ACGGGTACACCCTGGCCGACTGGTGGCTTCAGGAGGGCGACCGCAGCGAGCCGTTCGACCGTTCCTGCTTCATCA
CGCCGGCGCTGTGCGAGGCCAAGCGCTCGCCGCGGCTGCTCGACGGCTCCCTGTTCGCGGGCGGTCTCGACGGCT
CGCTGGGCCGCTCGACGCTGGAGGAGCAGCGGTCCCAGTTCCCGTACGCGTACCACTTCGACGCGGCCCTGCTGG
CGAAGTTCCTGACCAAGTACGGCACCGACCGCGGCGTCCGCCATGTCGTCGACGACGTGACCGAGGTGGGCCGCG
ACGAGCGCGGCTGGATCAGCCATGTCGCCACCCGTGAGCACGGTGACCTCACCGGTGACCTGTTCATCGACTGCAC
CGGCTTCAAGGGCATGCTGATCAACGAGACCCTGGAGGAGCCGTTCGAGTCCTTCCAGGACGTGCTGCCGAACAA
CCGCGCCGTCGCGCTGCGCGTCCCGCAGGACGACCAGGCCACCACCGGGATGAACCCGTACACCACGGCGACCGC
GATGGACGCCGGCTGGATCTGGAACATCCCGCTGTTCGGGCGGAACGGCAACGGCTATGTGTACTCCGACGAGTT
CTGCTCCCCGGAGGAGGCGGAGCGCACCCTGCGGAACCACGTCGCCCCCGGCCGGGACGACCTGGAGGCCAACC
ACATCCGGATGCGCATCGGGCGCAACCGCCGCTCGTGGGTCAACAACTGTGTGGCCATCGGCCTGTCCAGCGCCTT
CGTCGAGCCGCTGGAGTCCACCGGCATCTTCTTCATCCAGCACGGCATCGAGCAGCTGGTGAAGAACTTCCCGGAC
GAGCGCTGGGACCCGGCCCTGGCCGACGACTACAACAACCGGGTCGCCGAGGTCCTGGACGGCGTCAAGGAGTT
CCTGGTCCTGCACTACAAGGCGGCGCAGCGCGAGGACACCCCGTACTGGAAGGAGGCCAAGACCCGTGCCCTGCC
CGACGGGCTCGCCGAGCGCCTCGCCATCGGCACCTCGCACCTGCTCGACGAGCGCACCATCTACCAGCCGTACCAC
GGCTTCGAGCAGTACTCCTGGATCACGATGATGCTGGGCCTCGGCCATGAGCCGGAGCGGCCGCGCCCCTCCCTC
GCGCACATCGACCCGACGAACGCGCGCGCCGAGCTCGCGCGGCTCCGGGCGGACGCGGACGAGCTGGTCGCCGC
CCTGCCCAGCTGCTACGAGTACATCGCCTCGCTCAACAGCTGA

Figure S34: DNA sequence of the tryptophan halogenase enzyme *RebH*.

References

1. Ahmed, Y.; Rebets, Y.; Estévez, M. R.; Zapp, J.; Myronovskyi, M.; Luzhetskyy, A., Microbial cell factories 2020, 19, 1-16.
2. Myronovskyi, M.; Rosenkränzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A., Metabolic engineering 2018, 49, 316-324.
3. Shuai, H.; Myronovskyi, M.; Rosenkränzer, B.; Paulus, C.; Nadmid, S.; Stierhof, M.; Kolling, D.; Luzhetskyy, A., ACS Chemical Biology 2022, 17, 598-608.
4. Flett, F.; Mersinias, V.; Smith, C. P., FEMS microbiology letters 1997, 155, 223-229.
5. Grant, S. G.; Jessee, J.; Bloom, F. R.; Hanahan, D., Proceedings of the National Academy of Sciences 1990, 87, 4645-4649.
6. Zhang, Y.; Muyrers, J. P.; Testa, G.; Stewart, A. F., Nature biotechnology 2000, 18, 1314-1317.
7. Baker Brachmann, C.; Davies, A.; Cost, G. J.; Caputo, E.; Li, J.; Hieter, P.; Boeke, J. D., Yeast 1998, 14, 115-132.
8. Bilyk, O.; Sekurova, O. N.; Zotchev, S. B.; Luzhetskyy, A., PloS one 2016, 11, e0158682.
9. Oberhäuser, P.; Myronovskyi, M.; Stierhof, M.; Gromyko, O.; Luzhetskyy, A., Microbial Cell Factories 2025, 24, 1-11.
10. Gregory, M. A.; Till, R.; Smith, M. C., Journal of bacteriology 2003, 185, 5320-5323.

2.3 Atrevomycin, a non-ribosomally synthesized cyclopeptide from the non-categorized *Streptomyces* LV1-209GEK biosynthesized by two interfering NRPS gene cluster

Patrick Oberhäuser, Marc Stierhof, Liliya Horbal, Alina Zaiachkivska, Gregor Ullrich, Andriy Luzhetskyy

To be submitted

2.3.1 Abstract

Atrevomycins are newly identified cyclic depsipeptides produced by the uncharacterized Actinobacterium *Streptomyces* sp. LV1-209GEK. They are structurally related to atrovimycin whereas structural elucidation and genome analysis revealed that their biosynthesis involves two distinct but cooperative NRPS/PKS gene clusters. These clusters jointly assemble the peptide backbone and install a fatty acid side chain. Comparative analysis suggests that the two clusters may have evolved from an ancestral gene cluster related to glycopeptide production. This system exemplifies the modular and adaptive nature of bacterial BGCs evolution.

2.3.2 Introduction

Natural products (NP) have long been a valuable source for biologically active secondary metabolites with remarkable chemical diversity. One notable group are the depsipeptides, a class of NPs defined by the presence of both amide and ester bonds within their peptide backbone[1]. These compounds are mainly produced as macrocycles and often incorporate non-proteinogenic amino acids such as β-hydroxyphenylalanine, β-methylaspartate, and 4-hydroxyphenylglycine. The resulting chemical diversity contributes to their pharmacological applications that includes antibacterial properties, cytotoxic, immunosuppressive, and antitumor activities[2-4]. Skyllamycin A, for example, can block PDGF receptor interactions through a mechanism that differ from conventional tyrosine kinase inhibitors[3-4].

Depsipeptides are predominantly found in microbial sources, with actinobacterial species like *Streptomyces* standing out. Their biosynthesis is based on the modular design of non-ribosomal peptide synthetases (NRPSs). Specialized adenylation domains within these enzymes recruit unusual building blocks, while additional catalytic domains create ester linkages. Two main enzymatic strategies are known to support this process[5-7]. In one mechanism the condensation domain catalyzes the ester formation during chain elongation, either by accepting α-hydroxy acids directly or by reducing α-keto acids as seen in valinomycin and cereulide biosynthesis. The other mechanism facilitates thioesterase domains for cyclization through intramolecular transesterification which often involves amino acids with hydroxyl-containing side chains[6-8].

Nearly all characterized nonribosomal depsipeptides synthetases produce cyclic compounds with rare exceptions like asperphenamate[9-10]. Their cyclic nature offers them advantages including lower entropic costs during target binding, preorganization for ion coordination and enhanced resistance to proteolytic degradation[6, 11-13]. Further structural and biological activity

diversification is achieved through the action of tailoring enzymes like cytochrome P450 monooxygenases, methyltransferases or through hybrid NRPS/PKS systems able to install fatty acid side chains[14].

In this study, we report the identification and characterization of a new member of the CDP class of natural products, named atrevomycins. We identified these compounds in the extracts of the uncharacterized strain *Streptomyces* LV1-209GEK. Comprehensive structural elucidation using high-resolution mass spectrometry (HRMS) and detailed ^{1}H, ^{13}C, and ^{15}N NMR spectroscopy revealed that atrevomycin shares structural similarities with the known CDPs atrovimycin which belongs to a bigger class of CDPs containing cinnamoyl residues like skyllamycin and cinnapaptin. Based on genomic analysis and biosynthetic logic, the production of atrevomycin is driven by two distinct NRPS gene clusters, highlighting the potential of cluster evolution through transfer of genetic material between *Streptomyces* species.

2.3.3 Results and Discussion

Heterologous production of atrevomycin.

We routinely screen Streptomycetes strains derived from diverse environmental samples as part of our efforts to identify novel secondary metabolites. One of the strains from our soil-derived collection isolated in eastern Europe was the uncharacterized *Streptomyces* LV1-209GEK. This strain was cultivated in SG production medium and metabolites extracted with butanol. Subsequent LC-HRMS analysis revealed two major peaks with $[M+H]^+$ 1289.5965 m/z (atrevomycin A) and $[M+H]^+$ 1191.5598 m/z m/z (atrevomycin B) eluting at t_R = 5.2 min and t_R = 5.8 min, respectively, both showing UV absorption maxima at λ_{max} 226, 266 and 316 nm. Dereplication did not result in any matches, suggesting these compounds may represent new natural products (Fig. 1). A scale-up fermentation of *Streptomyces* LV1-209GEK was performed in 10 L SG medium under the same cultivation conditions in order to isolate these potentially novel compounds. The crude extract was subjected to a three-step purification process resulting in the purification 29.3 mg of atrevomycin A and 6.8 mg atrevomycin B.

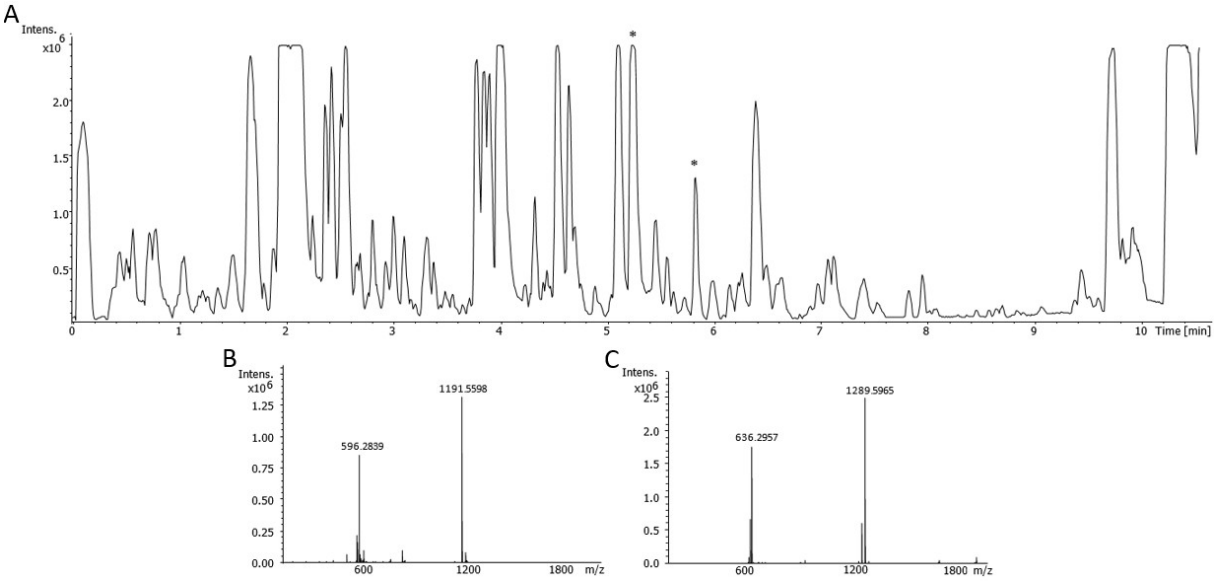


Figure 5: LC-MS-detection of atrevomycin A and B. A – Base peak chromatogram of crude extract from *S.* sp. LV1-209GEK. Peaks corresponding to atrevomycin are marked with an asterisk. B and C – Mass spectra of the peaks corresponding to atrevomycin A and B, respectively.

Structure elucidation of atrevomycin A and B

The molecular formula of atrevomycin A was determined to be $C_{62}H_{85}N_{10}O_{20}$, with a calculated $[M+H]^+$ = 1289.5936 m/z. For NMR measurements, the compound was dissolved in MeOH-d_3, which provided the best resolution by preserving NH protonation. Analysis of 1D and 2D NMR data (Fig. S2-S10) revealed spectra with characteristic chemical shifts of a peptide. Subsequent assignment of the amino acid residues using COSY, HSQC, ^{15}N-HSQC, ^{15}N-HMBC, and ^{13}C-HMBC correlations identified 2×Leu, 3×Ser, 2×Thr, Gly, hydroxyphenylglycine (HPG), and β-phenylserine. Additionally, a rare cinnamic acyl chain was identified, attached to the amino group of the N-terminal Thr-1. The peptide backbone was assigned using ^{13}C-HMBC and ROESY correlations (Table S11), revealing a cyclic peptide structure with an ester bond between the β-OH of Thr-1 and the carboxyl group of the C-terminal Leu-2 (Fig. 2, R1).

The structure of the minor compound, atrevomycin B, was determined based on the molecular formula $C_{57}H_{79}N_{10}O_{18}$, with a calculated $[M+H]^+$ = 1191.5568 m/z. Based on 2D NMR data (Figure S12-S17), the same cyclic peptide structure was assigned as previously determined for atrevomycin A. The difference lies in the N-terminal modification, which features 3-phenylpropanoic acid instead of the rare cinnamic acyl chain (Fig. 2, R2).

The absolute stereochemistry of the amino acid residues was determined using the advanced Marfey's method, revealing 2×L-Ser, D-Ser, L-allo-Thr, D-allo-Thr, 2×D-Leu, and L-threo-β-

phenylserine (Figure S18). The correct order of amino acids and the configuration of HPG were inferred from the presence of epimerase domains in the NRPS modules of the corresponding BGC of atrevomycin (see next section), which indicate L-to-D conversion of specific amino acids.

A structure similarity search revealed atrovimycin A, which differs only by a D-Val residue at position 9, while the stereochemical assignments of the other amino acids are identical. Due to the high similarity between the two compounds, the configuration of the cinnamic acyl chain was not determined but assumed to be equivalent.

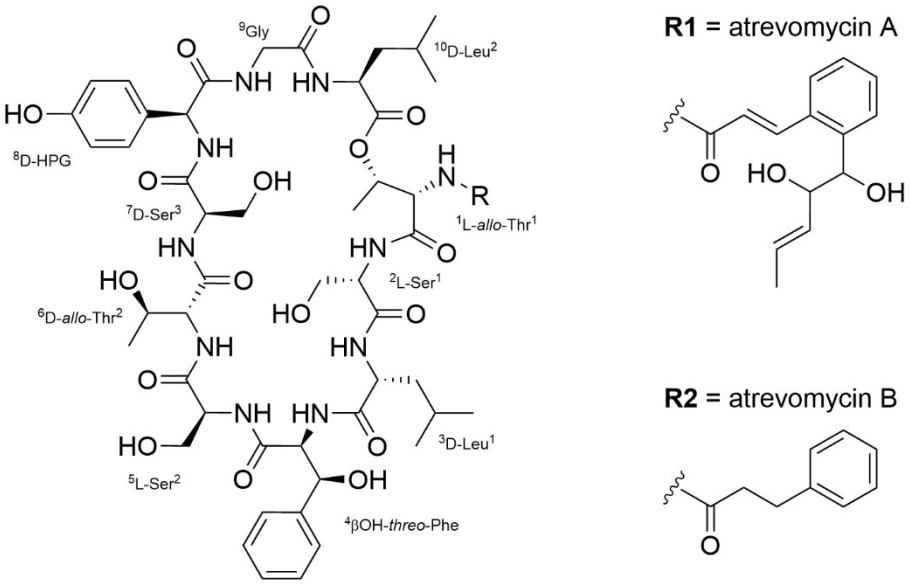


Figure 2: Structure of atrevomycin A and B.

