Exploring the Biosynthetic Potential of *Cystobacter fuscus* – Characterization of New Structures and Studies on their Biosynthesis

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

von
Lena Etzbach
Saarbrücken
2015

Tag des Kolloquiums:	17.04.2015
Dekan:	Prof. Dr. Dirk Bähre
Berichterstatter:	Prof. Dr. Rolf Müller
	Prof. Dr. Rolf Hartmann
Vorsitz:	Prof. Dr. Uli Kazmaier
Akad. Mitarbeiter:	Dr. Martin Frotscher

Diese Arbeit entstand unter der Anleitung von Prof. Dr. Rolf Müller in der Fachrichtung 8.2, Pharmazeutische Biotechnologie der Naturwissenschaftlich-Technischen Fakultät III der Universität des Saarlandes von August 2011 bis Februar 2015.

"Twenty years from now you will be more disappointed by the things you didn't do than by the ones you did. So throw off the bowlines, sail away from the safe harbor. Catch the trade winds in your sails. Explore. Dream. Discover." Marc Twain

Danksagung

An erster Stelle möchte ich meinem Doktorvater Prof. Dr. Rolf Müller für sein mir entgegengebrachtes Vertrauen und den Freiraum für mein Projekt danken, für die Bereitstellung der außergewöhnlichen instrumentellen Ausstattung sowie für die vielen Ratschläge und die gute Zusammenarbeit. Auch meinem Co-Betreuer Prof. Dr. Rolf W. Hartmann möchte ich für die Übernahme dieser Aufgabe danken.

A big thank you goes to Dr. Alberto Plaza who was a great mentor and friend in the last years. He always encouraged me to discover new opportunities and to follow my own ways.

Bedanken möchte ich mich auch bei der ganzen Arbeitsgruppe für die gute Zusammenarbeit. Bei Dr. Sascha Baumann für die spontane Übernahme der Betreuung im letzten Jahr meiner Arbeit. Bei Katrin Jungmann und Hilda Sucipto für das Beantworten meiner vielen Fragen rund um die Molekularbiologie sowie bei Dr. Thomas Hoffmann, der mir von Anfang an immer mit Rat und Tat zur Seite stand. Außerdem bei meinen anderen „supergeilen Kollegen“, die ein wichtiger Grund dafür waren, dass ich jeden Morgen gerne zur Arbeit gegangen bin, auch wenn es in der Forschung auf und ab ging. Danke für jedes TGIF-Bier, bei dem Erfolge gefeiert und Niederlagen vergessen wurden!

Ein besonderes Anliegen ist es mir, mich bei meinen Eltern Sylvia und Karl-Heinz sowie bei meiner Schwester Britta für die langjährige Unterstützung zu bedanken. Ohne euch wäre ich nicht da, wo ich heute bin.

Zu guter Letzt möchte ich mich bei meinem Verlobten Björn Keller bedanken, ohne den diese Promotion nicht möglich gewesen wäre. Dafür, dass er mir in den letzten Jahren immer den Rücken frei gehalten hat und sich unzählige Probevorträge von mir angehört hat. Danke für das Verständnis, die unbezahlbare Unterstützung und auch für das Korrekturlesen meiner Dissertation.

Summary

Natural products provide an important basis for the drug development in particular to fight the increasing resistance of microorganisms against existing drugs. Myxobacteria, gram negative ubiquitous soil bacteria belonging to the class of γ-proteobacteria, have proven an impressive capacity for the production of chemically intriguing natural products.

In a chemical screening approach, *Cystobacter fuscus* MCy9118 was used to isolate new natural products based on chemical and structural features where two new compound families were identified and characterized.

The cystomanamides are characterized by hydroxylated amino acids, an unusual branched chain fatty acid residue as well as N-glycosylation. The macyranones are unusual peptides belonging to the class of peptidic *α'*,*β'*-epoxyketones. Macyranone A exhibits potent and selective activity against *Trypanosoma brucei rhodesiense,* the parasitic agent causing the African sleeping sickness. The target was identified as 20S proteasome by enzymatic assays. A co-crystal structure of macyranone A with the yeast 20S proteasome revealed the binding mode underlying the compound-target interaction. Macyranone A represents the successful isolation of a new natural product with identified mode of action, which can be useful for the drug development process against the African sleeping sickness.

The biosynthetic gene clusters for both compound families were identified as PKS / NRPS hybrids by in silico analysis and confirmed by gene deletion.

Zusammenfassung

Naturstoffe bilden eine wichtige Grundlage zur Entwicklung neuer Wirkstoffe für die Arzneimitteltherapie, vor allem im Kampf gegen Antibiotikaresistenzen bei Pilzen, Viren, Parasiten und Bakterien. Myxobakterien, gramnegative ubiquitäre Bodenbakterien, haben eine beeindruckende Kapazität zur Produktion von chemisch faszinierenden Naturstoffen unter Beweis gestellt.

Das Myxobakterium *Cystobacter fuscus* MCy9118 wurde in einem chemischen Screening auf die produzierten Naturstoffe hin untersucht und dabei konnten zwei neue Substanzfamilien charakterisiert werden.

Die Cystomanamide weisen als besondere Strukturelemente hydroxylierte Aminosäuren, eine verzweigte Fettsäurekette sowie eine N-Glykosylierung auf. Bei der zweiten Substanzgruppe handelt es sich um die Macyranone, sehr ungewöhnliche Verbindungen aus der Gruppe der peptidischen *α'*,*β'*-Epoxyketone. Macyranone A zeigt biologische Aktivität gegen *Trypanosoma brucei rhodesiense*, dem parasitären Erreger der Afrikanischen Schlafkrankheit. Das 20S Proteasom konnte als Target identifiziert und durch eine Cokristallisation des Macyranone A mit dem 20S Proteasom der Hefe bestätigt werden. Damit konnte ein Naturstoff isoliert werden, der mit bekanntem Wirkmechanismus als Ansatzpunkt für die Entwicklung neuer Wirkstoffe gegen die Schlafkrankheit genutzt werden kann.

Die Biosynthesegencluster beider Substanzklassen wurden durch *in silico* Analyse als PKS-NRPS Hybrid identifiziert und durch den Knockout von beteiligten Genen bestätigt.

Vorveröffentlichungen der Dissertation

Teile dieser Arbeit wurden vorab mit Genehmigung der Naturwissenschaftlich-Technischen Fakultät III, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen veröffentlicht oder sind derzeit in Vorbereitung zur Veröffentlichung:

Publikationen

Etzbach L., Plaza A., Garcia R., Baumann S., Müller R. (**2014**) Cystomanamides: Structure and Biosynthetic Pathway of a Family of Glycosylated Lipopeptides from Myxobacteria *Org. Lett.*, **16** (9), 2414-2417.

Etzbach L., Plaza A., Dubiella C., Groll M., Kaiser M., Müller R. (**2015**) Macyranones: Structure, Biosynthesis and Binding Mode of an Antiparasitic Epoxyketone that Targets the 20S Proteasome, submitted

Tagungsbeiträge

Etzbach, L., Plaza A., Garcia R., Harmrolfs K., Müller R. (**2012**) New Natural products from Myxobacteria of the Recently Discovered Genus *„Aetherobacter"* (Poster) *International Congress on Natural Products Research*, NY, USA

Etzbach, L., Plaza A., Garcia R., Harmrolfs K., Müller R. (**2012**) New Natural products from a Myxobacterium of the Recently Discovered Genus *„Aetherobacter"* (Poster) *VAAM International Workshop 2012*, Braunschweig, Germany

Etzbach, L., Plaza A., Garcia R., Müller R. (**2013**) Macyranine A: An Unusual Peptide from the Myxobacterial Genus *Cystobacter* Discovered by LC-SPE-NMR (Poster) *1st European Conference on Natural Products*, Frankfurt am Main, Germany

Etzbach, L., Plaza A., Garcia R., Baumann S., Müller R. (**2014**) Cystomanamides: Structure and Biosynthetic Pathway of a Family of Glycosylated Lipopeptides from the Myxobacterial Genus *Cystobacter* (Poster) *GRC Marine Natural Products*, Ventura, CA, USA

Etzbach, L., Plaza A., Garcia R., Baumann S., Müller R. (**2014**) Cystomanamides: A Family of Novel Glycosylated Lipopeptides from Myxobacteria (Vortrag) *Summer Symposium of the Interdisciplinary Graduate School of Natural Products Research*, Saarbrücken, Germany

Etzbach, L., Plaza A., Garcia R., Baumann S., Müller R. (**2014**) Cystomanamides: Structure and Biosynthetic Pathway of a Novel Family of Glycosylated Lipopeptides from Myxobacteria (Vortrag) *VAAM Workshop Biology of Natural Products-Producing Microorganisms*, Dresden, Germany

Weitere Publikationen

Fu C., **Etzbach L.**, Bauer A., Brönstrup M., Froidbise A., Hammann P., Herrmann J., Mondesert G., Kurz M., Schiell M., Schummer D., Toti L., Wink J., Müller R. (**2015**) Biosynthetic Studies of Telomycin Reveal Unprecedented Lipopeptides with Enhanced Activity, submitted

Table of Contents

1 Introduction

1.1 Importance of Natural Products for Drug Discovery

The idea of using natural biotic resources to treat a wide spectrum of bodily and mental defects goes back to the beginning of civilization. But only starting with the discovery of penicillin by Alexander Fleming in 1928,[1] the concept of using natural products from microorganisms as starting material for the development of drugs has spread. Even if the number of pure natural products directly used as pharmaceuticals is comparatively small with 5% of the total amount of small molecule approved drugs between 1980 and 2010, the portion of natural product derived small molecules (34%) tells a different story. Also taking into account the compounds that are synthetically derived but simulate a natural product or its pharmacophore (30%) this adds up to 64% of the total number of 1073 small molecule approved drugs (Figure 1.1).[2] These numbers impressively underline the importance of natural products for the drug discovery process.

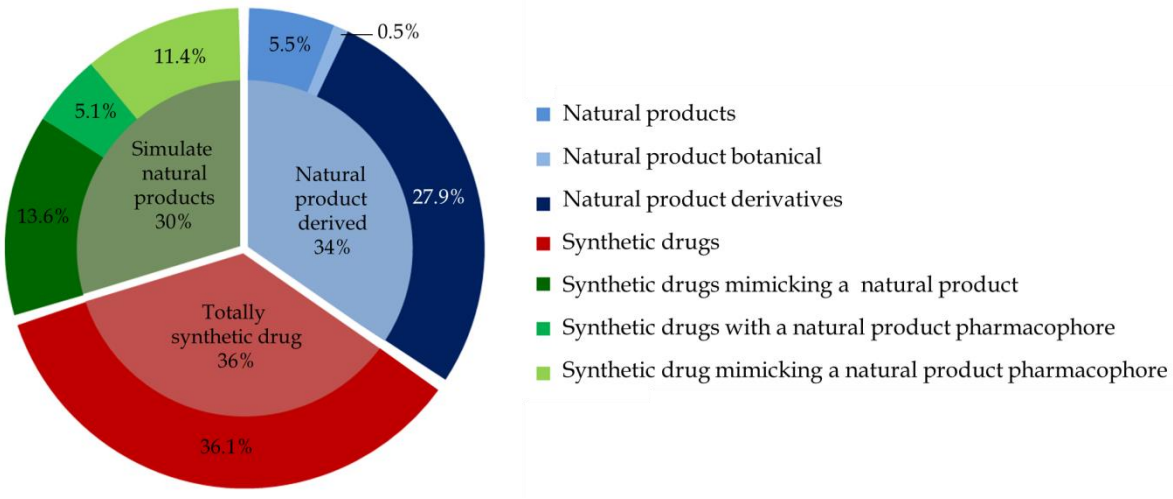


Figure 1.1: Source of small-molecule approved drugs (Figure modified from Newman & Cragg)[1]

Natural products play a particular role in the fight against infectious diseases that comprise infections caused by bacteria, parasites, viruses and fungi. The increasing resistance against used drugs is a concern in all areas of application, as for example against antimalarial drugs as well as HIV drugs, but the rapidly emerging resistance to antibacterial drugs commonly used to treat infections are of explicit concern.[3] Currently, few bacterial classes described as ESKAPE bugs (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumanii*, *Pseudomonas aeruginosa*, and

Enterobacter species) cause the majority of US hospital infections and are able to "escape" the effects of antibacterial drugs.[4] This development is compounded by the fact that, since the golden age of antibiotics in the 1970s, the number of new antibiotics in the drug discovery pipeline has dramatically decreased. Drug resistance is not only a big issue in the treatment of infectious diseases but also in the treatment of cancer. Multidrug resistance is a major factor in the failure of many forms of chemotherapy and it affects patients with a variety of blood cancers and solid tumors, including breast, ovarian, lung, and lower gastrointestinal tract cancers.[5] How can natural products help to fight against the increasing drug resistance? The term natural product in principle refers to any organic material of biological origin but generally means a secondary metabolite. Secondary metabolites are small molecules with a molecular weight lower than around 1500 Da that are not essential for the survival, growth or development of an organism but they confer an evolutionary benefit to the producing organism.[6] In contrast to chemically synthesized compounds, secondary metabolites are evolutionary designed to interact with a broad range of biological macromolecules and therefore we find a strong correlation from natural products to biological functionality together with a broad chemical diversity.[6] Natural products significantly enlarge the chemical space of synthetic libraries and they open the door for the development of new drugs as well as for the identification of new targets.

1.2 Discovery of Natural Products from Myxobacteria

Myxobacteria are Gram-negative δ-proteobacteria commonly found in soil samples. They are renowned for their social behavior and complex life cycle including the formation of colorful fruiting bodies under starving conditions as well as their micropredatory behavior.[7] Their gliding motility facilitates colony growth and expansion, as well as behaviors such as chemotaxis, predation, and development.[8] Besides this ability to move by gliding on solid surfaces, they are famous for their capability to produce a wide range of secondary metabolites.

The ability of myxobacteria to produce antibiotic substances was already observed in 1947[9] but systematic investigations concerning the secondary metabolites produced by myxobacteria have started with the formation of the Helmholtz Center for Infection Research in 1965 (Institute for Molecular Biology, Biochemistry and Biophysics (IMB) at that time). Since then, more than 100 natural product core structures and approximately

600 structural derivatives were reported in the literature.[10] That the full potential is still to be realized is already proven by the fact that only our group published seven new compounds or compound families from myxobacteria in 2014, including the macrolide glycoside antibiotics disciformycin A and B,[11] novel derivatives of the sponge metabolites microsclerodermins,[12] the glycosylated lipopeptides cystomanamides,[13] angiolactone[14] and hyafurones[15], the siderophores hyalachelins[16] and the cystobactamids[17] adding up to a total of more than 160 new scaffolds from myxobacteria (unpublished results). Myxobacteria have proven an enormous potential for the production of secondary metabolites that exhibit mainly antifungal, antibacterial and cytotoxic activity as well as immunosuppressive, antiviral, antiplasmodial and antimalarial activity.[18] This potential is of especial interest because myxobacterial compounds often target structures that are rarely hit by compounds from other producers. A novel mode of action bears the high chance to discover agents that circumvent resistance against existing drugs. A prominent example is the cytotoxic natural product epothilone that interacts with the eukaryotic cytoskeleton.[19] It was isolated from the myxobacterium *Sorangium cellulosum* and the semi-synthetic analogue ixabepilone (Ixempra™) was approved 2007 for clinical use in the US for the treatment of aggressive breast cancer and can also be used to treat paclitaxel-resistant tumors.[20]

The potential to produce such a chemical diversity is closely connected to the remarkable size of myxobacterial genomes. It was found that large genomes are enriched with secondary metabolite biosynthesis genes as well as genes involved in regulatory processes for secondary metabolite production. Konstantinos and Tiedje also found a correlation of the genome size and the habitat which made them conclude that "…bacteria with such [large] genomes are more ecologically successful in environments where resources are scarce but diverse and where there is little penalty for slow growth."[21] This is true for the soil habitat of myxobacteria and genome sizes that frequently comprise more than 10 Mbp and up to 14.8 Mbp from *Sorangium cellulosum* So0157-2.[22] Even if little is known about the natural function of many secondary metabolites from myxobacteria, their "microbial rich" habitat as well as their micropredatory lifestyle could be reasons for their extraordinary capability to produce structurally diverse natural products. The species needs to protect its ecological niche in the highly competitive terrestrial environment.

Most of the myxobacterial secondary metabolites belong to the classes of polyketides (PK), nonribosomal peptides (NRP) or hybrids of both. They are biosynthesized by supersized multimodular assembly lines composed of polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) multienzymes involving multistep biosynthetic processes. In general, these multienzymes are composed of domains where each domain carries out one biosynthetic step involving monomer selection and activation, binding of the substrates as thioesters, chain elongation and release. The domains are organized in functional modules and each module is typically responsible for the incorporation of one building block into the growing polyketide or peptide chain. The correlation between the enzyme architecture and secondary metabolite structure is known as the principle of colinearity. Simple building blocks such as dicarboxylic acids and amino acids are used to generate a wide range of complex natural products and the chemical diversity is expanded by enzymes introducing β-branching and post-PKS and NRPS reactions such as hydroxylation, glycosylation, and epimerization.[23] The colinearity between the activities of modular domains and structure of the product allows the targeted manipulation of secondary metabolites by deletion, replacement or alteration of the single domains to study the biosynthesis or generate novel compounds.[24–26]

The genes responsible for the biosynthesis of a secondary metabolite are usually located within a specific region of the bacterial chromosome, the so called biosynthetic gene cluster.[27] Only this fact allows and simplifies the identification and investigation of complete secondary metabolite biosynthesis pathways because whole gene clusters can be identified by using signature genes.[28] The arrangement of biosynthetic genes in clusters also allows the prediction of the products and *vice versa*, the chemical structure of a secondary metabolite gives an initial indication of the assembly of the biosynthetic gene cluster. Furthermore, the gene cluster structure of natural products biosynthesis genes enables the targeted manipulation of the secondary metabolites as well as its heterologous production in a convenient host.

1.2.1 Polyketide synthases (PKS)

Polyketides are an important class of natural products. Myxobacterial members of this diverse family of compounds are the already mentioned disciformycins[11] as well as stigmatellin[29], spirangien[30], jerangolid[31] and soraphen[32] (see Figure 1.2).

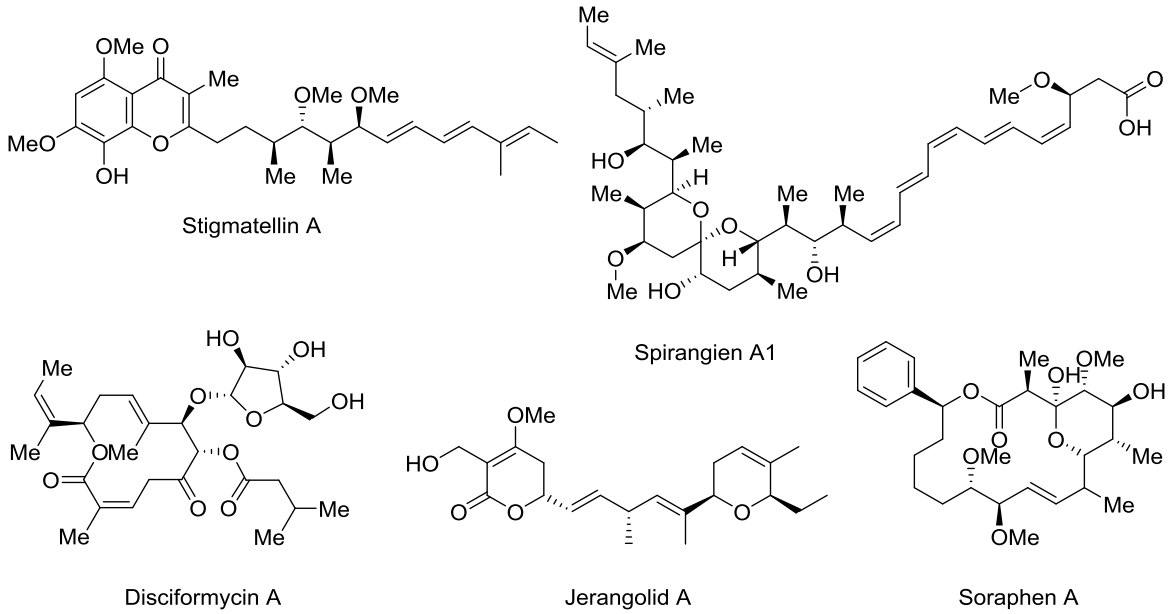


Figure 1.2: Chemical structures of polyketides from myxobacteria.

Looking at the structural diversity of polyketides, it is remarkable that they are the result of an assembly of some of the simplest biosynthetic building blocks: coenzyme-A (CoA)-activated malonate or methylmalonate. PKSs show a high degree of similarity to FAS, both in the nature of the chemistry used in chain extension from simple precursors and in the character of the enzymes that are used for chain assembly.[33] Research around fatty acid synthases (FAS) has greatly enhanced our understanding of polyketide synthase (PKS) mechanisms.[34] The differences are caused by the higher flexibility of PKS in terms of choice of starter and extender units as well as the control of the reductive cycle (see below).[35] PKS are divided into subgroups following the example set by FAS. Type I PKS are large multifunctional proteins each consisting of a series of catalytic domains. Type II PKS consist of discrete monofunctional enzymes with particular functions that mainly result in the biosynthesis of aromatic polyketides. The third type, the chalcone synthase type, consists of small homodimeric proteins.

Type I PKS can be divided into two subgroups, the iterative PKS that is primarily found in fungi and the modular PKS that form giant multimodular megasynthases which are mainly found in prokaryotes.[36] Polyketides can be seen as linear polymers where the polymer backbone consists of C2 units. During biosynthesis, every second carbon bears a ketone function, thus calling the compounds polyketides. The ketone groups can subsequently be processed via ketoreduction to hydroxy groups, via dehydration to the double bond and via enoyl reduction of the double bond to the fully reduced form.

Modular PKS pathways in homology to FAS pathways consist of several modules following the colinearity principle mentioned above. Each PKS module is responsible for the addition and manipulation of one polyketide subunit and the final number of ketide units in polyketides synthesized by regular type I (modular) PKS therefore equals the number of modules found in the PKS. Every module consists of several domains where each domain is responsible for one catalytic step. The action of three core domains – a keto synthase (KS) domain, an acyltransferase (AT) domain, and an acyl carrier protein (ACP domain) - is the minimum requirement for one elongation step (see Figure 1.4A) but further domains can be introduced into the module to modify the incorporated subunit.[36] In a typical elongation module consisting of KS-AT-ACP, the KS domain is acylated either with the starter unit or the growing polyketide chain from the previous module. The AT domain is responsible for the selection of the extender unit (most commonly malonyl- or methylmalonyl- CoA) and catalyzes its transfer to the acyl carrier protein (ACP). ATs of modular type I PKSs are often selective for extender units and are also able to select various unusual extender units like hydroxymalonyl-CoA, isobutyrylmalonyl-CoA, hydroxymalonyl-CoA and aminomalonyl-CoA.[37–40] Selection of extender units can partly be predicted based on specific residues in the active site of the AT domain.[41] Through a thioester linkage to the phosphopantetheine arm (ppant arm), the ACP domain is able to shuttle extender units and polyketide intermediates between the various domains. The ppant arm is a 4'-phosphopantetheine co-factor which is post-translationally attached to the active-site serine of the ACP apoenzyme via phosphopantetheinyl transferases (PPTases).[42] The ACP domain docks to the acylated ketosynthase (KS) domain, where the magic of type I PKS happens. The KS domain catalyzes a decarboxylative condensation to form a carbon-carbon bond in a Claisen condensation step and thus builds up the carbon skeleton of the growing polyketide. The ACP shuttles the intermediate to processing enzymes such as a methyltransferase (MT) domain, which can transfer a methyl group to the α-carbon, a ketoreductase (KR) domain, which can reduce the *β*-keto group, a dehydratase (DH) domain, which can catalyze dehydration to yield a double bond between the α and the β carbon, and an enoyl-reductase (ER) domain, which can stereoselectively reduce the double bond. The ACP subsequently forwards the processed polyketide to the KS of a downstream module or a chain-releasing thioesterase (TE) domain. Only recently, researchers from the University of Michigan uncovered the structure of a whole PKS module by electron cryo-microscopy and thereby explained the

structural mechanism underlying the ACP shuttle machinery. It has been shown that the whole PKS module creates a reaction chamber for the ACP domain to allow the transport of building blocks and intermediates between the catalytic domains of one module.[43,44]

The common starter units of type I PKS are malonyl- or methylmalonyl-CoA, but varying starter units enlarge the chemical diversity produced by PKS. Amino acid-derived short branched chain carboxylic acids like isobutyryl-CoA, 2-methylbutyryl-CoA or isovaleryl-CoA as shown for the biosynthesis of the myxobacterial natural products myxothiazol[45] and the herein presented cystomanamides[13] as well as cyclohexanecarboxylic acid as shown for the biosynthesis of manumycin-type antibiotics[46–48] are two representatives for the extreme diversity of PKS starter units.[49]

Although the modular PKS is the most abundant form in myxobacteria, myxobacterial biosynthetic pathways frequently show unusual genetic and biochemical features.[27] The iteratively acting module in a modular type I PKS pathway in the stigmatellin biosynthesis in *Stigmatella aurantiaca* Sg a15[50] is only one example for the unusual features found in myxobacterial pathways.

1.2.2 Nonribosomal peptide synthetases (NRPS)

Nonribosomally biosynthesized peptides are a class of peptide secondary metabolites that are characterized by some structural features. They can contain proteinogenic as well as non-proteinogenic amino acids like D-amino acids[13] and ornithine[51,52] and are often macrocyclic[53] or branched macrocyclic[54]. They contain heterocyclic rings like thiazoline and oxazoline and they bear modifications like *N*-methylation, *N*-formylation, halogenation, hydroxylation, and glycosylation.[55] Examples for unusual structures produced by NRPS in myxobacteria are the argyrins,[53] myxovalargins[56] and vioprolides[57] (see Figure 1.3).

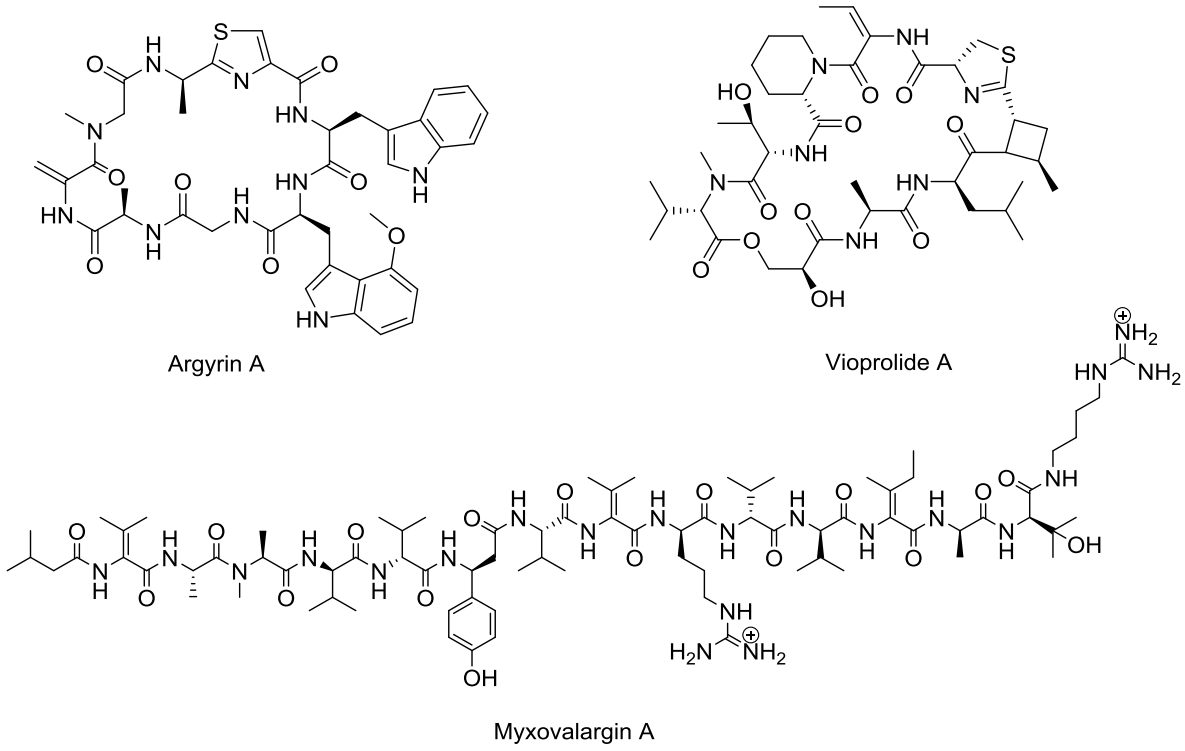


Figure 1.3: Chemical structures of nonribosomal peptides from myxobacteria.

NRPS show a modular structure with strong similarity to modular type I PKS but instead of carboxylic acids, amino acids serve as building blocks. In analogy to PKS, three domains are needed in a minimal NRPS module to accomplish one cycle of chain elongation (Figure 1.4B). The adenylation (A) domain is responsible for the selection of the substrate amino acid and for its activation as amino acid adenylate.[58] A domains were found to bear highly conserved core motifs and Stachelhaus *et al.* identified ten amino acids residues by analysis of the binding pocket and comparison with other A domains to be responsible for the substrate specificity. Bioinformatics tools were developed based on this "Stachelhaus code" to predict the substrate specificity of A domains.[59] The peptidyl carrier protein (PCP) domain, in analogy to the ACP domain from PKS, covalently binds the activated amino acid as a thioester to the ppant arm which then acts as a shuttle to reach the various catalytic centers.[42] The condensation (C) domain recognizes on the one hand the growing peptide chain bound to the previous PCP domain and on the other hand the newly activated amino acid and catalyzes the amide bond formation to result in an elongated peptide chain.[60] Various functional subtypes of the condensation domain exist that catalyze either the peptide bond between two L-amino acids, between D- and L-amino acids, heterocyclization domains that catalyze both the amide bond formation as

well as the subsequent cyclization of cysteine, serine or threonine residues and dual epimerization / condensation domains that catalyze both the epimerization and amide bond formation. Recently, starter C domains were described that acylate the first amino acid with a *β*-hydroxy-carboxylic acid.[61] To produce the several hundred different building blocks found in NRPS products, a huge number of optional domains as well as modifying enzymes were identified up to date including optional domains responsible for epimerization (E), cyclization (Cy), oxidation (Ox), reduction (R), N-methylation (Mt) and formylation (F) as well as modifying enzymes for glycosylation, halogenation and reduction.[62]

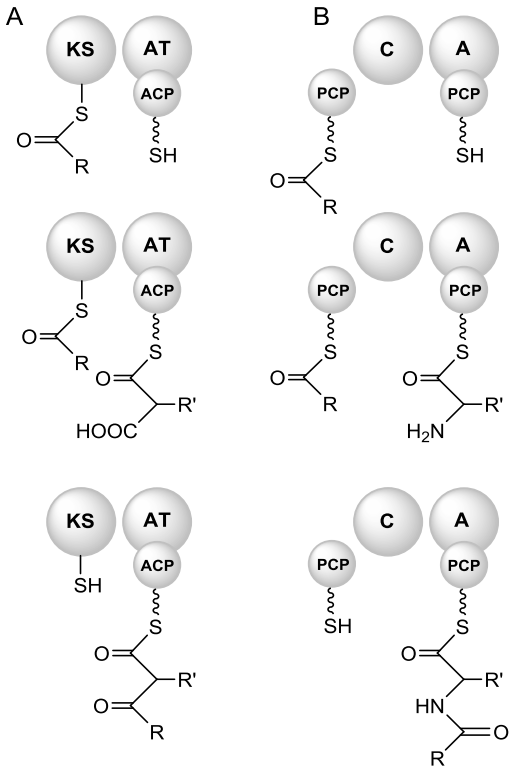


Figure 1.4: **A**: Three core domains are minimally required for one polyketide synthase module: an acyltransferase (AT) domain, an acyl carrier protein (ACP) domain and a ketosynthase (KS) domain. The AT domain selects the extender unit and attaches it to the phosphopantetheine arm (ppant arm) of the ACP domain. The acylated KS domain catalyzes the central Claisen condensation to form the C-C bond to connect the building blocks. **B**: In a similar fashion, the three minimally required domains in nonribosomal peptide synthetases are a condensation (C), an adenylation (A) and a peptidyl carrier protein (PCP) domain. The growing chain is recognized by the C domain while bound to the PCP domain of the previous module. The elongation unit is selected and activated by the A domain and attached to the ppant arm of the PCP domain. The condensation domain then catalyzes the amide bond formation (Figure modified from Khosla & Harbury, 2001[63]).

1.2.3 Mixed PKS/NRPS hybrids

One highlight in myxobacterial natural products is that their biosynthesis is frequently directed by mixed PKS–NRPS systems whereas for example actinomycetes and *Bacilli* predominantly synthesize pure PK or NRP natural products.[64] Myxobacteria use the combination of both strategies to create a structural diversity unique in the world of natural products. They take advantage of the full chemical diversity by combining unusual carbon skeletons produced by PKS with the diverse structures produced by NRPS. The most prominent representatives of the mixed PK/NRP are the cytotoxic epothilones,[65] but also the cystomanamides and macyranones presented in this thesis belong to the diverse and highly interesting group of mixed PKS/NRPS natural products (see Figure 1.5).

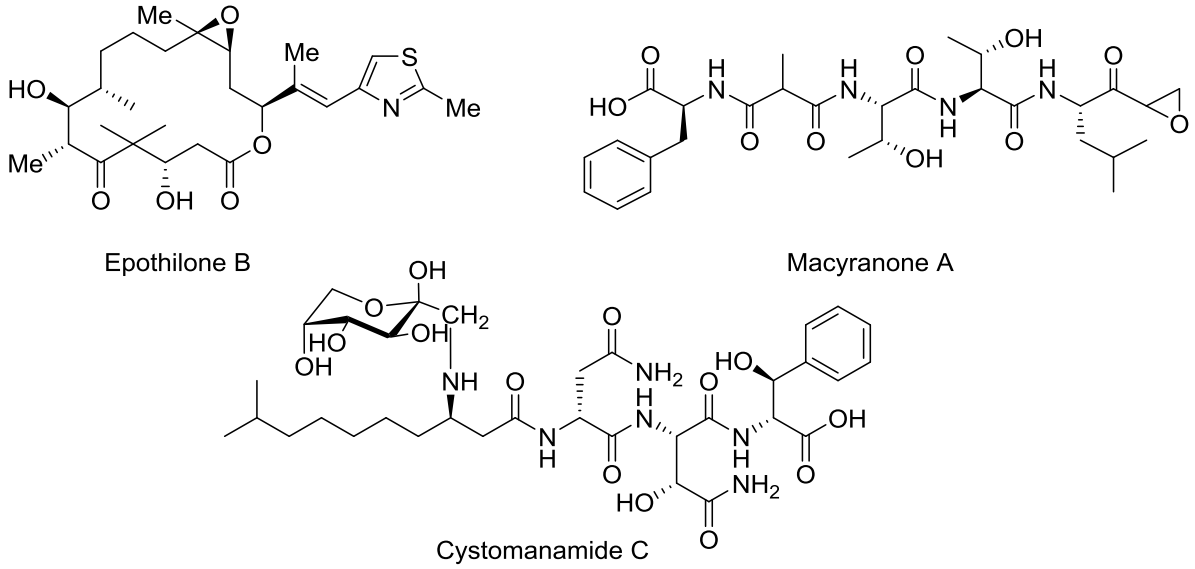


Figure 1.5: Chemical structure of mixed PK/NRP natural products from myxobacteria.

1.3 Improving Natural Product Discovery

The above mentioned reasons explain the need to isolate new natural products and researchers around the world have been successful to identify a vast number of natural products. Nowadays, natural product chemists are facing a challenge since reisolation rates of already known secondary metabolites or derivatives thereof are increasing and the discovery rate of novel compounds is decreasing using conventional bioactivity-guided isolation approaches. Low-abundant bioactive natural products, natural products with no or unusual bioactivity and products of silent gene clusters fall through the cracks

of a conventional screening approach. By the use of novel approaches that take advantage of recent developments in the diverse field of natural products research, novel scaffolds can still be isolated from myxobacteria and other sources that have not been analyzed before.

What are the main factors that have changed in the recent years influencing natural product discovery? Approaches that are used to isolate new natural products can be assigned to three generic principles:

I. Bioactive natural products can be isolated using a biological screening approach following the biological activity throughout the fractionation process (bioactivity-guided isolation).
II. Genome mining can be used that allows the identification of biosynthetic gene clusters in the genome of an organism followed by the determination of the associated compound(s).
III. The produced secondary metabolites can be screened by their chemical and structural features to allow the identification of new chemical entities. All three approaches share some working steps like the imperative necessity to isolate the new natural product and characterize it using different spectroscopic techniques including 1D and 2D NMR, HR-MS, and chemical derivatization but the way to get to this new natural product is widely differing (Figure 1.6).

All three approaches profit from a wide range of investigated phylogenetic diversity since biodiversity is closely linked to the discovery of chemical diversity.[66,67] The phylogenetic diversity is a measure for biodiversity based on evolutionary relationships and represents the variety of organisms. The screening of rare myxobacterial families and genera from geographically and ecologically promising habitats is therefore an encouraging method to discover novel natural products.[68]

Apart from that, researchers have witnessed major technological advances within the fields in the last years that have led to improvements in the realization of all three principles.

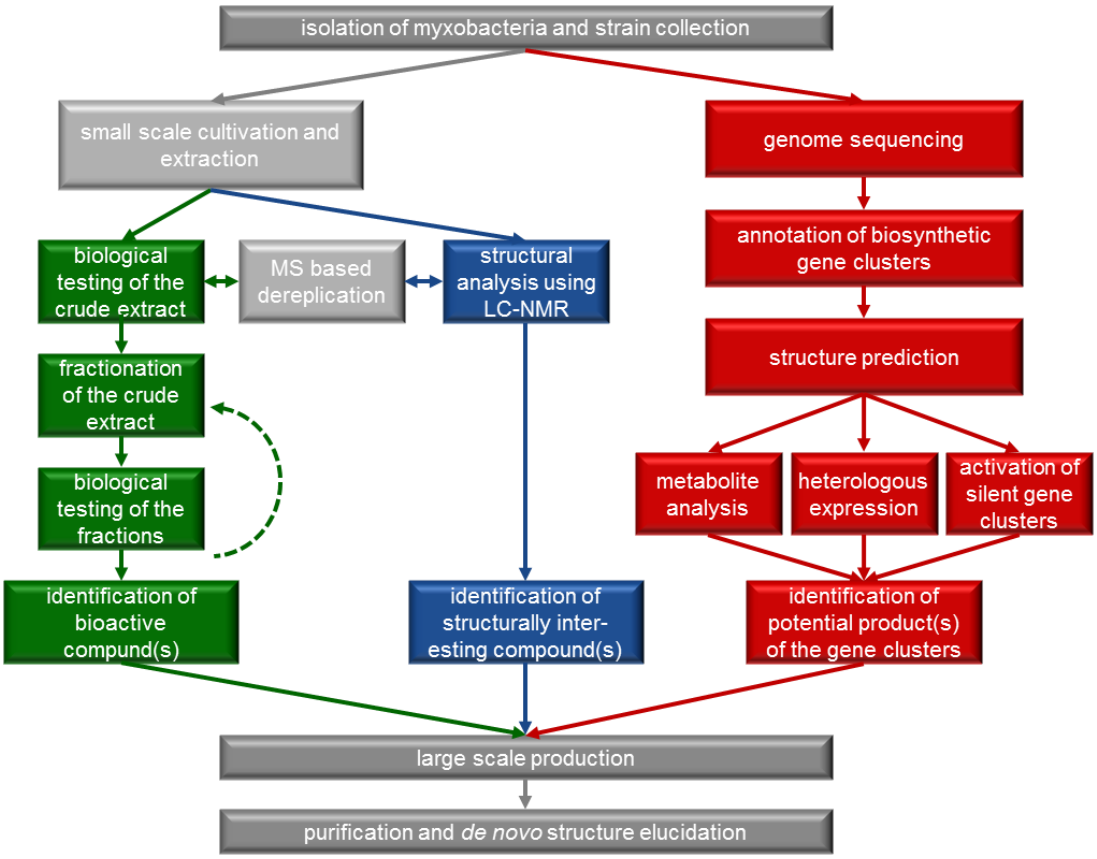


Figure 1.6: Schematic workflow of the work steps of the introduced approaches to isolate new natural products. Green: biological screening; blue: chemical screening; red: genome mining, light gray: shared by two approaches, dark gray: shared by all three approaches.

1.3.1 Biological screening

The conventional bioactivity-guided isolation has been successful throughout decades. Myxothiazol, one of the first secondary metabolites from myxobacteria, was isolated in 1980 by Reichenbach et al. using a paper-disc method against *Mucor hiemalis* to narrow down the antifungal activity of the crude extract.[69] In a classical workflow, crude extracts are prepared from small-scale cultivation. These extracts are subjected to the primary biological screening to test it for e. g. antimicrobial, cytotoxic, or antiviral properties. Extracts are then fractionated by chromatographic techniques and the fractions are tested again in biological assays. Successive rounds of bioactivity-guided fractionation are necessary to narrow down the biological activity to one specific compound. Once the active target compound is identified, the microorganism is fermented in larger scale for subsequent purification and *de novo* structure elucidation as well as further bioactivity testing.

Major progress was made with respect to high-throughput screenings (HTS) to assay the biological activity using a wide range of organisms, cells, viruses, organelles and enzymes. Not only the availability of targets was enhanced but also the detection methods were developed further to include for example fluorescence cellular imaging[70] and luciferase-expression[71] based systems. One of the main issues in the process is the complexity of crude extracts that complicates the development of robust HTS methods. Ingredients may interfere with the assay read-out (e. g. fluorescent or colored compounds), be insoluble or unstable or may cover the bioactivity of minor compounds. Another improvement in the field of bioactivity-guided fractionation represents the automated HPLC-MS fractionation where fractions can be trapped in HTS compatible 96-well plates, a process that reduces the laborious workload of this screening approach.

1.3.2 Genome mining

With the fast-growing number of fully sequenced microbial genomes starting with the first complete genome sequence of *Haemophilus influenzae* in 1995,[72] it became clear that the biosynthetic potential of microorganisms was largely underestimated. Even organisms that were previously not assumed to be secondary metabolite producers include huge biosynthetic gene clusters in their genomes and have the capacity to produce unknown natural products.[73] The bioinformatics annotation of biosynthetic gene clusters made clear that also for myxobacteria, there is a large gap between known secondary metabolites from a strain and the amount of present secondary metabolite gene clusters.[74] Under laboratory conditions, a large portion of secondary metabolites is either not produced or the amount is too small for identification. Besides the genome-independent approaches of varying the culture conditions[75] and co-cultivation with other organisms,[76] the genetic manipulation of genes responsible for gene regulation[77,78] is a potential target to enable production of the respective secondary metabolites. Another promising method is the heterologous expression of an unknown biosynthetic pathway in a convenient heterologous host to identify new natural products.[79] This approach requires a detailed knowledge of biosynthetic secondary metabolite pathways as well as the development of efficient tools to unravel the biosynthetic machinery.

1.3.3 Chemical screening

Natural products possess an enormous chemical and structural diversity so that they represent a rich source of novel molecular scaffolds. The chemical screening approach allows the search for interesting structural elements irrespectively of biological activity. The focus lies on the identification of new scaffolds in a structure-guided isolation since new scaffolds are closely related to finding novel modes of action as exemplified by the FabF inhibitor platensimycin isolated from *Streptomyces platensis*.[80] Another example where a natural product goes hand in hand with the discovery of a new mechanism of action is argyrin A, a cyclical peptide derived from *Archangium gephyra*.[53] The myxobacterial compound acts by stabilizing the cyclin kinase inhibitor $p27^{kip1}$, a tumor suppressor protein.[81]

The success of the chemical screening approach highly depends on the quality of the analytical techniques, it benefits from modern state-of-the-art LC-MS instrumentation as well as a highly sensitive NMR platform. The traditional way of studying chemical features of natural products includes fractionation of an extract based on ^{1}H-NMR, isolation of the compound by liquid chromatography and structure elucidation using various spectroscopic methods (NMR-guided isolation). More recently, the trend is towards hyphenated techniques. The development of hyphenated techniques that combine the separation of compounds using liquid chromatography (LC) with nuclear magnetic resonance (NMR) spectroscopy were already started in the 1970s, where the first stopped-flow application was shown.[82] Up to now, LC-NMR with on-flow, stopped-flow and loop-storage procedures were successfully installed,[83] but a major development in the field was the use of solid-phase extraction (SPE) systems at the interphase of LC and NMR.[84] The increasing NMR spectrometer field strength and methods for adequate solvent suppression also contributed substantially to the development of applicable systems.[85,86]

Liquid chromatography (LC)-solid-phase extraction (SPE)-nuclear magnetic resonance (NMR)-mass spectrometry (MS) coupling, short LC-SPE-NMR-MS, became a key technique used to screen complex crude extracts for novel natural products with interesting structural features. It allows for rapid insights into structural elements of an unknown compound present in minute quantities. Even if the technique is in most cases not suitable for the full structure elucidation of complex natural products, the laborious and time-consuming pre-fractionation and isolation of known or uninteresting

compounds can be circumvented by taking advantage of the full flexibility regarding chromatographic conditions and NMR acquisition methods.[84]

The concept of the hyphenated technique is to collect as many information as possible in one individual operation. The complex starting material is injected into the LC system, where a separation of the compounds is achieved in a reversed-phase HPLC system. Peaks are detected using UV as well as high-resolution MS detection. The process involves the addition of water to the mobile phase via a make-up pump after elution from the LC column to allow retention of the target compound on the cartridge. The system allows the trapping and enrichment of compounds on the cartridges by multiple subsequent chromatographic injections. The compounds are then dried with nitrogen gas and eluted with an appropriate deuterated solvent, typically acetonitrile-*d3* or methanol-*d4* to achieve a high spectral quality. Through a capillary system, the compound is pumped directly into the NMR flow probe for data acquisition (Figure 1.7). The development of cryogenic flow probes highly improved the technique.[87]

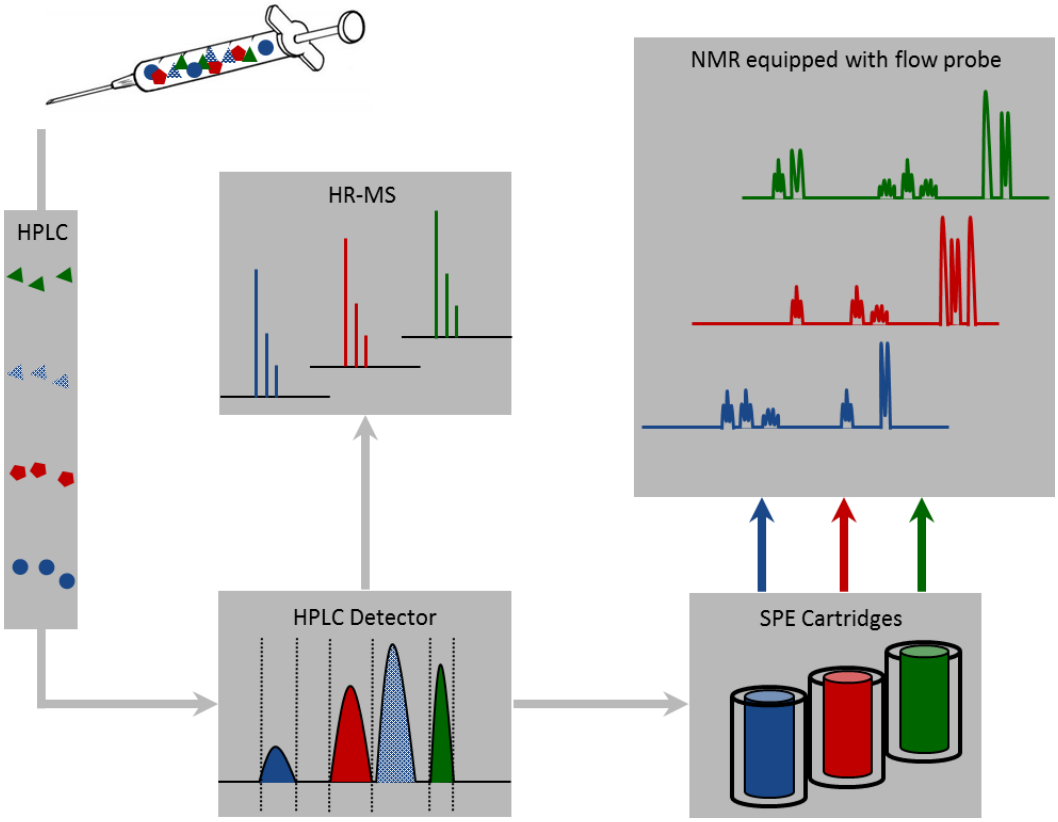


Figure 1.7: Schematic illustration of the LC-SPE-NMR-MS process

1.4 Structure Elucidation of Natural Products

50 years ago, structure elucidation of unknown compounds required decomposition into known fragments or total chemical synthesis to corroborate proposed structures. Gram quantities and years of tedious effort were needed to get insights into (often mistaken) structures.[88] Nowadays, the time to characterize an unknown compound has decreased dramatically and also the amount of sample required for this purpose has decreased, depending on the complexity of the structure sometimes to less than 1 mg.

Commonly, *de novo* structure elucidation is a task for nuclear magnetic resonance spectroscopy even though methods like infrared spectroscopy and X-ray crystallography are also useful to complement data obtained from NMR. Another standard analytical technique used in the full *de novo* structure elucidation of unknown secondary metabolites is mass spectrometry. It is useful to determine the elemental composition of a molecule as well as gaining insights into partial structures using MSn experiments.[89]

1.4.1 Nuclear magnetic resonance (NMR) spectroscopy

NMR is the most powerful and informative analytical tool in natural products chemistry with a wide range of applications. The spectroscopic technique underwent an enormous development since its first description by Isidor Rabi in 1938.[90] The technique uses the magnetic properties of certain nuclei and has become an indispensable part of pharmaceutical research.[91] Several advances of the technique have been awarded with the Nobel Prize (see Table 1.1).

Table 1.1: List of Nobel Prize winners with scientific contributions concerning NMR (http://www.nobelprize.org)

Nobel Prize winners	Year	Contribution
Otto Stern	1943	discovery of the magnetic moment of the proton
Isidor Rabi	1944	first NMR measurements
Felix Bloch Edward M. Purcell	1952	expanding the technique for use on liquids and solids
Richard R. Ernst	1991	development of high resolution NMR spectroscopy
Kurt Wüthrich	2002	development of NMR-methods for 3D protein structure determination
Paul C. Lauterbur Peter Mansfield	2003	discoveries concerning magnetic resonance imaging

Since the first commercial spectrometer from Varian with 30 MHz became available in 1952, the technique underwent manifold developments concerning pulse sequences, acquisition, processing methods and particularly hardware. Whereas the continuous wave (CW) mode was applied in the beginning of the technique that especially restricted the ability to obtain ^{13}C spectra, pulsed Fourier transform NMR[92,93] (FT-NMR) spectroscopy is the state of the art nowadays.[94] The resonances are not anymore measured using a fixed magnetic field and sweeping the frequency of the electromagnetic radiation but all nuclei are excited at the same time by a short radio frequency pulse. After the radio frequency pulse, the exited nuclei emit the absorbed radiation causing a signal. The free-induction decay (FID) is recorded as a function of 'time after the pulse'. The FID contains the sum of NMR responses from all the exited nuclei with a superposition of all excited frequencies, resulting in a time-domain signal. This signal cannot be interpreted directly but it is possible to extract the individual resonance frequencies and to convert the signal into the familiar frequency domain spectrum using a mathematical operation, the Fourier transformation (Figure 1.8). Since the signal contains information on all resonance frequencies in a single measurement in contrast to CW measurements, the signal can be repeatedly measured in a shorter time and summed up resulting in an increased signal to noise ratio compared to the CW measurements. This allows the structural analysis of much smaller amounts of compound as well as measurements of low abundant isotopes like carbon-13.[95] Another important benefit and a direct consequence from the introduction of FT-NMR was the development of multiple pulse sequences in 1976 that greatly increased the applicability of NMR for natural product studies.[96]

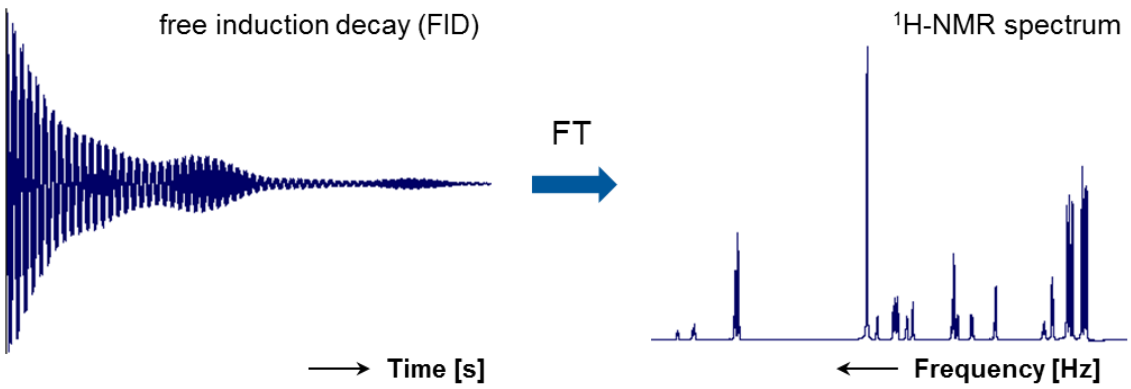


Figure 1.8: The diagram shows an FID signal from proton-nuclei obtained with the PFT technique. By Fourier transformation (FT), the time-domain signal is transformed into the frequency domain signal resulting in the ^{1}H-NMR spectrum.

Breton and Reynolds[97] nicely summed up the points that dramatically lowered the necessary amounts of samples for structure elucidation since 1984:

- New pulse sequences
- Higher field magnets due to the development of superconducting alloys
- Cryogenic probes with enhanced sensitivity
- Increased speed and power of computers
- Improvements in electronic components
- Development of superconducting alloys

As a result, high field magnets up to 1 GHz and cryogenic probes are nowadays improving the sensitivity of the technique to allow the full structure elucidation of natural products often accessible in minimum quantities.[98] The non-invasive and non-destructive nature of the technique raises the utility for natural products chemistry despite its inherent low sensitivity compared to optical spectroscopic methods. For the determination of the molecular structure of natural products, different NMR experiments are useful including one-dimensional NMR experiments as well as two-dimensional homo- and heteronuclear correlation experiments. Only the combination of different experiments allows it to piece together the structural information to unequivocally characterize a new natural product.

1.4.1.1 One-dimensional NMR (1D NMR)

For natural product chemists, there are two 1D NMR experiments commonly used to characterize small molecules, the proton (^{1}H) NMR and the carbon (^{13}C) NMR. 1D NMR spectra are relatively simple to analyze compared to two-dimensional spectra but provide high information content.

The ^{1}H NMR allows the identification of hydrogen atoms in an organic molecule. The chemically different protons in an organic molecule experience a different magnetic field depending on their chemical environment. There are basically three parameters[99] that can be extracted from the proton spectrum:

1) The chemical shift from a signal allows precise conclusions concerning the neighborhood of this hydrogen and its association with functional groups.
2) The splitting pattern, so the fine structure of a signal, gives information about the number of neighboring spin-coupled (hydrogen) nuclei of the protons causing the signal as well as about the geminal and vicinal H, H coupling constants. Since the

vicinal ^{3}J(H,H) coupling constants highly depend on the dihedral angle as formulated by Karplus, they provide important information for the conformational analysis of molecules (see also 1.4.1.2).[100]

3) The intensity of resonance signals gives information about the number of chemically equivalent protons that cause a signal and can be measured by integration. Besides the information about the number of protons, the intensity of the signals also allows the quantitative analysis for example of different components in natural products mixtures.

^{13}C-NMR is a useful tool but the experiment is much less sensitive than ^{1}H NMR for two reasons: The low natural abundance of the ^{13}C isotope with only 1.1 % of all natural carbon is the first reason for the low sensitivity. Since the most abundant ^{12}C isotope has a spin quantum number of zero and is therefore NMR silent, only the low abundant ^{13}C isotope with a spin quantum number of ½ can be observed. Secondly, the intrinsic sensitivity of the ^{13}C isotope, expressed by the gyromagnetic ratio, is only around 1/4th of the proton adding up to a relative sensitivity to the proton observation of about 1.76 X 10^{-4}.

To simplify analysis of the ^{13}C spectra and to avoid dividing the already low intensities of the ^{13}C signals into multiplets due to coupling with hydrogen nuclei, ^{1}H broad-band decoupling is commonly used. In a proton-decoupled spectrum, all carbons appear as single peaks. As a result, the intensity is largely increased and the spectra recorded by this method contain only one parameter, the chemical shift.[101]

The editing of ^{13}C spectra based on the number of attached protons can be very helpful to gain more information from the spectra. APT,[102] INEPT[103] or DEPT[104] spectral editing sequences can be used to allow the differentiation of methyl, methylene, methine and quaternary carbons. The APT sequence provides partial spectral editing to discriminate methylene or quaternary carbons from methyl or methine carbons but with a lower sensitivity of the experiment. INEPT and DEPT sequences only give peaks for protonated carbons but therefore, due to polarization transfer, are significantly more sensitive compared to standard ^{13}C measurements.[97]

1D ^{13}C NMR can often be avoided by using more sensitive proton-detected heteronuclear 2D experiments providing both carbon shift data and correlation information. The advantage of proton-detected experiments is increased by the use of an inverse probe that is optimized for the ^{1}H detection.[105]

1.4.1.2 Two-dimensional NMR (2D NMR)[99,101]

With the introduction of two-dimensional experiments,[96] structure elucidation using NMR has begun to replace the structure elucidation by chemical synthesis or decomposition into known fragments and became the method of choice for natural products chemistry. The use of 2D NMR experiments significantly increased the amount of information which can be extracted from the complex spectra.

Homonuclear ^{1}H-^{1}H correlation experiments are the most frequently used ones providing both a high degree of sensitivity and excellent information content. The ^{1}H,^{1}H-COSY (Correlation SpectroscopY), that was the first two-dimensional sequence proposed, correlates the chemical shifts of two protons which are separated from each other with two, three and rarely four chemical bonds.[106] A correlation assumes the spins to share a mutual *J*-coupling. The 2D-TOCSY (TOtal Correlation SpectroskopY) or HOHAHA (HOmonuclear HArtmann-HAhn) experiment allows to identify all protons belonging to a common coupled spin system.[107] The experiment is based on magnetization transfer along a sequence of coupled protons and depends on sizes of ^{1}H, ^{1}H couplings. In case of strongly overlapped proton signals, a hybrid HSQC-TOCSY experiment can be useful that relies on the ^{13}C resolution but at a large cost in sensitivity.[108]

Heteronuclear ^{1}H-^{13}C experiments are extremely useful in structure elucidation of natural products, since most compounds consist of a carbon-hydrogen skeleton. With a loss of sensitivity, the heteronuclear experiments can also be used for proton correlation to other heteroatoms like ^{15}N. Early heteronuclear experiments were acquired by ^{13}C detection (^{13}C-^{1}H experiments) using the HETCOR sequence.[109] However, because of the sensitivity advantages of ^{1}H detection, ^{1}H detected experiments like HSQC[110] and HMQC[111] have established themselves. The heteronuclear 2D NMR experiment ^{1}H,^{13}C-HSQC (Heteronuclear Single Quantum Coherence), that gives a better signal/noise ratio compared to the HMQC experiment, utilizes one-bond couplings and hence identifies directly connected nuclei. While a combination of ^{1}H,^{1}H COSY and ^{1}H,^{13}C HSQC correlations would already be sufficient to solve the structure of molecules exclusively composed of protonated carbons, one type of spectrum is missing to provide the final tool for the determination of complex natural products. The heteronuclear 2D experiment ^{1}H,^{13}C-HMBC (Heteronuclear Multiple Bond Correlation) closes this gap by providing correlation spectra for detection of protons and carbons separated from each other with two, three or in special cases four chemical bonds.[112]

Whereas the already mentioned 2D NMR experiments map specifically through-bond interactions, the so called *scalar couplings*, additional experiments can also provide information about correlations through space (*dipolar couplings*) giving rise to the nuclear Overhauser effect (NOE) that causes changes in the resonance intensities. Since the NOE are only observed between those spins that are close in space (< 5 Å apart), it is important in modern structure elucidation providing information about the 3D geometry of a molecule. 2D-NOESY[113] (Nuclear Overhauser Effect SpectroscopY) and 2D-ROESY[114] (Rotating frame nuclear Overhauser Effect SpectroscopY) are the most widely used techniques to uncover the spacial interactions. The 2D-ROESY is more applicable for natural products research due to its independency of the size of a molecule.

Both techniques are used for *J* based configuration analysis[115] to gather information about the relative configuration of two neighbored stereocenters in a flexible molecule. The method uses both, ^{1}H-^{1}H and ^{1}H-^{13}C coupling constants, to assign anti or gauche relationships of vicinally substituted carbon chains. The combination with NOE allows the determination of the conformation of adjacent asymmetric centers often found in natural products.

1.4.2 Dereplication

The successful identification of unknown secondary metabolites requires the recognition of already studied compounds to eliminate them from further processing within the screening and to avoid repeated characterization. The underlying operation called dereplication is a key process in natural product screenings. The goal is to detect the known compounds as early as possible in the screening of myxobacterial strains in order to concentrate on promising unknown natural products. This is becoming increasingly important with growing reisolation rates. Since the same metabolites are often produced by multiple strains, dereplication already plays a fundamental role in the strain selection for discovery programs to save time to concentrate on producers of novel secondary metabolites.

Dereplication is done in different ways beginning from ring binders containing printouts of proton NMR or UV spectra for a manual comparison (personal experience) through the use of commercial data bases like SciFinder[116] and MarinLit[117] for comparison of molecular formulas or partial structures. It goes up to the use of mass spectral libraries like the MassBank,[118] a public repository of mass spectra of small chemical compounds for

life sciences, or the Myxobase,[119] that was developed for the efficient dereplication of myxobacterial secondary metabolites.

Myxobase contains a chemical compound database, where the dereplication is done based on the comparison of different features measured in a defined analytical system. The retention time, high-resolution *m/z* value and isotope pattern fit are used to compare secondary metabolites to compounds in the database. The system highly depends on the reproducibility of the measured features and is only as good as the quality of the applied analytical techniques. A metabolite profiling using HPLC/HRMS in combination with an appropriate database can therefore be used for dereplication as well as for strain selection. Also the previously described LC-NMR technique is a tool to accomplish the dereplication of natural products and is especially useful for compounds were the data from LC-MS do not allow the confident identification.[120]

1.5 Outline of the Study

The main topic of this work was the metabolic profiling of the myxobacterial strain *Cystobacter fuscus* MCy9118. The myxobacterium showed a wide range of biological activity and was therefore prioritized in the screening process. The goal was to unravel the full chemical diversity provided by this natural product producer. Although the identification of derivatives of known compounds is of significant importance, our aim was to identify novel basic structures ideally exhibiting new modes of action. To realize this goal, a structure based screening approach was chosen that relies on the use of high-end analytical techniques like HR-MS-based dereplication and LC-SPE-NMR-MS for the identification of new structural scaffolds. In the course of this screening, two new compound classes were identified besides several known antibiotics like althiomycin, roimatacene and myxochelin A and B.

The first parts of both Chapters 2 and 3 report the studies concerning the identification and structure elucidation of the new compound classes named cystomanamides and macyranones. Their structures were elucidated by detailed analysis of the 1D and 2D NMR as well as MS/MS data and chemical derivatization. The cystomanamides are lipopeptides characterized by β-hydroxy amino acids along with 3-amino-9-methyldecanoic acid. The fatty acid residue is *N*-glycosylated in cystomanamide C and D. The macyranones are characterized by the presence of a rare 2-methylmalonamide moiety

in the peptidic backbone and an α-amino ketone fragment including an *α′,β′*-epoxyketone in macyranone A.

The next aim of this thesis was to increase the knowledge of bacterial secondary metabolite biosynthesis. In-depth genome sequence analysis of *C. fuscus* MCy9118 gave first insights into the biosynthesis of the new natural products. Confirmation of the biosynthetic gene clusters identified as PKS/NRPS hybrids was achieved by gene deletion using a single crossover homologous recombination strategy which resulted in the abolishment of the production in both cases. Studies concerning the identification and confirmation of the cystomanamide and macyranone biosynthetic gene clusters are reported in the second part of both Chapters 2 and 3.

In order to study the biological activity of the isolated compounds, they were tested against a broad range of enzymes, bacteria, fungi, parasites and mammalian cell lines. Macyranone A was found to inhibit the eukaryotic 20S proteasome and to be active against the parasitic agents *Trypanosoma brucei rhodesiense* and *Leishmania donovani*. The results of these studies concerning the biological activity of macyranone A are illustrated in Chapter 3.

1.6 References

(1) Fleming, A.: On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. *Bull. World Health Organ.* **2001**, *79*, 780–90.

(2) Newman, D.J.; Cragg, G.M.: Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* **2012**, *75*, 311–35.

(3) World Health Organization: Antimicrobial Resistance. Global Report on Surveillance. *Antimicrob. Resist.* **2014**, p 254.

(4) Boucher, H.W.; Talbot, G.H.; et al.: Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **2009**, *48*, 1–12.

(5) Persidis, A.: Cancer multidrug resistance. *Nat. Biotechnol.* **1999**, *17*, 94–5.

(6) Schmitt, E.K.; Moore, C.M.; et al.: Natural products as catalysts for innovation: a pharmaceutical industry perspective. *Curr. Opin. Chem. Biol.* **2011**, *15*, 497–504.

(7) Reichenbach, H.: Die Biologie der Myxobakterien. *Biol. unserer Zeit* **1974**, *4*, 33–45.

