

Heterologous Production of Natural Products: From Compound Discovery to Metabolic Engineering of the Host Strain

Dissertation

zur Erlangung des Grades

des Doktors der Naturwissenschaften

der Naturwissenschaftlich-Technischen Fakultät

der Universität des Saarlandes

von

Nils Gummerlich

Saarbrücken

2019

Tag des Kolloquiums:	16.06.2020
Dekan:	Prof. Dr. Guido Kickelbick
Berichterstatter:	Prof Dr. Andriy Luzhetskyy Prof. Dr. Rolf Müller
Vorsitz:	Prof. Dr. Uli Kazmaier
Akad. Mitarbeiter:	Dr.-Ing. Michael Kohlstedt

Diese Arbeit entstand unter Anleitung von Prof. Dr. Andriy Luzhetskyy in der Fachrichtung 8.2, Pharmazeutische Biotechnologie der Naturwissenschaftlich-Technischen Fakultät der Universität des Saarlandes von Januar 2016 bis Dezember 2019.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Prof. Dr. Andriy Luzhetskyy for the great opportunity to pursue my doctorate in his working group, for his valuable guidance throughout my PhD and his continuous support to tackle countless technical problems.

Furthermore, I would like to thank Dr. Yuriy Rebets for being a great mentor during my PhD. Many of the problems that occurred during the PhD would have been much harder to solve without his guidance. Additionally, I would also like to thank him for being a good friend and squash partner during these years.

I am very grateful to Prof. Dr. Rolf Müller for his valuable advice during my committee meetings.

I gratefully acknowledge Dr. Josef Zapp, Constanze Paulus, Marc Stierhof and Suvd Nadmid for introducing me to the world of NMR spectroscopy and their help and support with NMR problems.

I would also like to thank Nikolas for being a close friend and for making my PhD life a time full of joy with his cheerfulness and kind attitude.

Moreover, I would like to thank the rest of my colleagues. Thanks to all of you, I was able to fully enjoy my PhD. Thank you very much for all the laughter, the coffee breaks, the cake times, the out seminars, the scientific discussions and everything else I forgot to mention.

Weiterhin möchte ich meiner Familie danken. Vielen Dank Gisela, Jörn, Svenja und Liv. Ohne Euch wäre ich nicht derjenige, der ich heute bin. Ohne Eure Unterstützung wäre weder mein Studium noch meine Dissertation möglich gewesen und dafür bin ich euch endlos dankbar.

Zu guter Letzt möchte ich Angela danken. Vielen Dank, dass du mir seit 10 Jahren hilfst meine Träume zu realisieren. Vielen Dank für deine Hilfe während des Studiums, für tausende Stunden des gemeinsamen Lernens, vielen Dank für das Korrekturlesen von zahllosen Einleitungen, Diskussionen und Ergebnissen, vielen Dank für deine emotionale Unterstützung in jeder Situation, vielen Dank für dein Zuhören, deine Ratschläge, deine Liebe, deine Geduld und dein Weitblick an mich zu glauben. Vielen Dank für deine Liebe.

Publications & Conference Contributions

Rodríguez Estévez, M.; Myronovskyi, M.; Gummerlich, N.; Zapp, J.; Luzhetskyy, A.; “Benzanthric acid, a novel metabolite from *Streptomyces albus* Del14 expressing nybomycin gene cluster.” *Front Chem* 2019 (“in press”).

Rodríguez Estévez, M.; Myronovskyi, M.; Gummerlich, N.; Nadmid, S.; Luzhetskyy, A.; “Heterologous Expression of the Nybomycin Gene Cluster from the Marine Strain *Streptomyces albus* subsp. *chlorinus* NRRL B-24108.” *Mar. Drugs* 2018, 16, 435.

Gummerlich, N.; Rebets, Y.; Paulus, C.; Zapp, J.; Luzhetskyy, A.; “Isolation of two unique polyketides through genome mining of *Saccharothrix espanaensis*.” *10th HIPS Symposium*, Saarbrücken, 2019.

Gummerlich, N.; Manderscheid, N.; Rebets, Y.; Petzke, L.; Luzhetskyy, A.; “Modification of the produced spectrum of pamamycins by precursor supply modulation.” *3rd ECNP*, Frankfurt 2018.

Zusammenfassung

Naturstoffe aus Aktinobakterien sind eine wichtige Quelle für die Entwicklung neuer Medikamente. Der Fokus dieser Arbeit lag daher einerseits auf der Entdeckung neuer Naturstoffe und andererseits auf der Weiterentwicklung bereits bekannter, aber noch ungenutzter Naturstoffe.

Durch die Auswahl und Expression biosynthetischer Genclustern aus dem Stamm *Saccharothrix espanaensis* im optimierten Wirt *Streptomyces lividans* wurden zwei neue Polyketid-Naturstoffe, pentangumycin und SEK90, entdeckt. Beide Stoffe wurden erfolgreich isoliert, deren Strukturen und Biosynthesewege wurden aufgeklärt und deren biologische Aktivität bestimmt. Ein großer Mehrwert der heterologen Expression als universellen Ansatz für die Entdeckung neuer Naturstoffe auch aus nicht-*Streptomyces* Aktinobakterien konnte hiermit aufgezeigt werden.

Die Polyketid Naturstoffgruppe der Pamamycine zeigt eine Vielfalt an unterschiedlichen biologischen Aktivitäten, was sie zu einem interessanten Forschungsgebiet macht. Ein Hauptaspekt, der die Weiterentwicklung dieser Stoffe zu industriell genutzten Stoffen verhindert, ist die Komplexität deren Isolierung, da durch den Einbau verschiedener Primärmetaboliten mindestens 18 verschiedene Derivate produziert werden. Durch gezielte genetische Modifikation der primären Stoffwechselwege in *Streptomyces albus* J1074 konnte die Versorgung dieser Metaboliten modifiziert und die Produktion der Pamamycin Derivate erfolgreich gesteuert werden.

Abstract

Natural products from Actinobacteria are an important source for the development of new medicines. The focus of the presented work was therefore, on the discovery of new natural substances and on the further development of already known, but unexploited natural products.

By selecting and expressing biosynthetic gene clusters from the strain *Saccharothrix espanaensis* in the optimized host *Streptomyces lividans*, two new polyketides, pentangumycin and SEK90, were discovered. After their successful isolation, their structures and biosynthetic pathways were elucidated, and their biological activity was determined. With this application of the heterologous expression, we could demonstrate its great utility as universal approach for new natural products discovery also from non-*Streptomyces* Actinobacteria.

The group of polyketide natural products pamamycins possesses a variety of biological activities and is therefore, a very interesting research topic. A main aspect that prevents their further development to industrially used substances is the complexity of their isolation, since at least 18 different derivatives are produced by the incorporation of different precursors. Through genetic manipulations of primary metabolic pathways in *Streptomyces albus J1074*, the supply of precursors into the heterologously expressed pamamycin biosynthetic pathway was changed leading to a modified spectrum of accumulated pamamycins.

Table of Contents

1. Introduction.....	13
1.1. Natural Products are a main source of pharmacological lead compounds.	13
1.2. Groups of Natural Products.....	16
1.2.1. Polyketides.....	17
1.2.2. Type I Polyketides	18
1.2.3. Type II Polyketides	20
1.2.3.1. Angucyclines.....	21
1.2.3.2. Macro-poly-olide antibiotics	23
1.2.3.2.1. Macrotetrolides	24
1.2.3.2.2. Macrodiolides	24
1.3. Polyketides biosynthesis precursors supply pathways	26
1.3.1. Acetyl-CoA and Malonyl-CoA.....	27
1.3.2. Methylmalonyl-CoA, Propionyl-CoA and Succinyl-CoA	29
1.3.3. Ethylmalonyl-CoA.....	30
1.4. Genome mining for new Natural Products	30
1.5. Heterologous Hosts	32
1.5.1. <i>S. albus</i> J1074.....	33
1.5.2. <i>S. lividans</i> TK24.....	33
1.6. Outline	35
2. Discovery of polyketides through genome mining of <i>Saccharothrix espanaensis</i>	37
2.1. Abstract.....	37
2.2. Introduction.....	38
2.3. Results and Discussion.....	41
2.3.1. Selection of putative biosynthetic gene clusters.....	41
2.3.2. Success rate of the expression system	43
2.3.3. Isolation of pentangumycin and SEK90	45
2.3.4. Both (1) and (2) are the result of type II PKS gene cluster expression	47
2.3.5. Detailed analysis and border prediction of the biosynthetic gene cluster of pentangumycin.....	48
2.3.6. Origin of SEK90.....	50
2.3.7. Biosynthesis of pentangumycin's core structure	52
2.3.8. Tailoring steps of pentangumycin's biosynthesis.....	54

2.3.9.	Analysis of regulatory gene functions.....	55
2.4.	Conclusion	57
2.5.	Material and Methods.....	58
2.5.1.	General Biological procedures	58
2.5.2.	General analytical procedures	58
2.5.3.	In-silico analysis of gene clusters	59
2.5.4.	Extraction, Dereplication and Isolation	59
2.5.5.	Modification of BACs	61
2.5.6.	Construction of Plasmids	61
2.5.7.	Synthesis of 1E5_CMP1-2	61
2.5.8.	Feeding experiments.....	61
3.	Tuning the production of pamamycins through engineering of the precursors pool in the heterologous host <i>S. albus</i> J1074	63
3.1.	Abstract.....	63
3.2.	Introduction.....	64
3.3.	Materials and Methods	67
3.3.1.	Bacterial strains and culture	67
3.3.2.	General procedures	67
3.3.3.	Construction of recombinant cosmid clones and gene deletions in <i>S. albus</i>	67
3.3.4.	Pamamycin production and analysis	68
3.3.5.	Measurement of the CoA-Esters	69
3.3.6.	Incorporation of labelled amino acids	69
3.4.	Results.....	70
3.4.1.	Bioinformatics identification of pamamycin precursors supply pathways and corresponding genes in the genome of <i>S. albus</i>	70
3.4.2.	Metabolic pathways involving ethylmalonyl-CoA in <i>S. albus</i>	73
3.4.2.1.	Ethylmalonyl-CoA is supplied by Crotonyl-CoA Carboxylase/Reductase encoded by XNR_0456	73
3.4.2.2.	Ethylmalonyl-CoA mutase (MeaA).....	74
3.4.3.	Metabolic pathways involving methylmalonyl-CoA	75
3.4.3.1.	Methylmalonyl-CoA mutase (Mcm).....	75
3.4.3.2.	Propionyl-CoA carboxylase (PCC1-2).....	76
3.4.3.3.	Valine dehydrogenase (Vdh)	77
3.4.3.4.	Propionyl-CoA carboxylase (PCC3).....	78
3.4.3.5.	Remaining methylmalonyl-CoA supply pathways.....	78
3.5.	Discussion	79

3.6. Conclusion	81
4. Discussion.....	83
4.1. General scope of the presented work.....	83
4.2. Discovery of Natural Products from rare Actinobacteria.....	83
4.2.1. Heterologous expression of BGCs from the rare Actinobacteria <i>Saccharothrix espanaensis</i> DSM44229	85
4.2.2. The angucyclinone pentangumycin	87
4.2.3. The shunt product SEK90.....	88
4.3. Precursor flux modification	89
4.3.1. Monensin and Salinomycin.....	90
4.3.2. Pamamycins	91
4.4. Conclusion	95
5. References.....	96
6. Appendix.....	109
6.1. Section 2	109
6.2. Section 3	205

1. Introduction

1.1. Natural Products are a main source of pharmacological lead compounds.

Natural products provide lead structures for novel active pharmaceutical ingredients and are therefore, one of the key sources for future drugs.¹ The classic Waksman approach based on activity guided screening used since 1940 resulted in the discovery of numerous clinically relevant compounds. Its success kick-started the golden era of antibiotics that lasted until the 1980s.²⁻³ Apart from the treatment of bacterial infections, natural products are used for a wide variety of applications, amongst others for the treatment of cancer, diabetes and as lipid-lowering agents.⁴ In general, natural products are isolated from many different sources. Most identified natural products (~350.000) are derived from the Plant Kingdom. Research of the Animal Kingdom, especially the marine invertebrates, resulted in the discovery of 100.000 compounds up to now. Considering natural products of microorganisms origin (~70.000), the highest rate of bioactivity was observed for compounds derived from Actinobacteria.⁵ *Streptomyces*, the largest genus of Actinobacteria, are well-known for their capacity to produce structurally-diverse secondary metabolites⁶ as illustrated with an excerpt of compounds shown in Figure 1. Since their biosynthesis relies only on primary metabolites derived from catabolic pathways, two main factors contribute to the structural diversity of these secondary metabolites. Firstly, the enzymes that synthesize natural products assemble the utilized building blocks into a variety of core structures. For example, polyketides are derived from acyl-CoA-esters such as acetyl-CoA, propionyl-CoA, malonyl-CoA or methylmalonyl-CoA and many others, while Ribosomally-synthesized and post-translationally modified peptides (RiPPs) and Non-ribosomal Peptides (NRPs) are assembled from proteinogenic and non-proteinogenic amino acids that are specially synthesized for their assembly line.⁷ Secondly, post-translational enzymatic modifications result in drastic alterations of the produced core structure. These modifications can include reduction and oxidation of carbon-carbon bonds, glycosylation, hydroxylation, methylation, amination, carboxylation and more.⁸ While the classic Waksman approach of activity guided screening in Actinobacteria led to the discovery of a large number of bioactive natural products shortly after its implementation, its significant limitation became evident in recent years: already known natural products are constantly rediscovered resulting in an overall

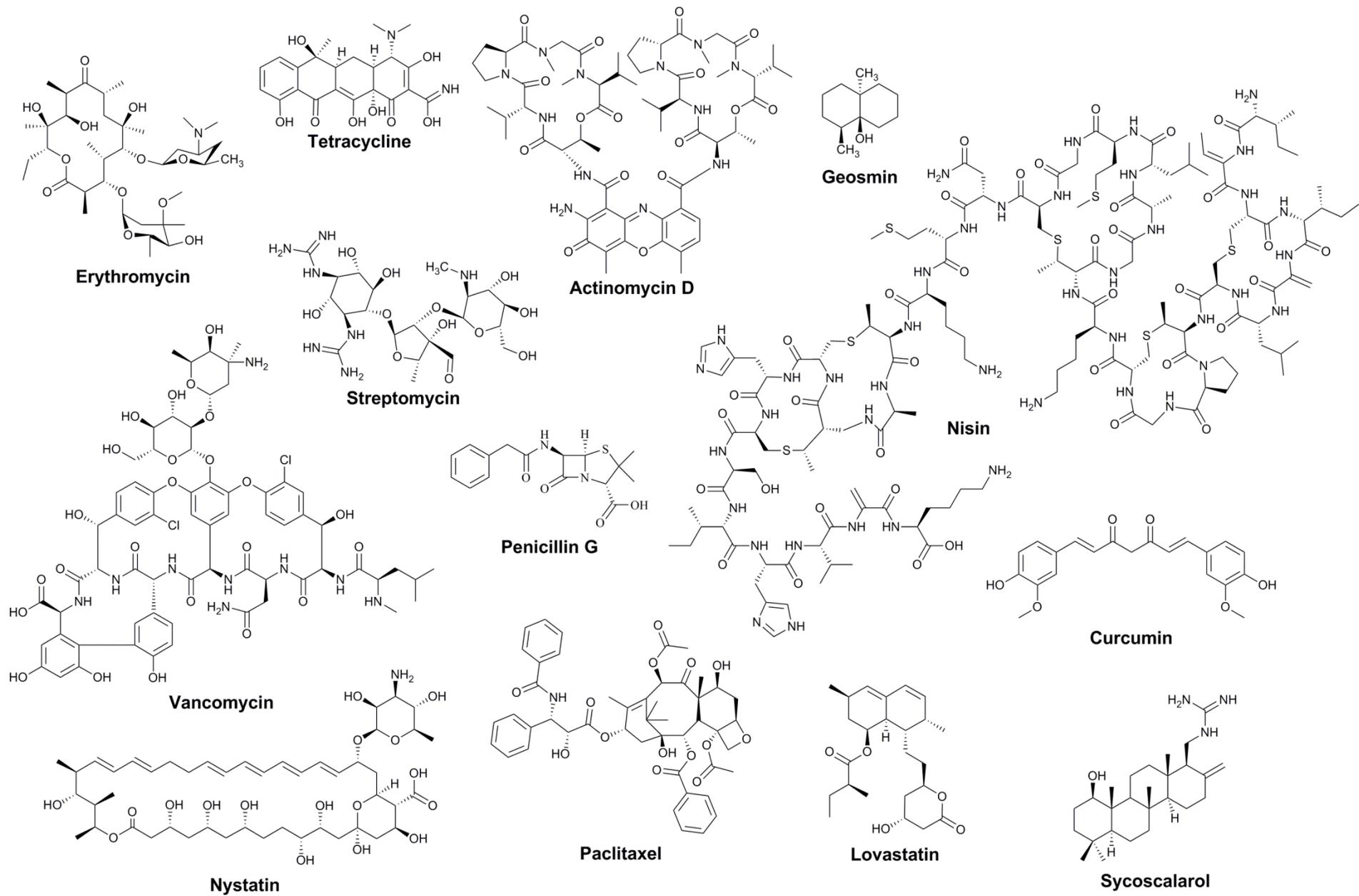


Figure 1: Natural products and their structural diversity.

deceleration of the discovery process.⁹⁻¹⁰ To overcome these challenges, scientists constructed databases of identified natural products describing their chemical and physical properties. These databases lead to the development of the procedure called dereplication. This procedure includes the analysis of extracted metabolites using high performance liquid chromatography coupled with high resolution mass spectrometry (LC-HRMS), followed by the systematic comparison of the respective chromatograms with natural product databases.

While the dereplication reduced the rediscovery rate of natural products, it presented its own limitations including a high expenditure of time due to lack of automation and error-proneness. For example, detected mass-to-charge ratios belonging to formed adduct ions or fragments of previously discovered natural products can mistakenly be considered as novel compounds. Additionally, mass-to-charge ratios of novel and previously discovered compounds can be identical, presenting the risk for novel compounds being overlooked, which is especially the case for smaller molecules.¹¹

Several methods have been developed in order to increase the general efficiency of the discovery of novel natural products. The “one strain many compounds” (OSMAC) approach, one of the earliest methods, included the modification of cultivation conditions (media composition, aeration, culture vessel, addition of enzyme inhibitors) and its first application resulted in the identification of 20 novel metabolites from different microorganisms.¹² Another approach is the targeted interaction screening, which includes the co-cultivation of two different microbial strains. A successful application of this approach included the co-cultivation of *Streptomyces coelicolor* M145 with *Amycolatopsis sp.* AA4, which resulted in the production of amycomycin, a potent antibiotic against *Staphylococcus aureus*.¹³ Recent publications of Seyedsayamdost *et al.* showed another powerful tool for the identification of natural products. They constructed a chemical library including 640 compounds and supplemented the media of the growing bacteria (*Burkholderia thailandensis*) with compounds from the chemical library. To screen the 640 differently grown colonies, they used high-throughput elicitor screening (HiTES) and observed activated natural product production in several cases. Interestingly, the media supplementation with sub-toxic

concentration of known antibiotics (cefotaxime, ceftazidime, trimetoprim, and piperacillin) showed the best results in terms of the activation of secondary metabolism.¹⁴⁻¹⁵

1.2. Groups of Natural Products

Natural products are generally classified according to their biosynthetic origin. Terpenes and terpenoids are the largest class of small molecule natural products.¹⁶ They derive from the C₅ substrates dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP). One molecule of DMAPP and one or more molecules of IPP are condensed in a “head-to-tail” fashion to form geranyl-diphosphate (C₁₀), farnesyl diphosphate (C₁₅) or geranylgeranyl diphosphate (C₂₀). Farnesyl- and geranylgeranyl diphosphate can then condense to the squalene, which is a precursor for a variety of important natural products including cholesterol and β-carotenoids. The different C₁₀-C₄₀ precursors are processed by terpene synthases to form the finalized terpenes and terpenoids.

Amino acid derived natural products can either be assembled by specialized assembly lines or via the translational apparatus of the ribosome. The large biosynthetic gene clusters (BGC) responsible for the non-ribosomal peptide (NRP) production encode a modular assembly line. Each module is responsible for the incorporation of a single amino acid. According to this rule of collinearity, for example a pentapeptide requires five modules to be built (Figure 2).¹⁷ The typical NRP-synthase (NRPS) carries at least 3 domains in each module. 1 – The adenylation domain (A) is responsible for the selection, activation and loading of the amino acid onto the thiolation domain. 2 – The thiolation domain (T) or peptidyl-carrier protein (PCP) carries a 4'-phosphopantetheine moiety and is responsible for the transfer of the growing amino acid chain between the domains and modules. 3 – The condensation domain (C) catalyzes the amide bond formation between the selected extender amino acid and the growing peptide chain. The incorporation of reductive (Re), oxidative (Ox), methylation (M) and other domains into the assembly line can drastically modify the peptide chain. The final domain usually carries a thioesterase (Te) activity that disconnects the oligopeptide from the enzyme.¹⁸ In contrast to the translational apparatus of the ribosome, non-ribosomal peptide synthetases can utilize non-proteinogenic amino acids that most often are specially synthesized by enzymes encoded in the NRPS BGC.¹⁹

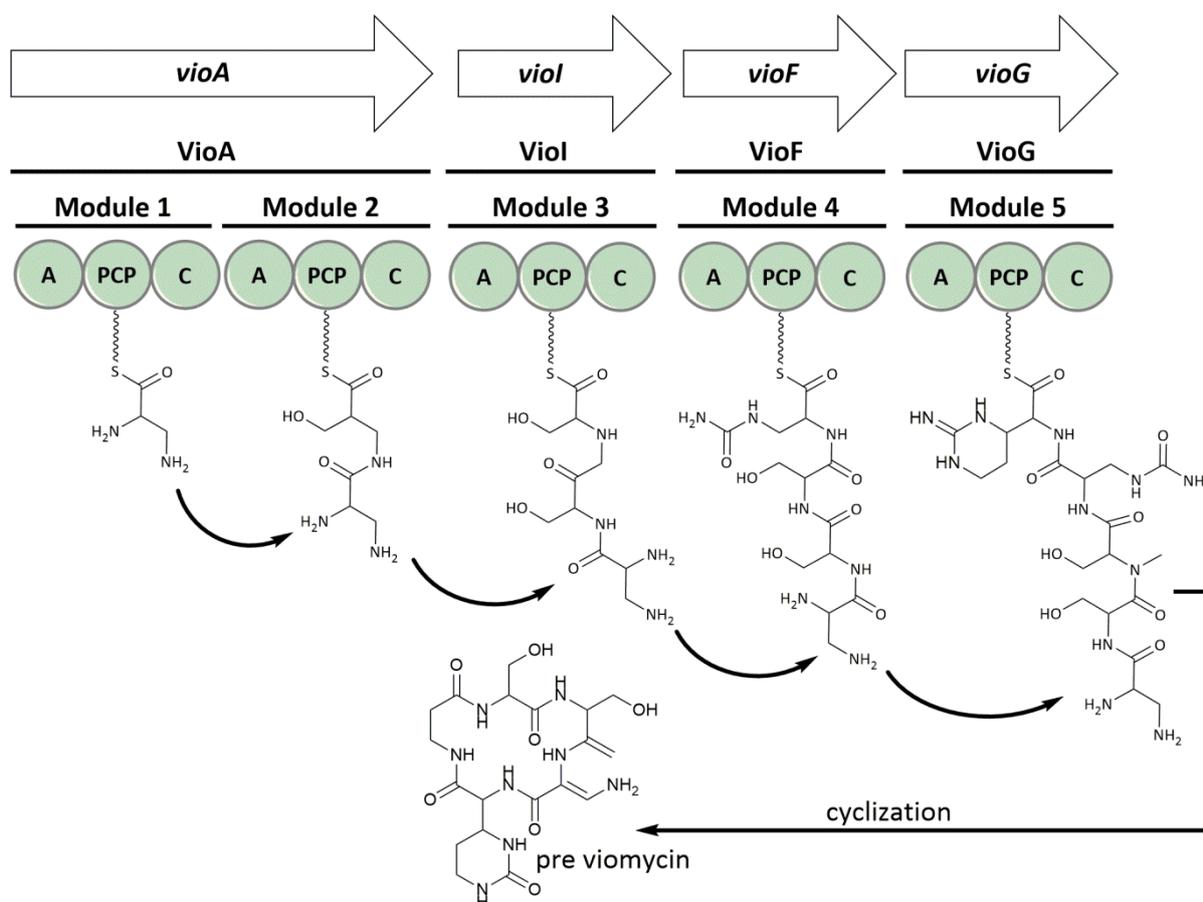


Figure 2: Schematic representation of the modular NRPS biosynthesis of viomycin.

The biosynthetic gene clusters for Ribosomally-synthesized and post translationally modified peptides (RiPPs) are considerably smaller than NRPS (15 kbp or less).²⁰ Even though there are more than 20 sub classes, their BGCs usually contain the genes encoding the precursor peptide, a regulatory element, a transporter protein and to some extent tailoring enzymes. Although the design of the precursor peptides varies in the different subclasses, they usually have a core peptide- and a leader peptide region. After the production of the precursor peptide by the ribosome, the leader peptide is recognized by tailoring enzymes, which subsequently modify the core peptide. A following proteolysis and export out of the cell will deliver the mature RiPP.²¹

1.2.1. Polyketides

Polyketides are a large group of natural products that, based on their structural properties, are divided into sub-group of polyphenols (or aromatic polyketides), macrolides, polyenes and polyethers.²² Similar to NRPs, their biosynthetic gene machinery works in an assembly line fashion. Polyketide synthases (PKS) are a family of multi-domain enzymes or large

enzymatic complexes that can be divided in three main groups depending on their organization and functional peculiarities. 1 – Type I PKS are large modular enzymes, which can work either in conveyor or iterative fashion. Similar to NRPS, the size of the produced polyketide is defined by the number of modules available (with some exceptions for the iterative Type I PKS). 2 – Type II PKS are aggregates of mono-functional enzymes (acyl-carrier protein, ketosynthase- α , ketosynthase- β). The size of the polyketide is defined by a channel formed between the α -unit and the β -unit of the complex.²³ 3 – Type III polyketide synthases form a homodimer and act directly on acyl-CoA ester, and thus, the group is defined by the absence of an acyl-carrier protein.²⁴

1.2.2. Type I Polyketides

Type I PKS can be divided by its mode of action into modular and iterative enzymes.²⁵⁻²⁶ The highly used antibiotic erythromycin is a modular-assembled polyketide and the synthesis of its core unit 6-deoxyerythronolide B is one of the best studied biosynthetic routes (Figure 3).²⁷ The 6-deoxyerythronolide B synthase (DEBS) is the prototype of non-iterative type I PKS. It consists of three distinct proteins DEBS 1, 2 and 3, each carrying two biosynthetic modules (DEBS 1 has an additional loading module). Each module contains a set of domains including an acyl-carrier protein (ACP), an acyltransferase (AT) and a ketosynthase (KS), combined with a different set of ketoreduction domains.²⁸ The first and last modules of type I PKS are an exception. The loading module usually contains only an AT and an ACP and starts the biosynthesis with the selection of the starter unit and its transfer to the following module. The last module contains additionally a thioesterase (TE) domain and terminates the reaction. Modules can contain a varying set of ketoreduction domains that are responsible for different reduction degrees of the keto-groups, which are formed during the condensation of the acyl units. These domains can include a ketoreductase (KR), an enoyl-reductase (ER) and a dehydratase (DH).²⁹⁻³⁰ Among all domains, the ACP holds a special function. Through a sulfur-carbon bond between its 4'-phosphopantetheine moiety and the growing polyketide chain it can transport the covalently bound chain between the single domains of the polyketide synthase. With a specialized active site, the AT chooses the extender unit (e.g., malonyl-CoA and methylmalonyl-CoA) and transfers it to the 4'-

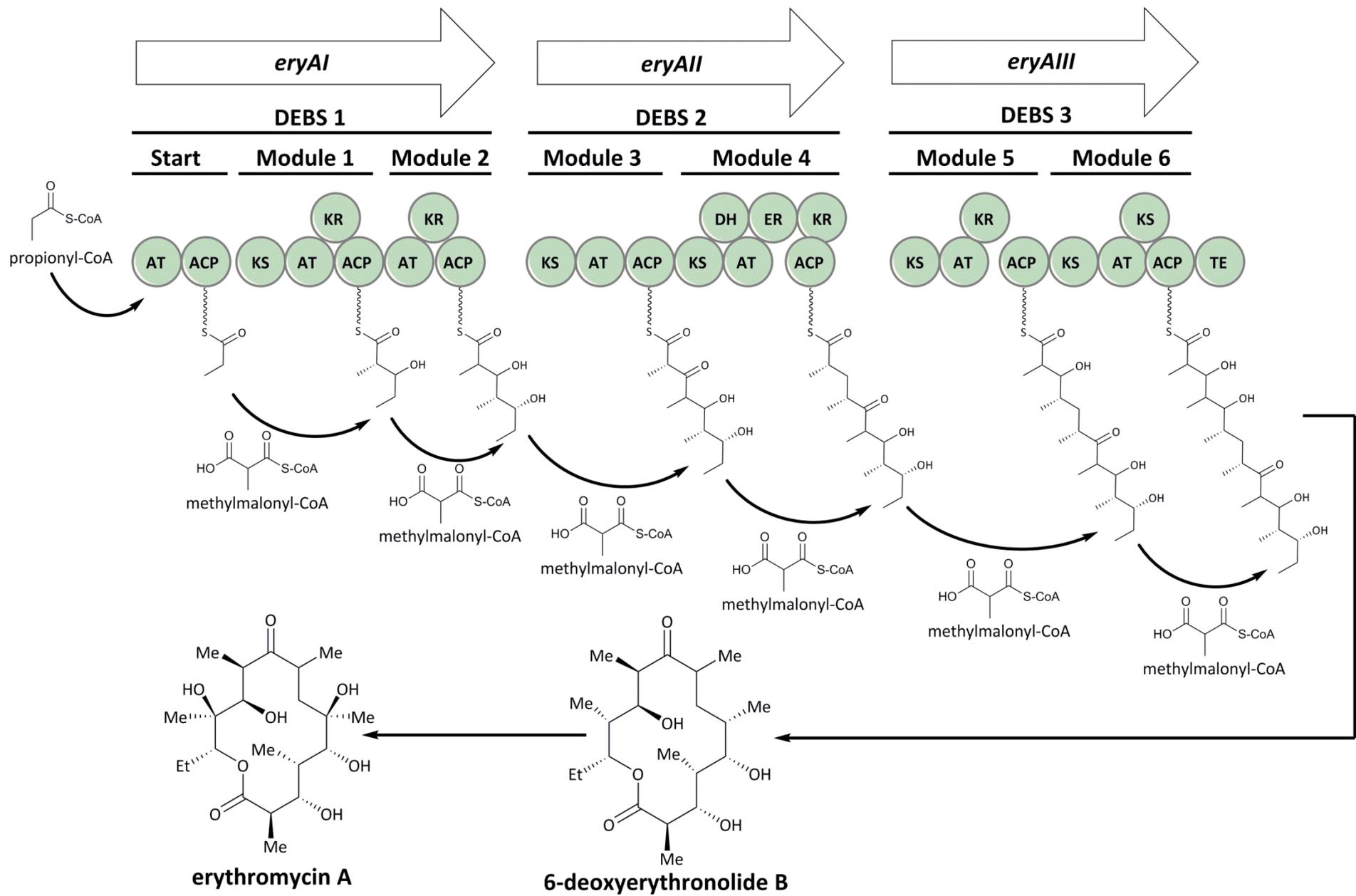


Figure 3: Schematic representation of the biosynthesis of erythromycin A.

phosphopantetheine moiety of the ACP.³¹ The ketosynthase carries the growing polyketide and catalyzes a decarboxylative Claisen condensation between the extender units delivered by the ACP and the polyketide chain. Subsequently, the chain is carried over to the KS of the next module by the ACP. Every additional module in the domain can further modify the growing chain. The KR reduces the ketone to a hydroxy group, the DH removes the hydroxy group and creates an alkene and the ER reduces the alkene to an alkane. The C-terminally located TE domain usually catalyzes the cleavage from the PKS, cyclizes the growing polyketide chain and subsequently forms a macrolactone.³² Structures generated by type I PKS can, to a certain degree, be predicted using bioinformatics tools, while the possibility of domain repetition and domain skipping renders absolute predictions of the chemical structure impossible.

Compared to modular type I PKS, the iterative type I PKS highly resembles the type I fatty acid synthases (FAS).²⁵ The fatty acid synthases utilizes acetyl-CoA and malonyl-CoA as building blocks and connects them via Claisen condensation. The products are generally fully reduced saturated fatty acids. In contrast to the FAS, type I PKS can produce fully reduced,³³ partially reduced³⁴ or non-reduced compounds.³⁵ Opposed to modular type I PKS, iterative type I PKS are composed of a single large enzyme, which is used in an iterative fashion. The encoded domains and the mode of assembly are identical to modular type I PKS. The different degrees of reduction by the iterative type I PKS are achieved through the absence of specific domains, inactive domains or domain skipping.^{22, 36}

1.2.3. Type II Polyketides

Similar to the iterative type I PKS, genes encoding the biosynthesis of type II polyketides resemble highly the type II FAS and work iteratively. However, structural differences between the derived products from type I and type II PKS are considerable.³⁷ The mono-enzymatic sub-unit KS α , the chain length determinant sub-unit KS β (also referred as chain length factor, CLF) and the ACP work iteratively in a three-step cycle (Figure 4).³⁸ First, the growing polyketide chain is loaded to the KS α . Second, a new building block is transferred to the 4'-phosphopantetheine of ACP. Third, the chain elongation is performed with a decarboxylative Claisen condensation by the KS α accompanied with the transfer of the growing chain on the ACP. At last, the elongated polyketide chain is moved back to the

catalytic cystein of the KS α . The number of iterations is determined by the length of a channel formed by the CLF subunit. After the described elongation cycle, tailoring steps are performed.³⁹⁻⁴¹ Usually, the nascent polyketide chain undergoes single reduction by a ketoreductase and a cyclization and aromatization by a series of cyclases and bifunctional cyclases/aromatases, resulting in the formation of polyphenolic compounds.⁴² Further modifications take also often place and include i.a. aromatization, oxygenation, oxidation, additional reductions, glycosylation, acylation, leading to the diversity of polyketides produced by type II system.

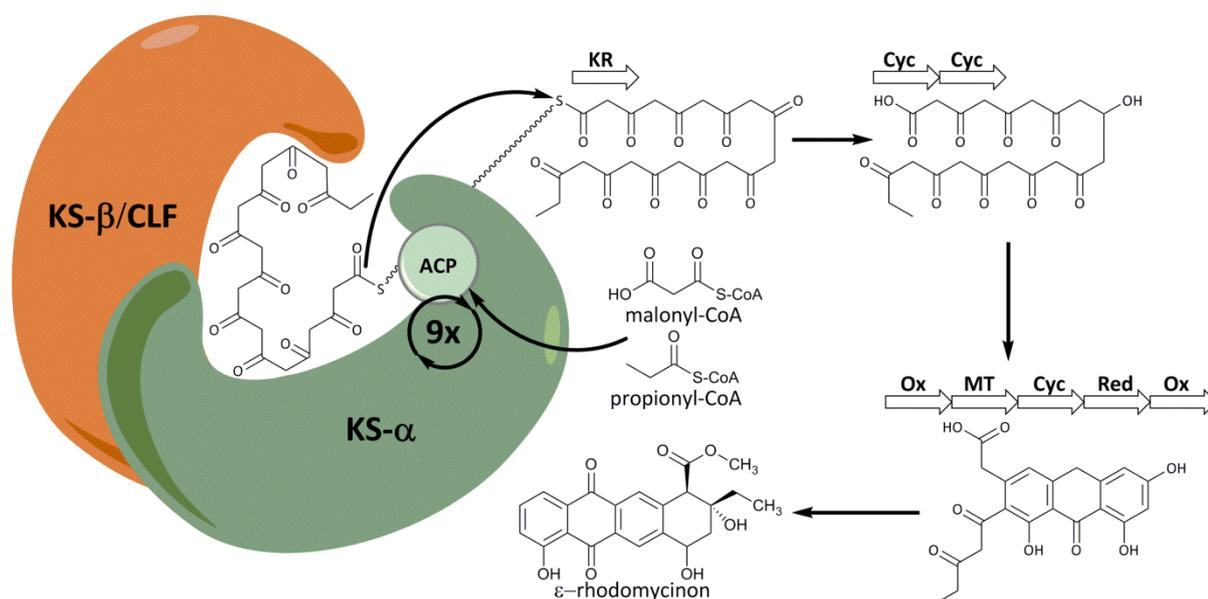


Figure 4: Schematic representation of the biosynthesis of ϵ -rhodomyconin.

1.2.3.1. Angucyclines

Angucyclines are the largest group of polycyclic aromatic type polyketides assembled by type II PKS. Since the discovery of tetrangomycin⁴³ in 1965, extensive natural product research resulted in the discovery of more than one hundred, mostly antibiotic and cytotoxic active, angucyclinones exclusively derived from Actinobacteria (Figure 5). Characteristically, members of this group contain a benz[a]anthracene moiety, which in some cases is drastically rearranged through extensive oxidative reactions and tailoring steps.^{13, 44} The biosynthesis is usually initiated with an acetyl-CoA starter unit and elongated with 9 malonyl-CoA units, resulting in the deca-ketide intermediate. Typically, a direct cyclization of the deca-ketide towards the benz[a]anthracene moiety is dictated by dedicated cyclases enzymes.^{8, 45} Nevertheless, in two cases an initial cyclization towards an anthracycline type

backbone with a subsequent rearrangement of ring A to the benz[a]anthracene moiety was reported for the biosynthesis of PD116198 and BE-7585A.⁴⁶⁻⁴⁷ A well-studied example of an angucycline biosynthesis is the jadomycin case. The ketoacyl-synthase JadA, the chain length factor JadB and the acyl carrier protein JadC form the initial decaketide, which is reduced by the ketoreductase JadE and cyclized by the cyclases JadD and JadI to the first known angucyclinone intermediate UWM6.⁴⁸⁻⁴⁹ The previously described steps are highly similar in the biosynthesis of the majority of angucyclinones. In the case of jadomycin, the formation

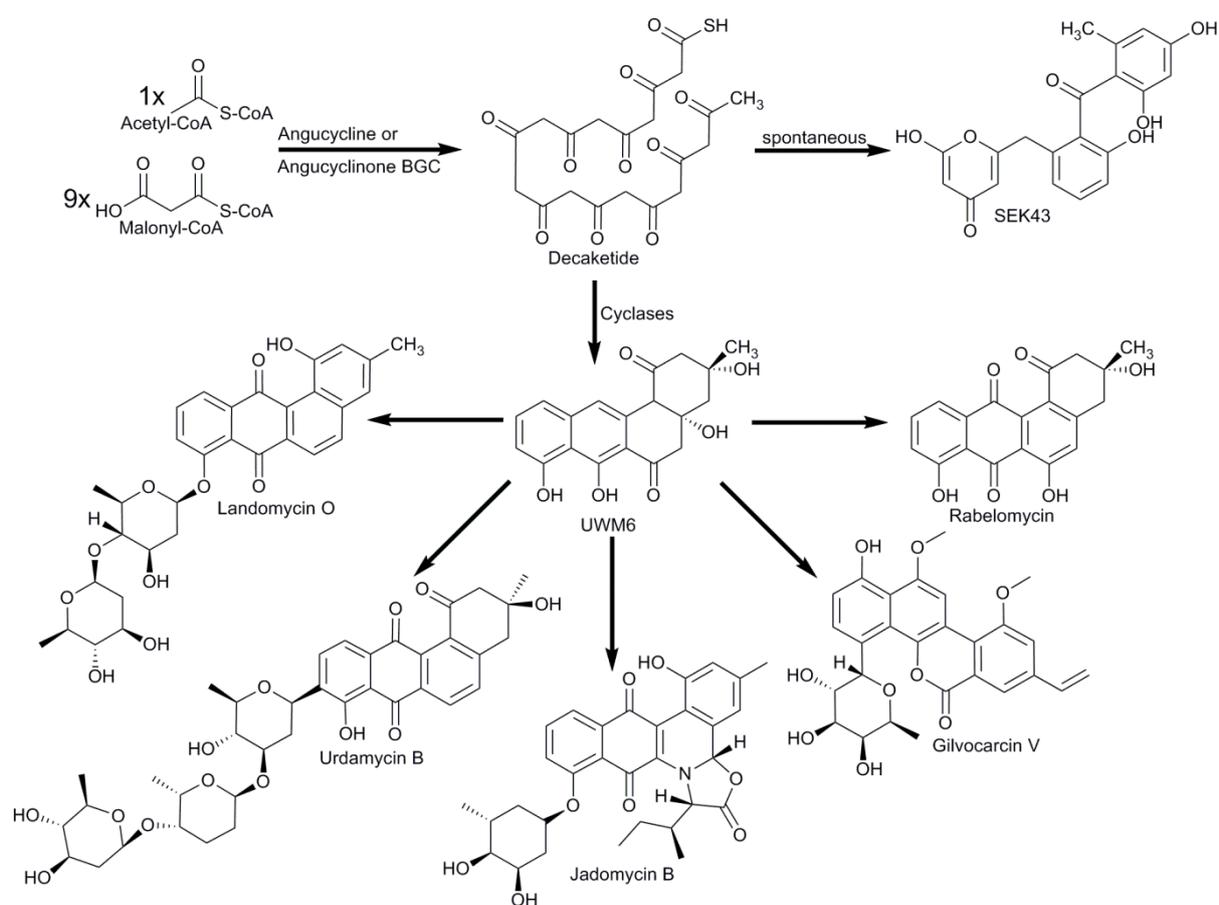


Figure 5 Examples of angucyclines and angucyclinones, important steps in their biosynthesis and commonly observed shunt products.

of UWM6 is then followed by a sequence of reductive and oxidative steps, resulting in the formation of dehydrorabelomycin and the oxidation and very rare cleavage of ring B. Thereafter, different amino acids can be incorporated in the B-ring, which to the best of our knowledge is unique for jadomycin.^{48, 50-54} The last step of jadomycin's biosynthesis is the glycosylation at the hydroxy group of ring D.⁵⁵ Considering the regulation of jadomycin's biosynthesis, five genes, known as *jadW*₁₋₃ and *jadR*₁₋₂, are responsible for its putative

control. While the *jadW* genes resemble gamma-butyrolactone autoregulators, a class of regulators known to be involved in the production of secondary metabolites and morphological differentiation, the genes *jadR₁* and *jadR₂* were identified as “atypical” response regulator (*jadR₁*) and “pseudo” GBL receptor (*jadR₂*). The influence of these genes on jadomycin’s production was investigated with mutational studies. The deletion of *jadW₁* had a negative impact on the growth and sporulation rate of *S. venezuelae* ISP5230 and on jadomycin’s production. Furthermore, a detailed study of *jadR₁* and *jadR₂* revealed them to be an interacting regulatory system, where *JadR₂* is controlling the expression of *jadR₁* and therefore, the expression of jadomycin’s biosynthetic genes.⁵⁶⁻⁶⁰ Additionally, the biosynthesis of Jadomycin B in *Streptomyces venezuelae* ISP5230 is controlled by an atypical mechanism and activated by environmental stress such as ethanol toxicity or heat shock.⁶¹

During the biosynthesis of jadomycin and other angucyclines and angucyclinones, errors in the cyclization of the nascent polyketide chain and processing of UWM6 can occur, resulting in the accumulation of different shunt products. Commonly observed shunt products are SEK43 and rabelomycin. The former was identified through the expression of tetracenomycins minimal PKS genes, actinorhodin’s ketoreductase and griseusin’s aromatase without a corresponding cyclase. It has been shown that the formed decaketide can spontaneously cyclize to SEK43.⁶² The latter was initially thought to be the main product of an angucyclinone BGC of *S. olivaceus* ATCC 21, but this is contradicted by the rabelomycin identification together with the native products in numerous heterologous expression experiments with for example gilvocarcins BGC.⁶³ Therefore, rabelomycin has been recognized as shunt product that occurs through the oxidation of the intermediate UWM6 followed by a spontaneous reduction.⁶⁴

1.2.3.2. Macro-poly-olide antibiotics

Macro-poly-olide antibiotics, a unique group of polyketides, can be divided in two sub-groups based on their structural features, macrotetrolide (nonactins) and the macrodiolides (pamamycins).⁶⁵⁻⁶⁹ Compounds of both classes possess an enormous variety of biological activity, including antifungal, antibacterial and anticancer activity, and contain a similar structural backbone that includes tetrahydrofuran moieties as well as four or two macrolactone forming ester bonds.⁷⁰⁻⁷²

1.2.3.2.1. Macrotetrolides

Macrotetrolides include nonactin and its derivatives monactin, dinactin, trinactin and tetranactin. Nonactin was the first macrotetrolide that has been identified.⁷² It is composed of four enantiomeric units of nonactic acid. Its biosynthetic gene cluster has been identified by cloning 55 kbp region using the *nonR* resistance gene as probe. An analysis of the BGC revealed the presence of five genes encoding ketosynthases (*nonJ, K, P, Q, U*) and four genes encoding ketoreductases (*nonE, M, N, O*), while no gene encoding an ACP has been found.^{70, 73-75} All ketosynthases of nonactin's BGC highly resemble ketosynthases involved in the biosynthesis of type II polyketides and fatty acids, which usually require an ACP to mediate the growing carbon backbone between the involved proteins. To exclude the involvement of a trans-ACP located in the genome of *S. griseus*, nonactin's BGC was successfully cloned and expressed in *Streptomyces lividans*. Extensive ²H, ¹³C and ¹⁸O labelled precursor feeding studies have been carried out in order to elucidate the biosynthesis of nonactin (Figure 6). It has been shown that the initial step for the formation of nonactic acid is the assembly of succinate and malonate (**N1**). Subsequently, the growing polyketide chain is elongated with malonate and methylmalonate (**N2, N3 & N4**). The ketoreductases (NonE, M, N, O) reduce the ketogroups to two hydroxy groups and one alkene. The CoA-ligase NonL and the hydratase NonS form the tetrahydrofuran moiety out of one hydroxy group and the alkene, resulting in nonactic acid. As a next step, the ketosynthases NonJ (**N5**) and NonK assemble four nonactic acid units to form nonactin and its derivatives. The derivatives of nonactin are formed by the incorporation of ethylmalonate instead of methylmalonate into the positions R₁-R₄.

1.2.3.2.2. Macrodiolides

The macrodiolides pamamycins are a group of natural products that contains more than 18 derivatives.⁶⁵⁻⁶⁹ The BGC of pamamycins was first identified by Rebets *et al.* through alignment of antiSMASH predicted BGCs in the genome of the two pamamycin producing strains *S. alboniger* DSMZ40043 and *Streptomyces sp.* HKI1 118. The 23 kbp large BGC is highly similar to nonactin's BGC with only a few differences: Pamamycin's BGC consists of two core regions left and right and contains two additional KS, an aminotransferase (*pamX*), a methyltransferase (*pamY*) and two regulatory elements (*pamR₁* and *pamR₂*).⁷⁶

Furthermore, an ACP has been identified in the left core of the cluster, indicating the necessity of an ACP as opposed to the biosynthesis of nonactin. An *in vitro* analysis of PamA

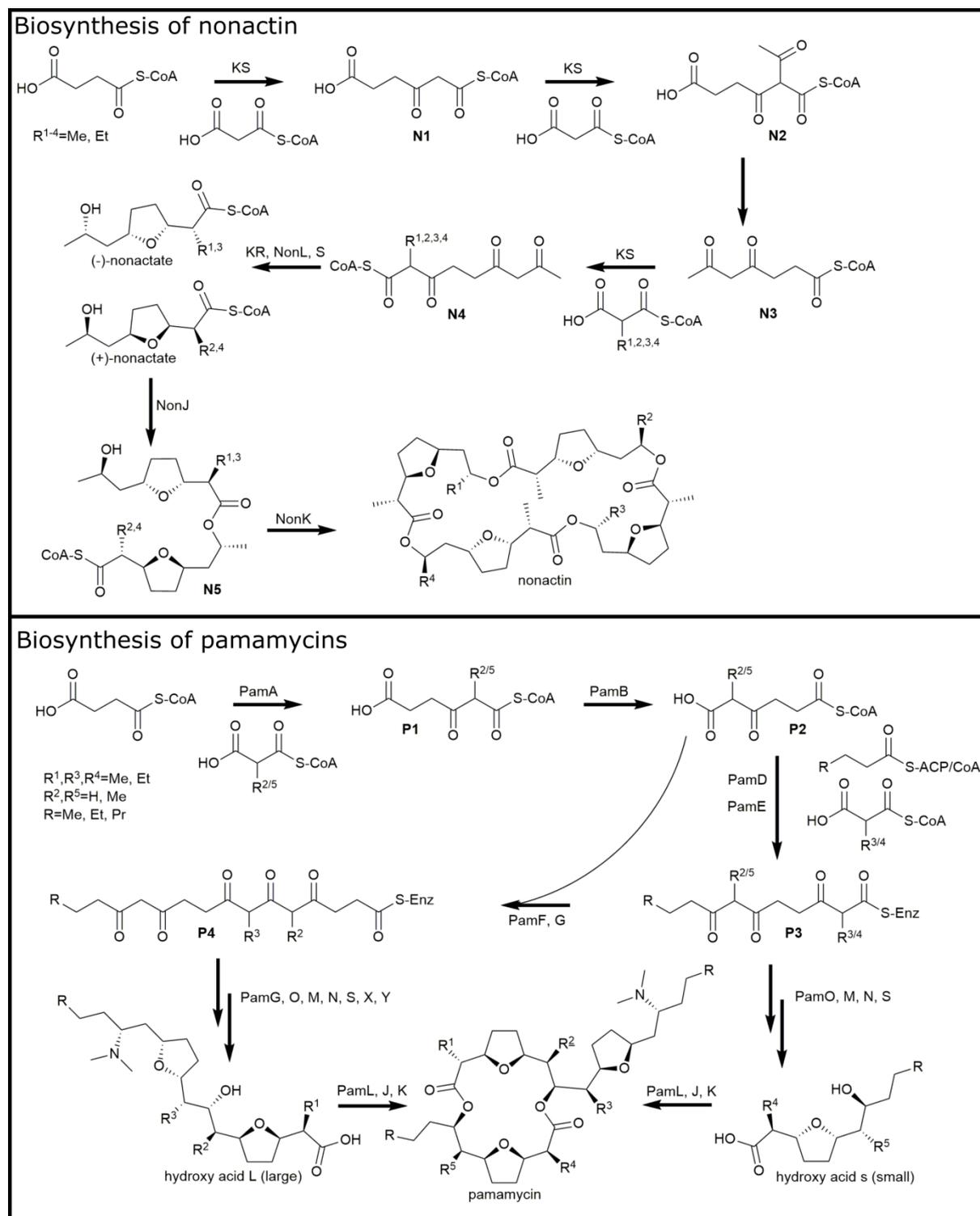


Figure 6 Structures of the macro-poly-olide antibiotics, Nonactin and Pamamycin, and their biosynthetic pathway.

confirmed its function as ketosynthase in the formation of the initial intermediate 3-oxoadipyl-CoA or 2-methyl-3-oxoadipyl-CoA (**P1**) through condensation of malonyl-CoA or

methylmalonyl-CoA and succinyl-CoA (Figure 6).⁷⁶ The product is rotated on the molecule of Co-enzyme A by the acyltransferase PamB. The resulting compound (**P2**) is elongated through a condensation with an acyl-CoA derivative (acetyl-CoA, propionyl-CoA or butyryl-CoA) followed by an elongation with methylmalonyl-CoA or ethylmalonyl-CoA by the KS PamD and PamE (**P3**). The resulting product is further modified in two different ways by pamamycin's biosynthetic enzymes. In the first branch, the ketosynthases PamF and PamG elongate **P3** to **P4** by adding one more molecule of **P2** and one molecule of methyl- or ethylmalonate, correspondingly. **P4** is subsequently modified towards the intermediate named hydroxy acid L (large). In the second branch, **P3** is used by pamamycin's biosynthetic enzymes (PamO, M, N, S) to form hydroxy acid S (small). The resulting intermediates, hydroxy acids S and L, are activated by the acyl-CoA ligase PamL and assembled by the ketosynthases, PamJ and PamK. With the elucidation of pamamycin's biosynthesis, the reason for the high derivative count has been identified: The involved ketosynthases utilize different malonyl-CoA derivatives, which results in sidechains of different lengths (R^{1-5}). Furthermore, non-methylated and partially-methylated amine derivatives were identified and through incorporation of different acyl-CoA units the peripheral side chains (R) can vary as well.⁷⁶ This variation in incorporation and the resulting high derivative count is one of the major obstacles for the development of downstream applications of pamamycins.

1.3. Polyketides biosynthesis precursors supply pathways

Despite the enormous variety of the chemical structures of natural products, their biosynthesis pathways are utilizing only a limited number of precursors that with some exceptions originate from the primary metabolism of the producing organism. In the case of polyketides only several common acyl-CoA derivatives are used. Mostly, these acyl-CoA esters are intermediates or products in the common metabolic pathways. Only some special elongation and starter units for the biosynthesis of polyketides are synthesized by dedicated enzymes encoded by genes within particular BGCs. Others, e.g., acetyl-CoA, malonyl-CoA, methylmalonyl-CoA, butyryl-CoA are accumulated within the network of anabolic and catabolic reactions (Figure 7). In order to understand the supply of precursors into the biosynthesis of secondary metabolites, the entire network of metabolic pathways within the producing organism should ideally be considered.

1.3.1. Acetyl-CoA and Malonyl-CoA

Due to the high similarity of PKS to FAS, it is not surprising that malonyl-CoA is one of the most commonly used precursors for the biosynthesis of polyketides. For the formation of malonyl-CoA, two possible routes are generally accepted. In the first pathway, acetyl-CoA is carboxylated by the acetyl-CoA carboxylase (ACC). ACC is an essential enzyme and crucial for the survival of the cell and usually multiple copies of its gene can be found in genomes.⁷⁷⁻⁷⁸ The second pathway includes the direct conversion from malonate to malonyl-CoA by the malonyl-CoA synthase (MatB).⁷⁹ Malonate as a precursor is supplied by the degradation of pyrimidines.⁸⁰⁻⁸¹ Considering the malonyl-CoA concentration, the overall influence of the second pathway is unknown and expected to be minimal.⁷ Usually, the gene encoding the malonyl-CoA decarboxylase can be found in proximity to the gene encoding the malonyl-CoA synthase. While the malonyl-CoA synthase activates malonate to malonyl-CoA, the malonyl-CoA decarboxylase transforms malonyl-CoA to acetyl-CoA. The proximity of the genes encoding both enzymes indicates the direct conversion of malonyl-CoA, derived from the pyrimidine pathway, to acetyl-CoA.⁸⁰

Acetyl-CoA is one of the most important primary metabolites and its supply is very complex. One of its main sources is the glycolysis. As a final step of the glycolysis, pyruvate is decarboxylated to acetate and activated to a CoA-ester by the pyruvate dehydrogenase complex.⁸²⁻⁸³ An additional supply of acetyl-CoA is through the β -oxidation of fatty acids.⁸⁴ During this process, the acyl-CoA dehydrogenase reduces the fatty acid between position 2 and 3 to an alkene. Subsequently, the enzymes enoyl-CoA hydratase and hydroxyacetyl-CoA dehydrogenase oxidize the fatty acid to a 3-keto-fatty acid. At last, the thiolase reaction results in the formation of acetyl-CoA and a shortened acyl-CoA derivative.⁸⁵ At low glucose levels acetyl-CoA can be supplied by various other pathways. First, acetate, if available, can be activated with a CoA-ester by the acetyl-CoA synthetase.⁸⁶ Second, ethanol can be a putative source through repetitive oxidation catalyzed by an alcohol dehydrogenase and an aldehyde dehydrogenase.⁸⁷ Third, branched chain amino acids, such as valine, leucine and isoleucine can be reduced to isobutyryl-CoA and mutated by the isobutyryl-CoA mutase to butyryl-CoA. Butyryl-CoA can further be degraded to acetyl-CoA through β -oxidation.⁸⁸

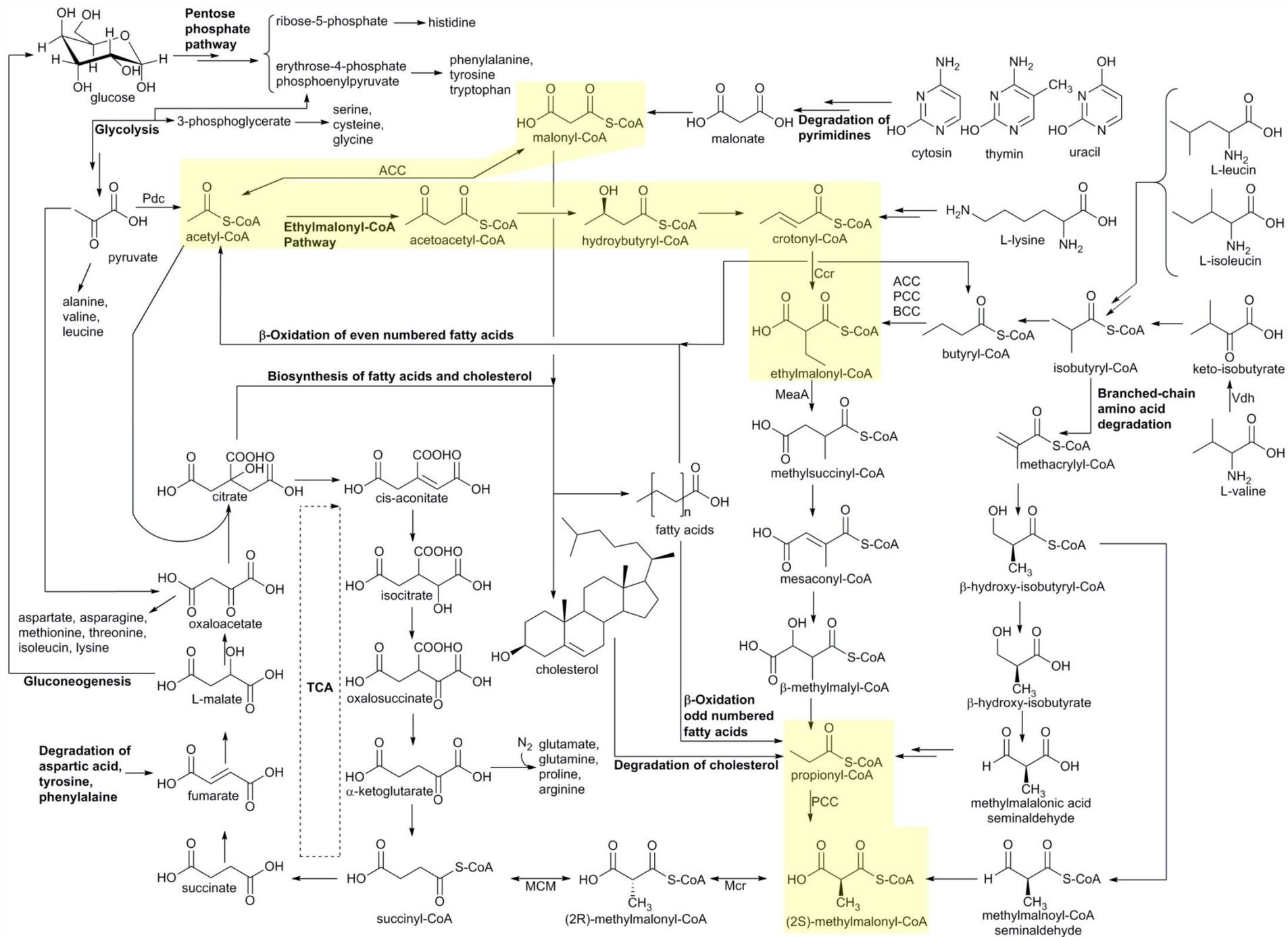


Figure 7: Important steps in the primary metabolism of precursor supply in prokaryotes.

1.3.2. Methylmalonyl-CoA, Propionyl-CoA and Succinyl-CoA

Methylmalonyl-CoA is commonly incorporated by type I PKS and is one of the main building blocks used in the biosynthesis of erythromycin.²⁸ Additionally, it has been reported to be used in the formation of specialized fatty acids within *Mycobacterium tuberculosis*.⁸⁹

In terms of the supply of methylmalonyl-CoA, three biosynthetic pathways have mainly been reported. The most influential source for the intracellular concentration of methylmalonyl-CoA is the carboxylation of propionyl-CoA by the propionyl-CoA carboxylase (PCC).⁹⁰ Propionyl-CoA can be derived from a variety of sources, including the degradation of odd-numbered and branched chain fatty acids, methionine, valine and isoleucine and the catabolism of cholesterol.⁹¹⁻⁹² Additionally, propionyl-CoA can be formed by the direct activation of propionate through an acyl-CoA ligase.⁹³ Once formed, propionyl-CoA can be carboxylated by the biotin dependent propionyl-CoA carboxylase.⁹⁰ Furthermore, loose substrate specificity of ACCs and butyryl-CoA carboxylases (BCC) have been reported and can therefore, also accept propionyl-CoA as substrate for carboxylation.⁹⁴

Succinyl-CoA is another putative predecessor of methylmalonyl-CoA. It is a key intermediate in the tricarboxylic acid cycle (TCA) and formed by the decarboxylation of α -ketoglutarate, catalyzed by the α -ketoglutarate dehydrogenase. The conversion of succinyl-CoA to (2S)-methylmalonyl-CoA occurs in two steps. First, the Methylmalonyl-CoA Mutase (Mcm) catalyzes a complex radical mutation from succinyl-CoA to (2R)-methylmalonyl-CoA. Thereafter, the methylmalonyl-CoA racemase converts the stereochemistry of the methyl-group to (2S)-methylmalonyl-CoA.⁹⁵⁻⁹⁷ The chemical balance of this two-step reaction is highly dependent on the growth state of the bacteria and the carbon source available in the medium.⁹⁸

Studies in *Streptomyces* species using labelled valine revealed a second pathway for the degradation of valine to methylmalonyl-CoA. The newly proposed pathway, like the original one, is initiated with the deamination of valine to 3-methyl-2-oxobutanoate by the valine dehydrogenase.⁹⁹⁻¹⁰⁰ The subsequent steps in the pathways ultimately lead to β -hydroxyisobutyryl-CoA. After the formation of this intermediate, the newly proposed pathway is branching from the classical pathway for valine degradation. In the original pathway, the

CoA-ester is firstly cleaved and secondly the hydroxy group is oxidized to an aldehyde. The pathway is finalized by the decarboxylation and the activation of the aldehyde group with a CoA-ester to propionyl-CoA. Considering the newly proposed pathway, the hydroxy group is immediately oxidized to an aldehyde, while the initial CoA-ester of hydroxy-isobutyryl-CoA is not cleft. The aldehyde is subsequently oxidized to a carboxyl group. Hence, the formation of propionyl-CoA is omitted and the pathway results in the direct formation of methylmalonyl-CoA.¹⁰⁰

1.3.3. Ethylmalonyl-CoA

Compared to its derivatives malonyl-CoA and methylmalonyl-CoA, ethylmalonyl-CoA is rarely incorporated into polyketides. Its biosynthetic roots can be reduced to two possible pathways. In the first pathway, a condensation of two acetyl-CoA units by the β -ketothiolase to acetoacetyl-CoA is followed by a reduction to hydroxy-butyryl-CoA by the acetoacetyl-CoA reductase. The subsequent dehydration results in crotonyl-CoA, which is reduced and carboxylated by the crotonyl-CoA-reductase/carboxylase to ethylmalonyl-CoA. This ethylmalonyl-CoA pathway assimilates acetate and supplies the tricarboxylic acid cycle (TCA) with succinyl-CoA. The pathway is an alternative to the glyoxylate- and methylaspartate cycle.¹⁰¹

Another putative source for ethylmalonyl-CoA is the carboxylation of butyryl-CoA. Butyryl-CoA can derive from a variety of pathways, including the degradation of fatty acids and several amino acids.¹⁰² Specialized butyryl-CoA carboxylases have been reported.⁷ Nevertheless, the relaxed substrate specificity of ACCs, discussed previously in Section 1.3.2, opens the possibility of the butyryl-CoA carboxylation by an ACC, which could additionally result in the formation of ethylmalonyl-CoA.⁹⁴

1.4. Genome mining for new Natural Products

In the early 2000s, the genomes of *Streptomyces coelicolor* and *Streptomyces avermitilis* were sequenced.¹⁰³⁻¹⁰⁴ Since the number of BGCs in both strains outpaces the number of the corresponding produced natural products, the idea of genome mining was initiated: This included the use of available genomic information for the prediction of BGCs; the activation of their expression and thus, the production, isolation and biological activity testing of the

corresponding natural product.¹⁰⁵ Later on, the identification of biosynthetic gene clusters of known natural products became another core interest of the genome mining approach.¹⁰⁵ Before genomic data became widely available, researchers had to take conserved regions of identified BGCs and use them as probes in southern blot.¹⁰⁶ With the emergence of next generation sequencing and expanding accessibility of genomic information, this tedious laboratory work was mostly replaced by *in silico* predictions of BGCs. The deepening of the understanding of biochemical mechanisms in the biosynthesis of particular groups of natural products led to the emergence of several bioinformatics tools dedicated to the identification of corresponding BGCs. The logical development of these tools led to their combination resulting in powerful and versatile softwares like antiSMASH and PRISM, which allows the prediction of many different types of BGCs within sequenced genomes.¹⁰⁷⁻¹⁰⁹ Additionally, the construction of arranged genomic libraries became more reasonable due to the significant progress in sequencing technology, generally leading to a decrease in the cost of the procedure.¹⁰⁷

Both aforementioned factors enabled the cluster-first approach: the screening of promising genomes for unique BGCs, the construction of genomic libraries and the heterologous expression of the identified BGCs. The cluster first approach includes the selection of promising candidate genomes. Doroghazi *et al.* showed that *Pseudonocardia* possess a similar number of BGCs per genome compared to *Streptomyces* and therefore, *Pseudonocardia* are an underutilized source for the discovery of new natural products.¹¹⁰

The strain *Saccharothrix espanaensis*, a member of the *Pseudonocardia*, was discovered by Labeda and Lechevalier in Spain in 1989¹¹¹. It forms yellow-brown mycelium and does not produce spores during his growth cycle. *Saccharothrix espanaensis* is known as producer of saccharomicins A & B, two derivatives with antibiotic activity¹¹². Their unique structure contains a small aglycon with 17 sugar moieties attached.¹¹²⁻¹¹³ The corresponding BGC was cloned and the biosynthesis of its aglycon has previously been elucidated¹¹⁴. Furthermore, the efficient glycosylation of polyphenolic compounds by this strain has been reported¹¹⁵. The genome of *Saccharothrix espanaensis* is 9.3 Mbp in size and codes for 8501 CDCs with an antiSMASH analysis predicting 31 putative biosynthetic gene clusters.¹¹³

1.5. Heterologous Hosts

Microbial heterologous hosts are usually unicellular organisms with advantageous characteristics, such as short multiplication time, genetic amenability and simple cultivation conditions. Initially developed heterologous hosts were mostly derivatives of *Escherichia coli* and used for the expression of proteins.¹¹⁶ Nowadays, heterologous hosts are used for a variety of purposes such as the expression and isolation of proteins for industrial and research purposes, the activation of silent BGCs and the generation of new natural product derivatives or increased production titers. Furthermore, with genome mining approaches they can be used for natural products discovery.¹¹⁷⁻¹²⁰ In contrast to native strains, heterologous hosts offer several advantages including rapid cell multiplication, well-analyzed metabolic networks and production conditions, implemented standard genetic manipulation protocols and cluster-free strains with a clean metabolic background that facilitates the downstream processing and increases the production titers.^{119, 121} The genome *S. coelicolor* A3(2), a commonly used host, was sequenced for the first time in 2002 and covers 8.6 Mbp including 7825 genes and 20 predicted BGCs.¹⁰³ Its plasmid free derivative *S. coelicolor* M145 was found to be unable to produce methylenomycin.¹²² M145 was used to develop the improved hosts M512, M1152 and M1154. The metabolic background of M512 was further reduced by deletion of pathway-specific activator genes (*actII-orf4* and *redD*). Therefore, the constructed derivative was unable to produce actinorhodin and prodiginine. M1152 and M1154 were specifically developed for the heterologous expression of BGCs and the subsequent screening for antibiotic activity.¹²² Hence, most BGCs responsible for the production of active natural products were deleted from the genome (actinorhodin, prodiginine, calcium-dependent antibiotic and a type I PKS cluster). Furthermore, M1152 and M1154 contain a mutation in the β -polymerase encoding gene *rpoB* [C1298T, S443L], while M1154 additionally contains a mutation in the ribosomal protein S12 encoding gene *rpsL* [A262G, K88E]. Both mutations in the aforementioned genes have been reported to increase the level of the secondary metabolite production. All modifications in the described strains lead to a decreased metabolic background and increased secondary metabolite production.¹²²

1.5.1. *S. albus* J1074

S. albus J1074 is a *Sall* deficient derivative of *S. albus* G. It was described by Chater and Wilde in 1976.¹²³ Its short multiplication time, good genetic amenability and production capabilities spurred the interest in *S. albus* J1074 as a heterologous host.¹²¹ At the same time, the development of efficient genetic modification tools, simple conjugation protocols and the development of a promotor library lead to its status as one of the most important *Streptomyces* hosts. Its 6.8 MPb linear genome covers 25 BGCs and is one of the shortest of the commonly used *Streptomyces* strains.^{121, 124} In 1981, *S. albus* J1074 was first used as cloning host and later as host for heterologous expression of steffimycin's BGC.¹²⁵⁻¹²⁷ Since then, its suitability as heterologous host has been proven by the expression of multiple BGCs, including fredericamycin, isomigrastatin, napyradiomycin, cyclooctatin, thiocoraline and moenomycin.¹²⁸⁻¹³¹ Recently, further attempts of host improvement were carried out by Myronovskyi *et al.* which resulted in the deletion of 15 native BGCs.¹³² In this study, 7.3 % (~500.000 bp) of the overall genome were deleted and additional plasmid/cosmid integration sites were added. These implemented changes led to a drastic simplification of the overall metabolic background, which resulted in an optimized availability of precursors, increased BGC copy numbers and thus, an improved production titer of numerous natural products (Tunicamycin B2, Didesmethylnensacarcin, Griseorhodin A, Aloesaponarin II, Pyridinopyrona A).¹³²

1.5.2. *S. lividans* TK24

Streptomyces lividans is phylogenetically closely related to *S. coelicolor* with a drastically altered metabolic profile. The genome of *S. lividans* contains the silent BGCs for actinorhodin, undecyleprodigiosin and calcium-dependent antibiotics, which are all produced by the *S. coelicolor*.¹³³ It has been shown that the mutation of *rpsL* and *rsmG* strongly activates the expression of actinorhodin's gene cluster. Furthermore, certain mutations of *rpoB* led to the activation of actinorhodin, undecylprodigiosin and calcium-dependent antibiotic.¹³⁴⁻¹³⁵ Even though *S. lividans* has been used as host for the expression of isomigrastatin, novobiocin, 6-desoxyerythronolide, puromycin, oxytetracycline, staurosporine, macrotetrolide, daptomycin, capreomycin and nikkomycin, only few host improvements were carried out until recently.^{74, 136-141} Bai *et al.* constructed *S. lividans* SBT5

out of *S. lividans* TK24 through the deletion of actinorhodin's, undecylprodigiosin and calcium-dependent antibiotic (CDA) BGCs.¹⁴² Furthermore, *S. lividans* was optimized for the production of mithramycin A through the deletion of up to 4 different BGCs.¹⁴³ An improved *S. lividans* TK24 host system was developed by Luzhetskyy *et al.* through the deletion of twelve native BGCs and the addition of two *attB*-sites into the genome. The suitability of the developed strain was proven by the expression of a variety of clusters and the comparison of production levels of the corresponding natural products (unpublished data).

1.6. Outline

The aim of the PhD project was to utilize heterologous hosts and BGC expression in order to address two problems: 1 – To access the chemical potential of the rare Actinobacteria *Saccharothrix espanaensis* for natural product discovery; 2 – To control the derivatization of pamamycins through engineering of the heterologous hosts precursor supply in order to facilitate the downstream processing, namely the isolation of pamamycins.

The work described in Section 2 approaches the common rediscovery issue of Actinobacteria derived natural products by the systematic expression of BGCs in heterologous strains and the identification of the two polyketides pentangumycin and SEK90. The genome of *Saccharothrix espanaensis* was analyzed with antiSMASH and the predicted BGC's were aligned with its sequenced genomic library. The genomic fragments encoded in the library were selected for heterologous expression based on the BGCs they encode. The utilization of improved heterologous hosts facilitated the natural product isolation and enabled the analysis of the biosynthetic pathways of both compounds.

Section 3 is focused on the promising group of polyketides pamamycins. After the identification of the corresponding BGC and its successful heterologous expression in *S. albus* J1074, the tuning of its produced derivatives was targeted. This should facilitate the isolation process and therefore, the development of pamamycins as putative drug candidate. Enzymes involved in pamamycins biosynthesis incorporate malonyl-CoA, methylmalonyl-CoA and ethylmalonyl-CoA according to their intracellular availability. We successfully showed that targeted knockouts of genes involved in the precursor supply in *S. albus* J1074 results in drastic changes of the intracellular CoA-ester concentrations and a modified spectrum of produced pamamycin derivatives.

2. Discovery of polyketides through genome mining of *Saccharothrix espanaensis*

2.1. Abstract

Natural products are an important source of novel investigational drugs in drug discovery. Especially in the field of antibiotics, Actinobacteria have proven to be a reliable source for lead structures. The discovery of these natural products with activity- and structure-guided screenings has been impeded by the constant rediscovery of previously identified compounds. Additionally, a large discrepancy between produced natural products and biosynthetic potential in Actinobacteria, including representatives of the order *Pseudonocardiales*, has been revealed using genome sequencing. To turn this genomic potential into novel natural products, we used an approach including the *in-silico* pre-selection of unique biosynthetic gene clusters followed by their systematic heterologous expression. As a proof of concept, fifteen *Saccharothrix espanaensis* genomic library clones covering predicted biosynthetic gene clusters were chosen for expression in two heterologous hosts, *Streptomyces lividans* and *Streptomyces albus*. As a result, two novel natural products, an unusual angucyclinone pentangumycin and a new type II polyketide synthase shunt product SEK90, were identified. After purification and structure elucidation, the biosynthetic pathways leading to the formation of pentangumycin and SEK90 were deduced using mutational analysis of the respective gene cluster and ¹³C-labelled precursor feeding experiments.

2.2. Introduction

The high importance of natural products in drug discovery is reflected in the fact that a high proportion of drugs that gained a marketing authorization between 1981 and 2010 originated from natural products.¹⁴⁴⁻¹⁴⁵ Amongst all natural products, microbial compounds present a large fraction (~70,000) with an impressive variety of chemical structures and biological activities.⁵ Actinobacteria, especially *Streptomyces*, represent one of the most important sources of natural products.¹⁴⁶ Nevertheless, during the last decades the discovery process in Actinobacteria was significantly slowed down through the re-isolation of known compounds in activity- and structure-guided screenings.⁹⁻¹⁰ As an early solution for the rediscovery issue, natural product databases have been created. Metabolic profiles recorded with high performance liquid chromatography coupled with high resolution mass spectrometry (LC-HRMS) can be compared with these databases. This process, known as dereplication, is problematic due to it being time consuming and error prone. First, the formation of adduct ions, the loss of hydroxyl groups and the occurrence of other fragmentations are commonly observed in mass spectrometry and can lead to false positives. Second, molecular masses, especially those of small molecules, can be identical and therefore lead to false negatives.¹¹ An additional problem of the described “compound first” approach is the need for laborious studies in order to identify the biosynthetic gene clusters (BGCs) that are responsible for the production of the isolated compounds. In contrast, the “cluster to compound” approach aims to obtain new natural products after cloning and expression of a particular BGC of interest in heterologous hosts. This approach became increasingly reasonable due to the low cost of next generation sequencing¹⁰⁷ which has resulted in a flood of genomic information.¹⁴⁷ The obtained information revealed a discrepancy in Actinobacteria between the genome encoded BGCs and produced compounds.¹¹⁰ It has been shown that *Streptomyces* contain an average of 21.9 BGCs per genome, with 40-48% of these BGCs being unique.¹⁴⁸ The “cluster to compound” approach aims to tap this genomic potential and offers decisive advantages over the “compound first” approach. First, genome analysis and BGCs prediction tools such as antiSMASH¹⁴⁹ offer automated cluster homology comparisons within hours. In turn, this allows cataloguing and prioritization of BGCs with the unique features. Second, the utilization of optimized heterologous hosts and the expression of BGCs from genomic libraries offer advantages over

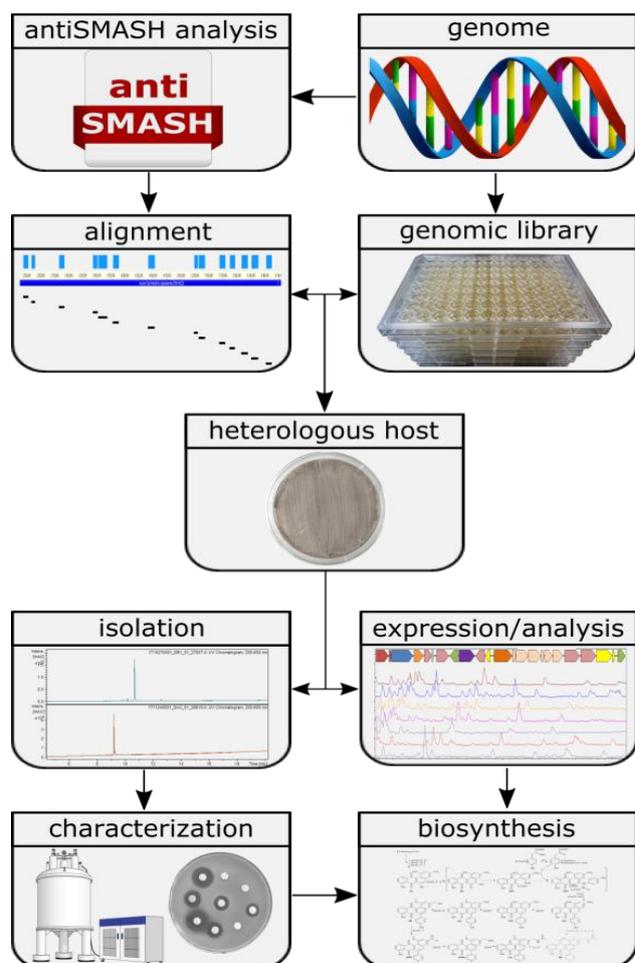


Figure 8: Workflow for the identification of novel natural products using genome mining.

the use of native producing strains.^{122, 132} The specialized hosts typically are genetically tractable, ideal growth and production conditions can be easily identified and standard genetic procedures, such as conjugation,¹⁵⁰ the λ Red-mediated gene deletions¹⁵¹ and the overexpression of single genes and BGCs, are well established. Furthermore, the produced compounds can be identified by simple comparison of the metabolic profiles from the heterologous host with and without the expressed BGCs.¹³² The knowledge of the expressed BGC combined with the LC-HRMS analysis minimizes the chances for the rediscovery of natural products.

The aim of the current study was to assess the untapped potential for the

production of secondary metabolites hidden in the genomes of Actinobacteria. To achieve this goal, a widely applicable workflow was developed (Figure 8). This workflow includes the analysis of genome sequences utilizing algorithms such as antiSMASH¹⁰⁸ or PRISM,¹⁰⁹ the construction of a sequenced genomic library, *in silico* pre-selection and expression of the selected BGCs in heterologous hosts followed by the identification, isolation and characterization of natural products. To challenge the applicability of the designed approach, we aimed to express BGCs from a strain of Actinobacteria that is more distantly related to commonly used heterologous *Streptomyces* hosts. The large-scale analysis by Doroghazi and co-authors revealed that the order of *Pseudonocardiales* has an average of 19.8 BGCs per genome.¹¹⁰ The strain *Saccharothrix espanaensis*, discovered by Labeda and Lechevalier in 1989¹¹¹ and belonging to the order of *Pseudonocardiales*, harbored all of the desired elements to prove the potential of our approach. This strain has a 9.3 Mbp genome

containing at least 31 antiSMASH-predicted BGCs.¹¹³ Due to the strain being poorly genetically tractable, any genetic manipulations are almost impossible. Furthermore, only a single group of polysaccharide natural products, the saccharomicins¹¹², is known to be produced by *S. espanaensis*.