The whole genome of *Streptomyces* LV1-209GEK was sequenced and subsequently analyzed using the antiSMASH platform[15]. A total of 32 putative biosynthetic gene clusters associated with secondary metabolite production were identified. Notably, none of these BGCs showed significant similarity to the characterized atrovimycin A BGC despite the close structural relationship of the atrevomycins to atrovimycin. However, two BGCs were identified that share moderate similarity with the known depsipeptide BGCs of cinnapeptin (21%)[16] and skyllamycin (32%)[3], suggesting that these clusters may be responsible for atrevomycin biosynthesis. While most natural products are encoded by a single biosynthetic gene cluster, an increasing number of examples reveal that multiple BGCs can act synergistically to produce a single compound. This phenomenon, though often overlooked, has significant implications for genome mining and heterologous expression strategies, which may fail if an additional, functionally required BGC is absent. For instance, the biosynthesis of actinomycin L in *Streptomyces* sp. MBT27 involves a canonical actinomycin BGC highly similar to that of *S.* antibioticus. However, the

characteristic anthranilamide moiety of actinomycin L is not encoded within the primary cluster and the enzyme responsible for converting anthranilic acid to anthranilamide remains unidentified[17-18]. Even more intricate is the biosynthesis of the catecholate-hydroxamate siderophores qinichelin and griseobactin in *Streptomyces* sp. MBT76, which involves coordinated activity of at least four distinct BGCs. These clusters cooperate to expand structural diversity from a limited genetic repertoire, a strategy reminiscent of the functional crosstalk observed in the assembly of erythrochelin in *Saccharopolyspora erythraea* and rhodochelin in *Rhodococcus jostii* RHA1[18].

Bioinformatic analysis of the skyllamycin-like cluster (designated cluster 1) revealed a total of 68 open reading frames (ORFs), with putative functions assigned through BLASTp searches (see Supplementary Table S3). Several of these genes, including *avo18* (NRPS), *avo21* (ACP S-malonyltransferase), and *avo23–avo24* (acyl-CoA carboxylase subunits), show high sequence similarity (up to 91%) to the corresponding genes *avm3*, *avm6*, *avm10*, and *avm11* from the atrovimycin BGC. Four β-ketosynthase-encoding genes (*avo38*, *avo41-avo43*) also display significant homology (65–85%) to *avm17* and *avm20–avm22* and are proposed to be involved in the biosynthesis of the cinnamic acyl side chain. Furthermore, three NRPS genes (*avo18*, *avo49*, *avo50*) likely encode the core peptide assembly line, including a condensation starter (C_{start}) domain that could enable acylation with the cinnamoyl moiety. In contrast to the atrovimycin BGC, cluster 1 contains several additional genes without clear homologs in the *avm* cluster, including numerous predicted regulatory proteins (*avo5*, *avo15*, *avo19*, *avo20*, *avo22*, *avo25*, *avo27*, *avo29*, *avo56*, *avo59*, *avo64*). These transcriptional regulators may play a role in orchestrating gene expression across both depsipeptide-associated BGCs and could be involved in the regulatory crosstalk required for coordinated biosynthesis of the atrevomycins. Additional genes encoding putative transporters and enzymes of unknown function may further contribute to the structural diversification or export of the final products

The second depsipeptide-associated cluster, showing similarity to the cinnapeptin BGC and designated as cluster 2, comprises 37 ORFs (see Supplementary Table S4). Among them encode the two genes *avo16* and *avo17* NRPSs that show high similarity to *avm30* (85%) and *avm31* (80%) from the atrovimycin BGC, suggesting that this cluster provides the remaining NRPS modules involved in atrevomycin biosynthesis. In addition, *avo14*, encoding a cytochrome P450 enzyme, corresponds to *avm28* with 92% similarity.

Unlike the atrovimycin cluster, cluster 2 contains a number of genes without clear counterparts in the *avm* BGC, including several transcriptional regulators (*avo1*, *avo2*, *avo5*, *avo7*, *avo27*,

avo28, *avo31*, *avo37*) and transport-related proteins (*avo6*, *avo12*, *avo18*, *avo19*). Two genes in the *avo* cluster (*avo20* and *avo21*) encode predicted acyltransferases. In contrast, the *avm* cluster contains only one homolog. This difference may point to a gene expansion event. The duplicated enzymes could have taken on specialized roles in acyl group modification or activation. They likely interact functionally with the polyketide synthase machinery of cluster 1, potentially contributing to structural variation or enhanced metabolic efficiency. This supports the hypothesis that both clusters act cooperatively during atrevomycin biosynthesis and may be co-regulated to ensure coordinated expression.

Our results suggests that atrevomycin biosynthesis follows the logic of atrovimycin but is divided between two distinct gene clusters. The skyllamycin-like cluster likely initiates biosynthesis by assembling and incorporating the cinnamic acid-derived acyl side chain into the nascent peptide. The cinnapeptin-like cluster provides the remaining NRPS modules to complete peptide elongation and tailoring. Both clusters also contain regulatory and transporter genes, which likely coordinate expression and facilitate substrate exchange. This arrangement supports a cooperative biosynthetic model (Fig. 3).

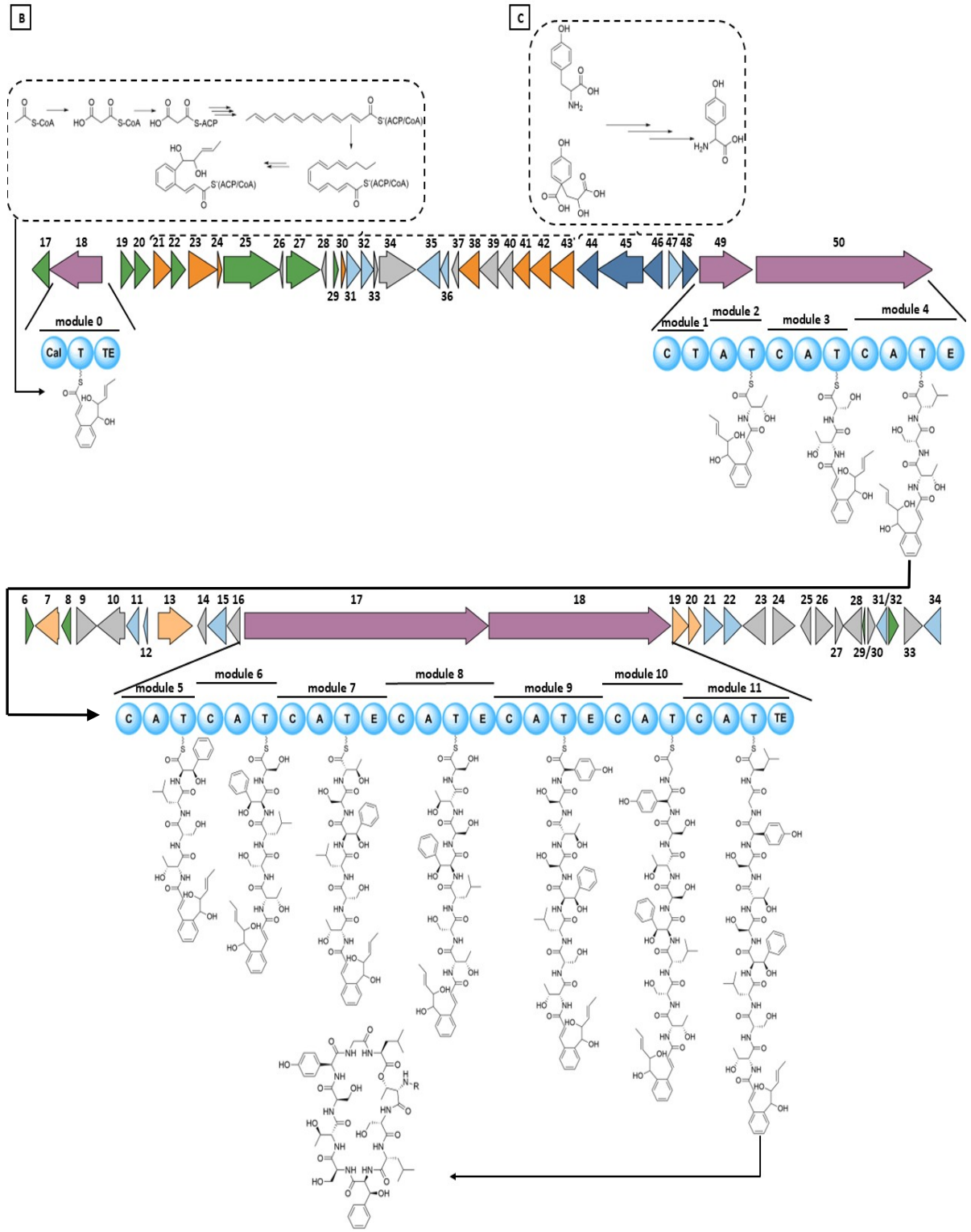


Figure 3: Biosynthetic pathway model of atrevomycin in *S.* sp. 1-209GEK. The functional cross-talk among two entirely separate gene clusters ensures the assembly of the peptide core and fatty acid side chain. Cluster 1 initiates the biosynthesis by producing the cinnmatic acid chain (**B**) and acylating it to threonine on the first NRPS module. This precursor gets elongated by the NRPS set encoded in cluster 2 until the final product releases the chain. The biosynthetic set of genes needed to produce HPG are encoded in cluster 1 (**C**).

Analyses of the adenylation domain organization and annotation by antiSMASH revealed a discrepancy in the predicted amino acid sequence at the C-terminus of the NRPS assembly line. Structure elucidation by NMR determined that the final three residues of the peptide are incorporated in the order X (HPG), glycine, and leucine, whereas bioinformatic predictions suggest a reversed order. Furthermore, the presence of an epimerization domain in module 8 suggests the conversion of L- to D-leucine, which aligns with the structural data but not with the predicted organization of the NRPS assembly line. One explanation for this discrepancy could be that the predicted function of the adenylation domain is incorrect. This would also explain why HPG is found in the D-configuration in atrovimycin A. It suggests that the epimerization domain in module 9 converts L-HPG to D-HPG, rather than converting leucine. The absence of an epimerization domain in module 11, which would be expected for the L- to D-leucine conversion, further supports this interpretation. Therefore, we hypothesize that the terminal thioesterase domain may play a compensatory role by catalyzing the epimerization of leucine from the L- to D-form during or after peptide release. A comparable mechanism has been reported for nocardicin A biosynthesis. There, the bifunctional TE domain *NocTE* catalyzes both L-to D epimerization and subsequent hydrolytic cleavage of the terminal HPG residue[19]. While no direct evidence currently confirms such activity in the atrevomycin biosynthetic pathway, this scenario is consistent with other reported cases where TE domains perform non-canonical functions in the absence of epimerization domains. For example, in the biosynthesis of chitinimides catalyzes the TE domain of *ChmC* a transesterification on peptide substrates[20]. Similarly does the TE domain in the production of the depsipeptides FR900359 mediate an intermolecular transesterification reaction essential for side chain installation[21].

The atrevomycin BGC, though functionally equivalent to the atrovimycin BGC, is uniquely split across distinct genomic loci in *Streptomyces* LV1-209GEK highlighting the genomic plasticity of *Streptomyces* and other Actinobacteria. This plasticity is a key driver of BGC diversification and enables metabolic adaptability. Comparative genomic analysis have revealed that the evolution of BGC is strongly influenced by horizontal gene transfer (HGT), homologous recombination and the integration of mobile genetic elements such as phage DNA and plasmids[22-23]. Especially HGT facilitates the acquisition and rearrangement of biosynthetic modules across species. This event often results in non-canonical or split gene clusters that retain or adapt their original biosynthetic function. In the case of atrevomycin, the identification of two discrete but interacting gene clusters corresponding to the biosynthesis of its polyketide chain and peptide core supports the evolutionary process of BGCs. This phenomenon likely reflects both the ecological versatility and the high frequency of recombination events in

Streptomyces driven by selective pressure such as interspecies competition. Split or fragmented BGCs as functionally unified systems in genome mining and biosynthetic pathway reconstruction needs to be considered.

2.3.4 Materials and Methods

General Experimental Procedure

Table S5 provides an overview of all strains and plasmids used in this study. *Streptomyces* strains were cultivated on soya flour mannitol agar (MS agar)[24] for sporulation and in liquid tryptic soy broth prior to metabolite production (TSB, Sigma-Aldrich, St. Louis, MO, USA). Main culture for secondary metabolite production was done in liquid SG medium (20 g/L glucose, 5 g/L yeast extract, 10 g/L soytone, and 2 g/L calcium carbonate).

Isolation of Chromosomal DNA from LV1-209GEK, Genome mining and Bioinformatic Analysis

Chromosomal DNA from *Streptomyces* LV1-209GEK was isolated using standard protocols[24-26]. Molecular biology reagents were used as per the manufacturer's protocol (NEB, Ipswich, UK; Thermo Fischer Scientific, Waltham, MA, USA). The LV1-209GEK genome was screened for secondary metabolite biosynthetic gene cluster using the antiSMASH online tool[15]. Genetic data was analyzed with the software Geneious prime 2022.2.2[28]. Selected genes were further analyzed using the BLAST online tool from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST/) and the Universal Protein Resource (UniProt) (https://www.uniprot.org/blast).

Metabolic Extraction and Analysis

Streptomyces LV1-209GEK was precultured in 25 mL TSB for 24 h before 1 mL of this seed culture was used to inoculate 100 mL of SG production medium. After the cultures were incubated for 7 days at 28 °C, the secreted metabolites were extracted from the supernatant with equal amounts of butanol, followed by solvent evaporation and dissolving in methanol. From the achieved methanol extract was 1 µL separated on a Dionex Ultimate 3000 UPLC (Thermo Fischer Scientific, Waltham, MA, USA), a 10-cm ACQUITY UPLC BEH C18 column, 1.7 µm

(Waters, Milford, MA, USA) and a linear gradient of 0.1 % formic acid solution in acetonitrile against 0.1 % formic acid solution in water from 5 % to 95 % in 18 min at a flow rate of 0.6 ml/min to enable HPLC-MS analysis. Samples were analyzed using an amaZon speed mass spectrometer or maXis high resolution LC-QTOF system (Bruker, USA). Data were collected and analyzed with the Bruker Compass Data Analysis software version 4.1 (Bruker, Billerica, MA, USA). Monoisotopic mass was searched in the natural products database DNP (Dictionary of Natural Products[27]).