(8) Nan, B.; Zusman, D.R.: Uncovering the mystery of gliding motility in the myxobacteria. *Annu. Rev. Genet.* **2011**, *45*, 21–39.

(9) Oxford, A.E.: Observations Concerning the Growth and Metabolic Activities of Myxococci in a Simple Protein-free Liquid Medium. *J. Bacteriol.* **1947**, *53*, 129–38.

(10) Plaza, A.; Müller, R.: Myxobacteria: Chemical Diversity and Screening Strategies. *Natural Products: Discourse, Diversity, and Design* **2014**, (Osbourn, A., Goss, R. J., and Carter, G. T., Eds.) John Wiley & Sons, Inc., Hoboken, NJ, USA,.

(11) Surup, F.; Viehrig, K.; et al.: Disciformycins A and B: 12-Membered Macrolide Glycoside Antibiotics from the Myxobacterium Pyxidicoccus fallax Active against Multiresistant Staphylococci. *Angew. Chem. Int. Ed. Engl.* **2014**, 1–5.

(12) Hoffmann, T.; Müller, S.; et al.: Microsclerodermins from terrestrial myxobacteria: an intriguing biosynthesis likely connected to a sponge symbiont. *J. Am. Chem. Soc.* **2013**, *135*, 16904–11.

(13) Etzbach, L.; Plaza, A.; et al.: Cystomanamides: structure and biosynthetic pathway of a family of glycosylated lipopeptides from myxobacteria. *Org. Lett.* **2014**, *16*, 2414–7.

(14) Raju, R.; Garcia, R.; Müller, R.: Angiolactone, a new Butyrolactone isolated from the terrestrial myxobacterium, Angiococcus sp. *J. Antibiot.* **2014,.**

(15) Okanya, P.W.; Mohr, K.I.; et al.: Hyafurones, hyapyrrolines, and hyapyrones: polyketides from Hyalangium minutum. *J. Nat. Prod.* **2014**, *77*, 1420–9.

(16) Nadmid, S.; Plaza, A.; et al.: Hyalachelins A-C, unusual siderophores isolated from the terrestrial myxobacterium Hyalangium minutum. *Org. Lett.* **2014**, *16*, 4130–3.

(17) Baumann, S.; Herrmann, J.; et al.: Cystobactamids: Myxobacterial Topoisomerase Inhibitors Exhibiting Potent Antibacterial Activity. *Angew. Chem. Int. Ed. Engl.* **2014**, 14605–14609.

(18) Weissman, K.J.; Müller, R.: Myxobacterial secondary metabolites: bioactivities and modes-of-action. *Nat. Prod. Rep.* **2010**, *27*, 1276–95.

(19) Reichenbach, H.; Höfle, G.: Epothilone, a Myxobacterial Metabolite with Promising Antitumor Activity. *Anticancer Agents from Natural Products* **2005**, (Kingston, D., Cragg, G., and Newman, D., Eds.) CRC Press.

(20) Hunt, J.T.: Discovery of ixabepilone. *Mol. Cancer Ther.* **2009**, *8*, 275–81.

(21) Konstantinidis, K.T.; Tiedje, J.M.: Trends between gene content and genome size in prokaryotic species with larger genomes. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 3160–5.

(22) Han, K.; Li, Z.; et al.: Extraordinary expansion of a Sorangium cellulosum genome from an alkaline milieu. *Sci. Rep.* **2013**, *3*, 2101.

(23) Walsh, C.T.; Chen, H.; et al.: Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr. Opin. Chem. Biol.* **2001**, *5*, 525–34.

(24) Baerga-Ortiz, A.; Popovic, B.; et al.: Directed Mutagenesis Alters the Stereochemistry of Catalysis by Isolated Ketoreductase Domains from the Erythromycin Polyketide Synthase. *Chem. Biol.* **2006**, *13*, 277–285.

(25) McDaniel, R.; Thamchaipenet, A.; et al.: Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 1846–1851.

(26) Kao, C.; Luo, G.; Katz, L.: Manipulation of macrolide ring size by directed mutagenesis of a modular polyketide synthase. *J. Am. Chem. Soc.***1995**, *117*, 9105–9106.

(27) Wenzel, S.C.; Müller, R.: Myxobacterial natural product assembly lines: fascinating examples of curious biochemistry. *Nat. Prod. Rep.* **2007**, *24*, 1211–24.

(28) Medema, M.H.; Blin, K.; et al.: antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **2011**, *39*, W339–46.

(29) Gaitatzis, N.; Silakowski, B.; et al.: The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. *J. Biol. Chem.* **2002**, *277*, 13082–90.

(30) Frank, B.; Knauber, J.; et al.: Spiroketal polyketide formation in Sorangium: identification and analysis of the biosynthetic gene cluster for the highly cytotoxic spirangienes. *Chem. Biol.* **2007**, *14*, 221–33.

(31) Julien, B.; Tian, Z.-Q.; et al.: Analysis of the ambruticin and jerangolid gene clusters of Sorangium cellulosum reveals unusual mechanisms of polyketide biosynthesis. *Chem. Biol.* **2006**, *13*, 1277–86.

(32) Ligon, J.; Hill, S.; et al.: Characterization of the biosynthetic gene cluster for the antifungal polyketide soraphen A from Sorangium cellulosum So ce26. *Gene* **2002**, *285*, 257–267.

(33) Staunton, J.; Weissman, K.J.: Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **2001**, *18*, 380–416.

(34) Wakil, S.J.: Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry* **1989**, *28*, 4523–4530.

(35) Hopwood, D.A.: Genetic Contributions to Understanding Polyketide Synthases. *Chem. Rev.* **1997**, *97*, 2465–2498.

(36) Hertweck, C.: The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 4688–716.

(37) Wenzel, S.C.; Williamson, R.M.; et al.: On the biosynthetic origin of methoxymalonyl-acyl carrier protein, the substrate for incorporation of "glycolate" units into ansamitocin and soraphen A. *J. Am. Chem. Soc.* **2006**, *128*, 14325–36.

(38) Wilson, M.C.; Nam, S.-J.; et al.: Structure and biosynthesis of the marine streptomycete ansamycin ansalactam A and its distinctive branched chain polyketide extender unit. *J. Am. Chem. Soc.* **2011**, *133*, 1971–7.

(39) Wu, K.; Chung, L.; et al.: The FK520 gene cluster of Streptomyces hygroscopicus var. ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene* **2000**, *251*, 81–90.

(40) Chan, Y.A.; Boyne, M.T.; et al.: Hydroxymalonyl-acyl carrier protein (ACP) and aminomalonyl-ACP are two additional type I polyketide synthase extender units. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 14349–54.

(41) Yadav, G.; Gokhale, R.S.; Mohanty, D.: Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *J. Mol. Biol.* **2003**, *328*, 335–63.

(42) Walsh, C.T.; Gehring, A.M.; et al.: Post-translational modification of polyketide and nonribosomal peptide synthases. *Curr. Opin. Chem. Biol.* **1997**, *1*, 309–15.

(43) Dutta, S.; Whicher, J.R.; et al.: Structure of a modular polyketide synthase. *Nature* **2014**, *510*, 512–7.

(44) Whicher, J.R.; Dutta, S.; et al.: Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature* **2014**, *510*, 560–4.

(45) Silakowski, B.; Schairer, H.U.; et al.: New Lessons for Combinatorial Biosynthesis from Myxobacteria: THE MYXOTHIAZOL BIOSYNTHETIC GENE CLUSTER OF Stigmatella aurantiaca DW4/3-1. *J. Biol. Chem.* **1999**, *274*, 37391–37399.

(46) Hu, Y.; Floss, H.G.: Further studies on the biosynthesis of the manumycin-type antibiotic, asukamycin, and the chemical synthesis of protoasukamycin. *J. Am. Chem. Soc.* **2004**, *126*, 3837–44.

(47) Moore, B.S.; Cho, H.; et al.: Biosynthetic studies on ansatrienin A. Formation of the cyclohexanecarboxylic acid moiety. *J. Am. Chem. Soc.* **1993**, *115*, 5254–5266.

(48) Thiericke, R.; Zeeck, A.; et al.: Biosynthesis of the manumycin group antibiotics. *J. Am. Chem. Soc.* **1990**, *112*, 3979–3987.

(49) Moore, B.S.; Hertweck, C.: Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat. Prod. Rep.* **2002**, *19*, 70–99.

(50) Gaitatzis, N.; Silakowski, B.; et al.: The biosynthesis of the aromatic myxobacterial electron transport inhibitor stigmatellin is directed by a novel type of modular polyketide synthase. *J. Biol. Chem.* **2002**, *277*, 13082–90.

(51) Mootz, H.D.; Marahiel, M.A.: The tyrocidine biosynthesis operon of Bacillus brevis: complete nucleotide sequence and biochemical characterization of functional internal adenylation domains. *J. Bacteriol.* **1997**, *179*, 6843–50.

(52) Wu, X.; Ballard, J.; Jiang, Y.W.: Structure and biosynthesis of the BT peptide antibiotic from Brevibacillus texasporus. *Appl. Environ. Microbiol.* **2005**, *71*, 8519–30.

(53) Sasse, F.; Steinmetz, H.; et al.: Argyrins, Immunosuppressive Cyclic Peptides from Myxobacteria. I. Production, Isolation, Physico-chemical and Biological Properties. *J. Antibiot.* **2002**, *55*, 543–551.

(54) Sheehan, J.C.; Mania, D.; et al.: The structure of telomycin. *J. Am. Chem. Soc.* **1968**, *90*, 462–70.

(55) Schwarzer, D.; Finking, R.; Marahiel, M. a.: Nonribosomal peptides: from genes to products. *Nat. Prod. Rep.* **2003**, *20*, 275.

(56) Irschik, H.; Gerth, K.; et al.: The myxovalargins, new peptide antibiotics from Myxococcus fulvus (Myxobacterales). I. Cultivation, isolation, and some chemical and biological properties. *J. Antibiot.* **1983**, *36*, 6–12.

(57) Schummer, D.; Höfle, G.; et al.: Antibiotics from Gliding Bacteria, LXXVI. Vioprolides: New Antifungal and Cytotoxic Peptolides from Cystobacter violaceus. *Liebigs Ann.* **2006**, *1996*, 971–978.

(58) Dieckmann, R.; Lee, Y.-O.; et al.: Expression of an active adenylate-forming domain of peptide synthetases corresponding to acyl-CoA-synthetases. *FEBS Lett.* **1995**, *357*, 212–216.

(59) Röttig, M.; Medema, M.H.; et al.: NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **2011**, *39*, W362–7.

(60) Marahiel, M.A.; Stachelhaus, T.; Mootz, H.D.: Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chem. Rev.* **1997**, *97*, 2651–2674.

(61) Rausch, C.; Hoof, I.; et al.: Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evol. Biol.* **2007**, *7*, 78.

(62) Schoenafinger, G.; Marahiel, M.A.: Natural Products in Chemical Biology. *Nat. Prod. Chem. Biol.* **2012**, (Civjan, N., Ed.) John Wiley & Sons, Inc., Hoboken, NJ, USA,.

(63) Khosla, C.; Harbury, P.B.: Modular enzymes. *Nature* **2001**, *409*, 247–52.

(64) Weissman, K.J.; Müller, R.: A brief tour of myxobacterial secondary metabolism. *Bioorg. Med. Chem.* **2009**, *17*, 2121–36.

(65) Höfle, G.; Bedorf, N.; et al.: Epothilon A und B – neuartige, 16gliedrige Makrolide mit cytotoxischer Wirkung: Isolierung, Struktur im Kristall und Konformation in Lösung. *Angew. Chemie* **1996**, *108*, 1671–1673.

(66) Plaza, A.; Garcia, R.; et al.: Aetheramides A and B, Potent HIV-Inhibitory Depsipeptides from a Myxobacterium of the New Genus "Aetherobacter." *Org. Lett.* **2012**, *14*, 2010–2013.

(67) Garcia, R.; Gerth, K.; et al.: Expanded phylogeny of myxobacteria and evidence for cultivation of the "unculturables." *Mol. Phylogenet. Evol.* **2010**, *57*, 878–887.

(68) Garcia, R.O.; Krug, D.; Müller, R.: Chapter 3. Discovering natural products from myxobacteria with emphasis on rare producer strains in combination with improved analytical methods. *Methods Enzymol.* **2009**, *458*, 59–91.

(69) Gerth, K.; Irschik, H.; et al.: Myxothiazol, an antibiotic from myxococcus fulvus. (myxobacterales) I. cultivation, isolation, physico-chemical and biological properties. *J. Antibiot.* **1980**, *33*, 1474–1479.

(70) Zanella, F.; Lorens, J.B.; Link, W.: High content screening: seeing is believing. *Trends Biotechnol.* **2010**, *28*, 237–45.

(71) Martinez, J.P.; Hinkelmann, B.; et al.: Identification of myxobacteria-derived HIV inhibitors by a high-throughput two-step infectivity assay. *Microb. Cell Fact.* **2013**, *12*, 85.

(72) Fleischmann, R.; Adams, M.; et al.: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science (80-.).* **1995**, *269*, 496–512.

(73) Bode, H.B.; Müller, R.: The impact of bacterial genomics on natural product research. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 6828–46.

(74) Wenzel, S.C.; Müller, R.: The impact of genomics on the exploitation of the myxobacterial secondary metabolome. *Nat. Prod. Rep.* **2009**, *26*, 1385–407.

(75) Bode, H.B.; Bethe, B.; et al.: Big effects from small changes: possible ways to explore nature's chemical diversity. *ChemBioChem* **2002**, *3*, 619–27.

(76) Scherlach, K.; Hertweck, C.: Triggering cryptic natural product biosynthesis in microorganisms. *Org. Biomol. Chem.* **2009**, *7*, 1753–60.

(77) Bergmann, S.; Schümann, J.; et al.: Genomics-driven discovery of PKS-NRPS hybrid metabolites from Aspergillus nidulans. *Nat. Chem. Biol.* **2007**, *3*, 213–7.

(78) Laureti, L.; Song, L.; et al.: Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in Streptomyces ambofaciens. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 6258–63.

(79) Yamanaka, K.; Reynolds, K. a; et al.: Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 1957–62.

(80) Wang, J.; Soisson, S.M.; et al.: Platensimycin is a selective FabF inhibitor with potent antibiotic properties. *Nature* **2006**, *441*, 358–361.

(81) Nickeleit, I.; Zender, S.; et al.: Argyrin A Reveals a Critical Role for the Tumor Suppressor Protein p27kip1 in Mediating Antitumor Activities in Response to Proteasome Inhibition. *Cancer Cell* **2008**, *14*, 23–35.

(82) Watanabe, N.; Niki, E.: Direct-Coupling of FT-NMR to High Performance Liquid Chromatography. *Proc. Japan Acad. Ser. B Phys. Biol. Sci.* **1978**, *54*, 194–199.

(83) Exarchou, V.; Krucker, M.; et al.: LC-NMR coupling technology: recent advancements and applications in natural products analysis. *Magn. Reson. Chem.* **2005**, *43*, 681–7.

(84) Schlotterbeck, G.; Ceccarelli, S.M.: LC-SPE-NMR-MS: a total analysis system for bioanalysis. *Bioanalysis* **2009**, *1*, 549–59.

(85) Smallcombe, S.H.; Patt, S.L.; Keifer, P.A.: WET Solvent Suppression and Its Applications to LC NMR and High-Resolution NMR Spectroscopy. *J. Magn. Reson. Ser. A* **1995**, *117*, 295–303.

(86) Liu, M.; Mao, X.; et al.: Improved WATERGATE Pulse Sequences for Solvent Suppression in NMR Spectroscopy. *J. Magn. Reson.* **1998**, *132*, 125–129.

(87) Spraul, M.; Freund, A.S.; et al.: Advancing NMR sensitivity for LC-NMR-MS using a cryoflow probe: application to the analysis of acetaminophen metabolites in urine. *Anal. Chem.* **2003**, *75*, 1536–41.

(88) Nicolaou, K.C.; Snyder, S. a: Chasing molecules that were never there: misassigned natural products and the role of chemical synthesis in modern structure elucidation. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 1012–44.

(89) Kind, T.; Fiehn, O.: Advances in structure elucidation of small molecules using mass spectrometry. *Bioanal. Rev.* **2010**, *2*, 23–60.

(90) Rabi, I.; Zacharias, J.; et al.: A New Method of Measuring Nuclear Magnetic Moment. *Phys. Rev.* **1938**, *53*, 318–318.

(91) Holzgrabe, U.; Diehl, B.W.K.; Wawer, I.: NMR spectroscopy in pharmacy. *J. Pharm. Biomed. Anal.* **1998**, *17*, 557–616.

(92) Ernst, R.R.: Application of Fourier Transform Spectroscopy to Magnetic Resonance. *Rev. Sci. Instrum.* **1966**, *37*, 93.

(93) Rabenstein, D.L.; Nakashima, T.T.: Spin-Echo Fourier Transform Nuclear Magnetic Resonance Spectroscopy. *Anal. Chem.* **1979**, *51*, 1465A–1474A.

(94) Bross-Walch, N.; Kühn, T.; et al.: Strategies and tools for structure determination of natural products using modern methods of NMR spectroscopy. *Chem. Biodivers.* **2005**, *2*, 147–77.

(95) Farrar, T.C.; Becker, E.D.: Pulse and Fourier Transform NMR. Introduction to Theory and Methods. **1971**, Academic Press, New York, p 118.

(96) Aue, W.P.: Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *J. Chem. Phys.* **1976**, *64*, 2229.

(97) Breton, R.C.; Reynolds, W.F.: Using NMR to identify and characterize natural products. *Nat. Prod. Rep.* **2013**, *30*, 501–24.

(98) Webb, A.: Increasing the sensitivity of magnetic resonance spectroscopy and imaging. *Anal. Chem.* **2012**, *84*, 9–16.

(99) Friebolin, H.: Basic One- and Two-Dimensional NMR Spectroscopy. **2011**, Fidth Edit. Wiley-VCH, Weinheim, p 418.

(100) Minch, M.J.: Orientational dependence of vicinal proton-proton NMR coupling constants: The Karplus relationship. *Concepts Magn. Reson.* **1994**, *6*, 41–56.

(101) Claridge, T.D.: High-Resolution NMR Techniques in Organic Chemistry. **2009**, Second edi. Elsevir, Oxford, UK, p 383.

(102) Patt, S.L.; Shoolery, J.N.: Attached proton test for carbon-13 NMR. *J. Magn. Reson.* **1982**, *46*, 535–539.

(103) Morris, G.A.; Freeman, R.: Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.* **1979**, *101*, 760–762.

(104) Doddrell, D.M.; Pegg, D.T.; Bendall, M.R.: Distortionless enhancement of NMR signals by polarization transfer. *J. Magn. Reson.* **1982**, *48*, 323–327.

(105) Molinski, T.F.; Morinaka, B.I.: Integrated Approaches To the Configurational Assignment of Marine Natural Products. *Tetrahedron* **2012**, *68*, 9307–9343.

(106) Bax, A.; Freeman, R.: Investigation of complex networks of spin-spin coupling by two-dimensional NMR. *J. Magn. Reson.* **1981**, *44*, 542–561.

(107) Bax, A.; Davis, D.G.: MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson.* **1985**, *65*, 355–360.

(108) Kover, K.E.; Prakash, O.; Hruby, V.J.: z-Filtered Heteronuclear Coupled-HSQC-TOCSY Experiment as a Means for Measuring Long-Range Heteronuclear Coupling Constants. *J. Magn. Reson. Ser. A* **1993**, *103*, 92–96.

(109) Bax, A.; Morris, G. a: An improved method for heteronuclear chemical shift correlation by two-dimensional NMR. *J. Magn. Reson.* **1981**, *42*, 501–505.

(110) Bodenhausen, G.; Ruben, D.J.: Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **1980**, *69*, 185–189.

(111) Bax, A.; Subramanian, S.: Sensitivity-enhanced two-dimensional heteronuclear shift correlation NMR spectroscopy. *J. Magn. Reson.* **1986**, *67*, 565–569.

(112) Bax, A.; Summers, M.F.: Proton and carbon-13 assignments from sensitivity-enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR. *J. Am. Chem. Soc.* **1986**, *108*, 2093–2094.

(113) Macura, S.; Huang, Y.; et al.: Two-dimensional chemical exchange and cross-relaxation spectroscopy of coupled nuclear spins. *J. Magn. Reson.* **1981**, *43*, 259–281.

(114) Bax, A.; Davis, D.G.: Practical aspects of two-dimensional transverse NOE spectroscopy. *J. Magn. Reson.* **1985**, *63*, 207–213.

(115) Matsumori, N.; Kaneno, D.; et al.: Stereochemical Determination of Acyclic Structures Based on Carbon-Proton Spin-Coupling Constants. A Method of Configuration Analysis for Natural Products. *J. Org. Chem.* **1999**, *64*, 866–876.

(116) https://scifinder.cas.org

(117) http://pubs.rsc.org/marinlit

(118) Horai, H.; Arita, M.; et al.: MassBank: A public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **2010**, *45*, 703–714.

(119) Krug, D.; Müller, R.: Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. *Nat. Prod. Rep.* **2014**, *31*, 768–83.

(120) Bobzin, S.C.; Yang, S.; Kasten, T.P.: LC-NMR: a new tool to expedite the dereplication and identification of natural products. *J. Ind. Microbiol. Biotechnol.* **2000**, *25*, 342–345.

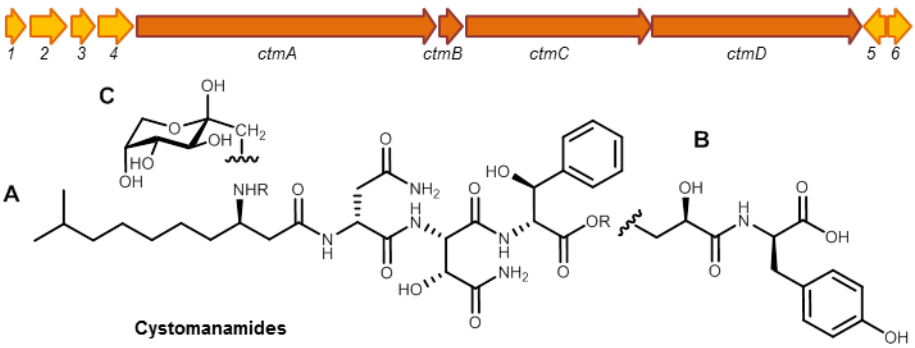
Chapter 2

Cystomanamides: Structure and Biosynthetic Pathway of a Family of Glycosylated Lipopeptides from Myxobacteria

Lena Etzbach, Alberto Plaza, Ronald Garcia, Sascha Baumann, Rolf Müller*

Department of Microbial Natural Products, Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research (HZI) and Pharmaceutical Biotechnology, Saarland University, Campus C2 3, 66123 Saarbrücken, Germany

German Center for Infection Research (DZIF), Partner site Hannover-Braunschweig 38124, Germany



Org. Lett., 2014, 16 (9), pp 2414–2417

DOI: 10.1021/ol500779s

Publication Date (Web): April 15, 2014

2 Cystomanamides

2.1 Abstract

Cystomanamides A-D were isolated as novel natural product scaffolds from *Cystobacter fuscus* MCy9118 and their structures were established by spectroscopic techniques including 2D NMR, LC-SPE-NMR/-MS and HR-MS. The cystomanamides contain β-hydroxy amino acids along with 3-amino-9-methyldecanoic acid that is N-glycosylated in cystomanamide C and D. The gene cluster for cystomanamide biosynthesis was identified by gene disruption as PKS/NRPS hybrid incorporating an iso-fatty acid as starter unit and including a reductive amination step at the interface of the PKS and NRPS modules.

2.2 Main Text

Myxobacteria have proven to exhibit a fascinating capacity to produce chemically intriguing natural products which often show unique structural elements rarely produced by other sources.[1] Myxobacterial secondary metabolites belong to multiple structural classes and many of these diverse compounds originate from mixed polyketide-nonribosomal peptide biosynthetic pathways.[2] Complex multimodular enzymes involving various catalytic and structural domains are responsible for the biosynthesis of these versatile structures.[3] The ability to generate a wide range of complex natural products is expanded by enzymes introducing β-branching and post-PKS and NRPS reactions such as hydroxylation, glycosylation and epimerization.[4]

Our discovery strategy to determine natural products with novel structural frameworks includes UHPLC/HRMS-based metabolomics for strain selection and dereplication as well as hyphenated chromatographic methods such as LC-SPE-NMR/-MS for isolation driven by chemical and structural features. This approach indicated the presence of unusual peptides in an extract of *Cystobacter fuscus* MCy9118 and finally led to the isolation of a new family of lipopeptides, the cystomanamides (ctm).

These new linear peptides comprise exclusively non-proteinogenic amino acids and bear an unusual 3-amino-9-methyldecanoic acid residue at the N-terminus. Cystomanamide C and D contain an *N*-linked glycosylation, which is one of the rare examples of late-stage modification reactions of myxobacterial metabolites.[5] In addition to these four new

myxobacterial natural products, the strain was also found to produce the known antibiotics althiomycin[6,7], roimatacene[8] and myxochelin A[9] and B[10].

HRESIMS of cystomanamide A **(1)** displayed an $[M+H]^+$ peak at *m/z* 609.3248 (calcd for $C_{28}H_{45}N_6O_9$, 609.3243), consistent with the molecular formula $C_{28}H_{44}N_6O_9$ containing 10 double-bond equivalents (DBE). The ^{1}H NMR spectrum of cystomanamide B in CD_3OD exhibited signals characteristic of a peptide including three α-proton signals at δ 4.77 (1H, dd, *J* = 7.4, 6.3 Hz), 4.80 (1H, d, *J* = 3.75 Hz) and 4.49 (1H, d, *J* = 3.25 Hz). Moreover, a downfield pair of triplets at δ 7.28 (2H, t, *J* = 7.45 Hz) and 7.20 (1H, t, *J* = 7.45 Hz) and a doublet at δ 7.34 (2H, d, *J* = 7.45 Hz) were observed (see Table S4). The HSQC spectrum revealed the presence of several methylenes between δ_H 1.2 and 2.9 and two oxygenated methines at δ_C 72.8 / δ_H 4.29 (β-OH-Asn) and δ_C 74.9 / δ_H 5.22 (β-OH-Phe).

A detailed analysis of the 2D NMR data obtained from HSQC, HMBC, COSY and TOCSY experiments indicated the presence of asparagine, β-hydroxy asparagine (β-OH-Asn) and β-hydroxy phenylalanine (β-OH-Phe). Additionally, a 3-amino-9-methyldecanoic acid residue (AMDA) was identified and its partial structure was deduced as follows. A sequential spin system starting from two methyl groups at δ 0.89 (Me-10 and Me-11$_{AMDA}$) attached to a methine at δ 1.55 (H-9$_{AMDA}$) and comprising five sequential methylene groups H_2-8 to H_2-4 was obtained from COSY and TOCSY correlations. COSY correlations extended this fragment by an additional aminomethine group at δ 3.54 (H-3$_{AMDA}$) followed by a methylene at δ 2.68, 2.50 (H_2-2$_{AMDA}$). Finally, HMBC correlations from H-3$_{AMDA}$ and H_2-2$_{AMDA}$ to a carbonyl resonance at δ 173.2 (C-1$_{AMDA}$) clearly indicated that AMDA represents a β-amino acid. The sequence was established through HMBC correlations from α-protons to carbonyl carbons of adjacent residues to result in the sequence: AMDA-Asn-(β-OH-Asn)-(β-OH-Phe). Long-range HBMC correlations from the α-proton at δ 4.77 (H-2$_{Asn}$) to C-1$_{AMDA}$ connected AMDA to the N-terminus of asparagine, thereby completing the structure of **1** as linear tetrapeptide.

HRESIMS of cystomanamide B (**2**) displayed an $[M+H]^+$ peak at *m/z* 860.4039 (calcd for $C_{40}H_{58}N_7O_{14}$, 860.4036), consistent with the molecular formula $C_{40}H_{57}N_7O_{14}$ containing 16 DBE. The ^{1}H NMR spectra of **2** in comparison to that of **1** exhibited additional signals for one α-proton at δ 4.43 (1H, dd, *J* = 7.5, 5.5 Hz) and two downfield doublets at δ 7.05 (2H, d, *J* = 8.4 Hz) and δ 6.68 (2H, d, *J* = 8.4 Hz). The HSQC spectrum revealed the additional presence of one methylene at δ_C 38.2 / δ_H 3.12, 2.97 (Tyr), one oxygenated methylene at δ_C 67.9 / δ_H 4.32, 4.16 (glyceric acid GA) and one oxygenated methine at δ_C 71.0 / δ_H 4.24

(GA). This evidence in combination with the DQF-COSY and HMBC correlations indicated that in comparison to **1**, cystomanamide B bears an additional glyceric acid (GA) and tyrosine residue. HMBC long range correlations from α-protons to carbonyl carbons of adjacent residues and from the methylene protons H_2-3_{GA} to C-$1_{\square\text{-OH-Phe}}$ resulted in the linear depsipeptide **2** with the sequence AMDA-Asn-(β-OH-Asn)-(β-OH-Phe)-GA-Tyr.

The ESI-MS/MS fragmentation patterns of cystomanamide C **(3)** and D **(4)** indicated a glycosylation. More precisely, **3** showed the loss of a 162 mass units fragment compared to **1** suggesting the presence of a hexose residue. The ^{1}H NMR and HSQC spectra in comparison with those of **1** showed various additional signals between 3.1 and 4.2 ppm with different intensities belonging to three rotamers of a sugar moiety. Based on HMBC and ROESY NMR data, the sugar residue was identified as fructose with β-D-fructopyranose as the most abundant rotamer, a sugar moiety also found in kwansonine A and B.[11] The HMBC spectrum showed a key correlation from the methylene at δ 3.27 (H-1_{FRU}) to the aminomethine at δ 3.59 (H-3_{AMDA}) indicating an *N*-linked glycosylation of the β-amino fatty acid chain. Brabantamides are another example for compounds showing a sugar attached to a fatty acid derived moiety.[12]

Cystomanamide D showed similar NMR and MS data with a mass difference of 16 Da compared to **3** that was explained by the absence of the β-hydroxylation of the phenylalanine residue.

The absolute configuration of the amino acid residues was elucidated by MS detected chromatographic analysis of the L- and D-FDLA (1-fluoro-2,4-dinitro-phenyl-5-L/D-leucinamide) derivatives of the acid hydrolyzate of cystomanamide A, B and C and comparison with respective standards.[13] The amino acid residues were assigned as D-asparagine, L-*erythro*-β-OH-asparagine, D-*threo*-β-OH-phenylalanine and D-tyrosine. The *R* configuration of C-3 of the AMDA residue was also determined by this method. β-Amino fatty acids derivatized with Marfey's reagent show a behavior analogue to α-amino acids where the L series elutes earlier than the D series.[14] The D configurations of glyceric acid in **2** and the fructose in **3** were both established by chiral HPLC of the acid hydrolyzate and comparison with respective standards.

Based on the chemical structures of the cystomanamides (see Figure 2.1), it seemed likely that these compounds are products of a PKS/NRPS hybrid megasynthetase. Genome sequence data of *C. fuscus* MCy9118 were generated using Illumina sequencing. A

retrobiosynthetic approach in combination with antiSMASH 2.0[15] analysis of the draft genome sequence led to the identification of the *ctm* biosynthetic gene cluster. The predicted gene cluster consists of 10 open reading frames (ORFs) and has an overall GC content of 69.8%.

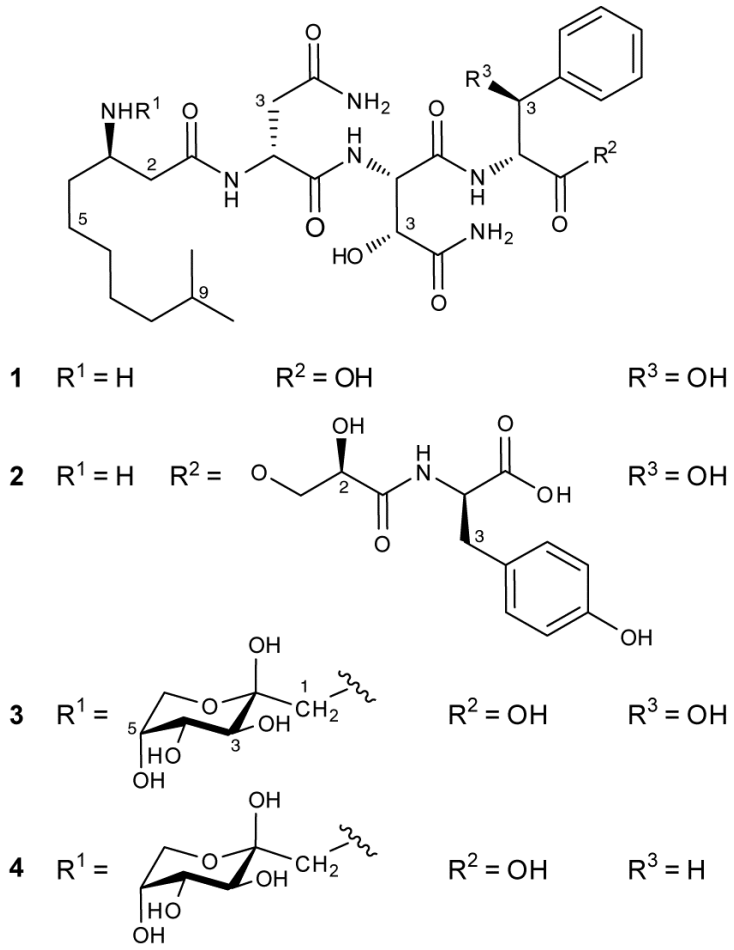


Figure 2.1: Structures of cystomanamide A (1), B (2), C (3), D (4)

To further analyze the catalytic domains and the A domain substrate specificity, the open reading frames were translated and analyzed using Pfam[16], NRPS predictor2[17], and PKS/NRPS Analysis[18]. This analysis led to the prediction of one loading module, six elongation modules and one termination module. Genes not resulting in predictions related to PKS or NRPS domains were analyzed via the BLAST algorithm[19] using the non-redundant sequence database at the National Center for Biotechnology Information (NCBI). *CtmA* encodes for a complex protein that contains functional domains belonging to fatty acid synthases, polyketide synthases, amino transferases and non-ribosomal peptide synthetases. The gene shows similarity to *mycA* encoding for mycosubtilin

synthase subunit A.[20] Feeding experiments with L-[methyl-$^2H_3$]leucine indicated a leucine-derived branched chain carboxylic acid starter unit.[21] The assembly line starts with a CoA ligase domain responsible for recognition and activation of the starter molecule 9-methyldecanoic acid, an iso-odd fatty acid (see Figure 2.2). The activated substrate is then transferred to the first acyl carrier protein in the loading module.[22] It undergoes one elongation step using malonate as a substrate to form the -ketothioester. The amino transferase domain (AMT) located in module 1 at the interface of the PKS and NRPS modules next reductively aminates the β-ketothioester which is then passed on to the NRPS module as shown for mycosubtilin biosynthesis.[23] This AMT shows significant (57 %) similarity to the AMT found in MycA. It was proven *in vitro* that the AMT catalyzes amine transfer from an amino acid to a protein-bound β-ketothioester to generate the corresponding protein-bound β-aminothioester dependent on pyridoxal 5′-phosphate (PLP).[23] Biosynthesis continues with three NRPS based reaction cycles. *In silico* analysis of the A domain specificities is consistent with the incorporated amino acids. Module 2, 4 and 6 contain epimerization domains that are responsible for the transformation of the L- into the respective D-amino acid. The incorporation of D-Asn, D-Phe and D-Tyr are in accordance with the structure and the absolute configuration of **2**. The second NRPS module, module 3, is split into two proteins, CtmA and CtmC, which can be regarded as unusual but is not unprecedented.[22] The condensation domain is encoded by *ctmA* whereas the adenylation domain and the peptidyl carrier protein are encoded by *ctmC*. The module bears an additional condensation domain encoded by *ctmC*, which seems to be inactive due to the missing catalytic histidine residues in the active site.[24] Genetically, the module is separated by *ctmB* whose product is homologous to TauD from *Streptomyces auratus* (44.9%), a well-studied non-heme iron hydroxylase. CtmB contains the conserved 2-His-1-carboxylate facial triad responsible for iron binding[25,26] as well as the conserved Arg residue that ligates β-ketoglutarate.[27] It also shows similarity (35.1%) to SyrP from *Pseudomonas syringae*, which is responsible for the β-hydroxylation of an aspartyl residue in syringomycin E biosynthesis.[28] This indicates that CtmB is most likely responsible for the hydroxylation of the second asparagine residue to form L-*erythro*-β-OH-asparagine. Module 5 has a similar organization as NRPS modules but instead of an adenylation domain it contains an unusual domain that exhibits the three conserved motifs specific for the HAD superfamily.[29]

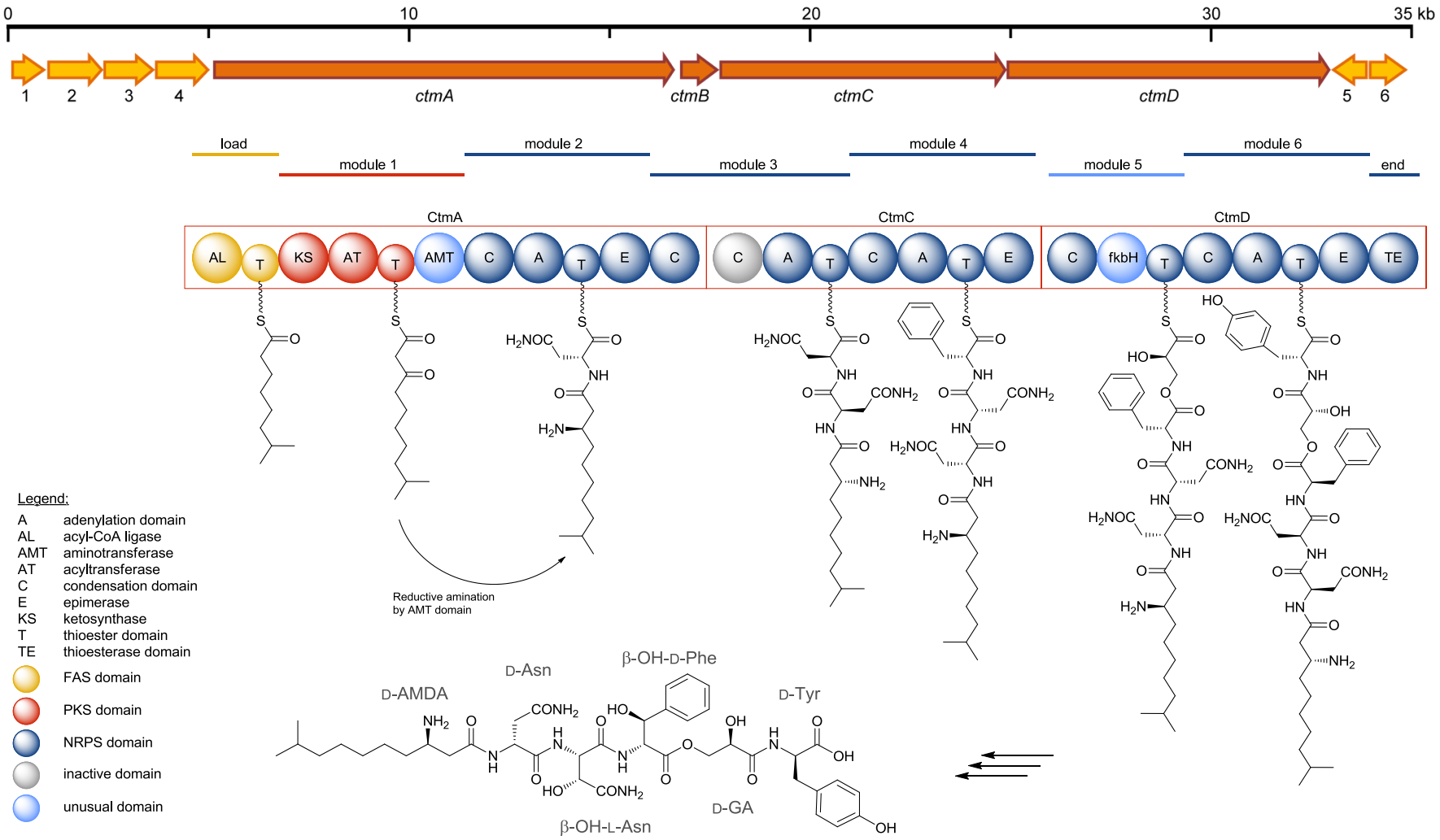


Figure 2.2: Gene cluster of cystomanamides in *Cystobacter fuscus* MCy9118 and the biosynthetic pathway of **2**.

BLAST results show similarity to FkbH[30] from *Streptomyces hygroscopicus* subsp. *ascomyceticus* (29.5 %) and OzmB[31] from *Streptomyces albus* (30.6 %) which are responsible for the formation of glyceryl-ACP in the biosynthesis of the polyketide natural products FK520 and oxazolomycin. For OzmB it was demonstrated that it acts as a bifunctional glyceryl transferase/phosphatase that first binds D-1,3-bisphosphoglycerate from the glycolytic pool to form the D-3-phosphoglyceryl-S-OzmB intermediate. In the next step it removes the phosphate group to receive the D-3-glyceryl-S-OzmB species (acting as a phosphatase), and finally it acts as glyceryl transferase by transferring the glyceryl group to an acyl carrier protein (ACP).[31] To the best of our knowledge, CtmD is the first protein containing such a domain integrated into an NRPS module underlining the enormous potential of these modular megaenzymes for combinatorial biosynthesis.

Although the condensation domain of module 5 is homologous to the typical amide-forming condensation domains, it most likely catalyzes the formation of an ester bond instead of a peptide bond. The biochemical evidence for condensation domains being able to catalyze ester bond formation was given in studies concerning the mycotoxins fumonisins and the antitumor antibiotic C-1027.[32,33] The same mechanism with C domains embedded in elongation modules that are responsible for chain extension via ester bond formation has been proposed for the biosynthesis of some other nonribosomal peptides.[34-36]

The last step in the biosynthesis of the cystomanamides is the incorporation and epimerization of tyrosine in module 6. The assembly line is terminated by a thioesterase domain that releases the product to receive the linear PKS/NRPS product **2**. The glycosylation of **3** and **4** requires enzymes encoded outside the aglycon cluster which could not be identified as 41 genes were annotated as glycosyltransferase in the genome of *Cystobacter fuscus* MCy9118.

To further confirm that the candidate gene cluster is responsible for biosynthesis of the cystomanamides, the PKS/NRPS gene *ctmA* was inactivated using a single crossover homologous recombination knockout strategy. The resulting *ctmA* mutant lost the ability to produce **1** as well as its analogues (see Figure S4 and S5), therefore validating that the proposed gene cluster is indeed responsible for cystomanamide biosynthesis.

To explore the relation of the surrounding open reading frames in the biosynthetic pathway, we inactivated *orf2, orf4* and *orf5*. Cultivation of the resulting *orf2* and *orf4*

mutants and subsequent HPLC analysis revealed that inactivation of both genes abolished production of **1** and its analogues (see Figure S4 and S5). The product of *orf1* shows similarity to JmjC domain containing proteins that are connected to transcription factors[37], whereas the products of *orf2*, *orf 3* and *orf4* show similarity to the very heterogeneous group of β-lactamases. *Orf1* to *orf4* are most likely involved in the regulation of gene expression. At the same time, HPLC analysis of a cultivation of *orf5* mutants showed a cystomanamide production comparable to the wild type strain, indicating no or only a minor role of *orf5* in the biosynthesis.

The compounds were tested in various bioactivity assays including cytotoxicity against HCT-116 and CHO-K1 cells, antibacterial tests against various gram negative and gram positive bacterial strains, antifungal assays against *Candida albicans* and *Mucor hiemalis* and HIV-1 inhibition. Up to date, they have not shown biological activity. We continue functional testing to find the often very specific biological activity of these natural products and to further evaluate their biological function.

In summary, we discovered a new family of glycosylated lipopeptides using a structure-guided approach by LC-SPE-NMR. The compounds were fully characterized and a gene cluster responsible for cystomanamide biosynthesis was identified. Inactivation of three independent genes in this cluster completely abolished cystomanamide production in the mutants, verifying their essential role during biosynthesis.

2.3 Supporting Information

2.3.1 General experimental procedures

Optical rotations were measured with a Jasco polarimeter, IR spectra were recorded on a Perkin-Elmer FT-IR Spectrum One spectrometer. NMR spectra were recorded in methanol-d_4 on a Bruker Ascend 700 spectrometer with a 5 mm TXI cryoprobe (^{1}H at 700 MHz, ^{13}C at 175 MHz). DQF-COSY, HOHAHA, HSQC, HMBC, and ROESY experiments were recorded using standard pulse programs. HSQC experiments were optimized for $^1J_{C\text{-}H}$ = 145 Hz, and HMBC spectra were optimized for $^{2,3}J_{C\text{-}H}$ = 6 Hz. LC-HRMS data was performed on a Dionex Ultimate 3000 RSLC system using a Waters BEHC18, 100 x 2.1 mm, 1.7 µm dp column. Separation of 2 µl sample was achieved by a linear gradient with (A) H_2O + 0.1 % FA to (B) ACN + 0.1 % FA at a flow rate of 600 µl/min and 45 °C. The gradient was initiated by a 0.5 min isocratic step at 5 % B, followed by an increase to 95%

B in 18 min to end up with a 2 min step at 95 % B before reequilibration with initial conditions. UV spectra were recorded by a DAD in the range from 200 to 600 nm. The LC flow was split to 75 µl/min before entering the maXis HR-ToF mass spectrometer (BrukerDaltonics, Bremen, Germany) using the standard ESI source. Mass spectra were acquired in centroid mode ranging from 150 - 2000 *m/z* at 2 Hz scan speed.

2.3.2 Cultivation of strain *Cystobacter fuscus* MCy9118

Myxobacterial strain MCy9118 was isolated in 2007 from a Philippine soil sample using rabbit dung as bait. From the appearance of fruiting bodies on dung, the strain was purified by standard myxobacterial method of repeated subcultivations of the swarm colony on agar. Morphological characterizations of the strain combined with molecular phylogenetics led to the identification of strain MCy9118 to *Cystobacter fuscus*. To transfer the strain from solid medium to liquid medium and to grow the mutants, actively growing cells were excised from agar plates, inoculated into a 50 mL buffled flask containing 20 mL M medium (1 % phytone, 1 % maltose x H_2O, 0.1 % $CaCl_2$, 0.1 % $MgSO_4$ x 7 H_2O, 25 mM HEPES, 8 mg/L Fe-EDTA adjusted to pH 7.2 with 10 N KOH) and incubated at 30°C for 2 days. For fast analysis of the secondary metabolite profile, 2 % (w/v) XAD adsorber resin (Amberlite XAD-7, Sigma) were added on day three and harvested after 4 h by centrifugation. The XAD/cell mixture was extracted with 5 mL methanol and analyzed by HPLC-MS.

For best production of the cystomanamides, the preculture was used to inoculate two 1 L flasks containing 100 mL VY/2S (0.5 % Baker's yeast, 0.5 % $CaCl_2$ x 2 H_2O, 1.0 % soluble starch, 5 mM HEPES adjusted to pH 7.0 with 10 N KOH) medium that were incubated at 30°C for 8 h. 300 mL VY/2S medium were added into each flask and incubated at 30°C overnight. The cultures were used to inoculate a 10 L fermentor containing 8 L VY/2S medium (pO_2 = 20 %; stirring rate 600rpm; 30°C; pH 7.0). 2 % (w/v) XAD adsorber resin (Amberlite XAD-7, Sigma) was added on day 2 of the fermentation and further incubated for 8 days. Cultures were harvested by sieving the culture to receive the XAD and subsequent centrifugation at 4000 rpm for 10 min at 20°C to receive the cells. The supernatant was discarded.

2.3.3 Isolation

Cells and XAD from the 8 L culture were freeze dried overnight and extracted separately. To remove impurities, 2 x 200 mL hexane and 3 x 300 mL EtOAc were used. Extraction was accomplished with 3 x 300 mL MeOH to obtain a cell and an XAD extract with 2.0 g and 1.5 g, respectively. Cystomanamide A, C, and D were extracted from the XAD crude extract starting with two successive runs of sephadex fractionation. The sephadex fractionation was performed using a GE Healthcare SR 25/100 column and a TAC15/750G0-SR.2 column, sephadex LH 20 was used for separation and methanol as mobile phase. Subsequently, the fractions containing the peptides were purified by semipreparative HPLC (Phenomenex Synergi™ 4 µm Fusion-RP 80 Å, LC Column 250 x 10 mm, DAD at 220 nm) eluting with a linear gradient of 10-100% MeOH/H_2O + 0.1% FA in 50 min to afford compounds **1** (0.10 mg/L, t_R = 27.1 min), 3 (0.09 mg/L, t_R = 25.9 min) and 4 (0.02 mg/L, t_R = 28.1 min).

2 was extracted from the cell crude extract using sephadex fractionation (GE Healthcare SR 25/100 column, sephadex LH 20 with methanol as mobile phase) and subsequent purification by semipreparative HPLC (Phenomenex Synergi™ 4 µm Fusion-RP 80 Å, LC Column 250 x 10 mm, DAD at 220 nm) eluting with a linear gradient of 30-40% MeOH/H_2O + 0.1% FA in 5 min followed by 40-80% MeOH/H_2O + 0.1% FA in 40 min to afford compound 2 (0.04 mg/L, t_R = 20.4 min).

Cystomanamide A (**1**): Colorless amorphous powder; $[\alpha]^{20}_D$ + 18.4 (c 0.21, MeOH); IR (film) ν_{max} 3299, 2960, 2931, 2866, 1664, 1593, 1393 cm^{-1}; ^1H and ^{13}C NMR data, see Table S4; HRESIMS *m/z* 609.3248 $[M+H]^+$ corresponding to a molecular formula $C_{28}H_{44}N_6O_9$ (calcd for $C_{28}H_{45}N_6O_9$, 609.3243, Δ = 0.82 ppm)

Cystomanamide B (**2**): Colorless amorphous powder; $[\alpha]^{20}_D$ + 13.2 (c 0.09, MeOH); IR (film) ν_{max} 3287, 2946, 2325, 1676, 1593, 1389, 1350, 1032 cm^{-1}; ^1H and ^{13}C NMR data, see Table S5; HRESIMS *m/z* 860.4039 $[M+H]^+$ corresponding to a molecular formula $C_{40}H_{57}N_7O_{14}$ (calcd for $C_{40}H_{58}N_7O_{14}$, 860.4036, Δ = 0.35 ppm)

Cystomanamide C (**3**): Colorless amorphous powder; $[\alpha]^{20}_D$ + 19.2 (c 0.07, MeOH); IR (film) ν_{max} 3334, 2949, 1669, 1596, 1397, 1085, 1027 cm^{-1}; ^1H and ^{13}C NMR data, see Table S6; HRESIMS *m/z* 771.3786 $[M+H]^+$ corresponding to a molecular formula $C_{34}H_{54}N_6O_{14}$ (calcd for $C_{34}H_{55}N_6O_{14}$, 771.3771, Δ = 1.94 ppm)

Cystomanamide D (**4**): Colorless amorphous powder; ^{1}H and ^{13}C NMR data, see Table S7; HRESIMS *m/z* 755.3829 $[M+H]^+$ corresponding to a molecular formula $C_{34}H_{54}N_6O_{13}$ (calcd for $C_{34}H_{55}N_6O_{13}$, 755.3822, Δ = 0.93 ppm)

2.3.4 Absolute configuration of glyceric acid

An aliquot (0.2 mg) of compound 2 was hydrolyzed with 6 N HCl (0.3 mL) for 16 h at 90 °C. The hydrolyzate was concentrated to dryness and subjected to chiral HPLC analysis (Phenomenex Chirex 3126 (d)-penicillamine (150 x 4.6 mm) HPLC column; flow rate 1 mL/min; UV detection at 254 nm; solvent 2 mM $CuSO_4$:MeOH 85:15). The retention time of glyceric acid from the hydrolyzate with 18.7 min was compared to authentic standards whose retention times were 14.2 min for L-glyceric acid and 18.4 min for D-glyceric acid.

2.3.5 Absolute configuration of fructose

An aliquot (0.4 mg) of compound 3 was hydrolyzed with 1.25 M HCl in methanol (0.5 mL) for 16 h at 35 °C. The hydrolyzate was neutralized using 1 $NaHCO_3$, concentrated to dryness and subjected to chiral HPLC analysis (Chiralpak IE (Daicel Chiral Technologies) 5 µm (250 x 4.6 mm) HPLC analytical column; flow rate 1 mL/min; MS detection ranging from 150 to 1500 *m/z*; solvent hexanes (+0.1 % FA): ethanol (0.1 % FA) 80:20). The retention time of fructose from the hydrolyzate with 23.3 min was compared to hydrolyzed D- and L- fructose whose retention times were 22.9 min and 28.8 min, respectively.

2.3.6 LC/MS analysis of L/D-FDLA derivatives

Approximately 0.2 mg of compound 1, 2, and 3 were hydrolyzed with 6 N HCl (0.8 mL) and shaken at 90°C for 16 h. Samples were dried under vacuum and dissolved in H_2O (100 µL). It was split into two 50 µL aliquots and 1 N $NaHCO_3$ (20 µL) and 1% 1-fluoro-2,4-dinitrophenyl-5-leucine-amide (L-FDLA or D-FDLA solution in acetone, 100 µL) were added respectively. The mixtures were heated to 40° C for 40 min, allowed to cool to RT and neutralized with 2 N HCl (20 µL). Solvents were evaporated to dryness under nitrogen and residues were dissolved in 1 mL CH_3CN and analyzed by LCMS. All measurements were performed on a Dionex Ultimate 3000 RSLC system using a Waters BEH C18, 100 x 2.1 mm, 1.7 µm column by injection of 1 µl sample. Separation was achieved by a gradient using (A) H_2O + 0.1% FA to (B) ACN + 0.1% FA at a flow rate of 550 µl/min and 45 °C. The gradient was as follows: starting at 5 % B to increase to 10 % B

in 1 min, from 1 to 15 min increase to 35 % B, from 15 to 22 min increase to 50% B, from 22 to 25 min increase to 80 % B. After a 1 min hold at 80 % B the system was reequilibrated with initial conditions for 5 minutes. UV data was acquired at 340 nm and MS–detection was performed simultaneously. Coupling the HPLC to the MS was supported by an Advion Triversa Nanomate nanoESI system attached to a Thermo Fisher Orbitrap. LC flow is split to 500 nL/min before entering the ion source. Mass spectra were acquired in centroid mode ranging from 150 to 1000 *m/z* at a resolution of R = 30000. Results can be found in Table S1.

Table S1. Results of Marfey's analysis of the cystomanamides compared to the respective standards

	detected *m/z*	Cystomanamide A		Cystomanamide B		Cystomanamide C		Standard	
		L-FDLA	D-FDLA	L-FDLA	D-FDLA	L-FDLA	D-FDLA	L-FDLA	D-FDLA
AMDA	296.28	25.84	24.69	25.84	24.69	25.85	24.69	--	--
L-Asn	428.14	13.75	13.02	13.80	13.00	13.80	13.01	13.02	13.79
***erythro*-β-OH-L-Asn**	444.14	12.51	12.07	12.54	12.08	12.50	12.08	12.47	12.03
***threo*-β-OH-L-Phe**	476.18	19.61	15.76	19.63	15.75	19.60	15.76	15.74	19.60
L-Tyr	476.18	--	--	13.30	13.15	--	--	13.15	13.31
***threo*- β -OH-L-Asn**	444.14	--	--	--	--	--	--	10.61	10.61
***erythro*-β-OH-L-Phe**	476.18	--	--	--	--	--	--	15.91	18.54

2.3.7 Feeding experiments

Cultivation of MCy9118 was performed in 50 mL VY/2S medium. The shake flask cultures were grown for 3 days. L-[5,5,5,-*d3*]leucine was dissolved in DMSO and 25 µL were added to the culture in two portions (after 1 h and 18 h) to a final concentration of 0.2 mM. On day 3, 2 % (w/v) XAD adsorber resin (Amberlite XAD-7, Sigma) was added to the culture and the culture was harvested after 2h by centrifugation. The sediment was extracted with 5 mL MeOH for 30 min and filtered. The solvent was evaporated *in vacuo* and the residue was redissolved in 200 µL MeOH and analyzed by LC-HRMS.

2.3.7.1 Feeding of labelled L-leucine-*5,5,5*-D3

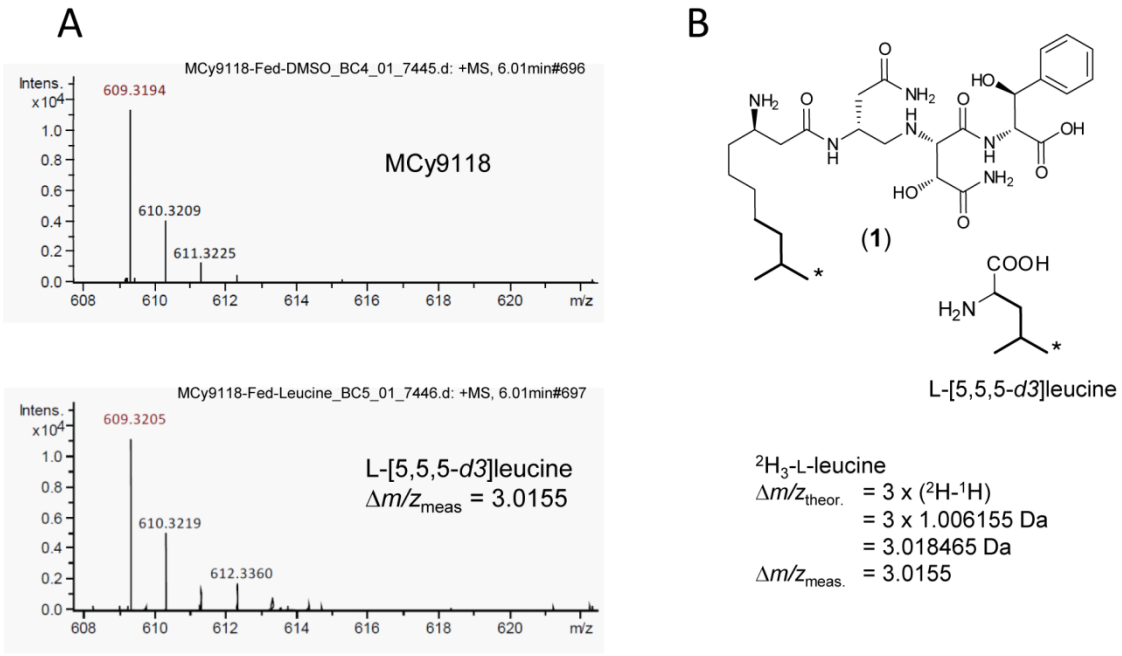


Figure S1: A: Isotopic peak pattern of the highly abundant $[M+H]^+$ signals for cystomanamide A (609 *m/z*) prove incorporation of labeled L-leucine in MCy9118. The observed mass shifts fit to the heavy isotopes that were incorporated. B: Location of the deuterium label in the fed leucine (*) and likely incorporation pattern of leucine skeleton into the fatty acid moiety of cystomanamide A (bold lines).

2.3.8 Gene disruption in MCy9118 via single crossover homologous recombination

A homologous fragment of each target gene with a size between 975 and 1222 bp was amplified from genomic DNA using the respective oligonucleotides that additionally inserted stop codons on each side of the fragment. The fragment was ligated into the vector pCR®II-TOPO®, resulting in the respective pTOPO derivative for gene disruption. The plasmids were transformed into *E. coli* strain SCS110 for amplification of the plasmid without Dam or Dcm methylation. Clones containing the plasmid were selected on LB-Agar (low salt) supplemented with 50 mg/L Kanamycin-sulfate (Kan) at 37°C. The plasmid was recovered from the strain by the alkaline lysis method (Qiagen MiniPrep Kit), and after restriction analysis it was introduced into MCy9118 by electroporation (650 V, 400 Ω). Mutant clones of MCy9118 were selected on M-Agar (1.5% agar) containing 50 mg/L Kan. Single clones were used to inoculate liquid medium (M medium + 50 mg/L Kan) for secondary metabolite analysis and isolation of genomic DNA for genetic verification by PCR. To genetically verify the mutants, two PCR reactions were performed for each mutant. The first PCR was performed with primers that bind outside of the

region of plasmid insertion. A product is only generated if the gene is not disrupted by the plasmid backbone so that the mutants should not show a product in contrast to the compared wild type. The second PCR was performed with one primer that binds outside of the region of plasmid insertion and one that binds the plasmid backbone. A product is only generated if the plasmid was inserted in the right position so that the mutants should show a product in contrast to the compared wild type were no product is expected. Mutants fulfilling both conditions were considered as genetically verified.