As proof of concept, we report the discovery of two novel natural products, which were obtained by applying the proposed genome mining approach. Neither of the isolated compounds was detectable in the metabolic profile of *S. espanaensis*. Additionally, we were able to elucidate the biosynthesis of both compounds using the available information about the respective BGC architecture, mutational analysis and feeding experiments with labelled precursors.

2.3. Results and Discussion

2.3.1. Selection of putative biosynthetic gene clusters

Initially, the biosynthetic potential of *S. espanaensis* was assessed. An antiSMASH analysis of its genome predicted the presence of 31 BGCs, while the only known products are saccharomicins (Table 1).¹¹² The BGC of saccharomicins was identified as cluster number 28.¹¹³ It was originally cloned as cosmid clone by Berner *et al.*, while attempts of its complete expression in heterologous hosts had failed. Nevertheless, its aglycon was successfully expressed and its biosynthesis was elucidated.^{114, 152} A bacterial artificial chromosome (BAC) library of *S. espanaensis* was constructed using the pSMART-BAC-S vector as backbone with an average inserted fragments size of 100 kbp. The library was end-sequenced and the sequences were mapped to the genome of *S. espanaensis*. This allowed the alignment of predicted BGCs to the clones of the genomic library. To minimize the laboratory effort and reduce the putative rediscovery rate of the described approach, BGCs were prioritized by their class and predicted products. First of all, eight BGCs were not covered by the genomic library and therefore, could not be included into further work. Of the remaining 23 BGCs present as whole in the BAC library, three were predicted to be responsible for the production of geosmin, melanin and bacteriocin, and thus were excluded. Three BGCs shared more than 75% similarity to clusters of assigned products and were therefore not used for heterologous expression. Finally, a predicted terpene cluster lacking its synthase and a lantipeptide BGC were excluded as well. Thereafter, 15 remaining BACs covering 17 BGCs were selected for heterologous expression in *S. lividans* Δ YA6 and *S. albus* J1074 (Table 1). Among those were two terpene clusters, five type I and type II polyketide synthase (PKS) clusters, five nonribosomal peptide synthetase (NRPS) clusters, two type I PKS/NRPS hybrid clusters, one lantipeptide cluster, one aminoglycoside cluster and one polysaccharide cluster encoding saccharomicins BGC.

Table 1: Biosynthetic gene clusters identified in *S. espanaensis*.

#	Prediction: ^a	Predicted Cluster: ^b	Homology: ^c	BAC: ^d	Expressed: ^e
1	Terpene	Geosmin	100%	known	No
2	Lanthipeptide	Erythreapeptin	75%	homology	No
3	Terpene	Isorenieratene	42%	ex: 1C15	No
4	Bacteriocine			known	No
5	Furan	Asukamycin	30%	ex: 3A24	No
6	NRPS			ex: 3E7	No
7	Other	A54145	3%	not covered	No
8	NRPS	Myxochelin	50%	ex: 1G5	No
9	Indole	Frankiamicin	14%	ex: 1F6	No
10	Ladderane, NRPS	Skyllamycin	22%	ex: 1F6	No
11	Linear azole containing Peptides	A201A	6%	ex: 1I20	No
12	Type I PKS	Tylactone	6%	ex: 3E19	No
13	NRPS	Tyrobetaine	53%	ex: 3K5	No
14	NRPS			not covered	No
15	NRPS, Type I PKS	Kedarcidin	18%	not covered	No
16	Lanthipeptide	Kinamycin	5%	ex: 1G11	No
17	Oligosaccharide	Teicoplanin	4%	not covered	No
18	Melanin			known	No
19	NRPS, Type I PKS	Leinamycin	15%	not covered	No
20	NRPS, Type I PKS	Lavendiol	35%	ex: 3C18	Yes
21	Terpene	Isorenieratene	85%	homology	No
22	NRPS	Cyclomarin	13%	ex: 1L8	No
23	Aminoglycoside			ex: 3M21	No
24	NRPS	Ficellomycin	3%	ex: 3M21	No
25	RIPP	Anantin C	75%	homology	No
26	Terpene, Type II PKS, Type I PKS, NRPS	Fluostatin	23%	ex: 1E5	Yes
27	terpene,NRPS,T1PKS	Ficellomycin	27%	ex: 1C7	No
28	Oligosaccharide	Desosamine	22%	ex: 3K17	No
29	Terpene			excluded	No
30	Terpene	SF2575	6%	excluded	No
31	Lanthipeptide	Olimycin A	8%	excluded	No

a: Prediction: Predicted class of BGC;

b: Predicted cluster: Cluster with the highest homology;

c: Homology: Homology between the predicted cluster and the BGC identified in the genome of *S. espanaensis*;

d: Work: known: Cluster was known and excluded; homology: due to high homology of the whole cluster, the cluster was excluded; not covered: Cluster was not covered by our library; ex: "XXX" Cluster was covered by the BAC "XXX" and chosen for expression in our host;

e: Expressed: Indication of successful expression in our heterologous hosts

2.3.2. Success rate of the expression system

The chosen BACs were conjugated in *S. lividans* Δ YA6 and *S. albus* J1074. A colony PCR (Polymerase chain reaction) verified exconjugant was cultivated in different production media and metabolites were extracted using different solvents. Their metabolic profiles were analyzed with LC-MS and compared to the metabolic profiles of the empty heterologous hosts. Singular peaks were identified and assumed to be the products, shunt-products or intermediates of the expressed BGC. Extracts containing singular peaks were further analyzed by LC-HRMS, and the obtained exact masses were compared to masses in common natural product databases (DNP [<http://dnp.chemnetbase.com>], Supernatural, StreptDB).¹⁵³⁻¹⁵⁴

The recombinant strain *S. lividans* Δ YA6_3C18, containing a 115 kbp insert of the *S. espanaensis* genome, revealed a singular peak (Figure S 1) with a mass of 308.1987 Da (m/z 309.2060 [M+H]⁺, R_t =11.5 min). BAC clone 3C18 carries a type I PKS cluster similar to lavendiols BGC¹⁵⁵ and an unknown NRPS BGC. Further work with the expressed compound was neglected due to its low production rate. The expression of 3C18 was unsuccessful in *S. albus* J1074.

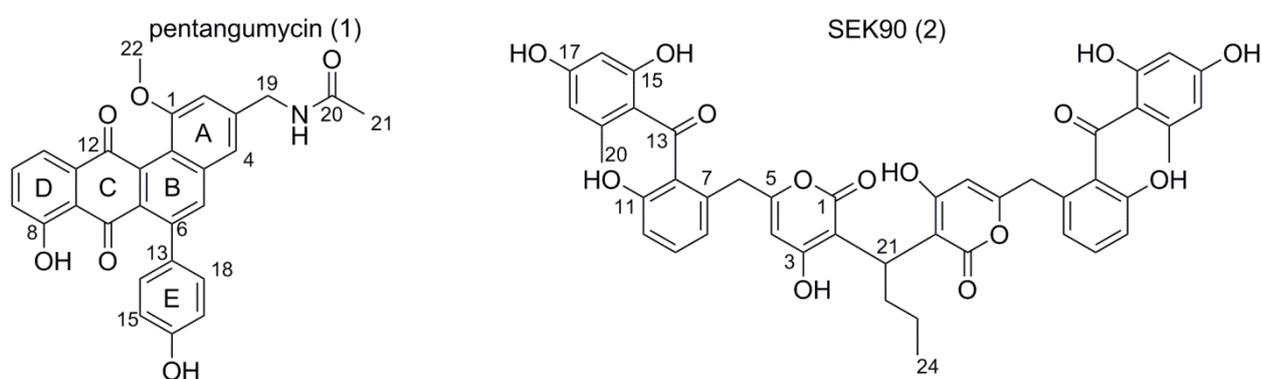


Figure 9: Structural formulas of pentangumycin and SEK90.

The recombinant strain *S. lividans* Δ YA6_1E5, containing a 116 kbp insert from the genome of *S. espanaensis*, produced two singular peaks with masses of 467.1368 Da (m/z 468.1438 [M+H]⁺, R_t =9.1 min, further designated as pentangumycin **(1)**) and 790.2261 Da (m/z 791.2339 [M+H]⁺, R_t =10.6 min; further designated as SEK90 **(2)**) as shown in Figure 9. *S. albus* J1074_1E5 produced compound **1** (Figure S 2), but its low production titer rendered

further work infeasible. Both compounds **1** and **2** have not been observed in extracts obtained from *S. espanaensis* (Figure S 3).

In the case of the remaining 13 BAC clones, no production was observed under the tested conditions. An overall success rate of 11% (corresponding to 2 out of 17 BGCs, BAC_3C18 and BAC_1E5) was observed when expressing BGCs from *Pseudonocardiales* in *Streptomyces* hosts. In similar experiments with BGCs from a *Streptomyces* strain, a success rate of 35% (6 out of 17 BGCs) has been observed (unpublished data). The generally low expression rate can be explained by a variety of factors: 1 – Differences in the regulatory network of strains can cause major shifts of natural product production. Prominent examples include the mutations in the genes [C1298T, S443L] and *rpsL* [A262G, K88E] in the strains *S. coelicolor* and *S. lividans*, which activated and increased the production of natural products in both strains.¹²² 2 – Codon bias of different strains can be restrictive for the expression of BGCs in heterologous hosts. In fact, it has been shown to be one of the most important factors for the expression of prokaryotic genes.¹⁵⁶⁻¹⁵⁷ In *E. coli* the heterologous expression of genes was increased by expanding the intracellular concentration of rare tRNA.¹⁵⁸ In *S. lividans*, the expression of a transglutaminase was increased by 73.6%, when using a codon optimized gene for expression.¹⁵⁹ 3 – Promoters and ribosomal binding sites not derived from the native strain can be unfitting for the polymerase of the host strain and therefore, disrupt the expression of BGCs. The utilization of synthetic and known strong promoters can activate the expression of native and heterologously expressed BGCs. Salas *et al.* successfully activated the expression of an NRPS and PKS-NRPS hybrid BGC by introducing the strong and constitutive *ermE**p promoter in front of the NRPS (*sshg_00313*) and PKS-NRPS (*sshg_05713*) genes. The activation of the NRPS BGC was confirmed by the production of a blue pigment. Furthermore, they identified the natural products 6-epi-alteramid A and B, which were produced by the PKS-NRPS BGC.¹⁶⁰ 4 – The toxicity of either expressed proteins or produced natural products is an additional issues to keep in mind when expressing BGCs in heterologous hosts. For example, a mutated version of an outer membrane protein (OmpA) of *E. coli* K12 caused a toxic lysis of the cell when expressed.¹⁶¹ Furthermore, the toxicity of avermectin was shown to be a restrictive factor in its own production in *S. avermitilis*. After the introduction of multiple copies of the ABC transporter AvtAB, the avermectin production increased from 3.3 g/l to 4.8 g/l.¹⁶² 5 – Environmental factors can

drastically influence the production of natural products. A well-studied example thereof is the production of jadomycin, since it was only produced after the strain was either treated with a heat shock or the production media was supplemented with ethanol.⁶¹ Furthermore, it has been proven that the addition of sugars¹³ and sub-toxic concentrations of antibiotics can activate and strongly alter the expression of BGCs in bacteria.¹⁴⁻¹⁵

2.3.3. Isolation of pentangumycin and SEK90

Ultimately, 3.5 mg of pentangumycin **1** and 14 mg of SEK90 **2** were isolated from 15 L of a *S. lividans* Δ YA6_1E5 production culture. Both compounds were purified using size exclusion chromatography and preparative HPLC. Through analysis of LC-HRMS data and NMR data (¹H-NMR, HSQC, HMBC, ¹H-¹H-COSY, ¹³C-NMR), the structures of both compounds were elucidated (Figure 9). NMR data and the detailed data analysis are presented in Table 2 and Figure S 4 - Figure S 50. All chemical shifts observed for **1** were within expected ranges as calculated for its structure. The connection between ring E and B was observed through HMBC correlations between H-5 and C-13 and between H-14/H-18 and C-6 (Figure 9, Figure S 4). **1** is a member of the angucyclinone group containing a unique fifth phenol ring attached through a direct carbon-carbon bond at position C-6. Additional ring systems, typically derived from amino acids, have previously been observed for the angucycline family of aromatic polyketides. While such structures have been found in several jadomycins and urdamycins, their position and type of connection are different.¹⁶³⁻¹⁶⁵ Additionally, formicamycins and fasamycins, both anthracycline antibiotics, carry a similar structural motif with the fifth aromatic ring at position C-7 of ring B.¹⁶⁶ Furthermore, **1** has a methylated hydroxy group at position C-1 in ring A, a modification that can be found only in a few angucyclines, like the chlorocyclinones.¹⁶⁷ In addition, the drastic structural feature of **1** is the presence of an aminated and subsequently, acetylated methyl group at position C-19, that has not been previously reported.

Table 2: NMR spectroscopic data for pentangumycin (**1**) (DMSO-d₆) and SEK90 (**2**) (DMSO-d₆).

pentangumycin (1)				SEK90 (2)			
Pos.	δ_C	δ_H (J in Hz)	HMBC	Pos.	δ_C	δ_H (J in Hz)	HMBC
1	156.72, C	-	-	1	166,5 C	-	-
2	108.73, CH	7.10, s	1, 4, 12b, 19	2	102,7 C	-	-
3	143.23, C	-	-	3	165,5 C	-	-
4	117.52, CH	7.44, s	1, 2, 4a, 5, 12b, 19	4	101,9 CH	5,68 (s br)	2, 5, 6
4a	136.62, C	-	-	5	160,6 C	-	-
5	135.12, CH	7.88, s	1, 4, 4a, 6a, 7, 12a, 12b, 13	6	36,1 CH ₂	3,55 (s br)	4, 5, 7, 8, 12
6	139.11, C	-	-	7	132,5 C	-	-
6a	131.27, C	-	-	8	120,1 CH	6,74 dd	9, 10, 11, 12, 13
7	188.02, C	-	-	9	130,9 CH	7,2 dd	7, 8, 10, 11, 12
7a	115.80, C	-	-	10	114,4 CH	6,77 dd	8, 9, 11, 12, 13
8	160.52, C	-	-	11	153,3 C	-	-
9	122.98, CH	7.26, d	7a, 11	12	130,6 C	-	-
10	136.66, CH	7.73, t (7.5)	8, 11a	13	199,9 C	-	-
11	117.20, CH	7.50, dd (7.5, 0.93)	7a, 9, 10, 8, 12	14	115,3 C	-	-
11a	135.51, C	-	-	15	164,9 C	-	-
12	185.53, C	-	-	16	100,5 CH	6,11 d (2,4)	14, 15, 17
12a	139.97, C	-	-	17	166,3 C	-	-
12b	119.48, C	-	-	18	111,4 CH	6,04 dd (0,64 Hz & 2,4)	13, 14, 15, 16, 20
13	132.18, C	-	-	19	142,7 C	-	-
14/18	129.94, CH	7.26, d	6, 14/18, 16	20	21,2 CH ₃	1,8 s	14, 18, 19
15/17	114.77, CH	6.78, dt (8.5, 2)	13, 15/17, 16	21	29,5 CH	4,36 t (8,2)	1, 2, 3, 22, 23
16	156.58, C	-	-	22	31,5 CH ₂	1,75 dt (7,3 Hz & 8,2)	2, 21, 23, 24
19	42.37, CH ₂	4.43, d (6)	2, 3, 4, 20	23	20,6 CH ₂	1,05 dq (7,3)	21, 22, 24
20	169.55, C	-	-	24	13,8 CH ₃	0,89 t(7,3)	22, 23
21	22.75, CH ₃	1.93, s	20	-	-	-	-
22	56.12, CH ₃	3.88, s	1	-	-	-	-
NH	-	8.53, t (br)	19, 20	-	-	-	-
OH ₁	-	-	-	-	-	12,67 s	14, 15, 16
OH ₂	-	-	-	-	-	11,49 s (br)	-
OH ₃	-	-	-	-	-	10,38 s	16, 18
OH ₄	-	-	-	-	-	9,78 s	9, 10, 11, 12

The NMR spectra of **2** contained less carbon and proton signals than anticipated regarding its observed exact mass. The ¹H-NMR signal of H-21 (CH) indicated the presence of a mirror plane, since all other integrated proton signals were a multiple of 1 (2-6). Another strong indicator was the low number of carbon peaks observed in the ¹³C-NMR (24 peaks in ¹³C-NMR), which only sums up to the corresponding exact mass of **2** with an unreasonable number of incorporated heteroatoms. After the identification of the mirror plane, the monomer SEK43 (Figure S 51) could easily be elucidated.⁶² The protons H-16, H-18 and H-20 as well as H-6, H-8, H-9 and H-10 showed all necessary HMBC and ¹H-¹H-cosy correlations for the elucidation of their corresponding aromatic system. The connecting ketone moiety was elucidated through the correlations of H-18 and H-10 to C-13 and the carbon shift of C-13 (199.9 ppm). The pyrone moiety of SEK43 was elucidated through correlations between H-6 to C-5 and C-4, H-4 to C-1, C-2, C-4 and C-6, H-21 to C-1 and C-3 and H-22 to C-2. The connecting butyl group showed all expected ¹H-¹H-cosy and HMBC correlations (Figure S 28). The putative axial stereochemistry of **2** can be assumed to be a racemic mixture due to its spontaneous formation as described below.

The biological activity of **1** and **2** was tested. Both compounds showed no antimicrobial activity over a range of tested concentration (0.2 µg/ml – 100 µg/ml; data not shown). In a cytotoxicity test, **1** showed an IC₅₀ of 18.77 µM against HuH7.5 cells and an IC₅₀ of 24.89 µM against HCT116 cells, whereas **2** was not active (Figure S 52 & Figure S 53). In a CAS assay, **2** showed a low iron-binding siderophore activity (Figure S 54).¹⁶⁸

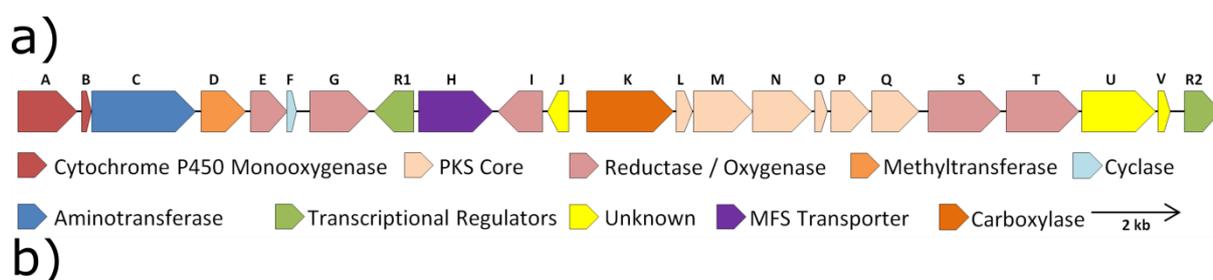
2.3.4. Both (1) and (2) are the result of type II PKS gene cluster expression

The antiSMASH analysis of the sequence of BAC 1E5 revealed the presence of three putative BGCs: a terpene cluster, a type-II PKS and a PKS–NRPS hybrid. From a structural perspective, it is obvious that neither **1** nor **2** are products of the terpene BGC, and that **1** is a product of the type II polyketide synthase. However, **2** cannot easily be associated with the type II PKS or the PKS-NRPS hybrid. Even though the SEK43 moiety of **2** is a known shunt product of type II polyketide synthases, the connection through a butyl-group has never been reported. In order to verify the biosynthetic origin of **1** and **2**, we aimed to identify the core genes of the corresponding BGCs and deleted them from BAC 1E5. A BLAST analysis of the type II PKS BGC

showed that gene BN6_54860 is coding for a β -ketoacyl-synthase (Figure 10) and gene BN6_55110 is encoding a type I polyketide synthase in the PKS-NRPS BGC. Both genes were deleted separately in frame from BAC 1E5. The resulting BACs, 1E5 Δ 54860 and 1E5 Δ 55110, were conjugated into *S. lividans* Δ YA6, and the metabolic profiles of both recombinant strains were analyzed. The production in the strain carrying mutant BAC 1E5 Δ 55110 remained unchanged, whereas the deletion of gene BN6_54860 resulted in the loss of production of **1** and **2** (Figure S 55). This clearly proves the origin of both compounds to be the type II PKS gene cluster.

2.3.5. Detailed analysis and border prediction of the biosynthetic gene cluster of pentangumycin

The putative borders of pentangumycins BGC are defined by the genes BN6_54730 (further



Gene#	Name	Homology to:	Annotation	Ident. [%]	Cover. [%]
BN6_54730	<i>penA</i>	WP_124773694	cytochrome P450	71.1	99
BN6_54740	<i>penB</i>	WP_124773693	ferredoxin	64.6	98
BN6_54750	<i>penC</i>	WP_124773692	aminotransferase class III-fold, pyridoxal phosphate-dependent	59.5	99
BN6_54760	<i>penD</i>	WP_124773691	methyltransferase	68.8	99
BN6_54770	<i>penE</i>	WP_015037167(<i>jadG</i>)*	antibiotic biosynthesis monooxygenase	58.8	90
BN6_54780	<i>penF</i>	WP_007904366	ester cyclase	46.3	96
BN6_54790	<i>penG</i>	WP_005164912	FAD-dependent oxidoreductase	41.7	90
BN6_54800	<i>penR1</i>	WP_034626020	XRE family transcriptional regulator	57.4	95
BN6_54810	<i>penH</i>	WP_124773686	MFS transporter	49.2	89
BN6_54820	<i>penI</i>	WP_124776097	aldo/keto reductase	61.1	76
BN6_54830	<i>penJ</i>	WP_124776095	nuclear transport factor 2 family protein	75.5	77
BN6_54840	<i>penK</i>	WP_015035778	acetyl-/propionyl-CoA carboxylase subunit alpha	69.8	99
BN6_54850	<i>penL</i>	AAD13535.1 (<i>lanF</i>)	Polyketide cyclase	74.1	98
BN6_54860	<i>penM</i>	AAD13536.1 (<i>lanA</i>)	Beta-ketoacyl-ACP synthase homolog	73.3	99
BN6_54870	<i>penN</i>	AAD13537.1 (<i>lanB</i>)	Polyketide chain length factor	68.2	99
BN6_54880	<i>penO</i>	AAD13538.1 (<i>lanC</i>)	acyl carrier protein	69.4	97
BN6_54890	<i>penP</i>	AAD13539.1 (<i>lanD</i>)	ketoreductase	77.0	99
BN6_54900	<i>penQ</i>	AAD13540.1 (<i>lanL</i>)	aromatase/cyclase	65.6	96
BN6_54910	<i>penS</i>	WP_015037166(<i>jadF</i>)*	oxidoreductase	69.7	72
BN6_54920	<i>penT</i>	AAV52248(<i>jadH</i>)*	oxidoreductase	66.8	99
BN6_54930	<i>penU</i>	WP_046087459	acyl-CoA carboxylase subunit beta	81.8	99
BN6_54940	<i>penV</i>	WP_062205789	acyl-CoA carboxylase subunit epsilon	38.3	73
BN6_54950	<i>penR2</i>	AAB36584(<i>jadR1</i>)	Phosphate regulon transcriptional regulatory protein PhoB	57.9	95

Genes bold: Part of Micromonospora LB.39

* putatively involved in ring opening of Ring B

Figure 10: Schematic representation of the biosynthetic gene cluster of pentangumycins (a) and) BLAST analysis of the pen genes (b).

as *penA*) on a left edge and BN6_54950 (further as *penR2*) on a right (Figure 10a). In order to predict their functions, individual *pen* genes were analyzed using BLASTx. Detailed results are illustrated in Figure 10b. The core of the *pen*-cluster is formed by genes *penM*, *penN* and *penO*, coding for the minimal polyketide synthase. Together with two cyclase genes, *penL* and *penF*, the cyclase/aromatase *penQ* and the ketoreductase *penP* are forming a minimal set of genes required for the biosynthesis of the angucyclinone core structure. The polyketide core region is surrounded by seven genes putatively involved in oxidative or reductive reactions. The gene *penA* encodes a cytochrome P450 enzyme and is adjacent to *penB*, which encodes a ferredoxin protein. Both genes have previously been described as functional unit.¹⁶⁹⁻¹⁷¹ The genes *penE*, *penG*, *penI*, *penS* and *penT* encode a monooxygenase, three oxidoreductases and an aldo-/ketoreductase, respectively. Except for PenE, PenS and PenT, the function of these enzymes cannot be predicted from the BLAST analysis. *penE*, *penS* and *penT* encode orthologues of JadG, JadF and JanH, respectively, which are three enzymes putatively involved in the oxidative opening of ring B of jadomycin.⁴⁹ Other genes such as *penC* (encoding putative aminotransferase) and *penD* (encoding putative methyltransferase) are also supposed to be involved in post-PKS modifications of **1**. *penK*, *penU* and *penV* encode α -, β - and ϵ -subunits of a carboxyltransferase, respectively. The function of these enzymes in the biosynthesis of **1** is unclear. Two regulatory genes, *penR1* encoding putative transcriptional regulator with the predicted helix-turn-helix motif at N-terminus and *penR2* coding for a SARP protein, are present in the cluster.⁶⁰ PenR1 controls most probably the transcription of the outward oriented *penH* (Figure 10) similar to landomycins lanK, lanJ regulatory system.¹⁷²⁻¹⁷⁴ The gene *penH* is encoding a putative major facilitator superfamily transporter.¹⁷⁵ In turn, PenR2 shows a high degree of similarity compared to many well-studied SARP proteins from angucycline biosynthetic gene clusters, including JadR1 from *S. venezuelae*.⁶⁰ Due to the similar structure of **1** and landomycin and jadomycin,^{163, 176} we aimed to identify homologues to genes involved in the biosynthetic pathways of both aforementioned compounds within the *pen*-gene cluster. The PKS core genes *penL*, *penN*, *penO*, *penP* and *penQ* showed a significant similarity towards landomycin biosynthetic genes *lanL*, *lanA*, *lanB*, *lanC* and *lanD*, respectively. As mentioned previously, the oxygenase *penE* and both oxidoreductases *penS* and *penT* showed a remarkable homology to *jadG*, *jadF* and *jadH*, respectively.⁵⁴ Additionally, a striking similarity between

individual genes including *penA*, *penB*, *penC*, *penD* and *penH* and overall organization of the entire *pen*-gene cluster and the putative biosynthetic gene cluster with unknown product from *Micromonospora* sp. LB39 was observed (Figure S 56).

We aimed to confirm the predicted borders (*penA* and *penR2*) of pentangumycins biosynthetic gene cluster. Therefore, two BACs (1E5_DEL_LEFT and 1E5_DEL_RIGHT) were designed. BAC 1E5_DEL_LEFT was constructed through deletion of 44.1 kbp covering the left flanking region of the *pen*-cluster upstream from the gene *penA* (Figure 10a), including the entire terpene cluster. The construction of 1E5_DEL_RIGHT was carried out in a similar way by deleting a 36.3 kbp large region downstream to the *penR2* gene, including the entire predicted PKS-NRPS gene cluster. The constructed BACs were conjugated into *S. lividans* ΔYA6, and the production of **1** was analyzed. The recombinant strains carrying the mutated BACs still produced **1**. On the other hand, deletions of the terminal genes *penA* and *penR2* caused complete cessation of pentangumycins biosynthesis in the recombinant strains carrying corresponding BAC clones 1E5Δ*penA* and 1E5Δ*penR2*. The metabolic profiles of *S. lividans* trans-conjugants showed that both genes are required for production of **1**. These findings confirm that all genes necessary for production of **1** are within the predicted borders of *pen*-gene cluster.

2.3.6. Origin of SEK90

From the structural perspective **2** is a dimer of SEK43 linked by a butyl group. It was shown by McDaniel that SEK43 is formed by spontaneous cyclisation of the polyketide chain in mutants lacking the cyclase.⁶² Since then, SEK43 has been reported as a shunt product in several studies of aromatic polyketide cyclases/aromatases.¹⁷⁷⁻¹⁷⁹ At the same time SEK43 was found in the extract of natural strains only in the case of aranciamycin producing *Streptomyces* sp. Tü6384.¹⁸⁰ On the other hand, its derivative named SEK43F, resulting from fusion of SEK43 with the pyrole-like moiety, was isolated from the recombinant *S. albus* expressing fluostatins biosynthetic gene cluster.¹⁸¹ Nevertheless, a dimerization of this compound has not been described thus far. SEK43 has the 4-hydroxypyrrone structural motif. The reactivity of 4-hydroxypyrrones towards saturated aldehydes was reported previously.¹⁸² It can be hypothesized that the dimerization towards SEK90 can occur spontaneously inside the cell after the biosynthetic formation of SEK43 (Figure 11). To verify this theory, an

experimental setup was designed. 6-Benzyl-4-hydroxy-2-pyrone that possesses the same structural motif as SEK43 was incubated with butanal or formaldehyde in distilled water, production medium inoculated with *S. lividans* Δ YA6 and production medium inoculated with *S. lividans* Δ YA6_1E5. If the dimerization of the pyrone with the corresponding aldehyde is performed by an enzymatic reaction, the expected dimers (1E5_CMP1 – 1E5_CMP2, see Figure S 51) should only be present in the extracts of *S. lividans* Δ YA6 or *S. lividans* Δ YA6_1E5, depending on the presence of the necessary enzymes in the heterologous host or the introduced BAC. If the reaction towards **2** is spontaneous, the dimerization of pyrone and aldehydes should occur in the distilled water mixture. After incubation, all reactions were extracted and analyzed with LC-HRMS. Peaks that correspond to dimerized pyrones were present in all mixtures including the distilled water based one, confirming our hypothesis about spontaneous nature of the 4-hydroxy-pyrone dimerization. Furthermore, we have synthesized 1E5_CMP1 and 1E5_CMP2 to verify their structure by NMR (Table S 1 and Table S 2; Figure S 57 - Figure S 70).¹⁸² Peaks observed in distilled water mixtures had

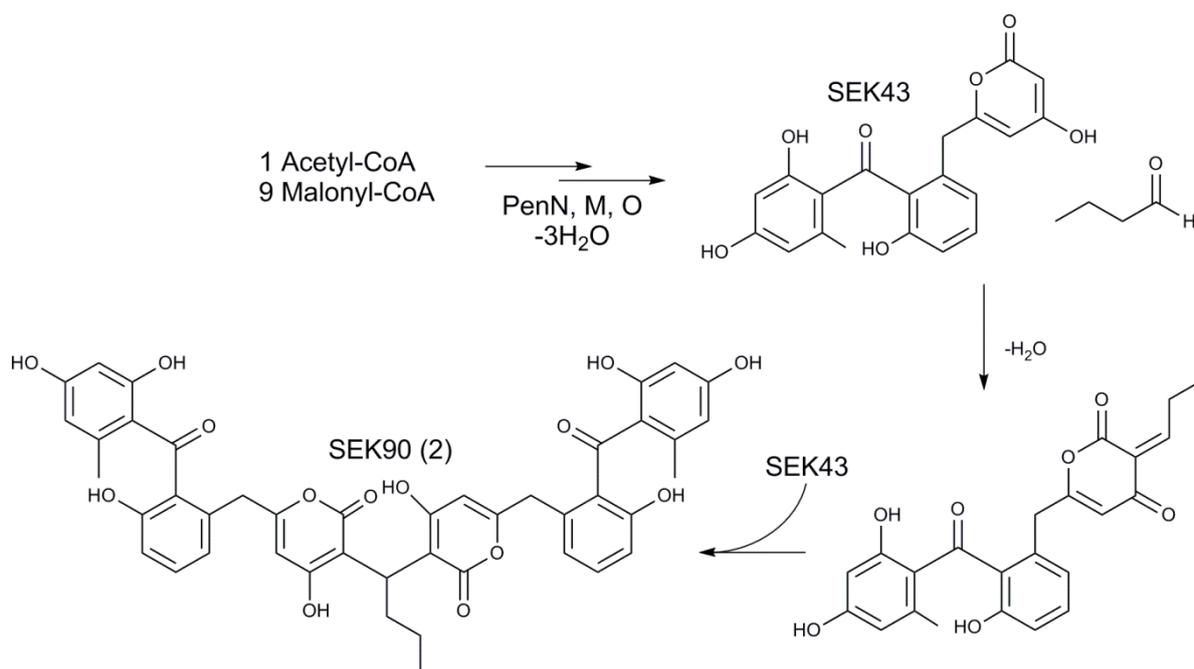


Figure 11: Origin of SEK90.

identical exact masses and HPLC retention times with the synthesized standards (Figures S71 and S72). The detailed analysis of the *S. lividans* Δ YA6 1E5 extract has revealed the presence of SEK43 (detected m/z 396.09679 $[\text{M}+\text{H}]^+$, $\Delta 0.242$ ppm to SEK43 with the calculated exact mass of 368.0860 Da) (Figure S 51 and Figure S 73) and a derivative of SEK90 connected by a

methyl- instead of a butyl-group named as SEK87 (detected m/z 749.1859 $[M+H]^+$, $\Delta 0.777$ ppm to SEK87 with the calculated exact mass of 748.1792 Da) (Figure S 51 and Figure S 74). None of the three compounds can be found in the extract of *S. espanaensis* (data not shown). Thus, it is obvious, that SEK90 as well as SEK87 derive from the combination of an unbalanced performance of pentangumycins minimal PKS, which seems to be mis-coordinated with the activity of the Pen cyclases, leading to accumulation of SEK43, and the primary metabolism of *S. lividans*. A similar situation was proposed for the assembly of SEK43F, which resulted from interplay between the primary metabolism of the host strain *S. albus* and the heterologously expressed fluostatins type II PKS gene cluster.¹⁸¹

2.3.7. Biosynthesis of pentangumycin's core structure

Based on the similarity of the *pen* genes to the genes involved in jadomycins and landomycins biosynthesis, a biosynthetic route towards **1** was proposed (Figure 12).¹⁸³⁻¹⁸⁴ The minimal PKS PenM, PenN and PenO utilize acetyl-CoA and malonyl-CoA to form the initial 20 carbon polyketide chain, which is reduced at position C-9 by the ketoreductase PenP and cyclized by the cyclases PenL and PenQ. Like in the case of all other angucycline type aromatic polyketides, the biosynthesis of **1** seems to proceed through the common known intermediate UWM6.¹⁷⁶ UWM6 is oxidized by the oxidoreductases PenE, PenS (analog of LanE)¹⁸⁴ and PenT, which leads to the formation of **I**. It is hypothesized that the oxidation of ring B is catalyzed by enzymes PenS, PenT and PenE, which are highly similar to JadG, JadF and JanH respectively, from the jadomycins biosynthesis.⁵⁴ The cleavage of ring B undergoes a Baeyer-Villiger-oxidation through the putative intermediates **II** and **III**. The latter one is proposed to have an aldehyde functional group at position C-5 and a carboxyl group at position C-6. Similar to *dauD* in daunomycins biosynthesis, the ester cyclase PenF could perform a Knoevenagel condensation, which results in the formation of the C-C bond between the β -position of 4-hydroxyl-phenylpyruvic acid, an intermediate of the tyrosine degradation catalyzed by an aromatic amino acid transaminase,¹⁸⁵ and the aldehyde of **III**, resulting in intermediate **IV**. An intramolecular nucleophilic reaction between the γ -hydroxyl and the α -keto group cleaves oxalic acid and thus creates a double bond between the former β and γ positions and a conjugated system. The subsequent reformation of ring B is catalyzed by the decarboxylation at position C-5 and directed through the six-membered

transition-state towards the former β -position of the 4-hydroxyl-phenylpyruvic acid, leading to intermediate **V**. The aromatase/cyclase PenQ is proposed to be responsible for the aromatization of ring B, resulting in intermediate **VI**. In order to prove that the fifth ring of **1** derives from the incorporation of 4-hydroxyl-phenylpyruvic acid, a culture of *S. lividans* Δ YA6_1E5 was supplemented with fully labeled ^{13}C - ^{15}N L-tyrosine. As result the mass shift from m/z 468 $[\text{M}+\text{H}]^+$ to m/z 475 $[\text{M}+\text{H}]^+$ was observed for the peak that corresponds to **1** (Figure S 75 - Figure S 80). The +7 mass shift clearly shows that the fifth ring of **1** derives

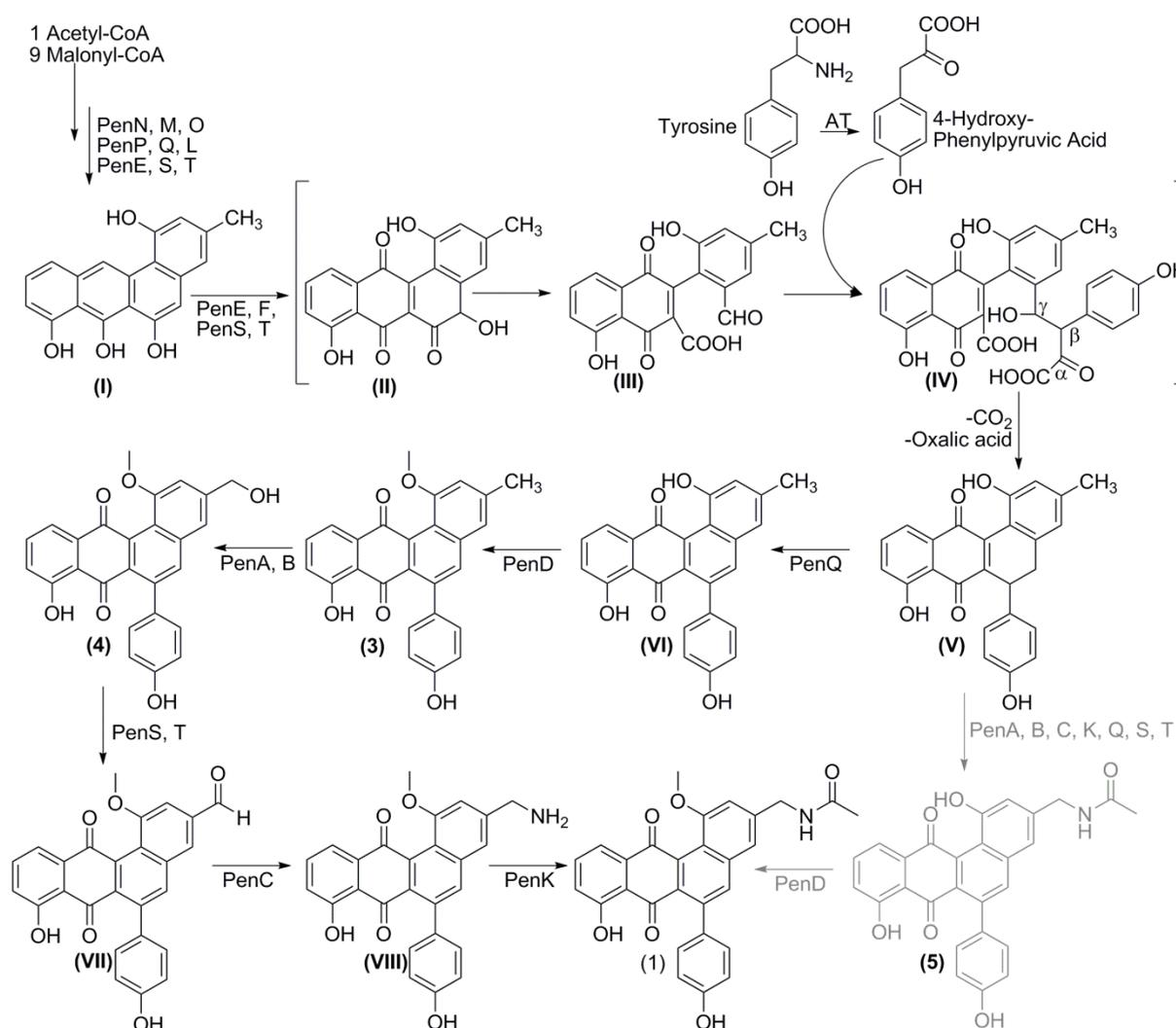


Figure 12: Biosynthetic pathway of pentangumycin; Roman numerals: proposed intermediates; Arabic numerals: intermediates with identified fitting masses after knockout; Black pathway: methylation as first biosynthetic step after aromatization; Grey pathway: methylation as last biosynthetic step to complete the formation of pentangumycin (**1**).

from incorporation of L-tyrosine as it is in the case of jadomycin and not from acetate as it was shown for formicamycin.¹⁶⁶ Furthermore, the lack of two tyrosine derived carbons in the final structure shows that the insertion of the fifth ring most probably occurs by the

proposed mechanism with the loss of α - and β - carbons of tyrosine and the exchange of carbon 6 of **IV** with one of L-tyrosine's in the re-cyclization process (Figure 12).

2.3.8. Tailoring steps of pentangumycin's biosynthesis

To elucidate the role of remaining genes encoding the putative tailoring enzymes involved in post PKS modifications of **1**, a set of single knockout mutant BACs (1E5 Δ *penA*; 1E5 Δ *penC*; 1E5 Δ *penD*; 1E5 Δ *penE*, 1E5 Δ *penG*; 1E5 Δ *penI*, 1E5 Δ *penJ*, 1E5 Δ *penU*, 1E5 Δ *penV*) was created and expressed in *S. lividans* Δ YA6. The production of **1** was not affected by deletions of *penG*, *penI*, *penJ*, *penU* and *penV* genes from BAC 1E5. At the same time, recombinant strains carrying 1E5 with deleted *penA*, *penC*, *penD* and *penE* genes lack the production of **1**, but were found (except *penE* mutant) to accumulate related angucyclinone compounds based on optical absorption and mass-spectrometry data (Figure 13). The inactivation of *penA*, coding for a cytochrome P450, resulted in the production of a new compound with m/z 411.1218 [M+H]⁺ that is very close to the calculated exact mass of 410.1154 Da of proposed intermediate **3** (Δ 2.190 ppm) (Figure S 81). The deletion of *penC*, coding for an aminotransferase, led to accumulation of the proposed intermediate **4** (detected m/z 427.1174 [M+H]⁺, calculated mass of 426.1103, Δ 0.409 ppm) (Figure S 82). The BAC lacking *penD* gene, coding for a methyltransferase, facilitated the production of another compound with m/z 454.1278 [M+H]⁺ that corresponds to the mass of the proposed intermediate **5** (calculated exact mass of 453.1212, Δ 1.506 ppm) (Figure S 83). With the information gained by the analysis of metabolic profiles of recombinant strains carrying 1E5 containing deletion of aforementioned genes, it can be hypothesized that the cytochrome P450 (PenA) catalyzes a hydroxylation of the methyl group at position C-19 to form a primary alcohol **4** which is subsequently oxidized by one of the oxidoreductases to an aldehyde giving predicted intermediate **VII**, similar how it is proposed for borrelidin biosynthetic pathway.¹⁸⁶ The aminotransferase (PenC) performs an amination of the aldehyde at this position, resulting in structure **VIII**. However, in the extract of *S. lividans* Δ YA6_1E5 Δ *penC* only proposed intermediate **4** can be found, most probably due to the fact that the aldehyde of **VII** is not stable and is spontaneously reduced to an alcohol. Lastly, the deletion of the methyltransferase gene *penD* leads to accumulation of **5**, the last intermediate before formation of **1** lacking methylation of the hydroxyl group at position C-1 (Figure 12 and

Figure 13). The presence of this intermediate, as well as methylated **3** and **4**, makes us believe that PenD can act on both substrates **5** and **VI** as well as that the presence of a methyl group at this position is not crucial for amination and acylation events.

The deletion of *penE* leads to a complete loss of production of **1**. Interestingly, the shunt product rabelomycin⁶³ can be identified in all our extracts obtained from knockout mutants (identified with an external standard), but its production is drastically reduced in the extracts of *S. lividans* ΔYA6 1E5Δ*penE* (Figure S 84). This result indicates an involvement of *penE* in the early steps of the biosynthesis of **1**.

2.3.9. Analysis of regulatory gene functions

To elucidate the functions of the identified regulatory genes on the biosynthesis of **1**, we constructed recombinant BACs with deletions of *penR1* and *penR2*. Both BACs were conjugated into *S. lividans* ΔYA6. The constructed recombinant strains *S. lividans* ΔYA6 1E5_Δ*penR1* and *S. lividans* ΔYA6 1E5_Δ*penR2* were analyzed for production of **1** (Figure S 85). While **1** was still detectable in the Δ*penR1* mutant with a 50-fold decrease in yield, it was no longer present in the *S. lividans* ΔYA6 1E5_Δ*penR2* extracts. From this observation and high similarity of *penR2* to other SARP encoding genes it becomes obvious that this gene is encoding a pathway specific regulator controlling the expression of structural *pen*-genes. Furthermore, the overexpression of either *penR1* or *penR2* might lead to the increase in the production of **1**. To test this idea both genes were cloned under the control of the moderately strong A3 promoter¹⁸⁷ and introduced into *S. lividans* ΔYA6 using a markerless vector system.¹⁸⁸ The resulting strains *S. lividans* ΔYA6_A3_ *penR1* and *S. lividans* ΔYA6_A3_ *penR2* were tested for the production of **1** and **2**. Surprisingly, the production level of **1** remained unchanged in both strains (Figure S 86). *S. lividans* ΔYA6_A3_ *penR1* 1E5 still produced **2**, but **2** was absent in extracts of *S. lividans* ΔYA6_A3_ *penR2* 1E5 (Figure S 87). This leads to the conclusion that the unbalanced expression of pentangumycins BGC is indeed the reason for accumulation of SEK90 and its derivatives by the heterologous strain carrying BAC 1E5. Furthermore, two novel peaks **6** and **7** were detected in the extract of *S. lividans* ΔYA6_A3_ *penR1* 1E5 (Figure S 88). Based on LC-HRMS data **6** could be proposed to be an analogue of **1** with tryptophan incorporated into ring B instead of tyrosine (detected *m/z* 491.15934 [M+H]⁺, calculated exact mass of 490.1529, Δ1.646), while **7** seems

to carry phenylalanine derived moiety at the same position (detected m/z 452.14903 $[M+H]^+$, calculated exact mass 451.1419, $\Delta 0.485$ ppm) (Figure S 88). To prove the possible nature of the identified derivatives **6** and **7** the culture of *S. lividans* Δ YA6_A3_penR1 1E5 was supplemented with either L-tryptophan $^{13}C_{11}^{15}N_2$ or L-phenylalanine (ring-D₅). As result, in the correspondingly fed cultures the mass of **6** shifted from m/z 491.16 $[M+H]^+$ to m/z 500.19 $[M+H]^+$ (m/z +9) (Figure S 89), and the mass of **7** shifted from m/z 452.16 $[M+H]^+$ to m/z 457.20 $[M+H]^+$ (m/z +5) (Figure S 90), that corresponds to incorporation of labelled L-tryptophan and L-phenylalanine.

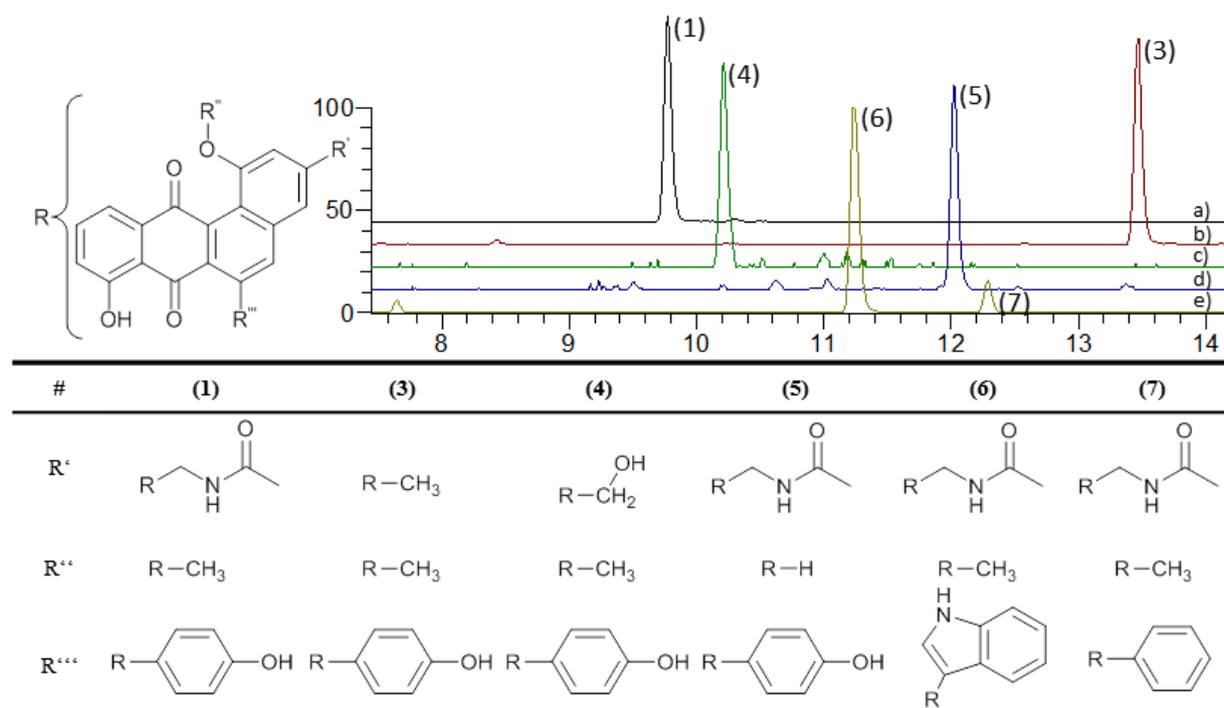


Figure 13: Produced derivatives through knockouts and overexpression of pentangumycin; a) *S. lividans* YAΔ6_1E5 (Extracted mass: 468); b) *S. lividans* YAΔ6_1E5ΔpenA (Extracted mass: 411); c) *S. lividans* YAΔ6_1E5ΔpenC (Extracted mass: 427); d) *S. lividans* YAΔ6_1E5ΔpenD (Extracted mass: 411); *S. lividans* YAΔ6_1E5ΔpenA (Extracted mass: 454); d) *S. lividans* YAΔ6_A3_R1_1E5 (Extracted masses: 452 & 491).

2.4. Conclusion

In conclusion, we have shown that our approach based on systematic analysis of genome of an Actinobacteria strain for unique BGCs combined with heterologous expression is a promising source of new natural products with unusual structures. As result, two novel compounds were isolated in sufficient amounts for structure elucidation and testing of their biological activity. Through genetic manipulation and feeding experiments the biosynthesis of **1** was elucidated. SEK90 represents a new shunt product in biosynthesis of aromatic polyketides that arises from interplay between expressed BGC and primary metabolism of the host strain. **2** cannot be found in *S. albus* J1074. At the same time, pentangumycin, the actual final product of the *pen*-gene cluster is a new member of the angucyclinone family with a heavily modified ring system and a new unusual tailoring modification, which was not observed in other members of this group of natural products before.

Importantly, we have demonstrated that *Streptomyces* species can be used as hosts for heterologous expression of biosynthetic gene clusters derived from distantly related Actinobacteria species, even if their expression is silent in the wild-type strain. However, still relatively low success rate should trigger the development of other than *Streptomyces* host strains for actinobacterial secondary metabolism gene cluster expression.

2.5. Material and Methods

2.5.1. General Biological procedures

All oligonucleotides, plasmids and strains used in this work can be found in Table S 3 and Table S 4. Standard procedures for *Escherichia coli* (transformation, plasmid preparations and BAC isolations) were performed as described by Sambrook.¹⁸⁹ Intergeneric conjugation between *Streptomyces* and *E. coli* was performed as described by Kieser¹⁹⁰ on MS-Agar plates using *E. coli* PUB307 as donor strain. Standard DNA manipulations (ligation, polymerase chain reactions and endonuclease digestion) were performed according to the corresponding manufacturer's protocol. All *Streptomyces* were cultivated in 15 mL TSB (17 g Tryptone, 3 g Peptone, 5 g NaCl, 2.5 g K₂HPO₄, 2.5 g glucose ad 1000ml H₂O_{dest}; pH=7.2) in a 100 mL flask (4 baffles; 5 g glass beads), inoculated from spores of an MS-Plate, until a dense culture was obtained. 50 mL of the corresponding production medium was inoculated with 1 mL of the seed culture in a 500 mL flask (3 baffles, 10 g glass beads) and incubated for 72-164h at 180 RPM and 29 °C. Production optimization steps were carried out by adjustment of different parameters (Growth time, pH, Medium, Extracting solvent, addition of XAD polymers)

2.5.2. General analytical procedures

All HPLC-MS spectra were recorded on a Dionex Ultimate 3000 (Thermo Fisher Scientific Waltham, Massachusetts, USA) coupled with an AmaZon ETD SL speed, Apollo II ESI (Bruker, Billerica, Massachusetts, USA) on a Waters BEH C18 column (100 mm x 2.1 mm, 1.7 µm) with a 18 minute linear gradient from 5% acetonitrile (0.1 % formic acid) to 95% acetonitrile (0.1 % formic acid). The mass spectra were recorded in centroid mode (200 to 2000 m/z) at a scan rate of 2 Hz. Prior to analysis, all samples were dissolved in Methanol and centrifuged for 10 minutes at 4 °C and 15.000 rpm. The data was analysed using Bruker Compass Data Analysis 4.2. High resolution masses were recorded on a Dionex Ultimate 3000 RSLC HPLC (Thermo Fisher Scientific) coupled with an LTQ Orbitrap (Thermo Fisher Scientific) on a BEH C18 100 x 2.1 mm, 1.7 µm column (Waters) with a 18 minute linear gradient from 5% acetonitrile (0.1 % formic acid) to 95% acetonitrile (0.1 % formic acid). Data analysis was carried out using Xcalibur 3.0. Preparative HPLC was carried out using a Waters

AutopurificationSystem equipped with a Waters 2545 Binary Gradient module, Waters SFO (System Fluidics organizer), Waters 2998 PAD (Photodiode array Detector) Waters 2767 Sample Manager and Waters SQ-Detector-2 on Nucleodur C18 Htec 250/21-C18-5 μ m column (Macherey-Nagel Düren, Germany). Data analysis was carried out using MassLynx. Structural elucidation was carried out using the Nuclear Magnetic Resonance (NMR) technology. The purified compounds were solved in 300 μ L deuterated solvent (DMSO- d_6 , MEOD- d_4 , $CDCl_3$) and measured in a corresponding 5 mm Shigemi-tube (DEUTERO GMBH; Kastellaun, Germany). NMR data (1H , HH-COSY; TOCSY; HMBC, HSQC, ^{13}C) were acquired either on a Bruker Ascend 700 spectrometer equipped with a 5 mm TXA Cryoprobe or a Bruker Avance 500 spectrometer equipped with a 5mm BBO Probe at 300K (Bruker, BioSpin, GmbH, Rheinstetten, Germany). The data was analyzed using Brukers TopSpin 3.5a software

2.5.3. In-silico analysis of gene clusters

The genome of *Saccharothrix espanaensis* was analyzed with antiSMASH¹⁰⁸ and the Geneious software. Identified BGCs were mapped to the BAC library. A detailed analysis of 1E5's BGC was performed with BLAST.¹⁹¹

2.5.4. Extraction, Dereplication and Isolation

The biomass was separated from the production medium (SG-Medium [20g glucose, 5g yeast extract, 10g bactosoytone, 2g $CaCO_3$ ad 1000ml H_2O_{dest} ; pH=7.2], 172 h, 29 °C, 180 rpm, pH=7.2, 1 % XAD-16 after 24h) by centrifugation. 20 mL of the supernatant was extracted with 20 mL ethyl acetate followed by 20 mL butanol for 20 minutes using a Laboshake LS500 (C. Gerhardt GmbH, Königswinter, Germany) at 160 rpm. The biomass was extracted with a 1:1 mixture of acetone and methanol for 60 minutes. The solvents were evaporated to dryness with a rotary evaporator (150 rpm, 60 °C, 240 mBar for ethylacetate or 25 mBar for butanol) or under a nitrogen stream at 40 °C. The extracts were analyzed using LC-MS and LC-HRMS. Singular peaks were compared to the databases Dictionary of Natural Products, version 27.1 (CRC Press, Boca Raton, FL, United States), Sci-Finder (CAS, Columbus, USA) and Chemspider (Royal Society of Chemistry, Raleigh, USA). For production, 15 Liter of SG-medium inoculated with the corresponding recombinant strain was cultivated. The extracts were evaporated to dryness and dissolved in 10 mL methanol. To purify single

compounds, a size exclusion chromatography (SEC) was performed. The column was packed with 600 mL Sephadex LH-20 resin (GE Healthcare Europe GmbH, 79111 Freiburg, Germany) solved in Methanol. At a flow rate of 10 mL per 15 min. a fraction collector was used to separate the extract within 24 hours. To identify fractions containing the novel substances every 4th fraction was analyzed by HPLC-MS. Fractions with the corresponding mass were combined and evaporated. Prior to NMR analysis a preparative HPLC was performed. Preparative HPLC information can be found in Table S 5 and Table S 6.

pentangumycin (**1**): ¹H-NMR (DMSO-d₆ 700 MHz) δ_H 7.10 (1H, s, H-2), 7.44 (1H, s, H-4), 7.88 (1H, s, H-5), 7.26^a (1H, d, H-9), 7.73 (1H, t, J=7.5 Hz, H-10), 7.50 (1H, dd, J=7.5, 0.93 Hz, H-11), 7.26^a (1H, d, H14/18), 6.78 (1H, dt, J=8.5, 2 Hz, H-15/17), 4.43 (2H, d, J=6 Hz, H-19), 1.93 (3H, s, H-21), 3.88 (3H, s, H-22), 8.53 (1H, t br, NH); ¹³C-NMR (DMSO-d₆ 175MHz) δ_C 156.72 (C, C-1), 108.73 (CH, C-2), 143.23 (C, C-3), 117.52 (CH, C-4), 136.62 (C, C-4a), 135.12 (CH, C-5), 139.11 (C, C-6), 131.27 (C, C-6a), 188.02 (C, C-7), 115.80 (C, C-7a), 160.52 (C, C-8), 122.98 (CH, C-9), 136.66 (CH, C-10), 117.20 (CH, C-11), 135.51 (C, C-11a), 185.53 (C, C-12), 139.97 (C, C-12a), 119.48 (C, C12b), 132.18 (C, C-13), 129.94 (CH, C-14/18), 114.77 (CH, C-15/17), 156.58 (C, C-16), 42.37 (CH₂, C-19), 169.55 (C, C-20), 22.75 (CH₃, C-21), 56.12 (OCH₃, C22); HRESMS: *m/z* 468.1438 [M+H]⁺ (calcd for C₂₈H₂₂NO₆⁺ 468.1442)

SEK90 (**2**): ¹H-NMR (DMSO-d₆ 500 MHz) δ_H 5.68 (1H, s br, H-4, H-4'), 3.55 (1H, s br, H-6, H-6'), 6.74 (1H, dd, J=7.2, 0.8 Hz H-8, H-8'), 7.2 (1H, dd, J=7.9 Hz, H-9, H-9'), 6.77 (1H, dd, 7.5 Hz, H-10, H-10'), 6.11 (1H, d, J=2.4 Hz, H-16, H-16'), 6.04 (1H, dd, J=2.4 Hz, 0.6Hz, H-18, H-18'), 1.8 (3H, s, H-20; H-20'), 4.36 (1H, t, J=8.2Hz, H-21), 1.75 (2H, dt, J=7.3, 8.2, H-22), 1.05 (2H, dq, J=7.3Hz, H-23), 0.79 (3H, t, J=7.3, H-24), 12.67 (1H, s, OH-1, OH-1'), 11.49 (1H, s br, OH-2, OH-2'), 10.38 (1H, s, OH-3, OH-3'), 9.78 (1H, s, OH-4, OH-4'). ¹³C-NMR (DMSO-d₆ 175MHz) δ_C 166.5 (C, C-1; C-1'), 102.7 (C, C-2; C-2'), 165.5 (C, C-3; C-3'), 101.9 (CH, C-4; C-4'), 160.6 (C, C-5; C-5'), 36.1 (CH₂, C-6; C-6'), 132.5 (C, C-7; C-7'), 120.1 (CH, C-8; C-8'), 130.9 (CH, C-9; C-9'), 114.4 (CH, C-10; C-10') 153.3 (C, C-11; C-11'), 130.6 (C, C-12; C-12'), 199.9 (C, C-13; C-13'), 115.3 (C, C-14; C-14'), 164.9 (C, C-15; C-15'), 100.5 (CH, C-16; C-16'), 166.3 (C, C-17; C-17'), 111.4 (CH, C-18; C-18'), 142.7 (C, C-19; C-19'), 21.2 (CH₃, C-20; C-20'), 29.5 (CH, C-21; C-21'), 31.5 (CH₂, C-22; C-22') 20.6 (CH₂, C-23; C-23'), 13.8 (CH₃, C-24; C-24'); HRESMS: *m/z* 791.2339 [M+H]⁺ (calcd for C₄₄H₃₉O₁₄⁺, 791.2334)

2.5.5. Modification of BACs

All constructed single knockout BACs were obtained by homologous recombination as previously described by Gust¹⁵¹ and Myronovskiy¹⁹² and confirmed by PCR. The recombinant BACs were transformed in *E. coli* PUB307 and conjugated into *S. lividans* ΔYA6. All recombinant strains were cultivated, and their metabolome was analyzed as described above.

2.5.6. Construction of Plasmids

For the construction of the plasmids pTOS_A3_R1 and pTOS_A3_R2 we amplified the genes using PCR, while introducing the sequence of the A3-Promotor¹⁸⁷ and a *KPN*I restriction site in the 3' oligonucleotide and a *Hind*III restriction site in the 5' oligonucleotide. The amplified PCR product and the integrative pTOS plasmid were cut using *Hind*III and *KPN*I and ligated by the T4 ligase. The obtained DNA fragment was transformed into *E. coli* PUB307 and conjugated in *S. lividans* ΔYA6. The backbone of pTOS was removed from the genome by expression of pUWL_DRE containing the DRE recombinase.

2.5.7. Synthesis of 1E5_CMP1-2

6-benzyl-4-Hydroxy-2-pyrone (Wuxi AppTec, Saint Paul, USA) was used for the synthesis of 1E5CMP1- 2 using the conditions described by de March.¹⁸² The pyrone was mixed (2:1) with either butyraldehyde or formaldehyde in 3 mL Ethanol containing 10μL Piperidine and 10μL glacial acetic acid. The reaction was performed at room temperature for 24 h. All compounds were purified using Waters Autopurification System and confirmed by NMR.

2.5.8. Feeding experiments

To prove the incorporation from L-Phenylalanine, L-Tyrosine and L-Tryptophan, a culture of either *S. lividans* ΔYA6_1E5 or *S. lividans* ΔYA6_1E5_A3_R1 was prepared as described in 2.5.1. Thereafter, 10 mg of the corresponding labelled amino acid were supplemented in 5 steps, 2 mg after 24 h, 48 h, 72 h, 96 h and 120 h. After 164 h of incubation, the supernatant of the production cultures was extracted, and the incorporation was quantified with LC-MS.

3. Tuning the production of pamamycins through engineering of the precursors pool in the heterologous host *S. albus* J1074

3.1. Abstract

Pamamycins, a group of polyketides originally discovered in *Streptomyces alboniger*, can induce sporulation in *Streptomyces* and inhibit the growth of Gram-positive bacteria, *Mycobacterium tuberculosis* and fungi. Pamamycins biosynthetic gene cluster encodes for 6 ketosynthases. Three of them (PamA, PamE and PamG) can utilize a variety of extender units such as malonyl-CoA, 2-S-methylmalonyl-CoA and 2-S-ethylmalonyl-CoA. PamA utilizes succinyl-CoA as starter unit and either malonyl-CoA or 2-S-methylmalonyl-CoA as extender units. PamE and PamG incorporate malonyl-CoA, 2-S-methylmalonyl-CoA or 2-S-ethylmalonyl-CoA as building blocks in the growing polyketide chain of pamamycin and thus, further increase the chemical variety. The utilization of different extender units results in a diversity of produced pamamycins, covering at least 18 different derivatives with molecular weights ranging from 579 g/mol up to at least 649 g/mol. For the isolation of pamamycins, we aimed to simplify the downstream processing of pamamycins through modification of the precursor supply, which results in an altered spectrum of produced derivatives. Eight core genes responsible for the supply of 2-S-methylmalonyl-CoA and 2-S-ethylmalonyl-CoA were identified using NCBI's Basic Local Alignment Search Tool (BLAST) in the heterologous host *S. albus* J1074. By applying the recombineering technology, we were able to construct knockout mutants of the described genes, to measure their corresponding CoA-ester concentration and to analyze the influence on the production of pamamycins.

3.2. Introduction

Polyketides form one of the most important groups of natural products known to mankind. They are produced by bacteria, fungi and plants and their chemical diversity is reflected in various clinical applications, including antibiotics,¹⁹³ antihelmintic drugs,¹⁹⁴ immunosuppressants,¹⁹⁵ antifungals,¹⁹⁶ cholesterol-lowering drugs¹⁹⁷ and cytotoxic agents.¹⁹⁸ Polyketides are assembled by simple Claisen condensations of malonate-derived building blocks.²⁸ The formation of primary poly- β -ketide chain is performed by an enzyme complex called polyketide synthase (PKS) in a process similar to the biosynthesis of fatty acids.²⁵ In a view of such simplicity of basic assembly principles, the enormous variety of polyketides is achieved in two steps. Firstly, the use of various acyl precursors as starter (acetyl-CoA, propionyl-CoA, isovaleryl-CoA, benzoyl-CoA, etc) and extender building blocks (methylmalonyl-CoA, ethylmalonyl-CoA, chloroethylmalonyl-CoA, methoxymalonyl-CoA, etc)⁷, together with a changing degree of ketoreduction leads to differences in the nascent scaffold.¹⁹⁹ Secondly, intensive post-PKS modifications such as glycosylation, amination and methylation strongly alter the basic polyketide structure.²⁰⁰

The majority of polyketides' precursors originates from the primary metabolism.⁷ The most commonly utilized extender unit, malonyl-CoA, is produced via two different pathways. The predominant way used by bacteria is the direct carboxylation of acetyl-CoA by the acetyl-CoA carboxylase.⁷⁷ Acetyl-CoA in turn is supplied from multiple catabolic pathways, including the β -oxidation of fatty acids. Additionally, malonyl-CoA can also be formed by an activation of malonate with CoA by the malonyl-CoA synthetase MatB.²⁰¹ Methylmalonyl-CoA in bacteria was shown to originate mainly from two reactions: carboxylation of propionyl-CoA by the propionyl-CoA carboxylase (PCC) and the reversible isomerization of succinyl-CoA by methylmalonyl-CoA mutase (Mcm).⁹⁸ Lastly, ethylmalonyl-CoA is generated by carboxylation of crotonyl-CoA by the crotonyl-CoA carboxylase/reductase (Ccr). The deep understanding of central metabolism routes, supplying these biosynthetic precursors, allows to manipulate the biosynthesis of polyketide antibiotics not only in terms of production yield but also to selectively modulate the particular derivatives accumulation. The prominent example of the latter case is the changing of the monensins spectra by affecting ethylmalonyl-CoA supply.²⁰² However, such examples are rather rare most probably due to the lack of a clear

understanding of interplays between primary and secondary metabolism. First of all, this is caused by the great redundancy and strong interdependence of the aforementioned central metabolism pathways that complicate the analysis and render the understanding of interactions between primary and secondary metabolism extremely difficult. The deep investigation of metabolic pathways for main polyketide precursors in commonly used heterologous strain will broaden the possibilities of metabolic engineering of these secondary metabolites production in Actinobacteria.

Pamamycins are a group of macrodiolide polyketides originally found in the extract of *Streptomyces alboniger*.⁷¹ They were shown to have an activity against Gram-positive bacteria, including *Mycobacterium tuberculosis* and fungi.⁷¹ Pamamycins are produced as a mixture of closely related derivatives with molecular weight ranging from 579 to 649 Da,

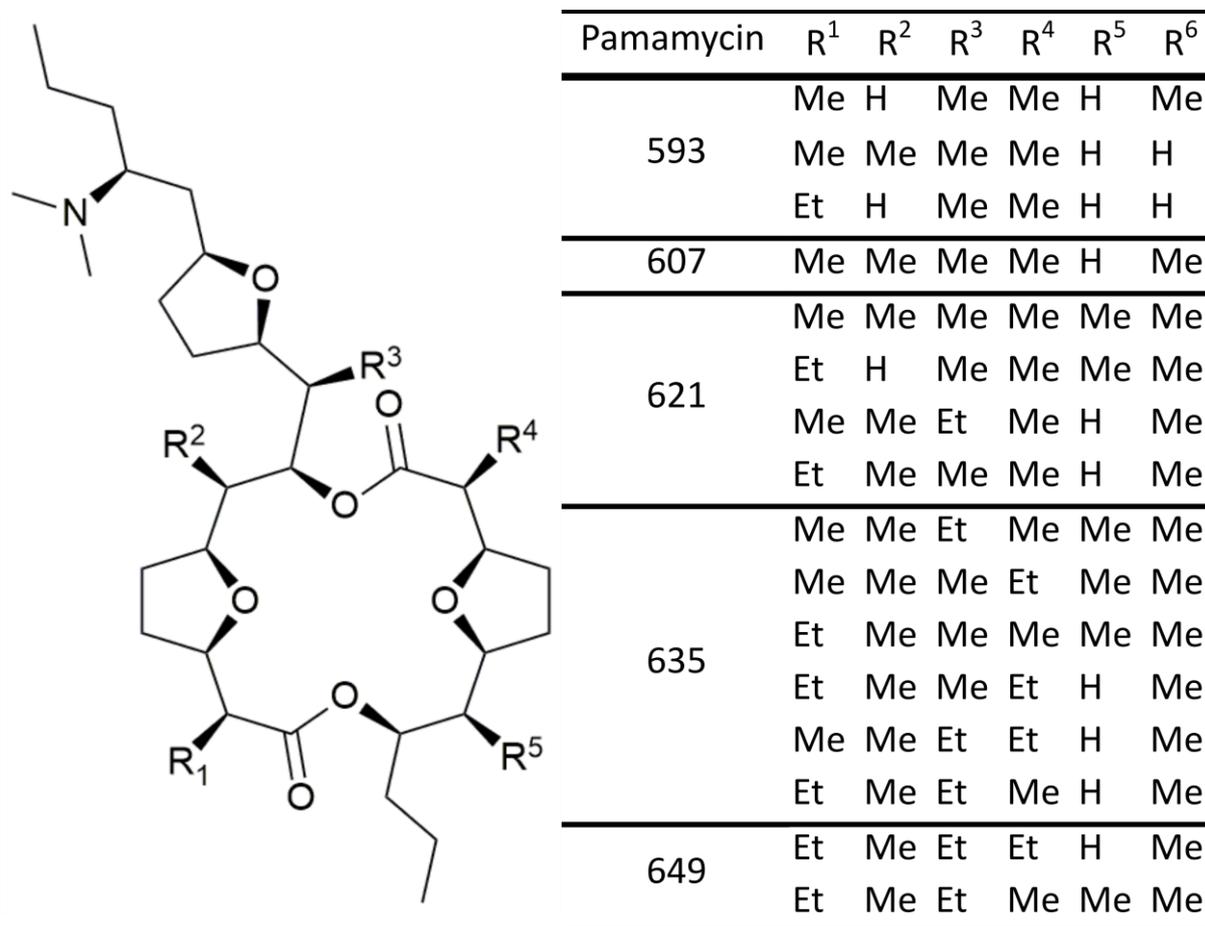


Figure 14 The core structure of Pamamycins, it's derivatives and their corresponding molecular masses

which differ by the side chain substituents in six different positions (Figure 14).⁶⁵ Such structural diversity is introduced during the primary polyketide chain assembly by unusual

PKS enzymes that equally accept malonyl-CoA, methylmalonyl-CoA and ethylmalonyl-CoA as extender units.⁷⁶ This makes the biosynthesis of pamamycins a great model to study the interactions between different polyketides precursor supply pathways. On the other hand, it was noted that low molecular weight pamamycins have a stronger antimycobacterial activity.⁶⁶ However, due to the high complexity of the produced mixture the isolation of a sufficient quantity of any specific derivative for further structure-activity studies is virtually impossible. To overcome this problem and to pursue the development of these highly promising compounds, we aimed to modify the produced spectrum of pamamycins by controlling the flux of precursors into their biosynthesis in the heterologous strain *S. albus* J1074. Herein we report the elucidation of biosynthetic routes for methylmalonyl-CoA and ethylmalonyl-CoA in *S. albus* J1074 and the use of this knowledge to modulate the production of pamamycins.

3.3. Materials and Methods

3.3.1. Bacterial strains and culture

Strains and vectors used in this study are presented in Table S 7. All *E. coli* cultures were grown at 37 °C with 180 RPM in LB medium. For general procedures, *Streptomyces* cultures were grown on MS agar medium.¹⁹⁰ For DNA isolation and pre-cultures, *Streptomyces* strains were cultivated in liquid TSB medium for 24-48 h at 29 °C and 230 rpm. SGG (10 g/L starch, 10 g/L glycerol, 2.5 g/L corn steep solids, 5 g/L Peptone, 2 g/L yeast extract, 1 g/L NaCl, 3 g/L CaCO₃, dH₂O, pH=7.2) was used as main pamamycins production medium and strains were incubated for 72 h at 29°C and 230 rpm. Feeding experiments were carried out in SGG and liquid Hopwood's¹⁹⁰ minimal medium under conditions described above. When needed, media were supplemented with the following antibiotics: apramycin 50 µg/mL, hygromycin 50 µg/mL, chloramphenicol 30 µg/mL, kanamycin 100 µg/mL, nalidixic acid 30 µg/mL.

3.3.2. General procedures

E. coli transformation and plasmid DNA isolation were performed according to standard protocols.¹⁸⁹ Total DNA isolation and other procedures with *Streptomyces* strains were performed as previously described.¹⁹⁰ Intergeneric conjugation between *Streptomyces* and *E. coli* was done according to a modified protocol of Flett *et al.*²⁰³ using MS agar medium and *E. coli* WM6026 as donor strain.²⁰⁴ DNA manipulations such as endonuclease restriction, ligation and polymerase chain reaction (PCR) were performed according to manufacturers' protocols (Thermo Fischer Scientific, USA; NEB, USA). The oligonucleotides used in this work are listed in Table S 7 (Eurofins, Germany). DNA sequencing was performed at Eurofins GATC Biotech (Eurofins, Germany).

3.3.3. Construction of recombinant cosmid clones and gene deletions in *S. albus*

S. albus cosmid library clones pSMARTgus_3D17, pSMARTgus_2J19, pSMARTgus_1M11, pSMARTgus_3H24, pSMARTgus_2E5, pSMARTgus_3J4 were modified by deleting the genes of interest using Red/ET recombineering as described earlier.¹⁵¹ To facilitate markerless deletions the iterative marker excision system was used.¹⁹² The resulting cosmids 3D17 Δ ccr2 Δ meaA, 3D17 Δ meaA, 3D17 Δ ccr2, 2J19 Δ vdh, 3H24 Δ pcc3, 1M11 Δ pcc2, 2E5 Δ pcc1 and

3J4 Δ *mcm* were verified by PCR and the sequencing of the deletion site. The obtained recombinant cosmids were introduced into *S. albus* by intergeneric conjugation. The screening for double crossover event was carried out by blue-white selection as described.²⁰⁵ The IMES resistance cassette was removed after each individual deletion by expressing *phiC31 int* gene.¹⁹² All mutant strains were verified by PCR to carry the correct deletion. Strains and their genotypes are listed in Table S 8.

3.3.4. Pamamycin production and analysis

The R2 cosmid, containing pamamycins biosynthetic gene cluster, was introduced into *S. albus* strains by intergeneric conjugation. For the pamamycin production, a pre-culture was inoculated from spores into 15 mL of TSB in 100 mL flasks and incubated as described above. 1 mL of each pre-culture was inoculated into 50 mL of SGG medium in 500 mL flasks with 1 baffle containing 30 g of glass beads and incubated as described in 3.3.1. In all cases at least three independent cultures of the same strain were examined.

The biomass was harvested by centrifugation (10 min, 4500 rpm, 4°C) and extracted with 7.5 mL of a 1:1 mixture of methanol and acetone for 60 min with shaking at 160 rpm. The biomass was subsequently separated by centrifugation (10 min, 4500 rpm, 4°C) and dried under a nitrogen stream. 20 mL of supernatant were extracted with equal volume of ethyl acetate for 30 min with shaking at 160 rpm. The extracts were evaporated using a rotary evaporator (240 mbar, 60 °C, 160 rpm). Subsequently, the extract of biomass and supernatant were dissolved in 150 μ L of a 1:1 mixture of methanol and DMSO and combined.

The extracts were analysed by HPLC-MS (Dionex Ultimate 3000, Thermo Fisher Scientific USA, AmaZon ETD SL speed with Apollo II ESI source, Bruker, USA) using a Waters BEH C18 column (100 mm x 2.1 mm, 1.7 μ m, Waters Corporation, USA). The volume of injection was 1 μ L. Details of solvents and the gradients used are shown in Table S 9. The mass spectra were recorded in centroid mode (200 to 2000 m/z) at a scan rate of 2 Hz. The data was collected and analysed using Bruker Compass Data Analysis software version 4.2 (Bruker, USA). The area under curve (AUC) of a smoothed BPC chromatogram was used as a

measure to compare the amount of pamamycins produced. A standard curve was obtained with the pure sample of pamamycin 607.

3.3.5. Measurement of the CoA-Esters

The CoA-ester measurement was carried out by Wittmann *et al.* with a modified protocol described by Peyraud *et al.*²⁰⁶

3.3.6. Incorporation of labelled amino acids

To analyse the incorporation of labelled amino acids into pamamycins, the corresponding *S. albus* culture was prepared as described in Section 3.3.1. 1 mg of single amino acids (L-Valine-2-¹³C; Sigma Aldrich, USA) and 2 mg of an amino acid mixture (Sigma Aldrich, St. Louis, Missouri, USA) were supplemented to the production medium, incubated with *S. albus* $\Delta 5$ R2 and *S. albus* $\Delta 7$ R2, respectively, after 12, 24, 36, 48 and 60 h after the incubation. Pamamycins and the corresponding incorporation were monitored by HPLC-MS as described in Section 3.3.4.

3.4. Results

3.4.1. Bioinformatics identification of pamamycin precursors supply pathways and corresponding genes in the genome of *S. albus*

The pamamycins assembly line utilizes malonyl-CoA, methylmalonyl-CoA and ethylmalonyl-CoA as extension units. Since *S. alboniger* is poorly genetically tractable, the heterologous system based on *S. albus* J1074 and the R2 cosmid carrying the entire *pam* gene cluster is the preferred production model. In order to identify the metabolic pathways that are

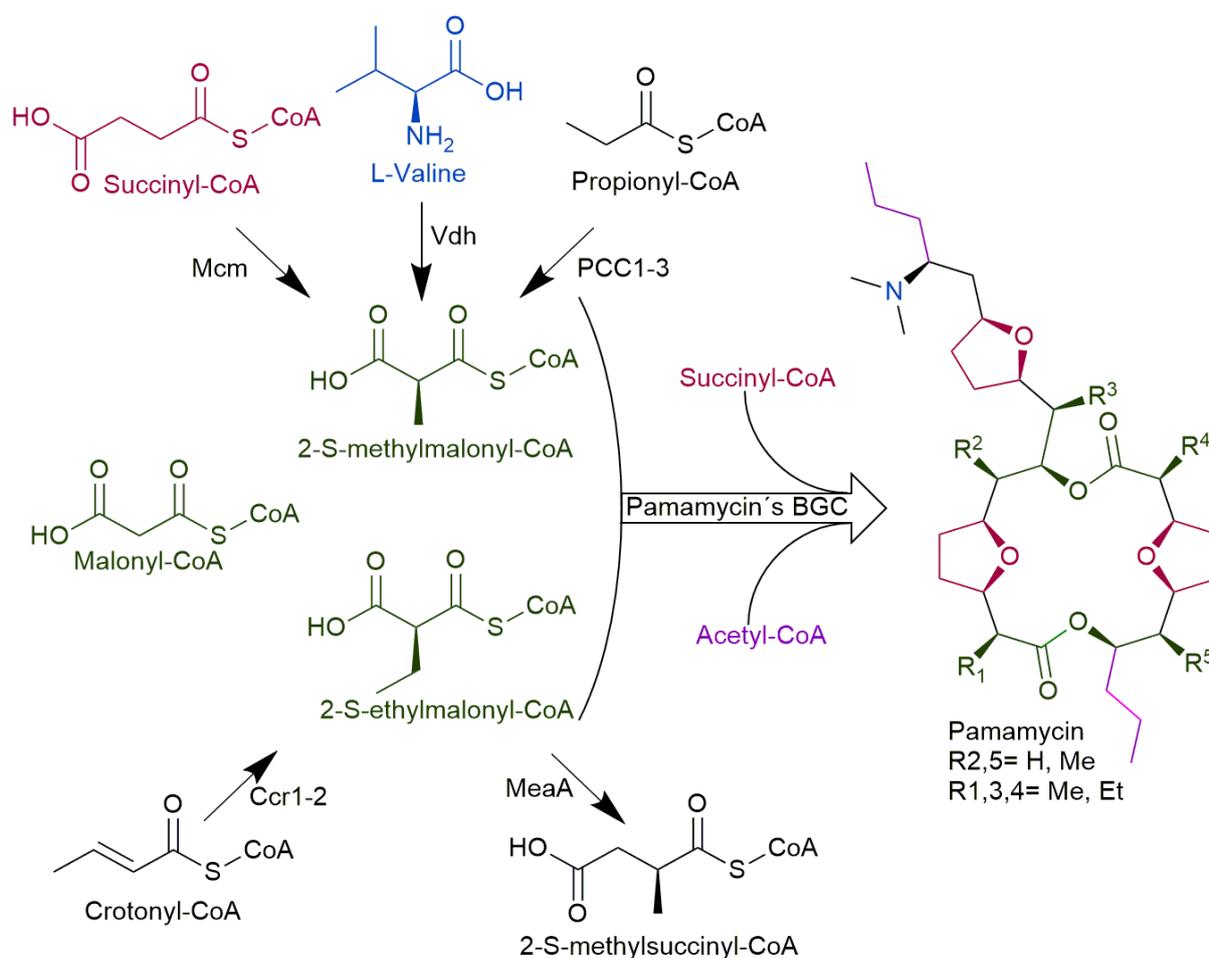


Figure 15 Biosynthetic origin of the precursors, which are involved in pamamycins biosynthesis.

involved in the supply of methylmalonyl-CoA and ethylmalonyl-CoA and genes encoding corresponding enzymes directly involved in conversion of these CoA-esters, we performed a KEGG Pathway database and KEGG Orthology analysis of the genome of *S. albus*.²⁰⁷ Indeed, the genome of *S. albus* is predicted to code for all three metabolic pathways known to employ methylmalonyl-CoA: 1 – Valine, leucine and isoleucine degradation (KEGG Pathway

ID salb00280); 2 – Glyoxylate and dicarboxylate metabolism (salb00630); and 3 – propanoate metabolism (salb00640) (Figure 15). Thus, methylmalonyl-CoA is proposed to be an intermediate in three central metabolism reactions in *S. albus*: 1 – The reversible interconversion from (R)-methylmalonyl-CoA to succinyl-CoA catalysed by methylmalonyl-CoA mutase (Mcm) and the (S)-(R) epimerization by methylmalonyl-CoA epimerase; 2 – The carboxylation of propionyl-CoA performed by propionyl-CoA carboxylase (PCC) and 3 – The oxidation by the aldehyde dehydrogenase (XNR_4007 or XNR_3418) of (S)-methylmalonate semialdehyde. Ethylmalonyl-CoA is predicted to participate in two major reactions: 1 – The reductive carboxylation of crotonyl-CoA that is catalysed by the crotonyl-CoA reductase/carboxylase (Ccr) and 2 – The conversion to (2S)-methylsuccinyl-CoA catalysed by the (2R)-ethylmalonyl-CoA mutase (MeaA).

