Genomics-assisted chemical exploration of the secondary metabolomes of *Massilia* and *Acidovorax* species

Zur Erlangung des akademischen Grades eines

Dr.rer.nat

von der Fakultät Bio- und Chemieingenieurwesen
der Technischen Universität Dortmund
genehmigte Dissertation

vorgelegt von

M. Sc. Jan Diettrich

aus

Witten

Tag der mündlichen Prüfung: 03.03.2022

1. Gutachter/-in: Prof. Dr. Markus Nett

2. Gutachter/-in: Prof. Dr. Dirk Tischler

Dortmund 2022

Table of Contents

List of Figures

List of Tables

Declaration on the reproduction of previously published content

Parts of this work have been published by the author of this thesis and/or are based on data which were obtained in BSc and MSc theses performed at the Laboratory of Technical Biology under the supervision of the author. The following list gives a precise overview of the individual contributions.

Chapter 4.2.1 in parts modified from [A]

Chapter 4.3.1 in parts modified from [A]

[A] „Genomics-inspired discovery of massiliachelin, an agrochelin epimer from *Massilia* sp. NR 4-1“, Diettrich, J.; Kage, H.; Nett, M. *Beilstein J. Org. Chem.* **2019,** *15,* 1298–1303. doi.org/10.3762/bjoc.15.128

List of Abbreviations

ACP	acyl carrier protein
AT	acyl transferase
ATP	adenosine triphosphate
A	adenylation
ABC	ATP binding cassette
BFB	bacterial fruit blotch
BGC	biosynthetic gene cluster
CoA	coenzyme A
C	condensation
COSY	correlated spectroscopy
DH	dehydratase
DNA	deoxyribonucleic acid
DMSO	dimethyl sulfoxide
ESI	electrospray ionisation
ER	enoyl reductase
E	epimerization
FAAL	fatty acyl-AMP ligase
FUR	ferric uptake regulator
F	Fungi
GC	guanine-cytosine
IC_{50}	half maximal inhibitory concentration

HMBC	heteronuclear multiple-bond correlation
HSQC	heteronuclear single quantum coherence
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
hbu	hydroxybutanoic acid
KR	ketoreductase
LC	Liquid chromatography
MAT	malonylacyl transferase
MS	mass spectrometry
m/z	mass-to-charge ratio
Mbp	megabase pair
MT	methylation
MIC	minimal inhibition concentration
NADPH	nicotinamide adenine dinucleotide phosphate
NRPS	nonribosomal peptide synthetase
NIS	NRPS-independent siderophore
NMR	nuclear magnetic resonance
NOESY	nuclear overhauser enhancement spectroscopy
OSMAC	one strain many compounds
AOS	oxamine synthase
Ox	oxidation
PSA	pentylsalicylic acid
PCP	peptidyl carrier protein

PKS	polyketide synthase
PCR	polymerase chain reaction
PLP	pyridoxal 5′-phosphate
Q-TOF	quadrupole-time of flight
RNA	ribonucleic acid
RiPP	ribosomally synthesized and posttranslationally modified peptide
SAM	S-adenosyl methionine
SM	secondary metabolite
SDR	short-chain dehydrogenases/reductases
THN	tetrahydroxynaphthalene
TE	thioesterase
T	thiolation
THNS	1,3,6,8-tetrahydroxynaphthalene synthase
tRNA	transfer ribonucleic acid
TCA	tricarboxylic acid cycle
T3SS	type III secretetion sytem
UV/VIS	ultraviolet/visible
ZUR	zinc uptake regulator

1 Introduction

1.1 The need for novel antibiotics

Bacteria, fungi and plants have been a rich source of natural products over the last centuries [1]. Natural product-derived antibiotics are often highly active and highly specific as their structures have been evolutionarily optimized [2]. The majority of antibiotics in clinical use were discovered between 1940 and 1960 from bacteria of the family *Actinomycetaceae* [3]. Even recently approved drugs like daptomycin or ceftobiprole derive from previously described chemical scaffolds [4].

In recent years the number of multidrug-resistant organisms has dramatically increased. These organisms have become a considerable threat to public health worldwide. The emergence of resistance to antimicrobial compounds is even more threatening considering that the number of newly approved drugs per year is decreasing within the last decades [5].

Many of the antibiotics were isolated from antibacterial screening programs by cultivating Gram-positive bacteria from soil samples. A downside of this approach is the high rediscovery rate of already identified compounds.The evolvement of genome sequencing technology led to a paradigm shift in natural product discovery with the start of the new century. On the one hand it became evident, that even well-investigated strains possess a higher capacity of producing secondary metabolites than originally assumed. On the other hand many bacteria that had been neglected in screening approaches before were identified as putative producers of novel secondary metabolites [4]. These findings led to a novel strategy where the genome sequence of a bacterium is analysed for the secondary metabolite production potential before cultivation. This approach is known as genome mining [6].

As a result the number of novel natural products from before neglected genera increased rapidly. Especially Gram-negative proteobacteria like *Burkholderia* or *Janthinobacterium* were identified as rich sources of novel secondary metabolites [7]. Motivated by these findings the question arises if there are further neglected genera that are chemically underinvestigated but bear yet unknown potential as producers of novel secondary metabolites.

1.2 Genome Mining

1.2.1 A short history of natural product discovery and genome mining

When the Italian physician Bartolomeo Gosio isolated for the first time an antibiotic from the fungus *Penicillium brevicompactum* in 1893, it was not the only time in the history of natural products discovery that serendipity played an important role [8]. Gosio was looking for a metabolite causing pellagra when he noticed that the purified compound, mycophenolic acid, inhibited the growth of *Bacillus anthracis* [9]. Thirty-five years later in September 1928 it was again a lucky stroke of destiny when Alexander Fleming returned from vacation and found a forgotten petri dish of *Staphylococcus aureus.* A contamination with *Penicillium notatum* showed a characteristic inhibitory effect and Fleming described the production of a bright yellow colour after four or five days. Penicillin was discovered [10]–[12].

The microbiologist Selman Waksman set another milestone by the discovery of streptomycin from *Streptomyces griseus* in 1942 [13]. From there on *Streptomyces* species received high attention as producers of secondary metabolites. A detailed review on antibiotics from *Streptomyces* spp. was published before [14]. The following time from the 1940s to the 1960s is called the golden era of antibiotics. About 150 novel antibiotic compounds were discovered during that period including compound classes like cephalosporins, macrolides and tetracyclines [15]–[17]. The discovery of novel compounds in this era was the result of extensive effort and time-consuming screening approaches. The company Eli Lilly raised the idea of asking missionaries to collect soil samples from exotic places around the world. One of these samples sent in from Borneo in 1952 harbored a strain of *Streptomyces orientalis* which was able to produce vancomycin [18], [19]. Another example of lucky circumstances was the finding of β-lactamase inhibitors, which were isolated as byproducts of *Streptomyces clavuligerus* in 1976 [20].

In the same year the production of the antibiotic methylenomycin A by the strain *Streptomyces coelicolor* A3(2) was linked to its plasmid SCP1. For the first time the biosynthesis of a compound was connected to its genetic basis [21]. In the following years the number of natural products linked to their gene clusters increased [22], [23].

Interestingly in the year 2002 it was again the strain *Streptomyces coelicolor* A3(2) that influenced the history of genome mining. The complete genome of this strain was sequenced and revealed 7825 genes and more than 20 predicted gene clusters coding for expressed or predicted natural compounds. Maybe it was the birth of the modern genome mining approach when the authors wrote: “The genome sequence will greatly increase our understanding of microbial life in the soil as well as aiding the generation of new drug candidates by genetic engineering” [24]. The idea of predicting novel bioactive compounds based on genome sequences without the use of extensive bioactivity screenings was born [6].

The DNA sequencing techniques improved rapidly and due to decreasing costs genome sequence data became more and more publicly available. A detailed history of DNA sequencing is given by Shendure et al. [25].

1.2.2 In silico genome mining

The in silico genome mining approach of analyzing DNA sequence data makes finding of secondary metabolites more efficient compared to bioactivity screening methods. There are, however, requirements that need to be fulfilled. Although secondary metabolites are structurally highly diverse, their producing enzymes often exhibit a significant degree of homology. The most classical approach of genome mining is the search for genes that encode core biosynthetic enzymes involved in the biosynthesis of secondary metabolites. Many of these machineries are highly conserved and harbor similar amino acid sequences. A known sequence of one or multiple enzymes can serve as a probe for a comparison with a genome sequence of an organism of interest [6]. Some compound classes like polyketides, nonribosomal peptides, ribosomally synthesized and post-translationally modifed peptides (RiPPs) turned out to show excellent suitability as subject for genome mining approaches for several reasons.

Nonribosomal peptide synthetases (NRPS) harbor modularly organized core enzymes that show a high degree of conservation and, furthermore, follow defined rules of biosynthesis. Each module is responsible for the incorporation of a single amino acid into the growing peptide chain. This one-to-one correspondence is known as principle of colinearity. In addition key residues in their adenylation domains allow a prediction of the substrates used and, hence, the structure of the final NRPS

product [6]. The biosynthesis of nonribosomal peptides and the principle of colinearity is discussed in detail in chapter 1.4.

Polyketides are a further compound class that were subject of several successful genome mining approaches. The modular type I polyketide synthases (PKS) follow, similar to the NRPS, rules of collinearity which allow to predict the biosynthetic product to a certain extent. Stambomycin is an illustrative example where the elucidated structure was in good accordance with the structure prediction [26]. Although the rules of collinearity only apply to the modularly organized PKS, the products of other PKS classes have also been successfully identified by utilizing genome mining approaches [6]. More insights in the biosynthesis are given in chapter 1.5.

A lot of computational genome mining tools are online available like ClustScan [27], NP.searcher [28] and SBSPKS [29]. The development of such tools is quite rapid and new tools arise regularly. An overview of the currently available tools can be found at http://www.secondarymetabolites.org. In 2011, the antibiotics and Secondary Metabolite Analysis SHell was launched [30]. Since then, antiSMASH has become one of the leading computational genome mining tools, which recognizes more than 20 different compound classes.

Even though in silico studies were facilitated with the development of computational tools, the concept of using conserved sequence data for identifying novel secondary metabolites was already known. The sequences of already described and highly conserved genes were used to design primers followed by PCR experiments. In a screening approach the labeled gene was used as a probe in Southern hybridization experiments [31]. An extensive list of successfully identified gene clusters of actinomycetes using these methods was described by Weber et al. [31].

The expanded strategy of searching for complete gene clusters instead of single genes is called comparative genome mining. On the one hand this approach provides the advantage of gaining information about associations between biosynthetic pathways and produced molecules. On the other hand the challenge is to find appropriate rules for defining a corresponding gene cluster [6]. An interesting approach of comparative genome mining was carried out in the marine strain *Streptomyces* sp. NA03103. In this study an orphan NRPS gene cluster was found,

which did not show any homology to gene clusters in public databases. However, the upstream and downstream genes flanking the gene cluster were highly conserved. The heterologous expression of the gene cluster led to the discovery of ashmides A and B [32].

The genome neighborhood network is a large scale bioinformatic method that is able to assign complete biosynthetic pathways [33]. Assigning a function of an enzyme by sequence comparison can lead to misannotation. In prokaryotic cells genes of metabolic pathways are typically organized in operons. Therefore, the short distance between the genes in such a neighborhood can be connected to a functional relationship of their expressed products if it is phylogenetically conserved over different species and thus helps increasing the annotation precision [34], [35].

A remarkable study, which involved the analysis of 1669 prokaryotic, 49 fungal and 80 metazoan genomes, was carried out in 2019. By introducing so called neighborhood function profiles the prediction power of genome neighborhood networks could be significantly increased. As an outcome it was also shown that it is quite common that certain pairs of genes with unrelated functions cluster together [36]. This phenomenon known as “genome hitchhiking” was described before [37]. Besides chemical aspects genome neighborhood network studies also give valuable insides in understanding evolutionary mechanisms [33].

Some approaches are focusing especially on the phylogenetic inference and on evolutionary analyses based on phylogenies. Not only do these approaches help understanding the mechanisms of how the structures of secondary metabolites developed but they can also guide the discovery of novel compounds. The phylogenetic approaches can in general be divided into two classes. The first one focuses on gene trees, whereas the second one investigates species trees [38]. An interesting application of the first class is the phylogenomic analysis of fungal type I PKS, which could be grouped into three clades corresponding to highly, partially and non-reducing enzymes [39]. In addition to the possibility of predicting enzyme functions, gene trees contribute to the understanding of gene history, loss and gain of single domains as well as horizontal gene transfer [38]. The second class, however, maps the trait on a phylogenetic tree representing the evolutionary relationship.

A genome mining technique is to map chemical properties on the species trees helping to predict promising groups of organisms. An extensive approach of that kind was carried out by Anzai et al. [40]. The authors compared 366 *Aspergillus* strains and could establish a correlation between the phylogenetic distribution and the biosynthesis of antimicrobial compounds [40]. A well-illustrated phylogenetic genome mining approach was recently carried out in *Streptomyces* spp. The authors identified biosynthetic gene clusters of 1100 *Streptomyces* genomes by using antiSMASH. From the analysis of these gene clusters, the authors could draw three conclusions. First, *Streptomyces* genomes possess a high abundance of so-called hybrid biosynthetic gene clusters (BGCs). The respective loci contain genes that code for more than one type of scaffold-synthesizing enzymes. Secondly, many *Streptomyces* gene clusters can be associated with the production of antitumor compounds. Thirdly, the number and type of BGCs can vary dramatically; even in strains belong to the same species. These conclusions led to the suggestion that within-species sequencing strategy for finding secondary metabolites from microbes provides a promising complementary approach compared to focusing on single representatives of a species [41].

1.2.3 Selective identification of gene clusters

Along with the increasing availability of genomic data and thousands of newly discovered and uncharacterized BGCs the question arises of how to pick reasonable candidates for identifying novel molecular structures [42], [43]. An approach that, in contrast to the aforementioned holistic methods, focuses on the selective identification of gene clusters is the resistance gene-directed genome mining. In many cases the produced antimicrobial compounds are toxic to the organism itself and therefore a resistance mechanism is required in order to avoid suicide [44]. One known mechanism is the duplication of the target gene. The enzyme encoded by the gene itself is affected by the produced secondary metabolite, whereas the duplicated gene expresses an enzyme which is not inhibited. The identification of these duplicated genes can lead to the identification of promising compounds [42]. This mechanism is known for metabolites like novobiocin [45], pentalenolactone [46] or salinosporamide A [47]. The first successful study of resistance guided genome mining investigated actinomycetes and led to the discovery of pekiskomycin [48]. In 2019 a resistance gene-directed genome mining of 50 *Aspergillus* species was carried out. The authors identified 72 unique families of putative resistance genes

showing the potency of a strategy, which is investigating BGCs in a selective manner [42].

Another approach, which does not directly focus on the mining of the core biosynthetic gene, is the genomic investigation of regulators. The regulation of metal homeostasis is essential for all microorganisms [49]. In the cell metals like iron and zinc act as cofactors in proteins. Therefore, the concentration of these metals must be tightly regulated. Two examples of the ferric uptake superfamily involved in this regulation process are the proteins FUR and ZUR [49]. The simplified mechanism of these proteins is the binding to palindromic sequences in the promoters of their DNA target using a co-repressor. This repressor induces a conformational change in the regulator protein allowing interaction with the DNA [49], [50]. The first investigation of that kind was done by Spohn et al. [51]. In this study a zinc uptake regulator was identified in the actinomycete *Amycolatopsis japonicum*. The binding motif of this regulator was analyzed and then used for the discovery of genes involved in the production of zinc-binding secondary metabolites. As a result the gene cluster encoding for the biosynthesis of a zinc complexing molecule was identified. Because the expression of this metabolite is already suppressed at low zinc concentrations, the gene cluster was cloned in a mutant strain lacking the regulation system. This led to the production of ethylenediamine-disuccinate [51]. FUR is often involved in the regulation of iron complexing compounds known as siderophores. These iron chelating molecules are produced under iron limiting conditions and have been subject to a multitude of genome mining approaches due to their versatile biochemical functions. Siderophores are discussed in detail in chapter 1.3.

1.2.4 From in silico to laboratory

The described genome mining approaches, although using different techniques, focus all on the identification of promising gene clusters. However, they cannot answer the question, if a predicted secondary metabolite is really produced under given laboratory conditions or if the gene cluster remains silent [52]. The basic idea behind the OSMAC concept (One Strain MAny Compounds) is, that a microorganism has the potential to produce several natural products under specific cultivation conditions. As a consequence, variation of the cultivation conditions can lead to the production of different metabolites [53]. Examples for parameters than can be

changed are temperature, nutrient composition, vessel types or shaking conditions [52], [54].

In an advanced OSMAC study, the bioinformatic substrate prediction directs the medium composition. A successful study of this kind was carried out by Lautru et al. [55]. The analysis of a gene cluster in the genome of *Streptomyces coelicolor* resulted in the prediction of hydroxamic acid functional groups. These characteristic functional groups suggested the produced compound to be a siderophore. Cultivation in an iron-deficient medium led to the production of the compound coelichelin [55]. Cupriachelin is a further example for the discovery of a siderophore by a combined in silico and guided OSMAC approach. The proximity of genes encoding lipoprotein receptors and ABC-type iron transporters provided evidence for a putative siderophore gene cluster. The cultivation under iron efficient conditions triggered the production of the photoreactive siderophore [56].

A mass spectrometry (MS)–guided genome-mining method that also connects the genotype of a compound to its chemotype was developed by Kersten et al. [57]. In the so called natural product peptidogenomics approach, peptidic natural products are analyzed via multi-dimensional mass spectrometry. The mass shifts of the sequential fragment ions are characteristic for the identified peptide natural product [57]. Another chemotype-to-genotype approach led to the discovery of the lipopeptide taiwachelin. Instead of mass spectroscopy a $^{1}H,^{15}N$ HMBC experiment was carried out to confirm the presence of hydroxamate groups. The corresponding correlations are quite specific and give an indication for the presence of a siderophore [58]. However, the instrumental analytics aided genome mining approaches must not be focused on a special compound class. In an untargeted approach the elucidation of two new myxobacterial secondary metabolite classes, the lipothiazoles and the ε-amidated tripeptides, via liquid chromatography - mass spectrometry / mass spectrometry (LC-MS/MS) was reported [59].

The rapid development of genome mining and DNA sequencing not only led to a more efficient evaluation of known microbes but allowed the application of these tools on less well-studied microorganisms. Various investigations showed that also rare actinomycetes, cyanobacteria, plant endosymbionts as well as several proteobacteria like *Burkholderia* and myxobacteria are rich sources of novel secondary metabolites [7], [60], [61]. Cyanobacteria are known as producers of non-ribosomal peptides,

polyketides and hybrid peptide-polyketide compounds [62], [63]. The potent cancer cell toxin curacin A was isolated from the cyanobacterium *Lyngbya majuscula* and shows a high level of collinearity between genes in the cluster and the predicted biosynthesis [64]. Over 600 metabolites with antibacterial, antifungal or cytostatic activity have been isolated from myxobacteria [65]. The compounds belong to classes like macrolactones, peptides or alkaloids. The development of drug discovery in myxobacteria was recently reviewed by Müller et al [66]. An example of a successful genome mining approach in myxobacteria is the discovery of myxoprincomides produced by the strain *Myxococcus xanthus* DK1622 using a combined method of targeted mutagenesis, liquid chromatography and high-resolution mass spectrometry [67]. The same strain was subject of a self-resistance-guided screening approach and revealed some unusual alkylpyrones, which act as topoisomerase inhibitors [68]. Kunakom et al. investigated the natural product discovery in *Burkholderia* sp. and reported that 32% of the discoveries are based on genome mining approaches. The biosynthetic pathway of spliceostatin, a natural compound that is produced by *Burkholderia* spp., was deduced by a genome mining based approach [69]. Further β-proteobacteria are discussed in chapter 1.6.

1.3 Siderophores

Siderophores are small metal-binding chelators with a broad variety of functions. They act as iron homeostasis regulators [70], antibiotics [71], inhibitors of metalloenzymes [72] or virulence factors [73]. They find application in the biotechnological, pharmaceutical or agricultural industry [74]. In recent years the number of successful genome mining approaches directed to siderophores rapidly increased [74]–[78].

Siderophores are suitable targets of genome mining approaches for several reasons. On the one hand the biosynthetic core enzymes belong to classes like NRPS, PKS or hybrids of both that are often highly conserved. Therefore, genome mining allows to predict novel natural products of this compound class. On the other hand the metal compexing properties of siderophores allow an unambigious analytical identification of these compounds. Common strategies are the cultivation under different iron supplementation conditions [79] or the MS-guided identification of metal-siderophore complexes [80].

1.3.1 The need for iron

Iron is essential for all eukaryotes and nearly all prokaryotes. Exceptions are certain lactobacilli and *Borrelia burgdorferi* with intracellular iron concentration lower than 10 atoms per cell [81], [82]. Iron is the fourth most abundant element in the earth's crust and the most abundant transition metal in the human body [83]–[85]. It plays an important role in metabolic and electron transport processes including tricarboxylic acid cycle, oxidative phosphorylation and photosynthesis [84], [86]. Furthermore, several proteins utilize iron as a cofactor to fulfill functions including DNA repair, DNA replication or metabolic catalysis [87], [88]. Due to the existence of interconvertible oxidation states iron plays a pivotal role in electron transfer processes [71]. In general, the transition metal occurs in the reduced ferrous, Fe(II), or the oxidized ferric, Fe(III), form. The oxidation states Fe(IV), Fe(V) and Fe(I) occur less frequently and mainly as intermediate states in catalytic reactions. Despite its ubiquity in the earth's crust, the bioavailability of iron is limited [71]. The oxidized form is highly insoluble with a solubility of 10^{-18} M at pH 7 [89]. Other sources state a concentration of $1.4 * 10^{-9}$ M in aqueous solutions [90]. Regardless the basis of calculation, the resulting concentration is far lower than required by bacteria. To achieve iron homeostasis bacteria, fungi and even some plants are known to produce ferric chelators [91]–[93]. These molecules are known as siderophores. The word siderophore originates from the Greek where *sideros* means iron and *phores* means bearer [94].

1.3.2 Chemistry and structure of siderophores

Between 1949 and 1952 the first three siderophores, mycobactin, coprogen and ferrichrome were isolated and identified as growth factors [95]–[97]. In 1956 a milestone was set by observing the increased production of ferrichrome A by the fungus *Ustilago sphaerogena* under iron-deficient conditions [96]. The same authors investigated in a following study the production levels of siderophores under different iron concentrations. A natural product of *Bacillus subitilis* showed an increased production level under low iron concentrations. For the first time a siderophore directed expression strategy was proposed when the authors stated: „This suggests that these compounds are involved in an iron sequestering and transferring system, the latter either initiated or magnified by the organism during iron deficiency” [98].

The respective iron chelator from *B. subtilis* was later identified as 2,3-dihydroxybenzoylglycine [99].

Siderophores in general possess a higher affinity towards the ferric form of iron than to the ferrous form. This selectivity is also important for the discrimination of other transition metals. There are not many biologically important trivalent cations in contrast to divalent cations like zinc(II), copper(II), nickel(II) or manganese(II). Only the kinetically inert Co(III) and aluminium(III) which has a smaller atom radius of 0.54 Å (compared to 0.64 Å of iron) compete for complexation [100]. Despite the lower affinity there are, however, siderophores that are capable of forming complexes with dipositive cations. A list of putative non-classical biological functions of siderophores was provided by Johnstone et al. [71]. An illustrative example of such a metallophore is micacocidin produced by the phytopathogen *Ralstonia solanacearum*. It tends to form stable complexes with Zn^{2+} and Cu^{2+} that were not displaced by iron after 24 h of incubation [101].

Iron(III) as a hard lewis acid tends to build strong interactions with oxygen atoms. The interaction is tighter the higher the charge of the oxygen is. Therefore oxygen is a very prominent donor atom, where the most common geometry is octahedral. This geometry favors the thermodynamically stable high-spin iron(III) species. An example of such a complex is $Fe(H_2O)_6^{3+}$ where six ligands are arranged around the iron center with minimal degree of ligand repulsion. Siderophores typically feature multiple ligand groups, which form the iron complex. Hydroxamate, α-hydroxycarboxylate and catechol groups are the major bidentate ligands, each having a high selectivity for iron(III). Hydroxyphenyloxazolone, α-aminocarboxylate and α-hydroxyimidazole groups are also found as ligands in some siderophores (**Figure 1A**). The participation of donor atoms like sulfur or nitrogen in siderophores is also observed although they tend to have a lower affinity for iron(III) [100]. These ferric complexing ligand groups originate from a variety of biosynthetic pathways like the shikimate pathway, the TCA cycle, the hydroxamate- or the 2,3-dihydroxybenzoate pathway. Detailed insights into the biosynthesis of the ligand groups are provided elsewhere [72].

The biosynthesis of the final siderophore structure requires the incorporation of these ligand groups into a molecular scaffold to allow the octahedral orientation of the iron-siderophore complex. Despite the huge variety of structurally different siderophores

the backbone is assembled by a small number of enzyme classes involving NRPS

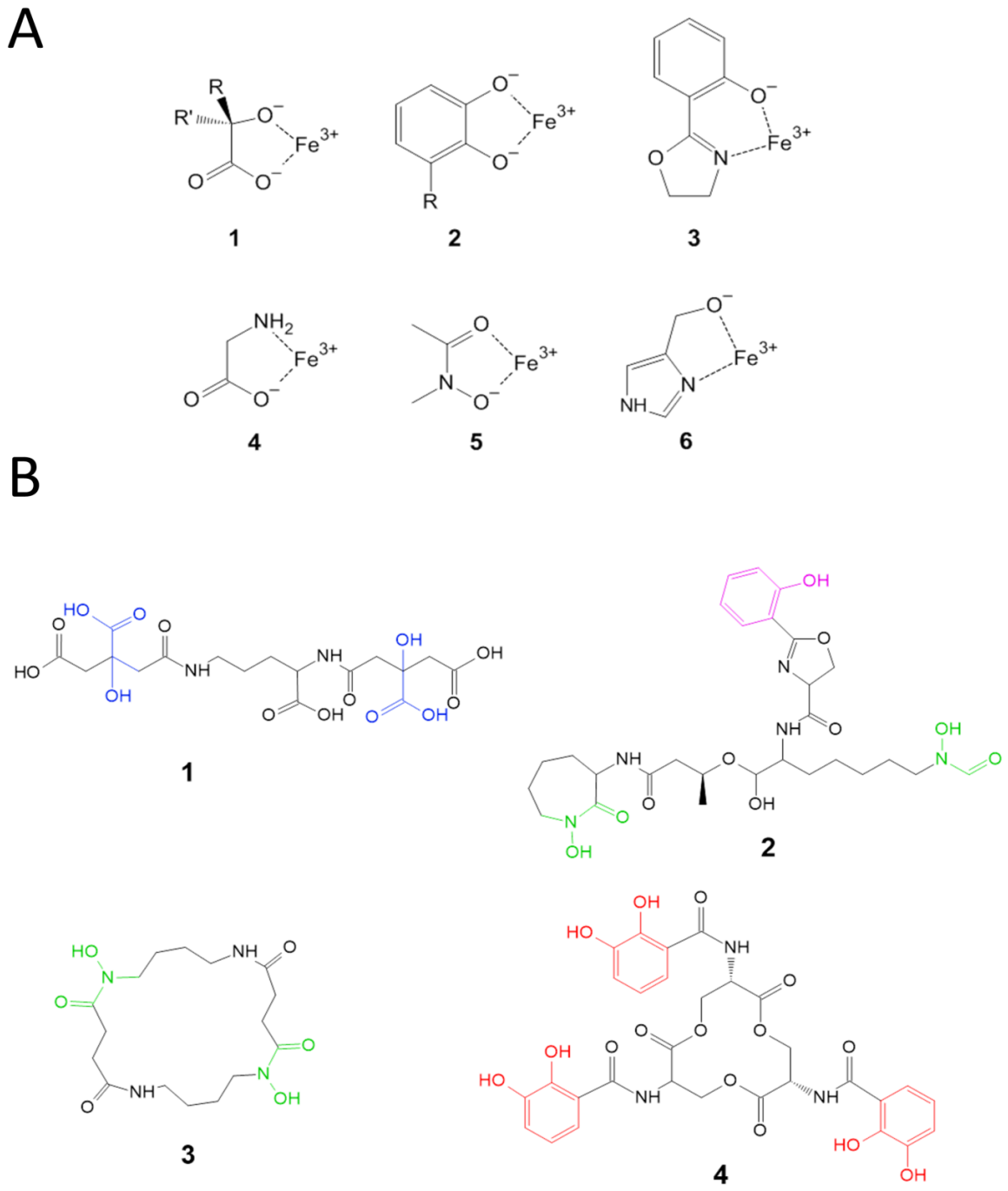


Figure 1: A) Structures of typical siderophore ligand groups (1) α-hydroxycarboxylate, (2) catecholate, (3) hydroxyphenyloxazolone, (4) α-aminocarboxylate, (5) hydroxamate, (6) α-hydroxyimidazole. B) Representative examples of bacterial siderophores: (1) staphyloferrin A, (2) mycobactin (3), putrebactin, (4) enterobactin.

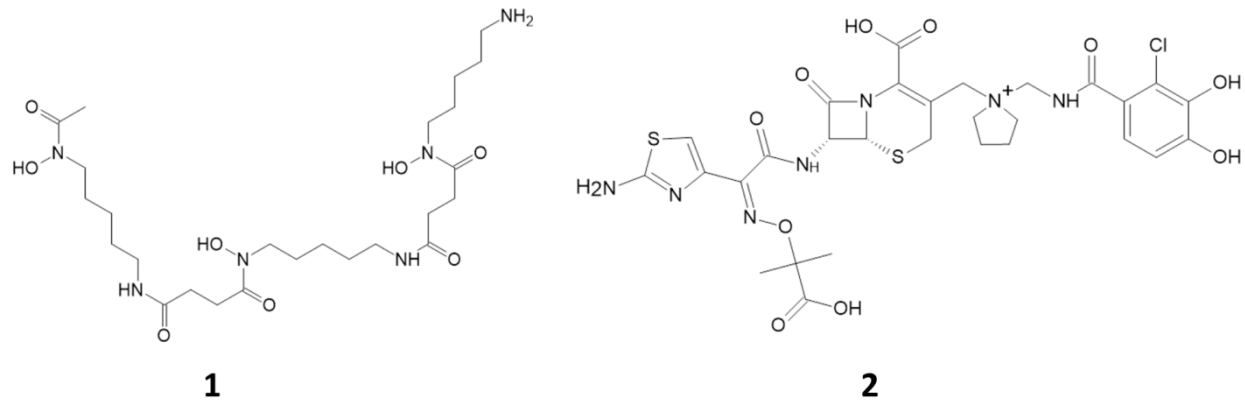
NRPS/PKS hybrids or NRPS-independent siderophore (NIS) synthetases [100], [102]. NPRS and PKS are discussed in detail in chapter 1.4 and 1.5. Some siderophores are assembled from alternating dicarboxylic acid, diamine and amino alcohol monomers. Although these building blocks derive from amino acids the

biosynthesis of the corresponding siderophores differs from the NRPS logic. Further information on NIS synthetized siderophores are provided elsewhere [103].

Hydroxamate siderophores, like putrebactin produced by *Shewanella putrefaciens,* are frequently found in nature (**Figure 1B**) [104]. While hydroxamate siderophores are known from both fungi and bacteria, catecholates are mostly produced by bacteria [104]. Enterobactin is an example of a siderophore that is produced by several *Streptomyces* species [105]. α-hydroxycarboxylate siderophores are produced by bacteria and fungi [104]. Staphyloferrrin A produced by *Staphylococcus hyicus* is a prominent example of an α-hydroxycarboxylate siderophore [79], [106]. The mixed type siderophores contain multiple ligand groups. An example of this type is mycobactin produced by *Mycobacterium tuberculosis* [95]. An extensive list of microorganisms and the corresponding produced siderophores was provided by Saha et al [104]. A review focusing especially on siderophores produced by fungi was published by Haas et al. [92]. Despite the limited number of ligand groups, their combination with different backbone structures leads to an enormous variety of siderophores in nature [72].

1.3.3 Medical application of siderophores

Siderophores are not only necessary for iron homeostasis in microorganisms (and graminaceous plants) but also play an important role in medical applications. It is estimated that millions of people suffer from iron overload including hereditary hemochromatosis, a disease caused by increased iron deposition [107], [108]. Siderophores find application in iron chelating therapy. Desferrioxamine B is a siderophore isolated from *Streptomyces pilosus* (**Figure 2**). It was one of the first complexing agents used for the treatment of iron overload. The application of desferrioxamine B showed an increase in iron excretion of patients with various types of iron overload when given parenterally [109].



1 **2**

Figure 2: Structures of desferrioxamine B (1) and cefiderocol (2)

A promising strategy in antibiotic drug design is to target siderophore-mediated iron acquisition [110]. The so called “trojan horse” strategy exploits iron uptake systems to enter the cell and unfold its antibiotic effect [111]. This strategy, however, can cause unwanted side effects due to the lack of selectivity and uncontrolled cell penetration activity [112]. A nature inspired approach to reduce the risk of the aforementioned effect is the application of so-called sideromycins. These siderophore–drug conjugates consist of a siderophore, a linker, and a drug. Sideromycins enter bacterial cells via energy-coupled transport. Due to this efficient uptake mechanism the minimal inhibitory concentration of sideromycins is at least 100-fold lower than diffusion transported antibiotics. Although several natural sideromycins are known like salmycin or albomycin, the majority of these compounds is nowadays synthesized chemically [113]. An extensive list of chemically synthesized sideromycins was provided by Fisher et al.[114].

Cefiderocol is an example that recently passed phase III clinical trials (**Figure 2**) [115]. This cephalosporin shows activity against carbapenem-resistant and multidrug-resistant Gram-negative bacilli. It consists of an iron-complexing catechol side chain and a ceftazidime related, antibacterial active component [116]–[118]. Liu et al. reported another promising approach by synthezising a siderophore-cephalosporin-oxazolidinone conjugate. The hydrolysis induced release of the oxazolidinone warhead provides activity against Gram-negative bacteria [119].

1.3.4 Iron uptake in bacteria

The cellular uptake of siderophores has been mainly investigated in Gram-negative bacteria. In contrast to Gram-positive bacteria they possess a cell wall consisting of

an outer membrane layer functioning as a molecular sieve and and a peptidoglycan layer increasing the mechanical stability. Due to their size siderophores barely permeate through the outer membrane. In general, they are transported into the periplasmic space via energy dependent and ligand-specific receptors. The energy is provided by the so called TonB enzyme complex. It consists of the three cytoplasmic membrane proteins TonB, ExbB and ExbD. These enzymes transfer the proton motive force of the cytoplasmic membrane to the outer membrane where direct contact between the C-terminal domain of TonB and the N-terminus of the outer membrane protein is required [111]. The next steps in the absorption process are the binding to a periplasmic binding protein followed by the transport across the cytoplasmic membrane via enzymes belonging to the ABC transporter superfamily [84], [100]. Once inside the cytosol, the iron is released from the siderophore by one of two possible mechanisms. Either the coordinated Fe^{3+} is reduced to Fe^{2+} by an iron reductase and subsequently transferred to acceptor molecules or the iron-siderophore complex is subjected to enzymatic hydrolysis.

The uptake mechanisms of Gram-postive bacteria are less well investigated. They do not possess an outer membrane and lack a periplasm as consequence [100]. Despite these structural differences Gram-positive bacteria are known to produce lipoprotein receptors and ABC transporters similar to those in Gram-negative bacteria. For *Staphylococcus aureus* it is known that it possesses siderophore periplasmic binding proteins (SPBP) as well as ABC transporters involved in the siderophore uptake mechanism. As Gram-positive do not possess an outer membrane, the SPBP are (in contrast to Gram-negative bacteria) membrane-anchored and recognize and transport siderophore-Fe(III) complexes [120].

1.4 Nonribosomal peptides

Peptides play essential roles in the biochemistry of life, e.g. as tachykinins, vasoactive intestinal peptides, endorphins, or antimicrobial peptides [121]–[124]. The majority of bioactive peptides are synthesized ribosomally. Though, there exist alternative biosynthetic routes for the formation of peptides.

In 1963 Tatum et al. investigated the biosynthesis of tyrocidine and draw the conclusion that this polypeptide is produced even under ribosome inhibition. The authors stated that a class of polypeptides “might be synthesized without an RNA

template by the stepwise enzymatic addition of amino acids" [125]. These so-called non-ribosomal peptides have gained increased attention in the following years. A prominent contributor to the research in this field was the Nobel Prize winner Fritz Lipmann, who discovered the ATP-dependent activation of amino acids [126]. Non-ribosomal peptides show a great variety in structure and function and are mainly produced by fungi and bacteria. They include many important therapeutics, such as antibiotics (penicillin), antitumor compounds (bleomycin) and immunosuppressants (cyclosporine) [126], [127].

The biosynthesis of nonribosomal peptides is mainly conducted by large enzymes, named nonribosomal peptide synthetases (NRPS). NRPS are modularly organized. Each module harbors different catalytic domains which together fuse amino acid building blocks successively to a growing peptide chain. The basic module includes a minimal set of three catalytic domains, which are responsible for one cycle of chain elongation. The adenylation domain (A) fulfills the function of selection, activation and loading of the amino acid to the thiolation domain (T). NRPS substrates include the 20 proteinogenic amino acids as well as a huge variety of nonproteinogenic amino acids, carboxylic acids and α-hydroxy or keto acids [126]. A single adenylation domain possesses substrate specificity and is responsible for the selection of a specific building block. Several approaches were carried out that identified essential residues in the A domain which allow to establish a correlation with the substrate specificity [128], [129]. Databases of these essential residues are increasing and are also available via online tools like the NRPS predictor [130]. The NRPS predictor uses a combination of existing databases [131] and further bacterial and fungal A domains of its own database [130]. However, a relaxed substrate specificity is observed in some NRPS and appears to be a strategy to reach a higher product diversity. A single NRPS is therefore capable of producing several products. The micromystin NRPS is an example that catalyzes the biosynthesis of several cyclic peptides with different amino acid sequences. [132].

Furthermore, the adenylation domain activates the amino acid monomer by forming an amino acyl-intermediate. The thiolation domain is also referred to as peptidyl carrier protein (PCP) and bears a phosphopantetheine moiety, which allows the intermediates to reach the catalytic centers in the corresponding module. The condensation domain (C) catalyses the coupling reaction of the amino acyl

intermediate to the α-amino group of the building block (**Figure 3**). In summary each module catalyses the connection of one amino acid to the growing peptide chain. This principle is known as collinearity rule and corresponds to type I polyketide synthases (PKS) discussed in 1.4.

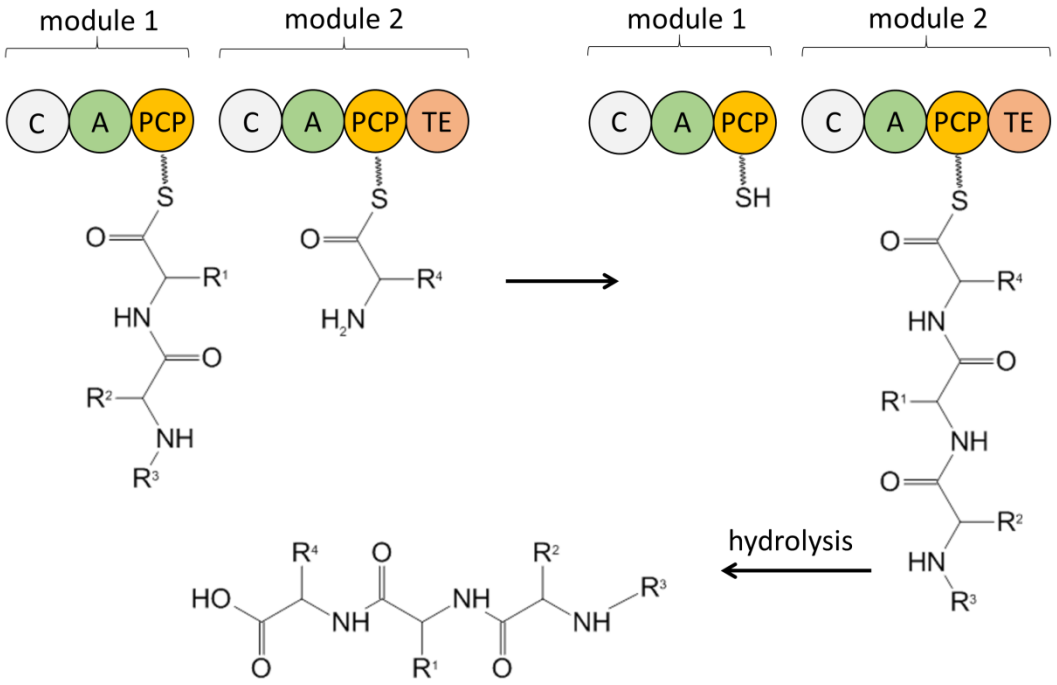


Figure 3: NRPS assembly line composed of two modules including condensation (C) adenylation (A) and peptidyl carrier protein (PCP) domains. The thioesterase (TE) domain catalyzes the final hydrolysis.

The described set of domains is able to synthesize peptidic structures. To achieve the great variety of functions and structures that occur in nature, further domains are required. Modification domains modify the NRP structure during or after the assembly. Epimerization domains (E) invert the configuration of L- amino acids into the corresponding D-form. Methylation domains (MT) catalyze the transfer of methyl groups. Cyclization domains catalyze the formation of thiazoline or oxazoline rings. These five-membered rings can be further oxidized by oxidation (Ox) domains [126].

1.5 Polyketides

The term polyketide was introduced by John Norman Collie in 1907. In his work "Derivatives of the multiple keten Group" he gave the ketomethylene structures the name ketides. Collie stated that the ketide group is the building block for a very large number of interesting compounds formed by the simplest reactions. Although Collies definition was not precise because his definition also included various lactones,

pyrone derivatives, benzenoid compounds and substituted pyridines, it already defined polyketides by the way in which they are assembled [133], [134].

Polyketides act as pigments, virulence factors, infochemicals or for defense. In a medicinal context they are used as antibiotics (e.g., erythromycin and tetracycline), immunosuppressants (e.g., rapamycin), antiparasitics, cholesterol-lowering (e.g., lovastatin) and antitumoral agents (e.g., daunomycin) [135]–[137].

Although polyketides have highly diverse structures they arise from surprisingly few building blocks like acetate and propionate. The biosynthesis is initiated with an activated acyl starter unit which is then elongated with malonyl-CoA-derived extender units. Chain elongation occurs by repetitive decarboxylative *thio*-Claisen condensations. Biosynthesis of polyketides is similar to fatty acid biosynthesis regarding the chemical mechanisms of chain extension and also in the use of precursors like acetyl-CoA and malonyl-CoA.

The fatty acid biosynthesis involves a (malonyl)acyl transferase (MAT/AT), which selects the starter and extender units and transfers them to a phosphopantetheinylated acyl carrier protein (ACP). The decarboxylative Claisen condensation reactions between an extender unit and the growing polyketide chain are catalyzed by a β-ketosynthase (KS). In fatty acid biosynthesis, the β-ketoacyl intermediate is reduced after each elongation by a ketoreductase (KR), dehydratase (DH) and an enoyl reductase (ER) leading to a fully saturated acyl structure. In polyketide biosynthesis, however, these reductive steps are optional and can also allow partly reduced intermediates. This difference contributes to the great variety among polyketide structures [138].

Polyketide synthases (PKS) are divided into three distinct classes (**Figure 4**). Type I PKS are multifunctional enzymes with covalently fused catalytic domains. They synthesize compounds like macrolides and polyethers. Type I PKS are subdivided into iterative and modular PKS. The latter are also called assembly-line PKS, because each module only catalyses one elongation step. In contrast, iterative type I PKS carry out a defined number of successive chain elongations. While assembly-line PKS are mainly found in bacteria, iterative PKS predominate in fungi [139]. Both iterative and modular PKS were reviewed before [139]–[141]. The modular type I PKS can be further subdivided into the so-called *trans*-AT PKS and *cis*-AT PKS. The

trans-AT PKS lack individual AT domains and are loaded by stand-alone acyl transferases. In contrast, *cis*-AT modules contain all three essential domains (KS, AT, and ACP).

Type II PKS are organized in a dissociable complex of discrete, monofunctional enzymes and synthesize products that are classified as anthracyclines, angucyclines, aureolic acids, tetracyclines, tetracenomycins, pradimicin-type polyphenols and benzoisochromanequinones. They are called “minimal PKS” as they consist only of two ketosynthase units and an acyl carrier protein. The two ketosynthase units, KS_{α} and KS_{β}, form a heterodimer, where the α-subunit catalyzes Claisen-type C-C bond formation reactions [142]. Decarboxylase activity towards malonyl-ACP and the determination of the chain length are functions of the β-subunit [143] .

The type III PKS are homodimeric synthases and consist only of a KS domain [144]. In contrast to type I and type II PKS their substrates are not bound to ACP domains. The “all in one” type III PKS catalyses the selection of the starter unit, the chain elongation and the cyclisation reaction. Type III PKS produce a wide range of compounds like chalcones, stilbenes, acridones or pyrones. These enzymes are widely distributed in plants, but occur also frequently in bacteria and fungi [68], [145]. Depending on their origin, they can utilize different cyclization modes in the synthesis of aromatic polyketides. An illustrative example is the pigment chrysophanol, which can be generated by bacterial or fungal type III PKS. In fungi, the folding is carried over the folding mode F (F for fungi), whereas, in bacteria, the folding acts in mode S (S for *Streptomyces*). Two intact C2 units are incorporated in the first ring of the natural compound by the folding mode F, whereas three C2 units are incorporated via the folding mode S [146]. Further insides into aromatic type III PKS products can be found elsewhere [138], [147].

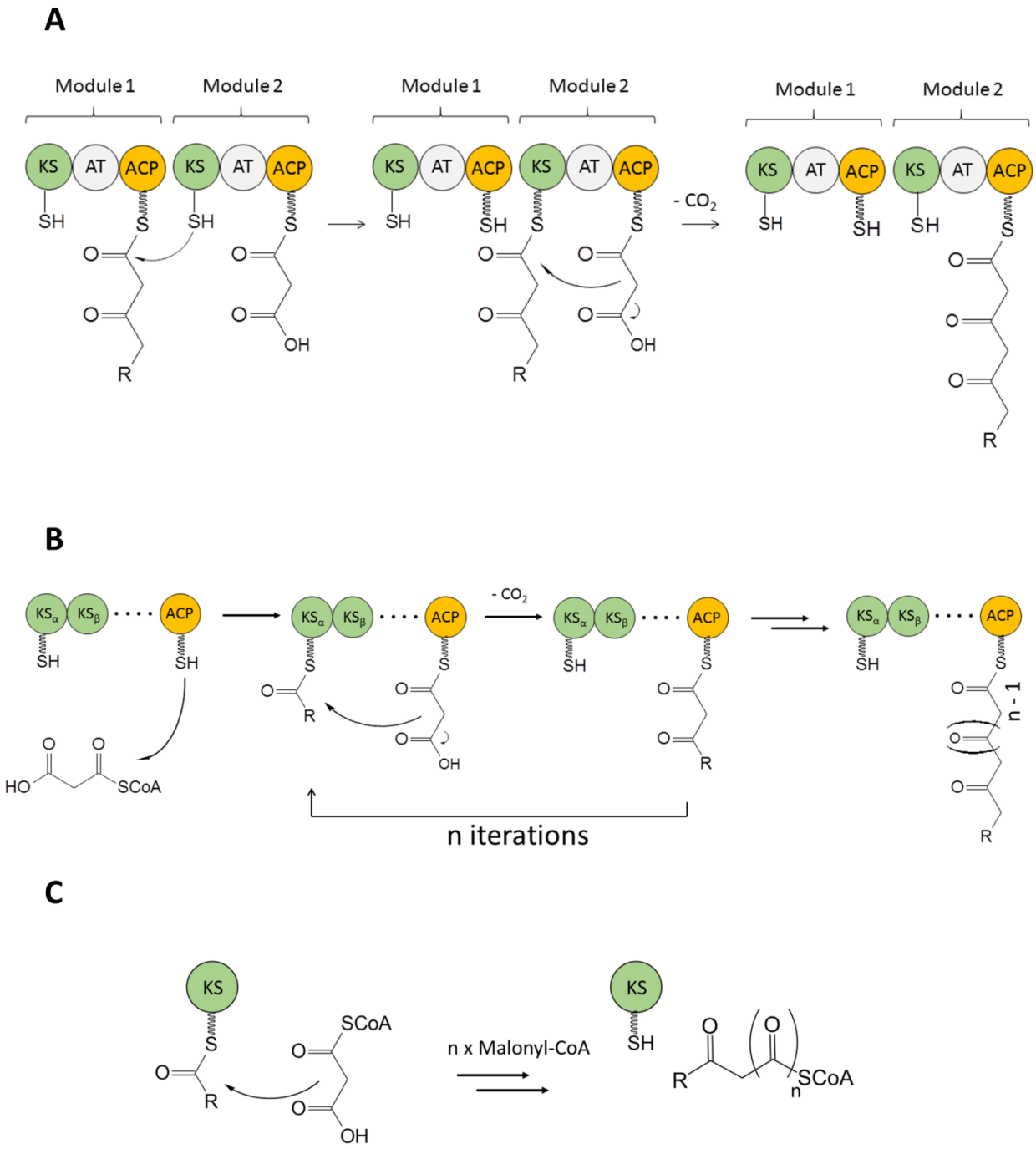


Figure 4: Scheme of type I (A), type II (B) and type III PKS (C) and their catalytic reactions. Domain notation: KS, ß-ketoacyl synthase; AT, acyltransferase; ACP, acyl carrier protein.