Table S2. Deduced functions of ORFs in cystomanamide biosynthesis

Orf	protein similarity	source organism	identity [%]	reference
1	JmjC domain-containing protein	*Nostoc* sp.	29	ADL59767
2	protease / beta-lactamase	*Streptomyces sviceus* ATCC 29083	36	WP_007382053
3	beta-lactamase	*Nakamurella multipartita* DSM 44233	33	YP_003202535
4	beta-lactamase	*Streptomyces canus*	30	WP_020124107
CtmA	putative Aspartate racemase	*Phaeospirillum molischianum*	40	WP_002726614
CtmB	tarine catabolism dioxygenase TauD/TfdA	*Nostoc punctiforme* ATCC 29133	43	YP_001866803
CtmC	NRPS	*Paenibacillus alvei*	36	WP_005546175
CtmD	amino acid adenylation protein	*Clostridium cellulolyticum* H10	30	YP_002506645
5	short-chain dehydrogenase/reductase	*Methylobacterium extorquens* AM1	54	YP_002964386
6	LysR family transcriptional regulator	*Rhizobium etli* CFN 42	58	YP_468471

* Within or in the close neighborhood of the cystomanamide biosynthetic gene cluster, no genes were identified showing significant homology to those in charge of the pro-drug activating mechanism described by Bode and coworkers (Reimer, D.; Bode, H.B. *Nat. Prod. Rep.* 2014, *31*, 154-9)

Table S3. Primers used for amplification and verification of the cystomanamide knockout mutants in MCy9118

***ctmA*KO**	
*ctmA*KOFor	TGATTGATTGAGTCTTCGTCCCCTACAAGCGGATC
*ctmA*KORev	TCAATCAATCAAGTCGCCCTTCACCTGGAACC
Verify_*ctmA*KO_for	GATCATGGAGCAGCAACTGGC
Verify_*ctmA*KO_rev	CACAGCAGGCGGTAGTACAGC
***orf5*KO**	
*orf5*KOFor	TGATTGATTGAATGATGCCCGTTCGTGGCGCGTTG
*orf5*KORev	TCAATCAATCATCACCCGACGTGGTGTTCGCCGTG
Verify_*orf5*KO_for	CTCCACCGGCTCGTCCATTGG
Verify_*orf5*KO_rev	GAGTAGGTCGCGTGTTCAGAAACC
***orf2*KO**	
*orf2*KOFor	TGATTGATTGA GAAGCCTCGCCGGACGAGTCC
*orf2*KORev	TCAATCAATCAGTTGCTCAACACCACGACCGTCAG
Verify_*orf2*KO_for	CTCGGAGCCTGTGCCACGACG
Verify_*orf2*KO_rev	CCCACGTGGTTGCCATCCAGG
***orf4*KO**	
*orf4*KOFor	TGATTGATTGA AGCTCGCTGGCAGGCCTGCTC
*orf4*KORev	TCAATCAATCACGTCTCGATGAACGTCTGCTGCTG
Verify_*orf4*KO_for	CGAACCCATCCATGAACCCTATCG
Verify_*orf4*KO_rev	CCACGAGGATGTTGTCGGTCG
pTOPO backbone	
pTOPOin	CCTCTAGATGCATGCTCGAGC
pTOPOout	TTGGTACCGAGCTCGGATCC

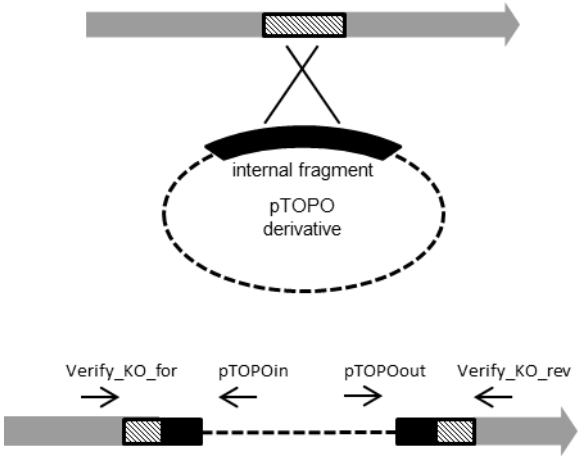


Figure S2: Principle of homologous recombination and PCR verification used for the *Cystobacter fuscus* MCy9118 mutants

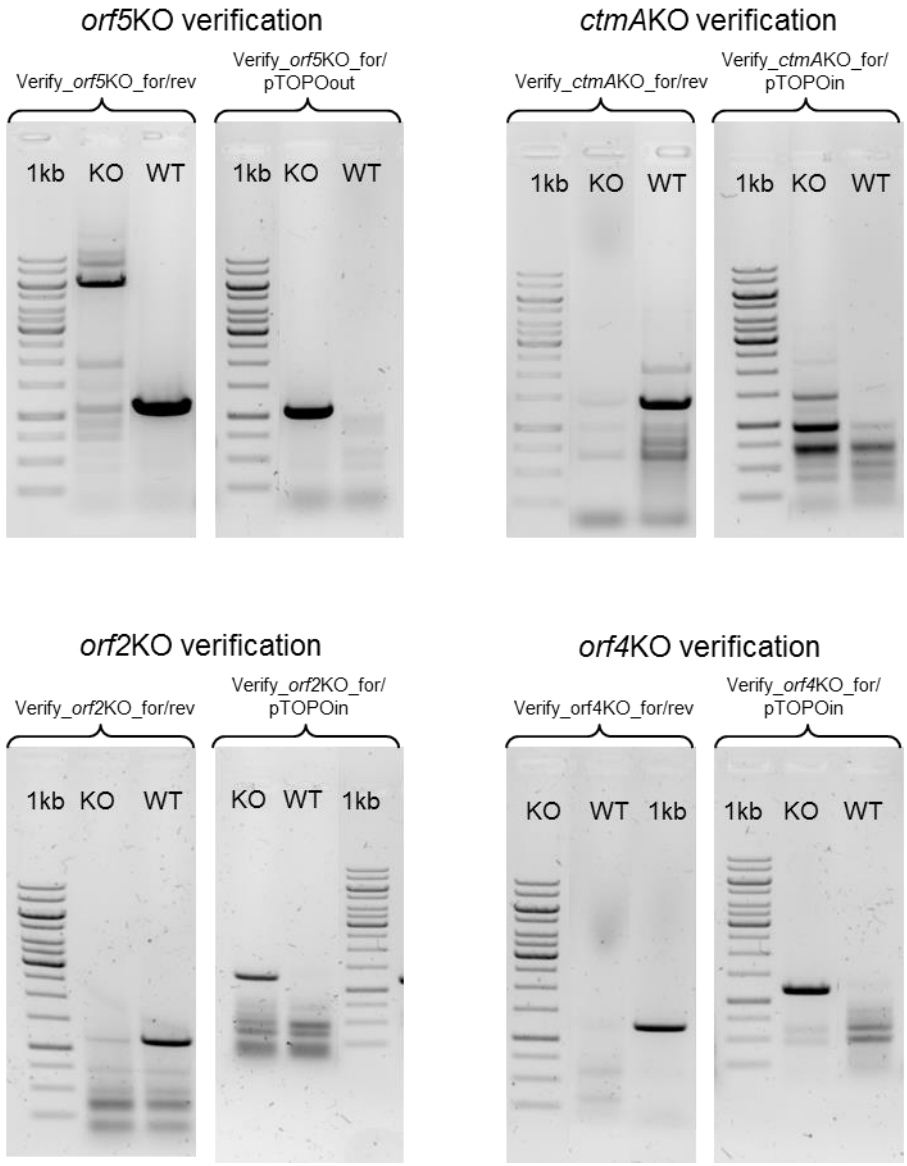


Figure S3. Verification of knockout mutants MCy9118::pLE_*ctmA*KO, MCy9118::pLE_*orf2*KO, MCy9118::pLE_*orf4*KO and MCy9118::pLE_*orf5*KO using PCR

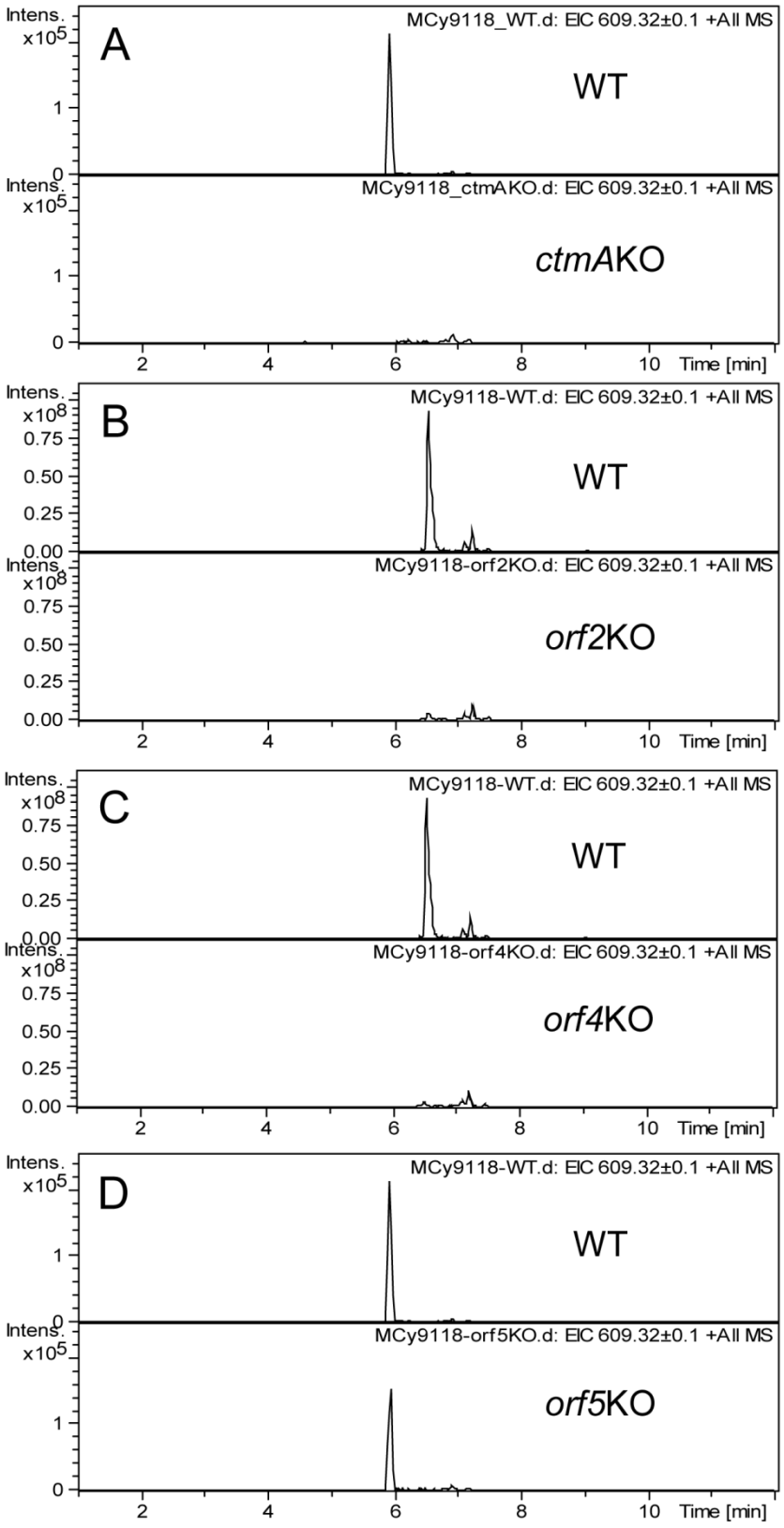


Figure S4. Loss of the cystomanamide A production as detected by comparative LC-MS analysis of MCy9118 wildtype and mutant strains. An extracted ion chromatogram for the $[M+H]^+$ signal of cystomanamide A is shown. A: *ctmA* knockout mutant; B: *orf2* knockout mutant; C: *orf5* knockout mutant; D: *orf4* knockout mutant

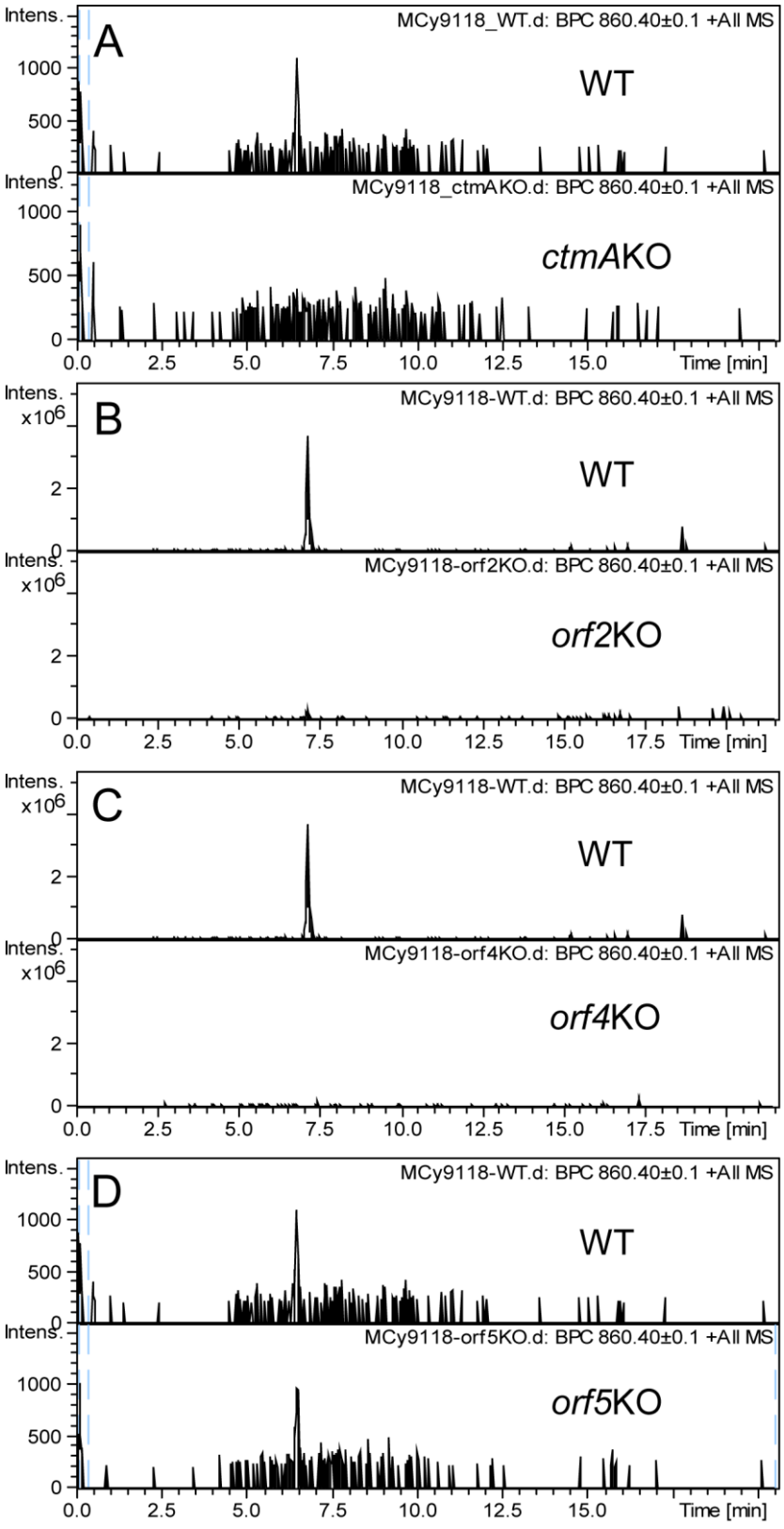


Figure S5. Loss of the cystomanamide B production as detected by comparative LC-MS analysis of MCy9118 wildtype and mutant strains. An extracted ion chromatogram for the $[M+H]^+$ signal of cystomanamide A is shown. A: *ctmA* knockout mutant; B: *orf2* knockout mutant; C: *orf5* knockout mutant; D: *orf4* knockout mutant

Table S4. NMR spectroscopic data for cystomanamide A (**1**) (methanol-d_4)

	δ_C[a]	δ_H[b]	mult (*J* in Hz)	HMBC[c]	COSY[d]	TOCSY[d]
	AMDA					
1	173.2					
2a	37.0	2.68	o	1,3	2b, 3	2b, 3, 4, 5, 6, 7, 8
2b		2.50	dd (17.0, 9.75)	1, 3, 4	2a, 3	2a, 3, 4, 5, 6, 7, 8
3	49.8	3.54	m	-	2a, 2b, 4	2, 4, 5, 6, 7, 8
4a	33.8	1.67	m	2, 3, 5, 6	3, 4b, 5	2, 3, 4b, 5, 6, 7, 8
4b		1.62	m	2, 3, 5	3, 4a, 5	2, 3, 4a, 5, 6, 7, 8
5	26.2	1.41	dd (7.3, 7.3)	6	4, 6	2, 3, 4, 6, 7, 8, 10, 11
6	30.5	1.35	o	5, 7, 8	5, 7	2, 3, 4, 5, 7, 8, 9, 10, 11
7	28.1	1.36	o	5, 6, 8	6, 8	2, 3, 4, 5, 6, 8, 9, 10, 11
8	40.0	1.21	m	6, 7, 9, 10, 11	7, 9	2, 3, 4, 5, 6, 7, 9, 10, 11
9	29.0	1.55	m	7, 8, 10, 11	8, 10, 11	5, 6, 7, 8, 10, 11
10	22.9	0.89	d (6.65)	8, 9, 11	9	7, 8, 9
11	22.9	0.89	d (6.65)	8, 9, 10	9	7, 8, 9
	Asn					
1	172.8					
2	51.6	4.77	dd (7.4, 6.3)	1, 3, 4, 1_{AMDA}	3	3
3a	36.4	2.89	dd (15.8, 6.3)	1, 2, 4	2, 3b	2, 3b
3b		2.68	o	1, 2, 4	2, 3a	2, 3a
4	174.9					
	β-OH-Asn					
1	170.1					
2	56.9	4.80	d (3.75)	1, 3, 4, 1_{Asn}	3	3
3	72.8	4.29	d (3.75)	1, 2, 4	2	2
4	176.3					
	β-OH-Phe					
1	176.2					
2	61.0	4.49	d (3.25)	1, 3, 4, $1_{β\text{-}OH\text{-}Asn}$	3	3
3	74.9	5.22	d (3.25)	1, 2, 4, 5, 9	2	2
4	143.3					
5	127.5	7.34	d (7.45)	3, 7, 9	6, 8	6, 7, 8
6	128.9	7.28	t (7.45)	4, 8	5, 9, 7	5, 7, 9
7	128.2	7.20	t (7.45)	5, 9	6, 8	5, 6, 8, 9
8	128.9	7.28	t (7.45)	4, 6	5, 7, 9	5, 7, 9
9	127.5	7.34	d (7.45)	3, 5, 7	6, 8	6, 7, 8

[a]Recorded at 175 MHz; referenced to residual methanol-d_4 at δ 49.15 ppm.
[b]Recorded at 700 MHz; referenced to residual methanol-d_4 at δ 3.31 ppm.
[c]Proton showing correlation to indicated carbon.
[d]Proton showing correlation to indicated proton.

Table S5. NMR spectroscopic data for cystomanamide B (**2**) (methanol-d_4)

	δ_C[a]	δ_H[b]	mult (*J* in Hz)	HMBC[c]	COSY[d]	TOCSY[d]
	AMDA					
1	172.6					
2a	37.3	2.70	dd (16.5, 3.7)	1, 3, 4	2b, 3	2b, 3, 4, 5, 6, 7
2b		2.54	dd (16.5, 9.2)	1, 3, 4	2a, 3	2a, 3, 4, 5, 6, 7
3	50.1	3.52	m	2, 4	2, 4	2, 4, 5, 6, 7
4a	33.8	1.66	m	2, 3, 5, 6	3, 5	2, 3, 4b, 5, 6, 7, 8
4b		1.62	m	2, 3, 5, 6	3, 5	2, 3, 4a, 5, 6, 7, 8
5	26.3	1.40	overlap	4, 6	4, 6	2, 3, 4, 6, 7, 8, 10, 11
6	30.6	1.33	overlap	4, 5, 7, 8	5, 7	2, 3, 4, 5, 7, 8, 9, 10, 11
7	28.2	1.34	overlap	6, 8, 9	6, 8	2, 3, 4, 5, 6, 8, 9, 10, 11
8	40.0	1.20	m	6, 7, 9, 10, 11	7, 9	4, 5, 6, 7, 9, 10, 11
9	29.0	1.53	m	7, 8, 10, 11	8, 10, 11	6, 7, 8, 10, 11
10	22.9	0.89	d (6.6)	8, 9, 11	9	5, 6, 7, 8, 9
11	22.9	0.89	d (6.6)	8, 9, 10	9	5, 6, 7, 8, 9
	Asn					
1	173.0					
2	52.1	4.72	dd (7.6, 5.8)	1, 3, 4, 1_{AMDA}	3	3
3a	37.0	2.80	dd (15.8, 7.6)	1, 2, 4	2, 3b	2, 3b
3b		2.76	dd (15.8, 5.8)	1, 2, 4	2, 3a	2, 3a
4	174.5					
	β-OH-Asn					
1	170.6					
2	57.2	4.77	d (4.2)	1, 4, 1_{Asn}	3	3
3	72.7	4.30	d (4.2)	1, 2, 4	2	2
4	176.0					
	β-OH-Phe					
1	171.0					
2	60.1	4.75	d (3.6)	1, 3, 4, $1_{β\text{-OH-Asn}}$	3	3
3	74.3	5.26	d (3.6)	1, 2, 4, 5, 9	2	2
4	141.5					
5	127.5	7.35	d (7.5)	3, 7, 9	6, 8	6, 7, 8
6	129.3	7.31	t (7.5)	4, 5, 8, 9	5, 7, 9	5, 7, 9
7	128.9	7.24	t (7.3)	5, 9	6, 8	5, 6, 8, 9
8	129.3	7.31	t (7.5)	4, 5, 6, 9	5, 7, 9	5, 7, 9
9	127.5	7.35	d (7.5)	3, 5, 7	6, 8	6, 7, 8
	GA					
1	172.2					
2	71.0	4.24	dd (6.0, 4.0)	1, 3a, 3b	3	3
3a	67.9	4.32	dd (11.1, 4.0)	1, $1_{β\text{-OH-Phe}}$	2, 3b	2, 3b
3b		4.16	dd (11.1, 6.0)	1, $1_{β\text{-OH-Phe}}$	2, 3a	2, 3a
	Tyr					
1	177.5					
2	57.2	4.43	dd (7.5, 5.5)	1, 3a, 3b, 4, 1_{GA}	3	3
3a	38.2	3.12	dd (13.8, 5.5)	1, 2, 4, 5	2, 3b	2, 3b
3b		2.97	dd (13.8, 7.5)	1, 2, 4, 5	2, 3a	2, 3a
4	129.5					
5/9	131.4	7.05	d (8.4)	6, 8	6, 8	6, 8
6/8	116.1	6.68	d (8.4)	5, 9	5, 9	5, 9
7	156.6					

[a]Recorded at 175 MHz; referenced to residual methanol-d_4 at δ 49.15 ppm.
[b]Recorded at 700 MHz; referenced to residual methanol-d_4 at δ 3.31 ppm.
[c]Carbon showing correlation to indicated proton.
[c]Proton showing correlation to indicated proton.

Table S6. NMR spectroscopic data for cystomanamide C (**3**) (methanol-d_4)

	δ_C[a]	δ_H[b]	mult (*J* in Hz)	HMBC[c]	TOCSY[d]	ROESY[d]
	AMDA					
1	173.4					
2a	34.3	2.71	o	1, 3	2b, 3, 4, 5, 6, 7, 8	
2b		2.60	dd (17.0, 9.0)	1, 3	2a, 3, 4, 5, 6, 7, 8	
3	57.1	3.59	m		2, 4, 5, 6, 7, 8	1_{Fru}
4a	31.2	1.77	m		2, 3, 4b, 5, 6, 7, 8	1_{Fru}
4b		1.62	m		2, 3, 4a, 5, 6, 7, 8	1_{Fru}
5a	26.4	1.43	m	6	2, 3, 4, 5b, 8, 9, 10, 11	
5b		1.35	o	6	2, 3, 4, 5a, 8, 9, 10, 11	
6	30.7	1.35	o	5, 7, 8	2, 3, 4, 8, 9, 10, 11	
7	28.3	1.35	o	5, 6, 8	2, 3, 4, 8, 9, 10, 11	
8	40.0	1.20	m	6, 7, 9, 10, 11	3, 4, 5, 6, 7, 9, 10, 11	
9	29.1	1.53	m	7, 8, 10, 11	7, 8, 10, 11	
10	22.9	0.89	d (6.5)	8, 9	7, 8, 9	
11	22.9	0.89	d (6.5)	8, 9	7, 8, 9	
	Asn					
1	172.6					
2	51.7	4.75	dd (7.5, 6.5)	1, 3, 4, 1_{AMDA}	3	
3a	36.5	2.87	dd (15.9, 6.5)	1, 2, 4	2, 3b	
3b		2.72	o	1, 2, 4	2, 3a	
4	174.7					
	β-OH-Asn					
1	169.9					
2	56.8	4.82	d (3.5)	1, 3, 4, 1_{Asn}	3	
3	72.8	4.24	d (3.5)	1, 2, 4	2	
4	176.4					
	β-OH-Phe					
1	176.3					
2	60.9	4.51	d (3.0)	1, 3, 4, $1_{β\text{-}OH\text{-}Asn}$	3	
3	74.8	5.25	d (3.0)	1, 4, 5, 9	2	
4	142.9					
5	127.4	7.33	d (7.5)	3, 7, 9	6, 7, 8	
6	128.9	7.28	t (7.5)	4, 5, 8, 9	5, 7, 9	
7	128.1	7.20	t (7.5)	5, 9	5, 6, 8, 9	
8	128.9	7.28	t (7.5)	4, 5, 6, 9	5, 7, 9	
9	127.4	7.33	d (7.5)	3, 5, 7	6, 7, 8	
	Fructose					
1	50.8	3.27	s	2, 3, 3_{AMDA}		3_{AMDA}
2	96.7					
3	71.0	3.74	d (9.6)	2, 4, 5	5	1
4	71.1	3.80	dd (9.6, 3.2)	3, 5	5	6a
5	70.8	3.86	broad signal	4, 6	3, 4, 6a, 6b	
6a	65.0	4.02	d (12.6)	2, 5	5, 6b	
6b		3.69	d (12.6, 1.6)	2, 4, 5	5, 6a	

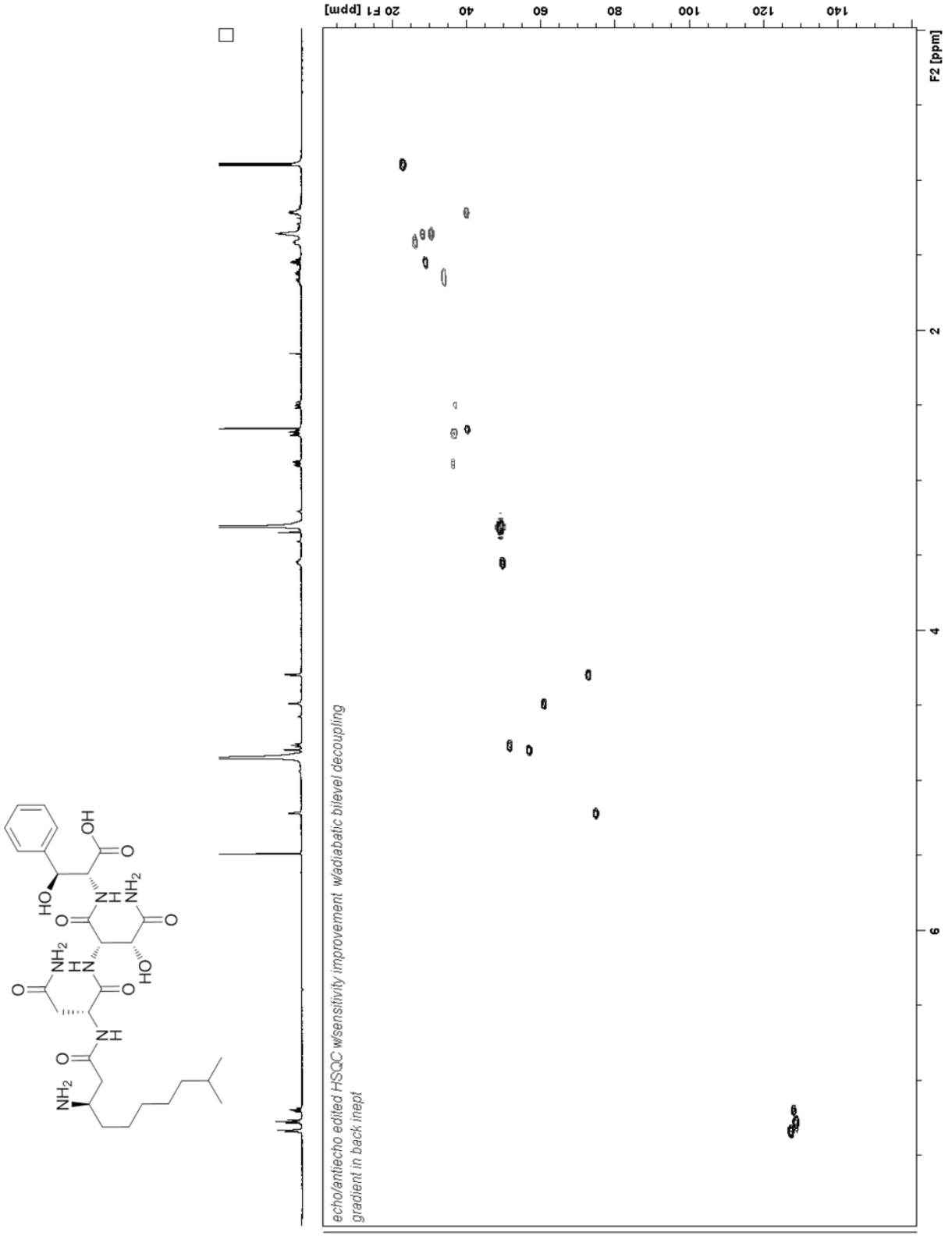
[a]Recorded at 175 MHz; referenced to residual methanol-d_4 at δ 49.15 ppm.
[b]Recorded at 700 MHz; referenced to residual methanol-d_4 at δ 3.31 ppm.
[c]Carbon showing correlation to indicated proton.
[c]Proton showing correlation to indicated proton.

Table S7. NMR spectroscopic data for cystomanamide D (**4**) (methanol-d_4)

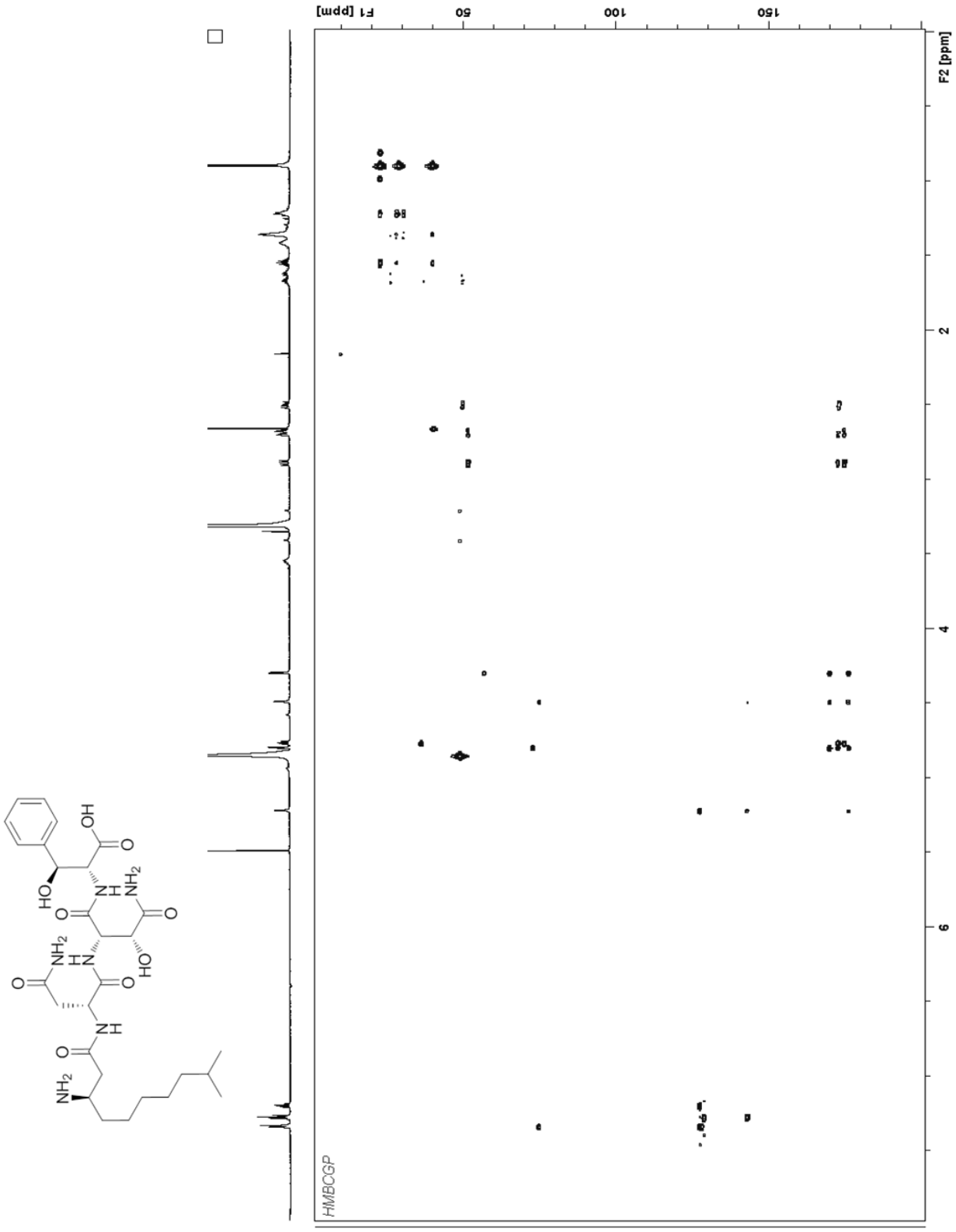
	δ_C[a]	δ_H[b]	mult (*J*, Hz)
	AMDA		
1	173.4		
2a	36.4	2.86	o
2b		2.64	o
3	56.8	3.49	m
4a	31.2	1.71	m
4b		1.54	
5	26.3	1.39	m
6	30.7	1.35	o
7	28.3	1.35	o
8	40.0	1.21	m
9	29.1	1.54	m
10	23.0	0.89	d (6.7)
11	23.0	0.89	d (6.7)
	Asn		
1	172.5		
2	51.5	4.78	o
3a	36.5	2.83	m
3b		2.64	m
4	174.6		
	β-OH-Asn		
1	169.6		
2	57.4	4.86	o
3	73.0	4.34	d (3.5)
4	176.3		
	Phe		
1	177.3		
2	57.2	4.40	m
3a	38.7	3.15	m
3b		3.00	dd (13.1, 6.1)
4	139.0		
5	130.8	7.17	d (7.5)
6	129.2	7.23	t (7.5)
7	127.3	7.16	o
8	129.2	7.23	t (7.5)
9	130.8	7.17	d (7.5)
	Fructose		
1	50.2	3.29	o
2	96.8		
3	71.7	3.71	d (9.8)
4	71.2	3.80	dd (9.8, 3.2)
5	70.1	3.86	broad signal
6a	65.2	4.02	d (12.5)
6b		3.68	dd (12.5, 2.6)

[a]Recorded at 175 MHz; referenced to residual methanol-d_4 at δ 49.15 ppm.
[b]Recorded at 700 MHz; referenced to residual methanol-d_4 at δ 3.31 ppm.

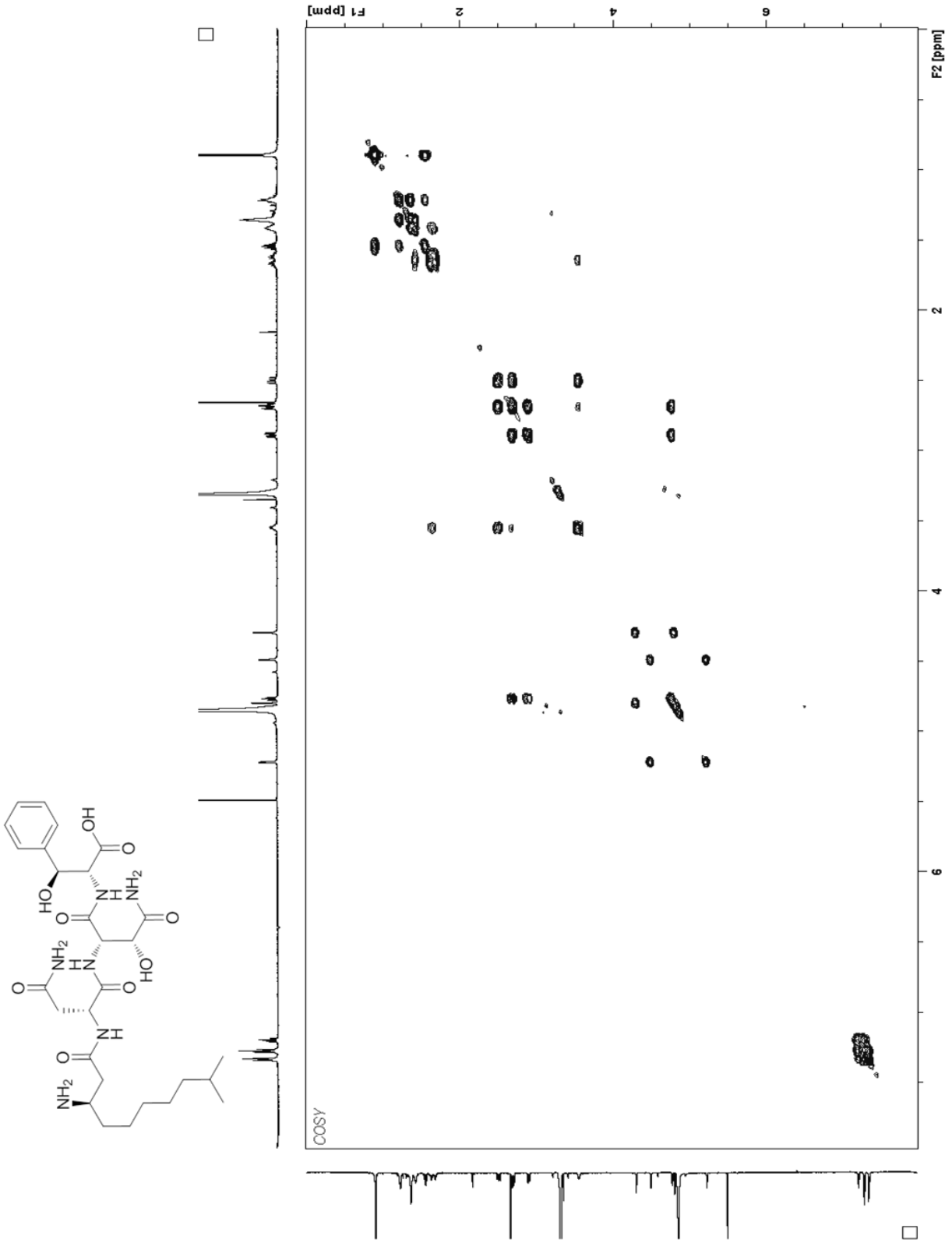
^{1}H NMR spectrum of cystomanamide A (**1**) in methanol-*d4*



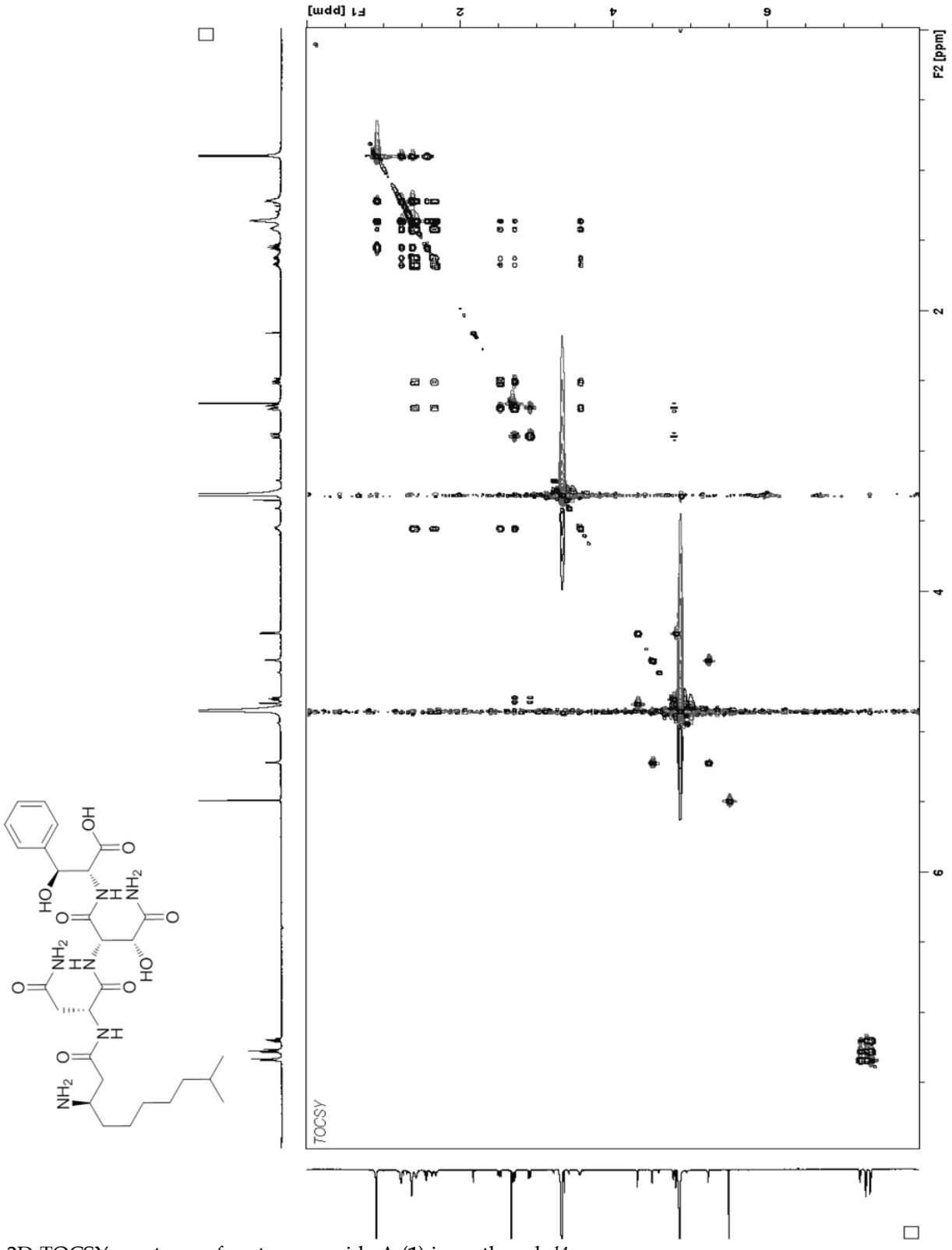
HSQC spectrum of cystomanamide A (**1**) in methanol d4



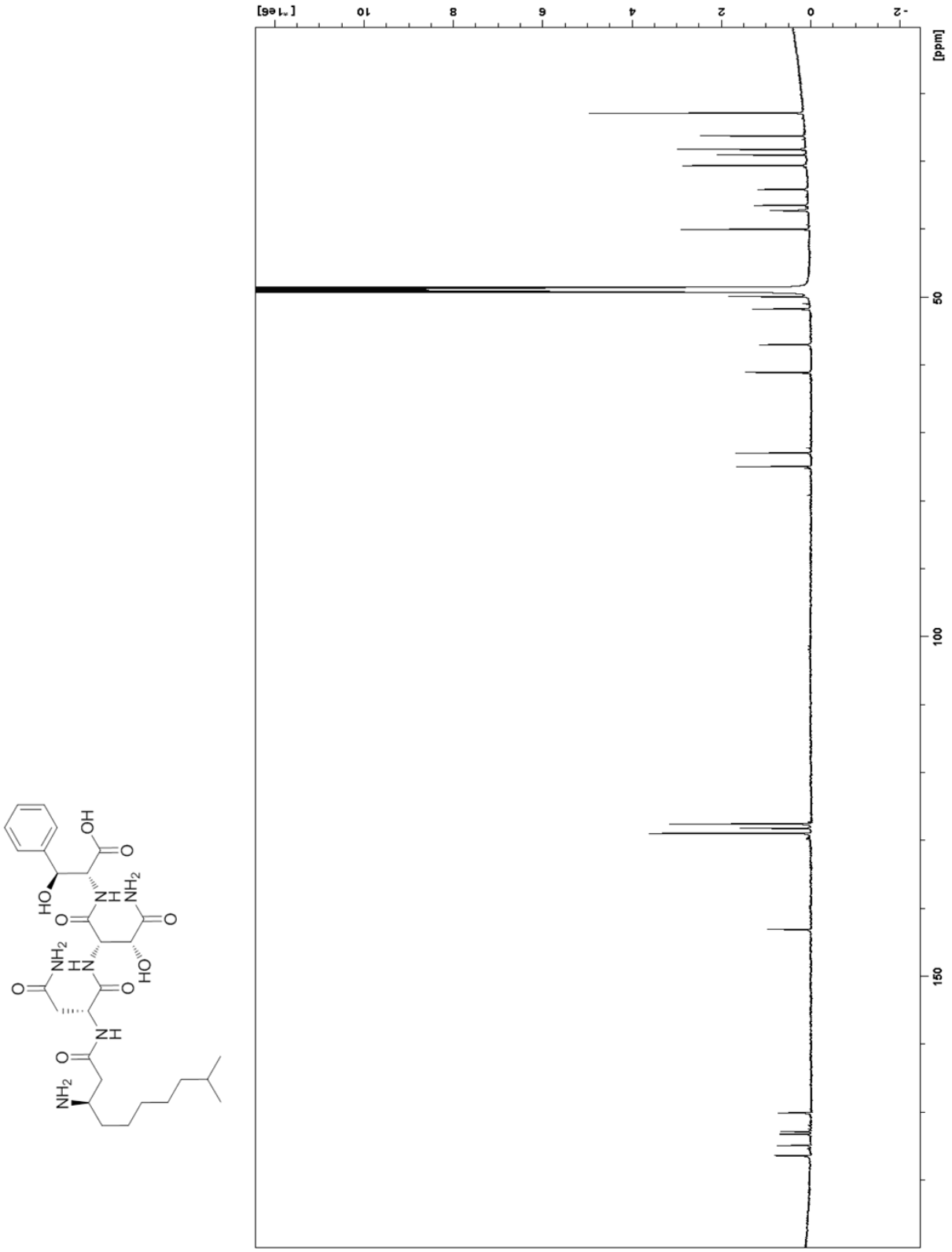
HMBC spectrum of cystomanamide A (**1**) in methanol-d4



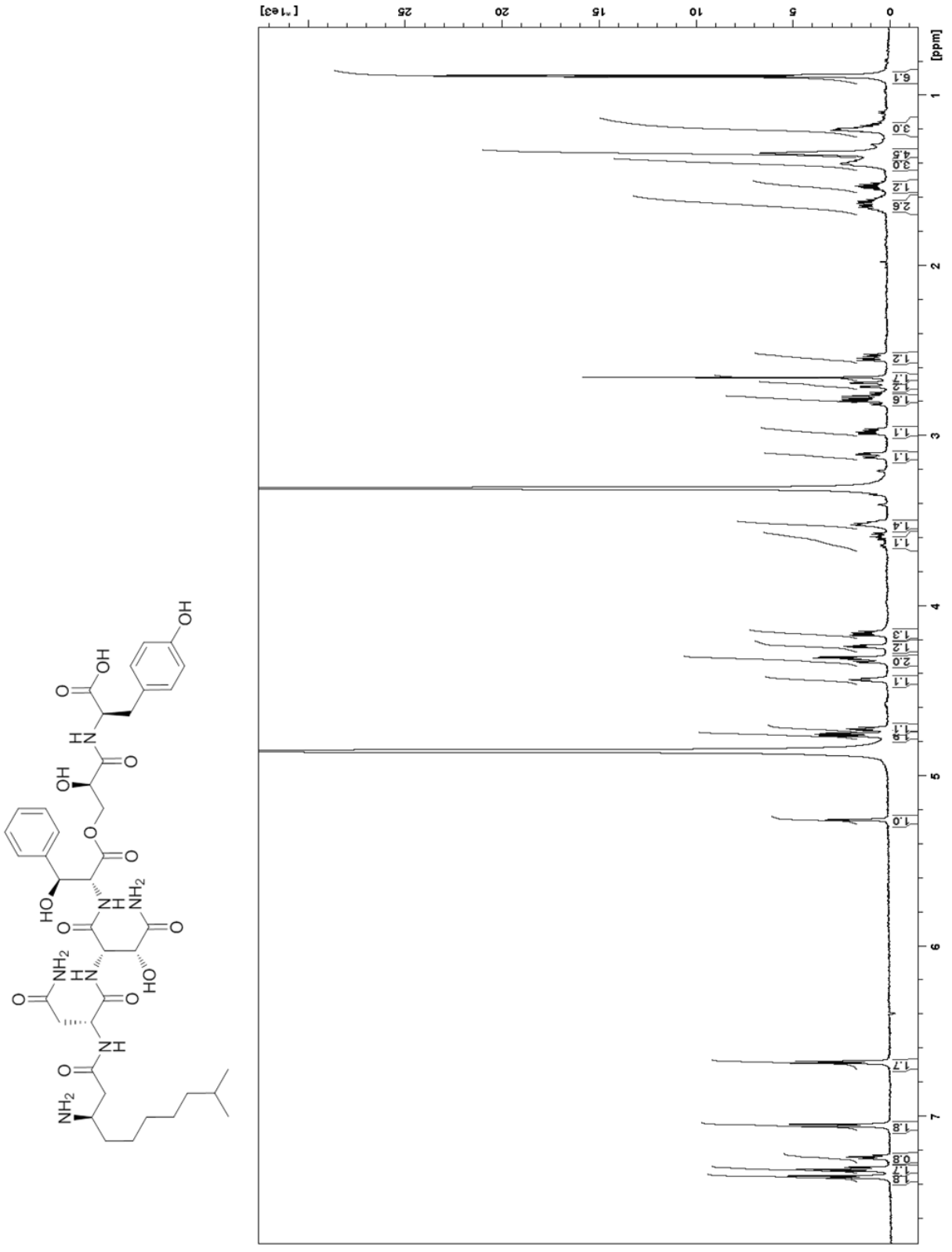
DQF-COSY spectrum of cystomanamide A (**1**) in methanol-d4



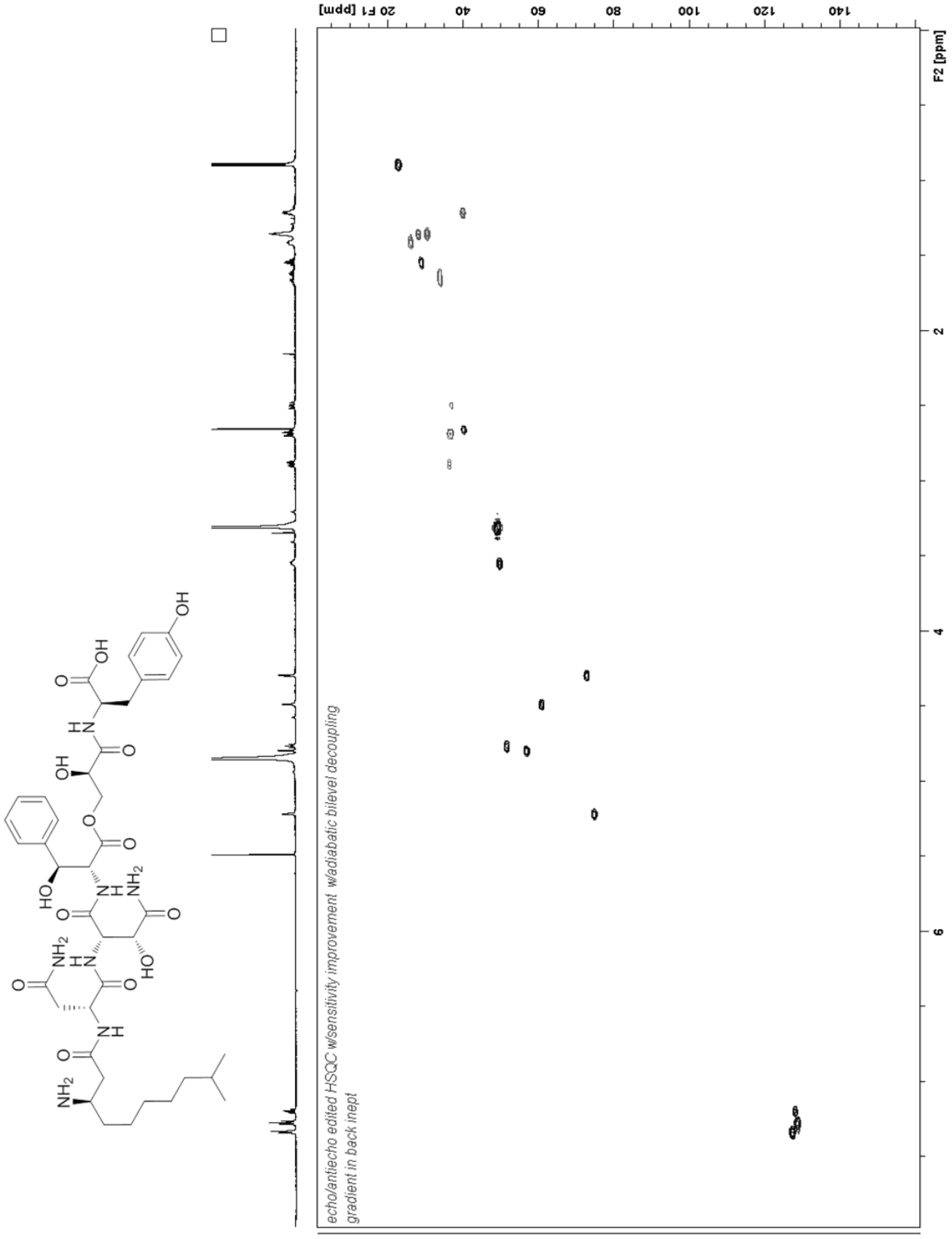
2D-TOCSY spectrum of cystomanamide A (**1**) in methanol-*d4*



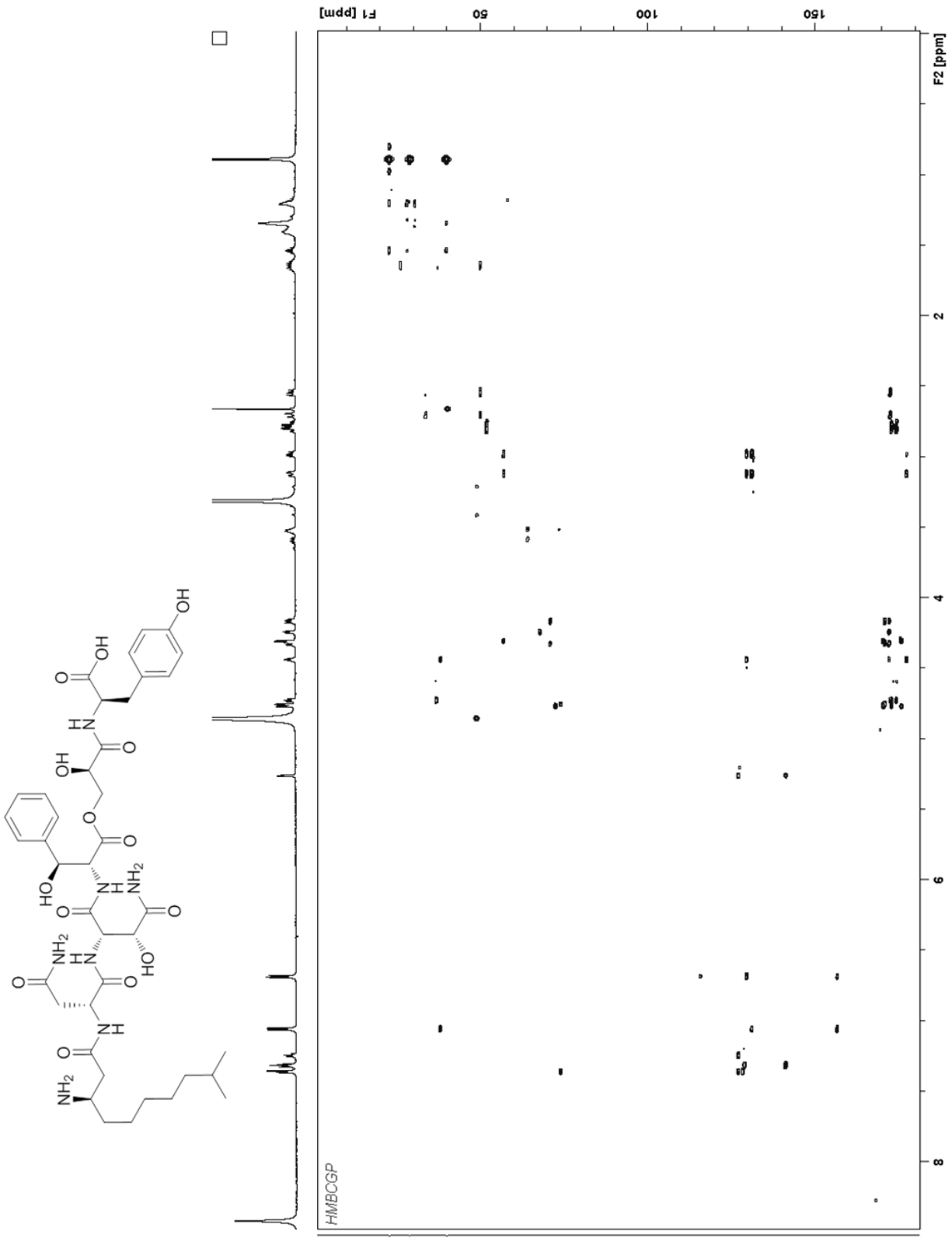
^{13}C spectrum of cystomanamide A (**1**) in methanol-*d4*



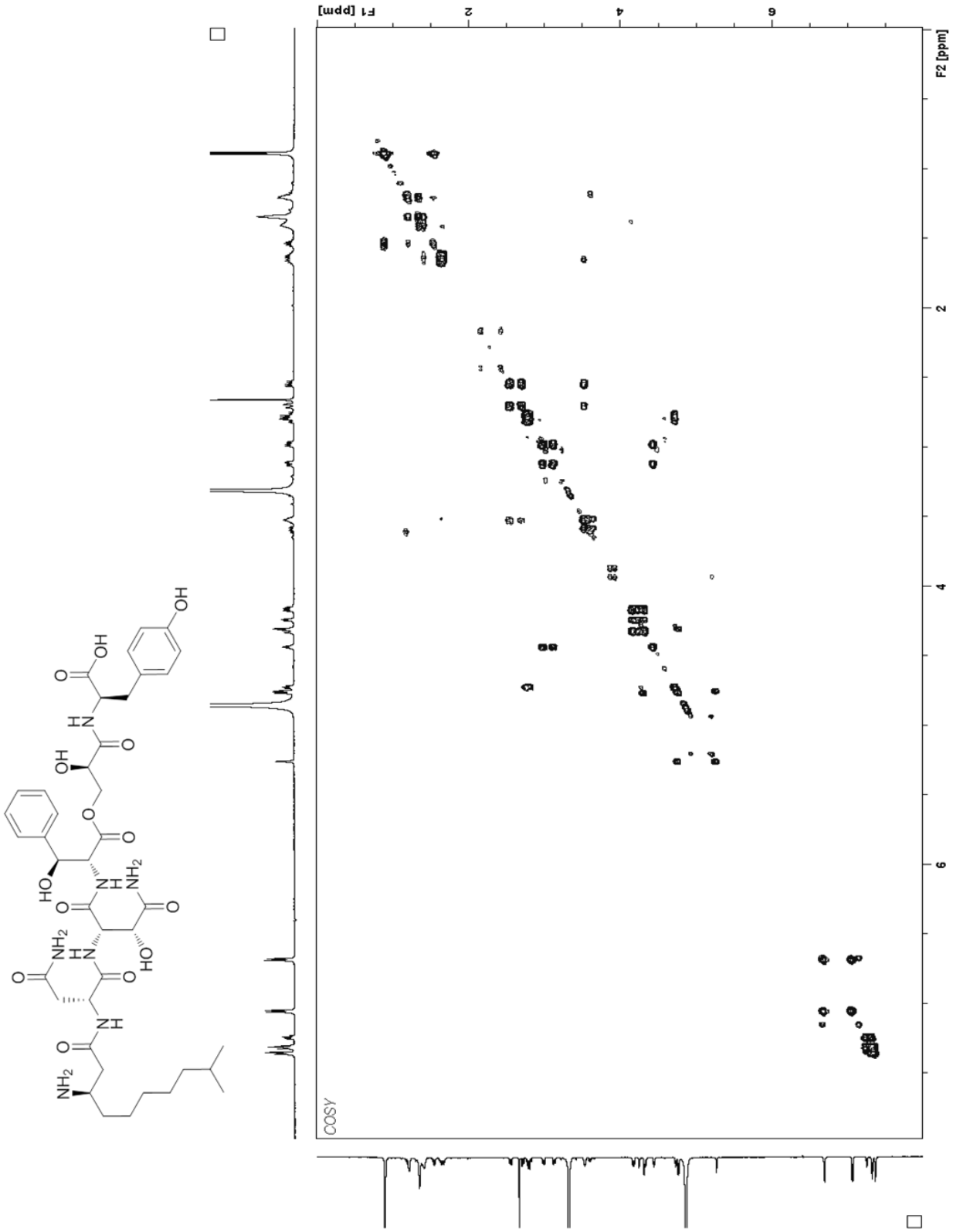
^{1}H NMR spectrum of cystomanamide B (**2**) in methanol-*d4*



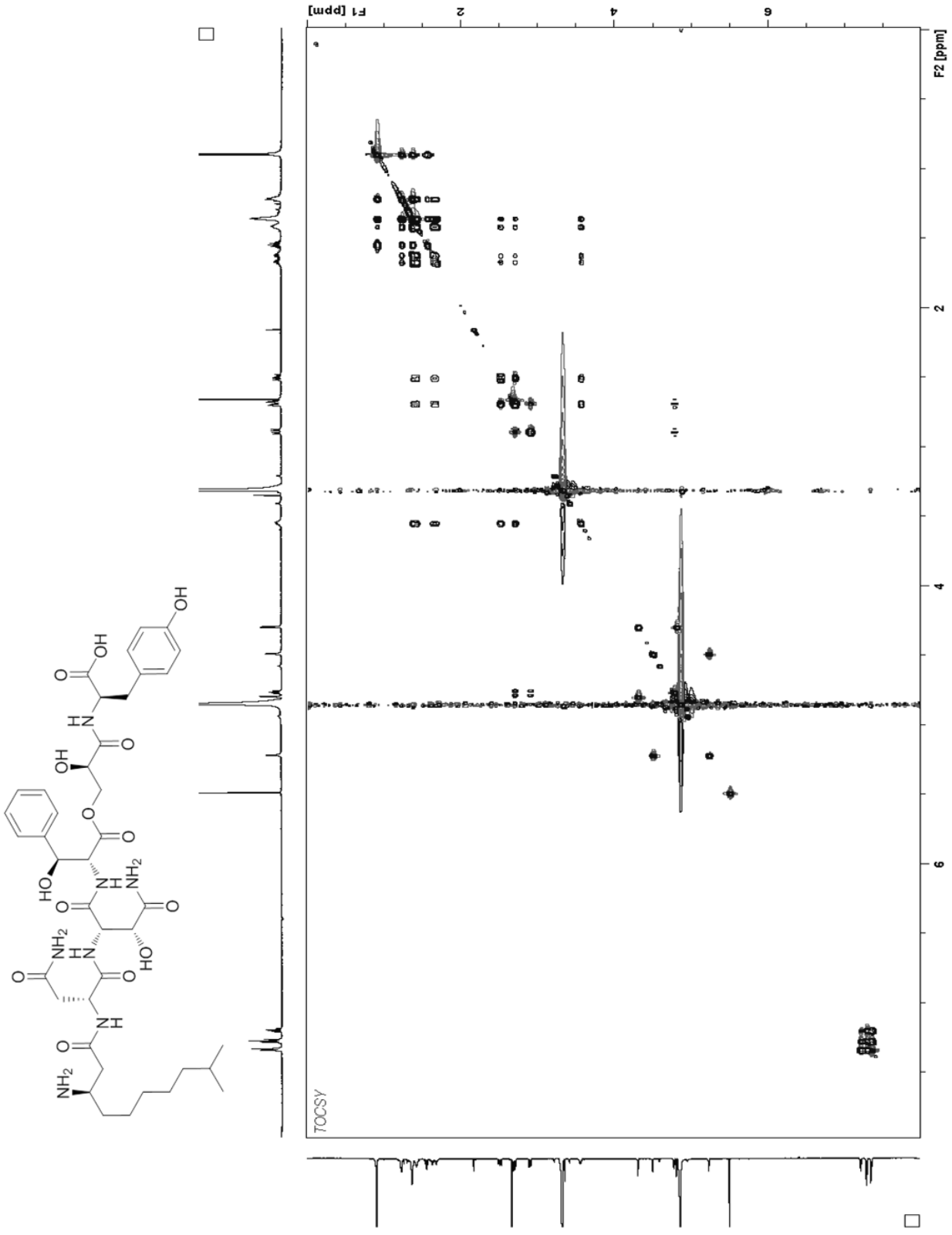
HSQC spectrum of cystomanamide B (**2**) in methanol-*d4*



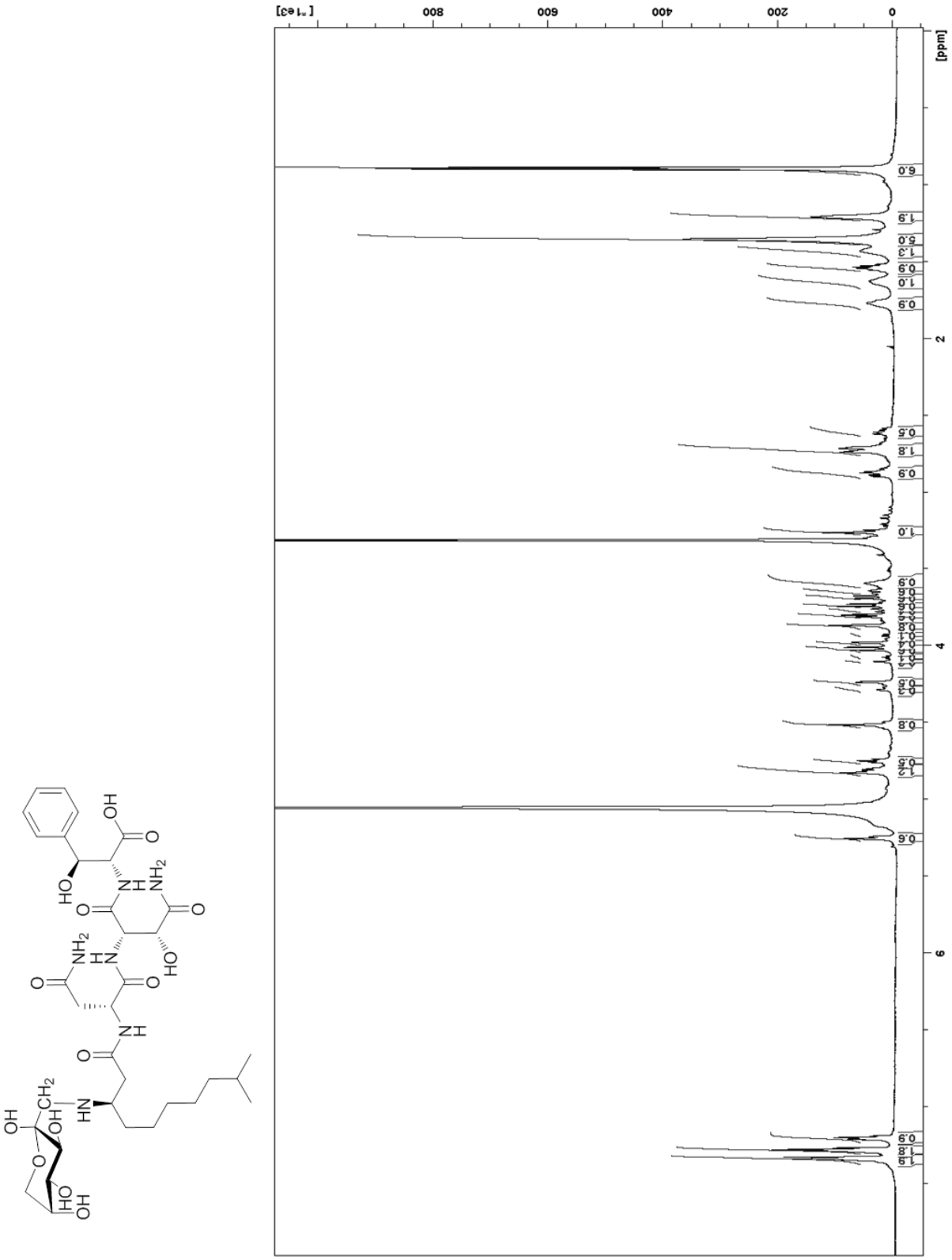
HMBC spectrum of cystomanamide B (**2**) in methanol-*d4*