Table 3: Identified genes involved in the biosynthesis of 2S-methylmalonyl-CoA and 2S-ethylmalonyl-CoA.

XNR	Gene	Abbr.	SCO	% of homology
XNR_5889	Crotonyl-CoA reductase	ccr1	sco6473	35%
XNR_0456	Crotonyl-CoA carboxylase/ reductase	ccr2	sco6473	93%
XNR_0457	Ethylmalonyl-CoA mutase, methylsuccinyl-CoA forming	meaA	sco6472	86%
XNR_2273	Propionyl-CoA carboxylase, alpha subunit	pcc1	sco2776	86%
XNR_2274	Propionyl-CoA carboxylase beta subunit			
XNR_4211	Propionyl-CoA carboxylase alpha subunit	pcc2	sco4380	89%
XNR_4212	Propionyl-CoA carboxylase			
XNR_4024	Propionyl-CoA carboxylase beta subunit	pcc3	sco4926	89%
XNR_4665	Methylmalonyl-CoA mutase large subunit	mcm	sco4869	50%
XNR_4666	Methylmalonyl-CoA small subunit			
XNR_2839	Valine dehydrogenase	vdh	sco4089	85%

In order to identify genes encoding corresponding enzymes, we performed a *S. albus* genome BLAST search for the orthologues from corresponding enzymes from *S. coelicolor*.¹⁹¹ In *S. coelicolor*, PCC is composed of two subunits: the α -subunit (AccA2, SCO6271), which carries the biotin carboxylase (BC) and biotin carboxyl carrier functions. This protein is also a part of the acetyl-CoA carboxylase complex. The β -subunit is performing the carboxylation reaction of specific substrate (PccB, SCO4926).⁹⁰ In the genome of *S. albus* J1074 genes

encoding three putative PCC complexes were identified: PCC1 (XNR_2273 and XNR_2274) that corresponds to PccA and PccB of *S. coelicolor* (SCO4381 and SCO4380); PCC2 (XNR_4211 and XNR_4212), orthologues of AccC and AccD1 (SCO2777 and SCO2776); PCC3 formed by XNR_4019 and XNR_4024, orthologues of AccA2 and PccB, respectively (Table 3). At the same time, we were not able to find any putative candidate genes to encode the propionyl-CoA transcarboxylase homologue of *Propionibacterium*, which is an alternative to the typical PCCs, utilizing oxaloacetate as a donor of the carboxylic group.²⁰⁸ Additionally, we were able to identify several genes putatively encoding acyl-CoA-carboxylases without defined substrate specificity, which were excluded from further work. The genome of *S. albus* is coding for three orthologues of *S. coelicolor* methylmalonyl-CoA mutase: XNR_4665 and XNR_4666 have a high degree of similarity to MutA (SCO6832) and MutA2 (SCO4869) and XNR_1417, which is annotated as Mcm and shows a 50% amino acid similarity to MutA from *S. coelicolor* (Table 3). However XNR_1417, most probably acts as isobutyryl-CoA mutase (89% aa identity to IcmA SCO5415). Interestingly, in *S. coelicolor* both genes encoding paralogues of *mcm mutA1* and *mutA2* are located distantly from each other. At the same time, in the genome of *S. albus* they are forming one operon with the gene encoding the methylmalonyl-CoA mutase-associated GTPase MeaB (SCO5400 orthologue). Furthermore, they are clustered together with genes that encode cobalamin's biosynthesis, the co-factor, necessary of the methylmalonyl-CoA mutase. The key enzyme in the ethylmalonyl-CoA pathway is the crotonyl-CoA reductase/carboxylase. The genome of *S. albus* is coding for two orthologous of *S. coelicolor* Ccr (SCO6473): XNR_0456 and XNR_5889 (Table 3). The latter one is located within the type I PKS-NRPS gene cluster and is most probably involved in the supply of precursors into respective biosynthesis pathway. We also were able to identify the ethylmalonyl-CoA mutase MeaA (SCO6472) orthologue encoded by the gene XNR_0457 in *S. albus*. As in *S. coelicolor*, XNR_0457 is forming one operon with *ccr2* (XNR_0456) and several other genes putatively involved in the ethylmalonyl-CoA pathway. Lastly, the genome of *S. albus* is coding for a valine dehydrogenase Vdh (SCO4089) orthologue (XNR_2839), putatively involved in initial step of branched chain amino acids degradation pathway that utilizes methylmalonyl-CoA as one of intermediates (Table 3).

3.4.2. Metabolic pathways involving ethylmalonyl-CoA in *S. albus*

3.4.2.1. Ethylmalonyl-CoA is supplied by Crotonyl-CoA Carboxylase/Reductase encoded by XNR_0456

A large part of pamamycins diversity is due to the incorporation of ethylmalonyl-CoA in certain derivatives.⁷⁶ As such, Pam635 and Pam649 variants as well as some Pam621 derivatives carry ethyl side chains. We reasoned that by influencing the supply of

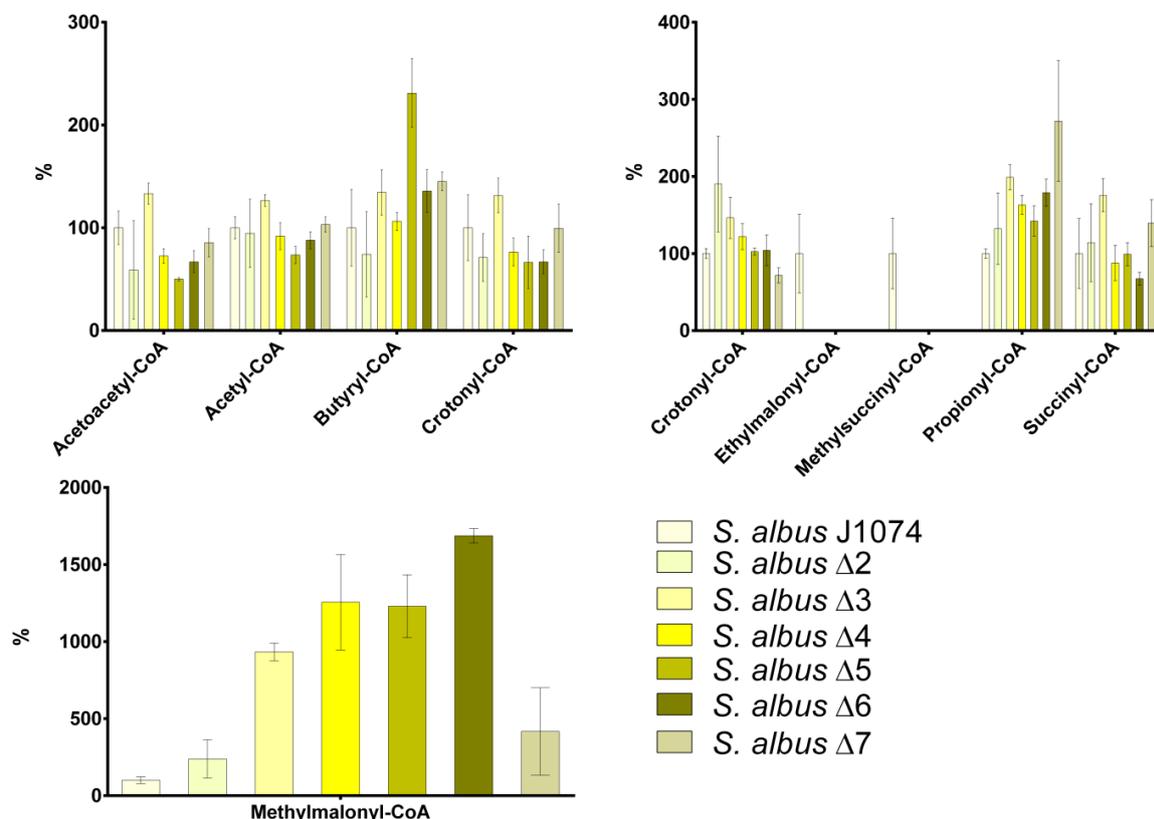


Figure 16: CoA-Ester levels in the constructed consecutive *S. albus* knockout strains compared to *S. albus* J1074.

ethylmalonyl-CoA the spectra of accumulated pamamycins can be simplified significantly. As described above, the genome of *S. albus* J1074 is coding for two paralogues of Ccr. Firstly, we analyzed the strain *S. albus* Del1. This strain has a deletion of two gene clusters responsible for the production of candidicin and frontalamides. The NRPS/Type 1 PKS Cluster includes the putative *ccr1* gene XNR_5889. After the introduction of the R2 Cosmid, which carries pamamycins entire biosynthetic gene cluster, no difference in either level or the produced spectrum of pamamycins was observed. As a next step, the gene XNR_0456 was markerless and in frame deleted from the chromosome of *S. albus* Del1. The CoA-Ester pool

as well as the spectrum of produced pamamycins from the resulting strain *S. albus* $\Delta 2$ was measured. As a result, the complete absence of ethylmalonyl-CoA and its subsequent product methylsuccinyl-CoA and a 1.9-fold increase of ethylmalonyl-CoA's predecessor crotonyl-CoA were observed, when compared to the wild type strain *S. albus* J1074. Additionally, a 2.3-fold increase of methylmalonyl-CoA was detected in *S. albus* $\Delta 2$ (Figure 16). After the introduction of the R2 cosmid into *S. albus* $\Delta 2$, the production level of pamamycins was compared to the wild type strain. The overall production of all pamamycins was reduced by 20% in the measured triplicate. The strain demonstrated a 92.5% decrease in the production of pam649 as well as a decrease of the smaller pamamycins that seem to incorporate ethylmalonyl-CoA. The accumulation of pam635 dropped by 80% when compared to the original production level, while the accumulation of Pam621 dropped by 5%. Meanwhile, the production of pam565, pam579, pam593 and pam607 increased from 0.16% to 0.62%, from 3.1% to 5.3%, from 11.4% to 16.26% and from 39.1% to 46.2% (Figure 17).

3.4.2.2. Ethylmalonyl-CoA mutase (MeaA)

The ethylmalonyl-CoA pathway is proposed to be involved in the assimilation of acetate in order to substitute the glyoxylate cycle for replenishing the tricarboxylic acid cycle during growth on this carbon source. This pathway is widely spread among *Streptomyces* and *S. albus*'s genome is predicted to code for almost all enzymes involved in the conversion of acetate to succinyl-CoA, including genes for a putative methylmalonyl-CoA/ethylmalonyl-CoA epimerase (XNR_1439) and a (2R)-ethylmalonyl-CoA carbonylmotase (XNR_0457). We assumed that blocking the conversion of ethylmalonyl-CoA to methylsuccinyl-CoA will increase the intracellular ethylmalonyl-CoA pool. However, the deletion of methylmalonyl-CoA/ethylmalonyl-CoA epimerase was predicted to also affect the methylmalonyl-CoA metabolism. To avoid this, we have blocked the next step in the ethylmalonyl-CoA consumption by deleting the gene XNR_0457 (MeaA) within the chromosome of *S. albus* Del1 and $\Delta 2$. The CoA-Ester pool of *S. albus* Del1 Δ meaA as well as the produced spectrum of pamamycins was measured. As a result, the product methylsuccinyl-CoA was no longer detectable, the ethylmalonyl-CoA pool increased 1.6-fold and the propionyl-CoA pool increased 1.3-fold (Figure 16). After introducing the R2 cosmid into *S. albus* Del1 Δ meaA and

S. albus $\Delta 2$, the produced spectrum of pamamycins was analyzed. While in *S. albus* Del1 the overall production of pamamycins was decreased by 30%, the knockout of *meaA* resulted in the accumulation of high molecular weight pamamycins. Pam621 was increased from 33.9% to 38.6%, Pam635 from 11.5% to 14.79%, and Pam649 from 0.6% to 0.9% (Figure 17). At the same time, a decreased production of low molecular weight pamamycins was observed. The knockout of *meaA* in $\Delta 2$ resulted in a reversed effect of the *ccr2* knockout. (Data not shown)

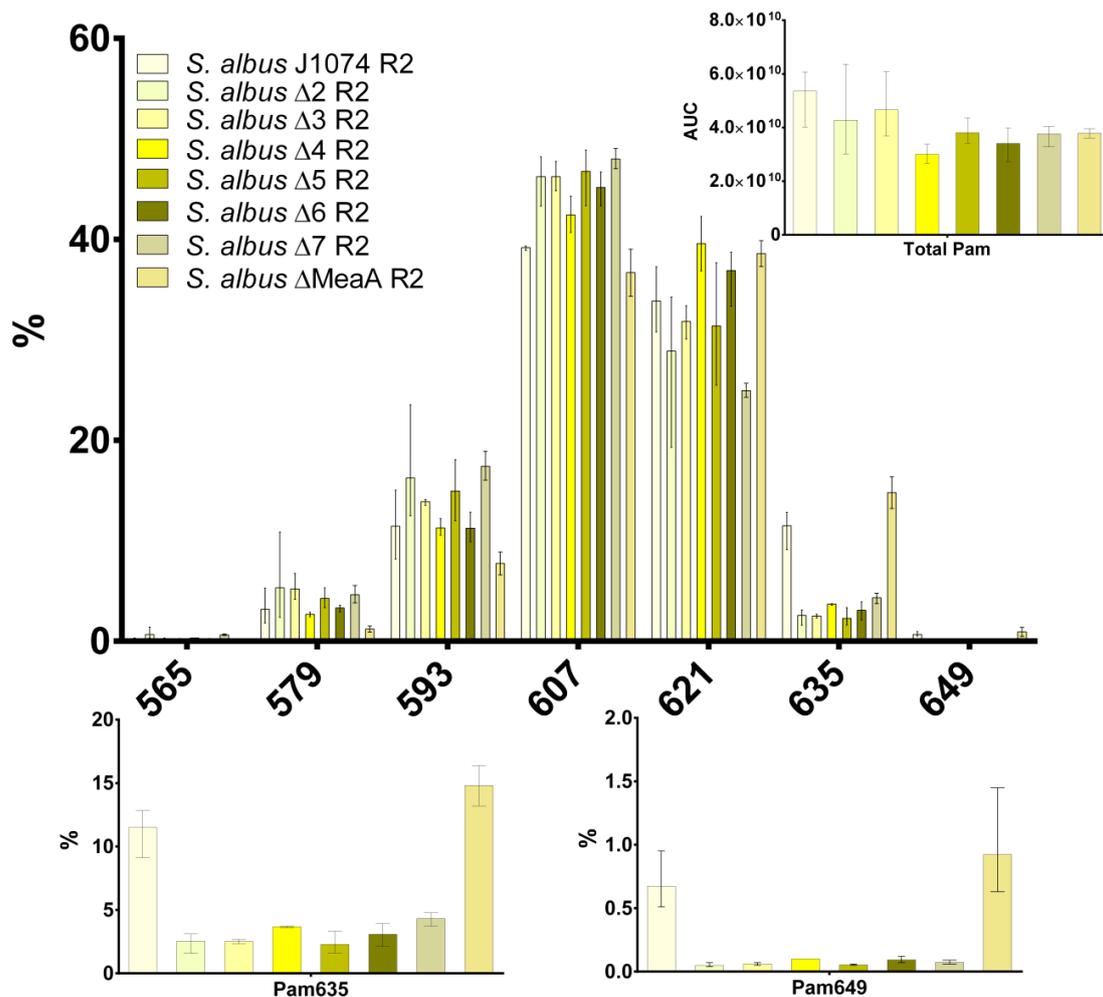


Figure 17: The production of pamamycins in the constructed *S. albus* strains.

3.4.3. Metabolic pathways involving methylmalonyl-CoA

3.4.3.1. Methylmalonyl-CoA mutase (Mcm)

The methylmalonyl-CoA mutase XNR_4665 (Mcm) is of special interest in the regulation of methylmalonyl-CoA within *Streptomyces*. The performed radical isomerization, which is dependent on the media conditions, can be performed in both directions. To elucidate its

potential role in the supply of methylmalonyl-CoA within *S. albus*, we deleted the coding gene in the constructed strain *S. albus* $\Delta 2$ resulting in the mutant *S. albus* $\Delta 3$. The CoA-ester pool of $\Delta 3$ was measured. When compared to *S. albus* $\Delta 2$, most esters remained unchanged in the newly obtained mutant strain. Nevertheless, we observed a 5.6-fold increase of methylmalonyl-CoA and a 1.7-fold increase of propionyl-CoA (Figure 16). After the introduction of the R2 cosmid, the production of pamamycins were measured and compared to its predecessor strain $\Delta 2$. Surprisingly, the production of pamamycins in the constructed mutant remained unchanged (Figure 17).

3.4.3.2. Propionyl-CoA carboxylase (PCC1-2)

It was originally reported that in *S. coelicolor*, PCC3 is responsible for the carboxylation of propionyl-CoA to methylmalonyl-CoA; orthologous of XNR_4024 (PCC3) can be found in the genome of *S. albus*. The genome of *S. albus* possesses two additional putative acyl-CoA carboxylase complexes with unknown functions similar to *S. coelicolors* PCC3 (Table 3). In order to elucidate their potential role in the production of methylmalonyl-CoA and ethylmalonyl-CoA we deleted the corresponding genes XNR_2273 (PCC1) and XNR_4211 (PCC2) within the chromosome of *S. albus* Del1 and *S. albus* $\Delta 3$ and, which lacks both *ccr* genes and the *mcm* gene. The mutant strains generated by the deletion of XNR_2273 (PCC1), in *S. albus* $\Delta 3$ resulted in the strain *S. albus* $\Delta 4$ and the subsequent inactivation of XNR_4211 (PCC3) in *S. albus* $\Delta 5$. The mutant strains generated by the deletion of either XNR_2273 (PCC1) or XNR_4211 (PCC2) in *S. albus* $\Delta 2$ resulted in the strains *S. albus* $\Delta 2 \Delta pcc1$ and *S. albus* $\Delta 2 \Delta pcc2$. The CoA-ester pools of all generated strains were compared to their corresponding predecessor. The single knockouts of either *pcc1* or *pcc2* in *S. albus* $\Delta 2$ had no major impact for precursors involved in pamamycins biosynthesis. Nevertheless, we observed a 1.8-fold increase of acetyl-CoA after the knockout of *pcc1*, as well as a 1.85-fold increase of succinyl-CoA after the knockout of *pcc2* (Figure 18). The analysis of the CoA-Ester pool of *S. albus* $\Delta 4$ and $\Delta 5$ revealed a 1.35-fold increase of methylmalonyl-CoA compared to mutant $\Delta 3$. Furthermore, $\Delta 5$ revealed a 2.17-fold increase of butyryl-CoA compared to $\Delta 4$ (Figure 16). After the introduction of the R2 cosmid into the mutant's $\Delta 4$ and $\Delta 5$, we compared the produced spectra of pamamycins. The overall production of pamamycins in $\Delta 4$ was decreased, when compared to $\Delta 3$, by 35%. However, the mutant $\Delta 5$, when compared

to $\Delta 4$, showed an increased overall production of 26%. Furthermore, it was observed that mutant $\Delta 4$ showed an increased production of high molecular weight pamamycins, while producing a decreased amount of low molecular weight pamamycins, when compared to its predecessor $\Delta 3$. While Pam649 remained unchanged, the production of Pam635 and Pam621 increased from 2.5% to 3.7% and from 31.8% to 39.8%, respectively. The constructed mutant $\Delta 5$ showed no difference in the production of pamamycins when compared to $\Delta 3$ and therefore, showed the reversed effects when compared to $\Delta 4$ (Figure 17).

3.4.3.3. Valine dehydrogenase (Vdh)

It has been hypothesized that the valine dehydrogenase pathway provides substantial amounts of methylmalonyl-CoA within *S. albus* J1074. To evaluate its influence, the gene XNR_2839 was deleted markerless and in frame from the chromosome of *S. albus* J1074 and

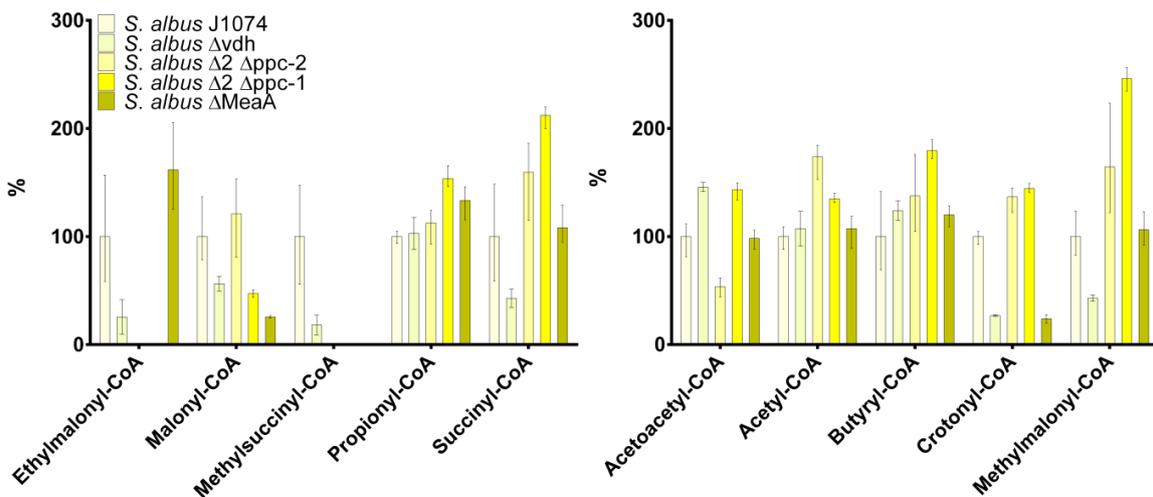


Figure 18: CoA-Ester levels in the constructed single gene knockout *S. albus* strains compared to *S. albus* J1074.

S. albus $\Delta 5$, resulting in the strains *S. albus* Δvdh and *S. albus* $\Delta 6$. The CoA-ester pools of both constructed mutants were analyzed and compared to their corresponding predecessors. Surprisingly, a decrease of a multitude of CoA esters in *S. albus* Δvdh was observed (Figure 18). While acetoacetyl-CoA, acetyl-CoA and butyryl-CoA remained unchanged in the constructed mutant, crotonyl-CoA was reduced by 76%, ethylmalonyl-CoA by 75%, methylsuccinyl-CoA by 82% and methylmalonyl-CoA by 58%. After introduction of the R2 cosmid, the produced spectrum of pamamycins was analyzed. Contrary to what the CoA-

ester analysis indicates, the amount and the composition of the produced pamamycins is identical when compared to the wild type strain. The CoA-ester analysis of the mutant *S. albus* $\Delta 6$ revealed a 1.37-fold increase of methylmalonyl-CoA, while all other measured CoA-esters remained unchanged (Figure 16). After the introduction of the R2 cosmid and the analysis of the production of pamamycins, no major changes were observed (Figure 17). To further evaluate the influence of Vdh on the supply of methylmalonyl-CoA in the heterologous host, we grew *S. albus* J1074 and *S. albus* $\Delta 5$ in the presence of ^{13}C -2-L-valine and monitored the incorporation of ^{13}C into pamamycins. As result, the derivatives Pam607, Pam621 and Pam635 showed an increase of the +1 ^{13}C isotope peak from 36-40% to 46-57% indicating the incorporation of methylmalonyl-CoA derived from the L-Valine degradation pathway.

3.4.3.4. Propionyl-CoA carboxylase (PCC3)

After the attempt to elucidate the function of PCC1 and PCC2 as described in 3.4.3.2, the function of PCC3 was analyzed. Therefore, XNR_4024 (PCC3) was deleted in frame and markerless from the chromosome of *S. albus* $\Delta 6$, resulting in the mutant *S. albus* $\Delta 7$. The CoA-ester pool of the constructed mutant was analyzed and revealed a 4.0-fold decrease of methylmalonyl-CoA as well as a 1.5-fold increase of propionyl-CoA (Figure 16). After the introduction of the R2-Cosmid into *S. albus* $\Delta 7$, the production of pamamycins was analyzed. The production of Pam621, containing only methylmalonyl-CoA units, was decreased from 36.9% to 25.0%, while the production of Pam593 increased from 11.2% to 17.4% (Figure 17).

3.4.3.5. Remaining methylmalonyl-CoA supply pathways

The previously described deletion experiments indicate that the supply of methylmalonyl-CoA in *S. albus* relies on more pathways than we were able to identify. To elucidate the possibility whether the remaining amount of methylmalonyl-CoA derives from amino acids, we grew *S. albus* J1074 R2, *S. albus* $\Delta 6$ R2 and *S. albus* $\Delta 7$ R2 in the presence of a ^{13}C labelled amino acid mixture and monitored the incorporation of ^{13}C into pamamycins. While only a moderate incorporation in pamamycins Pam607, Pam621 and Pam635 was observed in *S. albus* $\Delta 6$ R2 and *S. albus* $\Delta 7$ R2, resulting in ^{13}C isotope shifts from 39-43% to 44-47%, a huge incorporation into the same Pamamycins was observed in the wild type strain *S. albus* J1074 R2, where the ^{13}C isotope peak shifted from 41-42% to 53-56%.

3.5. Discussion

Our investigations of the precursor supply metabolism in the heterologous host *S. albus* revealed the impact of enzymes on the intracellular CoA-ester concentration as well as their subsequent influence on the production of pamamycins. The deletion of the Crotonyl-CoA carboxylase/reductase 1 (CCR1; XNR_5889) showed no significant impact on the intracellular concentration of ethylmalonyl-CoA. Its location within the NRPS/Type-I-PKS cluster as well as its lack of influence indicates a putative inactivity or a direct involvement in the biosynthesis of the frontalamides. In contrast, the deletion of Ccr2 resulted in the discontinuation of the ethylmalonyl-CoA production in *S. albus* $\Delta 2$. Therefore, it can be hypothesized that Ccr2 is the only source of ethylmalonyl-CoA in *S. albus*. Nevertheless, the strains lacking Ccr2 still produce minimal amounts of the putatively ethylmalonyl-CoA containing Pamamycins Pam635 and Pam649. Therefore, a secondary source of ethylmalonyl-CoA, producing small amounts, which are not detectable with our methods, seems plausible. It has been reported that *S. coelicolors* acetyl-CoA carboxylase can utilize acetyl-, propionyl- and butyryl-CoA with approximately the same specificity.²⁰⁹⁻²¹⁰ The carboxylation of butyryl-CoA results in the formation of ethylmalonyl-CoA and could provide the building block for the high molecular weight pamamycins. Another explanation for the presence of traces of Pam635 and Pam649 in the obtained extracts is the formation of derivatives with prolonged incorporated alkene side chains as described by Hartl *et al.*⁶⁷

The deletion of *meaA* (XNR_4665) resulted in the accumulation of ethylmalonyl-CoA as well as the deficit of methylsuccinyl-CoA. These findings confirm the postulated function as part of the degradation of ethylmalonyl-CoA. Our investigations revealed a direct correlation between the intracellular concentration of ethylmalonyl-CoA and the production level of the pamamycins Pam635 and Pam649. Consequently, it can be concluded that enzymes involved in the biosynthesis of pamamycins incorporate derivatives of malonyl-CoA solely because of their intracellular availability. Furthermore, the obtained results of the ethylmalonyl-CoA pathway indicate a logical and non-redundant function of the involved enzymes.

The methylmalonyl-CoA pathways show a greater complexity and redundancy. The functions of two enzymes with obvious impact on the intracellular concentration of methylmalonyl-CoA were identified. Since the deletion of Mcm resulted in a 5.62-fold increase of

methylmalonyl-CoA, the putatively bidirectional enzyme can be assumed to mutate methylmalonyl-CoA to succinyl-CoA under the chosen growth conditions. Since all CoA-measurements were carried out in the stationary bacterial growth phase, it can be hypothesized that Mcm is the main enzyme utilizing methylmalonyl-CoA and its deletion leads to the accumulation of the primary metabolite. The 4-fold reduction of the intracellular methylmalonyl-CoA concentration caused by the deletion of PCC3, clarified its function as the main supply for methylmalonyl-CoA within *S. albus*. Since the deletion of PCC1 and PCC2 from the genome of *S. albus* resulted in increased methylmalonyl-CoA titers, we can exclude the carboxylation of propionyl-CoA as their function. Nevertheless, their specific functions in the primary metabolism remain unclear. The single gene knockout of Vdh caused different effects, whether the gene was deleted from the chromosome of *S. albus* J1074 or *S. albus* $\Delta 5$. In *S. albus* J1074, an overall decrease of most measured CoA-esters was observed, while the knockout *S. albus* $\Delta 5$ resulted in an increased methylmalonyl-CoA pool. All obtained results highlight the complexity of the regulatory mechanism of the primary metabolism in *Streptomyces*. Since the increased as well as the decreased methylmalonyl-CoA titers did not affect the production of pamamycins as expected, it can be hypothesized, that the accumulation of methylmalonyl-CoA through the deletion of Mcm occurs during the late stage growth period, while pamamycins production starts after 12 h and reaches its maximum after 36 h. The observed negative effect on the overall production could be explained by additional stress exerted by the deletions. The remaining quantities of methylmalonyl-CoA in our final strain *S. albus* $\Delta 7$ indicate additional metabolic pathways leading to the formation of this primary metabolite. Feeding of a mixture of labelled amino acids resulted in their strongly decreased incorporation rates into pamamycins in *S. albus* $\Delta 6$ and *S. albus* $\Delta 7$, when compared to the wild type *S. albus* J1074. Therefore, the incorporation of amino acids degradation products other than those of L-Valine can be excluded as main route towards methylmalonyl-CoA. The remaining incorporation from 39-43% to 44-47% could be a result of pamamycins nitrogen methylation utilizing labelled methionine as a source. Another putative pathway is the unspecific carboxylation of propionyl-CoA through *S. albus*' Acetyl-CoA Carboxylase. The investigation of this possibility is rendered difficult by the essential nature of the mentioned enzyme.

3.6. Conclusion

Our investigations revealed the complexity of the primary metabolism in *Streptomyces*. Nevertheless, we were able to conclusively prove the influence of four enzymes in the pathways leading to either ethylmalonyl-CoA or methylmalonyl-CoA. The knockouts of the genes involved in the ethylmalonyl-CoA pathway resulted in obvious results regarding the production of the CoA esters and it was clearly possible to manipulate the spectrum of produced pamamycins through modification of the intracellular CoA-ester concentration. Such a simplified metabolic profile of the pamamycin producer *S. albus* enables further purification and activity/toxicity testing of single molecules. An analysis of the effects caused by the deletions of genes involved in the methylmalonyl-CoA pathway was more complex. The functions of two genes (Mcm and PCC3) resulted in comprehensible effects on the intracellular CoA ester concentration. However, these results were not reflected in the produced derivatives of pamamycins.

4. Discussion

4.1. General scope of the presented work

The presented work addressed two common problems in natural products research: 1. The constant supply of natural products with new chemical scaffolds and derivatives is one of the most important, if not the most important mission of natural products research. In order to avoid the rediscovery of known natural products, a widely applicable targeted genome mining approach was developed. Through its application, new polyketides with unique structural features were identified and their biosynthesis was elucidated. 2. Pamamycins biosynthesis was used as a model to explore the possibility of precursor supply modification mediated control of natural product's derivatives production. It was successfully shown that the targeted inactivation of genes encoding enzymes from primary metabolism responsible for the supply of different acyl-CoA has a strong influence on the production quantity of a particular natural product and the number of derivatives produced by a strain. Such manipulations can facilitate the downstream processing of natural products and enable the future development towards new active pharmaceutical ingredients.

4.2. Discovery of Natural Products from rare Actinobacteria

After decades of research on the genus *Streptomyces*, countless natural products have been identified. A statistical analysis of the Antibiotic Literature Database (ABL) showed that out of around twenty-three thousand microbial compounds approximately 32.1% (~7360) are derived from *Streptomyces* species.²¹¹ The number of discovered natural products from these bacteria raises the problem of re-discovery of the same natural products from different isolates. Strains sampled from different environments and at different distant locations can still produce the same compounds. One such example is the tetrahydroisoquinoline natural product perquinoline. Despite of its unique structure and biosynthesis, this compound is produced by *Streptomyces* sp. IB2014/016-6 isolated from samples collected at Lake Baikal and *S. odonellii* NRRL B24891 found from soil sampled in Brazil.²¹² Nevertheless, the phylum of Actinobacteria contains numerous genera (*Micromonospora*, *Actinoplanes*, *Actinomonodura*, *Pseudonocardia*)^{110, 211} that are rarely isolated from typical sources. Such underexplored species possess a great potential for the discovery of new chemical entities and thus are attracting more and more attention. At the

moment, there are around 2400 compounds in the ABL isolated from non-*Streptomyces* Actinobacteria²¹¹ and among those some are in clinical use such as rifamycin from *Amycolatopsis mediteranei*,²¹³ erythromycin from *Saccharopolyspora erythrea*,¹⁹³ teicoplanin from *Actinoplanes teichomyceticus*²¹⁴ and vancomycin from *Amycolatopsis orientalis*.^{211, 215} Despite of their obvious chemical potential, these strains are rather rare and working with them presents multiple challenges. While techniques for cultivating and genetic engineering are well developed for *Streptomyces*, many of the non-*Streptomyces* Actinobacteria species are difficult to manipulate.^{211, 216-218} To bypass potential cultivation issues with rare Actinobacteria, the heterologous expression of BGCs in *Streptomyces* hosts is an attractive alternative. Nevertheless, only few examples of successful rare Actinobacteria cluster expression in *Streptomyces* hosts have been described. In a targeted genome mining approach, 86 different *Salinispora* genomes were analyzed. This work resulted in the discovery of thiolactomycin's biosynthetic gene clusters. The BGC was found in four different *Salinispora* strains and its 26 kbp genomic region was cloned using a synthetic double-stranded DNA-mediated cloning strategy based on the transformation associated recombination (TAR) in *Saccharomyces cerevisiae*. After cloning, the BGC was conjugated and successfully expressed in *S. coelicolor* M1152.²¹⁹ In a similar attempt, the antibiotic taromycin A and its biosynthetic gene cluster were identified in *Saccharomonospora* sp. CNQ490. The similarity of taromycin's towards daptomycin's BGC spurred the authors' interest. After the successful cloning of the BGC, modifications in the regulatory elements of the cluster were carried out. 1 – A putative transcriptional repressor of the LuxR-regulator type (*tar20*) was deleted from the BGC due to its possible negative effects. 2 – A SARP regulator (*tar19*) was deleted in an identical fashion. While the expression of the unmodified cluster in *S. coelicolor* M1146 was unsuccessful, both constructed knockout mutants produced a series of unique chlorinated lipopeptides that are structurally highly similar to datpomycin.²²⁰ Another example of a successfully expressed BGC from rare Actinobacteria is enterocin. After the identification of enterocins BGC in *Salinispora pacifica* CNT-150 the cluster was cloned using TAR cloning and expressed in *S. lividans* and *S. coelicolor* M1146. Both *Streptomyces* strains showed production of enterocin at the level similar to the native producing strain.²²¹ Despite some obvious success in this direction, there is no systemic

approach described that exploits in full the genetic potential of these strains. Only selected clusters were targeted abandoning other BGCs.

4.2.1. Heterologous expression of BGCs from the rare Actinobacteria *Saccharothrix espanaensis* DSM44229

For the described targeted genome mining approach, the strain *S. espanaensis* was chosen as BGCs source. Even though this strain has a 9.3 Mbp large genome and 31 putative antiSMASH predicted BGCs, only the saccharomicins were isolated from its extracts. Furthermore, the laboratory work with *S. espanaensis* is highly challenging, as it does not form spores under laboratory conditions and it is not genetically amenable. To access the chemical potential encrypted in the genome of *S. espanaensis*, a genomic library was constructed with the pSMART-BAC-S vector as a backbone and end sequenced. To minimize the laboratory effort, BGCs were prioritized by type and degree of their homology to known clusters using antiSMASH. Altogether 17 unique BGCs on 15 BAC clones were conjugated into *S. albus* J1074 and *S. lividans* ΔYA6. Ultimately, two BACs (1E5 and 3C18) were successfully expressed in *S. lividans* and one construct (1E5) was active in *S. albus*. The described approach offers five major advantages, when compared to the screening of native strains. 1 – Novel compounds can easily be identified by comparison of the metabolic profiles of the native and recombinant host. 2 – The process of dereplication is minimized, since only the newly produced compounds need to be compared to the natural product databases and this can potentially be automatized. 3 – The chance of rediscovery is minimized since the exact mass for dereplication and the putative biosynthetic origin are known. 4 – Production and growth conditions for the heterologous hosts are standardized. Furthermore, the use of improved *Streptomyces* hosts can result in higher yields of natural products. They also have a minimized metabolic background, facilitating the compound isolation procedure. 5 – Genomic modification protocols of the BAC and the heterologous hosts are well-established and drastically simplify yield improvement and the elucidation of the biosynthetic pathway of the new natural product. Despite all mentioned advantages of the heterologous expression approach, several problems need to be mentioned. A major concern is the low expression rate of clusters. Only 11% (2 out of 17) of the chosen clusters were expressed in the heterologous hosts. Reasons for this can be manifold and include differences in the regulatory network of host and native strain, host resistance problems

towards proteins or the natural products, differences in promoter- and ribosomal binding regions and environmental factors, such as media and growth conditions, that could influence both, host and native strain differently. Yamanaka *et al.* activated the production of taromycin A in their heterologous hosts by influencing the regulatory elements of the expressed cluster. Even though this solution is applicable for single clusters, the laborious effort denies its application in a high throughput manner. A high throughput approach for natural product discovery was described by Seyedsayamdost.¹⁴ It was shown that the addition of sub-toxic concentration of antibiotics to the production medium strongly activated the secondary metabolism. In theory, this approach of supplementing sub toxic concentration of antibiotics could be combined with the expression of BGCs in heterologous hosts. The increased screening efforts can be carried out in a high throughput manner as described by Seyedsayamdost.¹⁴ Even though this though experiment could increase the expression rate, it only focuses on environmental factors and the global regulatory mechanism of the strains, leaving codon bias as well as resistance problems an remaining issue. The identification of a multidrug resistance protein such as the MdfA protein in *E. coli*²²² for *Streptomyces* and its expression in the host could increase host resistance and therefore, enable the expression of toxic compounds. Nevertheless, the expression of a multi-drug resistance protein will complicate the cloning process. Furthermore, modifications of the recognition of promoter- and ribosomal binding sites cannot be carried out easily in a high throughput scenario. Therefore, the identification and development of different host strains from a range of rare Actinobacteria would be ideal to express a variety of clusters from different sources. The host could then be chosen according to the phylogenetic relationship to the “cluster donor” strain, this should solve potential codon usage and promoter recognition issues and increase the expression rate drastically. Unpublished data from our laboratory showed that clusters derived from *Streptomyces* that are expressed in *Streptomyces* hosts have a much higher success rate of 35% (6 out of 17 BGCs). To achieve an optimal level of BGC expression a final theoretical experimental setup would be plausible. 1 – Isolation and development of different Actinobacteria hosts, ideally deriving from different species, including *Micromonospora*, *Actinoplanes*, *Actinomonodura* and *Pseudonocardia*. 2 – Identification of Actinobacteria strains with promising chemical potential through high cluster count and construction of sequenced genomic libraries from

the corresponding strains. 3 – Identification for the ideal host and “cluster donor” combination. 4 – Construction of the recombinant strains through intergeneric conjugation with chosen BGCs that were identified using antiSMASH. 5 – High throughput screening of the constructed recombinant strains, cultivated with the addition of sub-toxic levels of different antibiotics. 6 – Identification, dereplication, isolation and structural elucidation of unique peaks. This setup would circumvent codon biased and promoter issues through utilization of phylogenetically closely related host species, it would activate the global regulatory mechanism through addition of different chemicals and minimize the rediscovery chance, due to the known biosynthetic origin of the expressed compound.

Another concern of the described work in Section 2 was the constant decrease of the production rate of pentangumycin and SEK90. While the recombinant host *S. lividans* Δ YA6_1E5 initially produced approximately 0.35 mg*L⁻¹ of pentangumycin, the production dropped to 0.14 mg*L⁻¹ after several passages. This trend progressed even further with growing generation numbers to a point where pentangumycin was hardly detectable in extracts measured with LC-MS. In order to stabilize the production of the expressed compounds, the regulatory elements *penR1* and *penR2* were deleted from the pentangumycin BGC and, alternatively, over-expressed under the influence of the medium-strength A3 promoter.¹⁸⁷ Both changes of the regulatory elements did not stabilize the production. The deletion of *penR2* resulted in the cessation of pentangumycins and SEK90 biosynthesis. Due to these results, a clear conclusion of the role of each transcriptional regulator is hard to draw. Nevertheless, the cessation of the production of pentangumycin and SEK90 after knockout of *penR2* indicates its role as positive regulatory element. Since its expression under the constitutive A3 promoter did not affect the production of these natural products, a cross regulation between the two regulatory elements *penR1* and *penR2*, similar to jadomycins *jadR1* and *jadR2* seems plausible.⁵⁸

4.2.2. The angucyclinone pentangumycin

Pentangumycin is a member of the angucyclinone family with several unique features; one of the most drastic is the presence of the additional phenol unit at position C-6 of ring B. The modifications of ring B are quite rare and can only be found in few angucyclines, including jadomycins. Furthermore, the amination of position C-19 and its subsequent acetylation has

never been reported in angucyclinones and angucyclines before. The amination is hypothesized to be identical to the one reported in borrelidins biosynthesis.²²³ Initially, the methyl-group at position C-19 is oxidized by the cytochrome P450 enzyme PenA and the ferredoxin reductase PenB to a hydroxy group. The hydroxy group is subsequently oxidized to an aldehyde by one of the oxygenases encoded in pentangumycins BGC. The aminotransferase and carboxyltransferase PenC and PenK could finalize the formation of amide moiety of pentangumycins. The methylation of the hydroxy group at position C-1 was only reported in the chlorocyclinones and is putatively performed by the methyltransferase PenD.¹⁶⁷ From the biosynthetic perspective, the incorporation of ring E is the most interesting feature of pentangumycin. Due to the homology between enzymes, JadG, JadF and JadH that are supposedly involved in the Baeyer-Villiger oxidation and cleavage of jadomycin's ring B and the genes PenS, PenT and PenE an identical opening of pentangumycin's ring B can be hypothesized.⁴⁸ In contrast to jadomycin's biosynthesis no amino acid is incorporated, but the α -keto-4-hydroxy-phenylpyruvic acid, the first product in tyrosin degradation. This modification results, in contrast to jadomycins tertiary amine, in the unique C-C bond formation between ring B and ring E, which is described for the first time in the case of pentangumycin biosynthesis.

4.2.3. The shunt product SEK90

Another compound identified in the extract of *S. lividans* carrying pentangumycin BGC is multicyclic metabolite named SEK90. It was proposed to be a shunt product similar to SEK43. Indeed, SEK90 is a dimer of SEK43 connected through a butyryl-group. SEK43 was discovered through the expression of tetracenomycins minimal PKS genes, actinorhodin's ketoreductase and griseusin's aromatase without a corresponding cyclase in *S. coelicolor* A3(2).²²⁴ It is therefore, a product of the spontaneous cyclization of the minimal PKS derived decaketide. SEK90's production in *S. lividans* Δ YA6 after the introduction of the *pen* BGC indicates an aberrant functioning of Pen cyclases. Furthermore, the amount of SEK90 isolated from the production broth was almost four times higher when compared to pentangumycin (13 mg and 3.5 mg from SEK90 and pentangumycin, respectively). The large quantity of the shunt product could be related to an unbalance transcription of the cyclases and the minimal PKS genes or a malfunction of the cyclases due to misfolding of the enzymes.

4.3. Precursor flux modification

The complexity of the primary metabolism and the interdependence of metabolic pathways is a well-established fact.²²⁵ In order to summarize and classify the growing information about biosynthetic routes and pathways the Kyoto Encyclopedia of Genes and Genomes (KEGG) was developed. It provides information about identified pathways, including the genes that encode the respective enzymes as well as metabolic products of the corresponding reactions.²⁰⁷ At the same time, the reactions of the secondary metabolism, despite being very diverse and unique, use only a limited number of initial substrates. These essential building blocks originate from the primary metabolism. While polyketides are assembled from carboxylic acids (acetyl-CoA, propionyl-CoA, malonyl-CoA, methylmalonyl-CoA, ethylmalonyl-CoA), NRPs and RIPPs usually derive from proteinogenic and non-proteinogenic amino acids. As described in Section 1.3, these precursors come from a variety of sources and their supply is crucial for natural product yield, especially the supply of malonyl-CoA can be a restrictive factor for the assembly of polyketides.²²⁶

It was shown, that the expression of *Corynebacterium glutamicum* genes in *E. coli*, encoding an acetyl-CoA carboxylase and its partner biotin carboxylase, which provides the necessary cofactor, increased the (2s)-flavone production rate four-fold. Interestingly, the overexpression of *E. coli*'s native ACC had no influence on the production of flavones and even showed a negative effect on the vitality of the recombinant strains. Koffas *et al.* tested the influence of Gram-negative derived ACCs in a similar manner. The expression of the ACC from *Photobacterium luminescens* in *E. coli* under the control of the T7 promoter increased the pinocembrin production up to 5.5-fold. In order to provide sufficient amounts of the necessary cofactor, they co-expressed the *P. luminescens* biotin carboxylase. The co-expression resulted in an up to 13-fold increase in the flavonoid production.

The limiting factor in the production of erythromycin in *E. coli* was found to be the supply of methylmalonyl-CoA. As described in Section 1.3.2, methylmalonyl-CoA is mainly provided by the carboxylation of propionyl-CoA through the propionyl-CoA carboxylase (PCC). The expression of 13 different PCC homologues from different species in *E. coli*, resulted in a 6-desoxyerythronolide production ranging from 0.2 mg*L⁻¹ with the expression of *Chloroflexus aurantiacus* PCC and up to 6 mg*L⁻¹ with the expression of PCC from

Myxococcus fulvus.²²⁷ A comparative study between the influence of the methylmalonyl-CoA mutase pathway and the propionyl-CoA carboxylase pathway was carried out by Kealey *et al.*. In order to analyze the impact of each pathway they expressed PCC and Mcm derived from *S. coelicolor* in *E. coli* under control of the T7 promoter. While the recombinant strain with PCC produced $6.5 \text{ mg} \cdot \text{L}^{-1}$ 6-desoxyerythronolide, the expression of Mcm resulted in a production titer of $0.85 \text{ mg} \cdot \text{L}^{-1}$.²²⁸

4.3.1. Monensin and Salinomycin

Monensin is a polyether ionophoric polyketide antibiotic that was isolated from *S. cinnamonensis*.²²⁹⁻²³⁰ The derivatives monensin A and monensin B differ in a methyl-group due the incorporation of either methylmalonyl-CoA or ethylmalonyl-CoA in the polyketide backbone.²⁰² The native producer accumulates both derivatives equally and was therefore, a good target for the analysis of the precursor supply of ethylmalonyl-CoA. Reynolds *et al.* constructed strains with modified ethylmalonyl-CoA pathways and analyzed the influence on the distribution between monensin A and B. The recombinant strain *S. cinnamonensis* containing the *ccr* gene under control of *ermE* promoter showed no significant changes in the distribution of both derivatives. The mutant strain lacking the native *ccr* gene drastically shifted the production towards the methylmalonyl-CoA containing derivative monensin B. This shift was reversed through reintroduction of the *ermE* promoter controlled *ccr* gene. Furthermore, the authors were able to show a significant influence of the supplementation of different amino acids to the production medium on the distribution of monensins A and B in the *ccr* mutant. When valine was added as major component to a chemically defined medium, a 3:1 ratio of monensins A to B was observed. If either leucine or isoleucine was supplemented, the mutant strain mainly produced monensin B (>90 %). These results indicate a secondary pathway leading to ethylmalonyl-CoA, which is activated by valine. Unfortunately, the scientific paper only provided information about the distribution between the monensin derivatives and not the overall production levels.

Another example for the importance of precursor supply and its modification was shown with the polyether polyketide salinomycin, produced by *S. albus* DSM41398.²³¹⁻²³² The analysis of its BGC and the corresponding biosynthesis showed the utilization of one acetyl-CoA, five malonyl-CoA, six methylmalonyl-CoA and three ethylmalonyl-CoA during

salinomycin assembly.²³³ Among other genetic modifications, Bai *et al.* successfully improved salinomycin's production through the expression of *ccr* under control of the *ermE* promoter.²³⁴ This modification increased the intracellular ethylmalonyl-CoA concentration from 99.4 nmol*g⁻¹DCW (Dry cell weight) to 121.3 nmol*g⁻¹ DCW and thus, the salinomycin production from 6 g*L⁻¹ to 8 g*L⁻¹. In another example, the precursor supply of the high-yield salinomycin producer *S. albus* BK3-25 (18 g*L⁻¹) was analyzed.²³⁵ The concentrations of malonyl-CoA (130 nmol*g⁻¹ DCW), methylmalonyl-CoA (326 nmol*g⁻¹ DCW) and ethylmalonyl-CoA (59 nmol*g⁻¹ DCW) were measured. Due to these results, the authors hypothesized that increasing the ethylmalonyl-CoA pool would lead to improved salinomycin production levels. In contrast to the previously described approaches, every gene involved in the ethylmalonyl-CoA pathway was expressed under control of the *ermE* promoter. The overexpression of gene implicated in initial step of ethylmalonyl-CoA pathway, the formation of acetoacetyl-CoA from two acetyl-CoA units catalyzed by acetoacetyl-CoA synthetase, had no effect on salinomycins production. Nevertheless, the overexpression of either 3-oxoacyl-synthase III and 3-hydroxyacyl-CoA dehydratase or the crotonyl-CoA reductase increased the salinomycin production up to 22.4 g*L⁻¹ and 21.3 g*L⁻¹, respectively. Furthermore, authors constructed a strain that overexpressed all of three mentioned genes, which increased the production further up to 25.1 g*L⁻¹. The final strain had a higher intracellular ethylmalonyl-CoA concentration reaching 101.2 nmol*g⁻¹ DCW and a decreased concentration of malonyl-CoA and methylmalonyl-CoA, 110.5 nmol*g⁻¹ DCW and 223.4 nmol*g⁻¹ DCW, respectively.

4.3.2. Pamamycins

The biosynthesis pathway for macrodiolide polyketides pamamycins utilizes succinyl-CoA, acetyl-CoA, malonyl-CoA, methylmalonyl-CoA and ethylmalonyl-CoA as precursors. The apparently random incorporation of malonyl-CoA and its derivatives through the ketosynthases is responsible for the formation of more than 18 known pamamycin derivatives. Pamamycin's incredible bioactivity and the incorporation of different malonyl-CoA derivatives turn this natural product into an excellent object of study for two reasons. First, since the distribution of pamamycins derivatives is only controlled by the bioavailability of the corresponding precursors, it can indirectly be used to evaluate the intracellular level of methylmalonyl-CoA and ethylmalonyl-CoA through the production rate of different

pamamycin derivatives. Second, pamamycins' strong bioactivity against *Mycobacterium tuberculosis* and other microorganisms turn these natural products into a promising drug candidate, which due to its high derivative count and problematic chemical synthesis could not be further developed. In the presented work, we have shown that the production of different derivatives of pamamycins can be changed significantly through modifications of the precursors flux and therefore, this can simplify the purification of particular pamamycins without drastically reducing the overall yield.

The discussed modifications of the precursor supply in different *Streptomyces* species and *E. coli* mainly focused on a single precursor or a single pathway. The work described in Section 3 aimed to obtain a more global picture of the precursor supply of polyketides in the model actinobacteria strain *S. albus* J1074. Due to their part in the fatty acid synthesis, fatty acid degradation, glycolysis and involvement in the tricarboxylic acid cycle, malonyl-CoA, succinyl-CoA and acetyl-CoA cannot be freely modified without impairing the strains fitness and were, therefore, excluded.

An analysis of the genome of *S. albus* revealed two genes putatively encoding crotonyl-CoA carboxylase/reductase involved in the transformation of crotonyl-CoA to ethylmalonyl-CoA (*ccr1* and *ccr2*). In order to evaluate their influence on the intracellular concentration of ethylmalonyl-CoA, the corresponding genes were deleted from the chromosome of *S. albus*. The deletion of *ccr1* neither had an impact on the availability of ethylmalonyl-CoA nor the production of pamamycins. After the deletion of crotonyl-CoA carboxylase/reductase (*ccr2*) from the genome of *S. albus*, ethylmalonyl-CoA was no longer detectable. In contrast to this, the production of high molecular weight pamamycins Pam635 and Pam649, which usually contain at least one ethylmalonyl-CoA unit, was drastically reduced but did not completely cease. There are two possible ways to explain this contradiction. 1 – In Section 1.2.3.2.2 we described that the biosynthesis of pamamycins can utilize elongated acyl-CoA derivatives as elongation units and form thus far unknown pamamycin derivatives with modified side chains in position R or 2 – the remaining small amount of ethylmalonyl-CoA, which is left in the cell, is directly incorporated in pamamycins and the intracellular concentration of ethylmalonyl-CoA is too low to be detected with the used HPLC-MS based method.

The function of the putative ethylmalonyl-CoA mutase (MeaA) was analyzed through the knockout of the corresponding gene *meaA*. Its involvement in the degradation of ethylmalonyl-CoA to methylsuccinyl-CoA was verified, as the concentration of ethylmalonyl-CoA increased up to 200%. Furthermore, methylsuccinyl-CoA, the degradation product of ethylmalonyl-CoA, was no longer detectable in the constructed mutant strains and the high molecular weight pamamycins Pam635 and Pam649 showed increased production titers.

As described in Section 1.3.2, the supply of methylmalonyl-CoA is manifold and can occur through numerous pathways. The analysis of *S. albus* genome revealed three putative PCC complexes, the Mcm pathway and the valine degradation pathway. To analyze their influence on the production of pamamycins and the CoA-ester levels, consecutive deletions of all genes were carried out in the *ccr1* and *ccr2* deficient *S. albus* mutant. Interestingly, the modifications of all aforementioned methylmalonyl-CoA pathways showed no major influence on the production of pamamycins. Nevertheless, drastic shifts in the availability of methylmalonyl-CoA were observed. 1 – The *mcm* mutant *S. albus* $\Delta 3$ showed a 10-fold increased methylmalonyl-CoA pool when compared to *S. albus* J1074. This result indicates a conversion of methylmalonyl-CoA to succinyl-CoA under the chosen growth conditions through the Mcm and not as initially hypothesized a conversion of succinyl-CoA to methylmalonyl-CoA. Furthermore, the drastic increase indicates that *S. albus* does not possess another methylmalonyl-CoA degradation route and therefore, this precursor accumulates during the late stage of growth. 2 – The knockout of *pcc3* in *S. albus* $\Delta 7$ resulted in a 4-fold decreased methylmalonyl-CoA pool when compared to its predecessor *S. albus* $\Delta 6$, indicating that *pcc3* is the main route responsible for the formation of methylmalonyl-CoA in *S. albus*. Furthermore, it can be concluded that around 25% of the available methylmalonyl-CoA derives from a PCC independent pathway. Since *vdh*, *pcc1* and *pcc2* are deleted in *S. albus* $\Delta 7$, this pathway remains unknown. As the degradation pathway of leucine and isoleucine also results in the intermediate isobutyryl-CoA, it could replace the valine degradation pathway and reinstate the formation of methylmalonyl-CoA through the branched chain amino acid degradation described in Section 1.3.2 (Figure 7). In order to shut down the supply of methylmalonyl-CoA through the degradation of branched-chain amino acids, the deletion of the isobutyryl-CoA dehydrogenase appears reasonable, since multiple copies of the branched-chain amino acid dehydrogenase gene were found in the genome of

S. albus. Nevertheless, the deletion of this major pathway could interfere with the vitality of the corresponding strain. Another putative pathway for the formation of methylmalonyl-CoA lies in the increased propionyl-CoA concentration in *S. albus* $\Delta 7$, which is a result of the deletion of *pcc3*. As a result of the substrate specificity of the acetyl-CoA carboxylase, described in Section 1.3.2, ACC can synthesize methylmalonyl-CoA by carboxylation of propionyl-CoA. The high concentration of propionyl-CoA in *S. albus* $\Delta 7$ further increases the probability of this carboxylation. 3 – The deletions of *pcc1*, *pcc2* and *vdh* showed no major impacts on the CoA ester level in the consecutive knockout strains *S. albus* $\Delta 4$ - $\Delta 6$. Interestingly, the single deletion of *vdh* gene in *S. albus* J1074 had a different impact on the CoA ester levels. *S. albus* Δvdh shows a drastically reduced ethylmalonyl-CoA, malonyl-CoA, crotonyl-CoA and methylmalonyl-CoA pool, indicating a globally impaired fitness of this strain. Nevertheless, pamamycins production remained unchanged in this mutant.

4.4. Conclusion

In order to introduce new medicines to the market, a constant supply of the development pipeline with new lead structures is necessary. Since natural products have proven to be a great source of compounds with interesting pharmaceutical features, their continuous discovery is a necessity. The presented work provides a new way for the discovery of natural products derived from rare Actinobacteria using a systemic heterologous expression approach, which resulted in the identification of two novel polyketides. This universal approach can access the genome encoded chemical potential of diverse Actinobacteria. On the other hand, the heterologous expression in well-established hosts provides the possibilities for tuning the production of a desired compound using metabolic engineering. By affecting the acyl-CoA precursors' supply in *S. albus*, we achieved a modification of the spectra of accumulated pamamycins. Furthermore, we gained insights into the origin of different acyl-CoA esters in the strain and factors influencing their concentrations. This knowledge can be applied in order to tune the production of other polyketide antibiotics in this commonly used actinobacterial host strain.

Both projects underline the importance and versatility of Actinobacteria in the discovery of natural products and their putative development towards active pharmaceutical ingredients as well as emphasize the significance of the design of new approaches and techniques in order to access these biologically active compounds.

5. References

1. Newman, D. J.; Cragg, G. M.; Kingston, D. G. I., Chapter 5 - Natural Products as Pharmaceuticals and Sources for Lead Structures**Note: This chapter reflects the opinions of the authors, not necessarily those of the US Government. In *The Practice of Medicinal Chemistry (Fourth Edition)*, Wermuth, C. G.; Aldous, D.; Raboisson, P.; Rognan, D., Eds. Academic Press: San Diego, 2015; pp 101-139.
2. Schatz, A.; Bugie, E.; Waksman, S. A., Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. 1944. *Clin Orthop Relat Res* **2005**, (437), 3-6.
3. Davies, J., Where have All the Antibiotics Gone? *Can J Infect Dis Med Microbiol* **2006**, *17* (5), 287-90.
4. Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs from 1981 to 2014. *J Nat Prod* **2016**, *79* (3), 629-61.
5. Berdy, J., Thoughts and facts about antibiotics: where we are now and where we are heading. *J Antibiot (Tokyo)* **2012**, *65* (8), 385-95.
6. Pham, J. V.; Yilma, M. A.; Feliz, A.; Majid, M. T.; Maffetone, N.; Walker, J. R.; Kim, E.; Cho, H. J.; Reynolds, J. M.; Song, M. C.; Park, S. R.; Yoon, Y. J., A Review of the Microbial Production of Bioactive Natural Products and Biologics. *Front Microbiol* **2019**, *10*, 1404.
7. Chan, Y. A.; Podevels, A. M.; Kevany, B. M.; Thomas, M. G., Biosynthesis of polyketide synthase extender units. *Nat Prod Rep* **2009**, *26* (1), 90-114.
8. Kharel, M. K.; Pahari, P.; Shepherd, M. D.; Tibrewal, N.; Nybo, S. E.; Shaaban, K. A.; Rohr, J., Angucyclines: Biosynthesis, mode-of-action, new natural products, and synthesis. *Nat Prod Rep* **2012**, *29* (2), 264-325.
9. Walsh, C. T., A chemocentric view of the natural product inventory. *Nat Chem Biol* **2015**, *11* (9), 620-4.
10. Kong, D. X.; Guo, M. Y.; Xiao, Z. H.; Chen, L. L.; Zhang, H. Y., Historical variation of structural novelty in a natural product library. *Chem Biodivers* **2011**, *8* (11), 1968-77.
11. Cech, N. B.; Yu, K., Mass Spectrometry for Natural Products Research: Challenges, Pitfalls, and Opportunities. *LCGC North America* **2013**, *31* (11), 938-947.
12. Bode, H. B.; Bethe, B.; Hofs, R.; Zeeck, A., Big effects from small changes: possible ways to explore nature's chemical diversity. *ChemBiochem* **2002**, *3* (7), 619-27.
13. Pishchany, G.; Mevers, E.; Ndousse-Fetter, S.; Horvath, D. J., Jr.; Paludo, C. R.; Silva-Junior, E. A.; Koren, S.; Skaar, E. P.; Clardy, J.; Kolter, R., Amycomycin is a potent and specific antibiotic discovered with a targeted interaction screen. *Proc Natl Acad Sci U S A* **2018**, *115* (40), 10124-10129.
14. Seyedsayamdost, M. R., Toward a global picture of bacterial secondary metabolism. *J Ind Microbiol Biotechnol* **2019**, *46* (3-4), 301-311.
15. Moon, K.; Xu, F.; Zhang, C.; Seyedsayamdost, M. R., Bioactivity-HiTES Unveils Cryptic Antibiotics Encoded in Actinomycete Bacteria. *ACS Chem Biol* **2019**, *14* (4), 767-774.
16. Oldfield, E.; Lin, F. Y., Terpene biosynthesis: modularity rules. *Angew Chem Int Ed Engl* **2012**, *51* (5), 1124-37.
17. Barkei, J. J.; Kevany, B. M.; Felnagle, E. A.; Thomas, M. G., Investigations into viomycin biosynthesis by using heterologous production in *Streptomyces lividans*. *ChemBiochem* **2009**, *10* (2), 366-76.
18. Sussmuth, R. D.; Mainz, A., Nonribosomal Peptide Synthesis-Principles and Prospects. *Angew Chem Int Ed Engl* **2017**, *56* (14), 3770-3821.
19. Pfeifer, V.; Nicholson, G. J.; Ries, J.; Recktenwald, J.; Schefer, A. B.; Shawky, R. M.; Schroder, J.; Wohlleben, W.; Pelzer, S., A polyketide synthase in glycopeptide biosynthesis: the biosynthesis of the non-proteinogenic amino acid (S)-3,5-dihydroxyphenylglycine. *J Biol Chem* **2001**, *276* (42), 38370-7.

20. Trmčić, A.; Samelis, J.; Monnet, C.; Rogelj, I.; Matijašić, B. B., Complete nisin A gene cluster from *Lactococcus lactis* M78 (HM219853) — obtaining the nucleic acid sequence and comparing it to other published nisin sequences. *Genes & Genomics* **2011**, *33* (3), 217.
21. Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Goransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Muller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Sussmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep* **2013**, *30* (1), 108-60.
22. Hertweck, C., The biosynthetic logic of polyketide diversity. *Angew Chem Int Ed Engl* **2009**, *48* (26), 4688-716.
23. Shen, B., Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol* **2003**, *7* (2), 285-95.
24. Yu, D.; Xu, F.; Zeng, J.; Zhan, J., Type III polyketide synthases in natural product biosynthesis. *IUBMB Life* **2012**, *64* (4), 285-95.
25. Herbst, D. A.; Townsend, C. A.; Maier, T., The architectures of iterative type I PKS and FAS. *Nat Prod Rep* **2018**, *35* (10), 1046-1069.
26. Keatinge-Clay, A. T., The structures of type I polyketide synthases. *Nat Prod Rep* **2012**, *29* (10), 1050-73.
27. Bender, C. L.; Alarcon-Chaidez, F.; Gross, D. C., *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiol Mol Biol Rev* **1999**, *63* (2), 266-92.
28. Staunton, J.; Wilkinson, B., Biosynthesis of Erythromycin and Rapamycin. *Chem Rev* **1997**, *97* (7), 2611-2630.
29. Bevitt, D. J.; Cortes, J.; Haydock, S. F.; Leadlay, P. F., 6-Deoxyerythronolide-B synthase 2 from *Saccharopolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme. *Eur J Biochem* **1992**, *204* (1), 39-49.
30. Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L., Modular organization of genes required for complex polyketide biosynthesis. *Science* **1991**, *252* (5006), 675-9.
31. Byers, D. M.; Gong, H., Acyl carrier protein: structure-function relationships in a conserved multifunctional protein family. *Biochem Cell Biol* **2007**, *85* (6), 649-62.
32. Robbins, T.; Liu, Y. C.; Cane, D. E.; Khosla, C., Structure and mechanism of assembly line polyketide synthases. *Curr Opin Struct Biol* **2016**, *41*, 10-18.
33. Kaulmann, U.; Hertweck, C., Biosynthesis of polyunsaturated fatty acids by polyketide synthases. *Angew Chem Int Ed Engl* **2002**, *41* (11), 1866-9.
34. Campbell, C. D.; Vederas, J. C., Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers* **2010**, *93* (9), 755-63.
35. Yu, J., Current understanding on aflatoxin biosynthesis and future perspective in reducing aflatoxin contamination. *Toxins (Basel)* **2012**, *4* (11), 1024-57.
36. Smith, S.; Tsai, S. C., The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat Prod Rep* **2007**, *24* (5), 1041-72.
37. Chen, A.; Re, R. N.; Burkart, M. D., Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions. *Nat Prod Rep* **2018**, *35* (10), 1029-1045.
38. Niraula, N. P.; Kim, S. H.; Sohng, J. K.; Kim, E. S., Biotechnological doxorubicin production: pathway and regulation engineering of strains for enhanced production. *Appl Microbiol Biotechnol* **2010**, *87* (4), 1187-94.

39. Keatinge-Clay, A. T.; Maltby, D. A.; Medzihradzky, K. F.; Khosla, C.; Stroud, R. M., An antibiotic factory caught in action. *Nat Struct Mol Biol* **2004**, *11* (9), 888-93.
40. Korman, T. P.; Hill, J. A.; Vu, T. N.; Tsai, S. C., Structural analysis of actinorhodin polyketide ketoreductase: cofactor binding and substrate specificity. *Biochemistry* **2004**, *43* (46), 14529-38.
41. Javidpour, P.; Bruegger, J.; Srithahan, S.; Korman, T. P.; Crump, M. P.; Crosby, J.; Burkart, M. D.; Tsai, S. C., The determinants of activity and specificity in actinorhodin type II polyketide ketoreductase. *Chem Biol* **2013**, *20* (10), 1225-34.
42. Taguchi, T.; Awakawa, T.; Nishihara, Y.; Kawamura, M.; Ohnishi, Y.; Ichinose, K., Bifunctionality of ActIV as a Cyclase-Thioesterase Revealed by in Vitro Reconstitution of Actinorhodin Biosynthesis in *Streptomyces coelicolor* A3(2). *Chembiochem* **2017**, *18* (3), 316-323.
43. Dann, M.; Lefemine, D. V.; Barbatschi, F.; Shu, P.; Kunstmann, M. P.; Mitscher, L. A.; Bohonos, N., Tetrangomycin, a new quinone antibiotic. *Antimicrob Agents Chemother (Bethesda)* **1965**, *5*, 832-5.
44. Fischer, C.; Lipata, F.; Rohr, J., The complete gene cluster of the antitumor agent gilvocarcin V and its implication for the biosynthesis of the gilvocarcins. *J Am Chem Soc* **2003**, *125* (26), 7818-9.
45. Rohr, J.; Thiericke, R., Angucycline group antibiotics. *Nat Prod Rep* **1992**, *9* (2), 103-37.
46. Gould, S. J.; Halley, K. A., Biosynthesis of the benz[a]anthraquinone antibiotic PD 116198: evidence for a rearranged skeleton. *Journal of the American Chemical Society* **1991**, *113* (13), 5092-5093.
47. Sasaki, E.; Ogasawara, Y.; Liu, H. W., A biosynthetic pathway for BE-7585A, a 2-thiosugar-containing angucycline-type natural product. *J Am Chem Soc* **2010**, *132* (21), 7405-17.
48. Rix, U.; Wang, C.; Chen, Y.; Lipata, F. M.; Remsing Rix, L. L.; Greenwell, L. M.; Vining, L. C.; Yang, K.; Rohr, J., The oxidative ring cleavage in jadomycin biosynthesis: a multistep oxygenation cascade in a biosynthetic black box. *Chembiochem* **2005**, *6* (5), 838-45.
49. Chen, Y. H.; Wang, C. C.; Greenwell, L.; Rix, U.; Hoffmeister, D.; Vining, L. C.; Rohr, J.; Yang, K. Q., Functional analyses of oxygenases in jadomycin biosynthesis and identification of JadH as a bifunctional oxygenase/dehydrase. *J Biol Chem* **2005**, *280* (23), 22508-14.
50. Yang, K.; Han, L.; Ayer, S. W.; Vining, L. C., Accumulation of the angucycline antibiotic rabelomycin after disruption of an oxygenase gene in the jadomycin B biosynthetic gene cluster of *Streptomyces venezuelae*. *Microbiology* **1996**, *142* (Pt 1), 123-32.
51. Chen, Y.; Fan, K.; He, Y.; Xu, X.; Peng, Y.; Yu, T.; Jia, C.; Yang, K., Characterization of JadH as an FAD- and NAD(P)H-dependent bifunctional hydroxylase/dehydrase in jadomycin biosynthesis. *Chembiochem* **2010**, *11* (8), 1055-60.
52. Kharel, M. K.; Zhu, L.; Liu, T.; Rohr, J., Multi-oxygenase complexes of the gilvocarcin and jadomycin biosyntheses. *J Am Chem Soc* **2007**, *129* (13), 3780-1.
53. Kulowski, K.; Wendt-Pienkowski, E.; Han, L.; Yang, K.; Vining, L. C.; Hutchinson, C. R., Functional Characterization of the jadI Gene As a Cyclase Forming Angucyclinones. *Journal of the American Chemical Society* **1999**, *121* (9), 1786-1794.
54. Rix, U.; Zheng, J.; Remsing Rix, L. L.; Greenwell, L.; Yang, K.; Rohr, J., The dynamic structure of jadomycin B and the amino acid incorporation step of its biosynthesis. *J Am Chem Soc* **2004**, *126* (14), 4496-7.
55. Wang, L.; White, R. L.; Vining, L. C., Biosynthesis of the dideoxysugar component of jadomycin B: genes in the jad cluster of *Streptomyces venezuelae* ISP5230 for L-digitoxose assembly and transfer to the angucycline aglycone. *Microbiology* **2002**, *148* (Pt 4), 1091-103.
56. Wang, L.; Vining, L. C., Control of growth, secondary metabolism and sporulation in *Streptomyces venezuelae* ISP5230 by jadW(1), a member of the afsA family of gamma-butyrolactone regulatory genes. *Microbiology* **2003**, *149* (Pt 8), 1991-2004.
57. Yang, K.; Han, L.; Vining, L. C., Regulation of jadomycin B production in *Streptomyces venezuelae* ISP5230: involvement of a repressor gene, jadR2. *J Bacteriol* **1995**, *177* (21), 6111-7.
58. Yang, K.; Han, L.; He, J.; Wang, L.; Vining, L. C., A repressor-response regulator gene pair controlling jadomycin B production in *Streptomyces venezuelae* ISP5230. *Gene* **2001**, *279* (2), 165-73.

59. Xu, G.; Wang, J.; Wang, L.; Tian, X.; Yang, H.; Fan, K.; Yang, K.; Tan, H., "Pseudo" gamma-butyrolactone receptors respond to antibiotic signals to coordinate antibiotic biosynthesis. *J Biol Chem* **2010**, *285* (35), 27440-8.
60. Wang, L.; Tian, X.; Wang, J.; Yang, H.; Fan, K.; Xu, G.; Yang, K.; Tan, H., Autoregulation of antibiotic biosynthesis by binding of the end product to an atypical response regulator. *Proc Natl Acad Sci U S A* **2009**, *106* (21), 8617-22.
61. Doull, J. L.; Singh, A. K.; Hoare, M.; Ayer, S. W., Conditions for the production of jadomycin B by *Streptomyces venezuelae* ISP5230: effects of heat shock, ethanol treatment and phage infection. *J Ind Microbiol* **1994**, *13* (2), 120-5.
62. McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C., Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. *Nature* **1995**, *375* (6532), 549-54.
63. Liu, W. C.; Parker, L.; Slusarchyk, S.; Greenwood, G. L.; Gram, S. F.; Meyers, E., Isolation, characterization, and structure of rabelomycin, a new antibiotic. *J Antibiot (Tokyo)* **1970**, *23* (9), 437-41.
64. Kharel, M. K.; Pahari, P.; Lian, H.; Rohr, J., Enzymatic Total Synthesis of Rabelomycin, an Angucycline Group Antibiotic. *Organic Letters* **2010**, *12* (12), 2814-2817.
65. Kondo, S.; Yasui, K.; Natsume, M.; Katayama, M.; Marumo, S., Isolation, physico-chemical properties and biological activity of pamamycin-607, an aerial mycelium-inducing substance from *Streptomyces alboniger*. *J Antibiot (Tokyo)* **1988**, *41* (9), 1196-204.
66. Natsume, M.; Tazawa, J.; Yagi, K.; Abe, H.; Kondo, S.; Marumo, S., Structure-activity relationship of pamamycins: effects of alkyl substituents. *J Antibiot (Tokyo)* **1995**, *48* (10), 1159-64.
67. Hartl, A.; Stelzner, A.; Schlegel, R.; Heinze, S.; Hulsmann, H.; Fleck, W.; Grafe, U., Discovery of new homologous pamamycins by mass spectrometry and post mortem inhibitory action on autolysis of chicken embryo chorioallantoic membrane blood vessels. *J Antibiot (Tokyo)* **1998**, *51* (11), 1040-6.
68. Kozone, I.; Chikamoto, N.; Abe, H.; Natsume, M., De-N-methylpamamycin-593A and B, new pamamycin derivatives isolated from *Streptomyces alboniger*. *J Antibiot (Tokyo)* **1999**, *52* (3), 329-31.
69. Kozone, I.; Hashimoto, M.; Grafe, U.; Kawaide, H.; Abe, H.; Natsume, M., Structure-activity relationship of pamamycins: effect of side chain length on aerial mycelium-inducing activity. *J Antibiot (Tokyo)* **2008**, *61* (2), 98-102.
70. Shen, B.; Kwon, H. J., Macrotetrolide biosynthesis: a novel type II polyketide synthase. *Chem Rec* **2002**, *2* (6), 389-96.
71. McCann, P. A.; Pogell, B. M., Pamamycin: a new antibiotic and stimulator of aerial mycelia formation. *J Antibiot (Tokyo)* **1979**, *32* (7), 673-8.
72. Meyers, E.; Pansy, F. E.; Perlman, D.; Smith, D. A.; Weisenborn, F. L., The in Vitro Activity of Nonactin and Its Homologs: Monactin, Dinactin and Trinactin. *J Antibiot (Tokyo)* **1965**, *18*, 128-9.
73. Smith, W. C.; Xiang, L.; Shen, B., Genetic localization and molecular characterization of the nonS gene required for macrotetrolide biosynthesis in *Streptomyces griseus* DSM40695. *Antimicrob Agents Chemother* **2000**, *44* (7), 1809-17.
74. Kwon, H. J.; Smith, W. C.; Xiang, L.; Shen, B., Cloning and heterologous expression of the macrotetrolide biosynthetic gene cluster revealed a novel polyketide synthase that lacks an acyl carrier protein. *J Am Chem Soc* **2001**, *123* (14), 3385-6.
75. Hill, A. M., The biosynthesis, molecular genetics and enzymology of the polyketide-derived metabolites. *Nat Prod Rep* **2006**, *23* (2), 256-320.
76. Rebets, Y.; Brotz, E.; Manderscheid, N.; Tokovenko, B.; Myronovskiy, M.; Metz, P.; Petzke, L.; Luzhetskyy, A., Insights into the pamamycin biosynthesis. *Angew Chem Int Ed Engl* **2015**, *54* (7), 2280-4.
77. Tong, L., Acetyl-coenzyme A carboxylase: crucial metabolic enzyme and attractive target for drug discovery. *Cell Mol Life Sci* **2005**, *62* (16), 1784-803.
78. Cronan, J. E., Jr.; Waldrop, G. L., Multi-subunit acetyl-CoA carboxylases. *Prog Lipid Res* **2002**, *41* (5), 407-35.