1.6 β-Proteobacteria

The name Proteobacteria derives from the greek god Proteus who is gifted with the ability to assume many different shapes. This ability reflects the morphological diversity in this phylum. The Proteobacteria are often referred to as „purple bacteria and their relatives“ although only few of them produce purple pigments or are photosynthetic [148]. They are divided into six major phylogenetic branches named by the greek letters α to ζ and represent the largest bacterial phylum with over 116 known families [149]. They are Gram-negative, often facultatively or obligately anaerobic and populate a variety of ecosystems like aquatic, soil, plant and even animal niches [150]. Some Proteobacteria were also described as „gifted microbes“ due to their capability to produce a large variety of secondary metabolites [7]. The corresponding taxa harbor between 20 and 30 secondary metabolite BGCs in their genomes and, thus, devote a large percentage of their genetic potential to the production of such specialized molecules [7].

There are classes of proteobacteria that have been recognized as producers of secondary metabolites for a long time. *Pseudomonas* species belonging to the γ-proteobacteria or myxobacteria as representatives of the δ-proteobacteria are examples of prolific producers of antimicrobial compounds [65], [151]. In 1977, ambruticin was isolated as the first myxobacterial antibiotic [152]. Although these proteobacteria have been a rich source of novel secondary metabolites over the last decades, with rising genome sequence data it became evident that some promising bacterial taxa have been neglected [4]. *Burkholderia* is a group belonging to the phylum of the β-proteobacteria that has attracted increasing attention as producers of natural products [153]. Kunakom et al. published a curated compilation of *Burkholderia* sensu lato secondary metabolites. The produced natural products belong to the classes of the nonribosomal peptides, shikimate pathway derived metabolites, hybrid polyketide-nonribosomal peptides, polyketides, ribosomally synthesized and posttranslationally modified peptides (RiPPs), polyynes and terpenes [61].

Further promising β-proteobacteria were identified, which are phylogenetically related to *Burkholderia.* For instance, the bioplastic producer *Cupriavidus necator* H16, which was formerly known as *Ralstonia eutropha* H16, has been found to be capable of nonribosomal peptide biosynthesis [58]. The siderophore cupriachelin produced by

this strain is photoreactive and releases the bound iron under light radiation supplementing diatoms in a symbiotic manner [56]. *Variovorax boronicumulans* is another β-proteobacterium showing close phylogenetic relationship to *Burkholderia* [154]. This strain produces lipopeptide siderophores named variochelins [155].

Furthermore, the chemistry of several plant pathogens among the β-proteobacteria was explored by genome mining approaches to identify molecular determinants of virulence [156]. An example is the strain *Ralstonia solanacearum* GMI1000. Genome analysis revealed its potential to produce a variety of natural compounds. Subsequent chemical investigations led to the identification of ralfuranone, micacocidin, staphyloferrin B and ralsolamycin [101], [157]–[159].

1.6.1 The genus *Massilia*

In 1998 an unkown bacterium was isolated from an immunocompromised patient with meningoencephalitis. 16S rRNA analysis showed closest relation to *Duganella zoogloeoides* and *Telluria mixta.* Further analysis, including GC-content determination and electron microscopy, showed significant differences leading to the determination of the new taxon named *Massilia timonae* gen. nov., sp. nov. [160]. Members of the genus *Massilia* are aerobic, flagellated, nonspore forming rods and belong to the family *Oxalobacteraceae*. They can be isolated from a variety of environments like dust samples, freshwater, soils, rhizosphere and roots of different plant species [161]–[166]. However, some *Massilia* strains were even retrieved from harsh environments like glacier ice and desert soil [167], [168]. Although the number of newly identified *Massilia* species is rapidly increasing, not much is known about the secondary metabolism or the chemistry of the genus in general. Assuming a correlation between genome size and metabolic potential *Massilia* is a promising source of secondary metabolites with genome sizes ranging from 5.0 to 7.5 Mbp [169]. When this PhD project started, only *Massilia timonae* and *Massilia* sp. BS-1 had been identified as producers of *N*-acyl homoserine lactones and violacein, respectively [170], [171].

1.6.2 The genus *Acidovorax*

The genus *Acidovorax* was introduced in 1990 by the reclassification of several *Pseudomonas* strains [172]. Strains of the genus were isolated from a variety of habitats such as soil [173], sludge [174], plants [175], water [176] or clinical samples

[177]. Among the latter *Acidovorax citrulli* plays an important agricultural and economic role as it is the causative agent of bacterial fruit blotch (BFB) [178], [179]. This disease affects watermelons and melons as well as other cucurbits, such as cucumbers, squashs and pumpkins [179]. Numerous outbreaks of BFB have led to significant economic damage worldwide [180]. The history of BFB and the biology of the seed infection were described in detail before [181], [182].

Acidovorax citrulli strains can be divided into two distinctive classes (class I and II) that are based on DNA fingerprint analyses and biochemical properties [180]. Strains belonging to the first group were isolated mainly from non-watermelon cucurbits, whereas the strains belonging to the second group were associated with watermelons [179]. It was further shown that these two types can be clearly distinguished by sequences of type III-secreted virulence effectors [183].

The assembly of the strain M6 which belongs to the class I revealed the presence of a 53 kb large plasmid. Similar plasmids were found in several strains belonging to group I, but were absent in all tested group II *A. citrulli* strains. The exact role of these plasmids is still to be determined [184].

Despite the agricultural importance of *Acidovorax citrulli*, surprisingly little is known about the chemical host-pathogen interactions. Previous studies showed that type III secretetion sytem (T3SS) genes are critical for pathogenicity [179], [185]. It was further shown that the ferric uptake regulator gene *furA* is required for *Acidovorax citrulli* virulence on watermelons [186]. The *furA* gene that encodes the FurA protein has, next to the regulation of intracellular iron homeostasis, several functions that might effect the virulence of a pathogen. It is involved amongst others in the regulation of quorum sensing, biofilm formation and the oxidative stress response in plant-pathogenic bacteria [187], [188]. A *furA* deletion mutant of the strain AACC0-1 showed a significantly reduced virulence on melon seedlings as well as an increased siderophore production [186].

2 Scope of this thesis

The emergence of multidrug-resistant bacterial pathogens in combination with treatment failures urges the need for the development of novel therapeutic drugs [118]. In recent years the number of bacterial genome sequences which are publically available has increased a lot. With computational support public sequence databases can be systematically analyzed in order to detect new biosynthesis gene clusters of secondary metabolites. As a result of these efforts Gram-negative bacteria were identified as potent producers of secondary metabolites [189]. Although an increasing number of natural products from β-proteobacteria have been identified lately, many genera of this phylum still remain chemically underinvestigated. Despite the increasing number of sequenced strains of the genus *Massilia* in recent years, not much is known about its secondary metabolism. Bioinformatic analysis identified *Massilia* sp. NR 4-1 as a promising representative of the genus for further chemical investigation. The main objective of this thesis was to isolate and identify secondary metabolites from strain NR 4-1. This objective can be further specified:

- One BGC of *Massilia* sp. NR 4-1 showed similarities to a BGC from *Ralstonia solanacearum* GMI1000, which is involved in the biosynthesis of the siderophore micacocidin. Therefore, this BGC represents a suitable target for a guided genome mining approach. One scope of this thesis was the identification of adequate chemical stimuli to trigger the production of the predicted siderophore. Afterwards, the respective compound should be isolated, its structure should be determined and compared with the bioinformatic proposal.
- A pathway for violacein biosynthesis was identified in *Massilia* sp. NR 4-1. This blue-purple pigment bears great medical potential due to its antimicrobial, antiprotozoal and anticancer properties. Another goal involved the establishment of appropriate cultivation conditions for violacein production, as well as the development of an extraction and isolation protocol, which allows the separation of violacein from its byproducts.
- *Massilia* sp. NR 4-1 devotes a large proportion of its genome to the production of secondary metabolites. Several BGCs cannot be associated with known natural products. The final scope comprises the identification of laboratory conditions which allow the isolation of these compounds. In a further step their

structures should be elucidated using mass spectrometry as well as 1D and 2D NMR measurements.

The strain *Acidovorax citrulli* AAC00-1 was subject of a further objective of this thesis. *Acidovorax citrulli* strains play an important agricultural and economic role as they are known to cause BFB on watermelons [179]. Despite their importance, surprisingly little is known about the chemical host-pathogen interactions and the secondary metabolism of *Acidovorax citrulli* strains. The scopy of this study was to investigate the secondary metabolite potential of AAC00-1 and identify potential virulence factors by bioinformatic analysis. Based on the resulting prediction of a natural compound different cultivation conditions were applied to investigate the secondary metabolism based on analytical methods.

3 Material and methods

3.1 Bioinformatic analyses

The antibiotics and Secondary Metabolite Analysis Shell AntiSMASH was used for the annotations and analyses of the secondary metabolite biosynthesis gene clusters in the genomes of *Massilia* sp. NR 4-1 and *Acidovorax citrulli* AAC00-1 [190]. Nucleotide and protein sequences were obtained from the GenBank database [191]. Homologies of different sequences were analyzed with the Basic Local Alignment Search Tool (BLAST) [192]. Signature motifs of the adenylation domains were analyzed with the NRPS Predictor [130]. The software SnapGene was used for the illustration of the gene clusters.

3.2 Analytical methods

3.2.1 Liquid chromatography – mass spectrometry (LC-MS)

Liquid chromatography – high resolution mass spectrometry (LC-HRMS) analyses were performed with an Agilent 1260 Infinity HPLC system combined with a Compact quadrupole-time of flight (Q-TOF) mass spectrometer (Bruker Daltonics). The Q-TOF mass spectrometer was interfaced with an electrospray ionization source. All analyses were performed in positive ion mode using a capillary potential of 4.5 kV. The desolvation gas (N_2) temperature was 220 °C with a flow rate of 12 L/min. Compounds were infused into the electrospray source by the 1260 Infinity high performance liquid chromatography (HPLC) system at a flow rate of 0.4 mL/min. Sample volumes of 5 µL were injected using a linear gradient of acetonitrile in water supplemented with 0.1% (v/v) formic acid. The gradient conditions were as follows: from 5% acetonitrile to 95% over 7 min, 95% acetonitrile for 2 min, 5% acetonitrile for 5 min. The UV detection wavelength was set between 100 nm and 600 nm. Measured masses ranged from *m/z* 100 to 1000 and *m/z* 500 to 1500. The protonated molecular ions of test compounds were generated at an end plate offset of 500 V. Nitrogen was used as the collision gas, at a pressure of 5 mPa.

LC-LRMS analyses were performed with an Agilent 1260 Infinity II LC system which was coupled to an Agilent 6120 Single Quadrupole LC/MS System (Agilent Technologies Inc.,Santa Clara, CA, USA) and further equipped with a diode array detector, a multicolumn thermostat, a degasser, a multisampler, a binary pump and

an Agilent poroshell 120 EC-C18 column (4.6×100 mM, 2.7 µm particle size). Sample volumes of 5 µL were injected using a linear gradient of acetonitrile in water supplemented with 0.1% (v/v) formic acid. The gradient conditions were as follows: from 5% acetonitrile to 95% over 7 min, 95% acetonitrile for 2 min, 5% acetonitrile for 5 min. The flow rate was set to 1 mL/min at a column temperature of 40 °C and a detection wavelength between 100 nm and 600 nm. Measured masses ranged from *m/z* 100 to 1000 and *m/z* 500 to 1500. The gas chamber was heated to 350 °C with a gas flow of 12 L/min. The capillary voltage was set to 3 kV and the nebulizer pressure to 35 psi.

3.2.2 Nuclear magnetic resonance (NMR) spectroscopy

NMR spectra were acquired at 300 K on a Bruker AV 700 Avance III HD (CryoProbe) or on a Bruker AV600 Avance III HD (CryoProbe). Samples were dissolved in deuterated solvents for measurement. Pyridine-d_5 was referenced to 7.19, 7.55 and 8.71 ppm in the ^{1}H NMR spectrum and to 123.5, 135.5 and 149.5 ppm in the ^{13}C NMR spectrum. Chloroform-*d* was referenced to 7.24 ppm in the ^{1}H NMR and 77.0 ppm in the ^{13}C NMR spectrum. DMSO-d_6 was referenced to 2.50 ppm in the ^{1}H NMR spectrum and to 39.5 ppm in the ^{13}C NMR spectrum.

3.2.3 Spectrophotometric measurements

The optical density was measured at a wavelength of 600 nm with a Biochrom Libra S11 photometer. Samples were diluted until they reached an optical density between 0.1 and 1. Sarstedt polystyrol cuvettes (10 x 4 x 45 mm) were used.

3.2.4 HPLC measurements

HPLC experiments were conducted with a Shimadzu HPLC system (LC-20AD system with SPD-M20A detector) using a linear gradient of methanol in water supplemented with 0.1% (v/v) trifluoroacetic acid. The gradient conditions were as follows: 10% methanol for 5 min, from 10% to 100% over 35 min, 100% for 10 min, followed by 10% for 10 min. The isocratic mehod was carried out with 50% methanol, supplemented with 0.1% (v/v) trifluoroacetic acid.

The following columns and flow rates were used:

- AppliChrom OTU LipoMare C18 5 µm 250mm x 4.6 mm, 1 mL/min (massiliachelin and *A. citrulli* analysis)
- VP 250/8 Nucleodur C18 PAH, 3 µm, 2 mL/min (massiliachelin isolation)
- VP 250/10 Nucleodur C18 HTec, 5 µm, 2 mL/min (massargiline and violacein isolation)
- VP 250/10 Nucleodur C18 Pyramid, 5 µm, 2 mL/min (massargiline purification)

3.3 Cultivation and isolation

3.3.1 R2A Medium

Ingredients	Standard R2A	Reduced R2A	R2A without additional carbon source
Yeast extract	0.50 g/L	0.50 g/L	0.50 g/L
Proteose peptone	0.50 g/L	0.50 g/L	0.50 g/L
Casamino acids	0.50 g/L	0.50 g/L	0.50 g/L
Glucose	0.50 g/L	0.20 g/L	-
Soluble starch	0.50 g/L	0.20 g/L	-
Sodium pyruvate	0.30 g/L	0.30 g/L	0.30 g/L
K_2HPO_4	0.30 g/L	0.30 g/L	0.30 g/L
$MgSO_4$ x 7 H_2O	0.050 g/L	0.050 g/L	0.050 g/L

The medium is adjusted to pH 7.2 prior to autoclave sterilization (121 °C, 15 min).

If not stated differently all chemicals were obtained from Roth chemicals

3.3.2 H3 Medium

Solution A	
KH_2PO_4	2.3 g
Na_2HPO_4 x 2 H_2O	2.9 g
Distilled water	50 mL

Solution B	
NH_4Cl	1 g
$MgSO_4$ x 7 H_2O	0.5 g
$CaCl_2$ x 2 H_2O	0.01 g
$MnCl_2$ x 4 H_2O	0.005 g
$NaVO_3$ x H_2O	0.005 g
Trace element solution SL-6	5 mL
Distilled water	915 mL

Solution C	
Ferric ammonium citrate	0.05 g
Distilled water	20 mL

Solutions A, B, C are autoclaved separately for 15 min at 121 °C, cooled down to 50 °C and then mixed aseptically with 5.0 mL filter-sterilized standard vitamin solution (see below) and 10.0 ml filter-sterilized 5% $NaHCO_3$ (pH 7-8).The final pH of this medium should be 6.8 without adjustment.The medium is supplemented with 0.2% (m/v) glucose as a carbon source.

Standard vitamin solution	
Riboflavin	10 mg
Thiamine-HCl x 2 H_2O	50 mg
Nicotinic acid	50 mg
Pyridoxine-HCl	50 mg
Ca-pantothenate	50 mg
Biotin	0.1 mg
Folic acid	0.2 mg
Cobalamin	1 mg
Distilled water	100 mL

Trace element solution SL-6	
$ZnSO_4$ x 7 H_2O	0.1 g
$MnCl_2$ x 4 H_2O	0.03 g
H_3BO_3	0,3 g
$CoCl_2$ x 6 H_2O	0,2 g
$CuCl_2$ x 2 H_2O	0.01 g
$NiCl_2$ x 6 H_2O	0.02 g
Na_2MoO_4 x 2 H_2O	0.03 g
Distilled water	1000 mL

3.3.3 H3 Medium supplemented with watermelon

The medium was prepared according to 3.3.2 without the addition of glucose. Parts of a watermelon were shredded to a degree that parts of the shell were still intact (**Figure 5**, right). 10 g of the shredded watermelon were added to 1 L of the H3 medium and autoclaved.

Figure 5: Preparation of a watermelon for the preparation of supplemented H3 medium

3.3.4 LB medium

Tryptone	10.0 g
Yeast extract	5.0 g
NaCl	10.0 g
Distilled water	1000 mL

The medium is adjusted to pH 7.0 prior to autoclave sterilization (121 °C, 15 min).

3.3.5 ISP2 medium with 10% NaCl

Yeast extract	4.0 g
Malt extract	10.0 g
Glucose	4.0 g
NaCl	50.0 g

The medium is adjusted to pH 7.3 prior to autoclave sterilization (121 °C, 15 min).

3.3.6 Small scale cultivation of *Massilia* sp. NR. 4-1 and *A. citrulli* AAC00-1

Small scale cultivation experiments were used for the search of novel metabolites. Furthermore, the experiments using different iron concentrations, the ^{13}C-carbon source feeding experiments and the identification of violacein and its byproducts were carried out in a small scale. Strains NR 4-1 and AAC00-1 were cultivated in 20 mL of medium (the type of medium is stated in chapter 4) in a 100 mL buffled flask on a rotary shaker at 160 rpm and 30 °C for 24 h. This seed culture was transferred to a

a 500 mL buffled flask containing 200 mL medium on a rotary shaker at 160 rpm and 30 °C for 5 days. The fermentation broth was extracted three times with ethyl acetate and the solvent was removed in vacuo. The dried extracts were resuspended in 1.5 mL methanol, filtered and supplied to the HPLC.

3.3.7 Large scale cultivation of *Massilia* sp. NR. 4-1

Large scale cultivation conditions were used for the isolation and structure elucidation experiments of massiliachelin, massargiline and violacein. Strain NR 4-1 was cultivated in 20 mL of medium (the type of medium is stated in chapter 4) in a 100 ml buffled flask on a rotary shaker at 160 rpm and 30 °C for 24 h. This seed culture was transferred to a 500 mL buffled flask containing 100 mL of medium. In a second upscaling step this culture was again transferred after 24 h to a 3 L flask containing 1L of medium. With this inoculum strain NR 4-1 was cultured in three 5 L Simax flasks each containing 2 L of medium. The cultures were shaken at 160 rpm and 30 °C. After seven days of cultivation the fermentation broth was extracted three times with ethyl acetate and the solvent was removed in vacuo. For the isolation of violacein the *Massilia* culture was mixed 1:1 with methanol and centrifuged for 10 minutes at a rotation speed of 4500 rpm. The liquid phase was separated from the cell mass by decanting. The remaining cell mass was washed with methanol. The dried extracts were resuspended in 3 mL methanol, filtered and supplied to the HPLC.

3.3.8 Complexation study with gallium

A methanolic solution of massiliachelin was treated with a 15-fold excess of Ga(III) nitrate hydrate (Sigma Aldrich) and stirred for 24 h at room temperature.

3.3.9 ^{13}C feeding studies

[1-^{13}C]-glucose, and [1-^{13}C]-acetate (Sigma-Aldrich) were fed aseptically to the medium. In the feeding experiments marked as “fed batch” the labeled precursor was added aseptically after 20 h of incubation. The precursors were fed in a concentration of 100 mg/L if not stated differently.

3.3.10 Quantification of massargiline production

A stock solution containing 1.0 mg of massargiline dissolved in 10 mL methanol was prepared. Different volumes of this solution were injected into the HPLC in order to

generate a calibration curve (**Figure A 29**). The extract of 200 mL culture was dissolved in 1.5 mL methanol. Fifty µL of each sample were applied to the HPLC. The resulting peak areas and the calculated concentrations are listed in **Table A 2**. The first culture was extracted directly after inoculation to take the produced amount of massargiline by the preculture into account. This value is set to a concentration of 0 mg/L as a reference.

3.3.11 Antimicrobial testing

The determination of the MIC values was carried out by Dr. Florian Kloß, Hans-Knöll-Institute Jena.

The determination of the antiprotozoal testing was carried out by Dr. Marcel Kaiser, Swiss Tropical and Public Health Institute, Basel, Switzerland according to the following protocol [193]:

Activity against *Trypanosoma brucei rhodesiense* STIB900. This stock was isolated in 1982 from a human patient in Tanzania and after several mouse passages cloned and adapted to axenic culture conditions [194]. Minimum Essential Medium (50 µL) supplemented with 25 mM HEPES, 1 g/L additional glucose, 1% MEM non-essential amino acids (100x), 0.2 mM 2-mercaptoethanol, 1mM Na-pyruvate and 15% heat inactivated horse serum was added to each well of a 96-well microtiter plate. Serial drug dilutions of eleven 3-fold dilution steps covering a range from 100 to 0.002 µg/mL were prepared. Then 4x103 bloodstream forms of *T. b. rhodesiense* STIB 900 in 50 µL was added to each well and the plate incubated at 37 °C under a 5% CO_2 atmosphere for 70 h. 10 µL resazurin solution (resazurin, 12.5 mg in 100 mL double-distilled water) was then added to each well and incubation continued for 2–4 h [195]. Then the plates were read with a Spectramax Gemini XS microplate fluorometer (Molecular Devices Cooperation, Sunnyvale, CA, USA) using an excitation wavelength of 536 nm and an emission wavelength of 588 nm. Data were analyzed with the graphic programme Softmax Pro (Molecular Devices Cooperation, Sunnyvale, CA, USA), which calculated IC_{50} values by linear regression [196] and 4-parameter logistic regression from the sigmoidal dose inhibition curves. Melarsoprol (Arsobal Sanofi-Aventis, received from WHO) is used as control.

Activity against *T. cruzi.* Rat skeletal myoblasts (L-6 cells) were seeded in 96-well microtitre plates at 2000 cells/well in 100 µL RPMI 1640 medium with 10% FBS and

2 mM L-glutamine. After 24 h the medium was removed and replaced by 100 µL per well containing 5000 trypomastigote forms of *T. cruzi* Tulahuen strain C2C4 containing the β-galactosidase (Lac Z) gene [197]. After 48 h the medium was removed from the wells and replaced by 100 µL fresh medium with or without a serial drug dilution of eleven 3-fold dilution steps covering a range from 100 to 0.002 µg/mL. After 96 h of incubation the plates were inspected under an inverted microscope to assure growth of the controls and sterility. Then the substrate CPRG/Nonidet (50 µL) was added to all wells. A color reaction developed within 2–6 h and could be read photometrically at 540 nm. Data were analyzed with the graphic programme Softmax Pro (Molecular Devices), which calculated IC_{50} values by linear regression [196] and 4-parameter logistic regression from the sigmoidal dose inhibition curves. Benznidazole is used as control (IC_{50} 0.5±0.2 µg / mL).

Activity against *L. donovani* axenic amastigotes. Amastigotes of *L. donovani* strain MHOM/ET/67/L82 are grown in axenic culture at 37 °C in SM medium [198] at pH 5.4 supplemented with 10% heat-inactivated fetal bovine serum under an atmosphere of 5% CO_2 in air. One hundred µL of culture medium with 10^5 amastigotes from axenic culture with or without a serial drug dilution are seeded in 96-well microtitre plates. Serial drug dilutions of eleven 3-fold dilution steps covering a range from 100 to 0.002 µg/mL are prepared. After 70 h of incubation the plates are inspected under an inverted microscope to assure growth of the controls and sterile conditions. 10 µL of resazurin (12.5 mg resazurin dissolved in 100 mL distilled water) are then added to each well and the plates incubated for another 2 h. Then the plates are read with a Spectramax Gemini XS microplate fluorometer (Molecular Devices Cooperation, Sunnyvale, CA, USA) using an excitation wavelength of 536 nm and an emission wavelength of 588 nm. From the sigmoidal inhibition curves the IC_{50} values are calculated by linear regression [196] and 4-parameter logistic regression using SoftmaxPro software (Molecular Devices Cooperation, Sunnyvale, CA, USA).

Activity against *P. falciparum*. In vitro activity against erythrocytic stages of *P. falciparum* was determined using a ^{3}H-hypoxanthine incorporation assay [199], using the drug sensitive NF54 strain [200] and the standard drugs chloroquine (Sigma C6628) and artesunate (Sigma A3731). Compounds were dissolved in DMSO at 10 mg/mL and further diluted in medium before added to parasite cultures incubated in RPMI 1640 medium without hypoxanthine, supplemented with HEPES (5.94 g/L),

$NaHCO_3$ (2.1 g / L), neomycin (100 U/mL), Albumax^R (5 g/L) and washed human red cells A^+ at 2.5% haematocrit (0.3% parasitaemia). Serial drug dilutions of eleven 3-fold dilution steps covering a range from 100 to 0.002 µg/mL were prepared. The 96-well plates were incubated in a humidified atmosphere at 37 °C; 4% CO_2, 3% O_2, 93% N_2. After 48 h 50 µL of ^{3}H-hypoxanthine was added to each well of the plate. The plates were incubated for a further 24 h under the same conditions. The plates were then harvested with a Betaplate™ cell harvester (Wallac, Zurich, Switzerland), and the red blood cells transferred onto a glass fibre filter then washed with distilled water. The dried filters were inserted into a plastic foil with 10 mL of scintillation fluid, and counted in a Betaplate™ liquid scintillation counter (Wallac, Zurich, Switzerland). IC_{50} values were calculated from sigmoidal inhibition curves by linear regression [196] using Microsoft Excel. Chloroquine and artemisinin are used as control.

In vitro cytotoxicity with L-6 cells*.* Assays were performed in 96-well microtiter plates, each well containing 100 µL of RPMI 1640 medium supplemented with 1% L-glutamine (200 mM) and 10% fetal bovine serum, and 4000 L-6 cells (a primary cell line derived from rat skeletal myoblasts) [201], [202]. Serial drug dilutions of eleven 3-fold dilution steps covering a range from 100 to 0.002 µg/mL were prepared. After 70 h of incubation the plates were inspected under an inverted microscope to assure growth of the controls and sterile conditions. 10 µL of resazurin was then added to each well and the plates incubated for another 2 hours. Then the plates were read with a Spectramax Gemini XS microplate fluorometer (Molecular Devices Cooperation, Sunnyvale, CA, USA) using an excitation wavelength of 536 nm and an emission wave length of 588 nm. The IC_{50} values were calculated by linear regression [196] and 4-parameter logistic regression from the sigmoidal dose inhibition curves using SoftmaxPro software (Molecular Devices Cooperation, Sunnyvale, CA, USA). Podophyllotoxin (Sigma P4405) is used as control.

Boullion dilution assay. The MIC values were tested against *Mycobacterium vaccae* 10670, *Staphylococcus aureus* 134/94, *Enterococcus* 1528, *Pseudomonas aeruginosa* K799/61 *and Escherichia coli* SG458 and *Sporobolomyces salmonicolor.* The organism was cultured in 500 µL aliquots of tryptic soy broth (peptone casein 1.7% (w/v), peptone from soymeal 0.3% (w/v), D-glucose 0.25% (w/v), NaCl 0.5% (w/v), K_2HPO_4 0.25% (w/v), pH 7.3) at 37 °C for 16 h. Before incubation, massargiline was added in decreasing concentrations (100 µg/mL to 12.5 µg/mL). After incubation,

the OD_{600} was measured. The experiment was performed in triplicate. Ciprofloxacin was used as a positive control.

4 Results

4.1 AntiSMASH analysis of *Massilia* sp. NR 4-1

Massilia sp. NR 4-1 is a β-proteobacterium which had been isolated from a soil sample collected under a nutmeg tree in South Korea [203]. Its genome has a size of 6.36 Mbp and comprises 5285 coding DNA sequences with a GC content of 63.80%, which is consistent with the genus description [160]. The NCBI database lists 57 *Massilia* strains of which 7 are completely sequenced [204]. The average genome size of these bacteria is 5.6 Mbp. Furthermore, Miess et. al sequenced 6 *Massilia* strains with genome sizes ranging from 5.9 to 7.5 Mbp and 6 to 11 BGCs per genome [205]. Bioinformatic analysis of the genome of strain NR 4-1 using antiSMASH 4.0 revealed a total of 15 BGCs. Based on this result strain NR 4-1 shows a secondary metabolite (SM) coding capacity of 0.73 Mbp or 11% revealing an encouraging potential of secondary metabolites biosynthesis. The SM coding capacity of bacteria was investigated before [7]. Proteobacteria are considered to have a high SM capacity if they devote 9–17% of their genomes to the production of secondary metabolites. Within the β-proteobacteria three *Burkholderia* species ranging from 6.7 to 9.1 Mb showed the highest SM coding capacity with 9-14 %.

The antiSMASH analysis of the genome from *Massilia* sp. NR 4-1 reveals a diverse array of BGCs, including loci involved in the biosynthesis of polyketides, nonribosomal peptides, terpenes and hybrids thereof. The complete list is shown in **Table 1**. According to the prediction strain NR 4-1 has the genetic potential for the production of the purple pigment violacein. Three gene clusters showed moderate similarity (50-80%) whereas the remaining 11 gene clusters show low similarity (<50%) to known BGCs. The similarity describes the degree of likeness between two sequences [30]. The bioinformatics results suggest that *Massilia* sp. NR 4-1 could be a potential source for novel secondary metabolites.

Table 1: Results of the antiSMASH analysis of *Massilia* sp. NR 4-1

BGC	Type	From bp	To bp	Most similar known cluster	Similar-ity
1	NRPS, T1PKS	502,605	563,276	micacocidin	70%
2	NRPS	862,657	930,163	gacamide A	42%
3	NRPS, T1PKS	973,455	1,024,326	tubulysin A	6%
4	NRPS	1,039,439	1,112,542	glidopeptin	50%
5	T1PKS, NRPS	1,204,076	1,291,258	oxalomycin B	6%
6	transAT-PKS-like, transAT-PKS, PKS-like, NRPS	1,364,719	1,468,611	thailandamide / thailandamide lactone	73%
7	transAT-PKS, NRPS, T1PKS, acyl amino acids	1,482,690	1,655,838	crochelin A	46%
8	terpene	1,676,612	1,698,310		
9	bacteriocin	3,050,336	3,062,054		
10	T3PKS	3,312,023	3,353,093	ketomemicin B3 / ketomemicin B4	33%
11	Acyl amino acids	4,063,577	4,124,221	N-tetradecanoyl tyrosine	6%
12	homoserine lactone	4,322,112	4,343,796		
13	indole	4,596,587	4,619,604	violacein	100%
14	NRPS	4,955,982	5,001,666		
15	arylpolyene	6,230,842	6,274,474	aryl polyenes	44%

4.2 Isolation and structure elucidation

4.2.1 Genomic inspired discovery of massiliachelin

The BGC number 1 (**Table 1**) shows 70% similarity to the *mic* locus from the plant pathogenic bacterium *Ralstonia solanacearum* GMI1000, which is involved in the

biosynthesis of the siderophore micacocidin [101], [206]. Micacocidin is a thiazoline-containing natural product that shows significant activity against *Mycoplasma pneumonia* [206]. The *mic* locus covers 38.1 kbp of contiguous DNA and includes three core biosynthesis genes for enzymes with a function as PKS (MicG), NRPS (MicH) and PKS/NRPS hybrid (MicC). In a first step, MicC activates and loads hexanoic acid. Chain elongation, ring formation, and aromatization lead to the synthesis of 6-pentylsalicylate. Subsequently two cysteines are attached and cyclized [206]. MicG introduces a C2-unit whose carbonyl group at C1 is reduced. The position 2 is methylated twice in NADPH- and SAM-dependent steps. MicH condenses and cyclizes a third cysteine unit forming a thiazoline ring.

The locus in strain NR 4-1 with a size of 34.7 kbp includes ten genes (RS02190-RS02235), of which seven have homologs in the *mic* cluster [101]. Both clusters show the presence of a TonB-dependent receptor, as well as two iron-siderophore ABC transporters indicating an involvement in the biosynthesis of an iron complexing compound. A further comparison of the core biosynthetic genes shows a strong conservation of *micC* and *micG*. In contrast to the micacocidin BGC the NRPS encoding gene *micH* homolog is missing in the *Massilia* locus. The absence of this homolog inidicates a production of a truncated micacocidin derivative which lacks the second thiazoline ring. The micacocidin-type BGC from *Massilia* sp. NR 4-1 and its enzymatic assembly line are depicted in **Figure 6**.

Agrochelin is a metal-chelating thiazole alkaloid which was isolated from a marine bacterium of the genus *Agrobacterium*. It shows structural similarity to micacocidin, but is missing the terminal thiazoline ring [207]. The choice of cultivation conditions can be directed by the assumption that strain NR 4-1 produces a natural compound that has a similar constitution. A common approach to unambiguously identify siderophores is the cultivation under iron deficient conditions [208].

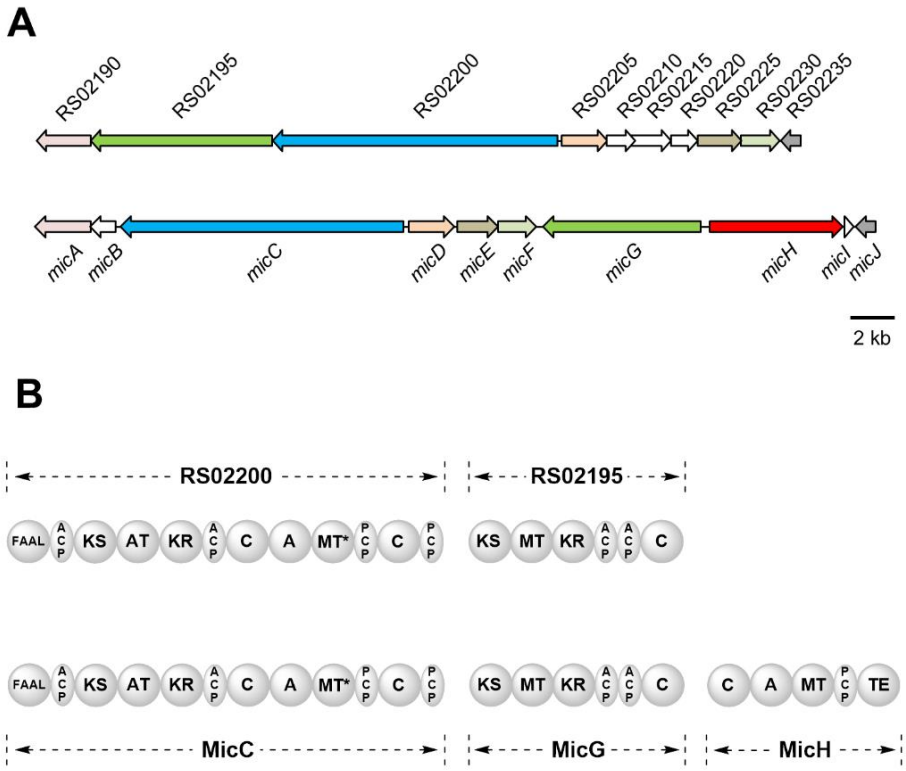


Figure 6: A) Organization of the micacocidin-type gene cluster from *Massilia* sp. NR 4-1 (top) and of the *mic* gene cluster from *R. solanacearum* GMI1000 (bottom). The open reading frames are color-coded according to their predicted function. B) Domain architecture of the massiliachelin (top) and the micacocidin (bottom) assembly lines. Domain notation: FAAL, fatty acyl-AMP ligase; ACP, acyl carrier protein; KS, β-ketoacyl synthase; AT, acyltransferase; KR, ketoreductase; C, condensation; A, adenylation; MT, methyltransferase; PCP, peptidyl carrier protein; TE, thioesterase. The asterisk indicates a methyltransferase-like epimerization domain. Modified from [190]

Massilia sp. NR 4-1 was cultured in R2A medium with the addition of 10 µM of $FeCl_3$ and without supplementation of iron. The ethyl acetate extract of the two cultures were applied to the HPLC. The comparison of the UV profiles revealed a peak that is only present under iron deficiency conditions (**Figure 7**). A 6 L culture of strain NR 4-1 was extracted and 12.0 mg of the corresponding metabolite, which is referred to as massiliachelin, were isolated.

High-resolution electrospray ionization (ESI-MS) of massiliachelin yielded a pseudomolecular ion peak at *m/z* 467.2033 $[M + H]^+$ (**Figure A 7**). This value matches exactly the calculated formula $C_{23}H_{34}N_2O_4S_2$ (calc: 467.2033), which is identical with the formula of agrochelin [207]. It accounts for eight degrees of unsaturation. For NMR spectroscopy, the substance was dissolved in deuterated chloroform. The signals in the ^1H NMR spectrum and the ^{13}C NMR spectrum accounted for 31 non-exchangeable protons and 23 carbons (**Table 2**).

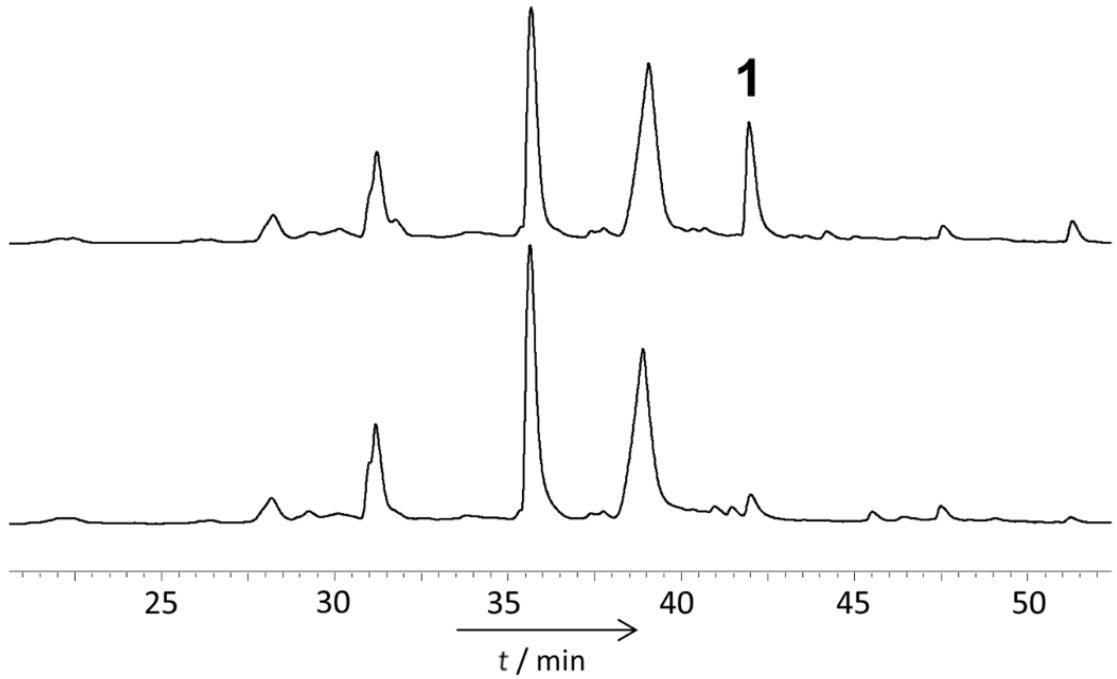


Figure 7: UV chromatograms of crude culture extracts from *Massilia* sp. NR 4-1 grown under iron-deficient (top) and iron-replete (bottom) conditions. The chromatograms were recorded at a wavelength of 274 nm. Modified from [190]

Eight carbon atoms must be sp^2-hybridized according to their chemical shifts and six of them (C-1 to C-6) can be attributed to a tri-substituted benzene moiety. The substituent at C-1 was assigned to a hydroxy group (δ_{C1} = 158.8 ppm). The *n*-pentyl side chain moiety in position 3 can be identified by COSY and HMBC data. Further evaluation of the COSY spectrum allows the identification of two further spin systems, C13-C14-C15 and C16-C17-C19. HMBC correlations from H-13 and H-14 to C-12 in combination with characteristic chemical shifts of C12, C13 and C14 known from other natural products like yersiniabactin and pyochelin allow the identification of a thiazoline substituent [209], [210]. HMBC correlations confirm the linkage to the benzene ring in *ortho* position to the hydroxy group. Likewise, the thiazolidine moiety, including C-15. C-16 and C-17, was determined. A long range correlation from H_3-18 to C-15 indicates *N*-methylation of this ring. Finally, the HMBC correlations from H_3-21 and H_3-22 to the quaternary carbon atoms C-20, C-19, and the last sp^2-hybridized carboxyl C-23, establish the terminal part of the molecule. The hydroxy group at C-19 and the carboxyl group at C-23 were deduced from the carbon chemical shift values.

Table 2: NMR spectroscopic data of massiliachelin in $CDCl_3$

pos.	δ_C, type	δ_H, mult. (*J* in Hz)	COSY	HMBC	NOESY
C1	158.8, C				
C2	112.6, C				
C3	144.7, C				
C4	122.2, CH	6.76, d (7.4)	5	1, 2, 3, 6, 7, 12	5, 7a, 7b, 8
C5	135.9, CH	7.33, t (8.2, 7.8)	4, 6	1, 2, 3	4, 6
C6	116.1, CH	7.22, d (8.5)	5	2, 4, 12	5
C7	35.1, CH_2	a: 2.92, ddd (14.4, 9.8, 6.4)	7b, 8	2, 3, 4, 8	4, 7b, 8, 9
		b: 2.76, ddd (14.4, 9.8, 6.3)	7a, 8	2, 3, 4, 8	4, 7a, 8, 9
C8	31.8, CH_2	1.58, m	7a, 7b, 9	3, 7, 9, 10	7a, 7b, 9, 10
C9	31.7, CH_2	1.33, m	8, 10	7, 11	7a, 7b, 8, 10, 11
C10	22.4, CH_2	1.33, m	9, 11	8	9, 11
C11	14.0, CH_3	0.89, t (7.1)	10	9, 10	9, 10
C12	182.3, C				
C13	33.5, CH_2	a: 3.72, dd (12.1, 9.9)	13b, 14	12, 14, 15	13b, 14
		b: 3.46, dd (12.1, 6.4)	13a, 14	12, 14, 15	13a, 14, 15
C14	70.7, CH	5.09, td (9.9, 6.4)	13a, 13b, 15	12, 15	13a, 13b, (15), 16
C15	80.4, CH	4.20, d (9.9)	14	13, 14, 16, 17, 18	13b, (14), 18
C16	36.1, CH_2	3.23, d (7.9)	17	15, 17	14, 17, 19
C17	73.1, CH	3.49, dt (7.9, 2.6)	16, 19	15, 16, 18	16, 18, 19, 22
C18	47.3, CH_3	2.66, s		15, 17	15, 17, 22
C19	78.6, CH	3.31, d (2.6)	17	16, 20, 21, 22, 23	16, 17, 21, 22
C20	44.3, C				
C21	26.5, CH_3	1.38, s		19, 20, 22, 23	19
C22	22.1, CH_3	1.25, s		19, 20, 21, 23	17, 18, 19
C23	181.3 C				

The structure of massiliachelin is depicted in **Figure 8**. Although massiliachelin possesses the same chemical constitution as agrochelin, a comparison of NMR data sets revealed differences in the chemical shift values and the coupling patterns [207]. The relative configuration of agrochelin was deduced by the evaluation of the NOESY spectrum. The observation of a strong NOE between H-14 and H-15 and

between H-17 and N-Me, but not between H-15 and H-17 led to the determination of an relative configuration of 14*R*,15*S*,17*R*,19*S* [207]. The NOESY spectrum of massiliachelin, however, showed only a weak correlation between H-14 and H-15. On the other hand, strong NOEs from the *N*-methyl protons in position 18 to H-17 and H-15 were observed. This observation is only possible if the latter protons are *syn*-oriented. Therefore, a configuration of 14*R*, 15*R*, 17*R*, 19*S* is deduced for massiliachelin. Thus, massiliachelin and agrochelin form a pair of epimers. The occurrence of two interconvertible epimers of natural products with a thiazoline–thiazolidine motif is known from other natural products like yersiniabactin, micacocidin and pyochelin (**Figure 8**) [211]–[213].

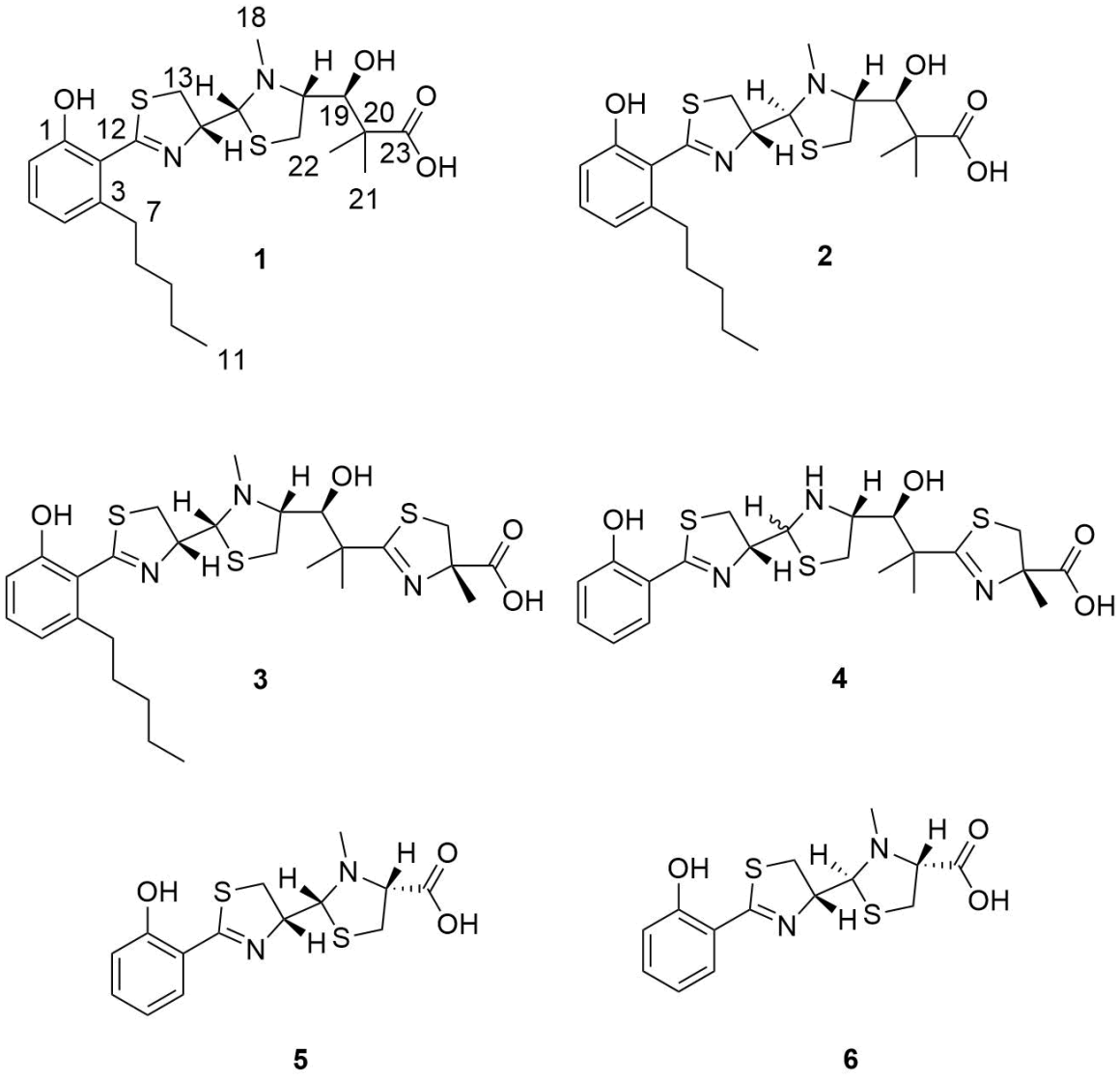
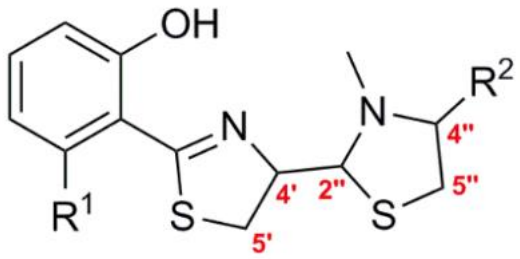


Figure 8: Structures of massiliachelin (1), agrochelin (2), micacocidin (3), yersiniabactin (4). pyochelin I (5), pyochelin II (6). Modified from [190]

The comparison of the ^{1}H NMR spectra of both epimers revealed a difference of the coupling pattern at position 14. The respective hydrogen of agrochelin is split into a quartet, whereas in massiliachelin a triplet of doublets is observed due to the different coupling constants (9.9 Hz and 6.4 Hz). These differences in the ^{1}H NMR spectrum are consistent with other known pairs of epimers containing a thiazoline–thiazolidine motif (**Table 3**) [207], [211]–[213].

