universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Discovery of novel biologically active molecules via conventional and genome-guided bioprospecting of actinomycete bacteria“

verfasst von / submitted by

Jaime Felipe Guerrero Garzón

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, 2019 / Vienna, 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 796610449
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Doktoratsstudium Pharmazie
Betreut von / Supervisor:	Univ.-Prof. Dr. Sergey B. Zotchev
Mitbetreut von / Co-Supervisor:	

Acknowledgements

The current project was carried out at the laboratory of Pharmaceutical Biotechnology, Pharmacognosy Department, University of Vienna.

My complete gratitude to Prof. Dr. Sergey B. Zotchev for helping me and encouraging me at any time to pursue this degree. Thank you for believing in me and giving me the opportunity to work on this fascinating project and offer me a place in your research group.

I would like to thank especially to Dr. Olga Sekurova for teaching me to work with Actinomycetes and for those long talks and discussions about science and more during my PhD.

Special thanks to Dr. Martin Zehl for his amazing help and expertise concerning mass spectrometry. Thank you for all the feed backs and for being involved in my project. Also, I would like to say thanks to Mag. Christina Sykora for helping me with mass spectrometry data analysis.

Thanks to all our collaborators during this project; Prof. Dr. Galina Selivanova and her research group for performing the bioassays on human cell lines, Prof. Dr. Ernest Urban and Prof. Dr. Finn Lillelund Aachmann for their help with NMR analysis and Prof. Dr. Jörn Kalinowski and his group for sequencing the genomes study in this project.

I would like to say thank you to all the group of the Pharmaceutical Biotechnology, especially to my colleagues Dr. Chris Wawrosch, Dr. Martina Oberhofer, Dr. Olha Schneider, Dr. Paulina Bekiesch, Dr. Andrea Kodym, emer. o.Univ.-Prof. Mag. Dr. Wolfgang Kubelka, Mag. Silvia Löbsch, Mag. Florian Gössnitzer, Frau Krista Kucevic, Mag. Fabian Malfent and Mag. Arthur Kaplan. All these people helped me in many ways to pave the way during my PhD. More than work colleagues, I consider you as family.

Many thanks to the Diploma students Johanna Höller and Sophie Graffius for their dedicated work and contribution to this project with the strains S1-1-ISP4-01 and RLA070 respectively.

There are not words to describe my gratitude to my lovely wife, Paulina, for trusting me and literally following me to the other side of the world, without you I could never have achieve it. You always be there for me in the good and bad times.

To my parents and sister for giving me their unconditional support. Despite the distance, I have never felt alone. Family is where the heart is.

Special thanks to Claudia, you offered me your sincere friendship and always were there to listen me.

I also express my gratitude to all those people that have already passed away, but left their finger prints in my life.

Table of contents

Abbreviations list

µF: micro Faraday
µL: microliter
µM: micromolar
A: Adenylation Domain
ACP: Acyl Carrier Protein
act: Actinorhodin
Amp: Ampicillin
Asp: Aspartic Acid
AT: Acyltransferase
ATP: Adenosine Triphosphate
BGCs: Biosynthetic Gene Clusters
bla: ampicillin resistance gene
bp: Base pair
c.f.u.: colony-forming unit
C: Condensation Domain
cda: calcium-dependent antibiotic
Clm: Chloramphenicol
cm: centimetre
CoA: Coenzyme A
COSY: Correlation Spectroscopy
cpk: coelimycin
CSCs: Cancer Stem Cells
C-terminal: Carboxyl-terminus
Cy: Heterocyclization Domain
Da: Dalton
DDA: Data-dependent acquisition
DMSO: Dimethyl sulfoxide
DNA: Deoxyribonucleic Acid
DOX: Doxorubicin
E: Epimerization Domain
EDTA: Ethylenediaminetetraacetic acid
ELSD: Evaporative Light Scattering Detector
ESI-Qq-TOF-MS: Electrospray-ionisation quadrupole time-of-flight tandem mass spectrometry
ESKAPE: Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, and Enterobacter species
F: Formylation Domain
g: gram
GC: Guanine-Cytosine
gDNA: Genomic DNA
Glu: Glutamic Acid
h: hour
HMBC: Heteronuclear Multiple Bond Correlation
HPLC-CAD/MS: High Resolution Liquid Chromatography - Collisionally Activated Dissociation/Mass Spectrometry
HPLC-MS: High Performance Liquid Chromatography - Mass Spectrometry
HRMS: High Resosutiom Mass Spectrometry
HSQC: Heteronuclear Single Guantum Coherence
Kan: Kanamycin
Kb: Kilobase
kDa: kilodalton
KS: Ketosynthase
kV: kilovolt
L: liter
LC-MS: Liquid Chromatography - Mass Spectrometry
m/z: mass-to-charge ratio
M: Methylation Domain
Mb: Mega bases
MCoA: Malonyl-CoA
mg: milligram
MICs: Minimal Inhibitory Concentrations
min: Minute
mL: milliliter
mm: millimetre
MOPS: 3-(N-morpholino)propanesulfonic acid
mRNA: Messenger RNA
ms: millisecond
MW: Molecular Weight
Nal: Nalidixic acid
ng: nanogram
nm: nanometre
NMR: Nuclear Magnetic Resonance
NPs: Natural Products
NRPs: Nonribosomal Peptides
NRPSs: Nonribosomal Peptide Synthetases
N-terminal: Amino-terminus
OD: Optical density
ORF: Open Reading Frame
OSMAC: One Strain Many Compounds
Ox: Oxidation Domain
PBS: Phosphate-buffered saline
PCD: Programmed Cell Death
PCP: Peptidyl-Carrier Protein
PKS: Polyketide Synthases
ppm: Parts per million
PTMs: Post-Translational Modifications
R: Reduction Domain
RE: Recognition Element
red: prodiginines
RiPPs: Ribosomally Synthesized and Post-translationally Modified Peptides

RNA: Ribonucleic Acid
rpm: revolutions per minute
rpoB: β subunit of RNA polymerase
rpsL: ribosomal protein S12
rRNA: Ribosomal RNA
SARP: Streptomyces Antibiotic Regulatory Protein
SDS: Sodium Dodecyl Sulfate
t(m)RNA: Transfer-messenger RNA
TAR: Transformation-Associated Recombination
Te: Thioesterase Domain
UV: Ultraviolet
V: volt
wt: Wild Type
µL: microliter
µM: micromolar
µm: micrometre
Ω: Ohm

1. Zusammenfassung

Entdeckung neuer biologisch aktiver Moleküle durch konventionelle und genomgesteuerte Bioprospektion von Aktinomyceten-Bakterien.

Actinomyceten-Bakterien sind ausgezeichnete Quellen für eine Vielzahl von biologisch aktiven Verbindungen.

Im Rahmen des aktuellen Projekts wurden mehrere von verschiedenen Herkünften isolierte Aktinomyceten zur Herstellung neuartiger bioaktiver Sekundärmetaboliten untersucht.

Zwei Ansätze wurden angewandt, um solche Moleküle zu finden und zu isolieren. Einerseits wurde der traditionelle Bioprospecting-Ansatz verwendet, indem die Stämme unter verschiedenen Wachstumsbedingungen kultiviert wurden, gekoppelt mit Bioaktivitätsprüfungen, HPLC-MS- und NMR-Analysen. Andererseits wurde ein "Genome Mining"-Ansatz in Verbindung mit gentechnischen Techniken wie der heterologen Expression von biosynthetischen Genclustern (BGCs) und der Überexpression von positiven Transkriptionsregulatoren zur Herstellung von Sekundärmetaboliten angewendet.

Die Bioprospektion von Streptomyceten, die aus einer äthiopischen Wüste isoliert wurden, ermöglichte die Reinigung von zwei antibakteriellen Anthrazyklinen, deren Strukturen durch NMR und HRMS aufgeklärt wurden. Die Analyse des Genoms dieses Bakteriums bestätigte sein Potenzial zur Biosynthese chemisch unterschiedlicher Naturprodukte und erlaubte es, bestimmte BGCs mit experimentell identifizierten Metaboliten zu verbinden.

Die Bioprospektion eines weiteren *Streptomyces* Stammes, der aus einem Tiefseesediment isoliert wurde, das im Pazifischen Ozean gesammelt wurde, ermöglichte die Identifizierung der antimykotischen Verbindung Cycloheximid und eines möglicherweise neuartigen Tripeptidsiderophors. Im Genom dieses Isolats wurden entsprechende BGCs gefunden.

Ein aus der Rhizosphäre von Edelweiß isolierter *Kitasatospora* Stamm produzierte antibakterielle Chlortetracycline zusammen mit zwei potenziell neuen Verbindungen.

Ein Stamm der Gattung *Amycolatopsis* aus der mongolischen Steppe lieferte Extrakte, die gegen ein grampositives Bakterium wirksam ist. Die bioaktive Verbindung wurde gereinigt und ihre Struktur durch NMR als 1,2,4-trimethoxynaphthalene bestimmt. Darüber hinaus wurde festgestellt, dass dieser Stamm verschiedene Tigloside produziert, die keine antimikrobielle Aktivität zeigten. Anschließend wurde das Genom dieses Stammes auf das Vorhandensein von Biosynthesegenen für Sekundärmetaboliten analysiert, wodurch mehrere einzigartige BGCs identifiziert werden konnten. Dieser Analyse folgte die Klonierung und

erfolgreiche heterologe Expression eines einzigartigen Lassopeptids BGC, wodurch zwei neuartige bioaktive Lassopeptide produziert wurden.

1. Abstract

Discovery of novel biologically active molecules via conventional and genome-guided bioprospecting of actinomycete bacteria

Actinomycete bacteria are excellent sources for a wide range of biologically active compounds.

In the current project, several actinomycetes isolated from different environments were investigated for production of novel bioactive secondary metabolites.

Two approaches were applied to find and isolate such molecules. On one hand, the traditional bioprospecting approach was used by cultivating the strains in different growth conditions coupled with bioactivity testing, HPLC-MS and NMR analyses. On the other hand, a “genome mining” approach coupled with genetic engineering techniques such as heterologous expression of biosynthetic gene clusters (BGCs) and overexpression of positive transcriptional regulators to produce secondary metabolites was applied.

Bioprospecting of *Streptomyces* isolated from an Ethiopian desert yielded the purification of two antibacterial anthracyclines, which structures were elucidated by NMR and HRMS. Analysis of this bacterium’s genome confirmed its potential to biosynthesize chemically diverse natural products, and allowed to connect certain BGCs with experimentally identified metabolites. Bioprospecting of another *Streptomyces* strain isolated from a deep-sea sediment collected in Pacific Ocean allowed identification of the antifungal compound cycloheximide and a potentially novel tripeptide siderophore. Corresponding BGCs were found in the genome of this isolate.

A *Kitasatospora* strain isolated from the rhizosphere of Edelweiß was shown to produce antibacterial chlortetracyclines along with two potentially novel compounds.

A strain of the genus *Amycolatopsis* from the Mongolian steppe yielded extracts active against a Gram-positive bacterium. The bioactive compound was purified, and its structure determined by NMR as 1,2,4-trimethoxynaphthalene. In addition, this strain was found to produce various tiglosides that did not display antimicrobial activity. Next, the genome of this strain was analyzed for the presence on secondary metabolite biosynthesis genes, allowing identification of several unique BGCs. This analysis was followed by cloning and successful heterologous expression of one unique lasso peptide BGC, yielding production of two novel bioactive lasso peptides.

2. Introduction

2.1. Microorganisms as an important source of natural products

Natural products (NPs) encompass an enormous group of structurally diverse chemical entities with a wide variety of bioactivities. They originate from bacterial, fungal, plant, marine and terrestrial animals [1].

Bérdi J. defined NPs as “chemical (carbon) compounds isolated from diverse living things” that may derive from primary or, most commonly, from secondary metabolism of living organisms. The primary metabolites (polysaccharides, proteins, nucleic and fatty acids) are common in all biological systems. The secondary metabolites are, however, low molecular weight (usually with MW<3000 Da) chemically and taxonomically extremely diverse compounds with not defined function, characteristic mainly to some specific, distinct types of organisms [2]. However, other authors defined NPs as certain secondary metabolites that are not required for the host´s survival, but which provide an advantage to the host in its native environment conditions [1].

Nevertheless, something that it is completely clear is the fact that many of the NPs, the group of small molecular secondary metabolites produced by microorganisms, normally are bioactive [1,2].

Since the beginning of human history, microorganisms have provided abundant sources of NPs that have found multiple uses, importantly for human medicine and agriculture. Many of them have been developed as commercial products, such as antibiotics, anti-cancer agents, immunosuppressants and pesticides with a huge market [1].

It is estimated that nearly 50% of all new approved drugs are either NPs or NP-based [3].

The most relevant medicinal uses of NPs has been in the area of anti-infectives, notably anti-bacterial and anti-fungal therapy [1], and cancer.

2.2. Antibiotics

In 1947, Selman Waksman published a definition of “antibiotic”: "An antibiotic is a chemical compound, produced by microorganisms, which has inhibits the growth of and even to destroy other bacteria and or microorganisms” [4].

It is known that ancient civilizations like Greeks, Egyptians, Incas, Chinese, Indians and Persians used a variety of naturally available treatments for infection, for example herbs, honey, molds and even animal faeces. This empirical knowledge was preserved through generations as part of our cultural heritage [4,5].

During the “Golden Age” of antibiotic discovery, after the discovery of penicillin, many microorganisms were isolated, especially from soil samples collected all around the world. It is estimated that tens of millions of soil microorganisms were screened [6], an enormous effort that provided the clear majority of microbial metabolites known today [2,7].

Antibiotics target bacterial vital cellular functions (*Fig. 1*), causing growth inhibition or microbial death. Classical antibiotic targets are the cell wall synthesis (β-lactams and glycopeptides; vancomycin produced by *Amycolatopsis orientalis*), protein biosynthesis via interaction with ribosomal subunits (aminoglycosides; streptomycin produced by *Streptomyces griseus*, and macrolides; erythromycin produced by *Saccharopolyspora erythraea*), DNA replication (aminocoumarins; produced by *Streptomyces niveus*), RNA synthesis (rifampicin produced by *Amycolatopsis rifamycinica*) metabolic pathways interference (folic acid antagonists; sulphonamides and trimethoprim) or bacterial membrane structure disruption (lipopeptides; daptomycin produced by *Streptomyces roseosporus*) [8–10].

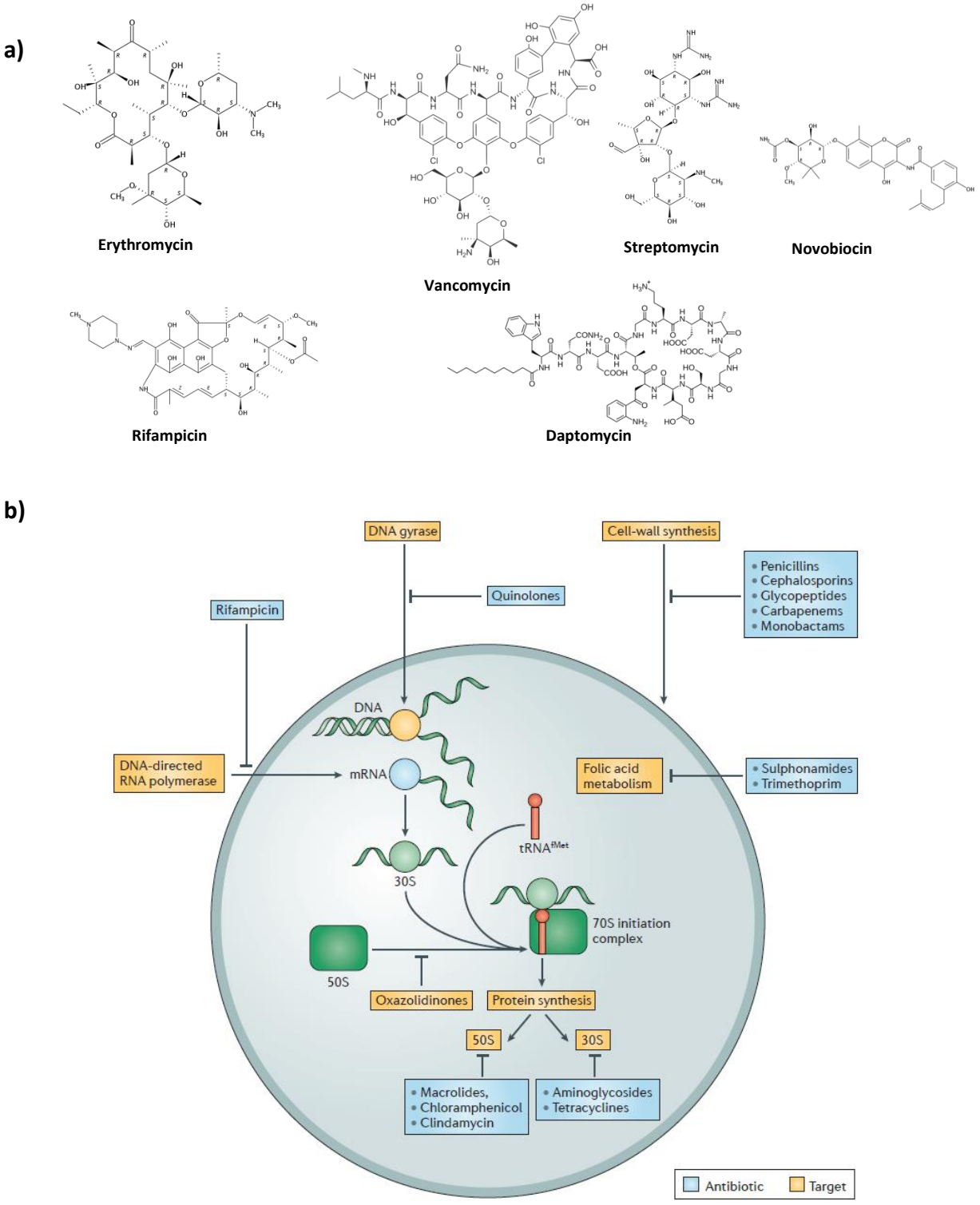


Figure 1. a) Chemical structures of different antibiotics produced by Actinomycete bacteria. b) Common targets of antibiotics. Antibiotics can interfere with the DNA replication, synthesis of RNA, protein synthesis, cell wall, cytoplasmic membrane or metabolic pathways. Adapted from Lewis, 2013.

2.2.1. Antibiotic resistance

Millions of lives have been saved since antibiotics were introduced into medical practice. Life expectancy increased from 47 years in 1900 up to 74 years for males and 80 years for females in the year 2000 in the United States [11].

However, the biggest advantage bacteria have over antibiotics is their incredible adaptability to changes in their environment [7].

Mass production and uncontrolled application of broad spectrum antibiotics induced defense mechanisms of bacteria and resulted in the development and spread of antibiotic resistance. Microorganisms replicate and mutate rapidly, exchanging their genetic material, which results in a wide range of acquired resistance. The number of multi-resistant pathogens is constantly increasing. Furthermore, the emergence of big cities and globalization are two main additional factors that intensify the spread of bacterial infections [7].

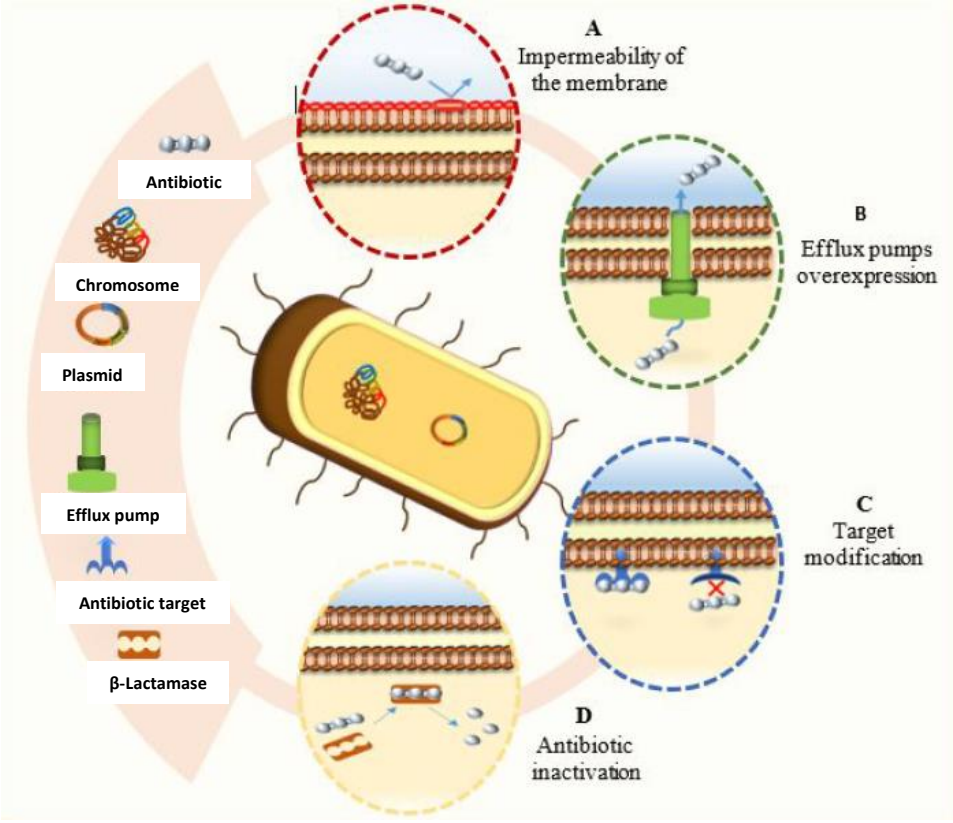


Figure 2. Bacterial antibiotic resistance mechanisms. A) Reduction of the membrane permeability prevents the entry of the antibiotic. B) Overexpression of pumps that expel the antibiotic outside of the cell. C) Modification of the antibiotic's target makes it unrecognizable. D) Enzymatic inactivation of the antibiotic and loss of its activity. Adapted from Douafer et al., 2019.

Antibiotic resistance may occur via four main mechanisms (*Fig. 2*) such as: a) target modification through a mutation or chemical modification, b) decreased permeability of the bacterial cell to the antibiotic, which is the most common cause of intrinsic resistance, c)

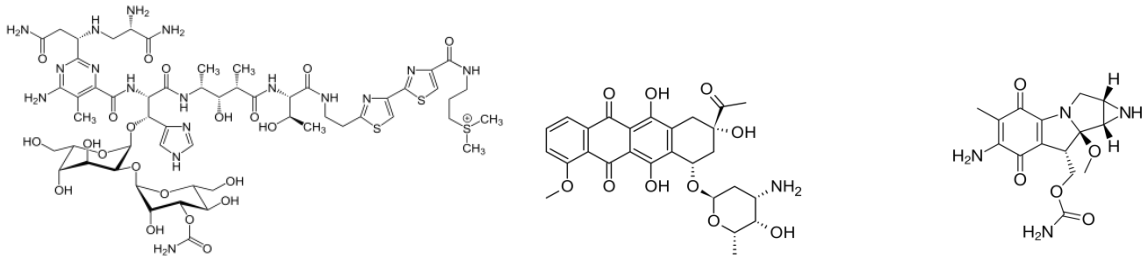
reduction of the antibiotic influx or increase of its efflux, to prevent it from reaching its target in sufficient quantity, and d) deactivating processes which degrade the antibiotic such as enzymatic hydrolysis or any other chemical modification that reduces the affinity of the antibiotic for its target [10,12].

Multidrug resistant bacterial pathogens known as ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* spp) are a serious global health threat because they can cause untreatable infectious diseases [13]. According to recent review, it is estimated that the number of deaths worldwide caused by infectious diseases will increase from 700,000 to 10 million annually by 2050 [14]. Hence, urgent action in terms of novel antibiotic discovery and strict regulation of antibiotic use are required in order to decrease this threat.

2.3. Cancer drug resistance

Cancer is the second leading cause of mortality worldwide. It was estimated that only in the year 2018, 18.1 million of new cancer cases and 9.6 million cancer deaths occurred. It has been shown that cancer incidence and mortality are rapidly growing around the world. The reasons are complex but they are mainly related to aging and growth of the global population. Unfortunately, cancer is a very complex disease which diagnosis and treatment are challenging [15,16].

Chemotherapy is one of the principal ways of treating cancer [17]. Several antitumor agents are bacteria-derived and currently used in clinics (*Fig. 3*): bleomycins (produced by *Streptomyces verticillus*) are used in combination with other drugs to treat several kinds of cancers. Their mechanism of action is through DNA cleavage [18]. Daunorubicin (produced by *Streptomyces peucetius*) is used in chemotherapy as the drug that interacts with DNA by intercalation and inhibits the enzyme topoisomerase II [19,20]. Mitomycin C (produced by *Streptomyces caespitosus*) is another example of antitumor drug that has been used in combinatorial chemotherapy in clinics for a variety of tumors. Mitomycin C is a potent DNA crosslinker [21].



Bleomycin A2 **Daunorubicin** **Mitomycin C**

Figure 3. Chemical structures of different antitumor compounds produced by Actinomycete bacteria.

Unfortunately, drug resistance is a limitation of chemotherapy effectiveness. The mechanisms of resistance to cytotoxic chemotherapeutics can be developed through alterations in the drug target, activation of pro-survival pathways and/or ineffective induction of cell death. Resistance can pre-exist in the tumor cells before the treatment or acquired during the chemotherapy [17]. Therefore, new anticancer agents with different mechanisms of action must be developed in order to generate novel chemotherapies with decreased chance to develop resistance. One strategy is by targeting cancer stem cells (CSCs), which comprise the subpopulation of tumor bulk and are known to acquire resistance to conventional therapies (chemotherapy and radiotherapy). It is suggested that CSCs stay inactive during chemotherapy and thus they can restart cancer development after the therapy is finished. Dormancy is a state in which tumor cells remain undetectable during the tumor progression or after the treatment. Furthermore, CSCs are considered as the primary tumor initiator cells and one of the main causes of tumor relapse and metastasis [22,23]. Therefore, developing anti-cancer drugs that can help in targeting CSC remains a challenge and a priority.

2.4. Actinomycete bacteria

Actinomycete bacteria encompass a huge and heterogeneous group of Gram-positive bacteria with a high GC genomic DNA content and include at least 350 genera [24]. They constitute one of the largest bacterial phyla and are ubiquitous in soil environments as well as in marine and fresh water ecosystems. They can survive in association with non-living organic substrates or interact with partners of higher eukaryotes. Free-living actinobacteria are not frequent [25].

Actinomycete bacteria usually grow by combination of tip extension and branching of the hyphae, therefore having a complex morphological differentiation. For many years, actinomycetes were considered transitional organisms between fungi and bacteria, because many of them produce mycelium and look like filamentous fungi, and some reproduce by sporulation. However, like all bacteria, actinomycetes' cells carry a chromosome that is organized in a prokaryotic nucleoid. They also contain a peptidoglycan cell wall. Most *Actinobacteria* are aerobic (with exceptions), can be heterotrophic or chemoautotrophic, but most are chemoheterotrophic and able to use different nutritional sources [26].

The filamentous *Actinobacteria*, which belong to the family Actinomycetaceae, are frequently called actinomycetes. They are highly versatile NPs producers, biosynthesizing two-thirds of all known antibiotics, many anticancer, antifungal and immunosuppressive agents. These bacteria are not only important for human health, but also agriculture and biotechnology [25,26].

By far, the best characterized genus of the Actinomycetaceae is *Streptomyces*, responsible for production of ca 50% of all known antibiotics and many other valuable NPs. However,

less studied so-called “rare actinomycetes” can be sources of unique secondary metabolites as well. Those actinomycete bacteria are more difficult to isolate compared to the members of genus *Streptomyces* [24,25].

2.4.1 Life cycle of actinomycetes

Actinomycetes have a wide variety of morphologies, differing mainly with respect to the presence or absence of a substrate mycelium or aerial mycelium, the color of the mycelium, the production of pigments, and the structure and form of their spores. However, the actinomycete bacteria with primarily mycelia lifestyles usually reproduce by forming asexual spores [26].

Understanding the life cycle of those microbes has been a main task for researches, not only to learn more about their biology and physiology, but also because of its link to the secondary metabolite production.

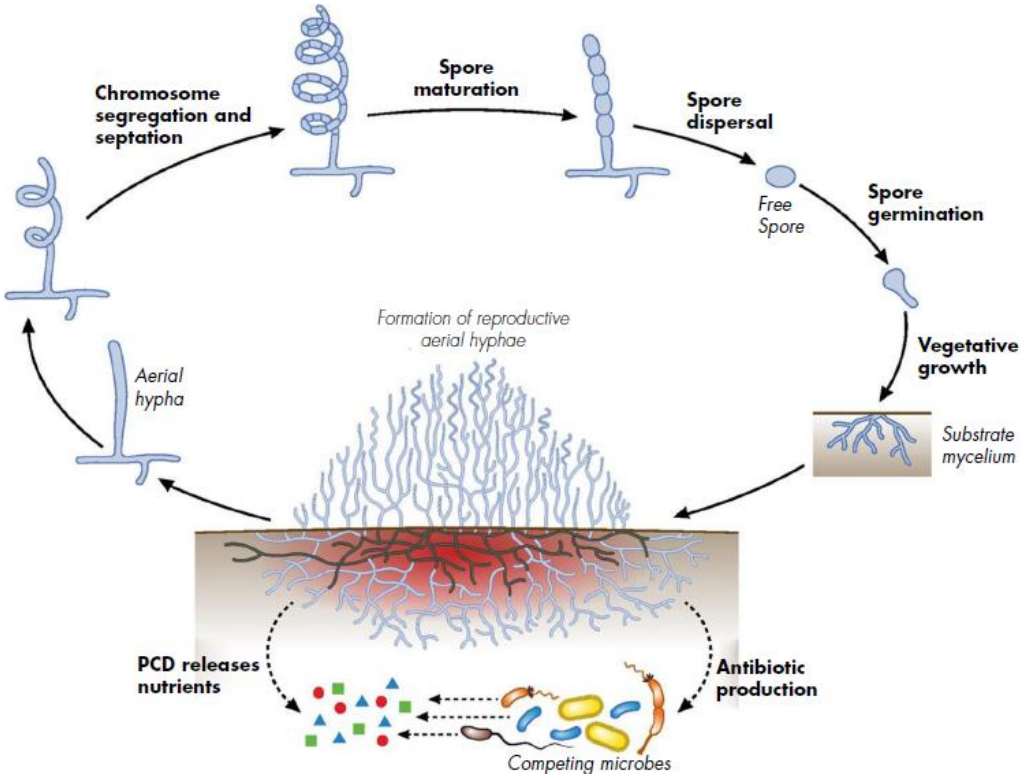


Figure 4. Life cycle of *Streptomyces* genus. After *Streptomyces* spores germinates, subtract mycelium is form. When environmental conditions are no longer optimal for growing (nutrients depletion), part of the mycelium will be degraded through programmed cell death (PCD) releasing nutrients that are going to be used for the formation of aerial hypha to start a new cycle of sporulation. Furthermore, during this step, antibiotic are produced to defend the nutrients from competing microbes. Adapted from van der Meij *et al.*, 2017.

In the case of the *Streptomyces* genus (*Fig. 4*), the life cycle starts when a spore germinates in favorable conditions, and develops into hyphae. Then hyphae grow by branching and tip extension, forming the substrate mycelium. The release of exo-enzymes allows breaking of

chitin and cellulose that may be present in the substrate to provide nutrients [25]. When nutrients become scarce, a part of the substrate mycelium is sacrificed through autolytic degradation via programmed cell death (PCD)**.** The release of nutrients is used for the formation of aerial mycelium, which starts to replicate extensively DNA and triggers cell division to form chains of spores, where each spore contains a single chromosome [27–29].

The initiation of sporulation coincides with secondary metabolites production, especially antibiotics. It is believed, that it provides protection against competing microorganisms attracted by the nutrients released during PCD [25].

