

Unravelling the Chemical Diversity of Actinobacteria Obtained from Unique Asian Ecological Niches

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Niraj Aryal
aus Rupandehi, Nepal

Tübingen
2021

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Prüfung:

28.07.2021

Dekan:

Prof. Dr. Thilo Stehle

1. Berichterstatter:

Prof. Dr. Harald Groß

2. Berichterstatter:

PD Dr. Bertolt Gust

Statement on the Originality of This Thesis

I hereby declare that I alone wrote the doctoral work submitted here under the title ‘Unravelling the Chemical Diversity of Actinobacteria Obtained from Unique Asian Ecological Niches’, that I only used the sources and materials cited in the work, and that all citations, whether word for word or paraphrased are given as such. I declare that I adhered to the guidelines set forth by the University of Tübingen to guarantee proper academic scholarship (Senate Resolution 25.05.2000). I declare that these statements are true and that I am concealing nothing. I understand that any false statements can be punished with a jail term of up to three years or a financial penalty.

Parts of this work have been published in:

Niraj Aryal, Saefuddin Aziz, Prajwal Rajbhandari, Harald Gross (2019) Draft genome sequence of *Nonomuraea* sp. strain C10, a producer of brartemycin, isolated from a mud dauber wasp nest in Nepal. *Microbiol Resour Announc* 8:e01109-19

Niraj Aryal, Saefuddin Aziz, Prajwal Rajbhandari, Harald Gross (2021) Draft Genome Sequence of the Sattazolin-Producing Strain *Pseudonocardia* sp. C8, Isolated from a Mud Dauber Wasp Nest in Nepal. *Microbiol Resour Announc*

.....

Place, Date

.....

Signature

Dedicated to
my
mother

“Imagination is more important than knowledge”

Albert Einstein.

Contribution of Other Scientists to This Work

Bioactivity analysis

The bioactivity of all the pure compounds was assessed by Mr. Jan Straetener and Dr. Anne Berscheid at Interfaculty Institute of Microbiology and Infection Medicine from the group of Prof. Heike Brötz-Oesterhelt (University of Tuebingen)

Stereochemistry determination

For the new streptocytosine and compound XI, the determination of the stereocenter of the side chain was conducted by Mr. Christian Geibel from the group Prof. Laemmerhofer group (University of Tuebingen).

Chemical facilities

HRMS measurements were performed by Dr. Dorothee Wistuba (Department of Mass Spectrometry, University of Tuebingen) and Dr. Markus Kramer facilitated the 700 MHz NMR measurements of novel streptocytosine, 4F-plicacetin, compound XI and compound 2.1 in this thesis.

16S rRNA Sequencing

Genomic DNA isolation and partial sequencing analysis of the Nepalese isolates was conducted by my colleague Dr. Saefuddin Aziz (member of the group of Prof. Gross).

Abstract

Actinobacteria play a vital role in pharmaceutical drug discovery therefore they are even considered as nature's pharmacists. This phylum of bacteria has been very prolific in the production of secondary metabolites and many thereof are used in the form of medicines serving the humankind. Most of the drugs that get prescribed today are mimics of nature that researchers around the globe were capable to understand. However, there is still a lot to discover and learn from the nature. In our study, we have relied on the concept that bioprospecting strains from underexplored geographical locations and unique Asian niches which will in return offer the opportunity of finding novel compounds for pharmaceutical purposes.

Following this approach, we designed a work to isolate strains from various ecological niches of Nepal, which is still a virgin land in terms of natural product research. This attempt represents therefore the pioneering work where we used Nepalese isolates to identify novel chemistry. Nepal has a varied climate and a broad bandwidth of altitudes (100 – 8849 meter) within in a short distance. These unique features make it one of the most biodiverse countries in all of Asia. Therefore, the first part of the thesis will mainly focus on soil-borne *Streptomyces* isolated from Nepal and their chemical analysis using a metabolomics-based workflow. Similarly, this study is continued with the investigation of the chemistry of rare actinomycetes isolated this time from a mud dauber wasp nests which are used as a traditional medicine in some parts of Nepal.

The second part of the thesis describes the isolation and structure elucidation of nucleoside antibiotics from the marine strain, *Streptomyces* sp. SHP-22-7 isolated from the underexplored Enggano Island, Indonesia. In this study four novel structures belonging to plicacetin/amicetin family were identified with a combined genomics and metabolomics approach. Upon bioactivity analysis, the isolated pure compounds possessed an anti-mycobacterial activity. Furthermore, we have also probed successfully the concept of precursor directed biosynthesis to generate novel unnatural plicacetin derivatives.

Zusammenfassung

Actinobakterien spielen eine wichtige Rolle bei der Entdeckung pharmazeutischer Wirkstoffe, weshalb sie sogar als „Pharmazeuten“ unter den Mikroben bezeichnet werden. Diese Bakterien-Klasse war in der Vergangenheit höchst ergiebig was die Entdeckung von Sekundärmetaboliten betrifft und einige davon werden bereits als Medikamente eingesetzt. Die meisten der Medikamente, die wir heute verschrieben bekommen, sind auf Naturstoffe zurückzuführen, die Forscher rund um den Globus entdeckt und aufgeklärt haben. Dennoch gibt es noch viel zu entdecken und von der Natur zu lernen. In dieser Studie vertreten wir die Hypothese, dass das Bioprospecting von Actinobakterien aus unerforschten geografischen Gebieten und einzigartigen asiatischen Nischen die Möglichkeit bietet, neuartige Verbindungen für pharmazeutische Zwecke zu finden.

Dieser Hypothese folgend, konzipierten wir eine Arbeit zur Isolierung von Actinobakterien aus verschiedenen ökologischen Nischen Nepals, welches hinsichtlich Naturstoffforschung noch als absolutes Neuland gilt. Dieser Ansatz stellt somit eine Pionierarbeit dar, bei der wir nepalesische Isolate zur Findung neuartiger Sekundärmetabolite verwendet haben. Nepal besitzt unterschiedliche Klimazonen und weist eine große Bandbreite an Höhenlagen auf (100 - 8849 Meter), welche sich beide auf engstem Raum abwechseln. Diese einzigartigen Eigenschaften machen es zu einem der artenreichsten Länder in ganz Asien.

Daher wird sich der erste Teil dieser Arbeit hauptsächlich auf die aus Nepal isolierten bodenbürtigen *Streptomyces*-Arten und deren chemische Analyse mit Hilfe eines Metabolomics-basierten Workflows konzentrieren. In ähnlicher Weise wird diese Studie mit chemischen Untersuchungen von seltenen Actinomyceten fortgesetzt, die diesmal aus den Nestern einer Grabwespe isoliert wurden, die in einigen Teilen Nepals als traditionelle Medizin verwendet wird.

Der zweite Teil der Arbeit beschreibt die Isolierung und Strukturaufklärung von Nukleosid-Antibiotika aus dem marinen Stamm *Streptomyces* sp. SHP-22-7, der wiederum auf der bisher wenig erforschten Insel Enggano, Indonesien gesammelt wurde. In dieser Studie wurden vier neue Sekundärmetaboliten, die zur Plicacetin/Amycetin-Familie gehören, mit einem kombinierten Genomik- und Metabolomik-Ansatz identifiziert. In Bioassays zeigten die isolierten Reinstoffe eine anti-mykobakterielle Aktivität. Darüber hinaus haben wir erfolgreich das Konzept der Präkursor-gesteuerten Biosynthese angewendet, um neue unnatürliche Plicacetinderivate zu erzeugen.

Acknowledgement

PhD is a journey not a race. And this journey of acquiring knowledge, patience and life skills is barely possible if someone does not get any support. In my journey from the beginning to this dissertation writing, I am thankful to so many people.

First of all, I would like to express my heartfelt thanks to my supervisor Prof. Harald Gross for believing in me. His experience, available funding and luxurious laboratory facility has always kept us afloat. His support in every aspect made this journey so easy despite of several challenges. Similarly, I would like to thank my second supervisor PD Dr. Bertolt Gust for constantly providing advice and recommendations whenever they were necessary. I would like to extend my thanks to Prof. Yvonne Mast from DSMZ and Prof. Heike Brötz-Oesterhelt from University of Tuebingen for accepting the offer to participate in my oral defence as examiners.

I am indebted to the DAAD, which provided me with the funding for 4 years and 4 months. Learning about culture, science, food, and society that all was made possible by this funding. They are the reason why I feel so good about Germany and all the German public (who were the taxpayers of the funding that I got). In addition, I would like to acknowledge the financial support that was provided by DAAD for printing of this thesis.

I would like to express my appreciation to the Research Institute for Bioscience and Biotechnology for providing the facility in Nepal for bacteria isolation. They also supported my visit to Nepal to attend an International Conference which was held in February 2020. In this row, I would also like to thank Amici Treks and Expedition, Nepal who has helped me with soil samples from Everest base camp. I would like to express my thanks to Prof. Yvonne Mast, Ira Handayani and Oliver Heinrich who has provided me with the strain *Streptomyces* SHP 22-7 and extracts.

I am thankful to all former members of the Gross group who helped me understand the lab in the very early days. Dr. Henrike Miess deserves special thanks for arranging the shipping for my Nepalese isolates to Germany. Dr. Julia and Ghazaleh were so calm and easygoing ladies. Humboldt scientists Prof. Hari Datta Bhattarai and Prof. Naheed Riaz were very inspiring and knowledgeable. I would like to thank Hamada Saad and wish him success as well. I had learned so much of science and academic skills from him in the early days. Fred and Aziz were good hearted and helpful. I thank Irina for taking good care of the NMR and helping me with measurements. Likewise, Mr. Kornberger, Patricia and recently Manuela facilitated the ordering process.

I would like to thank my countrymen Keshab. I have learnt some chemistry knowledge from him. Apart from all, Junjing Jiao was one of the close friends with whom I had lots for memories in Rebel Haus and in the lab. We talk so much of global politics while eating our lunch and admire each other.

I am thankful to all my group members, students who assisted me in my work and all who were connected directly and indirectly to this work. During my stay in Germany, I had enjoyed so many good cakes. I am thankful to all the guys who had birthday each year-round and offered us many tasty cakes “Kuchen”. I am thankful to Nhomsai for allowing me to join the coffee facility at sixth floor. Oh my! If I really start remembering each of memories and writing them in this book, I will probably write another dissertation on topic “THANK-YOU”.

Finally, I would like to thank my wife and entire family for being patient and allowing me to do, what I wanted.

Table of Contents

Contribution of Other Scientists to This Work	I
Abstract	II
Zusammenfassung	III
Acknowledgement	IV
Table of figures	VIII
List of tables	XII
Abbreviations	XIV
Chapter 1. Introduction	1
1.1 Microbes – Friend or Foe !.....	2
1.2 Microbial Natural Product (MNP)	3
1.1.2 Actinobacteria – nature’s pharmacist	4
1.1.3 Actinobacteria involved in this study	7
1.3 Modern approaches for analysis.....	8
1.1.4 Genome Mining and bioinformatics.....	8
1.1.5 LC-MS and metabolomics.....	9
1.1.6 OSMAC (One Strain MAny Compounds)	11
1.4 Mining Environment for chemical diversity	11
1.1.7 Geographical biodiversity	12
1.1.8 Diverse Ecological niches/habitats for strain collection	13
Aim of this research	17
Chapter 2. Material and methods	18
2.1 Material	19
2.2 Bacterial strains	25
2.3 Methods	25
2.3.1 Soil sampling and actinomycetes isolation and storage	25
2.3.2 Cultivation of bacterial strains	26
2.3.3 Measurement of growth curve and secondary metabolite production curve.....	28
2.3.4 Fractionation of crude extracts	30
2.3.5 Bioassays	30
2.3.6 DNA Extraction, 16S rRNA Amplification and 16S rRNA Sequencing	32
2.3.7 Genome sequencing	33
2.3.8 Chemical analysis.....	33
Chapter 3. NPs from Nepalese actinomycetes	40
3.1 <i>Streptomyces</i> sp. WL006, <i>Streptomyces</i> sp. NB004.....	41
3.1.1 Sampling and pre-screening	41
3.1.2 Antimicrobial assay of crude extracts	42
3.1.3 Strain characterization	44
3.1.4 Chemical Analysis.....	45
3.2 Non- <i>Streptomyces</i> - <i>Pseudonocardia</i> sp. C8 and <i>Nonomuraea</i> sp. C10).....	57
3.2.1 Sampling and pre-screening	57
3.2.2 Sequencing and Bioinformatic analysis	58
3.2.3 Chemical analysis of <i>Nonomuraea</i> sp. C10	62
3.2.4 Chemical analysis of <i>Pseudonocardia</i> sp. C8.....	72
3.3 Overall discussion and conclusion	75
Chapter 4. New nucleoside antibiotics from <i>Streptomyces</i> sp. SHP 22-779	
4.1 Bioinformatic analysis of <i>Streptomyces</i> sp. SHP 22-7	80
4.2 HR-LC-MS/MS based dereplication and target identification.....	83

4.3	Growth characteristics.....	86
4.4	Fractionation and isolation of compounds	87
4.5	Structure elucidation of Plicacetin and derivatives	90
4.5.1	Plicacetin/Amicetin	90
4.5.2	Compound I.....	92
4.5.3	Compound IV.....	95
4.5.4	Compound VI.....	96
4.5.5	Compound XI.....	98
4.5.6	Structure speculations for compound VII, VIII, IX, X and Xa.....	101
4.5.7	Compound XII	103
4.6	Structure elucidation of streptocytosines	105
4.6.1	Streptocytosine New*	107
4.7	Precursor directed biosynthesis (PDB)	108
4.7.1	Feeding of halogenated benzoic acid	110
4.8	Bioassay of isolated plicacetin and derivatives.....	116
4.9	Overall discussion	117
	References.....	122
	Supplementary Information	131

Table of figures

Figure 1-1 WHO priority pathogens list for R&D of new antibiotics	2
Figure 1-2 Drugs approved from 1981-2019.....	3
Figure 1-3 Taxonomy of Phylum Actinobacteria.....	4
Figure 1-4 a. Successful drugs from genus <i>Streptomyces</i> b. Timeline microbial drug discovery including <i>Streptomyces</i>	6
Figure 1-5 Drugs from non-streptomyces	7
Figure 1-6 Two workflows for NP based drug discovery	10
Figure 1-7 Top 10 most biodiverse countries	12
Figure 1-8 Sampling sites in Nepal and Indonesia, this study.....	13
Figure 2-1 Selective isolation of actinomycetes from soil	26
Figure 2-2 A) Enzyscreen minibioreactor used for small scale cultivation B) G,T,S,M,I,Sm as medium initials representing six different media	27
Figure 3-1 Map of Nepal showing districts of soil sample collection.	41
Figure 3-2 Antimicrobial activities shown by crude extracts of strain <i>Streptomyces</i> sp. WL006 and <i>Streptomyces</i> sp. NB004 against <i>Bacillus subtilis</i> , <i>Escherichia coli</i> , <i>Mycobacterium phlei</i>	44
Figure 3-3 BPC (up) and LC (down) of <i>Streptomyces</i> sp. WL006 crude extract in positive mode	45
Figure 3-4 Structure of nonactin and other macrotetralides ¹⁰²	46
Figure 3-5 LC-MS spectra for nonactin ($m/z=737.7$), monactin ($m/z=751.6$), dynactin ($m/z=765.4$), trinactin ($m/z=779.5$) and tetranactin ($m/z=793.7$) along with their ammonium, sodium and potassium adducts.	46
Figure 3-6 XIC and spectra for linear form of nonactin, monactin, dinactin, trinactin, tetranactin with m/z 755.8, 769.7, 783.7, 797.8 and 811.9.....	47
Figure 3-7 Molecular network clustering nonactin, its linear form and ammonium adduct. (Red=Crude, Sky blue=Fraction F, Purple=Fraction E and Gray=Medium)	48
Figure 3-8 Extracted ion chromatogram (EIC) of new trilactones eluting earlier than known trilactones	49
Figure 3-9 Structure of known trilactones	49
Figure 3-10 MN annotated for polyether ionophores	50
Figure 3-11 LC-MS showing nonactin dimers and their derivatives along with their unsaturated congeners.....	52

Figure 3-12 Griseorhodin and derivatives.....	53
Figure 3-13 Griseorhodin and hyaluromycin clustered together (Red=Crude, Purple=Fraction E, Sky Blue=Fraction F and Gray=Medium).....	54
Figure 3-14 Fragmentation of griseorhodin A and its fragment similarity charts to other derivatives from same family (including hyaluromycin).....	54
Figure 3-15 MN for <i>Streptomyces</i> sp. NB004 with actinomycin cluster and actinomycin D structure.....	56
Figure 3-16 a. <i>Sceliphron</i> wasp in Nepal (Original picture) b. Wasp Nest. c. <i>Sceliphron</i>	57
Figure 3-17 Antismash results for strain C8	60
Figure 3-18 Antismash results for strain C10	61
Figure 3-19 Predicted BGC for Orsellinic acid from <i>Nonomuraea</i> sp. C10.....	61
Figure 3-20 100% similarity of MSAS/OSAS AT domain(red) and ACP domain from query sequence to avilamycin and everminomycin.	61
Figure 3-21 <i>Nonomuraea</i> sp. C10 fractions (W=Water in Water:MeOH mixture)	63
Figure 3-22 HPLC chromatogram of fraction W20 showing brartemicin (3), orsellinic acid ester (4) and peak 1, peak 2.1 and 2.3 which are proposed new derivatives of brartemicin.	63
Figure 3-23 Absorption spectra (Top) of brartemicin and derivatives in W20 fraction (Pink=Baseline, Purple =Peak 1, Green = Peak 2.1, Blue = Brartemicin, Red = OA ester).....	64
Figure 3-24 MS ² showing brartemicin and their novel derivatives.	65
Figure 3-25 Structure of compound 1	65
Figure 3-26 Structure of new brartemicin derivative	68
Figure 3-27 Structure of Brartemicin.....	70
Figure 3-28 n-butyl orsellinate.....	71
Figure 3-29 HPLC chromatogram showing media (top) and crude extract (bottom) chromatogram of <i>Pseudonocardia</i> sp. C8.....	72
Figure 3-30 Structure of Sattazolin showing key correlations.....	73
Figure 3-31 A) General scheme for acyloin condensation. B) Biosynthesis of sattazolin catalysed by enzyme Cbei2730 (ThDP-dependent).....	73
Figure 3-32 NonR-catalyzed stereoselective hydrolysis creating dimers.	76
Figure 3-33 Comparison of A) NonR catalyzed dinactin dimer and B) bonactin.....	77
Figure 3-34 RiPP like masses detected in LC-MS analysis of fraction 1 from <i>Pseudonocardia</i> sp. C8 with fragment similarity. A) Type 1 B) Type II	78

Figure 4-1 Comparison of gene clusters coding for amicetin and transport-related proteins (green) between <i>Streptomyces vinaceusdrappus</i> NRRL 2363 (top panel) and <i>Streptomyces</i> sp. strain SHP 22-7 (bottom panel).....	82
Figure 4-2 The structure of fragments $[M+H]^+=174$ and $[M+H]^+=288$	83
Figure 4-3 Extraction Ion Chromatogram (EIC) of $[M+H]^+=174$ & $[M+H]^+=288$ (Red) and Base Peak Chromatogram (blue) of the crude extract from <i>Streptomyces</i> sp. SHP-22-7	84
Figure 4-4 The growth curve of <i>Streptomyces</i> sp. strain SHP 22-7 in NL-300 media.....	86
Figure 4-5 The plicacetin derivatives production profile at 30 time points	87
Figure 4-6 Fractionation scheme	88
Figure 4-7 HPLC chromatogram showing the pure compound I (21 min) and its decomposition product (14.5 min).....	89
Figure 4-8 HPLC chromatogram of Fraction 3 showing plicacetin, amicetin and other derivatives.....	89
Figure 4-9 Key correlations for plicacetin.....	92
Figure 4-10 HRMS spectrum of compound I in positive mode	92
Figure 4-11 Key correlations of compound I	94
Figure 4-12 Figure showing MS ¹ and MS ² data for compound IV along with structure.	95
Figure 4-13 Structure of Compound IV with key correlations.....	96
Figure 4-14 A) MS ¹ and MS ² spectra of compound VI. B) Explanation of fragments from MS ² spectra.....	97
Figure 4-15 Aromatic region H ¹ -H ¹ COSY (red), other correlations were similar to the plicacetin standard.....	98
Figure 4-16 HRMS spectrum of pure compound XI in positive and negative mode	98
Figure 4-17 Key correlations of the ¹ H- ¹³ C HMBC (arrow) and ¹ H- ¹ H COSY (bold bonds) spectra of compound XI	100
Figure 4-18 Base peak chromatogram and UV chromatogram showing elution of plicacetin derivatives at different retention times.	102
Figure 4-19 HPLC chromatogram showing compound XII	103
Figure 4-20 Key correlations for compound XII.....	105
Figure 4-21 Structure of streptcytosines isolated in this study.....	106
Figure 4-22 HPLC separation of streptcytosines.....	106
Figure 4-23 Structure showing key correlation for novel streptcytosine	108
Figure 4-24 43 precursors employed in the precursor-directed biosynthesis study.	109

Figure 4-25 The ratios of compound VI between cultivation under regular fermentation conditions (left) and with supplementation of the precursor-directed biosynthesis (right)	109
Figure 4-26 LC-MS chromatogram between sample fed with precursor 1a (arrow marked) and media control (blue) in positive mode. The arrow indicates compound VI.....	110
Figure 4-27 Comparative LC-MS profiling of a sample fed with precursor 2c (red) and media control (blue) in positive mode. The arrow represents precursor attached product.	111
Figure 4-28 Comparative LC-MS profiling of a sample fed with precursor 3c (red) and media control (blue) in positive mode. The arrow represents precursor attached product.	111
Figure 4-29 Comparative LC-MS profiling of a sample fed with precursor 4c (A), 4d (B) 4e (C) (red) and media control (blue) in positive mode. Isotope pattern is also shown in left. The arrow represents precursor attached product.....	112
Figure 4-30 ¹⁹ F NMR spectrum for 4F-plicacetin (red) compared to standard value and peak shape. Note: Difference in ppm observed is due to the use of different solvent.....	112
Figure 4-31 4F-plicacetin ¹³ C NMR compared to the plicacetin standard.....	113
Figure 4-32 Single fluorine substituent ¹³ C shifts and coupling constants ¹⁶⁸	113
Figure 4-33 ¹³ C NMR showing coupling constants for ortho-, meta- and para-carbon on a benzene ring with para-substituted fluorine atom.....	114
Figure 4-34 Structure of 4F-plicacetin with key correlations	114
Figure 4-35 Comparative LC-MS profiling of a sample fed with precursor 5c (red) and media control (blue) in positive mode. The arrow represents precursor attached product.....	115
Figure 4-36 Comparative LC-MS profiling of a sample fed with precursor 5d (red) and media control (blue) in positive mode. The arrow represents precursor attached product.....	115
Figure 4-37 Comparative LC-MS profiling of a sample fed with precursor 5e (red) and media control (blue) in positive mode. The arrow represents precursor attached product.....	115
Figure 4-38 Comparative LC-MS profiling of a sample fed with precursor 5h (red) and media control (blue) in positive mode. The arrow represents precursor attached product.....	116
Figure 4-39 Activity profile for plicacetin and derivatives (Green – high activity, parrot green – moderate activity, yellow – low activity).....	117
Figure 4-40 Possible methylation and demethylation of known compounds to produce new congener	118
Figure 4-41 Structural comparison of 40551-D and Compound XII	119
Figure 4-42 Tentative structures for compound XII congeners	119

List of tables

Table 2-1 Media used in OSMAC approach.	19
Table 2-2 Other Media and buffer solutions.....	20
Table 2-3 Chemicals used in precursor-directed biosynthesis (PDB)	21
Table 2-4 Chemicals	23
Table 2-5 Employed instruments.....	24
Table 2-6 Auxiliary materials	24
Table 2-7 The time schedule for growth and secondary metabolite production curve.....	29
Table 2-8 The concentration gradient used in SPE for <i>Streptomyces</i> SHP22-7	30
Table 2-9 Open column C18 fractionation scheme for <i>Nonomuraea</i> sp. C10.....	30
Table 2-10 HPLC method for plicacetin and its derivatives isolation.....	33
Table 2-11 HPLC method for Brartemicin and derivatives isolation	34
Table 2-12 HPLC columns used in this study	34
Table 2-13 Low resolution LC-MS parameters	35
Table 3-1 Soil habitats for sample collection	41
Table 3-2 Crude extract activity of 2 prioritized strains	42
Table 3-3 OSMAC approach and the resultant antimicrobial activity profile in six different culture media. Media*, G=GYM, T=TSB, S=S/N, M=MM, I=ISP4, Sm=Seed media.....	43
Table 3-4 16S rRNA of prioritized isolates	44
Table 3-5 Trilactones and their proposed new derivatives A-E	49
Table 3-6 Dimers of nonactic acid and homononactic acid.....	51
Table 3-7 Rubromycin class of compounds in identified in this study.....	55
Table 3-8 Actinomycin annotation table with details	56
Table 3-9 16S rRNA of 4 colonies isolated from mud dauber nest soil.	59
Table 3-10 NMR signals of compound 1 recorded in <i>d</i> ₆ -DMSO (400 MHz).....	66
Table 3-11 NMR data of compound 2.1 recorded in <i>d</i> ₆ -DMSO (700 MHz).....	68
Table 3-12 NMR signals of compound 3 recorded in <i>d</i> ₆ -DMSO (400 MHz).....	70
Table 3-13 NMR signals of compound 4 recorded in <i>d</i> ₆ -DMSO (400 MHz).....	71
Table 3-14 NMR data of C8P7 recorded in <i>d</i> ₆ -DMSO (400 MHz).....	74
Table 4-1 Information of contigs coding for secondary metabolite in <i>Streptomyces</i> sp. strain SHP 22-7	80
Table 4-2 Corresponding BGCs coding for amicetin between <i>Streptomyces</i> sp. strain SHP 22-7 and <i>Streptomyces vinaceusdrappus</i> NRRL 2363	81

Table 4-3 Nucleoside antibiotics characterized in this study	84
Table 4-4 Peaks and their designated masses (also refer to Figure 4-8). Gray highlighted are all fully characterized compound from this list	90
Table 4-5 NMR data of plicacetin recorded in <i>d</i> ₄ -MeOD (400 MHz)	91
Table 4-6 NMR data of compound I recorded in <i>d</i> ₆ -DMSO (400 MHz)	94
Table 4-7 NMR data of compound I recorded in <i>d</i> ₄ -MeOD (400 MHz).....	95
Table 4-8 NMR signals of compound XI recorded in <i>d</i> ₄ -methanol (700 MHz)	100
Table 4-9 The absorbance at specific wavelength and corresponding λ_{\max} (log ϵ) values for compound XI	101
Table 4-10 Derivatives with same mass but proposed different structures.....	102
Table 4-11 NMR data for compound XII recorded in <i>d</i> ₄ -MeOD (400 MHz).....	104
Table 4-12 Specific wavelength and corresponding λ_{\max} (log ϵ) values for compound XII....	105
Table 4-13 NMR data for streptcytosine (New*) recorded in <i>d</i> ₆ -DMSO (700 MHz).....	107
Table 4-14 The overview for all successfully incorporated precursors	116

Abbreviations

°C	degrees Celsius
1D	one dimensional
2D	two dimensional
$[\alpha]_D^T$	specific optical rotation (T = temperature; D = sodium D line (589 nm))
Δ	difference
δ	NMR chemical shift (ppm)
ϵ	(UV spectroscopy) molar attenuation coefficient
λ	wavelength
μ	micro (10^{-6})
Å	Angstrom
ACP	Acyl Carrier Protein
ATCC	American Type Culture Collection
A	Adenine
AT	acetyltransferase
BGC	Biosynthetic Gene Cluster
CID	Collision Induced Dissociation
d	(NMR) doublet
dd	(NMR) doublet of doublets
Da	Dalton
DAD	Diode Array Detector
DEPT	Distortionless Enhancement by Polarization Transfer
DMSO	dimethyl sulphoxide
ESI	electrospray ionization
EDTA	Ethylenediaminetetraacetic acid
FA	Formic Acid
FDA	Food and Drug Administration
g	gram
G	Guanine
GNPS	Global Natural Product Social molecular networking
HCl	hydrochloric acid
HMBC	Heteronuclear Multiple Bond Correlation
HR	High Resolution
HSQC	Heteronuclear Single Quantum Coherence
Hz	Hertz
IR	Infrared
J	Spin-spin coupling constant [Hz]
KR	ketoreductase
KS	ketosynthase
L	liter
LC	Liquid Chromatography
LR	(MS) low resolution
m/z	mass to charge ratio
mL	milliliter
mM	millimolar
m	(NMR) multiplet
mg	milligram
min	minutes

M	Molar
MeOH	methanol
MIC	Minimum Inhibition Concentration
MHz	Megahertz
MRSA	Methicillin-resistant Staphylococcus aureus
MS	Mass Spectroscopy
MS/MS or MS ²	Tandem Mass Spectrometry
NCBI	National Center for Biotechnology Information
NMR	Nuclear Magnetic Resonance
NP	Natural product
NOESY	Nuclear Overhauser Effect Spectroscopy
NRPS	Non-Ribosomal Peptide Synthetase
PKS	Polyketide Synthase
PCR	Polymerase Chain Reaction
RP	Reversed Phase
s	(NMR) singlet
sp	Species
SPE	Solid Phase Extraction
SAM	S-adenosyl methionine
t	(NMR) triplet
t _R	retention time
R & D	Research and Development
RPMI	Roswell Park Memorial Institute
rpm	revolution per minute
TFA	Tri Fluro Aceticacid
Thr	threonine
TOCSY	Total Correlation Spectroscopy
UV	ultraviolet (spectroscopy)
VIS	visible (spectroscopy)
WHO	World Health Organization

Chapter 1. Introduction

1.1 Microbes – Friend or Foe !

The world in which we live is inhabited by a diverse microscopic form of life that we commonly refer to as microbes. Microbes bear an intimate relation to the cycle of life in nature, and thus affect our health, our food supply, our shelter and clothing, and many of our industrial processes¹. Microbes in form of pathogens were considered human enemies for the past centuries. Infact some of them still are. In year 2017, World Health Organization (WHO) has published list of so called “priority pathogens” a catalogue of 12 families of bacteria that pose the greatest threat to human health² (see **Figure 1-1**). This situation is predicted to get much worse by 2050 making resistant pathogens a leading cause of mortality³.

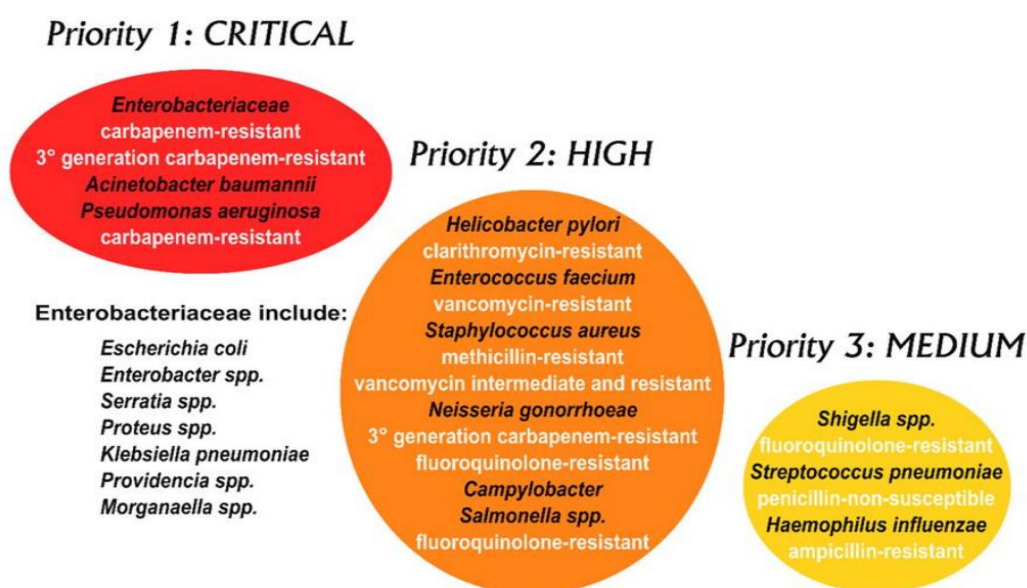


Figure 1-1 WHO priority pathogens list for R&D of new antibiotics^{2, 4}. Note that Mycobacteria are not included in this diagram, as they have, for some time, been considered separately as an absolute global priority.

Meanwhile on the other hand, microbes are considered our greatest of friend in the form of microbiomes. They are considered to be “essential organ” specially in the gut, which helps in digestion and in vivo defence by boosting our immune response⁵. In parallel, microbes have been extremely generous in offering human life-saving therapies and chemical entities. Understanding their biology and chemistry has been always fascinating. This study is more focused on microbial natural products (MNP).

1.2 Microbial Natural Product (MNP)

Natural products from microbes and plants make excellent drugs. Before the advent of high-throughput screening and the post-genomic era: more than 80% of drug substances were natural products or inspired by a natural compound⁶. This finding has been recently updated by Newmann and Cragg in 2020, stating that out of 1881 total FDA approved drugs until September 2019 more than 70 percent of drugs have still have roots to natural products.⁷ (see **Figure 1-2**)

In specific MNP were more successful as antibiotics, immunosuppressants, antiviral, antiparasitic and antitumor agents. These structurally diverse and complex molecules have demonstrated an impressive record of efficacy as pharmaceutical agents and are increasingly important in the treatment of a range of serious diseases, which eventually have played a substantial role in increasing the life expectancy of humans.

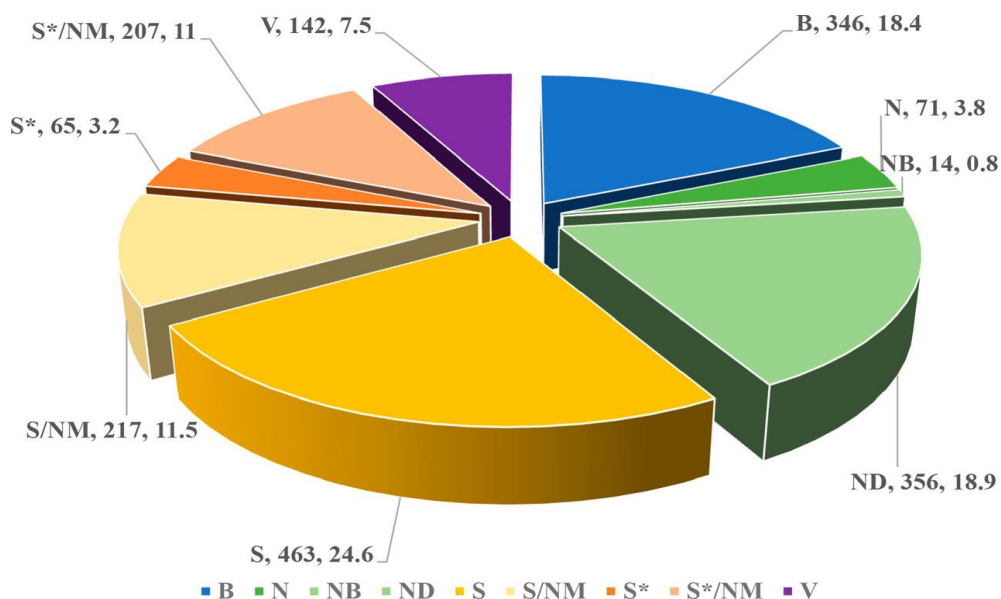


Figure 1-2 Drugs approved from 1981-2019, Codes used : B = biological macromolecule, N = unaltered natural product, NB = botanical drug (defined mixture), ND = natural product derivative, S = synthetic drug, S* = synthetic drug (NP pharmacophore), V = vaccine, /NM = mimic of natural product⁷

Approaching 21st century, MNP discovery faced serious challenges. As re-isolation of known compounds has become a major frustration in the drug discovery process, leading to fewer and fewer novel compounds being found. Many pharmaceutical companies started diverting from natural products research over the past 30 years and focus more on high throughput screening (HTS) technologies and molecular modelling for drug discovery. But, these types of screens

have largely been failures due to lack of viable hits found^{8, 9}. These failure attempt in turn increases again the cost of drug development. On an average, it takes \$985 million and 10–15 years to bring a new drug to market. It even gets worse for antibiotics as they also face resistance concerns from WHO and general public.

Despite these challenges, it is an undeniable fact that we need novel drug candidates to fight infectious diseases. In that context, MNP still holds potential. Among the bioactive compounds that have been obtained so far from microbes, 45 % are produced by actinomycetes, 38 % by fungi and 17 % by unicellular eubacteria¹⁰.

1.1.2 Actinobacteria – nature’s pharmacist

The phylum Actinobacteria, which is comprised mainly of Gram-positive organisms with a high G+C content (>55 mol% in genomic DNA), constitutes one of the largest phyla within the Bacteria). The different genera that are part of this phylum exhibit an enormous diversity in terms of their morphology, physiology, and metabolic capabilities¹¹. The morphologies of actinobacterial species vary from coccoid (e.g., *Micrococcus*) or rod-coccoid (e.g., *Arthrobacter*) to fragmenting hyphal forms (e.g., *Nocardia*) or highly differentiated branched mycelia (e.g., *Streptomyces*)¹². An updated taxonomy for Actinobacteria is published in the *Bergey's Manual of Systematic Bacteriology*¹³. (see **Figure 1-3**)

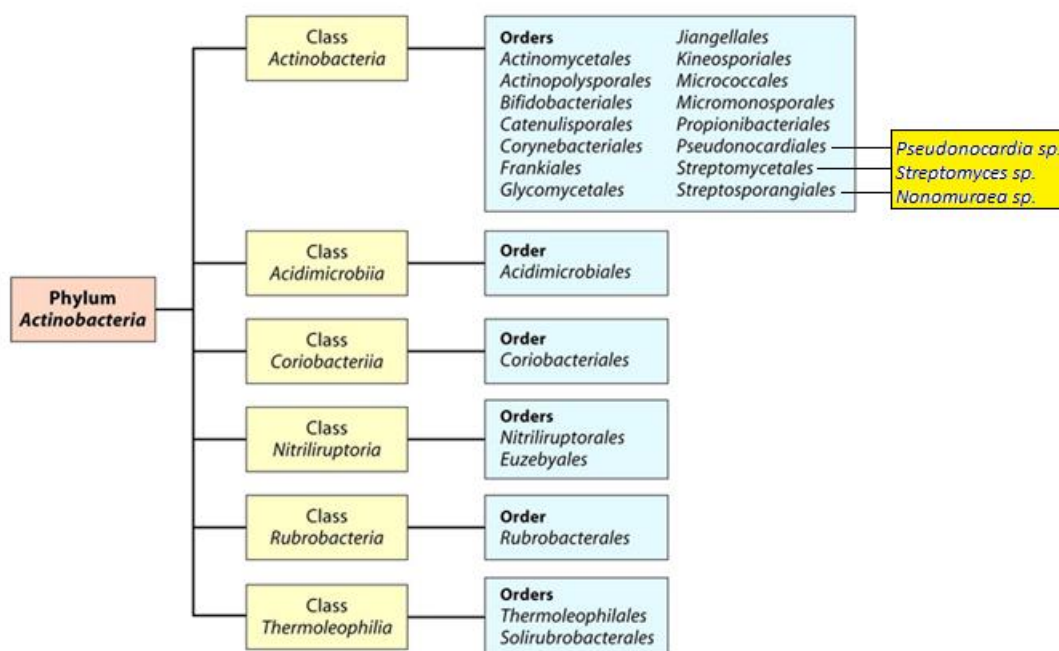


Figure 1-3 Taxonomy of Phylum Actinobacteria

Introduction

These bacteria are ubiquitous in the environment and are capable of both solitary inhabitation and forming symbioses, not only with other micro-organisms but also higher order creatures¹⁴. The largest genus of Actinobacteria, *Streptomyces* has been studied extensively in past decades.

1.2.1.1 The genus *Streptomyces*

Streptomyces are member of order Streptomycetales. *Streptomyces* genus have been proven as bountiful source of bioactive chemicals, contributing two-thirds of the clinically applicable antibiotics and a wide range of industrially important enzymes¹⁵. Actinomycin and streptomycin were among the first natural products to be approved by FDA from this genus. From Nobel Laureates Prof. Waksman actinomycin¹⁶ to Prof. Omura's ivermectin¹⁷, streptomycetes NPs has grasped great attention which led the golden age of antibiotic discovery. Successful drugs produced from genus *Streptomyces* and their discovery timeline are highlighted in **Figure 1-4**. Today, most of the research groups are deviating from explorative *Streptomyces* research because of rediscovery problem. Non-streptomycetes (**section 1.2.1.2**) and gram-negative bacteria are getting more and more attention in recent days. However, with modern computational tools, sophisticated techniques and careful observation one could squeeze more from this genus. In addition, special niches should be taken into consideration during the isolation of samples in order to enhance the novelty of isolates. Several new compounds from endophytic *Streptomyces* sp. were discovered by Ling Ding and her research team over the span of 7 years from mangrove isolated *Streptomyces* sp. This contains xiamycin¹⁸, kandenol A-E¹⁹, bacaryolanes A-C²⁰, divergolides A-D²¹ and recently azodyrecins A-C²². Several derivatives of other known class of molecules were also isolated during the investigation. Similarly, Currie and coworkers isolated several antimicrobials from insect associated *Streptomyces* sp²³. This indicates that there is still more chemistry hidden in *Streptomyces* sp.

In this work I have studied two *Streptomyces* isolates from diverse geographical location and during the chemical analysis I have successfully annotated several known chemical entities along with some new derivatives.

Introduction

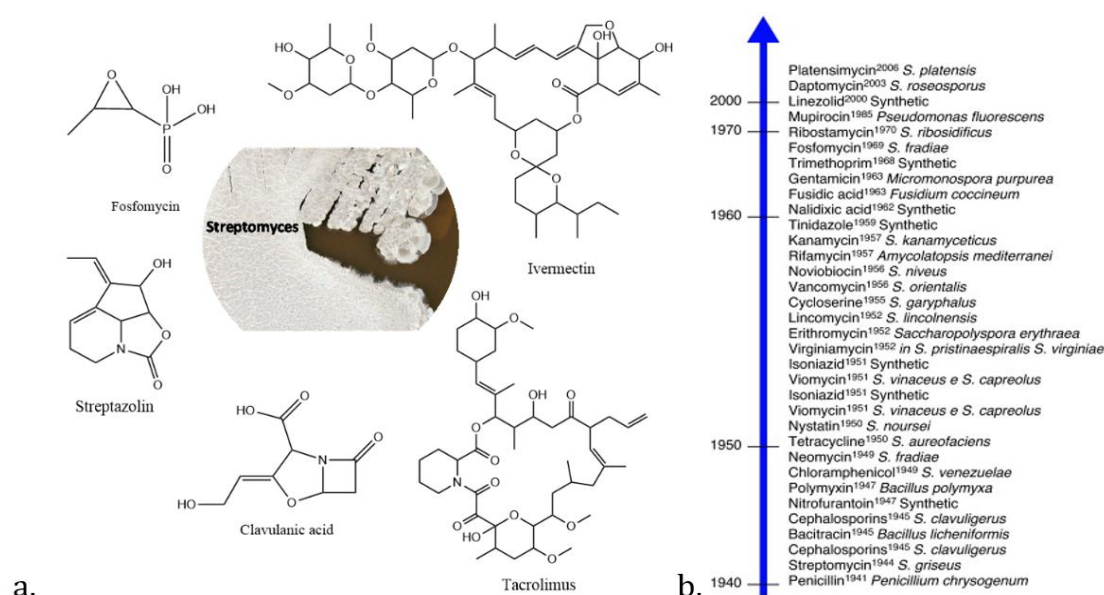


Figure 1-4 a. Successful drugs from genus *Streptomyces* b. Timeline microbial drug discovery including *Streptomyces*²⁴

1.2.1.2 Non-*Streptomyces* Actinobacteria

Since 1940s it is evident that there was a clear bias over one genus. *Streptomyces* natural product have been extensively studied and have also proven to be worthy as well. However, recently novel chemical scaffold has been also identified from non-streptomyces genus (see **Figure 1-5**). Among the non-streptomyces actinomycetes, members affiliated with the family *Micromonosporaceae* are more productive, followed by *Thermomonosporaceae*, *Streptosporangiaceae* and *Pseudonocardiales*. The non-streptomyces actinomycetes have been isolated from different environments, with varying diversity, endemicity, and potential to produce novel bioactive compounds. In laboratory conditions they grow slower, and most of them require a distinctive procedure for isolation, preservation, and cultivation²⁵. In past studies several research groups have recovered these bacteria from unique habitats such as caves, deserts²⁶, insects²⁷, deep sea sediments²⁸ etc.

Dalbavancin is a commercially available antibiotic which is synthesized from the A40926²⁹, a natural precursor produced from *Nonomuarea* sp. It is used to treat acute bacterial infections of the skin and structures of the skin (tissue under the skin) such as cellulitis (inflammation of the subcutaneous tissue), skin abscesses and wound infections.

Likewise, Gentamicin³⁰ is well known aminoglycoside antibiotic which were known to be produced from genus micromonospora. In this study, two non-streptomyces species are considered namely, *Nonomuraea* sp. C10 and *Pseudonocardia* sp. C8 for further chemical investigation.

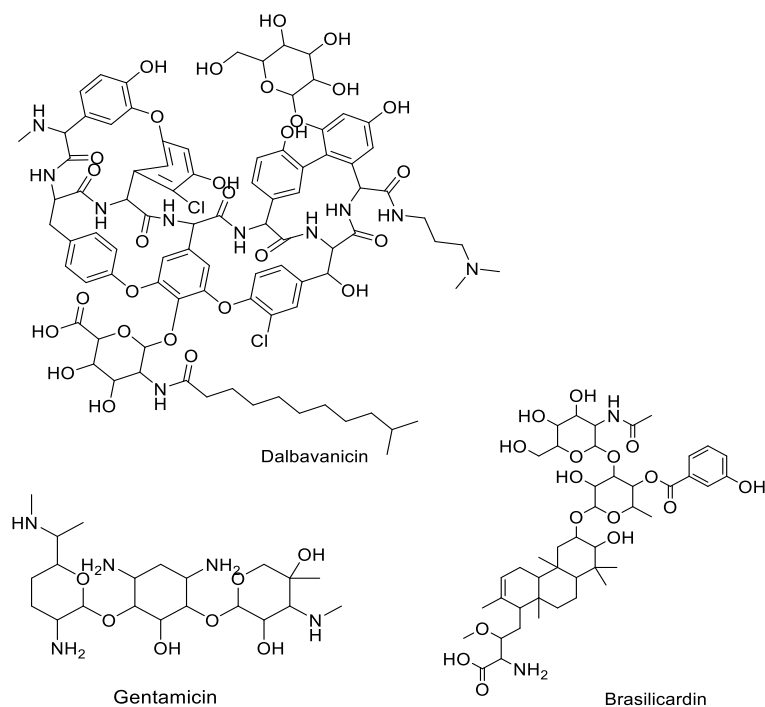


Figure 1-5 Drugs from non-streptomycetes

1.1.3 Actinobacteria involved in this study

1.2.1.3 Genus *Streptomyces* investigated in this study

1.2.1.3.1 *Streptomyces* sp. Strain SHP-22-7

Streptomyces sp. SHP 22-7 was obtained in collaboration with the Mast lab within the University of Tuebingen. It was originally collected from soil around mangrove roots on Enggano Island³¹. Using Pacific Biosciences RS II platform and Hierarchical Genome Assembly (HGAP) V.3.3 the 7.9-Mbp genome was sequenced with a 6-fold coverage and assembled into 146 contigs with a GC content of 72.20%. In sequence analysis it was evident that this strain harbors amicetin/plicacetin gene cluster (70% match). Mast Lab focused more on the regulation of plicacetin with *papR2* guided activation³². Meanwhile, we discovered novel derivatives from this strain during our chemical analysis. The good portion of this thesis focuses on isolation and characterization these novel nucleoside compounds.

1.2.1.3.2 Streptomyces sp. WL006 and Streptomyces sp. NB004

While on the other hand, during the early days, chemical investigations were carried out for several neplease isolates, of which isolate *Streptomyces sp. WL006* and *Streptomyces sp. NB004* will be more highlighted in this thesis. Collection of these isolates were done in collaboration with Research Institute of Bioscience and Biotechnology (RIBB), Nepal. Both Nepalese isolates were recovered from muddy soil near Bees Hazari Lake, Chitwan. During their initial bioactivity analysis, crude of both isolates showed prolific bioactivities. Of 22 bioactive prioritized isolates two best *Streptomyces sp.* were chosen for further chemical analysis which will be presented in this thesis.

1.2.1.4 Non-streptomyces in this study.

1.2.1.4.1 Nonomuraea sp. C10

As in other countries^{33,34}, in Nepal, the mud dauber wasp (*Screliphron sp.*) nest is used in folk medicine. To investigate if microbes are involved in its activity, nest-associated actinobacteria, *Nonomuraea sp. C10* was isolated. Mud dauber wasp nests were obtained from wooden houses in the Bardia district of the southwestern region of Nepal. During chemical analysis crude extracts of *Nonomuraea sp. C10* did not possess any antimicrobial activity. However, here we followed structure based approach of compound identification it was apparent that *Nonomuraea sp. C10* produced brartemicin, an antitumor compound³⁵.

1.2.1.4.2 Pseudonocardia sp. C8

Pseudonocardia sp. C8 was also obtained from the same niche where *Nonomuraea sp. C10* was isolated.

1.3 Modern approaches for analysis

1.1.4 Genome Mining and bioinformatics

The post-genomic world has revealed that the traditional culture-based, bioassay-guided strategies used to discover natural products have only provided access to a small fraction of the biosynthetic capacity encoded in genomes. It has become apparent that most biosynthetic pathways are expressed not at all, or scarcely, under laboratory conditions, or the products of

Introduction

these pathways have been overlooked. Genome mining seeks to address both of these issues and thus to exploit the hidden potential of biosynthetic pathways³⁶. This genes-to-molecules foundational knowledge has enabled the identification of cryptic natural product gene clusters from microbial genomes and facilitated the genomics guided discovery of novel natural products. NGS techniques in combination with genome mining approaches revolutionized the field of antibiotic research. Genome mining vastly depends on the quality of data i.e. sequence generated. Therefore, development of technologies for producing high quality genome sequences are also most vital. With (whole genome) sequencing becoming cheaper, more accurate and more efficient, it is a favoured technique for the characterisation of organisms and diseases on a systematic level. Most of the advantages were acquired in *Streptomyces* genome sequencing projects over the past decades. In 2002, the first *Streptomyces* genome sequence was published³⁷. This was the genome sequence of the model actinomycete *Streptomyces coelicolor*. Since then advances in genome sequencing and bioinformatic tools have permitted us to collect a wealth of knowledge and data on the biosynthesis of natural products, which has revealed the broad untapped biosynthetic diversity of microorganisms, especially actinomycetes³⁸. Among several in silico tools, PRISM³⁹ and antiSMASH Database^{40, 41} has been widely used for prediction of secondary metabolite biosynthetic gene cluster.

Genome mining of *Streptomyces curacoii* led to the discovery of new precursor peptide gene for ribosomally synthesized and post-translationally modified peptides (RiPPs). Purifying and determining the structure of the product of this RiPP BGC using ESI-MS and NMR resulted in the discovery of new cytotoxic compound, curacozole⁴². Similarly, a new lasso peptide, huascopeptin, was isolated following genome-mined discovery of a new biosynthetic gene cluster in extremotolerant *Streptomyces huasconensis* HST28T from Salar de Huasco, Atacama Desert, Chile⁴³.

1.1.5 LC-MS and metabolomics

Chemical diversity within an individual producer is massive with some of the key molecules produced in low amount. In the past, screening and isolation efforts are suspect of having harvested mainly the “low-hanging fruit”, leaving significant numbers of novel compounds to be discovered by means of improved methods for their detection and isolation⁴⁴.

Mass spectrometry has been primarily used in order to obtain m/z of compound and for deduction of the molecular formula. Nowadays it has evolved with wider applications. Incorporating mass spectrometry early in screening platforms could give us a wealth of structural information which

Introduction

can ease the process of dereplication. **Figure 1-6** describes two mostly used approaches. Bioactivity-based discovery are popular since the early times. Normally, it guides the isolation process for obtaining one or more bioactive molecule. It has been tremendously successful leading to discovery of several antibiotics. However, this process has been biased towards bioactive fractions leaving several other structural entities under explored. Another approach for isolation is Structure-based discovery, which emphasizes the use of spectrometric data early in the pipeline to quickly dereplicate known compounds and scaffolds, allowing researchers to save resources and focus on unknown compounds. The biggest advantage of time and sensitivity has been gained in the later approach. It has been used in synthetic biology efforts to link cryptic gene clusters to a natural product through genome mining efforts and to prioritize strains for natural product isolation⁴⁵.

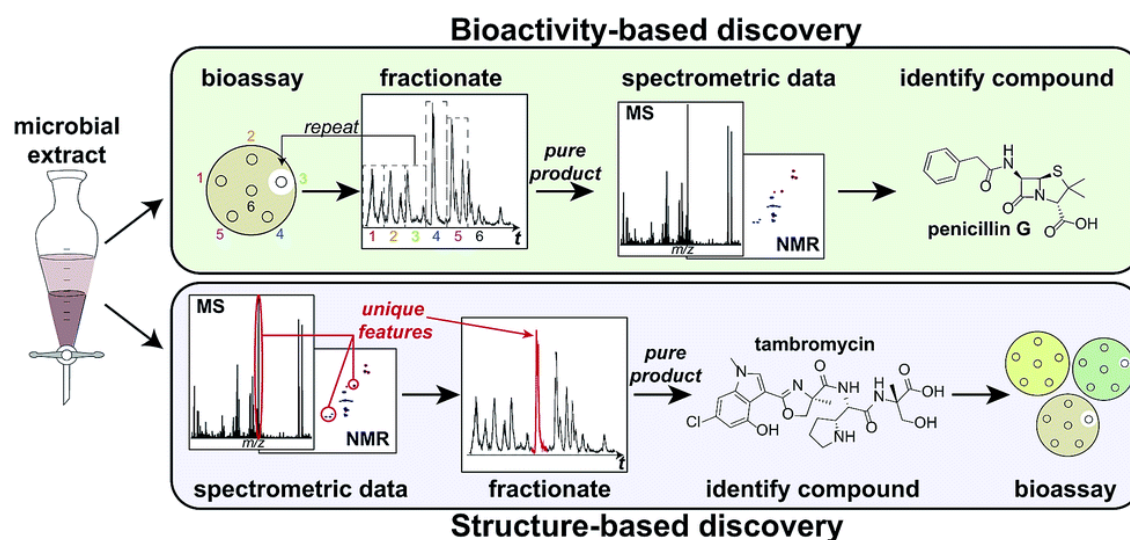


Figure 1-6 Two workflows for NP based drug discovery⁴⁵

The recent development in the field of metabolomics with several powerful algorithms and easy user interfaces has helped the scientific community to discover newer chemical scaffold. Starting with detection and ending up with annotation there are plenty of software and open-source platforms developed. The goals of these tools were - reducing screening time, reducing the number of false positives, and expanding the scope of each screen finally a trustworthy annotation of known compound. This will eventually allow for a shorter and less costly discovery time depending on the instrumentation used. One example is SIRIUS⁴⁶ (<https://bio.informatik.uni-jena.de/sirius/>) for which author claims identification rates of more than 70% on challenging metabolomics datasets. In addition, open source in silico tools like MZmine 2⁴⁷ and GNPS⁴⁸ could be also used to preprocess the raw MS² data and visualize the structural fingerprints within a molecular network.

1.1.6 OSMAC (One Strain MANY Compounds)

Natural product isolation starts with screening, followed by extraction, fractionation and finally leading to the isolation of pure compounds. Each of the mentioned steps are very complex enough. NP research has been flourished with two dominant approaches in the past. The first one is a taxonomical approach for NP isolation where, specific genera were heavily explored once their potentiality was verified. Then after, bioassay guided approaches in parallel led us to golden period antibiotics discovery. Overuse of such assays resulted in a rediscovery problem which in return hampered the discovery of new molecules. Therefore, use of novel culture techniques were inevitable.

Secondary metabolites are the end product of gene cluster present within the genome. But their expression under normal laboratory condition and fixed media supplement would led scientists to known scaffold of NPs. In order to diversify the NP library, OSMAC⁴⁹ approach was implemented in early 2000's. This technique principally varies cultivation conditions and has been successful for finding new compounds within the same organism. When the extracts were analysed with hyphenated techniques like, GC-MS, LC-PDA, LC-MS, LC-FTIR, LC-NMR and LC-NMR-MS, prioritization and dereplication of NP becomes more effective. Nocardamicin glucuronide⁵⁰, a new siderophore was isolated from *Streptomyces* sp. 80H647 by using this technique. Similarly, set of new butyrolactone was isolated from fungus *Bulgaria inquinans*⁵¹. More recently, OSMAC approach combined with varied temperatures resulted in isolation of a novel compound peucemycin from *Streptomyces peucetius*⁵².

1.4 Mining Environment for chemical diversity

The biodiversity we see today is the fruit of billions of years of evolution, shaped by natural processes and, increasingly, by the influence of humans. This diversity is often understood in terms of the wide variety of plants, animals, and microorganisms. In this study I have selected strains of these biodiverse countries and among the unique ecological niches. It is natural that researcher make attempt to studies biodiversities within their homeland, however, this is always subject to facilities and infrastructure provided in these developing countries. One of several researchers is Prof. Taifu Mahmud, who has been consistently producing new chemistry from Indonesian actinomycetes⁵³⁻⁵⁵. But, when it comes to Nepal, much of its biodiversity is underexplored. Despite of high biodiversity index ratio, most of its species are never studied for potential natural product.

1.1.7 Geographical biodiversity

Flora and fauna can be indigenous to certain niches or geographical location. On a reputed English language online magazine Mongabay, Earth's most biodiverse countries are listed. This list puts Indonesia and Nepal as one of the most biodiverse countries (see **Figure 1-7**). Bioprospecting NP relies on availability of biodiversity as a source material for discovery of new leads. Furthermore, geographically diverse samples tend to also be chemically diverse, and thus, the examination of the same or similar organisms from different locations has been a productive approach to the discovery of novel chemical entities⁵⁶⁻⁶⁰. In this context, microorganisms as well belong to very diverse communities encompassing archaea, bacteria and fungi. Depending on their geographical and ecological origin, they have adapted to a variety of environmental conditions, some of them to extreme ones, such as high salinity, alkalinity, acidity and high or low temperatures⁶¹, later discussed in section 1.4.2.4. Geographical locations selected for sampling is shown in **Figure 1-8**.

EARTH'S MOST BIODIVERSE COUNTRIES										
COUNTRY	Birds	Amphib	Mammals	Reptiles	Fish	Vascular plants	BioD Index	Rank	BioD index/Land area	Rank
Brazil	17.6%	13.6%	11.8%	7.9%	13.7%	20.8%	0.85	1	0.10	
Colombia	18.3%	10.2%	8.1%	5.9%	6.2%	19.0%	0.68	2	0.57	
Indonesia	16.2%	4.6%	12.2%	7.1%	14.1%	10.9%	0.65	3	0.34	
China	12.5%	5.5%	10.0%	4.7%	10.1%	11.9%	0.55	4	0.06	
Mexico	10.9%	5.0%	9.5%	8.9%	7.9%	9.7%	0.52	5	0.26	
Peru	18.1%	7.6%	8.5%	4.7%	4.7%	6.3%	0.50	6	0.41	
Australia	7.1%	3.2%	6.4%	10.1%	14.7%	5.8%	0.47	7	0.06	
India	11.9%	5.2%	7.5%	6.7%	7.4%	6.9%	0.46	8	0.14	
Ecuador	16.0%	7.2%	6.8%	4.3%	3.3%	7.2%	0.45	9	1.59	21
Venezuela	13.7%	4.8%	6.6%	3.9%	5.2%	7.8%	0.42	10	0.45	
United States	8.5%	4.0%	8.0%	5.2%	9.3%	7.2%	0.42	11	0.04	
Bolivia	14.3%	3.2%	6.6%	3.0%	1.2%	6.4%	0.35	12	0.31	
South Africa	7.6%	1.7%	5.4%	4.4%	6.2%	8.7%	0.34	13	0.27	
DR Congo	10.9%	3.2%	7.8%	2.9%	4.5%	4.1%	0.33	14	0.14	
Malaysia	7.1%	3.5%	6.1%	4.7%	5.8%	5.7%	0.33	15	0.97	32
Viet Nam	8.3%	3.0%	5.2%	4.5%	7.3%	3.9%	0.32	16	0.97	33
Papua New Guinea	7.2%	4.9%	4.9%	2.7%	8.5%	4.3%	0.32	17	0.69	
Thailand	9.3%	1.9%	5.7%	4.2%	6.4%	4.3%	0.32	18	0.60	
Tanzania	10.6%	2.7%	6.5%	3.5%	5.3%	3.7%	0.32	19	0.34	
Argentina	10.0%	2.3%	6.8%	4.3%	3.0%	3.5%	0.30	20	0.10	
Cameroon	8.8%	2.9%	6.1%	2.8%	3.1%	3.1%	0.27	21	0.55	
Kenya	10.4%	1.5%	6.9%	2.7%	3.2%	2.4%	0.27	22	0.47	
Panama	8.8%	2.8%	4.5%	2.6%	4.2%	3.7%	0.27	23	3.45	9
Philippines	5.7%	1.5%	3.5%	2.0%	9.9%	3.3%	0.26	24	0.87	
Costa Rica	8.6%	2.7%	4.1%	2.6%	3.3%	4.5%	0.26	25	4.89	6
Myanmar	10.2%	1.2%	5.4%	3.0%	3.1%	2.6%	0.25	26	0.37	
Japan	4.4%	1.0%	2.6%	1.0%	12.1%	2.1%	0.23	27	0.62	
Angola	9.1%	1.3%	5.2%	2.4%	2.7%	1.9%	0.23	28	0.18	
Madagascar	2.5%	4.1%	4.2%	4.0%	3.5%	3.5%	0.22	29	0.36	
Mozambique	6.7%	1.2%	4.3%	2.2%	5.3%	2.1%	0.22	30	0.26	
Guatemala	7.1%	2.2%	4.0%	2.6%	2.7%	3.2%	0.22	31	1.93	18
Guyana	7.9%	1.8%	4.3%	1.8%	3.0%	2.4%	0.21	32	0.98	31
Uganda	9.9%	0.8%	5.8%	1.7%	0.8%	2.4%	0.21	33	0.87	
Guinea	6.3%	1.0%	4.1%	5.8%	3.0%	1.1%	0.21	35	0.86	
Nigeria	8.6%	1.5%	5.2%	1.9%	2.3%	1.7%	0.21	34	0.24	
Honduras	7.0%	1.6%	3.9%	2.6%	3.1%	2.1%	0.20	36	1.78	20
Nicaragua	6.8%	0.9%	3.7%	1.9%	3.2%	2.8%	0.19	37	1.46	22
Laos	6.9%	1.3%	3.9%	1.7%	1.7%	3.1%	0.19	38	0.76	
Congo	6.1%	1.0%	3.6%	3.3%	2.3%	2.2%	0.19	39	0.53	
Ethiopia	8.1%	0.9%	4.9%	2.3%	0.5%	2.4%	0.19	40	0.17	
Sudan	9.2%	0.2%	5.1%	1.8%	1.4%	1.2%	0.19	41	0.10	
French Guiana	7.1%	1.4%	3.8%	1.6%	2.9%	2.1%	0.19	42	2.26	15
Ghana	6.8%	1.1%	4.7%	1.8%	2.1%	1.4%	0.18	43	0.75	
Suriname	7.0%	1.5%	3.8%	0.6%	3.1%	1.9%	0.18	44	1.09	28
Gabon	6.1%	1.3%	3.3%	1.3%	2.3%	2.5%	0.17	45	0.64	
Ivory Coast	6.7%	1.1%	4.6%	1.5%	2.0%	1.4%	0.17	46	0.53	
Zambia	7.3%	1.2%	4.3%	1.8%	1.2%	1.8%	0.17	47	0.23	
Iran	4.7%	0.3%	3.4%	3.2%	1.9%	3.0%	0.17	48	0.10	
Nepal	8.1%	0.6%	3.3%	1.4%	0.9%	2.6%	0.17	49	1.16	27
Paraguay	6.9%	1.0%	3.0%	1.8%	0.8%	2.9%	0.16	50	0.39	

Figure 1-7 Top 50 most biodiverse countries⁶²

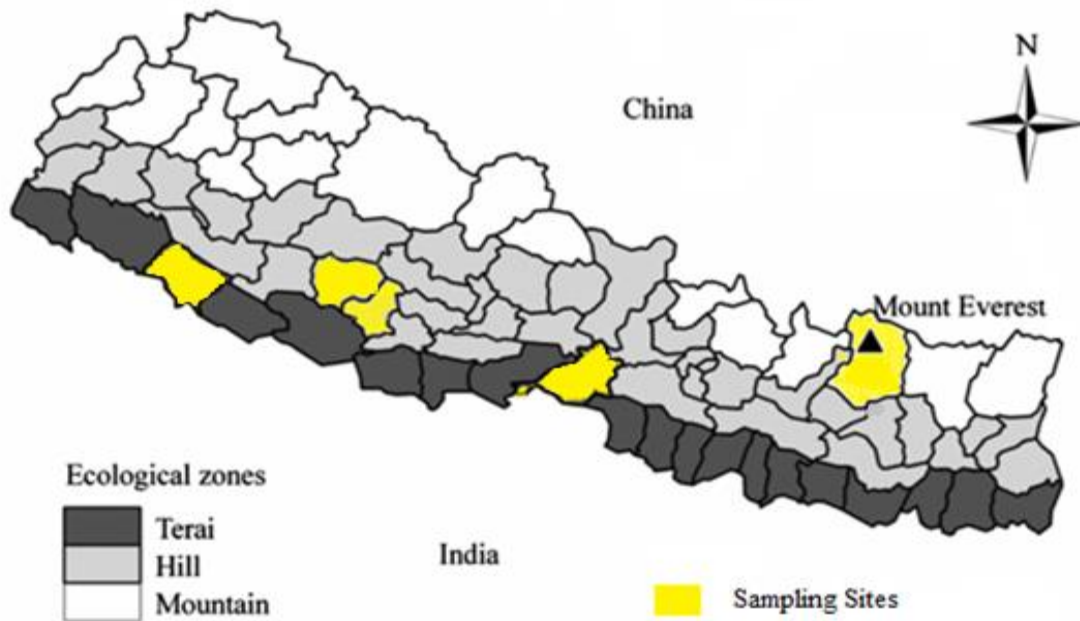


Figure 1-8 Sampling sites in Nepal and Indonesia, this study

1.1.8 Diverse Ecological niches/habitats for strain collection

1.4.1.1 Terrestrial habitat

Soil comprises most of the terrestrial habitat giving an organism a firm surface to dwell on. Soil micro-organisms have always been a great resource for discovering not only antibiotics but also a broad range of other bioactive natural products. Amongst soil bacteria, actinobacteria in general are the most prolific producers of bioactive natural products. They are ubiquitous and can be found in most environments on the Earth. There are several secondary metabolite producers which were isolated from soil.

Streptomyces avermitilis is one of the classic examples. It was isolated from soil sample close to a golf course in Kawana, on the southeast coast of Honshu, Japan. This *Streptomyces* sp. was capable of producing avermectin, later modified and developed as ivermectin, which showed antiparasitic activity⁶³. They were approved for both veterinary and human medicine for treatment of onchocerciasis and lymphatic filariasis. These success stories were well appreciated with the Nobel Prize in physiology or medicine in the year 2015⁶⁴.

1.4.1.2 Marine habitat

Marine habitats are also one of the richest sources of bioactive molecules. Marine habitats are even diverse within itself allowing us to isolate microbes from sponges, corals, animals, algae, mangrove, plants and marine sediments. Genus *Streptomyces* here also continues to be the predominant source of new chemistry, with 167 new metabolites reported during year 2018, representing >69% of the marine-sourced bacterial NPs⁶⁵. A comparative analysis by Kong and co-workers showed that marine natural products are superior to terrestrial natural products in terms of chemical novelty⁶⁶. With development of bioinformatics and genomics there has been a paradigm shift on the approach of bioprospecting marine resources. Cytarabine (Ara-C) and vidarabine (Ara-A), the first FDA-approved marine-derived drugs, were synthetic pyrimidine and purine nucleosides, respectively, inspired from naturally occurring nucleosides originally isolated from the Caribbean sponge *Tethya crypta*. Later in the year 2004, ziconotide gained FDA approval for the management of severe chronic pain.

Marine microbes were benefitted with advancement in genomics and bioinformatics. *Salinospora tropica* was the first marine actinomycetes to be sequenced in 2007⁶⁷. Results revealed a complex secondary metabolome containing high diversity of polyketide synthases (PKSs), as well as nonribosomal peptide synthetases (NRPSs) and hybrid PKS/NRPS pathways. Further investigation led to discovery of many compounds which includes, salinisporamide⁶⁸, sporolides⁶⁹, lymphostin⁷⁰ and salinilactam A. Precursor-directed biosynthesis and semisynthesis produced bromine and iodine-substituted derivatives⁷¹, while genetic engineering was used to produce fluoro salinisporamide derivatives⁷². Salinasporamide A (Marizomib) is currently in Phase III clinical trials for glioblastoma brain cancer⁷³. Nucleoside antibiotics, Streptocytosine A-D⁷⁴ were reported from *Streptomyces* strain TPU1236A isolated in Okinawa, Japan from sea water from which only streptocytosine A was shown to exhibit activity against *Mycobacterium smegmatis*.

1.4.1.3 Mangroove habitat

Mangrove swamps occupy about 181,000 km² or cover approximately 75% of the world's tropical and subtropical coastlines⁷⁵. Mangrove is a unique woody plant community of intertidal coasts in tropical and subtropical coastal regions. Global mangroves are mainly distributed in Asia (42%), Africa (20%), North and Central America (15%), Oceania (12%) and South America (11%)⁷⁶. The mangrove forests are among the world's most productive ecosystems which improves coastal water, produces commercial forest products, supports coastal fisheries, and protects coastlines. When it comes to natural products mangrove as a plant has been studied extensively for production of varieties of NP which includes naphthoquinones, glucosides, terpenoids, limonoids and many more⁷⁷. Due to its properties of high salinity, strong winds, extreme tides, high temperature, anaerobic soils, and muddiness; mangrove is inhabited with diverse array of microbes. Attractive structures such as salinosporamides⁶⁸, xiamycins⁷⁸ and novel indolocarbazoles⁷⁹ were isolated from mangrove associated bacteria. Actinobacteria were among the top isolated phyla even as mangrove symbionts. A series of nine-membered dilactone antimycins and its derivatives were isolated from mangrove actinomycetes. These molecules were shown to possess fungicidal and cytotoxic activities⁷⁶. Likewise, macrolide, divergolides A-D were isolated from an endophyte *Streptomyces* sp. HKI0576 of the mangrove tree *Bruguiera gymnorrhiza* in China.

1.4.1.4 Extreme habitats

A microorganism is classified as an extremophile if it grows optimally under conditions where so-called normal bacteria cannot live anymore. Although the first extremophilic organism, which was thermophilic (heat-loving), was isolated in the 1860's, later it was widely accepted that there were a significant number of organisms capable of surviving in extreme environments, such as high/low pH, temperature, salinity, oxygen and pressure. These organisms have been found to develop unique defences against their environment, leading to the biosynthesis of novel molecules ranging from simple osmolytes and lipids to complex secondary metabolites. Enzymes discovered such bacteria have significant value for biotechnological application. One of these marked discovery is Taq polymerase from *Thermus aquaticus*⁸⁰. However, compounds from extremophile rare also important.

Introduction

The mildly thermophilic soil actinomycete *Thermoactinomyces* strain TM-64 (opt. 45°C) has been found to produce a thiazole containing alkaloid, TM-64⁸¹. Recently, several actinobacterial strains were isolated from antarctica which show good antitumor activities⁸².

1.4.1.5 Insect symbionts and human microbiomes

Insects provide experimentally tractable and cost-effective model systems to investigate the evolutionary development and chemical basis of animal–bacterial interactions, and symbiosis in particular²⁷. Studying these interactions will shed light on equivalent processes in other animals, including humans. Furthermore, in depth analysis if such interactions could lead us to new bioactive structures. In insects, these symbioses are best exemplified in fungus-growing ant^{83,84}, solitary digger wasp⁸⁵, and southern pine beetle⁸⁶ systems, where Actinobacteria (typically *Streptomyces*) provide chemical defences, which could be used as antimicrobials.