Atrevomycin Isolation

Streptomyces LV1-209GEK was pre-cultured in 6 flasks each containing 25 mL of TSB for 24 h. Subsequently, 100 flasks each containing 100 mL of SG production medium were inoculated with 1 mL of the seed culture and incubated for 7 days at 28 °C. The mycelial part was separated by centrifugation and metabolites extracted from the supernatant as described above. The atrevomycins were purified on a Biotage Isolera One LC-system (Biotage, Uppsala, Sweden) using a Chromabond Flash RS330 C18 EC 360g (Macherey-Nagel, Düren, Germany) and water/methanol as the mobile phase at a flow rate of 100 mL/min. Different mixtures of A and B were used in a gradient (10-40 %, 40-80% and 80-100% A-B) and fractions containing atrevomycins were collected, pooled together, and used for preparative HPLC as the second separation step. Further purification steps were performed on a Waters autopurification system (Waters, Milford, MA, USA) equipped with a Waters 2545 Binary Gradient module using a Nucleodur C18 HTec column (5 µm, 250 x 21 mm, Macherey-Nagel, Düren, Germany) and a photodiode array detector (Waters 2998, Waters, Milford, MA, USA). Gradient elution from 5-95% water/methanol with 0.1 % formic acid was used to elute atrevomycins. Individual peaks were collected and analyzed by LC-MS as described above. Finally, all collected peaks were submitted to final purification step by semipreparative HPLC (Dionex UltiMate 3000, Thermo Fisher Scientific, USA) using a C18 column (Synergi 10 µm, 250 x 10 mm; Phenomenex, Aschaffenburg, Germany) and gradient elution from 5-95% water/acetonitrile with 0.1 % formic acid as eluent.

Atrevomycin A

White powder; 29.3 mg; $[\alpha]_D^{20}$ -88 (c 0.55, MeOH); UV (37% ACN in H_2O + 0.1% FA) λ_{max} 200 nm, 280 nm; ^{1}H and ^{13}C NMR data, see Table S1; ESI-TOF-MS m/z 1289.5965 $[M+2H]^+$ (calc. for $C_{62}H_{84}N_{10}O_{20}$ 1289.5936), see Figure S1.

Atrevomycin B

White powder; 6.8 mg; $[\alpha]_D^{20}$ -10 (c 0.68, MeOH); UV (37% ACN in H_2O + 0.1% FA) λ_{max} 200 nm, 276 nm; ^1H and ^{13}C NMR data, see Table S2; ESI-TOF-MS m/z 1191.5598 $[M+H]^+$ (calc. for $C_{57}H_{76}N_{10}O_{18}$ 1191.5568), see Figure S1.

NMR Data Acquisition and Optical Rotation (OR)

The chemical structures of all the compounds were determined via multidimensional NMR analysis. ^1H-NMR, ^{13}C-NMR, and 2D spectra were recorded at 500 MHz (^1H)/ 126 MHz (^{13}C), conducted in the Bruker Avance Neo 500 MHz, equipped with a Prodigy Cryo-probe. Samples were dissolved in methanol-d_4 or dimethyl sulfoxide-d_6. Chemical shifts are reported in ppm relative to tetramethylsilane; the solvent was used as the internal standard. Coupling constants are reported in Hertz (Hz). Multiplicity is reported with the usual abbreviations (s: singlet, br s: broad singlet, d: doublet, dd: doublet of doublets, ddd: doublet of doublet of doublets, t: triplet, dt: doublet of triplets, q: quartet, m: multiplet). Chiroptical measurements of all the compounds in H_2O ($[\alpha]_D^{20}$) were obtained on a model Jasco P-2000 Automatic Digital Polarimeter (JASCO, Easton, MD, USA) in a 3.5 x 50mm cell at 20 °C.
Optical rotations were measured using a JASCO P-2000 digital polarimeter (28600 Mary's Ct, Easton, MD, USA)

Marfey's Method

500 µg of atrevomycin A was hydrolyzed in 100 µL 6 N HCL at 100°C for 45 min in a closed vial filled with nitrogen. The sample was then dried for 15 min and dissolved in 110 µL of water before 50 µL was transferred to a 1.5 mL Eppendorf tube. After that, 20 µL of 1 N $NaHCO_3$ and 20 µL of 1% L-FDLA or D-FDLA in acetone was added to the hydrolysate. The amino acid standards were prepared in the same way using only L-FDLA. All reaction mixtures were incubated at 40°C for 2 hours at 700 rpm and subsequently quenched with 10 µL 2 N HCL to stop the reaction. The samples were diluted with 300 µL ACN ending in 400 µL total volume from which 1 µL of each sample was analyzed in a maXis high-resolution LC-QTOF system using aqueous ACN with 0.1 vol% FA 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. Sample detection was carried out at 340 nm.

Antimicrobial susceptibility test

Minimum inhibitory concentrations (MICs) were determined according to standard procedures. Single colonies of the tested strains were suspended in cation-adjusted Müller-Hinton broth to achieve a final inoculum of approximately 104 CFU mL-1. Serial dilutions of 409 (0.03 to 64 µg mL-1) were prepared in sterile 96-well plates before the strain suspension was added. Growth inhibition was assessed after overnight incubation (16-18 h) at 30-37°C. A panel consisting of the following strains was tested: *B. subtilis* DSM-10, *S. aureus* Newman, *Mycobacterium smegmatis* MC2155, *Citrobacter freundii* DSM-30039, *E. coli* BW25113 (wt), *E. coli* JW0451-2 (ΔacrB), *Pseudomonas aeruginosa* PA14 DSM-19882, *Acinetobacter baumanii* DSM-30008, *Mucor hiemalis* DSM-2656, *Pichia anomala* DSM-6766, *Cryptococcus neoformans* DSM-11959, *Candida albicans* DSM-1665, CHO-K1 and HepG2.

Author contributions

P.O., M.S., L.H. designed experiments. P.O., M.S., L.H.., A.Z. and G.U. performed experiments. M.S. solved NMR structures, P.O. and M.S. wrote the manuscript. All authors edited the manuscript. A.L. supervised the work.

2.3.5 References

1. Davies, J. S., The cyclization of peptides and depsipeptides. *Journal of peptide science: an official publication of the European Peptide Society* **2003,** *9* (8), 471-501.
2. Taevernier, L.; Wynendaele, E.; Gevaert, B.; De Spiegeleer, B., Chemical classification of cyclic depsipeptides. *Current Protein and Peptide Science* **2017,** *18* (5), 425-452.
3. Pohle, S.; Appelt, C.; Roux, M.; Fiedler, H.-P.; Süssmuth, R. D., Biosynthetic gene cluster of the non-ribosomally synthesized cyclodepsipeptide skyllamycin: deciphering unprecedented ways of unusual hydroxylation reactions. *Journal of the American Chemical Society* **2011,** *133* (16), 6194-6205.
4. Bracegirdle, J.; Hou, P.; Nowak, V. V.; Ackerley, D. F.; Keyzers, R. A.; Owen, J. G., Skyllamycins D and E, non-ribosomal cyclic depsipeptides from lichen-sourced *Streptomyces* anulatus. *Journal of Natural Products* **2021,** *84* (9), 2536-2543.
5. Wang, X.; Gong, X.; Li, P.; Lai, D.; Zhou, L., Structural diversity and biological activities of cyclic depsipeptides from fungi. *Molecules* **2018,** *23* (1), 169.
6. Alonzo, D. A.; Schmeing, T. M., Biosynthesis of depsipeptides, or Depsi: The peptides with varied generations. *Protein Science* **2020,** *29* (12), 2316-2347.
7. Farah, H. I.; Supratman, U.; Hidayat, A. T.; Maharani, R., An overview of the synthesis of biologically active cyclodepsipeptides. *ChemistrySelect* **2022,** *7* (1), e202103470.

8. Zeng, M.; Tao, J.; Xu, S.; Bai, X.; Zhang, H., Marine organisms as a prolific source of bioactive depsipeptides. *Marine Drugs* **2023,** *21* (2), 120.
9. Frisvad, J. C.; Houbraken, J.; Popma, S.; Samson, R. A., Two new Penicillium species Penicillium buchwaldii and Penicillium spathulatum, producing the anticancer compound asperphenamate. *FEMS Microbiology Letters* **2013,** *339* (2), 77-92.
10. Li, W.; Fan, A.; Wang, L.; Zhang, P.; Liu, Z.; An, Z.; Yin, W.-B., Asperphenamate biosynthesis reveals a novel two-module NRPS system to synthesize amino acid esters in fungi. *Chemical science* **2018,** *9* (9), 2589-2594.
11. Wei, G.; Xi, W.; Nussinov, R.; Ma, B., Protein ensembles: how does nature harness thermodynamic fluctuations for life? The diverse functional roles of conformational ensembles in the cell. *Chemical reviews* **2016,** *116* (11), 6516-6551.
12. Song, Q.; Cheng, Z.; Kariuki, M.; Hall, S. C.; Hill, S. K.; Rho, J. Y.; Perrier, S., Molecular self-assembly and supramolecular chemistry of cyclic peptides. *Chemical Reviews* **2021,** *121* (22), 13936-13995.
13. Lai, Z.; Yuan, X.; Chen, H.; Zhu, Y.; Dong, N.; Shan, A., Strategies employed in the design of antimicrobial peptides with enhanced proteolytic stability. *Biotechnology Advances* **2022,** *59*, 107962.
14. Sun, C.; Yang, Z.; Zhang, C.; Liu, Z.; He, J.; Liu, Q.; Zhang, T.; Ju, J.; Ma, J., Genome mining of Streptomyces atratus SCSIO ZH16: discovery of atratumycin and identification of its biosynthetic gene cluster. *Organic letters* **2019,** *21* (5), 1453-1457.
15. Blin, K.; Shaw, S.; Augustijn, H. E.; Reitz, Z. L.; Biermann, F.; Alanjary, M.; Fetter, A.; Terlouw, B. R.; Metcalf, W. W.; Helfrich, E. J., antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic acids research* **2023,** *51* (W1), W46-W50.
16. Zhang, C.; Seyedsayamdost, M. R., Discovery of a cryptic depsipeptide from streptomyces ghanaensis via MALDI-MS-guided high-throughput elicitor screening. *Angewandte Chemie International Edition* **2020,** *59* (51), 23005-23009.
17. Machushynets, N. V.; Elsayed, S. S.; Du, C.; Siegler, M. A.; de la Cruz, M.; Genilloud, O.; Hankemeier, T.; van Wezel, G. P., Discovery of actinomycin L, a new member of the actinomycin family of antibiotics. *Scientific Reports* **2022,** *12* (1), 2813.
18. Gubbens, J.; Wu, C.; Zhu, H.; Filippov, D. V.; Florea, B. I.; Rigali, S. b.; Overkleeft, H. S.; van Wezel, G. P., Intertwined Precursor Supply during Biosynthesis of the Catecholate–Hydroxamate Siderophores Qinichelins in Streptomyces sp. MBT76. *ACS Chemical Biology* **2017,** *12* (11), 2756-2766.
19. Patel, K. D.; d'Andrea, F. B.; Gaudelli, N. M.; Buller, A. R.; Townsend, C. A.; Gulick, A. M., Structure of a bound peptide phosphonate reveals the mechanism of nocardicin bifunctional thioesterase epimerase-hydrolase half-reactions. *Nature communications* **2019,** *10* (1), 3868.
20. Niu, W.; Liu, J.; Duan, Y.; Zhong, L.; Pang, L.; Zhong, G.; Zhang, Y.; Bian, X., Biosynthesis of Nonribosomal Peptides Chitinimides Reveal a Special Type of Thioesterase Domains. *Chemistry–A European Journal* **2024,** *30* (69), e202402763.
21. Hermes, C.; Richarz, R.; Wirtz, D. A.; Patt, J.; Hanke, W.; Kehraus, S.; Voß, J. H.; Küppers, J.; Ohbayashi, T.; Namasivayam, V., Thioesterase-mediated side chain transesterification generates potent Gq signaling inhibitor FR900359. *Nature communications* **2021,** *12* (1), 144.
22. Wheadon, M. J.; Townsend, C. A., Evolutionary and functional analysis of an NRPS condensation domain integrates β-lactam, D-amino acid, and dehydroamino acid synthesis. *Proceedings of the National Academy of Sciences* **2021,** *118* (17), e2026017118.
23. Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H., Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC evolutionary biology* **2007,** *7*, 1-15.

24. Kieser, T.; Bibb, M.; Buttner, M.; Chater, K.; Hopwood, D., Practical streptomyces genetics: John innes foundation. *Norwich Research Park, Colney* **2000**, 44-61.
25. Russell, D. W.; Sambrook, J., Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press New York: 2001.
26. Green, M. R.; Sambrook, J., Molecular cloning. *A Laboratory Manual 4th* **2012,** *448*.
27. Buckingham, J. E. Dictionary of Natural Products. https://dnp.chemnetbase.com.
28. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28:1647–9.

2.3.6 Supplementary information

Table S1: NMR data (500 MHz, MeOD-d_3) of 1289.

Subunit	Position	δ(^{13}C or ^{15}N) [ppm]	δ(^{1}H) [ppm], mult(J)	key-HMBC (H to C)	key-NOESY/ ROESY
Acyl	1	171.6	-		
	2	122.8	6.99, d (15.3)	1	
	3	140.4	8.05, d (15.3)	1	10
	4	134.4	-		
	5	127.7	7.75, d (8.1)		
	6	128.5	7.32, t (7.6)		
	7	130.3	7.42, t (7.5)		
	8	128.7	7.61, d (7.9)		
	9	142.8	-		
	10	73.5	5.09, bs		
	11	77.1	4.05, ovl		3
	12	131.2	5.41, ovl		
	13	129.1	5.42, ovl		
	14	18.1	1.54, d (5.3)		
Thr-1	CO	172.3	-		
	α	62.4	4.44, t (3.3)	1, Thr^{1}-CO	Ser1-NH
	β	70.7	5.30, dq (3.3, 6.4)		
	γ	14.7	1.24, d (6.7)		Leu2-α
	NH	113.6	8.98, bs		
Ser-1	CO	172.9	-		
	α	55.1	4.76, dt (5.0, 10.0)	Ser1-CO, Thr1-CO	Leu1-NH
	β	61.4	3.89, ovl 3.26, m		
	NH	113.6	8.90, d (8.8)		Thr1-α
Leu-1	CO	177.5	-		
	α	54.0	4.31, ovl	Ser1-CO, Leu1-CO	β-OH-Ph-NH
	β	41.3	1.80, m 1.47, m		
	γ	25.7	1.30, m		
	δ1	23.1	0.85, d (6.4)		
	δ2	22.9	0.79, d (6.4)		
	NH	121.7	7.88, d (6.0)		Ser1-α
β-OH-Phe	CO	173.9	-		
	α	63.5	4.31, ovl	β-OH-Ph-CO	Ser2-NH
	β	72.4	5.26, bs		
	1	142.6	-		
	2,3	127.0	7.51, d (7.4)		
	4,5	129.4	7.35, t (7.6)		
	6	128.6	7.26, t (7.4)		
	NH	112.9	8.38, d (6.0)		Ser2-NH, Leu1-α