Table 3: Stereochemistry of siderophores with a thiazoline-thiazolidine motif.

Natural product	Configuration	δ(H-4'), m (*J* in Hz) measured in $CDCl_3$	Reference
Pyochelin I	4'*R*, 2''*R*, 4''*R*	5.08, td (8.8, 5.0)	[212]
Yersiniabactin I	4'*R*, 2''*R*, 4''*R*	5.15, td (9.2, 4.8)	[211]
Massiliachelin	4'*R*, 2''*R*, 4''*R*	5.09, td (9.9, 6.4)	
Neopyochelin I	4'*S*, 2''*S*, 4''*R*	5.12, td (9.1, 4.6)	[212]
Pyochelin II	4'*R*, 2''*S*, 4''*R*	4.92, q (8.2)	[212]
Yersiniabactin II	4'*R*, 2''*S*, 4''*R*	4.78, q (n.d.)	[213]
Agrochelin	4'*R*, 2''*S*, 4''*R*	4.72, q (8.7)	[214]
Neopyochelin II	4'*S*, 2''*R*, 4''*R*	4.79, q (8.0)	[212]



The described pairs of epimers are known to underlie a spontaneous interconversion in most solvents. NMR studies of the pyochelins were carried out with a mixture of the two epimers due to the good distinction of data sets [213].

An inspection of the massiliachelin NMR spectra also revealed two distinct sets of data, which are easily distinguishable due to integration of the corresponding ^{1}H signals and evaluation of two dimensional NMR data (**Figure A 1**). The epimers massiliachelin and agrochelin are chromatographically separable by using an isocratic method with increased solvent flow (**Figure 9A**). Fractionation of the peaks

at t_R = 12 min and t_R = 12,6 min and instantaneous reinjection led to an increased concentration of the corresponding epimer (**Figure 9B**). After 24 h, the equilibrium state was nearly restored (**Figure 9C**).

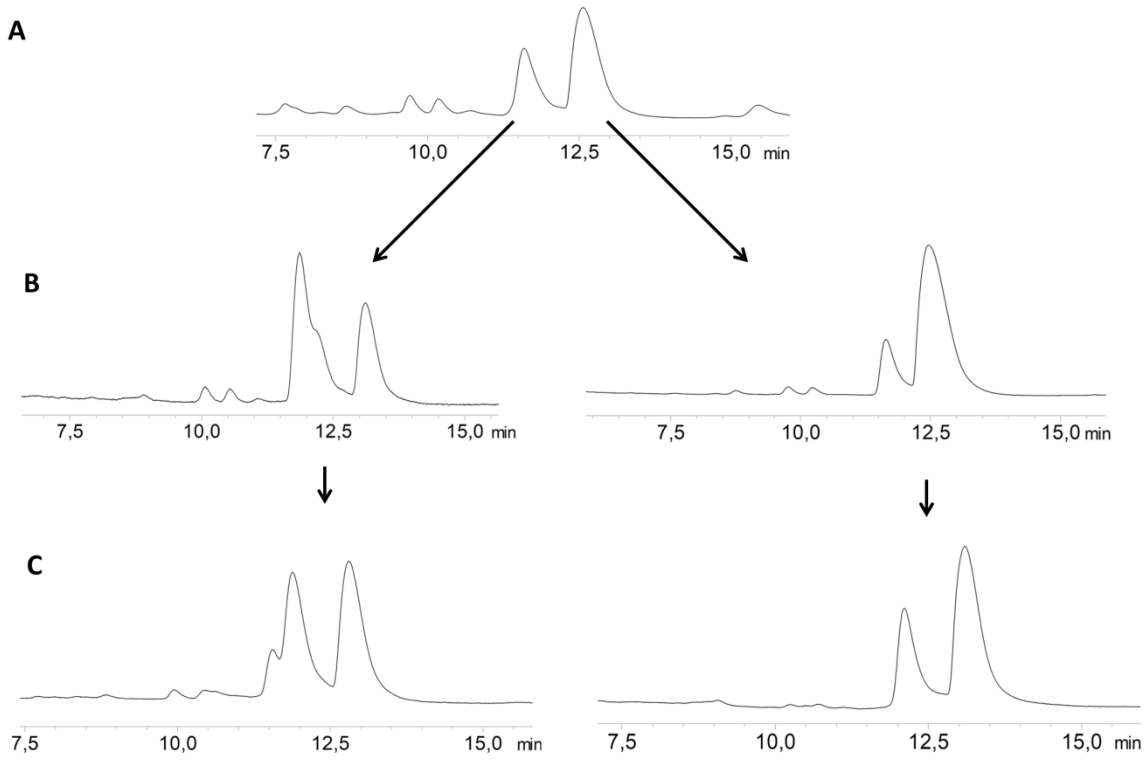


Figure 9: Interconversion reaction of massiliachelin. The two peaks at t_R = 12 min and t_R = 12,6 min were fractionated (A). The fractions were reinjected into the HPLC (B). The intensity of the peaks reverts to the equilibrium after 24 h (C).

To prevent the interconversion reaction of thiazolidine-containing siderophores, complexation with gallium can be carried out [215]. Massiliachelin was complexed with Ga^{3+} according to a protocol, which showed appropriate results for micacocidin [206]. The MS results revealed the formation of a massiliachelin-gallium complex and the presence of unbound massiliachelin (**Figure A 24**). Gallium complexes can be easily identified by the *m/z* value of the complex in combination with the distinctive isotope pattern of gallium (^{69}Ga: 60,1%; ^{71}Ga: 39,9%) [216]. The NMR data of the sample confirmed the presence of multiple data sets.

4.2.2 The isolation of violacein

Myeong et al. identified *Massilia* sp. NR 4-1 in the course of the genome sequencing as a potential producer of violacein. The antiSMASH result confirmed the presence of the violacein gene cluster including the genes *vioA* to *vioE* (BGC 13 in **Table 1**) [203]. Violacein has a molecular mass of 343.34 g/mol. Deoxyviolacein and oxyviolacein are byproducts of the violacein synthesis that are produced in lower concentrations. Their structures differ in the absence (deoxyviolacein) and the presence (oxyviolacein) of an additional hydroxyl group (**Figure 10**).

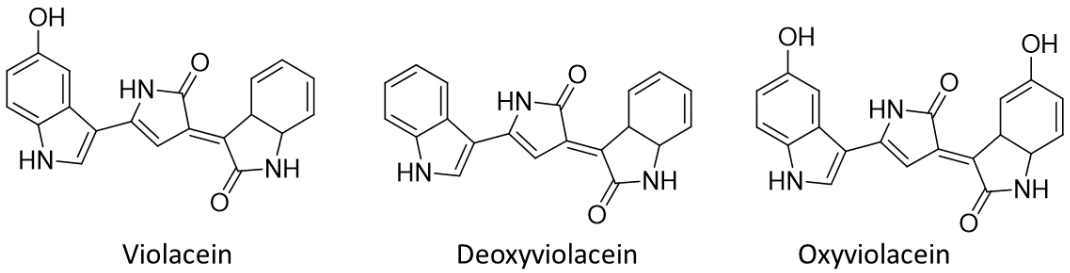


Figure 10: Structures of violacein, deoxyviolacein, and oxyviolacein

The mixture of violacein and deoxyviolacein is generally referred to as crude violacein [217].

The production of violacein by *Massilia* sp. NR 4-1 has not been experimentally confirmed prior to this PhD project. *Massilia* sp. BS-1 is the only *Massilia* strain, in which violacein production was investigated before [170]. In this study the authors used a minimal medium (MM2) under variation of different carbon sources for production. Additionally the dependence of the tryptophan and the histidine concentration was investigated. In summary, *Massilia* sp. BS-1 was found to produce the highest amounts of violacein by using a histidine concentration of 32 µM, a tryptophan concentration of 1 mM and glucose as a carbon source. By the usage of a 3-L jar fermentor with aeration (0.5 v/v/m) and agitation (300 rpm) at 25 °C, a yield of 1.3 mM crude violacein was observed after 24 h.

To investigate the violacein production of *Massilia* sp. NR 4-1, the production parameters of *Massilia* sp. BS-1 were used as a reference. Considering the molecular masses of histidine (M = 155.15 g/mol) and tryptophan (M = 204.23 g/mol) 11 mg and 200 mg of the corresponding amino acid were added aseptically to 1 L of

H3 medium. The prepared medium was inoculated with 50 mL of a 2 day old *Massilia* culture in a 5-L buffled flask. After 4 days, 500 mL of the culture were extracted three times with ethyl acetate. Violacein possesses a strong chromophore which allows to evaluate the efficiency of extraction by the remaining violet coloring of the aqueous phase. Rodriques et al. carried out an extraction after centrifugation with glass beans. The ethyl acetate extraction was performed “until no more pigment color accumulated in the organic phase” [218].

After the third extraction of the strain NR 4-1 extract, the aqueous phase remained purple due to residual violacein. Therefore, the remaining 500 mL of *Massilia* culture were mixed with 500 mL methanol and centrifuged for 10 minutes at a rotation speed of 4500 rpm. The liquid phase was separated from the cell mass by decanting. The remaining cell mass was washed with 500 mL methanol to dissolve the remaining violacein. The absence of purple coloring showed a high extraction efficiency using a 1:1 methanol/water mixture (**Figure 11**).

Figure 11: Comparison of the extraction efficiency of ethyl acetate (right) and methanol extraction (left).

The extract was dried in vacuo and dissolved in 1.5 mL methanol. Violacein shows an absorption maximum at a wavelength of 582 nm in methanol [219]. UV/VIS monitoring allowed an unambiguous assignment of the violacein peak in the HPLC chromatogram at a retention time of t_R = 43.5 minutes (**Figure 12**). The byproducts oxyviolacein (t_R = 35.4 min) and deoxyviolacein are detectable (t_R = 46.2 min) with lower intensity.

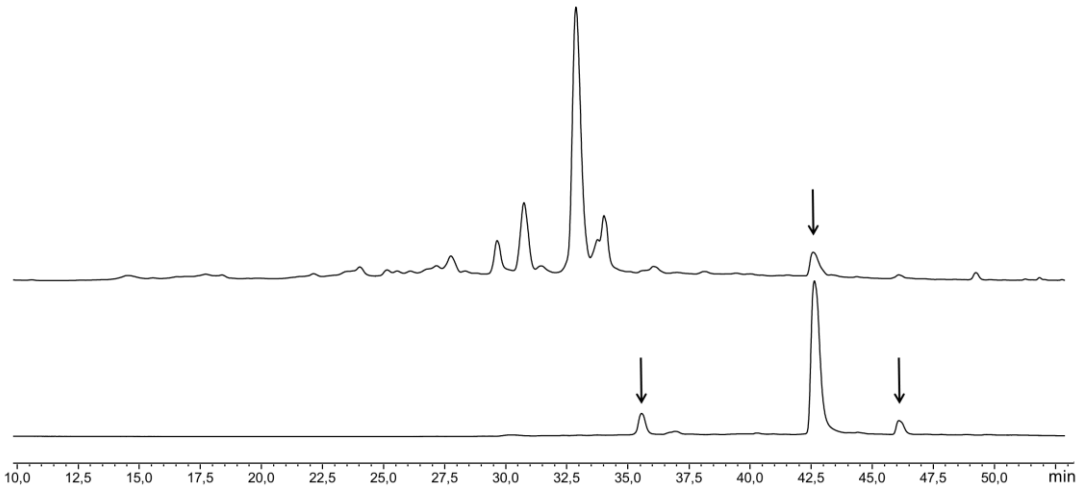


Figure 12: UV chromatograms of *Massilia* extract monitored at a wavelength of 270 nm (top) and 570 nm (bottom).

Many violacein concentrations stated in literature are calculated via the absorption coefficient and accompanying spectrophotometric analyses. However, a literature review by Choi et al. revealed that the reported values of the absorption coefficient range from 10.9 to 74.3 L / (g*cm) at a wavelengths between 570 and 575 nm in methanol [220]. Due to the ambiguity of the photometric quantification, both the extract and the violacein were analysed gravimetrically. The extraction of 1L *Massilia* culture and subsequent chromatographic separation led to 473 mg extract and 9.3 mg purified violacein.

The isolated violacein was dissolved in 1 mL methanol-d_4 and applied to ^{1}H and ^{13}C NMR measurement. According to the low signal intensity, the compound was not completely soluble in the solvent. In DMSO-d_6 the signal intensity significantly increased (**Figure A 22**). The observed chemical shifts are consistent with literature data [221].

4.2.3 The discovery of massargiline

The isolation of massiliachelin and violacein were successful examples of secondary metabolite isolation studies that were guided by genomic analyses. According to the antiSMASH results, *Massilia* sp. NR 4-1 harbors further potential for the biosynthesis of yet unidentified natural products (**Table 1**). In order to identify these compounds, strain NR. 4-1 was cultivated in complex R2A medium and extracted with ethyl acetate after 5 days of incubation. The analysis of the chromatogram revealed a

further distinctive peak with a UV maximum of 270 nm which was absent in the chromatogram of the medium control (**Figure A 28**). Three liters of *Massilia* were cultured and the marked peak was subsequently isolated gaining 8.9 mg of the unknown compound, which is referred to as massargiline (**Figure 13**).

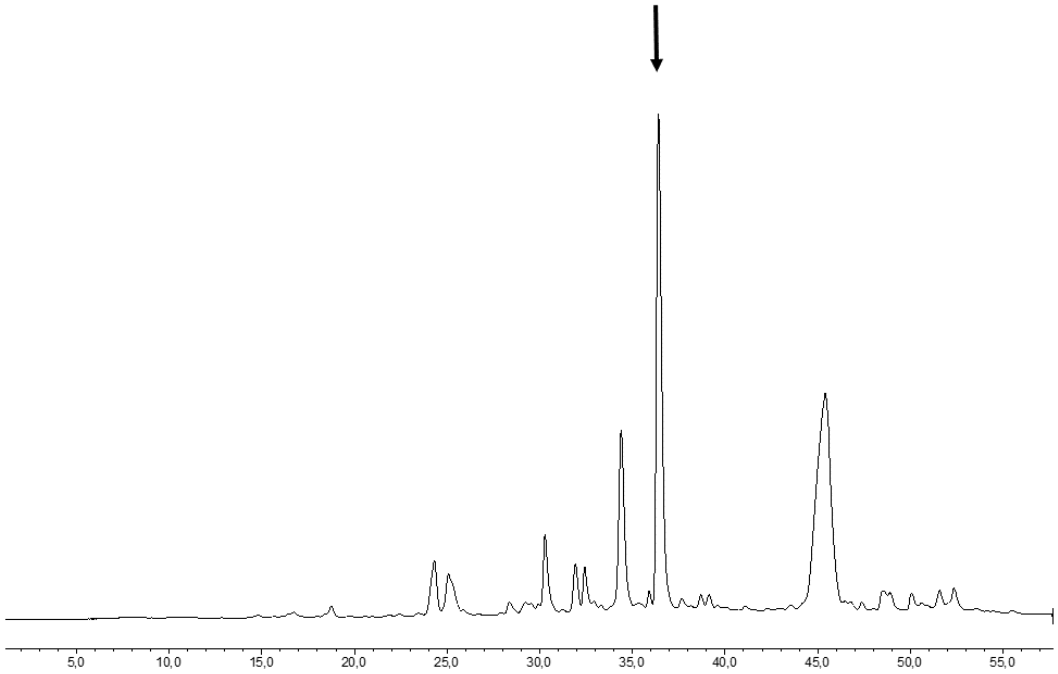


Figure 13: UV chromatogram of a crude culture extracts from *Massilia* sp. NR 4-1 recorded at 270 nm. The isolated peak at a retention time of 37 min is marked with an arrow.

The molecular formula of massargiline was determined as $C_{15}H_{17}N_3$ based on its molecular ion at *m/z* 240.1496 $[M + H]^+$ (**Figure A 8**). NMR data were recorded in DMSO-d_6. The results are shown in **Table 4**.

Nine degrees of unsaturation are calculated from the molecular formula. NMR measurements confirmed the number of 15 carbon atoms and revealed the presence of 17 protons in DMSO-*d6.* Eleven carbon atoms of massargiline are sp^2-hybridized according to their chemical shifts. Six of these atoms can be attributed to a monosubstituted benzene moiety. In accordance, the ^1H NMR spectrum shows five aromatic signals, including three triplets and two doublets. COSY data enabled an elucidation of the fragment C11-C6-C8, in which the carbon atoms C-6 and C-8 form an olefinic double bond with typical chemical shift values. The double bond at position 9 and 3 is connected to the COSY fragment featuring the methylene groups 10, 12 and 13 and to the fragment C11-C6-C8. This connection is confirmed by long

range correlations from H-13 to C-9 and from H-8 to C-3, respectively. HMBC correlations between CH-5 and CH_2-11 allow to connect the benzene moiety.

Table 4: NMR data of massiliargiline in DMSO-d_6

#	δ_C		δ_H, M (*J* in Hz)	COSY	HMBC ($^1H \rightarrow ^{13}C$)
1	142.9	C_q			
2	139.8	C_q			
3	128.9	C_q			
4, 4'	128.5	CH x 2	7.30, t (8.4, 6.9)	5, 7	2, 4
5, 5'	128.4	CH x 2	7.21, d (8.4)	4	5, 7, 11
6	127.9	CH	6.08, dt (16.0, 7.0)	8, 11	2, 3, 9, 11
7	126.1	CH	7.21, t (6.9)	4	5
8	116.6	CH	6.17, d (16.0)	6	2, 3, 9, 11
9	115.6	C_q			
10	44.0	CH_2	3.78, t (7.3)	12	1, 3, 12, 13
11	38.3	CH_2	3.47, d (7.0)	6	2, 5, 8
12	28.1	CH_2	2.45, quint (7.3)	10, 13	3, 10, 13
13	21.4	CH_2	2.76, t (7.3)	12	3, 9, 10, 12
14		NH	7.65, s		1
15		NH	12.22, s		1, 3, 9

The resulting structure accounts for six degrees of unsaturation and 14 carbon atoms. The remaining sp^2-hybridized carbon atom at position 1 shows a chemical shift value of 142.9 ppm. It must therefore be attached to a nitrogen atom via a double bond. Two double bond equivalents remain for the assignment of ring structures. The NH-group at position 15 can be deduced from HMBC correlations to the carbon atoms 3 and 9. Furthermore, long range correlations from H_2-10 to C-1 and C-33 indicate the presence of a pyrrolo[1,2-*c*]imidazole ring system. The proposed structure is depicted in **Figure 14**.

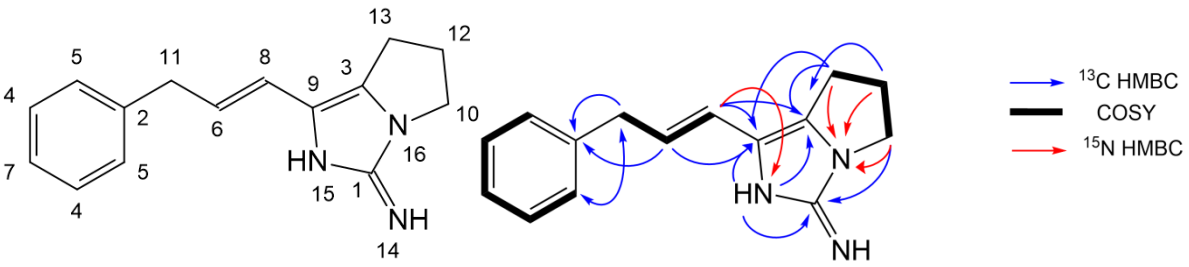


Figure 14: Structure of massargiline (left) and COSY, key ^{13}C HMBC and ^{15}N HMBC correlations in this natural product (right).

Subsequently, ^{1}H, ^{15}N HSQC and ^{1}H, ^{15}N HMBC spectra of massargiline were recorded in order to verify the proposed structure. To obtain sufficient material for this experiment, *Massilia* sp. NR 4-1 was cultivated in R2A medium at a 12-L scale. Initial fractionation of the obtained extract yielded 38 mg of crude massargiline. Since NMR analysis showed the presence of a contamination, the material was further purified with a Pyramid C18 column. This led to 24 mg of pure massargiline. The ^{15}N NMR results are shown in **Table A 1.** Evaluation of the HMBC correlations revealed the presence of four long range correlations from H-8 to N-15 allowing an unambiguous distinction between the nitrogen atoms 14 and 15. Correlations from H_2-10, H_2-11 and H_2-12 to N-14 complete the assignment of all three nitrogen atoms (**Figure 14**).

The quaternary guanidine carbon atom of massargiline (C-1) shows a chemical shift of 142.9 ppm in DMSO-d_6. Bedford et al. carried out NMR studies of 16 fused oxoheterocycles with different substitution patterns which contain a guanidine unit [222]. The reported chemical shift values range from 152.6 ppm to 165.4 ppm in DMSO-d_6. Compared to these literature data the observed value in massargiline is clearly shifted upfield. A commercially avalilable substance with structural similarity to massargiline is 6,7-dihydro-5*H*-pyrrolo[1,2-*c*]imidazol-3-amine (**Figure 15**).

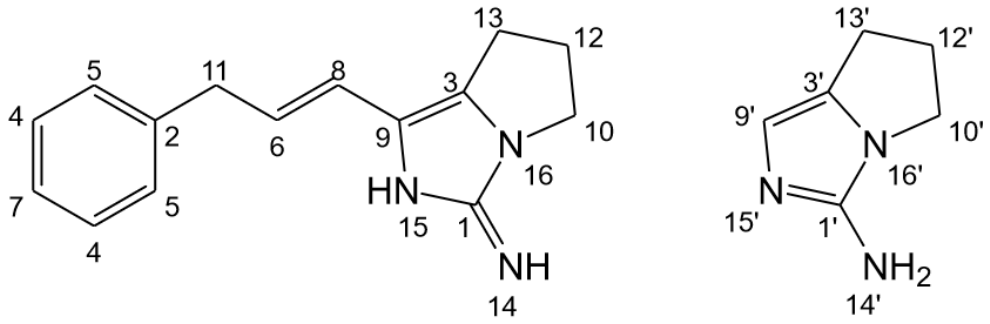


Figure 15: Comparison between the structures of massargiline (left) and 6,7-dihydro-5H-pyrrolo[1,2-c]imidazol-3-amine (right).

^{1}H, ^{13}C and ^{15}N NMR measurements of this reference were carried out to provide experimental evidence supporting the structural proposal for massargiline. Fifty milligrams of the compound were measured in DMSO-d_6. The results are consistent with the values observed for massargiline (**Table 5**) and, hence, support the structural proposal.

Table 5: Comparison of NMR data related to massargiline and 6,7-dihydro-5H-pyrrolo[1,2-c]imidazol-3-amine in DMSO-d_6

#	δ_C	δ_H, M (J in Hz)	#	δ_C	δ_H, M (J in Hz)
15			15‘		
1	144.6		1‘	145.2	
16			16'		
3	128.1		3‘	131.5	
9	117.0	6.55 t (1,3)	9‘	114.8	
10	43.7	3.78 t (7.3)	10‘	42.4	3.51 t (7.0)
12	28.3	2.45 q (7.3)	12‘	30.3	2.15 q (7.1)
13	21.3	2.76 t (7.3)	13‘	22.1	2.60 t (7.1)
14		7.65 s	14‘		5.2 s

4.3 Biosynthesis

4.3.1 Massiliachelin

In chapter 4.2.1 the absolute configuration of massiliachelin was deduced by spectroscopic methods. Agrochelin and massiliachelin were identified as pair of epimers. In addition to the spectroscopic data, a bioinformatic analysis might enable to predict the configurational assignment of natural products harboring the same thiazoline–thiazolidine motive.

Pyochelin, yersiniabactin, and micacocidin all harbor D-thiazoline rings, which participate in iron chelation. The corresponding heterocycles derive from D-cysteine. Usually the presence of epimerization (E) domains in NRPS assembly lines allows the prediction of a D-amino acid residue [210], [223]. It was shown before that pyochelin and yersiniabactin synthetases activate and load only L-cysteine, although sequence analysis of the corresponding gene clusters did not reveal any E domains [224], [225]. For the protein PchE of the pyochelin assembly line it was demonstrated that an unusual methyltransferase-like domain acts as an E domain. Experiments with deuterium led to the assumption that this domain catalyzes the formation of a carbanionic species. The reprotonation from solvent protons most likely occurs stereorandomly. The *N*-acyl-cysteine-S-enzyme or aryl-thiazoline-S-enzyme intermediates with D centers are subsequently selectively processed by the enzymatic assembly lines [223].

The presence or absence of an E domain in the biosynthetic assembly lines of micacocidin [101], yersiniabactin [226], and enantio-pyochelin [227] can be used as an indicator for the configuration of the generated thiazoline ring. By applying this method to the MicC homolog of *Massilia* sp. NR 4-1 the D configuration of the thiazoline motif can be deduced (**Table 6**). The configuration of the thiazolidine ring can be determined by a comparison of MicC of micacocidin, HMWP2 of yersiniabactin and RS02200 of massiliachelin. They all feature a single adenylation domain which is known to accept only L-cysteine [223]. The assembly line of pyochelin is an example, which features two adenylation domains (located in PchE and PchF) for the utilization of two cysteine residues. The configuration of C-17 therefore derives from L-cysteine and can be assigned as *R.*

Table 6: Epimerization (E) domains in pyochelin-type synthetases and configuration of the generated thiazoline ring

	Protein accession #	Methyltransferase-like E domain in PchE homolog	Proposed configuration of thiazoline ring
Pyochelin	AAD55800	Yes	D
Enantio-pyochelin	ABW70809	No	L
Yersiniabactin	P48633	Yes	D
Micacocidin	CAD15508	Yes	D
Massiliachelin	AKU20507	Yes	(D)

Domain swapping experiments showed that KR domains are responsible for the stereochemical outcome of ketoreduction reactions in PKS assembly lines [228]. The KR domains are classified as A-type or B-type. The classification depends on the resulting absolute configuration of the alcohol and the priority of the two carbon atoms in β-position. The exact combination of absolute configurations and priorities defining each type is well illustrated by Caffrey [229]. A correlation of KR domain motifs with alcohol stereochemistry revealed only few consistent differences. The motif LDD from residue 93 to 95 is indicative for the generation of a B-type alcohol stereochemistry. Only D95 is invariant, whereas L93 and D94 are less strictly conserved. On the other hand, none of the A-type KR domains has a hydrophobic residue followed by two acidic residues in the same positions. The other difference between the two groups can be found in the region from residue 141 to 144. In the B group, position 144 is typically a proline residue. Members of the A-type typically lack this residue, but show a strong conservation of a tryptophan residue at position 141.

The KR domains in the micacocidin, yersiniabactin and massiliachelin assembly lines all show the two motifs indicating the formation of B-type alcohol stereochemistry (**Table 7**). This informatics-driven assignment is validated by the observation of an absolute configuration of *S* in micacocidin and yersiniabactin at the corresponding position. Thus, the absolute configuration for massiliachelin at C-19 is deduced as *S*.

Table 7: Alignment of the residue 88–103 and 134–148 regions of the ketoreductase (KR) domains in MicG from *Ralstonia solanacearum* GMI1000, HMWP1 from *Yersinia pestis* KIM6+, and ACZ_RS02195 from *Massilia* sp. NR 4-1.

	Protein accession#	**88–103 region**	**134–149 region**	**KR-type**	**Configuration (proposed)**
MicG	CAD15512	HLAGIVRDAPLAAADWR	FSSAASACGAPGQGAH	B	S
HMWP1	AAC69588	HAAGVLADAPLQELDDH	YSSAAATLGAPGQSAH	B	S
RS02195	WP_050407226	HLAAVVRDATLAAINTE	FSSAATAFGAPGQGAY	B	(S)

From the bioinformatic analysis and the chemical structure of massiliachelin, a possible biosynthetic route was deduced. According to this model, the first two modules of RS02200 (FAAL-ACP-KS-AT-KR-ACP) initiate the biosynthesis by forming a δ-hydroxylated tetraketide intermediate. Subsequent intramolecular aldol condensation and aromatization leads to thioester-bound 6-pentylsalicylate corresponding to the formation of the same building block in micacocidin biosynthesis [206].

The first NRPS module in RS02200 catalyses the condensation of cysteine and the following cyclization. In the epimerization mechanism, deprotonation leads to the formation of a carbanion at position C-14, which results in sterically random reprotonation from the solvent. Intermediates with the D-thiazoline ring are selectively translocated to the next downstream module. Module 3 catalyses the formation of the second thiazolidine ring. *S*-adenosyl-L-methionine (SAM)-dependent methylation at position 18 is carried out by the MT domain in module 2. Module 4 of enzyme RS02195 defines the switching point from NRPS back to PKS logic. A Claisen condensation with malonyl-CoA followed by two SAM-dependent methylations leads to an intermediate that is released by the TE domain [226]. The reduction of the second thiazoline ring by a thiazolinyl imide reductase ACZ75_02195 leads to the final product massiliachelin (**Figure 16**). Further studies need to be done to determine the exact timing of this reduction. It is assumed that the action of the imine reductase PchG of pyochelin is required before the release from the enzymatic

assembly line [231]. The timing of the YbtU-mediated reduction in the biosynthesis of yersiniabactin is still unknown [226].

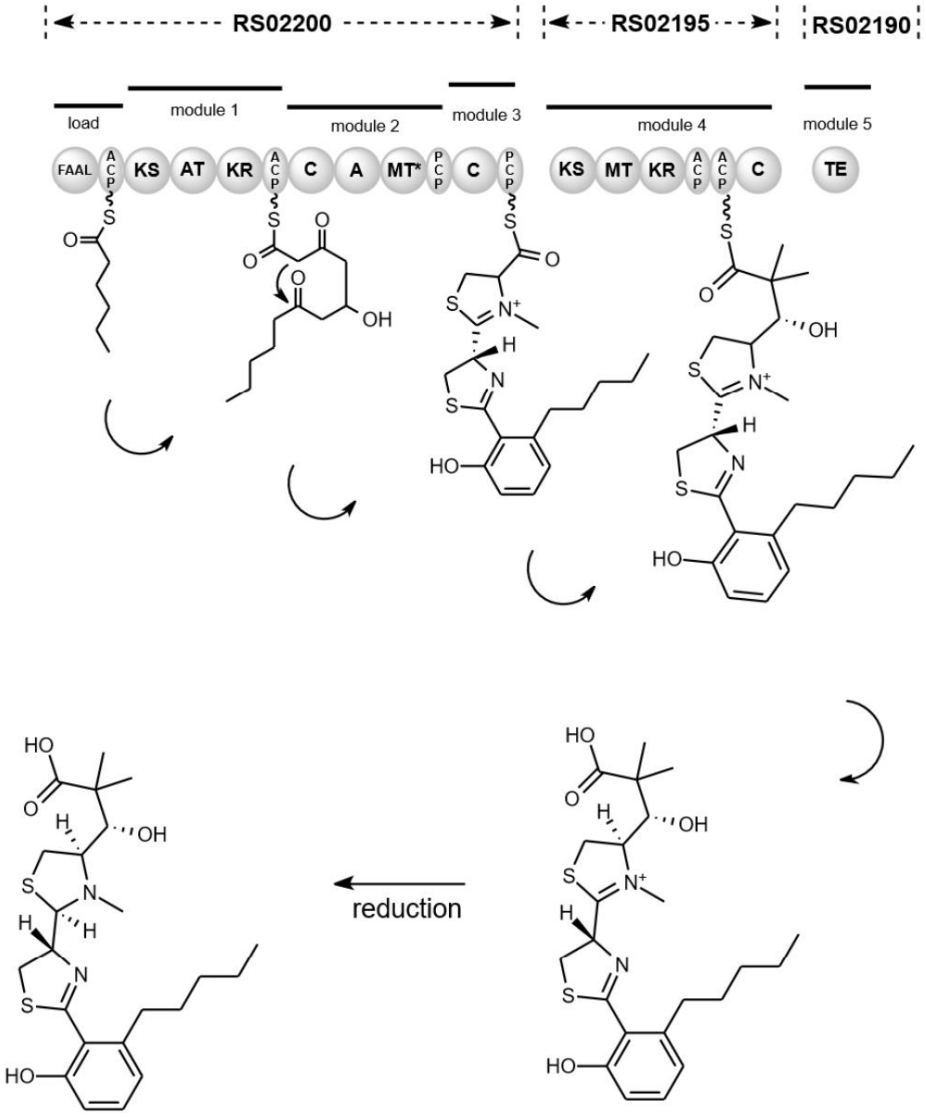


Figure 16: Proposed biosynthesis of massiliachelin

4.3.2 Massargiline

Massargiline features a guanidine moiety which is integrated into a pyrrolo[1,2-*c*]imidazole ring system. In nature, guanidine groups typically derive from the amino acid arginine or its degradation products (e.g., agmatine) [232]. The chemical structure of massargiline allowed to identify a putative gene cluster shown in **Table 1** responsible for the biosynthesis. After a closer inspection of the previously identified

BGCs in *Massilia* sp. NR 4-1 cluster 10 was identified as a promising candidate (**Table 8**).

Table 8: Putative massargiline gene cluster from *Massilia* sp. NR 4-1

Name	ORF Locus Tag	Proposed function	Protein ID
msaA	ACZ75_13490	arginine-tRNA synthetase	AKU22327.1
msaB	ACZ75_13495	hypothetical protein	AKU24886.1
msaC	ACZ75_13500	dehydratase	AKU22328.1
msaD	ACZ75_13505	ornithine cyclodeaminase	AKU22329.1
msaE	ACZ75_13510	enoyl-CoA hydratase/isomerase	AKU22330.1
msaF	ACZ75_13520	dehydrogenase/reductase	AKU22331.1
msaG	ACZ75_13530	aminotransferase class I and II	AKU22332.1
msaH	ACZ75_13535	T3PKS	AKU22333.1
msaI	ACZ75_13545	transporter	AKU22334.1
msaJ	ACZ75_13550	dihydrofolate reductase	AKU22335.1
msaK	ACZ75_13555	hypothetical protein	AKU22336.1
msaL	ACZ75_13560	hypothetical protein	AKU24887.1
msaM	ACZ75_13565	hypothetical protein	AKU22337.1
msaN	ACZ75_13570	hypothetical protein	AKU22338.1
msaO	ACZ75_13575	hypothetical protein	AKU22339.1
msaP	ACZ75_13580	Drug resistance transporter	AKU24888.1
msaQ	ACZ75_13585	peptidase S9	AKU22340.1
msaR	ACZ75_13595	thioesterase	AKU22341.1
msaS	ACZ75_13600	peptidase M20	AKU22342.1

The corresponding locus encodes a type III PKS and an arginyl-tRNA synthetase. Considering the aforementioned structural characteristics of a benzene moiety and a guanidine group, these enzymes suggest the involvement of this BGC in the biosynthesis. A possible biosynthetic route to massargiline involves seven steps (**Figure 17**). First, the type III PKS catalyzes the sequential decarboxylative condensations of five malonyl-CoA units followed by an intramolecular cyclization. A

pentaketide is also synthesized by the type III PKS 1,3,6,8-tetrahydroxynaphthalene synthase (THNS). The proposed resorcinol-like intermediate in massargiline biosynthesis is also considered as a favorable intermediate leading via S- or U-shaped cyclization patterns to tetrahydroxynaphthalene [233].

Subsequently, the polyketide intermediate is condensed with the amino acid arginine by the aminotransferase AKU22332.1. This enzyme shows similarity to a pyridoxal 5′-phosphate (PLP)-dependent amino acid C_{α}-acyltransferase. PLP is a coenzyme for a variety of amino acid reactions like decarboxylation, racemization or elimination which are usually catalyzed by a single PLP-dependent enzyme [234], [235]. An exceptional subgroup of the PLP-dependent enzymes is the α-oxamine synthase (AOS) family. Enzymes of this family catalyze the condensation between an amino acid and an acyl-CoA normally followed by a decarboxylation [236], [237]. Here, the enzyme catalyzes a C-C-bond formation between the Cα of arginine and the carbonyl carbon of the acyl-CoA-thioester, followed by a decarboxylation to generate an α-oxoamine structure.

ACZ75_13505 encodes a putative ornithine cyclodeaminase. These enzymes are known to catalyze the formation of proline from ornithine by releasing ammonium [238], [239]. Here, the enzyme is proposed to catalyze the cyclization and deamination of the arginine-containing intermediate leading to the formation of a pyrrolidine ring.

The second five membered ring is formed upon condensation. There are two putative enzymes present in the gene locus for the catalysis of this enamine-like structure. ACZ75_RS13555 codes for a protein of the M20/M25/M40 metallo-hydrolase family. There are metallo-hydrolases that are known to catalyze the hydrolysis of five-membered rings. Isatin hydrolase is an example that catalyses the reaction from isatin to isatinate [240], [241]. The second candidate is a peptidase of the S9 family encoded by ACZ75_13585. The condensation reaction between the carbonyl function and the amine leads to an enamine-like structure with the establishment of a double bond at position 3. Finally the reduction steps are putatively carried out by the reductase AKU22331 and the dehydratase AKU22328. MsaF belongs to the SDR (Short-chain dehydrogenases/reductases) superfamily [242]. Several representatives of this family are known to participate in the biosynthesis of polyketides [243], [244].

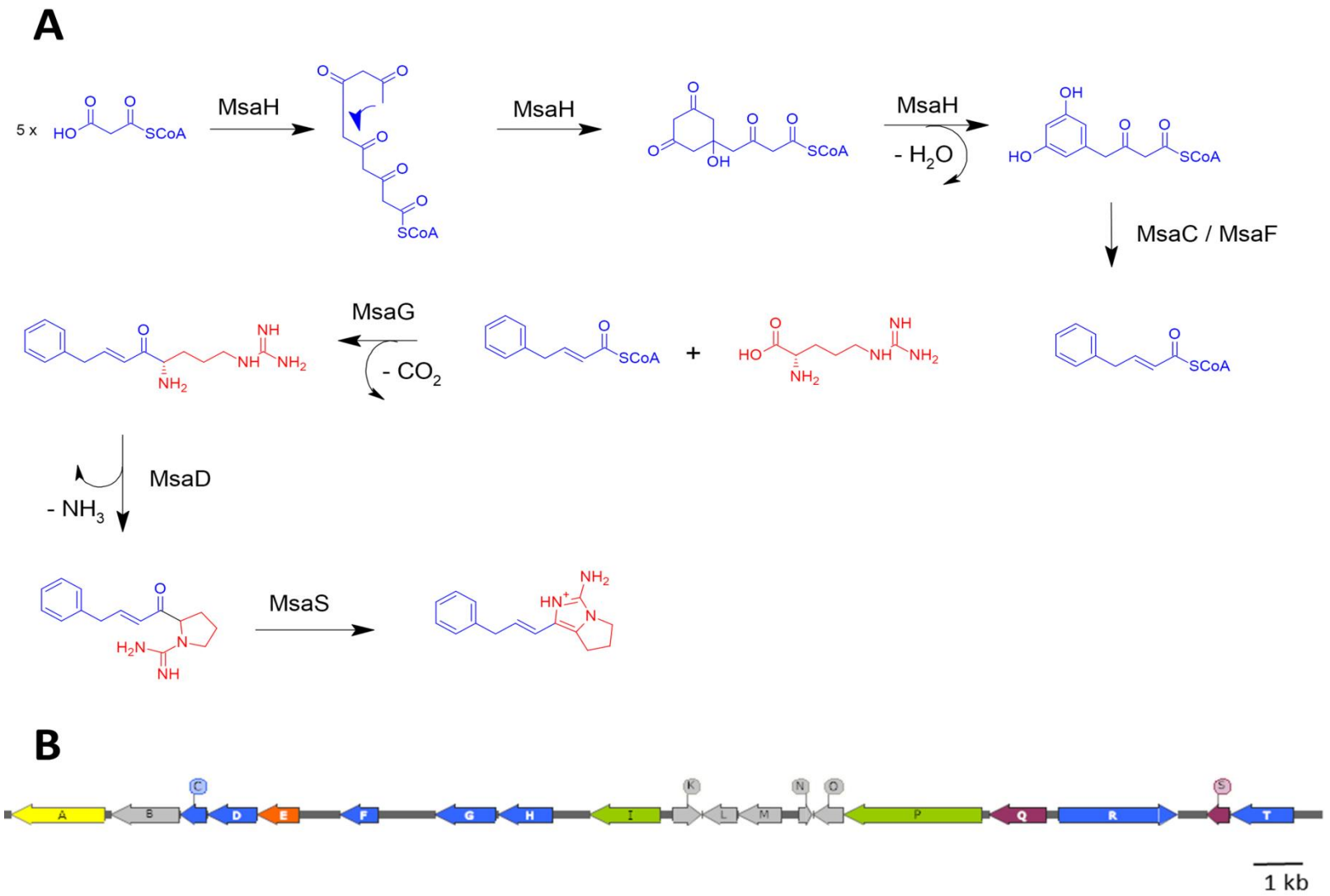


Figure 17: Proposed biosynthetic route of massargiline (A). The malonyl-CoA derived atoms are coloured in blue. The arginine derived atoms are coloured in red. Organization of massargiline gene cluster (B)

It must be noted, however, that the precise timing of the biosynthetic reactions might differ from the proposal depicted in **Figure 17**, e.g. regarding the MsaD-catalyzed cyclization. In order to confirm the proposed pathway, individual expression experiments can be carried out, as demonstrated in the ketomemicin pathway [245]. In this PhD study, feeding studies using isotopically labeled precursors were carried out in order to explore the biosynthesis of massargiline. For this, the kinetics of massargiline production had to be established first.

To investigate the production of massargiline, six 200 mL cultures of *Massilia* NR 4-1 were incubated in R2A medium over 4 days and samples were drawn in regular intervals. The massargiline concentration in the prepared extracts was determined via HPLC (**Figure 18**). The calibration curve and the peak areas are shown in **Table A 2** and **Table A 3**. In the same period, the optical densities of the different cultures were measured at a wavelength of 600 nm.

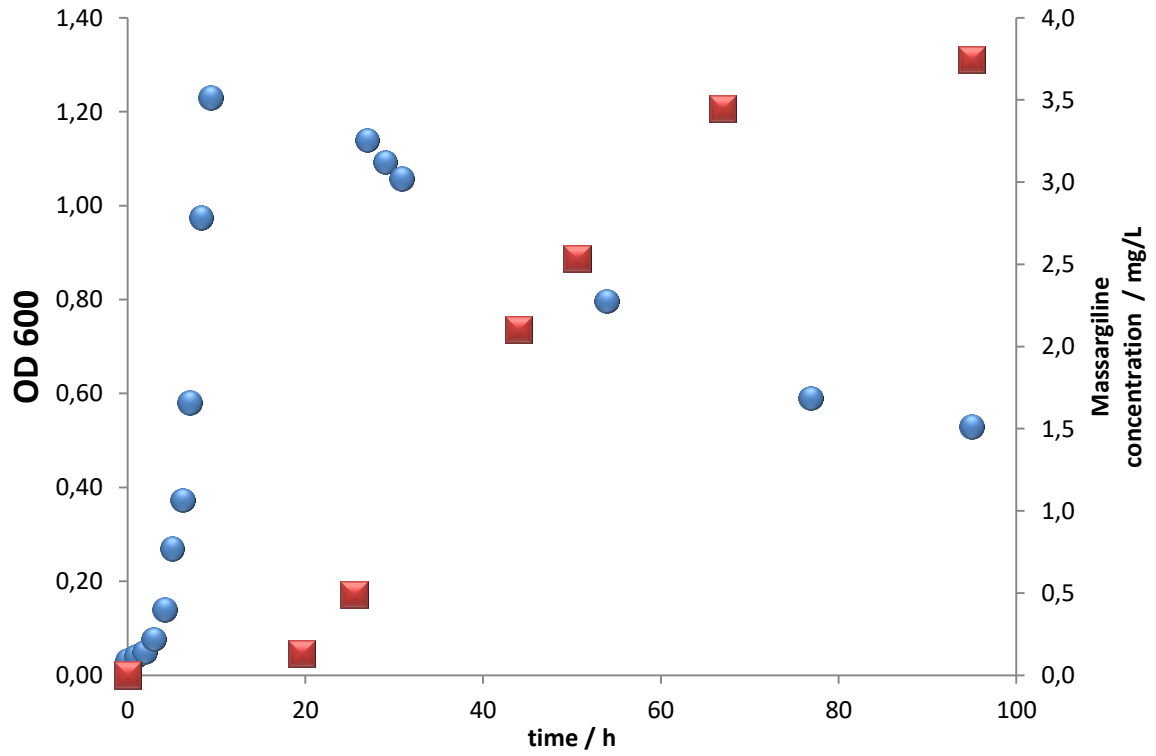


Figure 18: Optical density of *Massilia* sp. NR 4-1 (blue dots) and production level of massargiline (red squares) measured over 96 hours

In this study, strain NR 4-1 reached the stationary growth phase after 10 hours of incubation, while massargiline production started 20 hours after inoculation. Interestingly, the evaluation of the chromatograms revealed an already significant

production level of the putative siderophore massiliachelin at this time point (**Figure A 30**). After 96 hours the production level of massargiline raised to a value of 3.7 mg/L. To verify the assumed polyketide origin of massargiline, a feeding study with [1-^{13}C]-acetate was initiated. According to the postulated biosynthetic proposal, ^{13}C enrichment was expected in five carbon atoms of massargiline, i.e. at positions 2, 4, 4', 6 and 9. A comparison of MS data of *Massilia* sp. NR 4-1 crude extracts revealed that massargiline is produced in complex R2A medium, but not in minimal H3 medium (**Figure 19**).

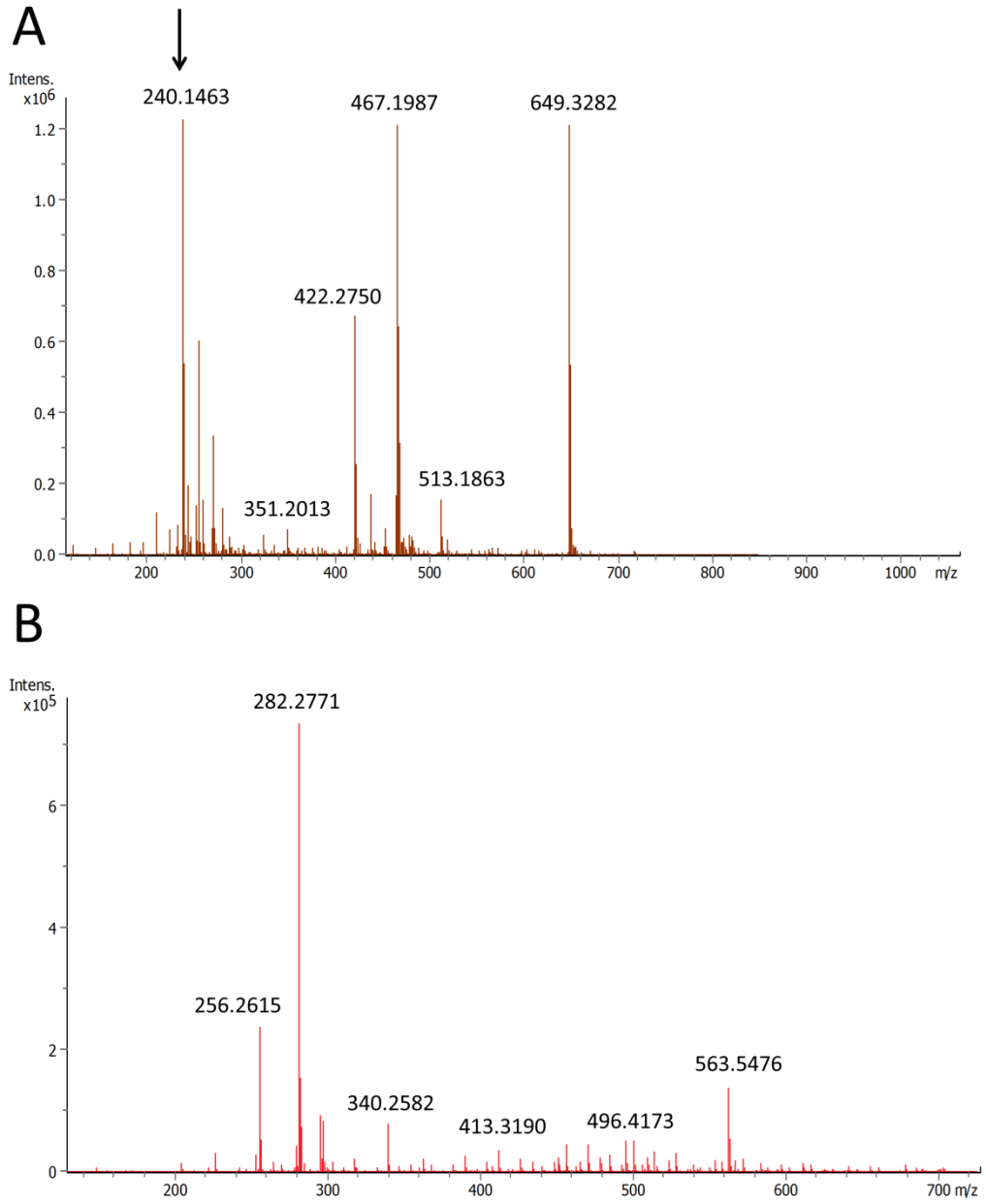


Figure 19: Comparison of *Massilia* sp. NR 4-1 HR-ESI-MS spectra of extracts cultivated in R2A medium (A) and H3 minimal medium (B). The *m/z* value of massargiline is marked with an arrow.