79. Kim, Y. S., Malonate metabolism: biochemistry, molecular biology, physiology, and industrial application. *J Biochem Mol Biol* **2002**, *35* (5), 443-51.
80. An, J. H.; Kim, Y. S., A gene cluster encoding malonyl-CoA decarboxylase (MatA), malonyl-CoA synthetase (MatB) and a putative dicarboxylate carrier protein (MatC) in *Rhizobium trifolii*--cloning, sequencing, and expression of the enzymes in *Escherichia coli*. *Eur J Biochem* **1998**, *257* (2), 395-402.
81. An, J. H.; Lee, G. Y.; Jung, J. W.; Lee, W.; Kim, Y. S., Identification of residues essential for a two-step reaction by malonyl-CoA synthetase from *Rhizobium trifolii*. *Biochem J* **1999**, *344 Pt 1*, 159-66.
82. J. M. Berg, J. L. T., L. Stryer, *Biochemie - Glykolyse*. 7 ed.; 2013; Vol. Glycolyse und Gluconeogenese, p 46.
83. Quail, M. A.; Haydon, D. J.; Guest, J. R., The *pdhR-aceEF-lpd* operon of *Escherichia coli* expresses the pyruvate dehydrogenase complex. *Mol Microbiol* **1994**, *12* (1), 95-104.
84. J. M. Berg, J. L. T., L. Stryer, *Biochemie - Fettsäurestoffwechsel*. 7 ed.; 2013; Vol. Glycolyse und Gluconeogenese, p 35.
85. Weeks, G.; Shapiro, M.; Burns, R. O.; Wakil, S. J., Control of fatty acid metabolism. I. Induction of the enzymes of fatty acid oxidation in *Escherichia coli*. *J Bacteriol* **1969**, *97* (2), 827-36.
86. Starai, V. J.; Escalante-Semerena, J. C., Acetyl-coenzyme A synthetase (AMP forming). *Cell Mol Life Sci* **2004**, *61* (16), 2020-30.
87. Krivoruchko, A.; Zhang, Y.; Siewers, V.; Chen, Y.; Nielsen, J., Microbial acetyl-CoA metabolism and metabolic engineering. *Metab Eng* **2015**, *28*, 28-42.
88. Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M., KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* **1999**, *27* (1), 29-34.
89. Gokhale, R. S.; Saxena, P.; Chopra, T.; Mohanty, D., Versatile polyketide enzymatic machinery for the biosynthesis of complex mycobacterial lipids. *Nat Prod Rep* **2007**, *24* (2), 267-77.
90. Rodriguez, E.; Gramajo, H., Genetic and biochemical characterization of the alpha and beta components of a propionyl-CoA carboxylase complex of *Streptomyces coelicolor* A3(2). *Microbiology* **1999**, *145 (Pt 11)*, 3109-19.
91. Shaw, J. H.; Wolfe, R. R., Response to glucose and lipid infusions in sepsis: a kinetic analysis. *Metabolism* **1985**, *34* (5), 442-9.
92. Van der Geize, R.; Yam, K.; Heuser, T.; Wilbrink, M. H.; Hara, H.; Anderton, M. C.; Sim, E.; Dijkhuizen, L.; Davies, J. E.; Mohn, W. W.; Eltis, L. D., A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into *Mycobacterium tuberculosis* survival in macrophages. *Proc Natl Acad Sci U S A* **2007**, *104* (6), 1947-52.
93. Horswill, A. R.; Escalante-Semerena, J. C., *Salmonella typhimurium* LT2 catabolizes propionate via the 2-methylcitric acid cycle. *J Bacteriol* **1999**, *181* (18), 5615-23.
94. Arabolaza, A.; Shillito, M. E.; Lin, T. W.; Diacovich, L.; Melgar, M.; Pham, H.; Amick, D.; Gramajo, H.; Tsai, S. C., Crystal structures and mutational analyses of acyl-CoA carboxylase beta subunit of *Streptomyces coelicolor*. *Biochemistry* **2010**, *49* (34), 7367-76.
95. Leadlay, P. F., Purification and characterization of methylmalonyl-CoA epimerase from *Propionibacterium shermanii*. *Biochem J* **1981**, *197* (2), 413-9.
96. Marsh, E. N.; McKie, N.; Davis, N. K.; Leadlay, P. F., Cloning and structural characterization of the genes coding for adenosylcobalamin-dependent methylmalonyl-CoA mutase from *Propionibacterium shermanii*. *Biochem J* **1989**, *260* (2), 345-52.
97. McKie, N.; Keep, N. H.; Patchett, M. L.; Leadlay, P. F., Adenosylcobalamin-dependent methylmalonyl-CoA mutase from *Propionibacterium shermanii*. Active holoenzyme produced from *Escherichia coli*. *Biochem J* **1990**, *269* (2), 293-8.
98. Reeves, A. R.; Brikun, I. A.; Cernota, W. H.; Leach, B. I.; Gonzalez, M. C.; Weber, J. M., Effects of methylmalonyl-CoA mutase gene knockouts on erythromycin production in carbohydrate-based and oil-based fermentations of *Saccharopolyspora erythraea*. *J Ind Microbiol Biotechnol* **2006**, *33* (7), 600-9.

99. Tang, L.; Zhang, Y. X.; Hutchinson, C. R., Amino acid catabolism and antibiotic synthesis: valine is a source of precursors for macrolide biosynthesis in *Streptomyces ambofaciens* and *Streptomyces fradiae*. *J Bacteriol* **1994**, *176* (19), 6107-19.
100. Reynolds, K. A.; O'Hagan, D.; Gani, D.; Robinson, J. A., Butyrate metabolism in streptomycetes. Characterization of an intramolecular vicinal interchange rearrangement linking isobutyrate and butyrate in *Streptomyces cinnamonensis*. *Journal of the Chemical Society, Perkin Transactions 1* **1988**, (12), 3195-3207.
101. Erb, T. J.; Berg, I. A.; Brecht, V.; Muller, M.; Fuchs, G.; Alber, B. E., Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway. *Proc Natl Acad Sci U S A* **2007**, *104* (25), 10631-6.
102. Vrijbloed, J. W.; Zerbe-Burkhardt, K.; Ratnatilleke, A.; Grubelnik-Leiser, A.; Robinson, J. A., Insertional inactivation of methylmalonyl coenzyme A (CoA) mutase and isobutyryl-CoA mutase genes in *Streptomyces cinnamonensis*: influence on polyketide antibiotic biosynthesis. *J Bacteriol* **1999**, *181* (18), 5600-5.
103. Bentley, S. D.; Chater, K. F.; Cerdeño-Tárraga, A. M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C. H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabbinowitsch, E.; Rajandream, M. A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A., Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **2002**, *417* (6885), 141-147.
104. Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S., Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* **2003**, *21* (5), 526-31.
105. Ziemert, N.; Alanjary, M.; Weber, T., The evolution of genome mining in microbes - a review. *Nat Prod Rep* **2016**, *33* (8), 988-1005.
106. Weber, T.; Welzel, K.; Pelzer, S.; Vente, A.; Wohlleben, W., Exploiting the genetic potential of polyketide producing streptomycetes. *J Biotechnol* **2003**, *106* (2-3), 221-32.
107. National-Human-Genome-Research-Institute; K.A., W., DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). www.genome.gov/sequencingcostsdata **2019**.
108. Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de Los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H., antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* **2017**, *45* (W1), W36-W41.
109. Skinnider, M. A.; Merwin, N. J.; Johnston, C. W.; Magarvey, N. A., PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res* **2017**, *45* (W1), W49-W54.
110. Doroghazi, J. R.; Albright, J. C.; Goering, A. W.; Ju, K. S.; Haines, R. R.; Tchalukov, K. A.; Labeda, D. P.; Kelleher, N. L.; Metcalf, W. W., A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat Chem Biol* **2014**, *10* (11), 963-8.
111. Labeda, D. P.; Lechevalier, M. P., Amendment of the Genus *Saccharothrix* Labeda et al. 1984 and Descriptions of *Saccharothrix espanaensis* sp. nov., *Saccharothrix cryophilis* sp. nov., and *Saccharothrix mutabilis* comb. nov. *International Journal of Systematic and Evolutionary Microbiology* **1989**, *39* (4), 420-423.
112. Singh, M. P.; Petersen, P. J.; Weiss, W. J.; Kong, F.; Greenstein, M., Saccharomicins, novel heptadecaglycoside antibiotics produced by *Saccharothrix espanaensis*: antibacterial and mechanistic activities. *Antimicrob Agents Chemother* **2000**, *44* (8), 2154-9.
113. Strobel, T.; Al-Dilaimi, A.; Blom, J.; Gessner, A.; Kalinowski, J.; Luzhetska, M.; Puhler, A.; Szczepanowski, R.; Bechthold, A.; Ruckert, C., Complete genome sequence of *Saccharothrix espanaensis* DSM 44229(T) and comparison to the other completely sequenced Pseudonocardiaaceae. *BMC Genomics* **2012**, *13*, 465.

114. Berner, M.; Krug, D.; Bihlmaier, C.; Vente, A.; Muller, R.; Bechthold, A., Genes and enzymes involved in caffeic acid biosynthesis in the actinomycete *Saccharothrix espanaensis*. *J Bacteriol* **2006**, *188* (7), 2666-73.
115. Linnenbrink, A. Gewinnung glykosylierter Naturstoffe durch Biotransformation in *Saccharothrix espanaensis* sowie funktionelle Charakterisierung dreier Polyketidbiosynthesegencluster aus dem Mensacarcinproduzenten *Streptomyces* sp. Gö C4/4. Albert-Ludwing Universität, Freiburg, 2009.
116. Makrides, S. C., Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev* **1996**, *60* (3), 512-38.
117. Gomes, A.; Byregowda, S.; Veeregowda, B.; Vinayagamurthy, B., An Overview of Heterologous Expression Host Systems for the Production of Recombinant Proteins. *Advances in Animal and Veterinary Sciences* **2016**, *4*, 346-356.
118. Zhang, M. M.; Wang, Y.; Ang, E. L.; Zhao, H., Engineering microbial hosts for production of bacterial natural products. *Nat Prod Rep* **2016**, *33* (8), 963-87.
119. Huo, L.; Hug, J. J.; Fu, C.; Bian, X.; Zhang, Y.; Muller, R., Heterologous expression of bacterial natural product biosynthetic pathways. *Nat Prod Rep* **2019**.
120. Myronovskyi, M.; Luzhetskyy, A., Heterologous production of small molecules in the optimized *Streptomyces* hosts. *Nat Prod Rep* **2019**, *36* (9), 1281-1294.
121. Baltz, R. H., *Streptomyces* and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters. *J Ind Microbiol Biotechnol* **2010**, *37* (8), 759-72.
122. Gomez-Escribano, J. P.; Bibb, M. J., Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* **2011**, *4* (2), 207-15.
123. Chater, K. F.; Wilde, L. C., *Streptomyces albus* G mutants defective in the SalGI restriction-modification system. *J Gen Microbiol* **1980**, *116* (2), 323-34.
124. Zaburanyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A., Insights into naturally minimised *Streptomyces albus* J1074 genome. *BMC Genomics* **2014**, *15*, 97.
125. Fernández, E.; Lombó, F.; Méndez, C.; Salas, J. A., An ABC transporter is essential for resistance to the antitumor agent mithramycin in the producer *Streptomyces argillaceus*. *Molecular and General Genetics MGG* **1996**, *251* (6), 692-698.
126. Rodriguez, A. M.; Olano, C.; Vilches, C.; Mendez, C.; Salas, J. A., *Streptomyces antibioticus* contains at least three oleandomycin-resistance determinants, one of which shows similarity with proteins of the ABC-transporter superfamily. *Mol Microbiol* **1993**, *8* (3), 571-82.
127. Olano, C.; Abdelfattah, M. S.; Gullon, S.; Brana, A. F.; Rohr, J.; Mendez, C.; Salas, J. A., Glycosylated derivatives of steffimycin: insights into the role of the sugar moieties for the biological activity. *Chembiochem* **2008**, *9* (4), 624-33.
128. Chen, Y.; Wendt-Pienkowski, E.; Shen, B., Identification and utility of FdmR1 as a *Streptomyces* antibiotic regulatory protein activator for fredericamycin production in *Streptomyces griseus* ATCC 49344 and heterologous hosts. *J Bacteriol* **2008**, *190* (16), 5587-96.
129. Feng, Z.; Wang, L.; Rajska, S. R.; Xu, Z.; Coeffet-LeGal, M. F.; Shen, B., Engineered production of iso-migrastatin in heterologous *Streptomyces* hosts. *Bioorg Med Chem* **2009**, *17* (6), 2147-53.
130. Winter, J. M.; Moffitt, M. C.; Zazopoulos, E.; McAlpine, J. B.; Dorrestein, P. C.; Moore, B. S., Molecular basis for chloronium-mediated meroterpene cyclization: cloning, sequencing, and heterologous expression of the napradiomycin biosynthetic gene cluster. *J Biol Chem* **2007**, *282* (22), 16362-8.
131. Lombo, F.; Velasco, A.; Castro, A.; de la Calle, F.; Brana, A. F.; Sanchez-Puelles, J. M.; Mendez, C.; Salas, J. A., Deciphering the biosynthesis pathway of the antitumor thiocoraline from a marine actinomycete and its expression in two *streptomyces* species. *Chembiochem* **2006**, *7* (2), 366-76.
132. Myronovskyi, M.; Rosenkranzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A., Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab Eng* **2018**, *49*, 316-324.

133. Ruckert, C.; Albersmeier, A.; Busche, T.; Jaenicke, S.; Winkler, A.; Friethjonsson, O. H.; Hreggviethsson, G. O.; Lambert, C.; Badcock, D.; Bernaerts, K.; Anne, J.; Economou, A.; Kalinowski, J., Complete genome sequence of *Streptomyces lividans* TK24. *J Biotechnol* **2015**, *199*, 21-2.
134. Ochi, K., From microbial differentiation to ribosome engineering. *Biosci Biotechnol Biochem* **2007**, *71* (6), 1373-86.
135. Tala, A.; Wang, G.; Zemanova, M.; Okamoto, S.; Ochi, K.; Alifano, P., Activation of dormant bacterial genes by *Nonomuraea* sp. strain ATCC 39727 mutant-type RNA polymerase. *J Bacteriol* **2009**, *191* (3), 805-14.
136. Lacalle, R. A.; Tercero, J. A.; Jimenez, A., Cloning of the complete biosynthetic gene cluster for an aminonucleoside antibiotic, puromycin, and its regulated expression in heterologous hosts. *EMBO J* **1992**, *11* (2), 785-92.
137. Binnie, C.; Warren, M.; Butler, M. J., Cloning and heterologous expression in *Streptomyces lividans* of *Streptomyces rimosus* genes involved in oxytetracycline biosynthesis. *J Bacteriol* **1989**, *171* (2), 887-95.
138. Onaka, H.; Taniguchi, S.; Igarashi, Y.; Furumai, T., Cloning of the staurosporine biosynthetic gene cluster from *Streptomyces* sp. TP-A0274 and its heterologous expression in *Streptomyces lividans*. *J Antibiot (Tokyo)* **2002**, *55* (12), 1063-71.
139. Penn, J.; Li, X.; Whiting, A.; Latif, M.; Gibson, T.; Silva, C. J.; Brian, P.; Davies, J.; Miao, V.; Wrigley, S. K.; Baltz, R. H., Heterologous production of daptomycin in *Streptomyces lividans*. *J Ind Microbiol Biotechnol* **2006**, *33* (2), 121-8.
140. Felnagle, E. A.; Rondon, M. R.; Berti, A. D.; Crosby, H. A.; Thomas, M. G., Identification of the biosynthetic gene cluster and an additional gene for resistance to the antituberculosis drug capreomycin. *Appl Environ Microbiol* **2007**, *73* (13), 4162-70.
141. Bormann, C.; Mohrle, V.; Bruntner, C., Cloning and heterologous expression of the entire set of structural genes for nikkomycin synthesis from *Streptomyces tendae* Tu901 in *Streptomyces lividans*. *J Bacteriol* **1996**, *178* (4), 1216-8.
142. Bai, T., Yu, Y., Xu, Z., Tao, M., Construction of *Streptomyces lividans* SBT5 as an efficient heterologous expression host. *J. Huazhong Agric. Univ.* **2014**, *33*, 1-6.
143. Novakova, R.; Nunez, L. E.; Homerova, D.; Knirschova, R.; Feckova, L.; Rezuchova, B.; Sevcikova, B.; Menendez, N.; Moris, F.; Cortes, J.; Kormanec, J., Increased heterologous production of the antitumoral polyketide mithramycin A by engineered *Streptomyces lividans* TK24 strains. *Appl Microbiol Biotechnol* **2018**, *102* (2), 857-869.
144. Demain, A. L., Importance of microbial natural products and the need to revitalize their discovery. *J Ind Microbiol Biotechnol* **2014**, *41* (2), 185-201.
145. 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* (3), 311-35.
146. Berdy, J., Bioactive microbial metabolites. *J Antibiot (Tokyo)* **2005**, *58* (1), 1-26.
147. Ahmed, N., A flood of microbial genomes-do we need more? *PLoS One* **2009**, *4* (6), e5831.
148. Tokovenko, B.; Rebets, Y.; Luzhetskyy, A., *Automating Assessment of the Undiscovered Biosynthetic Potential of Actinobacteria*. 2016.
149. Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Muller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H., antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* **2015**, *43* (W1), W237-43.
150. Mazodier, P.; Petter, R.; Thompson, C., Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. *Journal of bacteriology* **1989**, *171* (6), 3583-3585.
151. Gust, B.; Chandra, G.; Jakimowicz, D.; Yuqing, T.; Bruton, C. J.; Chater, K. F., Lambda red-mediated genetic manipulation of antibiotic-producing *Streptomyces*. *Adv Appl Microbiol* **2004**, *54*, 107-28.

152. Brenner, M., Molekularbiologische Untersuchungen zur Biosynthese der Oligosaccharid-Antibiotika Saccharomicin A und B in *Saccharothrix espanaensis*. *Albert-Ludwig-Universität Freiburg* **2016**, *urn:nbn:de:bsz:25-opus-27680*.
153. Banerjee, P.; Erehman, J.; Gohlke, B. O.; Wilhelm, T.; Preissner, R.; Dunkel, M., Super Natural II--a database of natural products. *Nucleic Acids Res* **2015**, *43* (Database issue), D935-9.
154. Klementz, D.; Doring, K.; Lucas, X.; Telukunta, K. K.; Erxleben, A.; Deubel, D.; Erber, A.; Santillana, I.; Thomas, O. S.; Bechthold, A.; Gunther, S., StreptomeDB 2.0--an extended resource of natural products produced by streptomycetes. *Nucleic Acids Res* **2016**, *44* (D1), D509-14.
155. Pait, I. G. U.; Kitani, S.; Roslan, F. W.; Ulanova, D.; Arai, M.; Ikeda, H.; Nihira, T., Discovery of a new diol-containing polyketide by heterologous expression of a silent biosynthetic gene cluster from *Streptomyces lavendulae* FRI-5. *J Ind Microbiol Biotechnol* **2018**, *45* (2), 77-87.
156. Lithwick, G.; Margalit, H., Hierarchy of sequence-dependent features associated with prokaryotic translation. *Genome Res* **2003**, *13* (12), 2665-73.
157. Gustafsson, C.; Govindarajan, S.; Minshull, J., Codon bias and heterologous protein expression. *Trends Biotechnol* **2004**, *22* (7), 346-53.
158. Kane, J. F., Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Current Opinion in Biotechnology* **1995**, *6* (5), 494-500.
159. Liu, S.; Wang, M.; Du, G.; Chen, J., Improving the active expression of transglutaminase in *Streptomyces lividans* by promoter engineering and codon optimization. *BMC Biotechnol* **2016**, *16* (1), 75.
160. Olano, C.; Garcia, I.; Gonzalez, A.; Rodriguez, M.; Rozas, D.; Rubio, J.; Sanchez-Hidalgo, M.; Brana, A. F.; Mendez, C.; Salas, J. A., Activation and identification of five clusters for secondary metabolites in *Streptomyces albus* J1074. *Microb Biotechnol* **2014**, *7* (3), 242-56.
161. Freudl, R.; Braun, G.; Hindennach, I.; Henning, U., Lethal mutations in the structural gene of an outer membrane protein (OmpA) of *Escherichia coli* K12. *Mol Gen Genet* **1985**, *201* (1), 76-81.
162. Qiu, J.; Zhuo, Y.; Zhu, D.; Zhou, X.; Zhang, L.; Bai, L.; Deng, Z., Overexpression of the ABC transporter AvtAB increases avermectin production in *Streptomyces avermitilis*. *Appl Microbiol Biotechnol* **2011**, *92* (2), 337-45.
163. Doull, J. L.; Ayer, S. W.; Singh, A. K.; Thibault, P., Production of a novel polyketide antibiotic, jadomycin B, by *Streptomyces venezuelae* following heat shock. *J Antibiot (Tokyo)* **1993**, *46* (5), 869-71.
164. Drautz, H.; Zahner, H.; Rohr, J.; Zeeck, A., Metabolic products of microorganisms. 234. Urdamycins, new angucycline antibiotics from *Streptomyces fradiae*. I. Isolation, characterization and biological properties. *J Antibiot (Tokyo)* **1986**, *39* (12), 1657-69.
165. Robertson, A. W.; MacLeod, J. M.; MacIntyre, L. W.; Forget, S. M.; Hall, S. R.; Bennett, L. G.; Correa, H.; Kerr, R. G.; Goralski, K. B.; Jakeman, D. L., Post Polyketide Synthase Carbon-Carbon Bond Formation in Type-II PKS-Derived Natural Products from *Streptomyces venezuelae*. *J Org Chem* **2018**, *83* (4), 1876-1890.
166. Qin, Z.; Munnoch, J. T.; Devine, R.; Holmes, N. A.; Seipke, R. F.; Wilkinson, K. A.; Wilkinson, B.; Hutchings, M. I., Formicamycins, antibacterial polyketides produced by *Streptomyces formicae* isolated from African *Tetraponera* plant-ants. *Chem Sci* **2017**, *8* (4), 3218-3227.
167. Potterat, O.; Puder, C.; Wagner, K.; Bolek, W.; Vettermann, R.; Kauschke, S. G., Chlorocyclinones A-D, chlorinated angucyclinones from *Streptomyces* sp. strongly antagonizing rosiglitazone-induced PPAR-gamma activation. *J Nat Prod* **2007**, *70* (12), 1934-8.
168. Loudon, B. C.; Haarmann, D.; Lynne, A. M., Use of Blue Agar CAS Assay for Siderophore Detection. *J Microbiol Biol Educ* **2011**, *12* (1), 51-3.
169. Girvan, H. M.; Munro, A. W., Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. *Curr Opin Chem Biol* **2016**, *31*, 136-45.
170. Waskell, L.; Kim, J.-J. P., Electron Transfer Partners of Cytochrome P450. In *Cytochrome P450: Structure, Mechanism, and Biochemistry*, Ortiz de Montellano, P. R., Ed. Springer International Publishing: Cham %@ 978-3-319-12108-6, 2015; pp 33-68.

171. Sasaki, M.; Akahira, A.; Oshiman, K.; Tsuchido, T.; Matsumura, Y., Purification of cytochrome P450 and ferredoxin, involved in bisphenol A degradation, from *Sphingomonas* sp. strain AO1. *Appl Environ Microbiol* **2005**, *71* (12), 8024-30.
172. Novakova, R.; Homerova, D.; Feckova, L.; Kormanec, J., Characterization of a regulatory gene essential for the production of the angucycline-like polyketide antibiotic auricin in *Streptomyces aureofaciens* CCM 3239. *Microbiology* **2005**, *151* (Pt 8), 2693-706.
173. Ostash, I.; Ostash, B.; Walker, S.; Fedorenko, V., Proton-dependent transporter gene *IndJ* confers resistance to landomycin E in *Streptomyces globisporus* 1912. *Russian Journal of Genetics* **2007**, *43* (8), 852-857.
174. Rebets, Y.; Dutko, L.; Ostash, B.; Luzhetskyy, A.; Kulachkovskyy, O.; Yamaguchi, T.; Nakamura, T.; Bechthold, A.; Fedorenko, V., Function of *lanI* in regulation of landomycin A biosynthesis in *Streptomyces cyanogenus* S136 and cross-complementation studies with *Streptomyces* antibiotic regulatory proteins encoding genes. *Arch Microbiol* **2008**, *189* (2), 111-20.
175. Pickens, L. B.; Tang, Y., Oxytetracycline biosynthesis. *The Journal of biological chemistry* **2010**, *285* (36), 27509-27515.
176. Westrich, L.; Domann, S.; Faust, B.; Bedford, D.; Hopwood, D. A.; Bechthold, A., Cloning and characterization of a gene cluster from *Streptomyces cyanogenus* S136 probably involved in landomycin biosynthesis. *FEMS Microbiol Lett* **1999**, *170* (2), 381-7.
177. Meurer, G.; Gerlitz, M.; Wendt-Pienkowski, E.; Vining, L. C.; Rohr, J.; Hutchinson, C. R., Iterative type II polyketide synthases, cyclases and ketoreductases exhibit context-dependent behavior in the biosynthesis of linear and angular decapolyketides. *Chem Biol* **1997**, *4* (6), 433-43.
178. Zawada, R. J.; Khosla, C., Domain analysis of the molecular recognition features of aromatic polyketide synthase subunits. *J Biol Chem* **1997**, *272* (26), 16184-8.
179. Bao, W.; Sheldon, P. J.; Wendt-Pienkowski, E.; Hutchinson, C. R., The *Streptomyces peucetius* *dpsC* gene determines the choice of starter unit in biosynthesis of the daunorubicin polyketide. *J Bacteriol* **1999**, *181* (15), 4690-5.
180. Nachtigall, J., Strukturaufklärung von Naturstoffen aus Bakterien der Gattung *Streptomyces*. *Fakultät II – Mathematik und Naturwissenschaften der Technischen Universität Berlin* **2012**.
181. Huang, C.; Yang, C.; Zhu, Y.; Zhang, W.; Yuan, C.; Zhang, C., Marine Bacterial Aromatic Polyketides From Host-Dependent Heterologous Expression and Fungal Mode of Cyclization. *Front Chem* **2018**, *6*, 528.
182. de March, P.; Moreno-Mañas, M.; Casado, J.; Pleixats, R.; L. Roca, J.; Trius, A., *The reactivity of 4-hydroxy-6-methyl-2-pyrone towards aliphatic saturated and α,β -unsaturated aldehydes*. 1984; Vol. 21, p 85-89.
183. Sharif, E. U.; O'Doherty, G. A., Biosynthesis and Total Synthesis Studies on The Jadomycin Family of Natural Products. *European J Org Chem* **2012**, *2012* (11).
184. Yushchuk, O.; Kharel, M.; Ostash, I.; Ostash, B., Landomycin biosynthesis and its regulation in *Streptomyces*. *Applied Microbiology and Biotechnology* **2019**.
185. Dickens, M. L.; Ye, J.; Strohl, W. R., Analysis of clustered genes encoding both early and late steps in daunomycin biosynthesis by *Streptomyces* sp. strain C5. *J Bacteriol* **1995**, *177* (3), 536-43.
186. Olano, C.; Moss, S. J.; Brana, A. F.; Sheridan, R. M.; Math, V.; Weston, A. J.; Mendez, C.; Leadlay, P. F.; Wilkinson, B.; Salas, J. A., Biosynthesis of the angiogenesis inhibitor borrelidin by *Streptomyces parvulus* Tu4055: insights into nitrile formation. *Mol Microbiol* **2004**, *52* (6), 1745-56.
187. Siegl, T.; Tokovenko, B.; Myronovskyy, M.; Luzhetskyy, A., Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes. *Metab Eng* **2013**, *19*, 98-106.
188. Herrmann, S.; Siegl, T.; Luzhetska, M.; Petzke, L.; Jilg, C.; Welle, E.; Erb, A.; Leadlay, P. F.; Bechthold, A.; Luzhetskyy, A., Site-specific recombination strategies for engineering actinomycete genomes. *Appl Environ Microbiol* **2012**, *78* (6), 1804-12.
189. Green, M. R.; Sambrook, J.; Sambrook, J., *Molecular cloning : a laboratory manual*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y., 2012.

190. Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A., Practical *Streptomyces* Genetics. **2010**.
191. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic local alignment search tool. *J Mol Biol* **1990**, *215* (3), 403-10.
192. Myronovskiy, M.; Rosenkranzer, B.; Luzhetskyy, A., Iterative marker excision system. *Appl Microbiol Biotechnol* **2014**, *98* (10), 4557-70.
193. Bunch, R. L. M., James M. Erythromycin, its salts, and method of preparation. 29.09.1953, 1953.
194. Egerton, J. R.; Ostlind, D. A.; Blair, L. S.; Eary, C. H.; Suhayda, D.; Cifelli, S.; Riek, R. F.; Campbell, W. C., Avermectins, new family of potent anthelmintic agents: efficacy of the B1a component. *Antimicrob Agents Chemother* **1979**, *15* (3), 372-8.
195. Calne, R. Y.; Collier, D. S.; Lim, S.; Pollard, S. G.; Samaan, A.; White, D. J.; Thiru, S., Rapamycin for immunosuppression in organ allografting. *Lancet* **1989**, *2* (8656), 227.
196. Harris, M. J.; Rao, N. U.; Spaulding, E. H.; Tyson, R. R.; Zubrzycki, L., Antifungal action of nystatin on the fecal flora during administration of neomycin-polymyxin. *Antibiot Annu* **1955**, *3*, 681-9.
197. Endo, A.; Kuroda, M.; Tanzawa, K., Competitive inhibition of 3-hydroxy-3-methylglutaryl coenzyme A reductase by ML-236A and ML-236B fungal metabolites, having hypocholesterolemic activity. *FEBS Lett* **1976**, *72* (2), 323-6.
198. Dimarco, A.; Gaetani, M.; Dorigotti, L.; Soldati, M.; Bellini, O., Daunomycin: A New Antibiotic with Antitumor Activity. *Cancer Chemother Rep* **1964**, *38*, 31-8.
199. Kornfuehrer, T.; Eustáquio, A. S., Diversification of polyketide structures via synthase engineering. *MedChemComm* **2019**, *10* (8), 1256-1272.
200. Staunton, J.; Weissman, K. J., Polyketide biosynthesis: a millennium review. *Nat Prod Rep* **2001**, *18* (4), 380-416.
201. Hughes, A. J.; Keatinge-Clay, A., Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem Biol* **2011**, *18* (2), 165-76.
202. Liu, H.; Reynolds, K. A., Precursor supply for polyketide biosynthesis: the role of crotonyl-CoA reductase. *Metab Eng* **2001**, *3* (1), 40-8.
203. Flett, F.; Mersinias, V.; Smith, C. P., High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. *FEMS Microbiol Lett* **1997**, *155* (2), 223-9.
204. Woodyer, R. D.; Shao, Z.; Thomas, P. M.; Kelleher, N. L.; Blodgett, J. A.; Metcalf, W. W.; van der Donk, W. A.; Zhao, H., Heterologous production of fosfomycin and identification of the minimal biosynthetic gene cluster. *Chem Biol* **2006**, *13* (11), 1171-82.
205. Myronovskiy, M.; Welle, E.; Fedorenko, V.; Luzhetskyy, A., Beta-glucuronidase as a sensitive and versatile reporter in actinomycetes. *Appl Environ Microbiol* **2011**, *77* (15), 5370-83.
206. Peyraud, R.; Kiefer, P.; Christen, P.; Massou, S.; Portais, J. C.; Vorholt, J. A., Demonstration of the ethylmalonyl-CoA pathway by using ¹³C metabolomics. *Proc Natl Acad Sci U S A* **2009**, *106* (12), 4846-51.
207. Du, J.; Yuan, Z.; Ma, Z.; Song, J.; Xie, X.; Chen, Y., KEGG-PATH: Kyoto encyclopedia of genes and genomes-based pathway analysis using a path analysis model. *Mol Biosyst* **2014**, *10* (9), 2441-7.
208. Hall, P. R.; Wang, Y. F.; Rivera-Hainaj, R. E.; Zheng, X.; Pustai-Carey, M.; Carey, P. R.; Yee, V. C., Transcarboxylase 12S crystal structure: hexamer assembly and substrate binding to a multienzyme core. *EMBO J* **2003**, *22* (10), 2334-47.
209. Rodriguez, E.; Banchio, C.; Diacovich, L.; Bibb, M. J.; Gramajo, H., Role of an essential acyl coenzyme A carboxylase in the primary and secondary metabolism of *Streptomyces coelicolor* A3(2). *Appl Environ Microbiol* **2001**, *67* (9), 4166-76.

210. Diacovich, L.; Peiru, S.; Kurth, D.; Rodriguez, E.; Podesta, F.; Khosla, C.; Gramajo, H., Kinetic and structural analysis of a new group of Acyl-CoA carboxylases found in *Streptomyces coelicolor* A3(2). *J Biol Chem* **2002**, *277* (34), 31228-36.
211. Lazzarini, A.; Cavaletti, L.; Toppo, G.; Marinelli, F., Rare genera of actinomycetes as potential producers of new antibiotics. *Antonie Van Leeuwenhoek* **2001**, *79* (3-4), 399-405.
212. Rebets, Y.; Nadmid, S.; Paulus, C.; Dahlem, C.; Herrmann, J.; Hubner, H.; Ruckert, C.; Kiemer, A. K.; Gmeiner, P.; Kalinowski, J.; Muller, R.; Luzhetskyy, A., Perquinolines A-C: Unprecedented Bacterial Tetrahydroisoquinolines Involving an Intriguing Biosynthesis. *Angew Chem Int Ed Engl* **2019**, *58* (37), 12930-12934.
213. Sensi, P.; Margalith, P.; Timbal, M. T., Rifomycin, a new antibiotic; preliminary report. *Farmaco Sci* **1959**, *14* (2), 146-7.
214. Parenti, F.; Beretta, G.; Berti, M.; Arioli, V., Teichomycins, new antibiotics from *Actinoplanes teichomyceticus* Nov. Sp. I. Description of the producer strain, fermentation studies and biological properties. *J Antibiot (Tokyo)* **1978**, *31* (4), 276-83.
215. Geraci, J. E.; Heilman, F. R.; Nichols, D. R.; Ross, G. T.; Wellman, W. E., Some laboratory and clinical experiences with a new antibiotic, vancomycin. *Proc Staff Meet Mayo Clin* **1956**, *31* (21), 564-82.
216. Qin, S.; Li, J.; Chen, H. H.; Zhao, G. Z.; Zhu, W. Y.; Jiang, C. L.; Xu, L. H.; Li, W. J., Isolation, diversity, and antimicrobial activity of rare actinobacteria from medicinal plants of tropical rain forests in Xishuangbanna, China. *Appl Environ Microbiol* **2009**, *75* (19), 6176-86.
217. Baltz, R. H., Marcel Faber Roundtable: is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *J Ind Microbiol Biotechnol* **2006**, *33* (7), 507-13.
218. Zotchev, S. B., Marine actinomycetes as an emerging resource for the drug development pipelines. *J Biotechnol* **2012**, *158* (4), 168-75.
219. Tang, X.; Li, J.; Millan-Aguinaga, N.; Zhang, J. J.; O'Neill, E. C.; Ugalde, J. A.; Jensen, P. R.; Mantovani, S. M.; Moore, B. S., Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-directed Genome Mining. *ACS Chem Biol* **2015**, *10* (12), 2841-2849.
220. Yamanaka, K.; Reynolds, K. A.; Kersten, R. D.; Ryan, K. S.; Gonzalez, D. J.; Nizet, V.; Dorrestein, P. C.; Moore, B. S., Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc Natl Acad Sci U S A* **2014**, *111* (5), 1957-62.
221. Bonet, B.; Teufel, R.; Crusemann, M.; Ziemert, N.; Moore, B. S., Direct capture and heterologous expression of *Salinispora* natural product genes for the biosynthesis of enterocin. *J Nat Prod* **2015**, *78* (3), 539-42.
222. Edgar, R.; Bibi, E., MdfA, an *Escherichia coli* multidrug resistance protein with an extraordinarily broad spectrum of drug recognition. *J Bacteriol* **1997**, *179* (7), 2274-80.
223. Moss, S. J.; Carletti, I.; Olano, C.; Sheridan, R. M.; Ward, M.; Math, V.; Nur, E. A. M.; Brana, A. F.; Zhang, M. Q.; Leadlay, P. F.; Mendez, C.; Salas, J. A.; Wilkinson, B., Biosynthesis of the angiogenesis inhibitor borrelidin: directed biosynthesis of novel analogues. *Chem Commun (Camb)* **2006**, (22), 2341-3.
224. McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C., Engineered biosynthesis of novel polyketides. *Science* **1993**, *262* (5139), 1546-50.
225. Klosik, D. F.; Grimbs, A.; Bornholdt, S.; Hutt, M. T., The interdependent network of gene regulation and metabolism is robust where it needs to be. *Nat Commun* **2017**, *8* (1), 534.
226. Miyahisa, I.; Kaneko, M.; Funai, N.; Kawasaki, H.; Kojima, H.; Ohnishi, Y.; Horinouchi, S., Efficient production of (2S)-flavanones by *Escherichia coli* containing an artificial biosynthetic gene cluster. *Appl Microbiol Biotechnol* **2005**, *68* (4), 498-504.
227. Vandova, G. A.; O'Brien, R. V.; Lowry, B.; Robbins, T. F.; Fischer, C. R.; Davis, R. W.; Khosla, C.; Harvey, C. J.; Hillenmeyer, M. E., Heterologous expression of diverse propionyl-CoA carboxylases affects polyketide production in *Escherichia coli*. *J Antibiot (Tokyo)* **2017**, *70* (7), 859-863.

228. Murli, S.; Kennedy, J.; Dayem, L. C.; Carney, J. R.; Kealey, J. T., Metabolic engineering of *Escherichia coli* for improved 6-deoxyerythronolide B production. *J Ind Microbiol Biotechnol* **2003**, *30* (8), 500-9.
229. Haney, M. E., Jr.; Hoehn, M. M., Monensin, a new biologically active compound. I. Discovery and isolation. *Antimicrob Agents Chemother (Bethesda)* **1967**, *7*, 349-52.
230. Markowska, A.; Kaysiewicz, J.; Markowska, J.; Huczynski, A., Doxycycline, salinomycin, monensin and ivermectin repositioned as cancer drugs. *Bioorg Med Chem Lett* **2019**, *29* (13), 1549-1554.
231. Kevin li, D. A.; Meujo, D. A.; Hamann, M. T., Polyether ionophores: broad-spectrum and promising biologically active molecules for the control of drug-resistant bacteria and parasites. *Expert Opin Drug Discov* **2009**, *4* (2), 109-46.
232. Gumila, C.; Ancelin, M. L.; Delort, A. M.; Jeminet, G.; Vial, H. J., Characterization of the potent in vitro and in vivo antimalarial activities of ionophore compounds. *Antimicrob Agents Chemother* **1997**, *41* (3), 523-9.
233. Jiang, C.; Wang, H.; Kang, Q.; Liu, J.; Bai, L., Cloning and characterization of the polyether salinomycin biosynthesis gene cluster of *Streptomyces albus* XM211. *Appl Environ Microbiol* **2012**, *78* (4), 994-1003.
234. Lu, C.; Zhang, X.; Jiang, M.; Bai, L., Enhanced salinomycin production by adjusting the supply of polyketide extender units in *Streptomyces albus*. *Metab Eng* **2016**, *35*, 129-137.
235. Zhang, X.; Lu, C.; Bai, L., Conversion of the high-yield salinomycin producer *Streptomyces albus* BK3-25 into a surrogate host for polyketide production. *Sci China Life Sci* **2017**, *60* (9), 1000-1009.
236. Maresca, M.; Erler, A.; Fu, J.; Friedrich, A.; Zhang, Y.; Stewart, A. F., Single-stranded heteroduplex intermediates in lambda Red homologous recombination. *BMC Mol Biol* **2010**, *11*, 54.
237. Strochlic, T. I.; Viaud, J.; Rennefahrt, U. E.; Anastassiadis, T.; Peterson, J. R., Phosphoinositides are essential coactivators for p21-activated kinase 1. *Mol Cell* **2010**, *40* (3), 493-500.
238. Myronovskyi, M.; Tokovenko, B.; Brotz, E.; Ruckert, C.; Kalinowski, J.; Luzhetskyy, A., Genome rearrangements of *Streptomyces albus* J1074 lead to the carotenoid gene cluster activation. *Appl Microbiol Biotechnol* **2014**, *98* (2), 795-806.
239. Blodgett, J. A.; Thomas, P. M.; Li, G.; Velasquez, J. E.; van der Donk, W. A.; Kelleher, N. L.; Metcalf, W. W., Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide. *Nat Chem Biol* **2007**, *3* (8), 480-5.
240. Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Rao, R. N.; Schoner, B. E., Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene* **1992**, *116* (1), 43-9.
241. Petzke, L.; Luzhetskyy, A.; Bechthold, A., Transgenese in Streptomycceten: Transposons, Rekombinasen und Meganukleasen. **2010**.

6. Appendix

6.1. Section 2

Table S 1: NMR Data 1E5_CMP1:

Pos.	δ_c	δ_H , m (J in Hz)	HMBC
1 & 1'	129,6	7,22 m	-
2 & 2'	128,8	7,33 m	-
3 & 3'	127,6	7,28 m	-
4 & 4'	120,8	5,90 s	-
5 & 5'	39,7	3,71 s	-
6	18,75	3,48 s	-
OH ₁ & OH ₁ '	-	10,69 s	-

Table S 2: NMR Data 1E5_CMP2.

Pos.	δ_c	δ_H , m (J in Hz)	HMBC
1 & 1'	170,29	-	-
2 & 2'	168,58	-	-
3 & 3'	163,33	-	-
4 & 4'	134,5	-	-
5 & 5'	129,23	7,22 m	C-10
6 & 6'	128,86	7,33 m	C-4
7 & 7'	127,44	7,28 m	C-5
8 & 8'	104,95	-	-
9 & 9'	102,83	5,86 m	C-10, C-8, C-3, C-2
10 & 10'	39,8	3,76 d (7,1 Hz)	C-9, C-5, C-4, C-3
11	31,42	4,13 t (7,6 Hz)	C-13, C-12, C-8, C-2, C-1
12	30,13	2,16 m	C-14, C-13, C-11, C-8
13	21,83	1,25 m	C-14, C-12, C-11
14	13,74	0,90 t (7,4 Hz)	C-13, C-12
OH ₁	-	11,37 s	-
OH ₁ '	-	10,55 s (br)	-

Table S 3: Oligonucleotides used in Section 2; all Oligonucleotides were produced by Eurofins (Eurofins Genomics Germany GmbH, Ebersberg, Germany).

Name	Sequence of the Oligonucleotide
f3E7test	GTCGTTGACCCTCAGCAAGT
r3E7test	GCCTACCACTTCTACGCGTT
f1G5test	GTAGGCGAACAACTGCAACG
r1G5test	CTTCATGCGCAAGGAGATGC
f1C15test	ACCCGATCGGCCTCAAGATC
r1C15test	AGATCACGATCATCGCGGTC
f3K5test	GCGAAGTACGACACGACGTA
r3K5test	GGGGTGTCCGGTTCTACAAG
f1F6tes	ATGCTCAGGATCGAATCGCC
r1F6test	ACAGCTACGCCTTCGACTTC
f1C7test	GAATCATCGCCGACAGCAAC
r1C7test	GTTCTCGTCGATCTTGCCGA
f3E19test	AGCCTGGTCAACCACTGTTC
r3E19test	AGCCCAACAGTTGATAGCCC
f1G11test	CCTTCGAGATCTCCGGGTTG
r1G11test	GACTCGGGAACTCCTTACGC
f3C18test	GAGTACCACATCGCCACCAA
r3C18test	AGAGGATCTCCAGGTCGTCTG
f1L8test	CCACTGACGTGACGCATACT
r1L8test	GTGAGCAGTGAGGTGTGGTT
f3M21test	AGACCGTCGAGCTGAACAAG
r3M21test	CACACCGTCTGGAGGATCTG
f1E5test	AAGCTGTGGTTGACTCCGG
r1E5test	GCTTTCGGGGTCAATTGTCTG
f3A24test	AGTTCGTGGAAGCGCTGGAC
r3A24test	AAGTCTCCAGGACCTTCACC
f3K17test	GAAGACGTGATGCTCCACA
r3K17test	ATCTTCGCGCGCTACAACAT
f-del-54860	CACCACGCCCCGGCACGACGACCCGGCCACCGCTCACGGC TTCCGGGGATCCGTCGACCC
r-del-54860	GGGCCGCGCCAGCACCATGGCGCTCTGGAAGCCGCCGAAAC TGTAGGCTGGAGCTGCTTCG
f-del-55110	ATGGAGCCCATCGCGATCATCGGTGTGGGGTCCCGGTTCCC TTCCGGGGATCCGTCGACCC
r-del-55110	CGCGACCTCGCGGGTGGTTCGTACGACGACGACGCGGCA TGTAGGCTGGAGCTGCTTCG
f-chk-54860	GCCTGCTTCGACGCGATCAA
r-chk-54860	GATCACGCCATGTCTGAACG
f-chk-55110	GATCACGCTCTCCAGACG
r-chk-55110	GACGAGTCTTCATCGCAA
f-del-54730	GTGACCACCAAGCCCCGCTCCACGATCGACCTCTATTCCGACGAATTTCCGGGGATCCGTCGACCC
r-del-54730	TCAGCGGTGCAGGGTGGTCCGGGACGCGCGGATGCCGCGGATGGTGTGTAGGCTGGAGCTGCTTCG
f-del-54740	ATGAGGCTGATCGTCGATCGCGACCGCTGCGAAGGGCACGGCGTGTGCGTTCGGGGATCCGTCGACCC
r-del-54740	TCAGCCTTCAGCTTGAGGGCGCGGACCGGACAGATGAGCACCGAGTGTAGGCTGGAGCTGCTTCG
f-del-54750	ATGGAGAGCACGCGCGAACGGTGTTCGCGCCGCGCTCGGCCGACCATTCGGGGATCCGTCGACCC
r-del-54750	CTAGGACACCCTGGTCCGGGCCGCGCCGCTTGAGCAGCCGGCCAGCTGTAGGCTGGAGCTGCTTCG
f-del-54760	ATGTCAGAGGAGAGCACCGAGAACACGACCGCCATGCGGTTGCTGCGGCTTCGGGGATCCGTCGACCC
r-del-54760	CTAGGCGCGCACTGCTTCTACGACGCTGCACAGCCCGCGGGGACGATCTGTAGGCTGGAGCTGCTTCG
f-del-54770+80	ATGCCGGTAATCGCGGCAACGACGTTGTCTGACCGTCTTCAACATGTTCCGGGGATCCGTCGACCC
r-del-54770+80	TCAGTGCGGCGGGTTCGACCCAGCCAGTTGGACGAGGATGCCATGTCTGTAGGCTGGAGCTGCTTCG
f-del-54790	ATGAGCACACCCCGTCACTGGTCTGCTCGGCGGCTCCTTGCCGGTCTTCGGGGATCCGTCGACCC
r-del-54790	TCAGTCGCGGTGCACCGACGTGACCGCGTCCGACAGGTACAGCCCCGGTGTAGGCTGGAGCTGCTTCG
f-del-54800	ATGATGGAAGGCGAACTGGGGGCTTCTGCGCAGTCGCCGTGAAGCCGTTCCGGGGATCCGTCGACCC
r-del-54800	TCAGCCGGAGCCGACCGCGCAACCCACCTGGTTACGTTACGCGGTGTAGGCTGGAGCTGCTTCG
f-del-54820	TCAGCCGAGGCGCTGTCCAATCGGGACAGTTCCGCGCAGTCAGCTCCGGGGATCCGTCGACCC
r-del-54820	ATGCGTCCGACTCCGCTGGGACCGCCGCTGGGACACCGGGTTGATGTAGGCTGGAGCTGCTTCG
f-del-54830	TCAGGCGGGGTGCGGCGACCCGGTTCGACAGATCAGGTGGAACGACAGCTTCGGGGATCCGTCGACCC
r-del-54830	ATGACAGCCGAGTCCGACTGCCGTCCCGTCCGACGTTGGGCGGCTGCTGTAGGCTGGAGCTGCTTCG
f-del-54930	ATCAGAGGACGGCCCGACCCGAGGCCACCGAGCGCCAGCACGCAAGGTTCCGGGGATCCGTCGACCC
r-del-54930	TCATTGCGGCGGGTTCGGTGTCTGCGGCTGGGCGGTCGCGTGTGTAGGCTGGAGCTGCTTCG
f-del-54940	GTGCAAGGTGCATTCGTGGAGTTGCGATGCGTTGCCGACCAATACCGGTTCCGGGGATCCGTCGACCC
r-del-54940	CTATTGCCAACTCGACGGGCTCGGTTAGGCCGTAACCGGTTCCGCGCGGTGTAGGCTGGAGCTGCTTCG

f-del-54950	GTGGTAGGGCAACAAGCCGGTGAGTTAGTCAGCACATGGGAGCCCGCGCTTCCGGGGATCCGTCGACCC
f-del-54950	TCAGCCCTCGCCGAGGCGGAAACCGACACCCGGAACGGTGACGATCCACTGTAGGCTGGAGCTGCTTCG
f-del-1E5-left	GTGGTGCCGCTCTCCCTCATAGTTTTGCCAACGACGGCGTCAGTGCCTTCCGGGGATCCGTCGACCC
r-del-1E5-left	CTGGCAGTGCCGGTCTCAGCTCGGCCCTCCCTACCGGCGACGGCTGCCATGTAGGCTGGAGCTGCTTCG
f-del-1E5-right	TCACCACGCTCCAAGGCGGAGCCGGGACGTGCCCGCCGCCACGGCCGCTTCCGGGGATCCGTCGACCC
r-del-1E5-right	TCACGTACGGTGAGGCGGATGGGTATGCCGAGGCTGGGGTTGCCCGGATGTAGGCTGGAGCTGCTTCG
f-del-54730-chk	TGAACTTCGGGACCTTGACG
r-del-54730-chk	CGATCGACGATCAGCCTCAT
f-del-54740-chk	GGTACAGCCGACCGACAC
r-del-54740-chk	ATCAGCTCGATGGTGCGG
f-del-54750-chk	ATGAGGCTGATCGTCGATCG
r-del-54750-chk	GTTCTCGGTGCTCTCCTCTG
f-del-54760-chk	CGCTGACGGTGATAAGCA
r-del-54760-chk	TCCCGCATCTCTTGAGCAG
f-del-54770+80-chk	TGGACCTCGACATGTTGGTC
r-del-54770+80-chk	TCCTCTCTGGGTCTGTTC
f-del-54790-chk	GAACAGGACCCAGGAGAGGA
r-del-54790-chk	GTTTCGAGACGATGCAACTGG
f-del-54800-chk	TCCCTACTTCTGGTCGGAC
r-del-54800-chk	GTGCTCGGTGATGCTTCCC
f-del-54820-chk	CATCGGCGTTGTCGAAAC
r-del-54820-chk	GTCGTTCCACCTGATCTGGT
f-del-54830-chk	GGTCGTGCAGGTAGTACAGG
r-del-54830-chk	GAGTGGGAGTTCTGCGACAA
f-del-54930-chk	GAAGGAGCAGTGCGTGAA
r-del-54930-chk	CTTCTCCACCAGAACCTCCG
f-del-54940-chk	TGGTCCGGCAGTACAAAGAC
r-del-54940-chk	TTTTCCGGTGTCATGACGGT
f-del-54950-chk	TGCGGTTCACTCTTGAGC
r-del-54950-chk	ATGTGGATGACTGCCCGTG
f-chk-1E5-le-int	GATCGCCATTGTGTCCGTTG
r-chk-1E5-le-int	CATCTCGATCACTCCACGG
f-chk-1E5-ri-int	TCACAAGGCACCTTACGGAC
r-chk-1E5-ri-int	CATCAGCGAAATCACTGCCG
f-chk-1E5_ri	CTCCGGGAAGTGATACTCGC
r-chk-1E5_ri	GACGTTCTCCTCAGGTGACG
f-chk-1E5_le	CACTCGACGTCCAAGTCA
r-chk-1E5_le	GAGTGCCCGGATATCTGAA
F-54800A3OE	AAAAAAGGTACCTAGCAGGGCTCCAAAACAAACGCCTGATGTAGGATCAGATGAAAAAAAAAAGGAGGA
R-54800A3OE	AAATACATATGATGGAAGGCGAACTGGG
F-54950A3OE	AAAAAAAAAGCTTTTCAGCCGGAGCCGACCCGCGCAACCCACCTGGTTCAGTT
R-54950A3OE	AAAAAAGGTACCTAGCAGGGCTCCAAAACAAACGCCTGATGTAGGATCAGATGAAAAAAAAAAGGAGGA
	AAATACATGTGGTAGGGCAACAAGCC
	AAAAAAAAAGCTTTGTGGGTGTTCACTCCGGA

Table S 4: Strains, Plasmids and BACs used in Section 2.

Streptomyces	Characteristics	Reference
<i>S. albus</i> J1074	<i>S. albus</i> G1 (DSM 41398) derivative with the defective <i>SalG I</i> restriction modification system heterologous host	(Chater and Wilde, 1980 ¹²³)
<i>Streptomyces lividans</i> Δ6	<i>S. lividans</i> TK24 derivative, deletion of 6 BGCs	(Paper in Revision)
<i>Saccharothrix espanaensis</i> DSM 44229 (T)	Producer of Saccharomicins	Labeda <i>et al.</i> ¹¹¹
<i>E. coli</i>	Characteristics	Reference
GB2005	General cloning host	Maresca <i>et al.</i> , 2010 ²³⁶
ET12567 pUZ8002	Strain used for intergeneric conjugation,	Kieser 2000 ¹⁹⁰
GB2005-red-rham	GB2005, RhamC-BAD- γ baA used for Red/ET	Strochlic <i>et al.</i> , 2010 ²³⁷
Plasmids	Characteristics	Reference
psmart-BAC-S	Apr ^R ; BAC Vector	Lucigen
pTOS	Apr ^R ; VWB-based <i>Streptomyces</i> integrative vector	Herrmann <i>et al.</i> (2012) ¹⁸⁸
pTOS_A3_R1	pTOS derivative; Integrated Gene 54800, influenced by A3 promotor cloned into the Kpn I and Hind III site	This work
pTOS_A3_R2	pTOS derivative; Integrated Gene 54950, influenced by A3 promotor cloned into the Kpn I and Hind III site	This work
patt-shyg	<i>Swa I/EcoR V</i> fragment of phyg-OK containing hyg cloned into the <i>EcoR V</i> site of patt	Myronovskyi <i>et al.</i> ¹⁹²
BACs	Characteristics	Reference
pSMART-BAC-S-1C15	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 1,670,099-1,769,322)	Intact genomics
pSMART-BAC-S-3A24	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 2,190,390-2,280,022)	Intact genomics
pSMART-BAC-S-3E7	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 2,332,949-2,404,154)	Intact genomics
pSMART-BAC-S-1G5	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 2,825,657-2,924,398)	Intact genomics
pSMART-BAC-S-1F6	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 3,429,623-3,516,631)	Intact genomics
pSMART-BAC-S-1I20	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 3,577,211-3,660,078)	Intact genomics
pSMART-BAC-S-3E19	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 3,526,632-3,686,841)	Intact genomics
pSMART-BAC-S-3K5	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 3,786,071-3,893,910)	Intact genomics
pSMART-BAC-S-1G11	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 4,412,386-4,533,983)	Intact genomics
pSMART-BAC-S-3C18	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 4,412,386-4,533,983)	Intact genomics
pSMART-BAC-S-1L8	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 5,681,922-5,790,602)	Intact genomics
pSMART-BAC-S-3M21	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 5,865,637-5,963,104)	Intact genomics
pSMART-BAC-S-1E5	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 6,074,530-6,190,776)	Intact genomics
pSMART-BAC-S-1C7	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 6,252,700-6,375,967)	Intact genomics
pSMART-BAC-S-3K17	pSMART-BAC-S derivative containing a fragment of the <i>S. espanaensis</i> chromosome (Locus in bp: 6,511,399-6,617,883)	Intact genomics

1E5Δ55110	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_55110	This work
1E5Δ54860	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54860	This work
1E5ΔpenA	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54730	This work
1E5ΔpenC	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54750	This work
1E5ΔpenD	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54760	This work
1E5ΔpenE/F	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54770 & BN6_54780	This work
1E5ΔpenG	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54790	This work
1E5ΔpenR1	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54800	This work
1E5ΔpenI	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54820	This work
1E5ΔpenJ	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54830	This work
1E5ΔpenV	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54940	This work
1E5ΔpenR2	pSMART-BAC-S-1E5 derivative; Knockout of gene BN6_54950	This work
1E5Δleft	pSMART-BAC-S-1E5 derivative; Deletion of left flanking region (Deletion locus in bp: 6,079,768 - 6,123,914; Deletion of 44,147 bp)	This work
1E5Δright	pSMART-BAC-S-1E5 derivative; Deletion of right flanking region (Deletion locus in bp: 6,148,356 - 6,184,748; Deletion of 36,393 bp)	This work

Table S 5: Preparative HPLC method for pentangumycin (CMP468) & SEK90 (CMP791); 20 mL/min flow rate.

Method pentangumycin			Method SEK90			Waters AutoPurification System
Min.	%A	%B	Min.	%A	%B	Waters 2545 Binary Gradient module
0	95	5	0	95	5	Waters SFO (System Fluidics organizer)
2	95	5	2	95	5	Waters 2998 PAD (Photodiode array Detector)
16	3	97	16	3	97	Waters SQ-Detector-2
17	3	97	20	3	97	Waters 2767 Sample Manager
18	95	5	21	95	5	Nucleodur C18 Htec 250/4,6 C18 5 μ M (analytical)
19	95	5	22	95	5	Nucleodur C18 Htec 250/21 C18 5 μ M (preparative)

Table S 6: Preparative HPLC method and yield for 1E5CMP1 - 1E5CMP2; Flow Rate: 20mL/min.

Method 1E5_CMP1-2			Waters AutoPurification System	Purified compounds
Min.	%A	%B	Waters 2545 Binary Gradient module	1E5CMP1: 6 mg
0	95	5	Waters SFO (System Fluidics organizer)	1E5CMP1: 7 mg
1	95	5	Waters 2998 PAD (Photodiode array Detector)	
5	50	50	Waters 2767 Sample Manager	
15	5	95	Nucleodur C18 Htec 250/4,6 C18 5 μ M (analytical)	
19	5	95	Nucleodur C18 Htec 250/21 C18 5 μ M (preparative)	
20	95	5		
21	95	5		

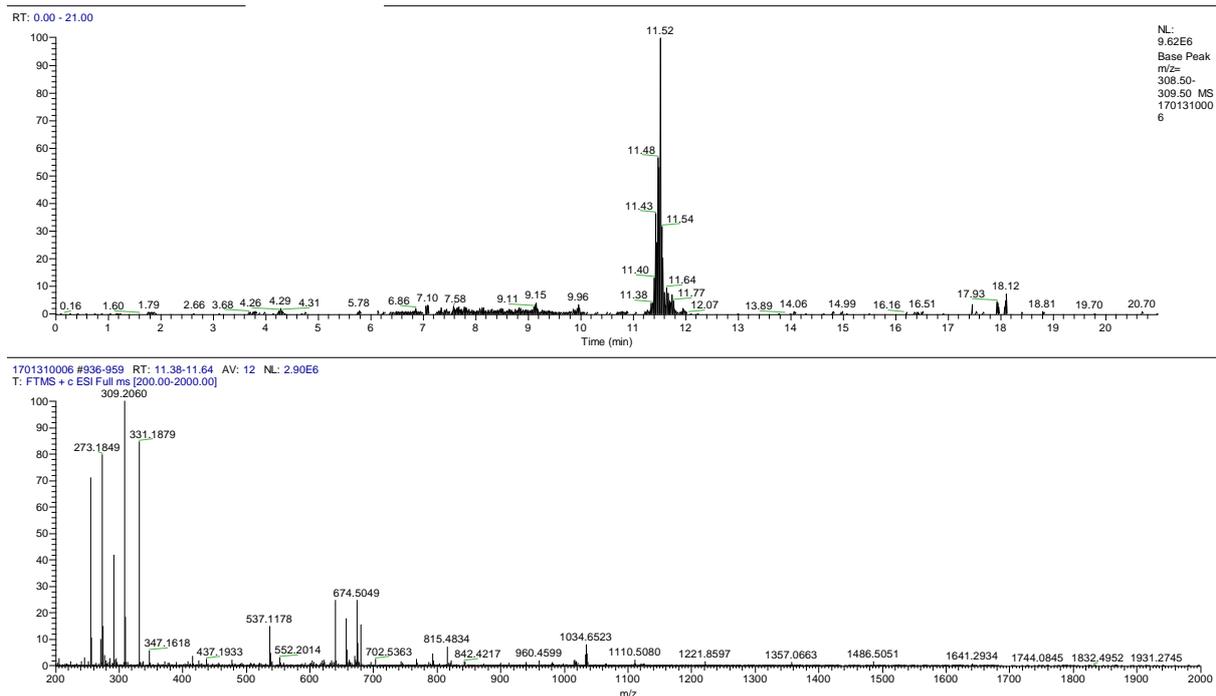


Figure S 1: HPLC-MS Extracted ion chromatogram (Extracted mass 309 ± 0.5) of *S. lividans* Δ YA6_3C18 and its ESI full ms chromatogram.

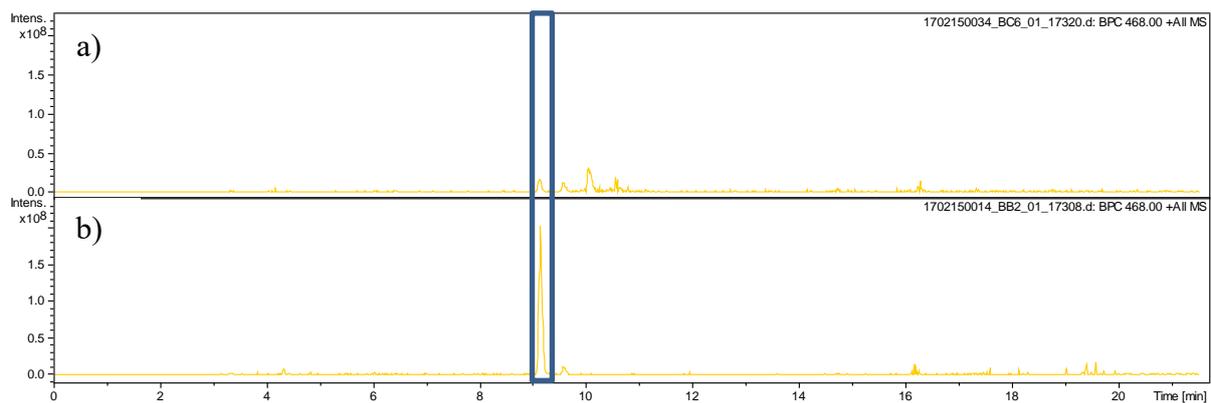


Figure S 2: HPLC-MS Extracted ion chromatogram (Extracted mass 468 ± 0.5) of a) *S. albus* J1074_1E5 and b) *S. lividans* $\Delta 6_1E5$.

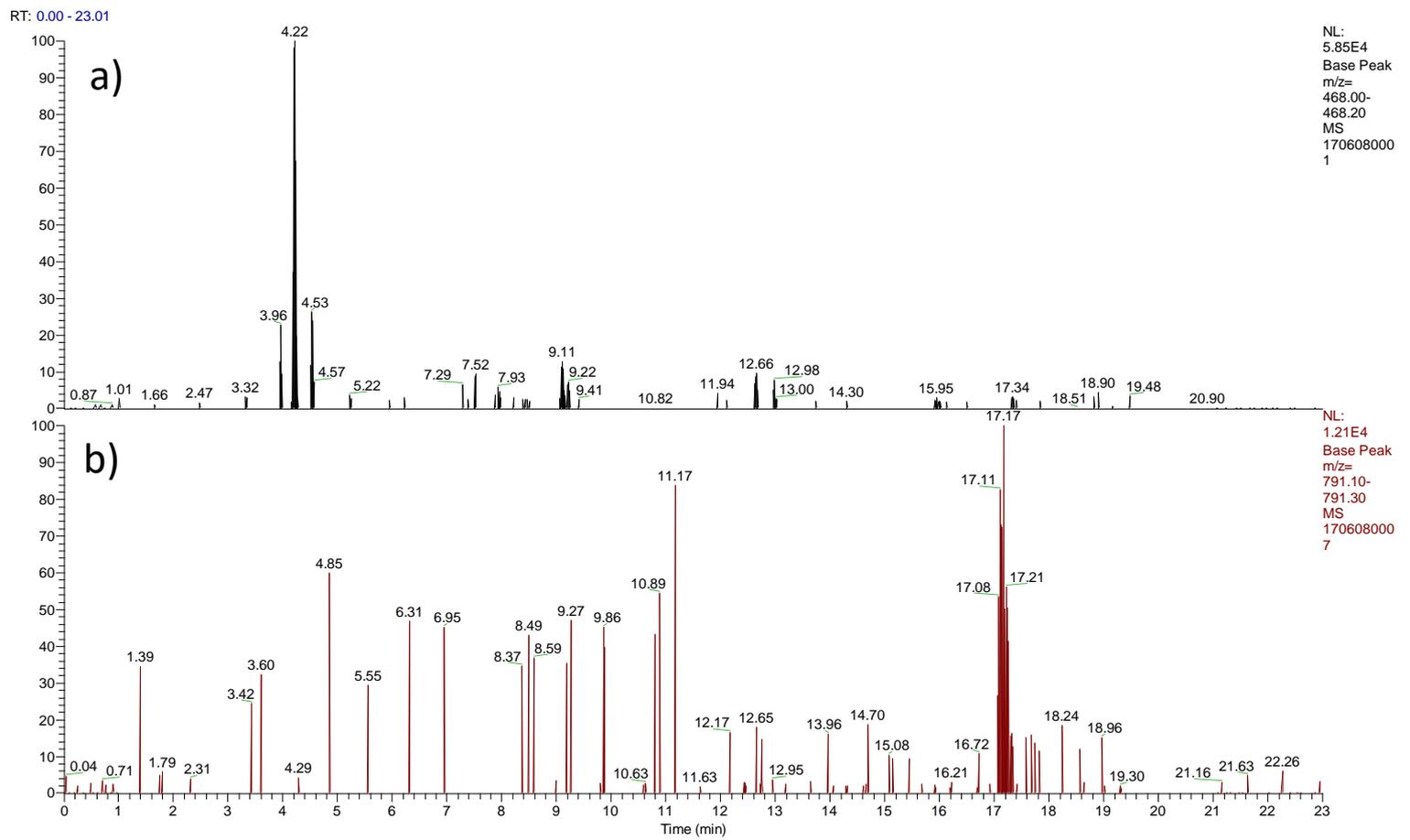


Figure S 3: HPLC-MS Extracted ion chromatogram; a) Extracted mass (468.00 – 468.20) BPC of an ethyl acetate extract of *S. espanaensis* grown in SG Medium; b) Extracted mass (791.10 – 791.30) BPC of a butanol extract of *S. espanaensis* grown in SG Medium.

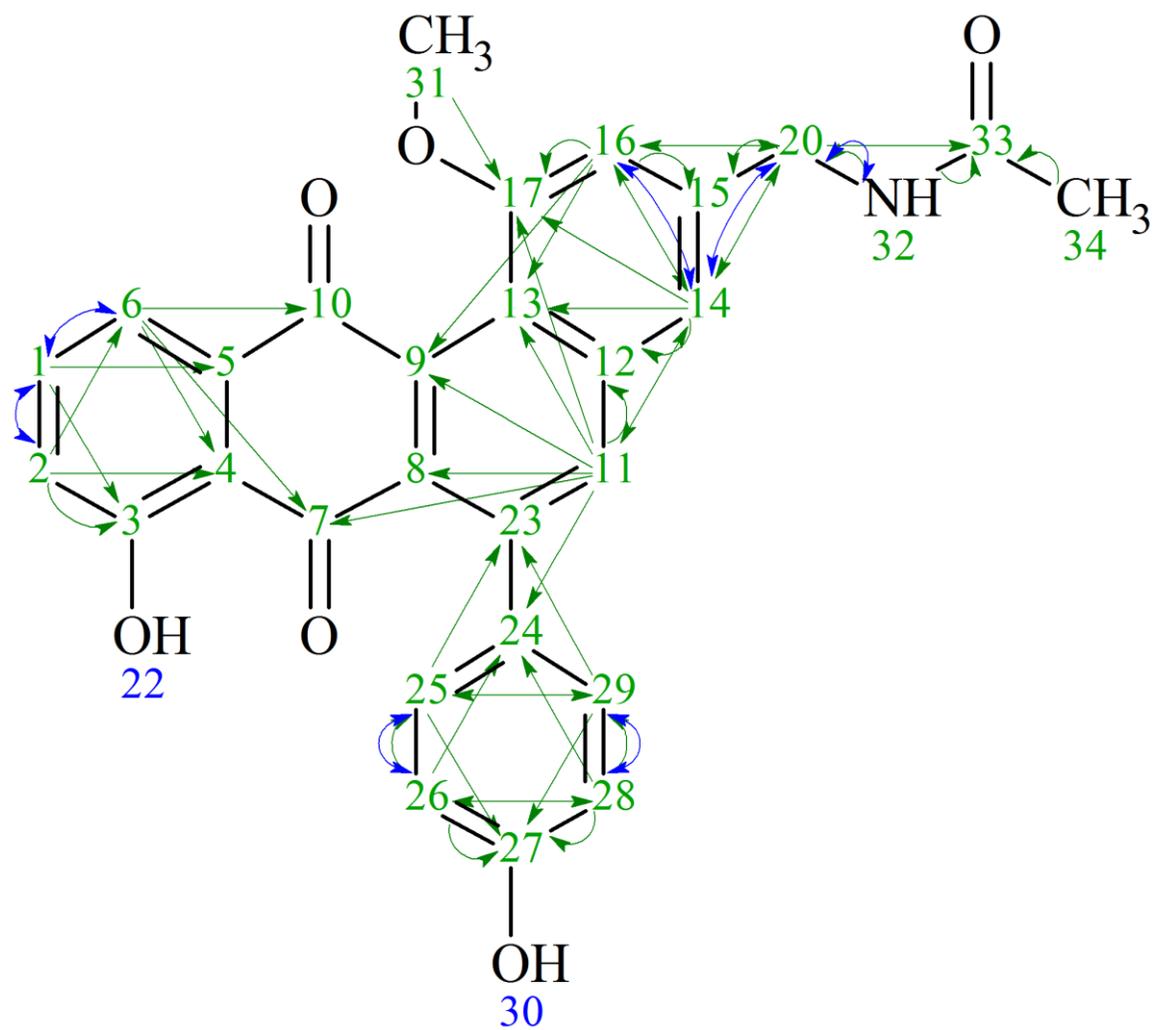


Figure S 4: Structure of pentangumycin with all observed correlations (green: HMBC correlations H → C; blue: ¹H-¹H-Cosy correlation H → H).

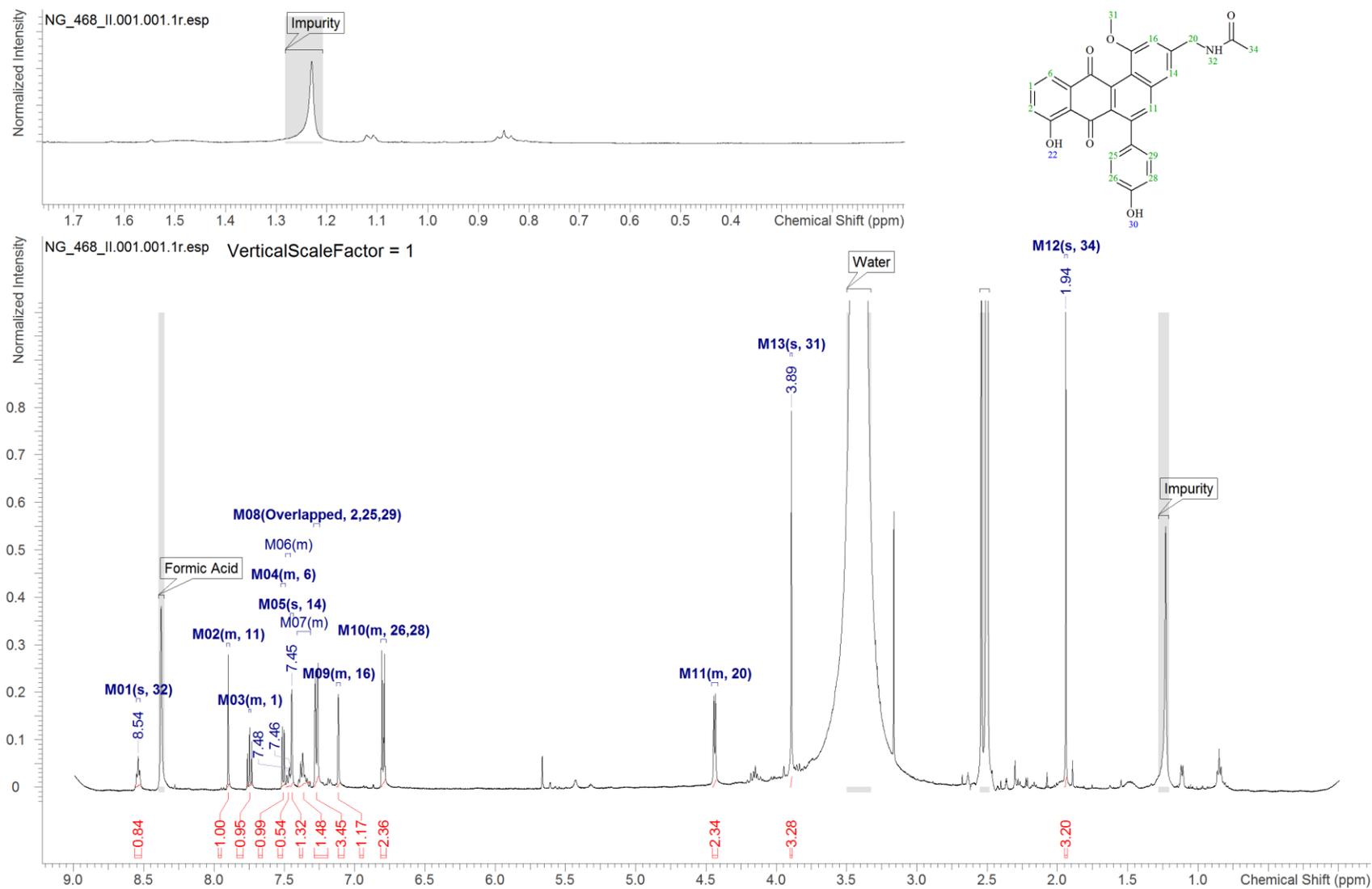


Figure S 5: $^1\text{H-NMR}$ spectrum (500 MHz, DMSO-d_6) of pentangumycin; complete Spectrum and zoom from 1.7 to 0.1 ppm.

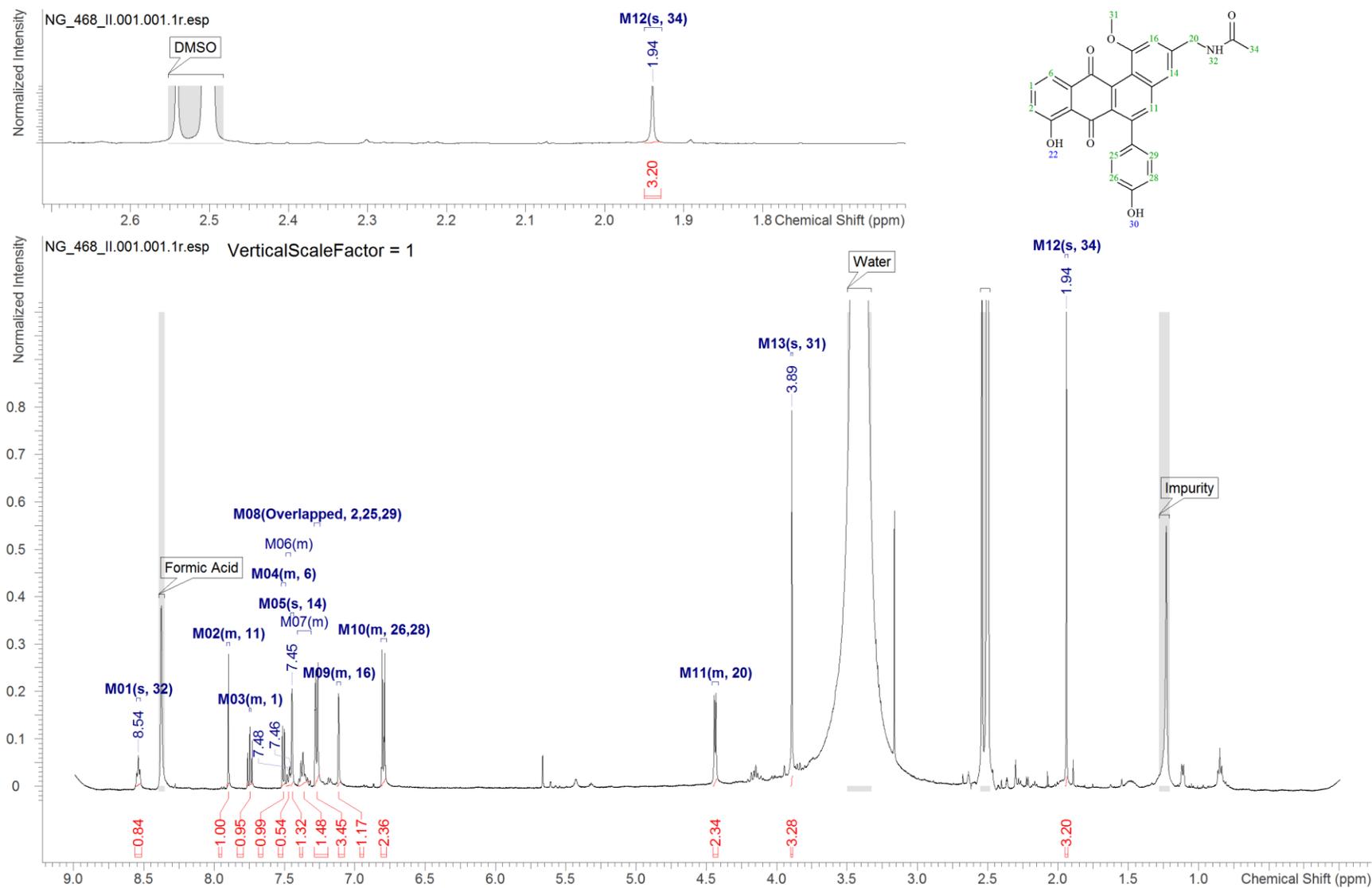


Figure S 6: $^1\text{H-NMR}$ spectrum (500 MHz, DMSO-d_6) of pentangumycin; complete Spectrum and zoom from 2.7 to 1.7 ppm.

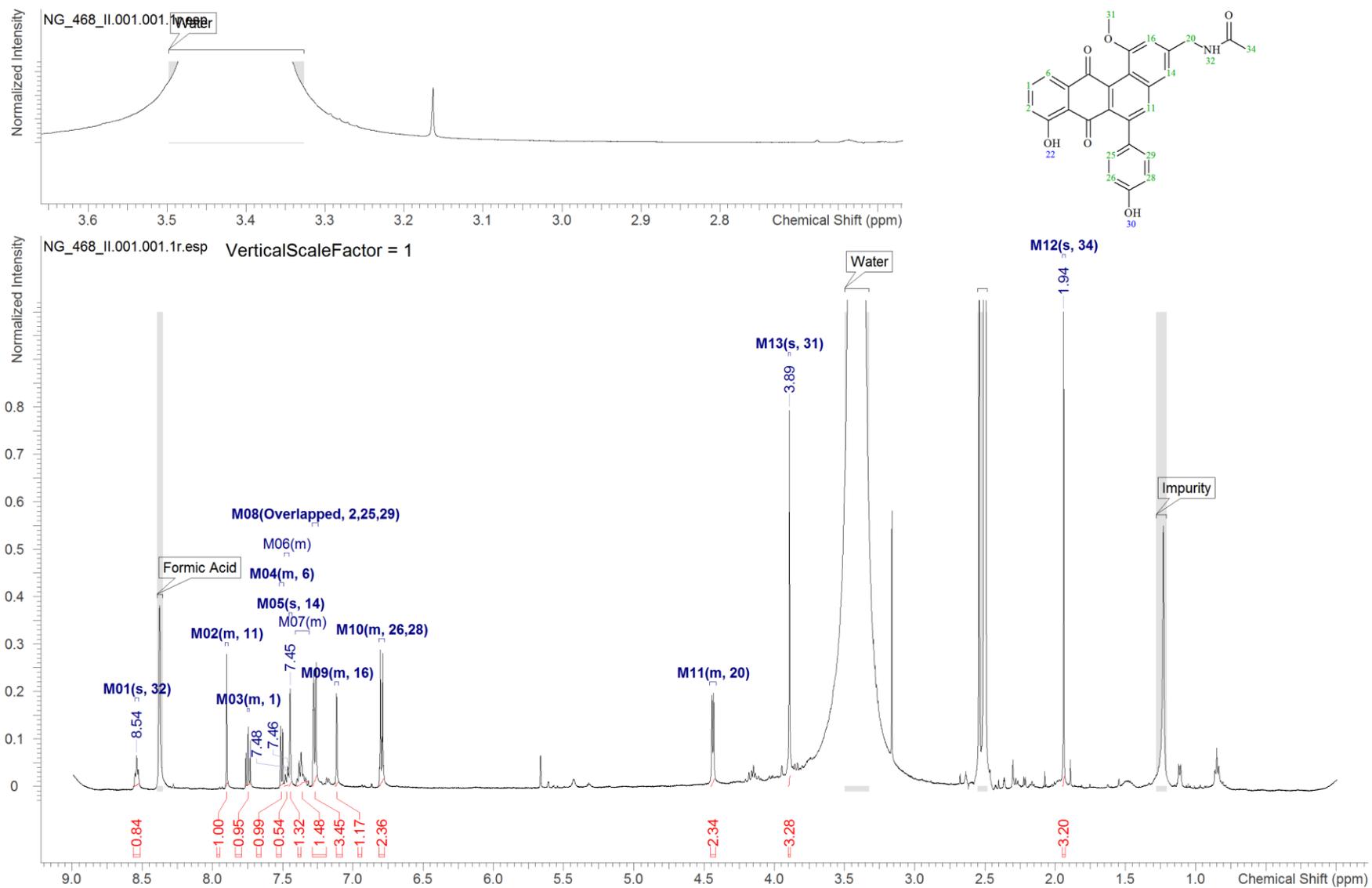


Figure S 7: $^1\text{H-NMR}$ spectrum (500 MHz, DMSO-d_6) of pentangumycin; complete Spectrum and zoom from 3.65 to 2.65 ppm.