In case of human microbiomes, functional metabolites produced during microbe-host interaction and microbe-microbe interaction are also characterized as bioactive molecules. Studying the metabolites produced by human-associated bacteria can reveal the language of bacteria-host communication. Such knowledge has scientific merit in itself but can also be used to guide new therapies that seek to modulate the human microbiota⁸⁷. The gut microbiota has a strong potential to biosynthesize large amounts of structurally distinct metabolites with a variety of biological activities that are promising drugs and drug candidates. *Streptomyces* are found to colonize human bodies, however natural product isolation from these genera has barely been reported. In a recent study, a *Streptomyces* sp. isolated from sputum of a senior male patient with a history of lung and kidney TB, recurrent respiratory infections and COPD was found to produce actinomycin X₂, fungichromin and filipin III⁸⁸. Moreover, in one study it has been suggested that these streptomycetes within gut microbiome have also been evolved as 'old friends' to suppress colon tumorigenesis by producing antiproliferatives and immunosuppressants⁸⁹. More research in future will guide us through this field. But several studies have indicated towards the abundance of actinobacteria associated with human body^{89,90} which have definitive function.

Aim of this research

Aim of this research.

Actinomycetes are still a good source of novel compounds. If they are approached in a non-traditional way with respect to isolation, culture techniques and innovative analysis of metabolites and genomes via modern computational tools, there is a lot more to discover. To justify this statement, in both projects we have applied either of these techniques or their combination to annotate actinobacterial metabolome, which ultimately led to the finding of new molecules, yet again from nature's pharmacist – Actinobacteria.

Two major aims of this thesis are stated below

1. Combined genomics and metabolomics-based approach for identification and annotation of secondary metabolites from four actinobacterial isolates from Nepal.
2. Isolation and characterization of new nucleoside antibiotics from the Indonesian mangrove strain *Streptomyces* sp. SHP-7-22

Chapter 2. Material and methods

2.1 Material

2.1.1 Media and chemicals

Table 2-1 Media used in OSMAC approach.

Media/initials	Ingredients/Liter, g=gram pH Adjusted – 7.2~7.5
GYM (G)	D-Glucose – 4g Yeast Extract – 4g Malt Extract – 10g
TSB (T)	Difco Tryptic Soy Broth
S/N (S)	Starch – 10g NaNO ₃ – 2g K ₂ HPO ₄ – 5.65 KH ₂ PO ₄ – 2.35 MgSO ₄ x7H ₂ O – 1g CaCl ₂ – 0.1g Trace Element – 1mL
MM (M)	(NH ₄) ₂ SO ₄ – 2g K ₂ HPO ₄ – 0.5g MgSO ₄ x7H ₂ O – 0.2g FeSO ₄ x7H ₂ O – 0.01g Glucose – 1.25g
ISP4 (I)	Soluble Starch – 10g K ₂ HPO ₄ – 1g MgSO ₄ – 1g NaCl – 1g (NH ₄) ₂ SO ₄ 2g CaCO ₃ – 2g
Seed Media (Sm)	Soluble Starch – 10g Glucose – 5g NZ-Case – 3g Yeast Extract – 2g K ₂ HPO ₄ – 1g MgSO ₄ – 0.5g

Table 2-2 Other Media and buffer solutions

Media	Suppliers
ISP2	Difco
ISP4	Difco
Actinomycetes Isolation Agar	Difco

NL-410 Medium composition

Ingredients	Composition	Supplier
CaCO ₃	1 g/L	TH. GEYER
Casamino acid	5 g/L	MP Biomedicals
Glucose	10 g/L	Sigma Aldrich
Glycerol	10 g/L	Sigma Aldrich
Oatmeal	5 g/L	Becton Dickinson
Soymeal	10 g/L	Hensel
Yeast Extract	5 g/L	Becton Dickinson
Add ddH ₂ O up to 1 L, Adjust pH 7.0 before autoclaving		

NL-300 medium composition

Ingredients	Composition	Supplier
Cotton seed	20 g/L	Sigma Aldrich
Mannitol	20 g/L	Sigma Aldrich
Add ddH ₂ O up to 1 L		
Adjust pH 7.5 before autoclaving		

Humic acid agar

Ingredients	Composition	Supplier
Humic acid	5 g/L	Carl Roth GmbH + Co. KG
Na ₂ HPO ₄ ·12H ₂ O	0.5 g/L	Merck
KCl	1.7 g/L	Sigma Aldrich
MgSO ₄ ·7H ₂ O	0.05 g/L	Sigma Aldrich
Yeast extract	0.2 g/L	Difco
FeSO ₄ ·7H ₂ O	0.01 g/L	Sigma Aldrich
Agar	18 g/L	Sigma Aldrich
Add ddH ₂ O up to 1 L, Adjust pH 7.0 before autoclaving		

Phosphate buffer solution

Ingredients	Composition	Supplier
Na ₂ HPO ₄	0.14598 g/L	Sigma
NaH ₂ PO ₄ ·H ₂ O	0.01856 g/L	Sigma
Adjust pH 7.35		

Table 2-3 Chemicals used in precursor-directed biosynthesis (PDB)

Nr.	Chemicals	0.5M solution	Supplier
1a	Benzoic acid	Ethanol	Acros Organics
2a	2-mercaptobenzoic acid (thiosalicylic acid)	DMSO	Sigma Aldrich
2b	2-hydroxybenzoic acid (salicylic acid)	Ethanol/DMSO/water (1:1:1)	Sigma Aldrich
2c	2-fluorobenzoic acid	DMSO	Sigma Aldrich
2d	2-chlorobenzoic acid	Ethanol	Acros Organics
2e	2-bromobenzoic acid	Ethanol	Acros Organics
2f	2-aminobenzoic acid (anthranilic acid)	DMSO	Sigma Aldrich
3a	3-mercaptobenzoic acid	DMSO	Sigma Aldrich
3b	3-hydroxybenzoic acid	DMSO/water (1:1)	Acros Organics
3c	3-fluorobenzoic acid	Ethanol/DMSO/water (1:1:1)	Sigma Aldrich
3e	3-chlorobenzoic acid	DMSO	Sigma Aldrich
3d	3-bromobenzoic acid	DMSO/ethanol (1:1)	Sigma Aldrich
3f	3-aminobenzoic acid	DMSO/water (1:1)	Sigma Aldrich
3g	3-nitrobenzoic acid	Ethanol	Alfa Aesar
3h	3-methylbenzoic acid	Ethanol	Merck
3i	3-acetoxybenzoic acid	Ethanol	Acros Organics
4a	4-mercaptobenzoic acid	DMSO	Sigma Aldrich
4b	4-hydroxybenzoic acid	DMSO/water (1:1)	Acros Organics
4c	4-fluorobenzoic acid	Ethanol/DMSO/water (1:1:1)	Alfa Aesar
4d	4-chlorobenzoic acid	DMSO	Sigma Aldrich

Material and methods

4e	4-bromobenzoic acid	DMSO	Sigma Aldrich
4f	4-aminobenzoic acid	DMSO	Sigma Aldrich
4g	4-nitrobenzoic acid	DMSO/ethanol (1:1)	Alfa Aesar
4h	4-(methylthio)-benzoic acid	DMSO	Sigma Aldrich
5a	2,3-dihydroxy benzoic acid	DMSO/water (1:1)	Sigma Aldrich
5b	3,4-dihydroxy benzoic acid	DMSO	Sigma Aldrich
5c	3,5-difluorobenzoic acid	DMSO	Sigma Aldrich
5d	2,3-difluorobenzoic acid	DMSO	Sigma Aldrich
5e	3,4-difluorobenzoic acid	DMSO	Sigma Aldrich
5f	2-hydroxy-4-methylbenzoic acid	DMSO	Sigma Aldrich
5g	4-hydroxy-3-methoxybenzoic acid	DMSO	Sigma Aldrich
5h	4-amino-3-methylbenzoic acid	DMSO	Sigma Aldrich
6a	2,3,4-trihydroxy benzoic acid	DMSO	Sigma Aldrich
6b	3,4,5-trihydroxy benzoic acid	DMSO	Merck
6c	2,3,4-trifluorobenzoic acid	DMSO	Sigma Aldrich
6d	3,4,5-trifluorobenzoic acid	DMSO	Sigma Aldrich
6e	4-hydroxy-3,5-dimethylbenzoic acid	DMSO	Sigma Aldrich
6f	3,5-dibromo-4-hydroxybenzoic acid	DMSO	Sigma Aldrich
7a	4-aminocyclohexane carboxylic acid	DMSO	Sigma Aldrich
7b	Phenylpropionic acid	DMSO/water (1:1)	Sigma Aldrich
7c	Octanoic acid	Ethanol	Sigma Aldrich
7d	Nonanoic acid	Ethanol	Sigma Aldrich
7e	Decanoic acid	Ethanol	Sigma Aldrich

Table 2-4 Chemicals

Chemical	Supplier
<i>d</i> ₄ -methanol for NMR	Sigma Aldrich
<i>d</i> ₆ -DMSO for NMR	Sigma Aldrich
DMSO	Carl Roth GmbH + Co. KG
Ethanol	TH. GEYER
Ethyl acetate	Chemical dispensary (Bulk material)
Methanol for HPLC	Merck
Methanol for LC-MS	TH. GEYER
Methanol for SPE	Chemical dispensary (Bulk material)
Trifluoroacetic acid	Sigma Aldrich
Acetonitrile for LC-MS	TH. GEYER
2-Methylbutyric acid	Sigma Aldrich
4-methylhexanoic acid	Sigma Aldrich
(S)-2-methylbutyric acid	Sigma Aldrich
rac-2-methylbutyric acid	Activate Scientific
(R)-2-methylbutyric acid	Chemspace
(R)-4-methylhexanoic acid	Chemspace
(S)-4-methylhexanoic acid	Chemspace
Agar	Sigma Aldrich
Humic acid	Carl Roth GmbH + Co. KG
Sodium chloride	Thermo

2.1.2 Instruments and Auxiliary materials**Table 2-5** Employed instruments.

General instruments	Manufacturer
Autoclave (Model-VX-150)	Systemec
Centrifuge (Model-5424 R)	Eppendorf
Centrifuge (Heraeus Multifuge 4KR)	ThermoScientific
Clean bench (safe 2020)	ThermoScientific
Fourier transform infrared spectrometer (FT/IR-4200)	Jasco
High Performance Liquid Chromatography (HPLC)	Waters
Fraction collector	Waters and Advantec CHF122SC
Incubation shaker Multitron Pro	Infors
Liquid chromatography mass spectrometry (LC-MS)	Agilent, AB Sciex
Laboratory freeze dryer (Alpha 3-4 LSC basic)	Martin Christ
Nuclear magnetic resonance spectroscopy (NMR)	Bruker
Oven	Binder
pH meter (FiveEasy)	Mettler Toledo
Pipettes (Discovery Comfort: 1000 µL; 100 µL; 10 µL; 2 µL)	HTL Lab Solutions
Refrigerator	Liebherr
Rotary evaporator system	Heidolph, Huber
Scale (B22002)	Oxford
Special accuracy weighing scale (BP210D)	Sartorius
Sonicator	Bandelin
UV/VIS Spectrometer (Lambda 25)	PerkinElmer
Water purification system (Elga™ Purelab™ Flex)	Elga
Mini-Bioreactor	Enzyscreen, NL

Table 2-6 Auxiliary materials

Materials	Manufacturer
Eppendorf tubes (20 µL / 200 µL / 1.5 mL / 2.0 mL)	Sarstedt
Falcon tubes (15 mL / 50 mL)	Sarstedt
Filter (single use, non-pyrogenic)	Sartorius stedim
SPE cartridge	Phenomenex
Syringe Injekt™ (10 mL / 20 mL)	ThermoScientific
Polygoprep 60-50 C18 stationary phase	Macherey-Nagel
Sephadex™ LH-20	GE Healthcare

2.2 Bacterial strains

Approximately 400 actinomycetes looking colonies were isolated at Research Institute for Bioscience and Biotechnology (RIBB), Nepal, of which 22 were prioritized. Here, we mention only those characterized strains which were further considered for chemical analysis.

2.2.1 Bacterial strains from Nepal

Strain Name	Source
<i>Streptomyces</i> sp. WL006	Isolated from Wetland
<i>Streptomyces</i> sp. NB004	Isolated from river island – Nagarban
<i>Nonomuraea</i> sp. C10	Isolated from mud dauber soil nest
<i>Pseudonocardia</i> sp. C8	Isolated from mud dauber soil nest

2.2.2 Bacterial strain from Indonesia

Strain Name	Source
<i>Streptomyces</i> sp. SHP22-7	Kindly provided by Prof. Yvonne Mast from Department of Microbiology / Biotechnology, University of Tübingen, Germany

2.3 Methods

2.3.1 Soil sampling and actinomycetes isolation and storage

Soil samples were collected from six different districts of Nepal. Soil samples were collected at a depth of 20 cm after removal of ~3 cm top into a sterile plastic box, whereas liquid samples were directly collected into sterile falcon tubes. Dry soil samples (this includes also mud dauber nest soil) were sieved through a 2 mm mesh sieve to exclude pebbles and debris. One gram of each soil sample was suspended in 10 mL of normal saline and 6-fold diluted after drying and heating (5-7 days). 1 mL was taken from each and plated. Alternatively, dry soil samples were sprinkled on agar plates by the method described by Hayakawa and co workers⁹¹ (see **Figure 2-1**). Liquid samples were randomly diluted and plated directly. After seven days several colonies were manually picked onto fresh agar plates (humic acid agar⁹¹ and actinomycetes isolation agar). Several rounds of subcultures were performed under sterile conditions to obtain

Material and methods

pure colonies. Axenic colonies were finally cryopreserved with standard protocol of streptomycetes maintenance⁹².

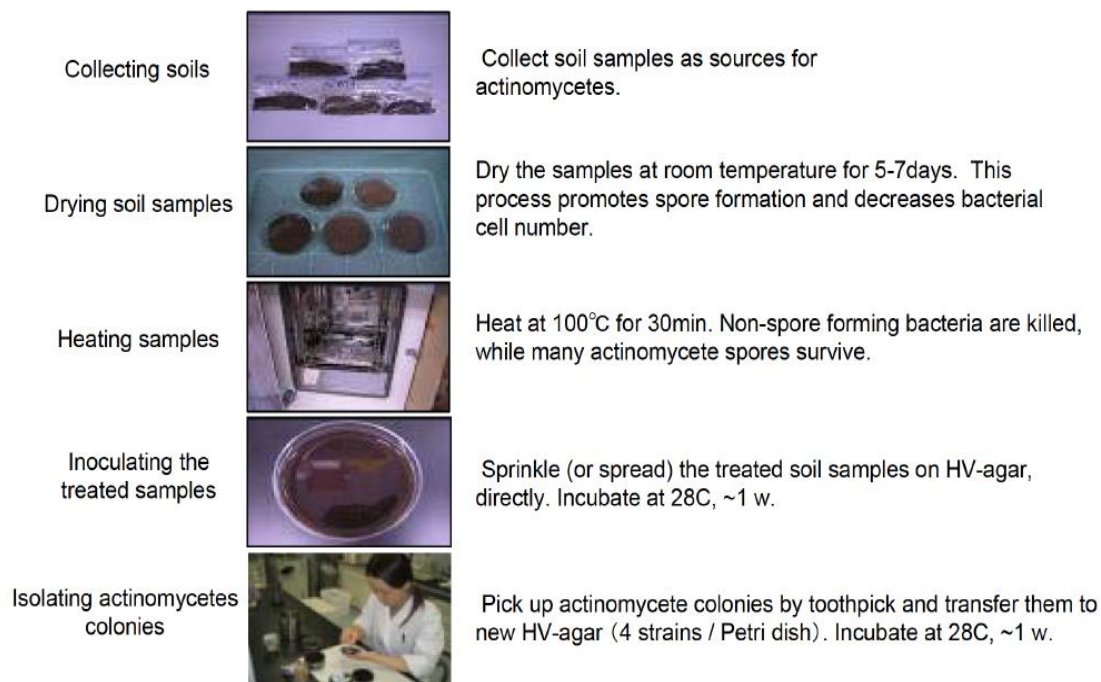


Figure 2-1 Selective isolation of actinomycetes from soil⁹¹

2.3.2 Cultivation of bacterial strains

2.3.2.1 OSMAC approach for Nepalese actinomycetes

For initial screening of bacteria, media optimization was carried out on 3 minimal media (G, T, S) and 3 rich media (M, I, Sm) see Table 2-1. Small-scale cultivation was carried out in 24 deep-well microtiter plates (MTPs) (11 mL/well) with sandwich covers^{93, 94} (see **Figure 2-2**). GYM medium was used for preparing the inoculum of which 100 µl was inoculated to the main cultures. In order to avoid cross contamination while shaking, minimum volume of 2.5 ml/well were used for cultivation. Subsequently, cultures were rotated at 300 rpm at 30°C for 9 days in an orbital shaker. Upscaling of the culture for *Streptomyces* sp. WL006 was done also in GYM medium and for *Streptomyces* sp. NB004, ISP4 medium was used for large scale cultivation.

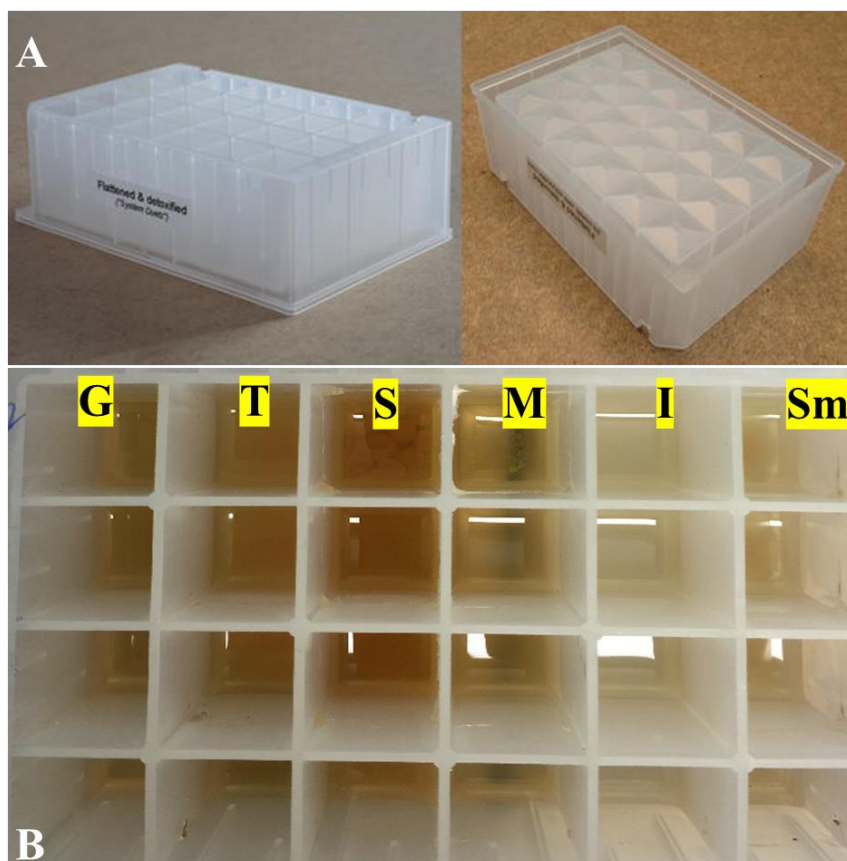


Figure 2-2 A) EnzyScreen mini-bioreactor used for small scale cultivation B) G,T,S,M,I,Sm as medium initials representing six different media

Cultivation on MTPs was ceased on the 10th day by adding an equal volume (2.5 mL) of methanol on the wells and sonicated for 15 min. Culture broths were then centrifuged for 10 min to remove cell pellets and the resultant supernatant was evaporated to dryness using rotary evaporator. The crude extract supernatant was then resuspended in 1 mL of water/methanol mixture (1:1). This was further used for downstream chemical analysis and antimicrobial assays. For *Streptomyces* sp. WL006, production batch (4L) was extracted with ethyl acetate (2 times). The crude was then subjected to further fractionation to generate 6 fractions (A-F) with decreasing polarity on C18 reverse phase open column. These fractions were also analysed via LC-MS after dissolving in methanol:water (50:50) mixture.

2.3.2.2 Cultivation and extraction of Non-*Streptomyces* Nepalese strains

Extraction of wasp nest was carried out at the very beginning. First of all, obtained wasp nest (soil) was opened up. Dead insect debris were cleared off and soil was further crushed to fine powder. Furthermore, we divided all mud nests in two parts. For the first part we crushed, dissolved the nest in water and applied heat (80°C) while extraction to obtain 3 extracts based

on polarity. Powder soil was first dissolved in water and extracted with n-butanol to obtain polar compounds. Remaining aqueous phase was re-extracted with ethyl acetate as well. This resulted in three extracts, which were subjected to LC-MS and for further bioassay. Secondly, actinomycetes strains were obtained from the nest soil as well using the method in section 2.3.1. Non-streptomyces strains, *Nonomuraea* sp. C10 and *Pseudonocardia* sp. C8 inoculum were prepared in Seed Medium (Sm). After 4 days, 1/10 of inoculum was subjected to freshly prepared main culture medium (also Sm medium) and was further cultivated until 10 days. Mycelia were centrifuged at 4400 rpm to obtain cell-free supernatant. The resultant supernatant was extracted with n-butanol while mycelia was extracted with methanol for further chemical analysis.

2.3.2.3 Cultivation of *Streptomyces* SHP-22-7

2.3.2.3.1 Preculture preparation

Single colony of the bacteria, which was obtained from the laboratory of Prof. Yvonne Mast. It was picked and then inoculated in 100 mL NL-410 media. The inoculum in baffled Erlenmeyer flasks (50mL) was shaken at 28°C with 120 rpm for 72 hours.

2.3.2.3.2 Large scale cultivation for further analysis

10 mL (10% of scale-up cultivation media) preculture of *Streptomyces* sp. strain SHP 22-7 was inoculated in 300 mL baffled flasks with metal spring including 100 mL NL-300 media. Subsequently, the culture was incubated at 28°C at 120 rpm for 96 hours.

2.3.3 Measurement of growth curve and secondary metabolite production curve

Preculture was prepared as described in section 2.3.2.3.1 and 5 mL homogeneous preculture was put into each baffled flask with metal spring including 50 mL NL-300 medium. According to the time schedule planned in advance (Table 2-7), 10 mL culture from corresponding baffled flask was taken out in triplicate, altogether 30 mL. After centrifugation at 4400 rpm for 10 min the supernatant was discarded. The wet weight was measured directly. Subsequently, the wet cells were dried for 12 hours in hot air oven at 65°C for the dry weight. In parallel, media controls were also measured similarly.

Material and methods

For obtaining secondary metabolite production curve, 10 mL aliquot was taken. This was then extracted with equal volume ethyl acetate. After solvent evaporation, resultant crude was further dissolved in 1 mL water:methanol (1:1) for LR-LC/MS measurements. To obtain production graph, semi-quantitative approach was employed. This method involves the measurement of peak area of target compound at specific retention time in an LC-MS software. Target compound was confirmed via its $[M+H]^+$ values.

Table 2-7 The time schedule for growth and secondary metabolite production curve

Number	Total time (h)	Interval (h)
	0	
1	8	8
2	11	3
3	14	3
4	17	3
5	20	3
6	23	3
7	26	3
8	29	3
9	32	3
10	35	3
11	38	3
12	41	3
13	44	3
14	47	3
15	55	8
16	63	8
17	71	8
18	79	8
19	87	8
20	95	8
21	103	8
22	111	8
23	119	8
24	127	8
25	135	8
26	143	8
27	151	8
28	159	8
29	167	8
30	175	8

2.3.4 Fractionation of crude extracts

Fractionation was one of the vital process in the workflow of compound isolation. Due to the use of complex medium, several media components, and primary metabolites have overcrowded the crude extracts. To decomplexify the extracts a fractionation step was crucial. Sephadex LH-20 and Polygoprep C18 stationary phase was used to achieve sub-fractions. Depending on the complexity and amount of crude obtained, further fractionations were carried out either with semi preparative HPLC or in a SPE cartridge (Phenomenex® Strata™-XL 100 µm Polymeric Reversed Phase, 5 g/20 mL, Giga Tubes) to facilitate purification of compound. The stepwise concentration gradient was shown as follows (**Table 2-8**).

Table 2-8 The concentration gradient used in SPE for *Streptomyces* SHP22-7

Fraction	Mobile phase	Volume
Wash	H ₂ O	50 mL
Equilibration	MeOH	50 mL
A	MeOH: H ₂ O = 3:7	50 mL
B	MeOH: H ₂ O = 5:5	30 mL
C	MeOH: H ₂ O = 7:3	30 mL
D	MeOH: H ₂ O = 9:1	30 mL
E	MeOH	100 mL

Table 2-9 Open column C18 fractionation scheme for *Nonomuraea* sp. C10

Fraction	Ratio (H ₂ O:MeOH)
A	100:0
B	80:20
C	60:40
D	40:60
E	20:80
F	0:100

2.3.5 Bioassays

2.3.5.1 Antibacterial assays

The minimal inhibitory concentration (MIC) was determined in a cation-adjusted Mueller-Hinton medium that contains casein, beef extract and starch by using a twofold serial

Material and methods

dilution method according to the standards and guidelines of the Clinical and Laboratory Standards Institute (CLSI)⁹⁵. For MIC testing of *Mycobacterium smegmatis*, Middlebrook 7H9 broth instead of cation-adjusted Müller Hinton medium was used. In brief, a twofold serial dilution of the test compound was prepared in microtiter plates and seeded using a final inoculum of bacteria of 5×10^5 colony-forming units per mL. After overnight incubation at 37°C in ambient air, the MIC was determined as the lowest compound concentration preventing visible bacterial growth. The strain panel included representative species of nosocomial pathogens, which are known as “ESKAPE” bacteria. Specifically, the following strains were used: *Enterococcus faecium* BM 4147–1, *Staphylococcus aureus* ATCC 29213, *Staphylococcus aureus* NCTC8325, *Klebsiella pneumoniae* ATCC 12657, *Acinetobacter baumannii* 09987, *Pseudomonas aeruginosa* ATCC 27853, *Enterobacter aerogenes* ATCC 13048. *Bacillus subtilis* 168, *Escherichia coli* ATCC 25922 and *Mycobacterium smegmatis* mc² 155 were used as further reference strains. The ATCC strains were provided by the American Type Culture Collection. *A. baumannii* 09987 was obtained from the University of Bonn, Germany.

2.3.5.2 Cytotoxicity assay

The cytotoxicity test against the HeLa human cervical carcinoma cell line was performed in RPMI cell culture medium supplemented with 10% fetal bovine serum using the 7-hydroxy-3H-phenoxazin-3-one-10-oxide (resazurin) assay. A twofold serial dilution of the test compounds was prepared in duplicates in a microtiter plate and seeded with trypsinized HeLa cells to a final cell concentration of 1×10^4 cells per well. After 24 h incubation at 37°C, 5 % CO₂, 95% relative humidity, resazurin was added at a final concentration of 200 µM, and cells were again incubated overnight. Cell viability was assessed by determining the reduction of resazurin to the fluorescent resorufin. Fluorescence was measured in a TECAN M200 reader at an excitation wavelength of 560 nm and an emission wavelength of 600 nm in relation to the untreated control.

2.3.5.3 Screening bioassays for extracts obtained from Nepalese actinomycetes

Disk diffusion assays were performed to evaluate the antibacterial activity on a qualitative level. *Escherichia coli*, *Bacillus subtilis* and *Mycobacterium phlei* were used as the test strains. The crude obtained from six different media for each 22 colonies (i.e 132 samples) were dissolved in methanol:water (1:1) and loaded into a dry sterile bioassay discs (diameter=6 mm). Similarly, same discs were also soaked in solvent methanol which were used as the negative control.

Altogether. All the disks were completely dried before placing it on agar plates. The test microorganisms, *E. coli* and *M. phlei* were grown in TSB agar at 37 °C while *Bacillus subtilis* was grown in 30°C for 24 hrs. A clear zone of inhibition (ZoI) was measured with simple ruler in between 24-36 hours of cultivation. Isolates with maximum ZoI were prioritized for further chemical analysis.

2.3.6 DNA Extraction, 16S rRNA Amplification and 16S rRNA Sequencing

Each bacterial culture broth, 5 mL was centrifuged (5,000 g/3 min), washed with STE buffer, and treated with 1.4 M NaCl, 20 mM EDTA, 100 mM Tris-HCl (pH 8.0), 5 mg/mL proteinase K, 2 mg/mL lysozyme, and 1 mg/mL RNase A. Samples were then frozen at -80°C for 3–5 min and thawed in a dry bath for 3 min at 65°C (this process was repeated twice). Next, an equal volume of a solution containing phenol/chloroform/isoamyl alcohol 25:24:1 was added to each tube, homogenized by inversion, and centrifuged (10 min, 14,000 g, 4°C). The aqueous phase was recovered, mixed with 400 mL chloroform, and centrifuged (10 min, 14,000 g, 4°C). Genomic DNA was precipitated with ammonium acetate and cold isopropyl alcohol and frozen at -20°C for 10 min and subsequently centrifuged. The supernatant was discarded, and the pellet formed was washed with cold ethanol (70%), centrifuged, and solubilized in TE buffer (10 mM Tris-HCl and 1 mM EDTA). To assess the integrity and purity of the DNA, an agarose gel (1%) was stained with peqGREEN™ and visualized in a transilluminator. The 16S rRNA gene for Nepalese samples was amplified by polymerase chain reaction (PCR) using primers, 27F (5-AGAGTTTGATCCTGGCTCAG-3) and 1492R (5-TACGGCTACCTTGTTACGACTT-3) ⁹⁶. The PCR mix contained PCR buffer, DNTPs, MgCl₂, and Taq polymerase. The protocol conditions included an initial denaturation period at 94°C, followed by 35 cycles of denaturation at 94°C for 60 s, annealing phase under 63°C for 60s, and extension at 72°C for 60 s. The cycling phase, a final extension at 72°C for 10 min, was subsequently performed. The PCR products were visualized in a 1% agarose gel electrophoresis and stained with peqGREEN™. The PCR product was purified by PCR Purification Kit (250) QIAGEN™. The sequencing of the 16S rRNA gene was performed by Eurofins Genomics (Ebersberg, Germany) with the ABI PRISM BigDye™ Terminator cycle sequencing kit (Applied Biosystems, United States), and the same primers were used for amplification, following the protocols provided by the manufacturer.

2.3.7 Genome sequencing

Nonomuraea sp. C10 and *Pseudonocardia* sp. C8 was grown in seed medium (see **Table 2-1**) for 8 days at 30°C on a rotary shaker (160 rpm). For genomic DNA isolation, a Quick-DNA fungal/bacterial DNA miniprep kit (Zymo Research, Irvine, CA, USA) was used according to the manufacturer's protocol except that the vortex-mixing step was reduced from 15 min to 5 min and was conducted at maximum speed. The DNA was sheared using a Covaris g-TUBE, and the genomic library was prepared according to the standard PacBio protocol, followed by size selection with the BluePippin size selection system (Sage Science, Inc.). The 10-kb library was sequenced on a PacBio Sequel instrument using one single-molecule real-time (SMRT) cell.

2.3.8 Chemical analysis

2.3.8.1 HPLC analysis

High performance liquid chromatography (HPLC) is a powerful chromatographic tool because of its high efficacy and accuracy in the pharmaceutical field, especially for the isolation and separation of natural products (NPs), which contains complex matrices⁹⁷.

Once fractions were obtained, isolation and purification of target products were carried out with the HPLC, which consists of a Waters binary pump 1525, a Waters 717 plus autosampler, a Waters photodiode array detector 2996 and an Advantec CHF122SC fraction collector. The full process was controlled by Millennium software.

Method optimization was key for isolation, therefore for each compounds a set of different HPLC method were needed. Mobile phases were also chosen as per need. For plicacetin and its derivatives a semi preparative column was employed to profile the fraction containing target compounds eluting with methanol (A) and 1M phosphate buffer solution (B). The concentration gradient is shown in **Table 2-10**. Further isolation was carried out on analytical columns to achieve better resolution and subsequently purity. Isolation of streptcytosine B-C, streptcytosine N and new streptcytosine was done using isocratic flow (0.6mL/min = A, 0.4mL/min =B) for obtained fraction using Kinetex EVO C18, and Kinetex XB-C18.

Table 2-10 HPLC method for plicacetin and its derivatives isolation

Time	Flow (mL/min)	A: MeOH (%)	B: 1M phosphate buffer solution (%)
0	3	60	40
26	3	70	30

Material and methods

29	3	95	5
36	3	95	5
37	3	60	40
42	3	60	40

The pure substance was collected in the clean glass vials over the time. Whenever necessary, a Luna C5 column (5 μ m, 250 x 4.6 mm) was applied for removing the lipid contaminant from the samples. Suitable HPLC columns were employed for resolving the compounds and for a desalting process. A list of columns used in this study is displayed in **Table 2-12**

For isolation of brartemicin and its derivatives from *Nonomuraea* sp. C10 a different HPLC method was employed as shown in **Table 2-11**

Table 2-11 HPLC method for Brartemicin and derivatives isolation

Time	Flow (mL/min)	A: ACN (%)	B: 0.1% TFA in Water
0	3	5	95
5	3	10	90
10	3	20	80
20	3	30	70
38	3	50	50
45	3	80	20
50	3	5	95
55	3	5	95

Table 2-12 HPLC columns used in this study

Column Name	Manufacturer
Synergi Hydro-RP, 80 Å, 250x10 mm, 4 μ m	Phenomenex
Synergi Polar-RP, 80 Å, 250x10 mm, 4 μ m	Phenomenex
Luna [®] Omega Polar C18, 100 Å, 250x10 mm, 5 μ m	Phenomenex
Kinetex [®] XB-C18, 100 Å, 250x4.6 mm, 5 μ m	Phenomenex
Luna [®] C18, 100 Å, 250x4.6mm, 5 μ m	Phenomenex

Luna [®] C5, 100 Å, 250x4.6mm, 5 µm	Phenomenex
Kinetex [®] PFP, 250x4.6mm, 5 µm	Phenomenex
Kinetex [®] EVO C18, 250x4.6mm, 5 µm	Phenomenex

2.3.8.2 Liquid Chromatography Mass Spectrometry (LC-MS)

Low resolution (LR) LC-MS analysis was performed applying an 1100 Series HPLC system (Agilent Technologies, Waldbronn, Germany) connected with an ABSCIEX 3200 Q TRAP LC-MS/MS mass spectrometer (AB Sciex, Germany GmbH, Darmstadt, Germany). The HPLC part comprises of G1322A degasser, G1312A binary pump, G1329A autosampler and G1315A diode array detector. LRLC-MS parameters are given in **Table 2-13**. A Phenomenex Luna C18 (2) column (5 µm, 250 x 2.0 mm) was used with a flow rate of 0.2 mL/min. Organic mobile phase (A-pump) was either methanol or acetonitrile and aqueous phase (B-pump) was 0.1% TFA in ddH₂O was applied during screening measurements in Q1/Q3 positive and negative scan mode with mass range between m/z at 100 to 1600. A linear gradient of 10% to 90% of A was set within 45 min. Once acquired data were analysed with Analyst[®] 1.6 software.

Table 2-13 Low resolution LC-MS parameters

Parameter	Value
Declustering potential	70
Entrance potential	10
Curtain gas	10
Collision gas	2
IonSpray Voltage	4500
Temperature	450
Ion Source Gas 1	50
Ion Source Gas 2	50

HR-LC-ESI-MS/MS was done for the crude extracts and pure samples. These were acquired using a Bruker TOF-MS MaXis Impact ESI-HR-MS. An LC-method was applied as follows: with 0.1% FA in H₂O as solvent A and Acetonitrile as solvent B, a linear gradient of 10% to 100% B for 35 min, 100% B for additional 10 min, using a flow rate of 0.5 mL/min; 3 µL injection volume und UV detector (UV/VIS) wavelength monitoring at 190-800 nm. The separation was carried out on a Luna Omega Polar C18 column (3 µm, 250 x 4.6 mm), and the range of MS acquisition was m/z 100-1800.

A capillary voltage of 4.5 kV, nebulizer gas pressure (nitrogen) of 2 bar, ion source temperature of 200°C, dry gas flow of 8 L/min source temperature, and scan rates of 1 Hz for MS¹ and 1 Hz for MS² were used. For acquiring MS² fragmentation, the 10 most intense ions per MS¹ were selected for subsequent CID with stepped CID energy applied. The employed parameters for tandem MS were applied as previously detailed⁹⁸. The samples were measured in ESI (+) and ESI (-) modes. Sodium formate was used as an internal calibrant. Data analysis was performed using Bruker Daltonics Data Analysis 4.2.

2.3.8.3 Preprocessing MS data and generation of Molecular Network (MN)

A bucket table was generated for crude extract and media by using positive measurements on MetaboScape 3.0. Settings used: intensity threshold: 10⁻⁴, minimum peak length: 7, minimum peak length recursive: 3, primary ion: [M+H] and *m/z* range 100 – 1800Da.

Then the resultant feature detection bucket (MS/MS) was exported as mascot generic format (.mgf) file, which was later used as an input to the GNPS webservice (<http://gnps.ucsd.edu>). Cosine adjusted to 0.7, ion tolerance 0.05 Da, parent mass tolerance 0.5Da were adjusted. Molecular network file was visualized on Cytoscape 3.7.1⁹⁹. Each node is labelled with the precursor mass and edge thickness is referring to cosine similarity score.

2.3.8.4 Compound Annotation using SIRIUS

Compound annotation was done by using Sirius 4.0.1 along with user interface CSI:FingerID⁴⁶. This interface allows us to compare our experimental data to that of *in silico* generated fragments and finds the similarity score and molecular formula. Molecular formulae were compared to the ones which was generated using Bruker Data Analysis 4.0. On platform CSIFinger-ID, specific collision energy was given as an input along with MS² raw data. This *in silico* generated fragment was matched to the experimental fragment. The compound in the database having the highest match was considered as closest possible structure. Multiple hits were generated. These hits were ranked in the order of most similar structure on the provided databases. The final formula was subjected to Sci-finder search for further confirmation. Derivatives were annotated by looking at the possible losses (in Dalton) to the main compound.

2.3.8.5 Determination of compound novelty

After isolation of target compound via HPLC, NMR was performed to confirm the structure. These data were further supported by HRMS which provided us with molecular formula and rdb

values. UV, IR, OD gave additional hit on structural features and other physical parameters. As this work was specific to actinobacterial metabolites, databases like Antibase, Streptome, NPAtlas were prioritized as a search engine in our very first step. In this workflow, we considered an isolated compound to be novel once there were no hits in SciFinder, a paid web-based database from CAS with is the biggest chemical structure bank containing over 179 million defined structures.

2.3.8.6 Determination of chirality of compounds

For the determination of the chirality, compounds were dissolved in 6N HCl and heated to 110°C overnight. With this hydrolysis step, the chiral carboxylic acids 2-methylbutyric acid and 4-methylhexanoic acid, respectively, were cleaved. After hydrolysis, a liquid-liquid extraction step was performed with hexane. The organic layer was separated and proceeded to derivatization. As derivatization agent, 1-naphthylamine was used in combination with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) dissolved in isopropanol. Derivatization was performed while shaking at 25°C for 12 hours. After drying of the product, it was resolved in the respective mobile phase.

HPLC-UV analysis was performed on an Agilent 1260 LC-system from Agilent Technologies (Waldbronn, Germany) equipped with a quaternary pump, a degasser, an autosampler, and a UV-DAD detector.

HPLC-MS analysis was performed on an Agilent 1290 LC-system with an API 4000 QQQ MS system from AB Sciex (Toronto, Ontario, Canada) and a HTC PAL autosampler from CTC Analytics (Zwingen, Switzerland). MS² product ion scan was performed equally for both compounds with the following settings: Declustering Potential: -70.0 V, Entrance Potential: -10.0 V, Collision Energy: -30.0 V, Collision Cell Exit Potential: -15.0 V, Collision gas: 6 psi, Curtain gas: 30 psi, Ion source gas 1: 50 psi, Ion source gas 2: 40 psi, IonSpray Voltage: -4500.0 V, Temp: 450 °C

Both chiral columns (Chiralpak IB-U and IH-U, 100 x 3 mm, 1.6 µm fully porous particles) were from Daicel (Tokyo, Japan).

HPLC-UV measurements for 2-methylbutyric acid on Chiralpak IB-U were performed in normal phase, using hexane/isopropanol (95/5; v/v) as mobile phase. Flow rate was set to 1 mL/min, column temperature was held constant at 25°C. Detection wavelength was 215 nm. MS measurements for this compound were conducted on the same column but in reversed phase gradient elution mode with water + 0.1% acetic acid (A) and acetonitrile + 0.1% acetic acid (B). Column temperature was set to 40°C, flow rate was 0.6 mL/min.

Material and methods

HPLC-UV measurements for 4-methylhexanoic acid on Chiralpak IH-U were performed in reversed phase, using a gradient between water + 0.1% acetic acid (A) and acetonitrile + 0.1% acetic acid (B). Flow rate was set to 0.2 mL/min, column temperature was held constant at 25°C. Detection wavelength was 215 nm. MS measurements for this compound were conducted on the same column under equal conditions.

2.3.8.7 UV spectroscopy

UV spectra were recorded on a Perkin-Elmer Lambda 25 UV/VIS spectrometer using a 1.0 cm quartz cell. Compounds were diluted in methanol or DMF. Dilutions were measured with the UV-scan mode from 600 to 200 nm. The molar absorption coefficient ϵ was calculated according to Beer-Lambert law.

$$\epsilon = \frac{A}{c * d} [L * mol^{-1} * cm^{-1}]$$

A = absorption at peak maximum (nondimensional)

c = concentration in mol / L

d = layer thickness of solution in cm

2.3.8.8 IR spectroscopy

IR spectra were recorded using a Jasco FTIR-4000 series of Fourier transform infrared spectrometers interfaced with a MIRacle ATR device (ZnSe crystal). Analysis and reporting were performed with Spectra Manager 2.10.01 software.

2.3.8.9 Optical rotation

Optical rotations were measured on a Jasco model P-2000 Series polarimeter (100 mm, 1cm³ cell) operating at $\lambda=589$ nm corresponding to the sodium D line. The specific optical rotation $[\alpha]_D^T$ was calculated according to the equation below.

$$[\alpha]_D^T = \frac{\alpha}{c \times l} \times 10^4$$

$[\alpha]_D^T$ = specific optical rotation

T = temperature in °C

D = sodium D line ($\lambda=589$)

Material and methods

A = rotation in degrees (°)

c = concentration

l = cell length

2.3.8.10 CD spectroscopy

CD spectra were recorded using Jasco J-720 spectropolarimeter. The samples were dissolved in MeOH and measured in the wavelength range of 400 to 200nm.

Chapter 3. NPs from Nepalese actinomycetes

3.1 Streptomyces sp. WL006, Streptomyces sp. NB004

3.1.1 Sampling and pre-screening

Diverse soil habitats were chosen for sampling. (see **Table 3-1**). This was achieved within the country in five different districts (**Figure 3-1**). The aim was to cover places with variation in altitude and to select unique niches for soil isolation.

Figure 3-1 Map of Nepal showing districts of soil sample collection.



Table 3-1 Soil habitats for sample collection

	Collection habitats/sites of collection	Distinct colony/ies	Altitude (in meters)
1	Mountains (Everest trail)/6	24	3440 – 5140
2	Hilly Regions/5	39	1000 - 3000
3	Wetland/5	28	Below 400
4	Lakes/1	10	Below 400
5	Protected Areas/2	9	Below 400
6	River/1	1	Below 400
7	Other/5	9	Below 400
	TOTAL:	120	

Of 120 distinct colonies screened, 22 were prioritized based on unique morphology. Slow-growing, chalky white and coarse colonies resembling to actinomycetes were prioritized. Twenty two selected isolates were further tested via the OSMAC approach⁴⁹ as described in

NPs from Nepalese actinomycetes

section 2.3.2.1. During our chemical analysis in various media a diverse range of metabolites were produced in each medium. Characterizing all of them was out of capacity, therefore we used such chemical screening rather to prioritize the working strain.

3.1.2 Antimicrobial assay of crude extracts

To further prioritize the most promising strains in terms of secondary metabolism, crude extracts of 22 strains in 6 different media were subjected to an antimicrobial assay. Different activity profiles were obtained in various media which further supported the concept of OSMAC. Strains with maximum zone of inhibition (ZoI) was selected for 16S rRNA sequencing and for in-depth chemical analysis. **Table 3-3** shows the antimicrobial assay carried out in six different media for all 22 strains. However, in this study we prioritized two bacterial strains for further LC-MS based dereplication based on their activity (see **Table 3-2**).

Streptomyces sp. WL006

MoC	G	T	S	M	I	Sm
ZoI _(B. sub)	1.5	0	0	0	0	1.5
ZoI _(E. col)	1.5	0	0	0	0	1.5
ZoI _(M.phl)	3	1.5	3	0.5	0	3

Streptomyces sp. NB004

MoC	G	T	S	M	I	Sm
ZoI _(B. sub)	0	1	8	0	9	7.5
ZoI _(E. col)	0	0	5	0	5	4
ZoI _(M.phl)	0	1	10	0	10	10

B.sub-*Bacillus subtilis*, E.col-*Escherichia coli*, M.phl-*Mycobacterium phlei*
 ZoI=Zone of Inhibition-in mm
 MoC=Medium of cultivation

Table 3-2 Bioactivity of crude extract of 2 prioritized *Streptomyces* strains

Table 3-3 OSMAC approach and the resultant antimicrobial activity profile in six different culture media. Media*, G=GYM, T=TSB, S=S/N, M=MM, I=ISP4, Sm=Seed medi

S.N	Strain Designation	Habitat, District	Antimicrobial Activity in different media*		
			<i>E. coli</i>	<i>B. subtilis</i>	<i>M. phlei</i>
1.	WL006	Lake, Chitwan	G,S,Sm	G,S,Sm	G,T,S,Sm
2.	NB004	Soil, Chitwan	S, I, Sm	T,S, I,Sm	T, S,I,Sm
3.	BK003	Insect Nest	G,T,S,I,Sm	G,I,Sm	G,T,S,Sm
4.	WL004	Wetland, Chitwan			
5.	RP009	Hilly soil, Rolpa	Sm	G,I,Sm	G,Sm
6.	RP016	Hilly soil, Rolpa			
7.	NB001	Nagarban, Chitwan	T,S,I,Sm	T,S,Sm	
8.	NP001	Soil, Solukhumbu			
9.	WL002	Wetland, Chitwan	G,T,S,Sm	G,Sm	G,T,S,Sm
10.	NB021	Soil, Chitwan	G,T,S,Sm	G,Sm	G,T,S,I,Sm
11.	RP008	Hilly soil, Rolpa			
12.	RP015	Hilly soil, Rolpa			
13.	RP007	Hilly soil, Rolpa	G,S,Sm	Sm	G,T,S,Sm
14.	WL017	Wetland, Chitwan	G,S,Sm	G,Sm	G,T,S,Sm
15.	WL009	Wetland, Chitwan	G,S,Sm	Sm	G,T,S,I,Sm
16.	RP025	Soil, Rolpa			T
17.	RP022	Soil, Rolpa			
18.	RP026	Soil, Rolpa			
19.	NarRiv1	River, Chitwan			
20.	WL033	Wetland, Chitwan			
21.	WL067	Wetland, Chitwan			
22.	NB008	Soil, Chitwan	G,T,S,Sm	G,Sm	G,T,S,M,I,Sm

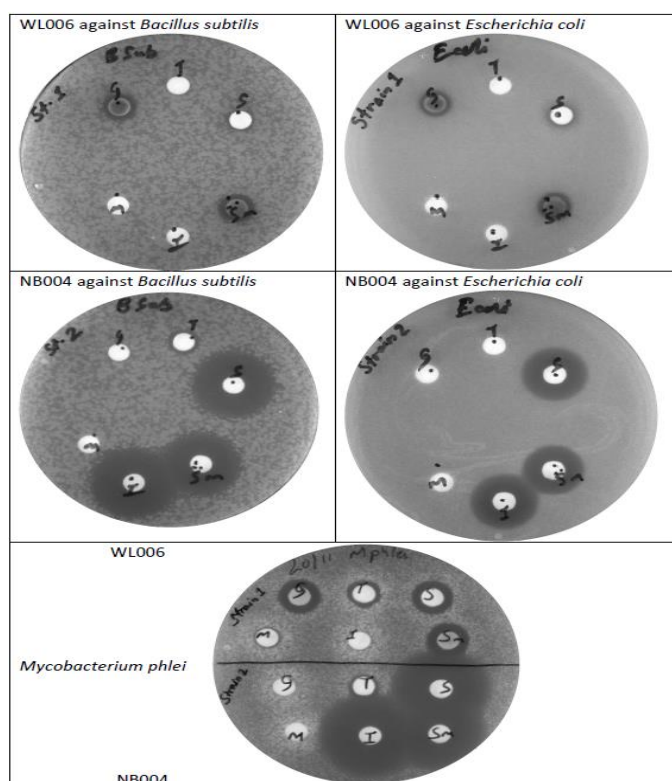


Figure 3-2 Antimicrobial activities shown by crude extracts of strain *Streptomyces* sp. WL006 and *Streptomyces* sp. NB004 against *Bacillus subtilis*, *Escherichia coli*, *Mycobacterium phlei*

3.1.3 Strain characterization

16S rRNA analysis was done for our two prioritized isolates based on their antimicrobial activity. Also, four others were chosen as backup. Upon sequence processing NCBI blast results showed that all the actinomycetes belong to genus *Streptomyces* (

Table 3-4).

Table 3-4 16S rRNA of prioritized isolates

Nr.	Representative isolate (accession no.)	Nearest strain (accession no.)	Sequence identity (%)	Location
1.	WL006 (MN707545)	<i>Streptomyces</i> sp. strain RB110 (KY558688.2)	99.91	Chitwan
2.	NB004 (MN707541)	<i>Streptomyces prasinus</i> (KR085840.1)	99.65	Chitwan
3.	BK003 (MN707540)	<i>Streptomyces californicus</i> (FJ481076.1)	97.73	Chitwan
4.	WL004 (MN707542)	<i>Streptomyces</i> sp. strain ISR_1 (KX035073)	96.54	Chitwan
5.	RP016 (MN707543)	<i>Streptomyces violacens</i> (KF876851.1)	95.80	Rolpa
6.	NP001 (MN707544)	<i>Streptomyces chattanoogensis</i> (NR_042829.1)	99.69	Solukhumbu

3.1.4 Chemical Analysis

A metabolomics based approach using mass spectrometry was employed for analysis of all data in conjunction with SIRIUS 4 platform. Data were acquired in both Positive and negative modes with specification provided in method section 2.3.8.2. Better ionization was observed in positive mode, i.e showing a high abundance of protonated molecules and adducts, as a result these data were considered for further annotation of metabolites.

3.1.4.1 *Streptomyces* sp. WL006

3.1.4.1.1 *Polyether ionophore annotation*

During our initial LC-MS analysis, we observed a strong positive mode ionization for a group of compounds ranging from m/z 737 to 815. But due to a lack of chromophores or weak chromophores these compounds were not detected by the diode array detector (DAD) **Figure 3-3**. These were later recognized as a class of polyether ionophores because of their ability to form complexes with alkali cations. They display pronounced bioactivity as antimicrobial, antitumor and immunosuppressive agent¹⁰⁰. A manual analysis and molecular networking approach were chosen to analyze their tandem MS data. Nonactin, monactin, dinactin, trinactin and tetranactin (**Figure 3-4**) were annotated along with their sodium, potassium and ammonium adducts. LR-LC/MS data are shown in **Figure 3-5** along with their linear form (**Figure 3-6**). These linear forms are product of resistance gene *nonR*¹⁰¹. Crude extract (Red), Fraction E (Purple), Fraction F (Sky Blue) and Media control (Gray) were subjected to molecular networking via GNPS platform to provide further evidence if the molecules with similar fragments build a cluster. Nonactin was confirmed, both in spectral library hit and in MN (**Figure 3-7**).

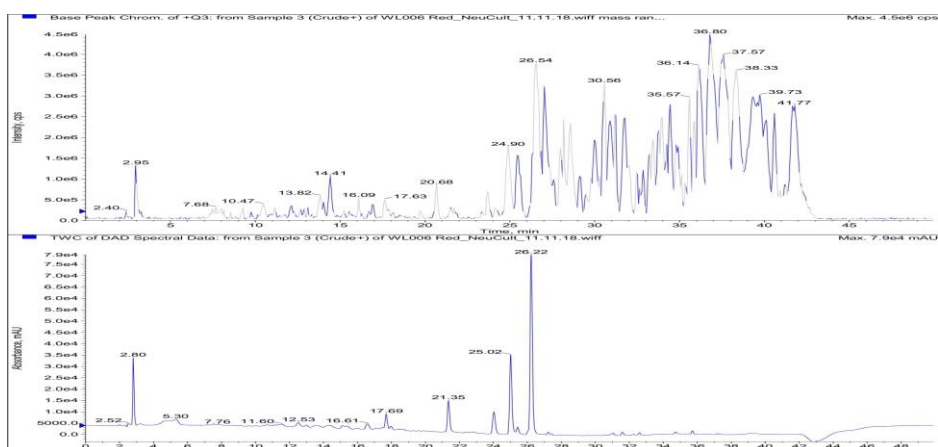


Figure 3-3 BPC (up) and LC (down) of *Streptomyces* sp. WL006 crude extract in positive mode

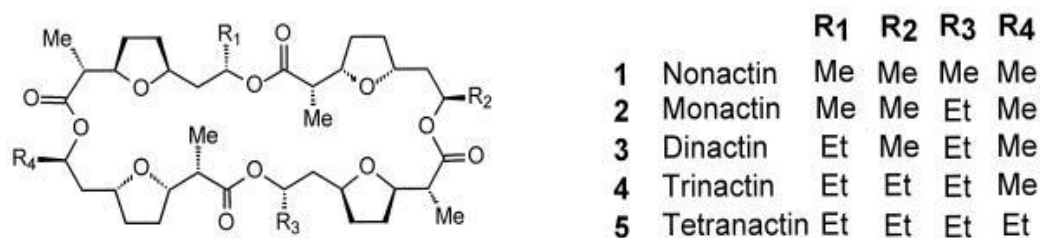


Figure 3-4 Structure of nonactin and other macrocyclic tetralides¹⁰²

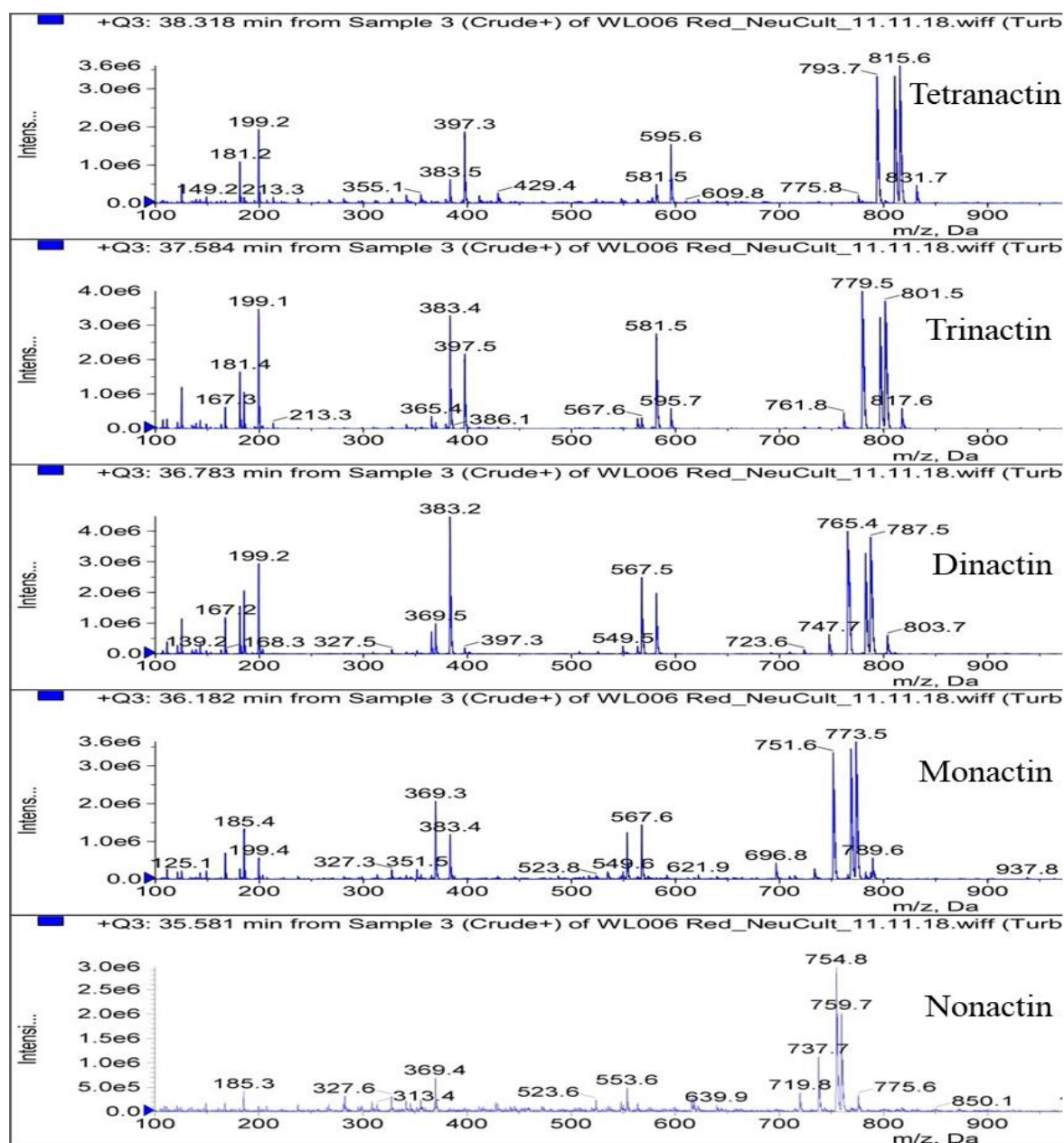


Figure 3-5 LC-MS spectra for nonactin ($m/z=737.7$), monactin ($m/z=751.6$), dynactin ($m/z=765.4$), trinactin ($m/z=779.5$) and tetranactin ($m/z=793.7$) along with their ammonium, sodium and potassium adducts.

NPs from Nepalese actinomycetes

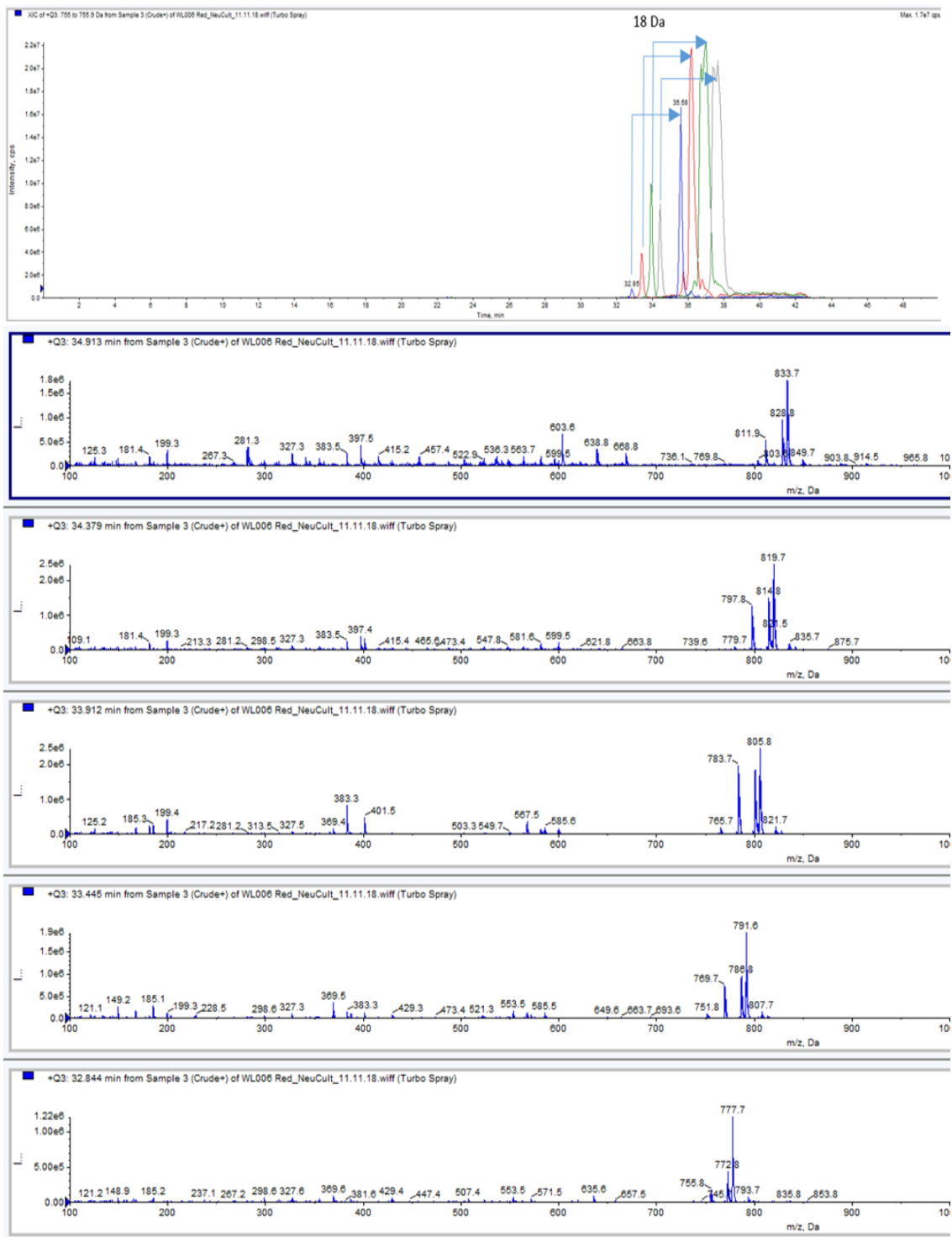


Figure 3-6 XIC and spectra for linear form of nonactin, monactin, dinactin, trinactin, tetranactin with m/z 755.8, 769.7, 783.7, 797.8 and 811.9.

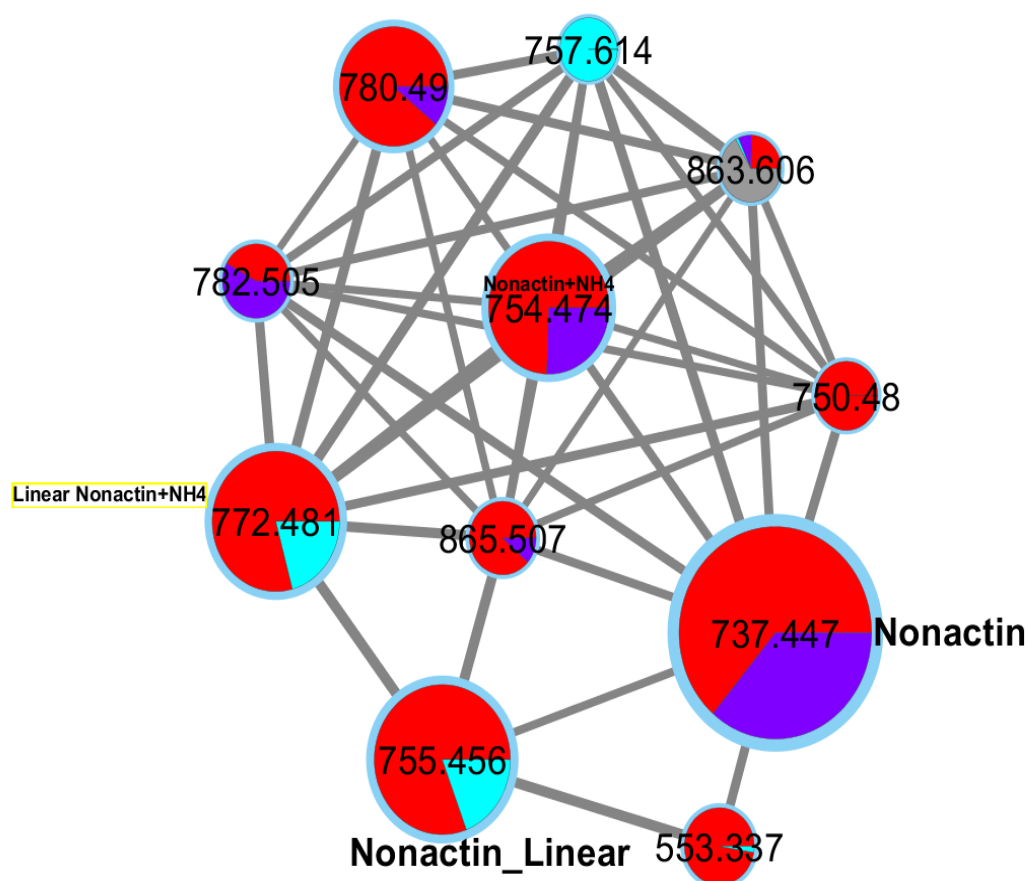


Figure 3-7 Molecular network clustering nonactin, its linear form and ammonium adduct. (Red=Crude, Sky blue=Fraction F, Purple=Fraction E and Gray=Medium)

Subsequently, MS/MS investigations found the mass range of 569-597 Da exhibiting similar fragments to that of macrotetrolides. These masses were in the range of trilactones reported previously by Rezanka and coworker in 2004¹⁰³ (see **Figure 3-9**). However, the set of compounds we investigated were 2 Da more than that of published trilactones and also shared the same chemical affinity for metal ions as like other polyether ionophores. To solve the puzzle, the query MS/MS spectrum was imported to CSI-FingerID for similarity hits⁴⁶. We found no hits for the m/z 569.3319, 583.3473 and 597.3630 on publicly available databases. Upon formula comparison 16 Da more to that of known trilactones as shown in **Table 3-5**. We could speculate that this difference of 16 Da might be from hydroxylation. This was also evident as their degree of unsaturation (rdb) was unchanged and also hydroxylated trilactones would elute earlier (see **Figure 3-8**).

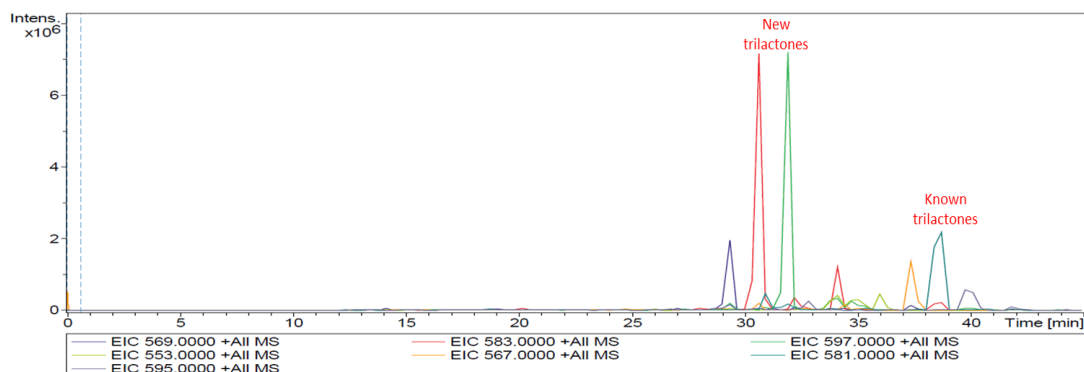
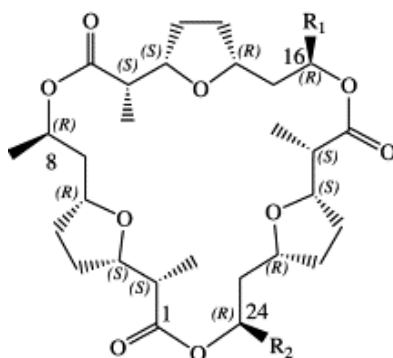


Figure 3-8 Extracted ion chromatogram (EIC) of new trilactones eluting earlier than known trilactones



$R_1=R_2=CH_3$ Trilactone 1 ($m/z=553.3375$)
 $R_1=CH_3, R_2=CH_2CH_3$ Trilactone 2 ($m/z=567.3532$)
 $R_1=R_2=CH_2CH_3$ Trilactone 3 ($m/z=581.3687$)

Figure 3-9 Structure of known trilactones

Table 3-5 Trilactones and their proposed new derivatives A-E

Trilactones by Rezanka <i>et al.</i> , 2004					Diff. in Dalton	Proposed new trilactones		
SN	Compound	Obs. m/z	rdb	Mol. formula		Mol. Formula	rdb	Obs. m/z
1.	Trilactone 1	553.3375	7	$C_{30}H_{48}O_9$	16 Da	$C_{30}H_{48}O_{10}$	7	569.3319- A
2.	Trilactone 2	567.3532	7	$C_{31}H_{50}O_9$	16 Da	$C_{31}H_{50}O_{10}$	7	583.3473- B
3.	Trilactone 3	581.3687	7	$C_{32}H_{52}O_9$	16 Da	$C_{32}H_{52}O_{10}$	7	597.3633- C
4.	Trilactone E (New)	595.3645	7	$C_{33}H_{54}O_9$	16 Da	$C_{33}H_{55}O_{10}$	7	611.3786- D

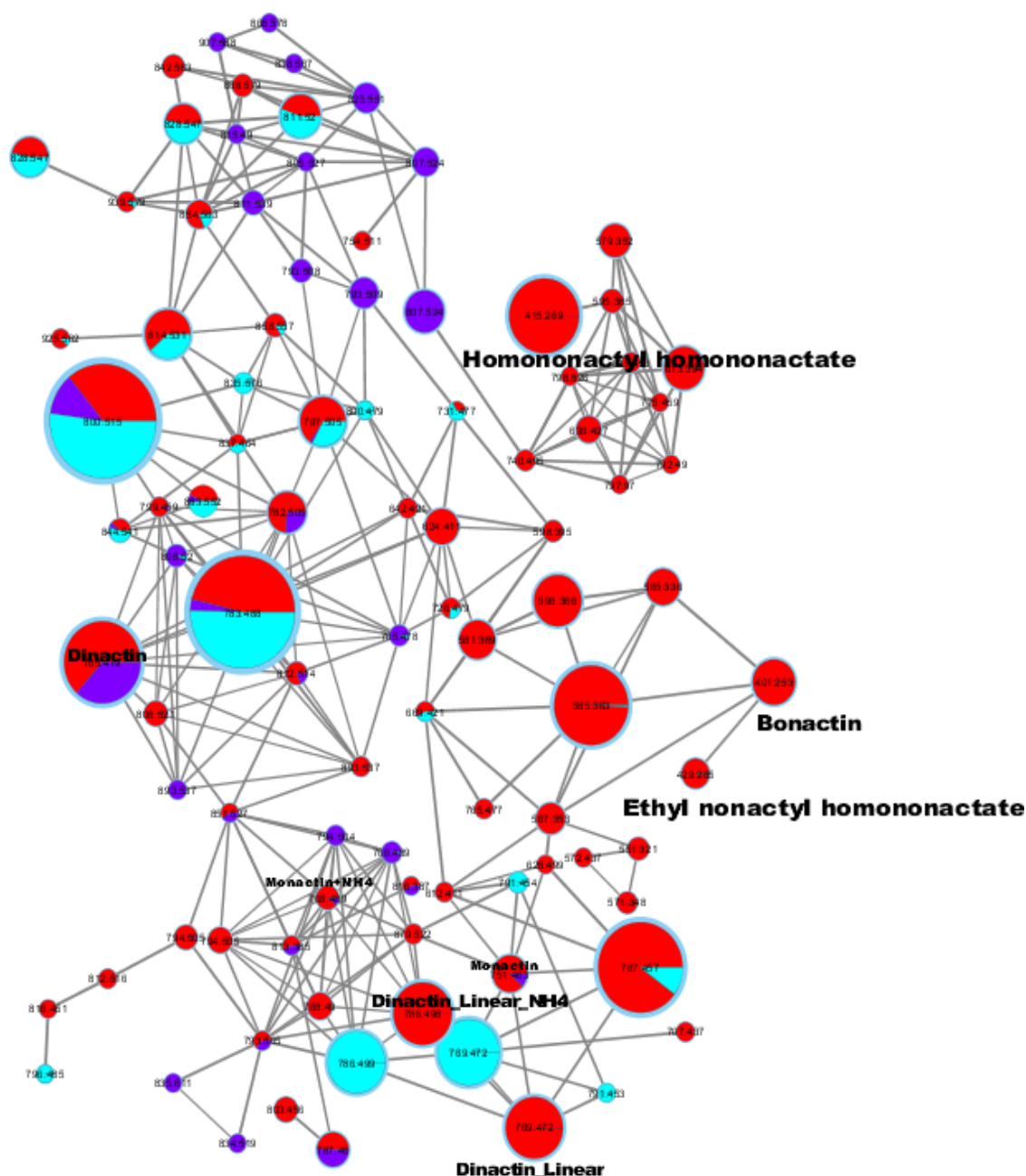
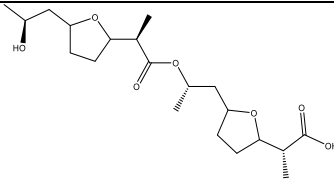
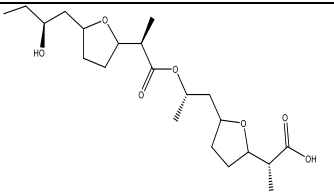
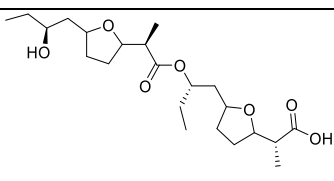
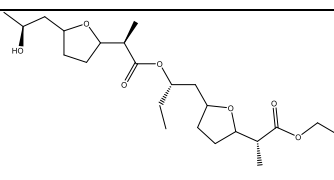
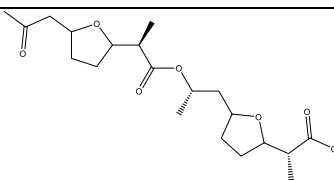
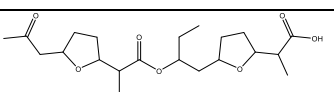


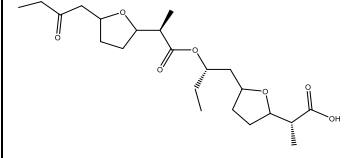
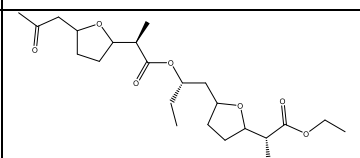
Figure 3-10 MN annotated for polyether ionophores

Similarly, we also annotated \pm nonactic acid and homomonactic acid dimers which presented masses in the range of 387-429 Da which were first reported in late 1990s¹⁰⁴. In addition, the corresponding unsaturated congener of these dimers were also detected in the LR-LC/MS which were proposed to be new (see **Figure 3-11**). Alongside we were also able to detect some of the dimers and their adduct in a MN (see **Figure 3-10**). These results were further supported by *in-silico* fragment comparison and hit generation in CSI-FingerID, a user interface of SIRIUS 4. In this context, with careful observation MS/MS data, partial structure of another new dimer was proposed (see **Table 3-5**).