Initially, feeding experiments using complex R2A with labeled [1-^{13}C]-glucose (**Figure A 31**) and [1-^{13}C]-acetate (**Figure A 32**) did not lead to the production of massargiline with a ^{13}C enriched isotopic pattern.

Complex R2A medium contains glucose and starch as carbon sources. As the MS data revealed the production with incorporation of unlabeled carbon atoms further experiments with R2A medium containing lower concentrations of the unlabeled carbon sources (= reduced R2A medium) were carried out. Previous measurements had revealed that massargiline biosynthesis started after 20 hours (**Figure 18**). Therefore, labeled [1-^{13}C]-acetate was fed at this timepoint to minimize its distribution in the primary metabolism. Subsequent MS analysis did not show the expected mass increase of the massargiline ion by 5 Da. Instead, a shift in the ratio of the isotopes with *m/z* 240.2 and 241.2 was observed (**Figure A 33**).

In another experiment the more lipophilic ethyl[1-^{13}C]-acetate was fed to a *Massilia* NR 4-1 culture after 20 hours in reduced R2A medium. MS evaluation of the ethyl acetate extract did not show a change in the isotopic pattern compared to the control (**Figure A 34**).

A further experiment was carried out in R2A medium without the addition of glucose and starch. [1-^{13}C]-acetate was added aseptically after 20 hours of incubation. The comparison of the mass spectra between this sample and a control with no additional labeled carbon source showed a gradual increase of the *m/z* value of 240.2 up to 5 Da to a value of 245.2 (**Figure 20**). This elemental composition of five ^{13}C incorporations matches perfectly the expected value and strengthens the hypothesis of a polyketide origin of massargiline.

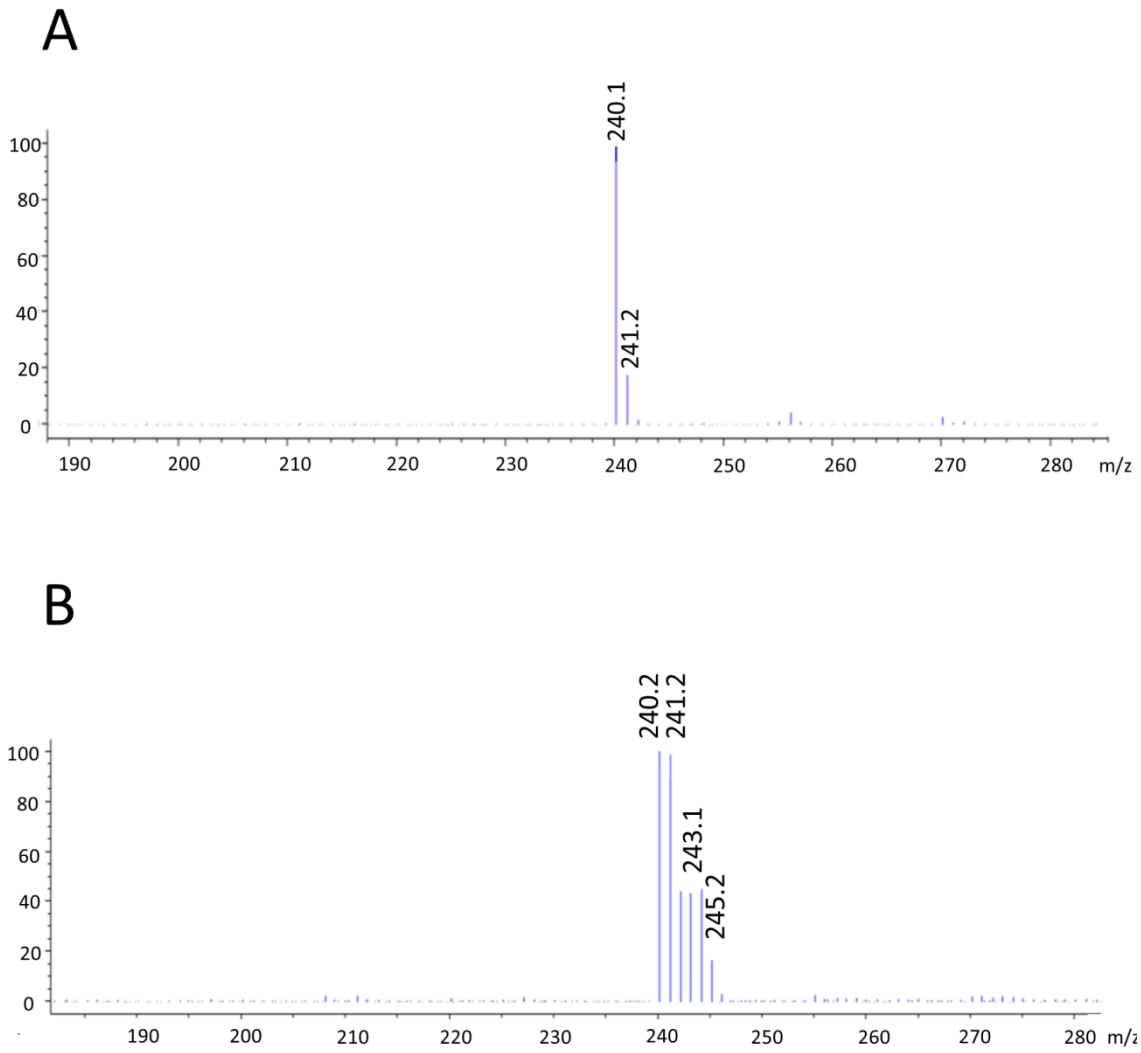


Figure 20: Isotopic pattern of massargiline obtained from *Massilia* sp. NR 4-1 cultures without (A) and with feeding of [1-13C]-acetate (B) in R2A medium without glucose and starch.

4.4 Testing of massargiline

The antimicrobial properties of massargiline were tested in a broth dilution assay. The tested strains included the Gram-positive *Mycobacterium vaccae* 10670, the multiresistant *Staphylococcus aureus* 134/94 (MRSA 134/94) and the vancomycin resistant *Enterococcus* 1528 (VRE 1528), the Gram-negative *Pseudomonas aeruginosa* K799/61 *and Escherichia coli* SG458, as well as the fungus *Sporobolomyces salmonicolor.* The determined MIC values are shown in **Table 9**.

Table 9: MIC values of massargiline

Test organism	MIC (mg/L)
Methicillin-resistant *Staphylococcus aureus* 134/94	100
Sporobolomyces salmonicolor	12.5
Escherichia coli SG458	50
Vancomycin-resistant *Enterococcus* 1528	100
Pseudomonas aeruginosa K799/61	25
Mycobacterium vaccae 10670	25

In comparison to known antibiotics [246], massargiline exhibits only moderate bioactivities. Approved antibiotics such as ciprofloxacin reach MIC values of 0.015 mg/L against *E. coli* NCTC 10418 and *E. coli* ATCC 25922 and 0.25 mg/L against *P. aeruginosa* NCTC 10662 and *P. aeruginosa* ATCC 27853.

Furthermore, antiprotozoal activities of massargiline were tested. (**Table 10**). The activities can be classified as "good" (IC_{50} < 20 µg/mL), "moderate" (IC_{50} 20-60 µg/mL), "weak" (IC_{50} 61-200 µg/mL) and "inactive" (IC_{50} > 200 µg/mL) [247]. Following this classification massargiline shows good activity against *Plasmodium falciparum* NF54, *Trypanosoma brucei rhodesiense* and *Leishmania donovani* axenic *amastigotes*. Moderate acitivty is exhibited against *Trypanosoma cruzi.* The cytotoxic effects of massargiline were assessed towards L6 cells. Approved antiprotozoal drugs like chloroquine reach IC_{50} values of 0.78 µg/mL against *Leishmania donovani* axenic amastigotes and 3.81 µg/mL against *Trypanosoma brucei rhodesiense* [248], [249].

Table 10: IC_{50} values of massargiline against protozoa and L6 cells

	IC_{50} (µg/mL)
Trypanosoma brucei rhodesiense	17.25
Trypanosoma cruzi	31.75
Leishmania donovani axenic amastigotes	17.20
Plasmodium falciparum NF54	2.87
L6 cells	28.10

4.5 Antismash analysis of *Acidovorax citrulli* AAC00-1

Strain AAC00-1 is a Gram-negative, mesophilic, aerobic and seed-borne bacterium belonging to the *A. citrulli* class II. Its genome has a size of 5.35 Mbp (NCBI Reference Sequence: NC_008752.1) with a GC content of 68.53%. The high GC content is in agreement with other *Acidovorax* strains [250]. Antismash analysis revealed the presence of 8 BGCs (**Table 11**) whereof three show low similarity to known clusters. Strain AAC00-1 shows a SM coding capacity of 0.28 Mbp or 5.2%, which is low compared to other β-proteobacteria [7].

Table 11 Results of the antiSMASH [23] analysis of *Acidovorax Citrulli* AAC00-1

Region	Type	From	To	Most similar known cluster	Similarity
Region 1	terpene	1,593,235	1,626,532		
Region 2	RiPP-like	3,370,980	3,381,777		
Region 3	NRPS, T1PKS	4,102,479	4,162,185	Crochelin A	42%
Region 4	homoserine lactone	4,222,995	4,243,630		
Region 5	phosphonate	4,290,828	4,331,700		
Region 6	NRPS-like	4,653,459	4,693,108		
Region 7	β-lactone	4,820,667	4,849,742	Fengycin	13%
Region 8	NRPS	5,173,956	5,218,305	Teixobactin	20%

BGC number 3 appeared to be a promising target for a genome mining study as the cluster involves multiple NPRS/PKS modules (**Table 12**). BGC 6 and 8 also suggest the involvement of NRPS, but they do not harbor genes for multimodular enzyme complexes. Both clusters harbor genes encoding standalone NRPS that are commonly involved in the generation of unusual building blocks [126] and, therefore, are not appropriate candidates for the search of siderophores.

BGC 3 consists of 16 genes of which 5 encode NRPS or PKS modules that allow to make a biosynthetic proposal of a putatively produced compound. The gene organization shows high similarity to other acyl peptide siderophore loci like the variochelins, taiwachelins or cupriachelin [56], [155], [251].

The FAAL domain of ABM34280 initiates the biosynthesis of the putative siderophore with the activation of typically dodecanonoic or tetradecanoic acid [155]. Aave_3733 also harbors the minimal NRPS set of the three domains C-A-PCP which activates and incorporates the first amino acid to the growing oligomer. The PKS Aave_3732 contains the minimal set of domains KS-AT-ACP and an additional KR domain putatively responsible for the formation of a hydroxy group. Four further amino acids are activated and condensed to the growing chain by the NRPS Aave_3731, Aave_3730 and Aave_3729. Aave_3729 and Aave_3731 harbor the minimal NRPS domain set of C-A-PCP, whereas Aave_3730 comprises two NRPS modules. Aditionally the first module habors an E domain which catalyzes a stereochemical inversion of the corresponding amino acid. The organization of the BGC and the molecular assembly line are depicted in **Figure 21**.

Furthermore, the gene cluster includes genes that encode for accessory proteins that are typical for NRPS clusters. Aave_3736 encodes a MbtH-like protein that putatively influences the amino acid activation [252]. Aave_3722 acts as a phosphopantetheinyl transferase [253]. The thioesterase Aave_3735 fulfills „housekeeping functions“ and removes aberrant intermediates blocking the assembly line [254]. Genes encoding transporters are a typical feature in lipopeptide BGCs [155]. Blast analysis identified Aave_3724 as a putative cyclic peptide transporter, which might indicate a macrocyclization of the NRPS/PKS product [255].

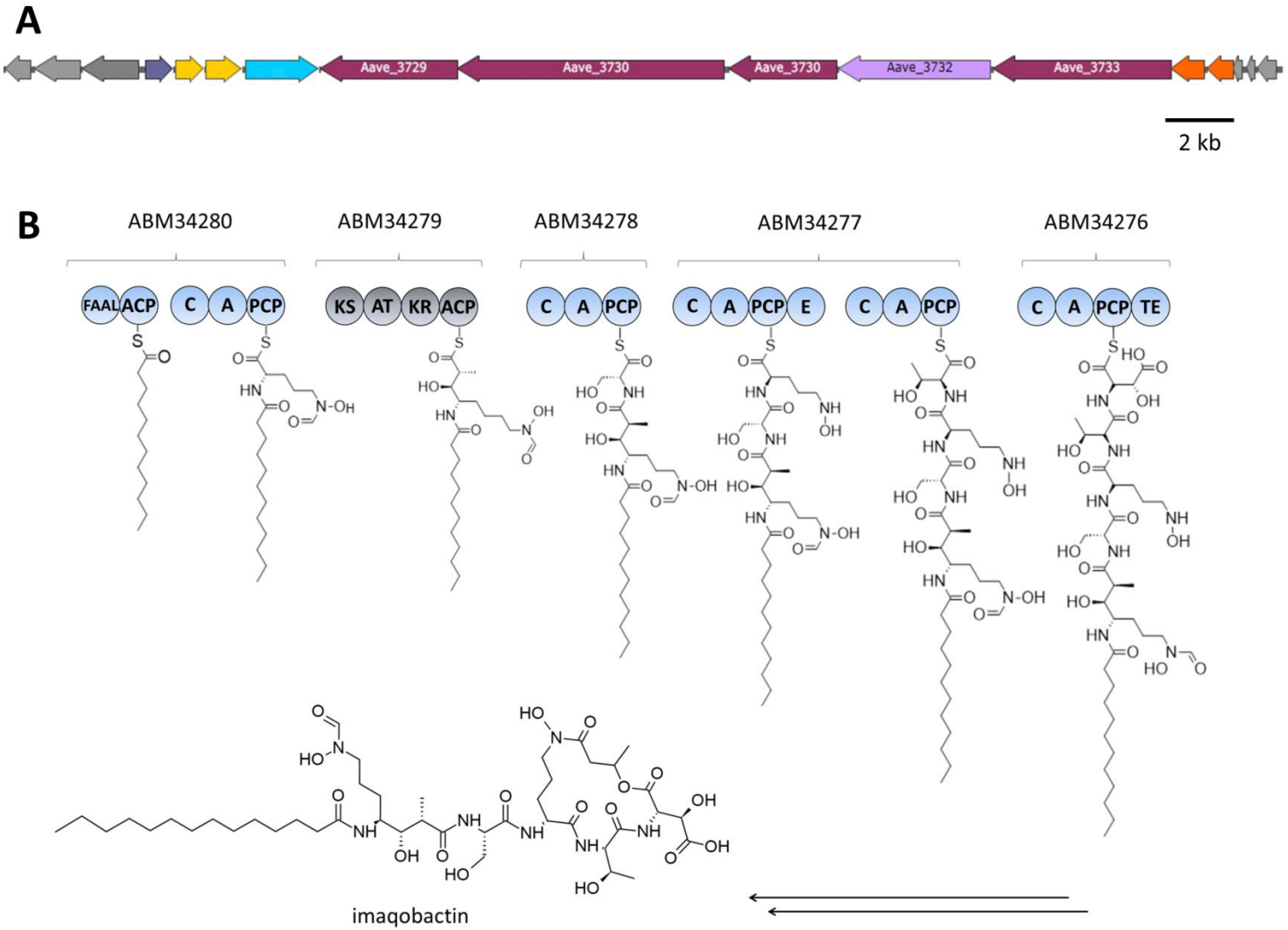


Figure 21: Organization of BGC 3 in **Table 11** (A). Molecular assembly line and proposed biosynthesis of imaqobactin

In addition, genes for putative tailoring enzymes were identified. Aave_3734 encodes a taurine dioxygenase-like enzyme, which might be involved in the β-hydroxylation of aspartate. This oxidation reaction occurs most likely after incorporation of the amino acid into the growing peptide chain [256]. The presence of β-hydroxyaspartate residues is a common motif in photoreactive acyl peptide siderophores [58], [155]. Aave_3723 belongs to a class of flavin-dependent monooxygenases involved in *N*-

Table 12: Annotation of gene cluster 3 from Acidovorax citrulli AAC00-1

Gene	Size (nt)	Protein accession number	Proposed function (domain architecture)
Aave_3722	774	ABM34269	4'-phosphopantetheinyl transferase
Aave_3723	1359	ABM34270	lysine N(6)-hydroxylase/L-ornithine N(5)-oxygenase family protein
Aave_3724	1695	ABM34271	cyclic peptide transporter
Aave_3725	831	ABM34272	ferric iron reductase protein FhuF
Aave_3726	942	ABM34273	formyl transferase
Aave_3727	1101	ABM34274	N-acetyltransferase
Aave_3728	2202	ABM34275	TonB-dependent siderophore receptor
Aave_3729	4107	ABM34276	NRPS (C-A-PCP)
Aave_3730	7956	ABM34277	NRPS (C-A-PCP-E C-A-PCP)
Aave_3731	3393	ABM34278	NRPS (C-A-PCP)
Aave_3732	4608	ABM34279	PKS (KS-AT-KR-ACP)
Aave_3733	5340	ABM34280	NRPS (FAAL-ACP-C-A-PCP)
Aave_3734	993	ABM34281	Taurine catabolism dioxygenase TauD/TfdA
Aave_3735	777	ABM34282	Thioesterase
Aave_3736	288	ABM34283	MbtH domain protein
Aave_3737	534	ABM34284	RNA polymerase, sigma-24 subunit

hydroxylation of ornithine and lysine [257]. Blast analysis of Aave_3726 showed close homology to a forymltransferase from *Acidovorax avenae*.

In order to deduce the amino acid building blocks of this NPRS assembly line, signature motifs of the corresponding A domains were analyzed. An analysis of Aave_3733 by the NRPS Predictor [130] revealed a high similarity to an A domain of anabaenopeptilide synthetase A [258] (6 of 8 identities, 75%) and coelichelin synthetase [259] (6 of 8 identities, 75%). The two enzymes are responsible for the incorporation of threonine and N^{δ}-formyl-N^{δ}-hydroxyornithine, respectively. A literature search revealed that the binding pocket motif of Aave_3733 is identical to A domains from *Variovorax paradoxus* B4 and *V. paradoxus* S110, which were predicted to incorporate N^{δ}-formyl-N^{δ}-hydroxyornithine [155].

Table 13: Binding pocket signatures and substrate predictions

Protein	binding motif	substrate prediction	source
ABM34280	D V W N I G L I H K	N^{δ}-formyl-N^{δ}-hydroxyornithine	[155], [259]
ABM34278	D V W H V S L I D K	serine	[260]
ABM34277	D G E G S G G V T K (A1)	5-hydroxy-ornithine	[128]
	D F W N I G M V H K (A2)	threonine	[259]
ABM34276	D L T K V G H V G K	aspartate	[261]

Aave_3731 shows 100% similarity to an A domain of pyoverdine synthetase which activates serine [260]. Aave_3730 harbors two A domains. The in silico prediction showed no hit for the first domain. A further comparison to databases [128] revealed closest homology to an A domain (protein Pa-367-47, module 4) that accepts 5-hydroxy-ornithine (5 of 8 similarities 62,5%). The second A domain possesses a signature that is 100% similar to the second module of CchH from the coelichelin synthetase which accepts threonine. Aave_3729 showed closest relation to the A domain of surfactin synthetase B (7 of 8 identities, 87%) which activates aspartate. The results of the substrate predicitions are summed up in **Table 13.**

A search for the proposed acyl peptide with the amino acid sequence shown in **Table 13** identified imaqobactin (**Figure 21**) as a potential product of BGC 3 in strain AAC00-1 [80].

This acyl peptide siderophore was isolated from the arctic bacterium *Variovorax s*p. RKJM285 and forms stable complexes with iron, gallium and aluminium. It shows structural similarity to the bioinformatics based prediction. Imaqobactin harbors a dodecanoic acid as well as a 4,7-diamino-3-hydroxy-2-methylheptanoic acid moiety that is similar to the partial structure of variochelin A [80]. Its biosynthetic origin is likely arising from a Claisen condensation of N^{δ}-formyl-N^{δ}-hydroxyornithine and methylmalonate. A following ketoreduction leads to the formation of the hydroxy group and is consistent with the domain architecture of the PKS. The incorporation of N^{δ}-formyl-N^{δ}-hydroxyornithine is in agreement with the amino acid predicition and the presence of the tailoring enzymes.

The following polypeptide partial structure consists of the amino acids L-serine, D-ornithine, L-threonine and L-erythro-*β*-hydroxyaspartate. The incorporation of these amino acids would meet the expectation due to the bioinformatic analysis including the stereochemical inversion of L-ornithine to D-ornithine by the E domain. Imaqobactin shows further structural modifications including an acetylation of ornithine with 3-hydroxybutanoic acid (hbu) which is known from other lipopeptide siderophores [56], [58]. Aave_3727 was identified as a potential *N*-acetyltransferase that catalyzes this reaction. Furthermore, imaqobactin shows the presence of a cyclic depsipeptide link between the 3-hydroxybutanoic acid and the *β*-hydroxyaspartate residue. Ring formations of acylated peptids are known to be carried out by TE domains [262]. The mechanism was studied in detail before [261].

4.6 Search for a putative siderophore in strain AAC00-1

In order to identify an imaqobactin-type siderophore in strain AAC00-1, the bacterium was cultured under different iron supplementation conditions. Iron free minimal H3 medium (200 mL) was supplemented with a $FeCl_3$ concentration of 0.1 mM or 1 mM. One experiment was carried out without the supplementation of iron. After 5 days of incubation the cultures were extracted three times with ethyl acetate, dissolved in 1.5 mL methanol and applied to the HPLC. The comparison of the chromatograms revealed a peak at a retention time of 53 min, of which the intensity decreases with increasing iron concentration (**Figure 22**).

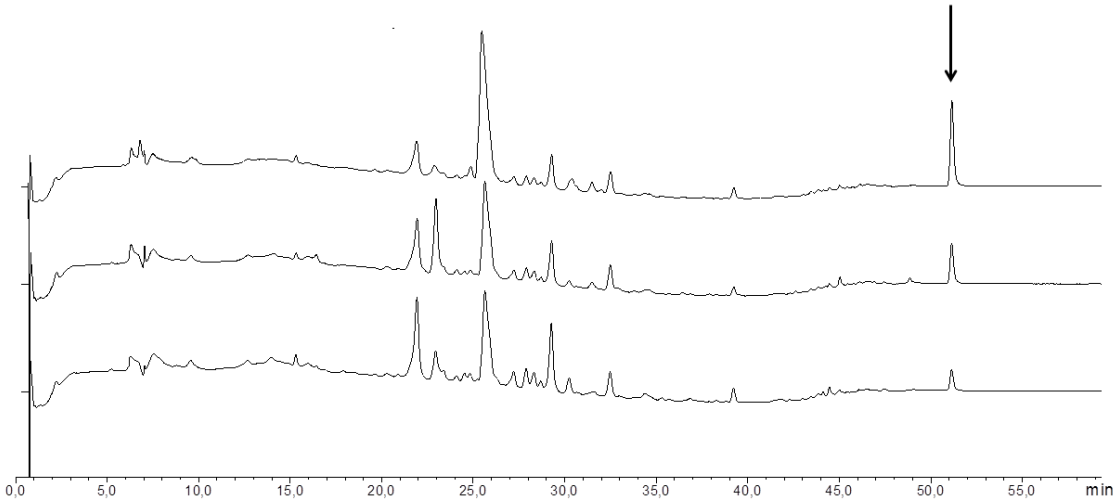


Figure 22: UV chromatograms of crude culture extracts from *A. citrulli* AAC00-1 supplemented with a $FeCl_3$ concentration of 0 mM (top), 0.1 mM (middle) and 1 mM (bottom). The extracts were measured at a wavelength of 250 nm. The isolated peak at a retention time of 52 min is marked with an arrow.

0.8 mg of the compound were isolated and applied to the MS. Imaqobactin possesses a *m/z* value of 934.5038 $[M+H]^+$ and its Fe(III)-complex exhibits a *m/z* value of 987.4168 $[M+H]^+$. The unknown compound revealed a *m/z* value of 447.3992 (**Figure A 35**) and is therefore not identical with imaqobactin. Kerr and coworkers reported that imaqobactin was produced only under a defined medium composition. *Variovorax* sp. RKJM285 is an arctic marine bacterium and therefore ISP2 complex medium supplemented with marine salts (ISP2m) was used for cultivation. In the absence of the marine salts, no siderophore production was observed [80].

A further MS-guided approach was carried out to evaluate the potential production of imaqobactin or a similar compound in strain AAC00-1. *A. citrulli* AAC00-1 was cultivated in three media including LB complex medium, ISP2m complex medium and H3 minimal medium supplemented with watermelon. As *A. citrulli strain* AAC00-1 is known to be a watermelon pathogen [179], the H3 medium supplemented with watermelon was chosen to set a potential chemical stimuli that might trigger the production of secondary metabolites.

It was taken into consideration that a derivative of imaqobactin might be produced. A derivative of imaqobactin that lacks the acylation with hbu and is not cyclized would show a calculated *m/z* value of 866.4717 $[M+H]^+$. A further difference in the expected molecular mass might occur be the acceptance of a different fatty acid starter unit by the FAAL-domain of ABM34280. The incorporation of tetradecanoic instead of

dodecanonoic would lead to a calculated *m/z* value of 962.5292 $[M+H]^+$ for the corresponding imaqobactin derivative.

Strain AAC00-1 was cultivated in 200 mL of each medium. After 5 days of incubation the cultures as well as the media controls were extracted three times with ethyl acetate. The samples were dissolved in 1.5 mL methanol and applied to the MS. The sample data were evaluated to find the *m/z* value of imaqobactin or a similar value that is not present in the media control. **Figure 23** shows the comparison of the sample cultivated in H3 medium supplemented with watermelon and the medium control.

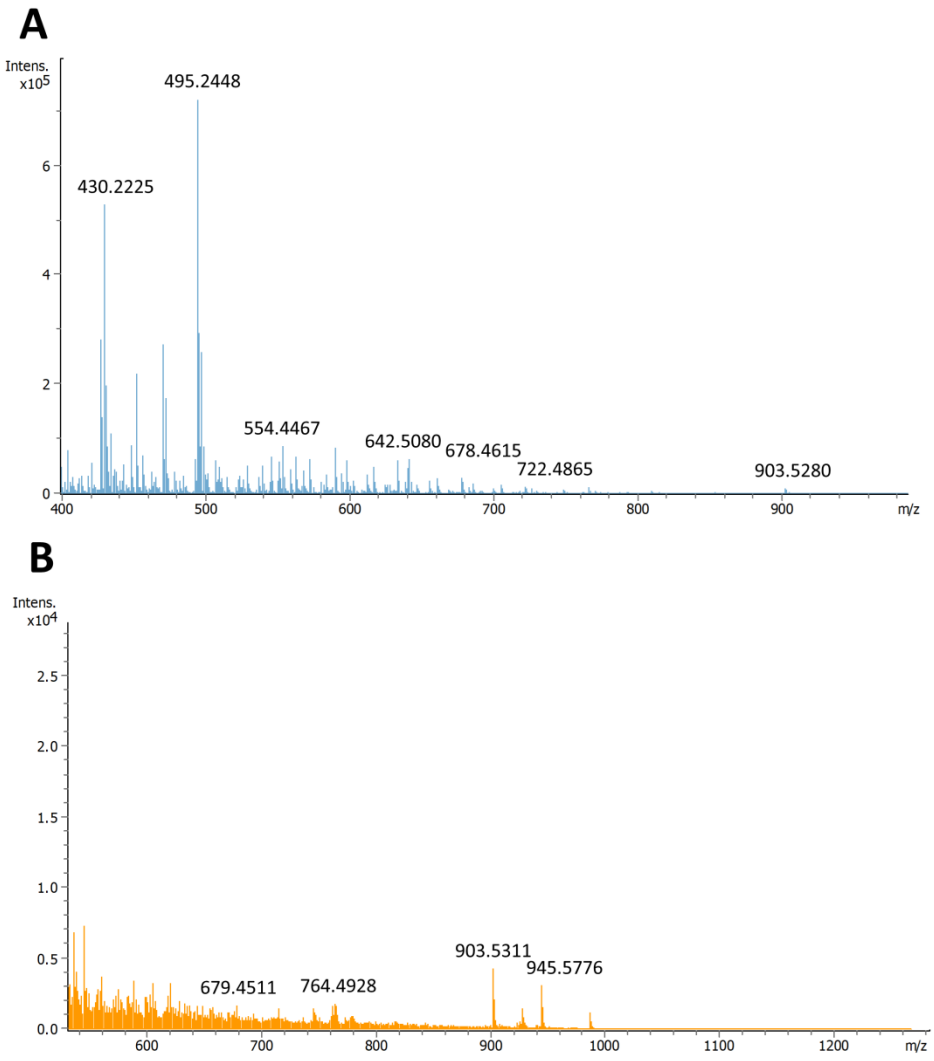


Figure 23: MS data of the strain AAC00-1 extract grown in H3 medium supplemented with watermelon (A) and the medium control (B).

The evaluation of the data did not reveal the mass of imaqobactin. The *m/z* value of 903.5280 that was found in the sample was also present in the medium control.

The evaluation of the MS data of the samples cultivated in LB medium, ISP2m as well as H3 minimal medium did not reveal a *m/z* value of 934.5038 $[M+H]^+$ or a signal in the range of possible derivatives of imaqobactin (**Figure A 36** - **Figure A 38**).

5 Discussion

5.1 Genome mining for biosynthetic gene clusters in *Massilia sp.* NR 4-1

For over a century natural product discovery from Gram-positive bacteria like actinomycetes led to important contributions in medicine [189]. Several milestones were set during the golden age of antibiotics [17]. However, in the last decades the number of natural products discovered by the pharmaceutical industry decreased [3]. With the beginning of the new millennium a paradigm shift occurred in natural product discovery driven by the increasing number of genome sequencing data. The idea of predicting novel secondary metabolites by analyzing genomic data instead of carrying out extensive screening approaches led to the recognition of afore neglected potential. This concept known as genome mining revealed a multitude of natural product BGCs, which exceed the number of discovered secondary metabolites in fermentation studies [7].

Next to the discovery of unknown gene clusters in well investigated strains it became evident that some bacterial taxa remained chemically underinvestigated despite their potential for the production of secondary metabolites. In recent years Gram-negative bacteria like myxobacteria or cyanobacteria were identified as rich sources and potent producers of natural products [62], [66]. Nowadays, they are known to produce several hundreds of antibiotics. Although Gram-negative bacteria have attracted increasing interest in natural product discovery, some taxa still bear undiscovered potential. For instance, the genus *Burkholderia* became only recently subject of extensive in silico genome-guided discovery approaches [4]. The identified gene clusters encode for compound classes like polyketides, nonribosomal peptides, lantipeptides and structurally unclassified molecules [61]. These promising findings raised the question of how large the unrevealed chemical potential of the phylum β-proteobacteria still is.

The genus *Massilia* is foremost known in a medical context with some members causing infections in immunocompromised patients [169]. However, in recent years the increasing numbers of identified strains showed that they are ubiquitous in the environment and settle in a variety of niches like soil, water, the rhizosphere of plants or even in harsh environments like glacier ice and desert soil [66], [161]–[166].

Despite the increasing number of sequenced *Massilia* genomes not much is known about the secondary metabolism of this genus.

In this PhD project the strain *Massilia* sp. NR 4-1 appeared as an excellent candidate for an intensive genome mining study as it provides all needed requirements. It combines a large 6.36 Mbp-sized genome with a secondary metabolite capacity of 0.73 Mbp or 11%. AntiSMASH analysis predicted a total of 15 BGCs. Some of these gene clusters possess high or at least moderate similarity to known gene clusters. Other BGCs with no related homologs bear a promising potential of finding structurally unknown compounds. In this study three natural products were isolated from *Massilia sp.* NR 4-1 and it therefore represents the first approach in which multiple natural products were isolated from a single strain of the genus *Massilia*.

5.2 The discovery of massiliachelin

One of the BGCs showed a high similarity to a locus from the plant pathogenic bacterium *Ralstonia solanacearum* GMI1000 which governs the biosynthesis of the siderophore micacocidin. Assuming that the corresponding molecule from *Massilia* sp. NR 4-1 also shows iron complexing properties the cultivation conditions were established. Siderophore biosynthesis can usually be triggered by choosing iron deficient cultivation conditions. Even before the era of genome mining this strategy led to the discovery of hundreds of siderophores. The first described approach of that kind was the discovery of 2,3-dihydroxybenzoylglycine in 1956 [99]. Although the analytical chemistry advanced this approach still allows an easy isolation procedure of siderophores compared to compounds for which cultivation conditions need to be established first. The discovery of massiliachelin showed the significance of this approach in combination with modern genome mining. A chemical prediction can guide the production and isolation process of the corresponding compound and also facilitate the structure determination.

Massiliachelin was elucidated as an epimer of the alkaloid agrochelin, which is known to have cytotoxic properties [207]. Both structures contain a thiazol-thiazolidine motif which is also present in yersiniabactin, pyochelin or micacocidin. These compounds are known to occur in nature as pair of diastereomers.

The interconversion of massiliachelin and agrochelin occurs rapidly in protic solvents. This was first observed for pyochelin. In chloroform, pyochelin I isomerizes rapidly to

a mixture of pyochelin II and pyochelin I in a ratio of ca. 3:1 [212]. During the structure elucidation of yersiniabactin, the compound was also described as unstable in nonacidic protic solvents. A second set of NMR signals was observed in the ratio of 2:3 assuming the presence of a diastereomer [209]. Further insights into the mechanism of intervonversion were provided by Schlegel et al. [213]. The authors stated that pyochelin I and II are only barely separable with chromatographic methods. The complexation with Fe^{3+} leads to a formation of complexes that are stable in DMSO, but not in methanol [213]. Therefore, NMR studies were carried out with a mixture of the two epimers due to the good distinction of data sets. The epimerization most likely occurs by an opening of the C-N-bond or the C-S-bond of the thiazolidine ring [213].

Complexation experiments with gallium were carried out. This metal is able to mimic iron because of the similar atom radius. Furthermore, it is redox-insensitive leading to the formation of strong complexes [214]. Massiliachelin was complexed with Ga^{3+} according to a protocol which showed appropriate results for micacocidin [206]. Although it was demonstrated that massiliachelin forms gallium complexes, NMR data revealed the presence of uncomplexed compound. Except for analytical purposes, siderophore-gallium complexes also bear potential in diagnostic and therapeutic applications [263]. Once the siderophore-gallium complex is taken up by a pathogen the radioactive isotope ^{68}Ga can act as a nuclear imaging tracer. A successful study was carried out investigating *Aspergillus fumigatus* causing invasive pulmonary aspergillosis [264]. The therapeutic application is based on two effects: On the one hand, gallium forms a firm complex with the siderophore and therefore reduces the concentration of uncomplexed siderophore for iron complexation [216]. The second application is related to the trojan horse strategy. After cellular uptake, gallium can mimic iron and replace it as a ligand in enzyme complexes. Due to its inability to react in redox reactions, it renders the gallium-loaded enzyme inactive [215], [216], [265].

It was further demonstrated that NRPS, non-iterative type I PKS and/or hybrids of these enzymes are ideal targets in genome mining studies, as they follow the principle of colinearity and therefore allow good prediction of the resulting chemical structure. As an outcome of this work it was shown that the structure of massiliachelin is consistent with the architecture of a biosynthetic assembly line,

which is encoded in the genome of *Massilia* sp. NR 4-1. Furthermore, it was possible not only to predict the chemical constitution of massiliachelin, but also to determine the configurational assignment. A comparison with other thiazol-thiazolidine confirmed the usefulness of bioinformatic analyses for the stereochemical analysis of this class of natural products.

5.3 The isolation of violacein

Violacein is a blue-purple pigment that is produced by Gram-negative bacteria like *Chromobacterium violaceum* or *Janthinobacterium lividum* [266]. A list of violacein producers with the corresponding yields was published before [267]. This pigment shows antibacterial properties especially against Gram-positive bacteria and significant biological activity against protozoans and metazoans [267]–[269]. Furthermore, violacein shows anticancer properties with IC_{50} values in the submicromolar concentration range [220].

Violacein consists of three partial structures: 2-pyrrolidone, 5-hydroxyindole and oxindole. Due to the two indole rings, violacein is referred to as bisindole. All carbon, nitrogen and hydrogen atoms derive from L-tryptophan whereas the oxygen atoms originate from molecular oxygen [270]. Violacein is produced in a five step enzymatic pathway by the enzymes VioA-VioE.

In this study, the violacein production was investigated in strain NR 4-1. The cultivation in minimal medium with the supplementation of tryptophan and histidine resulted in lower titers (9.3 mg/L) compared to strain BS-1 which was the first *Massilia* strain subjected to a violacein production study [170]. The production of violacein was analysed gravimetrically. In a study investigating the production of violacein in *Escherichia coli* the authors also evaluated the deoxyviolacein and violacein analytics. This study showed that photometric measurements lead to a strong overestimation of violacein (up to 680%) and deoxyviolacein (up to 320%) [218]. Potent producers of violacein like psychotropic bacterium RT102 can reach titers of 3.5 g/L [271]. Lower titers of 2 mg/L are reached by strains like *Alteromonas luteoviolacea* or *C. violaceum* ATCC 553 [272], [273].

Mixing the cells with methanol and centrifugation led to exhaustive extraction of the pigment. Isolation with HPLC allowed the separation of violacein and its byproducts. The structures were solved by NMR spectroscopy or mass spectroscopy.

5.4 The discovery of massargiline

Massargiline was isolated from the ethyl acetate extract of *Massilia* sp. NR 4-1. Its structure was elucidated by ^{1}H, ^{13}C and ^{15}N NMR spectroscopy. ^{15}N NMR is a potent application to further confirm structures of nitrogen containing molecules. Fused oxoheterocycles, which contain the guanidine group, have been successfully distinguished before via ^{15}N NMR spectroscopy [222]. Another example of the application of this powerful method is the analysis of tautomers of disubstituted guanidines and related structures [274].

The biosynthetic proposal suggests the involvement of six enzymes, including a type III PKS, an aminotransferase, an ornithine cyclodeaminase, a peptidase, a reductase and a dehydratase. Since a polyketidic origin was assumed, feeding experiments with 1-^{13}C-acetate were conducted. The mass spectrum showed a characteristic mass shift of 5 Da and therefore indicates the involvement of a type III PKS in the biosynthesis. The aminotransferase belongs to the α-oxamine synthase (AOS) family and catalyzes a C-C-bond formation followed by a decarboxylation.

In the pathway of massargiline a putative cyclodeaminase catalyzes the deaminating cyclisation of arginine or a derivative. The mechanism of this reaction is quite remarkable. Catabolic reactions from arginine to ornithine catalyzed by arginases (for example in *Bacillus subtilis* [275], [276]) or arginine amidinotransferases are well understood and extensively studied [277]. Ornithine is further catabolized to glutamate or proline in the bacterial metabolism. The latter reaction is catalyzed by ornithine cyclodeaminases [278]. These enzymes are known to catalyze the formation of proline from ornithine by releasing ammonium [238], [239]. Costilow et al. successfully isolated an ornithine cyclodeaminase from *Clostridium* strains in 1971 [279]. In 2004, the crystal structure was obtained and the mechanism of action was clarified [280]. AgrE from the cyanobacterium *Anabaena* is a recently identified, bifunctional ornithine cyclodeaminase, that is capable of catalyzing the reactions of arginine to ornithine with subsequent deaminating cyclisation [278]. However, there is no enzyme known yet that accepts arginine or arginine derivatives for direct deamination and subsequent cyclization.

Interestingly, the blastp analysis of the ornithine cyclodeaminase revealed closest identity to a gene (WP_130025750.1) of the tetracycline degrading bacterium *Pandoraea* sp. XY-2 sequenced in 2019 [281]. The corresponding gene cluster also

includes a type III PKS and a PLP-dependent amino-acid C_{α}-acyltransferase. Furthermore, it bears an asparaginase like gene, also suggesting the involvement of a nitrogen containing amino acid in the putative biosynthesis of the BGC product.

Bioactivity testing showed that massargiline possesses moderate antimicrobial and antiprotozoal activities. However, the biological function of massargiline remains unknown for now.

5.5 Genome mining for biosynthetic gene clusters in *A. citrulli* AAC00-1

The genus *Acidovorax* belongs to the β-proteobacteria and can be divided into two distinct classes. Strain *A. citrulli* AAC00-1 belongs to class two, whose representatives are associated with watermelons. Due to their ability to cause BFB on watermelons they play an important agricultural and economic role. Despite its importance not much is known about the secondary metabolism of *Acidovorax citrulli* strains.

A genome mining approach identified a lipopetide encoding BGC in strain AAC00-1. It includes 5 genes encoding 4 NRPS and 1 PKS as well as several accessory proteins and tailoring enzymes. A closer inspection of the domain architecture of the core enzymes and the substrate specitifity of the NRPS adenylation domains identified imaqobactin as potential natural product of the BGC.

This acyl peptide is produced by the marine arctic marine bacterium *Variovorax* sp. RKJM285. It was shown before that marine-like siderophores like the variochelins [155], the serobactins [282] and taiwachelin [251] are produced by plant-associated bacteria. Genome mining revealed that *A. citrulli* AAC00-1 harbors the genetic potential to produce lipopeptide siderophores.

In order to isolate a corresponding natural compound cultivation experiments under different iron concentrations in minimal H3 medium were carried out. A similiar approach using iron-deficient conditions in H3 minimal medium led to the discovery of taiwachelin [251]. However, the analytical evaluation of the extract did not reveal the production of a natural compound that could be linked to the identified BGC.

In a further MS-guided approach different media (ISP2m, LB medium, H3 supplemented with watermelon and H3 medium) compositions were tested. This strategy was inspired by the discovery of imaqobactin where Robertson et al. isolated

the natural product in complex ISP2m medium by identifing two coeluting ions with a difference in the *m/z* values that is consistent with the complexation of Fe(III). It was further observed that siderophore production is only initiated with marine salt supplementation [80].

As *A. citrulli* AAC00-1 is known to grow in complex LB medium [186], it was tested along with a minimal H3 medium that is supplemented with watermelon and without supplementation. The evaluation of the MS data did not reveal a mass that is in accordance with the bioinformatics based prediction.

Although the production of a siderophore was not observed it was demonstrated before that a *furA* deletion mutant strain AAC00-1 is capable of producing an iron complexing compound in a higher concentration. This result was quantified by measuring the halo diameter on chrome azurol S plates [186].

The molecular host–pathogen interactions of *A. citrulli* AAC00-1 are not completely understood yet. It was shown that T3SS genes as well as the *furA* gene are crucial for the virulence of this strain. The elucidation of a putative siderophore might, hence, improve the understanding of the virulence of strain AAC00-1 and the mechanisms of BFB in general.

6 Summary

In this thesis the strains *Massilia* sp. NR 4-1 and *Acidovorax citrulli* AAC00-1 were subjected to a bioinformatics analysis (genome mining) to evaluate their potential for natural product biosynthesis.

Two prediction guided as well as an unguided investigation were carried out. The structure of massiliachelin was elucidated by the application of mass spectrometry and NMR spectroscopy. Since a function as siderophore was predicted, *Massilia* sp. NR 4-1 was grown under iron-deficient conditions to induce the production of this iron chelator. A combined approach of spectroscopic and bioinformatic analyses allowed the determination of the absolute configuration of massiliachelin.

Bioinformatic analysis identified the complete violacein pathway in strain NR 4-1. Cultivation in minimal medium with the supplementation of tryptophan and histidine led to the isolation of the natural product via HPLC. The structures of violacein and its byproducts oxyviolacein and deoxyviolacein were solved by NMR spectroscopy and mass spectroscopy.

Massargiline is a novel natural product that was isolated from *Massilia* sp. NR 4-1. Its structure was elucidated by ^{1}H, ^{13}C and ^{15}N NMR spectroscopy. Based on the structure a putative gene cluster was identified by bioinformatic analysis and a biosynthetic proposal was made. Feeding studies using isotopically labeled precursors were carried out in order to support the proposal. The kinetics of massargiline production and its bioactivity were determined.

A further prediction guided genome mining approach was directed to the watermelon pathogen *Acidovorax citrulli* AAC00-1. The siderophore imaqobactin was identified as putative secondary metabolite. Cultivation experiments did not give any evidence for the production of this natural product.

This work is the first example of an extensive genome mining approach of a bacterium from the genus *Massilia*. It was demonstrated that *Massilia* sp. NR 4-1 bears a high potential of secondary metabolite capacity which led to the isolation of three natural compounds.

Zusammenfassung

In dieser Arbeit wurden die Stämme *Massilia* sp. NR 4-1 und *Acidovorax citrulli* AAC00-1 einer bioinformatischen Analyse unterzogen (genome mining), um ihr Potential zur Naturstoffbiosynthese zu untersuchen.

Zwei strukturvorhersage-basierte sowie eine ungerichtete Untersuchung wurden durchgeführt. Die Struktur des Massiliachelins wurde mittels Massenspektrometrie und Kernresonanzspektroskopie aufgeklärt. Da eine Funktion als Siderophor vorhergesagt wurde, wurde *Massilia* sp. NR 4-1 unter Eisenmangelbedingungen angezogen, um die Produktion dieses Eisenkomplexbildners zu induzieren. Eine Kombination aus spektroskopischen und bioinformatischen Analysen erlaubte die Bestimmung der absoluten Konfiguration von Massiliachelin.

Die bioinformatische Analyse bestätigte das Vorhandensein des vollständigen Violacein-Stoffwechselweges in Stamm NR 4-1. Eine Kultivierung in Minimalmedium unter Zugabe von Tryptophan und Histidin führte zur Isolierung des Naturstoffs mittels HPLC. Die Struktur des Violaceins, sowie die Struktur der Nebenprodukte Deoxyviolacein und Oxyviolacein wurden via Massenspektrometrie und Kernresonanzspektroskopie aufgeklärt.

Massargiline ist ein neuartiger Naturstoff der ebenfalls aus *Massilia* sp. NR 4-1 isoliert wurde. Die Struktur Massargilines wurde mittels ^{1}H, ^{13}C and ^{15}N NMR-Spektroskopie aufgeklärt. Basierend auf der Struktur wurde ein Gencluster identifiziert, das zur Vorhersage eines möglichen Biosyntheseweges diente. Fütterungsstudien mit Isotopen-markierten Präkursoren unterstützen diese Vorhersage. Zudem wurde die Kinetik der Massargiline- Produktion und die Bioaktivität dieses Naturstoffs bestimmt.

Eine weitere vorhersage-basierte Genome mining-Studie behandelte das Wassermelonenpathogen *Acidovorax citrulli* AAC00-1. Das Siderophor Imaqobactin wurde als potentieller Sekundärmetabolit identifiziert. Kultivierungsexperimente führten nicht zum Nachweis der Produktion dieses Naturstoffs.

In dieser Arbeit wird zum ersten Mal eine umfangreiche bioinformatische Analyse eines Bakteriums der Gattung Massilia durchgeführt. Es wurde gezeigt, dass *Massilia* sp. NR 4-1 ein hohes Potential zu der Bildung von Sekundärmetaboliten besitzt, welches zu der Isolierung dreier Naturstoffe führte.