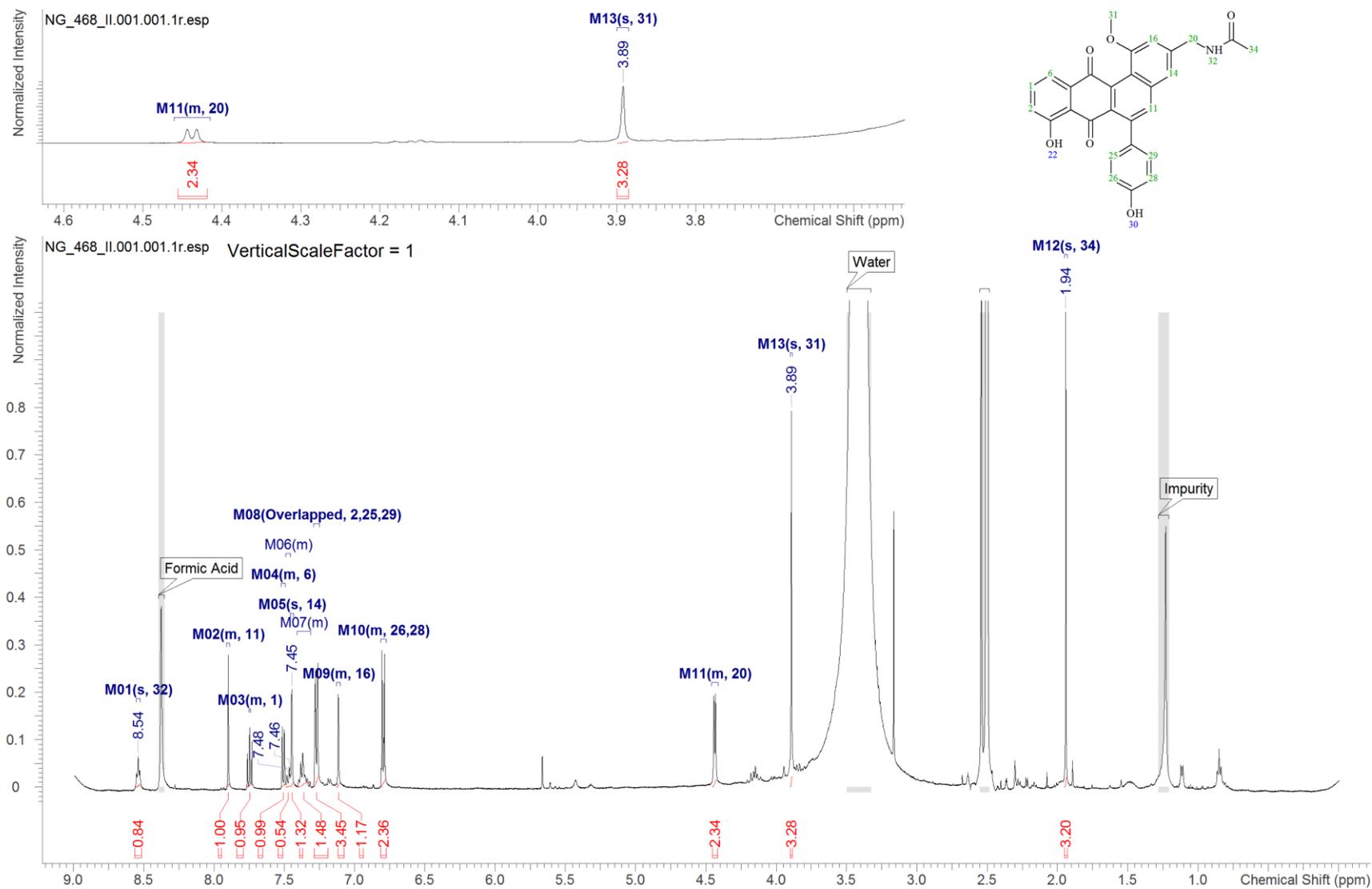


Figure S 8: ¹H-NMR spectrum (500 MHz, DMSO-d₆) of pentangumycin; complete Spectrum and zoom from 4.6 to 3.6 ppm.

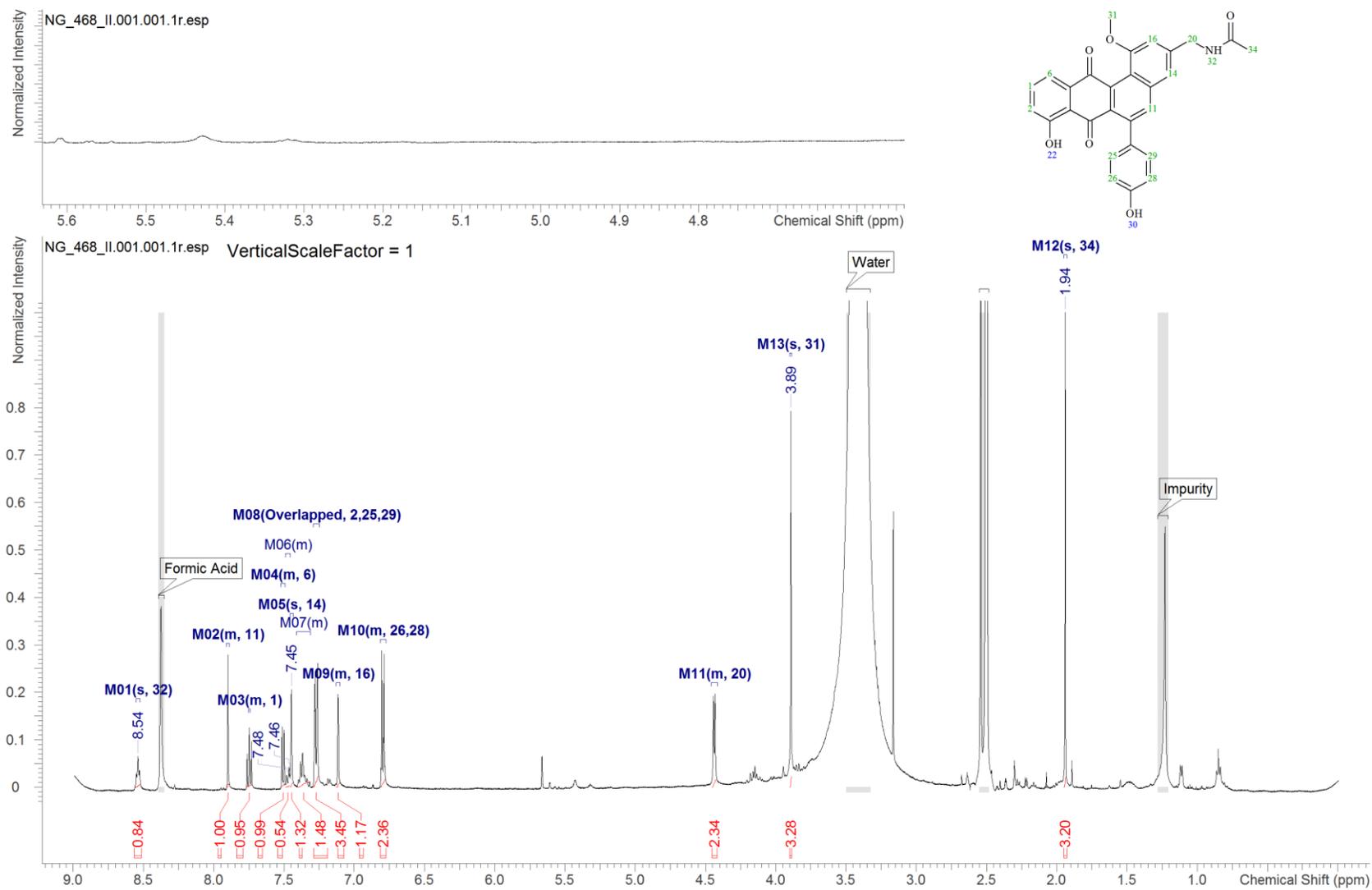


Figure S 9: $^1\text{H-NMR}$ spectrum (500 MHz, DMSO-d_6) of pentangumycin; complete Spectrum and zoom from 5.6 to 4.6 ppm.

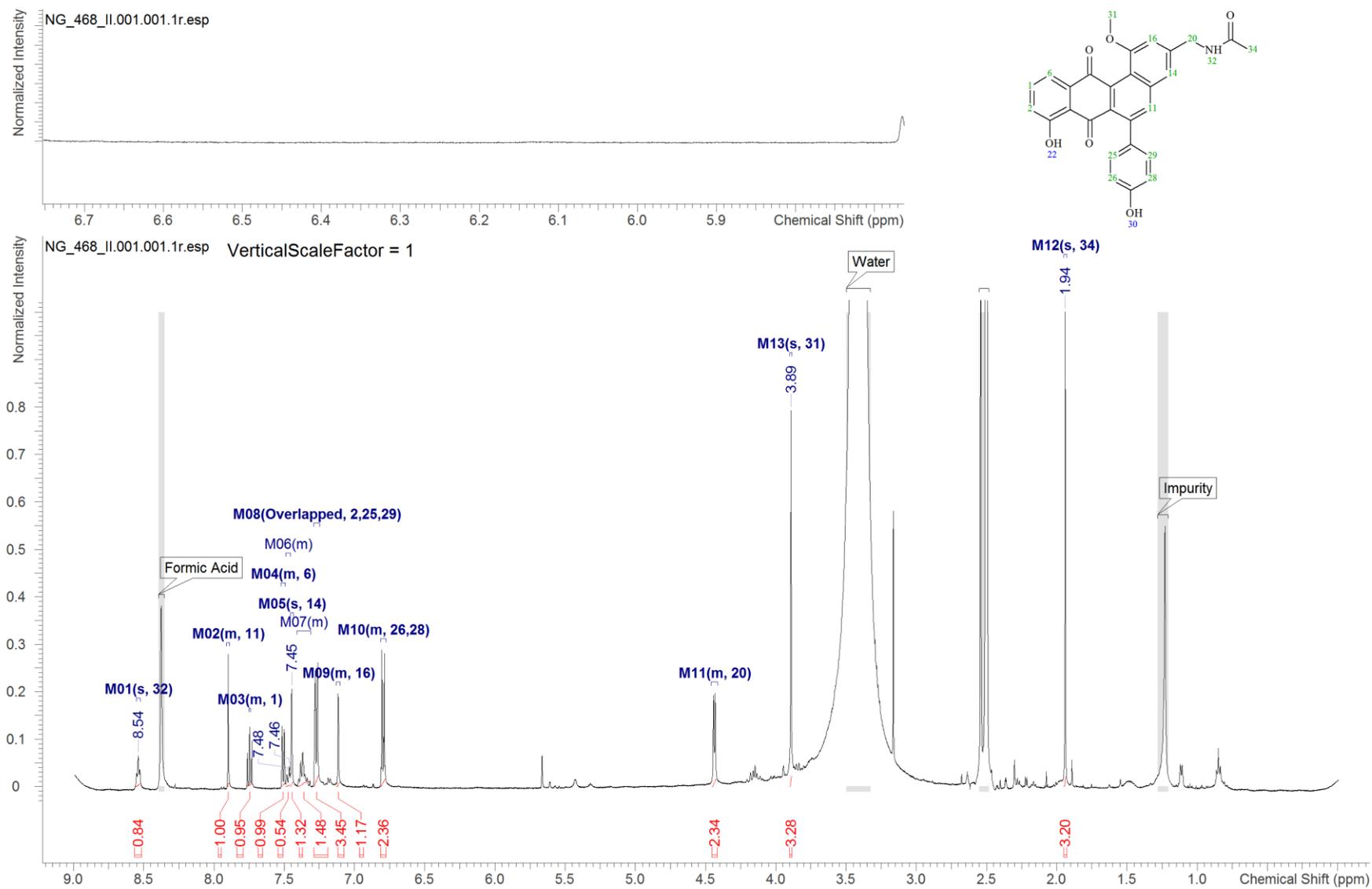


Figure S 10: $^1\text{H-NMR}$ spectrum (500 MHz, DMSO-d_6) of pentangumycin; complete Spectrum and zoom from 6.7 to 5.7 ppm.

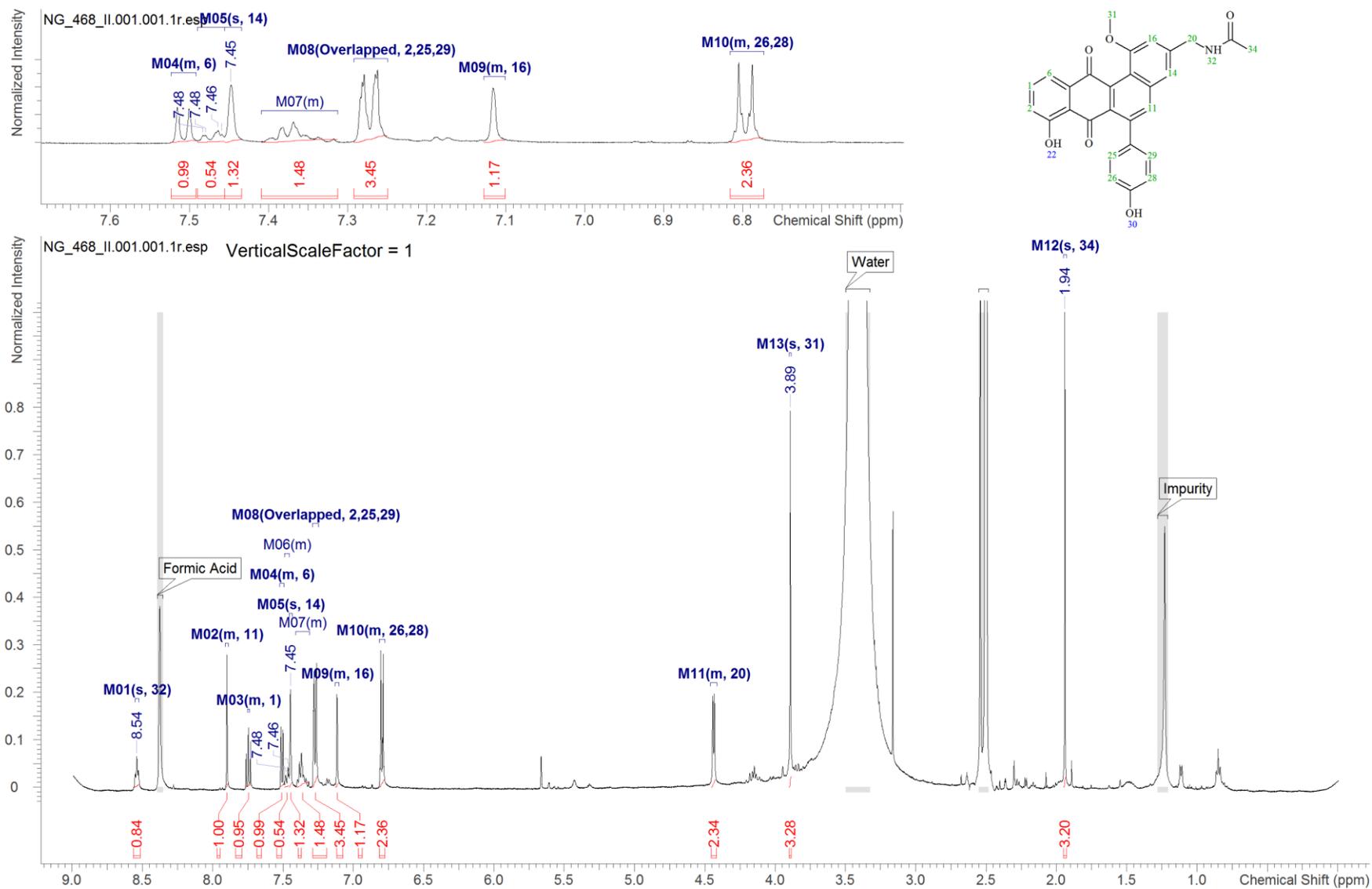


Figure S 11: $^1\text{H-NMR}$ spectrum (500 MHz, DMSO-d_6) of pentangumycin; complete Spectrum and zoom from 7.6 to 6.7 ppm.

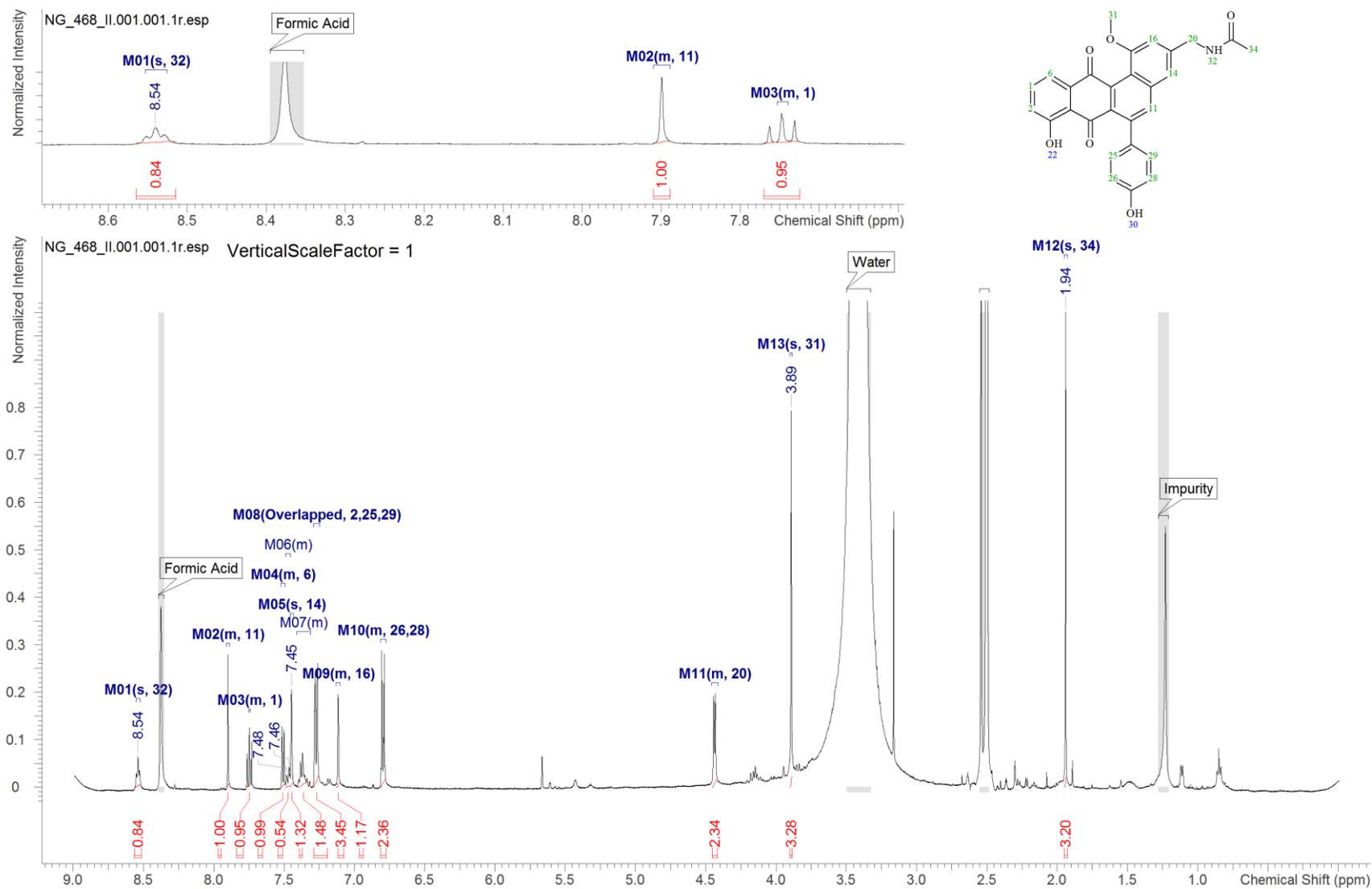


Figure S 12: $^1\text{H-NMR}$ spectrum (500 MHz, DMSO-d_6) of pentangumycin; complete Spectrum and zoom from 7.6 to 6.7 ppm.

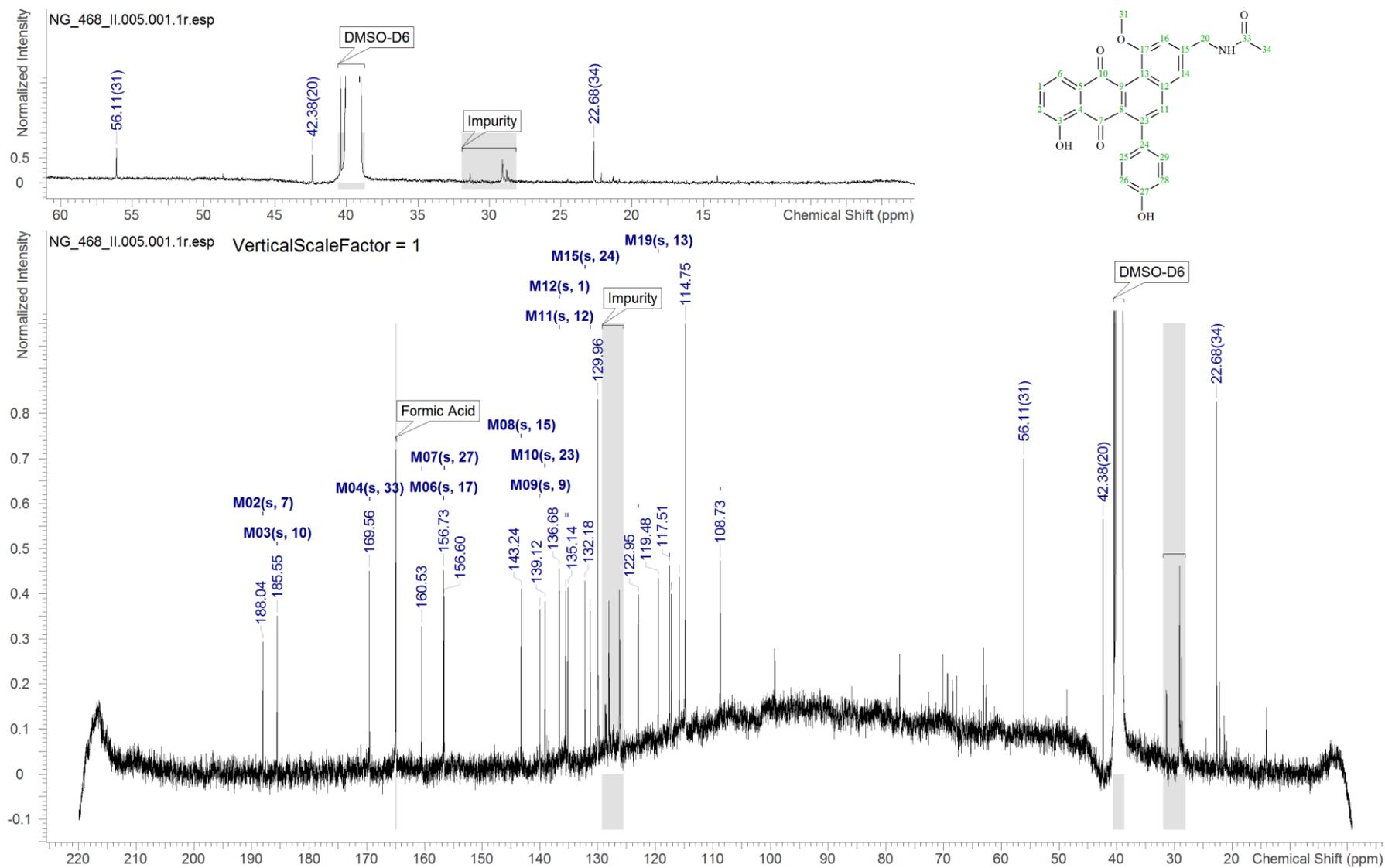


Figure S 13: ^{13}C -NMR (125 MHz, DMSO- d_6) of pentangumycin; complete spectrum and zoom from 60 to 0 ppm.

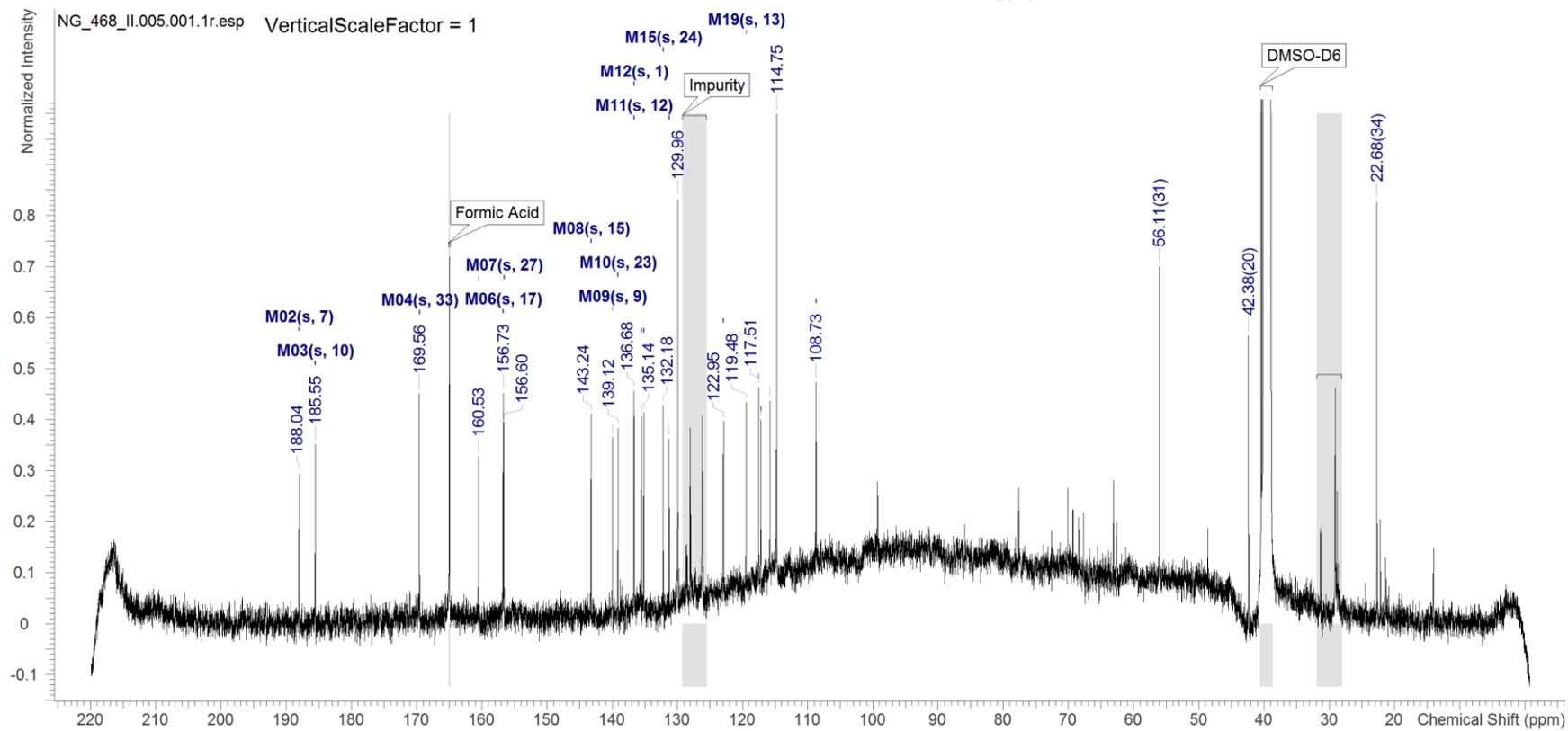
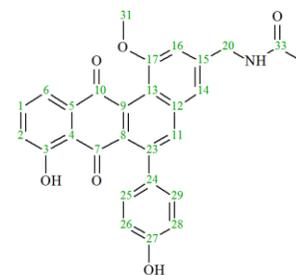
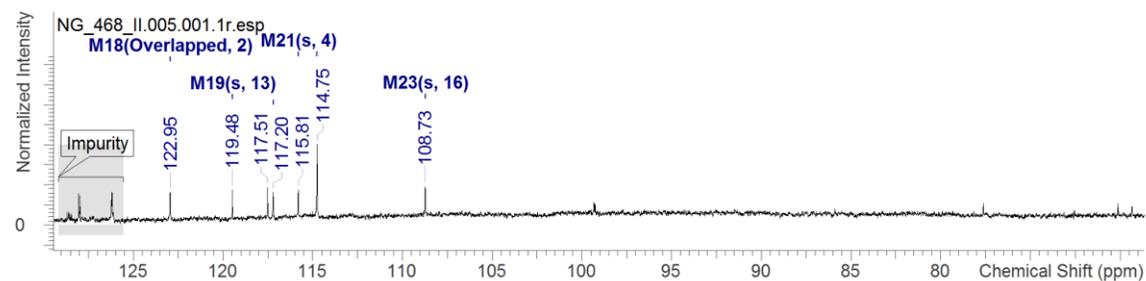


Figure S 14: ^{13}C -NMR (125 MHz, DMSO- d_6) of pentangumycin; complete spectrum and zoom from 130 to 70 ppm.

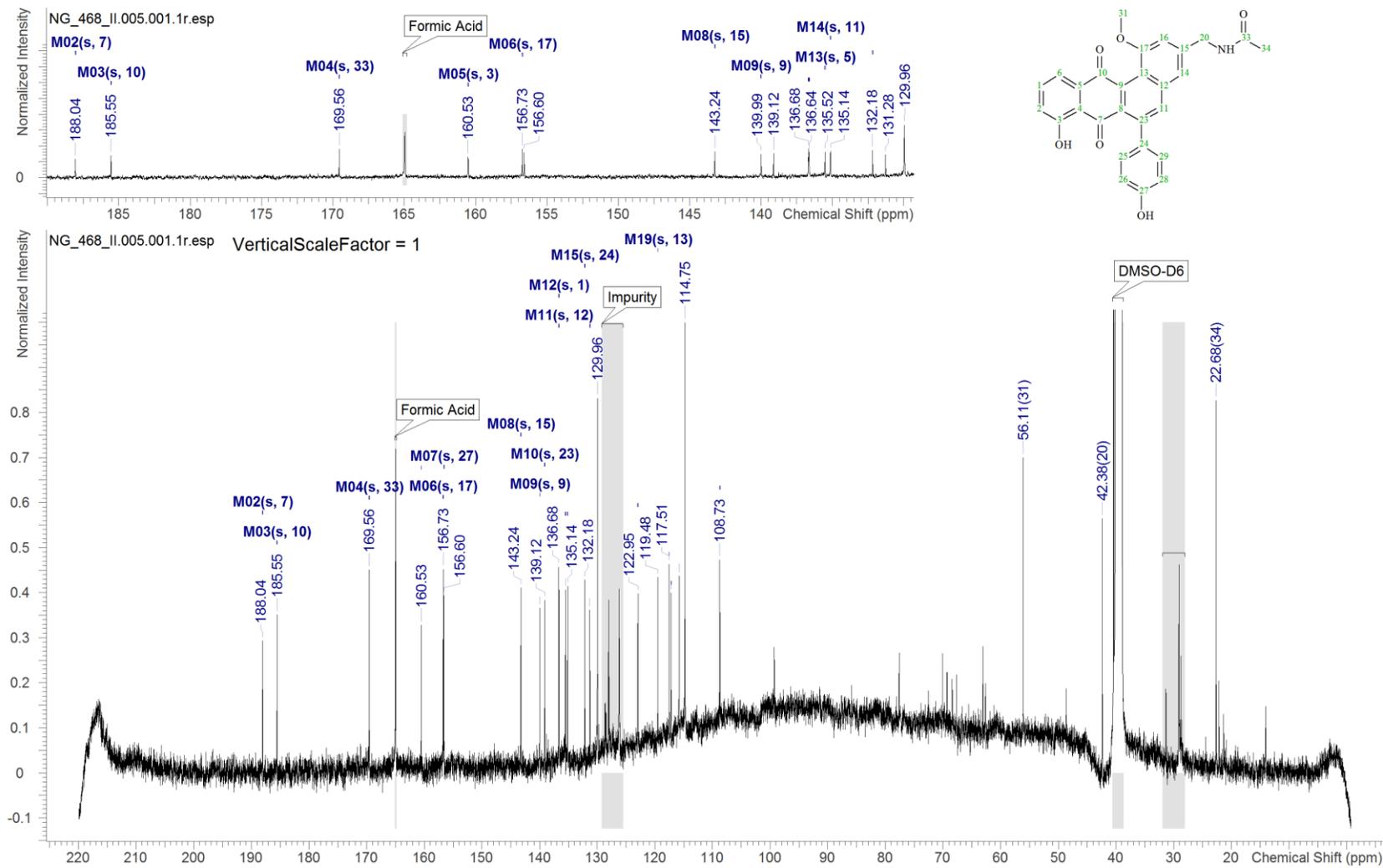
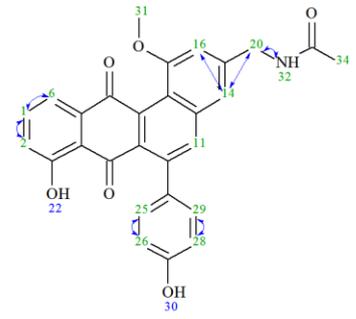


Figure S 15: ^{13}C -NMR (125 MHz, DMSO- d_6) of pentangumycin; complete spectrum and zoom from 190 to 130 ppm.



NG_468_171128_neu.002.001.2rr.esp

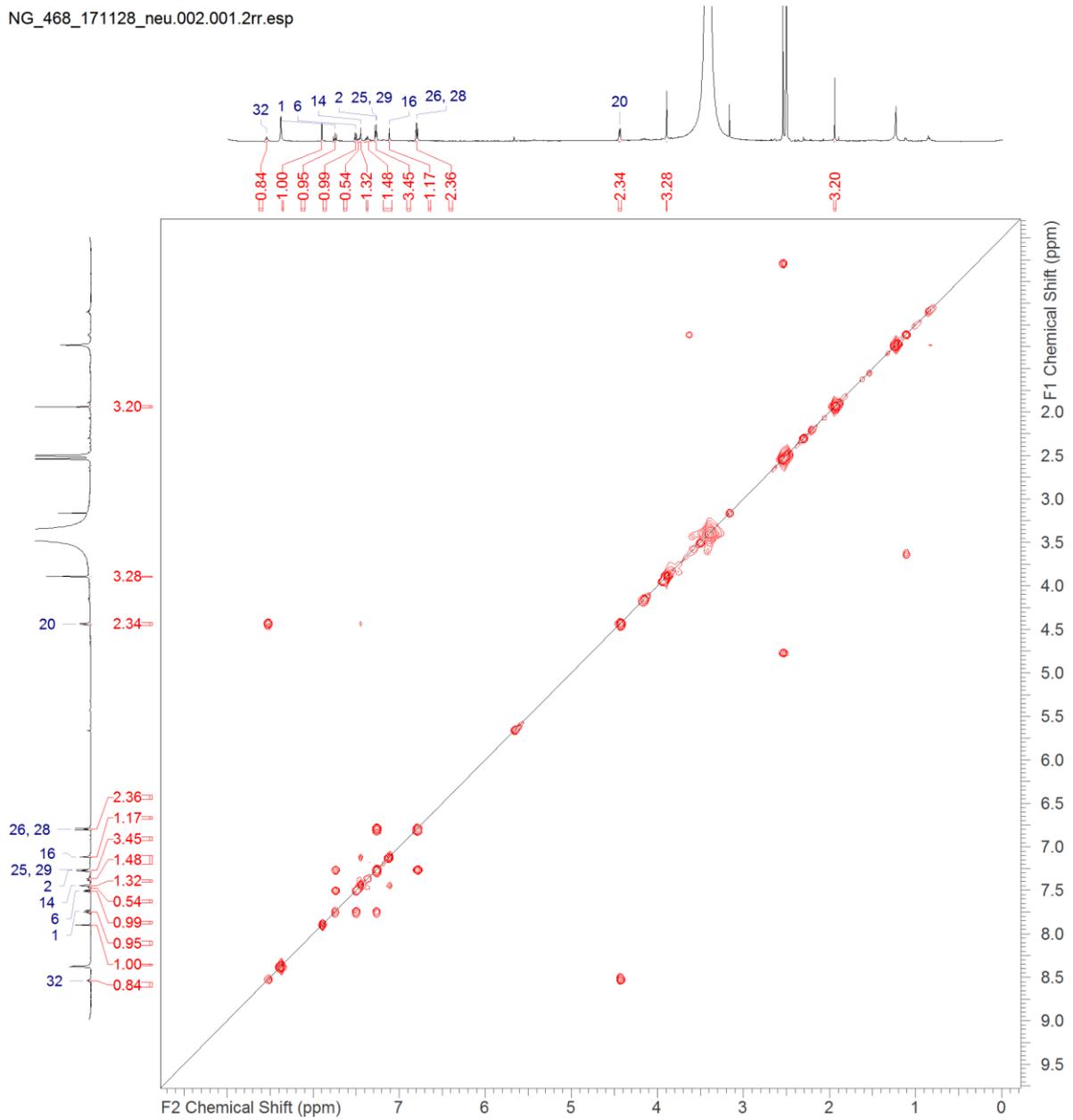


Figure S 16: ^1H - ^1H -COSY spectrum (500 MHz, DMSO-d_6) of pentangumycin.

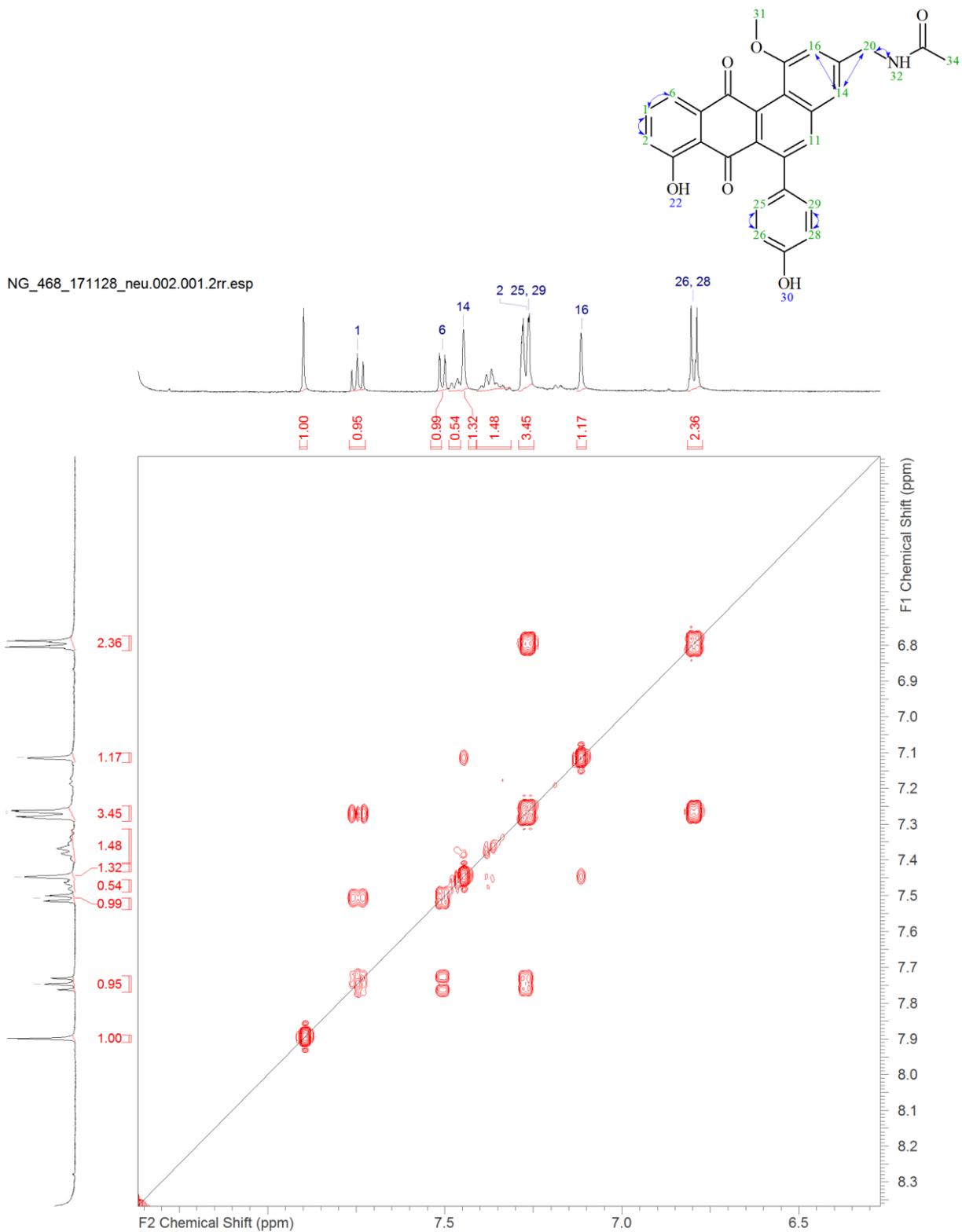


Figure S 17: ^1H - ^1H -COSY spectrum (500 MHz, DMSO-d_6) of pentangumycin; Zoom to aromatic region.

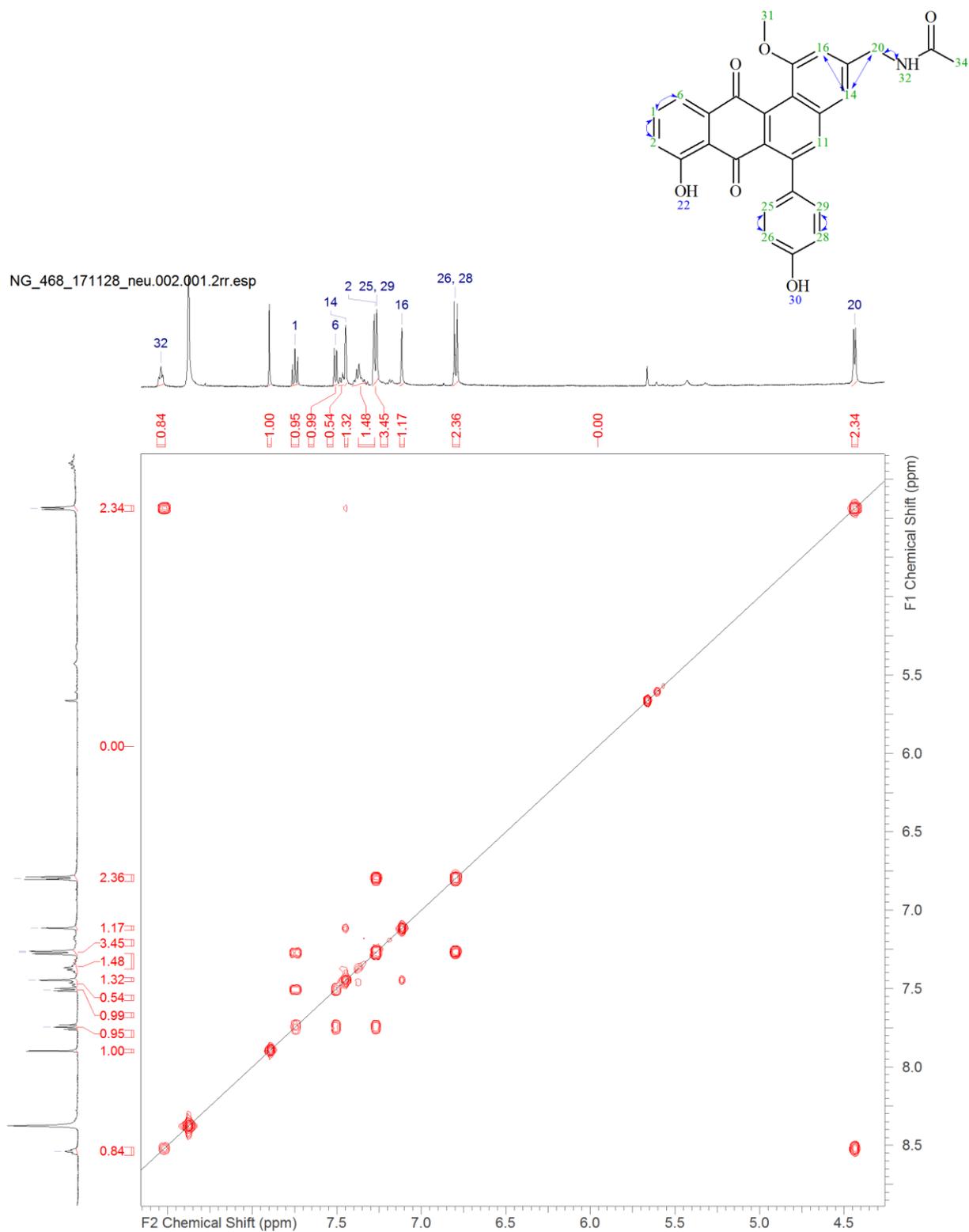
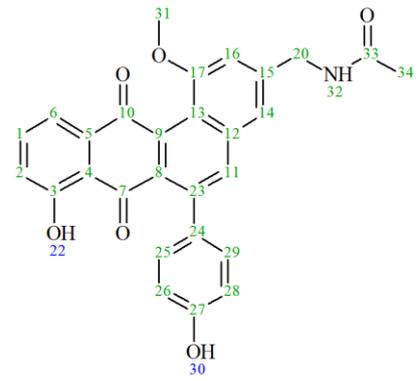


Figure S 18: ^1H - ^1H -COSY spectrum (500 MHz, DMSO-d_6) of pentangumycin; Zoom from 8.5 to 4.5 ppm.



NG_468_171128_neu.003.001.2rr.esp

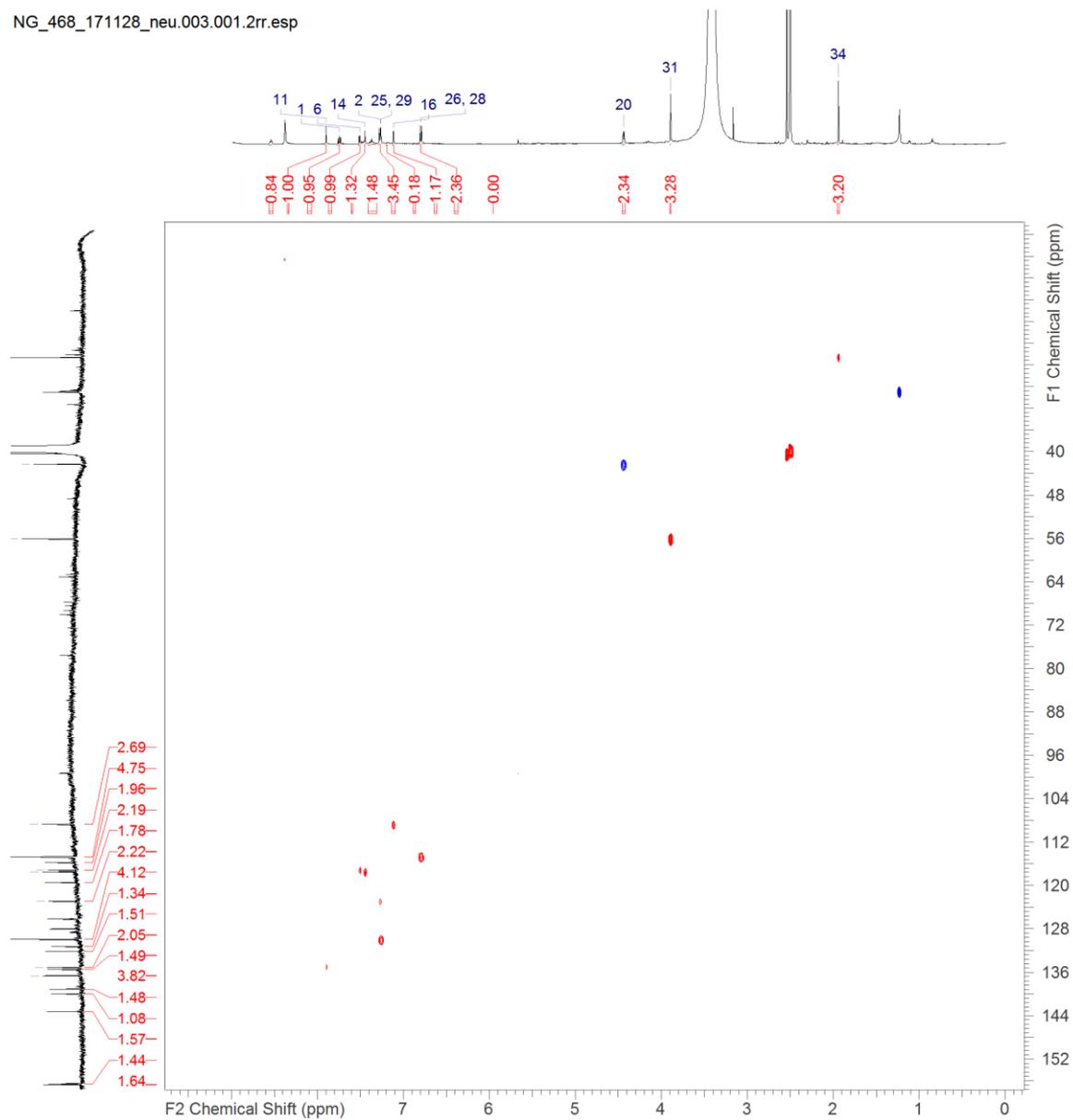


Figure S 19: HSQC-spectrum (500 MHz; 125 MHz, DMSO-d₆) of pentangumycin; complete spectrum.

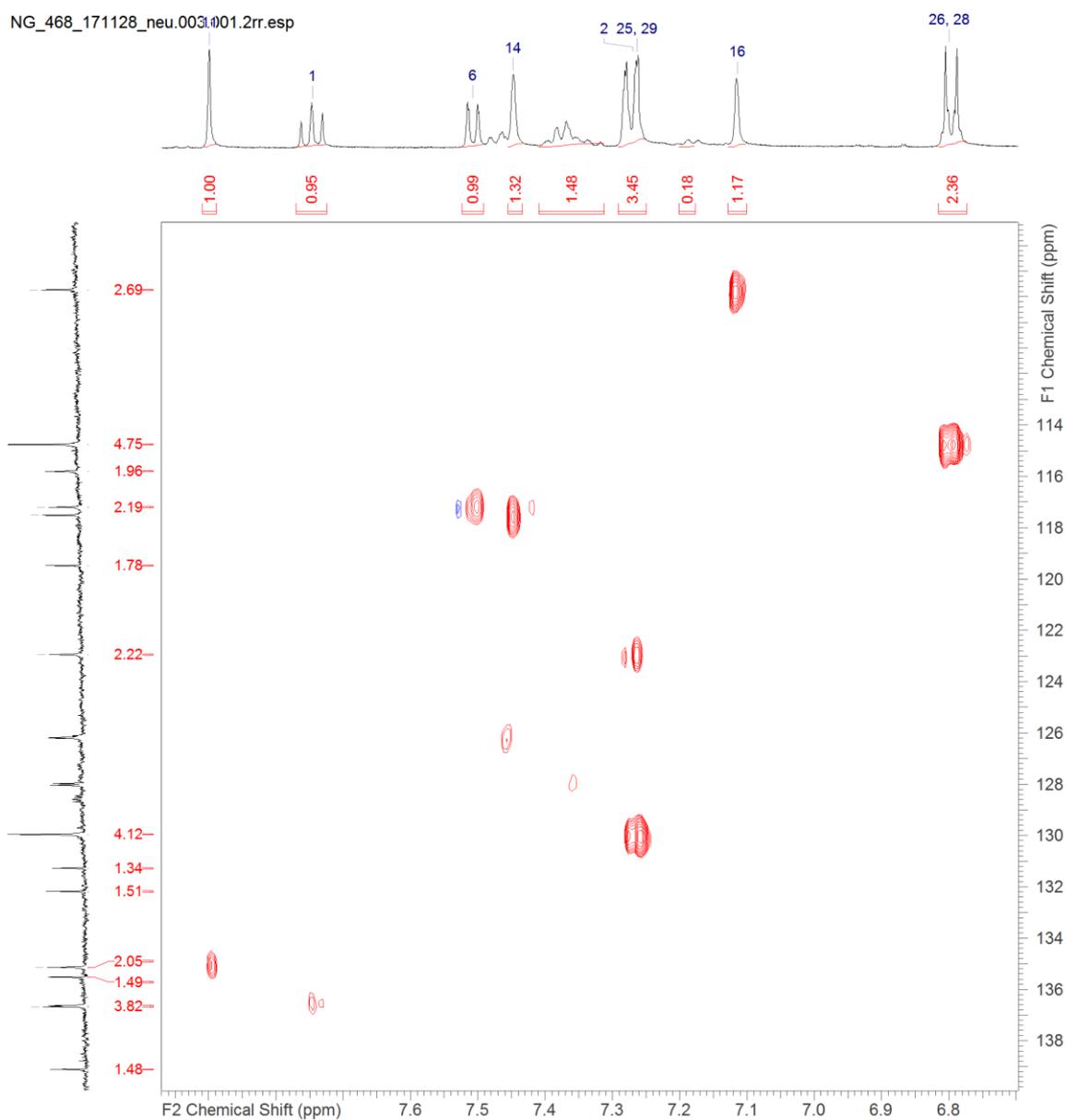
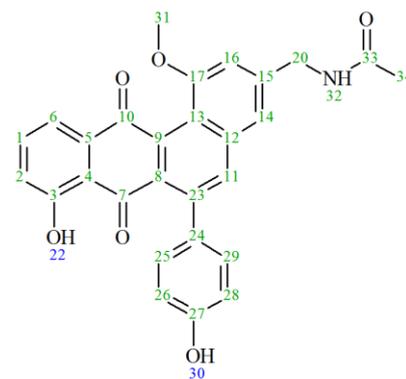
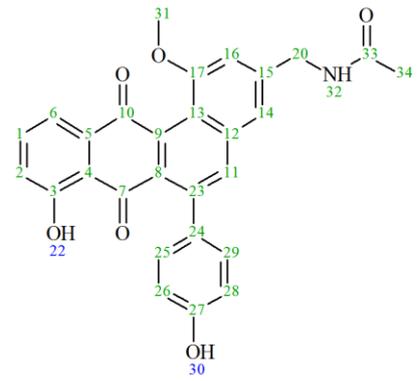


Figure S 20: HSQC-spectrum (500 MHz; 125 MHz, DMSO- d_6) of pentangumycin; zoom from 7.9 to 6.8 ppm and 139 to 107 ppm.



NG_468_171128_neu.003.001.2rr.esp

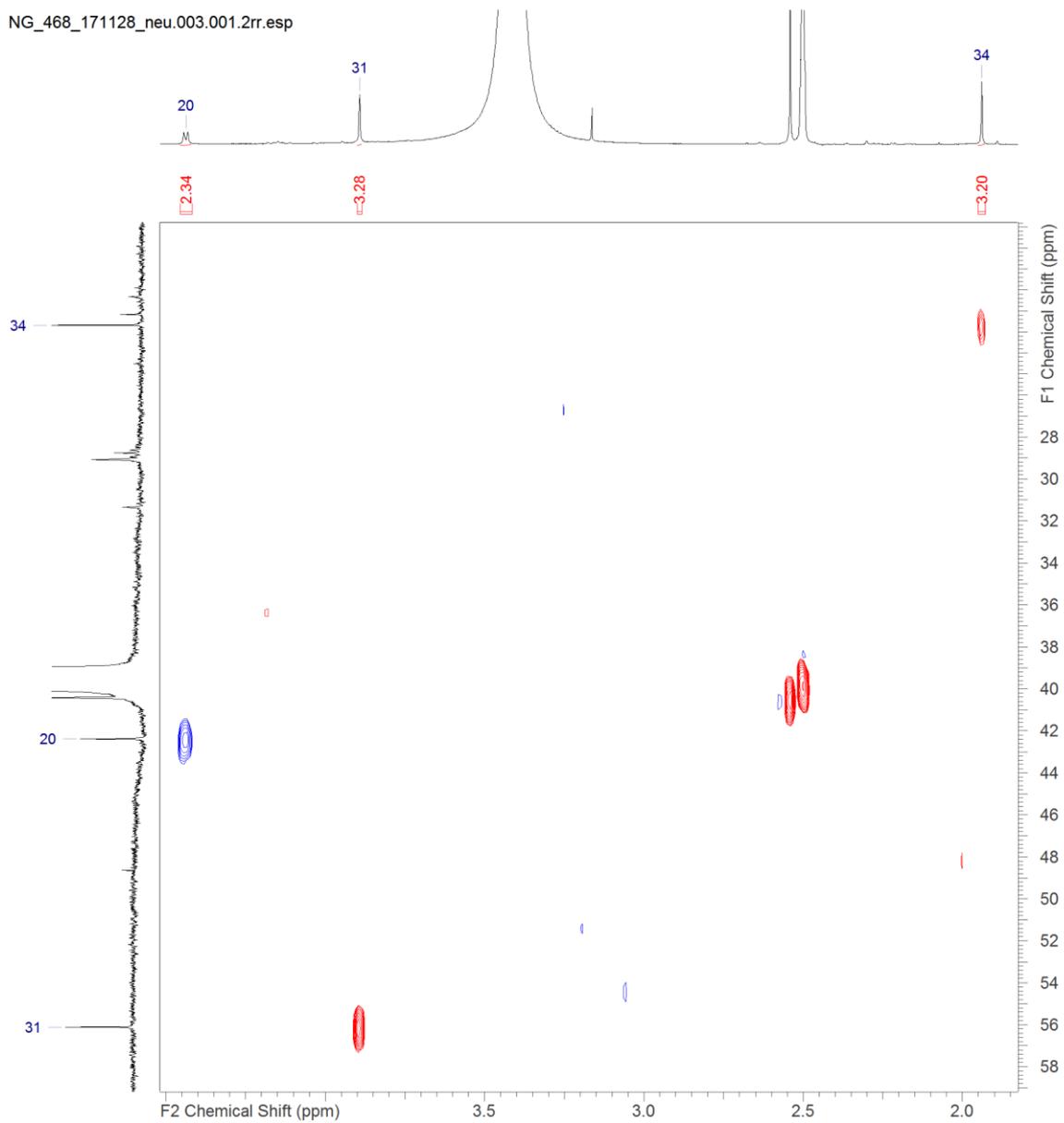
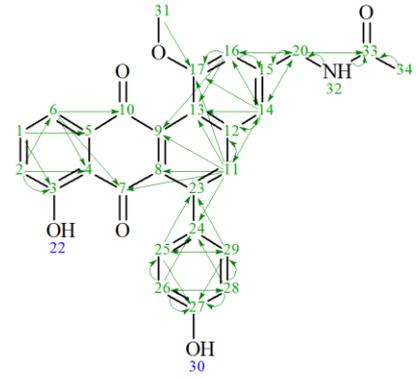
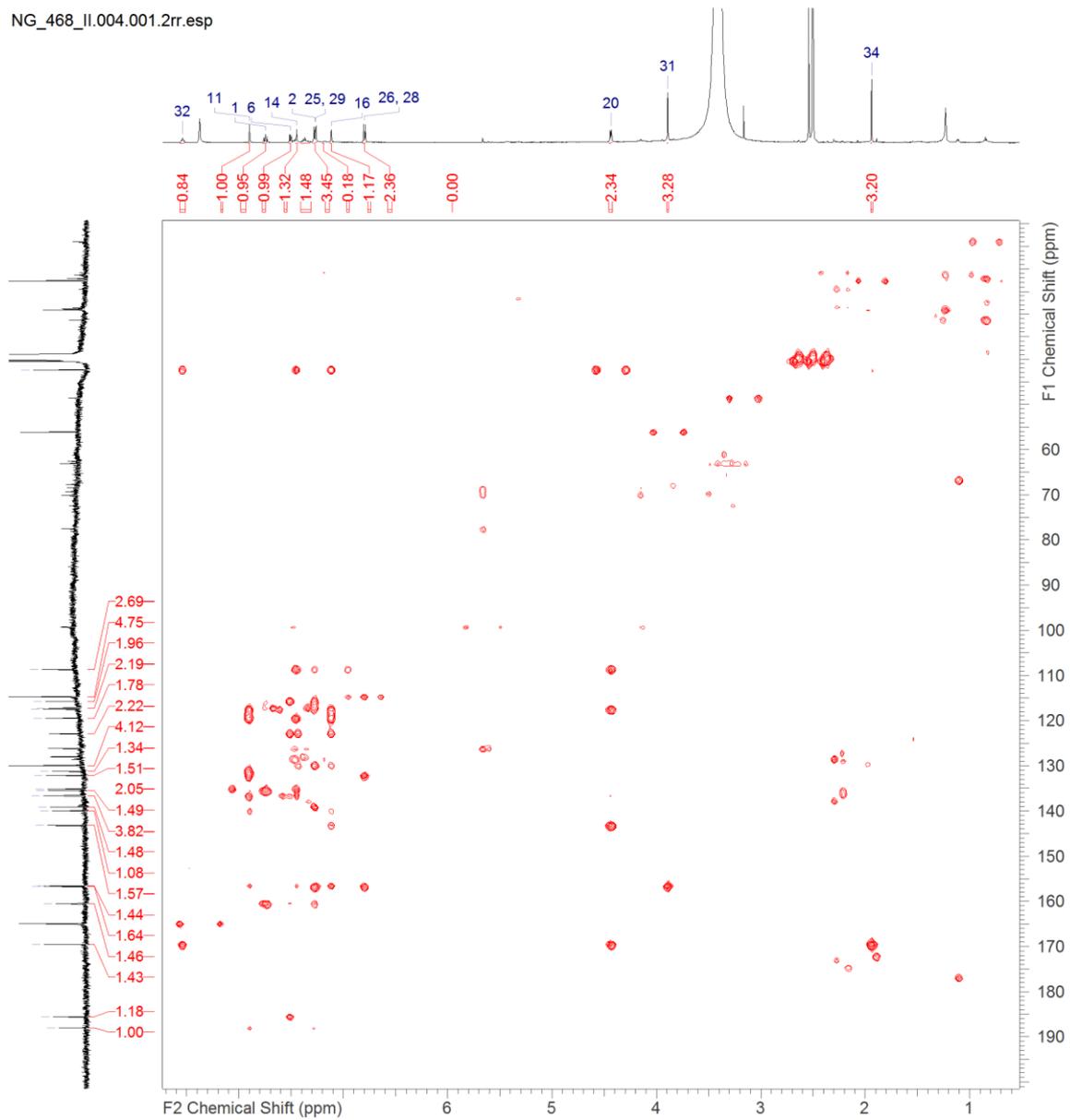


Figure S 21: HSQC-spectrum (500 MHz; 125 MHz, DMSO-d₆) of pentangumycin; zoom from 4 to 2 ppm and 59 to 19 ppm.



NG_468_II.004.001.2rr.esp



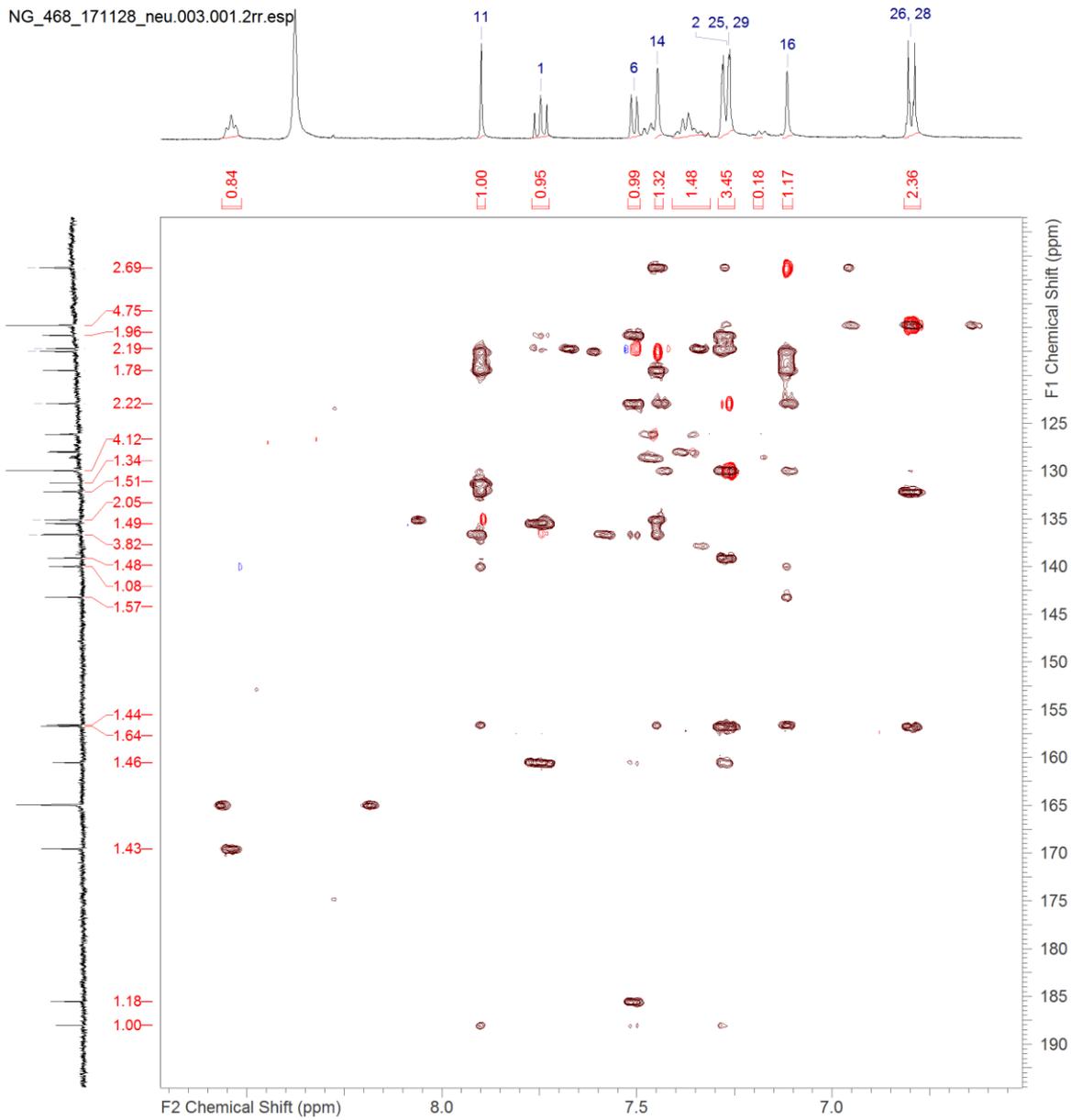
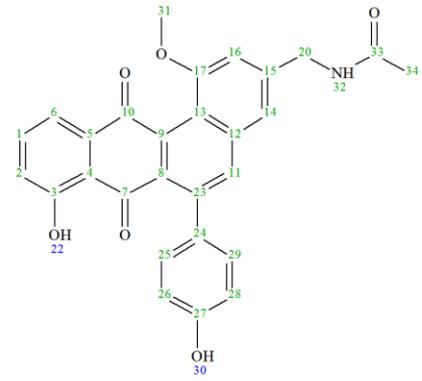


Figure S 23: HMBC - HSQC-Overlay spectrum (500 MHz; 125 MHz, DMSO- d_6) of pentangumycin (black: HMBC; red/blue: HSQC); zoom from 9.2 to 6.5 ppm and 195 to 110 ppm.

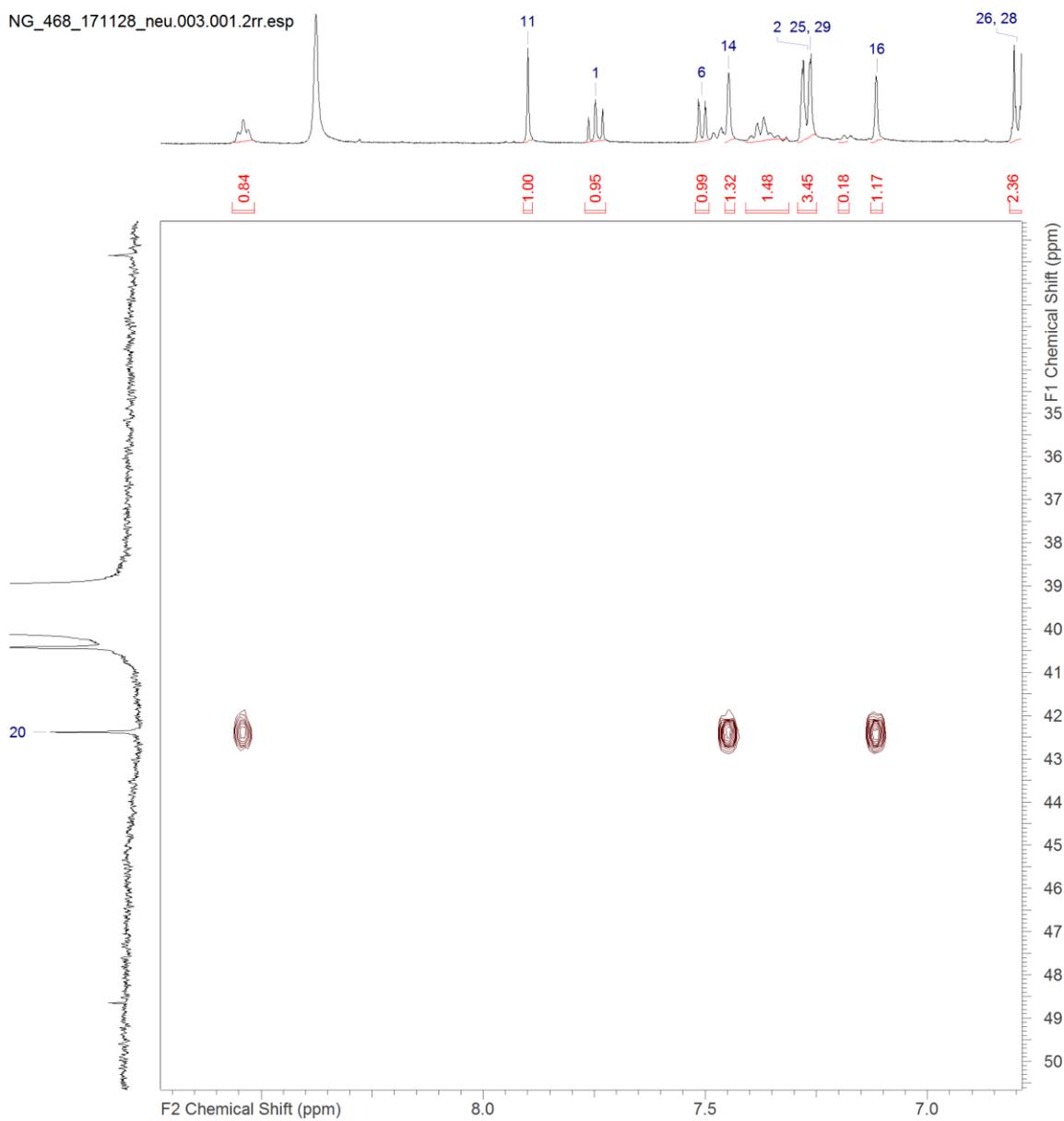
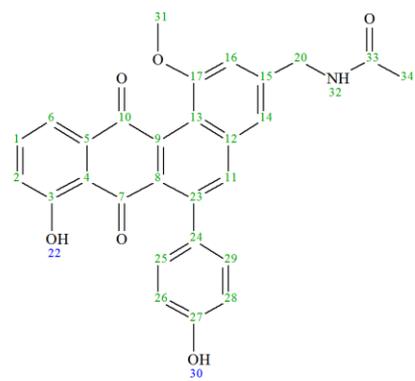
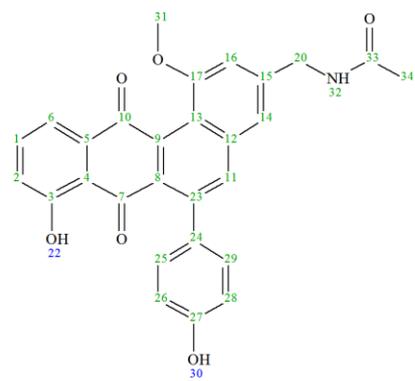


Figure S 24: HMBC - HSQC-Overlay spectrum (500 MHz; 125 MHz, DMSO- d_6) of pentangumycin (black: HMBC; red/blue: HSQC); zoom from 9.2 to 6.5 ppm and 50 to 31 ppm.



NG_468_171128_neu.003.001.2rr.esp

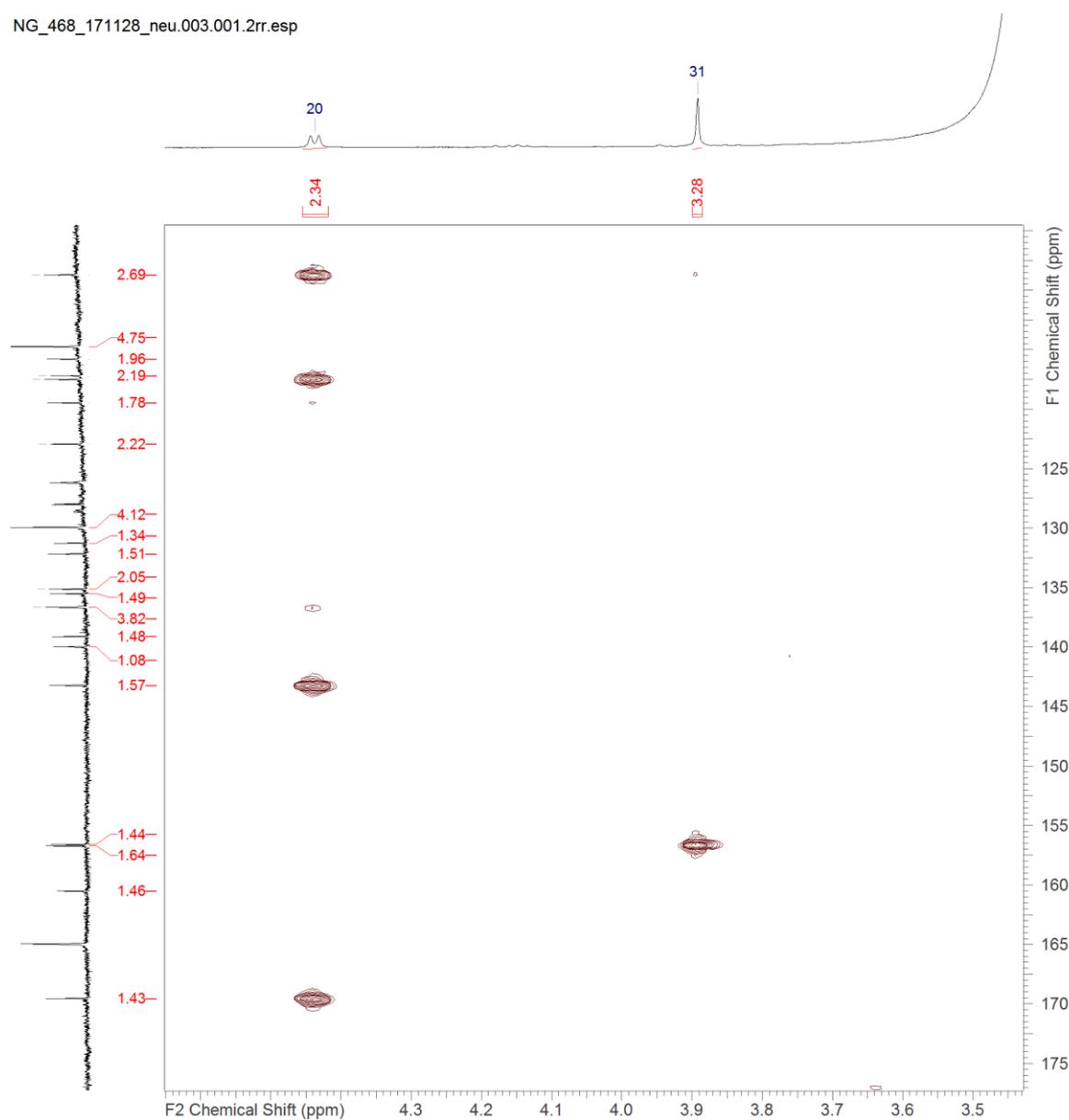
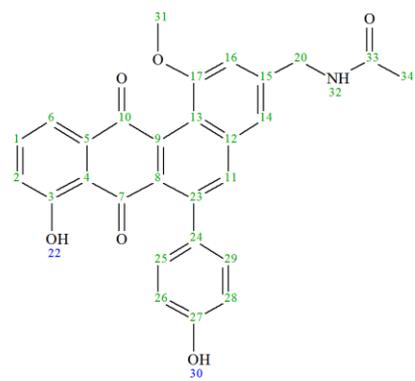


Figure S 25: HMBC - HSQC-Overlay spectrum (500 MHz; 125 MHz, DMSO- d_6) of pentangumycin (black: HMBC; red/blue: HSQC); zoom from 4.6 to 3.5 ppm and 110 to 175 ppm.



NG_468_171128_neu.003.001.2rr.esp

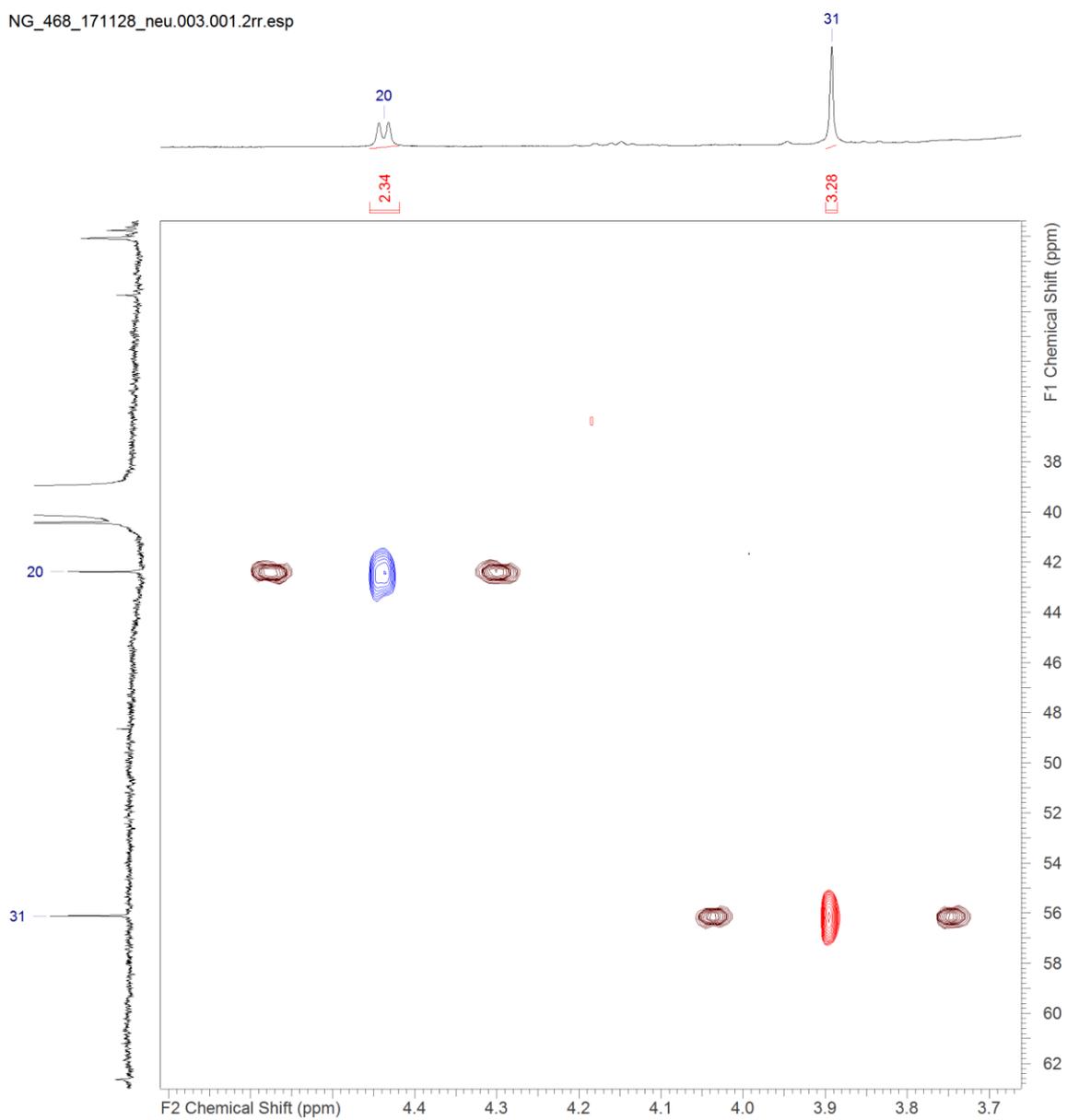


Figure S 26: HMBC - HSQC-Overlay spectrum (500 MHz; 125 MHz, DMSO- d_6) of pentangumycin (black: HMBC; red/blue: HSQC); zoom from 4.6 to 3.7 ppm and 62 to 34 ppm.