Table 3-6 Dimers of nonactic acid and homononactic acid

Compounds	Structure	Mol. Formula	ppm	Obs. m/z [M+H] ⁺	rdb	t _R
Known linear dimers						
Nonactyl-nonactate OR dimeric nonactic acid (Fleck <i>et al.</i> , 1996)		C ₂₀ H ₃₄ O ₇	0.6	387.2375	4	22.7
Nonactyl homononactate (Huang <i>et al.</i> , 2015) and Bonactin (Schumacher <i>et al.</i> , 2003)		C ₂₁ H ₃₆ O ₇	0.5	401.2349	4	24.4
Homononactyl homononactate (Huang <i>et al.</i> , 2015)		C ₂₂ H ₃₈ O ₇	1.1	415.2686	4	26.1
Ethyl Homononactyl homononactate (Han <i>et al.</i> , 2014)		C ₂₃ H ₄₀ O ₇	0.0	429.2847	4	28.3
Unsaturated (2Da less) New derivatives						
Unsaturated Nonactyl Nonactate Derivative	 (Proposed structure)	C ₂₀ H ₃₂ O ₇	0.7	385.2218	5	24.0
Unsaturated Bonactin	 (GNPS spectral library hit)	C ₂₁ H ₃₄ O ₇	0.5	399.2375	5	25.4

NPs from Nepalese actinomycetes

Unsaturated Homomonactyl homomonactate		$C_{22}H_{36}O_7$	0.1	413.2534	5	27.0
Unsaturated Ethyl homomonactyl homomonactate		$C_{23}H_{38}O_7$	0.4	427.2692	5	28.6

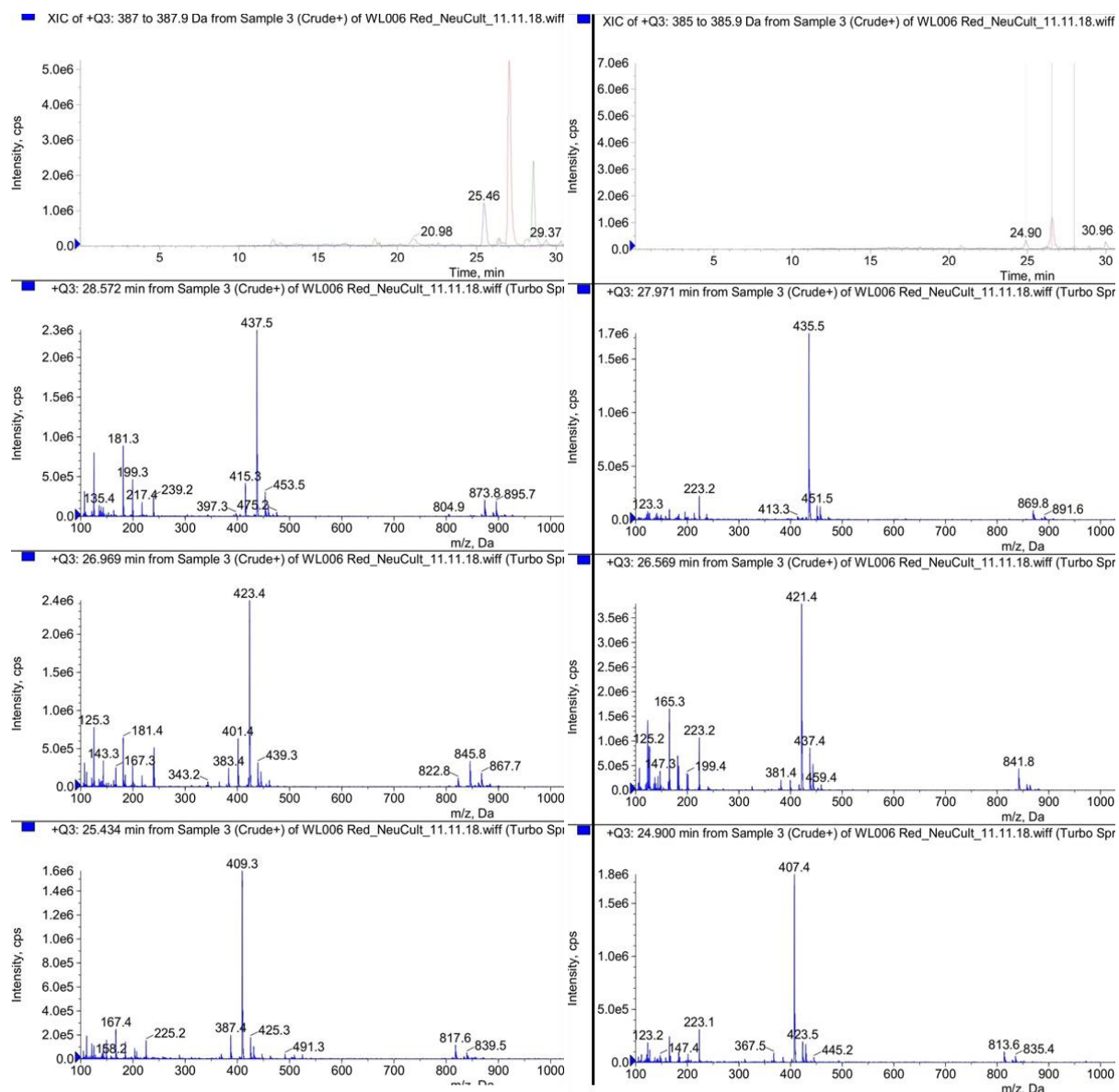
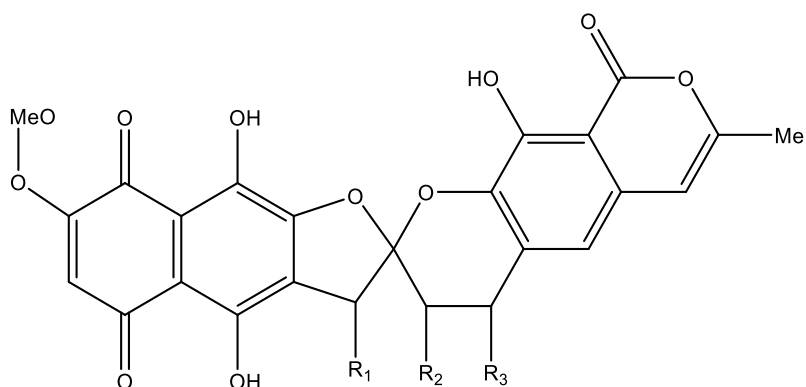


Figure 3-11 LC-MS showing nonactin dimers and their derivatives along with their unsaturated congeners

3.1.4.1.2 Griseorhodin annotation

Griseorhodins belong to rubromycin family and represent an interesting group of pigmented polyketides^{105, 106}. They are distinguished by the dense oxygen functionality located on the spiroketal core (**Figure 3-12**). These structurally diverse molecules exhibit antimicrobial, cytotoxic, human telomerase inhibition and HIV reverse transcriptase inhibition features¹⁰⁷⁻¹¹⁰. Crude extracts of *Streptomyces* sp. WL006 were dense red in color indicating the presence of pigments. Later during our HPLC analysis we observed a set of compounds with UV_(max) 232, 262, 300 nm. To confirm the mass, we performed LC-MS in parallel. Four griseorhodins could be visualized in a global molecular network within a cluster. Interestingly, we have also found new compound exhibiting the similar fragmentation to that of known grisorhodins. Observed *m/z* for this compound was 604.1092. It seems to be clustered together with other known griseorhodins in a MN (see **Figure 3-13**). Raw data analysis also supported the fact that the new mass shared similarity to the griseorhodinA (see **Figure 3-14**). Later, upon careful analysis of the fragmentation pattern and literature survey we annotated the molecule as hyaluromycin¹¹¹. This was possible once we integrated very recent NPAtlas¹¹² database to SIRIUS platform.




$R_1=OH, R_2=R_3=$		Griseorhodin A
$R_1= R_2= R_3= OH$		Griseorhodin C
$R_1= R_2= OH, R_3= H$		Griseorhodin G
$R_1= R_2= OH, R_3= OMe$		8-Methoxy Griseorhodin C
$R_1= OH, R_2= R_3= H$		7,8 Dideoxy Griseorhodin C
$R_1= ketone, R_2= R_3= H$		7,8 Dideoxy-6-oxo Griseorhodin C

Figure 3-12 Griseorhodin and derivatives

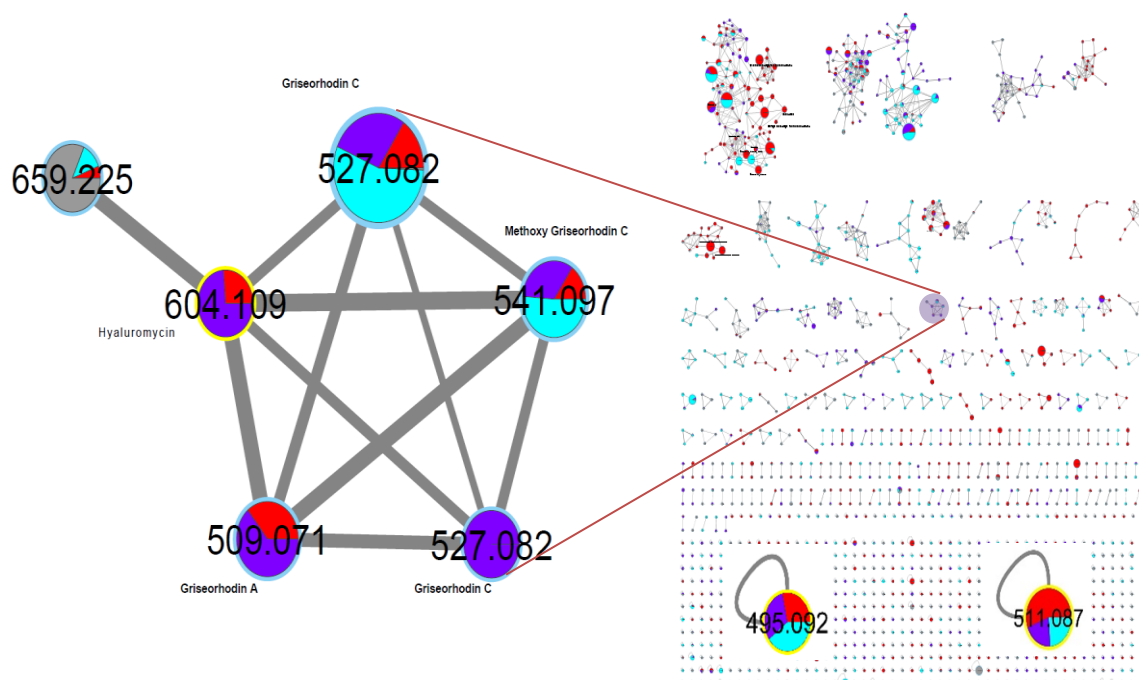


Figure 3-13 Griseorhodin and hyaluromycin clustered together (Red=Crude, Purple=Fraction E, Sky Blue=Fraction F and Gray=Medium).

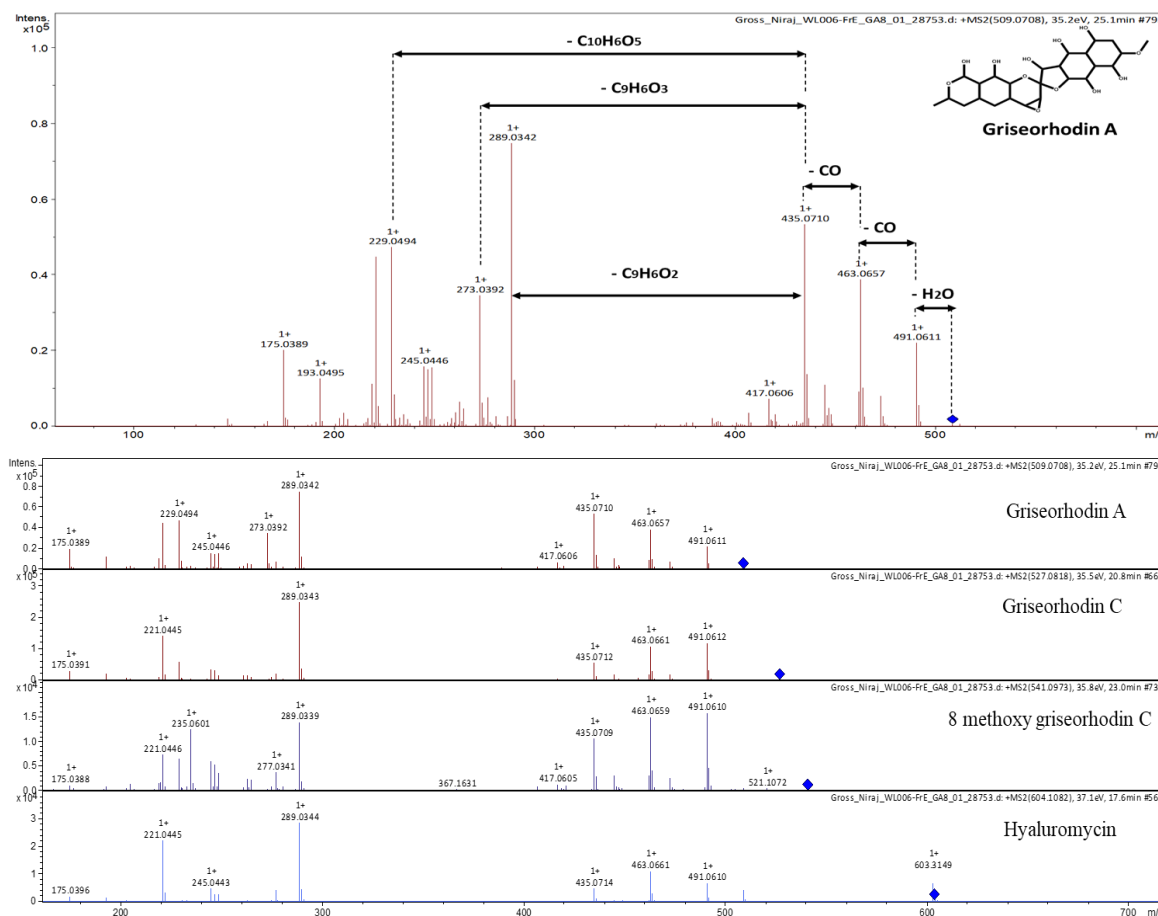
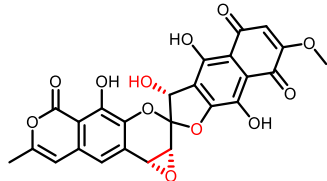
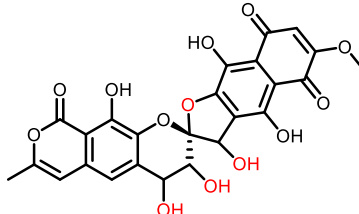
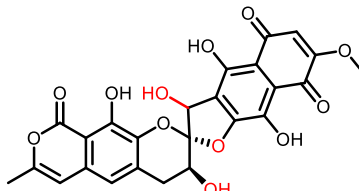
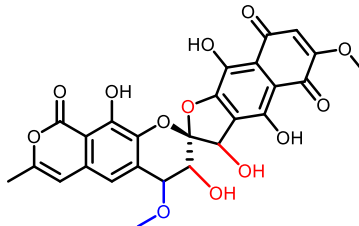
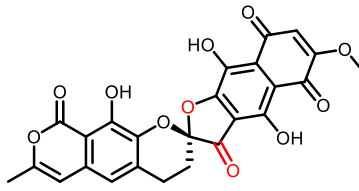
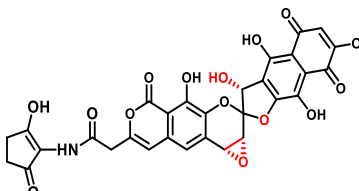


Figure 3-14 Fragmentation of griseorhodin A and its fragment similarity charts to other derivatives from same family (including hyaluromycin)

Table 3-7 Rubromycin class of compounds identified in this study

Compound Annotated	Structure	Mol. Formula	pp m	Obs. m/z	rdb	t _R
Griseorhodin A		C ₂₅ H ₁₆ O ₁₂	0.2 0.5	507.0569-N 509.0715-P	18	25.1
Griseorhodin C		C ₂₅ H ₁₈ O ₁₃	1.1 0.4	525.0681-N 527.0818-P	17	20.8
Griseorhodin G		C ₂₅ H ₁₈ O ₁₂	0.5 0.9	509.0728-N 511.0871-P	17	23.2
8-Methoxy-griseorhodin C		C ₂₆ H ₂₀ O ₁₃	1.0 0.8	539.0836-N 541.0973-P	17	22.8
Dideoxy griseorhodin C		C ₂₅ H ₁₆ O ₁₁	0.3 1.2	493.0763-N 495.0916-P	17	26.7
Hyaluromycin	-95 Da to Griseorhodin A 	C ₃₀ H ₂₁ NO ₁₃	0.6	604.1082-P Poor negative mode ionization	21	17.6

Retention time

P = acquisition in positive mode

N = acquisition in negative mode

t_R =

3.1.4.2 *Streptomyces* sp. NB004

Streptomyces sp. NB004 was isolated from soil of a river island called “Nagarban” in Chitwan district. Crude extract of this isolate showed promising antimicrobial bioactivities which convinced us for further MS and MS/MS analysis. Raw data and molecular network studies pointed towards the presence of actinomycin and analogues. Actinomycins are heavily studied molecules which were reported from several *Streptomyces* strain¹¹³. They possess an impressive bioactive profile with different mechanism of action¹¹⁴. Several publications have been focused in structural characterization of these molecules¹¹⁵. In our early activity analysis of crude extract this isolate possessed highest activity against both gram-positive and gram-negative strains. It was also evident from published data that actinomycin have such activities and could be explained by our dereplication. (see **Table 3-2**) was MN annotation is presented in **Figure 3-15**. Annotations were carried out using GNPS and CSI-FingerID (see **Table 3-8**).

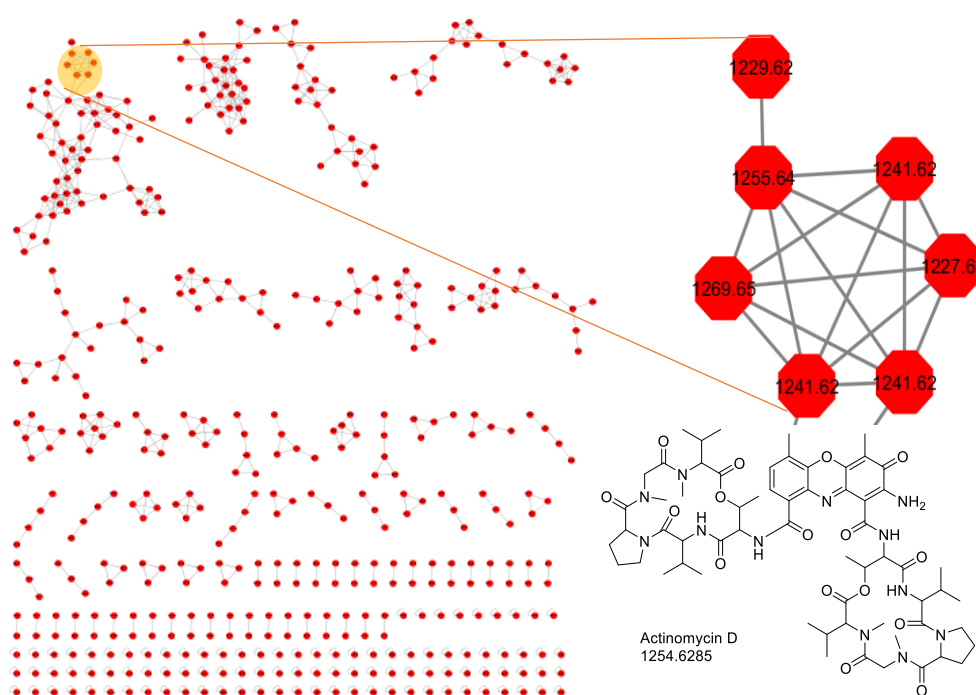


Figure 3-15 MN for Strain NB004 with actinomycin cluster and actinomycin D structure

Table 3-8 Actinomycin annotation table with details

S.N	Observed [M+H]	t _R [^] (min)	Molecular Formula	Error ppm	rdb*	Compound	Annotation source
1.	1269.6550	24.9	C ₆₃ H ₈₈ N ₁₂ O ₁₆	0.9	26	Actinomycin, 2A-D-Leucine	CSI-FingerID
2.	1255.6357	24.2	C ₆₂ H ₈₆ N ₁₂ O ₁₆	0.8	26	Actinomycin D	GNPS
3.	1241.6223	21.2	C ₆₁ H ₈₄ N ₁₂ O ₁₆	1.3	26	Actinomycin D ₀	CSI-FingerID, Antibase

4.	1241.6212	22.8	C ₆₁ H ₈₄ N ₁₂ O ₁₆	-	-	Actinomycin like compound	-
5.	1227.6389	24.1	C ₆₃ H ₈₈ N ₁₂ O ₁₆	-	-	Actinomycin like compound	-

*rdb= ring double bond equivalence, ^t_R= retention time

3.2 Non-Streptomyces - *Pseudonocardia* sp. C8 and *Nonomuraea* sp. C10)

3.2.1 Sampling and pre-screening

Both non-Streptomyces investigated in this study were also categorized as rare actinomycetes. They were isolated from the soil sample of a mud-dauber nest which is still used as traditional medicine in some areas of Nepal. These ubiquitous nests were obtained from a local house of Bardia district, Nepal in year 2017.

In general, mud nests are constructed by a variety of wasps representing the families *Vespidae*, *Sphecidae* and *Pompilidae*. *Sceliphron* genus, in the family *Sphecidae* is assumed to be the wasp involved in making the sampled mud nest in this study. Members of this genus are easily distinguishable from other mud-nest builders by the extraordinarily thin, straight, stalk-like first segment of the abdomen, also known as the petiole, indicated by the arrow in **Figure 3-16**. These solitary wasps hunt spiders as food for their larvae. They are usually not offensive.

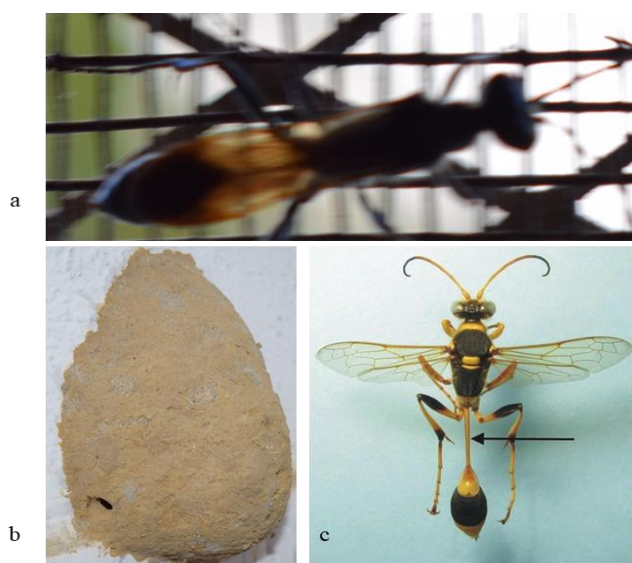


Figure 3-16 a. *Sceliphron* wasp in Nepal (Original picture) b. Wasp Nest. c. *Sceliphron*¹¹⁶

From personal experience and communication with elderly people it was confirmed that mud dauber nest was used in traditional practice in some regions of Nepal. Recently, communities living in Jau and Unini River, in the River Negro basin, Amazon, Brazil¹¹⁷ were also found to have similar practices. They were used to treat mumps and earaches in both regions¹¹⁷. In some regions of Nepal, the clear supernatant of mud dauber nests was also consumed after boiling. This was mainly given to someone who has respiratory problems like tonsillitis or sometimes in case of mild fever. The initial aim of the project was to verify or examine these claims with scientific evidence.

As a first step, morphological investigations of nests were carried out. Most of the nests measured 7-9 cm in length and a width of 3 cm. They were mostly oval in structure and with slight muddy smell. In order to identify the bioactive metabolite in the mud dauber nest itself, extraction and partitioning was done as described in material and method section 2.3.2.2. The obtained nest extracts (ethyl acetate, butanol, and water) however showed no antimicrobial activity. Following up, a LC-MS analysis of the mud dauber nest soil gave no substantial evidence for the presence of secondary metabolites.

Therefore, the approach was adopted and focused rather on the microbiome that were present in the nest. Kumar *et. al.* carried out screening to actinomycetes from mud dauber wasp nest and carried out antimicrobial activities test^{118, 119}. However, till date no chemical analysis has been reported. In our study firstly we isolate nest associated actinomycetes and screen them for secondary metabolites. For this purpose, nest soil was processed with a method described by Hayakawa⁹¹ previously. Upon rigorous and selective screening, four distinct actinomycetes looking colonies were isolated from the sampled mud dauber nest. Growth of other non-sporulating bacteria were discouraged with heating the soil samples at higher temperatures (100°C). Subsequently, the morphology was also checked within actinomycetes section of Bergey's manual of bacteriology¹²⁰. Growth characteristics, appearance of colonies pointed towards the presence of actinobacteria. Later, they were also confirmed with partial sequencing. This study relied more on the ecological standpoint. We assumed that unique niches would allow us the opportunity to isolate rare bacterial strains which will probably result in novel chemistry.

3.2.2 Sequencing and Bioinformatic analysis

A series of subcultures were required to achieve axenic cultures of actinomycetes. Initial priority for screening was based on morphology and growth characteristics from which we were able to prioritize 4 actinomyces like colonies. Only 3 strains were partially sequenced (16S rRNA) to confirm their identity while for the remaining strain extraction of genomic DNA was

problematic. As a result, two rare actinomycetes came to our knowledge (**Table 3-9**). To understand more on genomic level we performed whole genome sequencing (WGS)¹²¹ of these rare actinomycetes. The reads were assembled to a 6,241,701-nucleotide draft genome at 189-fold coverage. The resulting sequence consists of two contigs with a G+C content of 73.60 %. Gene functional annotation using PGAP v4.12¹²² identified 5,637 coding genes.

Table 3-9 16S rRNA of 4 colonies isolated from mud dauber nest soil.

S.N	Representative isolate (accession no.)	Nearest strain (accession no.)	Sequence identity (%)	Location
1.	COL1 (MN707538)	<i>Streptomyces</i> sp. (LC487847.1)	99.93	Bardia
2.	COL8 = <i>Pseudonocardia</i> sp. C8 (MN707539)	<i>Pseudonocardia</i> sp. (KR057433.2)	98.97	Bardia
3.	COL10 = <i>Nonomuraea</i> sp. C10 (MN400757)	<i>Nonomuraea</i> sp. strain 7K523	99.00	Bardia
4.	COL4	NA	NA	Bardia

An automated genome-based taxonomic analysis of strain C8, employing the Type Strain Genome Server (TYGS),¹²³ revealed that *Pseudonocardia ammonioxydans* CGMCC 4.1877^{T124} represents the closest related type strain of C8. In pairwise comparisons, independent of the applied Genome BLAST Distance Phylogeny (GBDP) formula, the digital DNA-DNA hybridization (dDDH) values d0, d4, and d6 did not exceed 32.1%. Since these values are well below the species threshold of 70 %, strain C8 represents a candidate new *Pseudonocardia* species. This finding was complemented by an analysis of the average nucleotide identity (ANI) using autoMLST¹²⁵, which revealed that the C8 genome sequence had 86.7% ANI to *Pseudonocardia ammonioxydans* CGMCC 4.1877^T. Since this ANI value is well below the boundary of 95 to 96 % ANI for species delineation, thereby corroborating the TYGS results. Similarly, partial and WGS were done for strain C10. Based on 16S rRNA gene sequence (1,367 bp) similarity, bacterium C10 or COL10 was identified as a *Nonomuraea* species. The strains most closely related to C10 are *Nonomuraea* sp. strain 7K523 (GenBank accession number MG770787) and *Nonomuraea harbinensis* Gsoil-1046 (KY078837), both with 99% sequence identity. Among the actinobacteria, the genus *Nonomuraea* has one of the largest genome, e.g. *Nonomuraea* sp. ATCC 55076¹²⁶ which possesses the genome of 13.1 Mbp. A genome of strain C10 consists of a total of 9,416,283 bp, and a G+C content of 71.4%. Gene functional annotation using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP v4.8) (12) identified 8,473 coding genes.

3.2.2.1 Antismash results for *Pseudonocardia* sp. C8

Automated secondary metabolism analysis using AntiSMASH 6.0.0¹²⁷ predicted 16 biosynthetic gene clusters (BGCs) in 15 regions indicated in **Figure 3-17**. Only one matched at the 100% level the known cluster encoding the compatible solute ectoine¹²⁸, while the remaining BGCs might code for new secondary metabolites. Unlike strain C10, C8 possessed only few PKS products and many RiPP clusters. In our preliminary chemical analysis, we could not observe metabolites which would answer the genomic potential of this strain. Therefore, an in-depth analysis of the genome is required in order to draw more conclusion and to design an appropriate medium for production.

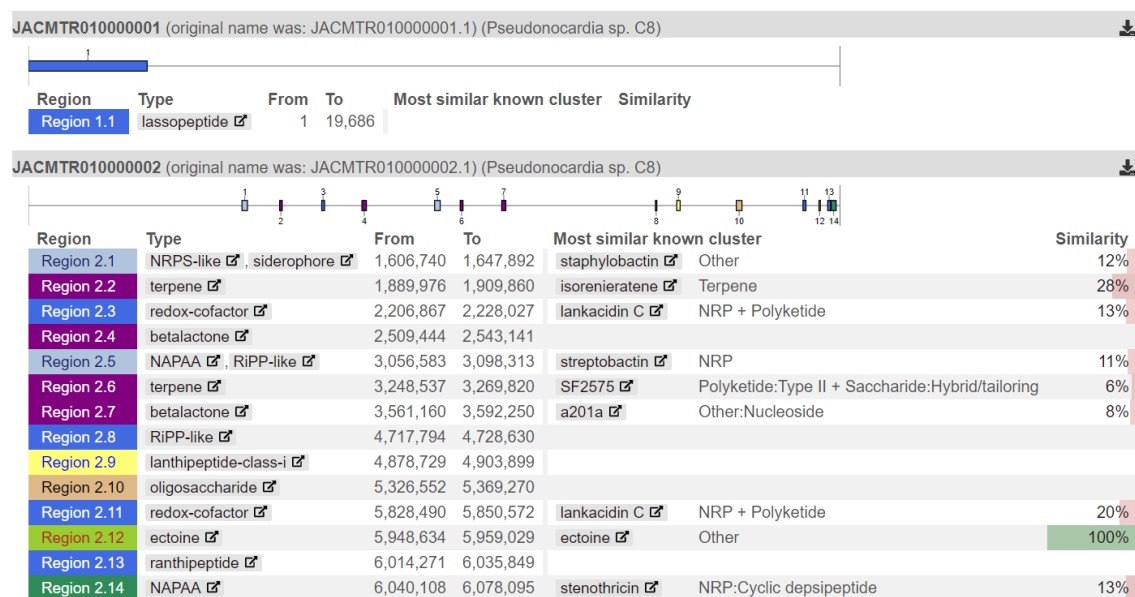


Figure 3-17 Antismash results for strain C8

3.2.2.2 Antismash results for *Nonomuraea* sp. C10

Antismash 5.0 identified 21 regions, however 5 thereof contained multiple pathways, which elevates the actual number of BGC's encoding SM's to 30 (**Figure 3-18**). Based on a 100% rate of similarity, cluster 1.11 and 3.3 could be readily assigned to the production of alkylresorcinol and geosmin. Biosynthetically, OA are simple iterative PKS with alkylresorcylic acid in it. It is known to be synthesized from one acetyl-CoA unit and three molecules of malonyl-CoA. Even though, orsellinic ester was isolated without prior genomic information, one goal of the bioinformatic analysis was to find the gene cluster responsible for its production. Upon closer analysis, we prioritized four possible candidates for its production. Cluster 1.5, a typeIII PKS system was also found to make such a product. These are normally seen in plants.¹²⁹ Also, cluster 1.10, 1.15 and 3.2 were also equally good candidate as they

NPs from Nepalese actinomycetes

were type I PKS. With analysis of adjacent genes, we finally reached to the conclusion that 1.15 is the prioritized candidate.

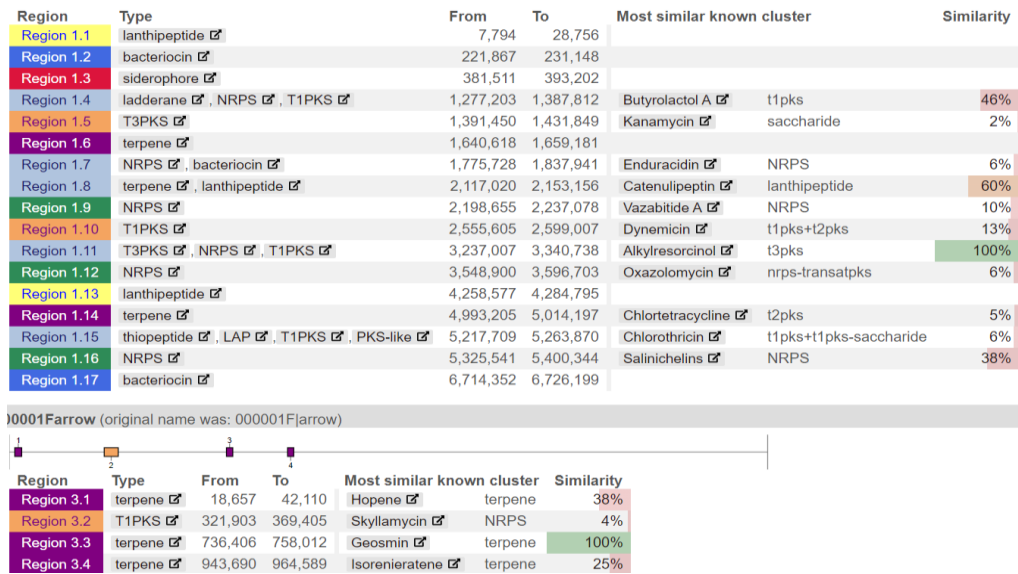


Figure 3-18 Antismash results for strain C10

From previous biosynthetic studies made on avilamycin it was evident that aviM gene is involved in biosynthesis of OA¹³⁰. Similar organization was observed for cluster 1.15 **Figure 3-19**. Further supporting our hypothesis, in **Figure 3-20**, the AT domain showed 100% similarity with avilamycin and everminomycin compounds which also harbours OA in their structure. Further molecular biology, for example heterologous expression or knock out studies could give additional evidence on this.



■ Structural Genes ■ Additional Biosynthetic Genes

Figure 3-19 Predicted BGC for Orsellinic acid from *Nonomuraea* sp. C10

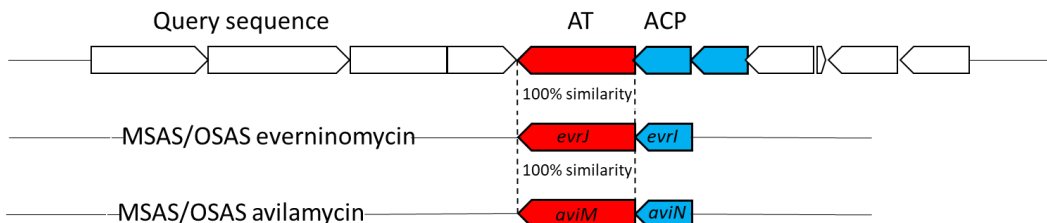


Figure 3-20 100% similarity of MSAS/OSAS AT domain (red) and ACP domain from query sequence to avilamycin and everminomycin.

NPs from Nepalese actinomycetes

Another cluster of interest was cluster 1.10, which pointed towards presence of an enediyne polyketide synthase. A NCBI blast analysis gave hits for EspE which encodes an enediyne structure in esperamicin¹³¹. This was supported by neighbouring genes which were annotated by antiSMASH 6.0.0 as enediyne biosynthesis genes. Also, esperamicin was reported to be produced by *Actinomadura verrucosospora*. In 2000, some of the bacteria of the genus *Actinomadura* were transferred to the new genus named *Nonomuraea* and both belong to the group rare actinomycetes class¹³². In-depth studies on these results are still required, however future efforts will be more directed towards the characterization of orsellinic acid gene cluster from strain *Nonomuraea* sp. C10 by knockout experiments.

3.2.3 Chemical analysis of *Nonomuraea* sp. C10

Butanolic extract of *Nonomuraea* sp. C10 supernatant was subjected to fractionation. Approximately 4 g crude extract was loaded to the C18 stationary phase and gradient-wise elution was performed as illustrated in **Figure 3-21**. As we obtained nearly pure substance in our W0 fraction, it was prioritized over others. Upon isolation and further spectroscopic analysis, orsellinic acid ester was identified and characterized from fraction W0, while brartemicin and other derivatives were recovered from fraction W20 as a major compound and in W40 as a minor compound. Since more polar fractions (W100, W80, W60) contained higher proportion of the media components, their analysis was not conducted.

The UV spectrum of orsellinic acid, with absorption maxima at (λ_{\max}) 216, 263, and 300 nm, was indicative of a benzoyl chromophore. (see **Figure 3-22** and **Figure 3-23**). It was evident that fraction W20 contains a set of compounds with similar structures or the same chromophore. HPLC equipped with semi-preparatory column (Synergi Polar-RP & Omega Polar) were used for the further purification of compounds as shown in **Figure 3-22**. Pure compounds were confirmed with NMR and MS data, which will be separately presented in sections below. Just from the HPLC analysis it was indicative that more structures which might contain the UV chromophore like orsellinic acid as and which could be potentially new.

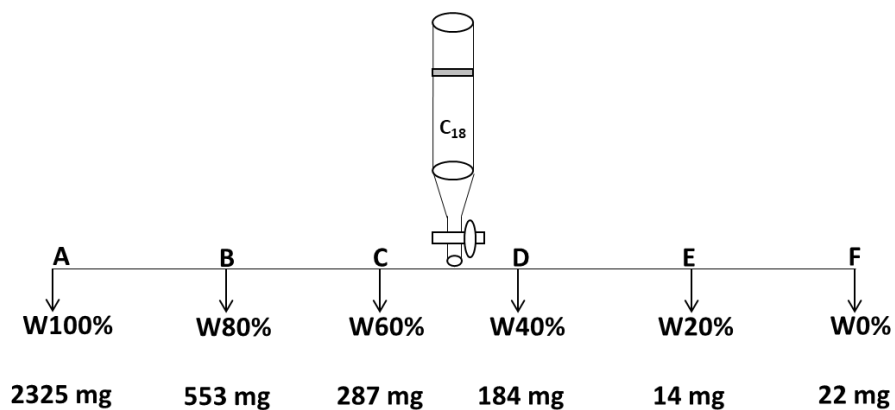


Figure 3-21 *Nonomuraea* sp. C10 fractions (W=Water in Water:MeOH mixture)

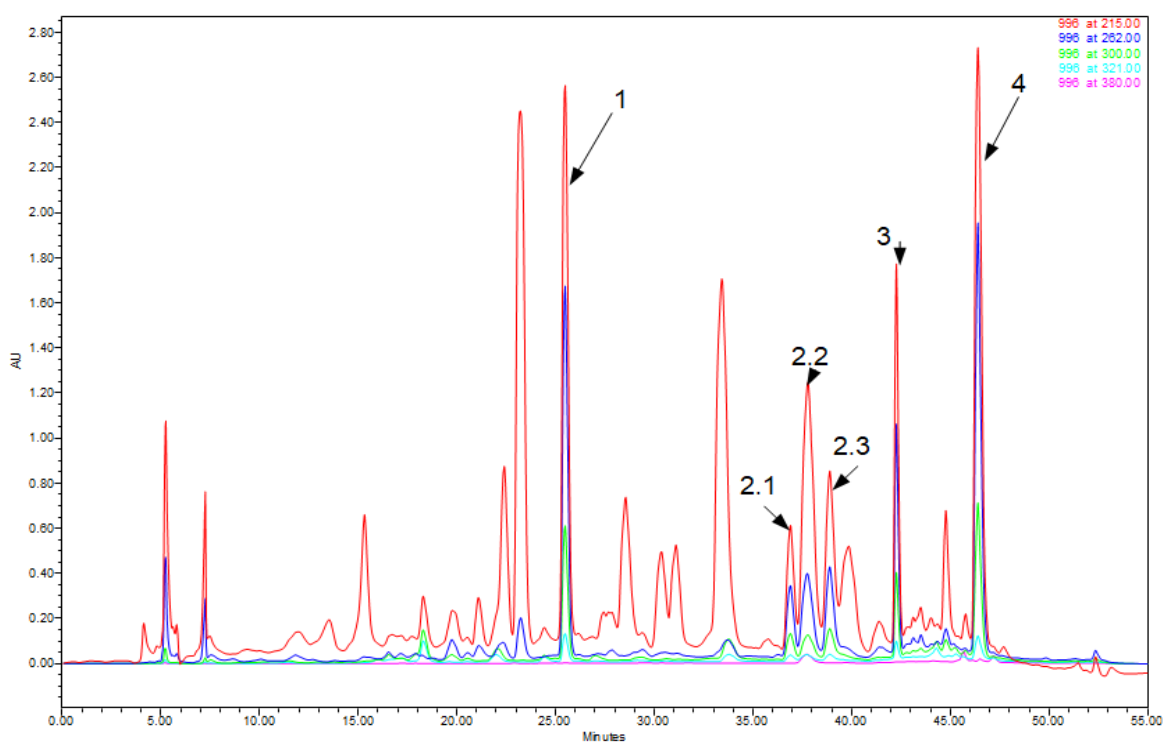


Figure 3-22 HPLC chromatogram of fraction W20 showing brartemicin (3), orsellinic acid ester (4) and peak 1, peak 2.1 and 2.3 which are proposed new derivatives of brartemicin.

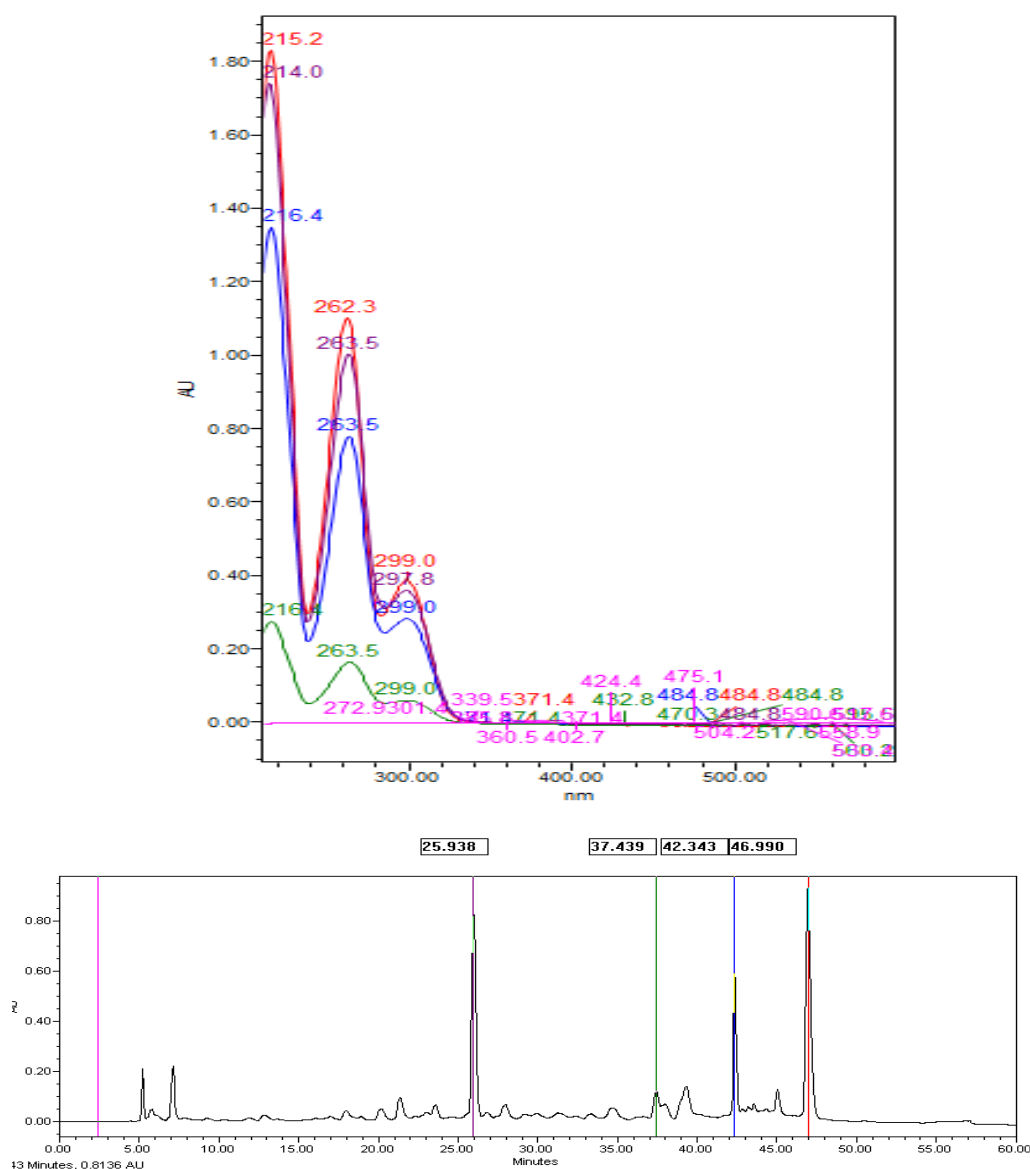


Figure 3-23 Absorption spectra (Top) of brartemicin and derivatives in W20 fraction (Pink=Baseline, Purple =Peak 1, Green = Peak 2.1, Blue = Brartemicin, Red = OA ester)

3.2.3.1 HRMS analysis for brartemicin and its related compounds

To provide more evidence for the presence of brartemicin and its new natural derivatives we performed HR-LCMS/MS analysis for fraction W20. This allowed us to understand the fragmentation pattern of the molecule and to detect derivatives on the basis of fragmentation similarity as well. For this purpose, the ion $m/z = 643$ was extracted in positive mode to see the EIC for brartemicin. To our surprise, this ion was detected at multiple retention time in EIC (see **Figure 3-24**). Also, tandem MS fragments of these peaks were very similar. This also indicated that several compound bearing an orsellinic acid moiety were present in the fraction To get more insight, we closely examined the fragments. Fragment $m/z = 151.039$ (red) is the product of ester bond breakage while $m/z = 313.092$ (blue) indicates the presence of a sugar moiety besides the

NPs from Nepalese actinomycetes

aromatic chain. These results indicated that fraction W20 might consist of many positional isomers of this class of compounds. Consequently, all derivatives that could be obtained in sufficient quantity were isolated.

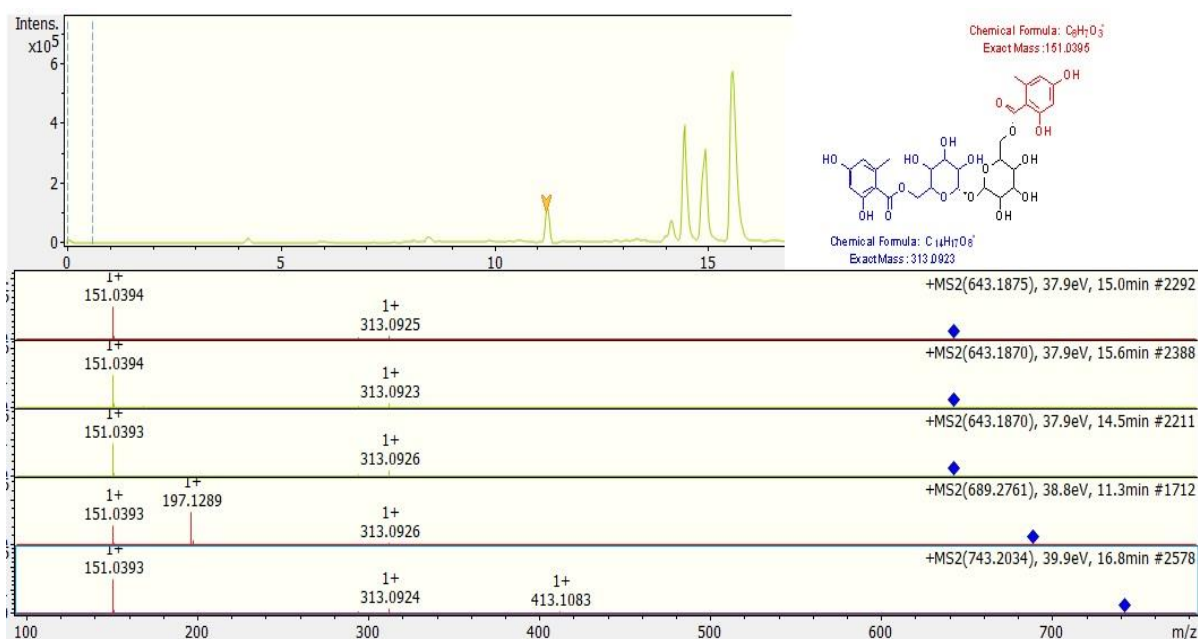


Figure 3-24 MS² showing brartemicin and their novel derivatives.

3.2.3.2 Compound 1

From HRMS molecular formula for compound **1** (1 in HPLC chromatogram) was established as C₂₀H₂₈O₁₄ exhibiting ring double bond seven. This compound was judged to be relatively polar and eluted early on a C18 column. HPLC isolation was achieved with Luna Omega Polar C18 column. A total of 5.01 mg of compound **1** was subjected to NMR measurements for full structure characterization, which revealed the structure as shown below in **Figure 3-25**. NMR table of compound **1** is presented below in **Table 3-10**.

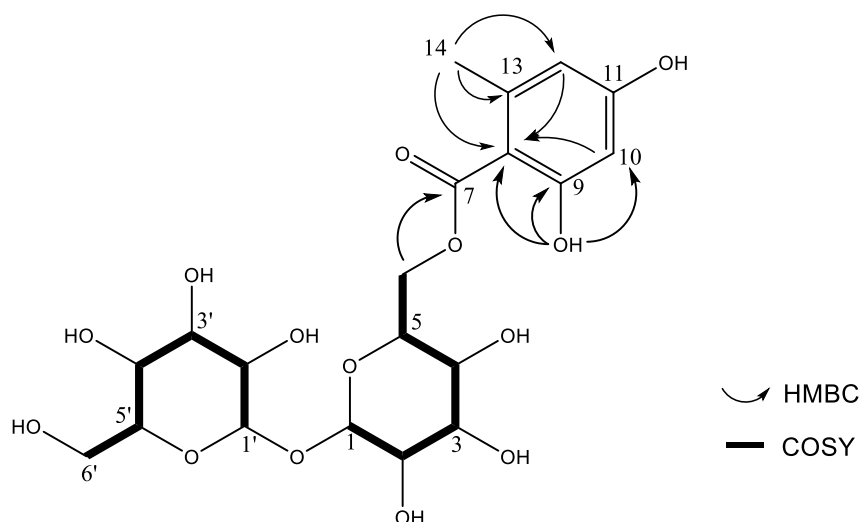


Figure 3-25 Structure of compound **1**

It was clear from the sum formula that compound **1** lacked some carbons in comparison to brartermicin. The reduction of ring double bond equivalent value from 12 to 7 indicated the loss of aromatic portion of the molecule. This was also seen by NMR experiments.

Table 3-10 NMR signals of compound **1** recorded in d_6 -DMSO (400 MHz)

Unit	Position	$\delta_{C/N}^a$, mult.	δ_H^b (mult., J in Hz)
Trehalose	1'	93.4, CH (overlap)	4.88, d (3.6)
	2'	71.4, CH	3.23, m
	3'	72.8, CH	3.55, m (overlap)
	4'	70.0, CH	3.14, dd (9.38)
	5'	72.6, CH	3.66, m
	6'	60.6, CH ₂	3.55, dd (overlap) 3.44, m
	1	93.4, CH (overlap)	4.90, (d 3.6)
	2	71.4, CH	3.27, (dd, 9.6, 3.6)
	3	72.5, CH	3.60, (t, 9.28)
	4	70.3, CH	3.20, (t, 9.52)
	5	69.6, CH	4.05, m
	6	64.4, CH ₂	4.44, dd (12.0, 1.93) 4.30, dd (12.0, 5.18)
Orsellinic acid	7	170.2, qC	
	8	105.9, qC	
	9	162.6, qC	
	10	100.5, CH	6.15, d (2.3)
	11	161.7, qC	
	12	110.8, CH	6.19, d (2.1)
	13	142.0, qC	
	14	23.0, CH ₃	2.37, s
	Aromatic OH		11.04, s
	Aromatic OH		10.16, s

NPs from Nepalese actinomycetes

Positive mesomeric effects were seen for protons $\delta_{\text{H}} = 6.15, 6.19$ (dd). Upon further analysis of recorded 2D spectra, we could clearly observe HMBC correlations from the sugar methylene to the carbonyl atom ($\delta_{\text{C}} = 170.2$). The proton rich sugar part was established with COSY correlations while the aromatic region was further confirmed with HMBC experiments as shown in **Figure 3-25**. Thus, the structure of compound **1** was confirmed.

Compound **1** represents a monoester brartemicin and was previously synthesized by Jacobsen and coworkers in 2015 with in the frame of structure activity relationship study.¹³³ It was identified as synthesis product, but it has been not reported from any natural sources yet. Though synthesis was already done for this compound, for biotechnological purposes it can be produced also from this strain whenever required. Previously, brartemicin was identified as the high affinity ligand for the carbohydrate-recognition domain of the macrophage receptor mincle. But its monoester derivative was slumped in activity, which indicates the importance for the benzoyl moiety in binding. In this study, we isolated this monoester compound from the bacterium itself. This finding leaves the question if this monoester derivative is a biosynthetic intermediate or a degradation product. Further studies are required in this direction along with the determination of activity.

3.2.3.3 Compound 2.1 and 2.3

During HPLC analysis peak **2.1** and **2.3** showed the same UV profiles, indicating the presence of benzoyl moiety. In MS² analysis, they fragmented identically and also had same molecular formula as brartemicin, which is C₂₈H₃₄O₁₇. However, the retention time was substantially different in both cases (see **Figure 3-24**). We speculated them to be positional isomers of brartemicin. In this case, we reach the limitation of mass spectrometry analysis which mainly relied on fragmentation. To shed more light on it, NMR spectra were recorded. Target compound **2.1** was purified from two upscaled batches. A total of 3.01 mg compound was subjected to NMR. During our NMR analysis it became apparent that compound **2.1** is a new brartemicin derivative (see **Figure 3-26**). Unlike brartemicin, the symmetry for compound **2.1** was disturbed which resulted in separate set of 28 ¹³C signals (**Table 3-11**). NMR measurements revealed the true structure of a novel brartemicin derivative as shown below, which was as hypothesized, a positional isomer of brartemicin. Though we speculated that compound **2.3** also represented another positional isomer, due to its low amount a full characterization was not possible.

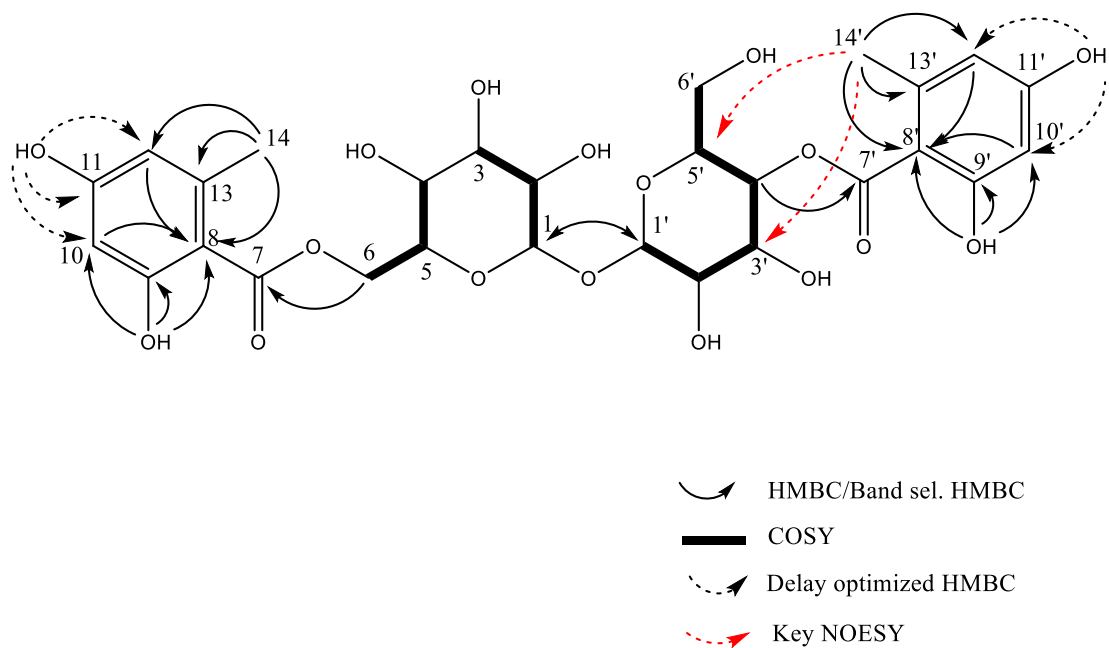


Figure 3-26 Structure of new brartemicin derivative

Table 3-11 NMR data of compound 2.1 recorded in d_6 -DMSO (700 MHz)

Unit	Position	$\delta_{C/N}^a$, mult.	δ_H^b (mult., J in Hz)
Trehalose	1'	93.4, CH	4.94, d
	2'	71.4, CH	3.28, dd
	3'	70.4, CH	3.86, t
	4'	72.1, CH	4.91, t
	5'	70.2, CH	4.00, m
	6'	60.4, CH ₂	3.40, m
	1	93.1, CH	4.97, d
	2	71.6, CH	3.42, m
	3	72.4, CH	3.61, t
	4	70.3, CH	3.21, t
	5	69.7, CH	4.04, m
	6	64.5, CH ₂	4.31, dd 4.46, dd
Orsellinic acid 1	7	170.2, qC	
	8	106.0, qC	
	9	162.5, qC	
	10	100.5, CH	6.16, d
	11	161.7, qC	

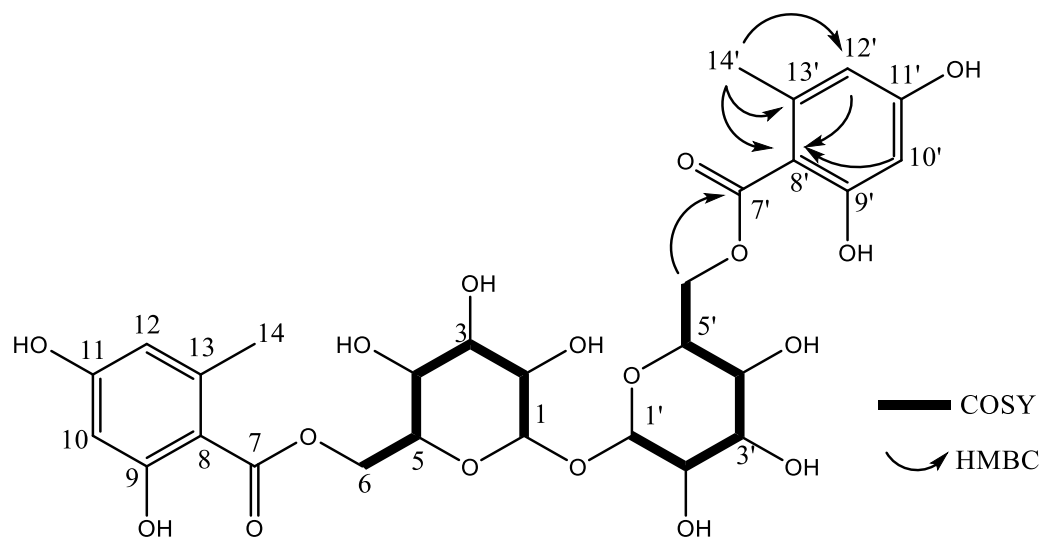
	12	110.8, CH	6.21, d
	13	141.9, qC	
	14	23.0, CH ₃	2.38, s
Orsellinic acid 2	7'	169.0, qC	
	8'	107.8, qC	
	9'	160.9, qC	
	10'	100.4, CH	6.16, d
	11'	161.0, qC	
	12'	110.4, CH	6.17, d
	13'	140.9, qC	
	14'	22.3, CH ₃	2.31, s

3.2.3.4 Compound 3

The benzoyl moiety containing compound 3 was collected via HPLC. 4 mg of colourless sample was subjected ¹³C and ¹H NMR experiments which revealed only 14 carbon signals indicating the smaller molecule. However, LC MS/MS positive mode analysis predicted the compound [M+H]⁺ as 643.1870 Da in size is associated with molecular formula as C₂₈H₃₄O₁₇ with rdb equivalents of 12. Later after full set NMR data, it became apparent that strain C10 produces a glycosyl glycoside derivative that consists of α, α-trehalose substituted at positions 6 and 6' by O-2,4-dihydroxy-6-methylbenzoyl groups, brartemicin (**Figure 3-27, Table 3-12**). It is a symmetrical compound reported with cell invasive activities³⁵. Due to its symmetry NMR correlations were established only for 1 sugar and 1 aromatic system. The aromatic region of brartemicin, termed as orsellinic acid (OA) is a 2,4-dihydroxybenzoic acid which belongs to the family of resorcinols. Though OA is a PKS product the actual route of biosynthesis OA in brartemicin still remains elusive. OA has been taken as precursor in many natural products. Fungi are among the favourites to harbour OA BGCs.¹³⁴⁻¹³⁷ However, plants and bacteria were also reported to generate this aromatic moiety^{129, 138}. In fungi and bacteria iterative type I PKSs are known to synthesize orsellinic acid while plants possess type III PKS to biosynthesize the same product.¹²⁹ OA synthase has been investigated during the biosynthesis studies on avilamycin (AviM) in *Streptomyces viridochromogenes* Tü57.¹³⁰

Table 3-12 NMR signals of compound 3 recorded in d_6 -DMSO (400 MHz)

Unit	Position	$\delta_{C/N}^a$, mult.	δ_H^b (mult., J in Hz)
Trehalose	1, 1'	93.9, CH	4.90, (d 3.6)
	2, 2'	71.4, CH	3.28, (dd, 9.6, 3.6)
	3, 3'	72.5, CH	3.60, (t, 9.28)
	4, 4'	70.3, CH	3.22, (t, 9.52)
	5, 5'	69.6, CH	4.07, m
	6, 6'	64.4, CH ₂	4.44, dd (12.0, 1.93) 4.30, dd (12.0, 5.18)
Orsellinic acid	7, 7'	170.2, qC	
	8, 8'	106.0, qC	
	9, 9'	162.6, qC	
	10, 10'	100.5, CH	6.15, d (2.3)
	11, 11'	161.7, qC	
	12, 12'	110.8, CH	6.19, d (2.1)
	13, 13'	142.0, qC	
	14, 14'	23.0, CH ₃	2.37, s

**Figure 3-27** Structure of Brartemicin

3.2.3.5 Compound 4

Unlike other brartemicin derivatives compound 4 was smaller in size. HRMS data revealed a mass of $m/z = 223.0971$ $[M - H]^-$ (calcd for $C_{12}H_{15}O_4$) with rdb equivalent of 5. The structure of compound 4 was fully confirmed by NMR data (**Table 3-13**). Upon SciFinder search, compound 4 was identified as n-butyl orsellinate (**Figure 3-28**). It was apparent from a literature survey that orsellinates were repeatedly isolated from various lichens and occasionally from fungus. They seem to possess moderate antifungal¹³⁹, antioxidant and mushroom tyrosinase activities¹⁴⁰. Despite of its activity and biosynthesis one thing that raises our attention is the source of isolation. Normally, orsellinates are not isolated from microbial origins and especially actinomycetes. Therefore, we doubted that orsellinate isolated in this study might be an extraction artifact because our culture supernatant was extracted with n-butanol. Therefore, we attempted extraction with other solvent like ethyl acetate and performed LC-MS analysis to look for the target mass. However, the result was not decisive as the compound showed a poor ionization.

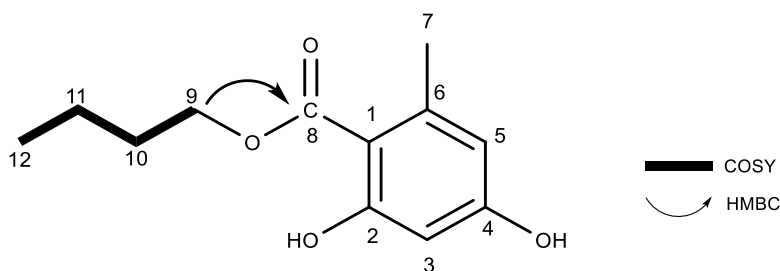


Figure 3-28 n-butyl orsellinate

Table 3-13 NMR signals of compound 4 recorded in d_6 -DMSO (400 MHz)

Unit	Position	$\delta_{C/N}^a$, mult.	δ_H^b (mult., J in Hz)
Aromatic	1	107.1, qC	
	2	161.5, qC	
	3	100.4, CH	6.14, d (2.4)
	4	161.2, qC	
	5	110.3, CH	6.16, d (2.4)
	6	140.8, qC	
	7	22.3, CH ₃	2.30, s
	8	170.1, qC, C=O	
	9	64.4, CH ₂	4.23, t (6.4)

Side chain	10	30.0, CH ₂	1.67, m
	11	18.7, CH ₂	1.40, m
	12	13.5, CH ₃	0.91, t (7.36)

3.2.4 Chemical analysis of *Pseudonocardia* sp. C8

During our HPLC profiling (Figure 3-29), it was evident that there were 3 outstanding peaks (C8P4, C8P5 and C8P7) in the crude extract of *Pseudonocardia* sp. C8 which were not present in the control SM medium. Isolation was carried out to obtain pure compounds and to characterize the structures.

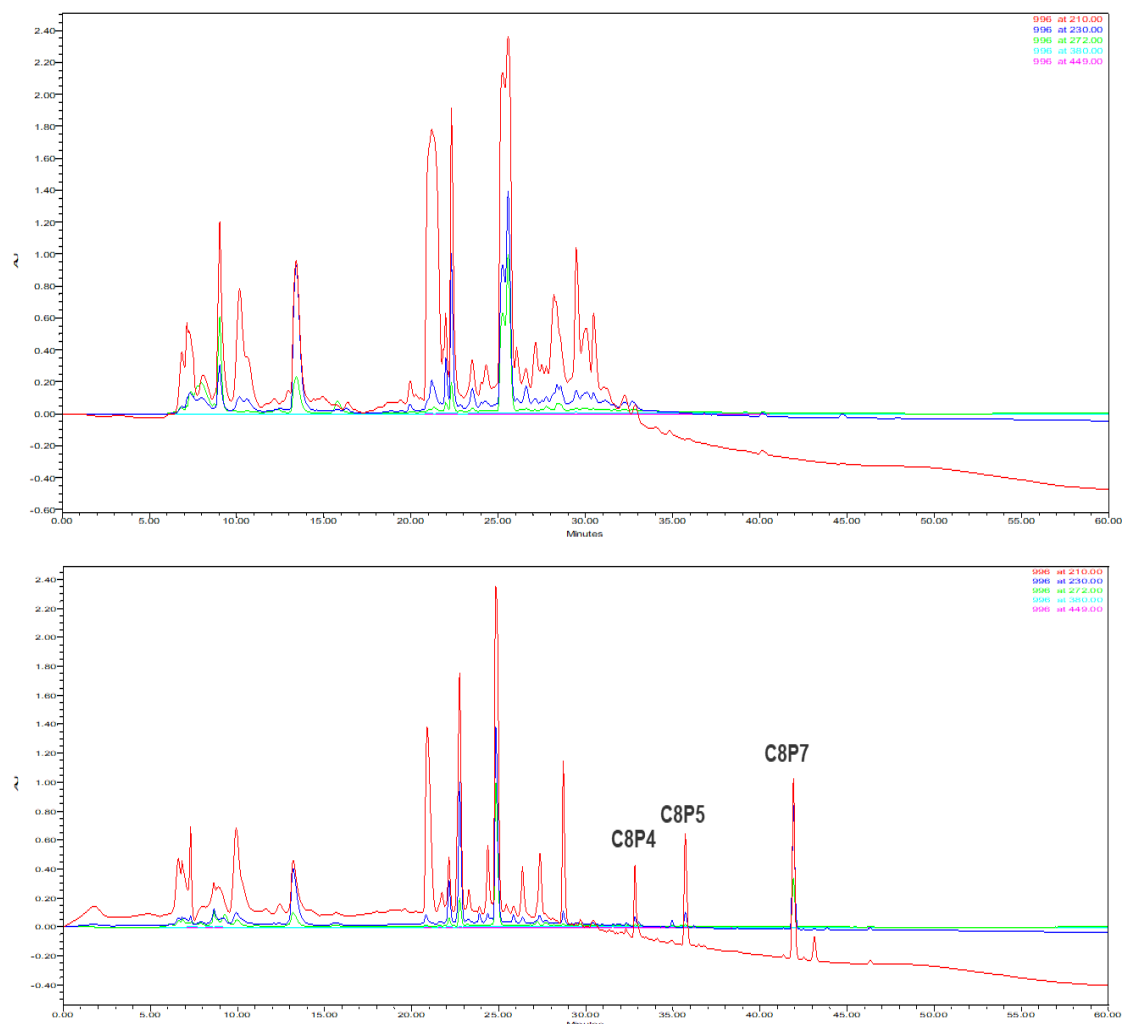
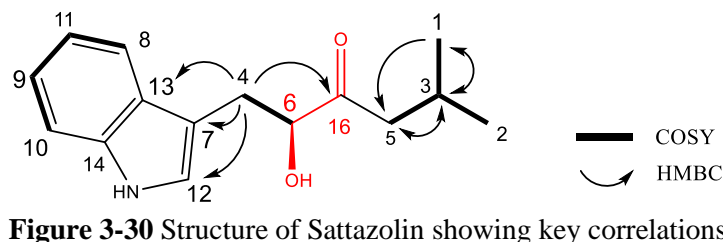


Figure 3-29 HPLC chromatogram showing media (top) and crude extract (bottom) chromatogram of *Pseudonocardia* sp. C8.

3.2.4.1 C8P7

Low resolution LC-MS were unable to confirm the exact mass for compound C8P7. Therefore, upon HPLC isolation the molecule was subjected to NMR measurement. It showed the presence of indole moiety and the aromatic amino acid, tryptophan. Later, further analysis using HRMS data yielded the protonated mass, $[M+H]^+=246.1484$ Da and molecular formula $C_{15}H_{19}NO_2$. It was finally identified by NMR as sattazolin^{141, 142} (see **Figure 3-30**, **Table 3-14**).