2.4.2. Genus *Streptomyces*

Streptomyces is a diverse genus of actinomycetes with around 600 described species [25] and the best studied *Actinobacteria* genus. They are abundant and important in the soil, where they play a key role in the carbon cycle. Furthermore, they are source of diverse bioactive secondary metabolites with great relevance for medicine, biotech industry and agriculture (*Fig. 5*) [26]. The variety of chemical entities produced by this genus includes non-ribosomal peptides, polyketides, lactams, lactones, terpenes, indoles, quinones, etc. [30].Generally, the genome size of *Streptomyces* species is 8 Mb on average, and it is well known that they possess many genes involved in secondary metabolites biosynthesis [9,31].

This genus yielded many important compounds in several antibiotic classes: macrolides (tylosin, spiramycin); aminoglycosides (neomycin, kanamycin); β-lactams (cephamycin, carbapenems); tetracyclines (tetracycline, chlortetracycline, oxytetracycline); polyenes (candicidin, amphotericin B, nystatin); peptides (actinomycin); and chloramphenicol. Other type of substances isolated from *Streptomyces* species that are also in clinical use include antifungals (amphotericin B, nystatin), antiparasitics (avermectin), antitumorals (daunorubicin) and immunosuppressives (rapamycin) [9,26].

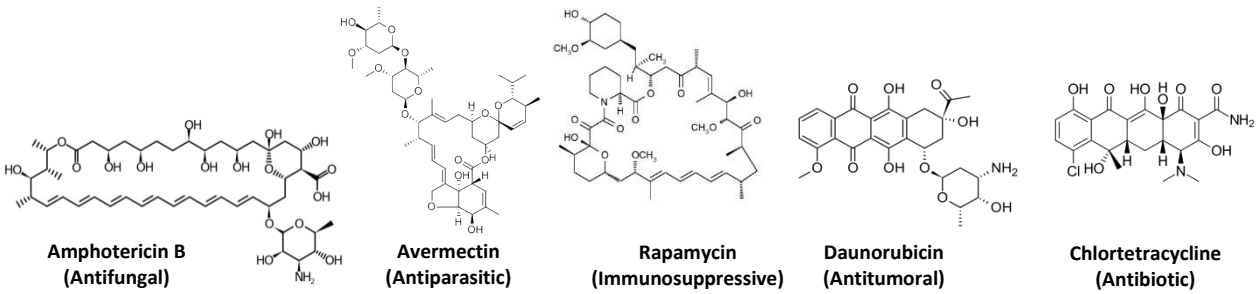
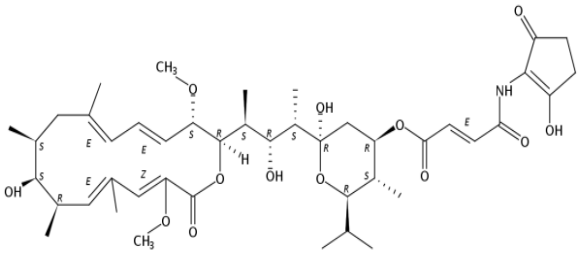


Figure 5. Bioactive compounds produced by *Streptomyces* species.

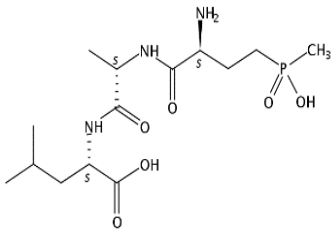
2.4.3. Genus *Kitasatospora*

For many years the strains of *Kitasatospora* were considered a part of the genus *Streptomyces*. However, Zhang *et al*. concluded that the *Kitasatospora* is a legitimate separate genus on the basis of phenotypic and genetic differences from *Streptomyces* [32]. Currently, 31 species have been validly proposed for this genus and the genomes of 15 strains were sequenced [33]. Normally, the genome size of *Kitasatospora* is about 7-9 Mb and contain many genes governing secondary metabolites [34].

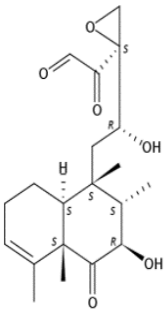
Several novel NPs have been discovered from genus *Kitasatospora* members, including compounds with unique chemical structures and interesting bioactivities (*Fig. 6*). For example, the nematocidal and antifungal compound setamycin, the herbicidal phosalacine, the antitumor compound terpentecin and the immunomodulator kifunensine [34].



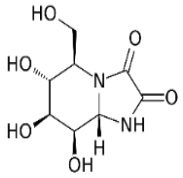
Setamycin
(Nematocidal, antifungal)



Phosalacine
(Herbicidal)



Terpentecin
(Antitumoral)



Kifunesine
(Immunomodulator)

Figure 6. Bioactive compounds produced by *Kitasatospora* species.

2.4.4. Genus *Amycolatopsis*

The genus *Amycolatopsis* is ubiquitously distributed and many members have been isolated mostly from soil, but also from aquatic environments, rock surfaces, and from clinical sources [35].

Initially, many members of the genus *Amycolatopsis* were classified as *Streptomyces* then shifted to *Nocardia* and finally a new genus *Amycolatopsis* was established that included those species in which mycolic acid was absent in their cell wall [36]. Today, 76 different *Amycolatopsis* species have been validly described [33]. *Amycolatopsis* species have comparatively large genomes (5-10 Mb) and there are publicly available as complete or draft versions 41 genome sequences representing 28 different species [35,37].

Amycolatopsis is a genus of special importance for its capacity to produce several medicinally important antibiotics (*Fig. 7*), such as balhimycin, vancomycin and rifamycin, and other metabolites with antibacterial, antifungal, antitumoral or antiviral properties such as as quartromicin, echinosporin, chelocardin, kigamicin and the macrotermycins A-D [35,38].

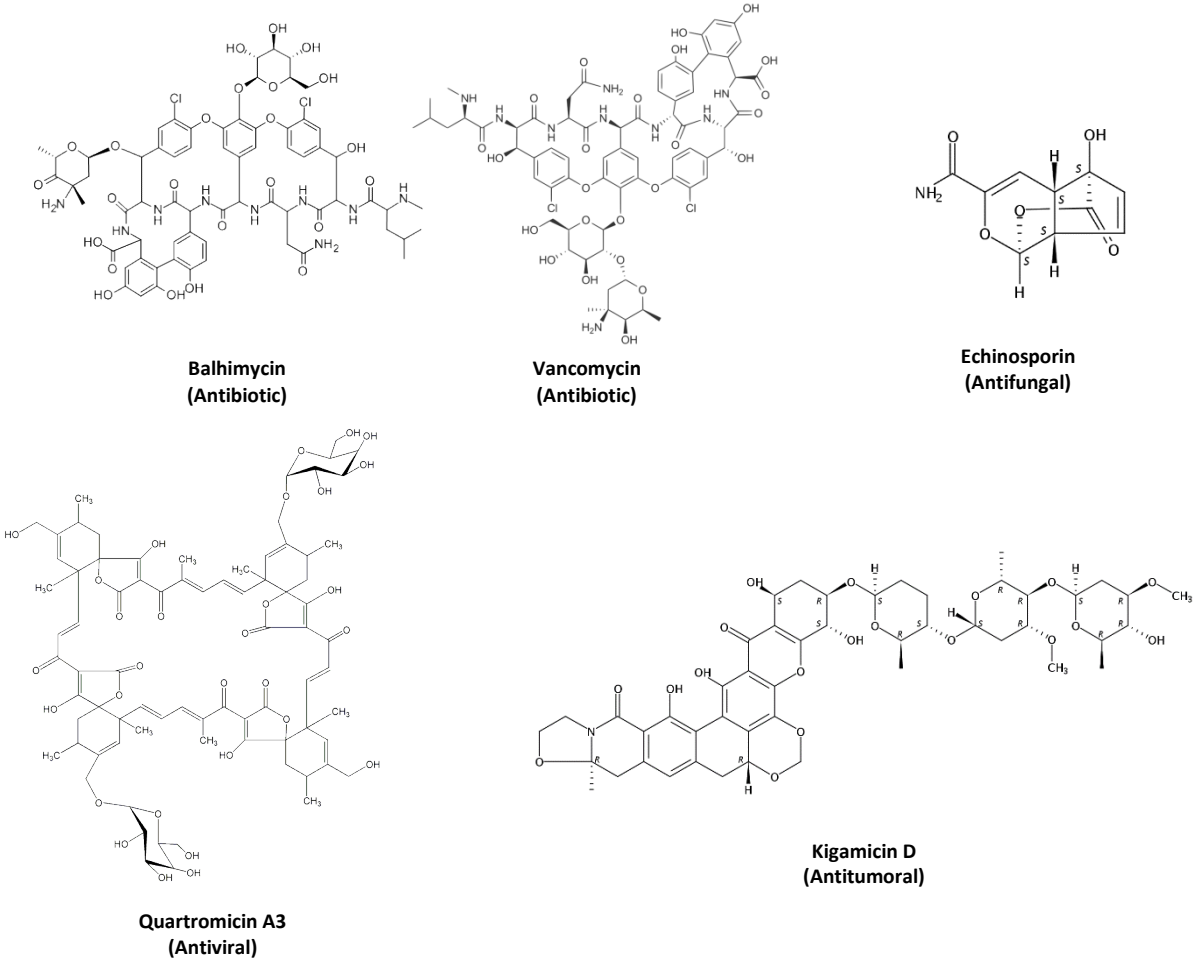


Figure 7. Bioactive compounds produced by *Amycolatopsis* species.

2.5. Biosynthetic gene clusters in Actinomycete bacteria

Most of the secondary metabolites are assembled by specific enzymes that usually use precursors coming from the primary metabolism. The biosynthesis in most of the cases has two stages. First stage included assembly of molecule scaffolds. In the second stage, these scaffolds are modified by specific enzymes that add chemical groups via glycosylation, methylation, acylation, hydroxylation, etc., and generating the final compound [39].

A biosynthetic gene cluster (BGC) is a group of genes that are clustered in a particular site of a genome that encode biosynthetic enzymes for the production of a specialized molecule and its chemical derivatives [40]. A typical BGC (*Fig. 8*) harbors genes encoding different types of enzymes. Normally, in the central part of the cluster, genes encoding enzymes responsible for scaffold biosynthesis are located. In their vicinity, genes for scaffold modification enzymes are usually found. It is also common to find in the BGCs some pathway-specific regulatory genes. Many BGCs also contain genes encoding enzymes that inactivate the endogenously accumulated final metabolite or modify the metabolite target within the producing organism in order to confer resistance to its own product (antibiotics). Another strategy to confer resistance to its own product is through the presence of genes encoding efflux pumps that export the molecules outside the cells and release them to the environment [39].

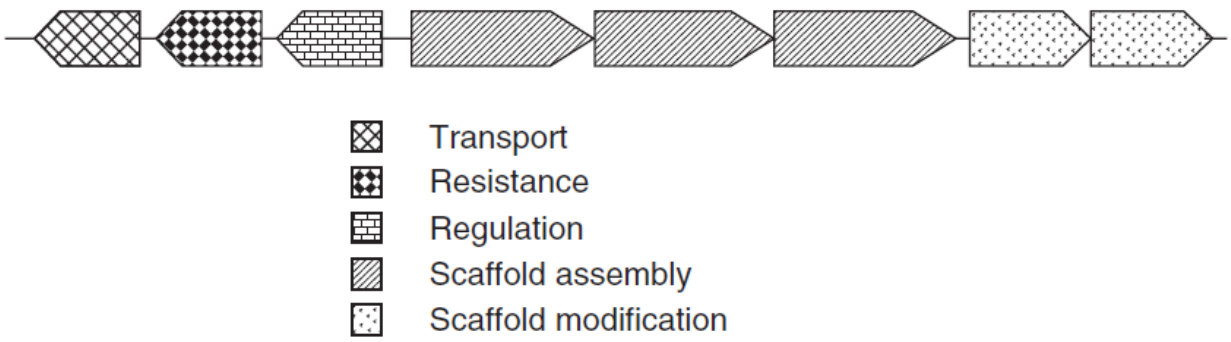


Figure 8. Organization of a common biosynthetic gene cluster (BGC). We can find in a BGC genes involved in different tasks; scaffold assembly, scaffold modification, regulation, resistance, transport. Adapted from Zotchev, 2008.

An important breakthrough in the history of secondary metabolites produced by microorganisms was achieved in 2002 via genome sequencing and analysis of actinomycete bacterium *Streptomyces coelicolor* A3(2) [41]. It was found that *S. coelicolor*, one of the first sequenced microbes, contains many more BGC encoding NP-like biosynthetic pathways than the known NPs produced by the organism under laboratory growth conditions. Since then, similar observations have been reported for many diverse sequenced

microorganisms, especially in the genomes of actinomycete bacteria where BGC are abundant [42,43].

Actinomycete bacteria harbor an extraordinary biosynthetic machinery in their BGC and therefore have the potential to produce an amazing diversity of NPs, including polyketides, pyrones, peptides, siderophores, γ-butyrolactones, butenolides, furans, terpenoids, fatty acids, oligopyrroles, etc. [44].

2.5.1. Polyketide synthases

Polyketides encompass a large class of compounds produced by bacteria, fungi and plants that shows different chemical structures. These metabolites vary in molecular weight, chemical moieties, and could be linear, polycyclic, or macrocyclic chemical structures. Currently, polyketide from microbial origin find clinical use as antibiotics (tetracycline and erythromycin), antiparasitic agents (avermectin), antifungals (amphotericin B), anticancer drugs (daunorubicin), and immunosuppressants (rapamycin) [26,45,46].

The polyketide biosynthesis occurs through repetitive decarboxylative Claisen thioester condensation reactions of activated acyl starter units, such as acetyl-coenzyme A (CoA) and malonyl-CoA (MCoA), with MCoA, producing a chain extension every time that a unit is added. The chain elongation depends on at least three enzymatic functions: a ketosynthase (KS), an acyltransferase (AT), and an acyl carrier protein (ACP). A set of those domains constitute a module, and generally, each module is responsible for only a single elongation cycle [45,47].

Three main types of bacterial polyketide synthases (PKS) are known to date, regarding architecture and functionality [46,47].

PKSs type I are large multifunctional enzymes which contain linearly arranged and covalently fused catalytic domains. Two different types of PKSs type I are known: modular type I PKSs, that act non-iteratively (present in bacteria), and iterative type I PKSs (mainly found in fungi and in some bacteria). The terms “iterative and non-iterative” refer to the ability of the KS domain to catalyze more than one round of elongation (iterative) or not (non-iterative). Reduced polyketides (macrolides, polyethers and polyenes) such as erythromycin A and amphotericin B are PKS I products [47,48].

PKSs type II are multienzyme complexes found exclusively in bacteria that constitute a single set of iteratively acting enzymes. Aromatic polyketides (often polycyclic) such as tetracenomycin C are PKS II products [49].

PKSs type III, also known as chalcone synthase-like PKS, are homodimeric enzymes that are iteratively acting condensing enzymes. Aromatic polyketides (often monocyclic or

bicyclic) such as flavolin are PKS III products [50]. PKS III have long been found in plants; lately related enzymes have also been discovered in bacteria and fungi [47].

2.5.2 Nonribosomal peptide synthetases

Nonribosomal peptide synthetases (NRPSs) constitute a unique class of multi-modular enzymes that produce peptides independently of the ribosomes [51,52]. Nonribosomal peptides (NRPs) are produced by many bacteria, fungi, and lower eukaryotes.

Some NRPs of microbial origin have different medical applications, such as anti-bacterials (vancomycin, teicoplanin), immunosuppressive (cyclosporine A), or antitumor agents (bleomycin A2) [53].

The sequence of amino acids in the peptide is determined by the order of catalytically active entities within NRPSs [51].

NRPs biosynthesis generally occurs in three main phases. First, building-block (amino acids) assembly. Second, NRPS-mediated peptide assembly. Third, post-NRPS modification and decoration. Generally, NRPS enzymes use proteinogenic amino acids and/or non-proteinogenic amino acids for the biosynthesis of NRP molecules [53].

The modules are the catalytic entities responsible for the incorporation of a distinct amino acid into the peptide. The minimal organization of these catalytic units consists of a condensation (C) domain, an adenylation (A) domain, and a peptidyl-carrier protein (PCP). Each module activates and binds a single building block to a growing peptide chain. The domain A performs the selection, activation and loading of the amino acid onto the domain PCP. The tethered amino acid is then shuttled to the domain C, where coupling to the upstream nascent peptide chain is established. Another possibility is that the building block could be shuttled to different domains, either modified in the respective module by domains or tailoring enzymes that could lead further modifications of the peptide, for example, epimerization (E), formylation (F), methylation (M), heterocyclization (Cy), reduction (R), and oxidation (Ox). Finally, a thioesterase (Te) domain disconnects the mature peptide from the NRPS machinery and often mediates macrocyclization during this release step [53–55].

2.5.3. Ribosomally synthesized and post-translationally modified peptides

Ribosomally synthesized and post-translationally modified peptides (RiPPs) constitute an interesting class of structurally diverse natural products. RiPPs are produced by members of the archaea, bacteria, and eukaryote domains and their biosynthetic genes are present in the available sequenced genomes [56,57].

Unlike NRPS, which require large multimodular enzyme complexes to incorporate non-proteinogenic amino acids into a peptide backbone[51,52], RiPPs are capable of accessing a similar degree of chemical diversity through extensive post-translational modifications (PTMs) of a ribosomally synthesized precursor peptide [58]. Therefore, RiPPs have a wide variety of structural features conferring them diverse biological activities such as antifungal (pinensins), antibacterial (microcin J25), and antiviral (labyrinthopeptin A1) [57,59]. Some RiPPs are currently used in clinics or are in clinical trials, for example, the lanthipeptide duramycin for treatment of cystic fibrosis, the thiopeptide LFF571 to treat *Clostridium difficile* infection, and the Ziconotide, a conotoxin peptide used for neuropathic pain treatment [60].

The biosynthesis of this class of compounds starts with the synthesis of a precursor peptide. Normally, RiPPs precursors contain an N-terminal leader region and a C-terminal core region respectively. The biosynthetic machinery recognizes the leader peptide, which is removed by one or more peptidases while the core peptide undergoes PTMs to ultimately yield the mature RiPPs. In some cases, additional PTMs occur after removal of the leader peptide [57,59].

RiPPs are classified into more than 20 different subfamilies based on their biosynthetic machinery and structural features [56].

2.5.3.1 Lasso peptides

Lasso peptides are NPs of the RiPPs type usually produced by bacteria.

The lasso peptides molecules normally have an average size range of 15–24 amino acid residues (molecular weights of 1500–2500 Da) and have a unique topology. They consist of a macrolactam ring comprised of 7 to 9 residues, where the amino group of the N-terminal residue reacts with the carboxyl side chain of a Glu or Asp to form a macrolactam ring, and a linear C-terminal peptide tail which is trapped within the ring either by bulky side chains (steric effect) or by one or two disulfide bonds, or by both. This lasso topology makes lasso peptide molecules extraordinarily stable [61–63].

Depending on the number of disulfide bridges, the lasso peptide family is subdivided into three subtypes, class II (without), class III (one) or class I (two) [62].

Generally, a core lasso peptide BGC (*Fig.* 9a) encodes a precursor peptide (A) that contains an N-terminal leader and a C-terminal core region sequence, a cysteine protease (B), a macrolactam synthase (C) and a RiPP recognition element (RE) that binds to the leader peptide sequence (E). Alternatively, many clusters can encode fused RE-B proteins (E+B) [64,65]. However, in some cases lasso peptide clusters can harbor additional genes that encodes for ABC transporters (D), isopeptidase or additional modification enzymes [66].

The biosynthesis of lasso peptides (*Fig. 9b*) was first understood with the lasso peptide archetype microcin J25 produced by *Escherichia coli*. In this model, the precursor peptide A is processed by the mature cysteine protease B that cleaves off the leader peptide and releases the prefolded core peptide. Next, the macrolactam synthetase C catalyzes the activation of the carboxyl side chain of the Glu/Asp residue allowing the macrocyclization reaction, generating the mature lasso peptide [67,68].

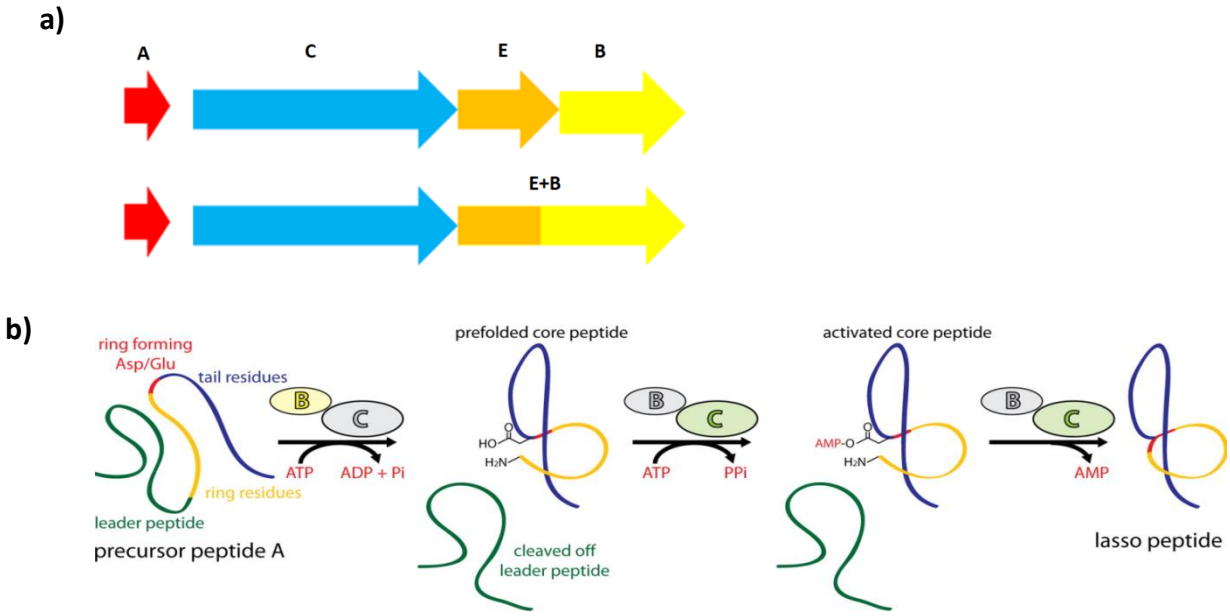


Figure 9. a) Core lasso peptide BGC. In this kind of cluster we can mainly find 3 classes of genes; A encodes for a precursor peptide, B encodes a protease and C encodes a macrolactam synthase. **b) Biosynthesis of lasso peptides.** The protease (B) cleave off the leader peptide from the precursor peptide (A), then the macrolactam synthase (C) catalyzes the formation of the macrolactam ring and assist the folding of the mature lasso peptide. Adapted from Mevaere et al., 2018 and Hegemann *et al.*, 2015.

Lasso peptides exhibit a versatile array of interesting biological activities, such as receptor antagonists (anantin produced by *Streptomyces coerulescens*), enzyme inhibitors (sungsanpin produced by *Streptomyces* sp.) and/or antibacterial (sviceucin produced by *Streptomyces sviceus*) or antiviral (siamycin produced by *Streptomyces* sp.) [62]. It is worthy to point out that most of the bioactive lasso peptides are isolated from *Actinobacteria*, and

regarding antibacterial activities, those molecules show a narrow-spectrum and frequently possess novel modes of action [65].

2.6. Bioprospecting of actinomycete bacteria

During the “golden age” of antibiotic discovery (1950s-1970s), it is estimated that tens of millions of soil microorganisms were screened, yielding most of the microbial secondary metabolite producers known today, with nearly 70 % of antibiotics being produced by *Streptomyces* species. Later, the significance of the non-*Streptomyces* actinomycete species (rare actinomycetes) in this regard was recognized. However, in the recent years the ratio of new actinomycete compounds had definitely decreased [2,7,31].

One of the major problems with bacterial bioprospecting was that some metabolites were very often rediscovered. This led to the assumption that microorganisms as sources for new therapeutics have been exhausted [69,70].

However, there are hopes for a “renaissance” in NPs discovery. It has been recognized that after a century of developing cultivation techniques, we are still unable to cultivate the vast majority of environmental bacteria. Given that industrial screening for antibiotics had involved pure cultures of microorganisms, it is clear that only a small fraction of the microbial world had been surveyed for bioactive secondary metabolites production. Furthermore, it also became evident that even in those microorganisms that were amenable to cultivation many compounds were not produced under routine screening conditions. Therefore, within the last 50 years, a minimal number of soil actinomycetes and their bioactive metabolites have been sampled [31,71].

The use of specific enrichments and selections can reduce the burden of screening by avoiding common actinomycetes; this approach might help to harness new microbial biodiversity by targeting bacteria from taxa that have not been studied extensively. Alternatively, another strategy might be the screening of complex or remote ecological sites including deep ocean, deserts, high mountains as well as plant endophytes, animal symbionts or human microbiota. Nevertheless, it still unknown if those unexplored taxa in terms of biosynthetic potential would be sufficiently different and could produce novel bioactive compounds [31,69–71].

2.7. Genome mining

Most of the known bioactive metabolites were discovered through empirical bioprospecting screening programs. In most cases, this process has been driven either by bioactivity guided fractionation of crude fermentation broth extracts, or via chemical screening (isolation of chromatographically resolvable metabolites with “interesting” spectroscopic properties). However, this strategy has a main disadvantage, since it selects mostly those compounds that had either a high potency and/or were produced in significant amounts. Furthermore, it is completely empirical, because is not possible to predict which kind of molecule would be produced [69,72].

The recent advances in DNA sequencing coupled with bioinformatics tools such as antiSMASH, PRISM or NaPDoS, have led to the rapid determination of bacterial genome sequences and the easy identification of BGCs [73–76].

It became quite clear that actinomycete genomes contain a large number of BGCs and therefore possess the potential to produce novel secondary metabolites [41–43]. However, under normal laboratory conditions only a small number of those compounds are produced in noticeable quantities, the rest of the BGCs are considered “silent” because they yield undetectable amounts of compounds and it is unknown which conditions trigger their expression [77,78].

Genome mining is a new research field that encompasses bioinformatics, molecular genetics, and NP analytical chemistry to produce secondary metabolites of BGCs found in the genome sequence of a certain organism [79].

Genome mining allows that new molecules could first be bioinformatically identified from de genomic DNA sequences. The latest might enable to choose a combination of targeted detection methods or simply screening different cultivation conditions that would eventually lead to identification of the predicted compound. Furthermore, it allows to plan genetic engineering strategies in order to activate silent BGCs and decreases considerably the chance of re-discovery of already known compounds [69,70,80].

Finally, another strategy to do genome mining is through the gene inactivation in a particular BGC predicted to be involved in the production of a novel secondary metabolite. Then, a comparative metabolite profiling of the wild type and mutant strains using HPLC-MS could lead to the identification of the compound encoded by this particular BGC [70].

2.8. Strategies for activation of silent biosynthetic gene clusters

During the last decades different strategies to trigger the production of cryptic metabolites in microorganisms have been developed. Many of those approaches enabled the purification and identification of novel compounds. Those strategies can be classified in two main categories; physiological triggers, and genetic engineering tools [77,78].

The OSMAC (One Strain Many Compounds) approach is traditionally used in NPs discovery, and is based on the fact that a single organism is capable of producing a diverse array of secondary metabolites. However, those are produced only under certain environmental conditions. Therefore, OSMAC encompasses the variation of different culture parameters, including nutritional components, pH, culture aeration, temperature etc. Even though OSMAC is not targeted to the activation of a specific BGC, the systematic alteration of culturing parameters could trigger the activation of random BGCs [81]. Nowadays, with the use of genome mining, it is possible to predict the structural and physico-chemical properties of some cryptic secondary metabolites and therefore tailor the cultivation conditions to produce the desired compound. An example of the latest was the discovery of the siderophore compound coelichelin from *Streptomyces coelicolor*. Originally, by examining the genome of this organism, a BGC was detected and suspected to be responsible for production of a tripeptide. The structure of the compound was suggested, allowing to speculate that it might be an iron siderophore. Based on this hypothesis, a medium depleted in iron was rationally selected to induce the production of this compound [69,70,78,82]. It has also been demonstrated that the use of chemical elicitors (antibiotics, salts of rare elements, etc.) and/or histone deacetytales inhibitors (sodium butyrate, valproic acid, etc.) can trigger the activation of BGCs [78,83,84].

Traditionally, bacteria were cultivated in axenic conditions, inevitably missing out many of the signals that trigger the production of NPs in the original habitat. In this sense, co-cultivation in the presence of other bacteria or eukaryotic microorganisms has revealed a new universe of NPs that were not observed in pure cultures [25,71].

Genetic modification of bacterial strains can be used as an alternative way to activate cryptic BGCs or to improve the titers of their cognate secondary metabolites. One of those approaches is based on the use of mutagenic compounds or radiation to induce non-specific mutations in the transcriptional (RNA polymerase) and/or translational (ribosome) machineries to generate mutants that could eventually trigger the overproduction of a novel compound. Although it is also possible to make mutations by using molecular biology tools [78].

Most of the BGCs contain pathway-specific regulatory genes encoding transcriptional activators and/or repressors. Therefore, the expression of a BGC can be manipulated by constitutive overexpression of key positive regulatory genes or by disruption of negative

regulatory genes. Examples of positive regulators are the members of the *Streptomyces* antibiotic regulatory protein (SARP) family and the large ATP-binding regulator of the LuxR family. On the other hand, TetR and GntR families are examples of transcriptional repressors. Furthermore, some regulator families, like LysR and MarR, can be either transcriptional activators or repressors [1,70,85,86]. Additionally, there exist global regulators, which are found normally outside the BGC and affect certain metabolic pathways involved in different cellular functions [78,87].

One of the main criteria that must be considered before selecting the optimal genetic engineering tool for the activation of certain BGC is whether the targeted microorganism is genetically amenable and/or can be cultivated at all. A strategy to circumvent the latest would be through the use of heterologous expression of an interesting BGC. However, the heterologous expression of BGC in an "unnatural host" can be challenging. The expression of a BGC relies on numerous cellular processes: transcription, translation, protein folding, post-translational modifications, metabolic pathways, immune mechanisms etc. All of those are highly dependent on the genetics and physiology of the producing organism. Thus, the use of host species as closely related as possible to the organism from which the BGC to be expressed was isolated is optimal.

Therefore, the heterologous expression of BGCs isolated from actinomycetes bacteria should be ideally done in another actinobacteria. In this regard, several *Streptomyces* strains have been engineered to be used as heterologous hosts. Many of those hosts have been engineered by eliminating competing endogenous pathways via deletion of corresponding BGCs. Therefore, some original secondary metabolites produced by the host will not be produced anymore, thus precluding channeling away precursors needed for the biosynthesis of an exogenous compound [70,79,88].

An engineered *Streptomyces* host example is *Streptomyces coelicolor* M1154, which was obtained from *S. coelicolor* M1146, an engineered strain from which four antibiotic BGCs were deleted (actinorhodin (*act*), prodiginines (*red*), coelimycin (*cpk*), and the calcium-dependent antibiotic (*cda*)). *S. coelicolor* M1154 additionally contains two point mutations into the genes *rpoB* (β subunit of RNA polymerase) and *rpsL* (ribosomal protein S12). Therefore, strain M114 possesses a simplified metabolic profile, and the production of heterologous metabolites is usually much higher in this host than in the parent strain M1146. Different BGCs encompassing a wide range of structural classes (RiPPs, NRPS, PKS, etc.) have been successfully heterologously expressed in strain M1154 [79,88].

Finally, in parallel with the development of heterologous hosts, different methods for cloning of entire BGCs have emerged, for example transformation-associated recombination in yeast (TAR) and Gibson ligation [89,90]. Large BGCs can be assembled either from several overlapping PCR products or from restriction fragments isolated from a genomic library

encompassing the entire BGCs. It is estimated that DNA synthesis will become cheaper in the future. Therefore, it will be an affordable alternative of DNA assembly technique [43,70].

3. Aims of the study

During the last decades the increases of antibiotic resistance among pathogens and cancer drug resistance have become two of the main health concerns worldwide. Therefore, new antibiotics and antitumor compounds are urgently needed.

Actinomycete bacteria have shown to be an extraordinary source of bioactive metabolites. However, high rates of compounds re-discovery along with other factors led many pharmaceutical companies to abandon their programs on natural products research.