Ser-2	CO	172.2	-		
	α	58.5	4.33, dt (4.1, 6.9)	β-OH-Ph-CO, Ser2-CO	Thr2-NH
	β	62.5	3.90, ovl		
	NH	112.4	7.77, d (6.7)		β-OH-Ph-α
Thr-2	CO	173.1	-		
	α	61.4	4.01, t (9.3)	Ser2-CO, Thr2-CO	Ser3-NH
	β	68.8	3.78, m		
	γ	20.4	0.71, d (5.7)		
	NH	114.7	7.18, d (7.9)		Ser2-α
Ser-3	CO	171.2	-		
	α	56.3	4.58, m	Thr2-CO, Ser3-CO	HPG-NH
	β	62.4	3.86, ovl		
	NH	119.9	8.16, d (8.8)		Thr2-α
HPG	CO	174.1	-		
	α	58.4	5.54, d (7.9)	Ser3-CO, HPG-CO	Gly-NH
	1	128.6	-		
	2,3	130.9	7.28, d (8.3)		
	4,5	116.9	6.77, d (8.1)		
	6	159.0	-		
	NH	122.5	7.81, d (7.4)		Ser3-α
Gly	CO	172.7	-		
	α	43.8	4.04, ovl 3.41, dd (5.8, 15.4)	HPG-CO, Gly-CO	Leu2-NH
	NH	106.6	8.43, t (5.8)		HPG-α
Leu-2	CO	172.9	-		
	α	53.9	4.20, m	Gly-CO, Leu2-CO	Thr1-γ
	β	40.5	1.64, m 1.56, m		
	γ	25.9	1.70, m		
	δ1	21.7	0.88, d (6.4)		
	δ2	23.2	0.94, d (6.4)		
	NH	118.3	8.49, d (8.5)		Glyc-α

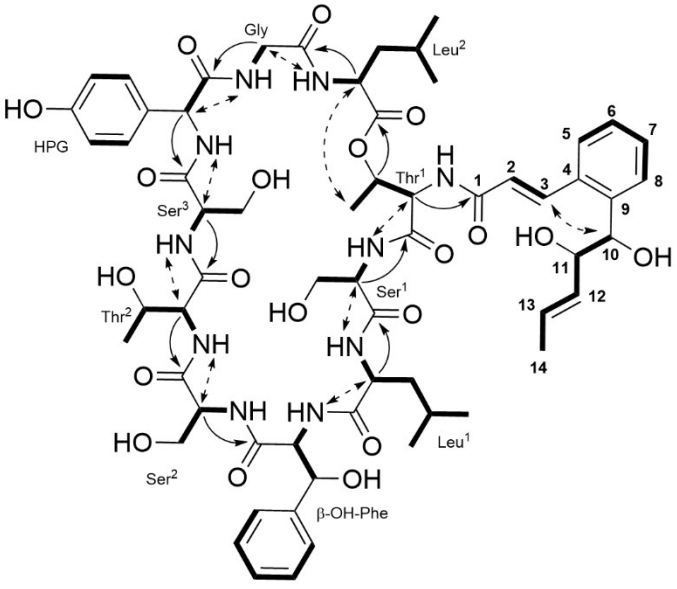


Figure S1: Structure of 1289 showing COSY ( ), HMBC ( ) and ROESY/NOESY ( ) correlations.

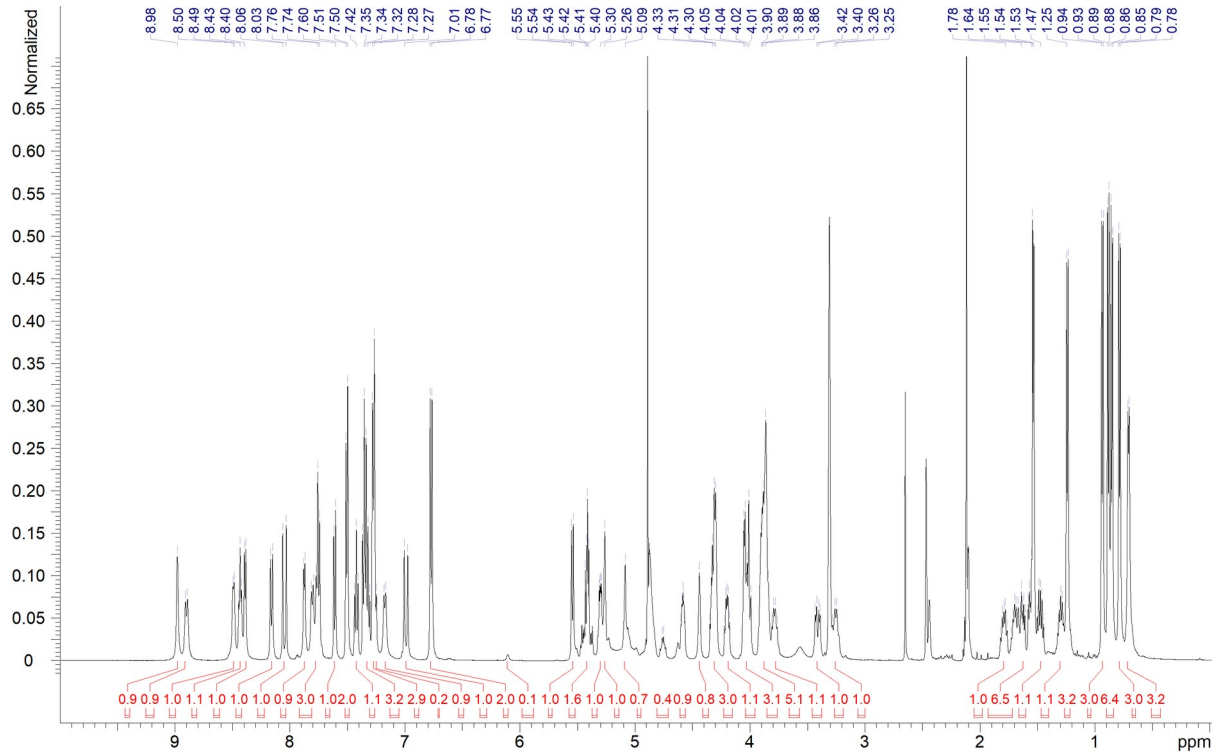


Figure S2: ^{1}H-NMR (500 MHz) spectrum of 1289 in MeOD-d_3.

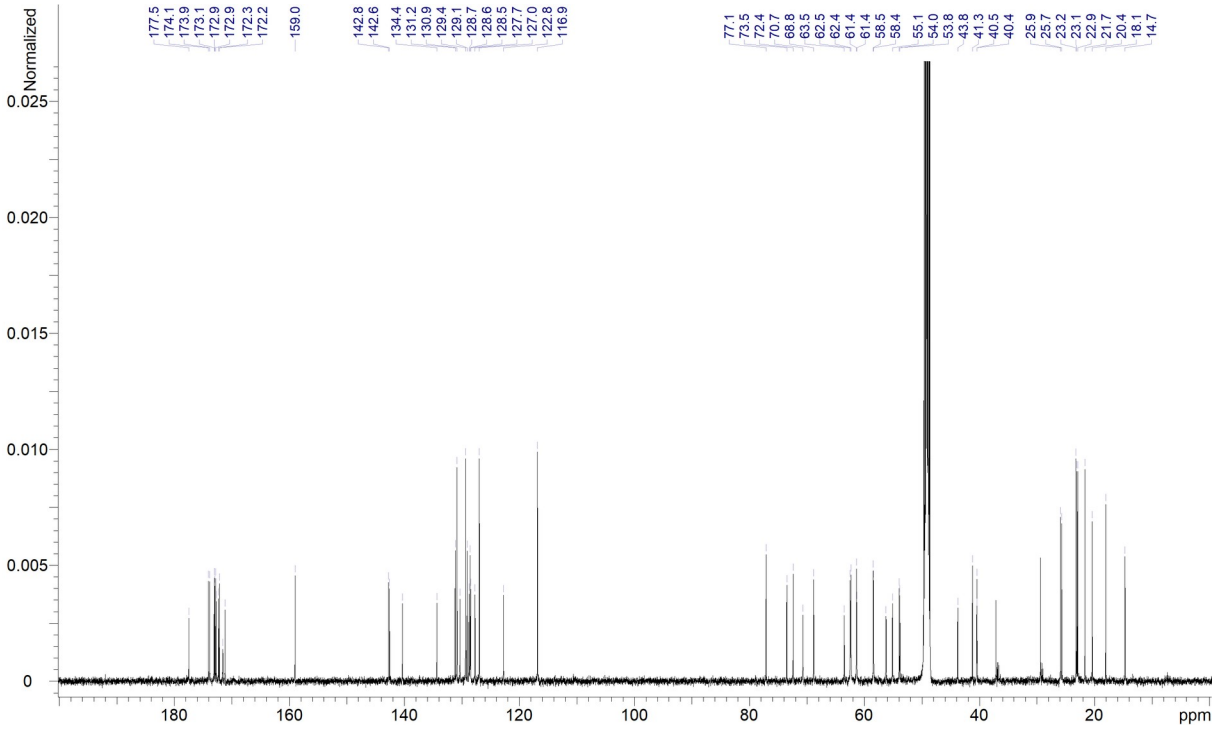


Figure S3: ^{13}C-NMR (125 MHz) spectrum of 1289 in MeOD-d_3.

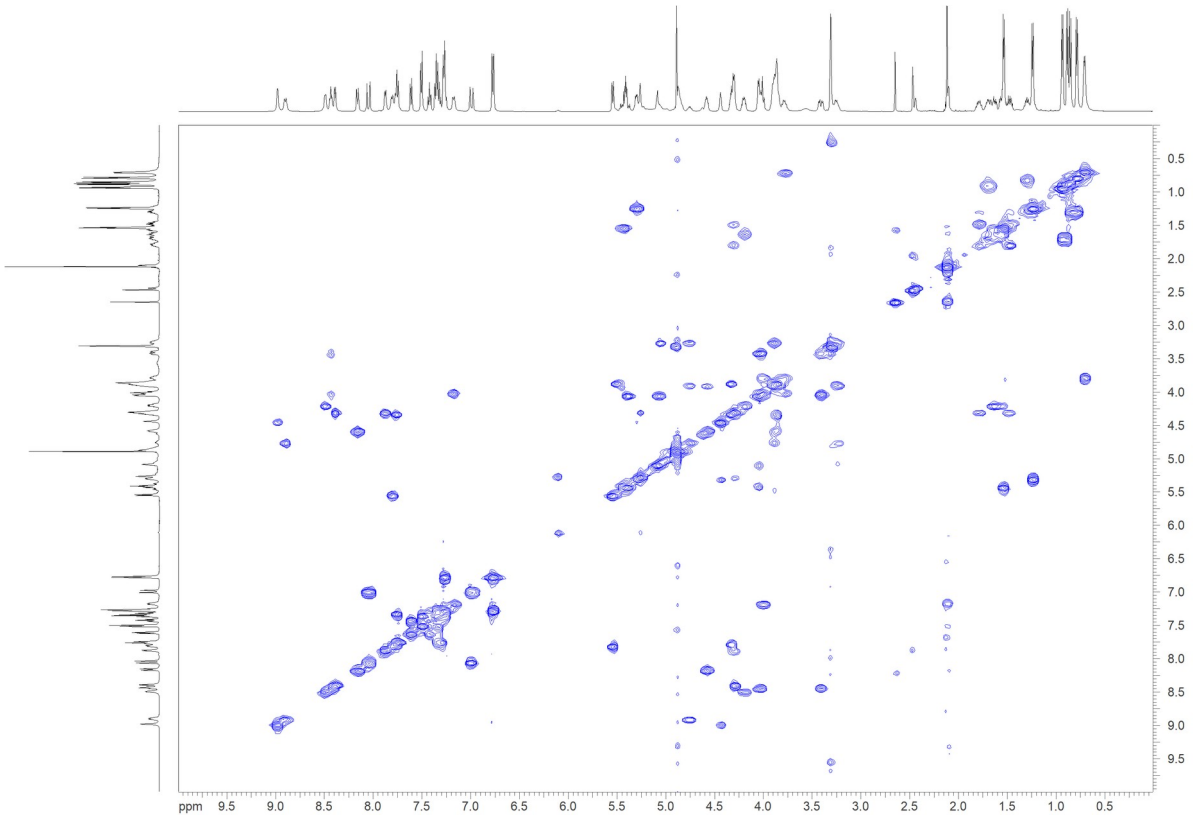


Figure S4: COSY spectrum of 1289 in MeOD-d_3.

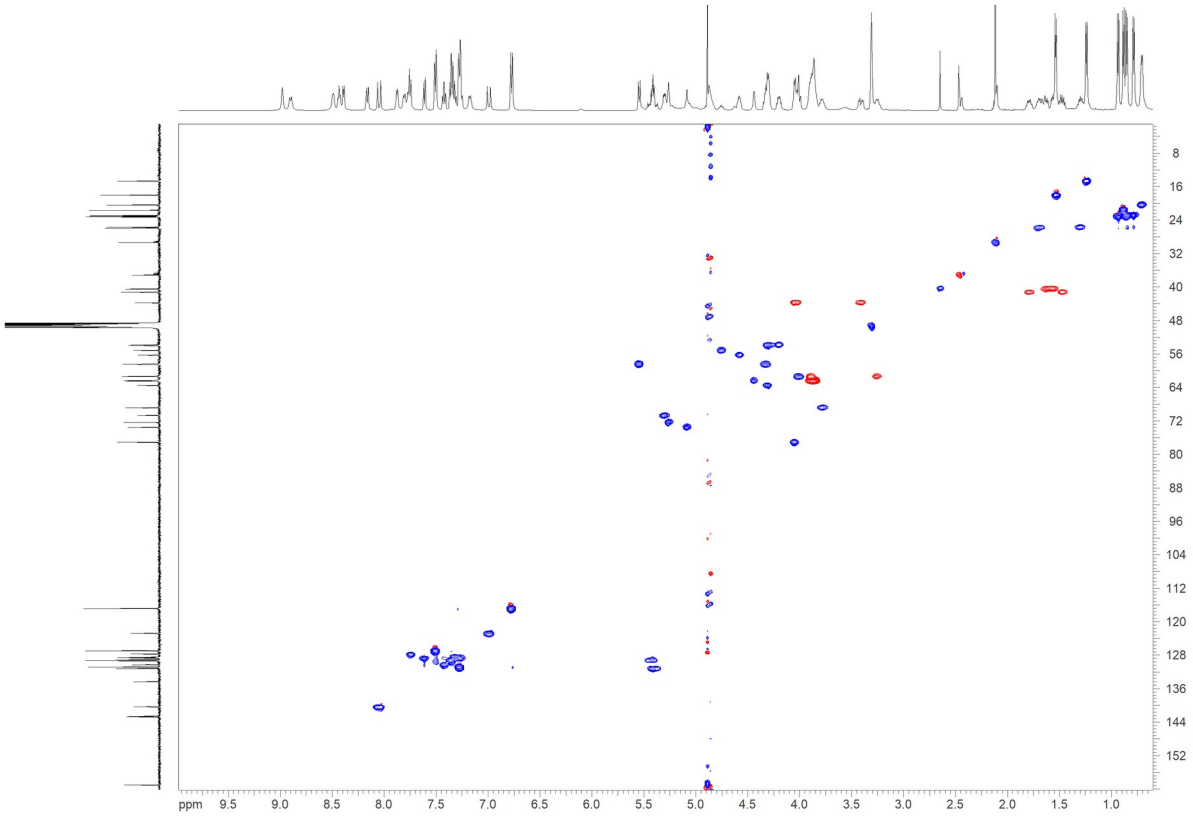


Figure S5: Edited-HSQC spectrum of 1289 in MeOD-d_3.