Sattabacins and sattazolin were firstly reported from *Bacillus* sp. strain B-60 by in mid 1990s. Later in 2014, *Xenorhabus boveni* was also found to produce similar structures. However, this time due to ease in sequencing technology, authors also found the BGC for production of these xenoclyoins. This work mainly pointed towards the involvement of ThDP-dependent acyloin-like enzymes (XclA) responsible for carrying acyloin condensation reaction for the formation of the product. Recently, sattazolins were recharacterized from anaerobic bacteria *Clostridium beijerinckii* from Hertweck laboratory. The thiamine diphosphate (ThDP)-dependent sattazolin producing synthase (Cbei2730) was also identified in silico and characterized both in vivo and in in vitro enzyme assays.¹⁴³ (see **Figure 3-31**)

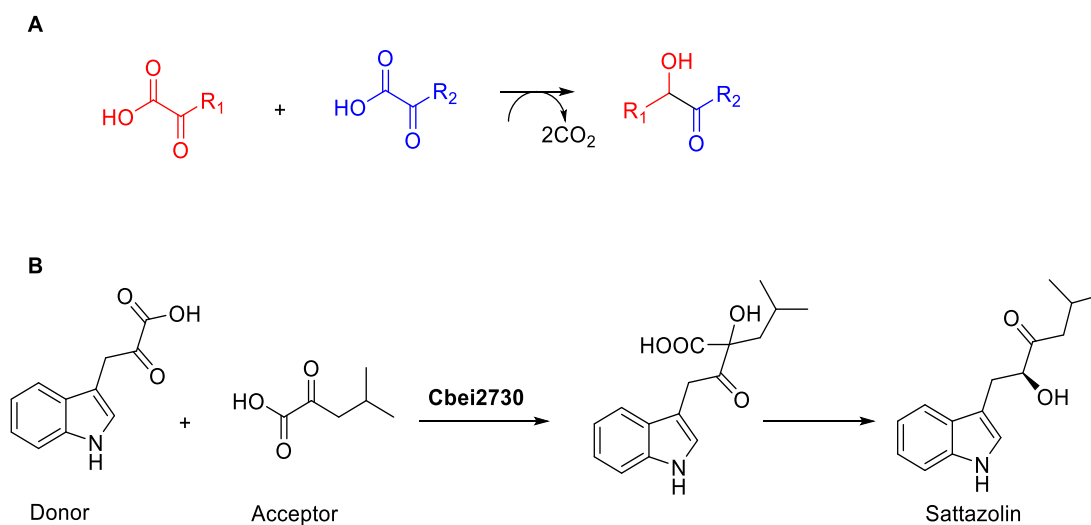


Figure 3-31 A) General scheme for acyloin condensation. B) Biosynthesis of sattazolin catalysed by enzyme Cbei2730 (ThDP-dependent)

Table 3-14 NMR data of C8P7 recorded in *d*₆-DMSO (400 MHz)

Position	$\delta_{C/N}^a$, mult.	δ_H^b (mult., J in Hz)	HMBC
1	22.4, CH ₃	0.78, (d, 6.9)	2, 3
2	22.5, CH ₃	0.80, (d, 6.9)	1, 3
3	23.3, CH	1.96, (m)	1,2
4	25.4, CH ₂	2.86 (dd, 14.8, 5.4) 3.01 (dd, 14.8, 7.6)	7, 12, 13
5	46.4, CH ₂	2.38, (t, 6.94)	1, 2, 3, 16
6	76.8, CH	4.16, (m)	
7	110.1, qC		
8	111.3, CH	7.31, (d, 8.1)	13, 9
9	118.2, CH	6.96, (t, 7.8, 0.7)	8, 13
10	118.5, CH	7.52, (d, 7.9)	7, 11, 14
11	120.8, CH	7.04, (t, 7.6, 1)	10, 14
12	123.7, CH	7.11, (s, 2.28)	7, 13, 14
13	127.4, qC		
14	136.0, qC		
15	212.9, qC, C=O		
16	OH -	5.39 (5.46)	
17	NH -	10.81 (s)	

3.2.4.2 C8P4 and C8P5

Along with sattazolin, two additional masses were detected during the LC-MS based screening. The molecular formula of these two metabolites, C8P4 and C8P5, were established as C₂₂H₃₄N₂O₆ (calculated for C₂₂H₃₅N₂O₆⁺: 423.2490, found: 423.2486) and C₂₂H₃₅N₂O₆P (calculated for C₂₂H₃₆N₂O₆P⁺: 503.2153, found: 503.2147), respectively, using positive ion ESI-HR-MS. SIRIUS 4, a java-based software for analysis of LC-MS/MS fragments predicted the loss of phosphorous with mass difference of 80 Da. This was further confirmed with ³¹P NMR signals at δ_P 0.47.

Meanwhile, similar results were also obtained by Dr. Junjing Jiao from our group when he was independently working with strain *Paenibacillus* sp. MW14, isolated from agricultural soil of

NPs from Nepalese actinomycetes

Iran. By comparison of ¹H-NMR and MS/MS data it was evident that both compounds were identical and also losing phosphate in similar manner. However, I was unable to reproduce the production of C8P4 and C8P5 leaving the speculations about the purity of strain itself. To further verify the data, we sequenced *Pseudonocardia* sp. C8 strain. However, BGC encoding for an amicoumacin type antibiotics were not observed. Therefore, further work in compound isolation and structure determination was carried out from another strain *Paenibacillus* sp. MW14 which will not be a part of this thesis.

3.3 Overall discussion and conclusion

This chapter mainly focuses on the strains that were isolated from various parts of Nepal. In search of novelty, parameters like altitude, habitat, temperature etc. were considered during our sampling. It has been proven that bioprospecting of environments with strong selecting factors has been an alternative strategy to enrich novelty in strains and subsequently in compounds as well.¹⁴⁴

In the past, most of the research conducted originally in Nepal was restricted to the screening level, both in terms of strains and their bioactivity. Previously, bacteria have been isolated from Khumbu region which has an altitude between 3000-6000 meters. However, scientific work has just been carried out on the crude level and no chemical analysis has ever been reported.^{145, 146} In many cases, the bacteria was identified based on their morphological data only.¹⁴⁷ In-depth chemical analysis has not been performed in the past due to a lack of scientific infrastructures and different priorities in research. In our work, we tried to address this gap. Therefore, this is a pioneer work which involves bioprospecting of the microbial diversity and their secondary metabolites. We have conducted a screening of bacteria from unique Nepalese ecological niches all over the country, which resulted in identification of novel actinobacteria along with the regular *Streptomyces*. Furthermore, we performed metabolomics-based workflow to annotate and characterize bioactive molecules.

In our first project, we have isolated mostly *Streptomyces*. Initial crude level activities were very convincing. Therefore, we prioritized them for our chemical analysis. Knowing the fact that the genus *Streptomyces* is ubiquitous and that it had been heavily explored all over the world, we set up a workflow where we could dereplicate the compounds just *in silico* without performing any isolation work. This technique not only saved us time but also brought some reliable results to present. Both, known and unknown compounds were identified in this process, mainly from two *Streptomyces* species. As a result, we were able to dereplicate three different classes of compounds along with their derivatives. LC-MS/MS techniques was effectively employed to

detect the subtle changes in the known structures eventually leading to new congeners from polyether ionophore family (**Figure 3-11**). To examine further, we closely analysed a nonactin cluster from *Streptomyces griseus*. In a study conducted by Cox and co-workers, the NonR protein was found to cleave at the ester linkage which not only generates open chain analogues but also results in the formation of such dimers (**Figure 3-32**). These dimers did not possess antimicrobial activities, supporting their hypothesis of bacterial resistance.¹⁰¹

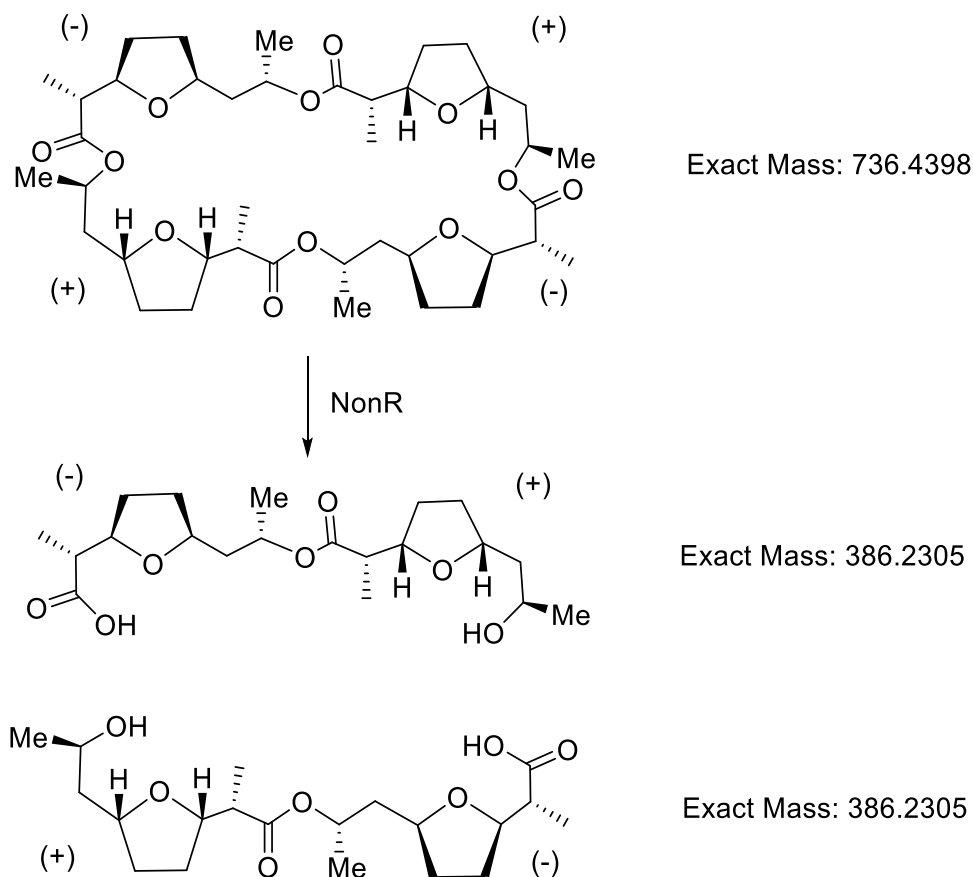
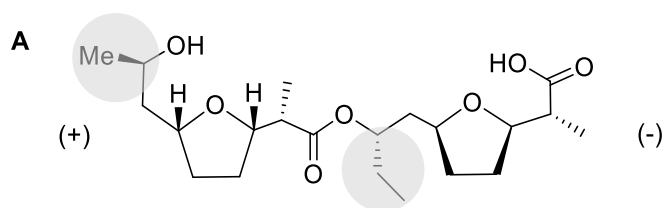
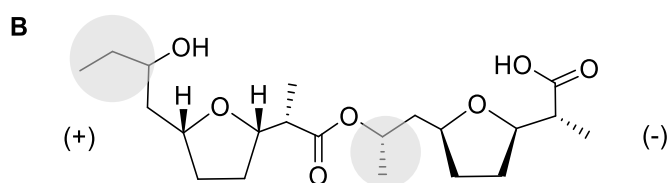


Figure 3-32 NonR-catalyzed stereoselective hydrolysis creating dimers.

Interestingly, another compound which has similar skeleton to that of nonactin dimers, bonactin, was found to possess antimicrobial activity¹⁴⁸ (see **Figure 3-33**). Mechanism of these different activities are not yet explained clearly although one could speculate that the changes in the side chain may have resulted in an altered activity. Further work on isolation of these new congeners is inevitable to shed more light on SAR study of these dimers.



NonR catalyzed dimer product of Dinactin
Exact Mass: 400.2461



Bonactin
Exact Mass: 400.2461

Figure 3-33 Comparison of A) NonR-catalyzed dinactin dimer and B) bonactin

In the context of these new dereplicated unsaturated congeners, we assume a reduction of the OH to a keto group on a (+) nonactin dimer resulting in a two Da less dimers. The mechanism of dimer formation would remain the same creating more dimeric compounds from other nonactin derivatives as well. However, without genomic information and further knockout experiments it is not possible to predict which of the enzymes is responsible for double bond insertion to these dimers. Another possibility for such double bond formation could be autocatalytic reactions which could result in unspecific double bond formation. Additionally, for stereochemistry determination isolation work is also a prerequisite which was not done in our studies.

In second part of this chapter, we focused more on obtaining novel genus or a species for our compound exploration. Several novel compounds in recent years are produced from rare actinomycetes. A lantibiotic, planosporicins are one of many examples that has been identified from the rare genus *Planomonospora*¹⁴⁹. In our second screening project we were able to isolate two rare actinomycetes. Upon initial chemical screening of *Nonomuraea* sp. C10, we characterized four compounds. However, all these derivatives did not possess antimicrobial, antifungal and proteasome activity (**Figure S58**). In addition, for brartemicin (compound 3), proteasome inhibition assay was conducted, showing no inhibition at all (data not presented). Similarly, from another strain (*Pseudonocardia* sp. C8), sattazolin was fully characterized. From our previous bioinformatic analysis, it became apparent that *Pseudonocardia* sp. C8 is a new species. However, taxonomic studies are inevitable to prove this finding. Meanwhile, the

NPs from Nepalese actinomycetes

OSMAC approach for finding a better cultivation medium can be continued. And, recently upon a closer look at the genomic data, we were able to identify several BGC's encoding unidentified RiPPs and lasso peptides. Expressing these gene cluster by changing the cultivation medium is one of the classic strategies however, it is still considered effective. Upon several such trials, we have eventually found a medium, modified TSB (mTSB), which can produce these large molecules in a higher amount. One litre of mTSB medium was extracted with n-butanol. Subsequently, the obtained crude extract was subjected to column chromatography to obtain five fractions. Fractions 1 and 2 thereof were found to contain m/z values as shown in **Figure 3-34**. Fragment analysis also indicated towards the presence of peptides.

In future, we aim to characterize these compounds along with genomic and chemical information. A powerful algorithm or a machine learning web based platform like RiPPMiner¹⁵⁰ or RODEO (<http://ripp.rodeo/advanced.html>) could be useful to analyse such class of compounds. In summary, with all these findings we conclude that actinomycetes have still a lot to offer. However, the approach of looking their genomes or metabolomes should be improved.

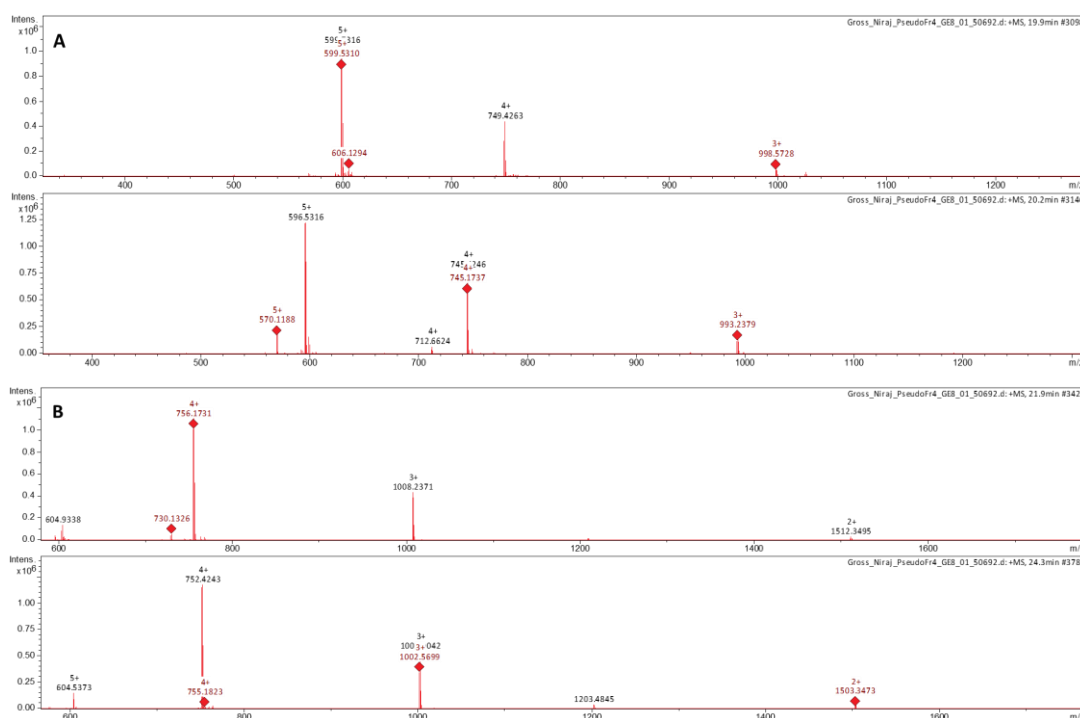


Figure 3-34 RiPP-derived masses detected in LC-MS analysis of fraction 1 from *Pseudonocardia* sp. C8 with fragment similarity. A) Type I B) Type II

In conclusion, our work with the Nepalese isolates gives a solid foundation in regard to how natural product isolation work can be carried out. In broader picture, it also highlights the importance of infrastructure and advantage of conducting collaborative research between low-income and developed countries.

Chapter 4. New nucleoside antibiotics from *Streptomyces* sp. SHP 22-7

4.1 Bioinformatic analysis of *Streptomyces* sp. SHP 22-7

AntiSMASH 4.0 was used for analysis of the genome sequence of *Streptomyces* sp. SHP 22-7. Twenty-three contigs coding for various compounds were predicted, one of which shows 70% similarity with the gene cluster coding for amicetin in *Streptomyces vinaceusdrappus* NRRL 2363 (Table 4-1), i.e. contig 10.1.

Table 4-1 Information of contigs coding for secondary metabolite in *Streptomyces* sp. strain SHP 22-7

Contig	Type	Most similar known cluster	Similarity
1.1	Terpene	Albaflavenone	100%
1.2	T2PKS	Spore pigment	66%
2.1	Furan	Methylenomycin	9%
3.1	Melanin	Istamycin	4e%
4.1	Ectoine	Ectoine	100%
5.1	NRPS	Phosphonoglycans	5%
7.1	Bacteriocin	Unknown	/
7.2	Terpene	Geosmin	100%
10.1	NRPS-like	Amicetin	70%
10.2	NRPS	Calcium-dependent antibiotics	37%
11.1	T3PKS	Herboxidiene	8%
13.1	Siderophore	Desferrioxamine	66%
17.1	Indole	Terfestatin	19%
17.2	Terpene	Carotenoid	36%
19.1	T2PKS	Fluostatins M-Q	62%
20.1	Terpene	Hopene	69%
21.1	Lanthipeptide	Unknown	/
22.1	NRPS	Antimycin	93%
24.1	NRPS	Lipopeptide 8D1-1/1-2	29%
24.2	Siderophore	Grincamycin	5%
42.1	hgIE-KS	Sanglifehrin A	13%
44.1	Bacteriocin	Infomatipeptin	42%
48.1	T1PKS	Ibomycin	15%

On contig 10.1, the functions of biosynthesis gene cluster in *Streptomyces* sp. SHP 22-7 were compared with BGCs in *Streptomyces vinaceusdrappus* NRRL 2363. 16 out of 21 genes were highly homologous (identity > 98.9%) to the genes of the amicetin BGC in *Streptomyces* sp. SHP 22-7 employing blastn and blastp. Based on sequences and size of genes, three pairs of

genes were matched for three gene clusters of *Streptomyces vinaceusdrappus* NRRL 2363, respectively. They are highlighted in red (**Table 4-2**). Due to lack of overlap between most of reads on contig 10.1 and contig 10.2, the remaining five amicetin BGCs in *Streptomyces vinaceusdrappus* NRRL 2363 were not detected (**Figure 4-1**). Therefore, only 70% similarity was observed.

Combing the preliminary data from LC-MS and genomic information, it was evident that plicacetin and other nucleoside antibiotics were produced by *Streptomyces* sp. SHP 22-7.

Table 4-2 Corresponding BGCs coding for amicetin between *Streptomyces* sp. strain SHP 22-7 and *Streptomyces vinaceusdrappus* NRRL 2363

Gene of SHP 22-7	Size (nt)	Putative function	Identity (%)	Gene of from NRRL(2363)	Size (nt)
<i>D3C59_17715</i>	735	NAD-dependent epimerase/dehydratase protein	99.86	<i>amiU</i>	1065
<i>D3C59_17720</i>	528	NAD-dependent epimerase/dehydratase protein	99.81		
<i>D3C59_17725</i>	864	Nonribosomal peptide synthetase (A-PCP)	99.88	<i>amiT</i>	2466
<i>D3C59_17730</i>	1317	serine hydroxymethyltransferase	99.92	<i>amiS</i>	1359
<i>D3C59_17735</i>	810	Malonyl CoA-acyl carrier protein transacylase	99.88	<i>amiR</i>	900
<i>D3C59_17740</i>	1192	MFS transporter	99.66	<i>amiQ</i>	1194
<i>D3C59_17745</i>	633	TetR family transcriptional regulator	98.90	<i>amiP</i>	639
<i>D3C59_17750</i>	1679	ABC transporter ATP-binding protein	99.94	<i>amiO</i>	1680
<i>D3C59_17755</i>	1318	lipopolysaccharide biosynthesis protein RfbH	99.85	<i>amiN</i>	1320
<i>D3C59_17760</i>	2034	aminodeoxychorismate synthase component I	99.95	<i>amiM</i>	2034
<i>D3C59_17765</i>	543	Benzoate-CoA synthase	99.82	<i>amiL</i>	1497
<i>D3C59_17770</i>	942	Benzoate-CoA synthase	99.89		

New nucleoside antibiotics from *Streptomyces* sp. SHP 22-7

<i>D3C59_17775</i>	957	NAD(P)-dependent oxidoreductase	99.90	<i>amiK</i>	948
<i>D3C59_17780</i>	405	Cytosylglucuronic acid synthase/glycosyltransferase	100	<i>amiJ</i>	1212
<i>D3C59_17785</i>	636	Cytosylglucuronic acid synthase/glycosyltransferase	99.69		
<i>D3C59_17790</i>	549	(Deoxy)cytidine deoxyribosyltransferase	100	<i>amiI</i>	549
<i>D3C59_17795</i>	1254	methyltransferase domain-containing protein	100	<i>amiH</i>	750
<i>D3C59_17800</i>	1488	glycosyltransferase family protein	1 100	<i>amiG</i>	1488
<i>D3C59_17805</i>	1359	GNAT family N-acetyltransferase	99.93	<i>amiF</i>	1407
/	/	Glucose-1-phosphate thymidyltransferase	/	<i>amiE</i>	774
/	/	NDP-hexose 3-ketoreductase	/	<i>amiD</i>	969
/	/	NDP-hexose 2,3-dehydratase	/	<i>amiC</i>	1377
/	/	NDP-sugar aminotransferase	/	<i>amiB</i>	1119
/	/	Putative-4-amino-4-deoxychorismate lyase	/	<i>amiA</i>	795

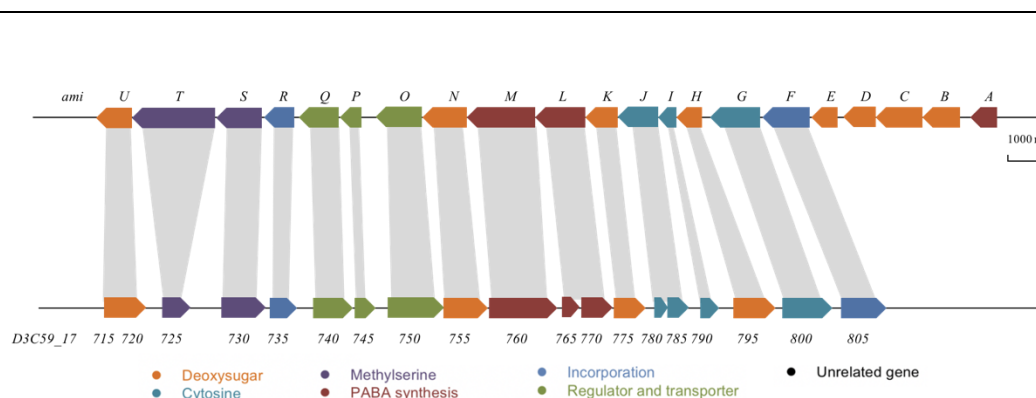


Figure 4-1 Comparison of gene clusters coding for amicetin and transport-related proteins (green) between *Streptomyces vinaceusdrappus* NRRL 2363 (top panel) and *Streptomyces* sp. strain SHP 22-7 (bottom panel)

4.2 HR-LC-MS/MS based dereplication and target identification

In order to sort out knowns from unknowns, a sensitive analytical technique HR-LC-MS/MS was chosen. The crude extracts from the whole fermentation broth of *Streptomyces* sp. SHP 22-7 were subjected to LC-MS analysis as described in the method section 2.3.8.2, and later analysed in combination with available public databases and in silico platforms. Several new masses within the range of 480-688 Da were discovered, which were tentatively assigned as plicacetin derivatives. On meticulous analysis of MS/MS data, two characteristic fragments were found to be consistent for most of the target compounds. Fragments $[M+H]^+=174$ and $[M+H]^+=288$, which resulted from the monosaccharide and disaccharide moiety (**Figure 4-2**). Throughout this study these conserved fragments were set as the foundation to track down new plicacetin/amicetin derivatives.

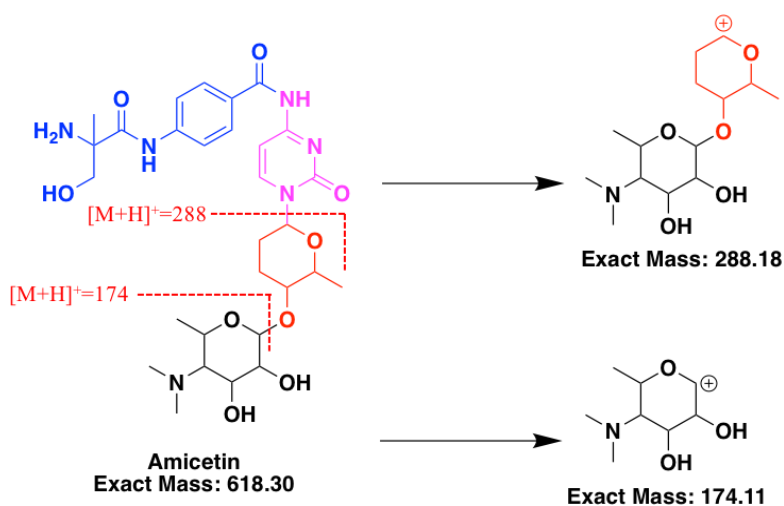


Figure 4-2 The structure of fragments $[M+H]^+=174$ and $[M+H]^+=288$

At first, extracted ion chromatograms (EIC) were generated to visualize all compounds that showed a fragmentation behaviour like plicacetin or amicetin. (see **Figure 4-3**). This strategy led to the detection of 21 derivatives, 12 thereof were fully characterized by NMR measurements, while the remaining 9 were partially characterized via MS² fragmentation studies and spectroscopic data. Summary of all characterized compound is presented in **Table 4-3**.

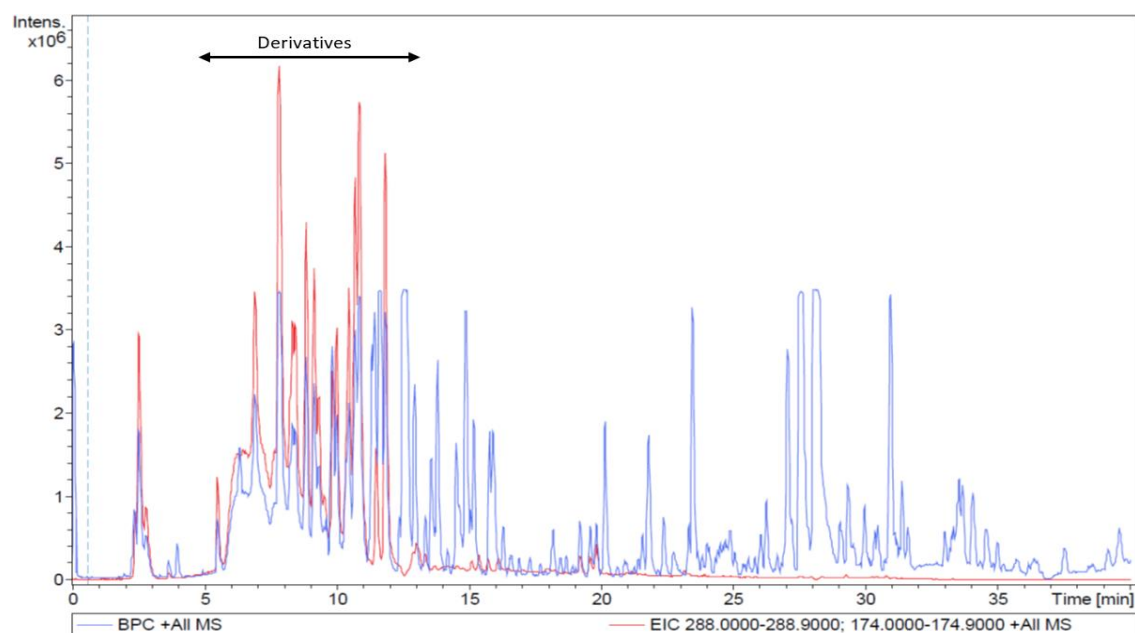
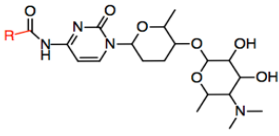
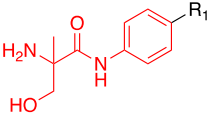
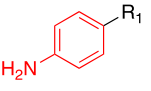
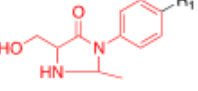
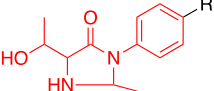
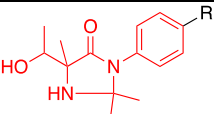
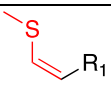
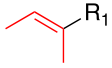
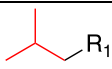
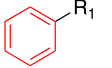
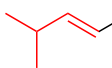
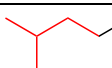
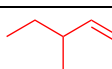
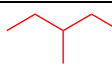
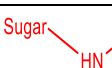
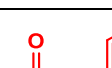
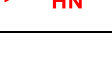
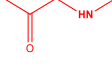
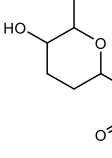
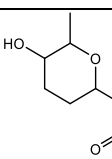


Figure 4-3 Extraction Ion Chromatogram (EIC) of $[M+H]^+=174$ & $[M+H]^+=288$ (Red) and Base Peak Chromatogram (blue) of the crude extract from *Streptomyces* sp. SHP-22-7

Table 4-3 Nucleoside antibiotics characterized in this study

S.N	$[M+H]^+$	Predicted molecular formula	Error in ppm/rdB	 Predicted Structure R-R₁	Compound annotated (CAS Nr.)
1	619.3083	C ₂₉ H ₄₂ N ₆ O ₉	0.4/12		Amicetin [#]
2	518.2602	C ₂₅ H ₃₅ N ₅ O ₇	1.2/11		Plicacetin*
3	631.3094	C ₃₀ H ₄₂ N ₆ O ₉	1.3/13		Streptcytosine A [#]
4	645.3252	C ₃₁ H ₄₄ N ₆ O ₉	1.5/13		New [#]
5	673.3556	C ₃₃ H ₄₈ N ₆ O ₉	0.1/13		New [#]
6	499.2217	C ₂₂ H ₃₄ N ₄ O ₇ S	0.5/8		CAS 1461743-58-7* Compound I

New nucleoside antibiotics from *Streptomyces* sp. SHP 22-7

7	481.2655	C ₂₃ H ₃₆ N ₄ O ₇	0.0/8		40551-I* Compound IV
8	483.2813	C ₂₃ H ₃₈ N ₄ O ₇	0.0/7		40551-K#
9	503.2497	C ₂₅ H ₃₄ N ₄ O ₇	0.4/11		CAS 51693-93-7* Compound VI
10	495.2816	C ₂₄ H ₃₈ N ₄ O ₇	0.1/8		40551-L#
11	497.2971	C ₂₄ H ₄₀ N ₄ O ₇	0.3/7		Cytosaminomycin E#
12	509.2972	C ₂₅ H ₄₀ N ₄ O ₇	0.5/8		40551-G#
13	511.3127	C ₂₅ H ₄₂ N ₄ O ₇	0.3/7		New* Compound XI
14	504.2453	C ₂₅ H ₃₃ N ₅ O ₇	0.1/11		Norplicacetin#
15	560.2720	C ₂₇ H ₃₇ N ₅ O ₈	0.9/12		New#
16	574.2868	C ₂₈ H ₃₉ N ₅ O ₈	0.5/12		New#
17	588.3035	C ₂₉ H ₄₁ N ₅ O ₈	1.8/12		New* Compound XII
Small streptocytosine molecules					
	[M+H]⁺	Predicted molecular formula	Error in ppm/rdb	Predicted Structure	Compound annotated
18	310.1766	C ₁₅ H ₂₃ N ₃ O ₄	1.2/6		New*
19	308.1610	C ₁₅ H ₂₁ N ₃ O ₄	1.6/7		Streptocytosine*

20	310.1733	C ₁₅ H ₂₃ N ₃ O ₄	1.0/6		Streptocytosine*
21	345.1604	C ₁₈ H ₂₂ N ₃ O ₄	0.3/10		Streptocytosine N*

#LC-MS/MS based characterization

*NMR and MS/MS based characterization

New* = New compound

New# = New compound

4.3 Growth characteristics

The basic idea behind measuring a growth curve was to monitor various growth phases and cell biomass measurements. Alongside, we also aim to obtain the production profile of secondary metabolites in order to conduct feeding experiments during precursor directed biosynthesis studies. A growth curve based on wet and dry weight at 30 continuous time points were measured, respectively (**Figure 4-4**).

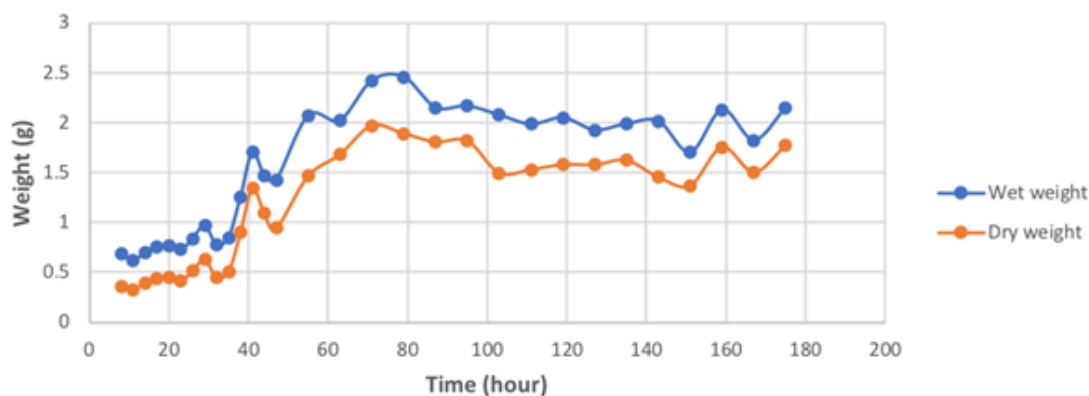


Figure 4-4 The growth curve of *Streptomyces* sp. strain SHP 22-7 in NL-300 media

We found that strain SHP 22-7 grown in NL-300, exhibits a lag phase until the first 32h. During the log phase there is spike in cell biomass and also the major production of secondary metabolites occurs (**Figure 4-5**). This finding sets the foundation for PDB studies and semi-quantification of metabolites produced.

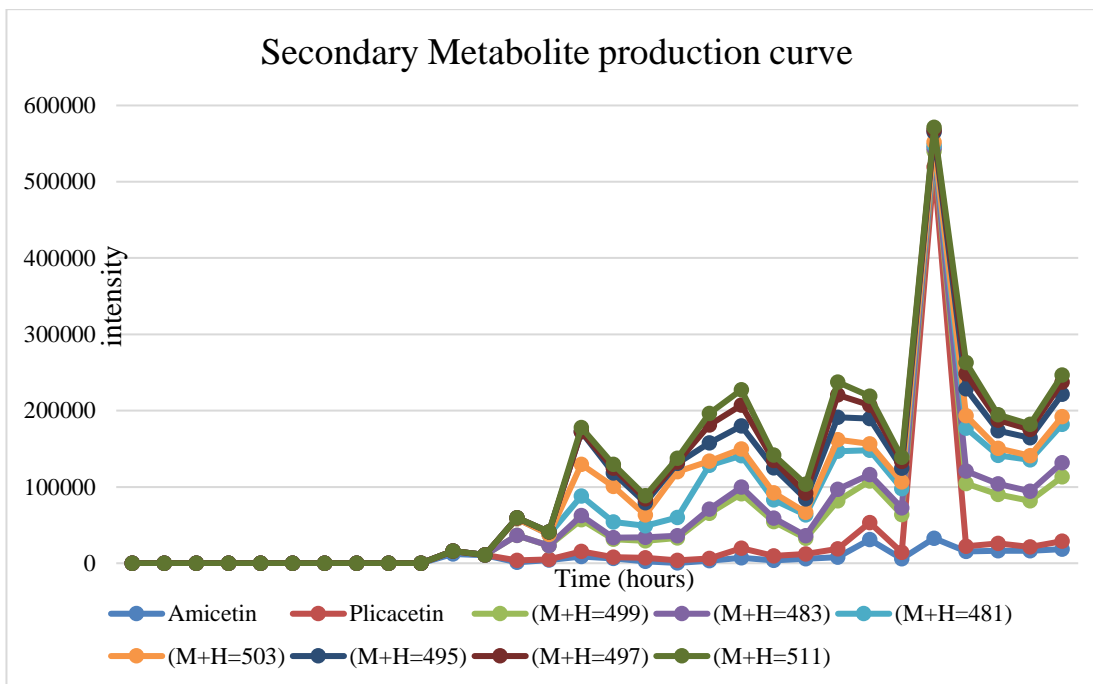


Figure 4-5 The plicacetin derivatives production profile at 30 time points

4.4 Fractionation and isolation of compounds

Due to the use of a rich medium for production and the application of a whole broth extraction, a complex crude extract was obtained upon ethyl acetate extraction. Diluted crude sample was subjected to LC-MS which indicated presence of similar compounds. Fractionation was the next step. A general fractionation scheme is shown in **Figure 4-6**. Once the fractions were obtained, they were subjected to LC-MS to ensure the production of the target compounds. Depending on the complexity of fraction obtained, sub fractions were subsequently generated with semi-preparative HPLC. As most of the plicacetin derivatives were unstable, a buffer solution was used as mobile phase during the HPLC-based isolations.

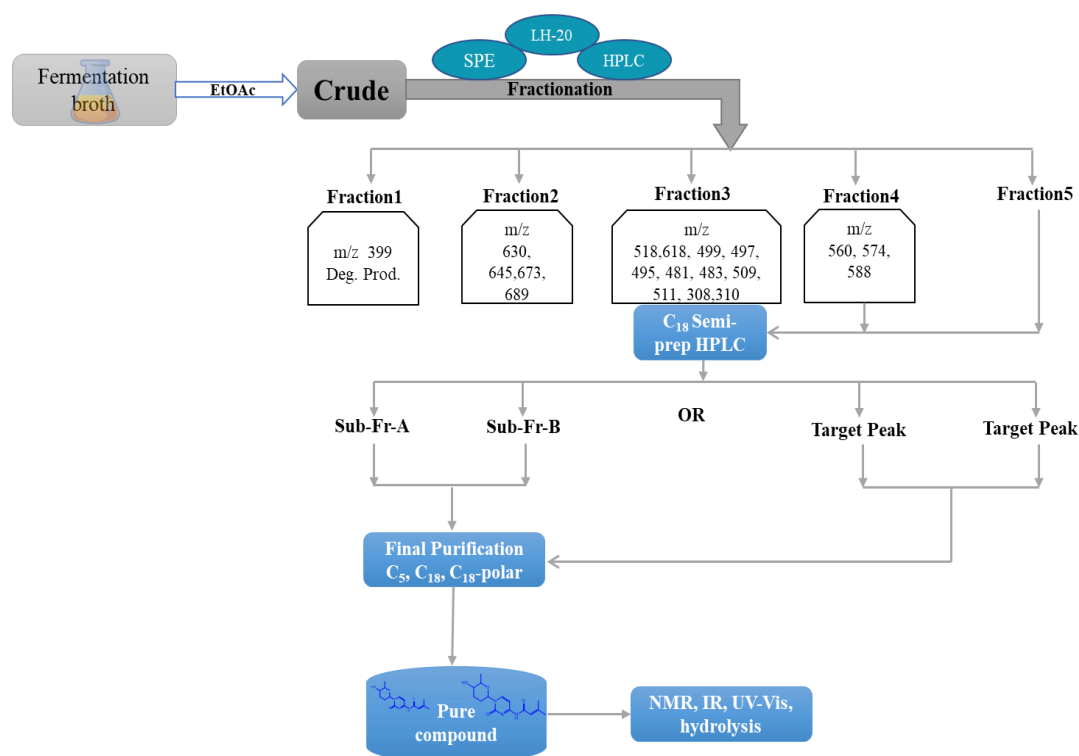


Figure 4-6 Fractionation scheme

However, this isolation scheme had to be adapted from batch to batch. Each time the crude extract batch was received, there was a huge fluctuation concerning the target compound (plicacetin/amicetin class) and their amounts. In majority batches, fraction 3 was found to contain most of the masses related to the plicacetin type fragmentation which is presented in figure above.

High performance liquid chromatography was used throughout the isolation process. Methanol and 0.1% trifluoroacetic acid (TFA) initially served as mobile phase. However, during the purification process, the instability of compound I was increasingly observed (**Figure 4-7**). Such instability was also observed during plicacetin purification (data not shown). Later, it came to our notice that the decomposition is located in the amide bond between cytosine and its side chain which was confirmed with LC-MS analysis after isolation of degraded compounds. Eventually, we speculated that those decompositions are due to the usage of acid mobile phases. Hence, a 1M phosphate buffer solution was employed in subsequent isolation procedures to increase the stability. By this way, the stability was improved tremendously, however additional desalting process of sample was necessary prior to activity analysis as it requires the accurate weight of samples. Occasionally, lipid contamination was observed in purified compounds. This component was then removed with an analytical column consisting of a shorter carbon chain length (C5). This not only removed lipids, but also allowed a better resolution of the target peak.

The HPLC chromatogram obtained by using a polar RP column for fraction 3 is shown in **Figure 4-8**. Masses of each peak were confirmed via LR-LCMS/MS. One advantage that allowed us to stick to the specific m/z value throughout several batches was ionization. Due to the good positive mode ionization, it was possible to trace the compounds. Once the target peak was collected by HPLC, it served later as a standard peak in order to trace the target compound in other batches. A list of m/z values corresponding to peaks in the HPLC along with their designation is shown in **Table 4-4**.

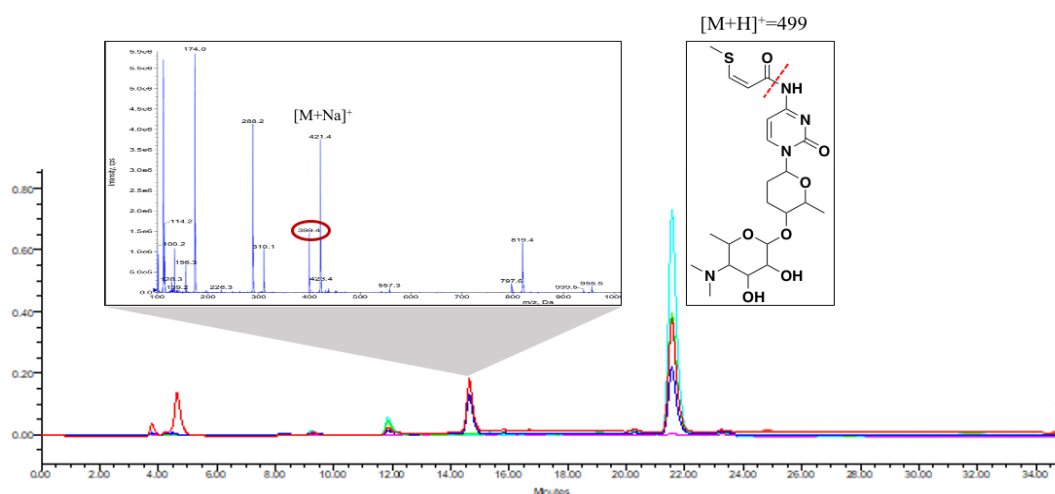


Figure 4-7 HPLC chromatogram showing the pure compound I (21 min) and its decomposition product (14.5 min)

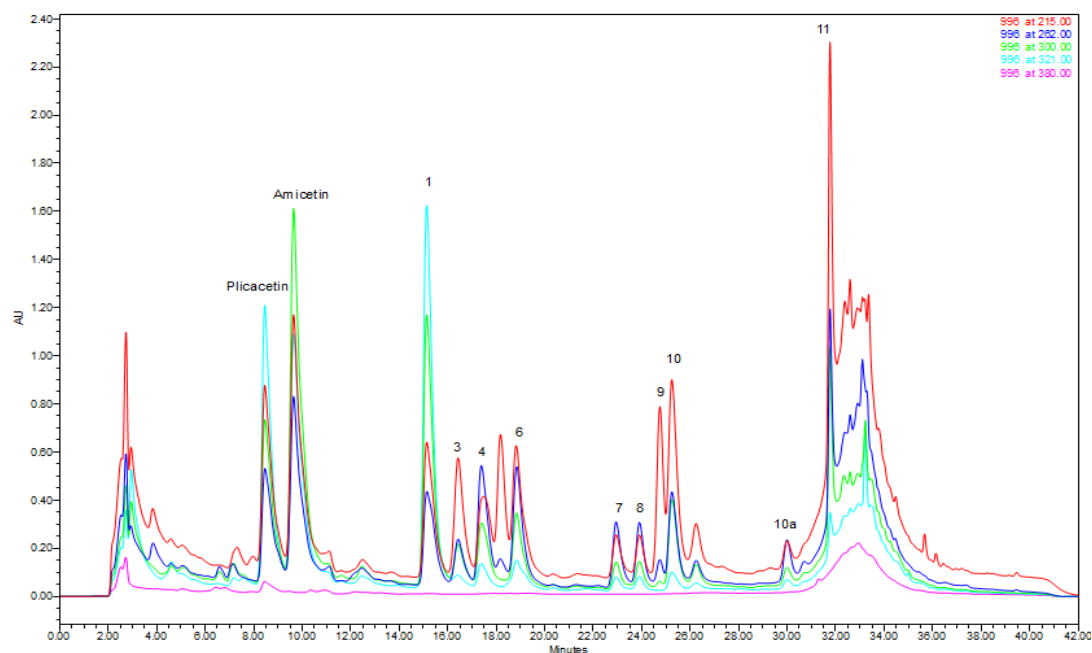


Figure 4-8 HPLC chromatogram of Fraction 3 showing plicacetin, amicetin and other derivatives.

Table 4-4 Peaks and their designated masses (also refer to Figure 4-8). Gray highlighted are all fully characterized compound from this list

Peak Nr.	Compound name	Retention time in HPLC(min)	[M+H] ⁺
	Plicacetin	8.1	518
	Amicetin	10	618
1	I	14.5	499
3	III	15.9	483
4	IV	16.9	481
6	VI	18	503
7	VII	22	495
8	VIII	23	495
9	IX	24	497
10	X	24.5	497
10a	Xa	29	509
11	XI	31.5	511

4.5 Structure elucidation of Plicacetin and derivatives

As listed above in **Table 4-3**, in total 21 compounds were characterized in this study. 10 of which were fully characterized by NMR and tandem MS data, while the rest were characterized by partial NMR data and/or MS/MS data. These characterizations were confirmed by available literatures and databases. Novelty of new compounds were confirmed by searching the chemical entity in Scifinder database (CAS). Supportive evidence was collected by matching fragments *in silico* platforms like SIRIUS 4 and Metaboscope 3.0 (see method section **2.3.8.4** and **2.3.8.5**).

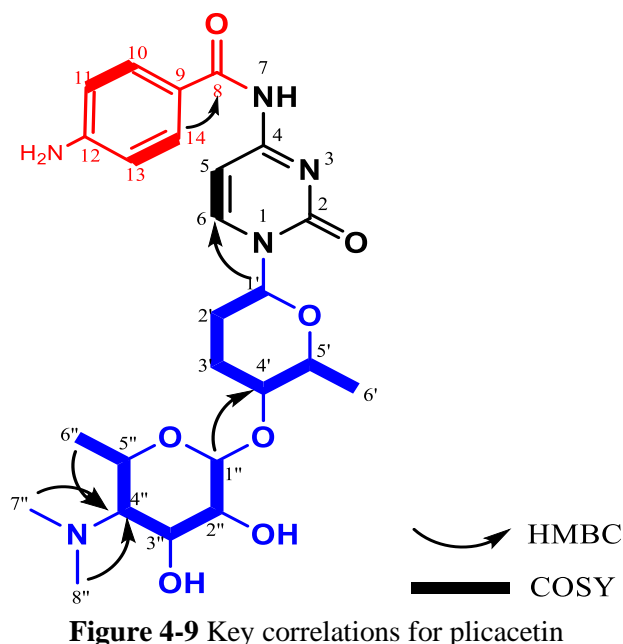
4.5.1 Plicacetin/Amicetin

Previous overexpression studies from Mast and coworkers had also revealed the presence of amicetin from the culture broth of *Streptomyces* sp. SHP-22-7^{32, 151}. To further confirm the known substances and to characterize the unknowns, isolation and structure elucidation was carried out in our group. Due to presence of strong chromophores (cytosine moiety, benzene ring system) in the structure, these compounds were well detected by a HPLC PDA detector. In most of the batches, production of plicacetin and amicetin was consistent. However, regarding the remaining derivatives this was not the case. Initially, plicacetin was isolated using an acidic

system (MeOH and 0.1% TFA/Water), but due to degradation problems it was soon shifted to buffer system. This seemed to slow down the degradation process but was unable to cease it. To confirm the structure of the purified plicacetin, we performed series of 1D and 2D NMR experiments including ^1H -NMR, ^{13}C -NMR, DEPT 135 NMR and 2D NMR, such as ^1H - ^1H COSY, ^1H - ^{13}C HSQC and ^1H - ^{13}C HMBC spectrum. These data were in agreement with published literatures values^{152, 153}. The NMR data obtained are presented in **Table 4-5** along with the structure of plicacetin (**Figure 4-9**).

Table 4-5 NMR data of plicacetin recorded in d_4 -MeOD (400 MHz)

Unit	Position	$\delta_{\text{C/N}}$ in ppm (mult.)	δ_{H} in ppm (mult., J in Hz)
Aromatic region	C=O	168.6, qC	-
	C10, C14	131.5, 2xCH	7.76 (d, 8.7, 2H)
	C11, C13	114.4, 2xCH	6.69 (d, 8.7, 2H)
	C12	155.0, qC	-
	C9	121.0, qC	-
Cytosine(black)	C=O	157.4, qC	-
	C4	164.8, qC	-
	C5	98.7, CH	7.59 (d, 7.4, 1H)
	C6	145.8, CH	8.11 (d, 7.4, 1H)
Sugar 1	C1'	84.6, CH	5.76 (dd, 10.8, 2.3, 1H)
	C2'	31.0, CH ₂	2.15 (m, 1H), 1.65 (m, 1H)
	C3'	27.7, CH ₂	2.38 (m, 1H), 1.65 (m, 1H)
	C4'	75.0, CH	3.41 (m, 1H)
	C5'	78.4, CH	3.75 (m, 1H)
	C6'	19.1 CH ₃	1.36 (d, 6.4, 3H)
Sugar 2	C1''	96.1, CH	4.92 (d, 3.7, 1H)
	C2''	74.6, CH	3.41 (m, 1H) overlap
	C3''	70.2, CH	3.87 (t, 9.6, 1H)
	C4''	71.8, CH	2.13 (t, 10.1, 1H) overlap
	C5''	67.4, CH	3.83 (m, 1H)
	C6''	19.6, CH ₃	1.24 (d, 6.4, 3H)
	C7''	42.3, CH ₃	2.47 (s, 3H)
	C8''	42.3, CH ₃	2.47 (s, 3H)



As plicacetin and amicetin were known compounds and described in several literature reports, we have collected and characterized it to function as a standard. Amicetin was characterized based on MS data and a milligram was collected for only activity analysis.

4.5.2 Compound I

HR/MS was conducted firstly, to identify the predicted molecular formula (**Figure 4-10**). Compound I exhibited the molecular formula $C_{22}H_{34}N_4O_7S$, with $[M+H]^+ = 498.2215$.

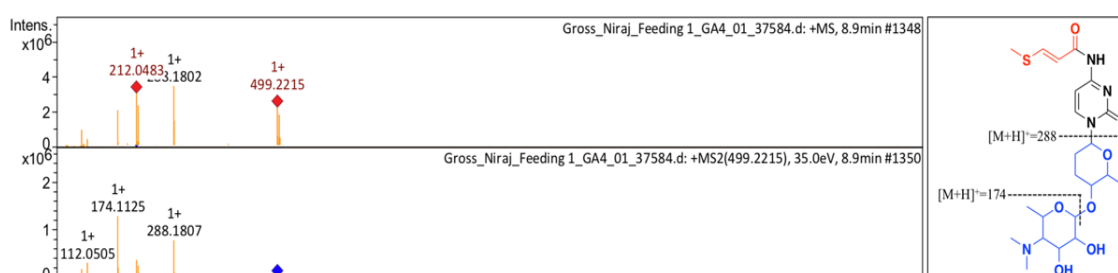


Figure 4-10 HRMS spectrum of compound I in positive mode

To confirm the structure of the pure compound I, 1D NMR including 1H -NMR, ^{13}C -NMR, DEPT 135 NMR and 2D NMR experiments, such as 1H - 1H COSY, 1H - ^{13}C HSQC and 1H - ^{13}C HMBC spectra were measured in d_6 -DMSO.

In the 1H -NMR spectrum, two separate signals located at chemical shift δ_H 7.32 and 8.11 with a doublet shape was observed representing two vinyl protons of cytosine. Due to an electron withdrawing effect of the nitrogen and sulphur atoms, the N-methyl group on amosamine

moiety demonstrated a higher chemical shift than the methyl directly attached on sugar moieties. Therefore, signals at δ_{H} 2.36 and 2.37 belonged to nitrogen and sulphur methyl, while another two signals at δ_{H} 1.15 and 1.24 were produced by methyl groups directly connected with sugar ring. Additionally, based on coupling the constant $J = 14.7$ Hz between δ_{H} 7.82 and 6.22, the double bond between C-9 and C-10 is identified as an interchangeable cis/trans configuration with a 1:3 ratio. This was confirmed also with HPLC by re-chromatography of the pure compound. The occurrence of two peaks at different t_{R} confirmed that the structure possessed a interchangeable cis/trans bond (data not shown).

The ^{13}C -NMR spectrum implied, that compound I contains two carbonyl groups located at δ_{C} 163.6 and 154.0 and three regular methyl groups ranging from 14.0 to 20.0 ppm. Since the height of signal at δ_{C} 42.2 is higher than the remaining signals, it was identified as two methyl groups attached to nitrogen. Interpretation of the DEPT 135 NMR experiment showed that, altogether twelve tertiary carbons and two secondary carbons belonged to the sugar moiety. Furthermore, the assignments of carbons and its matching hydrogens were confirmed with the aid of ^1H - ^{13}C HSQC data.

By a combination of resonances in ^1H - ^{13}C HMBC, the correlations between hydrogens and carbons over three bonds could be observed. Key correlations of the ^1H - ^{13}C HMBC spectrum are marked with arrows (**Figure 4-11**).

The ^1H - ^1H COSY spectrum illustrated several correlations between hydrogens and provided additional evidence to confirm the structure, such as the interaction between H-5' and H-6'. The remaining correlations of the ^1H - ^1H COSY spectrum are highlighted with bold bonds (**Figure 4-11**).

Combined with all data from HR/MS and NMR experiments, structure of compound I was able to be confirmed as presented, which comprises a disaccharide moiety with a cytosine moiety which is in turn bonded to a side chain including a sulphur-methyl. The exact NMR data are summarized in following table (**Table 4-6**).

After confirmation of the structure of compound I, SciFinder was used to ascertain its novelty. Unfortunately, compound I was isolated from marine *Actinomycetes* strain PM0895117/MTCC 5675 in 2013 by Kate and co-workers from India, which displayed anti-cancer activity¹⁵⁴.

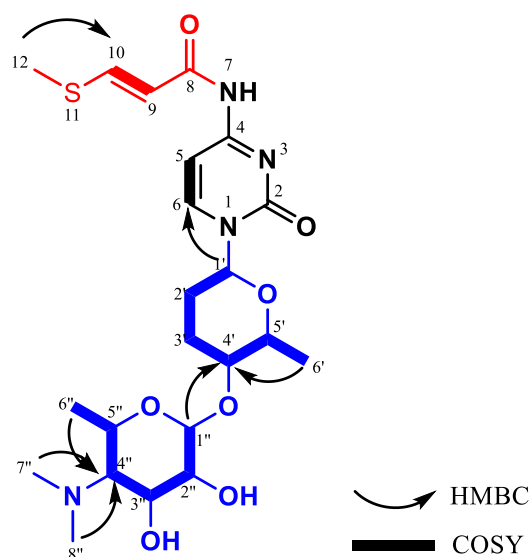


Figure 4-11 Key correlations of compound I

Table 4-6 NMR data of compound I recorded in d_6 -DMSO (400 MHz)

Unit	Position	δ_{CN} in ppm (mult.)	δ_H in ppm (mult., J in Hz)
Side Chain	C=O	163.6, qC	-
	C9	115.8, CH	6.22 (t, 5.0, 1H)
	C10	146.9, CH	7.82 (d, 14.7, 1H)
	-SCH ₃	14.1, CH ₃	2.36 (s, 3H)
Cytosine	C=O	153.9, qC	-
	C4	162.8, qC	-
	C5	95.9 CH	7.32 (d, 7.4, 1H)
	C6	145.5, CH	8.11 (d, 7.4, 1H)
Sugar 1	C1'	82.2, CH	5.69 (dd, 10.8, 2.3, 1H)
	C2'	29.5, CH ₂	1.93 (m, 1H), 1.70 (m, 1H)
	C3'	26.1, CH ₂	2.30 (m, 1H), 1.4 (m, 1H)
	C4'	73.1, CH	3.25 (m, 1H) overlap
	C5'	76.4, CH	3.63 (m, 1H) overlap
	C6'	18.5 CH ₃	1.24 (d, 3.7, 3H)
Sugar 2	C1''	94.7, CH	4.79 (d, 3.7, 1H)
	C2''	72.9, CH	3.25 (m, 1H) overlap
	C3''	68.3, CH	3.76 (t, 9.6, 1H)
	C4''	70.1, CH	2.00 (t, 10.1, 1H)
	C5''	65.7, CH	3.61 (m, 1H) overlap
	C6''	19.2, CH ₃	1.15 (d, 6.1, 3H)
	C7''	42.2, CH ₃	2.37 (s, 3H)
	C8''	42.2, CH ₃	2.37 (s, 3H)

4.5.3 Compound IV

The HR/MS measurement in positive mode indicated that compound IV has $[M+H]^+ = 481.2657$ with a molecular formula predicted as $C_{23}H_{36}N_4O_7$ ($\Delta = 0.0$ ppm). Ring double bond equivalency of 8 was predicted which was in agreement with our structure prediction. The fractions were obtained from several crude batches and their fractionation via the scheme presented in **Figure 4-6**. Four milligram of pure compound was obtained from HPLC isolations. These samples were subjected to 1D and 2D NMR measurements giving us the final structure as shown in **Figure 4-13**. Most of the structural entity shared the similarity to already annotated plicacetin and compound I, therefore similar values and correlation were obtained. Except for the side chain (red in **Figure 4-12**) which was assured by HMBC measurements shown below

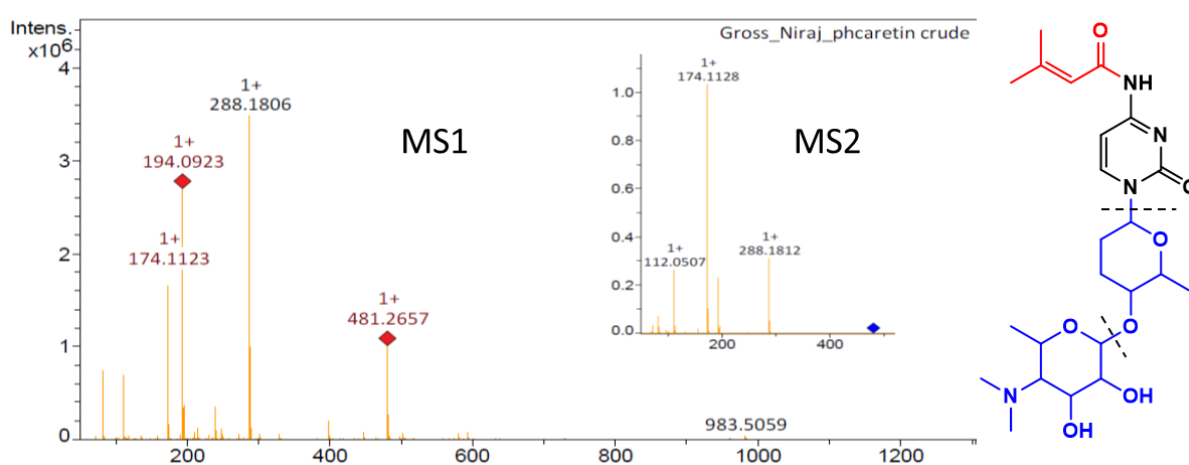


Figure 4-12 Figure showing MS¹ and MS² data for compound IV along with structure.

Table 4-7 NMR data of compound I recorded in *d*₄-MeOD (400 MHz)

Unit	Position	$\delta_{C/N}$ in ppm (mult.)	δ_H in ppm (mult., <i>J</i> in Hz)
Side Chain (red)	C=O	164.8, qC	-
	C9	118.8, CH	5.94 (t, 1.2, 1H)
	C10	159.5, qC	-
	C11	20.5, CH ₃	2.22 (s, 3H)
	C12	27.7, CH ₃ overlap	1.96 (s, 3H)
Cytosine (black)	C=O	157.4, qC	-
	C4	164.8, qC	-
	C5	98.1 CH	7.55 (d, 7.6, 1H)
	C6	145.5, CH	8.09 (d, 7.6, 1H)
Sugar 1	C1'	84.6, CH	5.76 (dd, 10.8, 2.3, 1H)
	C2'	31.0, CH ₂	2.13 (m, 1H), 1.66 (m, 1H)
	C3'	27.7, CH ₂	2.38 (m, 1H), 1.65 (m, 1H)
	C4'	75.0, CH	3.41 (m, 1H)

	C5'	78.4, CH	3.75 (m, 1H)
	C6'	19.1 CH ₃	1.36 (d, 3.7, 3H)
Sugar 2	C1''	96.1, CH	4.92 (d, 3.7, 1H)
	C2''	74.6, CH	3.41 (m, 1H) overlap
	C3''	70.2, CH	3.87 (t, 9.6, 1H)
	C4''	71.8, CH	2.13 (t, 10.1, 1H) overlap
	C5''	67.4, CH	3.83 (m, 1H)
	C6''	19.6, CH ₃	1.24 (d, 6.1, 3H)
	C7''	42.3, CH ₃	2.46 (s, 3H)
	C8''	42.2, CH ₃	2.46 (s, 3H)

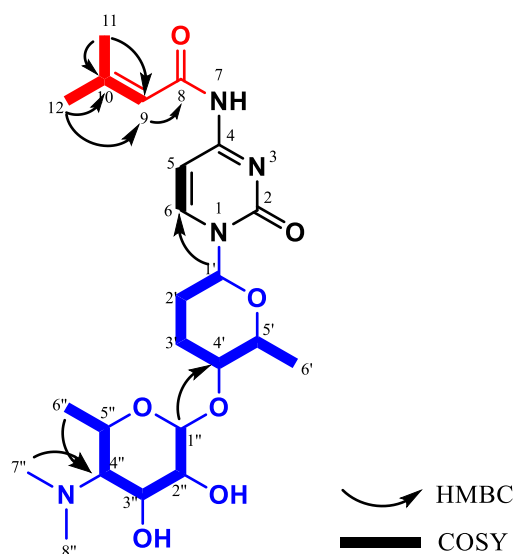


Figure 4-13 Structure of Compound IV with key correlations.

Once the full structure confirmation was accomplished via NMR, the compound was yet again searched in SciFinder in order to compare the novelty of the compound. Unfortunately, it was found to be known and recently isolated and characterized by Hiroshi and coworkers in 2019¹⁵⁵. Strong anti-mycobacterial activities were reported in a patent WO 2019/044941¹⁵⁵. for compound VI with MIC values of 0.78 ug/mL and 0.39 ug/mL against the strains *Mycobacterium avium* JCM 15430 and *Mycobacterium intracellulare* JCM 6384, respectively.

4.5.4 Compound VI

Compound VI was identified and isolated also from fraction 3. It was evident from mass spectrometry that it carries $[M+H]^+ = 503.2497$. With 11 ring double bond equivalents, its molecular formula was predicted to be C₂₅H₃₄N₄O₇ ($\Delta = 0.6$ ppm). Mass fragmentation studies revealed further evidence to the proposed structure (**Figure 4-14**), which were also supported

by H^1-H^1 COSY spectra (see **Figure 4-15**). Keeping the known skeleton intact to that of plicacetin rest was confirmed with partial NMR and MS^2 data.

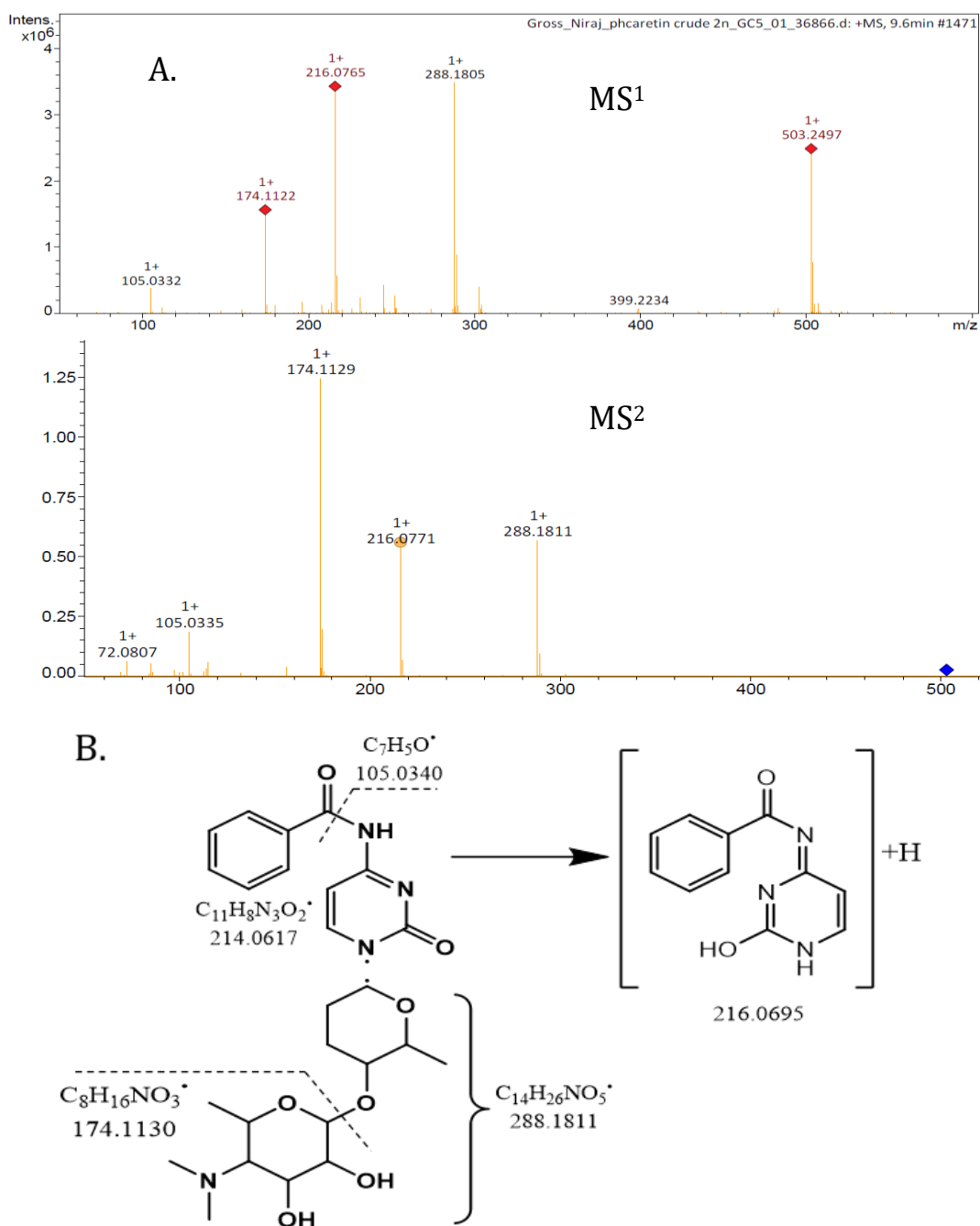


Figure 4-14 A) MS^1 and MS^2 spectra of compound VI. B) Explanation of fragments from MS^2 spectra.

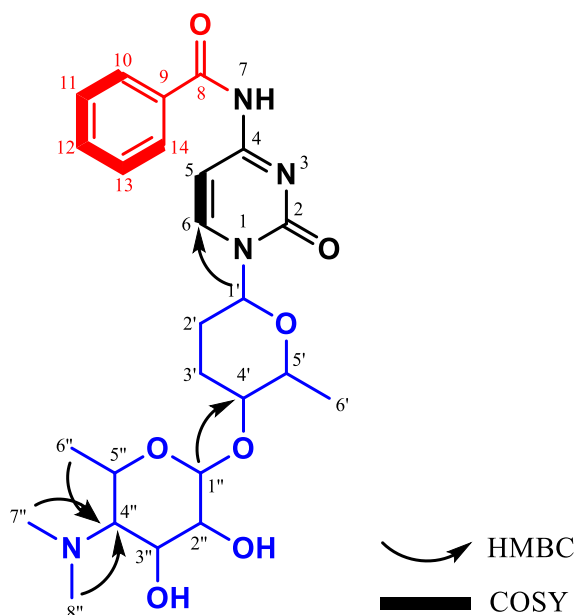


Figure 4-15 Aromatic region ^1H - ^1H COSY (red), other correlations were similar to the plicacetin standard.

4.5.5 Compound XI

The formula of compound XI was identified as $\text{C}_{25}\text{H}_{42}\text{N}_4\text{O}_7$, whose exact neutral mass is 510.3119 with a mean error 1.1 ppm (**Figure 4-16**). In positive mode, based on the plicacetin structure, the ion fragments $[\text{M}+\text{H}]^+=174$ and 288 were identified as monosaccharide and disaccharide as well. Meanwhile, relative peaks marked with $[\text{M}-\text{H}]^-=509.2985$ and $[2\text{M}-\text{H}]^-=1019.5053$ also existed in the negative mode, which were additional evidence for the confirmation of the molecule mass and its formula.

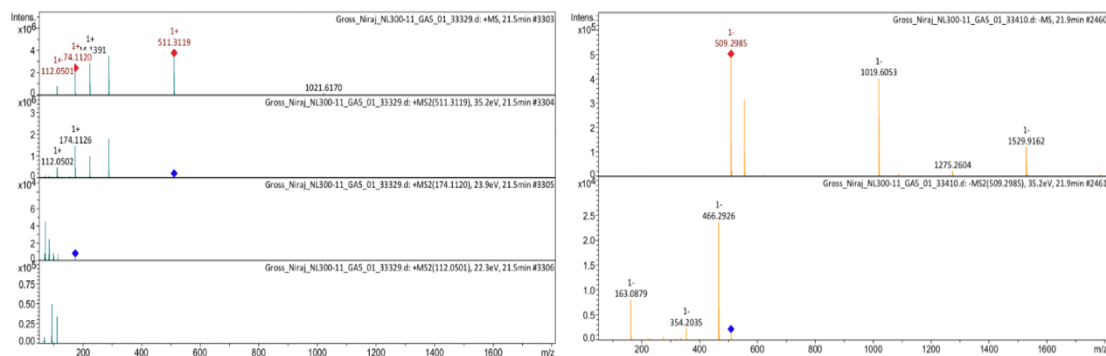


Figure 4-16 HRMS spectrum of pure compound XI in positive and negative mode

1D NMR experiments containing ^1H -NMR, ^{13}C -NMR, DEPT135 NMR and 2D NMR experiments, such as ^1H - ^1H COSY, ^1H - ^{13}C HSQC, ^1H - ^{13}C HMBC, NOESY and ^1H - ^{13}C HMBC-TOCSY spectra were recorded in d_4 -methanol.