7 Appendix

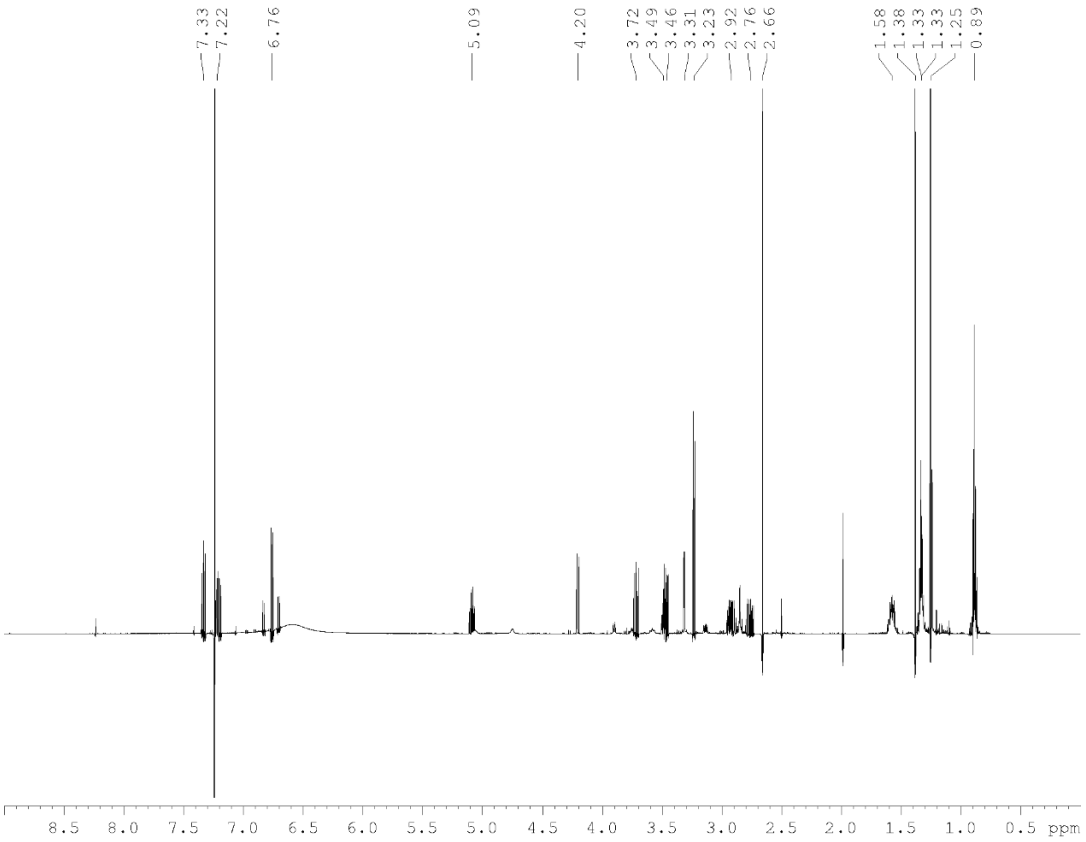


Figure A 1: ^{1}H NMR (600 MHz) spectrum of massiliachelin in chloroform-d

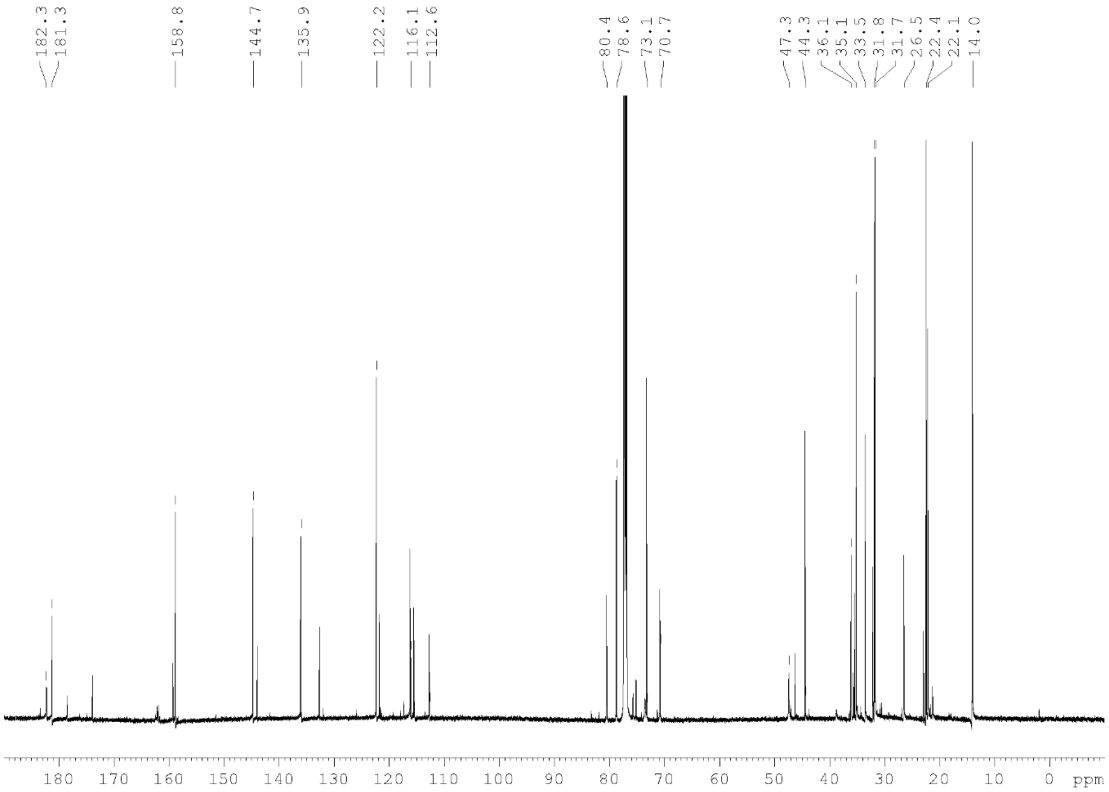


Figure A 2: 1H-decoupled ^{13}C NMR (600MHz) spectrum of massiliachelin in chloroform-d

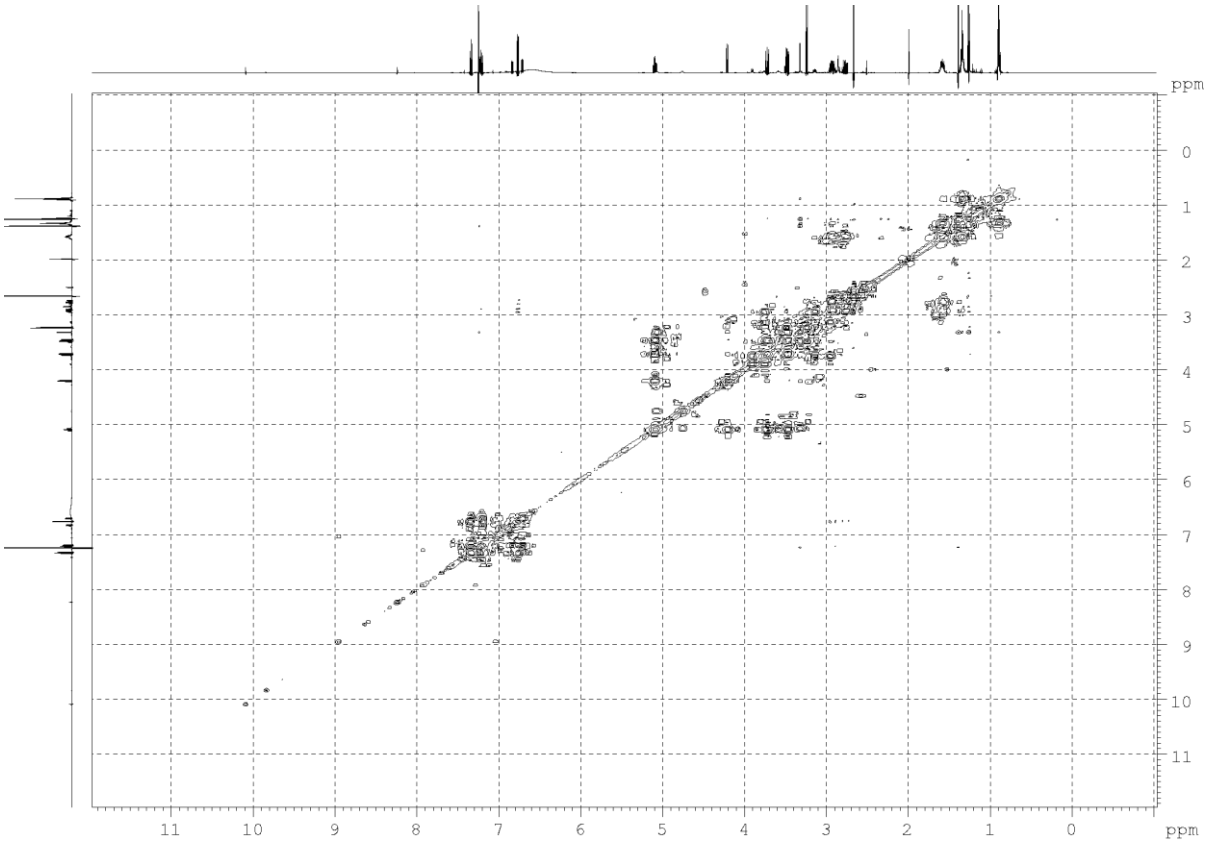


Figure A 3: ^{1}H, ^{1}H COSY spectrum (600MHz) of massiliachelin in chloroform-d

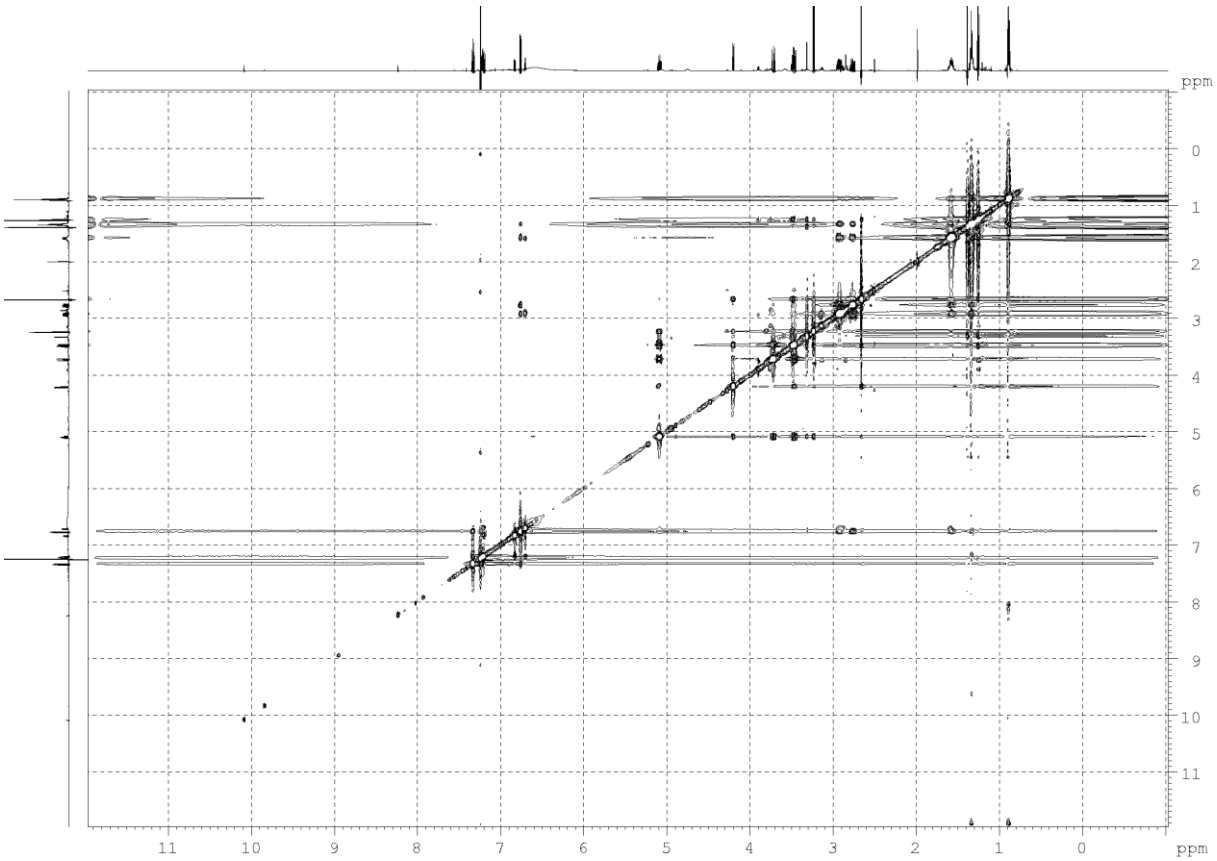


Figure A 4: ^{1}H, ^{1}H NOESY spectrum (600MHz) of massiliachelin in chloroform-d

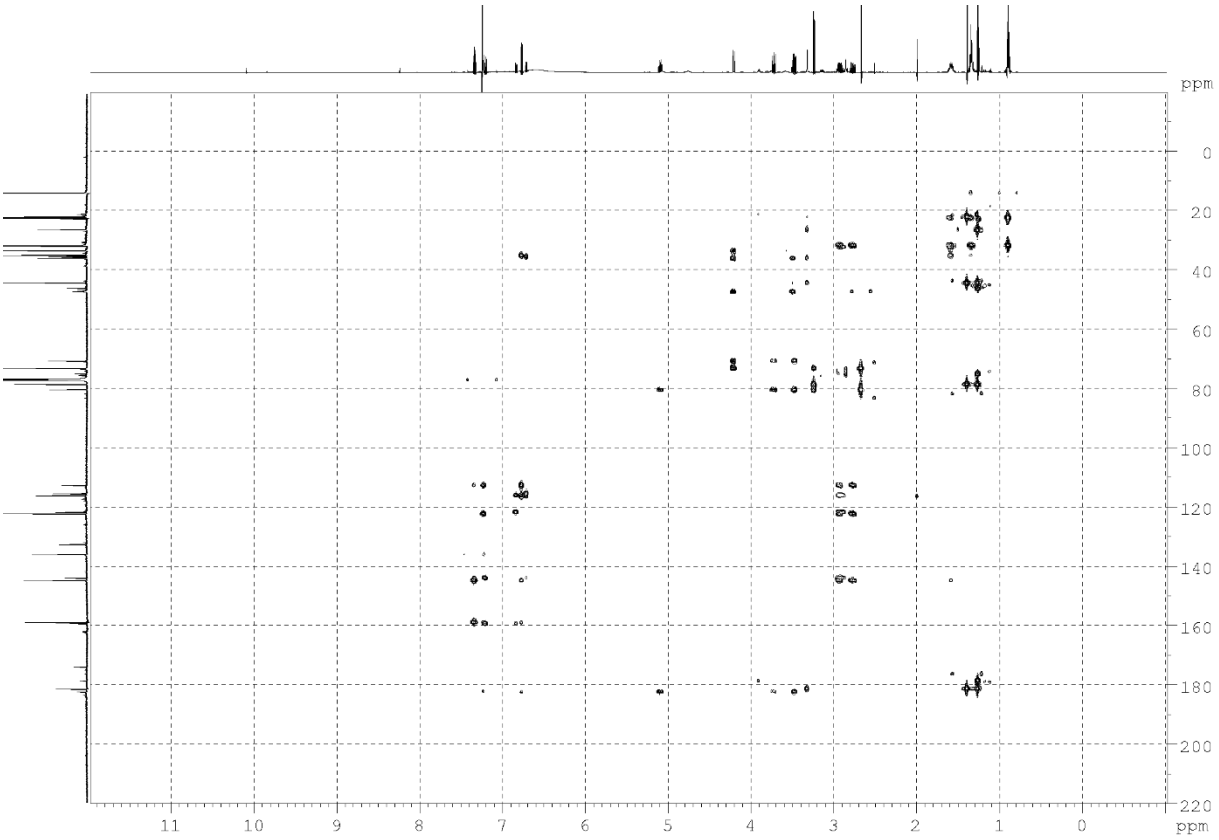


Figure A 5: ^{1}H, ^{13}C HMBC spectrum (600MHz) of massiliachelin in chloroform-d

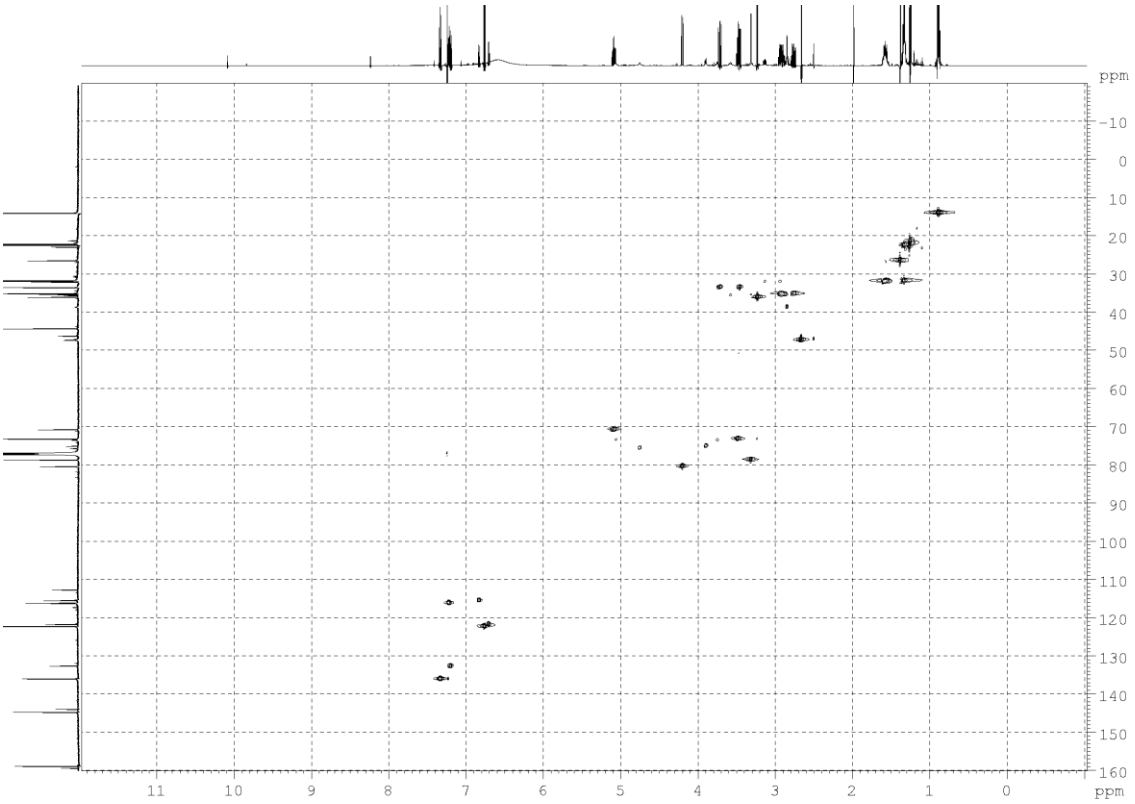


Figure A 6: ^{1}H, ^{13}C HSQC spectrum (600MHz) of massiliachelin in chloroform-d

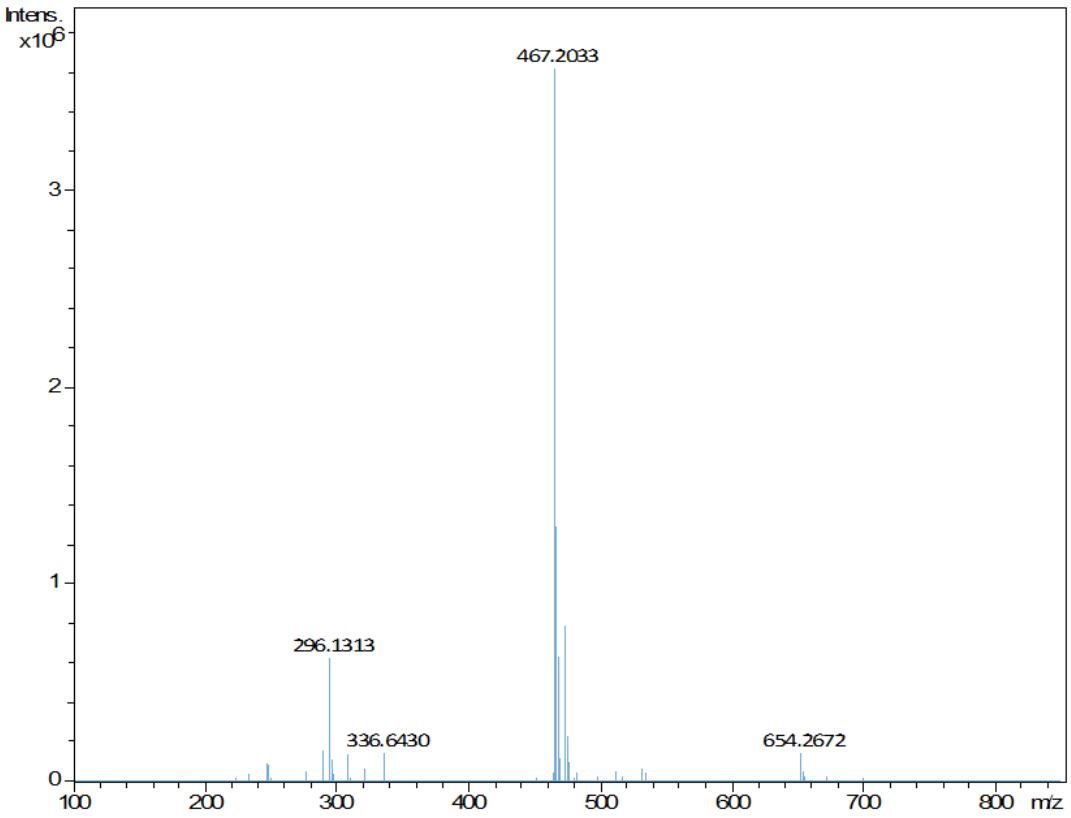


Figure A 7: HR-ESI-MS spectrum of massiliachelin

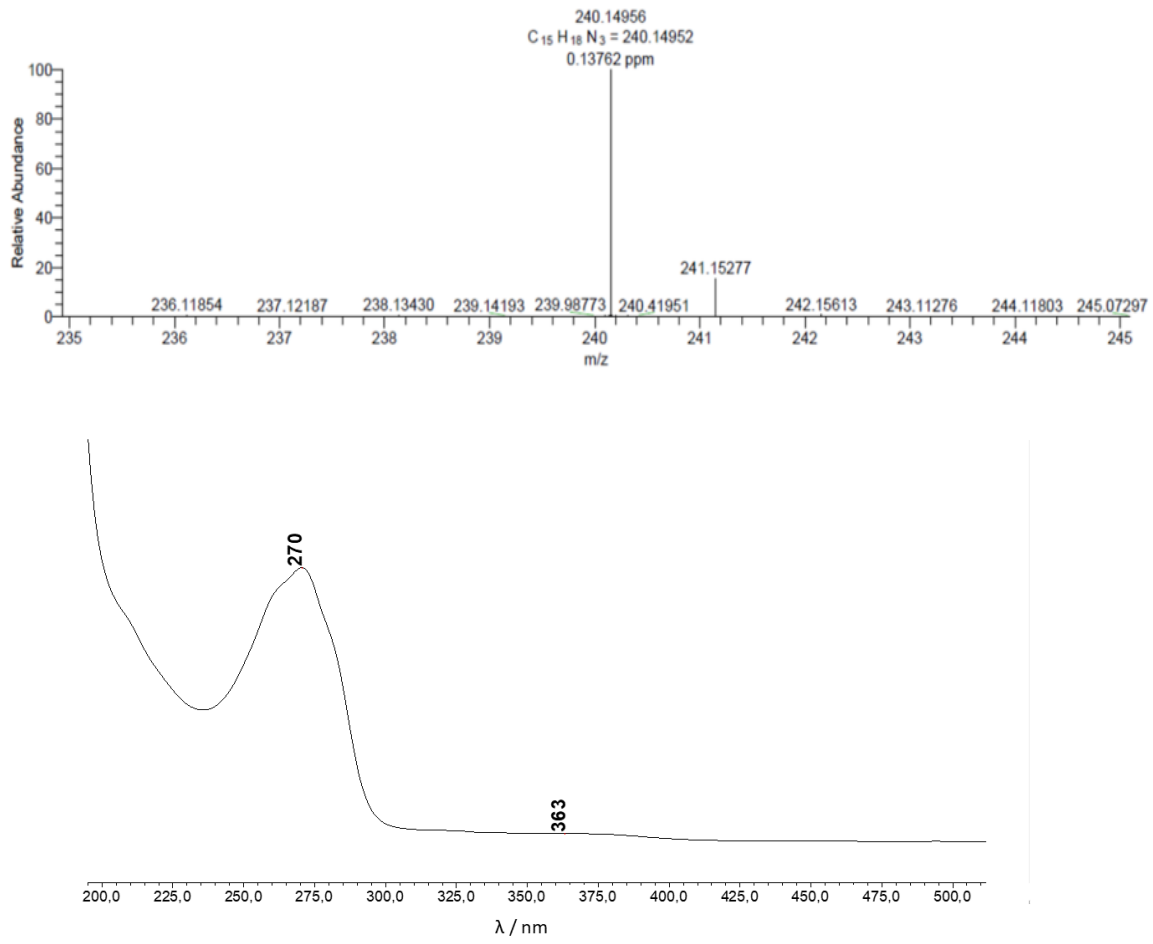


Figure A 8: HR-ESI-MS data (top) and UV-Profile of massargiline in methanol (bottom).