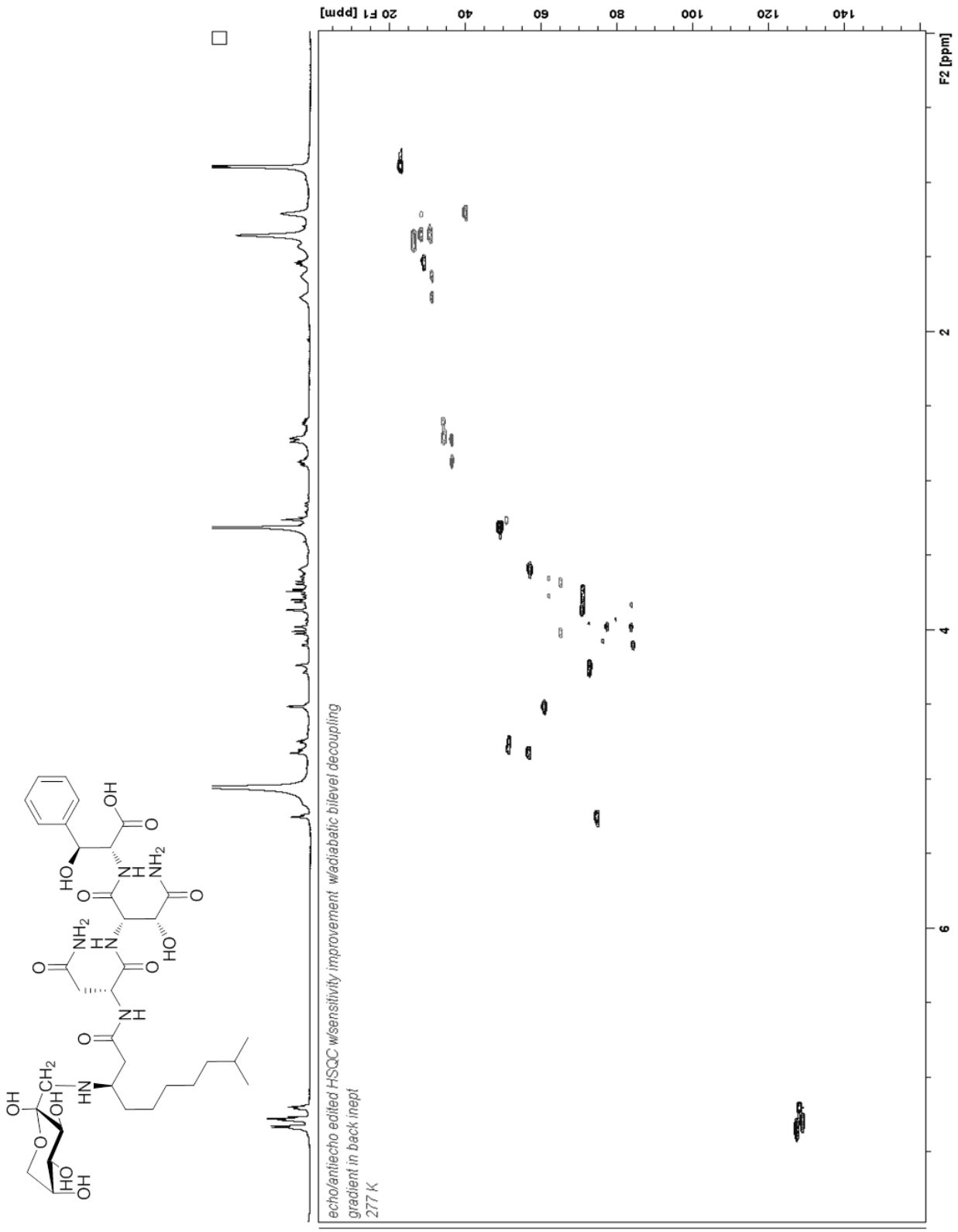
DQF-COSY spectrum of cystomanamide B (**2**) in methanol-*d4*



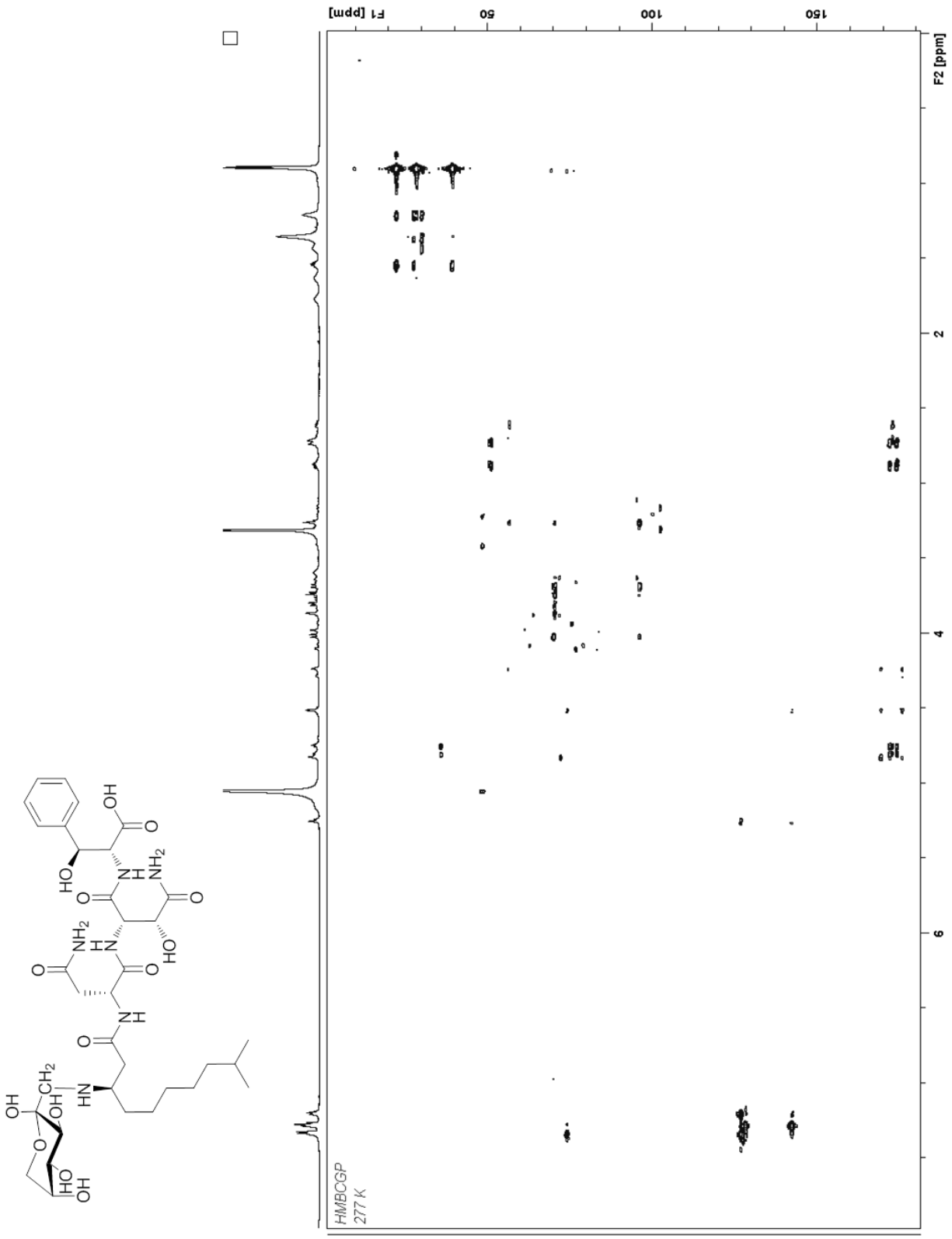
2D-TOCSY spectrum of cystomanamide B (**2**) in methanol-d4



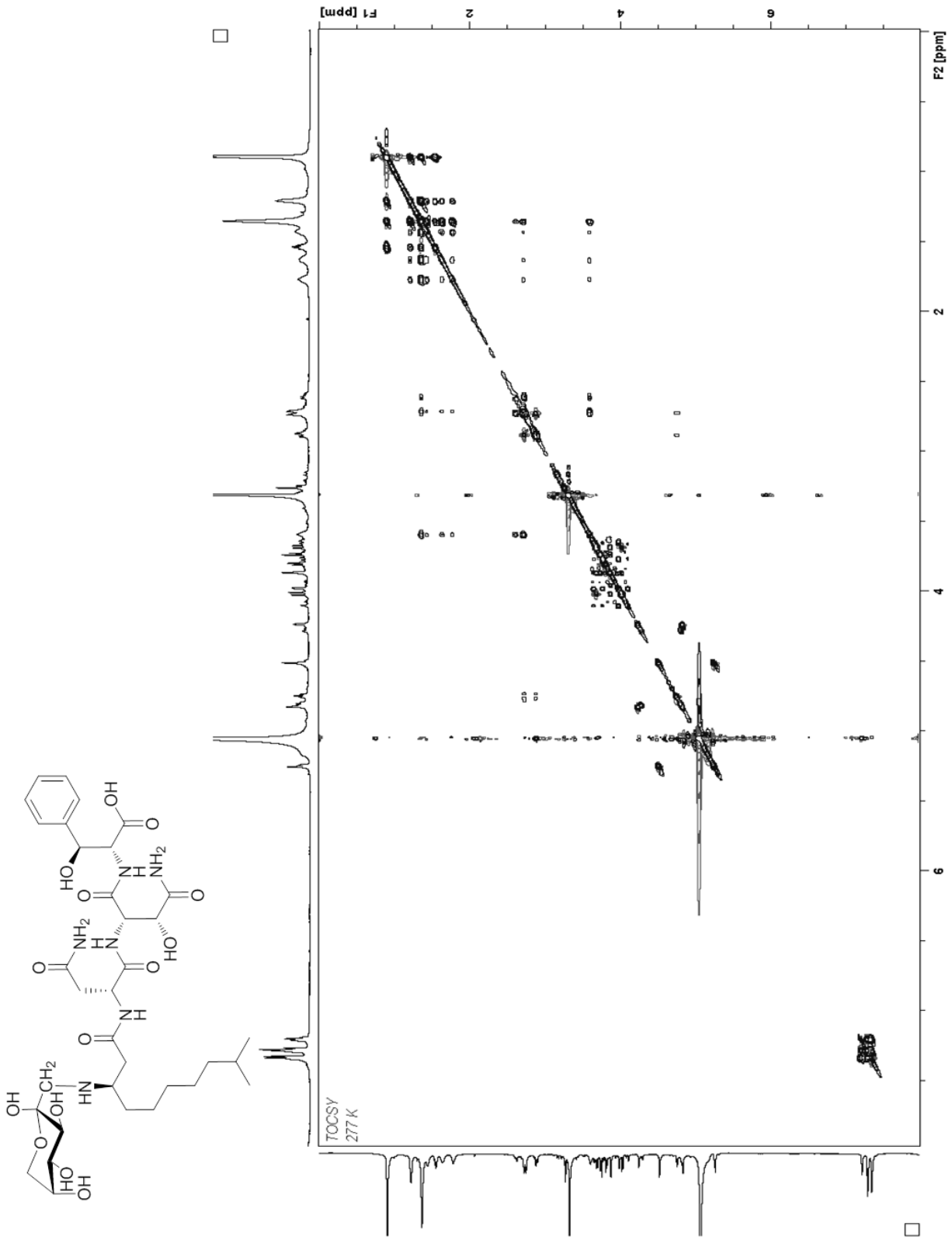
^{1}H NMR spectrum of cystomanamide C (**3**) in methanol-*d4*



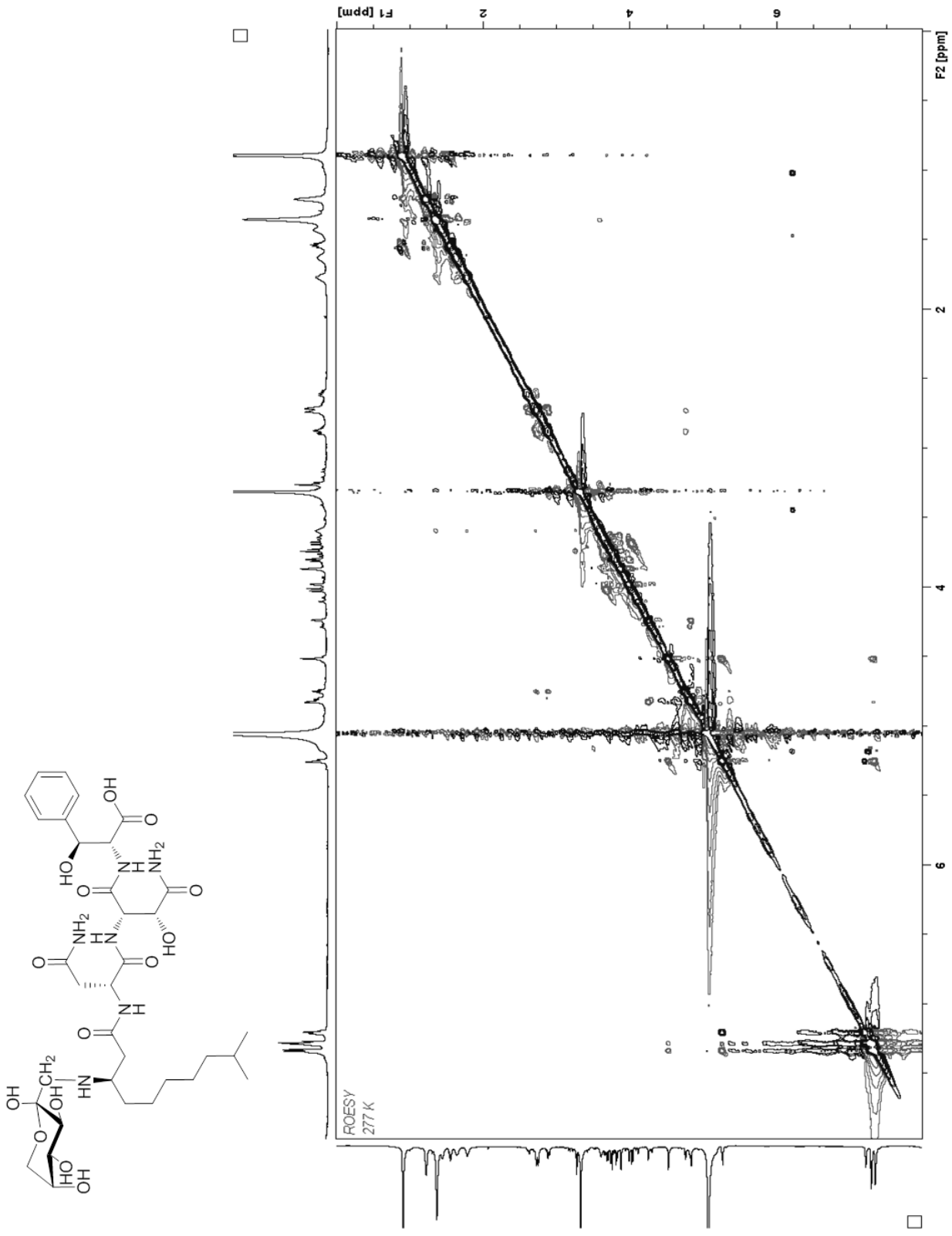
HSQC spectrum of cystomanamide C (**3**) in methanol-*d4*



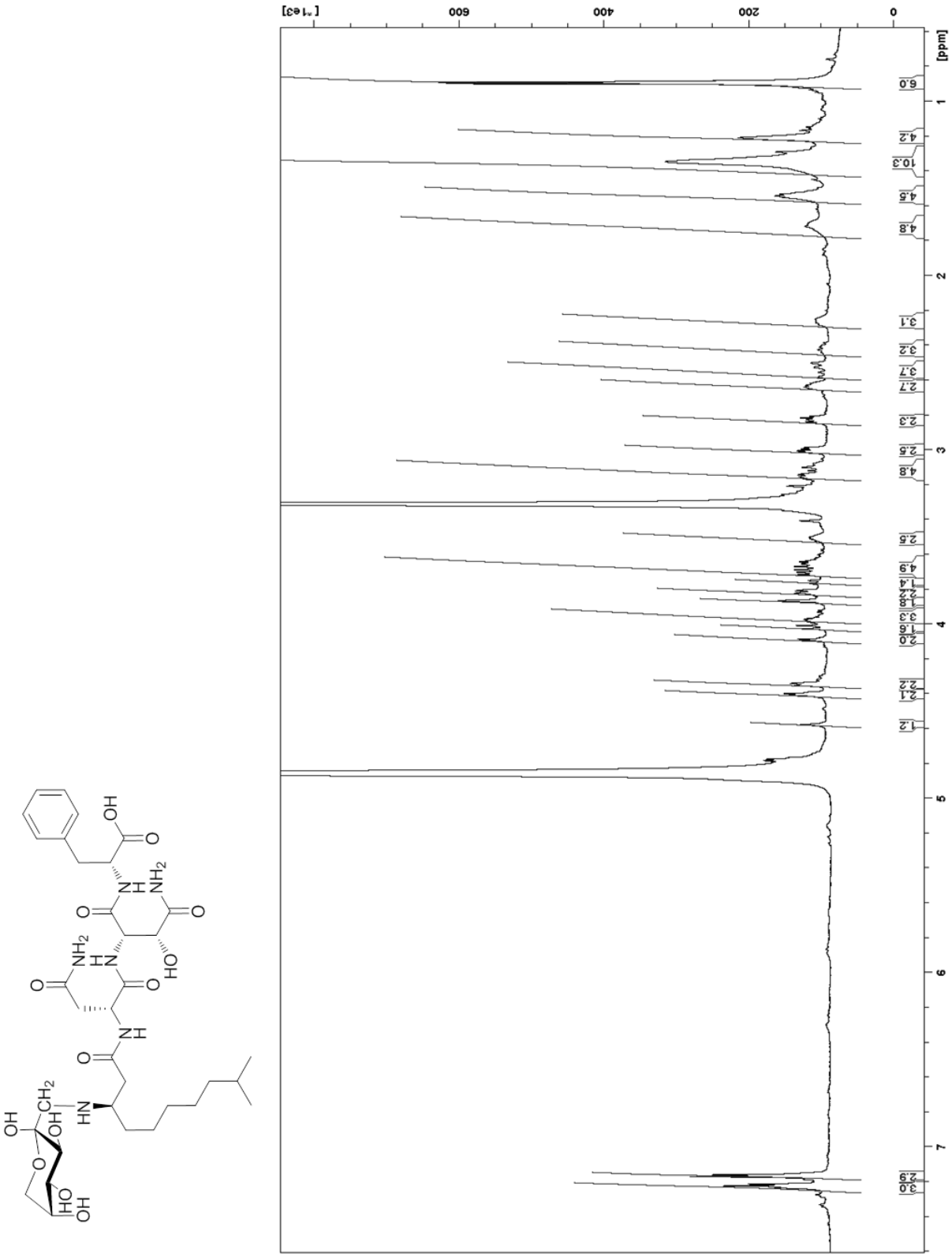
HMBC spectrum of cystomanamide C (**3**) in methanol-*d4*



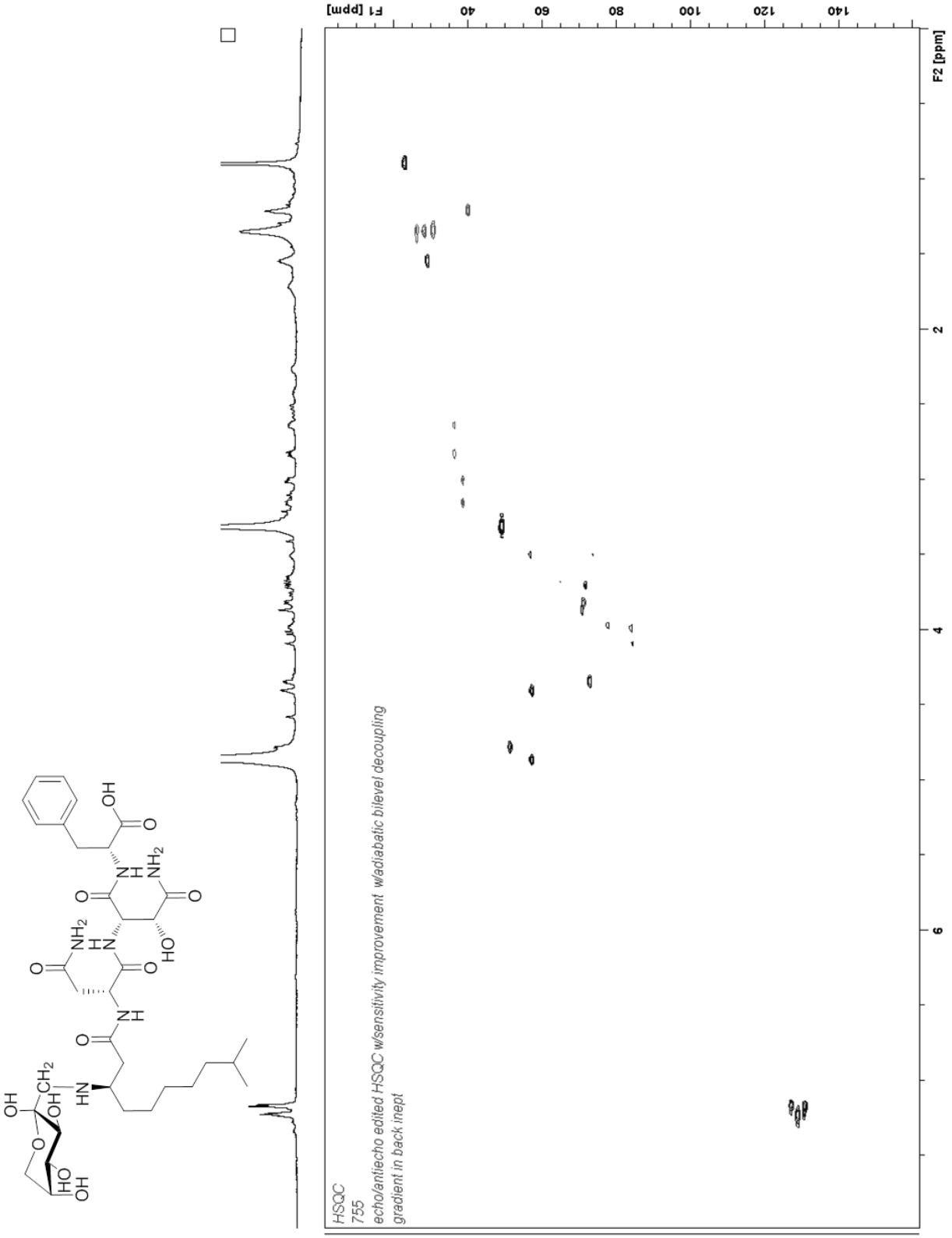
2D-TOCSY spectrum of cystomanamide C (**3**) in methanol-*d4*



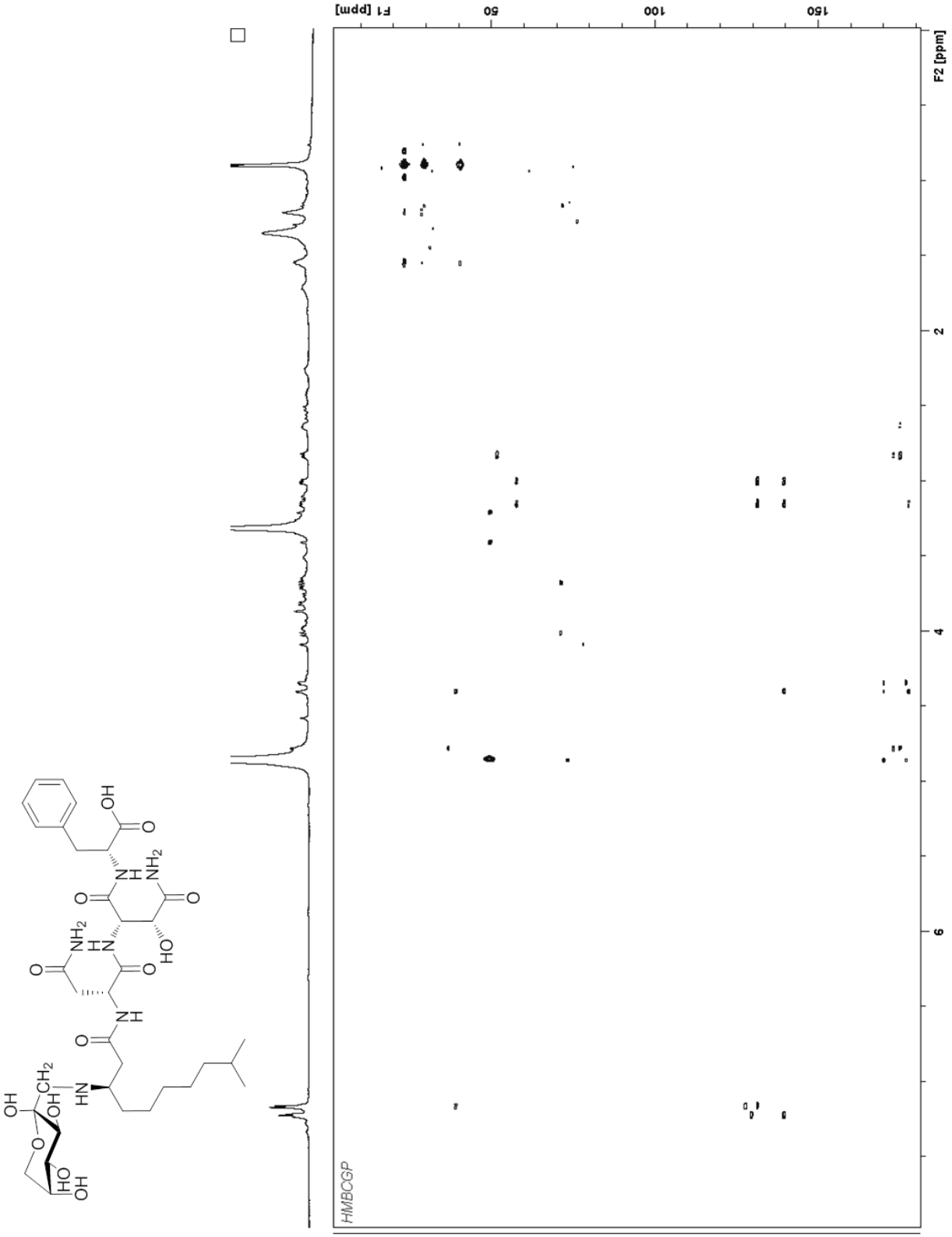
ROESY spectrum of cystomanamide C (**3**) in methanol-*d4*



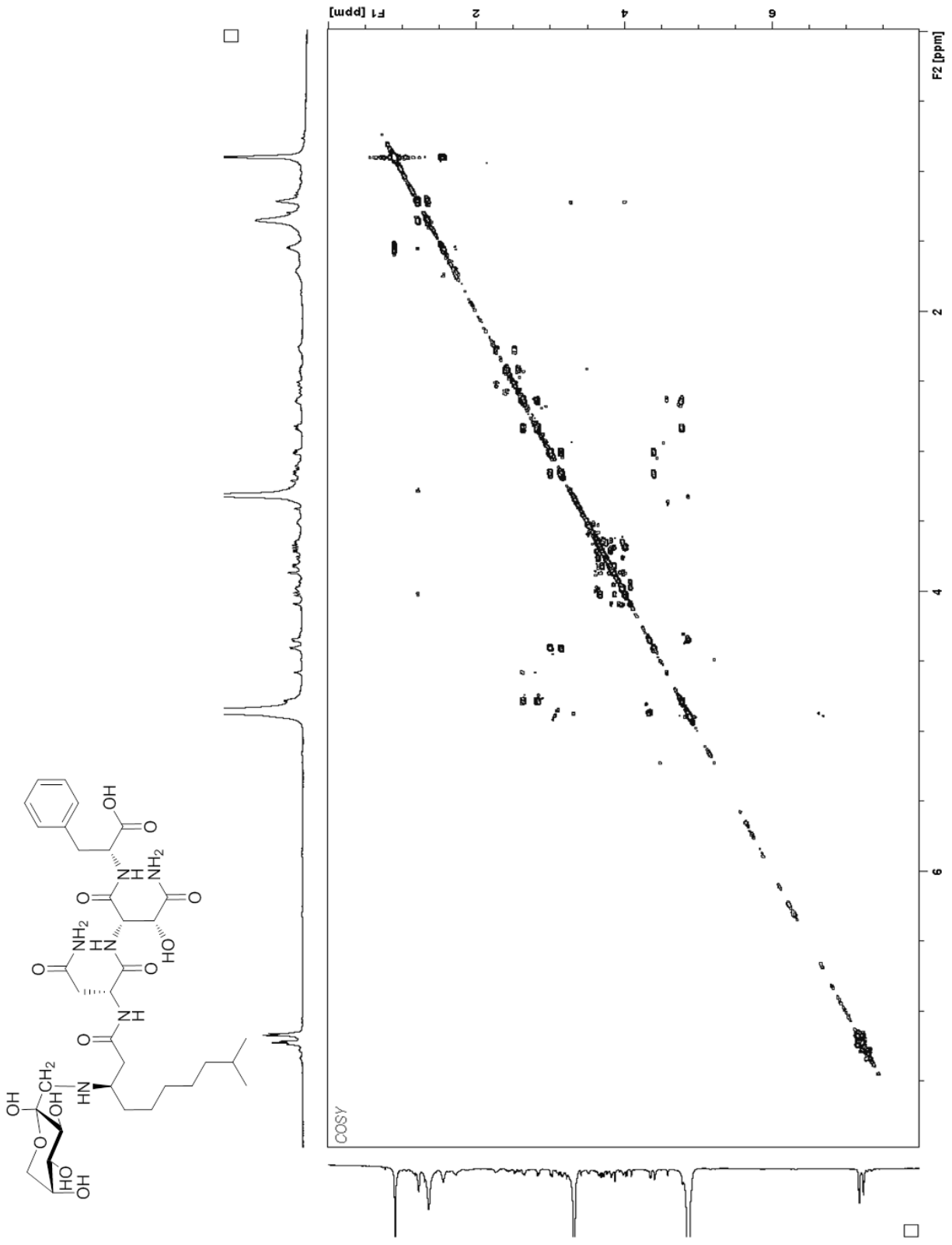
^{1}H NMR spectrum of cystomanamide D (**4**) in methanol-*d4*



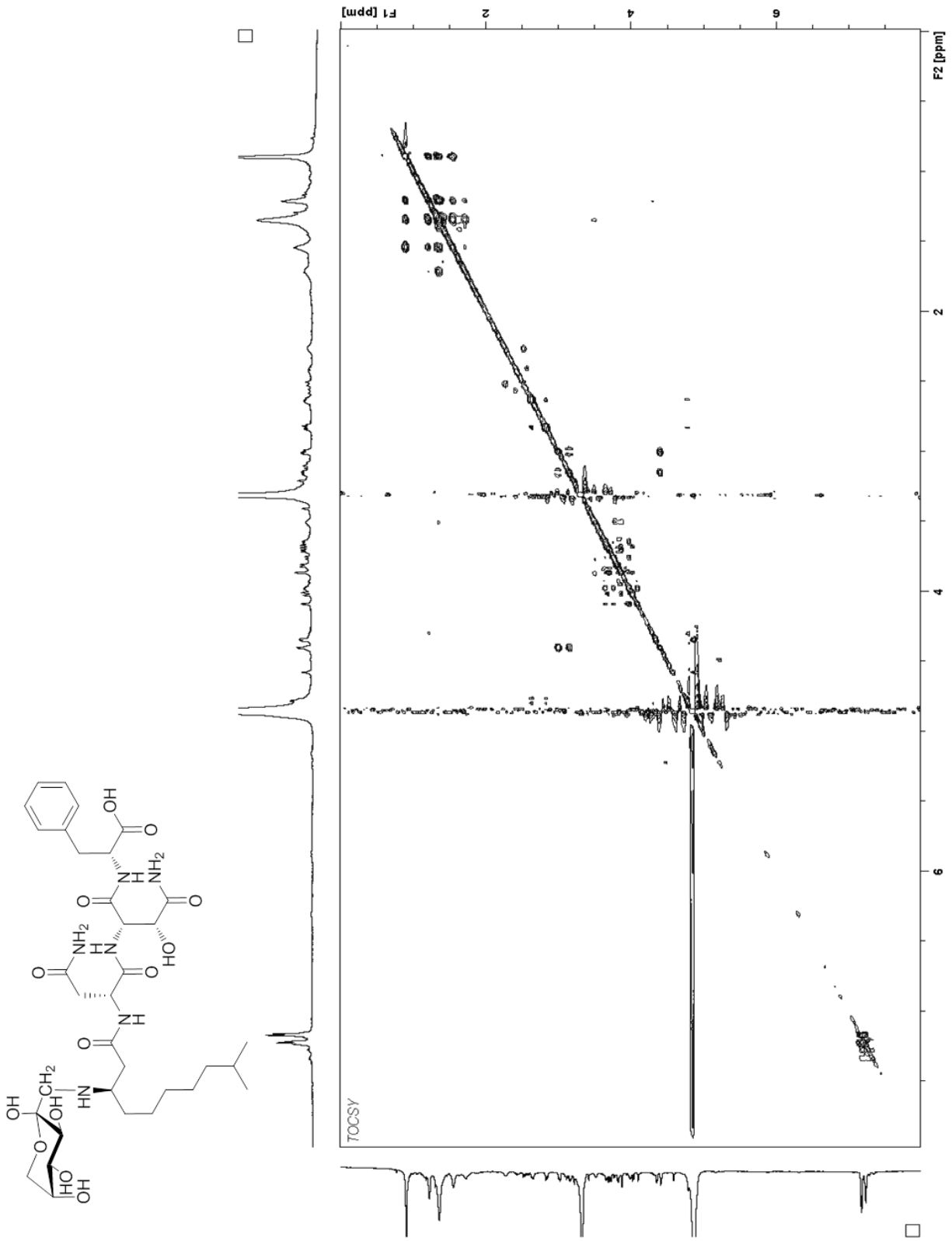
HSQC spectrum of cystomanamide D (**4**) in methanol-*d4*



HMBC spectrum of cystomanamide D (**4**) in methanol-*d4*



DQF-COSY spectrum of cystomanamide D (**4**) in methanol-*d4*



2D-TOCSY spectrum of cystomanamide D (**4**) in methanol-*d4*

2.4 References

(1) Wenzel, S.C.; Müller, R.: The biosynthetic potential of myxobacteria and their impact in drug discovery. *Curr. Opin. Drug Discov. Devel.* **2009**, *12*, 220–30.

(2) Weissman, K.J.; Müller, R.: Myxobacterial secondary metabolites: bioactivities and modes-of-action. *Nat. Prod. Rep.* **2010**, *27*, 1276–95.

(3) Fischbach, M.A.; Walsh, C.T.: Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* **2006**, *106*, 3468–96.

(4) Walsh, C.T.; Chen, H.; et al.: Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr. Opin. Chem. Biol.* **2001**, *5*, 525–34.

(5) Weissman, K.J.; Müller, R.: A brief tour of myxobacterial secondary metabolism. *Bioorg. Med. Chem.* **2009**, *17*, 2121–36.

(6) Yamaguchi, H.; Nakayama, Y.; et al.: A new antibiotic, althiomycin. *J. Antibiot.* **1957**, *10*, 195–200.

(7) Bycroft, B.W.; Pinchin, R.: Structure of althiomycin, a highly modified peptide antibiotic. *J. Chem. Soc. Chem. Commun.* **1975**, 121.

(8) Zander, W.; Gerth, K.; et al.: Roimatacene: an antibiotic against Gram-negative bacteria isolated from Cystobacter ferrugineus Cb G35 (Myxobacteria). *Chem. Eur. J.* **2011**, *17*, 7875–81.

(9) Kunze, B.; Bedorf, N.; et al.: Myxochelin A, a new iron-chelating compound from Angiococcus disciformis (Myxobacterales). Production, isolation, physico-chemical and biological properties. *J. Antibiot.* **1989**, *42*, 14–7.

(10) Ambrosi, H.-D.; Hartmann, V.; et al.: Myxochelins B, C, D, E and F: A New Structural Principle for Powerful Siderophores Imitating Nature. *Eur. J. Org. Chem.* **1998**, *1998*, 541–551.

(11) Ogawa, Y.; Konishi, T.: N-glycosides of amino acid amides from Hemerocallis fulva var. sempervirens. *Chem. Pharm. Bull.* **2009**, *57*, 1110–2.

(12) Schmidt, Y.; van der Voort, M.; et al.: Biosynthetic origin of the antibiotic cyclocarbamate brabantamide A (SB-253514) in plant-associated Pseudomonas. *ChemBioChem* **2014**, *15*, 259–66.

(13) Harada, K.; Fujii, K.; et al.: A method usingL/CMS for determination of absolute configuration of constituent amino acids in peptide --- advanced Marfey's method ---. *Tetrahedron Lett.* **1995**, *36*, 1515–1518.

(14) Gerwick, W.H.; Jiang, Z.D.; et al.: Total structure of hormothamnin A, A toxic cyclic undecapeptide from the tropical marine cyanobacterium hormothamnion enteromorphoides. *Tetrahedron* **1992**, *48*, 2313–2324.

(15) Blin, K.; Medema, M.H.; et al.: antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **2013**, *41*, W204–12.

(16) Punta, M.; Coggill, P.C.; et al.: The Pfam protein families database. *Nucleic Acids Res.* **2012**, *40*, D290–301.

(17) Röttig, M.; Medema, M.H.; et al.: NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **2011**, *39*, W362–7.

(18) Bachmann, B.O.; Ravel, J.: Chapter 8. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. *Methods Enzymol.* **2009**, *458*, 181–217.

(19) Altschul, S.F.; Gish, W.; et al.: Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–10.

(20) Duitman, E.H.; Hamoen, L.W.; et al.: The mycosubtilin synthetase of Bacillus subtilis ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 13294–9.

(21) Bode, H.B.; Dickschat, J.S.; et al.: Biosynthesis of iso-fatty acids in myxobacteria: iso-even fatty acids are derived by alpha-oxidation from iso-odd fatty acids. *J. Am. Chem. Soc.* **2005**, *127*, 532–3.

(22) Silakowski, B.; Nordsiek, G.; et al.: Novel features in a combined polyketide synthase/non-ribosomal peptide synthetase: the myxalamid biosynthetic gene cluster of the myxobacterium Stigmatella aurantiaca Sga15. *Chem. Biol.* **2001**, *8*, 59–69.

(23) Aron, Z.D.; Dorrestein, P.C.; et al.: Characterization of a new tailoring domain in polyketide biogenesis: the amine transferase domain of MycA in the mycosubtilin gene cluster. *J. Am. Chem. Soc.* **2005**, *127*, 14986–7.

(24) Stachelhaus, T.: Peptide Bond Formation in Nonribosomal Peptide Biosynthesis. CATALYTIC ROLE OF THE CONDENSATION DOMAIN. *J. Biol. Chem.* **1998**, *273*, 22773–22781.

(25) Hegg, E.L.; Que, L.: The 2-His-1-carboxylate facial triad- An emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur. J. Biochem.* **1997**, *250*, 625–9.

(26) Ryle, M.J.; Koehntop, K.D.; et al.: Interconversion of two oxidized forms of taurine/alpha-ketoglutarate dioxygenase, a non-heme iron hydroxylase: evidence for bicarbonate binding. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3790–5.

(27) Hausinger, R.P.: FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 21–68.

(28) Singh, G.M.; Fortin, P.D.; et al.: beta-Hydroxylation of the aspartyl residue in the phytotoxin syringomycin E: characterization of two candidate hydroxylases AspH and SyrP in Pseudomonas syringae. *Biochemistry* **2008**, *47*, 11310–20.

(29) Koonin, E. V; Tatusov, R.L.: Computer analysis of bacterial haloacid dehalogenases defines a large superfamily of hydrolases with diverse specificity. Application of an iterative approach to database search. *J. Mol. Biol.* **1994**, *244*, 125–32.

(30) Wu, K.; Chung, L.; et al.: The FK520 gene cluster of Streptomyces hygroscopicus var. ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene* **2000**, *251*, 81–90.

(31) Dorrestein, P.C.; Van Lanen, S.G.; et al.: The bifunctional glyceryl transferase/phosphatase OzmB belonging to the HAD superfamily that diverts 1,3-bisphosphoglycerate into polyketide biosynthesis. *J. Am. Chem. Soc.* **2006**, *128*, 10386–7.

(32) Zaleta-Rivera, K.; Xu, C.; et al.: A bidomain nonribosomal peptide synthetase encoded by FUM14 catalyzes the formation of tricarballylic esters in the biosynthesis of fumonisins. *Biochemistry* **2006**, *45*, 2561–9.

(33) Lin, S.; Van Lanen, S.G.; Shen, B.: A free-standing condensation enzyme catalyzing ester bond formation in C-1027 biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 4183–8.

(34) Magarvey, N.A.; Beck, Z.Q.; et al.: Biosynthetic characterization and chemoenzymatic assembly of the cryptophycins. Potent anticancer agents from cyanobionts. *ACS Chem. Biol.* **2006**, *1*, 766–79.

(35) Fujimori, D.G.; Hrvatin, S.; et al.: Cloning and characterization of the biosynthetic gene cluster for kutznerides. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16498–503.

(36) Xu, Y.; Orozco, R.; et al.: Biosynthesis of the cyclooligomer depsipeptide beauvericin, a virulence factor of the entomopathogenic fungus Beauveria bassiana. *Chem. Biol.* **2008**, *15*, 898–907.

(37) Clissold, P.M.; Ponting, C.P.: JmjC: cupin metalloenzyme-like domains in jumonji, hairless and phospholipase A2beta. *Trends Biochem. Sci.* **2001**, *26*, 7–9.

Chapter 3

Macyranones: Structure, Biosynthesis and Binding Mode of an Antiparasitic Epoxyketone Proteasome Inhibitor from Myxobacteria

Lena Etzbach [†Γ], Alberto Plaza [†Γ#], Christian Dubiella [ϕ], Michael Groll [ϕ], Marcel Kaiser [§‡], Rolf Müller [†Γ*]

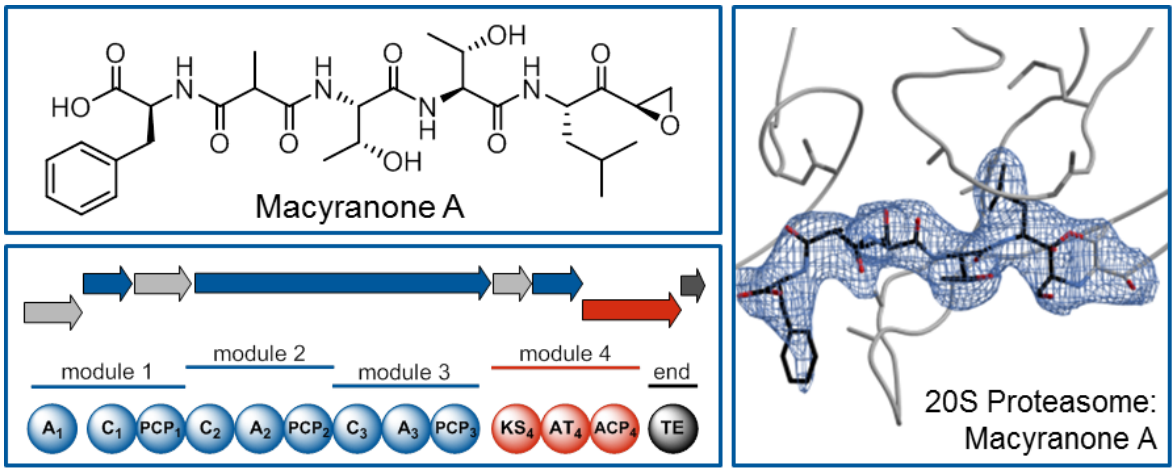
[†] Department of Microbial Natural Products, Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research (HZI) and Pharmaceutical Biotechnology, Saarland University, Campus C2 3, 66123 Saarbrücken, Germany

[Γ]German Center for Infection Research (DZIF), Partner site Hannover-Braunschweig 38124, Germany

[§] Swiss Tropical and Public Health Institute (Swiss TPH), Socinstraße 57, 4002 Basel, Switzerland

[‡] University of Basel, Petersplatz 1, 4003 Basel, Switzerland

[ϕ] Center for Integrated Protein Science Munich (CIPSM), Department für Chemie, Technische Universität München, Lichtenbergstraße 4, 85747 Garching, Germany



3 Macyranones

3.1 Abstract

In our screening efforts to identify unique scaffolds from myxobacteria for the drug discovery process, we used LC-SPE-NMR-MS techniques to isolate six linear peptides, termed macyranone A – F, from *Cystobacter fuscus* MCy9118. The macyranones are characterized by a rare 2-methylmalonamide moiety and an a-amino ketone fragment including an *α'*,*β'*-epoxyketone in macyranone A. Gene disruption experiments allowed to confirm the biosynthetic gene cluster of the macyranones as PKS/NRPS hybrid. Detailed *in silico* and phylogenetic analysis unraveled that the biosynthesis involves two conspicuous amide bond formations accomplished by an amidotransferase and a unique condensation domain. The gene cluster provides further insights into the formation of the powerful epoxyketone residue involving an acyl-CoA dehydrogenase and an unconventional free-standing thioesterase. Macyranone A exhibits a potent inhibitory effect against the parasites *Trypanosoma brucei rhodesiense* and *Leishmania donovani* with IC_{50} values of 1.55 and 0.22 µM, respectively. Unlike other epoxyketones, macyranone A exhibits low cytotoxicity against mammalian and cancer cell lines leaving a potential therapeutic window open for development as anti-parasitic agent. Macyranone A was found to inhibit the chymotrypsin-like activity of the yeast 20S proteasome with an IC_{50} of 5.9 nM and the human constitutive proteasome and immunoproteasome with IC_{50} of 21 nM and 15 nM, respectively. The β5 subunit of the 20S proteasome was characterized as target by X-ray crystallography revealing an irreversible binding mode similar to the natural product epoxomicin. The presence of the methylmalonamide residue facilitates the stabilization of macyranone A with the active β5 subunit of the proteasome.

3.2 Introduction

The proteasome aroused the interest as therapeutic target because of its involvement in many cellular processes like cell cycle control, transcriptional regulation and apoptosis.[1] Up to date, two proteasome inhibitors have been approved by the FDA for the treatment of blood cancers such as multiple myeloma and mantle cell lymphoma: the peptide boronate bortezomib and the peptidic epoxyketone carfilzomib. Unlike bortezomib, which belongs to the first class of synthetic proteasome inhibitors, carfilzomib directly

emerged from the natural product epoxomicin.[2,3] Another promising drug candidate with natural origin is salinosporamide A which was isolated from the marine actinomycete *Salinispora tropica*.[4] Moreover, the evaluation of proteasome type specific ligands gave rise to alternative strategies to treat chronic inflammation disorders like rheumatoid arthritis and multiple sclerosis,[5] as well as infectious diseases including malaria[6] and human sleeping sickness[7]. The common target of all these compounds is the 20S core particle of the proteasome, which is formed by two outer α-rings and two inner β-rings that are composed of seven α- and β-subunits, respectively. The two β-rings form a barrel-like architecture in which the substrate proteins are cleaved to smaller peptide fragments by the catalytically active subunits β1, β2 and β5.[1] Although these subunits all share an N-terminal threonine (Thr1) as active site nucleophile, β1, β2, and β5 differ in their substrate specificity for cleavage after acidic, basic, and hydrophobic residues, hence being attributed to caspase-like (C-L), trypsin-like (T-L), and chymotrypsin-like (CT-L) activity, respectively.[8,9] Currently, proteasome inhibitors bearing an *α′,β′*-epoxyketone pharmacophore represent the benchmark for specific proteasome blockage by exploiting both nucleophiles, Thr1O$^{\gamma}$ and Thr1N. Their inhibitory mechanism results in the irreversible formation of a morpholine ring, thereby discriminating against other serine-, cysteine- or aspartate proteases.[10] Up to date, epoxyketones have mainly been isolated from actinobacteria[3,11–13] and with the recent discovery of the carmaphycins[14,15] also from cyanobacteria. The biosynthetic gene clusters of the natural peptidyl-epoxyketones eponemycin (Epn) and epoxomicin (Epx) were characterized only recently and gave the first insights into the biosynthesis of the epoxyketone warhead.[16] Herein, we report the first isolation of a new epoxyketone proteasome inhibitor, termed macyranone A (**1**), together with five structural analogues (**2**-**6**) from myxobacteria. Since the discovery of the first antibiotic myxothiazol[17] from myxobacteria in 1980, they have become famous for their capability to produce a wide range of secondary metabolites. These gliding bacteria have proven to be a rich source of structurally intriguing and bioactive natural products,[18] many of which originate from mixed polyketide-nonribosomal peptide biosynthetic pathways.[19] Their underlying biosynthetic machineries are termed non-ribosomal peptide synthases (NRPSs) and polyketide synthetases (PKSs) and represent large multimodular enzymes in which each module is responsible for the selection, activation and incorporation of one building block. Such building blocks are usually simple activated short chain dicarboxylic acids such as malonyl-CoA or natural as well as non-natural

amino acids used to generate a wide range of structurally complex and diverse natural products.

In the search for new natural products from myxobacteria, our focus lies on the identification of new scaffolds in a structure-guided isolation approach. We have been applying LC-SPE-NMR to screen crude extracts for distinct structural elements. Using this approach, we recently isolated a number of new natural products bearing unusual scaffolds including the jahnellamides[20], the hyalachelins[21] and the cystomanamides.[22]

3.3 Results and Discussion

In the course of metabolic profiling of *Cystobacter fuscus* MCy9118 using the NMR screening approach based on structural features, an unknown compound family was detected together with the previously reported cystomanamides and the known antibiotics althiomycin, roimatacene and myxochelin A and B.[22] The small molecules, termed macyranones, were isolated from a large scale shake flask cultivation of the myxobacterial strain using fractionation via size-exclusion chromatography in combination with reversed-phase HPLC for purification of the compounds. Macyranone A was the most abundant derivative with a yield of 5.2 mg isolated from 10 L cultivation volume.

3.3.1 Structure elucidation

HRESIMS of macyranone A displayed an ion peak at *m/z* 607.2987 $[M+H]^+$ (calcd for $C_{29}H_{43}N_4O_{10}$, 607.2974, Δ = 1.97 ppm), consistent with the molecular formula $C_{29}H_{42}N_4O_{10}$ containing eleven double-bond equivalents (DBE). The ^{1}H NMR spectrum of **1** in methanol-d_4 exhibited signals characteristic of a peptide including four α-proton signals at δ 4.68 (1H, dd, *J* = 9.0, 4.8 Hz), 4.34 (1H, d, *J* = 4.0 Hz) 4.37 (1H, d, *J* = 6.2 Hz) and 4.50 (1H, dd, *J* = 10.4, 4.0 Hz) along with five methyl signals at δ 0.92 (3H, d, *J* = 6.5 Hz), 0.96 (3H, d, *J* = 6.6 Hz), 1.16 (3H, d, *J* = 6.4 Hz), 1.20 (3H, d, *J* = 6.4 Hz), and 1.26 (3H, d, *J* = 7.2 Hz) (see Table S1). Additionally, a downfield pair of triplets at δ 7.28 (2H, t, *J* = 7.5 Hz) and 7.21 (1H, t, *J* = 7.5 Hz), and a doublet at δ 7.23 (2H, d, *J* = 7.5 Hz) characteristic of a phenyl group were observed. The HSQC spectrum revealed the presence of two methines ($\delta_{C\text{-}4ekLeu}$ 25.8, $\delta_{H\text{-}4ekLeu}$ 1.72 and $\delta_{C\text{-}1'ekLeu}$ 52.7, $\delta_{H\text{-}1'ekLeu}$ 3.63), two oxygenated methines ($\delta_{C\text{-}3Thr}$ 68.1, $\delta_{H\text{-}3Thr}$ 4.17, $\delta_{C\text{-}3aThr}$ 68.6, $\delta_{H\text{-}aThr}$ 4.04) and three methylene groups ($\delta_{C\text{-}2'ekLeu}$ 47.5, $\delta_{H\text{-}2'ekLeu}$ 3.02/2.98, $\delta_{C\text{-}3ekLeu}$ 39.6, $\delta_{H\text{-}3ekLeu}$ 1.55/1.50 and $\delta_{C\text{-}3Phe}$ 38.2, $\delta_{H\text{-}3Phe}$ 3.26/2.98). A detailed analysis of the 2D NMR data obtained from HSQC, HMBC and DQF-COSY

experiments established the presence of phenylalanine, two threonine residues, a leucine derived epoxyketone residue (ekLeu) and a 2-methylmalonamide residue (Mma). The latter was identified by a COSY correlation from the methine proton at δ 3.33 (H-2_{Mma}) to methyl protons at δ 1.26 (Me-4_{Mma}) and HMBC correlations from the methine proton at δ 3.33 (H-2_{Mma}) to the carbon resonances at δ 173.2 (C-1_{Mma}) and 173.3 (C-3_{Mma}). The structure of the remaining $C_8H_{14}NO_2$ was deduced as follows. A sequential spin system starting from the α-proton at δ 4.50 (H-2_{ekLeu}), and followed by two methylene protons at δ 1.55/1.50 (H-3_{ekLeu}), the methine proton at δ 1.72 (H-4_{ekLeu}) and two methyl groups at δ 0.92 and 0.96 (Me-5_{ekLeu} and Me-6_{ekLeu}), which are typical signals for a leucine residue, were deduced from TOCSY and COSY spectra. However, an HMBC correlation from H-2_{ekLeu} to the ketone resonance at δ 207.7 was indicative of a modified leucine residue. Indeed, HMBC correlations from H-3_{ekLeu} and the methylene protons at δ 3.02/2.98 (H-$2'_{ekLeu}$) to the ketone resonance as well as from the methine at δ 3.63 (H-$1'_{ekLeu}$) to the α-carbon at δ 54.1 (C-2_{ekLeu}), and a COSY correlation between H-$1'_{ekLeu}$ and H-$2'_{ekLeu}$ clearly identified the presence of a leucine derived *α'*,*β'*-epoxyketone. The complete sequence was established as Phe-Mma-Thr-aThr-ekLeu through HMBC correlations between α-protons to carbonyl carbons of adjacent residues (Figure S1). In addition, the composition of **1** was supported by ESI-MS/MS measurements. The MS^2 spectrum of the major ion peak at *m/z* 607 displayed ions at *m/z* 432 $[M+H\text{-ekLeu-}H_2O]^+$, *m/z* 331 $[M+H\text{-aThr-ekLeu-}H_2O]^+$, and *m/z* 230 $[M+H\text{-Thr-aThr-ekLeu-}H_2O]^+$. Thus, the fragmentation patterns were in agreement with the structure of **1** determined by NMR.

Macyranone A was shown to be instable in buffered aqueous systems at room temperature. An intramolecular rearrangement of the epoxide ring results in the furanone ring that is found in macyranone B (**2**). Besides macyranone A and B, four additional analogues (**3**-**6**) were isolated from the crude extract (Figure 3.1A).

The molecular formula of macyranone C (**3**) was established to be $C_{29}H_{44}N_4O_9$ (*m/z* 593.3181 $[M+H]^+$). Its NMR data (Table S2) closely resemble to that of **1**, except for the replacement of signal belonging to the epoxide group for an ethyl group (etLeu) ($\delta_{C\text{-}2'etLeu}$ 8.4, $\delta_{H\text{-}2'etLeu}$ 0.99; $\delta_{C\text{-}1'etLeu}$ 33.7, $\delta_{H\text{-}1'etLeu}$ 2.56). Macyranone D and E (**4** and **5**) turned out to be composed of $C_{29}H_{40}N_6O_9$ (*m/z* 617.2938 $[M+H]^+$) and $C_{34}H_{43}N_5O_9$ (*m/z* 666.3121 $[M+H]^+$), respectively. The NMR data (Table S2) revealed that their amino acid sequence is identical to that of **3**, only differing by the occurrence of histidine in **4** and tryptophan in **5** instead of leucine. Macyranone F displays an ion peak at *m/z* 652.2933 $[M+H]^+$

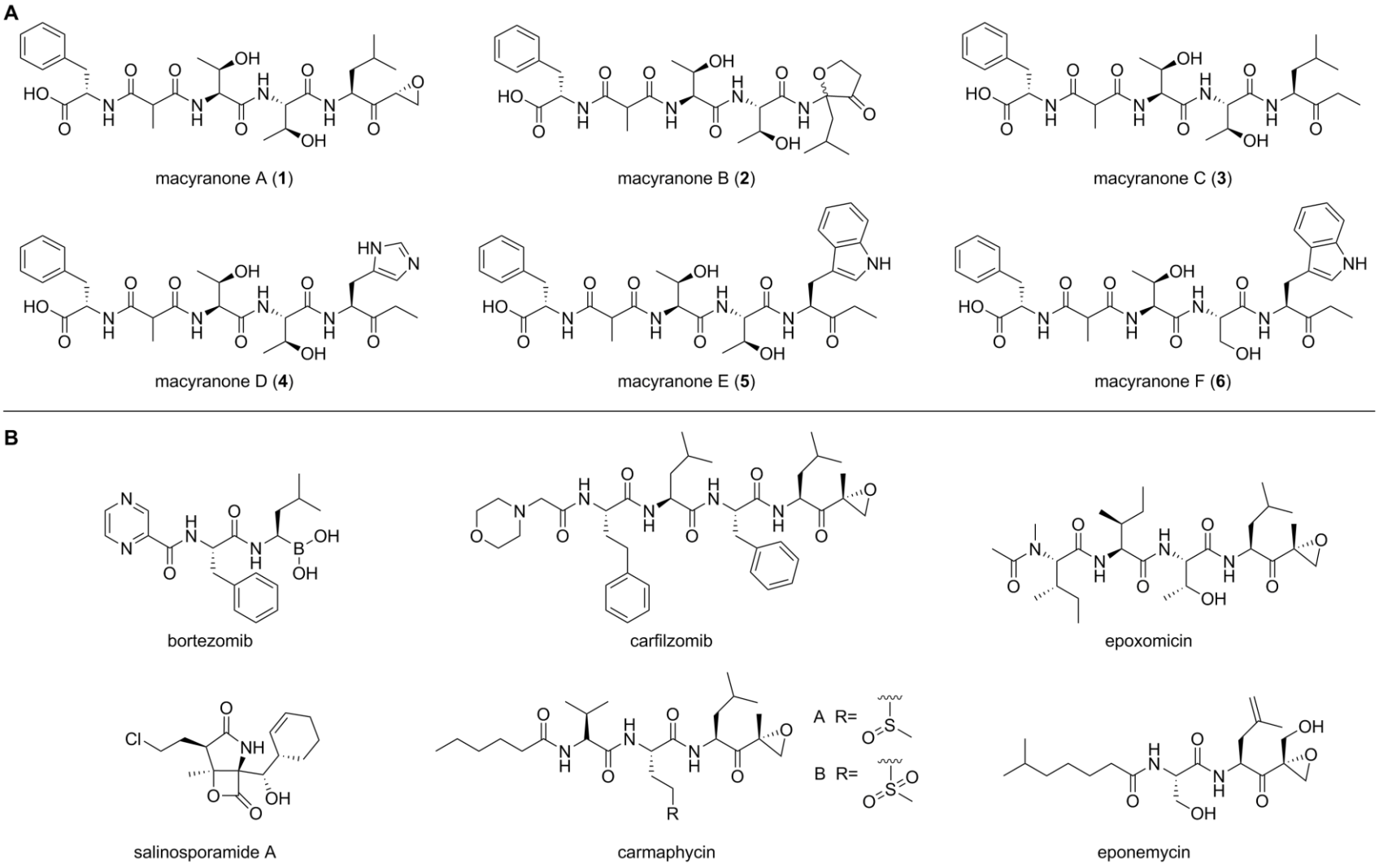


Figure 3.1. A) Structures of the macyranones A - F. B) Structures of other proteasome inhibitors.

(corresponding to a molecular formula $C_{33}H_{41}N_5O_9$) which lacks 14 Da compared to **5**. In fact, a threonine residue is replaced by a serine residue as revealed by NMR data (Table S2).

The absolute configurations of the amino acid residues in the natural products were determined by MS detected chromatographic analysis of the L- and D-FDLA (1-fluoro-2,4-dinitro-phenyl-5-L/D-leucinamide) derivatives of the acid hydrolyzate of macyranone A-F (Table S3).[23] It was found that all the amino acid residues present in the macyranones possess L-configuration. Moreover, LC-MS analysis indicated the presence of L-threonine and L-*allo*-threonine in compounds **1**-**5**. The absolute configurations of the two threonine residues were determined in **5** by *J* based configuration analysis.[24] A large coupling constant of $^3J_{H\text{-}H}$ = 6.5 Hz indicated an anti-orientation between H-2_{aThr} and H-3_{aThr}. Additionally, a ROESY spectrum recorded in DMSO-d_6 displayed a correlation from NH_{aThr} to Me-4_{aThr} (see Figure S2). Taken together, these data reveal an *erythro* configuration for C-2_{aThr} and C-3_{aThr} thereby establishing the absolute configuration of Thr and aThr as L-threonine, and aThr as L-*allo*-threonine. Identical configurations for **1**-**5** were assumed at comparable chiral centers based on the similarities of their structures and NMR data.

The *R* configuration of the epoxyketone warhead was deduced via the co-crystal structure of macyranone A in complex with the 20S proteasome as described for epoxomicin[10] which is illustrated by the structural superposition in Figure 3.4. For macyranone B, we observe that the conversion of **1** is not stereospecific since it yields a racemic mixture of **2**. Further, the 2-methylmalonamide residue present in all macyranones undergoes a fast proton exchange which prevented the assignment of its configuration.

3.3.2 Identification of the *myn* biosynthetic gene cluster

The chemical structures of the macyranones suggest that the compounds are products of a polyketide synthase (PKS) and non-ribosomal peptide synthetase (NRPS) hybrid. The building blocks of the macyranones are expected to include L-phenylalanine, methylmalonyl-CoA, L-threonine, L-*allo*-threonine or L-serine, a variable amino acid (leucine, histidine or tryptophan), and a malonyl-CoA unit. Using the bioinformatic prediction tool antiSMASH 2.0,[25] a candidate PKS/NRPS biosynthetic gene cluster (*myn*) was identified. The predicted gene cluster consists of eight open reading frames (ORFs) with an overall GC content of 70.0%. Analysis of the catalytic domains led to the

prediction of three NRPS modules and one PKS module (Figure 3.2). The adenylation domain specificity in the three NRPS modules matched the distinct pattern of building blocks for the core amino acid motif Thr-Thr-Leu (Table S5). Genes that did not yield a prediction for PKS or NRPS domains were analyzed via the BLAST algorithm[26] (blastx) against the non-redundant NCBI database (Table 3.1, Table S4).

Feeding experiments with labeled sodium propionate [1-^{13}C] and sodium acetate [$^{13}C_2$] depicted the incorporation of methylmalonyl-CoA and malonyl-CoA, respectively. Incorporation of leucine-*d3* confirmed the biosynthetic origin of the ekLeu residue as leucine-derived (Figure S3).

Intriguingly, the identified building blocks do not exhibit the typical composition reflecting a sequential N- to C-terminal biosynthesis. Based on the arrangement and presence of modules responsible for the incorporation of building blocks, it seems more likely that the assembly line starts with threonine. Subsequent attachment of phenylalanine and methylmalonyl-CoA occur via an alternative route. MynA shows sequence similarity to a family of asparagine synthetases with the closest similarity to PdmN from the pradimicin pathway[27] (41.6 % identity), to RubR from the rubromycin pathway (40.3 % identity), and to FdmV from the fredericamycin pathway[28] (39.5 % identity). Alignment of the three enzymes to the typical asparagine synthetase B-like amide synthetase (AS B) from *E. coli* (20.5 % identity) shows that all mentioned enzymes contain the conserved domains present in AS B: the N-terminal glutaminase domain and the C-terminal amide synthetase domain (Figure S4).[29] The glutaminase domain is responsible for the hydrolysis of L-Glu to supply a free ammonia group and the amide synthase domain catalyzes the conversion of aspartate to asparagine via a β-aspartyl-AMP intermediate.[30,31] For FdmV, it was proven that it catalyzes lactam ring formation with l-Glu, l-Asn or free NH_3 as nitrogen source.[28] All class II glutamine amidotransferases contain a conserved N-terminal cysteine residue (Cys2) that was shown to act as active-site nucleophile for the hydrolysis of glutamine to glutamate and ammonia.[32] This residue is retained in AS B and FdmV but is mutated to serine in MynA, PdmN, and RubR. The mutation leads to a loss of the glutaminase activity in the three enzymes whereas the active site residues responsible for the coordination of the AMP intermediate are preserved. RubR could not be assigned to a function but PdmN was suspected to be responsible for transferring an amino acid into the polyketide intermediate of pradimicin.[27] Therefore, the amide synthetase activity seems to be

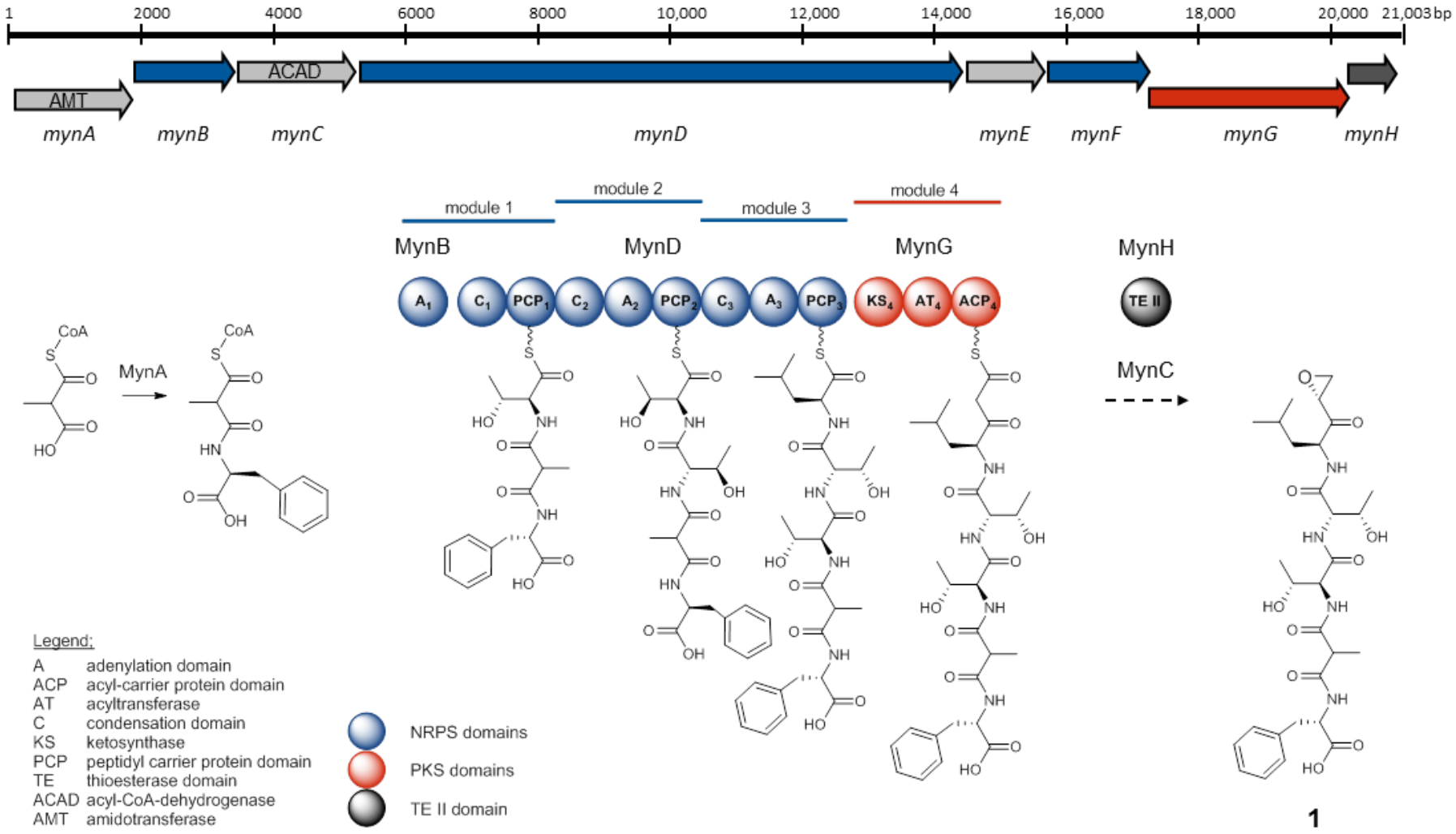


Figure 3.2. Gene cluster of the macyranones in *Cystobacter fuscus* MCy9118 and the deduced biosynthetic pathway of **1**.

retained in MynA and PdmN by using other substrates as nitrogen donor (L-Phe and D-Ala, respectively). Consequently, MynA could be accountable for amide bond formation between the methylmalonic acid residue and phenylalanine via an AMP intermediate with phenylalanine acting as amino donor. The first adenylation domain found in the identified *myn* biosynthetic gene cluster is a stand-alone adenylation domain with threonine substrate specificity predicted for *mynB*. The coupling of the methylmalonate-phenylalanine and the first threonine could be catalyzed by the condensation domain in the first module of the peptide synthetase as discussed for lipopeptide biosynthesis.[33] Phylogenetic analysis indeed demonstrated a highly unusual nature of this C domain since it does not group with other known types like $^{L}C_{L}$ type, $^{D}C_{L}$ type, starter, hybrid, cyclization, epimerization or dual function condensation domains (Figure S5).

