

**Novel Secondary Metabolites
from Myxobacteria
and their
Biosynthetic Machinery**

Dissertation
zur Erlangung des Grades
des Doktors der Naturwissenschaften
der Naturwissenschaftlichen-Technischen Fakultät III
Chemie, Pharmazie, Bio- und Werkstoffwissenschaften
der Universität des Saarlandes

von

Ole Revermann

Saarbrücken

2012

Tag des Kolloquiums:	01 März 2013
Dekan:	Prof. Dr. Volkhard Helms
Berichterstatter:	Prof. Dr. Rolf Müller Prof. Dr. Uli Kazmaier
Vorsitz:	Prof. Dr. Rolf W. Hartmann
Akad. Mitarbeiter:	Akad. Direktor Dr. Josef Zapp

Danksagung

Ich bedanke mich besonders bei Prof. Dr. Rolf Müller für die Aufnahme in seinen Arbeitskreis und die intensive Betreuung mit vielen hilfreichen Diskussionen und wertvollen Anregungen.

Ein herzlicher Dank geht an Herrn Prof. Dr. Uli Kazmaier für die Übernahme des Koreferates. Ich danke Daniel Krug für die individuelle Einführung in die verschiedenen Themengebiete, sowie die persönliche Betreuung mit vielen produktiven Gesprächen und hilfreichen Ideen. Ein Dank geht an Luke, Janet, Kookie und Daniel, die einzelne Teile dieser Dissertation Korrektur gelesen haben. Ein Dankeschön auch an alle Autoren / Coautoren der einzelnen Publikationen und für die gute Zusammenarbeit.

Meiner Diplomandin Anette und meinen Wahlpflichtpraktikantinnen Irina und Larissa danke ich für eine schöne und produktive Zeit im Labor.

Ich bedanke mich bei Thorsten, Kathrin, Dominik, Kookie, Tina und Ronald, die meine treuen Gefährten sowohl während der Arbeit, als auch in meiner Freizeit waren. Ich möchte mich bei allen Mitgliedern (inklusive allen Ehemaligen) des Arbeitskreises von Prof. Dr. Rolf Müller für eine schöne gemeinsame Zeit bedanken. Helge und seiner gesamten Gruppe, danke ich für eine gute Zusammenarbeit.

Für die große Unterstützung im Bereich NMR geht mein besonderer Dank an Josef Zapp.

Für meine angenehmen Kurzaufenthalte in Braunschweig, bedanke ich mich herzlich bei Heinrich Steinmetz und Rolf Jansen.

Ein herzliches Dankeschön geht natürlich ebenfalls an alle meine Freunde und WG-Mitbewohner, die für ein ausgeglichenes Freizeitprogramm gesorgt haben.

Nicht zuletzt geht ein besonderer Dank an meine Familie und meine Freundin Sabrina, die mich unterstützt und motiviert haben und immer hinter mir standen.

Ich danke von Herzen meinem Vater, der dies leider nicht mehr miterleben durfte.

Vorveröffentlichungen der Dissertation

Einige Teile dieser Arbeit wurden mit der Genehmigung des Fachbereichs, vertreten durch den Mentor der Arbeit, in den nachfolgenden Beiträgen vorab publiziert:

Publikationen

Cortina, N. S., Krug, D., Plaza, A., **Revermann, O.** and Müller, R. (2012) “Myxoprincomide: a natural product from *Myxococcus xanthus* discovered by comprehensive analysis of the secondary metabolome”, *Angewandte Chemie-International Edition*, 51 (3) 811-816.

Rachid, S., **Revermann, O.**, Dauth, C., Kazmaier, U., Müller, R. (2010): “Characterization of a novel type of oxidative decarboxylase involved in the biosynthesis of the styryl moiety of chondrochloren from an acylated tyrosine.” *J. Biol. Chem.*, 285 (17): 12482-12489.

Weitere Publikationen

Folgende Veröffentlichungen erschienen im Zeitraum dieser Arbeit, sind jedoch nicht Bestandteil dieser Dissertation:

Cortina, N. S., **Revermann, O.**, Krug, D. and Müller, R. (2011): “Identification and Characterization of the Althiomycin Biosynthetic Gene Cluster in *Myxococcus xanthus* DK897”, *ChemBioChem*, 12 (9), 1411-1416.

Krug, D., Zurek, G., **Revermann, O.**, Vos, M., Velicer, G. J., Müller, R. (2008): “Discovering the hidden secondary metabolome of *Myxococcus xanthus*: a study of intraspecific diversity”, *Applied and Environmental Microbiology*, 74 (10):3058-68.

Tagungsbeiträge

Vorträge

Revermann, O., Krug, D., Kegler, C., und Müller, R. (September 2008): “Biosynthesis of the Bicyclic Tetrapeptide Cittilin”, VAAM International Workshop „*Biology and Chemistry of Antibiotic-Producing Bacteria*“, Technische Universität Berlin

Revermann, O., Rachid, S. and Müller, R. (September 2007) „Biosynthese des Chondrochlorens aus *Chondromyces crocatus*“, 28. Tübinger-Göttinger Gespräche zur Chemie von Mikroorganismen, Haus Benediktushöhe in Zellingen-Retzbach.

Revermann, O., Krug, D., und Müller, R. (September 2006) „Screening nach neuen Metaboliten aus *Myxococcus xanthus*“, 27. Tübinger-Göttinger Gespräche zur Chemie von Mikroorganismen, Haus Benediktushöhe in Zellingen-Retzbach.

Posterpräsentationen

Revermann, O., Krug, D., und Müller, R. (Oktober 2007) “Myxobacteria as multiproducers of secondary metabolites”, International Workshop "*Biology of Bacteria Producing Natural Products*", European Academy Otzenhausen in Nonnweiler/Saarland.

Revermann, O., Krug, D., und Müller, R. (Juli 2007) “Myxobacteria as multiproducers of secondary metabolites”, 34th International conference on the biology of the myxobacteria Granada, Spanien.

Zusammenfassung

Die vorliegende Arbeit befaßt sich mit Naturstoffen aus Myxobakterien. Mittels LC-MS wurde in Extrakten von *Myxococcus xanthus* Kulturen nach neuen Sekundärmetaboliten gesucht. Insgesamt wurden 4 unterschiedliche Substanzfamilien isoliert und charakterisiert. Die Identifizierung der zugrundeliegenden Biosynthesegencluster zeigt interessante Einblicke in unterschiedliche Biosynthesewege von Myxobakterien.

Das Citteringencluster wurde im Genom von *M. xanthus* Stämmen identifiziert. Für einen myxobakteriellen Sekundärmetaboliten konnte durch heterologe Expression in *S. aurantiaca* Sga 15 erstmalig bewiesen werden, dass dieser ribosomalen Ursprungs ist. Die Modifikation des genetischen Codes führte zu neuen Citterinen.

Die neuen Myxoprincomide werden durch einen NRPS/PKS-Hybrid Biosyntheseweg gebildet. 10 Myxoprincomidderivate wurden aus *M. xanthus* DK897 isoliert. Mit 39 katalytischen Domänen ist es das bislang größte myxobakterielle Biosynthesegen. Aufgrund ungewöhnlicher Aminosäuren sowie seltener Strukturmerkmale, ergibt sich insbesondere im Vergleich zu *M. xanthus* DK1622 für die Myxoprincomide eine außergewöhnliche Biosynthese. Erstaunlich ist auch die Vielzahl von 10 Derivaten, die größtenteils als Intermediate der Biosynthese hydrolytisch freigesetzt werden.

Des Weiteren wurden Prechondrochloren von einer *C. crocatus* Cm c5 Mutante sowie neue Spirangiene mit einem verkürzten Chromophor von einer zufälligen *S. cellulorum* So ce90 Mutante isoliert und strukturell aufgeklärt.

Abstract

This thesis deals with natural products from myxobacteria. *Myxococcus xanthus* extracts were screened by LC-MS for novel secondary metabolites. Altogether 4 different substance families were isolated and characterized. Identification and analysis of the associated biosynthetic gene clusters allows interesting insights into diverse biosynthetic routes of myxobacteria.

The cittilin gene cluster was identified in the genome of *M. xanthus*. For the first time, in case of a myxobacterial secondary metabolite, it could be demonstrated by heterologous expression in *S. aurantiaca* Sga 15 that the cittilin precursor peptide is built by the ribosome. Modification of the genetic code yielded novel cittilin derivatives.

The novel myxoprincomides (mxp) core structure is produced by a NRPS/PKS hybrid biosynthetic pathway in *M. xanthus* DK897 and DK1622. 10 mxp derivatives from DK897 were isolated and correlated to the so far largest myxobacterial biosynthetic gene exhibiting 39 catalytic domains. The mxp biosynthesis is exciting due to incorporation of unusual amino acids and other rare structural features, particularly in comparison to mxp of DK1622. Strikingly, the number of 10 mxp derivatives mainly representing hydrolytically released intermediates of the biosynthesis.

Prechondrochloren was isolated from a *C. crocatus* Cm c5 mutant and structurally elucidated. Furthermore it was shown that a spontaneous mutant of *S. cellulosum* So ce90 produces novel spirangien derivatives with a shortened chromophore.

Table of contents

Danksagung	III
Vorveröffentlichungen der Dissertation	V
Zusammenfassung	VII
Abstract	VIII
Table of contents	IX
 Introduction	 1-28
1. Natural products as a source of pharmaceuticals	1
2. Common features of myxobacteria	3
3. Myxobacteria as multi-producers of secondary metabolites	5
4. Secondary metabolites of <i>Myxococcus xanthus</i>	9
5. Biosynthetic machinery for secondary metabolite formation in bacteria	11
5.1 Biochemistry of nonribosomal peptide synthetases and polyketide synthases	11
5.2 Post-assembly line modifications	18
5.3 Biochemistry of ribosomally-produced peptides	19
6. Discovery of novel myxobacterial metabolites, their structural characterization and elucidation of underlying biosynthetic pathways	22
7. Outline of this work	25
 Chapter 1 to 4	 29-278
 Chapter 1	
Cittilin, an unusual ribosomally produced secondary metabolite from <i>Myxococcus xanthus</i>	29
Chapter 2	
Chapter 2A	68
The structure elucidation and biosynthesis of novel Myxoprincomides from <i>Myxococcus xanthus</i> DK897	69

Chapter 2B

Myxoprincomide: A Natural Product from <i>Myxococcus xanthus</i> Discovered by Comprehensive Analysis of the Secondary Metabolome	169
---	-----

Chapter 3

Characterization of a Novel Type of Oxidative Decarboxylase Involved in the Biosynthesis of the Styryl Moiety of Chondrochloren from an Acylated Tyrosine	223
---	-----

Chapter 4

Natural evolution in the laboratory: Spirangien C – F produced by a spontaneous mutant of <i>Sorangium cellulosum</i> So ce90	251
---	-----

Discussion	279-298
-------------------	---------

1. Biosynthetic pathways	279
2. Structure and biosynthesis of cittelin	280
3. Myxoprincomide of <i>Myxococcus xanthus</i> DK1622 and DK897	284
4. Module skipping increases chemical diversity	292
5. Pre-chondrochloren, last intermediate in chondrochloren biosynthesis	294
6. Spirangien and spiroketal formations	295

References	299
-------------------	-----

Author's contributions from chapter 1 to 4	315
---	-----

Abbreviations	317
----------------------	-----

Introduction

1. Natural products as a source of pharmaceuticals

Microorganisms produce large numbers of structurally diverse secondary metabolites and thus constitute an invaluable source of new drugs. Some of these natural products (NPs) exhibit a wide range of biological activities like anti-cancer, antifungal, herbicidal, insecticidal or immunosuppressive.^[1-4] Many drugs are isolated from natural sources (e.g. taxol, rapamycin, vancomycin) or derived from NPs (β -lactams, steroids, macrolides). In other cases, only the pharmacophore is based on a natural compound, as in the case of Lipitor[®], a drug used to treat high cholesterol levels. Natural compounds are therefore a very useful launching point for drug development. As an example, small modifications to the structure of epothilone, a secondary metabolite originally isolated from the myxobacterium *Sorangium cellulosum*, recently led to the establishment of a new anticancer drug, Ixabepilone[®].^[5] That means if NPs are not used as found in nature, they often are used as lead structures for novel drug development. NPs have traditionally also been a rich source of antibiotics; the first antibiotic, penicillin, was a chance discovery from a fungi by Alexander Fleming. Some compounds have also been adopted for use in agriculture as herbicides or as insecticides.^[4]

Natural products often show huge structural diversity that is difficult to achieve via organic synthesis. The annual percentage is currently estimated at around 50% of pharmaceutical sales derived from natural sources.^[1;6] Assumptions on the overall influence of NPs to the discovery and development of pharmaceutical treatments are up to 75%.^[7] Prominent examples used in clinical therapy are the erythromycin (antibiotic) and rapamycin (immunosuppressant), salinosporamide (proteasome inhibitor), glycopeptides like vancomycin (antibiotic), and the lipopeptide daptomycin (antibiotic).^[1;8-11]

Pathogens are becoming increasingly multi-drug resistant. Standard therapies are no longer sufficient, and even second and third-line drugs are not active enough against emerging multi-drug resistant strains.^[4;12] As anti-microbial drugs lose their effectiveness, it becomes even more critical to search for novel biologically active NPs. As new sources of chemical diversity, NPs are essential to continuously drive new drug discovery or development.^[7]

Several major pharmaceutical companies have abandoned the field, focusing rather on chemical synthesis, and high-throughput screening of synthetic libraries, but smaller companies and research groups continue the search for new NPs.^[13] From 2000 to 2008, 300

secondary metabolites with antimicrobial activity were discovered.^[14] In 2008, 1065 new NPs of marine origin were reported.^[15]

Natural compounds perform better in bioactivity assays, showing higher hit rates compared to synthetic libraries. Combinatorial synthesis did not result in a large improvement in drug development, largely due to a lack of “structural complexity” compared to bioactive natural products.^[16] NPs generally have lower numbers of nitrogen, halogen, and sulfur atoms but higher amounts of oxygen.^[17] Another reason for better activity of NPs might be that they have been selected for through an evolutionary process.^[18]

Bioactive compounds mainly originate from four different groups: bacteria (33%), plants (27%), fungi (15%), and animals (13%).^[17] In the case of microorganisms, some bacterial orders, such as the *Actinomycetales*, have been intensively investigated. Thus, in order to increase the chance of finding a novel compound it is worth looking for other species to find novel structures and modes of action.^[19] So far, only around 1 % of the prokaryotes have been cultivated,^[4] and an even lower percentage is expected for bacteria from unusual habitats,^[20] organisms which are very difficult to culture in the lab,^[21] or pathogens.^[22] However, laboratory conditions are being established for more and more bacteria, such as cyanobacteria,^[23] marine actinomycetes,^[24] some bacterial endosymbionts,^[25;26] pseudomonads, and myxobacteria.^[27;28] Regarding myxobacteria marine environments or other less explored habitats are of great interest for the isolation of novel species.^[12;29]

2. Common features of myxobacteria

Myxobacteria are Gram-negative bacteria belonging to the delta subgroup of proteobacteria, which can be isolated from different habitats like soil, herbivore dung, decaying plant mass and tree bark, showing that they are able to adapt to diverse environments.^[30-32] Recently myxobacterial isolates from marine environments have been described and halophilic strains from terrestrial sources have been reported.^[33;34]

Their cells are typically rod-shaped and often rather large (4 – 12 µm long and 0.7 – 1.2 µm wide in their vegetative form). Some genera can be cultivated in dispersion, others will form clumps, even after several passages.^[35]

Compared to other bacteria, myxobacteria have some noticeable features:^[36] The capability to glide on solid surfaces, building swarms, degrading macromolecules, and preying on living microorganisms are their main characteristics. Myxobacteria possess a sort of “social” behavior as they swarm and collectively hunt for other microorganisms such as bacteria and fungi.^[37;38] For that reason, myxobacteria can be called “micropredators”.^[33] To achieve this, they excrete enzymes which lyse other organism so that proteins and lipids can be consumed. The species *Sorangium cellulosum* is not a predator but with the help of exo-enzymes it is even able to degrade cellulose as a nutrition source.^[35;38]

Under starvation conditions, myxobacteria start a multicellular development cycle. Cells aggregate, building characteristic multi-cellular fruiting bodies (Figure 1) containing myxospores^[38;39] which are resistant against extreme environmental conditions like temperature changes, desiccation, and UV radiation.^[33] As soon as environmental conditions become more favorable, the myxospores are retransformed into normal vegetative cells. The shape of the fruiting bodies varies a lot and can therefore be used for taxonomic classification.^[35;40] Myxobacteria belong to the order *Mycococcales*, which can be distinguished into the three suborders *Nannocystineae*, *Cystobacterineae* and *Sorangiiineae*.^[41]

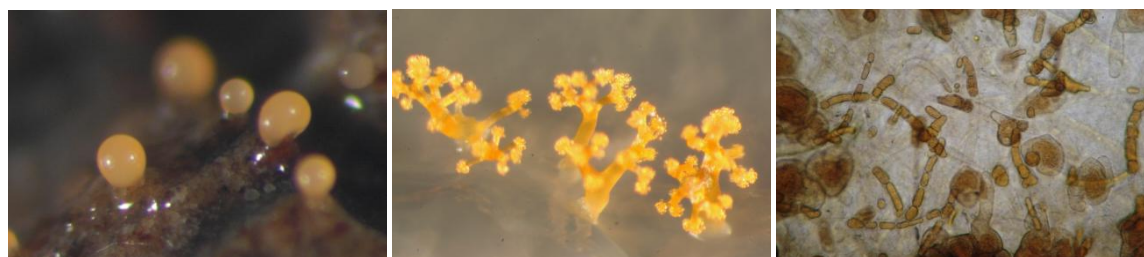


Figure 1: Fruiting bodies of *Myxococcus xanthus*, *Chondromyces crocatus*, *Sorangium cellulosum* (magnification: left 64x, middle 30x, right 60x; pictures R. O. Garcia)

Due to their fascinating behavior, myxobacteria have long been studied by microbiologists. More recently they have also been established to be multi-producers of secondary metabolites and therefore they represent a valuable source for natural products.^[42;43] Alongside to the two other significant wells for natural compounds, the actinomycetes and fungi. Most of the secondary metabolite research with myxobacteria has been performed at the German Research Center for Biotechnology (Braunschweig, Germany; recently renamed to Helmholtz Center for Infection Research). At this center, more than 7500 myxobacterial strains have been isolated and characterized, from which over 100 different core structures with around 500 derivatives have been described to date, with more continuously being found.^[30;38;44] Further work to increase the number of different species and families is ongoing, with a high expectancy for the uncovering of new substance classes.^[30;40] In the past, myxobacteria have been considered exotic organisms but over time they have become more understood and established.^[45;46]

Many of the isolated compounds show novel structural diversity. When bioactivity is detected, it can often be attributed to an outstanding mode of action.^[44-47] This finding is especially valuable for new lead structures for drug development, making myxobacterial research even more attractive.

3. Myxobacteria as multi-producers of secondary metabolites

To date, 67 different substance classes from myxobacteria have been published, but there is evidence of more than 100 diverse core structures. As many of them have further derivatives, there are over 500 substances known.^[30;48] Several different scaffolds of NPs can be produced by myxobacteria, some unique and others already known from other bacteria. The structural diversity varies from polyketides (PKs), non-ribosomal peptides (NRPs), and hybrids thereof, to alkaloids, terpenoids, and phenyl-propanoids.^[45;49]

A huge number of myxobacterial metabolites exhibit biological activities. The range is very broad from antibacterial and cytotoxic to pheromone-like effects. These diverse activities make myxobacterial metabolites highly relevant in clinical research, but we have only barely scratched the surface of myxobacterial secondary metabolism.^[30;50;51] However, not all substances can be assigned to a biological function or mode of action. 29% show antibacterial activity, and the most frequently observed bioactivity of myxobacterial metabolites is anti-fungal activity (54%).^[44]

Distinct structure types can have the same general bioactivity but different modes of action. Epothilone (**1**) and tubulysin (**2**),^[52;53] for example, are cytotoxic, both acting on the cytoskeleton of eukaryotic cells, but their effects are different. Both cause cell cycle arrest and induce apoptosis,^[52;54] but tubulysin (**2**) inhibits the polymerization of tubulin, while **1** stabilizes microtubules. It binds to β -tubulin and induces microtubule polymerization at the cytoskeleton of eukaryotic cells.^[55-57]

A derivative of **1** has been established as an anti-cancer drug. In 2007, Ixabepilone[®] (Ixempra) was approved for the US market as the first myxobacteria derived compound for clinical use against breast cancer. Other compounds interacting with the cytoskeleton produced from myxobacteria are disorazol (**4**),^[58] chondramide (**5**),^[59] and rhizopodin (**6**).^[60;61] A further potential antitumor drug is argyirin (**3**), which inhibits the proteasome.^[62]

Soraphen (**7**), a metabolite from *Sorangium cellulosum* has an anti-fungal effect and was evaluated to be used in plant protection.^[63] However, further studies showed a teratogenic effect, stopping further investigations. In eukaryotic cells, it interacts with acetyl-CoA carboxylase and inhibits its activity.^[64;65] As the acetyl-CoA carboxylase is important for fatty acid synthesis it could be useful for treatment of cancer^[66] and obesity^[67]. Stigmatellin (**8**), as well, has anti-fungal activity, with a different mode of action compared to **7**.^[65;68]

Thuggacins (**9**) are bioactive compounds against *Mycobacterium tuberculosis* and other Gram-positive bacteria, inhibiting the respiratory chain of the bacteria. ^[69;70] Further mentionable are myxothiazol (**10**) and melithiazols (**11**), targeting the mitochondrial electron transport chain. ^[71;72] Sorangicin (**13**), myxopyronin (**12**), and corralopyronin (**14**) affect eubacterial RNA polymerases. ^[73-75] Protein synthesis is targeted by myxovalargin (**18**) ^[76]. Chondrochloren (**15**), partly investigated in this work, has an antibiotic effect. ^[77] The typical earthy smell of myxobacteria is due to geosmin (**16**). ^[78]

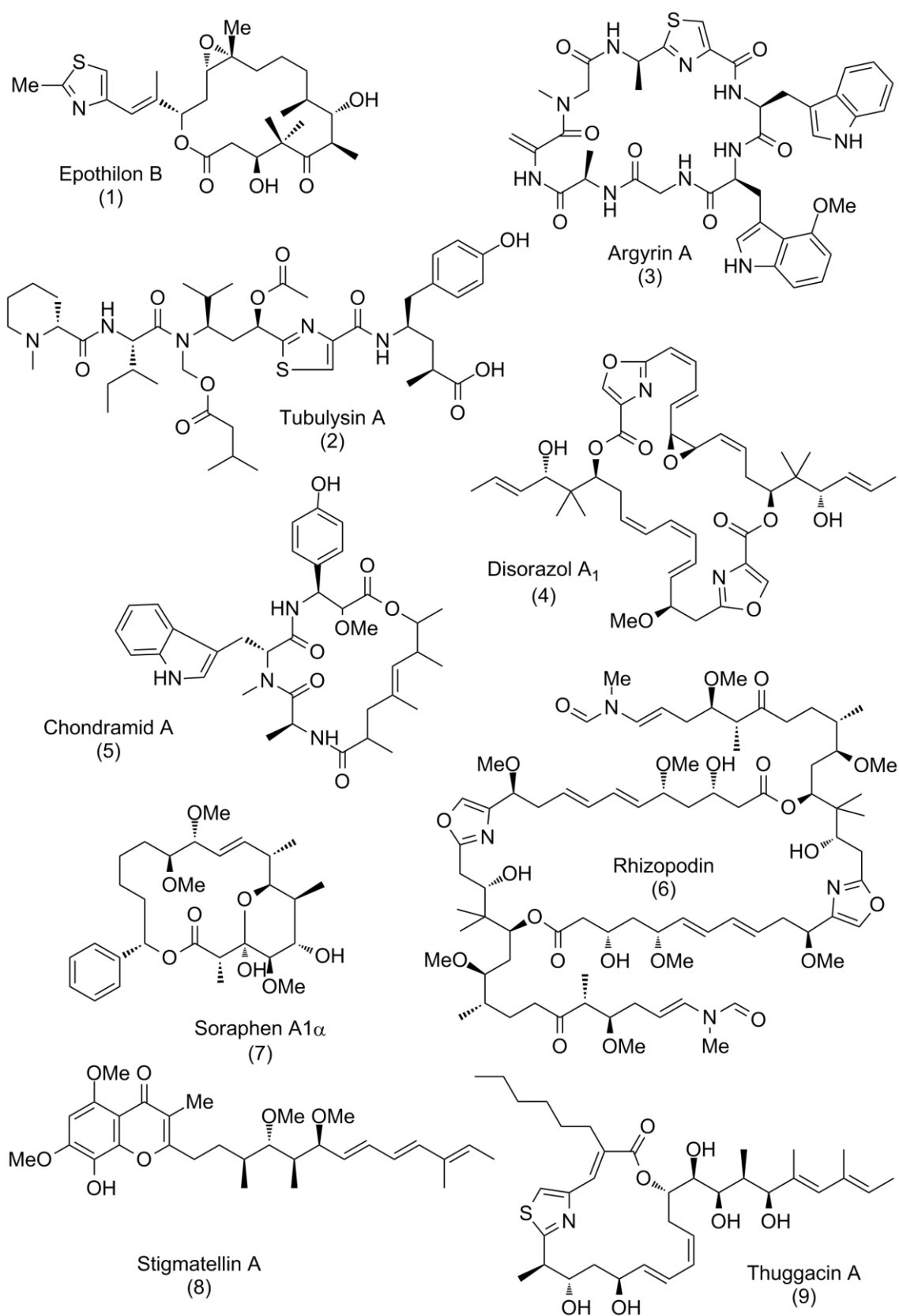


Figure 2: Epothilon B (1), Tubulysin A (2), Argyrin A (3), Disorazol A₁ (4), Chondramid A (5), Rhizopodin (6), Soraphen A1 α (7), Stigmatellin A (8), Thuggacin A (9).

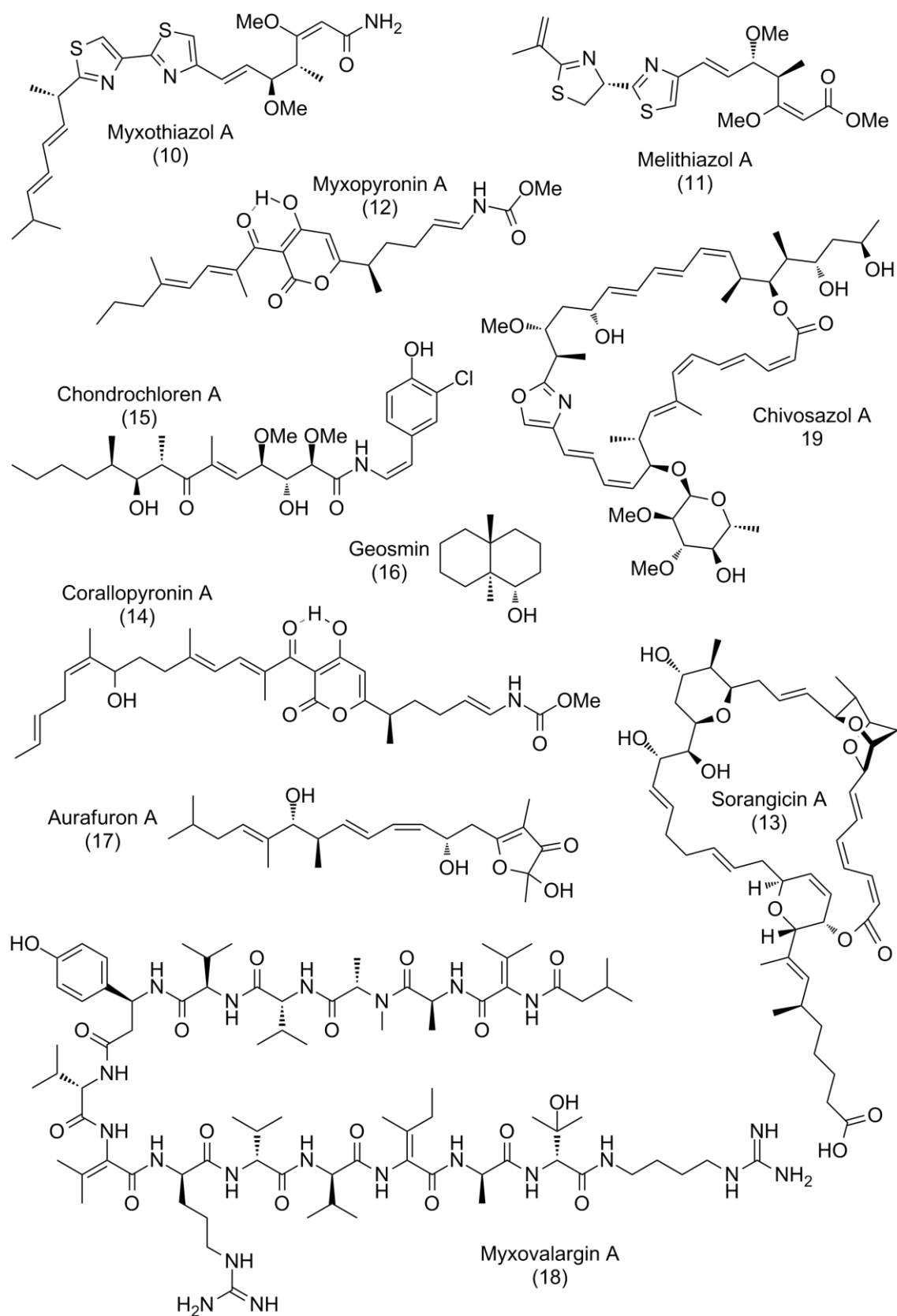


Figure 3: Myxothiazol A (10), Melithiazol A (11), Myxopyronin A (12), Sorangicin A (13), Corallopyronin A (14), Chondrochloren A (15), Geosmin (16), Aurafuron A (17), Myxovalargin A (18), Chivosazol A (19).

4. Secondary metabolites of *Myxococcus xanthus*

The myxobacterium *Myxococcus xanthus* DK1622 is a model strain for prokaryotic multicellular differentiation and thus has become the best-studied myxobacterium. Although it was used extensively for investigations of the bacterial life-cycle, it was long believed to be a non-producer of secondary metabolites. However, since the genome of DK1622 was sequenced, it is known to contain 18 gene clusters for secondary metabolite production.^[79] Biosynthesis of secondary metabolites is often encoded in gene clusters within the genome. The part of the genome in DK1622 dedicated to secondary metabolites (more than 8.5%) is much higher in comparison to other natural product producers such as *Streptomyces coelicolor* (4.5%) and *S. avermitilis* (6.6%).^[80;81] The strain DK1622 produces myxalamides (**20**)^[82], myxochelin (**21**)^[83], myxovirescin (**22**)^[84], myxochromides (**23**)^[85], DKxanthene (**24**)^[86] and two others that are the subject of investigation in the course of this thesis: cittilin and myxoprincomide. However, there are 18 different biosynthetic gene clusters in the genome and until now it was thus not possible to assign each single one to a detectable compound. It is possible that not all gene clusters are expressed under given cultivation conditions, as expression may need to be induced.^[87] Or production levels can be very low if a compound is not needed; for example, if an excess of iron is present, an iron chelator like **21** is not required for growth, meaning that production of that compound will be down-regulated by the cell. Another possibility for the lack of a compound could be its degradation as observed in the case of chivosazol (**19**) at later time points during cultivation.^[88] However, transcriptomic and proteomic investigations of DK1622 showed that almost all clusters are active during lab cultivation.^[89;90]

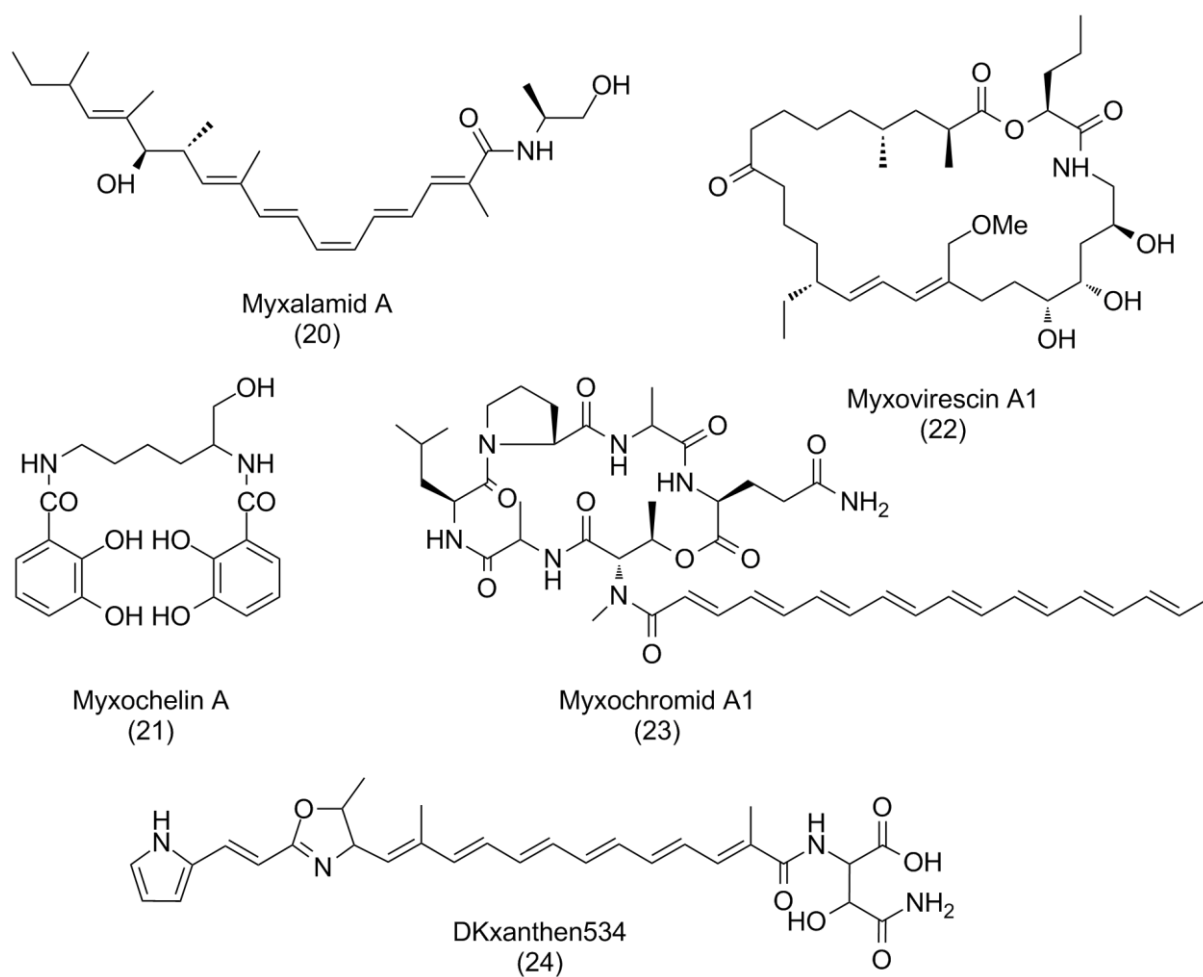


Figure 4: Myxalamide A (20), Myxochelin A (21), Myxovirescin A1 (22), Myxochromide A1 (23), DKxanthen534 (24).

5. Biosynthetic machinery for secondary metabolite formation in bacteria

In addition to the discovery of new structures, their corresponding biosynthetic pathways are attaining more and more interest.^[46] Knowledge of the biosynthetic pathway offers the possibility to manipulate it genetically and therewith to increase yields or obtain new derivatives.^[91] This information can also be used to express biosynthetic gene clusters in heterologous hosts.^[92]

Within the last decade, a large number of biosynthetic gene clusters have been identified. This genetic information enables the prediction of unknown pathways and subsequently new structures.^[93-96]

Myxobacteria have very large genomes compared to other bacteria, as revealed by genome sequencing of *Myxococcus xanthus* DK1622 and *Sorangium cellulosum* So ce56. A large genome size seems to be characteristic for myxobacteria and is littered with secondary metabolite gene clusters,^[97] making them very likely to be a rich source of currently unknown secondary metabolites.

Most myxobacterial products are polyketides (PKs), nonribosomal peptides (NRPs)^[98], or hybrids of both^[42;43]; recent investigations partly described in this thesis show as well ribosomally produced compounds. In addition to these NP groups, steroids have been found in myxobacteria, which is rather unusual for prokaryotes.^[99]

5.1 Biochemistry of nonribosomal peptide synthetases and polyketide synthases

The origin of two large groups of secondary metabolites, the polyketides and non-ribosomal peptides, are products of multi-step biosynthetic pathways. The underlying biosyntheses is directed by complex biosynthetic machineries, such biosynthetic machineries are polyketide synthases (PKS) and non-ribosomal peptide synthetases (NPRS). Aurafuron (**17**) and soraphen (**7**) for instance are myxobacterial metabolites generated by pure PKSs.^[100] While myxochelin (**21**) is a myxobacterial representative resulting of a NRPS pathway.

These PKS and NRPS multi-enzymes condense different substrate units (monomeric building blocks) with each other to form a final product. Multi-enzymes consist out of repeated functional units which are called modules. One module is in charge of the incorporation of

one building block. A module is further divided into domains which are responsible for single reaction steps.

This is analogous to an assembly line, where one module is responsible for one component here for the incorporation of one building block. ^[101;102] Single domains within each module perform each necessary enzymatic reaction.

PKS and NRPS use similar mechanisms for product chain elongation. The general mechanism is explained by the “multiple carrier thio-template mechanism” model. It shows how substrate and intermediates are activated and processed. ^[103;104] All building blocks are covalently tethered to a carrier protein (CP) via a thioester bond to the terminal thiol of 4'-phosphopantetheine (4'-ppant) arm in every module. NRPS CPs are peptidyl CPs (or PCPs) and PKS are acyl CPs (or ACPs). Each CP has to be converted from its inactive form (apo) to its active form (holo) by a phosphopantetheinyltransferase (PPTase) enzyme (Figure 5). ^[105;106] The coenzyme A is post-translationally attached by the PPTase to a highly conserved serine residue located in the carrier domain of PKS or NRPS proteins. ^[107]

The phosphopantetheine arm fulfills two functions. One is to tether the intermediate covalently to the module. And the second one is the transport of the intermediate to each active site within the module and to pass it on to the downstream module. ^[107] This can be done because of the high flexibility of the approx. 20 Å long ppant arm.

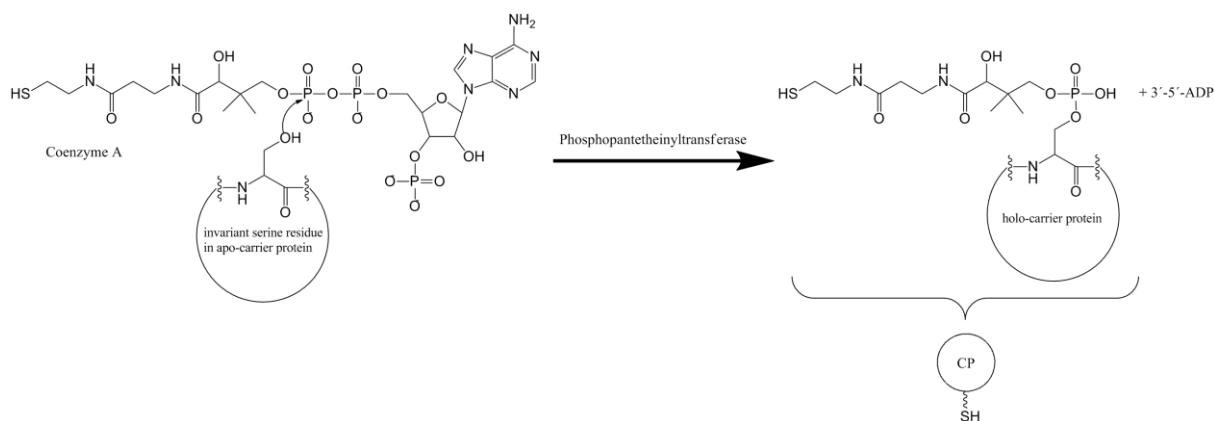


Figure 5: Posttranslational activation of the apo-carrier protein via the phosphopantetheinyltransferase into the active holo-carrier protein.

In some cases, NRPS or PKS work in an archetypal manner, with each module being used once (noniterative) and the reactions taking place in the order of domain arrangement in the genes, following the so-called “co-linearity rule”. ^[108] If so bioinformatic tools can be used to predict the specificity of the domains such that the utilized monomeric building blocks are known. ^[102;105;109] When the genes are arranged in the classical order, then in principle it is feasible to predict the final structure *in silico* and the bacterial extract can be analyzed for the

expected compound. However, this is only possible if the gene cluster fits to the “colinearity rule”, with every domain performing its predicted reaction one after another without exception.^[110] Unfortunately, in many cases, especially for myxobacteria, the correlation between gene and function is not straightforwardly identified, as iterative use of modules (more than one round of chain extension is catalyzed by a module) or skipping of modules (a module is omitted) occurs or unusual starter units or building blocks are incorporated.^[47;108]

5.1.1 Biochemistry of nonribosomal peptide synthetases

Nonribosomal peptide synthetases (NRPS) appear mainly in bacteria and fungi. In the biosynthesis of NRPS, single amino acids are attached to each other. The core NRPS module consists of an adenylation (A) domain, a condensation (C) domain, and a peptidyl carrier protein (PCP, or thiolation (T) domain). Within the loading (or starting) module the C domain is typically lacking.^[105] NRPS biosynthesis starts with the selection and activation of a specific amino acid chosen by the A domain. In comparison to ribosomal peptide synthesis, non-proteinogenic amino acids and aryl acids can also be selected, increasing the structural diversity of NRPS.^[105;106] The selected amino acid is activated by the A domain, a process which consumes ATP (Figure 6).

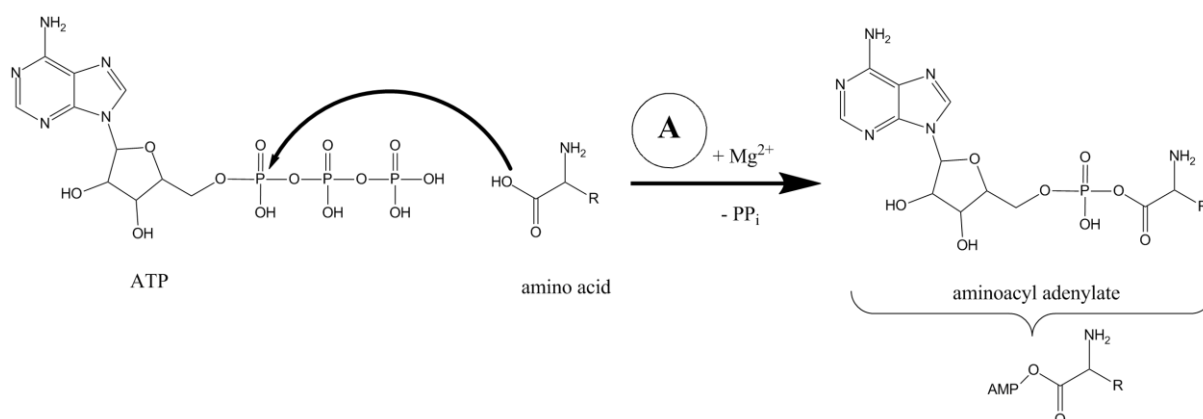


Figure 6: Substrate recognition and activation and by the adenylation domain.

The selected amino acid is loaded and covalent bound onto the PCP of the loading module. Another amino acid is selected and loaded by the A domain of the following module onto the PCP of this module. In the next step the condensation of both amino acids catalyzed by the C domain takes place. The N-terminal end of the PCP-bound amino acid attacks the acyl group of the substrate bound to the preceding PCP in the assembly line. By this nucleophilic

reaction, a new peptide bond is formed and the substrate is extended by one amino acid (Figure 7). These steps are repeated by every module until the last one in the assembly line is reached. The last module typically contains a thioesterase (TE) domain, which is in charge of releasing the complete product from its assembly line. Either by catalyzing the hydrolysis of the enzyme-bound thioester intermediate to build a free carboxylic acid, or by internal macrocyclization.^[108] The free acid is hydrolyzed in the presence of water. While for the ring closure, a nucleophile (typically a -OH or -NH₂ group) from somewhere in the backbone is required to form the ring at the PCP-bound acyl group.

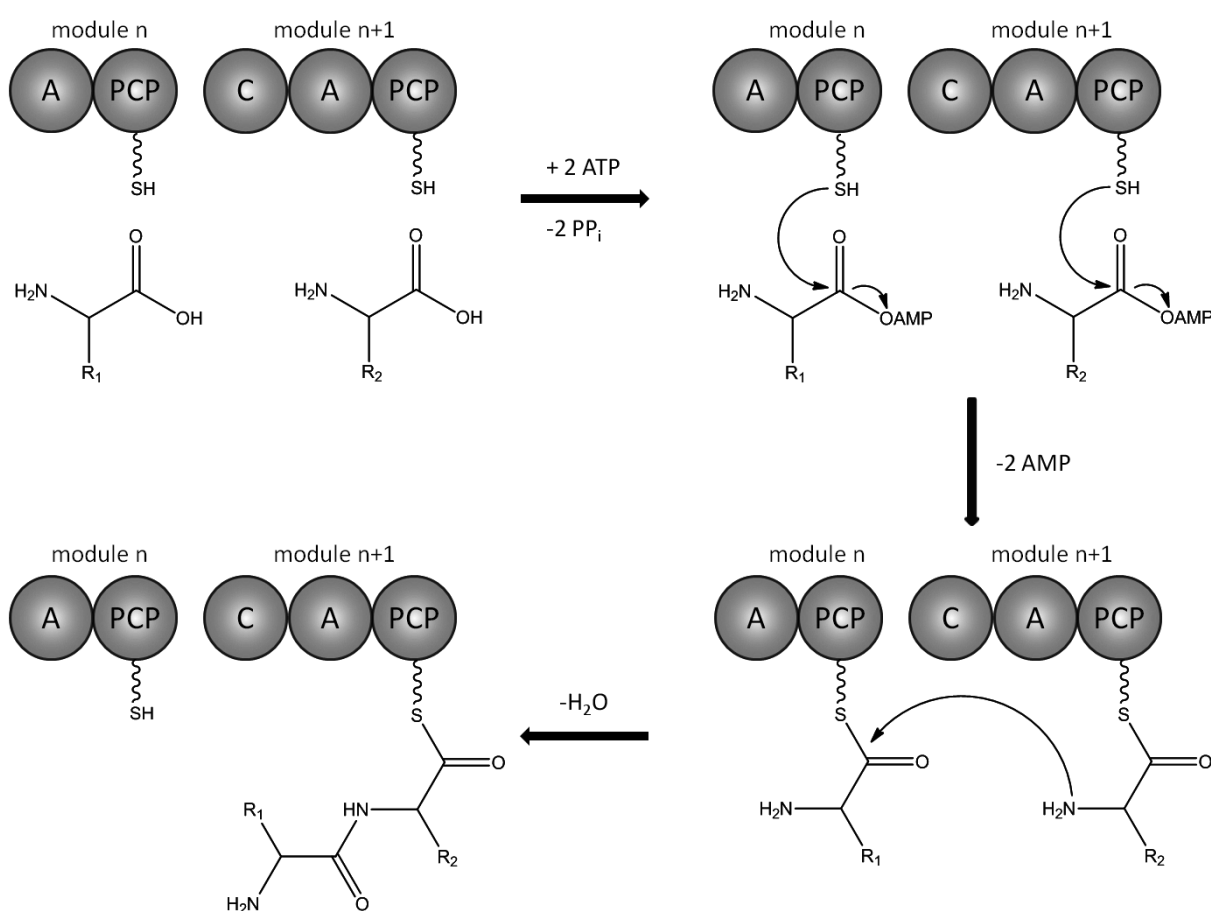


Figure 7: Schematic representation of reactions performed by nonribosomal peptide synthetases (NRPS). Selection and activation of the starter unit by the A domain of module n. And selection and activation of the elongation unit by the A domain of module n+1. After the starter unit and the elongation unit are covalently linked to the PCPs the elongation occurs via condensation catalyzed by the C domain. This results in a chain transfer to module n+1.

Aside from these minimal modules (containing C, A, PCP domain) there are several optional domains that are able to alter the product during assembly. Methyltransferase (MT) domains methylate nitrogens, oxygens, or carbons, and are often integrated into A domains.

Heterocyclization (HC) domains are able to form heterocycles by using the side chains of threonine, cysteine, serine and acyl groups from the backbone.^[106] Oxidation (Ox) domains can further oxidize the compound using FMN as cofactor. They often exist together with other domains and are integrated into A domains.^[106] Epimerase (E) domains can invert α -stereocenters of amino acids to the D-form. But there are also known cases where D-amino acids are directly incorporated, as in the case of cyclosporine.^[111] Or in other cases, the C domain is capable of both reactions the condensation and the epimerization.^[112]

The products can be further modified via tailoring enzymes, resulting in halogenation or cross-linking of amino acid side chains.^[108;113] Such post-NRPS modifications by additional enzymes are quite common. Noteworthy are glycosylations, hydroxylations, acylations, double bond introduction, removal, *O*-, *N*- or *C*-methylations, or epoxidations.^[106]

Since the crystal structure of the gramicidin S synthetase PheA was analyzed in detail the A domain specificity can be predicted from the primary sequence of the protein. From the crystal structure it was observed that 8 to 10 amino acid residues in the binding pocket are critical for the recognition of the incorporated amino acid.^[114-116] But structural data is only available for a limited number of A domains and therefore the deduced nonribosomal code is not yet fully understood. Thus it is still not possible to confidently predict every specific amino acid chosen by the each A domain.

5.1.2 Polyketide synthases

The compound backbone of a polyketide (PK) is constructed by acyl-CoAs. The biosynthetic principle of a polyketide synthase (PKS) is similar to fatty acid biosynthesis and comparable to NRPSs. There are three essential domains in a minimal module, as well as some additional ones. The three core domains are the acyltransferase (AT), the ketosynthase (KS) and the acyl carrier protein (ACP). The AT acts analogously to the A domain in NRPS biosynthesis. The starting AT selects the starter unit, usually an acetyl- or a propionyl-CoA. Exceptions such as branched-chain CoAs or aromatic carboxylic acids are also known.^[98;117-119] For chain extension, malonyl-CoA or methylmalonyl-CoA are typically employed. Nonetheless, hydroxymalonyl-CoA, methoxymalonyl-CoA, and aminomalonyl-CoA have also been reported.^[120-123] The selected building blocks of the AT domains can be predicted in silico, too, like for the A domains. This predication relies on conserved amino acid motifs within the sequence.^[109;124;125]

16

PKS systems are classified into different groups depending on their enzymatic pathways. In bacteria, the most common are type I PKSs. Here, we have modular organization, where each module fulfills one round of chain extension.^[106;110] The multifunctional enzymes contain non-iterative modules.^[110] Good examples of type I PKS products are erythromycin, monensin, and amphotericin.

Type II PKSs are iteratively-acting multienzyme complexes, often producing aromatic or polycyclic polyketides,^[126] such as tetracycline and daunorubicin. Similar to chalcone synthase are the type III PKSs. These type III PKSs are homodimeric enzymes acting iteratively. Flaviolin for example is a result of a type III PKS biosynthesis.

5.1.3 PKS-NRPS hybrids

NRPS-PKS hybrid metabolites are composed of short chain carboxylic acids and amino acids.^[127] The PKS and the NRPS machinery can be combined to produce mixed products. These hybrid products can either be produced in one assembly chain or in two. In the latter one PKS and one NRPS assembly produce each an intermediate which is then stitched together post-synthetically;^[128;129] coronatine for example is produced via this mechanism.^[130-132] In hybrid assembly lines the KS domains are modified to accept a peptidyl intermediate for condensation; conversely, the C domain must accept a ketide. The same holds true for PCPs and ACPs; they must interact with each other or at least submit the intermediates to each other, in order to form a functional assembly line. To overcome such difficulties, the different modules are connected by “linkers” and “docking domains”.

Chondrochloren (**15**), which is also a subject of investigation in the present work, is the product of a hybrid PKS-NRPS biosynthetic pathway (Figure 9).^[133] In total, the chondrochloren gene cluster exhibits 12 open reading frames, five of which are PKS genes and one NRPS gene. It is a neat example where different numbers of reductive domains are also present in each module. The biosynthesis begins with a butyrate starter unit chosen by coenzyme A ligase (CL). Module one houses a full reductive loop (DH, ER, KR); while module two contains only a DH and a KR domain; and module three has just one inactive DH domain. Module four is similar to module two and the last PKS module has a KR and a MT, followed by four ORFs with different functions. The NRPS module contains the core domains (C, A, PCP) and a TE. After the release of the compound by the TE, it is further modified by CndG which decarboxylates the pre-chondrochlorens.^[134]

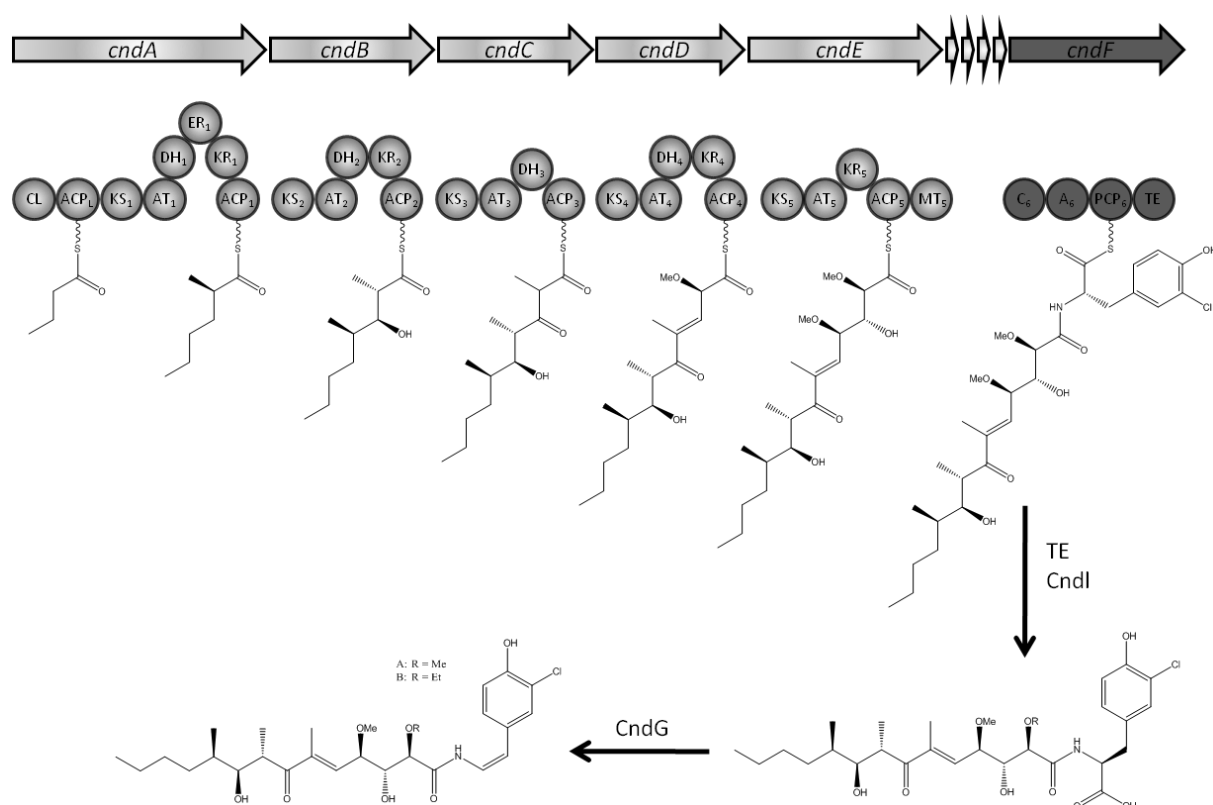


Figure 9: Mixed PKS-NRPS biosynthetic pathway of chondrochloren A and B. *cndA* to *cndE* are PKS genes, *cndF* is a NRPS gene; CL = CoA-ligase; ACP = acyl carrier protein; KS = ketosynthase; AT = acyltransferase; DH = dehydratase; ER = enoyl reductase; KR = β -ketoacyl reductases; MT = methyltransferase; C = condensation; A = adenylation; PCP = peptidyl carrier protein; TE = thioesterase.

5.2 Post-assembly line modifications

The released product can further be modified by other enzymes, yielding the final product.^[135;136] Possible modifications include methylation, glycosylation, halogenation, and hydroxylation.^[110;135;137] Halogenation and glycosylation are unusual for most myxobacterial compounds, but are known in some cases e.g. in the biosynthesis of chondrochloren (Figure 9) (halogenation), and chivosazol (**19**) (glycosylation).^[29;45;47;75;138] More common are acylations, hydroxylations and methylations.^[43] Often, these final modification steps are essential for the bioactivity of the molecule. Investigations on such enzymes are usually aimed at substrate specificity determination to find out whether they exhibit potential for use as biocatalysts.^[139] Glycotransferases and P450 enzymes have been shown to be useful for such applications.^[140-142] Genetic manipulation, for example, enabled different sugar substrates to be accepted and subsequently attached to the original core structure.^[140]

5.3 Biochemistry of ribosomally-produced peptides

NRPS and PKS are usually encoded by large, highly conserved genes that are easily detectable in the genome. In contrast, genes for ribosomally synthesized peptides are short and not as obvious.

In ribosomal peptide (RP) synthesis, only the 20 proteinogenic amino acids are incorporated, as encoded by the triplet in the mRNA. Normally, a precursor peptide is encoded in front of the peptide sequence, which codes for a recognition sequence for a protease.^[143;144] Further sequences for peptide-modifying enzymes involved in post-translational modifications are often arranged nearby. The general layout of an RP biosynthetic gene encodes a signal peptide, leader peptide, and a core peptide, followed by a recognition sequence (Figure 10). The N-terminal leader and the C-terminal core peptide are crucial for processing. The precursor peptide encodes a peptide leader extension at the N-terminus. The C-terminal side codes for the core peptide, leading to the mature natural compound. The stretch of sequence next to it is called recognition sequence. There are several hypotheses concerning the function of the leader peptide, though it most likely acts as a secretion signal. Several other functions have been postulated; one is that it has a recognition motif for post-translational modification enzymes. Moreover, it may also contribute to the folding of the precursor peptide and stabilize it against degradation. Another idea is that it inactivates the precursor peptide inside the cell. The exact function of the leader peptide likely differs for every compound class.^[145]

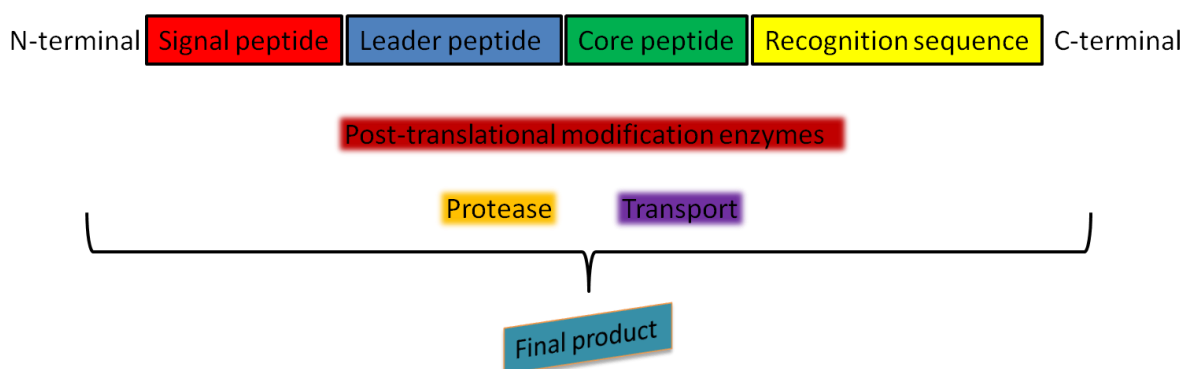


Figure 10: Scheme of a ribosomally produced peptide. The N-terminal leader and C-terminal core peptide are absolutely required. Often, a signal peptide is attached to the N-terminus and a recognition sequence to C-terminus. The ribosomally synthesized precursor peptide can be modified by post-translational modification enzymes. A protease releases the product, which is further transported out of the cell.

Natural products of ribosomal origin are described in literature as bacteriocins, ^[146;147] comprising of the lantibiotics and the microcins. Lantibiotics are described as antimicrobial substances containing intramolecular thioether cross-links found between dehydroalanine, dehydrobutyryne, and cysteine residues (Figure 11). ^[145;148] Lantibiotics are split up into four different subgroups. ^[149;150] And if there is no biological function known, they are mentioned in literature as lantipeptides. ^[149]

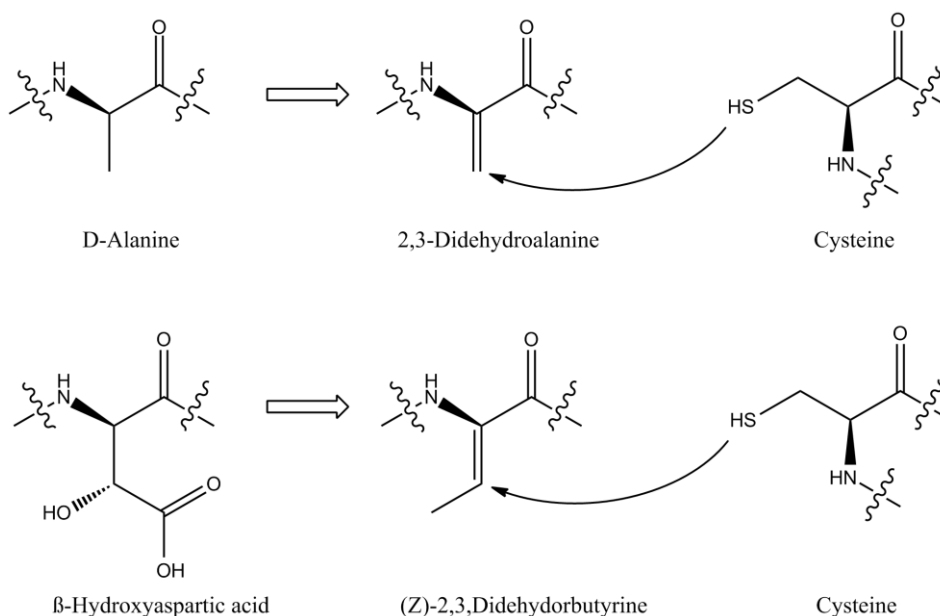


Figure 11: Thioether cross-link formation like it occurs in lantibiotics. At first D-alanine or rather β -hydroxyaspartic acid is dehydrated. Then these dehydro amino acids can form a thioether crosslink together with cysteine.

Microcins are low molecular mass antibacterial peptides from enterobacteria. They are secreted in defense against other prokaryotic species. They generally weigh less than 10 kDa and secretion into the extracellular medium involves ABC (ATP-binding cassette)-type transporters. ^[151]

Other notable RPs are the lasso peptides, thiostrepton ^[152], thiocillin ^[153], and thiopeptide ^[154]. RPs from plants are cyanopeptides and cylotides; from fungi, α -amanitin ^[155] is a common example. Further cyclic peptides called cyanobactins, with similar gene cluster architecture, were recently discovered from cyanobacteria. ^[156;157] With the help of genome mining, several gene clusters responsible for thiopeptides were identified. ^[158] In all these cases, the gene cluster contained a precursor peptide coding for the protease recognition sequence and further sequences for post-translational modification enzymes. The coding sequences are flanked by a leading recognition sequence and motifs for start and stop cyclization signals. ^[143;144]

The cyclic octapeptides patellamides (**25**)^[144] were the first RPs isolated which are not part of the already long-known group of bacteriocins. Following that, trichamide (**26**)^[159], microcyclamide (**27**)^[157] and the microviridins (**28**)^[160] were published. Lately, the pace of discovery of ribosomally-produced natural products has been rapid.^[144;152;155;160-163]

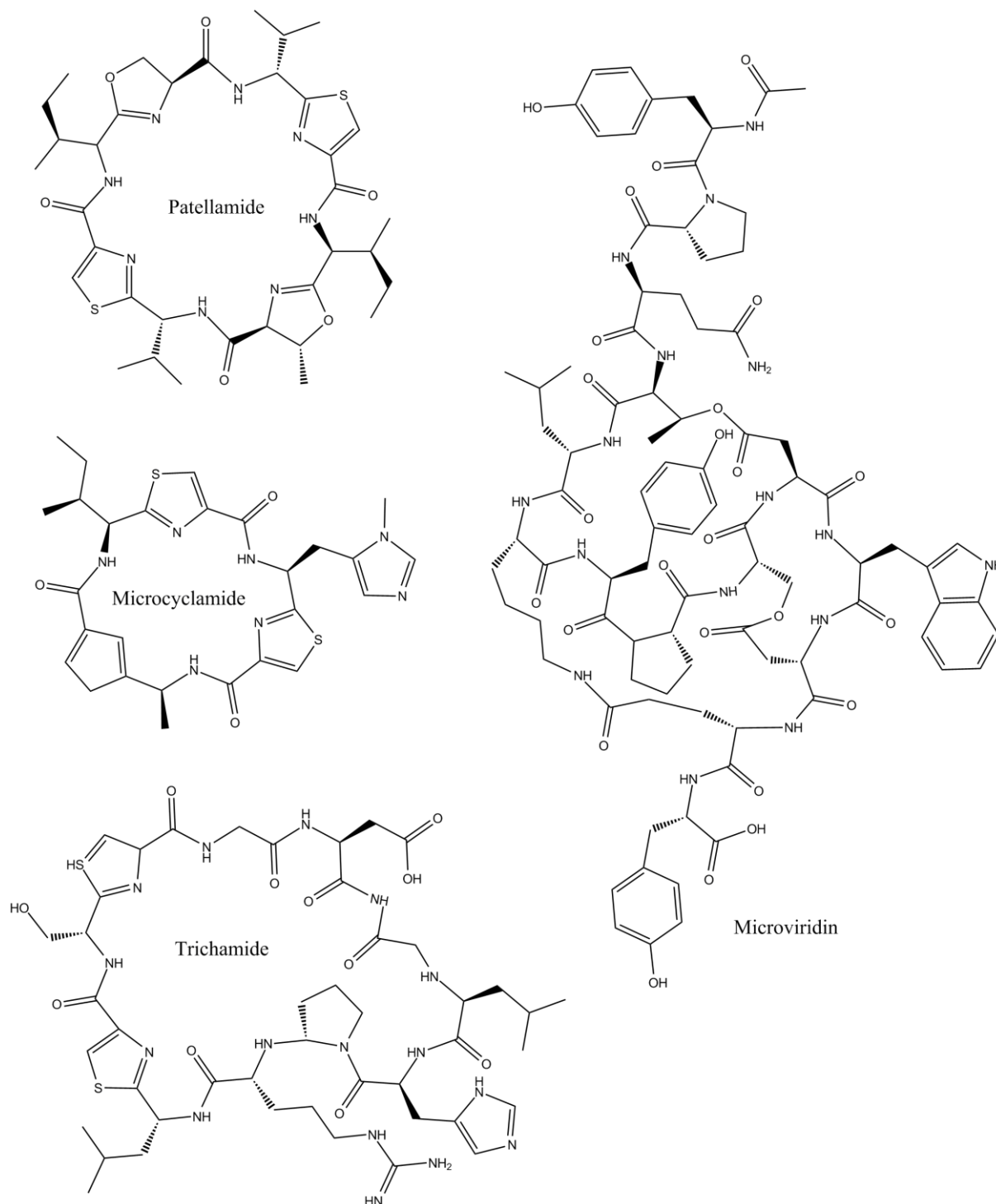


Figure 12: Patellamide (25), Trichamide (26), Microcyclamide (27), Microviridin (28).

6. Discovery of novel myxobacterial metabolites, their structural characterization and elucidation of underlying biosynthetic pathways

As a group, myxobacteria together produce a huge variety of secondary metabolites. Many of these discovered structure types are novel. The species *Sorangium cellulosum* for instance produces a host of antibiotics, many of them reveal unique structural features compared to NP from other microorganisms.^[27;48] Although over 100 different myxobacterial core structures are known up to date it is likely that this is only the surface of the whole myxobacterial secondary metabolism.^[48] To illustrate that the analysis of the genome of *M. xanthus* DK1622 revealed 18 PKS, NRPS, or hybrid gene clusters encoding for secondary metabolites.^[79] However, prior to this study only five published compounds have been characterized and correlated to their corresponding gene clusters.

That means 13 secondary metabolites have not been detected or at least not isolated and characterized. A possible reason for that might be the failure to find or to detect the missing compounds in a screening. This may not be solely due to unfavorable cultivation conditions,^[87] but it could also be a matter of detection sensitivity. Manual data analysis is not able to realize small differences within the data. Therefore, new screening platforms have to be developed with a low detection limit. Furthermore statistical methods can be implemented to decrease the manual burden of analysis.^[51;164]

The detection limit depends on the screening method. Investigations of secondary metabolite via mass spectrometry (MS) for example, require only a few micrograms of substance. MS can provide a wealth of information, allowing the prediction of a sum formula, as well as reveal structural information by fragmentation patterns. Structure elucidation is essential for further studies on the metabolite, like analysis of its biosynthetic gene clusters or for the chemical synthesis. However, larger amounts are required to obtain sufficient information for complete structure elucidation by other methods.

Together with the MS information nuclear magnetic resonance (NMR) spectroscopy or X-ray crystallography are used for complete structure elucidation. Compared to MS the NMR spectroscopy and the X-ray crystallography have the advantage that these are non-destructive methods. One distinct benefit of NMR spectrometry compared to X-ray crystallography is that it does not require any crystals, which need even more material.^[165] Furthermore NMR can be coupled to HPLC systems.

However if only small amounts of a compound are available sensitive methods are needed. To increase the sensitivity of NMR spectroscopy, the magnetic field can be increased or the probe can be modified. As changes in the magnetic field are limited due to technical reasons, it is more practical to change the probe. Capillary and micro-cryoprobes are valuable developments that help to increase the sensitivity in NMR spectroscopy. These probes result in a 10 – 20-fold increase in mass sensitivity compared to a conventional 5 mm probe. Of course one reason for the enhancement is a more concentrated sample due to less solvent (5 mm tube approx. 600 μ L; 1.7 mm tube approx. 35 μ L and 1 mm tube approx. 7 μ L). Another reason is that nearly all electronic noise is eliminated with the cryoprobes as they are chilled to 20 K. ^[165]

Typically NMR data interpretation leads to the correct molecular structure of a secondary metabolite. But if the interpretation of the NMR data is difficult *in silico* prediction of the structure might be helpful. For such prediction the biosynthetic gene cluster sequence has to be available. *In silico* Modules and domains will be assigned and substrate specificities will be predicted. But as these predictions are not always possible or reliable, finally the structure of the metabolite is still needed.

Another alternative to identify unknown structures of compound which are only available in insufficient amounts is to increase their production rate by genetic engineering, followed by subsequent compound isolation. For doing genetic engineering the knowledge of the biosynthetic gene cluster is needed. Such information is also of increasing relevance for all biologically active compounds as sufficient yields of product are often not achievable by total synthesis. As well, the underlying enzymatic functions may be of interest to create novel derivatives by manipulation of the biosynthetic pathway. ^[91] Often these compounds have the same mode-of-action but differ in efficacy. It may also be advantageous to transfer the complete biosynthetic gene cluster into a new host with easier cultivation properties. ^[92] For example, a good heterologous expression host may be simpler to handle or has faster growth than the native organism and therefore it can be used to produce larger amounts of a natural compound.

Recently, more and more gene cluster information is obtained by sequencing complete genomes. The genome is analyzed *in silico* for biosynthetic genes, followed by targeted gene inactivation within a cluster of interest. Subsequent comparative analysis of secondary metabolome between wild type and mutant leads to a correlation of the biosynthetic genes. A lack of metabolite production indicates the inactivation of the correct gene cluster.

As costs for genome sequencing decrease, the technique of genome mining is more often applied.^[166-168] Within the next few years an enormous boost of sequenced microorganisms is expected. By August 2009, around 800 bacterial genomes had been sequenced and about 700 incomplete ones have been deposited in the publicly accessible databases.^[145]

7. Outline of this work

The purpose of this thesis was to investigate the secondary metabolism of myxobacteria, with major focus on the search for novel natural products from these microorganisms. The goal was to discover novel secondary metabolites and to isolate and characterize them with spectroscopic methods. The secondary metabolite profiles of *M. xanthus* strains should be screened for novel substances via HPLC-MS (high performance liquid chromatography coupled with mass spectrometry).

As bacteria produce secondary metabolites via several different biosynthetic pathways, another goal was to examine how the compounds are generated, which genetic machinery is involved, and what possibilities there are to manipulate them to obtain unnatural derivatives.

Chapter 1 explores the topic of the bicyclic tetrapeptide cittilin (RP66453, Figure 13) produced by *M. xanthus*.^[169-171] Investigations of further *M. xanthus* isolates revealed that the majority of the strains produce cittilin.^[51] Its production is described from an *Actinomycetes* strain, too.^[172] Two cittilin derivatives are known, differing from each other by a methylation of the hydroxyl group of the first tyrosine. But the structural striking features of cittilin are the two aryl crosslinks between the tyrosine rings; one is a carbon link, but the other an ether connection. Cittilin B has been described as a neurotensin antagonist (neurotensin receptor from guinea pigs $IC_{50} = 30 \text{ mg / mL}$).^[170;172] Therefore, it might be used as a lead structure for the development of antipsychotic pharmaceuticals.

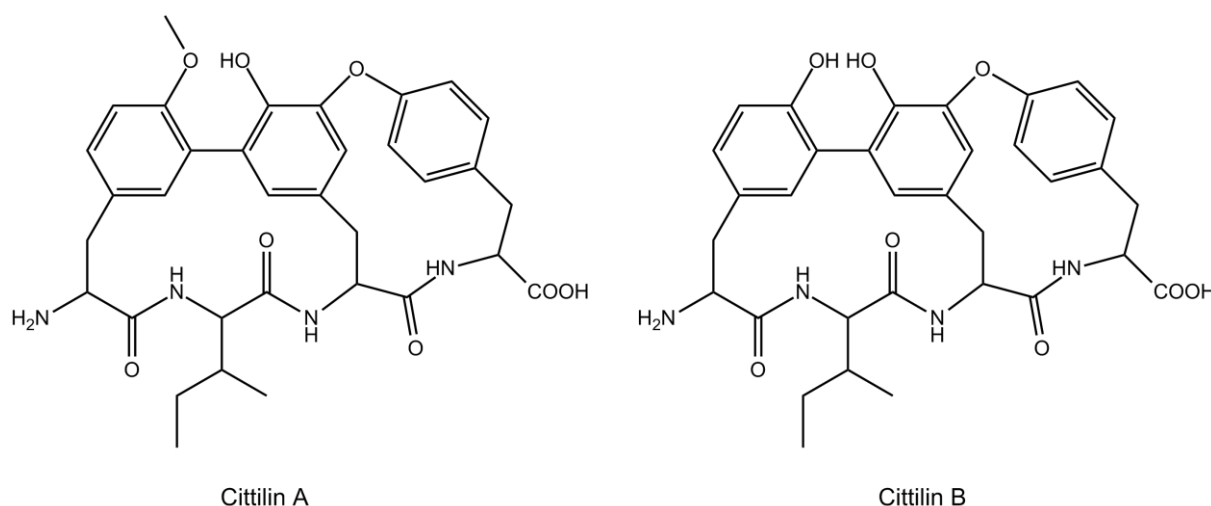


Figure 13: Cittilin A (29) and B (30), secondary metabolites produced by many *M. xanthus* strains.

The genetic origin of cittilin remained elusive so far. The question of an NRPS or a ribosomal biosynthetic locus emerged. Therefore, the first goal was to identify the genes responsible for

the citrillin pathway. The genome of DK1622 and DK896 were screened for the genetic origin of citrillin and the biosynthetic pathway was assessed.

Efforts were also made to exploit the biosynthetic machinery for the generation of novel citrillin derivatives by genetic engineering. The results were monitored by high performance liquid chromatography coupled to high resolution mass spectrometry (HPLC-HRMS) and HRMS/MS.

M. xanthus was shown to be a multi-producer of secondary metabolites. For DK1622 it was reported that it exhibits further gene clusters for unknown secondary metabolites. In a screening of 98 *M. xanthus* strains several unknown secondary metabolites were discovered but neither their structure nor their corresponding biosynthetic genes were known. ^[51] Therefore, the goal was to isolate and structurally characterize some of these unknown *M. xanthus* compounds with all structural variants. By feeding of isotope labeled precursors, HR-MS and HR-MS/MS analysis, in combination of NMR-based experiments, the novel compounds were characterized. The data from different methods were pooled to generate a molecular “dossier” used in suggesting a final structure.

Beside the structure, the correlation of the biosynthetic gene cluster was of special interest. After identification of the responsible genes *in silico* gene analysis, genetic manipulation of producers and feeding experiments led to further understanding of biosynthetic pathways. Furthermore it is possible to optimize the production rate by promoter insertion. That led to sufficient yields for isolation of compounds that were only produced in small quantities.

The novel myxoprincomides and their biosynthetic pathways are illustrated in **chapter 2**. Myxoprincomides of DK897 are described in **chapter 2A** while the analogous version of DK1622 was investigated in **chapter 2B**.

Chapter 3 deals with chondrochloren, a secondary metabolite with weak antibiotic activity produced by *Chondromyces crocatus* Cm c5. ^[77] Its biosynthesis has been described by Rachid et al. ^[133] But the exact formation of the final product remained elusive. Therefore, S. Rachid has inactivated the oxidative decarboxylase CndG in the chondrochloren gene cluster. The obtained mutant produced carboxylated chondrochloren A and B (pre-chondrochlorens) instead of the chondrochloren itself (Figure 14). This so called pre-chondrochloren were isolated so that it could be used for *in vitro* experiments to identify the molecular mechanisms underlying the decarboxylation.

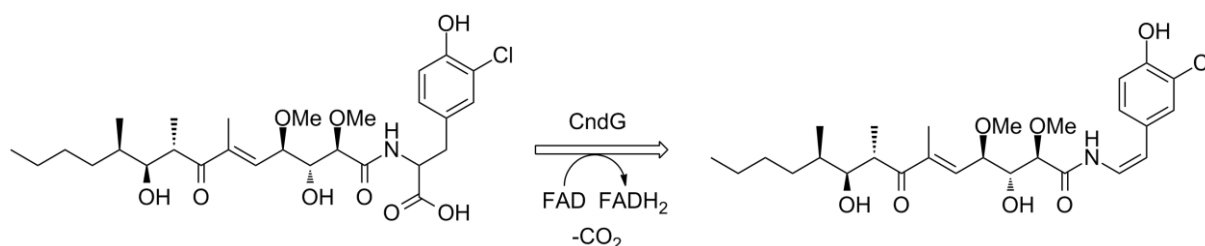


Figure 14: Pre-chondrochloren A (31, left) and B are converted by CndG to Chondrochloren A (32, right) and B.

Novel spirangiens are presented in **Chapter 4**. These secondary metabolites are produced by *Sorangium cellulosum* So ce90, best known as a producer of epothilone.^[173;174] Spirangien (Figure 15) has been previously shown to exhibit antifungal and highly cytotoxic properties. Spirangien contains a polyketide backbone with a spiroketal in the centre and a chromophore consisting of five conjugated double bonds.

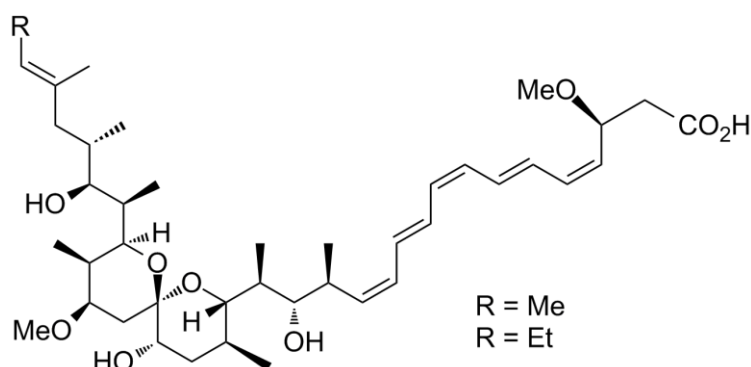


Figure 15: Spirangien A (33, R = Me) and Spirangien B (34, R = Et) produced by *Sorangium cellulosum* So ce90.

In the course on the identification of the biosynthetic gene clusters by Frank et al.^[175], a spontaneous mutant occurred. It lacked spirangien production but four novel derivatives with lower molecular weight were detected. To reveal the molecular mechanisms underlying in the production of the novel spirangien derivatives these light- and oxygen-sensitive compounds were isolated and characterized by NMR experiments.

Chapter 1

Cittilin, an unusual ribosomally produced secondary metabolite from *Myxococcus xanthus*

Abstract

Cittilins are bicyclic tetrapeptides which are produced by the majority of known *Myxococcus xanthus* strains. Unlike many other myxobacterial secondary metabolites the cittilins are no products of NRPS-based biosynthesis, instead they are formed via the bacterial ribosome. In this work the biosynthetic machinery underlying the formation of cittilins is investigated. Heterologous expression of the *cit* operon in the myxobacterium *S. aurantiaca* confirms the ribosome-dependent biosynthetic route, and knockout experiments provide evidence for the involvement of a methyltransferase and a P450 oxidoreductase in the post-translational modification of cittilins. Furthermore, the possibility to generate novel cittilin derivatives by directed mutagenesis of the gene encoding cittilin precursor is demonstrated.

Introduction

Myxobacteria are a rich source of natural products. More than 100 core structures of secondary metabolites from myxobacteria are known.^[38] The best studied myxobacterium is *Myxococcus xanthus* DK1622, which was intensely investigated because of its intriguing biological characteristics. But DK1622 has also been shown to produce several secondary metabolites: the myxalamides^[82], the myxochelins^[83], the myxovirescins^[84], the myxochromides^[85], DKxanthenes^[86] myxoprincomides^[176] and cittilins. The cittilins A (**35**) and B (RP-66453; **36**) are produced by more than half of all known *Myxococcus xanthus*

strains investigated in a *M. xanthus* screening ^[51] and by some *Streptomyces* sp.. The cittilins were originally isolated and their planar structures solved by Trowitzsch-Kienast and coworkers in the course of their work on the myxobacterial metabolite saframcin. ^[169-171] The structure of cittilin B (**36**) has previously been reported from *Streptomyces* strain 9738, where it was named RP-66453 (**36**). Studies have shown that **36** binds specifically to the neurotensin receptor from guinea-pig ($IC_{50} = 4.86 \times 10^{-5}$ M). ^[170;172] This drug target is very important for the development of antipsychotic pharmaceuticals. Chemical modification of the molecule indeed yielded in an effective and patented neurotensin antagonists, which may be useful for treatment of Alzheimer's and Parkinson's diseases. ^[177] Two natural cittilin derivatives are known, **35** and **36**, the difference being a methyl substitution at the hydroxyl group of the first tyrosine ring. The cittilins are bicyclic tetrapeptides, containing three tyrosine residues and one isoleucine amino acid residue. The tyrosine rings are connected via a C-C bridge and a C-O-C linkage, similarly to the substructure found in balhimycin. ^[178] Thus, the cittilins belong to a group of macrocycles with aryl – aryl and aryl – aryl - ether bonds like in vancomycin (**37**), kistamine ^[179] and chloropeptin ^[180]. It has been shown that formation of these aryl bonds can be catalyzed in nature by cytochrome P450 enzymes. The biosynthesis of vancomycin (**37**) constitutes a well-studied example for both types of bond-forming reactions. ^[181]

The absolute stereochemistry of the cittilins was not solved after the initial isolation and NMR analysis by Trowitzsch-Kienast et al. nor by Helynck et al.. Stereochemical determination was accomplished later by total synthesis; all five carbonic stereo centers and the atropisomerism of the biaryl axis have been determined by Boger et al. ^[182-185] The absolute configuration of (**36**) has been assigned for the carbonic stereo centers as all *S* configuration which means that all amino acids are in the L-configuration and the molecule axis has a *R* configuration.

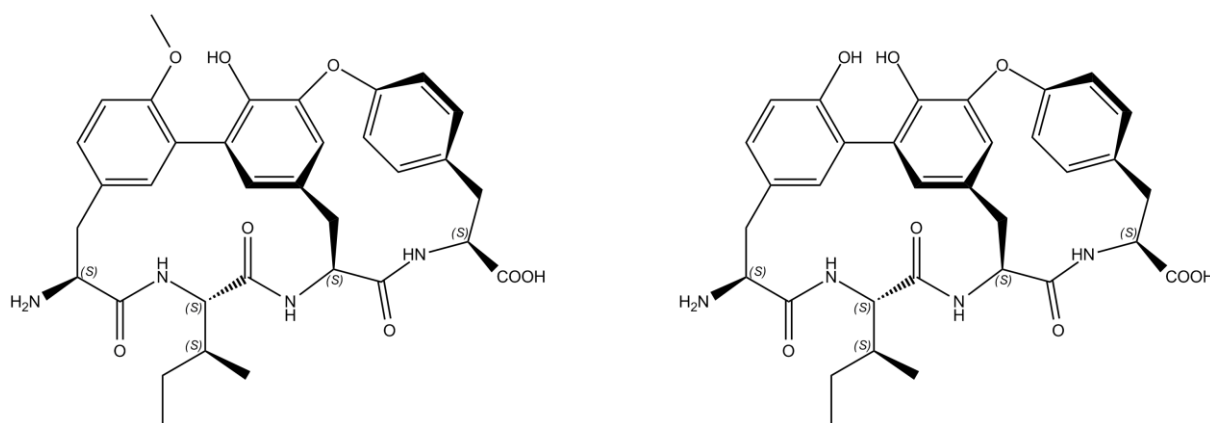


Figure 16: 35 Cittilin A

36 Cittilin B (RP-66453)

Many secondary metabolites produced by myxobacteria or other microorganisms are peptides. The structural class of peptides is biosynthetically often derived from non-ribosomal peptide synthetase (NRPS) assembly lines.^[105] A typical example for such an NRPS product is the antibiotic gramicidin.^[186;187] It is a linear peptide made from combinations of up to 15 D and L amino acids. However, in a growing number of recent reports it has been shown that such highly modified peptides like the patellamides (38)^[144] and the microviridins (28) can also be produced by a ribosomal pathway.^[157;160] Exhibiting various tailoring enzymes for peptide modifications like cyclases, methyltransferases to give some examples. The ribosomal peptide (RP) synthesis incorporates only the 20 proteinogenic amino acids which are encoded by the three letter code. The RP synthesis is the general way for the primary metabolism to synthesize peptides. For the secondary metabolism a typical RP gene cluster contains a signal peptide, a leader peptide, a core peptide followed by a recognition sequence and further modifying enzymes.^[143;144] Such biosynthetic architecture is known from microcins and lantibiotics, both are two long known classes of ribosomal peptides. Another big group comprises the recently published cyanobactins produced by cyanobacteria.^[156;157]

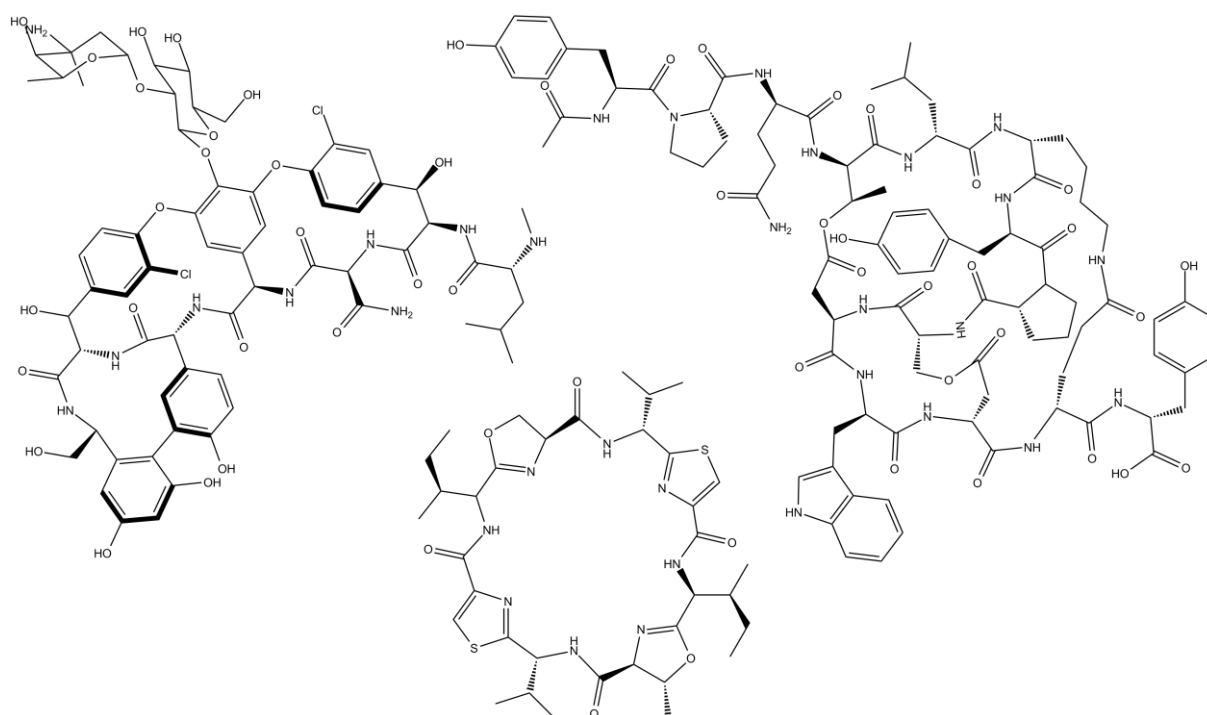


Figure 17: Vancomycin (37) Patellamide (38) Microviridin (39)

In this work the cittilins A (35) and B (36) are revealed as representatives of a novel class of ribosomal peptides. In comparison to the previously known groups, the cittilins exhibit a fascinating structure consisting of a bicyclic cross-linked tetrapeptide. The cyclization is unusual as it is neither formed by a typical peptide – peptide backbone ring closure as found

in **38** nor does it incorporate cysteine residues as in the case of the lantibiotics.^[145;148] The C-C and C-O connectivity in the cittilin structure thus resembles more the non-ribosomally generated vancomycin (**37**).^[181]

Here it could be shown by heterologous expression that cittilin is biosynthetically produced by a ribosome-dependent pathway. Cittilin is the first known myxobacterial compound for which the ribosomal synthesis could be demonstrated. Aside from a very short sequence region, only one P450 enzyme and one methyltransferase is required to generate cittilin. Efforts were also made to exploit the cittilin biosynthetic machinery for the generation of novel derivatives by genetic engineering.

Results and Discussion

During the LC-MS guided workup of an *M. xanthus* DK897 extract a compound giving rise to a signal of $m/z = 631.2754$ $[M+H]^+$, which matched the molecular formula of Cittilin A (calc. for $C_{34}H_{39}N_4O_8$, 631.2762 $[M+H]^+$). Its MS and MS² spectra are depicted in Figure 18. The corresponding peak showed a UV maximum at 290 nm (Supplementary figure 4) and was present in fair amounts to be isolated. The extract was partitioned by a methanolic Sephadex[®] LH20 column, followed by gradient silica gel column. A final methanol - water (3:2) Sephadex[®] LH20 column yielded the pure compound. The structure was subsequently solved by NMR spectroscopy (data presented in Table 1). A comparison of the final structure with literature data revealed the compounds as the known cittilins A and B.