The ^1H -NMR spectrum demonstrated signals which were highly similar with the other derivatives. The signals at δ_{H} 7.48 and 8.11 were determined quickly as the hydrogens of the cytosine moiety. Most of the hydrogens on two sugar units ranged from δ_{H} 3.00 to 4.00 ppm, while the anomeric ones influenced by an electron-withdrawing effect, has a higher chemical shift δ_{H} 5.69 and 4.79. The hydrogen located at δ_{H} 2.46 belonged to a nitrogen methyl on amosamine moiety. However, due to the predicted alkane side chain, the signals from δ_{H} 1.00 to 2.00 exhibited a higher degree of complexity. Hence, assignment of remaining signals was confirmed with 2D NMR data.

In combination with a DEPT135 NMR spectrum, three quaternary carbons were identified at δ_{C} 176.2, 164.4 and 157.5 in ^{13}C -NMR spectrum. Two of them represent carbonyl groups, which can be detected as a shoulder peak at 298 nm in the UV/VIS spectrum as well. The resonances of five secondary carbons at δ_{C} 36.0, 32.6, 30.8, 30.3 and 27.7 appeared in a negative phase and are derived from side chain part and sugar. During the elucidation process, all hydrogens were assigned to their carbons via the ^1H - ^{13}C HSQC spectrum.

In the ^1H - ^1H COSY spectrum, the signals at 1.24 and 1.36 ppm were assigned to two methyl residues of the sugar ring due to the interaction with signals at 3.82 and 3.74 ppm, which belonged to sugar hydrogens. The remaining key correlations of hydrogens on two sugar moieties are presented using bold bonds illustrated in **Figure 4-17**. Another four correlations revealed the connected type among hydrogens on the side chain, gathering the correlation between a methyl signal at δ_{H} 0.92 and δ_{C} 32.6 in ^1H - ^{13}C HMBC spectrum. Besides that, a hydrogen atom δ_{H} 1.47 showing two interactions with carbons at δ_{C} 35.4 and 30.3 ppm also indicated its assignment.

Hydroxy groups from compound XI could not be confirmed, since they can be covered by signals from d_4 -methanol. In this case, the absorbance at ν_{max} 3382.53 cm^{-1} can serve as the existing evidence for the existence hydroxy groups on sugar moieties. Gathering the absorbance band at ν_{max} 1672.95 cm^{-1} in IR spectrum (Supplementary information), the amide bond between side chain and cytosine could be supported. Additionally, a sharp peak at 248 nm in UV/Vis spectrum provided additional evidence for presence of a conjugated system or aromatic systems in compound XI.

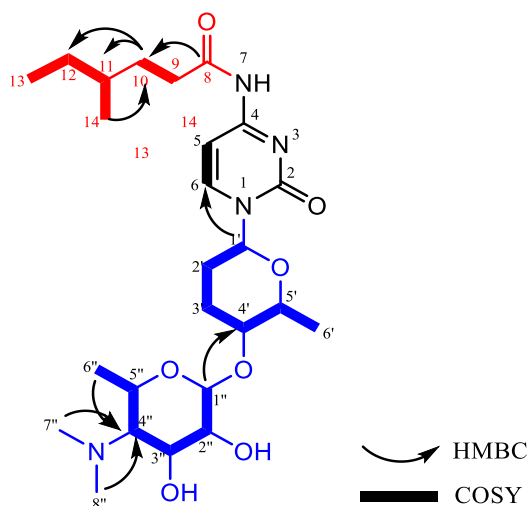


Figure 4-17 Key correlations of the ^1H - ^{13}C HMBC (arrow) and ^1H - ^1H COSY (bold bonds) spectra of compound XI

Combining all signals and correlations, the planar structure of compound XI was confirmed as follows (**Figure 4-17**). All of signals detected in ^1H -NMR and ^{13}C -NMR are listed as follows (**Table 4-8**). Results on the IH-U column showed an elution order of *R* before *S* for 4-methyl-*N*-1-naphthylhexanamide and a result of *S* as absolute configuration for the hydrolysis product of the natural product.

Table 4-8 NMR signals of compound XI recorded in d_4 -methanol (700 MHz)

Unit	Position	$\delta_{\text{C/N}}$ in ppm (mult.)	δ_{H} in ppm (mult., <i>J</i> in Hz)
Side Chain	C=O	176.2, qC	-
	C9	36.0, CH ₂	2.45 (s, 2H)
	C10	32.6, CH ₂	1.47 (m, 2H)
	C11	35.4, CH	1.39 (d, 6.1, 1H)
	C12	30.3, CH ₂	1.24 (d, 6.2, 2H)
	C13	11.6, CH ₃	0.90 (m, 3H)
	C14	19.3, CH ₃	0.92 (m, 3H)
Cytosine	C=O	157.5, qC	-
	C4	164.3, qC	-
	C5	98.4, CH	7.48 (d, 7.5, 1H)
	C6	146.3, CH	8.11 (d, 7.6, 1H)
Sugar 1	C1'	84.7, CH	5.76 (m, 1H)
	C2'	27.7, CH ₂	1.64, 2.37 (m, 2H) overlap
	C3'	30.8, CH ₂	1.64, 2.15 (m, 2H) overlap

	C4'	75.0, CH	3.40 (s, 1H)
	C5'	78.5, CH	3.74 (d, 3, 1H)
	C6'	19.1, CH ₃	1.36 (d, 6.2, 3H)
Sugar 2	C1''	96.2, CH	4.91 (s, 1H)
	C2''	74.7, CH	3.42 (dd, 9.5, 3.8, 1H)
	C3''	70.2, CH	3.87 (s, 1H)
	C4''	71.9, CH	2.13 (m, 1H)
	C5''	67.4, CH	3.82 (d, 4, 1H)
	C6''	19.7, CH ₃	1.24 (d, 6.2, 2H)
	C7''	42.4, CH ₃	2.46 (s, 3H)
	C8''	42.4, CH ₃	2.46 (s, 3H)

Furthermore, according to the Beer–Lambert law $A = \epsilon lc$, the molar attenuation coefficient at different wavelength was calculated and summarized below (**Table 4-9**). UV/Vis, IR and CD spectra are presented in the supplementary information. As molecule is structurally similar to also cytosaminomycin E, CD spectrum for this compound were in alignment to the data published by Xu and coworkers.

Table 4-9 The absorbance at specific wavelength and corresponding λ_{\max} ($\log \epsilon$) values for compound XI

Wavelength (nm)	Absorbance	λ_{\max} ($\log \epsilon$)
215	1.0008	4.13
246	0.7034	3.97
298	0.3347	3.65

4.5.6 Structure speculations for compound VII, VIII, IX, X and Xa

For compound masses 481, 483, 495 and 497 there were two structural possibilities. These were well illustrated by Hiroshi et. al.¹⁵⁵ and Xu et. al.¹⁵⁶ also in their patent and published results, respectively. During our LCMS screening, it became apparent that there are multiple BPC for one m/z value. See **Figure 4-18** and **Table 4-10**

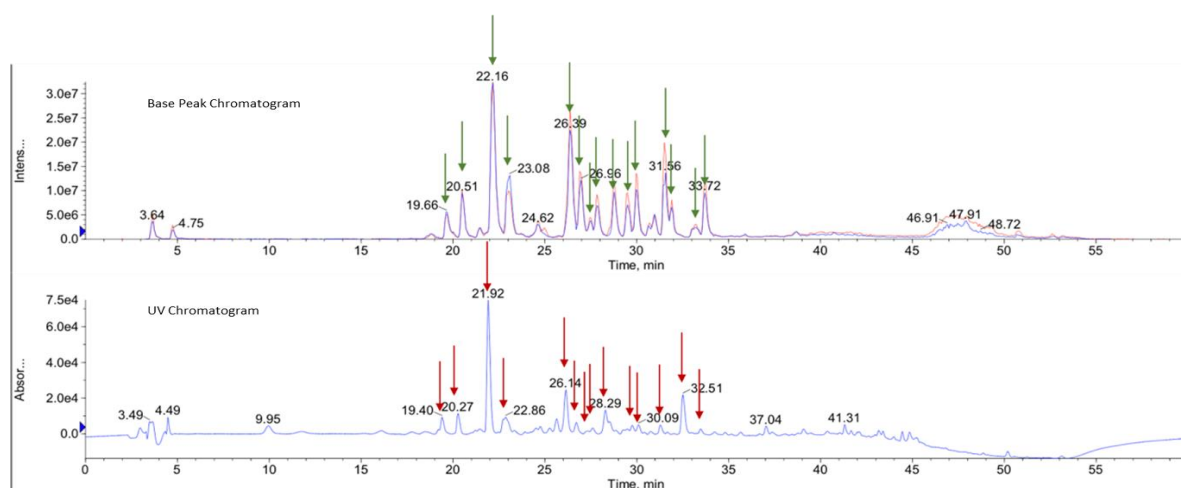
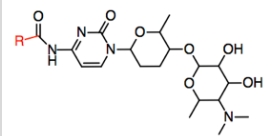
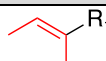
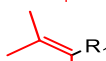

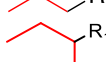
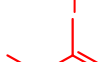


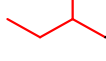
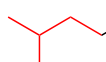


Figure 4-18 Base peak chromatogram and UV chromatogram showing elution of plicacetin derivatives at different retention times.

Table 4-10 Derivatives with same mass but proposed different structures.

Retention Time (min)	[M+H] ⁺	Predicted molecular formula	 Predicted Structure R- R ₁	Possibility (CAS Nr.)	Ratio in crude (/100)
26.14	481	C ₂₃ H ₃₆ N ₄ O ₇		40551-I	1.4176
26.96	481	C ₂₃ H ₃₆ N ₄ O ₇		40551-F	0.8134
27.43	483	C ₂₃ H ₃₈ N ₄ O ₇		40551-K	0.4891
27.83	483	C ₂₃ H ₃₈ N ₄ O ₇		Unknown	0.9274
29.50	495	C ₂₄ H ₃₈ N ₄ O ₇		Unknown	0.4543
29.97	495	C ₂₄ H ₃₈ N ₄ O ₇		40551-L	1.1885
31.51	497	C ₂₄ H ₄₀ N ₄ O ₇		Unknown	1.3192
31.91	497	C ₂₄ H ₄₀ N ₄ O ₇		Cytosaminomycin E	0.2372
33.17	509	C ₂₅ H ₄₀ N ₄ O ₇		40551-G	0.2548

Most of compounds in **Table 4-9** were not obtained in sufficient amounts during our isolation attempts, therefore a complete spectroscopic analysis was not possible. However, HR/MS data showed identical molecular formulae and the MS² fragments are also similar for both compounds. This is due to the fine structural changes in the side chain which ultimately gave 2

equal structural possibilities for these molecules, of which one is known while other is unknown. With this, we reached the limit of HR/MS based characterization as well. NMR data are required to shed more light on these. For example, the compound with m/z 495 was found to be eluted in two different retention times. One could speculate that this can be due to presence of *cis/trans* configurations which was also observed with compound I ($m/z = 499$) or two slightly different structures offering a double bond in different positions. Further work is needed to confirm these data. Also, the variability in the branching and chain length was also seen in some of these derivatives with $m/z = 497$ and 483.

4.5.7 Compound XII

HR/MS data confirmed the molecular formula for $[M+H]^+ = 588.3035$ to be $C_{29}H_{41}N_5O_8$ with an error of 1.8 ppm and rdb was found to be 12. During cultivation of the 6th batch, compound XII was produced in a collectable amount. In earlier batches it was found to be produced in lower amounts. But, unlike other derivatives compound XII exhibited different UV profile which was indicative of change in chromophore. The absorbance at specific wavelength and corresponding λ_{max} ($\log \epsilon$) values for compound XII is presented in **Table 4-12**. Compound XII was collected over HPLC with same system as for other derivatives. A total of 13.1 mg of sample was subjected to NMR to figure out key changes in the structure.

Compared to other plicacetin/amicetin derivatives, compound XII was stable. Isolation of compound XII was carried out in HPLC using RP column Luna Omega Polar. Methylated derivatives of XII were also observed with $m/z = 574, 560$. But, due to low abundance, spectroscopic data could not be presented. Structure speculations of these derivatives will be discussed later in discussion section below.

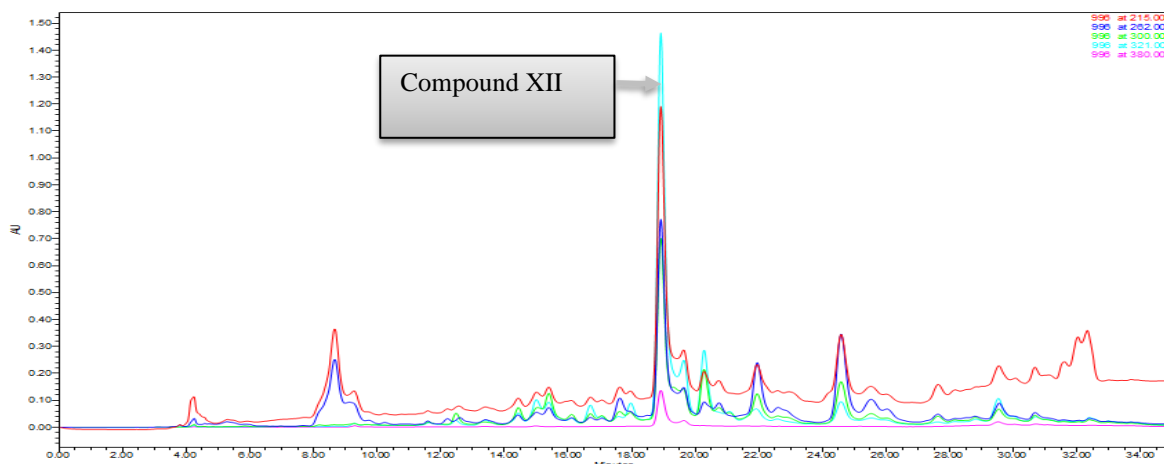


Figure 4-19 HPLC chromatogram showing compound XII

Table 4-11 NMR data for compound XII recorded in *d*₄-MeOD (400 MHz)

Unit	Position	δ_{CN} in ppm (mult.)	δ_H in ppm (mult., <i>J</i> in Hz)
Side chain (brown)	C16	59.2, CH	4.17 (q, 7.0, 1H)
	C17	212.3, qC	-
	C18	25.4, CH ₃	2.17 (s)
	C19	17.3, CH ₃	1.42 (d, 7.0, 3H)
Aromatic region (red)	C=O	168.6, qC	-
	C10, C14	131.5, 2CH	7.82 (d, 8.7, 2H)
	C11, C13	114.4, 2CH	6.64 (d, 8.7, 2H)
	C12	155.0, qC	-
	C9	121.0, qC	-
Cytosine (black)	C=O	157.4, qC	-
	C4	164.8, qC	-
	C5	98.7, CH	7.58 (d, 7.4, 1H)
	C6	145.8, CH	8.12 (d, 7.4, 1H)
Sugar 1	C1'	84.6, CH	5.77 (dd, 10.8, 2.3, 1H)
	C2'	31.0, CH ₂	2.15 (m, 1H), 1.65 (m, 1H)
	C3'	27.7, CH ₂	2.38 (m, 1H), 1.68 (m, 1H)
	C4'	75.0, CH	3.41 (m, 1H)
	C5'	78.4, CH	3.75 (m, 1H)
	C6'	19.1 CH ₃	1.36 (d, 6.4, 3H)
Sugar 2	C1''	96.1, CH	4.92 (d, 3.7, 1H)
	C2''	74.6, CH	3.41 (m, 1H) overlap
	C3''	70.2, CH	3.87 (t, 9.6, 1H)
	C4''	71.8, CH	2.13 (t, 10.1, 1H) overlap
	C5''	67.4, CH	3.83 (m, 1H)
	C6''	19.6, CH ₃	1.24 (d, 6.4, 3H)
	C7''	42.3, CH ₃	2.47 (s, 3H)
	C8''	42.3, CH ₃	2.47 (s, 3H)

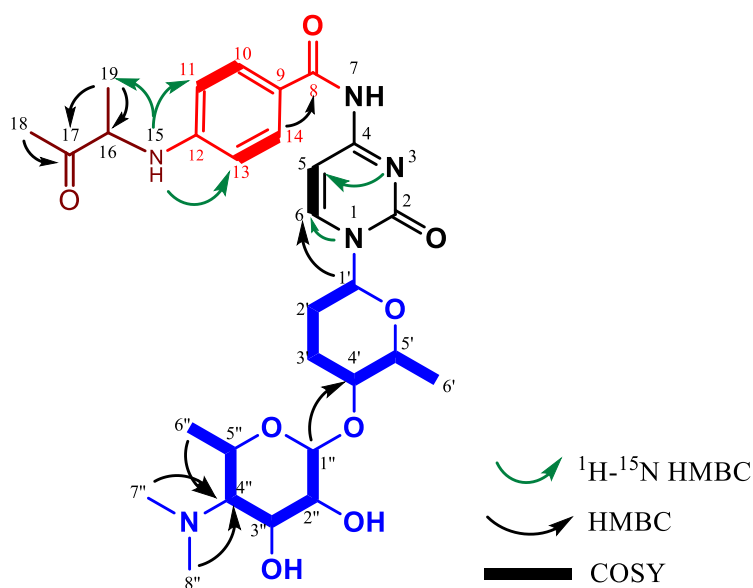


Figure 4-20 Key correlations for compound XII

Table 4-12 Specific wavelength and corresponding λ_{max} ($\log \epsilon$) values for compound XII

Wavelength (nm)	Absorbance	λ_{max} ($\log \epsilon$)
332	0.90684	4.59
255	0.47594	4.31

4.6 Structure elucidation of streptcytosines

Amicetins were discovered in golden era of antibiotics. They are found be produced by various actinomycetes¹⁵⁷⁻¹⁶¹. These antibiotics share a common disaccharide pyrimidine nucleoside motif, namely cytosamine, in which amosamine is α - (1 \rightarrow 4)-linked to amicitose and itself in turn β - (C \rightarrow N)-linked to a cytosine nucleus. The NH_2 group of the amosamine residue could be mono- or dimethylated, while the C-3 of the amicitose residue could be hydroxylated, and the N7 on cytosine could be acylated, thus constituting the molecular diversity of these compounds¹⁵³. Structurally, streptcytosine A was reported from the culture broth of *Streptomyces* sp. TPU1236A collected in Okinawa, Japan⁷⁴ and the same compound was later isolated from *Streptomyces rochei* 06CM016 collected in the Mediterranean Sea and named rocheicoside A. Though streptcytosine A had similar structure to that of of amicetin, other streptcytosines lacked amosamine moiety (**Figure 4-21**). These molecules possess antibacterial and antiviral activity and are active against tuberculosis^{74, 153, 162, 163}.

In this study, known streptcytosines were isolated along with one novel streptcytosine congener (**Figure 4-21**). The structure of new streptcytosine congener was confirmed with spectroscopic data (HR/MS, IR, UV, and 2D NMR experiments including ^1H - ^1H COSY, ^1H - ^{13}C HSQC, ^1H - ^{13}C HMBC, ^1H - ^{15}N HMBC and ^1H - ^1H NOESY spectra).

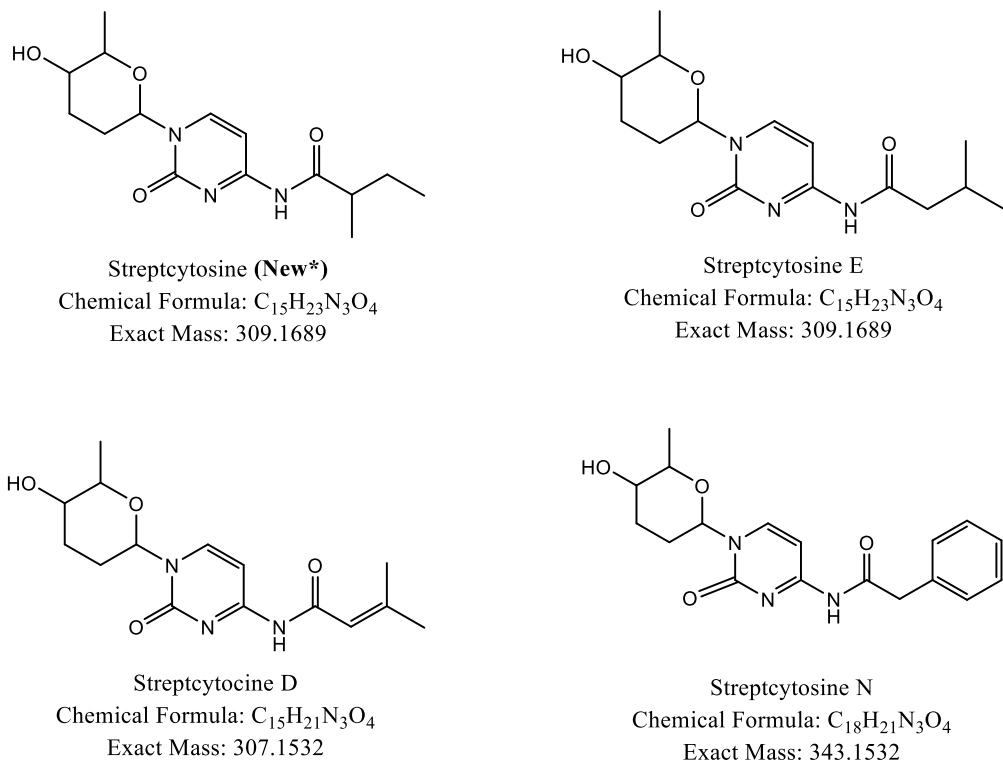


Figure 4-21 Structure of streptcytosines isolated in this study.

Streptcytosines eluted in the earlier fractions during the fractionation process with sephadex LH-20. All coeluting peaks which contained our target masses were first collected via preparatory HPLC and were later resolved with an analytical column employing an isocratic run (see **Figure 4-22**)

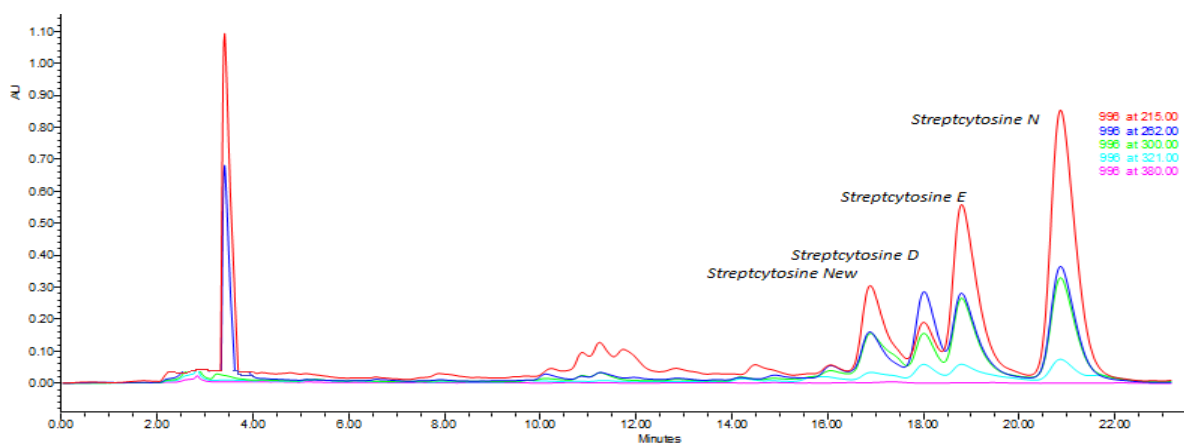


Figure 4-22 HPLC separation of streptcytosines

4.6.1 Streptcytosine New*

In this study, a new streptcytosine derivative was isolated along with other known streptcytosines. HPLC separation of streptcytosines were carried out as mentioned in method section 2.3.8.1 using an isocratic flow. Except for streptcytosine A, all other consisted of one sugar moiety. In our case, it consisted of amicitose as a sugar unit bonded to cytosine via C-N bond. It was again extended via a C-N bond to a carbonyl and branched side chain. We attempted to solve the stereochemistry of branched butyric side chain. Results on the IB-U column showed an elution order of *S* before *R* for 2-methyl-*N*-1-naphthylbutanamide and a result of *S* as absolute configuration for the hydrolysis product of the natural product. A slight amount of *R* could be detected, presumably due to racemization during reconditioning. This evidence were supported by IR, UV-Vis, optical rotation and CD data which are presented in supplementary information.

Table 4-13 NMR data for streptcytosine (New*) recorded in *d*₆-DMSO (700 MHz)

Unit	Position	$\delta_{C/N}$ in ppm (mult.)	δ_H in ppm (mult., <i>J</i> in Hz)
Side chain(red)	C1	11.4, CH ₃	0.82 (t, 7.4, 3H)
	C2	26.4, CH ₂	1.37, 1.55 (m)
	C3	41.7, CH	2.55 (m)
	C4	177.3, qC	-
	C5	16.77, CH ₃	1.04 (d, 6.6, 3H)
Cytosine(black)	C=O	162.4, qC	-
	C4	153.9, qC	-
	C5	98.8, CH	7.27 (d, 7.5, 1H)
	C6	145.4, CH	8.07 (d, 7.5, 1H)
Sugar 1(blue)	C1'	82.1, CH	5.62 (dd, 10.8, 2.2, 1H)
	C2'	29.7, CH ₂	1.88 (m, 1H), 1.66 (m, 1H)
	C3'	31.3, CH ₂	2.00 (m, 1H), 1.51 (m, 1H)
	C4'	68.5, CH	3.11 (m, 1H)
	C5'	78.5, CH	3.38 (m, 1H)
	C6'	18.1 CH ₃	1.20 (d, 6.2, 3H)

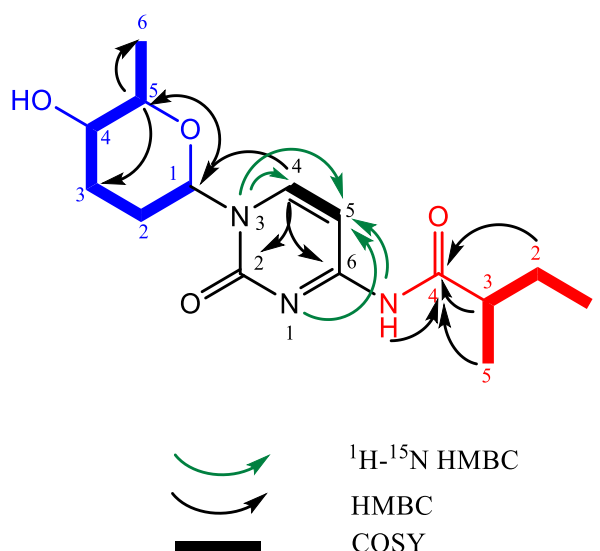


Figure 4-23 Structure showing key correlation for novel streptocytosine

4.7 Precursor directed biosynthesis (PDB)

Having observed the structural diversity of natural plicacetin derivatives, we probed the concept of precursor-directed biosynthesis study (PDB). In this strategy, the growth medium of a microorganism is supplemented with an artificial precursor that can be incorporated into the molecule of interest instead of the natural substrate¹⁶⁴. Such an approach was successfully applied to generate clickable microcystins by supplementing the culture medium with various azide- or terminal alkyne-containing amino acid analogues. Similarly, novel glycopeptides were generated by feeding synthetic β -hydroxy amino acids¹⁶⁵.

In this experiment, we fed various benzoic acid derivatives. Based on previous experiences from our laboratory and published literature, benzoic acid derivatives with at least one halogen substitution in different positions were probed, since most of the compounds containing a halogen empirically have a higher bioactivity but also exhibit sometimes higher toxicity. In order to applying bioisosterism strategy, several bioisosters were tested instead of para-aminobenzoic acid, such as mercaptobenzoic acid and hydroxybenzoic acid. In this way, the replacement of a hydrogen atom provides the possibility for testing the flexibility of the enzyme involved for changing the bioavailability and for altering the metabolism of known compounds later on. Also, precursors comprising two oxidation states of nitrogens or triple bonds. Furthermore, nonaromatic residues and several fatty acids are added as precursor candidates to perform a comprehensive analysis.

Ten out of 43 precursors tested in this project were taken up by the wild type *Streptomyces* sp. strain SHP 22-7 strain and transformed into plicacetin derivatives, which were confirmed by LRLC-MS and HR/MS. They are marked in grey in the following figure (**Figure 4-24**).

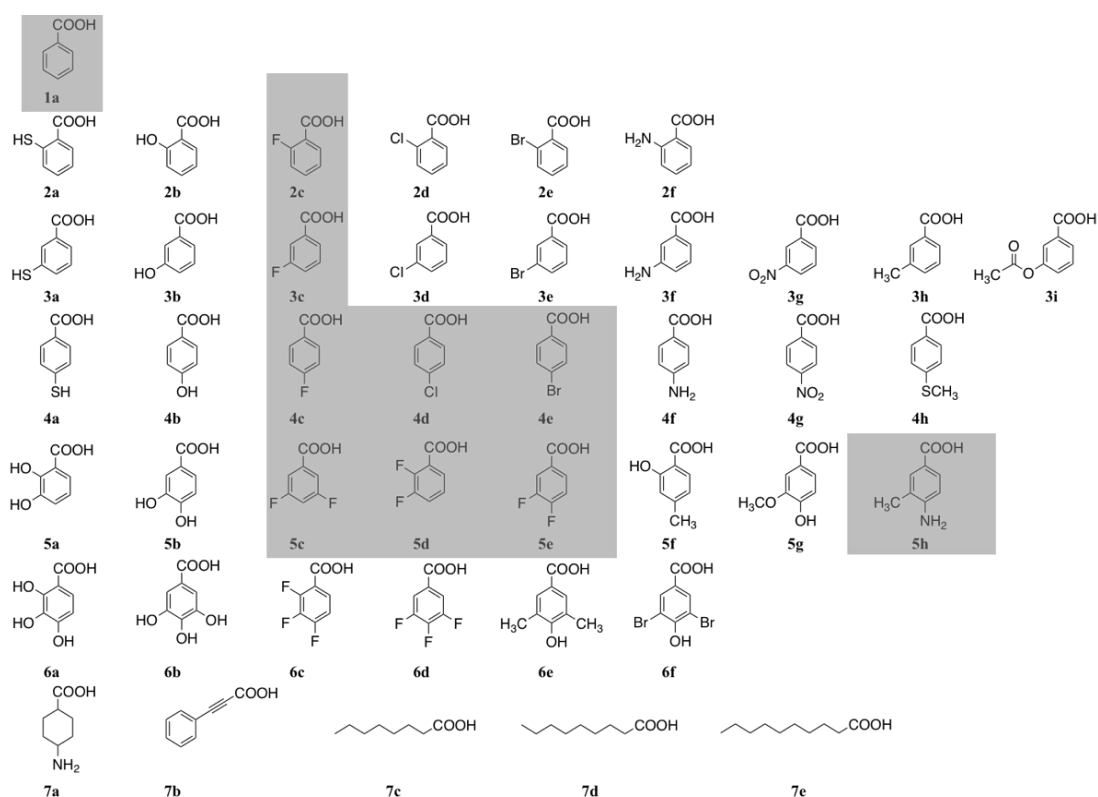


Figure 4-24 43 precursors employed in the precursor-directed biosynthesis study.

Compound VI was produced in low amounts in the wild type strain. Upon feeding of 1a, we could observe a sharp increase in production of compound VI. This led to the conclusion that precursor 1a has been successfully taken up. However, semi quantitative LC-MS result showed a tremendous concentration difference between cultivation under regular fermentation conditions and with supplementation of the specific precursors. By increasing the precursor pool the product was increased by approximately ten folds (**Figure 4-25**). Compound VI eluted at 22 min in the corresponding HPLC analysis (**Figure 4-26**).

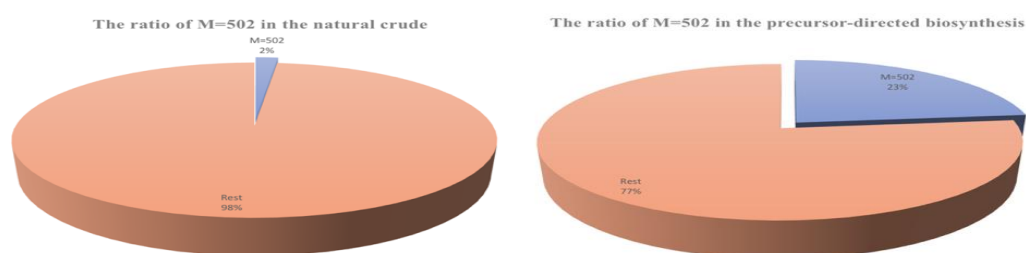


Figure 4-25 The ratios of compound VI between cultivation under regular fermentation conditions (left) and with supplementation of the precursor-directed biosynthesis (right)

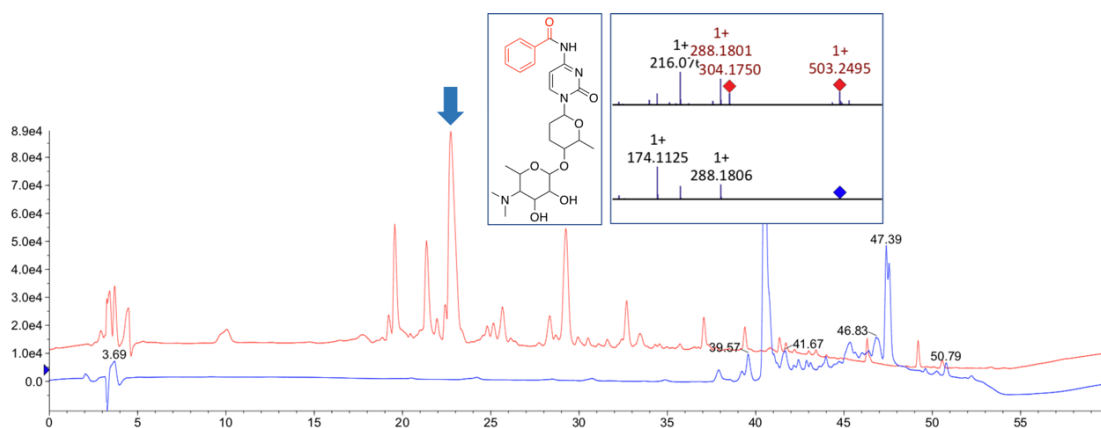


Figure 4-26 LC-MS chromatogram between sample fed with precursor 1a (arrow marked) and media control (blue) in positive mode. The arrow indicates compound VI

4.7.1 Feeding of halogenated benzoic acid

In natural products, a carbon-halogen bond is taken with great interest. Compounds with halogen-dependent activity include vancomycin, rebeccamycin, salinosporamide A and many more. Over 4000 halogenated products have been isolated from natural sources.¹⁶⁶ Among halogenation also, chlorination is a predominant event followed by bromination. While iodination and fluorination are rare in nature. Examples of halogenated natural products include the remarkably potent anticancer agents β -lactone salinosporamide A, the macrocyclic lactone polyether spongistatin, the indolocarbazole rebeccamycin, an enediyne calicheamicin γ II, and a number of antibiotics, such as vancomycin, chlortetracycline, and chloramphenicol^{167, 168}.

Similarly, we applied PDB concept to insert halogens into the plicacetin skeleton and isolated them to improve their bioassay profile. When we fed fluoro-benzoic acids to the strain, ortho- and meta-position, precursor 2c and 3c were taken in and transformed into novel derivatives (**Figure 4-27** and **Figure 4-28**), separately. Whereas the remaining precursors with various substituent groups in ortho- and meta-position cannot be taken up. The acceptance of precursor 2c and 3c can be ascertained by the analysis of the fragmentation patterns of the compounds, which matched the classical characteristic ion fragments according to the preceding experience.

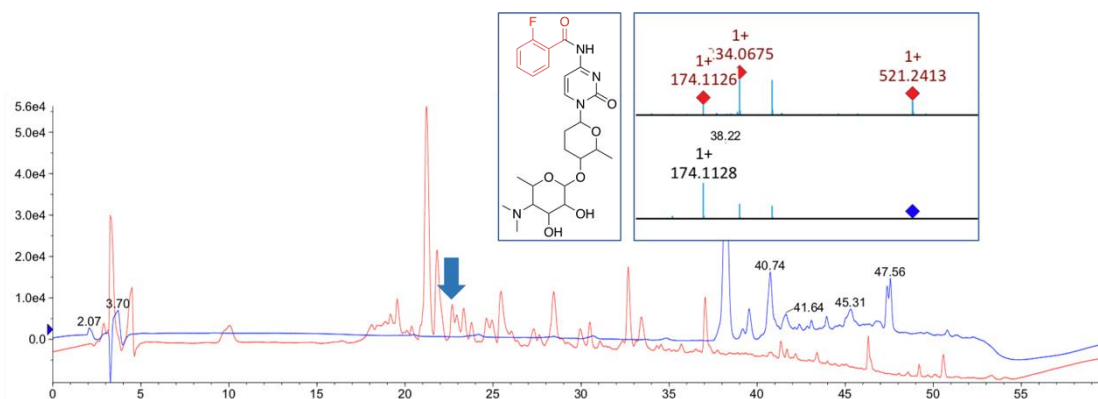


Figure 4-27 Comparative LC-MS profiling of a sample fed with precursor 2c (red) and media control (blue) in positive mode. The arrow represents precursor attached product.

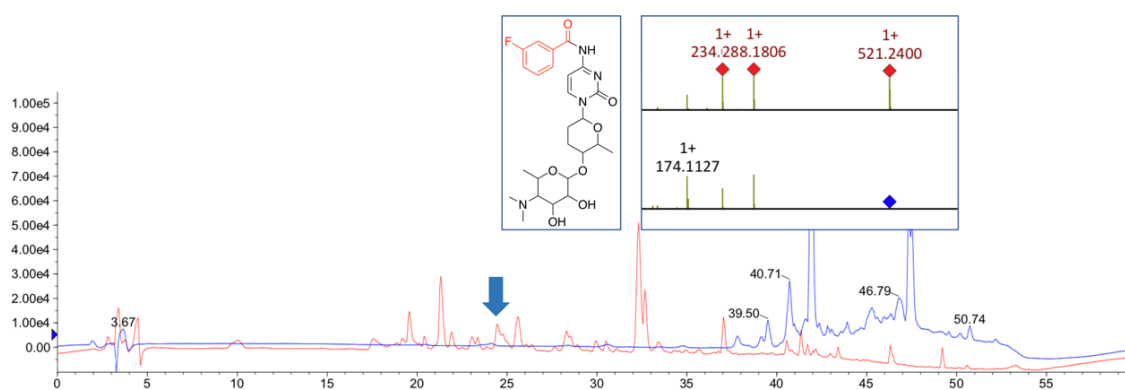
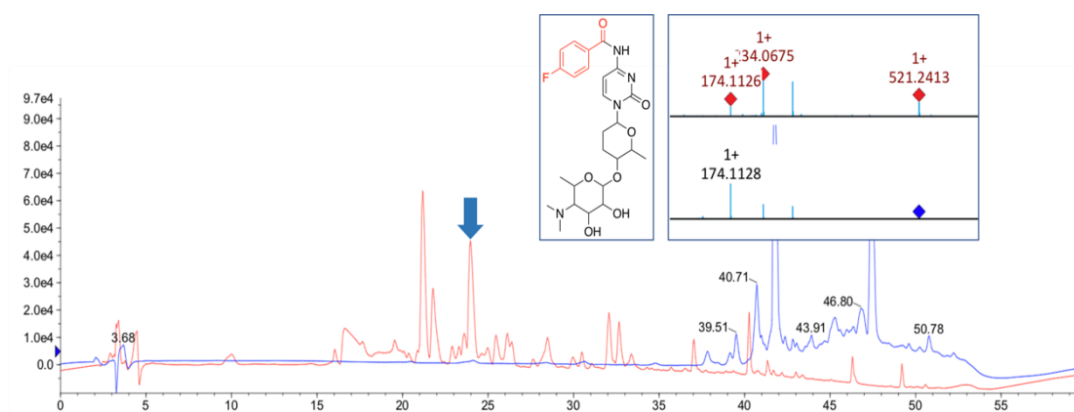


Figure 4-28 Comparative LC-MS profiling of a sample fed with precursor 3c (red) and media control (blue) in positive mode. The arrow represents precursor attached product.

All halogenated precursors in para-position were accepted (**Figure 4-29**). Meanwhile, the isotope patterns of chlorine and bromine was also detected in HR/MS, which served as an additional evidence.



A

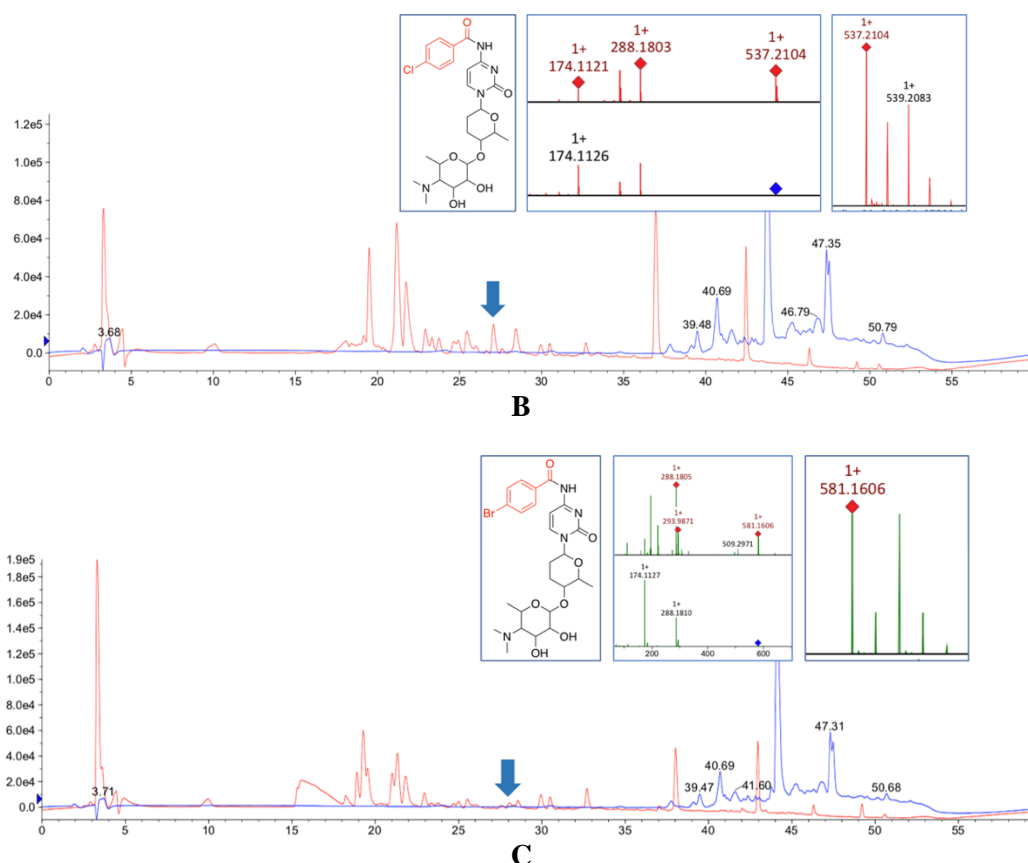


Figure 4-29 Comparative LC-MS profiling of a sample fed with precursor 4c (A), 4d (B) 4e (C) (red) and media control (blue) in positive mode. Isotope pattern is also shown in left. The arrow represents precursor attached product.

Para-substituted fluoro-benzoic acid was not only successfully incorporated but also produced in collectable amounts. For all fluoro-substituted plicacetin upscaling was performed (from 50 mL to 4L). However, in this study only para positioned fluoro-plicacetin derivative (4F-plicacetin) was fully characterized by 1D and 2D spectroscopic analysis. First of all, presence of fluorine atom was confirmed by a ^{19}F NMR experiment (**Figure 4-30**).

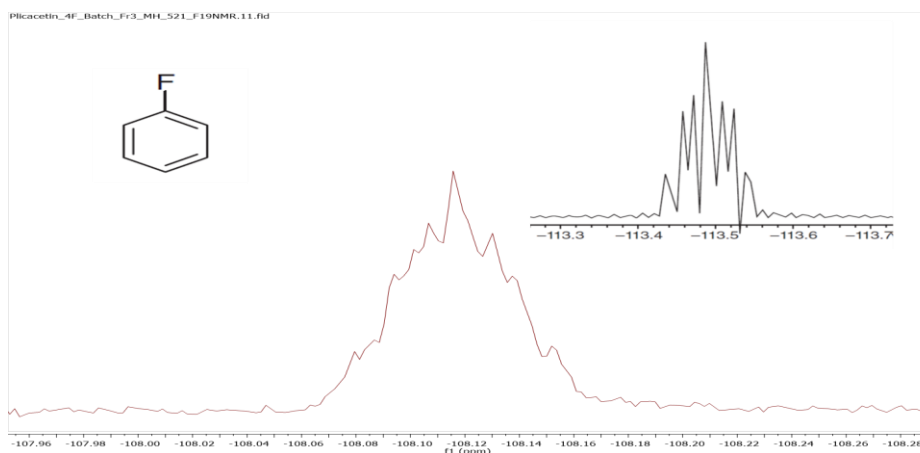


Figure 4-30 ^{19}F NMR spectrum for 4F-plicacetin (red) compared to standard value and peak shape. Note: Difference in ppm observed is due to the use of different solvent

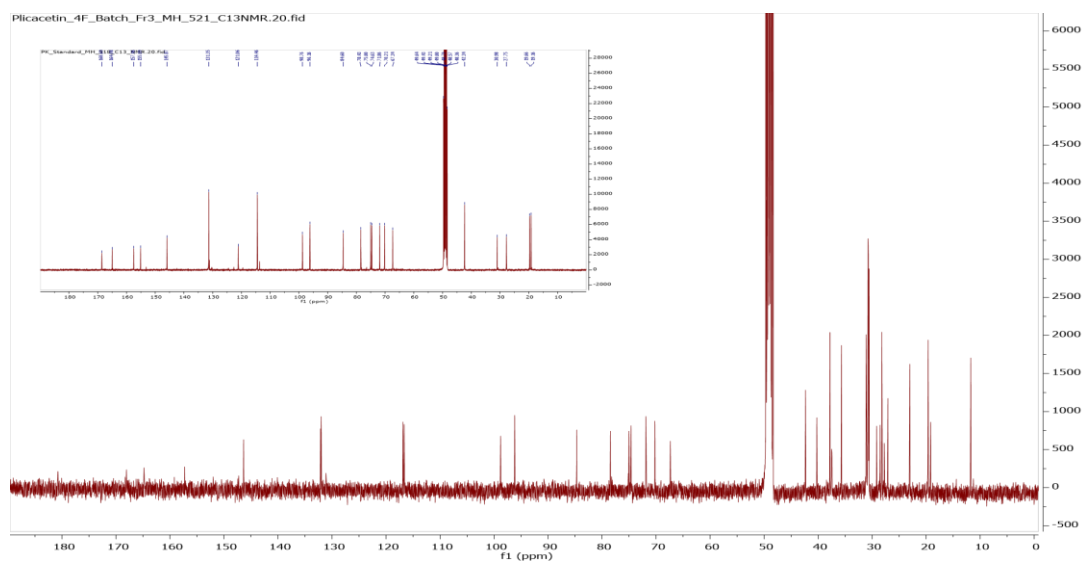


Figure 4-31 4F-plicacetin ^{13}C NMR compared to the plicacetin standard.

The majority of the structure of 4F-plicacetin resembled normal plicacetin, therefore we focused only on the differences which were given on the fluorine attached aromatic moiety. Anisotropic effects do not play a significant role in fluorine NMR. Therefore, fluorine substituents on a benzene ring absorb in the general region of fluoroalkenes. However, fluorine substituent on benzene has a characteristic effect upon the ^{13}C spectrum of benzene, and it couples in a distinctive and highly consistent manner with the ipso-, ortho-, meta-, and para-carbons¹⁶⁹ (**Figure 4-32**). In our ^{13}C analysis, we observed similar shifts (**Figure 4-33**). We have found coupling constants with similar values and characteristic carbon splitting which was definitive evidence for the generation of fluorinated plicacetin.

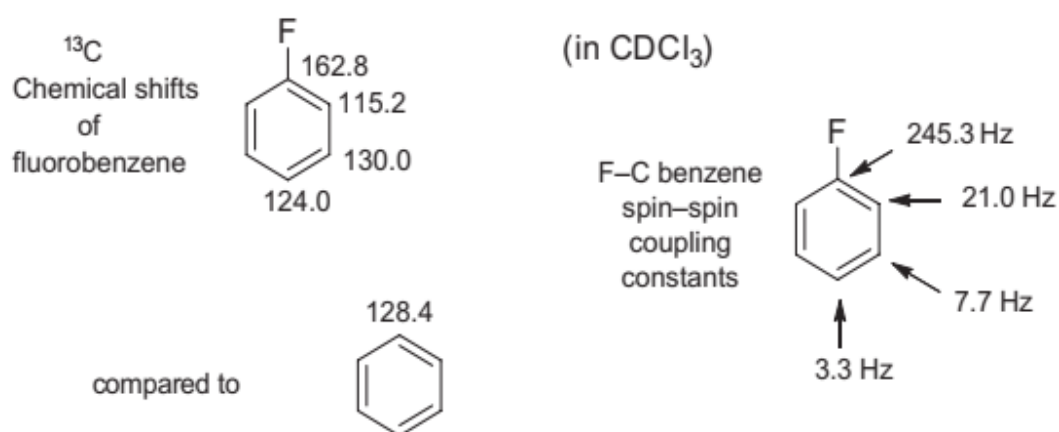


Figure 4-32 Single fluorine substituent ^{13}C shifts and coupling constants¹⁶⁹

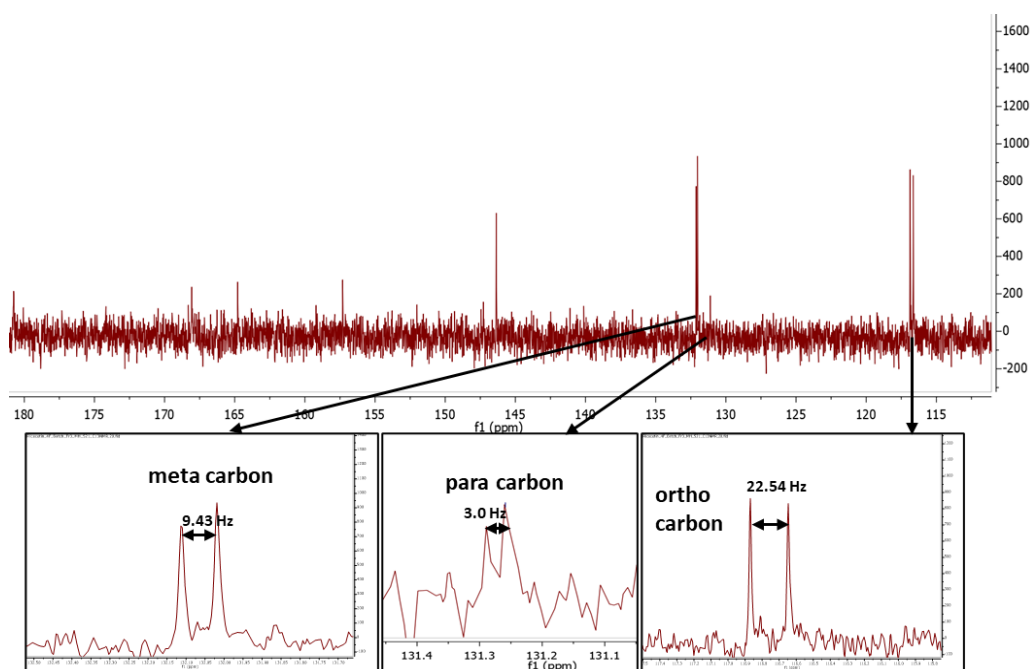


Figure 4-33 ^{13}C NMR showing coupling constants for ortho-, meta- and para-carbon on a benzene ring with para-substituted fluorine atom

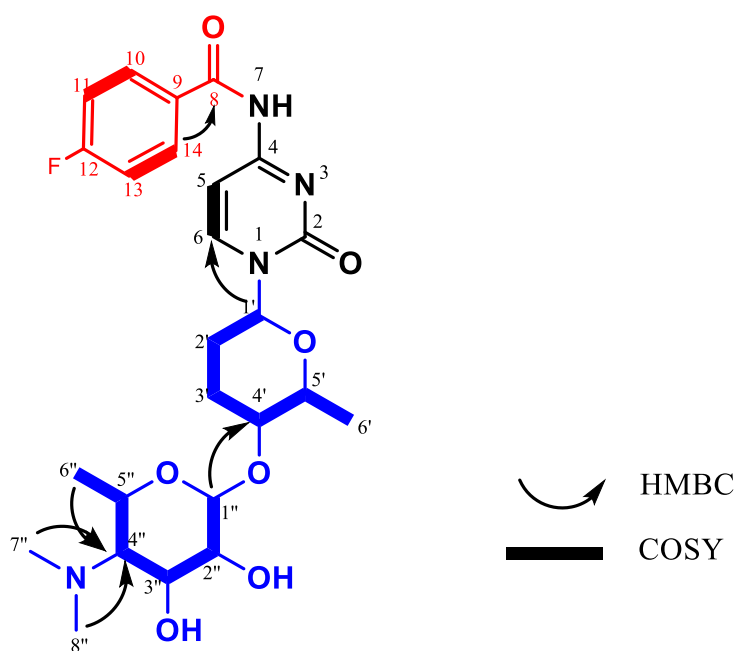


Figure 4-34 Structure of 4F-plicacetin with key correlations

Compared to mono-substituted chemicals, all di-substituted fluorine in various positions, 3,5-difluorobenzoic acid (**Figure 4-35**), 2,3-difluorobenzoic acid (**Figure 4-36**) and 3,4-difluorobenzoic acid (**Figure 4-37**), were absorbed and integrated into the molecule.

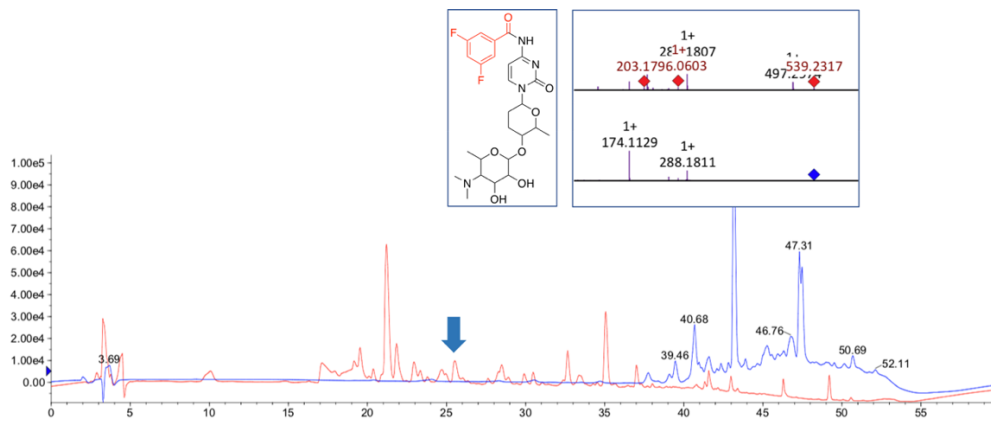


Figure 4-35 Comparative LC-MS profiling of a sample fed with precursor 5c (red) and media control (blue) in positive mode. The arrow represents precursor attached product.

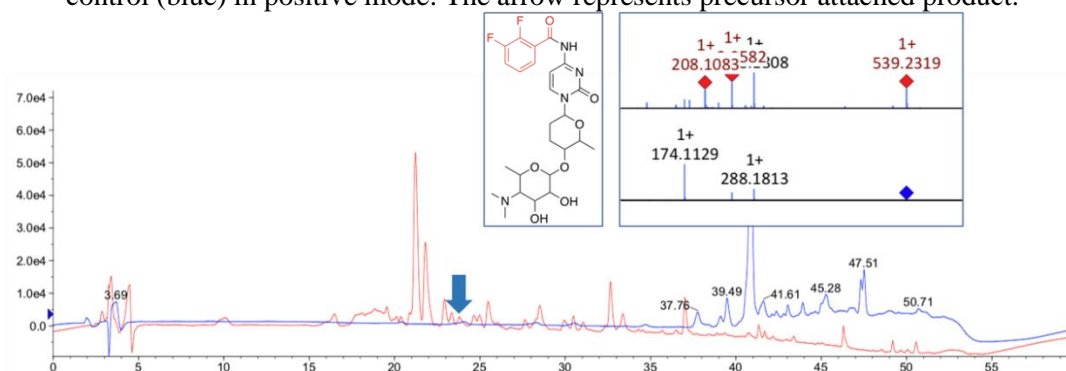


Figure 4-36 Comparative LC-MS profiling of a sample fed with precursor 5d (red) and media control (blue) in positive mode. The arrow represents precursor attached product.

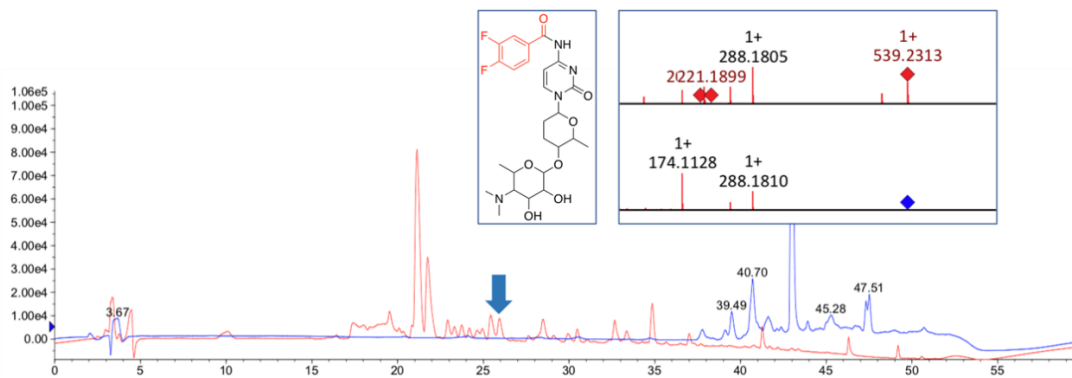


Figure 4-37 Comparative LC-MS profiling of a sample fed with precursor 5e (red) and media control (blue) in positive mode. The arrow represents precursor attached product.

As an exception, a small amount of 4-amino-3-methylbenzoic acid was accepted and transformed into plicacetin derivatives (**Figure 4-38**).

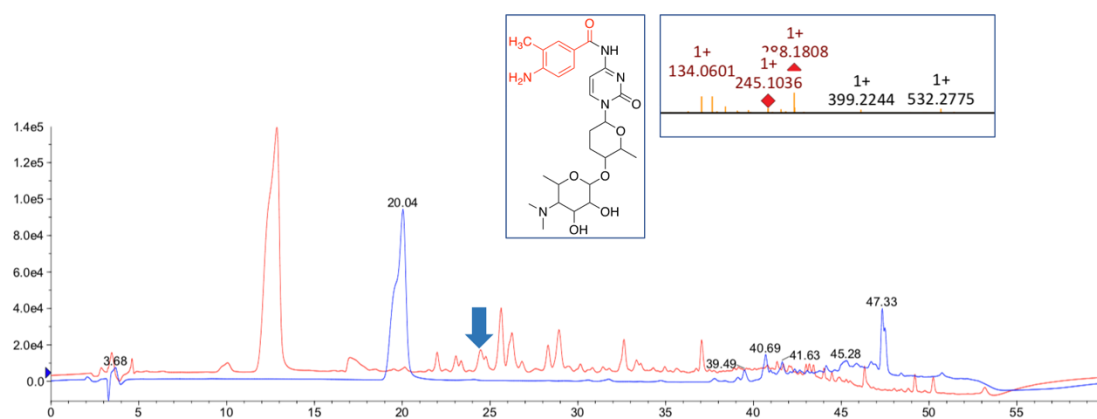


Figure 4-38 Comparative LC-MS profiling of a sample fed with precursor 5h (red) and media control (blue) in positive mode. The arrow represents precursor attached product.

Additionally, all of trisubstituted precursors, nonaromatic precursors and three chain fatty acids cannot be incorporated by this wild type strain. A summarized overview for all successfully taken chemicals is presented in **Table 4-14**.

Table 4-14 The overview for all successfully incorporated precursors

Nr .	Precursors	Compound	Chemical formula	Predicted [M+H] ⁺	Measured [M+H] ⁺
1a	Benzoic acid	F1	C ₂₅ H ₃₄ N ₄ O ₇	503.57	503.2495
2c	2-fluorobenzoic acid	F2	C ₂₅ H ₃₃ FN ₄ O ₇	521.56	521.2413
3c	3-fluorobenzoic acid	F3	C ₂₅ H ₃₃ FN ₄ O ₇	521.56	521.2400
4c	4-fluorobenzoic acid*	F4	C ₂₅ H ₃₃ FN ₄ O ₇	521.56	521.2413
4d	4-chlorobenzoic acid	F5	C ₂₅ H ₃₃ ClN ₄ O ₇	537.01	537.2104
4e	4-bromobenzoic acid	F6	C ₂₅ H ₃₃ BrN ₄ O ₇	582.46	581.1606
5c	3,5-difluorobenzoic acid	F7	C ₂₅ H ₃₂ F ₂ N ₄ O ₇	539.55	539.2317
5d	2,3-difluorobenzoic acid	F8	C ₂₅ H ₃₂ F ₂ N ₄ O ₇	539.55	539.2319
5e	3,4-difluorobenzoic acid	F9	C ₂₅ H ₃₂ F ₂ N ₄ O ₇	539.55	539.2313
5h	4-amino-3-methylbenzoic acid	F10	C ₂₆ H ₃₇ N ₅ O ₇	532.61	532.2775

*isolated and characterized

4.8 Bioassay of isolated plicacetin and derivatives

Amicetin have sparked the interest recently with their potent anti-mycobacterim activity. Amicetin's binding to the 70S ribosomes of *Thermus thermophilus* has been unambiguously

determined by crystallography and reveals it to occupy the peptidyl transferase center P-site of the ribosome¹⁷⁰. A highly similar compound streptocytosine A has also shown good potential⁷⁴. Lately, Hiroshi *et al.* isolated six amicetin/plicacetin type compounds which also possessed anti mycobacterium activity¹⁵⁵. Also in same year a group of Chinese researchers reported cytosaminomycin E, which was found to be cytotoxic¹⁵⁶. Based on these evidence and literature surveys we setup an activity assessment in collaboration with Prof. Heike Broetz Oesterhelt. Beside the known compound, several newly isolated plicacetin derivatives were tested against various gram positive and negative strains. Cytotoxicity was also assessed (**Figure 4-39**).

	MIC in µg/ml						
	Plicacetin	4F-Plicacetin	Norplicacetin	Compound XII	Compound XI	Known 499	Known 481
<i>Enterococcus faecium</i> BM4147-1	>32	>32	>32	>32	>32	>32	>32
<i>Staphylococcus aureus</i> ATCC29213	>32 (32)	>32	>32	>32	>32	16 (8)	>32
<i>Klebsiella pneumoniae</i> ATCC12657	>32	>32	>32	>32	>32	>32	>32
<i>Acinetobacter baumannii</i> 09987	>32	>32	>32	>32	>32	>32	>32
<i>Pseudomonas aeruginosa</i> ATCC27853	>32	>32	>32	>32	>32	>32 (32)	>32
<i>Enterobacter aerogenes</i> ATCC13048	>32	>32	>32	>32	>32	>32	>32
<i>Escherichia coli</i> ATCC25922	>32	>32	>32	>32	>32	>32	>32
<i>Bacillus subtilis</i> 168	32 (16)	>32 (32)	>32 (32)	>32	>32	8	32
<i>Staphylococcus aureus</i> NCTC8325	>32 (16)	>32 (32)	>32 (32)	>32	>32	16 (8)	>32 (32)
<i>Mycobacterium smegmatis</i> mc ² 155	2 (1)	8 (2)	>32 (32)	2	16	4 (2)	4 (2)
	IC ₅₀ in µg/ml						
Hela	16 - 32	>64	>64	16 - 32	>64	1 - 2	2 - 4

Figure 4-39 Activity profile for plicacetin and derivatives (Green – high activity, parrot green – moderate activity, yellow – low activity)

4.9 Overall discussion

The biosynthetic gene cluster coding for amicetin in *Streptomyces* sp. strain SHP 22-7 only demonstrated a 70% similarity with the cluster in *Streptomyces vinaceusdrappus* NRRL 2363. All corresponding genes showed more than 98% identity. Hence, the remaining 30% differences was located in another five genes, which are not detected in *Streptomyces* sp. strain SHP 22-7 by existing sequencing. Therefore, having a good quality of sequencing and less contigs could definitely give more confidence during careful genome mining. Amicetin-type antibiotics have been isolated during 1950s. Nevertheless, none of them is approved to be used in clinical practice

so far, since further biological studies are suspended due to a lack of quantities, which results from difficult chemical synthesis¹⁷¹. One of the big disadvantages of this class of molecules is their instability. In addition, activity analysis clearly pointed towards both antimicrobial activity and cytotoxicity. A potential drug having both of these activities might face some problem in the selection process. This could explain that despite their prolific activity and ubiquitous production why they were not considered for clinical trials. During our isolation, we have faced a serious issue of spontaneous compound degradation. Using the sodium phosphate buffer was counterproductive, though it had slightly minimized degradation, because at the end desalting step for required if compound needs an activity analysis. Recently, researchers at University of Utah have performed a synthesis attempt to stabilize amicetin and they were found to be successful.¹⁷⁰

Amicetin/plicacetin are produced among their several derivatives. We reported the characterization of altogether 21 derivatives. Masses between the range of 480-688 Da (> 21 derivatives) eluting at a narrow chromatogram had negative effects on the purity of the collected substance. This was because of very subtle side chain modifications that are given within the group. Even after obtaining a pure compound, it went through series of further purification steps (desalting, fat contamination, degradation products etc.). This also results in loss of substance along with time. Likewise, reproducibility was another huge challenge while working with this wild type strain. The amount of specific compounds within different batches were always fluctuating. But still the effort was worth because amicetin-based compounds exhibited good activities. More derivatives towards amicetin might give better understanding of the binding domains. During our separation we have also reported mass 645.234 Da. We speculate that this mass represents a methylated derivative of streptcytosine A or is generated by a demethylation event on 40551-H (see **Figure 4-40**).

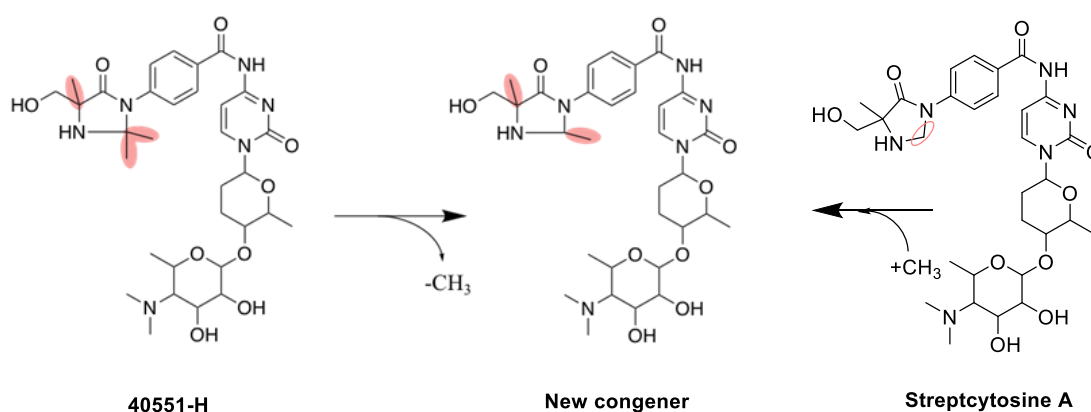


Figure 4-40 Possible methylation and demethylation of known compounds to produce new congeners

Additionally, Compound XII shared a very similar structure to that of 40551-D, where N-methylbutyramide exhibited an extended side chain. In our case, it is seemed to be flipped. Instead of iso-methyl we have side chain with butan-2-one. Interestingly, compound 40551-D does not show any inhibition towards *Mycobacterium smegmatis* MC² 155, but our compound XII shows substantial inhibition. More studies are required in this direction to find how the compound binds to its target. One could speculate the exposure of oxygen atom might provide the possibility for establishing a further hydrogen bond with the target.

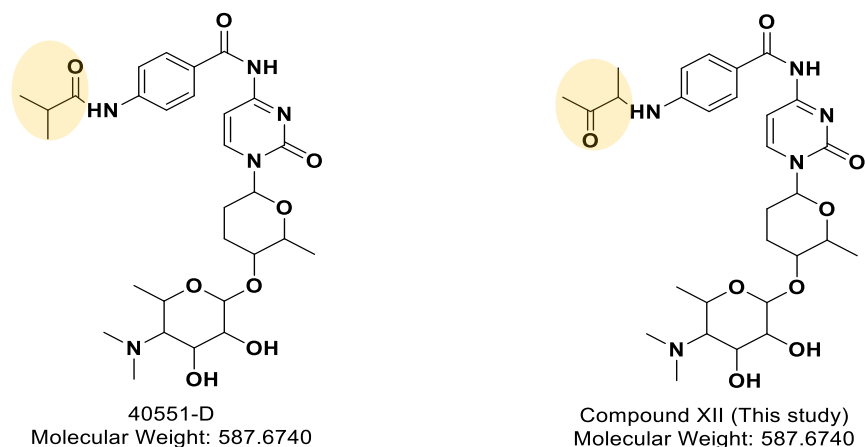


Figure 4-41 Structural comparison of 40551-D¹⁵⁵ and Compound XII

Moreover, during our isolation attempts, there were structures that just carried a methylation event or demethylation event as shown in **Figure 4-40**. Similar demethylation in the side chain was observed also for compound XII. Despite numerous attempts, the purification of such minor peaks was difficult particularly given that the compound has degradation tendencies. But, with partial NMR data and very convincing MS/MS data, we here speculate on their structure (data not presented).

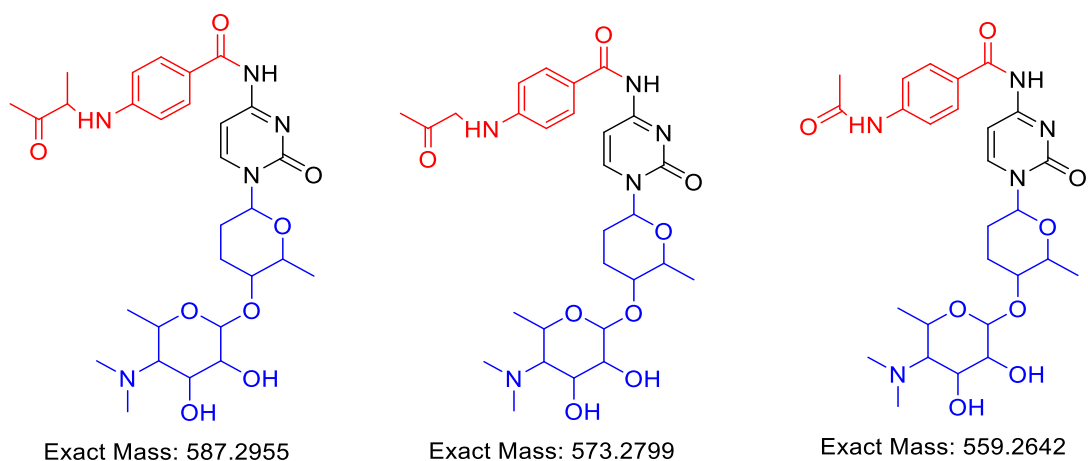


Figure 4-42 Tentative structures for compound XII congeners

Except for streptocytosine A, no activity has been reported for other streptocytosines. While streptocytosine A has been found with a disaccharide unit, streptocytosines B-E seem to lack an amosamine moiety. Loss of amosamine has a substantial effect in activity of these molecules. Similarly, streptocytosines G-O were isolated later and characterized which were also found to be inactive.¹⁵⁶ Streptocytosines isolated in this study were also found to be inactive. This highlights already the importance of the amosamine moiety in the structure. Recently, Serrano *et. al.* found supportive evidence on this. Previously, it was assumed that the methyl serine moiety is crucial for activity¹⁷². However, a recent study contradicts with this finding and disproves it. Serrano *et. al.* has performed binding studies which prevails the importance of the amino-hexopyranose in amicetin activity¹⁷⁰.

One of the vital question is “Why were so many derivatives observed with this strain?”.

