Expanding the chemical space of microbial specialized metabolites: structure elucidation and biosynthesis of novel bioactive natural products from actinomycetes

Daniel Carretero Molina

PhD Thesis
Granada
2024

UNIVERSIDAD
DE GRANADA
Doctoral Programme in Pharmacy

Supervisors:
José Fernando Reyes Benítez
Francisco Javier Ortiz López

UNIVERSIDAD DE GRANADA

DOCTORAL PROGRAMME IN PHARMACY

Expanding the chemical space of microbial specialized metabolites: structure elucidation and biosynthesis of novel bioactive natural products from actinomycetes

Ampliación del espacio químico de metabolitos especializados de origen microbiano. Elucidación estructural y biosíntesis de nuevos productos naturales bioactivos procedentes de actinomicetos

MEDINA

Chemistry Department at Fundación MEDINA

Daniel Carretero Molina

Supervisors:

José Fernando Reyes Benítez

Francisco Javier Ortiz López

Doctoral Thesis
Granada, 2024

The present international doctoral thesis has been conducted at the Chemistry Department of Fundación MEDINA under the supervision of Dr. José Fernando Reyes Benítez and Dr. Francisco Javier Ortiz López, within the research area of New Therapeutic Targets of the Doctoral Program in Pharmacy (B15.56.1) at the School of Doctoral Studies in Health Sciences of the University of Granada. This research work has been primarily funded by (1) the IIMENA project (Integrating Informatics and Metabolic Engineering for the Biosynthesis of Novel Antibiotics), awarded by the Challenge program from the Novo Nordisk Foundation, Denmark [NNF16OC0021746] and partially funded by the (2) European Union's Seventh Framework Programme for research, technological development, and demonstration under grant agreement no 312184 (PharmaSea) and the (3) Ministerio de Economía y Competitividad under the Youth Employment Initiative of the European Social Fund (PEJ-2014-A-29071).

To obtain the International Mention, a 3-month research stay has been conducted at the Novo Nordisk Foundation Center for Biosustainability (DTU-Biosustain), under the supervision of Prof. Dr. Tilmann Weber and Dr. Tetiana Gren, where studies were carried out on:

o The identification of the Biosynthetic Gene Cluster (BCG) responsible for producing the gargantulide macrolides using the bacterial version of the antiSMASH software. Familiarization with other bioinformatics tools and databases for genome analysis, such as MiBIG and BIG-SCAPE.

o The complementary combination of NMR results with the genome-based bioinformatic analysis of the Type I PKS gene cluster of gargantulides to fully assign their absolute configurations.

The results from this thesis were published in the following peer-reviewed journals:

1. (2022) **Carretero-Molina, D.;** Ortiz-López, F. J.; Gren, T.; Oves-Costales, D.; Martín, J.; Román-Hurtado, F.; Jørgensen, T.S.; de la Cruz, M.; Diaz, C.; Vicente, F.; Blin, K.; Reyes, F.; Weber, T.; Genilloud, O. Discovery of gargantulides B and C, new 52-membered macrolactones from *Amycolatopsis* sp. Complete absolute stereochemistry of the gargantulide family. *Org. Chem. Front.*, **2022**, *9*, 462-470. doi: 10.1039/D1QO01480C.

2. (2021) **Carretero-Molina, D.;** Ortiz-López, F.J.; Martín, J.; González, I.; Sánchez-Hidalgo, M.; Román-Hurtado, F.; Díaz, C.; de la Cruz, M.; Genilloud, O.; Reyes, F. Pentaminomycins F and G, Nonribosomal Peptides Containing 2-Pyridylalanine. *J. Nat Prod.*, **2021**, *84*(4):1127-1134. doi: 10.1021/acs.jnatprod.0c01199.

3. (2020) **Carretero-Molina, D.;** Ortiz-López, F.J.; Martín, J.; Oves-Costales, D.; Díaz, C.; de la Cruz, M.; Cautain, B.; Vicente, F.; Genilloud, O.; Reyes, F. New Napyradiomycin Analogues from *Streptomyces* sp. Strain CA-271078. *Mar Drugs.* **2020**, 18(1):22. doi: 10.3390/md18010022.

Other scientific publications during PhD period:

1. (2023) Ortiz-López, F.J.; Oves-Costales, D.; **Carretero-Molina, D.;** Martín, J.; Díaz, C.; de la Cruz, M.; Román-Hurtado, F.; Álvarez-Arévalo, M.; Jørgensen, T.S.; Reyes, F.; Weber, T.; Genilloud, O. Crossiellidines A-F, Unprecedented Pyrazine-Alkylguanidine Metabolites with Broad-Spectrum Antibacterial Activity from *Crossiella* sp. *Org Lett.* **2023**. doi: 10.1021/acs.orglett.3c01088.

2. (2022) Aldholmi, M.; Ahmad, R.; **Carretero-Molina, D.**; Pérez-Victoria, I.; Martín, J.; Reyes, F.; Genilloud, O.; Gourbeyre, L.; Gefflaut, T.; Carlsson, H.; Maklakov, A.; O'Neill, E.; Field, R. A.; Wilkinson, B.; O'Connell, M.; Ganesan, A. Euglenatides, Potent Antiproliferative Cyclic Peptides Isolated from the Freshwater Photosynthetic Microalga Euglena gracilis. *Angew. Chem. Int. Ed.* **2022**, *61*, e202203175; *Angew. Chem.* **2022**, *134*, e202203175. doi: 10.1002/anie.202203175.

3. (2021) Beck C.; Gren, T.; Ortiz-López, F.J.; Jørgensen, T.S.; **Carretero-Molina, D.**; Martín Serrano, J.; Tormo, J.R.; Oves-Costales, D.; Kontou, E.E.; Mohite, O.S.; Mingyar, E.; Stegmann, E.; Genilloud, O.; Weber, T. Activation and Identification of a Griseusin Cluster in *Streptomyces* sp. CA-256286 by Employing Transcriptional Regulators and Multi-Omics Methods. *Molecules.* **2021**, 26(21):6580. doi: 10.3390/molecules26216580.

4. (2021) Román-Hurtado, F.; Sánchez-Hidalgo, M.; Martín, J.; Ortiz-López, F.J.; **Carretero-Molina, D.**; Reyes, F.; Genilloud, O. One Pathway, Two Cyclic Non-Ribosomal Pentapeptides: Heterologous Expression of BE-18257 Antibiotics and Pentaminomycins from *Streptomyces cacaoi* CA-170360. *Microorganisms*, **2021**, *9*(1), 135. doi: 10.3390/microorganisms9010135.

5. (2020) Ortiz-López, F. J.; **Carretero-Molina, D.**; Sánchez-Hidalgo, M.; Martín, J.; González, I.; Román-Hurtado, F.; de la Cruz, M.; García-Fernández, S.; Reyes, F.; Deisinger, J. P.; Müller, A.; Schneider, T.; Genilloud, O. First Member of the New Lanthidin RiPP Family. *Angew. Chem. Int. Ed.* **2020**, *59*, 12654. doi: 10.1002/anie.202005187.

Contributions to scientific conferences and meetings:

1. (2023) Ortiz-López, F.J.; Oves-Costales, D.; **Carretero-Molina, D.**; Martín, J.; Díaz, C.; de la Cruz, M.; Román-Hurtado, F.; Álvarez-Arévalo, M.; Jørgensen, T.S.; Reyes, F.; Weber, T.; Genilloud, O. Crossiellidines A-F, Unprecedented Pyrazine-Alkylguanidine Metabolites with Broad-Spectrum Antibacterial Activity from *Crossiella* sp. 31st International Symposium on the Chemistry of Natural Products & 11th International Congress on Biodiversity. October 15th to 19th, 2023, Naples, Italy

2. (2023) Ortiz-López, F.J.; Oves-Costales, D.; **Carretero-Molina, D.**; Martín, J.; Díaz, C.; de la Cruz, M.; Román-Hurtado, F.; Álvarez-Arévalo, M.; Jørgensen, T.S.; Reyes, F.; Weber, T.; Genilloud, O. Crossiellidines A–F, Unprecedented Pyrazine-Alkylguanidine Metabolites with Broad-Spectrum Antibacterial Activity from *Crossiella* sp. RICT 2023, 57th International Conference on Medicinal Chemistry. Drug Discovery and Selection. July 5th to 7th, 2023, Lille, France.

3. (2019) Ortiz-López, F.J.; **Carretero-Molina, D.**; Martín, J.; de la Cruz, M.; Sánchez, M.; Díaz, C.; González, I.; Morosini, M.I.; Vicente, F.; Reyes, F.; Deisinger, J.; Müller, A.; Schneider, T.; Genilloud, O. Discovery of MDN-0207, a novel glycosylated lanthipeptide with unusual structural features and potent antibacterial activity. First International Conference on RiPPs (Ribosomally synthesized and Post-translationally modified Peptides). April 24th to 26th, 2019, Granada, Spain.

4. (2019) **Carretero-Molina, D.**; Ortiz-López, F.J.; Díaz, C.; de la Cruz, M.; González, I.; Reyes, F.; Vicente, F.; Genilloud, O. New analogues of nosokomycin and polycyclic xanthones with antibacterial properties from MEDINA´s actinomycetes collection. I Congreso de Investigadores del PTS. February 13th to 15th, 2019, Granada, Spain.

5. (2018) Ortiz-López, F.J.; **Carretero-Molina, D.**; Díaz, C; de la Cruz, M; González, I; Reyes, F; Vicente, F; Genilloud, O. New antibacterial compounds from MEDINA´s actinomycetes collection. 3rd European Conference on Natural Products (DECHEMA). September 2nd to 5th, 2018, Frankfurt am Main, Germany.

6. (2018) **Carretero-Molina, D.**; Ortiz-López, F.J.; Oves-Costales, D.; Román, F.; Reyes, F.; Vicente, F.; Lee, S.Y.; Palsson, B.; Weber, T.; Genilloud, O. The IIMENA project: Integration of Informatics and Metabolomic Engineering for the Discovery of Novel Antibiotics". I Euroindoamerican Natural Products Meeting (EIAMNP). May 29th to June 1st, 2018, Madrid, Spain.

7. (2018) Ortiz-López, F.J.; **Carretero-Molina, D.**; Martín, J; de la Cruz, M.; Sánchez, M.; Díaz, C.; González, I.; Morosini, M.I.; Vicente, F.; Reyes, F.; Deisinger, J.; Müller, A.; Schneider, T.; Genilloud, O. Discovery of MDN-0207, an unprecedented glycosylated lanthipeptide with unusual cell wall mode of action. Gordon Research Conferences 2018. March 4th to 9th, 2018, Ventura, California, USA.

A Luca Hidalgo Carretero,
famoso en el mundo entero.

"In scientific thinking are always present elements of poetry. Science and music requires a thought homogeneous."

— Albert Einstein

Acknowledgments

A lo largo de mi etapa como doctorando he tenido la oportunidad de compartir diferentes partes del trayecto con muchas personas inspiradoras, tanto para el desarrollo profesional como personal. Os quiero dar las gracias a todos los que habéis formado parte de esta aventura por ayudarme a alcanzar la meta.

Me gustaría comenzar agradeciendo a mis directores de tesis, Javier y Fernando. Gracias, Fernando, por brindarme la oportunidad de sumergirme en el mundo de la investigación de los productos naturales. Ha sido una inmersión que ha marcado significativamente la dirección de mi carrera en los últimos años, y el punto de partida fue ese primer voto de confianza. Gracias, paisano, por tu supervisión, los viajes a cursos y congresos y las cenas en la Bodeguilla. Que de esto último sigamos disfrutando aún después de la tesis. Gracias, Javi, por haberme enseñado a usar el paracaídas en el salto al vacío que fue comenzar en este mundillo. Has sido un excelente mentor, al que admiro como investigador, y a quien debo mucho en esta tesis. Gracias por tus consejos, tu total implicación, tu -casi- infinita paciencia, tus minuciosas revisiones y tu capacidad para motivarme. Para mí se quedan las largas horas de trabajo compartidas en el laboratorio, los mano a mano con el Gilson, el curso avanzado para descifrar tu caligrafía encriptada, pero también, los productivos ratos de charlas y discusiones con las que revisitábamos el pasado y descubríamos el futuro, a menudo intercaladas con risas, bromas y humor gaditano. Has sido una pieza clave para que haya llegado hasta aquí y, después de todos estos años de trabajo, más allá de todo lo aprendido, me llevo una amistad que valoro mucho. Gracias a los dos por haber pensado en mí para cada proyecto que surgía, lo que me ha permitido culminar esta etapa, y por vuestro apoyo y comprensión cuando todo se puso cuesta arriba.

Quiero expresar también el más sincero agradecimiento a mi tutora, Mayca. Gracias por tus palabras de ánimo, tu disposición en todo momento para prestarme tu ayuda y tu tiempo. Fuiste de mis mejores profesoras y has sido una excelente tutora.

Agradecer a todos y cada uno de mis colegas de Fundación MEDINA por su profesionalidad, pero sobre todo por su compañerismo. La ciencia en general, y la química en particular, ya son suficientemente desafiantes por sí solas, estar rodeado de buena gente es el mejor material con el que se debe equipar un laboratorio. A pesar de ser muchos los nombres quiero agradecer a un buen puñado de ellos: Eli, Patri, Teresa, Carlos, Paqui, Isa, Fer, Adolfo, Manuela, Rubén, Gloria, Mercedes, Pepe, Freddie, Pili, Rachel, Clara, Berna y todos los que caben en este etcétera. Gracias en particular, Dani Oves, por tu disposición a resolver cualquier consulta que pudiera surgirme y por haberte tomado la molestia de revisar y mejorar este manuscrito. Gracias, Jesús e Ignacio, por vuestra ayuda y por la innumerable cantidad de muestras que me habéis puesto tanto en el HRMS como en el RMN. Gracias, Germán, por tu paciencia cada vez que te he dado la lata con problemas en la VPN.

En especial, quiero mencionar a aquellas personas que empezaron siendo compañeras y se han convertido en grandes amistades. Gracias, Cari y Ángeles, que con los años igualemos con cervezas el número de placas de ensayo con las que os hemos bombardeado durante tanto tiempo. A ti, María, primera alumna a la que tutoricé, espero haberte enseñado algo, desde luego juntos hemos aprendido, sufrido y disfrutado mucho. Carmen, gracias por ese soplo revitalizante que llevas allá por donde vas, y por cargarme de energía con cada abrazo que me has dado.

I would also like to thank my colleagues at DTU for the fruitful collaboration during the IIMENA project and their warm welcome during my time in Copenhagen. I want to extend special thanks to Tilmann Weber and Tetiana Gren for their exceptional guidance and mentorship. I am also very grateful to Tue Sparholt, Kai Blin, Omkar Mohite and Darko Kjiproski for their expertise and assistance. Thanks to the entire team for contributing to the excellent academic and social work environment. It was a very rewarding experience where very good results were achieved. Jeg takker jer for det.

Expresar el más profundo agradecimiento a mi familia por su apoyo incondicional dentro y fuera de mi trayectoria académica, empezando por mis padres, Daniel y Pastora. Gracias, mamá, por tu amor infinito y tu dedicación incansable hacia los tuyos. Eres mi mayor ejemplo de bondad y generosidad. Papá, espero que tu empeño constante en transmitirnos e inculcarnos los valores del esfuerzo, trabajo y perseverancia se vea recompensado de alguna manera con esta tesis. Si he llegado hasta aquí es gracias a vosotros. Espero tener la oportunidad de devolveros al menos un pequeño porcentaje de todo lo que me habéis dado.

Gracias, tata, por haber sido la hermana mayor, con todo lo que eso implica, y mi figura de mujer fuerte y empoderada más cercana. Fuiste la principal influencia en mi trayectoria académica y pude aprovechar tus consejos, apuntes y clases particulares. Gracias por haberme facilitado la apertura de tantas puertas en lo personal, académico y profesional. José Pedro, gracias por asumir la temeridad de entrar en esta familia. Pero, sobre todo, gracias a ambos por el sobrino que me habéis dado. Lucrecio, gracias por esa sonrisa y el amor más puro que he conocido. Poder jugar contigo como el niño que aún vive dentro de mí es un regalo. Vete acostumbrando al 'maluco' de tu tío, que tienes *pa´* rato.

Quiero acordarme también de mis abuelos y abuelas (*in memoriam*). Antonia, Teresa, Manolo e Isidoro, me habría encantado poder celebrar esta victoria con vosotros. En especial, gracias a mi abuelo Isidoro, no hay duda de que hay una fuerte impronta genética suya, eso que tanto me costaba explicarle, en mi ADN. Gracias al resto de titas, titos, primas y primos por el cariño y el cuidado, sobre todo en tiempos difíciles. Vuestra presencia ha sido un gran consuelo.

Gostaria de expressar minha sincera gratidão à minha família brasileira por terem acolhido com tanto amor e carinho um "intruso" como eu. Obrigado, Sergio e Regina, pela confiança e por terem influenciado sua filha pra que ela se tornasse esta pessoa tão maravilhosa.

Siguiendo con la familia, quiero agradecer y felicitarme por el grupo de amigos del que he tenido el privilegio de rodearme a lo largo de los años. Sin duda mi mayor logro, del que me siento corresponsable. Gracias Víctor, Andrés Jimmy, Yisus, Antonello, Alonso, Jésus, Juan, Guillermo, López, Diego, Borja, Daviyo, Charles y algunos más, por todos los viajes, rutas en bici y travesías en la Gittana que hemos disfrutado juntos y por las que nos embarcaremos en el futuro. Gracias también al Consejo, porque Mija siga resurgiendo siempre por Navidad.

Por otro lado, quiero agradecer a todos mis compañeros de la Tuna de la Facultad de Ciencias de la Universidad de Granada. De los grandes aciertos durante mi vida universitaria, que me ha brindado, y lo sigue haciendo, memorables ratos de música, viajes e historias para el recuerdo. Que su andadura perdure muchos años más. ¡Aúpa Tuna!

Ich möchte mich auch bei all meinen Kollegen vom Musikverein Feuerwehrkapelle für die herzliche Aufnahme bedanken. Ihr habt uns das Gefühl gegeben, ein Teil von euch zu sein. Ich hätte nie gedacht, dass das Saxophon das beste Instrument sein würde, um aus Nußloch unser neues Zuhause zu machen. Gesegnete Musik als universelle Sprache. Vielen Dank an euch alle.

No quiero olvidarme de la grupeta de 'Portugal Corchopanera': Víctor, Alonso, Denise, Pepe y Luci, gracias. Si tiene que ocurrir una tragedia, que al menos ocurra rodeado de gente como vosotros. Gracias por no haberos desentendido ni un segundo en la que ha sido una de mis rutas más duras.

8

Una parte de mis agradecimientos la dedico a todo el personal sanitario que me ha acompañado, tanto de España como de Alemania, pues es también gracias a ellos que he podido culminar esta etapa. Gracias a Juan Tormo, por mantenerme siempre optimista y darme ganas de luchar. A José María Torralba, Eva Bayo, Javier Olmedo, Mabel Vega, Peter Schmidt y a muchos otros por vuestra gran profesionalidad y vuestra cercanía. A veces, por complicado que parezca, pueden conseguirse resultados que desafían a los diagnósticos más desalentadores.

Gracias a mis bichos peludos, Maggie, Jamonito y Pucho, por haber sido la mejor compañía durante la cantidad de horas que he invertido en escribir esta tesis.

Pero, sobre todo, mi mayor agradecimiento es para ti, Denise. Has sido mi pareja, mi enfermera, mi terapeuta y mi cosupervisora de tesis, todo en uno. Te estaré eternamente agradecido por tu amor, tu compañía y tu cariño, que han sido sin duda el bálsamo más sanador de todos. Gracias por haber recorrido conmigo un camino pedregoso y convertirlo en un paseo por la arena a la orillita del mar. Sé que eso que has hecho de forma tan natural ha supuesto un gran sacrificio, ahora es tiempo de tomarle juntos la revancha a estos dos últimos años. Lo mejor que me llevo de mi etapa de doctorado es a ti como compañera de vida. No era fácil destapar el tesoro que esconden los productos naturales, pero lo he encontrado, venía de Brasil y trabajaba en mi departamento. Esta tesis posiblemente no se habría finalizado sin tu ayuda. Mil gracias. Va por ti y por nuestras próximas aventuras. Te quiero.