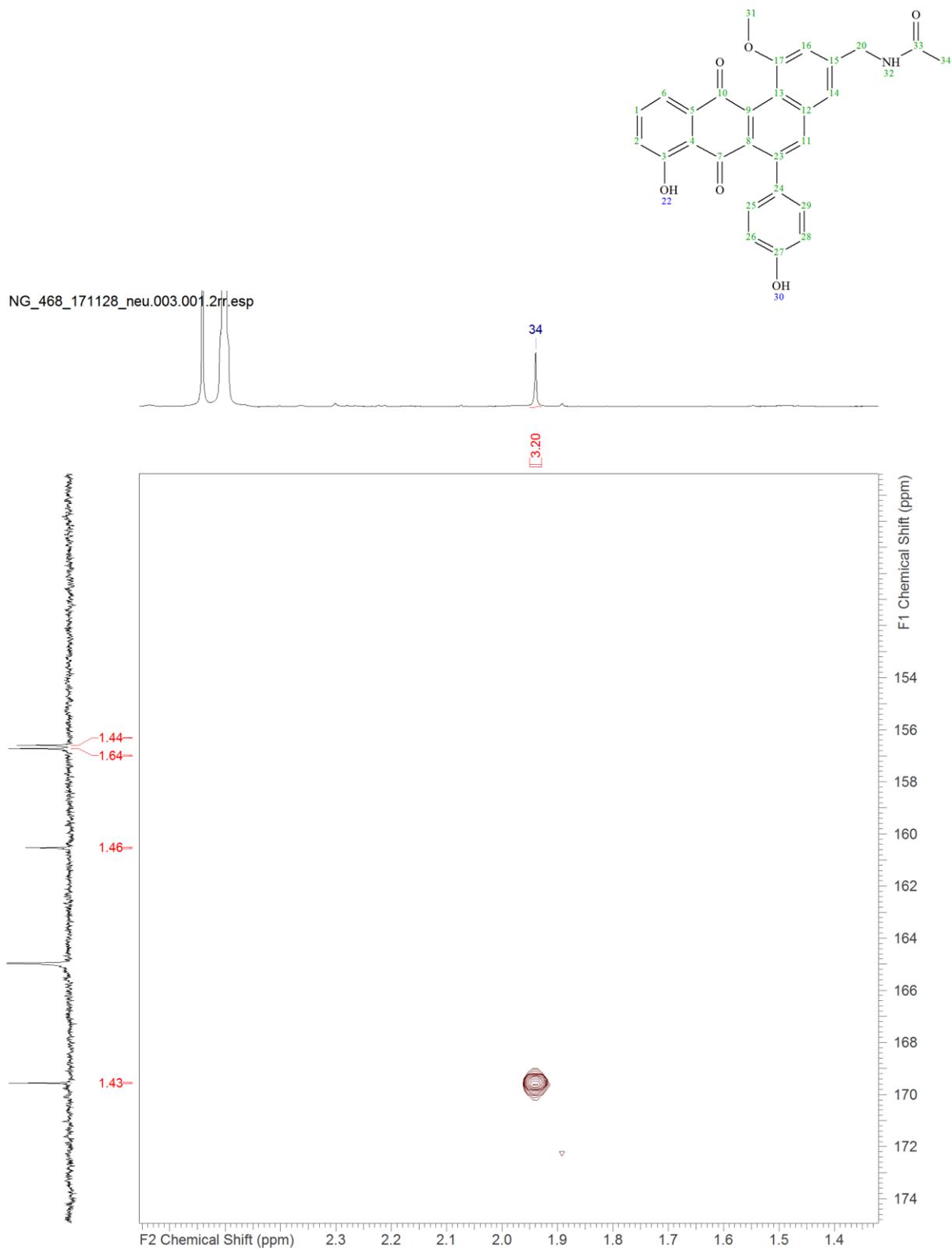


Figure S 27: HMBC - HSQC-Overlay spectrum (500 MHz; 125 MHz, DMSO- d_6) of pentangumycin (black: HMBC; red/blue: HSQC); zoom from 2.6 to 1.4 ppm and 174 to 151 ppm.

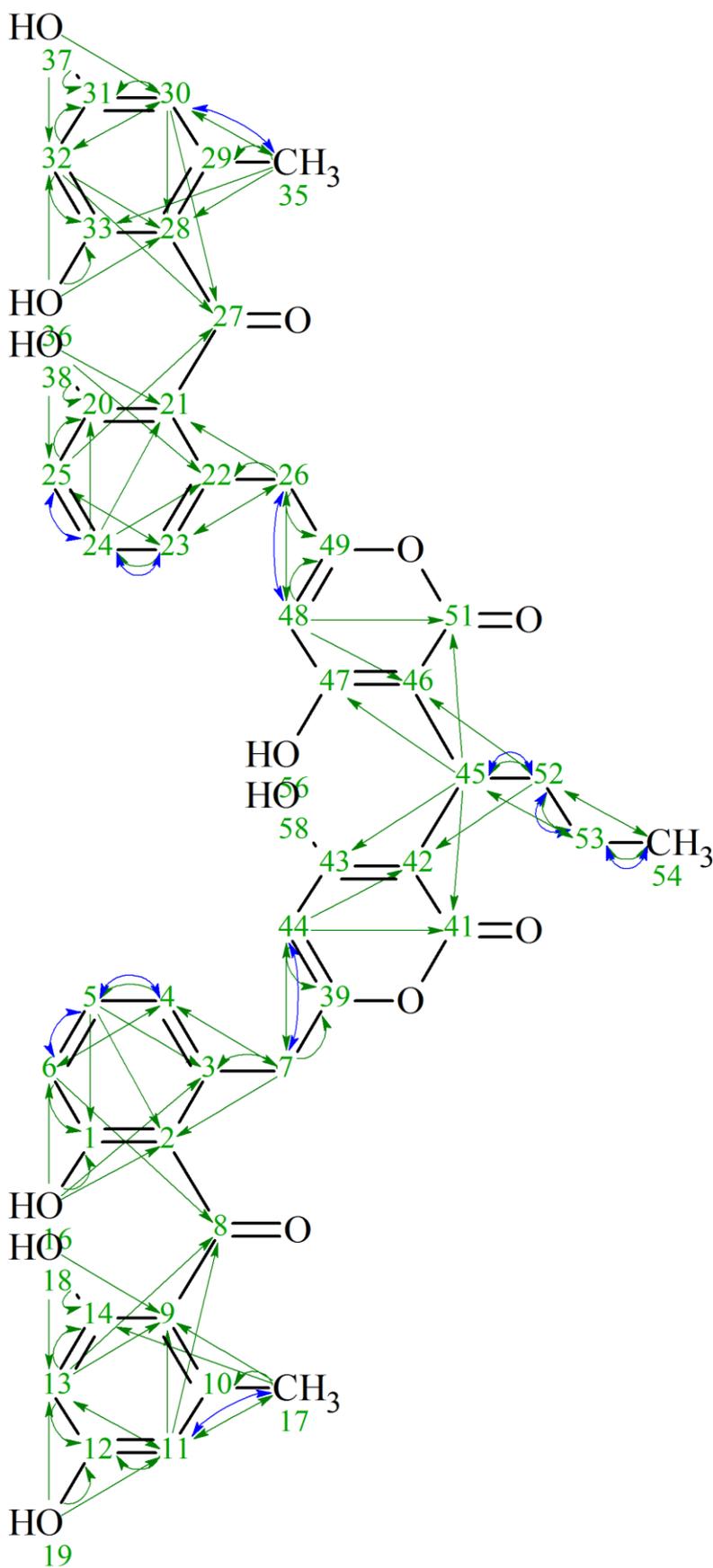


Figure S 28: Structure of SEK90 with all observed correlations (green: HMBC correlations H → C; blue: ¹H-¹H-Cosy correlation H → H).

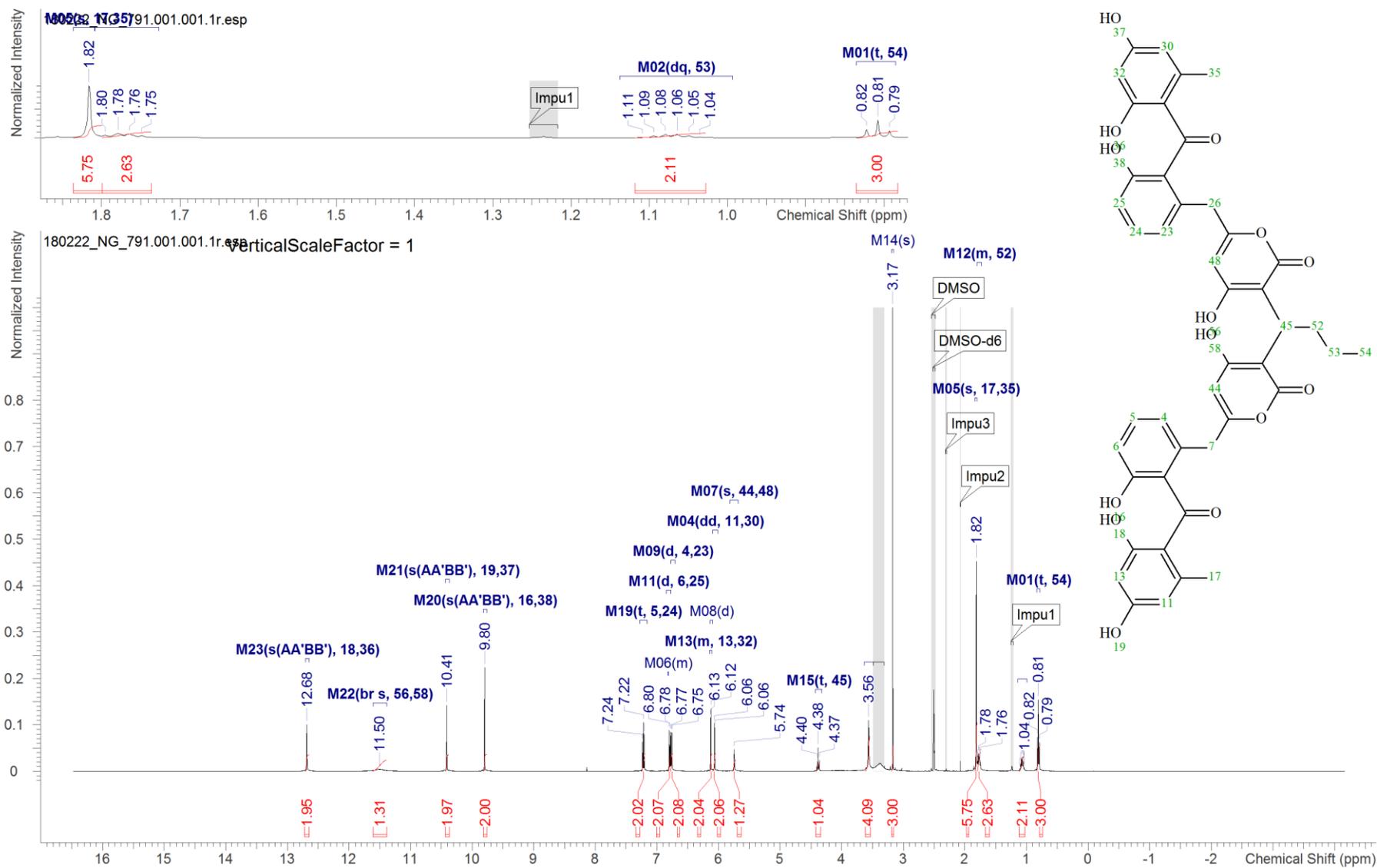


Figure S 29: ¹H-NMR spectrum (500 MHz, DMSO-d₆) of SEK90; complete Spectrum and zoom from 1.8 to 0.7 ppm.

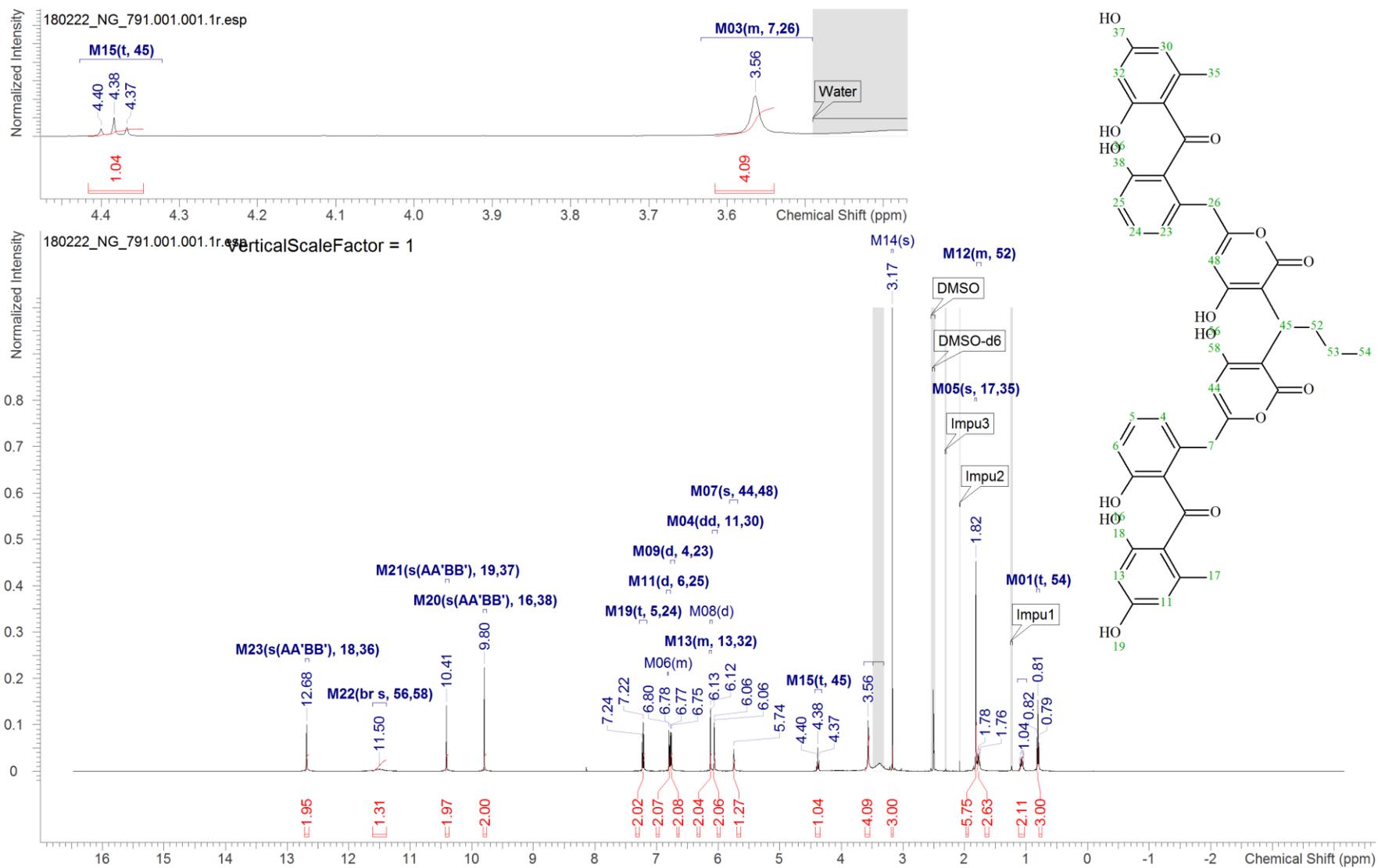


Figure S 30: $^1\text{H-NMR}$ spectrum (500 MHz, DMSO-d_6) of SEK90; complete Spectrum and zoom from 4.4 to 3.4 ppm.

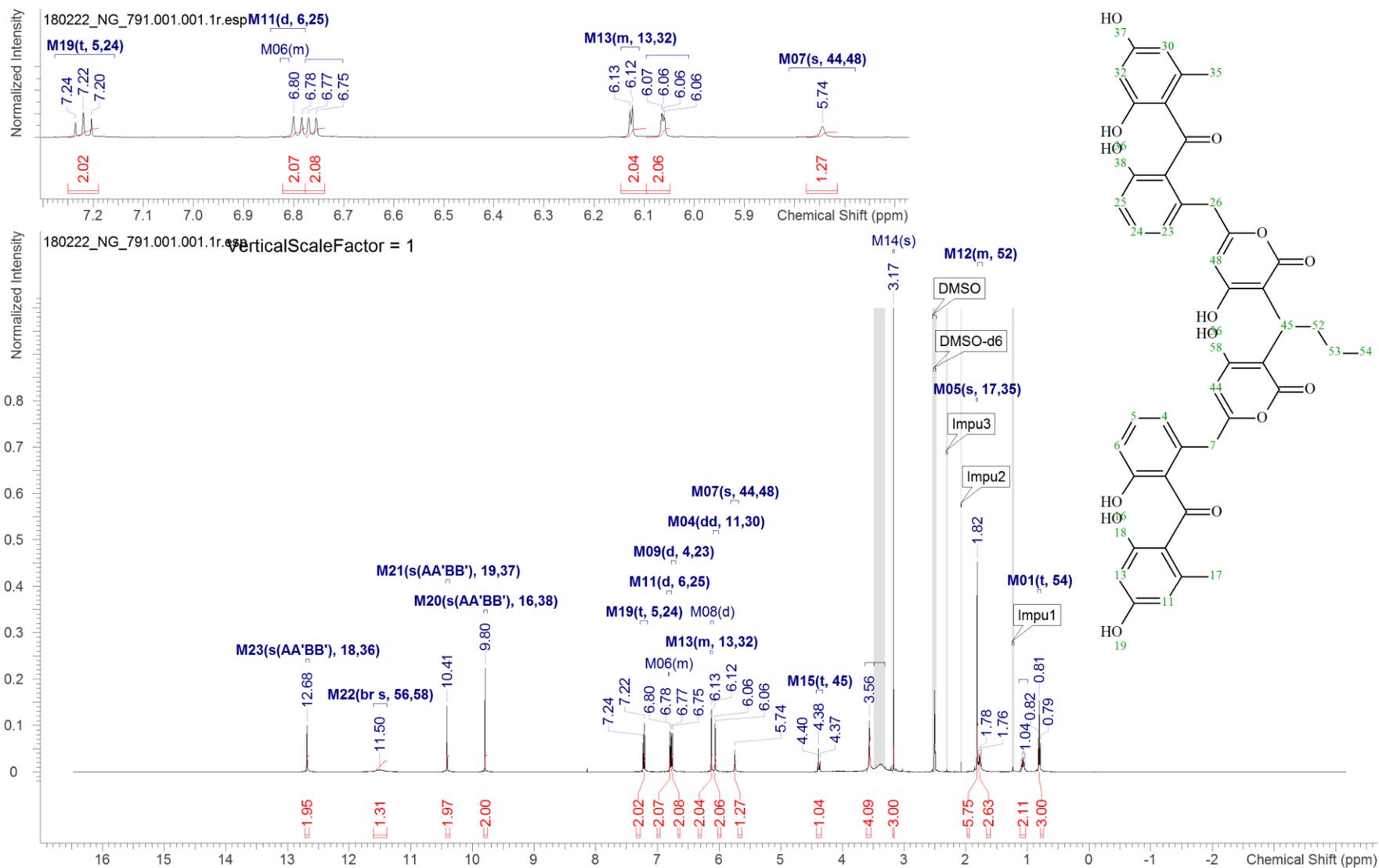


Figure S 31: $^1\text{H-NMR}$ spectrum (500 MHz, DMSO-d_6) of SEK90; complete Spectrum and zoom from 7.2 to 5.7 ppm.

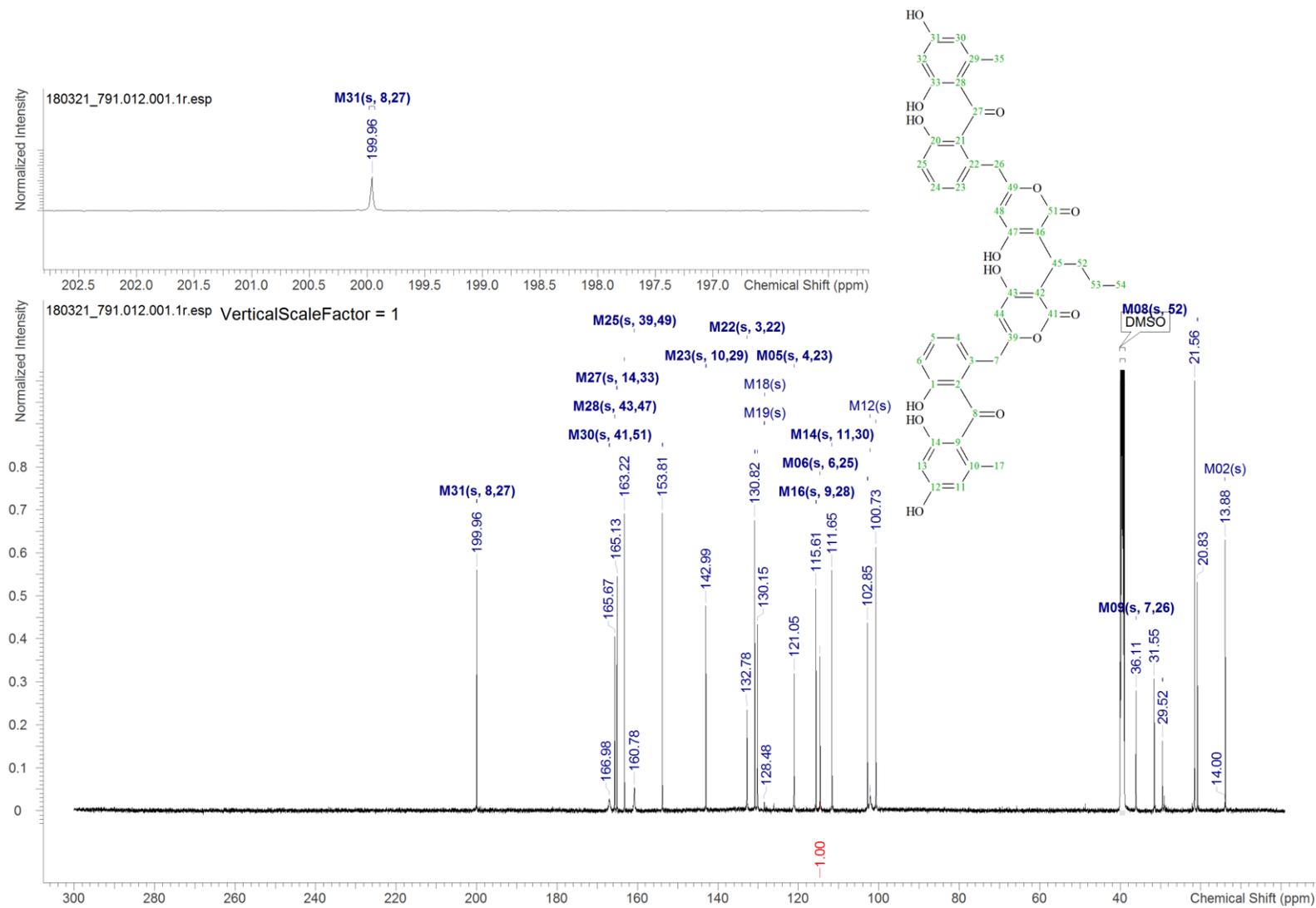


Figure S 33: ^{13}C -NMR spectrum (125 MHz, DMSO-d_6) of SEK90; complete Spectrum and zoom from 202 to 196 ppm.

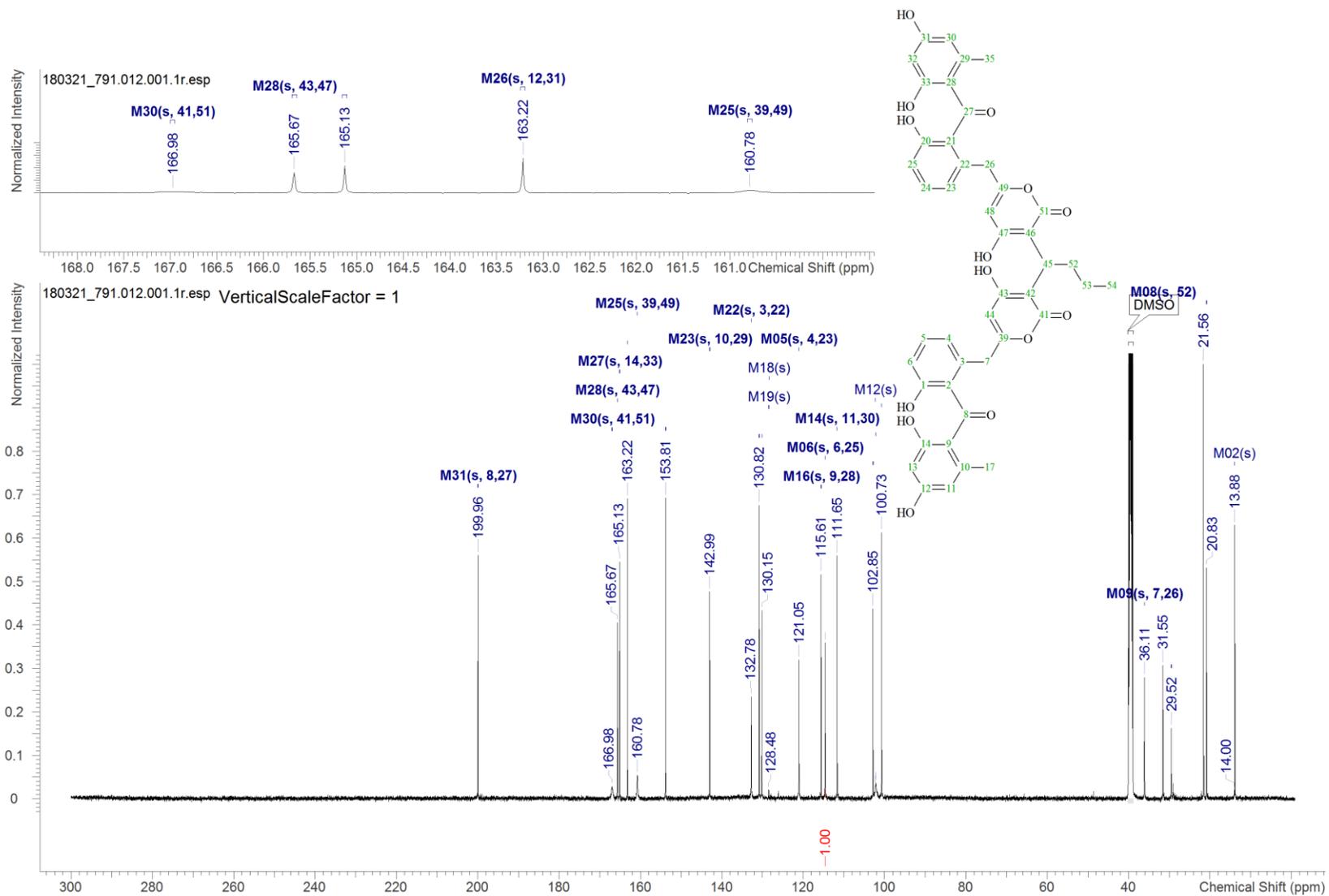


Figure S 34: ^{13}C -NMR spectrum (125 MHz, DMSO-d_6) of SEK90; complete Spectrum and zoom from 168 to 60 ppm.

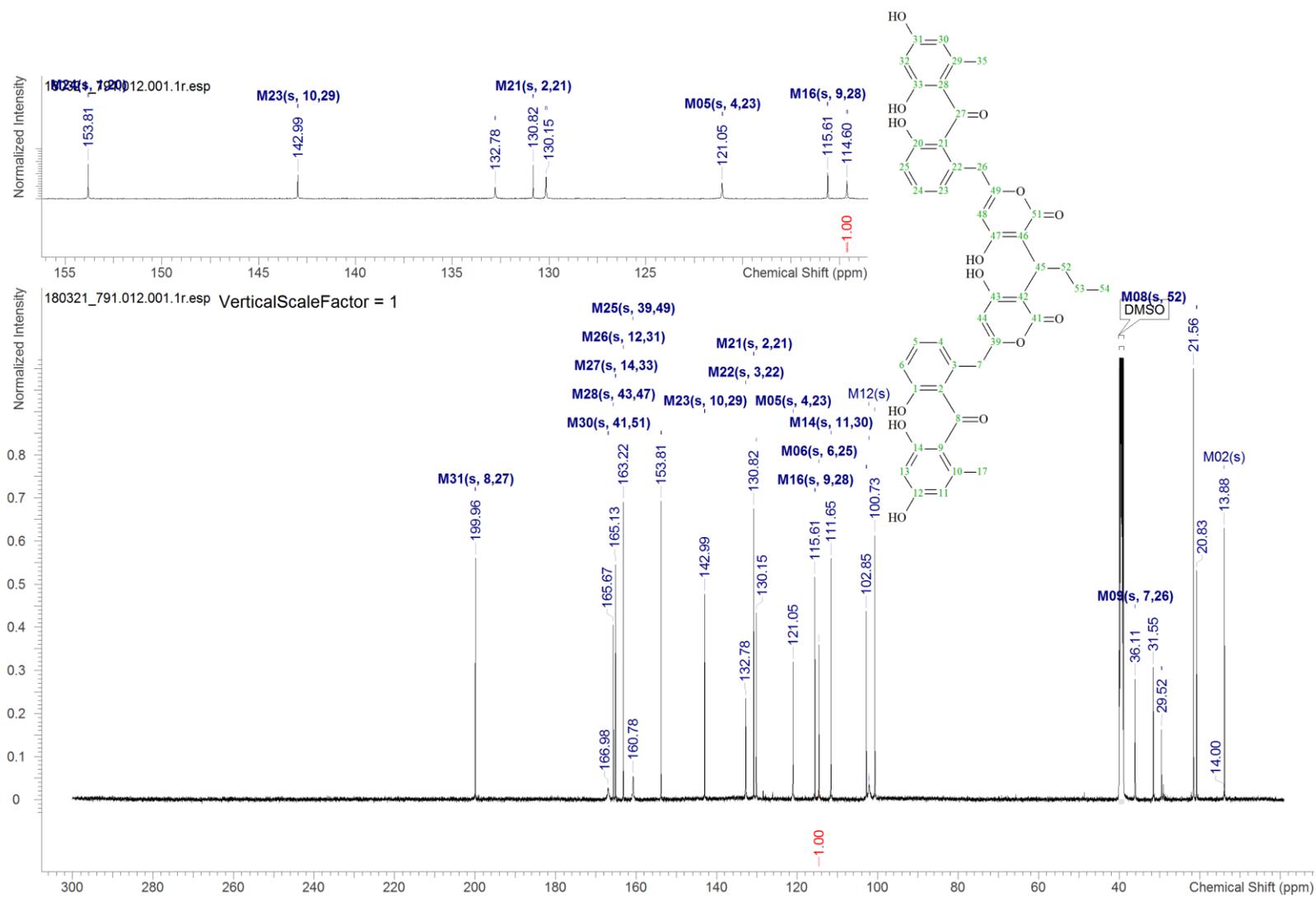


Figure S 35: ^{13}C -NMR spectrum (125 MHz, DMSO-d_6) of SEK90; complete Spectrum and zoom from 155 to 113 ppm.

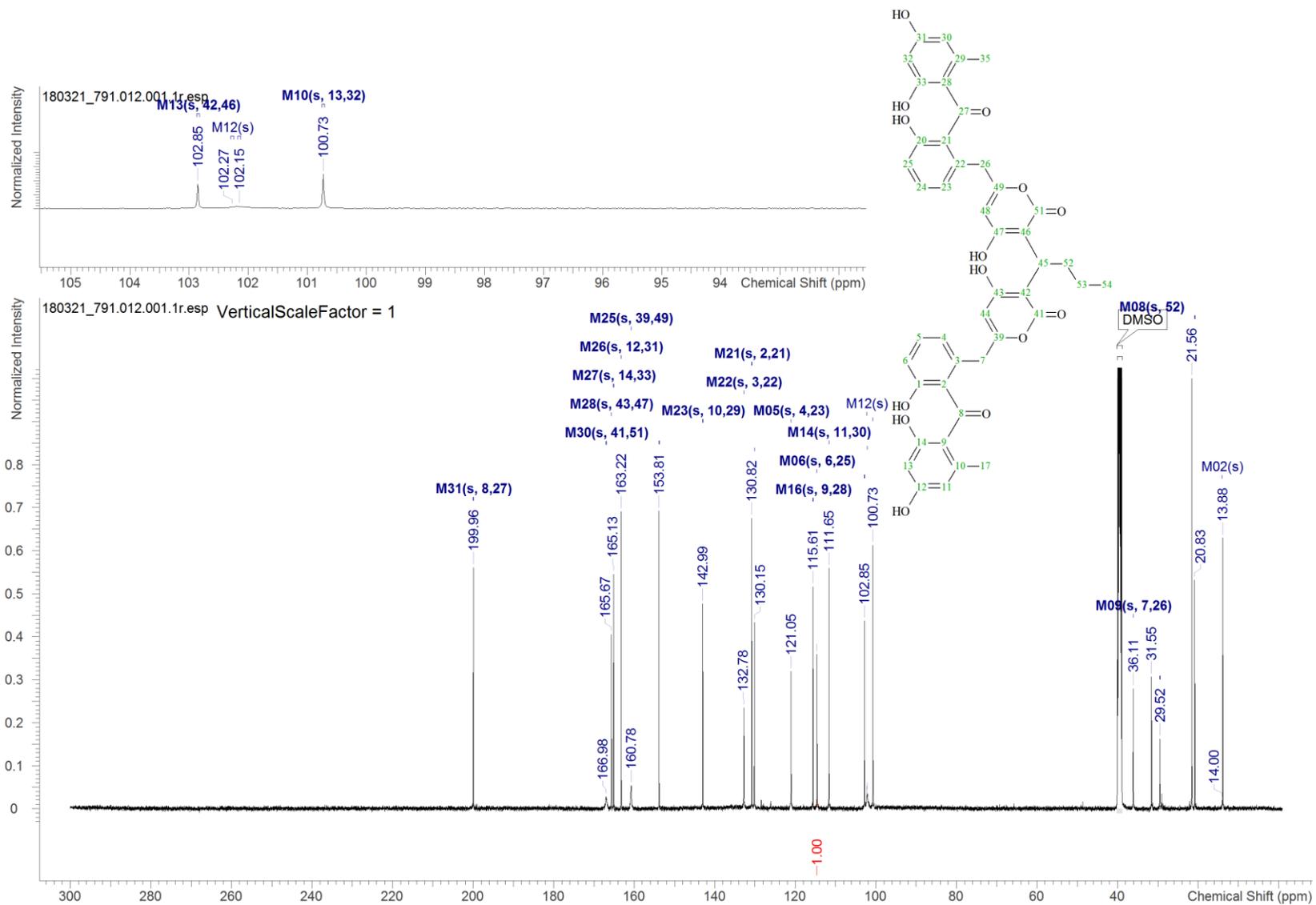


Figure S 36: ^{13}C -NMR spectrum (125 MHz, DMSO-d_6) of SEK90; complete Spectrum and zoom from 105 to 92 ppm.

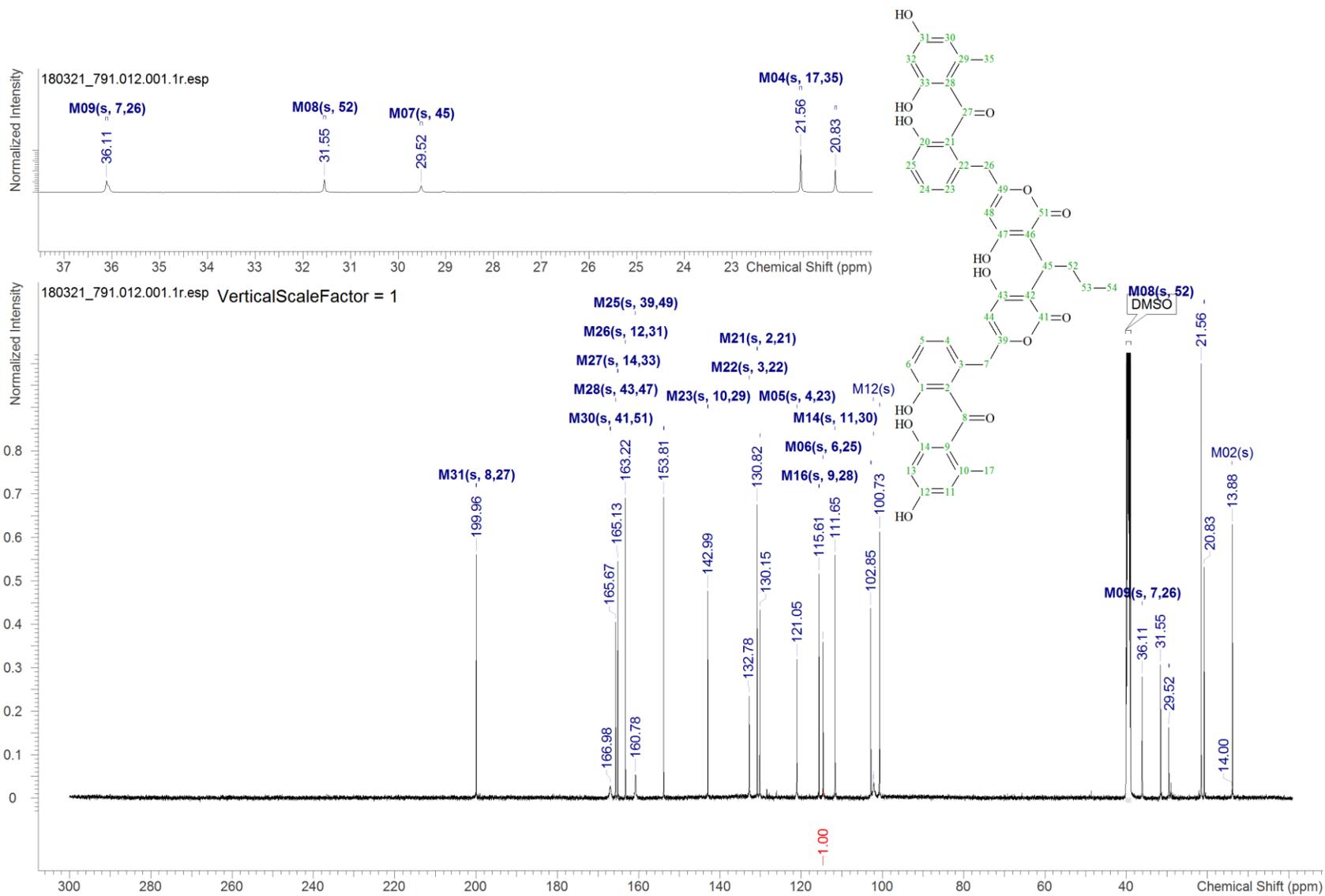


Figure S 37: ^{13}C -NMR spectrum (125 MHz, DMSO-d_6) of SEK90; complete Spectrum and zoom from 37 to 20 ppm.

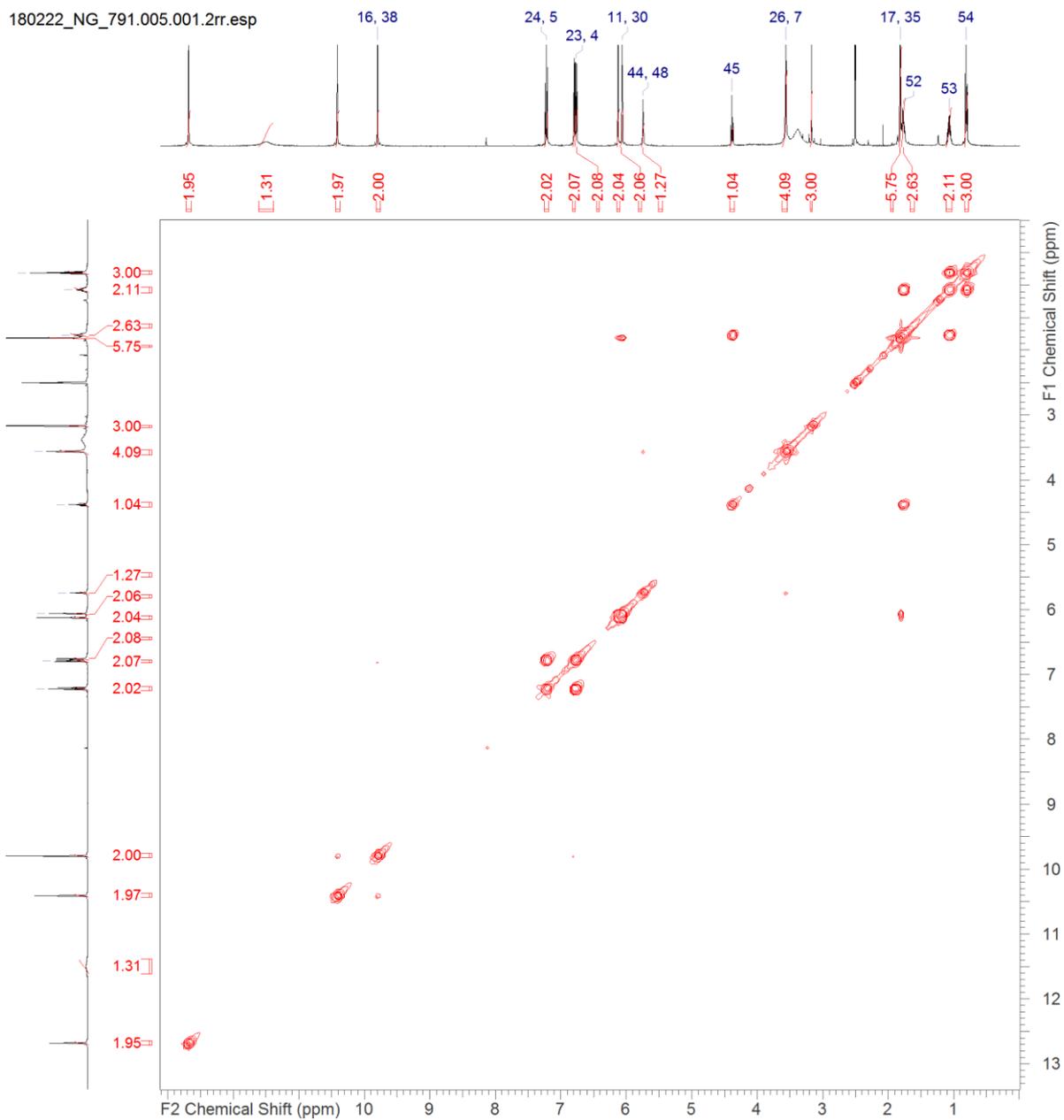


Figure S 39: ^1H - ^1H - COSY spectrum (500 MHz, DMSO-d_6) of SEK90.

180222_NG_791.005.001.2rr.esp

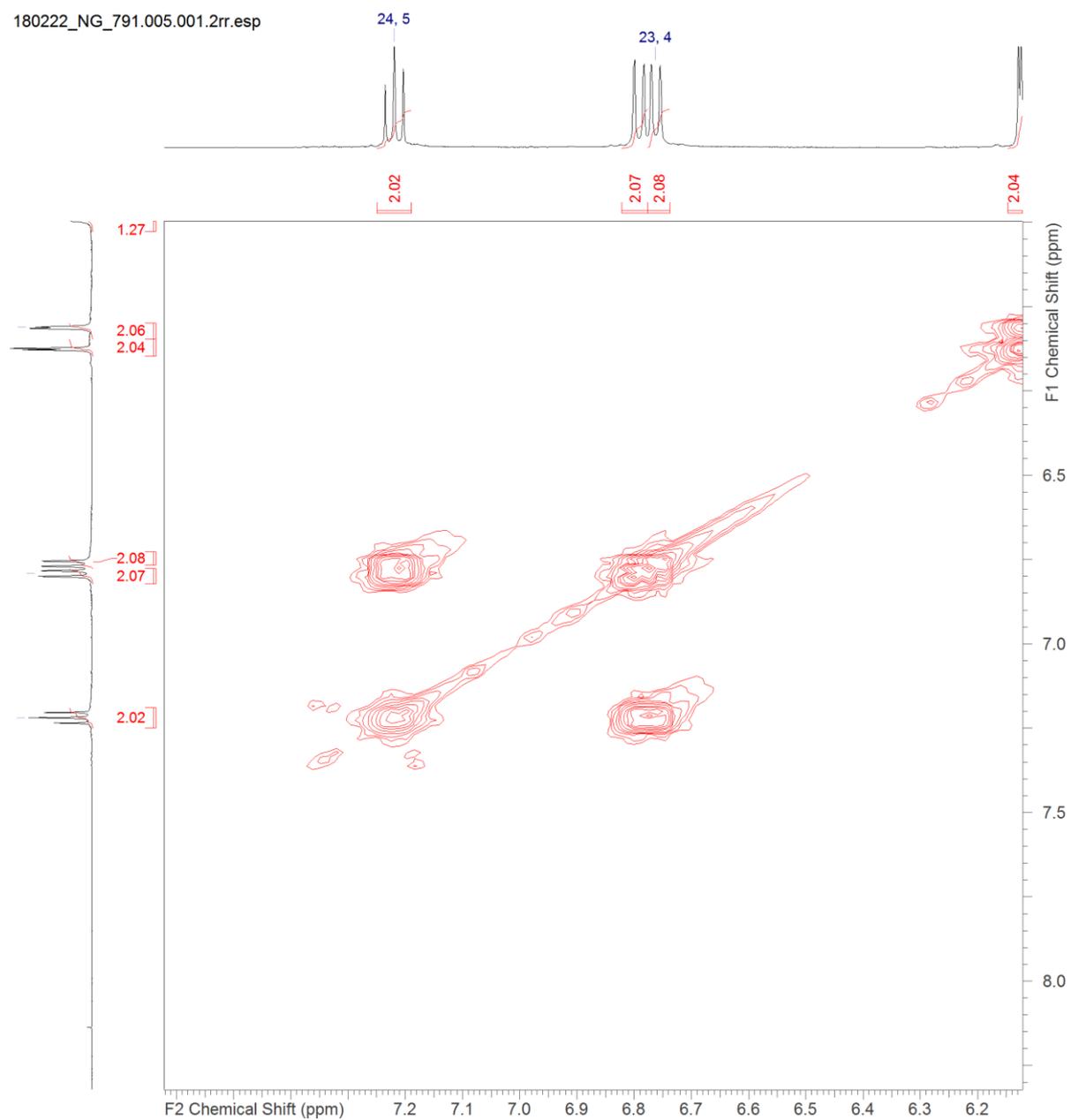


Figure S 40: ^1H - ^1H - COSY spectrum (500 MHz, DMSO-d_6) of SEK90; zoom from 8 to 6.2 ppm.

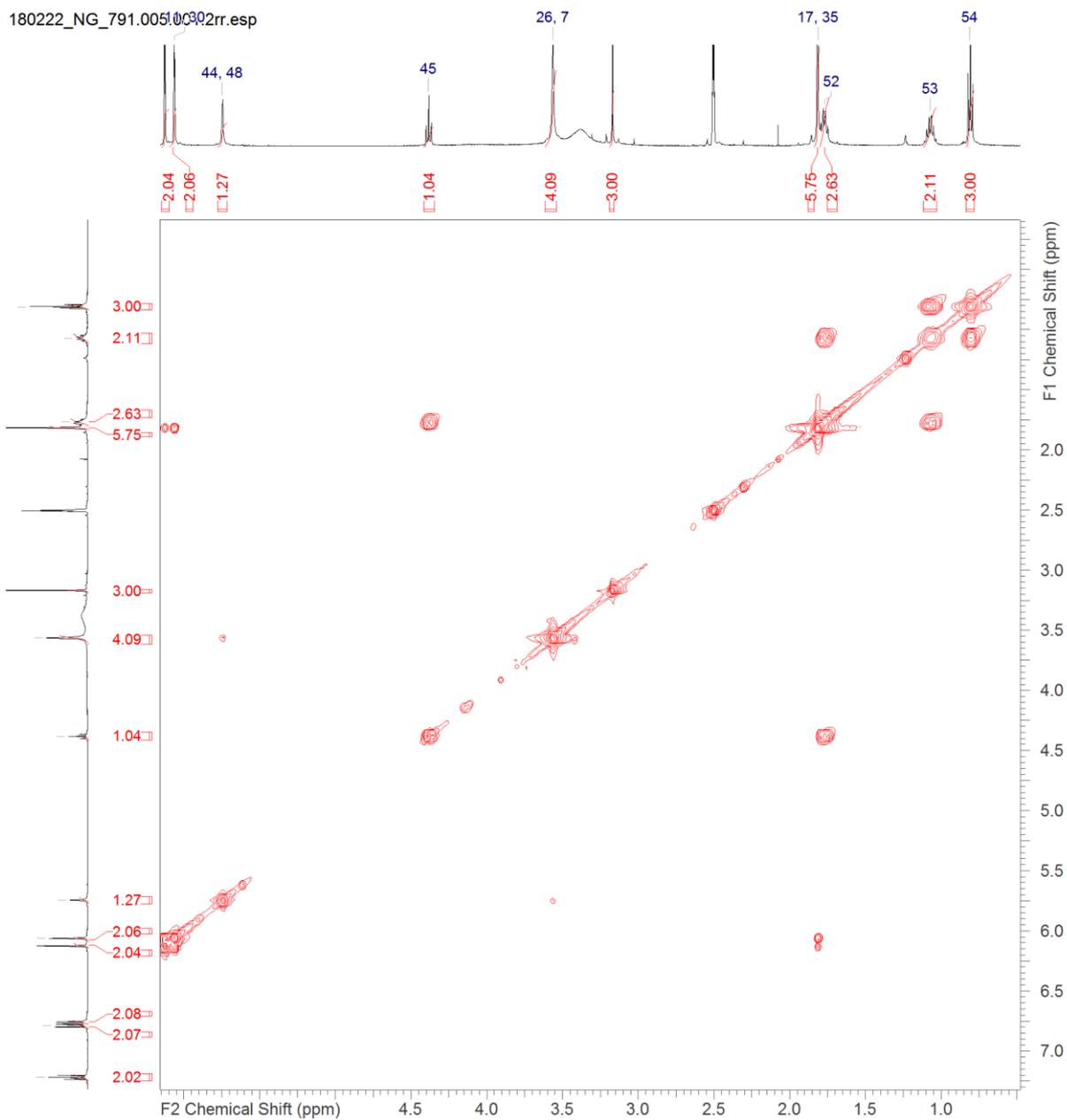


Figure S 41: ^1H - ^1H -COSY spectrum (500 MHz, DMSO-d_6) of SEK90; zoom from 7 to 0.5 ppm.

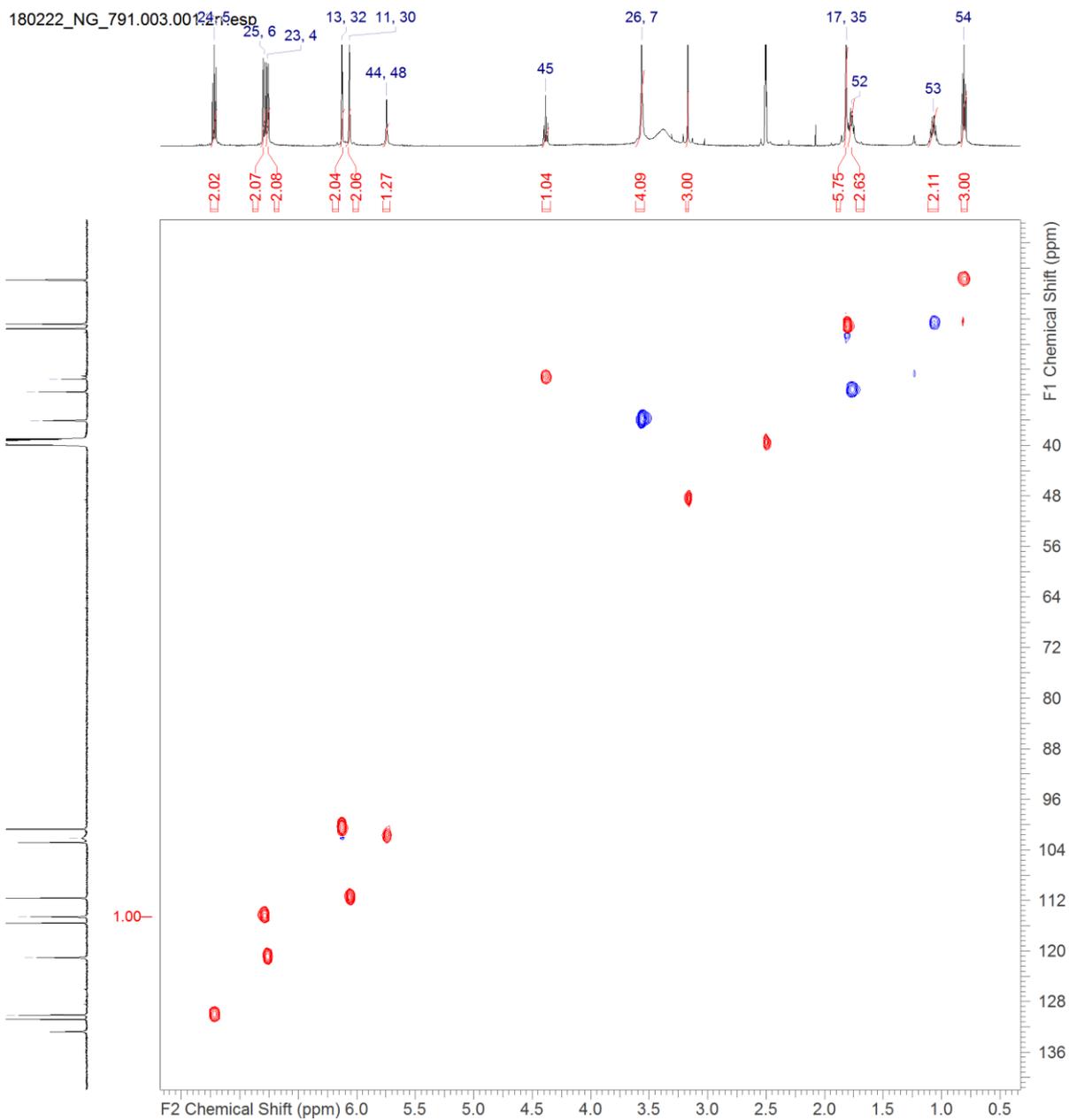


Figure S 42: HSQC-spectrum (500 MHz; 125 MHz, DMSO-d₆) of SEK90.

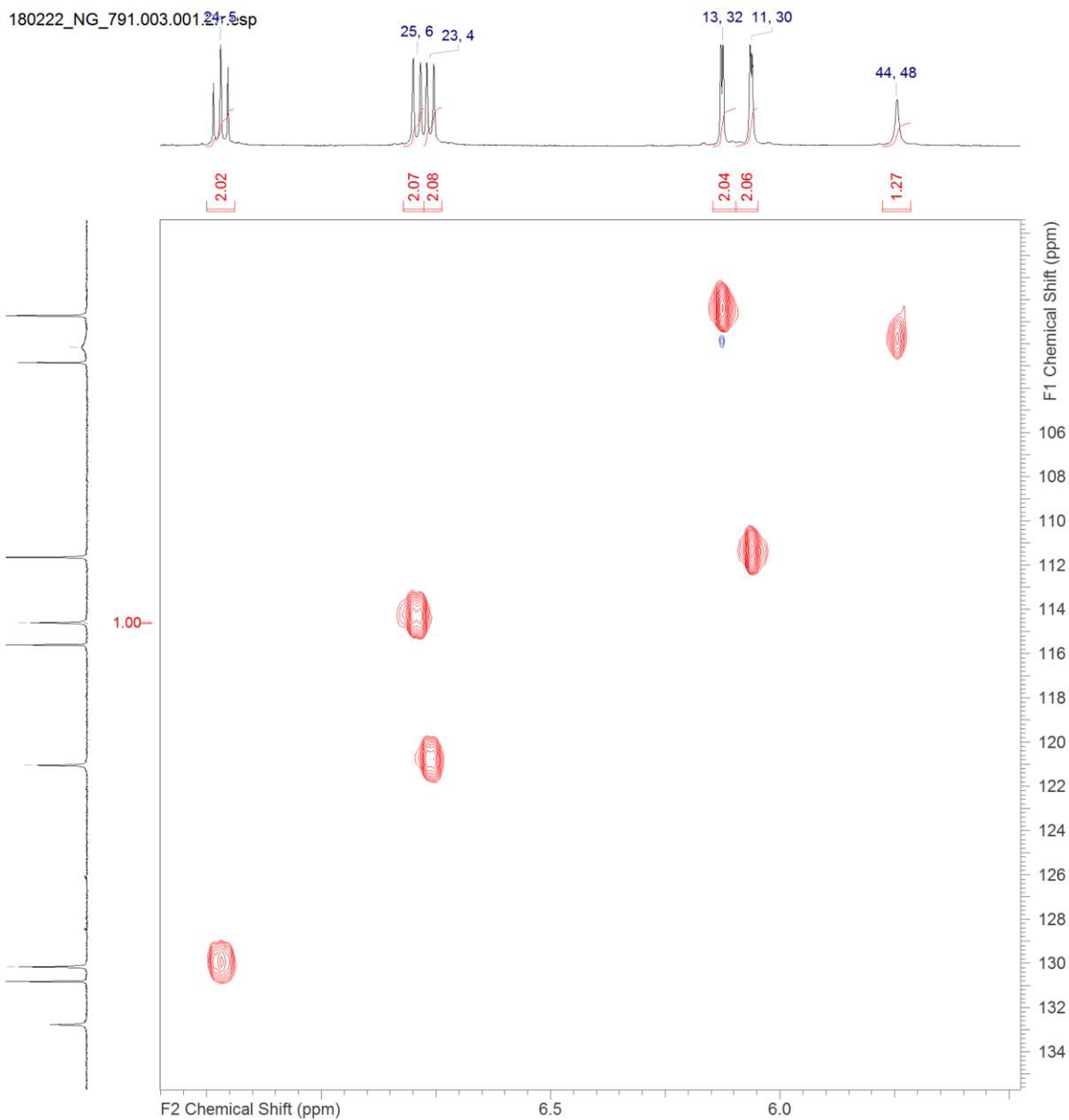


Figure S 43: HSQC-spectrum (500 MHz; 125 MHz, DMSO- d_6) of SEK90; zoom from 7.5 to 5.5 ppm and 135 to 99 ppm.

180222_NG_791.003.001.2rr.esp

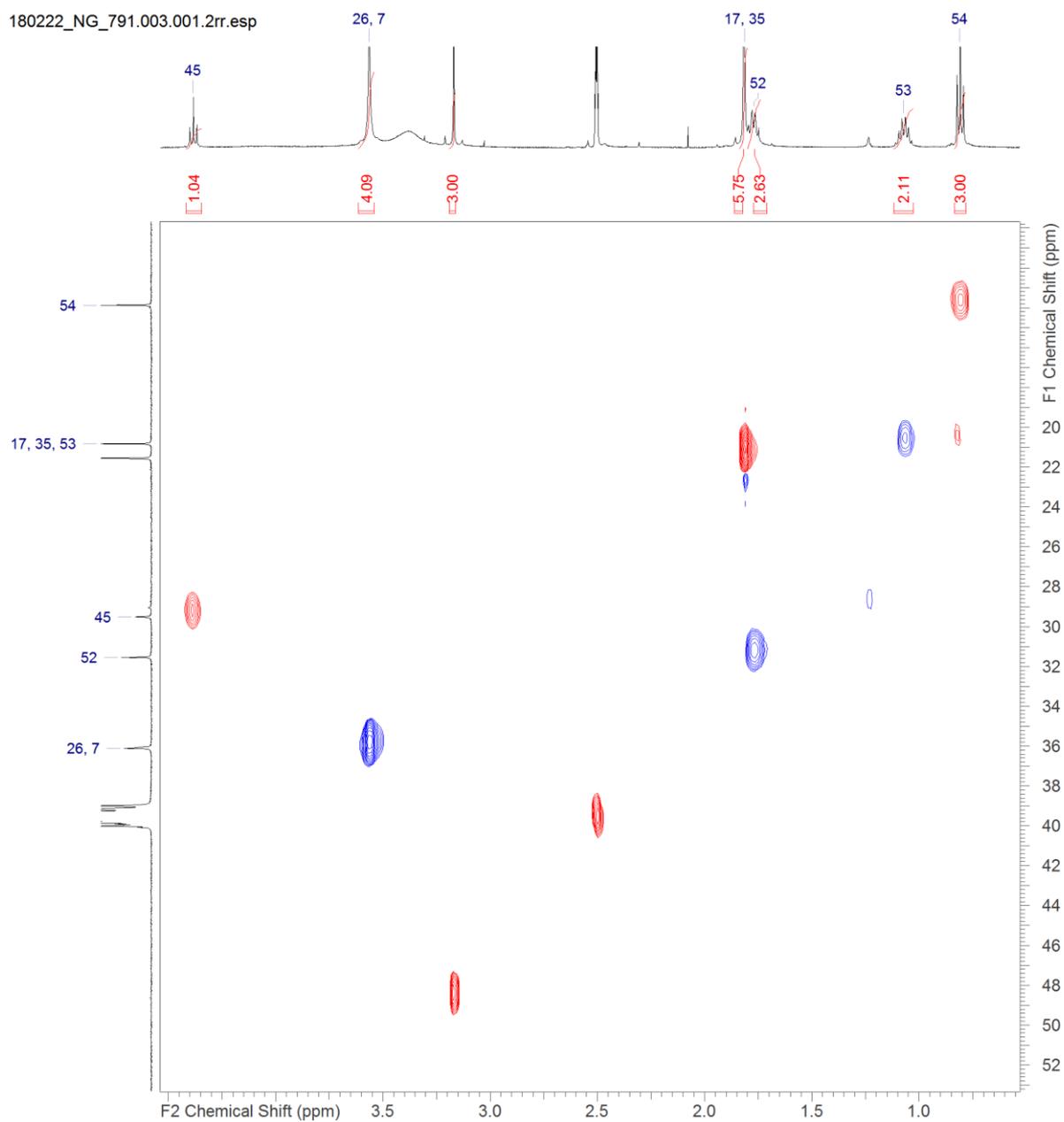


Figure S 44: HSQC-spectrum (500 MHz; 125 MHz, DMSO-d₆) of SEK90; zoom from 4 to 0,5 ppm and 53 to 16 ppm.

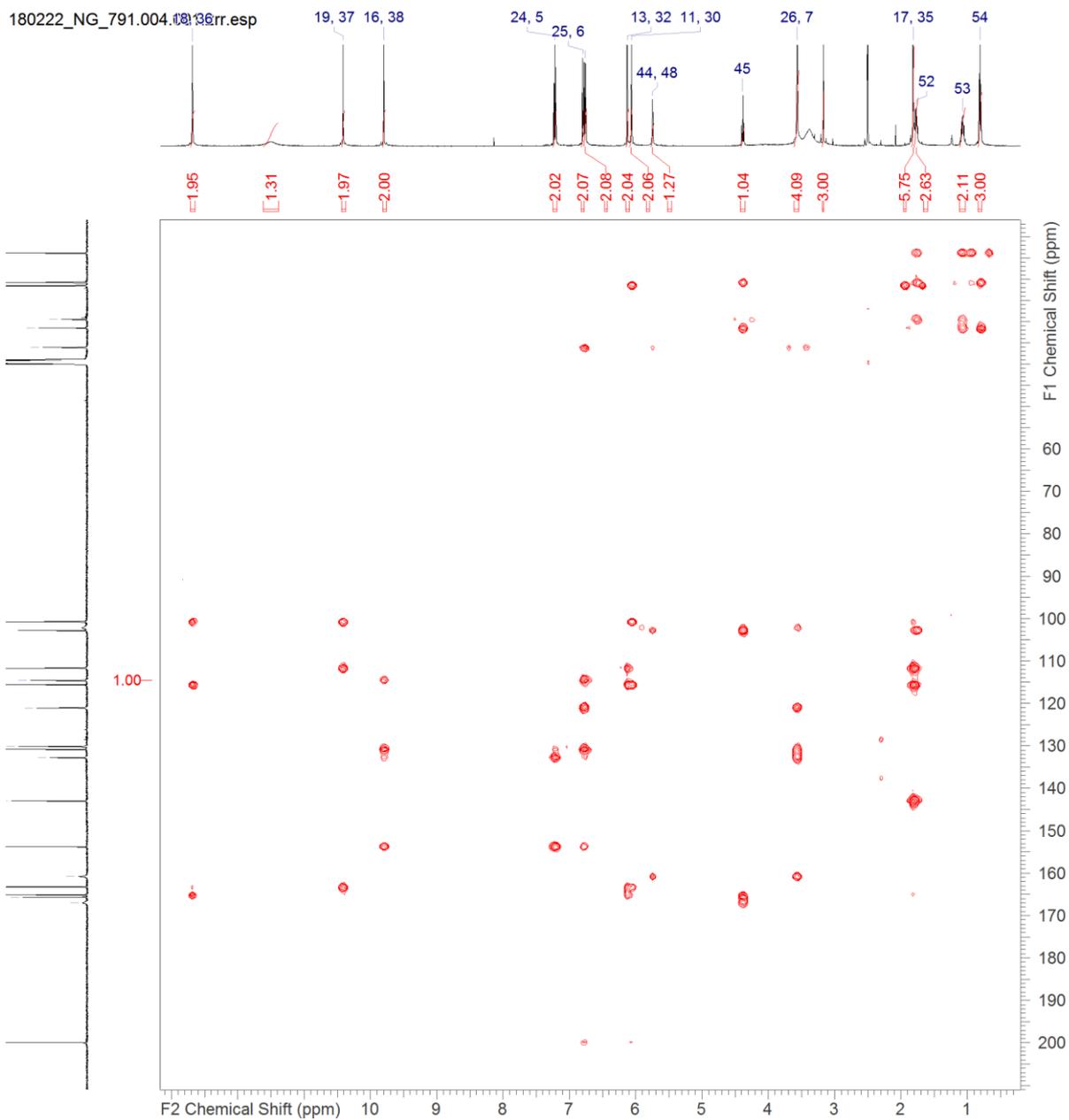


Figure S 45: HSQC-spectrum (500 MHz; 125 MHz, DMSO- d_6) of SEK90; zoom from 4 to 0.5 ppm and 53 to 16 ppm.

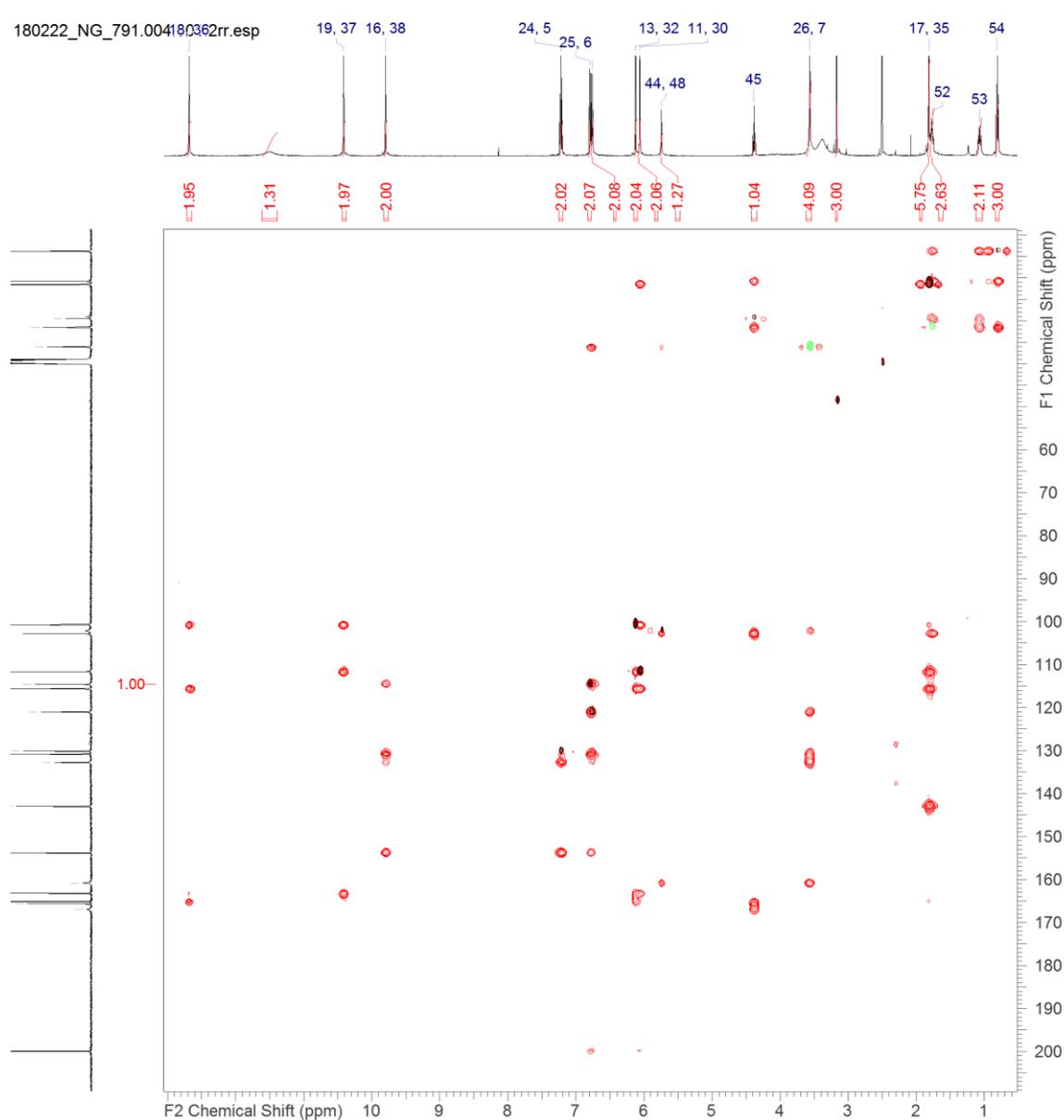
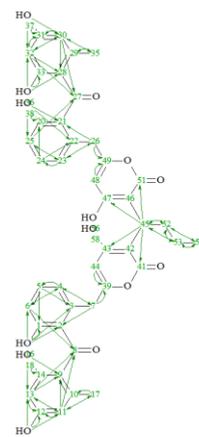


Figure S 46: HMBC - HSQC-Overlay spectrum (500 MHz; 125 MHz, DMSO-d₆) of SEK90 (red: HMBC; black/green: HSQC).

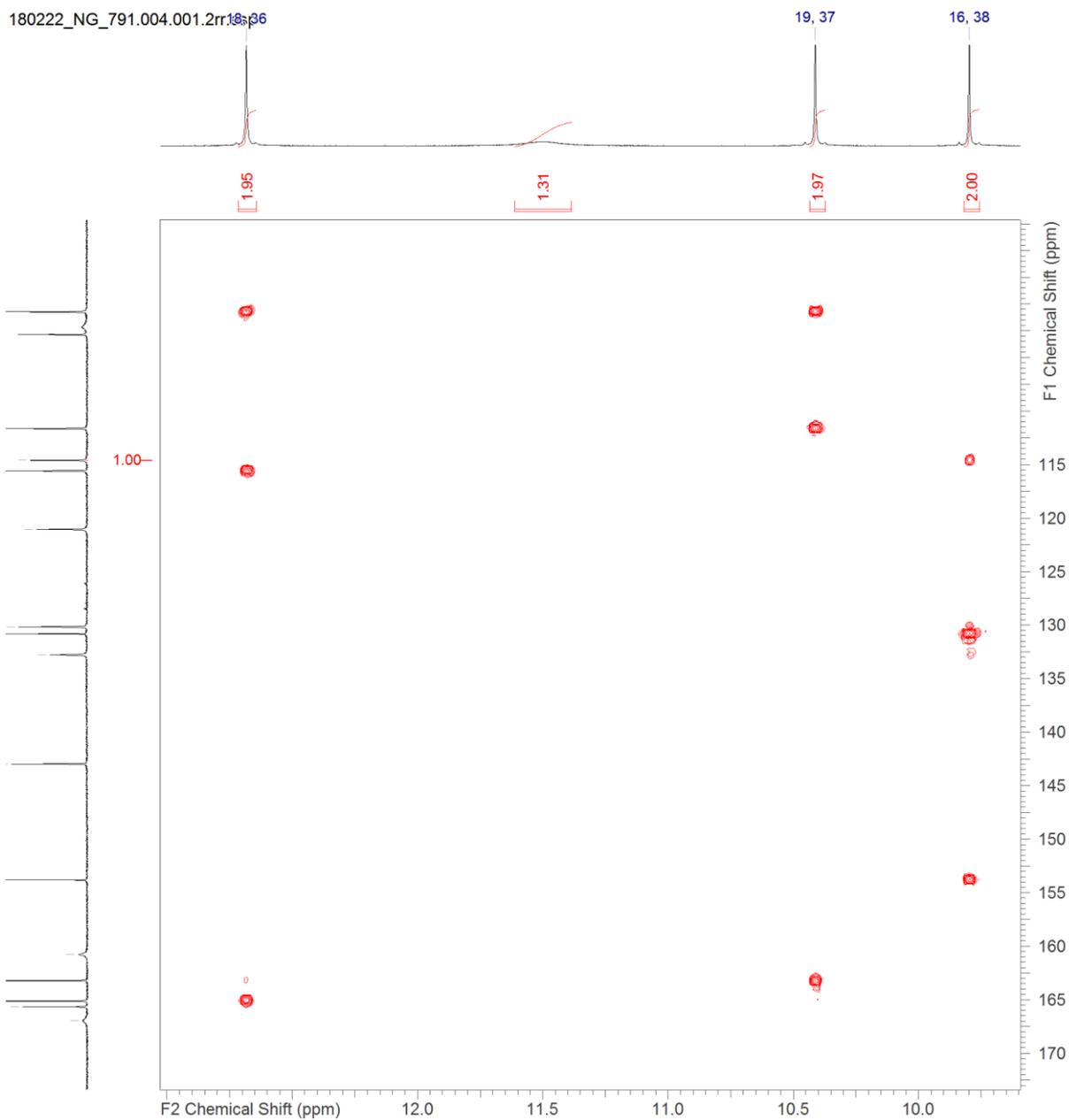


Figure S 47: HMBC - HSQC-Overlay spectrum (500 MHz; 125 MHz, DMSO-d₆) of SEK90 (black: HMBC; red/blue: HSQC); zoom from 13 to 10 ppm and 170 to 105 ppm.

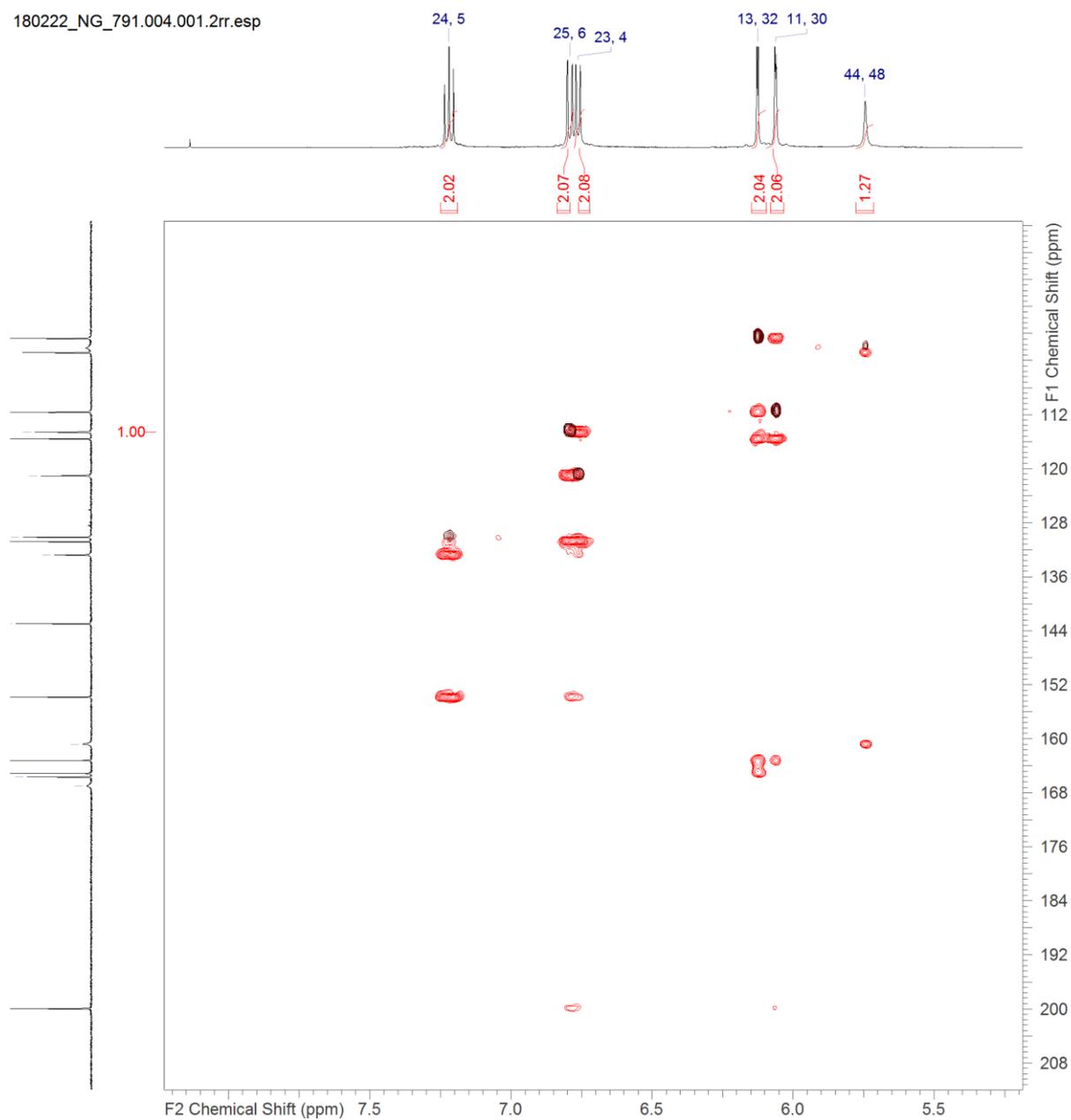


Figure S 48: HMBC - HSQC-Overlay spectrum (500 MHz; 125 MHz, DMSO-d₆) of SEK90 (black: HMBC; red/blue: HSQC); zoom from 13 to 10 ppm and 170 to 105 ppm.

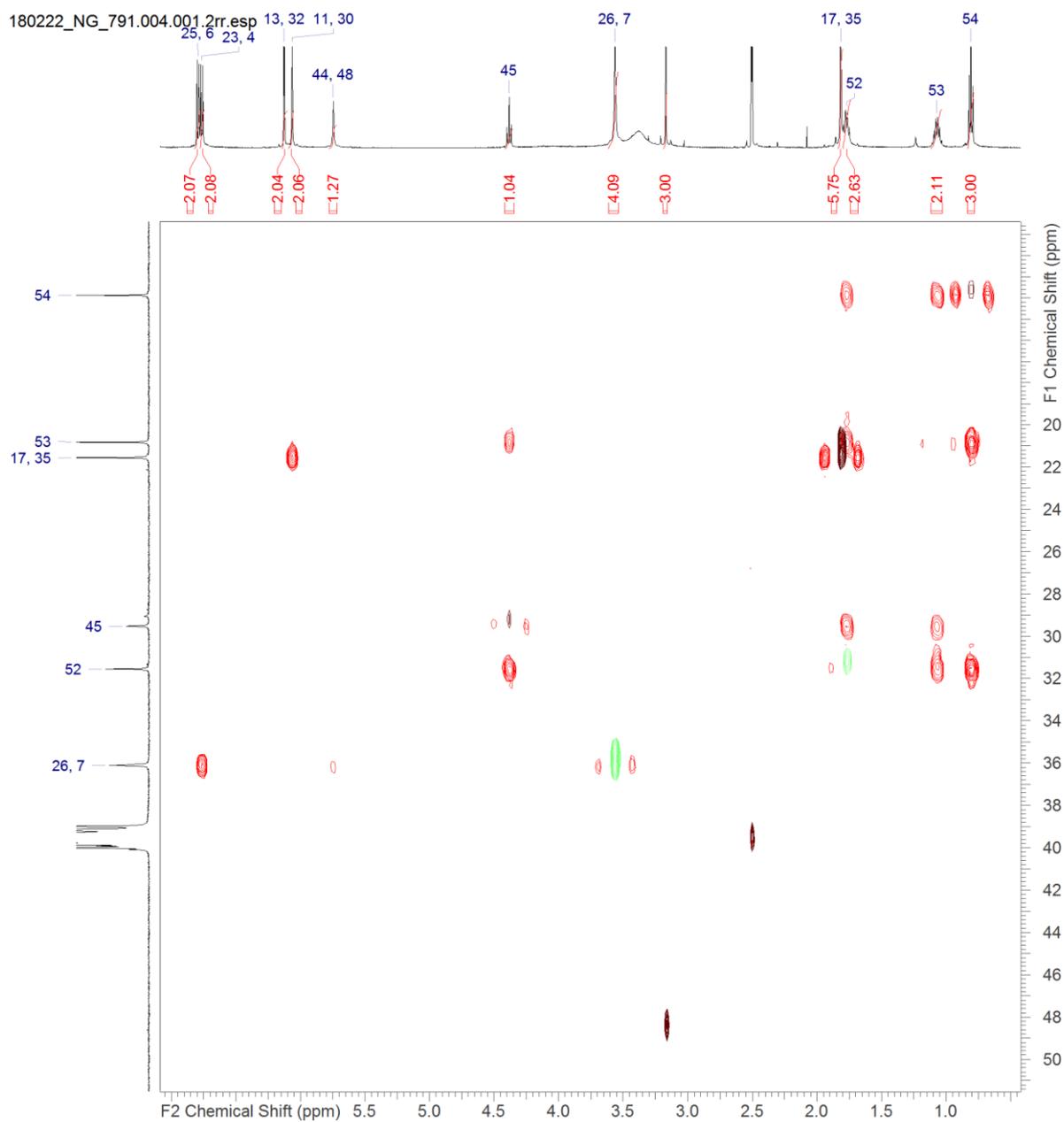


Figure S 49: HMBC - HSQC-Overlay spectrum (500 MHz; 125 MHz, DMSO- d_6) of SEK90 (black: HMBC; red/blue: HSQC); zoom from 6.5 to 0.5 ppm and 51 to 10 ppm.