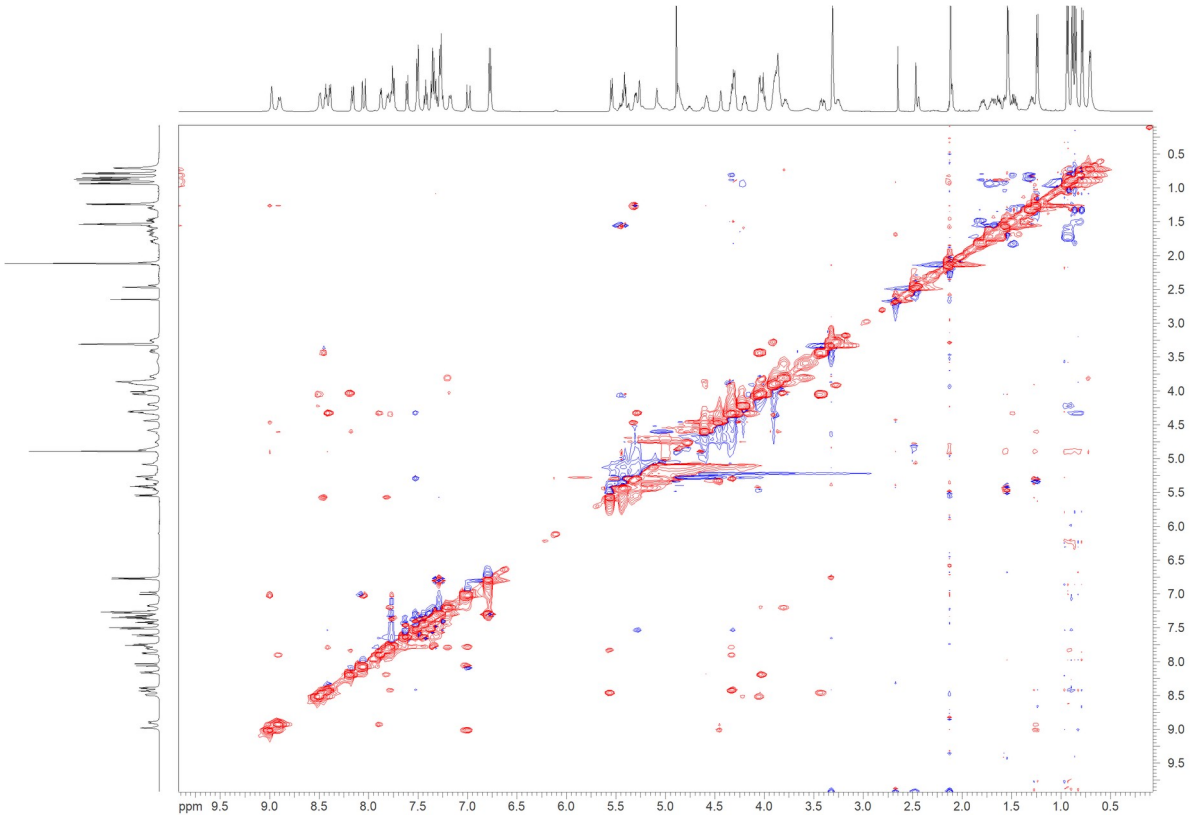


Figure S6: NOESY spectrum of 1289 in MeOD-d_3.

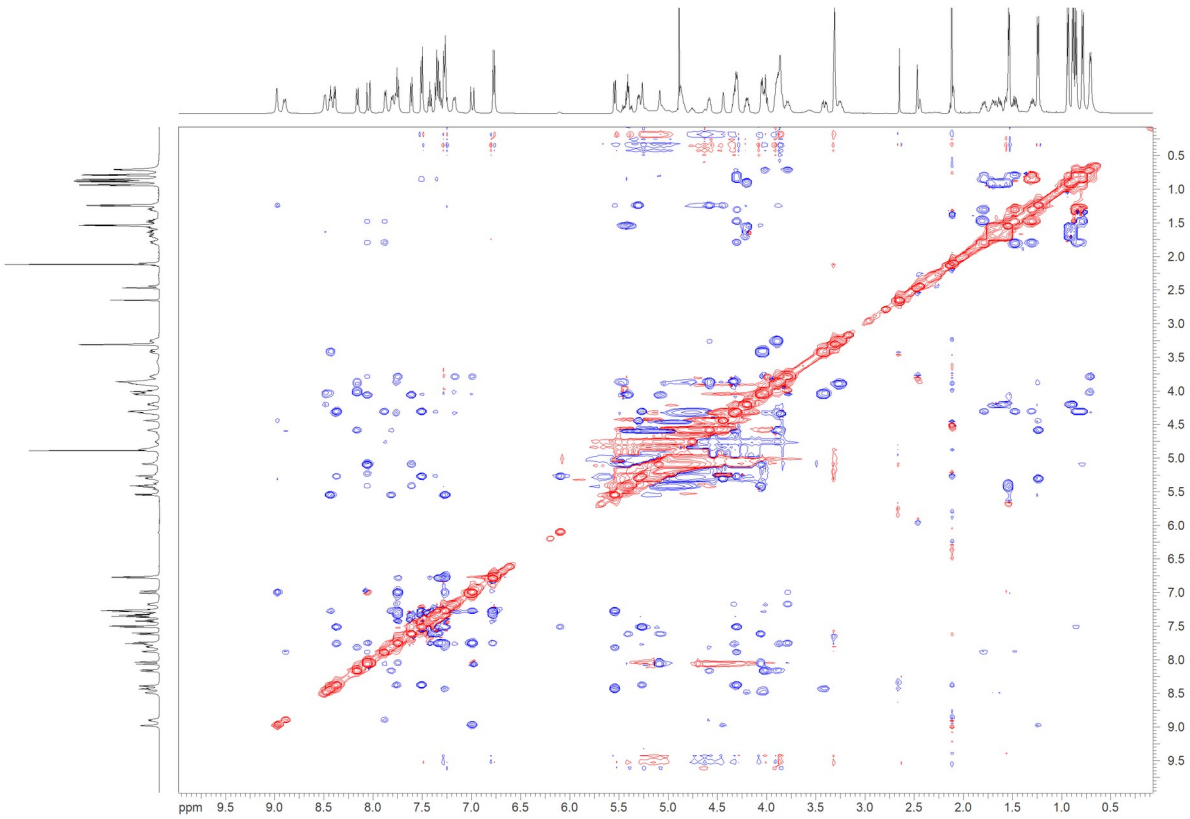


Figure S7: ROESY spectrum of 1289 in MeOD-d_3.

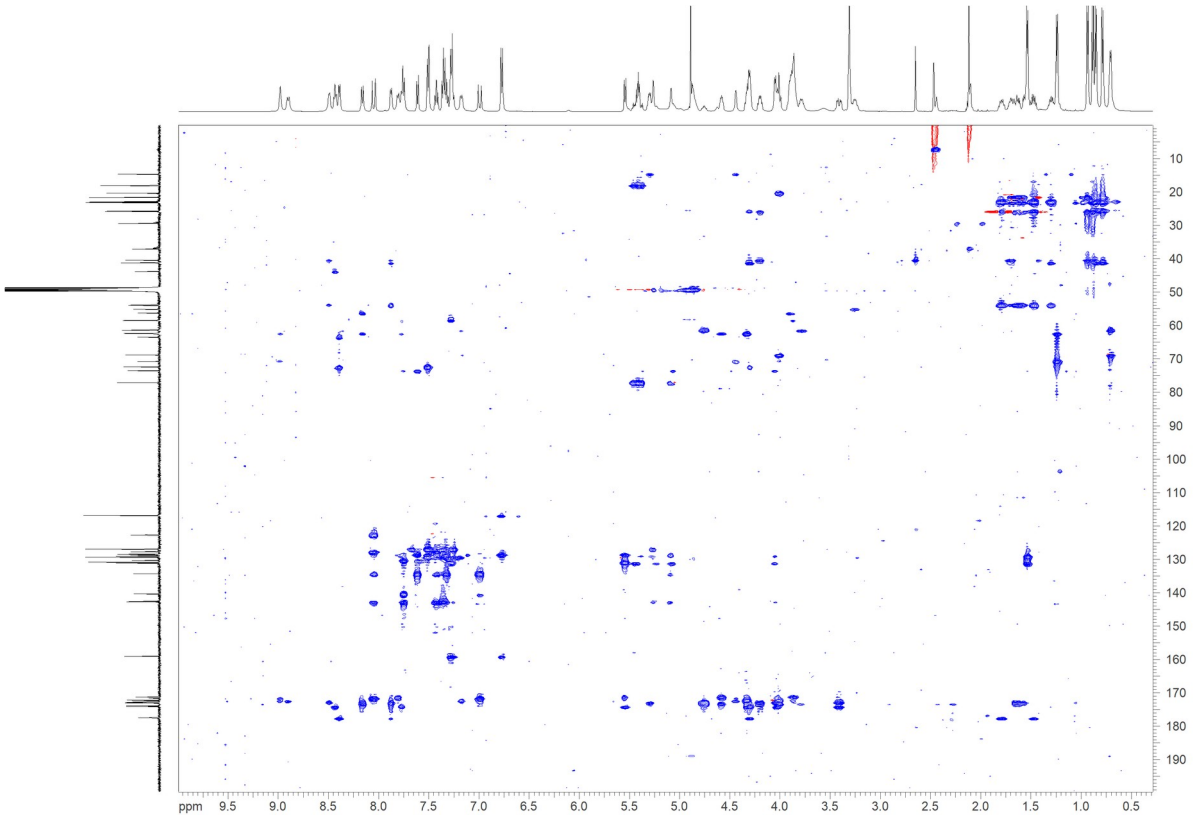


Figure S8: HMBC spectrum of 1289 in MeOD-d_3.

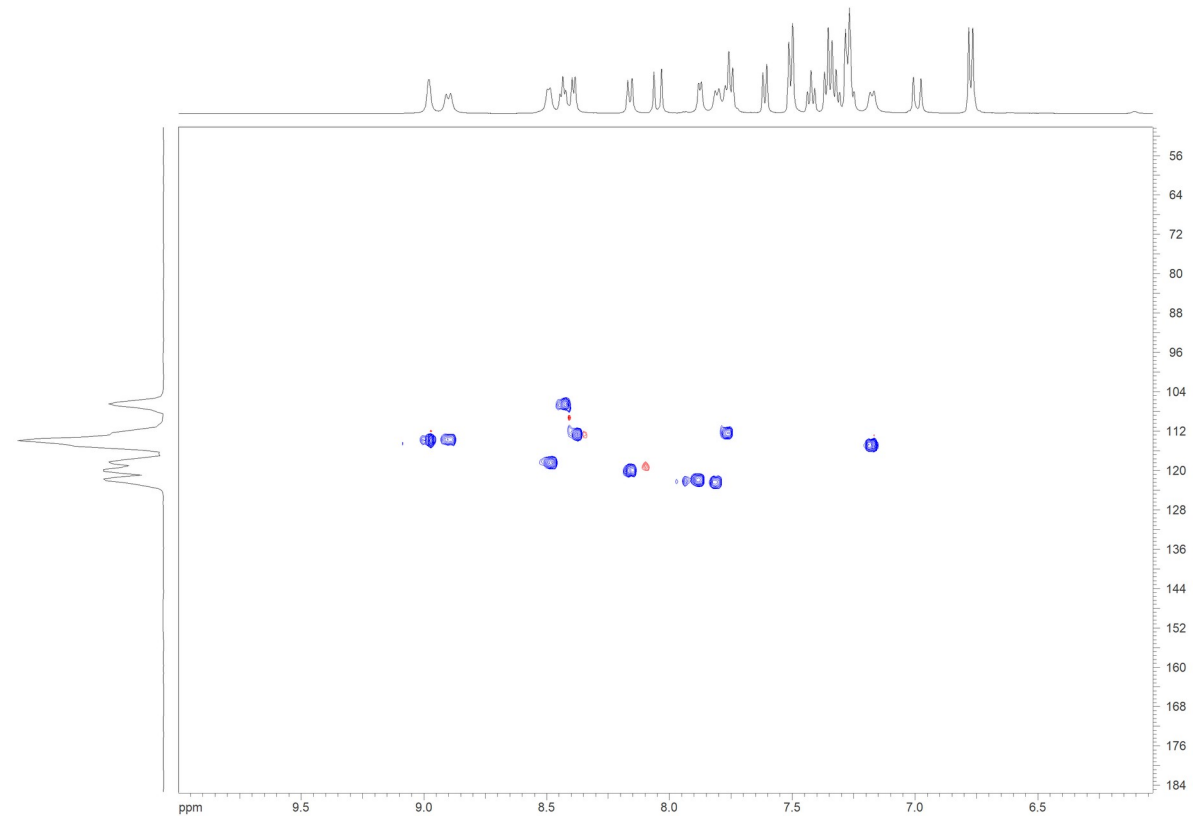


Figure S9: ^{15}N-HSQC spectrum of 1289 in MeOD-d_3.

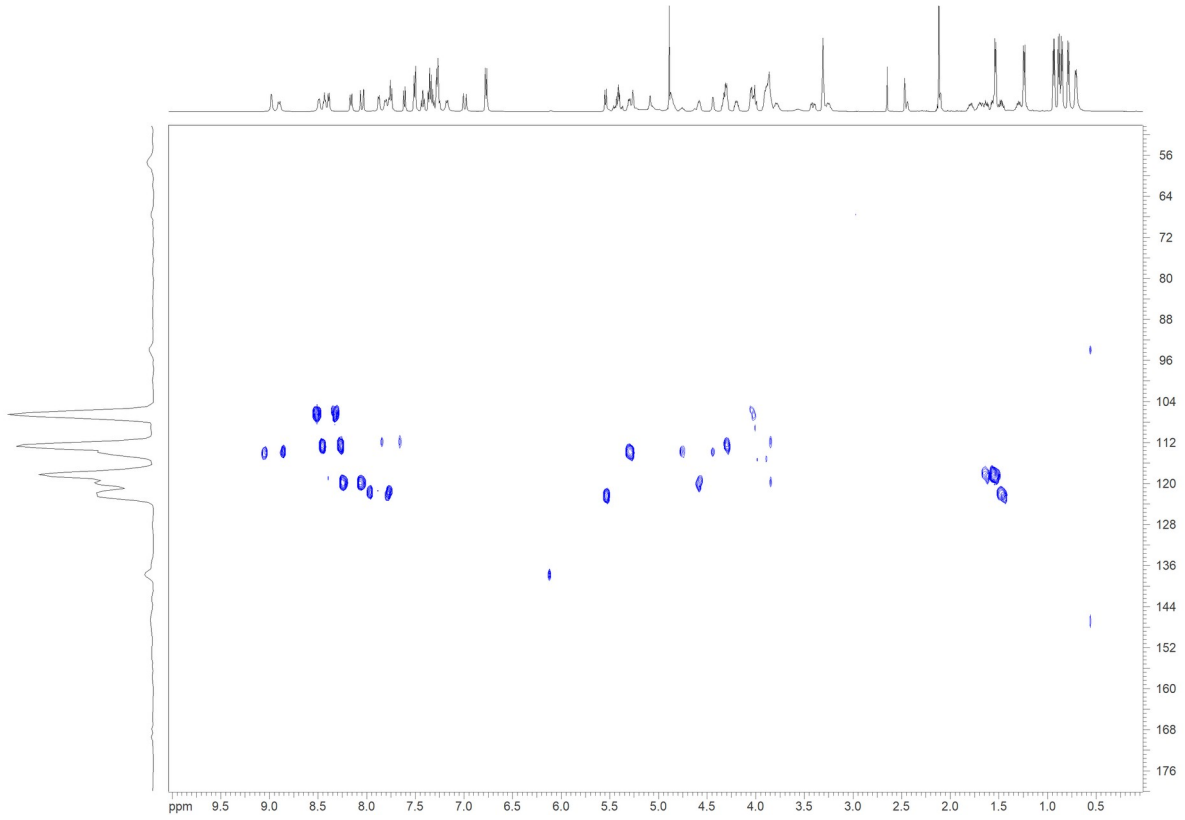


Figure S10: ^{15}N-HMBC spectrum of 1289 in MeOD-d_3.