Abbreviations

AcN	Acetonitrile
ACP	Acyl Carrier Protein
ADC	Antibody Drug Conjugate
AMR	Antimicrobial Resistance
AMT	Aminotransferases
Arg	Arginine
AROs	Aromatases
AT	Acyl transferase
BGC	Biosynthetic Gene Cluster
BiG-SCAPE	Biosynthetic Gene Similarity Clustering and Prospecting Engine
CC	Combinatorial Chemistry
CHS	Chalcone Synthase
CNP	Chemistry of Natural Products
COSY	Correlated Spectroscopy
CYCs	Cyclases
CYP450	Cytochrome P450
Cys	Cysteine
DA	Differential Analysis
DAD	Diode Array Detector
DH	Dehydratase
DMAPP	Dimethylallyl Pyrophosphate
DMSO	Dimethylsulfoxide
ER	Enoyl Reductase
ESI	Electrospray Ionization
ESKAPE	*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa,* and *Enterobacter spp*
EtOAc	Ethyl Acetate
FAD	Flavin Adenine Dinucleotide
FAS	Fatty Acid Synthase
FDA	Food and Drugs Administration
L-FDVA	*N*-a-(2,4-dinitro-5-fluorophenyl)-L-valinamide
FPP	Farnesyl Diphosphate
G3P	Glyceraldehyde-3-phosphate
GC	Gas Chromatography
GGPP	Geranylgeranyl diphosphate
GPP	Geranyl Diphosphate
HGT	Horizontal Gene Transfer
HMBC	Heteronuclear Multiple Bond Correlation

HOX	Hypohalous Acid
HRMS	High-Resolution Mass Spectrometry
HSQC	Heteronuclear Single Quantum Coherence
HTS	High-Throughput Screening
IPP	Isopentenyl pyrophosphate
IR	Infrared
JBCA	*J*-Based Configuration Analysis
KR	Ketoreductase
KS	Ketoacyl synthase
L-2-PAL	3-(2-pyridyl)-L-alanine
LC	Liquid Chromatography
Leu	Leucine
MDR	Multidrug-resistant
MEP	Methyl Erythritol Phosphate
MIBiG	Minimum Information about a Biosynthetic Gene cluster
MIC	Minimum Inhibitory Concentration
MMAE	Monomethylauristatin E
MPLC	Medium Pressure Liquid Chromatography
MRSA	Methicillin-resistant *Staphylococcus aureus*
MSSA	Methicillin-susceptible *Staphylococcus aureus*
MS	Mass Spectrometry
MT	Methyltransferase
NCBI	National Center for Biotechnology Information
NPDs	Napyradiomycins
NPD-As	Type A Napyradiomycins
MVA	Mevalonic Acid Pathway
NOESY	Nuclear Overhauser Effect Spectroscopy
N5-OH-Arg	N^5-hydroxyarginine
NP	Natural Products
NMR	Nuclear Magnetic Resonance
NRPs	Non-Ribosomal Peptides
NRPS	Non-Ribosomal Peptide Synthetase
OSMAC	One Strain, Many Compounds
Ox	Oxidases
PBP	Penicillin Binding Protein
PCD	Programmed Cell Death
Phe	Phenilalanine
PKs	Polyketides
PKSs	Polyketide Synthases
PLSDA	Partial Least Squares-Discriminant Analysis (PLSDA)

QTOF	Quadrupole Time-of-flight
RiPP	Ribosomally Synthesized and Post-translationally modified Peptide
RMN	Resonancia Magnética Nuclear
ROESY	Rotating frame Overhauser Enhancement Spectroscopy
SAR	Structure-Activity Relationship
SDA	Sabouraud Dextrose Agar
TFA	Trifluoroacetic Acid
T1PKS	Multimodular Type I Polyketide Synthase
TCs	Terpene Cyclases
TE	Thioesterase
TNH	1,3,6,8-tetrahydroxynaphthalene
TOCSY	Total Correlation Spectroscopy
TOF	Time-of-flight
TPS	Terpene Synthases
Trp	Triptophan
UHPLC	Ultra High-Pressure Liquid Chromatography
UV	Ultraviolet
Val	Valine
VCPO	Vanadium-dependent Chloroperoxidase
VRE	Vancomycin-resistant *Enterococcus*
VHPO	Vanadium-dependent Haloperoxidases
WHO	World Health Organization

INDEX

Index

Figure Index

Resumen

La incesante aparición de resistencias bacterianas a los antibióticos actuales representa una amenaza significativa para la salud, siendo de vital importancia el desarrollo de nuevos tratamientos eficaces. En este contexto, la diversidad estructural de los productos naturales y sus funciones biológicas optimizadas por la evolución, ofrecen una amplia gama de posibles fármacos con potencial antimicrobiano. Históricamente estas estructuras químicas privilegiadas han jugado un papel inestimable en el descubrimiento de fármacos. Muchos de los medicamentos más importantes, incluyendo antitumorales, inmunosupresores y, en particular, antibióticos, se descubrieron originalmente a partir de fuentes naturales. A pesar de que los metabolitos secundarios muestran una ilimitada diversidad química y biológica, solo conocemos una pequeña fracción del vasto reservorio disponible en la naturaleza, donde continuamente se siguen descubriendo nuevos compuestos. Con el propósito de expandir el espacio químico en torno a los diferentes grupos de productos naturales bioactivos, se ha llevado a cabo en la presente tesis doctoral una serie de estudios destinados al descubrimiento de nuevos metabolitos secundarios a partir de la colección de microorganismos de la Fundación MEDINA. A tal efecto, se han aplicado en estos trabajos i) técnicas cribado de alto rendimiento, ii) desreplicación química de compuestos, iii) aislamiento de productos naturales microbianos, iv) elucidación estructural mediante espectroscopia de RMN y espectrometría de masas de alta resolución (HRMS), así como v) análisis de genomas para identificar microorganismos capaces de biosintetizar nuevos compuestos con características estructurales novedosas y actividad biológica prometedora. Concretamente, esta tesis describe el descubrimiento de: i) cuatro nuevos meroterpenoides de la familia de las napiradiomicinas obtenidos a partir de una cepa de *Streptomyces* de origen marino, incluyendo la napiradiomicina D1, primer miembro de una nuevo subtipo estructural caracterizado por un anillo de 14 miembros que mostró actividad significativa para inhibir el crecimiento de *Staphylococcus aureus* resistente a la meticilina, *Mycobacterium tuberculosis* y la línea celular tumoral humana HepG2; ii) tres nuevos pentapéptidos cíclicos análogos de las pentaminomicinas (familia que incluye miembros con actividad antibacteriana selectiva contra *A. baumannii*) aislados de un *Streptomyces* terrestre, dos de los cuales constituyen los primeros ejemplos reportados en la literatura de péptidos no ribosomales conteniendo un residuo de 2-piridil-alanina; y iii) dos nuevas macrolactonas glicosiladas de 52 miembros, gargantulidas B y C, con una potente acción antibacteriana frente a bacterias Gram-positivas multirresistentes (ej. MRSA) y actividad moderada contra *A. baumannii*, producidas por una cepa del género *Amycolatopsis* y cuya biosíntesis implica a un clúster de genes biosintéticos con un tamaño extraordinario de 216 kbp. Mediante la exploración de fuentes microbianas de hábitats marinos y terrestres, estos hallazgos aportan nuevos conocimientos que ayudan a expandir la diversidad química de los metabolitos secundarios de origen microbiano. Estos resultados ahondan en el vasto potencial de los productos naturales como un pilar prometedor en el desarrollo de nuevos compuestos bioactivos.

Abstract

The relentless emergence of bacterial resistance to current antibiotics poses a significant health treat, making the development of new effective treatments of vital importance. In this context, the unique structural diversity and biologically optimized functions shaped by the evolution of natural products offer a wide range of drug leads with antimicrobial potential. Historically, these privileged chemical structures have played a valuable role in inspiring drug discovery. Many of the world's most important drugs, including antitumoral, immunosuppressive, and particularly antibiotics were originally discovered from natural sources. Despite the unlimited chemical and biological diversity displayed by secondary metabolites, only a small fraction of the vast reservoir has been surveyed, in which novel compounds are being continuously discovered. With the aim of expanding the chemical space around different groups of bioactive natural products, the present doctoral thesis describes a series of research studies to discover new secondary metabolites from the Fundación MEDINA microorganisms' collection. To this end, the following approaches have been employed: i) high-throughput screening (HTS), ii) chemical dereplication, iii) isolation of microbial natural products, iv) structural elucidation through NMR spectroscopy and high-resolution mass spectrometry (HRMS), and v) genome mining to identify talented producers that can biosynthesize new promising bioactive compounds with remarkable structural features. Specifically, this thesis report the discovery of: i) four new meroterpenoids from the napyradiomycin family obtained from a marine-derived *Streptomyces*, including napyradiomycin D1, the first member of a new structural subtype characterized by a 14-membered ring, which exhibited significant bioactivity against methicillin-resistant *Staphylococcus aureus* (MRSA), *Mycobacterium tuberculosis*, and the human tumour cell line HepG2; ii) three new cyclic pentapeptides analogous to pentaminomycins (family of natural products including members with selective antibacterial activity against *A. baumannii*) isolated from a terrestrial *Streptomyces*, two of which were revealed as the first reported examples of non-ribosomal peptides containing a 2-pyridyl-alanine residue and iii) two new 52-membered glycosylated macrolactones, gargantulides B and C, showing potent growth inhibitory activity against multi-resistant Gram-positive bacteria (*e.g.*, MRSA) and moderate activity against *A. baumannii*, produced by an *Amycolatopsis* strain with an extraordinarily large biosynthetic gene cluster of 216 kbp. By exploring microbial sources from both marine and terrestrial habitats, these findings contribute to expanding the chemical diversity of microbial secondary metabolites. These results underscore the vast potential of these compounds as a promising cornerstone for drug development of new bioactive compounds.

INTRODUCTION

Introduction

1. Natural Products - Origin, Importance, and Description

Natural resources have always been essential for the development and maintenance of quality of life. The diversity of products obtained from this source mainly includes foods, oils, carbohydrates, proteins, fibers, and biofuels.[1,2] The significance of natural products as raw materials has led to an increasing chemical profiling of plants, fungi, bacteria, and animals in search of special substances, the so-called specialized or "secondary" metabolites. Methods for the isolation of these compounds serve as the basis for the agrochemical industry, including the production of pesticides and herbicides, the manufacture of cosmetic products, perfumes, and dyes, and development of drugs. Within this context, the chemistry of natural products (CNP) has sought and investigated active substances in terrestrial and marine biota. Historically, CNP has a strong therapeutic appeal, and it comes as no surprise that natural products are the main source of chemical diversity for the discovery of new drugs.[3–5]

Newman and Cragg have released over the past 25 years six review articles (1997, 2003, 2007, 2012, 2016 and 2020) in which they consistently show that approximately 50% of the drugs for clinical use on the pharmaceutical market are of natural origin (Fig. 1).[3,5–9] In general, these compounds are found as purified, derived, analogous, or otherwise obtained through processes of semi-synthesis. Ultimately, metabolites found in natural sources have served as inspiration for the synthesis of new chemical entities.[3,6,7,10] In the case of drugs used in the treatment of cancer or infectious diseases this percentage is even higher, with about two-thirds of these agents having a natural origin. The fascinating chemical diversity and biological complexity of this bioactive specialized metabolites can be exemplified with (1) antibiotic (e.g., tetracyclines, vancomycin, clindamycin), (2) antiparasitic (e.g., berberine, avermectin), (3) antimalarial (e.g., artemisinin, artesunate, quinine), (4) lipid control agents (e.g., statin analogues), (5) immunosuppressant use for organ transplantation (e.g., rapamycin, cyclosporine), and (6) anticancer drugs (e.g., vincristine, irinotecan, paclitaxel, doxorubicin).[3,11,12] All these drugs resulted in a significant rise in life expectancy and life quality.[13]
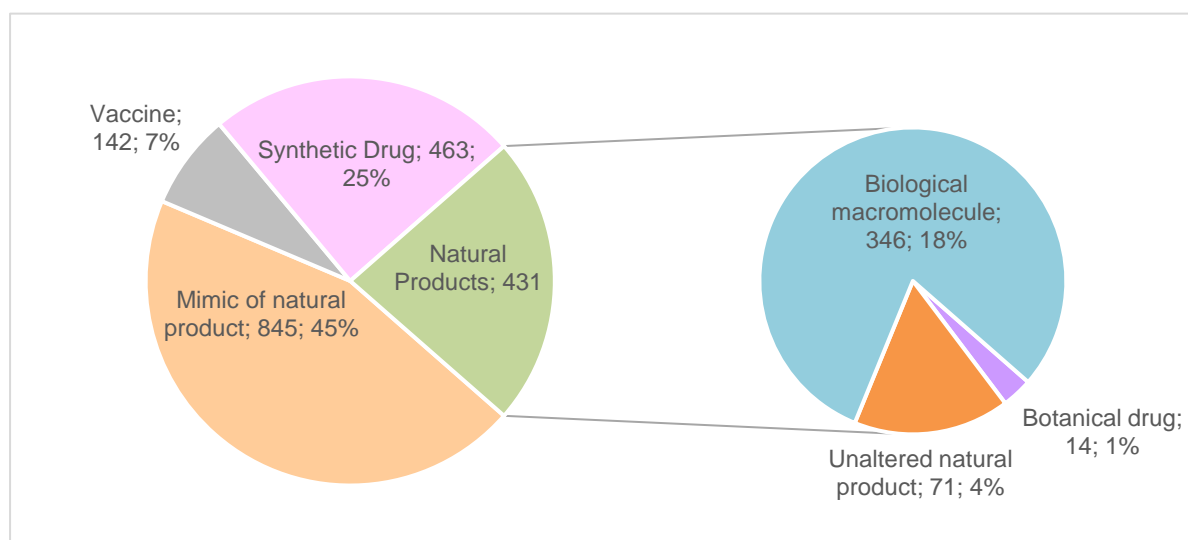
Figure 1. All newly approved drugs from January/1981 to Sep/ 2019 (n= 1881), data from Newman and Cragg, 2020.[3] Mimic of natural products include synthetic drugs with NP pharmacophore, synthetic drugs that mimic NP and NP derivatives.

The term natural products (NP) broadly refers to chemical compounds derived from a myriad of living sources including plants, animals, marine organisms, and microorganisms.[14,15] In particular, microbial natural products are one of the richest and most diverse sources of bioactive compounds with diverse chemical structures and a wide range of bioactivity.[16] For instance, bacterial metabolism involves the biosynthesis of primary and secondary metabolites that differ in their respective roles and functions (Table 1). Primary metabolites are usually biosynthesized during the growth phase and

execute physiological duties in the organism. These compounds include amino acids, organic acids, and nucleosides and play a crucial role in central metabolism, microbial growth, and proliferation. On the other hand, secondary or specialized metabolites are mainly extracellular compounds generated near or after the stationary growth phase that guide microbial interaction with their surroundings. As a result, these molecules serve an important part in ecological activities such as defense mechanisms, signaling, protection, and pigment production.[17]

Primary metabolism is characterized by its specificity, producing a limited number of well-known end-products that have been chemically and biologically studied in the last decades.[18,19] The opposite scenario is seen for secondary metabolism, with gaps of uncertainty in the understanding of its production and biochemical roles.[20,21] Often, secondary metabolism is promiscuous, involving enzymes with low substrate specificity and bottleneck enzymes that produce slightly differing components of a particular chemical family.[22]

	Primary metabolites	Secondary metabolites
Definition	Metabolites required for basic cellular processes	Metabolites not directly required for basic cellular processes
Function	Essential for growth, development, and reproduction	Specialized roles such as defense, attraction, or communication
Biosynthesis	Produced via well-defined metabolic pathways	Produced via complex and often unique biosynthetic pathways
Pathway organization	Intermediates are interconnected, allowing for tracer studies and pathway elucidation	Pathways for each secondary metabolite are mostly unrelated to one another and often without diffusible intermediates
Regulation	Production regulated by feedback inhibition and other metabolic control mechanisms	Production is regulated by numerous factors such as environmental cues and genetic regulation
Evolutionary conservation	Generally conserved across all organisms	Often specific to certain species or groups of organisms
Expression	Genes and pathways active and expressed at high concentrations	Biosynthetic genes are often silent or cryptic; often not expressed under standard physiological conditions
End products	Convergent: few known end-products	Divergent: many unknown end-products
Molecular properties	Charged to prevent crossing the cell membranes	Lowly charged and non-polar, facilitating diffusion through cell membranes; molecular weight ranging from ~100 to 1500 Daltons (typically, low molecular weight compounds)
Examples	Organic acids, amino acids, nucleosides	Polyketides, terpenoids, non-ribosomal peptides
Concentration	Typically present in high concentrations	Often present in low concentrations
Location	Found inside the cell	Predominantly extracellular
Challenges for study	All major metabolic pathways are active, making them easier to study	Developing methods to activate silent biosynthetic gene clusters and identifying new natural products

Table 1. Comparison of Primary and Secondary Metabolites: Characteristics and Function.

Although thousands of specialized metabolites are produced by microorganisms, they typically derive from only a few essential building blocks, with pathways that contain a limited number of reactions branching off from primary metabolism. The most important precursors are acetyl coenzyme A and propionyl-CoA, together with tricarboxylic acid, shikimic acid, and amino acids pathways.[23] These cycles can be combined in thousands of different ways, giving rise to a wide variety of chemical structures

including polyketides, terpenes, non-ribosomal peptides, steroids, and fatty acid derivatives. Furthermore, unlike primary metabolism, most biosynthetic pathways encoding secondary metabolites remain inactive or "silent" when exposed to standard laboratory growth conditions,[24,25] thereby posing a significant challenge in the study of these compounds. Even in highly productive or genetically talented bacteria, the full characterization of many of these pathways has yet to be achieved and well-known bacterial strains are still being reported as the producers of novel chemical structures.[26] Many pathways, like those for non-ribosomal peptides or polyketides, lack diffusible intermediates due to their assembly-line mechanism.[27] Thus, the main obstacles to studying secondary metabolism are its silent nature and the capacity to investigate its varied natural products chemically and biologically.

2.   Historical use of Natural Products for human health

Throughout history, nature has always been the primary source of drugs for human health. Traditional medicine relies heavily on natural products, with about 80% of the world's countries depending on them for healthcare according to the World Health Organization (WHO).[28] The earliest known records of natural products as therapeutic agents can be traced back to ancient Mesopotamia, where plant-based medicines were prevalent.[29] Plants-derived extracts and ointments have been extensively documented in drug encyclopedias worldwide as effective remedies for the treatment of ailments.[4] The heritage of this past wisdom and experimentation, inbred in our society, has been the basis of evidence-based modern medicine and remains a significant source of inspiration for drug discovery.[30]

The identification of therapeutic properties from plant sources led to the isolation of the first natural products-based active compounds in the 1800s such as morphine (1806), quinine (1820), caffeine (1821), atropine (1831), ephedrine (1885) and acetylsalicylic acid (1899).[31,32] These milestones paved the way for the essential structural characterization of the growing variety of chemical scaffolds. Within this context, different physical techniques including mass spectrometry (MS), nuclear magnetic resonance (NMR) spectroscopy, and X-ray crystallography saw significant advancements and greatly enhanced the field of analytical chemistry in the 1940s.[33] Today, these techniques have become essential tools for the discovery of new natural products, allowing the identification of their structures and the understanding of the mechanisms of action that underpin their biological activity.

The chemistry of natural products has since been extensively investigated in both terrestrial and marine biotas to identify active substances that may address unmet clinical needs. Its flourishing as a chemo-pharmaceutical frontier science occurred during World War II, in which the demand for antibiotics boosted research programs within pharmaceutical companies with new metabolic screening methodologies. The discovery of penicillin as a result of those efforts is undoubtedly on of the most important scientific discoveries in our recent history. This antibiotic, isolated in 1928 and structurally elucidated in 1945,[34] marked a shift towards using microorganisms as a source of natural products.[35] Afterwards, the systematic screening of the first actinomycetes obtained from the soil in the 1940s[36] and the subsequent discovery of other invaluable antibiotics such as actinomycin,[37] streptothricins,[37] and most notably, streptomycin,[38] greatly expanded the arsenal of life-saving compounds. During the "*golden age of antibiotics*" (1950-1960) chemistry and microbiology operated in tandem to extract secondary metabolites from cultivated bacteria. This process enabled the production of a variety of novel and effective drug targets, opening a range of possibilities for treatments and pharmaceuticals that would not have been possible without their synergy.

The success of NP discovery in the 40s, 50s, and 60s prompted pharmaceutical companies to intensively investigate nature as a potential source of new drugs, providing the development of innovative methods for isolation and structure elucidation of NPs. During this golden age of natural products (1930-1970), the continuous arsenal of new antibiotics was so massive that it was assumed that infectious diseases would be controlled and annihilated by the end of the 20th century.[39]

Between the 1970s and early 2000s, the discovery of novel antibiotics stalled. The rediscovery of previously known compounds was considered a major drawback in the drug development field, together with increasing legal issues on intellectual property, defense of biodiversity (fauna and flora heritage), and subsequent royalties for traditional communities in which possible bioactive compounds were

extracted. The pharmaceutical industry slowly shifted to alternative drug discovery methods like automated high-throughput screening technology (HTS) toward synthetic chemical libraries[40] and combinatorial chemistry (CC), which were believed to result in faster and cheaper drug development. Complementarily, the chemical profile of natural sources widely varied according to seasonality and environmental variations, decreasing reproducibility and proper supply.

CC particularly promised to fulfill this demand by creating an immensity of compounds for testing in HTS, being highly reproducible and of low complexity. CC would be fully transposable to an industrial scale, in case of approval by the Food and Drugs Administration (FDA) or other regulatory bodies. From the political-economic point of view, CC would eliminate all the diplomatic and bureaucratic issues involved in sampling and extraction from foreign natural resources. In summary, the cost of investment in discovery and subsequent operational development would be much lower and less bureaucratic than the use of the NP.[41,42] With such expectations on the horizon, many "big pharma" companies including GlaxoSmithKline, Hoffman-La Roche, Pfizer, Bristol Myers Squibb, Sanofi-Aventis, Schering-Plough and Eli Lilly embraced the CC approach and limited their activities in NP discovery.[41,42] Notable exceptions such as Novartis, Merck, and Bayer maintained their NP-based drug discovery programs.

Even though CC provided a significant number of compounds, the results fell short of expectations and, until mid- in the 1990s there were doubts about the usefulness of this methodology in the discovery of lead compounds. Most companies applied the same, simple synthetic steps, leading to similar core structures. This redundancy created libraries with poor structural novelty, low coverage of the chemical space, and the predominance of molecules that were unable to cross the cell membrane due to their lipophilicity, decreasing the discovery of novel chemical entities with medicinal properties.[11] This failure, alongside increased barriers set for natural products research, resulted in a collapse of the drug discovery and development pipelines,[43] which particularly affected the anti-infective field due to the lack of new antibiotic scaffolds in a context of growing antimicrobial resistance.