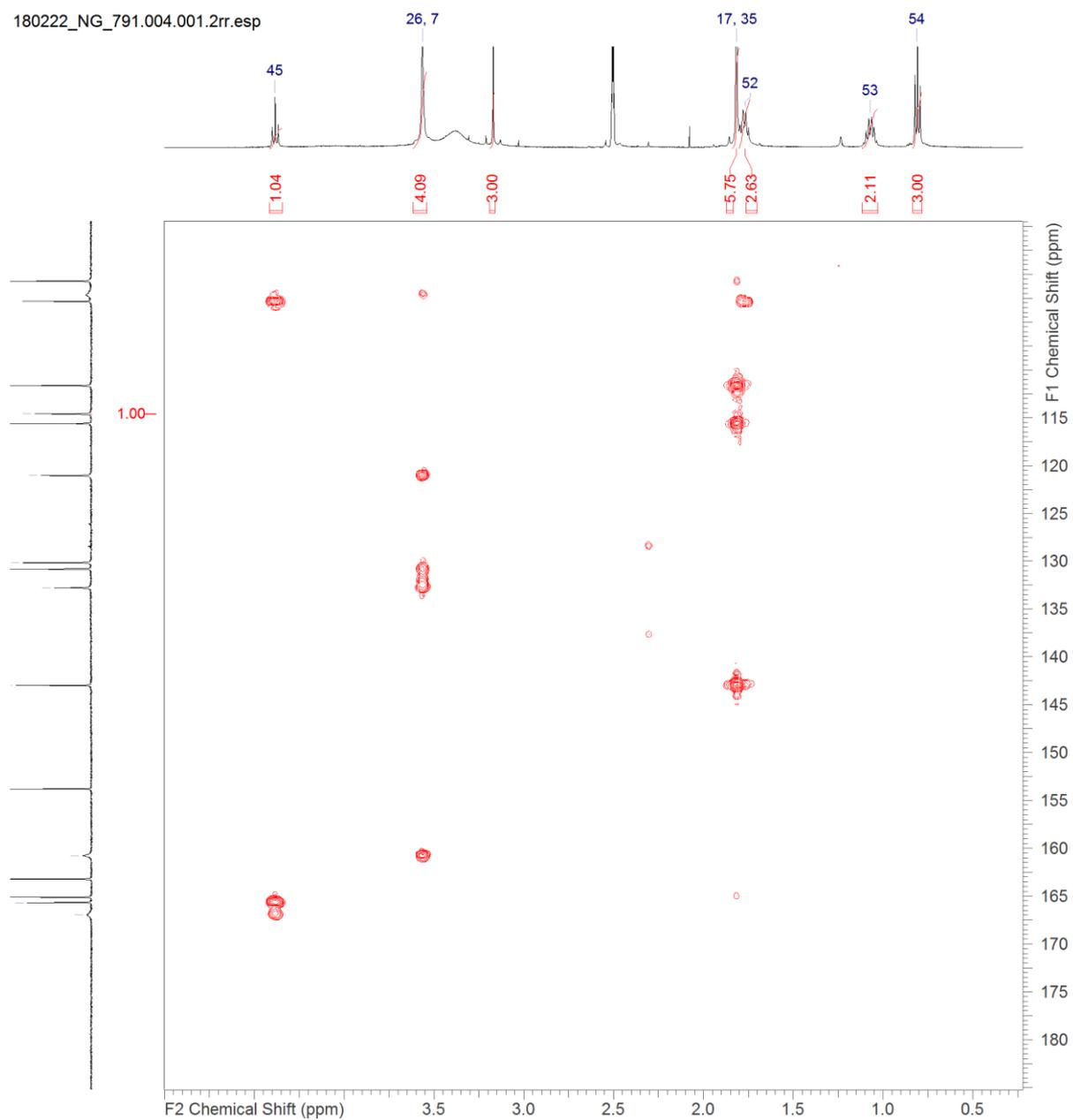


Figure S 50: HMBC - HSQC-Overlay spectrum (500 MHz; 125 MHz, DMSO- d_6) of SEK90 (black: HMBC; red/blue: HSQC); zoom from 4.5 to 0.5 ppm and 185 to 100 ppm.

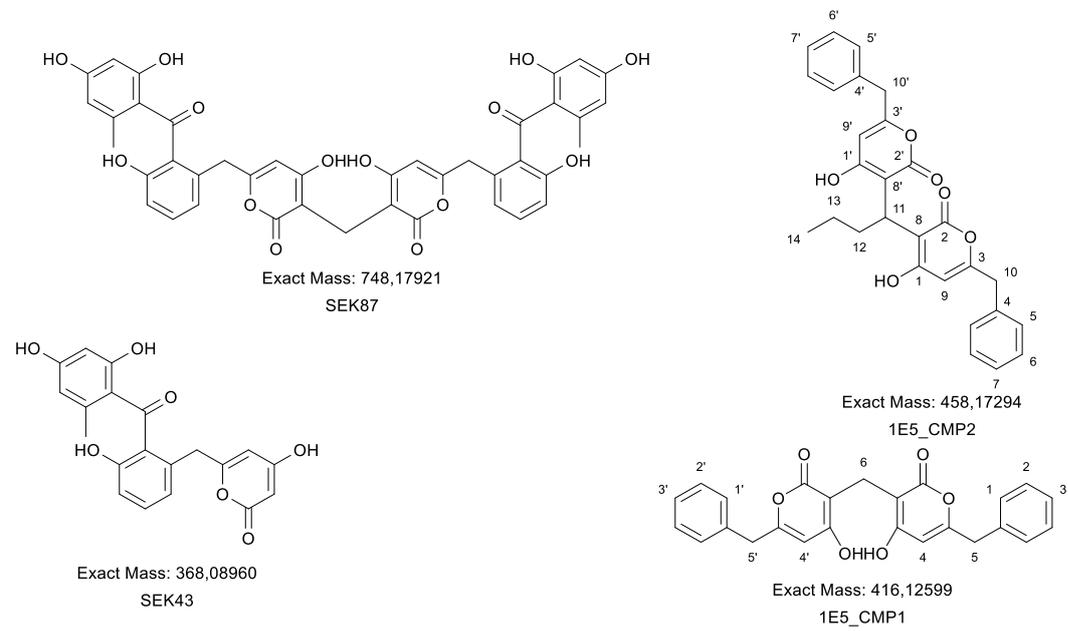


Figure S 51: Structures of SEK87, SEK43 and 1E5_CMP1 - 1E5_CMP2.

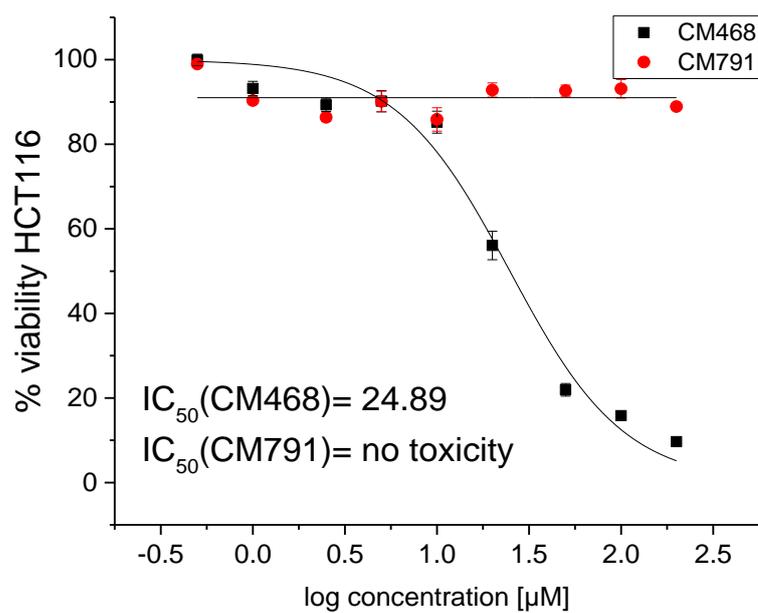


Figure S 52: Viability assay of **(1)** (CMP468/pentangumycin) and **(2)** (CMP791/SEK90) for HCT116 cells.

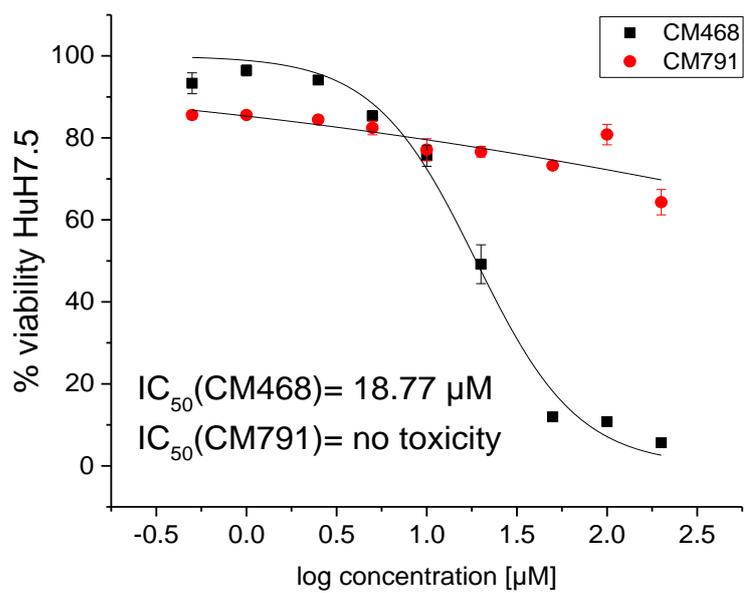


Figure S 53: Viability assay of **(1)** (CMP468/pentangumycin) and **(2)** (CMP791/SEK90) for HuH7.5 cells.

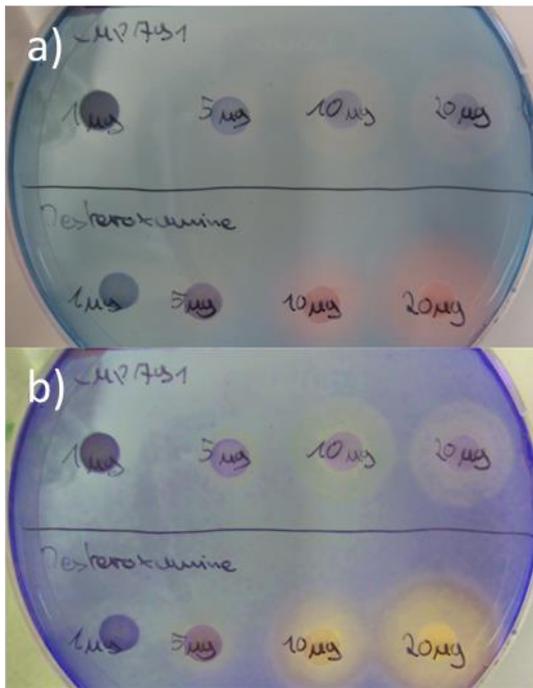


Figure S 54: CAS-Assay for SEK90 (CMP791) on agar diffusion discs compared to desferrioxamine; from left to right: 1 μ g, 5 μ g, 10 μ g, 20 μ g a): Original picture; b) Enhanced colors for better visibility.

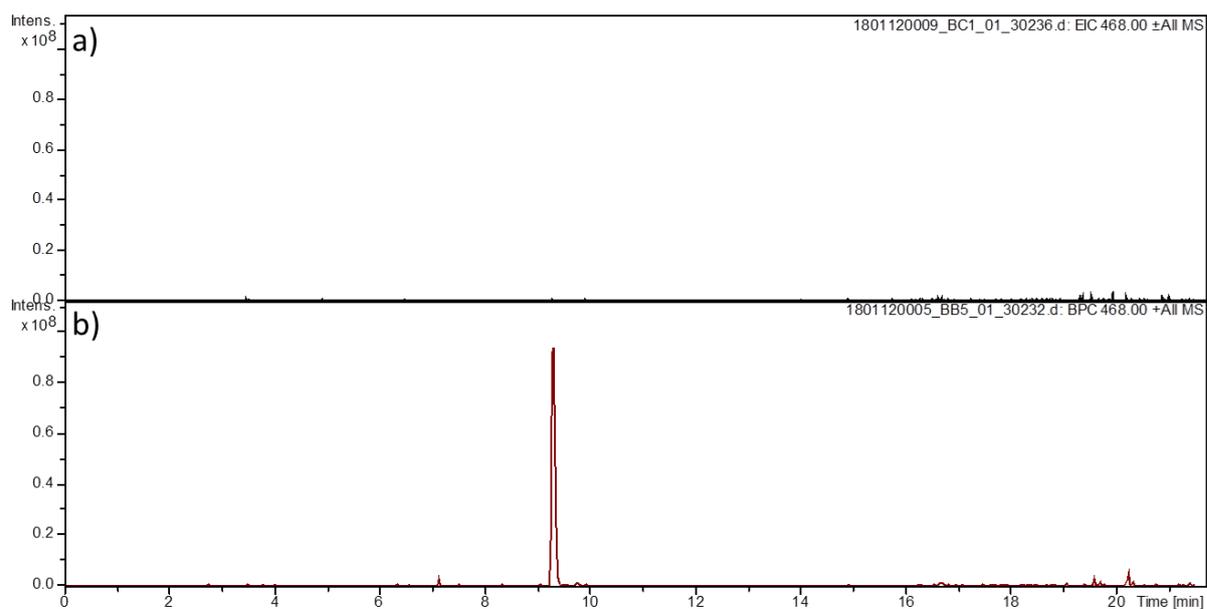


Figure S 55: HPLC-MS Extracted ion chromatogram (Extracted mass 468 ± 0.5) of a) *S. lividans* Δ YA6_1E5 Δ 54860 and b) *S. lividans* Δ YA6_1E5 Δ 55110; The production is unchanged in the deletion of the NRPS/PKS I hybrid core gene, while the production is abolished in the created mutant lacking the type II PKS gene.

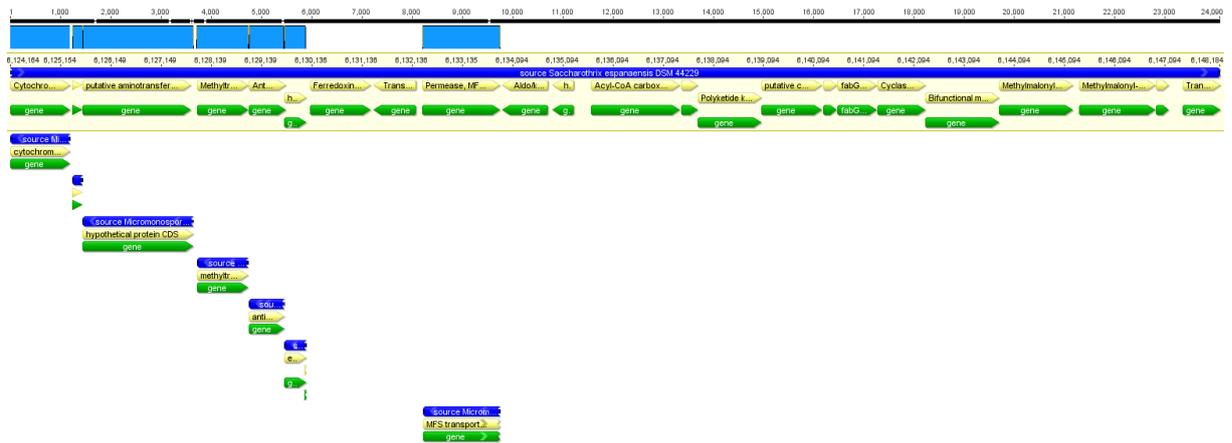


Figure S 56: Genes (WP_124773686 & WP_124773691-WP124773694) of *Micromonospora* sp. LB39 mapped to pentangumycins biosynthetic gene cluster.

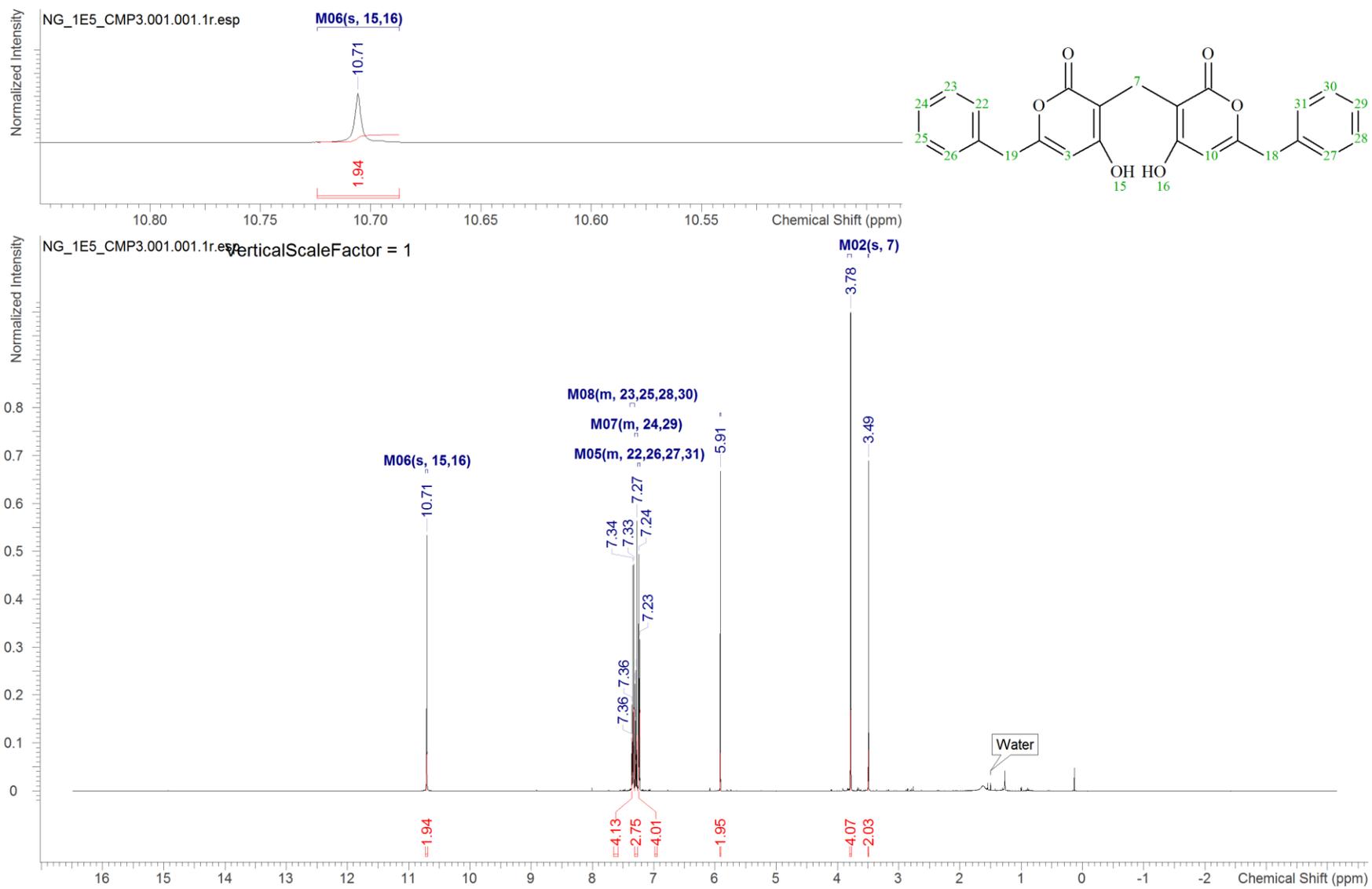


Figure S 57: ¹H-NMR spectrum (500 MHz, CDCl₃) of 1E5_CMP1; zoom from 10.8 ppm to 10.5 ppm.

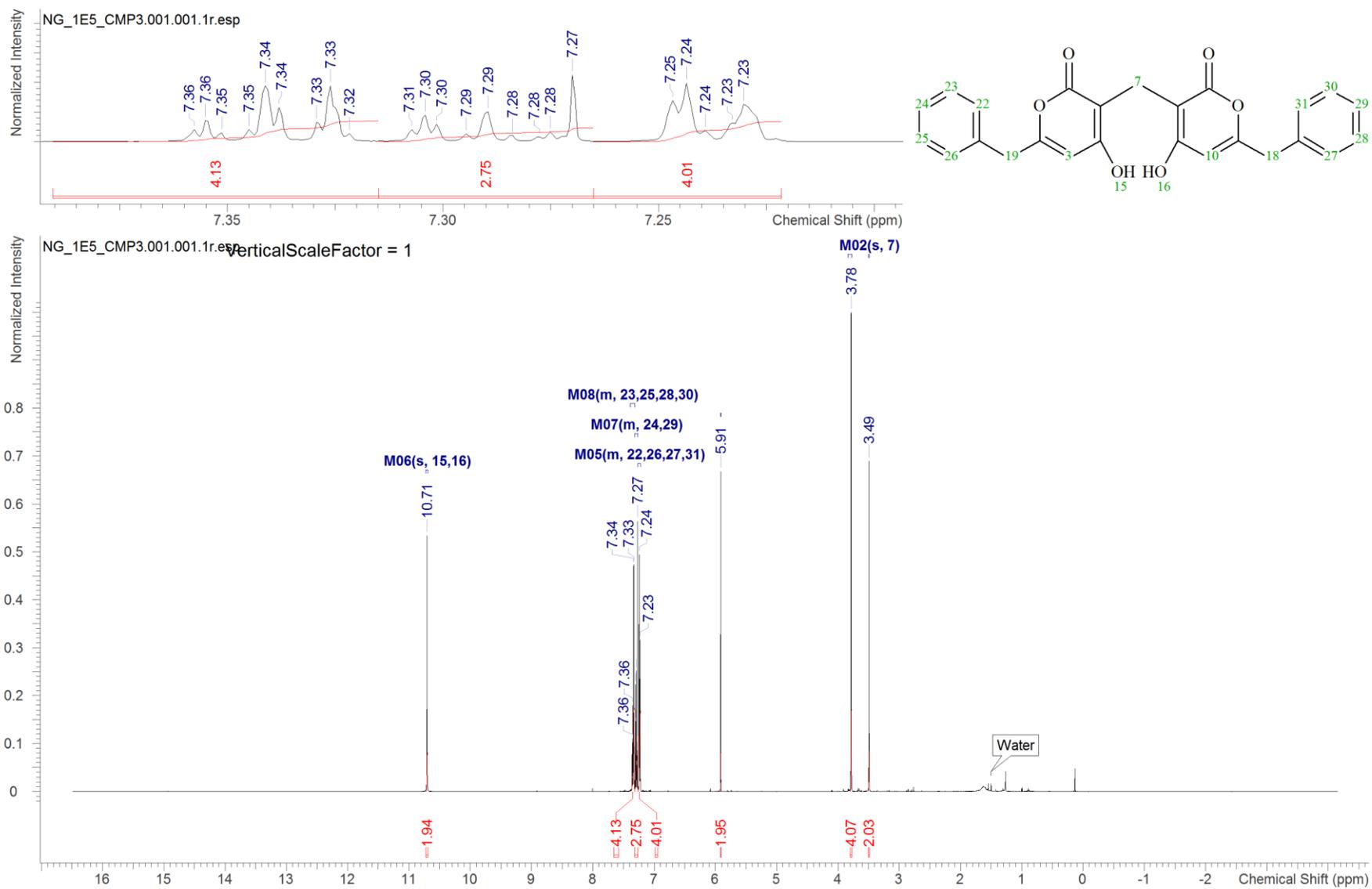


Figure S 58: ¹H-NMR spectrum (500 MHz, CDCl₃) of 1E5_CMP1; zoom from 7.4 ppm to 7.2 ppm.

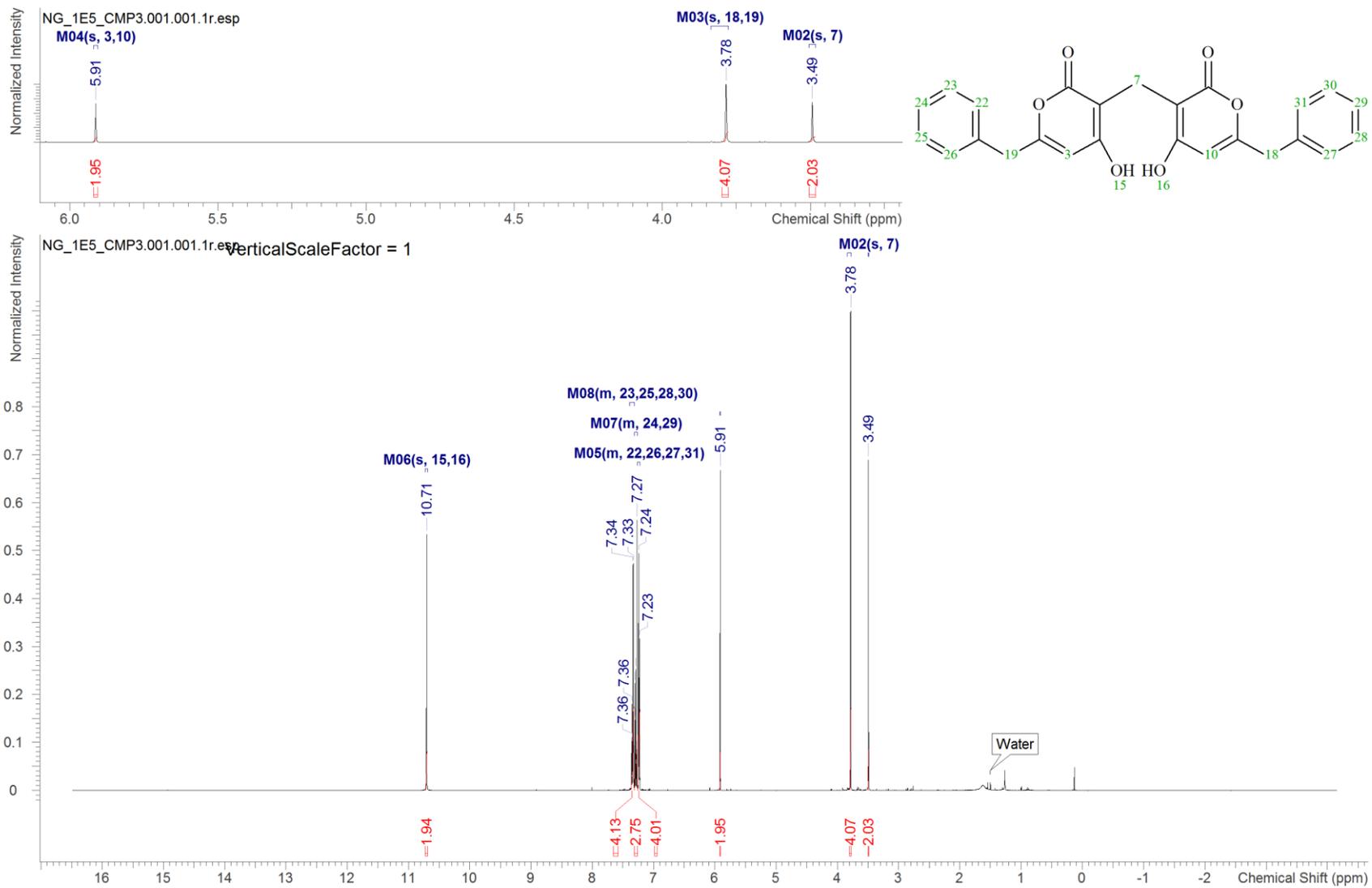


Figure S 59 $^1\text{H-NMR}$ spectrum (500 MHz, CDCl_3) of 1E5_CMP1; zoom from 6 ppm to 3.5 ppm

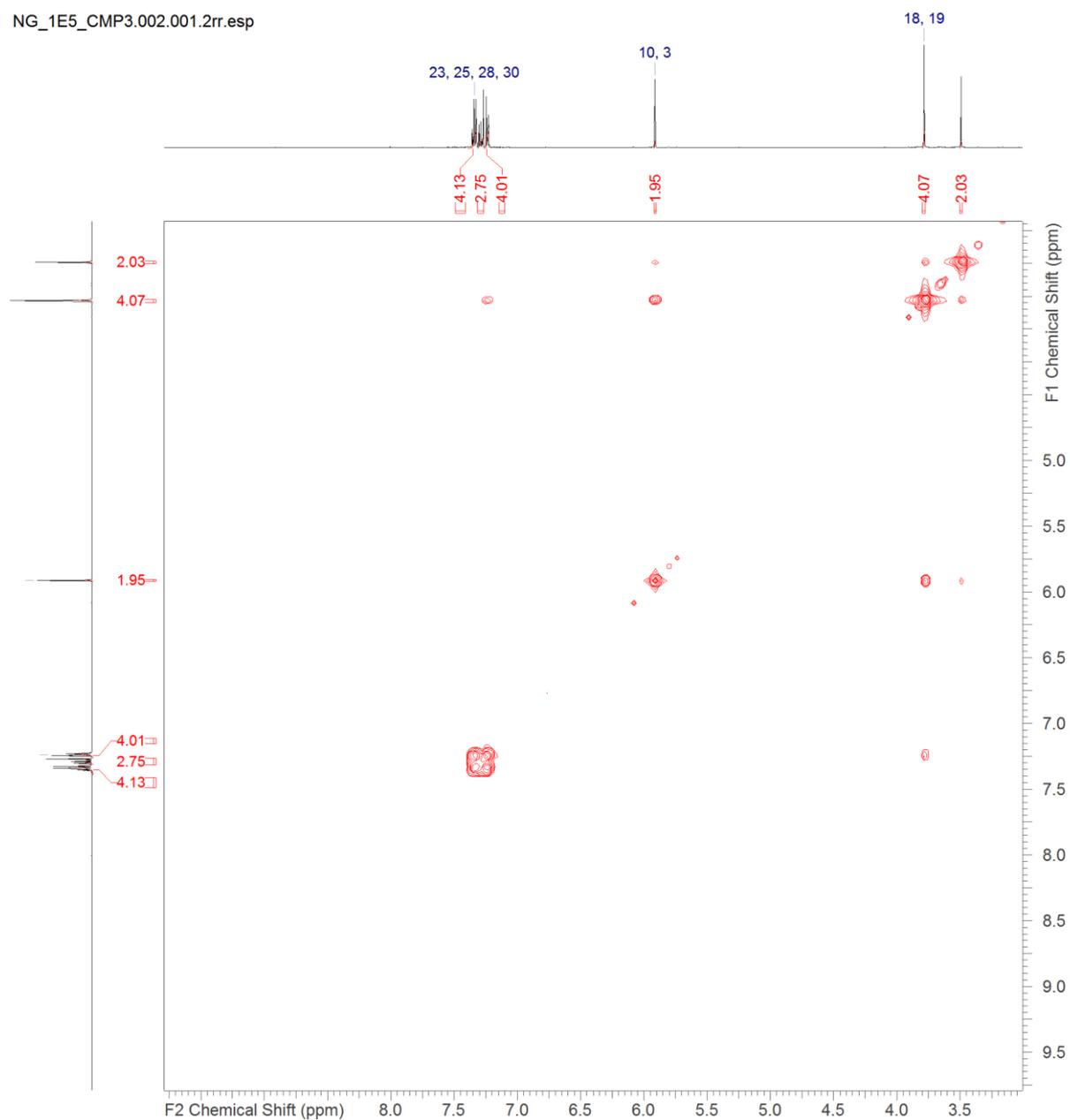


Figure S 60: ^1H - ^1H - COSY spectrum (500 MHz, CDCl_3) of 1E5_CMP1.

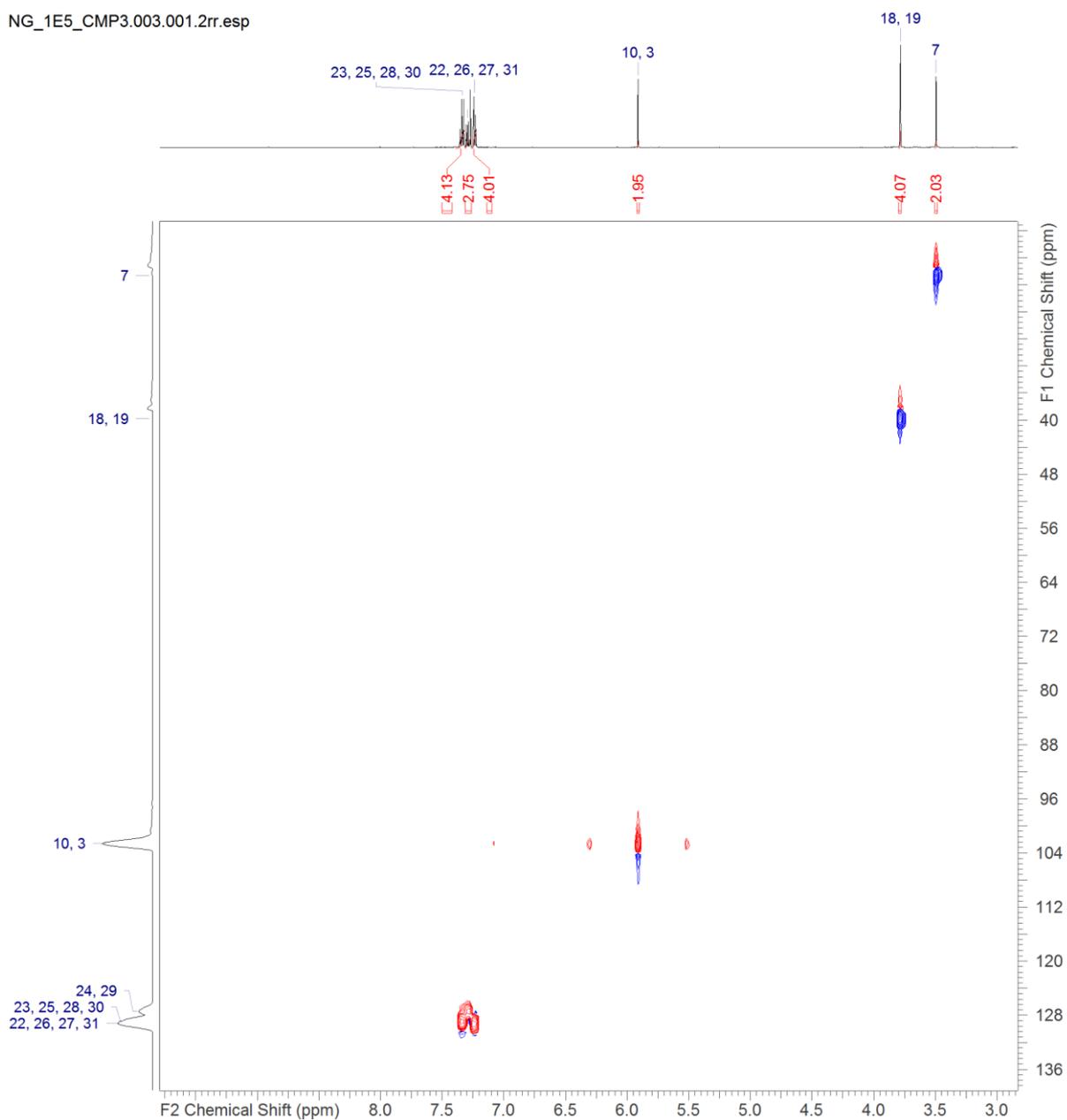


Figure S 61: HSQC-spectrum (500 MHz; 125 MHz, CDCl₃) of 1E5_CMP1.

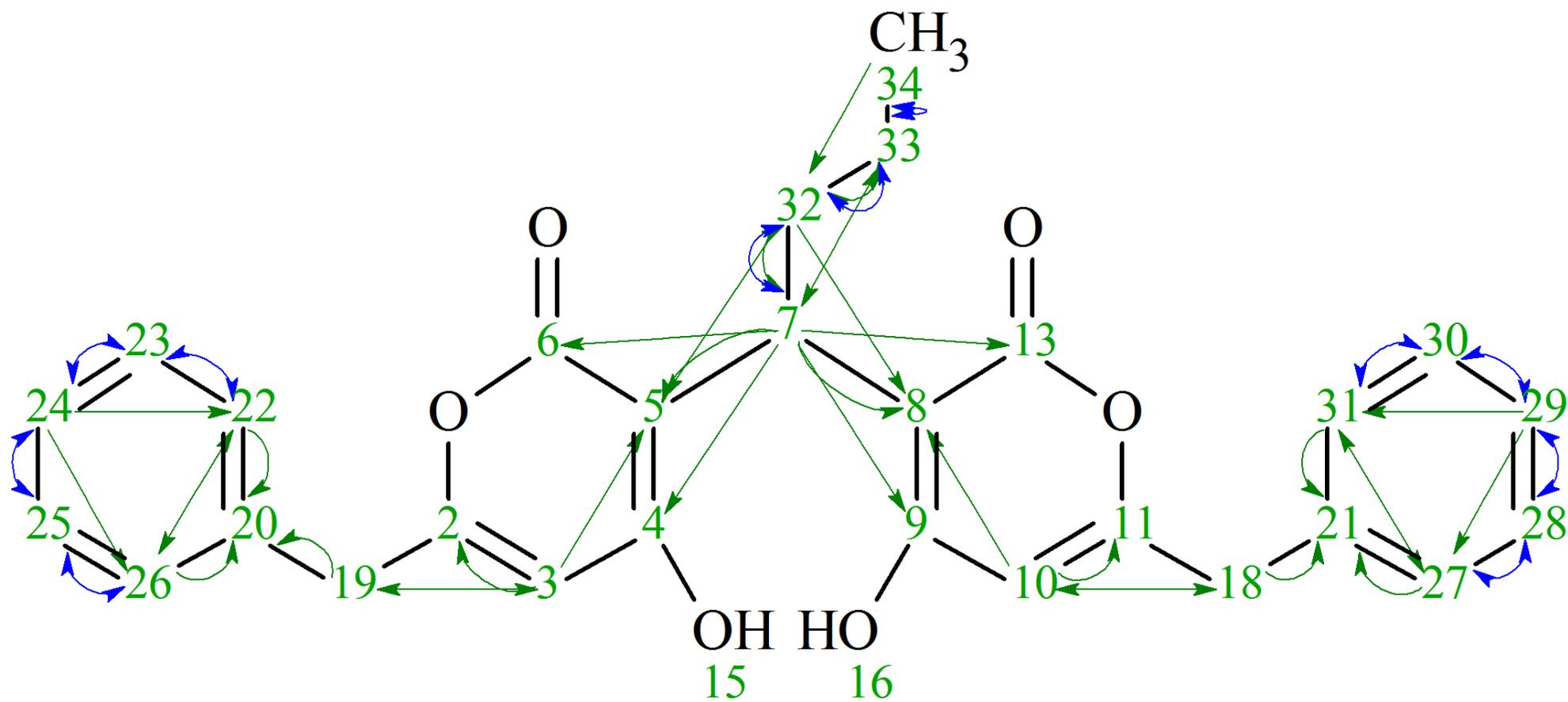


Figure S 62: Structure of 1E5_CMP2 with all observed correlations (HMBC: green C → H; ¹H-¹H-Cosy: blue H → H).

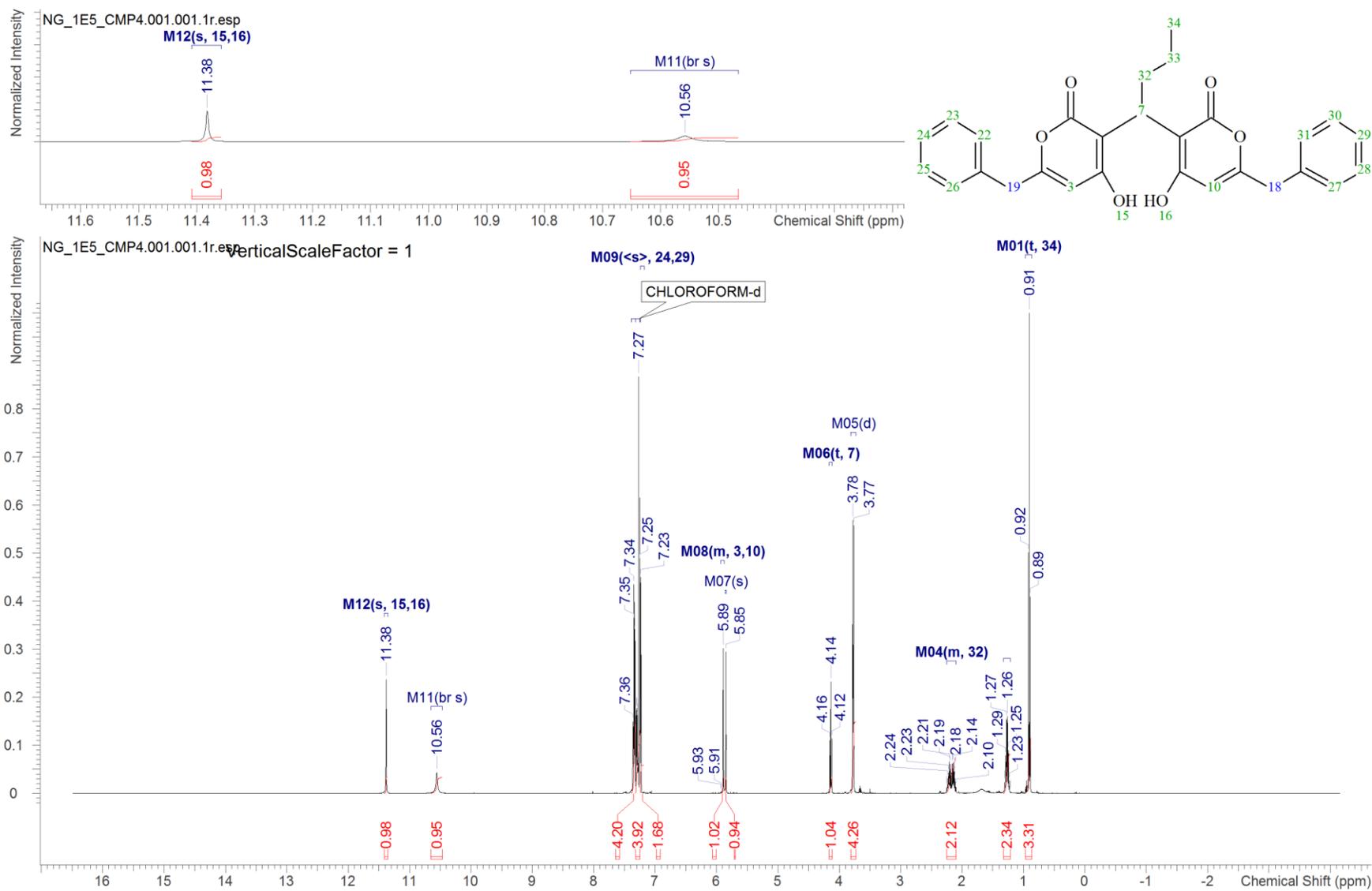


Figure S 63: $^1\text{H-NMR}$ spectrum (500 MHz, CDCl_3) of 1E5_CMP2; zoom from 11.6 to 9.5 ppm.

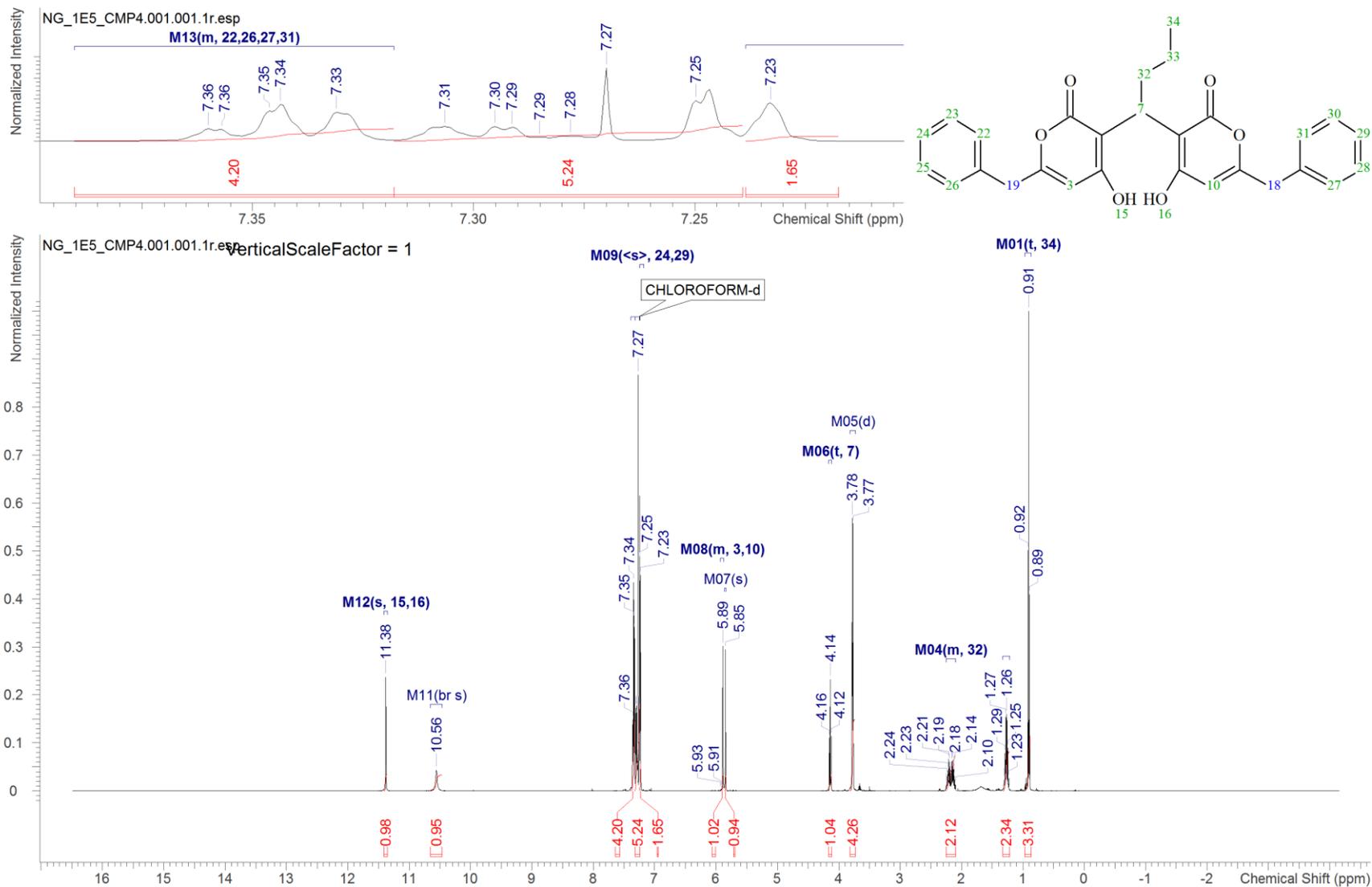


Figure S 64: $^1\text{H-NMR}$ spectrum (500 MHz, CDCl_3) of 1E5_CMP2; zoom from 7.4 to 7.2 ppm.

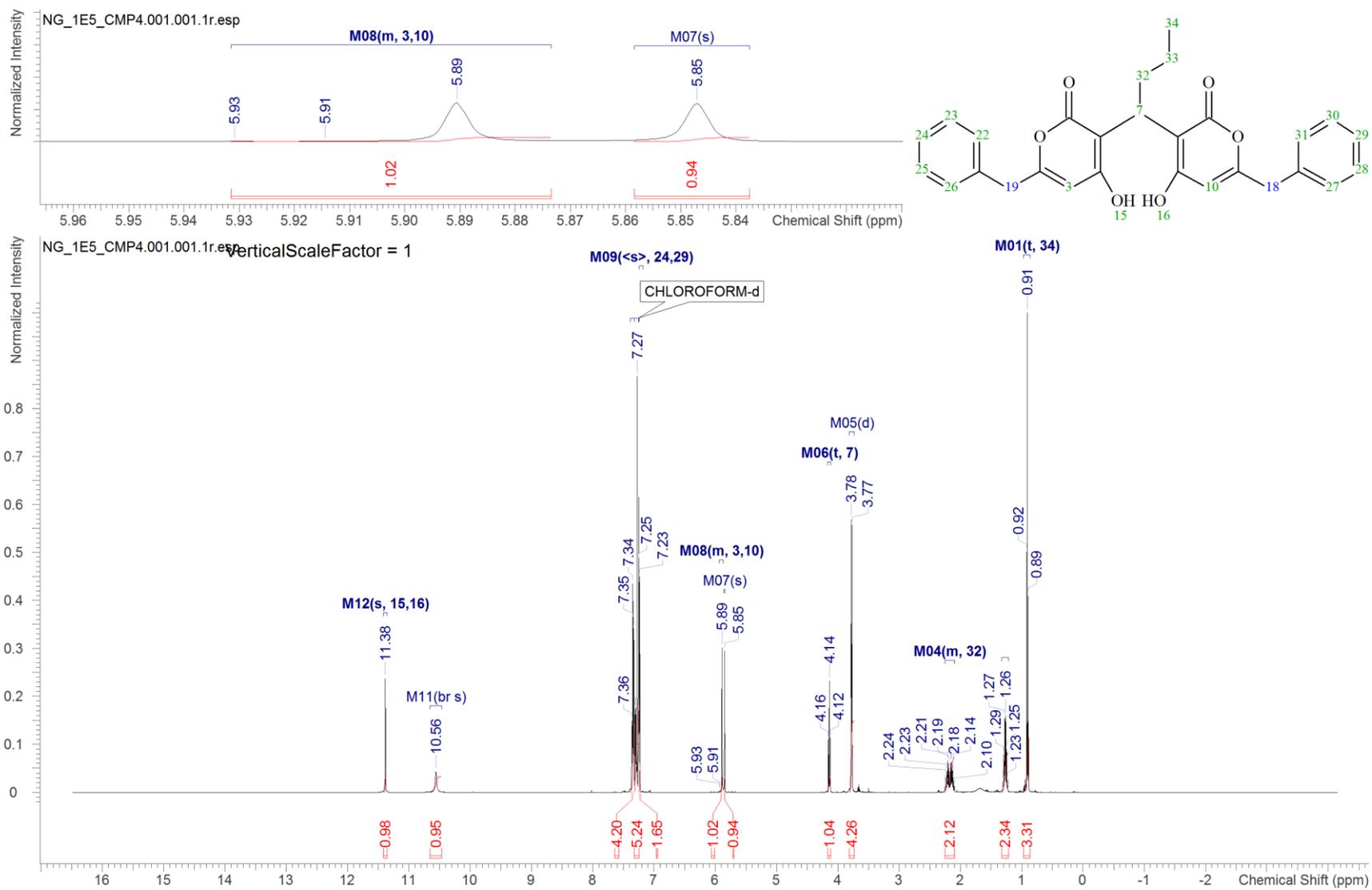


Figure S 65: $^1\text{H-NMR}$ spectrum (500 MHz, CDCl_3) of 1E5_CMP2; zoom from 5.96 to 5.8 ppm.

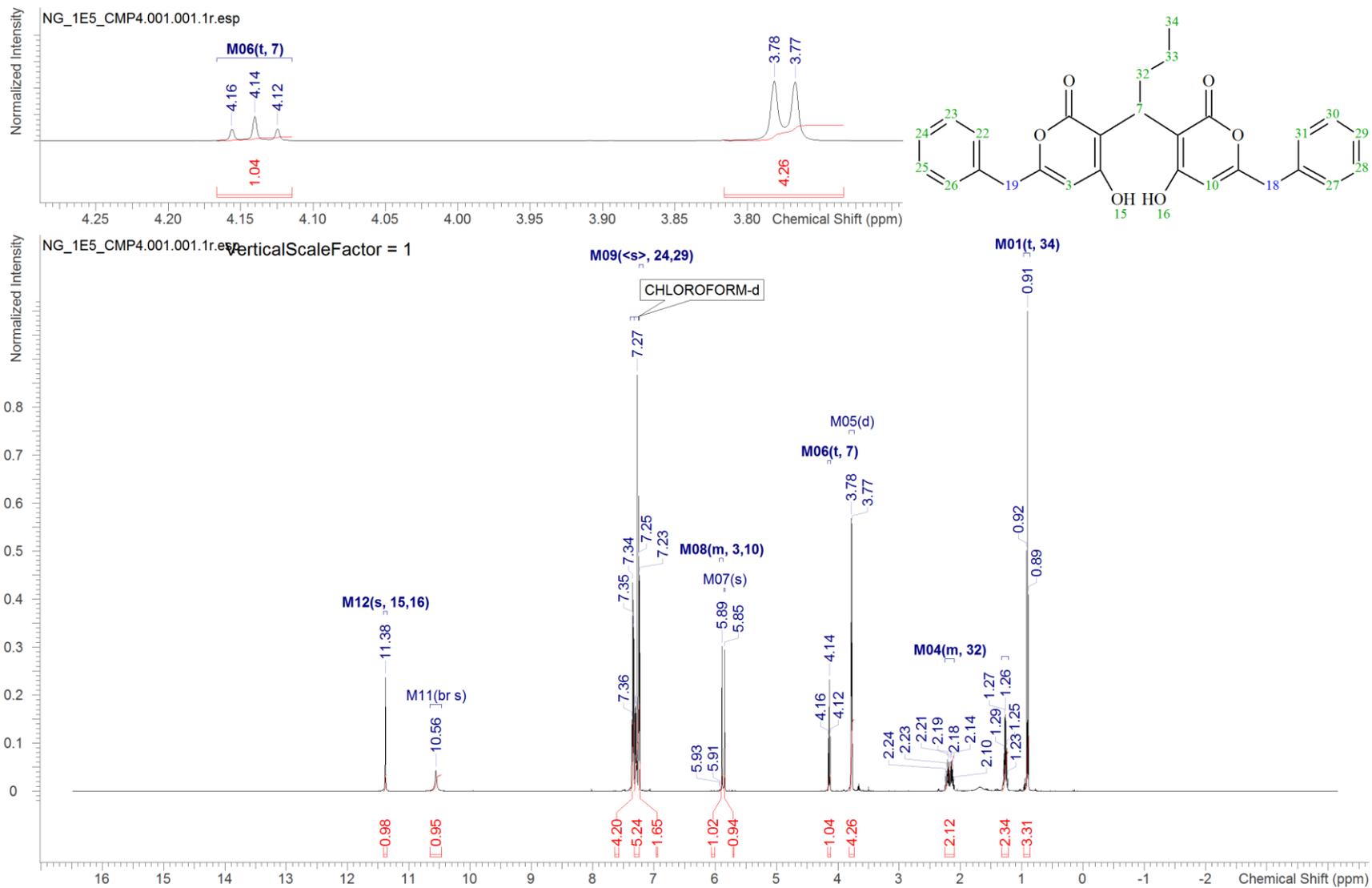


Figure S 66: $^1\text{H-NMR}$ spectrum (500 MHz, CDCl_3) of 1E5_CMP2; zoom from 4.25 to 3.75 ppm.

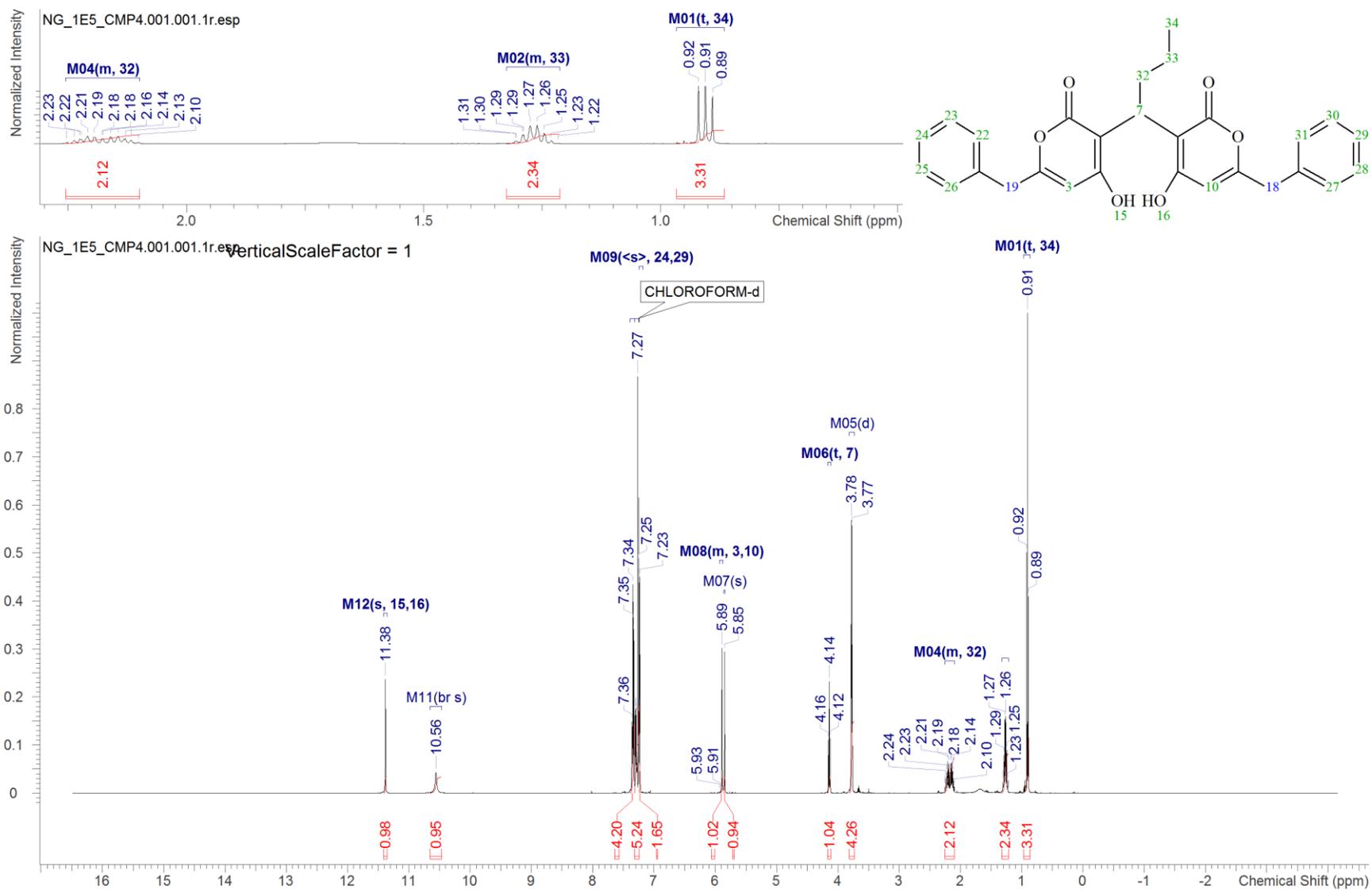


Figure S 67: $^1\text{H-NMR}$ spectrum (500 MHz, CDCl_3) of 1E5_CMP2; zoom from 2.5 to 0.5 ppm.

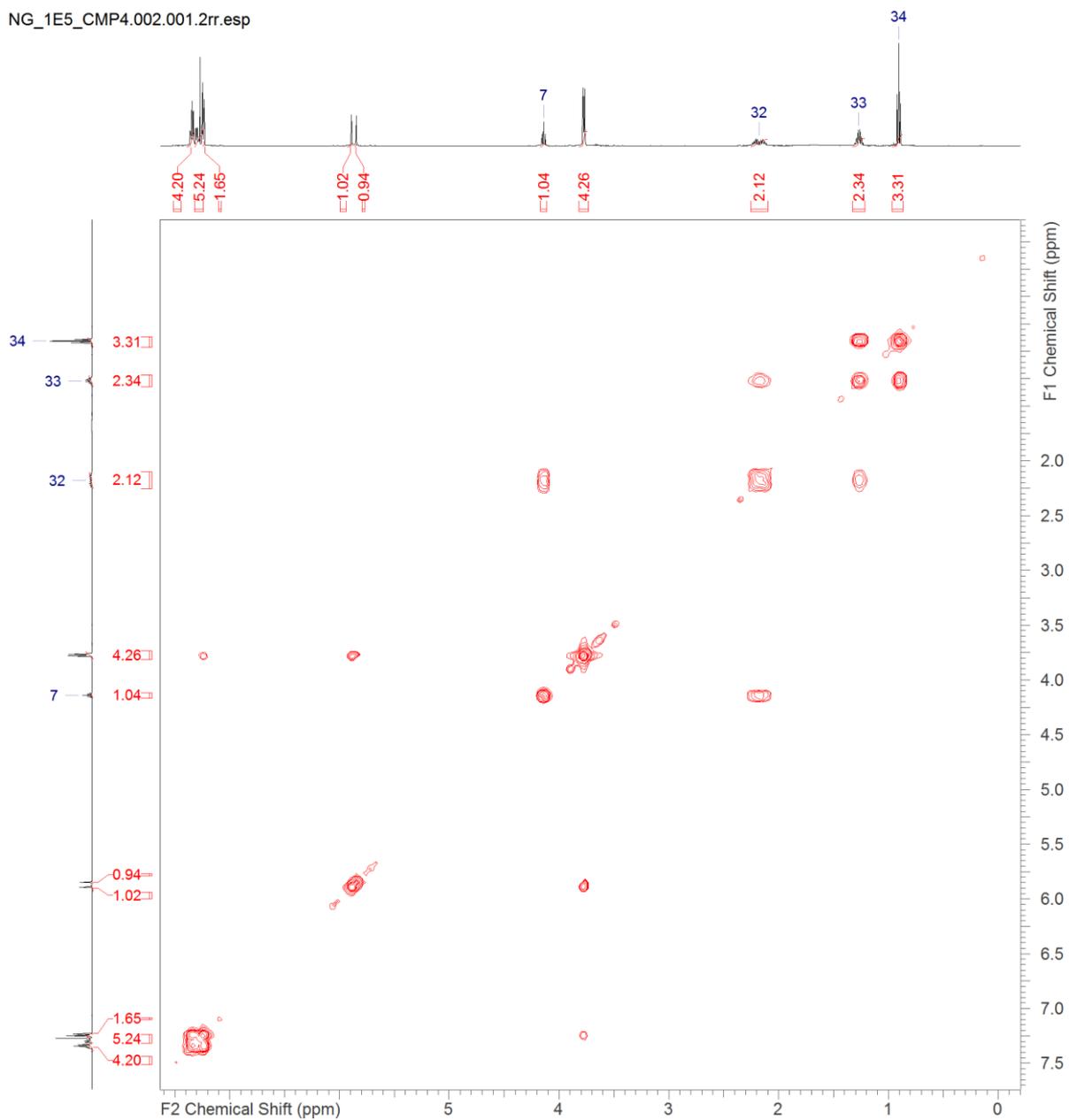


Figure S 68: ^1H - ^1H - COSY spectrum (500 MHz, CDCl_3) of 1E5_CMP2.

NG_1E5_CMP4.003.001.2rr.esp

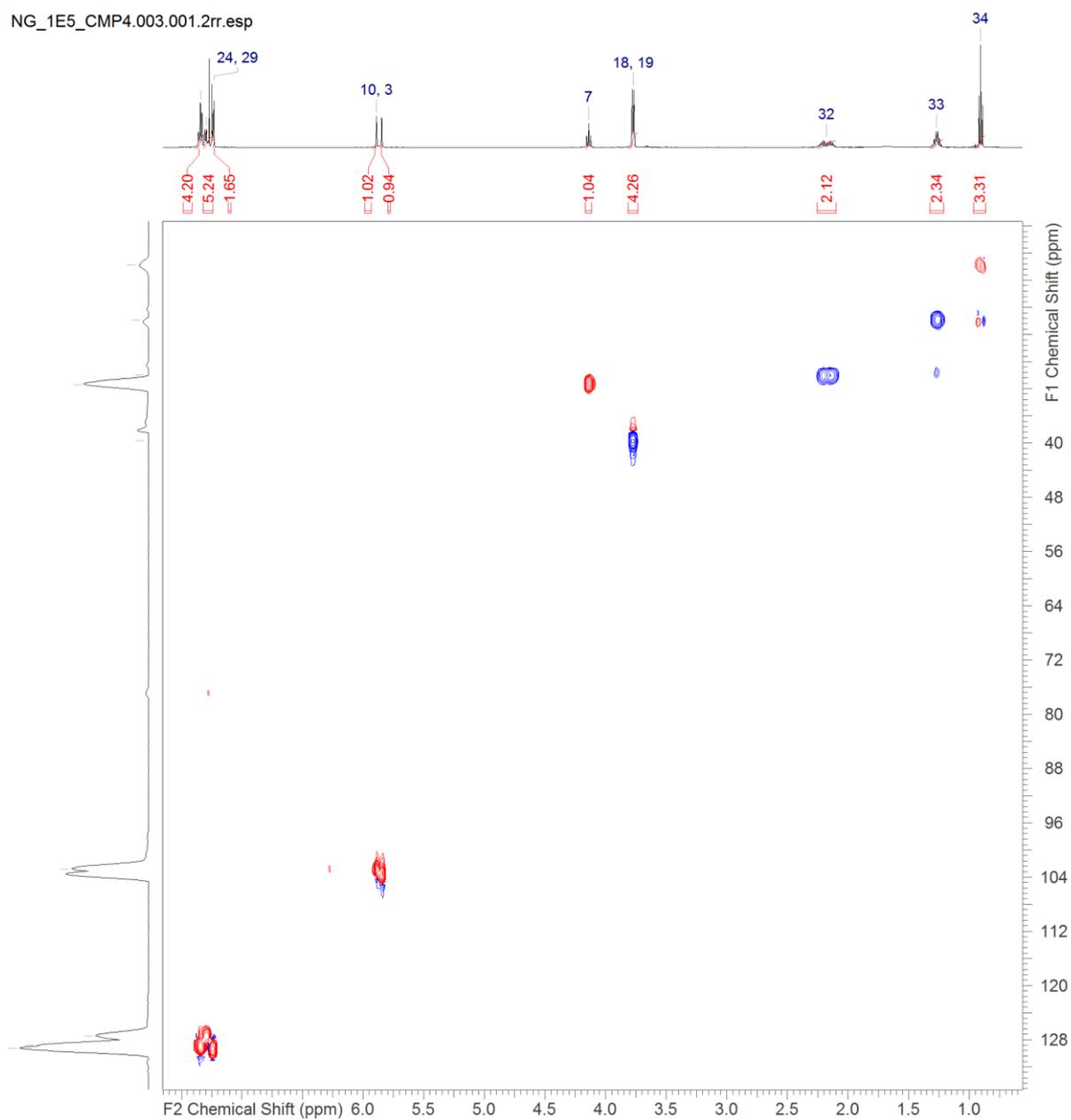


Figure S 69: HSQC-spectrum (500 MHz; 125 MHz, CDCL₃) of 1E5_CMP2.

NG_1E5_CMP4.004.001.2rr.esp

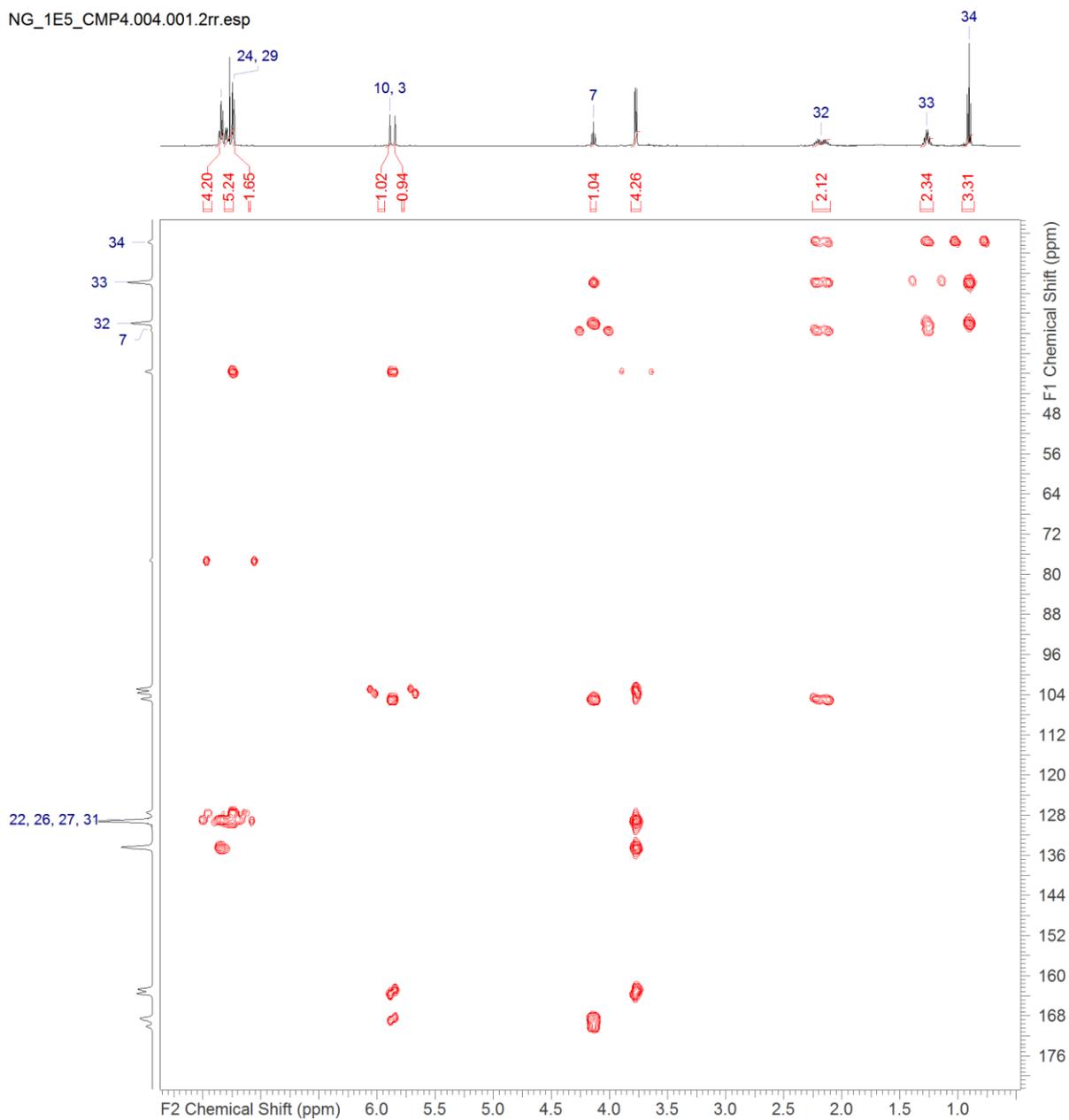


Figure S 70: HMBC-spectrum (500 MHz; 125 MHz, CDCl₃) of 1E5_CMP2.

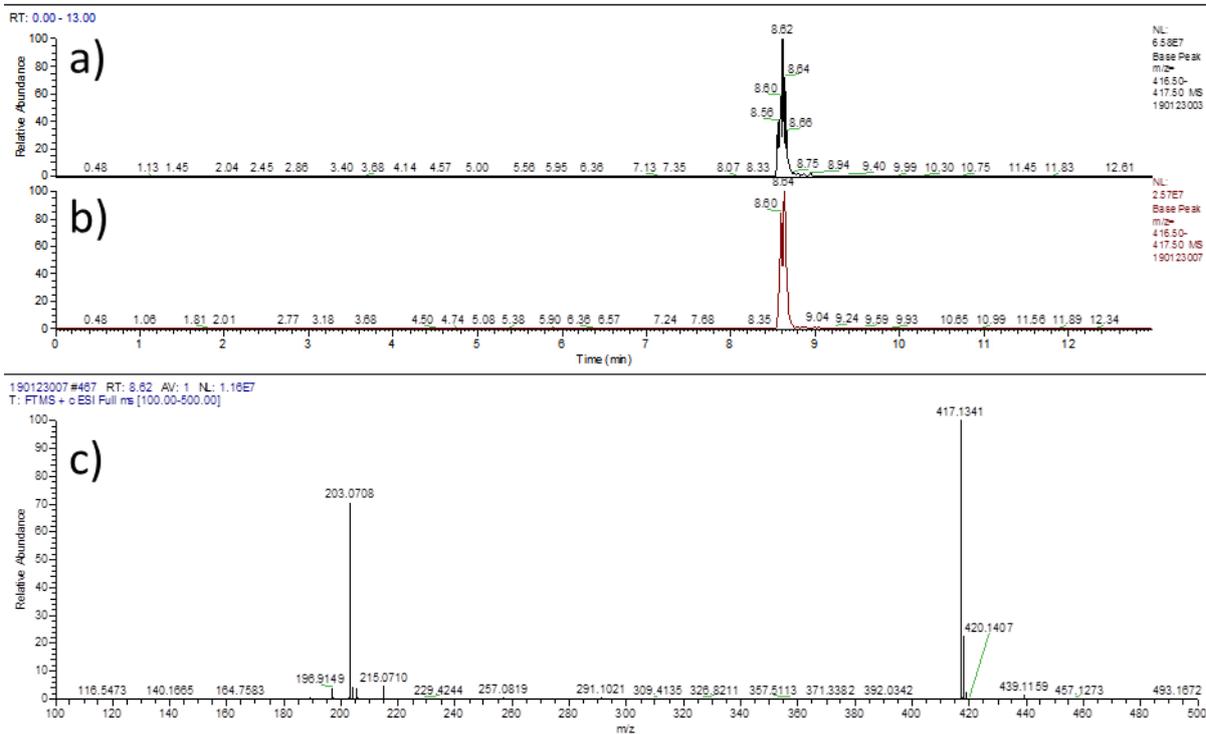


Figure S 71: HPLC-MS Extracted ion chromatogram (Extracted mass 417 ± 0.5) of a) Spontaneously formed 1E5_CMP1; b) synthesized 1E5_CMP1; c) ESI full ms chromatogram of a).

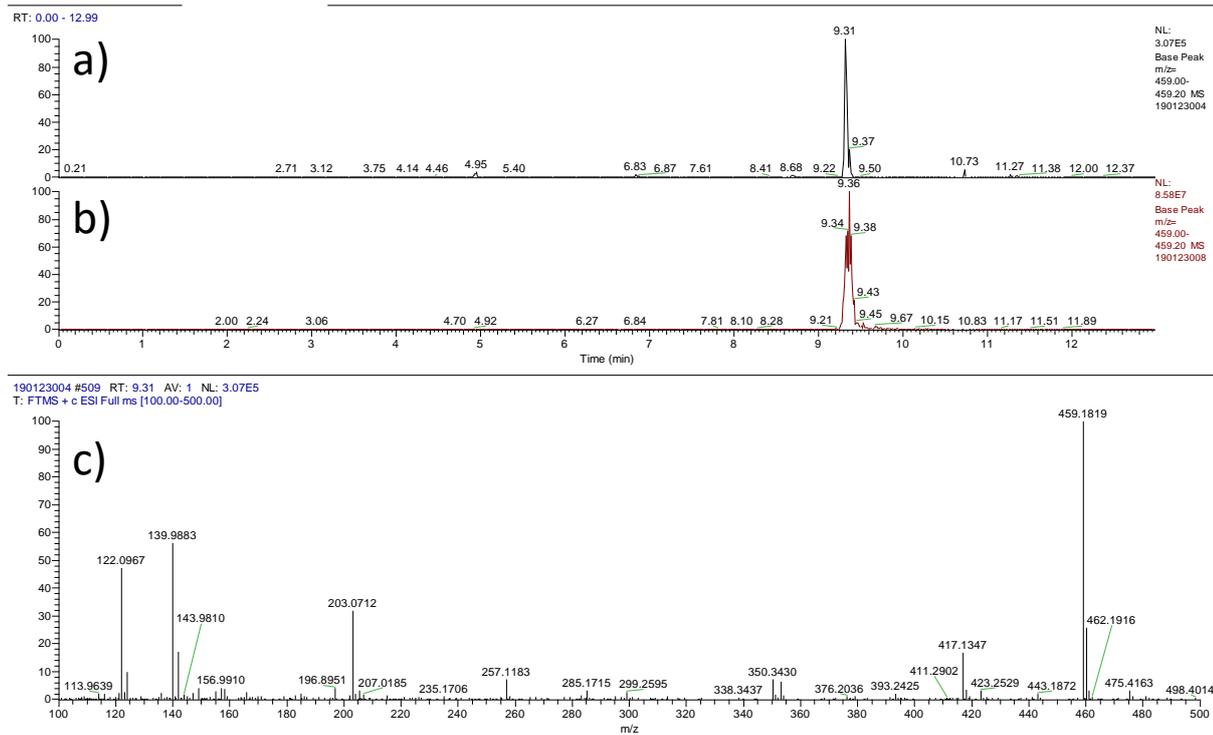


Figure S 72: HPLC-MS Extracted ion chromatogram (Extracted mass 459.1 ± 0.1) of a) Spontaneously formed 1E5_CMP2; b) synthesized 1E5_CMP2; c) Esi full ms chromatogram of a).

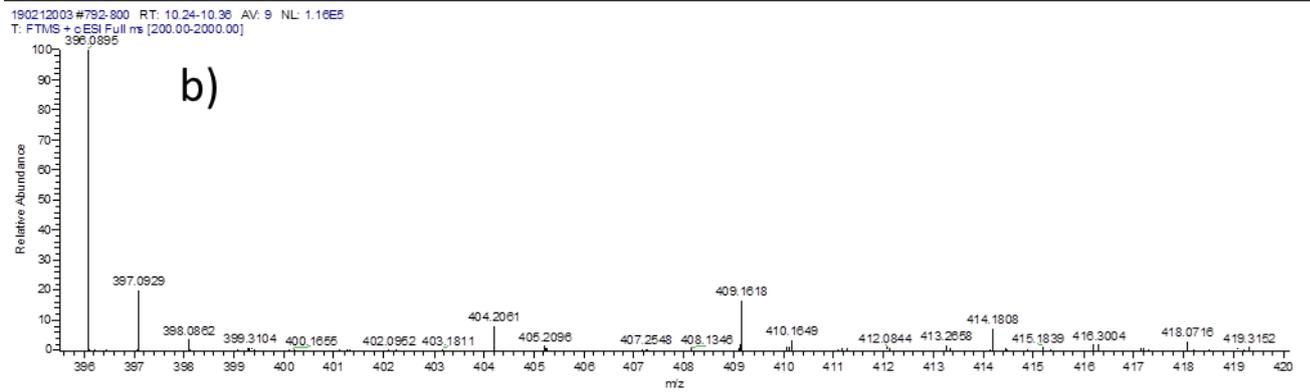
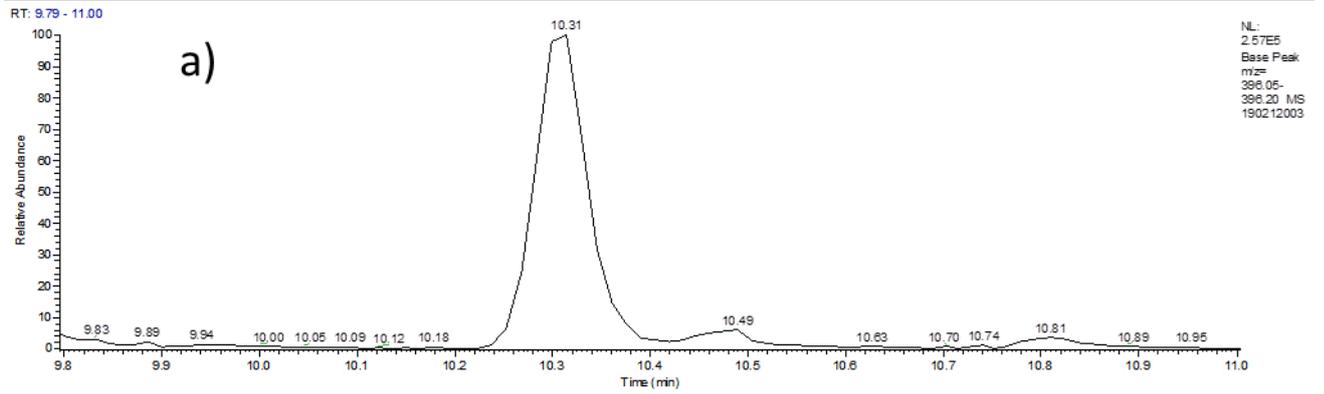
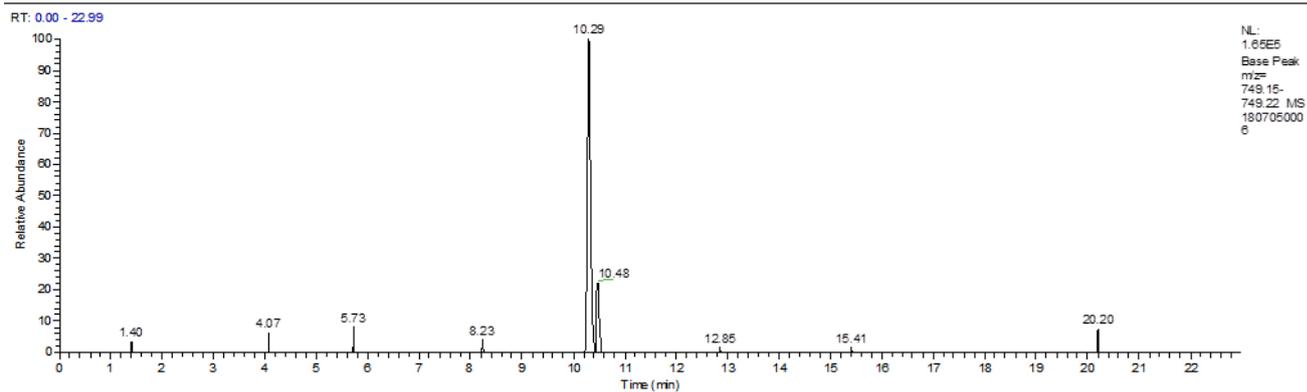


Figure S 73: HPLC Extracted ion chromatogram (Extracted mass 396.05-696.20) of *S. lividans* Δ YA6 1E5 b) Esi full ms chromatogram with the exact mass of SEK43.