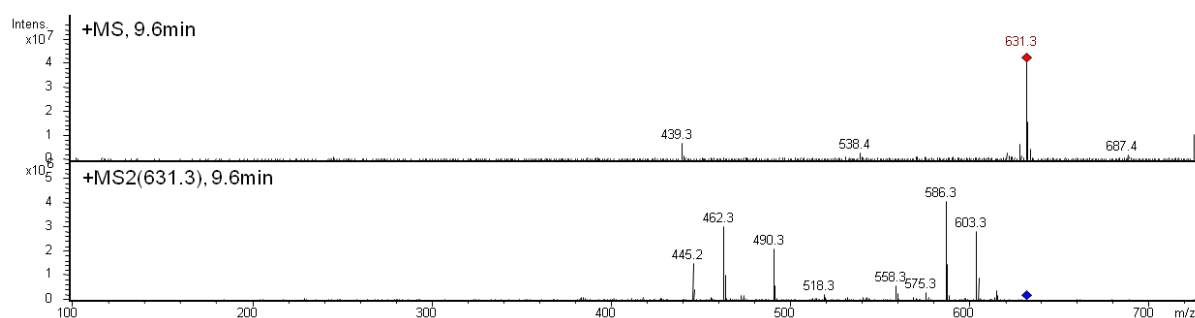


Figure 18: MS spectrum of cittilin A (top) and its MS² spectrum (bottom).

Amino acid	Position	Cittilin A ^1H (<i>J</i>) ppm (Hz)	^{13}C ppm	Cittilin B ^1H (<i>J</i>) ppm (Hz)	^{13}C ppm	m
Tyrosine 1	1-CO		169.9		170.1	
	2-CH	4.21	53.8	4.08	54.1	m
	3a-CHH	3.46	36.7	3.41	37.6	m
	3b-CHH	3.05 (6 / 15)	36.7	2.99 (6 / 15)	37.6	dd
	4-C		129.7		128.0	
	5-CH	6.92	137.0	6.90	137.0	m
	6-C		125.9		125.6	
	7-C		157.9		155.9	
	8-CH	7.21 (2 / 8)	131.7	7.19 (2 / 8)	131.7	dd
	9-CH	7.01 (9)	112.0	6.98 (8)	111.8	d
	O-CH ₃	3.88	56.3			s
Isoleucine	1-CO		169.9	-	171.7	
	2-CH	4.20	60.1	4.16	60.0	m
	3-CH	1.77	37.1	1.77	37.1	m
	4a-CHH	1.61	26.3	1.62	26.4	m
	4b-CHH	1.27	26.3	1.26	26.4	m
	5-CH ₃	0.95 (7)	15.7	0.96 (6)	15.9	d
	6-CH ₃	0.92 (7)	11.0	0.92 (7)	10.9	t
Tyrosine 2	1-CO		173.0		171.5	
	2-CH	3.68	59.8	3.68	59.8	m
	3a-CHH	3.10 (2 / 14)	40.7	3.16 (2 / 13)	40.7	dd
	3b-CHH	2.47 (4 / 13)	41.0	2.47 (4 / 14)	40.9	dd
	4-C		136.6		135.4	
	5-CH	5.78 (2)	120.7	5.80 (2)	120.8	d
	6-C		154.2		152.6	
	7-C		145.2		143.3	
	8-C		129.7		129.6	
Tyrosine 3	9-CH	6.62 (2)	126.3	6.63 (2)	126.4	d
	1-CO		176.8		175.2	
	2-CH	4.30	54.2	4.26	54.7	m
	3a-CHH	3.61	37.3	3.61	37.4	m
	3b-CHH	2.98 (7 / 14)	37.3	2.94 (m)	37.4	dd
	4-C		135.5		133.9	
	5-CH	7.22	135.4	7.19	135.4	m
	6-CH	6.83 (3 / 8)	124.8	6.79 (3 / 9)	124.6	dd
	7-C		164.5		162.5	
	8-CH	7.47 (2 / 8)	127.0	7.45 (3 / 9)	127.0	dd
	9-CH	7.35 (2 / 8)	130.7	7.36 (2 / 8)	131.0	dd

Table 1: ^1H and ^{13}C NMR spectral data of the cittilin A and B; recorded at 500 and 125 MHz in CD₃OD; d, doublet; dd, doublet of doublets; s, singlet; m, multiplet.

Aiming at the analysis of intra-species metabolite diversity, a worldwide LC-MS based screen of 98 *M. xanthus* isolates has been performed previously and revealed altogether 69 strains as producers of cittilins.^[51] LC-MS analysis of *M. xanthus* DK1622 extracts highlighted this strain as a cittilin producer as well. Therefore we set out to search the genome sequence data available for DK1622^[79] for a potential cittilin biosynthesis gene cluster. We initially performed a query for typical NRPS genes based on a retro-biosynthetic analysis of cittilin. If cittilins were produced by a typical NRPS pathway, the gene cluster would be expected to contain four modules: a loading module and three extension modules with the C-terminus containing a thioesterase (TE) domain. The adenylation (A) domain specificities should then match the incorporation of three tyrosine units and one isoleucine. Furthermore a methyltransferase (MT) and genes facilitating the cyclization would be expected. Aromatization and heterocyclization processes are commonly observed for natural products and some well described examples exist.^[188] In the case of PKS or NRPS-directed biosynthesis dedicated cyclization domains effect the ring formation. When a peptide is cyclized at the end of an assembly, the intramolecular cyclization is usually catalyzed by a thioesterase. Aryl – aryl linkages as they are found in the vancomycin type of peptides, however, have been shown to be introduced by the action of P450 enzymes.^[181;189]

Yet another cyclization mode is found in the lantibiotics, where cyclizations happen in two sequential steps. Following the removal of a hydroxyl group of a threonine or serine by dehydration, a second cyclization step links the unsaturated group with a thiol group of a cysteine. These dehydration cyclization processes are often performed by a bifunctional synthetase such as LanM.^[190]

All A domains in NRPS clusters in the genome of DK1622 were analyzed, but no NRPS cluster matching the retrobiosynthetic prediction could be spotted within the genome. Therefore we reasoned that an alternative biosynthetic pathway should exist. As cittilin contains only proteinogenic amino acids we considered the possibility that it could be assembled in ribosomal manner with additional post-translational modifications. Therefore, the genome was searched for a sequence encoding for ribosomal peptide biosynthesis like for bacteriocin-like ribosomal biosynthesis.^[191] To search for a potentially matching genetic region the degenerated genetic code for the amino acid sequence YIYY was used as a query term on both strands in the DK1622 genome. Exactly one small open reading frame (orf) of 231 bp length was detected which was followed by genes encoding for a P450 and a methyltransferase. Additionally, a set of ABC transporter genes was identified upstream of

this region. Such short orfs are difficult to detect during automatic genome annotation, as one typically expects much longer orfs for regions encoding biosynthetic pathways. Sequence analysis of the short orf reveals a hypothetical leader peptide encoding region followed by the 12 bp stretch encoding the YIYY core peptide, terminated by a stop codon. This arrangement is highly unusual for ribosomal biosynthetic gene clusters, since the core peptide region usually continues with a recognition sequence for processing the peptide. But no typical motifs like the double glycine motif ^[190] for a recognition site for a processing protease could be located neither before nor after the core peptide.

Therefore the region encoding the cittilin leader peptide was further investigated. For in-silico statements more sequence data from four strains cittilin producing strains (DK1622, DK897, A47, ST200611) were analyzed. None of them contained typical cleavage motifs, known from other RP, like a GG or GA ^[190] within the precursor peptide.

Among the cittilin producing strains with known cittilin gene sequences, one strain (ST200611) was found where the hypothetical leader peptide sequence is different from the other strains. In ST200611 the leader peptide is 310 bp long instead of 231 bp in the other strains DK897, DK1622 and A47. These three other strains have exactly the same length and the identity on the protein level is 95 % which is a highly conserved region (Figure 19).

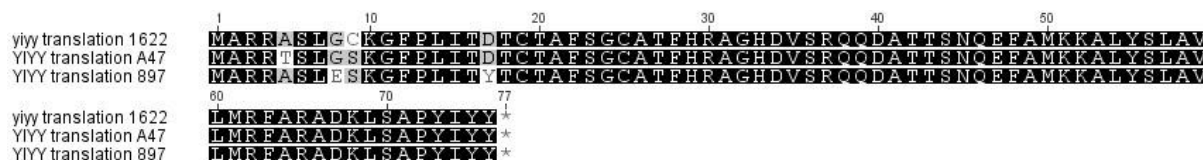


Figure 19: Cittylin leader peptide alignment of *M. xanthus* DK1622, A47 and DK897.

Only within the first 17 amino acids are some modifications. Compared to the sequence of ST200611 the identity is much less (Figure 20). ST200611 contains several stop codons within the sequence. As all the other strains have very high similarity it could be a matter of sequencing errors. Therefore the region was re-sequenced via amplification by PCR. The sequence analysis revealed that the original sequence is indeed correct. Thus, the shortest encoding region without a stop codon would be 78 bp long. That means that the other encoding sequence before that region is apparently not absolute necessary. Consequently, a maximum of 26 amino acids (EFAMKKALYSLAVLMRFARADKLSAPYIYY* see Figure 20) could be already sufficient to form a functional cittilin leader peptide.

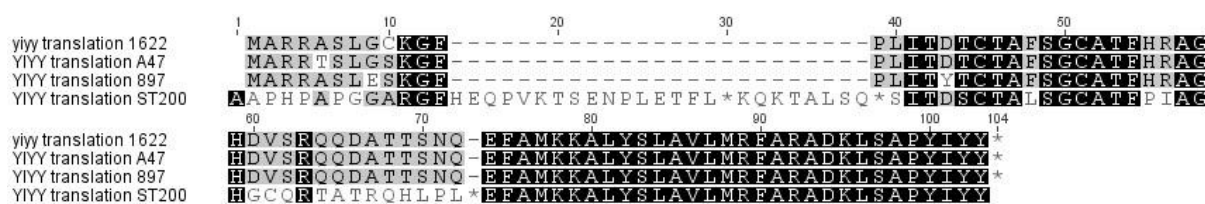


Figure 20: Citalin leader peptide alignment of DK1622, A47, DK897 and ST200611.

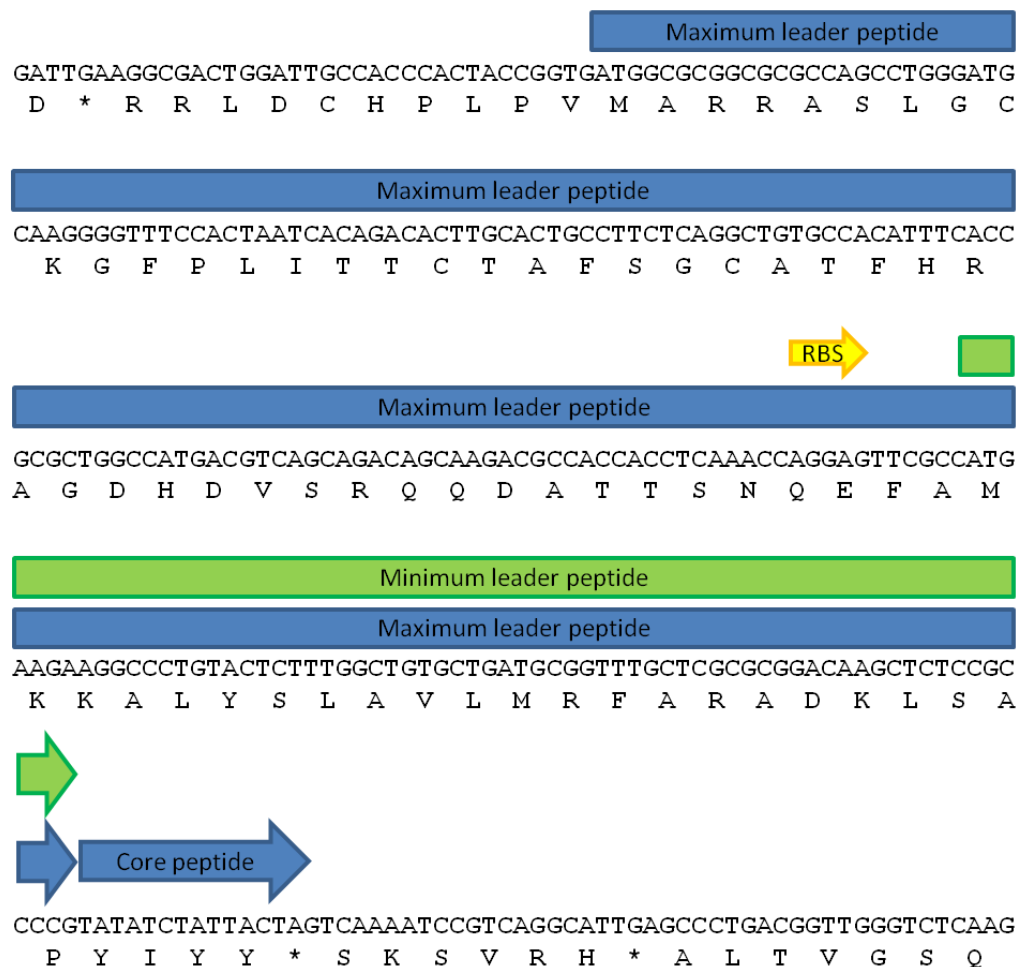


Figure 21: Orf of YIYY region in genome of *M. xanthus* DK1622, maximum leader peptide (216 bp), ribosomal binding site (RBS), minimum leader peptide (69 bp), core peptide (12 bp) followed by a stop codon.

Immediately upstream of these 78 bp stretch a possible ribosomal binding site (RBS, AGGAG) is present. That means that the maximum leader peptide can be reduced to 69 bp (Figure 21). But no further statement can be made about the function of the leader peptide. But it suggests that this unique sequence is involved in the processing of the tetrapeptide, or it stabilizes the precursor peptide against degradation.^[145] To get more information about it could be useful to analyze the leader peptide of the other known cittilin *M. xanthus* producers,

as several are noted in a big screening project ^[51] e.g.: A2, A66, A98. Of course the sequence of *Streptomyces* producer strains would be very interesting as well, but finally it has to be proven by biochemical experiments.

Upstream of the leader peptide region a set of ABC transporter genes are located (four orfs). Genes encoding possible modifying enzymes are located directly downstream of the core peptide sequence: a P450 and a methyltransferase. The predicted orfs in the vicinity of the core peptide locus are shown in Figure 22.

The orf encoding a putative P450 enzyme starts only a few base pairs after the core peptide region. A protein blast search on “<http://blast.ncbi.nlm.nih.gov/Blast.cgi>” identified it as cytochrome P450 family protein, and the closest relatives are from bacterium *Ellin514*, *Nostoc punctiforme* PCC 73102 and *Bacillus halodurans* C-125 as. The typical function of P450 enzymes is to catalyze regiospecific and stereospecificly the oxidation of non-activated hydrocarbons. ^[192] The next open reading frame encodes for an *S*-adenosylmethionine-dependent methyltransferase (MT). ^[193] Those MTs are typically able to attach methyl groups to hydroxyl groups. Here the best similarity to other methyltransferases is to genes from *Methanoculleus marisnigri* JR1, *Frankia* sp. EUN1f and *Streptomyces* sp. AA4.

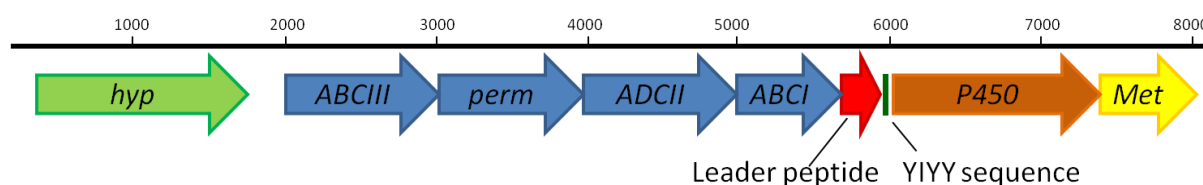


Figure 22: Orfs around the citrillin core peptide: *hyp* = FAD-binding domain protein, monooxygenase; *ABC III* = periplasmic ferric siderophore-binding protein; *perm* = permease protein; *ABC II* = permease protein; *ABC I* = ATP-binding protein

The P450 enzyme could take part in the cyclization of citrillin. It has two C-C linked tyrosine rings and two C-O-C bound tyrosines. These structural fragments are analogous to those found in vancomycin (37). In the vancomycin (37) biosynthesis the oxygenases OxyA, -B and -C are responsible for the connection. OxyA and OxyB are in charge of C-C connection and OxyC for the phenolic coupling. ^[181;189] These oxygenases are P450 enzymes and they perform the ring closure during the formation of the backbone. ^[189] Modifying enzymes in the RP biosynthesis are comparable enzymes and often similar to the ones in the NRPS biosynthesis. ^[194] Therefore it is most likely that the two rings of citrillin are formed via a similar mechanism. The methylation might be performed by the methyltransferase which is encoded adjacent to the P450.

In order to prove the involvement of the putative P450 enzyme and the methyltransferase in the biosynthesis of cittilin, knockout mutants were created by insertional inactivation in DK1622. For the P450 knockout mutant a single crossover experiment was performed. An internal fragment of approximately 500 bp was amplified by PCR, cloned into the pCR2.1-TOPO vector (Invitrogen) and the resulting plasmid was introduced into the strain like it is often described in literature.^[195-198] In the case of the methyltransferase which has a size of 615 bp a stop codon and a frame shift were additionally added into its DNA sequence to ensure that the gene in the resulting mutant is inactive.

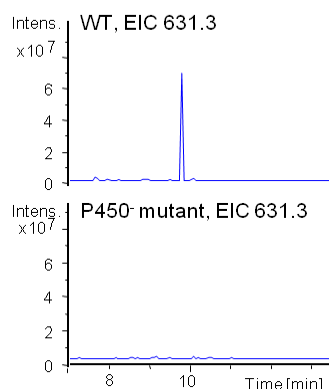


Figure 23: Extracted ion chromatogram for the mass of cittilin from crude extracts of wild type (above), P450- mutant (below).

The P450 mutant showed complete abolishment of cittilin production as judged on the basis of LC-MS measurements (Figure 23). However, no evidence for production of the unmodified tetrapeptide YIYY was found. In order to exclude the possibility that failure of detection was due to insufficient sensitivity during the analysis, samples were compared to a synthetic reference for YIYY, and thereby the absence of the unmodified tetrapeptide was confirmed. Reasons for that might be degradation, or no secretion out of the cell, or unmodified tetrapeptide is only separated after it was processed by the P450. A second P450 enzyme could not be detected directly in the region of the cluster. But an oxidoreductase was detected 2.7 kbp downstream of the methyltransferase (refer to Supplementary figure 10). Investigations of its effect on cittilin biosynthesis are ongoing. As well as the influence of a peptidase upstream of this oxidoreductase. However the heterologous expression showed that both linkages can be build if only the P450 enzyme is present.

This is surprising as the cross links are different, the first one is a C-O linkage and the second one a C-C bond. Furthermore, in the case of the vancomycin (**37**) biosynthesis two different P450 enzymes are known to be responsible for these two reactions.^[181;189] Therefore it cannot

be excluded that another enzyme encoded somewhere else in the DK1622 genome might do this task.

The outcome of the methyltransferase inactivation experiment was not as unambiguous; here cittilin A production was still noticeable, albeit with a strong reduction in yield. It may be concluded that other methyltransferases in the genome are able to complement its function in the knockout mutant. That is known from other natural products, like in the case of myxochelins, where other methyltransferases encoded in the genome complement the knockout of the methyltransferase in the myxochelins biosynthetic gene cluster.^[199]

In order to narrow down which genes are essential for cittilin biosynthesis, a heterologous expression of the cittilin gene locus was preformed. The gene cluster was initially amplified by two PCRs (1st leader and core peptide; 2nd P450 and methyltransferase), subcloned and then stitched together in the final expression vector pCK (Supplementary figure 1) to yield the plasmid pCK-Cittilin. The pCK plasmid^[90] contains an Mx8-phage derived attP site and an intP site, and thus is able to integrate into the phage attachment site of various myxobacteria. The cloned cittilin cluster in the plasmid pCK-Cittilin is under the control of the T7A1 promoter. T7A1 has been shown to successfully drive expression in myxobacteria.^[200] As host strain *Stigmatella aurantiaca* Sg a15 was chosen, because it was expected that this host would supply similar cellular machinery^[92] to *M. xanthus* and evidently no cittilin gene cluster is present in its genome. After transformation of the pCK-Cittilin plasmid into *S. aurantiaca* Sg a15, the heterologous expression clones were cultivated and extracts analyzed by LC-HR-MS measurements. Cittilin production could be detected as judged on the basis of high-resolution LC-MS data and tandem-MS analysis (Figure 24). Retention time and fragmentation pattern are clearly identical with those recorded from cittilin of the natural producer strain DK897 (Figure 24 and Figure 18) although the observed production level in the heterologous expression experiment was rather low.

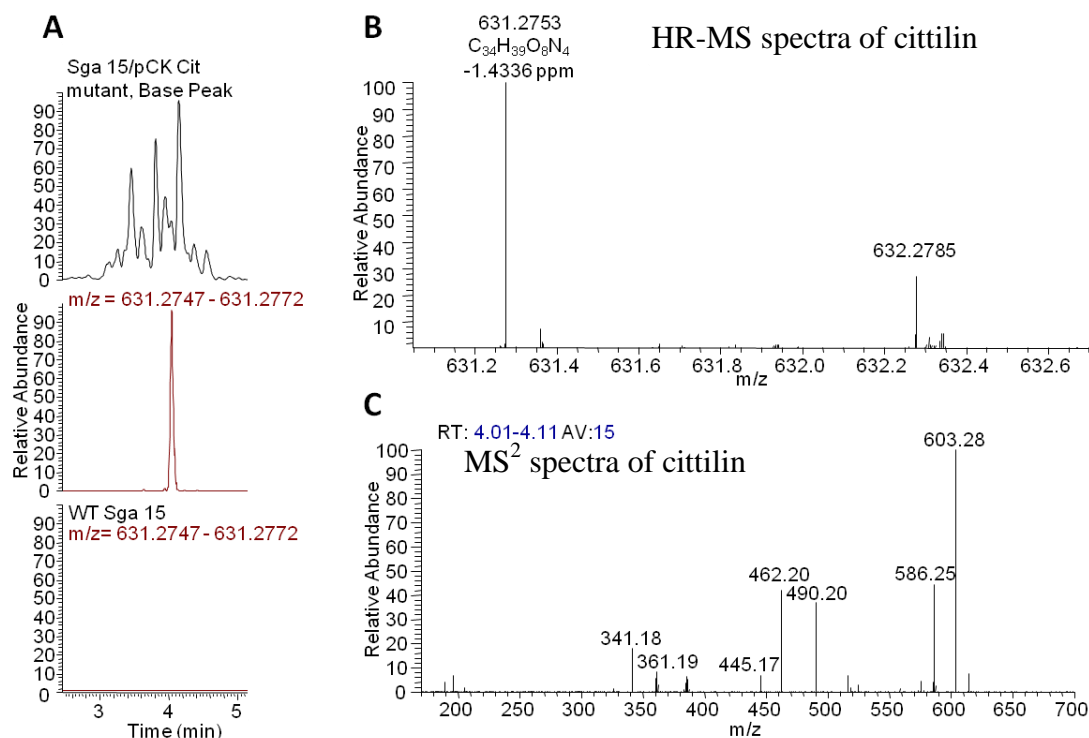


Figure 24: **A:** top base peak chromatogram of an extract from *S. aurantiaca* Sg a15 mutant comprising the cittilin genes, middle extracted ion chromatogram (EIC) of the cittilin of the same extract, bottom: EIC of cittilin of Sg a15 wild type extract. **B:** monoisotopic mass of cittilin produced by the Sg a15 mutant at the same retention time as the cittilin from DK897. **C:** Fragmentation pattern of cittilin produce by the Sg a15 mutant.

This heterologous expression of cittilin shows that the essential genes for cittilin biosynthesis have been identified. According to the heterologous expression results, fundamental functions are the leader peptide, the core peptide, a P450 enzyme and a methyltransferase.

In comparison to other known gene clusters of ribosomal peptides, the cittilin gene cluster exhibits some striking differences. For example, the coding sequence for a recognition site for a protease could not be discovered, moreover there are currently no detectable motifs for start or stop cyclization. ^[143;144] In the case of the cyanobactins the core peptide is flanked by putative start and stop recognition sequences. ^[156] Furthermore, normally the leader and the core peptide are separated by a GG or GA cleavage motif. ^[190] This indicates that a different recognition sequence is used in case of cittilins. Whether the leader peptide has a role to stabilize the precursor peptide against degradation or inactivates it inside the cell remains elusive. In general it has been noted that the exact function of leader peptides differs for every compound class. ^[145]

Besides *Stigmatella aurantiaca* Sg a15, *M. xanthus* A47 was tested as heterologous host for expression experiments of cittilin as it is a close relative to the native host, DK1622.

Phenotypically A47 does not produce cittilin, but compared to the other heterologous strains *Stigmatella aurantiaca* Sg a15 and *M. xanthus* A47 produced cittilin amounts in the same range of the wild type strain DK897. This big difference is not easily explainable, as Sg a15 seems to meet all requirements for the production as well and the introduced plasmid (pCK-Cittilin) was exactly the same. Therefore available contigs containing genome sequence of strain A47 were analyzed for the cittilin cluster and revealed that parts of the cluster are present (Figure 25). The ABC transporter genes, the leader and the core peptide exist, and an apparently truncated P450-encoding gene is present, too. However, genes downstream of the P450 gene including the methyltransferase are absent. Hence no cittilin can be produced, but the immature peptide which is not cyclized and methylated might be released. A matching YIYY tetrapeptide could however not be detected using sensitive HR-LC-MS measurements, a result which is in agreement to the analysis of the P450 knockout mutant. For the heterologous expression experiment, the pCK-Cittilin plasmid integrates into the phage attachment site of the bacteria, which leads to a genotype in which the genes for the leader and the core peptide are then present twice in the genome. This could possibly increase the intracellular level of the tetrapeptide peptide, but more importantly it complements the A47 strain with functional P450- and methyltransferase-encoding genes. A possible reason for increased production could be the natural presence of the ABC transporter genes in strain A47 as they might play a role in processing the mature peptide. They are lacking in case of the heterologous expression in Sg a15, offering an explanation for minor cittilin production (ABC transporter genes see later on in this chapter).

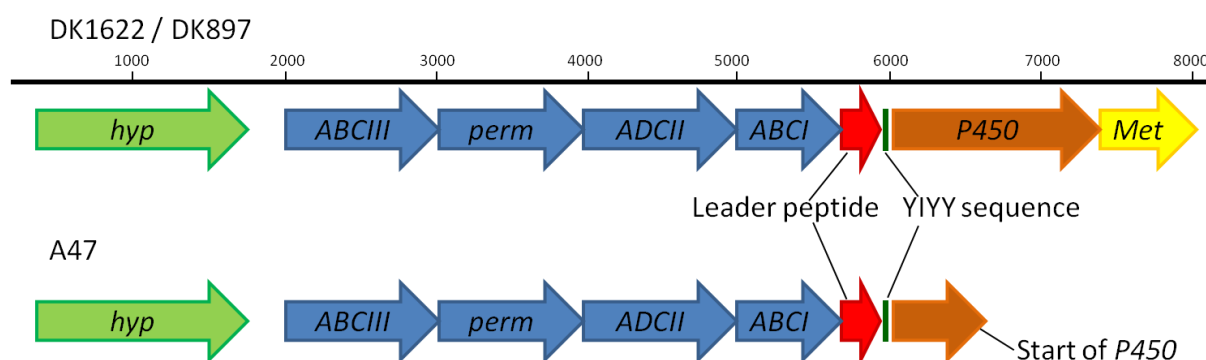


Figure 25:

Above complete cittilin gene cluster in DK897 and DK1622, below cittilin cluster in A47 with defective P450 and no methyltransferase. *hyp* = FAD-binding domain protein, monooxygenase; *ABC III* = periplasmic ferric siderophore binding protein; *perm* = permease protein; *ABC II* = permease protein; *ABC I* = ATP-binding protein; leader peptide; YIYY = core peptide; P450; Met = methyltransferase.

As major constituent parts of the cittilin gene cluster appear to be intact in strain A47, it should be possible to activate it by addition of the missing genes. When the P450 and the methyltransferase are introduced, the gene cluster should in principle contain all necessary genes to resume cittilin production. The two missing genes were amplified from the pCKCittilin expression construct and cloned into the pCK vector. Here the genes are under the control of a T7A1 promoter. The resulting pCK-P450-Met plasmid was introduced into the strain A47. Obtained clones were cultivated and analyzed via HPLC-MS measurements. No cittilin production was detected. The reasons for the lack of production are not clear. One possibility might be that the genes are not directly behind the gene cluster as in this strategy the plasmid integrates into the phage attachment site. Another explanation could be that the P450 is shorter than predicted (the protein start could be at a different position / later on in the sequence) and therefore not under the control of the T7A1 promoter. Possible start codons are 102 bp or 360 bp after the so far expected gene start. But if that is the case the gene should be still under the control of the natural promoter.

Further attempts for gene cluster reactivation in A47 should be done. Alternative start codons for the P450-encoding gene should be tested. As an option, the genomic integration via single-crossover recombination downstream of the cittilin-encoding region could be performed by using a different plasmid which does not integrate into the phage attachment site. This way, it should be possible to reactivate the biosynthetic pathway, if no additional non-obvious mutations are present which could hypothetically render the genes naturally present in A47 as non-functional.

Schmidt et al. have shown in their study of the patellamide biosynthetic pathway that it is possible to create a peptide library by site directed mutagenesis of the encoding sequence by PCR.^[143] Making use of the knowledge that the essential functions of the cittilin gene cluster are now established, new cittilin derivatives were created by amplifying the core and leader peptide whereat the sequence of the core peptide was modified by the oligonucleotides. For this method the natural *SpeI* site and the *NdeI* restriction site for cloning the cluster into the pCK vector were used. The modified regions were sequenced for correctness. To keep the main architecture the three tyrosines remained untouched, and only the isoleucine was exchanged against glycine, alanine, asparagine and valine as a proof of principle. As the production rate in the heterologous host Sg a15 is very low, the pCK-cittilin vectors were transferred into the natural host strain *M. xanthus* DK897. The recombinant cittilin clones

produced, beside the normal cittilin, also new "unnatural" cittilin derivatives (Figure 26, Figure 27).

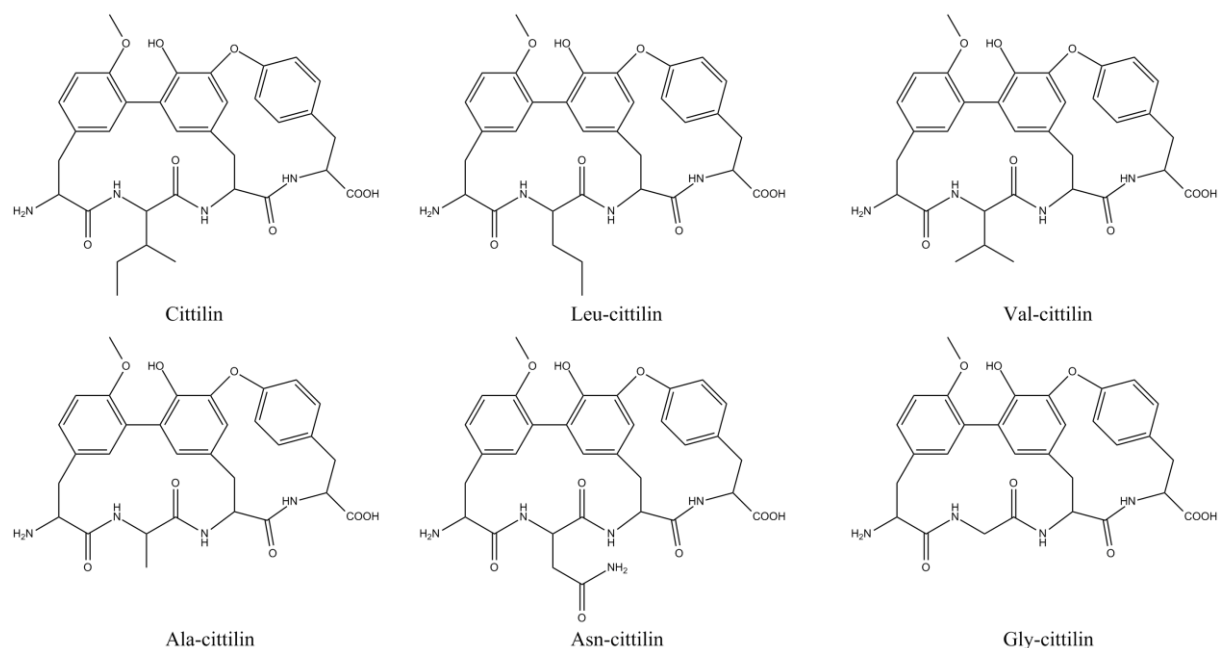


Figure 26: Cittilin A and the new cittilin derivatives where the isoleucine is exchanged against: leucine, valine, alanine, asparagine and glycine.

Production of novel cittilin derivatives was unambiguously proven by LC-HR-MS and tandem MS and it was confirmed that none of the new substances are produced by the wild type DK897 (Figure 27). For val-cittilin we virtually detect two matching peaks at distinct retention times, as it has the same mass as the natural cittilin B. All measured masses agree well with the proposed molecular formulas for the new derivatives (Figure 28). As isoleucine can be exchanged by five other structurally diverse amino acids, it is supposable that the other proteinogenic amino acids could also be incorporated at the same position. However, when one of the tyrosine would be replaced the complete core structure is likely to undergo an unforeseeable change, and thus the respective experiment was not attempted in this study. In addition, it is unclear how in the case of Tyr substitution a cyclization would be performed by the biosynthetic machinery. This issue remains to be clarified in future studies.

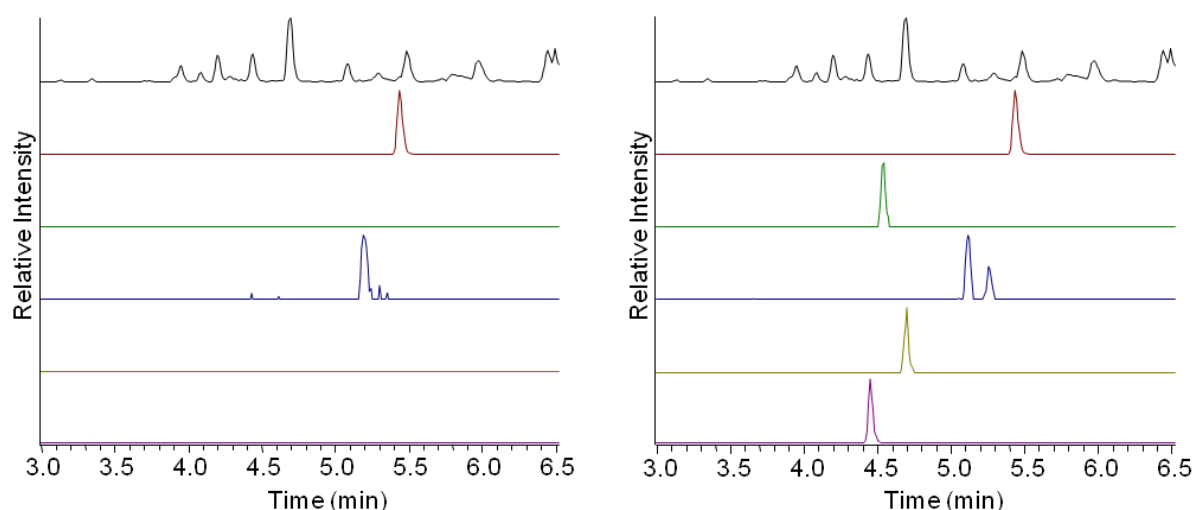


Figure 27: HR-MS chromatograms: left DK897 wild type, right DK897 mutants; top BPC, EIC on masses of cittilin (calc. 631.2762), gly-cittilin (calc. 575.2136), val-cittilin / cittilin B (calc. 617.2606), ala-cittilin (calc. 589.2293) and asn-cittilin (calc. 632.2351).

Novel cittilin derivatives were not only generated in *M. xanthus* DK897 but also in *M. xanthus* A9, A23 and *Stigmatella aurantiaca* Sg a15. The production rates were very low and routine LC-MS analysis did not show much production. Therefore a selective ion monitoring (SIM) scan from 550 - 650 m/z was performed. With this method it was possible to detect some of the expected compounds in small traces, namely the leu-cittilin, asp-cittilin and gly-cittilin derivatives. To prove the identity of these compounds MS² experiments were executed. This experiment showed for the leucine derivative the same fragmentation pattern as the one of cittilin (Figure 29).

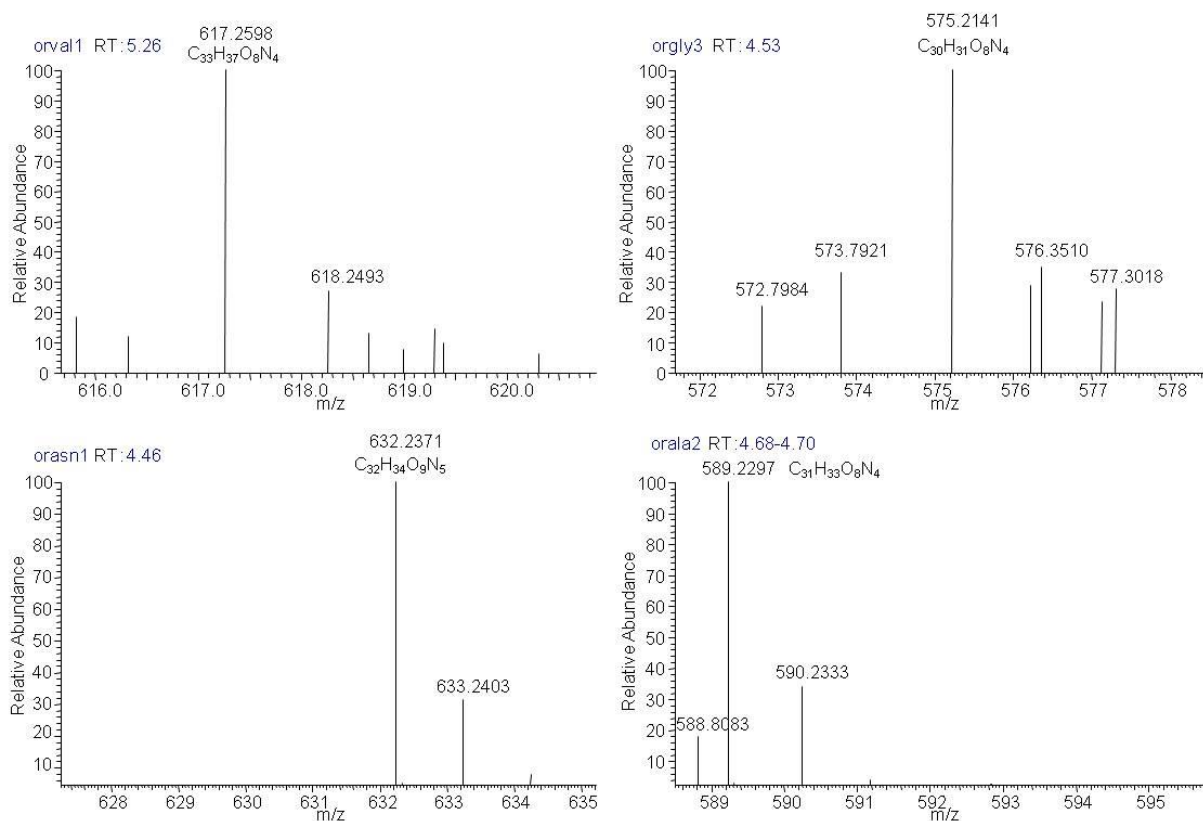


Figure 28: HR-MS spectra of the new cittilin derivatives: Val-cittilin (m/z measured 617.2598), Gly-cittilin (m/z measured 575.2141), Asn-cittilin (m/z measured 632.2371), Ala-cittilin (m/z measured 589.2297).

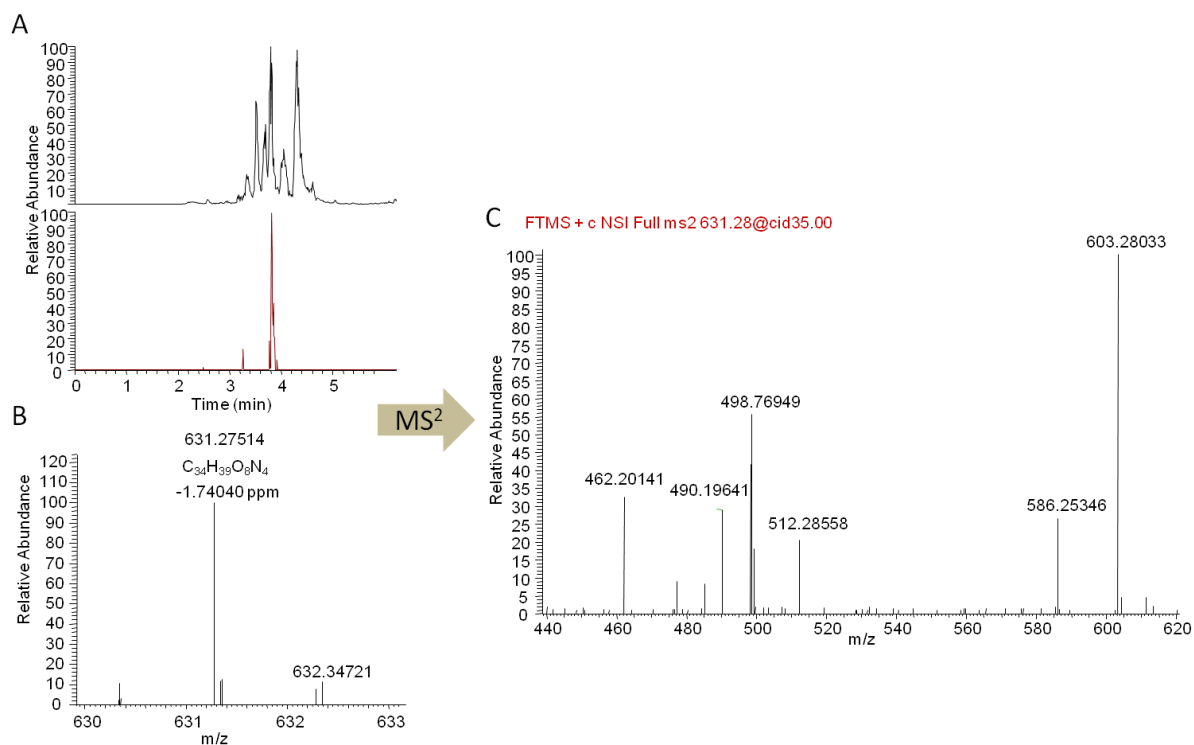


Figure 29: A: SIM scan 550 -650 m/z of Leu A23 mutant extract; B: mass spectra of leu-cittilin; C: MS^2 of leu-cittilin

These results show in general that cittilin cannot only be modified artificially in the wild type strain, it is possible in heterologous host, too.

By the heterologous expression of the cittilin gene cluster it was shown that one P450 is apparently sufficient for the bicyclization of the peptide. But in other cases like in the vancomycin two different enzyme are in charge of that. ^[181;189] It was thought, that the *hyp* gene in front of the ABC transporter might encode for another P450 enzyme, but in strain ST200611 it does not exist (Figure 30). ST200611 has a different *hyp2* gene in opposite direction at that position. The closest relative in a blast search is a dipeptidyl aminopeptidase / acylaminoacyl peptidase from *Rhodococcus erythropolis* SK121. Therefore it can be expected that *hyp* is not an essential part in the cittilin pathway.

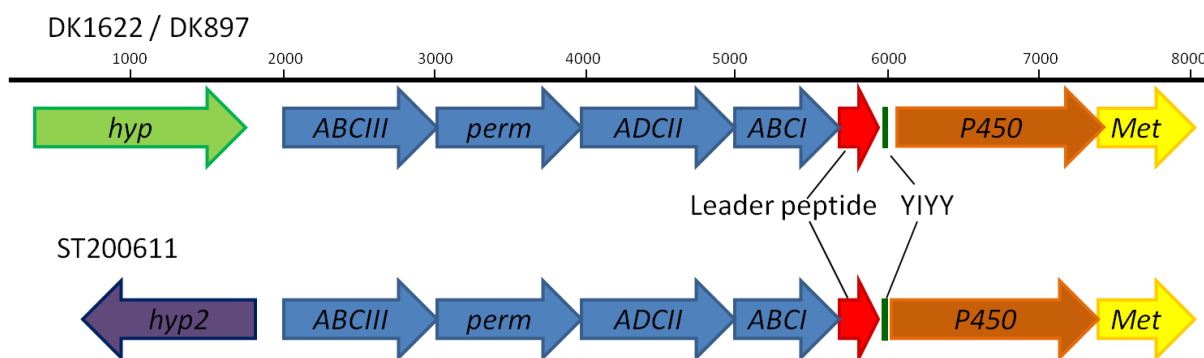


Figure 30: Top: cittilin gene cluster in DK897 and DK1622, below cittilin cluster in ST200611. *hyp* = FAD-binding domain protein, monooxygenase; *ABC III* = periplasmic ferric siderophore binding protein; *perm* = permease protein; *ABC II* = permease protein; *ABC I* = ATP-binding protein; leader peptide; YIYY = core peptide; *P450*; *Met* = methyltransferase *hyp2* = dipeptidyl aminopeptidase / acylaminoacyl peptidase.

Another possibility for a second P450 enzyme would be that a different enzyme is taking over this job, but this second enzyme would then have to be present as well in the heterologous host strain Sg a15.

To answer the question if one P450 enzyme is able to perform two reactions, a C-C and a phenolic connection, the enzyme should be expressed and an *in vitro* assay with the precursor peptide should be carried out. Hence at first an expression construct for protein expression had to be designed, so that the protein can be isolated by chromatographic methods. Therefore a tag has to be fused to the protein which can be removed later on. Furthermore it had to be decided at which side the tag should be attached. For it, the P450 was compared to the vancomycin types of P450, they contain there characteristic sequence motive at the C-terminus ^[189;201], so it was assumed that it is the same case here. Therefore the tag should be

placed at the N-terminus. Primers were designed to amplify the P450 gene and to introduce restriction sites which enable it to be introduced in the pSUMO ck4 vector (LifeSensors) (Supplementary figure 2) and the pet-28a vector (Supplementary figure 3). Both vectors contain a tag sequence which is arranged at the N-terminus of the protein. The P450 gene of DK1622 was amplified by Phusion[®] polymerase (Finnzymes) and cloned into both vectors. The proteins were extracted from the received clones and loaded on a SDS-Gel. No band with a correct size could be detected on the gel. Therefore the plasmid was transformed into the *BL21*⁺ and examined in the same manner. A band with the correct size at 54 kDa was detected in the pellet but not in the supernatant (Figure 31). The band was analyzed via MALDI and proven to be the correct P450 protein.

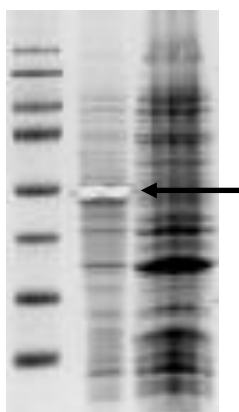


Figure 31: SDS gel: left marker; middle pellet; right supernatant; the 54 kDa band (marked with arrow and cut out for analysis) was analyzed via MALDI.

To get the P450 protein into the soluble phase it was expressed together with some chaperones pGro7 (TaKaRa), pKJE7 (TaKaRa) and pL1SL2^[202] in *BL21*⁺ and the latter in *Rosetta* as well. Clones with the TaKaRa plasmids were induced with *L*-Arabinose and IPTG, pL1SL2 only with IPTG. The cultures were treated as described above. The result is shown in Figure 32 no P450 enzyme in the soluble phase only in the pellet. To make sure that the slightly bigger bands are not the intended proteins they were analyzed by MALDI as well. But the bands were identified as chaperones.

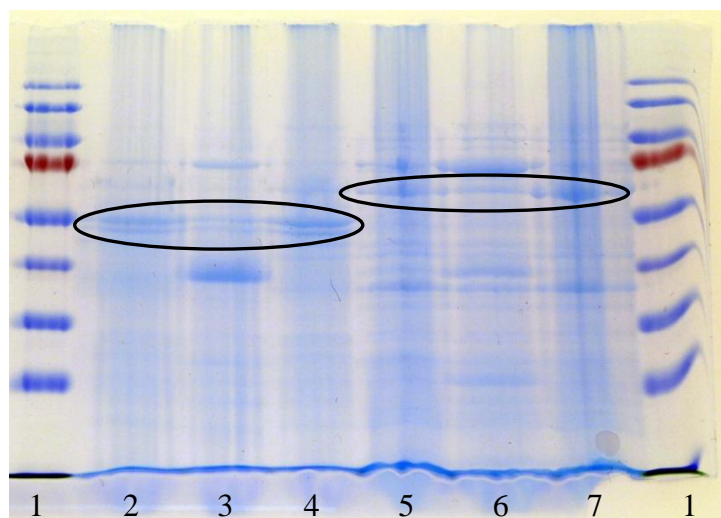


Figure 32: SDS gel: 1. Marker; pellets of 2. pGro7, 3. pKJE7, 4. pL1SL2; supernatant of: 5. pGro7, 6. pKJE7 and 7. pL1SL2; the marked region were analyzed by MALDI.

These results demonstrate that it is not easy to bring the P450 protein from the cittilin cluster into a soluble form. Here more advanced techniques are required. As a workaround, an assay with the cells of these expression mutants and the tetrapeptide YIYY was attempted. Cells were sonicated and the linear tetrapeptide was added. The mixture was extracted with acetonitrile and analyzed in the HPLC-MS to see if YIYY is converted into cittilin. But no cittilin was found, and after incubation the tetrapeptide could not be detected anymore. This finding hints at a rapid degradation processes. In this case it might be useful to express the whole leader peptide as a fusion protein as this might protect it from degradation.

As no pure enzyme could be isolated and the assay with the crude cells failed, no conclusion can be made about the enzymatic function of the cyclization procedure of cittilin. Therefore this interesting point remains open for further investigations whether the P450 enzyme is able to perform two reactions a C-C and a phenolic connection or not. If it performs both, it might be a simultaneous reaction. Alternatively a different enzyme which is present as well in the heterologous host strain is performing this job, although this possibility is considered to be rather unlikely.

Due to the low production rate in the heterologous expression strain Sg a15 compared to the natural producer *M. xanthus* DK897, the impact of ABC transporter genes was addressed. Such putative ABC transporter system exists upstream of the leader peptide region (refer to Figure 22). ABC I is an ATP-binding protein, ABC II and perm are permeases, the ABC III is a periplasmic ferric siderophore-binding protein. Upstream of that a hypothetical protein is encoded, exhibiting similarity to other genes with unknown function from *Rhodococcus*

opacus B4 and *Anabaena variabilis* ATCC 29413. Furthermore it shows partial similarity to a monooxygenase FAD-binding protein from *Nitrosococcus halophilus* Nc4, which belongs to the family of 2-polyprenyl-6-methoxyphenol hydroxylases and is related to FAD-dependent oxidoreductases.

These genes in front of the cittilin gene cluster might have an important role in processing the peptide. In view of literature reports ^[157;160;203] from other RP biosynthetic systems, such a role would be conceivable. The ABC (ATP-binding cassette)-type transporters are similar to the microcin biosynthetic gene cluster, where they are involved in secretion of the product into the extracellular medium. ^[151] Additional evidence for the involvement of the ABC transporter in cittilin production comes from the fact that these genes are present in every cittilin producing strain (where the cittilin sequence is known). Furthermore the Dittmann group could show that if they express the genes of the RP microviridin without an ABC transporter, the production level is very low. Moreover they were able to detect unprocessed compounds which contained amino acids of the leader peptide as well. However, during heterologous production of cittilin no unprocessed peptides were detected in the present study, here the transporter genes were absent. To find out if the transporter genes have an effect on the biosynthesis, in first instance for processing the mature peptide, an attempt was made to clone these genes upstream of the gene cluster and thus co-express them together with the cittilin cluster. As heterologous hosts *M. xanthus* A9 (a none cittilin producing *M. xanthus* strain) and *Stigmatella aurantiaca* Sg a15 were used.

By PCR it was tried to amplify the genes in one piece (~ 5 kbp) by different polymerases (Taq, Phusion, HotStartTaq, KOD, GoTaq), different sets of primers and different qualities of genomic DNA of diverse strains. Sometimes some DNA bands were spotted but never in the correct size. A reason for this might be several repeats in this genome region. Hence the four ABC transporter genes in front of the cittilin encoding area were amplified separately to get small fragments at around 1 kb. To stitch the sections together on both sides of each gene, restriction sites had to be introduced. As only one clear ribosomal binding site (RBS) could be found (between *ABCIII* and *Perm*), artificial RBSs were introduced by the plasmid. The advantage of this strategy is that it can be shown if all four genes are necessary, typically ABC transporters are encoded in four genes. The single genes were cloned into the pCK-cittilin plasmid. Restriction enzymes which do not cut in the pCK-cittilin vector, nor in the ABC genes and in the cittilin cluster had to be found. 12 enzymes were possible: *AsiSI*, *AvrII*, *BamHI*, *BspTI*, *Eco105I*, *KspAI*, *PacI*, *PmeI*, *PsiI*, *SnaBI*, *SspI*, *SwaI*. As it is not clear what the regions between the genes are good for (and some genes are overlapping) any

modifications of intergenic region was to be avoided. Therefore mutations had to be integrated into the second or third amino acid of the genes. At the 5' end an *NdeI* restriction site was always introduced followed by another restriction site (Figure 33). On the other side the same restriction enzyme which was introduced in the gene before was added by a primer. When the cluster was stitched together the *NdeI* site was always deleted inside the encoding sequence.

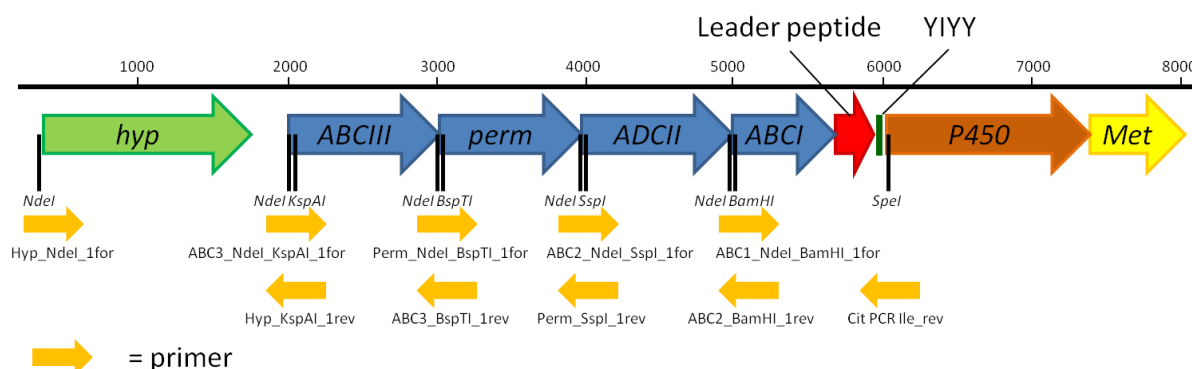


Figure 33: Citalin gene cluster with ABC transporter genes, restriction sites indicated by black bar, yellow arrows: primers which were used to introduce restriction sites.

By introducing the restriction sites at the beginning of the genes following changes will occur in the amino acid sequence (marked in bold; original sequence => mutated sequence): *ABCI*: MIEAK => MIGSK, *ABCI*: VPAN => VNIN, *Perm*: VKRAF => VKLKF, *ABCI*: VSPSN => VVNSN and *Hyp*: no changes.

To find out which ABC gene set is really needed for the best expression the amount of citalin was quantified by LC-MS measurements. For that purpose the mutants were cultivated in duplicates (Sg a15) or rather triplicates (A9). The cultures were extracted and quantified by HPLC-HR-MS measurements. An internal standard (Coumarin 314) was used for quantitation. The peak area of citalin-derived MS signals were integrated and quantified by the Bruker QuantAnalysis software. Furthermore the data were corrected by cell density of the cultures.

All mutants of *M. xanthus* A9 produced more or less the same amount of citalin (Figure 34). *ABCI* clones a little bit more than the others but this is within the common fault tolerance. Here it looks like there is no influence of the ABC transporter on the citalin production. In the case of the Sg a15 mutants the picture is a different one (Figure 35), here the production remains at the same level if only the first three genes are present. But the mutant with all four genes has an increase of production of around three fold. This can be considered as reasonable

because only here all four proteins are expressed which normally constitute an ABC transporter. This result shows clearly that the ABC transporter does have an influence for cittilin production. On the other hand it is no dramatic increase of production. A possible explanation for that might be that the ABC transporter is not involved in processing the peptide, only in transporting it out of the cell. In the absence of the transporter genes, this function is possibly substituted by other transport mechanisms. As a careful speculation, one could hypothesize that the impact on productivity should be greater if there would be an involvement of the transporter in processing the mature peptide.

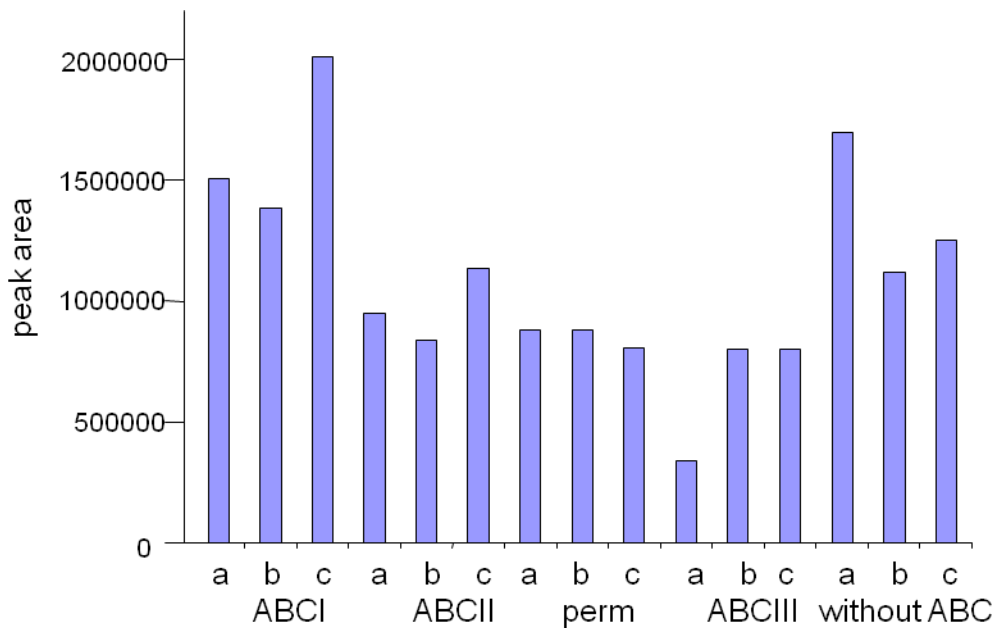


Figure 34: Quantification of A9 Cittilin transporter mutants with different cluster length, a, b and c are different clones.

A reason for no effect in the case of *M. xanthus* A9 might be, that parts of the gene cluster may be present in the genome although A9 is a none producer of cittilin. It could be similar to the case of A47 where a defective cittilin gene cluster was detected, however all ABC transporter genes being present. Therefore an addition of the already existing genes would not produce a noticeable effect. Here it would be mandatory to check if A9 has such an imperfect gene cluster, but whole genome sequence information is not currently available for this strain and this problematic was not known at that time point.

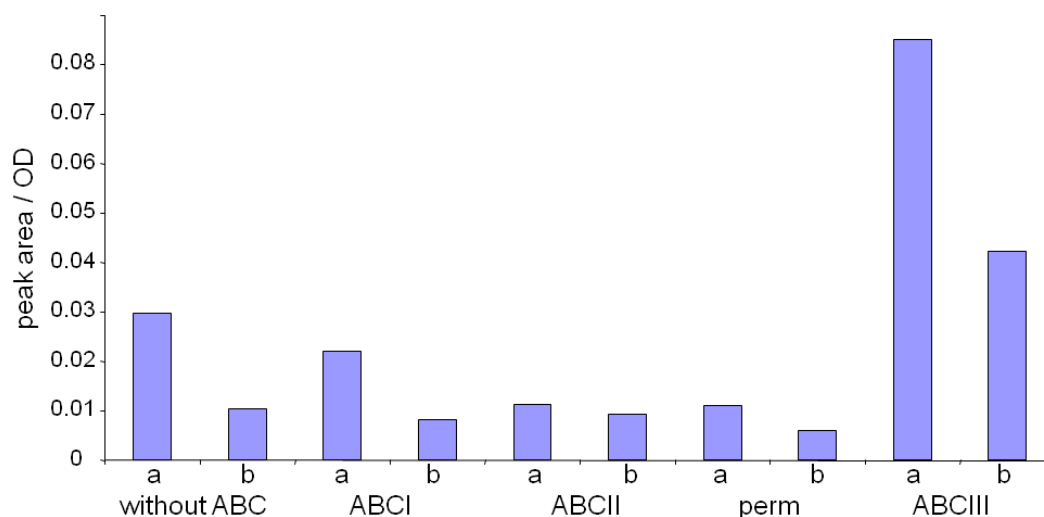


Figure 35: Quantification of Sg a15 cittilin mutants with different cluster length, a and b are two different clones.

To sum up cittilin was isolated from the crude organic extract of *M. xanthus* DK897. It was structurally characterized and afterwards cittilin B was confirmed to be identical with RP-66453 from a *Streptomyces* sp.^[172]. Cittilin could also be detected in DK1622; a strain for which complete genome information was available.^[79] Here the in silico search lead to the identification of the biosynthetic pathway which is based on the ribosome and not as initially expected, from NRPS origin. A heterologous expression of the cittilin gene cluster in *Stigmatella aurantiaca* Sg a15 provided evidence for the ribosomal metabolic pathway. It could be shown that the leader peptide, the core peptide, a P450 and a methyltransferase are absolute required for the biosynthesis of the molecule. Furthermore, the simple modifiability of ribosomal synthesis makes it possible to successfully exchange isoleucine against other amino acids. Several nonnatural cittilin derivatives were created by genetic modification of the heterologous expression construct in the original host strain and in Sg a15.

To conclude cittilin represents the first secondary metabolite isolated from myxobacterial extracts for which the ribosomal synthesis has actually been proven. The discovery of the biosynthetic machinery underlying its formation supports the notion that there are further pathways for natural product production in myxobacteria which remain to be explored. Even though myxobacteria are well known to be ‘superproducers’ of polyketides, non-ribosomal polypeptides, hybrids thereof, alkaloids, terpenoids and phenyl-propanoids.^[45;49]

Methods

Cultivation and extraction: the *M. xanthus* strains were cultivated in CTT media (10 g Casitone (Difco), 10 ml 0.8 M MgSO₄, 10 ml 1 M Tris-HCl pH 8.0, 1 ml K₂HPO₄ pH 7.6 and 1 liter distilled water, pH adjusted to 7.6, for mutants kanamycin 50 µg / ml were added). *Stigmatella aurantiaca* grew in Tryptone-Starch (TS) medium (1 % tryptone, 1.19 % HEPES, 0.4 % soluble starch, 0.2 % MgSO₄ x 7 H₂O; pH adjusted to 7.2, for mutants kanamycin 50 µg / ml). Bacteria strains and mutant strains were grown 50-ml medium supplemented with 1 % XAD-16 resin (Sigma-Aldrich), for 72 h at 180 rpm and 30 °C in 300 ml Erlenmeyer flasks. The cultures were centrifugation (10 min, 8000 rpm), supernatant was discarded and the remaining pellet was extracted with methanol (2 x 30 ml). The solvent was removed under vacuum and resuspended in 1 ml methanol.

For cittilin isolation 10 l of DK897 was cultivated in large scale in 3 l Erlenmeyer flasks similar as described for small amounts. The obtained extract was run on a methanolic Sephadex[®] LH20 column. Cittilin containing was placed on a silica gel column which was run with a chloroform methanol gradient (start 9:1, end pure methanol). For further purification another Sephadex[®] LH20 column was performed with H₂O / MeOH (3:2).

Mass spectrometry: For LC-HR-MS an Accela UPLC system coupled to a LTQ Orbitrap (Thermo Scientific) with nano electrospray ionization (ESI) (TriVersa NanoMate (Advion)) was used. To separate the extract a gradient was run on a RP-18 column (50 mm x 2.1 mm, particle size 1.7 µm; Waters Acquity BEH) with water (+ 0.1 % formic acid) and acetonitrile (+ 0.1 % formic acid) as solvent, flow rate of 400 µl / min, gradient: 0 min 1 % B; 1 min 1 % B; 10 min 99 % B; 12 min 99 % B; 12.5 min 1 % B; 14 min 1 % B. For fragmentation experiments 30 % collision energy was used.

For LC-MS same solvents were used under following conditions: RP-18 column (Luna, 2.5 µm), t = 32 °C, flow rate 400 µl / min; 2 min 5 % B, 20 min 95 %, 3 min 95 %.

NMR spectroscopy: NMR measurements were performed on a 500 MHz Bruker Avance magnet. Using Topspin2 and deuterated solvents (methanol-d₄).

The tetrapeptide YIYY was synthesized by the company Gen Script.

Search for the cittilin cluster: For analyzing the A domains in DK1622 the following web tools have been used: <http://nrps.igs.umaryland.edu/nrps/index.html> and <http://www->

ab.informatik.uni-tuebingen.de/toolbox/index.php?view=domainpred. Cittilin contains four amino acids: 3 x tyrosine and 1 x isoleucine hence four possibilities were considered for a ribosomal search although only tyr-tyr-ile-tyr has the correct amino acid order. The codons for Tyrosine are: TAT and TAC, for Isoleucine: ATT, ATC and ATA. Using the degenerated code: Y = T or C and H = T, C or A. TAYTAYTAYATH for tyr-tyr-tyr-ile, TAYTAYATHTAY for tyr-tyr-ile-tyr, TAYATHTAYTAY for tyr-ile-tyr-tyr, ATHTA YTAYTAY for ile-tyr-tyr-tyr. It was search in the genome on both strands.

Gene knockouts: The PCR fragments were amplified from genomic DNA of *M. xanthus* DK1622 with Taqpolymerase. For the methyltransferase knockout an overlap extension PCR was made at 58 °C with following oligonucleotides: 1. PCR CitMet1 = TCAGTCCTCTGTCGATGACGAAGG, CitMet3 = CACTACCATGCGTAGTACTTCCGC GACGGA, product length: 192 bp. 2. PCR at 58.0°C, CitMet2 = GTCGCGGAAGTACT ACGCATGGTAGTGCC, Cit Met4 = TGGGCGCAACCCTGGGACATCGGAC, product length: 390 bp. 3. PCR, 58.0°C with the two first PCR products as template and CitMet1 and Cit Met4 as oligonucleotides, product length 555bp. For P450 knockout: PCR at 72°C with CitP450-1F = CGTGTAGCGCAGGCGAGGAAGGTC, CitP450-1R = GTGGTGGCCCTC ACCGAAACGATG. The PCR product was cloned into the pCR2.1topo vector (Invitrogen) and electro competent *E. coli* (DH10b) cells. The colonies were selected on Luria–Bertani (LB)-agar containing kanamycin sulfate (Kan, 50 mg/mL), single colonies were cultivated in LB medium (1.5 mL, with Kan 50 mg/mL). From these cultures plasmid was isolated using the Nucleospin plasmid kit (Macherey & Nagel). The obtained plasmids were used to create *M. xanthus* mutants as described previously. ^[195-198]

Heterologous expression: to express the cittilin gene cluster, it was transferred into the pCK vector, restriction sites were introduced by oligonucleotides (The pCK plasmid provided Dr. C. Kegler). *NdeI* in front of the precursor peptide, *HindIII* before the twelve bp encoding the cittilin, the core peptide has a natural *SpeI* site, and behind the P450 an *EcoRI* site was introduced. Citt-EcoRI_rev2 = GCCTG AGGAATTCAGTCCTCTGTCGATGA, Citt_HindIII_rev = CGGATTTACTAGTCTAGTAATAGATATACGGGGCGGAAAGCTT GTCCGCG, Citt_NdeI_for3 = CATATGATGGCG CGGCGCGCCAGCCTGGGATGCAA. The two fragments short and long were amplified from genomic DNA of DK1622 (Figure 36).

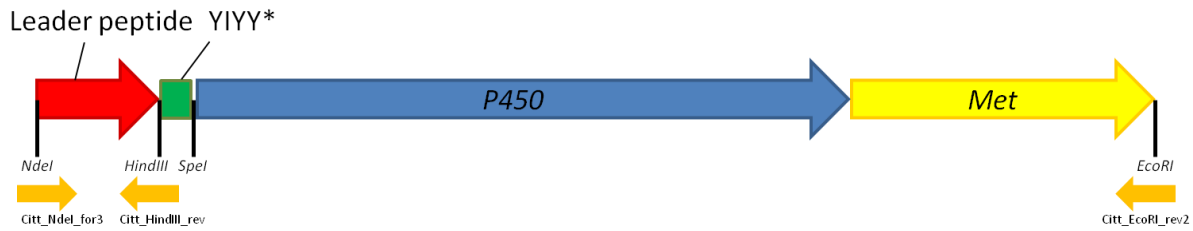


Figure 36: Citterlin cluster with introduced restriction sites and oligonucleotides.

The short fragment was cloned into the pSKII+ vector (pBluescript II SK(+)) vector (addgene) (Supplementary figure 1). The long fragment was placed into the pTOPO vector (pCR2.1topo vector (Invitrogen)). To open pSKII+ Citterlin short *SpeI* was used, the long fragment was cleaved out of pTOPO Citterlin long with *SpeI* and *XbaI* and stitched into the open pSKII+ Citterlin vector. As *XbaI* is compatible with *SpeI* two orientations are possible. The obtained clones were screened for the correct direction via PCR (*Taq*, 58°C) with the primers CitP450-1F = CGTGTAGCGCAGGCGAGGAAGGTC and M13rev = GAGCGGATAACAA TTTCACACAGG.

The citterlin cluster was transferred from pSKII+ Citterlin via *NdeI* and *EcoRI* into the pCK Citterlin vector. The pCKCitterlin vector was cloned into electro competent *E. coli* (DH10b) cells and preceded like above (gene knockout). The resulting plasmid was electroporated into *Stigmatella aurantiaca* Sg a15 as described for *M. xanthus*.^[195-198]

New citterlin derivatives: the amino acid sequence was exchanged by vary the oligonucleotides, considering the codon usage of Myxobacteria. The sequence of Ile was exchanged against one other amino acid sequence (Ala, Asn, Gly, Leu, Val). Cit PCR Asn rev = TTTGACTAGTCTAGTAATA**GTT**ATACGGGGCGGAAAG, Cit PCR Val rev = TTTGACTAGTCTAGTAATA**CAC**ATACGGGGCGGAAAG, Cit PCR Leu rev = TTTGACTAGTCTAGTAATA**CAG**ATACGGGGCGGAAAG, Cit PCR Gly rev = TTTGACTAGTCTAGTAATA**GCC**ATACGGGGCGGAAAG, Cit PCR Ala rev = TTTGACTAGTCTAGTAATA**GGC**ATACGGGGCGGAAAG. Another primer was designed over the *NdeI* side of the pSKII+ Citterlin vector: Cit PCR Nde for = CTCGAGCCACTACCATATGATGG. PCR was done with both primers at 58 °C, Phusion polymerase, pSK – Citterlin vector as template. The resulting fragment of 260 bp was digested with *SpeI* and *NdeI*, as well as the pCK Citterlin vector. Both are stitched together and the new plasmids were electroporated into *E.coli* DH10b. The plasmids of the clones were isolated like described above, sequenced for correctness and transferred into the wild type stain DK897 (displayed before).

Investigation of the leader peptide: for resequencing of the leader peptide of ST200611 primers were designed to amplify this region Citvorl 2F = CAGATTGAGGGCGACTGGATTGC, Citvorl 1R = GTCTTCCGACCCAGCCGTCAG. The PCR was done with Phusion[®] polymerase (Finnzymes) at 58 °C, the amplicate was cloned into the pJet vector (Fermentas) and send for sequencing (seq-it).

Re-activation of A47 gene cluster: Primers were designed for the amplification of the missing cittilin genes. Cit_Met_EcoRI_rev2 = TCTGCAGAATTCAGTCCTCTGTCGATGACGA, Cit_P450 Nde_1for = TGGAGCATATGGTTCGCTCCACCTGCGCCA. A PCR was performed with Phusion[®] polymerase (Finnzymes), the amplicate digested with *EcoRI* and *NdeI* as well as the designated pCK vector. Both were ligated together to result in a new vector pCK P450 Met (Figure 37).

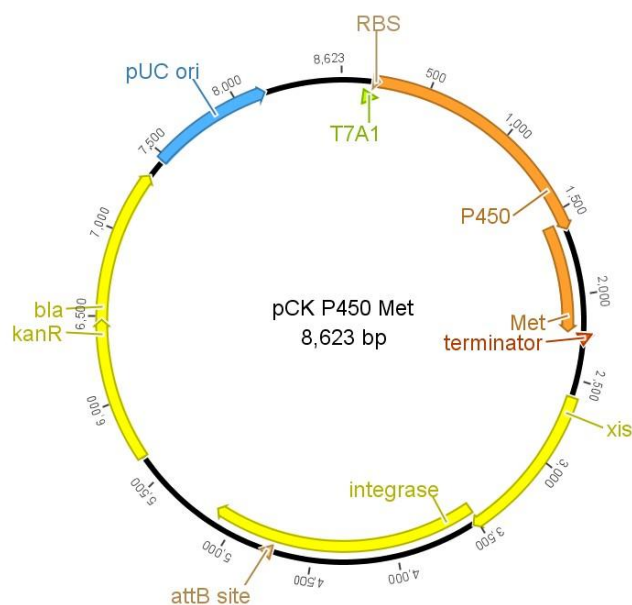


Figure 37: Vector for the complementation of strain A47 with P450- and MT-encoding genes.

This pCK P450 Met plasmid was electroporated into the strain A47; clones were picked and genetically verified. Two correct mutants were then cultivated with XAD, harvested after three days and their extracts were analyzed with the HPLC-MS system.

P450 expression: Primers were designed to amplify the P450 gene and to introduce restriction sites which enable it to be introduced in the pSUMO ck4 vector (LifeSensors) (Supplementary

figure 2) and the pet-28a vector (Supplementary figure 3). The following primers for amplification of the P450 enzyme were used:

Cit ex P450 *EcoRI* = TGCAGGAATTCTCATGCTCCCGCCGCATGCCGCA

Cit ex P450 *NheI* = GTTCGCTCCACCTGCGCCAGGCTAGCGGCAT

Cit ex p450 rev2 = CCGCTGAATTCTCATGCTCCCGCCGCATGCCGCACCTC

Cit ex p450 for2 = AGCCCGCTAGCTTGGGTCTCAAGAGCTGGTCGACGTGGGT

Genomic DNA of DK1622 was used for amplifying the P450 gene by Phusion[®] polymerase (Finnzymes). The amplificate was digested with *EcoRI* and *NheI*. The pSUMO ck4 vector was opened with *SpeI* and *MunI*, pET-28a vector with *EcoRI* and *NheI*. *SpeI* is ligatable with *NheI* and *EcoRI* with *MunI*. The PCR product was cloned into both vectors; clones were verified by digestion and send for sequencing. The construct was then transformed into chemical competent *BL21* cells.

A correct *BL21* clone was cultivated in a 1 l flask (250 ml LB media) to an OD of 0.8, the culture was placed at 16 °C, after 20 min induced by IPTG (50 µl, 1 M, final conc. 0.2 mM) and cultivated over night. 1 ml of culture was centrifuged, the pellet was resuspended in 100 µl BugBuster Master Mix (Novagen) shook at 500 rpm for 30 min, centrifuged at 4 °C for 45 min at 13000 rpm. Pellet and supernatant were placed on a 12 % SDS-Gel. Same procedure was accomplished for the *BL21*⁺ clones.

ABC transporter: The fragments were amplified by Phusion[®] polymerase (Finnzymes), cloned into the pJet vector (CloneJET[™], Fermentas) and sequenced for their correctness. The plasmids were digested with *NdeI* and then with the corresponding other enzyme. The *ABCI* fragment was ligated into the pCK-Cit vector in front of the cittilin cluster. The resulting plasmid was reopened by *NdeI* and the next restriction enzyme. And the next gene was added and so on to finally end up with five different constructs. These pCK Cit plasmids containing the region up to *ABCI*, *ABCII*, *perm*, *ABCIII* and *hyp* were introduced into A9 and Sg a15. The clones were analyzed for their genetically correctness by control PCR. All attempts failed to obtain mutants containing the *hyp* genes. The mutants were cultivated and the cittilin production was checked by HPLC-MS.

The following primers where designed:

Cit PCR Ile rev = TTTGACTAGTCTAGTAATAGATATACGGGGCGGAAAG

ABC1_*NdeI*_BamHI_1for = ACCATATGATCGGATCCAAGAACGTCAGCAAGCGCT

ABC2_BamHI_1rev = TTCTTGGATCCGATCATCTCGTGGCCTTTC

ABC2_*NdeI*_SspI_1for = CGCATATGAATATTAACGCCCCGCCCCGTCA

Perm_SspI_1rev = TTAATATTCACGGGCGCCTCGCTTCA

Perm_NdeI_BspTI_1for = CCCATATGAAGCTTAAGTTCGCCGCCGCCGCCGT

ABC3_BspTI_1rev = AACTTAAGCTTCACGGTGGGGAGGGACT

ABC3_NdeI_KspAI_1for = CTCATATGAGTGTTAACAACAGGCGTTCCCTTC

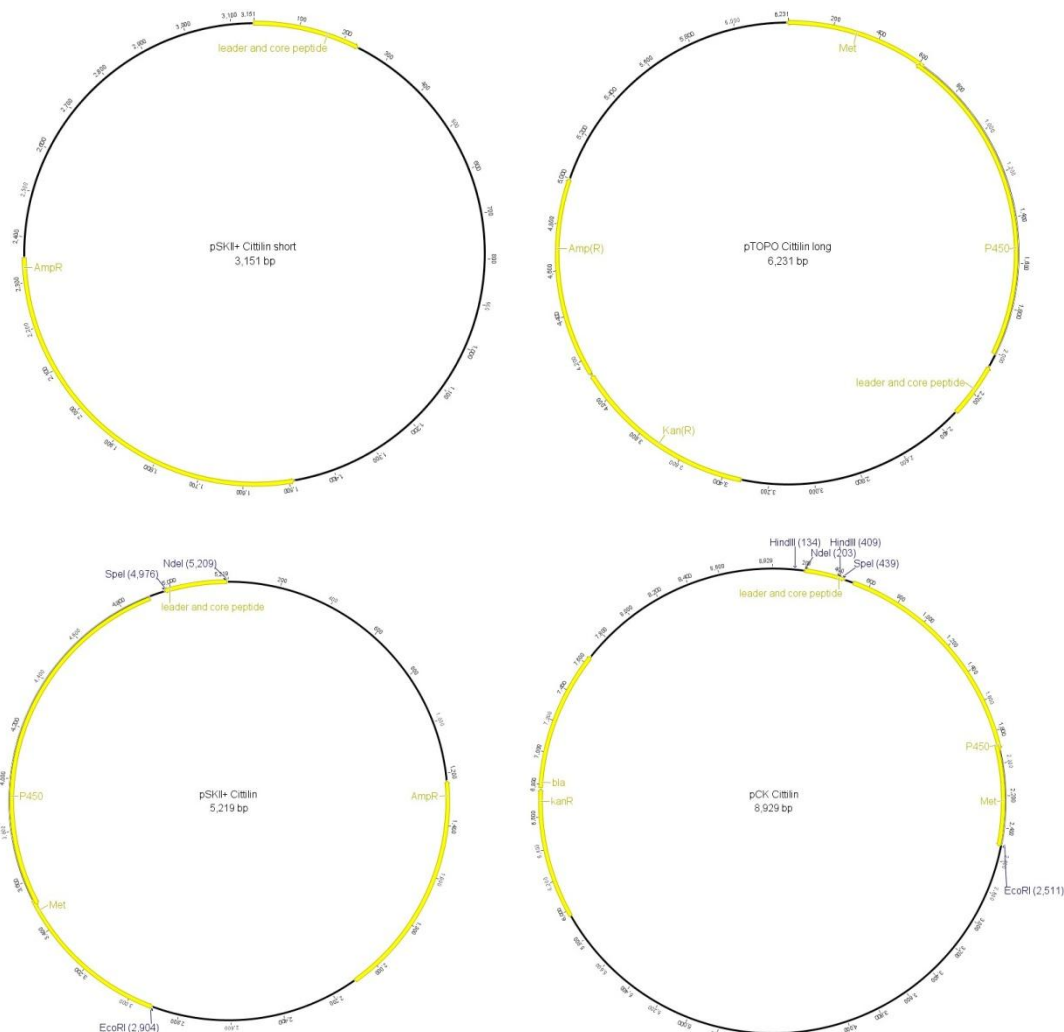
Hyp_KspAI_1rev = GAGTTAACCACGGTAGTCGTTGGAAAGGGTG

Hyp_NdeI_1for = TACATATGATGGCGGGGCTGCTGAG.

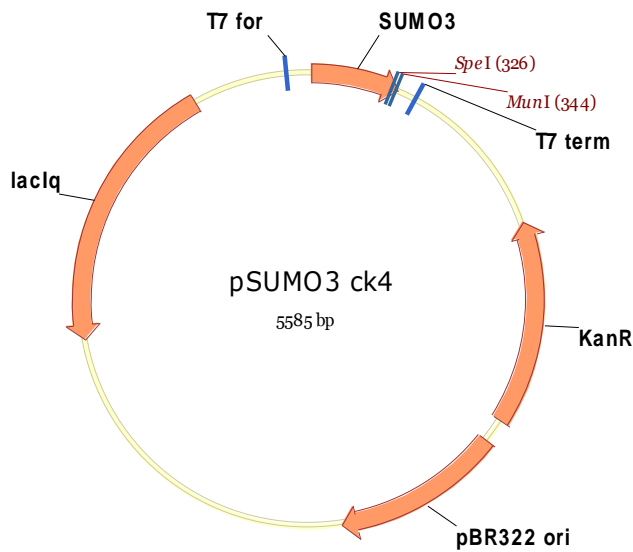
PCR conditions: Phusion, HF-Buffer, 68 °C 20 s, 1.5 min 72 °C. 30 cycles, genomic DNA of DK897. The obtained amplicates are ligated into the pJet vector (CloneJET™ PCR Cloning Kit, Fermentas), sequenced (seq-it) and placed in front of the cittilin cluster, in the pCK-Cit plasmid. Creation of *E.coli* and myxobacterial mutants as described before.

For the quantification 15 ml TS or (A9: CTT) media plus Kan50 and 0.25 ml XAD were inoculated with liquid preculture. After 7 days (A9: 4 days) the OD₆₀₀ was measured, cultures were harvested and the XAD was extracted twice with MeOH, always under same conditions for each sample. The received extracts were measured together with Coumarin 314 (standard program, see before) at the LTQ-Orbitrap for HR-MS quantification. The peak areas of cittilin were integrated and quantified by the Bruker QuantAnalysis software. The quantification is related to OD of the cultures.

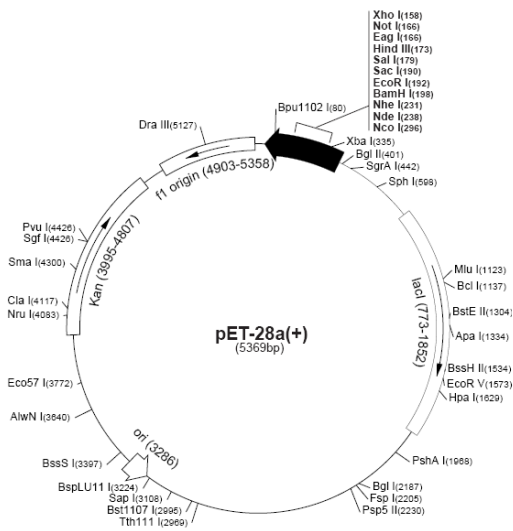
Supplementary



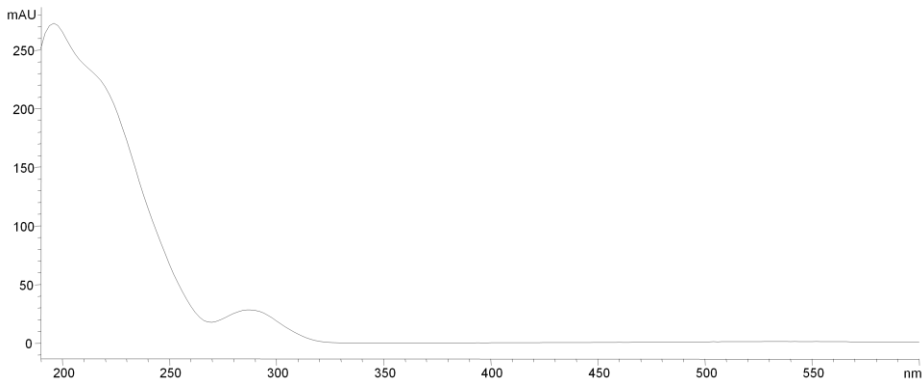
Supplementary figure 1: The four plasmids: pSKII+ Cittilin short containing: leader, core peptide and ampicillin resistance; pTOPO Cittilin long harbors: cittilin cluster, ampicillin and kanamycin resistance; pSKII+ encloses: cittilin cluster and ampicillin resistance; pCK Citlin with: complete cittilin cluster, ampicillin and kanamycin resistance.



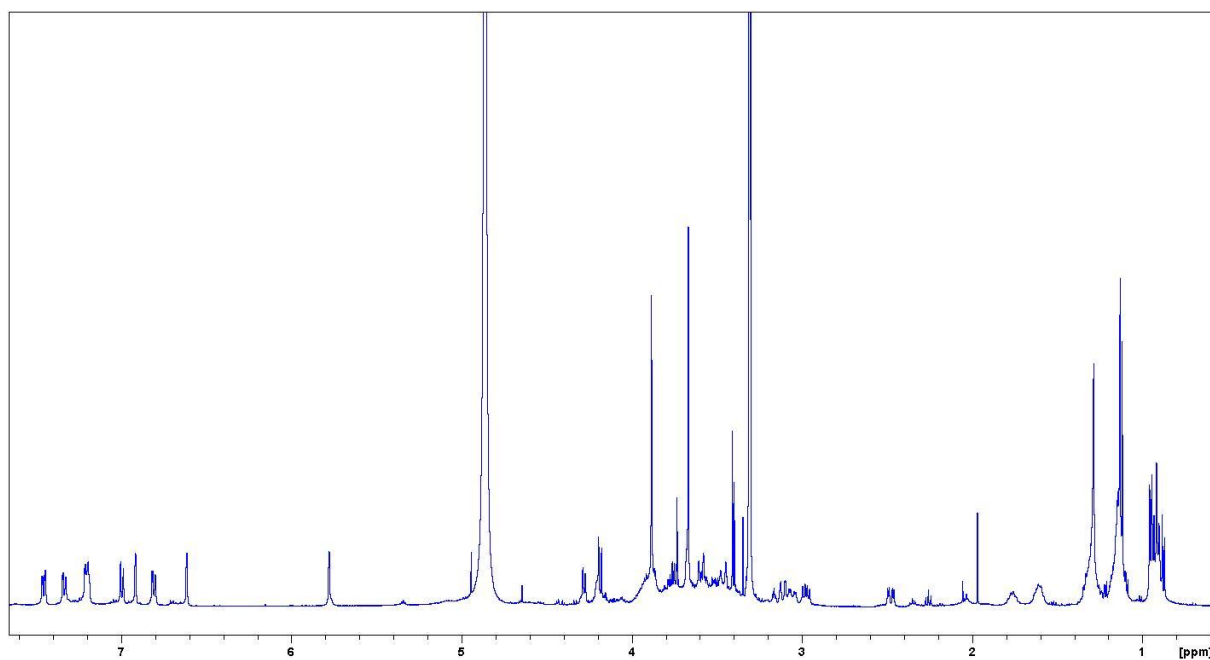
Supplementary figure 2: vector pSUMO ck4



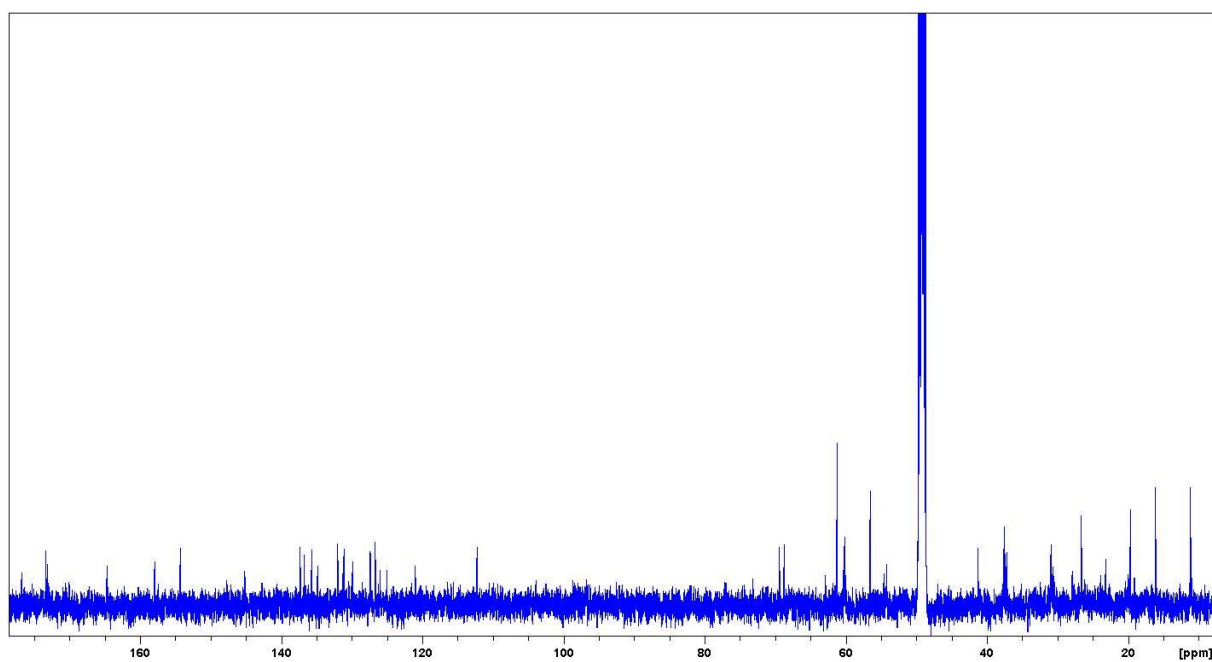
Supplementary figure 3: vector pet-28a(+) (Novagen, picture Novagen)



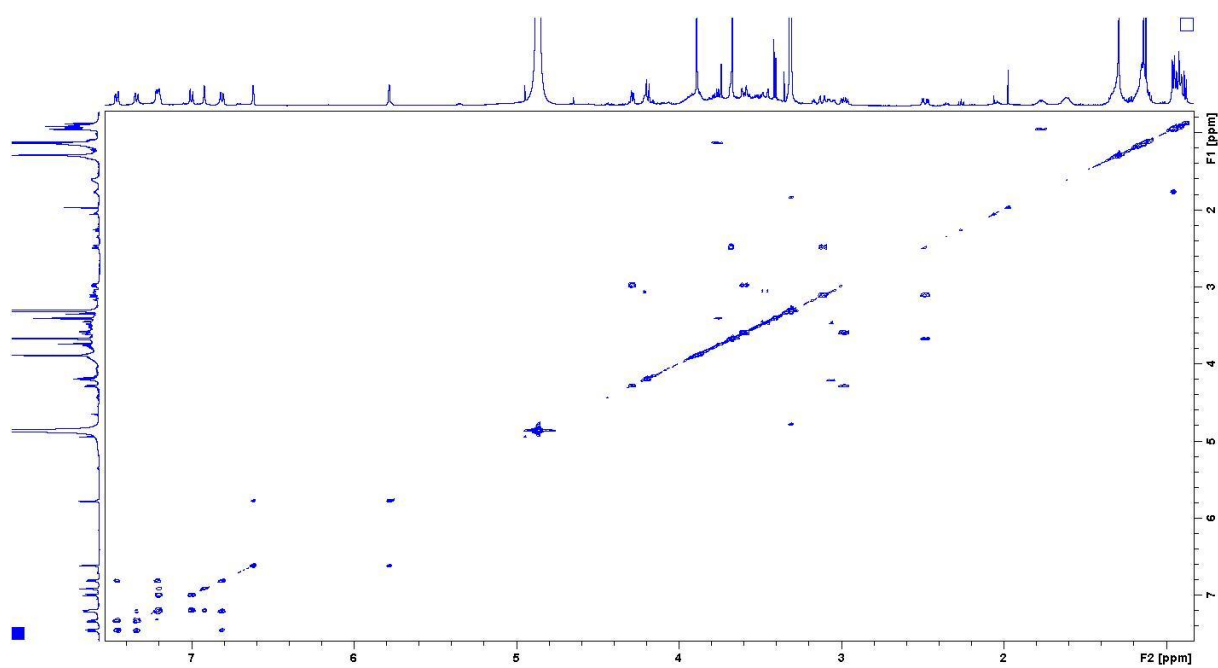
Supplementary figure 4: UV spectrum of cittilin, UV_{max} = 290 nm.