Currently, antimicrobial resistance (AMR) is the second leading cause of death, with 6.22 million deaths associated with or directly attributable to this phenomenon in 2019.[44] The WHO predicts that by 2050 this major public health threat could surpass cancer as the primary cause of death by causing 10 million deaths annually.[45] In Spain, antimicrobial resistance is responsible for more than 35,000 deaths, and causes four million serious infections per year, according to the Spanish Society of Infectious Diseases and Clinical Microbiology.[46]

Nowadays, several factors continue to contribute to the emergence of antibiotic resistance. These include inappropriate prescription and indiscriminate use of antibiotics in clinical and animal production, self-medication, poor hygiene and sanitation, climate change, and poor surveillance of their use and prescription. Consequently, these factors have propelled the development of multidrug resistance among pathogens, turning bacterial infections into a global public health concern.[47,48] The phenomenon of AMR affects all classes of antibiotics, and the emergence of cross-resistances within the same antibiotic classes is increasing rapidly.[49] Research efforts were primarily focused on combating pathogenic Gram-positive bacteria, further exacerbating the shortage of effective treatment options available against multidrug-resistant Gram-negative bacteria. This is particularly concerning as these types of infections, often acquired in hospitalized patients and associated with high mortality rates, have reached the pandemic scale.[49]

The COVID-19 pandemic has also resulted in a surge in the use of antibiotics, with approximately 70% of hospitalized patients receiving them.[50] There is evidence that bacterial co-infections, caused by nosocomial drug-resistant microorganisms such as *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, or *Acinetobacter baumannii*, have played a crucial role in the severity and survival of patients with severe COVID-19.[51] Due to the urgency of the situation, antibiotics are often administered without a confirmed microbiological diagnosis. This indiscriminate use of antibiotics has led to their depletion, and as a result, there has been a rise in the use of broad-spectrum antibiotics as alternatives. This practice may potentially increase resistance to some antibiotics, further aggravating the problem.[52,53] In this sense, the COVID-19 pandemic has highlighted even more the importance of developing treatments and vaccines with appropriate urgency to combat global health threats like AMR.

In this context of such a global health threat, it is crucial to prioritize research and development of new antibiotics and implement strategies to prevent the spread of resistant bacteria. Without new and effective antibiotics, the world faces a bleak future in which many infections that were once treatable could become deadly once again. While combinatorial synthesis produces molecules that obey Lipinski's rule of five in terms of size and property, the structural diversity of NPs, especially concerning the variety of scaffolds and chiral centers, may explain why approximately 50% of drugs introduced to the market in the last decades are directly or indirectly derived from NPs.[3] Therefore, NPs continue to serve as a rich and abundant reservoir of potential lead compounds for medicinal chemistry, emphasizing the importance of integrating this effort into global drug discovery programs. Recent progress in microbial genomics, new culturing techniques, natural product biosynthesis, bioinformatics, and analytical technologies are just some examples of rational screening of natural products.[54] This systemic and holistic view is revitalizing natural product-based drug discovery, making them an unbeatable choice for developing new drugs due to their inherent potentiality.[55]

3. Microbial Sources of Secondary Metabolites– roles of Actinomycetes in Drug Discovery

Microbial secondary metabolites, resulted from several million years of evolutionary biosynthetic optimization, have tremendous potential to provide new therapeutic agents.[56] In a recent review, Newman and Cragg have shown that, even with the challenges related to unlocked genomes and unculturable strains, microbial bioactive compounds are the future for drug discovery programs,[5] accounting for over 58 thousand natural compounds already reported from different microbial families.[57]

The potential of these microbial secondary metabolites comes not only from extremely diverse chemical structures but also from an extensive and still little-studied microbial population. Indeed, recent whole-genome sequencing of several fungi and bacteria shows that the potential of microbial biosynthetic machinery to produce secondary metabolites is fairly underestimated, meaning that a much broader range of compounds could be produced if the silent genes are induced by different methods.[58–62] These organisms can regulate and maintain chemical diversity at a low energy cost, which means that a variety of secondary metabolites can be produced by one single strain when environmentally needed. Thus, in competitive and stressful environments, microbes engage in constant interactions with the host and other organisms (neighbors and competitors), resulting in biological and ecological effects that shape the community and ensure survival.[63]

Apart from their ecological roles, many of these specialized metabolites have been shown to display a wide range of biological activities, including antibacterial (e.g., erythromycin, tetracyclines) immunosuppressants, (e.g., cyclosporine, tacrolimus), antifungal (amphotericin B, candicidin) antiviral (acyclovir, actinomycins), anticancer (doxorubicin, anthraquinones) and antiparasitic (avermectin, prodiginine).[3] In fact, approximately three-quarters of all known anti-infective drugs used in the clinic are derived from microbial sources, predominantly actinomycetes (Fig. 2).[64]

Actinomycetes are a highly heterogeneous group of aerobic, Gram-positive filamentous bacteria belonging to the Phylum *Actinobacteria*.[65] They are saprophytic microorganisms, (*i.e.*, they feed on decaying organic matter), being able to colonize a wide range of terrestrial, marine, and extreme ecosystems. As a result, they are diversely distributed in sediments, soils, and mainly in the rhizospheres of plants, playing an important ecological role in nutrient recycling.[64] They also possess a unique biological cycle within the bacterial kingdom, sharing similarities with eukaryotic fungi. Their complex life cycle encompasses shifts between reproductive aerial hyphae and vegetative mycelial forms. These hyphae eventually differentiate into conglomerated single-celled spore packets, which are the organism's primary means of reproduction. During this cycle, morphological development and secondary metabolism often take place at the same time.[66]
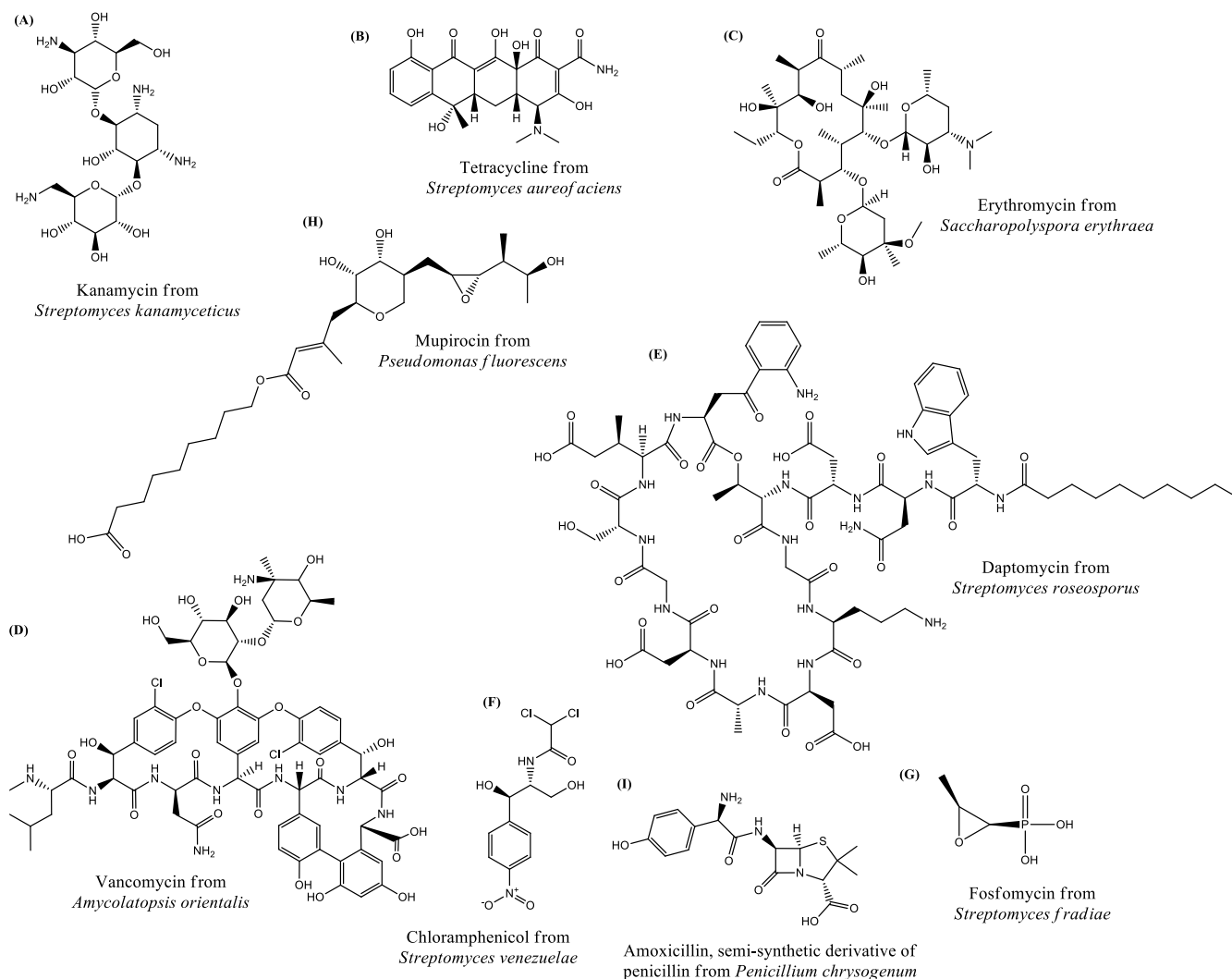
Figure 2. Drugs from microbial sources. Actinomycetes are known to produce several scaffolds with antibiotic activity, such as (A) aminoglycosides; (B) tetracyclines; (C) macrolides; (D) glycopeptides; (E) lipopetides; (F) amphenicols; (G) phosphonates. Additional significant classes of antibiotics derived from other bacterial sources and fungi include, among many others, (H) mupirocin or (I) penicillins.

Actinomycetes are known as "*the richest microbial source of bioactive specialized metabolites*". This title comes from their reported ability to activate numerous and diverse Biosynthetic Gene Clusters (BGCs) according to the environmental conditions, assembling a wide variety of secondary metabolites.[67] Indeed, genome sequencing of numerous Actinomycetes strains revealed that their genome size is significantly greater than what is required for fundamental operations. Remarkably, many actinomycetes harbor over 20 BCGs encoding the biosynthesis of secondary metabolites, which means they dedicate over 5% of their coding capacity to these processes.[67]

Among the Actinomycetes family, *Streptomyces* stands out as the most studied genus, with more than 800 species identified to date,[68] and has proven to be a prolific and impactful origin of invaluable chemical substances.[69] While a high guanine-cytosine (GC) content in their DNA is a common feature within the Actinomycetes family (>70%), *Streptomyces* excels with its notably large and linear genomes (8-10 Mb) typically containing 20 or more gene clusters.[70] They are mainly soil-dwelling organisms, which are characterized by the production of geosmin, a volatile metabolite that gives them their "earthy" odor. Spore germination marks the beginning of a complex life cycle for the multicellular mycelial *Streptomyces*.[71]

The spore germination process directly depends on the environmental conditions and correlates with the production of antibiotics and other secondary metabolites due to the stress induced by

undergoing programmed cell death (PCD).[71] Under physiological conditions, *Streptomyces* spore germination commences with the arthrospore, otherwise known as substrate mycelium. However, when confronted with stressful situations (such as scarcity of nutrients or microbial competition), growth initiates from the aerial mycelium instead. In simpler terms, when conditions are favorable, fully developed mycelia are generated, but under challenging conditions, the aerial mycelium undergoes subdivision through septa formation, eventually giving rise to spores (Fig. 3).[66]
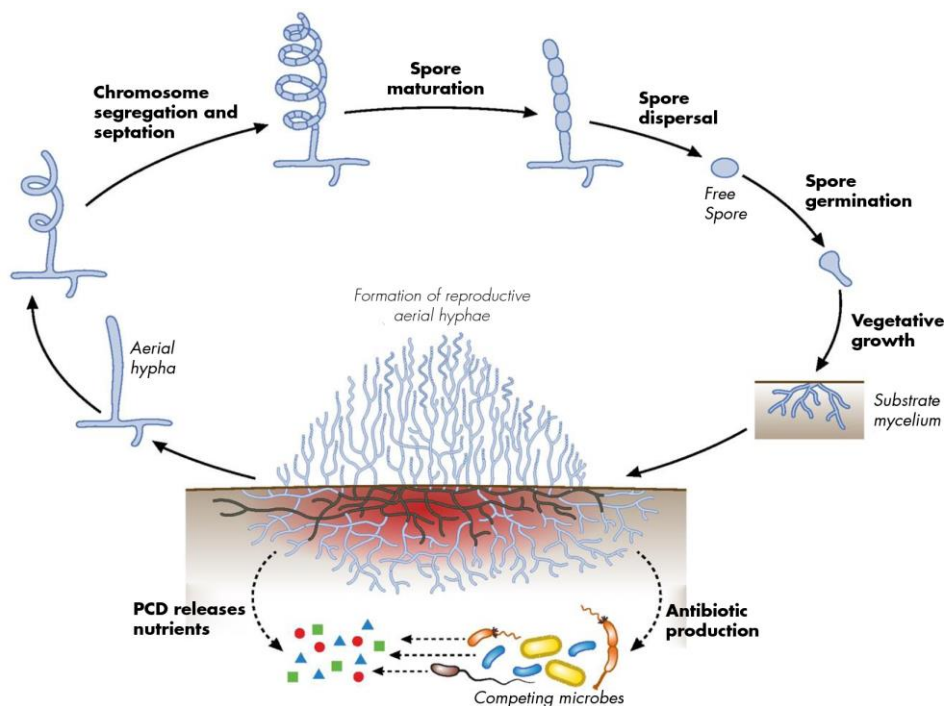

Figure 3. The developmental life cycle of *Streptomyces*. Figure from van der Meij et al., 2017.

Numerous secondary metabolites that exhibit antibiotic properties are synthesized by different strains of *Streptomyces*. These antibiotics are typically classified based on their mode of action, chemical structure, or range of effectiveness. For example, relevant aminoglycosides, such as neomycin,[72] streptomycin,[38] and kanamycin[73] are produced by *Streptomyces griseus, Streptomyces fradiae,* and S*treptomyces kanamyceticus*, respectively. Other antibacterial NPs derived from *Streptomyces* include tetracycline, produced by *S. rimosus* and *S. aureofaciens*; chloramphenicol[74] produced by *S. venezuelae*; or thienamycin[75] produced by *S. cattleya*.

Although most antibiotics of microbial origin are streptomyces-derived and many *Streptomyces* strains produce bioactive compounds, only a limited number have reached clinical use.[69] As a result, in the last decades the search for new NP sources has been broadened beyond soil towards underexplored habitats and minor taxa. These include non-*Streptomyces* species, also known as minor actinomycetes and novel *Streptomyces* found in marine or extreme environments.[76] In general, minor actinomycetes have a slower growth rate and smaller population and demand specific physiological conditions for growth. These features pose a greater challenge in the isolation of minor actinomycetes from soil or other substrates.[77]

While compounds from minor actinomycetes have been obtained less frequently, they have already yielded novel antimicrobials with unique chemical structures. Over the past twenty years, the percentage of known antibiotics that are produced by minor actinomycetes has risen to approximately 25-30%, primarily from the Micromonosporaceae family, with minor contributions originating from the Thermomonosporaceae and Pseudonocardiaceae families.[78] Some notable examples include erythromycin from *Sacchapolyspora erythraea*,[79] teicoplanin from *Actinoplanes teichomyceticus*,[80] fidaxomicin from *Dactylosporangium aurantiacum*,[81] gentamycin from *Micromonospora purpurea*,[82] dalbavancin from *Nonomuraea* sp.[83] or the potent lantibiotic microbisporicin, from *Microbispora*

corallina.[84] Recently, we have discovered in our group a novel family of unprecedented pyrazine-alkylguanidine metabolites, crossiellidines A–F, being the first specialized metabolites reported from the minor actinomycete genus *Crossiella.* Although not detailed in this thesis, these compounds exhibited significant broad-spectrum antibacterial activity, including clinically relevant ESKAPE pathogens such as *Klebsiella penumoniae*.[85]

Among these minor actinomycetes with the potential to produce antibiotic drug targets, the genus *Amycolatopsis* holds significant research value. This genus consists of 100 reported species and 5 subspecies, which form a unique lineage in the evolutionary tree of Pseudonocardiaceae.[68] Within the wide range of compounds produced for this genus, the most notable antibiotics are vancomycin and rifamycin. Vancomycin was first extracted from *Amycolatopsis orientalis* in the early 1950s and has been sparsely used in the clinic since 1958.[86] Similarly, rifamycin represents the first group of antibiotics targeting RNA polymerase and was isolated from *Amycolatopsis mediterranei* in 1959.[87] Other secondary metabolites from *Amycolatopsis* include polyphenols (kigamicins,[88] pradimicin-IRD),[89] linear polyketides (ECO-0501,[90] vancoresmycin),[91] macrolides (amycolatopsins)[92] macrolactams (rifamycinosides,[93] rifamorpholines),[94] thiazolyl peptides (thioamycolamides,[95] thiazomycins),[96] cyclic peptides (pargamicins,[97] valgamicins),[98] glycopeptides (chloroorienticins,[99] eremomycins),[100] amide derivatives (dipyrimicins, echinosporins),[101] glycoside derivatives (tigloside,[102] actinotetraoses),[103] enediyne derivatives (amycolamycins)[104] and sesquiterpenes (abscisic acid type sesquiterpenes).[105] In the last decade, more than 150 compounds have been isolated from *Amycolatopsis* species collected from soil, marine environments, lichen or plants.[106] Among those NPs, 41 compounds exhibited antimicrobial activities, thus highlighting the great potential of this genus as a source of new antibiotics.

Marine actinomycetes have also gained attention as a promising source of novel bioactive secondary metabolites due to the unique environmental conditions in which they survive.[107] These microorganisms cover all marine ecosystems, flourishing on different ranges of light, temperature, pressure, oxygen, nutrients, and salinity. Given that over 70% of Earth's surface is covered by water (95% of the earth's biosphere), sea/ocean habitats are major ecosystems where novel chemical structures absent in the terrestrial environment can be found.[108] Indeed, the success of marine natural products in drug development is now well established, starting with cytarabine isolated from *Cryptotheca crypta*, which was the first marine-derived natural product to become a successful drug for leukemia and lymphoma in 1969.[109,110]

Technological advancements in the collection of deep-sea samples and large-scale production via aquaculture and extraction have also contributed significantly to the growth of this field.[107,111] There are currently more than ten marine natural products-derived drugs approved for treating different types of cancer.[112,113] These drugs include among others (1) the depsipeptide plitidepsin,[114] extracted from the Mediterranean tunicate *Aplidium albicans*, which shows high effectiveness against various types of myelomas, (2) the macrolide polyketide eribulin mesylate,[114] active against metastatic breast cancer and first obtained from the discovery of halichondrin B, isolated from *Halichondria okadai* and (3) the alkaloid trabectedin[115,116] obtained from the ascidian *Ecteinascidia turbinata* and used in the treatment of soft tissue sarcomas and ovarian cancer.

Several other marine specialized metabolites have progressed from early clinical evaluation and are currently undergoing phase III clinical trials for various oncological indications.[117] A key member of this group is Salinosporamide A (also known as NPI-0052 or marizomib) isolated from *Salinispora* genus, a salt-water marine actinomycete. This compound has just started phase III clinical trials in 2020 for the treatment of glioblastoma.[118,119] Lastly, the small niche of marine minor actinomycetes is also a potential source for a wide spectrum of antimicrobials. In a continuously growing trend, a total of 32 antibiotic compounds from marine-sourced bacteria were reported in 2021.[108] Notably, anthracimycin, first isolated from a species of marine *Streptomyces*,[120] showed excellent *in vitro* antibiotic activity against many *Bacillus anthracis* (anthrax), methicillin-resistant *Staphylococcus aureus* (MRSA), and *Mycobacterium tuberculosis,* being currently in pre-clinical trials.[121]

Great progress has been made in the research of secondary metabolites from bacteria by the use of specialized media and metagenomics. Until recently, microbiology research faced constraints in their study of natural microbial ecosystems as a result of their incapacity to cultivate the majority of microorganisms. Current estimates suggest that < 1% of bacterial species have been cultured so far, strongly suggesting the existence of countless undiscovered microbial diversity.[122] Moreover, even among the species that are already identified, their potential for bioactivity or natural product discovery has only been partially explored. The "genomic revolution" has uncovered that actinobacteria possess a greater abundance of biosynthetic gene clusters within their genetic makeup compared to the secondary metabolites typically observed in standard laboratory conditions.[123] This finding illustrates how underestimated is their potential to produce novel compounds. Lastly, the full potential of actinomycetes for the production of novel natural products is still underexplored.[124] To fully leverage this untapped potential, it is essential to combine different strategies available in microbiology, molecular biology, and analytical chemistry.

4. Analytical and Computational Techniques in Drug Discovery: identification and structure elucidation of NPs

The success of NPs research in drug discovery is partially owed to the convergence of various technological advances. These include improvements in HTS analysis encompassing both detection and separation techniques, computational breakthroughs featuring powerful processors and memory storage capable of handling massive volumes of data, and the applicability of pattern recognition and statistical methods. Initially crafted for economic applications, these statistical methods were subsequently tailored to facilitate the acquisition, analysis, and interpretation of biological datasets.[125–128]

Specifically, in natural products research, the use of computational tools has acquired the utmost significance to attain precise and detailed chemical information, aiming to alleviate the complexity of crude extracts and streamline the process of identifying promising strains or potential drugs. Within these strategies, chemical dereplication is a swift method for the early-stage identification of already known compounds in crude extracts, thereby reducing the ratio of redundant isolation of secondary metabolites that have already been characterized. Compared to univariate approaches such as differential analysis (DA) and partial least squares-discriminant analysis (PLSDA), dereplication can be regarded as a rapid and cost-effective alternative. It offers an improved and more accurate procedure for the prioritization of bioactive samples.[129,130]

Chemical dereplication relies on the comparison of experimental data with those contained in comprehensive databases of spectrometric, spectroscopic, and physicochemical features -either experimental or calculated- for previously isolated compounds. Some examples of relevant databases for NPs dereplication are summarized in Table 2 and discussed below.