Table 3.1. Deduced functions of open reading frames in the macyranone biosynthetic gene cluster

Protein	Proposed function
MynA	amidotransferase
MynB	stand-alone $A1_{Thr}$ domain
MynC	acyl-CoA dehydrogenase
MynD	NRPS (C-PCP, C-$A2_{Thr}$-PCP, C-$A3_{Leu}$-PCP)
MynE	transporter
MynF	stand-alone $A4_{His}$ domain
MynG	PKS (KS-AT-ACP)
MynH	thioesterase type II

Alternatively, the stand-alone adenylation domain encoded by *mynB* is a possible candidate for the amide bond formation as shown for *orf19,* which encodes an amidating stand-alone A domain in streptothricin biosynthesis.[34] Comparison of the adenylation domains did not provide valuable information about the possible function of MynB.

The biosynthesis continues with MynD. Besides the condensation domain and the PCP from the first module, the protein contains two additional modules where the adenylation domains show specificity for threonine and leucine, respectively. These two modules correspond to the sequence of the amino acid residues in the main product macyranone A. *MynF* encodes a second stand-alone adenylation domain with histidine substrate specificity (Table S5).

Since some of the macyranones contain a leucine residue (macyranone A to C) and some contain a histidine (D) or tryptophan residue (E and F), MynF might act as an alternative

adenylation domain for module 3. The PKS which is responsible for the attachment of the C terminal malonyl-CoA unit is localized on *mynG*. Although active site analysis of the respective AT domain predicted the incorporation of an unusual substrate (Table S6), feeding experiments with sodium acetate [$^{13}C_2$] indicated the incorporation of malonyl-CoA (Figure S3). *mynE* seems to encode a transporter whereas *mynH* encodes a thioesterase type II. The latter protein exhibits some unusual features as revealed by phylogenetic analysis (Figure S7); in general, type I thioesterase domains remove the final product through hydrolysis and are usually integrated into the final module of the assembly line complex, whereas type II thioesterases are discrete proteins that can remove intermediates from any module in the complex. The TE shows similarity to a group of thioesterase domains characterized on the basis of the ajudazol TE (AjuTE) by Buntin et al. in 2010.[35] The AjuTE is integrated into a PKS module but phylogenetic analysis associated it with type II thioesterases. Therefore, the authors suggested that the TE was once a stand-alone type II TE domain, but was fused genetically to the end of the ajudazol assembly line and acts there as hydrolytic type I TE. The AjuTE is expected to be involved in isochromanone ring formation. As a consequence of this similarity, we believe that the type II TE domain present at the terminus of the *myn* gene cluster plays a crucial role in the release of the substrate from the assembly line and therefore in the formation of the epoxyketone unit. In comparison to the biosynthesis of the epoxyketones epoxomicin and eponemycin (**12**)[16], only eponemycin biosynthesis involves a similar free-standing TE (EpnB) that also shows 37.5 % similarity to MynH on the protein level. In contrast to the *myn* cluster, both clusters contain an integrated type I TE domain downstream of the final PKS module. Comparing the recently identified biosynthetic gene clusters of epoxomicin (*epx*) and eponemycin (*epn*) to the identified cluster of macyranone, it is striking that all three clusters encode an acyl-CoA dehydrogenase (*mynC*, *epxF*, and *epnF*), indicating a major role for this protein in the biosynthesis of the epoxyketone warhead. For epoxomicin and eponemycin, the authors suggested[16] a release of the intermediates as carboxylic acids by the C-terminal thioesterase domain with subsequent reduction and epoxidation accomplished by the acyl-CoA dehydrogenase (ACAD EpxF/EpnF) and a cytochrome P450 (CYP) monooxygenase (EpxC/EpnI), respectively. Alternatively, a mechanism that involves a putative C-methyltransferase domain found in both clusters with subsequent decarboxylation that initiates the epoxide formation by the ACAD and CYP P450 was taken into consideration. However, in contrast to the *epx* and *epn* clusters,

the gene cluster of macyranone does not contain a gene that encodes for a CYP P450 enzyme. Additionally, no C-methyltransferase domain was identified in the final PKS module (as reasoned by the unsubstituted epoxide ring of **1**) compared to epoxomicin and eponemycin.

Based on our findings, we take another mechanism for epoxyketone formation into consideration: MynC as well as EpxF and EpnF resemble very-long chain acyl-CoA dehydrogenases (VLCAD) isolated from mammalian sources. VLCAD are responsible for α,β-dehydrogenation of long-chain fatty acid acyl-CoA conjugates.[36] Members of the ACAD superfamily belong to the flavoproteins and were shown to catalyze various redox reactions including α,β-dehydrogenation, oxidation of nitroalkanes[37] as well as desulfurization[38]. Acyl-CoA dehydrogenases in general act on CoA esters rather than on free fatty acids.[39] Thus, we suggest as alternative an acyl-carrier-bound mechanism involving the acyl-CoA dehydrogenase MynC. A reductive mechanism similar to such catalyzed by reductase domains[40,41] in combination with simultaneous oxidation of the α-position of the ketone catalyzed by MynC could lead to a release of the compound as epoxyketone. Chain release may be spontaneous or involve the free-standing thioesterase MynH resulting in the epoxide (**1**) and, to a smaller extent, furanone (**2**) ring formation. Full reduction of the product might lead to macyranone C-F. Obviously, formation of the epoxyketone requires further analysis by biochemical studies. Figure 3.2 illustrates the genetic organization of the macyranone PKS/NRPS hybrid biosynthetic gene cluster as determined by *in silico* analysis including a biosynthetic hypothesis.

3.3.3 Gene disruption in MCy9118

To confirm that the identified candidate gene cluster is responsible for macyranone production, a single crossover knockout strategy was established. Gene disruption in MCy9118 via single crossover homologous recombination was accomplished according to a previously described protocol.[22] HPLC analysis of the resulting *mynA*, *mynC* and *mynD* mutants revealed that inactivation of all three genes abolished the production of the macyranones (Figure 3.3). These results confirm that the identified candidate biosynthetic gene cluster is indeed responsible for macyranone production.

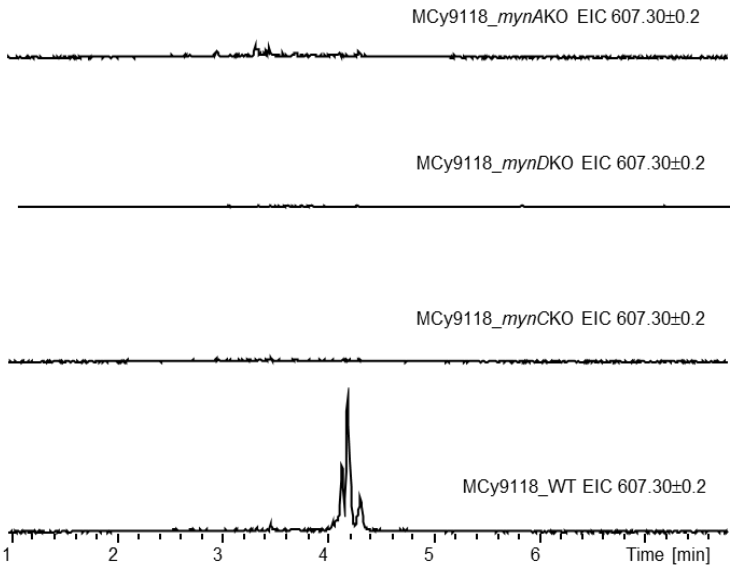


Figure 3.3. Loss of the macyranone A and B production in the MCy9118 knockout mutants *mynA*KO, *mynD*KO and *mynC*KO as detected by comparative HPLC-MS analysis of MCy9118 wildtype and mutant strains. An extracted ion chromatogram for the $[M+H]^+$ signal of macyranone A and B is shown.

3.3.4 Biological activity

Macyranone A inhibits the CT-L activity of the yeast 20S proteasome with an IC_{50} of 5.9 nM. It does not show an effect on the C-L activity whereas it modulates the T-L activity with an IC_{50} of 400 nM. Macyranone B, D, and E did not block any active site up to a tested concentration of 10 µM. Next, **1** was tested against the human constitutive proteasome (cCP) and immunoproteasome (iCP). Both subunits were simultaneously inhibited with an IC_{50} of 21 nM and 15 nM, respectively. However surprisingly, macyranone A does not show the expected high cytotoxicity against mammalian cell lines such as the human colon carcinoma HCT-116 cell line, the human leukemia cell lines THP-1 and HL-60, and the rat myoblast cell line L-6. Moreover, macyranone A shows a potent antiparasitic activity against *Trypanosoma brucei rhodesiense*, the causative agent of the African sleeping sickness and against extracellular grown amastigotes of *Leishmania donovani* causing leishmaniasis (Table 3.2). The activity against *L. donovani* is lost in the intracellular amastigote assay tested up to a concentration of 30 µM. Thus, due to the close similarity of the proteasomes in all eukaryotic cells, we expect that the selectivity towards the parasitic cells is a transport phenomenon or a metabolism-related specificity.[42] Macyranone A shows a poor stability in buffered aqueous systems that could be an explanation for the discrepancy between *in vitro* inhibition of the 20S proteasome and *in vivo* activity against eukaryotic cells.

Table 3.2: Biological activity of macyranone A and epoxomicin

Bioactivity assay	Macyranone A IC_{50} [µM]	Epoxomicin IC_{50} [µM]
Constitutive proteasome (Human, 10µg/mL), β5 subunit	0.021	
Immunoproteasome (Human, 10µg/mL), β5 subunit	0.015	
20S Proteasome (Yeast, 0.5 nM), β5 subunit	0.006	
Trypanosoma brucei rhodesiense	1.55	
Leishmania donovani	0.22	
Plasmodium falciparum	41.6	
HCT-116	29.2	0.0028
THP-1	38.3	0.0116
HL-60	21.9	0.0042
L-6	> 100	

3.3.5 Crystal structure of the 20S proteasome:macyranone A complex

To characterize the binding mode of macyranone A, we determined the crystal structure of the yeast 20S proteasome in complex with macyranone A (2.8 Å resolution, R_{free} = 21.9 %, PDB ID 5AHJ). Consistent with the proteasome:epoxomicin complex structure, macyranone A binds to the catalytically active Thr1 of all active subunits β1, β2 and β5 (Figure 3.4).[10] The mode of action comprises the nucleophilic attack of both nucleophiles present at the active site, Thr1O^{γ} and Thr1N, resulting in an irreversible morpholine ring adduct formation. Due to the bivalent binding to the N-terminal Thr1, macyranone A fulfills the requirement to selectively block the small class of N-terminal threonine nucleophile (Ntn) hydrolases, potentially preventing any off-target activity as shown in the case of the clinically relevant carfilzomib.[43] Furthermore, macyranone A is stabilized by antiparallel β-sheet formation of the peptide backbone in the substrate binding channel combined with distinct stabilization in the substrate specificity pockets. In this regard, the unusual 2-methylmalonamide unit causes a shift at the end of the peptide scaffold, which allows the C-terminal phenylalanine to hydrophobically interact with Ala22 of subunit β5 to achieve further stabilization. Moreover, the X-ray crystallographic studies complement the results of the NMR spectroscopic analysis supporting the peptide sequence as well as the assigned stereochemistry of macyranone A.

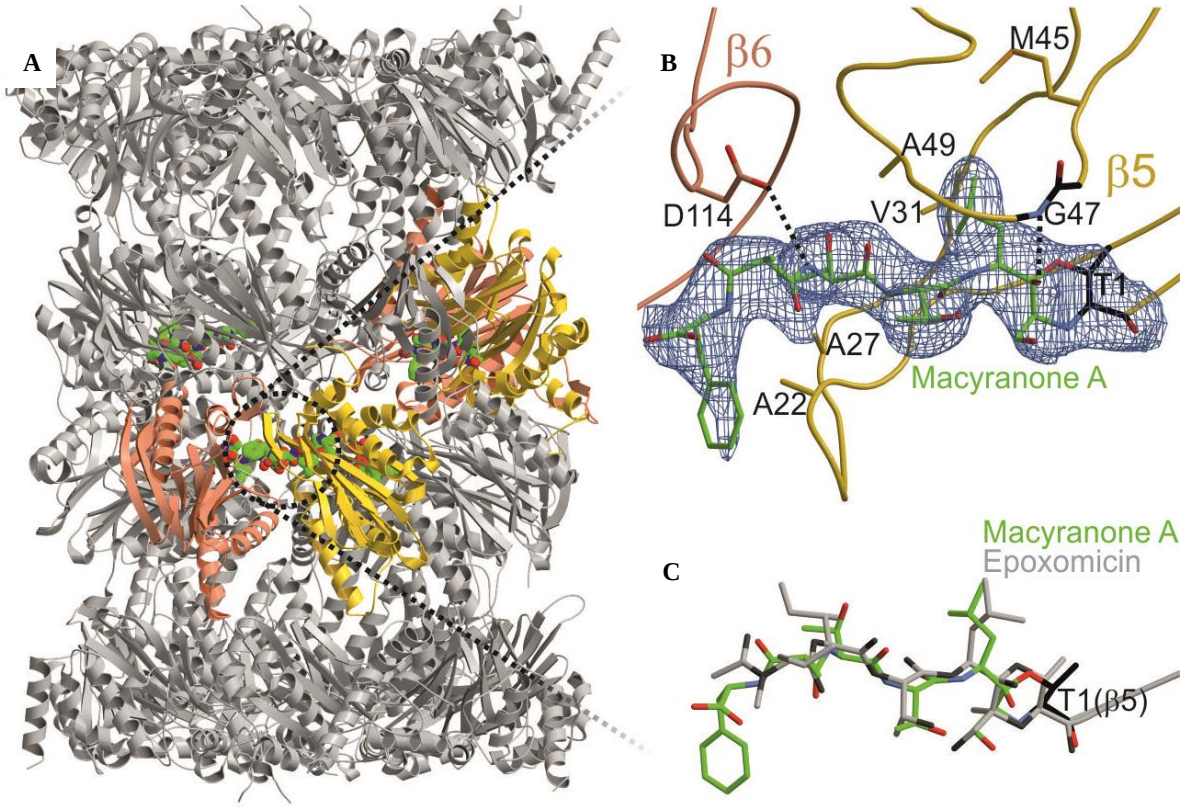


Figure 3.4. A: X-ray crystallographic binding analysis of macyranone A bound to the yeast 20S proteasome (yCP, PDB ID 5AHJ). B: The $2F_o–F_c$ electron density map (cyan mesh, contoured at 1σ) shows bound macyranone A (green) at Thr1 (black) and the oxyanion hole formed by G47 (black). Hydrogen-bonding of the ligand to subunit β5 (yellow) and subunit β6 (pale orange) is indicated by black dashed lines. C: Structural superposition of bound macyranone A (green) and epoxomicin (grey) at the Thr1 of subunit β5 of yCP.

3.4 Conclusions

The macyranones represent a new family of natural products isolated from myxobacteria. Their core structure consists of three amino acid residues, an unprecedented 2-methylmalonamide moiety, and an α-amino ketone fragment formed by condensation of an amino acid and an acetate-derived ethyl unit. In particular, macyranone A bears the powerful epoxyketone warhead found in the group of *α′,β′*-epoxyketones.

The biosynthesis of the macyranones deviates widely from the textbook logic of PKS/NRPS hybrids. During macyranone formation, the amidotransferase and the unconventional condensation domain are involved in two conspicuous amide bond formations comprising unprecedented biochemical reactions. Additionally, a acetyl-CoA dehydrogenase homologue and possibly also the type II thioesterase are involved in the formation of the epoxyketone residue representing highly unusual features where *in silico* analysis can currently only provide an idea of the real biosynthetic mechanism. The

identification of the biosynthetic gene cluster, detailed *in silico* analysis and the confirmation of the gene cluster via gene deletion thus set the stage for further detailed biochemical studies.

Moreover, macyranone A shows a potent activity *in vitro* against the β5 subunit of the human constitutive proteasome as well as the immunoproteasome. Additionally, **1** is active *in vivo* against the parasites *Trypanosoma brucei rhodesiense* and *Leishmania donovani*. With a surprisingly low cytotoxicity to mammalian cell lines, macyranone A exhibits a potential therapeutic window in contrast to other proteasome inhibitors that translate their 20S proteasome inhibition into high cytotoxicity. Therefore, the highly potent natural product may contribute to the development of new anti-parasitic agents for the treatment of infectious diseases like the African sleeping sickness.

X-ray analysis of macyranone A in complex with the yeast 20S proteasome confirms the covalent binding mechanism of the epoxyketones at the β5 subunit. Furthermore, the unusual 2-methylmalonamide residue was identified to further stabilize the ligand in the substrate binding channel allowing a hydrophobic interaction between the C-terminal phenylalanine and subunit β5.

In closing, these results contribute to our understanding of the selective blockage of the ubiquitous eukaryotic 20S proteasome and could be of help to the development of new anti-parasitic agents. Also, the identification of the biosynthetic gene cluster, detailed *in-silico* analysis and the confirmation of the gene cluster via gene deletion set the stage for further studies in the formation of the epoxyketone functionality. Finally, the potential development of macyranone A as antiparasitic drug provides another example of the importance of myxobacteria in drug discovery programs.

3.5 Experimental Section

3.5.1 General experimental procedures

NMR spectra were recorded in methanol-d_4 on a Bruker Ascend 700 spectrometer with a 5 mm TXI cryoprobe (^{1}H at 700 MHz, ^{13}C at 175 MHz) or a Bruker Avance III 500 spectrometer with a 5 mm TXI cryoprobe (^{1}H at 500 MHz, ^{13}C at 125 MHz). DQFCOSY, HOHAHA, HSQC, HMBC, and ROESY experiments were recorded using standard pulse programs. HSQC experiments were optimized for $^{1}J_{C\text{-}H}$ = 145 Hz, and HMBC spectra were optimized for $^{2,3}J_{C\text{-}H}$ = 6 Hz. The samples were dissolved in methanol-d_4 and the chemical

shifts of the solvent signals at 3.31 ppm (δ_H) and 49.15 ppm (δ_C) were considered as internal standard (reference signal). The observed chemical shift (δ) values were given in ppm and the coupling constants (*J*) in Hz. 5mm Shigemi tubes (Shigemi Inc., Allison Park, PA 15101, USA) were used to increase sensitivity. A Dionex Ultimate 3000 RSLC system was used for LC-HRMS measurements with a Waters BEHC18, 100 x 2.1 mm, 1.7 µm dp column. Separation of 2 µl sample was achieved by a linear gradient with (A) H_2O + 0.1 % FA to (B) ACN + 0.1 % FA at a flow rate of 600 µl/min and 45 °C. The gradient was initiated by a 0.5 min isocratic step at 5 % B, followed by an increase to 95% B in 18 min to end up with a 2 min step at 95 % B before reequilibration with initial conditions. UV spectra were recorded by a DAD in the range from 200 to 600 nm. The LC flow was split to 75 µl/min before entering the maXis HR-ToF mass spectrometer (Bruker Daltonics, Bremen, Germany) using the standard ESI source. Mass spectra were acquired in centroid mode ranging from 150 - 2000 *m/z* at 2 Hz scan speed.

3.5.2 Cultivation of strain MCy9118

The myxobacterial strain *Cystobacter fuscus* MCy9118 was isolated and cultivated using previously reported methods.[22] For the production of the macyranones, the preculture was used to inoculate 300 ml shake flasks containing 50 mL VY/2S-DY medium (0.4 % dry yeast, 0.05 % $CaCl_2$ x 2 H_2O, 1.0 % soluble starch, 5 mM HEPES adjusted to pH 7.0 with 10 N KOH) that was incubated at 30 °C for 8 h. Always 20 mL of the cultures were used to inoculate six times 2 L VY/2S-DY medium containing 8 mg/mL Fe-EDTA in 5 L shake flasks (shaking rate 160rpm; 30°C). 2 % XAD adsorber resin (Amberlite XAD-7, Sigma) was added to the shake flasks on day 2 of the cultivation. The cultures were incubated for 7 days. Cultures were harvested by sieving the culture to receive the XAD.

3.5.3 Isolation procedure

XAD was washed with *dd*H_2O to remove salts prior to freeze drying. The XAD was extracted exhaustively with MeOH to obtain the crude extract. Sephadex fractionation was performed using a GE Healthcare SR 25/100 column, sephadex LH 20 was used for separation and methanol as mobile phase. The fractions containing the respective peptides were combined and the solvent was evaporated. The compounds were separated by preparative HPLC. A Waters Xbridge C18, 150x19 mm, 5 µm dp column was used and a linear gradient from 5 to 95% B in 30min with (A) H2O + 0.1 % FA and (B)

MeOH + 0.1 % FA at a flow rate of 25 mL/min at room temperature. Each injection consisted of up to 150 mg of the combined fractions dissolved in 500 µl MeOH. Fractions were collected by time in test tubes. The compounds were repurified by semipreparative HPLC at 30°C eluting with a linear gradient and column according to the compound description and a DAD at 220nm.

Macyranone A (**1**): ^{1}H and ^{13}C NMR data, see Table S1; HRESIMS *m/z* 607.2967 $[M+H]^+$ corresponding to a molecular formula $C_{29}H_{42}N_4O_{10}$ (calcd for $C_{29}H_{43}N_4O_{10}$, 607.2974, Δ = 1.2 ppm) Isolation: Phenomenex Synergi™ 4 µm Fusion-RP 80 Å, LC Column 250 x 10 mm; t_R = 39.5 min; flow rate = 2.5 ml/min; 0-1min 10% B, 1-45min 10-90% B; A: H_2O + 0.1% FA, B: MeOH + 0.1% FA

Macyranone B (**2**): ^{1}H and ^{13}C NMR data, see Table S2; HRESIMS *m/z* 607.2963 $[M+H]^+$ corresponding to a molecular formula $C_{29}H_{42}N_4O_{10}$ (calcd for $C_{29}H_{43}N_4O_{10}$, 607.2974, Δ = 1.8 ppm) Isolation: Jupiter® 4 µm Proteo 90 Å, LC Column 250 x 10 mm; t_R = 43.9 min; flow rate = 2.5 ml/min; 0-7min 34% B, 7-47min 34-75% B; A: H_2O + 0.1% FA, B: MeOH + 0.1% FA

Macyranone C (**3**): ^{1}H and ^{13}C NMR data, see Table S2; HRESIMS *m/z* 593.3181 $[M+H]^+$ corresponding to a molecular formula $C_{29}H_{44}N_4O_9$ (calcd for $C_{29}H_{45}N_4O_9$, 593.3181, Δ = 0 ppm) Isolation: Jupiter® 4 µm Proteo 90 Å, LC Column 250 x 10 mm; t_R = 25.1 min; flow rate = 2.5 ml/min; 0-1min 10% B, 1-45min 10-90% B; A: H_2O + 0.1% FA, B: ACN + 0.1% FA

Macyranone D (**4**): ^{1}H and ^{13}C NMR data, see Table S2; HRESIMS *m/z* 617.2938 $[M+H]^+$ corresponding to a molecular formula $C_{29}H_{40}N_6O_9$ (calcd for $C_{29}H_{41}N_6O_9$, 617.2930, Δ = 1.3 ppm) Isolation: Synergi™ 4 µm Polar-RP 80 Å, LC Column 250 x 10 mm; t_R = 23.3 min; flow rate = 3.5 ml/min; 0-1min 10% B, 1-41min 10-90% B; A: H_2O + 0.1% FA, B: MeOH + 0.1% FA

Macyranone E (**5**): ^{1}H and ^{13}C NMR data, see Table S2; HRESIMS *m/z* 666.3121 $[M+H]^+$ corresponding to a molecular formula $C_{34}H_{43}N_5O_9$ (calcd for $C_{34}H_{44}N_5O_9$, 666.3134, Δ = 2.1 ppm) Isolation: Jupiter® 4 µm Proteo 90 Å, LC Column 250 x 10 mm; t_R = 40.7 min; flow rate = 2.5 ml/min; 0-1min 10% B, 1-45min 10-90% B; A: H_2O + 0.1% FA, B:ACN + 0.1% FA

Macyranone F (**6**): ^{1}H and ^{13}C NMR data, see Table S2; HRESIMS *m/z* 652.2933 $[M+H]^+$ corresponding to a molecular formula $C_{33}H_{41}N_5O_9$ (calcd for $C_{33}H_{42}N_5O_9$, 652.2977, Δ = 6.7 ppm) Isolation: Jupiter® 4 µm Proteo 90 Å, LC Column 250 x 10 mm; t_R = 39.6 min; flow

rate = 2.5 ml/min; 0-1min 10% B, 1-45min 10-90% B; A: H_2O + 0.1% FA, B: ACN + 0.1% FA

3.5.4 Advanced Marfey's method[44]

Approximately 0.2 mg of the compound were hydrolyzed with 6 N HCl (0.8 mL) and shaken at 90°C for 16 h in Eppendorf tubes. Samples were dried under vacuum and dissolved in H_2O (100 µL). They were split into two 50 µL aliquots and 1 N $NaHCO_3$ (20 µL) and 1% 1-fluoro-2,4-dinitrophenyl-5-leucine-amide (L-FDLA or D-FDLA solution in acetone, 100 µL) were added respectively. The mixtures were heated to 40°C for 40 min, cooled down to RT and the reaction was quenched by addition of 2 N HCl (20 µL). Solvents were evaporated under nitrogen and residues were redissolved in 1 mL CH_3CN and subsequently analyzed by LCMS. All measurements were performed on a Dionex Ultimate 3000 RSLC system using a Waters BEH C18, 100 x 2.1 mm, 1.7 µm column by injection of 1 µl sample. Separation was achieved by a gradient using (A) H2O + 0.1% FA to (B) ACN + 0.1% FA at a flow rate of 550 µl/min and 45 °C. The gradient was as follows: starting at 5 % B to increase to 10 % B in 1 min, from 1 to 15 min increase to 35 % B, from 15 to 22 S5 min increase to 50% B, from 22 to 25 min increase to 80 % B. After a 1 min hold at 80 % B the system was reequilibrated with initial conditions for 5 minutes. UV data was acquired at 340 nm and MS-detection was performed simultaneously. Coupling the HPLC to the MS was supported by an Advion Triversa Nanomate nanoESI system attached to a Thermo Fisher Orbitrap. LC flow is split to 500 nl/min before entering the ion source. Mass spectra were acquired in centroid mode ranging from 150 to 1000 *m/z* at a resolution of R = 30000. Results can be found in Table S3.

3.5.5 Sequence analysis

Routine DNA *in silico* sequence analysis was carried out using the Geneious® 6.1.6 software. Annotation of the catalytic domains and characterization of the A domain substrate specificity was performed using the tools antiSMASH 2.0[25], Pfam[45], NRPSpredictor2[46], PKS/NRPS Analysis[47] and NaPDoS[48].

3.5.6 General procedures for DNA manipulation and PCR

Chromosomal DNA from MCy9118 was prepared using the Puregene ® Core Kit A (Qiagen) according to the manufacturer's instructions. PCR was carried out using Phusion High Fidelity DNA Polymerase (Thermo Scientific) according to the

manufacturer's protocol with addition of 5% DMSO to the reaction mixture. For the amplification, conditions were: Initial denaturation step at 98°C for 2 min, denaturation for 10 sec at 98°C, annealing for 15 sec at 60 to 70°C according to the primer, extension at 72°C for 2min (30 cycles) and a final extension step at 72°C for 10 min. PCR products were purified via agarose gel electrophoresis and subsequent extraction from the gel using the peqGOLD Gel Extraction Kit (Peqlab). Ligation of DNA fragments with sticky ends was done using T4 DNA Ligase (Thermo Scientific) according to the manufacturer protocol using a 4:1 molar ratio of insert and vector. Restriction digest was done using enzymes from Thermo Scientific according to the manufacturer's protocol and recommendations regarding double digest. For pTOPO cloning, the TOPO TA Cloning Kit with pCR® 2.1 TOPO vector (Invitrogen) was used. Plasmids were introduced into chemically competent *E. coli* strains via heat shock transformation and into *C. fuscus* MCy9118 via electroporation at 650 V and 400 Ω. *E. coli* Clones containing a plasmid were selected on LB agar (1.5% agar) supplemented with 50 mg/L Kanamycin-sulfate (Kan) at 37°C. Mutant clones of MCy9118 were selected on M-agar (1.5% agar) containing 50 mg/L Kan.

3.5.7 Single crossover homologous recombination strategy

Gene disruption in MCy9118 via single crossover homologous recombination was accomplished according to a previously described protocol[22]. A homologous fragment of each target gene with a size around 1 kb was amplified from genomic DNA using the respective oligonucleotides listed in Table S7 that additionally inserted stop codons on each side of the fragment. The fragment was ligated into the vector pCR®II-TOPO®, resulting in the respective pTOPO derivative for gene disruption. The plasmids were transformed into *E. coli* strain SCS110 for amplification of the plasmid without Dam or Dcm methylation. The plasmid was recovered from the strain by the alkaline lysis method, and after restriction analysis it was introduced into MCy9118 by electroporation. Single clones were used to inoculate liquid medium (M medium + 50 mg/L Kan) for secondary metabolite analysis and isolation of genomic DNA for genetic verification by PCR. Subsequent HPLC analysis of the resulting *mynA*, *mynC* and *mynD* mutants revealed that inactivation of the three genes abolished production of the macyranones.

To genetically verify the mutants, two PCR reactions were performed for each mutant (see Figure S8). The first PCR was performed with primers that bind outside of the region of plasmid insertion. A product is only generated if the gene is not disrupted by the

plasmid back-bone so that the mutants should not show a product in contrast to the compared wild type. The second PCR was performed with one primer that binds outside of the region of plasmid insertion and one that binds the plasmid backbone. A product is only generated if the plasmid was inserted in the right position so that the mutants should show a product in contrast to the compared wild type were no product is expected. Mutants fulfilling both conditions were considered as genetically verified.

3.5.8 Activity against *Trypanosoma brucei rhodesiense* STIB900

This stock was isolated in 1982 from a human patient in Tanzania and after several mouse passages cloned and adapted to axenic culture conditions.[49] Minimum Essential Medium (50 µl) supplemented with 25 mM HEPES, 1 g/L additional glucose, 1 % MEM non-essential amino acids (100x), 0.2 mM 2-mercaptoethanol, 1 mM Na-pyruvate and 15 % heat inactivated horse serum was added to each well of a 96-well microtiter plate. Serial drug dilutions of eleven 3-fold dilution steps covering a range from 100 to 0.002 µg/ml were prepared. Then $4x10^3$ bloodstream forms of *T. b. rhodesiense* STIB 900 in 50 µl was added to each well and the plate incubated at 37 °C under a 5 % CO_2 atmosphere for 70 h. 10 µl Alamar Blue (resazurin, 12.5 mg in 100 mL double-distilled water) was then added to each well and incubation continued for a further 2–4 h.[50] Then the plates were read with a Spectramax Gemini XS microplate fluorometer (Molecular Devices Cooperation, Sunnyvale, CA, USA) using an excitation wave length of 536 nm and an emission wave length of 588 nm. The IC_{50} values were calculated by linear regression[51] from the sigmoidal dose inhibition curves using SoftmaxPro software (Molecular Devices Cooperation, Sunnyvale, CA, USA). Melarsoprol (Arsobal Sanofi-Aventis, received from WHO) is used as control.

3.5.9 Activity against *Leishmania donovani* axenic amastigotes

Amastigotes of *L. donovani* strain MHOM/ET/67/L82 were grown in axenic culture at 37 °C in SM medium[52] at pH 5.4 supplemented with 10 % heat-inactivated fetal bovine serum under an atmosphere of 5 % CO_2 in air. 100 µL of culture medium with 10^5 amastigotes from axenic culture with or without a serial drug dilution were seeded in 96-well microtiter plates. Serial drug dilutions of eleven 3-fold dilution steps covering a range from 90 to 0.002 µg/mL were prepared. After 70 h of incubation the plates were inspected under an inverted microscope to assure growth of the controls and sterile

conditions. 10 µl of Alamar Blue (12.5 mg resazurin dissolved in 100 mL distilled water)[53] were then added to each well and the plates incubated for another 2 h. Then the plates were read with a Spectramax Gemini XS microplate fluorometer (Molecular Devices Cooperation, Sunnyvale, CA, USA) using an excitation wave length of 536 nm and an emission wave length of 588 nm. Data were analyzed using the software Softmax Pro (Molecular Devices Cooperation, Sunnyvale, CA, USA). Decrease of fluorescence (= inhibition) was expressed as percentage of the fluorescence of control cultures and plotted against the drug concentrations. From the sigmoidal inhibition curves the IC_{50} values were calculated.

3.5.10 Activity against *L. donovani* intracellular amastigotes: macrophage assay[54]

Mouse peritoneal macrophages (4 x 10^4 in 100 µl RPMI 1640 medium with 10% heat-inactivated FBS) were seeded into wells of Lab-tek 16-chamber slides. After 24 hrs 1.2 x 10^5 amastigote *Leishmania donovani* in 100 µl were added. The amastigotes were taken from an axenic amastigote culture grown at pH 5.4. Four hrs later the medium containing free amastigote forms was removed and replaced by fresh medium. Next day the medium was replaced by medium containing different compound dilutions. Parasite growth in the presence of the drug was compared to control wells. After 96 hours of incubation the medium was removed and the slides fixed with methanol for 10 min followed by a staining with a 10% Giemsa solution. Infected and non-infected macrophages were counted for the control cultures and the ones exposed to the serial drug dilutions. The infection rates were determined. The results were expressed as % reduction in parasite burden compared to control wells, and the IC_{50} calculated by linear regression analysis.

3.5.11 Activity against *P. falciparum*

In vitro activity against erythrocytic stages of *P. falciparum* was determined using a ^{3}H-hypoxanthine incorporation assay[55,56], using the drug sensitive NF54 strain[57] and the standard drug chloroquine (Sigma C6628). Compounds were dissolved in DMSO at 10 mg/mL and added to parasite cultures incubated in RPMI 1640 medium without hypoxanthine, supplemented with HEPES (5.94 g/L), $NaHCO_3$ (2.1 g/L), neomycin (100 U/mL), Albumax^R (5 g/L) and washed human red cells A^+ at 2.5 % haematocrit (0.3 % parasitaemia). Serial drug dilutions of eleven 3-fold dilution steps covering a range from 100 to 0.002 µg/ml were prepared. The 96-well plates were incubated in a

humidified atmosphere at 37 °C; 4 % CO_2, 3 % O_2, 93 % N_2. After 48 h 50 µl of ^{3}H-hypoxanthine (= 0.5 µCi) was added to each well of the plate. The plates were incubated for a further 24 h under the same conditions. The plates were then harvested with a Betaplate™ cell harvester (Wallac, Zurich, Switzerland), and the red blood cells transferred onto a glass fibre filter then washed with distilled water. The dried filters were inserted into a plastic foil with 10 ml of scintillation fluid, and counted in a Betaplate™ liquid scintillation counter (Wallac, Zurich, Switzerland). IC_{50} values were calculated from sigmoidal inhibition curves by linear regression[51] using Microsoft Excel. Chloroquine and artemisinin are used as control.

3.5.12 In vitro cytotoxicity with L-6 cells

Assays were performed in 96-well microtiter plates, each well containing 100 µl of RPMI 1640 medium supplemented with 1 % l-glutamine (200 mM) and 10 % fetal bovine serum, and 4000 L-6 cells (a primary cell line derived from rat skeletal myoblasts).[58,59] Serial drug dilutions of eleven 3-fold dilution steps covering a range from 100 to 0.002 µg/ml were prepared. After 70 h of incubation the plates were inspected under an inverted microscope to assure growth of the controls and sterile conditions. 10 µl of Alamar Blue was then added to each well and the plates incubated for another 2 h. Then the plates were read with a Spectramax Gemini XS microplate fluorometer (Molecular Devices Cooperation, Sunnyvale, CA, USA) using an excitation wave length of 536 nm and an emission wave length of 588 nm. The IC50 values were calculated by linear regression[51] from the sigmoidal dose inhibition curves using SoftmaxPro software (Molecular Devices Cooperation, Sunnyvale, CA, USA). Podophyllotoxin (Sigma P4405) is used as control.

3.5.13 Inhibition assay with 20S proteasome from yeast

Compounds were assayed for inhibition of the 20S proteasome as previously described,[60] but using 0.5 nM of 20S proteasome from *Saccharomyces cerevisiae* (Enzo Life Sciences).

3.5.14 Crystallization and structure elucidation

yCP crystals were grown in hanging drop plates at 20 °C as previously described,[8,61] using a protein concentration of 40 mg/mL in Tris/HCl (20 mm, pH 7.5) and EDTA (1 mm). The drops contained 1 µL of the reservoir solution consisting of 30 mm $MgAc_2$, 100 mm morpholino-ethane-sulfonic acid (MES) (pH 7.2) and 10% (v/v) 2-methyl-2,4-pentanediol as well as 1 µL of protein. Crystals appeared after 48 h and were then soaked with

macyranone A in DMSO at final concentrations of 10 mm for 12 h following complementation of the droplets with cryoprotecting buffer consisting of 30% (w/v) 2-methyl-2,4-pentanediol, 20 mm $MgAc_2$, 100 mm MES (pH 6.9). The crystals were supercooled in a stream of liquid nitrogen gas at 100 K (Oxford Cryo Systems). The dataset of yCP:macyranone A structure was collected with 2.8 Å resolution using synchrotron radiation (λ = 1.0 Å) at the X06SA-beamline (Swiss Light Source, Villingen, Switzerland, Table S8). X-ray intensities were assessed with the program XDS[62] and data reduction was carried out using XSCALE[62]. The molecular replacement started with the coordinates of yCP (pdb entry code: 1RYP) and Translation/Libration/Screw (TLS) refinements were performed with REFMAC5 in the CCP4i suite.[63] Model building was carried out using the program package MAIN.[64]

3.5.15 Proteasome purification

20S proteasome from *Saccharomyces cerevisiae* (yCP) was purified as previously described.[8] The yeast cells were lysed in a continuous cell disruption system and centrifuged at 40,000 g. The suspension was applied to a phenyl sepharose HIC column after precipitation in aqueous 40% $(NH_4)_2SO_4$. Eluted fractions with 20S proteasome activity were pooled and purified using a FPLC system with hydroxyapatite column. The polishing was performed via a Resource-Q anion exchange column and Superose 6 size exclusion chromatography. Pooled fractions were concentrated to 40 mg/mL in 20 mm Tris (pH 7.5) and used for vitro assays and crystallization trials.

3.5.16 IC_{50} value determination

The in vitro proteasome inhibition assays were performed with fluorescence assays in 96-well plates. The assay mixtures contained 10 µg/mL of commercially available purified human cCP (Boston Biochem) and human iCP (Boston Biochem) in 100 mm Tris/HCl (pH 7.5) buffer. Inhibitors were dissolved in DMSO and added at various concentrations with three repetitions each, with a final concentration of less than 10% (w/v) DMSO. After the incubation time of 60 min at RT, the fluorogenic substrate Suc-Leu-Leu-Val-Tyr-AMC (Bachem) was added to measure the residual activity of the chymotrypsin-like site. Afterwards, the assay mixture was incubated for another hour at RT and fluorescence was determined on a Varian Cary Eclipse photofluorometer with excitation and emission wavelengths of λ_{exc} = 360 and λ_{em} = 460 nm, respectively.

3.6 Supporting Information

3.6.1 Structure elucidation

Table S1. NMR spectroscopic data for macyranone A (1) (methanol-d_4)

		δ_C[a]	δ_H[b]	mult (*J*, Hz)	HMBC[c]	COSY[d]
Phe	1	174.7				
	2	55.1	4.68	dd (9.0, 4.8)	1, 3, 4, 1_{MMA}	3a, 3b
	3a	38.3	3.26	dd (14.1, 4.8)	1, 2, 4, 5	2, 3b
	3b		2.98	o	1, 2, 4, 5	2, 3a
	4	138.3				
	5	130.2	7.23	d (7.5)	3, 7	6
	6	129.4	7.28	t (7.5)	4, 6	5, 7
	7	127.8	7.21	t (7.5)	5	6
Mma	1	173.2				
	2	48.2	3.33	d (7.2)	1, 3, 4	4
	3	173.2				
	4	15.9	1.26	d (7.2)	1, 2, 3	2
Thr	1	172.3				
	2	60.2	4.34	d (4.0)	1, 3, 4, 3_{MMA}	3
	3	68.2	4.17	dq (6.4, 4.0)	1, 2, 4	2, 4
	4	19.7	1.16	d (6.4)	2, 3	3
aThr	1	172.4				
	2	59.8	4.37	d (6.2)	1, 3, 4, 1_{Thr}	3
	3	68.6	4.04	dq (6.3)	1, 2, 4	2, 4
	4	19.6	1.20	d (6.3)	2, 3	3
Leu	1	207.6				
	2	54.1	4.50	dd (10.4, 4.0)	1, 3a, 3b, 4, 1_{aThr}	3a, 3b
	3a	39.7	1.55	m	1, 2, 4, 5, 6	2, 3b, 4
	3b		1.50	m	1, 2, 4, 5, 6	2, 3a, 4
	4	25.9	1.72	m	2, 3, 5, 6	3, 5, 6
	5	21.5	0.92	d (6.5)	3, 4, 6	4
	6	23.5	0.96	d (6.7)	3, 4, 5	4
	1'	52.8	3.63	dd (4.8, 2.4)	1, 2, 2'a, 2'b	2'a, 2'b
	2'a	47.5	3.02	dd (4.9, 6.1)	1, 1'	1', 2'b
	2'b		2.98	dd (2.4, 6.1)	1, 1'	1', 2'a

[a]Recorded at 175 MHz; referenced to residual methanol-d_4 at δ 49.15 ppm.
[b]Recorded at 700 MHz; referenced to residual methanol-d_4 at δ 3.31 ppm.
[c]Proton showing correlation to indicated carbon.
[d]Proton showing correlation to indicated proton.

Table S2. NMR spectroscopic data for macyranone B-F (2-6) (methanol-d_4)

	Macyranone B		Macyranone C		Macyranone D		Macyranone E		Macyranone F	
	δ_C[a]	δ_H[b]	δ_C[a]	δ_H[b]	δ_C[a]	δ_H[b]	δ_C[a]	δ_H[b]	δ_C[a]	δ_H[b]
Phe										
1	174.8		--		177.9		174.6		175.4	
2	55.8	4.69	55.5	4.70	57.2	4.59	55.0	4.70	55.5	4.67
3a	38.5	3.26	39.0	3.26	38.8	3.32	38.2	3.25	38.3	3.27
3b		2.96		2.98		2.93		2.98		2.96
4	138.3		--		139.4		138.2		138.5	
5	130.2	7.23	131.3	7.23	130.3	7.26	130.2	7.23	130.3	7.23
6	129.2	7.27	130.2	7.28	129.3	7.26	129.4	7.27	129.3	7.27
7	127.5	7.21	128.5	7.21	127.5	7.19	127.7	7.21	127.7	7.20
Mma										
1	173.0		--		173.2		173.2		173.3	
2	47.9	3.33	49.0	3.34	48.7	3.35	48.2	3.33	48.3	3.32
3	173.1		--		173.4		173.2		173.4	
4	15.7	1.24	16.8	1.26	15.5	1.25	16.0	1.26	15.8	1.24
Thr										
1	172.3				172.6		172.3		172.6	
2	60.4	4.33	61.2	4.33	60.3	4.32	60.1	4.34	60.2	4.34
3	68.0	4.17	68.9	4.17	68.3	4.20	68.2	4.15	68.2	4.17
4	19.6	1.16	20.5	1.16	19.9	1.12	19.7	1.12	19.6	1.15
aThr										
1	171.9		--		172.4		172.0		172.2	
2	58.8	4.37	60.9	4.37	61.0	4.20	59.9	4.36	56.7	4.43
3	68.4	4.00	69.3	4.06	68.3	4.02	68.5	4.00	62.6	3.76
4	19.2	1.17	20.5	1.22	20.3	1.18	19.5	1.09		
	Abf		**Leu**		**His**		**Trp**		**His**	
1	212.0		--		209.8		212.1		212.2	
2	87.4		59.3	4.40	58.5	4.68	60.1	4.72	60.3	4.70
3a	45.0	1.66	40.9	1.52	26.5	3.25	27.7	3.22	27.6	3.22
3b		1.62				2.95		3.10		3.12
4	24.8	1.67	26.6	1.67	131.6		110.6		110.6	
5	24.3	0.97	24.4	0.94	118.6	7.19	128.4		128.5	
6	24.1	0.88	22.3	0.89	134.8	8.43	119.1	7.53	119.1	7.53
7							119.8	7.01	119.7	7.01
8							122.4	7.08	122.3	7.08
9							112.2	7.32	112.3	7.32
10							137.9		138.0	
11							124.4	7.07	124.4	7.08
1`a	65.9	4.34	33.7	2.56	33.3	2.67	34.4	2.46	34.3	2.45
1`b		4.18				2.52		2.35		2.35
2`a	37.5	2.78	8.4	0.99	7.7	1.01	7.4	0.89	7.4	0.86
2`b		2.50								

[a]Recorded at 175 MHz; referenced to residual methanol-d_4 at δ 49.15 ppm.
[b]Recorded at 700 MHz; referenced to residual methanol-d_4 at δ 3.31 ppm.

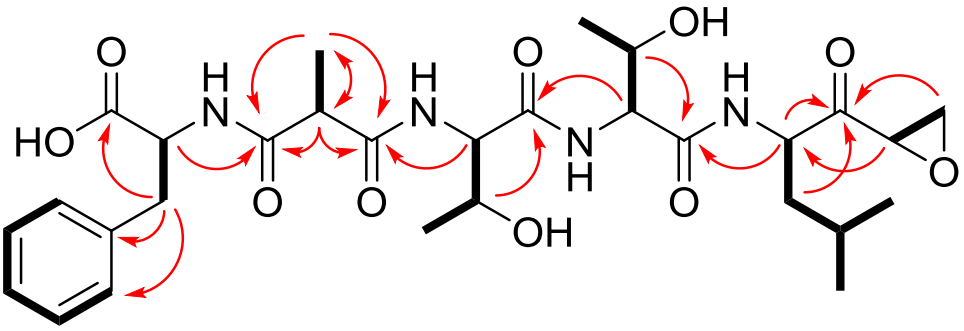


Figure S1. Key HMBC (red arrows) and COSY (thick bonds) correlations for macyranone A (**1**)

Table S3. Retention times of the D- and L-FDLA derivatized amino acid residues received from hydrolysis of the macyranones compared to the respective standards

amino acid	detected m/z	Macyranone A		Macyranone B		Macyranone C		Standard	
		D-FDLA	L-FDLA	D-FDLA	L-FDLA	D-FDLA	L-FDLA	D-FDLA	L-FDLA
L-Phe	460.18	22.16	18.95	22.13	18.93	22.18	18.94	21.49	18.26
L-Thr	414.16	16.08	13.03	16.05	13.00	16.10	13.03	15.36	12.38
L-*allo*-Thr	414.16	14.81	13.59	14.79	13.55	14.83	13.59	14.09	12.91
L-Ser	400.15	--	--	--	--	--	--	13.07	12.55
L-Leu	426.20	22.69/ 22.85	18.53/ 18.79	22.67/ 22.84	18.49/ 18.75	22.72/ 22.89	18.54/ 18.78	22.05/ 22.19	17.84/ 18.10
L-His	450.17	--	--	--	--	--	--	9.35	10.42

amino acid	detected m/z	Macyranone D		Macyranone F		Macyranone F	
		D-FDLA	L-FDLA	D-FDLA	L-FDLA	D-FDLA	L-FDLA
L-Phe	460.18	21.52	18.25	22.18	18.93	21.52	18.27
L-Thr	414.16	15.38	12.37	16.09	13.04	15.38	12.39
L-*allo*-Thr	414.16	14.15	12.91	14.83	13.58	--	--
L-Ser	400.15	--	--	--	--	13.09	12.55
L-Leu	426.20	--	--	--	--	--	--
L-His	450.17	9.29	10.30	--	--	--	--

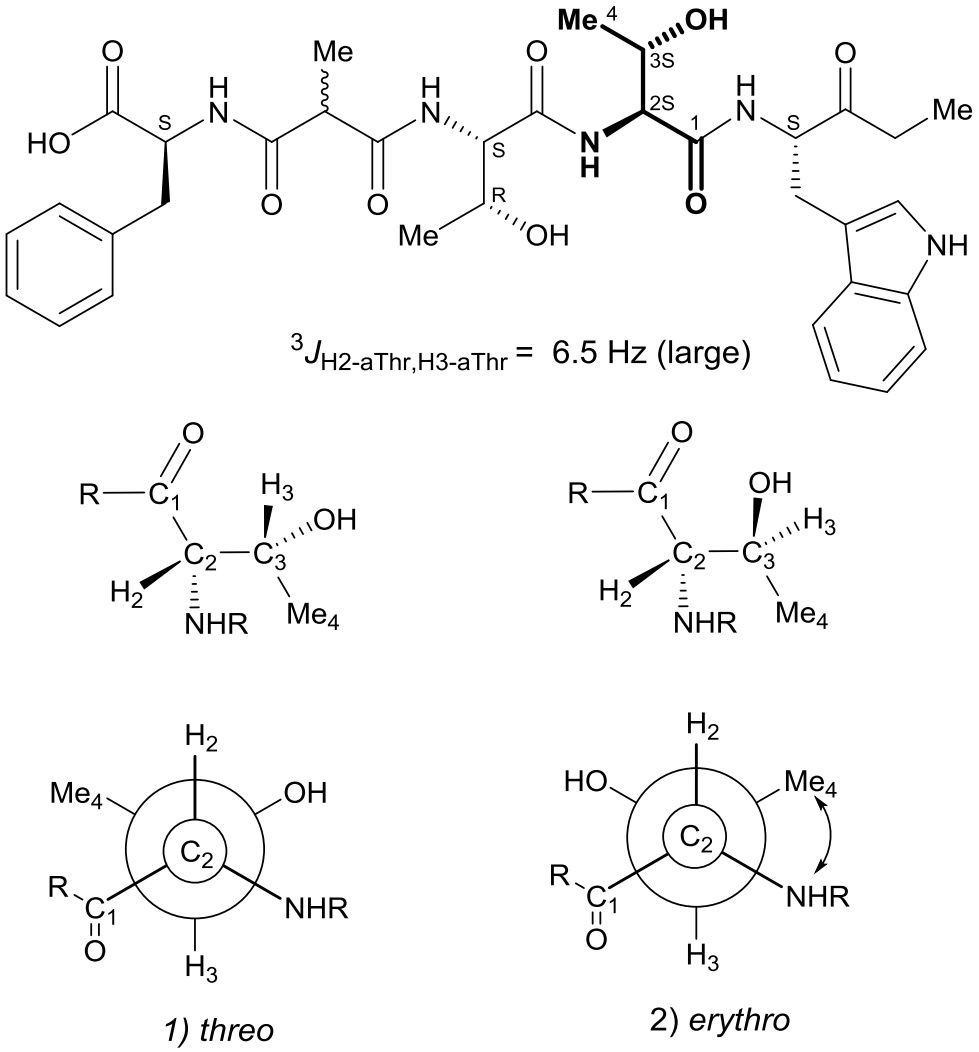
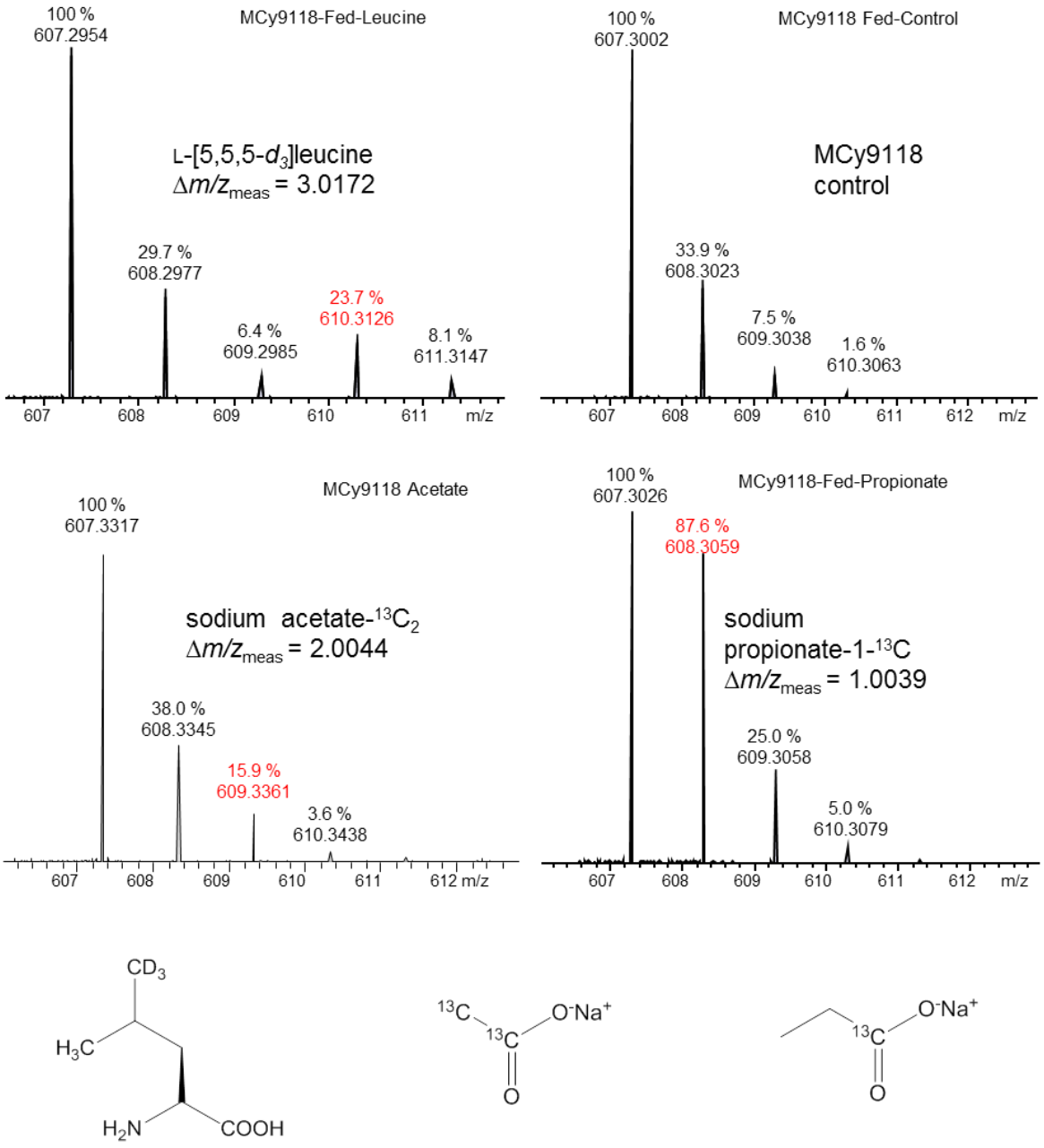


Figure S2. *J* based configuration analysis of macyranone E to confirm the L-*allo*-configuration of aThr.[24] The large coupling constant between H2-aThr and H3-aThr gives rise to the two possible rotamers 1 and 2 with anti-orientation of the respective protons and allows their discrimination and the assignment of the *erythro* configuration based on a key ROE correlation between the methyl protons Me4-aThr and the amine proton NH-aThr.

3.6.2 Feeding experiments using labeled precursors



$^{2}H_3$-L-leucine
$\Delta m/z_{calc}$ = 3 x (^{2}H -^{1}H)
= 3 x 1.006155 Da
= 3.018465 Da
$\Delta m/z_{meas}$ = 3.0172

$^{13}C_1$-sodium propionate
$\Delta m/z_{calc}$ = 2 x (^{13}C -^{12}C)
= 2 x 1.003355 Da
= 2.00671 Da
$\Delta m/z_{meas}$ = 2.0044

$^{13}C_1$-sodium propionate
$\Delta m/z_{calc}$ = (^{13}C -^{12}C)
= 1 x 1.003355 Da
= 1.003355 Da
$\Delta m/z_{meas}$ = 1.0039

Figure S3. Isotopic peak pattern of the highly abundant $[M+H]^+$ signals for macyranone A (607 *m/z*) prove incorporation of labeled L-leucine as well as sodium propionate and acetate in MCy9118. The observed mass shifts fit to the heavy isotopes that were incorporated.

3.6.3 Biosynthetic gene cluster analysis

The biosynthetic gene cluster was identified using the antiSMASH 2.0 software tool.[26] The biosynthetic gene cluster spans 21,003 bp and consists of 8 open reading frames (ORFs) named *mynA* to *mynH*. It has an overall GC content of 70.0%. The catalytic domains were annotated using antiSMASH 2.0, Pfam[46], NRPS predictor2[47], and PKS/NRPS Analysis[48]. Genes that did not allow a prediction for PKS or NRPS domains were analyzed via the BLAST algorithm.[27]

Table S4. Deduced functions of ORFs in the macyranone biosynthetic gene cluster

Gene	AA	Protein homolog	Similarity	reference	Proposed function
mynA	603	PdmN, *Actinomadura hibisca*	42 %	ADB23384	amidotransferase
mynB	509				stand-alone A_{Thr} domain
mynC	588	Butyryl-CoA dehydrogenase, *Streptomyces* sp. NRRL S-1022	67 %	WP_030347517	acyl-CoA dehydrogenase
mynD	3058				NRPS (C-PCP, C-A_{Thr}-PCP, C-A_{Leu}-PCP)
mynE	411	MFS transporter permease, *Saccharothrix* sp. NRRL B-16314	38 %	WP_033439102	transporter
mynF	527	Peptide synthetase, *Streptomyces griseus*	51 %	WP_003964446	stand-alone A_{His} domain
mynG	997				PKS (KS-AT-ACP)
mynH	258	thioesterase, *Amycolatopsis mediterranei*	44 %	WP_013230025	thioesterase type II

A: adenylation domain, ACP: acyl-carrier-protein domain AT: acyltransferase domain, C: condensation domain, KS: ketosynthase domain, PCP: peptidyl-carrier-protein domain

3.6.3.1 MynA analysis

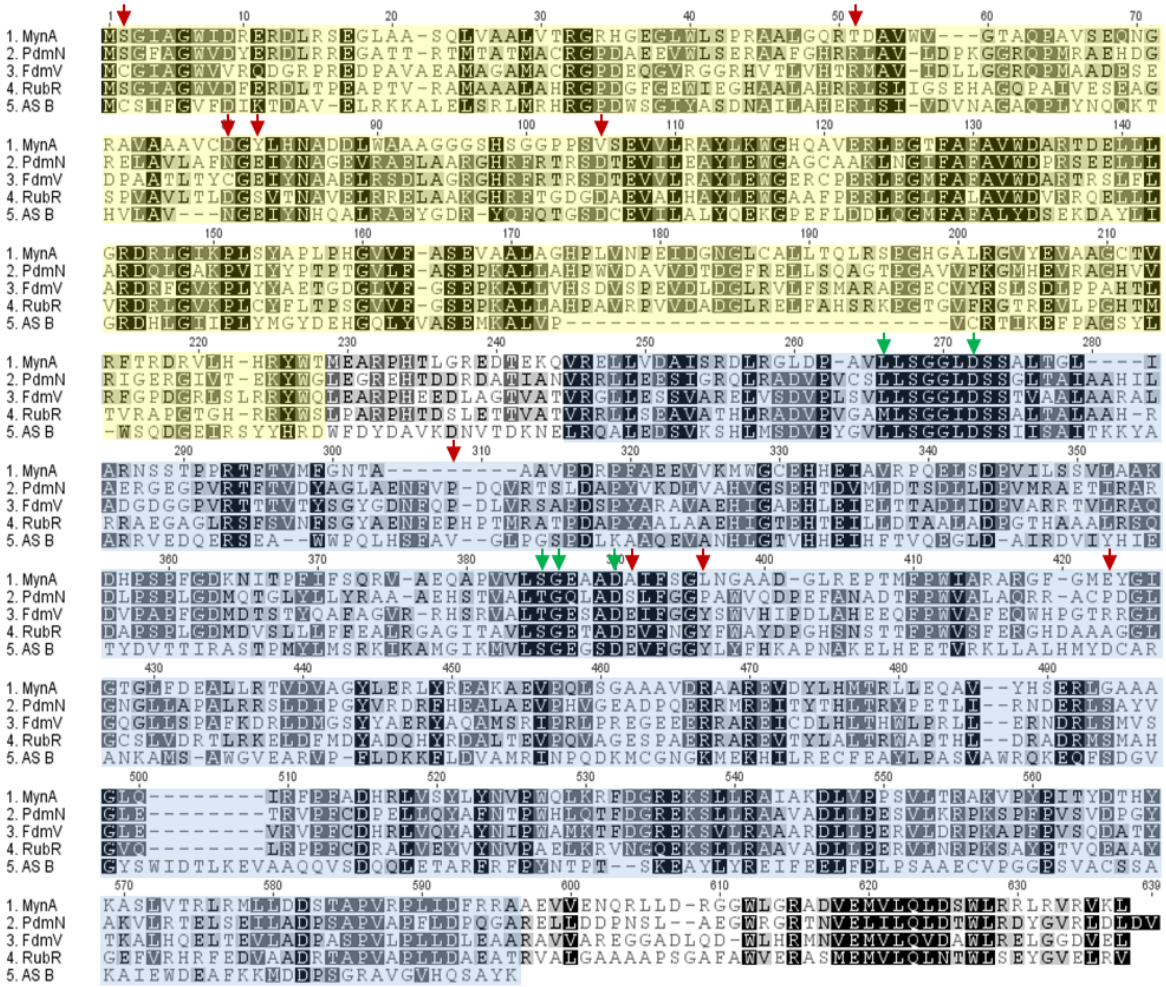


Figure S4. Amino acid sequence alignment of MynA with selected homologs. Yellow is the N-terminal glutaminase and blue the C-terminal amide synthetase domain. Arrows indicate active side residues that are consistent (green) or not (red) between MynA and AS B. Accession numbers for each of the enzymes are: RubR: AAM97368; PdmN: ABK58686; FdmV: AAQ08933; AS B: NP_415200.

3.6.3.2 A domain analysis

Three NRPS modules were identified in the biosynthetic gene cluster but in total four adenylation domains. Two of the adenylation domains represent discrete proteins. The A domain specificity was determined using the online tool NRPSpredictor2[47] based on the extracted Stachelhaus code.[62] The results are listed in Table S5. The predicted specificities of A domains 1-3 is consistent with the observed amino acids incorporated into the main product macyranone A. The fourth adenylation domain shows substrate specificity for histidine. Since histidine is incorporated in macyranone D, the domain might act as alternative adenylation domain for module three. Still, no adenylation domain with tryptophan specificity could be identified to explain the occurrence of macyranone E and F.