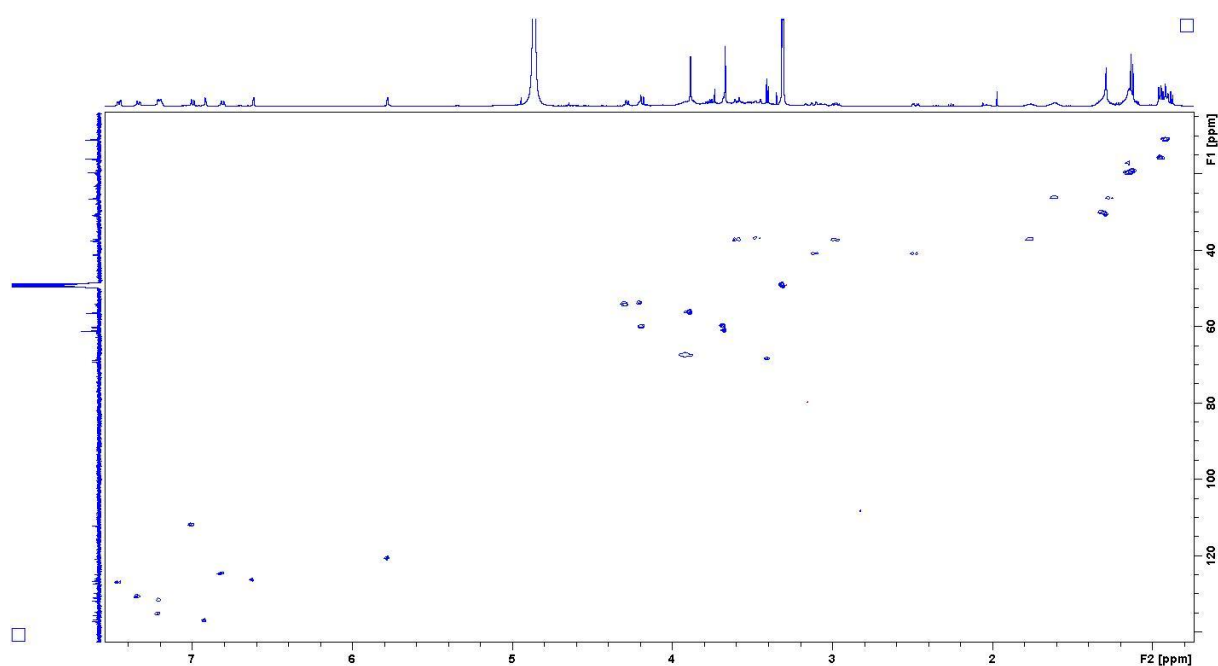
Supplementary figure 5: ^1H spectrum of cittilin A in CD_3OD at 500 MHz.



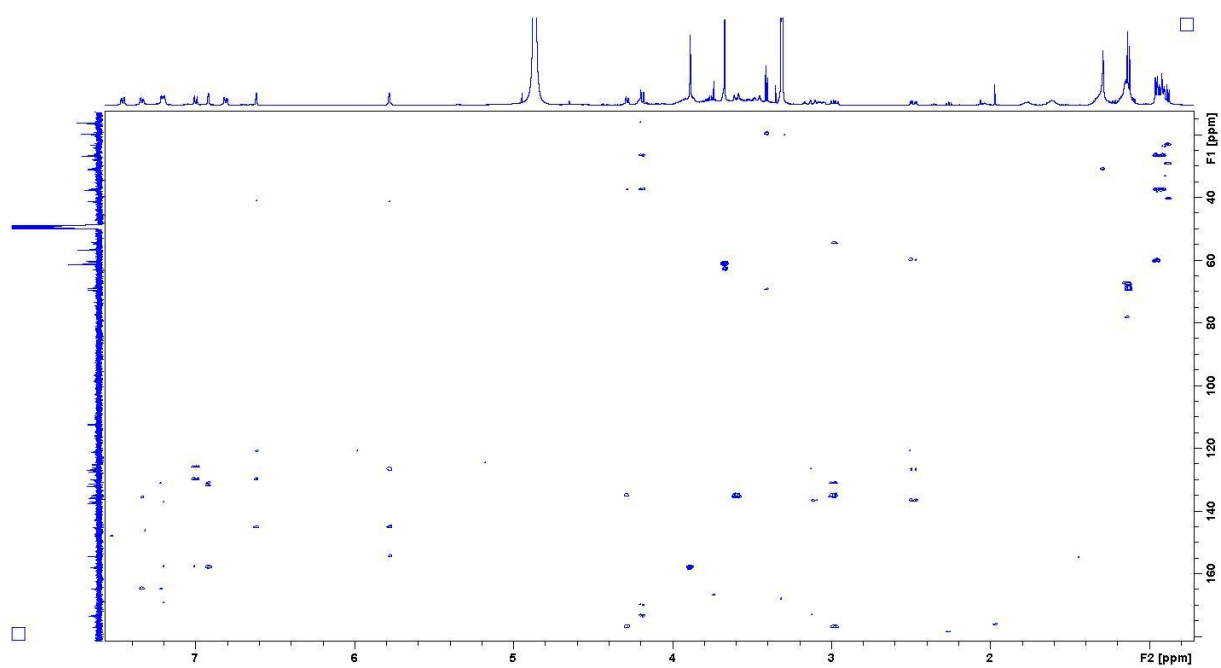
Supplementary figure 6: ^{13}C spectrum of cittilin A in CD_3OD at 125 MHz.



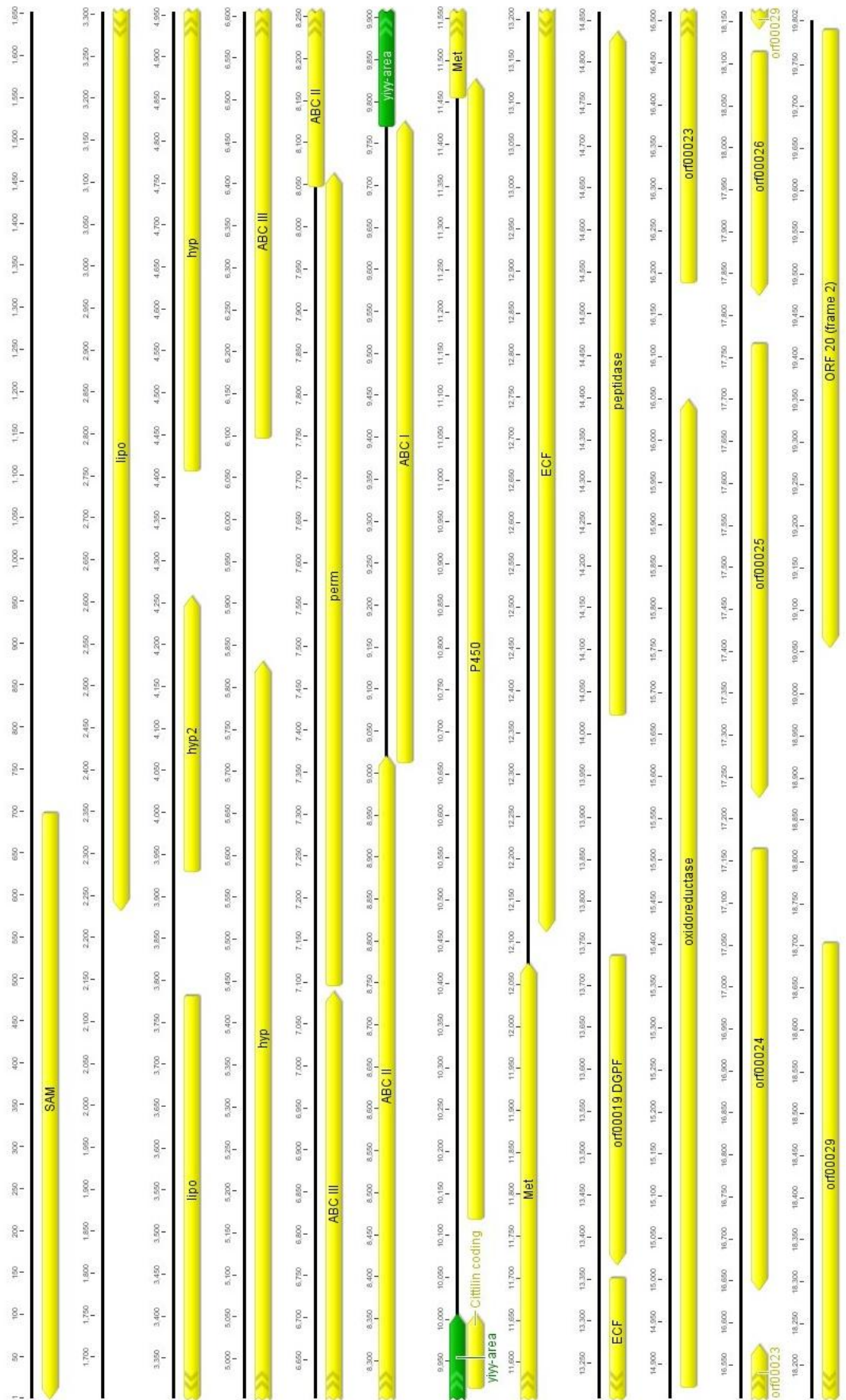
Supplementary figure 7: COSY spectrum of cittilin A in CD₃OD at 500 MHz.



Supplementary figure 8: HSQC spectrum of cittilin A in CD₃OD at 500 MHz.



Supplementary figure 9: HMBC spectrum of citalin A in CD₃OD at 500 MHz.



Supplementary figure 10: Genetic region around cittilin sequence

Cittilin cluster sequence in *M. xanthus* DK1622

```

1          10          20          30          40          50          60          70
|          |          |          |          |          |          |
ATGGCGGGGCTGCTGAGCGCCCGCGCGCTCGCGGACCACTTCGAGAAAGTCGTCATCCTGGAGAGGGACTCGCTC
CCCGGCATGCACGCGGCGCGCAAAGGTGCTCCCCAGGGCACACACGTCCACGTGATGCTGGACGCGGGTCACCGC
CTCCTGGAGCGCTTCTTCCCCGGCCTCCTCCAGGACCTCCAGGACCAGGGGGCCGCGCTCATCGATTTCGAGCCGG
GACGTGCGCTGGCACCCTTCGGCGTCTGGAAGTCGCGCGTCCCCCAAGGCCTTCCCCTGCTGGTGTGTACGCGT
CCCTTCCTGGAATGGCACGTGCTGCGCCGGGTGCTGGCGCTGCCCAACGTTCGAGTTTCACGGCGGCGTCTCGGTC
GAAGGGCTGCTCACGGACGCCTCCACACAGCGCGTCACCGGCGTGCGGTTGAAGAAGGCCGGCGTGAGGAACCG
CTGGAGGCCGCGCTGGTGGTGGATGCCACGGGCGGGGCTCGCGTGCGCCCCAGTGGCTGGAGGCCCTGGGCCAT
GCACGTCCCGACGAGGAACAGGTGCGGGTGGACCTCGCGTACACCACCCGCTCTACGAACCTCCCGCGCACCGC
CAGGAGGACTGGAAGGTGCTCATGCAGTACCCGTGTCCGCGCGTGAACCTGGCGCGCCGGCTTCATCTCCCGCGTA
GAGGGGGACCGCTGGATTGTACCCCTCAATGGGTACTTCGGCGAGCACGCGCCCCGCGGACGACGAAGGCTTCCTC
GCGTTTCGCGCGCTCACTGCCCCAGCCGGACCTGTACGCCTGCCTGCGGGAGGCCCCGGCCGTTGGGGCCGCTCACC
CTGCACAAGGTGAAGGAGAGCCGCTGGCGGCACTATGAACGCCTCGCGCGCTTCCCGGAGAACTGGATCATCCTC
GGTGACGCCGTGTGTGCCTTCAACCCTGTCTTCGGCCAAGGCATGTGCGTGGCGGCCAGGGCGCGGCGCAGCTC
CTCGCTGCATCGAAGAGCAGGCCAGGCGCTTCCCCACGACGCTCGACGGACTCGCCCAGCGCTTCGCAAGCGC
CTGTCCGACACCATCCGCCTGCCCTGGTTCATGGGCACGAACATCGACCTGCAGTACCCGAAGGCCGTTCGGCCAG
CGGAAGCCCCGGCGTGGGAGTGCTGCACTGGTACATCCGGCGGATGATGGAGCGCAGCTCCCGGGACGCCGCCGTG
CACCGGCAGTTCAACCGACTCCTGCACCTCCAGGCGGGGCTGGGCGCCGTCTCCAGCCTTCGGTGGCCCTGCCA
GTGCTCGCGGACGGCGCCCGCGCGCTCTTCATGCCCCTGCATGCGCGGGCGAACACGGACATCCGCCCCGGCCCCG
CCCTCCTCATCTCCATGACGGACCAGCGCGCAATGCACGCAGCGGCCGAGCCGCTCCGCGAAGAAGTCGAACAG
ACTCACGTCCGTGAGCGCGCGGCCCAATGGCCGGATTCAAACAGCCGGTCTGGTGGCCGAAGCCCGGCGCAAGG
ACCGCCCTGGAGCGAGACAGACTCAAGGGCAGCCGTCTCTCCCTTCAGAAGAAACACCCAGGCTTTGACGGTGCG
GACGTAAGCATCTATGATGCGAACTGATTGCGAGAATCAAAATCACCCCTTTCCAACGACTACCGTGAGTCCATC
CAACAGGCGTTCCCTTCCGCTCTTCGCCCTGCTTTCCGTGCTCGTCTGCTGGCCGTCTGTGGGTGGGCGTGAG
GTTCCAGCGCCCCGACGCCGCGCGTCCACTCCCGCCGCGGAGGCTGCCGCGCCGAGGGCCGGACGGTGACGCA
CGGACAAGGCACCACCGTCGTCCCGCTGAAGCCGCGGCGGGTGGTGGTGTTCGACCTGGTGGCGCTGGACATCCT
CCAGGCCCTGGAGGTGGACGTCCACGGCGTGGCGGGCGACATGTTCCCCCAGCACCTGACCCGGTTCGGGACGC
GAAGTACCCGCGCATGGGCACCCTGTTTCGAGCCTGACTACGAGGCACTCCAGGCCGCGAAGCCGGACCTCATCAT
CACCGGGGGCCGCTCCAGCGCGAAGTACTCGAACCTCTCCCGCATCGCGCCCACCATCGACGTGCCGATGAGCGG
CAAGGACTTCATCGCCTCGGTGATTGCCAACACGGAGATGCTAGCCAGCGTCTTTGGCAAGGAGGAGAAGGCGCG
CGGCCTGATTGAGGACCTGCGCAAGTCGGTGGCGGACCTCCAACAGATGACGGCGACACGCGGCAAGGGGCTGGT
CGTGCTCGTCACCGGAGGGCGCATGAGCGCCTATGGCCCGGGTTCGCGCTTTGGCGTCATCCATGGTGACTTCGG
CGTGCCCGAGGCCGTTCGAGGGCCTGCGCACCTCGCTGCACGGCGAGTCCATCAGCGCGGAGTTCATCCGGGAGAA
GAATCCCGACTGGCTGTTTCGTTCATCGACCGGGACGCCGCCACCGGTGAGAAGGAAGGCAATGCCCGGCAGGTGCT
GGACAATGAATTGGTGCGGCAGACGACGGCCTGGCAGAAGAACCAGGTGATCTACCTGGACCCCGGCGTCACCTA
CCTGACCGGCGGCGGCATCCAGTCCGTCAAGCAGCTCCGCGACCAGGTTCGCGAGCGCTACACGCAGGCCCCAGTA
GTCCCTCCCCACCGTGAAGCGCGCGTTTCGCCGCCGCCGCGCTCTCGTCACGTTGGCGGTGGTCAGCCTCCTGAT

```

TGGCGCCAGTGACGTGTCCTGGCGCGCCCTCTTCGCTCCCGAGCCGGATGAGCGCGCCCTCCAGGTGCTGGTCAT
CAGCCGGCTGCCGCGCCTGTTTCGCCGTACATGCTGGCGGGCACATCCCTGGGCGTCGCGGGCCTCATCATGCAGAT
GATTGCCCGCAACCGCTTCGTGGAGCCCACCACAGCGGGCACCGCGGAGTCCGCCAGCCTGGGCCTGCTGACCGC
CACCTGCTGGCGCCTGGCCTTCCGGTGCTCGGGAAGATGATGGTGGCCACCGTCTTCGCCCTCGCGGGGACGGC
GCTGTTCTGCTGGTGCTGCGGCGCATTCCTTCGCTCGGCCCTCATCGTCCCGGTGGTGGGCCTGATTCTGGG
CGGCATCATCGATTGCGGACGACCTTCTTCGCCTACCGGTACGACCTGCTTCAGACGGTCAACGCGTGGACCAC
GGGTGACTTCTCCACCGTGCTTCGCGGCCGTTACGAGCTGCTGTGGGTGACGCTGGGCCTCACCTGCGCCGCCA
CACTGCCGCGGACCGCTTCACCGTGCGCGGCATGGGCGAGACGTTACACCACCAACCTGGGGCTGAACTACCCGCG
CATCGTCGCGCTGGGGCTGGTCATCGTCGCCCTGGTCACCGCGATGGTGGTTCGTCACGGTGGGCATGATTCCGTT
CCTGGGACTGATTGTCCCCAACCTGGTCAGCATGGTCATGGGCGACAACGCGCGCAGGTCCATTCCCTGGGTGCG
GGTGAGCGGCGCGCCCTTCGTGCTGCTCTGCGACATCGTGGGCGCGGTGGTGCGCCACCCGTATGAGATTCCCGG
AGGCACCATCGCCGGCGTCATCGGCAGCGTGCTCTTCTGTACCTCCTGCTGAAGCGAGGCGCCCGTGCCGGCTA
ACGCCCCGCCCCGTACCCCCCTGCGGCCCTGCTCGTGCTGGGCTGCGTGCCGTGGCCTTCATGGCCCTGTTCA
TGCTCGTCGACGTGAGCGGCCCGTGGGACTTCGTGCTGCCCTTCCGCGGGAAGAAGGTGCGCCACCGCGTTGCTGG
TGGGCTACGCCGTGCGCGTCTCCACGGTGCTCTTCCAGACGGTGACGGGCAACCGCGTGCTGACGCCCCCATCA
TGGGGTTCGACTACCTGTACGTGCTCATCCAGACGTGCCTGGTGTCTTCTTCTGGGCTCCACCACCGTGGCGGGG
TGGACCCCCGTTTCTGTTTCGGCGCGGAGGTACATCATATGGTGGCCTTCTCCGCCATGCTGCACGGCTGGCTCT
TCGGCATGGCCCGGGGCAACGTCCACCTGCTACTGCTCACTGGCGTGGTGATGGGCGTGCTGTTCCGGAGCCTGT
CGTCGTTTCGTCCAGCGCGTCATCGAACCCAACGAGTTCATCTTCTGTCAGGACCGCTTCTTCGCCAGCTTCAACG
ACCCGGAGCACGACCTGCTGCTGCTCTCCGCGGTGCTGACGCTGGGGGTCTCCGTGCTGGGCCTGCGGCTGCTGC
GCGCCTGTGACGTCCTGGTGCTGGGCGGGAGGCCGCCATCAACCTGGGCGTGGACTACCAGCGGACGGTGTCAT
GGGTGCTGGCGCTGGTGGCCATCCTGGTCGCCGTGTCCACCGCGCTGGTGGGGCCCGTGACGTTCTGGGGTTGC
TGGTGGCCAACATCGCCTACGCGGTGATGCGGACGTACCAGCACGCGTTTCGTCTCCCCGCCGCCGCGCTCATCG
CCGCCATCGCCCTGATCGCCGGGACGTTTCATGCTGGAGCGCGTCTTCCGCCTGGACACCAATCCACGGGTTCATCA
TCGAGTTCGTGCGCGGGATGGTGTTTCATCGCGATGCTCATGAGAAAGGCCACGAGATGATCGAGGCGAAGAACGT
CAGCAAGCGCTACGGTGAAACGCTGGTGGTGGATGGCGTCACCTGAGCCTGCCCCGTGGGCGGCGTCACTCCAT
CATCGGCCCCAACGGCGCGGGCAAGTCGACCCTGCTGTCCATGATGAGCCGGGTGATGCCCATGTTCATCGGGCAA
CGTTTGGGTGGACGGGCTGGACGTACCAACACGCCGGGGGACACGCTCGCGAAGCGGCTGGCCATCCTCCGGCA
GGACAACCACATCTCCGCCCGGCTGACGGTGCGTGAGCTGGTGACCTTCGGCCGCTATCCGCACACCAAGGGCCG
CCCCACCGTGACGACCGCGAACACGTGGAGCGGGCCATCGAGCACATGGGGCTGAGCGCGCTCGCCGACCGCTT
CCTGGATGAGATGTCCGGCGGCCAGCGTCAGCGCGCCTTCGTGGCCATGGTGCTCTGCCAGGACACGGACACGT
GCTGCTGGACGAGCCGCTCAACGGCCTGGACCTGAAACACGCGGTGTCCATGATGAAGCGGCTCCGGCACGCCGC
GGACACGCTGGGCAAGAGCTTCGTGCTGGTGCTGCACGACATCAACTTCGCCTCCTGCTACTCCGACCACATCGT
CGCCATGCGCGACGGCAAGGTGCGGTTCCAGGGCCGCCCGGAGGACATCATGCGCCCCGACGTCTGCGCGCCAT
CTACGAGTTGGACATCTCCGTGCAGCAGATTGAAGGCGACTGGATTGCCACCCACTACCGGTGATGGCGGGCGC
GCCAGCCTGGGATGCAAGGGGTTTCCACTAATCACAGACACTTGCACTGCCTTCTCAGGCTGTGCCACATTTAC
CGCGCTGGCCATGACGTACGACAGACGAAGACGCCACCACCTCAAACCAGGAGTTCGCCATGAAGAAGGCCCTG
TACTCTTTGGCTGTGCTGATGCGGTTTGCTCGCGCGGACAAGCTCTCCGCCCCGTATATCTATTACTAGTCAAAA
TCCGTACGGCATTGAGCCCTGACGGTTGGGTCTCAAGAGCTGGTCGACGTGGGTGGAGCGCATGGTTTCGTCCAC
CTGCGCCAGCGGTAGTTCCCCCTCCAGGACCCTTTTCATGTACAGGTGCGCCTTCCGCCGGGGCCCCAGGGCCA
CTTCATTGCTGGCAACCTGGTGGAGTTCTCCGAGGACCCGCTGGGGTTCTCACCCTGCGCCAGGGAGTACGG
CGACGTGGTGCGCTGGGGAAGCGCAACTTCTCCTCAATCACCCGGACCTGATTGAACGCGTCTGGTCAACGG

CGACGGCAACTTTGTGAAGCTCGCCGGAGTGGGCCAAGGCAAGCGTCACAAAGGCGGCTTTCCCGAAGCGATGAT
GAACAGCGAGGGCGAGGACTGGCTGCGCAAGCGGCGCCTCGTCCAGCCCGCGTTCCATCGCAAGCACGTGGCCGC
GTGTGGTGACACCGTGGTGGCCCTACCGAAACGATGCTCCAGACGTGGCGCCCGGGCGACGCGCGGGACGTGCA
CGCGGACGTGTCCGCGCTGGCGCTCGACATCGTCAGCCGATTCTCTTCCACACGCCCATCGACGACGAGGCCCCG
CCACGTGGCCGACGCCGTGGACGCTGTCATGCGGCACACCGACAGCCCCCTGCGGCCGCCCATCTGGGTGCCAC
GCCCACCAACCTCCGCCTGCGCCGCGCCCTGGGCCGGCTCAACACGCTGCTCGCGACGCTCGTGCGCCGCTATCG
GGAGCAACCGGAGTCGCGCACGGACCTGCTCGCCCTGCTGCTGTCCGCTCCGGTCCCCCTCTCCGAGAATCAACT
GCGCGATGAGCTGGCGACGATGATCATGTCCGGGCACGAGACGACGGCGGACGCGCTCGTGTGGGCCTGGTACCT
GCTGGCCCAGCACCCGGAGGCCGAGGCCCGGCTGGTGGCGGAGCTGGAGACGGTGCTCGGGGGCCGGCTCCCTGG
CGCCGAGGACCTTCCTCGCCTGCGCTACACGGAGGCGGTGGTGAAGGAGGCCATGCGCCTGTACTCACCGGCGTG
GATCACCAGCCGGGAGGCGCTGCGCGACTGTGAGCTCGGCGGGTTCCACGTCCCAGCGGGAACGATGCTGGCGGT
GAGCCAGTGGGTACCCACCGCGACGCGCGCTACTTCGACGCCCCCGAGTCCTTCCGTCCAGACCGCTGGCTGTC
CGAGGACGCCCAACGCATGCACCGGTACGTCTACTTCCCGTTCGGCGGCGGCCCCCGGTTCTGCATCGGCTCGGC
GCTGGCGATGATGGAGACGGTGCTCATCACCGCTGCGTGGCACGGCGGTTCGGCTGGAGCTGGCACCGGGCTG
CGTGGTGCGGCCCCGGCCGGCGCTCGCGCTCCAGCCGCTCGGGGTCTGGCTGATTCCGCGGCATCGCTCTCACAC
AACACAGGAAGGCGAGGTGCGGCATGCGGCGGGAGCATGAAGGGGACGCGGTCCGGCGCGAGTTCTTCCGCTCGG
CATACTGCTGGGCGCACCTGGGACATCGGACGGCCCCAACAGGCCTTCGTCCAGTTGTGGGAGGCCGGCGCGA
TTTCCGGCGAGGTGCTCGACGTGGGCTGCGGCTTCGCGGAGAACGCCCTCTTTCTGGCGGCCAAGGGCCTGCCGG
TGTGCGGCGTGGACATGATGGAGCCGGCCATCCTGCGAGCACGCGAGACGGCCTCTCTGCGCGGGCTCGACGTGG
ACCTGCGTGTGGGCAACGCCCTGGAGCTGGCCACGCTGGGCCGGCGCTTCGACACCATCCTTGATTGCGCCCTGC
TCCACGTCTTCGAGCCTGGAGACCGGCCCGCGTTGCGCGCCAGCCTCGCGAGCGTGCTCCGGCCCCGGCGGGCACT
ACCATGCGCTGTACTTCCGCGACGGACCGCGCGCCGTTCCACCGGAGGCACTGAACGCGACCTTCGGGGAAGGCT
GGCGCGTGAAGTCCATCCAGGAGGCGCACTACGAGCAGACCGACCCGGAGGGCGCGCAGGCCTGGCTCGCCACCA
TCGAACGGGTGCCTCCTTCGTTCATCGACAGAGGACTGA

Chapter 2

Chapter 2A

**The structure elucidation and biosynthesis of novel
Myxoprincomides from *Myxococcus xanthus* DK897**

Chapter 2B

**Myxoprincomide: A Natural Product from
Myxococcus xanthus Discovered by Comprehensive
Analysis of the Secondary Metabolome**

Chapter 2A

The structure elucidation and biosynthesis of novel Myxoprincomides from *Myxococcus xanthus* DK897

Abstract

Myxococcus xanthus is a multiproducer of natural products. There are common secondary metabolites found in all *M. xanthus* strains, while others are only produced by a few strains. From *M. xanthus* DK897 extracts, novel myxoprincomides could be detected. Myxoprincomides are apparently produced by every *M. xanthus* strain analyzed to date, but two different types exist. The myxoprincomide family of peptides produced by *M. xanthus* DK897 comprises ten different derivatives which were isolated and characterized in this study. The biosynthetic genes were identified by transposon mutagenesis and the myxoprincomide gene cluster (*mxp*) was assembled using draft whole-genome sequence information. The *mxp* locus is largely similar to the related MXAN_3779 gene from *M. xanthus* DK1622, but distinct differences in the sequence exist which seem to lead to increased compound diversity. Interestingly, the structures of several myxoprincomides correspond well to intermediates expected to be formed on the NRPS-PKS assembly-line during myxoprincomide production. Thus, myxoprincomide biosynthesis is an intriguing example where release of intermediates from a giant biosynthetic multidomain enzyme can be observed, providing insights into the action of a biosynthetic assembly line *in-vivo*.

Introduction

Myxobacteria are a potent source of secondary metabolites. More than 100 core structures and some 500 derivatives thereof have been identified. Most notable are the epothilones, whose semisynthetic analogue Ixabepilone was recently FDA approved for the US market as an anticancer drug. A huge number of myxobacterial natural products show remarkable biological activities.^[16;204;205] These compounds are often products of polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs). In the case of myxobacteria NRPS/PKS hybrid synthetases are very common and frequently produce novel core structures.

Often the large myxobacterial genomes harbor several biosynthetic gene clusters encoding for secondary metabolites. The bioinformatic analysis of whole-genome information from *Myxococcus xanthus* DK1622 for example suggests that its biosynthetic potential is much larger than expected on the basis of known secondary metabolites from this strain. DK1622 was long time used as a model organism for studies of myxobacterial social motility and multicellular differentiation.^[36] It was initially not even recognized as a producer of secondary metabolites, until analysis of its genome revealed 18 biosynthetic gene clusters encoding for PKS, NRPS and hybrids thereof.^[91] Eight compound classes have since been assigned to biosynthetic gene clusters in DK1622: the DKxanthenes, the myxochromides, the myxalamides, the myxochelins, myxovirescins, myxoprincomides and two substances with unknown structure.^[90;176;195;206-209]

Due to the striking discrepancy between numbers of genome-encoded biosynthetic pathways and known metabolites, it is most likely that more secondary metabolites can be isolated, not only in DK1622. Other *M. xanthus* strains are assumed to be secondary metabolite multiproducers as well. Hence a large chemical screening was performed on 98 *M. xanthus* strains isolated from locations worldwide, as well as several strains attained from single habitat.^[51] The target was to find unknown metabolites, beside those already identified ones like DKxanthenes^[209], myxochelin^[83] from *M. xanthus* DK1622. The analysis was based on an HPLC-HR-MS screening, assisted by principal component analysis (PCA), a statistical tool to limit the manual effort.^[210] PCA facilitates the analysis of multivariate data sets. To compare each data point in the HPLC-MS chromatogram with the received chromatographic data of other strains the computational methodology creates n-dimensional models based on molecular features. The software extracts significant information from complex data sets. The received information here resulting in outliers was verified by analyzing the raw LC-MS data.

With this method several potential novel secondary metabolite candidates have been detected in *M. xanthus*.^[51] The next goal was on the one hand to isolate and characterize these supposedly metabolites and on the other hand to identify the genes underlying their biosynthesis. A detected group of candidates was isolated and correlated to their biosynthetic origin. We present the characterization of the myxoprincomide NRPS – PKS hybrid biosynthetic pathway in *M. xanthus* DK897. It is in charge of at least ten novel myxoprincomides which were structurally elucidated. Several of these structures are identified as biosynthetic intermediates of the myxoprincomide formation. The analysis of the *mvp* gene cluster in *M. xanthus* DK897 allows insights into an "imperfectly" working NRPS-PKS biosynthetic assembly line with unusual biosynthetic features.

Results and Discussion

Screening for novel metabolites and their corresponding genes

In a screening of *M. xanthus* strains 18 potential novel compounds beside myxochelin and cittilin were reported.^[51] Apart from the known substance groups myxalamides, myxochromides, DKxanthenes, one novel group comprising three metabolites was detected in strain DK897, which is not produced by the model strain DK1622 (Figure 38).

In order to correlate biosynthetic gene clusters to the potentially novel metabolites predicted from strain DK897, genes were randomly inactivated via transposon mutagenesis. The mutants were subsequently analyzed for the lack of these substances.

In search for the corresponding gene cluster, an existing transposon mutagenesis library consisting of *M. xanthus* DK897 mutants was screened. Mutants have been cultivated in microtiter plates and 2000 mutants were analyzed by HPLC-MS by D. Krug and A. Sandmann. This analysis identified two knockout mutants deficient in production of the compounds with $m/z = 382$, $m/z = 800$ and $m/z = 814$. Transposon recovery was then performed to identify the inactivated genes in this mutant, followed by primer walking to obtain initial sequence information. Moreover, additional genetic information was acquired with the help of further knockout mutants and subsequent plasmid recovery. The recovered plasmids were sequenced; using this approach, a 10460 bp stretch was sequenced. Concurrently, the genome of DK897 was subjected to whole-genome sequencing, yielding more than 500 contigs. The partial cluster sequence as obtained from plasmid recovery was

aligned with these contigs. Five contigs were found with 100% identity to the partial gene cluster sequence (Figure 39).

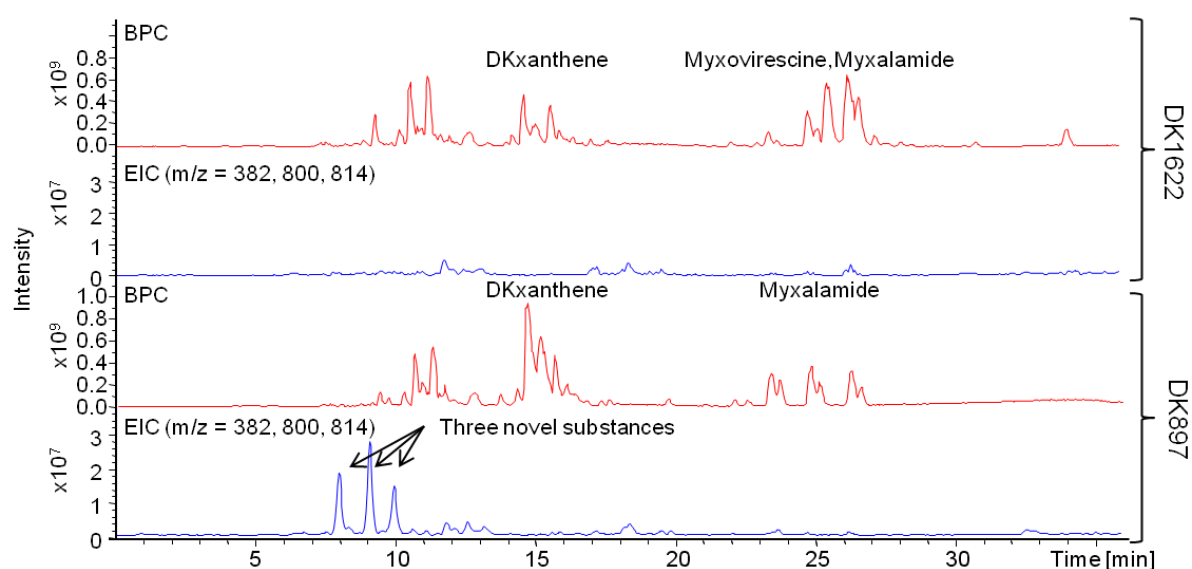


Figure 38: HPLC-MS chromatograms of extracts of *M. xanthus* strains. Base peak (BPC) and extracted ion chromatogram (EIC) of the two strains DK1622 and DK897.

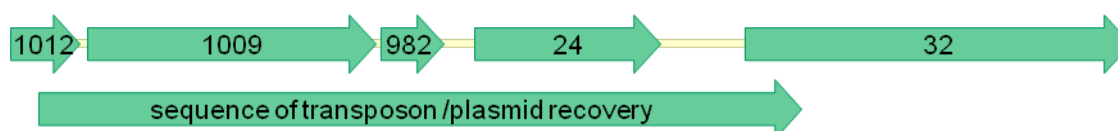


Figure 39: Order of the gene cluster sequence of transposon and plasmid recovery aligned to the contigs: contig1012 (950 bp), contig1009 (3947 bp), contig982 (860 bp) contig24 (2555 bp), contig32 (5276 bp) in total 15342 bp.

To confirm that these contigs were definitely part of one gene cluster multiple PCRs were performed, and used to close the gaps by sequencing of the obtained amplified fragments. By combining the sequence information from contigs, transposon recovery plasmids and PCR products, altogether the sequenced region added up to 15342 bp.

Next, the DK897 genome was subjected to paired-end sequencing to yield 139 scaffolds. The 15 kbp sequence previously established was aligned with the new scaffolds of DK897. One large scaffold with a length of 26843 bp (scaffold 0024) containing the previously identified 15 kbp segment was found (Figure 40). Gaps within this scaffold were closed by PCR and subsequent sequencing.

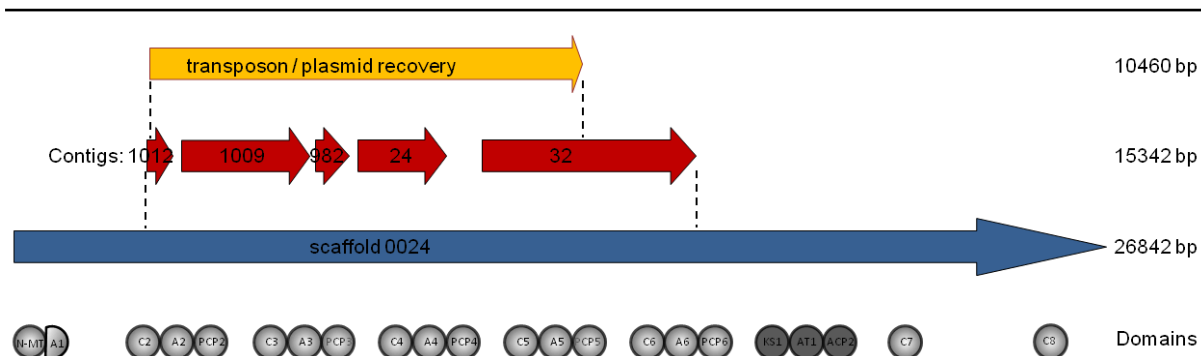


Figure 40: Scaffold 0024 in comparison to the contigs and the known sequence from primer walking.

Bioinformatic prediction of biosynthetic domains in the large reading frame spanning scaffold0024 revealed intriguing similarity to a different biosynthetic gene cluster in *M. xanthus* DK1622 (MXAN_3779), considering the order of domains and their predicted substrate specificity (Figure 41). MXAN_3779 encodes for a hybrid NRPS-PKS megasynthetase which consists of 13 modules. The loading module contains an acyl-CoA ligase domain and a PCP domain. Module 1 contains an *N*-methyltransferase, the A domain has predicted L-serine specificity. Four further NRPS modules are following with specificities for: an unknown amino acid, L-valine, L-serine and for another unknown amino acid. Module 6 is a PKS module followed by a module containing an oxidation domain (Ox) within a presumed A domain. The assembly line continues with further NRPS modules.

In the case of scaffold0024 from DK897 there is an *N*-methyltransferase, followed by five modules with similar specificity predictions as described for the DK1622 gene cluster. A PKS module and one module with an Ox domain within an A domain follow.

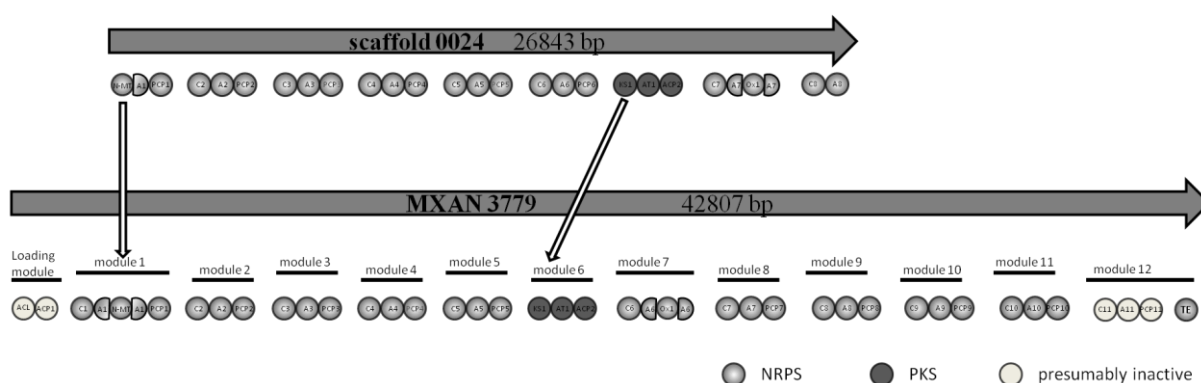


Figure 41: MXAN_3779 of DK1622 in comparison to scaffold 0024 of DK897.

In a study concerned with identification of metabolites that correlate to the MXAN_3779 gene cluster in DK1622, a novel compound with m/z $[M+2H]^{2+} = 506$ was identified (see chapter

gmxxorf2004

1 5,000 10,000 15,000 20,000 25,000 30,000 35,000 40,000 42,825

NRPS

A? C A. A C A. C A. C A. C ? C A C A A. A C A.

scaffold 00036

A. Pat

scaffold 00037

20... A? C

scaffold 00110

C A. C

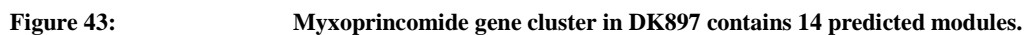
A.

scaffold 0024

C A. C A. A A K C ?? C

A.

The gaps within these scaffolds could be estimated by using the MXAN_3779 sequence. PCR amplification and subsequent sequencing was used to connect the scaffolds with the missing sequence. Finally the gene cluster from DK897 consists of 46959 bp (Figure 43).



74

A-Domain	Prediction with A ^[211]	Prediction with B ^[212]
A1	Ser	Ser
A2	Orn	Orn
A3	No	No
A4	Hpg	Val
A5	Ser	Ser
A6	No	No
A7	No	No
A8	Ser	Ser
A9	No	Tyr
A10	Ahp	Lys-b
A11	No	No
A12	No	Val

Table 2: A domain prediction using the web tools of A: J. Ravel et al. and B: T. Weber et al. (the substrate specificity code is depicted in Table 8).

For the A domains three times serine, twice valine, one tyrosine, one lysine, one isoleucine and four unknown amino acids were predicted. If the structure was presumed to be collinear with domain organization, the predicted small molecule outcome should be similar to the structure depicted in Figure 44. Unknown amino acid residues are marked with R. However the chemical formula for this theoretical structure would be at least $C_{51}H_{83}N_{13}O_{18}$ which is not in agreement with the proposed molecular formula (based on high-resolution MS) of the largest detected molecule, calculated as $C_{54}H_{83}N_{11}O_{18}$. This indicates that two nitrogens are dispensable, and thus at least one module could possibly be inactive.

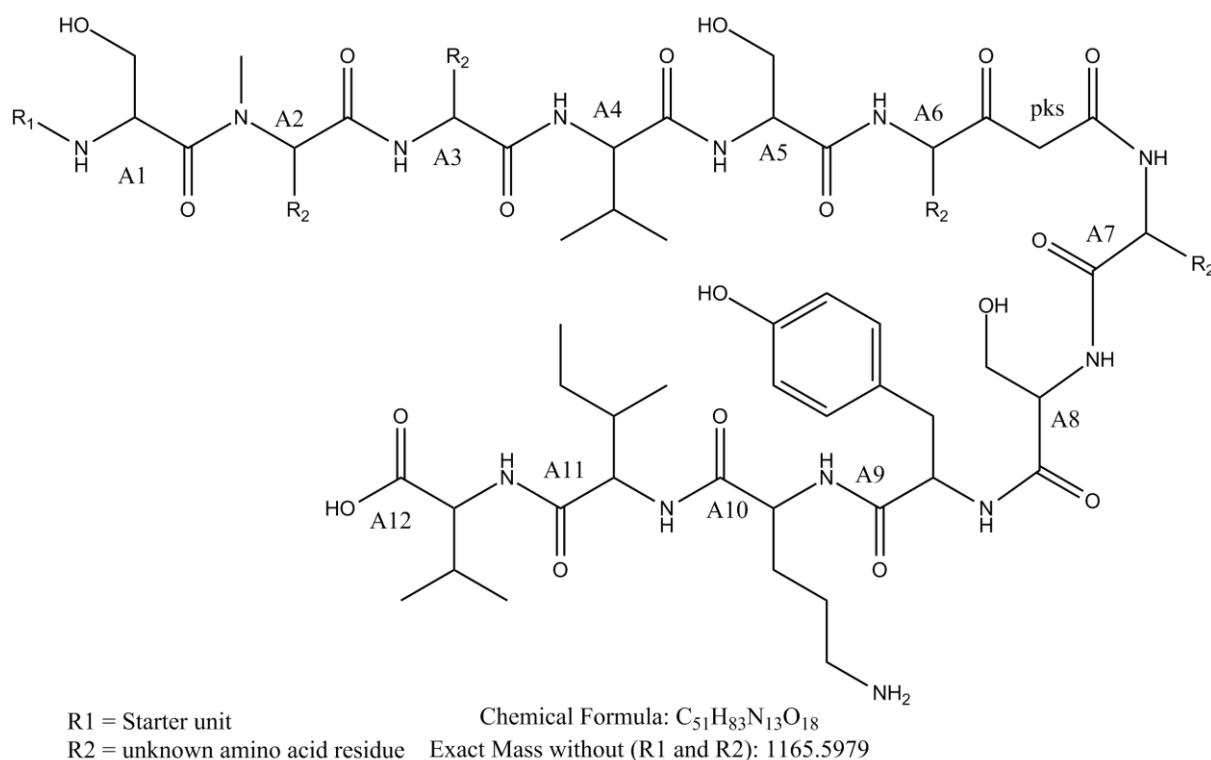


Figure 44: Predicted structure if standard textbook biochemistry would be applied by gene cluster. The minimal sum formula would be $C_{51}H_{83}N_{13}O_{18}$.

On protein level the different domains can be identified by highly conserved sequence motifs (core motifs or conserved regions).^[102] Diverse domains have different numbers of core motifs (e.g. A domains have 10 and C domains 7). If such regions are missing it is often a reason for the inactivity of the domain. Therefore an attempt was made to predict directly from the sequence data which domain might be inactive through an alignment of the A domains (Supplementary figure 81); the 10 conserved regions were checked for major differences. In general all core motifs of the A domains are present. But in the case of A1 the core motifs are present but region 9 and 10 are located after the *N*-methyltransferase domain, like it is known from other NRPS clusters.^[136;213]

Similar as with the A domains, the conserved regions of all C domains were also analyzed (Supplementary figure 82).^[102] Some irregularity was located in module one and six; these condensation domains possess unusual C5 and C7 motifs. This might hint at no usage of module 1 and 6. Furthermore module eight exhibits an oxidation domain within the A domain and the PCP is lacking. That means module eight cannot function as a common module.

The analysis of all PCPs revealed that the conserved motifs are present and intact (Supplementary figure 83).

Module 7 is a minimal polyketide synthase (PKS) module consisting of a ketosynthase (KS), an acyltransferase (AT) and an acyl carrier protein (ACP). The question for the PKS module is if it uses malonyl-CoA or methylmalonyl-CoA as substrate. For a prediction a phylogenetic tree of malonyl-CoA and methylmalonyl-CoA specific ATs of different myxobacterial PKS clusters was performed (Figure 45). The phylogenetic tree revealed that the AT domain of the myxoprincomide cluster groups slightly more with methylmalonyl-CoA specific ATs than with the malonyl-CoA specific ones, but it is still separate. Which might mean either that it is inactive or it is a novel type of AT where the type of substrate specificity cannot be predicted at the moment.

To conclude, the domain analysis did not yield in striking information to make a reliable substance prediction. Therefore the compounds had to be isolated.

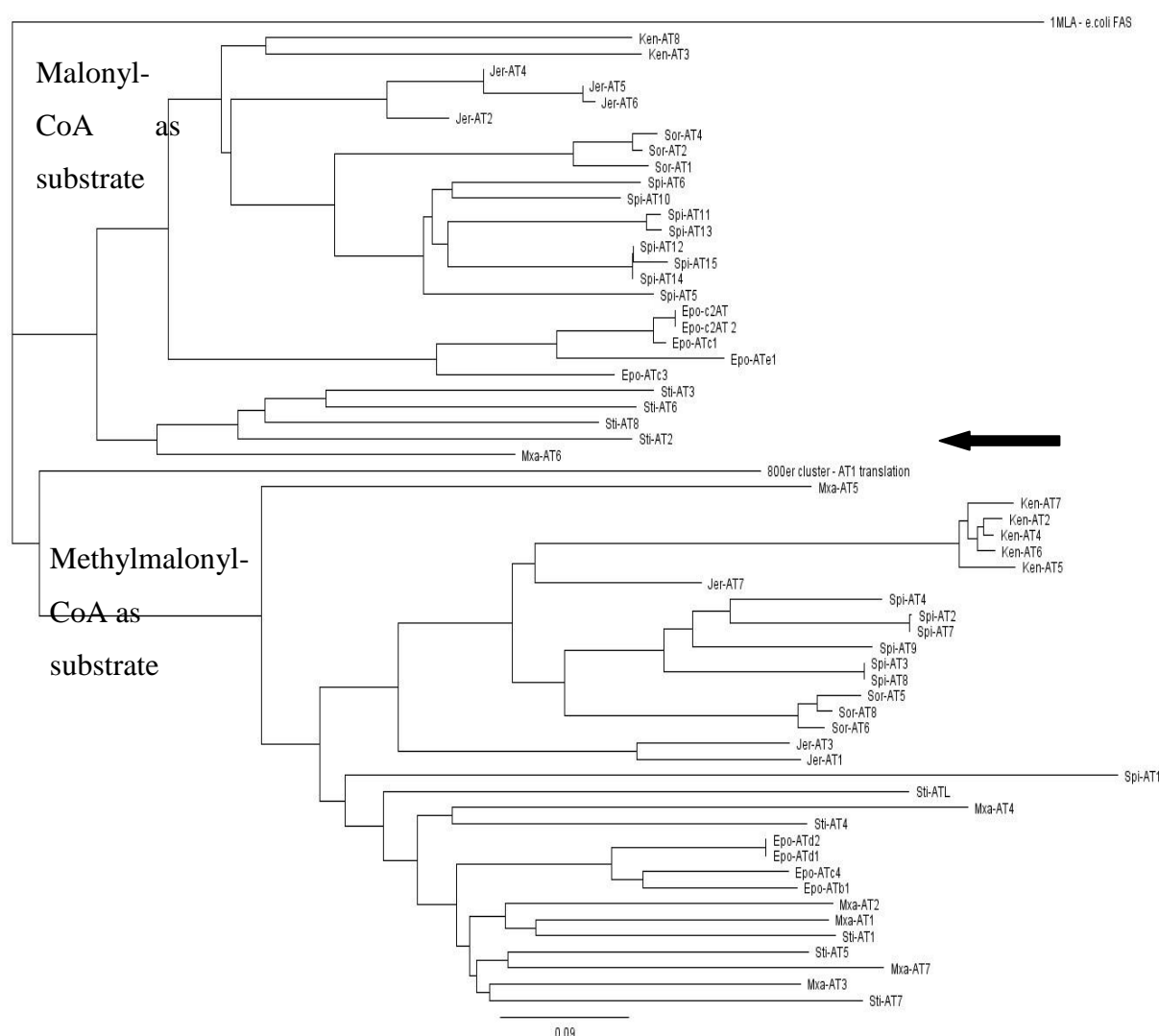


Figure 45: Phylogenetic tree of malonyl-CoA and methylmaloyl-CoA specific ATs domains of different myxobacterial PKS clusters compared the one AT in the myxoprincomide cluster (marked with arrow).

Substance characteristics

At the beginning it was thought to be a structure class of only three derivatives (c382, c800, c814). But later two further compounds c580 and c587 could be correlated to the biosynthetic gene cluster (refer to Figure 46). Further depth comparison of wild type and mutant extracts revealed finally ten different derivatives produced by the *mvp*_{DK897} cluster: c382, c683, c798, c800, c812, c814, c580, c581, c587 and c588. The substances exhibit a typical peptide UV-spectra. HR-MS measurements lead to the possible molecular formulas given in Table 3. To prove the accuracy of these formulae a reversed feeding experiment was carried out. The strain DK897 was cultivated in ¹³C / ¹⁵N background media. The extracts were analyzed with HR-MS, and compared to a mixture of the extracts of CTT with ¹³C, CTT with ¹⁵N, and a mixture of CTT, ¹⁵N and ¹³C. Using this technique all sum formulae could be confirmed by mass shifts depending on the number of carbons and nitrogens (Table 4). However, due to a very bad growth of the strain in the isotope background media (differs to the typical media composition for myxobacteria) only three myxoprincomides could be analyzed.

Compound	[m/z]	Calculated [m/z]	Sum formula
c382	382.1963	382.1973	C ₁₈ H ₂₇ N ₃ O ₆
c683	683.3602	683.3610	C ₃₁ H ₅₀ N ₆ O ₁₁
c798	798.3879	798.3880	C ₃₅ H ₅₅ N ₇ O ₁₄
c800	800.4187	800.4036	C ₃₅ H ₅₇ N ₇ O ₁₄
c812	812.4035	812.4036	C ₃₆ H ₅₇ N ₇ O ₁₄
c814	814.4189	814.4193	C ₃₆ H ₅₉ N ₇ O ₁₄
c580	580.7949	580.7954* (1160.5834)	C ₅₃ H ₈₁ N ₁₁ O ₁₈
c581	581.8027	581.8031* (1162.5990)	C ₅₃ H ₈₃ N ₁₁ O ₁₈
c587	587.8028	587.8031* (1174.5990)	C ₅₄ H ₈₃ N ₁₁ O ₁₈
c588	588.8107	588.8110* (1176.6147)	C ₅₄ H ₈₅ N ₁₁ O ₁₈

Table 3: Molecular formulas of myxoprincomides in DK897; * doubly charged.

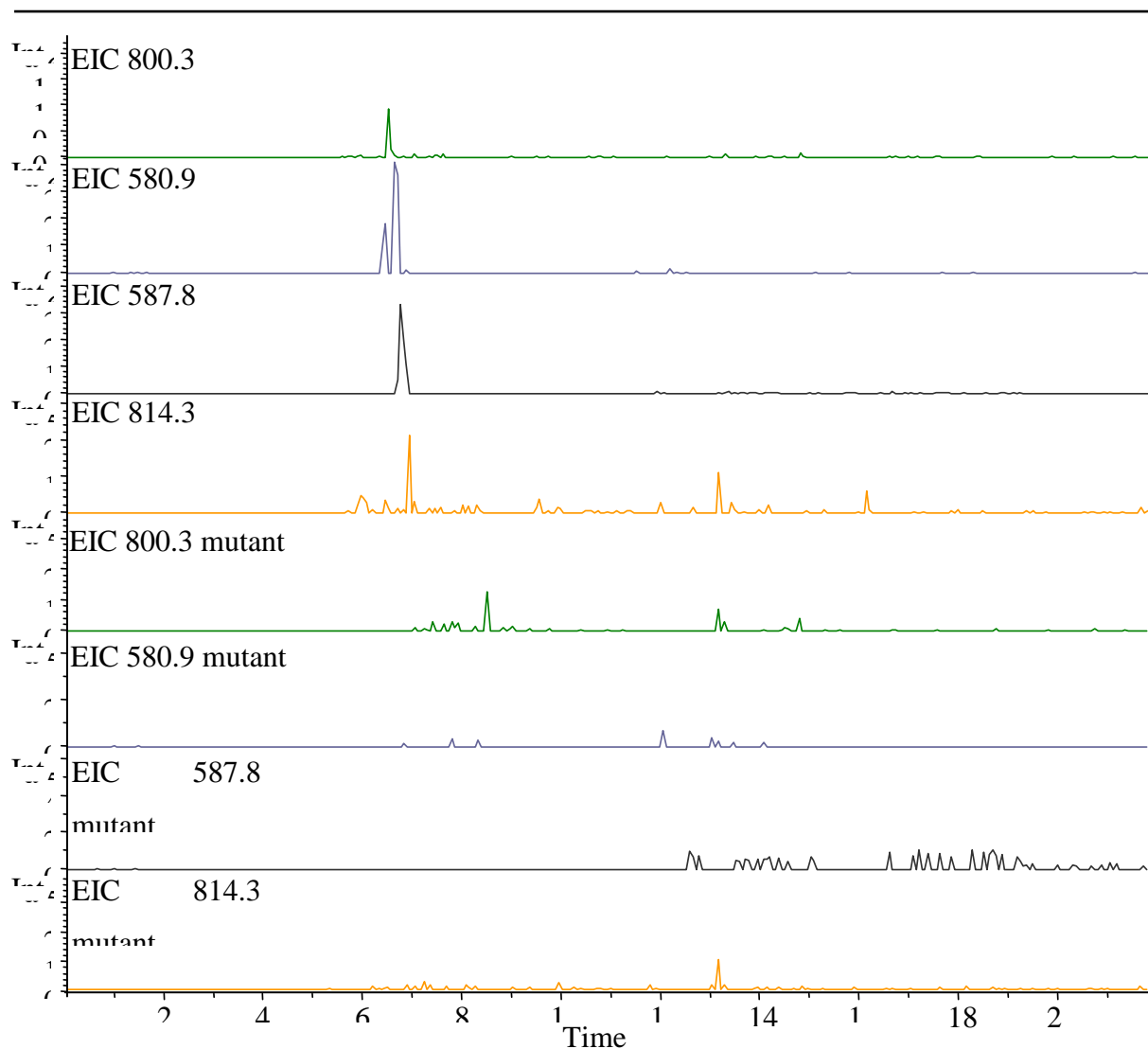


Figure 46: Extracted ion chromatograms of wild type and myxoprincomide knockout mutant showing that c800, c580, c587 and c587 are all compounds of the origin of one biosynthetic gene cluster.

Compound	[m/z]	[m/z] ¹³ C isotopes	No. of C	[m/z] ¹⁵ N isotopes	No. of N	Sum formula
c382	382.19735	400.25827	18	385.18881	3	C ₁₈ H ₂₈ O ₆ N ₃
c800	800.40381	835.51959	35	807,38257	7	C ₃₅ H ₅₈ O ₁₄ N ₇
c814	814.41965	850,53760	36	821,39923	7	C ₃₅ H ₅₈ O ₁₄ N ₇

Table 4: Confirmation of the molecular formulae of myxoprincomides in DK897 by incorporation of ¹³C and ¹⁵N.

In a next step the predicted A-domain specificity was checked. Therefore feeding experiments with the following amino acids were performed: L-Methionine-(methyl- D_3), L-Phenylalanine (-ring- $^{13}C_6$), L-Serine-($^{13}C_3$, ^{15}N), L-Valine- D_8 . The amino acids were added in small portions to the culture to a final concentration of 1 mM. The culture extracts were analyzed by HPLC-MS. Valine is incorporated once, but with a mass shift of +7. Serine is incorporated several times but the exact amount was not clear as the mass shifts were not clearly defined for multiple incorporations. In contrast phenylalanine feeding showed a clear mass shift of 6 units that means exact one phenylalanine is embedded. The feeding of methionine resulted in a mass shift of +3. By feeding labeled methionine a labeled SAM pool is build which is then utilized for *N*-methylations. If methionine would be incorporated as whole amino acid the characteristic isotopic pattern would change because of the sulphur isotopes. Here only the methyl group of methionine is incorporated. The results of the amino acid feedings are shown in Table 5. The findings were in good agreement to the predictions for the A domains of the myxoprincomide gene cluster.

Compound*	L-Methionine- (methyl- D_3)	L-Phenylalanine (-ring- $^{13}C_6$)	L-Valine (D_8)	L-Serine ($^{13}C_3$, ^{15}N)
c382	1	1	1	1
c581	1	1	1	3 - 4
c587	1	1	1	3 - 4
c683	1	1	1	3 - 4
c800	1	1	1	2 - 4
c814	1	1	1	3 - 4

Table 5: Number of incorporated labeled amino acids in the myxoprincomides. Methionine is incorporated as *N*-methyl- D_3 . * Not every compound was produced.

Purification procedure

Since the compounds could not be predicted from the genes it was planned to isolate and to characterize them. But the first problem was to obtain sufficient material for compound purification and structural elucidation. Isolation of the myxoprincomides from an extract of a

70 L cultivation of DK897 was attempted. All isolation methods failed to yield measurable amounts of myxoprincomides. Reasons for that were the low production rate and the strong polarity of the substance. Major parts of the compound were not absorbed by the absorber resin XAD16. After extractions with ethyl acetate or butanol the myxoprincomides stayed mainly in the water phase.

As soon as the complete sequence of the gene was established and it was known that the megasynthetase is encoded by a single gene, production of the new compounds could be improved by introducing a promoter upstream of the gene. A promoter was inserted upstream of the gene by a single cross over experiment resulting in the mutant called DK897_800erstart. The chosen promoter was known to increase the production rate in myxobacteria. The same techniques as used by N. Socorro Cortina et al. were employed.^[176] The clones were tested for their production by HPLC-MS. The promoter insertion yielded in a high production of myxoprincomides.

Beside that different media (like TS, CYE, M) were tested for improved myxoprincomide yield. Furthermore the addition of Fe(III)-EDTA was used to reduce the amount of produced myxochelin as it coelutes in the same range as the targeted compound. Using these conditions some of the myxoprincomides could be detected in the MS-BPC which was never possible in the wild type before. That means that the production rate was significantly increased. Therefore the DK897_800erstart mutant was cultivated in a 300 L fermenter.

The obtained XAD was transferred into a big glass column, extracted with water, followed by a mixture of water / MeOH (1:1), and pure MeOH. Major parts of the compound myxoprincomides were detected in the methanol water mixture and in the first pure methanol fractions. The solvent was removed, diluted in MeOH and defatted with n-heptane. The crude methanol extract was dissolved in water and portioned with ethyl acetate. The remaining crude extract was loaded onto a methanolic Sephadex[®] LH-20 column.

The fractions positive for myxoprincomides were further purified using a preparative HPLC-MS system equipped with an Xbridge C18 OBD column (Waters). First purification round using methanol / water eluent (pH = 3), second round same eluent but pH = 9.4 and a last purification step with acetonitrile as eluent (pH = 3). Impure fractions were further separated via a methanol Sephadex[®] LH-20 column to obtain pure substances of c382, c683, c798, c800, c812, c814, c582 and c587.

Structure elucidation of the myxoprincomides

For structure elucidation one and two dimensional NMR spectra were generated (spectra depicted in Supplementary figure 11 to Supplementary figure 71). The structure was solved step by step through identifying each single amino acid in the peptide chain. COSY and TOCSY spectra were used for spin system analysis. The correlations are depicted in Figure 47. CH correlations were received from HSQC spectra. For long range C - H correlations HMBC data were used (Figure 48). All experiments were carried out in d_6 -DMSO. The detailed NMR-data for the myxoprincomides are shown in Supplementary table 1 to Supplementary table 8.

As all ten peptides have similar core structure their structure elucidation is exemplified by c800 (**43**) as the best NMR data were obtained here; data shown in Table 6. In total seven amino acids (AA1 to AA7) could be identified for **43**:

AA1: The myxoprincomide structure begins with a methylated serine (*N*MeSer). In this spin system of *N*MeSer two protons are connected to a hydroxylated carbon atom. Next to that group a single proton is bound to a carbon with a shift of 64.4 ppm which shows an HMBC correlation (Figure 48) to a methyl group with a chemical shift of 33.1 ppm indicating an *N*-methyl group.

AA2: Adjacent to AA1 a β -oxidized valine (OH-Val1) is bound. This is reflected in two methyl groups showing a singlet in the proton spectra. These methyl groups are in the neighborhood of a quaternary carbon with a shift of 71.7 ppm. A chemical shift of 71.7 ppm indicates a hydroxyl group linked to a carbon. This carbon displays a HMBC correlation (Figure 48) to the α -proton of the amino acid and the two methylgroups mentioned before.

AA3: After OH-Val1 (AA2) an amino acid with an aromatic spin system follows in the peptide sequence. This amino acid exhibits five aromatic protons showing a characteristic coupling pattern of a phenyl ring. These protons are in the neighborhood of a CH₂ group followed by an α -proton. Hence it could be identified as a phenylalanine.

AA4: Similar shifts to the ones of AA2 appear in the next amino acid, it is another β -oxidized valine (OH-Val2).

AA5: The α -proton of the next spin system shows a COSY correlation to a CH_2 group. This carbon has a chemical shift of 62.0 ppm which originates from a connected hydroxyl group. Thus this amino acid is a serine.

AA6: The subsequent spin system in the COSY spectrum contains two methyl groups and three CH groups in a row. The first CH carries the two methyl groups to build an isopropyl unit. This isopropyl group is linked to a CH which is further connected to nitrogen atom and the third CH group. The latter is neighbored to a hydroxyl group, and a carbonyl verified by HMBC correlations (Figure 48). That means this amino acid is 3-amino-2-hydroxy-4-methylpentanoic acid (2OH- β -Val).

AA7: The last amino acid in the structure of c800 (**43**) shows further characteristic shifts of a serine like AA5.

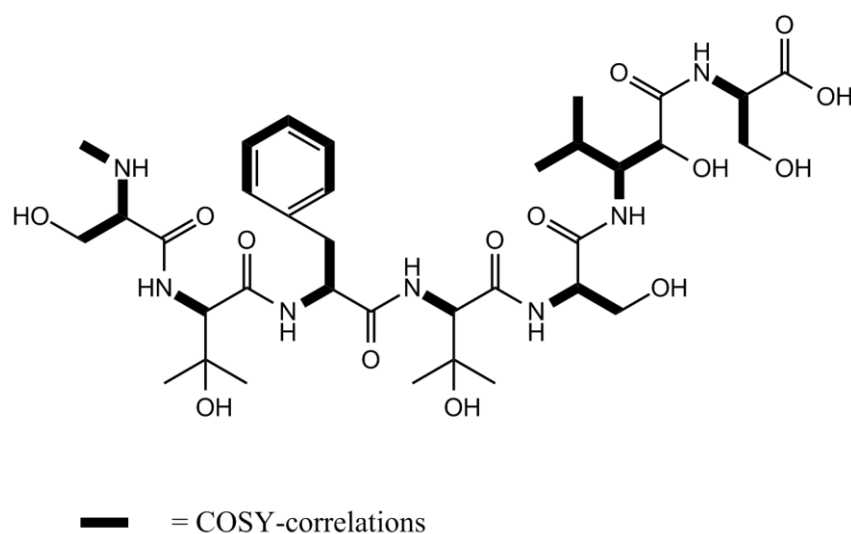


Figure 47: COSY correlations in c800.

The amino acid sequence of the peptide was determined by correlations from the carbonyl groups to the nitrogen protons and the α -protons of the amino acids (HMBC correlations, depicted in Figure 48). Beside that NOESY correlations were used, too. The outcome of the above described NMR analysis was the structure of c800 (**43**) as depicted in Figure 47 to Figure 49.

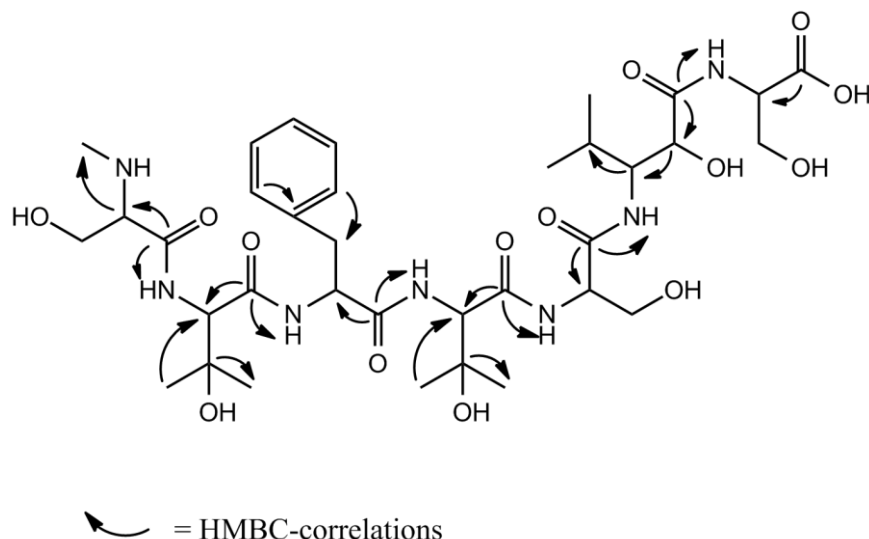


Figure 48: HMBC correlations of myxoprincomide c800; only important ones are shown.

Myxoprincomides c382 (**40**) and c683 (**41**) are smaller fragments of c800 (**43**). Therefore the chemical shifts are similar and could be compared to c800 (**43**). c382 (**40**) consists out of the first three amino acids *N*MeSer, OH-Val1 and Phe (AA1 to AA3). c683 (**41**) is a hexapeptide, it consist of the same first five amino acids (AA1 – AA5) which are part of c800 (**43**) while the last amino acid is a normal valine instead of 2OH- β -Val. The valine is reflected by two methyl groups which show a COSY correlation to a CH group that is connected to the α -CH of the amino acid.

Myxoprincomide c798 (**42**) has two protons less than c800 (**43**). The reason for that is a diketo group in (AA6). Here the 2-hydroxy group of 2OH- β -Val is oxidized to a keto group, resulting in 2-oxo- β -Leu.

Beside those four myxoprincomides five larger ones were isolated: c812 (**44**), c814 (**45**), c581 (**46**), c582 (**47**), c587 (**48**) and c588 (**49**). In c812 (**44**) and c814 (**45**) the pentanoic amino acid (AA6) is extended by one CH₂ unit to result in 3-amino-2-hydroxy-5-methylhexanoic amino acid and 3-amino-5-methyl-2-oxohexanoic amino acid respectively.

c581 (**46**), c582 (**47**), c587 (**48**) and c588 (**49**) are decapeptides. The amino acid sequence of c800 (**43**) continues with a tyrosine (AA8). Here the aromatic spin system shows a para disubstituted phenyl ring connected to the peptide backbone. Tyrosine is followed by a β -lysine (AA9) which is identified by the analysis of a spin system containing eight different proton signals bound to five different carbons. The final spin system (AA10) shows a methyl group neighbored to the α -proton of an amino acid. The chemical shifts are characteristic for an alanine.

The differences between c581 (**46**) and c582 (**47**) are again the variation of the oxidized stage in the pentanoic acid (AA6; 2OH- β -Val vs. 2-oxo- β -Leu).

While c587 (**48**) and c588 (**49**) consist out of a hexanoic acid at AA6, instead of a pentanoic acid. The difference is once again the oxidation stage of AA6, although c588 (**49**) could not be isolated in sufficient amounts for complete NMR analysis, but its mass indicates the keto hydroxyl form of **48**.

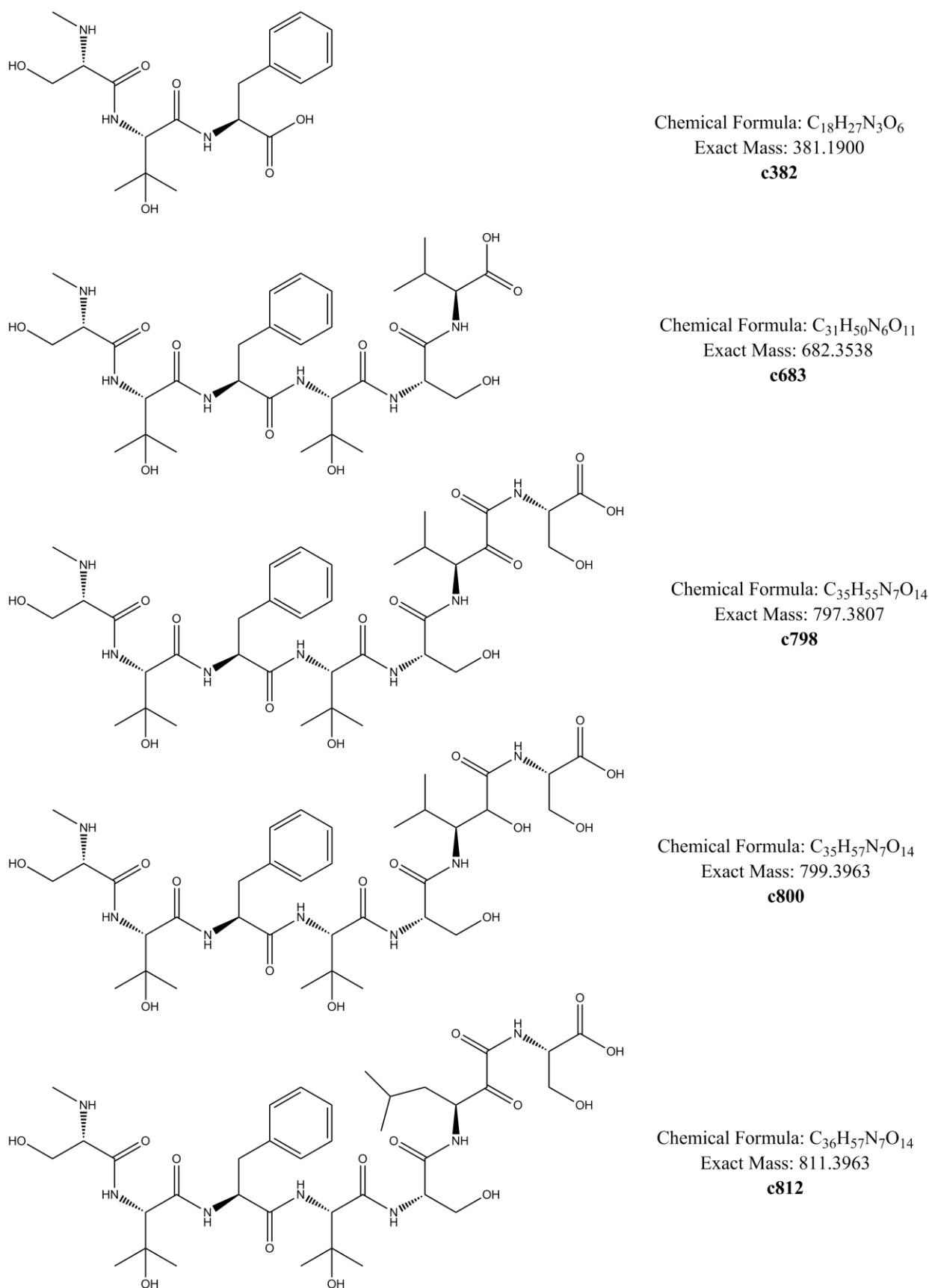


Figure 49: Structure of the myxoprincomides in DK897: 40 = c382, 41 = c683, 42 = c798, 43 = c800, 44 = c812

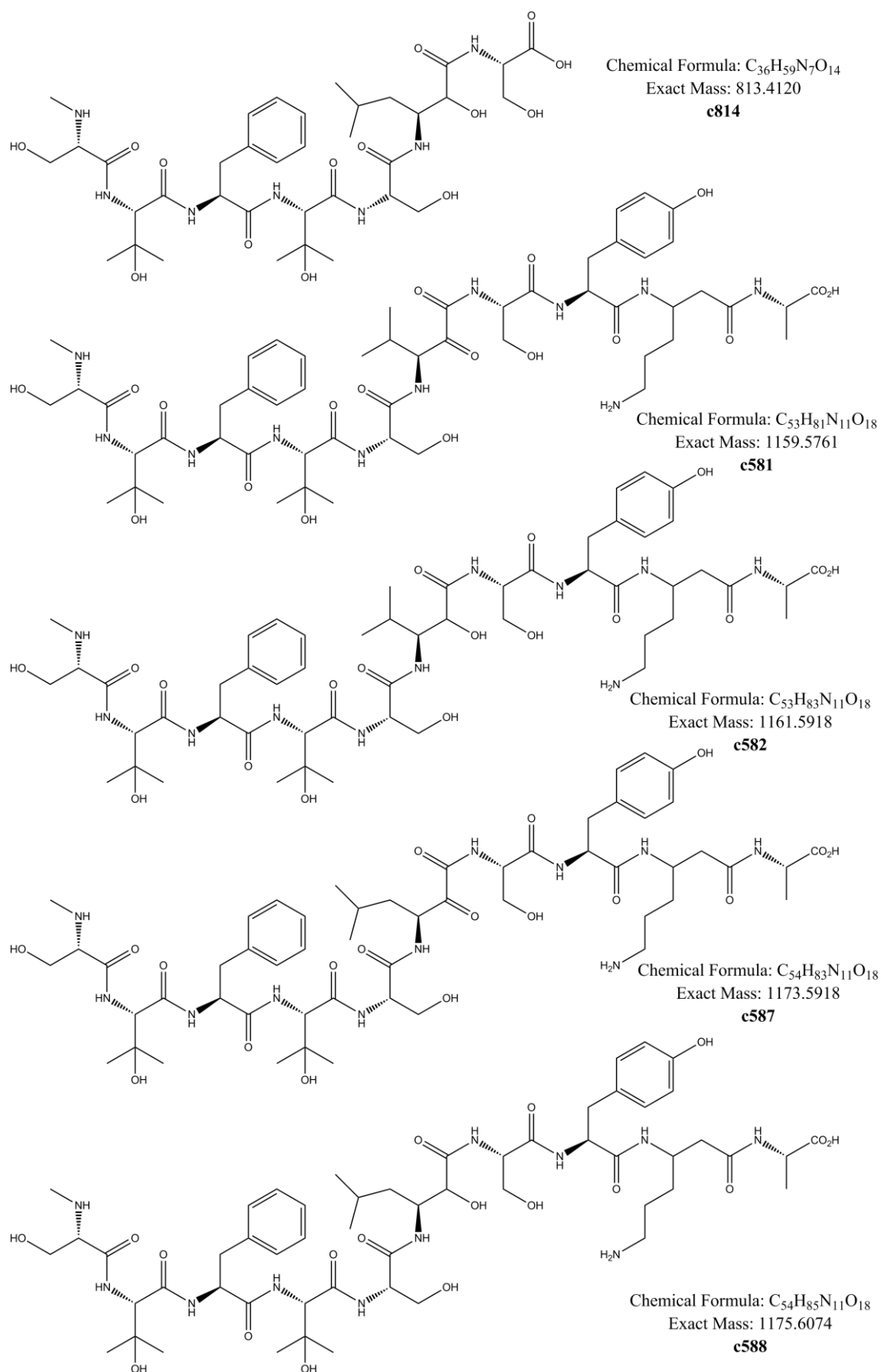


Figure 50: Structures of the myxoprincomides from DK897: 45 = c814, 46 = c581, 47 = c582, 48 = c587, 49 = c588

	δ_C	δ_H (J in Hz)	HMBC
NMeSer			
1	169.3		
2	64.4	3.53 m	1, 3
3	61.1	3.62 m	1, 2, NMe
		3.55 m	1, 2
NH		*	
NMe	33.1	2.31 s	2
OH-Val1			
1	170.2		
2	55.9	4.28 m	1, 3, 4, 5, 1NMe
3	71.8		
4	26.0	1.05 s	3, 5
5	25.9	0.98 s	3, 4
NH		8.47 m	2
Phe			
1	170.4		
2	54.8	4.58 m	1, 3, 1'
3	37.7	3.01 m	1, 2, 1', 2', 6'
		2.75 m	2, 1', 2', 6'
1'	138.0		
2', 6'	128.8	7.17 m	3, 1', 3', 4', 5'
3', 5'	129.7	7.17 m	1', 2', 4', 6'
4'	127.0	7.11 m	1', 3', 5'
NH		8.42 d (8.0)	1OH-Val1
OH-Val2			
1	171.9		
2	61.0	4.31 m	1, 3, 4, 5, 1Phe
3	72.0		
4	27.6	1.09 s	2, 3, 5
5	27.9	1.00 s	2, 3, 4
NH		8.09 d (8.0)	1
Ser1			
1	170.3		
2	55.9	4.27 m	3
3	62.0	3.49 m	1
		3.49 m	1
NH		8.20 d (5.9)	1
2OH-β-Val			
1	172.9		
2	70.9	4.09 m	1, 3, 4
3	57.5	3.79 m	4, 6, 1Ser1
4	29.7	1.75 m	3, 6
5	20.2	0.84 d (6.4)	3, 4, 6
6	19.9	0.76 d (6.4)	3, 4, 5,
NH		7.68 d (6.3)	1Ser
Ser2			
1	174.4		
2	56.9	3.86 m	1, 3
3	62.8	3.58 m	1, 2
		3.58 m	1, 2
NH		7.78 d	1, 1 2OH- β -Val

Table 6: NMR data of myxoprincomide c800 measured at 500 MHz in DMSO (* = not assigned).

By NMR experiments the structure was determined, but not the absolute configuration of the peptide. The absolute configuration of amino acids can be determined by different methods. One commonly used method is the Mosher ester technique; it identifies the absolute configuration of unknown stereogenic center of secondary alcohols or amines by NMR. ^[214;215] Mosher's acid (α -methoxy- α -trifluoromethylphenylacetic acid (MTPA)) reacts with an alcohol or amine to form an ester or amide. Here one uses the fact that the protons display different chemical shifts depending on an *R* or *S* configuration.

An alternative would be NMR measurements using the coupling constants and using NOESY or ROESY data. But even here one stereo center has to be known before. It is possible to determine the configuration by the combined analysis of homonuclear (H-H) and heteronuclear (C-H) $^{2,3}J$ couplings and NOEs. ^[216] But as the observed myxoprincomide spectra were already complicated due to signal overlapping it would be hard and probably not possible to assign the configuration for all stereo centers. Therefore NMR structure elucidation was only used for the identification of the amino acids and for determining the amino acid sequence. But the stereochemistry of the amino acids remained to be solved after the NMR analysis. To determine if the amino acids had L- or D- configuration the myxoprincomides were hydrolyzed and amino acids were assigned by advanced Marfey's method. ^[217-221] The advanced Marfeys reagent (D/L-FDLA) reacts with the amino acid and depending on its configuration it results in a LL or LD configuration and DL or DD, respectively (Figure 51). The diastereomers LL and RR elute in HPLC analytic at one retention time, same for DL and LD. The amino acids are derivatized once with D-FDLA and once with L-FDLA. With the knowledge of the configuration from the Marfeys reagent it possible to assign the amino acid configuration (Table 7). All amino acids in the myxoprincomides are L-amino acids.

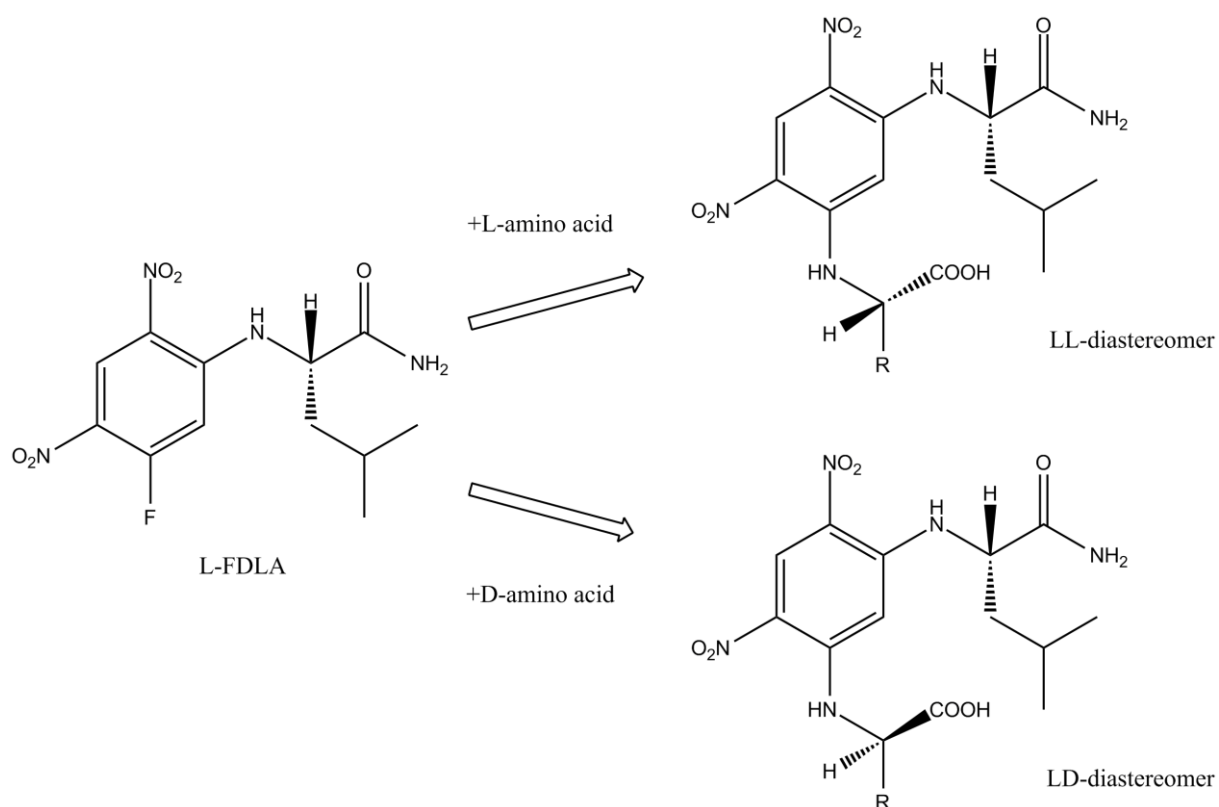


Figure 51: Marfeys reagent reacts with the amino acid to the LL- of LD-diastereomer.

Component	Mass of derivatized amino acid	Retention time of L-FDLA + L-aa	Retention time of D-FDLA + L-aa
L-Alanine	384.2	11.7	12.5
L-Tyrosine	476.3	11.9	12.4
L-Valine	412.3	12.4	13.7
L-Phenylalanine	460.3	13.1	14.2
Met-Ser	414.2	10.5	10.7
OH-Val	428.2	11.1	12.2
2OH-β-Val	442.2	12.1	12.9
2OH-β-Leu	456.2	12.2	13.9

Table 7: Retention times of the amino acids in the myxoprincomides with the Marfeys reagent.

Biosynthesis of the myxoprincomides

Knowing the structure of the myxoprincomides, the biosynthesis can be further analyzed and inactive modules can be assigned. The complete biosynthesis is schematically shown in Figure 53.

Based on the structure the biosynthesis begins with module 1 instead of starting with the loading module. A reason for it might be as described before, that the C domain in module 1 has deviations in its core motifs possibly rendering it inactive. The metabolite starts with L-serine which is methylated by the N-methyltransferase of module 1. In the next step the amino acid is condensed with a β -hydroxylated L-valine (module 2). Alternatively it is condensed with a valine which is hydroxylated later on during biosynthesis. An incorporation of L-phenylalanine and a chain extension by another β -hydroxylated L-valine or rather valine with later hydroxylation follows (modules 3 and 4). The prediction for the A domains in module 2 and 4 did expect ornithine and hydroxyl-phenyl-glycine and not valine or rather the uncommon β -hydroxylated L-valine. Module 5 extends the intermediate by an L-serine as predicted. The following three modules are special. For c683 (41) module 6 connects the intermediate with a valine. But for the longer metabolites at this position there is a hexanoic or a pentanoic amino acid. If the pentanoic or rather the hexanoic amino acids are directly attached by module 6 then module 7 and 8 would be inactive. This is very unlikely because the substrate specificity of the A domain of module 6 would be uncommonly wide it would be able to accept at least five different amino acids. From the genetic point of view the PKS module (module 7) seems to be active and the NRPS module 8 exhibits an oxidation domain within the A domain and no PCP. Such embedded oxidation domains within A domains are described for example for the biosynthesis of myxothiazol in MtaD^[119] and for epothilone in EposP^[222;223]. The Ox domain are located between the A8 and A9 motifs of an A domain.^[224] A suggestion for the involvement of module 7 and 8 is made for the myxoprincomide biosynthesis of DK1622 (see chapter 2B).^[176] In this proposal both modules are active. The PKS module (module 7) adds a malonyl-CoA to the intermediate (refer to Figure 52). And the oxidation domain of module 8 possibly performs an alpha oxidation of the acetyl group. The obtained functional groups tautomerise and are attacked by the serine introduced by module 9. During this process a CO is lost. Parts of the intermediates are further oxidized (by module 9) to build a diketo group; others remain as 1-keto-2-hydroxy group. This is a very unusual step and not fully understood. But feeding experiments of different ¹³C-labeled acetates in DK1622 supported the above described hypothesis because of an incorporation of one ¹³C-

labeled carbon (see chapter 2B).^[176] The mechanism will be further addressed in chapter 2B^[176].

After the unusual building block the assembly continues according to textbook NRPS biosynthesis. The amino acids: L-serine, L-tyrosine, beta-lysine, and L-alanine are added by module 9 to 12. The specificity for serine, tyrosine and beta-lysine was predicted before.