Table S2: NMR data (500 MHz, MeOD-d_3) of 1191.

Subunit	Position	δ(^{13}C) [ppm]	δ(^{1}H) [ppm], mult(J)	key-HMBC (H to C)
Acyl	1	178.8	-	
	2	37.5	3.03, ovl 2.90, ovl	
	3	31.9	3.04, ovl 2.93, ovl	
	4	143.2	-	
	5, 5'	129.7	7.24, ovl	
	6, 6'	129.7	7.30, ovl	
	7	127.2	7.23, ovl	
Thr-1	CO	172.4	-	
	α	61.9	4.35, t (3.51)	1, Thr^{1}-CO
	β	71.0	5.25, dq (6.9, 3.5)	Leu2-CO
	γ	14.8	1.22, d (6.3)	
	NH	-	8.94, d (1.4)	
Ser-1	CO	173.3	-	
	α	54.6	4.78, dd (9.5, 5.3)	Ser1-CO, Thr1-CO
	β	61.3	3.91, ovl 3.20, t (10.2)	
	NH	-	8.76, d (6.2)	
Leu-1	CO	177.8	-	
	α	54.5	4.24, m	Ser1-CO, Leu1-CO

	β	40.9	1.65, ovl	
	γ	25.6	1.15, m	
	δ1	23.3	0.70, d (6.4)	
	δ2	22.9	0.79, d (6.3)	
	NH	-	7.72, d (4.9)	
β-OH-Phe	CO	173.7	-	
	α	63.8	4.32, dd (2.0, 6.4)	Leu1-CO, β-OH-Ph-CO
	β	72.2	5.38, d (1.7)	
	1	142.9	-	
	2,3	126.9	7.52, d (7.3)	
	4,5	129.5	7.36, t (4.5)	
	6	128.7	7.27, ovl	
	NH	-	8.52, ovl	
Ser-2	CO	172.2	-	
	α	58.6	4.47, dd (7.5, 4.6)	β-OH-Ph-CO, Ser2-CO
	β	62.5	3.92, ovl	
	NH	-	7.85, d (7.4)	
Thr-2	CO	173.2	-	
	α	61.4	4.08, t (4.5)	Ser2-CO, Thr2-CO
	β	68.6	4.10, m	
	γ	20.7	1.10, d (5.2)	
	NH	-	7.34, ovl	
Ser-3	CO	171.1	-	
	α	56.2	4.55, m	Thr2-CO, Ser3-CO
	β	62.2	3.87, ovl	
	NH	-	8.26, d (8.5)	
HPG	CO	173.7	-	
	α	58.8	5.36, d (7.2)	Ser3-CO, HPG-CO
	1	128.3	-	
	2,3	130.8	7.10, d (8.2)	
	4,5	116.9	6.63, d (8.2)	
	6	159.0	-	
	NH	-	7.47, d (7.1)	
Glyc	CO	172.6	-	
	α	44.0	3.40, dd (6.2, 15.8)	HPG-CO, Gly-CO
	NH	-	8.32, d (8.5)	
Leu-2	CO	173.0	-	
	α	53.8	4.22, m	Gly-CO, Leu2-CO
	β	40.7	1.57, m	
	γ	26.1	1.72, ovl	
	δ1	23.4	0.96, d (6.4)	
	δ2	21.8	0.90, d (6.3)	
	NH	-	8.34, ovl	

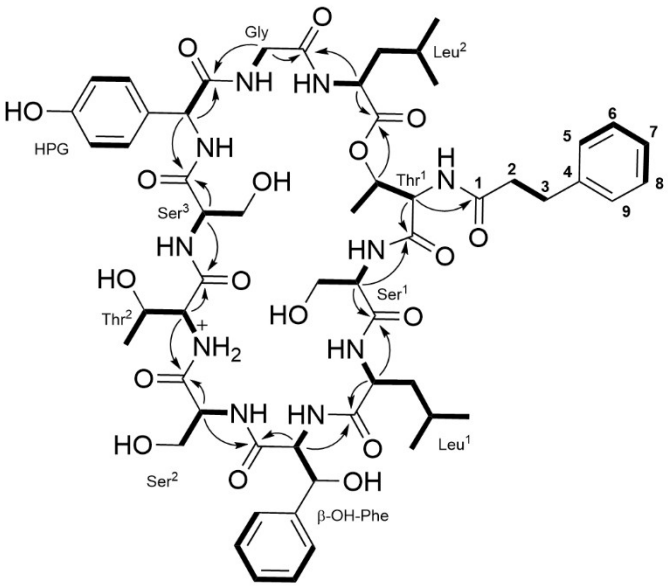


Figure S11: Structure of 1289 showing COSY (▬) and key HMBC (⌒) correlations.

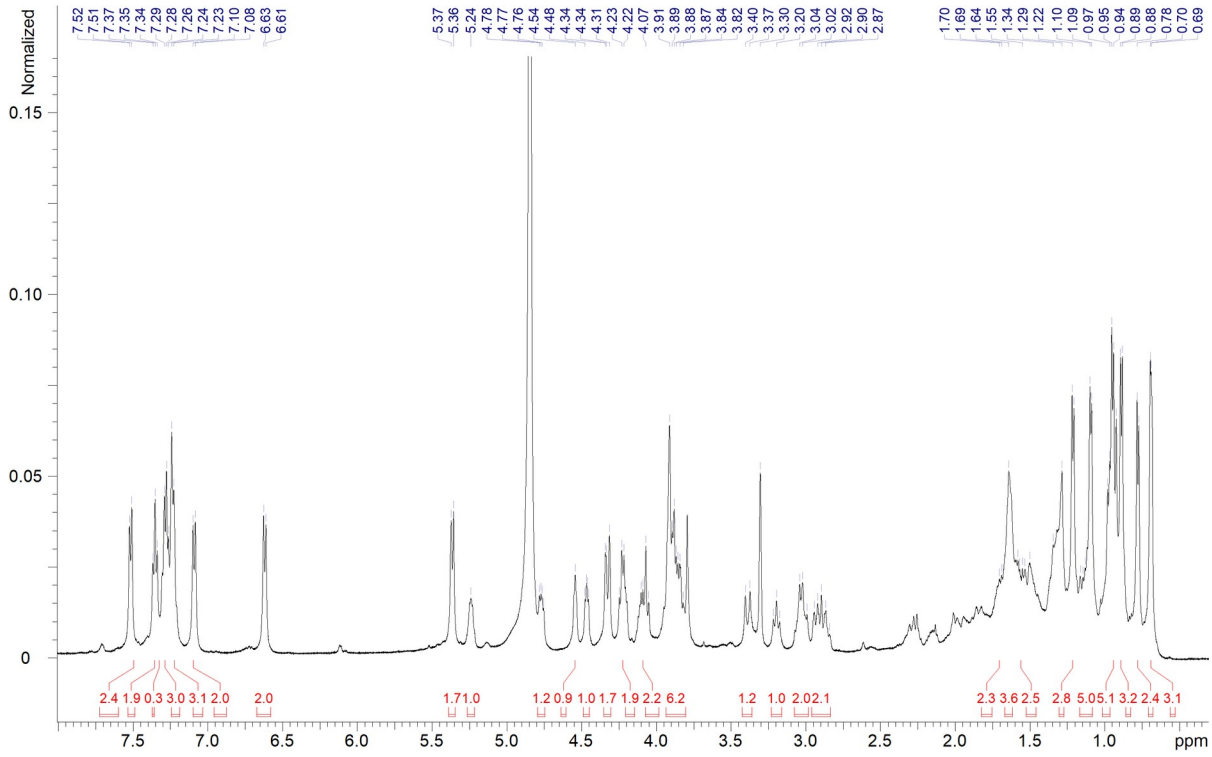


Figure S12: ^{1}H-NMR (500 MHz) spectrum of 1191 in MeOD-d_4.

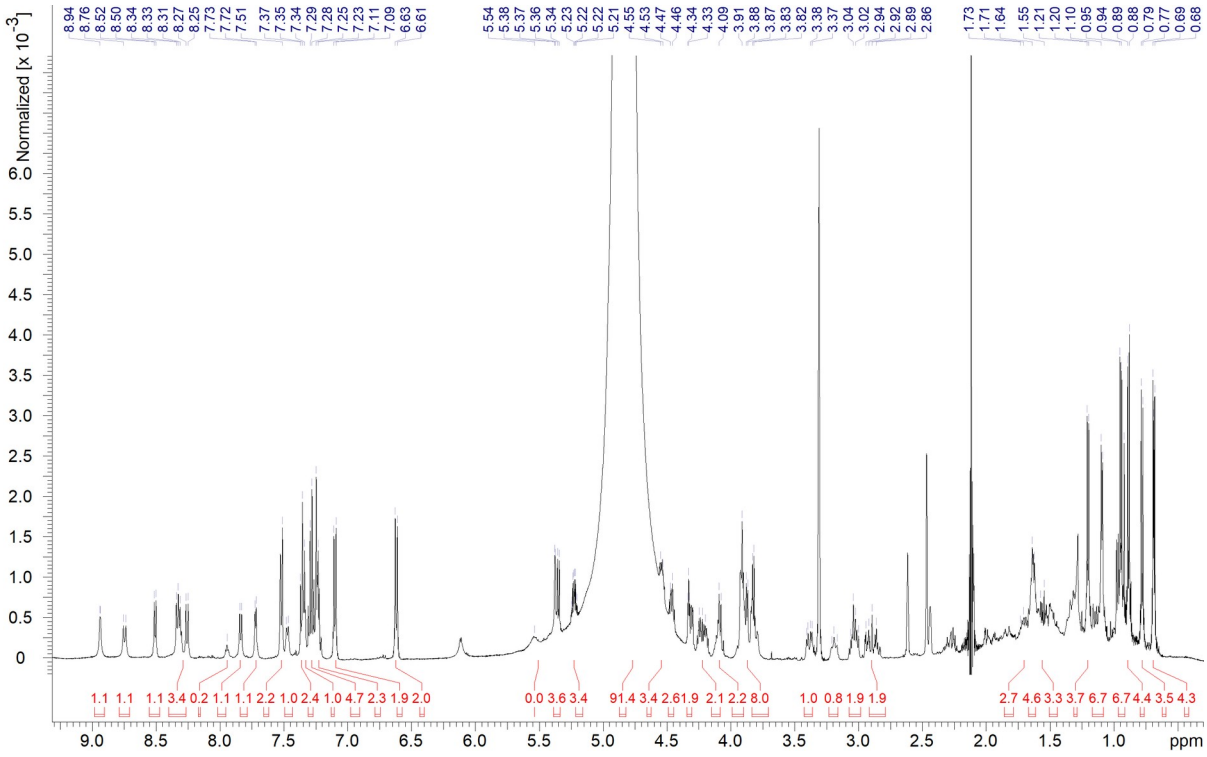


Figure S13: ^{1}H-NMR (500 MHz) spectrum of 1191 in MeOD-d_3.

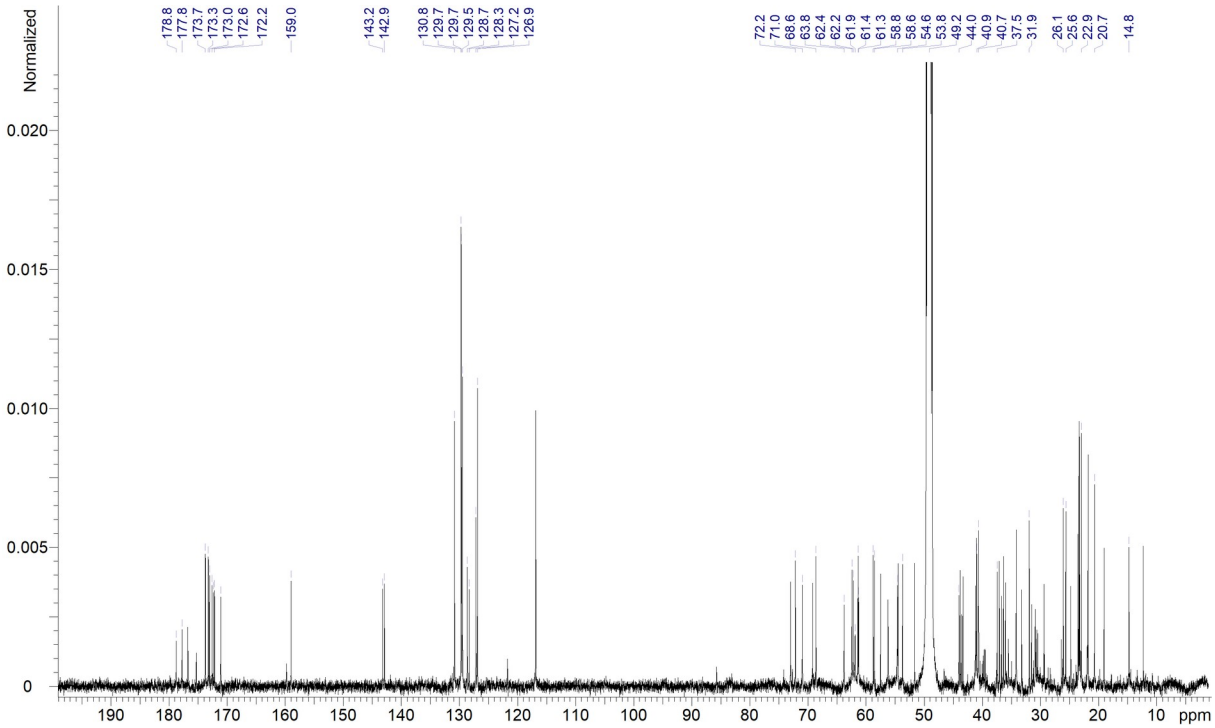


Figure S14: ^{13}C-NMR (125 MHz) spectrum of 1191 in MeOD-d_3.