Two answers are possible to this question. One possibility lies in the genome sequence itself, which is, if there are some possible transferases present in the amicetin gene cluster. Upon re-skimming of the available genomic data, no such evidence was found. This opens up the possibility for another argument. Production of many derivatives also indicated that the bacteria has a broad substrate specificity. With a rich culture medium and a diverse primary pool, the uptake of these branched chain could be possible. This substrate flexibility was also used at a later stage of our study to generate unnatural derivatives by PDB approach. It is one of the easiest techniques to generate derivatives with specific functionality changes. However, this vastly depends on the bacteria and their enzyme promiscuity. There are several examples where such an approach has also failed. Nonactic acid, a monomer of nonactin, was synthesized and fed to the *Streptomyces griseus* culture, which in turn, was not incorporated to the final product. Whereas, it was found to be toxic to the bacteria with $IC_{50} = 100 \mu\text{m}$.¹⁷³

In parallel, there are many successful examples as well. In two different studies conducted by the Nett Lab, TU Dortmund, successful production of halogenated aurachins, pseudochelins and myxochelins were observed¹⁷⁴ from various *Myxobacteria*. Furthermore, their structure activity relationship was also studied. This kind to study not only generates new unnatural drug candidates but also provides insights to their biosynthetic route.

In this study, precursor-directed biosynthesis was carried out to enrich further structure diversity and increase selectivity. In our case, fluorobenzoic acid very easily taken as it has stronger electron withdrawing effect and a less steric hindrance, which results from its smaller molecule size compared to that of chlorine and bromine. Through the acceptance for all difluoro substituted precursors, the electron withdrawing effect can be verified further. On the contrary, all halogenated precursors in para-position were accepted due to para-position containing a

lower steric effect. Additionally, 4-amino-3-methylbenzoic acid was taken in on account of the similarity with 4-amino benzoic acid. Therefore, the enzyme in the wild type strain has a lower tolerance and flexibility.

While precursor-directed biosynthesis brings us lots of benefits, several bottlenecks can be speculated. The culture scale is the main limitation due to the strict culture conditions and the high contamination possibility in *Streptomyces* strains. In addition, the absorptivity and conversion efficiency of precursors to products are low, due to the substrate competition. In such cases, an advanced isolation method might be explored further to get higher yield.

References

1. Waksman, S. A., The Microbe—Friend and Enemy of Man. *Chemical & Engineering News Archive* **1946**, 24, (10), 1372-1374.
2. World Health Organization. (2017, February 27). "WHO publishes list of bacteria for which new antibiotics are urgently needed" Retrieved from <https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed>.
3. Dadgostar, P., Antimicrobial Resistance: Implications and Costs. *Infect Drug Resist* **2019**, 12, 3903-3910.
4. de Miguel, T.; Rama, J. L. R.; Sieiro, C.; Sánchez, S.; Villa, T. G., Bacteriophages and Lysins as Possible Alternatives to Treat Antibiotic-Resistant Urinary Tract Infections. *Antibiotics (Basel)* **2020**, 9, (8), 466.
5. Wang, B.; Yao, M.; Lv, L.; Ling, Z.; Li, L., The Human Microbiota in Health and Disease. *Engineering* **2017**, 3, (1), 71-82.
6. Harvey, A. L., Natural products in drug discovery. *Drug Discovery Today* **2008**, 13, (19), 894-901.
7. Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, 83, (3), 770-803.
8. Murphy, K. E.; Sloan, G. F.; Lawhern, G. V.; Volk, G. E.; Shumate, J. T.; Wolfe, A. L., Advances in antibiotic drug discovery: reducing the barriers for antibiotic development. *Future Med. Chem.* **2020**, 12, (22), 2067-2087.
9. Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L., Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* **2007**, 6, (1), 29-40.
10. Mahajan, G. B.; Balachandran, L., Antibacterial agents from actinomycetes - a review. *Front Biosci (Elite Ed)* **2012**, 4, 240-53.
11. Gao, B.; Gupta, R. S., Phylogenetic framework and molecular signatures for the main clades of the phylum Actinobacteria. *Microbiology and molecular biology reviews : MMBR* **2012**, 76, (1), 66-112.
12. Gao, B.; Gupta, R. S., Phylogenetic Framework and Molecular Signatures for the Main Clades of the Phylum Actinobacteria. *Microbiol. Mol. Biol. Rev.* **2012**, 76, (1), 66-112.
13. Wolfgang Ludwig, J. E., Peter Schumann, Hans-Jürgen Busse, Martha E. Trujillo, Peter Kämpfer, William B. Whitman, *Road map of the phylum Actinobacteria. In: Goodfellow M. et al. (eds) Bergey's Manual® of Systematic Bacteriology*. Springer: New York, NY., 2012.
14. Prudence, S. M. M.; Addington, E.; Castaño-Espriu, L.; Mark, D. R.; Pintor-Escobar, L.; Russell, A. H.; McLean, T. C., Advances in actinomycete research: an ActinoBase review of 2019. *Microbiology* **2020**, 166, (8), 683-694.
15. van Wezel, G. P.; Krabben, P.; Traag, B. A.; Keijser, B. J. F.; Kerste, R.; Vijgenboom, E.; Heijnen, J. J.; Kraal, B., Unlocking Streptomyces spp. for use as sustainable industrial production platforms by morphological engineering. *Appl. Environ. Microbiol.* **2006**, 72, (8), 5283-5288.
16. Waksman, S. A., On the Classification of Actinomycetes. *J. Bacteriol.* **1940**, 39, (5), 549-558.
17. Crump, A.; Ōmura, S., Ivermectin, 'wonder drug' from Japan: the human use perspective. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **2011**, 87, (2), 13-28.
18. Ding, L.; Münch, J.; Goerls, H.; Maier, A.; Fiebig, H. H.; Lin, W. H.; Hertweck, C., Xiamycin, a pentacyclic indolosesquiterpene with selective anti-HIV activity from a bacterial mangrove endophyte. *Bioorg. Med. Chem. Lett.* **2010**, 20, (22), 6685-7.
19. Ding, L.; Maier, A.; Fiebig, H. H.; Lin, W. H.; Peschel, G.; Hertweck, C., Kandenols A-E, eudesmenes from an endophytic Streptomyces sp. of the mangrove tree Kandelia candel. *J. Nat. Prod.* **2012**, 75, (12), 2223-7.
20. Ding, L.; Goerls, H.; Dornblut, K.; Lin, W.; Maier, A.; Fiebig, H. H.; Hertweck, C., Bacaryolanes A-C, Rare Bacterial Caryolanes from a Mangrove Endophyte. *J. Nat. Prod.* **2015**, 78, (12), 2963-7.
21. Ding, L.; Maier, A.; Fiebig, H. H.; Görls, H.; Lin, W. H.; Peschel, G.; Hertweck, C., Divergolides A-D from a mangrove endophyte reveal an unparalleled plasticity in ansa-macrolide biosynthesis. *Angew. Chem. Int. Ed. Engl.* **2011**, 50, (7), 1630-4.
22. Wibowo, M.; Gotfredsen, C. H.; Sasseti, E.; Melchiorsen, J.; Clausen, M. H.; Gram, L.; Ding, L., Azodyrecins A–C: Azoxides from a Soil-Derived Streptomyces Species. *J. Nat. Prod.* **2020**.
23. Chevrette, M. G.; Carlson, C. M.; Ortega, H. E.; Thomas, C.; Ananiev, G. E.; Barns, K. J.; Book, A. J.; Cagnazzo, J.; Carlos, C.; Flanigan, W.; Grubbs, K. J.; Horn, H. A.; Hoffmann, F. M.; Klassen, J. L.; Knack, J. J.; Lewin, G. R.; McDonald, B. R.; Muller, L.; Melo, W. G. P.; Pinto-Tomás, A. A.; Schmitz, A.; Wendt-Pienkowski, E.; Wildman, S.; Zhao, M.; Zhang, F.; Bugni, T. S.; Andes, D. R.; Pupo, M. T.; Currie, C. R., The antimicrobial potential of Streptomyces from insect microbiomes. *Nature Communications* **2019**, 10, (1), 516.

References

24. de Lima Procópio, R. E.; da Silva, I. R.; Martins, M. K.; de Azevedo, J. L.; de Araújo, J. M., Antibiotics produced by Streptomyces. *The Brazilian Journal of Infectious Diseases* **2012**, 16, (5), 466-471.
25. Jose, P. A.; Sivakala, K. K.; Jha, B., Chapter 11 - Non-Streptomyces Actinomycetes and Natural Products: Recent Updates. In *Studies in Natural Products Chemistry*, Atta ur, R., Ed. Elsevier: 2019; Vol. 61, pp 395-409.
26. Mohammadipanah, F.; Wink, J., Actinobacteria from Arid and Desert Habitats: Diversity and Biological Activity. *Front. Microbiol.* **2016**, 6, (1541).
27. Beemelmans, C.; Guo, H.; Rischer, M.; Poulsen, M., Natural products from microbes associated with insects. *Beilstein J. Org. Chem.* **2016**, 12, 314-327.
28. Skropeta, D.; Wei, L., Recent advances in deep-sea natural products. *Nat. Prod. Rep.* **2014**, 31, (8), 999-1025.
29. Dalmastrì, C.; Gastaldo, L.; Marcone, G. L.; Binda, E.; Congiu, T.; Marinelli, F., Classification of *Nonomuraea* sp. ATCC 39727, an actinomycete that produces the glycopeptide antibiotic A40926, as *Nonomuraea gerenzanensis* sp. nov. *Int. J. Syst. Evol. Microbiol.* **2016**, 66, (2), 912-921.
30. Weinstein, M. J.; Luedemann, G. M.; Oden, E. M.; Wagman, G. H.; Rosselet, J. P.; Marquez, J. A.; Coniglio, C. T.; Charney, W.; Herzog, H. L.; Black, J., Gentamicin, I a New Antibiotic Complex from *Micromonospora*. *J. Med. Chem.* **1963**, 6, (4), 463-464.
31. Handayani, I.; Ratnakomala, S.; Lisdiyanti, P.; Fahrurrozi; Kusharyoto, W.; Alanjary, M.; Ort-Winklbauer, R.; Kulik, A.; Wohlleben, W.; Mast, Y., Complete Genome Sequence of *Streptomyces* sp. Strain SHP22-7, a New Species Isolated from Mangrove of Enggano Island, Indonesia. *Microbiology Resource Announcements* **2018**, 7, (20), e01317-18.
32. Krause, J.; Handayani, I.; Blin, K.; Kulik, A.; Mast, Y., Disclosing the Potential of the SARP-Type Regulator PapR2 for the Activation of Antibiotic Gene Clusters in Streptomyces. *Front. Microbiol.* **2020**, 11, (225).
33. Meyer-Rochow, V. B., Therapeutic arthropods and other, largely terrestrial, folk-medicinally important invertebrates: a comparative survey and review. *Journal of Ethnobiology and Ethnomedicine* **2017**, 13, (1), 9.
34. Rodrigues, E.; Lago, J. H.; de FL Santos, J.; Bitencourt, A. L. V., Nests of “caba-leão” wasps (*Sceliphron* sp., Sphecidae) used in traditional medicine by riverine communities of the Jaú and Unini Rivers, Amazon, Brazil: ethnopharmacological, chemical and mineralogical aspects. *Revista Brasileira de Farmacognosia* **2018**, 28, (3), 352-357.
35. Igarashi, Y.; Mogi, T.; Yanase, S.; Miyanaga, S.; Fujita, T.; Sakurai, H.; Saiki, I.; Ohsaki, A., Brartemicin, an inhibitor of tumor cell invasion from the actinomycete *Nonomuraea* sp. *J. Nat. Prod.* **2009**, 72, (5), 980-982.
36. Gross, H., Genomic mining--a concept for the discovery of new bioactive natural products. *Curr Opin Drug Discov Devel* **2009**, 12, (2), 207-19.
37. 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.
38. Mast, Y.; Stegmann, E., Actinomycetes: The Antibiotics Producers. *Antibiotics (Basel)* **2019**, 8, (3), 105.
39. 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.
40. Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Brucoleri, R.; Lee, S. Y.; Fischbach, M. A.; Müller, 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-W243.
41. Kautsar, S. A.; Blin, K.; Shaw, S.; Navarro-Muñoz, J. C.; Terlouw, B. R.; van der Hooft, J. J. J.; van Santen, J. A.; Tracanna, V.; Suarez Duran, H. G.; Pascal Andreu, V.; Selem-Mojica, N.; Alanjary, M.; Robinson, S. L.; Lund, G.; Epstein, S. C.; Sisto, A. C.; Charkoudian, L. K.; Collemare, J.; Lington, R. G.; Weber, T.; Medema, M. H., MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* **2019**, 48, (D1), D454-D458.
42. Kaweewan, I.; Komaki, H.; Hemmi, H.; Hoshino, K.; Hosaka, T.; Isokawa, G.; Oyoshi, T.; Kodani, S., Isolation and structure determination of a new cytotoxic peptide, curacozole, from *Streptomyces curacoi* based on genome mining. *The Journal of Antibiotics* **2019**, 72, (1), 1-7.

References

43. Cortés-Albayay, C.; Jarmusch, S. A.; Trusch, F.; Ebel, R.; Andrews, B. A.; Jaspars, M.; Asenjo, J. A., Downsizing Class II Lasso Peptides: Genome Mining-Guided Isolation of Huascopeptin Containing the First Gly1-Asp7 Macrocyclic. *The Journal of Organic Chemistry* **2020**, *85*, (3), 1661-1667.
44. Krug, D.; Müller, R., Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. *Nat. Prod. Rep.* **2014**, *31*, (6), 768-783.
45. Henke, M. T.; Kelleher, N. L., Modern mass spectrometry for synthetic biology and structure-based discovery of natural products. *Nat. Prod. Rep.* **2016**, *33*, (8), 942-50.
46. Duhrkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Bocker, S., SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat Methods* **2019**, *16*, (4), 299-302.
47. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Oresic, M., MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010**, *11*, 395.
48. Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kaponov, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrew, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya, P., C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. Ø.; Pogliano, K.; Lington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N., Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, (8), 828-837.
49. 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.
50. Lopez, J. A. V.; Nogawa, T.; Futamura, Y.; Shimizu, T.; Osada, H., Nocardamin glucuronide, a new member of the ferrioxamine siderophores isolated from the ascamycin-producing strain *Streptomyces* sp. 80H647. *The Journal of Antibiotics* **2019**, *72*, (12), 991-995.
51. Ariantari, N. P.; Daletos, G.; Mándi, A.; Kurtán, T.; Müller, W. E. G.; Lin, W.; Ancheeva, E.; Proksch, P., Expanding the chemical diversity of an endophytic fungus *Bulgaria inquinans*, an ascomycete associated with mistletoe, through an OSMAC approach. *RSC Advances* **2019**, *9*, (43), 25119-25132.
52. Pham, V. T. T.; Nguyen, H. T.; Nguyen, C. T.; Choi, Y. S.; Dhakal, D.; Kim, T.-S.; Jung, H. J.; Yamaguchi, T.; Sohng, J. K., Identification and enhancing production of a novel macrolide compound in engineered *Streptomyces peucetius*. *RSC Advances* **2021**, *11*, (5), 3168-3173.
53. Zhou, W.; Posri, P.; Abugrain, M. E.; Weisberg, A. J.; Chang, J. H.; Mahmud, T., Biosynthesis of the Nuclear Factor of Activated T Cells Inhibitor NFAT-133 in *Streptomyces pactum*. *ACS Chem. Biol.* **2020**, *15*, (12), 3217-3226.
54. Eida, A. A.; Mahmud, T., The secondary metabolite pactamycin with potential for pharmaceutical applications: biosynthesis and regulation. *Appl. Microbiol. Biotechnol.* **2019**, *103*, (11), 4337-4345.
55. Sheng, Y.; Lam, P. W.; Shahab, S.; Santosa, D. A.; Proteau, P. J.; Zabriskie, T. M.; Mahmud, T., Identification of Elaiophyllin Skeletal Variants from the Indonesian *Streptomyces* sp. ICBB 9297. *J. Nat. Prod.* **2015**, *78*, (11), 2768-2775.
56. Bravo-Monzón, A. E.; Ríos-Vásquez, E.; Delgado-Lamas, G.; Espinosa-García, F. J., Chemical diversity among populations of *Mikania micrantha*: geographic mosaic structure and herbivory. *Oecologia* **2014**, *174*, (1), 195-203.
57. Papke, R. T.; Ramsing, N. B.; Bateson, M. M.; Ward, D. M., Geographical isolation in hot spring cyanobacteria. *Environ. Microbiol.* **2003**, *5*, (8), 650-9.
58. Whitaker, R. J.; Grogan, D. W.; Taylor, J. W., Geographic barriers isolate endemic populations of hyperthermophilic archaea. *Science* **2003**, *301*, (5635), 976-8.

References

59. Keller, M.; Zengler, K., Tapping into microbial diversity. *Nature Reviews Microbiology* **2004**, *2*, (2), 141-150.
60. Naman, C. B.; Leber, C. A.; Gerwick, W. H., Chapter 5 - Modern Natural Products Drug Discovery and Its Relevance to Biodiversity Conservation. In *Microbial Resources*, Kurtböke, I., Ed. Academic Press: 2017; pp 103-120.
61. Sarethy, I. P.; Pan, S.; Danquah, M. K., Modern Taxonomy for Microbial Diversity. In *Biodiversity-The Dynamic Balance of the Planet*, Oscar Grillo, P., Ed. InTech: 2014.
62. Butler, R. A., Top 10 most biodiverse countries. *Mongabay*. Retrieved from <https://news.mongabay.com/2016/05/top-10-biodiverse-countries/>. **2016, 21 May**.
63. González Canga, A.; Sahagún Prieto, A. M.; Diez Liébana, M. J.; Fernández Martínez, N.; Sierra Vega, M.; García Vieitez, J. J., The Pharmacokinetics and Interactions of Ivermectin in Humans—A Mini-review. *The AAPS Journal* **2008**, *10*, (1), 42-46.
64. Van Voorhis, W. C.; Hooft van Huijsduijnen, R.; Wells, T. N. C., Profile of William C. Campbell, Satoshi Ōmura, and Youyou Tu, 2015 Nobel Laureates in Physiology or Medicine. *Proceedings of the National Academy of Sciences* **2015**, *112*, (52), 15773-15776.
65. Carroll, A. R.; Copp, B. R.; Davis, R. A.; Keyzers, R. A.; Prinsep, M. R., Marine natural products. *Nat. Prod. Rep.* **2020**, *37*, (2), 175-223.
66. Kong, D. X.; Jiang, Y. Y.; Zhang, H. Y., Marine natural products as sources of novel scaffolds: achievement and concern. *Drug Discov Today* **2010**, *15*, (21-22), 884-6.
67. Udvary, D. W.; Zeigler, L.; Asolkar, R. N.; Singan, V.; Lapidus, A.; Fenical, W.; Jensen, P. R.; Moore, B. S., Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proceedings of the National Academy of Sciences* **2007**, *104*, (25), 10376-10381.
68. Feling, R. H.; Buchanan, G. O.; Mincer, T. J.; Kauffman, C. A.; Jensen, P. R.; Fenical, W., Salinosporamide A: a highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus *salinispora*. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, (3), 355-7.
69. Buchanan, G. O.; Williams, P. G.; Feling, R. H.; Kauffman, C. A.; Jensen, P. R.; Fenical, W., Sporolides A and B: Structurally Unprecedented Halogenated Macrolides from the Marine Actinomycete *Salinispora tropica*. *Organic Letters* **2005**, *7*, (13), 2731-2734.
70. Aotani, Y.; Nagata, H.; Yoshida, M., Lymphostin (LK6-A), a novel immunosuppressant from *Streptomyces* sp. KY11783: structural elucidation. *J. Antibiot. (Tokyo)* **1997**, *50*, (7), 543-5.
71. Gulder, T. A. M.; Moore, B. S., Chasing the treasures of the sea—bacterial marine natural products. *Curr. Opin. Microbiol.* **2009**, *12*, (3), 252-260.
72. Eustáquio, A. S.; O'Hagan, D.; Moore, B. S., Engineering Fluorometabolite Production: Fluorinase Expression in *Salinispora tropica* Yields Fluorosalinosporamide. *J. Nat. Prod.* **2010**, *73*, (3), 378-382.
73. Fenical, W., Marine microbial natural products: the evolution of a new field of science. *The Journal of Antibiotics* **2020**, *73*, (8), 481-487.
74. Bu, Y. Y.; Yamazaki, H.; Ukai, K.; Namikoshi, M., Anti-mycobacterial nucleoside antibiotics from a marine-derived *Streptomyces* sp. TPU1236A. *Mar. Drugs* **2014**, *12*, (12), 6102-12.
75. Yang, J.; Gao, J.; Cheung, A.; Liu, B.; Schwendenmann, L.; Costello, M. J., Vegetation and sediment characteristics in an expanding mangrove forest in New Zealand. *Estuarine, Coastal and Shelf Science* **2013**, *134*, 11-18.
76. Xu, D.-B.; Ye, W.-W.; Han, Y.; Deng, Z.-X.; Hong, K., Natural products from mangrove actinomycetes. *Mar. Drugs* **2014**, *12*, (5), 2590-2613.
77. Wu, J.; Xiao, Q.; Xu, J.; Li, M.-Y.; Pan, J.-Y.; Yang, M.-h., Natural products from true mangrove flora: source, chemistry and bioactivities. *Nat. Prod. Rep.* **2008**, *25*, (5), 955-981.
78. Kim, S.-H.; Ha, T.-K.-Q.; Oh, W. K.; Shin, J.; Oh, D.-C., Antiviral Indolosesquiterpenoid Xiamycins C–E from a Halophilic Actinomycete. *J. Nat. Prod.* **2016**, *79*, (1), 51-58.
79. Sánchez, C.; Méndez, C.; Salas, J. A., Indolocarbazole natural products: occurrence, biosynthesis, and biological activity. *Nat. Prod. Rep.* **2006**, *23*, (6), 1007-45.
80. Chien, A.; Edgar, D. B.; Trela, J. M., Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J. Bacteriol.* **1976**, *127*, (3), 1550-1557.
81. Omura, S.; Suzuki, Y.; Kitao, C.; Takahashi, Y.; Konda, Y., Isolation of a new sulfur-containing basic substance from a Thermo actinomyces species. *J. Antibiot. (Tokyo)* **1975**, *28*, (8), 609-10.
82. Silva, L. J.; Crevelin, E. J.; Souza, D. T.; Lacerda-Júnior, G. V.; de Oliveira, V. M.; Ruiz, A. L. T. G.; Rosa, L. H.; Moraes, L. A. B.; Melo, I. S., Actinobacteria from Antarctica as a source for anticancer discovery. *Sci. Rep.* **2020**, *10*, (1), 13870.
83. Oh, D.-C.; Poulsen, M.; Currie, C. R.; Clardy, J., Dentigerumycin: a bacterial mediator of an antifungal symbiosis. *Nat. Chem. Biol.* **2009**, *5*, (6), 391-393.

References

84. Van Arnam, E. B.; Ruzzini, A. C.; Sit, C. S.; Horn, H.; Pinto-Tomás, A. A.; Currie, C. R.; Clardy, J., Selvamycin, an atypical antifungal polyene from two alternative genomic contexts. *Proceedings of the National Academy of Sciences* **2016**, 113, (46), 12940-12945.
85. Kroiss, J.; Kaltenpoth, M.; Schneider, B.; Schwinger, M. G.; Hertweck, C.; Maddula, R. K.; Strohm, E.; Svatos, A., Symbiotic Streptomycetes provide antibiotic combination prophylaxis for wasp offspring. *Nat. Chem. Biol.* **2010**, 6, (4), 261-3.
86. Scott, J. J.; Oh, D. C.; Yuceer, M. C.; Klepzig, K. D.; Clardy, J.; Currie, C. R., Bacterial protection of beetle-fungus mutualism. *Science* **2008**, 322, (5898), 63.
87. Milshteyn, A.; Colosimo, D. A.; Brady, S. F., Accessing Bioactive Natural Products from the Human Microbiome. *Cell Host Microbe* **2018**, 23, (6), 725-736.
88. Herbrík, A.; Corretto, E.; Chroňáková, A.; Langhansová, H.; Petrásková, P.; Hrdý, J.; Čihák, M.; Křišťůfek, V.; Bobek, J.; Petříček, M.; Petříčková, K., A Human Lung-Associated Streptomyces sp. TR1341 Produces Various Secondary Metabolites Responsible for Virulence, Cytotoxicity and Modulation of Immune Response. *Front. Microbiol.* **2020**, 10, (3028).
89. Li, J.; Li, Y.; Zhou, Y.; Wang, C.; Wu, B.; Wan, J., Actinomyces and Alimentary Tract Diseases: A Review of Its Biological Functions and Pathology. *Biomed Res Int* **2018**, 2018, 3820215-3820215.
90. Bolourian, A.; Mojtahedi, Z., Streptomycetes, shared microbiome member of soil and gut, as 'old friends' against colon cancer. *FEMS Microbiol. Ecol.* **2018**, 94, (8).
91. Hayakawa, M.; Nonomura, H., Humic acid-vitamin agar, a new medium for the selective isolation of soil actinomycetes. *Journal of Fermentation Technology* **1987**, 65, (5), 501-509.
92. Shepherd, M. D.; Kharel, M. K.; Bosserman, M. A.; Rohr, J., Laboratory maintenance of Streptomyces species. *Curr. Protoc. Microbiol.* **2010**, Chapter 10, Unit 10E.1.
93. Duetz, W. A.; Witholt, B., Oxygen transfer by orbital shaking of square vessels and deepwell microtiter plates of various dimensions. *Biochem. Eng. J.* **2004**, 17, (3), 181-185.
94. Siebenberg, S.; Bapat, P. M.; Lantz, A. E.; Gust, B.; Heide, L., Reducing the variability of antibiotic production in Streptomyces by cultivation in 24-square deepwell plates. *J. Biosci. Bioeng.* **2010**, 109, (3), 230-4.
95. National Committee for Clinical Laboratory Standards. Methods for Dilution Anti-microbial Susceptibility Tests for Bacteria that Grow Aerobically. Approved Standards. NCCLS Document M7-A4 4th ed. 1997.
96. Gontang, E. A.; Fenical, W.; Jensen, P. R., Phylogenetic diversity of gram-positive bacteria cultured from marine sediments. *Appl. Environ. Microbiol.* **2007**, 73, (10), 3272-82.
97. Wolfender, J. L., HPLC in natural product analysis: the detection issue. *Planta Med.* **2009**, 75, (7), 719-34.
98. Garg, N.; Kapon, C.; Lim, Y. W.; Koyama, N.; Vermeij, M. J.; Conrad, D.; Rohwer, F.; Dorrestein, P. C., Mass spectral similarity for untargeted metabolomics data analysis of complex mixtures. *Int J Mass Spectrom* **2015**, 377, 719-717.
99. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, 13, (11), 2498-504.
100. Wu, Y.; Sun, Y.-P., Synthesis of Nonactin and the Proposed Structure of Trilactone. *Organic Letters* **2006**, 8, (13), 2831-2834.
101. Cox, J. E.; Priestley, N. D., Nonactin Biosynthesis: The Product of the Resistance Gene Degrades Nonactin Stereospecifically to Form Homochiral Nonactate Dimers. *J. Am. Chem. Soc.* **2005**, 127, (22), 7976-7977.
102. Woo, A. J.; Strohl, W. R.; Priestley, N. D., Nonactin biosynthesis: the product of nonS catalyzes the formation of the furan ring of nonactic acid. *Antimicrob. Agents Chemother.* **1999**, 43, (7), 1662-1668.
103. Řezanka, T.; Spížek, J.; Přikrylová, V.; Prell, A.; Dembitsky, V. M., Five new derivatives of nonactic and homo-nonactic acids from Streptomyces globisporus. *Tetrahedron* **2004**, 60, (22), 4781-4787.
104. Fleck, W. F.; Ritzau, M.; Heinze, S.; Gräfe, U., Isolation of dimeric nonactic acid from the nonactin-producing Streptomyces spec. JA 5909-1. *J. Basic Microbiol.* **1996**, 36, (4), 235-238.
105. Ortega, H. E.; Batista, J. M., Jr.; Melo, W. G. P.; Clardy, J.; Pupo, M. T., Absolute Configurations of Griseorhodins A and C. *Tetrahedron Lett.* **2017**, 58, (50), 4721-4723.
106. Atkinson, D. J.; Brimble, M. A., Isolation, biological activity, biosynthesis and synthetic studies towards the rubromycin family of natural products. *Nat. Prod. Rep.* **2015**, 32, (6), 811-40.
107. Stroshane, R. M.; Chan, J. A.; Rubalcaba, E. A.; Garretson, A. L.; Aszalos, A. A.; Roller, P. P., Isolation and structure elucidation of a novel griseorhodin. *J. Antibiot. (Tokyo)* **1979**, 32, (3), 197-204.
108. Li, A.; Piel, J., A Gene Cluster from a Marine Streptomyces Encoding the Biosynthesis of the Aromatic Spiroketal Polyketide Griseorhodin A. *Chem. Biol.* **2002**, 9, (9), 1017-1026.

References

109. Lin, Z.; Zachariah, M. M.; Marett, L.; Hughen, R. W.; Teichert, R. W.; Concepcion, G. P.; Haygood, M. G.; Olivera, B. M.; Light, A. R.; Schmidt, E. W., Griseorhodins D–F, Neuroactive Intermediates and End Products of Post-PKS Tailoring Modification in Griseorhodin Biosynthesis. *J. Nat. Prod.* **2014**, *77*, (5), 1224-1230.
110. Ueno, T.; Takahashi, H.; Oda, M.; Mizunuma, M.; Yokoyama, A.; Goto, Y.; Mizushima, Y.; Sakaguchi, K.; Hayashi, H., Inhibition of Human Telomerase by Rubromycins: Implication of Spiroketal System of the Compounds as an Active Moiety. *Biochemistry* **2000**, *39*, (20), 5995-6002.
111. Kohi, S.; Sato, N.; Koga, A.; Hirata, K.; Harunari, E.; Igarashi, Y., Hyaluronidase Inhibitor, Attenuates Pancreatic Cancer Cell Migration and Proliferation. *J. Oncol.* **2016**, *2016*, 9063087-9063087.
112. van Santen, J. A.; Jacob, G.; Singh, A. L.; Aniebok, V.; Balunas, M. J.; Bunsko, D.; Neto, F. C.; Castaño-Espriu, L.; Chang, C.; Clark, T. N.; Cleary Little, J. L.; Delgadillo, D. A.; Dorrestein, P. C.; Duncan, K. R.; Egan, J. M.; Galey, M. M.; Haeckl, F. P. J.; Hua, A.; Hughes, A. H.; Iskakova, D.; Khadilkar, A.; Lee, J.-H.; Lee, S.; LeGrow, N.; Liu, D. Y.; Macho, J. M.; McCaughey, C. S.; Medema, M. H.; Neupane, R. P.; O'Donnell, T. J.; Paula, J. S.; Sanchez, L. M.; Shaikh, A. F.; Soldatou, S.; Terlouw, B. R.; Tran, T. A.; Valentine, M.; van der Hooft, J. J. J.; Vo, D. A.; Wang, M.; Wilson, D.; Zink, K. E.; Linington, R. G., The Natural Products Atlas: An Open Access Knowledge Base for Microbial Natural Products Discovery. *ACS Central Science* **2019**, *5*, (11), 1824-1833.
113. Vater, J.; Crnovčić, I.; Semsary, S.; Keller, U., MALDI-TOF mass spectrometry, an efficient technique for in situ detection and characterization of actinomycins. *J. Mass Spectrom.* **2014**, *49*, (3), 210-222.
114. Koba, M.; Konopa, J., [Actinomycin D and its mechanisms of action]. *Postepy Hig Med Dosw (Online)* **2005**, *59*, 290-8.
115. Thomas, D.; Morris, M.; Curtis, J. M.; Boyd, R. K., Fragmentation mechanisms of protonated actinomycins and their use in structural determination of unknown analogues. *J. Mass Spectrom.* **1995**, *30*, (8), 1111-1125.
116. Terry, H. Slender mud-dauber wasps: genus sceliphron. <http://museum.wa.gov.au/research/collections/terrestrial-zoology/entomology-insect-collection/entomology-factsheets/sceliphron>
117. Rodrigues, E.; Lago, J. H. G.; Santos, J. d. F. L.; Bitencourt, A. L. V., Nests of caba-leão wasps (*Sceliphron* sp., Sphecidae) used in traditional medicine by riverine communities of the Jaú and Unini Rivers, Amazon, Brazil: ethnopharmacological, chemical and mineralogical aspects. *Revista Brasileira de Farmacognosia* **2018**, *28*, 352-357.
118. Kumar, V.; Bharti, A.; Gupta, V. K.; Gusain, O.; Bisht, G. S., Actinomycetes from solitary wasp mud nest and swallow bird mud nest: isolation and screening for their antibacterial activity. *World Journal of Microbiology and Biotechnology* **2012**, *28*, (3), 871-880.
119. Kumar, V.; Naik, B.; Gusain, O.; Bisht, G. S., An actinomycete isolate from solitary wasp mud nest having strong antibacterial activity and kills the *Candida* cells due to the shrinkage and the cytosolic loss. *Front. Microbiol.* **2014**, *5*, (446).
120. Michael Goodfellow, e. v.; William B. Whitman, d. o. t. e. o.; Aidan C. Parte, m. e., *Bergey's manual of systematic bacteriology. Volume 5, The actinobacteria*. Second edition. New York : Springer, [2012] ©2012: 2012.
121. Aryal, N.; Aziz, S.; Rajbhandari, P.; Gross, H., Draft Genome Sequence of *Nonomuraea* sp. Strain C10, a Producer of Brartemicin, Isolated from a Mud Dauber Wasp Nest in Nepal. *Microbiol Resour Announc* **2019**, *8*, (45).
122. Tatusova, T.; DiCuccio, M.; Badretdin, A.; Chetvernin, V.; Nawrocki, E. P.; Zaslavsky, L.; Lomsadze, A.; Pruitt, K. D.; Borodovsky, M.; Ostell, J., NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **2016**, *44*, (14), 6614-24.
123. Meier-Kolthoff, J. P.; Göker, M., TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications* **2019**, *10*, (1), 2182.
124. Liu, Z. P.; Wu, J. F.; Liu, Z. H.; Liu, S. J., *Pseudonocardia ammonioxydans* sp. nov., isolated from coastal sediment. *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, (Pt 3), 555-558.
125. Alanjary, M.; Steinke, K.; Ziemert, N., AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res.* **2019**, *47*, (W1), W276-w282.
126. Nazari, B.; Forneris, C. C.; Gibson, M. I.; Moon, K.; Schramma, K. R.; Seyedsayamdost, M. R., *Nonomuraea* sp. ATCC 55076 harbours the largest actinomycete chromosome to date and the kistamicin biosynthetic gene cluster. *MedChemComm* **2017**, *8*, (4), 780-788.
127. Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T., antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **2019**, *47*, (W1), W81-w87.

References

128. Bursy, J.; Kuhlmann, A. U.; Pittelkow, M.; Hartmann, H.; Jebbar, M.; Pierik, A. J.; Bremer, E., Synthesis and Uptake of the Compatible Solutes Ectoine and 5-Hydroxyectoine by *Streptomyces coelicolor* A3(2) in Response to Salt and Heat Stresses. *Appl. Environ. Microbiol.* **2008**, *74*, (23), 7286-7296.
129. Taura, F.; Iijima, M.; Yamanaka, E.; Takahashi, H.; Kenmoku, H.; Saeki, H.; Morimoto, S.; Asakawa, Y.; Kurosaki, F.; Morita, H., A Novel Class of Plant Type III Polyketide Synthase Involved in Orsellinic Acid Biosynthesis from *Rhododendron dauricum*. *Frontiers in Plant Science* **2016**, *7*, (1452).
130. Weitnauer, G.; Mühlenweg, A.; Trefzer, A.; Hoffmeister, D.; Süßmuth, R. D.; Jung, G.; Welzel, K.; Vente, A.; Girreser, U.; Bechthold, A., Biosynthesis of the orthosomycin antibiotic avilamycin A: deductions from the molecular analysis of the *avi* biosynthetic gene cluster of *Streptomyces viridochromogenes* Tü57 and production of new antibiotics. *Chem. Biol.* **2001**, *8*, (6), 569-81.
131. Liu, W.; Ahlert, J.; Gao, Q.; Wendt-Pienkowski, E.; Shen, B.; Thorson, J. S., Rapid PCR amplification of minimal enediyne polyketide synthase cassettes leads to a predictive familial classification model. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, (21), 11959-63.
132. Gyobu, Y.; Miyadoh, S., Proposal to transfer *Actinomadura carminata* to a new subspecies of the genus *Nonomuraea* as *Nonomuraea roseoviolacea* subsp. *carminata* comb. nov. *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, (3), 881-889.
133. Jacobsen, K. M.; Keiding, U. B.; Clement, L. L.; Schaffert, E. S.; Rambaruth, N. D. S.; Johannsen, M.; Drickamer, K.; Poulsen, T. B., The natural product brartemicin is a high affinity ligand for the carbohydrate-recognition domain of the macrophage receptor mincle. *MedChemComm* **2015**, *6*, (4), 647-652.
134. Sanchez, J. F.; Chiang, Y.-M.; Szewczyk, E.; Davidson, A. D.; Ahuja, M.; Elizabeth Oakley, C.; Woo Bok, J.; Keller, N.; Oakley, B. R.; Wang, C. C. C., Molecular genetic analysis of the orsellinic acid/F9775 gene cluster of *Aspergillus nidulans*. *Mol. Biosyst.* **2010**, *6*, (3), 587-593.
135. Braesel, J.; Fricke, J.; Schwenk, D.; Hoffmeister, D., Biochemical and genetic basis of orsellinic acid biosynthesis and prenylation in a stereaceous basidiomycete. *Fungal Genet. Biol.* **2017**, *98*, 12-19.
136. Yu, P.-W.; Cho, T.-Y.; Liou, R.-F.; Tzean, S.-S.; Lee, T.-H., Identification of the orsellinic acid synthase PKS63787 for the biosynthesis of antroquinonols in *Antrodia cinnamomea*. *Appl. Microbiol. Biotechnol.* **2017**, *101*, (11), 4701-4711.
137. Jørgensen, S. H.; Frandsen, R. J. N.; Nielsen, K. F.; Lysøe, E.; Sondergaard, T. E.; Wimmer, R.; Giese, H.; Sørensen, J. L., *Fusarium graminearum* PKS14 is involved in orsellinic acid and orcinol synthesis. *Fungal Genet. Biol.* **2014**, *70*, 24-31.
138. Weitnauer, G.; Mühlenweg, A.; Trefzer, A.; Hoffmeister, D.; Süßmuth, R. D.; Jung, G.; Welzel, K.; Vente, A.; Girreser, U.; Bechthold, A., Biosynthesis of the orthosomycin antibiotic avilamycin A: deductions from the molecular analysis of the *avi* biosynthetic gene cluster of *Streptomyces viridochromogenes* Tü57 and production of new antibiotics. *Chem. Biol.* **2001**, *8*, (6), 569-581.
139. Tuan, N. T.; Dam, N. P.; Van Hieu, M.; Trang, D. T. X.; Danh, L. T.; Men, T. T.; De, T. Q.; Bach, L. T.; Kanaori, K., Chemical Constituents of the Lichen *Parmotrema tinctorum* and their Antifungal Activity. *Chemistry of Natural Compounds* **2020**, *56*, (2), 315-317.
140. Lopes, T. I. B.; Coelho, R. G.; Honda, N. K., Inhibition of Mushroom Tyrosinase Activity by Orsellinates. *Chem. Pharm. Bull. (Tokyo)* **2018**, *66*, (1), 61-64.
141. Lampis, G.; Deidda, D.; Maullu, C.; Madeddu, M. A.; Pompei, R.; Delle Monachie, F.; Satta, G., Sattabacins and sattazolins: new biologically active compounds with antiviral properties extracted from a *Bacillus* sp. *J. Antibiot. (Tokyo)* **1995**, *48*, (9), 967-72.
142. Li, J.; Chen, G.; Webster, J. M.; Czyzewska, E., Antimicrobial metabolites from a bacterial symbiont. *J. Nat. Prod.* **1995**, *58*, (7), 1081-6.
143. Schieferdecker, S.; Shabuer, G.; Letzel, A.-C.; Urbansky, B.; Ishida-Ito, M.; Ishida, K.; Cyruilies, M.; Dahse, H.-M.; Pidot, S.; Hertweck, C., Biosynthesis of Diverse Antimicrobial and Antiproliferative Acyloins in Anaerobic Bacteria. *ACS Chem. Biol.* **2019**, *14*, (7), 1490-1497.
144. Genilloud, O., Actinomycetes: still a source of novel antibiotics. *Nat. Prod. Rep.* **2017**, *34*, (10), 1203-1232.
145. Pandey, B.; Ghimire, P.; Agrawal, V. P., Studies on the antibacterial activity of the Actinomycetes isolated from the Khumbu Region of Nepal. *J. Biol. Sci.* **2004**, *23*, 44-53.
146. Yadav, J.; Shrestha, U. T.; Tiwari, K. B.; Sahukhal, G. S.; Agrawal, V. P., Streptomycin-like antibiotic from *Streptomyces* spp. isolated from Mount Everest base camp. *Nepal Journal of Science and Technology* **2008**, *9*, 73-77.
147. Gurung, T. D.; Sherpa, C.; Agrawal, V. P.; Lekhak, B., Isolation and characterization of antibacterial actinomycetes from soil samples of Kalapatthar, Mount Everest Region. *Nepal Journal of Science and Technology* **2009**, *10*, 173-182.

References

148. Schumacher, R. W.; Talmage, S. C.; Miller, S. A.; Sarris, K. E.; Davidson, B. S.; Goldberg, A., Isolation and Structure Determination of an Antimicrobial Ester from a Marine Sediment-Derived Bacterium. *J. Nat. Prod.* **2003**, 66, (9), 1291-1293.
149. Sherwood, E. J.; Bibb, M. J., The antibiotic planosporicin coordinates its own production in the actinomycete *Planomonospora alba*. *Proceedings of the National Academy of Sciences* **2013**, 110, (27), E2500-E2509.
150. Agrawal, P.; Khater, S.; Gupta, M.; Sain, N.; Mohanty, D., RiPPMiner: a bioinformatics resource for deciphering chemical structures of RiPPs based on prediction of cleavage and cross-links. *Nucleic Acids Res.* **2017**, 45, (W1), W80-w88.
151. Handayani, I.; Saad, H.; Ratnakomala, S.; Lisdiyanti, P.; Kusharyoto, W.; Krause, J.; Kulik, A.; Wohlleben, W.; Aziz, S.; Gross, H.; Gavriilidou, A.; Ziemert, N.; Mast, Y., Mining Indonesian Microbial Biodiversity for Novel Natural Compounds by a Combined Genome Mining and Molecular Networking Approach. *Mar. Drugs* **2021**, 19, (6), 316.
152. Aksoy, S.; Uzel, A.; Bedir, E., Cytosine-type nucleosides from marine-derived *Streptomyces rochei* 06CM016. *J. Antibiot. (Tokyo)* **2016**, 69, (1), 51-6.
153. Fu, J.; Laval, S.; Yu, B., Total Synthesis of Nucleoside Antibiotics Plicacetin and Streptocytosine A. *The Journal of Organic Chemistry* **2018**, 83, (13), 7076-7084.
154. Kate Abhijeet Sudhir, G. S. D., Sonawane Sailendra , Periyasamy Giridharan, , WO2013/144894A1. **2013**.
155. Hiroshi. Tomoda, N. K., Akihiko. Kanamoto, Junko. Hashimoto, Ikuko. Kozone., WO 2019/044941. **2019**.
156. Xu, C.-D.; Zhang, H.-J.; Ma, Z.-J., Pyrimidine Nucleosides from *Streptomyces* sp. SSA28. *J. Nat. Prod.* **2019**, 82, (9), 2509-2516.
157. Flynn, E. H.; Hinman, J.; Caron, E.; Woolf Jr, D., The Chemistry of Amicetin, a New Antibiotic1, 2. *J. Am. Chem. Soc.* **1953**, 75, (23), 5867-5871.
158. Itoh, J.; Miyadoh, S., SF2457, a new antibiotic related to amicetin. *J. Antibiot. (Tokyo)* **1992**, 45, (6), 846-53.
159. Haskell, T. H.; Ryder, A.; Frohardt, R. P.; Fusari, S. A.; Jakubowski, Z. L.; Bartz, Q. R., The isolation and characterization of three crystalline antibiotics from *Streptomyces plicatus*. *J. Am. Chem. Soc.* **1958**, 80, (3), 743-747.
160. Shammas, C.; Donarski, J. A.; Ramesh, V., NMR structure of the peptidyl transferase RNA inhibitor antibiotic amicetin. *Magn. Reson. Chem.* **2007**, 45, (2), 133-141.
161. Tomita, K.; Uenoyama, Y.; Fujisawa, K.; Kawaguchi, H., Oxamicetin, a new antibiotic of bacterial origin. 3. Taxonomy of the oxamicetin-producing organism. *J. Antibiot. (Tokyo)* **1973**, 26, (12), 765-70.
162. DeBoer, C.; Caron, E. L.; Hinman, J., Amicetin, a new streptomyces antibiotic. *J. Am. Chem. Soc.* **1953**, 75, (2), 499-500.
163. Evans, J.; Weare, G., Norplicetin, a new antibiotic from *Streptomyces plicatus*. *The Journal of antibiotics* **1977**, 30, (7), 604-606.
164. Thiericke, R.; Rohr, J., Biological variation of microbial metabolites by precursor-directed biosynthesis. *Nat. Prod. Rep.* **1993**, 10, (3), 265-289.
165. Stegmann, E.; Bischoff, D.; Kittel, C.; Pelzer, S.; Puk, O.; Recktenwald, J.; Weist, S.; Süßmuth, R.; Wohlleben, W. In *Precursor-Directed Biosynthesis for the Generation of Novel Glycopolides*, Biocombinatorial Approaches for Drug Finding, Berlin, Heidelberg, 2005//, 2005; Wohlleben, W.; Spellig, T.; Müller-Tiemann, B., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2005; pp 215-232.
166. Gribble, G. W., Natural organohalogens: a new frontier for medicinal agents? *J. Chem. Educ.* **2004**, 81, (10), 1441.
167. Neumann, C. S.; Fujimori, D. G.; Walsh, C. T., Halogenation Strategies In Natural Product Biosynthesis. *Chem. Biol.* **2008**, 15, (2), 99-109.
168. Fenical, W.; Jensen, P. R., Developing a new resource for drug discovery: marine actinomycete bacteria. *Nat. Chem. Biol.* **2006**, 2, (12), 666-673.
169. The Single Fluorine Substituent. In *Guide to Fluorine NMR for Organic Chemists*, 2016; pp 55-132.
170. Serrano, C. M.; Kanna Reddy, H. R.; Eiler, D.; Koch, M.; Tresco, B. I. C.; Barrows, L. R.; VanderLinden, R. T.; Testa, C. A.; Sebahar, P. R.; Looper, R. E., Unifying the Aminohexopyranose- and Peptidyl-Nucleoside Antibiotics: Implications for Antibiotic Design. *Angew. Chem. Int. Ed. Engl.* **2020**, 59, (28), 11330-11333.
171. Fu, J.; Xu, P.; Yu, B., Total Synthesis of Nucleoside Antibiotics Amicetin, Plicacetin, and Cytosaminomycin A-D. *Chinese Journal of Chemistry* n/a, (n/a).
172. Lichtenthaler, F. W.; Černá, J.; Rychlik, I., The effect of oxamicetin and some amicetin analogs on ribosomal peptidyl transferase. *FEBS Lett.* **1975**, 53, (2), 184-187.

References

173. Kusche, B. R.; Phillips, J. B.; Priestley, N. D., Nonactin biosynthesis: Setting limits on what can be achieved with precursor-directed biosynthesis. *Bioorg. Med. Chem. Lett.* **2009**, 19, (4), 1233-1235.
174. Sester, A.; Stier-Patowsky, K.; Hiller, W.; Kloss, F.; Lütz, S.; Nett, M., Biosynthetic Plasticity Enables Production of Fluorinated Aurachins. *Chembiochem : a European journal of chemical biology* **2020**, 21, (16), 2268-2273.

Supplementary Information





Map of District	Collection Sites	Colony codes	Genus
	Mud Dauber Nest “Kumal koti”	COL 1-COL 4	<i>Pseudonocordia</i> <i>Nonomuraea</i> <i>Streptomyces</i>
	Narayani River	Narayani River 1	<i>Streptomyces</i>
	Nagar Baan	NB001-NB004	
	Bis Hazari Lake	WL001-WL067	
	Batuli Lake	BAT001-BAT007	
	Termite Nest 'Dhamiro'	BK001-BK003	
	Swargadwari area	PYU001-PYU004	Not Assigned
	Sorakhutte	RP001-RP029	<i>Streptomyces</i>
	Liwang		
	Eriwang		
	Thawang		
	Namche Bazaar	NP001-NP006	<i>Streptomyces</i>
	Tengboche		
	Panboche		
	Dingboche		
	Lobuche		
	Gorakshep		

Table S1. Collection habitats and colonies designation.

Supplementary

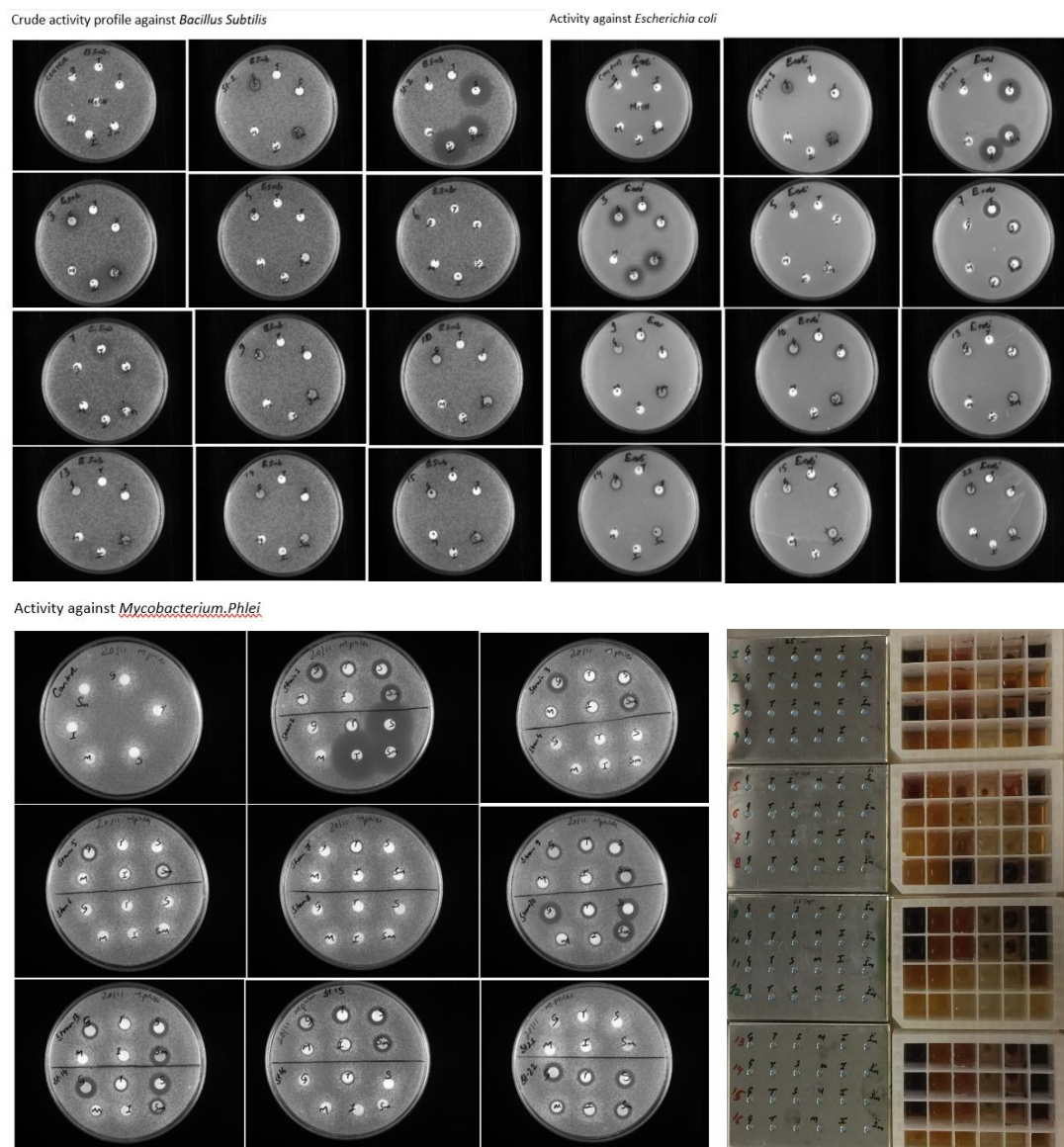


Figure S1 Antimicrobial activity accessed for crude from various media using disk diffusion assay and OSMAC approach in a mini bioreactor.

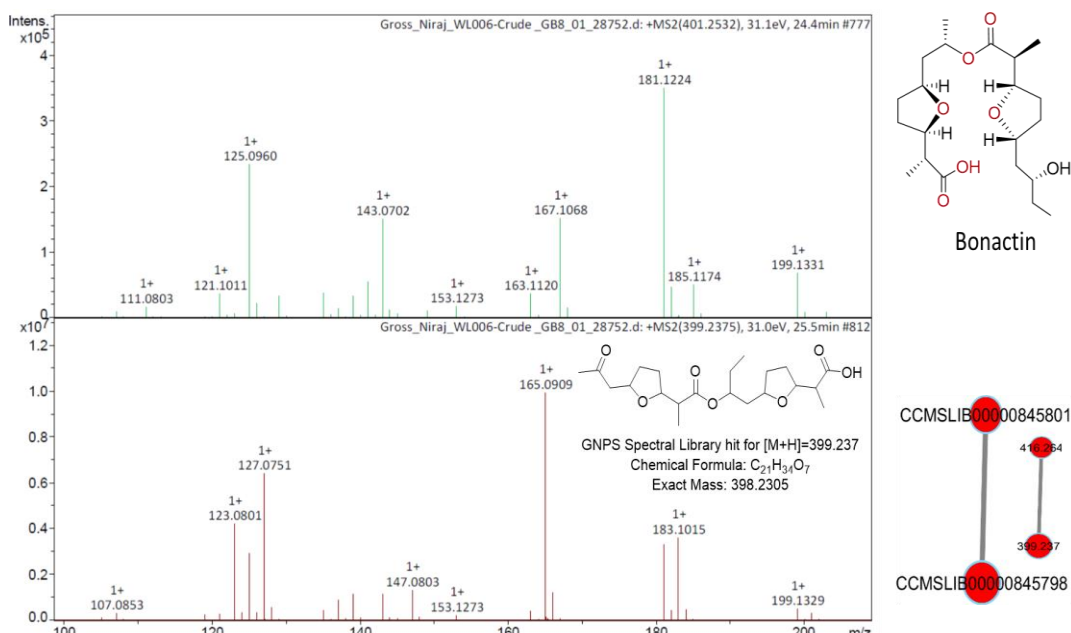


Figure S2. Lower MS/MS fragments comparison for bonactin (green) and its unsaturated congener (red) Spectrum ID from GNPS spectral library hit on MN (CCMSLIB00000845798=[M+H], CCMSLIB00000845801 [M+NH₄])

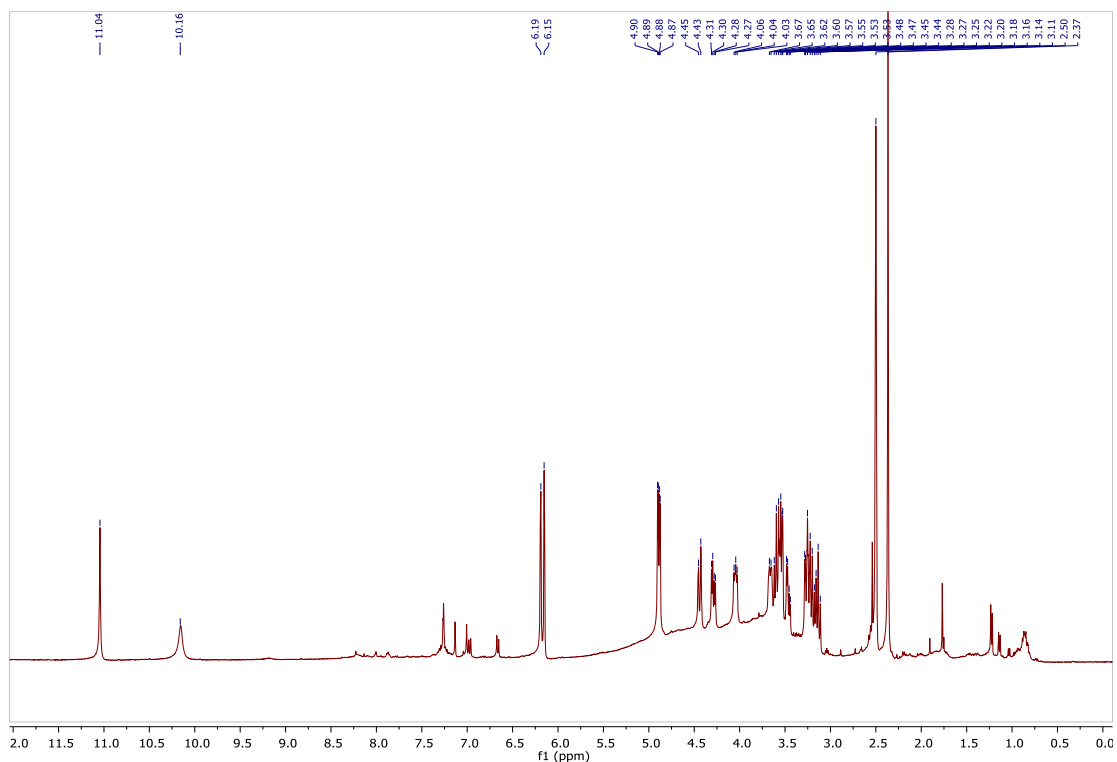


Figure S3. ¹H NMR spectrum of **1** (400 MHz, d₆-DMSO)

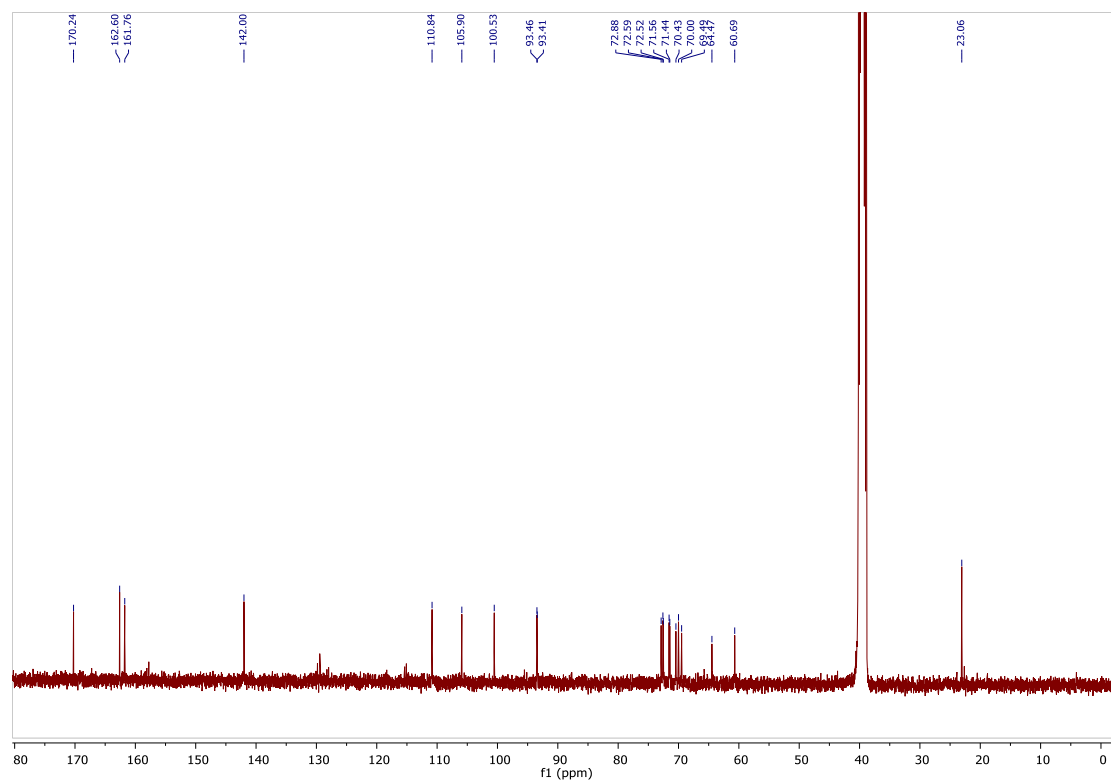


Figure S4. ^{13}C NMR spectrum of **1** (100 MHz, d_6 -DMSO)

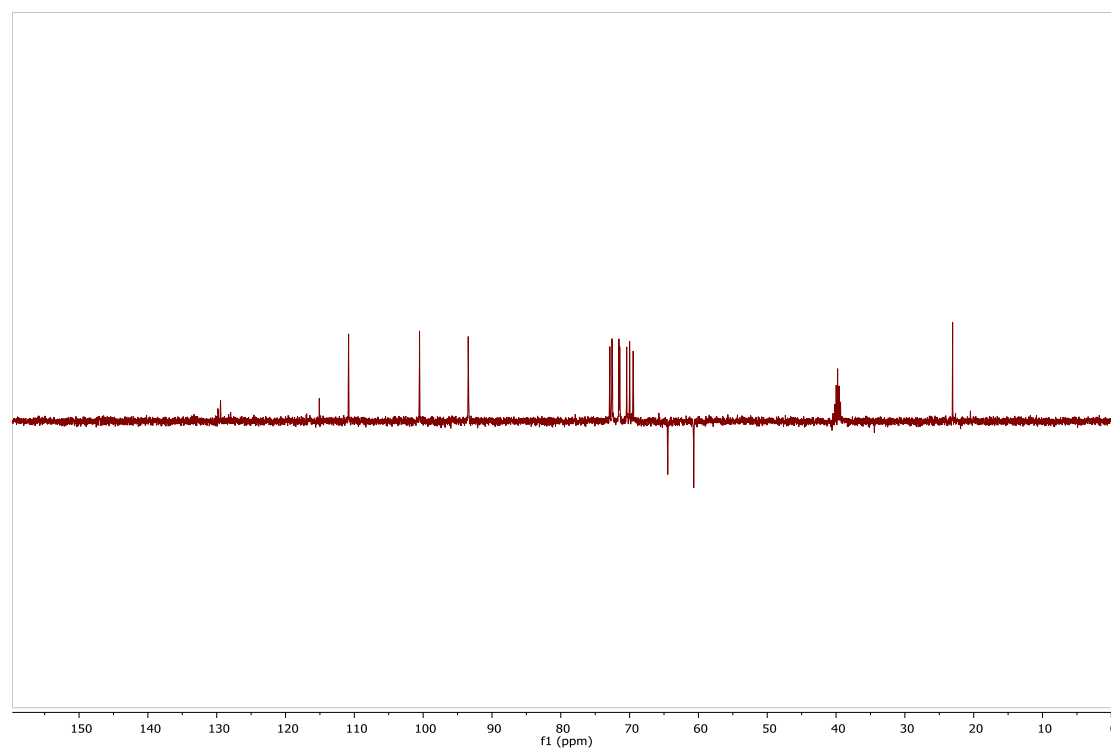


Figure S5. DEPT135 NMR spectrum of **1** (400 MHz, d_6 -DMSO)

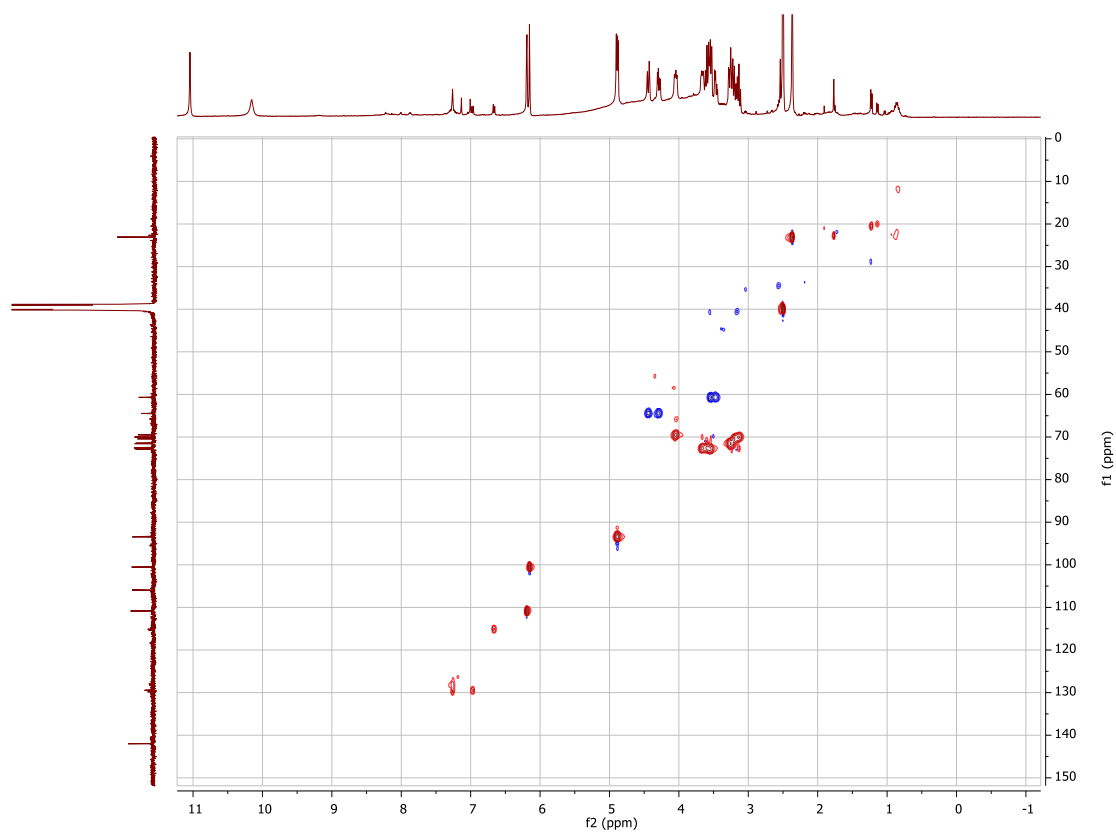


Figure S6. ^1H - ^{13}C multiplicity edited HSQC NMR spectrum of **1**

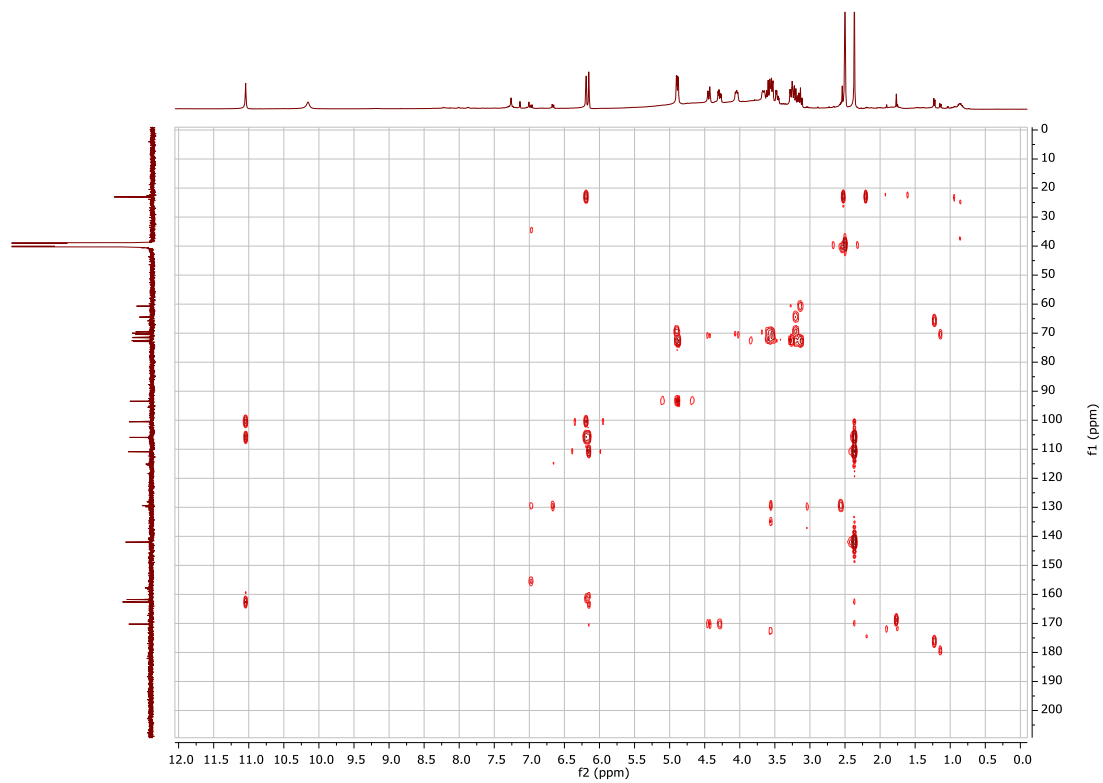


Figure S7. ^1H - ^{13}C HMBC NMR spectrum of **1** (400 MHz, d_6 -DMSO)

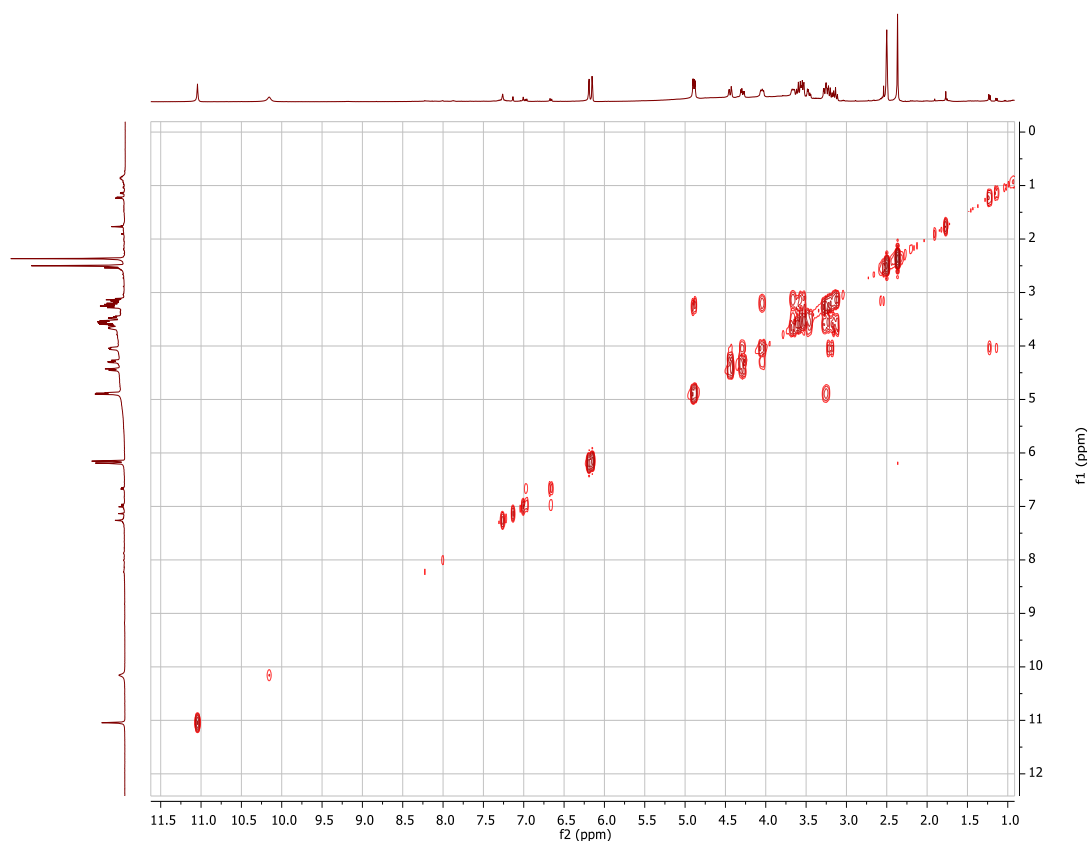


Figure S8. $^1\text{H} - ^1\text{H}$ COSY NMR spectrum of **1** (400 MHz, d_6 -DMSO)