The advent of the genomic era demonstrated that actinomycetes possess in their genomes the potential to produce previously undiscovered molecules. Furthermore, it is well known that the vast majority of microorganisms have not been cultivated yet. The latter may suggest that actinomycete bacteria from unique environments can still be attractive in terms of drug discovery.

In this project, four actinomycete strains belonging to three different genera were studied: two *Streptomyces* strains, one *Amycolatopsis* strain and one *Kitasatospora* strain. The four actinomycete bacteria were isolated from unique environments, including Ethiopian deserts, rare earth mine, marine sediments and rhizosphere of Edelweiß.

The main goal of this project was to discover novel bioactive compounds from these actinobacteria. In this sense, two approaches were selected.

On one hand, the classical bioprospecting approach was used. Most of the time, the production of secondary metabolites requires certain unknown environmental triggers. Therefore, the microorganisms must be cultivated under different growth conditions in order to find the best condition for production of bioactive molecules. On the other hand, the genomes of the actinomycete strains can be sequenced, followed by the genome mining approach to identify unique biosynthetic gene clusters (BGCs). Selected BGCs then must be activated either through the manipulation of cluster-specific transcriptional regulators or via heterologous expression. The cultures generated by both approaches must then be extracted with suitable organic solvents and antimicrobial activity tested using paper disk diffusion assay. The bioactive compounds shall then be purified and characterized by means of HPLC, HRMS and NMR, followed by bioactivity testing.

4. Results

4.1. *Streptomyces* sp. Go-475

4.1.1. Prospecting for bioactivities from *Streptomyces* sp. Go-475

Strain Go-475 was isolated from a soil sample collected in the desert of Gode region in Ethiopia. It was identified as a member of the genus *Streptomyces* using 16S rRNA gene sequencing.

Originally, Go-475 was cultivated on solid medium and its methanolic extract displayed antifungal activity and antibiotic activity against Gram-positive bacteria.

Strain Go-475 was cultivated in liquid medium 5288 and extracted with methanol. Interestingly, only bioactivity against the Gram-positive bacterium *Bacillus subtilis* was detected.

An upscaled batch of bioactive methanol extract was produced and fractionated using Flash chromatography (*Fig. 10*) under standard conditions (see Materials and methods section).

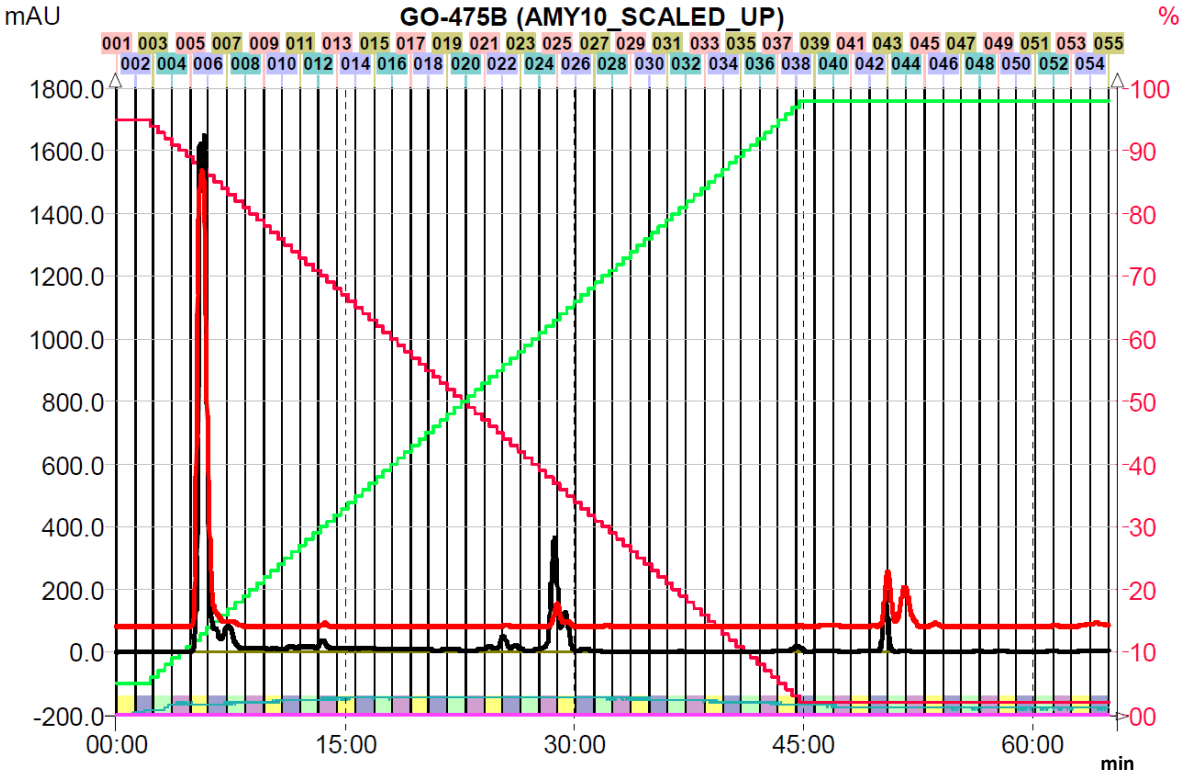


Figure 100. Go-475 Flash chromatography chromatogram. 55 fractions were obtained after Flash chromatography. 190 nm UV detector (black). ELSD detector (red). Methanol gradient (green). Water gradient (red).

The fractions were tested against *B. subtilis*, *Escherichia coli* and *Saccharomyces cerevisiae* by paper disc diffusion assay. Only fractions 24 and 42 showed bioactivity against *B. subtilis* (*Fig. 11*). However, neither fraction was bioactive against *E. coli* or *S. cerevisiae*.

The bioactive fractions were analyzed using HPLC-CAD/MS and high resolution LC-MS as described in Material and methods section. Fraction 42 contained several (branched) fatty acids and not further analyzed. In fraction 24, on the other hand, only one major constituent was detected.

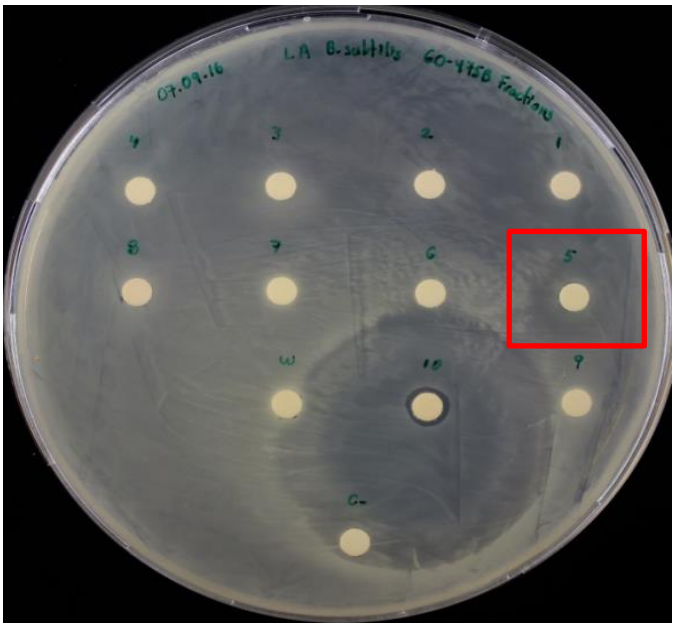
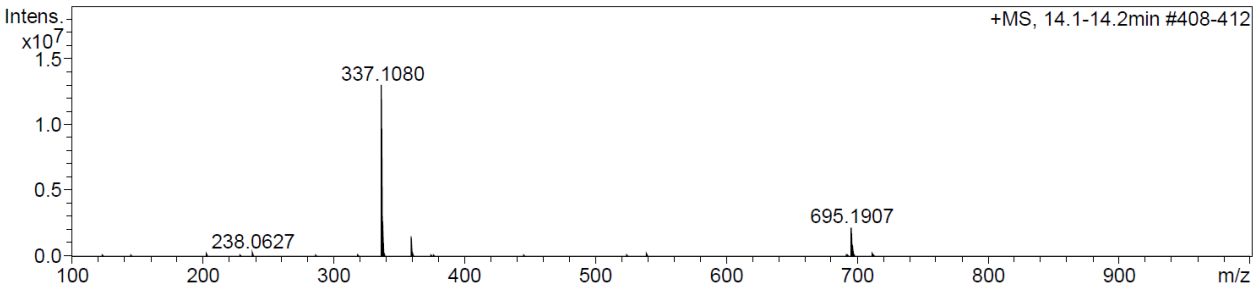


Figure 11. Bioassay of Go-475 fractions against *B. subtilis*. The 55 fractions were pooled into 11 fractions for bioassays. F5, in red color, is the original fraction 24 from Flash chromatography. F10 is the original fraction 42 from Flash chromatography.

High-resolution ESI-Qq-TOF-MS spectra, obtained by direct injection of fraction 24, showed the main compound as $[M+H]^+$ ion at m/z 337.1073 (*Fig. 12*) (calculated for $C_{20}H_{17}O_5^+$, m/z 337.1071, Δ = 0.8 ppm) and as $[M+Na]^+$ ion at m/z 359.0891 (calculated for $C_{20}H_{16}O_5Na^+$, m/z 359.0890, Δ= 0.2 ppm). The HRMS/MS-spectra of the $[M+H]^+$ ion were not very informative (*Fig. 13*), suggesting the presence of a stable aromatic systems, like in angucycline antibiotics. All these data would fit with the known compound 8-O-methyltetrangomycin ($C_{20}H_{16}O_5$). However, it could not be confirmed because there are no published MS/MS spectra of this compound.



Meas. m/z	#	Ion Formula	m/z	err [ppm]	mSigma	# mSigma	Score	rdb	e¯ Conf	N-Rule
337.1080	1	C20H17O5	337.1071	-2.8	8.6	1	98.04	12.5	even	ok

Figure 12. HPLC-HRMS positive mode data of Go-475 fraction 24.

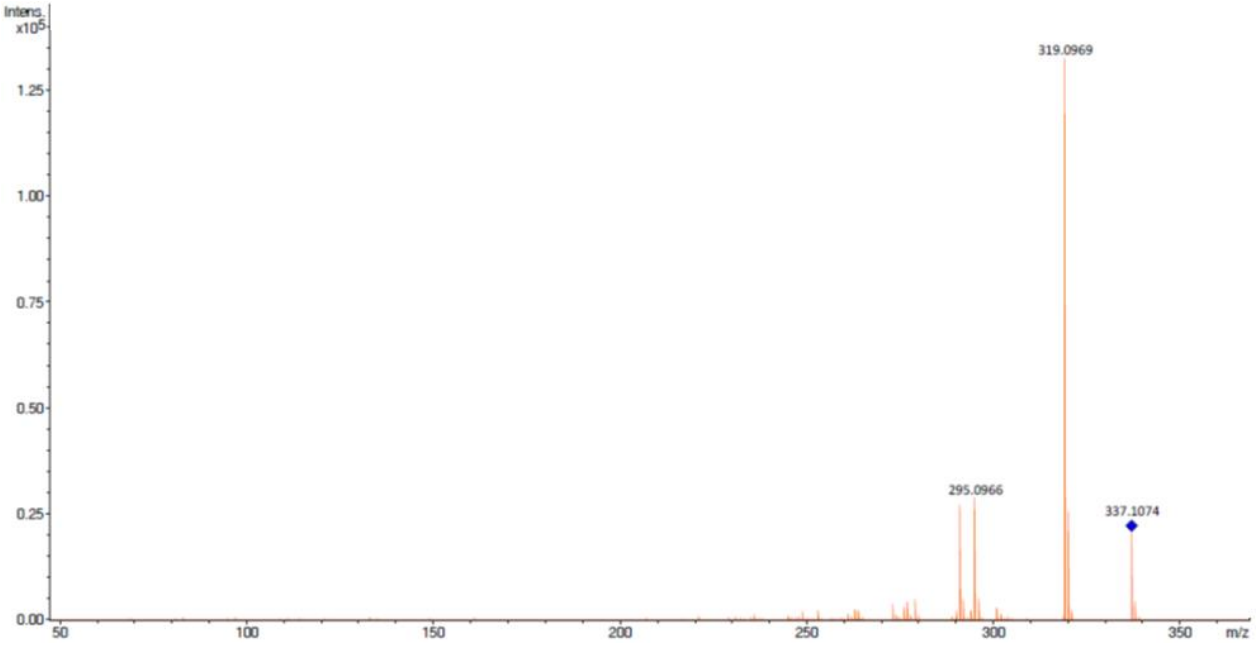


Figure 13. MS/MS spectrum of the $[M+H]^+$ ion of 8-O-Methyltetrangomycin (6-deoxy-8-O-methylrabelomycin). At m/z 337.1074 obtained on an ESI-Qq-TOF mass spectrometer with a collision energy of 25 eV.

The structure was confirmed by 1D and 2D NMR spectroscopy. When NMR analysis was performed in deuteromethanol, spectra characteristics for 8-O-methyltetrangomycin (1) were obtained. However, when the analysis was done using deuterochloroform, there was a complete conversion of the compound to 8-O-methyltetrangulol (2) (*Fig. 14*). It is suggested that traces of hydrochloric acid, present in the second solvent, catalyzed the conversion reaction. Assignment of the signals was confirmed by COSY, HSQC, and HMBC measurements. The structures of both compounds were confirmed when compared with data reported in the literature[91,92].

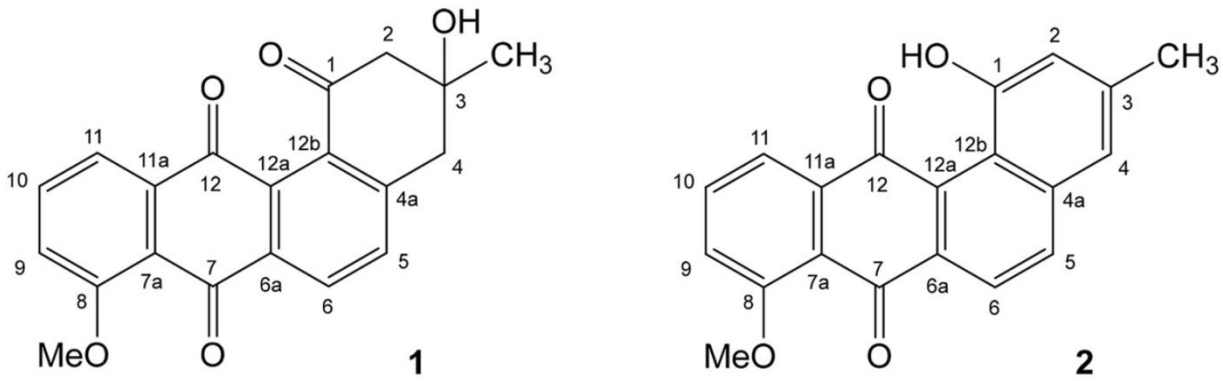


Figure 14. Chemical structures of 8-O-methyltetrangomycin (1) and 8-O-methyltetrangulol (2) produced by *Streptomyces* sp. Go-475.

4.1.2. Genome analysis

A complete genome sequence of strain Go-475 was obtained. The genome consists of a linear chromosome of 8,570,609 bp, has 71.96% G+C content, encodes 83 t(m)RNA genes, six *rrn* operons, 19 types of RNAs, and has unusually short inverted terminal repeats of 6,902 bp.

Interestingly, the Go-475 genome showed significant similarity to *Streptomyces olindensis* DAUFPE 5622 genome (draft quality sequence).

36 BGCs were identified in the Go-475 genome sequence by using antiSMASH 4.02 followed by manual curation (*Table 1*).

27 of those BGCs were present in the genomes of both *S.* sp. Go-475 and *S. olindensis*. *S. olindensis* was isolated from a soil sample collected in Brazil, and was shown to produce cosmomycins, glycosylated anthracyclines with cytotoxic activity[93].

No	Cluster type	Presence in another bacterium	Putative product
1	PKSI-NRPS	*Streptomyces olindensis* DAUFPE 5622	Glycosylated PK-NRS peptide hybrid
2	Melanin	Multiple *Streptomyces* spp.	Melanin
3	Nrps	*Streptomyces olindensis* DAUFPE 5622	NRS peptide
4	Unknown	*Streptomyces olindensis* DAUFPE 5622	Amino acid-based product
5	Terpene	*Streptomyces olindensis* DAUFPE 5622	Putative sesquiterpene
6	Lantipeptide	*Streptomyces viridochromogenes* DSM 40736	SapB/AmfS family lantipeptide
7	PKSIII	*Streptomyces* sp. XY006	1,3,6,8-tetrahydroxynaphthalene
8	Lassopeptide	*Streptomyces olindensis* DAUFPE 5622	Lassopeptide
9	Ectoine	*Streptomyces olindensis* DAUFPE 5622	Betaine
10	Terpene	*Streptomyces olindensis* DAUFPE 5622	Putative isoprenoid
11	PKSI-NRPS	*Streptomyces chartreusis* NRRL 12338	PK-NRS peptide hybrid
12	NRPS	*Streptomyces olindensis* DAUFPE 5622	Glycosylated NRS peptide
13	Terpene	*Streptomyces olindensis* DAUFPE 5622	Isoprenoid
14	Melanin	*Streptomyces* sp. S10(2016)	Melanin
15	Lasso peptide	*Streptomyces chartreusis* NRRL 12338	Lasso peptide
16	Siderophore	Multiple *Streptomyces* spp.	**Desferrioxamine B**
17	Butyrolactone	*Streptomyces olindensis* DAUFPE 5622	Butyrolactone
18	PKSI-NRPS	*Streptomyces olindensis* DAUFPE 5622	PK-NRS peptide hybrid
19	Phosphonate	*Streptomyces olindensis* DAUFPE 5622	Phosphonate metabolite
20	Terpene	-	Terpenoid
21	Aminocyclitol	*Streptomyces olindensis* DAUFPE 5622	Aminocyclitol
22	NRPS	*Streptomyces vitaminophilus* DSM 41686	NRS peptide
23	PKSII	*Streptomyces olindensis* DAUFPE 5622	**Cosmomycins**
24	Terpene	*Streptomyces olindensis* DAUFPE 5622	Albaflavenone
25	Terpene	-	Terpenoid (polyprenyl)
26	Siderophore	*Streptomyces olindensis* DAUFPE 5622	Siderophore
27	Bacteriocin	*Streptomyces olindensis* DAUFPE 5622	Bacteriocin
28	Terpene	Multiple *Streptomyces* spp.	**Geosmin**
29	PKSI-NRPS	*Streptomyces africanus* DSM 41829	PK-NRS peptide hybrid
30	Siderophore	*Streptomyces olindensis* DAUFPE 5622	Siderophore
31	PKSII	-	**Anthraquinones**
32	Terpene	*Streptomyces olindensis* DAUFPE 5622	Hopanoids
33	PKSI	*Streptomyces olindensis* DAUFPE 5622	PK
34	Bacteriocin	Multiple *Streptomyces* spp.	Bacteriocin
35	Lantipeptide	Multiple *Streptomyces* spp.	SapB/AmfS family lantipeptide
36	PKSI	*Streptomyces olindensis* DAUFPE 5622	PK

Table 1. BGCs identified in the genome of *Streptomyces* sp. Go-475 with antiSMASH 4.02 and manually curated. Products in bold text have >90% identity of the gene product with known clusters. In white color those BGCs which are present in both *S*. sp. Go-475 and *S. olindensis*. In gray color those BGCs which are present only in *S*. sp. Go-475.

By comparing the genomes, cluster 23 from *S.* sp. Go-475 was predicted to be responsible for the production of cosmomycins. Additionally, clusters 9 and 28 are most probably governing production of ectoine and geosmin, respectively.

Gene	Putative product	Putative function
orf1	TetR family transcriptional regulator	Efflux regulation
orf2	MFS family transporter	Compounds' efflux
orf3	Response regulator, LuxR family	Regulation of biosynthesis
orf4	Acetyl-CoA carboxylase	Biosynthesis of malonyl-CoA
orf5	Polyketide cyclase	Cyclization of the polyketide chain
orf6	Beta-ketoacyl synthase	Polyketide biosynthesis
orf7	Beta-ketoacyl synthase CLF	Polyketide chain length factor
orf8	Acyl carrier protein	Polyketide biosynthesis
orf9	3-oxoacyl-ACP reductase	1st reductive step in angucycline biosynthesis
orf10	Polyketide cyclase	1st cyclization/dehydration
orf11	acyl-CoA carboxylase subunit beta	Biosynthesis of malonyl-CoA
orf12	acyl-CoA carboxylase subunit epsilon	Biosynthesis of malonyl-CoA
orf13	Monooxygenase FAD-binding	Oxygenation at C12?
orf14	Monooxygenase FAD-binding	Oxygenation at C12?
orf15	SDR oxidoreductase	unknown
orf16	MFS family transporter	Compounds' efflux
orf17	Antibiotic biosynthesis monooxygenase	unknown
orf18	O-methyltransferase	Conversion of Tetrangomycin to 8-O-methyltatrangomycin
orf19	Antibiotic biosynthesis monooxygenase	unknown
orf20	SDR oxidoreductase	unknown
orf21	Antibiotic biosynthesis monooxygenase	unknown
orf22	S-adenosylmethionine synthetase	Provision of a methyl donor for O-methyltransferase (ORF18)

Table 2. Annotation of cluster 31 from *Streptomyces* sp. Go-475. BGC responsible of angucyclines biosynthesis.

Cluster 31 was identified as the most likely candidate to be responsible for the biosynthesis of the angucyclines (*Table 2*). It contains genes encoding PKS type II and other modification enzymes. Apparently, this cluster is absent in the *S. olindensis* genome.

The BGC for both angucyclines were already described and the core genes for this class of compounds previously identified [94,95]. However, the encoding genes from cluster 31 showed a homology lower than 80% compared with characterized clusters. Furthermore, cluster 31 contains additional genes which are absent in other angucycline BGCs. Therefore, it might

be that 8-O-methyltetrangomycin represents precursor for more complex angucycline molecules [96].

4.1.2.1. Proposed biosynthesis of angucyclines in Go-475

A biosynthetic pathway to produce 8-O-methyltetrangomycin was proposed based on the cluster 31 annotation and current knowledge on angucyclines biosynthesis (*Fig. 15*).

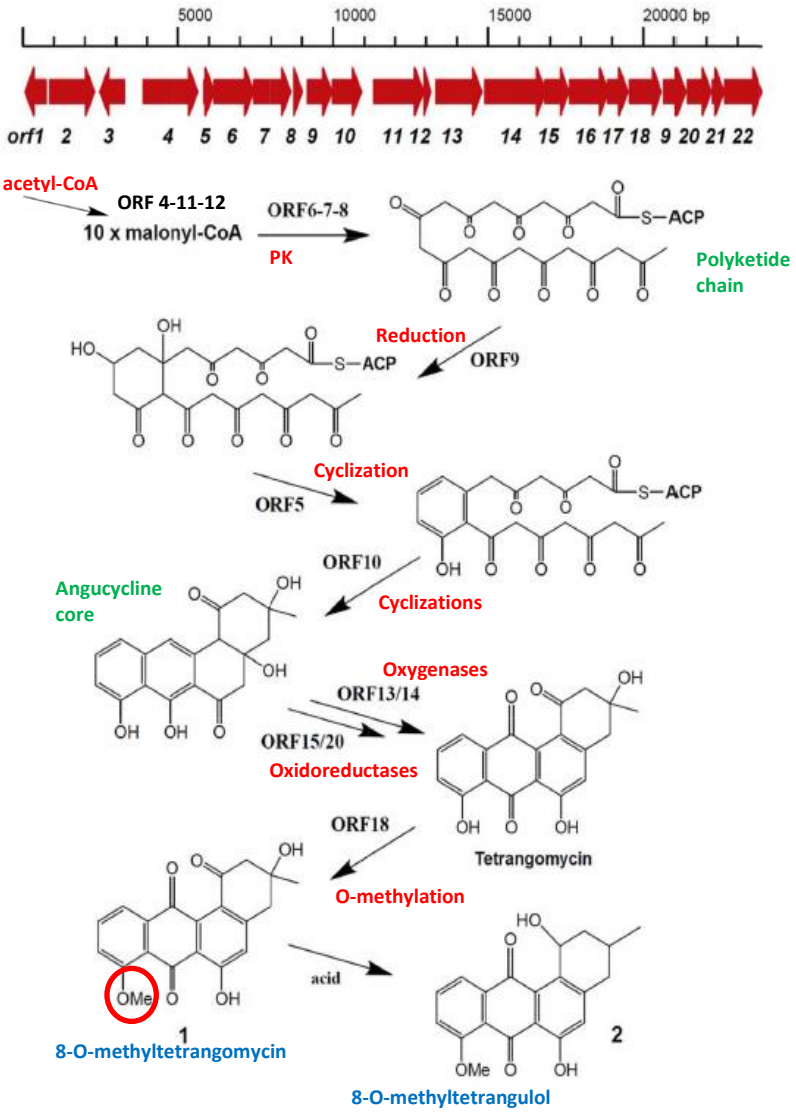


Figure 15. Proposed biosynthesis of angucycline and suggested 8-O-methyltetrangomycin. Adapted from Kibret et al., 2018.

It is suggested that enzymes encoded by ORF4-11-12 through carboxylation of acetyl-CoA would produce malonyl-CoAs, the building block molecules for angucyclines biosynthesis. The PKSs II (ORF6-7-8) biosynthesize a polyketide chain through decarboxylative condensation of the acetate units. ORF9 performs a step of reduction and ORF5 the

formation of the first aromatic ring. The cyclase ORF10 makes the further cyclization steps until the angucycline core is formed. Further reactions performed by oxygenases ORF13-14 and oxidoreductases ORF15-20 would produce the known compound tetrangomycin. Finally, ORF18 would catalyze an O-methylation at position C8 using S-adenosylmethionine, hypothetically produced by ORF22 product, to generate 8-O-methyltetrangomycin. It is suggested that an acidic environment would trigger the conversion of 8-O-methyltetrangomycin to 8-O-methyltetrangulol.

4.2 *Streptomyces* sp. S1-1-ISP4-01

4.2.1 Prospecting for bioactivities from *Streptomyces* sp. S1-1-ISP4-01

Strain S1-1-ISP4-01 was isolated from a deep-sea marine sediment collected in the North Pacific Ocean. It was identified as a member of the genus *Streptomyces* by using 16S rRNA gene sequencing.

S1-1-ISP4-01 was cultivated in different liquid media, and methanolic extract from GYM medium exhibited bioactivity against the yeast *S. cerevisiae*.

An attempt was made to establish a gene transfer system in this *Streptomyces* strain, which was found to be sensitive to antibiotic apramycin. Therefore, different plasmids carrying apramycin resistance gene were used to conjugate them into S1-1-ISP4-01 strain. For one of the plasmids, pSOK806 [97], transconjugants were obtained. However, when the recombinant strain S1-1-ISP4-01(pSOK806) was cultivated under the same conditions as the wild type (wt), the bioactivity against yeast disappeared (*Fig. 16*).

Figure 16. Bioassay of strain S1-1-ISPA-01 and pSOK806 mutant against *S. cerevisiae*. 1, wt. 2, pSOK806. 3, GYM medium control. 4, methanol blank.

The methanol extracts from both strains were analized by HPLC and two peaks were identified in the wt strain extract, which were absent in the S1-1-ISP4-01(pSOK806) strain and medium control extracts (*Fig. 17*).

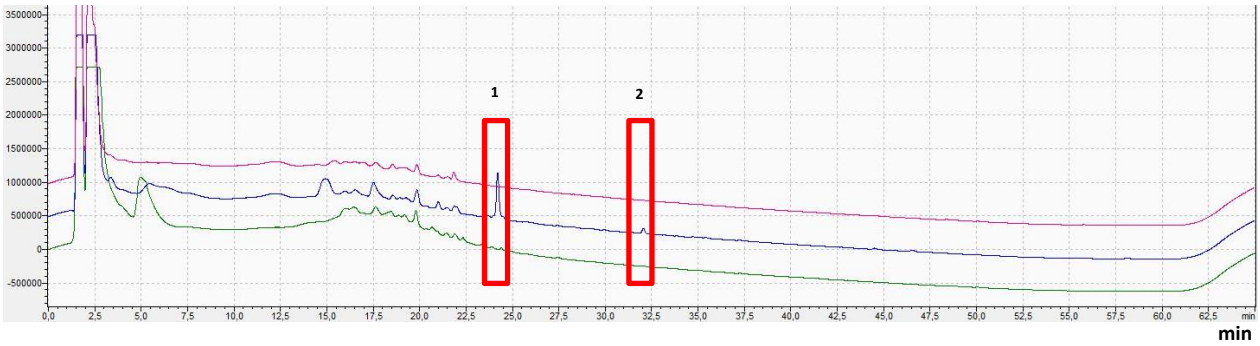


Figure 17. HPLC chromatograms of strain S1-1-ISPA-01 and S1-1-ISP4-01(pSOK806). Pink color: S1-1-ISP4-01(pSOK806) strain methanol extract. Blue color: S1-1-ISP4-01 wt strain (bioactive) methanol extract. Green color: GYM medium methanol extract. In red boxes: peak 1

Bioactive methanol extract from S1-1-ISP4-01 wt strain was fractionated by Flash chromatography (*Fig. 18*) under standard conditions (see Materials and methods section).

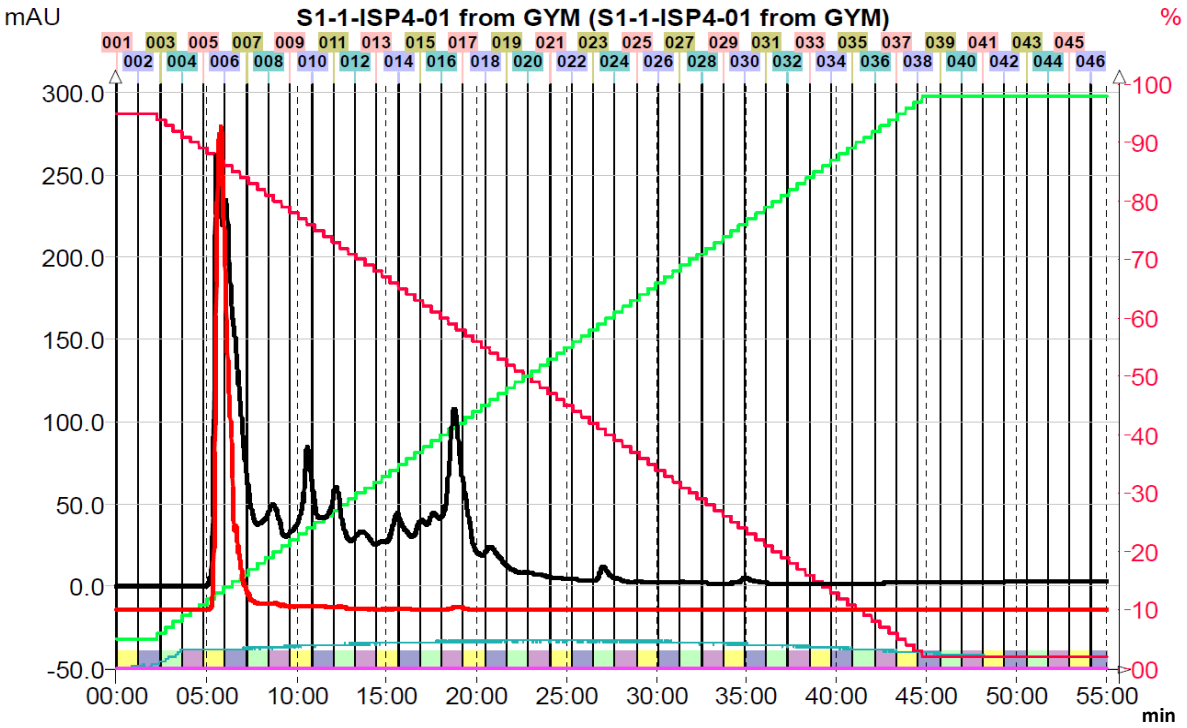


Figure 18. S1-1-ISP4-01 wt strain Flash chromatography chromatogram. 46 fractions were obtained after Flash chromatography. 190 nm UV detector (black). ELSD detector (red). Methanol gradient (green). Water gradient (red).