NPs Data Base	cmpds/ structures	NPs	Microbial NPs	Mass Search	MS information	NMR peak data	Structural features	Structure / substructure / similarity	UV	Taxonomy	Bioactivity
a)[131] SciFinder https://scifinder.cas.org	279 x 10^6	N.D.	N.D.	Y MF	N	Y	Y^σ	Y	N	Y	Y
b)[132] PubChem https://pubchem.ncbi.nlm.nih.gov/	115 x 10^6	N.D.	N.D.	Y MF	N	N	N	Y	N	N	Y
c)[133] ChemSpider https://www.chemspider.com/	100 x 10^6	> 75000	N.D.	Y MF	N	N	N	Y	N	N	N
d)[134] Reaxys https://www.reaxys.com/	110 x 10^6	310305	N.D.	Y MF	N	N	N	Y	Y	Y	Y
e) Dictionary of Natural Products CHEMnetBASE https://dnp.chemnetbase.com	N.D.	>340000	40575 bacteria + fungi	Y	N	N	N	Y	Y	Y	Y
f)[57] ANTIBASE The Natural Compound Identifier https://sciencesolutions.wiley.com/solutions/technique/screening/wiley-identifier-of-natural-products/	N.D.	> 58000	ca. 43000	Y MF	Y HRMS	Y	Y	Y	Y	Y	Y
g)[135] npatlas https://www.npatlas.org	33375	33375	33375 bacteria + fungi	Y EM	N	N	N	Y	N	Y	N
h)[136] GNPS http://gnps.ucsd.edu	$^\pi$ca. 160000	ca. 122000	N.D.	N	Y HRMS, MS^2	N	N	N	N	N	N
i) MarinLit https://marinlit.rsc.org/	N.D.	40543 marine NPs	N.D.	Y EM	Y HRMS	Y	Y NMR ft.	Y	Y	Y	Y
j)[137] ISDB https://oolonek.github.io/ISDB/	N.D.	170602 (ISDB)	N.D.	N	Y Pr. MS^2	N	Y^δ	N	N	N	Y
k)[138] StreptomeDB http://www.pharmbioinf.uni-freiburg.de/streptomedb	6524	6524	6524	Y RMF	Y HRMS	Y	Y^σ	Y	N	Y	Y
l)[139] COCONUT https://coconut.naturalproducts.net/	407270	407270	N.D.	Y MF	N	N	Y No NMR ft.	Y	N	Y	N
m)[140] LOTUS https://lotus.naturalproducts.net/	276518	276518	N.D.	Y MF	N	N	Y^σ	Y	N	N	N

NPs Data Base	cmpds/ structures	NPs	Microbial NPs	Mass Search	MS information	NMR peak data	Structural features	Structure / substructure / similarity	UV	Taxonomy	Bioactivity
n)[141] NPASS Natural Product Activity & Species Source Database https://bidd.group/NPASS/	N.D.	94413	3783 bacteria + fungi	N	N	N	Y^{σ}	Y	N	Y	Y
o)[142] NP-MRD https://np-mrd.org/	280010	40102	N.D.	N	N	Y	N	Y	N	N	N
p)[143] CMNPD https://www.cmnpd.org/	31561	31561	7578 marine bacteria + fungi	Y MF, MR	N	N	N	Y	N	Y	Y
q)[144] CyanoMetDB secondary metabolites https://zenodo.org/record/7922070	2605	2605	2605 cyano-bacteria	N	N	N	N	N	N	N	N
t)[145] MiMeDB https://mimedb.org/	24254	16018	14494	Y MW	Y HRMS, MS^2, EI-MS	Y	N	Y	N	Y	Y
u) NP structural features MEDINA In-house microbial natural products library	N.D.	ca. 423000	N.D.	Y MF, RMF, MR	Y HRMS	N	Y NMR ft.	Y	N	N	N

Table 2. Relevant Databases for Natural Product Dereplication. Public domain databases: b), c), g) h), j), k), l), m), n), o), p), q), and t). $^{\delta}$ISDB uses Database for Rapid Dereplication of Known Natural Products (DEREP-NP).[146] $^{\pi}$MassBank and HMDB (Human Metabolome Database) excluded. $^{\sigma}$NMR data is limited to a few general molecular descriptors atom types, and rotatable bonds. Abbreviations: Y= Yes, N= No, MR= Mass Range, RMF= Range of Molecular Formula, EM = Exact or accurate Mass, MF = Molecular Formula, MW = Molecular Weight, HRMS = High-Resolution Mass, EI-MS = Electron Impact Mass data, MS2 = MS/MS, Pr. MS^2 = Predicted MS/MS, N.D.= Not Described, NMR ft. = NMR features.

Typically, the acquisition of these experimental data involves liquid chromatography hyphenated to high-resolution mass spectrometry (LC-UV-HRMS), -often coupled in addition to diode array (DAD) detection (LC-UV-HRMS)-, and/ or Nuclear Magnetic Resonance (NMR) experiments. The combination of these techniques is a powerful approach to identify known compounds. Dereplication by mass spectrometry data is based mainly on accurate or exact mass, molecular formula (when it can be reliably deduced), and isotopic patterns. Also, the great development during the last years of valuable tools such as the Global Natural Products Social Molecular Networking (GNPS),[136] has expanded the application of tandem mass spectrometry (MS/MS) to chemical dereplication, as it allows to compare experimental fragmentation spectra with those contained in chemically annotated repositories.

The application of NMR in dereplication has revolutionized the field during the last decade due to improvement in sensitivity brought about by the development of new probe technologies (e.g., MicroCryoProbe™).[147] When using NMR for dereplication purposes, experimental spectra or chemical shift lists are searched in libraries of experimental or calculated spectra (or shifts). Another very useful approach is to search for structural features deduced from experimental NMR spectra (usually ^{1}H and HSQC) in databases of compounds that allow such queries (see Table 2). These methods benefit greatly from the deposition of raw NMR data into public natural product databases like NP-MRD,[142]

which offer more standardized and detailed information than just a list of values, images, or tables. This access to FAIR (Findable, Accessible, Interoperable, Reusable) data significantly streamlines the process of compound dereplication with NMR, enhancing both accuracy and efficiency.[148]

Other valuable approaches to be used in downstream dereplication are substructure-based or similarity-based searches when the structure of the compound has been partially determined. This allows the compound to be identified as a known compound or as an analogue compound belonging to a previously described NP family. Among the different databases, -either public or private-, useful for NP dereplication (see Table 2), those focused or enriched in microbial natural products such as the Natural Products Atlas,[135] the Dictionary of Natural Products, the Wiley Identifier of Natural Products (formerly Antibase)[57] or StreptomeDB,[138] are particularly relevant in our case. Other specialized databases are for example those focusing on marine NPs, such as MarinLit or the Comprehensive Marine Natural Products Database[143] (CMNPD). Some of these databases allow queries regarding accurate mass or molecular formula, NMR chemical shifts, or structural features (molecular descriptors of varying complexity), while others contain valuable search metadata such as taxonomy of the producing strain or biological activity.

At the chemistry department of MEDINA, we have developed a proprietary tool named "NP structural features" comprising different public databases of natural products (NP Atlas,[135] UNPD, StreptomeDB,[138] CMNPD,[143] NPASS,[141,] and others) and a total number of ca. 423000 compounds. This tool allows us to dereplicate NPs by molecular formula (or a range of), mass range, NMR structural features, and searches based on substructure or similarity.

Natural product drug discovery has also been greatly aided by the development of improved chromatographic separation approaches. Using stable and chemically non-reactive isolation and purification techniques, minor compounds present in extracts can be isolated on an unprecedented scale. The use of chromatographic techniques such as MPLC (Medium Pressure Liquid Chromatography), preparative and semi-preparative LC systems, UHPLC (Ultra High-Pressure Liquid Chromatography), ion exchange chromatography, and gel filtration have helped to solve numerous challenging aspects in natural products purification. These challenges include dealing with complex mixtures of secondary metabolites (addressed through size exclusion and polarity-based methods) and managing ionic and non-ionic interactions.

Apart from dereplication, isolation, and purification challenges, another major hurdle in the process of NPs drug discovery is the structure determination of new molecular entities. This field has been significantly transformed by advances in spectroscopic techniques, with ultra-high-resolution NMR technologies playing a key role.[149] Among the numerous advances in this field, the inception of two-dimensional NMR methods and improvements in sensitivity are particularly important for the determination of natural product structures. Multidimensional pulse methods provide an assessment of scalar ^{1}H-^{1}H bonds, correlations of ^{1}H with other atoms (^{13}C, ^{15}N, ^{31}P), and ^{1}H-^{1}H space intramolecular connectivity data that can map out the compound's structure.[150,151] In terms of sensitivity, superconducting magnets, cryogenic electronics, and micro-probe technologies have provided stronger magnetic fields that require less amount of material for structural analysis, operating even in the nanogram range.[152,153] The combination of cryogenic probe electronics with correlation spectroscopy techniques empowers the advancement of highly potent experiments that were previously unavailable using conventional hardware, including correlation targeting low-abundance ^{13}C and ^{15}N nuclei.[154] The integration of HRMS and 2D NMR has sped up structure elucidation, significantly reducing the amount of material required for analysis.

Recent trends in the determination of the stereochemistry of natural products involve the integration of conformational analysis and NMR experimental data. Methods such as DP4+,[155] J-DP4,[156] ML-J-DP4,[157,] and MESSI[158] enable increasingly faster stereochemical proposals for small molecules or molecular fragments with reduced computational resources. By leveraging statistical inference and machine learning approaches, they offer improved accuracy over traditional methods. These tools represent a significant leap forward in natural product chemistry, allowing for more efficient structure elucidation.

In recent years, the synergistic combination of genomic data with NMR spectroscopy has become increasingly prominent in the structural determination of complex natural products, particularly in defining their absolute stereochemistry. This approach has been successfully applied in this work for the gargantulides family— stereochemically intricate large polyol macrolides- allowing the assignments of the absolute configuration for all their chiral centers (Chapter 3). There are notable examples that demonstrate the robustness of this interdisciplinary approach in the field of complex macrolide stereostructural assignments. The full structure and stereochemistry of compounds such as neaumycin B[159] and caniferolides[160] have been accurately assigned integrating gene cluster mining with NMR-based structural analysis rapidly and efficiently. This toolkit has been successfully applied for the structure elucidation of NPs beyond macrolides. For instance, in the study of krisynomycins,[161] a combination of experimental results from the application of Marfey´s advanced method and bioinformatics-based gene cluster analysis, allowed the proposition of the absolute configuration for these cyclic nonribosomal peptides. Another notable application is observed for the class V lanthipeptide cacaoidin,[162] where the interpretation of the genomic data from the ribosomal sequence of the precursor peptide has been pivotal in establishing the placement of D and L-alanine residues within the molecule.

5. Microbiology Techniques

Traditionally, microbiology has relied on the conventional use of single-strain cultivation to screen secondary metabolites. However, this approach falls short of replicating the diverse biotic and abiotic interactions experienced by microorganisms in their natural habitats.[56,163] As a result, the growth conditions in monocultures diverge significantly from the complex environmental conditions encountered in nature imposing limitations on the achievable chemical diversity through a single strain. In the natural environment, microbial metabolic pathways are often subjected to intricate regulatory cascades, wherein their chemical profiles are governed by genes and influenced by biotic interactions.[63]

In practice, the screening of monocultures frequently results in a growing redundancy, yielding chemically impoverished profiles and the repetitive isolation of previously known secondary metabolites.[164,165] This means that current research has to maximize the biosynthetic potential of microorganisms by different strategies to increase the likelihood of discovering useful molecules, ultimately leading to a greater chance of success.[22]

In recent decades, the study of genomes has demonstrated that numerous biosynthetic pathways within microbial strains remain inactive when subjected to traditional laboratory cultivation methods.[166] This phenomenon highlights the existence of a multitude of potentially valuable bioactive compounds that are ripe for exploration. As a result, it becomes imperative to implement diverse strategies aimed at unlocking the genuine biosynthetic capabilities of these microorganisms, especially in the realm of uncovering pharmaceutical candidates sourced from microbial natural products.

Over the past years, numerous methods have emerged with the aim of simulating physiological conditions capable of activating cryptic genes. These efforts are designed to stimulate biosynthetic pathways and unlock the production of previously unexpressed chemical diversity.[59,167] These strategies have proven successful in genomic activation, involving simultaneous modifications across various levels of the cellular machinery to regulate the production of distinct classes of secondary metabolites.[167–169]

Among those, genetic-dependent strategies such as BGC refactoring,[170] heterologous expression,[171,] and metabolic engineering[60] have emerged as highly potent methodologies. They have consistently demonstrated their effectiveness in inducing gene expression and optimizing the production of known compounds. Typically, these techniques are employed for known structures and require prior knowledge of the microbial genome.

In addition to strategies that directly manipulate the genes, alternative approaches known as post-genomic strategies or cultivation-dependent methodologies, have emerged as viable options to expand chemodiversity. These strategies aim to achieve a more impartial regulation of secondary metabolites by modifying various aspects of the cellular machinery. Notably, these methodologies do not rely on prior knowledge of the organism's genome or biosynthetic pathways and include substrate feeding,[172]

co-culture,[173] elicitor feeding,[174,] and the One Strain, Many Compounds (OSMAC) approach.[175–178] These cultivation-dependent methodologies offer a practicable avenue for unbiasedly enhancing chemodiversity without the need for extensive genetic information or specialized resources and have been commonly applied to the screening of microbial species in drug discovery programs.

6. Biosynthesis of Microbial Natural Products

Significant progress has been made in understanding the molecular mechanisms involved in the biosynthesis of NPs. In brief, these specialized metabolites are produced by biosynthetic gene clusters (BGC), which are groups of two or more physically adjacent genes within a genome that encode the biosynthetic pathway for a specific secondary metabolite class, which may encompass various chemical forms or derivatives. These BGCs encode not only for the biosynthetic enzymes that catalyze the formation of metabolites from simple building blocks but also for regulatory, resistance, and self-transporting proteins, ensuring the production and release of the specialized compounds outside of the cell.[179]

These genes within BGCs are capable of horizontal gene transfer (HGT) due to their association with mobile genetic elements, aggregation into clusters, and their identification within plasmids.[123] These genetic packages tend to undergo exchange events more frequently among closely related strains, resembling "selfish operons", a phenomenon explained by the DNA mismatch repair and maintenance mechanism.[180] Therefore, it is possible to investigate BGCs within the framework of these mobile elements and treat them as distinct evolutionary units relative to their host organisms.[181]

In actinomycetes, BGCs are located in diverse regions of the chromosome, called genomic islands.[67] Recently, the advancement of genomics and bioinformatics tools has led to an increase in the discover of these BGCs and their respective bioactive compounds.[182–185] For instance, 58, 21, and 30 BGCs were discovered in *Streptomyces clavuligerus*, *S. bottropensis*, and *S. avermitilis*, respectively.[185,186] Indeed, since 2013, several Actinobacterial genomes have been rapidly sequenced for the drug discovery race and a total of 1,749 *Streptomyces* genomes have been deposited and are available in the RefSeq database into 2020.[187]

The genome sequencing of different *Streptomyces* sp. strains has also revealed that many identified BGCs are orphans, meaning they have not been associated with the natural products they encode.[25] Even the genome of the widely studied *S. coelicolor* was shown to contain 18 BGCs that had not yet been linked to their corresponding NPs.[188] All these factors contributed to the concept of natural product genome mining, an approach that uses genes to uncover the biosynthetic pathways responsible for the production of natural products.[184]

The studies on the biosynthesis of model secondary metabolites have revealed universal biosynthetic principles that apply to the most frequently studied classes of natural products. While the biosynthetic pathways within a particular class of natural products can vary significantly in terms of the reactions involved and result in a wide range of products, they all share a set of core biosynthetic enzymes that have been identified through genome sequencing. The genes that encode these enzymes share enough homology degree to be identified using bioinformatics tools that target BGCs.[189]

The bioinformatics analysis of biosynthetic gene clusters provides valuable insights into the chemical nature of the specialized metabolites they produce. Although structural prediction of natural products solely based on *in silico* data of the bacterial genome is still not reliable, BGCs can be particularly informative given that they reveal not only the class of natural products but also their specific chemical scaffold and features. Computational tools, such as antiSMASH[190] (the most widely used tool for automatically identifying and analyzing BGCs), PRISM,[191] ClustScan,[192] or CLUSEAN[193] have been widely used to efficiently compare the genomes submitted by the user with those from public genome sequence databases (mainly, NCBI) and provide structural predictions of secondary metabolites. To accomplish this prediction, they identify signature enzymes that are involved in the biosynthesis of these metabolites, which can be used to pinpoint relevant gene clusters using genome-mining techniques. As a result, natural products can be classified according to the types of enzymes involved in their biosynthesis.[194] Initially, genome mining was accomplished by identifying and exploring central enzymes of biosynthetic pathways related to non-ribosomal peptide synthetases (NRPs) and polyketide

synthases (PKSs). However, nowadays these computational tools have evolved and can now identify partially or complete BCGs of many different NP structural classes.[185]

Among the major classes of microbial secondary metabolites, polyketides (PKs), non-ribosomal peptides (NRPs), ribosomally synthesized and post-translationally modified peptides (RiPPs) and terpenoids have been extensively studied regarding their biosynthetic machinery due to their crucial roles in drug discovery as bioactive compounds. Understanding the biosynthetic pathways of these NPs is essential for the development of new drugs and the optimization of existing ones through engineering and combinatorial biosynthesis.

In brief, nature has developed two fundamental approaches for the biosynthesis of NPs: the first process entails independent enzymes acting on unconstrained substrates, whereas the second process involves modular enzymes that sequentially construct product chains by assembling smaller, covalently linked building blocks.[189] Terpenes fall into the independent-enzyme category, where discrete enzymes carry out individual biosynthetic steps. On the other hand, NRPS and PKS biosynthesis operate through a thiotemplated mechanism, where thioesters fulfill the role of activating acyl groups of monomers and facilitating their subsequent capture by nucleophiles. Throughout the biosynthetic process, all substrates, intermediates, and end products remain covalently attached to the enzyme.[185]

The biosynthesis of RiPPs, in turn, is based on a leader peptide-guided strategy. In this, a precursor peptide (that includes the leader peptide, LP) contains a core region that is transformed into a mature product after the release of the leader peptide. The LP serves as an allosteric effector to activate the biosynthetic enzymes, keeps the maturing peptide inactive until extracellular export, and is required to ensure that the post-translational modifications (PTMs) are carried out in the correct order. It is noteworthy that this strategy results in highly evolvable pathways because many of the post-translational processing enzymes recognize the leader peptide and are highly permissive concerning mutations in the core peptide. As a result, more than 40 different RiPP families have been identified so far.[195] For example, in 2020, we reported the discovery of cacaoidin, a lantibiotic that is a ribosomally synthesized and post-translationally modified peptide (RiPP) characterized by the presence of lanthionine or methyllanthionine rings and their antimicrobial activity.[162]

In this doctoral thesis, we have focused on three of the NP classes: polyketides, NRPs, and hybrid terpenolid/polyketide metabolites. For this reason, the biosynthesis of RiPPs is not further discussed below.

Polyketides

Polyketides are widespread natural products present in different kingdoms and organisms including plants, animals (e.g., sponges, mollusks), prokaryotes (e.g., actinobacteria), and fungi. These NPs showcase a diverse array of bioactivities encompassing anti-inflammatory, anticancer, antibacterial, immune-suppressing antifungal, antiviral, and anti-cholesterol properties. Within Actinomycetes, polyketides form a crucial class of compounds and have been recognized for producing numerous bioactive metabolites such as actinorhodin (a blue pigment), daunorubicin, rapamycin, oleandomycin, and caprazamycin, among others.[196]

The biosynthesis of microbial polyketides is a highly intricate process that relies on multifunctional enzymes known as polyketide synthases (PKSs). PKSs operate through a mechanism akin to that of fatty acid synthase (FAS), wherein they combine multiple acyl-thioester units, such as malonyl-CoA and methylmalonyl-CoA, to generate diverse polyketide structures containing keto groups, hydroxy groups, and/or double bonds at distinct positions within the molecule.[197]

PKSs can be mainly categorized into three different types according to their architecture and functionality: (1) Multimodular Type I PKSs,[198] (2) Iterative Type II PKSs[199] and (3) Type III PKSs[200] (Fig. 4). Despite their variances in structure and mechanism, the synthesis of different polyketide types involves a stepwise decarboxylative condensation of acyl CoA precursors facilitated by the ketoacyl synthase (KS) domain. Acyl carrier protein (ACP) plays a crucial role in type I and II PKSs by activating acyl CoA substrates and guiding the progression of polyketide intermediates. Conversely, type III PKSs operate independently of ACP and directly target the acyl CoA substrates for catalysis.[201]

(1) Multimodular Type I PKSs[198] are characterized by one or more large multi-enzyme complexes, which are composed of catalytic domains arranged linearly and covalently attached. In this process, the progressive elongation of the polyketide chain occurs through a sequential transfer from one active site to another, a process referred to as "Lego-ization."[202] This modular arrangement enables each module to possess a unique set of non-repetitive activities that catalyze a single elongation cycle. As a result, diverse chemical variations and complexities are generated in a stepwise manner, determined by the specific catalytic domains within each module (Fig. 5). For instance, the trans-acyltransferase (AT) PKSs has free-standing proteins acting in *trans* instead of AT domains, forming polyketides in a mosaic-like fashion. As a result, the ATs create enormous architectural diversity, including features such as non-elongating modules, intermodular domain activity, uncommon domain orders, and split modules. Several PKs used as antibiotics were recently found to be trans-AT PKS products, including mupirocin[203] and streptogramins.[204]

(2) The iterative Type II PKSs[199] comprises discrete and monofunctional enzymes that assemble PKs with a specific chain length by the repetitive use of the same active sites iteratively through a single set of catalysts enzymes. Biosynthetic reactions involved in Type II PKSs include the condensation of building blocks to produce polyketone chains, followed by cyclization and aromatization processes to yield phenolic aromatic compounds such as tetracenomycin C.[205]

(3) Type III PKSs are acyl carrier protein (ACP) independent and iterative. The chalcone synthase (CHS)[200] superfamily is both, the most frequently occurring and the best studied within type III PKSs. Type III PKSs are distinguished by the presence of a polyketide chain that is never directly attached to a protein.[206] In these homodimeric enzymes, the growing polyketide chain remains near the enzyme active site, enabling the highly selective and efficient catalysis of polyketide intermediates. Type III PKSs have a simple structure and catalyze various reactions in a single catalytic center, yielding aromatic polyketides (often monocyclic or bicyclic), as exemplified by flaviolin.[207]
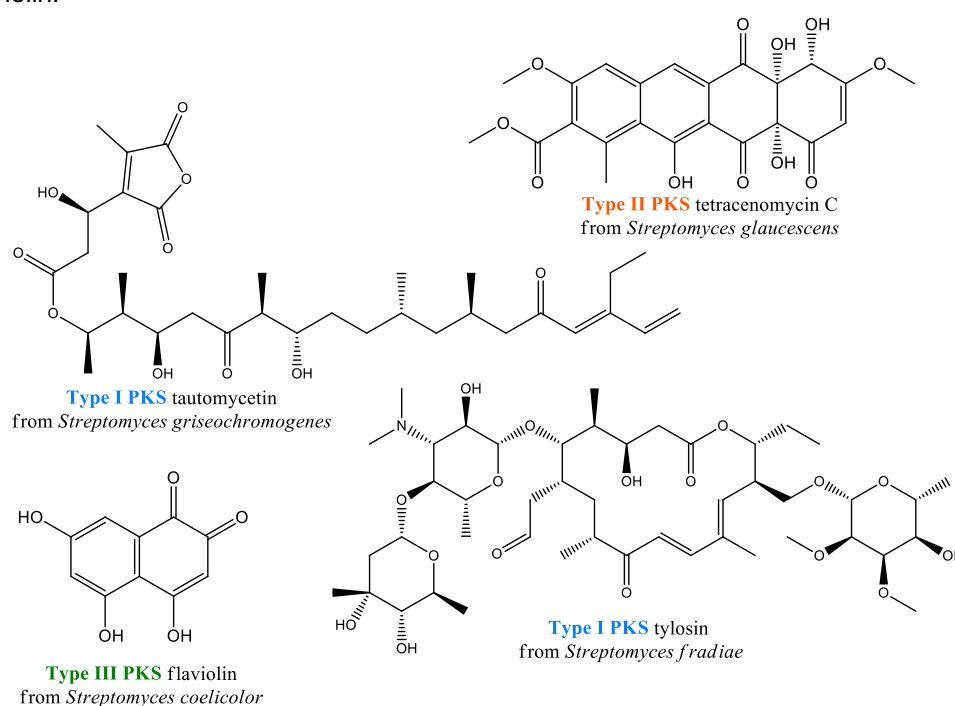

Type II PKS tetracenomycin C
from *Streptomyces glaucescens*

Type I PKS tautomycetin
from *Streptomyces griseochromogenes*

Type I PKS tylosin
from *Streptomyces fradiae*

Type III PKS flaviolin
from *Streptomyces coelicolor*

Figure 4. Schematic representation of the chemical structures for each of the three major types of polyketides mentioned above, according to their architecture and functionality. Tylosin and tautomycetin, exemplify the elaborate and multifaceted nature of type I PKS. Tetracenomycin C, illustrating the condensed aromatic structure characteristic of type II PKS. Flaviolin showcases the simplicity and fundamental architecture of type III PKS.