Module 13 can be assumed to be inactive as the PCP domain has some proline residues in the sequence that are not present in other PCP domains of the assembly line. These proline residues could have an effect on the proper folding of the domain and leads to inactivity. From the last module only the TE domain is needed. It is in charge of releasing the final product. For the release of the intermediates a bail out of the intermediates directly from the assembly line is imaginable. For such event normally a type II TE is needed. Typically these trans-acting TEs remove aberrant intermediates on biosynthetic assembly lines.^[225] Here such a TE was not identified so far, therefore it might be possible that there is a product release with unknown mechanism or a hydrolytical spontaneous release.^[225]

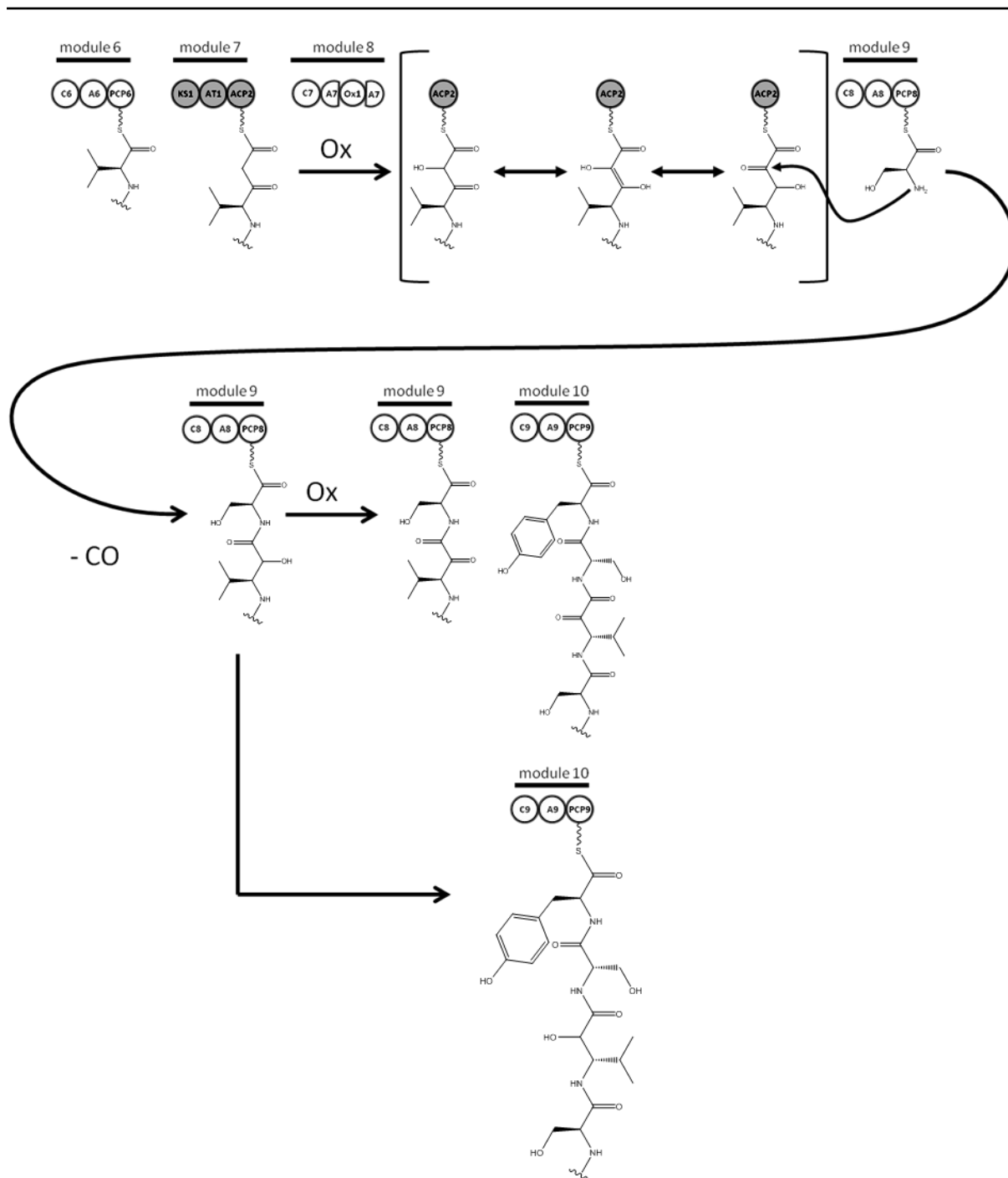


Figure 52: Proposed mechanism for the α -hydroxy moiety formation or rather α -keto intermediate formation (only modules 6 to 10 are shown).

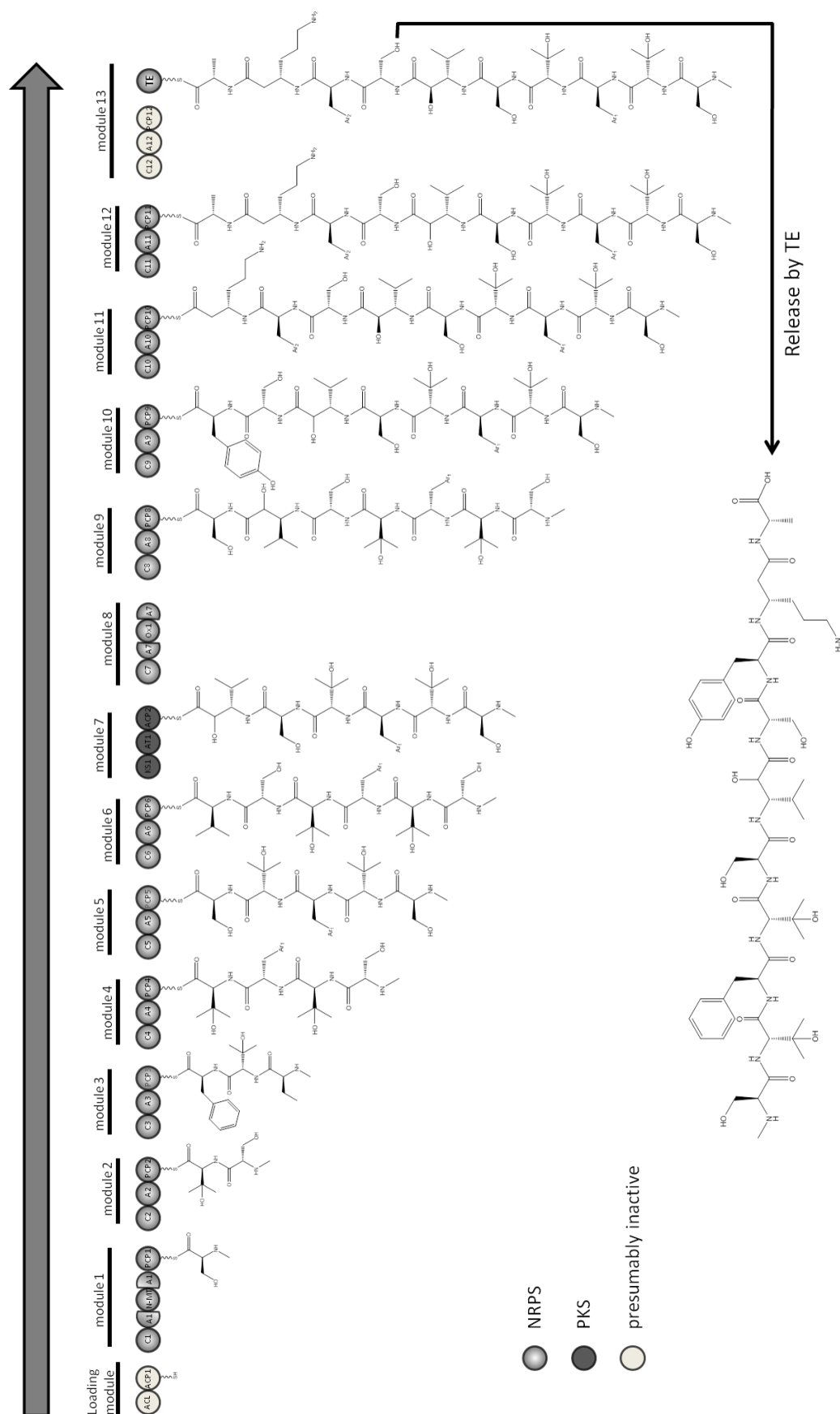


Figure 53: Scheme for the biosynthesis of the myxoprincomides in *M. xanthus* DK897 exemplified with the final structure of c581.

Comparison of the *mvp* biosynthetic pathways from DK897 and DK1622

The comparison of the two different myxoprincomide structures of DK897 and DK1622 (Figure 54) already suggests a biosynthetic relationship. 7 amino acids out of 9 are identical and in the same amino acid order. A view at the two biosynthetic pathways (Figure 53 and Figure 55) reveals that the two *mvp* clusters of the *M. xanthus* strains are very close relatives. As the domain order in the clusters is similar, apart from one additional module. Probably one module was inserted or lost during an evolutionary process.

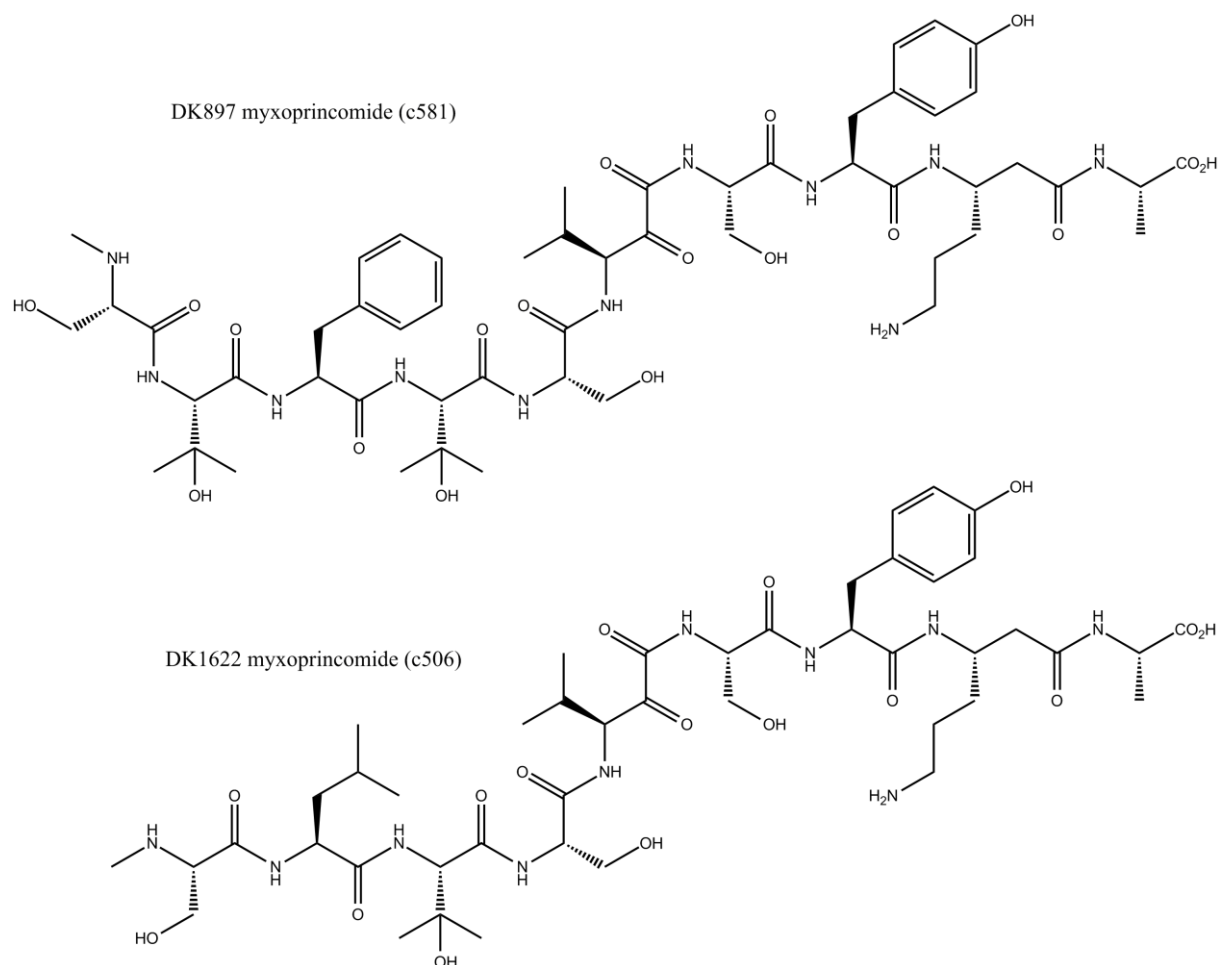


Figure 54: Structures of myxoprincomides of DK897 (c582) and of DK1622 (c506).

That the A domains in both biosynthetic gene cluster are very similar to each other can be shown by a phylogenetic tree (Figure 56). Each A domain has exactly one very close relative from the other strain, except the A domain A2 of *mvp*_{DK897}.

The A domain specificities within the cluster of DK1622^[176] and DK897 were compared (Table 8). Domains A2 (extra), A3 and A6 of DK897 have some differences to the corresponding A domain of DK1622 but a prediction by the residues was not possible.

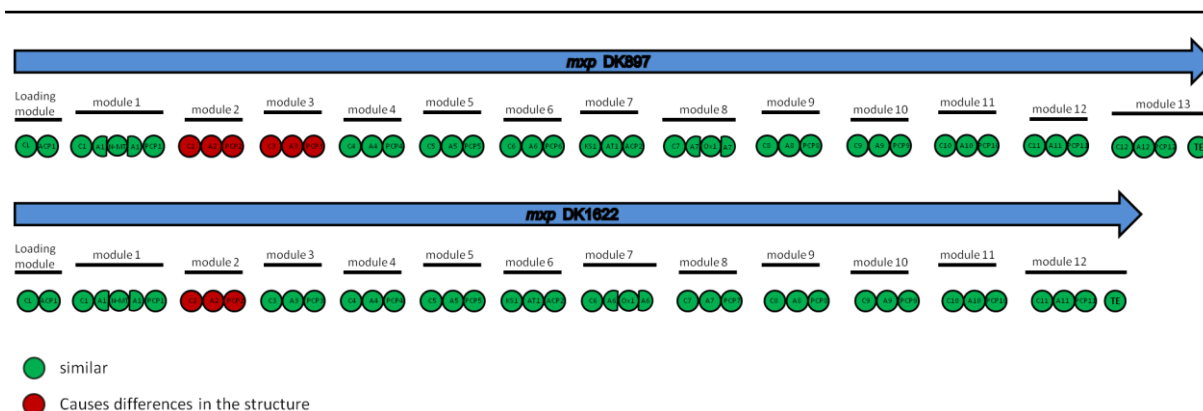


Figure 55: Comparison of the myxoprincomide clusters of DK897 and DK1622.

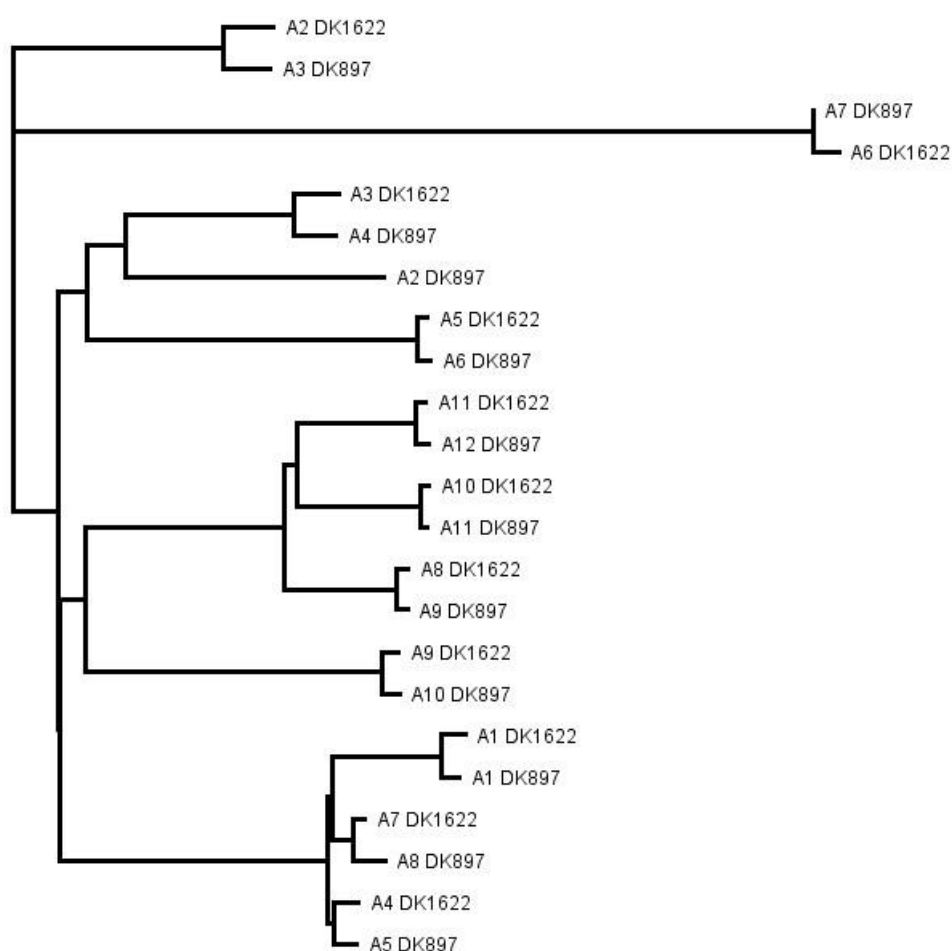


Figure 56: Phylogenetic tree of A-domains from myxoprincomide cluster of DK1622 and DK897

Furthermore every domain was aligned to its possible analogue of the other strain (Table 9). The identity is always above 87 % except for PCP3 (47.9%), C4 (82.4%), C5 (80.5%) and the domains of module 2 of DK897 which do not have an analogue in DK1622. That means here are the major differences on protein level between the two clusters.

To conclude, the gene cluster of DK897 exhibits one additional module (module 2) which incorporates an OH-Val or rather a valine with subsequent oxidation. While module 3 employs a phenylalanine instead of a leucine. Hence these two modules are in charge of the structural differences of the myxoprincomide of DK897 and DK1622.

A-Domain	Analysis with web tools of J. Ravel et al. ^[211]		Analysis with web tools of T. Weber et al. ^[212]	
	residues	pred.	residues	pred.
A1 DK1622	DVWHVSLV	ser	DVWHVSLVD-	ser
A1 DK897	DVWHVSLV	ser	DVWHVSLVD-	ser
A2 DK1622	DVVLMGAT	no	DVVLMGATMK	no
A2 DK897	DVGELVSI	orn	DVGEIGSIDK	orn
A2 DK1622	DVVLMGAT	no	DVVLMGATMK	no
A3 DK897	DVILLGGT	no	DVIVLGGTMK	no
A3 DK1622	DALWLGLG	hpg	DALWLGGTFK	val
A4 DK897	DALWLGLG	hpg	DALWLGGTFK	val
A4 DK1622	DVWHVSLV	ser	DVWHVSLVDK	ser
A5 DK897	DVWHVSLV	ser	DVWHVSLVDK	ser
A5 DK1622	DGMHNVGI	no	DGMHNVGIK	no
A6 DK897	DGMHHVGI	no	DGMHHVGIK	no
A6 DK1622	Mono-oxygenase	no	Mono-oxygenase	no
A7 DK897	Mono-oxygenase	no	Mono-oxygenase	no
A7 DK1622	DVWHVSLV	ser	DVWHVSLVDK	ser
A8 DK897	DVWHVSLV	ser	DVWHVSLVDK	ser
A8 DK1622	DG-DVAAV	no	DGEDVAAVCK	tyr
A9 DK897	DG-DVAAV	no	DGEDVAAVCK	tyr
A9 DK1622	DTENVGTA	ahp	DTENVGTAVK	lys-b
A10 DK897	DTENVGTA	ahp	DTENVGTAVK	lys-b
A10 DK1622	DG-DFGVV	no	DGEDFGVVCK	ile
A11 DK897	DG-DFGVV	no	DGEDFGVVC-	no
A11 DK1622	DV-DLAAV	no	DVDDLAAVYK	val
A12 DK897	DV-DLAAV	no	DVDDLAAVYK	val

Table 8: Analysis of the A domains of the DK897 cluster and *gmxf2004* of DK1622; differences in the residues are marked in bold

DK897	Identity [%]	DK1622	DK897	Identity [%]	DK1622
ACL	98.0	ACL	KS1	98.4	KS1
ACP1	100.0	ACP1	AT1	97.3	AT1
C1	97.7	C1	ACP2	96.9	ACP2
A1	94.9	A1	C7	91.9	C6
M	98.0	M	A7	97.8	A6
PCP1	100.0	PCP1	C8	87.6	C7
C2	90.8	C2	A8	95.3	A7
A2	50.1	A2	PCP8	100.0	PCP7
PCP2	47.9	PCP2	C9	97.7	C8
C3	87.6	C2	A9	97.3	A8
A3	89.1	A2	PCP9	95.4	PCP8
PCP3	47.9	PCP2	C10	98.2	C9
C4	82.4	C3	A10	96.3	A9
A4	90.0	A3	PCP10	100.0	PCP9
PCP4	92.3	PCP3	C11	98.2	C10
C5	80.5	C4	A11	98.1	A10
A5	94.5	A4	PCP11	97.0	PCP10
PCP5	95.4	PCP4	C12	96.6	C11
C6	91.7	C5	A12	97.1	A11
A6	97.3	A5	PCP12	100	PCP11
PCP6	98.5	PCP5	TE	97.2	TE

Table 9: Identity between the domains of *mxp*_{DK897} and *mxp*_{DK1622}.

Conclusion

To conclude, novel myxoprincomides¹ have been discovered in *M. xanthus* DK897. Structurally the myxoprincomides are interesting due to numerous striking features like unusual amino acids and other rare structural elements. Up to now, ten compounds, which were correlated to the DK897 myxoprincomide gene cluster, were detected: c382 (**40**), c683 (**41**), c800 (**43**), c812 (**44**), c814 (**45**), c581 (1160) (**46**), c582 (1162) (**47**), c587 (1174) (**48**) and c588 (1176) (**49**). The last four are probably products of the complete assembly line, while the others are possibly chunk products, released early from the assembly line. It is likely that further intermediates of the myxoprincomide family, like a tetra, penta, octa or nona peptides, are produced but not detected up to date. However it is a fascinating example of an NRPS-PKS assembly line, which releases several intermediates. With its 39 catalytic domains *mvp₈₉₇* is the largest myxobacterial biosynthetic gene identified so far. Furthermore it is particularly exciting due to its analog in *M. xanthus* DK1622. The analysis of 98 *M. xanthus* strains revealed that either the myxoprincomide type of DK897 or the variants of DK1622 are built, but never both.^[51] That means, either the strains possess the DK897 gene cluster or the DK1622 one. Therefore the myxoprincomides must exhibit a fundamental role for the *M. xanthus* strains.

The characterization of the biosynthetic machinery revealed, that it does not follow the biosynthetic textbook model. Several deviations of *in silico* predictions were identified. It shows once again that textbook biochemistry cannot always be applied for myxobacteria.

Methods

Cultivation:

Myxococcus xanthus DK897 and mutant strains were grown in CTT medium (1 l water, 10 g Casitone (Difco), 10 ml 1 M Tris-HCl pH = 8.0, 10 ml of 0.8 M MgSO₄, 1 ml K₂HPO₄ pH = 7.6, final pH adjusted to 7.6). For mutant strains kanamycin sulfate was added (50 µg / ml). Cultures were grown in 50 ml CTT medium containing 2 % amberlite XAD-16 resin (Sigma-Aldrich). Cultures were grown for 72 h at 180 rpm, 30 °C in 300 ml Erlenmeyer flasks. Cultures were harvested by centrifugation (10 min, 8000 rpm). The XAD / cell pellet was

¹ The name myxoprincomide is derived from myxobacteria (myxo-), its discovery by principal (-prin-) component (-co-) analysis and the amide (-mide) bond.

extracted with methanol (2 x 50 ml), dried in vacuo. For HPLC-MS (and HR-MS) analysis it was resuspended in 1 ml methanol.

MS analysis:

HPLC-HR-MS: Accela UPLC system coupled to a TriVersa NanoMate (Advion) ion source and a LTQ Orbitrap (Thermo Scientific). Separation: RP-18 column (50 mm x 2.1 mm, particle size 1.7 μ m; Waters Acquity BEH), mobile phase was water (A) and acetonitrile B (each plus 0.1 % formic acid), flow rate of 400 μ l / min, gradient: 0 min 1 % B; 1 min 1 % B; 10 min 99 % B; 12 min 99 % B; 12.5 min 1 % B; 14 min 1 % B.

HPLC-MS: Positive / negative mode, m/z = 100-1100, auto MSn, runtime 24 min, column Luna C18 2.5 μ m, T = 32 $^{\circ}$ C, flow rate = 0.400 ml/min, solvent A: water with 0.1 % formic acid, B: acetonitrile with 0.1 % formic acid. The different gradients are depicted in Table 10 to Table 12.

Time [min]	Solvent A [%]	Solvent B [%]
0	95	5
1	95	5
16	5	95
18	5	95
19	95	5
24	95	5

Table 10: Gradient of 1STDKURZ

Time [min]	Solvent A [%]	Solvent B [%]
0	95	5
2	95	5
22	5	95
25	5	95
27	95	5
32	95	5

Table 11: Gradient of 1Std_neu

Time [min]	Solvent A [%]	Solvent B [%]
0	95	5
2	95	5
32	5	95
36	5	95
39	95	5
44	95	5

Table 12: Gradient of 1STDLANG

Gene inactivation:

A fragment of ~550 bp was amplified using Taq polymerase (Fermentas). The resulting PCR product was cloned into pCR2.1-TOPO (Invitrogen). *E. coli* (DH10b) cells were used for cloning. The colonies were selected on Luria–Bertani (LB)-agar containing kanamycin sulfate (50 µg / ml), single colonies were cultivated in LB medium (1.5 ml, with Kan 50 µg / ml). From these cultures plasmid was isolated using the Nucleospin plasmid kit (Macherey & Nagel). The obtained plasmid was transformed into *M. xanthus* DK897 (650 V, 400 Ω, 25 µF, creation of *M. xanthus* mutants described previously ^[195-198]).

Transposon-mutagenesis and transposon recovery:

The plasmid pMycoMar ^[226] was used for transposon mutagenesis. It contains the mini - transposon *magellan4*, kanamycin resistance and the oriR6K. Because of the TA dinucleotide-recognition sequence it is able to integrate randomly into several host genomes. 45000 clones were obtained, 2688 of them were cultivated (each clone in 250 µl CTT-media containing kanamycin, in microtiter plates at 30 °C for 3 days). For HPLC-MS analysis the clones were cultivated under the same conditions together with XAD resin for 72 h. The cultures were centrifuged and the precipitate was extracted with methanol. The obtained extract was then analyzed by HPLC-MS. The extracts were screen for the lack of c382, c800, c814. In mutant 5D9 these compounds were missing. For transposon recovery the chromosomal DNA was isolated (Puregene™ DNA Purification Kit), digested with *Sma*I, re-ligated, transformed in *E. coli* DH5α-λpir. Plasmids were isolated and sequenced with primers (K388 and K390 for the kanamycin resistance gene). For obtaining further sequence information new primers were subsequently designed to obtain the complete sequences information of the plasmid. For obtaining more sequence information a TOPO plasmid (Invitrogen) was integrated at the end

of the known sequence of the myxoprincomide cluster in DK897. DNA was isolated, digested with *HindIII*, relegated, and transformed into *E. coli* DH5 α - λ pir, followed by plasmid isolation and again subsequent sequencing of the plasmid. By this method 10460 bp of sequence information were gathered. This information was later on used when contigs of the strain were available, to identify contigs which were part of the gene cluster.

Contigs:

A PCR at 65 °C with Phusion polymerase (Finnzymes) was performed between the gaps of the contigs. Primers:

1012_e_1for = CTGGGGGCTGAGGACTTCGTG

0024_a_1rev = GACGCGAGGCGCTGGGAGCAC

1012_a_1rev = GTAGCCGTTACACAGCCCCAC

0032_e_1for = CAACTCGCCGGGCATGCCGAG

0032_a_1rev = GTCTTCGCGAACCACGGCGAC

0982_e_1for = CTGGACTTGCTGGTGGGCCTG

0982_a_1rev = GTTGACGAAGAAGCCAATGAG

0024_e_1for = GAGGTGACCATCCACGCGCTG

1009_e_1for = GAACGAGACGTCGAAGATC

1009_a_1rev = GTCCCGCGGTCAGTAGCGCCAAG.

The resulting amplicates were cloned into the pJet vector (Fermentas), electroporated into *E.coli* (DH10b) and the plasmids were sequenced. The missing gap could be closed between 1009 and 1012 (324 bp), between 0982 and 0024 (639 bp), between 1009 and 0982 (236 bp), for the fourth gap there was no correct clone and it has to be repeated.

Scaffolds:

High quality genomic DNA was prepared (because of several PCR problems) and primers were designed to amplify these regions:

scaf37_1f = CGTTCGAGTCAGAGCGGCTGA

scaf37_2f = GAGCTCATCGAAGCGCAAGTC

scaf37_3f = TGACCACCCTGCCGTTGATG

scaf24_1r = CATCAACGGCAGGGTGGTCA

scaf24_2r = CACGTCTCCGGTGCGGTACAG

scaf24_3r = CACCCAACTCGATGCGGAAGC

A2_1f = ATCACGCCGCATCAGCGCGGT

A2_2f = GCTGACACCACCGTCCATCA
A2_3f = CTGGATGCGAGGGCCAACCA
A2_1r = TGCTGCATCTGCTCGAAGAGT
A2_2r = CGATGGACACGGAGTCACCCAC
A2_3r = AGGCCCCGAAGGGATTGGGCAC
A4_1f = TGCACCGGGGCCTGGACTTGCTG
A4_2f = CTCAGTGCGCTGCCCATGATG
A4_1r = TGAGGTGCGAACC GGCGGGCAG
A4_2r = CACGAAGCGAGAGCCGATGAG
A?_1f = CTCGGAGCGCATCTCCCGGTTG
A?_2f = CTGGAGTACGCCACCGAGCTG
A?_1r = GACACCCCAGAGGACGGCCAG
A?_2r = CTCCAGCGCGTCCTCCACATGG
A?2_1f = CTGGGAACGGAGGTCCTGCTC
A?2_2f = CTCATCGACTTCGGCCTGGACAC
A?2_1r = GGAAGTTGGAGCCGTTGTTGTAC
A?2_2r = GACCTCGTGCCGCTTCACCAC
A6_1f = CTGCAACGTGTCGCCATCCTCTC
A6_2f = CTCCAAGTTCGACCTGAGCCTG
A6_3f = CACTTCGGGTCAGGAGGACTTCG
A6_1r = GTGCGCGGAGCGACGTAGAG
A6_2r = TCGCGTCTCCTCACGCACCAC
A6_3r = GACACCCAGGTGAACGAAGGTC
A9_1f = GGACGCTTTCAGTGGACAGC
A9_2f = GAGCTGTGCATCGCGGGAGAAGG
A9_3f = CCTTCAACGTCTACGGCCCCGAC
A9_1r = GACCGACTGGAAGGCCGCGAG
A9_2r = GTTCACGAAGAAGCCGATGAG
A9_3r = GTCCAGGCTCAGCTCGAACAG
A10_1f = GATGCTGGCGTGGCTGGAGCA
A10_2f = CGTCATCTACACGTCCGGCT
A10_1r = GAGCCGCAGCTCGAGCCCATC
A10_2r = CTTGCGGTCCACCTTTCCAC

The obtained fragments (Supplementary figure 84 and Supplementary) were ligated into the pJet-vector (Fermentas), plasmid DNA was isolated and send for sequencing.

Annotation:

Using the PKS / NRPS analysis software of <http://nrps.igs.umaryland.edu/nrps/index.html> ^[211] annotations of the clusters were performed. For the substrate prediction of A domains the NRPSpredictor software (<http://www-ab.informatik.uni-tuebingen.de/software/NRPSpredictor>) ^[212] was used. The published genome sequence of the *M. xanthus* DK1622 is available under <http://ncbi.nlm.nih.gov>.

Reverse feeding:

The Casitone in the CTT media was exchanged against the enriched ¹⁵N respectively ¹³C powder (ISOGRO). 0.5 ml of DK897 CTT culture were centrifuged, the pellet was washed, and dissolved in 5 ml ¹⁵N / ¹³C media. Each sample was grown in 25 ml flasks, in duplicates as well as the control with CTT media. After 3 days the culture have been harvested and extracted and analyzed as described before.

Feeding of labeled amino acids:

The amino acids were dissolved in water and added in small portions to the culture, every morning and every evening to end with a final concentration of 1 mM in 25 ml. The cultures were extracted and measured in the hct and in the Orbitrap and compared to the control. In the hct only the 580.8 could be detected. In the Orbitrap measurements small traces of 800 and 814 were found, too.

Promoter insertion:

The first 804 bp of the gene cluster were amplified with Phusion (Finnzymes) polymerase using the following primers:

start800_1f = CGCGCTCATATGCACGAGACTCCGCGAAC

stop800_1r = CTCGAATTCGAAGTTGGGTCCACCGCTGATG

58 °C, GC buffer (Finnzymes). The PCR product and the pSBtn5Kan plasmid (Supplementary figure 79) were digested with *Nde*I and *Eco*RI and ligated. The resulting plasmid (pSBtn5Kan-800erstart) was transformed into the strain DK897. The plasmid should integrate directly in front of the cluster which is then under the control of the tn5 promoter

and has a RBS directly before the start codon. After confirmation of the appropriate integration by PCR the strain was cultivated in large scale.

Large scale fermentation:

The mutant DK897_800erstart was cultivated for 144 h in a 300 l fermenter with PS medium (10 g / l Casein Peptone Type S (Marcor), 2 g / l MgSO₄, 80 µg / ml Fe(III)EDTA , and 4 g / l starch) pH = 7.1). After 72 h, 2 % XAD-16 resin was added. At the end XAD and the cells were separated from the liquid medium. The fermentation was performed by W. Kessler.

Purification:

The XAD (2 kg) was transferred into a big glass column, extracted with water (2.5 l), a mixture of water / MeOH 1:1 (2.5 l), MeOH extraction (6 l). The cells (200 g) were extracted twice with MeOH and once with CHCl₃. The different fractions were analyzed for their compounds. Myxoprincomides were obtained in different fractions see Table 13.

Substance	c382	c580	c588	c683	c800	c814
Retention time [min]	6.8	7.9	8.2	8.0	7.7	8.3
H ₂ O	+	-	-	+	+	-
H ₂ O / MeOH	+	+	+	+	+	+
MeOH 1	+	+	+	+	+	+
MeOH 2	-	+	-	+	+	+
MeOH 3	-	+	-	+	-	-
MeOH 4	-	-	-	-	-	-
Cells 1	+	+	+	+	+	+
Cells 2	+	+	+	+	+	+
CHCl ₃	-	-	-	-	-	-

Table 13: Myxoprincomide fractions after elution from XAD.

The solvent was removed; the H₂O / MeOH and MeOH 1 fraction was again diluted in MeOH and extracted against heptane to get rid of fats. MeOH removal and dissolving in water, partition with ethyl acetate, here mayor parts of the DKxanthenes stayed. The obtained fraction had high concentration of myxoprincomides. Next step was methanolic Sephadex[®] LH-20 column. Myxoprincomide fractions were dried and dissolved in a water methanol mixture 1:1 and injected into the Waters Auto Purification System. Preparative LC runs were

performed using an Xbridge Prep C18 OBD column (19 x 150 mm, 5 µm particle size). As mobile phase water with 0.1 % formic acid (A) and methanol with 0.1% formic acid (B) were used. The following gradient was used:

Time [min]	A %	B %
0	73.0	27.0
3.00	73.0	27.0
9.00	50.0	50.0
9.50	5.0	95.0
12.50	5.0	95.0
13.00	73.0	27.0
16.00	73.0	27.0

Table 14: Gradient used at the Waters Auto Purification System.

Detection was performed by a 3100 Mass Detector (Waters) and fractions were collected on a 2767 Sample Manager (Waters) and concentrated afterwards.

Second purification was run at the same system, same gradient using different solvent system A2: H₂O + NH₄CO₂ 50 mM, pH = 9.4, B2: MeOH + H₂O 5% + NH₄CO₂ 50 mM, pH = 9.4.

Third purification was done with acetonitrile as solvent again under acidic conditions pH = 3, gradient:

Time [min]	A %	B %
0	95	5
2	95	5
9.00	65	35
9.50	5.0	95.0
10.50	5.0	95.0
11.00	95	5
14.00	95	5

Table 15: Gradient used for third purification step at the Waters Auto Purification System.

Still impure fractions were further separated via a methanol Sephadex[®] LH-20 column.

NMR:

The pure substances were measured with the microprobe at 500 MHz Bruker Avance NMR spectrometer in DMSO- d_6 .

Marfey:

The myxoprincomides were hydrolyzed by 5 M HCL at 100 °C over night. The solvent was removed, redissolved in 100 μ l water and split into two portions. 30 μ l of NaHCO₃ and 200 μ l of D- or L-FDLA were added. The solution was kept on 37 °C for 1 h. 800 μ l of acetonitrile was added, centrifuged and injected into the HPLC.

Production kinetics:

In 100 ml Erlenmeyer flasks 25 ml of CTT with 0.5 % XAD were inoculated with 50 μ l preculture (washed with CTT media before). At certain time points triplicates were harvested until 99 h after inoculation. The cultures were centrifuged and cell / XAD pellet was separated from the supernatant, samples were frozen at -20 °C. The pellet was extracted with 50 ml MeOH and the supernatant was freeze dried, the residue was resumed in MeOH.

Supplementary

¹³C-NMR was recorded at 125 MHz with DMSO- d_6 ; reference at ppm δ 39.51. ¹H-NMR was recorded at 500 MHz with DMSO- d_6 ; reference at ppm δ 2.50 ppm. Protons assigned in HMBC correlate to indicated carbon. Allocated NOESY protons showing correlation to indicated proton.

	δ_C	δ_H (<i>J</i> in Hz)	HMBC	NOESY
NMeSer				
1	169.2			
2	64.6	3.45 m	1, 3, N-Me	
3	61.1	3.57 dd (6.3, 11.4)	1, 2	
		3.66 dd (4.2, 11.6)	1, 2	
NH		*		
NMe	33.3	2.33 s	2	
OH-Val1				
1	169.7			
2	60.9	4.21 m	3, 4, 5, 1	
3	71.8			
4	27.6	1.05 s	2, 3, 5	
5	25.7	1.03 s	2, 3, 4	
NH		8.27 d (9.7)		
Phe				
1	174.6			
2	56.1	4.23 m	1, 3, 1'	NH
3	37.6	3.08 dd (4.3, 13.9)	1, 2, 1', 2', 6'	3
		2.82 dd (8.6, 13.5)	1, 2, 1', 2', 6'	3
1'	139.1			
2', 6'	129.7	7.17 m	3, 1', 3', 4', 5'	
3', 5'	128.4	7.18 m	1', 2', 4', 6'	
4'	126.4	7.13 m	1', 3', 5'	
NH		7.64 d (7.8)		2OH-Val

* not assigned

Supplementary table 1: NMR Spectroscopic data for c382 (DMSO-d6)

	δ_C	δ_H (J in Hz)	HMBC
NMeSer			
1	171.2		
2	66.4	3.04 m	1, 3, N-Me
3	61.9	3.54 m	1, 2
		3.40 dd (7.3, 10.7)	1, 2
NH		*	-
NMe	34.4	2.15 s	2
OH-Val1			
1	167.7		
2	59.6	4.27 m	1
3	71.7		
4	27.7	1.02 s	2, 3, 5
5	25.3	0.97 s	2, 3, 4
NH		7.34 m	
Phe			
1	171.2		
2	54.3	4.59 m	
3	37.3	3.05 m	2, 1', 2', 6'
		2.77 dd (10.4, 13.8)	1, 2, 1', 2', 6'
1'	138.0		
2', 6'	129.3	7.21 m	3, 1', 3', 4', 5'
3', 5'	128.2	7.18 m	1', 4',
4'	126.4	7.11 m	2', 6'
NH		8.55 m	1OH-Val2
OH-Val2			
1	170.1		
2	60.5	4.40 d (9.0)	1, 3, 4, 5, 1Phe
3	71.9		
4	28.2	1.10 s	2, 3, 5
5	25.0	1.10 s	3
NH		8.10 m	1Phe
Ser1			
1	167.7		
2	55.9	4.27 m	3
3	61.9	3.65 dd (5.4, 10.7)	1, 2
		3.54 m	1, 2
NH		8.06 m	

Supplementary table 2: NMR Spectroscopic data for c683 (DMSO- d_6); * not assigned

	δ_C	δ_H (J in Hz)	HMBC
Val			
1	174.9		
2	59.9	3.87 dd (4.9, 8.6)	1, 3, 4, 5, 1Ser
3	30.8	1.95 m	1, 2, 4, 5
4	19.8	0.76 t (7.2)	2, 5
5	18.3	0.76 t (7.2)	3, 4
NH		7.52 m	2, 1Ser

Supplementary table 2: continued

	δ_C	δ_H (J in Hz)	HMBC	NOESY
NMeSer				
1	170.7			
2	66.0	3.12 m	1, 3, NMe	1, 3, NMe
3	61.7	3.54 m	1, 2	2, 3
		3.43 dd (6.8, 10.5)	1, 2	2, 3
NH		*		
NMe	34.2	2.19 s	2	2
OH-Val1				
1	170.2			
2	60.4	4.46 m	1, 3, 5	
3	72.3			
4	28.0	1.12 s	2, 3, 5	2
5	24.7	1.06 s	2, 3	5
NH		8.07 m	1Phe	
Phe				
1	171.3			
2	54.2	4.59 m	1, 1AS,	3
3	37.4	3.05 m	1, 2, 1', 2', 6'	2, 3
		2.77 dd (10.3, 14.0)	1, 2, 1', 2', 6'	2, 3
1'	138.0			
2', 6'	129.3	7.20 m	3, 3', 5'	
3', 5'	128.3	7.18 m	1', 2', 6'	
4'	126.4	7.11 m	2', 6'	
NH		8.49 d (8.5)		

Supplementary table 3: NMR Spectroscopic data for c798 (DMSO- d_6); * not assigned

	δC	δH (J in Hz)	HMBC	NOESY
OH-Val2				
1	*			
2	59.8	4.27 m	3	4, 5, 1Ser1
3	71.8			
4	27.7	1.01 s	2, 3, 5	2
5	25.6	0.97 s	2, 3, 4	2
NH		8.11 m		
Ser1				
1	*			
2	55.7	4.31 m		3
3	63.3	3.67 m		3
		3.59 m	2	3
NH		8.67 d (7.8)	170.2	
Val				
1	*			
2	*			
3	57.7	5.06 m		
4	34.9	2.12 m		
5	19.6	0.83 d (6.8)	3, 6	
6	19.6	0.78 d (6.8)	3, 5	
NH		8.14 m		
Ser2				
1	*			
2	56.6	3.88 m		
3	61.7	3.63 m	2	
		3.54 m		
NH		8.18 m / 8.00 m		

Supplementary table 3: continued

	δ_C	δ_H (J in Hz)	HMBC
NMeSer			
1	169.3		
2	64.4	3.53 m	1, 3
3	61.1	3.62 m	1, 2, NMe
		3.55 m	1, 2
NH		*	
NMe	33.1	2.31 s	2
OH-Val1			
1	170.2		
2	55.9	4.28 m	1, 3, 4, 5, 1NMe
3	71.8		
4	26.0	1.05 s	3, 5
5	25.9	0.98 s	3, 4
NH		8.47 m	2
Phe			
1	170.4		
2	54.8	4.58 m	1, 3, 1'
3	37.7	3.01 m	1, 2, 1', 2', 6'
		2.75 m	2, 1', 2', 6'
1'	138.0		
2', 6'	128.8	7.17 m	3, 1', 3', 4', 5'
3', 5'	129.7	7.17 m	1', 2', 4', 6'
4'	127.0	7.11 m	1', 3', 5'
NH		8.42 d (8.0)	1OH-Val1
OH-Val2			
1	171.9		
2	61.0	4.31 m	1, 3, 4, 5, 1Phe
3	72.0		
4	27.6	1.09 s	2, 3, 5
5	27.9	1.00 s	2, 3, 4
NH		8.09 d (8.0)	1
Ser1			
1	170.3		
2	55.9	4.27 m	3
3	62.0	3.49 m	1
		3.49 m	1
NH		8.20 d (5.9)	1

Supplementary table 4: NMR Spectroscopic data for c800 (DMSO- d_6); * not assigned

	δ_C	δ_H (<i>J</i> in Hz)	HMBC
2OH-β-Val			
1	172.9		
2	70.9	4.09 m	1, 3, 4
3	57.5	3.79 m	4, 6, 1Ser1
4	29.7	1.75 m	3, 6
5	20.2	0.84 d (6.4)	3, 4, 6
6	19.9	0.76 d (6.4)	3, 4, 5,
NH		7.68 d (6.3)	1Ser
Ser2			
1	174.4		
2	56.9	3.86 m	1, 3
3	62.8	3.58 m	1, 2
		3.58 m	1, 2
NH		7.78 d	1, 1 β Val

Supplementary table 4: continued

	δ_C	δ_H (J in Hz)	HMBC
NMeSer			
1	169.4		
2	65.2	3.19 m	
3	61.0	3.61 m	1
		3.47 m	
NH		7.96 m / 8.05 m / 8.15 m / 8.63 m*	
NMe	33.5	2.26 s	2
OH-Val1			
1	169.8		
2	59.8	4.40 d (9.3)	1, 3
3	70.9		
4	27.0	1.06 s	2, 3, 5
5	24.9	1.02 s	2, 3, 4
NH		8.15 m	1/NMe
Phe			
1	170.6		
2	53.7	4.67 m	1, 1OH-Val1
3	36.8	3.07 m	1, 1', 2', 6'
		2.82 m	1, 1', 2', 6'
1'	137.2		
2', 6'	128.8	7.24 m	4, 2', 4', 6'
3', 5'	127.6	7.22 m	1', 3', 5'
4'	125.8	7.15 m	2', 6'
NH		8.20 m	
OH-Val2			
1	169.1		
2	59.4	4.49 d (9.9)	1, 3
3	71.3		
4	27.4	1.11 s	2, 3, 5
5	25.1	1.08 s	2, 3, 4
NH		7.96 m / 8.05 m / 8.15 m / 8.63 m*	
Ser1			
1	170.0		
2	54.8	4.33 d (9.0)	1, 1OH-Val2
3	60.8	3.71 m	
		3.62 m	
NH		7.96 m / 8.05 m / 8.15 m / 8.63 m*	

Supplementary table 5: NMR Spectroscopic data for c812 (DMSO- d_6); * not assigned

	δC	δH (J in Hz)	HMBC
β-Leu			
1	*		
2	*		
3	51.5 / 52.1	5.14 m / 4.10 m	
4	37.0	1.37 m	
		1.23 m	
5	23.4	1.53 m	
6	23.7	0.79 d (6.5)	4, 5, 7
7	21.4	0.84 d (6.5)	4, 5, 6
NH		8.17 m	
Ser2			
1	169.2		
2	55.9	4.19 m	
3	60.8	3.66 m	
		3.66 m	
NH		7.96 m / 8.05 m / 8.15 m / 8.63 m*	

Supplementary table 5: continued

	δ_C	δ_H (<i>J</i> in Hz)	HMBC	NOESY
NMeSer				
1	170.6			
2	66.4	2.97 dd (4.5, 7.1)	1, 3, NMe	1
3	61.7	3.54 m	1	1
		3.39 dd (7.2, 10.8)	1, 2	1
NH		*		
NMe	34.2	2.15 s	2	
OH-Val1				
1	170.0			
2	59.2	4.26 m		1, 1Phe
3	71.1			
4	27.3	1.03 s	2, 3, 5	
5	24.8	0.97 s	2, 3, 4	
NH		7.98 d (7.7)		
Phe				
1	171.0			
2	53.8	4.62 s		1, 1OH-Val1
3	36.7	3.06 dd (3.7, 14.2)	2, 1', 2', 6'	1
		2.78 dd (10.3, 14.2)	1, 2, 1', 2', 6'	1
1'	137.9			
2', 6'	128.9	7.24 m	3, 3', 4', 5'	
3', 5'	127.9	7.19 m	1', 3', 5'	
4'	126.0	7.13 m	2', 6'	
NH		8.47 d (8.6)	1OH-Val1	1OH-Val1
OH-Val2				
1	170.1			
2	60.1	4.35 m	1, 3, 5, 1Ser1	1, 1Phe
3	71.0			
4	27.3	1.11 s	2, 3, 5	
5	25.1	1.08 s	2, 3, 4	
NH		8.15 m	1Phe	1Phe
Ser1				
1	169.4			
2	54.8	4.24 m		1
3	61.4	3.54 m	1	1
		3.47 m	1	1
NH		8.14 m	1OH-Val2	

Supplementary table 6: NMR Spectroscopic data for c814 (DMSO-*d*₆); * not assigned

	δ_C	δ_H (<i>J</i> in Hz)	HMBC	NOESY
β-Leu				
1	171.7			
2	72.0	3.85 d (2.9)	1	1, 1Ser1
3	49.6	4.12 m		1
4	39.7	1.35 m	3, 7	
		1.25 m	3, 7	
5	23.8	1.53 m		
6	23.2	0.8 d (6.5)	5, 7	
7	21.7	0.8 d (6.5)	5, 6	
NH		7.61 d (9.0)	1Ser1	
Ser2				
1	172.8			
2	55.4	3.82 m	1	1
3	62.3	3.58 m	1, 2	1
		3.49 m	1, 2	1
NH		7.69 d (6.8)	1, 1 β Leu	1, 1 β -Leu

Supplementary table 6: continued

	δ_C	δ_H (J in Hz)	HMBC
NMeSer			
1	168.4		
2	64.2	3.52 m	NMe
3	60.9	3.67 m	
		3.57 m	1, 2
NH		8.29 m	
NMe	32.9	2.35 s	2
OH-Val1			
1	170.0		
2	60.2	4.34 m	1, 3, 4, 5, 1NMe
3	71.6		
4	27.5	1.07 s	2, 3, 5
5	25.5	1.04 s	2, 3, 4
NH		7.84 m / 8.59 m	
Phe			
1	171.3		
2	54.4	4.66 m	1, 3
3	37.4	3.05 m	2, 1', 2', 6'
		2.81 m	2, 2', 6'
1'	137.8		
2', 6'	129.4	7.23 m	1', 2', 6'
3', 5'	128.3	7.21 m	1', 3', 5'
4'	126.5	7.15 m	2', 6'
NH		8.24 m	1OH-Val1
OH-Val2			
1	*		
2	60.1	4.34 m	
3	72.1		
4	28.1	1.15 s	2, 3, 5
5	24.9	1.10 s	2, 3, 4
NH		7.84 m / 8.59 m	
Ser1			
1			
2	60.3	4.48 m	
3		3.72 m	
		3.54 m	
NH		7.66 m	

Supplementary table 7: NMR Spectroscopic data for c581 (DMSO- d_6); * not assigned

	δC	δH (<i>J</i> in Hz)	HMBC
2OH-β-Val			
1	170.5		
2	196.7		
3	57.9	5.15 m	1, 2, 4, 5, 6
4	29.6	2.16 m	5, 6
5	19.5	0.85 m	3, 4, 6
6	17.9	0.78 m	3, 4, 5
NH		8.03 m	
Ser2			
1	*		
2	55.4	4.38 m	
3	61.8	3.71 m	
		3.60 m	
NH		8.21 m	
Tyr			
1	*		
2	55.2	4.33 m	
3	36.8	2.93 m	
		2.70 m	
1'	128.2		
2', 6'	130.3	7.00 m	3, 2', 3', 4', 5', 6'
3', 5'	115.2	6.62m	1', 3', 4', 5'
4'	156.1		
NH		8.36 m	
β-Lys			
1	169.7		
2	40.9	2.21 m	
		2.21 m	
3	45.9	4.05 m	
4	30.7	1.52 m	
		1.36 m	
5	23.8	1.53 m	
		1.53 m	
6	38.8	2.74 m	4, 5
		2.74 m	4, 5
NH		7.89 m	
Ala			
1	175.7		
2	48.9	4.09 m	1, 3, 1 β -Lys
3	18.2	1.22 d (6.6)	1, 2
NH		7.94 m	1 β -Lys

Supplementary table 7: continued

	δ_C	δ_H (<i>J</i> in Hz)	HMBC	NOESY
NMeSer				
1	169.7			
2	65.9	3.11 m		NMe
3	61.7	3.59 m	1	
		3.45 m		
NH		*		
NMe	34.2	2.22 s	1,2	
OH-Val1				
1	170.0			
2	59.4	4.28 m	1, 4	5
3	71.5			
4	25.3	1.01 s	2, 3, 5	2
5	27.5	1.05 s	2, 3, 4	2
NH		8.03 m	1NMe	
Phe				
1	*			
2	54.1	4.65 m		3, NH
3	37.2	3.05 m	2', 6'	2, 3
		2.78 m	2, 2', 6'	2, 3
1'	137.9			
2', 6'	129.3	7.23 m	2', 4', 6'	
3', 5'	128.1	7.21 m	1', 3', 5'	2, 3
4'	126.4	7.15 m		
NH		8.20 d (6.9)	1 OH-Val1	2, 3
OH-Val2				
1	170.0			
2	60.1	4.40 d (9.1)	1, 3, 5, 1Phe	NH
3	71.8			
4	27.8	1.13 s	2, 3, 5	
5	24.9	1.10 s	2, 3, 4	2
NH		8.09 m		2Phe, 3Phe
Ser1				
1				
2	55.6	4.31 m		
3	61.7	3.59 m		
		3.54 m		
NH		8.10 m		

Supplementary table 8: NMR Spectroscopic data for c582 (DMSO-*d*₆); * not assigned

	δ_C	δ_H (J in Hz)	HMBC	NOESY
2OH-β-Val				
1	173.3			
2	70.5	4.12 m		NH, 4, 5, 6
3	57.2	3.80 m		NH, 4, 5, 6
4	29.2	1.80 m		
5	19.9	0.90 d (6.1)	3, 4, 6	NH, 2, 3, 4, 6
6	19.4	0.82 d (6.1)	3, 4, 5	NH, 2, 3, 4, 5
NH		7.63 d (8.4)		3, 4, 5, 6
Ser2				
1	170.0			
2	55.1	4.19 m	1	NH
3	61.7	3.59 m		NH
		3.48 m	1	
NH		7.89 m	12OH- β -Val	
Tyr				
1	170.7			
2	55.0	4.32 m	1	NH, 3, 2', 6'
3	36.4	2.96 m		2, 3, 2', 6', NH
		2.75 m	2', 6'	2, 3, 2', 6', NH
1'	128.1			
2', 6'	130.1	7.00 d (8.4)	3, 2', 4', 6'	2, 3, 3', 5', NH
3', 5'	115.1	6.63 d (8.4)	1', 3', 4', 5'	2', 6'
4'	155.9			
NH		8.10 m	1, 1AS	2
β-Lys				
1	169.6			
2	40.8	2.22 m		4
		2.16 m		
3	45.6	4.04 m		2, 4, 6, NH
4	30.4	1.51 m		3
		1.34 m		4
5	23.6	1.48 m		
		1.52 m		6
6	38.6	2.74 m	4	4
NH		7.73 d (9.1)		
Ala				
1	175.4			
2	49.0	4.03 m	1, 3, 1 β -Lys	3, NH
3	18.2	1.21 d (7.6)		2, NH
NH		7.86 m	1 β -Lys	3

Supplementary table 8: continued

	δ_C	δ_H (J in Hz)	HMBC
NMeSer			
1	*		
2	59.9	4.46 m	
3	64.7	3.39 m	
		3.39 m	
NH		*	
NMe	33.2	2.31 s	3
OH-Val1			
1	169.8		
2	59.8	4.51 m	1, 3
3	71.8		
4	24.6	1.13 s	2, 3, 5
5	24.9	1.10 s	3, 4
NH		8.10 m	
Phe			
1	*		
2	54.3	4.64 m	
3	37.3	3.04 m	2, 1', 2', 6'
		2.79 m	
1'	137.7		
2', 6'	129.3	7.22 m	3, 2', 4', 6'
3', 5'	128.2	7.20 m	1', 3', 5'
4'	126.4	7.15 m	2', 6'
NH		8.23 m	1OH-Val1
OH-Val2			
1	*		
2	59.9	4.46 m	
3	71.4		
4	27.5	1.06 s	2, 3, 5
5	25.6	1.03 s	2, 3, 4
NH		*	
Ser1			
1	*		
2	64.5	3.42 m	
3	61.0	3.63 m	
		3.52 m	2
NH		*	

Supplementary table 9: NMR Spectroscopic data for c587 (DMSO- d_6); * not assigned

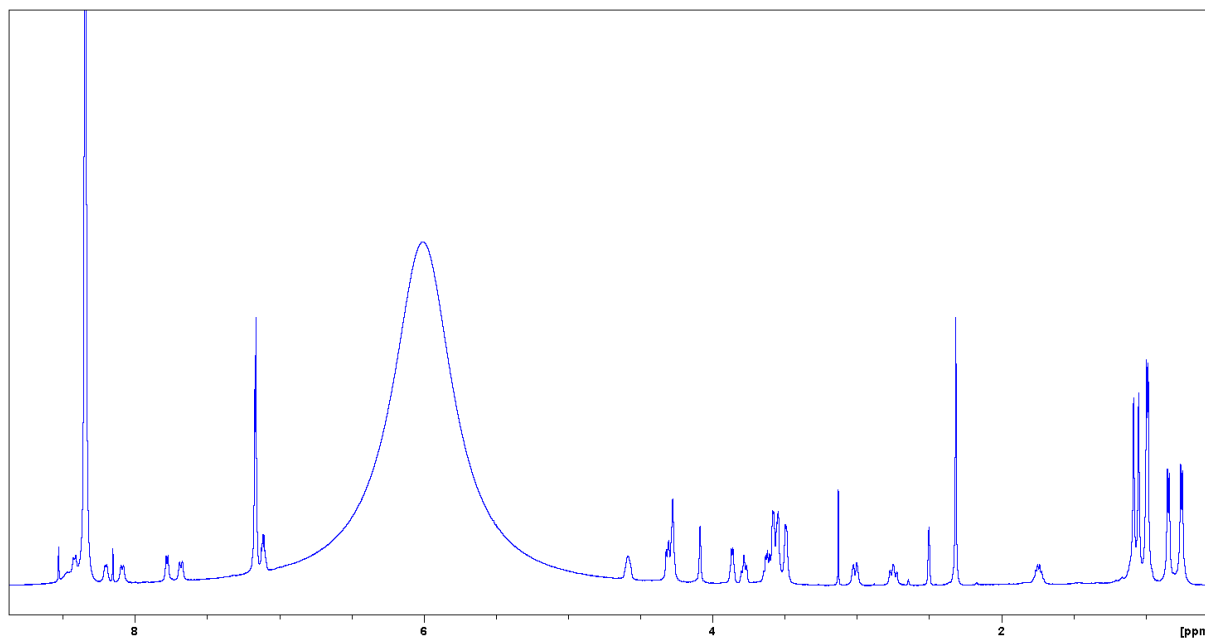
	δ_C	δ_H (J in Hz)	HMBC
β-Leu			
1	169.8		
2	52.0	5.09 m	
3	59.9	4.46 m	
4	38.4	3.31 m	
5	28.3	1.45 m	
6	24.0	0.79 m	5, 7
7	21.3	0.84 m	4, 6
NH		8.06 m	
Ser2			
1	168.9		
2	59.9	4.30 m	1, 1 β -Leu
3	61.6	3.70 m	
		3.58 m	
NH		8.27 m	
Tyr			
1	*		
2	55.1	4.33 m	
3	36.8	2.91 m	2', 6'
		2.69 m	2', 6'
1'	127.6		
2', 6'	130.2	6.99 m	3, 2', 3', 4', 5', 6'
3', 5'	115.1	6.63 m	1', 3', 4', 5'
4'	155.9		
NH		8.33 m	
β-Lys			
1	169.3		
2	40.7	2.20 m	
		2.20 m	
3	45.9	4.03 m	
4	30.5	1.50 m	
		1.35 m	
5	23.7	1.51 m	
		1.51 m	
6	38.6	2.74 m	4, 5
		2.74 m	4, 5
NH		7.84 m	

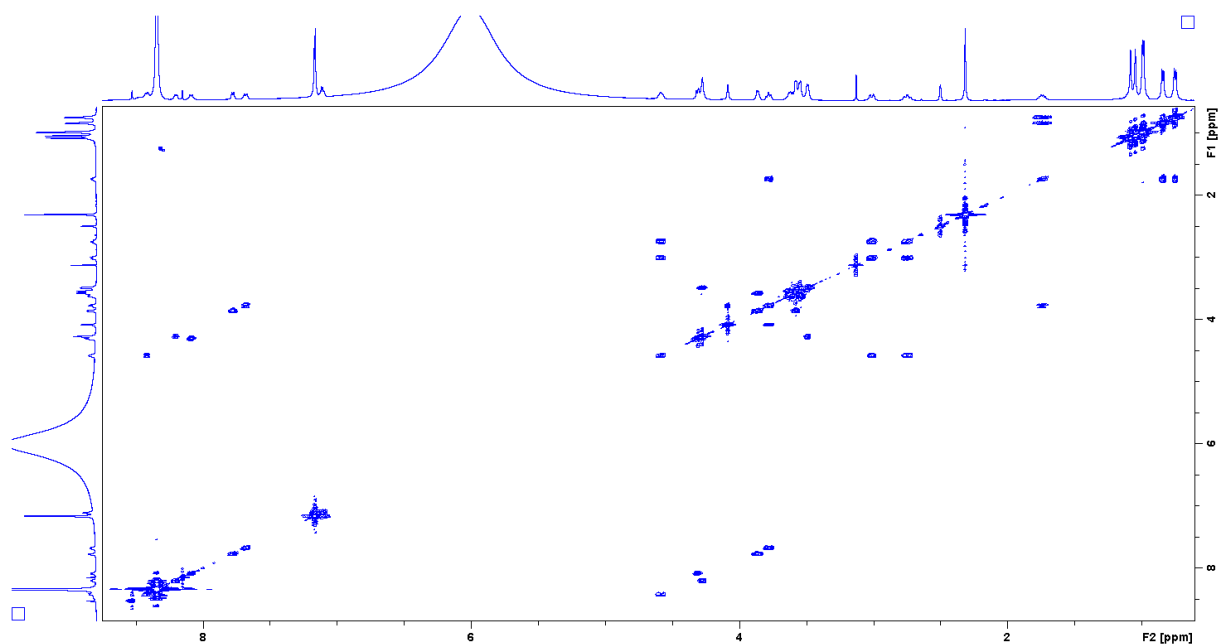
Supplementary table 9: continued

	δ_C	δ_H (J in Hz)	HMBC
Ala			
1	175.3		
2	48.9	4.08 m	1, 3, 1 β -Lys
3	18.0	1.22 d (6.9)	1, 2
NH		7.97 m	1 β -Lys

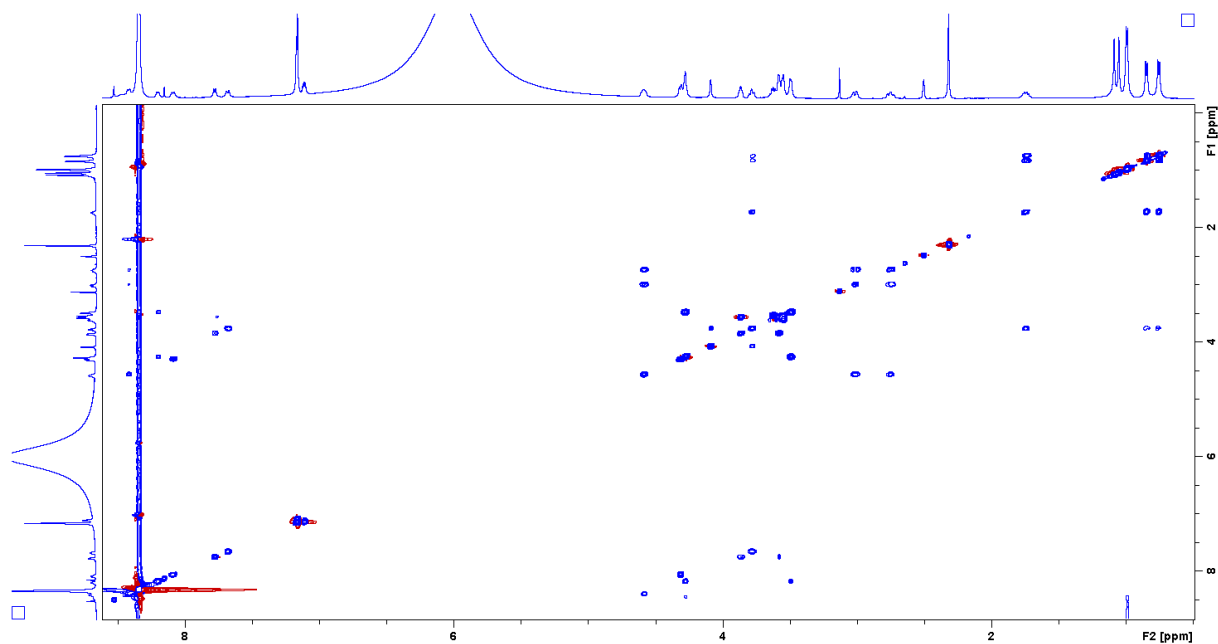
Supplementary table 9: continued

The NMR spectra of the isolated myxoprincomides from DK897 are depicted in the following figures:

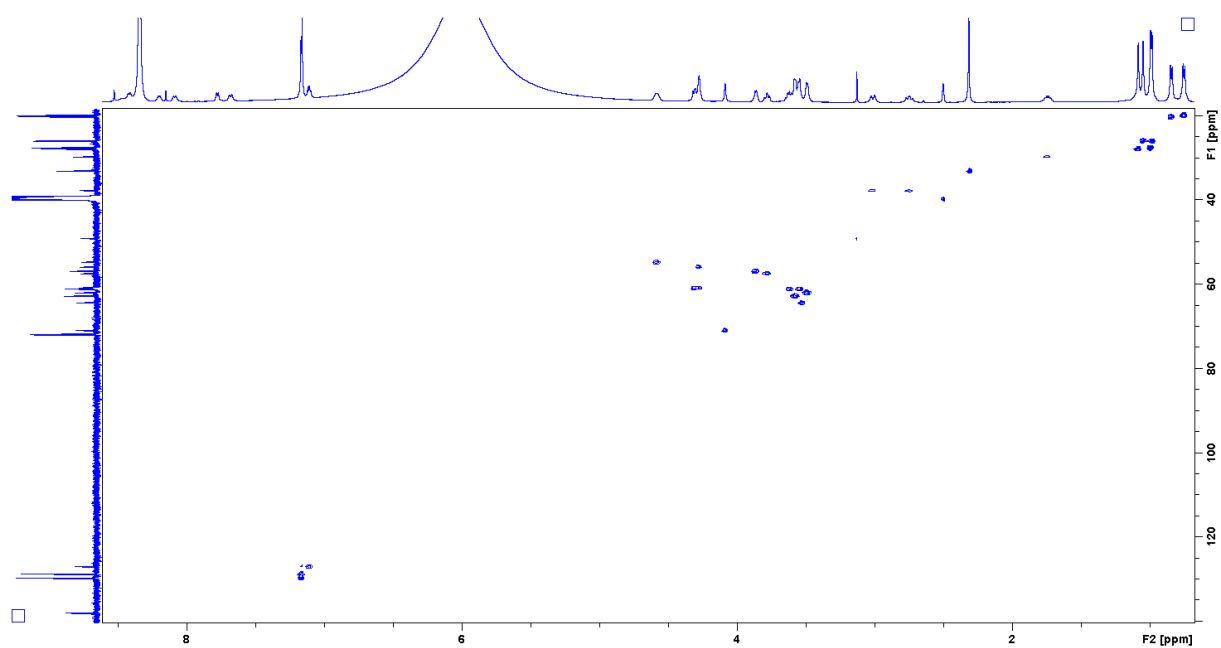
Supplementary figure 11: ^1H NMR spectrum of myxoprincomide c800 in $\text{DMSO}-d_6$.



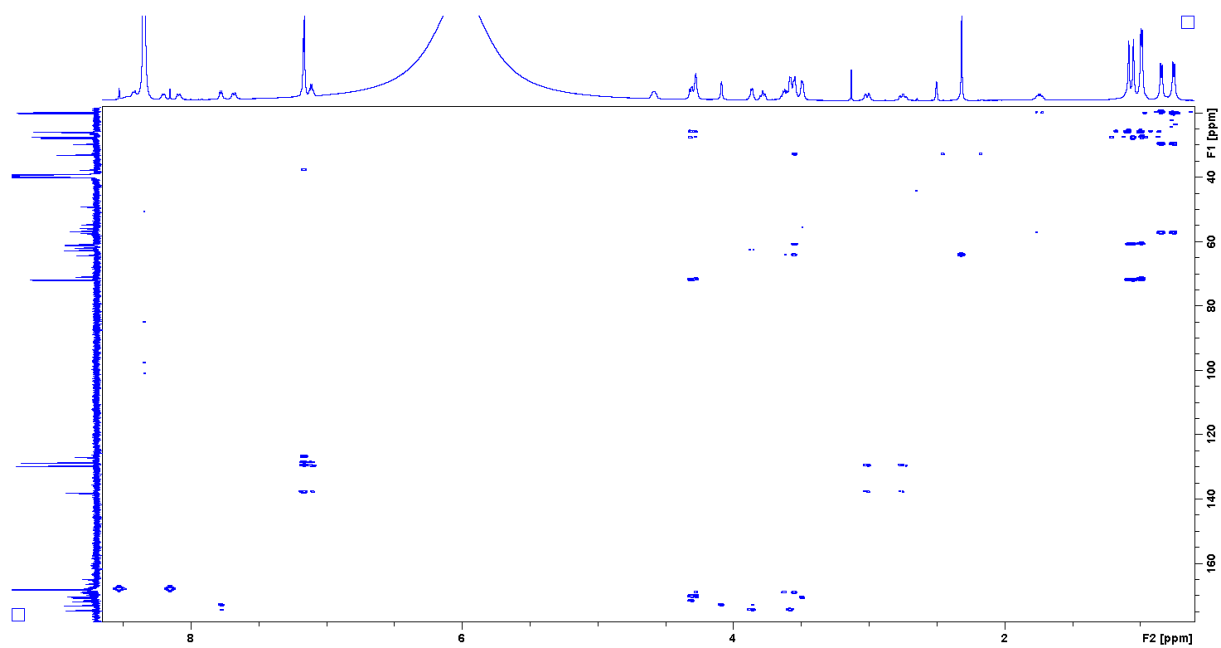
Supplementary figure 12: COSY spectrum of myxoprincomide c800 in DMSO- d_6 .



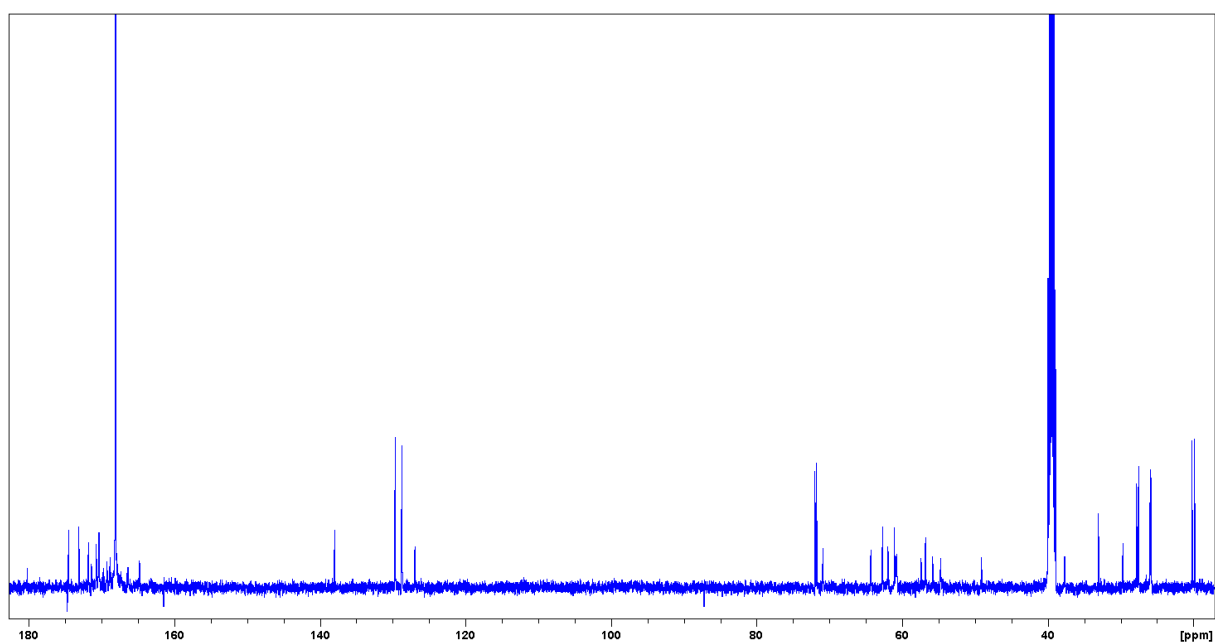
Supplementary figure 13: TOSY spectrum of myxoprincomide c800 in DMSO- d_6 .



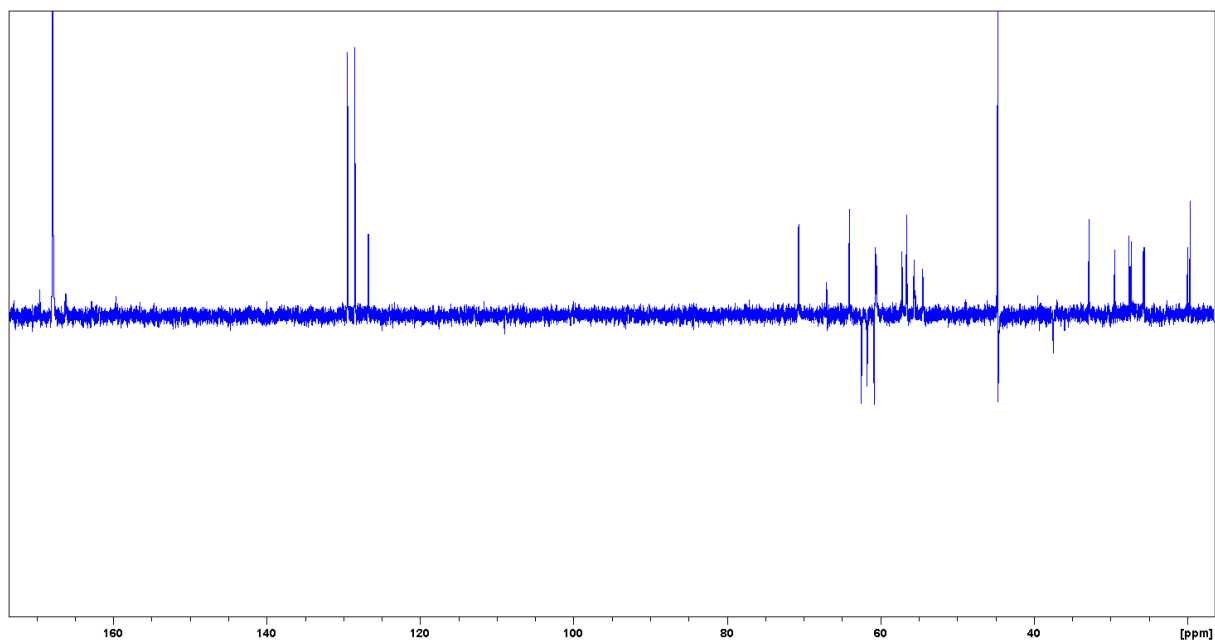
Supplementary figure 14: HSQC spectrum of myxoprincomide c800 in DMSO-*d*₆.



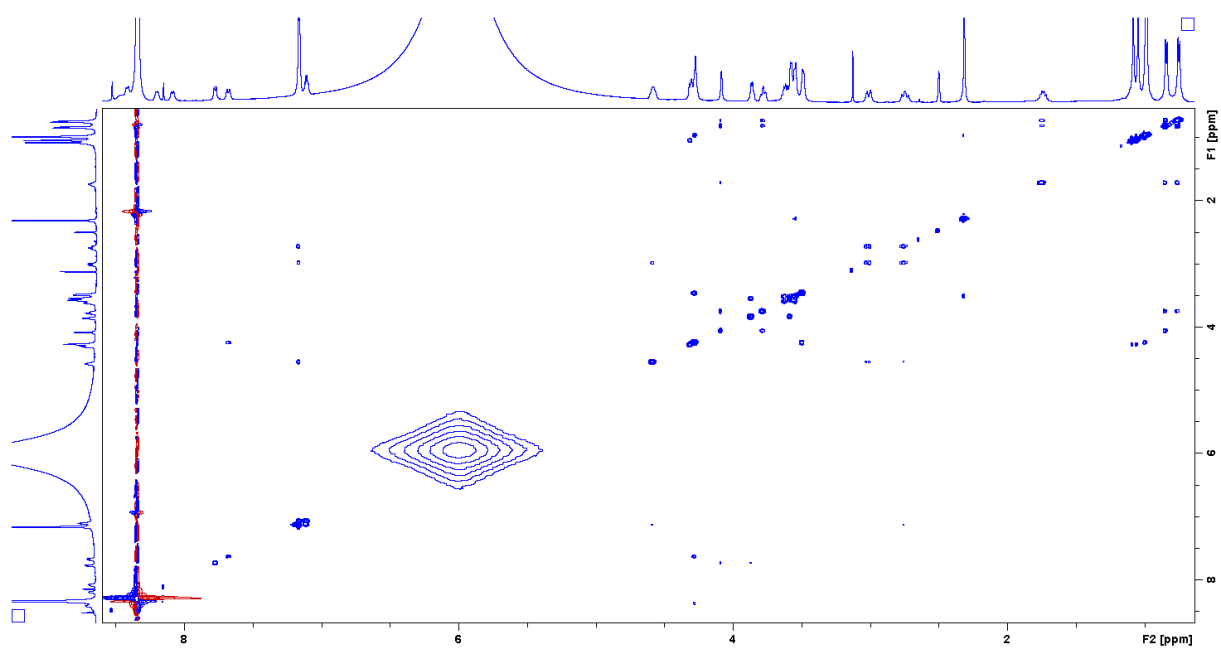
Supplementary figure 15: HMBC spectrum of myxoprincomide c800 in DMSO-*d*₆.



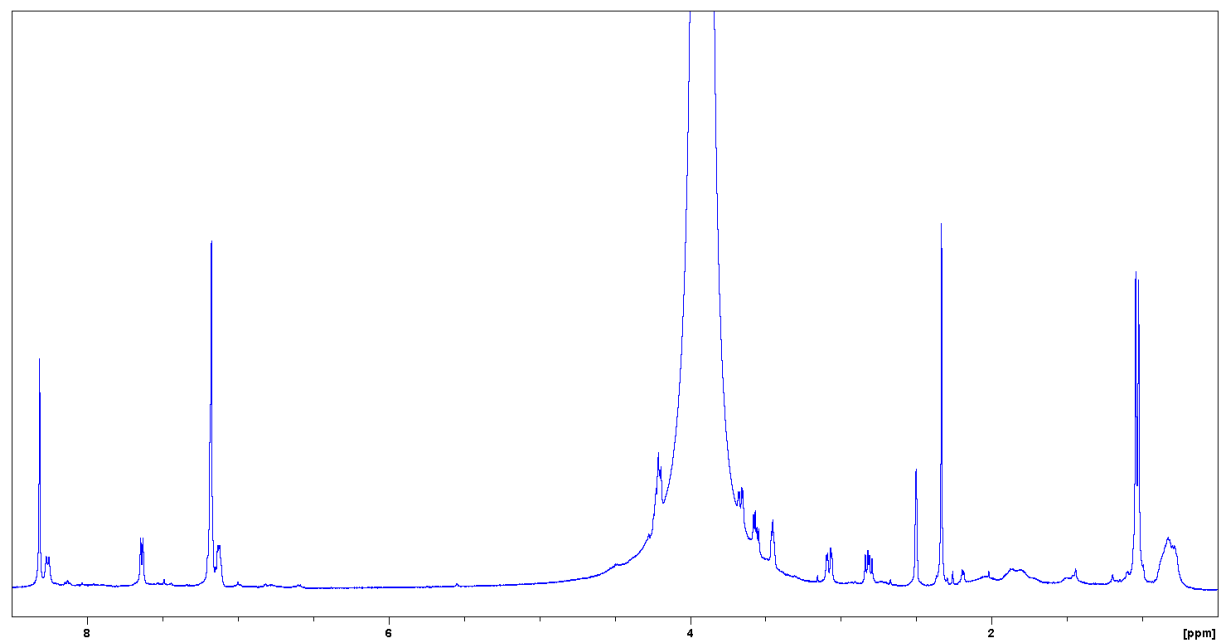
Supplementary figure 16: ^{13}C spectrum of myxoprincomide c800 in $\text{DMSO}-d_6$.



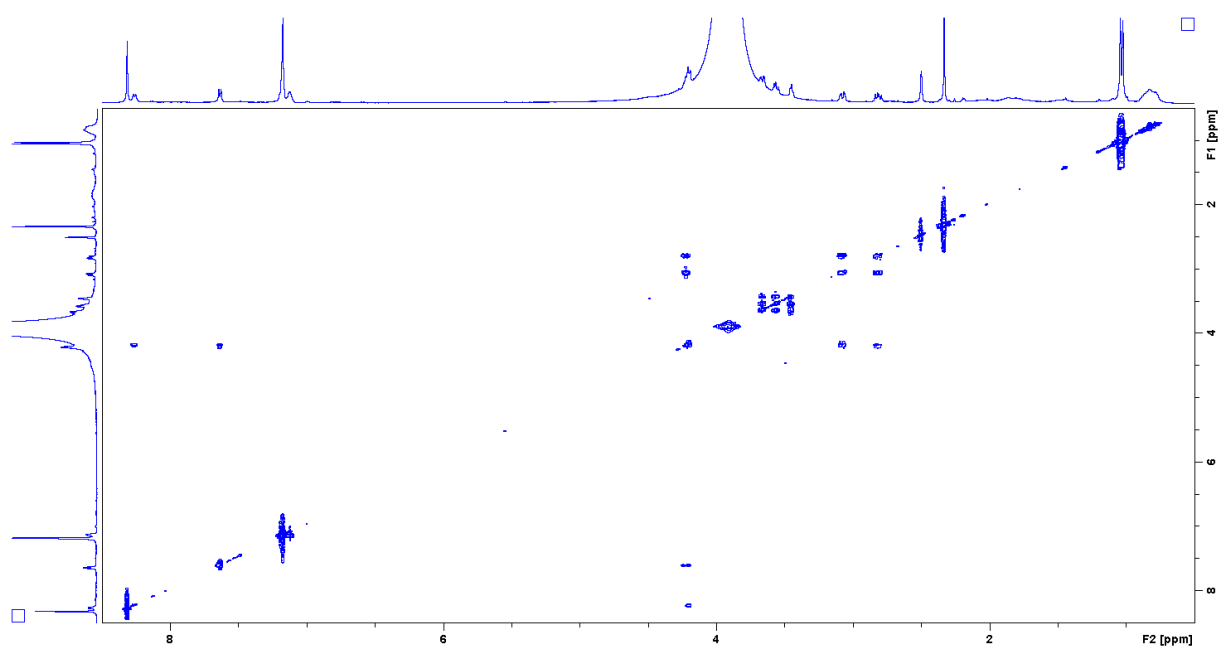
Supplementary figure 17: DEPT135 spectrum of myxoprincomide c800 in $\text{DMSO}-d_6$.



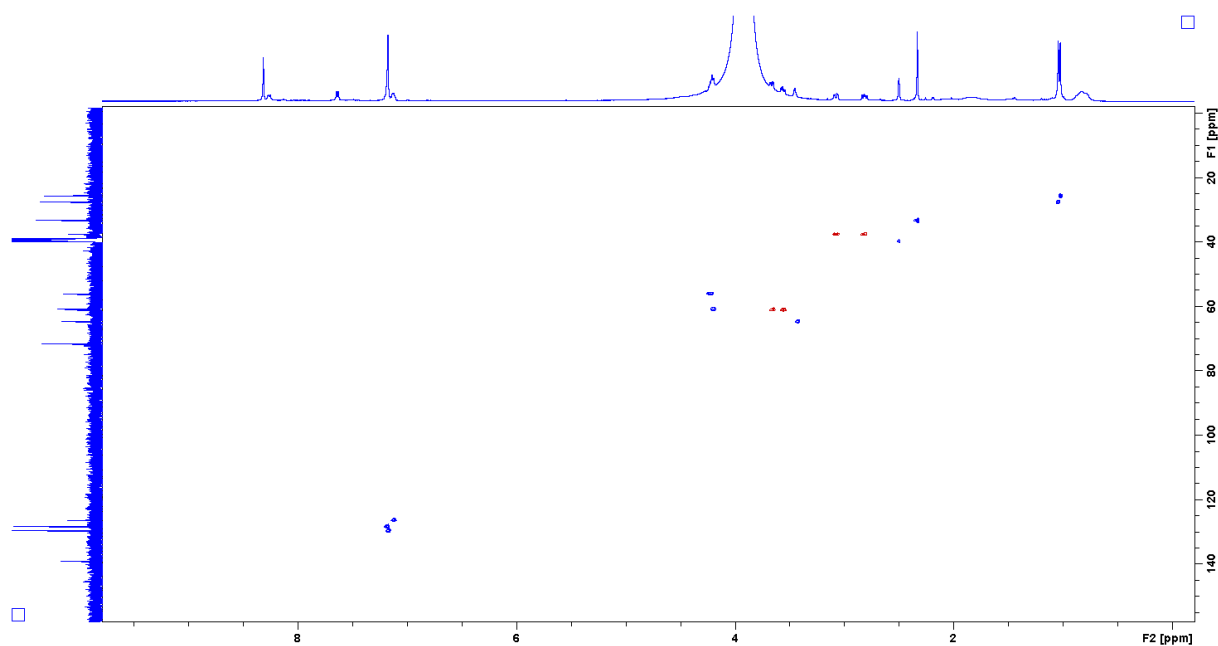
Supplementary figure 18: NOESY spectrum of myxoprincomide c800 in DMSO- d_6 .



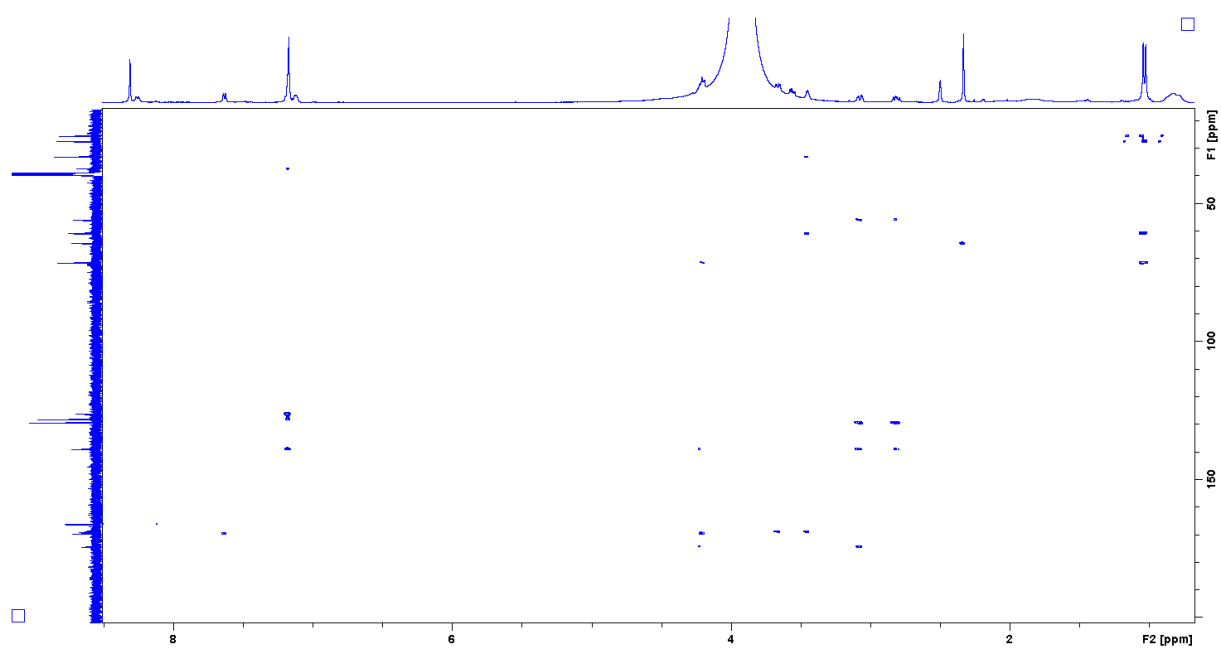
Supplementary figure 19: ^1H NMR spectrum of myxoprincomide c382 in DMSO- d_6 .



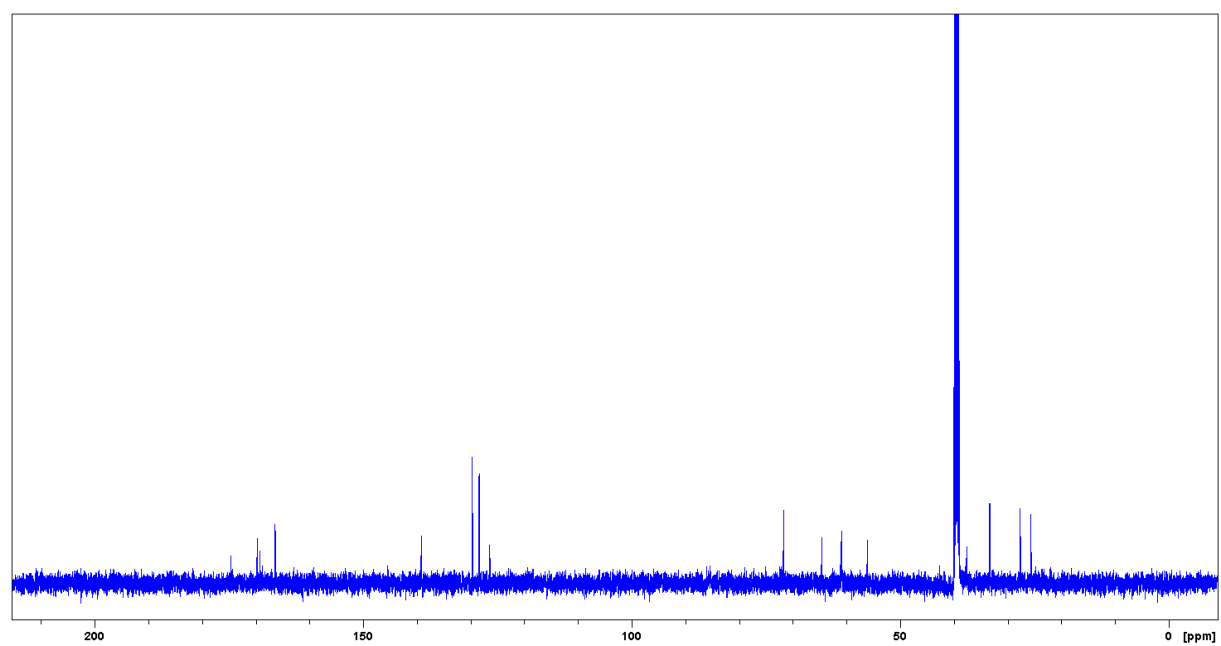
Supplementary figure 20: COSY spectrum of myxoprincomide c382 in DMSO- d_6 .