Fractions were tested against *S. cerevisiae* in paper disc diffusion assays. It was found that only F11, F12 and F13 showed bioactivity, where F12 was the most bioactive. The fractions were analyzed by HPLC-MS. A major peak contained in F12 showed a mass in the positive mode of $[M+H]^+$ m/z 282.1700 with a predicted formula of $C_{15}H_{23}NO_4$. Data were compared with databases and they perfectly fit with the antifungal compound cycloheximide (*Fig. 19*),

which is known to be produced by the genus *Streptomyces*. This mass was also found in the less bioactive fractions, F11 and F13 respectively.

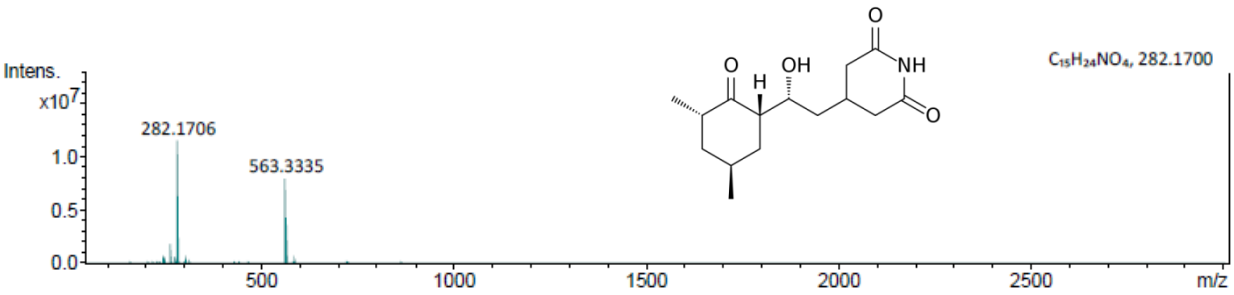


Figure 19. HPLC-HRMS positive mode data of *S*. sp. S1-1-ISP4-01 fraction 12 and chemical structure of the antifungal compound cycloheximide.

4.2.2. Genome analysis

A draft genome of 9.9 Mb from *S*. sp. S1-1-ISP4-01 was obtained, and 34 BGCs were identified in this genome by using antiSMASH 4.02 (*Table 3*).

No	Cluster type	Presence in another bacterium	Putative product
1	PKSII	*Streptomyces* Gö66	Isatropolones
2	PKSI-NRPS	*Streptomyces* sp. MMG1522	-
3	Bacteriocin	-	-
4	Bacteriocin	-	-
5	NRPS	*Streptomyces nodosus* strain ATCC 14899	Coelichelin
6	Terpene	-	-
7	Terpene	*Streptomyces stelliscabiei* strain P3825	Hopene
8	NRPS	*Amycolatopsis decaplanina* DSM 44594	Mirubactin
9	Siderophore	*Streptomyces* sp. Root1319	-
10	Lantipeptide	-	-
11	Lantipeptide	*Streptomyces pratensis* ATCC 33331	-
12	Bacteriocin	-	-
13	PKSI-NRPS	*Streptomyces clavuligerus* ATCC 27064	Frontalamides
14	PKSII-NRPS	*Streptomyces* sp. Root1319	Spore pigment
15	Bacteriocin	-	-
16	Other	-	-
17	Bacteriocin	*Streptomyces* sp. WM6372	-
18	PKSII	*Streptomyces* sp. MUSC119T	Lamdomycin
19	Butyrolactone	-	-
20	Bacteriocin	-	-
21	Trans AT-PKS	*Streptomyces anulatus* strain ATCC 11523	Cycloheximide
22	Terpene	-	-
23	NRPS	*Streptomyces* sp. Tu6071	-
24	Ectoine	*Streptomyces fulvissimus* DSM 40593	Ectoine
25	Ladderane-NRPS	*Streptomyces collinus* Tu 365	-
26	PKSI	*Streptomyces* sp. NRRL S-4	-
27	PKSI	-	-
28	NRPS	-	-
29	NRPS	-	-
30	NRPS	-	-
31	NRPS	*Streptomyces hygroscopicus* subsp. jinggangensis 5008	-
32	PKSI	-	
33	PKSI	-	-
34	Terpene	*Streptomyces* sp. JS01	-

Table 11. BGCs identified in the genome of Streptomyces sp. S1-1-ISP4-01 with antiSMASH 4.02. In gray color those BGCs which are present only in *S*. Sp. S1-1-ISPA-01.

Based on the literature data, and according to the MiBIG database, cluster 21 was found to be the one responsible for cycloheximide biosynthesis. This prediction perfectly fit with the bioassays and mass spectrometry results previously obtained.

Some other, and potentially interesting and unique BGCs were identified in this strain. One of them is cluster 1 (PKSII) which may be responsible for the biosynthesis of aromatic polyketide. In this cluster 3 positive transcriptional regulators belonging to the SARP family (SARP75, SARP79 and SARP95) were identified.

Each SARP regulator was independently cloned into the plasmid pSOK806 under the control of the strong constitutive P_{ermE*} promoter. The constructs were successfully conjugated into *S*. sp. S1-1-ISP4-01 to generate three recombinant strains that overexpress independently the SARP regulators.

These recombinant strains were cultivated in different fermentation media. Interestingly, in medium SM17 some of the strains overexpressing SARP genes yielded methanol extracts that were bioactive against *E. coli* and *B. subtilis* (*Fig. 20*). However, bioactivity was absent in the extracts from wt and S1-1-ISP4-01(pSOK806) strains.

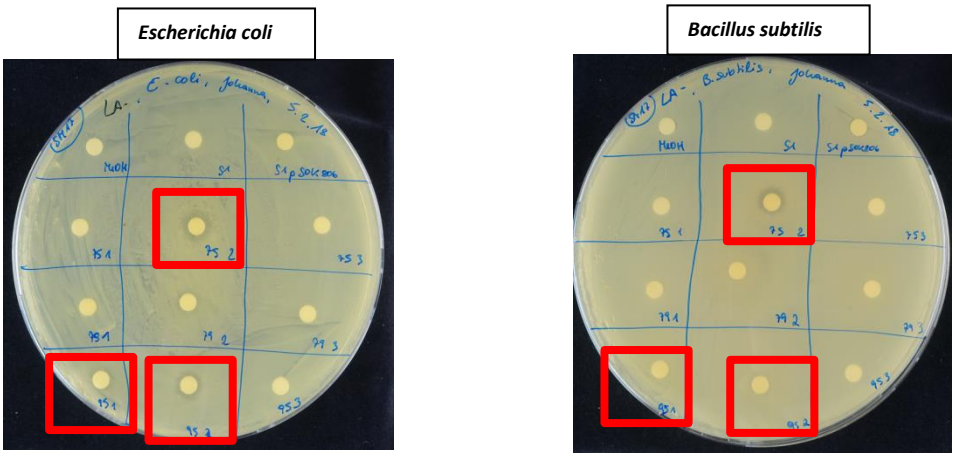


Figure 2012. Bioassays against *E. coli* and *B. subtilis* of methanol extracts from S1-1-ISP4-01 SARP mutants cultivated in SM17 medium. 3 SARP75 mutants (75). 3 SARP79 mutants (79). 3 SARP95 mutants (95). Wt S1-1-ISP4-01 (S1). Empty pSOK806 mutant (S1 pSOK806). In red box the bioactive SARP mutants.

Based on the results from bioactivity tests, the methanol extract of mutant strain SARP75_2, which showed the highest bioactivity against *E. coli* and *B. subtilis*, was fractionated by Flash chromatography under standard conditions (*Fig. 21*).

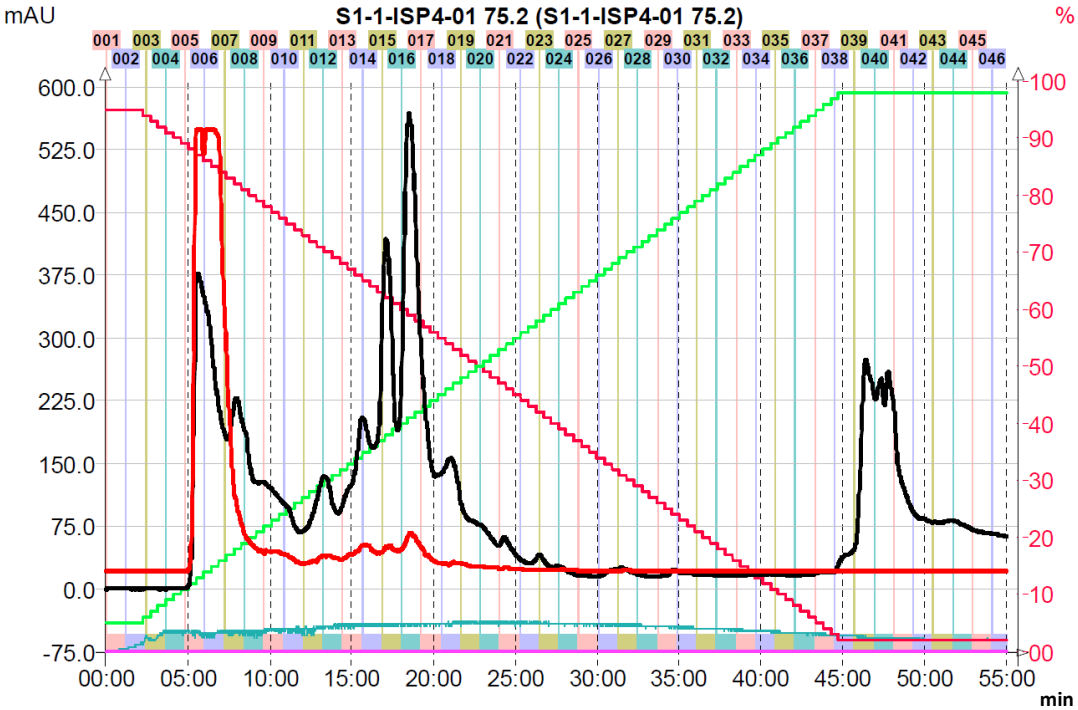


Figure 21. S1-1-ISP4-01 (SARP75_2) strain Flash chromatography chromatogram. 46 fractions were obtained after Flash chromatography. 190nm UV detector (black). ELSD detector (red). Methanol gradient (green). Water gradient (red).

Fraction 6 and 7 showed both bioactivities against *E. coli* and *B. subtilis*. However, they were not further investigated because they contain very polar compounds that could not be separated by reverse phase chromatography.

On the other hand, fractions 19, 20 and 21 showed bioactivity only against *B. subtilis*. They were pooled to make F8 fraction.

An up-scaled batch of S1-1-ISP4-01 (SARP75_2) strain in medium SM17 was generated to produce more of F8 fraction. The latter was further fractionated by semi-preparative HPLC as described in Material and methods section (*Fig. 22*).

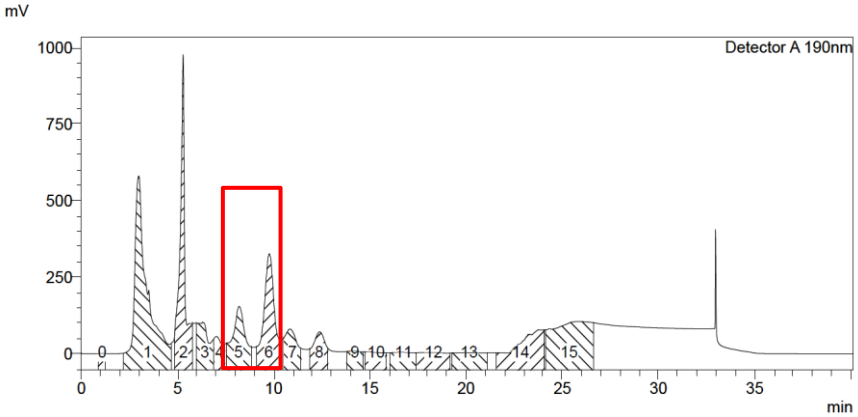


Figure 22. Semi preparative HPLC chromatogram of bioactive F8 fraction from S1-1-ISP4-01 (SARP75_2). 15 sub-fractions were generated. In red box are the bioactive sub-fractions.

After this fractionation, sub-fractions 5 and 6 showed a weak bioactivity against *B. subtilis*. Both sub-fractions were analyzed by LC-MS and found to contain the same main compound with a mass of 410.18 Da and a sum formula of $C_{18}H_{26}N_4O_7$ (*Fig. 23*).

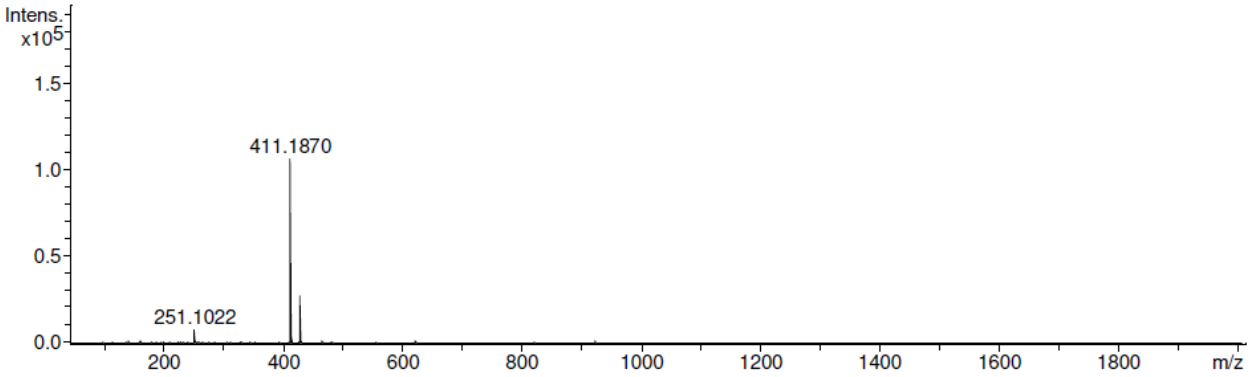


Figure 23. HPLC-HRMS positive mode data of main compound found in sub-fractions 5 and 6 from S1-1-ISP4-01 SARP75_2 mutant.

MS/MS data analysis suggested that the bioactive molecule is most likely linear and consists of the following building blocks: $C_7H_4O_3$ - NH_3 - C_5H_7NO - $C_6H_{12}N_2O_3$. It was proposed that the molecule is most probably an acylated tripeptide (ornithine-alanine-alanine) with a dihydroxybenzoic acid attached to one of the two nitrogens of the ornithine residue (*Fig. 24*). Nevertheless, even if the composition of the molecule was clearly to determine, many isomers of the molecule could be possible.

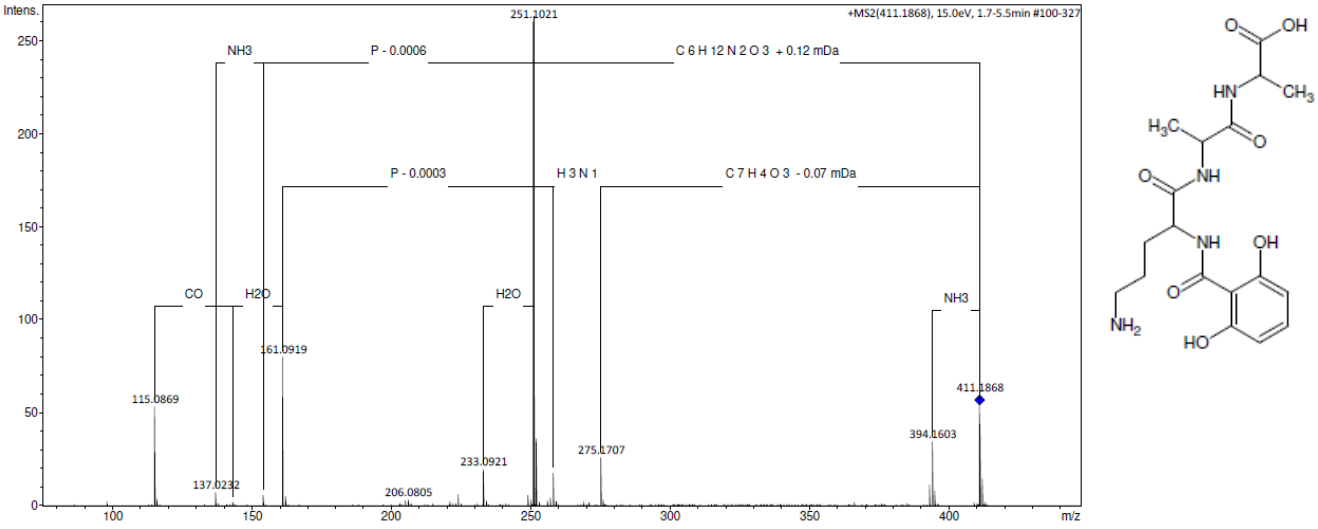


Figure 24. MS/MS data and proposed chemical structure of main compound found in sub-fractions 5 and 6 from S1-1-ISP4-01 SARP75_2 mutant.

Unfortunately, purification of enough material for NMR analysis and structural confirmation of this molecule turned out to be very challenging and was abandoned.

After genome analysis, it was suggested that cluster 8, a NRPS type cluster, was may be responsible for the biosynthesis of this compound. antiSMASH results suggested that it might be a siderophore related to mirubactin (*Fig. 13*).

Since connection between the overexpression of SARP75 and the production of this acylated tripeptide could not be established, a comparison of secondary metabolomes by HPLC-HRMS between the methanolic extracts of S1-1-ISP4-01(SARP75_2) and wt strain was performed. It was found that indeed, both strains produce the acylated tripeptide, cycloheximide and actiphenol, a precursor in the cycloheximide biosynthesis. However, differences of the hydrophilic compounds content were found. The latter could explain results from bioactivity tests, since the polar fractions of the methanolic extract from S1-1-ISP4-01(SARP75_2) showed bioactivity against *E. coli* and *B. subtilis* but the extract from the wt strain was not active. Therefore, it might be that the SARP regulator is activating the production of some bioactive polar compounds.

Recently, it became apparent that the cluster targeted for activation (cluster 1) was already characterized. This cluster was connected to the production of isatropolones, a group of compounds with activity against *Leishmania donovani* [98]. The crude extracts of both of S1-1-ISP4-01 (SARP75_2) and wt strain were analyzed again by HPLC-HRMS. However, we could not identify the production of isatropolones.

4.3. *Kitasatospora* sp. RLA070

4.3.1. Prospecting for bioactivities from *Kitasatospora* sp. RLA070

Strain RLA070 was isolated recently by Dr. Martina Oberhofer in our laboratory from the rhizosphere of the medicinal plant *Leontopodium nivale* (Edelweiss). It was identified as a member of the genus *Kitasatospora* by using 16S rRNA gene sequencing.

The strain was cultivated in different solid and liquid media. It was observed that the strain is able to sporulate on oat meal agar medium (OA). Furthermore, when cultivated in liquid ISP2 medium, its methanolic extracts exhibited potent antibiotic activity against different Gram-positive and Gram-negative bacteria.

An up-scaled batch of fermentation was produced and bioactive methanolic extract generated. The extract was fractionated using Flash chromatography (*Fig. 25*) under standard conditions (see Materials and methods section).

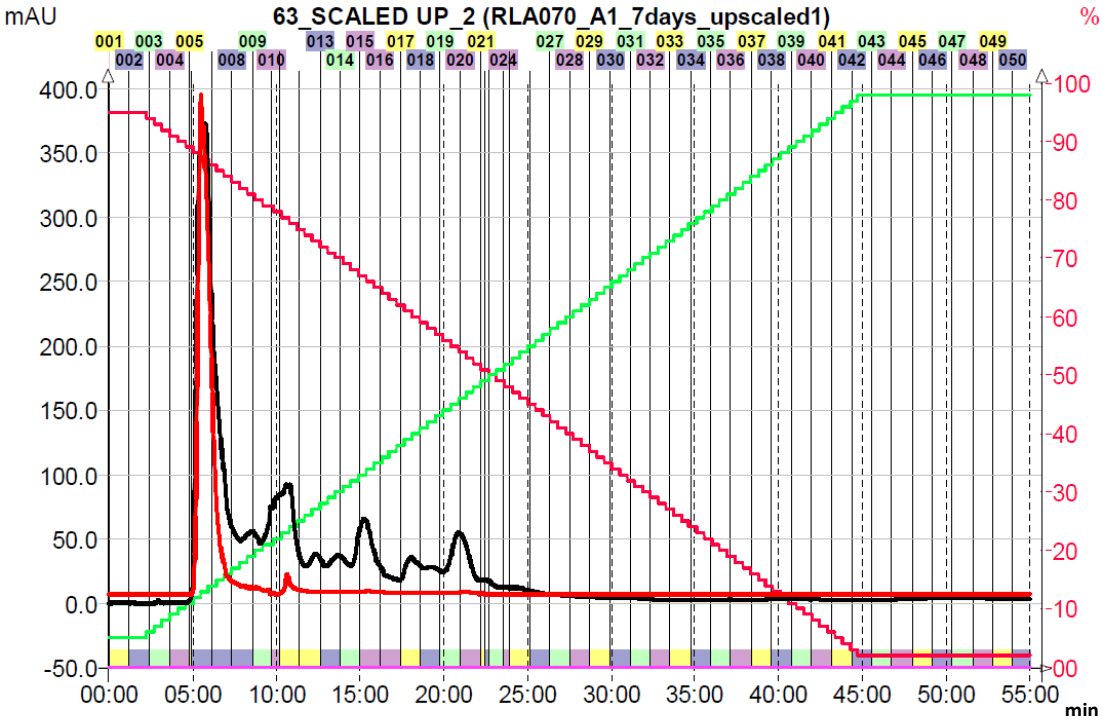


Figure 25. Extract from RLA070 strain Flash chromatography chromatogram. 50 fractions were obtained after Flash chromatography. 190nm UV detector (black). ELSD detector (red). Methanol gradient (green). Water gradient (red).

Fractions were pooled in 11 major fractions for bioassays, and bioactivity tests demonstrated antibiotic activity in F6 (fractions 19-20), F7 (fractions 19-24) and F8 (fractions 25-30).

The bioactive RLA070 methanolic crude extract and fractions F6, F7 and F8 were analyzed by HPLC-MS. The presence of halogenated compounds was confirmed when HRESIMS

analysis was performed. The presence of three peaks corresponding to a compound with a mass of 479.11 Da in the positive mode and 477.11 Da in the negative mode was detected. Proposed chemical formula $C_{22}H_{23}N_2O_8Cl$ was predictedfor this compound (*Fig. 26*).

Those results perfectly fit with the mass of the broad-spectrum antibiotic chlortetracycline.

F6 mainly contains the isomers, presumably epi-, iso- and/or epi-isochlortetracyclines (Peaks 1 and 2), while the chlortetracycline itself (peak 3) is mainly present in F7 and F8 (*Fig. 26*).

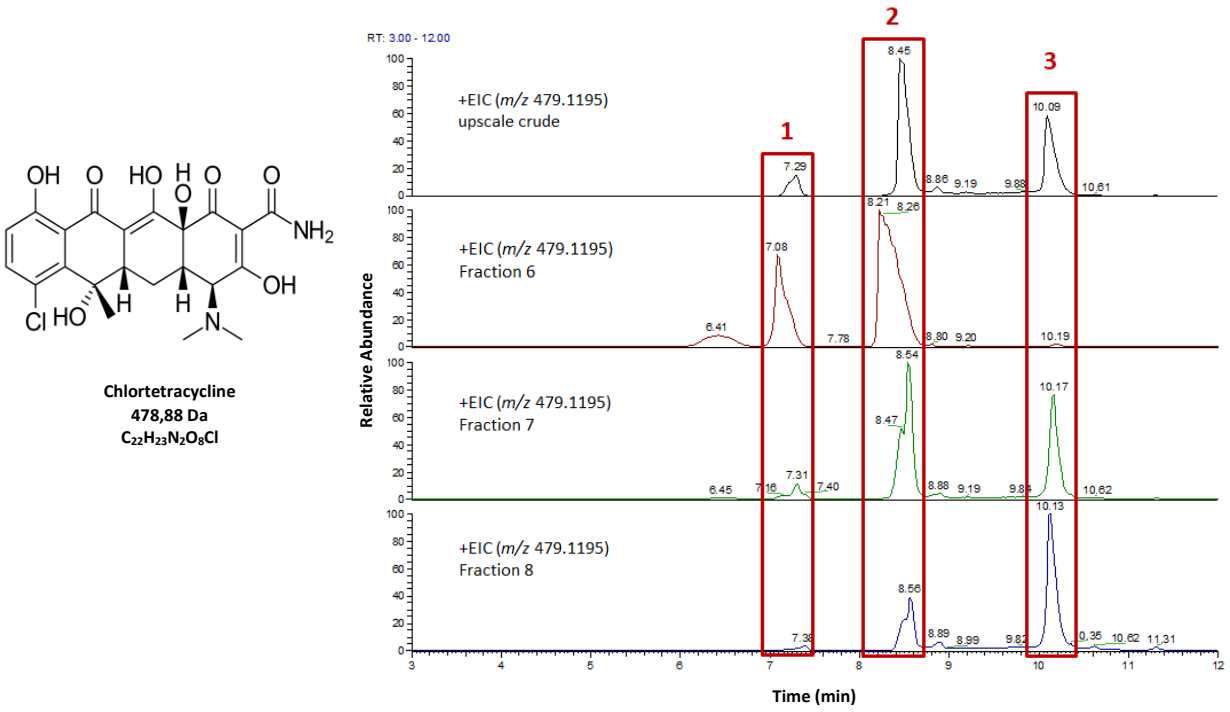


Figure 26. Chlortetracycline chemical structure and Positive Extracted Ion Chromatogram (+EIC) (m/z 479.1195) of RLA070 upscale crude extract, F6, F7 and F8. Peaks 1, 2 and 3 are indicated in red boxes.

This is not the first report of tetracyclines production by *Kitasatospora* species. Recently, the genome of the tetracycline producer *Streptomyces viridifaciens* DSM 40239 revealed that this organism belongs to the *Kitasatospora* genus [99].

Further analysis of F7 and F8 revealed the presence of at least 2 potentially novel compounds in both fractions, which masses did not fit with those in the databases, such as the Dictionary of Natural Products (*Fig. 27*).

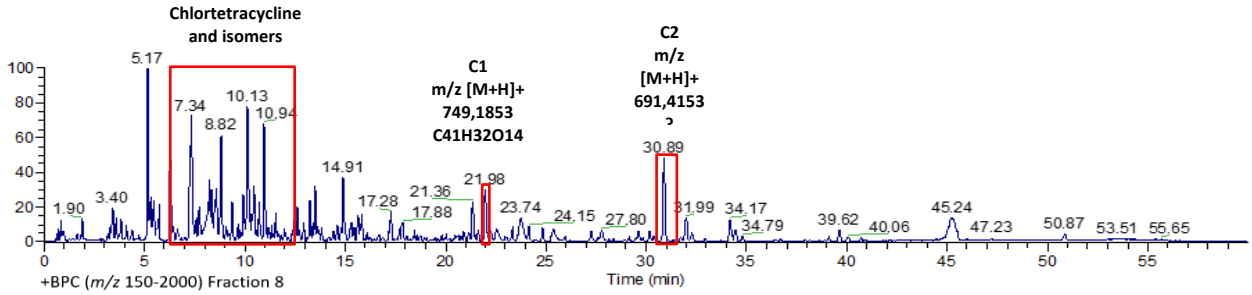


Figure 27. RLA070 fraction F8 Positive Base Peak Chromatogram (+BPC) (m/z 150-2000). In red boxes: chlortetracycline and isomers, predicted novel compound 1 (C1), and predicted novel compound 2 (C2).

Compound 1 (C1) has a mass of $[M+H]^+$ 749.1853 Da in the positive mode and $[M-H]^-$ 747.1713 Da in the negative mode. Its chemical formula was proposed as $C_{41}H_{32}O_{14}$ and it was suggested to represent a new anthraquinone dimer. However, further analyses with NMR must to be done to confirm this.

Compound 2 (C2) has a mass $[M+H]^+$ 791.4153 Da in the positive mode and $[M-H]^-$ 689.4016 Da in the negative mode. With just MS data, it was not possible to assign a formula to this compound.

F7 was further fractionated by a semi-preparative HPLC as described in Material and methods section (*Fig. 28*).

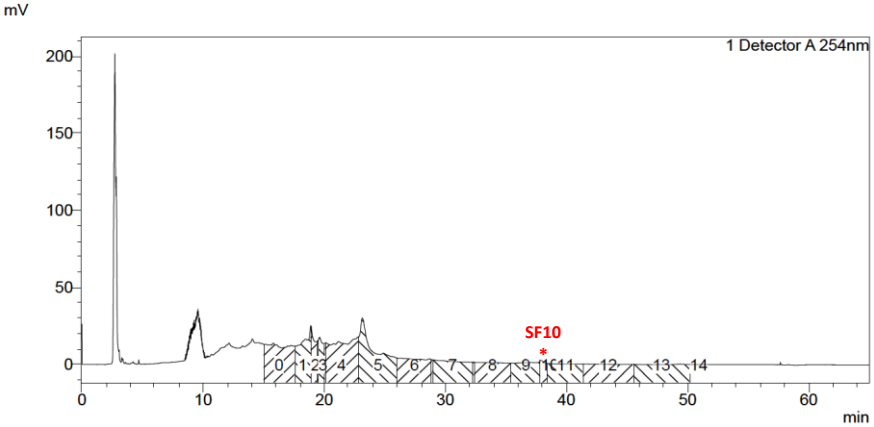


Figure 28. Semi preparative HPLC chromatogram of bioactive F8 fraction from RLA070 strain. 13 sub-fractions were generated. In red color is the sub-fraction SF10.

13 sub-fractions from F7 were collected and tested for antibiotic activity. Only those that contained chlortetracycline were found to be active. The latter might suggest that the whole antibiotic activity found in the methanolic extract from *Kitasatospora* sp. RLA070 cultivated in these conditions is due to the production of chlortetracycline and its isomers.

4.3.2. Genome analysis

A complete genome sequence of *Kitasatospora* sp. RLA070 was obtained. It consists of one linear chromosome of 10 Mb and 6 plasmids (both linear and circular).

The genome was analyzed with antiSMASH 5.0, and 48 BGCs were detected (*Table 4*). Interestingly, all the BGCs were found only in the chromosome.

No	Cluster type	Presence in another bacterium	Putative product
1	PKSI-NRPS	-	-
2	Ectoine	*Kibdelosporangium phytohabitans* strain KLBMP1111	-
3	Butyrolactone	*Streptomyces* sp. RSD-27	-
4	PKSI-NRPS	*Kitasatospora setae* KM-6054	-
5	PKSII	*Streptomyces aureofaciens* strain F3	Chlortetracycline
6	Terpene	*Streptomyces tsukubaensis* NRRL18488	-
7	Terpene	*Streptomyces* sp. Mg1	-
8	Terpene	-	-
9	NRPS	*Streptomyces* sp. NRRL S-495	-
10	NRPS	-	-
11	PKSI-NRPS	-	-
12	Terpene	*Kitasatospora* sp. Root107	-
13	Terpene	*Kitasatospora setae* KM-6054	Hopene
14	Other	*Kitasatospora* sp. Root187	-
15	Fused	-	-
16	Aminoglycoside	*Kitasatospora* sp. Root107	-
17	NRPS	*Streptomyces* sp. XY431	Livipeptin
18	PKSI-NRPS	-	-
19	Bacteriocin	*Streptomyces rubellomurinus* strain ATCC 31215	-
20	Siderophore	*Kitasatospora setae* KM-6054	-
21	Terpene	*Streptomyces* sp. NRRL S-495	Geosmin
22	PKSI-NRPS	-	-
23	Butyrolactone	-	-
24	NRPS	*Streptomyces roseochromogenus* subsp. oscitans DS 12.976	Arylomycin
25	Lanthipeptide	*Kitasatospora cheerisanensis* KCTC 2395	-
26	Terpene	*Streptomyces aureofaciens* strain ATCC 10762	-
27	PKSIII	-	-
28	Bacteriocin	-	-
29	Arylpolyene-NRPS	*Kitasatospora cheerisanensis* KCTC 2395	Qinichelins
30	Lanthipeptide	*Streptomyces rubellomurinus* strain ATCC 31215	-
31	PKSI	-	-
32	NRPS	-	-
33	PKSI-NRPS	*Saccharothrix* sp. ST-888	Sporolide

34	NRPS	-	-
35	Lanthipeptide	*Streptomyces* sp. SirexAA-E	-
36	Thiopeptide	*Streptomyces lactacystinaeus*	Lactazole
37	LAP-thiopeptide	*Kitasatospora* sp. Root187	-
38	Indol-lanthipeptide	*Streptomyces griseochromogenes*	Blasticidin
39	NRPS	-	-
40	NRPS	*Streptomyces venezuelae* ATCC 10712	-
41	NRPS	-	-
42	NRPS	*Streptomyces rubellomurinus* strain ATCC 31215	-
43	NRPS	*Actinopolyspora erythraea* strain YIM 90600	-
44	Lanthipeptide	-	-
45	PKSI	*Kitasatospora setae* KM-6054	Bafilomycin
46	NRPS	-	-
47	Ectoine	-	-
48	PKSI-NRPS	-	-

Table 4. BGCs identified in the genome of *Kitasatospora* sp. RLA070 with antiSMASH 5.0. In gray color those BGCs which are present only in *K*. sp. RLA070.