As discussed above, type I polyketide synthases (type I PKSs) are characterized by their large multifunctional proteins, consisting of numerous modules that house specific enzymatic domains responsible for distinct reactions.[196] In *Streptomyces*, type I PKSs are responsible for synthesizing reduced polyketides, including polyethers, polyenes, and macrolides such as tylosin[208] and erythromycin,[209] but are also involved in the biosynthesis of linear polyketides such as tautomycetin.[210]

Type I PKSs contain key domains within each module, including acyltransferase (AT), acyl carrier protein (ACP), and keto synthase (KS) working together to generate β-keto ester intermediates.[196] AT domains select and load the appropriate acyl-CoA building block onto the phosphopantetheine moiety of the ACP domain via thioester bond formation, where it undergoes a series of Claisen-type condensation reactions catalyzed by the KS domain between the growing polyketide chain and an extender unit (Fig. 5).[211] These collective actions lead to the formation of a polyketide chain of defined length in which two carbon atoms are added per module, generating a highly reactive ACP-bound β-ketoacyl intermediate. Remarkably, the order of PKS genes within the BGC typically follows the order of PKS proteins in the assembly line: This fact is commonly referred to as "colinearity" and enables the resulting polyketide backbone to be predicted solely based on the genome sequence.[212]

Expanded structural and stereochemical diversity can be introduced to the developing polyketide framework through the involvement of supplementary accessory domains, which modify the substituents and oxidation states of intermediates attached to the acyl carrier protein (ACP). These accessory domains encompass enzymes such as β-ketoreductases (KRs), dehydratase (DH), aromatases (AROs), cyclases (CYCs), enoyl reductase (ER), dimerases, P450 monooxygenases, methyltransferases, and glycosyltransferases. All these can modify the polyketide chain before and after the release from the PKS assembly line to yield natural products with unique properties and functions.[213] For example, the β-keto group can be modified by KR, DH, and ER domains, yielding functional groups such as alcohols and ketones. Ketoreductases (KR) are responsible for the reduction of the PKs, catalyzing the formation of hydroxy groups, which are essential for the formation of lactones, epoxides, and ethers. Dehydratases (DH) are responsible for the dehydration of β-hydroxy groups in the polyketide intermediate. This reaction is important for the formation of double bonds and conjugated systems in the final polyketide structure. Enoyl reductases (ER) catalyze the reduction of the double bond in the α,β-unsaturated intermediate, resulting in the formation of α,β-saturated acyl intermediates. This reduction step is important for the eventual presence of saturated, non-functionalized carbons, in the final polyketide structure. Overall, the KR, DH, and ER domains work together to modify the β-keto group and ultimately produce a diverse array of polyketide structures with unique functional groups and properties (Fig. 5).[212]

PKS systems usually contain thioesterase domains positioned at the end of the enzymatic chain.[189] In the final stage of assembly line processing, the polyketide intermediate bound to the acyl carrier protein (ACP) undergoes transfer to the carboxyl-terminal thioesterase (TE) domain. Their primary role involves facilitating the hydrolysis or macrocyclization of the polyketide chain, resulting in the release from the PKS assembly line (Fig. 5). Hydrolysis involves the cleavage of the ester bond between the polyketide chain and the TE domain, while macrocyclization refers to the formation of a cyclic structure from the linear polyketide intermediate. This can occur through different mechanisms such as ester bond formation (yielding macrolactones) or amide bond formations (leading to macrolactams). The specific mechanism depends on the structure of the polyketide intermediate and the thioesterase domain.[214]

Lastly, the production of certain types of polyketides involves the use of building blocks that are not derived from the decarboxylation of malonyl-CoA or methylmalonyl-CoA into acetate/propionate units. Thus, for instance, the biosynthesis of marginolactones (e.g., desertomycins, azalomycins) requires the use of guanidinobutanoate as a starter unit, which results in polyketides containing terminal alkylguanidino or alkylamino (when a final deamidation stage occurs) moieties.[215] The nature of these starter units, as well as the modifications that take place after the PKS process, can significantly impact the biological properties of the resulting polyketide.[216] All this diversity of natural products produced by PKS systems is a testament to the versatility of these biosynthetic pathways and the importance of

tailoring enzymes in natural product biosynthesis, which contributes significantly to the molecule's antibiotic activity.
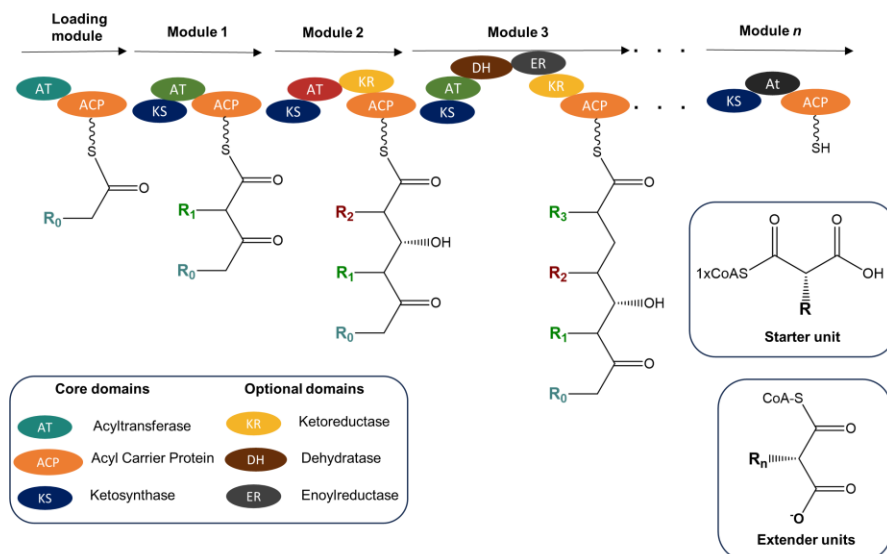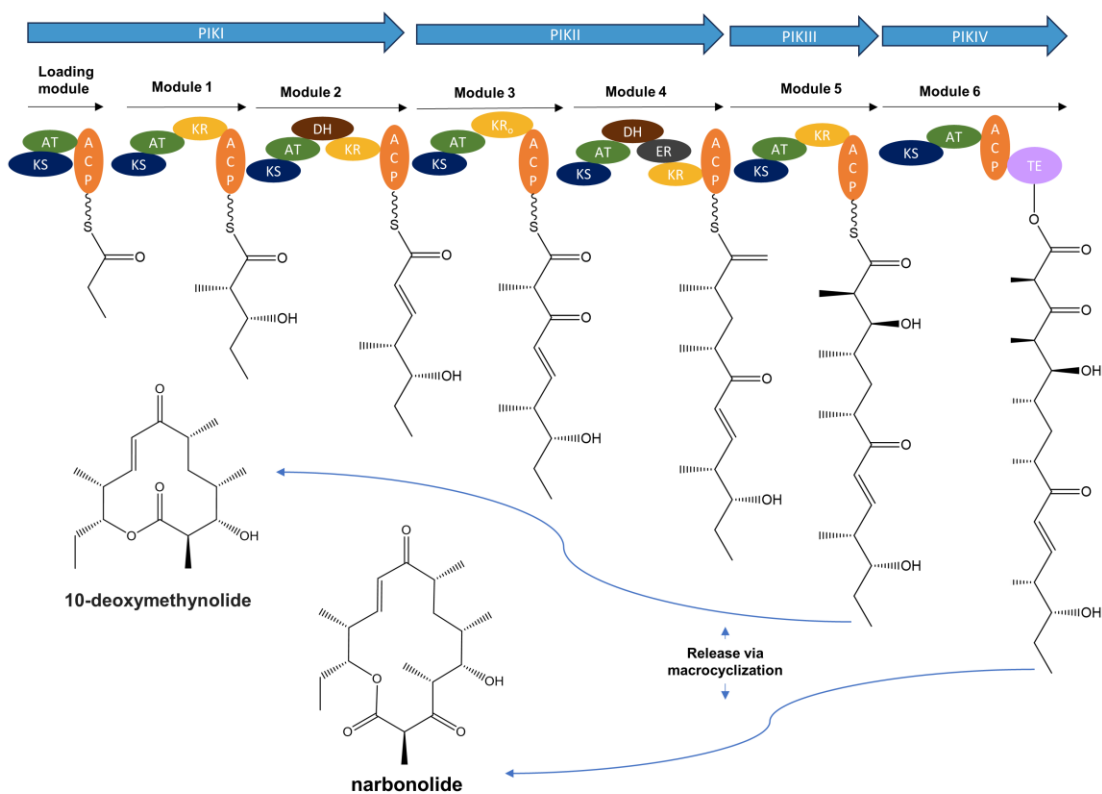

Figure 5. Schematics of multi-modular machinery of Type I Polyketides Biosynthesis. The illustration demonstrates the progression of the pikromycin PKS through six distinct modules, consecutively extending and altering a polyketide intermediate. This process yields two possible outcomes when released via macrocyclization (catalyzed by the C-terminal TE domain of the polyketide chain): 10-deoxymethylnolide, originating from module 5, or narbonolide, derived from module 6. Each module is visually distinguished by its unique color, while the protein domains are represented by circles. Figure adapted from Kornfuehrer et al 2019[217] and Duta et al 2014.[218]

Nonribosomal Peptides (NRP)

Analogous to type I PKSs, nonribosomal peptide synthetases (NRPSs) also share multimodular enzymatic assembly lines and play a crucial role in the production of various bioactive NPs.[189] The structural and functional variability of nonribosomal peptides (NRPs) results from the combinatory utilization and template-guided extension of approximately 550 amino acids, encompassing the 20 proteinogenic amino acids, a diverse array of nonproteinogenic amino acids, and additional carboxylic acids (e.g., aryl acids), which are incorporated into the expanding acyl chains.[219]

According to this multimodular enzymatic assembly model, a minimal NRPs domain is composed of three catalytically independent cores that extend the growing peptide chain by an extender unit (single amino acid or carboxylic acid). The activation of building blocks is catalyzed by the adenylation domain (A), which selects and activates the appropriate amino acid by attaching it to an adenosine monophosphate (AMP) molecule (similar to the function performed by acyltransferase (AT) domains in Type I polyketide synthase assembly lines). Subsequently, the aminoacyl-AMP intermediate is loaded onto the phosphopantetheine moiety of the thiolation domain (T) via thioester bond formation. The condensation domain (C) catalyzes then the formation of a peptide bond between the amino group of the recently loaded amino acid residue and the acyl group of immediately preceding residue in the evolving peptide chain, analogous to the function of the KS domain in Type I PKS systems (Fig. 6).[220]

Apart from the structural variations attained through the use of diverse noncanonical amino acid substrates, an additional crucial mechanism to produce both, structural and stereochemical diversity in NRPs, is the implementation of peptide tailoring reactions. These reactions are catalyzed by tailoring enzymes and/or additional NRPS accessory domains that introduce several functional groups such as hydroxyls, methyl groups, or carbonyls, into the peptide chain. Among these strategies, the inclusion of D-amino acids and methylated amide bonds enhances peptide stability against proteolytic degradation and promotes the emergence of distinct conformations, which play a crucial role in biological activity. Apart from epimerization and *N*-methylation, additional modifications can be introduced into the peptide sequence. Oxidases (Ox), methyltransferases (MT), and aminotransferases (AMT) are responsible for chemical modifications in the developing peptidyl chain.[27,221]

Once the mature peptide reaches the end of the assembly line, it needs to be cleaved from the enzyme complex. Typically, this cleavage reaction is facilitated by a thioesterase (TE) domain that is fused to the NRPS's C-terminal module. Termination modules typically follow a C-A-T-TE organization, similar to Type I PKSs.[214] During the concluding phase of peptide synthesis, a serine residue within the TE domain performs a nucleophilic attack on the PCP-peptidyl thioester, leading to the formation of a covalent acyl-enzyme intermediate. The fate of this intermediate is determined by the specific NRPS template and its corresponding TE domain. It can be released through hydrolysis, resulting in the formation of a linear peptide, or it can undergo an intramolecular reaction with an internal nucleophile, giving rise to a cyclic peptide. Alternatively, a TE-independent mechanism of chain release can occur via reductase-catalyzed thioester reduction, which produces a C-terminal aldehyde.[27,222]

Upon liberation from their assembly pathways, NRPs can still undergo diverse chemical modifications, including reduction, oxidation, glycosylation, or methylation, facilitated by specialized enzymes that are encoded within their biosynthetic gene clusters. These modifications are crucial for the biological efficacy of the peptide and modulate the hydrophobicity of the nonribosomal peptide framework, ultimately enabling the NRP to effectively fulfill its intended role and function.[220]
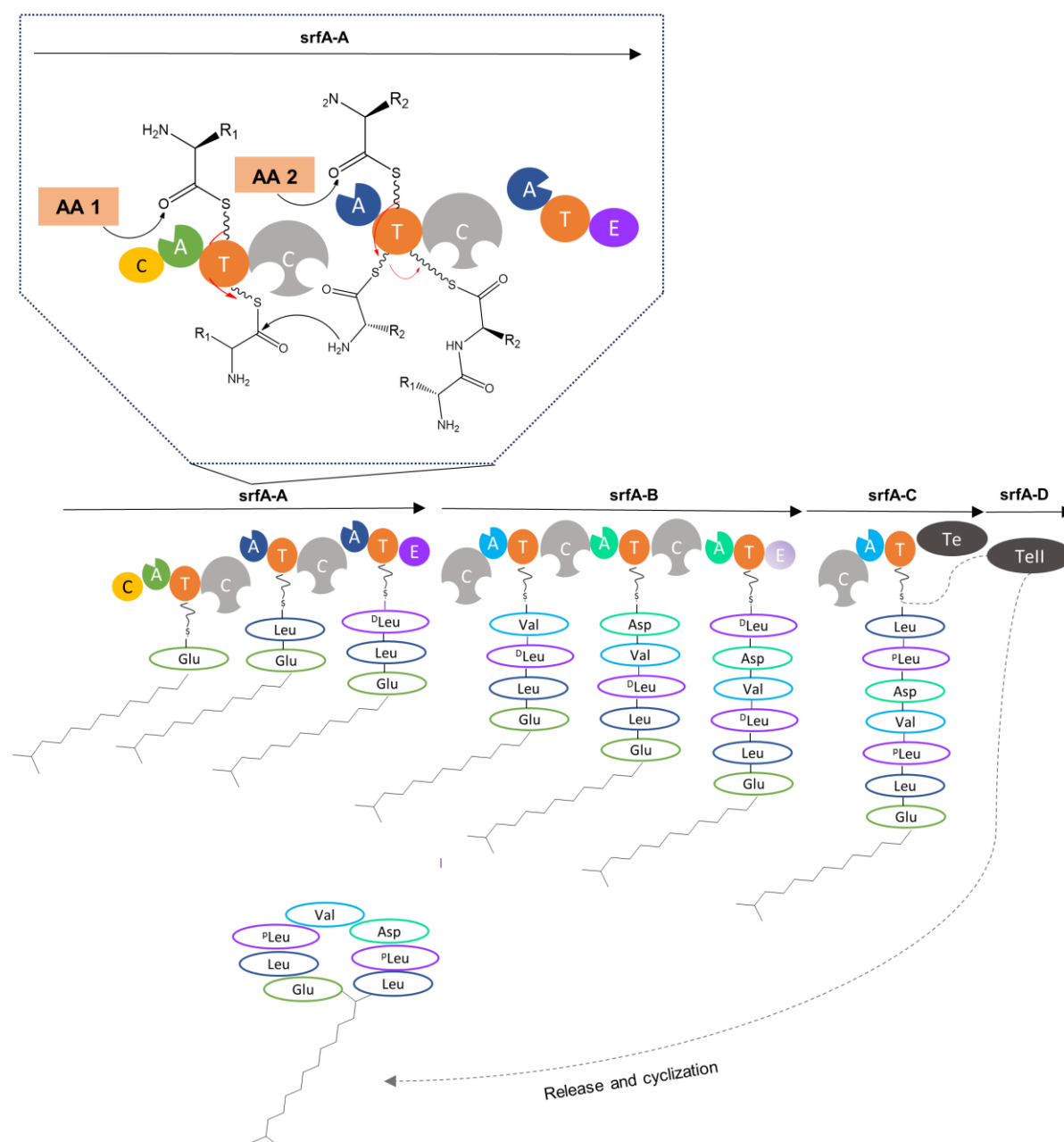
Figure 6. Schematics of multi-modular machinery of nonribosomal peptide (NRP) biosynthesis. Surfactin is used as an example. Its biosynthesis is catalyzed by four synthetase proteins (*srfA-A-D*) organized in modules. The first module is known as the initiation module and subsequent modules are known as elongation modules. Each module incorporates a single amino acid, therefore there are as many modules required as amino acids are present in the final peptide product. Domains are as follows: (A) adenylation, (C) condensation, (E) epimerization, (T) thiolation, and (Te) thioesterase. The amino acids incorporated are circled with the respective A domain or E domain color. Figure adapted from Winn et al., 2015[223] and Duban et al 2022.[224]

Terpenoids

With around 80,000 identified compounds belonging to 400 distinct structural groups, terpenoids represent the most expansive class of natural products in terms of chemical diversity. Unlike thiotemplate-based assembly pathways, terpenes are synthesized through the action of individual enzymes that facilitate separate biosynthetic reactions. Despite their chemical complexity, terpenes are synthesized using basic C5 isoprenoid units and two main phases.[225] Firstly, dimethylallyl phosphate starter units are polymerized through oligoprenyl synthetases using one or more isopentenyl diphosphate extender units, resulting in a linear oligoprenyl precursor. Secondly, terpene cyclases, the signature enzymes for this process, convert the resulting polyisoprenoid intermediates into one or more complex cyclized hydrocarbon backbones. During the second phase, tailoring enzymes heavily modify the backbone, resulting in the generation of several products from a single scaffold.[226]

The beginning stages of terpene biosynthesis involve the combination of dimethylallyl pyrophosphate (DMAPP) and isopentenyl pyrophosphate (IPP), which can be synthesized via the methyl-erythritol phosphate (MEP) or the mevalonic acid (MVA) pathways (Fig. 7). The cytosolic MVA pathway forms the precursor mevalonic acid from acetyl-CoA, whereas the other uses MEP derived from pyruvate and glyceraldehyde-3-phosphate (G3P). The utilization of distinct precursors and enzymes characterizes each pathway, and their presence varies among different organisms, which may employ one of them or both.[226]

As outlined above, the generation of terpenoid scaffolds involves enzymatic hydrocarbon backbone assembly and cyclization, performed by oligoprenyl synthetases and terpene cyclases (TCs), respectively. Terpene synthases (TPS) play a crucial role as sentinels in the biosynthesis of C10–C20 terpenoids, catalyzing the essential conversion of specific prenyl diphosphate substrates into a diverse array of hydrocarbon or oxygenated structures, thereby forming the fundamental scaffolds. The condensation of IPP and DMAPP results in geranyl diphosphate (GPP), serving as the building block for monoterpenoids (C10). The fusion of GPP with an additional IPP molecule produces farnesyl diphosphate (FPP), the precursor for sesquiterpenoids (C15). By combining FPP with IPP, geranylgeranyl diphosphate (GGPP) is generated, which serves as the precursor for diterpenoids (C20). Additionally, the condensation of two FPP or two GGPP molecules forms the core substrates for triterpenoids (C30) and carotenoids (C40), respectively.[227]