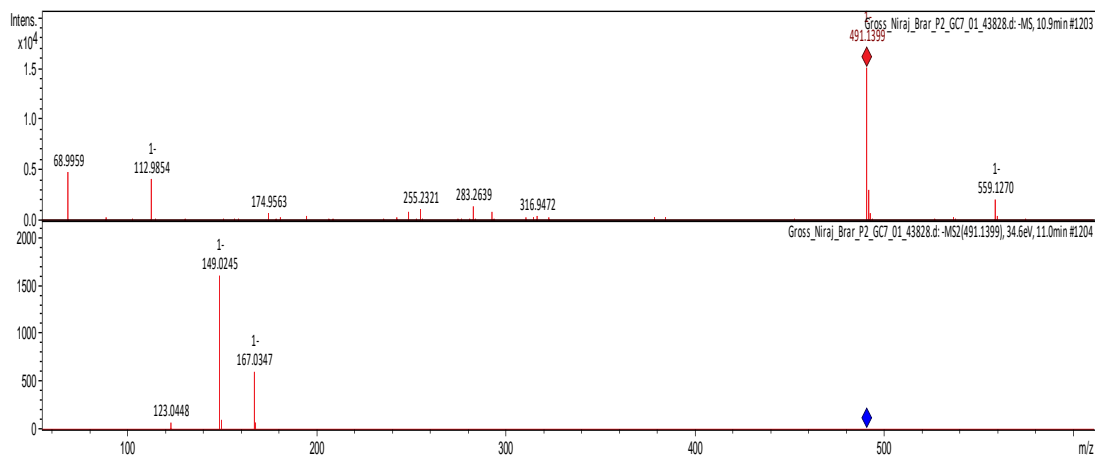


Figure S9. MS1 and MS2 spectra for m/z 492 or compound **1** in negative mode

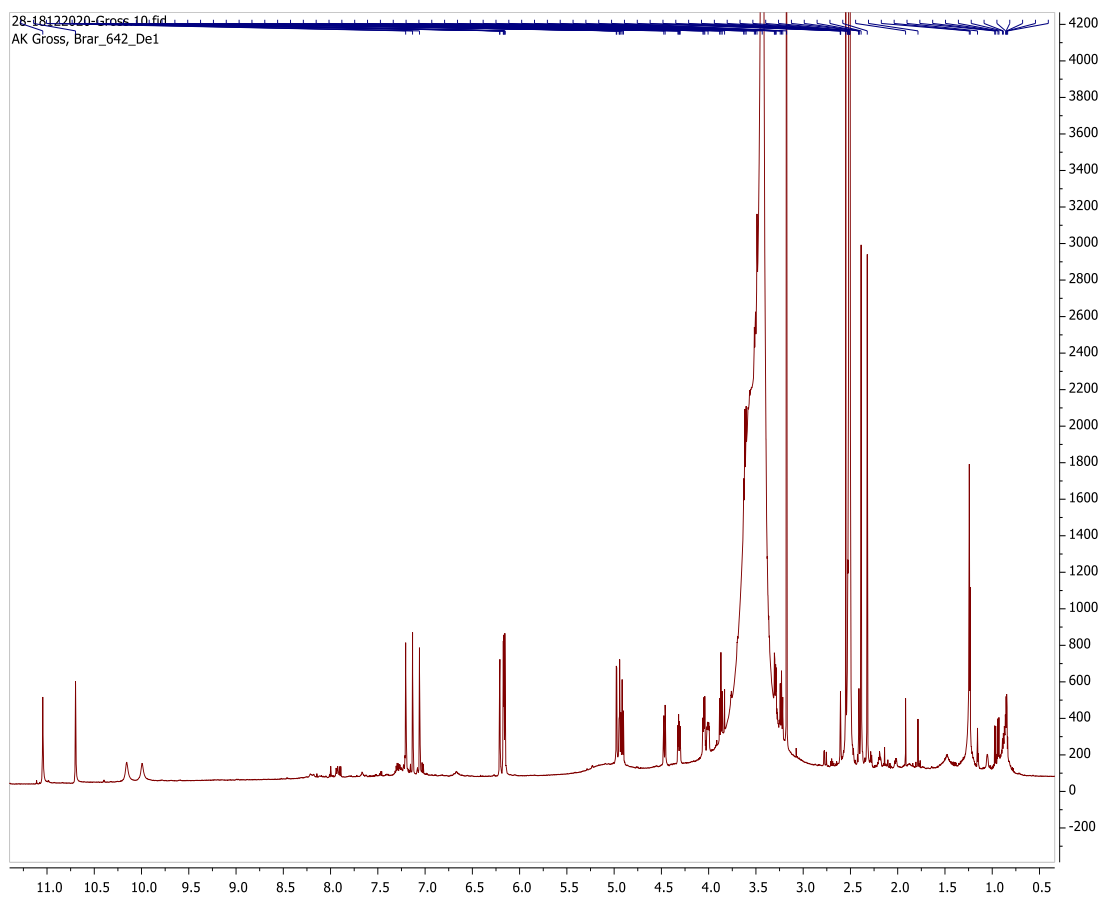


Figure S10. ¹H NMR spectrum of **compound 2.1** (700 MHz, *d*₆-DMSO)

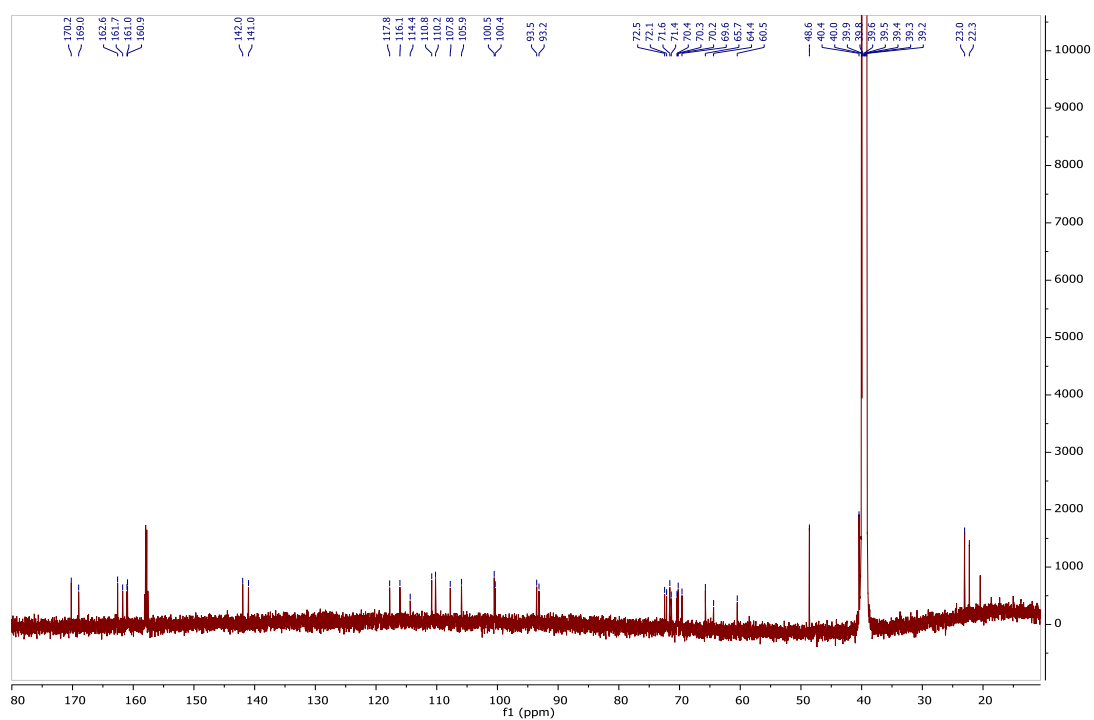


Figure S11. ¹³C NMR spectrum of **compound 2.1** (175 MHz, *d*₆-DMSO)

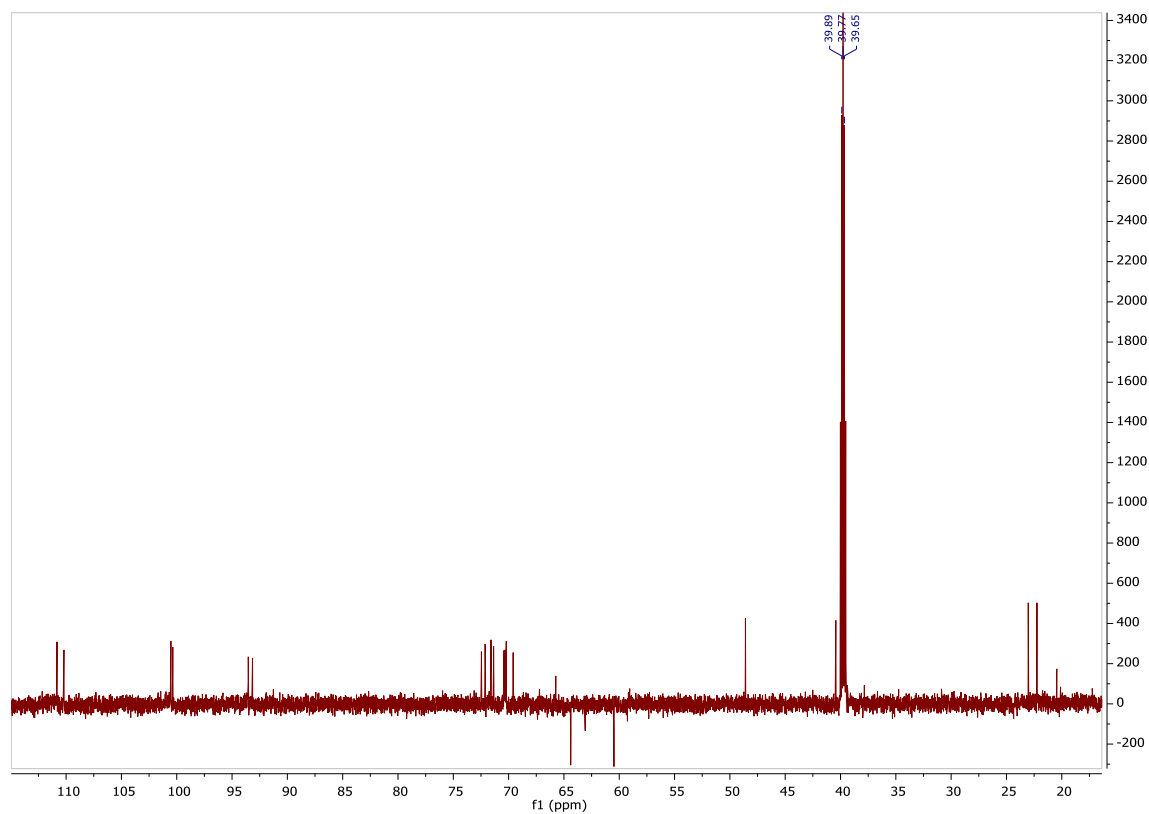


Figure S12. DEPT135 NMR spectrum of **compound 2.1**

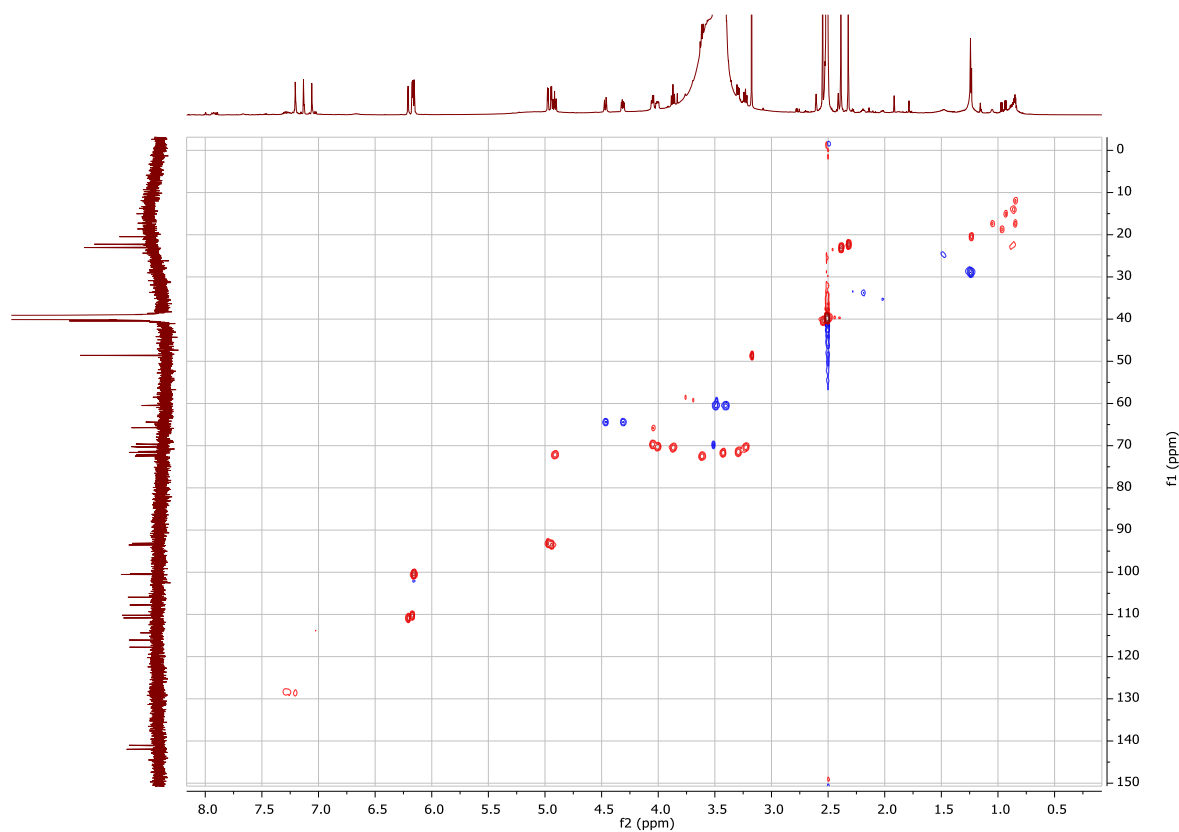


Figure S13. ^1H - ^{13}C multiplicity edited HSQC NMR spectrum of **compound 2.1**

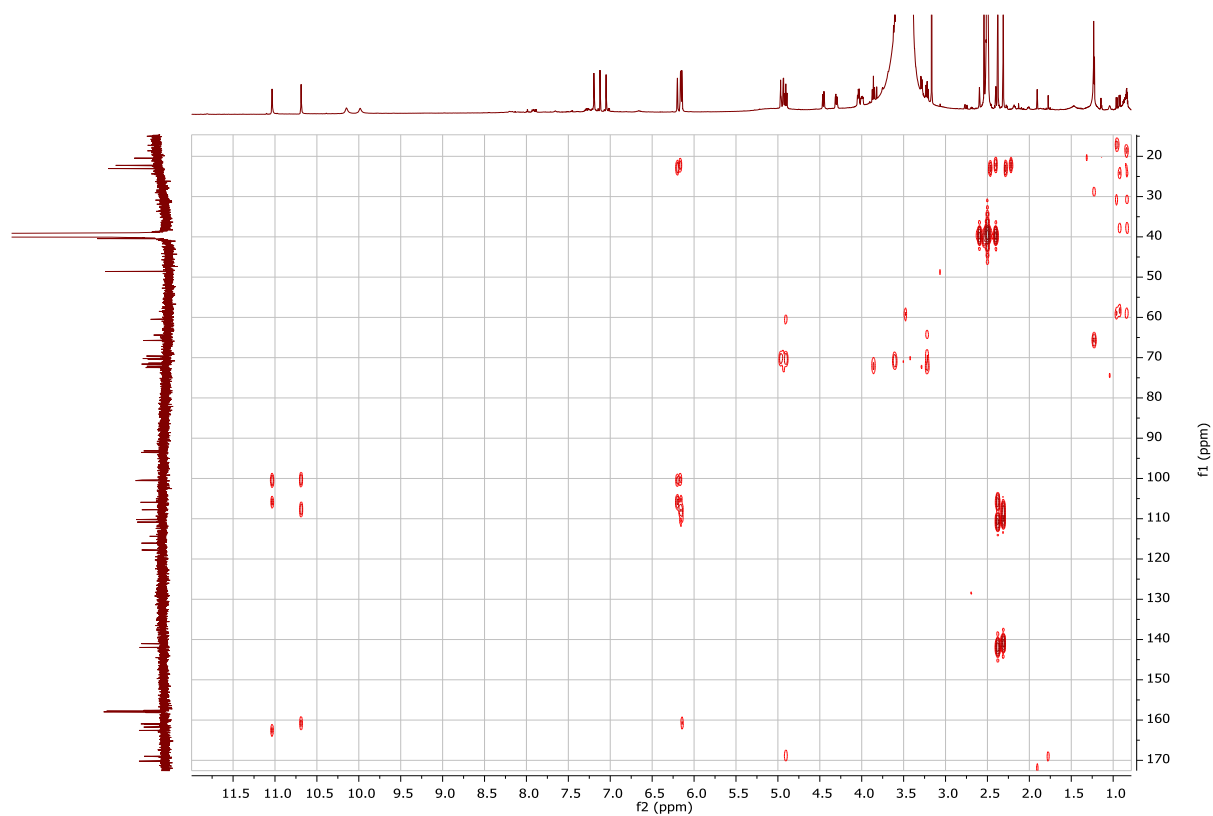


Figure S14. ^1H - ^{13}C HMBC NMR spectrum of **compound 2.1**

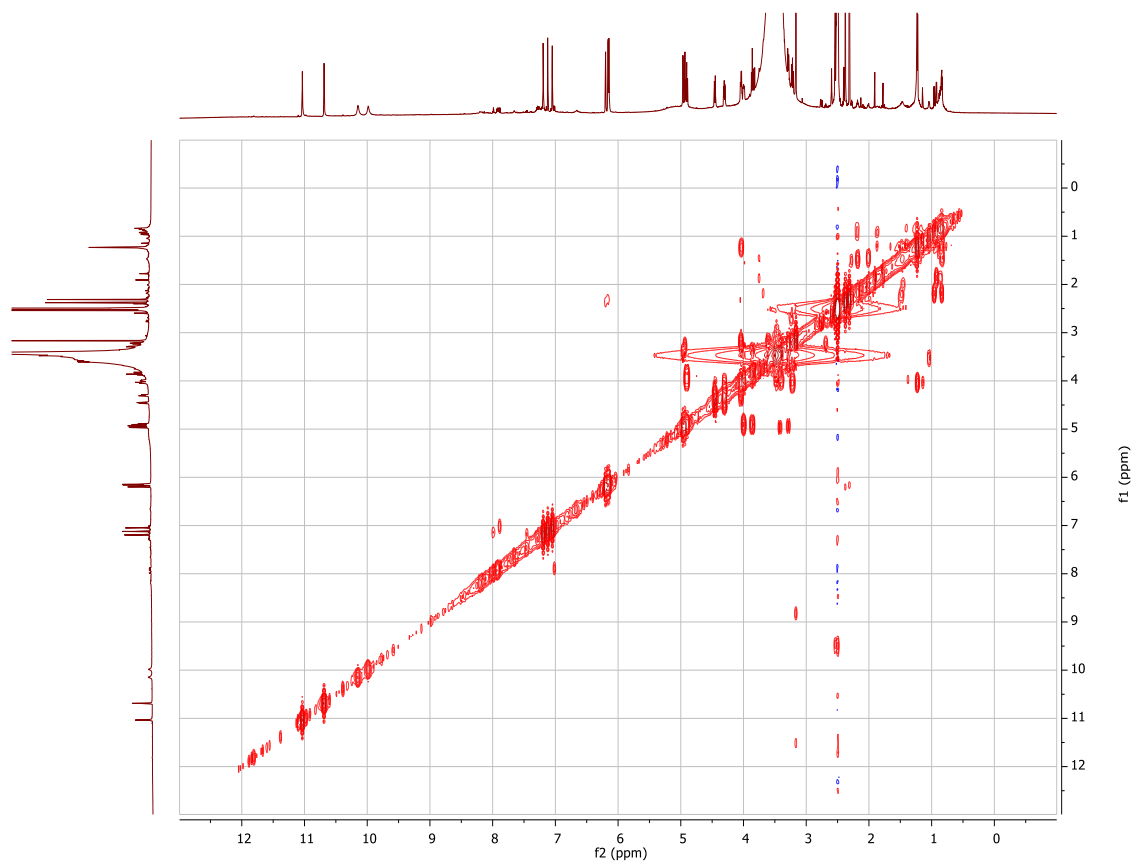


Figure S15. ^1H - ^1H COSY NMR spectrum of **compound 2.1**

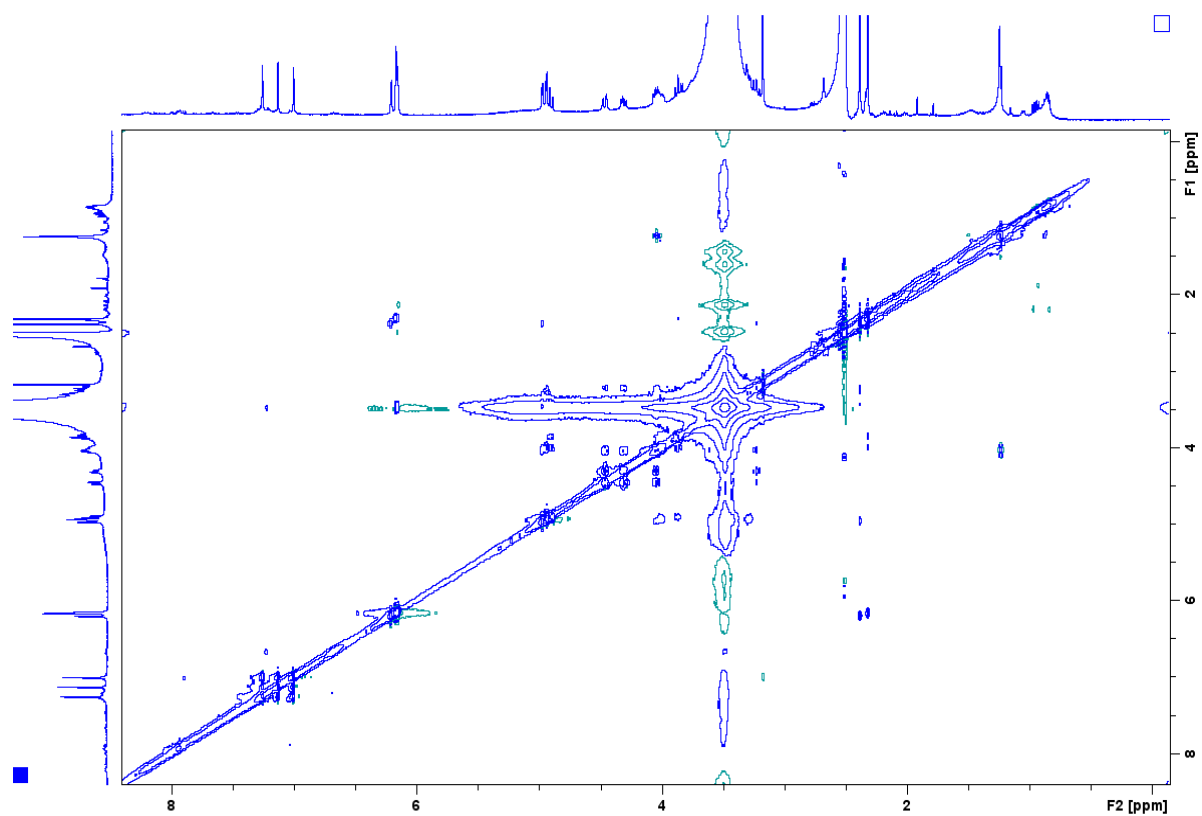


Figure S16. NOESY spectra of **compound 2.1** (400 MHz, d_6 DMSO)

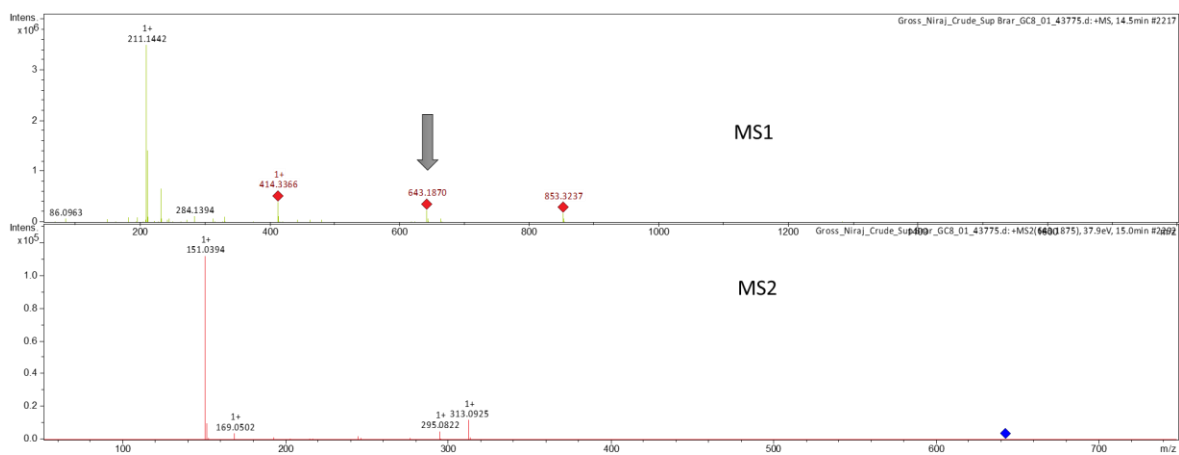


Figure S17. MS1 and MS2 spectra for **compound 2.1**

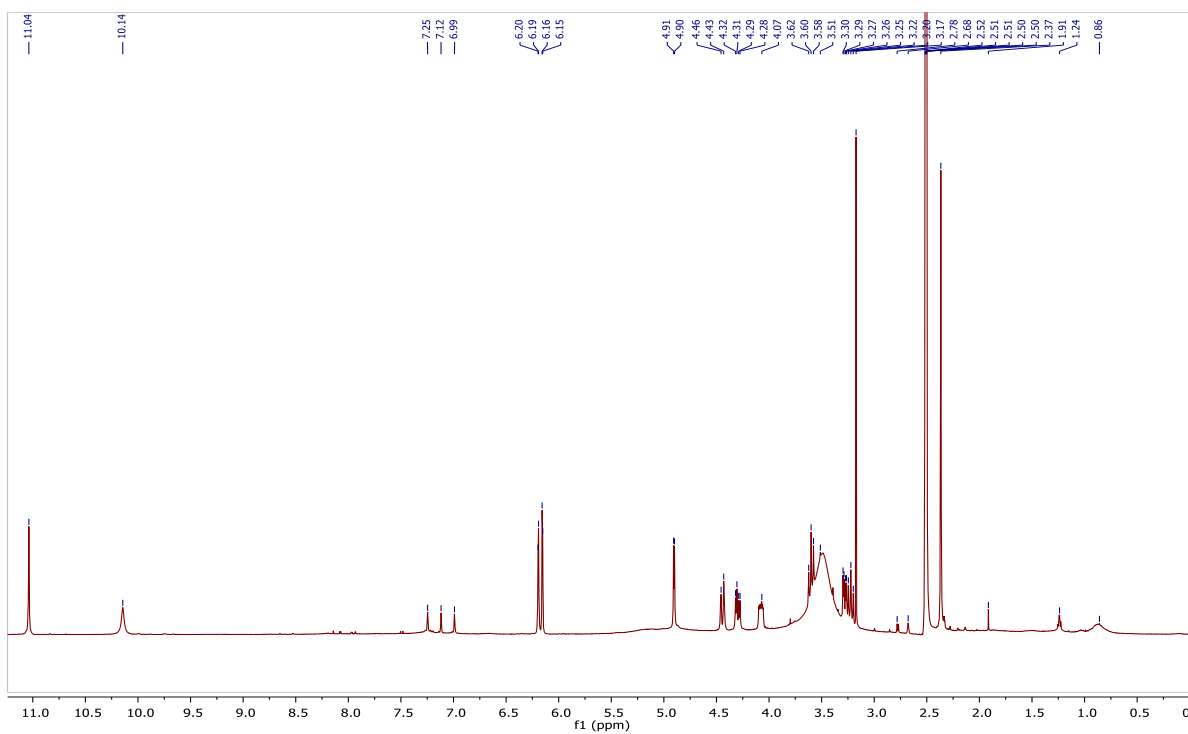


Figure S18. ^1H NMR spectrum of **Brartemicin** (400 MHz, d_6 -DMSO)

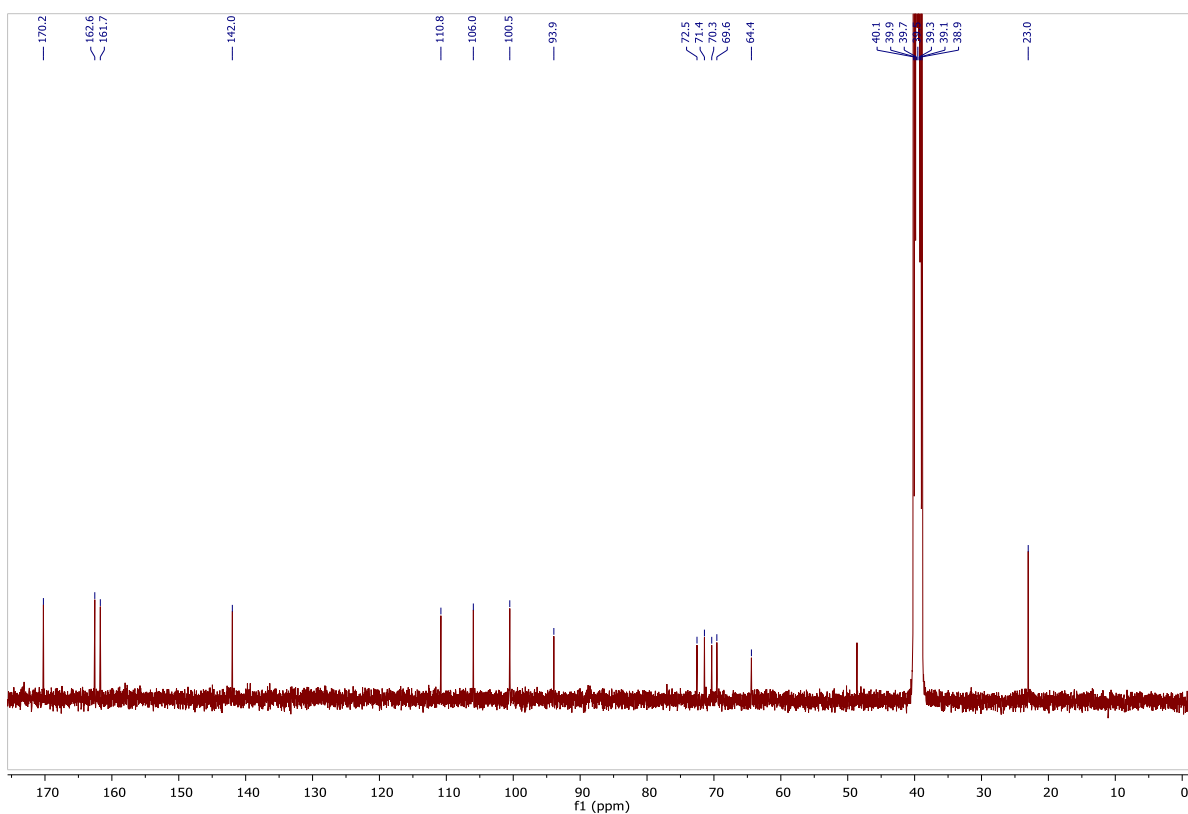


Figure S19. ^{13}C NMR spectrum of **Brartemicin** (100 MHz, d_6 -DMSO)

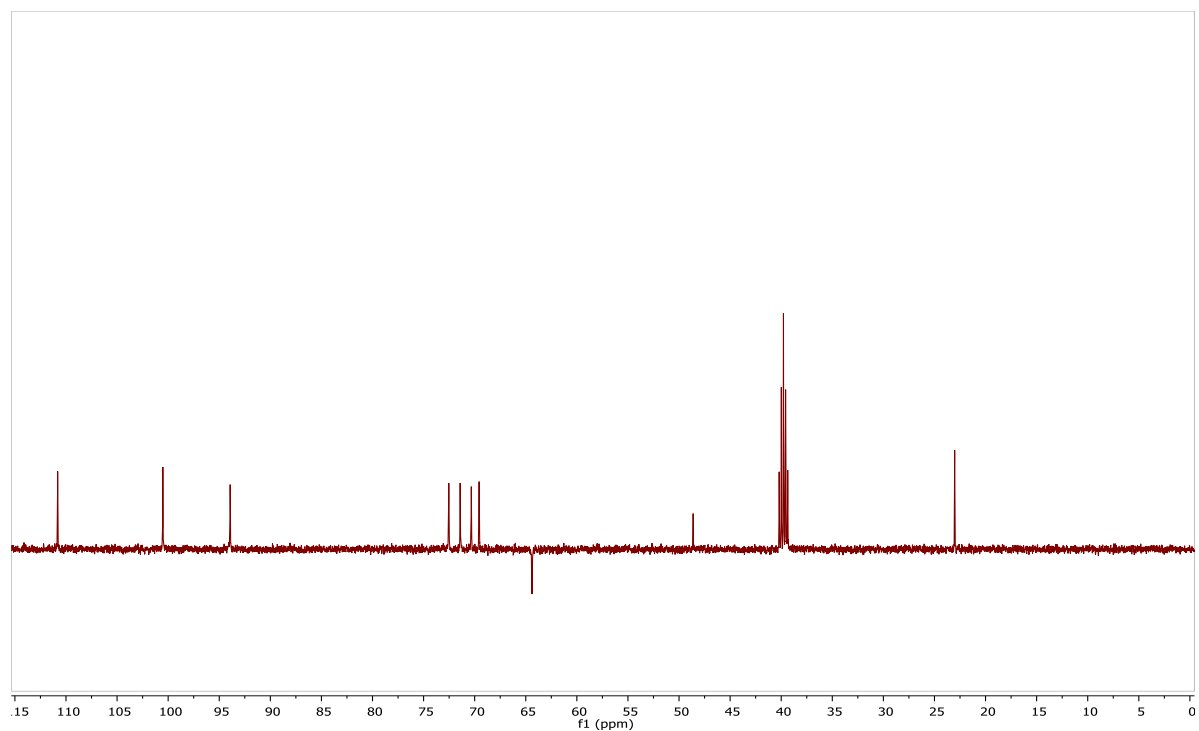


Figure S20. DEPT135 NMR spectrum of **Brartemicin** (400 MHz, d_6 -DMSO)

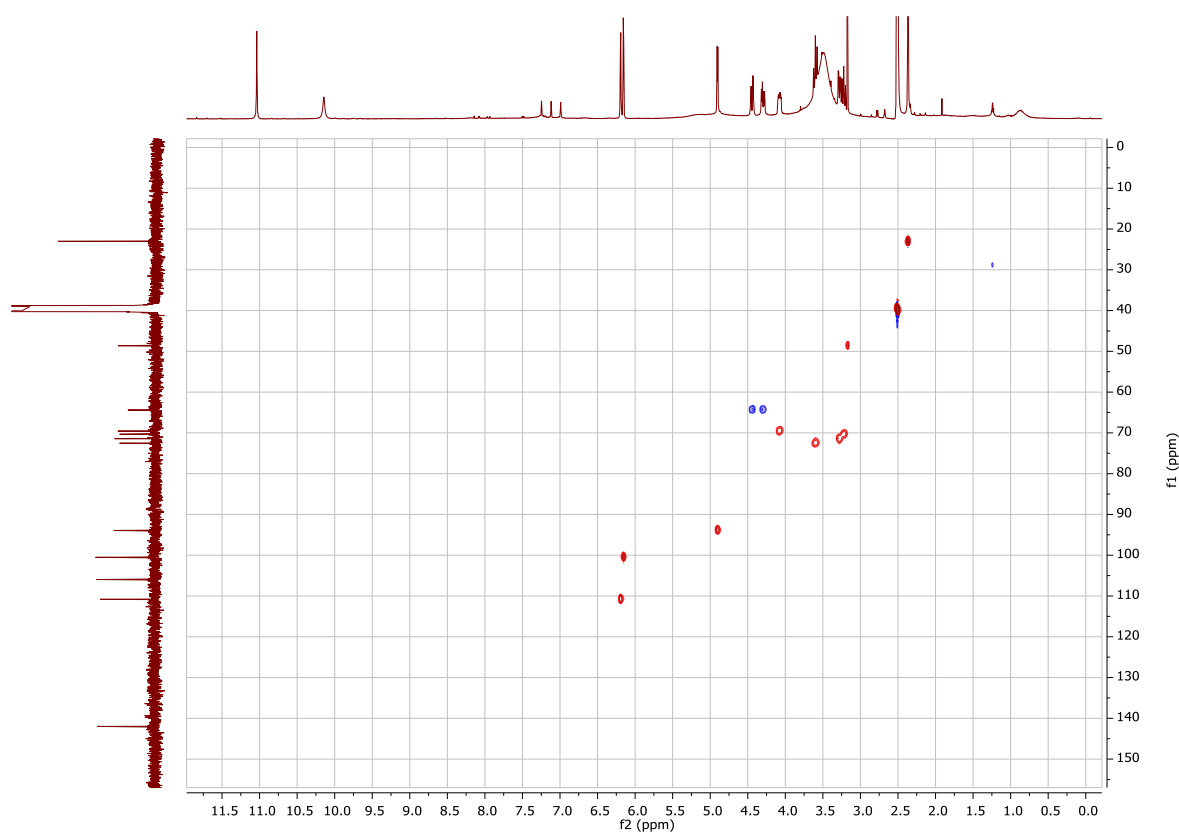


Figure S21. ^1H - ^{13}C multiplicity edited HSQC NMR spectrum of **Brartemicin** (400 MHz, d_6 -DMSO)

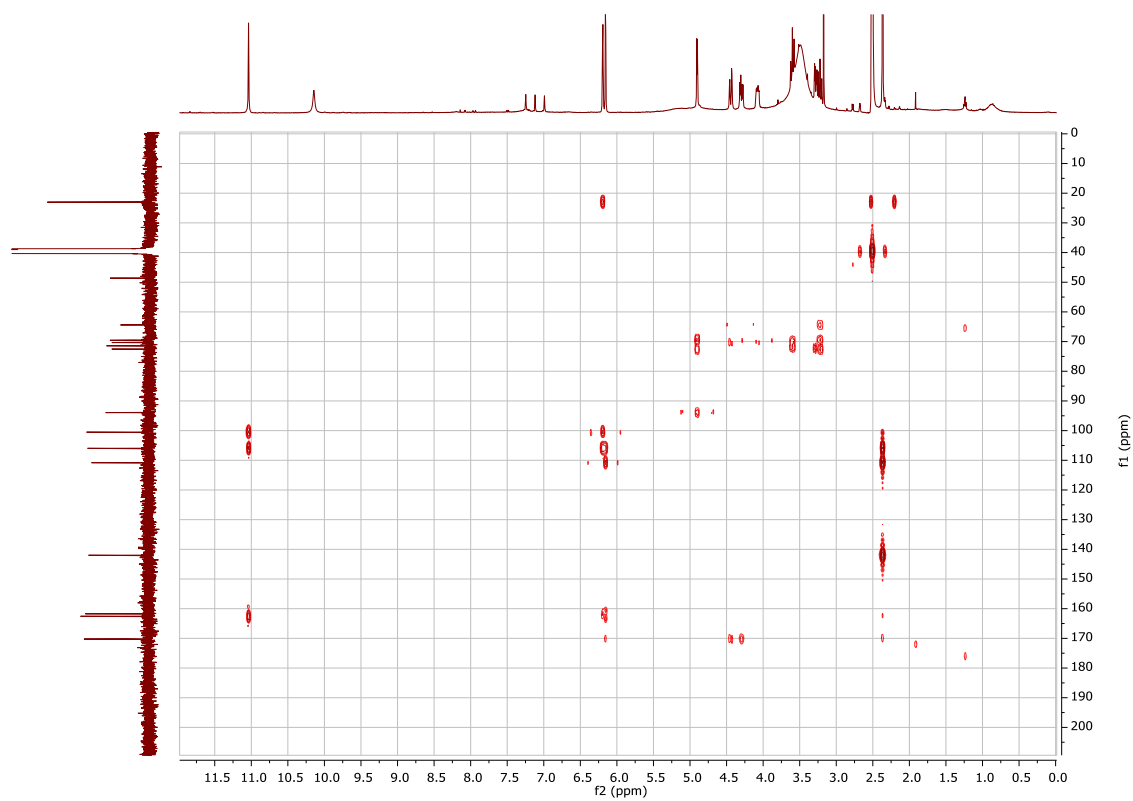


Figure S22. ^1H - ^{13}C HMBC NMR spectrum of **Brartemicin** (400 MHz, d_6 -DMSO)

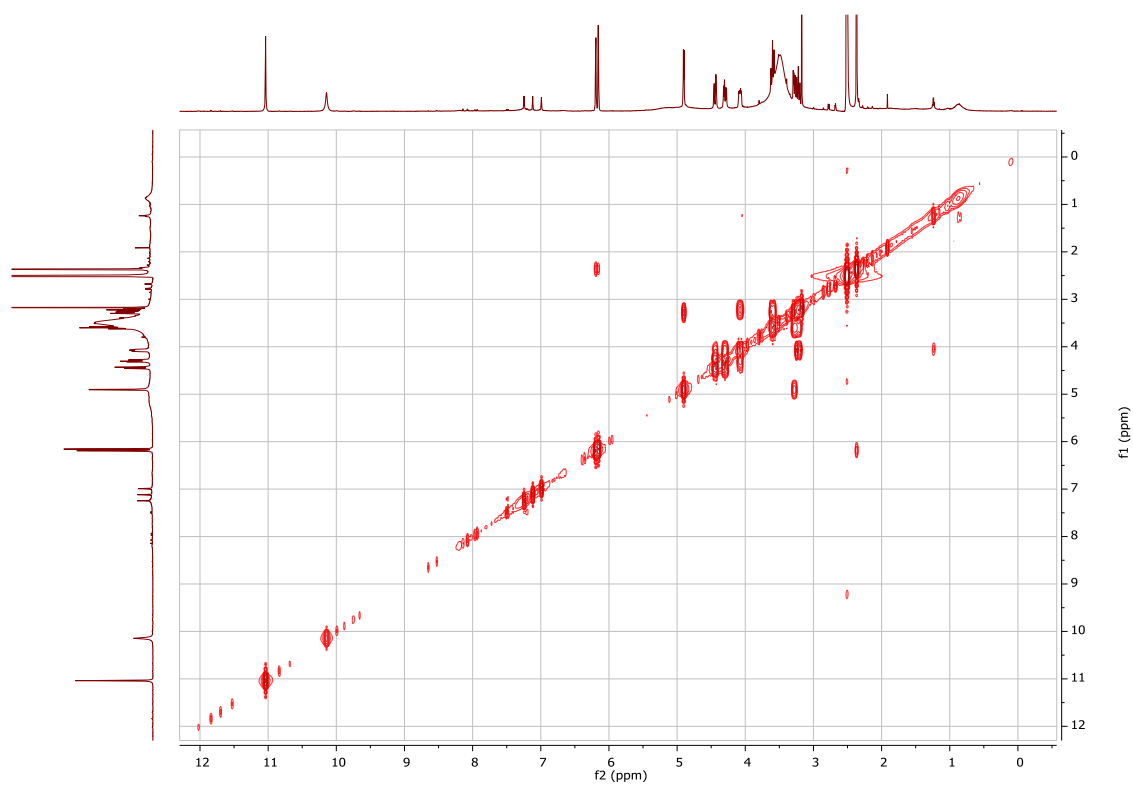


Figure S23. ^1H - ^1H COSY NMR spectrum of **Brartemicin** (400 MHz, d_6 -DMSO)

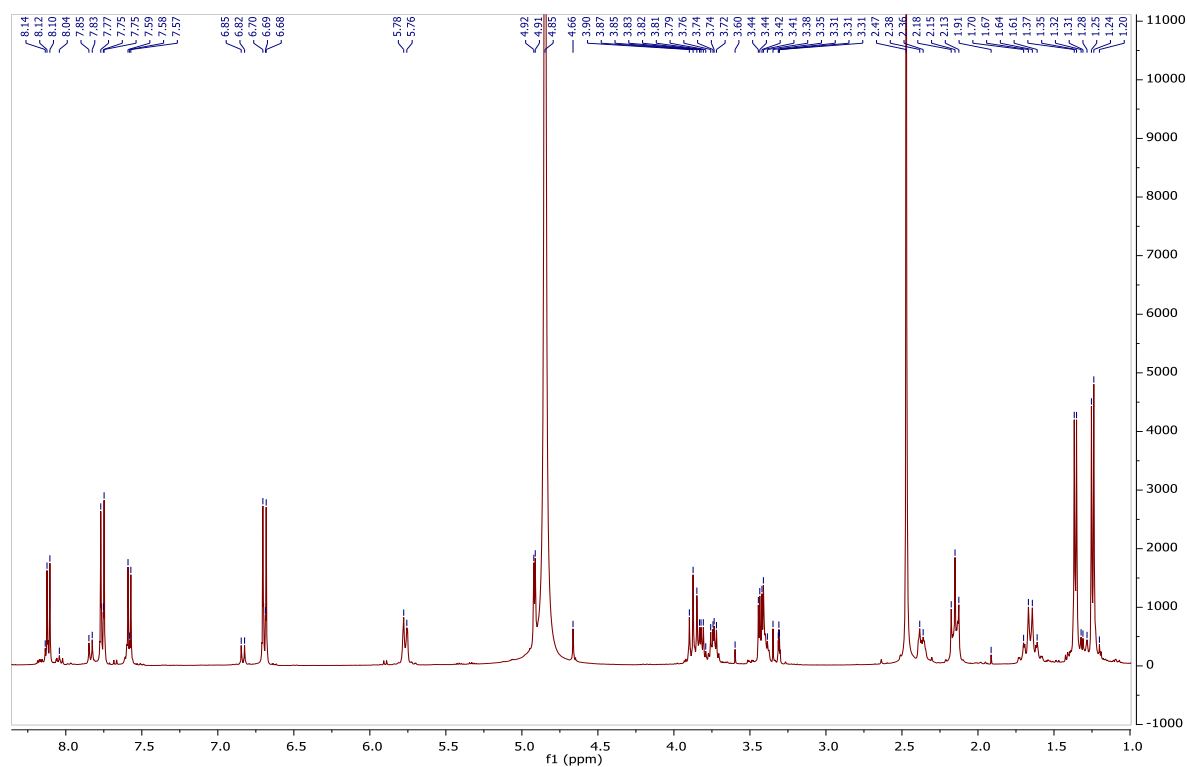


Figure S24. ^1H NMR spectrum of **Plicacetin** (400 MHz, d_4 -MeOD)

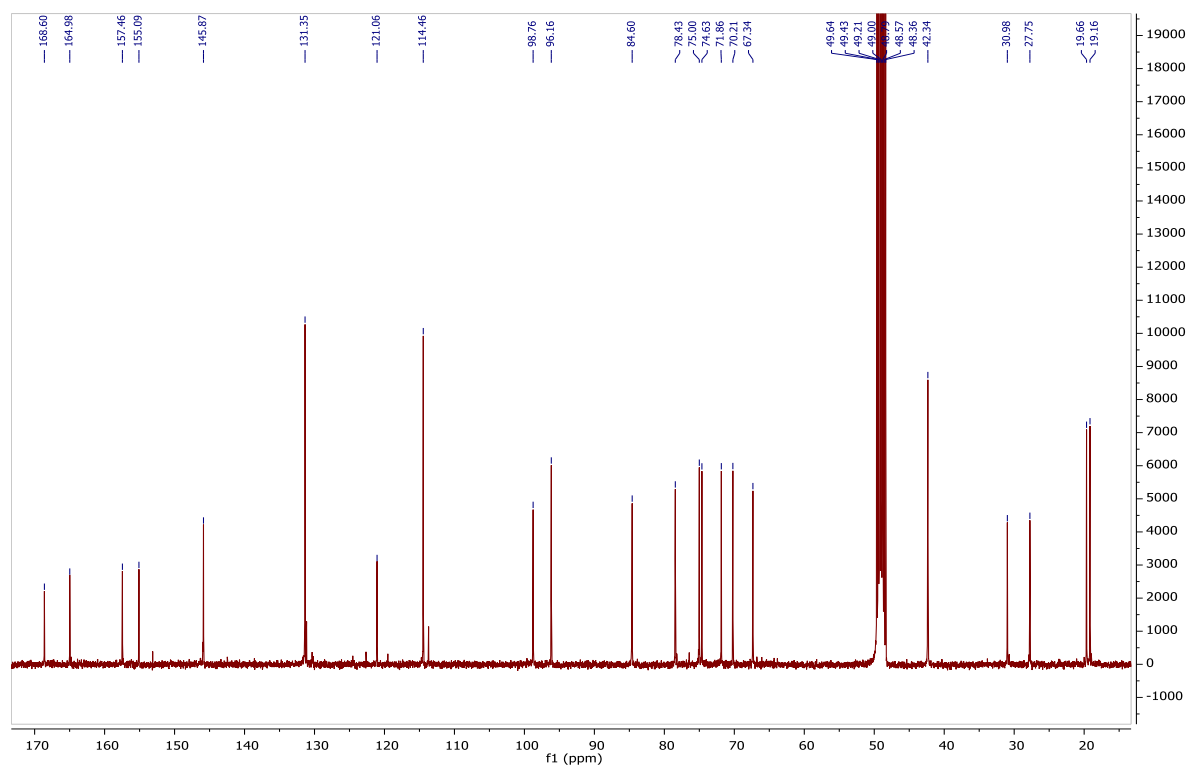


Figure S25. ^{13}C NMR spectrum of **Plicacetin** (100 MHz, d_4 -MeOD)

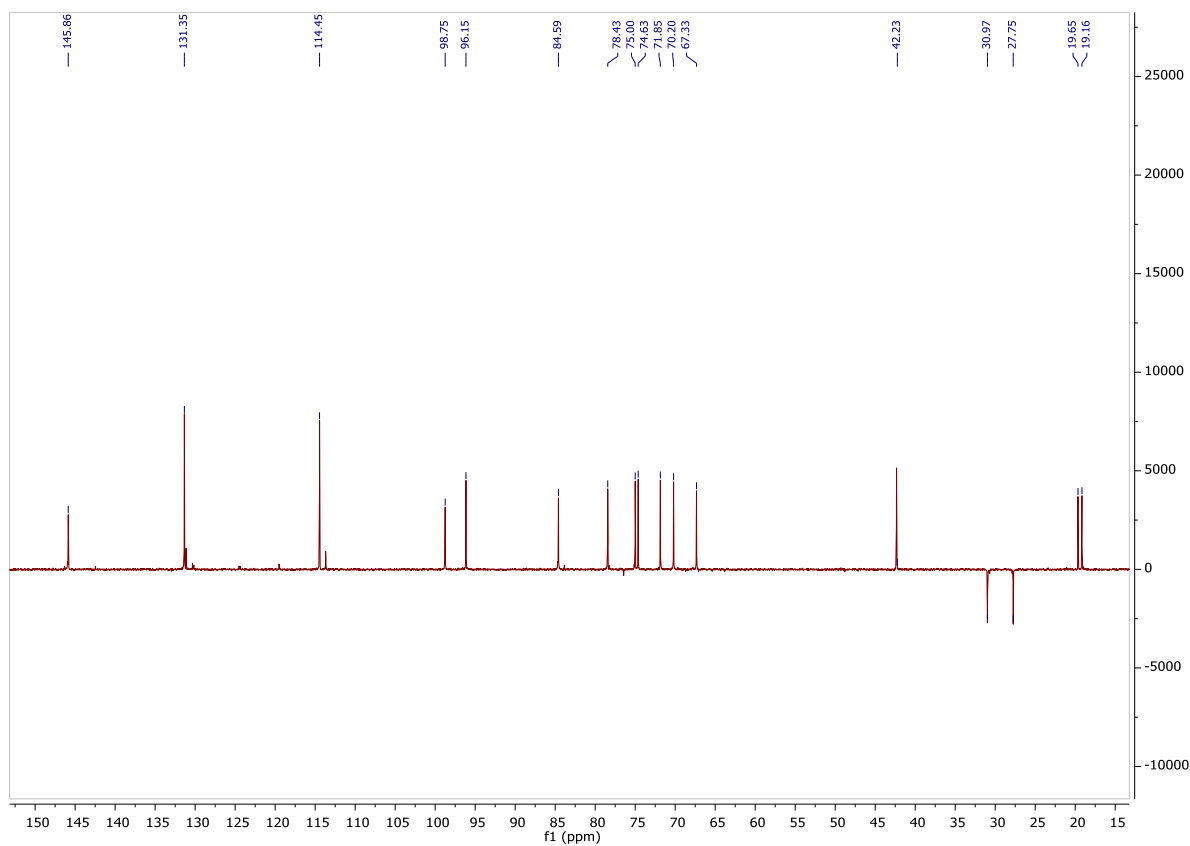


Figure S26. ^{13}C DEPT spectrum of Plicacetin

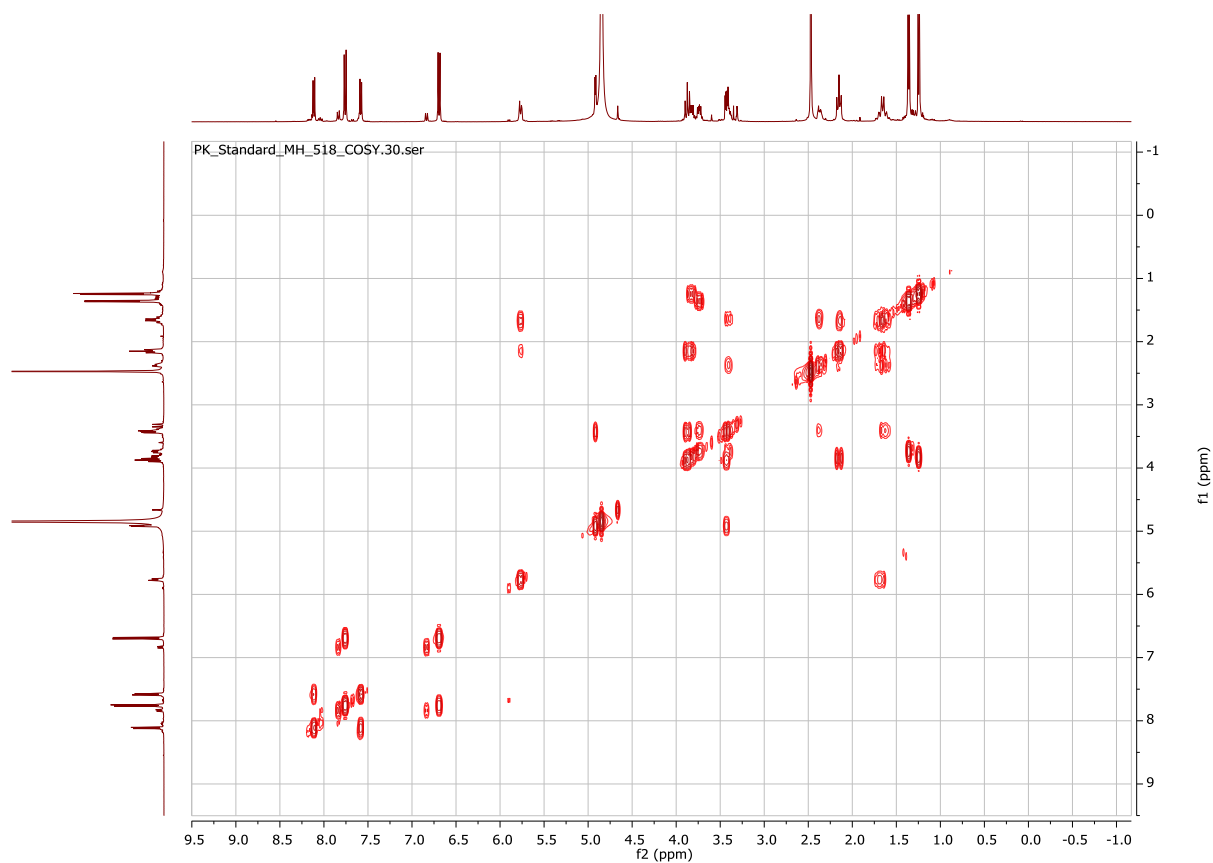


Figure S27. ^1H - ^1H COSY spectrum of Plicacetin

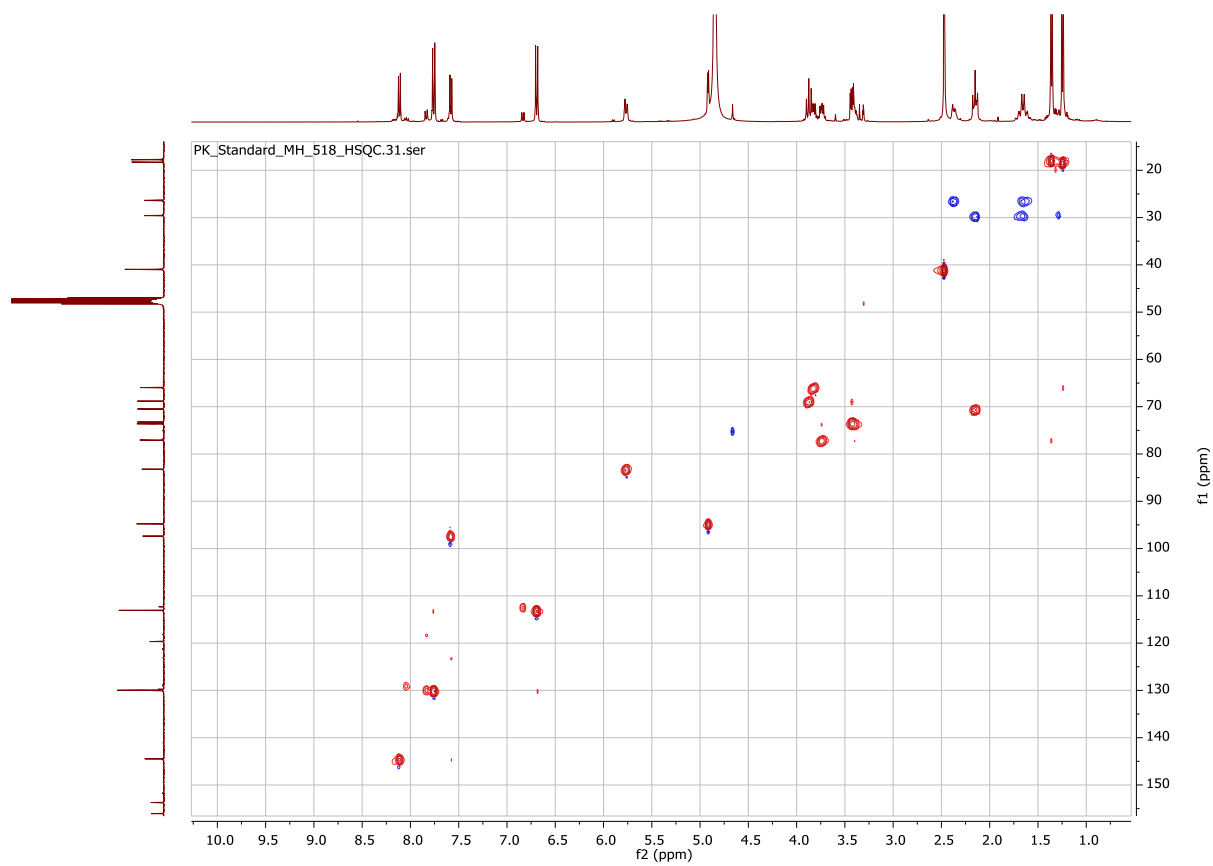


Figure S28. HSQC spectrum of Plicacetin

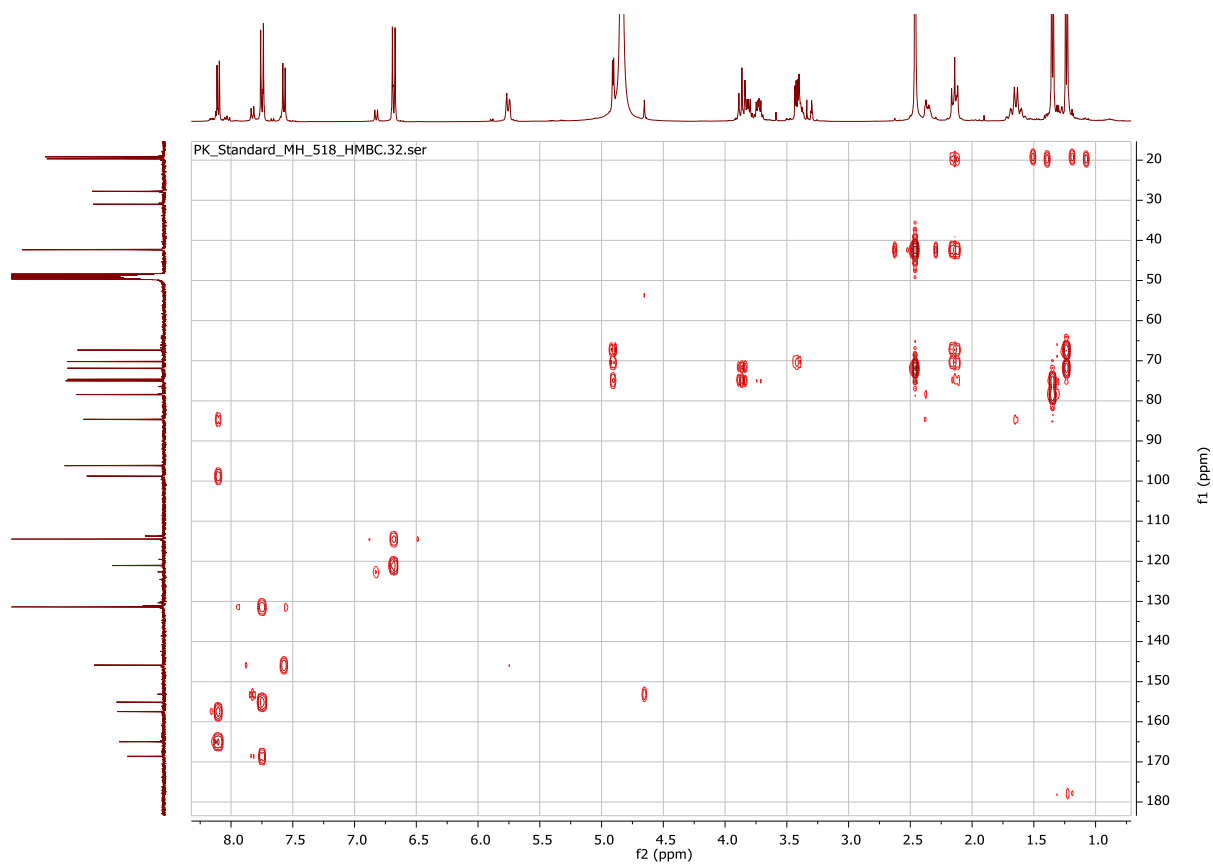


Figure S29. HMBC spectrum of Plicacetin

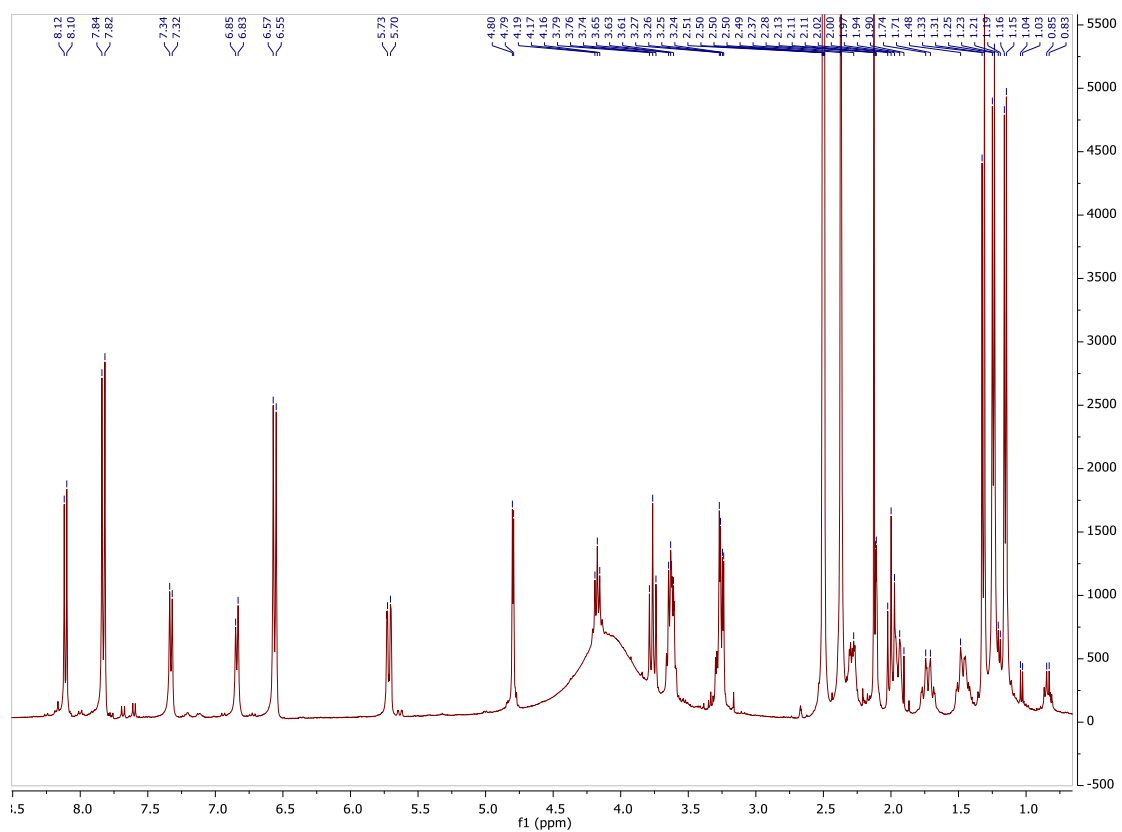


Figure S30. ^1H NMR spectrum of **Compound XII** (400 MHz, d_6 -DMSO)

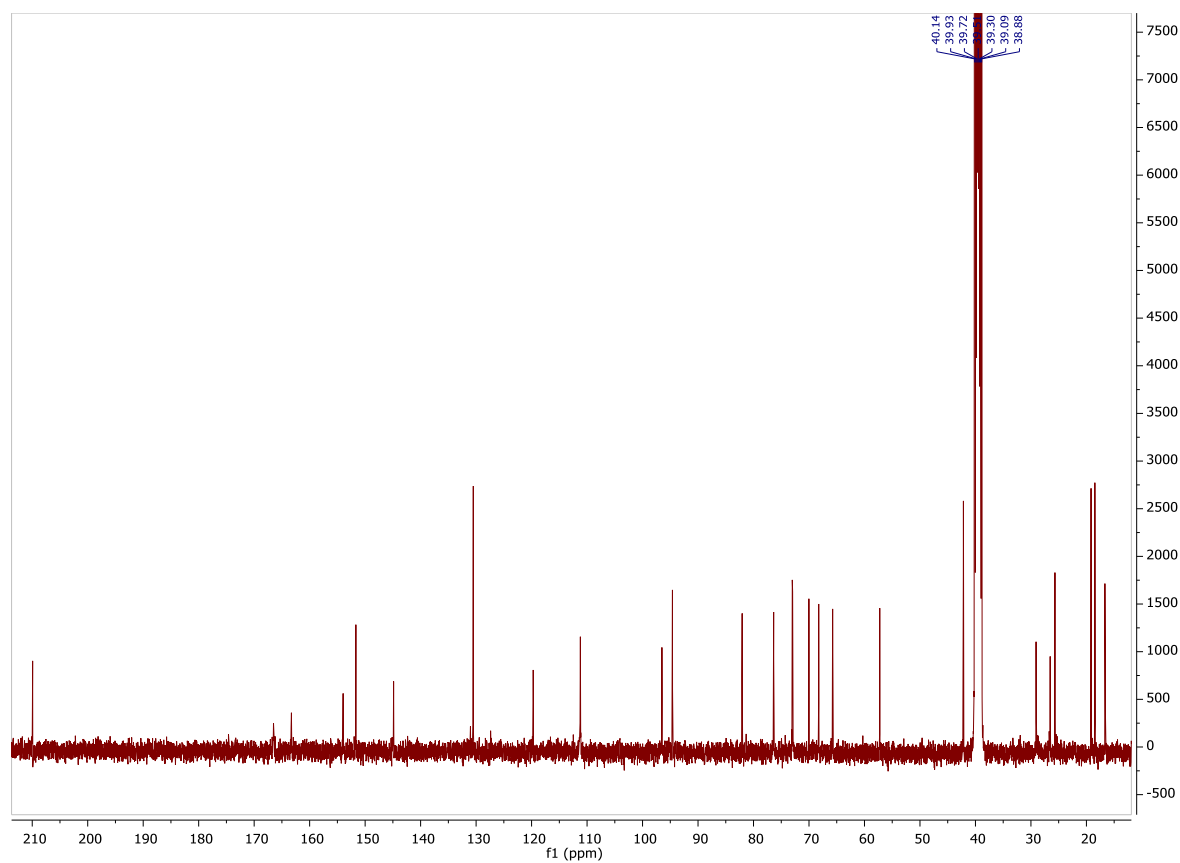


Figure S31. ^{13}C NMR spectrum of **Compound XII** (100 MHz, d_6 -DMSO)