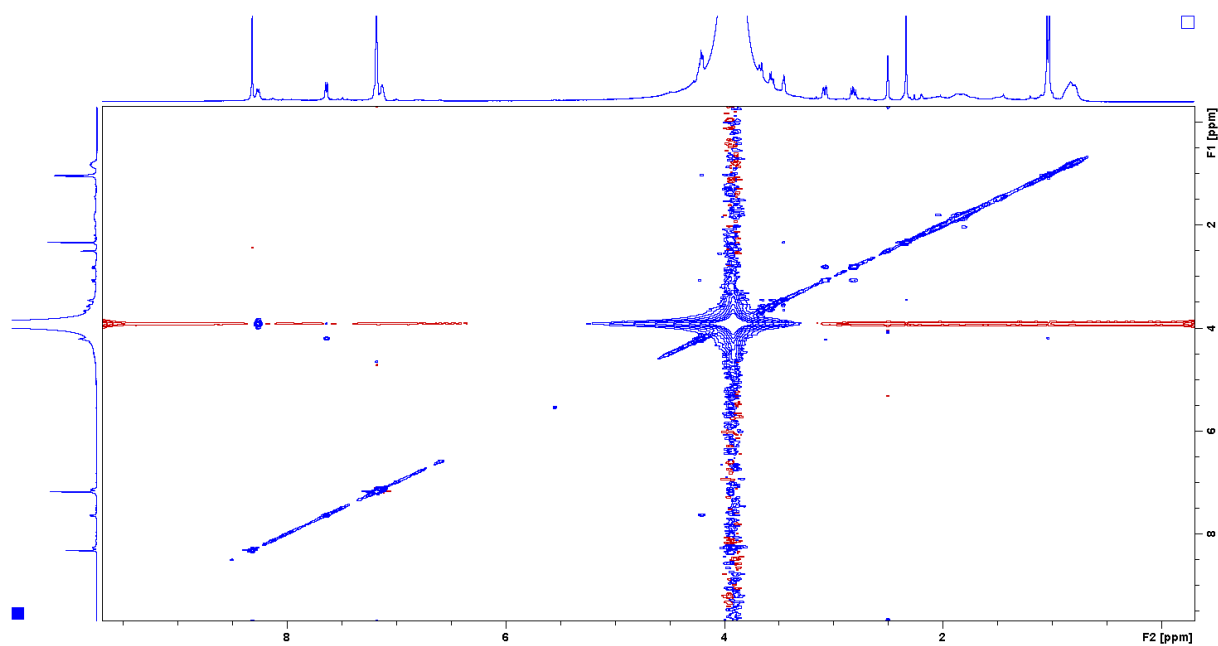
Supplementary figure 21: HSQC spectrum of myxoprincomide c382 in DMSO- d_6 .



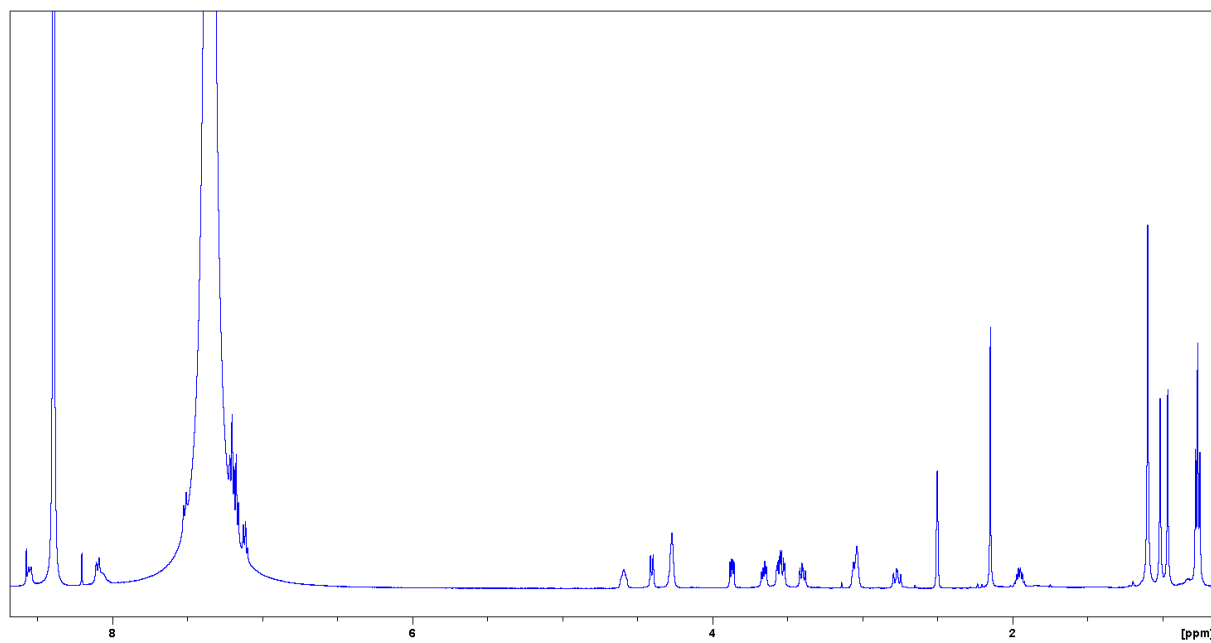
Supplementary figure 22: HMBC spectrum of myxoprincomide c382 in DMSO-*d*₆.



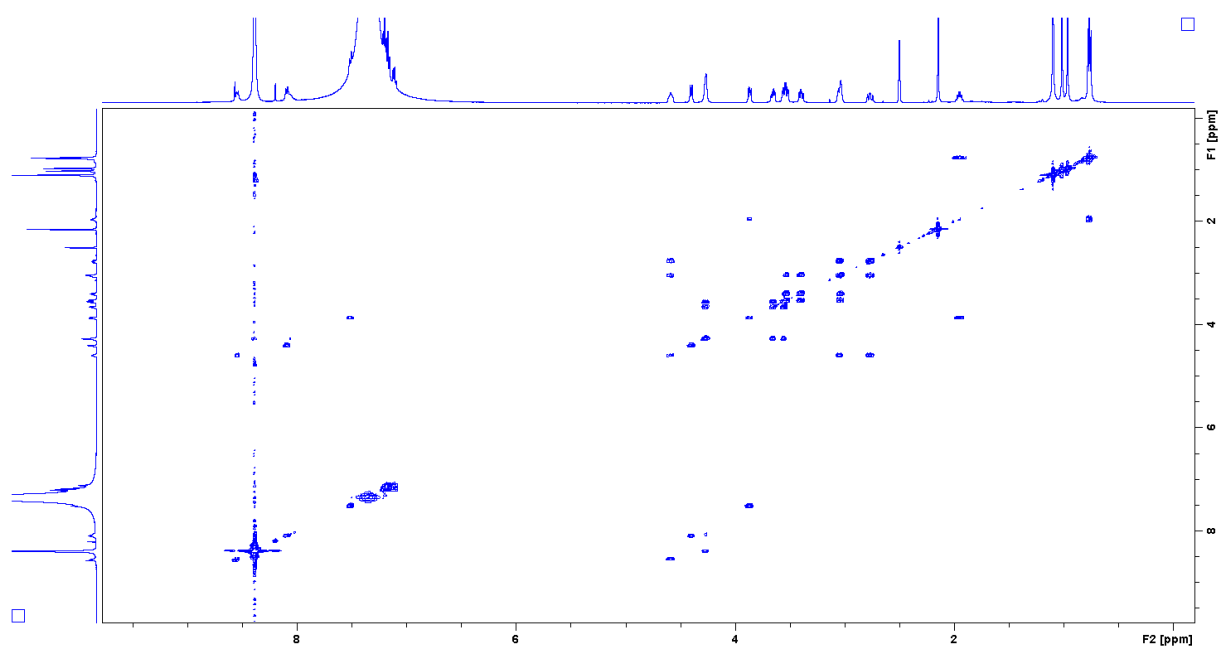
Supplementary figure 23: ¹³C spectrum of myxoprincomide c382 in DMSO-*d*₆.



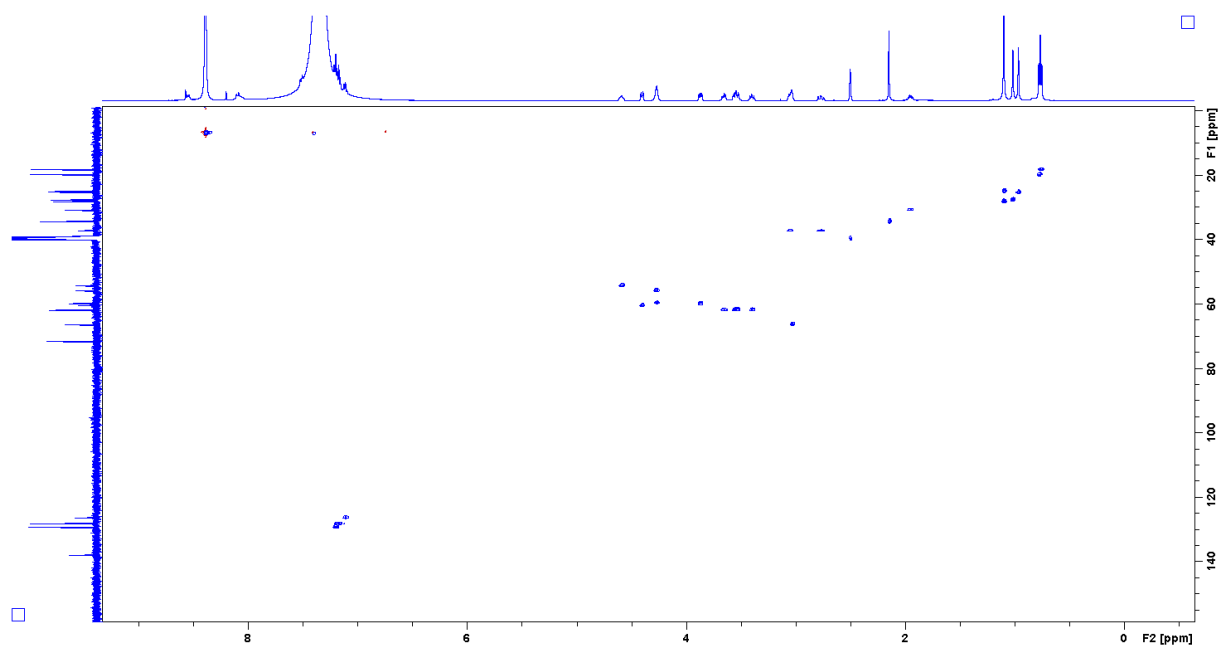
Supplementary figure 24: NOESY spectrum of myxoprincomide c382 in DMSO- d_6 .



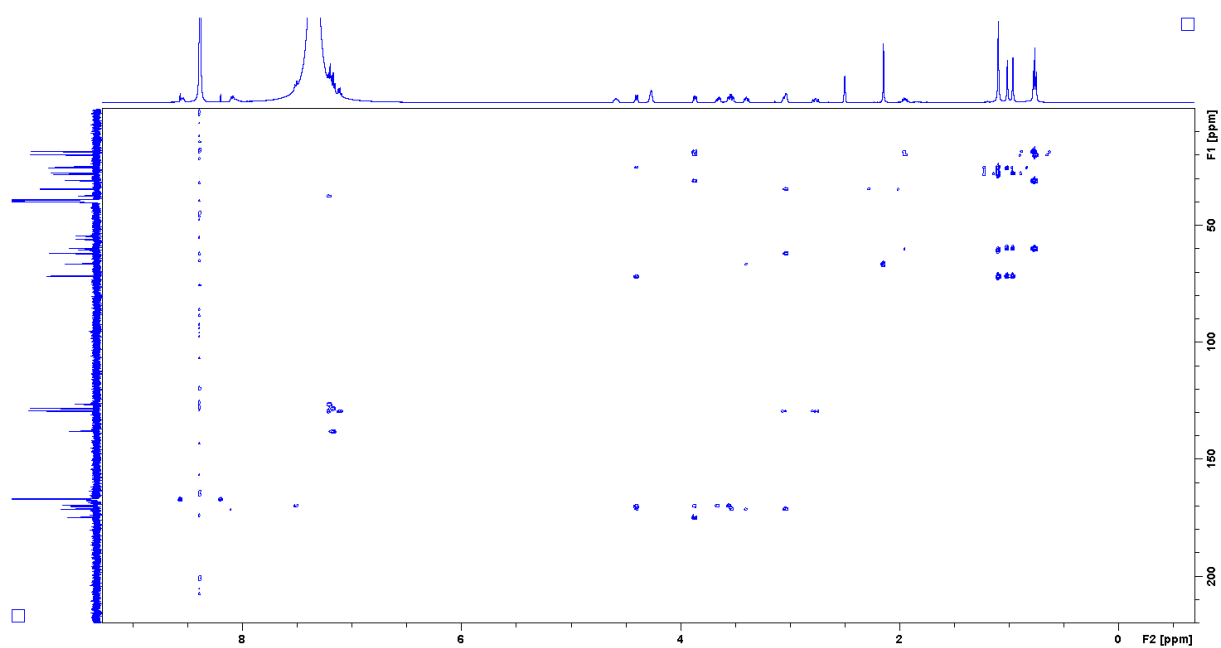
Supplementary figure 25: ^1H NMR spectrum of myxoprincomide c683 in DMSO- d_6 .



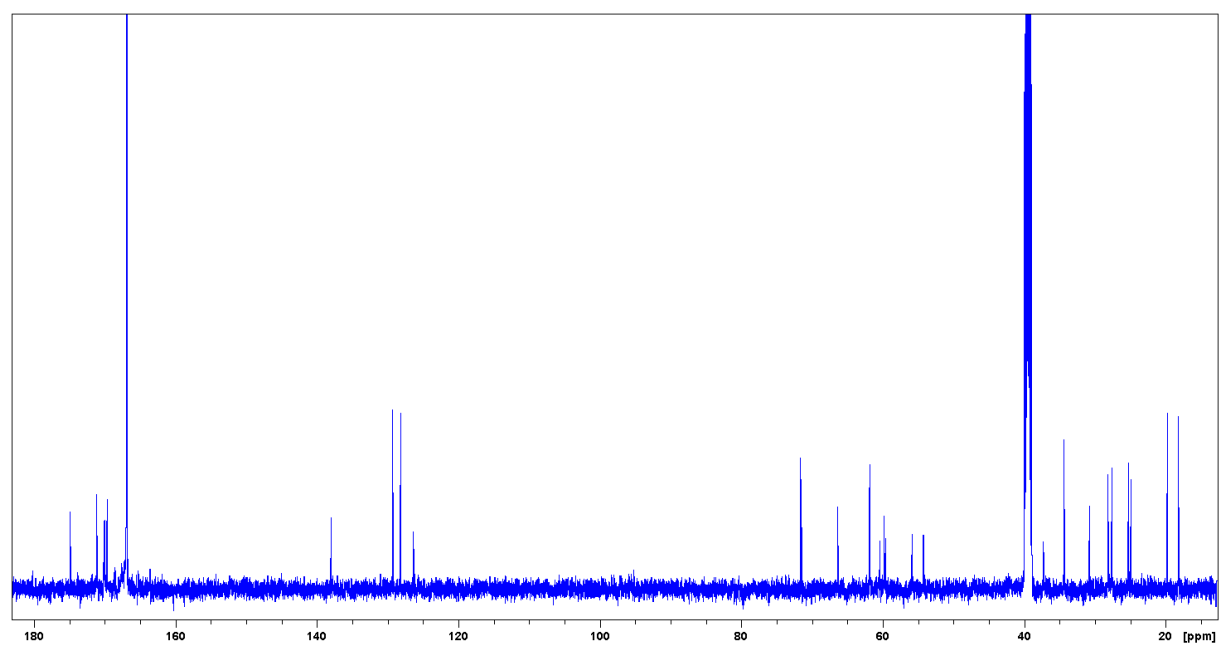
Supplementary figure 26: COSY spectrum of myxoprincomide c683 in DMSO- d_6 .



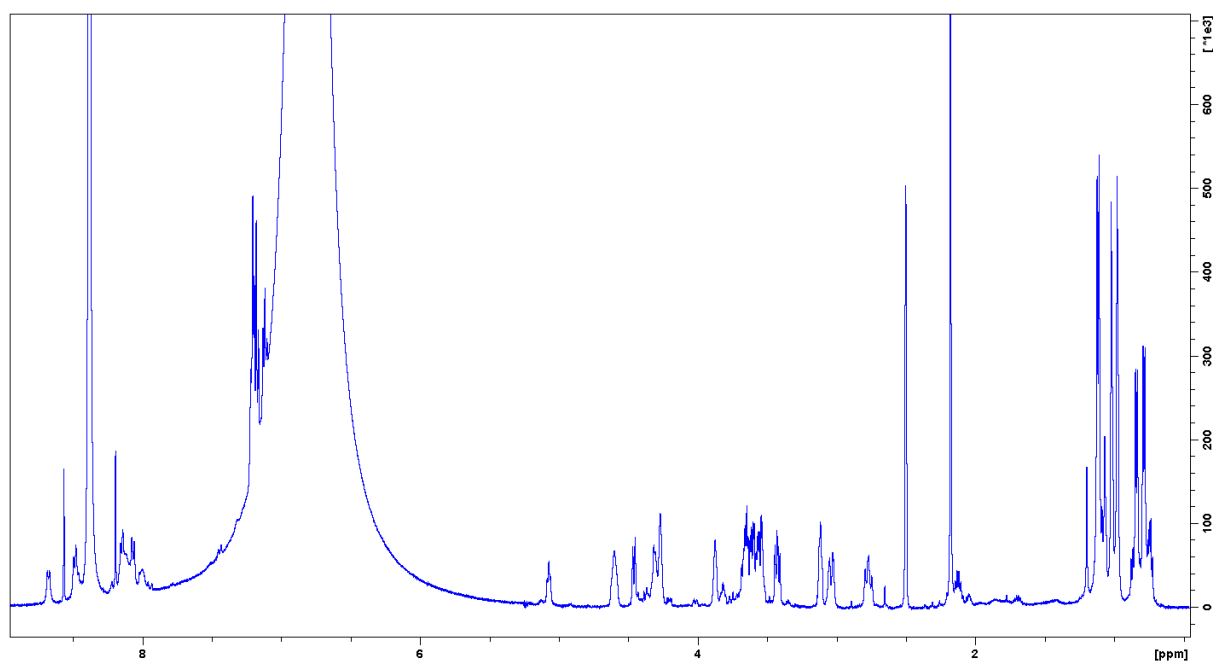
Supplementary figure 27: HSQC spectrum of myxoprincomide c683 in DMSO- d_6 .



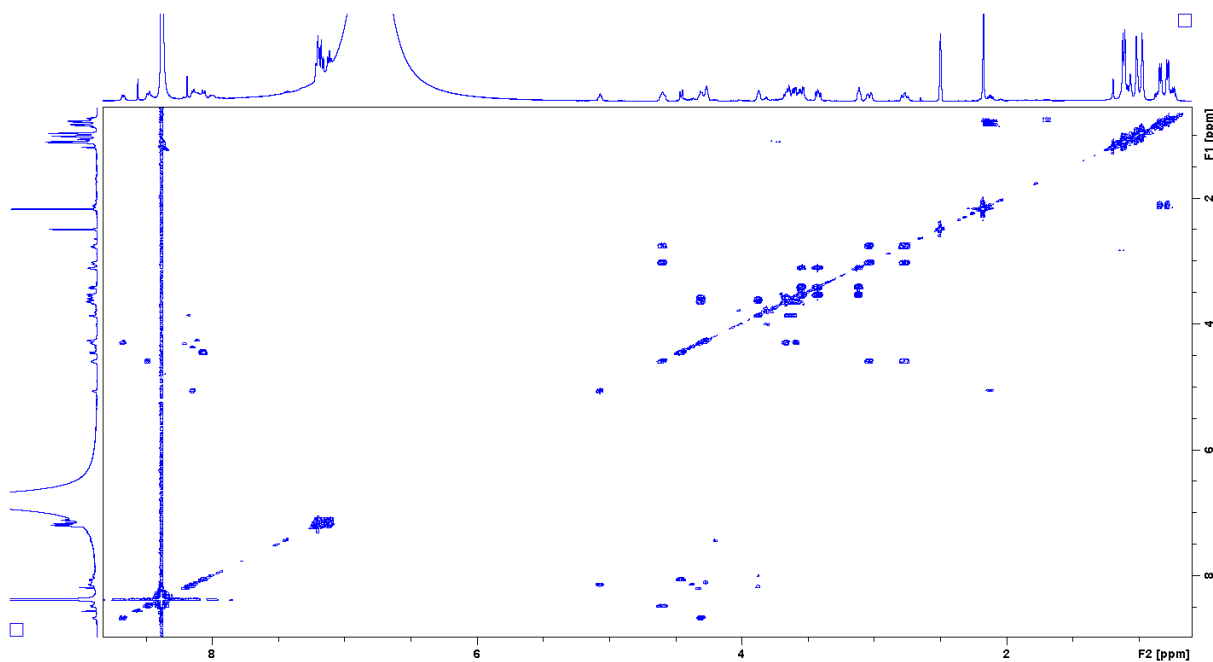
Supplementary figure 28: HMBC spectrum of myxoprincomide c683 in $\text{DMSO}-d_6$.



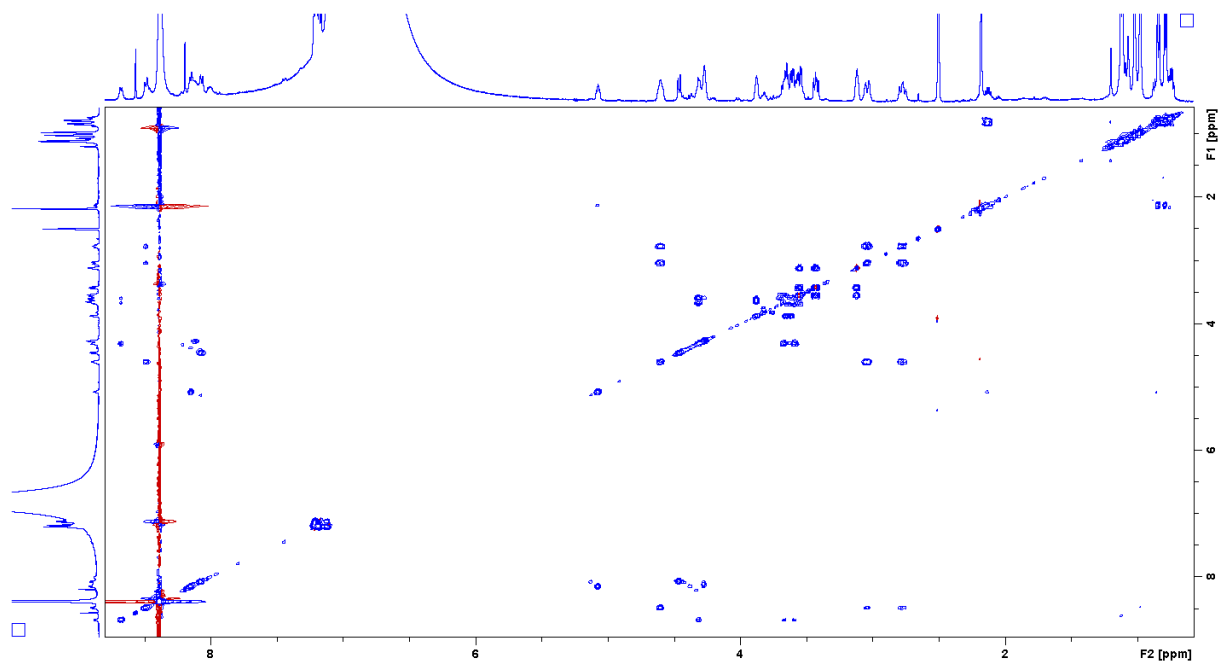
Supplementary figure 29: ^{13}C spectrum of myxoprincomide c683 in $\text{DMSO}-d_6$.



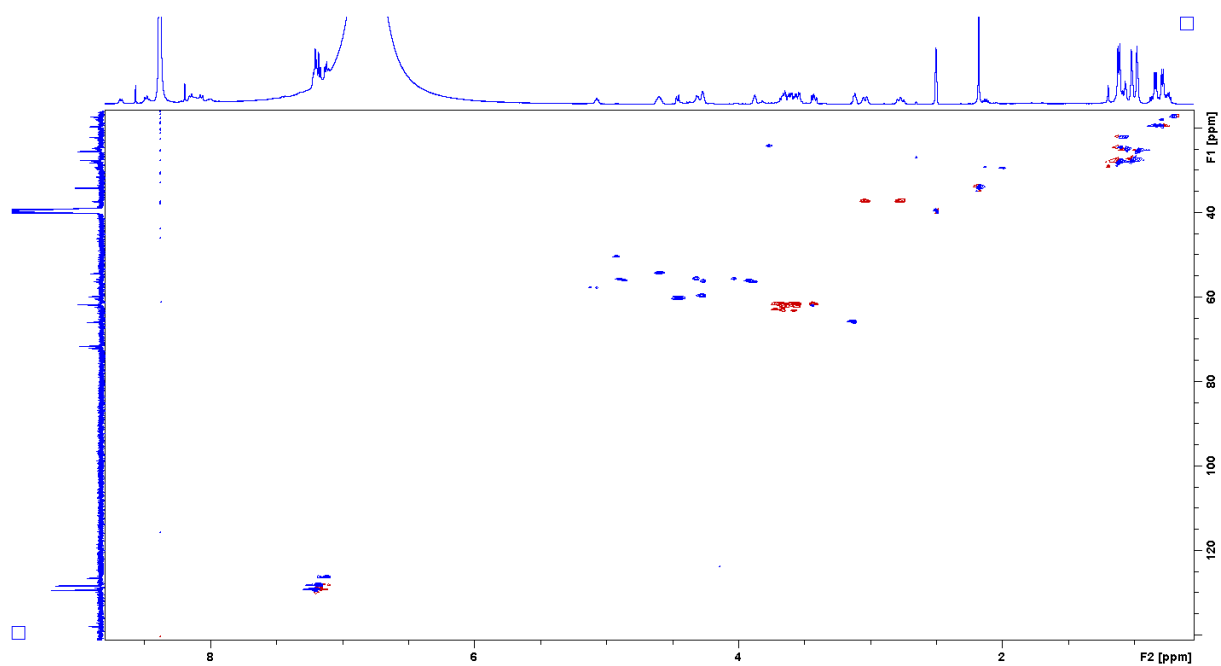
Supplementary figure 30: ¹H NMR spectrum of myxoprincomide c798 in DMSO-*d*₆.



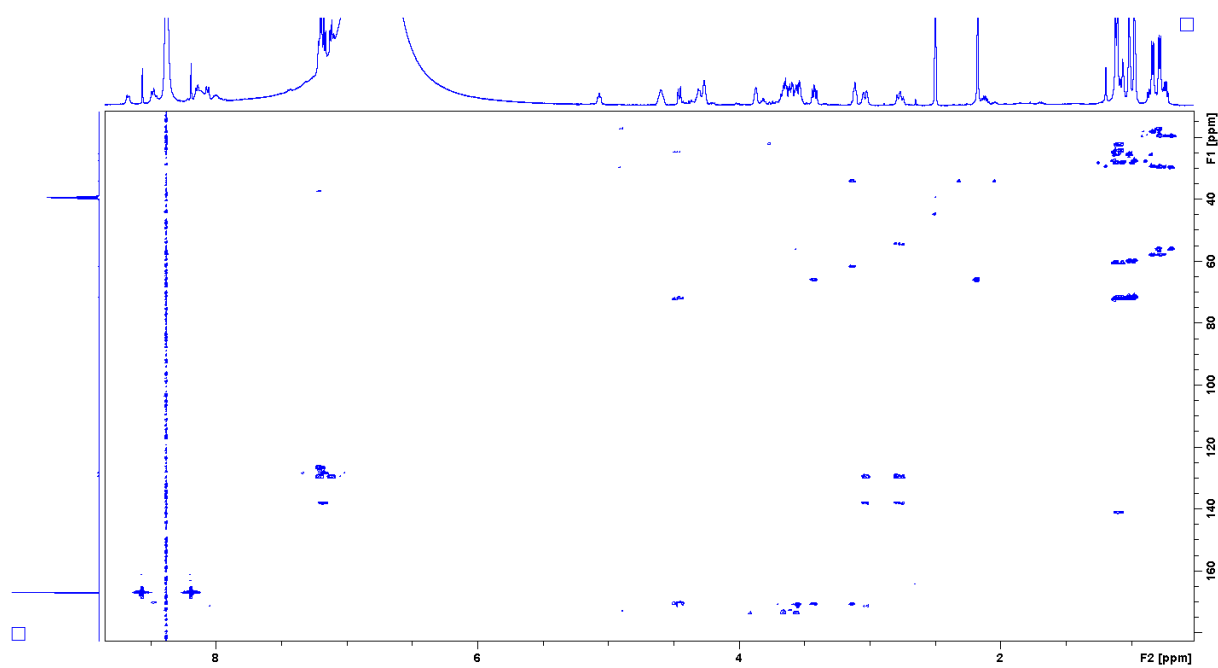
Supplementary figure 31: COSY spectrum of myxoprincomide c798 in DMSO-*d*₆.



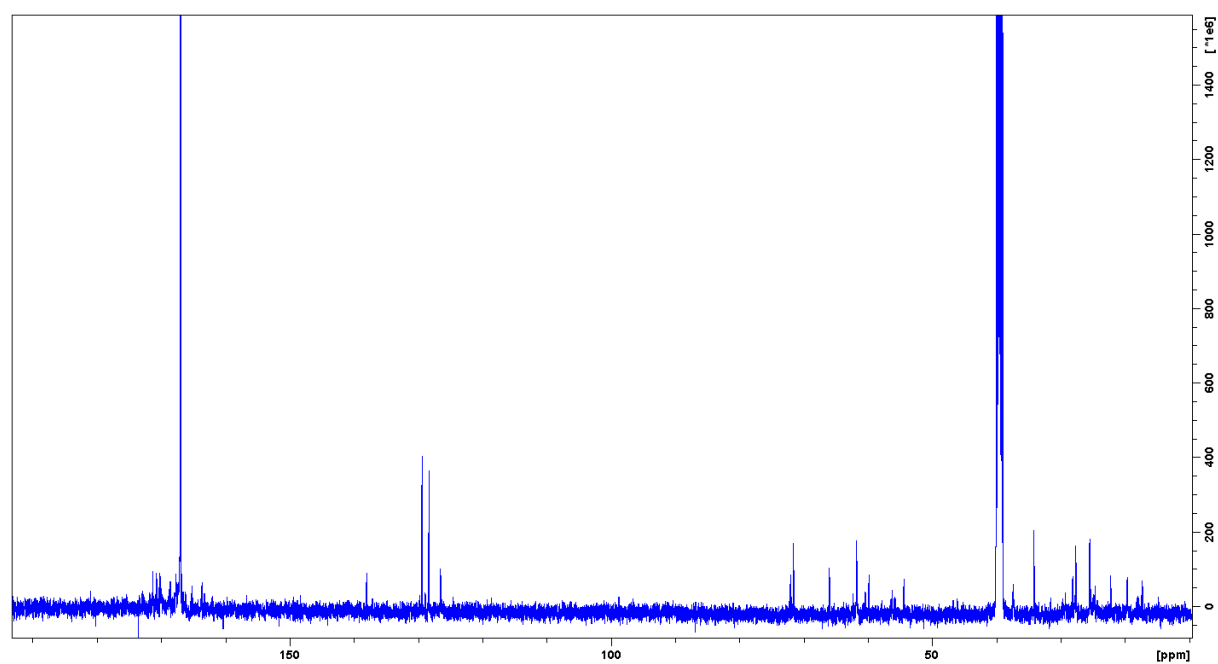
Supplementary figure 32: TOCSY spectrum of myxoprincomide c798 in DMSO- d_6 .



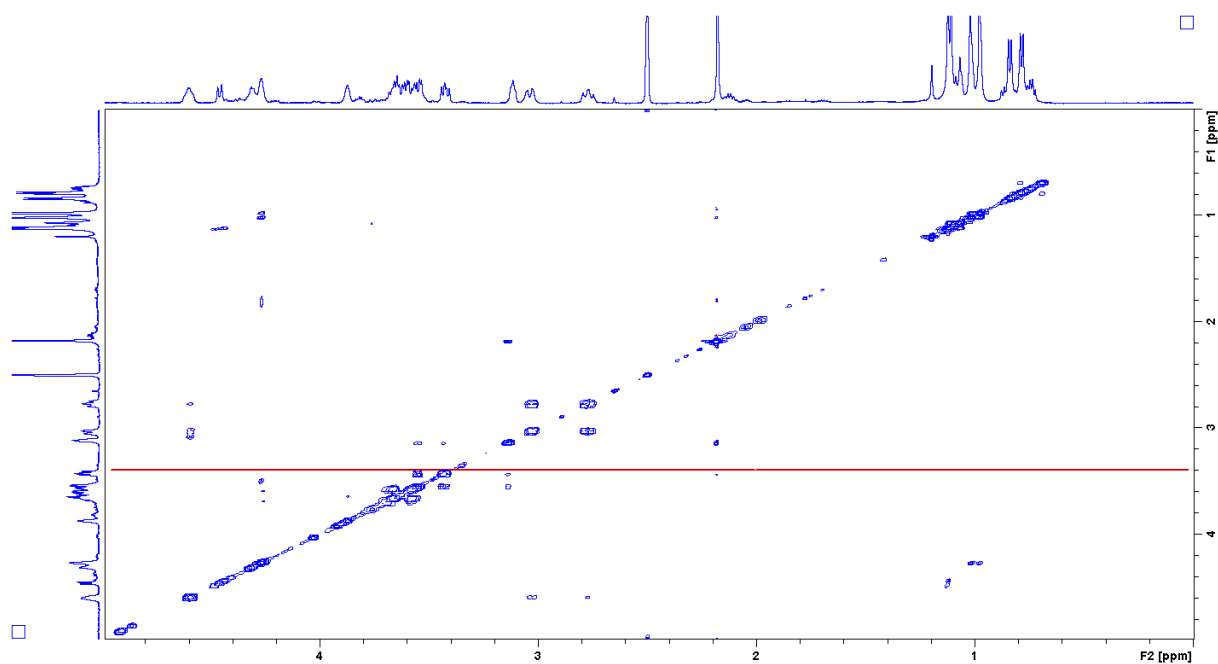
Supplementary figure 33: HSQC spectrum of myxoprincomide c798 in DMSO- d_6 .



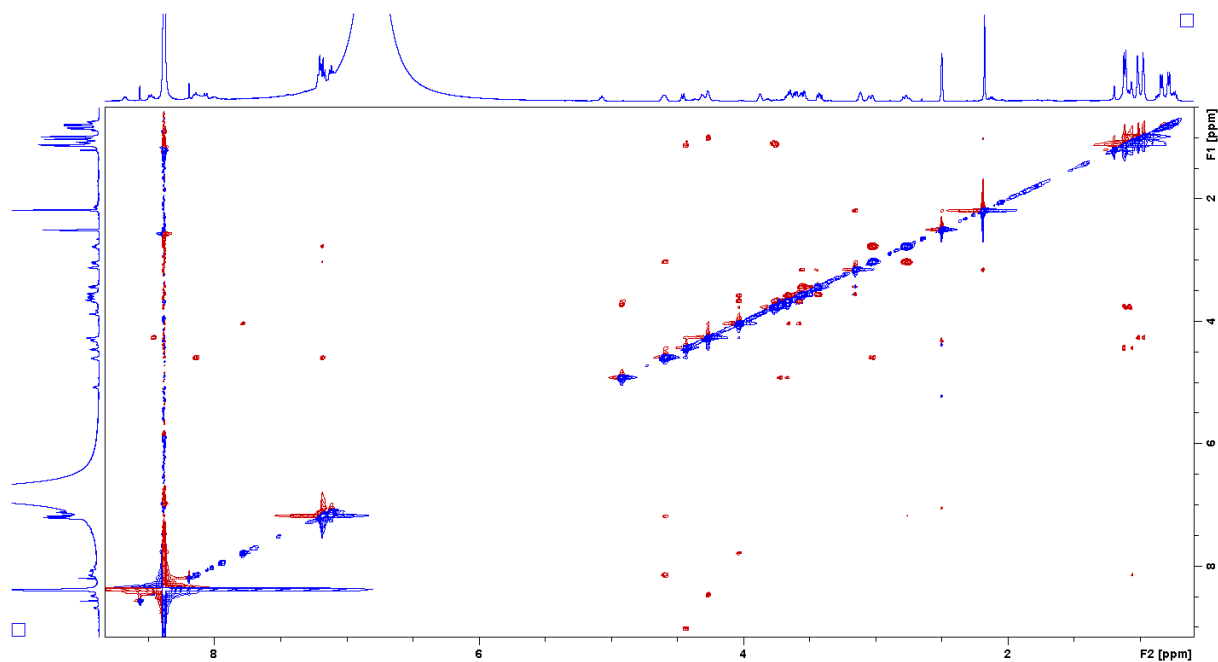
Supplementary figure 34: HMBC spectrum of myxoprincomide c798 in $\text{DMSO}-d_6$.



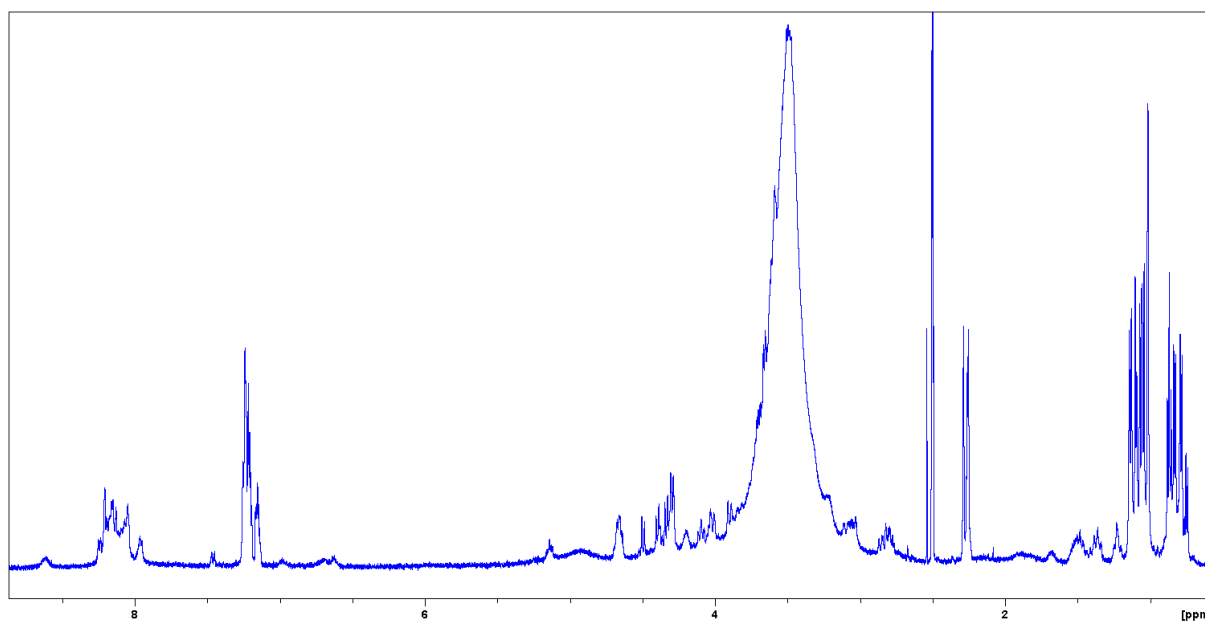
Supplementary figure 35: ^{13}C spectrum of myxoprincomide c798 in $\text{DMSO}-d_6$.



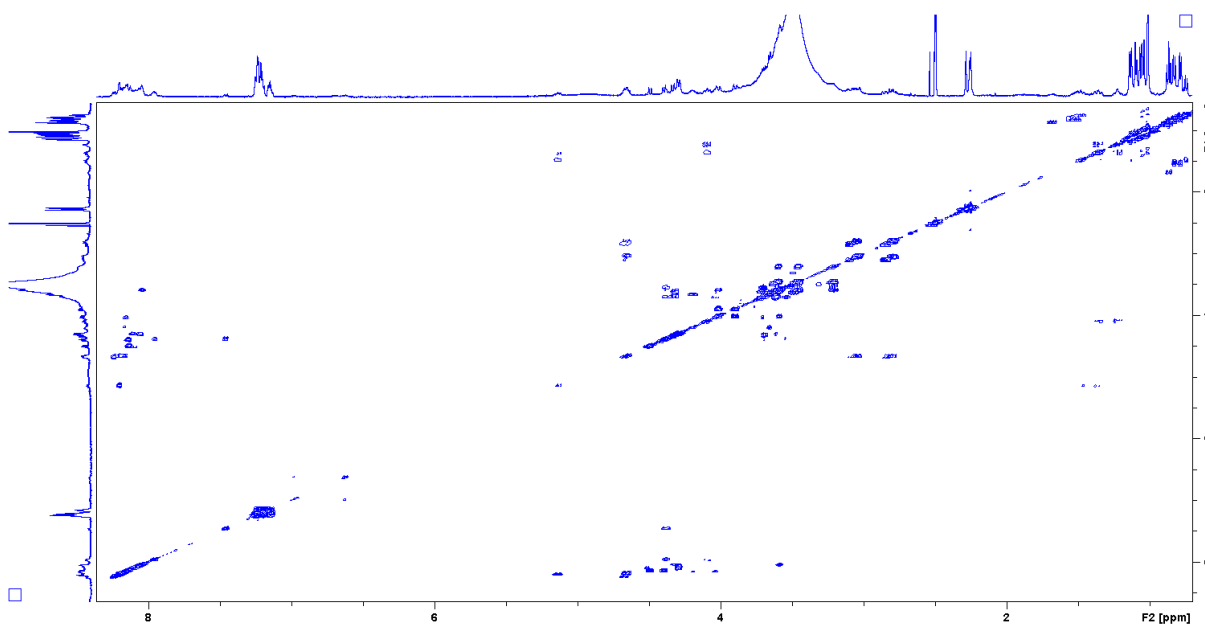
Supplementary figure 36: NOESY spectrum of myxoprincomide c798 in DMSO- d_6 .



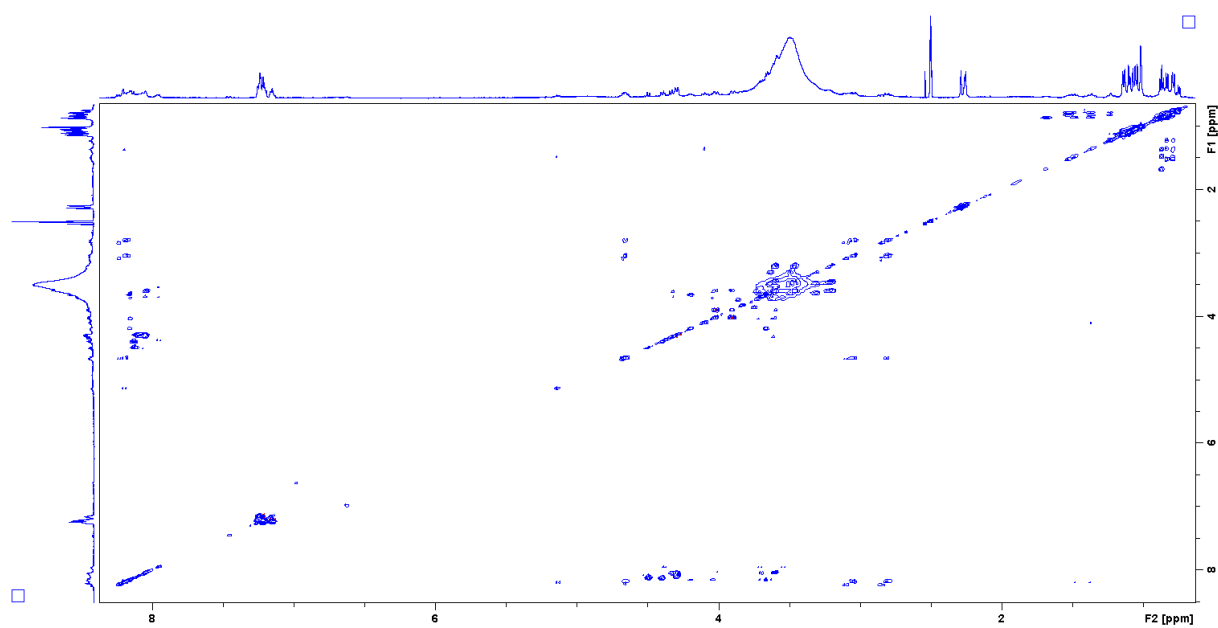
Supplementary figure 37: ROESY spectrum of myxoprincomide c798 in DMSO- d_6 .



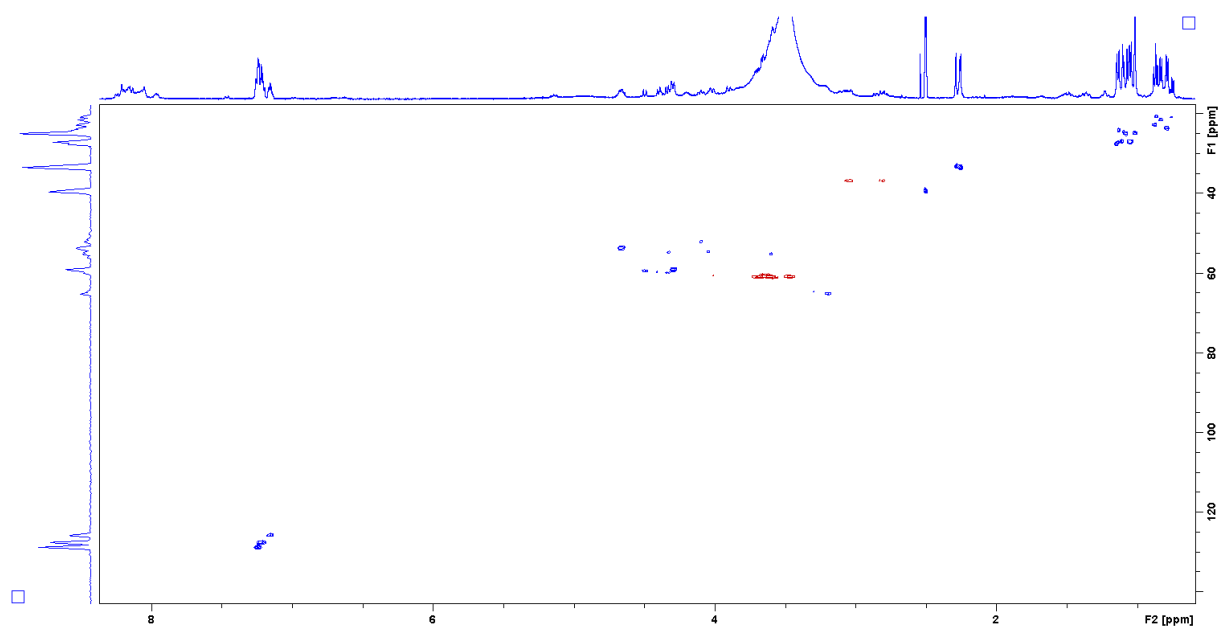
Supplementary figure 38: ^1H NMR spectrum of myxoprincomide c812 in $\text{DMSO}-d_6$.



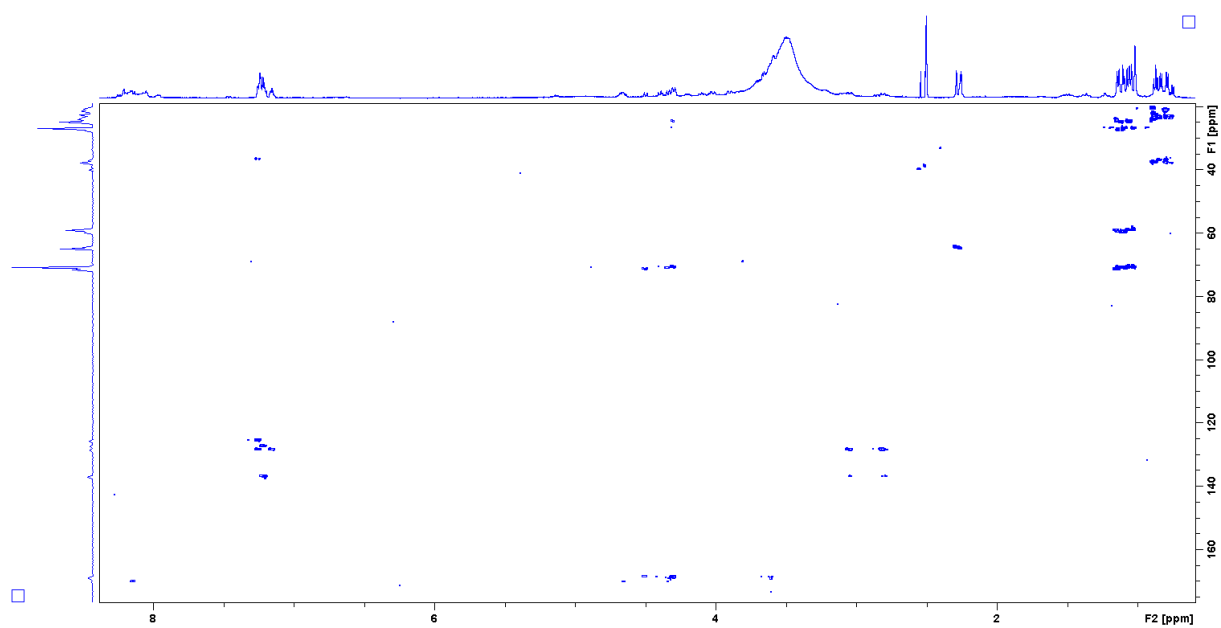
Supplementary figure 39: COSY spectrum of myxoprincomide c812 in $\text{DMSO}-d_6$.



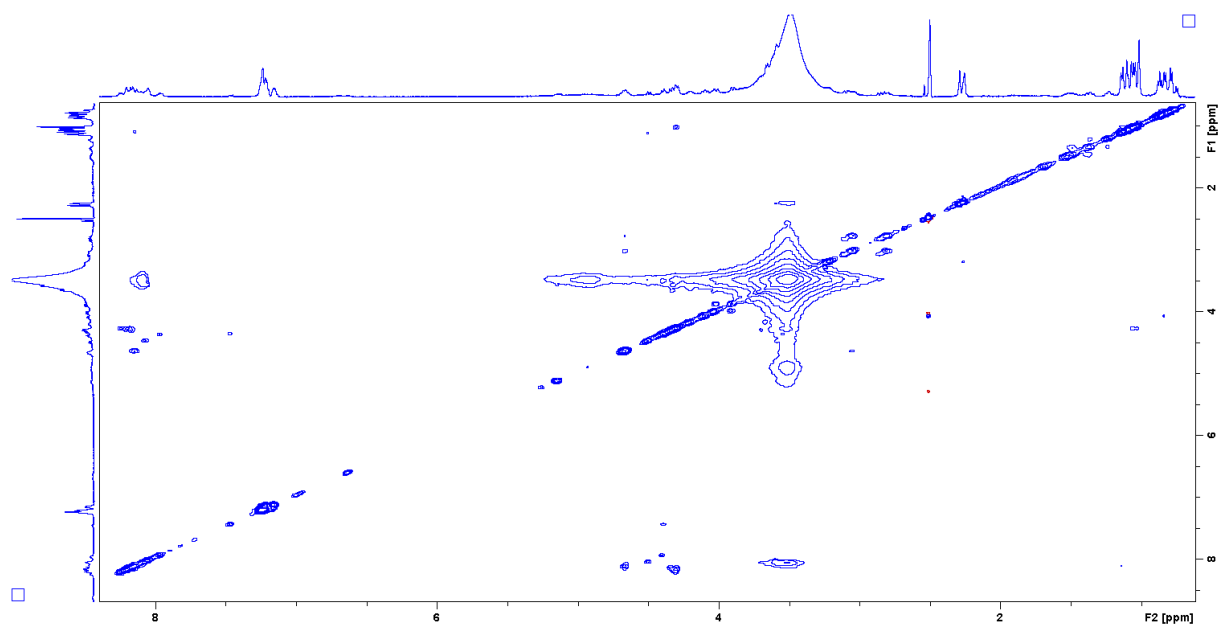
Supplementary figure 40: TOSY spectrum of myxoprincomide c812 in DMSO- d_6 .



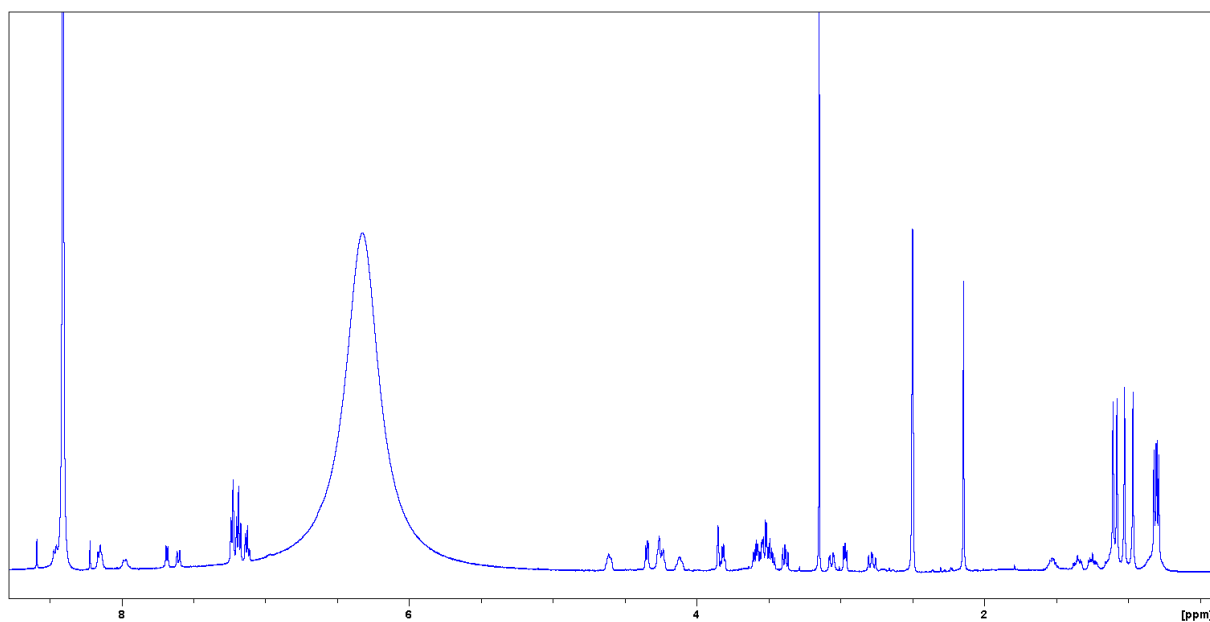
Supplementary figure 41: HSQC spectrum of myxoprincomide c812 in DMSO- d_6 .



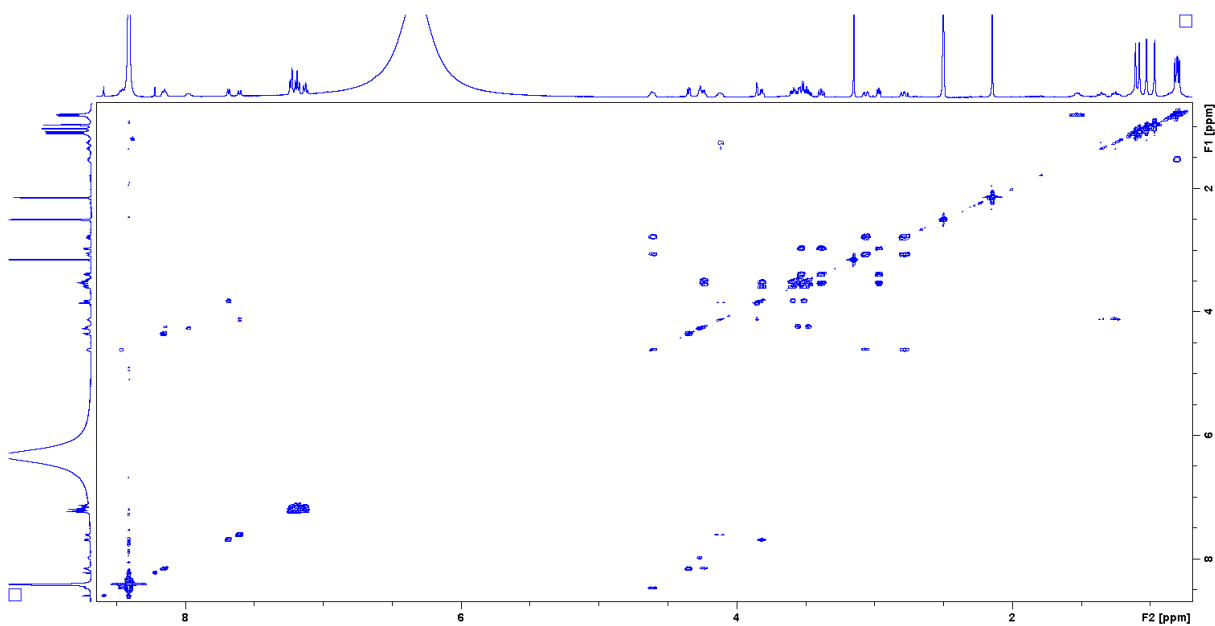
Supplementary figure 42: HMBC spectrum of myxoprincomide c812 in DMSO-*d*₆.



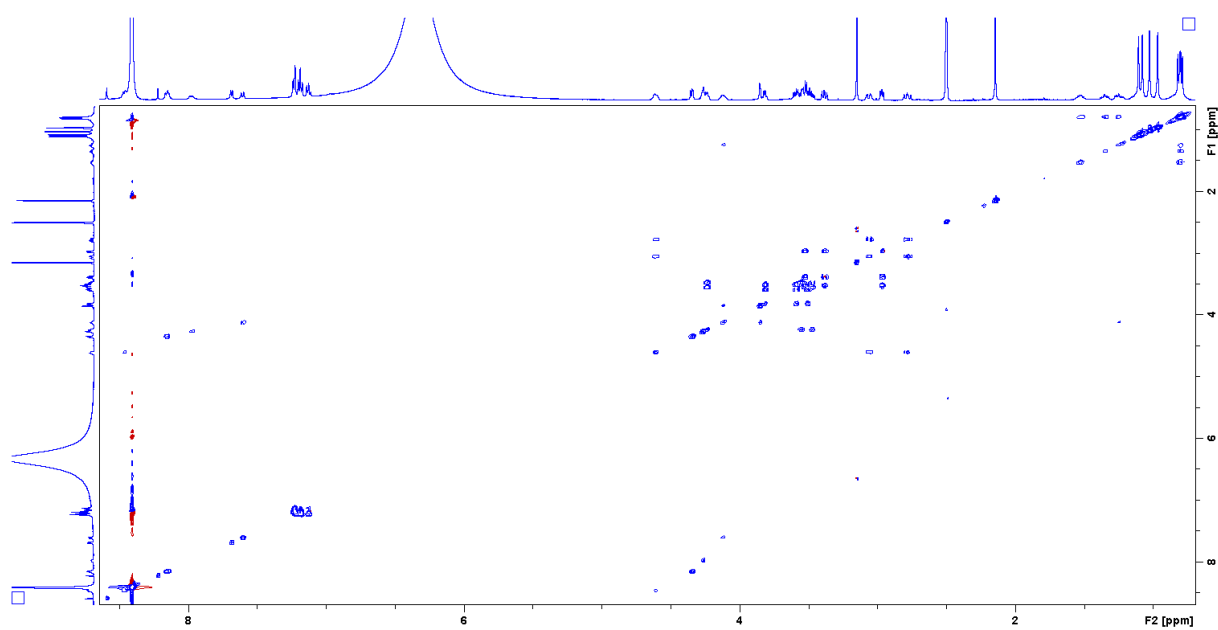
Supplementary figure 43: NOESY spectrum of myxoprincomide c812 in DMSO-*d*₆.



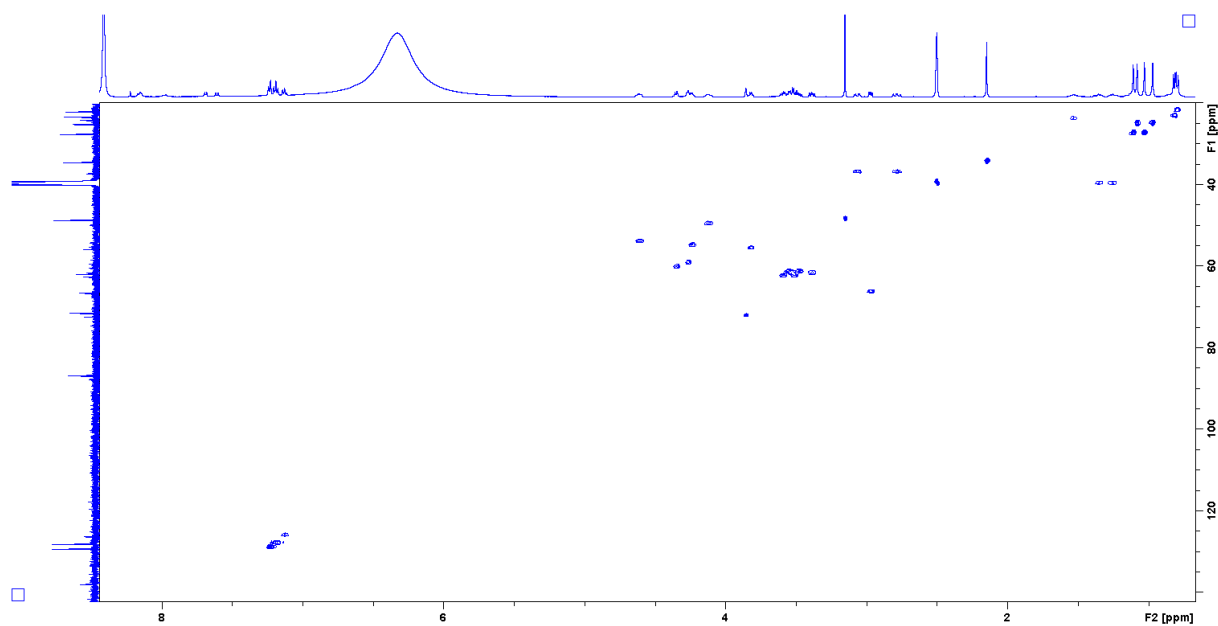
Supplementary figure 44: ^1H NMR spectrum of myxoprincomide c814 in $\text{DMSO}-d_6$.



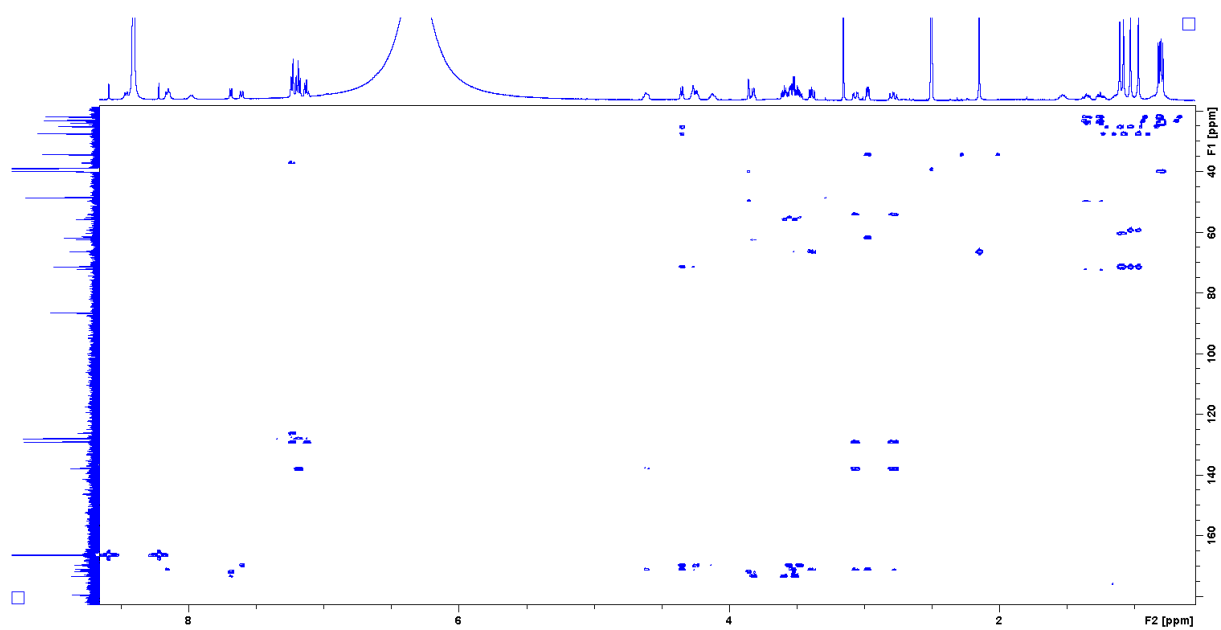
Supplementary figure 45: COSY spectrum of myxoprincomide c814 in $\text{DMSO}-d_6$.



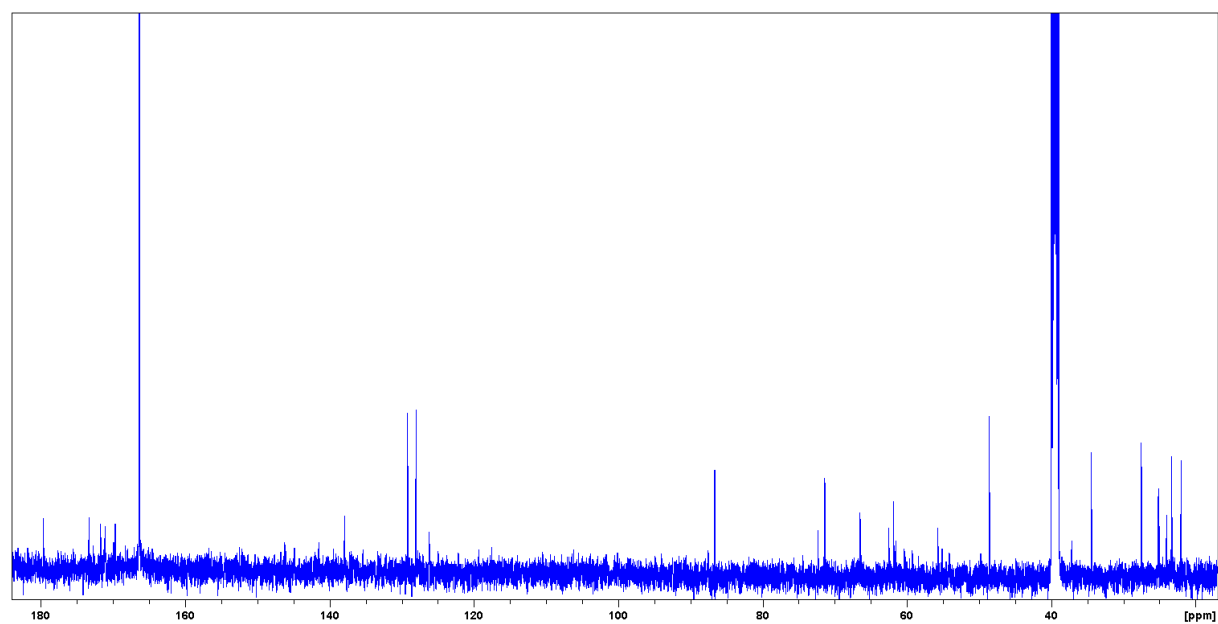
Supplementary figure 46: TOSY spectrum of myxoprincomide c814 in DMSO- d_6 .



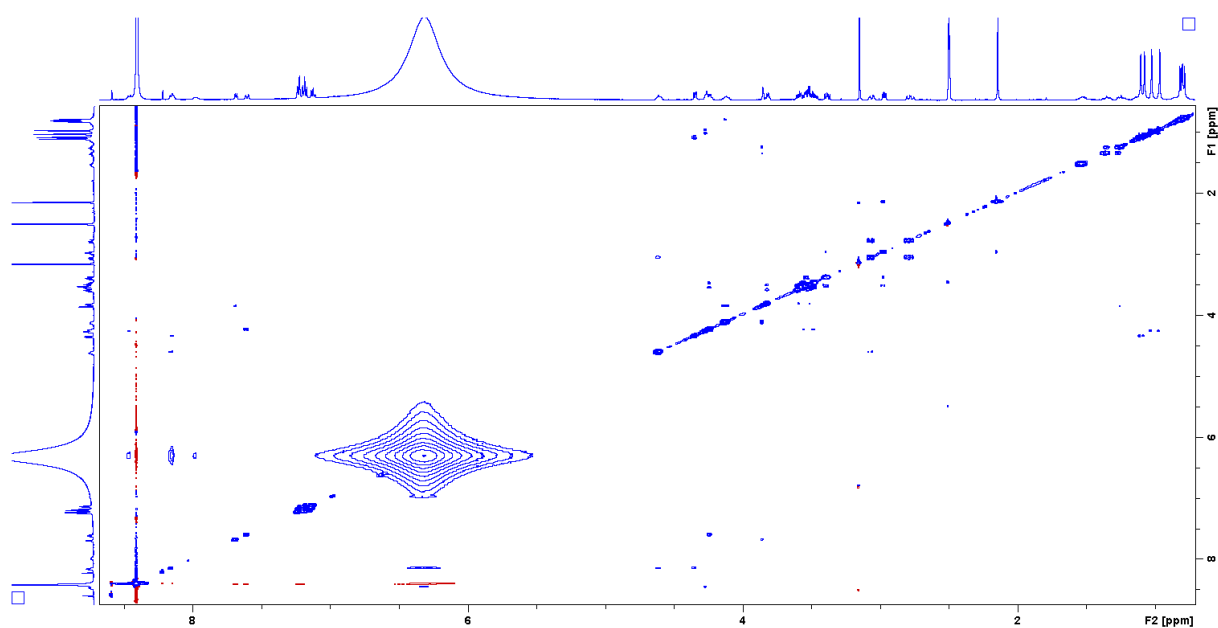
Supplementary figure 47: HSQC spectrum of myxoprincomide c814 in DMSO- d_6 .



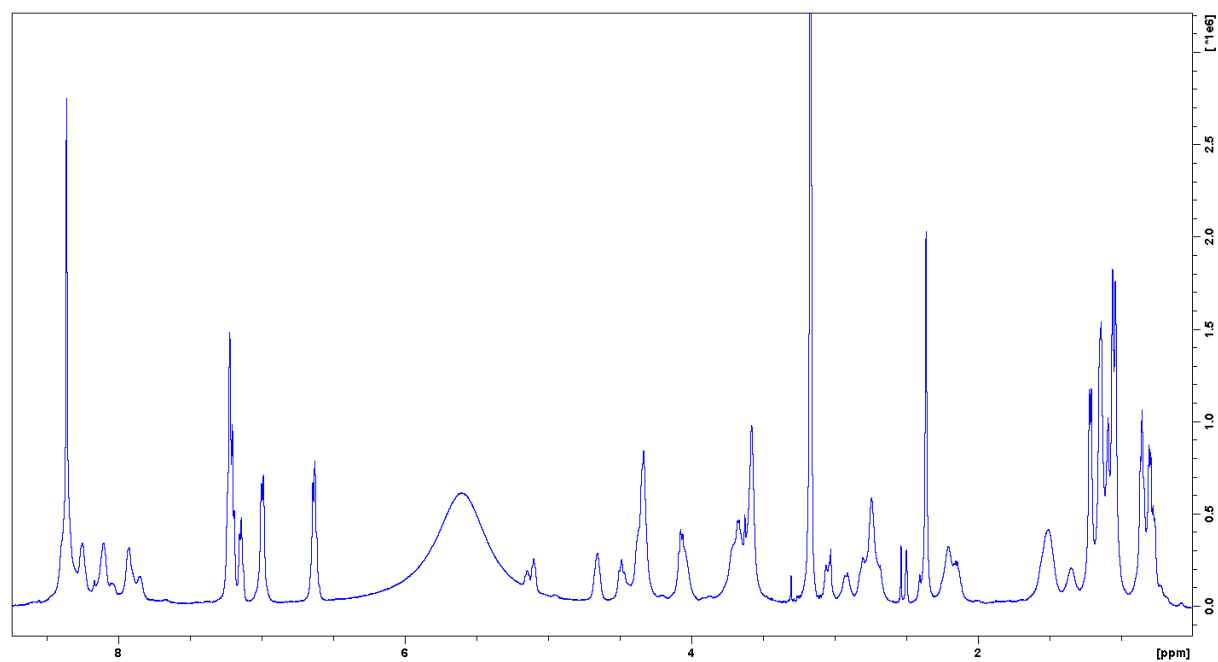
Supplementary figure 48: HMBC spectrum of myxoprincomide c814 in $\text{DMSO-}d_6$.



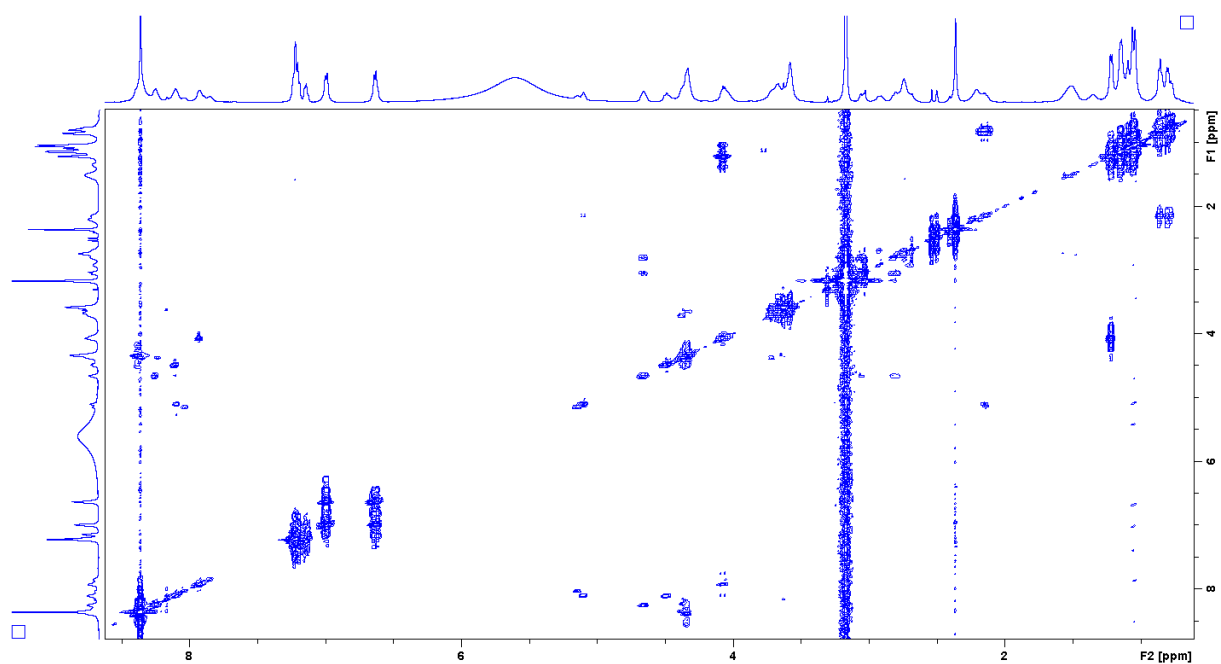
Supplementary figure 49: ^{13}C spectrum of myxoprincomide c814 in $\text{DMSO-}d_6$.



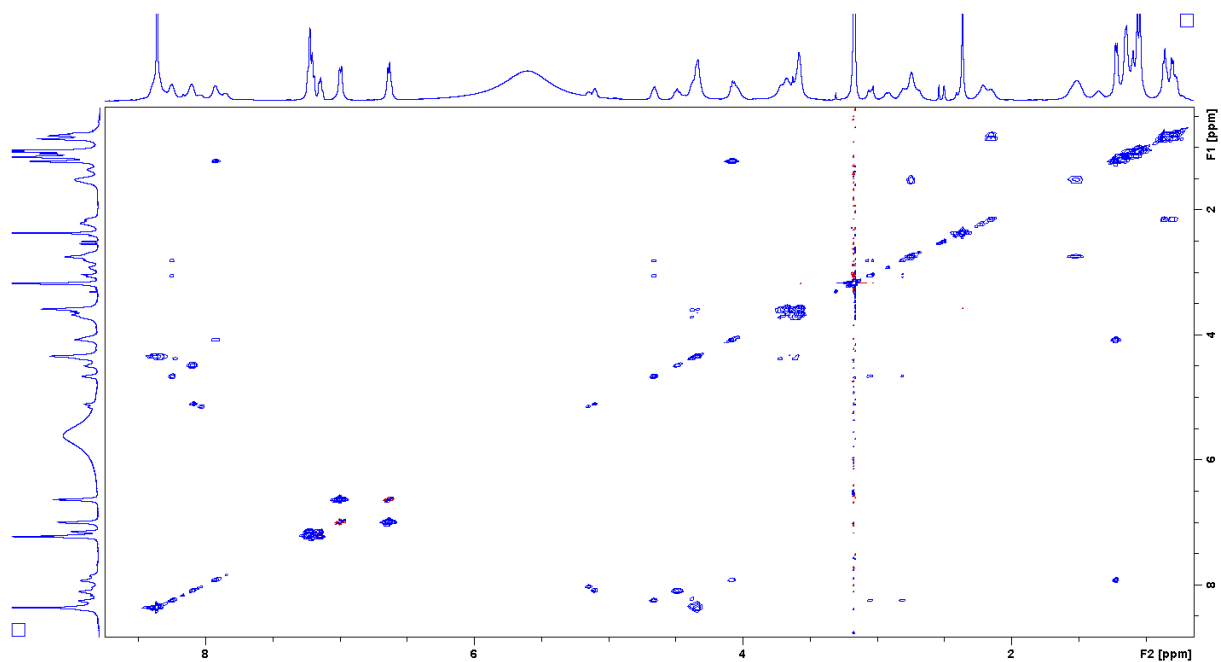
Supplementary figure 50: NOESY spectrum of myxoprincomide c814 in DMSO- d_6 .



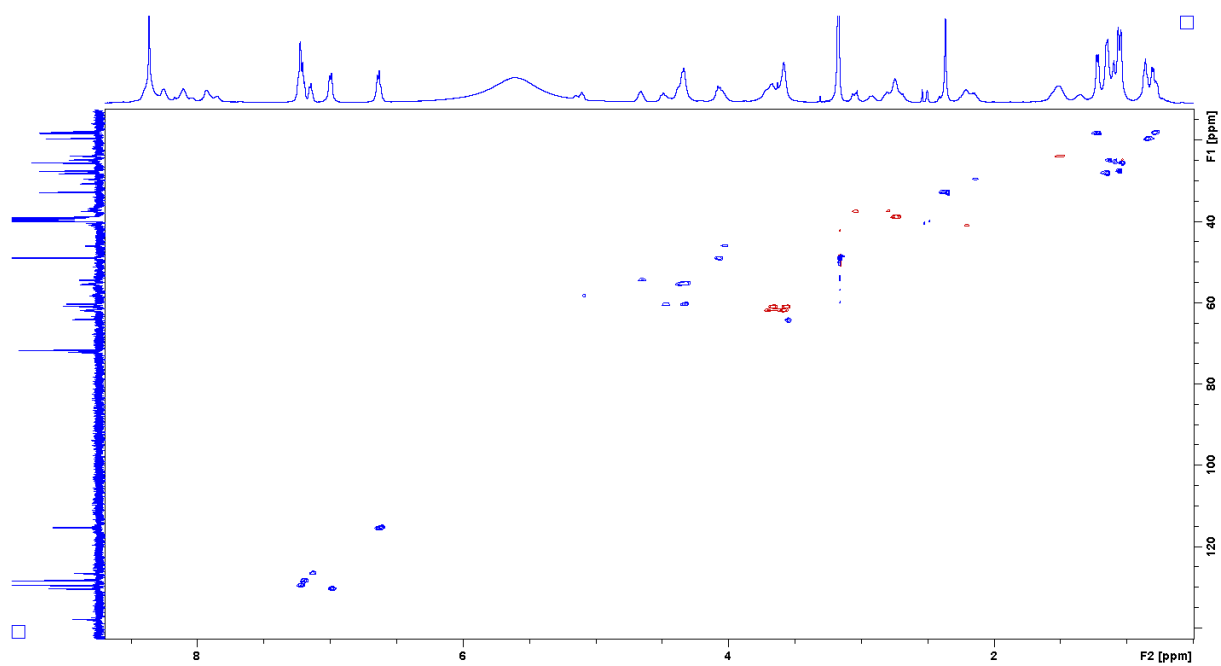
Supplementary figure 51: 1H NMR spectrum of myxoprincomide c581 in DMSO- d_6 .



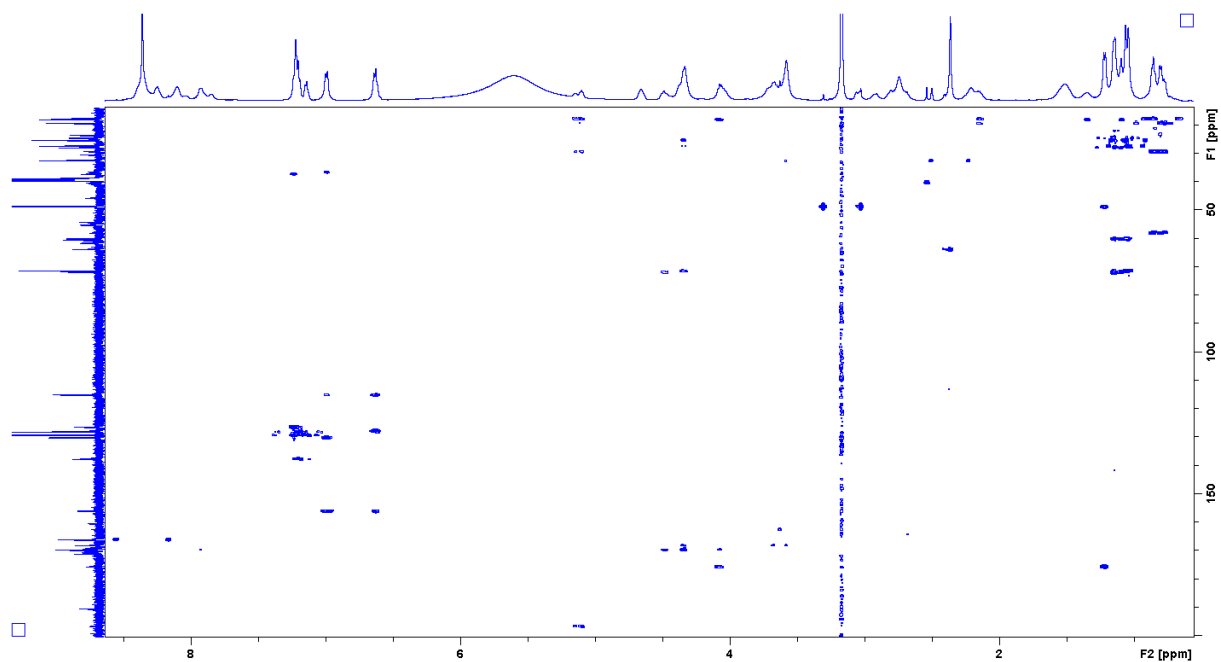
Supplementary figure 52: COSY spectrum of myxoprincomide c581 in DMSO- d_6 .



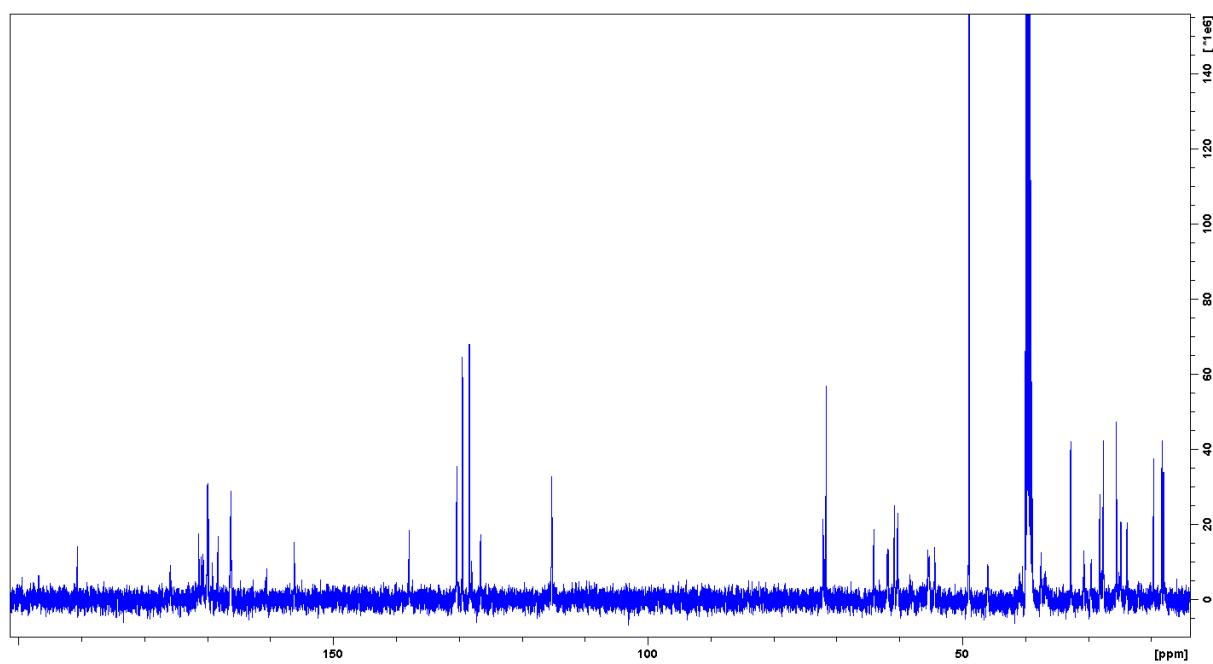
Supplementary figure 53: TOSY spectrum of myxoprincomide c581 in DMSO- d_6 .



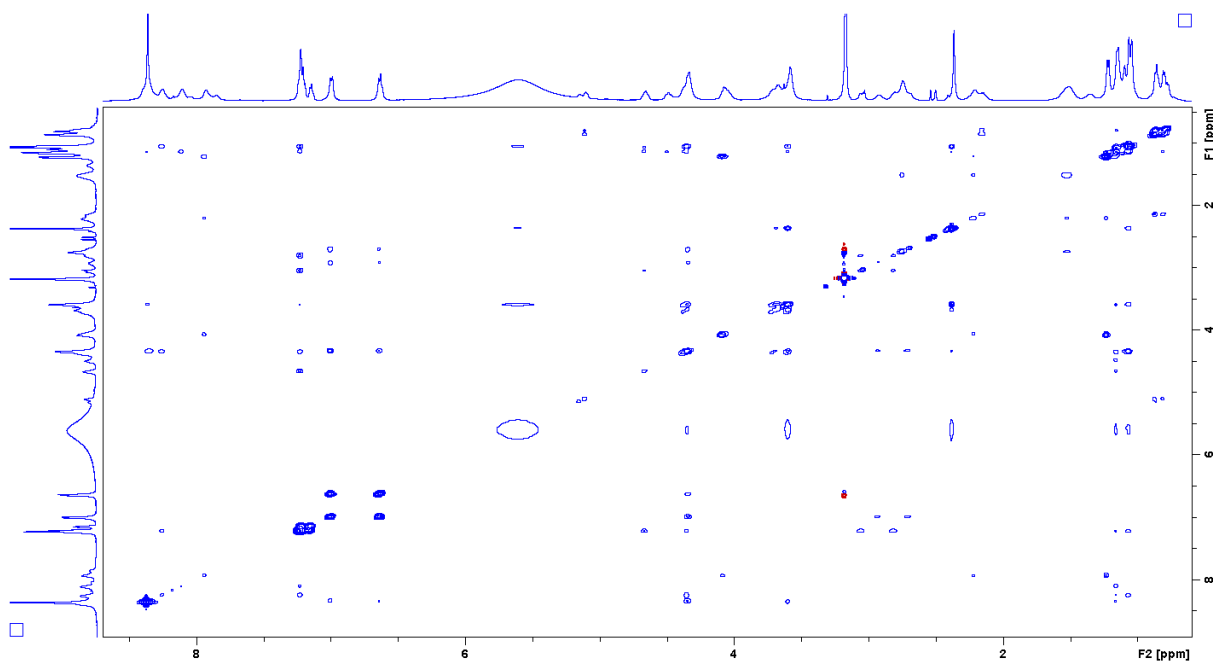
Supplementary figure 54: HSQC spectrum of myxoprincomide c581 in DMSO-*d*₆.



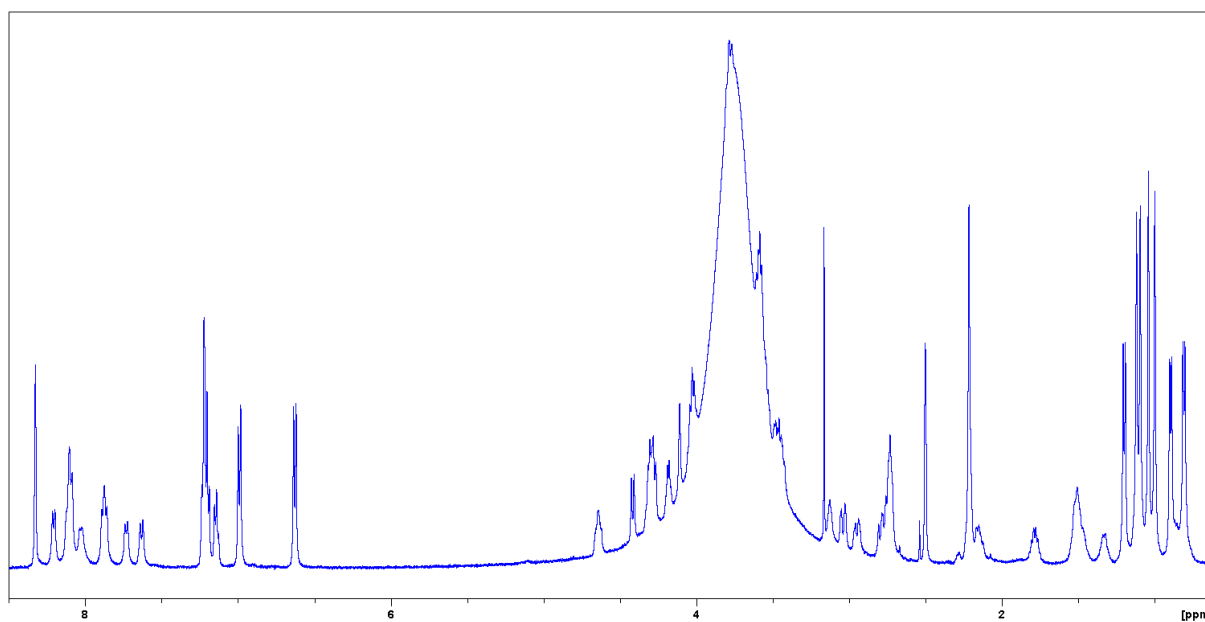
Supplementary figure 55: HMBC spectrum of myxoprincomide c581 in DMSO-*d*₆.