Table S5. Analysis of the adenylation domain substrate specificity based on the Stachelhaus code using NRPSpredictor2[47]

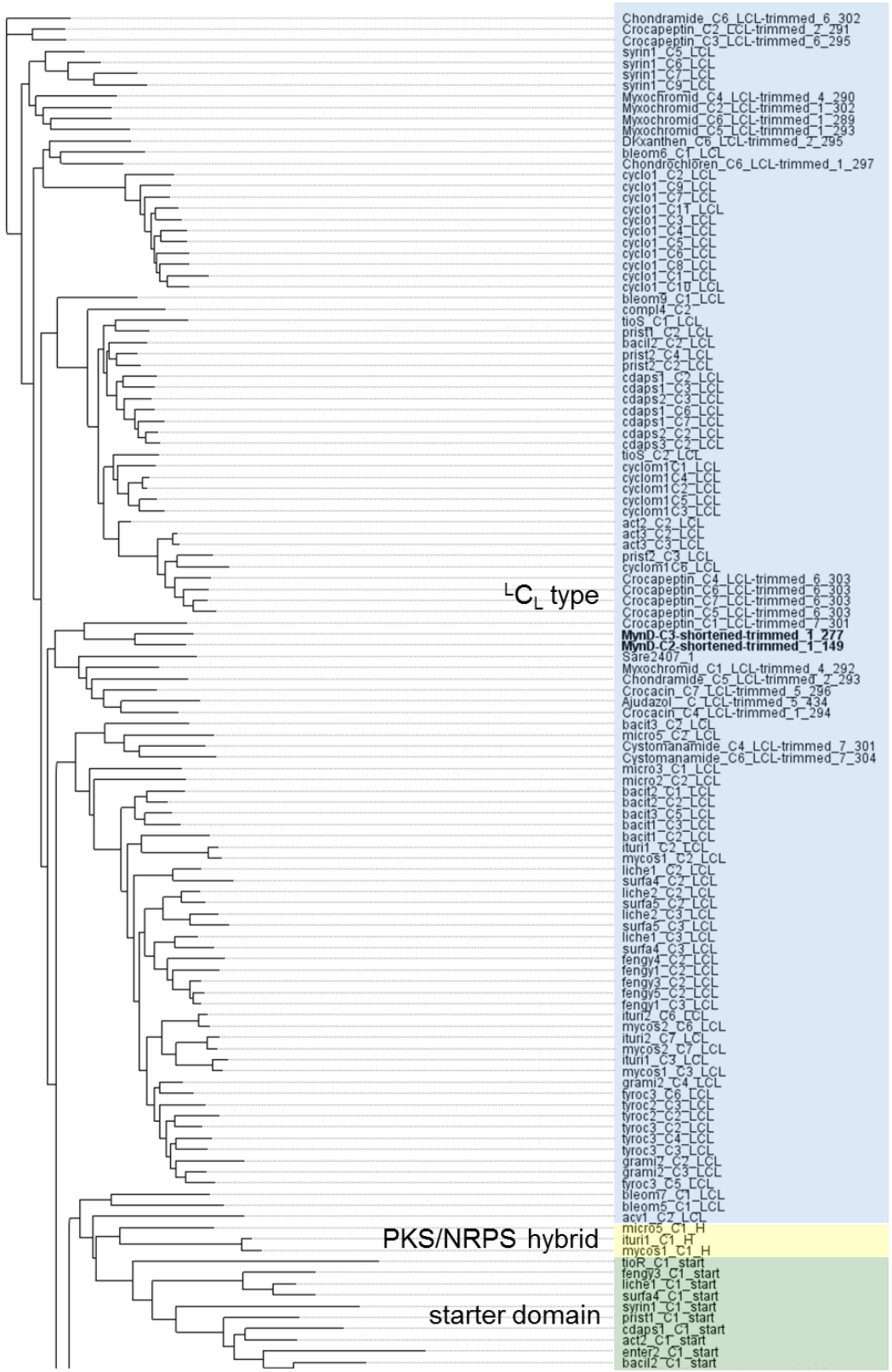
Adenylation domain	**The extracted Stachelhaus code using NRPSpredictor 2**										**Predicted amino acid[a]**
MynB-A1	D	F	F	N	I	G	T	V	F	K	Thr [b]
MynD-A2	D	F	W	N	I	G	M	V	H	K	Thr [c]
MynD-A3	D	A	W	F	M	I	A	V	V	K	Leu [b]
MynF-A4	D	S	A	L	I	A	E	V	W	K	His [b]

[a] substrate specificities as predicted by comparison of the 10 amino acid code deefindet by Stachelhaus et al.[62]
[b] predictions with lower confidence (70%)
[c] predictions with high confidence (100%)

3.6.3.3 C domain analysis

The C domains *myn* biosynthetic gene cluster were analyzed using NaPDoS.[49] MynD-C2 and –C3 are characterized as $^{L}C_{L}$-type condensation domains. This result confirms that both domains catalyze the formation of a peptide bond between two l-amino acids.

MynD-C1 does not group with any of the known C domain types. The result emphasizes the hypothesis, that the first condensation domain is involved in the biosynthesis of the unusual starter unit. It is a likely candidate for the catalysis of amide bond formation between the methylmalonate-phenylalanine biolding block and the first threonine that is selected and activated by the first adenylation domain identified in the gene cluster.


Chondramide_C6_LCL-trimmed_6_302
Crocapeptin_C2_LCL-trimmed_2_291
Crocapeptin_C3_LCL-trimmed_6_295
syrin1_C5_LCL
syrin1_C6_LCL
syrin1_C7_LCL
syrin1_C9_LCL
Myxochromid_C4_LCL-trimmed_4_290
Myxochromid_C2_LCL-trimmed_1_302
Myxochromid_C6_LCL-trimmed_1_299
Myxochromid_C5_LCL-trimmed_1_293
DKxanthen_C6_LCL-trimmed_2_295
bleom6_C1_LCL
Chondrochloren_C6_LCL-trimmed_1_297
cyclo1_C2_LCL
cyclo1_C9_LCL
cyclo1_C7_LCL
cyclo1_C11_LCL
cyclo1_C3_LCL
cyclo1_C4_LCL
cyclo1_C5_LCL
cyclo1_C6_LCL
cyclo1_C8_LCL
cyclo1_C1_LCL
cyclo1_C10_LCL
bleom9_C1_LCL
compl4_C2
tioS_C1_LCL
prist1_C2_LCL
bacil2_C2_LCL
prist2_C4_LCL
prist2_C2_LCL
cdaps1_C2_LCL
cdaps1_C3_LCL
cdaps2_C3_LCL
cdaps1_C6_LCL
cdaps1_C7_LCL
cdaps2_C2_LCL
cdaps3_C2_LCL
tioS_C2_LCL
cyclom1C1_LCL
cyclom1C4_LCL
cyclom1C2_LCL
cyclom1C5_LCL
cyclom1C3_LCL
act2_C2_LCL
act3_C2_LCL
act3_C3_LCL
prist2_C3_LCL
cyclom1C6_LCL
Crocapeptin_C4_LCL-trimmed_6_303
Crocapeptin_C6_LCL-trimmed_6_303
Crocapeptin_C7_LCL-trimmed_6_303
Crocapeptin_C5_LCL-trimmed_6_303
Crocapeptin_C1_LCL-trimmed_7_301
MynD-C3-shortened-trimmed_1_277
MynD-C2-shortened-trimmed_1_149
Sare2407_1
Myxochromid_C1_LCL-trimmed_4_292
Chondramide_C5_LCL-trimmed_2_293
Crocacin_C7_LCL-trimmed_5_296
Ajudazol_C_LCL-trimmed_5_434
Crocacin_C4_LCL-trimmed_1_294
bacit3_C2_LCL
micro5_C2_LCL
Cystomanamide_C4_LCL-trimmed_7_301
Cystomanamide_C6_LCL-trimmed_7_304
micro3_C1_LCL
micro2_C2_LCL
bacit2_C1_LCL
bacit2_C2_LCL
bacit3_C5_LCL
bacit1_C3_LCL
bacit1_C2_LCL
ituri1_C2_LCL
mycos1_C2_LCL
liche1_C2_LCL
surfa4_C2_LCL
liche2_C2_LCL
surfa5_C2_LCL
liche2_C3_LCL
surfa5_C3_LCL
liche1_C3_LCL
surfa4_C3_LCL
fengy4_C2_LCL
fengy1_C2_LCL
fengy3_C2_LCL
fengy5_C2_LCL
fengy1_C3_LCL
ituri2_C6_LCL
mycos2_C6_LCL
ituri2_C7_LCL
mycos2_C7_LCL
ituri1_C3_LCL
mycos1_C3_LCL
grami2_C4_LCL
tyroc3_C6_LCL
tyroc2_C3_LCL
tyroc2_C2_LCL
tyroc3_C2_LCL
tyroc3_C4_LCL
tyroc3_C3_LCL
grami2_C2_LCL
grami2_C3_LCL
tyroc3_C5_LCL
bleom7_C1_LCL
bleom5_C1_LCL
acv1_C2_LCL
micro5_C1_H
ituri1_C1_H
mycos1_C1_H
tioR_C1_start
fengy3_C1_start
liche1_C1_start
surfa4_C1_start
syrin1_C1_start
prist1_C1_start
cdaps1_C1_start
act2_C1_start
enter2_C1_start
bacil2_C1_start
$^{L}C_{L}$ type
PKS/NRPS hybrid
starter domain

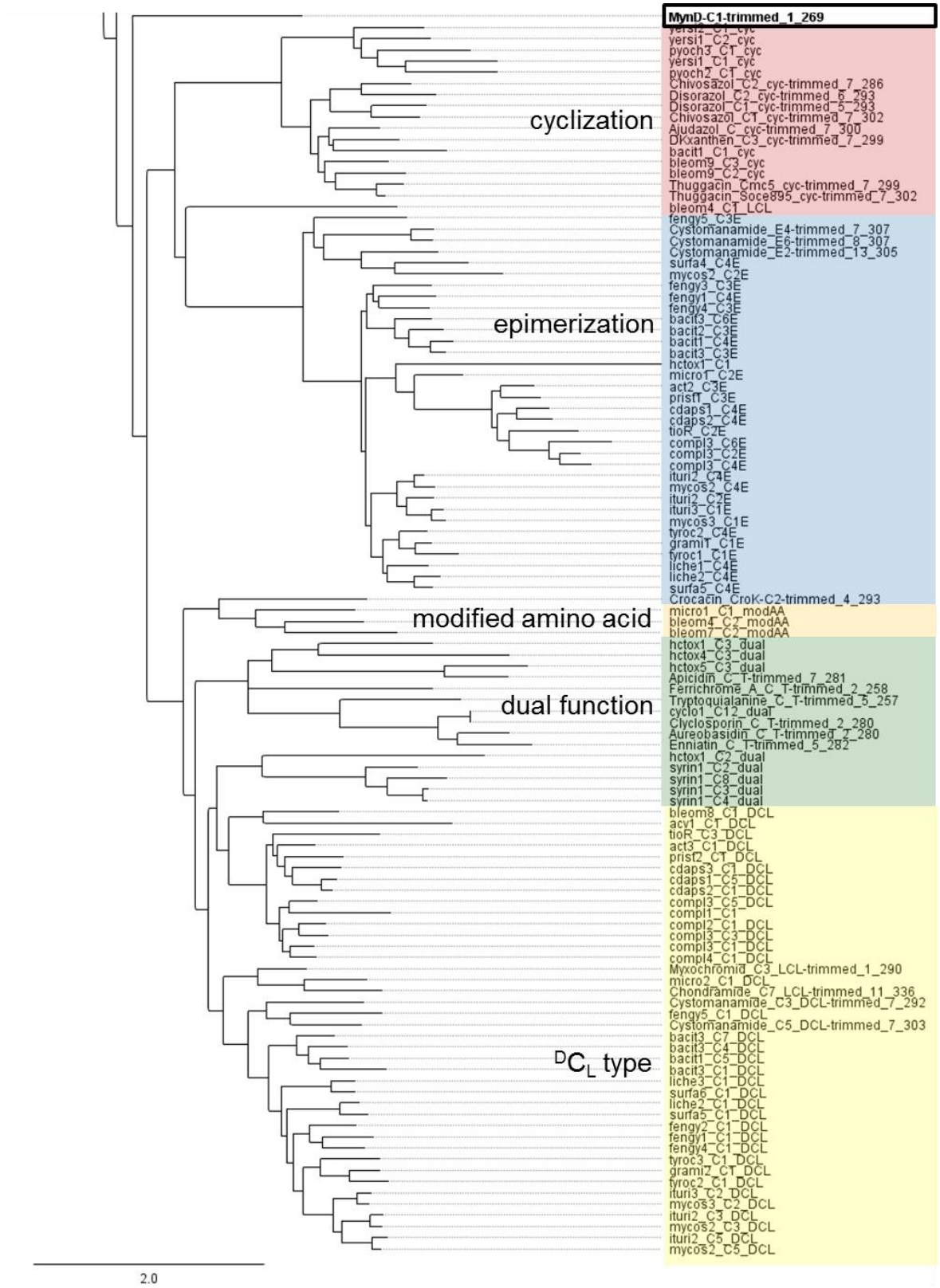
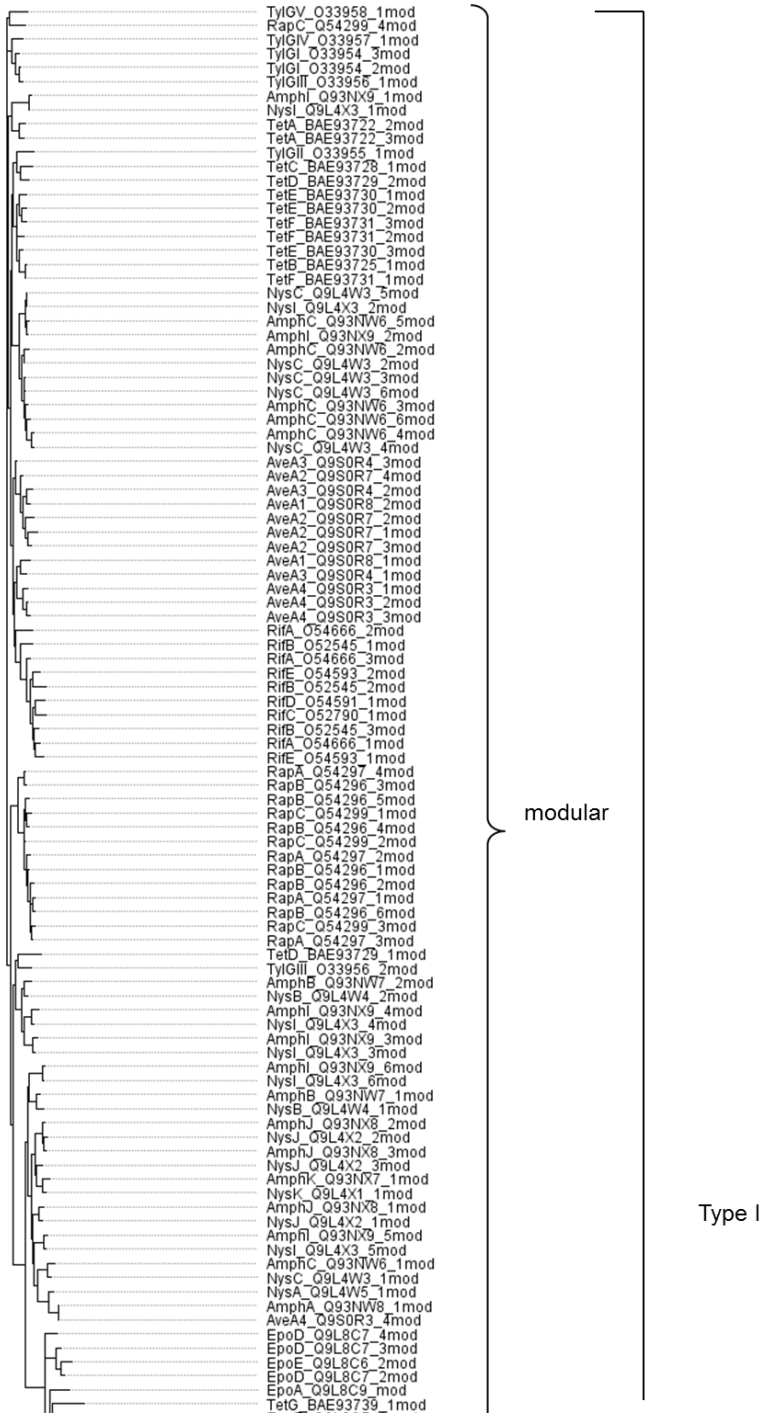


Figure S5. Phylogenetic tree highlighting different C domain types and the C-domains found in the *myn* biosynthetic gene cluster in *Cystobacter fuscus* MCy9118.

3.6.3.4 KS domain analysis

The ketosynthase (KS) domain of the myn was extracted and analyzed using NaPDoS.[49] The resulting phylogenetic tree shown in Figure S6 allows the assignment of the KS domain as PKS/NRPS hybrid KS domain. This assignment confirms the gene cluster analysis that the KS domain is situated downstream of a PCP domain (PCP3 on MynD) and catalyzes the condensation reaction between an amino acid and an acyl precursor.



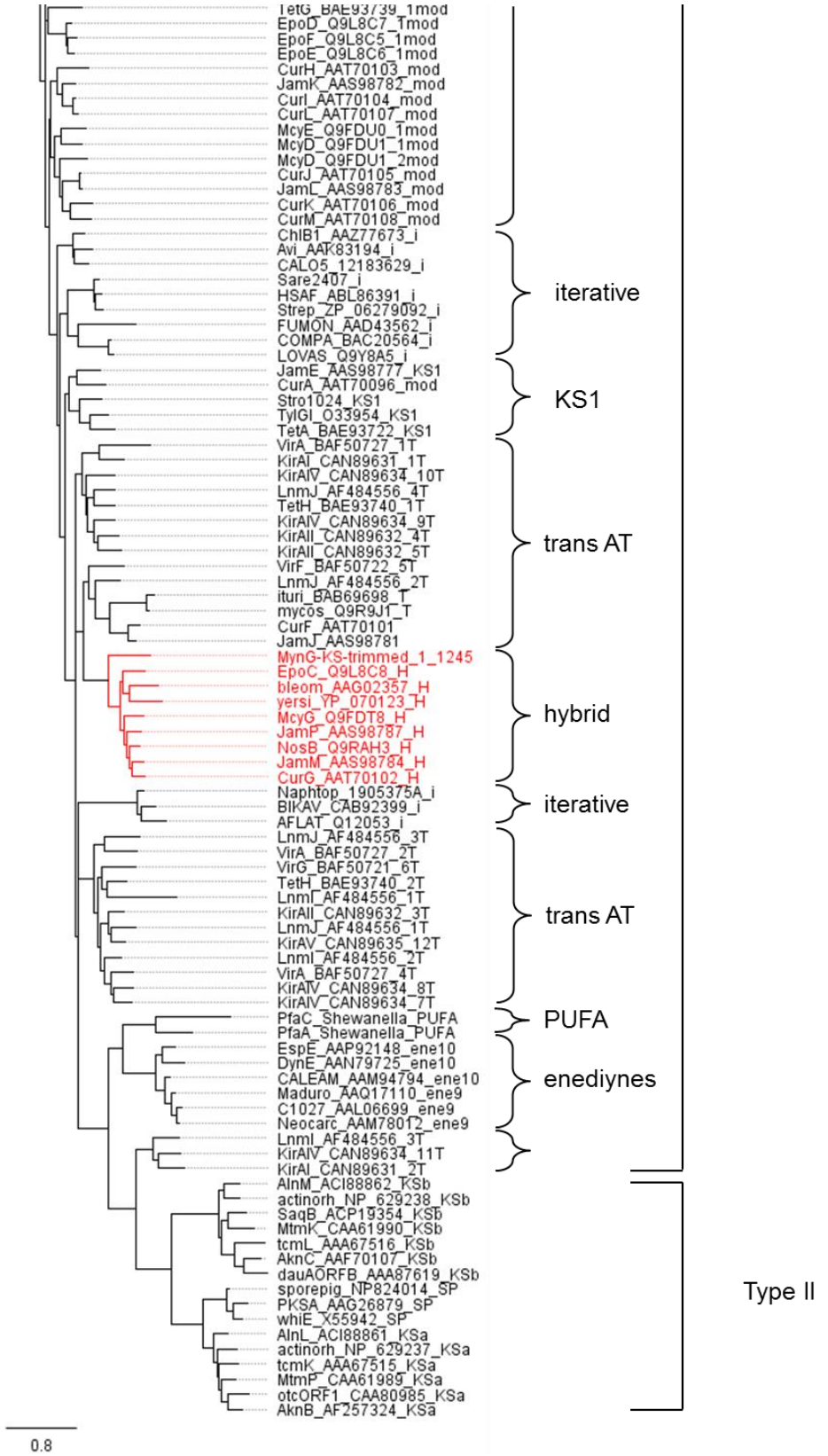


Figure S6. Phylogenetic tree highlighting different KS domain types and the KS domain found in the myn biosynthetic gene cluster in *Cystobacter fuscus* MCy9118 confirming it as hybrid KS domain

3.6.3.5 AT domain analysis

The substrate specificity of the acyl transferase (AT) domain encoded by *mynG* was analyzed using conserved motif analysis and comparison to the AT domains from the known epoxyketone biosynthetic gene clusters from epoxomicin and eponemycin. The domains were aligned to a reference AT from *E.coli* FAS, 1MLA (PDB 1MLA, UniProtKB P0AAI9) (Table S6). The data show that the three AT contain the GxSxG consensus motif with the catalytic serine residue. Even if the three AT domains are suspected to incorporate the same building block (most likely malonyl-CoA), the active site analysis to predict the used extender unit based on the amino acid position 200 in 1MLA is inconsistent. Ser200 as in EpnH supports mm-CoA and Phe200 as in EpxE supports m-CoA whereas various different residues like Gln200 could not be unequivocally assigned to a substrate. The incorporation of an unusual residue and conversion to the epoxyketone residue can therefore not be excluded.

Table S6. Active site analysis of the acyl transferase domains (AT) according to Yadav et al.[63] from the three PKS ancoded by the *myn*, the *epn* and *epx* biosynthetic gene clusters from eponemycin and epoxomicin, respectively

AT domain	observed	predicted	11	63	90	91	92	93	94	117	200	201	231	250	255
MynG-AT	m-CoA	?	Q	Q	G	H	S	L	G	R	G	H	D	H	I
EpnH-AT	m-CoA	mm-CoA	H	Q	G	N	S	V	G	R	S	H	T	H	V
EpxE-AT	m-CoA	m-CoA	Q	Q	G	N	S	I	G	R	F	H	N	H	V

3.6.3.6 TE domain analysis

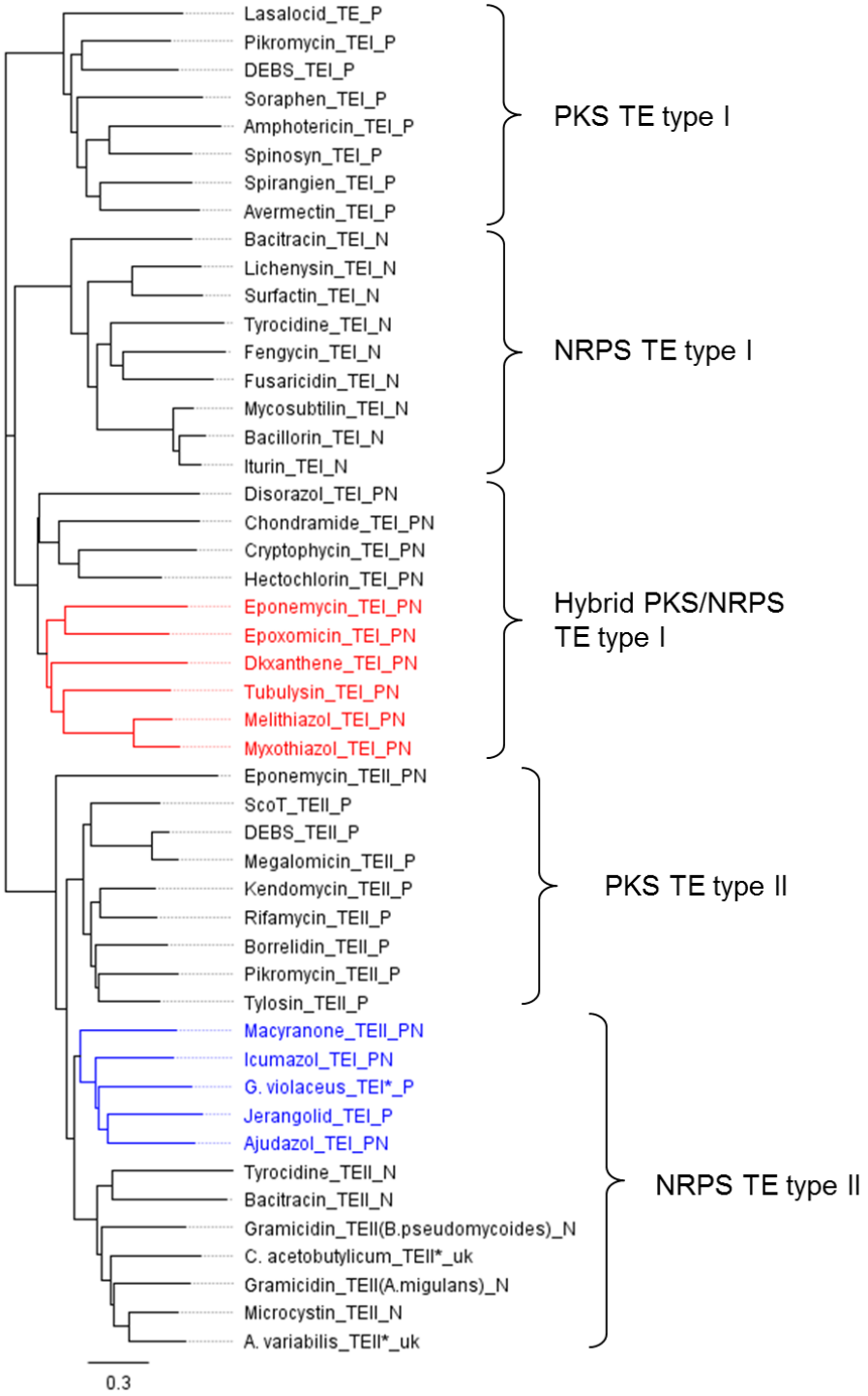


Figure S7. Phylogenetic tree analysis of various type I and II TE proteins from PKS and NRPS systems. The characterization as type I or type II thioesterase reflect whether the TE exists as an integral domain within a PKS or NRPS subunit (type I) or as discrete proteins (type II). The type of cluster to which each TE belongs is given (P, PKS; N, NRPS; PN, PKS/NRPS hybrid; *function assigned on the basis of genome annotation). Branch length indicates the number of inferred amino acid changes per position.

3.6.4 Targeted inactivation of the *myn* locus in MCy9118

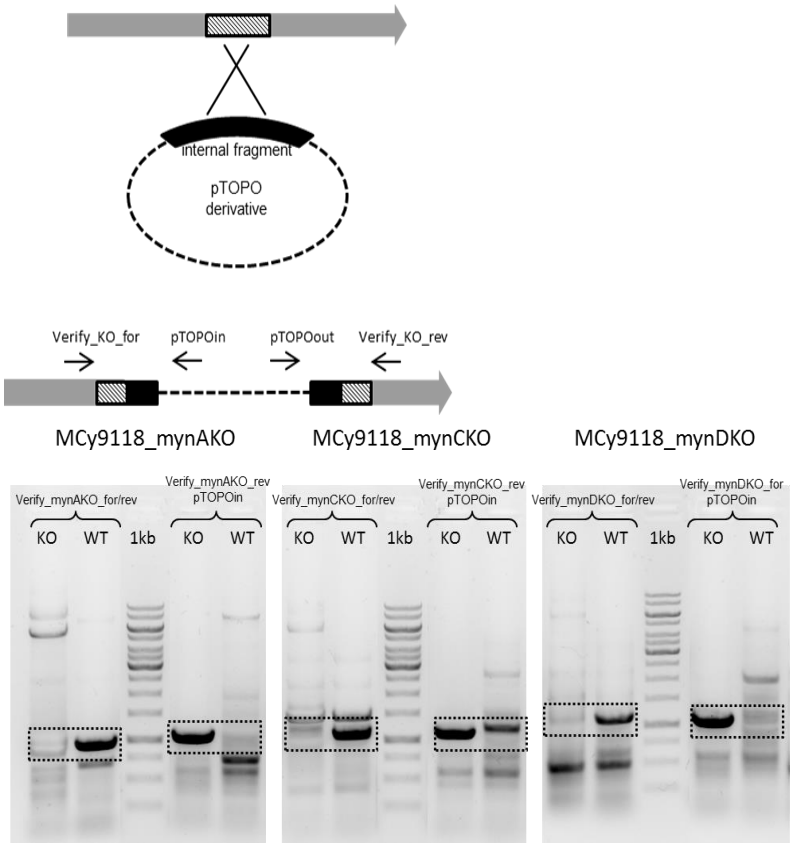


Figure S8. Left side: Principle of homologous recombination and PCR verification used for the *Cystobacter fuscus* MCy9118 knockout mutants; Right: Verification of mutants MCy9118::*mynA*KO, MCy9118::*myn*CKO and MCy9118::*mynD*KO using PCR

Table S7. Primers used for amplification and verification of the macyranone single crossover knock-out mutants in Mcy9118

MynD KO	
*mynD*KOFor	TGATTGATTGAGGTGGTCGACCGGGCCTCCCTG
*mynD*KORev	TCAATCAATCACGGCAGCAGCAAGGCCACCAGTTC
Verify_*mynD*KO_for	GAGGTTGGCTGGCACTGGTTG
Verify_*mynD*KO_rev	CAGCATCGCCTGGACCTGCTC
MynA KO	
*mynA*KOFor	TGATTGATTGATCCAGCGCGCTGACGGGCCTG
*mynA*KORev	TCAATCAATCATACGCGCAATCGACGGAGCCAGC
Verify_*mynA*KO_for	CGGTCCTGCTCTCTGGTGGTC
Verify_*mynA*KO_rev	GTGAAGAGACTAGAGCTTGACGCG
MynC KO	
*myn*CKOFor	TGATTGATTGAGACCGCCAGCTCCCTGACGG
*myn*CKORev	TCAATCAATCACAGGAACCGCTCGACCGGGAC
Verify_*myn*CKO_for	CTGCTTCGCGAGCGAATCGAC
Verify_*myn*CKO_rev	GTAGTACGACAGGATCGACATCCG

3.6.5 X-ray analysis of macyranone A in complex with yeast 20S proteasome

Table S8. Crystallographic data collection and refinement statistics of yCP in complex with macyranone A.

Crystallographic data	yCP:macyranone A*
Crystal parameters Space group Cell constants (dataset was collected from 1 crystal / 1 CP per AU)	 $P2_1$ a = 133.67 Å b = 300.70 Å c = 143.99 Å β = 112.52 °
Data collection Beamline Wavelength, (Å) Resolution range, (Å)† No. observations No. unique reflections‡ Completeness, (%)† R_{merge}, (%)†,§ I/σ (I)†	 X06SA, SLS 1.0 30-2.8 (2.9-2.8) 773,359 250,454 97.4 (97.0) 6.0 (49.0) 14.6 (2.7)
Refinement (REFMAC5) Resolution range, (Å) No. reflections working set No. reflections test set No. non-hydrogen No. ligand atoms Water, Mg^{2+} R_{work}/R_{free} (%)¶ RMSD bond (Å)/(°)** Average B-factor (Å²) Ramachandran plot, %***	 15-2.8 237,931 12,523 49,929 300 368 19.4/21.9 0.005/0.96 72.8 97.5/2.3/0.2
PDB accession code	5AHJ

*Dataset has been collected on a single crystal.

†Values in parentheses of resolution range, completeness, R_{merge}, and I/σ (I) correspond to the last resolution shell.

‡Friedel pairs were treated as identical reflections.

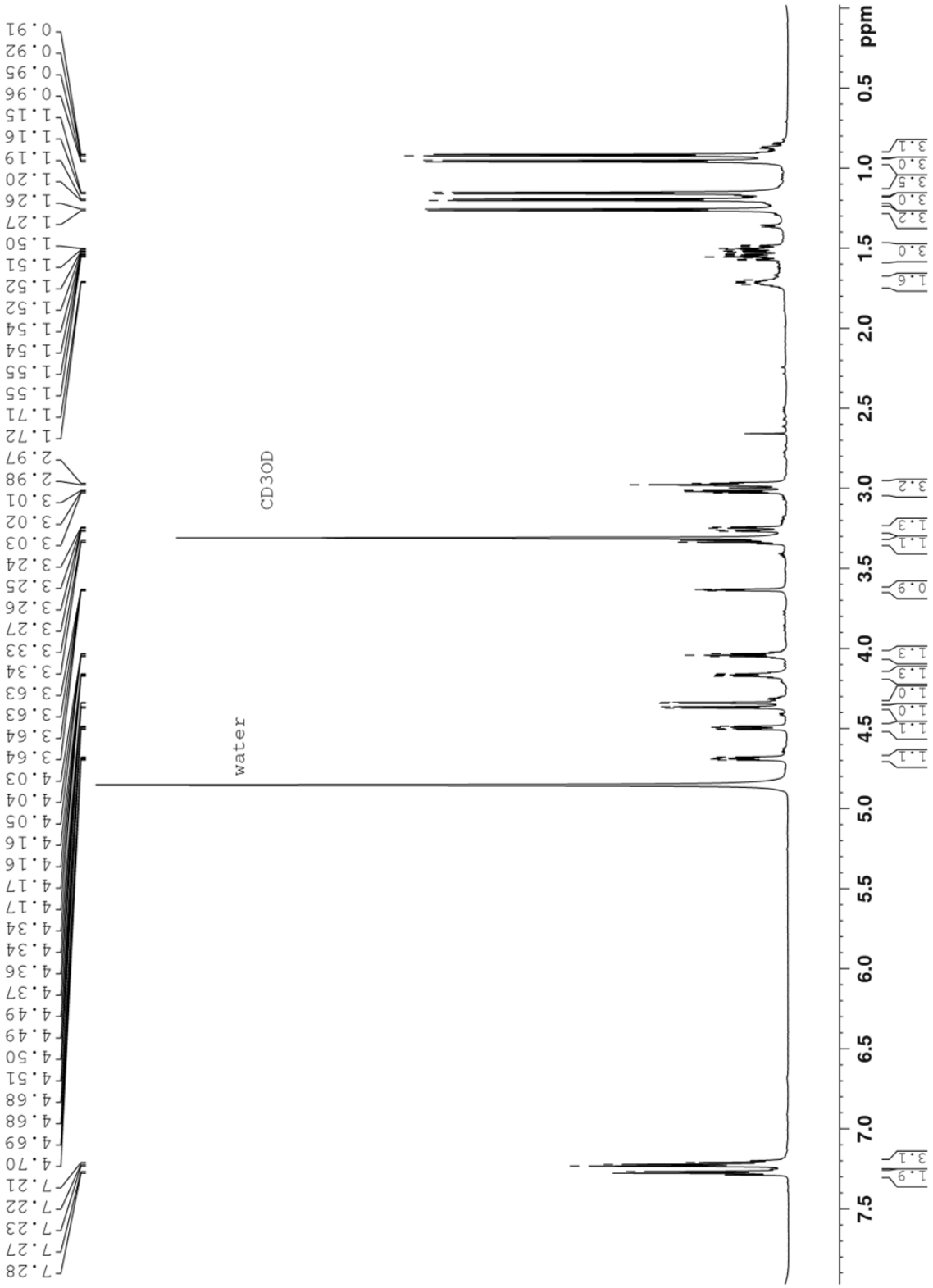
§R_{merge} $(I) = \Sigma_{hkl}\Sigma_j |[I(hkl)_j - I(hkl)]| / \Sigma_{hkl} I_{hkl}$, where $I(hkl)_j$ is the measurement of the intensity of reflection hkl and <I(hkl)> is the average intensity.

¶$R = \Sigma_{hkl} ||F_{obs}| - |F_{calc}|| / \Sigma_{hkl} |F_{obs}|$, where R_{free} is calculated without a sigma cut off for a randomly chosen 5% of reflections, which were not used for structure refinement, and R_{work} is calculated for the remaining reflections.

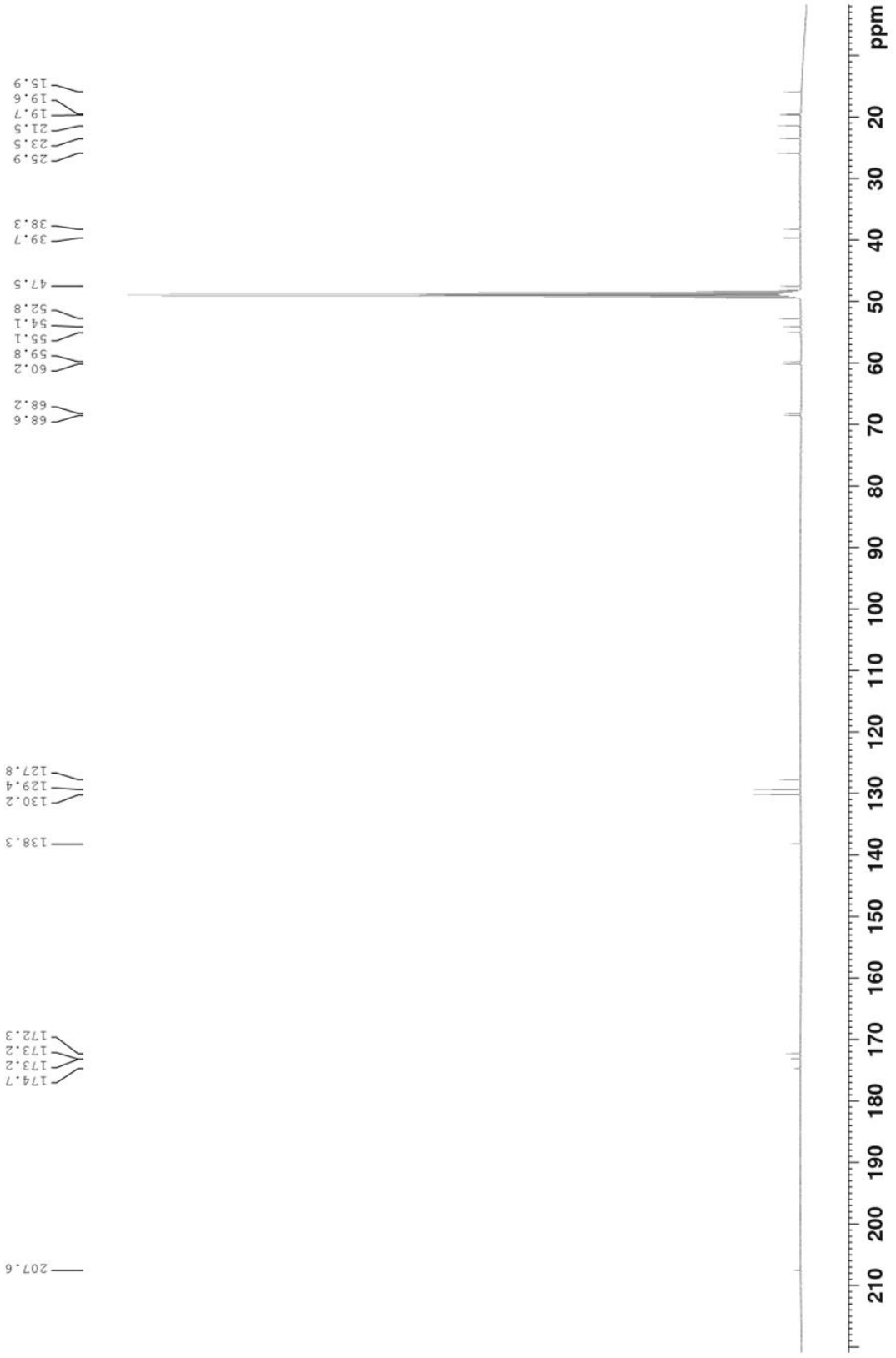
**Deviations from ideal bond lengths/angles.

***Number of residues in favored region/allowed region/outlier region.

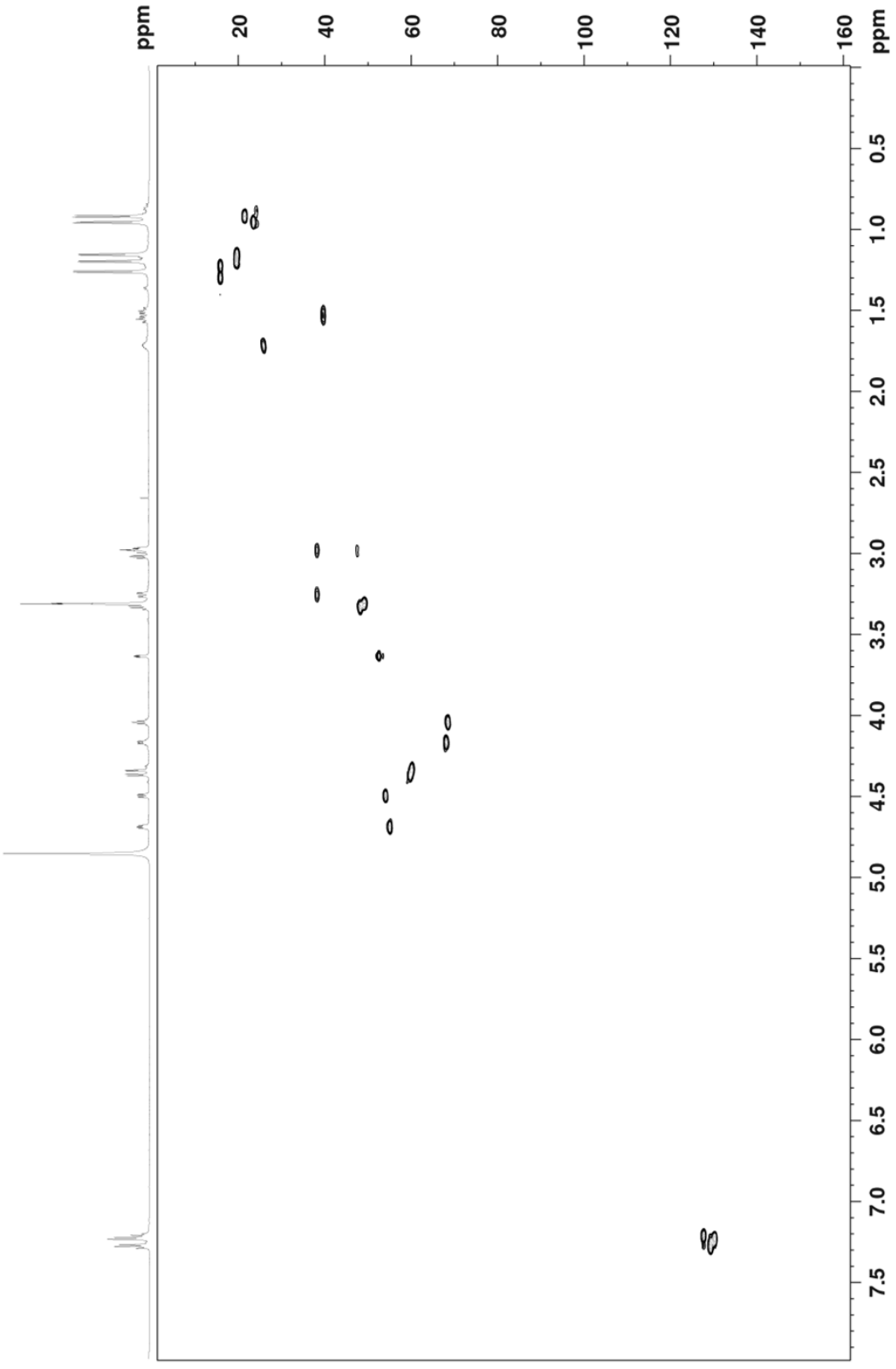
3.6.6 NMR spectra



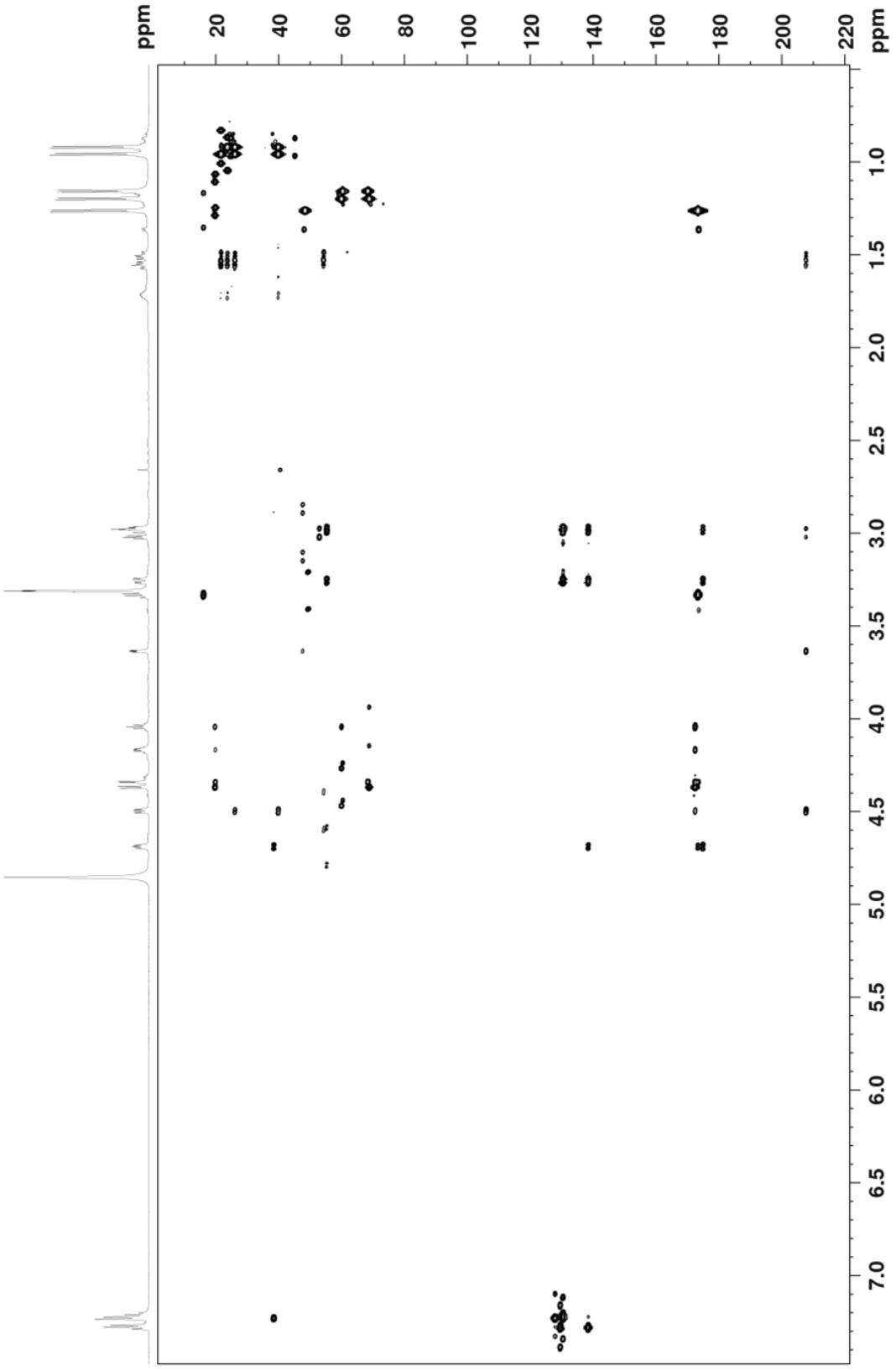
^{1}H NMR spectrum of macyranone A (**1**) in CD_3OD, 700 MHz



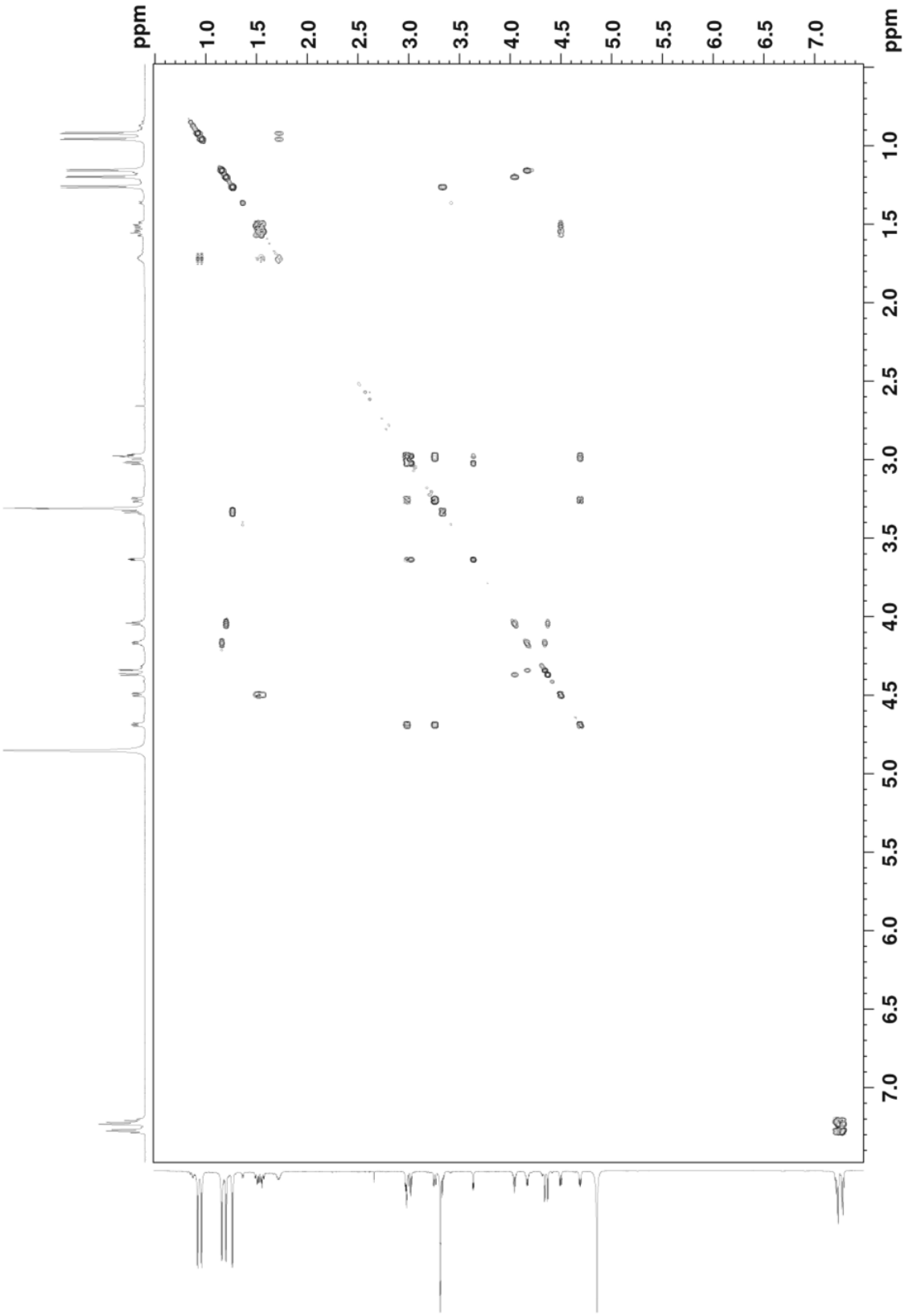
^{13}C NMR spectrum of macyranone A (**1**) in CD_3OD, 125 MHz



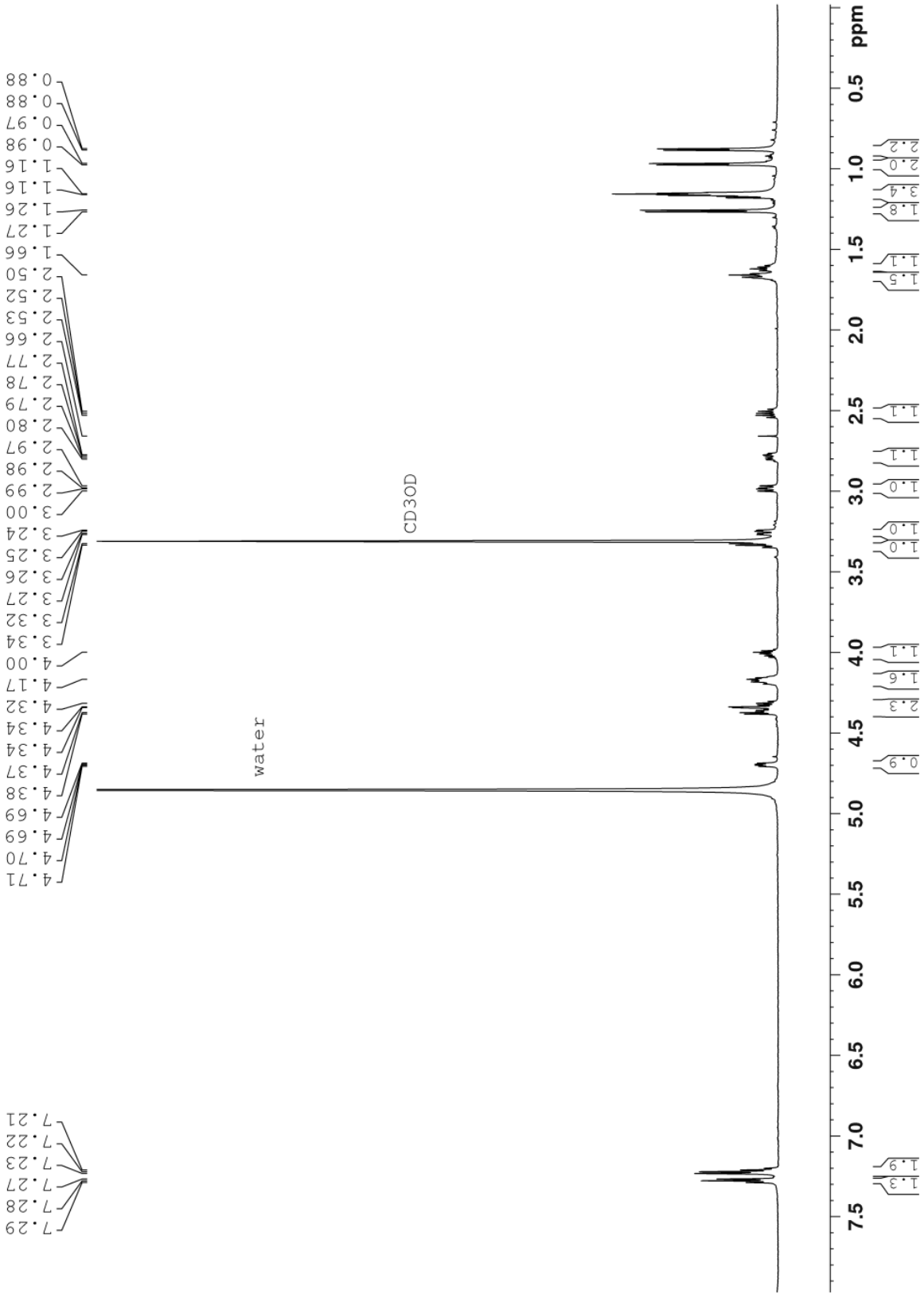
HSQC spectrum of macyranone A (**1**) in CD_3OD, 700 MHz



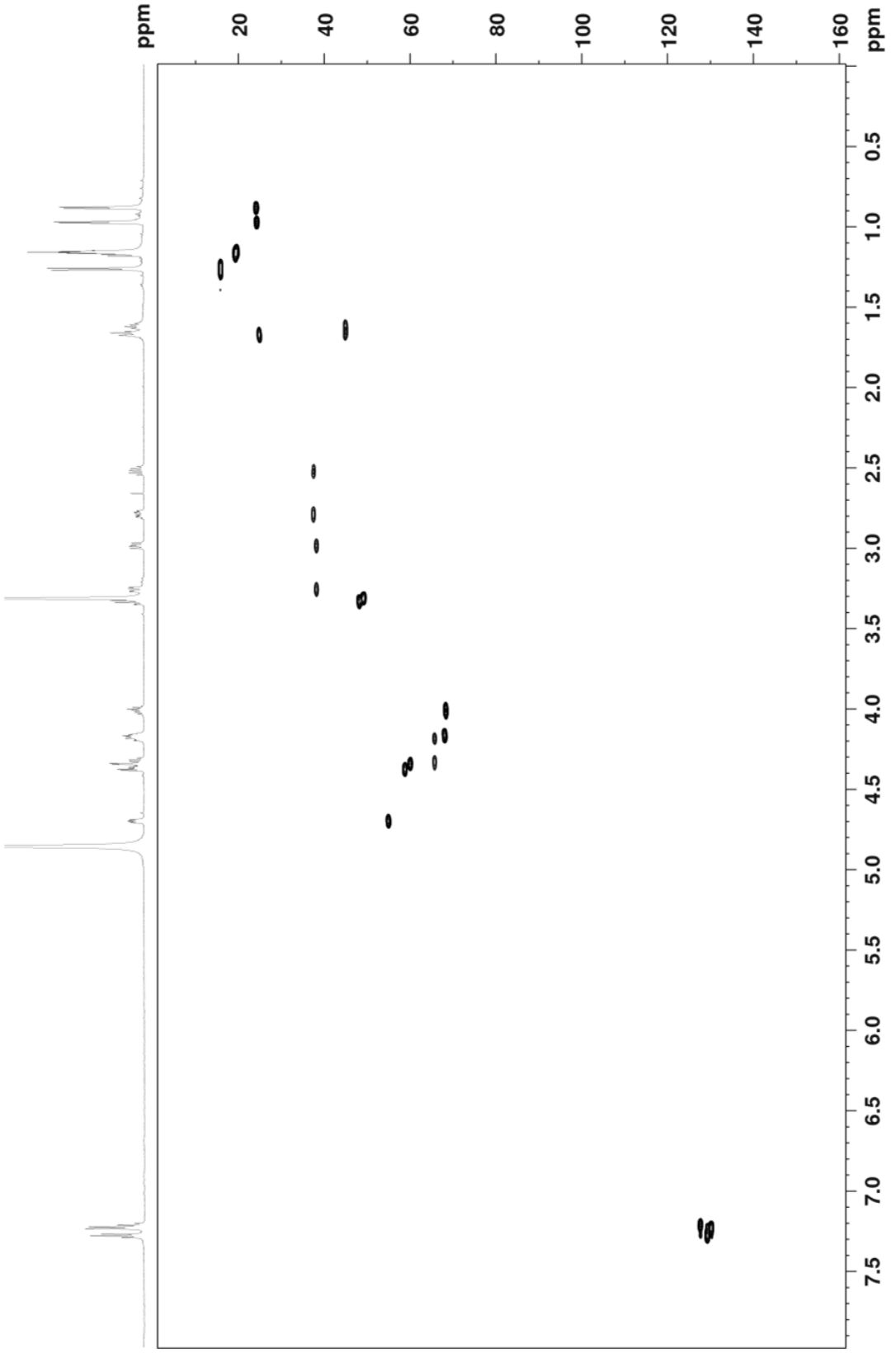
HMBC spectrum of macyranone A (**1**) in CD_3OD, 700 MHz



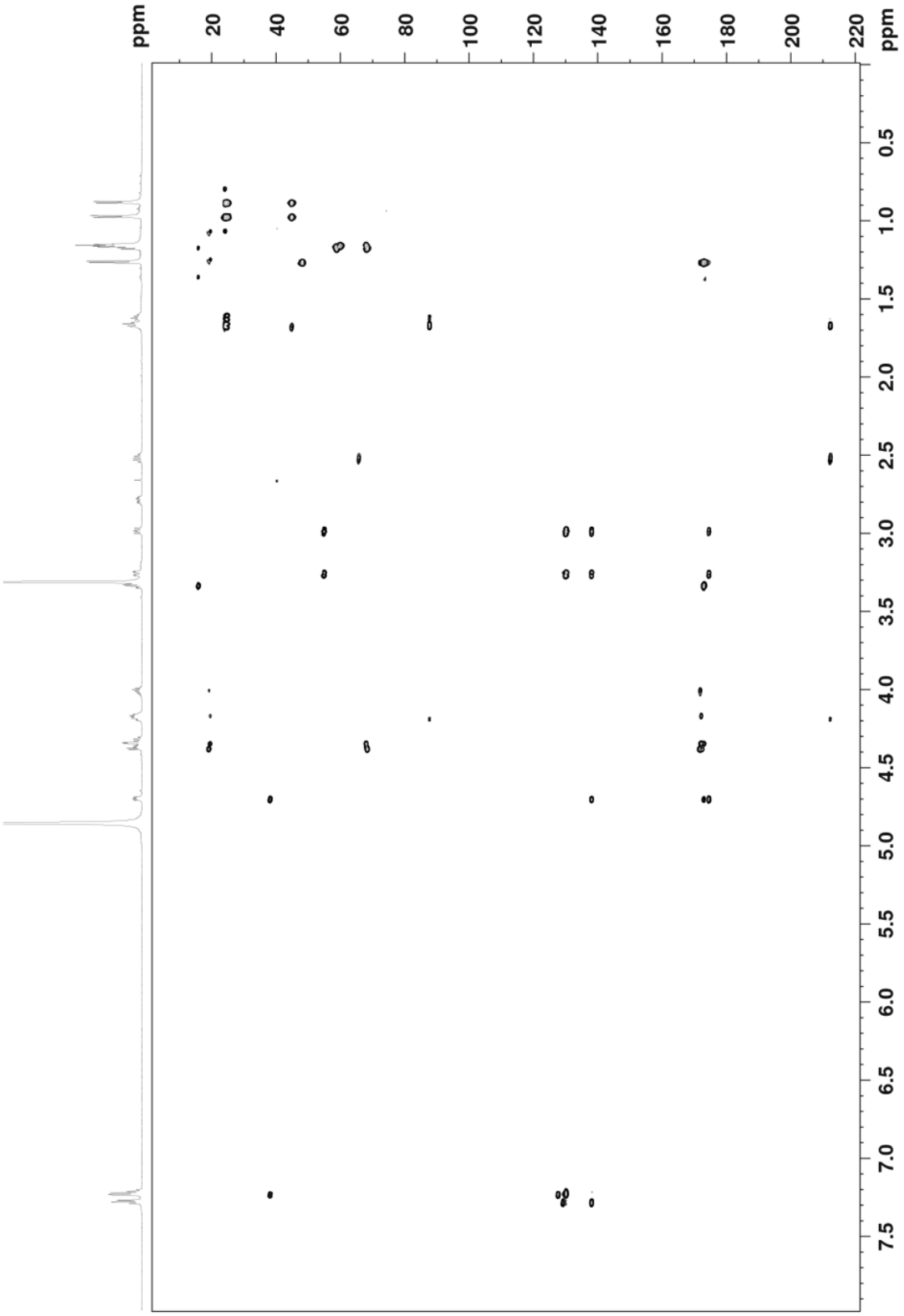
COSY spectrum of macyranone A (**1**) in CD_3OD, 700 MHz



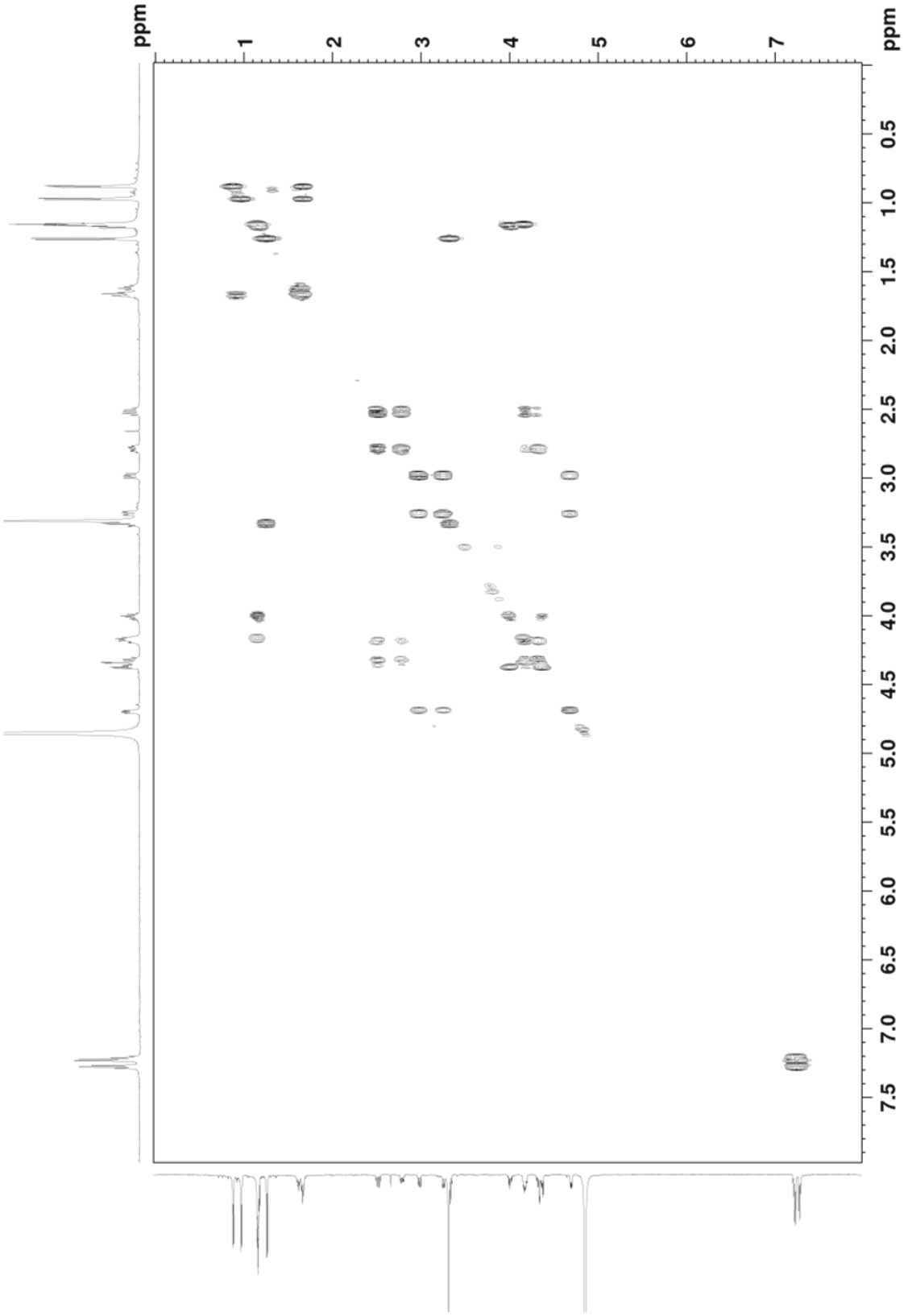
^{1}H NMR spectrum of macyranone B (**2**) in CD_3OD, 700 MHz