Many of the BGCs found in the genome of *Kitasatospora* sp. RLA070 are presumably unique. Therefore, this microorganism may have a good potential for the biosynthesis of novel compounds. Further research must to be done in terms of bioprospecting and genetic engineering of this strain. Indeed, we have experimental evidence that this organism could produce at least two novel compounds.

4.4. *Amycolatopsis* sp. YIM10

4.4.1 Prospecting for bioactivities from *Amycolatopsis* sp. YIM10

Strain YIM10 was isolated from rare earth mine of Bayan Obo, Inner Mongolia, China. It was identified as a member of the genus *Amycolatopsis* using 16S rRNA gene sequencing.

YIM10 was cultivated in different liquid media and, when cultivated in the soy flour-based medium 5288 (see Material and methods), it yielded methanolic extract with antibiotic activity against *B. subtilis*.

The extract was fractionated using Flash chromatography (*Fig. 29*) under standard conditions (see Material and methods section).

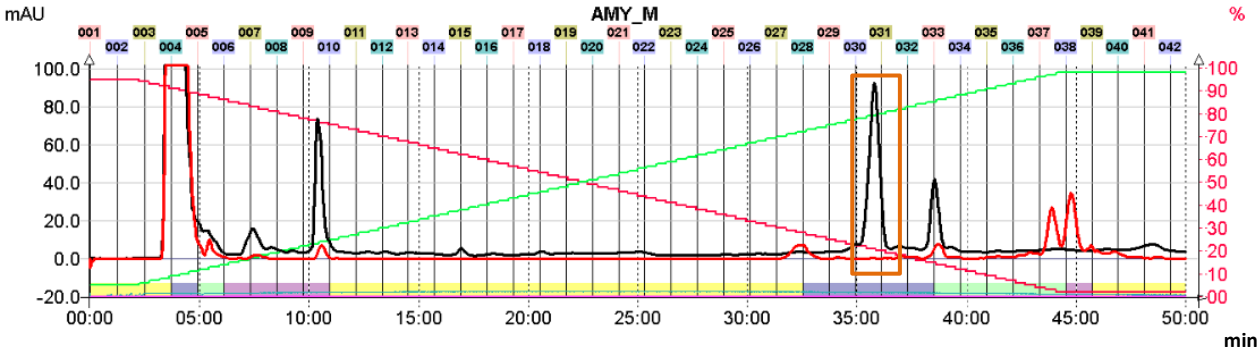


Figure 29. YIM10 strain's methanolic extract Flash chromatography chromatogram. 42 fractions were obtained after Flash chromatography. 190nm UV detector (black). ELSD detector (red). Methanol gradient (green). Water gradient (red). The bioactive fractions are boxed in orange.

42 fractions were obtained and pooled in 16 main fractions to perform antibiotic paper disc diffusion assay. It was found that F11 (fractions 30-31) is the only bioactive fraction againts *B. subtilis* (*Fig. 30*).

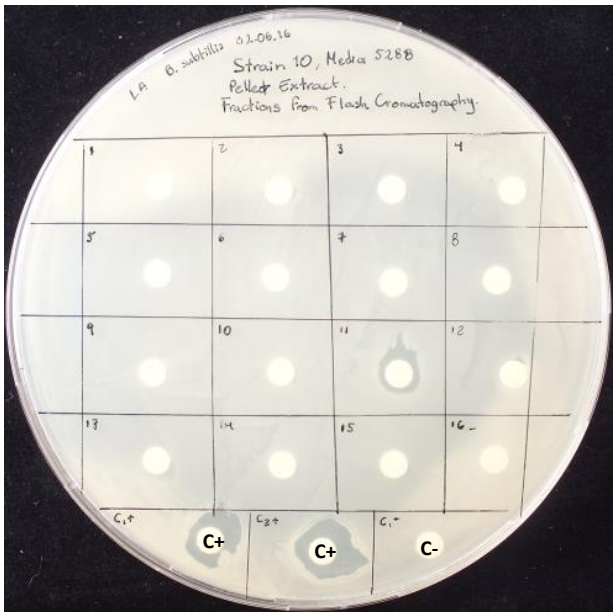


Figure 30. Bioassay against *B. subtilis* of fractions obtained after Flash chromatography of methanol crude extract from *Amycolatopsis* sp. YIM10 cultivated in 5288 medium. Numbers indicate the fractions (16 in total). Bioactive methanol extract from *Amycolatopsis* sp. YIM10 cultivated in 5288 medium (C+). Methanol blank (C-).

F11 was analyzed by HRMS, and the major compound with a mass $[M+H]^+$ of 219.1025 Da in the positive mode and a proposed sum formula $C_{13}H_{14}O_3$ (*Fig. 31*) was identified. Those results were compared with the Dictionary of Natural Products database. However, a proper

candidate produced by actynomycete bacteria that could fit these characteristics could not be identified. The latter suggested that this is probably a new compound.

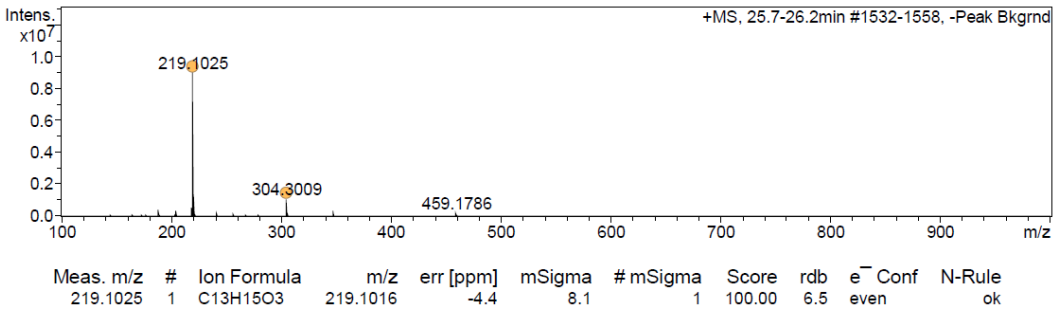


Figure 31. HPLC-HRMS positive mode data and proposed sum formula of main compound found in bioactive F11 fraction of *Amycolatopsis* sp. YIM10.

An up-scaled fermentation batch was prepared to generate more of the F11 fraction. 4 mg were obtained and was enough material to perform 1D and 2D NMR structure elucidation. The NMR results revealed that the bioactive compound is the small molecule 1,2,4-trimethoxynaphthalene (*Fig. 32*). Interestingly, this compound have been isolated previously only from the liverwort *Adelanthus decipiens* [100]. Furthermore, nobody has associated this molecule with a bioactivity before. However, the most interesting questions are: i) how an *Amycolatopsis* strain can produce a compound that is produced by a plant, and ii) how it is biosynthesized.

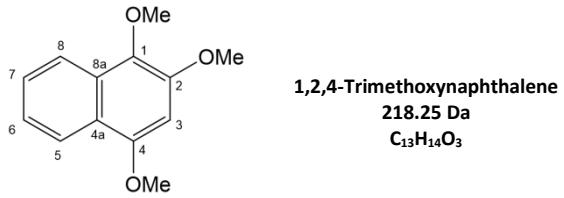


Figure 32. Chemical structure of the 1,2,4-Trimethoxynaphthalene molecule. Molecular weight and chemical formula.

The pellet of *Amycolatopsis* sp. YIM10 cultivated in 5288 medium was extracted with methanol and fractionated by semi preparative HPLC.

The latter lead to the identification and purification of several compounds that turned out to be different tigloside congeners.

The main compound related to tigloside was analyzed by HRMS and found to have a mass $[M+H]^+$ of 1159.4784 Da in the positive ionization mode (*Fig. 33*) and a proposed sum formula $C_{54}H_{78}O_{27}$.

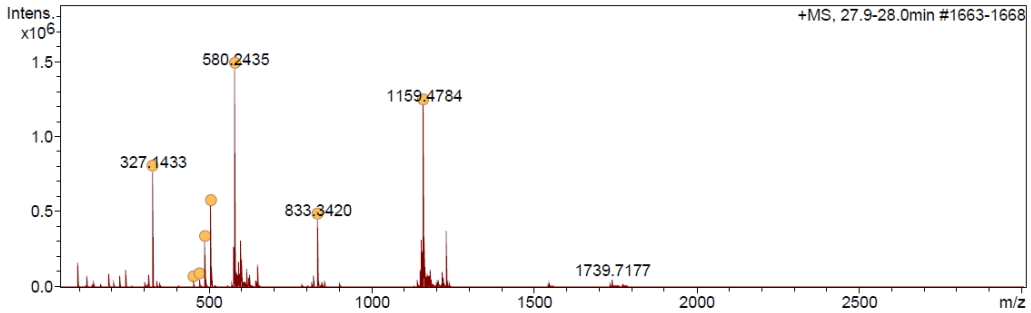


Figure 33. HPLC-HRMS positive mode data of main tigloside compound found produced by *Amycolatopsis* sp. YIM10.

The compound was confirmed by NMR analysis as tigloside, a tetrasaccharide tailored with tiglic acid moieties (*Fig. 34*). Originally, this compound was isolated from other *Amycolatopsis* strain [101]. Interestingly, YIM10 produces several minor derivatives that have not been described yet. Even though no bioactivity has been reported for this molecule, there are reports of other tetrasaccharide compounds from *Amycolatopsis* strains that display interesting bioactivities, like actinotetraose E that has immunosupresive activity [102]. For the cytoprotective trehangelins (trehalose angelates) that appear to be structurally related to tiglosides, BGC has been already identified [103].

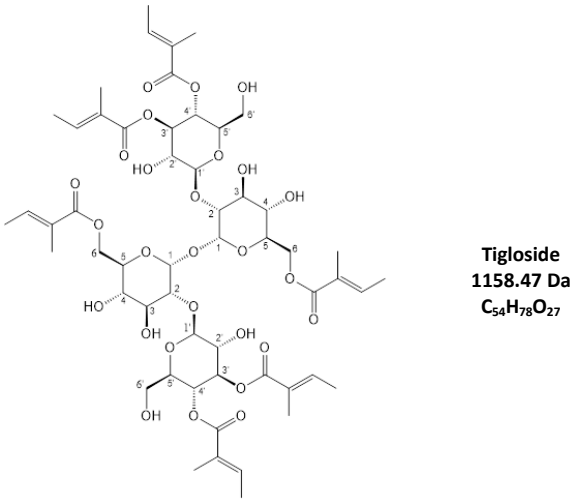


Figure 34. Chemical structure of tigloside, the main compound produced by *Amycolatopsis* sp. YIM10 cultivated in 5288 medium.

Keeping this in mind, the genome of YIM10 was scanned for genes products homologous to those encoded by the trehangelins BGC. However, not plausible matches were identified. The latter suggested that perhaps an alternative biosynthetic pathway is involved in tigloside production.

4.4.2. Genome analysis

The genome of *Amycolatopsis* sp. YIM10 was completely sequenced. It consists of a circular chromosome of 10.31 Mb and a plasmid of 39.9 Kb.

The genome was analyzed with antiSMASH 3.0.5 and 40 BGCs were identified, many of which appear unique, because they are not found in the genomes of other bacteria (*Table 5*).

No	Cluster type	Putative product	Presence in another bacterium
1	Ectoine	Ectoine	*Amycolatopsis decaplanina* DSM 44594
2	PKSI	Glycosylated polyketide	-
3	Terpene	Carotenoid	*Amycolatopsis japonica* MG417-CF17
4	Terpene	Glycosylated diterpenoid	*Saccharopolyspora erythraea* NRRL2338
5	Other	Halogenated NRPS	-
6	Bacteriocin	Bacteriocin	-
7	Terpene	Lipoprotein-terpenoid	-
8	PKSI-NRPS	Glycosylated PKSI-NRPS (threonine based)	-
9	Other	NRPS (aromatic amino acid)	-
10	Terpene	Carotenoid	*Amycolatopsis azurea* DSM 43854
11	NRPS	Glycosylated NRPS	-
12	PKSIII	Putative 1,2,4-Trimethoxynaphthalene	-
13	Oligosaccharide-NRPS	Halogenated-glycosylated NRPS	*Saccharomonospora glauca* K62
14	PKSI-NRPS	Hybrid PKSI-NRPS	-
15	NRPS	Anthramycin-like	*Nocardia brasiliensis* NBRC 14402
16	PKSI-NRPS	Hybrid PKSI-NRPS	-
17	Ladderane-NRPS	Carbamoylated ladderane-NRPS	-
18	NRPS	Mirubactin-like	*Saccharomonospora cyanea* NA-134
19	Arylpolyene-Ladderane	Halogenated arylpolyene-ladderane	-
20	NRPS	Halogenated NRPS	-
21	Lasso peptide	Lasso peptide	-
22	PKSI	Unknown	*Streptomyces cattleya* DSM 46488
23	Phosphonate	Unknown phosphonate	*Micromonospora sp.* NRRL B-16802
24	Indole	Indole	-
25	Siderophore	Unknown siderophore	*Amycolatopsis nigrescens* CSC17Ta-90
26	Bacteriocin	Bacteriocin	*Amycolatopsis sp.* MJM2582
27	Other	Halogenated NRPS	*Amycolatopsis nigrescens* CSC17Ta-90
28	PKSI	Glycosylated polyketide	-

29	NRPS	Carbamoylated NRPS	-
30	NRPS	Glycosylated nrps	-
31	PKSI	Unknown polyketide	*Amycolatopsis nigrescens* CSC17Ta-90
32	Other	Unknown	-
33	NRPS	Unknown cephalosporin	*Streptomyces cattleya* NRRL 8057
34	PKSI	Nystatin-like	*Streptomyces sp.* PVA 94-07
35	NRPS	Scabichelin-like	*Amycolatopsis sp.* MJM2582
36	PKSI	Quartromicin-like	*Amycolatopsis orientalis*
37	PKSI-NRPS	Unknown	*Amycolatopsis mediterranei* S699
38	PKSI	Glycosylated polyketide	*Amycolatopsis mediterranei S699*
39	Lantipeptide	Erythreapeptin-like	-
40	Butyrolactone-PKSI	Unknown	*Saccharomonospora viridis* DSM 43017

Table 5. BGCs identified in the genome of *Amycolatopsis* sp. YIM10 with antiSMASH 3.0.5. In gray color those BGCs which homologous BGC could not be identified in other genomes.

4.4.3. Heterologous expression

Given that the genome of *Amycolatopsis* sp. YIM10 contains very interesting BGCs and therefore may have a potential to produce novel compounds, it was regarded as an excellent candidate for genome mining and application of genetic engineering approaches.

However, the this *Amycolatopsis* strain was found to be resistant to all of the antibiotics used as selectable markers in actinomycetes, in particular apramycin, hygromycin, thiostrepton, kanamycin and puromycin. Thus, establishing a gene transfer system for this bacterium appeared problematic. Considering this, cloning and expression of BGCs in a heterologous host seemed like a logical strategy to circumvent the problem. Therefore, a fosmid (pCC1FOS-based) genome library was constructed (see Materials and methods section).

We were interested in two particular BGCs; cluster 21, a predicted lasso peptide cluster, and cluster 12, a predicted PKSIII type cluster. Both clusters were relatively small (less than 40 Kb) and the chance of having the complete cluster in a single fosmid library clone was considered relatively high.

Screening of the genome library using pooled PCR with primers designed for flanking and central regions of the clusters led to identification of both BGC (12 and 21) in single fosmids. The cassette containing integration site int-attP$^{\varphi C31}$ and *oriT sequence* was incorporated into identified fosmids carrying BGCs 12 and 21 as described in Materials and methods. The latter allowed to conjugate the recombinant fosmids in different *Streptomyces* hosts for integration into the genomes and heterologous expression.

Since cluster 12 is the only PKSIII type cluster identified in the genome of YIM10 by antiSMASH, it was proposed that this BGC could be the responsible for the biosynthesis of the 1,2,4-trimethoxynaphthalene. Heterologous expression of this BGC in different *Streptomyces* hosts coupled to the overexpression of a positive transcriptional regulator belonging to the LuxR family found in this BGC failed to activate the production of the targeted compound. Many reasons could explain the latest situation, but most probably hypothesis would be due to the lack of the biosynthetic machinery that would supply a proper precursor in the heterologous hosts used.

The lasso peptide BGC, cluster 21, is around 10 kb long and contains all the main genes for the biosynthesis of this class of peptides [64,65]. In particular, it contains genes presumed to be involved in the leader peptide cleavage (*filB1* and *filB2*), the macrolactam ring formation (*filC*), putative oxidoreductase reactions (*filE*), transporters (*filD1* and *filD2*), and transcriptional regulation (*filR1*) (*Fig. 35*).

This cluster contains 2 genes encoding for precursor peptides (*filA1* and *filA2*), suggesting the possibility to produce two lasso peptide molecules.

Gene	Gene product
filR1	transcriptional regulator (SARP family, 476 aa)
filA1	felipeptin A1 lasso peptide precursor (42 aa)
filA2	felipeptin A2 lasso peptide precursor (41 aa)
filC	asparagine synthase (lasso cyclase, 603 aa)
filB1	PqqD family protein (leader peptide binding, 85 aa)
filB2	transglutaminase-like superfamily protein (leader peptidase, 136 aa)
filE	LLM class flavin-dependent oxidoreductase (oxidoreductase, 348 aa)
filD1	ABC transporter (570 aa)
filD2	ABC transporter (572 aa)

Figure 35. Annotation of BGC 21, a lasso peptide type cluster from *Amycolatopsis* sp. YIM10.

The recombinant fosmid carrying cluster 21 was introduced into *Streptomyces coelicolor* M1154 and *Streptomyces albus* J1074, and recombinant strains cultivated in different liquid and solid media. However, no lasso peptide production could be detected.

Next, the gene *filR1*, encoding a transcriptional regulator of the SARP family, was cloned in the plasmid pSOK806[97] under the control of the strong P_{ermE*} promoter. The construct was conjugated into the *Streptomyces* hosts that already contained the fosmid with lasso peptide cluster 21. The constitutive overexpression of this SARP regulator allowed the production of both lasso peptides in the two *Streptomyces* hosts when cultivated in liquid MYM medium (*Fig. 36*), as described in Material and methods section.

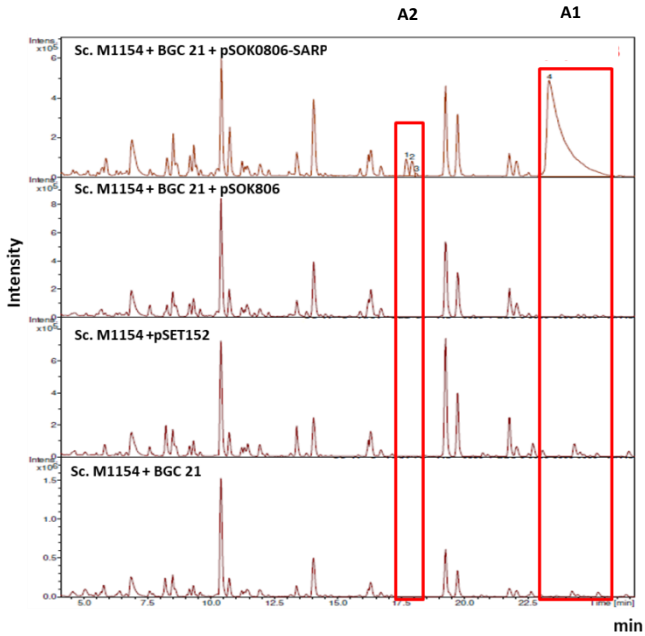
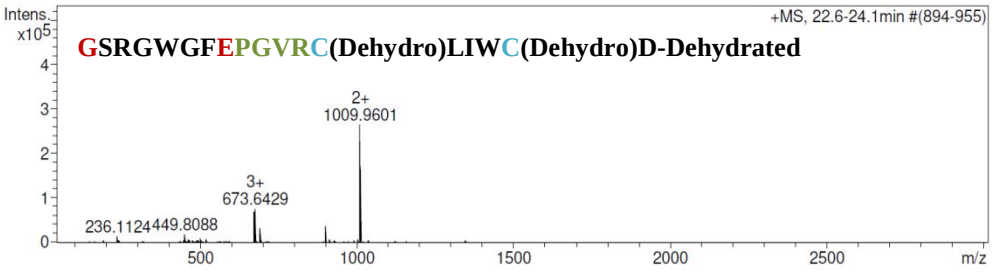


Figure 36. Positive Base Peak Chromatogram (+BPC) of *Streptomyces* M1154 (Sc. M1154) recombinant strains carrying cluster 21 and overexpressing SARP regulator gene *filR1* cultivated in MYM medium. In red box the peaks of lasso peptide A1 and A2 respectively.

Lasso peptide A1 was named as felipeptin A1 and predicted from the genome sequence as a 18 amino acid peptide. On the other hand, lasso peptide A2 was named as felipeptin A2 and predicted as a 17 amino acid peptide. Both lasso peptides are new molecules that never have been reported before.

The masses, chemical formulas and amino acid sequences were determined for both compounds through HRMS and fragmentation analysis, felipeptin A1 has a mass of 2019.3283 Da and a chemical formula $C_{91}H_{129}N_{26}O_{23}S_2$, felipeptin A2 has a mass of 1846.1093 Da and a chemical formula $C_{81}H_{118}N_{23}O_{23}S_2$ (*Fig. 37*).

Felipeptin A1: MW 2019.3283 Da, $C_{91}H_{129}N_{26}O_{23}S_2$



Felipeptin A2: MW 1826.1093 Da, $C_{81}H_{118}N_{23}O_{23}S_2$

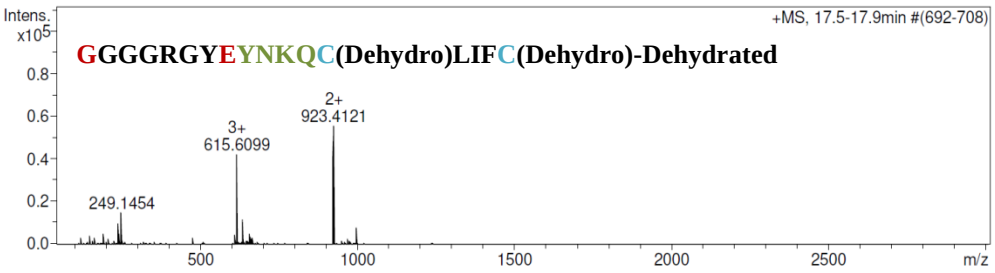


Figure 37. HPLC-HRMS positive mode data, proposed sum formulas and amino acid sequences of felipeptin A1 and felipeptin A2 respectively. In red colour the amino acid residues involved in the macrolactam ring formation. In blue colour the cysteine residues involved in the formation of putative a disulphide bridge. In green colour the amino acids between the two microcycles.

Given that *S. coelicolor* M1154 (Sc. M1154) host has a cleaner metabolic background than *S. albus* J1074 and basically no difference in lasso peptides yields were found between the two strains, it was decided to work further only with *S. coelicolor* M1154 strain.

Up-scaled fermentation and optimization of purification protocol allowed production yields of 12 mg/L of felipeptin A1, and 7 mg/L of felipeptin A2 (detailed information is found in Materials and methods section).

MS/MS analysis of both lasso peptides revealed that the two molecules had two microcycles in their structures. Both have a macrolactam ring encompassing eight amino acid residues from the glycine 1 (G) to the glutamic acid 8 (E) residue. Additionally, a second microcycle is formed between the two cysteine residues by a disulphide bridge formation. Both lasso peptides contain four amino acid residues between the two microcycles.

Recently, in Prof. Aachmann laboratory at Norwegian University of Science and Technology (NTNU), larger parts of felipeptin A1 and felipeptin A2 could be assigned through NMR analysis and a preliminary structure of felipeptin A2 has been calculated (*Fig. 38*).

Even though the proposed NMR structure of felipeptin A2 did not show the disulphide bridge between the two cysteine residues, there is evidence that those residues were reduced, clearly indicating the presence of a disulfide bridge connecting them.

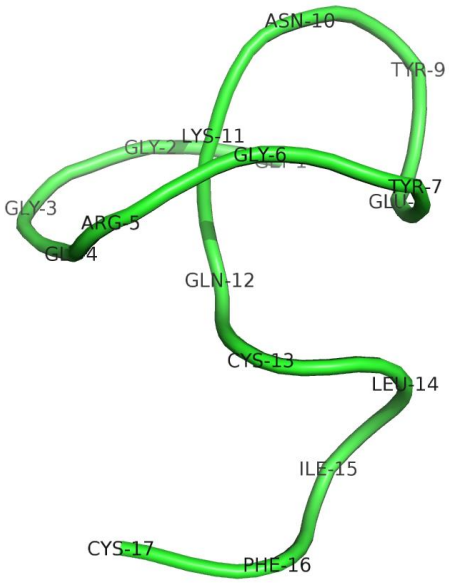


Figure 38. Preliminary structure of felipeptin A2 obtained by NMR analysis.

4.4.3.1. Bioassays for felipeptin A1 and A2

Given that there are examples of some lasso peptides originated from actinomycete bacteria that show antibiotic activities, and because the BGC of felipeptins A1 and A2 contains two ABC transporters, it was suggested that those compounds would have antibiotic properties. Thus, they were tested against different Gram-positive and -negative bacteria using paper disc diffusion assay, (see Materials and methods section).

It was found that both peptides have a weak activity against *B. subtilis*. Interestingly, when lasso peptides were tested in combination 1:1, a synergistic effect was detected (*Fig. 39*).

Compound	Amount (µg)
Felipeptin A1 (FA1)	100
Felipeptin A2 (FA2)	100
Mixture 1:1 FA1/FA2	50/50

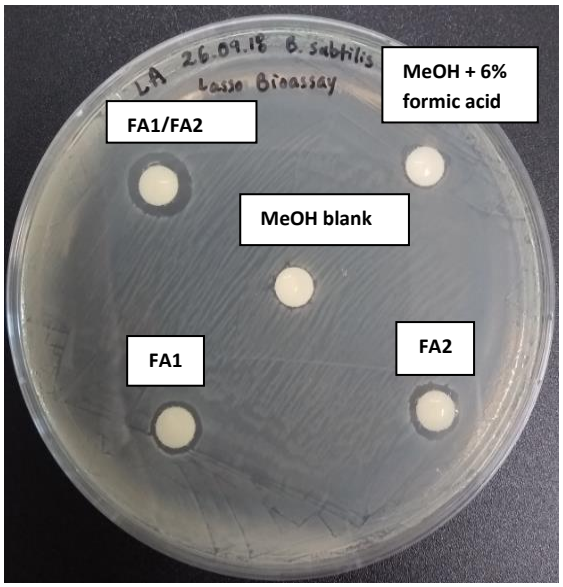


Figure 39. Bioassay against *B. subtilis* of felipeptin A1 (FA1), felipeptin A2 (FA2), mixture 1:1 (FA1/FA2). Blanks; MeoH, MeOH + formic acid 6%.

Felipeptins A1 and A2 were tested by MEDINA Foundation (Spain) against a panel of different Gram positive pathogens in liquid media-based assays in order to determine minimal inhibitory concentrations (MICs).
The pathogens used in the anti-infective assay were:

1. *Staphylococcus aureus* MRSA MB5393
2. *Staphylococcus aureus* MSSA ATCC29213
3. *Staphylococcus epidermidis*, CIinical isolate Iinezolid resistant
4. *Enterococcus faecium*, Clinical isolate vancomycin sensitive- 1
5. *Enterococcus faecium*, Clinical isolate VanA resistant
6. *Enterococcus faecium*, Clinical isolate VanB resistant
7. *Enterococcus gallinarum*, Clinical isolate VanC resistant
8. *Streptococcus mutans* ATCC25 17 5
9. *Streptococcus pyogenes*, Clinical isolate sensitive
10. *Streptococcus pneumoniae*, ATCC49619, penicillin medium resistance

The results obtained in the study pointed out that felipeptins A1, A2 and their combination did not show remarkable properties as anti-infective compounds against the selected Gram positive pathogens. Only in the cases of *S. pyogenes* and *S. pneumoniae* ATCC49619 felipeptin A1 and the 1:1 mixture showed moderate activity (*Table 6*). Interestingly, in the case of *S. pyogenes*, only a mixture of lasso peptides was active.

MIC [µg/mL]

Compound ID	*S. pyogenes*	*S. pneumoniae* ATCC 49619
FA2	>128	>128
FA1	>128	64
FA1 + FA2 (1:1)	64	64
Damptomycin	8	32
Erythromycin	<1	<1
Vancomycin	<1	<1

Table 6. Minimal inhibitory concentrations (MIC) (µg/mL) of felipeptin A1 (FA1), felipeptin A2 (FA2) and mixture 1:1 (FA1/FA2) against *S. pyogenes* and *S. pneumoniae* ATCC49619. Controls: daptomycin, erythromycin and vancomycin. Cases of slight activity are marked in color.

Next, felipeptins A1 and A2 were tested for bioactivity on human cells. Therefore, collaboration with the group of Prof. G. Selivanova at Karolinska Institute was established and both lasso peptides were tested in different human cell lines.

Initially, it was observed that the combinatorial treatment with felipeptins A1 and A2 significantly stimulates the proliferation of colon carcinoma HCT116 cells, especially in a ratio 1:1 in a concentration of 6.25 µM (*Fig. 40*).

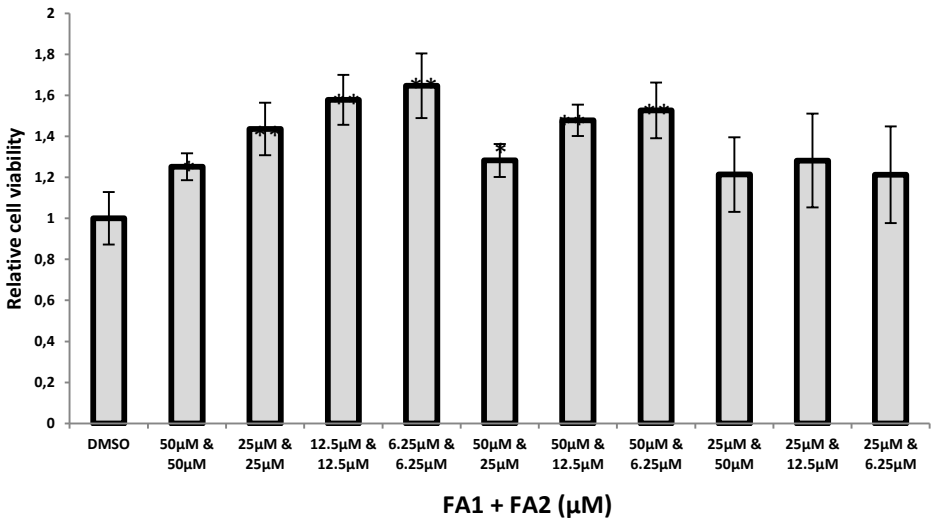


Figure 40. Synergic effect of felipeptins A1 (FA1) (lower position) and A2 (FA2) (upper position) on HCT116 cell viability assay. Resazurin assay after 72 h. DMSO used as control.