In contrast to the conserved core biosynthetic enzymes and domains found in PKSs and NRPSs, bacterial terpene cyclases (TCs) exhibit minimal overall sequence similarity. This lack of conservation in the primary sequence has hindered our understanding of terpene cyclization processes. These terpene cyclases are capable of liberating diphosphate from their linear substrates, enabling a wide range of cyclization and rearrangements that give rise to an extensive array of monocyclic and polycyclic hydrocarbon frameworks.[225] After the cyclization process, terpenoid scaffolds often undergo additional modifications to enhance their structural complexity through the incorporation of diverse functional groups. These modifications may include acetylation, glycosylation, and methylation, among others. The diversity in the size of terpene BGCs is directly correlated to the inclusion or exclusion of tailoring genes, which significantly impacts their complexity.[222] Within these clusters, the cytochrome P450 (CYP450) superfamily of enzymes plays a prominent role in expanding the range of terpene structures through oxidative modifications.[225] Specifically, the products of TPS can undergo various oxygenation reactions, followed by additional functional modifications, ultimately resulting in the generation of over 80,000 unique natural products.
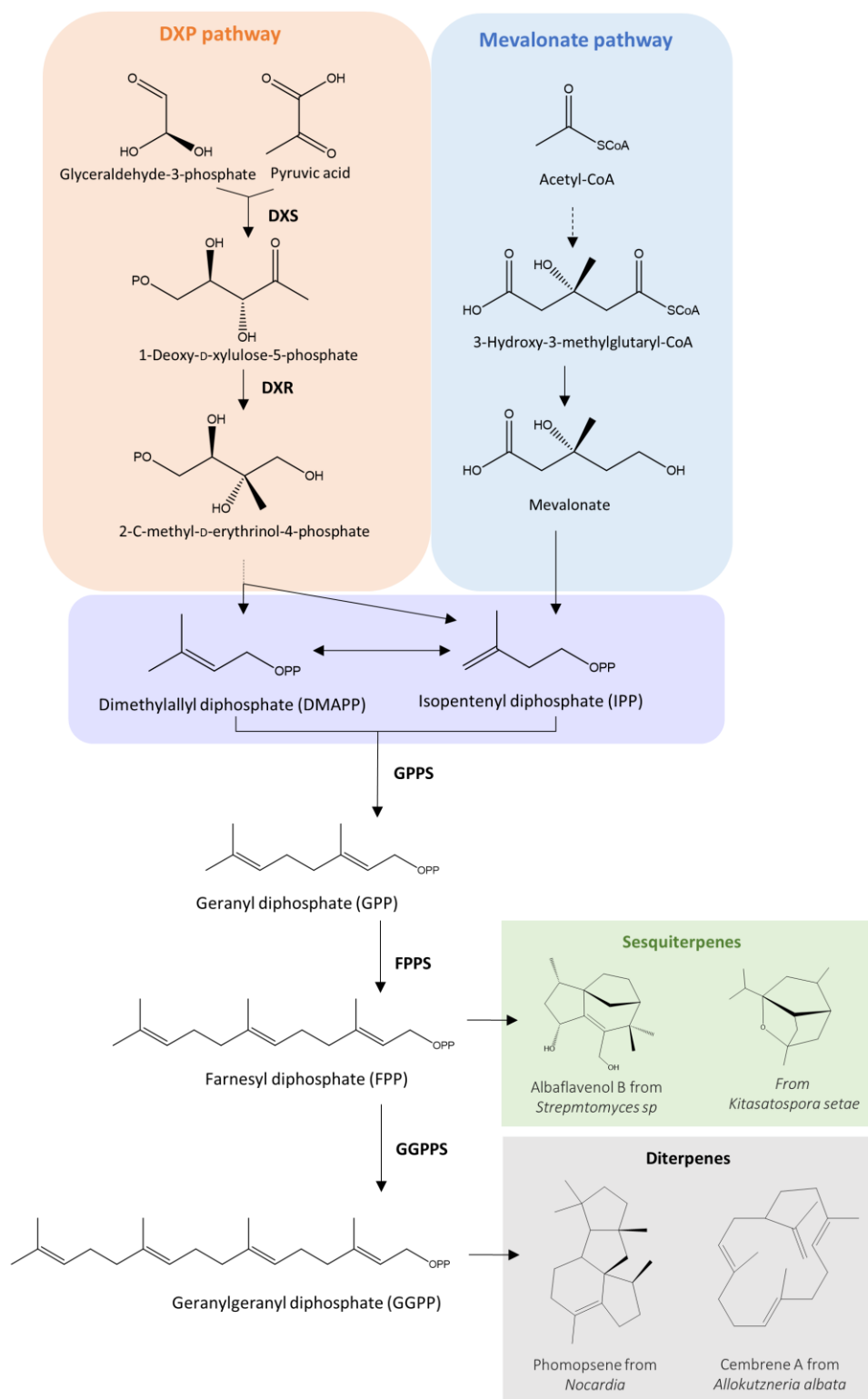
Figure 7. Overview of the terpene biosynthetic pathways. The 1-deoxy-D-xylulose-5-phosphate (DXP) and mevalonate pathways are responsible for the production of IPP and DMAPP from central metabolites. Key enzymes of each pathway are shown: 1-deoxy-D-xylulose-5-phosphate synthase (DXS), 1-deoxy-D-xylulose-5-phosphate reductase (DXR), 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR), and IPP-DMAPP isomerase (IDI). IPP and DMAPP are then converted to terpene synthase precursors through the action of the prenyltransferases GPP synthase (GPPS), FPP synthase (FPPS), and GGPP synthase (GGPPS). One example of a terpene synthase reaction and downstream processing reaction(s) is given for sesquiterpenes and diterpenes from Actinomycetes. Multiple steps are indicated by dashed lines. Figure adapted from Kirb and Keasling, 2009.[228]

In addition to the major classes of specialized metabolites, Nature has evolved to combine privileged scaffolds, thus providing "hybrid NPs". This hybridization approach may be viewed as the combination of biosynthetic strategies to produce new molecules that explore new areas of biologically relevant chemical space. Such hybrid NPs are sometimes the product of homo- or heterodimerizations of smaller NPs in which the monomers may have, but not necessarily, different biosynthetic origins (e.g., thiomarinol[229] or vincristine).[230] Many other times, these hybrid compounds arise from the genomic combination of unrelated biosynthetic pathways. After combination, biosynthetic cascades can further diverge or merge the original metabolic units into complex scaffolds that do not intuitively resemble the original parts.[231] Polyketide-non ribosomal peptide (PKS-NRPS) and polyketide-terpenoid (PKS-terpene) hybrids are the most frequent combinations, which result in exotic and biologically relevant NPs such as rapamycin[232] or xenovulene A,[233] respectively.

Among the PKS-terpene hybrids, the napyradiomycin family of bacterial meroterpenoids is a prominent example of hybrid NPs with cytotoxic and antibacterial bioactivities.[234–241] These NPs result from the combination of type III PKS (tetrahydronaphtalene synthase) and terpene biosynthetic machinery, with the involvement of vanadium-dependent chloroperoxidases (VCPOs) as key enzymes responsible for the class-defining chloronium-induced cyclizations (Fig. 8).[242,243] This family of bacterial meroterpenoids has been the subject of study in this work and the results are discussed in Chapter 1.
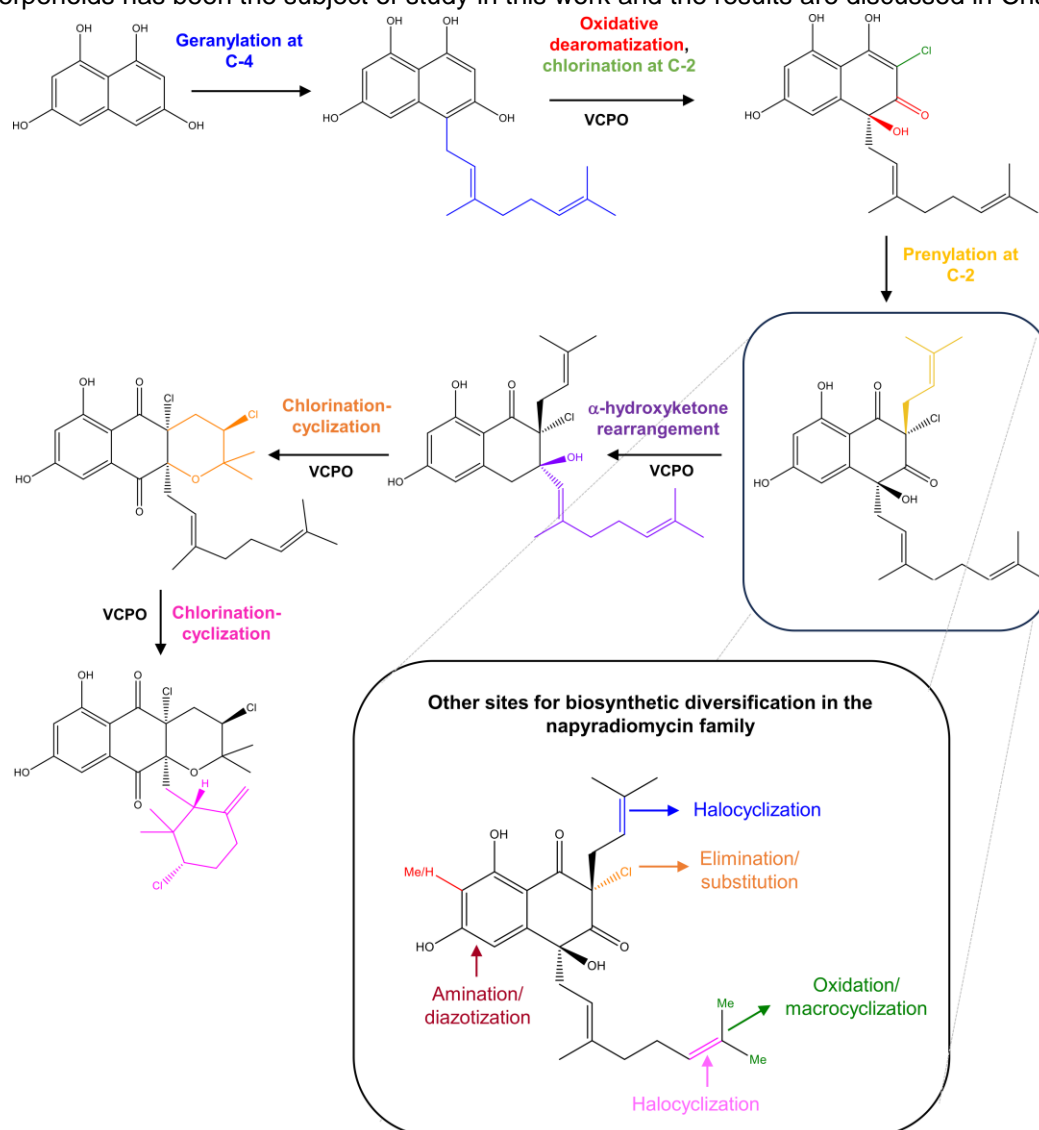
Figure 8. Proposed biosynthetic pathway of napyradiomycin. Legend: VCPO = Vanadium-Dependent Chloroperoxidases Family Enzymes. Figure adapted from McKinnie et al., 2019.[243]

AIMS

Aims

The overall aim of this Doctoral Thesis is to explore and expand the chemical space of microbial specialized metabolites produced by strains from the MEDINA actinomycetes collection, in order to provide novel natural products that may satisfy the need for new bioactive compounds or improve our knowledge of their biosynthetic pathways.

To achieve this main goal, we set the following **Objectives**:

- Select talented strains with remarkable biosynthetic potential and explore conditions to produce novel bioactive compounds of different families of natural products.

- Explore different classes of NPs as potential antibiotic candidates through an antibacterial screening campaign aimed at finding active compounds against a set of Gram-negative pathogens.

- Isolate and elucidate the structures of new bioactive or structurally novel natural products using spectroscopic approaches, notably (tandem) mass spectrometry and NMR spectroscopy.

- Identify the putative biosynthetic gene clusters (BGCs) responsible for producing the selected metabolites obtained by using bioinformatics tools (e.g., antiSMASH), as well as expand the knowledge on their biosynthetic pathways.

- When applicable, combine genome-based bioinformatics analysis and NMR spectroscopy to fully determine the absolute configuration of the isolated natural products.

- Characterize the antimicrobial activity profile of the newly isolated compounds against clinically relevant human pathogens including bacteria of the ESKAPE panel and other multidrug-resistant (MDR) pathogens such as *E. coli*, methicillin-resistant *S. aureus* (MRSA), vancomycin-resistant enterococci (VRE), *Mycobacterium tuberculosis* and *Aspergillus fumigatus*.

PREFACE

Preface

The methodologies applied to Natural Products Drug Discovery can be broadly classified into two groups, known as top-down and bottom-up approaches.[244] The top-down strategy goes from metabolome to genome and entails identification through chemical signatures and/or bioassay-guided isolation. Contrarily, the bottom-up approach revolves around genomics-driven exploration and engineering and relies on the activation and expression of the BGC. These two paradigms broadly delineate the pathways through which NPs are unearthed and comprehended (Fig. 1).

Both strategies have significantly enriched the landscape of natural product discovery and are led by recent advancements in genomics and metabolomics platforms and tools. Genomics-driven techniques have augmented our capacity to glean genetic information, predict biosynthetic pathways, and engineer organisms for the production of target compounds. Simultaneously, diverse sampling techniques, innovative culturing conditions (e.g., OSMAC), and advanced screening methodologies have amplified the efficacy of top-down strategies. These advancements have enabled the creation of microbial libraries that are far richer and more extensive, surpassing previous limitations and setting a new benchmark for biodiversity exploration and bioprospecting.
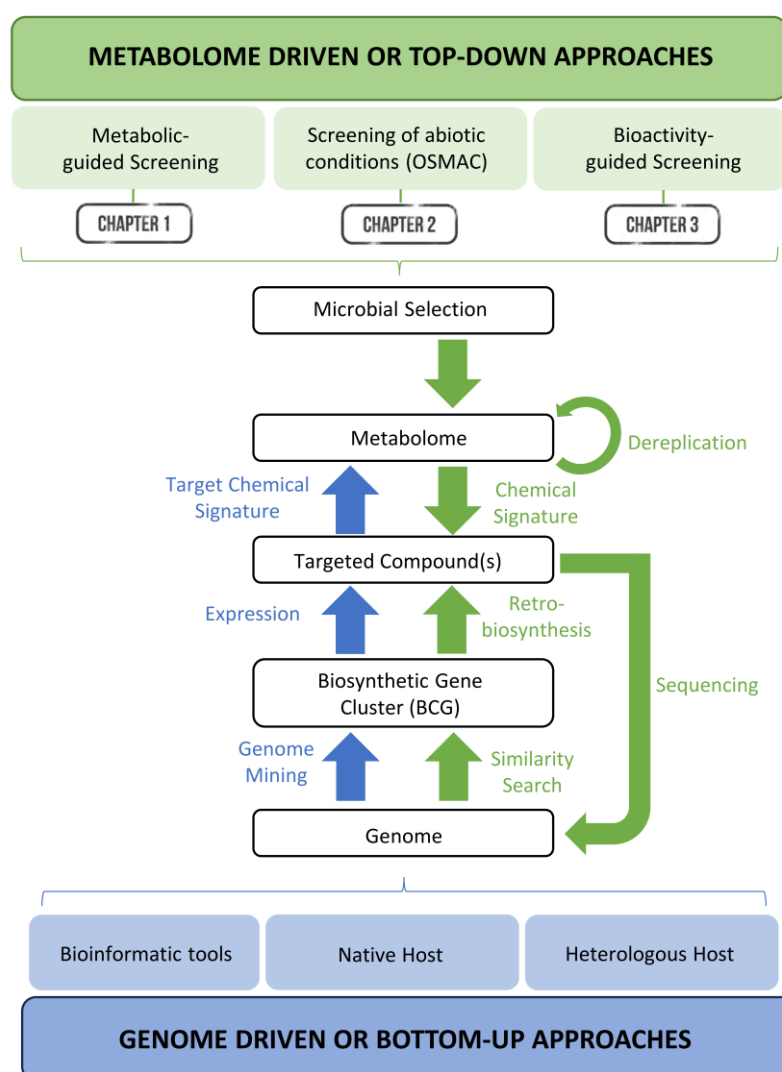
Figure 1. Schematic illustration of top-down vs. bottom-up approaches applied in NP discovery. Figure adapted from Subko, 2020.[245]

The top-down approach starts at the (micro)organism level, involving the direct extraction of samples from the environment or cultivation under diverse laboratory conditions (OSMAC). The

compounds of interest are subsequently isolated based on bioassay results and chemical signatures. This approach may also incorporate retrobiosynthetic analysis to infer core enzymatic activities and associated biosynthetic pathways, ultimately culminating in a comprehensive characterization of the target compound. Conversely, the bottom-up approach initiates at the (meta)genomic level, harnessing the potential of genomics and bioinformatics. Sequencing of environmental DNA (eDNA) or DNA extracted from specific microorganisms facilitates the identification of highly promising biosynthetic gene clusters (BGCs). These BGCs are then expressed in their native or heterologous hosts, with the identification of the target compound relying on chemical signature searches and metabolome analysis.

Top-down advances are among the most used strategies for the screening of new antibiotic drugs and are particularly noteworthy for institutions such as MEDINA, which harbors one of the largest microbial collections in the world, with about 190.000 strains, covering filamentous fungi, actinomycetes, and other bacteria. Particularly, the actinomycete collection, about 120.000 strains, encompasses a broad range of strains from the genus *Streptomyces* and many other minor taxa, including about ten percent of marine actinomycetes isolated from marine invertebrates and sediment samples collected along the seafloor.

To accomplish the overall aim of this doctoral thesis work, we have employed three different top-down approaches:

I)   First, we carried out chemical investigations of a marine *Streptomyces* sp. strain previously known to produce napyradiomycin meroterpenoids,[241] in order to isolate and characterize, -structurally and biologically-, new congeners of this interesting family of natural products. In addition, the genome sequencing of the producing strain allowed us to identify the corresponding BGC and hypothesize the biosynthetic pathways for the novel specialized metabolite. This work is summarized in Chapter 1 and was carried out at MEDINA facilities, except the whole genome sequencing (WGS) of the producing strain, which was performed at the Center for Biosustainability (DTU-Biosustain) in Copenhagen (Denmark).

II)   Second, the identification of the BGC responsible for producing pentaminomycins in a *Streptomyces cacaoi* strain from our microbial collection, prompted us to explore the production of novel members of this remarkable family of cyclic pentapeptides, by applying different fermentation conditions (Chapter 2). The experimental work of this chapter was carried out entirely at MEDINA´s facilities.

III)   Third, a bioassay-guided isolation process in the context of a screening campaign led to the discovery of two new 52-membered macrolactones, gargantulides B and C, with antibacterial activities against clinically relevant pathogens. The genome sequencing of the producing strain, *Amycolatopsis* sp, identified the BGC responsible for the biosynthesis of these huge macrolides. Bioinformatic analysis of the gene cluster, combined with extensive NMR analyses, led to the complete determination of the absolute configuration of this family of exceptionally complex polyketides (Chapter 3). The experimental work of this chapter was carried out at MEDINA and during a secondment at the Novo Nordisk Foundation Center for Biosustainability (DTU-Biosustain).