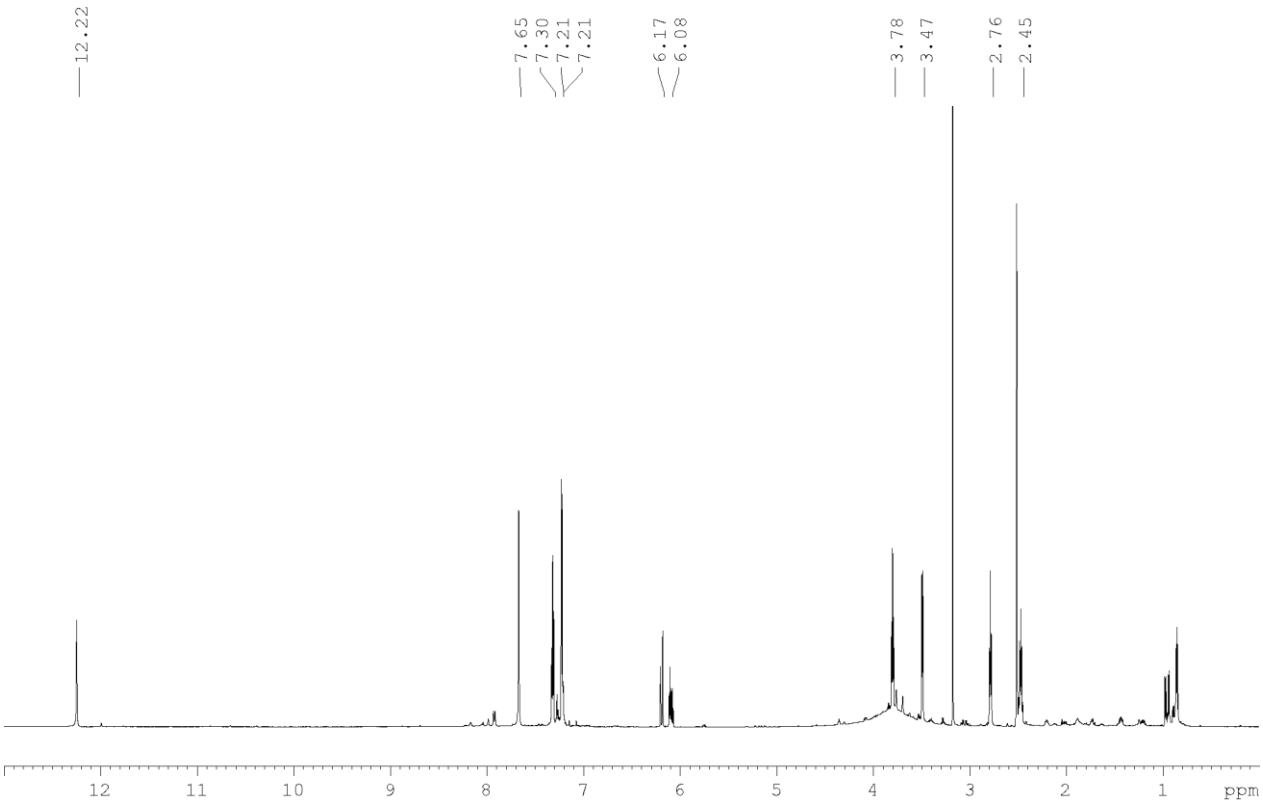


Figure A 9: ^{1}H NMR (600MHz) spectrum of massargiline in DMSO-d_6

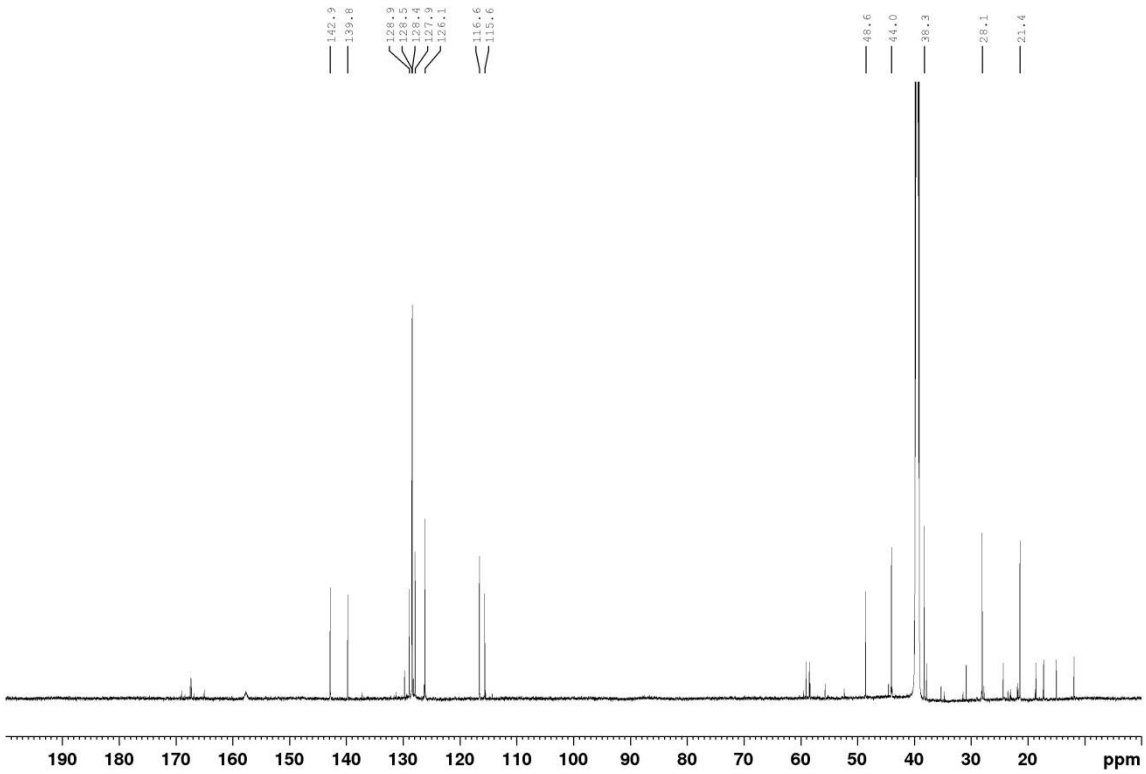


Figure A 10: ^{13}C NMR (600MHz) spectrum of massargiline in DMSO-d_6

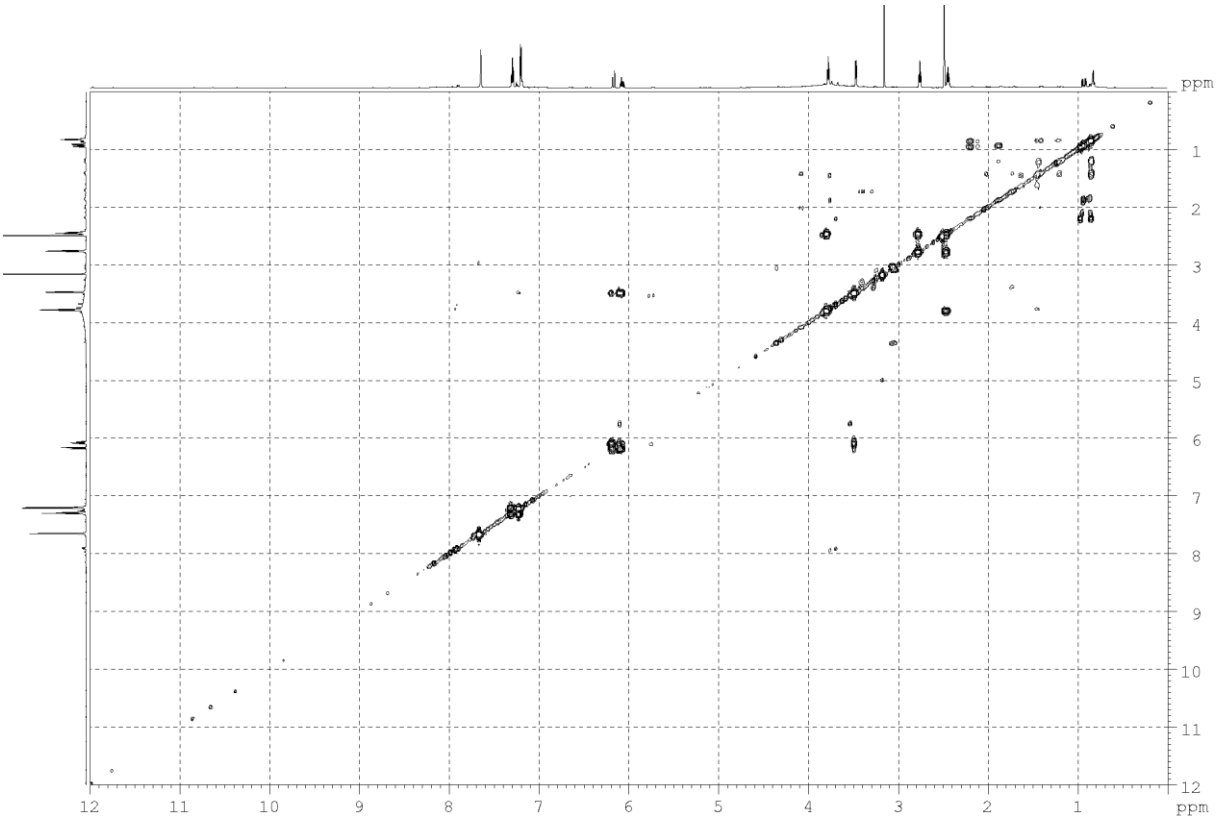


Figure A: 11: ^{1}H, ^{1}H COSY spectrum of massargiline in DMSO-d_6

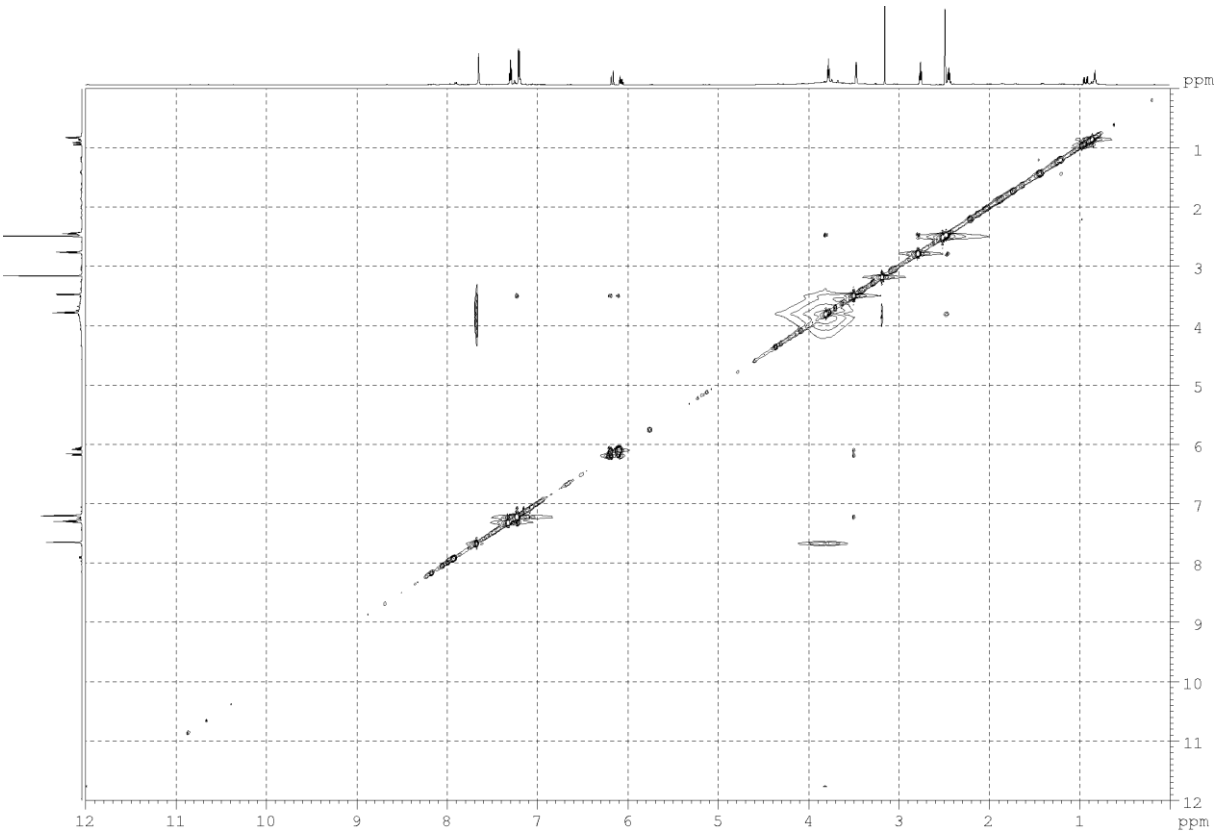


Figure A 12: ^{1}H, ^{1}H NOESY spectrum of massargiline in DMSO-d_6

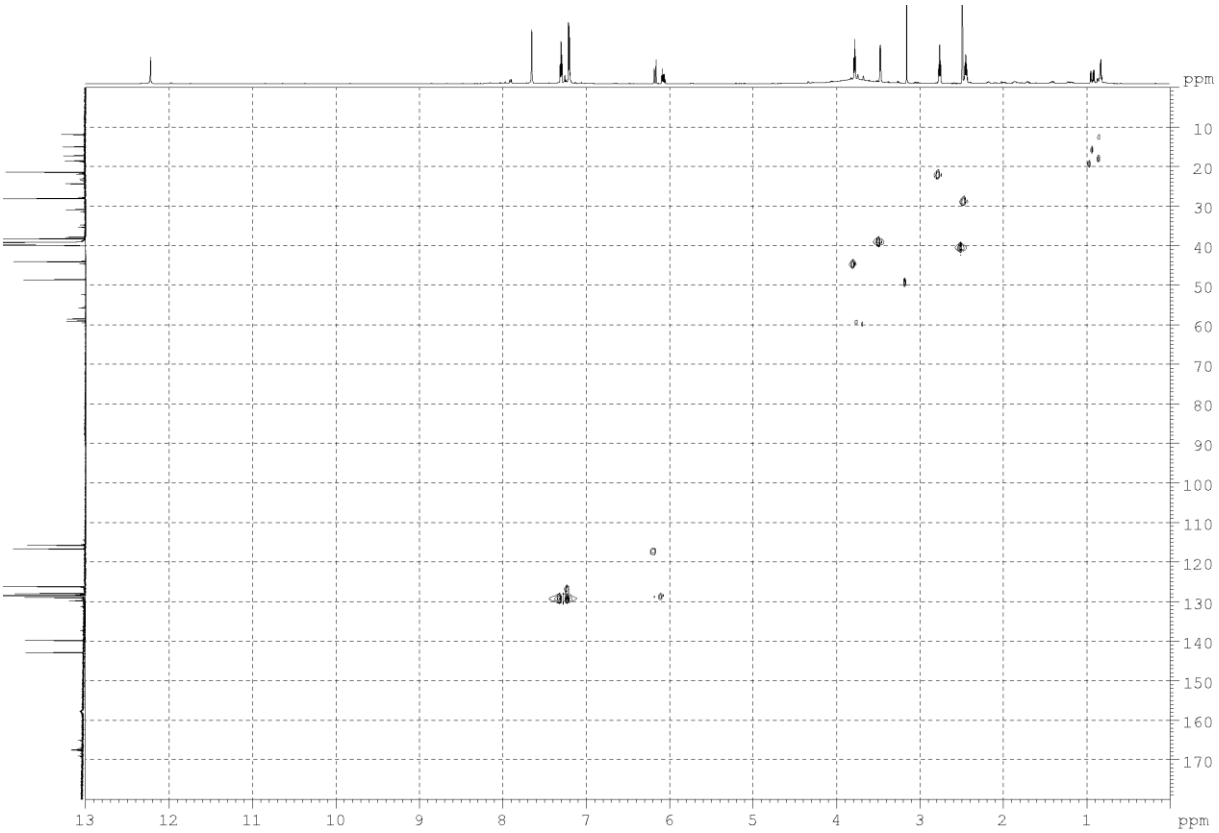


Figure A 13: ^{1}H, ^{1}C HSQC spectrum of massargiline in DMSO-d_6

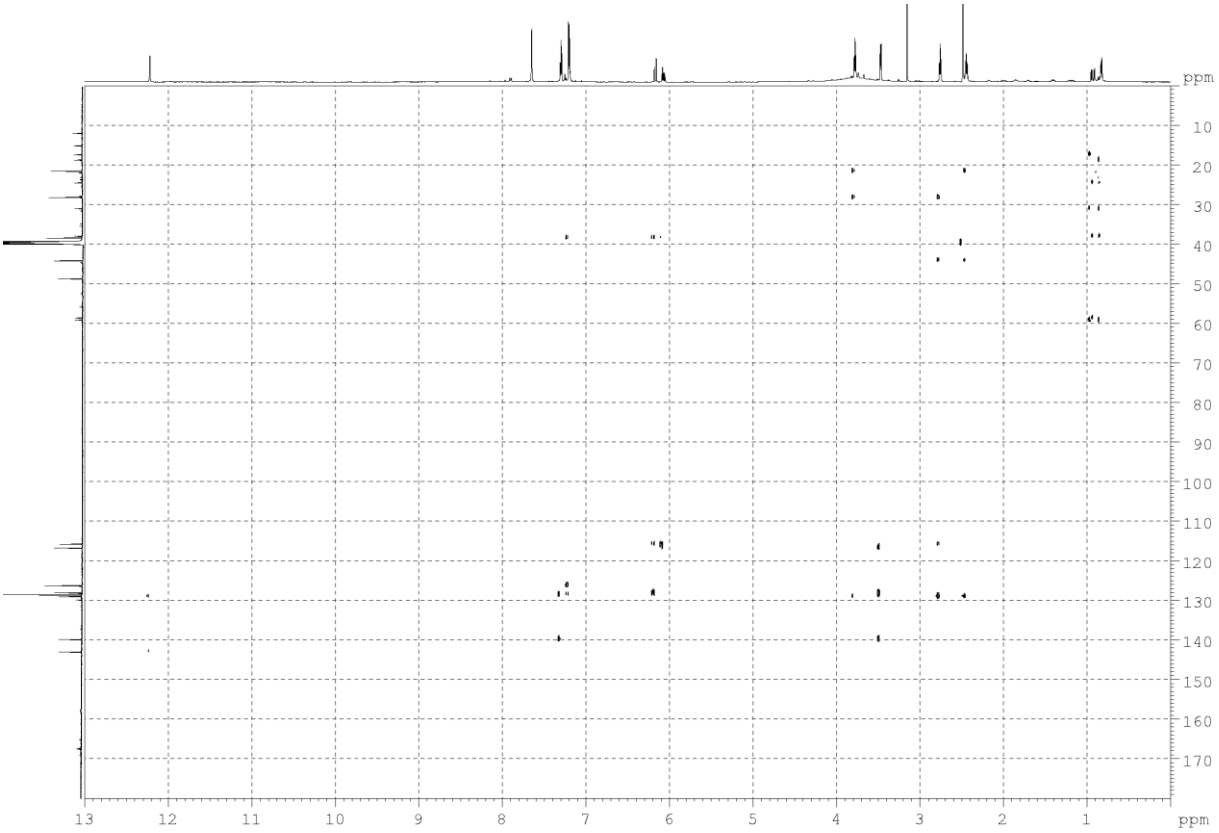


Figure A 14: ^{1}H, ^{13}C HMBC spectrum of massargiline in DMSO-d_6

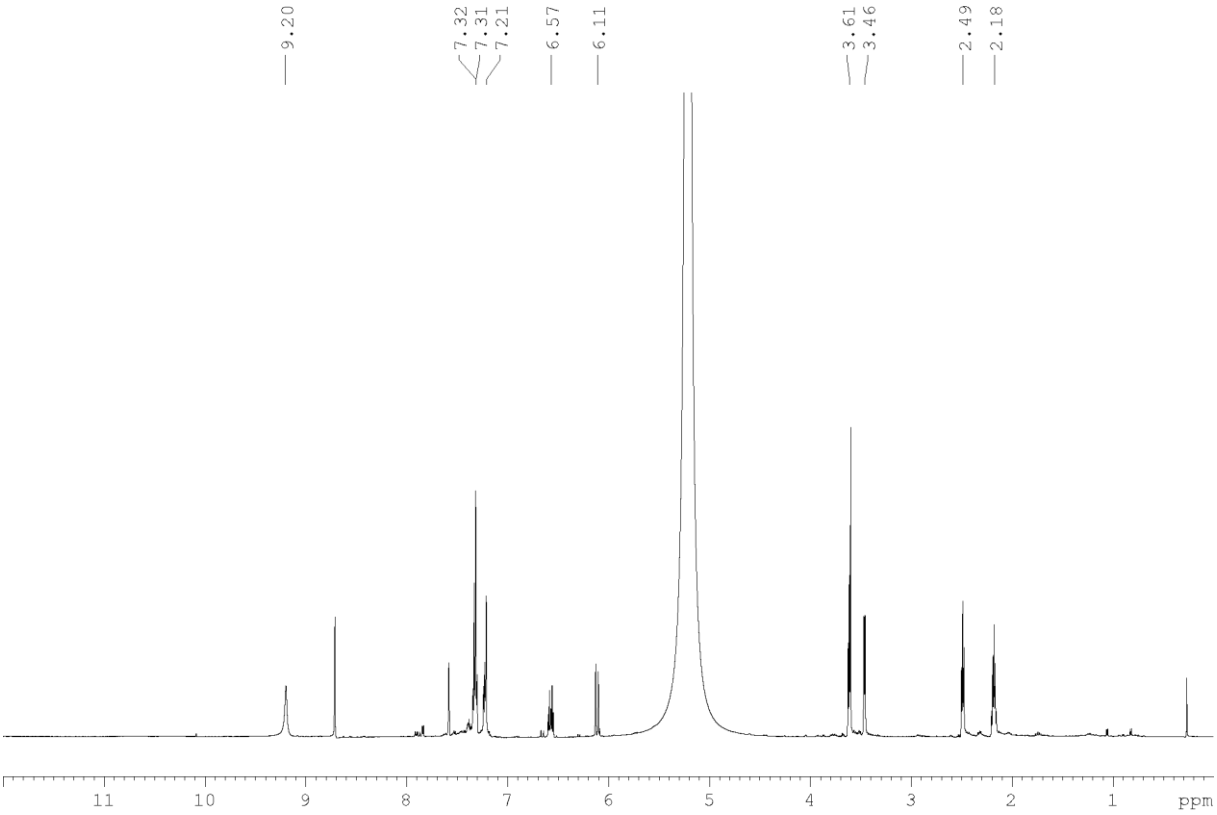


Figure A 15: ^{1}H NMR (600MHz) spectrum of massargiline in pyridine-d_5

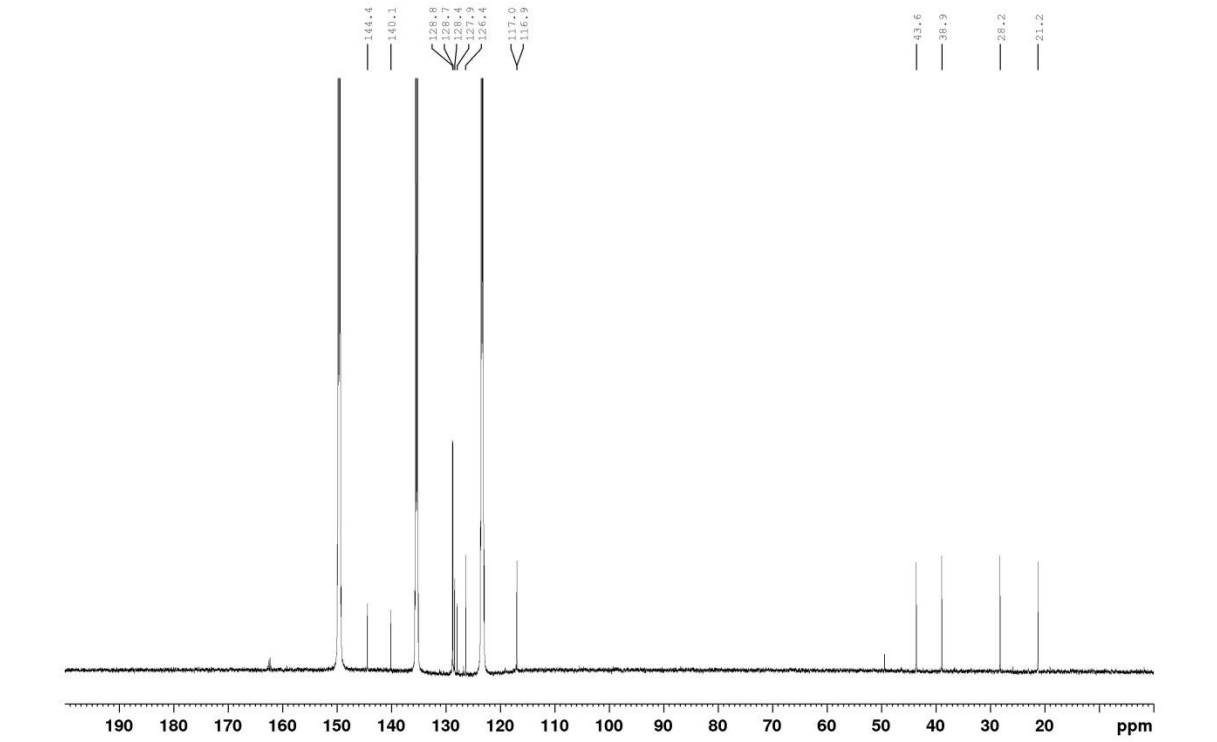


Figure A 16: ^{13}C NMR (600MHz) spectrum of massargiline in pyridine-d_5

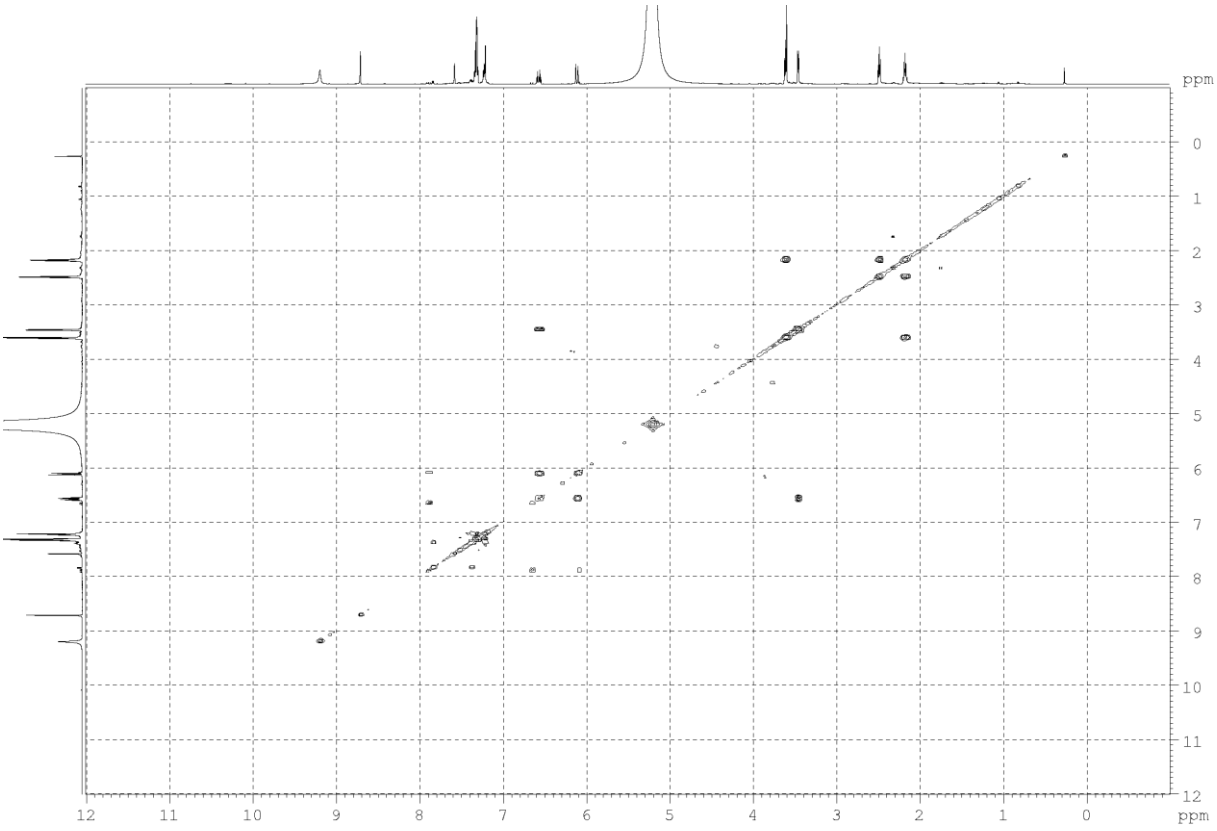


Figure A 17: ^{1}H, ^{1}H COSY spectrum of massargiline in pyridine-d_5

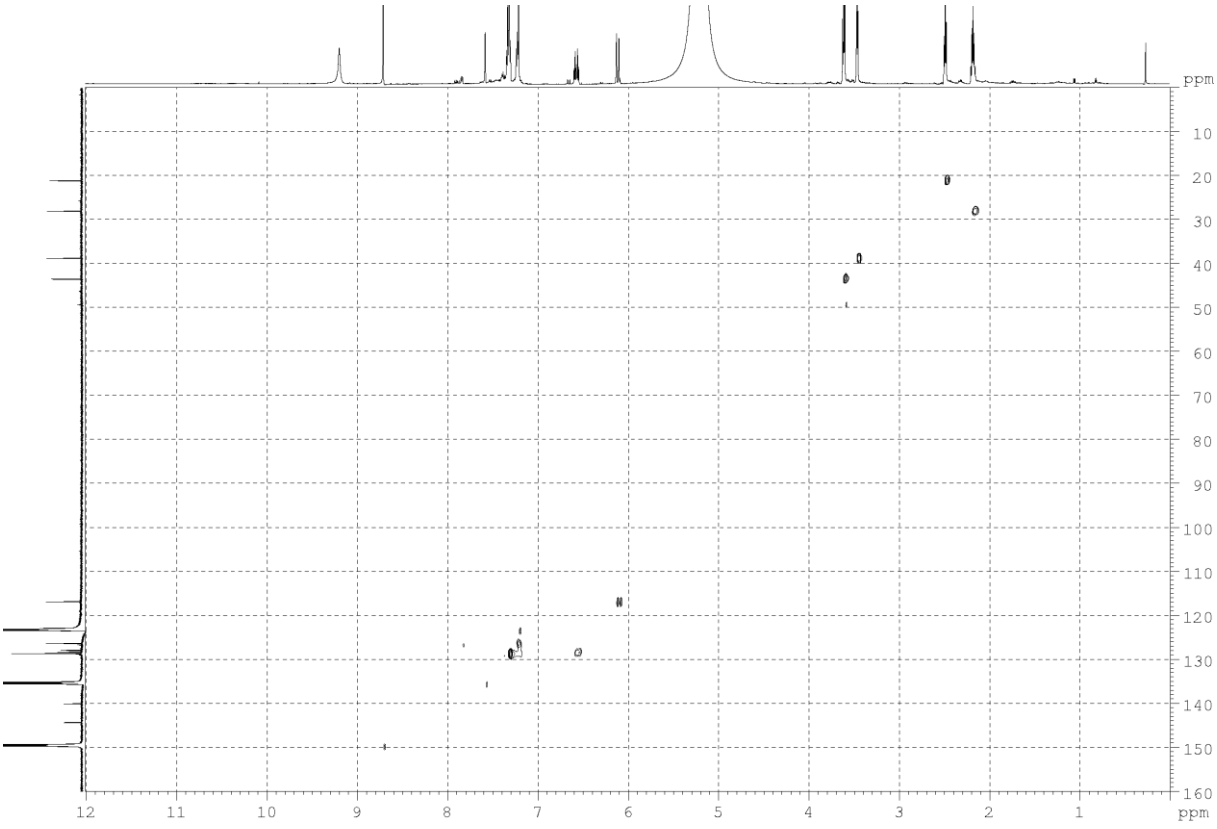


Figure A 18: ^{1}H, ^{13}C HSQC spectrum of massargiline in pyridine-d_5

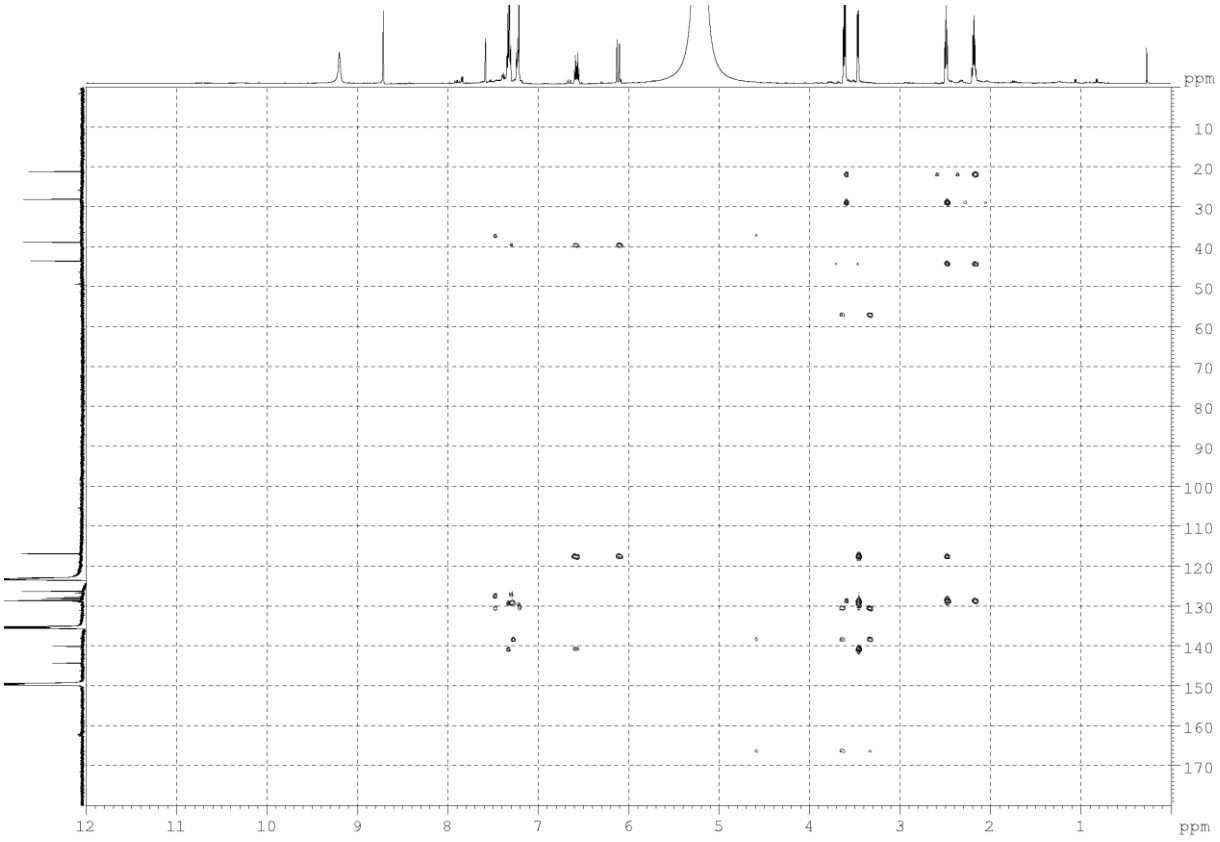


Figure A 19: ^{1}H, ^{13}C HMBC spectrum of massargiline in pyridine-d5

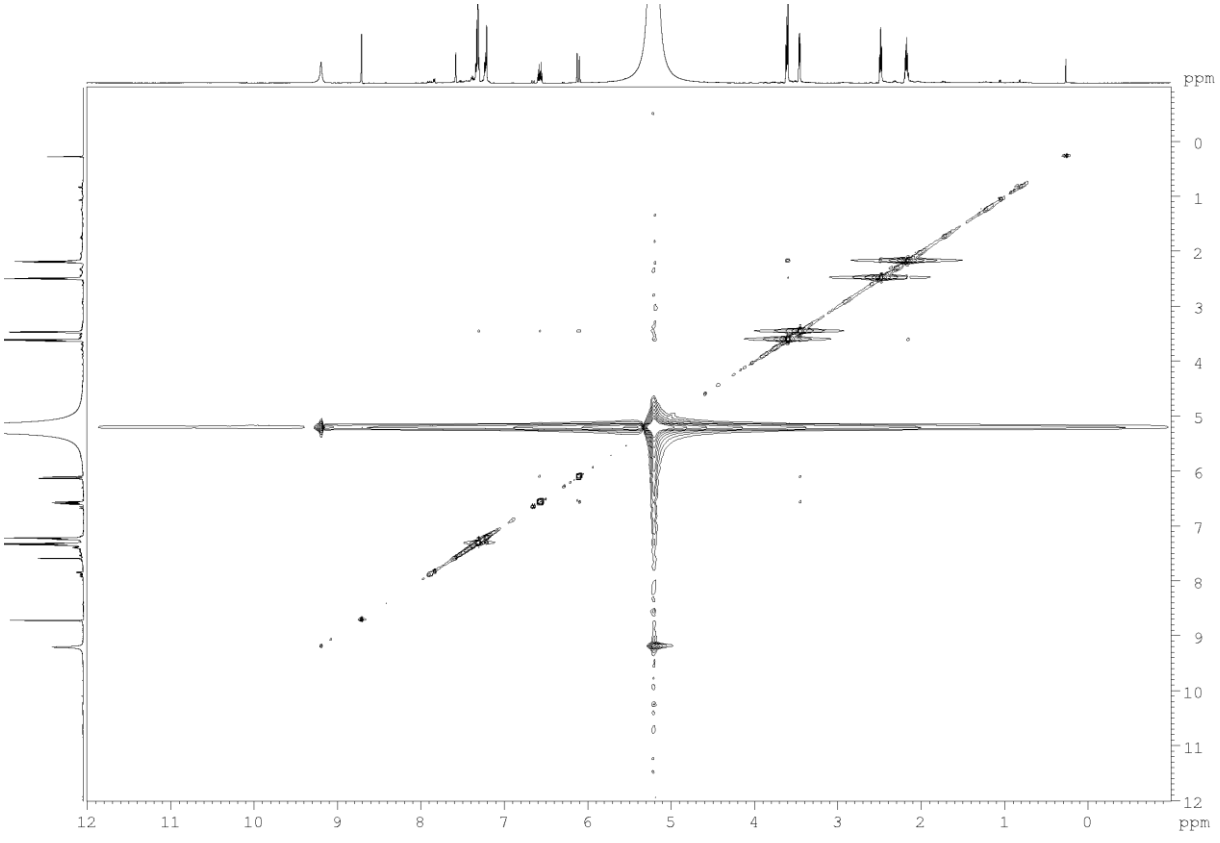


Figure A 20: ^{1}H, ^{1}H NOESY spectrum of massargiline in pyridine-d_5

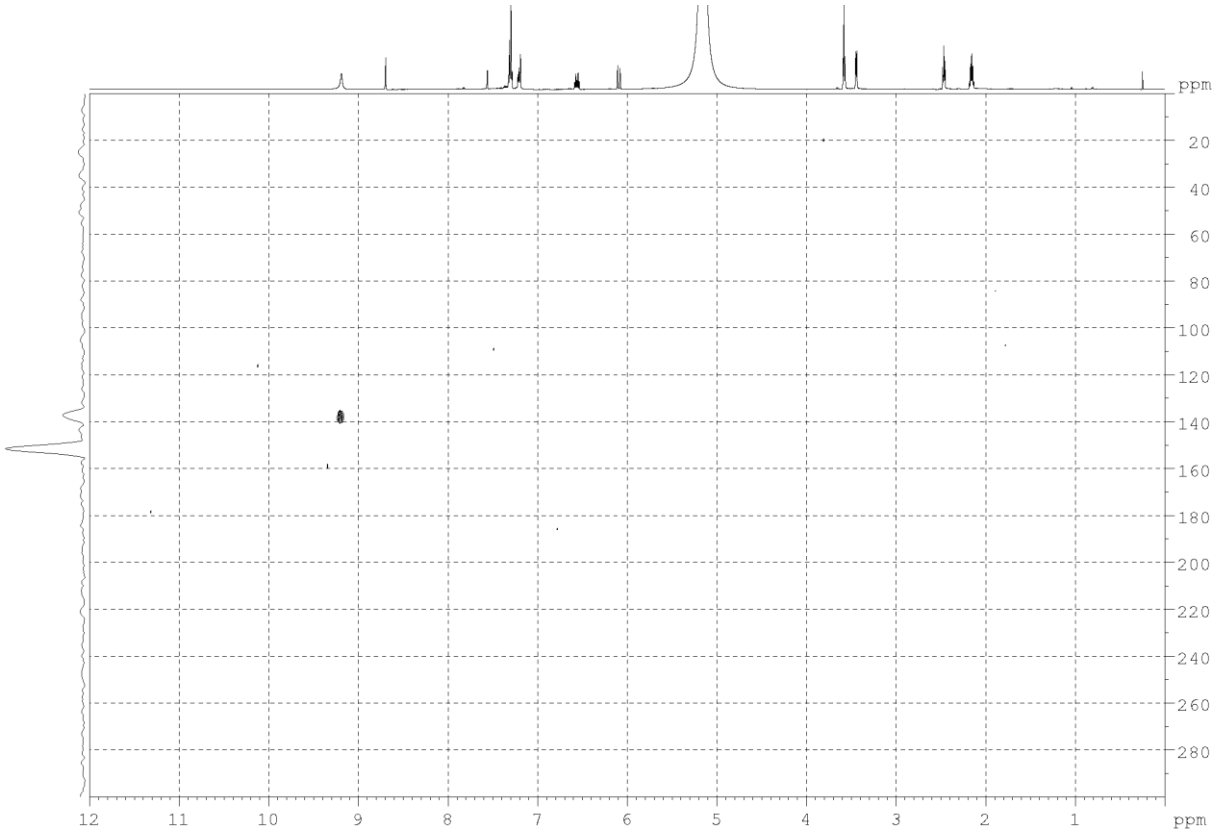


Figure A 21: ^{1}H, ^{15}N HSQC spectrum of massargiline in pyridine-d_5

Table A 1: NMR data of massargiline in pyridine-d5

#	δ_C [a]		δ_H, M (*J* in Hz) [a]	COSY	HMBC ($^{1}H \rightarrow ^{13}C$)	HMBC ($^{1}H \rightarrow ^{15}N$)
1	144.6	C_q				
2	140.2	C_q				
3	128.1	C_q				
4	128.9	CH 2	x 7.28, not resolved	7		
5	128.8	CH 2	x 7.28, not resolved			
6	128.4	CH	6.55, dt (16.0, 7.0)	8, 11	2, 9, 11	
7	126.5	CH	7.18, not resolved	4	5	
8	117.1	CH	6.07, d (16.0)	6	3, 9, 11	15
9	117.0	C_q [b]				
10	43.7	CH_2	3.56, t (7.2)	12	(1), 3, 12, 13	16
11	39.0	CH_2	3.43, d (7.0)	6	2, 5, 8	
12	28.3	CH_2	2.14, quint (7.3)	10, 13	3, 10, 13	16
13	21.3	CH_2	2.45, t (7.4)	12	3, 9, 10, 12	16

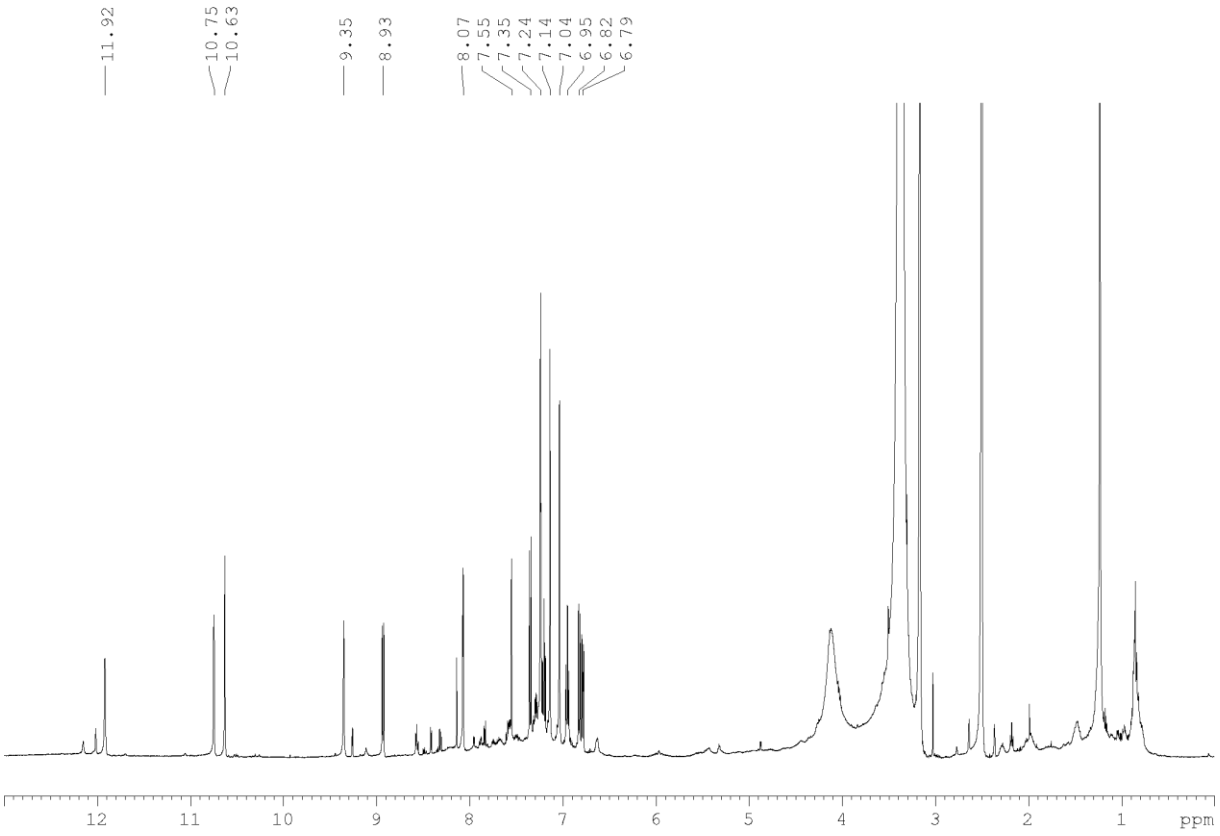


Figure A 22: ^{1}H NMR (500 MHz) spectrum of violacein in DMSO-d_6

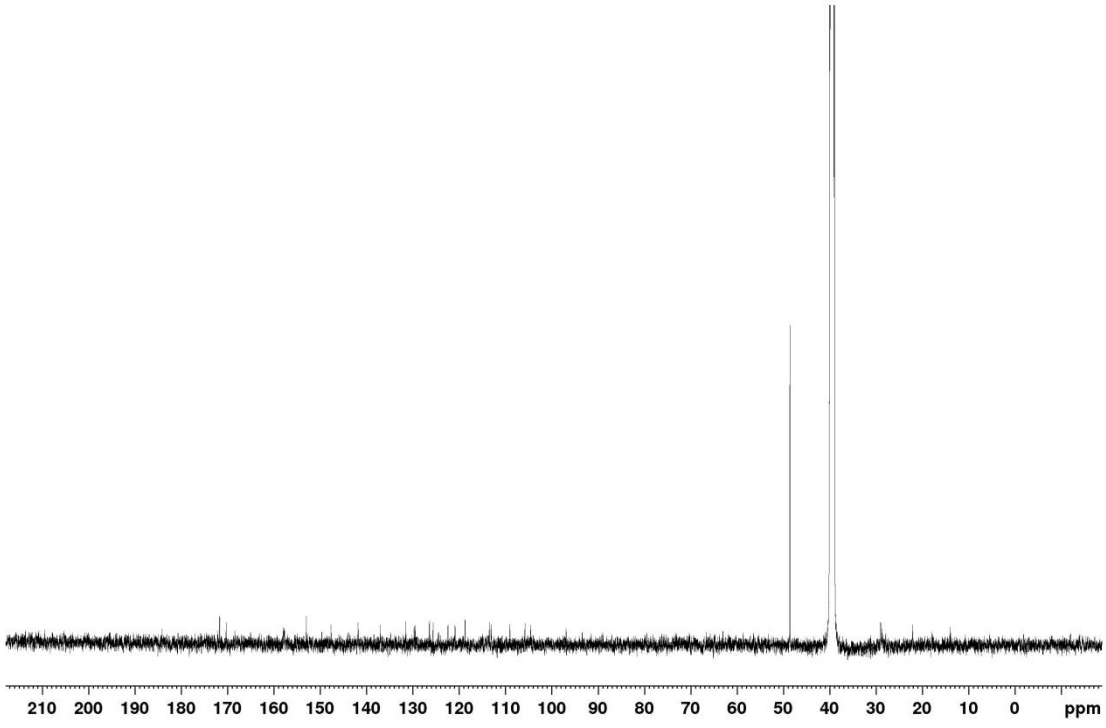


Figure A 23: ^{13}C NMR (500 MHz) spectrum of violacein in DMSO-d_6

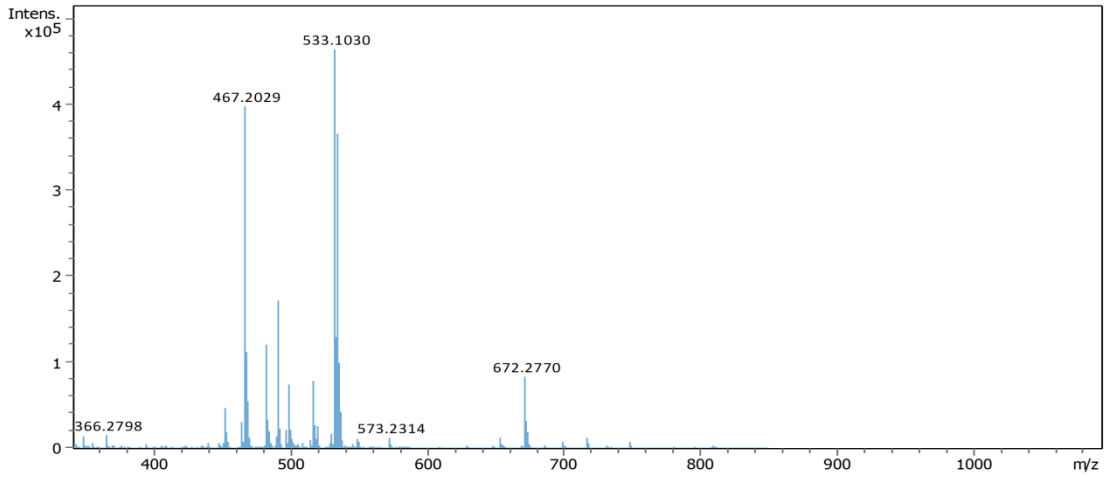


Figure A 24: HR-ESI-MS spectrum of massiliachelin + gallium

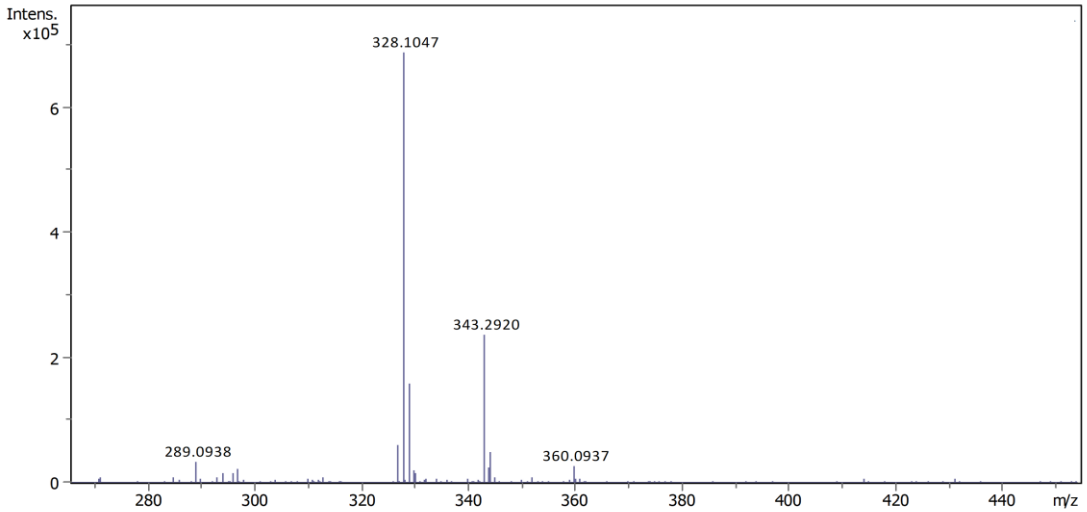


Figure A 25: HR-ESI-MS spectrum of fraction deoxyviolacein

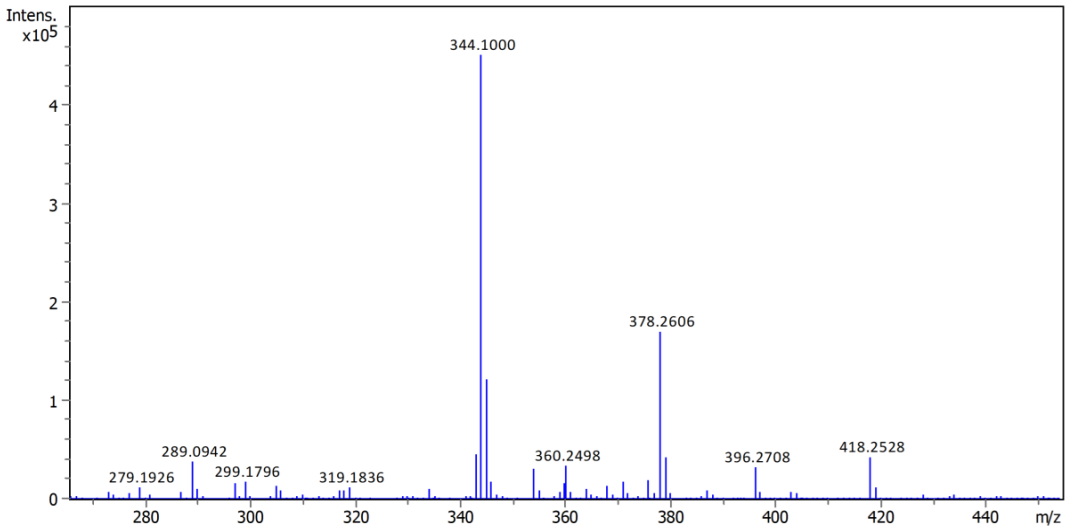


Figure A 26: HR-ESI-MS spectrum of violacein

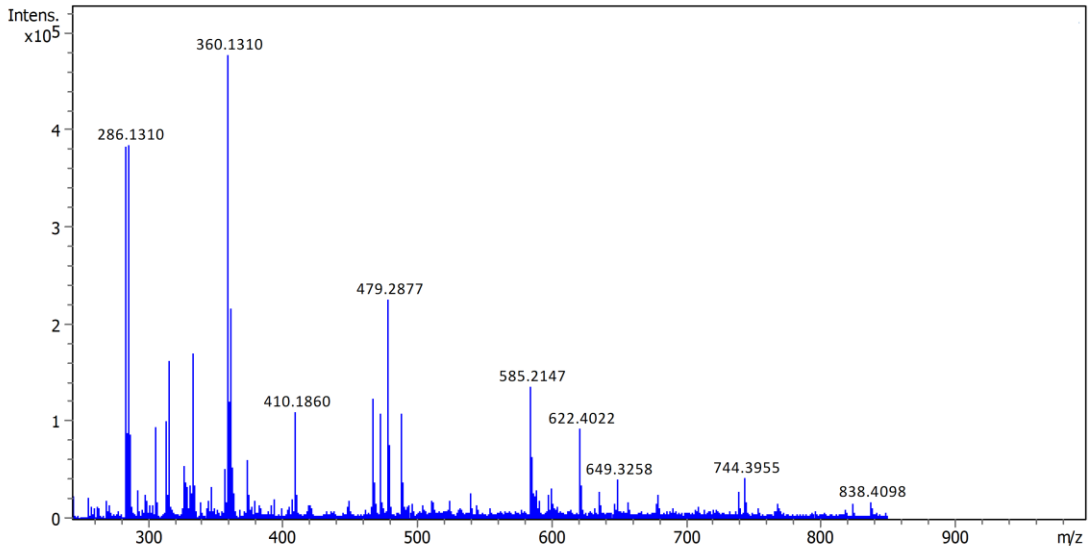


Figure A 27: HR-ESI-MS spectrum of oxyviolacein

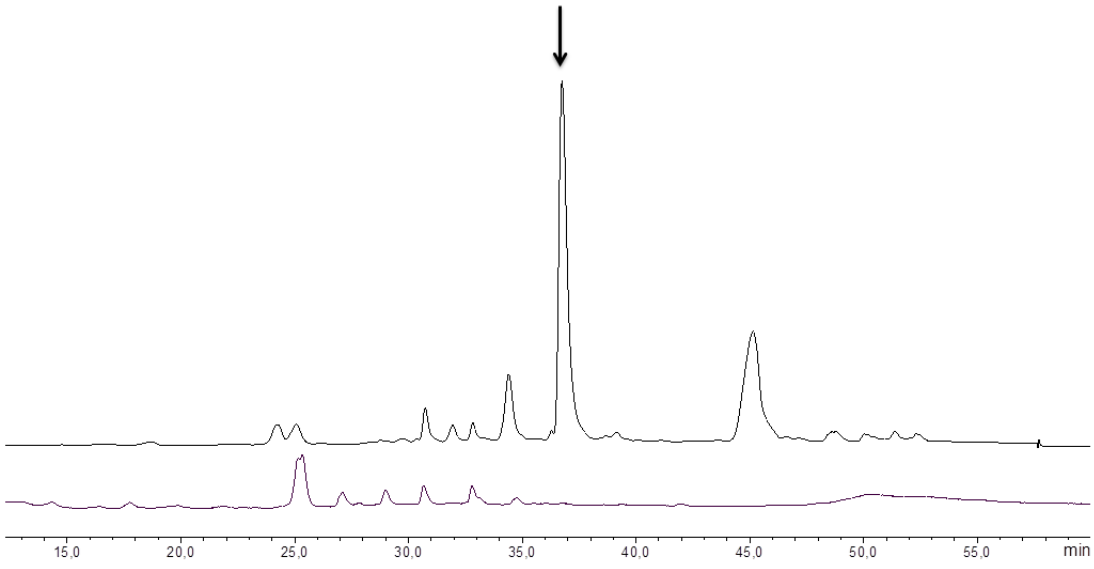


Figure A 28: Chromatograms of *Massilia* sp. NR. 4-1 extract (top) and the R2A medium control (bottom). The peak of massargiline is marked with an arrow.

Table A 2: calibration curve of massargiline production	
Injection volume V_{inj}	Peak area / (mAU * min)
3 µl	15104
10 µl	55290
50 µl	272656
100 µl	546535
120 µl	640204
150 µl	826311
180 µl	1016817
250 µl	1386561

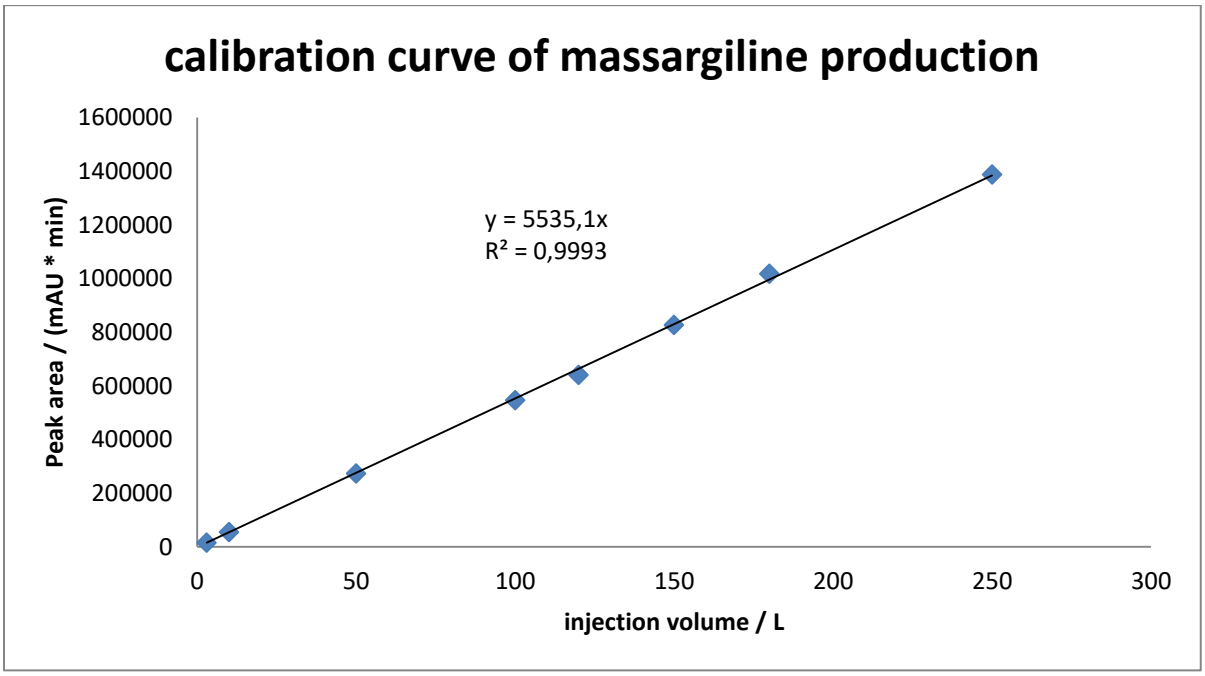


Figure A 29: calibration curve of massargiline production

Table A 3: Peak areas of massargiline measured over 96 h			
sample	Time	Peak area / (mAU * min)	Concentration (mg / L)
0	0	11086	0
1	19,6	56617	0,1316
2	25,5	181089	0,4914
3	44	739082	2,1044
4	50,6	888465	2,5362
5	67	1203019	3,4455
6	95	1307169	3,7465

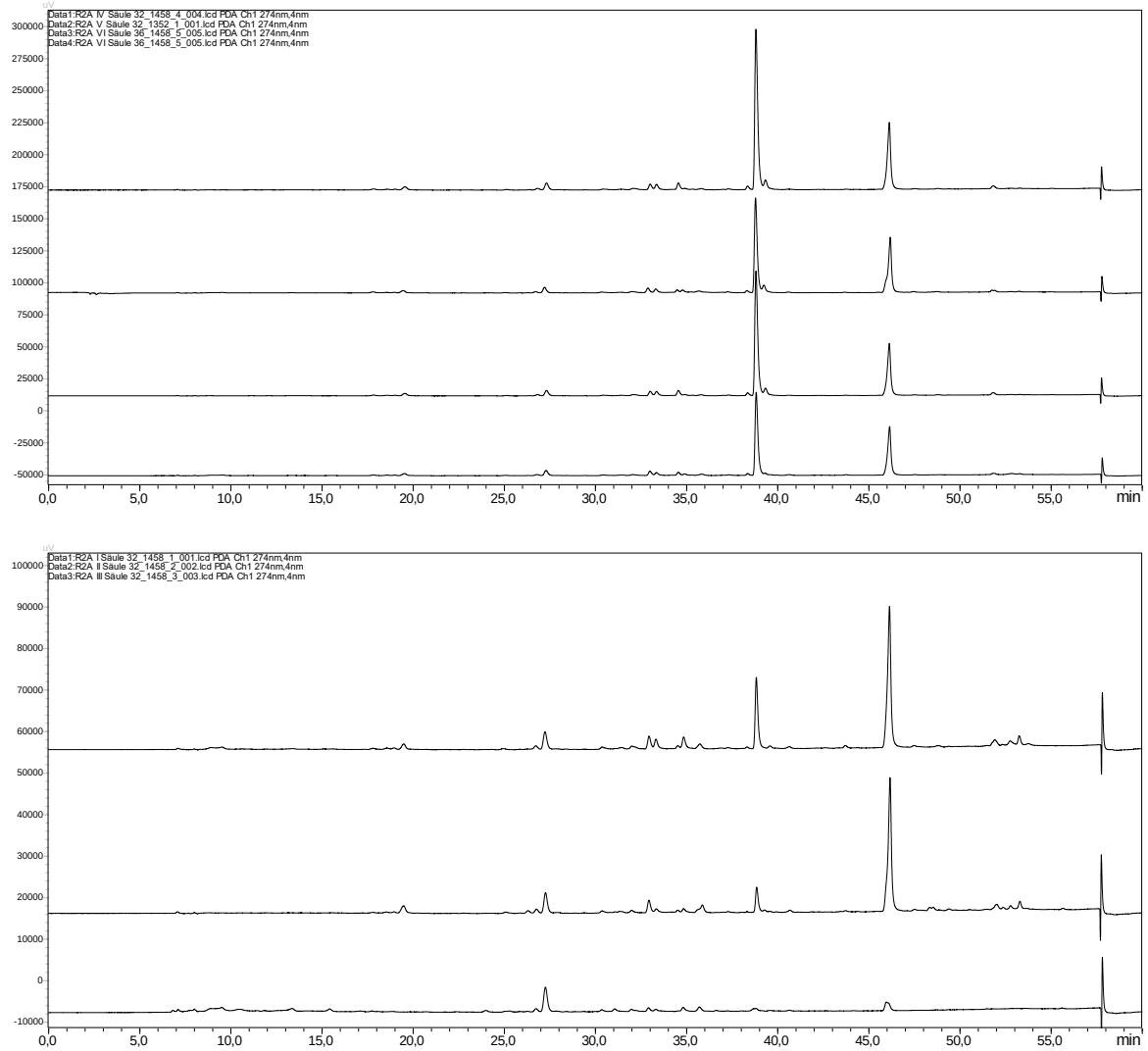


Figure A 30: Chromatograms of strain NR 4-1 over 96 hours

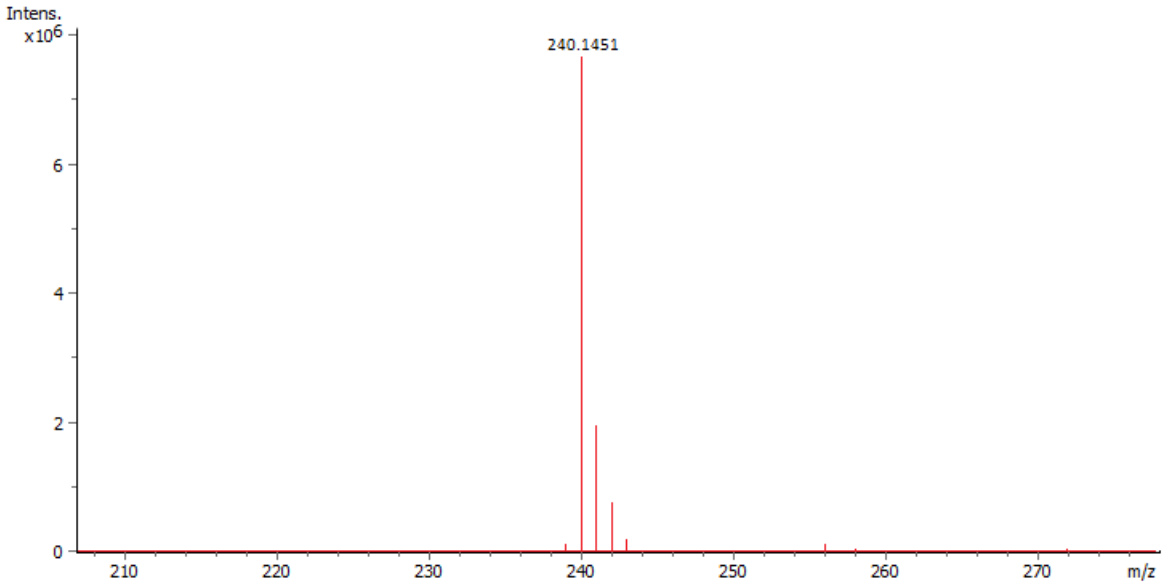


Figure A 31: HR-ESI-MS spectrum of massargiline after feeding [1-^{13}C]-acetate in complex R2A medium

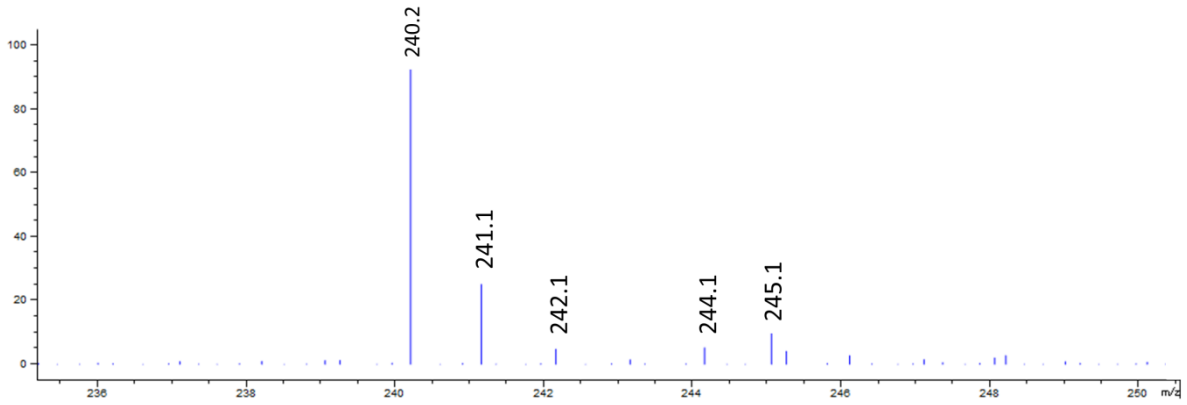


Figure A 32: MS data of massargiline (bottom) with [1-^{13}C]-glucose feeding.

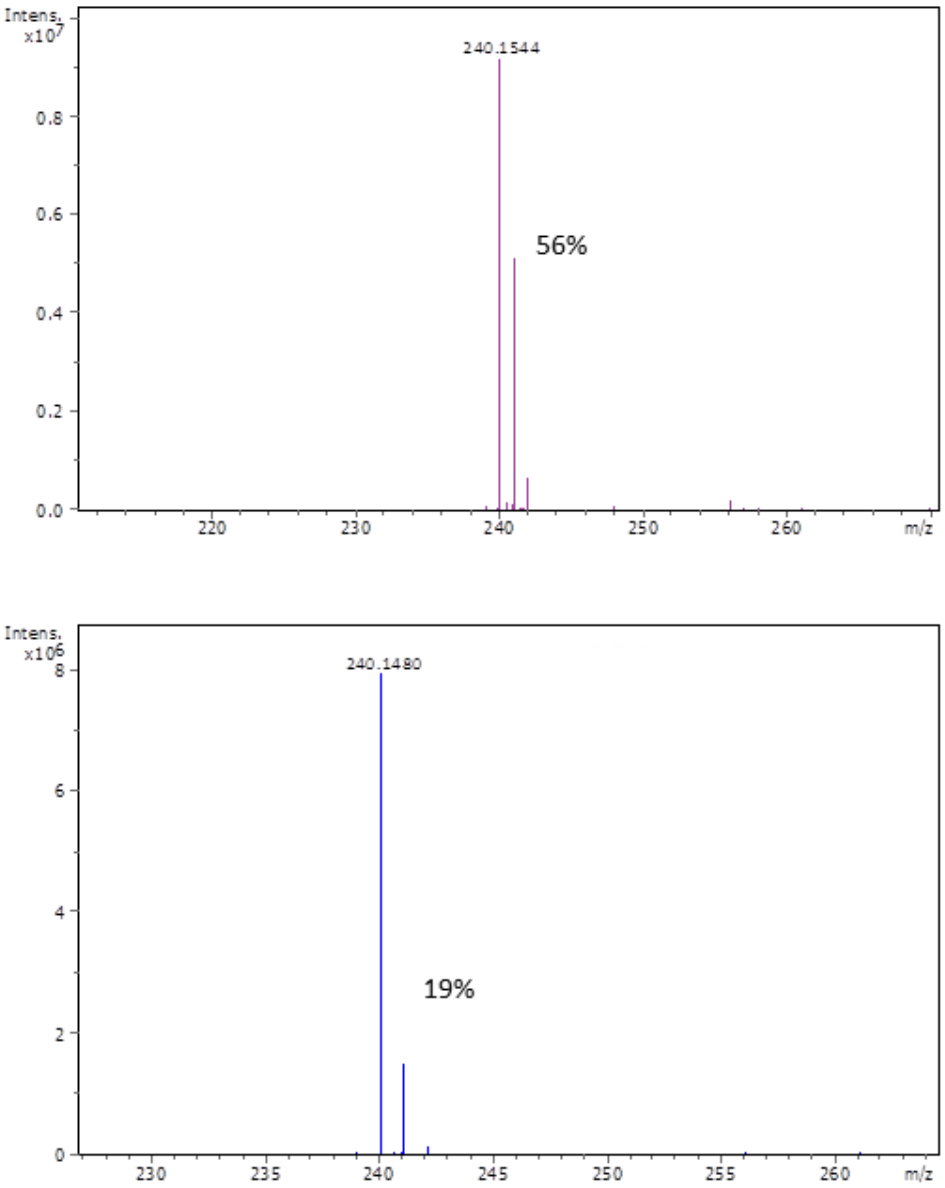


Figure A 33: HR-ESI-MS spectrum of massargiline after feeding [1-13C]-acetate in reduced R2A medium after 20 hours (top) and HR-ESI-MS spectrum of massargiline without the addition of labeled precursor (bottom). The intensity of the signal of with the a *m/z* value of 241 is given in percent relative to the value of the *m/z* value of 240.