When felipeptins A1 and A2 were tested individually on the HCT116 cells, effect on cell proliferation was lower compared to the combinatorial treatment (*Fig. 41*). These results

confirmed, that there is a clear synergistic effect of felipeptins A1 and A2 on the colon carcinoma HCT116 cells.

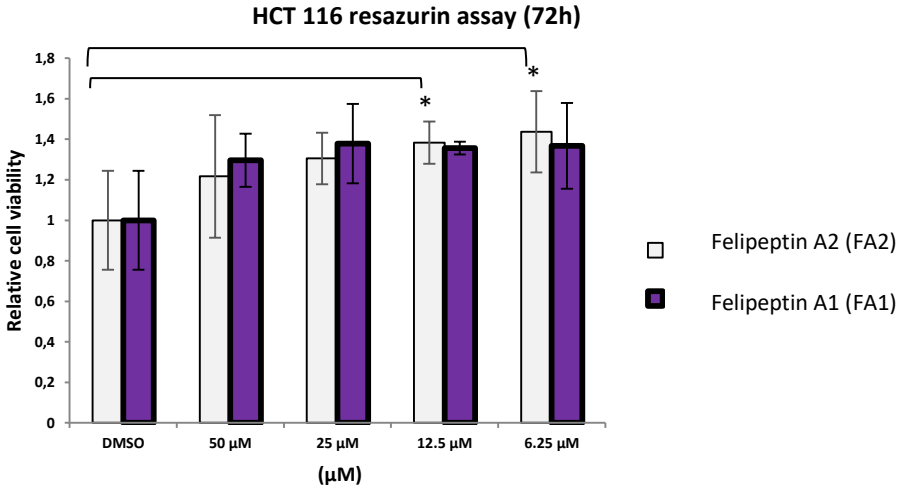


Figure 41. Felipeptins A1 and A2 individual treatments on HCT116 cell viability assay. Resazurin assay after 72 h. DMSO used as control.

The next step was to find out whether felipeptinsA1 and A2 have an effect on normal human cells. To address this question, normal BJ fibroblast were used.

Various concentrations and ratios of felipeptins A1 and A2 in combinatorial treatments were tested and cell viability with the different treatments evaluated. Interestingly, no stimulation of growth was observed; instead, slight inhibition was detected in some cases (*Fig. 42*).

The latter might suggest that the combinatorial treatment of felipeptins A1 and A2 influence cell proliferation of tumor cells without significant effects on normal cells.

In order to further investigate this hypothesis, felipeptins A1 and A2 were tested on a second tumor cell line, A375 melanoma cells.

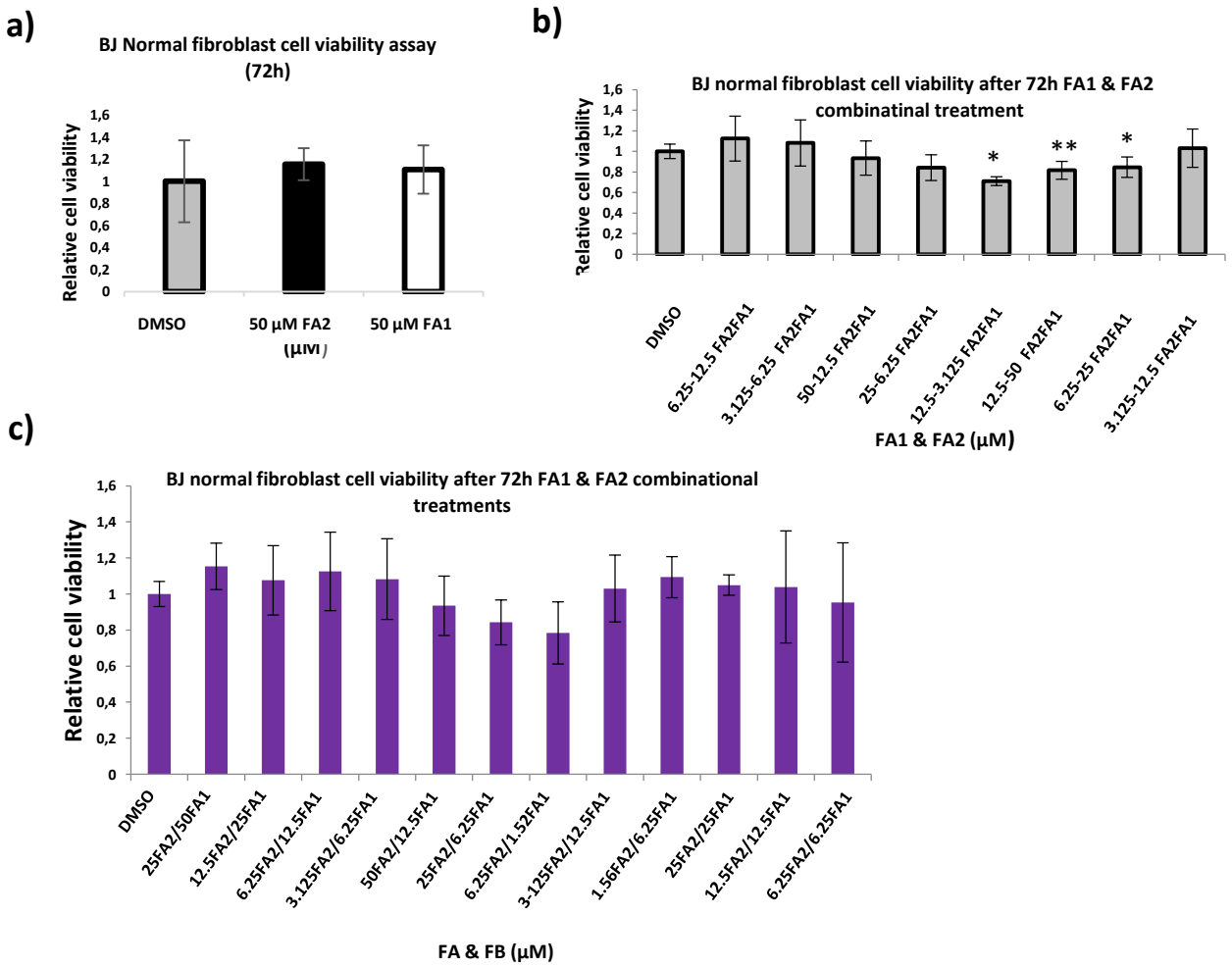


Figure 42. a) Felipeptins A1 (FA1) and A2 (FA2) individual treatments on BJ normal fibroblast cell viability assay. b) and c) Combinatorial treatments of felipeptins A1 and A2 on BJ normal fibroblast cell viability assay. Resazurin assay after 72 h. DMSO used as control.

Initially, felipeptins A1 and A2 were tested individually using different concentrations. Although no significant difference could be detected when compared to the control, it was possible to see a trend of slight stimulation of cell proliferation with both treatments at lower concentration (*Fig. 43*).

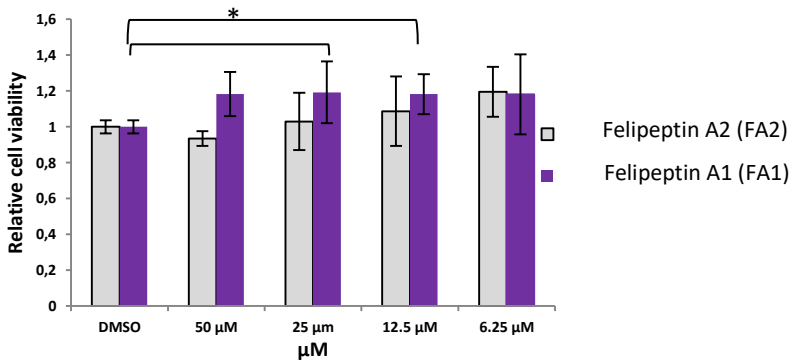


Figure 43. Felipeptins A1 and A2 individual treatments on A375 melanoma cell viability assay. Resazurin assay after 72 h. DMSO used as control.

When felipeptins A1 and A2 were tested in combinatorial treatments in the A375 melanoma cell viability test, it was found out that the cell proliferation significantly increases compared to the control (*Fig. 44*). This result further corroborated the synergistic effect the felipeptins A1 and A2 exert on the cell proliferation of the tumor cell lines.

Interestingly, the best result was obtained when felipeptins A1 12.5 µM and A2 6.25 µM (2:1) were used in combinatorial treatment. It seems that the effect not only depends on the concentration of felipeptins A1 and A2 but also on the ratio. Therefore, this condition was used for further experiments.

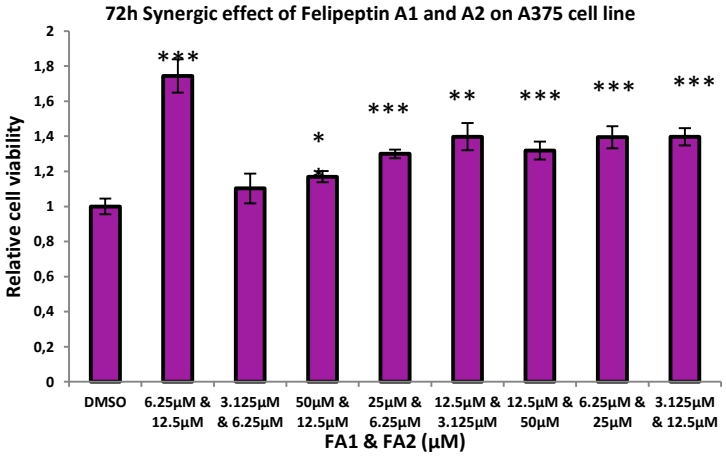


Figure 44. Synergic effect of felipeptins A1 (FA1) (lower position) and A2(FA2) on A375 melanoma cell viability assay. Resazurin assay after 72 h. DMSO used as control.

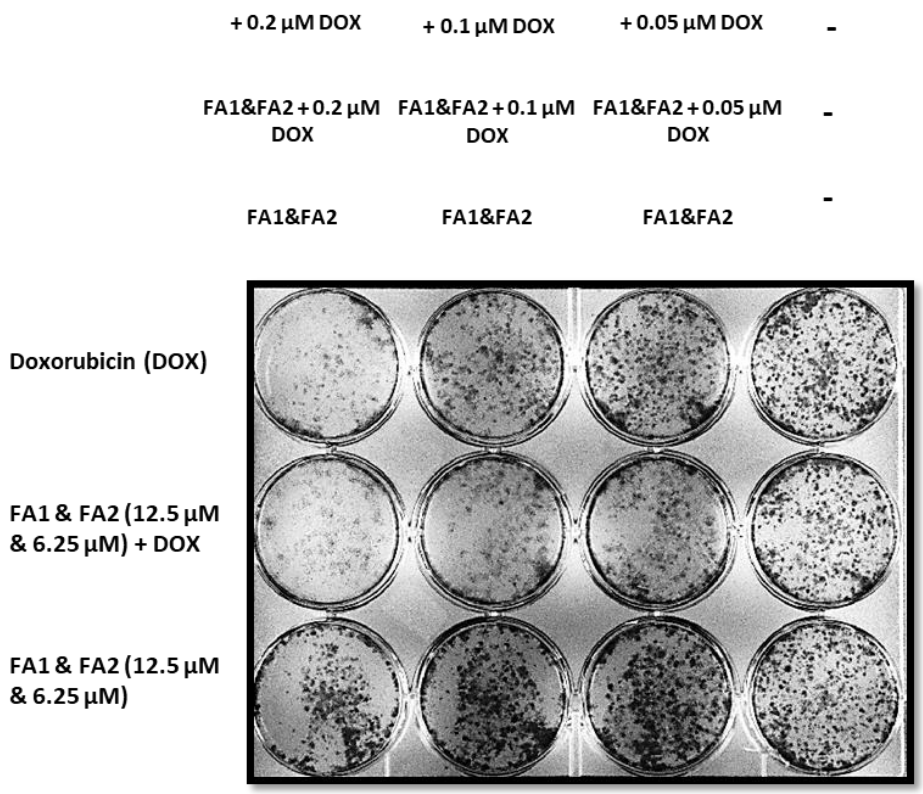


Figure 45. Colony formation assay for combinatorial treatment of felipeptins A1 (FA1) and A2 (FA2) plus doxorubicin (DOX) on A375 melanoma cells. 6 days colony formation assay. DMSO (-) used as control.

After the abovementioned results, it was hypothesized that the combinatorial treatment of felipeptins A1 and A2 may sensitize tumor cells to chemotherapy treatments. The idea behind this was that tumor cells that are in an active proliferation stage would be more sensitive to an anti-tumor drug treatment.

In this regard, doxorubicin, an anti-tumor drug currently used in clinics [104], was selected for the experiments. Doxorubicin binds to DNA-associated enzymes, intercalates DNA and targets multiple molecular targets to produce cytotoxic effects[104]. Colony formation assays were performed to determine the effect of combinatorial treatment of felipeptins A1 and A2 plus doxorubicin on A375 melanoma cells. It was observed that effectively, felipeptins A1 and A2 combinatorial treatment not only stimulates the proliferation of A375 melanoma cells, but clearly sensitizes them to doxorubicin treatment (*Fig. 45*).

To confirm the latest results, the experiment was performed but using different tumor cells, the breast carcinoma MCF7 cell line.

It was observed that the felipeptins A1 and A2 combinatorial treatment increases breast carcinoma MCF7 cell sensitivity to doxorubicin after 5 and 6 days in the colony formation assay (*Fig. 46*).

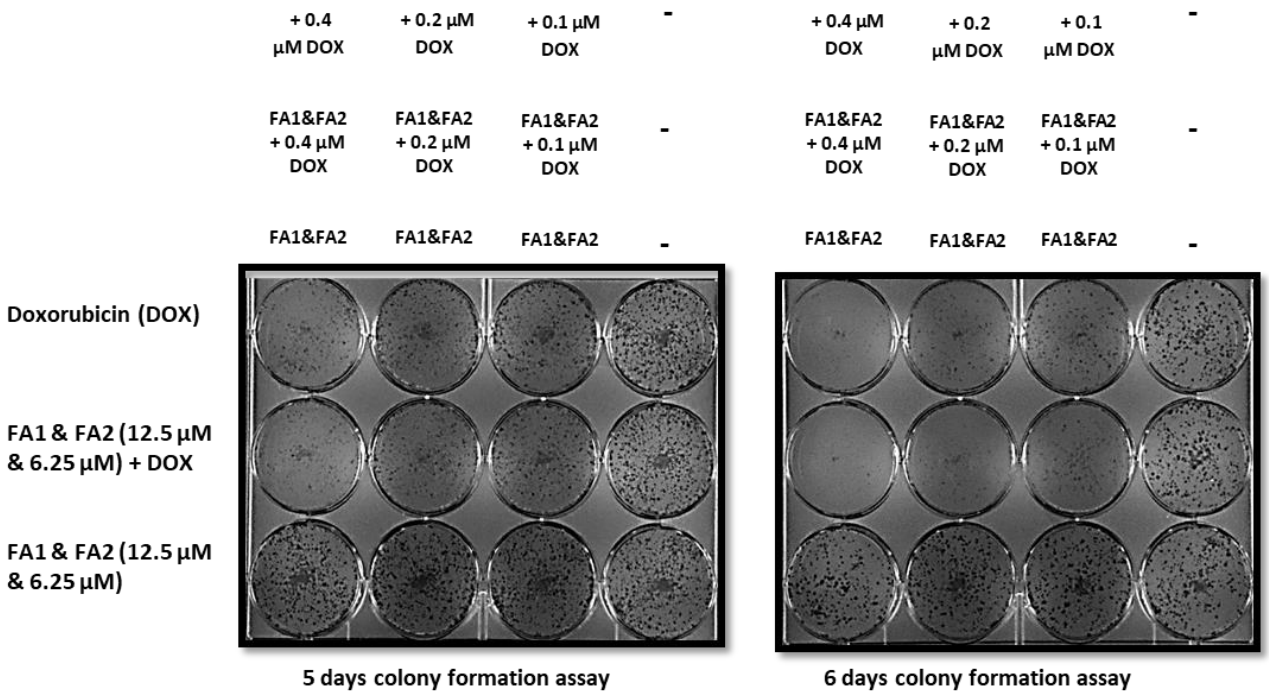


Figure 46. Colony formation assay for combinatorial treatment of felipeptins A1 (FA1) and A2 (FA2) plus doxorubicin (DOX) on breast carcinoma MCF7 cells. 5- and 6-days colony formation assay. DMSO (-) used as control.

Additionally, resazurin assays showed that the combinatorial treatment with felipeptins A1 and A2 increases breast carcinoma MCF7 cell sensitivity to doxorubicin after 5 days (*Fig. 47*).

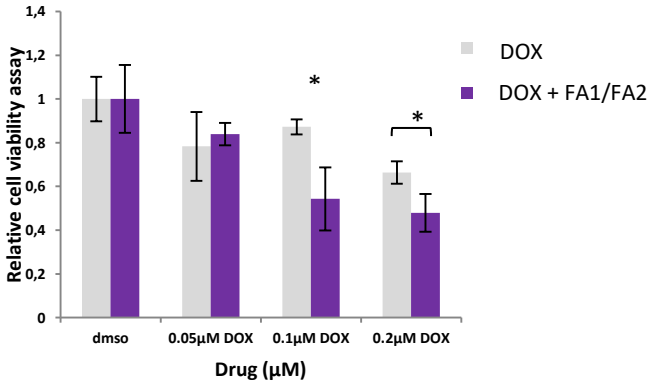


Figure 47. Effect of doxorubicin (DOX) plus felipeptins A1 (FA1) and A2 (FA2) combinational treatment on MCF7 cell viability after 5 days. Resazurin assay. DMSO used as control.

It still remains unclear how felipeptins A1 and A2 specifically stimulate cell proliferation of different tumor cell lines and why the effect depends on the concentration and ratio used. However, it is clear that the mixture exerts a synergistic effect.

5. General discussion

The increase in the incidences of antibiotic resistance among bacterial pathogens and chemotherapy treatment resistance in different types of cancer are two of the major current public health concerns worldwide [13,17].

Despite the fact that natural products (NPs) have historically been proved as an extraordinary source of bioactive compounds, especially those from actinomycete bacteria [25,26], many pharmaceutical companies abandoned their programs on NP research due to different reasons during the last decades [7]. Apparently, the high rates of molecules re-discovery led to the false conjecture that the metabolic potential of microorganisms was exhausted [105].

However, since the beginning of the genomic era, it was demonstrated that microorganisms harbor previously undiscovered genes for secondary metabolite biosynthesis in their genomes and thus their potential as producers of new NPs is far from being exhausted [41]. Moreover, if it is considered that less than 1% of the microbial species has been cultivated, the probability to isolate new strains that can produce eventually novel bioactive molecules with different mechanisms of action is very high [71]. In the current work, strains of actinomycete bacteria belonging to different genera and isolated from unique environments were investigated for the production of novel secondary metabolites. This ”classical” bioprospecting approach led to the identification of several bioactive molecules: the antibiotic angucycline compounds 8-O-methyltetrangomycin produced by *Streptomyces* sp. Go-475, the antifungal cycloheximide and a new bioactive acylated tripeptide produced by *Streptomyces* sp. S1-1-ISP4-01, the broad spectrum antibiotic chlortetracyclines produced by *Kitasatospora* sp. RLA070, and the bioactive 1,2,4-trimethoxynaphthalene produced by *Amycolatopsis* sp. YIM10 and isolated from a bacterium for the first time. Additionally, compounds that we could not associate with a bioactivity with the tests used were identified: several tigloside-related compounds produced by *Amycolatopsis* sp. YIM10, and at least two proposed novel compounds produced by *Kitasatospora* sp. RLA070.

Sequencing and analyses of the above mentioned actinobacteria’s genomes allowed to connect several of their secondary metabolites with their cognate BGCs. Furthermore, many unique BGCs were identified in the analyzed genomes, especially in the strains belonging to

Amycolatopsis and *Kitasatospora* genera. The latter might suggest that further efforts shall to be pursued in terms of rare actinomycete strains isolation and detailed investigation.

A genome mining attempted with *Streptomyces* sp. S1-1-ISP4-01 and focused on the constitutive overexpression of one of the SARP regulators, triggered the production of polar antibiotic compounds. Unfortunately, they could not be purified and characterized in the course of this PhD project. However, it opens the opportunity for future research, especially considering the fact that this antibiotic activity targets Gram-negative bacteria which are the primary concern with regard to antibiotic resistance.

In terms of novelty, the genome of *Amycolatopsis* sp. YIM10 is perhaps the most interesting, because many identified BGCs appear to be unique. Cluster 12 encoding a PKSIII enzyme presumably associated with 1,2,4-trimethoxynaphthalene production, and cluster 21, a novel lasso peptide-like cluster, were identified in a fosmid genome library constructed for *Amycolatopsis* sp. YIM10. Both clusters were subcloned and introduced into different *Streptomyces* hosts for heterologous expression. Initially, heterologous expression failed in both cases. However, the constitutive overexpression of a gene encoding SARP regulator in the lasso peptide cluster allowed the production of two novel lasso peptides named felipeptins A1 and A2. Same strategy was used in an attempt to activate cluster 12, using a LuxR regulator, but no product of this cluster could be identified.

Felipeptins A1 and A2 did not have pronounced antibiotic activity, while a certain degree of synergistic effect was detected in the assays against Gram-positive bacteria. On the other hand, both lasso peptides stimulated the proliferation of different cancer cell lines but did not significantly affect normal fibroblast cells. Interestingly, the major effect was achieved when a combinatorial treatment of felipeptins A1 and A2 was tested. Furthermore, it was observed that the response to the combinatorial treatment depends on the concentration and ratio of felipeptins A1 and A2. Additionally, it was observed that felipeptins A1 and A2 combinatorial treatment not only stimulates the proliferation of cancer cell lines, but clearly sensitizes them to doxorubicin treatment. Considering the latter, felipeptins A1 and A2 provide a very interesting system for future research concerning the mechanism of action, interactions at the molecular level, and potential application in combinatorial treatment of cancer.

Preliminary NMR results obtained by Prof. Aachman suggested that felipeptin A1 and A2 interact and form a complex. Further studies must to be done in this regard, since the structure of an A1+A2 complex is not yet completely solved, and the question of the complex stoichiometry remains open.

The next question to address would be the mechanism of action. Whether felipeptins A1 and A2 interact with the same receptor and/or different receptors in the cell membrane and/or act intracellularly. Given that the combinatorial treatment with felipeptins A1 and A2 sensitizes different tumor cell lines to doxorubicin, it would be interesting to test this combination with

other anti-tumor drugs. Finally, the combinatorial treatment of felipeptins A1 and A2 has to be tested on more tumor cell lines, especially cancer stems cells (CSCs), which are known to acquire a dormancy state during chemotherapy, and therefore show resistance to the current therapies[22]. It might be that the combinatorial treatment with felipeptins A1 and A2 can activate CSCs proliferation as well, making them sensitive to the conventional anti-tumor drug therapies and thus decreasing the chance of tumor relapse and metastasis development.

The results obtained in this project clearly suggest that actinomycete bacteria still remain an extraordinary source of bioactive molecules. However, in order to prevent re-discovery of known molecules, a combination of different approaches must be used. Isolation of new strains from underexplored sources and bioactivity screening combined with various types of rational genome mining and advanced analytical tools must be pursued. In addition, novel bioassays must be designed to find molecules with novel mechanisms of action.

6. Conclusion

This project demonstrates that actinomycete bacteria isolated from unique environments have the potential to produce novel bioactive compounds. In this regard, bioprospecting coupled with LC-MS and NMR analyses allowed the identification and isolation of several bioactive molecules. Furthermore, analysis of the genome sequences led to the identification of many unique BGCs and made possible to connect some of those to the isolated molecules. Finally, the use of genome mining together with genetic engineering allowed the production of two novel bioactive lasso peptides with potential application in cancer chemotherapies.

7. Materials and methods

7.1. Materials

7.1.1. List of microorganisms

Strain	Genus	Origin
GO-475	*Streptomyces*	Soil sample from a desert in Ethiopia
S1-1-ISP4-01	*Streptomyces*	Marine sediment from Pacific ocean
RLA070	*Kitasatospora*	Rhizosphere from Edelweiss
YIM10	*Amycolatopsis*	Soil sample from rare earth mine of Bayan Obo, Inner Mongolia, China

Table 7. Actinomycete bacteria

Strain	Function	Genotype/phenotype
Escherichia coli DH5α	General cloning host	luxS supE44 ΔlacU169 (ϕ80 lacZΔM15) hsdR17, recA1, endA1, gyrA96, thi-1, relA1
Escherichia coli ET12567 (pUZ8002)	Non-methylating strain with pUZ8002 helper plasmid and RP4 oriT for conjugation	dam− , dcm− , hsdM−, KanR, CmlR
Escherichia coli BW25113/pKD46	λ RED recombination system (pKD46)	rrnB3, ΔlacZ4787, hsdR514, Δ(araBAD)567, Δ(rhaBAD)568 rph-1, repA101(ts), araBp-gam-bet-exo, oriR101, AmpR
Escherichia coli EPI300-T1	Genome library construction and overproduction of fosmids.	F– mcrA Δ(mrr-hsdRMS-mcrBC) (StrR), φ80dlacZΔM15, ΔlacX74, recA1, endA1, araD139 Δ(ara, leu)7697, galU, galK, λ– rpsL nupG, trfA, tonA,

		dhfr
Escherichia coli DH10B	Strain used for conjugation	F- mcrA Δ(mrr-hsdRMS-mcrBC), φ80dlacZΔM15, ΔlacX74, endA1, recA1, deoR Δ(ara,leu)7697, araD139, galU, galK, nupG, rpsL λ-
Streptomyces coelicolor M1154	Host for heterologous expression	Δact Δred Δcpk Δcda rpoB[C1298T] rpsL[A262G]
Streptomyces albus J1074	Host for heterologous expression	Wild type

Table 8. Bacteria strains used for molecular biology and hosts for heterologous expression. *All the Escherichia coli strains were grown over the night at 200 rpm and 37 °C, except Escherichia coli BW25113/pKD46 which requires 30 °C for the induction of the λ RED recombination system.

Organism	**Grow conditions**
Escherichia coli DH5α	LB medium, 37 °C, over night
Bacillus subtilis 168	LB medium, 37 °C, over night
Pseudomonas fluorescens ATCC 1352	LB medium, 28 °C, over night
Staphylococcus carnosus DSMZ 20501	TSB medium, 37 °C, over night
Kocuria rhizophila DSMZ 348	LB medium, 28 °C, over night
Enterococcus mundtii DSMZ 4840	TSB medium, 37 °C, over night
Micrococcus luteus DSMZ 1790	TSB medium, 28 °C, over night
Bacillus subtilis DSMZ 10	TSB medium, 28 °C, over night
Erwinia persicina DSMZ 19328	TSB medium, 28 °C, over night
Pseudomonas putida KT2440	LB medium, 37 °C, over night
Saccharomyces cerevisiae BY4743	YPD medium, 28 °C, over night
Aspergillus niger 110.271	PDA medium, 28 °C, 2-3 days
Fusarium graminearum 112.300	PDA medium, 28 °C, 2-3 days

Table 9. Test organism used for bioactivity test

7.1.2. List of media

7.1.2.1. Media for fermentations

5288 medium

Soy flour 10 g, glycerol 15 g, NaCl 5 g, $CaCO_3$ 1 g, $CoCl_2$*$7H_2O$ 1 mg, in 1 L of distilled water, pH 6.8.

SM17 medium

Glucose 2 g, glycerol 40 g, starch 2 g, soy flour 5 g, peptone 5 g, yeast extract 5 g, NaCl 5 g, $CaCO_3$ 2 g, in 1 L of tap water.

GYM medium

Glucose 4 g, yeast extract 4 g, malt extract 10 g, in 1 L of tap water.

ISP2 medium

Yeast extract 4 g, malt extract 10 g, glucose 4 g, in 1 L of tap water.

MYM medium

Maltose 4 g, yeast extract 4 g, malt extract 10 g, MOPS 1.9 g, in 1 L of distilled water.

7.1.2.1 Other media

Luria broth (LB) medium

Tryptone 10 g, yeast extract 5 g, NaCl 5 g, in 1 L of distilled water.

L agar (LA) medium

Agar 10 g, tryptone 10 g, yeast extract 5 g, NaCl 5 g, in 1 L of distilled water.

YPD medium

Peptone 20 g, yeast extract 10 g, glucose 20 g, in 1 L of distilled water.

*Agar 15 g added for solid media.

Tryptone soy broth (TSB) medium

TSB powder (Sigma Aldrich) 30 g in 1 L of distilled water.

2 X YT medium

Tryptone 16 g, yeast extract 10g, NaCl 5 g, in 1 L of distilled water.

Yeast extract-malt extract (YEME) medium

Yeast extract 3 g, peptone 5 g, malt extract 3 g, glucose 10 g, sucrose 340 g (34% final), in 1 L of distilled water.

*After autoclaving, add $MgCl_2.6H_2O$ (2.5 M) 25 mg/L (5 mM final).

Soy flour-mannitol (SFM) medium

Soy flour 20 g in 600 ml of tap water, heat with stirring and boil it during 30 s. Mannitol 20 g in 400 ml of tap water. Mix both solutions; add agar 20 g/L and autoclave.

HA agar medium

Malt extract 10 g, yeast extract 4 g, glucose 4 g, $CaCl_2$ 1.46 g, in 1 L of distilled water, pH 7.3.

Oatmeal agar (OA) medium

Oatmeal 30 g in 1 L of distilled water. Heat with stirring and boil it during 1 min. Add agar 12 g/L and autoclave.

Potato glucose agar (PDA) medium

PDA powder (Sigma Aldrich) 39 g in 1 L of distilled water.

*All media were autoclaved at 121°C for 20 minutes.

7.1.3. Antibiotic solutions

Antibiotic	Abbreviation	Stock solution concentration (mg/mL)	Solvent	Concentration for E. coli (µg/mL)	Concentration for Actinomycetes (µg/mL)
Apramycin	Am	100	Water	100	50
Kanamycin	Kan	50	Water	25-50	25-50
Chloramphenicol	Clm	30	Absolut ethanol	15	25
Ampicillin	Amp	100	Water	100	-
Nalidixic acid	Nal	30	0.1 M NaOH	30	30

Table 10. Antibiotic solutions. *All water based antibiotic solutions were sterile filtered with 0.22 µm pore size sterile filters. Aliquots were store at -20 °C.

7.1.4. Buffer solutions

Solution I for fosmids isolation

50 mM tris/HCl, pH8; 10 mM EDTA (ethylenediaminetetraacetic acid), 100 µg/mL RNase (add fresh RNase solution every time before using solution I).

Solution I for fosmids isolation

200 mM NaOH; 1% SDS (sodium dodecyl sulfate).

Solution III for fosmids isolation

3 M potassium acetate, pH 5.5.

Tris HCL 100 mM buffer

12.114 g Tris-base, 56.86 mL 1 M HCl, in 1 L of distilled water, pH 8.0.

Phosphate solution for dialysis

50mM sodium phosphate (1 M Na_2HPO_4 46.6 mL; 1 M NaH_2PO_4 3.4 mL), 100 mM NaCl, in 1 L of distilled water, pH 8.0.

TBE buffer 10x

Tris base 108 g, boric acid 55 g, 0.5 M EDTA (pH 8.0), in 1 L of distilled water.

TSS buffer

PEG (polyethylene glycol) 8000 5 g, 1 M $MgCl_2$ 1.5 mL, DMSO (Dimethyl sulfoxide) 2.5 mL, in 50 mL of LB medium.

7.1.4.1. Other solutions

Agarose gel (0.8%) for DNA electrophoresis

TBE buffer (1x) 400 mL, Agarose 3.2 g, Gel Red (10000x) 20 µL.