As a general workflow in these top-down approaches, selected actinomycetes from MEDINA strains' collection were cultivated in liquid media and, following incubation, their metabolites were extracted with liquid solvents. The choice of extraction method determines the classes of metabolites that will be present in the extracts. For instance, polar solvents such as methanol, ethanol, or acetone result in the extraction of polar compounds, such as alkaloids and polyketides. Conversely, apolar solvents such as chloroform and hexane are more suitable for the extraction of lipophilic compounds such as terpenoids, steroids, or lipids. This was typically followed by HPLC-MS and NMR analysis of the crude extracts, and the chemical information was then subjected to dereplication using different NPs databases and applied as a first filter to avoid the re-isolation of

known compounds. Targeted or bioassay-guided isolation was subsequently conducted to obtain the pure compounds, whose structures were elucidated by high-resolution mass spectrometry (HRMS) and NMR spectroscopy. Biological activity was evaluated in the search for antibacterial or antifungal activity by serial dilution of pure compounds into microbial pathogenic strains. Subsequently, whole genome sequencing (WGS) of the producing strains was performed in collaboration with DTU-Biosustain, except in the case of pentaminomycins (Chapter 2), where the genome sequence was already available. Bioinformatics analysis of the genome sequence was performed in each case by using antiSMASH or PRISM to identify the BGC responsible for the production of the new compounds (Fig. 2).
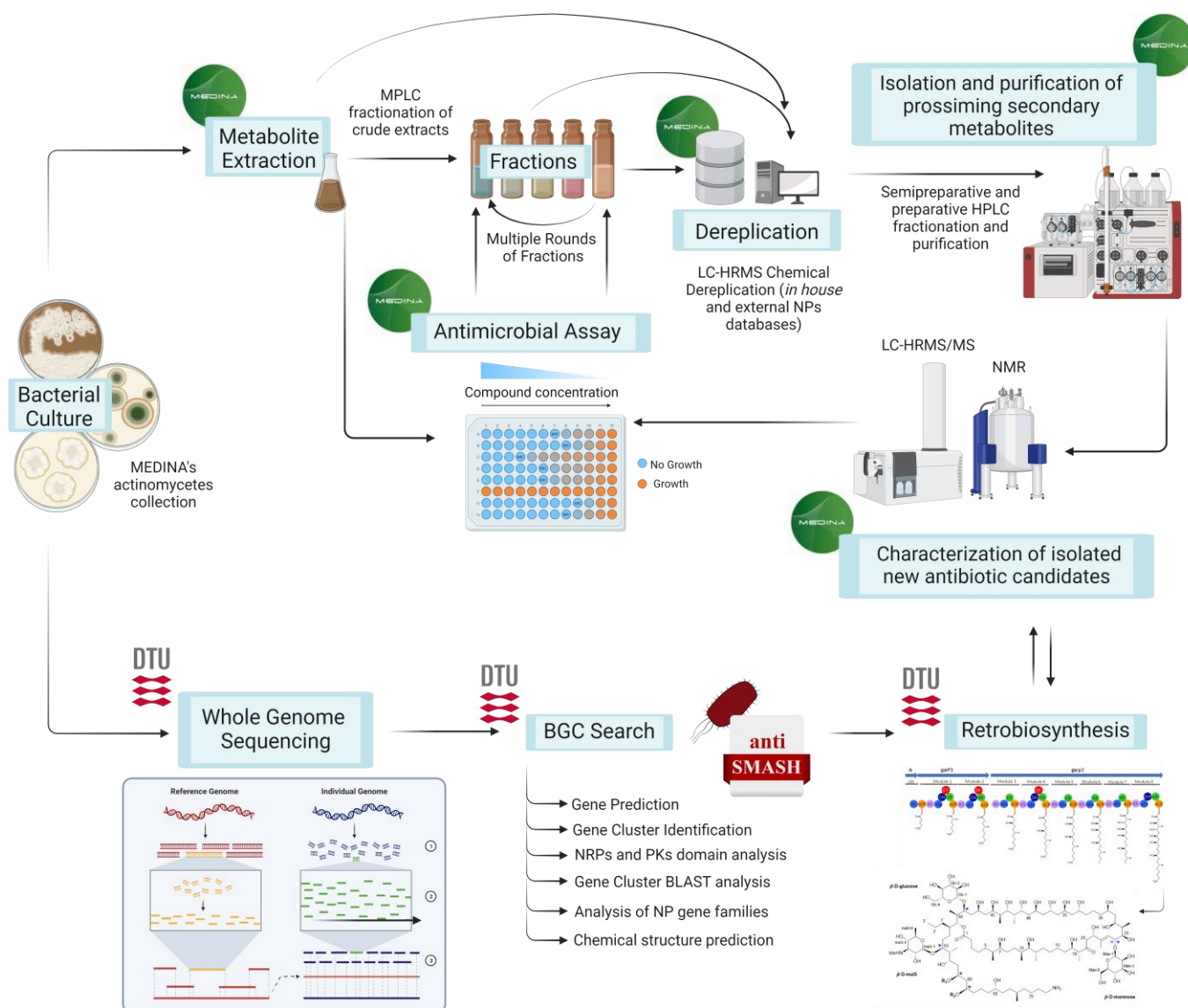
Figure 2. General workflow of bioactivity guided or targeted isolation of new natural products applied in this work, including further genomic characterization of the producing strains. Created with BioRender.com.
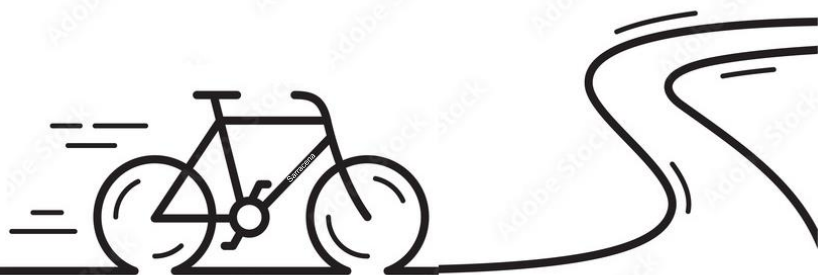
CHAPTER 1

Chapter 1. New Napyradiomycin Analogues from *Streptomyces* sp*.* Strain CA-271078

Daniel Carretero-Molina,[1,2] Francisco Javier Ortiz-López,[1,*] Jesús Martín,[1] Daniel Oves-Costales,[1] Caridad Díaz,[1] Mercedes de la Cruz,[1] Bastien Cautain,[1] Francisca Vicente,[1] Olga Genilloud[1] and Fernando Reyes[1,*]

Resumen: Como parte de nuestro continuo esfuerzo para descubrir nuevos compuestos bioactivos a partir de fuentes microbianas, una nueva investigación de extractos de cultivos a mayor escala de la cepa de origen marino *Streptomyces* sp. CA-271078, dió como resultado el aislamiento y la elucidación estructural de cuatro nuevas napiradiomicinas (**1**–**3**,**5**), la napiradiomicina SC (**4**), cuyos detalles estructurales no se habían reportado previamente, y otros diez análogos conocidos (**6**-**15**). Las estructuras de las nuevas napiradiomicinas se caracterizaron mediante HRMS y espectroscopía de RMN 1D y 2D, y sus configuraciones relativas se establecieron mediante una combinación de modelado molecular y análisis de RMN de constantes de acoplamiento y *nOe*. Se propone también la configuración absoluta de cada compuesto basándonos en argumentos biosintéticos y en la comparación de datos de rotación específicos con los de compuestos relacionados. Entre los nuevos compuestos, se determinó que **1** era el primer miembro no halogenado de la serie A de napiradiomicina que contenía una cadena lateral de prenilo funcionalizada, mientras que **2**–**4** albergaban en sus estructuras el anillo de clorociclohexano característico de la serie B de napiradiomicinas. Notablemente, el compuesto **5** muestra un anillo de éter cíclico de 14 miembros sin precedentes entre la cadena lateral de prenilo y el cromóforo, lo que representa el primer miembro de una nueva clase de napiradiomicinas que hemos designado como napiradiomicina D1. Las propiedades antiinfecciosas y citotóxicas de todos los compuestos aislados se evaluaron frente a un conjunto de microorganismos patógenos y la línea celular HepG2, respectivamente. Entre los nuevos compuestos, la napiradiomicina D1 exhibió una importante actividad inhibidora del crecimiento frente a *Staphylococcus aureus* resistente a la meticilina, *Mycobacterium tuberculosis* y HepG2

.

Affiliations

[1] Fundación MEDINA, Centro de Excelencia en Investigación de Medicamentos Innovadores en Andalucía, Avda. del Conocimiento 34, 18016, Armilla (Granada), Spain; daniel.carretero@medinaandalucia.es (D.C.-M.); jesus.martin@medinaandalucia.es (J.M.); daniel.oves@medinaandalucia.es (D.O.-C.); caridad.diaz@medinaandalucia.es (C.D.); mercedes.delacruz@medinaandalucia.es (M.d.l.C.); bastien.cautain@medinaandalucia.es (B.C.); francisca.vicente@medinaandalucia.es (F.V.); olga.genilloud@medinaandalucia.es (O.G.)

[2] Doctoral Programme in Pharmacy (B15.56.1), Doctoral School in Health Sciences, University of Granada, Spain.

* Correspondence: javier.ortiz@medinaandalucia.es (F.J.O.-L.); fernando.reyes@medinaandalucia.es (F.R.); Tel.: +34-958-993-965 (F.J.O.-L. and F.R.)

Abstract: As part of our continuing efforts to discover new bioactive compounds from microbial sources, a reinvestigation of extracts of scaled-up cultures of the marine-derived *Streptomyces* sp. strain CA-271078 resulted in the isolation and structural elucidation of four new napyradiomycins (**1**–**3,5**), the known napyradiomycin SC (**4**), whose structural features had not been previously described, and another ten related compounds (**6-15**). The structures of the new napyradiomycins were characterized by HRMS and 1D- and 2D-NMR spectroscopies and their relative configurations were established through a combination of molecular modeling with *nOe* and coupling constants NMR analysis. The absolute configuration of each compound is also proposed based on biosynthetic arguments and the comparison of specific rotation data with those of related compounds. Among the new compounds, **1** was determined to be the first non-halogenated member of napyradiomycin A series containing a functionalized prenyl side chain, while **2**–**4** harbor in their structures the characteristic chloro-cyclohexane ring of the napyradiomycin B series. Remarkably, compound **5** displays an unprecedented 14-membered cyclic ether ring between the prenyl side chain and the chromophore, thus representing the first member of a new class of napyradiomycins that we have designated as napyradiomycin D1. Anti-infective and cytotoxic properties for all isolated compounds were evaluated against a set of pathogenic microorganisms and the HepG2 cell line, respectively. Among the new compounds, napyradiomycin D1 exhibited significant growth-inhibitory activity against methicillin-resistant *Staphylococcus aureus*, *Mycobacterium tuberculosis,* and HepG2.

Keywords: napyradiomycins; marine actinomycetes; structural elucidation; antimicrobial activity; cytotoxicity

1. Introduction

The napyradiomycins are a large class of unique meroterpenoids with different halogenation patterns whose structures consist of a semi-naphthoquinone chromophore, a prenyl unit attached at C-4a that is cyclized to form a tetrahydropyran ring in most cases and a monoterpenoid subunit attached to C-10a.[234–237,246,247] Many of the structural variants within this family reside in this C-10a attached side chain. Hitherto, about 50 napyradiomycin (NPDs) derivatives have been discovered. They have been sorted into three different types according to their structural features: Type A (NPD-As) with a linear terpenoid side chain; Type B (NPD-Bs) where the side chain is cyclized to form a cyclohexane ring; and Type C (NPD-Cs), whose monoterpenoid subunit is cyclized between C7 and C10a of the naphthoquinone core to form a 14-membered ring.[234–238,246–253] Compounds belonging to this structural class display a wide range of biological activities, including cytotoxic and antibiotic properties, as well as ATPases inhibition or estrogen receptor antagonization.[239,254–256] Napyradiomycins were first isolated from the soil-derived bacterium *Chainia rubra* in Japan in 1986 (later transferred to the genus *Streptomyces*).[234–240,246–256] As time passed, a series of congeners were isolated from different actinomycetes including marine-derived strains.[236–238,246–249]

Multi-drug resistance is one of the emergent threats in the healthcare area due to the loss of effective activity of some drugs against multi-resistant bacteria.[257] The discovery and development of new and safer sources of antibiotics have therefore become an essential matter. Microorganisms from the marine environment are an important source of structurally diverse and biologically active secondary metabolites, as evidenced by the growing number of new marine natural products isolated yearly from different biological sources, but research into the pharmacology of marine organisms is limited and most of it remains unexplored.[258,259]

MEDINA´s collection of actinomycetes and filamentous fungi is one of the biggest microbial collections worldwide and contains about ten percent of marine actinomycetes isolated from sediment samples collected along the seafloor.[260] As part of our continuous efforts to discover new bioactive compounds from microbial sources, we initiated a more in-depth study of scaled-up cultures of *Streptomyces* sp. CA-271078, a marine-derived actinomycete strain producing MDN-0170 in whose extracts we had also observed the presence of minor structurally related napyradiomycin congeners by LC/MS analysis.[241] Herein, we report the isolation, structure elucidation, and biological activities of four new napyradiomycin congeners (**1**–**3,5**) together with

the description of the spectroscopic features of the known napyradiomycin SC (**4**), all isolated from the ethyl acetate extract of a culture broth of this strain.

2. Results

2.1. Isolation and Taxonomy of the Producing Microorganism

The isolation and taxonomy of the producing strain, CA-271078, were reported previously based on nearly complete 16S rRNA gene sequences (1359 bp, 94.2% coverage).[241] The data obtained strongly indicated that strain CA-271078 is a member of the genus *Streptomyces* and showed the closest relatedness with *Streptomyces aculeolatus* NBRC 14824(T) (99.34% similarity).

2.2. Extraction, Dereplication, and Bioassay-Guided Isolation

The producing strain CA-271078 was fermented at 28°C in 3 L of R358 medium for 6 days. Extraction with an equal volume of acetone and evaporation of the organic solvent after centrifugation and filtration to discard the mycelial debris afforded an acetone crude extract, which was subsequently subjected to liquid–liquid extraction with EtOAc (3 × 1 L). Analysis of this extract by LC/HRMS revealed the presence of some compounds that were not included in our in-house microbial natural products library[261] nor the Chapman and Hall Dictionary of Natural Products.[262]

The extract was subsequently chromatographed on a reversed-phase C18 column using a gradient of acetonitrile in water to afford five fractions: A–E. LC-DAD-HRMS analysis allowed us to establish that all these fractions contained possible known and bioactive NPDs bearing chlorine (according to their isotopic pattern), such as napyradiomycin B6[263] and 18-hydroxynapyradiomycin A1.[247] Additionally, these fractions also contained minor amounts of related NPDs whose molecular formulae suggested their novelty as natural products since they were not found in the Dictionary of Natural Products.[262] Further chromatographic separation on semipreparative reversed-phase HPLC using a gradient of CH$_3$CN/H$_2$O, allowed us to isolate fifteen compounds (Figure S43). Napyradiomycins A3 (**1**), B7a (**2**), B7b (**3**), and D1 (**5**) (Figure 1) were identified as new compounds based on ESI-TOF and NMR analyses. Additionally, we have also isolated napyradiomycin SC (**4**)[264] (Fig. 1), and its hitherto undescribed spectroscopic features will be reported here.
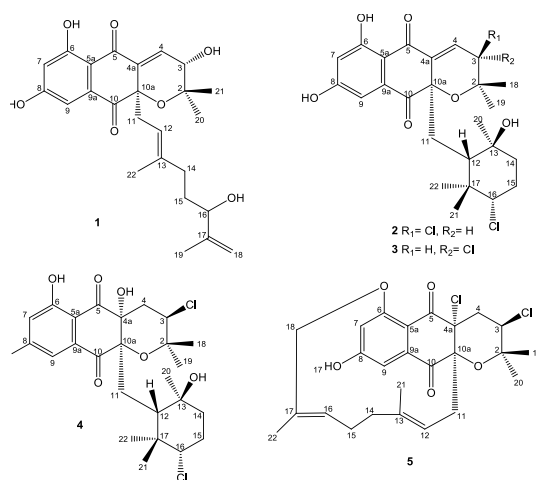
Figure 1. Compounds 1–5 isolated from culture broths of *Streptomyces* sp. CA-271078.

Finally, the spectroscopic data of the rest of the compounds isolated (**6**–**15**) were identical to those previously reported for MDN-0170 (**6**),[241] 3-chloro-6, 8-dihydroxy-8-α-lapachone (**7**),[250] 3-chloro-6-hydroxy-8-methoxy-α-lapachone (**8**),[247] napyradiomycin B6 (**9**),[263] 18-hydroxynapyradiomycin A1 (**10**),[247] napyradiomycin A2a (**11**),[247] napyradiomycin A2b (**12**),[247] napyradiomycin B4 (**13**),[238] napyradiomycin B2 (**14**),[234] and napyradiomycin B5 (**15**)[263] (Fig. S43).

2.3. Structural Elucidation

Compound **1** was obtained as a white powder and was assigned the molecular formula $C_{25}H_{30}O_7$ (11 degrees of unsaturation) by analysis of the sodium adduct present in its ESI-TOF mass spectrum (*m/z* 465.1878 $[M + Na]^+$, calcd. for $C_{25}H_{30}NaO_7^+$, 465.1884) (Fig. S1-A). The UV absorption pattern (Fig. S1-B) with maxima at 258, 314, and 362 nm along with the IR spectrum of **1** (broad absorption bands at around 3294 cm^{-1} for multiple hydroxyl groups, and at 1701 cm^{-1} for a conjugated carbonyl functionality) suggested that this compound possessed the dihydronaphthoquinone moiety typically present in napyradiomycin metabolites.

The ^1H NMR spectrum of **1** (Table 1 and Fig. S2) in DMSO-d_6 exhibited three deshielded signals attributable to two aromatic protons at δ_H 6.57 ppm (1H, *d*, 2.1, H-7) and δ_H 6.84 ppm (1H, *d*, 2.1, H-9), and one downfield olefinic proton signal at δ_H 6.89 ppm (1H, *d*, 6.4, H-4). This suggested the presence of a trisubstituted double bond attached to an electron-deactivating group. Signals for one olefin methine at δ_H 4.98 ppm (1H, br *t*, 7.25, H-12), three methylene protons at δ_H 2.43 ppm (2H, *m*, H-11), δ_H 1.74 ppm, δ_H 1.83 ppm (2H, *m*, H-14) and δ_H 1.31 ppm (2H, *m*, H-15), and two sp^2 methylene protons at δ_H 4.71 ppm and δ_H 4.85 ppm (2H, *br s*, H-18) were also present. On the other hand, four methyl groups in the aliphatic region at δ_H 1.61 ppm (3H, *s*, H-19), δ_H 1.27 ppm (3H, *s*, H-20), δ_H 0.88 ppm (3H, *s*, H-21), δ_H 1.23 ppm (3H, *s*, H-22), and one exchangeable OH signal (δ_H 12.52 ppm), could also be differentiated.

The ^{13}C NMR spectrum of **1** (Table 2 and Fig. S3) exhibited 25 signals: two carbonyl signals at δ 189.7 and 195.5 ppm, two phenolic carbons (δ_C 164.7 and 165.7 ppm), ten sp^2 methine or quaternary carbon signals resonating between δ_C 107.7 and 148.3 ppm, one sp^2 methylene at δ_C 110.3 ppm, one sp^3 oxygenated quaternary carbon at δ_C 82.2 ppm, two oxygenated methine carbons at δ_C 65.8 and 73.4 ppm, as well as other seven aliphatic methylene or methyl carbon signals with chemical shifts below δ_C 40.4 ppm. Summing up, according to a heteronuclear single quantum coherence spectroscopy (HSQC) experiment (multiplicity edited) (Fig. S5), the ^{13}C NMR spectrum evidenced the presence of 4 methyl, 4 methylene, 6 methine, and 11 quaternary carbons. These NMR spectroscopic data suggested that compound **1** was structurally related to the napyradiomycin family of antibiotics, more precisely to dihydronapthoquinones with a 10-carbon monoterpenoid branched side chain (NPDs A series).[237] A comparison of the NMR spectroscopic data of **1** with those of napyradiomycin A2a/A2b[247] revealed that **1** only differed from napyradiomycin A2a/A2b in the substitution at C-3, with the replacement of the chlorine atom in the latter compound by a hydroxy group in **1** and the presence of an additional olefinic bond at C-4/C-4a. This double bond was easily located since the signals for H_2-4 were lost and replaced by an sp^2 methine signal at δ_H 6.89 ppm. Consequently, the ^{13}C NMR spectrum (Table 2 and Fig. S3) showed two new olefinic carbons for C-4 and C-4a at δ_C 134.7 ppm (CH) and δ_C 137.6 ppm (C). A comprehensive analysis of 2D NMR data (Figs. 2 and S4–S7) allowed the full planar structure of **1** to be assigned, being the first member of the napyradiomycin A-series bearing a hydroxy group rather than a chlorine at position C-3 of the dihydropyran ring (Fig. 1).

	δ ^{1}H (mult, J, Hz)				
No.	1	2	3	4	5
2					
3	3.72, *d* (6.4)	4.98, *d* (1.6)	4.70, *d* (6.9)	4.38, *dd* (11.8, 4.4)	4.40, *dd* (11.8, 4.4)
4	6.89, *d* (6.4)	6.73, *d* (1.6)	7.01, *d* (6.9)	2.12, *dd* (13.4, 4.4)	2.25[a], *dd* (14.4, 4.4)
				2.19, *dd*, (13.4, 11.8)	2.52, *dd* (14.4, 11.8)
4a				6.74, s	
5					
5a					
6	12.52, *s*	12.66, *s*	12.56, *s*	11.94, *s*	
7	6.57, *d* (2.1)	6.59, *d* (2.2)	6.60, *d* (1.9)	6.66, *d* (2.0)	7.04, *d* (2.4)
8	--	--	--	--	--
9	6.84, *d* (2.1)	6.88, *d* (2.2)	6.88, *d* (1.9)	6.99, *d* (2.0)	7.18, *d* (2.4)
9a					
10					
10a					
11	2.43, *m*	1.52[a], *dd* (13.3, 3.4)	1.68, *dd* (15.2, 2.4)	1.27, br *d* (16.4, 2.3)	2.26[a], *m*
		1.85[b], dd (13.3, 3.20)	1.93, dd (15.2, 6.0)	2.32, dd (16.4, 6.0)	2.77, dd (13.5,8.9)
12	4.98, br *t* (7.25)	1.31, br *t* (3.3)	1.60[a], *dd* (6.0, 2.4)	1.55, *dd* (6.0, 2.3)	4.26, br *t* (8.9)
13		4.75, *s*	4.40, *s*	5.04, *s*	
14	1.74, *m*	1.46, *m*	1.37, *m*	1.47, *m*	1.42, *m*
	1.83, *m*	1.51[a], *m*	1.60[a], *m*	1.67, *m*	1.93[b], *m*
15	1.31, *m*	1.70, *dd* (13.5, 12.1)	1.70, *dd* (13.2, 12.1)	1.72, *m*	1.95[b], *m*
		1.81[b], *dd* (13.5, 3.8)	1.83, *dd* (13.2, 3.1)	1.82, *m*	
16	3.78, *t* (6.2, 6.2)	4.02, *dd* (12.1, 3.8)	3.77, *dd* (12.1, 3.1)	3.81, *dd* (11.9, 3.9)	4.91, br *t* (11.5, 6.5)
17					
18	4.71, br *s*	0.94[c], *s*	1.02, *s*	1.15, *s*	4.67, *s*
	4.85, br *s*				4.73, *s*
19	1.61, *s*	1.40, *s*	1.42, *s*	1.35, *s*	1.37, *s*
20	1.27, *s*	0.79, *s*	0.91, *s*	1.02, *s*	1.11, *s*
21	0.88, *s*	0.94[c], *s*	0.77, *s*	0.38, *s*	1.32, *s*
22	1.23, *s*	0.64, *s*	0.63, *s*	0.59, *s*	1.52, *s*

Table 1. ^{1}H NMR (500 MHz in DMSO-d_6) data for compounds 1-5. a, b, c overlapping signals.
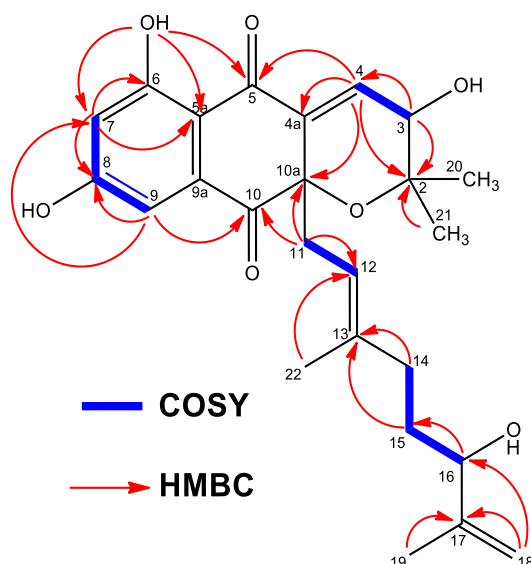
Figure 2. Key COSY and HMBC correlations observed in the spectra of compound **1**.