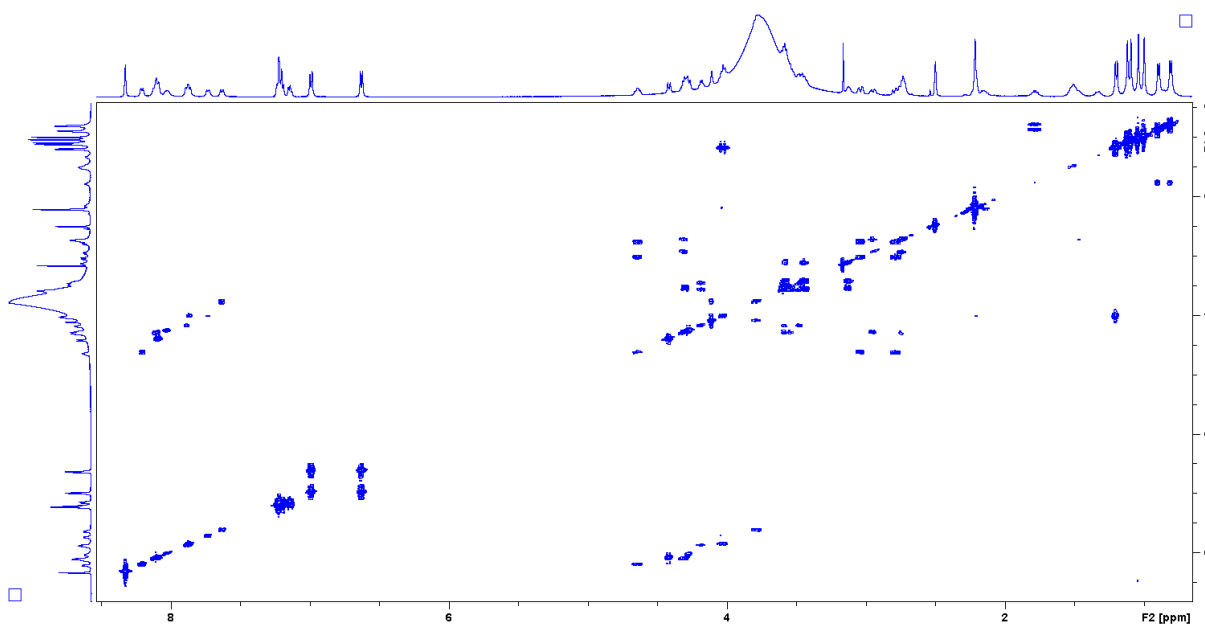
Supplementary figure 56: ^{13}C spectrum of myxoprincomide c581 in $\text{DMSO}-d_6$.



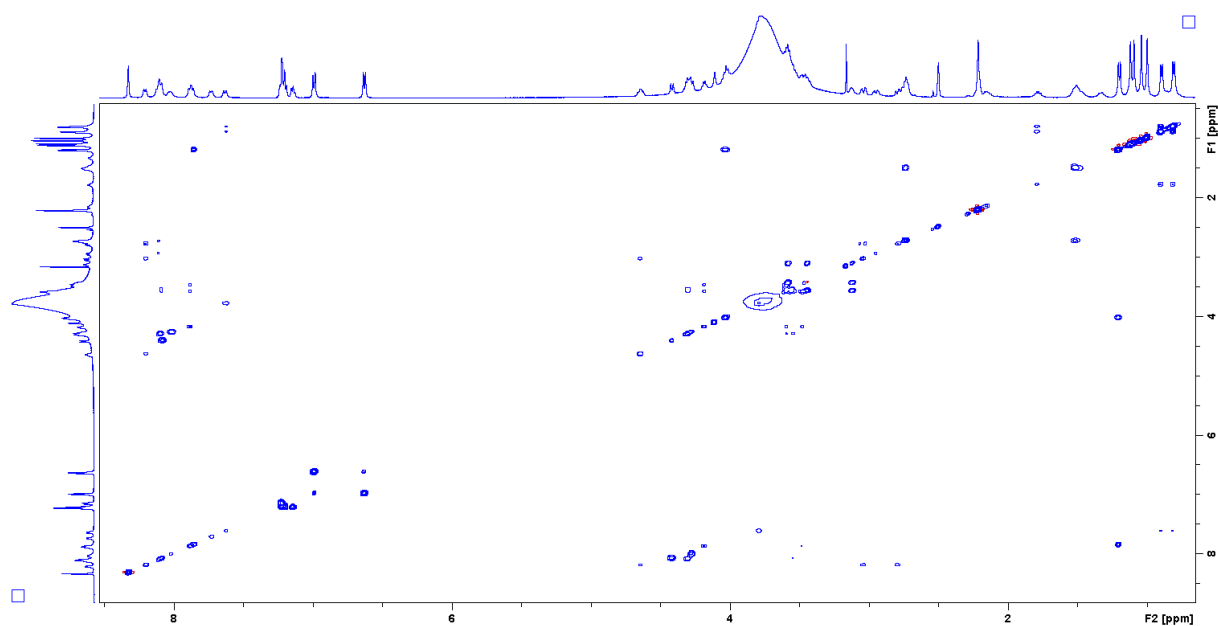
Supplementary figure 57: NOESY spectrum of myxoprincomide c581 in $\text{DMSO}-d_6$.



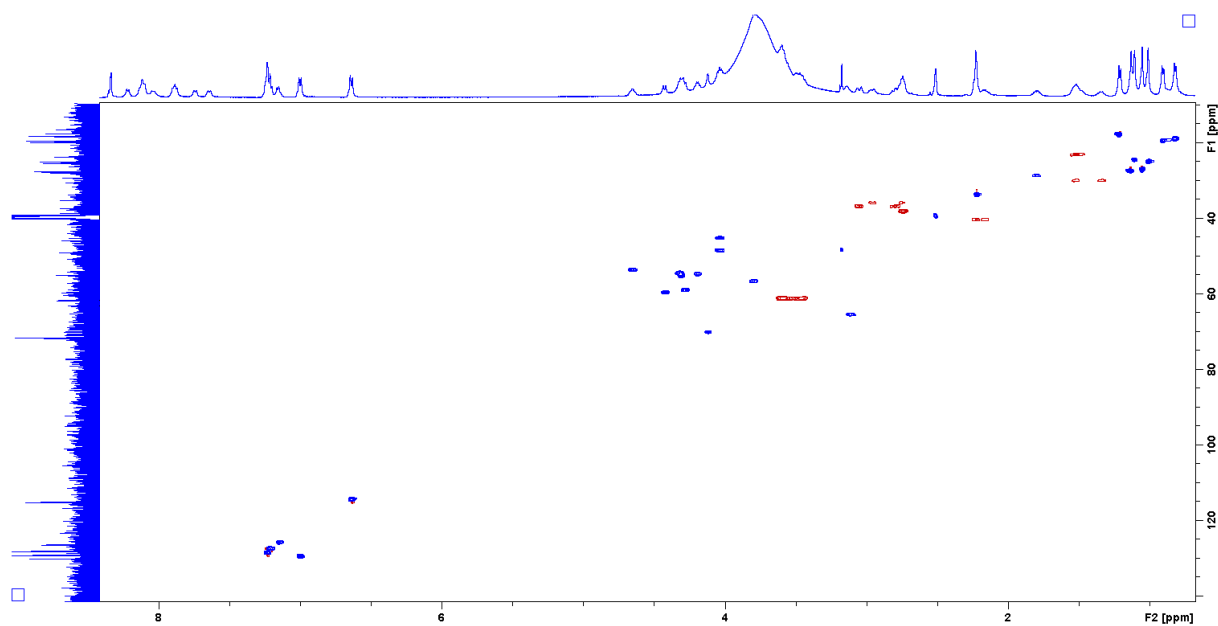
Supplementary figure 58: ^1H NMR spectrum of myxoprincomide c582 in $\text{DMSO}-d_6$.



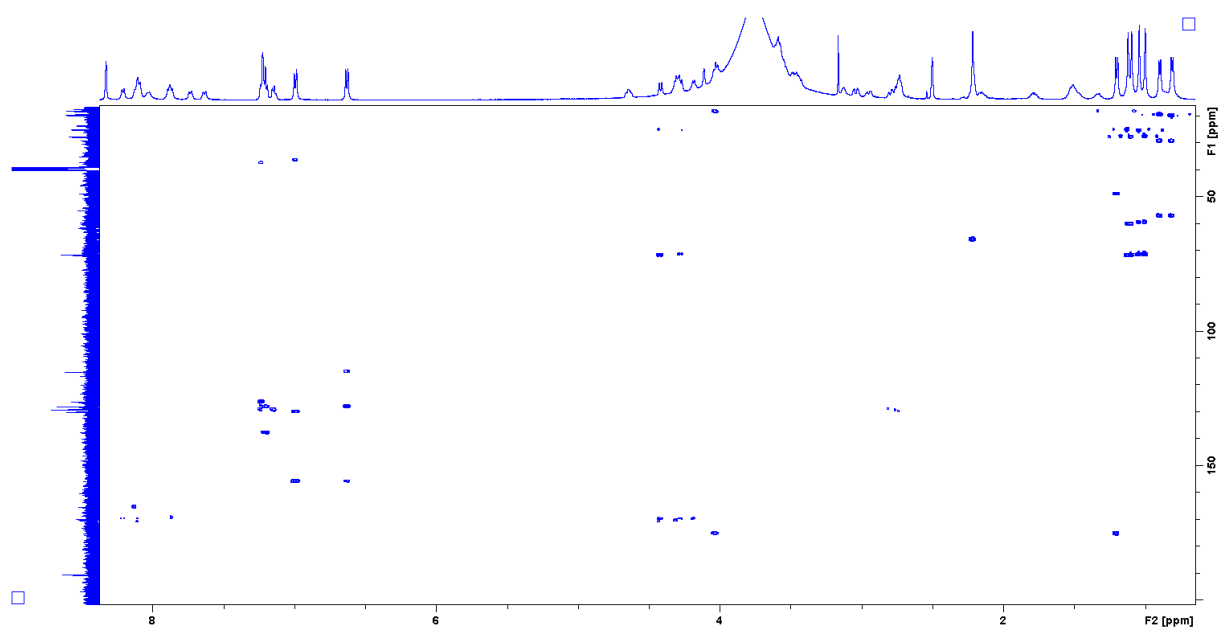
Supplementary figure 59: COSY spectrum of myxoprincomide c582 in $\text{DMSO}-d_6$.



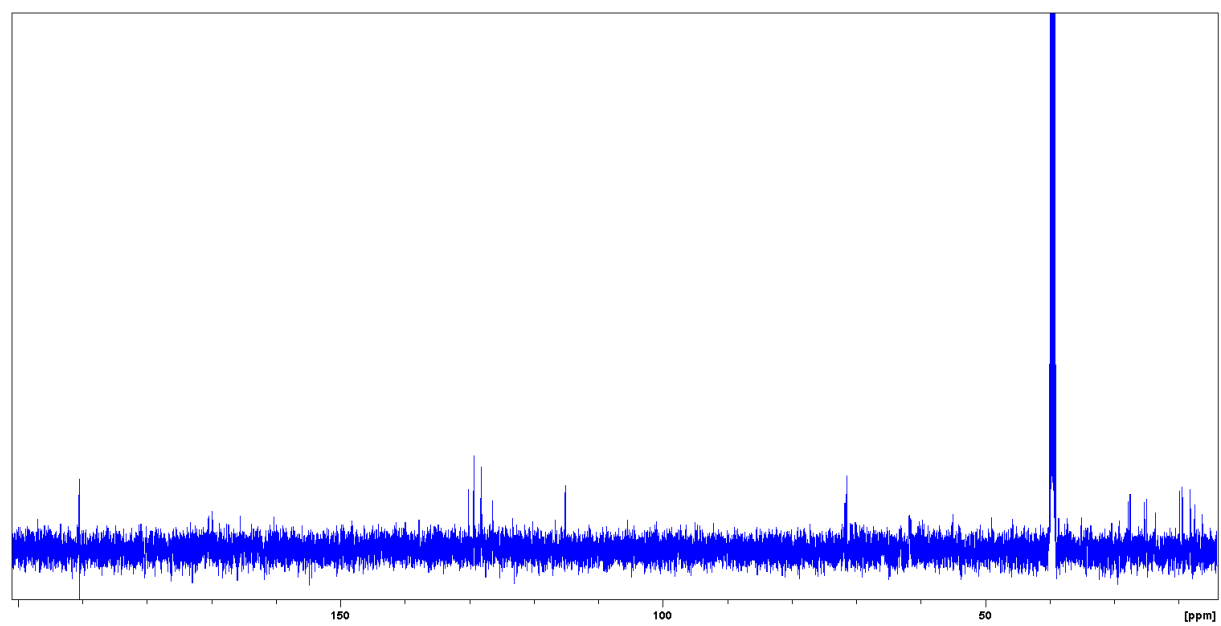
Supplementary figure 60: TOSY spectrum of myxoprincomide c582 in DMSO- d_6 .



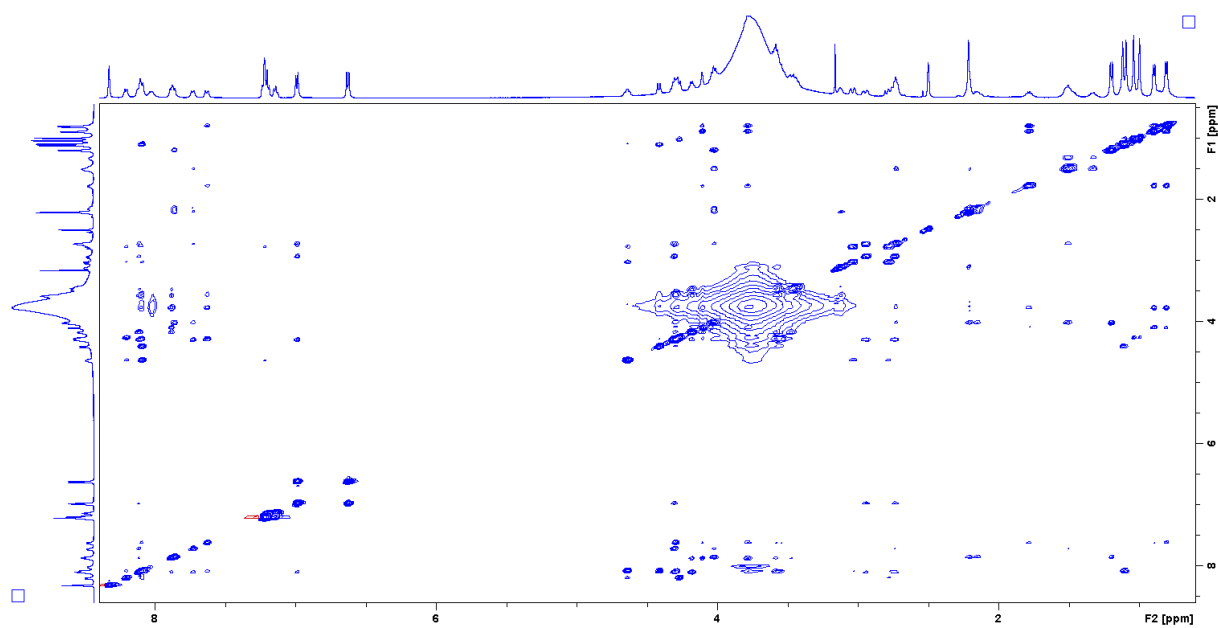
Supplementary figure 61: HSQC spectrum of myxoprincomide c582 in DMSO- d_6 .



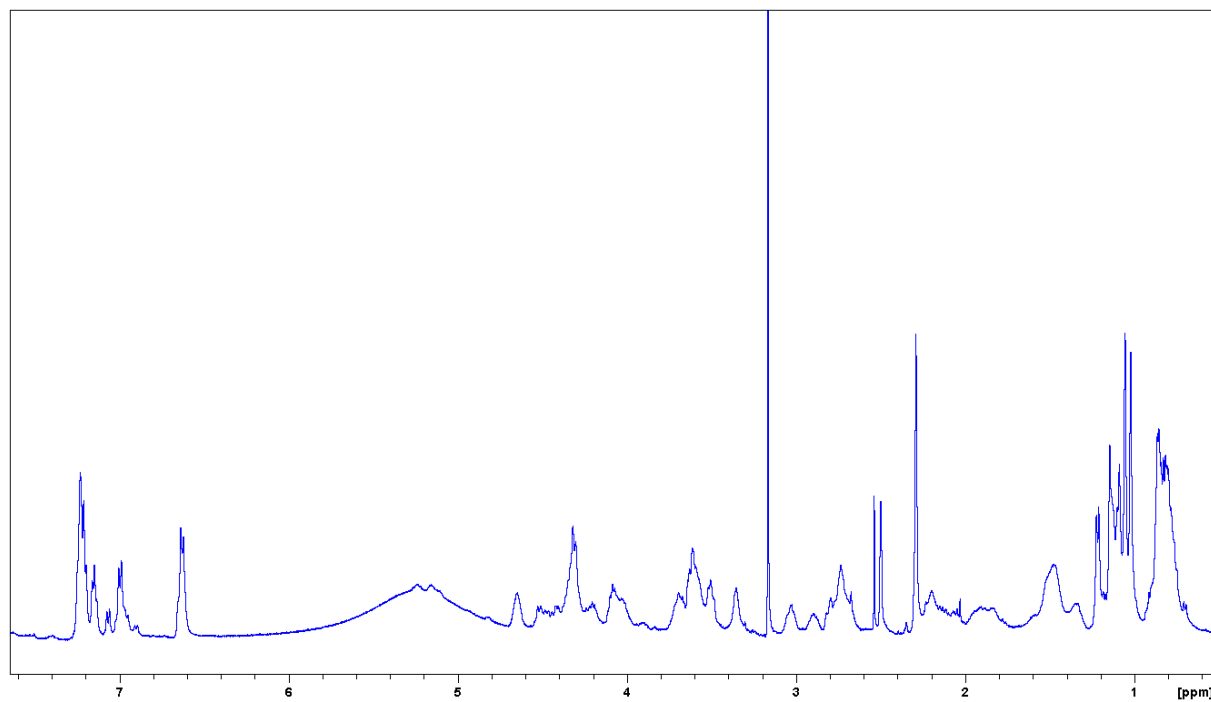
Supplementary figure 62: HMBC spectrum of myxoprincomide c582 in DMSO-*d*₆.



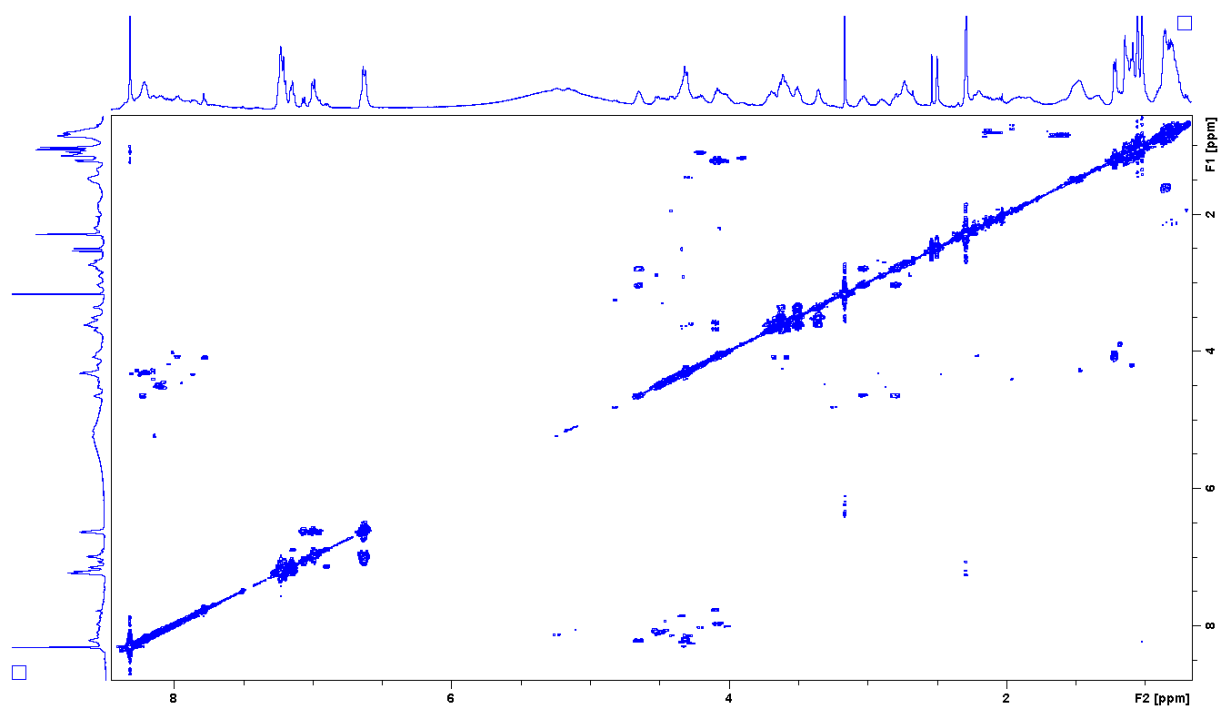
Supplementary figure 63: ¹³C spectrum of myxoprincomide c582 in DMSO-*d*₆.



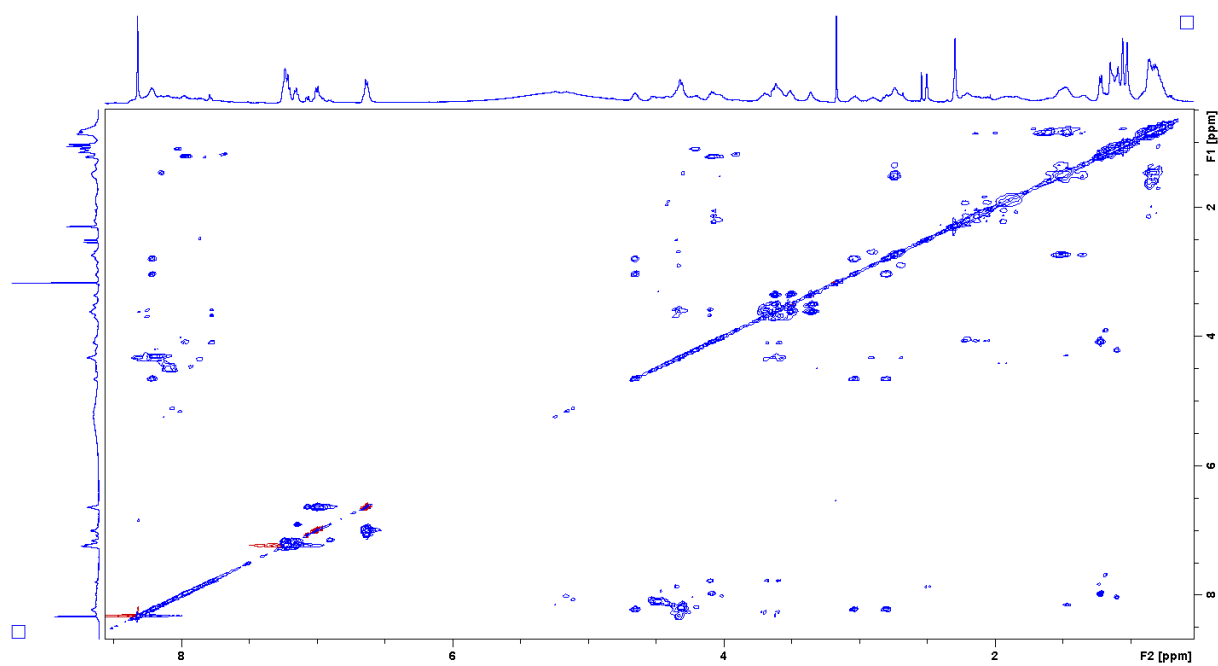
Supplementary figure 64: NOESY spectrum of myxoprincomide c582 in DMSO- d_6 .



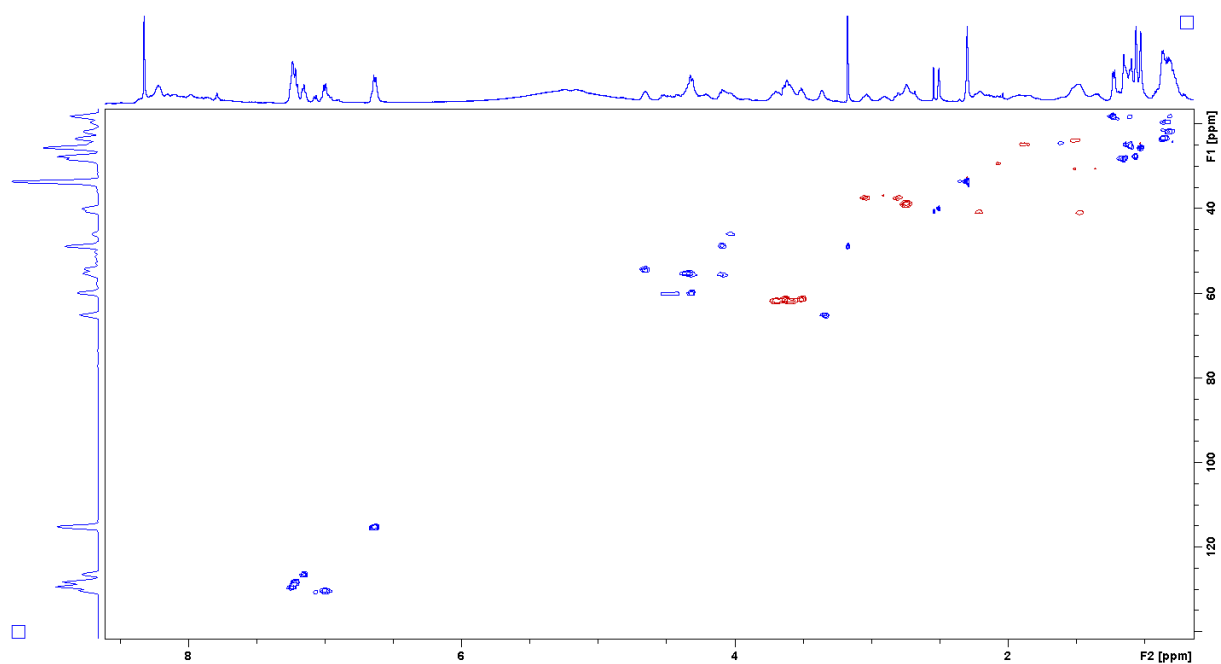
Supplementary figure 65: ^1H NMR spectrum of myxoprincomide c587 in DMSO- d_6 .



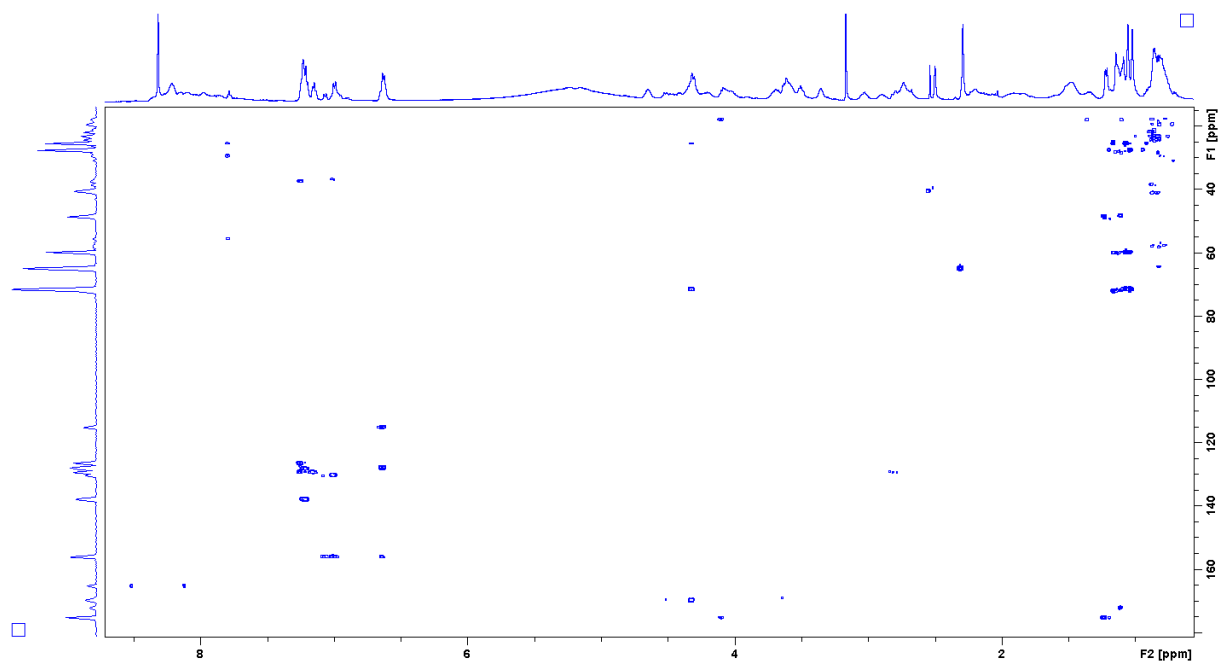
Supplementary figure 66: COSY spectrum of myxoprincomide c587 in DMSO- d_6 .



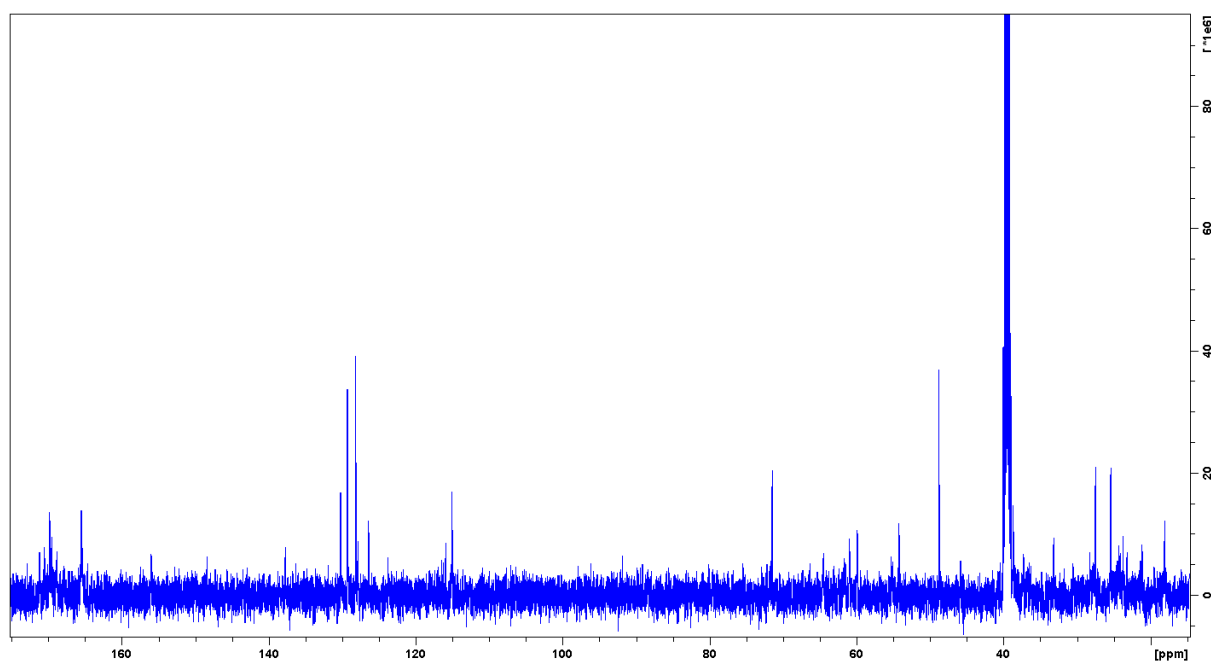
Supplementary figure 67: TOSY spectrum of myxoprincomide c587 in DMSO- d_6 .



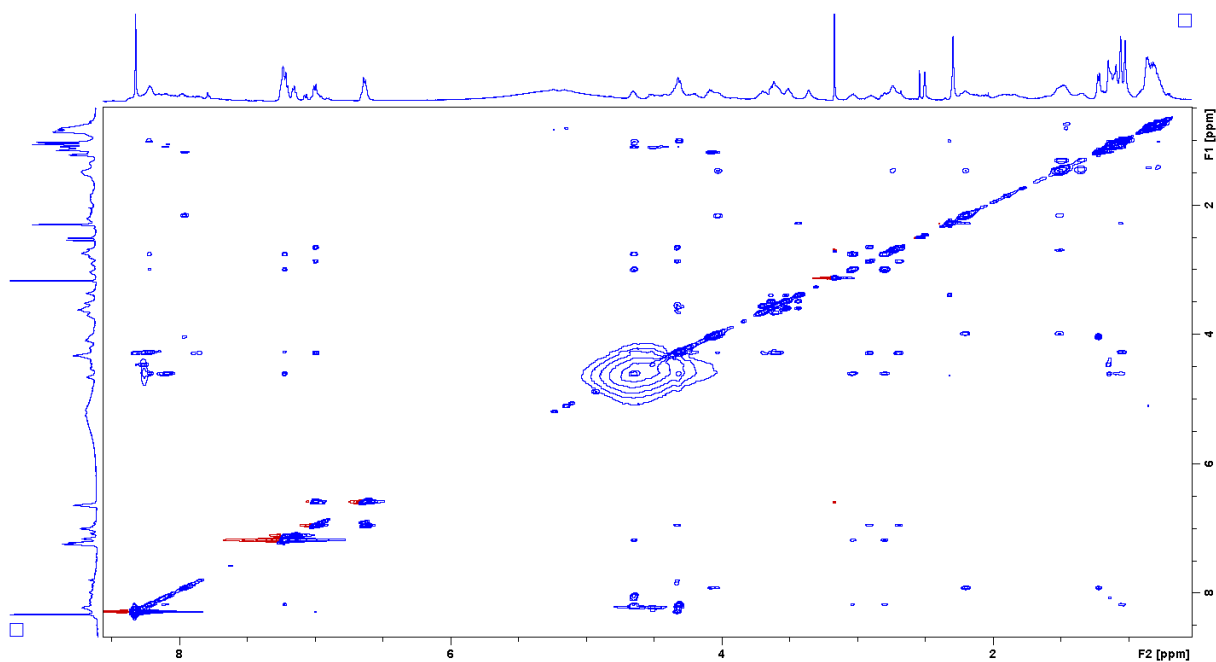
Supplementary figure 68: HSQC spectrum of myxoprincomide c587 in DMSO-*d*₆.



Supplementary figure 69: HMBC spectrum of myxoprincomide c587 in DMSO-*d*₆.



Supplementary figure 70: ^{13}C spectrum of myxoprincomide c587 in $\text{DMSO-}d_6$.



Supplementary figure 71: NOESY spectrum of myxoprincomide c587 in $\text{DMSO-}d_6$.

From the myxoprincomides MS^2 experiments were measured (Supplementary table 10) and the spectra are depicted in Supplementary figure 72. The main fragments of the double

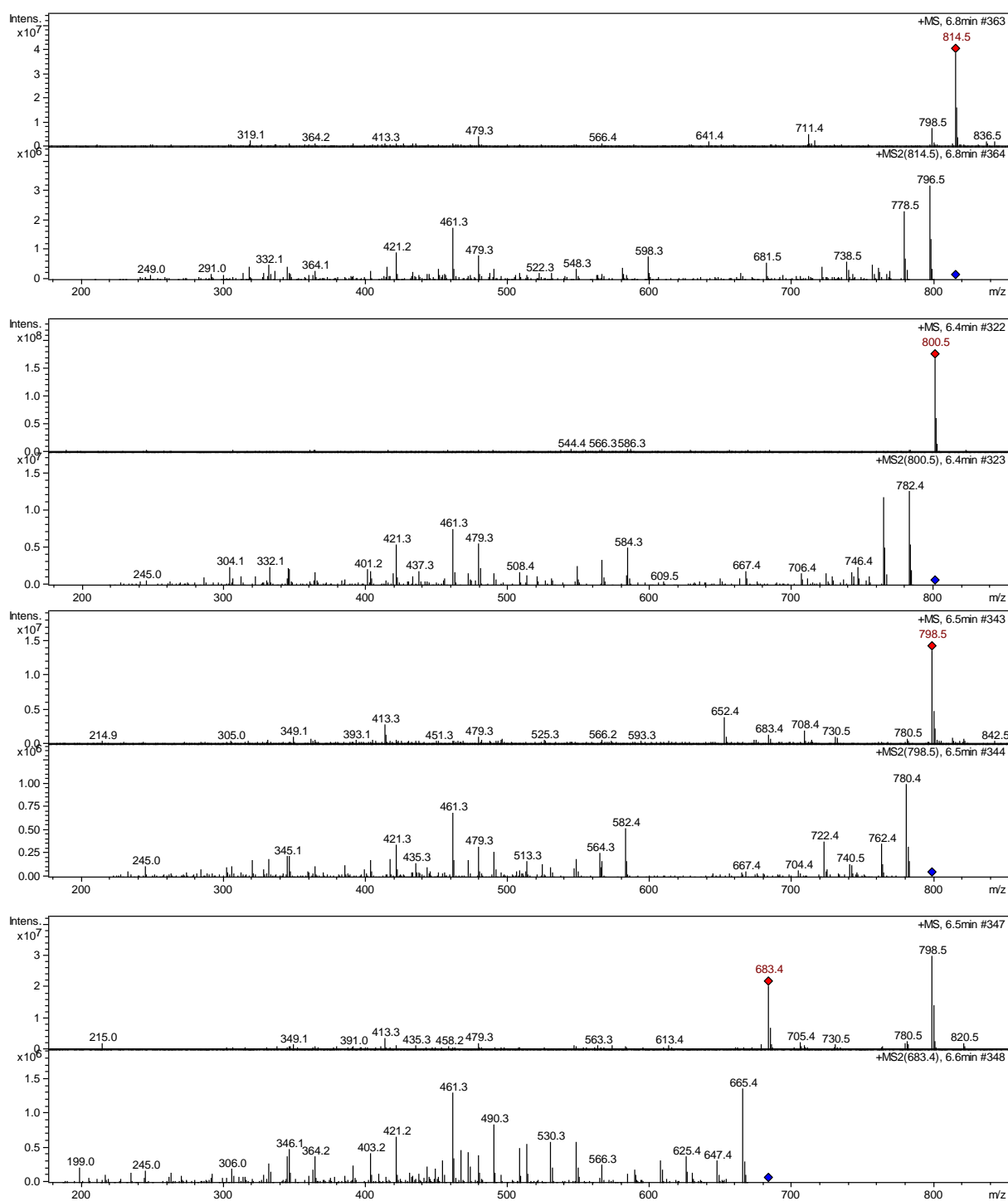
charged compounds are listed in Supplementary table 11 and the spectra are shown in Supplementary figure 73.

c814	c800	c798	c683	c382
814.5	800.5	798.5	683.4	
796.5	782.4	780.4	665.4	364.1
778.5	764.5	762.4	647.4	346.2
768.5	754.4			
756.5	742.4	740.5	625.4	
738.5	724.4	722.4	607.4	
720.6	706.4	704.4	589.3	
681.5	667.4	667.4		
	663.4			
663.4	649.3			
598.3	584.3	582.4	467.3	166
	566.4	564.3	566.3	
580.4	538.4			
566.4	530.3	530.2	530.3	
548.3	520.4	524.3	548.3	
540.4	513.4	513.3	513.3	
522.3	508.4	508.3	508.3	
513.3	503.3			
508.3	496.3		495.2	
499.3	490.3	490.3	490.3	189
495.2	479.3	479.3	479.4	
490.4	472.3	472.3	472.3	171.1
479.3	461.3	461.3	461.3	
473.2	455.3		454.3	
461.3	437.2	443.3	443.3	
451.2	437.3	435.3	437.2	
443.3	433.2		431.2	
433.3	421.1	421.3	421.2	120
421.2	414.3	417.3	409.2	
	403.2	403.3		
415.3	401.2	399.2		
403.2	394.1		403.2	
391.2	385.3	385.2	391.2	
373.1	364.2	364.1	364.2	
364.1	345.1	345.1	345.2	
345.1	332.1	332.1	332.1	
332.1	322.2	320.1	328.1	
	312	306.1	306	
318.1	304.1	302	292.1	
300.1	286	284	263.1	
291	279			291.1
	274			281.1
283	269.3			
273.4	262			263
259.1	245	245	245	
	240.8		234.9	
249	235.1	233	217	217
241.1	227		199	199

Supplementary table 10: MS2 comparison of 814, 800, 798, 683, and 382.

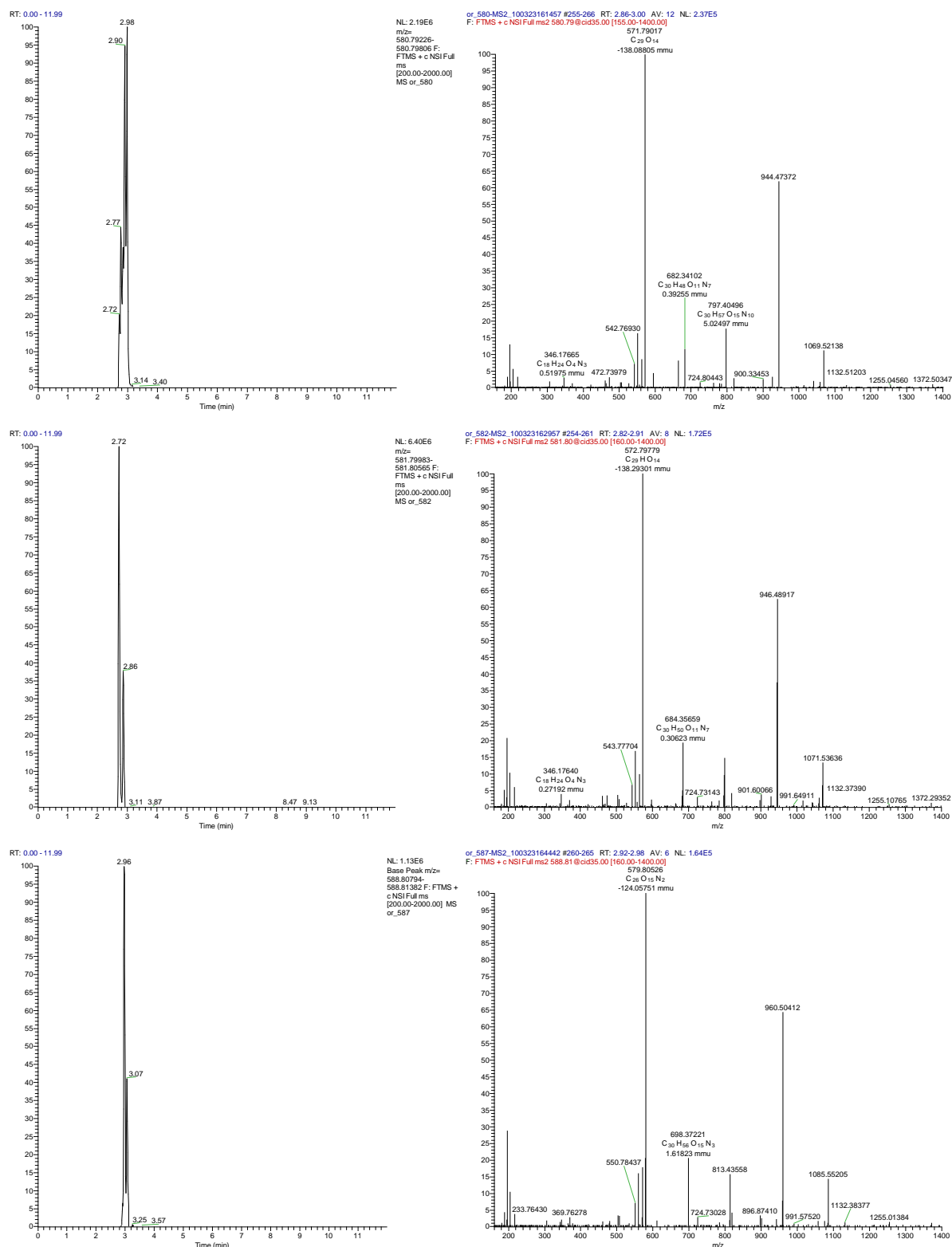
c580	c582	c587
944.4	946.5	960.5
797.4	799.4	813.4
682.3	684.3	698.3
571.8	572.8	579.8
551.7	552.8	559.8
472.7	473.8	480.7
346.2	346.2	346.2
306.1	306.1	306.1

Supplementary table 11: Fragmentation pattern of C580, c582, c587, blue = double charged.



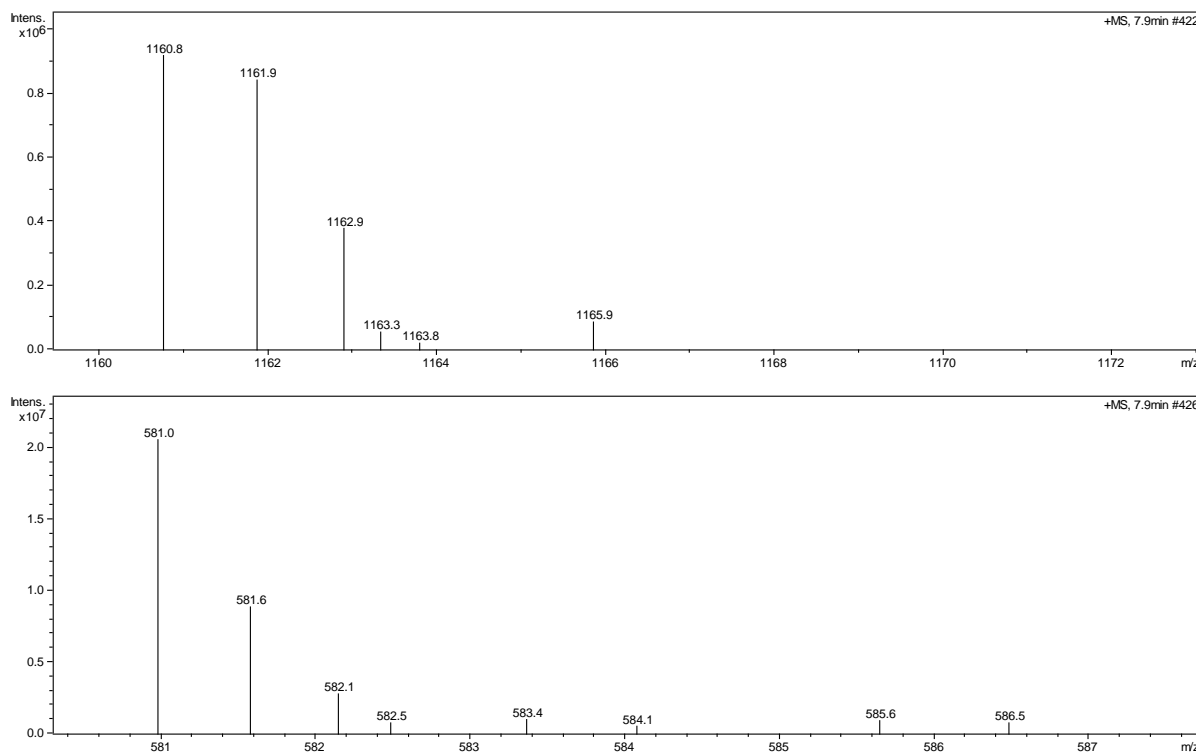
Supplementary figure 72: MS and MS² spectra of c814, c800, c798 and c683.

Chapter 2A: DK897 Myxoprincomides

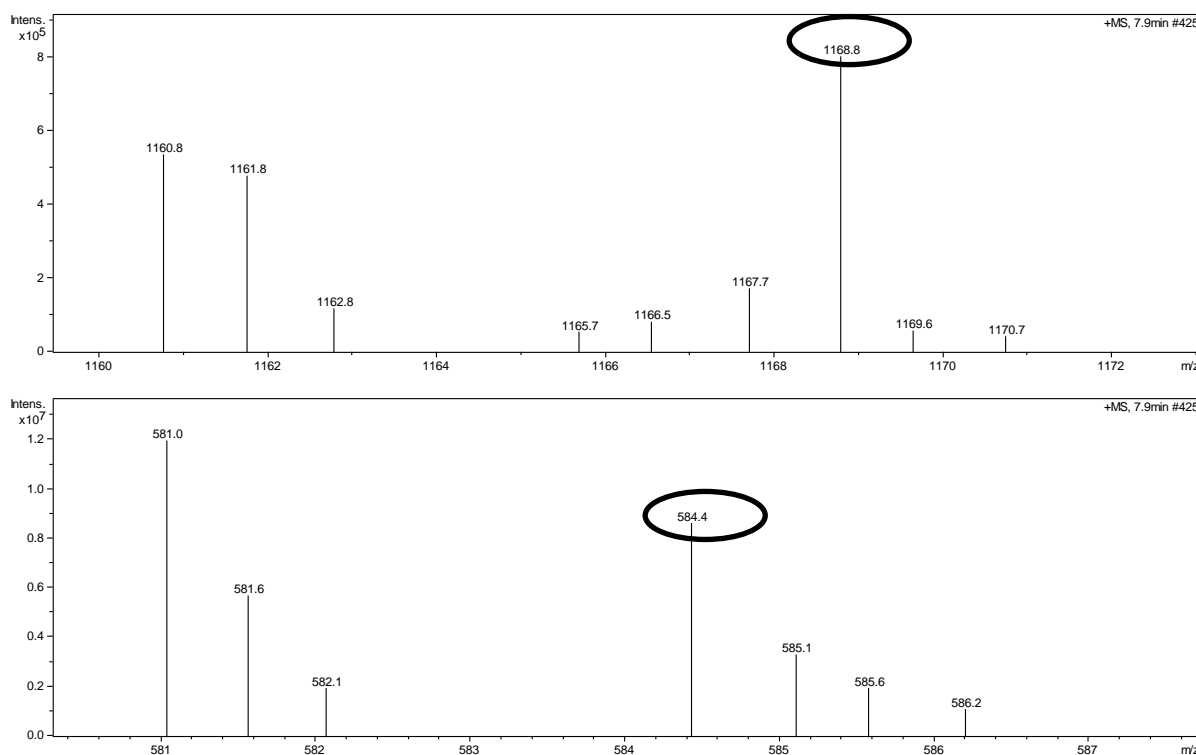


Supplementary figure 73: Chromatogram and fragmentation pattern of c581, c582 and c587.

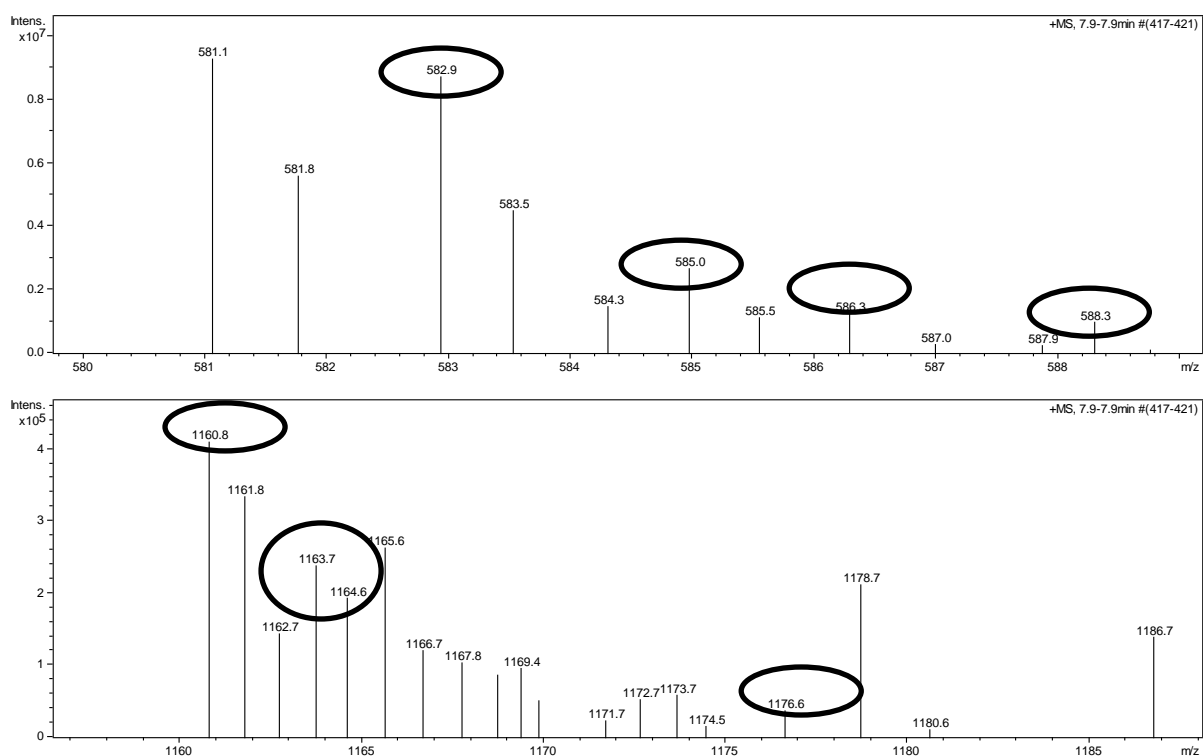
MS spectra of feeding studies are displayed in Supplementary figure 74 to Supplementary figure 78.



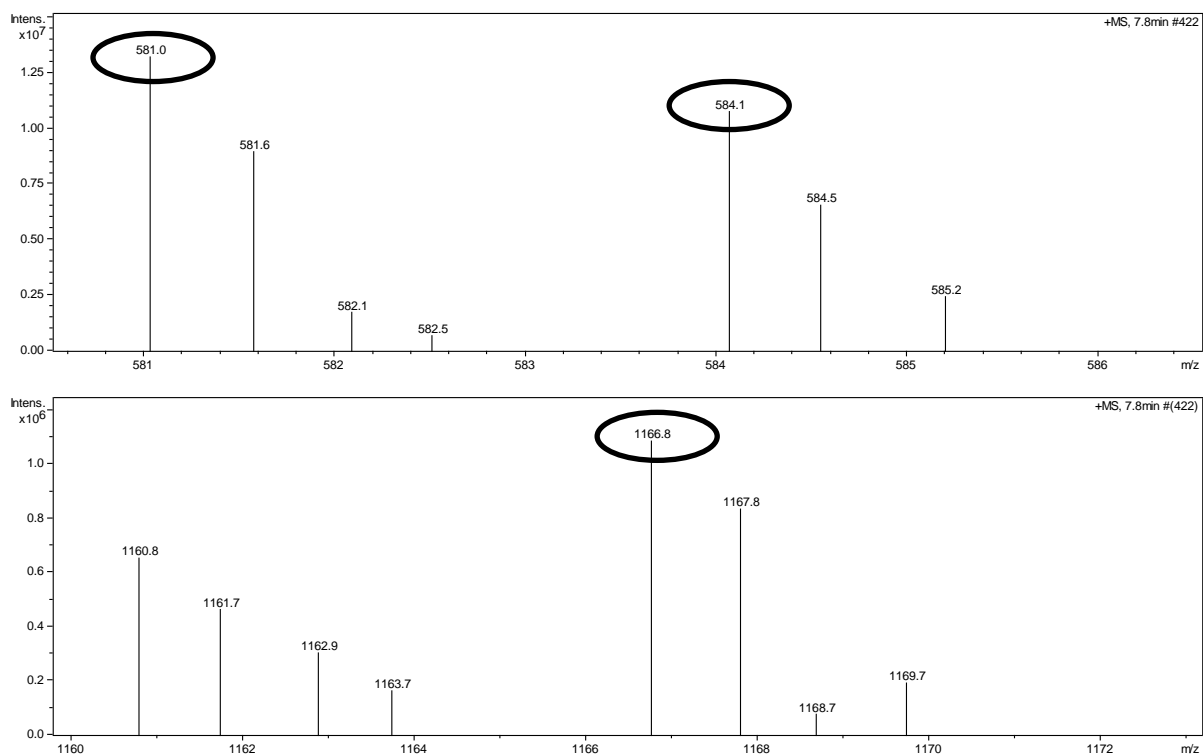
Supplementary figure 74: Isotopic pattern of c580 without labeled amino acid feeding.



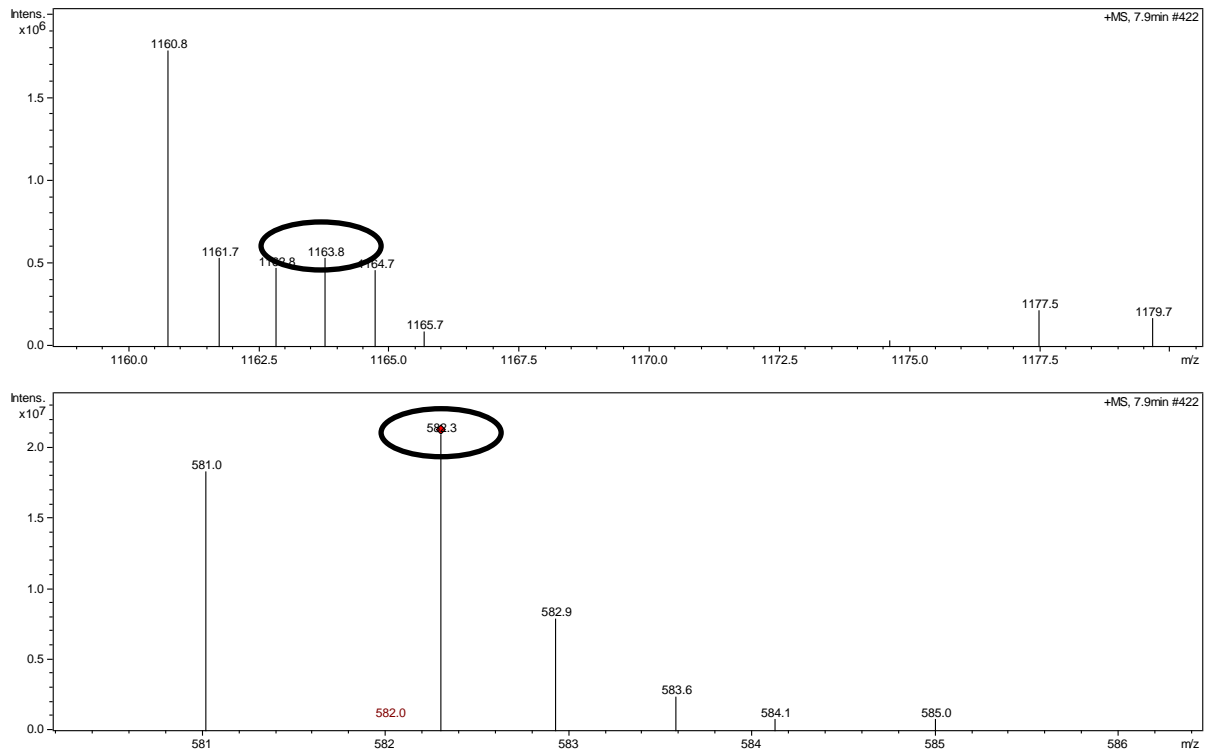
Supplementary figure 75: Isotopic pattern of c580 after feeding with isotopic labeled valine.



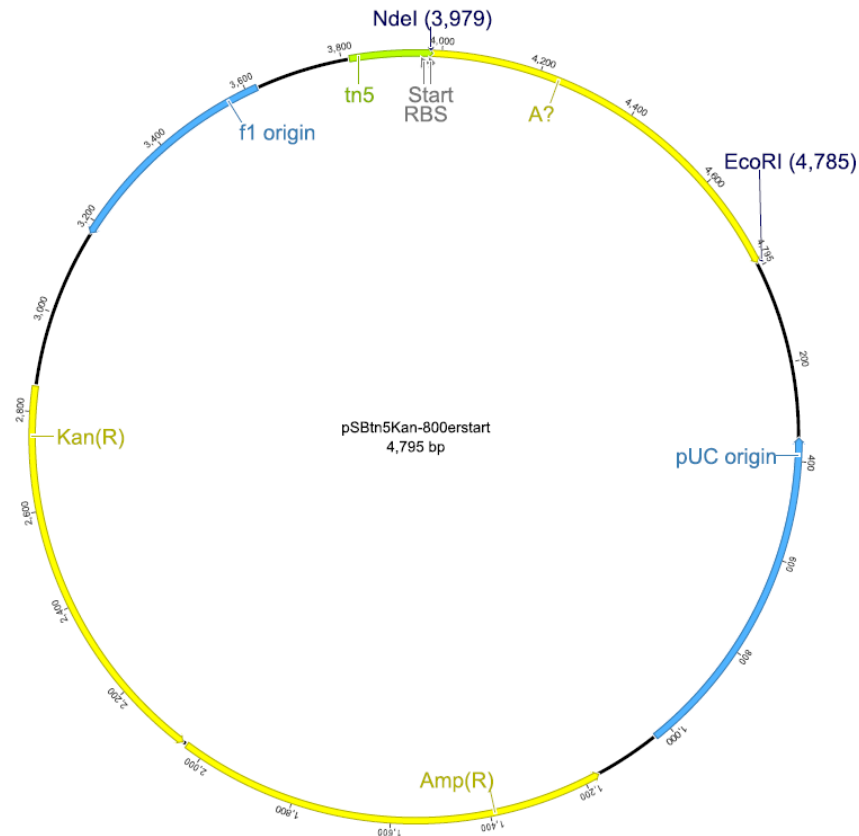
Supplementary figure 76: Isotopic pattern of c580 after feeding with isotopic labeled serine.



Supplementary figure 77: Isotopic pattern of c580 after feeding with isotopic labeled phenylalanine.



Supplementary figure 78: Isotopic pattern of c580 after feeding with isotopic labeled methionine.

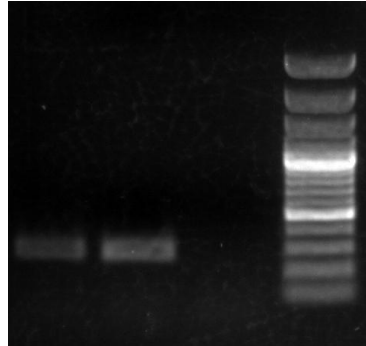


Supplementary figure 79: pSBtn5Kan-800erstart plasmid for promoter activation of the 800er cluster.

The obtained clones were analyzed by colony PCR with following primers:

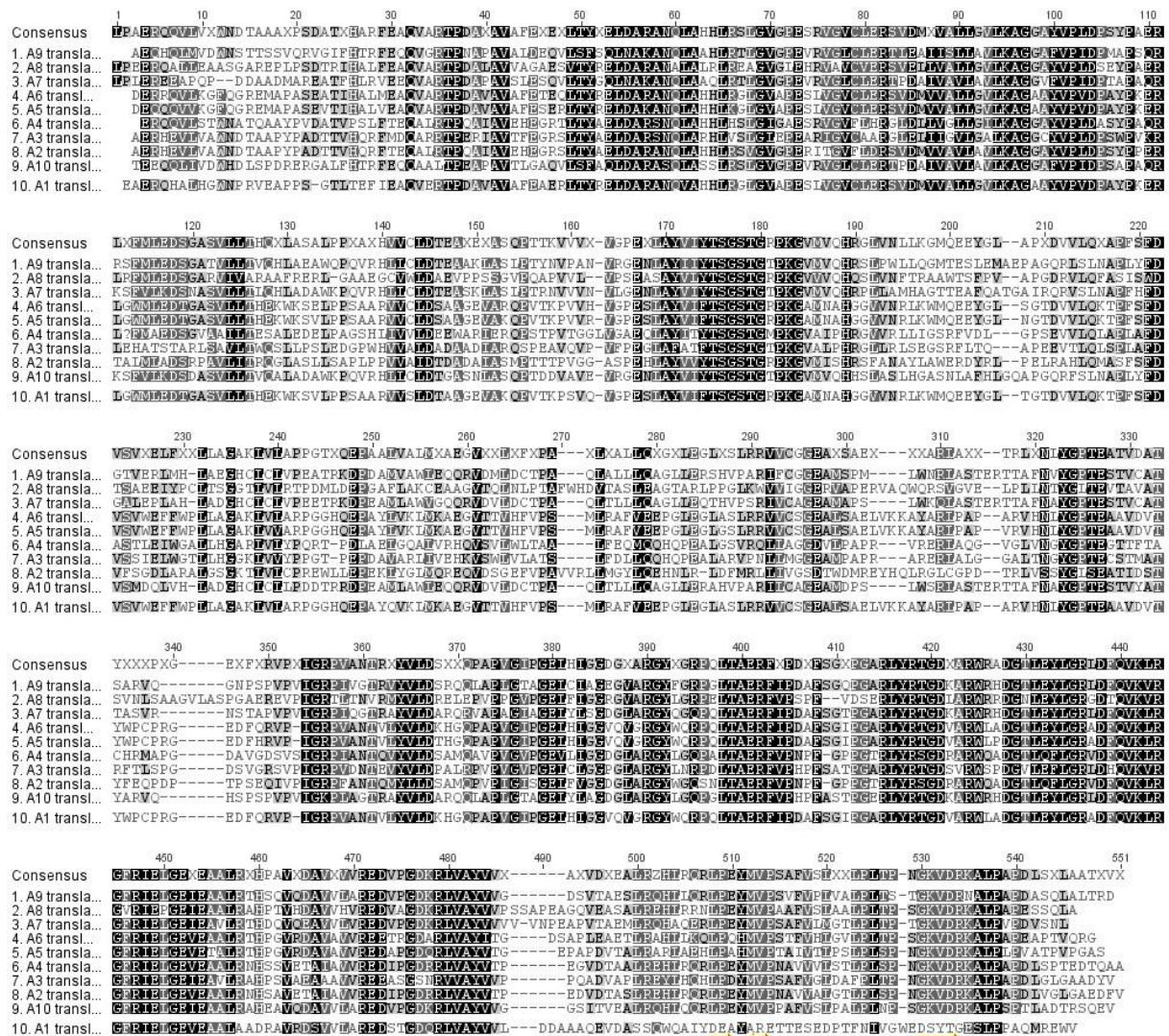
800startkon_1r = GAGGGTTCGCGGAGTCTCGTG

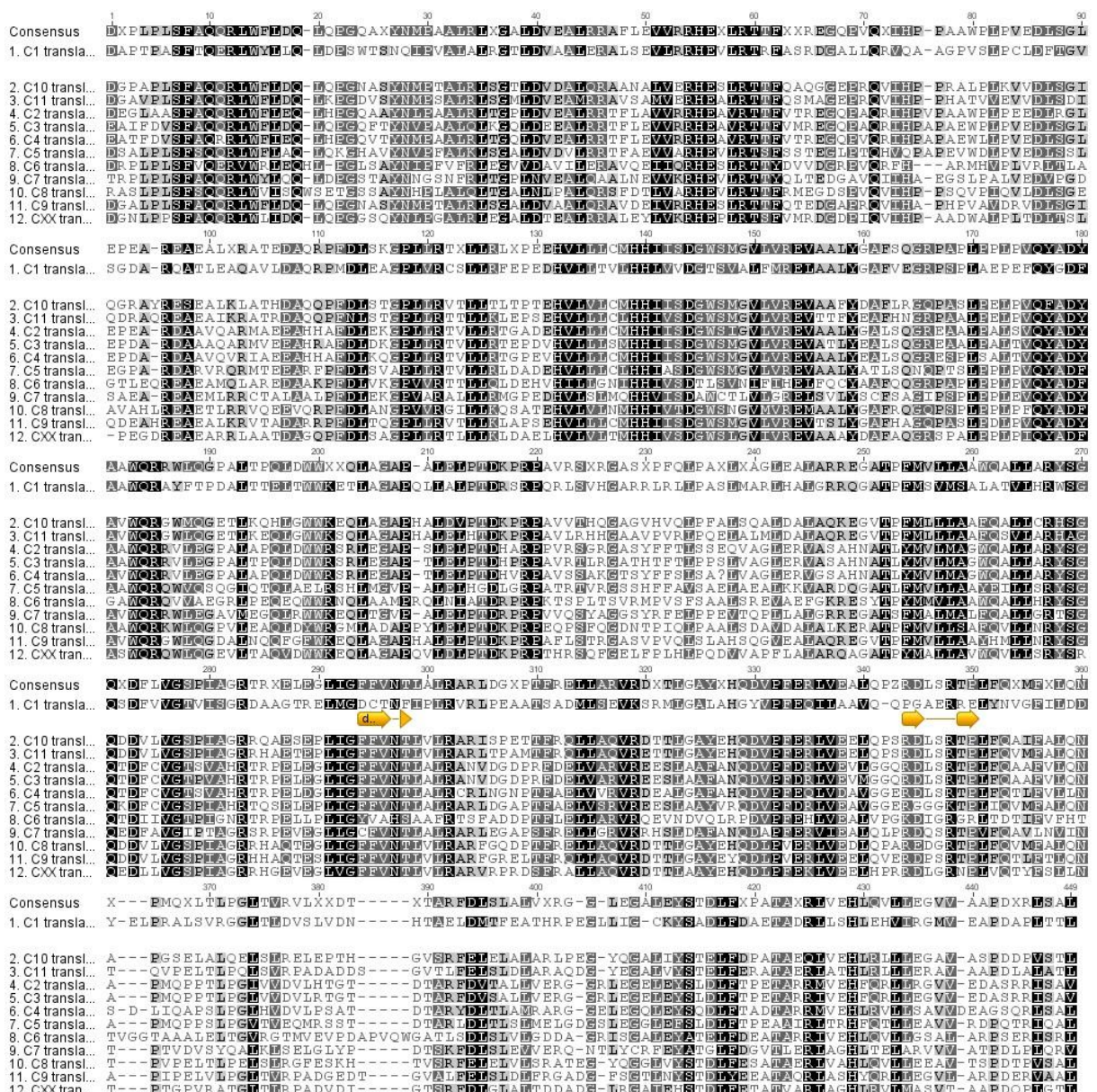
800startkon_1f = CGATTAAGTTGGGTAACGCCAGG



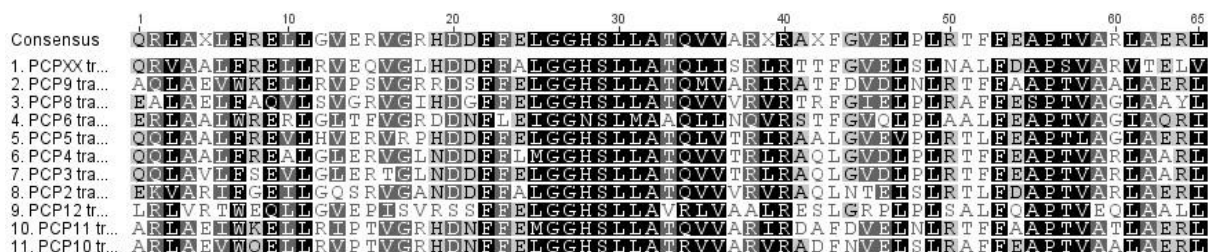
Clone 1, Clone 2, DK897, marker

Supplementary figure 80: Control gel of two DK897start clones and WT, 100 bp marker, expected size 288 bp.

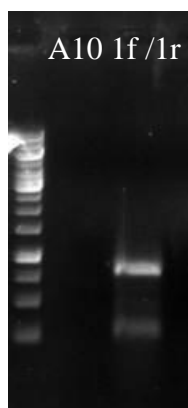




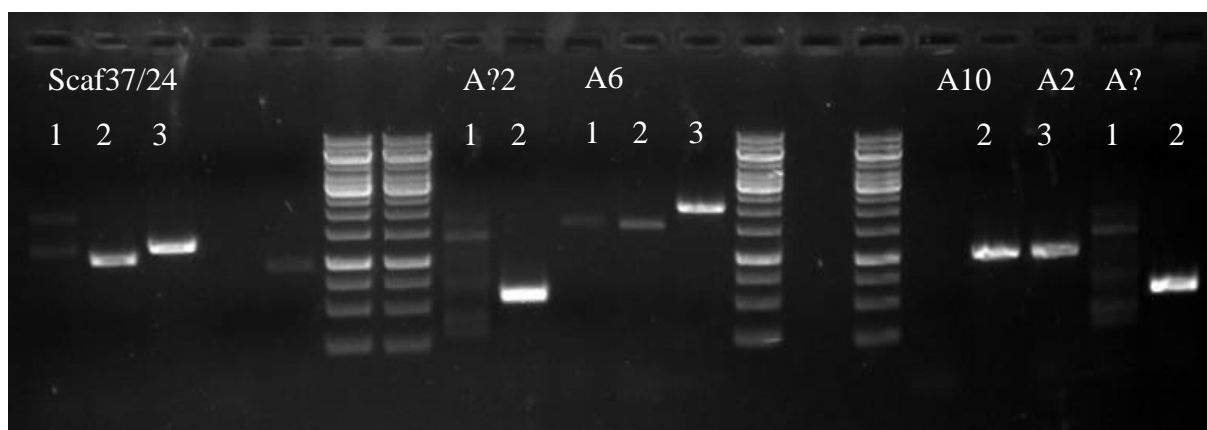
Supplementary figure 82: Alignment of C domains from the 800er cluster.



Supplementary figure 83: Alignment of PCP domains from the 800er cluster.



Supplementary figure 84: Gel after PCR with Phusion, GC buffer, 58 °C and primers for closing the gaps in the 800er cluster sequence.



Supplementary figure 85: Gel after PCR with Phusion, GC buffer, 63 °C and primers for closing the gaps in the 800er cluster sequence.

Production kinetics

To identify the optimal time point for compound isolation a production kinetic was performed. At defined time points triplicates were harvested until 99 h after inoculation. The samples were analyzed and quantified by HPLC-MS.

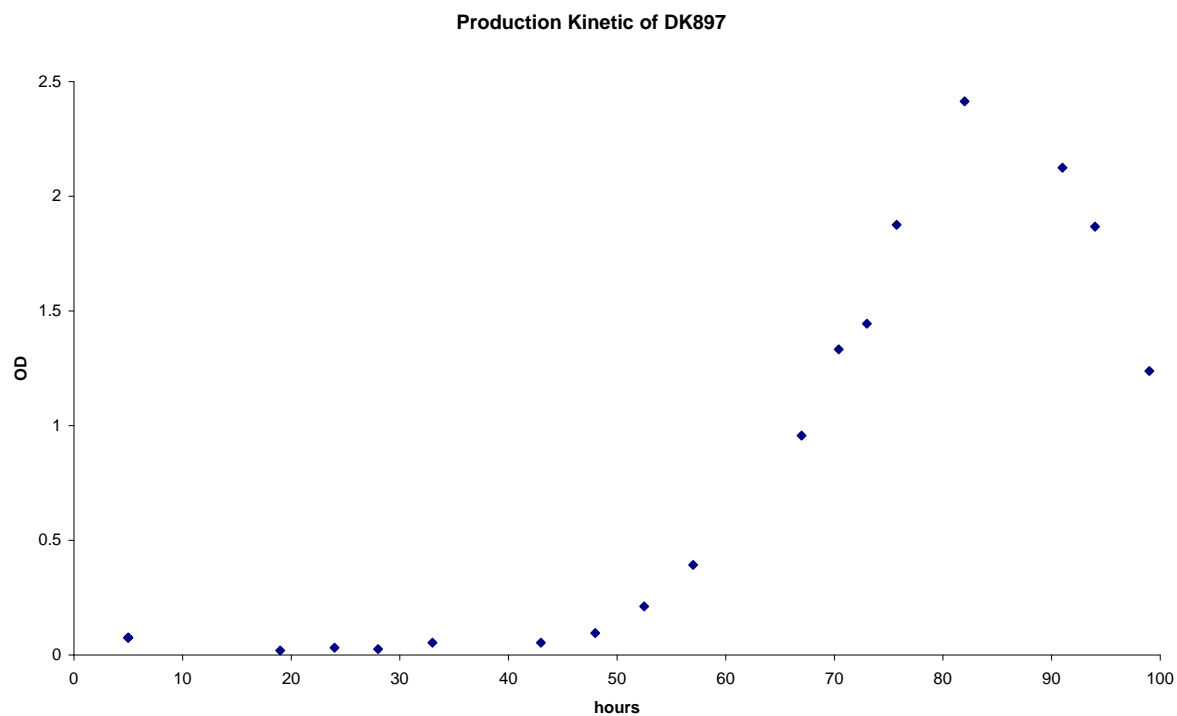


Figure 57: Growth curve of *M. xanthus* DK897.

The cell maximum is somewhere between 82 h and 91 h afterwards the culture begins to die (Figure 57).

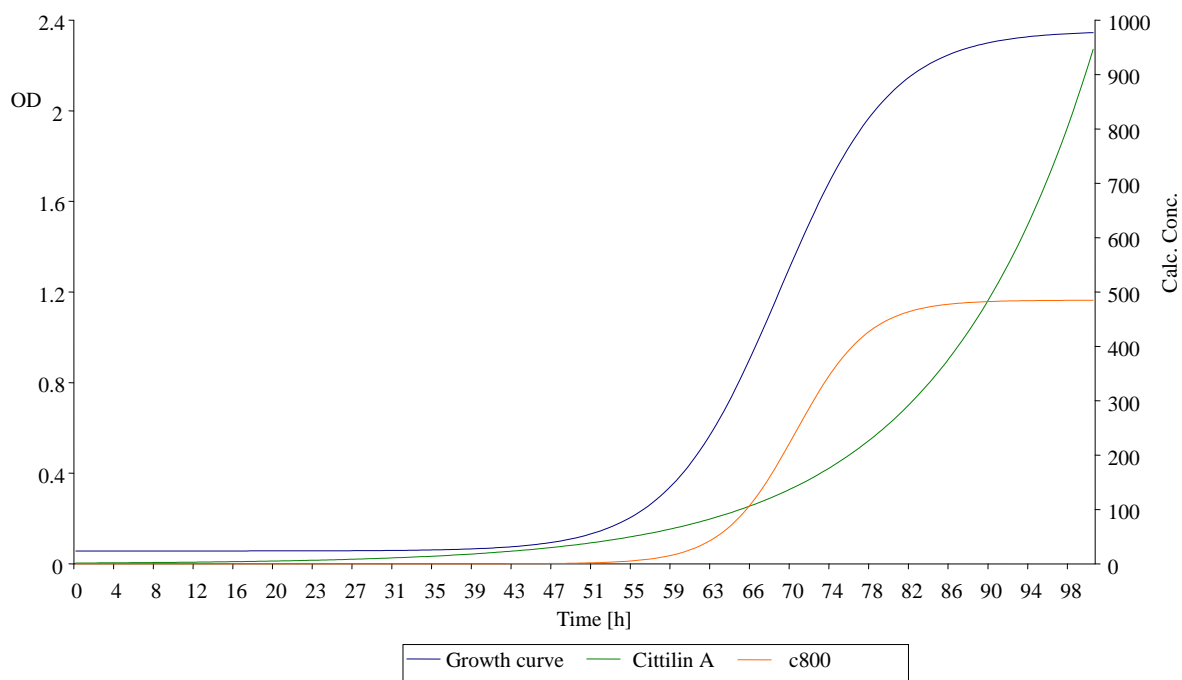


Figure 58: Production curve of cittilin A and c800 compared to cell growth.

Bioactivity tests of fractionated crude extracts

20 mg of extract of the strains DK897 with a 25 min long gradient at the waters machine. 48 fraction of each 11 ml (0.44 min) volume were collected. The solvent was removed and the remaining solid was transferred into a 48 well plate. Each fraction has been analyzed at the Orbitrap refer to section HPLC-HR-MS. Fractions were sent to Riken in Japan for following bioassays:

1. HL-60: Growth inhibition of human leukemia cell line, HL-60 for 48 h.
2. K562: Growth inhibition of human leukemia cell line, K562 for 48 h.
3. tsFT210: Growth inhibition of mouse cdc2 mutant cell line, tsFT210 for 48 h.
4. tsNRK: Morphological change of temperature-sensitive Rous sarcoma virus-infected NRK cells, tsNRK for 72 h.
5. *E. coli*: Growth inhibition of *Escherichia coli* (HO141 strain) for 6 h.
6. Yeast: Growth inhibition of *Saccharomyces cerevisiae* (MIC30M strain) for 18 h.
7. *M. oryzae*: Growth inhibition of *Magnaporthe Oryzae* (Kita strain) for 48 h.
8. p38: p38 MAP kinase inhibitor screening method using growth recovery of *E. coli* BL21(DE3) transformed by His-p38 expression plasmid as an index.
9. PDI: Enzymatic inhibition of insulin reductase activity of PDI for 3 h.
10. Heparanase: Enzymatic inhibition of heparan sulfate degradation assay for 24 h.
11. Plk1-PBD: Inhibition of polo box domain (PBD) of polo like kinase 1 (Plk1) dependent protein–protein interaction.
12. Plk3-PBD: Inhibition of polo box domain (PBD) of polo like kinase 3 (Plk3) dependent protein–protein interaction.

The following fractions showed activity (+: moderate activity; ++: strong activity), when possible it was correlated to masses.

1. DK897 (04) inhibited the growth of HL-60 (++) and K562 (+++) cells.

No clear MS peaks detected.

2. DK897 (15) inhibited heparanase activity (++).

Possible candidates: $m/z = 421, 485, 997, 800, 652, 417, 613, 576$

3. DK897 (24) inhibited PDI activity (++).

Possible candidates: $m/z = 602, 647, 552, 363, 289$

4. DK897 (37) inhibited the growth of *E. coli* (++).

Possible candidates: Myxalamides and $m/z = 344, 332, 358, 649$

5. DK897 (38) inhibited the growth of *M. oryzae* (+++).

Possible candidates: $m/z = 346, 334, 360$

6. DK897 (39) inhibited the growth of several cells (++ ~ +++) and inhibited Plk1/PBD and Plk3/PBD interactions (++)

No clear MS peaks detected

Chapter 2B

Myxoprincomide:

A Natural Product from *Myxococcus xanthus* Discovered by Comprehensive Analysis of the Secondary Metabolome

Cortina, N. S., Krug, D., Plaza, A., Revermann, O. and Müller, R. (2012)
Angew. Chem. Int. Ed., **51**, 811–816.

This article is available online at:

<http://onlinelibrary.wiley.com/doi/10.1002/anie.201106305/pdf>

The supporting information is available online at:

http://onlinelibrary.wiley.com/store/10.1002/anie.201106305/asset/supinfo/anie_201106305_s_m_miscellaneous_information.pdf?v=1&s=d6c3b5184ef7ed27dd47db6aecbd5d0632e0876b

Chapter 3

Characterization of a Novel Type of Oxidative Decarboxylase Involved in the Biosynthesis of the Styryl Moiety of Chondrochloren from an Acylated Tyrosine

Shwan Rachid, Ole Revermann, Christina Dauth, Uli Kazmaier, and Rolf Müller (2010)

J. Biol. Chem., **285**, 12482-12489.

This article is available online at:

<http://www.jbc.org/content/285/17/12482.full.pdf+html>

The supporting information is available online at:

<http://www.jbc.org/content/suppl/2010/01/15/M109.079707.DC1/jbc.M109.079707-1.pdf>

Chapter 4

Natural evolution in the laboratory: Spirangien C – F produced by a spontaneous mutant of *Sorangium cellulosum* So ce90

Abstract

Four new spirangiens C – F were isolated from extracts of a spontaneous mutant of *Sorangium cellulosum* So ce90. The novel structures were elucidated by NMR spectroscopy. All four compounds contain a shortened conjugated chromophore compared to the known spirangien. The mutation of the strain occurred during investigations of spirangien biosynthesis. To verify the mutated sequence region in the biosynthetic gene cluster of the spirangien a cosmid library was prepared. Sequence analysis revealed that the second last module exhibits mutations which might lead to module skipping. The results give further insights into spiroacetal formations.

Introduction

A vast number of natural products are produced by microorganisms. Within this huge amount of secondary metabolites there is enormous structure diversity. Many of these substances possess biological activity and show for example antibacterial or cytotoxic effects.^[1-4] Hence natural products serve as lead structures in medicinal chemistry^[4] but often natural compounds are a challenge for synthetic chemistry.^[227] To obtain new derivatives of natural products combinatorial biosynthesis can be employed.^[93] In nature novel substances emerge by evolutionary processes, each compound passed through a pre-selection for biological activity.^[18] Here it is reported from some kind of evolutionary process in *Sorangium cellulosum* which took place in the laboratory and conducted to novel secondary metabolites. The *Sorangium* species produces several different kinds of bioactive compounds: antifungals such as soraphens^[228] antibacterials like sorangicins^[75] or cytotoxic substances like ratjadons^[229]. Here the strain *Sorangium cellulosum* So ce90 was addressed. It produces famous anti cancer drug epothilone^[173;174] but it also generates another interesting compound which acts antifungal and is also highly cytotoxic, the spirangien A (**50**) (IC₅₀ 7 ng / mL against L929) and B (**51**).^[174;230] To date the mode of action still remains unknown^[231]. The natural producer builds only small amounts of Spirangiens. The metabolites are sensitive towards UV light, oxygen and heat.

The spirangiens have a polyketide backbone with a spiroketal core structure. The structure of Spirangien contains 14 stereo centers, a conjugated pentaenic chromophore with a terminal carboxylic acid. On the other side of the molecule it ends with a single double bond. At that position there is the difference between **50** and **51**. **50** ends with a methyl and **51** with an ethyl group. The absolute stereo chemistry was determined by chemical degradation and GC analysis.^[230] Furthermore a crystal structure of the spiroacetal fragment led to the relative configuration. The final absolute configuration was determined by total synthesis.^[232;233]

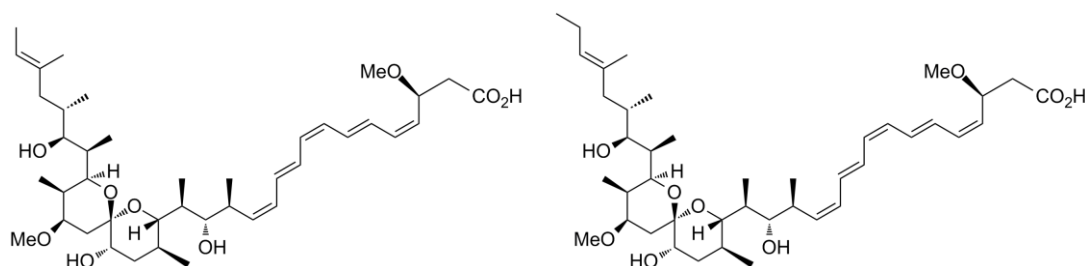


Figure 66: Structures of spirangien A (**50**; left), spirangien B (**51**; right).

Results and Discussion

The wild type strain So ce90 produces spirangien A (**50**) and B (**51**), but during conjugation experiments for the identification and the analysis of the its biosynthetic genes a spontaneous mutant (BFA3KO) appeared which exhibits the production of novel spirangien derivatives.^[175] The phenotype of the mutant differed from the wild type strain by the spirangiens. Compared to it the BFA3KO mutant produced no **50** and **51** but instead four new derivates with lower masses (Figure 67).

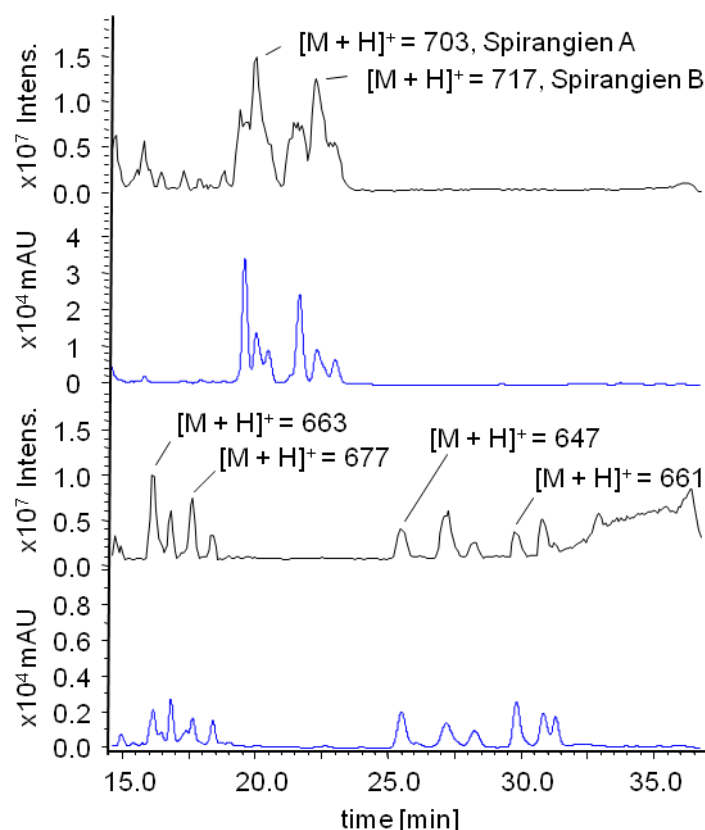


Figure 67: HPLC-MS chromatogram extract of wild type BPC⁻ (top) and UV chromatogram (upper middle), BFA3KO-3 mutant BPC⁻ (lower middle) and UV chromatogram (bottom).

High resolution mass spectrometry indicated molecular formulae as follows: C₃₈H₆₂O₉ (measured m/z = 663.44653, calculated m/z = 663.4467, spirangien C), C₃₉H₆₄O₉ (measured m/z = 677.46187, calculated m/z = 677.4623, spirangien D), C₃₈H₆₂O₈ (measured m/z = 647.45166, calculated m/z = 647.4517, spirangien E) and C₃₉H₆₄O₈ (measured m/z = 661.46619, calculated m/z = 661.4674, spirangien F). In contrast to **50** and **51** it appeared that these compounds lost one C₃H₄ unit plus one non-hydroxylated position (loss of O) in the case of spirangien E and F. Furthermore the UV-maxima exhibited a hypochromic shift of

30 nm (Figure 68) which indicates that conjugation to one double bond is lost in the unsaturated chain. ^[234]

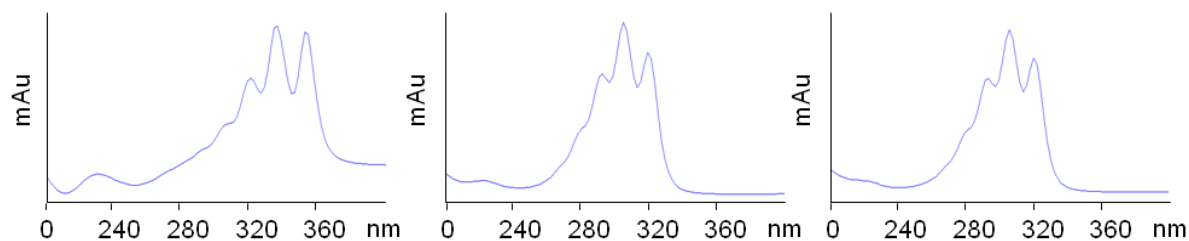


Figure 68: UV-spectra of spirangien A and of new derivatives spirangien C and E (order right to left).

To proof this hypothesis of a shortened chromophore the new derivatives had to be isolated. The BFA3KO mutant was fermented in large scale (by K. Gerth, Helmholtz Centre for Infection Research (HZI), Braunschweig) and the obtained crude extract was used for spirangien isolation. As the spirangiens are very sensitive to light and oxygen, the purification process was performed under light and oxygen exclusion. Crude extract of BFA3KO mutant was extracted with ethyl acetate. The received mixture was fractionated on a Sephadex[®] LH-20 (MeOH) column. The obtained fractions were separated by reverse phase HPLC to lead to the isolation of spirangien C, E and F.

The pure substances were used for NMR experiments. Their NMR spectra (see Supplementary figure 86 to Supplementary figure 102) showed a lack of one double bond compared to the published data of spirangien A and B. ^[230] The NMR-spectra revealed the presence of 10 sp^2 -hybridized carbons whereof two are already assigned for C-30 and C31. Therefore only four double bonds are left for the chromophore region which is in good agreement with the shift of 30 nm in the UV-spectra (Figure 68).