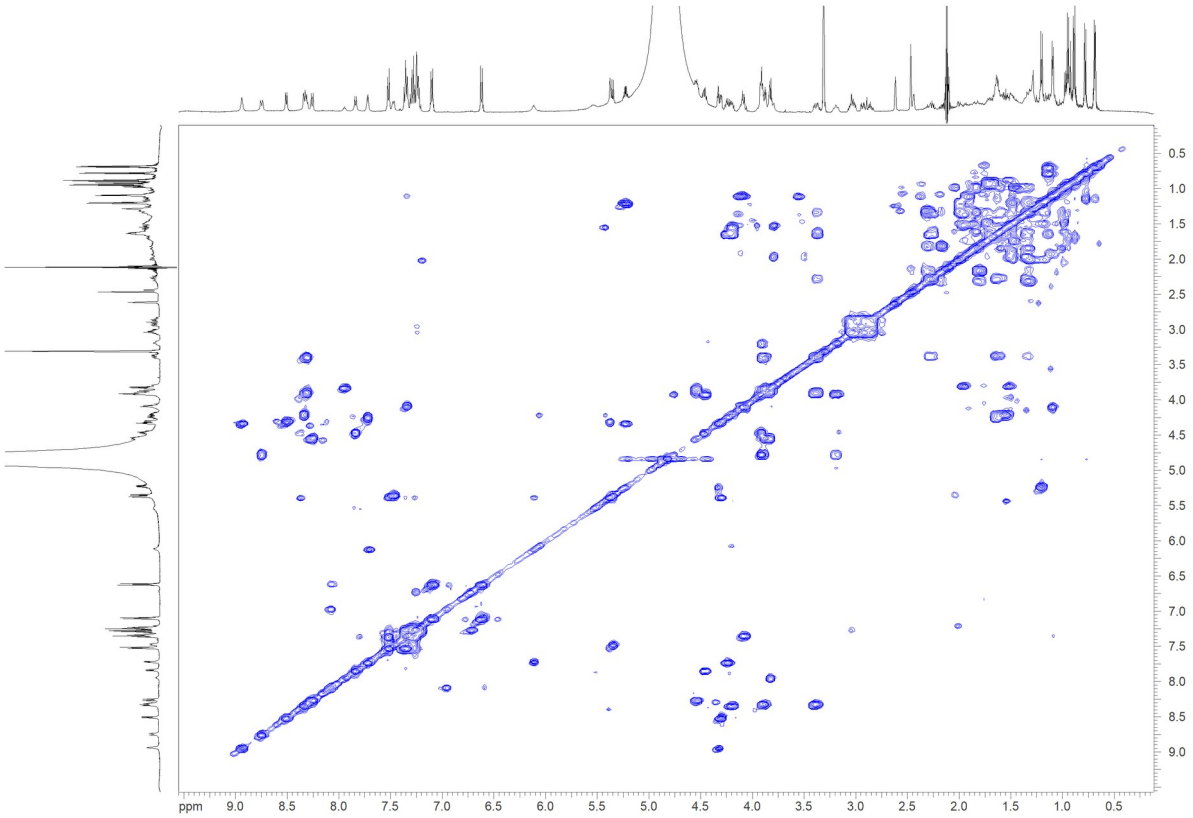


Figure S15: COSY spectrum of 1191 in MeOD-d_3.

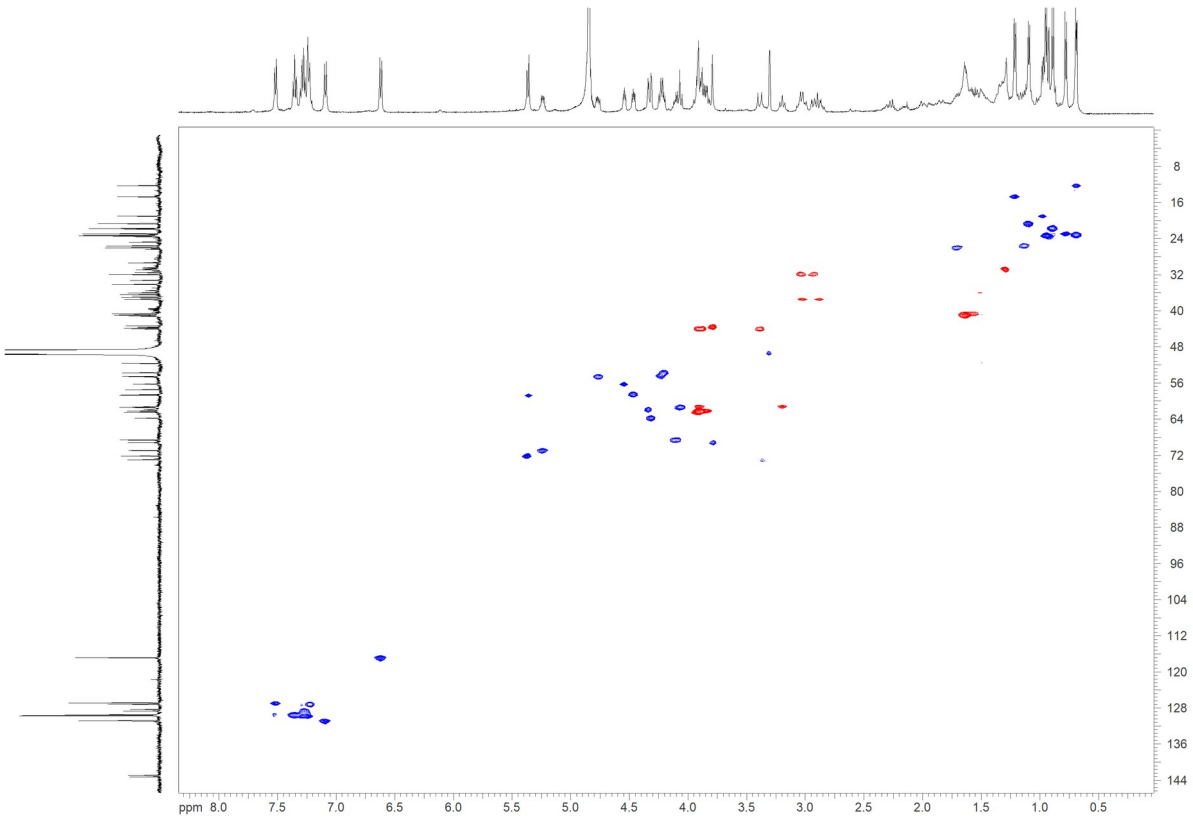


Figure S16: Ed-HSQC spectrum of 1191 in MeOD-d_3.

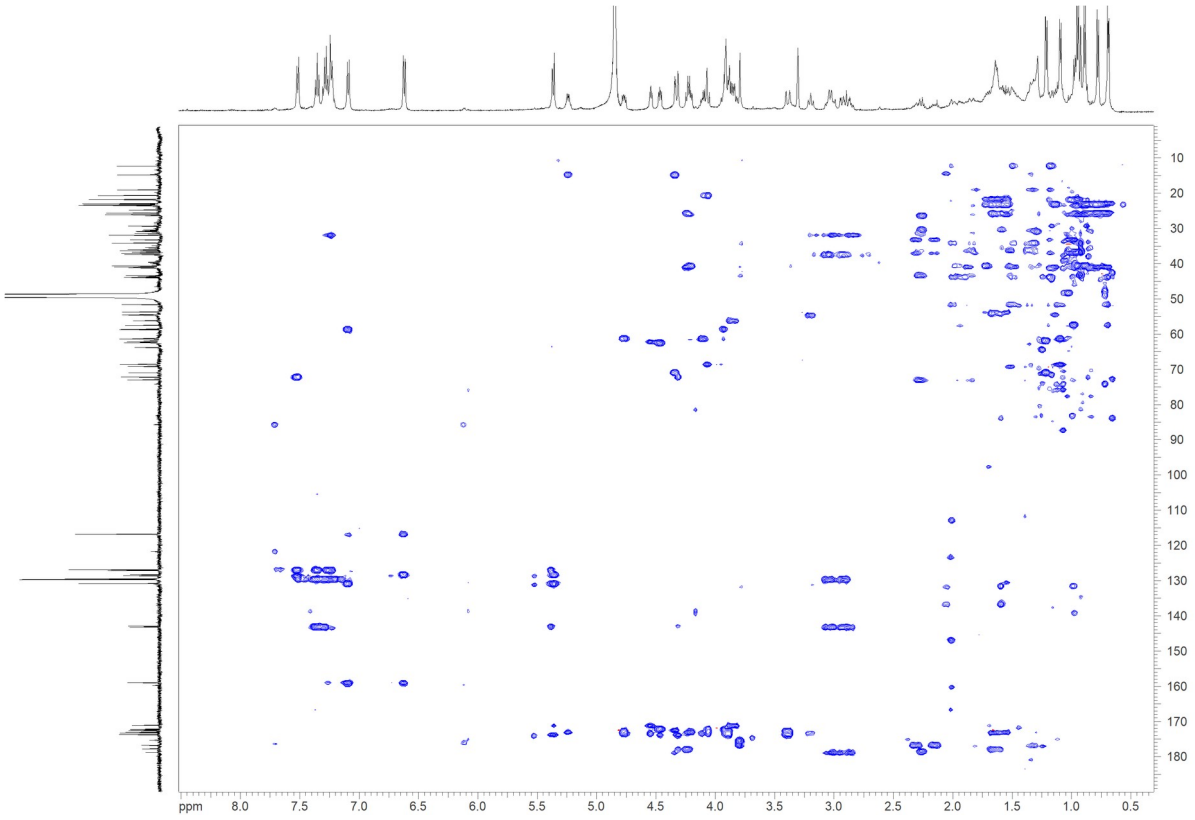


Figure S17: HMBC spectrum of 1191 in MeOD-d_3.

Table S3: Proposed function of the genes from cluster 2 and similarity comparison to the *avm* gene cluster.

cluster 2 cinnapeptin-like		***avm cluster***	
orf	**proposed function**	**corresponding genes**	**% Similarity to cluster 2**
1	Transcriptional regulator		
2	Sigma-70 family RNA polymerase factor		
3	Questin oxidase		
4	Class III extradiol ring-cleavage dioxygenase		
5	MarR family transcriptional regulator		
6	MFS Transporter		
7	TetR/AcrR family transcriptional regulator		
8	M6 family metalloprotease		
9	Class I adenylate forming enzyme		
10	Thioesterase II		
11	Acyl carrier protein		
12	MMPL family transporter		
13	Uncharacterized protein		
14	**Cytochrom P450**	avm28	92
15	Alpha/beta fold hydrolase		
16	**NRPS**	avm30	85
17	**NRPS**	avm31	80

18	ATP-binding cassette domain containing protein		
19	ABC Transporter		
20	Acyltransferase		
21	Acyltransferase		
22	PAS domain containing protein		
23	Phytase		
24	Pyridoxamine 5'-phosphate oxidase		
25	Citrate synthase 2		
26	VOC family protein		
27	Histidine kinase		
28	Response regulator		
29	DM13 domain containing protein		
30	methyltransferase		
31	XRE family transcriptional regulator		
32	Alanine-tRNA ligase-related protein		
33	Methyltransferase		
34	Uncharacterized protein		
35	Uncharacterized protein		
36	Uncharacterized protein		
37	Helix-turn-helix domain containing protein		

Table S4: Proposed function of the genes from cluster 1 and similarity comparison to the *avm* gene cluster.

cluster 1 skyllamycin-like		***avm cluster***	
orf	**proposed function**	**corresponding genes**	**% Similarity to cluster 1**
1	Sodium:solute symporter		
2	SDR family oxidoreductase		
3	Hypothetical protein		
4	DUF397 domain containing protein		
5	Helix-turn-helix transcriptional regulator		
6	GPP34 family phosphoprotein		
7	Serine hydrolase		
8	MFS Transporter		
9	adhesin		
10	ATP-grasp domain containing protein		
11	Aminoglycoside phosphotransferase		
12	MMPL transporter		
13	Polyprenyl synthetase		
14	FMN-dependent NADH azoreductase		
15	LysR family transcriptional regulator		
16	Nuclear transport factor 2		
17	NAD(P) dependent oxidoreductase		
18	**NRPS**	avm3	79

19	LuxR C-terminal-related transcriptional regulator		
20	LuxR family transcriptional regulator		
21	**ACP S-malonyltransferase**	avm6	79
22	IclR family transcriptional regulator		
23	**Acyl-CoA carboxylase subunit beta**	avm10	91
24	**Acyl-CoA carboxylase subunit epsilon**	avm11	57
25	SARP Regulator		
26	Hypothetical protein		
27	SARP regulator		
28	DUF6269 family protein		
29	LysR family transcriptional regulator		
30	Acyl carrier protein		
31	**3-oxoacyl reductase**	avm16	62
32	**2-hydroxychromene-2-carboxylate isomerase**	avm15	54
33	MbtH family protein		
34	AarF/ABC1/UbiB kinase		
35	**Cytochrom P450**	avm43	91
36	**3-hxdroxylacyl-ACP dehydratase FabZ**	avm42	92
37	Hypothetical protein		
38	**Beta-ketoacyl synthase**	avm17	70
39	Hypothetical protien		
40	Alpha/beta fold hydrolase		
41	**Beta-ketoacyl synthase**	avm20	65
42	**Beta-ketoacyl synthase**	avm21	75
43	**Beta-ketoacyl synthase**	avm22	85
44	**Prephenate dehydrogenase**	avm41	83
45	**Aminotransferase classI/II-fold pyridoxal phosphate dependent enzyme**	avm40	86
46	**4-hydroxyphenylpyruvate dioxygenase**	avm39	85
47	Thioesterase II		
48	**SDR oxidoreductase**	avm37	95
49	**NRPS**	avm36	81
50	**NRPS**	avm35	79
51	Transposase		
52	ABC transporter permease		
53	ABC transporter ATP binding		
54	Efflux RND transporter		
55	Hypothetical protein		
56	Response regulator transcription factor		
57	ATP-binding protein		
58	SIS domain-containing protein		
59	Metal-dependent transcriptional regulator		
60	Alpha/beta fold hydrolase		
61	transporter		
62	ABC transporter		
63	Bifunctional DNA primase/polymerase		
64	Transcriptional regulator		

65	Hypothetical protein		
66	ABC transporter substrate-binding		
67	Hypothetical protein		
68	formyltetrahydrofolate deformylase		

Table S5: Bacterial strains used in this study

Strain	**Description**	**Source**
S. lividans Del8	Cluster free derivative of *S. lividans* TK24	Yousra A., et al.[1]
S. albus Del14	Cluster free derivative of *S. albus* J1074	Myronovskyi, M., et al.[2]
Streptomyces LV1-209GEK	The wild-type strain	this study.
Escherichia coli ET12567 pUB307	Donor strain for intergeneric conjugation	Flett, F. et al.[3]
Escherichia coli DH10β	General cloning strain	Grant, S.G., et al[4]
Escherichia coli GB05-red	*E. coli* strain used for the Red/ET recombination	Zhang et al.[5]

References

1. Ahmed, Y.; Rebets, Y.; Estévez, M. R.; Zapp, J.; Myronovskyi, M.; Luzhetskyy, A., Microbial cell factories 2020, 19, 1-16.
2. Myronovskyi, M.; Rosenkränzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A., Metabolic engineering 2018, 49, 316-324.
3. Flett, F.; Mersinias, V.; Smith, C. P., FEMS microbiology letters 1997, 155, 223-229.
4. Grant, S. G.; Jessee, J.; Bloom, F. R.; Hanahan, D., Proceedings of the National Academy of Sciences 1990, 87, 4645-4649.
5. Zhang, Y.; Muyrers, J. P.; Testa, G.; Stewart, A. F., Nature biotechnology 2000, 18, 1314-1317.