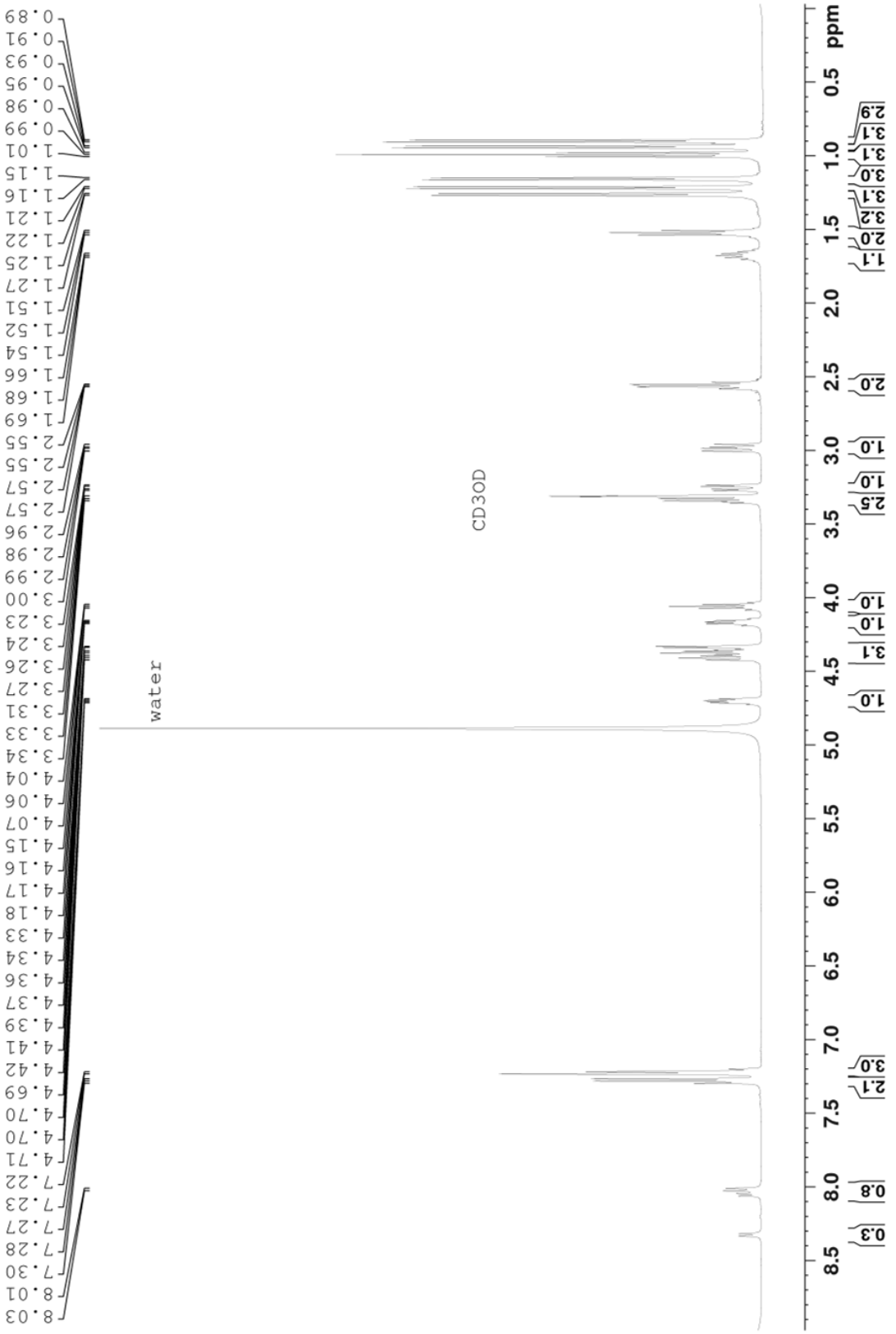
HSQC spectrum of macyranone B (**2**) in CD_3OD, 700 MHz



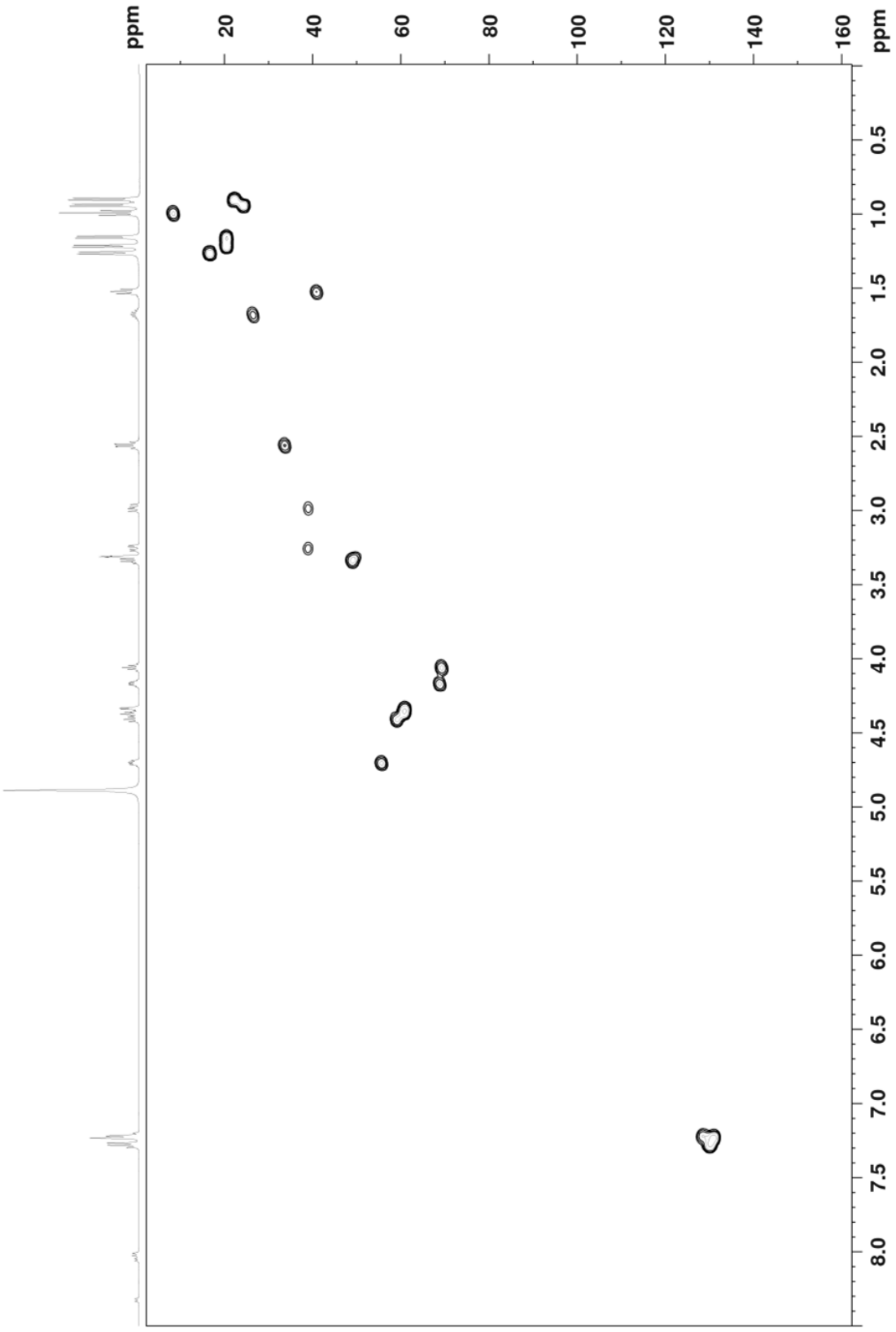
HMBC spectrum of macyranone B (**2**) in CD_3OD, 700 MHz



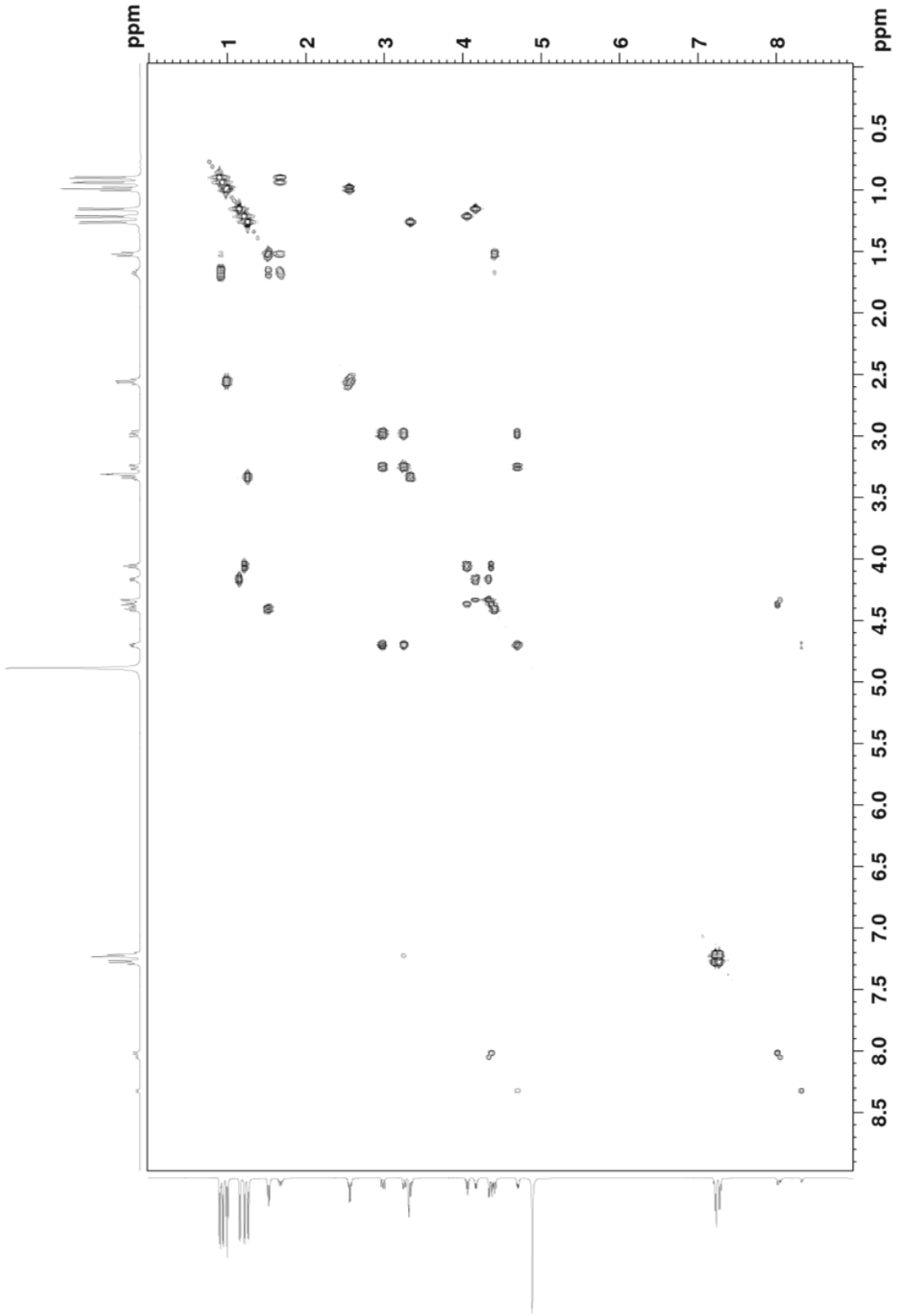
COSY spectrum of macyranone B (**2**) in CD_3OD, 700 MHz



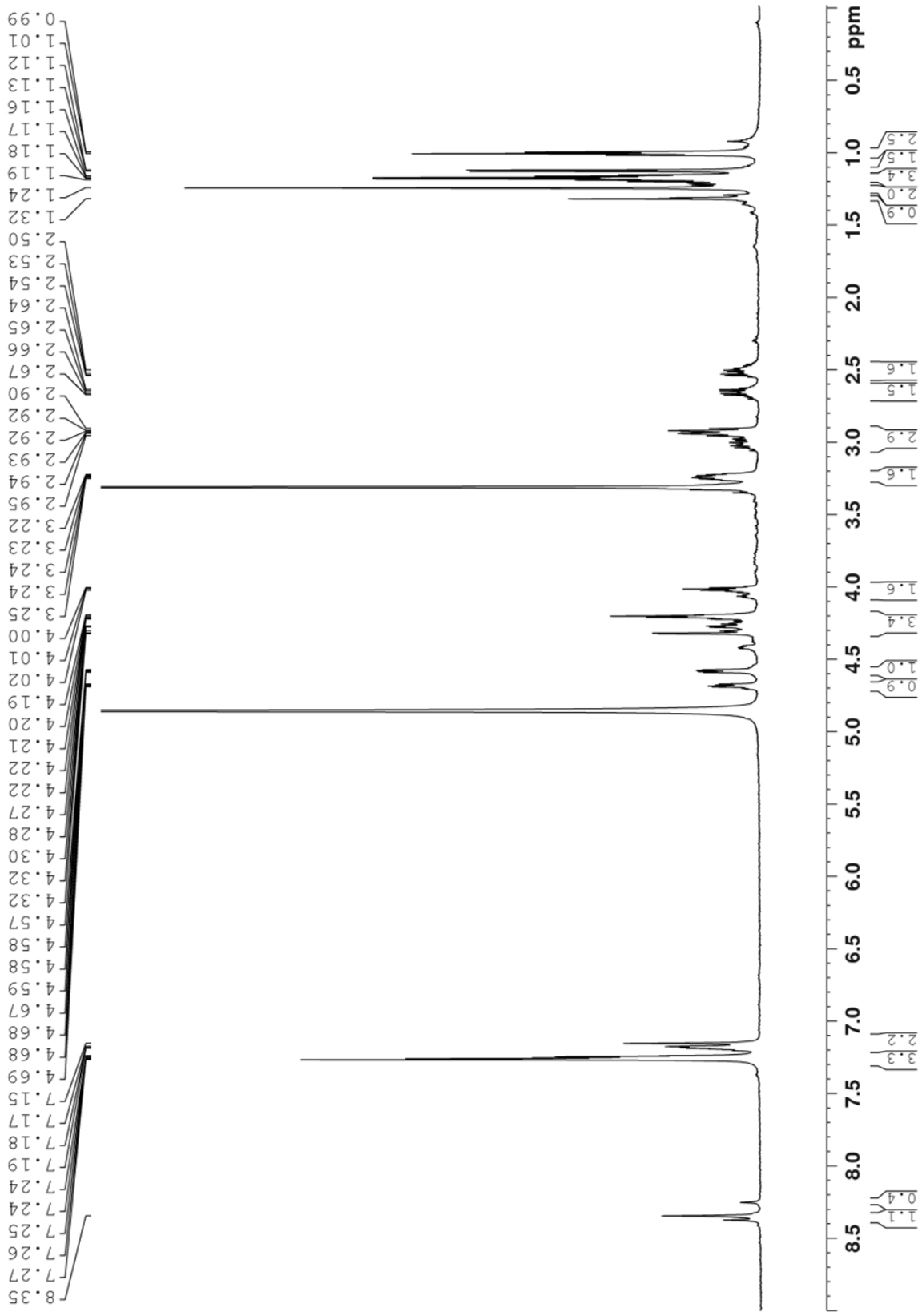
^{1}H NMR spectrum of macyranone C (**3**) in CD_3OD, 700 MHz



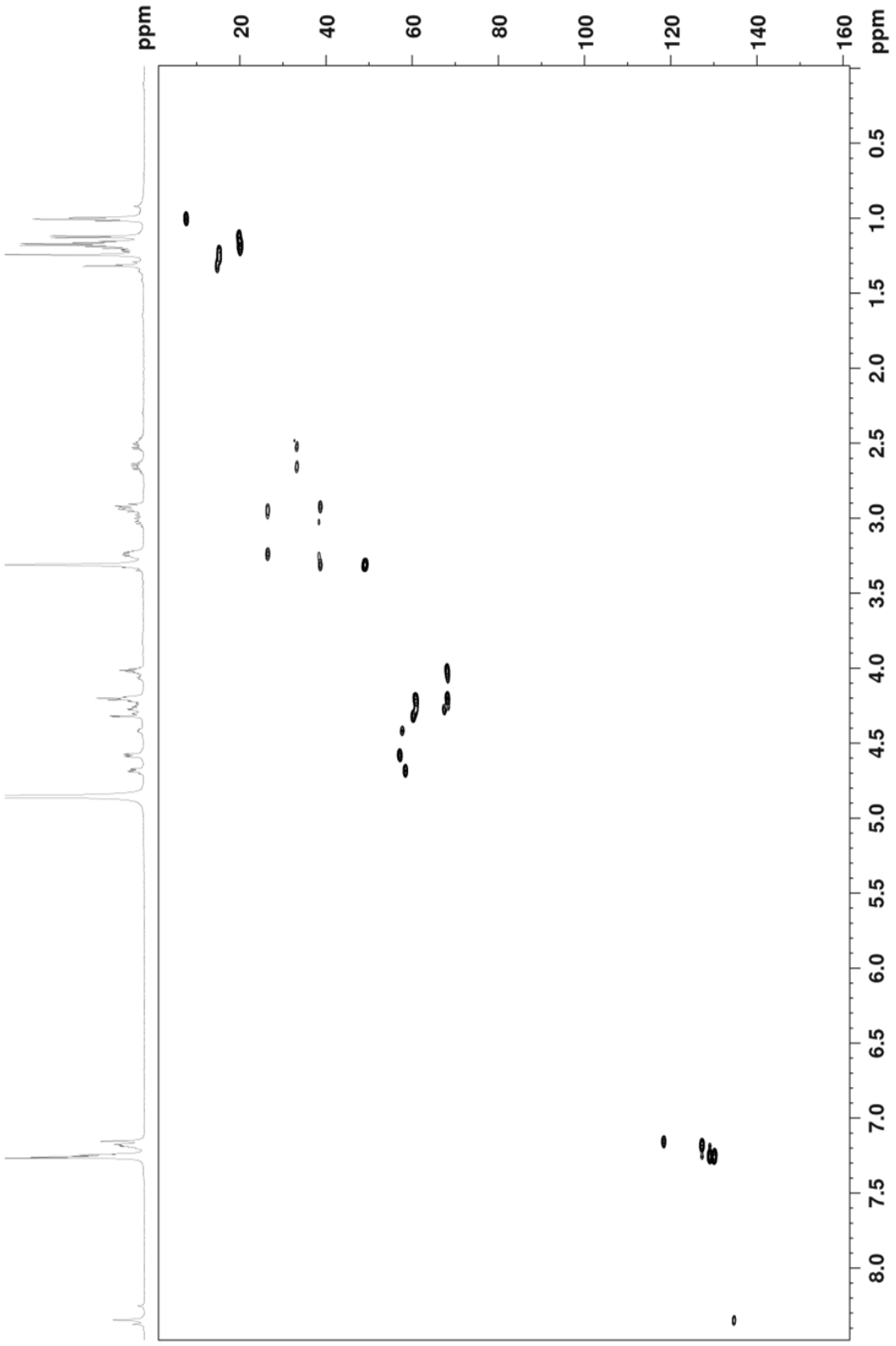
HSQC spectrum of macyranone C (**3**) in CD_3OD, 700 MHz



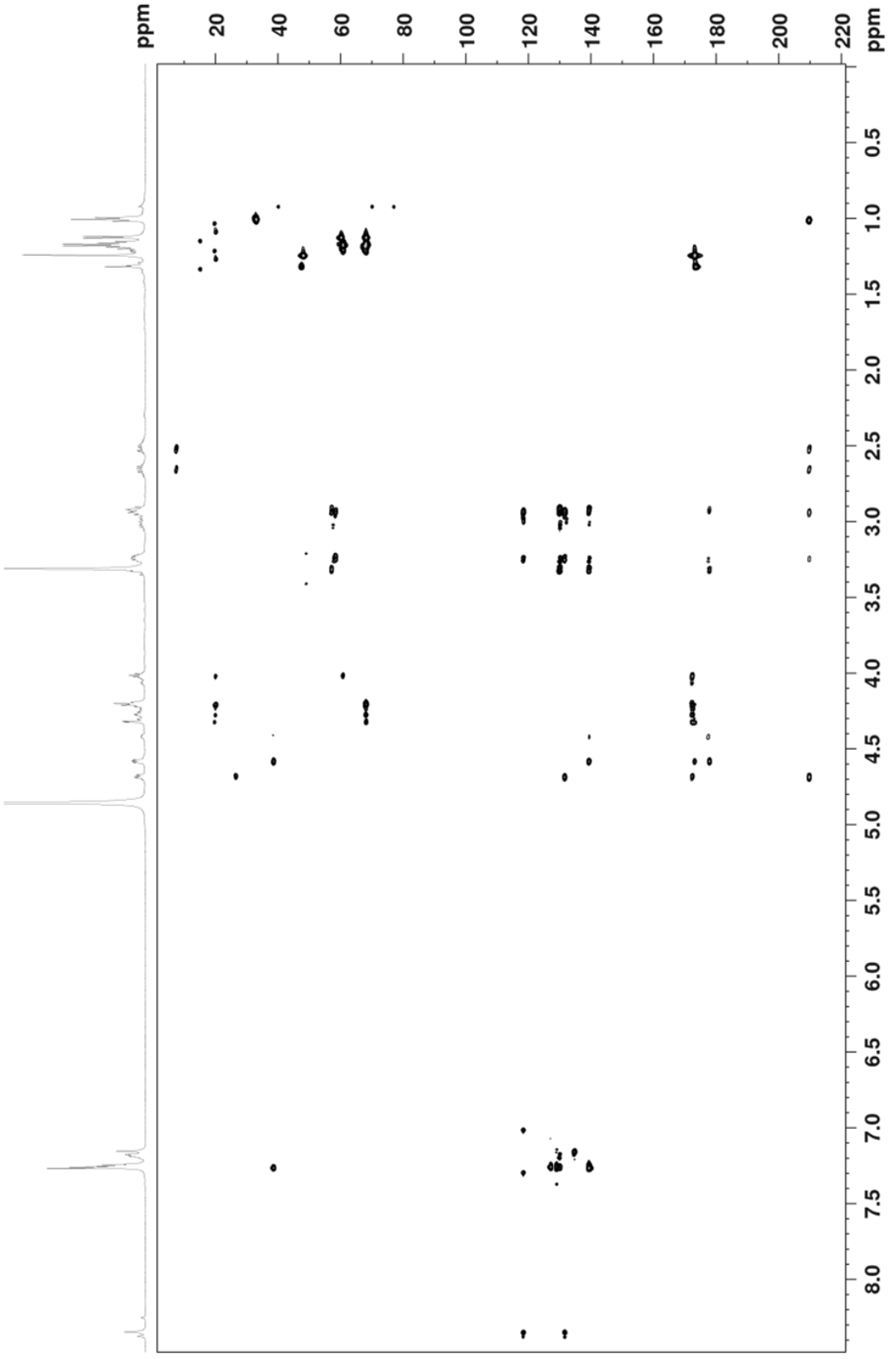
COSY spectrum of macyranone C (**3**) in CD_3OD, 700 MHz



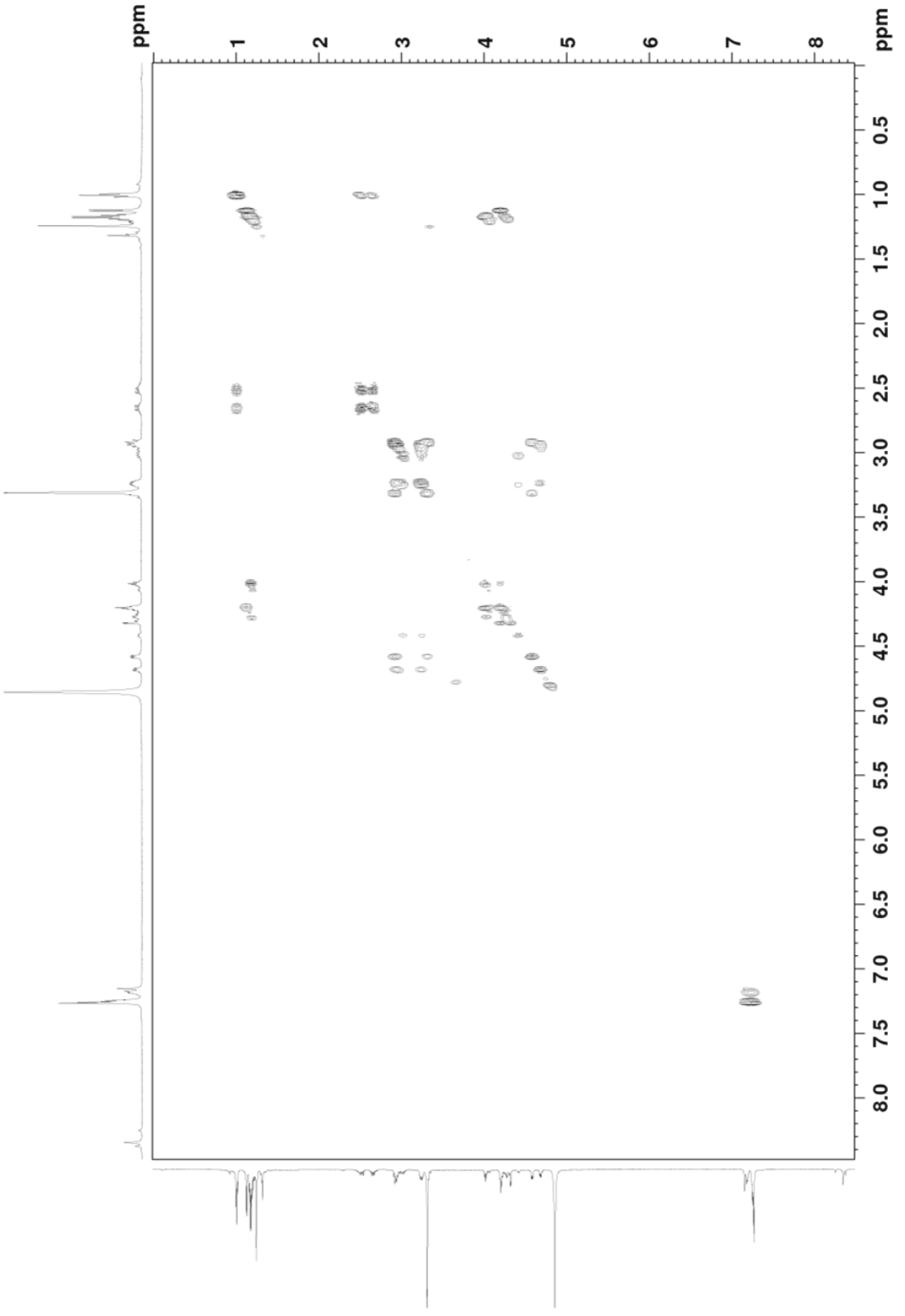
^{1}H NMR spectrum of macyranone D (**4**) in CD_3OD, 700 MHz



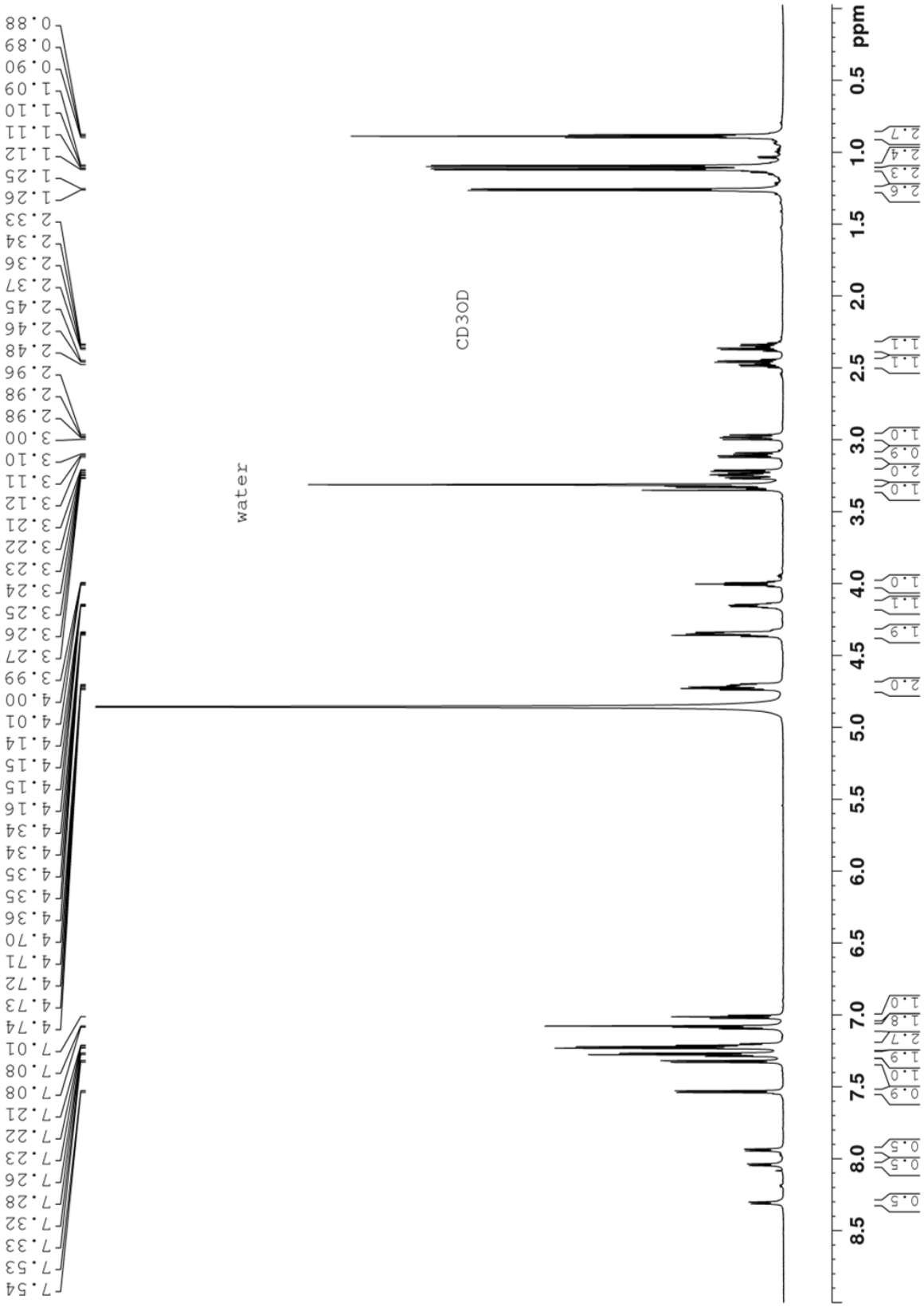
HSQC spectrum of macyranone D (**4**) in CD_3OD, 700 MHz



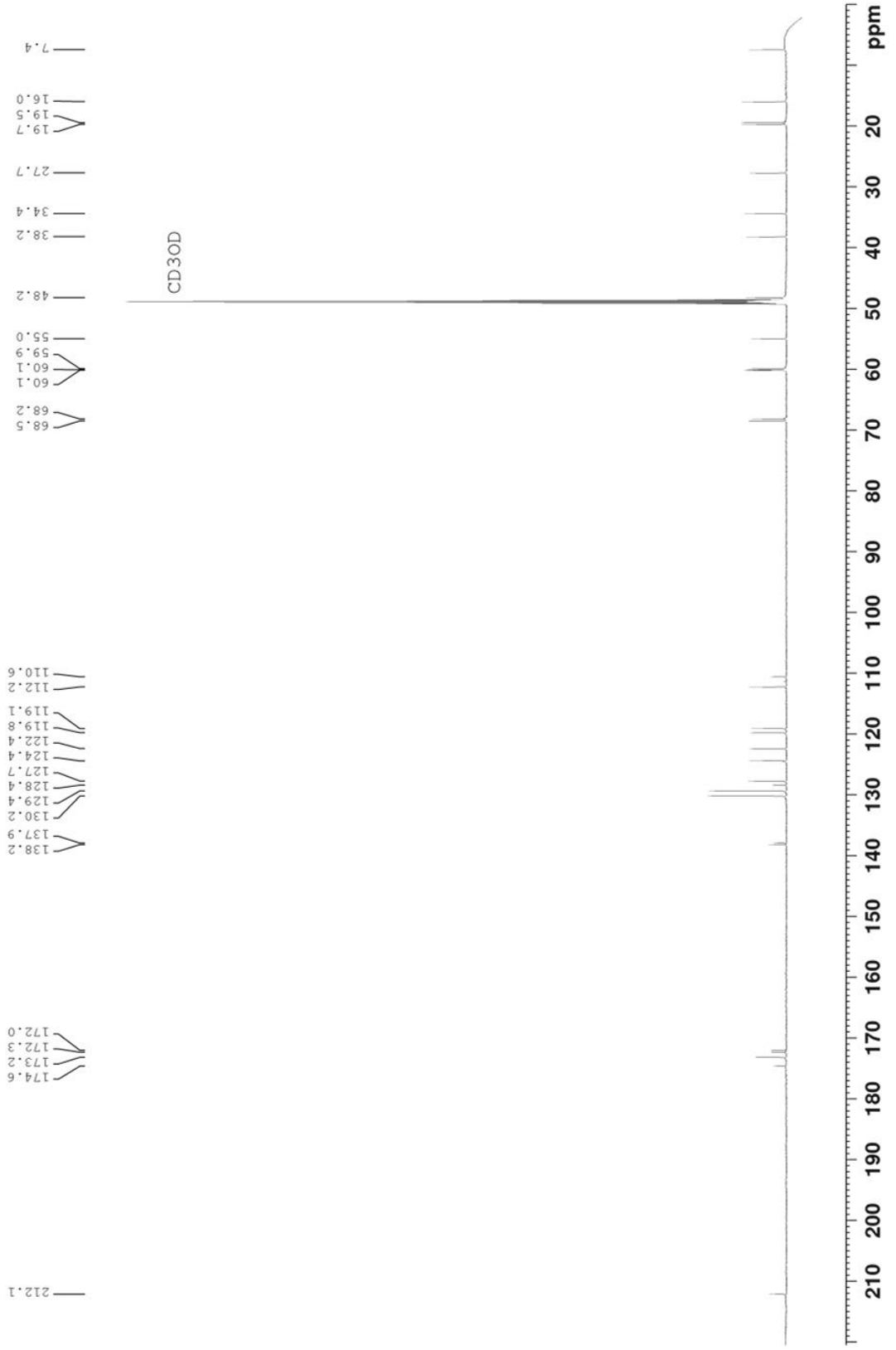
HMBC spectrum of macyranone D (**4**) in CD_3OD, 700 MHz



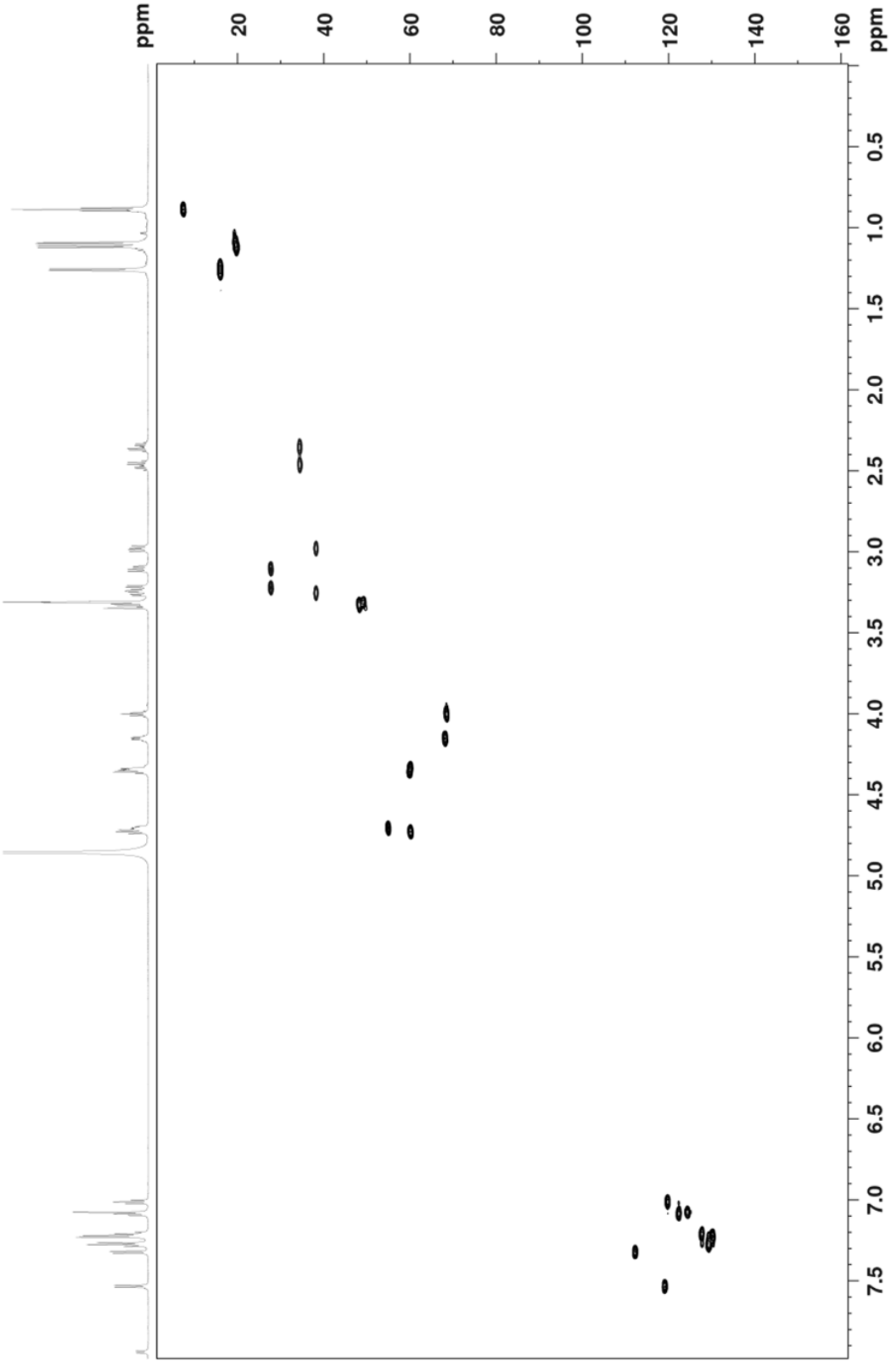
COSY spectrum of macyranone D (**4**) in CD_3OD, 700 MHz



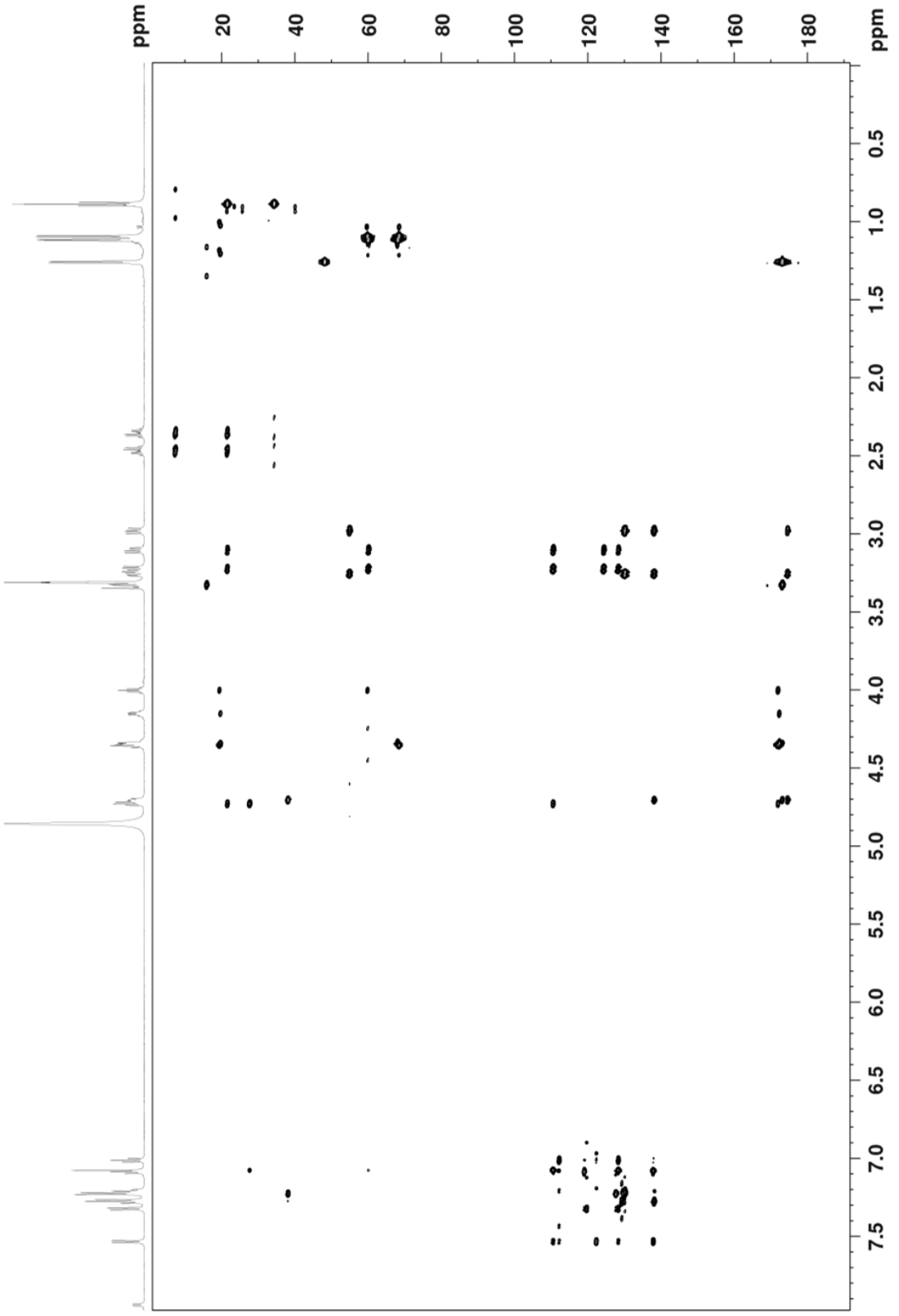
^{1}H NMR spectrum of macyranone E (**5**) in CD_3OD, 700 MHz



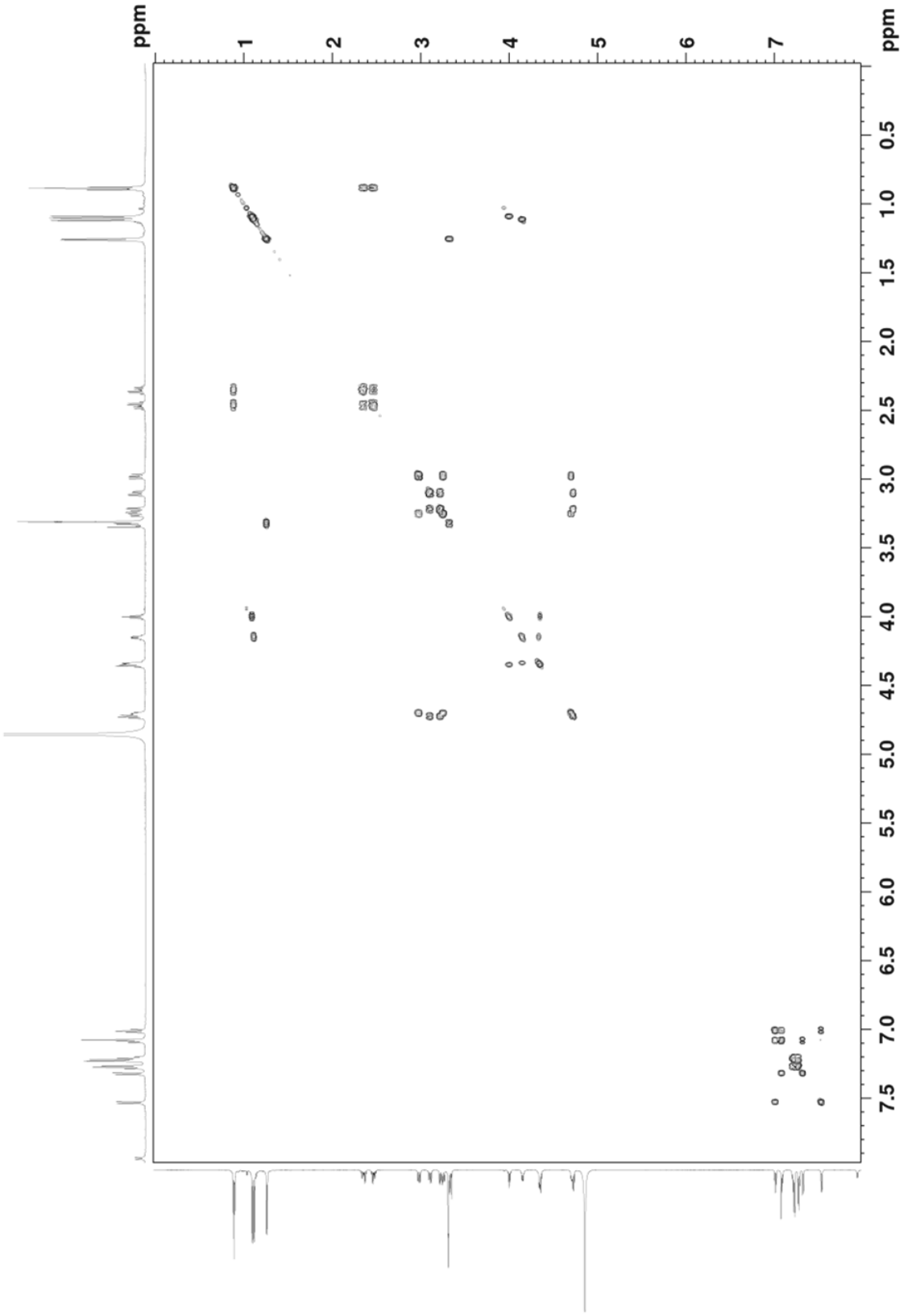
^{13}C NMR spectrum of macyranone E (**5**) in CD_3OD, 125 MHz



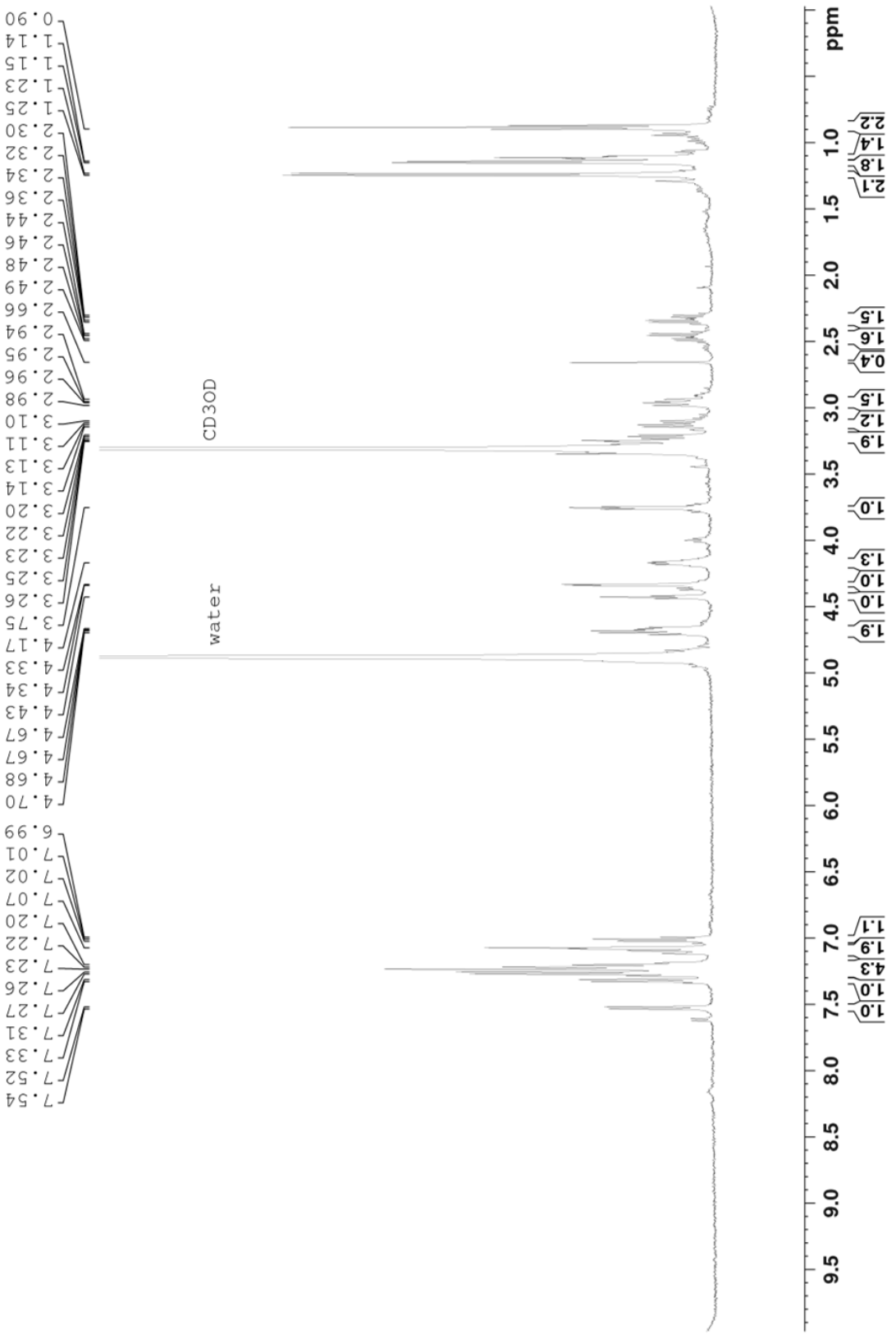
HSQC spectrum of macyranone E (**5**) in CD_3OD, 700 MHz



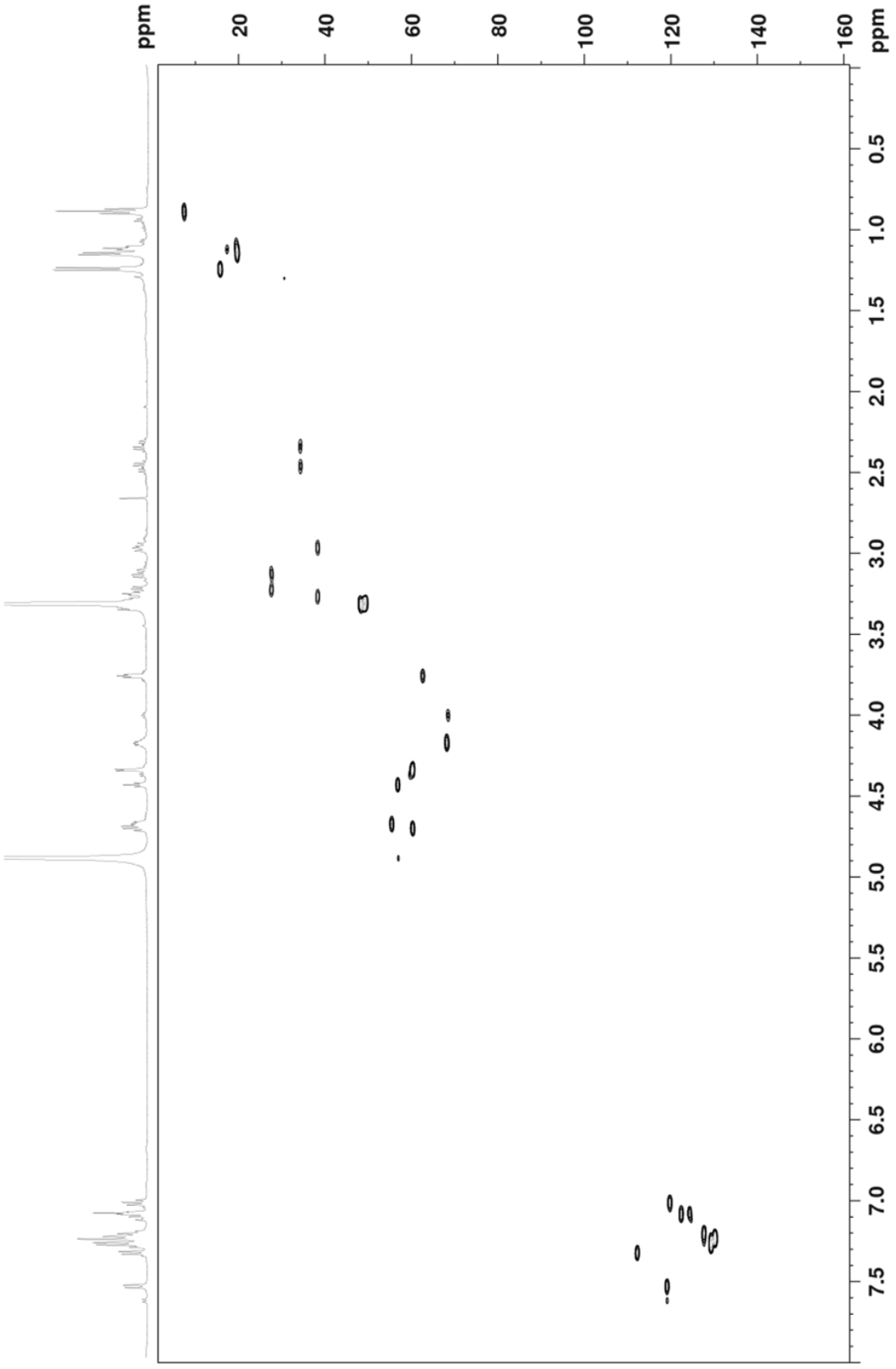
HMBC spectrum of macyranone E (**5**) in CD_3OD, 700 MHz



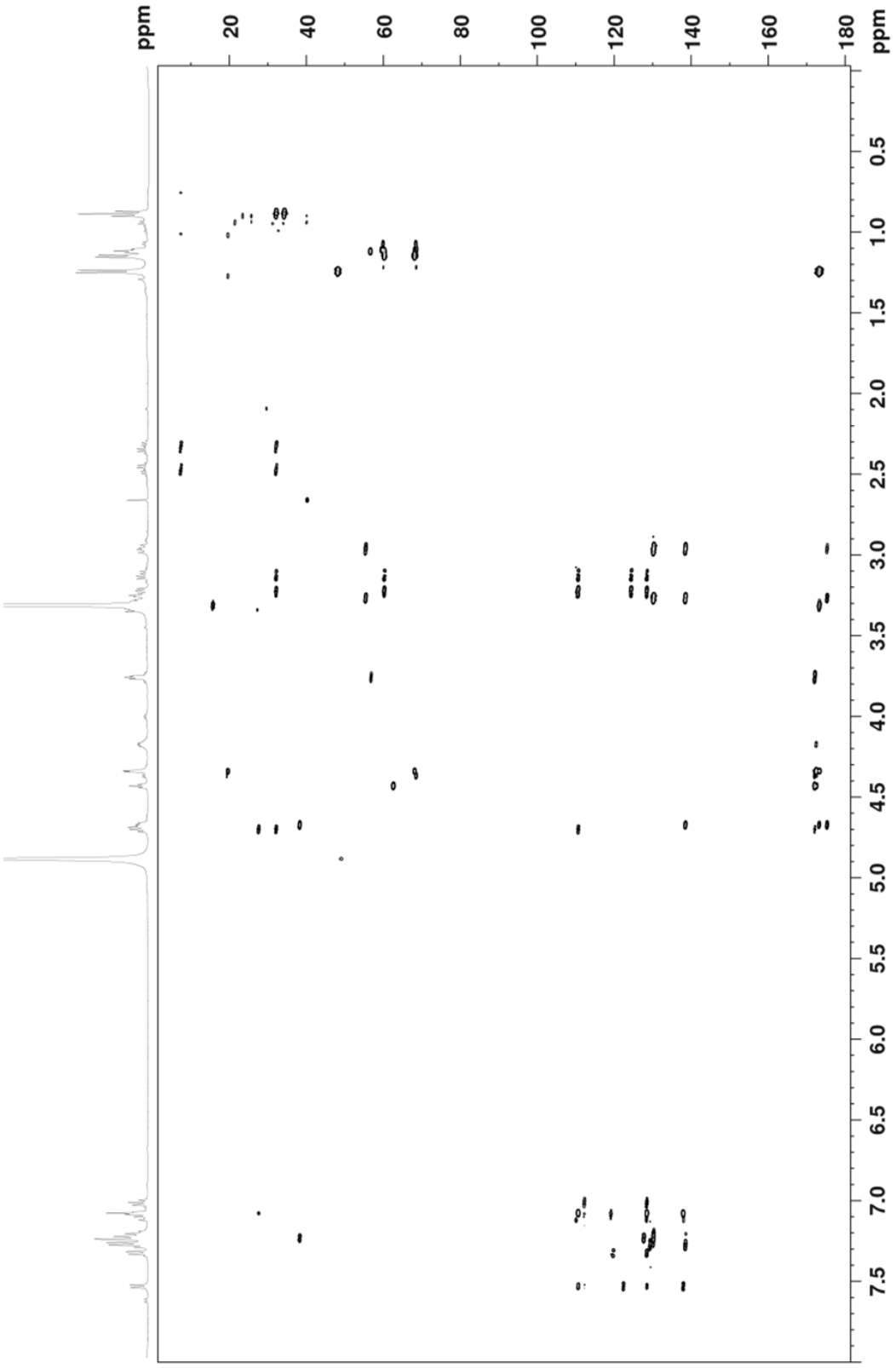
COSY spectrum of macyranone E (**5**) in CD_3OD, 700 MHz



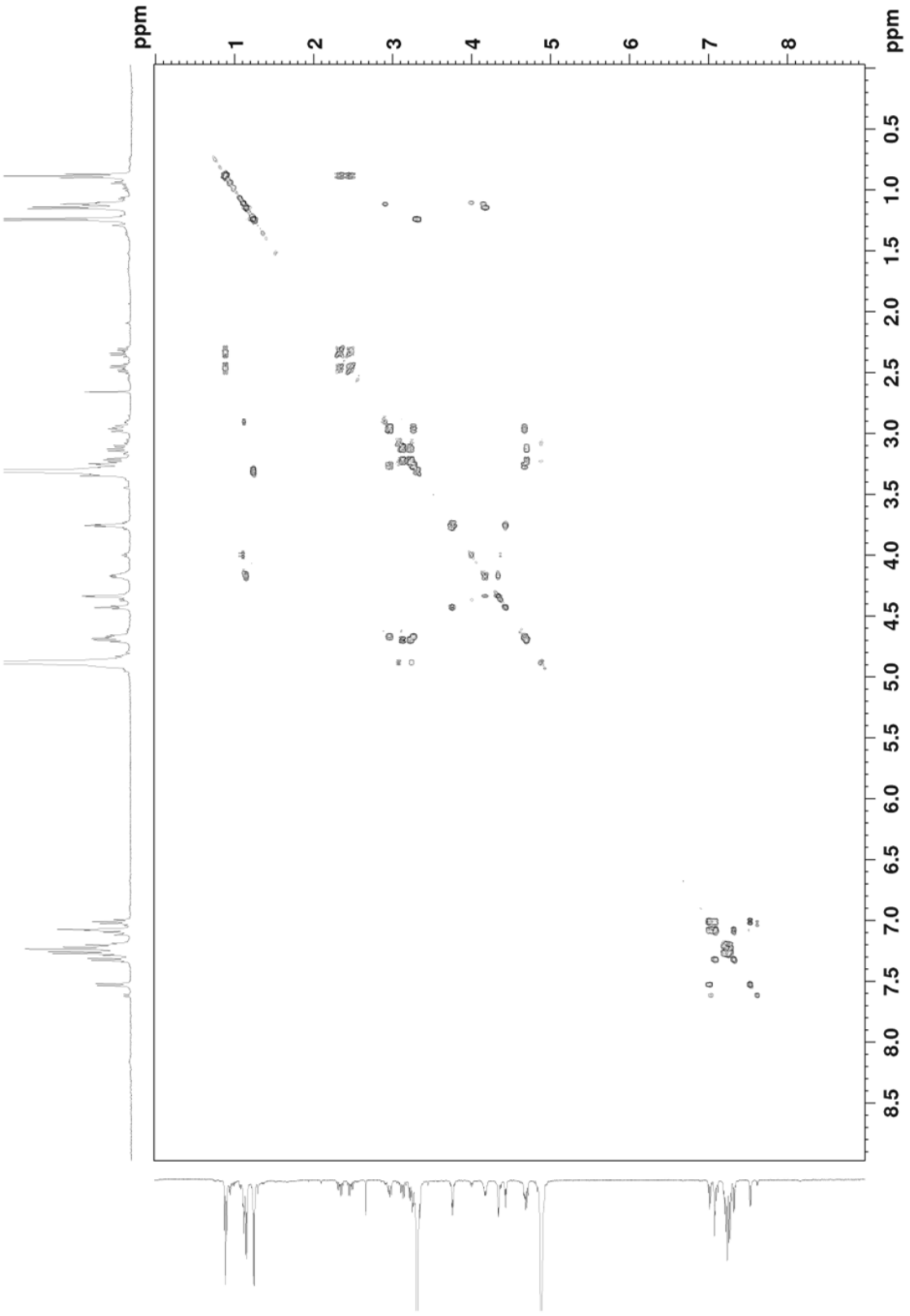
^{1}H NMR spectrum of macyranone F (**6**) in CD_3OD, 700 MHz



HSQC spectrum of macyranone F (**6**) in CD_3OD, 700 MHz



HMBC spectrum of macyranone F (**6**) in CD_3OD, 700 MHz



COSY spectrum of macyranone F (**6**) in CD_3OD, 700 MHz

3.7 References

(1) Huber, E.M.; Groll, M.: Inhibitors for the immuno- and constitutive proteasome: Current and future trends in drug development. *Angew. Chemie - Int. Ed.* **2012,.**

(2) Myung, J.; Kim, K.B.; et al.: Lack of proteasome active site allostery as revealed by subunit-specific inhibitors. *Mol. Cell* **2001**, *7*, 411–20.

(3) Hanada, M.; Sugawara, K.; et al.: Epoxomicin, a new antitumor agent of microbial origin. *J. Antibiot.* **1992**, *45*, 1746–52.

(4) Fenical, W.; Jensen, P.R.; et al.: Discovery and development of the anticancer agent salinosporamide A (NPI-0052). *Bioorganic Med. Chem.* **2009**, *17*, 2175–2180.

(5) Moore, B.S.; Eustáquio, A.S.; McGlinchey, R.P.: Advances in and applications of proteasome inhibitors. *Curr. Opin. Chem. Biol.* **2008**, *12*, 434–40.

(6) Kreidenweiss, A.; Kremsner, P.G.; Mordmüller, B.: Comprehensive study of proteasome inhibitors against Plasmodium falciparum laboratory strains and field isolates from Gabon. *Malar. J.* **2008**, *7*, 187.

(7) Glenn, R.J.; Pemberton, A.J.; et al.: Trypanocidal effect of alpha',beta'-epoxyketones indicates that trypanosomes are particularly sensitive to inhibitors of proteasome trypsin-like activity. *Int. J. Antimicrob. Agents* **2004**, *24*, 286–9.

(8) Groll, M.; Ditzel, L.; et al.: Structure of 20S proteasome from yeast at 2.4 A resolution. *Nature* **1997**, *386*, 463–471.

(9) Huber, E.M.; Basler, M.; et al.: Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity. *Cell* **2012**, *148*, 727–738.

(10) Groll, M.; Kim, K.B.; et al.: Crystal Structure of Epoxomicin:20S Proteasome Reveals a Molecular Basis for Selectivity of α',β'-Epoxyketone Proteasome Inhibitors. *J. Am. Chem. Soc.* **2000**, *122*, 1237–1238.

(11) Sugawara, K.; Hatori, M.; et al.: Eponemycin, a new antibiotic active against B16 melanoma. I. Production, isolation, structure and biological activity. *J. Antibiot.* **1990**, *43*, 8–18.

(12) Koguchi, Y.; Nishio, M.; et al.: TMC-89A and B, new proteasome inhibitors from streptomyces sp. TC 1087. *J. Antibiot.* **2000**, *53*, 967–972.

(13) Koguchi, Y.; Kohno, J.; et al.: TMC-86A, B and TMC-96, new proteasome inhibitors from Streptomyces sp. TC 1084 and Saccharothrix sp. TC 1094. I. Taxonomy, fermentation, isolation, and biological activities. *J. Antibiot.* **1999**, *52*, 1069–76.

(14) Pereira, A.R.; Kale, A.J.; et al.: The carmaphycins: new proteasome inhibitors exhibiting an α,β-epoxyketone warhead from a marine cyanobacterium. *ChemBioChem* **2012**, *13*, 810–7.

(15) Trivella, D.B.B.; Pereira, A.R.; et al.: Enzyme inhibition by hydroamination: design and mechanism of a hybrid carmaphycin-syringolin enone proteasome inhibitor. *Chem. Biol.* **2014**, *21*, 782–91.

(16) Schorn, M.; Zettler, J.; et al.: Genetic basis for the biosynthesis of the pharmaceutically important class of epoxyketone proteasome inhibitors. *ACS Chem. Biol.* **2014**, *9*, 301–9.

(17) Gerth, K.; Irschik, H.; et al.: Myxothiazol, an antibiotic from myxococcus fulvus. (myxobacterales) I. cultivation, isolation, physico-chemical and biological properties. *J. Antibiot.* **1980**, *33*, 1474–1479.

(18) Wenzel, S.C.; Müller, R.: The biosynthetic potential of myxobacteria and their impact in drug discovery. *Curr. Opin. Drug Discov. Devel.* **2009**, *12*, 220–30.