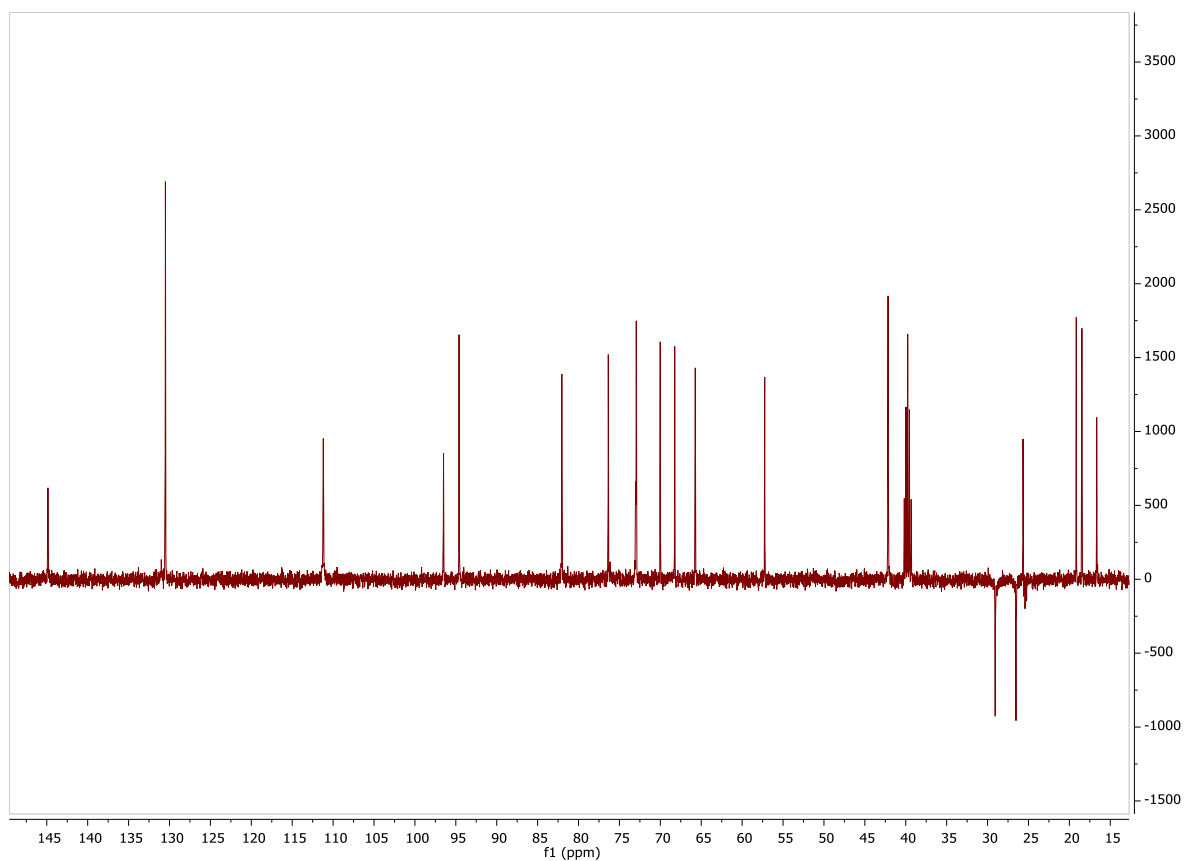


Figure S32. 135DEPT spectrum of Compound XII

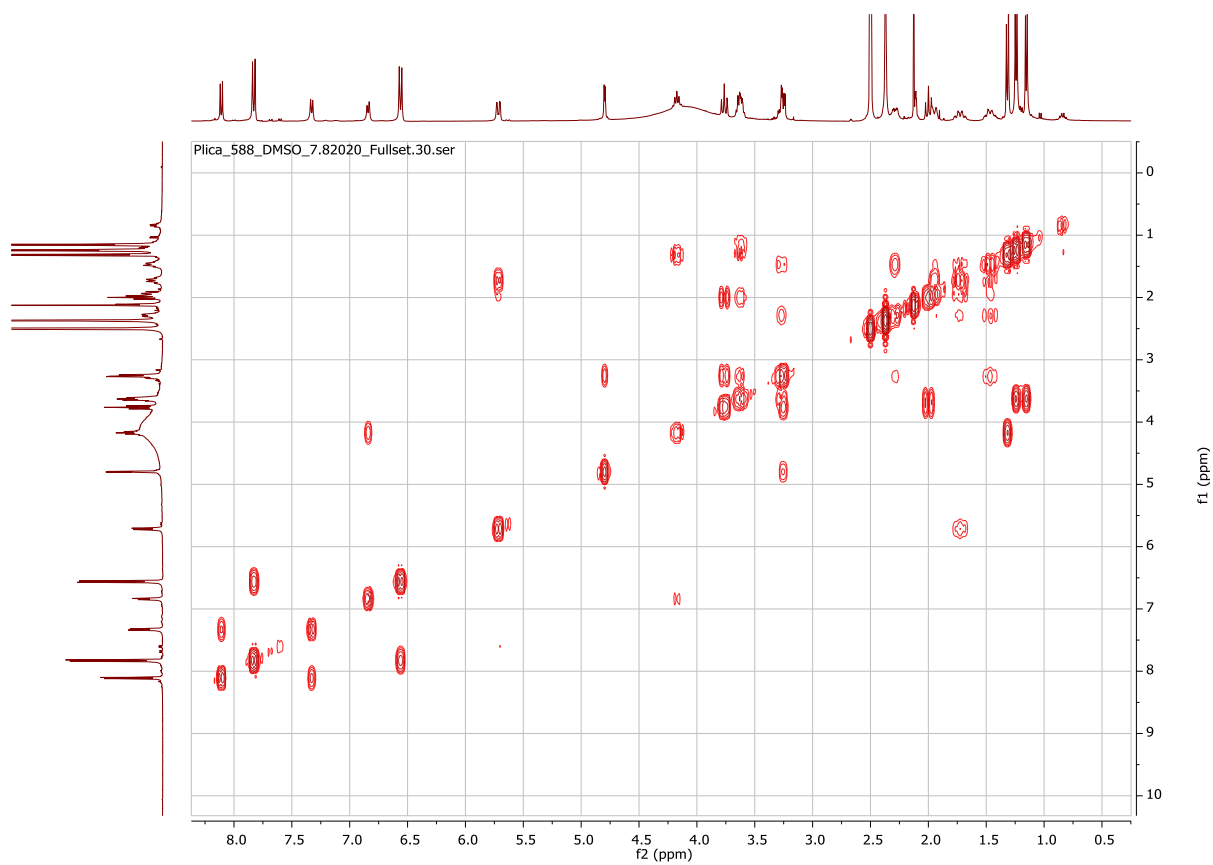


Figure S33. 1H-1H COSY spectrum of Compound XII

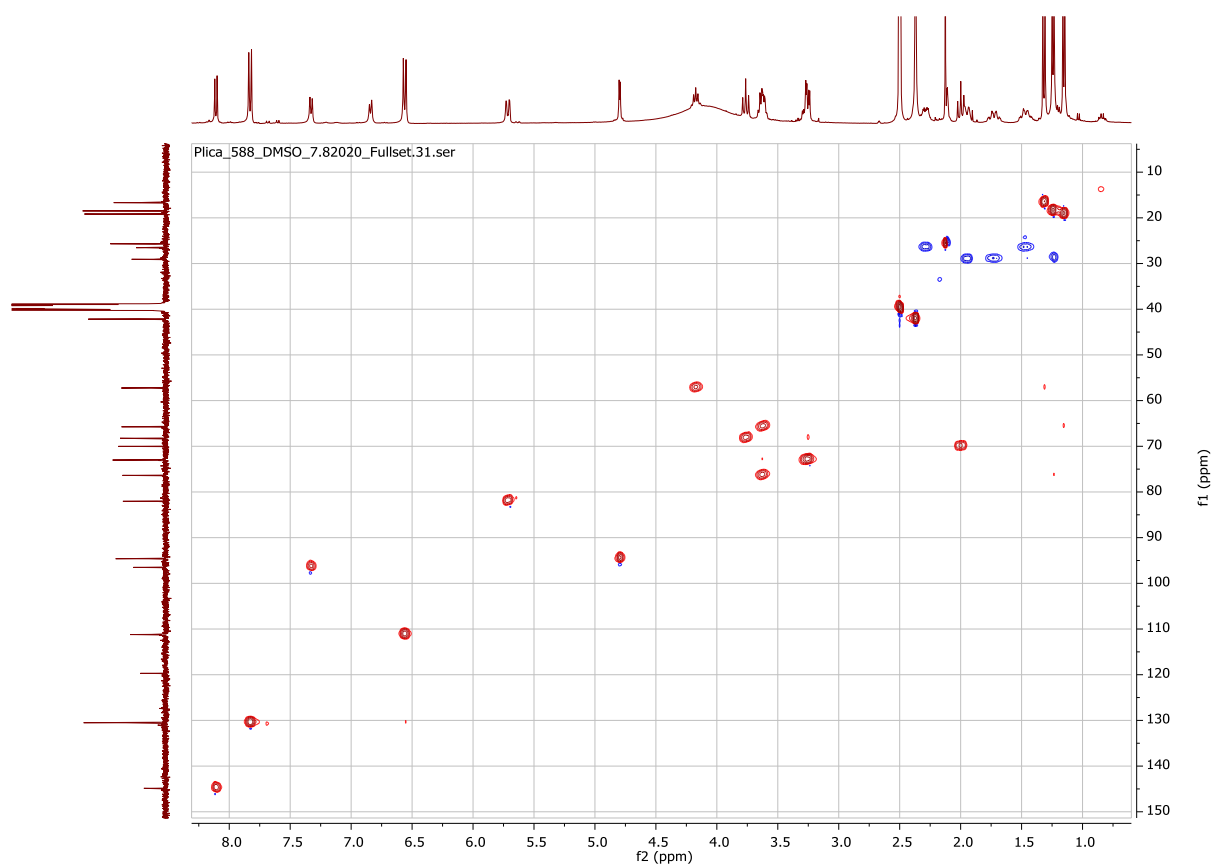


Figure S34. HSQC spectrum of Compound XII

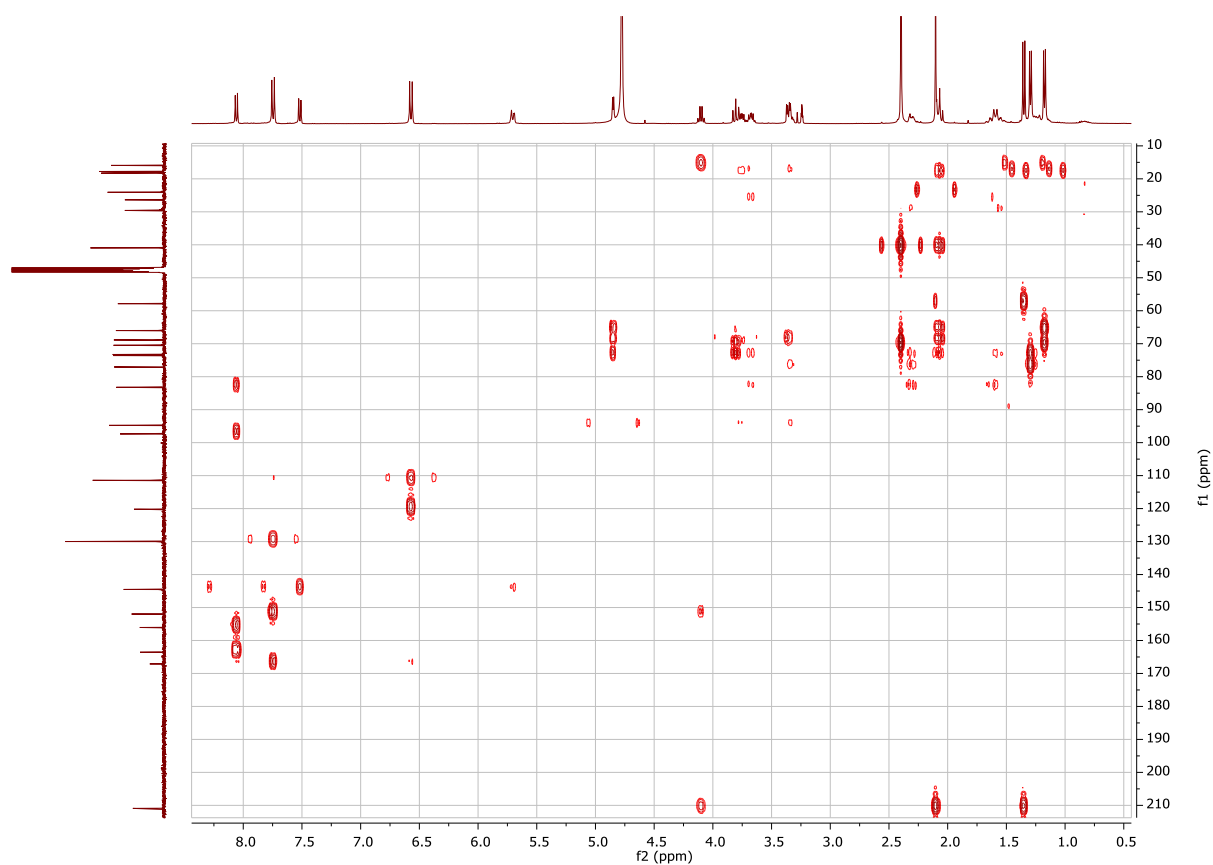


Figure S35. HMBC spectrum of Compound XII (in MeOD)

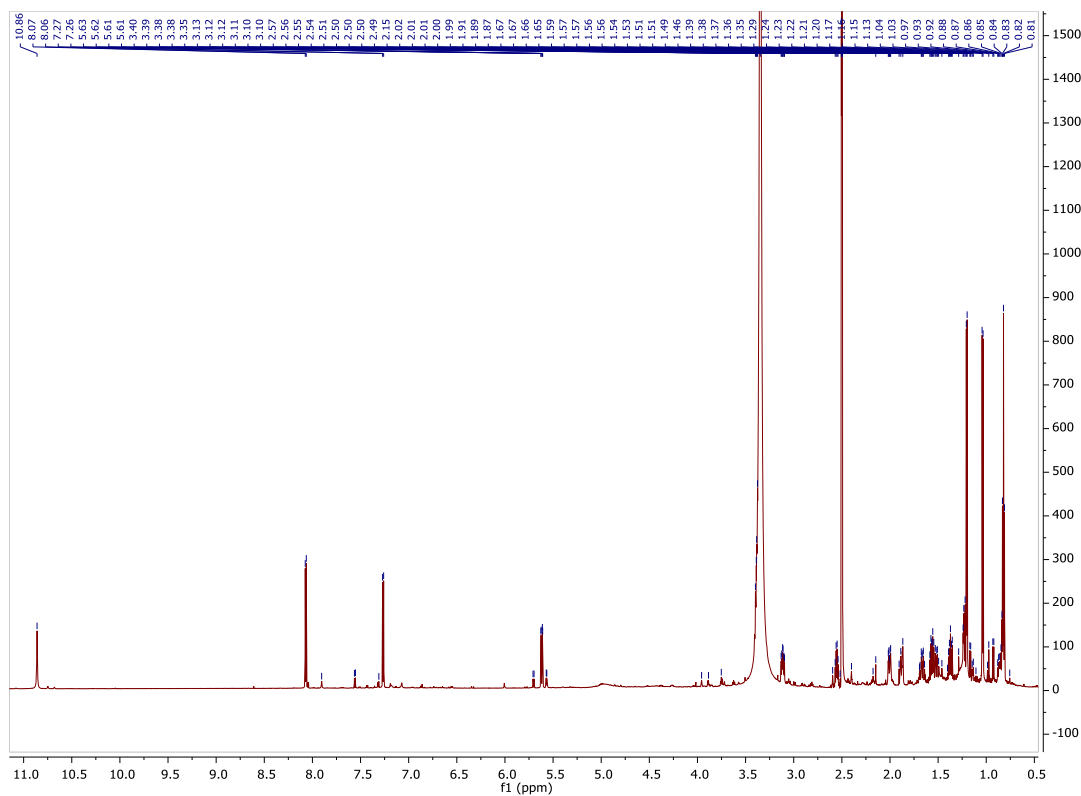


Figure S36. ^1H NMR spectrum of new streptocytosine (700 MHz, d_6 -DMSO)

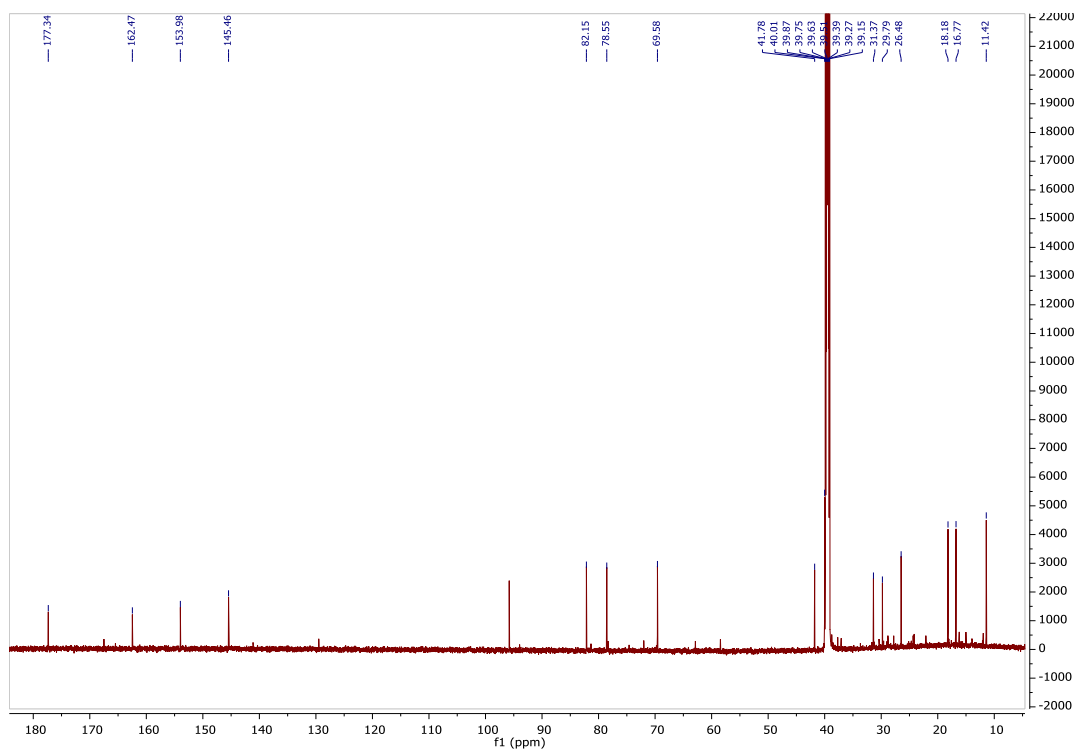


Figure S37. ^{13}C NMR spectrum of new streptocytosine (700 MHz, d_6 -DMSO)

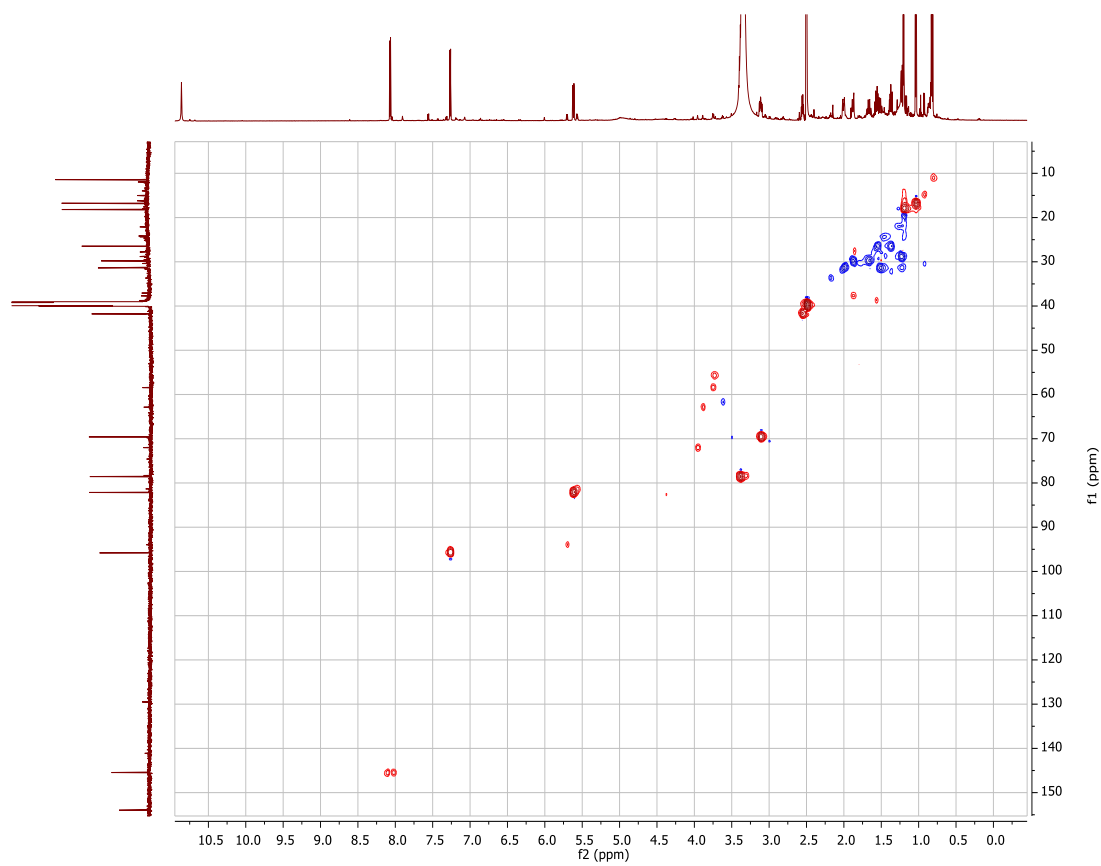


Figure S38. HSQC spectrum of **new streptocytosine** (700 MHz, d_6 -DMSO)

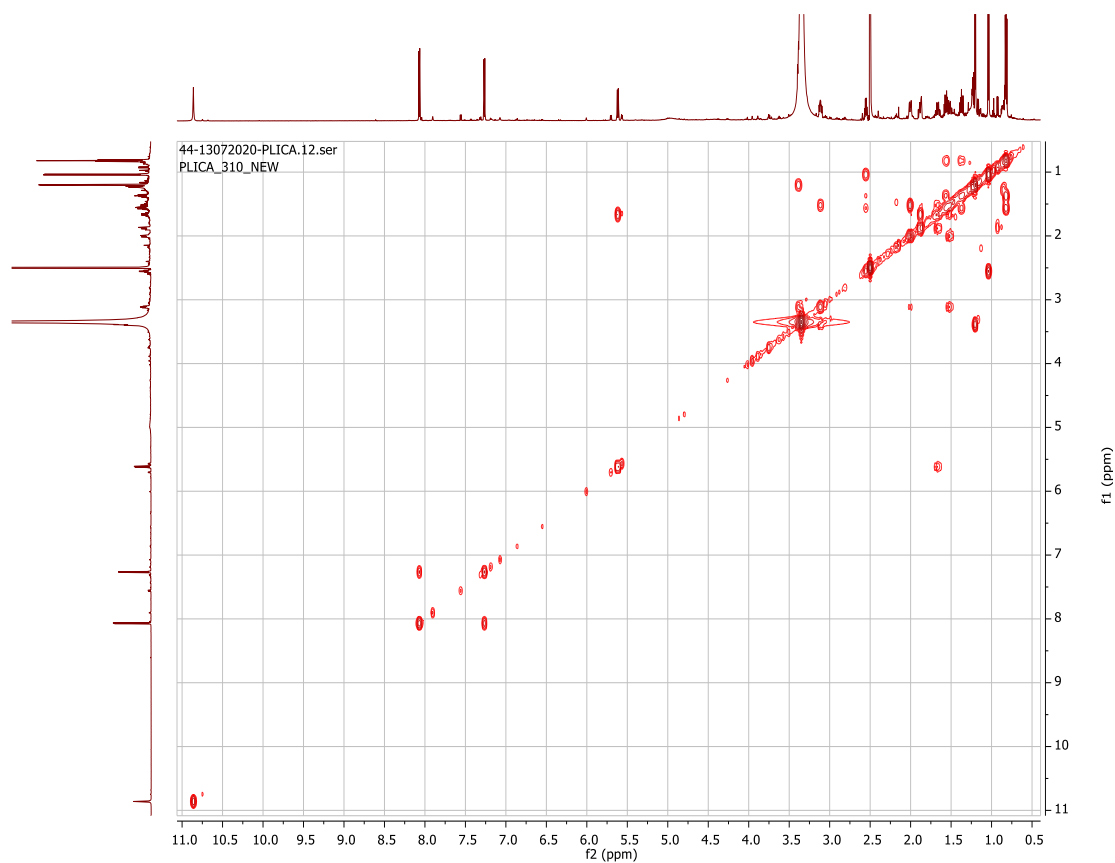


Figure S39. COSY spectrum of **new streptocytosine** (700 MHz, d_6 -DMSO)

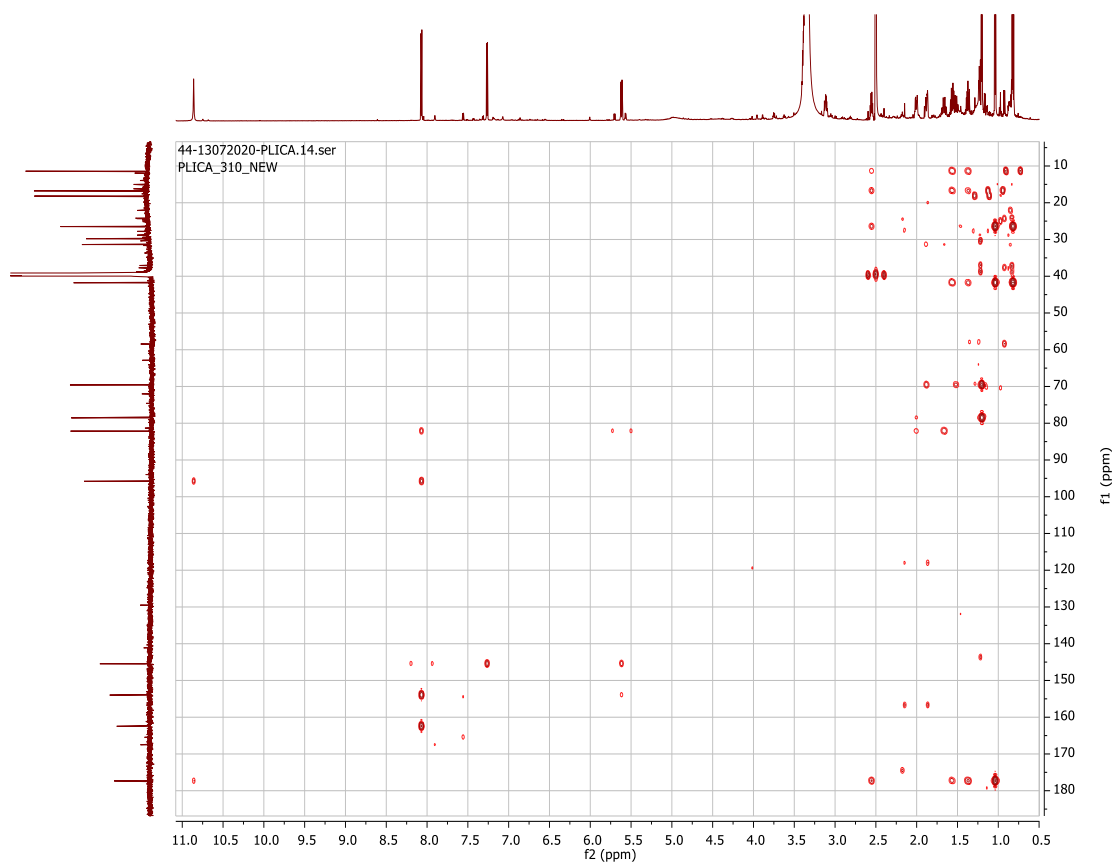


Figure S40. HMBC spectrum of new streptocytosine (700 MHz, d_6 -DMSO)

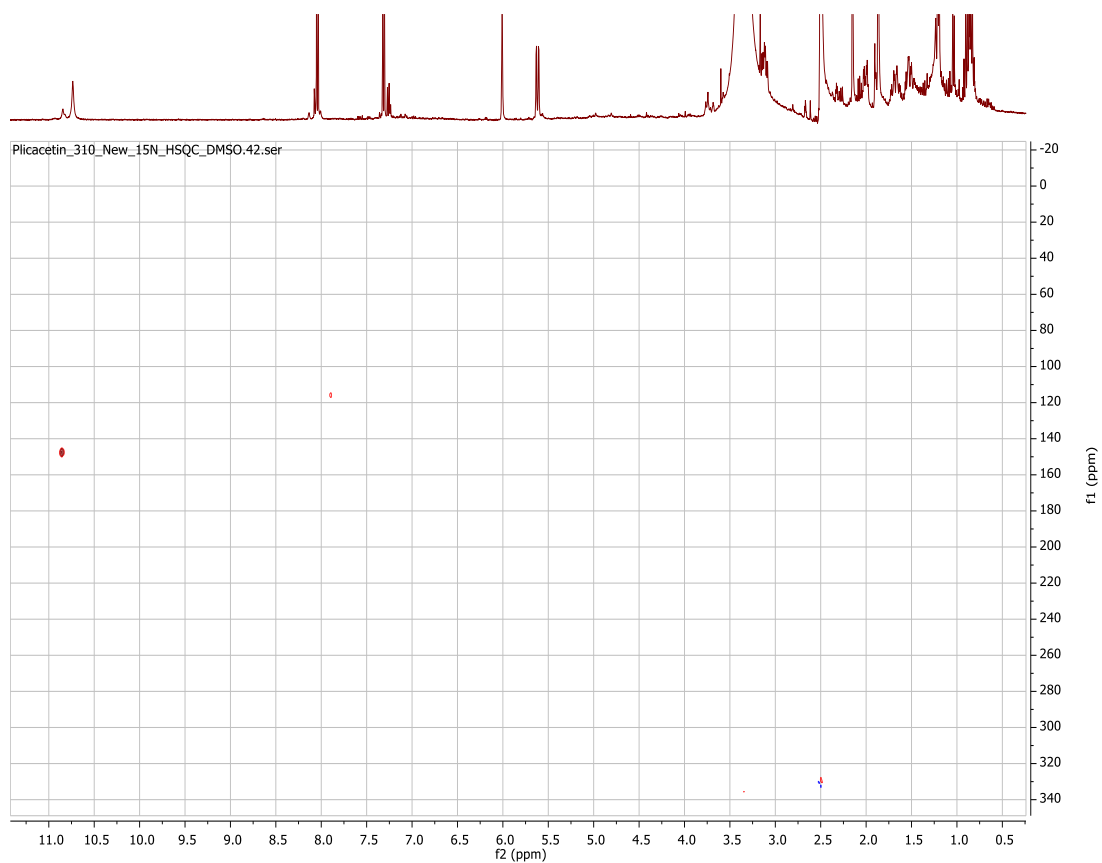


Figure S41. ^{15}N - ^1H spectrum of new streptocytosine (700 MHz, d_6 -DMSO)

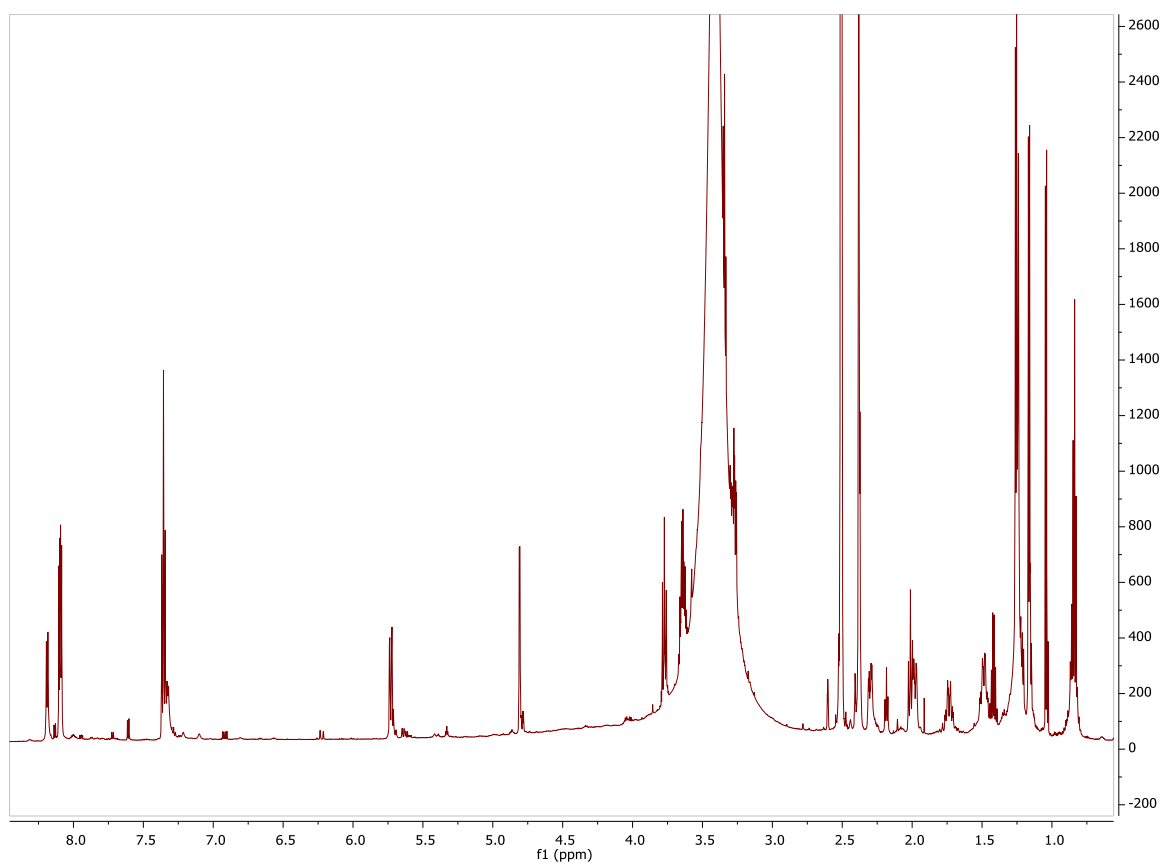


Figure S42. ^1H NMR spectrum of **4F-plicacetin** (700 MHz, d_6 -DMSO)

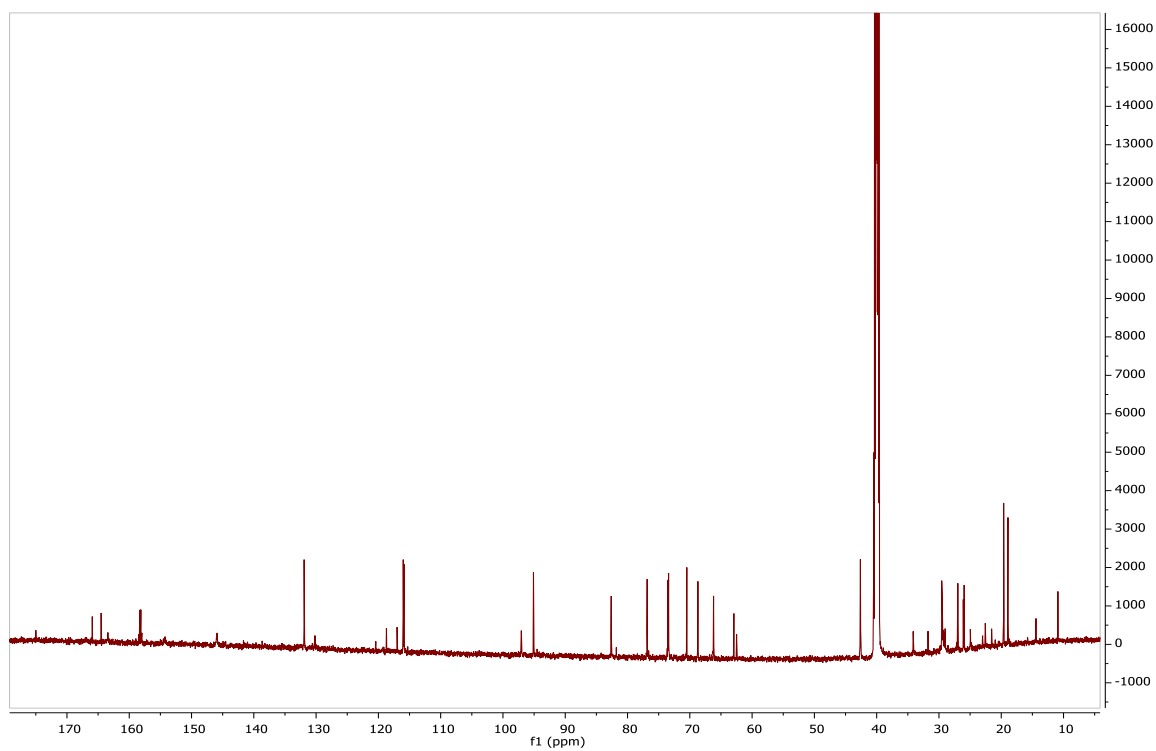


Figure S43. ^{13}C NMR spectrum of **4F-plicacetin** (700 MHz, d_6 -DMSO)

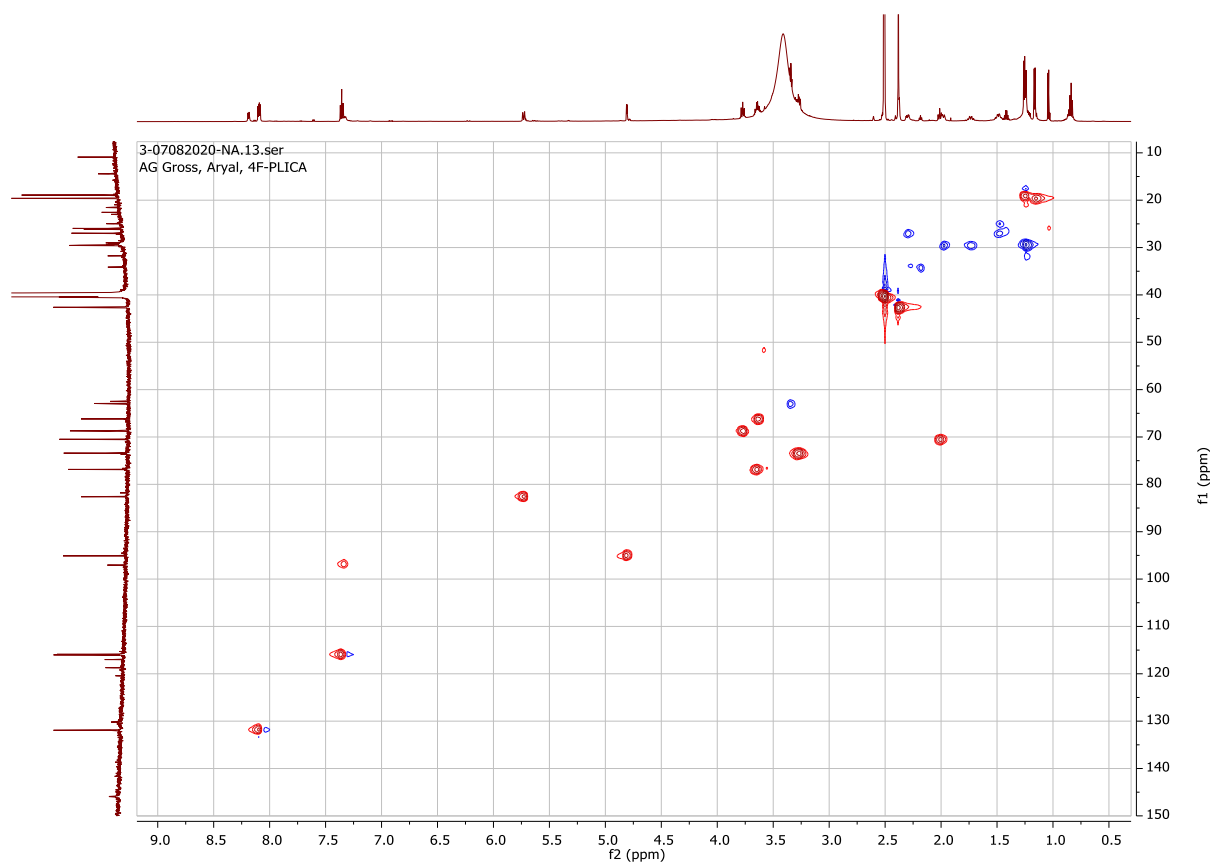


Figure S44. ¹H-¹³C HSQC spectrum of 4F-plicacetin (700 MHz, *d*₆-DMSO)

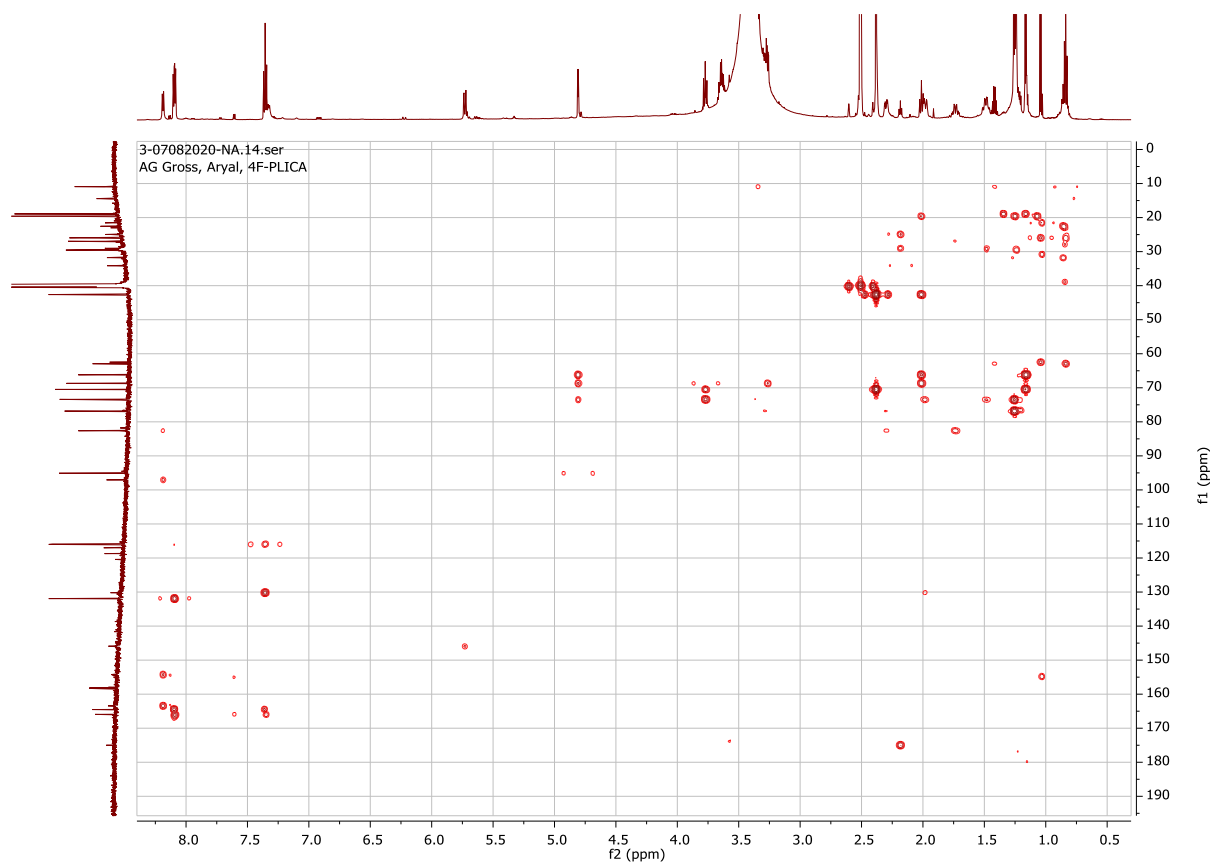


Figure S45. HMBC spectrum of 4F-plicacetin (700 MHz, *d*₆-DMSO)

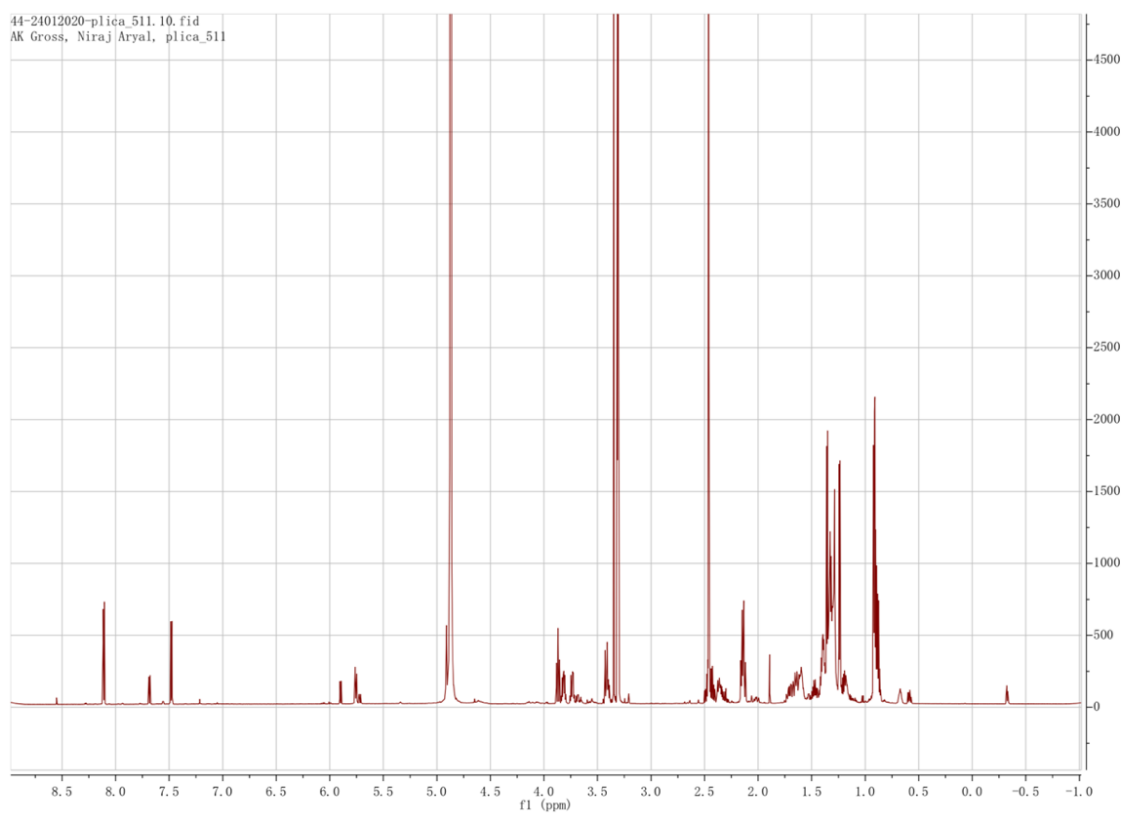


Figure S46. ^1H -NMR spectrum of **compound XI** (700 MHz, d_4 -methanol)

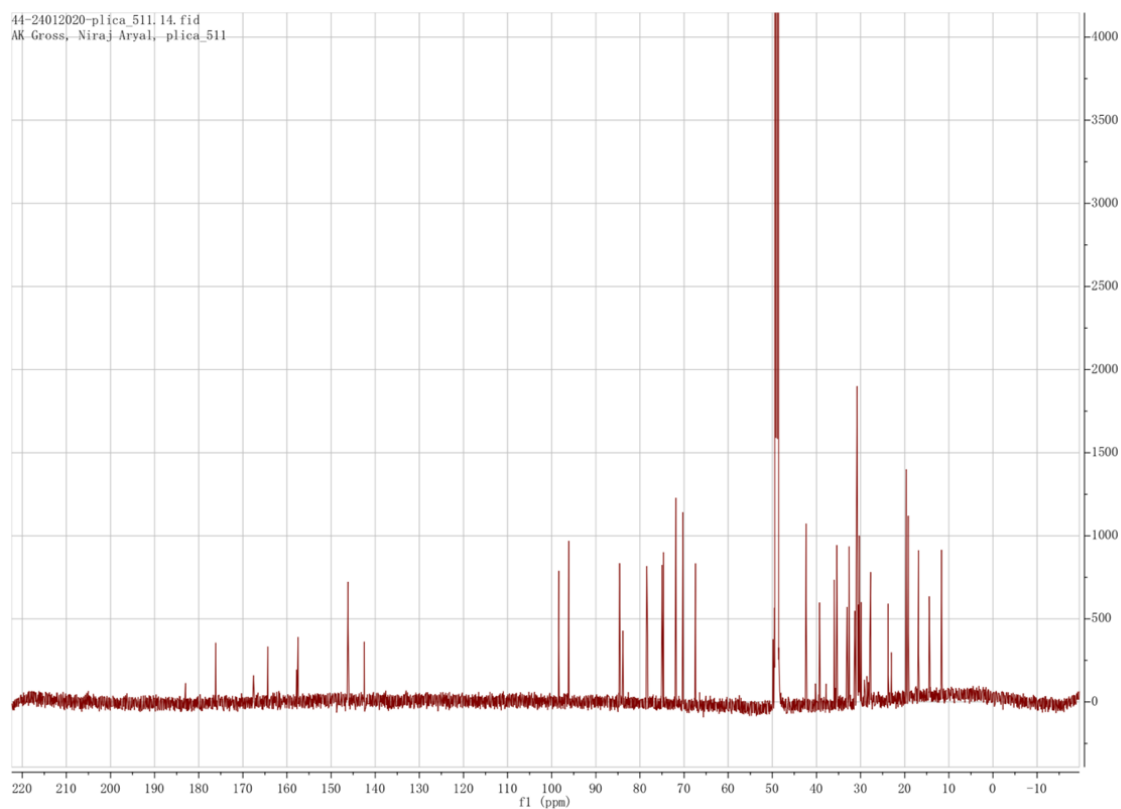


Figure S47. ^{13}C -NMR spectrum of **compound XI** (175 MHz, d_4 -methanol)

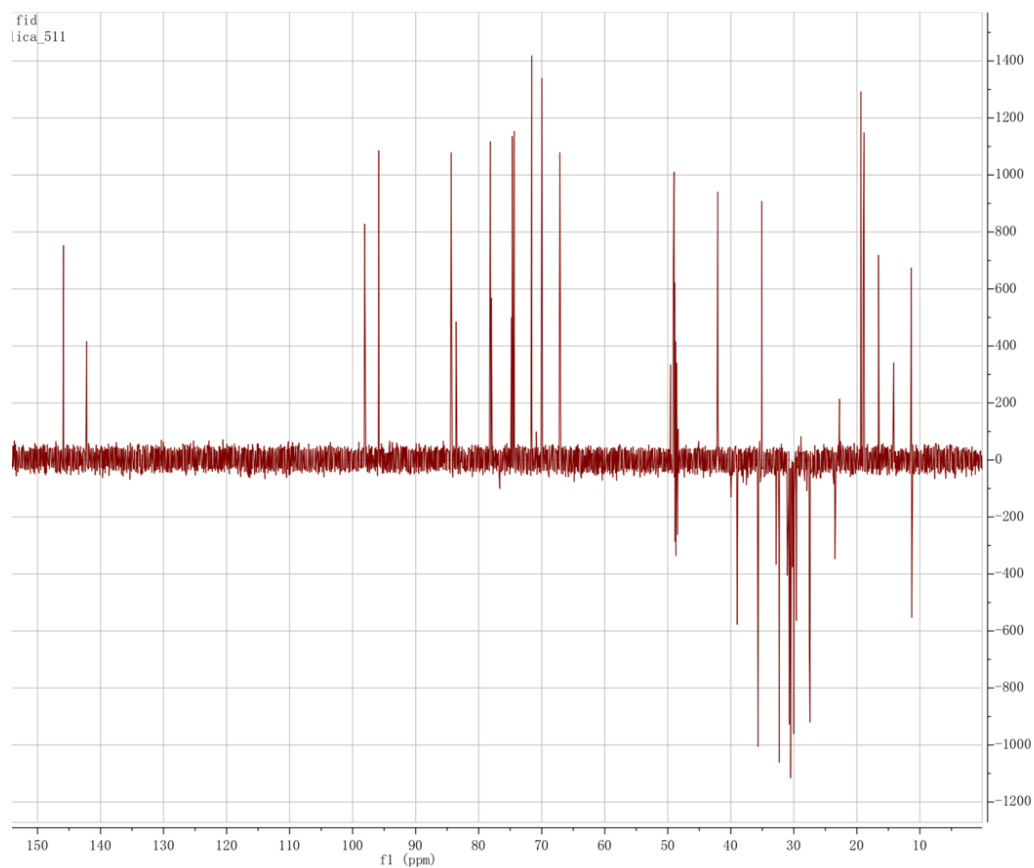


Figure S48. DEPT135 NMR spectrum of **compound XI** (700 MHz, d_4 -methanol)

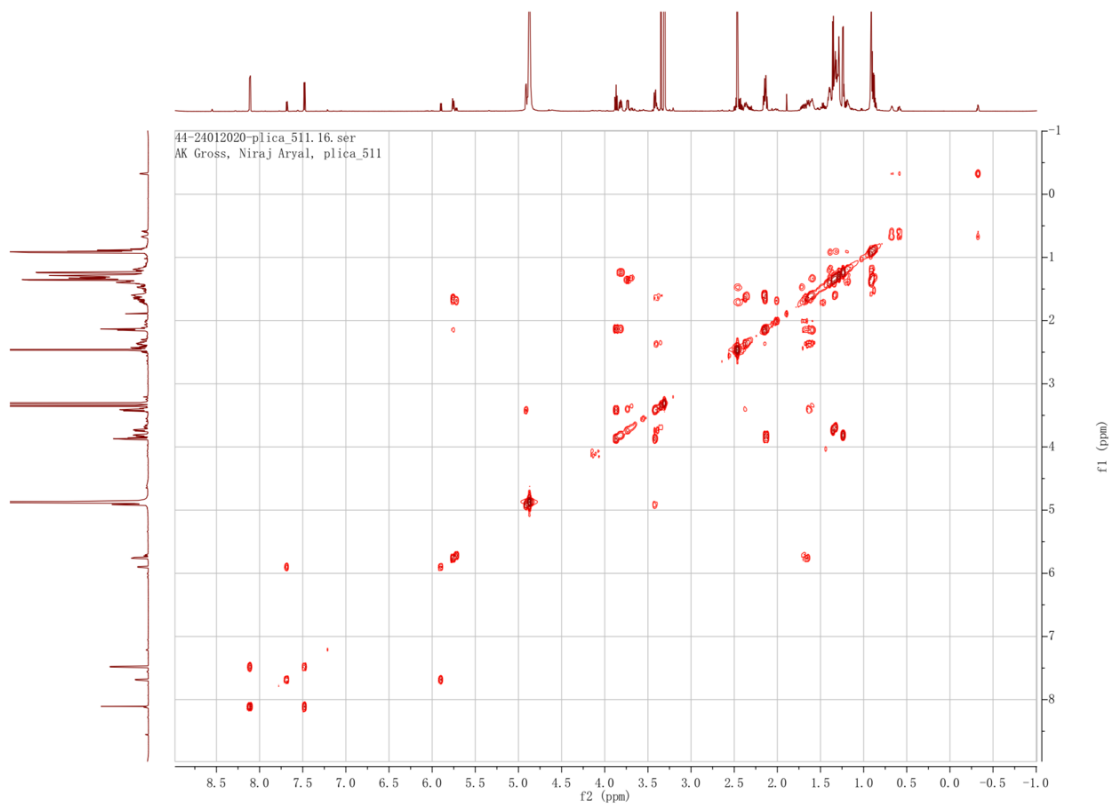


Figure S49. ^1H - ^1H COSY spectrum of **compound XI** (700 MHz, d_4 -methanol)

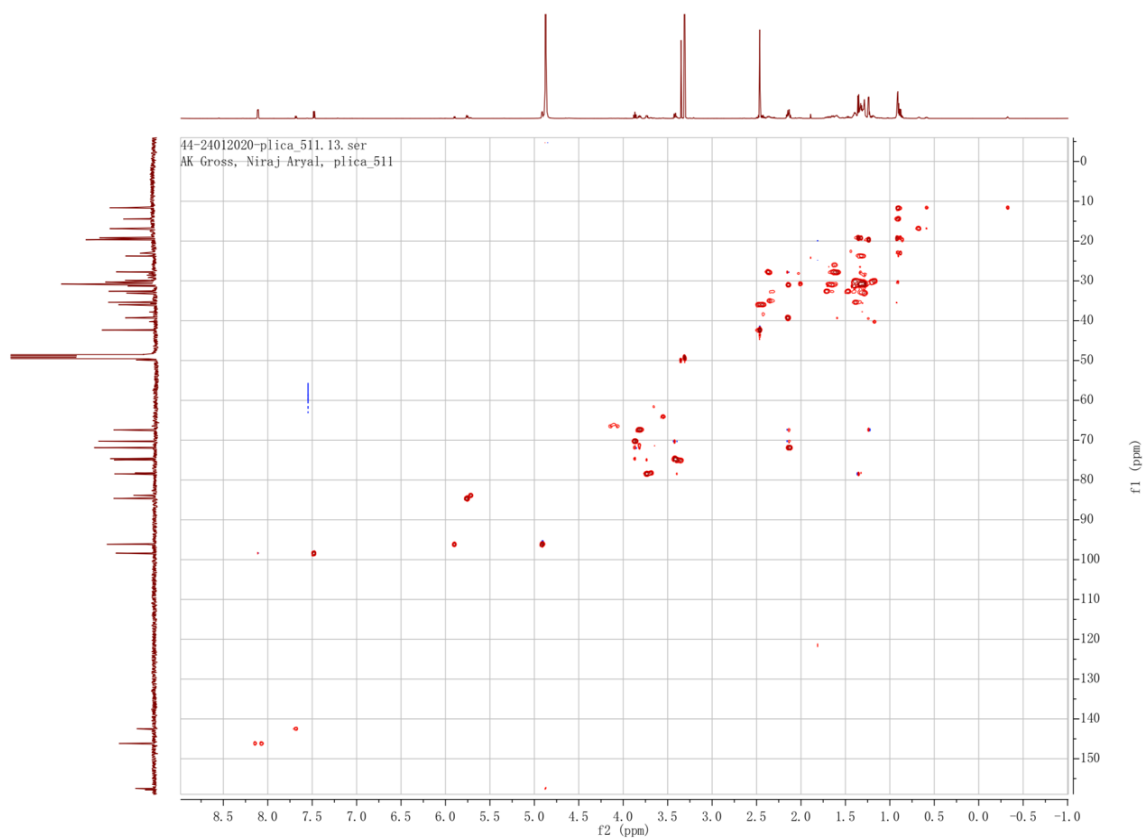


Figure S50. ^1H - ^{13}C HSQC spectrum of **compound XI** (700 MHz, d_4 -methanol)

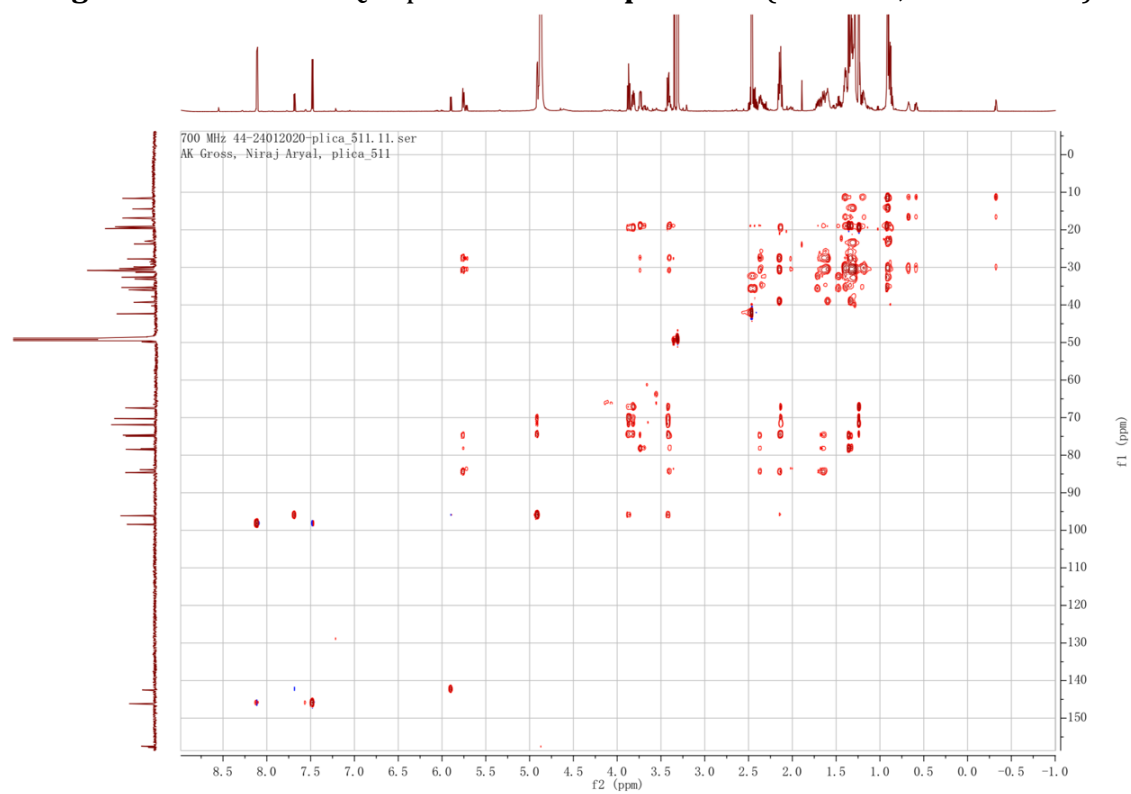


Figure S51. ^1H - ^{13}C HMBC spectrum of **compound XI** (700 MHz, d_4 -methanol)

Supplementary

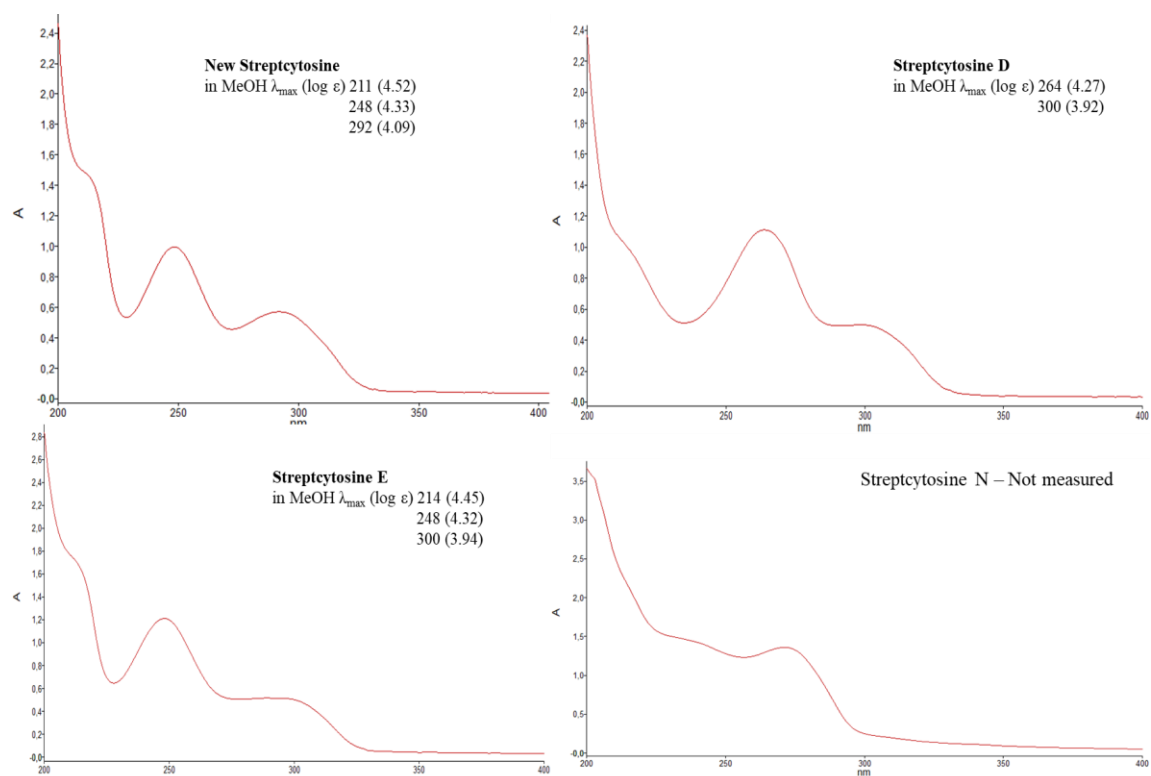


Figure S52. λ_{\max} (log ϵ) values for streptocytosines

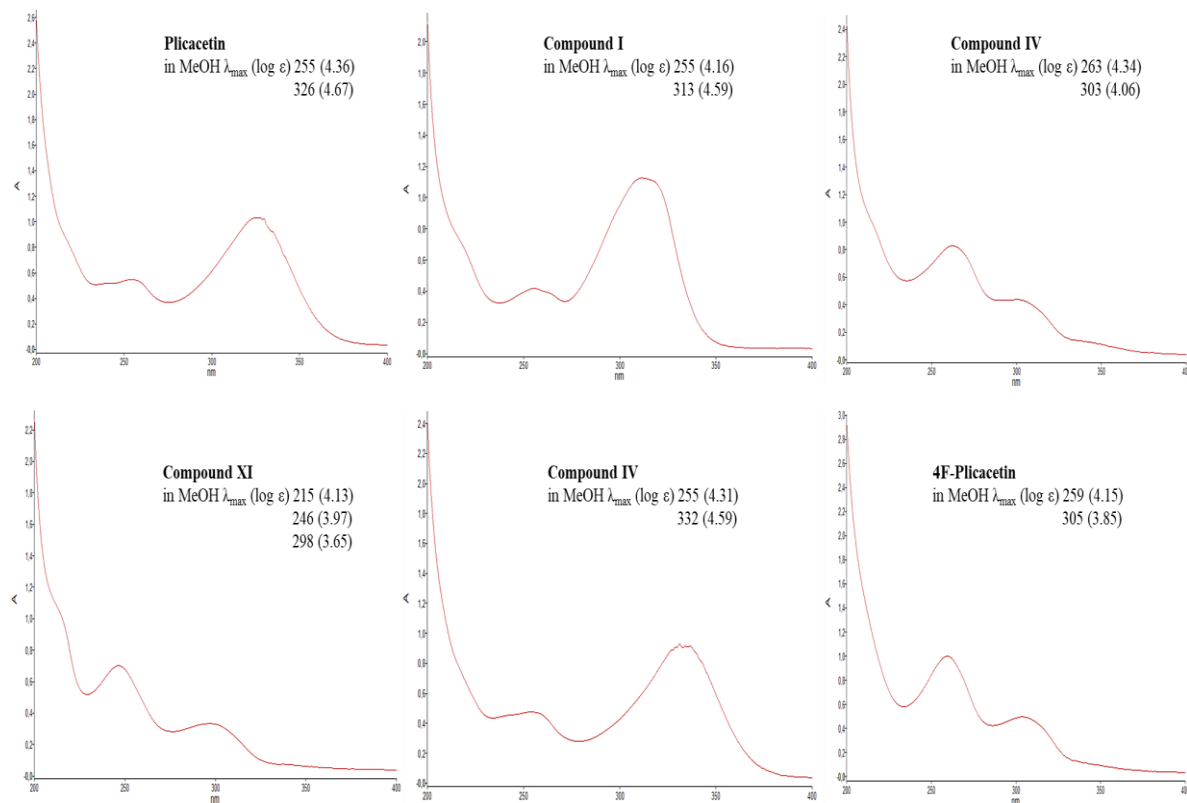
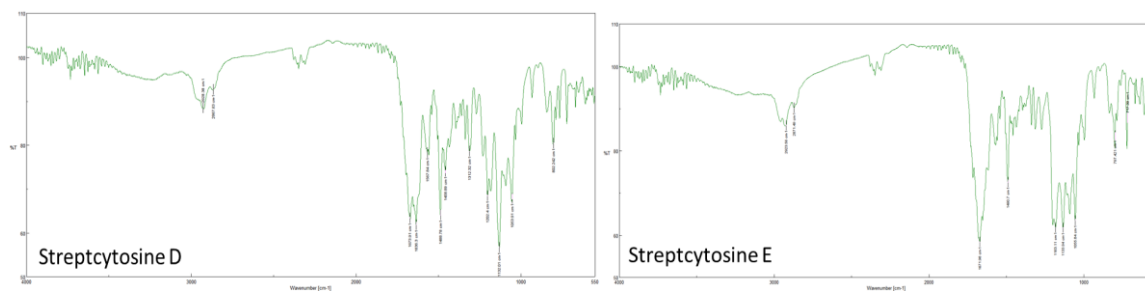


Figure S53. λ_{\max} (log ϵ) values for Plicacetin and isolated new derivatives

Compound Name	Specific rotation = α^{23}_D
Plicacetin	+92.30
Compound I	+95.12
Compound IV	+100.56
Compound XI	+29.76
Compound XII	+92.30
Streptcytosine E	+50
Streptcytosine D	+33.33
Streptcytosine New*	Not measured
4F-Plicacetin	+47.42

Table S2. Optical Rotation for all isolated compounds**Figure S54.** FT-IR spectrum of Streptcytosine D and E

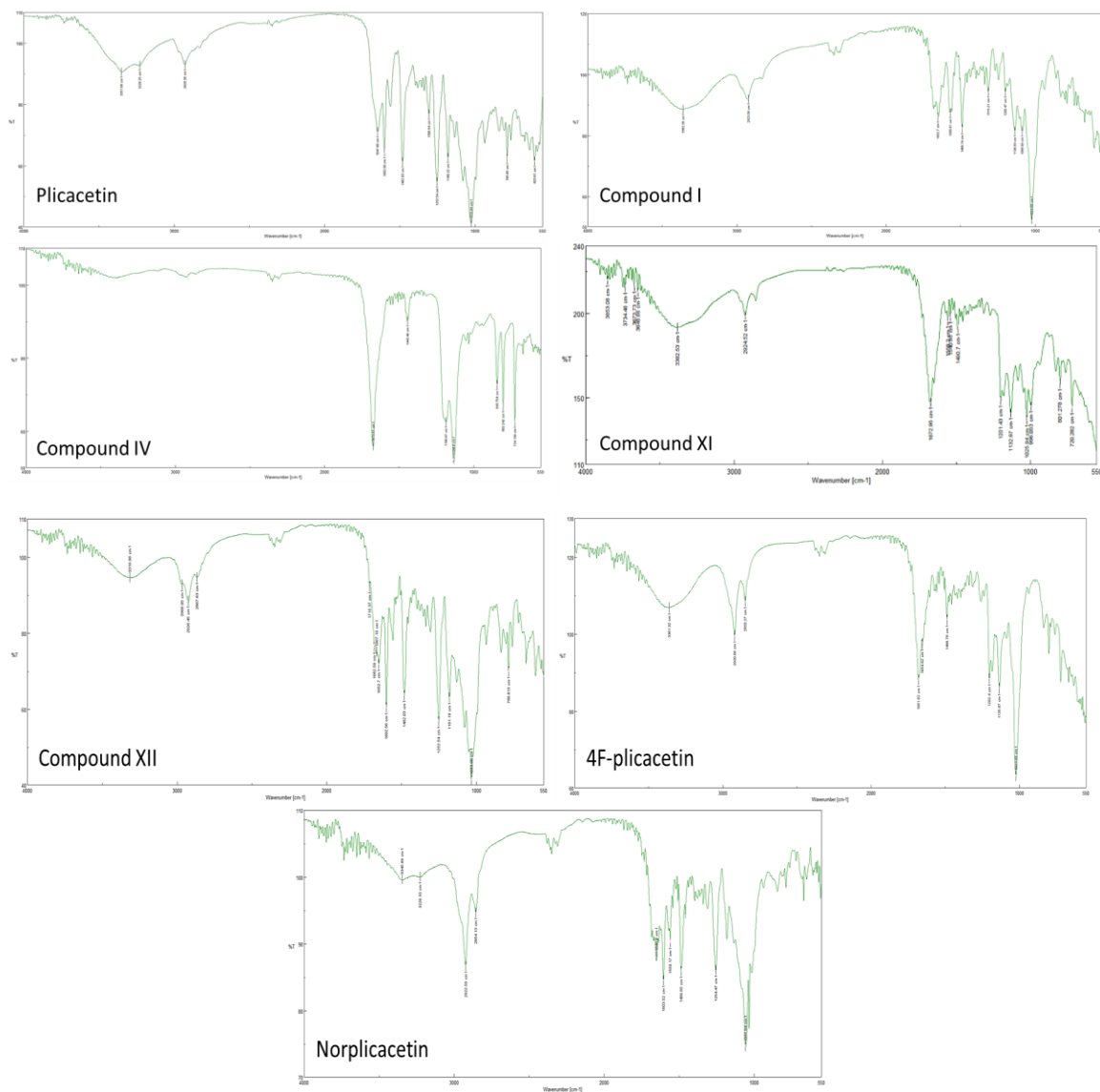


Figure S55. FT-IR spectrum of Plicacetin and derivatives

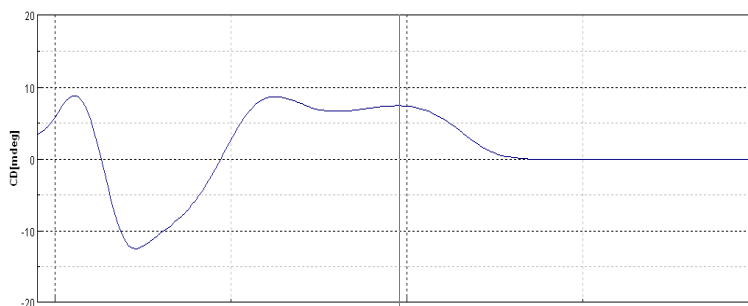


Figure S56. CD spectrum of New Streptocytosine

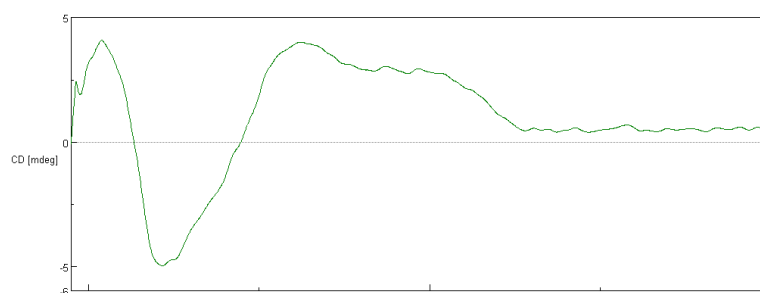


Figure S57. CD spectrum of compound XI

A.

Concentration [$\mu\text{g/ml}$]	Tu01 – <i>C. albicans</i>	Tu02 – <i>C. albicans</i>	Tu03 – <i>C. albicans</i>	Tu04 – <i>C. glabrata</i>	Tu05 – <i>C. tropicalis</i>
Brartemycin	> 32	> 32	> 32	> 32	> 32
New Brartemycin derivat	> 32	> 32	> 32	> 32	> 32
Orsellinic acid ester	> 32	> 32	> 32	> 32	> 32
Brartemycin 492	> 32	> 32	> 32	> 32	> 32
Caspofungin	0,06 – 0,5	0,06 – 0,25	0,125	0,25 – 0,5	0,25 – 0,5

B.

	MIC in $\mu\text{g/ml}$			
	Brartemycin Compound 3	Compound 2.1	Compound 4	Compound 1
<i>E.faecium</i> BM4147-1	>32	>32	>32	>32
<i>S.aureus</i> ATCC29213	>32	>32	>32	>32
<i>K.pneumoniae</i> ATCC12657	>32	>32	>32	>32
<i>A.baumannii</i> 09987	>32	>32	>32	>32
<i>P.aeruginosa</i> ATCC27853	>32	>32	>32	>32
<i>E.aerogenes</i> ATCC13048	>32	>32	>32	>32
<i>E.coli</i> ATCC25922	>32	>32	>32	>32
<i>B.subtilis</i> 168	>32	>32	>32	>32
<i>M.smegmatis</i> mc² 155	>32	>32	>32	>32
	IC ₅₀ in $\mu\text{g/ml}$			
HeLa	>64	>64	>64	>64

Figure S58. Antimicrobial and antifungal assay of brartemycin and its derivative