7.1.5. Vectors list

Vector	Function	Length (Kb)	Characteristics
pCC1FOS (Epicentre)	Genome library construction. Fosmid used for heterologous expression.	8.1	ClmR, lacZ, loxP, cos, parC, parB, parA, repE, ori2, oriV, redF
pSOK806	Integrative plasmid with strong ermE promoter.	5.7	ermE*p, ColEI oriR, AmR, , $oriT^{RP4}$, int-$attP^{VWB}$
pSET152	Integrative plasmid. Use as a control for heterologous expression.	5.5	lacZα ori^{pUC19}, $oriT^{RP4}$, int-$attP^{\varphi C31}$, aac(3)IV

Table 11. Vectors list

7.1.6. List of primers

Oligo name	Sequence (5´-3´)	Characteristics
27F	AGAGTTTGATCMTGGCTCAG	16S rRNA gene forward sequencing primer
1492R	CGGTTACCTTGTTACGACTT	16S rRNA gene reverse sequencing primer
pCC1FOS-FP	GGATGTGCTGCAAGGCGATTAAGTTGG	pCC1FOS forward sequencing primer
pCC1FOS-RP	CTCGTATGTTGTGTGGAATTGTGAGC	pCC1FOS reverse sequencing primer
Fw_Cluster21_A	TTGGTGCGCTGGATGTC	Lasso peptide cluster screening part A forward sequencing primer
Rv_Cluster21_A	TCTGTGGTGGCTGATTC	Lasso peptide cluster screening part A reverse sequencing primer
Fw_Cluster21_B	AGTACCGCCGCGTATTC	Lasso peptide cluster screening part B forward sequencing primer
Rv_Cluster21_B	AGCTACCGCCTGAAGTG	Lasso peptide cluster screening part B reverse sequencing primer
Fw_Cluster21_C	GGCGGATCAGGTAGATG	Lasso peptide cluster screening part C forward sequencing primer
Rv_Cluster21_C	ACGTCAGCGAAGCCATC	Lasso peptide cluster screening part C reverse sequencing primer

SARP_Fw	GTCA***GAATTC***GTGGCAATCGCGTTGCAC	SARP regulator forward sequencing primer (EcoRI restriction site)
SARP_Rv	GTCA***GATATC***GCGACGTACGACGACGAATC	SARP regulator reverse sequencing primer (EcoRV restriction site)

Table 12. List of primers. *In bold/italic letters are the endonuclease restriction site sequences.

7.1.7. Commercial kits

Genomic DNA extraction

Wizard® SV Genomic DNA Purification System (Promega, enhanced).

Isolation of plasmid

DNA Wizard® Plus SV Minipreps DNA Purification System (Promega).

Purification of DNA

DNA Clean & Concentrator™-5 (Zymo Research).

DNA recovery from gel

Zymoclean™ Gel DNA Recovery Kit (Zymo Research).

Genome library construction

CopyControl™ Fosmid Library. Production Kit with pCC1FOS™ Vector (epicentre).

Lasso peptide dialysis

20 mL 1 kDa MWCO Mega Pur-A-Lyzer kit (Sigma-Aldrich).

7.1.8. Enzymes

Enzyme	Company	Function
Taq polymerase	Provided by the laboratory	PCR
Phue polymerase	Provided by the laboratory	PCR
Q5 polymerase	New England Biolabs	PCR
EcoRV HF	New England Biolabs	Enzyme restriction assay
EcoRI HF	New England Biolabs	Enzyme restriction assay
T4 DNA ligase	New England Biolabs	Ligation reaction
RNAse A	New England Biolabs	RNA degradation
Lysozyme	Sigma Aldrich	Bacterial cell membrane disruption

Table 13. Enzymes.

7.2. Methods

7.2.1. Cultivation and storage of Actinomycetes bacteria

For growth of Actinomycetes bacteria different solid media were used depending on the strain.

Strain	**Genus**	**Solid medium**	**Incubation time (days)**
GO-475	*Streptomyces*	SFM	7-10
S1-1-ISP4-01	*Streptomyces*	SFM	4-7
RLA070	*Kitasatospora*	OA	4-7
YIM10	*Amycolatopsis*	SFM or HA	10-14
Streptomyces coelicolor M1154 and mutants	*Streptomyces*	SFM	7-10
Streptomyces albus J1074 and mutants	*Streptomyces*	SFM	7-10

Table 14. Actinomycetes bacteria growth conditions in solid media

An inoculum was taken with a sterile loop from a dense spore's suspension glycerol stock and carefully spread on 10 cm diameter sterile Petri dishes containing approximately 30 mL of corresponding solid medium. The plates were incubated at 28 °C for a period of time depending on the strain.

7.2.1.1. Spore's suspension glycerol stock preparation

5 mL of sterile glycerol 20% were pipetted onto a plate containing sporulated Actinomycetes bacteria.

The surface of the plate was gently scratched with the tip of a sterile 5 mL pipette to suspend the spores in glycerol. The suspension was filtered through a sterile syringe containing cotton wool into a sterile tube to get rid of the mycelium. Aliquots of the resulting spore suspension were stored in Cryo tubes at -80 °C.

7.2.1.2. Seeding culture preparation

200 µL from a dense spore's suspension glycerol stock were transferred to a sterile 100 mL flask, containing 15 mL of liquid medium, and were incubated at 200 rpm and 28 °C. Different incubation times were used depending on the strain.

Strain	**Genus**	**Medium**	**Incubation time (days)**
GO-475	*Streptomyces*	TSB	2
S1-1-ISP4-01	*Streptomyces*	TSB	1
RLA070	*Kitasatospora*	TSB	1
YIM10	*Amycolatopsis*	TSB	10
Streptomyces coelicolor M1154 and mutants	*Streptomyces*	YEME	5
Streptomyces albus J1074 and mutants	*Streptomyces*	YEME	1

Table 15. Actinomycete bacteria growth conditions in seed culture

7.2.1.3. Fermentation procedure

A sterile 250 mL baffled flask, containing 50 mL of a certain fermentation medium, was inoculated with 3 mL of seeding culture and incubated for 200 rpm and 28 °C. Fermentation medium and time of incubation depends on the strain.

Strain	**Genus**	**Medium**	**Incubation time (days)**
GO-475	*Streptomyces*	5288	7
S1-1-ISP4-01	*Streptomyces*	SM17, GYM	7
RLA070	*Kitasatospora*	ISP2	7
YIM10	*Amycolatopsis*	5288	14
Streptomyces coelicolor M1154 and mutants	*Streptomyces*	MYM	7
Streptomyces albus J1074 and mutants	*Streptomyces*	MYM	7

Table 16. Actinomycete bacteria growth conditions in fermentative media

7.2.1.4. Up-scaled fermentations

For up-scaled fermentation conditions the parameters were increased 10 times. A sterile 2 L baffled flask, containing 500 mL of fermentation medium, was inoculated with 30 mL of seeding culture and incubated under same conditions as for normal fermentation volume.

*Up-scaled cultures of *Amycolatopsis* YIM10 strain had different volumes depending on the production of certain secondary metabolite. 500 mL for the production of tigloside compound and 250 mL for the production of 1,2,4-Trimethoxynaphthalene compound.

7.2.2. Extraction of secondary metabolites from cultures

Cultures were extracted to concentrate secondary metabolites under reduced pressure conditions generating crude extracts.
The whole culture (50 mL) was transferred to a 100 mL round-bottom flask and frozen using a cold ethanol bath.
After freezing, the flask was connected to freeze-drier equipment. The material was completely freeze-dried approximately after 24 h.
30 mL of pure methanol was used to extract a dry culture (50 mL). The solvent was added to the 100 mL round flask containing the freeze-dried material and extracted for 1 h at 100 rpm and room temperature.
The extract was then transferred to a 50 mL falcon tube and centrifuged for 10 min at 4000 rpm and room temperature.
The supernatant was depleted to a cleaned 100 mL flask and the solvent was evaporated on a rotary evaporator at 40 °C.
The dried extract was dissolved in 5 mL pure methanol and the crude extract was centrifuged for 10 min at 4000 rpm and room temperature. The methanol extracts were stored at -20 °C until further analysis.

7.2.2.1. Up-scaled extraction

A 500 mL culture was centrifuged for 10 min at 4000 rpm and room temperature.
The pellet was extracted with 300 mL of pure methanol and briefly vortexed. Then it was extracted for 1 h at 100 rpm and room temperature.
The extract was then transferred to 50 mL falcon tubes and centrifuged for 10 min at 4000 rpm and room temperature.
The supernatant was depleted to a cleaned 500 mL flask and the solvent was evaporated on a rotary evaporator at 40 °C.
The dried extract was dissolved in 10 mL pure methanol and centrifuged for 10 min at 4000 rpm and room temperature. The methanol extracts were stored at -20 °C until further analysis.

7.2.3. Microorganisms cell suspension glycerol stock preparation

For cultivation of *E. coli* strains or test organisms for antimicrobial tests, 100 µL of cell suspension glycerol stock were fractionated spread on a 10 cm plate containing 25 mL of medium and appropriated antibiotics if needed.
After incubation, a single colony was picked using a toothpick and transferred to 3-10 mL of liquid medium and appropriated antibiotics if needed for propagation. After incubation at 200 rpm, the culture was centrifuged 10 min at 4000 rpm and room temperature. The cell pellet was suspended in 1-5 mL sterile glycerol 20%. Aliquots of the resulting microorganism cell suspension were stored in Cryo tubes at -80 °C.
* The filamentous fungi were scratched from the solid medium plate and suspended in sterile glycerol 20%. Aliquots were stored in Cryo tubes at -80 °C.

7.2.4. Paper disc diffusion assays for antimicrobial test

To test the antimicrobial activity of crude extracts, fractions and pure compounds, growth inhibition assays were performed.
200 µL of test organism cell suspension were spread on the corresponding solid media. Then 50 µL of extract were pipetted on a sterile paper discs and let them dry in a sterile fume hood. As negative controls 50 µL of solvent and 50 µL of medium crude extract were used.
The dry paper discs were placed on the inoculated plate and incubated depending on the test microorganism's growth conditions.

After incubation, the plates were check. It was considered as antimicrobial activity when a halo of growth inhibition surrounded a paper disc appeared. A clear halo was considered as a complete growth inhibition. A blurred halo was considered as a partial growth inhibition. The halo's diameter was measured in millimeter.

7.2.5. Flash chromatography fractionation

The crude extracts (approximately 500 mg) were mixed with Silica gel 60 from MERCK (particle size 0.063–0.200 nm for column chromatography) in a proportion 1:2 and packaged in cartridge. For fractionation the PuriFlash column 15 C18 HQ 35G−35.0 g (22 bar) was used with a gradient of 5–98% Methanol/H2O in 50min and a flow of 15 mL/min using an equipment PuriFlash 420. Fractions obtained were collected, freeze dried, and concentrated in methanol for antimicrobial bioassays.

7.2.6. High Performance Liquid Chromatography (HPLC) analysis

All crude extracts, fractions and pure compounds obtained during this project were analyzed using a Shimadzu HPLC system LC-20AT.

Some of the separations were carried out on an Acclaim 120 C18, 2.1 mm x 150 mm, 3 µm HPLC column (Dionex/Thermo Fisher Scientific) using 0.1% aqueous formic acid and acetonitrile as mobile phase A and B, respectively. The general gradient used in this project was: 5–95% B in 45 min followed by a washing (10 min at 95% B) and re-equilibration step (10 min at 5% B), the gradient was slightly modified depending on the compound of interest. The flow rate was 0.5 mL/min and the column oven temperature was set to 25°C.

Some extracts were separated using a Luna C18, 4.6 mm x 250 mm, 5 µm HPLC column (Phenomenex). HPLC separation conditions were exactly the same as with the Acclaim 120 C18. The only difference was a 1.0 mL/min flow rate.

PDA (photo diode array) detector, which scans from 190 – 800 nm and an ELSD (Evaporative Light Scattering Detector) were used.

7.2.7. Semi-preparative HPLC

Two semi-preparative HPLC equipment were used to purify some secondary metabolites during this project.

7.2.7.1. Semi-preparative HPLC Shimadzu LC-8A

Tigloside compound isolation

For the tigloside compound purification, 2 mL of crude extract was separated using a Hibar® 250-25 LiChrospher® 100 RP-18e, 5 µm column (MERK). 0.1% aqueous formic acid and acetonitrile were used as mobile phase A and B, respectively. The gradient used was: 42–65% B in 46 min followed by a washing (10 min at 95% B) and re-equilibration step (10 min at 42% B). The flow rate was 15.5 mL/min. Detection wavelength was 190 nm.

7.2.7.2. Semi-preparative HPLC Shimadzu LC-20AR

Felipeptins A1 and A2 isolation

In the case of felipeptins A1 and A2 purification, 2 mL of lasso peptides concentrated solution was separated using a Shim-pack GIS C18, 250 mm x 20 mm, 5 µm column (Shimadzu). 0.1% aqueous formic acid and acetonitrile were used as mobile phase A and B, respectively. The gradient used was: 10–55% B in 20 min followed by a washing (10 min at 95% B) and re-equilibration step (10 min at 10% B). The flow rate was 20 mL/min. Detection wavelength was 190 nm.

Acylated tripeptide isolation

For the purification of the acylated tripeptide molecule produced by the S1-1-ISP4-01 *Streptomyces* strain, 2 mL of F8 bioactive fraction was separated using the same column and solvent system as with Felipeptins A and B purification. The gradient used was: 18–22% B in 20 min followed by a washing (10 min at 95% B) and re-equilibration step (10 min at 18% B). The flow rate was 20 mL/min. Detection wavelength was 190 nm.

Chlortetracyclines compounds fractionation

Sub-fractions containing chlortetracyclines were obtained by separation of 1 mL bioactive F6 fraction (*Kitasatospora* RLA070) using a Luna C18, 10 mm x 250 mm, 5 µm HPLC semi preparative column (Phenomenex). 0.1% aqueous formic acid and acetonitrile were used as mobile phase A and B, respectively. The gradient used was: 5–95% B in 45 min followed by a washing (10 min at 95% B) and re-equilibration step (5 min at 10% B). The flow rate was 5 mL/min. Detection wavelength was 190 nm.

7.2.8. Liquid Chromatography-Mass Spectrometry (LC-MS)

The Liquid Chromatography-Mass Spectrometry (LC-MS) analyses were done by collaboration with Dr. Martin Zehl at Department of Analytical Chemistry, Faculty of Chemistry, University of Vienna, Vienna, Austria.

Samples were first analyzed by high-performance liquid chromatography (HPLC) with charged aerosol detection (CAD) and HPLC-MS. These analyses were performed on an UltiMate 3000 RSLC-series system (Dionex/Thermo Fisher Scientific, Germering, Germany) coupled in parallel to a Corona ultra RS charged aerosol detector (CAD, Dionex/Thermo Fisher Scientific) and an HCT 3D quadrupole ion trap mass spectrometer equipped with an orthogonal ESI source (Bruker Daltonics, Bremen, Germany). Separation was carried out on an Acclaim 120 C18, 2.1 mm × 150 mm, 3 µm HPLC column (Dionex/Thermo Fisher Scientific) as it was explained before on HPLC section.

After passing the DAD, the eluate flow was split 4:1 between the CAD and the MS, respectively. The CAD

nebulizer temperature was 35 °C and the ESI ion source was operated as follows: capillary voltage: +3.5/−3.7 kV, nebulizer: 26 psi (N_2), dry gas flow: 9 L/min (N_2), and dry temperature: 340 °C.

Positive and negative ion mode multistage mass spectra up to MS^3 were obtained in automated data-dependent acquisition (DDA) mode using helium as collision gas, an isolation window of Δm/z = 4, and a fragmentation amplitude of 1.0 V.

In the next step, high-resolution mass spectra were recorded on a maXis HD ESI-Qq-TOF mass spectrometer (Bruker Daltonics) that was also connected to an UltiMate 3000 RSLC-series system. The separation was performed with the above described HPLC methods. The eluate flow was split approximately 1:8 and the following ESI ion source settings were applied: capillary voltage: ±4.5 kV, nebulizer: 0.8 bar (N_2), dry gas flow: 7.0 L/min (N_2), and dry temperature: 200 °C. The sum formulas of the detected ions were determined using

Bruker Compass DataAnalysis 4.2 based on the mass accuracy (Δm/z ≤ 10 ppm) and isotopic pattern matching (SmartFormula algorithm).
The Dictionary of Natural Products, which can be found on the web under: http://dnp.chemnetbase.com/, were used to identify putative microbial compounds by comparing the obtained masses and sum formulas with this database.

7.2.9. Nuclear magnetic resonance (NMR)-Based structure elucidation

The Nuclear magnetic resonance (NMR)-Based structure elucidation for tigloside, 1,2,4-Trimethoxynaphthalene, 8-O-methyltetrangomycin and 8-O-methyltetrangulol molecules were done by collaboration with Prof. Dr. Ernst Urban at Department of Pharmaceutical Chemistry, University of Vienna, Vienna, Austria.
NMR spectra were recorded on a Bruker Avance 500 NMR spectrometer (UltraShield) using a 5 mm switchable probe (TCI Prodigy Kryo-probe head, 5 mm, tripel resonance-inversdetection probe head) with z axis gradients and automatic tuning and matching accessory (Bruker BioSpin). The resonance frequency for 1H NMR was 500.13 MHz and for 13C NMR 125.75 MHz. All measurements were performed for a solution in fully deuterated chloroform or methanol at 298 K. Standard 1D and gradient-enhanced (ge) 2D experiments, like double quantum filtered (DQF) COSY, HSQC, and HMBC, were used as supplied by the manufacturer. Chemical shifts are referenced internally to the residual, non-deuterated solvent signal for chloroform 1H (δ 7.26 ppm) or methanol (δ 3.31 ppm) and to the carbon signal of the solvent for chloroform 13C (δ 77.00 ppm) or methanol (δ 49.00 ppm). Chemical shifts for compounds analyzed in this project, as well as recorded NMR spectra are given in Supplementary Materials.

7.2.10. Lasso peptides NMR structure elucidation

The NMR structures of felipeptins A1 and A2 were done by collaboration with Prof. Dr. Finn Lillelund Aachmann at Department of Biotechnology and Food Science, Norwegian University of Science and Technology (NTNU), Trondheim, Norway.
The sample was prepared by dissolving 0.5-1.2 mg felipeptins A1 or A2 in 160 µL DMSO-d6 (99.9% D; Isotec, USA) and for Felipeptide A1 also 95% methanol-d4 (99.8% D; CIL, USA) and 5% formic acid-d2 (99%+ D, Sigma-Aldrich/Merck). 1.5 mg of compound in 160 µL DMSO-d6 (99.9% d). Hereafter the samples were transferred to a 3 mm NMR tube. All homo and heteronuclear NMR spectra were recorded on a Bruker Ascend 800 MHz Avance III HD NMR spectrometer (Bruker BioSpin AG, Fälladen, Switzerland) equipped with 5mm with

cryogenic CP-TCI probe. All NMR recording were performed at 25 °C. Shifts were determined relative to TMS, using the residual DMSO or methanol signals for spectra calibration. For chemical shift assignment of compound, the following spectra were recorded: 1D proton, 2D double quantum filtered correlation spectroscopy (DQF-COSY), 2D total correlation spectroscopy (TOCSY) with 70 ms mixing time 2D nuclear Overhauser effect correlation spectroscopy (NOESY) with 100 ms mixing time, 2D 13C heteronuclear single quantum coherence (HSQC) with multiplicity editing, 2D 13C HSQC-[1H,1H]TOCSY with 70 ms mixing time on protons, 2D heteronuclear multiple bond correlation (HMBC) with BIRD filter to suppress first order correlations and 2D 15N HSQC. The spectra were recorded, processed and analyzed using TopSpin 3.5 software (Bruker BioSpin).
The structures are calculated with CYANA based on distances extracted from the NOE assignment. The structures have been energy refined with YASARA (www.YASARA.org)

7.2.11. Genomic DNA (gDNA) isolation from Actinomycetes bacteria

Genomic DNA was isolated with Wizard® Genomic DNA Purification Kit (Promega).
This protocol was designed to get gDNA for doing PCR. However, it can be also used to get enough gDNA for genome sequencing. If the latest is the case, all amounts must to be up-scaled 10 times.
Each strain was cultivated to produce 1 mL seeding culture. Then it was centrifuged 5 min at 10000 rpm and 4 °C. The pellet was washed twice with 0.5 mL 50 mM EDTA and centrifuged as in last step.
The whole gDNA isolation procedure can be found in detail in the Kit protocol.
The only modification to the protocol was done when the culture was up-scaled for getting gDNA for genome sequencing. In this case an additional step was implemented. It was necessary to wash twice the aqueous phase after the protein precipitation step with 3 mL phenol-chloroform-isoamylalkohol (25:24:1, v/v; Rotiphenol®, Roth, Karlsruhe).

7.2.12. Genome sequencing

The genomes sequenced during this project were done by collaboration with the group of Prof. Dr. Jörn Kalinowski at Center for Biotechnology, Bielefeld University, Bielefeld, Germany.

7.2.13. DNA sequencing

All DNA sequencing services (except genome sequencing) were done by the company Eurofins Genomics.

7.2.14. Primer design

The software Clone Manager was used to design primers and all of them were synthetized by Eurofins Genomics.

7.2.15. Polymerase Chain Reaction (PCR)

All PCR performed during this project were carried out on a Master Cycler Nexus X2 from Eppendorf® thermocycler.
Parameters were adjusted depending on the primers
Setup of PCR reaction (40 µL)

Component	**Volume (µL)**
Taq buffer (10x)	4
dNTP's solution mix	0.5
27F primer (10 mM)	1
1492R primer (10 mM)	1
Taq DNA polymerase	0.5
Sample gDNA	1
dH_2O	32

Table 17. Setup of PCR reaction (40 µL) for 16S rRNA gene

PCR condition's program

Stage	Temperature (°C)	Time (sec)	Cycles
Initial denaturation	95	120	1
Denaturation	95	30	30
Annealing	60	30	
Elongation	72	180	
Final extension	72	300	1
Hold	4	∞	1

Table 18. PCR condition's program for 16S rRNA gene

7.2.15.2. Genome library screening by PCR

For screening of lasso peptide cluster 21 in genome library 3 pairs of primers were used:
Fw_Cluster21_A and Rv_Cluster21_A primers to amplify part A (659 bps).
Fw_Cluster21_B and Rv_Cluster21_B primers to amplify part B (645 bps).
Fw_Cluster21_C and Rv_Cluster21_C primers to amplify part C (792 bps).

Setup of PCR reaction (40 µL)

Component	Volume (µL)
Taq buffer (10x)	4
dNTP's solution mix	0.5
Fw primer (10 mM)	1
Rv primer (10 mM)	1
Taq DNA polymerase	1
Phu polymerase	0.3
DMSO	2
Sample gDNA	2
dH_2O	28.2

Table 19. Setup of PCR reaction (40 µL) for genome library screening

PCR condition's program

Stage	Temperature (°C)	Time (sec)	Cycles
Initial denaturation	95	120	1
Denaturation	95	30	25
Annealing	62	45	
Elongation	72	120	
Final extension	72	420	1
Hold	4	∞	1

Table 20. PCR condition's program for genome library screening

7.2.15.3. SARP regulator PCR

Setup of PCR reaction (40 µL)

Component	Volume (µL)
Q5 reaction buffer (5x)	8
Q5 High GC Enhancer (5x)	8
SARP_Fw primer (10 mM)	1
SARP_Rv primer (10 mM)	1
dNTP's solution mix	2
Sample DNA (pCC1-FOS-Lasso-intneo)	2
dH_2O	17

Table 21. Setup of PCR reaction (40 µL) for SARP regulator

PCR condition's program

Stage	Temperature (°C)	Time (sec)	Cycles
Initial denaturation	95	120	1
Denaturation	95	30	25
Annealing	62	45	
Elongation	72	240	
Final extension	72	600	1
Hold	4	∞	1

Table 22. PCR condition's program for SARP regulator

7.2.16. PCR-product Purification

All PCR products were purified either for enzyme restriction assays or for DNA sequencing. Different purification kits were used:

• Purification of specific amplification fragments from PCR reaction:
When the PCR reaction specifically amplified a single PCR product, it was used the DNA Clean & Concentrator™ kit (Zymo Research). This procedure was also used to concentrate PCR fragments after restriction assay analysis. For detailed procedure check the protocol of this kit.

• Purification of PCR product/DNA fragments from agarose gel:
If the PCR reaction amplified non-specifically DNA fragments, it was necessary to purify the correct fragment. Therefore, the PCR fragments had to be separated by electrophoresis in an agarose gel and the right size DNA fragment was cut from the gel. The fragment was purified therefore with Zymoclean™ Gel DNA Recovery Kit (Zymo Research). The kit was also used to purified other DNA fragments after enzyme restriction reactions (i.e. Plasmids). For detailed procedure check the protocol of this kit.

All DNA fragments were stored at -20 °C.

7.2.17. Plasmid DNA Isolation from *E. coli*

For isolation of plasmid DNA from *E. coli* strains, the pellet of 1.5 mL of overnight culture was obtained. The further steps were done using the Wizard® Plus SV Minipreps Purification System (Promega).
Detailed information of this procedure can be check in the protocol of this kit.

7.2.18. Isolation of fosmid DNA of *E.coli*

A large colony was inoculated in 5 mL LB medium containing the appropriate antibiotics. It was grown overnight at 37°C, shaking at 200 rpm. Cells are recovered by centrifugation at 13000 rpm for 1 min at 4 °C in a micro centrifuge. Repeat to recover 4 mL culture in a 2 mL tube.
The cell pellet was suspended by vortexing in 200 µL ice-cold solution I (50 mM Tris/HCl, pH 8; 10 mM EDTA, 100 µg/mL RNase (add RNase after autoclaving into small aliquots and store at 4°C)).
Immediately 400 µL solution II were added (200 mM NaOH; 1% SDS, store at room temperature) and mixed by inverting the tubes 10 times.
Immediately 300 µL of ice-cold solution III were added (3 M potassium acetate, pH 5.5) and mixed by inverting the tubes 5 times.
The mixture was centrifuged at full speed in a micro centrifuge for 5 min at 4°C.
The clear supernatant was transferred into a fresh tube, and 1 Vol. phenol-chloroform-isoamylalkohol (25:24:1, v/v; Rotiphenol®, Roth, Karlsruhe) (work in a hood with Phenol) added. The mixture was vortexed for 1 min. It was centrifuged at full speed in a micro centrifuge for 5 min at 4°C and the clear supernatant was transferred into a new tube and it was added 0.8 Vol. 2-propanol. Tubes were incubated on ice for 10 min. The mixture was spined as above and the pellet washed with 500 µL 70% ethanol.
Sample was spined as above and the tube was opened for 5 min at room temperature to dry the pellet. Pellet was suspended in 100 µL 10mM Tris/HCl (pH 8) and used 2 - 8 µL for restriction digest. It is important to remove all ethanol, as it would inhibit enzymes used on the DNA in following steps (polymerases, restriction enzymes, etc.)
For isolation of fosmid DNA from *E.coli* Epi 300 cells you need to induce the replication of the fosmid with “inducing solution” (L-arabinose solution) from the manufacturer. Follow the instructions from the manufacturer for this. If a cooling centrifuge is not used, keep all solutions and all tubes on ice in between the centrifugation steps (not solution II). Solution II has to be renewed after 3 months at the latest.

7.2.19. Enzyme restriction digestion

Enzyme restriction digestions were performed with restriction endonucleases.

Isolated plasmids or purified PCR products were incubated at 37 °C during 3 h with the restriction enzymes. After cutting, enzymes were inactivated at 65 °C for 20 min.

Component	**Reaction mix for plasmid DNA (µL)**	**Reaction mix for PCR-product (µL)**
DNA sample	10	5
Enzyme 1	1	1
Enzyme 2	1	1
Cut Smart buffer (10x)	2	2
Nuclease free H_2O	6	11
Total volume	20	20

Table 23. Setup of enzyme restriction digestion reaction

7.2.20. Ligation reaction with T4-DNA ligase

Enzyme digested plasmid fragment and insertion fragment with complementary restriction sites could be cloned by a ligation reaction. Concentration ratio vector/fragment (1:10) was estimated by gel electrophoresis. The ligation reactions took place in 1.5ml Eppendorf® tubes in a rack, which was incubated on an ice box overnight.

Component	**Volume (µL)**
Cut DNA vector	0.5
Cut DNA fragment	5
T4 DNAL buffer	2
T4 DNA ligase	1
Nuclease free H_2O	11.5
Total volume	20

Table 24. Setup of ligation reaction with T4-DNA ligase

7.2.21. Preparation of chemically competent cells

E. coli chemically competent cells were prepared to transform plasmids. During this project two strains were used; *E. coli* DH5α and *E. coli* ET12567 (pUZ8002).
0.5 ml of an overnight culture (10 mL) was used to inoculate 50 mL LB medium in a 250 mL baffled flask.
The culture was cultivated for 2-3 h at 37 °C shaking at 200 rpm to an OD_{600} of approximately 0.6.
Culture was divided in two 50 mL falcon tubes and kept on ice for 10 min. Then they were centrifuged at 3000 rpm for 10 min at 4°C.
Supernatant was removed, and the cell pellet was suspended in 2.5 mL ice cold TSS buffer each.
100 µL aliquots of the resulting suspension were pipetted fast into ice cold 1.5 mL eppendorf tubes and immediately frozen in liquid nitrogen before storing at -80 °C.

7.2.22. Preparation of electrocompetent cells

E. coli electrocompetent cells were prepared to transform either fosmids, plasmids or DNA cassettes. During this project several *E. coli* strains were used.
0.5 ml of an overnight culture (10 mL) was used to inoculate 50 mL LB medium plus antibiotics if it was needed in a 250 mL baffled flask.
The culture was cultivated for 2-4 h at 37 °C (30 °C for *E. coli* BW25113/pKD46) shaking at 200 rpm to an OD_{600} of approximately 0.6.
Culture was centrifuged at 4000 rpm for 5 min at 4°C. Supernatant was decanted and cell pellet was gentle mixed in 30 mL ice-cold 10% glycerol. It was centrifuged as above and cell pellet treated in the same way. The cell suspension was centrifuged as above one more time and supernatant decanted. The cell pellet was suspended in the remaining drops of 10% glycerol and kept it on ice.
50 µL of fresh cell suspension were used for electroporation.

7.2.23. Transformation of *E. coli* strains

7.2.23.1. Heat shock of chemically competent cells

1 µL of plasmid DNA or 10 µL of ligation reaction mix were added to ice melted chemically competent *E. coli* cell aliquot in eppendorf tubes and gentle mixed.
The mix was incubated on ice for 30 min. Then it was heat shocked for 1 min at 42 °C using a thermo-block. The aliquot was immediately put on ice and incubated for 5 min. The cells were recovered using 1 mL of room temperature LB medium in a test tube and incubated at 37 °C for 1 h with shaking at 200 rpm. The cell suspension was spreaded on a plate containing LA medium plus the corresponding antibiotics. Plates were incubated overnight at 37°C and colonies appeared on next day.

7.2.23.2. Electroporation of electrocompetent cells

1-2 µL (approx. 100 ng) of either fosmid, plasmid or DNA cassette were added to an ice-cold fresh 50 µL electrocompetent cells aliquot and gentle mixed. Electroporation was carried out in a 0.2 cm ice-cold electroporation cuvette (Biorad) using a BioRad GenePulser II electroporator (or equivalent) set to 200 Ω, 25 µF and 2.5 kV. The expected time constant is 4.5 – 5.6 ms. After electroporation, 1 mL ice cold LB medium was immediately added, the mixture was transferred to a test tube. Cells were incubates for 1 h at 37 °C (30 °C for E. coli BW25113/pKD46) with shaking at 200 rpm. The cell suspension was spreader on a plate containing LA medium plus the corresponding antibiotics. Plates were incubated overnight at 37 °C (30 °C for *E. coli* BW25113/pKD46) and colonies appeared normally on next day but sometimes after 48 – 72 h.