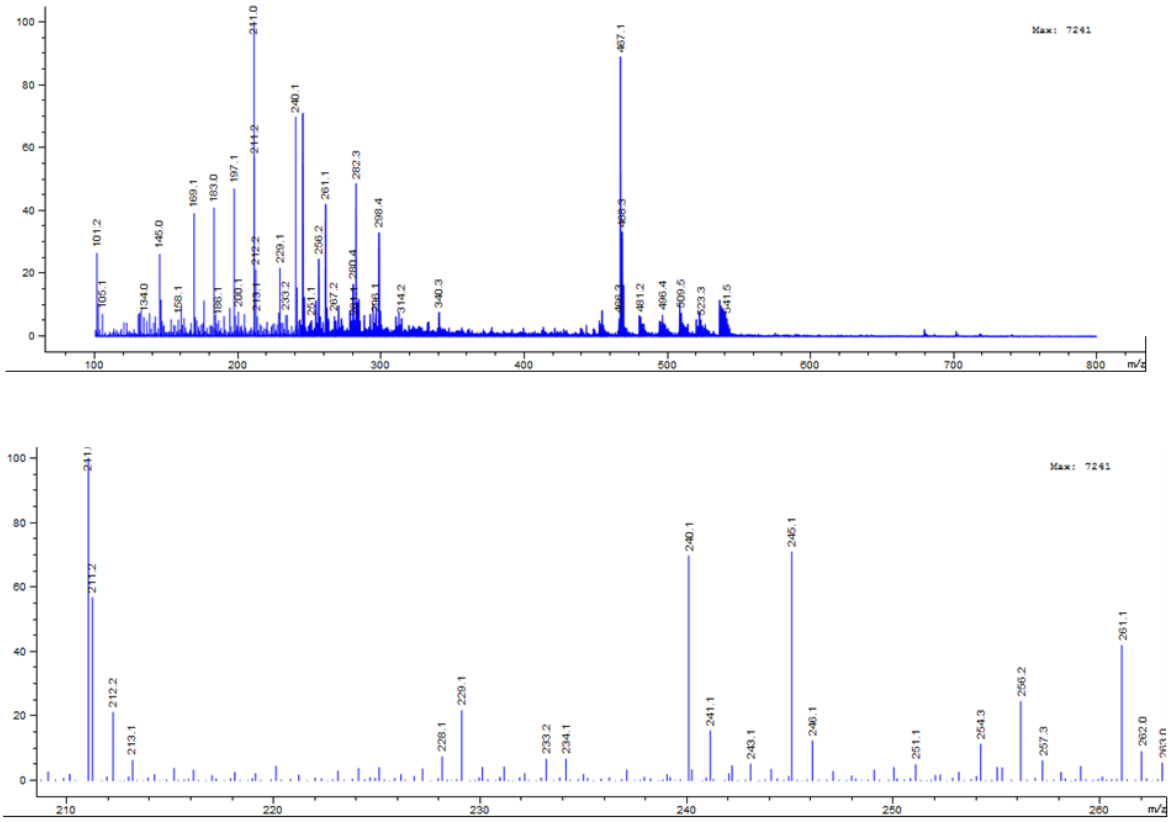


Figure A 34: MS data of crude strain NR 4-1 extract (top) and massargiline (bottom) with ethyl-[1-13C]-acetate after 20 hours in reduced R2A medium.

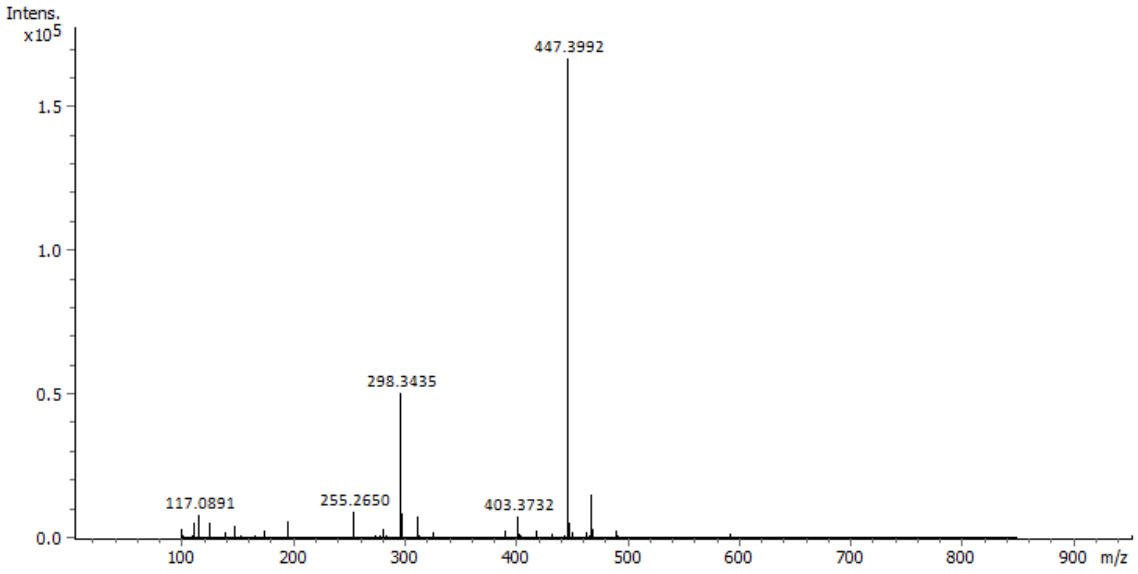


Figure A 35: HR-ESI-MS spectrum of *A. citrulli* AAC00-1 unknown compound

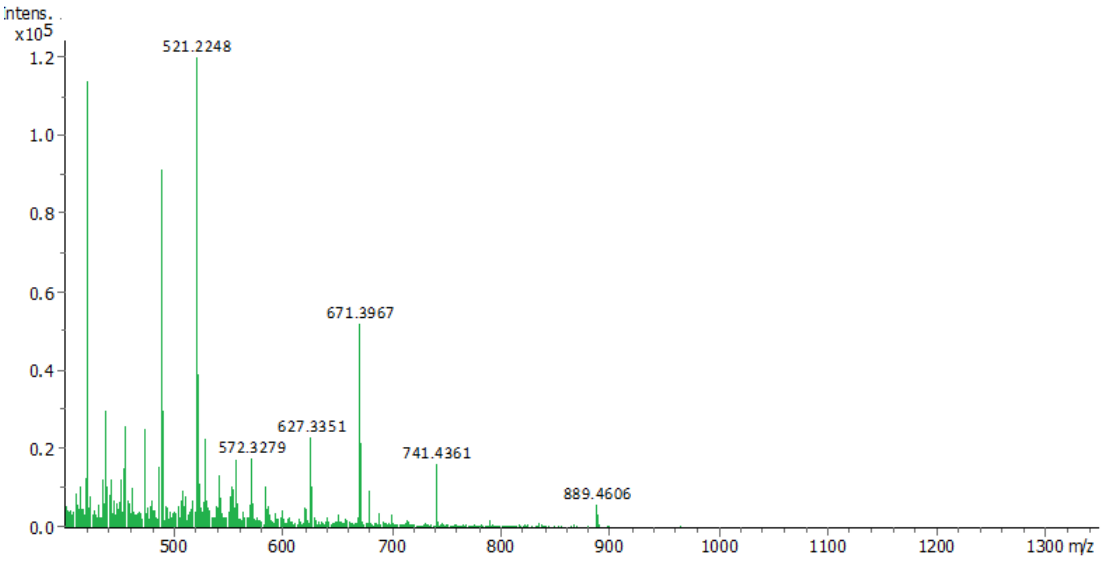


Figure A 36: HR-ESI-MS spectrum of *A. citrulli* AAC00-1 in LB medium

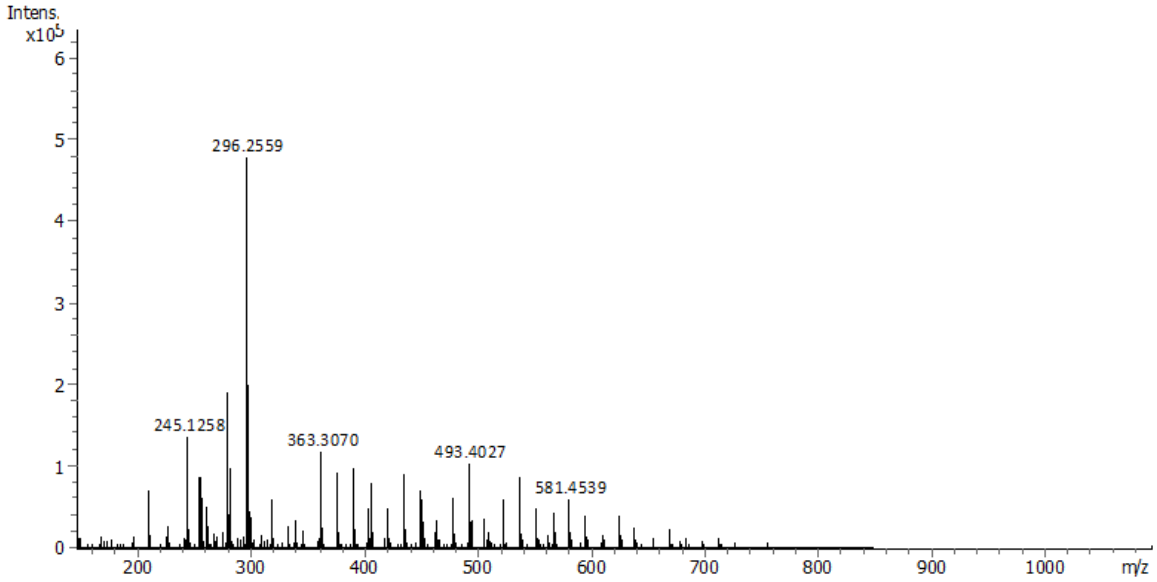


Figure A 37: HR-ESI-MS spectrum of *A. citrulli* AAC00-1 in H3 medium

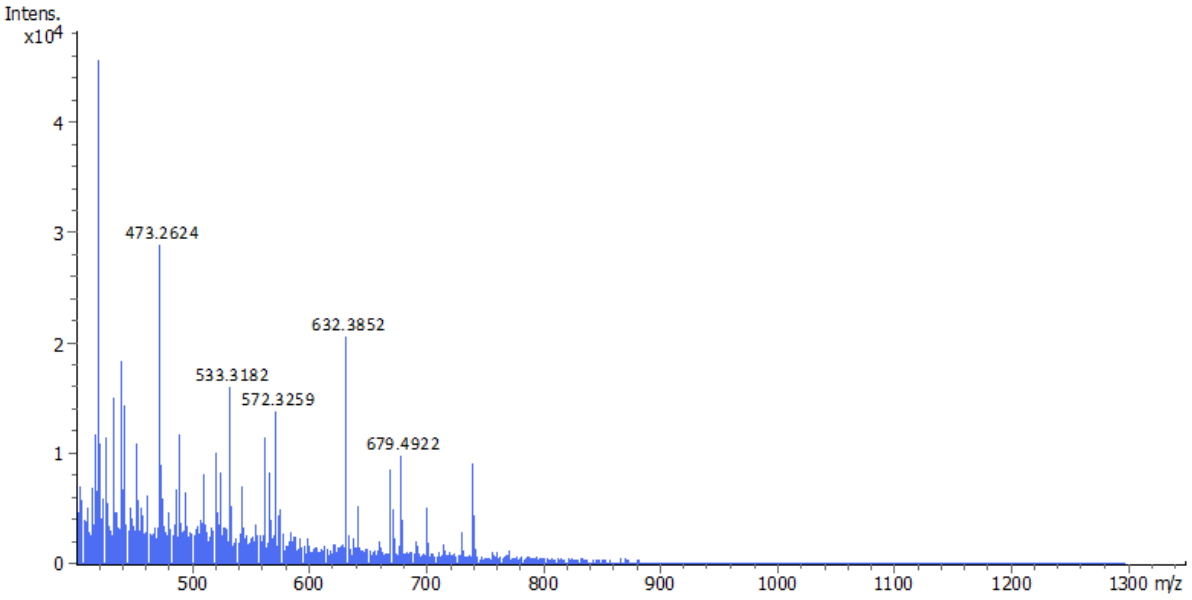


Figure A 38: HR-ESI-MS spectrum of *A. citrulli* AAC00-1 in ISP2 salt medium

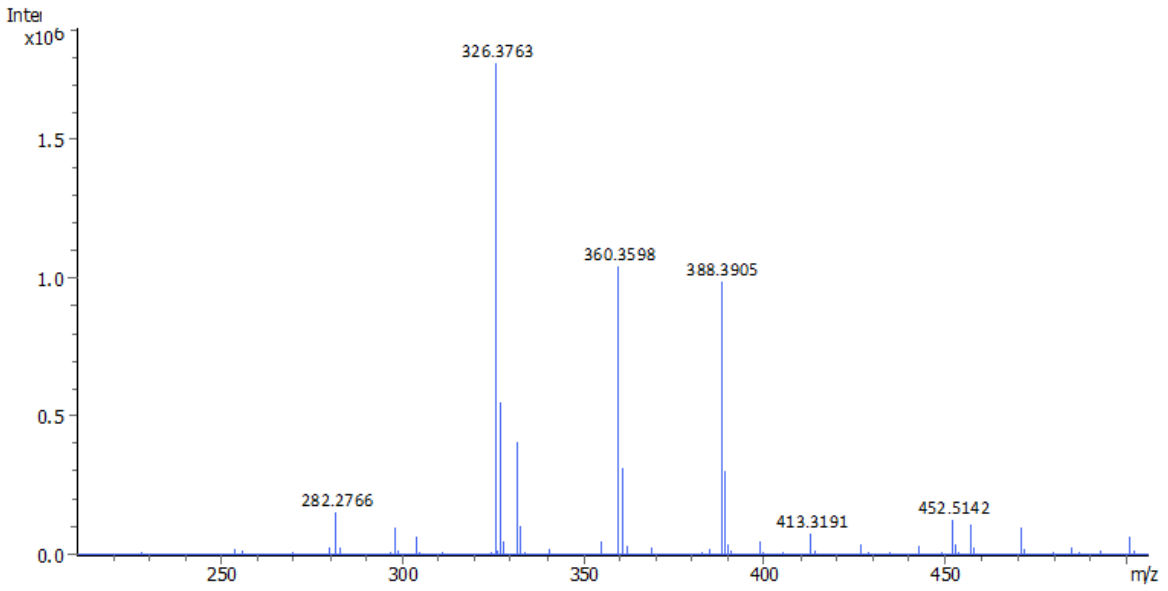


Figure A 39: HR-ESI-MS spectrum of LB medium control

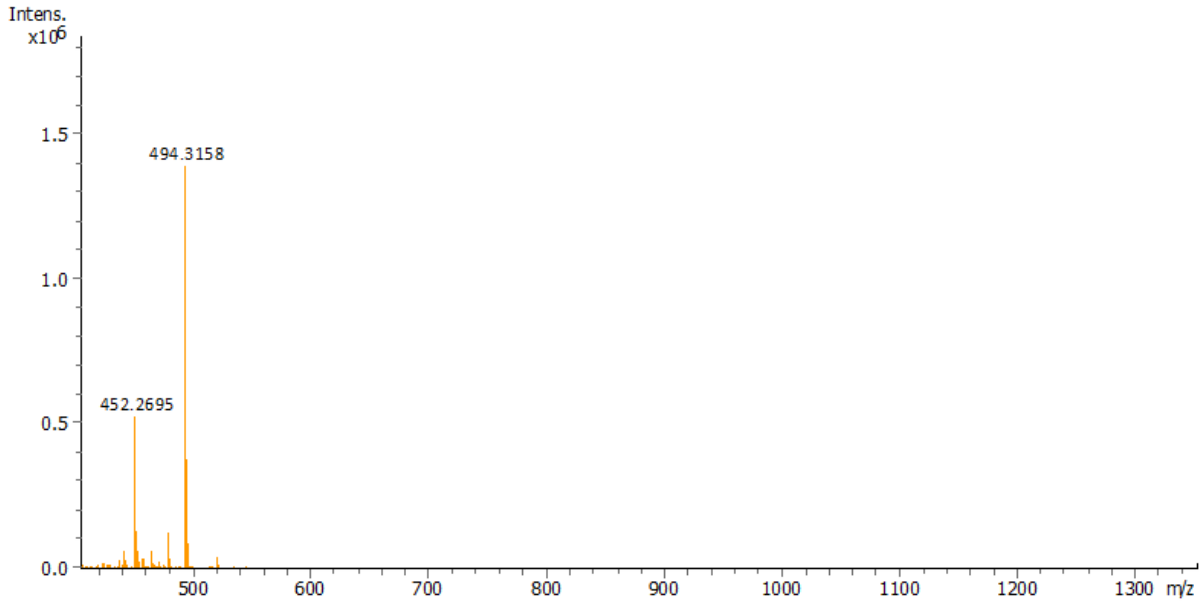


Figure A 40: Figure A 48: HR-ESI-MS spectrum of H3 medium control

8 Literature

[1] D. J. Newman and G. M. Cragg, “Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019,” *J. Nat. Prod.*, vol. 83, no. 3, pp. 770–803, 2020, doi: 10.1021/acs.jnatprod.9b01285.

[2] F. E. Koehn and G. T. Carter, “The evolving role of natural products in drug discovery,” *Nat. Rev. Drug Discov.*, vol. 4, no. 3, pp. 206–220, 2005, doi: 10.1038/nrd1657.

[3] L. Katz and R. H. Baltz, “Natural product discovery: past, present, and future,” *J. Ind. Microbiol. Biotechnol.*, vol. 43, no. 2–3, pp. 155–176, 2016, doi: 10.1007/s10295-015-1723-5.

[4] S. J. Pidot, S. Coyne, F. Kloss, and C. Hertweck, “Antibiotics from neglected bacterial sources,” *Int. J. Med. Microbiol.*, vol. 304, no. 1, pp. 14–22, 2014, doi: 10.1016/j.ijmm.2013.08.011.

[5] A. P. Magiorakos *et al.*, “Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: An international expert proposal for interim standard definitions for acquired resistance,” *Clin. Microbiol. Infect.*, vol. 18, no. 3, pp. 268–281, 2012, doi: 10.1111/j.1469-0691.2011.03570.x.

[6] N. Ziemert, M. Alanjary, and T. Weber, “The evolution of genome mining in microbes-a review,” *Nat. Prod. Rep.*, vol. 33, no. 8, pp. 988–1005, 2016, doi: 10.1039/c6np00025h.

[7] R. H. Baltz, “Gifted microbes for genome mining and natural product discovery,” *J. Ind. Microbiol. Biotechnol.*, vol. 44, no. 4–5, pp. 573–588, 2017, doi: 10.1007/s10295-016-1815-x.

[8] B. Gosio, “Contributo all’etiologia della pellagra; ricerche chimiche e batteriologiche sulle alterazioni del mais,” *G. R. Accad. Med. Torino*, vol. 61, p. 484, 1893.

[9] R. Bentley, “Mycophenolic acid: A one hundred year odyssey from antibiotic to immunosuppressant,” *Chem. Rev.*, vol. 100, no. 10, pp. 3801–3825, 2000, doi: 10.1021/cr990097b.

[10] K. C. Nicolaou and S. Rigol, “A brief history of antibiotics and select advances in their synthesis,” *J. Antibiot. (Tokyo).*, vol. 71, no. 2, pp. 153–184, 2018, doi: 10.1038/ja.2017.62.

[11] S. Y. Tan and Y. Tatsumura, “Alexander Fleming (1881–1955): Discoverer of penicillin,” *Singapore Med. J.*, vol. 56, no. 7, pp. 366–367, 2015, doi: 10.11622/smedj.2015105.

[12] A. Fleming, “On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of b. Influenzae,” *Br J Exp Pathol*, vol. 10, no. 2, pp. 226–236, 1929.

[13] a. S. Waksman and M. Tishler, “Chemical Nature of Actinomycin,” *J. Biol. Chem.*, vol. 142, pp. 519–528, 1942.

[14] R. E. de Lima Procópio, I. R. da Silva, M. K. Martins, J. L. de Azevedo, and J. M. de Araújo, “Antibiotics produced by Streptomyces,” *Brazilian J. Infect. Dis.*, vol. 16, no. 5, pp. 466–471, 2012, doi: 10.1016/j.bjid.2012.08.014.

[15] G. Dantas and M. O. A. Sommer, “How to fight back against antibiotic resistance,” *Am. Sci.*, vol. 102, no. 1, pp. 42–51, 2014, doi: 10.1511/2014.106.42.

[16] R. I. Aminov, “A brief history of the antibiotic era: Lessons learned and challenges for the future,” *Front. Microbiol.*, vol. 1, no. DEC, pp. 1–7, 2010, doi: 10.3389/fmicb.2010.00134.

[17] G. S. Bbosa, N. Mwebaza, J. Odda, D. B. Kyegombe, and M. Ntale, “Antibiotics/antibacterial drug use, their marketing and promotion during the post-antibiotic golden age and their role in emergence of bacterial resistance,” *Health (Irvine. Calif).*, vol. 06, no. 05, pp. 410–425, 2014, doi: 10.4236/health.2014.65059.

[18] K. Gould, “Antibiotics: From prehistory to the present day,” *J. Antimicrob. Chemother.*, vol. 71, no. 3, pp. 572–575, 2016, doi: 10.1093/jac/dkv484.

[19] D. P. Levine, “Vancomycin : A History,” vol. 42, no. Suppl 1, pp. 5–12, 2006.

[20] S. M. Drawz and R. A. Bonomo, “Three decades of β-lactamase inhibitors,” *Clin. Microbiol. Rev.*, vol. 23, no. 1, pp. 160–201, 2010, doi:

10.1128/CMR.00037-09.

[21] L. F. Wright and D. A. Hopwood, “Identification of the antibiotic determined by the SCP1 plasmid of Streptomyces coelicolor A3(2),” *J. Gen. Microbiol.*, vol. 95, no. 1, pp. 96–106, 1976, doi: 10.1099/00221287-95-1-96.

[22] S. Pelzer *et al.*, “Identification and analysis of the balhimycin biosynthetic gene cluster and its use for manipulating glycopeptide biosynthesis in Amycolatopsis mediterranei DSM5908,” *Antimicrob. Agents Chemother.*, vol. 43, no. 7, pp. 1565–1573, 1999, doi: 10.1128/aac.43.7.1565.

[23] S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, and L. Katz, “Modular Organization of Genes Required for Complex Polyketide Biosynthesis,” *Science (80-.).*, vol. 252, pp. 675–679, 1991.

[24] N. R. Thomson *et al.*, “Colicolor Genom.Pdf,” vol. 3, no. 2, 2002.

[25] J. Shendure *et al.*, “DNA sequencing at 40: Past, present and future,” *Nature*, vol. 550, no. 7676, 2017, doi: 10.1038/nature24286.

[26] L. Laureti *et al.*, “Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in Streptomyces ambofaciens,” *Proc. Natl. Acad. Sci.*, vol. 108, no. 15, pp. 6258 LP – 6263, Apr. 2011, doi: 10.1073/pnas.1019077108.

[27] A. Starcevic, J. Zucko, J. Simunkovic, P. F. Long, J. Cullum, and D. Hranueli, “ClustScan: An integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures,” *Nucleic Acids Res.*, vol. 36, no. 21, pp. 6882–6892, 2008, doi: 10.1093/nar/gkn685.

[28] M. H. T. Li, P. M. U. Ung, J. Zajkowski, S. Garneau-Tsodikova, and D. H. Sherman, “Automated genome mining for natural products,” *BMC Bioinformatics*, vol. 10, pp. 1–10, 2009, doi: 10.1186/1471-2105-10-185.

[29] S. Anand *et al.*, “SBSPKS: Structure based sequence analysis of polyketide synthases,” *Nucleic Acids Res.*, vol. 38, no. SUPPL. 2, pp. 487–496, 2010, doi: 10.1093/nar/gkq340.

[30] K. Blin *et al.*, “antiSMASH 5.0: updates to the secondary metabolite genome

mining pipeline," *Nucleic Acids Res.*, vol. 47, no. W1, pp. W81–W87, 2019, doi: 10.1093/nar/gkz310.

[31] T. Weber, K. Welzel, S. Pelzer, A. Vente, and W. Wohlleben, "Exploiting the genetic potential of polyketide producing streptomycetes," *J. Biotechnol.*, vol. 106, no. 2–3, pp. 221–232, 2003, doi: 10.1016/j.jbiotec.2003.08.004.

[32] J. Shi *et al.*, "Comparative genome mining and heterologous expression of an orphan NRPS gene cluster direct the production of ashimides," *Chem. Sci.*, vol. 10, no. 10, pp. 3042–3048, 2019, doi: 10.1039/c8sc05670f.

[33] S. Zhao *et al.*, "Prediction and characterization of enzymatic activities guided by sequence similarity and genome neighborhood networks," *Elife*, vol. 3, 2014, doi: 10.7554/elife.03275.

[34] J. D. Rudolf, X. Yan, and B. Shen, "Genome neighborhood network reveals insights into enediyne biosynthesis and facilitates prediction and prioritization for discovery," *J. Ind. Microbiol. Biotechnol.*, vol. 43, no. 2–3, pp. 261–276, 2016, doi: 10.1007/s10295-015-1671-0.

[35] F. Kandlinger, M. G. Plach, and R. Merkl, "AGeNNT: Annotation of enzyme families by means of refined neighborhood networks," *BMC Bioinformatics*, vol. 18, no. 1, pp. 1–13, 2017, doi: 10.1186/s12859-017-1689-6.

[36] M. Mihelčić, T. Šmuc, and F. Supek, "Patterns of diverse gene functions in genomic neighborhoods predict gene function and phenotype," *Sci. Rep.*, vol. 9, no. 1, pp. 1–16, 2019, doi: 10.1038/s41598-019-55984-0.

[37] I. B. Rogozin, "Connected gene neighborhoods in prokaryotic genomes," *Nucleic Acids Res.*, vol. 30, no. 10, pp. 2212–2223, 2002, doi: 10.1093/nar/30.10.2212.

[38] I. Schmitt and F. K. Barker, "Phylogenetic methods in natural product research," *Nat. Prod. Rep.*, vol. 26, no. 12, pp. 1585–1602, 2009, doi: 10.1039/b910458p.

[39] S. Kroken, N. L. Glass, J. W. Taylor, O. C. Yoder, and B. G. Turgeon, "Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 26, pp.

15670–15675, 2003, doi: 10.1073/pnas.2532165100.

[40] K. Anzai *et al.*, “Comparison of groupings among members of the genus Aspergillus based on phylogeny and production of bioactive compounds,” *Biosci. Biotechnol. Biochem.*, vol. 72, no. 8, pp. 2199–2202, 2008, doi: 10.1271/bbb.80123.

[41] K. C. Belknap, C. J. Park, B. M. Barth, and C. P. Andam, “Genome mining of biosynthetic and chemotherapeutic gene clusters in Streptomyces bacteria,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–9, 2020, doi: 10.1038/s41598-020-58904-9.

[42] I. Kjærbølling, T. Vesth, and M. R. Andersen, “Resistance Gene-Directed Genome Mining of 50 Aspergillus Species,” *mSystems*, vol. 4, no. 4, pp. 1–13, 2019, doi: 10.1128/msystems.00085-19.

[43] X. Tang *et al.*, “Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-directed Genome Mining,” *ACS Chem. Biol.*, vol. 10, no. 12, pp. 2841–2849, 2015, doi: 10.1021/acschembio.5b00658.

[44] G. D. Wright, “Molecular mechanisms of antibiotic resistance,” *Chem. Commun.*, vol. 47, no. 14, pp. 4055–4061, 2011, doi: 10.1039/c0cc05111j.

[45] Z. Wang, S. Li, M. Steffensky, A. Mu, and L. Heide, “Identification of the Novobiocin Biosynthetic Gene Cluster of Streptomyces spheroides NCIB 11891,” vol. 44, no. 5, pp. 1214–1222, 2000.

[46] C. N. Tetzlaff, Z. You, D. E. Cane, S. Takamatsu, S. Omura, and H. Ikeda, “A Gene Cluster for Biosynthesis of the Sesquiterpenoid Antibiotic Pentalenolactone in Streptomyces a V ermitilis †,” pp. 6179–6186, 2006, doi: 10.1021/bi060419n.

[47] A. J. Kale, R. P. McGlinchey, A. Lechner, and B. S. Moore, “Bacterial self-resistance to the natural proteasome inhibitor salinosporamide A,” *ACS Chem. Biol.*, vol. 6, no. 11, pp. 1257–1264, 2011, doi: 10.1021/cb2002544.

[48] M. N. Thaker *et al.*, “Identifying producers of antibacterial compounds by screening for antibiotic resistance,” *Nat. Biotechnol.*, vol. 31, no. 10, pp. 922–927, 2013, doi: 10.1038/nbt.2685.

[49] M. F. Fillat, “The fur (ferric uptake regulator) superfamily: Diversity and

versatility of key transcriptional regulators," *Arch. Biochem. Biophys.*, vol. 546, pp. 41–52, 2014, doi: 10.1016/j.abb.2014.01.029.

[50] N. Baichoo and J. D. Helmann, "Recognition of DNA by Fur: A reinterpretation of the Fur box consensus sequence," *J. Bacteriol.*, vol. 184, no. 21, pp. 5826–5832, 2002, doi: 10.1128/JB.184.21.5826-5832.2002.

[51] M. Spohn, W. Wohlleben, and E. Stegmann, "Elucidation of the zinc-dependent regulation in Amycolatopsis japonicum enabled the identification of the ethylenediamine-disuccinate ([S,S]-EDDS) genes," *Environ. Microbiol.*, vol. 18, no. 4, pp. 1249–1263, 2016, doi: 10.1111/1462-2920.13159.

[52] S. Romano, S. A. Jackson, S. Patry, and A. D. W. Dobson, "Extending the 'one strain many compounds' (OSMAC) principle to marine microorganisms," *Mar. Drugs*, vol. 16, no. 7, pp. 1–29, 2018, doi: 10.3390/md16070244.

[53] H. B. Bode, B. Bethe, R. Höfs, and A. Zeeck, "Big effects from small changes: Possible ways to explore nature's chemical diversity," *ChemBioChem*, vol. 3, no. 7, pp. 619–627, 2002, doi: 10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9.

[54] R. T. Hewage, T. Aree, C. Mahidol, S. Ruchirawat, and P. Kittakoop, "One strain-many compounds (OSMAC) method for production of polyketides, azaphilones, and an isochromanone using the endophytic fungus Dothideomycete sp.," *Phytochemistry*, vol. 108, pp. 87–94, 2014, doi: 10.1016/j.phytochem.2014.09.013.

[55] S. Lautru, R. J. Deeth, L. M. Bailey, and G. L. Challis, "Discovery of a new peptide natural product by streptomyces coelicolor genome mining," *Nat. Chem. Biol.*, vol. 1, no. 5, pp. 265–269, 2005, doi: 10.1038/nchembio731.

[56] M. F. Kreutzer, H. Kage, and M. Nett, "Structure and biosynthetic assembly of cupriachelin, a photoreactive siderophore from the bioplastic producer cupriavidus necator H16," *J. Am. Chem. Soc.*, vol. 134, no. 11, pp. 5415–5422, 2012, doi: 10.1021/ja300620z.

[57] R. D. Kersten *et al.*, "A mass spectrometry-guided genome mining approach for natural product peptidogenomics," *Nat. Chem. Biol.*, vol. 7, no. 11, pp. 794–

802, 2011, doi: 10.1038/nchembio.684.

[58] M. F. Kreutzer and M. Nett, “Genomics-driven discovery of taiwachelin, a lipopeptide siderophore from Cupriavidus taiwanensis,” *Org. Biomol. Chem.*, vol. 10, no. 47, pp. 9338–9343, 2012, doi: 10.1039/c2ob26296g.

[59] T. Hoffmann, D. Krug, S. Hüttel, and R. Müller, “Improving natural products identification through targeted LC-MS/MS in an untargeted secondary metabolomics workflow,” *Anal. Chem.*, vol. 86, no. 21, pp. 10780–10788, 2014, doi: 10.1021/ac502805w.

[60] J. Korp, M. S. Vela Gurovic, and M. Nett, “Antibiotics from predatory bacteria,” *Beilstein J. Org. Chem.*, vol. 12, pp. 594–607, 2016, doi: 10.3762/bjoc.12.58.

[61] S. Kunakom and A. S. Eustáquio, “Burkholderia as a Source of Natural Products,” *J. Nat. Prod.*, vol. 82, no. 7, pp. 2018–2037, 2019, doi: 10.1021/acs.jnatprod.8b01068.

[62] S. Vijayakumar and M. Menakha, “Pharmaceutical applications of cyanobacteria-A review,” *J. Acute Med.*, vol. 5, no. 1, pp. 15–23, 2015, doi: 10.1016/j.jacme.2015.02.004.

[63] M. L. Micallef, P. M. D’Agostino, D. Sharma, R. Viswanathan, and M. C. Moffitt, “Genome mining for natural product biosynthetic gene clusters in the Subsection V cyanobacteria,” *BMC Genomics*, vol. 16, no. 1, pp. 1–20, 2015, doi: 10.1186/s12864-015-1855-z.

[64] Z. Chang *et al.*, “Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium Lyngbya majuscula,” *J. Nat. Prod.*, vol. 67, no. 8, pp. 1356–1367, 2004, doi: 10.1021/np0499261.

[65] K. Mohr, “Diversity of Myxobacteria—We Only See the Tip of the Iceberg,” *Microorganisms*, vol. 6, no. 3, p. 84, 2018, doi: 10.3390/microorganisms6030084.

[66] J. Herrmann, A. A. Fayad, and R. Müller, “Natural products from myxobacteria: Novel metabolites and bioactivities,” *Nat. Prod. Rep.*, vol. 34, no. 2, pp. 135–160, 2017, doi: 10.1039/c6np00106h.

[67] N. S. Cortina, D. Krug, A. Plaza, O. Revermann, and R. Müller, “Myxoprincomide: A natural product from myxococcus xanthus discovered by comprehensive analysis of the secondary metabolome,” *Angew. Chemie - Int. Ed.*, vol. 51, no. 3, pp. 811–816, 2012, doi: 10.1002/anie.201106305.

[68] J. J. Hug, F. Panter, D. Krug, and R. Müller, “Genome mining reveals uncommon alkylpyrones as type III PKS products from myxobacteria,” *J. Ind. Microbiol. Biotechnol.*, vol. 46, no. 3–4, pp. 319–334, 2019, doi: 10.1007/s10295-018-2105-6.

[69] A. S. Eustáquio, J. E. Janso, A. S. Ratnayake, C. J. O’Donnell, and F. E. Koehn, “Spliceostatin hemiketal biosynthesis in Burkholderia spp. is catalyzed by an iron/α-ketoglutarate-dependent dioxygenase,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 33, 2014, doi: 10.1073/pnas.1408300111.

[70] A. Khan, P. Singh, and A. Srivastava, “Synthesis, nature and utility of universal iron chelator – Siderophore: A review,” *Microbiol. Res.*, vol. 212–213, no. October 2017, pp. 103–111, 2018, doi: 10.1016/j.micres.2017.10.012.

[71] T. C. Johnstone and E. M. Nolan, “Beyond iron: Non-classical biological functions of bacterial siderophores,” *Dalt. Trans.*, vol. 44, no. 14, pp. 6320–6339, 2015, doi: 10.1039/c4dt03559c.

[72] C. Kurth, H. Kage, and M. Nett, “Siderophores as molecular tools in medical and environmental applications,” *Org. Biomol. Chem.*, vol. 14, no. 35, pp. 8212–8227, 2016, doi: 10.1039/c6ob01400c.

[73] V. I. Holden and M. A. Bachman, “Diverging roles of bacterial siderophores during infection,” *Metallomics*, vol. 7, no. 6, pp. 986–995, 2015, doi: 10.1039/c4mt00333k.

[74] M. Hofmann, T. Heine, V. Schulz, S. Hofmann, and D. Tischler, “Draft genomes and initial characterization of siderophore producing pseudomonads isolated from mine dump and mine drainage,” *Biotechnol. Reports*, vol. 25, p. e00403, 2020, doi: https://doi.org/10.1016/j.btre.2019.e00403.

[75] A. Crits-Christoph, N. Bhattacharya, M. R. Olm, Y. S. Song, and J. F. Banfield, “Transporter genes in biosynthetic gene clusters predict metabolite

characteristics and siderophore activity," *Genome Res.*, vol. 31, no. 2, pp. 239–250, 2021, doi: 10.1101/GR.268169.120.

[76] Z. L. Reitz, C. D. Hardy, J. Suk, J. Bouvet, and A. Butler, "Erratum: Genomic analysis of siderophore β-hydroxylases reveals divergent stereocontrol and expands the condensation domain family (Proceedings of the National Academy of Sciences of the United States of America (2019) 116 (19805-19814) DOI: 10.1073/pnas.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 39, p. 24599, 2020, doi: 10.1073/pnas.2018013117.

[77] A. C. W. Franken *et al.*, "Genome mining and functional genomics for siderophore production in Aspergillus niger," *Brief. Funct. Genomics*, vol. 13, no. 6, pp. 482–492, 2014, doi: 10.1093/bfgp/elu026.

[78] H. Bruns *et al.*, "Function-related replacement of bacterial siderophore pathways," *ISME J.*, vol. 12, no. 2, pp. 320–329, 2018, doi: 10.1038/ismej.2017.137.

[79] S. KONETSCHNY-RAPP, G. JUNG, J. MEIWES, and H. ZÄHNER, "Staphyloferrin A: a structurally new siderophore from staphylococci," *Eur. J. Biochem.*, vol. 191, no. 1, pp. 65–74, 1990, doi: 10.1111/j.1432-1033.1990.tb19094.x.

[80] A. W. Robertson *et al.*, "Isolation of Imaqobactin, an Amphiphilic Siderophore from the Arctic Marine Bacterium Variovorax Species RKJM285," *J. Nat. Prod.*, vol. 81, no. 4, pp. 858–865, 2018, doi: 10.1021/acs.jnatprod.7b00943.

[81] F. Archibald, "Lactobacillus plantarum, an organism not requiring iron," *FEMS Microbiol. Lett.*, vol. 19, no. 1, pp. 29–32, 1983, doi: 10.1111/j.1574-6968.1983.tb00504.x.

[82] J. E. Posey and F. C. Gherardini, "Lack of a role for iron in the Lyme disease pathogen," *Science (80-.).*, vol. 288, no. 5471, pp. 1651–1653, 2000, doi: 10.1126/science.288.5471.1651.

[83] K. Pantopoulos, S. K. Porwal, A. Tartakoff, and L. Devireddy, "Mechanisms of mammalian iron homeostasis," *Biochemistry*, vol. 51, no. 29, pp. 5705–5724, 2012, doi: 10.1021/bi300752r.

[84] S. Fardeau, C. Mullié, N. Audic, A. Sasaki, and P. Sonnet, “Bacterial iron uptake : a promising solution against multidrug resistant bacteria,” *Sci. against Microb. Pathog. Commun. Curr. Res. Technol. Adv.*, no. January, pp. 695–705, 2011.

[85] D. L. Huber, “Synthesis, properties, and applications of iron nanoparticles,” *Small*, vol. 1, no. 5, pp. 482–501, 2005, doi: 10.1002/smll.200500006.

[86] A. N. N. J. M. Messenger and R. Barclay, “BACTERIA , IRON AND PATHOGENICITY Table 1 Roles of iron in bacteria Affected Area,” vol. 11, no. 2, p. 1983, 1983.

[87] C. Zhang, “Essential functions of iron-requiring proteins in DNA replication, repair and cell cycle control,” *Protein Cell*, vol. 5, no. 10, pp. 750–760, 2014, doi: 10.1007/s13238-014-0083-7.

[88] R. Gozzelino and P. Arosio, “Iron homeostasis in health and disease,” *Int. J. Mol. Sci.*, vol. 17, no. 1, pp. 2–14, 2016, doi: 10.3390/ijms17010130.

[89] W. H. Koppenol, “The Haber-Weiss cycle - 70 years later,” *Redox Rep.*, vol. 6, no. 4, pp. 229–234, 2001, doi: 10.1179/135100001101536373.

[90] J. R. Chipperfield and C. Ratledge, “Salicylic acid is not a bacterial siderophore: A theoretical study,” *BioMetals*, vol. 13, no. 2, pp. 165–168, 2000, doi: 10.1023/A:1009227206890.

[91] Ruben Martin and Stephen L. Buchwald, “基因的改变NIH Public Access,” *Bone*, vol. 23, no. 1, pp. 1–7, 2008, doi: 10.1038/jid.2014.371.

[92] H. Haas, M. Eisendle, and B. G. Turgeon, “Siderophores in fungal physiology and virulence,” *Annu. Rev. Phytopathol.*, vol. 46, pp. 149–187, 2008, doi: 10.1146/annurev.phyto.45.062806.094338.

[93] C. Curie *et al.*, “Metal movement within the plant: Contribution of nicotianamine and yellow stripe 1-like transporters,” *Ann. Bot.*, vol. 103, no. 1, pp. 1–11, 2009, doi: 10.1093/aob/mcn207.

[94] M. Saha, S. Sarkar, B. Sarkar, B. K. Sharma, S. Bhattacharjee, and P. Tribedi, “Microbial siderophores and their potential applications: a review,” *Environ. Sci.*

Pollut. Res., vol. 23, no. 5, pp. 3984–3999, 2016, doi: 10.1007/s11356-015-4294-0.

[95] J. FRANCIS, H. M. MACTURK, J. MADINAVEITIA, and G. A. SNOW, “Mycobactin, a growth factor for Mycobacterium johnei. I. Isolation from Mycobacterium phlei.,” *Biochem. J.*, vol. 55, no. 4, pp. 596–607, 1953, doi: 10.1042/bj0550596.

[96] J. A. Garibaldi and J. B. Neilands, “Isolation and Properties of Ferrichrome A,” *J. Am. Chem. Soc.*, vol. 77, no. 9, pp. 2429–2430, 1955, doi: 10.1021/ja01614a021.

[97] C. Pidacks *et al.*, “Coprogen, the Isolation of a New Growth Factor Required by Pilobolus Species,” *J. Am. Chem. Soc.*, vol. 75, no. 23, pp. 6064–6065, 1953, doi: 10.1021/ja01119a529.

[98] J. A. GARIBALDI and J. B. NEILANDS, “Formation of Iron-binding Compounds by Micro-organisms,” *Nature*, vol. 177, no. 4507, pp. 526–527, 1956, doi: 10.1038/177526a0.

[99] T. Ito and J. B. Neilands, “Products of ‘Low-iron Fermentation’ with Bacillus subtilis: Isolation, Characterization and Synthesis of 2,3-Dihydroxybenzoylglycine,” *J. Am. Chem. Soc.*, vol. 80, no. 17, pp. 4645–4647, 1958, doi: 10.1021/ja01550a058.

[100] R. C. Hider and X. Kong, “Chemistry and biology of siderophores,” *Nat. Prod. Rep.*, vol. 27, no. 5, pp. 637–657, 2010, doi: 10.1039/b906679a.

[101] M. F. Kreutzer *et al.*, “Biosynthesis of a complex yersiniabactin-like natural product via the mic locus in phytopathogen Ralstonia solanacearum,” *Appl. Environ. Microbiol.*, vol. 77, no. 17, pp. 6117–6124, 2011, doi: 10.1128/AEM.05198-11.

[102] C. Kurth, H. Kage, and M. Nett, “Siderophores as molecular tools in medical and environmental applications,” *Org. Biomol. Chem.*, vol. 14, no. 35, pp. 8212–8227, 2016, doi: 10.1039/c6ob01400c.

[103] G. L. Challis, “A widely distributed bacterial pathway for siderophore biosynthesis independent of nonribosomal peptide synthetases,”

ChemBioChem, vol. 6, no. 4, pp. 601–611, 2005, doi: 10.1002/cbic.200400283.

[104] R. Saha, N. Saha, R. S. Donofrio, and L. L. Bestervelt, “Microbial siderophores: A mini review,” *J. Basic Microbiol.*, vol. 53, no. 4, pp. 303–317, 2013, doi: 10.1002/jobm.201100552.

[105] K. N. Raymond, E. A. Dertz, and S. S. Kim, “Enterobactin: An archetype for microbial iron transport,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 7, pp. 3584–3588, 2003, doi: 10.1073/pnas.0630018100.

[106] M. J. Smith, B. Schwyn, J. B. Neilands, J. N. Shoolery, and I. Holden, “Rhizobactin, a Structurally Novel Siderophore from Rhizobium meliloti,” *J. Am. Chem. Soc.*, vol. 107, no. 6, pp. 1739–1743, 1985, doi: 10.1021/ja00292a047.

[107] A. T. Taher, J. B. Porter, A. Kattamis, V. Viprakasit, and M. D. Cappellini, “Efficacy and safety of iron-chelation therapy with deferoxamine, deferiprone, and deferasirox for the treatment of iron-loaded patients with nontransfusion-dependent thalassemia syndromes,” *Drug Des. Devel. Ther.*, vol. 10, pp. 4073–4078, 2016, doi: 10.2147/DDDT.S117080.

[108] J. N. Feder *et al.*, “A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis,” *Nat. Genet.*, vol. 13, no. 4, pp. 399–408, 1996, doi: 10.1038/ng0896-399.

[109] R. M. Bannerman, S. T. Callender, and D. L. Williams, “Effect of desferrioxamine and D.T.P.A. in iron overload,” *Br. Med. J.*, vol. 2, no. 5319, pp. 1573–1577, 1962, doi: 10.1136/bmj.2.5319.1573.

[110] R. Golonka, B. S. Yeoh, and M. Vijay-Kumar, “The Iron Tug-of-War between Bacterial Siderophores and Innate Immunity,” *J. Innate Immun.*, vol. 11, no. 3, pp. 249–262, 2019, doi: 10.1159/000494627.

[111] A. Górska, A. Sloderbach, and M. P. Marszałł, “Siderophore-drug complexes: Potential medicinal applications of the ‘Trojan horse’ strategy,” *Trends Pharmacol. Sci.*, vol. 35, no. 9, pp. 442–449, 2014, doi: 10.1016/j.tips.2014.06.007.

[112] Y. Huang *et al.*, “Curb challenges of the ‘Trojan Horse’ approach: Smart strategies in achieving effective yet safe cell-penetrating peptide-based drug

delivery," *Adv. Drug Deliv. Rev.*, vol. 65, no. 10, pp. 1299–1315, 2013, doi: 10.1016/j.addr.2012.11.007.

[113] V. Braun, A. Pramanik, T. Gwinner, M. Köberle, and E. Bohn, "Sideromycins: Tools and antibiotics," *BioMetals*, vol. 22, no. 1, pp. 3–13, 2009, doi: 10.1007/s10534-008-9199-7.

[114] J. F. Fisher and M. J. M. Editors, *Antibacterials*, vol. II. .

[115] R. G. Wunderink *et al.*, "Cefiderocol versus high-dose, extended-infusion meropenem for the treatment of Gram-negative nosocomial pneumonia (APEKS-NP): a randomised, double-blind, phase 3, non-inferiority trial.," *Lancet. Infect. Dis.*, vol. 21, no. 2, pp. 213–225, Feb. 2021, doi: 10.1016/S1473-3099(20)30731-3.

[116] J. J. Choi and M. W. McCarthy, "Cefiderocol: a novel siderophore cephalosporin," *Expert Opin. Investig. Drugs*, vol. 27, no. 2, pp. 193–197, 2018, doi: 10.1080/13543784.2018.1426745.

[117] G. G. Zhanel *et al.*, "Cefiderocol: A Siderophore Cephalosporin with Activity Against Carbapenem-Resistant and Multidrug-Resistant Gram-Negative Bacilli," *Drugs*, vol. 79, no. 3, pp. 271–289, 2019, doi: 10.1007/s40265-019-1055-2.

[118] K. H. Negash, J. K. S. Norris, and J. T. Hodgkinson, "Siderophore-Antibiotic Conjugate Design: New Drugs for Bad Bugs?," *Molecules*, vol. 24, no. 18, Sep. 2019, doi: 10.3390/molecules24183314.

[119] R. Liu, P. A. Miller, S. B. Vakulenko, N. K. Stewart, W. C. Boggess, and M. J. Miller, "A Synthetic Dual Drug Sideromycin Induces Gram-Negative Bacteria to Commit Suicide with a Gram-Positive Antibiotic," *J. Med. Chem.*, vol. 61, no. 9, pp. 3845–3854, 2018, doi: 10.1021/acs.jmedchem.8b00218.

[120] T. A. Wencewicz, T. E. Long, U. Möllmann, and M. J. Miller, "Trihydroxamate siderophore-fluoroquinolone conjugates are selective sideromycin antibiotics that target staphylococcus aureus," *Bioconjug. Chem.*, vol. 24, no. 3, pp. 473–486, 2013, doi: 10.1021/bc300610f.

[121] G. Deng and L. Jin, "The effects of vasoactive intestinal peptide in

neurodegenerative disorders," *Neurol. Res.*, vol. 39, no. 1, pp. 65–72, Jan. 2017, doi: 10.1080/01616412.2016.1250458.

[122] B. Ottesen and J. Fahrenkrug, "Vasoactive intestinal polypeptide and other preprovasoactive intestinal polypeptide-derived peptides in the female and male genital tract: Localization, biosynthesis, and functional and clinical significance," *Am. J. Obstet. Gynecol.*, vol. 172, no. 5, pp. 1615–1631, May 1995, doi: 10.1016/0002-9378(95)90506-5.