3. Summary and Conclusion

Antimicrobial resistance is climbing worldwide and has become a serious health crisis in the 21st century. Drug-resistant infections caused by pathogens such as methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococci* are increasingly hard to treat.[1] If no new treatment options are developed, projections warn that annual deaths could reach 10 million by 2050.[2] Finding antibiotics with novel molecular scaffolds and new modes of action is therefore of high interest. Not only for direct applications but also as lead structures for synthetic drug design. Natural products have played a central role in this search throughout history. Actinobacteria, and especially the genus *Streptomyces*, have supplied bioactive compounds or chemical backbones for over two-thirds of today's clinically used antibiotics. Recent progress in genome sequencing and bioinformatics has revealed an abundance of untapped genetic material in these organisms. This points to many silent or condition-dependent pathways that could produce previously unknown secondary metabolites.[1,3] Recent successes, such as the discovery of indanopyrrole A from a marine Streptomyces with antimicrobial properties, exemplify the untapped potential of these bacteria and the value of genome-guided NP discovery.[4-5] We combined dereplication, genome mining and heterologous expression in this study to identify new natural products with the goal to further extend the pipeline of new scaffolds. In doing so, we also linked the identified new compounds to their respective BGC and gained deeper insights into their biosynthesis, with the potential to uncover new enzymatic functions and biosynthetic mechanisms. A collection of underexplored *Streptomyces* strains was systematically analysed through metabolomic profiling and genomic analysis. From this collection, the strains *Streptomyces* sp. LV45-129 and *Streptomyces* sp. LV1-209 were prioritized based on their production of previously unreported compounds.

The uncharacterized *Streptomyces* sp. LV45-129, obtained from soil samples in Ukraine, was screened for novel NPs using dereplication techniques. This revealed the production of known compounds, including the aminoglycoside antibiotic puromycin and the macrolide pamamycin. Puromycin can cause premature termination of nascent polypeptide chains in eukaryotes and prokaryotes by mimicking the aminoacyl-tRNA and binding to the acceptor site of translating ribosomes and thus, blocking protein synthesis.[6] Pamamycins can be responsible to inhibit the growth of gram-positive bacteria and fungi.[7] In addition to these, three putatively new compounds were identified in the culture broth. Isolation and structure elucidation led to the discovery of the ichizinones, new trisubstituted members of the pyrazinone class of natural products that are structurally related to the reported pyrazinones JBIR-56 and JBIR-57.[8] Ichizinones are cyclic tetrapeptides derivatives, incorporating three proteinogenic amino acids common across all three compounds and a variable non-proteinogenic amino acids in their chain. The general sequence from the C- to the N-terminal side is L-valine, X, D-leucine and L-valine, where X represents the variable non-proteinogenic amino acid. The non-proteinogenic components were identified as 3-amino-pentanoic acid (ichizinone A), 3-amino-4-phenylbutanoic acid (ichizinone B), and

3-aminobutanoic acid (ichizinone C). Since no BGC and respective biosynthetic pathway had previously been reported for similar pyrazinones, we constructed a cosmid library of the strain to identify this. Genome mining revealed an NRPS cluster with no close homologs in existing databases. Heterologous expression of the cosmid covering this cluster led to the production of ichizinones. With the combination of enzymatic logic, solved structure and gene deletions we were able to describe the biosynthesis of the ichizinones. Our findings indicate a distinct biosynthetic mechanism, substantially diverging from that of smaller dipeptides, and suggest a potential role for a ketosynthase in amino acid extension within NRPS assembly lines. These insights expand current understanding of NRPS-PKS interactions and may inform future predictions of similar hybrid biosynthetic pathways.[8]

Bioinformatic analysis of *Streptomyces* sp. LV45-129 genome revealed a significant biosynthetic capacity with the presence of 38 predicted BGCs. The strain showed a vast variety of BGCs predicted to produce different natural product types such as terpenes, type I / type III PKS, NRPS, lassopeptides, nucleosides, siderophores and hybrid compounds of these types. Several of these BGCs corresponded to known metabolites including puromycin and pamamycin identified during the initial dereplication process. Others coded for ubiquitous Actinobacteria products like geosmin, a terpene responsible for the characteristic earthy smell of soil, and BGCs coding for secondary metabolites necessary for maintaining cell survival and adaption to environmental conditions.[9] For example, desferrioxamine B is a hydroxamate-type siderophore with the primary role to chelate iron from the environment and facilitate its uptake under iron-limiting conditions to ensure strain growth.[10-11] It was shown that deletion of this BGC in *S. albus* is lethal to the strain.[12] Clinically is this metabolite used as an iron-chelating agent for treating iron overload.[10-11] Some of the predicted BGCs were excluded from further investigations since the constructed cosmid library did not or just partially cover the whole set of biosynthetic genes. To overcome these limitations in cosmid library coverage, transformation-associated recombination (TAR) cloning in *Saccharomyces cerevisiae* was employed to reconstruct several BGCs that were fragmented across multiple overlapping cosmids.[13] This allowed the successful recovery of incomplete clusters and enabled their functional analysis alongside the BGCs fully captured on a single cosmid in heterologous expression experiments. As a result, the expression of one cluster was successful in the heterologous host *S. albus* Del14. This cluster was predicted to produce a terpene type compound, but the production level was too low to allow for isolation. Therefore, the production over the course of eight days were measured alongside feeding studies with pyruvate. Pyruvate is a key metabolic intermediate at the intersection of several central carbon metabolism pathways and plays in terpene biosynthesis a role primarily in the MEP (methylerythritol phosphate) pathway, which is one of the two main pathways for terpene precursor biosynthesis.[14-15] Thereby, we observed that the concentration of the target metabolite was the highest after five days of cultivation and the added pyruvate had no effect on the production. The idea of promoter exchange was excluded since the present genes were not orientated in one operon making this a tedious and challenging task. However, several genes coding for positive regulators were present, with

one of them belonging to the SARP family. The co-expression of this regulatory gene with the BGC increased the production level up to five-times and enabled successful isolation of mansevorone. Surprisingly, mansevorone is a pyranopyridine alkaloid rather than the predicted terpene type product of the cluster. The activation of the mansouramycin pathway in *S. albus* Del14 was not anticipated and points to complex interactions between native and foreign biosynthetic systems. When a heterologous BGC is introduced, it appears that the host can simultaneously trigger expression of own silent clusters. This suggests that transcriptional control and precursor supply are shared between the two systems and that shifts in one pathway can directly affect the other. This means, that the metabolic profile of the strain may change in ways that are not solely determined by the introduced cluster. This type of interplay is a well-recognized feature of *Streptomyces* biology and likely contributes to the chemical diversity that makes these bacteria such prolific producers of natural products.[4, 16] Mansevorone features a pyranopyridine (or azachromone) core structure, a scaffold that has primarily been accessed through synthetic routs to date. Its biosynthetic origin from a heterologously expressed gene cluster represents, to our knowledge, one of the first examples of this rare heterocyclic framework being generated enzymatically in a microbial host.[17-18]

Two compounds were discovered in the culture broth of the second strain *Streptomyces* sp. LV1-209GEK, which was also obtained from soil sampled collected in Ukraine. These compounds were identified by distinct mass signals and not present in spectral databases. Subsequent isolation and structure elucidation led to the discovery of two new cyclic depsipeptides named atrevomycin A and B. Both molecules consist of ten amino acid residues forming a macrocylic peptide ring with a fatty acid chain attached to it. The amino acid sequence was determined as follows: L-threonine, L-serine, L-leucine, L-phenylalanine, L-serine, L-threonine, L-serine, L-leucine, L-glycine and L-hydroxyphenylglycine. Genome mining revealed two distinct NRPS clusters sharing similarities with the BGCs for the depsipeptides skyllamycin, cinnapeptin and atrovimycin.[19-21] Comparative *in silico* analysis with BGCs from atrovimycin, skyllamacin and cinnapeptin suggested that these clusters are involved in the biosynthesis of atrevomycin, with one cluster likely producing the precursor for the other cluster. Bioinformatic tools such as antiSMASH and BLAST enabled annotation and functional prediction of the individual modules and tailoring domains, enabling reconstruction of the biosynthetic pathway and assignment of key enzymatic steps in atrevomycin assembly. The presence of a split biosynthetic gene cluster producing a single compound shows the dynamic nature of microbial genome evolution. The structural arrangement suggests derivation from a common ancestral cluster, followed by divergence through horizontal gene transfer, recombination, or insertion of foreign DNA such as bacteriophage DNA. Due to the potent activity of the structurally related depsipeptide atrovimycin against *Mycobacterium tuberculosis* and *Physarum polycephalum* are currently activity assays with atrevomycin underway. These tests will help to clarify whether the substitution of a single amino acid residue, glycine instead of valine, changes the activity of atrevomycin. If the compound remains active,

it could still serve as a lead for further studies. If activity is lost, the result will still be valuable since it will show how a single residue exchange can shape the structure – activity profile of this class.

In conclusion do *Streptomyces* remain a valuable source of bioactive and chemical diverse natural compounds, despite decades of research. By combining dereplication, genome mining, and heterologous expression strategies, we were able to identify novel metabolites and link them to their biosynthetic gene clusters. We identified several unusual features including fragmented biosynthetic gene clusters, uncommon metabolite scaffolds and rare catalytic functions. The pyranopyridine scaffold in mansevorone is unusual, as it has been mostly achieved through synthetic chemistry and only rarely observed in NPs. The discovery of the ichizinones and atrevomycin further reveals the hidden potential of underexplored *Streptomyces* strains. Our results also demonstrate that activating cryptic or poorly expressed pathways remains a major challenge. Achieving detectable metabolite production often depends on engineered regulatory elements and customized expression hosts. In this context does the regulatory crosstalk observed in the heterologous host *S. albus* Del14 reveal the complex interplay between native and introduced gene clusters. This shows that turning on a biosynthetic pathway is rarely a simple task and often depends on the host's metabolic environment. Even with modern genome sequencing, advanced cloning methods, and powerful bioinformatic tools, many promising BGCs remain silent in the laboratory. Our results reflect the need for expression systems that can adapt to different hosts and for careful adjustment of both regulatory and metabolic settings to switch these clusters on. In an era of rising antibiotic resistance, accessing the untapped chemical diversity stored in microbial genomes is not just valuable but urgent. The strategies we explored here illustrate a focused, genome-guided route to natural product discovery.

References

1. Alam, K.; Mazumder, A.; Sikdar, S.; Zhao, Y.-M.; Hao, J.; Song, C.; Wang, Y.; Sarkar, R.; Islam, S.; Zhang, Y., Streptomyces: The biofactory of secondary metabolites. Frontiers in Microbiology 2022, 13, 968053.
2. Walsh, T. R.; Gales, A. C.; Laxminarayan, R.; Dodd, P. C., Antimicrobial resistance: addressing a global threat to humanity. Public Library of Science San Francisco, CA USA: 2023; Vol. 20, p e1004264.
3. Jose, P. A.; Maharshi, A.; Jha, B., Actinobacteria in natural products research: Progress and prospects. Microbiological Research 2021, 246, 126708.
4. Liu, Z.; Zhao, Y.; Huang, C.; Luo, Y., Recent advances in silent gene cluster activation in Streptomyces. Frontiers in Bioengineering and Biotechnology 2021, 9, 632230.
5. Sweeney, D.; Bogdanov, A.; Chase, A. B.; Castro-Falcón, G.; Trinidad-Javier, A.; Dahesh, S.; Nizet, V.; Jensen, P. R., Pattern-based genome mining guides discovery of the antibiotic indanopyrrole A from a marine Streptomycete. Journal of Natural Products 2024, 87 (12), 2768-2778.

6. Aviner, R., The science of puromycin: From studies of ribosome function to applications in biotechnology. Computational and structural biotechnology journal 2020, 18, 1074-1083.
7. Eckert, N.; Rebets, Y.; Horbal, L.; Zapp, J.; Herrmann, J.; Busche, T.; Müller, R.; Kalinowski, J.; Luzhetskyy, A., Discovery and overproduction of novel highly bioactive pamamycins through transcriptional engineering of the biosynthetic gene cluster. Microbial Cell Factories 2023, 22 (1), 233.
8. Oberhäuser, P.; Myronovskyi, M.; Stierhof, M.; Gromyko, O.; Luzhetskyy, A., Microbial Cell Factories 2025, 24, 1-11.
9. Gerber, N.; Lechevalier, H., Geosmin, an earthy-smelling substance isolated from actinomycetes. Applied microbiology 1965, 13 (6), 935-938.
10. Bellotti, D.; Remelli, M., Deferoxamine B: a natural, excellent and versatile metal chelator. Molecules 2021, 26 (11), 3255.
11. Codd, R.; Richardson-Sanchez, T.; Telfer, T. J.; Gotsbacher, M. P., Advances in the chemical biology of desferrioxamine B. ACS Chemical Biology 2018, 13 (1), 11-25.
12. Myronovskyi, M.; Rosenkränzer, B.; Luzhetskyy, A., Iterative marker excision system. Applied microbiology and biotechnology 2014, 98, 4557-4570.
13. Kouprina, N.; Larionov, V., Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. Chromosoma 2016, 125 (4), 621-632.
14. Kumari, M.; Checker, V. G.; Kathpalia, R.; Srivastava, V.; Singh, I. K.; Singh, A., Metabolic engineering for enhanced terpenoid production: Leveraging new horizons with an old technique. Plant Physiology and Biochemistry 2024, 108511.
15. Yang, L.; Liu, H.; Jin, Y.; Liu, J.; Deng, L.; Wang, F., Recent advances in multiple strategies for the synthesis of terpenes by engineered yeast. Fermentation 2022, 8 (11), 615.
16. Zhang, X.; Zhang, F.; Li, C.; Li, J.; Xu, X.; Zhu, T.; Che, Q.; Li, D.; Zhang, G., Heterologous Expression of Type II PKS Gene Cluster Leads to Diversified Angucyclines in Streptomyces albus J1074. Marine Drugs 2024, 22 (11), 480.
17. Ye, S.; Ballin, G.; Pérez-Victoria, I.; Braña, A. F.; Martín, J.; Reyes, F.; Salas, J. A.; Méndez, C., Combinatorial biosynthesis yields novel hybrid argimycin P alkaloids with diverse scaffolds in Streptomyces argillaceus. Microbial Biotechnology 2022, 15 (12), 2905-2916.
18. Malets, Y. S.; Vashchenko, B. V.; Moskvina, V. S.; Golovchenko, O. V.; Brovarets, V. S.; Grygorenko, O. O., Parent 5 (7)-azachromones and their partially hydrogenated derivatives: synthesis and physicochemical properties. Chemistry of Heterocyclic Compounds 2023, 59 (6), 494-499.
19. Pohle, S.; Appelt, C.; Roux, M.; Fiedler, H.-P.; Süssmuth, R. D., Biosynthetic gene cluster of the non-ribosomally synthesized cyclodepsipeptide skyllamycin: deciphering unprecedented ways of unusual hydroxylation reactions. Journal of the American Chemical Society 2011, 133 (16), 6194-6205.
20. Liu, Q.; Liu, Z.; Sun, C.; Shao, M.; Ma, J.; Wei, X.; Zhang, T.; Li, W.; Ju, J., Discovery and biosynthesis of atrovimycin, an antitubercular and antifungal cyclodepsipeptide featuring vicinal-dihydroxylated cinnamic acyl chain. Organic Letters 2019, 21 (8), 2634-2638.
21. Zhang, C.; Seyedsayamdost, M. R., Discovery of a cryptic depsipeptide from streptomyces ghanaensis via MALDI-MS-guided high-throughput elicitor screening. Angewandte Chemie International Edition 2020, 59 (51), 23005-23009.