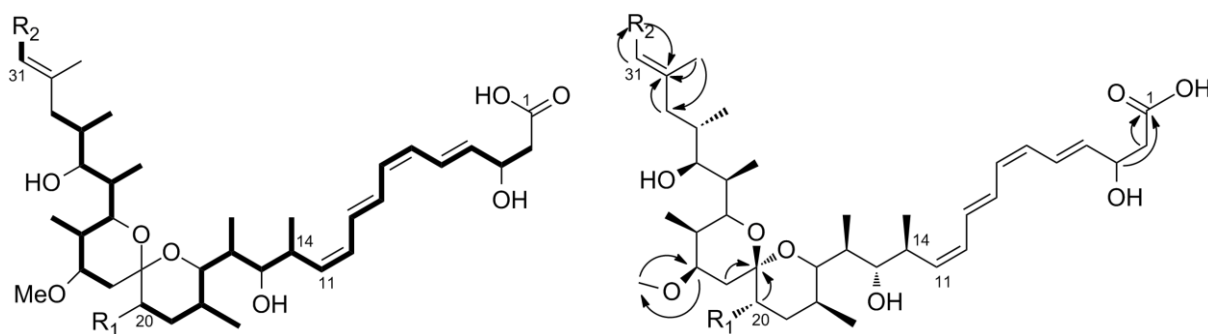


Figure 69: COSY fragments were assigned (marked in black); HMBC correlations as arrows (only fundamental ones); $R_1 = -OH$ or $-H$; $R_2 = -CH_3$ or $-CH_2CH_3$.

Spirangien F has a shift at position C-20 from 71.5 ppm to 37.5 ppm, suggesting here a lack of oxygen. COSY and HMBC spectra proofed that the spiroketal backbone is unchanged (Figure 69). The single double bond at position C-30 was verified by HMBC-correlations. In an analogous manner the methyl ether at position C-23 (78.6 ppm) is attested and the methyl group at position C-14 is present at 1.11 ppm. In contrast to that no hydroxyl methyl was found at position C-3.

To identify which of the double bonds of the chromophore is lacking compared to the original spirangien, the coupling constants of the conjugated double bonds were analyzed (Figure 70). By using them it is possible to decide if it is a *Z* or *E* double bond. Only C-4 / C-5 and C-8 / C-9 have large coupling constants of 14 to 15 *Hz* which is indicative of *trans* configuration. Hence the shortened spirangiens have an order of *Z*, *E*, *Z* and *E* double bonds, respectively. The known spirangien has a *Z*, *E*, *Z*, *E*, *Z* configuration. Aware of the existence of the methyl group at C-14 and biosynthetic consideration the double bond of C-12 / C-13 of the common spirangien is missing in the novel derivatives.

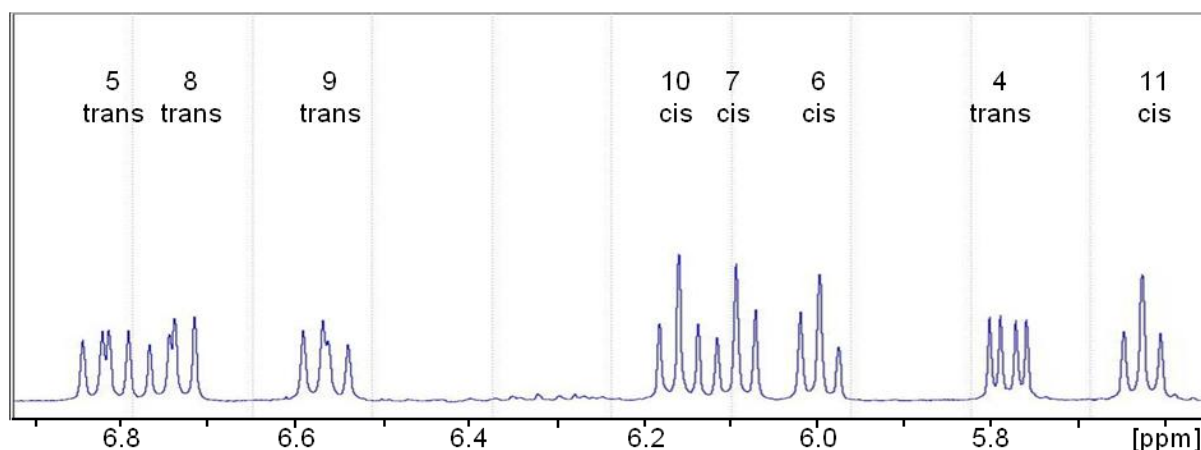


Figure 70: Double bond region of the ^1H -NMR-spectra of shortened spirangien.

Four new spirangien derivatives could be isolated; three of them could be characterized completely by NMR. Their final structure is depicted in Figure 71, the NMR data are listed in Table 19 to Table 21 and the spectra are shown in Supplementary figure 86 to Supplementary figure 102.

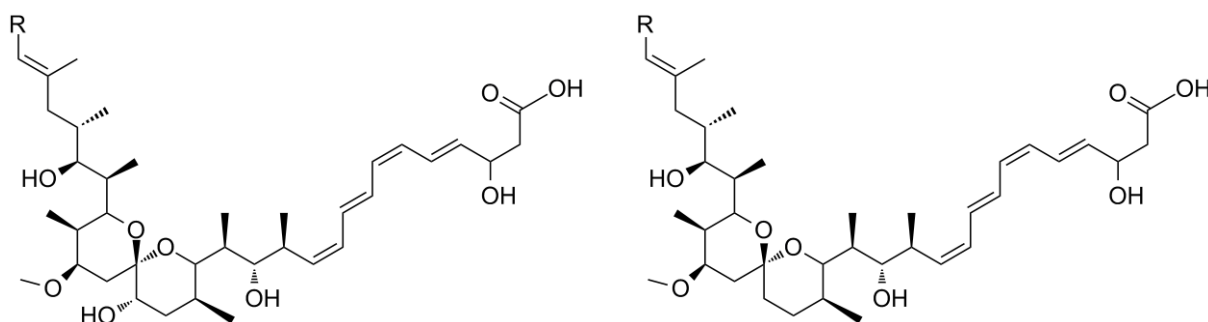


Figure 71: Shortened spirangien derivatives, left side: spirangien C (R = Me) and spirangien D (R = Et); right side: spirangien E (R = Me) and spirangien F (R = Et).

As further proof of the structures feeding experiments with L-[methyl-3D]-methionine to the BFA3KO mutant and to the wild type have been executed (by B. Frank). L-methyl-methionine is needed for the SAM pool, which is used for methylation of hydroxyl groups. After feeding of L-[methyl-3D]-methionine the spirangien of the wild type showed a mass shift of + 3 and + 6, the spirangiens of the mutant in contrast had only a mass shift of + 3. This result indicates that one methyl group is lost; this is in agreement with the NMR structure elucidation.

To identify the reason for the shortened spirangiens one has to observe the biosynthetic gene cluster of spirangien. It is 88 kb in size, encodes seven polyketide synthases *spiD* to *spiJ* and six further genes (*spiA*, *spiB*, *spiC* and *spiK*, *spiL* and *spiZ*) refer to Figure 72. ^[235] The function and the domains of each gene are given in Table 17. And a biosynthetic scheme for the spirangien assembly by Frank et al. is depicted in Figure 73. ^[235]

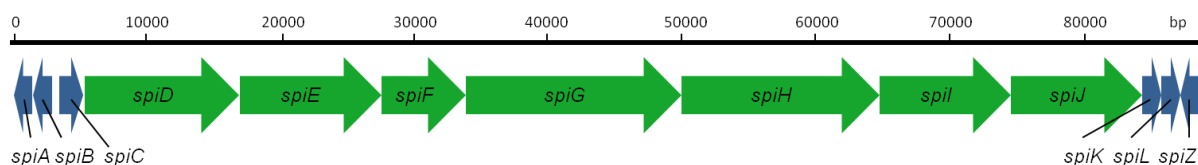


Figure 72: Biosynthetic gene cluster of spirangien; seven polyketide synthases (green) flanked on both sides by adjacent genes (blue).

Gene	Function	Protein Domains
spiA	Thioesterase	TE
spiB	Methyltransferase	Me
spiC	Cytochrome P450 monooxygenase	P450
spiD	PKS	ACP, KS, AT, AT, DH, KR, ACP, KS, AT, DH, ER, KR, ACP
spiE	PKS	KS, AT, KR, ACP, KS, AT, DH, KR, ACP
spiF	PKS	KS, AT, KR, ACP
spiG	PKS	KS, AT, KR, ACP, KS, AT, DH, ER, KR, ACP, KS, AT, KR, ACP
spiH	PKS	KS, AT, KR, ACP, KS, AT, KR, ACP, KS, AT, DH, KR, ACP
spiI	PKS	KS, AT, KR, ACP, KS, AT, DH, KR, ACP
spiJ	PKS	KS, AT, DH, KR, ACP, KS, AT, KR, ACP, TE
spiK	Methyltransferase	Me
spiL	cytochrome P450 monooxygenase	P450
spiZ	sigma-24 factor	Σ

Table 17: Biosynthetic genes of spirangien and their function. AT (acyl transferase), KS (ketosynthase), ACP (acyl carrier), KR (ketoreductase), DH (dehydratase), ER (enoyl reductase), TE (thioesterase).

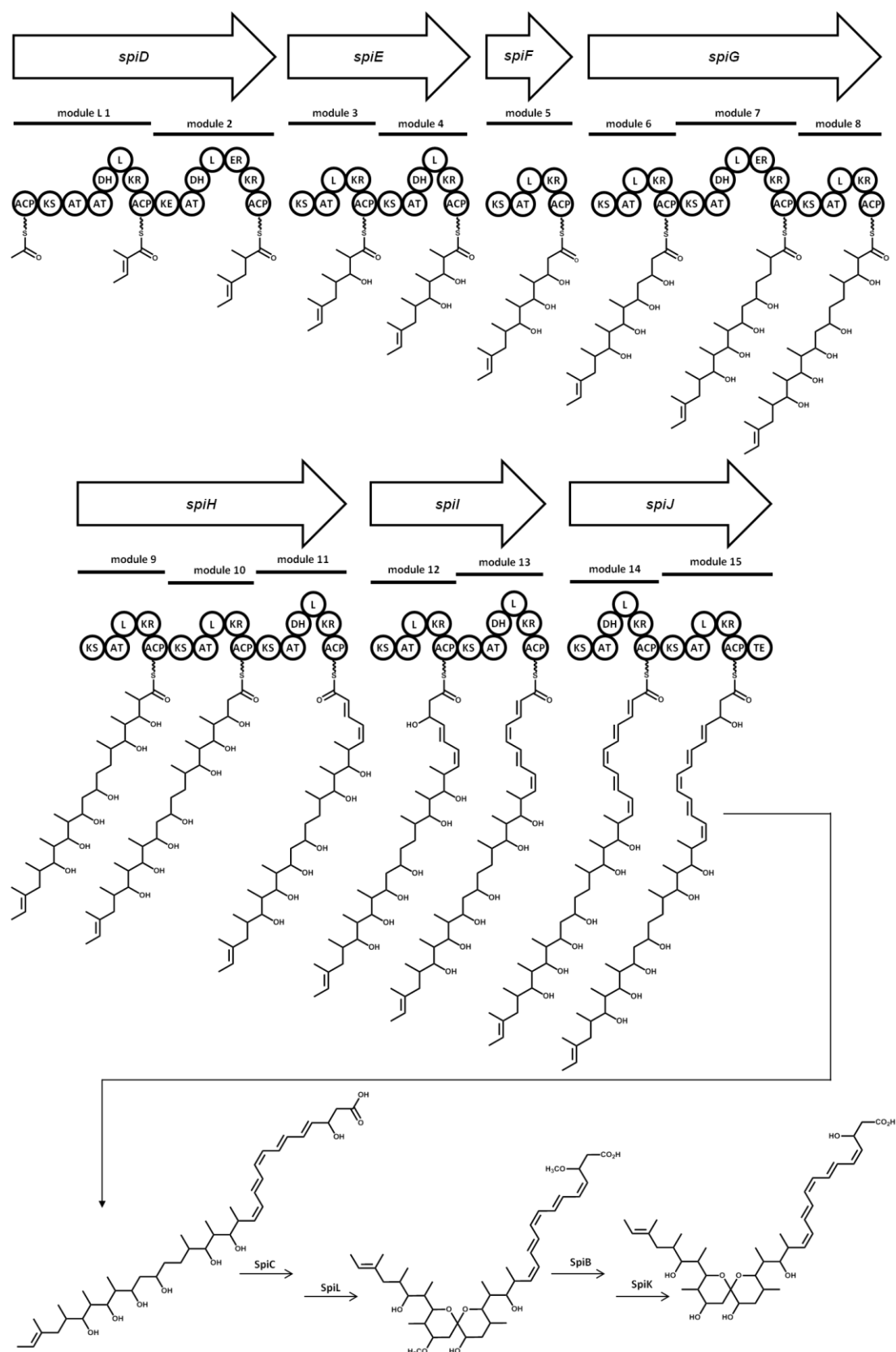


Figure 73: Biosynthesis model of Spirangien A according to Fank et al. AT (acyl transferase), KS (ketosynthase), ACP (acyl carrier), KR (ketoreductase), DH (dehydratase), ER (enoyl reductase), TE (thioesterase), L (linker regions).

The BFA3KO mutant (= BFR-J40) occurred during inactivation experiments of a gene encoding putative transcription factor (*deoRB*) in *S. cellulosum* So ce90. ^[175] It is unlikely that this inactivation altered the metabolites as it is assumed that *deoRB* is not directly involved in the disappearance of spirangien and in the production of new derivatives of that compound. It is more likely that a mutation or a deletion in the sequence of the biosynthetic gene cluster of the spirangien took place. Knowing the structures of the shortened spirangiens one can assume that the sequence in gene *spiJ* in module 14 differs compared to the one described in Figure 73. In order to identify any change in the sequence of the spirangien gene cluster during *deoRB* inactivation, the assumed region of *spiJ* was tried to be amplified but several attempts failed presumably because of extensive sequence repeats in the respective region. Furthermore, a comparative Southern blot analysis of genomic DNA of the mutant and the wild type indicates that no deletion of a larger region took place in the mutant (details in PhD thesis of B. Frank ^[175]). A further test to generate a double mutant and subsequent vector recovery failed because the respective construct (pRK415 derivative with thiostreptone resistance gene) could not be cloned. One way of getting information what exactly happened in the cluster was the generation of a cosmid library from the mutant genome and subsequent screen by PCR / Southern hybridization. Alternatively, cloning of the cluster by redET-technology could be performed. ^[236-238] But finally a cosmid library of the BFA3KO mutant was generated by S. Rachid to re-sequence the gene cluster.

To screening the cosmid library, hybridization with homologous DNA fragments from the ketosynthase enzyme of the module 9 of the known spirangien gene cluster as well as a fragment of the thioesterase in *spiJ* gene was carried out by S. Rachid (Figure 74). Primers BF_SH1 and BF_SH9 were used for amplification of a DIG-labeled KS probe, and BFST1 and CFST2 were used to generate a DIG-labeled fragment of the spirangien TE.

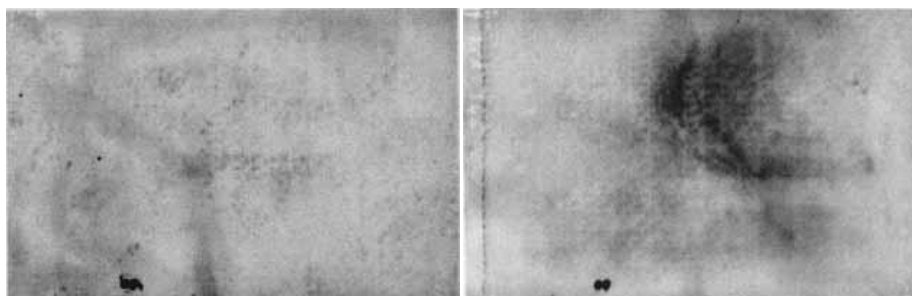


Figure 74: Hybridization of the BFA3KO cosmid library membrane with a KS probe of the module 9 (left) and with a TE probe from the *spiJ* gene, resulted in identification of B:P6 cosmid.

It has been confirmed by a PCR, that the cosmid B:P6 carries the 3' end of the gene cluster, the KS encoding gene and the TE region. Thus the insert of the cosmid was sequenced on both strands (Maren Scharfer, HZI Braunschweig), by using a shotgun library with DNA fragments of approximately 1.5 – 2.0 kb in length. Sequence analysis showed an insert length of 36053 bp. The sequence was aligned with the wild type sequence. Each difference in the sequence compared to the wild type was re-sequenced. Finally the alignment showed several mutated regions or point mutations in the sequence of *spiJ*, module 15, in an acyltransferase (AT). As these mutations occur at the same time it could be possible that there are some sequencing mistakes in the original wild type sequence. In the DNA sequence of the *spiJ* seven regions within the second AT domain with either nucleotide substitution, nucleotide deletion or insertion have been observed (Figure 75, latest status of ongoing investigations of S. Rachid).

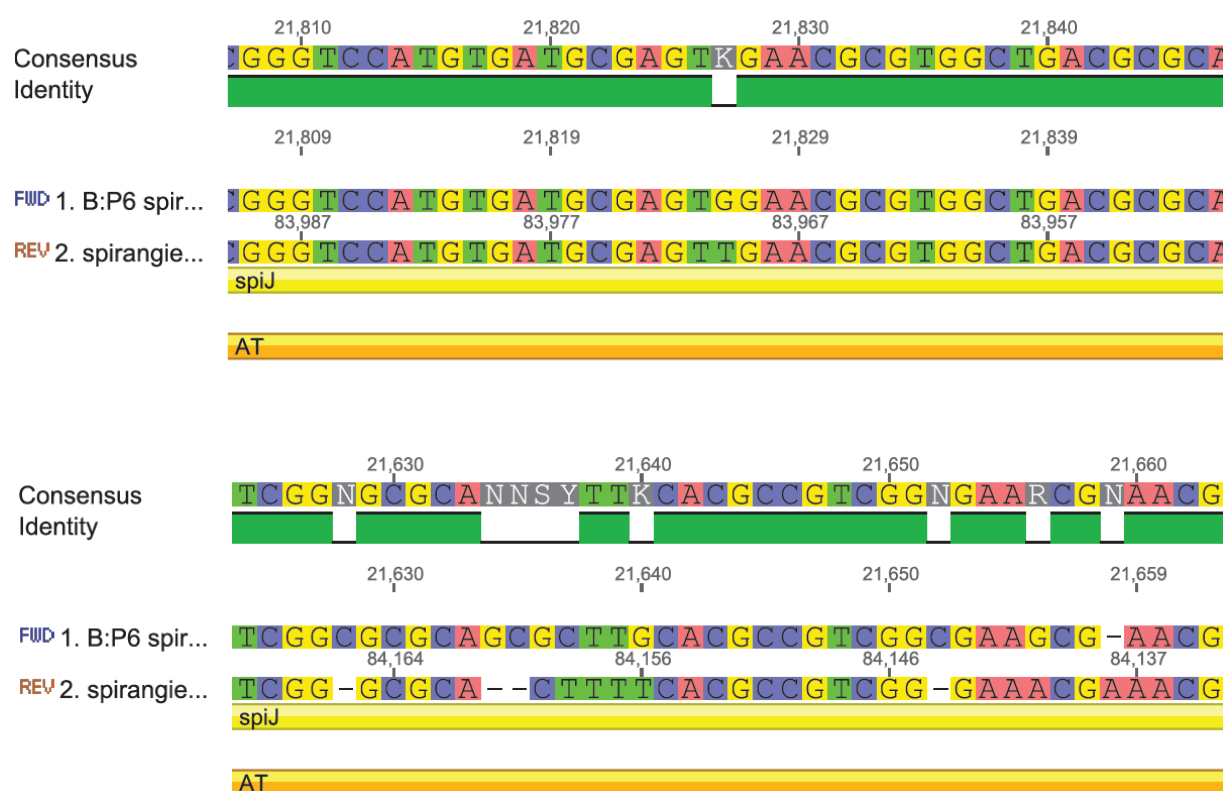


Figure 75: Alignment of the spirangien cluster and of the mutated cluster in cosmid B:P6. Only regions with differences in *spiJ* are shown.

Such a genetic modification might have caused a module skipping. Module skipping means to bypass the utilization of a complete module or of a single domain of a biosynthetic pathway. ^[239] The biosynthetic intermediate is directly transferred from a module to the after next one that means a complete module with all its enzyme activities is mistreated. ^[240] In

another case the ACP of the module is involved but the other domains are skipped and no extender unit is added to the intermediate due to an inactive AT domain.^[207]

The mutations here might have resulted in the inactivation of the AT domain of *SpiJ* and it can be assumed that a module skipping occurred within the final PKS extension module to lead to a production of shorter spirangiens.

The production of shorter derivatives without the hydroxyl group at position C-20 demonstrates that the substrate recognition for *SpiL* is tough, as spirangien C and D are oxidized in contrast to spirangien E and F. The lack of the oxidation here also sheds light onto the spiroketal formation. The biosynthetic spiroacetal formation is poorly understood as intermediates are often unstable and responsible enzymes were difficult to identify.^[241] For the spiroketal monensin it is described; it is formed by an S_N2 reaction catalyzed by an epoxide-hydrolase/cyclase enzymes.^[242] In contrast to that Osada et al. recently identified two responsible enzymes (RevG and RevJ) which are involved in the spiroacetal formation of reveromycin.^[241] RevG (a dihydroxy ketone synthase) conducts the dehydrogenation and RevJ (a spiroacetal synthase) accomplishes the cyclization.

The spiroacetal formation of avermectin^[243] and tautomycin is explained via a dehydrative cyclization mechanism as it is suggested for spirangien, too. Frank et al. describe two possible ways for the construction of the spiroketals (Figure 76).^[235] In the one case first the oxidation of C-21 takes place then spiroketal formation follows and the hydroxylation at C-20 occurs at the end. Alternatively C-20 is oxidized at first, then C-21 and afterwards the first ring is closed followed by the second ring closure. But the presence of non-hydroxylated derivatives at C-20 indicates that this last proposed pathway has to be wrong and that the hydroxyl group is not essential for spiroketal formation. Therefore, the left pathway of Figure 76 has to be the correct mechanism and the other hypothetical pathway can be excluded.

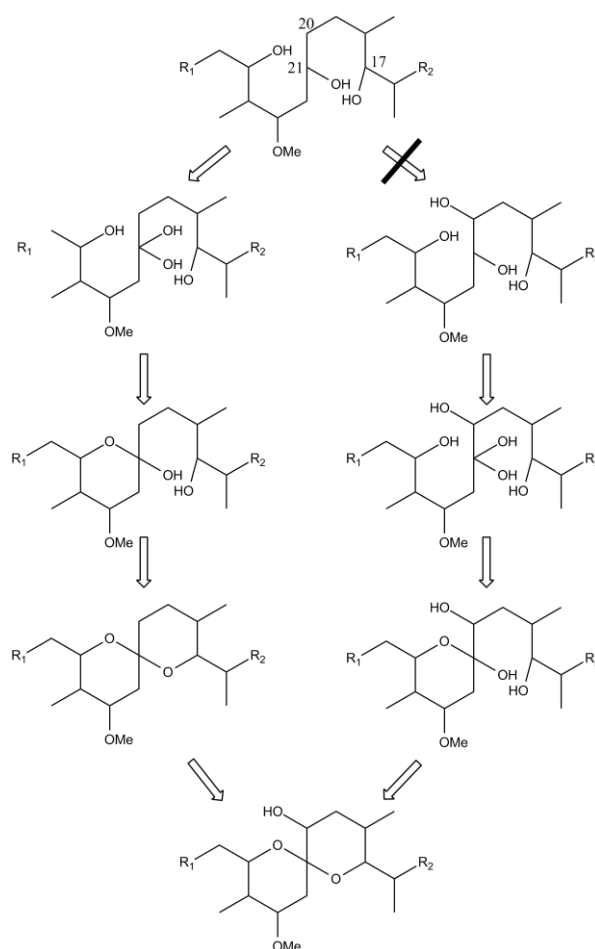


Figure 76: Proposed spiroketal formation during the biosynthesis of spirangien; ketogroup at C-21 shown as hydrated form.

To summarize some kind of evolution imitated accidental in the lab occurred to the strain So ce90. Spontaneous mutations appeared within the biosynthetic gene cluster of spirangien as a consequence novel spirangien derivatives emerged. It might be seen as evolution in the lab in front of the researches eye. Seven variabilities have occurred in the sequence of the second AT domain in *spiJ*. At least one of those mutations apparently caused a module skipping. As result of the skipping shorter spirangien derivatives have been produced by the “natural” mutant. These were isolated and characterized by NMR analysis. Their structure fits perfectly to the genetic modifications within the biosynthetic gene cluster. Furthermore it shows that methyltransferase (SpiK) has very tight substrate specificity as no further methyl ether occurs. The cytochrome P450 (SpiL) still accepts the novel derivatives but the error ratio is high as not the complete spirangien E and F are converted.

The evidence of the structure spirangien E and F and the interpretation of their biosynthesis contributes to the understanding of the spiroketal formation.

Experimental Section

Purification process of the spirangien derivatives

The BFKO mutant of *So ce90* was cultivated together with XAD in a 9 l fermenter (K. Gerth), for cultivation details see Frank et al. ^[235]. As the spirangien especially the shorten derivatives are light sensitive the whole work up process was performed with dimmed light, whenever possible under nitrogen and stored in MeOH at -20 °C. The XAD was eluted in an open glass column with MeOH (1.25 l). The organic solvent was removed; some brine was added to the water phase and extracted three times with ethyl acetate. The ethyl acetate contained 2.02 g of extract which was dissolved in MeOH and a partition (2 x) against heptane reduced the extract to 1.88 g. The crude extract was placed on a MPLC column (MeOH / H₂O (85 : 15), 50 mM K₂HPO₄ / NaH₂PO₄, pH = 7, flow rate = 50 ml / min, d = 10 cm, 10 bar) this led to six fractions which contained mainly: epothilone A, epothilone B, spirangien C, spirangien D, spirangien E and spirangien F. The MeOH was removed and the water phase extracted with ethyl acetate (3 x), the combined ethyl acetate was extracted against water, the solvent was dried with Na₂SO₄. The spirangien containing fractions were purified with preparative HPLC (16 x 250 mm Nucleodur C18 Machery – Nagel, MeOH / H₂O (85:15), 50 mM K₂HPO₄ / NaH₂PO₄, pH = 7, flow rate = 15 ml / min, detection: 320 nm). The appropriate peaks were collected, the MeOH was removed and the water phase extracted (3x) with EE which then was distributed against water and dried over Na₂SO₄. Four novel spirangien derivatives were isolated: spirangien C (2.1 mg), spirangien D (less than 1 mg), spirangien E (12.4 mg), spirangien F (7.6 mg).

Structure analysis

The structure analysis was performed by NMR spectroscopy in CD₃OD using a Bruker Avance 500 instrument. ¹H-, ¹H,¹H COSY-, HSQC-, HMBC- spectra were measured and when it was possible ¹³C spectra. The data are listed Table 19, Table 20 and Table 21.

Bioactivity

The three spirangien derivatives were tested against following cell lines: L929, Sw480 and U037. The new derivatives showed IC₅₀ values between 0.0007 to 0.005 µg / ml (see Table 18. The biological assay were performed by Y.A. Elnakady.

	Spirangien C IC50 [µg / ml]	Spirangien E IC50 [µg / ml]	Spirangien F IC50 [µg / ml]
L929	0.005	0.002	0.002
Sw480	not available	0.003	0.003
U937	0.0007	0.002	0.001

Table 18: IC50 values for the new spirangien derivatives.

Creation of cosmid library

Chromosomal DNA of *So ce90 BFA3KO* was isolated and partially digested with *Sau3AI*, dephosphorylated and subsequently ligated into SuperCos1 (Stratagene) pre-treated with *XbaI*, dephosphorylated, and restricted with *BamHI*. Packaging of the ligation mixture with Gigapack III Gold (Stratagene) packaging extract and transducing the resulting phages into *E. coli* HS699 generated a genomic library consisting of 2304 clones.

All single colonies were transferred into 384-well microtiter plates, grown in LB medium overnight, and replicated twice. Glycerol (10 %) was added to one copy of the library, and the plates were frozen at - 80 °C. For colony hybridization, the colonies of the second copy were transferred twice onto a 22.2 x 22.2 cm nylon membrane (Biodyne B, Pall) with a Qbot robot (Genetix). The creation of the cosmid library was done by S. Rachid.

Identification of the B:P6 cosmid

Primers BF_SH1 (5'-CGGAGATCGCATCGATGTC-3'), and BF_SH9 (5'-GTTCGGGACC GTGATCCC-3') were used to amplify a 952 bp DIG-labeled KS probe, and primers BFST1 and CFST2 (sequence by B. Frank) were used to generate a 576 bp DIG-labeled fragment of the TE (Supplementary figure 103).

No.	¹ H [ppm]		<i>J</i> [Hz]	¹³ C [ppm]	H-H-COSY	HMBC
1	-	-	-	174.6	-	-
2	2.51	d	6.4	43.1	3	1, 3, 4
3	4.60	m	-	69.8	2, 4	1, 2, 4
4	5.77	dd	6.2, 14.8	136.7	3, 5	2, 3
5	6.80	dd	11.5, 14.8	126.8	4, 6	6
6	5.99	t exp. dd	11.1	129.6	5, 7, 8	4, 5, 7
7	6.09	t exp. dd	11.5	131.0	6, 8	6, 8, 9
8	6.74	dd	11.5, 14.3	129.3	7, 9	7, 9
9	6.57	t exp. dd	11.8 14.3	130.4	8, 10, 11	7, 8, 10
10	6.16	t exp. dd	11.0	129.9	9, 11	9
11	5.62	t exp. dd	10.6	134.3	10, 14	9, 14 1 4Me, 15
14	2.94	m	-	35.8	11, 14Me, 15	10, 11, 14Me
14 Me	1.11	d	6.9	19.7	14	11, 14, 15
15	3.55	dd	2.2, 9.7	76.4	14, 16	11, 14, 16, 17
16	1.61	m	-	39.8	15, 16Me, 17	15, 16Me, 17
16 Me	0.79	d	6.9	9.3	16	15, 16
17	3.59	m	-	74.8	16, 18	15, 16, 16Me
18	1.91	m	-	25.4	18Me, 19ab	17, 18Me
18 Me	0.75	d	6.7	17.9	18	17, 18, 19
19	1.73	m	-	37.5	18, 18Me, 20	18, 20
19b	1.67	m	-	37.5	18, 18Me, 20	20, 21
20a	3.38	m	-	71.0	19ab	18, 19
20b	-	-	-	-	-	-
21	-	-	-	99.5	-	-
22a	2.04	m	-	33.7	22b, 23	21, 23, 24
22b	1.37	m	-	33.7	22a, 23	21, 23
23	3.65	m	-	78.6	22a, 22b, 24	23oMe, 24Me
23 OMe	3.36	s	-	55.1	-	23
24	2.12	m	-	32.5	23, 24Me, 25	22, 23, 24Me
24 Me	0.77	d	7.0	3.9	24	23, 24, 25
25	3.63	m	-	72.3	24, 26	23, 24Me, 26, 27
26	1.82	m	-	37.6	25, 26Me, 27	25, 26Me
26 Me	0.79	d	7.0	7.9	26	25, 26, 27
27	3.68	m	-	75.9	26, 28	25, 26Me, 28, 29
28	1.73	m	-	35.1	29ab, 27	27, 28Me, 29
28 Me	0.71	d	6.7	15.5	28Me	27, 28, 29
29a	2.66	m	-	45.6	31, 29b, 28	28, 28Me, 30
29b	1.60	m	-	45.6	29a, 31	28, 30
30	-	-	-	135.4	-	-
30 Me	1.60	m	-	14.7	-	29, 30, 31
31	5.21	q	6.5	120.9	29ab, 32	29, 30Me
32	1.58	m	-	13.4	31	30, 31
33	-	-	-	-	-	-

Table 19: NMR Spectroscopic data of the spirangien C (500 MHz, methanol-d₄), for better comparison to spirangien A and B the same numbering ^[230] is use, therefore C-12 and C-13 do not exist.

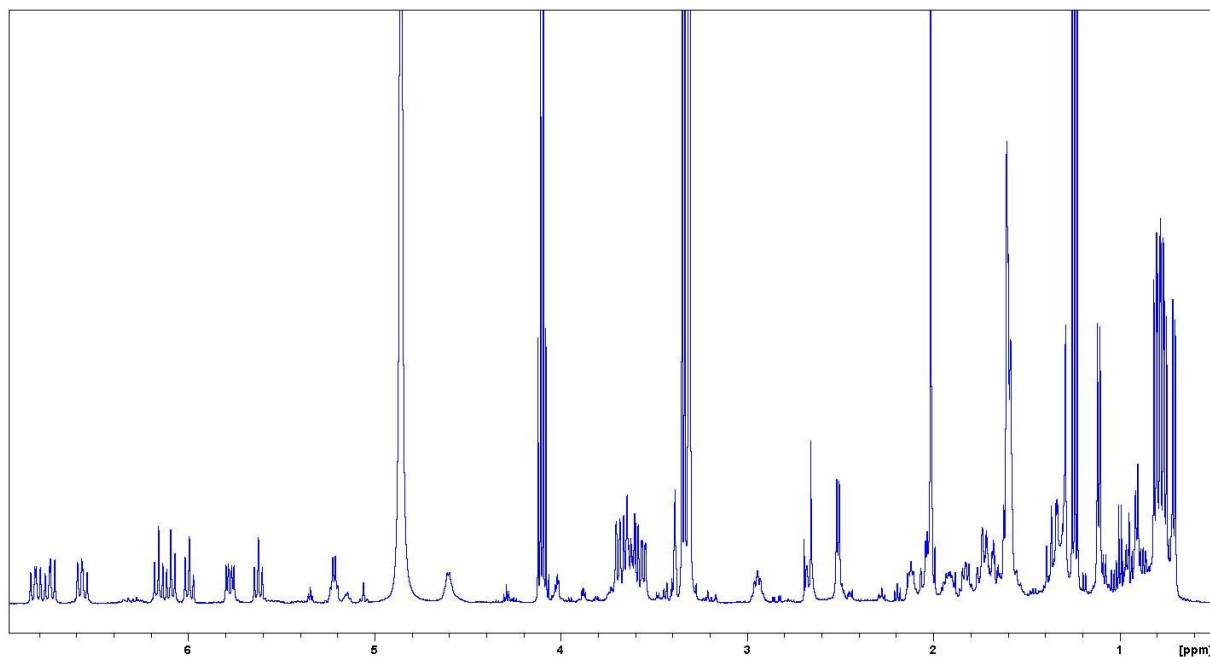
No.	¹ H [ppm]		<i>J</i> [Hz]	¹³ C [ppm]	H-H-COSY	HMBC
1	-	-	-	175.2	-	-
2	2.50	d	6.8	43.3	3	1, 3
3	4.60	m	-	69.5	2, 4, 5	1, 2, 4, 5
4	5.77	dd	6.3, 15.1	136.8	3, 5, 6	2, 3, 6
5	6.82	dd	11.2, 15.0	127.0	3, 4, 6, 7	3, 6
6	5.99	t exp. dd	11.0	129.6	5, 7	4, 5, 8
7	6.09	t exp. dd	11.0	131.0	5, 6, 8	5, 6, 8
8	6.74	dd	11.8, 14.7	129.2	6, 7, 9	9
9	6.56	t exp. dd	11.2, 14.2	130.3	8, 10, 11	7, 10, 11
10	6.15	t exp. dd	11.4	129.8	9, 11	9, 14
11	5.62	t exp. dd	10.3	134.1	9, 10, 14	10, 14, 14Me, 15
14	2.93	m	-	35.7	11, 14Me, 15	11, 14Me
14 Me	1.11	d	6.7	19.7	14	14, 15, 17
15	3.54	dd	2.0, 9.9	76.6	14, 16	11, 14Me, 17
16	1.61	m	-	40.3	15, 16Me, 17	15, 16Me
16 Me	0.74	d	4.4	9.5	16	15, 16, 17
17	3.60	d	9.8	75.4	16, 18	16, 16Me, 18Me, 19, 21
18	1.51	m	-	32.3	17, 18Me, 19	17, 19
18 Me	0.78	d	5.9	18.4	18	17, 18, 19
19	1.56	m	-	29.9	18, 20ab	18, 18Me, 20, 21
19b	1.56	m	-	29.9	18, 20ab	18, 18Me, 20, 21
20a	1.65	m	-	37.5	19, 20b	18, 19, 21
20b	1.49	m	-	37.5	19, 20a	18, 19, 21
21	-	-	-	98.1	-	-
22a	1.72	m	-	38.0	22b, 23	21, 23, 24
22b	1.43	m	-	38.0	22a, 23	21, 23
23	3.66	m	-	78.6	22ab, 24	21, 23
23 OMe	3.32	s	-	55.5	-	23
24	2.12	m	-	33.1	23, 24, 25	23, 24Me, 26
24 Me	0.77	d	6.7	4.2	24	23, 24, 25
25	3.64	m	-	72.2	26	24Me, 26
26	1.85	m	-	38.0	24, 26	25, 26Me, 27
26 Me	0.81	d	7.0	8.2	26	25, 26, 27
27	3.70	m	-	76.3	26, 28	25, 26, 28, 29
28	1.73	m	-	35.6	27, 28Me, 29ab	28Me, 29
28 Me	0.72	d	4.1	15.7	28	27, 28, 29
29a	2.68	m	-	46.2	28	28, 30
29b	1.60	m	-	46.2	28	28 30
30	-	-	-	134.8	-	-
30 Me	1.61	m	-	16.1	-	28, 30, 31
31	5.16	t	6.8	129.6	29ab, 32	29, 30Me, 32
32	2.03	m	-	22.4	31, 33	30, 31, 33
33	0.96	t	7.1	14.9	32	31, 32

Table 20: NMR Spectroscopic data of the spirangien F (500 MHz, methanol-d₄), for better comparison to spirangien A and B the same numbering ^[230] is use, therefore C-12 and C-13 do not exist.

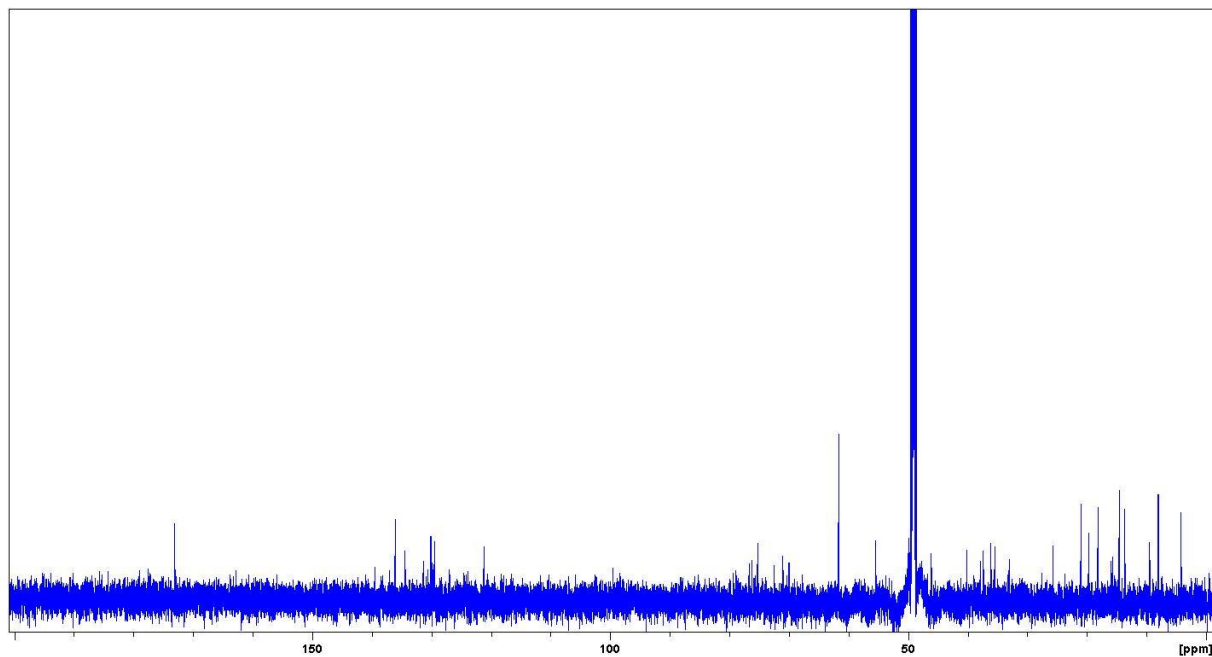
No.	¹ H [ppm]		<i>J</i> [Hz]	¹³ C [ppm]	H-H-COSY	HMBC
1	-	-	-	174.9	-	-
2	2.51	d	6.7	43.4	3	1, 3, 4
3	4.60	q exp. dt	6.4	70.0	2, 4, 5	1, 2, 4, 5
4	5.77	dd	6.4, 14.9	137.0	3, 5, 6	2, 3, 5
5	6.81	dd	11.7, 14.8	127.0	3, 4, 6, 7	3, 6, 7
6	5.99	t exp. dd	11.1	129.9	5, 7	4, 5, 7, 8
7	6.09	t exp. dd	11.1	131.3	6, 8	5, 9
8	6.73	dd	11.7, 14.2	129.5	7, 9	6, 7, 9
9	6.76	t exp. dd	11.3	130.6	8, 10	7, 8, 10, 11
10	6.15	t exp. dd	11.3	130.1	9, 11	8, 9, 11
11	5.62	t exp. dd	10.4	134.5	9, 10, 14	10, 14, 14Me, 15
14	2.82	m	-	36.2	11, 14Me, 15	10, 11, 14Me
14 Me	1.11	d	7.0	19.6	14	11, 14, 15
15	3.54	d exp. ddd	1.7, 10.1	76.8	14, 16	11, 14, 14Me, 16, 17
16	1.61	-	-	40.2	15, 16Me, 17	15, 16, Me
16 Me	0.72	m	-	9.5	16	15, 16, 17
17	3.60	m	-	75.4	16, 18	15, 16, 16Me, 18, 18Me
18	1.49	m	-	32.3	18Me	17, 19
18 Me	0.78	d	5.1	18.4	18	17, 18, 19
19	1.55	m	-	29.8	18, 20ab	17, 18, 20, 21
19b	1.55	m	-	29.8	18, 20ab	17, 18, 20, 21
20a	1.65	m	-	37.4	19, 20b	19, 21
20b	1.48	m	-	37.4	19, 20a	19, 21
21	-	-	-	98.1	-	-
22a	1.70	m	-	38.0	22b, 23	21, 23
22b	1.41	m	-	38.0	22a, 23	21, 23
23	3.66	m	-	78.6	22a, 24	24, 23OMe
23 oMe	3.32	s	-	55.5	-	23
24	2.12	m	-	32.9	23, 24Me, 25	24Me
24 Me	0.76	d	6.6	4.3	24	23, 24, 25
25	3.65	m	-	72.2	24, 26	23, 24, 24Me, 26, 27
26	1.84	m	-	37.8	25, 26Me, 27	25, 26Me
26 Me	0.80	d	7.1	8.7	26	25, 26, 27
27	3.70	m	-	76.2	26, 28	25, 26, 28, 28Me, 29
28	1.72	m	-	35.6	29ab, 27	27, 29
28 Me	0.71	m	-	15.6	28	27, 28, 29
29a	2.66	m	-	46.1	31, 29b, 28	28, 28Me, 30
29b	1.63	m	-	46.1	29a, 31	28, 28Me, 30
30	-	-	-	136.1	-	-
30 Me	1.61	m	-	16.0	31	29, 30, 31
31	5.22	q	6.4	121.2	29, 32	29, 30Me, 32
32	1.59	m	-	13.6	31	30, 31
33	-	-	-	-	-	-

Table 21: NMR Spectroscopic data of the spirangien E (500 MHz, methanol-d₄), for better comparison to spirangien A and B the same numbering ^[230] is use, therefore C-12 and C-13 do not exist.

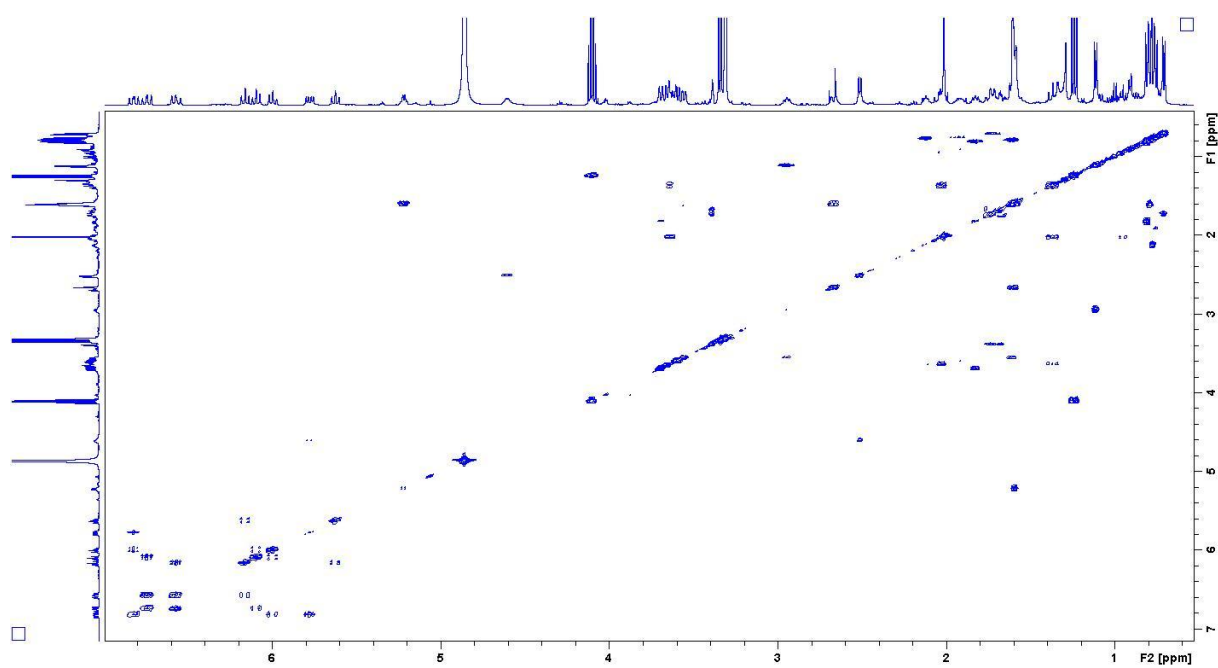
The cosmid C4L22 carrying the spirangien gene cluster of the So ce90 wild type was used as template DNA for the PCR amplifications. Hybridization of the southern membrane of the mutant (So ce90 BF3AKO) with each probe was performed under stringent condition (68 °C). PCR analysis of the cosmid B:P6: to test whether B:P6 carry the 3' end of the gene cluster a PCR amplification of the KS encoding gene (952 bp) in module 9, and TE region (576 bp) was performed using BF_SH1, and BF_SH9 and BFST1 and CFST2, respectively (Supplementary figure 103). The cosmid sequencing was done by M. Scharfer and the identification was performed by S. Rachid therefore it is not further described in details.

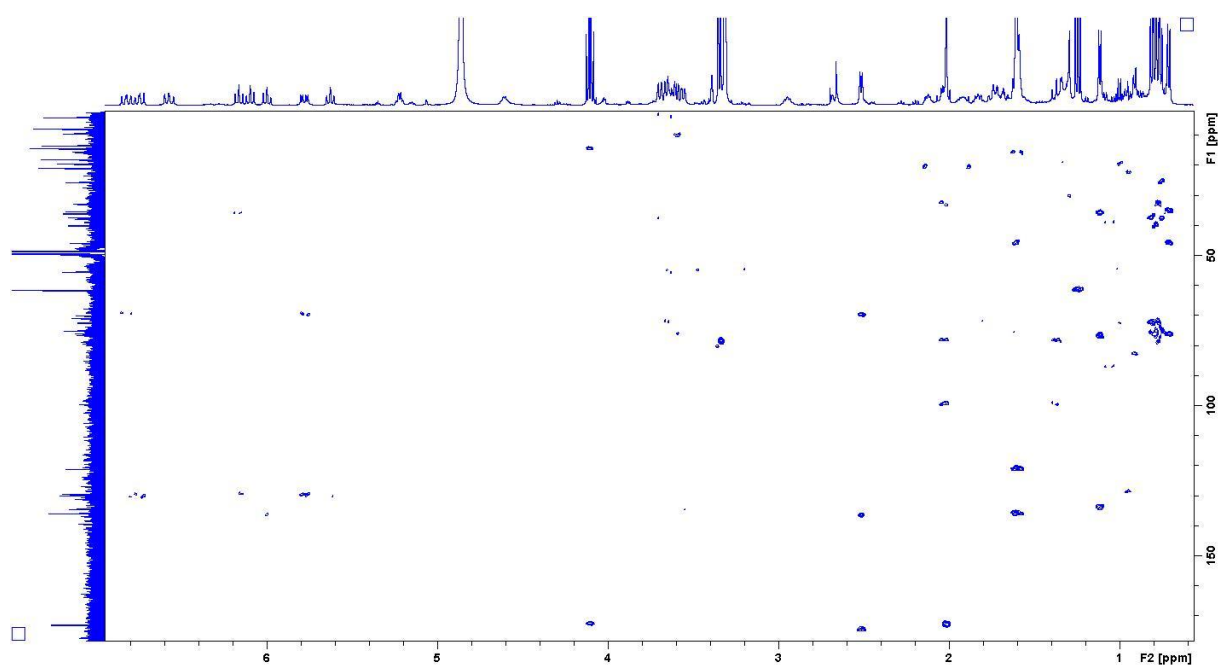
Supporting Information

Supplementary figure 86: ^1H spectrum of spirangien C in methanol- d_4 .

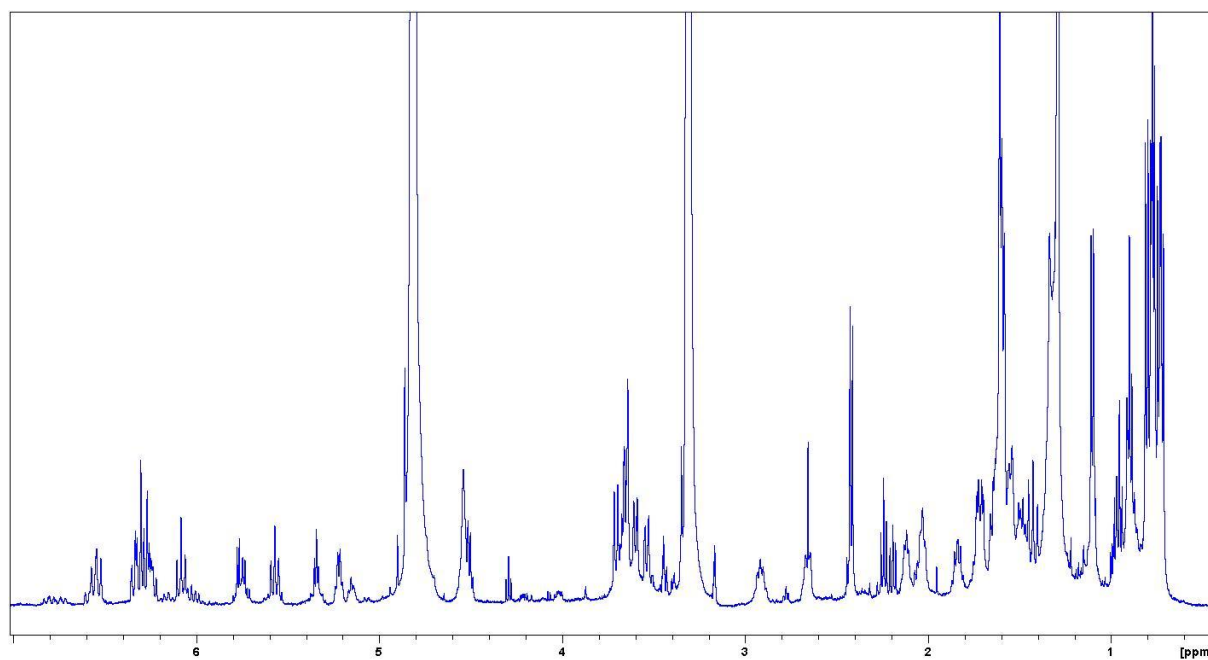


Supplementary figure 87: ^{13}C spectrum of spirangien C in methanol- d_4 .

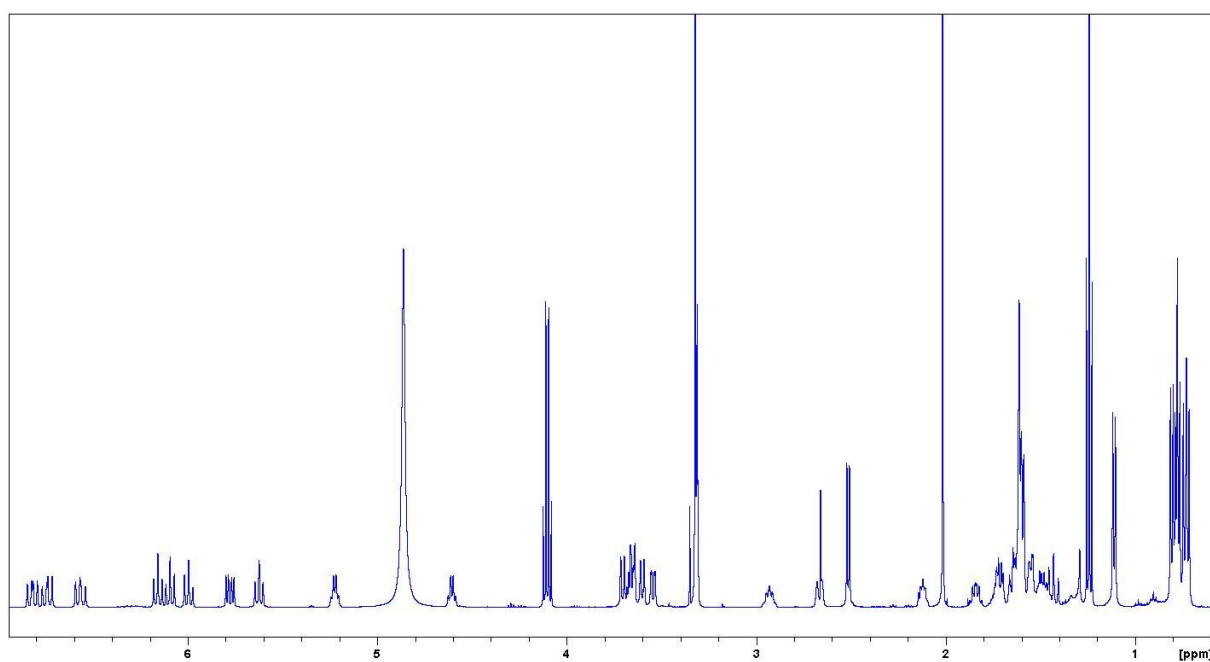




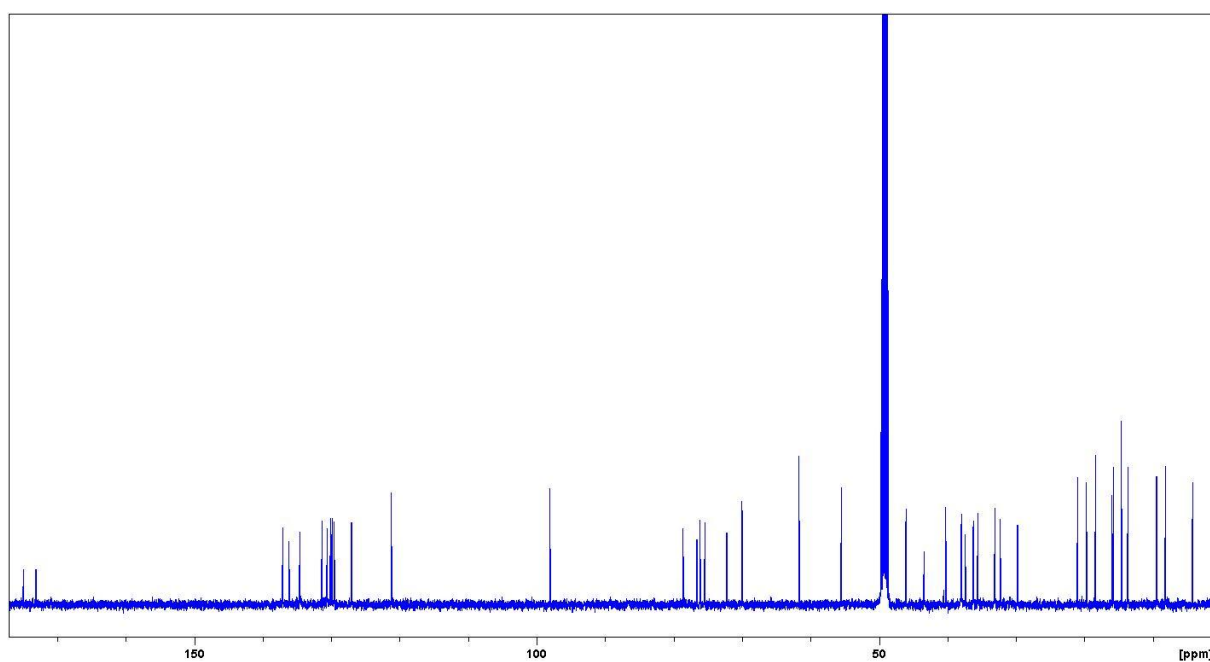
Supplementary figure 90: HMBC spectrum of spirangien C in methanol- d_4 .



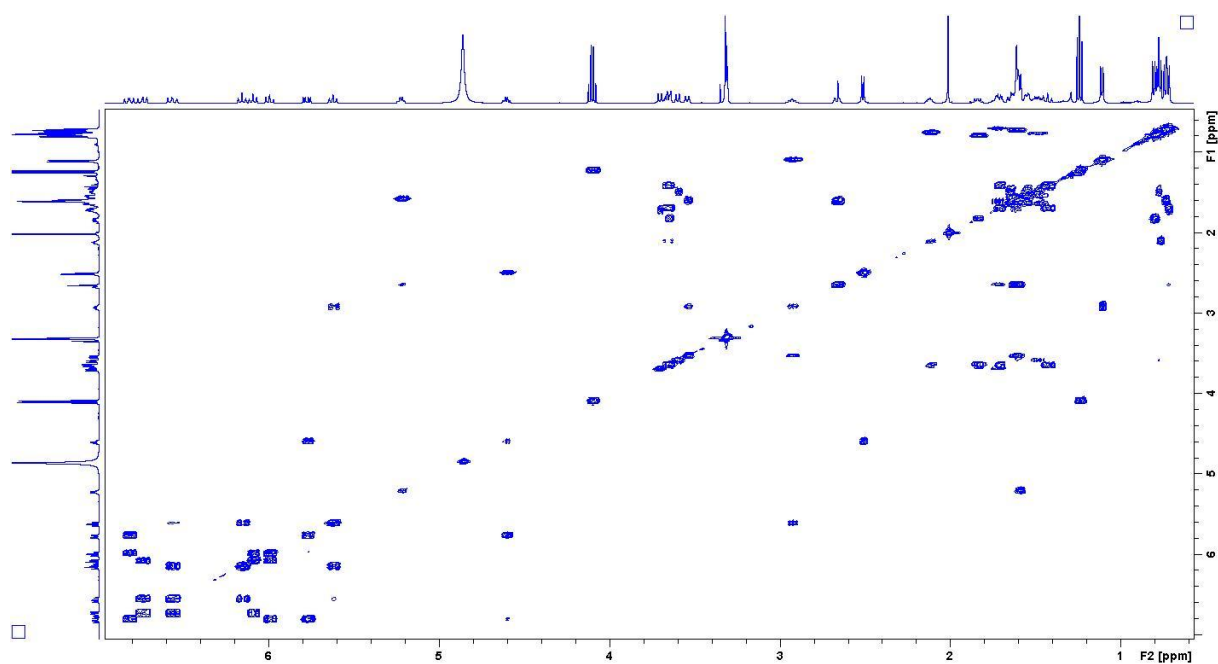
Supplementary figure 91: ^1H spectrum of spirangien D in methanol- d_4 .



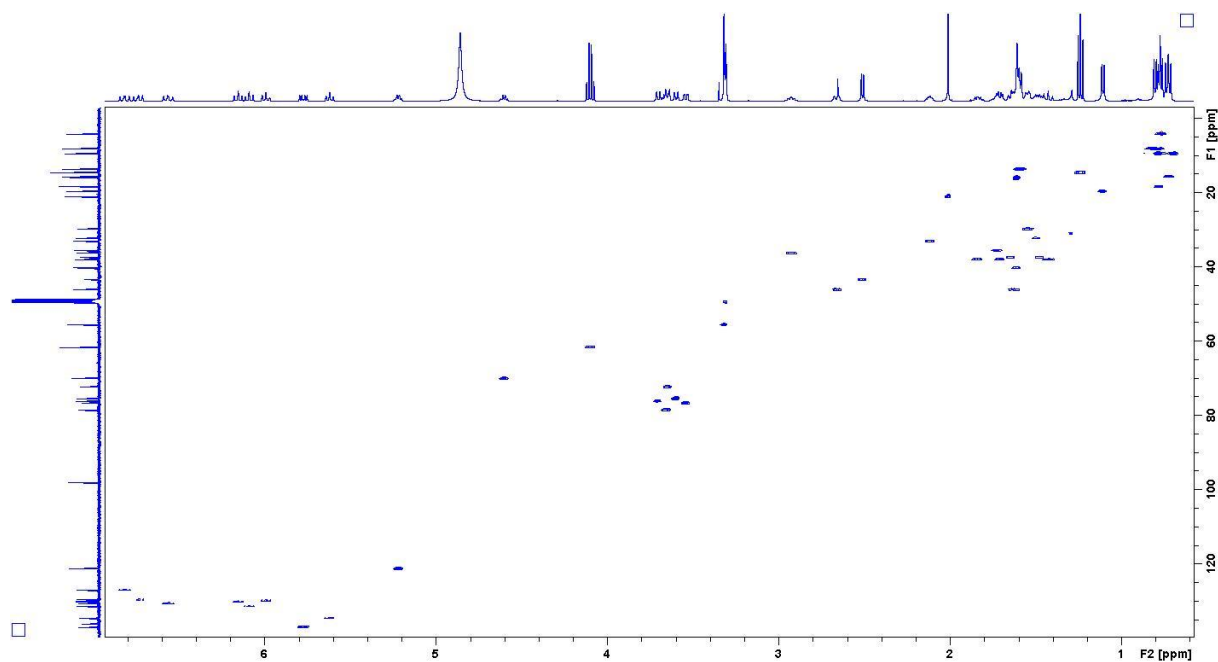
Supplementary figure 92: ^1H spectrum of spirangien E in methanol- d_4 .



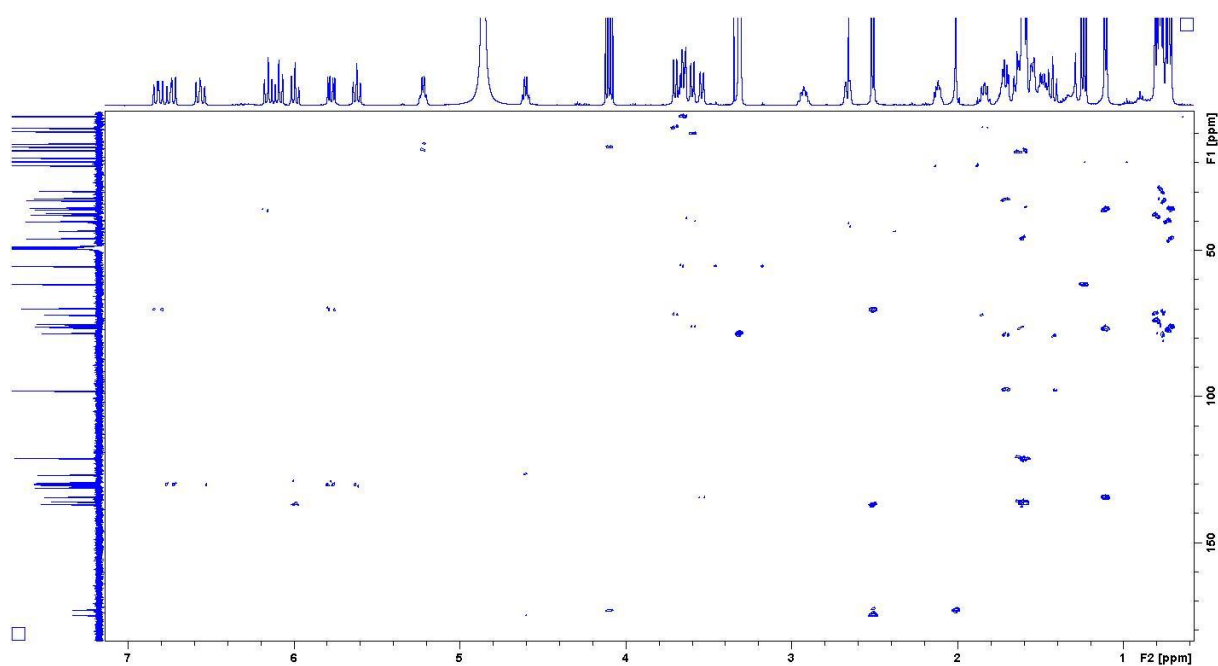
Supplementary figure 93: ^{13}C spectrum of spirangien E in methanol- d_4 .



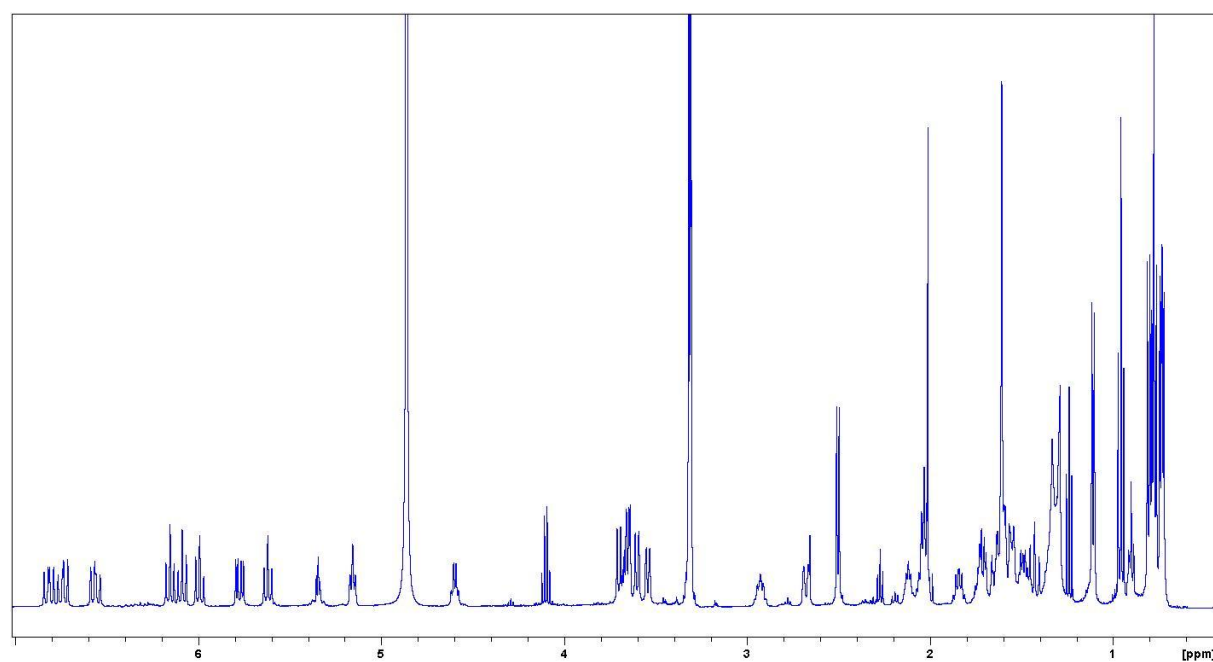
Supplementary figure 94: COSY spectrum of spirangien E in methanol- d_4 .



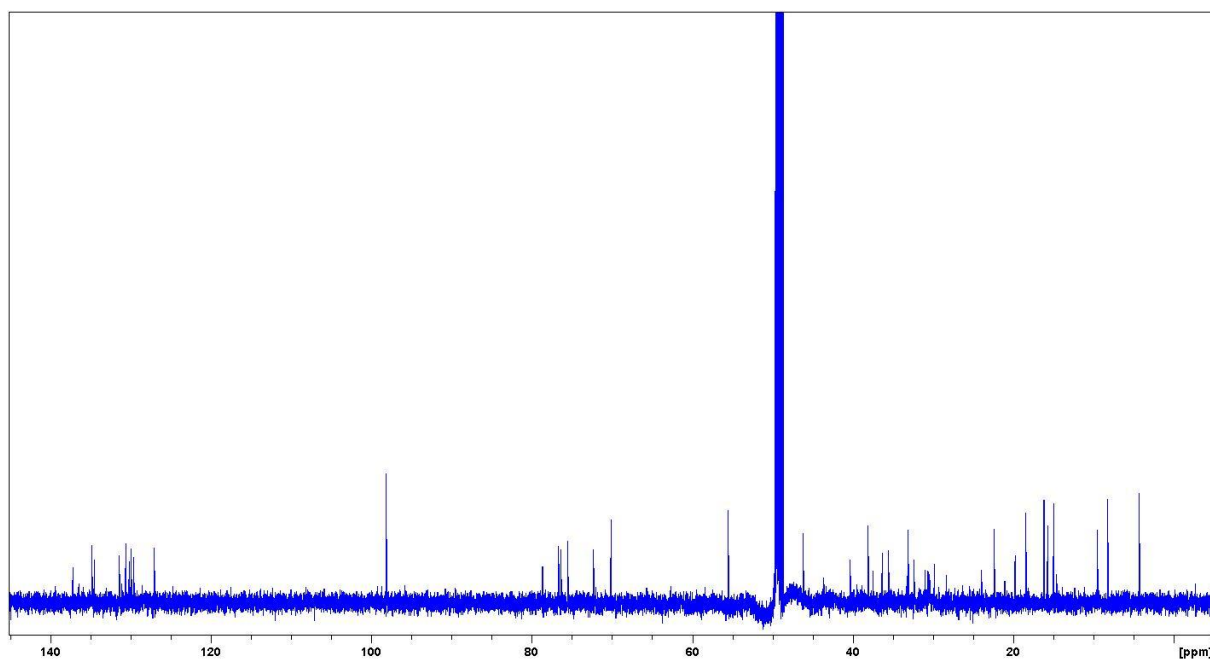
Supplementary figure 95: HSQC spectrum of spirangien E in methanol- d_4 .



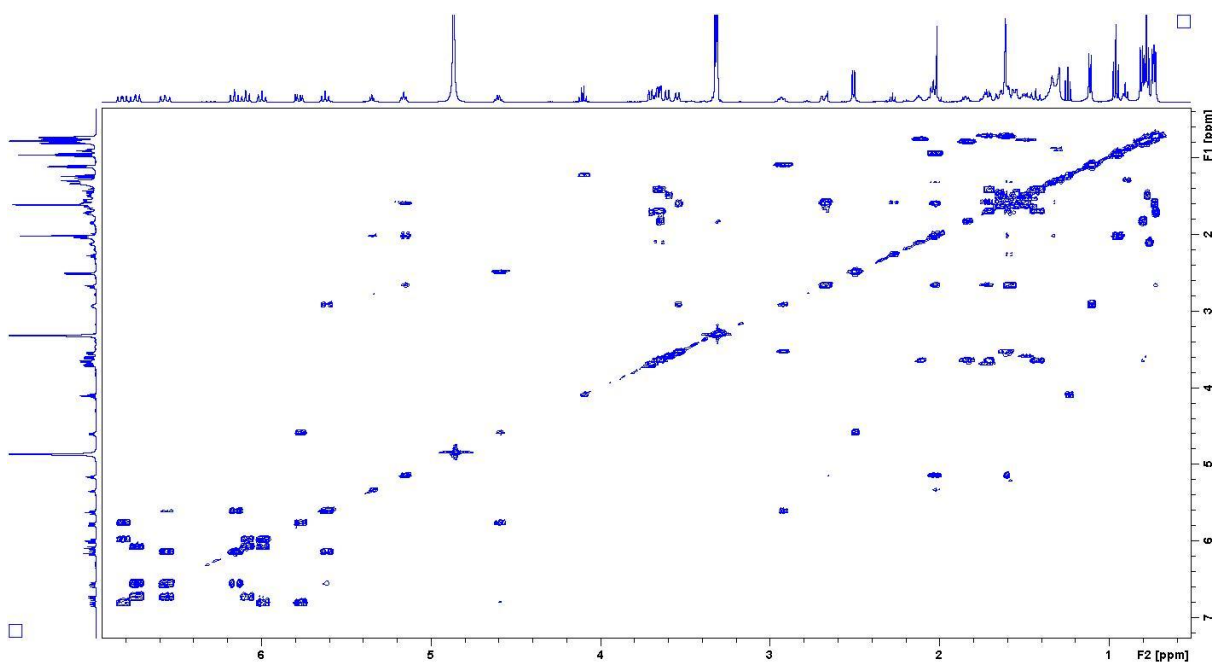
Supplementary figure 96: HMBC spectrum of spirangien E in methanol- d_4 .



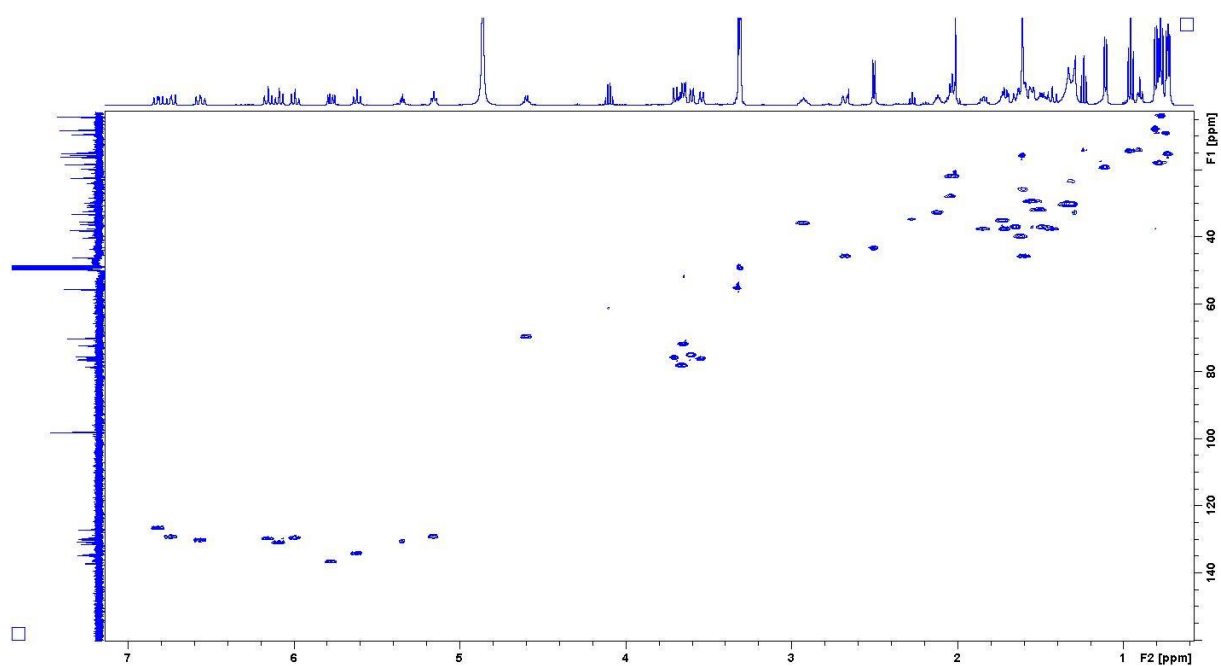
Supplementary figure 97: ^1H spectrum of spirangien F in methanol- d_4 .



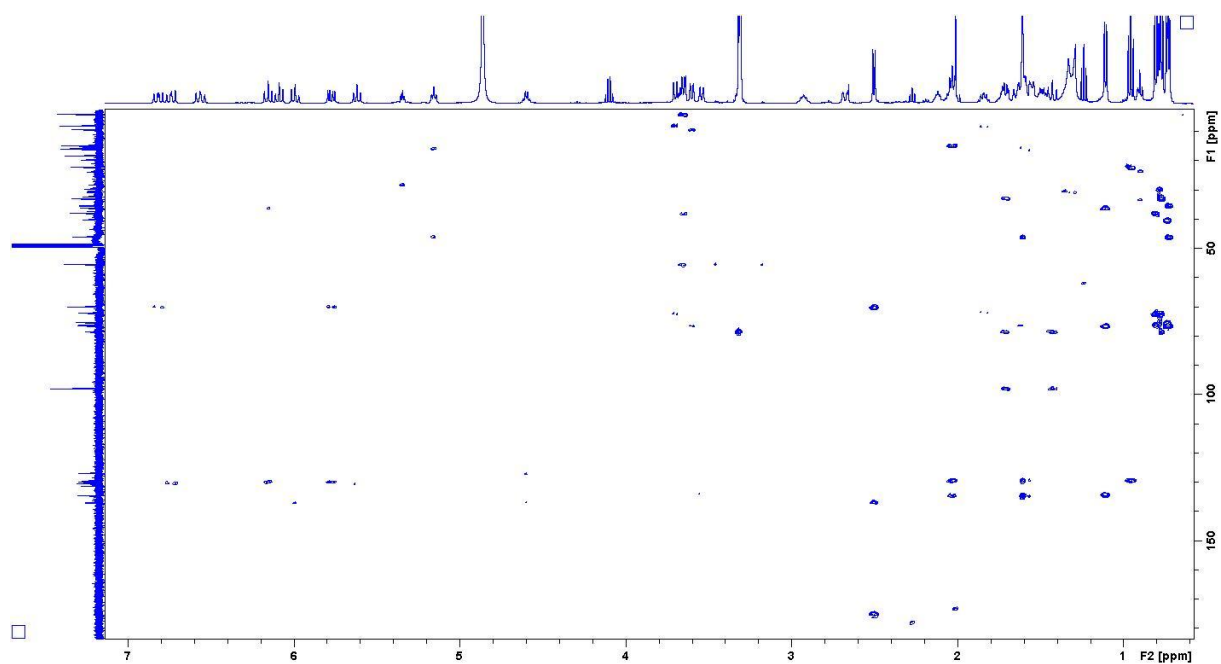
Supplementary figure 98: ^{13}C spectrum of spirangien F in methanol- d_4 .



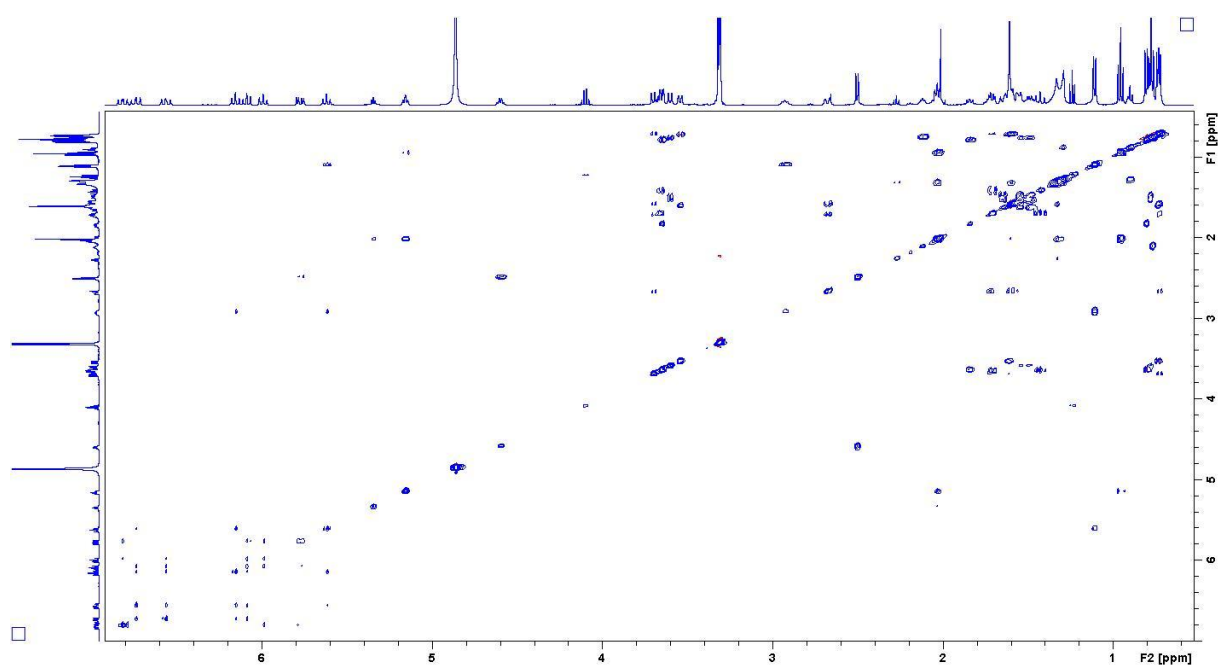
Supplementary figure 99: COSY spectrum of spirangien F in methanol- d_4 .



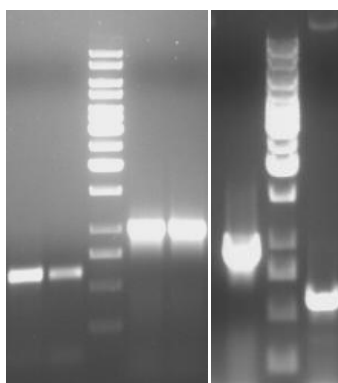
Supplementary figure 100: HSQC spectrum of spirangien F in methanol- d_4 .



Supplementary figure 101: HMBC spectrum of spirangien F in methanol- d_4 .



Supplementary figure 102: TOCSY spectrum of spirangien F in methanol- d_4 .



Supplementary figure 103: left: Dig-labeled PCR fragments of KS (lane 1 and 2), and TE fragment (lane 4 and 5). Lane 3 is DNA marker. Right: PCR analysis showing generation of a KS fragment (lane 1), and TE fragment (lane 3) using B:P6 as DNA template.

Discussion

1. Biosynthetic pathways

Bacteria utilize several different biosynthetic pathways for the production of secondary metabolites. The present work discusses biosynthetic pathways based on PKS / NRPS mechanisms on the one side and a highly unusual ribosomal pathway on the other side. In this thesis the PKS biosynthesis is represented by spirangien, while myxoprincomide and chondrochloren are a result of a PKS–NRPS hybrid systems. In contrast to that, the tetrapeptide cittilin is produced by a ribosomal pathway.

While PKS, NRPS and hybrid NRPS-PKS biosynthetic pathways are commonly known as biosynthetic routes for natural compounds of myxobacteria, ^[45;49] the ribosomal production of a secondary metabolite is proven here for the first time from that order. Various ribosomal biosynthetic pathways are known from other bacteria; the class of bacteriocins ^[146;147] is known since long for example, but recently more and more other ribosomal secondary metabolites are being published. ^[156;157]

In the last decade the methodology for identification and characterization of secondary metabolite biosynthetic pathways changed dramatically due to the availability of whole genome sequence information. More and more bacterial genomes are deciphered and bioinformatic analysis reveals the presence of secondary metabolite biosynthetic pathways. Astonishingly the number of known natural products (NP) is often much lower than the genetic capacity of the organism. ^[97;204] Thus, the discovery of novel NP from genetically proficient organisms currently constitutes a bottleneck, rather than obtaining the sequence information for a biosynthetic gene cluster.

For detection and functional analysis of biosynthetic genes, bioinformatic tools have been developed which are able to perform the annotation of biosynthetic pathways in genomes routinely. However, knowing the gene cluster does not automatically mean that the corresponding metabolite can be identified. Particular if the biosynthesis does not conform to the so called textbook biosynthetic model (described in the introduction). In some cases the biosynthesis follows the colinearity rule where each module is performing one function, but often biosynthesis in myxobacteria does not follow this model. ^[97] Frequently *in silico* analysis alone is not able to predict the final product from its genes, like in the case of the myxoprincomides.

2. Structure and biosynthesis of cittilin

Cittilin is a common substance for *M. xanthus*, since 69 strains of 98 various strains from differing locations worldwide are producing it.^[51] The biological function for the bacteria is not clear as an antibiotic or antifungal activity has not been found. Nevertheless cittilin B is medically interesting as it has been described as a neurotensin antagonist (neurotensin receptor from guinea pigs $IC_{50} = 30 \text{ mg / mL}$).^[170;172] Such antagonists are useful for treating depression, psychoses, Alzheimer and Parkinson diseases. Therefore RP-66453 which is structurally identical to cittilin B was patented by Rhone-Poulenc in 1995. The company used it as a lead structure for the development of antipsychotic pharmaceuticals. Chemical modification at the structure ended with the discovery of potent neurotensin antagonists.^[244]

The two aryl crosslinks between the tyrosine residues are the characteristic features of the cittilin tetrapeptide. The carbon - carbon bond together with the ether connection of phenyl rings are known from a small number of other natural compounds. The most famous example is vancomycin where both types of connections are present. Same structural features are known from teicoplanin (**52**)^[245], balhimycin^[178], kistamine^[179], and chloropectin^[180], too. Structures with the existence of only an ether bridge between two tyrosine rings are K 13 (**54**), deoxybouvardin (**55**), aceroside IV (**56**) and the structures of the RA family (for example RA VII (**53**), it is an anti – tumor agent)^[246]. Endo aryl-aryl bonds exist in many bioactive macrocycles like biphenomycin (**57**) and acerogenin K (**58**). Biphenomycin B is highly potent antibiotic against Gram-negative and β -lactam-resistant microorganisms. It was isolated from *Streptomyces griseorubiginosus* (NO. 43708) culture.^[247;248] These compounds demonstrate that aryl - aryl connections as present in cittilin are employed by different organisms for diverse structures.

From the chemical point of view the synthesis of such complex aryl - aryl architecture was a challenge. The synthesis of vancomycin for instance led to numerous impetuses in the development of synthetic methodologies.^[249] In the case of cittilin B the total synthesis solved its long time unknown absolute configuration.^[182-185]

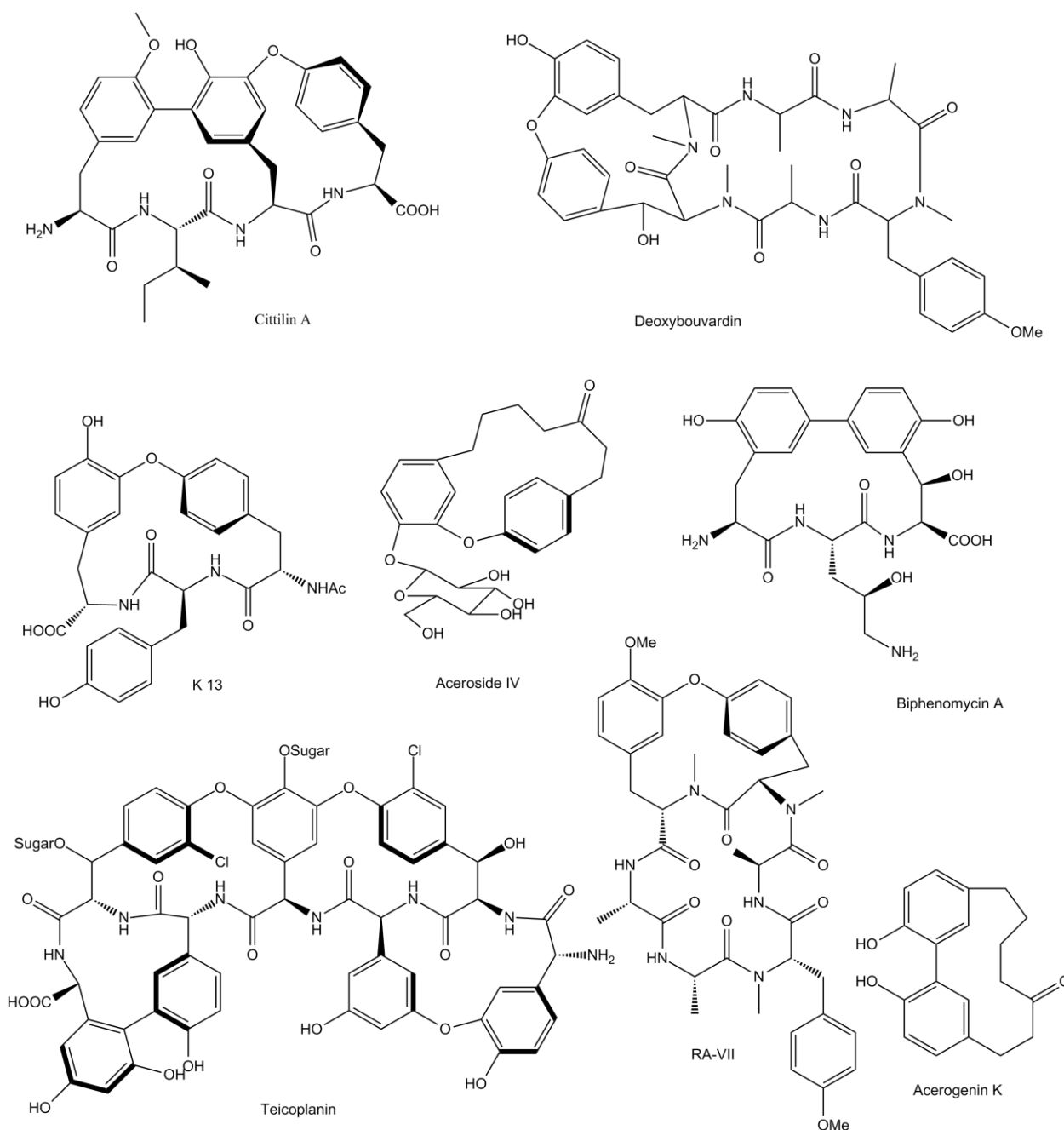


Figure 77: Cittilin, teicoplanin (52), RA VII (53), K 13 (54), deoxybouvardin (55), aceroside IV (56), biphinomycin A (57), acerogenin K (58).

The biosynthesis for the NRPS product vancomycin was well studied, too. Here the aryl - aryl connections are built by different cytochrome P450 enzymes. Each linkage is performed by a specialized P450 enzyme (further details see cittilin chapter).^[181] In contrast to vancomycin cittilin is not a NRPS product it is built by a ribosome-dependent pathway. For both, the ribosomal and the NRPS pathway there are enzymes which are able to modify the peptide core structures. These modifying enzymes of the ribosomal biosynthesis are similar to NRPS post modifying enzymes.^[194] However definitely more NRPS than ribosomal secondary

metabolites are known. But frequently natural peptides are thought to be of nonribosomal and not of ribosomal origin.^[144;152;155;162] After isolation cittelil was presumed to be an NRPS product, too. And therefore a search within the genome for typical NRPS A-domains with tyrosine and isoleucine specificities was performed. Only the lack of them led to the discovery of the encoding precursor peptide sequence.