[123] T. Van Loy, H. P. Vandersmissen, J. Poels, M. B. Van Hiel, H. Verlinden, and J. Vanden Broeck, "Tachykinin-related peptides and their receptors in invertebrates: A current view," *Peptides*, vol. 31, no. 3, pp. 520–524, 2010, doi: https://doi.org/10.1016/j.peptides.2009.09.023.

[124] H. K. Kang, C. Kim, C. H. Seo, and Y. Park, "The therapeutic applications of antimicrobial peptides (AMPs): a patent review," *J. Microbiol.*, vol. 55, no. 1, pp. 1–12, 2017, doi: 10.1007/s12275-017-6452-1.

[125] B. Mach, E. Reich, and E. L. Tatum, "Separation of the Biosynthesis of the Antibiotic Polypeptide Tyrocidine From Protein Biosynthesis," *Proc. Natl. Acad. Sci.*, vol. 50, no. 1, pp. 175–181, 1963, doi: 10.1073/pnas.50.1.175.

[126] R. D. Süssmuth and A. Mainz, "Nonribosomal Peptide Synthesis—Principles and Prospects," *Angew. Chemie - Int. Ed.*, vol. 56, no. 14, pp. 3770–3821, 2017, doi: 10.1002/anie.201609079.

[127] R. Finking and M. A. Marahiel, "Biosynthesis of nonribosomal peptides," *Annu. Rev. Microbiol.*, vol. 58, pp. 453–488, 2004, doi: 10.1146/annurev.micro.58.030603.123615.

[128] G. L. Challis, J. Ravel, and C. A. Townsend, "Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains," *Chem. Biol.*, vol. 7, no. 3, pp. 211–224, 2000, doi: 10.1016/S1074-5521(00)00091-0.

[129] T. Stachelhaus, H. D. Mootz, and M. A. Marahiel, "The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases," *Chem. Biol.*, vol. 6, no. 8, pp. 493–505, 1999, doi: 10.1016/S1074-5521(99)80082-9.

[130] M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch, and O. Kohlbacher, “NRPSpredictor2 - A web server for predicting NRPS adenylation domain specificity,” *Nucleic Acids Res.*, vol. 39, no. SUPPL. 2, pp. 362–367, 2011, doi: 10.1093/nar/gkr323.

[131] C. Rausch, T. Weber, O. Kohlbacher, W. Wohlleben, and D. H. Huson, “Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs),” *Nucleic Acids Res.*, vol. 33, no. 18, pp. 5799–5808, 2005, doi: 10.1093/nar/gki885.

[132] S. Meyer *et al.*, “Biochemical Dissection of the Natural Diversification of Microcystin Provides Lessons for Synthetic Biology of NRPS,” *Cell Chem. Biol.*, vol. 23, no. 4, pp. 462–471, Apr. 2016, doi: 10.1016/j.chembiol.2016.03.011.

[133] C. JN, “Derivates of the multiple ketene group,” *Proc. Chem. Soc*, no. 23, pp. 230–21, 1907.

[134] C. JN, “Derivatives of multiple keten group,” *J. Chem. Soc.*, vol. 91, pp. 1806–1813, 1907.

[135] D. E. Cane, “ Introduction: Polyketide and Nonribosomal Polypeptide Biosynthesis. From Collie to Coli ,” *Chem. Rev.*, vol. 97, no. 7, pp. 2463–2464, 1997, doi: 10.1021/cr970097g.

[136] R. Bentley and J. W. Bennett, “From Collie to Combinatorial Biosynthesis,” *Annu. Rev. Microbiol.*, vol. 53, pp. 411–446, 1999.

[137] M. A. Fischbach and C. T. Walsh, “Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic machinery, and mechanisms,” *Chem. Rev.*, vol. 106, no. 8, pp. 3468–3496, 2006, doi: 10.1021/cr0503097.

[138] C. Hertweck, “The biosynthetic logic of polyketide diversity,” *Angew. Chemie - Int. Ed.*, vol. 48, no. 26, pp. 4688–4716, 2009, doi: 10.1002/anie.200806121.

[139] A. Nivina, K. P. Yuet, J. Hsu, and C. Khosla, “Evolution and Diversity of Assembly-Line Polyketide Synthases,” *Chem. Rev.*, vol. 119, no. 24, pp. 12524–12547, 2019, doi: 10.1021/acs.chemrev.9b00525.

[140] H. Chen and L. Du, “Iterative polyketide biosynthesis by modular polyketide

synthases in bacteria," *Appl. Microbiol. Biotechnol.*, vol. 100, no. 2, pp. 541–557, 2016, doi: 10.1007/s00253-015-7093-0.

[141] C. Khosla, D. Herschlag, D. E. Cane, and C. T. Walsh, "Assembly line polyketide synthases: Mechanistic insights and unsolved problems," *Biochemistry*, vol. 53, no. 18, pp. 2875–2883, 2014, doi: 10.1021/bi500290t.

[142] C. Hertweck, A. Luzhetskyy, Y. Rebets, and A. Bechthold, "Type II polyketide synthases: Gaining a deeper insight into enzymatic teamwork," *Nat. Prod. Rep.*, vol. 24, no. 1, pp. 162–190, 2007, doi: 10.1039/b507395m.

[143] C. Bisang *et al.*, "A chain initiation factor common to both modular and aromatic polyketide synthases," *Nature*, vol. 401, no. 6752, pp. 502–505, 1999, doi: 10.1038/46829.

[144] J. Staunton and K. J. Weissman, "Polyketide biosynthesis: A millennium review," *Nat. Prod. Rep.*, vol. 18, no. 4, pp. 380–416, 2001, doi: 10.1039/a909079g.

[145] D. Yu, F. Xu, J. Zeng, and J. Zhan, "Type III polyketide synthases in natural product biosynthesis," *IUBMB Life*, vol. 64, no. 4, pp. 285–295, 2012, doi: 10.1002/iub.1005.

[146] G. Bringmann *et al.*, "Different polyketide folding modes converge to an identical molecular architecture," *Nat. Chem. Biol.*, vol. 2, no. 8, pp. 429–433, 2006, doi: 10.1038/nchembio805.

[147] M. B. Austin and J. P. Noel, "The chalcone synthase superfamily of type III polyketide synthases," *Nat. Prod. Rep.*, vol. 20, no. 1, pp. 79–110, 2003, doi: 10.1039/b100917f.

[148] E. Stackebrandt, R. G. E. Murray, and H. G. Truper, "Proteobacteria classis nov., a name for the phylogenetic taxon that includes the 'purple bacteria and their relatives,'" *Int. J. Syst. Bacteriol.*, vol. 38, no. 3, pp. 321–325, 1988, doi: 10.1099/00207713-38-3-321.

[149] C. D. Moon, W. Young, P. H. Maclean, A. L. Cookson, and E. N. Bermingham, "Metagenomic insights into the roles of Proteobacteria in the gastrointestinal microbiomes of healthy dogs and cats," *Microbiologyopen*, vol. 7, no. 5, pp. 1–

20, 2018, doi: 10.1002/mbo3.677.

[150] P. Bradley, S. Nayfach, and K. Pollard, *Phylogeny-corrected identification of microbial gene families relevant to human gut colonization*. 2017.

[151] H. Gross and J. E. Loper, “Genomics of secondary metabolite production by Pseudomonas spp.,” *Nat. Prod. Rep.*, vol. 26, no. 11, pp. 1408–1446, 2009, doi: 10.1039/B817075B.

[152] S. M. Ringel *et al.*, “Ambruticin (W7783), a new antifungal antibiotic,” *J. Antibiot. (Tokyo).*, vol. 30, no. 5, pp. 371–375, 1977, doi: 10.7164/antibiotics.30.371.

[153] X. Liu and Y. Q. Cheng, “Genome-guided discovery of diverse natural products from Burkholderia sp.,” *J. Ind. Microbiol. Biotechnol.*, vol. 41, no. 2, pp. 275–284, 2014, doi: 10.1007/s10295-013-1376-1.

[154] S.-L. Sun, W.-L. Yang, W.-W. Fang, Y.-X. Zhao, L. Guo, and Y.-J. Dai, “The Plant Growth-Promoting Rhizobacterium Variovorax boronicumulans CGMCC 4969 Regulates the Level of Indole-3-Acetic Acid Synthesized from Indole-3-Acet,” *Appl. Environ. Microbiol.*, vol. 84, no. 16, pp. e00298-18, Aug. 2018, doi: 10.1128/AEM.00298-18.

[155] C. Kurth, S. Schieferdecker, K. Athanasopoulou, I. Seccareccia, and M. Nett, “Variochelins, Lipopeptide Siderophores from Variovorax boronicumulans Discovered by Genome Mining,” *J. Nat. Prod.*, vol. 79, no. 4, pp. 865–872, 2016, doi: 10.1021/acs.jnatprod.5b00932.

[156] F. Baldeweg, D. Hoffmeister, and M. Nett, “A genomics perspective on natural product biosynthesis in plant pathogenic bacteria,” *Nat. Prod. Rep.*, vol. 36, no. 2, pp. 307–325, 2019, doi: 10.1039/c8np00025e.

[157] B. Wackler *et al.*, “Ralfuranone biosynthesis in ralstonia solanacearum suggests functional divergence in the quinone synthetase family of enzymes,” *Chem. Biol.*, vol. 18, no. 3, pp. 354–360, 2011, doi: 10.1016/j.chembiol.2011.01.010.

[158] G. Bhatt and T. P. Denny, “Ralstonia solanacearum iron scavenging by the

siderophore staphyloferrin B is controlled by PhcA, the global virulence regulator," *J. Bacteriol.*, vol. 186, no. 23, pp. 7896–7904, 2004, doi: 10.1128/JB.186.23.7896-7904.2004.

[159] F. Baldeweg, H. Kage, S. Schieferdecker, C. Allen, D. Hoffmeister, and M. Nett, "Structure of Ralsolamycin, the Interkingdom Morphogen from the Crop Plant Pathogen Ralstonia solanacearum," *Org. Lett.*, vol. 19, no. 18, pp. 4868–4871, 2017, doi: 10.1021/acs.orglett.7b02329.

[160] B. La Scola, R. J. Birtles, M. N. Mallet, and D. Raoult, "Massilia timonae gen. nov., sp. nov., isolated from blood of an immunocompromised patient with cerebellar lesions," *J. Clin. Microbiol.*, vol. 36, no. 10, pp. 2847–2852, 1998, doi: 10.1128/jcm.36.10.2847-2852.1998.

[161] V. Gallego, C. Sánchez-Porro, M. T. García, and A. Ventosa, "Massilia aurea sp. nov., isolated from drinking water," *Int. J. Syst. Evol. Microbiol.*, vol. 56, no. 10, pp. 2449–2453, 2006, doi: 10.1099/ijs.0.64389-0.

[162] M. Ofek, Y. Hadar, and D. Minz, "Ecology of root colonizing Massilia (Oxalobacteraceae)," *PLoS One*, vol. 7, no. 7, 2012, doi: 10.1371/journal.pone.0040117.

[163] R. A. I. Abou-Shanab, P. van Berkum, and J. S. Angle, "Heavy metal resistance and genotypic analysis of metal resistance genes in gram-positive and gram-negative bacteria present in Ni-rich serpentine soil and in the rhizosphere of Alyssum murale," *Chemosphere*, vol. 68, no. 2, pp. 360–367, 2007, doi: 10.1016/j.chemosphere.2006.12.051.

[164] A. B. Dohrmann and C. C. Tebbe, "Effect of elevated tropospheric ozone on the structure of bacterial communities inhabiting the rhizosphere of herbaceous plants native to Germany," *Appl. Environ. Microbiol.*, vol. 71, no. 12, pp. 7750–7758, 2005, doi: 10.1128/AEM.71.12.7750-7758.2005.

[165] M. Kuffner *et al.*, "Culturable bacteria from Zn- and Cd-accumulating Salix caprea with differential effects on plant growth and heavy metal availability," *J. Appl. Microbiol.*, vol. 108, no. 4, pp. 1471–1484, 2010, doi: 10.1111/j.1365-2672.2010.04670.x.

[166] J. Pakarinen *et al.*, “Predominance of Gram-positive bacteria in house dust in the low-allergy risk Russian Karelia,” *Environ. Microbiol.*, vol. 10, no. 12, pp. 3317–3325, 2008, doi: 10.1111/j.1462-2920.2008.01723.x.

[167] H. Wang, X. Zhang, S. Wang, B. Zhao, K. Lou, and X. H. Xing, “Massilia violaceinigra sp. Nov., a novel purple-pigmented bacterium isolated from glacier permafrost,” *Int. J. Syst. Evol. Microbiol.*, vol. 68, no. 7, pp. 2271–2278, 2018, doi: 10.1099/ijsem.0.002826.

[168] M. Ren, X. Li, Y. Zhang, Y. Jin, S. Li, and H. Huang, “Massilia armeniaca sp. Nov., isolated from desert soil,” *Int. J. Syst. Evol. Microbiol.*, vol. 68, no. 7, pp. 2319–2324, 2018, doi: 10.1099/ijsem.0.002836.

[169] H. Miess, A. Frediansyah, M. Göker, and H. Gross, “ Draft Genome Sequences of Six Type Strains of the Genus Massilia ,” *Microbiol. Resour. Announc.*, vol. 9, no. 18, pp. 3–5, 2020, doi: 10.1128/mra.00226-20.

[170] H. Agematu, K. Suzuki, and H. Tsuya, “Massilia sp. BS-1, a novel violacein-producing bacterium isolated from soil,” *Biosci. Biotechnol. Biochem.*, vol. 75, no. 10, pp. 2008–2010, 2011, doi: 10.1271/bbb.100729.

[171] C. D’Angelo-Picard, D. Faure, I. Penot, and Y. Dessaux, “Diversity of N-acyl homoserine lactone-producing and -degrading bacteria in soil and tobacco rhizosphere,” *Environ. Microbiol.*, vol. 7, no. 11, pp. 1796–1808, 2005, doi: 10.1111/j.1462-2920.2005.00886.x.

[172] D. Ley, “Acidovorax , a New Genus for Pseudomonas facilis , Pseudomonas Several Clinical Isolates , with the Species Acidovorax,” *Most*, vol. 40, no. 4, pp. 384–398, 1990.

[173] J. H. Choi, M. S. Kim, S. W. Roh, and J. W. Bae, “Acidovorax soli sp. nov., isolated from landfill soil,” *Int. J. Syst. Evol. Microbiol.*, vol. 60, no. 12, pp. 2715–2718, 2010, doi: 10.1099/ijs.0.019661-0.

[174] K. Heylen, L. Lebbe, and P. de Vos, “Acidovorax caeni sp. nov., a denitrifying species with genetically diverse isolates from activated sludge,” *Int. J. Syst. Evol. Microbiol.*, vol. 58, no. 1, pp. 73–77, 2008, doi: 10.1099/ijs.0.65387-0.

[175] L. Gardan, C. Dauga, P. Prior, M. Gillis, and G. S. Saddler, “Acidovorax anthurii

sp. nov., a new phytopathogenic bacterium which causes bacterial leaf-spot of anthurium," *Int. J. Syst. Evol. Microbiol.*, vol. 50, no. 1, pp. 235–246, 2000, doi: 10.1099/00207713-50-1-235.

[176] D. Pal, N. Kaur, S. K. Sudan, B. Bisht, S. Krishnamurthi, and S. Mayilraj, "Acidovorax kalamii sp. Nov., isolated from a water sample of the river Ganges," *Int. J. Syst. Evol. Microbiol.*, vol. 68, no. 5, pp. 1719–1724, 2018, doi: 10.1099/ijsem.0.002736.

[177] M. Vaneechoutte, M. J. Janssens, V. Avesani, M. Delmée, and P. Deschaght, "Description of Acidovorax wautersii sp. nov. to accommodate clinical isolates and an environmental isolate, most closely related to Acidovorax avenae," *Int. J. Syst. Evol. Microbiol.*, vol. 63, no. PART6, pp. 2203–2206, 2013, doi: 10.1099/ijs.0.046102-0.

[178] N. W. Schaad *et al.*, "Reclassification of subspecies of Acidovorax avenae as A. Avenae (Manns 1905) emend., A. cattleyae (Pavarino, 1911) comb. nov., A. citrulli Schaad et al., 1978) comb. nov., and proposal of A. oryzae sp. nov.," *Syst. Appl. Microbiol.*, vol. 31, no. 6–8, pp. 434–446, 2008, doi: 10.1016/j.syapm.2008.09.003.

[179] I. Jiménez-Guerrero *et al.*, "Show me your secret(ed) weapons: a multifaceted approach reveals a wide arsenal of type III-secreted effectors in the cucurbit pathogenic bacterium Acidovorax citrulli and novel effectors in the Acidovorax genus," *Mol. Plant Pathol.*, vol. 21, no. 1, pp. 17–37, 2020, doi: 10.1111/mpp.12877.

[180] R. R. Walcott, A. Fessehaie, and A. C. Castro, "Differences in pathogenicity between two genetically distinct groups of Acidovorax avenae subsp. citrulli on cucurbit hosts," *J. Phytopathol.*, vol. 152, no. 5, pp. 277–285, 2004, doi: 10.1111/j.1439-0434.2004.00841.x.

[181] S. Burdman and R. Walcott, "Acidovorax citrulli: Generating basic and applied knowledge to tackle a global threat to the cucurbit industry," *Mol. Plant Pathol.*, vol. 13, no. 8, pp. 805–815, 2012, doi: 10.1111/j.1364-3703.2012.00810.x.

[182] A. A. A. Husni, S. I. Ismail, N. Jaafar, and D. Zulperi, "Etiology, diagnostic approaches and management strategies of Acidovorax citrulli, a bacterial fruit

blotch pathogen of cucurbits," *Plant Prot. Sci.*, vol. 57, no. 2, pp. 75–94, 2021, doi: 10.17221/52/2020-PPS.

[183] N. Eckshtain-Levi *et al.*, "Comparative analysis of type III secreted effector genes reflects divergence of Acidovorax citrulli strains into three distinct lineages," *Phytopathology*, vol. 104, no. 11, pp. 1152–1162, 2014, doi: 10.1094/PHYTO-12-13-0350-R.

[184] R. Yang *et al.*, "Complete assembly of the genome of an Acidovorax citrulli strain reveals a naturally occurring plasmid in this species," *Front. Microbiol.*, vol. 10, no. JUN, pp. 1–17, 2019, doi: 10.3389/fmicb.2019.01400.

[185] S. Burdman, N. Kots, G. Kritzman, and J. Kopelowitz, "Molecular, physiological, and host-range characterization of Acidovorax avenae subsp. citrulli isolates from watermelon and melon in Israel," *Plant Dis.*, vol. 89, no. 12, pp. 1339–1347, 2005, doi: 10.1094/PD-89-1339.

[186] J. Liu *et al.*, "Ferric Uptake Regulator (FurA) is Required for Acidovorax citrulli Virulence on Watermelon," *Phytopathology*, vol. 109, no. 12, pp. 1997–2008, 2019, doi: 10.1094/PHYTO-05-19-0172-R.

[187] C. K. Tanui, D. Y. Shyntum, S. L. Priem, J. Theron, and L. N. Moleleki, "Influence of the ferric uptake regulator (Fur) protein on pathogenicity in Pectobacterium carotovorum subsp. brasiliense," *PLoS One*, vol. 12, no. 5, pp. 1–21, 2017, doi: 10.1371/journal.pone.0177647.

[188] H. Liu *et al.*, "Functional analysis of the ferric uptake regulator gene fur in Xanthomonas vesicatoria," *PLoS One*, vol. 11, no. 2, pp. 1–13, 2016, doi: 10.1371/journal.pone.0149280.

[189] J. Masschelein, M. Jenner, and G. L. Challis, "Antibiotics from Gram-negative bacteria: A comprehensive overview and selected biosynthetic highlights," *Nat. Prod. Rep.*, vol. 34, no. 7, pp. 712–783, 2017, doi: 10.1039/c7np00010c.

[190] K. Blin *et al.*, "AntiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline," *Nucleic Acids Res.*, vol. 47, no. W1, pp. W81–W87, 2019, doi: 10.1093/nar/gkz310.

[191] K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers,

"GenBank," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D67–D72, 2016, doi: 10.1093/nar/gkv1276.

[192] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *J. Mol. Biol.*, vol. 215, no. 3, pp. 403–410, 1990, doi: https://doi.org/10.1016/S0022-2836(05)80360-2.

[193] V. K. Dua, G. Verma, D. D. Agarwal, M. Kaiser, and R. Brun, "Antiprotozoal activities of traditional medicinal plants from the Garhwal region of North West Himalaya, India," *J. Ethnopharmacol.*, vol. 136, no. 1, pp. 123–128, 2011, doi: 10.1016/j.jep.2011.04.024.

[194] T. Baltz, D. Baltz, C. Giroud, and J. Crockett, "Cultivation in a semi-defined medium of animal infective forms of Trypanosoma brucei, T. equiperdum, T. evansi, T. rhodesiense and T. gambiense.," *EMBO J.*, vol. 4, no. 5, pp. 1273–1277, 1985, doi: 10.1002/j.1460-2075.1985.tb03772.x.

[195] B. Räz, M. Iten, Y. Grether-Bühler, R. Kaminsky, and R. Brun, "The Alamar Blue® assay to determine drug sensitivity of African trypanosomes (T.b. rhodesiense and T.b. gambiense) in vitro," *Acta Trop.*, vol. 68, no. 2, pp. 139–147, 1997, doi: https://doi.org/10.1016/S0001-706X(97)00079-X.

[196] W. Huber and J. C. Koella, "A comparison of three methods of estimating EC50 in studies of drug resistance of malaria parasites.," *Acta Trop.*, vol. 55, no. 4, pp. 257–261, Dec. 1993, doi: 10.1016/0001-706x(93)90083-n.

[197] F. S. Buckner, C. L. Verlinde, A. C. La Flamme, and W. C. Van Voorhis, "Efficient technique for screening drugs for activity against Trypanosoma cruzi using parasites expressing beta-galactosidase.," *Antimicrob. Agents Chemother.*, vol. 40, no. 11, pp. 2592–2597, Nov. 1996, doi: 10.1128/AAC.40.11.2592.

[198] I. Cunningham, "New culture medium for maintenance of tsetse tissues and growth of trypanosomatids.," *J. Protozool.*, vol. 24, no. 2, pp. 325–329, May 1977, doi: 10.1111/j.1550-7408.1977.tb00987.x.

[199] R. E. Desjardins, C. J. Canfield, J. D. Haynes, and J. D. Chulay, "Quantitative Activity Semiautomated Technique," *Antimicrob. Agents Chemother.*, vol. 16,

no. 6, pp. 710–718, 1979.

[200] T. Ponnudurai, A. D. Leeuwenberg, and J. H. Meuwissen, “Chloroquine sensitivity of isolates of Plasmodium falciparum adapted to in vitro culture.,” *Trop. Geogr. Med.*, vol. 33, no. 1, pp. 50–54, Mar. 1981.

[201] B. Page, M. Page, and C. Noel, “A new fluorometric assay for cytotoxicity measurements in-vitro.,” *Int. J. Oncol.*, vol. 3, no. 3, pp. 473–476, Sep. 1993.

[202] S. A. Ahmed, R. M. J. Gogal, and J. E. Walsh, “A new rapid and simple non-radioactive assay to monitor and determine the proliferation of lymphocytes: an alternative to [3H]thymidine incorporation assay.,” *J. Immunol. Methods*, vol. 170, no. 2, pp. 211–224, Apr. 1994, doi: 10.1016/0022-1759(94)90396-4.

[203] N. R. Myeong, H. J. Seong, H. J. Kim, and W. J. Sul, “Complete genome sequence of antibiotic and anticancer agent violacein producing Massilia sp. strain NR 4-1,” *J. Biotechnol.*, vol. 223, no. 4, pp. 36–37, 2016, doi: 10.1016/j.jbiotec.2016.02.027.

[204] “No Title.” https://www.ncbi.nlm.nih.gov/genome/browse#!/prokaryotes/15130/ (accessed Mar. 17, 2021).

[205] H. Miess, A. Frediansyah, M. Göker, and H. Gross, “Draft Genome Sequences of Six Type Strains of the Genus Massilia,” *Microbiol. Resour. Announc.*, vol. 9, no. 18, pp. 3–5, 2020, doi: 10.1128/mra.00226-20.

[206] H. Kage, M. F. Kreutzer, B. Wackler, D. Hoffmeister, and M. Nett, “An iterative type i polyketide synthase initiates the biosynthesis of the antimycoplasma agent micacocidin,” *Chem. Biol.*, vol. 20, no. 6, pp. 764–771, 2013, doi: 10.1016/j.chembiol.2013.04.010.

[207] L. M. Cañedo, J. A. De la Fuente, C. Gesto, M. J. Ferreiro, C. Jiménez, and R. Riguera, “Agrochelin, a new cytotoxic alkaloid from the marine bacteria Agrobacterium sp.,” *Tetrahedron Lett.*, vol. 40, no. 37, pp. 6841–6844, 1999, doi: 10.1016/S0040-4039(99)01377-5.

[208] T. Adilakshmi, P. D. Ayling, and C. Ratledge, “Mutational analysis of a role for salicylic acid in iron metabolism of Mycobacterium smegmatis,” *J. Bacteriol.*, vol. 182, no. 2, pp. 264–271, 2000, doi: 10.1128/JB.182.2.264-271.2000.

[209] H. Drechsel *et al.*, “Structure elucidation of yersiniabactin, a siderophore from highly virulent Yersinia strains,” *Liebigs Ann.*, vol. 1995, no. 10, pp. 1727–1733, 1995, doi: 10.1002/jlac.1995199510243.

[210] C. D. Cox, K. L. Rinehart, M. L. Moore, and J. C. Cook, “Pyochelin: novel structure of an iron-chelating growth promoter for Pseudomonas aeruginosa.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 78, no. 7, pp. 4256–4260, 1981, doi: 10.1073/pnas.78.7.4256.

[211] A. Ino, Y. Hasegawa, and A. Murabayashi, “Synthetic studies of thiazoline and thiazolidine-containing natural products - 2. Total synthesis of the antimycoplasma antibiotic micacocidin,” *Tetrahedron*, vol. 55, no. 34, pp. 10283–10294, 1999, doi: 10.1016/S0040-4020(99)00583-9.

[212] K. L. Rinehart, A. L. Staley, S. R. Wilson, R. G. Ankenbauer, and C. D. Cox, “Stereochemical Assignment of the Pyochelins,” *J. Org. Chem.*, vol. 60, no. 9, pp. 2786–2791, 1995, doi: 10.1021/jo00114a029.

[213] K. Schlegel, K. Taraz, and H. Budzikiewicz, “The stereoisomers of pyochelin, a siderophore of Pseudomonas aeruginosa,” *BioMetals*, vol. 17, no. 4, pp. 409–414, 2004, doi: 10.1023/B:BIOM.0000029437.42633.73.

[214] A. Souto *et al.*, “Structure and biosynthetic assembly of piscibactin, a siderophore from photobacterium damselae subsp. piscicida, predicted from genome analysis,” *European J. Org. Chem.*, no. 29, pp. 5693–5700, 2012, doi: 10.1002/ejoc.201200818.

[215] A. Ross-Gillespie, M. Weigert, S. P. Brown, and R. Kümmerli, “Gallium-mediated siderophore quenching as an evolutionarily robust antibacterial treatment,” *Evol. Med. Public Heal.*, vol. 2014, no. 1, pp. 18–29, 2014, doi: 10.1093/emph/eou003.

[216] N. Kircheva and T. Dudev, “Gallium as an Antibacterial Agent: A DFT/SMD Study of the Ga3+/Fe3+ Competition for Binding Bacterial Siderophores,” *Inorg. Chem.*, vol. 59, no. 9, pp. 6242–6254, 2020, doi: 10.1021/acs.inorgchem.0c00367.

[217] S. Y. Choi, K. H. Yoon, J. Il Lee, and R. J. Mitchell, “Violacein: Properties and

production of a versatile bacterial pigment," *Biomed Res. Int.*, vol. 2015, 2015, doi: 10.1155/2015/465056.

[218] A. L. Rodrigues *et al.*, "Systems metabolic engineering of Escherichia coli for production of the antitumor drugs violacein and deoxyviolacein," *Metab. Eng.*, vol. 20, pp. 29–41, 2013, doi: 10.1016/j.ymben.2013.08.004.

[219] V. K. Nathan and M. E. Rani, "Current Research in Microbiology Surface Culturing of Chromobacterium Production and Prospecting its Bio-Activities," no. September 2019, 2018.

[220] S. Y. Choi, K. H. Yoon, J. Il Lee, and R. J. Mitchell, "Violacein: Properties and production of a versatile bacterial pigment," *Biomed Res. Int.*, vol. 2015, 2015, doi: 10.1155/2015/465056.

[221] C. A. Aruldass, Rubiyatno, C. K. Venil, and W. A. Ahmad, "Violet pigment production from liquid pineapple waste by Chromobacterium violaceum UTM5 and evaluation of its bioactivity," *RSC Adv.*, vol. 5, no. 64, pp. 51524–51536, 2015, doi: 10.1039/c5ra05765e.

[222] T. P. J. W. G. A. Bedford G.R., "15N NMR studies of guanidines," *Magn. Reson. Chem.*, vol. 33, no. October 1994, pp. 383–388, 1995.

[223] H. M. Patel, J. Tao, and C. T. Walsh, "Epimerization of an L-cysteinyl to a D-cysteinyl residue during thiazoline ring formation in siderophore chain elongation by pyochelin synthetase from Pseudomonas aeruginosa," *Biochemistry*, vol. 42, no. 35, pp. 10514–10527, 2003, doi: 10.1021/bi034840c.

[224] L. E. N. Quadri, T. A. Keating, H. M. Patel, and C. T. Walsh, "Assembly of the Pseudomonas aeruginosa nonribosomal peptide siderophore pyochelin: In vitro reconstitution of aryl-4,2-bisthiazoline synthetase activity from PchD, PchE, and PchF," *Biochemistry*, vol. 38, no. 45, pp. 14941–14954, Nov. 1999, doi: 10.1021/bi991787c.

[225] A. M. Gehring, I. I. Mori, R. D. Perry, and C. T. Walsh, "The nonribosomal peptide synthetase HMWP2 forms a thiazoline ring during biogenesis of yersiniabactin, an iron-chelating virulence factor of yersinia pestis.," *Biochemistry*, vol. 37, no. 48, p. 17104, Dec. 1998, doi: 10.1021/bi9850524.

[226] D. A. Miller, L. Luo, N. Hillson, T. A. Keating, and C. T. Walsh, “Yersiniabactin synthetase: A four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of Yersinia pestis,” *Chem. Biol.*, vol. 9, no. 3, pp. 333–344, 2002, doi: 10.1016/S1074-5521(02)00115-1.

[227] Z. A. Youard, G. L. A. Mislin, P. A. Majcherczyk, I. J. Schalk, and C. Reimmann, “Pseudomonas fluorescens CHA0 produces enantio-pyochelin, the optical antipode of the pseudomonas aeruginosa siderophore pyochelin,” *J. Biol. Chem.*, vol. 282, no. 49, pp. 35546–35553, 2007, doi: 10.1074/jbc.M707039200.

[228] C. Khosla, R. S. Gokhale, J. R. Jacobsen, and D. E. Cane, “T Olerance and S Pecificity,” pp. 219–253, 1999.

[229] P. Caffrey, “Conserved Amino Acid Residues Correlating With Ketoreductase Stereospecificity in Modular Polyketide Synthases,” pp. 654–657, 2003, doi: 10.1002/cbic.200300581.

[230] P. Caffrey, “Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases,” *ChemBioChem*, vol. 4, no. 7, pp. 654–657, 2003, doi: 10.1002/cbic.200300581.

[231] C. Reimmann, H. M. Patel, L. Serino, M. Barone, C. T. Walsh, and D. Haas, “Essential pchG-dependent reduction in pyochelin biosynthesis of Pseudomonas aeruginosa,” *J. Bacteriol.*, vol. 183, no. 3, pp. 813–820, 2001, doi: 10.1128/JB.183.3.813-820.2001.

[232] S. M. Morris, “Arginine metabolism revisited,” *J. Nutr.*, vol. 146, no. 12, pp. 2579S-2586S, 2016, doi: 10.3945/jn.115.226621.

[233] M. B. Austin *et al.*, “Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediates,” *J. Biol. Chem.*, vol. 279, no. 43, pp. 45162–45174, 2004, doi: 10.1074/jbc.M406567200.

[234] H. Hayashi, “Pyridoxal enzymes: Mechanistic diversity and uniformity,” *J. Biochem.*, vol. 118, no. 3, pp. 463–473, 1995, doi: 10.1093/oxfordjournals.jbchem.a124931.

[235] A. C. Eliot and J. F. Kirsch, “Pyridoxal phosphate enzymes: Mechanistic,

structural, and evolutionary considerations," *Annu. Rev. Biochem.*, vol. 73, pp. 383–415, 2004, doi: 10.1146/annurev.biochem.73.011303.074021.

[236] B. A. Yard *et al.*, "The Structure of Serine Palmitoyltransferase; Gateway to Sphingolipid Biosynthesis," *J. Mol. Biol.*, vol. 370, no. 5, pp. 870–886, 2007, doi: 10.1016/j.jmb.2007.04.086.

[237] H. Ikushiro and H. Hayashi, "Mechanistic enzymology of serine palmitoyltransferase," *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1814, no. 11, pp. 1474–1480, 2011, doi: 10.1016/j.bbapap.2011.02.005.

[238] J. V. K. Jensen and V. F. Wendisch, "Ornithine cyclodeaminase-based proline production by Corynebacterium glutamicum," *Microb. Cell Fact.*, vol. 12, no. 1, pp. 1–10, 2013, doi: 10.1186/1475-2859-12-63.

[239] U. Schindler, N. Sans, and J. Schröder, "Ornithine cyclodeaminase from octopine Ti plasmid Ach5: identification, DNA sequence, enzyme properties, and comparison with gene and enzyme from nopaline Ti plasmid C58.," *J. Bacteriol.*, vol. 171, no. 2, pp. 847–854, 1989, doi: 10.1128/jb.171.2.847-854.1989.

[240] T. Sommer *et al.*, "A fundamental catalytic difference between zinc and manganese dependent enzymes revealed in a bacterial isatin hydrolase," *Sci. Rep.*, vol. 8, no. 1, pp. 1–11, 2018, doi: 10.1038/s41598-018-31259-y.

[241] K. Bjerregaard-Andersen, T. Sommer, J. K. Jensen, B. Jochimsen, M. Etzerodt, and J. P. Morth, "A proton wire and water channel revealed in the crystal structure of isatin hydrolase," *J. Biol. Chem.*, vol. 289, no. 31, pp. 21351–21359, 2014, doi: 10.1074/jbc.M114.568824.

[242] K. L. Kavanagh, H. Jörnvall, B. Persson, and U. Oppermann, "Medium- and short-chain dehydrogenase/reductase gene and protein families: The SDR superfamily: Functional and structural diversity within a family of metabolic and regulatory enzymes," *Cell. Mol. Life Sci.*, vol. 65, no. 24, pp. 3895–3906, 2008, doi: 10.1007/s00018-008-8588-y.

[243] A. Li and J. Piel, "A gene cluster from a marine Streptomyces encoding the biosynthesis of the aromatic spiroketal polyketide griseorhodin A," *Chem. Biol.*,

vol. 9, no. 9, pp. 1017–1026, 2002, doi: 10.1016/S1074-5521(02)00223-5.

[244] X. Xie, A. Garg, A. T. Keatinge-Clay, C. Khosla, and D. E. Cane, “Epimerase and Reductase Activities of Polyketide Synthase Ketoreductase Domains Utilize the Same Conserved Tyrosine and Serine Residues,” *Biochemistry*, vol. 55, no. 8, pp. 1179–1186, Mar. 2016, doi: 10.1021/acs.biochem.6b00024.

[245] J. Kawata, T. Naoe, Y. Ogasawara, and T. Dairi, “Biosynthesis of the Carbonylmethylene Structure Found in the Ketomemicin Class of Pseudotripeptides,” *Angew. Chemie*, vol. 129, no. 8, pp. 2058–2061, 2017, doi: 10.1002/ange.201611005.

[246] J. M. Andrews, “Determination of minimum inhibitory concentrations,” *J. Antimicrob. Chemother.*, vol. 48, no. suppl_1, pp. 5–16, Jul. 2001, doi: 10.1093/jac/48.suppl_1.5.

[247] F. Calzada *et al.*, “Antiprotozoal activities of tiliroside and other compounds from Sphaeralcea angustifolia (Cav.) G. don,” *Pharmacognosy Res.*, vol. 9, no. 2, pp. 133–137, 2017, doi: 10.4103/0974-8490.204644.

[248] V. P. C. Rocha, F. R. Nonato, E. T. Guimarães, L. A. R. de Freitas, and M. B. P. Soares, “Activity of antimalarial drugs in vitro and in a murine model of cutaneous leishmaniasis,” *J. Med. Microbiol.*, vol. 62, no. PART7, pp. 1001–1010, 2013, doi: 10.1099/jmm.0.058115-0.

[249] M. Kaiser, P. Mäser, L. P. Tadoori, J. R. Ioset, R. Brun, and D. J. Sullivan, “Antiprotozoal activity profiling of approved drugs: A starting point toward drug repositioning,” *PLoS One*, vol. 10, no. 8, pp. 1–16, 2015, doi: 10.1371/journal.pone.0135556.

[250] N. Eckshtain-Levi *et al.*, “Insights from the genome sequence of Acidovorax citrulli M6, a group I strain of the causal agent of bacterial fruit blotch of cucurbits,” *Front. Microbiol.*, vol. 7, no. APR, pp. 1–12, 2016, doi: 10.3389/fmicb.2016.00430.

[251] M. F. Kreutzer and M. Nett, “Genomics-driven discovery of taiwachelin, a lipopeptide siderophore from Cupriavidus taiwanensis,” *Org. Biomol. Chem.*, vol. 10, no. 47, pp. 9338–9343, 2012, doi: 10.1039/c2ob26296g.

[252] E. A. Felnagle *et al.*, “MbtH-Like Proteins as Integral Components of Bacterial Nonribosomal Peptide Synthetases,” *Biochemistry*, vol. 49, no. 41, pp. 8815–8817, Oct. 2010, doi: 10.1021/bi1012854.

[253] R. S. Flugel, Y. Hwangbo, R. H. Lambalot, J. E. Cronan Jr., and C. T. Walsh, “Holo-(Acyl Carrier Protein) Synthase and Phosphopantetheinyl Transfer in Escherichia coli *,” *J. Biol. Chem.*, vol. 275, no. 2, pp. 959–968, Jan. 2000, doi: 10.1074/jbc.275.2.959.

[254] M. Kotowska and K. Pawlik, “Roles of type II thioesterases and their application for secondary metabolite yield improvement,” *Appl. Microbiol. Biotechnol.*, vol. 98, no. 18, pp. 7735–7746, 2014, doi: 10.1007/s00253-014-5952-8.

[255] F. J. Sharom, P. Lu, R. Liu, and X. Yu, “Linear and cyclic peptides as substrates and modulators of P-glycoprotein: Peptide binding and effects on drug transport and accumulation,” *Biochem. J.*, vol. 333, no. 3, pp. 621–630, 1998, doi: 10.1042/bj3330621.

[256] M. P. Kem, H. Naka, A. Iinishi, M. G. Haygood, and A. Butler, “Fatty acid hydrolysis of acyl marinobactin siderophores by Marinobacter acylases,” *Biochemistry*, vol. 54, no. 3, pp. 744–752, 2015, doi: 10.1021/bi5013673.

[257] M. D. Angelica and Y. Fong, “Mechanistic and structural studies of the N-hydroxylating flavoprotein monooxygenases,” *Bioorg. Chem.*, vol. 39, no. 5–6, pp. 171–177, 2011, doi: 10.1016/j.bioorg.2011.07.006.Mechanistic.

[258] L. Rouhiainen *et al.*, “Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in Anabaena strain 90,” *Mol. Microbiol.*, vol. 37, no. 1, pp. 156–167, 2000, doi: 10.1046/j.1365-2958.2000.01982.x.

[259] G. L. Challis and J. Ravel, “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.*, vol. 187, no. 2, pp. 111–114, 2000, doi: 10.1016/S0378-1097(00)00184-1.

[260] I. J. Schalk and L. Guillon, “Pyoverdine biosynthesis and secretion in Pseudomonas aeruginosa: Implications for metal homeostasis,” *Environ. Microbiol.*, vol. 15, no. 6, pp. 1661–1673, 2013, doi: 10.1111/1462-2920.12013.

[261] S. D. Bruner *et al.*, “Structural Basis for the Cyclization of the Lipopeptide Antibiotic Surfactin by the Thioesterase Domain SrfTE,” *Structure*, vol. 10, no. 3, pp. 301–310, 2002, doi: https://doi.org/10.1016/S0969-2126(02)00716-5.

[262] V. Miao *et al.*, “Daptomycin biosynthesis in Streptomyces roseosporus: Cloning and analysis of the gene cluster and revision of peptide stereochemistry,” *Microbiology*, vol. 151, no. 5, pp. 1507–1523, 2005, doi: 10.1099/mic.0.27757-0.

[263] A. Pandey, C. Savino, S. H. Ahn, Z. Yang, S. G. Van Lanen, and E. Boros, “Theranostic Gallium Siderophore Ciprofloxacin Conjugate with Broad Spectrum Antibiotic Potency,” *J. Med. Chem.*, vol. 62, no. 21, pp. 9947–9960, 2019, doi: 10.1021/acs.jmedchem.9b01388.

[264] M. Petrik *et al.*, “68Ga-siderophores for PET imaging of invasive pulmonary aspergillosis: Proof of principle,” *J. Nucl. Med.*, vol. 51, no. 4, pp. 639–645, 2010, doi: 10.2967/jnumed.109.072462.

[265] E. Frangipani, C. Bonchi, F. Minandri, F. Imperi, and P. Visca, “Pyochelin potentiates the inhibitory activity of gallium on Pseudomonas aeruginosa,” *Antimicrob. Agents Chemother.*, vol. 58, no. 9, pp. 5572–5575, 2014, doi: 10.1128/AAC.03154-14.

[266] N. Durán *et al.*, “Advances in Chromobacterium violaceum and properties of violacein-Its main secondary metabolite: A review,” *Biotechnol. Adv.*, vol. 34, no. 5, pp. 1030–1045, 2016, doi: 10.1016/j.biotechadv.2016.06.003.

[267] M. Durán, A. N. Ponezi, A. Faljoni-Alario, M. F. S. Teixeira, G. Z. Justo, and N. Durán, “Potential applications of violacein: A microbial pigment,” *Med. Chem. Res.*, vol. 21, no. 7, pp. 1524–1532, 2012, doi: 10.1007/s00044-011-9654-9.

[268] C. K. Venil, Z. A. Zakaria, and W. A. Ahmad, “Bacterial pigments and their applications,” *Process Biochem.*, vol. 48, no. 7, pp. 1065–1079, 2013, doi: 10.1016/j.procbio.2013.06.006.

[269] M. Numan *et al.*, “Therapeutic applications of bacterial pigments: a review of current status and future opportunities,” *3 Biotech*, vol. 8, no. 4, pp. 1–15, 2018, doi: 10.1007/s13205-018-1227-x.

[270] N. Durán *et al.*, “Advances in Chromobacterium violaceum and properties of violacein-Its main secondary metabolite: A review,” *Biotechnol. Adv.*, vol. 34, no. 5, pp. 1030–1045, 2016, doi: 10.1016/j.biotechadv.2016.06.003.

[271] Y. Nakamura, C. Asada, and T. Sawada, “Production of antibacterial violet pigment by psychrotropic bacterium RT102 strain,” *Biotechnol. Bioprocess Eng.*, vol. 8, no. 1, pp. 37–40, 2003, doi: 10.1007/BF02932896.

[272] R. D. DEMOSS and M. E. HAPPEL, “Nutritional requirements of Chromobacterium violaceum.,” *J. Bacteriol.*, vol. 77, no. 2, pp. 137–141, 1959, doi: 10.1128/jb.77.2.137-141.1959.

[273] H. Laatsch, R. H. Thomson, and P. J. Cox, “Spectroscopic properties of violacein and related compounds: Crystal structure of tetramethylviolacein,” *J. Chem. Soc. Perkin Trans. 2*, no. 8, pp. 1331–1339, 1984, doi: 10.1039/p29840001331.

[274] I. Ghiviriga, B. E. D. M. El-Gendy, P. J. Steel, and A. R. Katritzky, “Tautomerism of guanidines studied by 15N NMR: 2-hydrazono-3-phenylquinazolin-4(3H)-ones and related compounds,” *Org. Biomol. Chem.*, vol. 7, no. 19, pp. 4110–4119, 2009, doi: 10.1039/b907577a.

[275] L. Xiong, J. L. L. Teng, M. G. Botelho, R. C. Lo, S. K. P. Lau, and P. C. Y. Woo, “Arginine metabolism in bacterial pathogenesis and cancer therapy,” *Int. J. Mol. Sci.*, vol. 17, no. 3, pp. 1–18, 2016, doi: 10.3390/ijms17030363.

[276] A. Maghnouj, T. F. De Sousa Cabral, V. Stalon, and C. Vander Wauven, “The arcABDC gene cluster, encoding the arginine deiminase pathway of Bacillus licheniformis, and its activation by the arginine repressor ArgR,” *J. Bacteriol.*, vol. 180, no. 24, pp. 6468–6475, 1998, doi: 10.1128/jb.180.24.6468-6475.1998.

[277] H. Machado, E. C. Sonnenschein, J. Melchiorsen, and L. Gram, “Genome mining reveals unlocked bioactive potential of marine Gram-negative bacteria,” *BMC Genomics*, vol. 16, no. 1, pp. 1–12, 2015, doi: 10.1186/s12864-015-1365-z.

[278] M. Burnat, S. Picossi, A. Valladares, A. Herrero, and E. Flores, “Catabolic

pathway of arginine in Anabaena involves a novel bifunctional enzyme that produces proline from arginine," *Mol. Microbiol.*, vol. 111, no. 4, pp. 883–897, 2019, doi: 10.1111/mmi.14203.

[279] W. L. Muth and R. N. Costilow, "Ornithine Cyclase (Deaminating)," *J. Biol. Chem.*, vol. 249, no. 23, pp. 7463–7467, 1974, doi: 10.1016/s0021-9258(19)81261-9.

[280] J. L. Goodman, S. Wang, S. Alam, F. J. Ruzicka, P. A. Frey, and J. E. Wedekind, "Ornithine cyclodeaminase: Structure, mechanism of action, and implications for the μ-crystallin family," *Biochemistry*, vol. 43, no. 44, pp. 13883–13891, 2004, doi: 10.1021/bi048207i.

[281] X. Wu *et al.*, "Whole genome sequencing and comparative genomics analyses of Pandoraea sp. XY-2, a new species capable of biodegrade tetracycline," *Front. Microbiol.*, vol. 10, no. JAN, pp. 1–19, 2019, doi: 10.3389/fmicb.2019.00033.

[282] F. Rosconi *et al.*, "Identification and structural characterization of serobactins, a suite of lipopeptide siderophores produced by the grass endophyte Herbaspirillum seropedicae," *Environ. Microbiol.*, vol. 15, no. 3, pp. 916–927, 2013, doi: 10.1111/1462-2920.12075.

Publications

1. J. Diettrich, H. Kage, M. Nett, Genomics-inspired discovery of massiliachelin, an agrochelin epimer from *Massilia* sp. NR 4-1. *Beilstein J. Org. Chem.* **2019***,* 15 (4), 1298–1303, 2019, doi: 10.3762/bjoc.15.128.

2. J. Diettrich, H. Kage, M. Nett, Massiliachelin, an antibiotic from *Massilia* sp. NR 4-1. In *Scientific Highlights 2019*; Tiller, J. C., Ed.; Fakultät Bio- und Chemieingenieurwesen, TU Dortmund, **2020**; p 66.

Curriculum vitae

Personal Information

Name	Jan Diettrich
Date of birth	12.01.1987
Place of birth	Witten, Germany
Nationality	German

Education, qualification and professional experience

2016 – 2020	PhD student, TU Dortmund University, Dortmund, Germany Department of Biochemical and Chemical Engineering, Chair of Technical Biology
2015 – 2016	ARVOS GmbH, Division Schmidtsche Schack, Process engineer
2013 – 2014	Technical Biochemistry, University of Applied Science Lübeck, Germany (Master of Science, 1.4)
2009 – 2013	Chemical and environmental engineering, University of Applied Science Lübeck, Germany (Bachelor of Science, 1.5)
2007 – 2009	Bioengineering TU Dortmund University, Dortmund, Germany
1997 – 2006	Amplonius Gymnasium, Rheinberg Degree: Abitur (2.2)
1993 – 1997	Paul-Gerhardt Grundschule, Rheinberg, Germany

Declaration (Affidavit)

I declare in lieu of oath that I have completed the present doctoral thesis with the following title independently and without any unauthorized assistance. I have not used any other sources or aids than the ones listed and have documented quotations and paraphrases as such. The thesis in its current or similar version has not been submitted to an auditing institution.

I declare in lieu of oath that I have not attempted to graduate in any doctoral procedure before.