7.2.24. Insertion of intneo disruption cassette (*neo*, int-attP$^{\varphi C31}$) in Clm resistant (*cat*) gene of pCC1-FOS fosmid

pCC1-FOS fosmid must to be electroporate into *E. coli* BW25113/pKD46 (λ RED recombination plasmid) strain.
pKD46 contain the resistance marker *bla* (ampicillin resistance) and a temperature sensitive origin of replication (requires 30 °C for replication).
Electrocompetent cells of *E. coli* BW25113/pKD46/pCC1-FOS strain had to be generated.
50 mL LB medium containing Amp (100 µg/mL) and Clm (12.5 µg/mL) were inoculated with 1 mL of overnight *E. coli* BW25113/pKD46/pCC1-FOS culture. 500 µL 1M L-arabinose stock

solution (final concentration is 10 mM, induces *red* genes) were added. Culture was incubated for approximately 3.5 – 4 h at 30 °C shaking at 4000 rpm to an OD_{600} of approximately 0.6. Again, 500 µL 1M L-arabinose stock solution were added and culture was incubated longer for 30 min as above.

The next steps were done as it was explained before in preparation of electrocompetent cells and electroporation of electrocompetent cells sections. Transformants were selected with Kan (50 µg/mL).

If not further gene disruptions will be made on this fosmid, the strain was incubated overnight at 37 °C to promote the loss of pKD46.

The final pCC1-FOS-intneo fosmid had to be purified and concentrated in approx. 15-20 µL of 10 mM Tris buffer pH 8.0 as explained before in the fosmid isolation section. The fosmid was then electroporated in the commercial electrocompetent cells of *E. coli* Epi 300 strain (Epicentre) in order to induce the fosmid multi-copy production as described by the manufacturer.

7.2.25. Conjugative DNA transfer from *E. coli* ET12567 (pUZ8002) into *Streptomyces* strains

E. coli ET12567/pUZ8002 (ET cells) is a non-methylating host (*dam*-, *dcm*-) carrying pUZ8002 helper plasmid which provides transfer functions from RK2. This strain allows mobilization of any plasmid carrying RK2 (RP4) *oriT* into the recipient host. The strain is resistant to both Clm (25 µg/mL) and Kan (25 µg/mL).

For conjugation, approximately 2 cm^2 of ET cells containing a vector of interest grown overnight on LA medium plus the corresponding antibiotics were scratched and suspended in 500 µL 2XYT medium.

Separately, 50 µL (approx. $5x10^9$ c.f.u./mL) of *Streptomyces* strain's spore suspension melted on ice were gentle mixed with 350 µL 2XYT medium and incubated at 50 °C for 10 min to germinate the spores.

Spore suspension mixture was cooled down at room temperature (approx. 10 min) and gentle mixed with 100 µL of ET cells suspension.

The mixture was centrifuged at 4000 rpm for 1 min and 250 µL of the supernatant were removed.

The pellet was carefully suspended in the remaining 2XYT medium and spreaded on a plate containing 30 mL SFM medium supplemented with 1M $MgCl_2$ (final concentration 10 mM).

The plate was incubated at 28 °C for 16-18 h and then first antibiotic selection was performed.

1 mL sterile distilled H_2O containing the selection's antibiotic (depending on vector selection's marker) and Nal (30µg/mL, selection against *E. coli*) were mixed and carefully homogenously spread on the SFM medium surface until dry. Plates were incubated longer at 28 °C for 4-7 days until transconjugants appeared. Single transconjugants were then transferred to a new SFM plate containing the selection's antibiotic and Nal (30 µg/mL).
After second antibiotic selection, spores from a single clone were spreaded on a new plate containing SFM medium supplemented with the same antibiotics and incubated at 28 °C until dense sporulation appeared (4-7 days). Spores were harvested and spores suspension prepared. Finally, they were stored at -80 °C.

7.2.26. Genome library construction of *Amycolatopsis* YIM10 strain

For building a genome library of *Amycolatopsis* YIM10 strain the CopyControl™ Fosmid Library. Production Kit with pCC1FOS™ Vector (Epicentre) was used.
gDNA of the strain was purified as it was mentioned before. The gDNA was mechanically sheared by pipetting to generate approximately 40 Kb DNA fragments size as explained in the company kit protocol.
The number of clones required to ensure a 99% probability of a given DNA sequence of *Amycolatopsis* YIM10 strain (genome = 10.37 Mb) being contained within a fosmid library composed of 40-kb inserts is approximately 1100 regarding the mathematical formula contained in the company kit protocol. Therefore, the genome library used in this project consists of twelve 96-well plates holding single colonies of *E. coli* Epi 300 strain, where each colony contains a single copy of fosmid pCC1-FOS harboring an approx. 40 Kb sized DNA fragment of the *Amycolatopsis* YIM10 strain genome. To induce the multi-copy fosmid production by *E. coli* Epi 300 strain, autoinduction solution from kit was used.
For further detailed information of this procedure check the kit protocol.

7.2.27. Dialysis procedure

For dialysis of lasso peptides, a 20 mL 1 kDa MWCO Mega Pur-A-Lyzer kit (Sigma-Aldrich) was used.
20 mL of concentrated lasso peptides methanolic extract was transfer into the dialysis container and dialyzed against 2 L of buffer solution (50 mM sodium phosphate, 100 mM NaCl, pH 8.0) for dialysis during 2 days at 4 °C with slowly stirring. For detailed information check the protocol from manufacturer.

7.2.28. Bioinformatic Genome Analysis by antiSMASH

AntiSMASH is a bioinformatic tool, which can predict bacterial gene clusters in genome sequences [74,75]. The genome sequences used in this project were uploaded to the antiSMASH server and analyzed for gene clusters by this software.
The online version is available at https://antismash.secondarymetabolites.org/.
The output was a list of all detected clusters and the functions of single genes, thus the type of the encoded enzymes were identified. Predicted biosynthetic genes and mainly core genes were verified by BLAST searches on the NCBI website (http://blast.ncbi.nlm.nih.gov/Blast.cgi) [106].

7.2.29. Clone Manager Software

Clone Manager is a software used in this project for *in-silico* construction of vector maps, PCR simulations, cloning simulations and primer design.

7.2.30. Lab Solutions Software

This Software was used during this project to analyze HPLC chromatograms.

7.2.31. Bruker Compass Data Analysis 4.2

Bruker Compass Data Analysis software was used to analyze mass spectra to calculate masses and the chemical sum formulas based on mass accuracy and isotopic pattern matching (SmartFormula algorithm).

7.2.32. Production protocol of felipeptins A1 and A2

7.2.32.1. Pre-inoculum preparation

100 mL flasks containing 10 mL of YEME medium were inoculated with 200 µL of *Streptomyces coelicolor* M1154_pCC1-FOS-Felipeptins-cluster_pSOK806-SARP strain spore suspension. The cultures were incubated at 200 rpm and 28 °C for 5 days. Then, 250 mL baffled flasks containing 50 mL of YEME medium were inoculated with 3 mL of the well grown cultures above mentioned. The cultures were incubated at 200 rpm and 28 °C for 3 days until dense growth was obtained.

7.2.32.2. Fermentation

2 L baffled flasks containing 500 mL of MYM medium were inoculated with 20 mL of pre-inoculum. The fermentation was carried out at 200 rpm and 28 °C for 7 days.

7.2.32.3. Extraction

Fermentation cultures were harvested and freeze dried during 2 days. 300 mL of methanol plus 6 % formic acid (18 mL of concentrated acid) were used to extract the freeze dried material generated from 500 mL fermentation culture. Extraction was carried out at 200 rpm and room temperature for 3 hours.

The methanol extract was centrifuged at 400 rpm and room temperature for 10 min to get rid of debris.

The methanol extract was concentrated under reduced pressure in a minimum volume of methanol (from 300 to 15 mL approximately). 1 mL of formic acid was added per 50 mL of concentrated methanol extract. Then, the extract was transferred to tubes for dialysis, as explained in dialysis procedure.

After dialysis, the whole material inside the tube was carefully transferred to 50 mL falcon tubes (20 mL of dialyzed liquid per falcon tube). The material was centrifuged at 10,000 rpm and 4 °C for 10 minutes. The supernatant was discarded. Each pellet was suspended in 2 mL methanol plus 6% formic acid (all pellet was dissolved). The whitish-transparent pellet mostly contains felipeptins A1 and A2 (can be check by HPLC).

The felipeptins were further purified by semi preparative HPLC, as already explained. Felipeptin A1 has a retention time of 14 minutes, and felipeptin A2 a retention time of 11 minutes approximately.

The already explained purification protocol should yield approximately 12 mg of pure felipeptin A1 and 7 mg of pure felipeptin A2 per liter of fermentation culture.

7.2.33. Bioassays of felipeptins A1 and A2 on human cell lines

Those experiments were performed by collaboration with Prof. Dr. Galina Selivanova at the Department of Microbiology, Tumor and Cell Biology, Karolinska Institute, Stockholm, Sweden.
During those studies four humans cell lines were used: colon carcinoma HCT116 cells, BJ normal fibroblast cells, A375 melanoma cells, and breast carcinoma MCF7 cells.

7.2.33.1. Cell viability assay using Resazurin

Day 1: Seed 2000-3000 cells in 96 well plates using 100 µL of complete medium in each well.
Day 2: Prepare a range of concentrations of felipeptins by adding the appropriate volumes of the compound solutions (from 5 mM stocks dissolved in DMSO) in complete medium. Remove the medium from the wells and add 100 µL of the compounds in respective wells.
Note: prepare at least 3 wells for each concentration and 3 wells for DMSO (DMSO dissolved in medium) as controls.
Check the plate using microscope and put the plate in 37 °C incubator with 5 % CO_2. Every 24 h check the cells to see their responses to the compounds.
Day 5: Terminate the treatment and check the relative cell viability using Resazurin assay

Resazurin assay

Prepare 5 µM resazurin solution using resazurin stock solution and complete medium. Remove the medium from the wells. Add 100 µL of 5 µM resazurin solution in the wells containing the cells. Additionally, add 100 µL of 5 µM resazurin solution in 3 wells without cells. Incubate the plate at 37 °C in incubator for 2-4 h.
Note: Incubation time depends on the cell type and cell number. Take the plate and measure the relative fluorescent units, using a plate reader.

7.2.33.2. Colony formation assay for combinational treatment of felipeptins A1 and A2 plus doxorubicin

Seed 1000-1500 cells in 12 well plates. After 24h, remove the medium and add 1 mL of medium containing the combination of felipeptins A1 (FA1) and A2 (FA2), in respective wells.
Note: for the control well, add medium containing DMSO.
Let the plates incubate at 37 °C for 72 h. After 72 h, prepare the concentrations of doxorubicin (DOX) in medium and add them to the wells that you have considered for DOX treatment and felipeptins mixture FA1/FA2 plus DOX combinational treatment
Note: In this step do not remove the medium from the wells. Just add 1 mL from the medium containing DOX in the considered wells.
For the wells that contain only felipeptins mixture FA1/FA2 treatment, add 1 mL of complete medium without removing already added medium.
Put the plates in incubator at 37 °C for 48-72 h. After that, remove the medium, wash 2 times with ice cold PBS, fix with ice cold 70% ethanol and continue with crystal violet staining. After crystal violet staining, take a picture using capture device.

8. References

1. Katz, L. & Baltz, R. H. Natural product discovery: past, present, and future. *Journal of Industrial Microbiology and Biotechnology* (2016). doi:10.1007/s10295-015-1723-5
2. Bérdy, J. Bioactive microbial metabolites. *J. Antibiot. (Tokyo).* (2005). doi:10.1038/ja.2005.1
3. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products* (2016). doi:10.1021/acs.jnatprod.5b01055
4. Mohr, K. I. History of antibiotics research. *Curr. Top. Microbiol. Immunol.* (2016). doi:10.1007/82_2016_499
5. Gould, K. Antibiotics: From prehistory to the present day. *Journal of Antimicrobial Chemotherapy* (2016). doi:10.1093/jac/dkv484
6. Baltz, R. H. Antimicrobials from Actinomycetes: back to the future. *Microbe* (2007).
7. Bérdy, J. Thoughts and facts about antibiotics: Where we are now and where we are heading. *Journal of Antibiotics* (2012). doi:10.1038/ja.2012.27
8. Kohanski, M. A., Dwyer, D. J. & Collins, J. J. How antibiotics kill bacteria: From targets to networks. *Nature Reviews Microbiology* (2010). doi:10.1038/nrmicro2333
9. de Lima Procópio, R. E., da Silva, I. R., Martins, M. K., de Azevedo, J. L. & de Araújo, J. M. Antibiotics produced by Streptomyces. *Brazilian Journal of Infectious Diseases* (2012). doi:10.1016/j.bjid.2012.08.014
10. Lewis, K. Platforms for antibiotic discovery. *Nature Reviews Drug Discovery* (2013). doi:10.1038/nrd3975
11. Martens, E. & Demain, A. L. The antibiotic resistance crisis, with a focus on the United States. *Journal of Antibiotics* (2017). doi:10.1038/ja.2017.30
12. douafer, hana, Andrieu, V., Phanstiel, O. & Brunel, J. M. Antibiotic adjuvants: Make antibiotics great again! *J. Med. Chem.* acs.jmedchem.8b01781 (2019). doi:10.1021/acs.jmedchem.8b01781
13. Mulani, M. S., Kamble, E. E., Kumkar, S. N., Tawre, M. S. & Pardesi, K. R. Emerging Strategies to Combat ESKAPE Pathogens in the Era of Antimicrobial Resistance: A Review. *Front. Microbiol.* 10, (2019).
14. O'neill, J. *Tackling drug resistant infections globally: Final report and recommendations*. *The Review on AMR* (2016). doi:10.1016/j.jpha.2015.11.005
15. Hassanpour, S. H. & Dehghani, M. Review of cancer from perspective of molecular. *J. Cancer Res. Pract.* (2017). doi:10.1016/j.jcrpr.2017.07.001
16. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* (2018). doi:10.3322/caac.21492

17. Holohan, C., Van Schaeybroeck, S., Longley, D. B. & Johnston, P. G. Cancer drug resistance: An evolving paradigm. *Nature Reviews Cancer* (2013). doi:10.1038/nrc3599
18. Hecht, S. M. Bleomycin: New Perspectives on the Mechanism of Action 1. *J. Nat. Prod.* 63, 158–168 (2000).
19. Chaires, J. B. Molecular Recognition of DNA by Daunorubicin. in 156–167 (1994). doi:10.1021/bk-1995-0574.ch010
20. Di Marco, A., Arcamone, F. & Zunino, F. Daunomycin (Daunorubicin) and Adriamycin and Structural Analogues: Biological Activity and Mechanism of Action. in *Mechanism of Action of Antimicrobial and Antitumor Agents* (2012). doi:10.1007/978-3-642-46304-4_8
21. Paz, M. M., Zhang, X., Lu, J. & Holmgren, A. A new mechanism of action for the anticancer drug Mitomycin C: Mechanism-based inhibition of thioredoxin reductase. *Chem. Res. Toxicol.* (2012). doi:10.1021/tx3002065
22. Kuşoğlu, A. & Biray Avcı, Ç. Cancer stem cells: A brief review of the current status. *Gene* (2019). doi:10.1016/j.gene.2018.09.052
23. Tabassum, D. P. & Polyak, K. Tumorigenesis: it takes a village. *Nat. Rev. Cancer* 15, 473–483 (2015).
24. Takahashi, Y. & Nakashima, T. Actinomycetes, an Inexhaustible Source of Naturally Occurring Antibiotics. *Antibiotics* (2018). doi:10.3390/antibiotics7020045
25. 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). doi:10.1093/femsre/fux005
26. Barka, E. A. *et al.* Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiol. Mol. Biol. Rev.* (2016). doi:10.1128/mmbr.00019-15
27. Flärdh, K. & Buttner, M. J. Streptomyces morphogenetics: Dissecting differentiation in a filamentous bacterium. *Nature Reviews Microbiology* (2009). doi:10.1038/nrmicro1968
28. Jakimowicz, D., Zydek, P., Kois, A., Zakrzewska-Czerwińska, J. & Chater, K. F. Alignment of multiple chromosomes along helical para scaffolding in sporulating Streptomyces hyphae. *Mol. Microbiol.* (2007). doi:10.1111/j.1365-2958.2007.05815.x
29. Jakimowicz, D. & Van Wezel, G. P. Cell division and DNA segregation in Streptomyces: How to build a septum in the middle of nowhere? *Molecular Microbiology* (2012). doi:10.1111/j.1365-2958.2012.08107.x
30. Adegboye, M. F. & Babalola, O. O. Actinomycetes : a yet inexhaustive source of bioactive secondary metabolites. in *Microbial pathogens and strategies for combating them: science, technology and education* (2013).

31. Baltz, R. H. Renaissance in antibacterial discovery from actinomycetes. *Current Opinion in Pharmacology* (2008). doi:10.1016/j.coph.2008.04.008
32. Zhang, Z., Wang, Y. & Ruan, J. A Proposal To Revive the Genus Kitasatospora (Omura, Takahashi, Iwai, and Tanaka 1982). *Int. J. Syst. Bacteriol.* (2009). doi:10.1099/00207713-47-4-1048
33. Parte, A. C. LPSN - List of prokaryotic names with standing in nomenclature (Bacterio.net), 20 years on. *International Journal of Systematic and Evolutionary Microbiology* (2018). doi:10.1099/ijsem.0.002786
34. Takahashi, Y. Genus Kitasatospora, taxonomic features and diversity of secondary metabolites. *Journal of Antibiotics* (2017). doi:10.1038/ja.2017.8
35. Adamek, M. *et al.* Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in Amycolatopsis species. *BMC Genomics* (2018). doi:10.1186/s12864-018-4809-4
36. LECHEVALIER, M. P., PRAUSER, H., LABEDA, D. P. & RUAN, J.-S. Two New Genera of Nocardioform Actinomycetes: Amycolata gen. nov. and Amycolatopsis gen. nov. *Int. J. Syst. Bacteriol.* 36, 29–37 (1986).
37. Kumari, R., Singh, P. & Lal, R. Genetics and Genomics of the Genus Amycolatopsis. *Indian J. Microbiol.* (2016). doi:10.1007/s12088-016-0590-8
38. Chen, S., Wu, Q., Shen, Q. & Wang, H. Progress in Understanding the Genetic Information and Biosynthetic Pathways behind Amycolatopsis Antibiotics, with Implications for the Continued Discovery of Novel Drugs. *ChemBioChem* (2016). doi:10.1002/cbic.201500542
39. Zotchev, S. B. Antibiotics: Biosynthesis. in *Wiley Encyclopedia of Chemical Biology* (John Wiley & Sons, Inc., 2008). doi:10.1002/9780470048672.wecb014
40. Medema, M. H. *et al.* Minimum Information about a Biosynthetic Gene cluster. *Nature Chemical Biology* (2015). doi:10.1038/nchembio.1890
41. Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature* (2002). doi:10.1038/417141a
42. Challis, G. L. Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology* (2008). doi:10.1099/mic.0.2008/018523-0
43. Zotchev, S. B., Sekurova, O. N. & Katz, L. Genome-based bioprospecting of microbes for new therapeutics. *Current Opinion in Biotechnology* (2012). doi:10.1016/j.copbio.2012.04.002
44. Van Keulen, G. & Dyson, P. J. Production of specialized metabolites by streptomyces coelicolor A3(2). in *Advances in Applied Microbiology* (2014). doi:10.1016/B978-0-12-800259-9.00006-8
45. Buchholz, T. J., Kittendorf, J. D. & Sherman, D. H. Polyketide Biosynthesis: Modular

Polyketide Synthases. in *Wiley Encyclopedia of Chemical Biology* (John Wiley & Sons, Inc., 2008). doi:10.1002/9780470048672.wecb459

46. Shen, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology* (2003). doi:10.1016/S1367-5931(03)00020-6
47. Hertweck, C. The biosynthetic logic of polyketide diversity. *Angewandte Chemie - International Edition* (2009). doi:10.1002/anie.200806121
48. Staunton, J. & Weissman, K. J. Polyketide biosynthesis: A millennium review. *Natural Product Reports* (2001). doi:10.1039/a909079g
49. Shen, B. Biosynthesis of Aromatic Polyketides. in (2007). doi:10.1007/3-540-48146-x_1
50. Moore, B. S. Discovery of a new bacterial polyketide biosynthetic pathway. *ChemBioChem* (2001). doi:10.1002/1439-7633(20010105)2:1<35::AID-CBIC35>3.0.CO;2-1
51. Schoenafinger, G. & A Marahiel, M. Nonribosomal Peptides: Biosynthesis. in *Wiley Encyclopedia of Chemical Biology* (John Wiley & Sons, Inc., 2008). doi:10.1002/9780470048672.wecb398
52. Finking, R. & Marahiel, M. A. Biosynthesis of Nonribosomal Peptides. *Annu. Rev. Microbiol.* (2004). doi:10.1146/annurev.micro.58.030603.123615
53. Süssmuth, R. D. & Mainz, A. Nonribosomal Peptide Synthesis—Principles and Prospects. *Angewandte Chemie - International Edition* (2017). doi:10.1002/anie.201609079
54. Payne, J. A. E., Schoppet, M., Hansen, M. H. & Cryle, M. J. Diversity of nature's assembly lines-recent discoveries in non-ribosomal peptide synthesis. *Mol. Biosyst.* (2017). doi:10.1039/c6mb00675b
55. McErlean, M., Overbay, J. & Van Lanen, S. Refining and expanding nonribosomal peptide synthetase function and mechanism. *Journal of Industrial Microbiology and Biotechnology* (2019). doi:10.1007/s10295-018-02130-w
56. Arnison, P. G. *et al.* Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Natural Product Reports* (2013). doi:10.1039/c2np20085f
57. Ortega, M. A. & Van Der Donk, W. A. New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products. *Cell Chemical Biology* (2016). doi:10.1016/j.chembiol.2015.11.012
58. McIntosh, J. A., Donia, M. S. & Schmidt, E. W. Ribosomal peptide natural products: Bridging the ribosomal and nonribosomal worlds. *Natural Product Reports* (2009). doi:10.1039/b714132g

59. Hudson, G. A. & Mitchell, D. A. RiPP antibiotics: biosynthesis and engineering potential. *Current Opinion in Microbiology* (2018). doi:10.1016/j.mib.2018.02.010
60. Hetrick, K. J. & van der Donk, W. A. Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era. *Current Opinion in Chemical Biology* (2017). doi:10.1016/j.cbpa.2017.02.005
61. Maksimov, M. O., Pan, S. J. & James Link, A. Lasso peptides: Structure, function, biosynthesis, and engineering. *Natural Product Reports* (2012). doi:10.1039/c2np20070h
62. Li, Y., Zirah, S. & Rebuffat, S. *Lasso Peptides*. (Springer New York, 2015). doi:10.1007/978-1-4939-1010-6
63. Hegemann, J. D., Zimmermann, M., Xie, X. & Marahiel, M. A. Lasso Peptides: An Intriguing Class of Bacterial Natural Products. *Acc. Chem. Res.* (2015). doi:10.1021/acs.accounts.5b00156
64. Burkhart, B. J., Hudson, G. A., Dunbar, K. L. & Mitchell, D. A. A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol.* (2015). doi:10.1038/nchembio.1856
65. Mevaere, J. *et al.* An orthogonal system for heterologous expression of actinobacterial lasso peptides in Streptomyces hosts. *Sci. Rep.* (2018). doi:10.1038/s41598-018-26620-0
66. Maksimov, M. O. & Link, A. J. Prospecting genomes for lasso peptides. *Journal of Industrial Microbiology and Biotechnology* (2014). doi:10.1007/s10295-013-1357-4
67. Solbiati, J. O. *et al.* Sequence analysis of the four plasmid genes required to produce the circular peptide antibiotic microcin J25. *J. Bacteriol.* (1999).
68. Duquesne, S. *et al.* Two Enzymes Catalyze the Maturation of a Lasso Peptide in Escherichia coli. *Chem. Biol.* (2007). doi:10.1016/j.chembiol.2007.06.004
69. Monciardini, P., Iorio, M., Maffioli, S., Sosio, M. & Donadio, S. Discovering new bioactive molecules from microbial sources. *Microb. Biotechnol.* (2014). doi:10.1111/1751-7915.12123
70. Sekurova, O. N., Schneider, O. & Zotchev, S. B. Novel bioactive natural products from bacteria via bioprospecting, genome mining and metabolic engineering. *Microbial Biotechnology* (2019). doi:10.1111/1751-7915.13398
71. Kolter, R. & van Wezel, G. P. Goodbye to brute force in antibiotic discovery? *Nat. Microbiol.* (2016). doi:10.1038/nmicrobiol.2015.20
72. Bachmann, B. O., Van Lanen, S. G. & Baltz, R. H. Microbial genome mining for accelerated natural products discovery: Is a renaissance in the making? *Journal of Industrial Microbiology and Biotechnology* (2014). doi:10.1007/s10295-013-1389-9
73. Ziemert, N. *et al.* The natural product domain seeker NaPDoS: A phylogeny based

bioinformatic tool to classify secondary metabolite gene diversity. *PLoS One* (2012). doi:10.1371/journal.pone.0034064

74. Medema, M. H. *et al.* AntiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* (2011). doi:10.1093/nar/gkr466
75. Blin, K. *et al.* AntiSMASH 4.0 - improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* (2017). doi:10.1093/nar/gkx319
76. Skinnider, M. A. *et al.* Genomes to natural products PRediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Res.* (2015). doi:10.1093/nar/gkv1012
77. Gross, H. Strategies to unravel the function of orphan biosynthesis pathways: Recent examples and future prospects. *Applied Microbiology and Biotechnology* (2007). doi:10.1007/s00253-007-0900-5
78. Zarins-Tutt, J. S. *et al.* Prospecting for new bacterial metabolites: A glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Natural Product Reports* (2016). doi:10.1039/c5np00111k
79. Gomez-Escribano, J. P. & Bibb, M. J. Heterologous expression of natural product biosynthetic gene clusters in Streptomyces coelicolor: From genome mining to manipulation of biosynthetic pathways. *Journal of Industrial Microbiology and Biotechnology* (2014). doi:10.1007/s10295-013-1348-5
80. Wohlleben, W., Mast, Y., Stegmann, E. & Ziemert, N. Antibiotic drug discovery. *Microb. Biotechnol.* (2016). doi:10.1111/1751-7915.12388
81. Bode, H. B., Bethe, B., Höfs, R. & Zeeck, A. Big effects from small changes: Possible ways to explore nature's chemical diversity. *ChemBioChem* (2002). doi:10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9
82. Challis, G. L. & Ravel, J. Coelichelin, a new peptide siderophore encoded by the Streptomyces coelicolor genome: Structure prediction from the sequence of its non-ribosomal peptide synthetase. *FEMS Microbiol. Lett.* (2000). doi:10.1016/S0378-1097(00)00184-1
83. Xu, D. *et al.* Bioprospecting deep-sea actinobacteria for novel anti-infective natural products. *Front. Microbiol.* (2018). doi:10.3389/fmicb.2018.00787
84. Moore, J. M., Bradshaw, E., Seipke, R. F., Hutchings, M. I. & McArthur, M. Use and discovery of chemical elicitors that stimulate biosynthetic gene clusters in streptomyces bacteria. *Methods Enzymol.* (2012). doi:10.1016/B978-0-12-404634-4.00018-8
85. Oh, S. Y., Shin, J. H. & Roe, J. H. Dual role of OhrR as a repressor and an activator in

response to organic hydroperoxides in Streptomyces coelicolor. *J. Bacteriol.* (2007). doi:10.1128/JB.00632-07

86. Maddocks, S. E. & Oyston, P. C. F. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* (2008). doi:10.1099/mic.0.2008/022772-0
87. Zhu, H., Sandiford, S. K. & Van Wezel, G. P. Triggers and cues that activate antibiotic production by actinomycetes. *Journal of Industrial Microbiology and Biotechnology* (2014). doi:10.1007/s10295-013-1309-z
88. Baltz, R. H. Genetic manipulation of secondary metabolite biosynthesis for improved production in Streptomyces and other actinomycetes. *J. Ind. Microbiol. Biotechnol.* (2016). doi:10.1007/s10295-015-1682-x
89. Kim, J. H. *et al.* Cloning large natural product gene clusters from the environment: Piecing environmental DNA gene clusters back together with TAR. *Biopolymers* (2010). doi:10.1002/bip.21450
90. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* (2009). doi:10.1038/nmeth.1318
91. Kesenheimer, C. & Groth, U. Total Synthesis of (−)-8- O -Methyltetrangomycin (MM 47755). *Org. Lett.* 8, 2507–2510 (2006).
92. Ding, C. *et al.* Synthesis study on marmycin A: Preparation of the C3′-desmethyl analogues. *J. Org. Chem.* (2009). doi:10.1021/jo9011078
93. Rojas, J. D. *et al.* Genome Sequence of Streptomyces olindensis DAUFPE 5622, Producer of the Antitumoral Anthracycline Cosmomycin D. *Genome Announc.* (2014). doi:10.1128/genomea.00541-14
94. Hong, S. T., Carney, J. R. & Could, S. J. Cloning and heterologous expression of the entire gene clusters for PD 116740 from Streptomyces strain WP 4669 and tetrangulol and tetrangomycin from Streptomyces rimosus NRRL 3016. *J. Bacteriol.* (1997). doi:10.1128/jb.179.2.470-476.1997
95. Palmu, K., Ishida, K., Mäntsälä, P., Hertweck, C. & Metsä-Ketelä, M. Artificial reconstruction of two cryptic angucycline antibiotic biosynthetic pathways. *ChemBioChem* (2007). doi:10.1002/cbic.200700140
96. Kibret, M. *et al.* Streptomyces spp. From Ethiopia producing antimicrobial compounds: Characterization via bioassays, genome analyses, and mass spectrometry. *Front. Microbiol.* (2018). doi:10.3389/fmicb.2018.01270
97. Sekurova, O. N., Zhang, J., Kristiansen, K. A. & Zotchev, S. B. Activation of chloramphenicol biosynthesis in Streptomyces venezuelae ATCC 10712 by ethanol shock: Insights from the promoter fusion studies. *Microb. Cell Fact.* (2016). doi:10.1186/s12934-016-0484-9

98. Cai, X. *et al.* Structure and Biosynthesis of Isatropolones, Bioactive Amine-Scavenging Fluorescent Natural Products from Streptomyces Gö66. *Angew. Chemie - Int. Ed.* (2017). doi:10.1002/anie.201701223
99. Ramijan, K., van Wezel, G. P. & Claessen, D. Genome Sequence of the Filamentous Actinomycete Kitasatospora viridifaciens . *Genome Announc.* (2017). doi:10.1128/genomea.01560-16
100. Rycroft, D. S., Cole, W. J. & Rong, S. Highly oxygenated naphthalenes and acetophenones from the liverwort Adelanthus decipiens from the British isles and south America. *Phytochemistry* (1998). doi:10.1016/S0031-9422(97)00806-6
101. Breinholt, J. *et al.* Tigloside: a New Tigloylated Tetrasaccharide from Amycolatopsis sp. *Acta Chem. Scand.* (2008). doi:10.3891/acta.chem.scand.52-1239
102. Guo, Z. K. *et al.* Actinotetraoses A-H: Tetrasaccharide derivatives from a grasshopper-associated amycolatopsis spHCa1. *Planta Med.* (2012). doi:10.1055/s-0031-1298592
103. Inahashi, Y. *et al.* Biosynthesis of Trehangelin in Polymorphospora rubra K07–0510: Identification of Metabolic Pathway to Angelyl-CoA. *ChemBioChem* (2016). doi:10.1002/cbic.201600208
104. Tacar, O., Sriamornsak, P. & Dass, C. R. Doxorubicin: An update on anticancer molecular action, toxicity and novel drug delivery systems. *Journal of Pharmacy and Pharmacology* (2013). doi:10.1111/j.2042-7158.2012.01567.x
105. Baltz, R. H. Natural product drug discovery in the genomic era: realities, conjectures, misconceptions, and opportunities. *J. Ind. Microbiol. Biotechnol.* (2019). doi:10.1007/s10295-018-2115-4
106. Boratyn, G. M. *et al.* BLAST: a more efficient report with usability improvements. *Nucleic Acids Res.* (2013). doi:10.1093/nar/gkt282