Ribosomal peptide encoding sequences, like the one of cittelil, can be very short and are hard to discover in the course of routine secondary metabolite gene cluster analysis as such sequences do not need to be within large operons. Cittelil for instance is inside of an open reading frame which is only 231 base pairs long. Up to date the core peptide and the modifying enzymes of the cittelil gene cluster are identified. A signal peptide and a recognition sequence are not present or not verified so far. The exact length and the function of the leader peptide are not known but leader peptides often differ for each ribosomal peptide.^[145] Furthermore the backbone ring closure of cittelil is unusual; it is not like in the case of the patellamides where the first amino acid is connected with the final one, to close the ring. And the ring closure is not similar to the lantibiotic ring formation.^[145;148] There, cysteine threonine and serine residues are used for cyclization. Lantibiotic cyclization happens in two sequential steps. At first threonine or serine are dehydrated. In a second step the dehydroxylated group is linked with a thiol group of a cysteine. Both processes are performed by a bifunctional synthetase LanM (refer to introduction chapter).^[190] In contrast to lantibiotics and patellamides, the cittelil rings are formed by a post-translational P450 enzyme. Involvement of the P450-dependent oxygenase in cittelil biosynthesis was unambiguously confirmed by construction of its oxygenase knockout mutant, which showed lack of cittelil production. Such P450 enzymes are implemented in the biosynthesis of vancomycin. There the C-C connections are build by two P450 enzymes (OxyA and OxyB) while OxyC, another P450, is in charge of the phenolic coupling.^[181;189]

Except for the modifying enzymes cittelil biosynthetic pathway is a promising system for a genetic library by site directed mutagenesis of the encoding sequence by PCR. By simple mutagenesis in the heterologous expression construct novel cittelil derivatives were generated. A similar approach was used by Schmidt et al. in a study with the patellamides, demonstrating that it is possible to create a peptide library by mutating the encoding sequence by PCR.^[143] But with the cittelil system even the obtained peptides are further modified. In ribosomal systems structural variation can be made easily compared to NRPS ones. For NRPS machineries such library creation is very difficult and only possible under high effort. Here complete A domains have to be exchanged / created, which then have to be able to interact

with the other domains / modules. But the domains are encoded in huge protein sequences and not in the three letter code for one amino acid like in the case of the ribosomal biosynthesis.

To further increase the diversity of the citteliln derivatives by site directed mutagenesis the core ring system could be changed. This is only possible if the N and C terminal sequences are not necessary for the recognition or the processing for the modifying enzymes. If so it is likely that it is even possible to produce a pentapeptide or a multi-peptide instead of the tetrapeptide, although the exact role of the precursor peptide is not known. Furthermore the P450 might not accept the pre-citteliln as substrate if one of the tyrosines is exchanged. However for all these modifications it looks like if the ABC transporter system is present the production rate is higher, even if the exact function could not be clarified, this should be addressed in further studies.

To conclude, the work presented in the citteliln chapter, revealed that myxobacteria are using their ribosomal machinery, beside the typical described PKS and NRPS pathways, for the production of secondary metabolites.

Citteliln is the first known myxobacterial representative for which the biosynthetic machinery could be shown to rely on a ribosomal mechanism. There is a number of other recently published secondary metabolites from other bacteria which are produced via ribosomal pathways like the patellamides ^[144], the microviridins ^[157;160] and the cyanobactins ^[156;157]. As more and more ribosomal peptides are identified from other bacteria it looks most likely that further interesting substances are as well built in ribosomal manner by myxobacteria. To increase the body of knowledge of ribosomal secondary metabolites further ribosomal pathways should be analyzed.

3. Myxoprincomides of *Myxococcus xanthus* DK1622 and DK897

The myxoprincomides are novel secondary metabolites produced by the *M. xanthus*. Two different types of myxoprincomides were isolated, one structure type from DK897 and the other one from DK1622. In the screening of 98 *M. xanthus* strains every strain produces myxoprincomides,^[51] only the type varies, either the DK1622 type or the DK897 variants are present. 73% of the screened strains produced the DK897 type of myxoprincomides (at least one of the 10 compounds was detected). While the DK1622 myxoprincomide variant (c506) was found in 36 strains. But both types are always produced in very small amounts. The biological function of the myxoprincomides remains elusive so far. In general the range of biological activities from known myxobacterial secondary metabolites is wide. Chondrochloren produced by *Chondromyces crocatus* Cm c5 for example shows weak anti-bacterial and anti-fungal activity.^[77] Spirangien a secondary metabolite of *Sorangium cellulosum* So ce90 is cytotoxic and anti-fungal.^[230] For *M. xanthus* secondary metabolites diverse biological activities have been reported, too. Cytotoxic effects for example are described from myxochromides^[85] while the myxalamides^[82] are acting anti-fungal. Anti-bacterial activity is known from myxovirescin [84] and myxochelin^[83]. The later is a siderophore which has as well antiviral activity. Furthermore the above discussed cittilin is described as neurotensin antagonist.^[170;172] In the case of the myxoprincomides, no detectable antifungal, antibiotic or cytotoxic function was found. But as every *M. xanthus* strain produces myxoprincomides its role should be significant for the producing organism. It might be something which it is normally not tested and hard to detect for instance like in the case of the DKxanthenes which play an important pheromone-like role for sporulation.^[86;196] Roles in microbial differentiation or function that are related between primary and secondary metabolism are often difficult to identify. Hence it is problematic to demonstrate competitive advantage to the producer.

At the beginning of the thesis the myxoprincomides of DK897 and DK1622 have been run as different projects, as the biosynthetic cluster of DK897 was not known. Therefore no correlation between the two clusters was obvious. Furthermore the masses and mass fragments of the myxoprincomides and of course the strains were different. During the ongoing process and the finish of the genome sequencing of DK897, it became more and more obvious that the biosynthetic machinery and as well the substances are close relatives.

But the strategy of correlating the biosynthetic genes to the compound was completely different.

In the case of DK987 first no genome information was available. Therefore a transposon mutagenesis was performed and the obtained transposon mutants were screened by LC-MS for lacking myxoprincomide production. Here approximately 2000 mutants were screened to identify the gene metabolite correlation. While for DK1622 the complete genome was known and a targeted mutagenesis was used to inactivate potential biosynthetic gene clusters. The comparatively low number of created mutants were screened for lack in production of any possible compound. In this approach the discovery of novel secondary metabolites from genome sequenced strain has been assisted by modern analytical techniques, combined with statistical tools to reduce manual efforts. And targeted mutagenesis experiments led to the correlation of novel compounds. Gene inactivation studies have been merged with liquid chromatography coupled to high resolution mass spectrometry (LC-HR-MS) analysis. The metabolome has been analyzed with principal component analysis (PCA). Yielding in statistic findings which point at differences between wild type and mutant strains. Manual comparison of LC-MS data would be very tedious, and failed before to assign novel metabolites. The huge amount of data including media-derived signals, low abundance of compounds make it impossible to detect any novel substance produced in trace amounts manually. Here implementation of statistical tools is a necessary assistance.

After confirmation of the gene metabolite correlation the gene cluster could be analyzed in silico and a substance isolation procedure was started. But in the event of DK897 a transposon recovery of the non myxoprincomide producing clone was carried out. Followed by subsequent plasmid sequencing to obtain first biosynthetic gene cluster information. That means to assign the myxoprincomide gene cluster in DK897 an older strategy was applied. In the case of the genome sequenced strain DK1622 the myxoprincomides and the correlating gene cluster have been detected with a novel approach, the “secondary metabolome mining”.

For DK1622 18 biosynthetic gene clusters (PKS, NRPS and hybrids thereof) have been located but only five of them were correlated to a natural product (at the beginning of this thesis).^[90;195;206-209] But in an evaluation it was found out, that at the transcriptomic and proteomic levels 13 gene clusters are active under laboratory conditions.^[89;90] The high number of unassigned biosynthetic gene clusters in DK1622 translates into a resource of novel natural products which should be exploited. Here three novel classes of secondary metabolites have been detected and connected to genes. With the discovery of the cittilin

biosynthetic genes, a cluster which is not NRPS or PKS based, in total 9 secondary metabolites of *M. xanthus* DK1622 are correlated to their genes.

However the genes revealed that there have to be more unknown compounds in the bacterial extracts. One possibility for not detecting them is that the products of these none correlated biosynthetic routes might be produced early and degraded during the life cycle of *M. xanthus*. From *Sorangium cellulosum* So ce56 such a case (chivosazol) has been observed.^[88] Chivosazol production can be detected in So ce56 after 100 h of growth. Further accumulation follows for further 200 h then the degradation process is faster than the chivosazol production.^[250] Similar cases could be investigated by a kinetic study involving harvesting and measuring of samples at different time points. Another factor might be a weak stability of the natural products. Furthermore a problem can be the adsorption characteristics with regards to the XAD-16 adsorber resin. Or if the metabolites do not have UV absorption and are very poorly ionizing it would be a detection problem.

Assuming that the molecules are present in the bacterial extracts it would be a matter of identification within the extract. Small quantities (under laboratory conditions) or low abundance at all may make their detection difficult. This means that the bottleneck is the analytical discovery of the secondary metabolites. It is an analytical challenge to solve this difficulty. To overcome this problematic the “secondary metabolome mining” approach has been developed (refer to chapter 2B).^[176]

The myxoprincomide structure shows several rare and unprecedented features beside common amino acids very uncommon structural elements are incorporated. For instance either (2*S*,3*S*)-3-amino-2-hydroxy-5-methylhexanoic acid (**59**) or (2*S*,3*S*)-3-Amino-2-hydroxy-4-methylpentanoic acid (**60**) are part of the myxoprincomides. Alternatively there is the diketo variant of **59** and **60** ((*S*)-3-amino-5-methyl-2-oxohexanoic acid (**62**) and (*S*)-3-amino-4-methyl-2-oxopentanoic acid (**63**)). An α -keto-valine building block is known from the structures of oriamide^[251], cyclotheonamides^[252], keramamides^[253] and scleritodormin^[254] which are all marine natural products. **59** is known from the structure of amastatin isolated from a streptomyces species.^[255-257] Beside that it is present in several synthetic derivatives of natural compounds.

Another uncommon part is (*S*)-Hydroxy-valine or (2*S*)-2-Amino-3-hydroxy-3-methylbutanoic acid (OH-Val; **61**). It is twice inside the structure of the long myxoprincomides of the DK897 type and only once in the DK1622 type. It occurs in nature as single amino acid^[258] or as glycerin ether^[259]. Therefore it is likely that no oxidation domain for the oxidation of a valine

is needed within the corresponding modules. **61** is part of the natural compounds zorbamycin^[260], luzopeptin^[261], and yaku'amides^[262]. Furthermore it is known from the myxobacterial compound myxovalargin^[76].

The fourth unusual amino acid is β -lysine. As secondary metabolite component it has been described within the structure of the antibiotics viomycin^[263] and nourseothricin^[264].

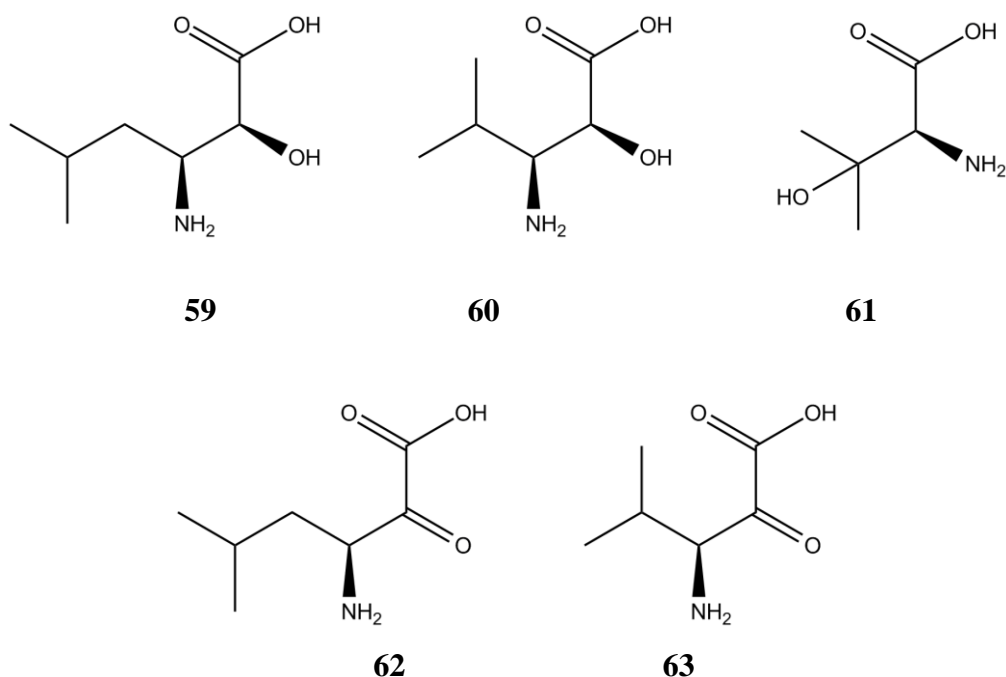


Figure 78: Structures of (2S,3S)-3-amino-2-hydroxy-5-methylhexanoic acid (59), (2S,3S)-3-Amino-2-hydroxy-4-methylpentanoic acid (60), (2S)-2-Amino-3-hydroxy-3-methylbutanoic acid, (S)-3-amino-5-methyl-2-oxohexanoic acid (62) and (S)-3-amino-4-methyl-2-oxopentanoic acid (63).

All these striking features reveal once again, that myxobacterial biosynthetic pathways often do not follow the textbook biochemistry.^[97] Furthermore it is typical for myxobacteria to produce a large family of metabolites by several variations during biosynthesis.^[45] From the evolutionary point of view such diversity of structures generated by one biosynthetic pathway is called as the ‘screening hypothesis’ meaning to produce different compounds at low costs.^[265]

The incorporation of these unusual building blocks makes *in silico* structure prediction nearly impossible. Further deviations from the textbook biochemistry are the missing of domains and the lacking of conserved motifs which might have the consequence of module skipping (refer to section module skipping; later on) which makes analysis even more complicated.

Otherwise such uncommon metabolites like the myxoprincomides are a thrilling occasion to explore unusual building blocks and their biosynthetic routes. One can assume that there are still a significant number of secondary metabolites which needed to be explored.

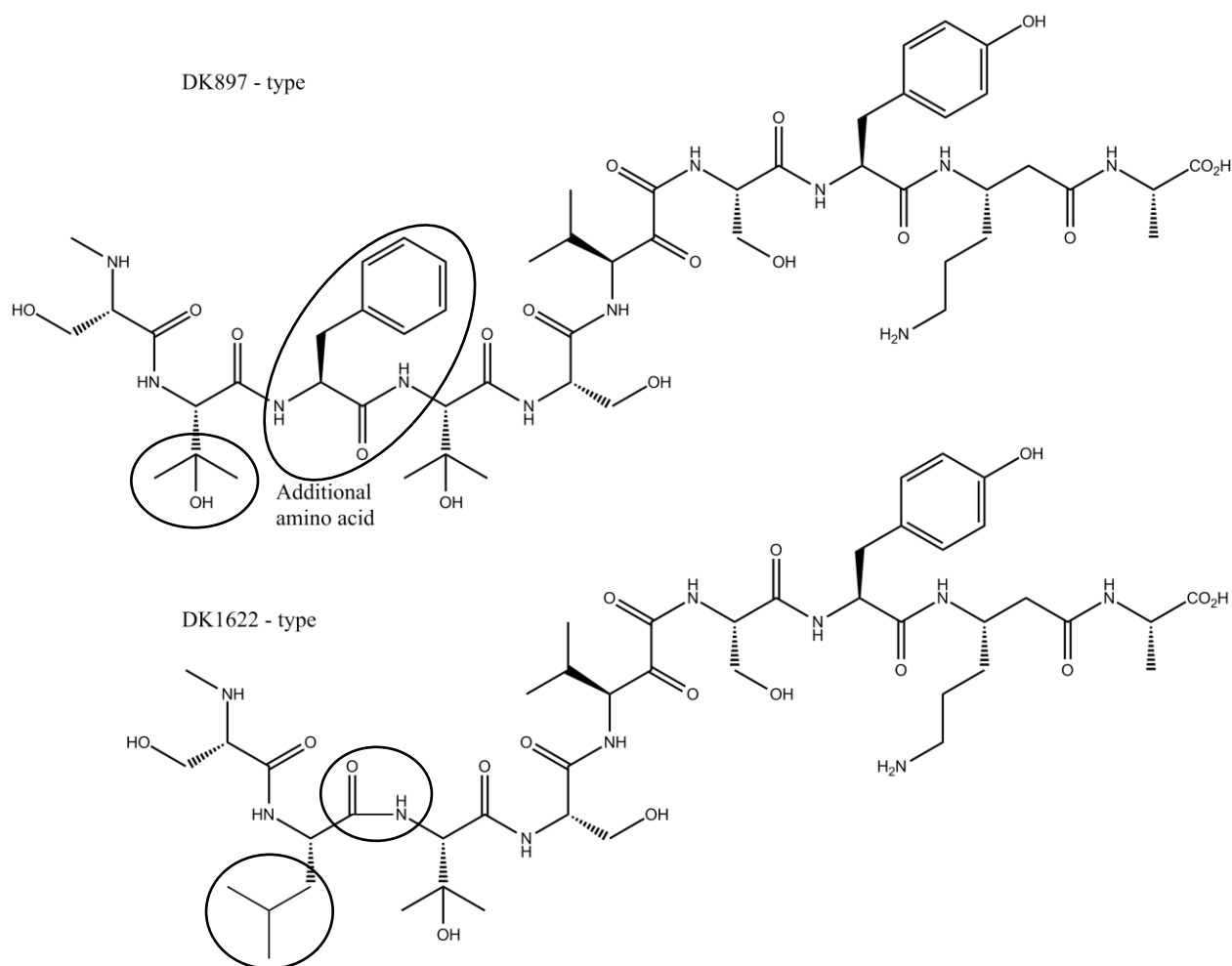


Figure 79: Structures of myxoprincomides in strain DK897 (top) and DK1622 (bottom), differences marked in circles.

The structural differences of the DK897 type of myxoprincomides and the DK1622 variants are depicted in Figure 79. The major disparities are two varying amino acids. One is an additional phenylalanine in the DK897 type of myxoprincomides. And another variation is the second amino acid. The DK897 variants exhibit an OH-Val at this position while the DK1622 types contain a leucine. Despite differences in one amino acid and an extra one, the molecular core structure and the amino acid order are the same.

The structural distinctions of the two myxoprincomide types are reflected in the biosynthetic gene clusters, too. The myxoprincomide gene is 42.8 kb (DK1622) respectively 46.9 kb

(DK897) in size. It consists of 13 (DK1622) or rather 14 (DK897) modules, and ends with a thioesterase domain. *mxp₈₉₇* is the largest myxobacterial biosynthetic gene exhibiting 39 catalytic domains. And in DK1622 its length is unique, too, as it encodes the largest single and continuous hybrid enzyme. Remarkably, the myxoprincomide gene cluster in DK897 has one module more which is in charge of an incorporation of one OH-Val. Furthermore there is an incorporation of one phenylalanine instead of a leucine. This might give insights into the evolutionary development of the biosynthetic gene clusters. Whereas at that point it is not possible to say which of the existing gene cluster was present at first. Either a module deletion took place or a module insertion. Another possibility would be that both gene clusters developed from one antecessor. In general there are different settings discussed in literature for the origin of a gene cluster.^[266] One described possibility is the duplication of existing gene cluster.^[267] Another explanation is the establishment of a novel cluster by relocation of scattered genes within the genome.^[268] The third way is the horizontal gene transfer from one genome to another^[267;269] but this still does not really answer the question how the gene cluster developed in the donor species. If not the complete cluster is from foreign origin single parts of a cluster might have their origin in horizontal gene transfer. The KS domain of module *epoB* in the biosynthesis of epothilone for instance has its origin in horizontal gene transfer.^[270] This might be the case for the extra module in the DK897 myxoprincomide biosynthesis, too.

However in both gene clusters, the PKS module and the following one show novel and to date unexplained biochemistry. Furthermore the assembly line has a broad intermediate acceptance as **62** and **63** and maybe even **59**, **60** are incorporated, if no oxidation of **59** and **60** is performed. It is not clear if some intermediates are further oxidized by oxidation domain in the module which follows the PKS one or if different amino acids are incorporated. The enzymatic machineries apparently tolerate a wide range of intermediates. Otherwise partly the intermediates are released from the assembly before the final product is reached. However, the final results are several intermediates and four different end products build by one single gene cluster.

At this point the premature release from the assembly line is not clear, it could be a spontaneous hydrolytical, or alternatively an enzyme catalyzed release, for instance by a thioesterase (TE). A TE domain (type I) is normally located at the end of the assembly line and it usually cleavages of the full-length product. For this final product release there are two major mechanisms.^[225;271] One possibility, catalyzed by the TE, is the intermolecular release by hydrolysis. Here water acts as nucleophile to release a linear product. This occurs for

example in the biosynthesis of vancomycin ^[9] or in chondrochloren ^[133]. In the latter the pre-chondrochloren (refer to chapter 7) is build before it is further modified by CndG into the final product. ^[134] The second major TE-catalyzed mechanism is the intramolecular release by macrocyclization; represented for example by the biosynthesis of epothilone or erythromycin 6-DEB. Here the hydroxyl (or in other cases the amino group of the peptidyl-acyl-chain) reacts as internal nucleophile to release a macrocyclic product.

Beside these two release mechanisms there are also several others. ^[225] Indanomycin for instance is a TE-less clusters where probably an extra module, containing a cyclase domain, releases the product. ^[272] Furthermore there are unknown release mechanisms where the release method is not obvious and not investigated so far. ^[225] For myxobacteria unusual release mechanisms are common, examples are the biosynthesis of melithiazol ^[273] stigmatellin ^[274] and aurafuron ^[275].

Intermediates which are stuck on the chain due to biosynthetic errors are typically removed by type II TE ^[225] but such a domain is not present in the myxoprincomide cluster. It is more likely that the release of the myxoprincomide intermediates is similar to the one of the pikromycin biosynthesis. There is a directly offload of an intermediate by the terminal TE domain. ^[276] The intermediates must either be passed through the modules to the terminal TE domain without undergoing extension, or directly transferred to the terminal TE domain. In the last case a conformational flexibility is needed to directly offload the diverse intermediates. To understand the release mechanisms deeper analysis or molecular studies are necessary.

Another open question in the biosynthesis of myxoprincomide is the PKS module and the subsequent one. Here a proposal for the DK1622 biosynthesis has been made where the intermediate is extended by a malonyl-CoA. Followed by an alpha oxidation of the acetyl group performed by the oxidation domain of the proceeding module. In this proposal an α -keto functional group is formed by tautomerization. This functional group is nucleophilically attacked by the amino group of the subsequent serine resulting in a loss of CO and a C-1 unit originating from a malonyl-CoA. This mechanism has been supported by ¹³C-labeled acetate feedings. Furthermore different myxoprincomide derivatives with different oxidation stages have been isolated in DK897. There are the alpha-hydroxy variants (**59** / **60**) and the diketone forms (**62** / **63**); both exist with two different amino acid residues.

To conclude, the myxoprincomides constitute a novel group of natural compounds identified from *M. xanthus*. Myxoprincomide production seems to be a characteristic feature of the species *M. xanthus* with at least two different types. Their structure contains unusual features which is typical for myxobacterial biosynthesis as the biosynthetic machineries most frequently do not follow the “textbook” model. It shows once again that *in silico* prediction of the compound`s structure does not function correctly for the complete molecule. ^[97] The specificities could not be predicted in every case and missing conserved motifs and missing domains seem to result in module skipping. The detailed characterizations of such biosynthetic routes are necessary to improve biosynthetic predictions. Beside that myxoprincomide biosynthesis is very fascinating as several intermediates and end products are released during the NRPS-PKS assembly of the myxoprincomide pathway.

Moreover a novel mechanism for the formation of diketo groups in peptide chains has been suggested. Here a C1 unit originating from acetate is incorporated by a PKS module. Oxidation and further tautomerization with subsequent extension of another amino acid leads to a loss of CO and formation of an α -hydroxy intermediate. This can further be oxidized into an α - keto moiety.

The myxoprincomide biosynthetic gene clusters in DK897 and DK1622 are slightly different but both contain several unusual modules which are in charge of this interesting biosynthesis. Two different strategies were applied to identify the compound gene correlation. Furthermore it could be demonstrated, that the ability of connecting biosynthetic gene clusters to its metabolite assists its isolation. As the natural abundance of the myxoprincomides is very low, genetic manipulation was used to create optimized producer strains.

4. Module skipping increases chemical diversity

Some of the reported biosynthetic assembly lines are deviations from textbook biochemistry. Such exceptions made it difficult to predict the final product of biosynthetic assembly lines from the primary sequence. In theory each module performs one discrete elongation step. And the order of different steps is given by the module sequence. More and more exceptions of the text book biochemistry are discovered ^[108]; especially from myxobacteria ^[45]; therefore this cannot be considered as individual case. For example there are reports of occasions of stuttering of modules (a module is used multiple times for chain extension) or module skipping. ^[277] These things might have been relevant for the evolution of PKSs, ^[278] as such stuttering increases the diversity of natural compounds. The biochemical knowledge of stuttering, if fully understood, can be used for engineering of metabolic assembly lines to create novel unnatural natural products. ^[277] Examples for module stuttering are stigmatellin and borellidin. For stigmatellin ten condensation cycles are necessary and these cycles are performed by nine PKS modules. ^[274] For borellidin eight condensation steps are necessary for the final product, but the biosynthetic assembly line contains only six extension modules. ^[279]

Module skipping is to bypass the utilization of a single domain or a complete module of a biosynthetic assembly. ^[239] When a complete module is skipped, the biosynthetic intermediate is directly transferred from one module to the after next one. That means that a complete module with all its enzyme activities is ignored. ^[240] Alternatively the ACP or PCP of the skipped module is still involved, but no extender unit is added to the intermediate because the AT or rather A domain is inactive. ^[207] Such module skipping can be irregular or programmed. For the aberrant skipping the epothilone derivative K is an example. ^[280] A well investigated example of such module skipping was reported by Leadlay et al. ^[278] They report of a hybrid tetraketide synthase of erythromycin containing a foreign module of rapamycin which resulted in different products: triketides as major product and tetraketides as minor products.

Programmed skipping is known from pikromycin synthase from *Streptomyces venezuelae*. Depending on the cultivation conditions either the full-length product or a truncated version is build. ^[281] Myxobacterial examples with intentional module skipping are myxochromide S ^[207], disorazol ^[282] and chivosazol ^[283].

For a NRPS-PKS hybrid system module skipping was shown by site-directed mutagenesis experiments in vivo in the biosynthesis of leinamycin. ^[284]

In this work such skipping is reported for NRPS-PKS hybrid product myxoprincomide. Here module 13 (DK897) or rather module 12 (DK1622) is skipped.

Another skipping example in this thesis is the PKS product spirangien. Here point mutations lead to module skipping. A spiroketal compound described in literature with module skipping is nemadectin (Figure 80).^[108] Here the penultimate methylmalonyl-CoA extender unit is not always incorporated by the PKS. This skipping leads to nine additional products.

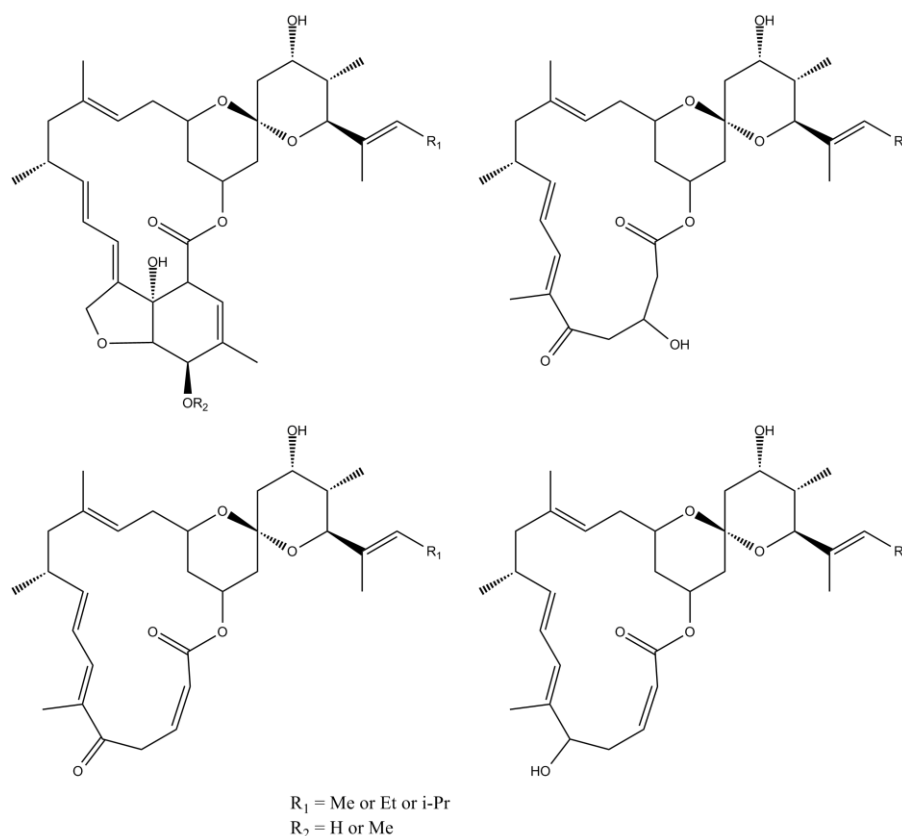


Figure 80: Nemadectin (top left) and products that result from module skipping

5. Pre-chondrochloren, last intermediate in chondrochloren biosynthesis

The strain *Chondromyces crocatus* Cm c5 is a very potent secondary metabolite producer; it generates at least six distinct classes. It is known to produce the highly antifungal ajudazols which inhibit the mitochondrial electron transport. ^[285] Another interesting metabolite is the macrolide thuggacin, it acts against tuberculosis. The biosynthesis of both metabolites is well discussed in the thesis of K. Buntin. ^[286]

The biosynthesis of chondrochloren ^[77] which is also generated by *Chondromyces crocatus* Cm c5 was described by Rachid et al. ^[133]. Particularly the post-assembly enzymatic reactions, the chlorination and the oxidative decarboxylation (by CndG), are very interesting. CndG operates post-PKS / NRPS and it attains the biological activity of the chondrochloren. Decarboxylations are widespread in nature; they specifically modify the substrates. ^[287;288] Three different enzyme classes are known which are distinguished by their used cofactors. One decarboxylase group uses the cofactors: biotin, flavin, and NAD⁺ / NADP⁺. ^[289-291] The second class employs inorganic cofactors. ^[292] And enzymes without a known cofactor build the third group. ^[287;293]

The oxidative decarboxylase (CndG) of the chondrochloren gene cluster was heterologously expressed in *Escherichia coli* by S. Rachid. To show that CndG is FAD-dependent; the pre-chondrochloren was needed as substrate to convert it by the enzyme into chondrochloren. A knockout of CndG leads to a carboxylated and saturated derivative of chondrochloren A and B, named as pre-chondrochloren. It is the last intermediate within the biosynthesis of chondrochloren. To proof the hypothesis of the pre-chondrochloren the mutant was cultivated in larger scale to isolate the novel created metabolite.

Biological activity tests of pre-chondrochloren revealed that the decarboxylation is essential for the bioactivity. This shows how important each single step in biosynthesis can be. In the case of the shortened spirangien variants structural variations did not exhibit strong difference in comparison to the normal spirangiens.

But it demonstrates that evolution which leads to structural differences makes sense for the bacteria. As huge structural diversity by small variations in the structure might lead to strong differences in bioactivity. That is also the reason why natural products are used as leads structures for drugs. ^[4]

6. Spirangien and spiroketal formations

Spirangien is only one of several interesting secondary metabolites of *Sorangium cellulosum* So ce90. Epothilone^[173;174] is the most famous representative of that strain. Small modifications to its structure led to the development of a new anticancer drug, Ixabepilone[®].^[5]

Spirangien is a product of pure PKS machinery. One common characteristic of type I PKS is the colinearity of the gene cluster with its structure. But in the described spirangien chapter a spontaneous mutant occurred which lead to an inactivation of a module (*spiJ*, module 14) and enforces a skipping of a module. This does not stop the complete spirangien production but it results in four novel shortened derivatives. The mutation in the spirangien cluster^[235] represents some kind of evolutionary processes in the laboratory. Hence it shows insights into the evolutionary development of biosynthetic gene clusters and how effective slight mutations can be in the progress of novel natural compounds.

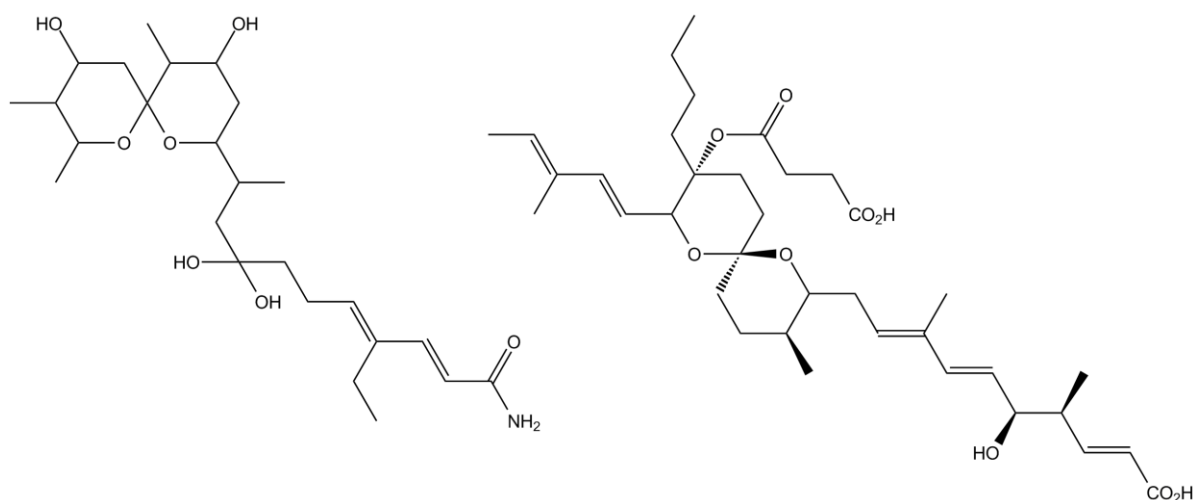


Figure 81: Structure of spirotamide A (left) and reveromycin A (right)

Natural compounds containing a spiroacetal group are ubiquitous in nature. Due to their specific structure spiroacetals exhibit various pharmaceutical activities. Recently a novel spiroacetal, the spirotamide (Figure 81) was isolated and structurally characterized.^[294] It showed cytotoxic, antibacterial and antifungal activity. Other spiroacetal examples with biological activity are tautomycin (anticancer), monensin (antibiotic), avermectin (antiparasitic) and here mentioned spirangien (antifungal). The diverse biological activities

and the stereospecific structures of spiroacetals made them attractive goals for organic synthesis.^[295] Compared to chemical synthesis, the biosynthetic spiroacetal formation has not been completely characterized. Although the knowledge, about the biosynthetic machinery and single biosynthetic steps, is important for the development of novel derivatives that may have higher biological activity. But often responsible enzymes could not be identified or unstable intermediates could not be analyzed.^[241]

For some biosynthetic spiroketal formations there are assumptions but the involved enzymes for the cyclization step could not be identified. Tautomycin^[296] and avermectin B1a^[297] are presumably formed through none enzymatic dehydrative cyclization mechanism^[235] (refer to Figure 82 I). The spiroacetal formation of monensin A is well studied. It is formed by an S_N2 reaction of an epoxide and ketone intermediate^[242;298] (refer to Figure 82 II). While for spirangien A a P450 is likely to be involved in the dehydrative cyclization process. It is assumed that a P450 builds a keto-diol which then cyclizes spontaneously to the spiroacetal. The spiroacetal formation is thermodynamically favored and not controlled by an enzyme.^[299] Due to the new shortened spirangien derivatives, which were isolated and its structure completely characterized by NMR spectroscopy, the biosynthetic spiroketal formation of spirangien could be described in more detail. A proposed pathway by Frank et al.^[175] could be affirmed. The existence of the new spirangien derivatives proves that a previous oxidation step for the spiroketal formation is not necessary.

Recently Osada et al. identified the responsible enzymes which are involved in the spiroacetal formation of reveromycin (refer to Figure 81 and Figure 82 III). Here two enzymes are performing spiroacetal formation. RevG (a dihydroxy ketone synthase) conducts the dehydrogenation and RevJ (a spiroacetal synthase) accomplishes the cyclization. This spiroacetal formation was reconstructed by *in vitro* experiments with the precursor and RevG and RevJ.^[241]

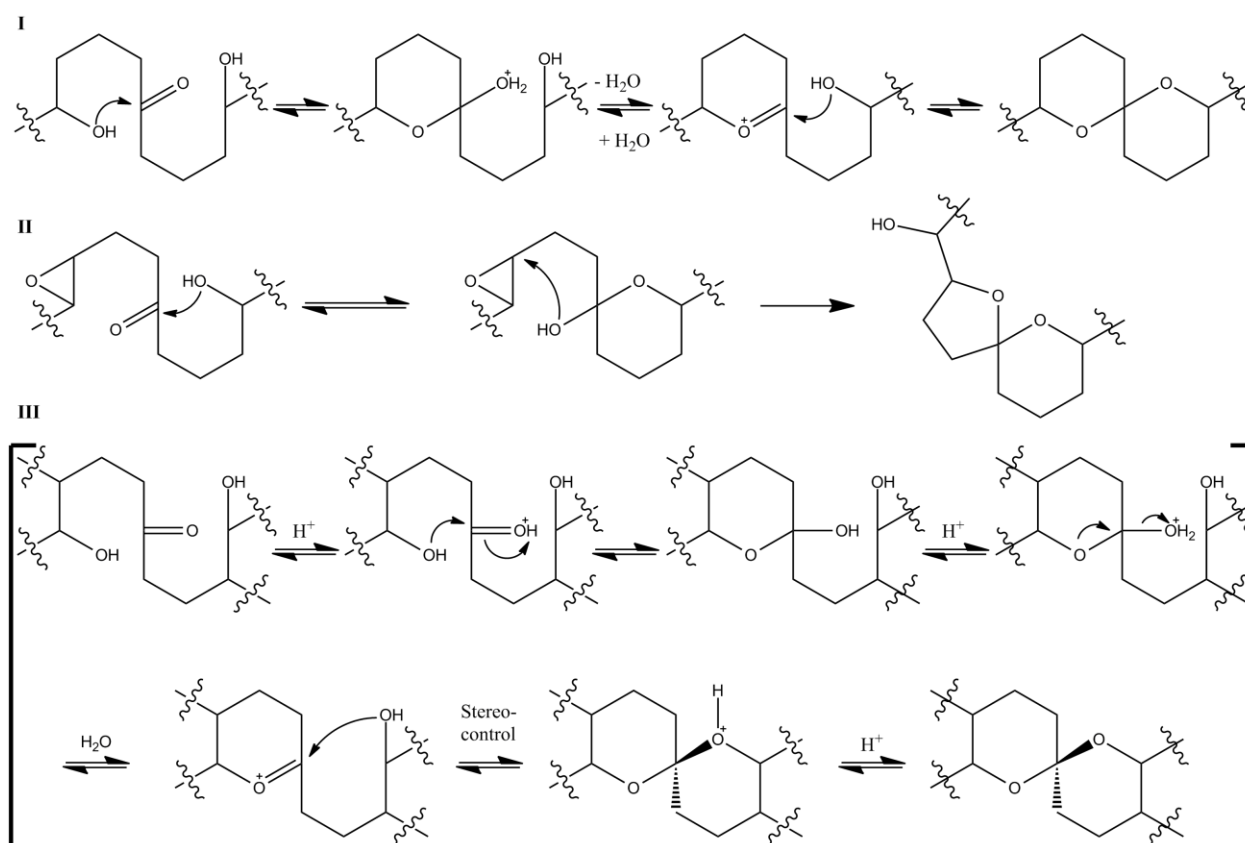


Figure 82: Schema of different spiroacetal formations: I via a dehydrative cyclization like in the case of tautomycin or avermectin B1a; II via an epoxide as in monensin A; III via enzymatic control shown for reveromycin

References

- [1.] A. L. Demain, *Appl Microbiol Biotechnol* **1999**, 52 455-463.
- [2.] P. G. Williams, G. O. Buchanan, R. H. Feling, C. A. Kauffman, P. R. Jensen, W. Fenical, *J.Org.Chem.* **2005**, 70 6196-6203.
- [3.] L. Brandi, A. Fabbretti, A. La Teana, M. Abbondi, D. Losi, S. Donadio, C. O. Gualerzi, *P.Natl.Acad.Sci.USA* **2006**, 103 39-44.
- [4.] A. L. Demain, *J.Ind.Microbiol.Biotechnol.* **2006**, 33 486-495.
- [5.] S. Danishefsky, *Nat.Prod.Rep.* **2010**, 27 1114-1116.
- [6.] D. D. Baker, M. Chu, U. Oza, V. Rajgarhia, *Nat.Prod.Rep.* **2007**, 24 1225-1244.
- [7.] J. Young, R. E. Taylor, *Nat.Prod.Rep.* **2008**, 25 651-655.
- [8.] D. J. Newman, G. M. Cragg, *J.Nat.Prod.* **2007**, 70 461-477.
- [9.] B. K. Hubbard, C. T. Walsh, *Angew.Chem.Int.Ed.* **2003**, 42 730-765.
- [10.] R. H. Baltz, V. Miao, S. K. Wrigley, *Nat.Prod.Rep.* **2005**, 22 717-741.
- [11.] J. A. Washington, W. R. Wilson, *Mayo Clin.Proc.* **1985**, 60 189-203.
- [12.] M. A. Fischbach, C. T. Walsh, *Science* **2009**, 325 1089-1093.
- [13.] C. Nathan, *Nature* **2004**, 431 899-902.
- [14.] M. Saleem, M. Nazir, M. S. Ali, H. Hussain, Y. S. Lee, N. Riaz, A. Jabbar, *Nat.Prod.Rep.* **2010**, 27 238-254.
- [15.] J. W. Blunt, B. R. Copp, M. H. Munro, P. T. Northcote, M. R. Prinsep, *Nat.Prod.Rep.* **2010**, 27 165-237.
- [16.] Di Santo R., *Nat.Prod.Rep.* **2010**, 27 1084-1098.
- [17.] T. Henkel, R. M. Brunne, H. Müller, F. Reichel, *Angew.Chem.Int.Ed.* **1999**, 38 643-647.
- [18.] C. Cordier, D. Morton, S. Murrison, A. Nelson, C. O'Leary-Steele, *Nat.Prod.Rep.* **2008**, 25 719-737.
- [19.] R. H. Baltz, *Journal of Industrial Microbiology & Biotechnology* **2006**, 33 507-513.
- [20.] E. Busti, P. Monciardini, L. Cavaletti, R. Bamonte, A. Lazzarini, M. Sosio, S. Donadio, *Microbiology* **2006**, 152 675-683.
- [21.] K. Harada, *Chem Pharm Bull (Tokyo)* **2004**, 52 889-899.
- [22.] V. Behal, *Folia Microbiol.* **2003**, 48 563-571.
- [23.] L. T. Tan, *Phytochemistry* **2007**, 68 954-979.
- [24.] K. S. Lam, *Curr.Opin.Microbiol.* **2006**, 9 245-251.

-
- [25.] K. Scherlach, L. P. Partida-Martinez, H. M. Dahse, C. Hertweck, *J.Am.Chem.Soc.* **2006**, *128* 11529-11536.
 - [26.] J. Li, K. Hu, J. M. Webster, *Chemistry of Hereocyclic Compounds* **1998**, *34* 1331-1339.
 - [27.] H. Reichenbach, G. Höfle, in *Drug Discovery from Nature* Eds.: S. Grabley, R. Thiericke), Springer, Berlin **1999**, p. pp. 149-179.
 - [28.] I. de Bruijn, M. J. D. de Kock, M. Yang, P. de Waard, T. A. van Beek, J. M. Raaijmakers, *Mol.Microbiol.* **2007**, *63* 417-428.
 - [29.] T. A. M. Gulder, B. S. Moore, *Curr.Opin.Microbiol.* **2009**, *12* 252-260.
 - [30.] R. O. Garcia, D. Krug, R. Müller, *Methods Enzymol.* **2009**, *458* 59-91.
 - [31.] K. Ravenschlag, K. Sahm, J. Pernthaler, R. Amann, *Appl.Environ.Microbiol.* **1999**, *65* 3982-3989.
 - [32.] H. Reichenbach, *Environ.Microbiol.* **1999**, *1* 15-21.
 - [33.] W. Dawid, *FEMS Microbiol.Rev.* **2000**, *24* 403-427.
 - [34.] T. Iizuka, Y. Jojima, R. Fudou, S. Yamanaka, *FEMS Microbiol.Lett.* **1998**, *169* 317-322.
 - [35.] L. Shimkets, M. Dworkin, H. Reichenbach, in *The Prokaryotes* ,Vol. 7, (Ed.: M. Dworkin), Springer, Berlin **2006**, p. pp. 31-115.
 - [36.] *Myxobacteria: Multicellularity and differentiation*, (Ed.: D. Whitworth) ASM Press, Chicago **2007**.
 - [37.] D. Kaiser, *Nat.Rev.Microbiol.* **2003**, *1* 45-54.
 - [38.] H. Reichenbach, *J.Ind.Microbiol.Biotechnol.* **2001**, *27* 149-156.
 - [39.] H. B. Kaplan, *Curr.Opin.Microbiol.* **2003**, *6* 572-577.
 - [40.] R. O. Garcia, H. Reichenbach, M. W. Ring, R. Müller, *International Journal of Systematic and Evolutionary Microbiology* **2009**, *59* 1524-1530.
 - [41.] H. Reichenbach, in *Bergey's manual of systematic bacteriology* Eds.: D. J. Brenner, N. R. Krieg, J. T. Staley), Springer, **2005**, p. pp. 1059-1144.
 - [42.] S. C. Wenzel, R. Müller, *Mol.Biosyst.* **2009**, *5* 567-574.
 - [43.] S. C. Wenzel, R. Müller, *Curr Opin Drug Discov Devel* **2009**, *12* 220-230.
 - [44.] K. Gerth, S. Pradella, O. Perlova, S. Beyer, R. Müller, *J.Biotechnol.* **2003**, *106* 233-253.
 - [45.] K. J. Weissman, R. Müller, *Bioorg.Med.Chem.* **2009**, *17* 2121-2136.
 - [46.] H. B. Bode, R. Müller, *J.Ind.Microbiol.Biotechnol.* **2006**, *33* 577-588.
 - [47.] S. C. Wenzel, R. Müller, *Nat.Prod.Rep.* **2007**, *24* 1211-1224.

-
- [48.] K. J. Weissman, R. Müller, *Nat.Prod.Rep.* **2010**, 27 1276-1295.
- [49.] M. Nett, G. M. König, *Nat.Prod.Rep.* **2007**, 24 1245-1261.
- [50.] D. Krug, G. Zurek, B. Schneider, R. Garcia, R. Müller, *Anal.Chim.Acta* **2008**, 624 97-106.
- [51.] D. Krug, G. Zurek, O. Revermann, M. Vos, G. J. Velicer, R. Müller, *Appl.Environ.Microbiol.* **2008**, 74 3058-3068.
- [52.] M. W. Khalil, F. Sasse, H. Lünsdorf, Y. A. Elnakady, H. Reichenbach, *ChemBioChem* **2006**, 7 678-683.
- [53.] F. Sasse, H. Steinmetz, J. Heil, G. Höfle, H. Reichenbach, *J.Antibiot.* **2000**, 53 879-885.
- [54.] J. Mulzer, *The Epothilones - An Outstanding Family of Anti-Tumour Agents: From Soil to the Clinic*, (Ed.: J. Mulzer) Springer, New York **2009**.
- [55.] S. Goodin, M. P. Kane, E. H. Rubin, *J Clin Oncol* **2004**, 22 2015-2025.
- [56.] D. M. Bollag, P. A. McQueney, J. Zhu, O. Hensens, L. Koupal, J. Liesch, M. Goetz, E. Lazarides, M. Woods, *Cancer Res.* **1995**, 55 2325-2333.
- [57.] H. Reichenbach, G. Höfle, *Drugs in R&D* **2008**, 9 1-10.
- [58.] Y. A. Elnakady, F. Sasse, H. Lünsdorf, H. Reichenbach, *Biochem.Pharmacol.* **2004**, 67 927-935.
- [59.] F. Sasse, B. Kunze, T. M. Gronewold, H. Reichenbach, *J.Natl.Cancer Inst.* **1998**, 90 1559-1563.
- [60.] G. Hagelueken, S. C. Albrecht, H. Steinmetz, R. Jansen, D. W. Heinz, M. Kalesse, W. D. Schubert, *Angew Chem Int Ed Engl* **2009**, 48 595-598.
- [61.] T. M. Gronewold, F. Sasse, H. Lünsdorf, H. Reichenbach, *Cell Tissue Res* **1999**, 295 121-129.
- [62.] I. Nückeleit, S. Zender, F. Sasse, R. Geffers, G. Brandes, I. Sörensen, H. Steinmetz, S. Kubicka, T. Carlomagno, D. Menche, I. Gütgemann, J. Buer, A. Gossler, M. P. Manns, M. Kalesse, R. Frank, N. P. Malek, *Cancer Res.* **2008**, 14 23-35.
- [63.] H. Reichenbach, G. Höfle, in *Scientific Annual Report*, Gesellschaft für Biotechnologische Forschung mbH, Braunschweig **1994**, p. pp. 5-22.
- [64.] L. Pridzun, F. Sasse, H. Reichenbach, in *Antifungal agents* Eds.: G. K. Dixon, C. L.G., D. W. Hollomon), BIOS Scientific Publishers Ltd, Oxford, UK **1995**, p. pp. 99-109.
- [65.] H. F. Vahlensieck, L. Pridzun, H. Reichenbach, A. Hinnen, *Curr.Genet.* **1994**, 25 95-100.

-
- [66.] A. Beckers, S. Organe, L. Tinunermans, K. Scheys, A. Peeters, K. Brusselmans, G. Verhoeven, J. V. Swinnen, *Cancer Res.* **2007**, *67* 8180-8187.
- [67.] Y. Shen, S. L. Volrath, S. C. Weatherly, T. D. Elich, L. Tong, *Mol.Cell* **2004**, *16* 881-891.
- [68.] G. Thierbach, B. Kunze, H. Reichenbach, G. Höfle, *Biochim.Biophys.Acta* **1984**, *765* 227-235.
- [69.] H. Irschik, H. Reichenbach, G. Höfle, R. Jansen, *J.Antibiot.* **2007**, *60* 733-738.
- [70.] H. Steinmetz, H. Irschik, B. Kunze, H. Reichenbach, G. Höfle, R. Jansen, *Chemistry* **2007**, *13* 5822-5832.
- [71.] K. Gerth, H. Irschik, H. Reichenbach, W. Trowitzsch, *J.Antibiot.* **1980**, *33* 1474-1479.
- [72.] F. Sasse, B. Böhlendorf, M. Herrmann, B. Kunze, E. Forche, H. Steinmetz, G. Höfle, H. Reichenbach, M. Hermann, *J.Antibiot.* **1999**, *52* 721-729.
- [73.] H. Irschik, R. Jansen, G. Höfle, K. Gerth, H. Reichenbach, *J.Antibiot.* **1985**, *38* 145-152.
- [74.] H. Irschik, K. Gerth, G. Höfle, W. Kohl, H. Reichenbach, *J.Antibiot.* **1983**, *36* 1651-1658.
- [75.] H. Irschik, R. Jansen, K. Gerth, G. Höfle, H. Reichenbach, *J.Antibiot.* **1987**, *40* 7-13.
- [76.] H. Irschik, H. Reichenbach, *J.Antibiot.* **1985**, *38* 1237-1245.
- [77.] R. Jansen, B. Kunze, H. Reichenbach, G. Höfle, *Eur.J.Org.Chem.* **2003**, 2684-2689.
- [78.] W. Trowitsch, L. Witte, H. Reichenbach, *FEMS Microbiol.Lett.* **1981**, *12* 257-260.
- [79.] B. S. Goldman, W. C. Nierman, D. Kaiser, S. C. Slater, A. S. Durkin, J. Eisen, C. M. Ronning, W. B. Barbazuk, M. Blanchard, C. Field, C. Halling, G. Hinkle, O. Iartchuk, H. S. Kim, C. Mackenzie, R. Madupu, N. Miller, A. Shvartsbeyn, S. A. Sullivan, M. Vaudin, R. Wiegand, H. B. Kaplan, *P.Natl.Acad.Sci.USA* **2006**, *103* 15200-15205.
- [80.] S. D. Bentley, K. F. Chater, A. M. Cerdono-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, D. A. Hopwood, *Nature* **2002**, *417* 141-147.
- [81.] H. Ikeda, J. Ishikawa, A. Hanamoto, M. Shinose, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori, S. Omura, *Nat.Biotechnol.* **2003**, *21* 526-531.

-
- [82.] R. Jansen, G. Reifenstahl, K. Gerth, H. Reichenbach, G. Höfle, *Liebigs Ann.Chem.* **1983**, 7 1081-1095.
- [83.] B. Kunze, N. Bedorf, W. Kohl, G. Höfle, H. Reichenbach, *J.Antibiot.* **1989**, 42 14-17.
- [84.] W. Trowitzsch, V. Wray, K. Gerth, G. Höfle, *J.Chem.Soc., Chem.Comm.* **1982**, 1340-1341.
- [85.] S. C. Wenzel, B. Kunze, G. Höfle, B. Silakowski, M. Scharfe, H. Blöcker, R. Müller, *ChemBioChem* **2005**, 6 375-385.
- [86.] P. Meiser, H. B. Bode, R. Müller, *P.Natl.Acad.Sci.USA* **2006**, 103 19128-19133.
- [87.] H. B. Bode, B. Bethe, R. Höfs, A. Zeeck, *ChemBioChem* **2002**, 3 619-627.
- [88.] R. Müller, K. Gerth, *J.Biotechnol.* **2006**, 121 192-200.
- [89.] C. Schley, M. O. Altmeyer, R. Swart, R. Müller, C. G. Huber, *J Proteome Res J Proteome Res* **2006**, 5 2760-2768.
- [90.] H. B. Bode, M. W. Ring, G. Schwär, M. O. Altmeyer, C. Kegler, I. R. Jose, M. Singer, R. Müller, *ChemBioChem* **2009**, 10 128-140.
- [91.] H. B. Bode, R. Müller, *Angew.Chem.Int.Ed.* **2005**, 44 6828-6846.
- [92.] S. C. Wenzel, R. Müller, *Curr.Opin.Biotechnol.* **2005**, 16 594-606.
- [93.] H. G. Menzella, C. D. Reeves, *Curr.Opin.Microbiol.* **2007**, 10 238-245.
- [94.] H. G. Floss, *J.Biotechnol.* **2006**, 124 242-257.
- [95.] K. J. Weissman, P. F. Leadlay, *Nat.Rev.Microbiol.* **2005**, 3 925-936.
- [96.] B. Shen, *Sci STKE* **2004**, 2004 e14.
- [97.] S. C. Wenzel, R. Müller, *Nat.Prod.Rep.* **2009**, 26 1385-1407.
- [98.] B. Silakowski, G. Nordsiek, B. Kunze, H. Blöcker, R. Müller, *Chemistry & Biology* **2001**, 8 59-69.
- [99.] H. B. Bode, B. Zeggel, B. Silakowski, S. C. Wenzel, H. Reichenbach, R. Müller, *Mol.Microbiol.* **2003**, 47 471-481.
- [100.] B. Kunze, H. Reichenbach, R. Müller, G. Höfle, *J.Antibiot.* **2005**, 58 244-251.
- [101.] K. J. Weissman, *Ernst Schering Res Found Workshop* **2005**, 43-78.
- [102.] D. Schwarzer, R. Finking, M. A. Marahiel, *Nat.Prod.Rep.* **2003**, 20 275-287.
- [103.] D. Schwarzer, M. A. Marahiel, *Naturwissenschaften* **2001**, 88 93-101.
- [104.] T. Stein, J. Vater, V. Kruft, A. Otto, B. Wittmann-Liebold, P. Franke, M. Panico, R. McDowell, H. R. Morris, *J.Biol.Chem.* **1996**, 271 15428-15435.
- [105.] R. Finking, M. A. Marahiel, *Annual Review of Microbiology* **2004**, 58 453-488.
- [106.] M. A. Fischbach, C. T. Walsh, *Chem.Rev.* **2006**, 106 3468-3496.
- [107.] J. R. Lai, A. Koglin, C. T. Walsh, *Biochemistry-US* **2006**, 45 14869-14879.

-
- [108.] S. J. Moss, C. J. Martin, B. Wilkinson, *Nat.Prod.Rep.* **2004**, 21 575-593.
- [109.] G. Yadav, R. S. Gokhale, D. Mohanty, *Journal of Molecular Biology* **2003**, 328 335-363.
- [110.] J. Staunton, K. J. Weissman, *Nat.Prod.Rep.* **2001**, 18 380-416.
- [111.] G. Weber, E. Leitner, *Curr.Genet.* **1994**, 26 461-467.
- [112.] K. Hoffmann, E. Schneider-Scherzer, H. Kleinkauf, R. Zocher, *J.Biol.Chem.* **1994**, 269 12710-12714.
- [113.] C. S. Neumann, D. G. Fujimori, C. T. Walsh, *Chemistry & Biology* **2008**, 15 99-109.
- [114.] E. Conti, T. Stachelhaus, M. A. Marahiel, P. Brick, *EMBO J.* **1997**, 16 4174-4183.
- [115.] T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem.Biol.* **1999**, 6 493-505.
- [116.] G. L. Challis, J. Ravel, C. A. Townsend, *Chem.Biol.* **2000**, 7 211-224.
- [117.] J. Ligon, S. Hill, J. Beck, R. Zirkle, I. Molnar, J. Zawodny, S. Money, T. Schupp, *Gene* **2002**, 285 257-267.
- [118.] H. Ikeda, T. Nonomiya, S. Omura, *J.Ind.Microbiol.Biotechnol.* **2001**, 27 170-176.
- [119.] B. Silakowski, H. U. Schairer, H. Ehret, B. Kunze, S. Weinig, G. Nordsiek, P. Brandt, H. Blöcker, G. Höfle, S. Beyer, R. Müller, *J.Biol.Chem.* **1999**, 274 37391-37399.
- [120.] C. D. Reeves, L. M. Chung, Y. Liu, Q. Xue, J. R. Carney, W. P. Revill, L. Katz, *J.Biol.Chem.* **2002**, 277 9155-9159.
- [121.] Y. A. Chan, M. T. Boyne, A. M. Podevels, A. K. Klimowicz, J. Handelsman, N. L. Kelleher, M. G. Thomas, *P.Natl.Acad.Sci.USA* **2006**, 103 14349-14354.
- [122.] B. J. Carroll, S. J. Moss, L. Q. Bai, Y. Kato, S. Toelzer, T. W. Yu, H. G. Floss, *J.Am.Chem.Soc.* **2002**, 124 4176-4177.
- [123.] S. C. Wenzel, R. M. Williamson, C. Grünanger, J. Xu, K. Gerth, R. A. Martinez, S. J. Moss, B. J. Carroll, S. Grond, C. J. Unkefer, R. Müller, H. G. Floss, *J.Am.Chem.Soc.* **2006**, 128 14325-14336.
- [124.] F. Del Vecchio, H. Petkovic, S. G. Kendrew, L. Low, B. Wilkinson, R. Lill, J. Cortes, B. A. Rudd, J. Staunton, P. F. Leadlay, *J.Ind.Microbiol.Biotechnol.* **2003**, 30 489-494.
- [125.] S. F. Haydock, J. F. Aparicio, I. Molnar, T. Schwecke, A. König, A. F. A. Marsden, I. S. Galloway, J. Staunton, P. F. Leadley, *FEBS Lett.* **1995**, 374 246-248.
- [126.] B. Shen, *Top.Curr.Chem.* **2000**, 209 1-51.
- [127.] C. M. Rath, J. B. Scaglione, J. D. Kittendorf, D. H. Sherman, in *Comprehensive Natural Products II. Chemistry and Biology*, Vol. 1:Natural Products Structural Diversity I - Secondary Metabolites: Organization and Biosynthesis Eds.: L. Mander, H. W. Liu), Elsevier, Oxford **2010**, p. pp. 453-492.

-
- [128.] L. Du, B. Shen, *Curr Opin Drug Discov Devel* **2001**, *4* 215-228.
- [129.] L. H. Du, C. Sanchez, B. Shen, *Metab.Eng.* **2001**, *3* 78-95.
- [130.] C. L. Bender, D. A. Palmer, A. Peñaloza-Vázquez, V. Rangaswamy, M. Ullrich, *Subcell.Biochem.* **1998**, *29* 321-341.
- [131.] V. Rangaswamy, R. Mitchell, M. Ullrich, C. Bender, *J.Bacteriol.* **1998**, *180* 3330-3338.
- [132.] V. Rangaswamy, S. Jiralerspong, R. Parry, C. L. Bender, *P.Natl.Acad.Sci.USA* **1998**, *95* 15469-15474.
- [133.] S. Rachid, M. Scharfe, H. Blöcker, K. J. Weissman, R. Müller, *Chem.Biol.* **2009**, *16* 70-81.
- [134.] S. Rachid, O. Revermann, C. Dauth, R. Mueller, *J.Biol.Chem.* **2010**, *285* 12482-12489.
- [135.] S. A. Samel, M. A. Marahiel, L. O. Essen, *Molecular Biosystems* **2008**, *4* 387-393.
- [136.] C. T. Walsh, H. W. Chen, T. A. Keating, B. K. Hubbard, H. C. Losey, L. S. Luo, C. G. Marshall, D. A. Miller, H. M. Patel, *Curr.Opin.Chem.Biol.* **2001**, *5* 525-534.
- [137.] U. Rix, C. Fischer, L. L. Remsing, J. Rohr, *Nat.Prod.Rep.* **2002**, *19* 542-580.
- [138.] R. Jansen, H. Irschik, H. Reichenbach, G. Höfle, *Liebigs Ann Chem* **1997**, 1725-1732.
- [139.] H. Zhou, X. Xie, Y. Tang, *Curr.Opin.Biotechnol.* **2008**, *19* 590-596.
- [140.] Y. J. Yoon, B. J. Beck, B. S. Kim, H. Y. Kang, K. A. Reynolds, D. H. Sherman, *Chem.Biol.* **2002**, *9* 203-214.
- [141.] C. Zhang, B. R. Griffith, Q. Fu, C. Albermann, X. Fu, I. K. Lee, L. Li, J. S. Thorson, *Science* **2006**, *313* 1291-1294.
- [142.] C. S. Zhang, C. Albermann, X. Fu, J. S. Thorson, *J.Am.Chem.Soc.* **2006**, *128* 16420-16421.
- [143.] M. S. Donia, B. J. Hathaway, S. Sudek, M. G. Haygood, M. J. Rosovitz, J. Ravel, E. W. Schmidt, *Nat Chem Biol Nat Chem Biol* **2006**, *2* 729-735.
- [144.] E. W. Schmidt, J. T. Nelson, D. A. Rasko, S. Sudek, J. A. Eisen, M. G. Haygood, J. Ravel, *P.Natl.Acad.Sci.USA* **2005**, *102* 7315-7320.
- [145.] T. J. Oman, W. A. van der Donk, *Nat.Chem.Biol.* **2010**, *6* 9-18.
- [146.] E. W. Schmidt, *BMC Biol.* **2010**, *8* 83.
- [147.] D. H. Haft, M. K. Basu, D. A. Mitchell, *BMC Biol.* **2010**, *8* 70.
- [148.] N. Schnell, K. D. Entian, U. Schneider, F. Gotz, H. Zahner, R. Kellner, G. Jung, *Nature* **1988**, *333* 276-278.

-
- [149.] Y. Goto, B. Li, J. Claesen, Y. X. Shi, M. J. Bibb, W. A. van der Donk, *Plos Biology* **2010**, 8 .
- [150.] W. M. Müller, T. Schmiederer, P. Ensle, R. D. Sussmuth, *Angew.Chem.Int.Ed Engl.* **2010**, 49 2436-2440.
- [151.] S. Duquesne, V. Petit, J. Peduzzi, S. Rebuffat, *J.Mol.Microbiol.Biotechnol.* **2007**, 13 200-209.
- [152.] W. L. Kelly, L. Pan, C. Li, *J.Am.Chem Soc.* **2009**.
- [153.] L. C. Wieland Brown, M. G. Acker, J. Clardy, C. T. Walsh, M. A. Fischbach, *Proc.Natl.Acad.Sci.U.S A* **2009**, 106 2549-2553.
- [154.] Y. Yu, L. Duan, Q. Zhang, R. Liao, Y. Ding, H. Pan, E. Wendt-Pienkowski, G. Tang, B. Shen, W. Liu, *ACS Chem Biol* **2009**.
- [155.] H. E. Hallen, H. Luo, J. S. Scott-Craig, J. D. Walton, *Proc.Natl.Acad.Sci.U.S A* **2007**, 104 19097-19101.
- [156.] M. S. Donia, J. Ravel, E. W. Schmidt, *Nat Chem Biol Nat Chem Biol* **2008**, 4 341-343.
- [157.] N. Ziemert, K. Ishida, P. Quillardet, C. Bouchier, C. Hertweck, N. T. de Marsac, E. Dittmann, *Appl.Environ.Microbiol.* **2008**, 74 1791-1797.
- [158.] C. Li, W. L. Kelly, *Nat.Prod.Rep.* **2010**, 27 153-164.
- [159.] S. Sudek, M. G. Haygood, D. T. Youssef, E. W. Schmidt, *Appl.Environ.Microbiol.* **2006**, 72 4382-4387.
- [160.] N. Ziemert, K. Ishida, A. Liaimer, C. Hertweck, E. Dittmann, *Angew.Chem.Int.Ed.* **2008**, 47 7756-7759.
- [161.] A. L. McClerren, L. E. Cooper, C. Quan, P. M. Thomas, N. L. Kelleher, W. A. van der Donk, *P.Natl.Acad.Sci.USA* **2006**, 103 17243-17248.
- [162.] S. W. Lee, D. A. Mitchell, A. L. Markley, M. E. Hensler, D. Gonzalez, A. Wohlrab, P. C. Dorrestein, V. Nizet, J. E. Dixon, *P.Natl.Acad.Sci.USA* **2008**, 105 5879-5884.
- [163.] R. Liao, L. Duan, C. Lei, H. Pan, Y. Ding, Q. Zhang, D. Chen, B. Shen, Y. Yu, W. Liu, *Chem Biol* **2009**, 16 141-147.
- [164.] Krug, D., Universität Saarbrücken, **2009**.
- [165.] T. F. Molinski, *Nat.Prod.Rep.* **2010**, 27 321-329.
- [166.] M. Zerkly, G. L. Challis, *ChemBioChem* **2009**, DOI: 10.1002/cbic.200800389 [167.] B. Ostash, A. Saghatelian, S. Walker, *Chem.Biol.* **2007**, 14 257-267.
- [168.] Y. Hu, V. Phelan, I. Ntai, C. M. Farnet, E. Zazopoulos, B. O. Bachmann, *Chemistry & Biology* **2007**, 14 691-701.

-
- [169.] H. Irschik, W. Trowitzsch-Kienast, K. Gerth, G. Höfle, H. Reichenbach, *J.Antibiot.* **1988**, *41* 993-998.
- [170.] W. Trowitzsch-Kienast, Citalins: Bicyclic Isotritirosines from *Myxococcus xanthus* *Citalins: Bicyclic Isotritirosines from Myxococcus xanthus*, pp. 496-497 (93).
- [171.] G. Höfle, Biological active secondary metabolites from myxobacteria - isolation and structure elucidation *Biological active secondary metabolites from myxobacteria - isolation and structure elucidation*, pp. 65-68 .
- [172.] G. Helynck, C. Dubertret, D. Frechet, J. Leboul, *J.Antibiot.* **1998**, *51* 512-514.
- [173.] G. Höfle, N. Bedorf, H. Steinmetz, D. Schomburg, K. Gerth, H. Reichenbach, *Angew.Chem.Int.Ed.* **1996**, *35* 1567-1569.
- [174.] Höfle, G., Bedorf, N., Gerth, K., and Reichenbach, H. Spirangiene, Herstellungsverfahren und diese Verbindungen enthaltende Mittel. 4211056[Patentschrift DE 4211056 C1]. 1993. Germany. 15-7-1993.
Ref Type: Patent
- [175.] Frank, B., Naturwissenschaftlich-Technische Fakultät III der Universität des Saarlandes, **2007**.
- [176.] N. Socorro Cortina, D. Krug, A. Plaza, O. Revermann, R. Müller, *Angew.Chem.Int.Ed.* **2012**, *51* 811-816.
- [177.] M. Bois-Choussy, P. Cristau, J. P. Zhu, *Angewandte Chemie-International Edition* **2003**, *42* 4238-4241.
- [178.] S. R. Nadkarni, M. V. Patel, S. Chatterjee, E. K. Vijayakumar, K. R. Desikan, J. Blumbach, B. N. Ganguli, M. Limbert, *J.Antibiot.* **1994**, *47* 334-341.
- [179.] N. Naruse, M. Oka, M. Konishi, T. Oki, *J.Antibiot.* **1993**, *46* 1812-1818.
- [180.] K. Matsuzaki, H. Ikeda, T. Ogino, A. Matsumoto, H. B. Woodruff, H. Tanaka, S. Omura, *The Journal of antibiotics* **1994**, *47* 1173-1174.
- [181.] O. Puk, D. Bischoff, C. Kittel, S. Pelzer, S. Weist, E. Stegmann, R. D. Sussmuth, W. Wohlleben, *J.Bacteriol.* **2004**, *186* 6093-6100.
- [182.] S. Boisnard, J. P. Zhu, *Tetrahedron Lett.* **2002**, *43* 2577-2580.
- [183.] S. Boisnard, A. C. Carbonnelle, J. P. Zhu, *Organic Letters* **2001**, *3* 2061-2064.
- [184.] P. J. Krenitsky, D. L. Boger, *Tetrahedron Lett.* **2003**, *44* 4019-4022.
- [185.] P. J. Krenitsky, D. L. Boger, *Tetrahedron Lett.* **2002**, *43* 407-410.
- [186.] J. Kratzschmar, M. Krause, M. A. Marahiel, *J.Bacteriol.* **1989**, *171* 5422-5429.
- [187.] K. M. Hoyer, C. Mahlert, M. A. Marahiel, *Chem.Biol.* **2007**, *14* 13-22.
- [188.] C. T. Walsh, *Science* **2004**, *303* 1805-1810.

-
- [189.] E. Stegmann, S. Pelzer, D. Bischoff, O. Puk, S. Stockert, D. Butz, K. Zerbe, J. Robinson, R. D. Sussmuth, W. Wohlleben, *J.Biotechnol.* **2006**, *124* 640-653.
 - [190.] B. Li, D. Sher, L. Kelly, Y. Shi, K. Huang, P. J. Knerr, I. Joewono, D. Rusch, S. W. Chisholm, W. A. van der Donk, *Proc.Natl.Acad.Sci.U.S A* **2010**, *107* 10430-10435.
 - [191.] S. Duquesne, D. stoumieux-Garzon, J. Peduzzi, S. Rebuffat, *Nat.Prod.Rep.* **2007**, *24* 708-734.
 - [192.] D. Werck-Reichhart, R. Feyereisen, *Genome Biol.* **2000**, *1* REVIEWS3003.
 - [193.] H. L. Schubert, R. M. Blumenthal, X. Cheng, *Trends Biochem.Sci.* **2003**, *28* 329-335.
 - [194.] J. A. McIntosh, M. S. Donia, E. W. Schmidt, *Nat.Prod.Rep.* **2009**, *26* 537-559.
 - [195.] H. B. Bode, P. Meiser, T. Klefisch, N. Socorro D.J.Cortina, D. Krug, A. Göhring, G. Schwär, T. Mahmud, Y. A. Elnakady, R. Müller, *ChemBioChem* **2007**, *8* 2139-2144.
 - [196.] P. Meiser, H. B. Bode, R. Müller, *Proc Natl Acad Sci U S A* **2006**, *103* 19128-19133.
 - [197.] H. B. Bode, M. W. Ring, G. Schwär, R. M. Kroppenstedt, D. Kaiser, R. Müller, *J.Bacteriol.* **2006**, *188* 6524-6528.
 - [198.] H. B. Bode, M. W. Ring, D. Kaiser, A. C. David, R. M. Kroppenstedt, G. Schwär, *J.Bacteriol.* **2006**, *188* 5632-5634.
 - [199.] Klefisch, T., Saarland University, **2007**.
 - [200.] S. Rachid, K. Gerth, I. Kochems, R. Müller, *Mol.Microbiol.* **2007**, *63* 1783-1796.
 - [201.] S. Pelzer, R. Sussmuth, D. Heckmann, J. Recktenwald, P. Huber, G. Jung, W. Wohlleben, *Antimicrob.Agents Chemother.* **1999**, *43* 1565-1573.
 - [202.] L. Betancor, M. J. Fernandez, K. J. Weissman, P. F. Leadlay, *ChemBioChem* **2008**, *9* 2962-2966.
 - [203.] personal communication from Dittmann et al. 2008.
Ref Type: Unpublished Work
 - [204.] C. T. Walsh, M. A. Fischbach, *J.Am.Chem.Soc.* **2010**.
 - [205.] G. M. Cragg, P. G. Grothaus, D. J. Newman, *Chem Rev.* **2009**, *109* 3012-3043.
 - [206.] C. M. Ronning, W. C. Nierman, in *Myxobacteria: Multicellularity and differentiation* Ed.: D. Whitworth), ASM Press, Chicago **2007**, p. pp. 285-298.
 - [207.] S. C. Wenzel, P. Meiser, T. Binz, T. Mahmud, R. Müller, *Angew.Chem.Int.Ed.* **2006**, *45* 2296-2301.
 - [208.] V. Simunovic, J. Zapp, S. Rachid, D. Krug, P. Meiser, R. Müller, *ChemBioChem* **2006**, *7* 1206-1220.
 - [209.] P. Meiser, K. J. Weissman, H. B. Bode, D. Krug, J. S. Dickschat, A. Sandmann, R. Müller, *Chem.Biol.* **2008**, *15* 771-781.

-
- [210.] D. Krug, G. Zurek, B. Schneider, C. Bässmann, R. Müller, *LC-GC Europe, The Applications Book* **2007**, March 2007 41-42.
- [211.] B. O. Bachmann, J. Ravel, *Methods in Enzymology* **2009**, 458 181-217.
- [212.] C. Rausch, T. Weber, O. Kohlbacher, W. Wohlleben, D. H. Huson, *Nucleic Acids Res.* **2005**, 33 5799-5808.
- [213.] T. Velkov, A. Lawen, in *Biotechnology Annual Review, Vol 9* **2003**, p. pp. 151-197.
- [214.] J. M. Seco, E. Quinoa, R. Riguera, *Chem.Rev.* **2004**, 104 17-117.
- [215.] T. R. Hoye, C. S. Jeffrey, F. Shao, *Nature Protocols* **2007**, 2 2451-2458.
- [216.] G. B. M. M. J. R. L. J. L. K. P. L. C. J. N. A. H. L. J. B. a. C. A. B. Alberto Plaza, *J.Org.Chem.* **2010**, 75 4344-4355.
- [217.] K. Fujii, Y. Ikai, H. Oka, M. Suzuki, K. Harada, *Anal.Chem.* **1997**, 69 5146-5151.
- [218.] K. Fujii, Y. Ikai, T. Mayumi, H. Oka, M. Suzuki, K. Harada, *Anal.Chem.* **1997**, 69 3346-3352.
- [219.] C. B'Hymer, M. Montes-Bayon, J. A. Caruso, *Journal of Separation Science* **2003**, 26 7-19.
- [220.] R. Bhushan, H. Bruckner, *Amino Acids* **2004**, 27 231-247.
- [221.] D. R. Goodlett, P. A. Abuaf, P. A. Savage, K. A. Kowalski, T. K. Mukherjee, J. W. Tolan, N. Corkum, G. Goldstein, J. B. Crowther, *J.Chromatogr.* **1995**, 707 233-244.
- [222.] I. Molnar, T. Schupp, M. Ono, R. Zirkle, M. Milnamow, B. Nowak-Thompson, N. Engel, C. Toupet, A. Stratmann, D. D. Cyr, J. Gorlach, J. M. Mayo, A. Hu, S. Goff, J. Schmid, J. M. Ligon, *Chem.Biol.* **2000**, 7 97-109.
- [223.] L. Tang, S. Shah, L. Chung, J. Carney, L. Katz, C. Khosla, B. Julien, *Science* **2000**, 287 640-642.
- [224.] L. Du, M. Chen, C. Sanchez, B. Shen, *FEMS Microbiol.Lett.* **2000**, 189 171-175.
- [225.] L. Du, L. Lou, *Nat.Prod.Rep.* **2010**, 27 255-278.
- [226.] E. J. Rubin, B. J. Akerley, V. N. Novik, D. J. Lampe, R. N. Husson, J. J. Mekalanos, *P.Natl.Acad.Sci.USA* **1999**, 96 1645-1650.
- [227.] E. A. Peterson, L. E. Overman, *P.Natl.Acad.Sci.USA* **2004**, 101 11943-11948.
- [228.] K. Gerth, N. Bedorf, H. Irschik, G. Höfle, H. Reichenbach, *J.Antibiot.* **1994**, 47 23-31.
- [229.] M. Köster, S. Lykke-Andersen, Y. A. Elnakady, K. Gerth, P. Washausen, G. Hofle, F. Sasse, J. Kjems, H. Hauser, *Exp.Cell Res.* **2003**, 286 321-331.
- [230.] J. Niggemann, N. Bedorf, U. Flörke, H. Steinmetz, K. Gerth, H. Reichenbach, G. Höfle, *Eur.J.Org.Chem.* **2005**, 2005 5013-5018.
- [231.] M. Lorenz, M. Kalesse, *Org.Lett.* **2008**, 10 4371-4374.

-
- [232.] I. Paterson, A. D. Findlay, E. A. Anderson, *Angewandte Chemie-International Edition* **2007**, 46 6699-6702.
- [233.] I. Paterson, A. D. Findlay, C. Noti, *Chem.Commun.(Camb.)* **2008**, 6408-6410.
- [234.] Virtual Textbook of Organic Chemistry. 22-4-2012.
Ref Type: Online Source
- [235.] B. Frank, J. Knauber, H. Steinmetz, M. Scharfe, H. Blöcker, S. Beyer, R. Müller, *Chem.Biol.* **2007**, 14 221-233.
- [236.] J. P. P. Muyrers, Y. Zhang, A. F. Stewart, *Trends Biochem.Sci.* **2001**, 26 325-331.
- [237.] Y. Zhang, J. P. P. Muyrers, G. Testa, A. F. Stewart, *Nat.Biotechnol.* **2000**, 18 1314-1317.
- [238.] Y. Zhang, F. Buchholz, J. P. Muyrers, F. A. Stewart, *Nat.Genet.* **1998**, 20 123-128.
- [239.] Ed.: D. A. Hopwood) **2009**, p. pp. 1-581.
- [240.] C. Rowe, I. Bohm, I. Thomas, B. Wilkinson, B. Rudd, G. Foster, A. Blackaby, P. Sidebottom, Y. Roddis, A. Buss, J. Staunton, P. Leadlay, *Chem.Biol.* **2001**, 8 475-485.
- [241.] *Nat.Chem.Biol.* **2011**, 7 461-468.
- [242.] A. R. Gallimore, C. B. Stark, A. Bhatt, B. M. Harvey, Y. Demydchuk, V. Bolanos-Garcia, D. J. Fowler, J. Staunton, P. F. Leadlay, J. B. Spencer, *Chem.Biol.* **2006**, 13 453-460.
- [243.] H. Ikeda, S. Omura, *Chem.Rev.* **1997**, 97 2591-2609.
- [244.] Clerc, F. F., Dubroeuq, M. C., Helynck, G., Leboul, J., and Martin, J. P. FR2720066. 1-11-1995. Ref Type: Patent
- [245.] M. Sosio, H. Kloosterman, A. Bianchi, P. de Vreugd, L. Dijkhuizen, S. Donadio, *Microbiology* **2004**, 150 95-102.
- [246.] H. Itokawa, K. Takeya, Y. Hitotsuyanagi, H. Morita, *The Alkaloids* **1997**, 49 301-387.
- [247.] R. Kannan, D. H. Williams, *J.Org.Chem.* **1987**, 52 5435-5437.
- [248.] U. Schmidt, R. Meyer, V. Leitenberger, A. Lieberknecht, H. Griesser, *Journal of the Chemical Society-Chemical Communications* **1991**, 275-277.
- [249.] K. C. Nicolaou, C. N. C. Boddy, S. Brase, N. Winssinger, *Angewandte Chemie-International Edition* **1999**, 38 2097-2152.
- [250.] C. Kegler, K. Gerth, R. Müller, *J.Biotechnol.* **2006**, 121 201-212.
- [251.] L. Chill, Y. Kashman, M. Schleyer, *Tetrahedron Tetrahedron* **1997**, 53 16147-16152.
- [252.] N. Fusetani, S. Matsunaga, H. Matsumoto, Y. Takebayashi, *J.Am.Chem.Soc.* **1990**, 112 7053-7054.

-
- [253.] F. Itagaki, H. Shigemori, M. Ishibashi, T. Nakamura, T. Sasaki, J. Kobayashi, *J.Org.Chem.* **1992**, 57 5540-5542.
- [254.] E. W. Schmidt, C. Raventos-Suarez, M. Bifano, A. T. Menendez, C. R. Fairchild, D. J. Faulkner, *J.Nat.Prod.* **2004**, 67 475-478.
- [255.] T. Aoyagi, H. Tobe, F. Kojima, M. Hamada, T. Takeuchi, H. Umezawa, *J.Antibiot.* **1978**, 31 636-638.
- [256.] S. Shak, N. O. Reich, I. M. Goldstein, P. R. Ortiz de Montellano, *J.Biol.Chem.* **1985**, 260 13023-13028.
- [257.] D. H. Rich, B. J. Moon, S. Harbeson, *J.Med.Chem.* **1984**, 27 417-422.
- [258.] Y. Aoyagi, T. Sugahara, *Phytochemistry Phytochemistry* **1988**, 27 3306-3307.
- [259.] T. S. Y. K. S. K. H. H. K. N. a. H. K. T.Kawaguchi, *Tetrahedron* **2010**, 66 504-507.
- [260.] L. Y. Wang, B. S. Yun, N. P. George, E. Wendt-Pienkowski, U. Galm, T. J. Oh, J. M. Coughlin, G. D. Zhang, M. F. Tao, B. Shen, *J.Nat.Prod.* **2007**, 70 402-406.
- [261.] M. S. Searle, J. G. Hall, L. P. G. Wakelin, *The Biochemical Journal* **1988**, 256 271-278.
- [262.] R. Ueoka, Y. Ise, S. Ohtsuka, S. Okada, T. Yamori, S. Matsunaga, *J.Am.Chem.Soc.* **2010**, 132 17692-17694.
- [263.] X. Yin, K. L. McPhail, K. J. Kim, T. M. Zabriskie, *ChemBioChem* **2004**, 5 1278.
- [264.] N. Grammel, K. Pankevych, J. Demydchuk, K. Lambrecht, H. P. Saluz, U. Keller, H. Krugel, *Eur.J.Biochem.* **2002**, 269 347-357.
- [265.] R. D. Firn, C. G. Jones, *Nat.Prod.Rep.* **2003**, 20 382-391.
- [266.] N. Khalid, K. H. Wolfe, *International Journal of Evolutionary Biology* **2011**, 2011 Article ID 423821.
- [267.] N. Khaldi, J. Collemare, M. H. Lebrun, K. H. Wolfe, *Genome Biology* **2008**, 9 article R18.
- [268.] S. Wong, K. H. Wolfe, *Nature Genetics* **2005**, 37 777-782.
- [269.] N. J. Patron, R. F. Waller, A. J. Cozijnsen, D. C. Straney, D. M. Gardiner, W. C. Nierman, B. J. Howlett, *BMC Evolutionary Biology* **2007**, 7 article 174.
- [270.] J. V. Lopez, *Mol Gen Genomics* **2003**, 270 420-431.
- [271.] K. J. Weissman, R. Müller, *ChemBioChem* **2008**, 9 826-848.
- [272.] C. Li, K. E. Roeger, W. L. Kelly, *ChemBioChem* **2009**, 10 1064-1072.
- [273.] I. Müller, S. Weinig, H. Steinmetz, B. Kunze, S. Veluthoor, T. Mahmud, R. Müller, *ChemBioChem* **2006**, 7 1197-1205.

- [274.] Gaitatzis, N., Silakowski, B., Kunze, B., Nordsiek, G., Blöcker, H., Höfle, G., and Müller, R. The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. *Journal of Biological Chemistry* **277**[15], 13082-13090. 2002.
- [275.] B. Frank, S. C. Wenzel, H. B. Bode, M. Scharfe, H. Blöcker, R. Müller, *Journal of Molecular Biology* **2007**, *374* 24-38.
- [276.] J. D. Kittendorf, B. J. Beck, T. J. Buchholz, W. Seufert, D. H. Sherman, *Chemistry & Biology* **2007**, *14* 944-954.
- [277.] S. C. Wenzel, R. Müller, *Curr.Opin.Chem.Biol.* **2005**, *9* 447-458.
- [278.] I. Thomas, C. J. Martin, C. J. Wilkinson, J. Staunton, P. F. Leadlay, *Chem.Biol.* **2002**, *9* 781-787.
- [279.] C. Olano, B. Wilkinson, C. Sanchez, S. J. Moss, R. Sheridan, V. Math, A. J. Weston, A. F. Brana, C. J. Martin, M. Oliynyk, C. Mendez, P. F. Leadlay, J. A. Salas, *Chem.Biol.* **2004**, *11* 87-97.
- [280.] I. H. Hardt, H. Steinmetz, K. Gerth, F. Sasse, H. Reichenbach, G. Hofle, *J.Nat.Prod.* **2001**, *64* 847-856.
- [281.] B. J. Beck, Y. J. Yoon, K. A. Reynolds, D. H. Sherman, *Chem.Biol.* **2002**, *9* 575-583.
- [282.] M. Kopp, H. Irschik, S. Pradella, R. Müller, *ChemBioChem* **2005**, *6* 1277-1286.
- [283.] O. Perlova, K. Gerth, A. Hans, O. Kaiser, R. Müller, *J.Biotechnol.* **2006**, *121* 174-191.
- [284.] G. L. Tang, Y. Q. Cheng, B. Shen, *J.Nat.Prod.* **2006**, *69* 387-393.
- [285.] B. Kunze, R. Jansen, G. Höfle, H. Reichenbach, *J.Antibiot.* **2004**, *57* 151-155.
- [286.] Buntin, K., Doktor der Naturwissenschaften (Dr.rer. nat.) thesis, Naturwissenschaftliche-Technische Fakultät III Chemie, Pharmazie, Bio- und Werkstoffwissenschaften an der Universität des Saarlandes, **2010**.
- [287.] A. M. Liu, H. Zhang, *Biochemistry-US* **2006**, *45* 10407-10411.
- [288.] Y. Kaminaga, J. Schnepf, G. Peel, C. M. Kish, G. Ben-Nissan, D. Weiss, I. Orlova, O. Lavie, D. Rhodes, K. Wood, D. M. Porterfield, A. J. Cooper, J. V. Schloss, E. Pichersky, A. Vainstein, N. Dudareva, *J.Biol.Chem.* **2006**, *281* 23357-23366.
- [289.] M. Blaesse, T. Kupke, R. Huber, S. Steinbacher, *Acta Crystallogr.D Biol Crystallogr.* **2003**, *59* 1414-1421.
- [290.] E. S. Rangarajan, Y. Li, P. Iannuzzi, A. Tocilj, L. W. Hung, A. Matte, M. Cygler, *Protein Sci.* **2004**, *13* 3006-3016.
- [291.] R. Studer, P. Dahinden, W. W. Wang, Y. Auchli, X. D. Li, P. Dimroth, *J.Mol.Biol* **2007**, *367* 547-557.

-
- [292.] D. Martynowski, Y. Eyobo, T. Li, K. Yang, A. Liu, H. Zhang, *Biochemistry-US* **2006**, 45 10412-10421.
- [293.] M. Yoshida, N. Fukuhara, T. Oikawa, *J.Bacteriol.* **2004**, 186 6855-6863.
- [294.] T. Nogawa, S. Takahashi, *The Journal of Antibiotics* **2012**, 65 123-128.
- [295.] I. Paterson, H. Hirai, C. Noti, *Chem.Asian J.* **2009**, 594-611.
- [296.] W. Li, Y. Luo, J. Ju, S. R. Rajski, H. Osada, B. Shen, *J.Nat.Prod.* **2009**.
- [297.] H. Ikeda, T. Nonomiya, M. Usami, T. Ohta, S. Omura, *P.Natl.Acad.Sci.USA* **1999**, 96 9509-9514.
- [298.] M. Oliynyk, C. B. W. Stark, A. Bhatt, M. A. Jones, Z. A. Hughes-Thomas, C. Wilkinson, Z. Oliynyk, Y. Demydchuk, J. Staunton, P. F. Leadlay, *Mol.Microbiol.* **2003**, 49 1179-1190.
- [299.] F. Perron, K. F. Albizati, *Chem.Rev.* **1989**, 89 1617-1661.

References

Author's contributions from chapter 1 to 4

Chapter 1:

The writer performed major parts of the work and wrote the manuscripts.

Chapter 2A:

Ole Revermann did most of the work and wrote the section.

Chapter 2B:

The author participated in the purification process and was involved in NMR analysis.

Chapter 3:

Ole Revermann contributed the purification of the prechondrochloren B and its structure elucidation by NMR.

Chapter 4:

The author performed the isolation of the shortened spirangiens and carried out the complete structure elucidation and wrote the document.

Abbreviations

4'-PPant	4'-phosphopantetheine
A	adenylation
ACL	acyl-CoA ligase
ACP	acyl carrier protein
AT	acyltransferase
bp	base pairs
BPC	base peak chromatogram
C	condensation
CL	coenzyme A ligase
CP	carrier protein
DH	β -hydroxyacyl dehydratase
E	epimerase
EIC	extracted ion chromatogram
ER	enoyl reductase
HC	Heterocyclization
HPLC	high performance liquid chromatography
HR-MS	high resolution mass spectrometry
Kan	kanamycin sulfate
KR	β -ketoacyl reductases
KS	ketosynthase
LC-HR-MS	liquid chromatography high resolution mass spectrometry
LB	Luria–Bertani
MALDI	matrix-assisted laser desorption ionization
MS	mass spectrometry
MT	methyltransferase
NMR	nuclear magnetic resonance
NMR	nuclear magnetic resonance
NPs	natural products
NPs	natural products
NRPs	non-ribosomal peptides

Abbreviations

NRPS	nonribosomal peptide synthetase
ORF	open reading frame
Ox	Oxidation
PCA	principal component analysis
PCP	peptidyl carrier protein
PCR	polymerase chain reaction
PKs	polypeptides
PKS	polyketide synthase
PKSs	polyketide synthases
PPTase	phosphopantetheinyltransferase
RBS	ribosomal binding site
RP	ribosomal peptide
R _t	retention time
SAM	<i>S</i> -adenosylmethionine
SIM	selective ion monitoring
TE	thioesterase
wt	wild type