1807050006 #904-919 RT: 10.20-10.37 AV: 16 NL: 6.15E4
T: FTMS + c ESI Full ms [200.00-2000.00]

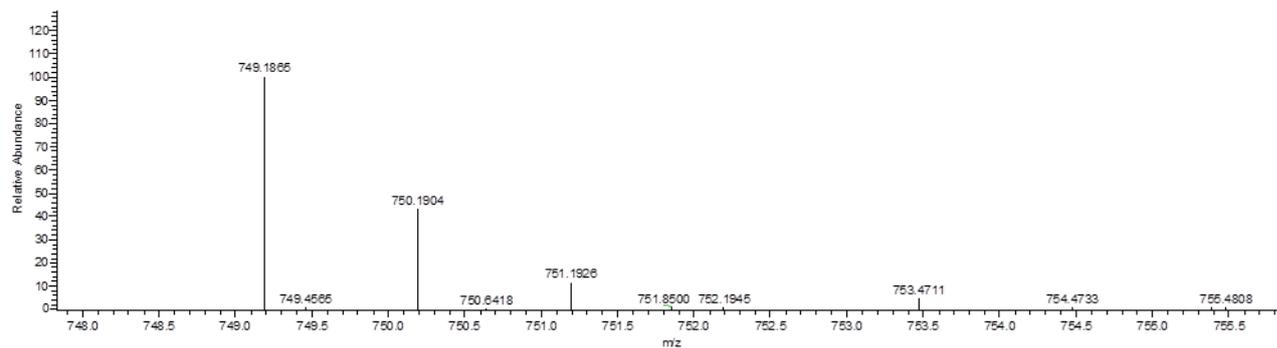


Figure S 74: HPLC Extracted ion chromatogram (Extracted mass 749.15-749.22) of *S. lividans* Δ YA6 1E5 b) Esi full ms chromatogram with the exact mass of SEK87.

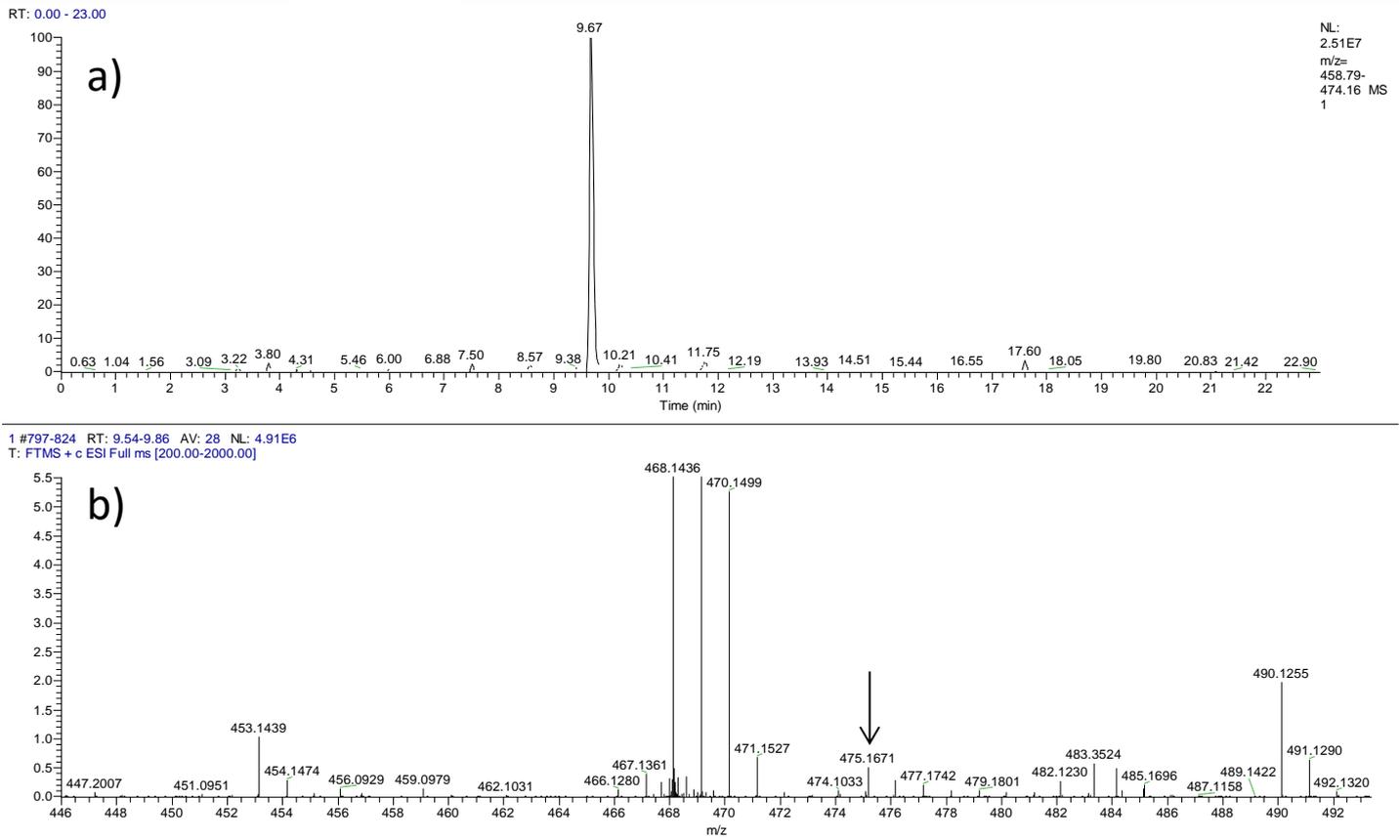


Figure S 75: HPLC-MS Extracted ion chromatogram (Extracted mass 468 ± 0.5) of a) *S. lividans* Δ YA6_1E5 with ^{13}C - 9 - ^{15}N -1-labelled L-tyrosine added to the medium b) MS-Chromatogram of a), incorporation of ^{13}C - 9 - ^{15}N -1-labelled L-tyrosine visible; First repeat of a triplicate.

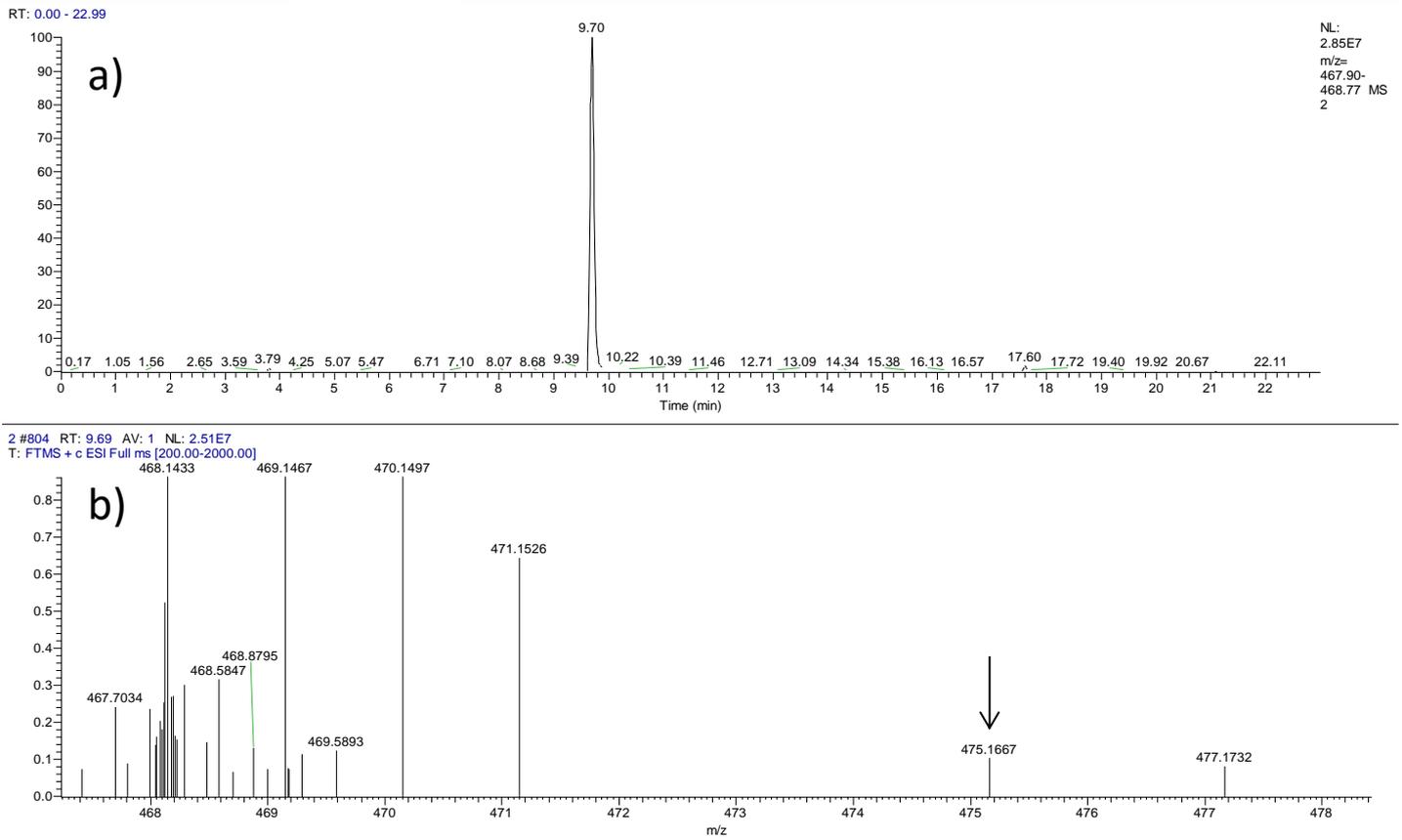


Figure S 76: HPLC-MS Extracted ion chromatogram (Extracted mass 468 ± 0.5) of a) *S. lividans* Δ YA6_1E5 with ^{13}C - 9 - ^{15}N -1-labelled L-tyrosine added to the medium b) MS-Chromatogram of a), incorporation of ^{13}C - 9 - ^{15}N -1-labelled L-tyrosine visible; second repeat of a triplicate.

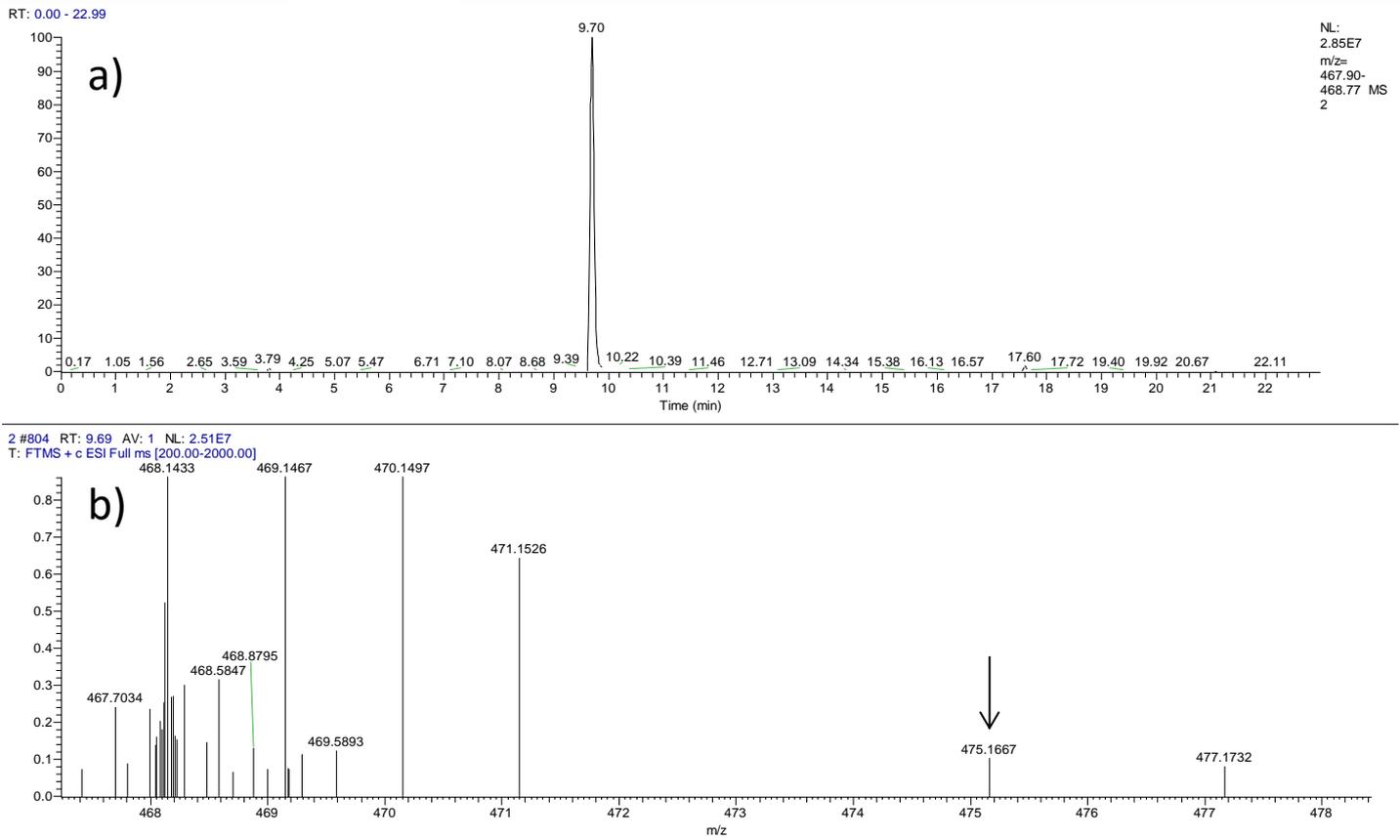


Figure S 77: HPLC-MS Extracted ion chromatogram (Extracted mass 468 ± 0.5) of a) *S. lividans* Δ YA6_1E5 with ^{13}C - 9 - ^{15}N -1-labelled L-tyrosine added to the medium b) MS-Chromatogram of a), incorporation of ^{13}C - 9 - ^{15}N -1-labelled L-tyrosine visible; Third repeat of a triplicate.

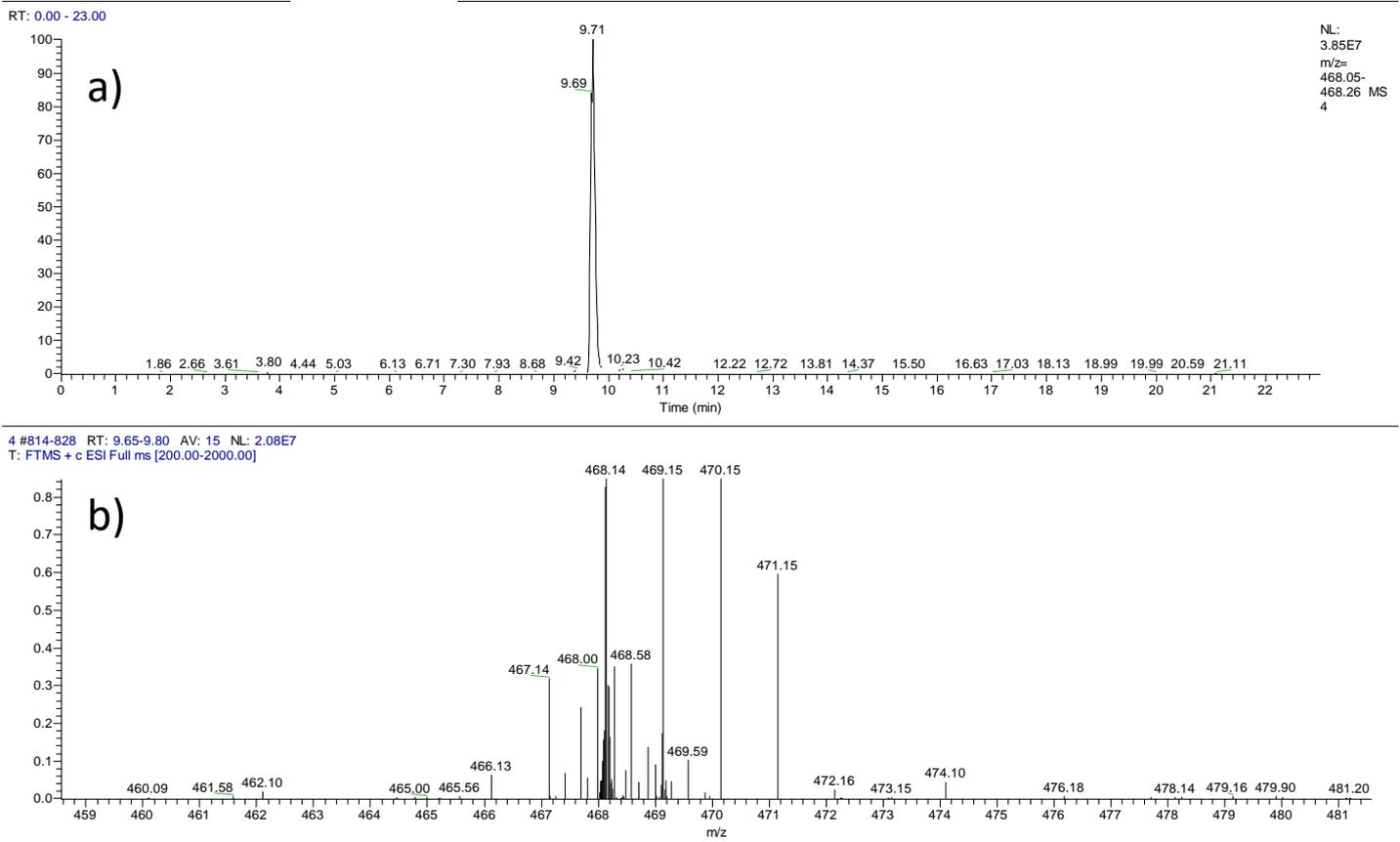


Figure S 78: HPLC-MS Extracted ion chromatogram (Extracted mass 468 ± 0.5) of a) *S. lividans* Δ YA6_1E5 without ^{13}C - ^{9-15}N -1-labelled L-tyrosine added to the medium b) MS-Chromatogram of a) no incorporation of ^{13}C - ^{9-15}N -1-labelled L-tyrosine visible; first repeat of a triplicate.

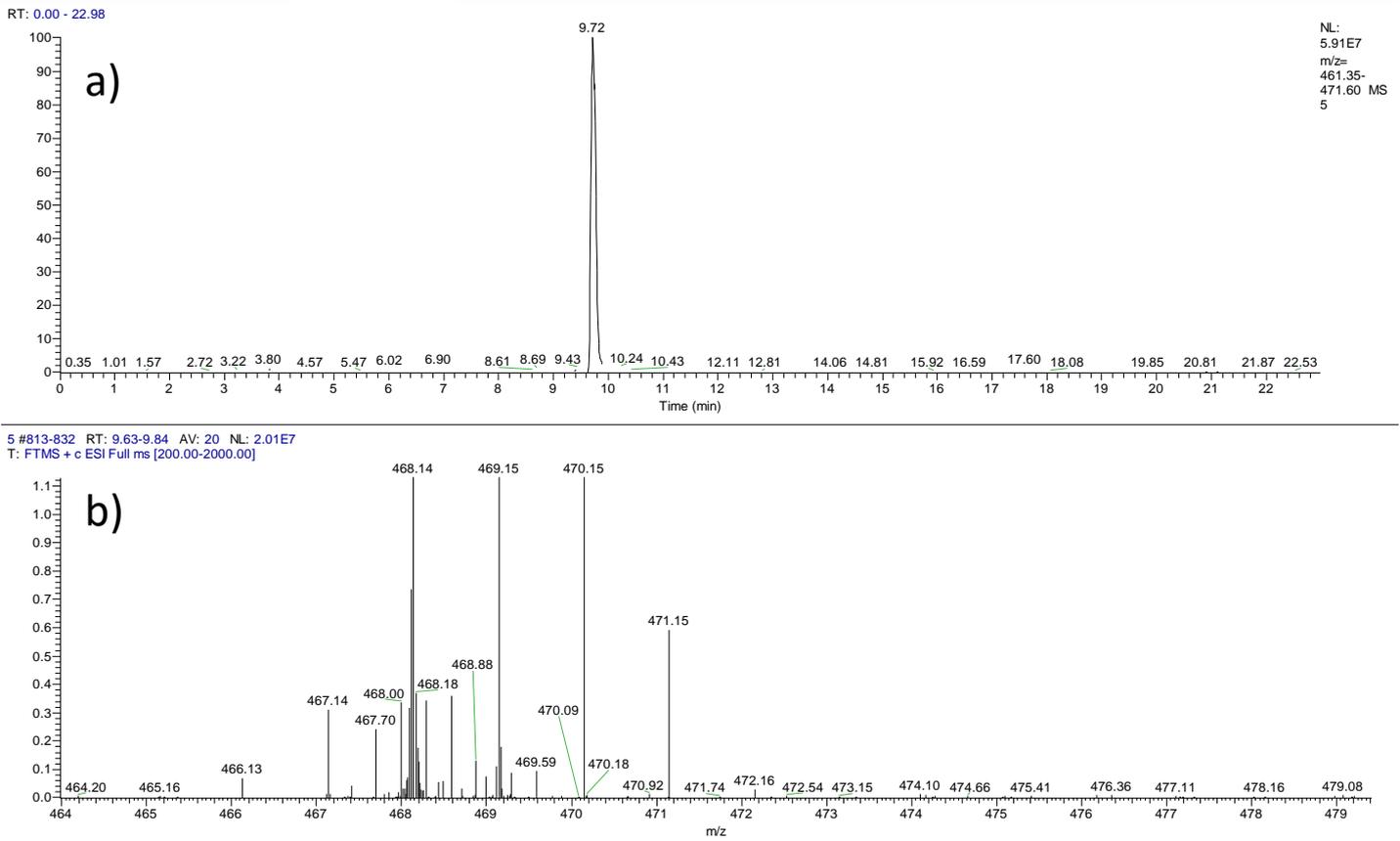
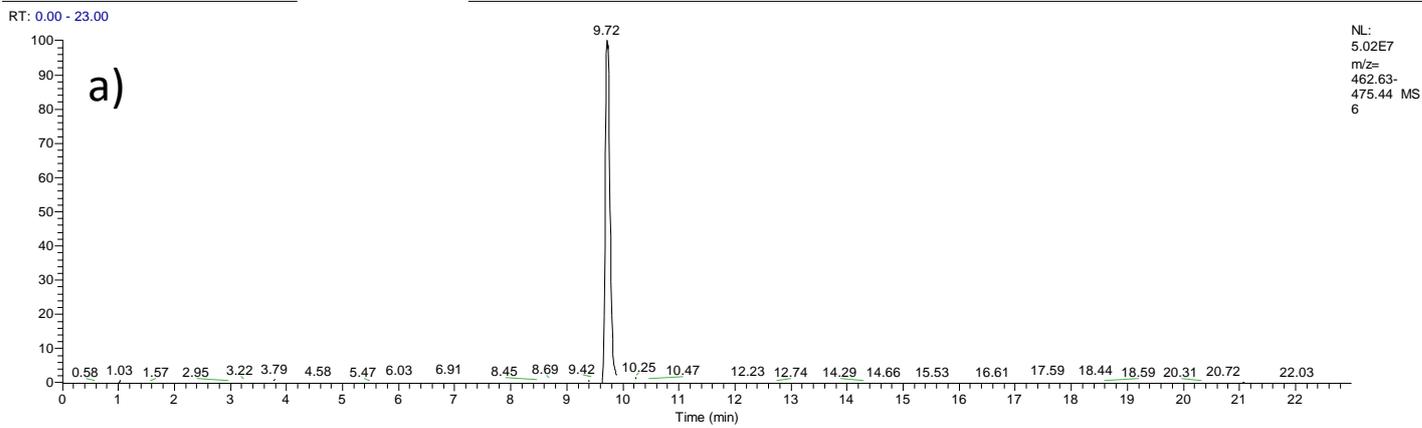


Figure S 79: HPLC-MS Extracted ion chromatogram (Extracted mass 468 ± 0.5) of a) *S. lividans* Δ YA6_1E5 without ^{13}C - ^{9-15}N -1-labelled L-tyrosine added to the medium b) MS-Chromatogram of a) no incorporation of ^{13}C - ^{9-15}N -1-labelled L-tyrosine visible; second repeat of a triplicate.



6 #811-831 RT: 9.64-9.86 AV: 21 NL: 1.60E7
T: FTMS + c ESI Full ms [200.00-2000.00]

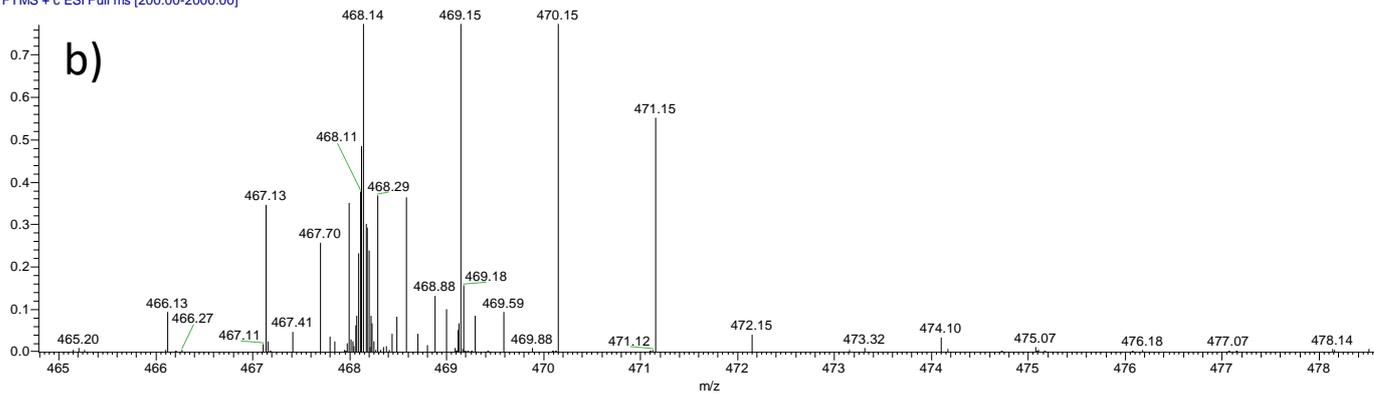
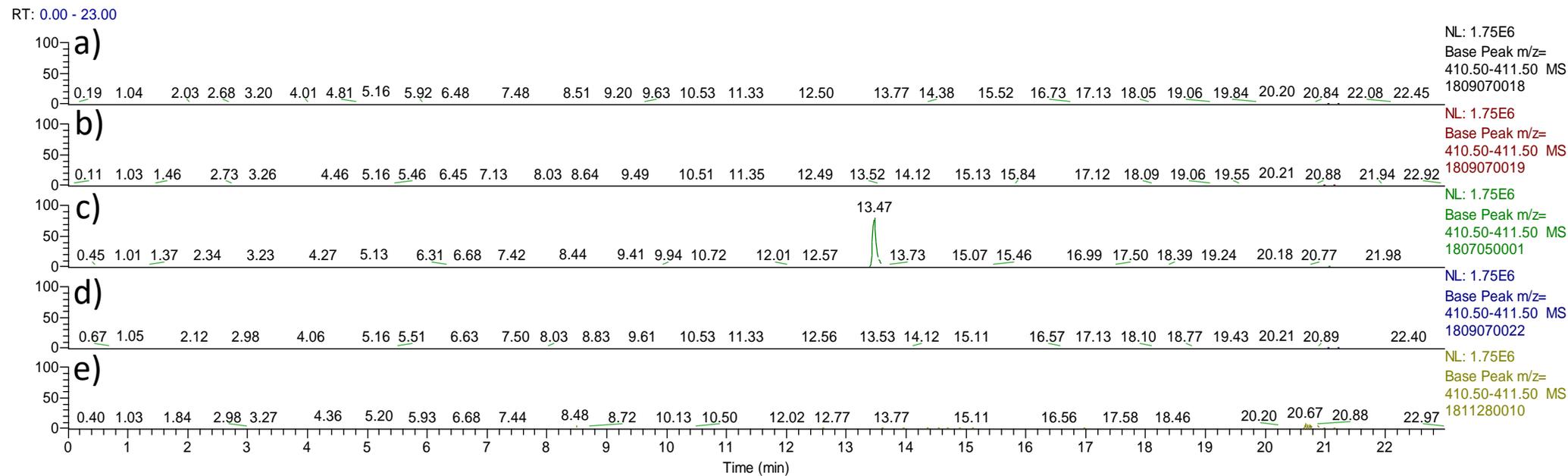


Figure S 80: HPLC-MS Extracted ion chromatogram (Extracted mass 468 ± 0.5) of a) *S. lividans* Δ YA6_1E5 without ^{13}C - ^{15}N -1-labelled L-tyrosine added to the medium b) MS-Chromatogram of a) no incorporation of ^{13}C - ^{15}N -1-labelled L-tyrosine visible; third repeat of a triplicate.



1807050001 #1135 RT: 13.46 AV: 1 NL: 1.37E6
 T: FTMS + c ESI Full ms [200.00-2000.00]

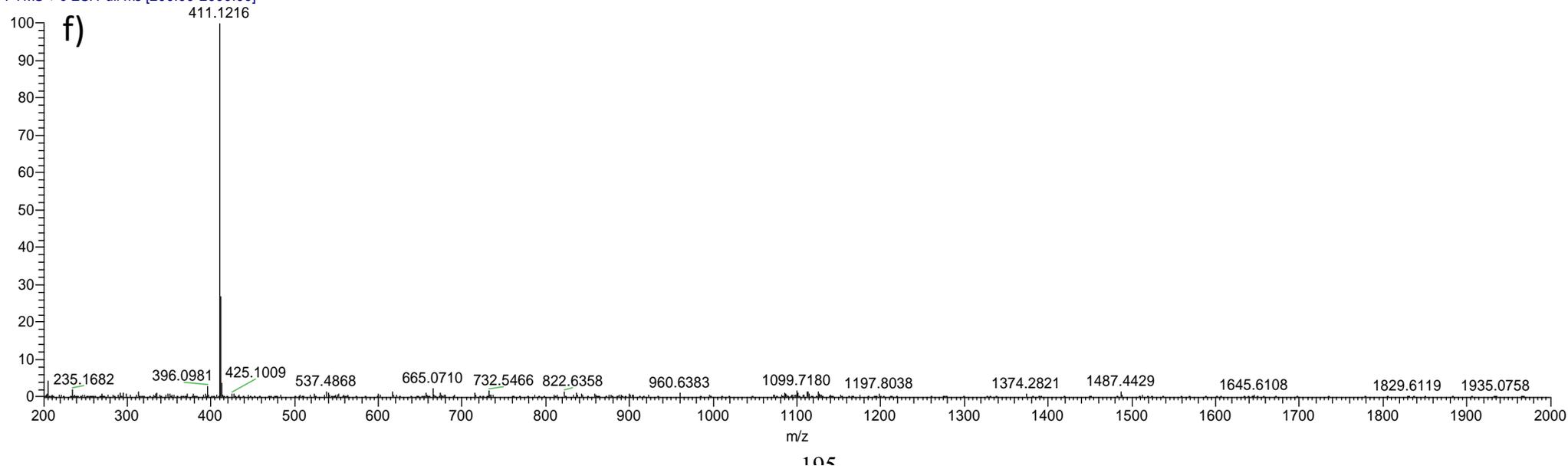
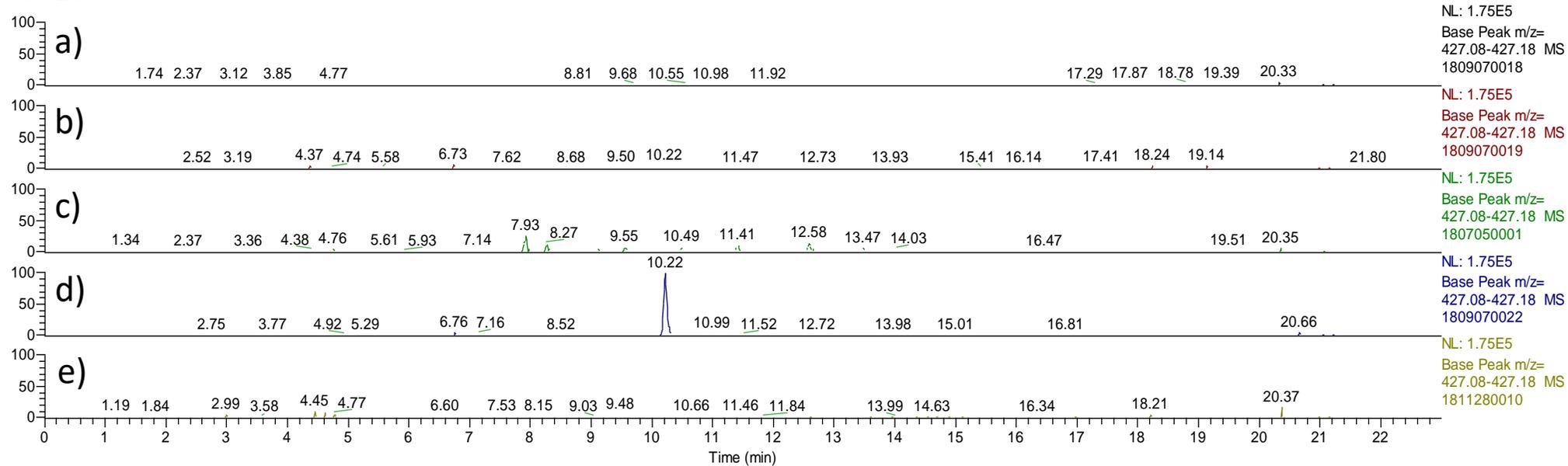


Figure S 81: HPLC-MS Extracted ion chromatogram (Extracted mass 411 ± 0.5 ; corresponding mass to (3)) of a) *S. lividans* Δ YA6; b) *S. lividans* Δ YA6_1E5; c) *S. lividans* Δ YA6 Δ penA; d) *S. lividans* Δ YA6 Δ penC; e) *S. lividans* Δ YA6 Δ penD; peak visible in c); f) ESI Full MS chromatogram of the peak visible in c) $R_t=13.47$.

RT: 0.00 - 23.00



1809070022 #822 RT: 10.22 AV: 1 NL: 1.72E5
 T: FTMS + c ESI Full ms [200.00-2000.00]

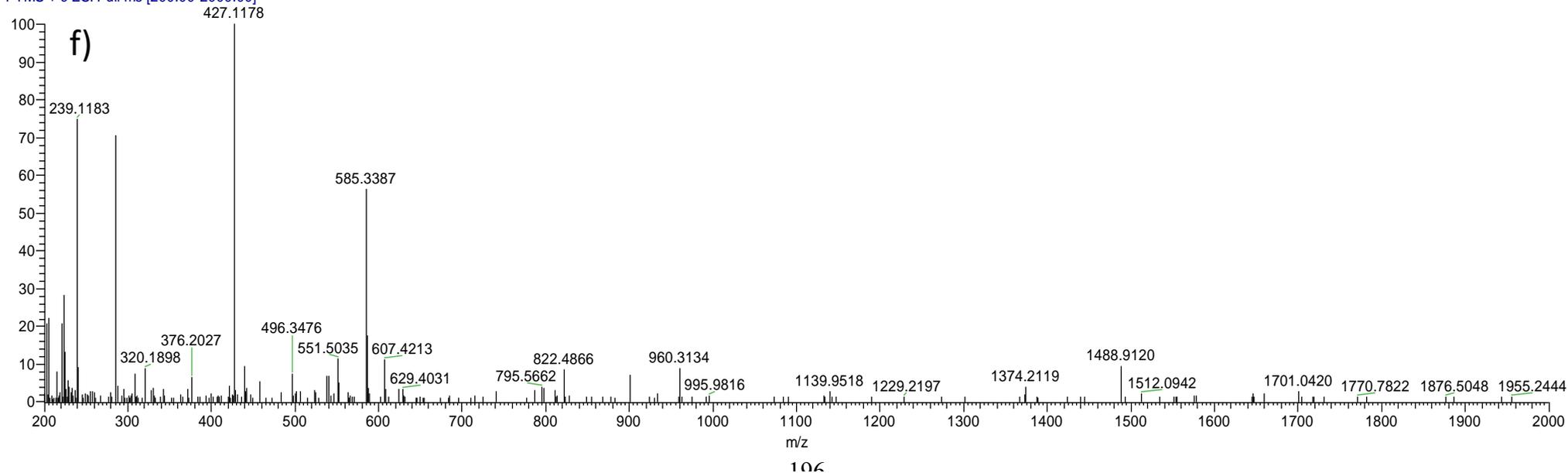
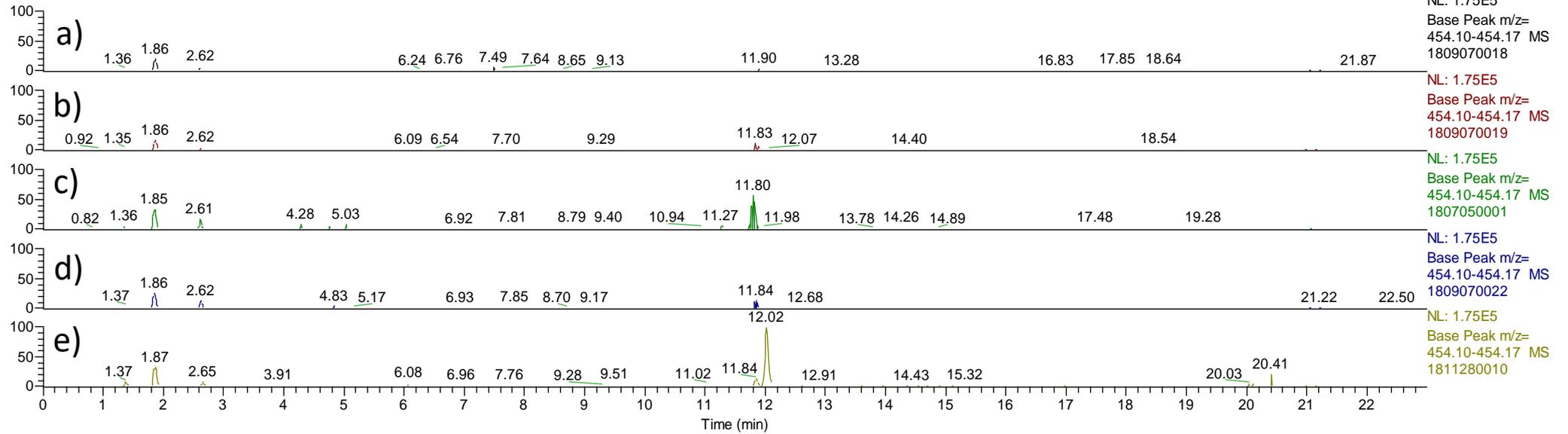


Figure S 82: HPLC-MS Extracted ion chromatogram (Extracted mass 427.09-427.18; corresponding mass to (4)) of a) *S. lividans* Δ YA6; b) *S. lividans* Δ YA6_1E5; c) *S. lividans* Δ YA6 Δ penA; d) *S. lividans* Δ YA6 Δ penC; e) *S. lividans* Δ YA6 Δ penD; peak visible in d); f) ESI full MS of the peak visible in d) ($R_t=10.22$).

RT: 0.00 - 23.00



1811280010 #938 RT: 12.02 AV: 1 NL: 2.04E5
T: FTMS + c ESI Full ms [200.00-2000.00]

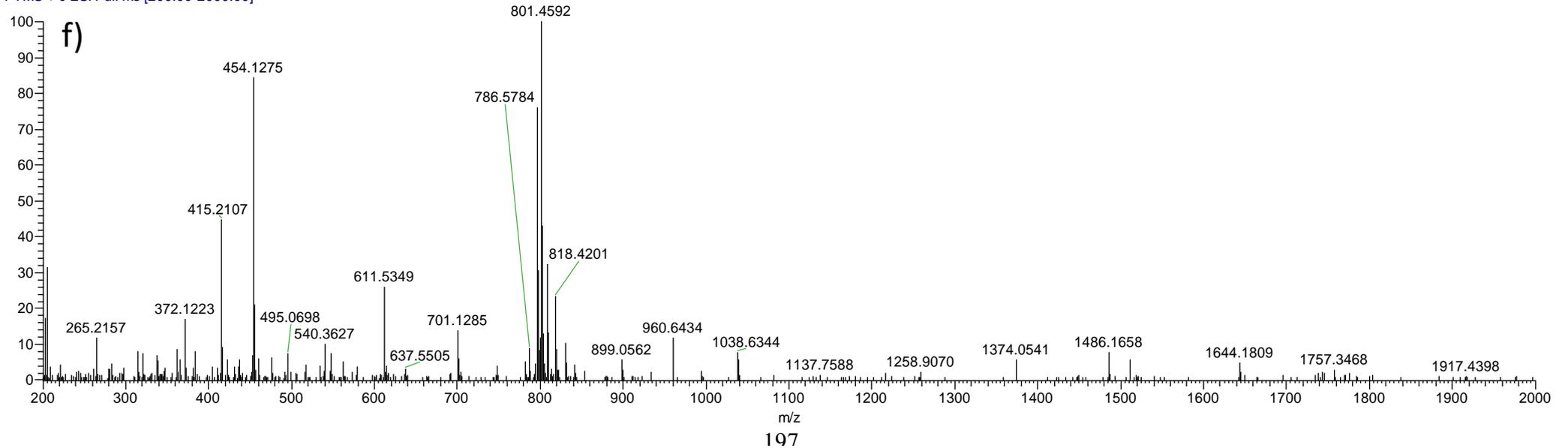


Figure S 83: HPLC-MS Extracted ion chromatogram (Extracted mass 454.10-454.17; corresponding mass to (5)) of a) *S. lividans* Δ YA6; b) *S. lividans* Δ YA6_1E5; c) *S. lividans* Δ YA6 Δ penA; d) *S. lividans* Δ YA6 Δ penC; e) *S. lividans* Δ YA6 Δ penD; peak visible in e); f) ESI full MS of the peak visible in e) ($R_t=12.02$).

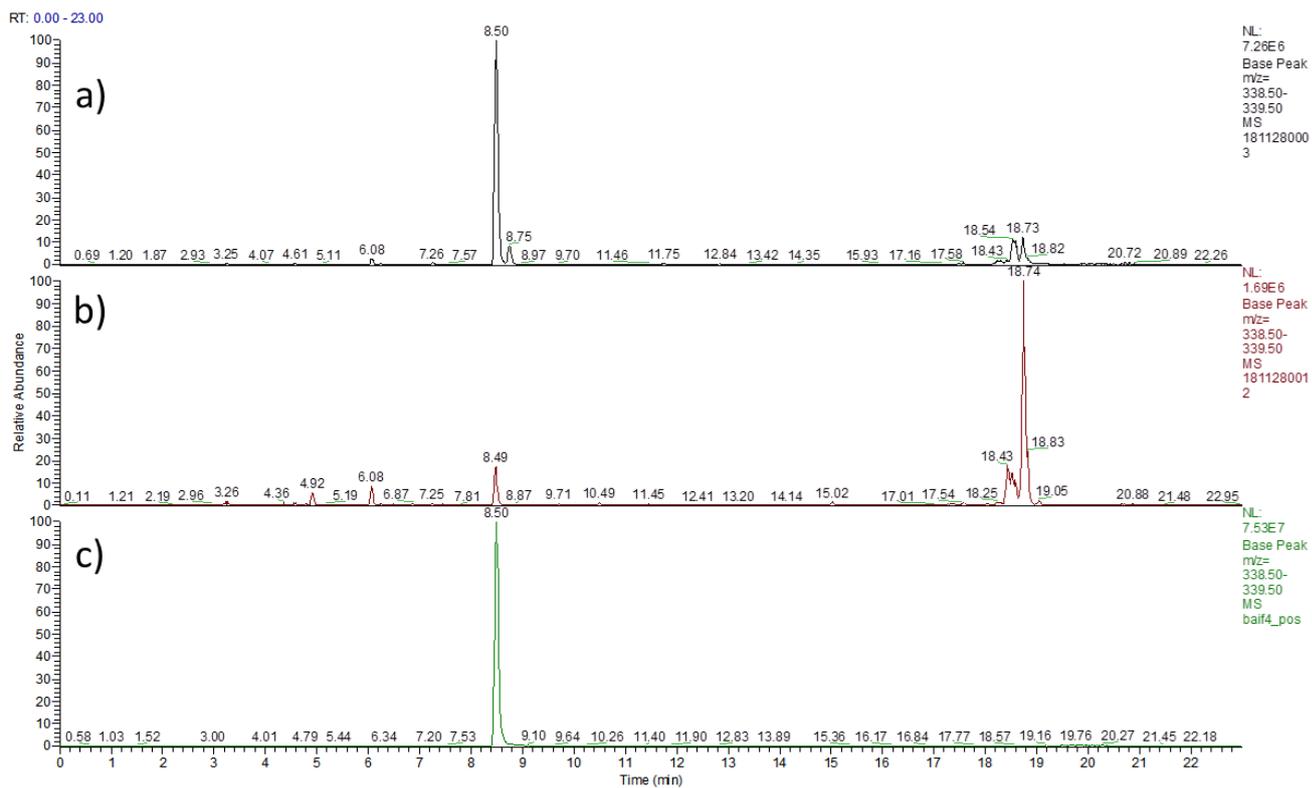


Figure S 84: HPLC-MS Extracted ion chromatogram (Extracted mass 339 ± 0.5 , Rabelomycin) of a) *S. lividans* Δ YA6 1E5, b) *S. lividans* Δ YA6 1E5 Δ penE c) pure rabelomycin as external standard.

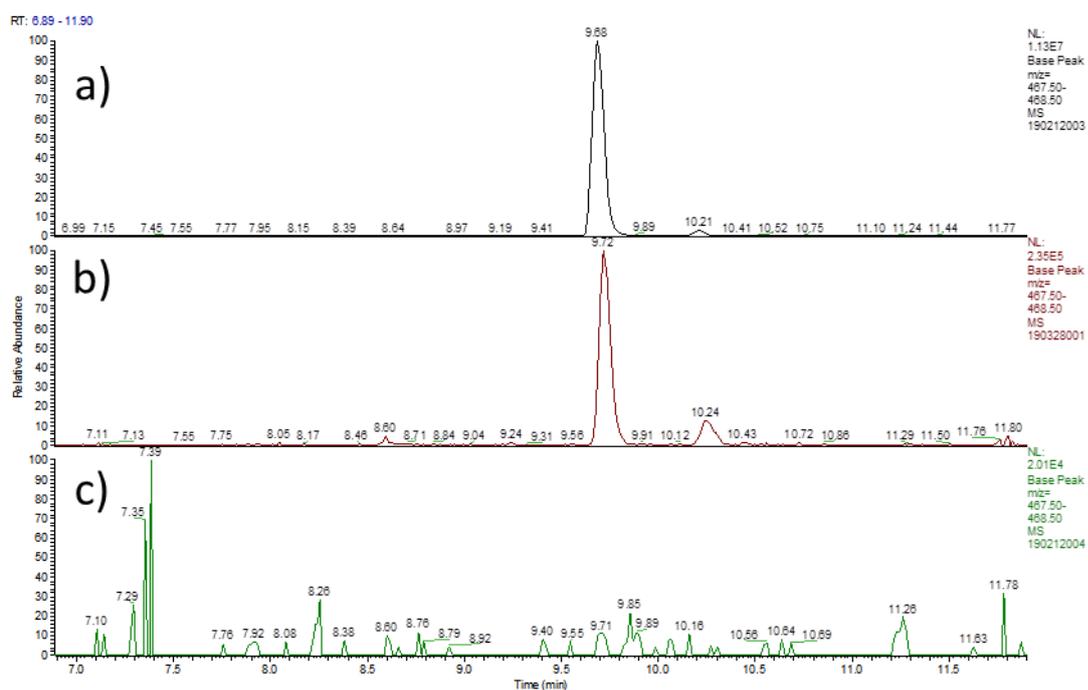


Figure S 85: HPLC-MS Extracted ion chromatogram (Extracted mass 468 ± 0.5 , **(1)**) of a) *S. lividans* Δ YA6 1E5, b) *S. lividans* Δ YA6 1E5 Δ penR1 and c) *S. lividans* Δ YA6 1E5 Δ penR2.

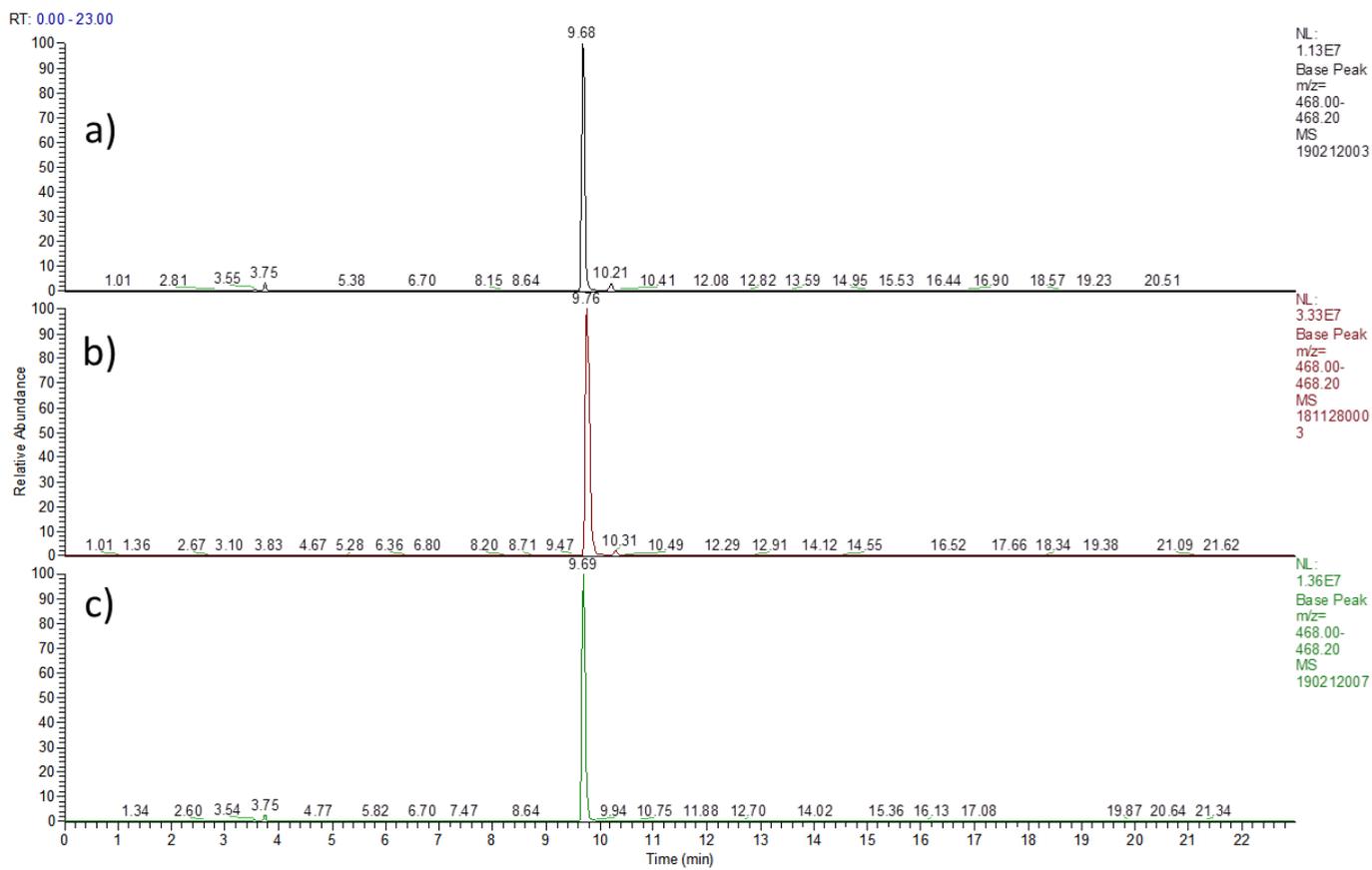


Figure S 86: HPLC-MS Extracted ion chromatogram (Extracted mass 468 ± 0.5 (**1**)) of a) *S. lividans* Δ YA6 1E5, b) *S. lividans* Δ YA6 A3_penR1 1E5 and c) *S. lividans* Δ YA6 A3_penR2 1E5.

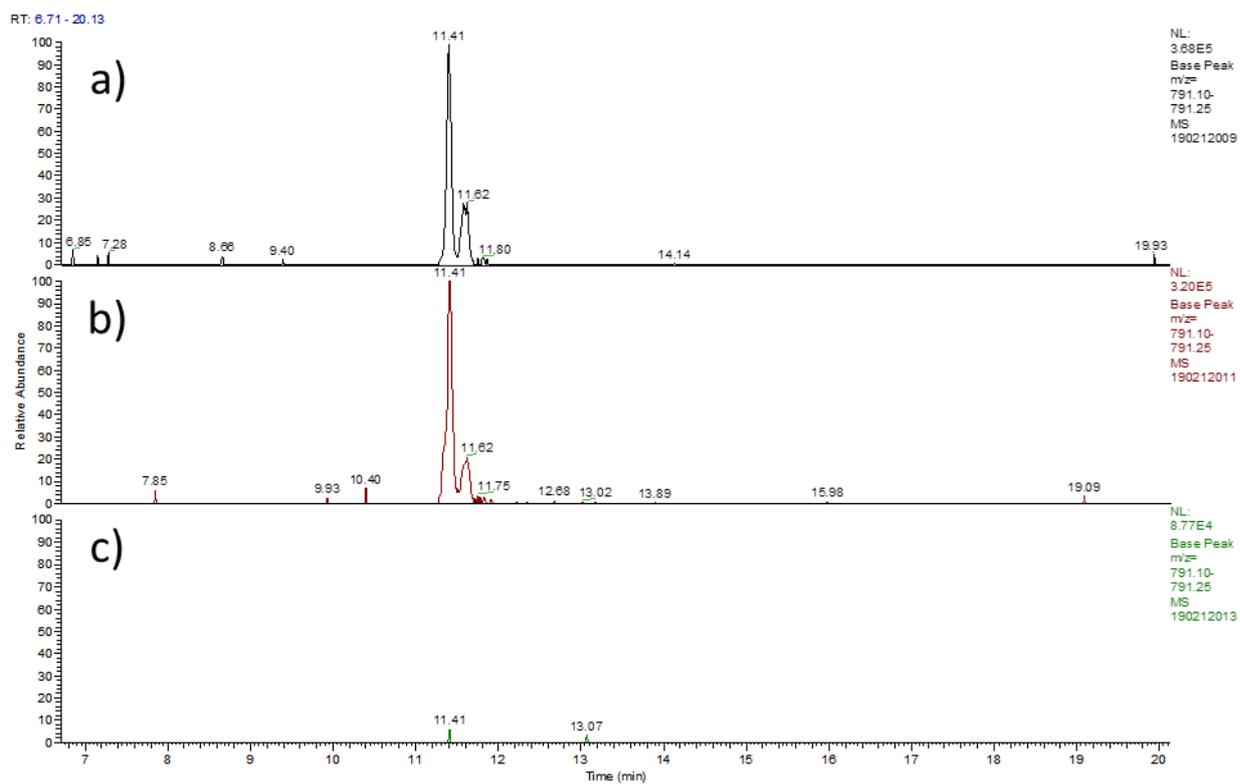


Figure S 87: HPLC-MS Extracted ion chromatogram (Extracted mass 791.10-791.25 (**2**)) of a) *S. lividans* Δ YA6 1E5, b) *S. lividans* Δ YA6 A3_penR1 1E5 and c) *S. lividans* Δ YA6 A3_penR2 1E5.

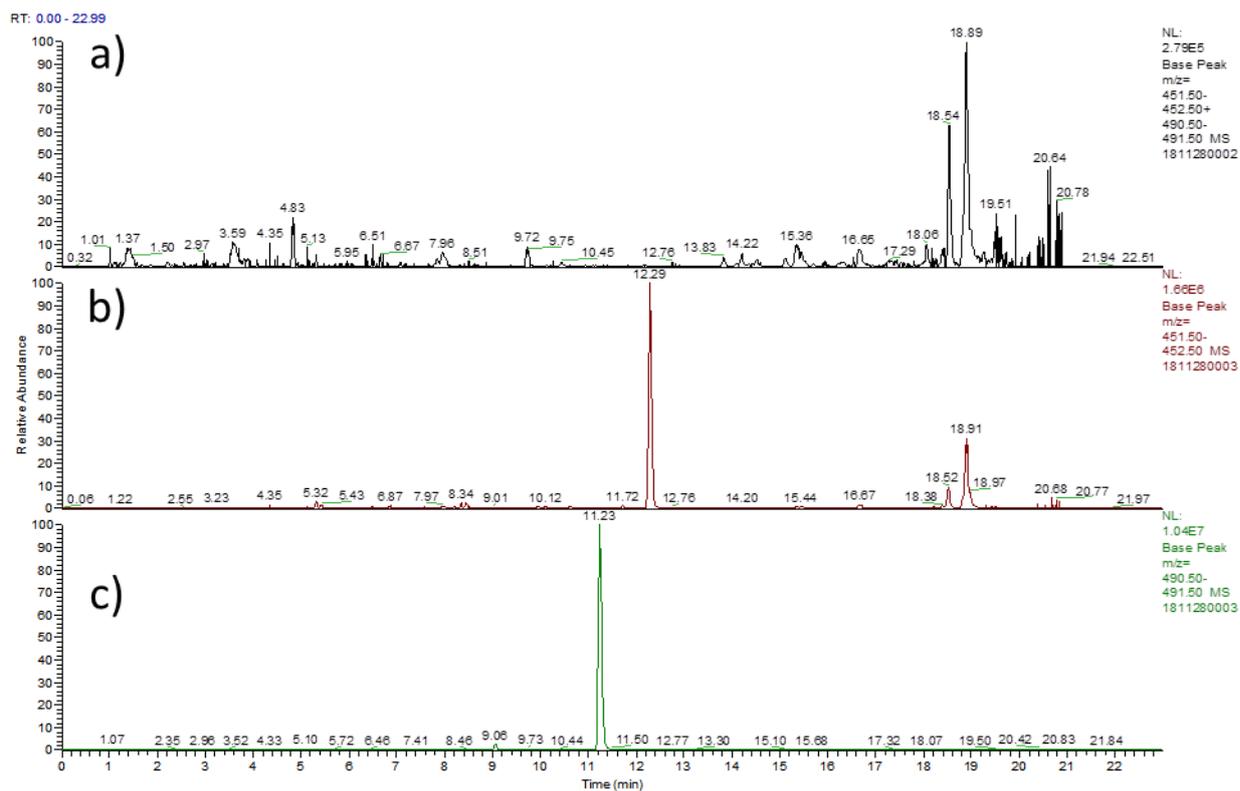


Figure S 88: HPLC-MS Extracted ion chromatogram a) *S. lividans* Δ YA6 A3_penR1 (Extracted mass 491 ± 0.5 (6) and 452 ± 0.5 (7)), b) *S. lividans* Δ YA6 A3_penR1 1E5 (Extracted mass 452 ± 0.5 (7)) c) *S. lividans* Δ YA6 A3_penR1 1E5 (Extracted mass 491 ± 0.5 (6)).

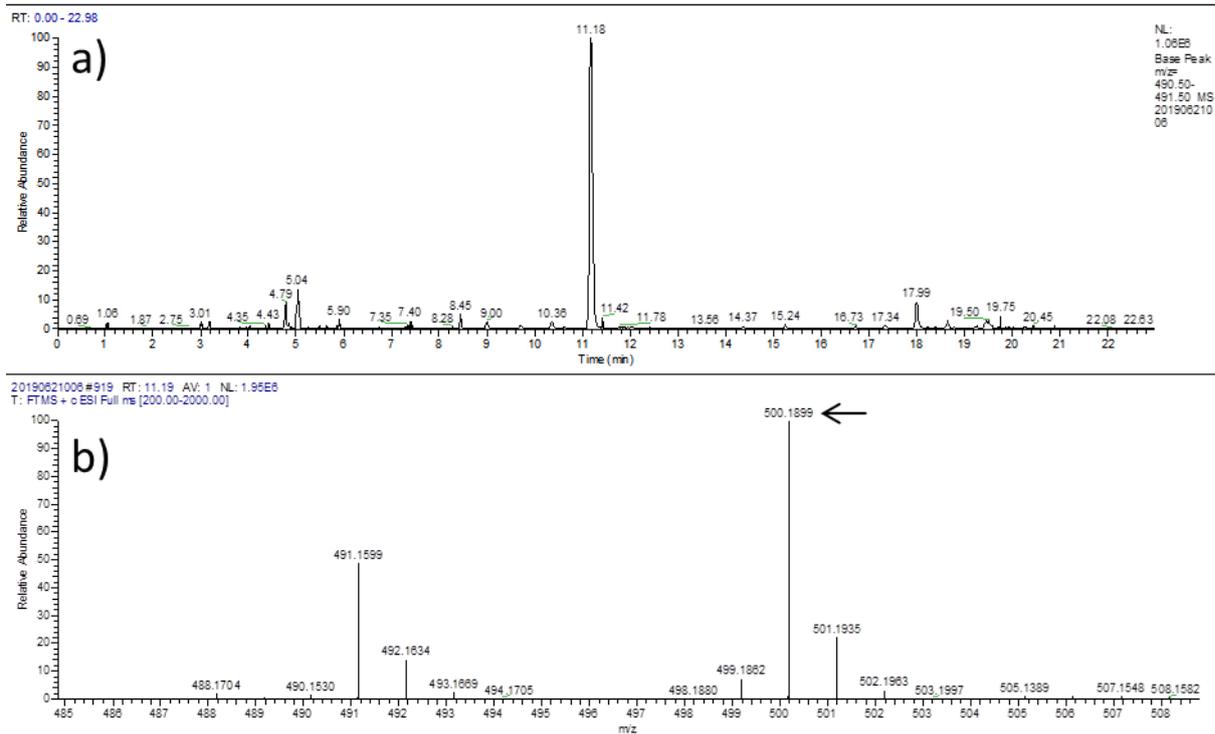


Figure S 89: HPLC-MS Extracted ion chromatogram (Extracted mass 491 ± 0.5) of a) *S. lividans* Δ YA6_A3_penR1 1E5 with L-tryptophan 13C-11 added to the medium b) MS-Chromatogram of a), incorporation of L-tryptophan 13C-11 visible.

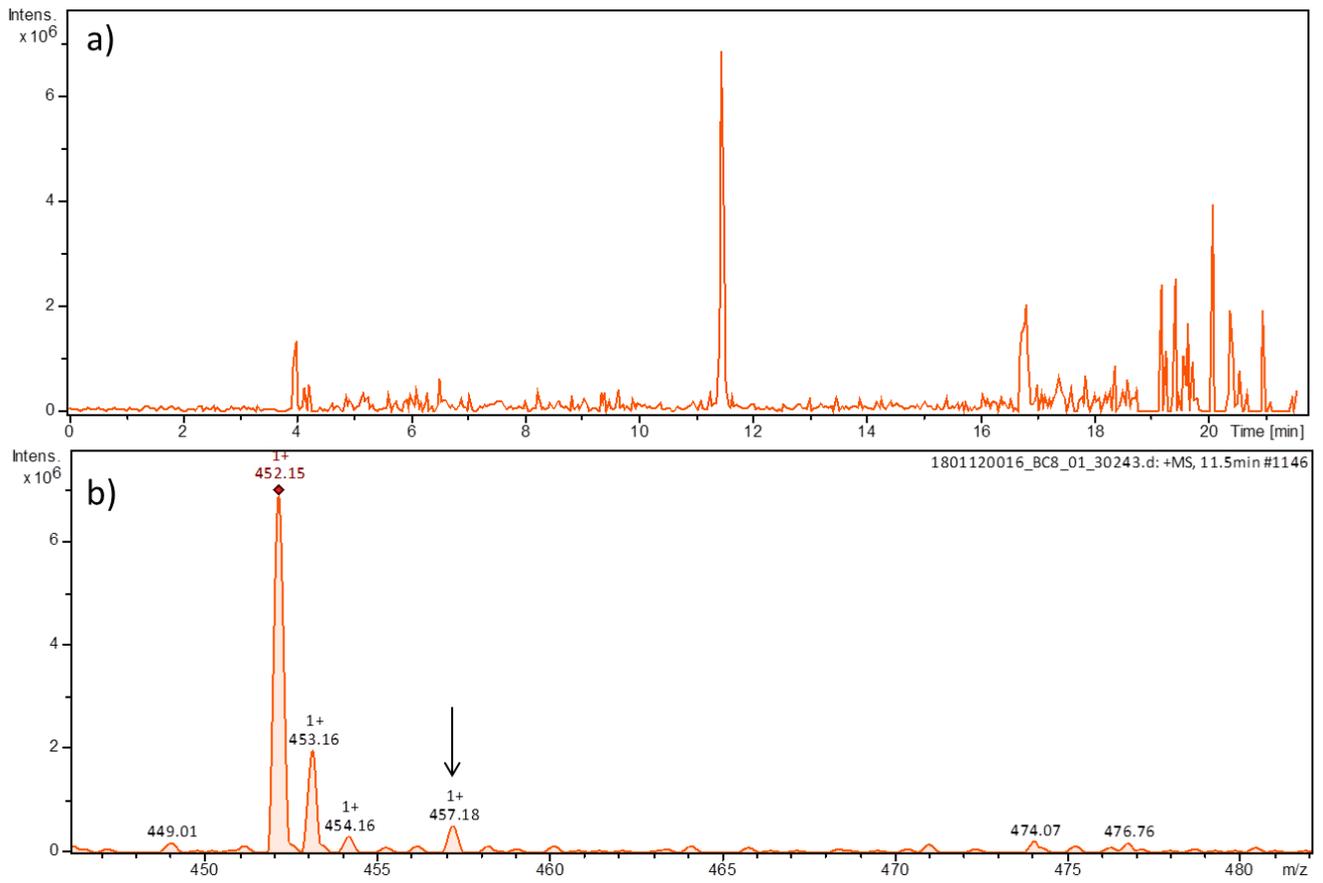


Figure S 90: HPLC-MS Extracted ion chromatogram (Extracted mass 452 ± 0.5) of a) *S. lividans* Δ YA6_A3_penR1 1E5 with 2D-5-labelled L-Phenylalanine added to the medium b) MS-Chromatogram of a), incorporation of 2D-5-labelled L-Phenylalanine visible.

6.2. Section 3

Table S 7: Oligonucleotides used in Section 3, supplied by Eurofins GENOMICS, Ebersberg, Germany.

Name	Sequence of Oligonucleotides
f-vdhhindERME	ATAAGCTTATCTACCGGGAGTCACCAC
r-vdhbamERME	ATGGATCCATTGGCCGAAATGATCCTCC
f-ccr2hindERME	ATAAGCTTTCAGCCCTCACTACGCCCTT
r-ccr2bamERME	ATGGATCCCTTCTGGCGCTCTGTCATGG
f-meaahindERME	AAAAAAAAAGCTTCATCAACCGCTTCCGGAACATCTG
r-meaabamERME	AAAAAAAAAGGATCCTTCTCCAGGAAGCGCGGGTTG
f-ccr2Del	CACTACGCCCTTCGCCGGAGGCACCACCGTGAAGGAAATATTCCGGGGATCCGTCGACC
r-ccr2Del	TGGCGCTCTGTCATGGAGGGAACCTCAGATGTTCCGGAATGTGTAGGCTGGAGCTGCTT C
f-mcmDel	CCGTCAGCGGCTCACAGCTCGTGGCCGAGCGACGCGGCAAGCCGCTCGACTTCCGGGG ATCCGTCGACCC
r-mcmDel	GGGCAAGGTCACAGGCGCGGGTACTGCTGCACGCCTGCCGCGCTGCCTATGTAGGCTG GAGCTGCTTCG
f-pcc1Del	TCATGCCTGGGCTCCTCGGCGGGGTGGCGGGAGCCTGCCGTACGACGGTTCGGGG ATCCGTCGACCC
r-pcc1Del	TCATTAACGCCGAGCGGTGCCCCACCGAGGCCGCGCCGGTCACGGCATGTAGGCT GGAGCTGCTTCG
f-pcc2Del	GGCTGCGCGCCGCCACCTCGACATCCCCGAAGCCGTCCTCAGCCAGGAGTTCGGGGGA TCCGTCGACCC
r-pcc2Del	TCACGCGTTTTCTCCTGGACGACGGCGAGCAGGGCACCGACCTCGACCTTGTAGGCTG GAGCTGCTTCG
f-pcc3Del	TCAGAGCGGGATGTTGCCGTGCTTCTTCGGCGGCAGGCTCTCCCGTTGTTCCGGGGAT CCGTCGACCC
r-pcc3Del	ATGTCCGCGCCGAAAACGCCACCACCCGCCAACCCCGAGCTGCACATGTAGGCTGG AGCTGCTTCG
f-vdhDel	CACCTTACGGACTCCGCCATCCGCCGATCTCACCGGGAGTCACCACCTTCGGGGAT CCGTCGACCC
r-vdhDel	CGGCGACGGGGGAAAATTGGCCGAAATGATCCTCCGTCCGGGGTGCGGGTGTAGGCT GGAGCTGCTTCG
f-meaaDel	ATGACAGAGCGCCAGAAGGACCGGCCGTGGCTCATGCGGACGTACTTCCGGGGATCCG TCGACCC
r-meaaDel	TCATGCGGGGACCTCCAGAGGGTCGAGCTTGTTCACGAGGCGGATTGTAGGCTGGAGC TGCTTCG
f-chk-ccr2	TCAGCCCTCACTACGCCCTT
r-chk-ccr2	TTCTGGCGCTCTGTCATGGA
f-chk-mcm	GATCGATCGTCGCCCCAT
r-chk-mcm	ATCAGCAGCGTTGCGGAGA
f-chk-pcc1	GTGGCGGGGGTGGGTGATT
r-chk-pcc1	GGTCGTTAGCGGTCGTTAA

f-chk-pcc2	ATCAGGGTGGTGGCGGTGG
r-chk-pcc2	CTGCTCGGTGACGGTGTCTG
f-chk-pcc3	GTACGGATACTCAAGCGCCT
r-chk-pcc3	CAAACACCGCACTGACGAAG
f-chk-vdh	CTCATGCCACGGCGTCCGC
r-chk-vdh	AATCCCCCGCCGCACACCT
f-chk-meaa	AACCACGAGATGCGTGCCCA
r-chk-meaa	TCCTCCAGGTGAGGAACAC

Table S 8: Bacterial strains and vectors used and developed in Section 3.

<i>Streptomyces</i>	Characteristics	Reference
<i>S. albus</i> J1074	<i>S. albus</i> G1 (DSM 41398) derivative with the defective SalGI restriction modification system heterologous host	(Chater and Wilde, 1980) ¹²³
<i>S. albus</i> Del1	<i>S. albus</i> J1074 with the deletion of 204 kB including <i>ccr1</i>	(Myronovskyi <i>et al.</i> , 2014b) ²³⁸
<i>S. albus</i> Δ2	<i>S. albus</i> Del1 with the deletion of <i>ccr2</i>	This work
<i>S. albus</i> Δ3	<i>S. albus</i> Del1 with the deletion of <i>ccr2</i> and <i>mcm</i>	This work
<i>S. albus</i> Δ4	<i>S. albus</i> Del1 with the deletion of <i>ccr2</i> , <i>mcm</i> and <i>pcc1</i>	This work
<i>S. albus</i> Δ5	<i>S. albus</i> Del1 with the deletion of <i>ccr2</i> , <i>mcm</i> , <i>pcc1</i> and <i>pcc2</i>	This work
<i>S. albus</i> Δ6	<i>S. albus</i> Del1 with the deletion of <i>ccr2</i> , <i>mcm</i> , <i>pcc1</i> , <i>pcc2</i> and <i>vdh</i>	This work
<i>S. albus</i> Δ7	<i>S. albus</i> Del1 with the deletion of <i>ccr2</i> , <i>mcm</i> , <i>pcc1</i> , <i>pcc2</i> , <i>mcm</i> , <i>vdh</i> and <i>pcc3</i>	This work
<i>S. albus</i> Del1 - Δ <i>pcc2</i>	<i>S. albus</i> Del1 with the deletion of <i>ccr2</i> and <i>pcc2</i>	This work
<i>S. albus</i> Del1 – Δ <i>pcc1</i>	<i>S. albus</i> Del1 with the deletion of <i>ccr2</i> and <i>pcc1</i>	This work
<i>S. albus</i> Del1 - Δ <i>meaA</i>	<i>S. albus</i> Del1 with the deletion of <i>meaA</i>	This work
<i>S. albus</i> Del1 - Δ <i>ccr2</i> Δ <i>meaA</i>	<i>S. albus</i> Del1 with the deletion of <i>meaA</i> and <i>ccr2</i>	This work
<i>S. albus</i> Δ <i>vdh</i>	<i>S. albus</i> J1074 with the deletion of <i>vdh</i>	This work
<i>E. coli</i>	Characteristics	Reference
GB2005	General cloning host	(Maresca <i>et al.</i> , 2010) ²³⁶
WM6026	Strain used for intergeneric conjugation, 2,6 Diaminopimelic acid auxotroph	(Blodgett <i>et al.</i> , 2007) ²³⁹
GB2005-red-rham	GB2005, RhamC-BAD- <i>ybaA</i> used for Red/ET	(Strochlic <i>et al.</i> , 2010) ²³⁷
Plasmids	Characteristics	Reference
patt-saac-OriT	Swal fragment of <i>paac-OK</i> containing <i>aac(3)IV</i> and <i>oriT</i> cloned into the <i>EcoRV</i> site of <i>patt</i>	(Myronovskyi <i>et al.</i> , 2014a) ¹⁹²
pUWL_oriT	Replicative vector for actinomycetes; pIJ101 replicon, <i>oriT</i> , <i>tsr</i> , <i>bla</i> , <i>ermE</i>	(Bierman <i>et al.</i> , 1992) ²⁴⁰
pUWL-H	pUWLoriT with Hygr (<i>hph</i>) instead of <i>Tsrr(tsrr)</i> ; Vector used to express genes under the influence of the <i>ermE</i> Promotor	(Petzke <i>et al.</i> , 2010) ²⁴¹

pUWL-H_CCR2	pUWL-H containing the gene <i>ccr2</i> under the influence of the <i>ermE</i> Promotor	This work
pUWL-H_VDH	pUWL-H containing the gene <i>vdh</i> under the influence of the <i>ermE</i> Promotor	This work
pUWL-oriT_meaA	pUWL-H containing the gene <i>meaA</i> under the influence of the <i>ermE</i> Promotor	This work
patts-HYG-OriT	Swal/Smal fragment of <i>phyg</i> -OK containing <i>hyg</i> and <i>oriT</i> cloned into the <i>EcoRV</i> site of <i>patt</i>	(Myronovskyi <i>et al.</i> , 2014a) ¹⁹²
pJET1.2	Cloning Vector; AmpR	Thermo Fisher Scientific
Cosmids	Characteristics	Reference
R2	Apr Cosmid containing the biosynthetic gene cluster of pamamycin based on the <i>poj463</i> plasmid, <i>int(vwb)</i> and <i>aac(3)IV</i>	(Rebets <i>et al.</i> , 2015) ⁷⁶
pSMARTgus	Derivative of pSMART containing the <i>gusA</i> gene	(Myronovskyi <i>et al.</i> , 2014b) ²³⁸
pSMARTgus_3D17	pSMARTgus based cosmid containing a part of the chromosomal DNA from <i>S. albus</i> J1074, including the genes <i>meaA</i> and <i>ccr2</i> . Size: 84kB	Personal communication with Prof. Dr. A. Luzhetskyy
3D17 Δ <i>ccr2</i> Δ <i>meaA</i>	Derivative of pSMARTgus_3D17 with the deletion of <i>meaA</i> and <i>ccr2</i>	This work
3D17 Δ <i>meaA</i>	Derivative of pSMARTgus_3D17 with the deletion of <i>meaA</i>	This work
3D17 Δ <i>ccr2</i>	Derivative of pSMARTgus_3D17 with the deletion of <i>ccr2</i>	This work
pSMARTgus_2J19	pSMARTgus based cosmid containing a part of the chromosomal DNA from <i>S. albus</i> J1074, including the gene <i>vdh</i> . Size: 54kB	Personal communication with Prof. Dr. A. Luzhetskyy
2J19 Δ <i>vdh</i>	Derivative of pSMARTgus_2J19 with the deletion of <i>vdh</i>	This work
pSMARTgus_3H24	pSMARTgus based cosmid containing a part of the chromosomal DNA from <i>S. albus</i> J1074, including the gene <i>pcc3</i> . Size: 72kB	Personal communication with Prof. Dr. A. Luzhetskyy
3H24 Δ <i>pcc3</i>	Derivative of pSMARTgus_2H24 with the deletion of <i>pcc3</i>	This work
pSMARTgus_1M11	pSMARTgus based cosmid containing a part of the chromosomal DNA from <i>S. albus</i> J1074, including the gene <i>pcc2</i> . Size: 66kB	Personal communication with Prof. Dr. A. Luzhetskyy
1M11 Δ <i>pcc2</i>	Derivative of pSMARTgus_1M11 with the deletion of <i>pcc2</i>	This work

pSMARTgus_2E5	pSMARTgus based cosmid containing a part of the chromosomal DNA from <i>S. albus</i> J1074, including the gene pcc3. Size: 75kB	Personal communication with Prof. Dr. A. Luzhetskyy
2E5Δpcc1	Derivative of pSMARTgus_2E5 with the deletion of pcc1	This work
pSMARTgus_3J4	pSMARTgus based cosmid containing a part of the chromosomal DNA from <i>S. albus</i> J1074, including the gene mcm. Size: 87kB	Personal communication with Prof. Dr. A. Luzhetskyy
3J4Δmcm	Derivative of pSMARTgus_3J4 with the deletion of mcm.	This work

Table S 9: HPLC-MS method to analyze pamamycins.

Time	% Buffer A	% Buffer B	Composition Buffer A	Composition Buffer B
0 min	80%	20%		
0.2 min	80%	20%	90mMol	200 ml 100mM
3 min	3%	97%	Ammonium formate in	Ammonium formate in H2OMiliQ 800 ml
10 min	0%	100%	H2OMiliQ	Acetonitrile
11 min	0%	100%		
12 min	80%	20%	Additional Information	
15 min	80%	20%	Column: 45 °C; Flow:0.55 ml/min,	