(19) Bode, H.B.; Müller, R.: Analysis of myxobacterial secondary metabolism goes molecular. *J. Ind. Microbiol. Biotechnol.* **2006**, *33*, 577–88.

(20) Plaza, A.; Viehrig, K.; et al.: Jahnellamides, α-keto-β-methionine-containing peptides from the terrestrial myxobacterium Jahnella sp.: structure and biosynthesis. *Org. Lett.* **2013**, *15*, 5882–5.

(21) Nadmid, S.; Plaza, A.; et al.: Hyalachelins A-C, unusual siderophores isolated from the terrestrial myxobacterium Hyalangium minutum. *Org. Lett.* **2014**, *16*, 4130–3.

(22) Etzbach, L.; Plaza, A.; et al.: Cystomanamides: structure and biosynthetic pathway of a family of glycosylated lipopeptides from myxobacteria. *Org. Lett.* **2014**, *16*, 2414–7.

(23) Harada, K.; Fujii, K.; et al.: A method usingL/CMS for determination of absolute configuration of constituent amino acids in peptide --- advanced Marfey's method ---. *Tetrahedron Lett.* **1995**, *36*, 1515–1518.

(24) Matsumori, N.; Kaneno, D.; et al.: Stereochemical Determination of Acyclic Structures Based on Carbon-Proton Spin-Coupling Constants. A Method of Configuration Analysis for Natural Products. *J. Org. Chem.* **1999**, *64*, 866–876.

(25) Blin, K.; Medema, M.H.; et al.: antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **2013**, *41*, W204–12.

(26) Altschul, S.F.; Gish, W.; et al.: Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–10.

(27) Kim, B.C.; Lee, J.M.; et al.: Cloning, sequencing, and characterization of the pradimicin biosynthetic gene cluster of Actinomadura hibisca P157-2. *J. Microbiol. Biotechnol.* **2007**, *17*, 830–839.

(28) Chen, Y.; Wendt-Pienkowski, E.; et al.: Characterization of FdmV as an amide synthetase for fredericamycin A biosynthesis in Streptomyces griseus ATCC 43944. *J. Biol. Chem.* **2010**, *285*, 38853–60.

(29) Richards, N.G.; Schuster, S.M.: Mechanistic issues in asparagine synthetase catalysis. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1998**, *72*, 145–98.

(30) Riley, M.; Abe, T.; et al.: Escherichia coli K-12: a cooperatively developed annotation snapshot--2005. *Nucleic Acids Res.* **2006**, *34*, 1–9.

(31) Zalkin, H.: Advances in Enzymology and Related Areas of Molecular Biology. **1993**, (Meister, A., Ed.) John Wiley & Sons, Inc., Hoboken, NJ, USA,.

(32) Larsen, T.M.; Boehlein, S.K.; et al.: Three-Dimensional Structure of Escherichia coli Asparagine Synthetase B: A Short Journey from Substrate to Product. *Biochemistry* **1999**, *38*, 16146–16157.

(33) Duitman, E.H.; Hamoen, L.W.; et al.: The mycosubtilin synthetase of Bacillus subtilis ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 13294–9.

(34) Maruyama, C.; Toyoda, J.; et al.: A stand-alone adenylation domain forms amide bonds in streptothricin biosynthesis. *Nat. Chem. Biol.* **2012**, *8*, 791–7.

(35) Buntin, K.; Weissman, K.J.; Müller, R.: An unusual thioesterase promotes isochromanone ring formation in ajudazol biosynthesis. *ChemBioChem* **2010**, *11*, 1137–46.

(36) Zhang, J.; Zhang, W.; et al.: Cloning and functional characterization of ACAD-9, a novel member of human acyl-CoA dehydrogenase family. *Biochem. Biophys. Res. Commun.* **2002**, *297*, 1033–1042.

(37) Daubner, S.C.; Gadda, G.; et al.: Cloning of nitroalkane oxidase from Fusarium oxysporum identifies a new member of the acyl-CoA dehydrogenase superfamily. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 2702–2707.

(38) Lei, B.; Shiao-Chun, T.U.: Gene overexpression, purification, and identification of a desulfurization enzyme from Rhodococcus sp. strain IGTS8 as a sulfide/sulfoxide monooxygenase. *J. Bacteriol.* **1996**, *178*, 5699–5705.

(39) Ghisla, S.; Thorpe, C.: Acyl-CoA dehydrogenases. A mechanistic overview. *Eur. J. Biochem.* **2004**, *271*, 494–508.

(40) Gaitatzis, N.; Kunze, B.; Müller, R.: In vitro reconstitution of the myxochelin biosynthetic machinery of Stigmatella aurantiaca Sg a15: Biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 11136–11141.

(41) Du, L.; Lou, L.: PKS and NRPS release mechanisms. *Nat. Prod. Rep.* **2010**, *27*, 255–78.

(42) Luscher, A.; de Koning, H.; Maser, P.: Chemotherapeutic Strategies Against Trypanosoma brucei: Drug Targets vs. Drug Targeting. *Curr. Pharm. Des.* **2007**, *13*, 555–567.

(43) Arastu-Kapur, S.; Anderl, J.L.; et al.: Nonproteasomal targets of the proteasome inhibitors bortezomib and carfilzomib: a link to clinical adverse events. *Clin. Cancer Res.* **2011**, *17*, 2734–43.

(44) Fujii, K.; Ikai, Y.; et al.: A Nonempirical Method Using LC/MS for Determination of the Absolute Configuration of Constituent Amino Acids in a Peptide: Combination of Marfey's Method with Mass Spectrometry and Its Practical Application. *Anal. Chem.* **1997**, *69*, 5146–5151.

(45) Punta, M.; Coggill, P.C.; et al.: The Pfam protein families database. *Nucleic Acids Res.* **2012**, *40*, D290–301.

(46) Röttig, M.; Medema, M.H.; et al.: NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **2011**, *39*, W362–7.

(47) Bachmann, B.O.; Ravel, J.: Chapter 8. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. *Methods Enzymol.* **2009**, *458*, 181–217.

(48) Ziemert, N.; Podell, S.; et al.: The natural product domain seeker NaPDoS: a phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. *PLoS One* **2012**, *7*, e34064.

(49) Baltz, T.; Baltz, D.; et al.: Cultivation in a semi-defined medium of animal infective forms of Trypanosoma brucei, T. equiperdum, T. evansi, T. rhodesiense and T. gambiense. *EMBO J.* **1985**, *4*, 1273–7.

(50) Räz, B.; Iten, M.; et al.: The Alamar Blue assay to determine drug sensitivity of African trypanosomes (T.b. rhodesiense and T.b. gambiense) in vitro. *Acta Trop.* **1997**, *68*, 139–47.

(51) Huber, W.; Koella, J.C.: A comparison of three methods of estimating EC50 in studies of drug resistance of malaria parasites. *Acta Trop.* **1993**, *55*, 257–61.

(52) Cunningham, I.: New culture medium for maintenance of tsetse tissues and growth of trypanosomatids. *J. Protozool.* **1977**, *24*, 325–9.

(53) Mikus, J.; Steverding, D.: A simple colorimetric method to screen drug cytotoxicity against Leishmania using the dye Alamar Blue. *Parasitol. Int.* **2000**, *48*, 265–9.

(54) Yang, M.; Arai, C.; et al.: Fluorinated rhodacyanine (SJL-01) possessing high efficacy for visceral leishmaniasis (VL). *J. Med. Chem.* **2010**, *53*, 368–73.

(55) Desjardins, R.E.; Canfield, C.J.; et al.: Quantitative assessment of antimalarial activity in vitro by a semiautomated microdilution technique. *Antimicrob. Agents Chemother.* **1979**, *16*, 710–8.

(56) Matile, H.; Pink, J.R.L.: Plasmodium falciparum malaria parasite cultures and their use in immunology. in *Immunological Methods* **1990**, (Lefkovits, I., and Pernis, B., Eds.) Academic Press, San Diego, pp 221–234.

(57) Ponnudurai, T.; Leeuwenberg, A.D.; Meuwissen, J.H.: Chloroquine sensitivity of isolates of Plasmodium falciparum adapted to in vitro culture. *Trop. Geogr. Med.* **1981**, *33*, 50–4.

(58) Page, B.; Page, M.; Noel, C.: A new fluorometric assay for cytotoxicity measurements in-vitro. *Int. J. Oncol.* **1993**, *3*, 473–6.

(59) Ahmed, S.A.; Gogal, R.M.; Walsh, J.E.: A new rapid and simple non-radioactive assay to monitor and determine the proliferation of lymphocytes: an alternative to [3H]thymidine incorporation assay. *J. Immunol. Methods* **1994**, *170*, 211–24.

(60) Rachid, S.; Huo, L.; et al.: Mining the cinnabaramide biosynthetic pathway to generate novel proteasome inhibitors. *ChemBioChem* **2011**, *12*, 922–31.

(61) Groll, M.; Huber, R.: Purification, crystallization, and x-ray analysis of the yeast 20S proteasome. *Methods Enzymol.* **2005**, *398*, 329–336.

(62) Kabsch, W.: XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125–132.

(63) Potterton, E.; Briggs, P.; et al.: A graphical user interface to the CCP4 program suite. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **2003**, *59*, 1131–1137.

(64) Turk, D.: MAIN software for density averaging, model building, structure refinement and validation. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1342–1357.

(65) Stachelhaus, T.; Mootz, H.D.; Marahiel, M.A.: The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* **1999**, *6*, 493–505.

(66) Yadav, G.; Gokhale, R.S.; Mohanty, D.: Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *J. Mol. Biol.* **2003**, *328*, 335–63.

4 Discussion

The main achievement of this thesis was the identification and full characterization of two novel natural product classes from the myxobacterium *Cystobacter fuscus* MCy9118 as well as their association with the corresponding biosynthetic pathways. It includes several aspects of the field of natural products (NP) research. The myxobacterial strain was deeply analyzed with regard to the produced compounds to profile the biosynthetic potential of the strain. As a consequence, the thesis deals with the isolation of natural products using a chemical screening approach based on LC-MS-SPE-NMR as well as the full *de novo* structural characterization of the identified natural products. This goes along with the identification and characterization of the biosynthetic machinery that underlies the production of the diverse compounds and studies concerning their biological activity.

4.1 The NP Isolation Approach – a Question of Faith?

The effectivity of a chemical screening approach for new molecular structures using hyphenated chromatographic techniques was clearly proven in this work but as illustrated in the introduction, several basic approaches are applied in the search for new natural products. So what are the advantages and disadvantages of a biological screening approach, a chemical screening approach or a genome mining approach and what are the explicit reasons to apply the one or the other?

The search for new natural products can nicely be compared to the well-known concept of a needle in a haystack without even knowing how the needle looks like. In microbial crude extracts, the diverse target molecules represent only a tiny fraction. Many approaches have been applied to isolate new natural products but not every approach can be equally promising in all cases and it is a matter of fact that the applied approach is the decisive factor that differentiates between success and failure. The central question is which approach is the one to apply in a specific case? Therefore, some fundamental considerations have to be taken into account which evaluate the factual circumstances and therefore allow a decision:

1) The objective needs to be clearly defined whether for example the isolation of a compound that shows a specific bioactivity, that acts on a specific target or that exhibits a new chemical scaffolds is the aim.

2) Considerations about the local circumstances are necessary like the available techniques since for example a chemical approach depends more on high-quality analytical techniques whereas for a biological screening one needs a constant access to biological testing.
3) The knowledge about the producing organism is crucial. Is the organism culturable and does it produce secondary metabolites under laboratory conditions? Are already secondary metabolites known from the organism? Are genome sequence data available or accessible which would be a requirement for a genome-based approach?

If the objective, in our case the isolation of natural products that exhibit new chemical scaffolds, is defined and the local circumstances are considered, the third point indicates that the NP isolation approach is still a case-to-case decision for every organism. It strongly depends on the prerequisites that every organism, that is a candidate for the screening process, provides.

4.1.1 The NP isolation approach in case of *Cystobacter fuscus* MCy9118

The decision-making process concerning the NP isolation strategy is preceded by initial testing of the organisms for culturability and bioactivity as well as MS based dereplication to identify already known compounds.

C. fuscus was found to exhibit a wide range of biological activities against Gram-positive and Gram-negative bacteria as well as fungi. Together with a good culturability allowing the production of sufficient material, this would enable a traditional bioactivity-guided isolation approach. However, the dereplication process using the "Myxobase" based on retention time in a defined analytical system, high-resolution *m/z* value and isotope pattern fit[1] showed the presence of already known secondary metabolites. This method identified the typical myxobacterial catechole siderophores myxochelin A[2] and B[3], that show weak activity against some Gram-positive bacteria.

Additionally, the strain produced the unstable roimatacene, a compound with moderate activity only against Gram-negative bacteria.[4] Due to the instability of the compound, the MS based dereplication method failed in the first line and only manual inspection of the produced LC-MS data in combination with NMR data from LC-NMR resulted in a correct assignment of the compound. The last known antibiotic produced by the strain was

althiomycin exhibiting a broad and potent antimicrobial activity against several Gram-positive and Gram-negative bacteria (Figure 4.2-C).[5,6]

Even if the production of these diverse compounds with such biological activities is exciting with respect to the enormous biosynthetic potential of this specific strain, particularly the wide range of bioactivities of the known compounds made an assignment of the bioactivity to a specific compound impossible. Therefore, the application of a bioactivity-guided isolation approach was unfavorable. Additionally, the genome sequence of the strain had not been known from the outset, so that the decision was made in favor of a chemical screening approach based on LC-SPE-NMR-MS (Figure 4.1).

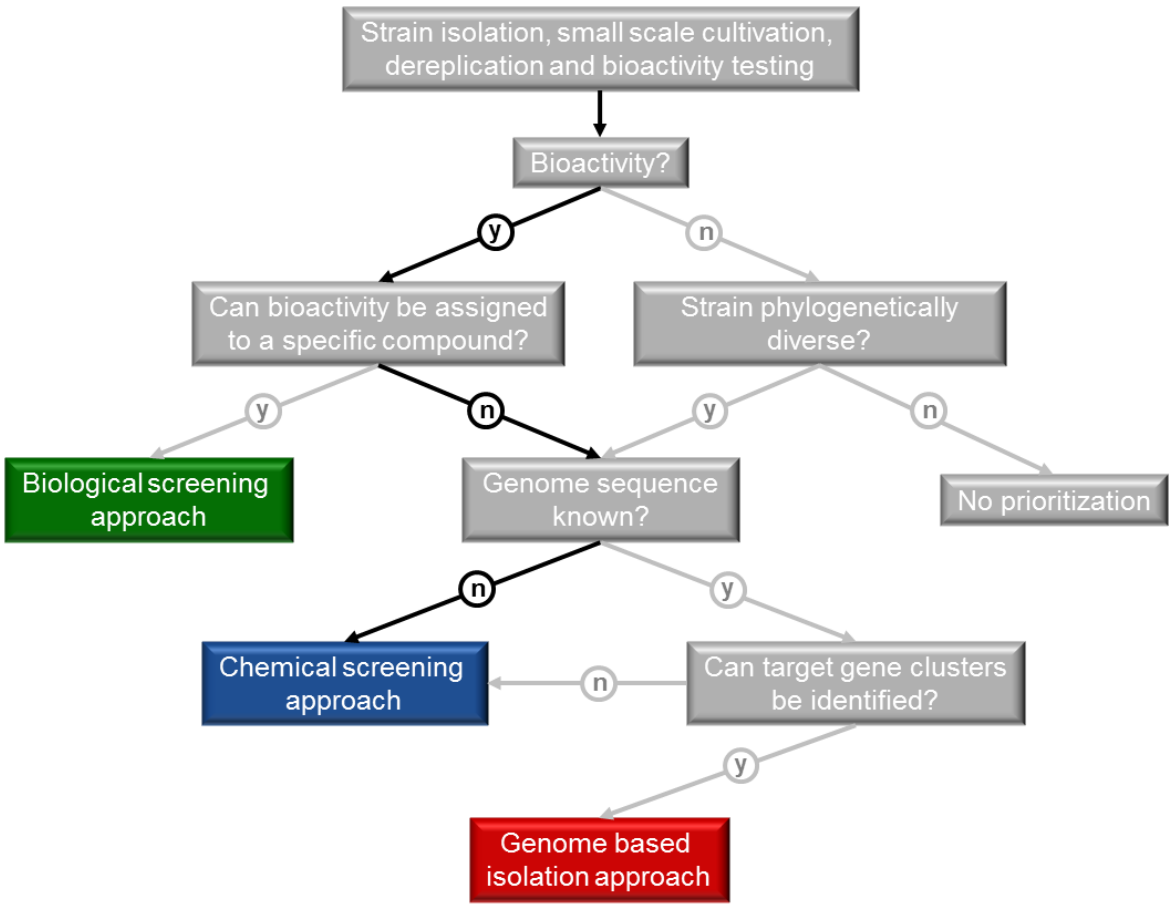


Figure 4.1 Decision tree of natural product isolation approaches with regard to the strategy used for *Cystobacter fuscus* MCy9118. Green: biological screening; blue: chemical screening; red: genome mining: y, yes; n, no.

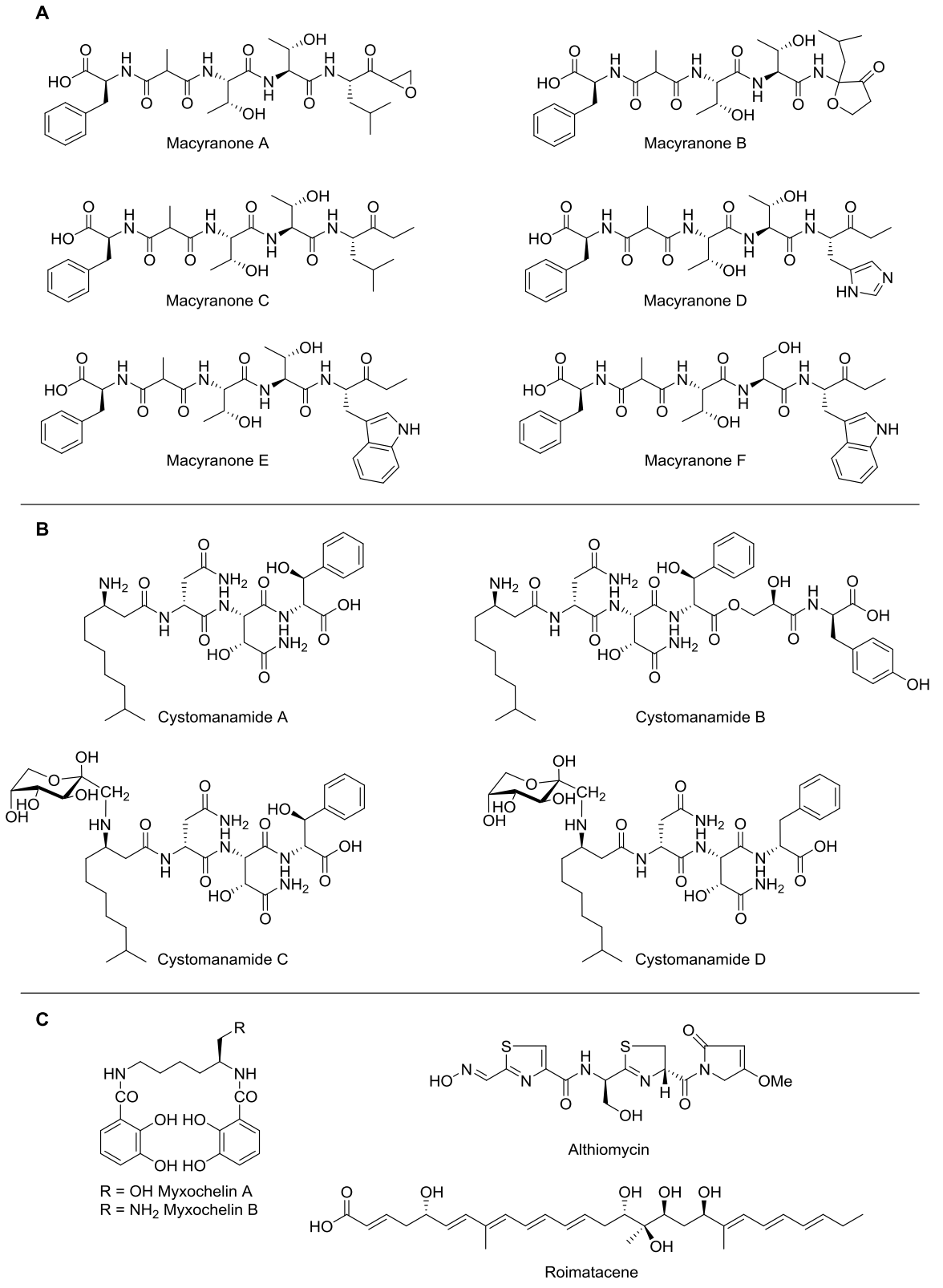


Figure 4.2: Representation of the full chemical potential of *Cystobacter fuscus* MCy9118. A: Chemical structures of the macyranones. B: Chemical structures of the cystomanamides. C: Chemical structures of known secondary metabolites produced by the strain.

In the scope of MS based dereplication, various peaks belonging to unknown compounds were identified. In this approach, the hyphenated LC-NMR technique was used to identify interesting target compounds out of these unknown compounds produced by the strain. Because of the work with small-scale cultivation as well as the intrinsic low sensitivity of NMR spectroscopy even with a cryogenic flow probe, the amount of compounds present in the crude extract was critical for the analysis by LC-NMR. The crude extract was analyzed and as mentioned above, the first success was the rapid dereplication of the known antibiotic roimatacene. Besides that, several peaks were analyzed and two compounds were chosen for further analysis based on their interesting NMR profile exhibiting signals for peptidic as well as aliphatic residues. These two compounds were later named macyranone B and cystomanamide A. Using the NMR data obtained from LC-SPE-NMR, it was possible to identify isolated fragments of the compounds and get an impression of their overall nature. Nevertheless, it was necessary to scale-up the cultivation conditions to receive a sufficient amount of extract for further processing. The crude extract underwent a fractionation procedure composed of size-exclusion chromatography and successive rounds of reversed-phase liquid chromatography to isolate the target compounds determined in the LC-NMR run. Structure elucidation of both compounds showed that both describe new molecular structures with unusual chemical features proving that the chemical screening approach was undoubtedly the right choice in case of *C. fuscus* MCy9118.

4.2 From Bacteria to Compounds

The main issue in the characterization of new natural products from myxobacteria is often the low yield and therefore limited amount of available compound.[7-9] Even if our NMR system with a 700 MHz magnet and a proton optimized 'inverse' cryoprobe is an excellent basis for the measurements of minimum quantities, the extremely low amounts still complicate the compound characterization. In addition to the high-tech instrumentation, we use high-quality magnetic susceptibility matched 5mm NMR tubes, so called Shigemi tubes, to minimize the required amount of solvent and maximize the concentration of the compound for the measurements.[10] Nevertheless, optimization of cultivation conditions and repeated isolation procedures are most of the time unavoidable to accumulate sufficient compound quantities.

4.2.1 Finding the best cultivation conditions

In case of the major compounds of the two new compound classes, named macyranone B and cystomanamide A, the yield was sufficient from the beginning to isolate enough material from a 6 L cultivation to analyze the compounds and elucidate the planar structures as described in detail in Chapter 2 and 3. With less than 0.1 mg/L, the yield was still inadequate for further studies and required an optimization of the production conditions to allow the full characterization of the compounds in terms of configuration as well as biological activity. Inspection of the MS^2 fragmentation data of the crude extract also indicated the presence of related compounds of both core structures being present in trace amounts. This kind of compound classes covering various chemical variants of one basic scaffold is commonly found in myxobacteria.[11] Examples of such diverse compound classes from myxobacteria are the chondramides,[12,13] the tubulysins[14,15], the recently published cystobactamids[16] as well as the microsclerodermins.[17] To increase the production of the compounds and to enable their isolation, varying cultivation conditions were tested using different media, cultivation times, shake flask cultivation in baffled and normal flasks as well as fermentation. The best cultivation conditions were different for the two compound classes, but a yeast-based production medium[18] turned out to be the best choice in both cases. Even if the best growth of the strain was observed in M-medium containing phytone and maltose as main components, the presence of other cells (autoclaved yeast cells) as well as the limited amount of nutrition in the yeast-based VY/2S medium could act as a stress factor for the myxobacterium and therefore induce the production of secondary metabolites.[19] An increase of the iron concentration in the production medium to 8 mg/L Fe-EDTA resulted in a significant decrease of the produced amount of the siderophores myxochelin A and B, circumventing the tedious separation of these compounds.

The macyranones reached their optimum production in shake flask cultivation with 0.4 % dry yeast instead of 0.5 % baker's yeast (VY/2S-DY medium) at a shaking rate of 160 rpm. A shorter cultivation time of 6 to 8 days was sufficient for production of the compounds. Surprisingly, using the shorter cultivation time macyranone B was replaced by macyranone A as most abundant derivative. Using a fermenter instead of shake flask cultivation, the macyranone production dropped significantly. Therefore, the shake-flask cultivation was applied for the production of larger amounts to avoid the tedious optimization of the fermentation conditions.

Cystomanamide A and B were produced almost equally in all cultivation conditions but the use of a fermenter triggered the production of the glycosylated derivatives cystomanamide C and D. The cystomanamides reached their highest level on day 10 of the fermentation. Cultivation in a 100 L fermenter using the tested conditions conducted by Wolfgang Kessler at the Helmholtz-Centre for Infection Research in Braunschweig unfortunately did not result in the desired production so that a 10 L fermentation was the method of choice.

4.2.2 Isolation of the target compounds

Coming to the isolation of natural products, Richard J. P. Cannell found the right words to describe the challenge: "It can seem a formidable task, faced with a liter of fermentation broth - a dark, viscous sludge - knowing that in there is one group of molecules that has to be separated from all the rest."[20] Even if the fermentation broth of *C. fuscus* is more of a bright orange to pinkish color, the task remains the same.

The use of an adsorber resin is one step to facilitate the extraction procedure since it allows solid-phase extraction. The target compounds stick to the adsorber (given that it is the right one) and can therefore be easily separated from the fermentation broth. For both compound families, the addition of 2 % Amberlite XAD-7 adsorber resin was even advantageous for the yield. It increased the production of the secondary metabolites suggesting a feedback inhibition by the end products that can be prevented or reduced by the addition of an adsorber resin.[21] A similar feedback inhibition by the end-product was found in the biosynthesis of the epothilone.[22] All compounds but cystomanamide B were found attached to the XAD, only cystomanamide B was yielded by extraction of the centrifuged cell mass.

4.2.3 Characteristics of the compounds

The new natural products of the class of cystomanamides are characterized by an unusual 3-amino-9-methyldecanoic acid (AMDA) residue, which is *N*-glycosylated in case of cystomanamide C and D, and several non-proteinogenic amino acid residues. These comprise β-hydroxylated amino acids namely β-hydroxy asparagine (β-OH-Asn) and β-hydroxy phenylalanine (β-OH-Phe) as well as D-amino acids like D-asparagine. D-tyrosine is found in cystomanamide B. Additionally, cystomanamide B bears an unusual glyceric acid (GA) residue (Figure 4.2-B). There are only a few peptides known where a glyceric

acid residue is present[23–26] but among them, cystomanamide B is the only peptide, where the GA residue is integrated into the peptide chain and does not represent the N-terminal residue. Also the unprecedented 3-amino-9-methyldecanoic acid cannot be found in any known natural product. The residue is a branched chain fatty acid with a modification of the chain in the β-position. The incorporation of branched-chain fatty acids is a more common motif in lipopeptidic secondary metabolites as exemplified by tauramamide and polymyxin.[27,28] The presence of an amine in the β-position of an acyl chain is unusual but was shown and studied in detail for mycosubtilin, a natural product that belongs to the iturin family.[29]

The macyranones are a diverse class of peptidic compounds with an unusual 2-methylmalonamide (Mma) residue found in all macyranones. The compounds contain mainly L-amino acids sharing a phenylalanine (Phe) as well as two threonine residues Thr and Thr' whereas Thr' is substituted by serine (Ser) in macyranone F. Thr' reveals the unusual *erythro* 2S, 3S configuration. The macyranones all bear an additional variable amino acid residue that is substituted with an ethyl unit at the C terminus. Macyranone A to C contain L-leucine as C-terminal amino acid and in case of the most abundant derivative macyranone A, the α-amino ketone fragment is converted into an *a',β'*-epoxyketone. This compound was shown to be instable and undergoes a rearrangement of the epoxide ring resulting in the furanone ring found in macyranone B. Macyranone C-F bear the unsubstituted ethyl unit with L-histidine found in macyranone D and L-tryptophan in macyranone E and F (Figure 4.2-A).

Several other *a',β'*-epoxyketones are found in the diverse group of microbial natural products. Up to date, epoxyketones have mainly been isolated from actinobacteria[30–33] and with the recent discovery of the carmaphycins[34] also from cyanobacteria. The epoxyketone natural products are discussed in detail in chapter 3. As another structural element, the macyranones bear the 2-methylmalonamide residue. This highly unusual residue is only found in the proteasome inhibitors TMC-89A and B, isolated from *streptomyces* sp. TC 1087.[32] The macyranones are the first natural products where both functional groups of this residue are derivatized with an amino acid residue.

In total, both compound families represent unusual natural products with unprecedented structural elements.

4.3 From Compounds to Genes

To link a compound to its biosynthetic gene cluster, a retrobiosynthetic approach can be helpful to get an impression how the cluster could be assembled following textbook logic. Compound-specific structural elements like polyketide units or amino acid building blocks are considered and a hypothetic gene cluster is developed.

The power of *in silico*-based analysis of genomes has largely increased over the past years. Several tools are available for the detailed analysis of *in-silico* genome data. As the most useful tool, antiSMASH 2.0 allows the rapid identification, annotation and analysis of secondary metabolite biosynthetic gene clusters.[35,36] Besides the gene prediction and the gene cluster identification, the tool can provide details about NRPS and PKS functional annotation and predict a chemical structure as a possible product of the NRPS / PKS biosynthetic machinery.

Figure 4.3 shows the antiSMASH 2.0 structure prediction received for the first analysis of the genome data from *Cystobacter fuscus* MCy9118 for the candidate gene clusters from the cystomanamides and macyranones. As described on the antiSMASH webpage, "the structure is a rough prediction of the core scaffold based on assumed PKS/NRPS colinearity where tailoring reactions are not taken into account".[37] Comparing the predicted structures with the actual products, it is, on the one hand, striking that the prediction of cystomanamide is much more detailed than the macyranone structure prediction. On the other hand, both predictions show a sufficient similarity to the products to consider the respective gene clusters as candidates for the biosynthesis of both compound families. Detailed analysis of both gene clusters revealed that the cystomanamide biosynthesis follows more the textbook rules than the macyranone cluster, and therefore facilitating the correct structure prediction of the cystomanamides.

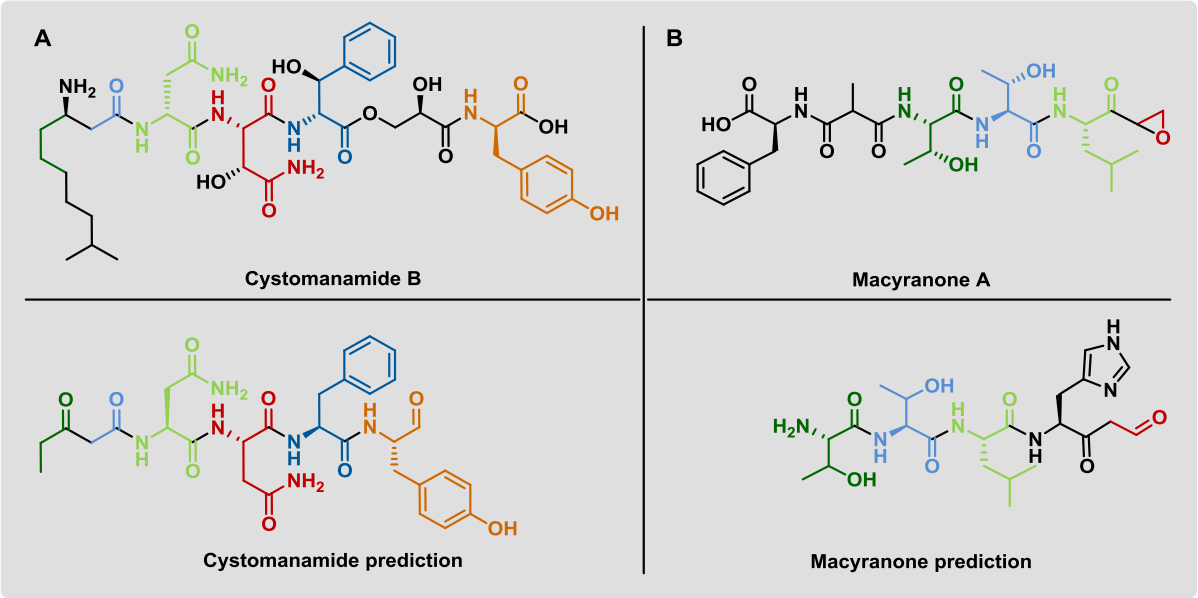


Figure 4.3: Comparison of the chemical structures of the actual product with the predicted structure for the biosynthetic gene clusters for A: cystomanamide and B: macyranone from antiSMASH 2.0 analysis. The colors reflect the correctly predicted building blocks.

To further analyze the catalytic domains and the A domain substrate specificity, the open reading frames were translated and analyzed using Pfam[38], NRPS predictor2[39], and PKS/NRPS Analysis[40]. Genes not resulting in predictions related to PKS or NRPS domains were analyzed via the BLAST algorithm[41] using the non-redundant sequence database at the National Center for Biotechnology Information (NCBI).

4.3.1 Cystomanamide biosynthetic gene cluster

The antiSMASH structure prediction based on the cystomanamide biosynthetic gene cluster was highly above average. The identified biosynthetic genes were in a good agreement with the predicted biochemical steps required for cystomanamide biosynthesis which is by far not always the case in myxobacterial biosynthetic systems.[42,43] The tool was capable to identify all regular PKS and NRPS domains as well as the fatty acid loading domains (see Chapter 2). Taking into account the identified domains that were not implemented into the predicted structure, all amino acids as well as their stereochemistry were concluded correctly. The stereochemical prediction is derived from the fact that NRPS modules incorporate mainly naturally occurring L-amino acids whereas epimerization domains present in the modules are responsible for a conversion into the respective D-amino acid. The identification of epimerization domains in the

respective modules fits perfectly with the presence and position of the D-amino acids D-asparagine, D-phenylalanine and D-tyrosine in cystomanamide B.

Deviations between the predicted structure and the actual product result from the involvement of unusual biosynthetic steps as well as tailoring reactions involved in the cystomanamide biosynthesis. This specifically refers to the branched-chain fatty acid starter unit, an aminotransferase domain at the interphase of the PKS and NRPS modules, an unusual FkbH-like domain that incorporates glyceric acid in an NRPS-like manner and tailoring reactions involving the hydroxylation and glycosylation of the compounds.

4.3.1.1 Starter module

Biosynthesis starts with incorporation of a fatty acid as starter unit as pointed out by the identification of an acyl-CoA ligase and an acyl carrier protein. Initial feeding experiments showed the incorporation of labeled leucine indicating a branched chain carboxylic acid as starter unit.[44] Based on the cystomanamide structures, the incorporated fatty acid is 9-methyl octanoic acid. The nature of the incorporated fatty acid could not be predicted by anti-SMASH 2.0 resulting in the first deviation between the predicted structure and the real one.

4.3.1.2 Aminotransferase (AMT) domain

The subsequent PKS elongation step is succeeded by an unusual reductive amination step conducted by an amino transferase at the interphase of the PKS and NRPS units. The *β*-ketothioester, emerged from the elongation with one acetate-derived C2-unit by the PKS module consisting of the domains KS-AT-ACP, is transformed into the *β*-amino thioester before the next elongation step by the first NRPS module incorporating D-asparagine into the growing chain. Such an amino transferase domain was investigated by Aron et al.[45] for the biosynthesis of mycosubtilin in a detailed study of the pyridoxal 5′-phosphate (PLP)-dependent enzyme operating *in cis* within the PKS and NRPS biosynthetic paradigm. It was found that the amine transfer happens in two steps: First, the amine is transferred from an amino acid (with glutamine as preferred amino source) to the protein-bound PLP where pyridoxamine 5′-phosphate (PMP) is formed. The second step involved the transfer of the amine from PMP to the protein bound *β*-ketothioester to generate the *β*-aminothioester under regeneration of PLP (Figure 4.4).

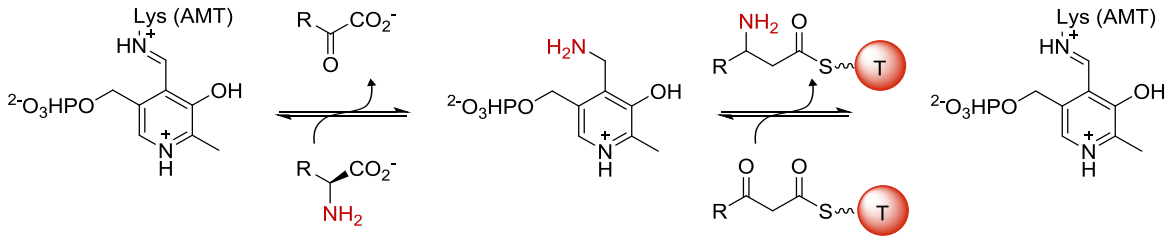


Figure 4.4. Mechanism of PLP mediated amine transfer as hypothesized for the amino transferase (AMT) domain located on CtmA based on the results for the AMT domain of mycosubtilin biosynthesis (Figure modified from Aron et al.[45])

4.3.1.3 FkbH-like domain

AntiSMASH prediction of the biosynthetic gene cluster identified a condensation domain as well as a peptidyl carrier protein between modules 4 and 6 located on CtmD, but was not able to assign them to a module or elongation step since no adenylation domain was identified. BLAST analysis revealed a region of the protein to be similar to proteins like FkbH[46] and OzmB[47] that are involved into the formation of glyceryl-ACP. Based on these results, this region or domain most likely acts in two steps (Figure 4.5): First D-1,3-bisphosphoglycerate from the glycolytic pool is bound and the phosphate group is removed to receive the D-3-glyceryl-S-CtmD species (acting as a phosphatase), and subsequently it acts as glyceryl transferase and transfers the glyceryl group to the following peptidyl carrier protein (PCP). For OzmB, glycerate was assigned to be covalently attached to the side chain of a cysteine residue as a thioester. Therefore, this FkbH-like domain could substitute the adenylation domain and take over the selection and activation of the building block. Up to now, this was only demonstrated for polyketide synthases as basis for the incorporation of unusual methoxymalonyl-CoA extender units.[46–48]

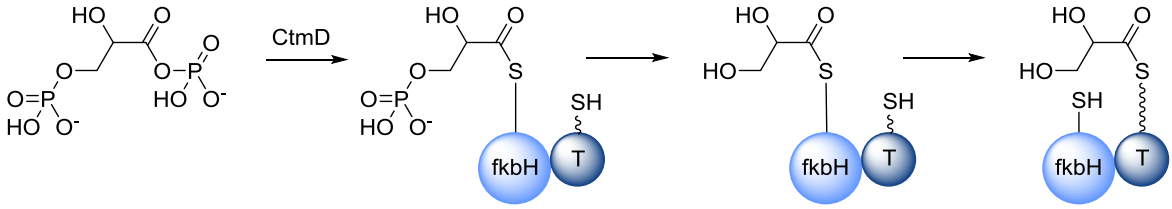


Figure 4.5: Selection and activation of D-1,3-bisphosphoglycerate from the glycolytic pool into the assembly line of cystomanamide B based on the results for OzmB in oxazolomycin biosynthesis (figure modified from Dorrestein et al.[47])

4.3.1.4 Further tailoring reactions

The biosynthetic gene cluster of the cystomanamides provides only scarce information about the tailoring reactions that involve the hydroxylation and glycosylation of the compounds. CtmB is homologous to the well-studied non-heme iron hydroxylase TauD and contains the conserved 2-His-1-carboxylate facial triad responsible for iron binding[49,50] as well as the conserved Arg residue that ligates *β*-ketoglutarate.[51] Its similarity to SyrP from *Pseudomonas syringae*, which is responsible for the *β*-hydroxylation of an aspartyl residue in syringomycin E biosynthesis[52] indicates that CtmB is most likely responsible for the hydroxylation of the second asparagine residue to form L-*erythro*-*β*-OH-asparagine. No further candidate proteins are encoded in the biosynthetic gene cluster that come into question for the hydroxylation of the phenylalanine residue. The hydroxylation of cystomanamide A-C as well as the glycosylation of C and D requires enzymes encoded outside the aglycon cluster.

4.3.2 Macyranone biosynthetic gene cluster

The prediction of the macyranone biosynthetic gene cluster was decidedly more complicated than of the cystomanamide cluster. The macyranone cluster is far from following textbook logic so that a prediction based on rules and linear assembly is highly overstrained in such a case. Nonetheless, it was helpful in terms that it unraveled the specific pattern of building blocks for the three core amino acids Thr-Thr-Leu and therefore allowed the identification of a candidate biosynthetic gene cluster that was later confirmed by gene deletion. As mentioned before, the biosynthesis of the macyranones deviates widely from textbook logic and multiple different unusual biochemical reactions are involved that are partly subject to speculation. I want to elaborate on some likely biochemical steps.

4.3.2.1 Alternative amide bond formation

The cluster prediction identified three NRPS elongation modules with specificity for two threonine building blocks and one leucine building block as well as one PKS elongation module. In addition to the three complete NRPS modules, another adenylation domain was found that shows an uncertain specificity for histidine. However, no modules and domains responsible for the incorporation of the phenylalanine and the methylmalonic acid could be found. This gives rise to the assumption, that the assembly line starts with

threonine and phenylalanine and methylmalonic acid are attached in an alternative way so that the formation of two amide bonds remains to be explained. Amide bonds represent fundamental linkages for numerous natural products and are normally formed by the ribosome or by the condensation domain of nonribosomal peptide synthetases. Since none of both seems to be true for the two amide bonds in question, other mechanisms have to be considered. Identification of a shunt product of the macyranone assembly line, that contains the shortened sequence Phe-Mma-Thr, indicates that the assembly line starts with the bonding of these three residues. As mentioned in chapter 3, only one likely candidate comes into question for the connection of the phenylalanine and the methylmalonic acid residue via an amide bond. MynA shows similarity to a family of asparagine synthetases, more detailed to a subgroup of asparagin synthetase (AS) B-like amide synthetases. Some members of the group were proven to be responsible for amide bond formation in secondary metabolites. [53–55]

MynA shows the closest similarity to PdmN from the pradimicin pathway[53] (41.6 % identity), to RubR from the rubromycin pathway (40.3 % identity) and to FdmV from the fredericamycin pathway[55] (39.5 % identity). The enzymes were compared to the typical AS B from *E. coli* (Chapter 3, Figure S4), which belongs to the class II glutamine amido-transferases (20.5 % identity) and converts aspartate into asparagine via a *β*-aspartyl-AMP intermediate with glutamine or free ammonia (NH_3) as nitrogen source.[56,57] All mentioned enzymes contain the conserved domains present in AS B: the N-terminal glutaminase domain and the C-terminal amide synthetase domain.[58] In principle, the glutaminase domain is responsible for the hydrolysis of L-glutamine to supply a free ammonia group. The amide synthase domain catalyzes the conversion of aspartic acid to asparagine.

The best characterized representative of the enzymes involved in secondary metabolite formation is FdmV, an amide synthetase involved in the formation of the aromatic pentadecaketide natural product fredericamycin.[55] *In vitro* and *in vivo* studies of the protein have proven that it catalyzed a lactam ring formation. The protein is capable of using L-Glu, L-Asn and free NH_3 as nitrogen source. PdmN is suspected to catalyze an amide bond formation by using D-Ala as amino donor and therefore to attach an amino acid to the aromatic polyketide intermediate of pradimicin.[53] No amidating function can be assigned to RubR since none of the known rubromycins contains an amide linkage.[54] Analyzing the alignment of the homolog enzymes and taking into account the results of studies concerning the three-dimensional structure of the *E. coli* AS B,[59] some interesting

facts could be used to explain the function of MynA: All class II glutamine amidotransferases contain a conserved N-terminal cysteine residue (Cys^2) that was shown to act as active-site nucleophile for the hydrolysis of glutamine to glutamate and ammonia. A site-mutation where the cysteine residue Cys^2 was exchanged for an alanine resulted in a catalytically inactive protein.[59] This residue is retained in AS B and FdmV but is mutated to serine in MynA, PdmN and RubR. The mutation leads to a loss of the glutaminase activity in the three enzymes. For MynA and PdmN, the amide synthetase activity seems to be retained by using other substrates as nitrogen donor (L-Phe and D-Ala, respectively). Additionally, four further conserved residues in the N-terminal glutaminase domain, that were shown to form hydrogen bonds with the glutamine ligand in AS B, were not retained in MynA (red arrows in yellow area in Chapter 3, Figure S4). This suggests a varied catalytic activity of the domain. The C-terminal amide synthetase domain, that is responsible for the production of the aspartyl-AMP intermediate, contains more active-site residues that match between AS B and MynA. The residues for the direct AMP binding are all conserved in MynA whereas residues for the coordination of AMP via metal ions (most likely Mg^{2+}) are only conserved for one metal ion instead of two. These results suggest that the formation of the amide bond between the methylmalonic acid residue and phenylalanine occur via an AMP intermediate with phenylalanine acting as amino donor (Figure 4.6).

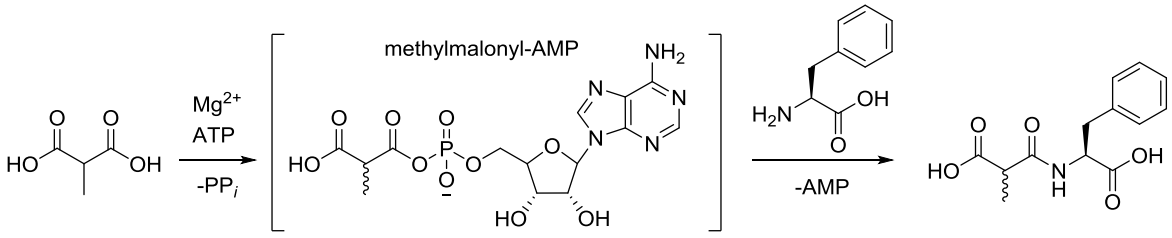


Figure 4.6: Activation of methylmalonate by the formation of a methylmalonyl-AMP intermediate. Subsequent conversion of the methylmalonyl-AMP intermediate via a nucleophilic attack by the phenylalanine adapted from the principle of the conversion of aspartate to asparagine by the AS B from *E. coli*[59]

For the formation of the second outstanding amide bond, two likely candidates come into question, both being part of the first NRPS module of macyranone biosynthesis. Either the adenylation domain or the condensation domain found in the cluster of the module could catalyze the connection. Module 1 is already unusual from the point that it actually contains a condensation domain. In general, NRPS starter modules do not contain a

condensation domain since the C domain catalyzes the condensation reaction of the growing peptide chain bound to the upstream PCP domain and the newly activated amino acid bound to the downstream PCP.[60] Exceptions are starter C domains that acylate the first amino acid with a *β*-hydroxy-carboxylic acid.[61] It is a common motif of lipopeptide synthetases comprising the surfactin, fengycin and lichenysin synthetases, that the first module is preceded by a condensation domain.[62–65] These condensation domains could catalyze the transfer of the fatty acid to the first amino acid. Since no lipid chain is involved in the macyranones, the occurrence of this first condensation domain already indicates an unusual priming of the assembly line. Phylogenetic analysis of the condensation domain using the online tool NaPDoS[66] revealed its highly unusual nature since it represents a new clade in the phylogenetic tree of condensation domains and does not group with other known types like $^{L}C_{L}$ type, $^{D}C_{L}$ type, starter, hybrid, cyclization, epimerization or dual function condensation domains (Chapter 3, Figure S5). Therefore, the condensation domain could be responsible for a new biochemical step which connects the Phe-Mma intermediate to the threonine residue selected by the first adenylation domain.

This first adenylation domain found in the *myn* biosynthetic gene cluster is a stand-alone adenylation domain with threonine substrate specificity localized on *mynB* and at the same time, it represents the second way how the formation of this amide bond could be accomplished. An amide-forming stand-alone adenylation domain was identified and characterized for streptothricin biosynthesis. Enzymatic assays with purified protein strongly suggested that amide-bond formation for the oligopeptide streptothricin was directly catalyzed by the stand-alone A domain Orf19. Therefore, the adenylation domain encoded by *mynB* itself could also be responsible for the amide bond formation.[67] Protein alignment between Orf19, MynB and other stand-alone adenylation domains did not result in an evaluable result since the identities are all in a similar range between 20 and 28 % (24.0 % identity between Orf19 and MynB). To get further insights into the formation of this amide bond, studies on the protein level would be essential.

Summing up the above results concerning the analysis of the cystomanamide and macyranone biosynthetic gene clusters, *in silico* predictions should be seen as a useful tool for the identification of a candidate biosynthetic gene cluster based on a known structure, but it cannot substitute a detailed analysis of the respective genes.

4.4 From Natural Products to Drugs

The risk and disadvantage of a screening for new compounds based on chemical and structural features is that we might end up with compounds not exhibiting any kind of biological activity as it was the case for the cystomanamides. Nevertheless, the compounds and the understanding of their biosynthesis contribute to our understanding of the bacteria, their secondary metabolism and the chemical diversity produced by them. Further screening of the compounds might still hit a hidden target in the future.

The macyranones show potent activity against the parasitic agents *Trypanosoma brucei rhodesiense* and *Leishmania donovani* and therefore they might aid in the fight against the African sleeping sickness and Leishmaniasis. With their low cytotoxicity against mammalian cell lines, the compounds are good candidates for the drug development as anti-parasitic agents. This selectivity is surprising in terms that the identified target, the 20S proteasome, is present in all eukaryotic cells so that a toxic effect on mammalian and parasitic cells would be expected. If we are able to decipher the mechanism of their selectivity for the parasitic cells, this might enable the selective use of targets shared by parasitic and mammalian cells. Possible explanations could be a transport-related selectivity so that the compounds can only enter the parasitic cells or a metabolism-related selectivity where the compounds are for example degraded by mammalian cells.[68]

For the use as a drug, also stability problems of macyranone A must be overcome since it has a low stability in buffered aqueous systems. The stability of all epoxyketones is an issue and also the approved drug carfilzomib possesses a very short systemic half-life limiting its applicability in the cancer treatment. Since previous studies[69] as well as the co-crystal structure of 20S proteasome and macyranone A clearly showed that the epoxyketone residue is crucial for the binding of the compound to its target, improvements of the structure are largely limited to the peptidic backbone. Replacement, addition and deletion of amino acids as well as making the peptide structure more rigid to improve the molecular stability of the peptide could be possible starting points for the process of a further drug development.

4.5 Final Thoughts

The isolation of new secondary metabolites from microorganisms has changed dramatically in the last half century. We cannot afford the aberration that new

compounds with a high UV absorbance and extraordinary biological activity will furthermore pop up in bacterial extracts. The exploitation of new bacterial sources with a high phylogenetic diversity, the development of enhanced detection and isolation methods and the expansion of our tools for the genetic manipulation of organisms are good starting points to carry forward the success of natural products. We need to strike out in new directions to still be able to isolate novel natural products from microbial sources and to make our contributions to the search for new anti-infectives to fight the increasing resistance of microorganisms against existing drugs. This work proves that if we are open for new approaches, myxobacteria still provide a valuable source for new natural products. With two new compound classes isolated from *Cystobacter fuscus*, we are one step further to unravel the chemical space produced by these extraordinary bacteria.

4.6 References

(1) Krug, D.; Müller, R.: Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. *Nat. Prod. Rep.* **2014**, *31*, 768–83.

(2) Kunze, B.; Bedorf, N.; et al.: Myxochelin A, a new iron-chelating compound from Angiococcus disciformis (Myxobacterales). Production, isolation, physico-chemical and biological properties. *J. Antibiot.* **1989**, *42*, 14–7.

(3) Ambrosi, H.-D.; Hartmann, V.; et al.: Myxochelins B, C, D, E and F: A New Structural Principle for Powerful Siderophores Imitating Nature. *Eur. J. Org. Chem.* **1998**, *1998*, 541–551.

(4) Zander, W.; Gerth, K.; et al.: Roimatacene: an antibiotic against Gram-negative bacteria isolated from Cystobacter ferrugineus Cb G35 (Myxobacteria). *Chem. Eur. J.* **2011**, *17*, 7875–81.

(5) Yamaguchi, H.; Nakayama, Y.; et al.: A new antibiotic, althiomycin. *J. Antibiot.* **1957**, *10*, 195–200.

(6) Bycroft, B.W.; Pinchin, R.: Structure of althiomycin, a highly modified peptide antibiotic. *J. Chem. Soc. Chem. Commun.* **1975**, 121.

(7) Rachid, S.; Gerth, K.; Müller, R.: NtcA: a negative regulator of secondary metabolite biosynthesis in Sorangium cellulosum. *J. Biotechnol.* **2009**, *140*, 135–42.

(8) Rachid, S.; Gerth, K.; et al.: Deciphering regulatory mechanisms for secondary metabolite production in the myxobacterium Sorangium cellulosum So ce56. *Mol. Microbiol.* **2007**, *63*, 1783–96.

(9) Wenzel, S.C.; Müller, R.: Myxobacteria--'microbial factories' for the production of bioactive secondary metabolites. *Mol. Biosyst.* **2009**, *5*, 567–74.

(10) Voehler, M.W.; Collier, G.; et al.: Performance of cryogenic probes as a function of ionic strength and sample tube geometry. *J. Magn. Reson.* **2006**, *183*, 102–9.

(11) Weissman, K.J.; Müller, R.: Myxobacterial secondary metabolites: bioactivities and modes-of-action. *Nat. Prod. Rep.* **2010**, *27*, 1276–95.

(12) Herrmann, J.; Hüttel, S.; Müller, R.: Discovery and biological activity of new chondramides from Chondromyces sp. *ChemBioChem* **2013**, *14*, 1573–80.

(13) Kunze, B.; Jansen, R.; et al.: Chondramides A-D, new antifungal and cytostatic depsipeptides from Chondromyces crocatus (myxobacteria). Production, physico-chemical and biological properties. *J. Antibiot.* **1995**, *48*, 1262–6.

(14) Steinmetz, H.; Glaser, N.; et al.: Isolation, crystal and solution structure determination, and biosynthesis of tubulysins--powerful inhibitors of tubulin polymerization from myxobacteria. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 4888–92.

(15) Chai, Y.; Pistorius, D.; et al.: Discovery of 23 natural tubulysins from Angiococcus disciformis An d48 and Cystobacter SBCb004. *Chem. Biol.* **2010**, *17*, 296–309.

(16) Baumann, S.; Herrmann, J.; et al.: Cystobactamids: Myxobacterial Topoisomerase Inhibitors Exhibiting Potent Antibacterial Activity. *Angew. Chem. Int. Ed. Engl.* **2014**, 14605–14609.

(17) Hoffmann, T.; Müller, S.; et al.: Microsclerodermins from terrestrial myxobacteria: an intriguing biosynthesis likely connected to a sponge symbiont. *J. Am. Chem. Soc.* **2013**, *135*, 16904–11.

(18) Garcia, R.O.; Krug, D.; Müller, R.: Chapter 3. Discovering natural products from myxobacteria with emphasis on rare producer strains in combination with improved analytical methods. *Methods Enzymol.* **2009**, *458*, 59–91.

(19) Scherlach, K.; Hertweck, C.: Triggering cryptic natural product biosynthesis in microorganisms. *Org. Biomol. Chem.* **2009**, *7*, 1753–60.

(20) Cannell, R.J.P.: How to Approach the Isolation of a Natural Product. *Methods in Biotechnology, Vol.4 Natural Products Isolation* **1998**, Humana Press, Hertfordshire, UK, p 473.

(21) Gerth, K.; Pradella, S.; et al.: Myxobacteria: proficient producers of novel natural products with various biological activities—past and future biotechnological aspects with the focus on the genus Sorangium. *J. Biotechnol.* **2003**, *106*, 233–253.

(22) Gerth, K.; Steinmetz, H.; et al.: Studies on the Biosynthesis of Epothilones. The PKS and Epothilone C/D Monooxygenase. *J. Antibiot.* **2001**, *54*, 144–148.

(23) Taori, K.; Matthew, S.; et al.: Lyngbyastatins 5-7, potent elastase inhibitors from Floridian marine cyanobacteria, Lyngbya spp. *J. Nat. Prod.* **2007**, *70*, 1593–1600.

(24) Zafrir-Ilan, E.; Carmeli, S.: Eight novel serine proteases inhibitors from a water bloom of the cyanobacterium Microcystis sp. *Tetrahedron* **2010**, *66*, 9194–9202.

(25) Plaza, A.; Bewley, C. a: Largamides A-H, unusual cyclic peptides from the marine cyanobacterium Oscillatoria sp. *J. Org. Chem.* **2006**, *71*, 6898–6907.

(26) Schummer, D.; Höfle, G.; et al.: Antibiotics from Gliding Bacteria, LXXVI. Vioprolides: New Antifungal and Cytotoxic Peptolides from Cystobacter violaceus. *Liebigs Ann.* **2006**, *1996*, 971–978.

(27) Desjardine, K.; Pereira, A.; et al.: Tauramamide, a lipopeptide antibiotic produced in culture by Brevibacillus laterosporus isolated from a marine habitat: Structure elucidation and synthesis. *J. Nat. Prod.* **2007**, *70*, 1850–1853.

(28) Zavascki, A.P.; Goldani, L.Z.; et al.: Polymyxin B for the treatment of multidrug-resistant pathogens: A critical review. *J. Antimicrob. Chemother.* **2007,.**

(29) Duitman, E.H.; Hamoen, L.W.; et al.: The mycosubtilin synthetase of Bacillus subtilis ATCC6633: a multifunctional hybrid between a peptide synthetase, an amino transferase, and a fatty acid synthase. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 13294–9.

(30) Hanada, M.; Sugawara, K.; et al.: Epoxomicin, a new antitumor agent of microbial origin. *J. Antibiot.* **1992**, *45*, 1746–52.

(31) Sugawara, K.; Hatori, M.; et al.: Eponemycin, a new antibiotic active against B16 melanoma. I. Production, isolation, structure and biological activity. *J. Antibiot.* **1990**, *43*, 8–18.

(32) Koguchi, Y.; Nishio, M.; et al.: TMC-89A and B, new proteasome inhibitors from streptomyces sp. TC 1087. *J. Antibiot.* **2000**, *53*, 967–972.

(33) Koguchi, Y.; Kohno, J.; et al.: TMC-86A, B and TMC-96, new proteasome inhibitors from Streptomyces sp. TC 1084 and Saccharothrix sp. TC 1094. I. Taxonomy, fermentation, isolation, and biological activities. *J. Antibiot.* **1999**, *52*, 1069–76.

(34) Pereira, A.R.; Kale, A.J.; et al.: The carmaphycins: new proteasome inhibitors exhibiting an α,β-epoxyketone warhead from a marine cyanobacterium. *ChemBioChem* **2012**, *13*, 810–7.

(35) Medema, M.H.; Blin, K.; et al.: antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **2011**, *39*, W339–46.

(36) Blin, K.; Medema, M.H.; et al.: antiSMASH 2.0--a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **2013**, *41*, W204–12.

(37) http://antismash.secondarymetabolites.org/

(38) Punta, M.; Coggill, P.C.; et al.: The Pfam protein families database. *Nucleic Acids Res.* **2012**, *40*, D290–301.

(39) Röttig, M.; Medema, M.H.; et al.: NRPSpredictor2--a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **2011**, *39*, W362–7.

(40) Bachmann, B.O.; Ravel, J.: Chapter 8. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. *Methods Enzymol.* **2009**, *458*, 181–217.

(41) Altschul, S.F.; Gish, W.; et al.: Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–10.

(42) Wenzel, S.C.; Kunze, B.; et al.: Structure and biosynthesis of myxochromides S1-3 in Stigmatella aurantiaca: evidence for an iterative bacterial type I polyketide synthase and for module skipping in nonribosomal peptide biosynthesis. *ChemBioChem* **2005**, *6*, 375–85.

(43) Wenzel, S.C.; Müller, R.: Formation of novel secondary metabolites by bacterial multimodular assembly lines: deviations from textbook biosynthetic logic. *Curr. Opin. Chem. Biol.* **2005**, *9*, 447–58.

(44) Bode, H.B.; Dickschat, J.S.; et al.: Biosynthesis of iso-fatty acids in myxobacteria: iso-even fatty acids are derived by alpha-oxidation from iso-odd fatty acids. *J. Am. Chem. Soc.* **2005**, *127*, 532–3.

(45) Aron, Z.D.; Dorrestein, P.C.; et al.: Characterization of a new tailoring domain in polyketide biogenesis: the amine transferase domain of MycA in the mycosubtilin gene cluster. *J. Am. Chem. Soc.* **2005**, *127*, 14986–7.

(46) Wu, K.; Chung, L.; et al.: The FK520 gene cluster of Streptomyces hygroscopicus var. ascomyceticus (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene* **2000**, *251*, 81–90.

(47) Dorrestein, P.C.; Van Lanen, S.G.; et al.: The bifunctional glyceryl transferase/phosphatase OzmB belonging to the HAD superfamily that diverts 1,3-bisphosphoglycerate into polyketide biosynthesis. *J. Am. Chem. Soc.* **2006**, *128*, 10386–7.

(48) Wenzel, S.C.; Williamson, R.M.; et al.: On the biosynthetic origin of methoxymalonyl-acyl carrier protein, the substrate for incorporation of "glycolate" units into ansamitocin and soraphen A. *J. Am. Chem. Soc.* **2006**, *128*, 14325–36.

(49) Hegg, E.L.; Que, L.: The 2-His-1-carboxylate facial triad- An emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur. J. Biochem.* **1997**, *250*, 625–9.

(50) Ryle, M.J.; Koehntop, K.D.; et al.: Interconversion of two oxidized forms of taurine/alpha-ketoglutarate dioxygenase, a non-heme iron hydroxylase: evidence for bicarbonate binding. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3790–5.

(51) Hausinger, R.P.: FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.* **2004**, *39*, 21–68.

(52) Singh, G.M.; Fortin, P.D.; et al.: beta-Hydroxylation of the aspartyl residue in the phytotoxin syringomycin E: characterization of two candidate hydroxylases AspH and SyrP in Pseudomonas syringae. *Biochemistry* **2008**, *47*, 11310–20.

(53) Kim, B.C.; Lee, J.M.; et al.: Cloning, sequencing, and characterization of the pradimicin biosynthetic gene cluster of Actinomadura hibisca P157-2. *J. Microbiol. Biotechnol.* **2007**, *17*, 830–839.

(54) Martin, R.; Sierner, O.; et al.: Collinone, a New Recombinant Angular Polyketide Antibiotic Made by an Engineered Streptomyces Strain. *J. Antibiot.* **2001**, *54*, 239–249.

(55) Chen, Y.; Wendt-Pienkowski, E.; et al.: Characterization of FdmV as an amide synthetase for fredericamycin A biosynthesis in Streptomyces griseus ATCC 43944. *J. Biol. Chem.* **2010**, *285*, 38853–60.

(56) Riley, M.; Abe, T.; et al.: Escherichia coli K-12: a cooperatively developed annotation snapshot--2005. *Nucleic Acids Res.* **2006**, *34*, 1–9.

(57) Zalkin, H.: Advances in Enzymology and Related Areas of Molecular Biology. **1993**, (Meister, A., Ed.) John Wiley & Sons, Inc., Hoboken, NJ, USA,.

(58) Richards, N.G.; Schuster, S.M.: Mechanistic issues in asparagine synthetase catalysis. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1998**, *72*, 145–98.

(59) Larsen, T.M.; Boehlein, S.K.; et al.: Three-Dimensional Structure of Escherichia coli Asparagine Synthetase B: A Short Journey from Substrate to Product. *Biochemistry* **1999**, *38*, 16146–16157.

(60) Marahiel, M.A.; Stachelhaus, T.; Mootz, H.D.: Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chem. Rev.* **1997**, *97*, 2651–2674.

(61) Rausch, C.; Hoof, I.; et al.: Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evol. Biol.* **2007**, *7*, 78.

(62) Cosmina, P.; Rodriguez, F.; et al.: Sequence and analysis of the genetic locus responsible for surfactin synthesis in Bacillus subtilis. *Mol. Microbiol.* **1993**, *8*, 821–831.

(63) Tosato, V.; Albertini, A.M.; et al.: Sequence completion, identification and definition of the fengycin operon in Bacillus subtilis 168. *Microbiology* **1997**, *143*, 3443–3450.

(64) Konz, D.; Doekel, S.; Marahiel, M.A.: Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin. *J. Bacteriol.* **1999**, *181*, 133–40.

(65) Chooi, Y.-H.; Tang, Y.: Adding the lipo to lipopeptides: do more with less. *Chem. Biol.* **2010**, *17*, 791–3.

(66) Ziemert, N.; Podell, S.; et al.: The natural product domain seeker NaPDoS: a phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. *PLoS One* **2012**, *7*, e34064.

(67) Maruyama, C.; Toyoda, J.; et al.: A stand-alone adenylation domain forms amide bonds in streptothricin biosynthesis. *Nat. Chem. Biol.* **2012**, *8*, 791–7.

(68) Luscher, A.; de Koning, H.; Maser, P.: Chemotherapeutic Strategies Against Trypanosoma brucei: Drug Targets vs. Drug Targeting. *Curr. Pharm. Des.* **2007**, *13*, 555–567.

(69) Groll, M.; Kim, K.B.; et al.: Crystal Structure of Epoxomicin:20S Proteasome Reveals a Molecular Basis for Selectivity of α′,β′-Epoxyketone Proteasome Inhibitors. *J. Am. Chem. Soc.* **2000**, *122*, 1237–1238.