No.	δ ^{13}C				
	1	**2**	**3**	**4**	**5**
2	75.4, C	76.1, C	74.8, C	79.1, C	78.2, C
3	65.8, CH	59.9, CH	57.3, CH	58.9, CH	59.8, CH
4	134.7, CH	136,2 CH	132.1, CH	40.4, CH$_2$	41.6 CH$_2$
4a	137.6, C	136.3, C	137.8, C	79.2, C	81.7, C
5	189.7, C	187.7, C	188.3, C	194.4, C	183.4, C
5a	110.8, C	110.5, C	110.4, C	107.9, C	116.7, C
6	164.7, C	164.8, C	164.8, C	163.9, C	162.7, C
7	107.8, CH	108.1, CH	108.1, CH	108.5, CH	113.5, CH
8	165.7, C	165.3, C	165.6, C	165.7, C	163.8, C
9	107.7, CH	107.7, CH	107.9, CH	107.8, CH	108.3, CH
9a	136.5, C	135.9, C	135.9, C	135.1, C	136.4, C
10	195.5, C	195.3, C	194.9, C	199.5, C	196.0, C
10a	82.2, C	82.9, C	82.6, C	83.7, C	82.9, C
11	40.4, CH$_2$	37.0, CH$_2$	37.8, CH$_2$	33.7, CH$_2$	41.5, CH$_2$
12	117.7, CH	50.5, CH	50.1, CH	48.6, CH	116.9, CH
13	138.9, C	70.5, C	70.1, C	70.1, C	140.6, C
14	35.5, CH$_2$	41.2, CH$_2$	41.3, CH$_2$	41.2, CH$_2$	39.7, CH$_2$
15	32.9, CH$_2$	31.1, CH$_2$	30.8, CH$_2$	30.5, CH$_2$	23.7, CH$_2$
16	73.4, CH	72.2, CH	72.4, CH	71.8, CH	126.5, CH
17	148.3, C	40.4, C	40.4, C	40.2, C	129.2, C
18	110.3, CH$_2$	20.5, CH$_3$	25.8, CH$_3$	21.6, CH$_3$	76.3, CH$_2$
19	17.8, CH$_3$	26.5, CH$_3$	26.5, CH$_3$	28.7, CH$_3$	28.9, CH$_3$
20	24.9, CH$_3$	23.6, CH$_3$	24.2, CH$_3$	24.4, CH$_3$	22.5, CH$_3$
21	25.7, CH$_3$	30.1, CH$_3$	29.0, CH$_3$	28.4, CH$_3$	15.2, CH$_3$
22	16.1, CH$_3$	16.6, CH$_3$	16.7, CH$_3$	16.0, CH$_3$	15.1, CH$_3$

Table 2. ^{13}C NMR (125 MHz in DMSO-d_6) data for compounds **1-5**.

The relative stereochemistry of the dihydropyran ring of compound **1** (Fig. 3) was assigned by interpretation of Rotating frame Overhauser Enhancement SpectroscopY (ROESY) data (Fig. S7) and the coupling constants observed in its ^{1}H NMR spectrum (Table 1, Fig. S2) in combination with molecular modeling using Chem3D 12.0 Pro. The biosynthetic route for all napyradiomycins described to date was also considered.[236,242,243] The almost equally intense ROESY correlations observed between H-3 and both geminal methyl groups (C-20 and C-21) together with a coupling constant value of 6.4 Hz between protons H-3 and H-4 (Fig. 3), in good agreement with a dihedral

56

angle of 28.7° measured in the energy-minimized molecular model (Fig. S8), confirmed the relative configuration at C-3 of the dihydropyran ring in **1**. Assuming an *R* absolute configuration at C-10a based on the common biosynthetic origin described for the napyradiomycin A series described to date, the absolute configuration at C-3 was also proposed to be *R*. Finally, the low amount of **1** obtained prevented the determination of the absolute configuration of the chiral center at C-16 using Mosher´s approaches.
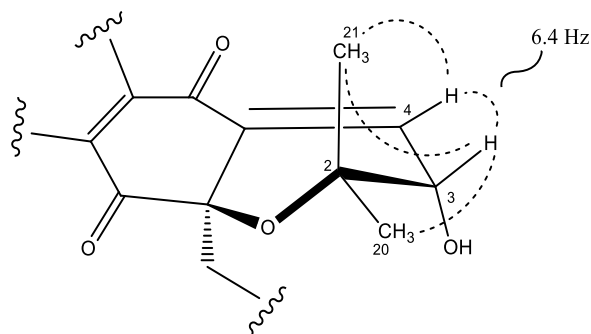
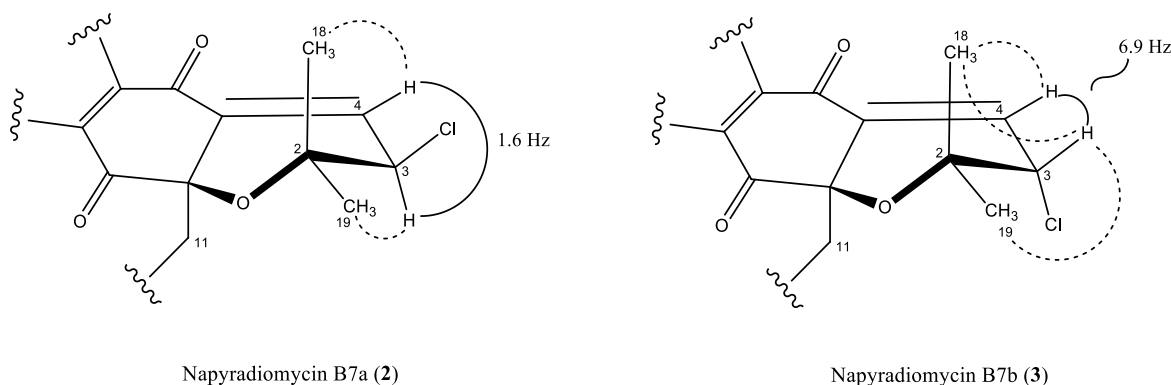**Figure 3.** Key ROESY correlations (dashed lines) and coupling constant that determine the relative configuration of the dihydropyran ring in napyradiomycin A3 (**1**).

Compounds **2** and **3** were isolated as white powders and their molecular formula was determined to be $C_{25}H_{30}Cl_2O_6$ based on HRMS measurements (Figs. S9-A, S18-A) (ESI-TOF-MS *m/z* 479.1389 [M + H − H_2O]^+ for compound **2** and 479.1394 [M + H − H_2O]^+ for compound **3,** calcd. for $C_{25}H_{29}{}^{35}Cl_2O_5{}^+$, 479.1387). This formula requires 10 degrees of unsaturation. Their UV, IR, and NMR spectroscopic data were almost identical and shared common features. The UV absorption pattern (Figs. S9-B, S18-B) along with the IR spectrum of both (broad absorption bands at around 3294 cm^{-1} for multiple hydroxyl groups, and at 1701 cm^{-1} for a conjugated carbonyl functionality) suggested that these compounds also possessed the dihydronaphthoquinone moiety present in napyradiomycin metabolites. Comprehensive analysis of 1D and 2D NMR data of compounds **2** (Tables 1 and 2; Figs. S10–S16) and **3** (Tables 1, 2; Figs S19–S25) revealed a strong similarity between them and evidenced that they were epimers at C-3 position of the dihydropyran ring.

Interpretation of 2D NMR data of compound **2** allowed all the protons and carbons to be assigned and according to an HSQC experiment (multiplicity edited) (Fig. S13), we distinguished the presence of 5 methyl, 3 methylene, 6 methine, and 11 quaternary carbons. These NMR spectroscopic data suggested that compound **2** was structurally related to dihydronapthoquinones with the monoterpenoid substituent cyclized to a 6-membered ring (NPDs B series).[4] The dihydronaphtoquinone ring was constructed based on Heteronuclear Multiple Bond Correlation (HMBC) correlations (Figs. 4, S14) from H-9 (δ_H 6.88 ppm) to C-5a, C-7, C-8, and C-10, and from OH-6 (δ_H 12.66 ppm) to C-5a, C-6, and C-7. In addition, correlations in the HMBC experiment between the two geminal methyl groups H$_3$-18 and H$_3$-19 (δ_H 0.94 and δ_H 1.40 ppm) and C-2 and C-3, and between the olefinic proton H-4 at δ_H 6.73 ppm and C-2, C-3, C-4a, and C-10a indicated the presence of a dihydropyran ring and confirmed the position of a chlorine substituent at C-3 (δ_H 4.98, s; δ_C 59.9, CH) and a trisubstituted double bond at $\Delta^{4,4a}$. Analysis of the overall NMR data set (Tables 1 and 2; Figs. S10–S16) for the monoterpene unit (C-11 to C-17 and C-20, C-21, and C-22) in **2** showed that it was a cyclohexane ring with a chair conformation. Based on these NMR features, the planar structure of **2** was established (Fig. 4). A similar analysis of the NMR data set (Tables 1 and 2; Figs. S19–S25) for compound **3** rendered the same planar structure and evidenced the close structural similarity of **2** and **3** with the previously reported antibiotic CNQ525.510A, which is additionally methylated at C-7.[236]
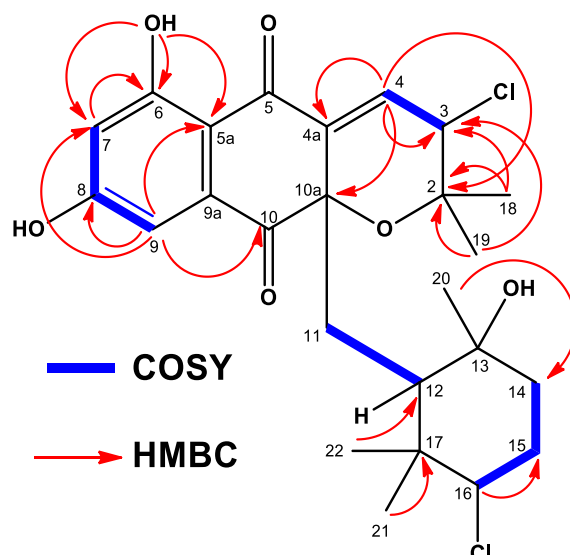
Figure 4. Key Correlated Spectroscopy (COSY) and HMBC correlations observed in the spectra of compounds **2** or **3**.

The relative stereochemistry of **2** and **3** was assigned by analysis of NOESY and ROESY NMR data (Figs. S15, S16, S24, and S25) and the coupling constants observed in their ^{1}H NMR spectra. Nuclear Overhauser Effect Spectroscopy (NOESY) correlations between the H$_3$-21 methyl protons and the methine protons H-12 and H-16 showed these protons were on the same face of the cyclohexane ring in both compounds (Fig. 5). A ROESY correlation between the H$_3$-20 and H$_3$-22 methyl groups determined their axial orientation on the bottom face of the ring. Finally, based on the existence of a strong ROESY correlation, H-12 and H-16 were positioned in a 1,3-diaxial position on the top face of the ring (Fig. 5)
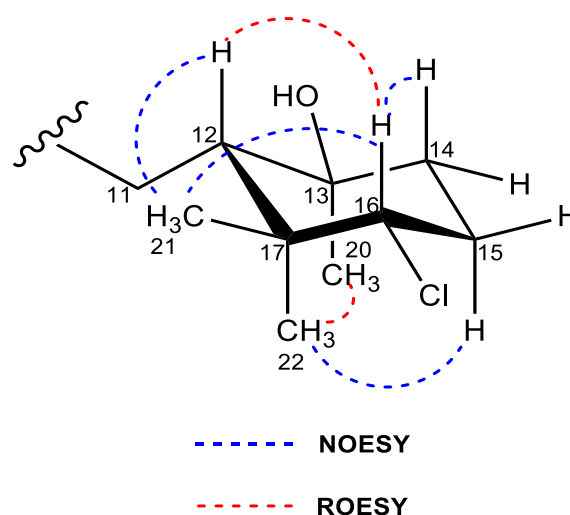
Figure 5. Key NOESY/ROESY correlations that determine the relative configuration of the cyclohexane ring in compounds **2** and **3**.

Regarding the dihydropyran ring, we used the vicinal $^3J_{HH}$ spin–spin coupling constants and key NOESY correlations to establish the relative configuration of this moiety. The coupling constant between the olefinic proton at H-4 and its vicinal proton H-3 in **2** had a value of 1.6 Hz (Fig. 6), which is in good agreement with a dihedral angle of 91.4° measured in the energy-minimized molecular model (Fig. S17). Furthermore, only one NOESY correlation is observed between H-3 and one of the geminal methyl groups (H$_3$-19), due to the antiperiplanar position of

the H$_3$-18 methyl with respect to H-3 (Figs. 6 and S15). Thus, the relative configuration on the dihydropyran ring for compound **2** was confirmed as depicted in Figure 1. Conversely, the almost equally intense and strong *nOe* correlations observed between H-3 and both geminal methyl groups (C-18 and C-19) in the spectra of **3** (Figs. 6 and S24), together with a coupling constant value of 6.9 Hz between protons H-3 and H-4 in accordance with a dihedral angle of 25.2° measured in the energy-minimized molecular model (Fig. S26), confirmed the opposite relative configuration at C-3 for compound **3**.

Napyradiomycin B7a (**2**)   Napyradiomycin B7b (**3**)

Figure 6. Key NOESY correlations (dashed lines) and coupling constants that determine the relative configuration of the dihydropyran ring in compounds **2** and **3**.

The absolute configurations of **2** and **3** were assumed to be the same as those reported for all the napyradiomycins in the B series. Apart from the common biosynthetic origin, this proposal was additionally supported by the comparison of specific rotation data. The sign of the specific rotation for compound **2** is negative ([α]$_D^{25}$ -41.0°, *c* 0.4, MeOH), as is that reported for CNQ525.510A,[236] which is a C-7 methylated version of **2**. On the contrary, the sign of [α]$_D^{25}$ value for the epimeric compound **3** is positive ([α]$_D^{25}$ +32.0°, *c* 0.13, MeOH). The previously reported B-type napyradiomycin MDN-0170[241] displays the same relative configuration at C-3 (but bearing a hydroxy group instead of a chlorine atom), and showed a positive value of specific rotation, providing evidence of the same absolute configuration.

Compound **4** was obtained as a white powder. The complex pattern of the ion clusters in the adducts of its ESI mass spectrum (Fig. S27-A) indicated the presence of two chlorine atoms in the molecule. ESI-TOF-MS analysis (Fig. S27-B) suggested the molecular formula C$_{25}$H$_{32}^{35}$Cl$_2$O$_7$ (*m/z* 497.1498 [M + H − H$_2$O]$^+$, calcd. for C$_{25}$H$_{31}^{35}$Cl$_2$O$_6^+$, 497.1492), indicating nine degrees of unsaturation. Analysis of combined ^{1}H and ^{13}C NMR spectroscopic data (Tables 1 and 2; Figs. S28–S33) showed signals similar to those of napyradiomycin metabolites. The presence of a tetrahydropyran ring fused to the dihydronaphthoquinone moiety in **4** was evidenced by the absence of the olefinic proton signal at C-4 present in compounds **1–3**, now replaced by two new diastereotopic protons at δ_H 2.12 and 2.19 at that position. Furthermore, the presence of a hydroxy substituent at C-4a, easily assigned based on HMBC correlations (Fig. S32) of that OH-4a at δ_H 6.74 with C-4, C-4a, C10, and C-10a that corroborates the presence of this tetrahydropyran ring. Comprehensive NMR analyses allowed all protons and carbons to be assigned, and compound **4** was confirmed to be napyradiomycin SC, whose structure was incompletely reported by Kamimura and co-workers in a Japanese patent in 1997.[264] Details on how the structure was assigned, spectroscopic data, and the absolute stereochemistry of compound **4** were never reported.

The relative stereochemistry of **4** was assigned by analysis of NOESY data (Figs. 7 and S33). Correlations of the protons on the cyclohexane ring were identical to those observed in the NOESY/ROESY experiments for **2** and **3**. Key NOESY correlations between the exchangeable OH signal at C-4a with both H-11 protons and the methine H-12 indicated that the tetrahydropyran

ring was *cis*-fused to the dihydroquinone, as observed for other napyradiomycins. A correlation observed between the OH-4a hydroxy group and the methine H-3 indicated that these were both oriented on the bottom face of the tetrahydropyran ring. The typical axial-equatorial coupling constants measured between H-3 (J = 11.8, 4.4 Hz) and the methylene proton pair H₂-11 are in good agreement with dihedral angles of –172.6° and –55.3° measured in the energy-minimized molecular model (Fig. S34) and confirm that the tetrahydropyran ring was in a chair form, identical to the configuration of this ring in the crystal structure of napyradiomycin B4,[238] a version of compound **4** bearing a chlorine instead of a hydroxy group at C-4a. The full stereostructure of napyradiomycin B4 was assigned by X-ray diffraction methods[238] and hence, the absolute stereochemistry is assumed identical to that of napyradiomycin B4 based on comparable specific rotation values. The sodium D line-specific rotations of compound **4** ([α]$_D^{25}$ -43.0°, *c* 0.4, MeOH) and napyradiomycin B4 ([α]$_D^{25}$ -190°, *c* 0.031, CHCl₃) are both negative, providing evidence that both compounds should possess the same absolute configuration.
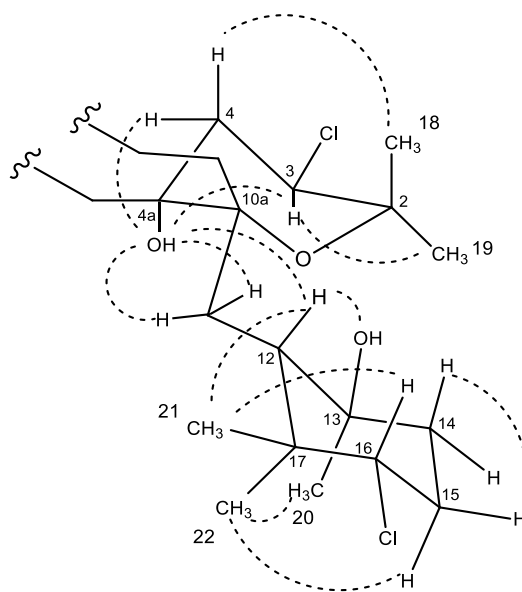
Figure 7. Conformation in solution of compound **4** based upon NOESY analysis.

Compound **5** was obtained as a white powder and was assigned the molecular formula C₂₅H₂₈Cl₂O₅ by evaluation of HRMS data (Fig. S35-A) (ESI-TOF-MS *m/z* 479.1384 [M+H]⁺, calcd. for C₂₅H₂₉³⁵Cl₂O₅⁺, 479.1387). The overall analysis of its NMR spectroscopic data (Tables 1 and 2; Figs. S36–S41) clearly showed that the compound possesses the characteristic dihydronaphthoquinone moiety of napyradiomycins, fused to a tetrahydropyran ring and bearing a linear monoterpenoid attached to its C-10a carbon. The ¹H NMR spectrum of **5** (Table 1, Fig. S36) exhibited, among other signals, four singlet methyl groups (δ_H 1.11, 1.32, 1.37, and 1.52 ppm), two doublets from aromatic protons (δ_H 7.04 and 7.18 ppm), one methine proton geminal to a chlorine atom (δ_H 4.40 ppm) and two signals from olefinic methine groups (δ_H 4.26 and 4.91 ppm). Interestingly, compound **5** lacks the characteristic deshielded singlet signal of OH-6. HMBC correlations (Fig. S40-A) from H-3 (δ_H 4.40 ppm) to C-19 / C-20, C-2, and C-4a, and from H₂-4 (δ_H 2.52, 2.25 ppm) to C-2, C-3, C-4a, C-5, and C-10a defined the presence of the archetypal tetrahydropyran ring with two chlorine substituents at C-3 and C-4a positions reported for other napyradiomycins such as 18-hydroxynapyradiomycin A1 (**10**),[247] napyradiomycin A2a (**11**),[247] napyradiomycin A2b (**12**),[247] or napyradiomycin B4 (**13**).[238] COSY NMR spectroscopic data (Fig. S38) allowed the identification of three key proton spin systems within the monoterpenoid moiety: H₂-11/H-12, H₂-14/H₂-15/H-16, and H₂-18 (Fig. 8). These three proton sequences were connected as a linear monoterpenoid side chain by interpretation of HMBC correlations (Fig. 8 and S40) from H₃-21 to C-12, C-13, and C-14, and from H₃-22 to C-16, C-17, and C-18. As expected, additional

HMBC correlations from both H_2-11 and H-12 established the attachment of this side chain to C-10a of the dihydronaphthoquinone moiety (Fig. 8). All these assignments accounted for ten of the eleven degrees of unsaturation and indicated that **5** was composed of a tetracyclic ring system, evidencing that **5** is closely related to napyradiomycin C1 and other compounds within the C series.[240,263]

The chemical shift of the methylene carbon signal C-18 (δ_C 76.3 ppm) strongly supported the presence of an oxygen substituent at this position. Moreover, the deshielding of this carbon signal in comparison with that found for 18-hydroxynapyradiomycin A1 (**10**) (δ_C 66.6 ppm) suggested the etherification of the 18-hydroxy groups in **5**. Finally, the strong key HMBC correlation from H_2-18 to C-6 (δ_C 162.7 ppm; Fig. S40-B) allowed us to establish unambiguously the existence of such ether link between C-6 and C-18, which results in a 14-membered bridging macrocycle between C-6 and C-10a (Fig. 8). This kind of O-linked cyclization has no precedent within the napyradiomycin metabolites, and therefore compound **5** represents the first member of a new subfamily of napyradiomycins, the D series. Thus, we propose the name napyradiomycin D1 for compound **5**.
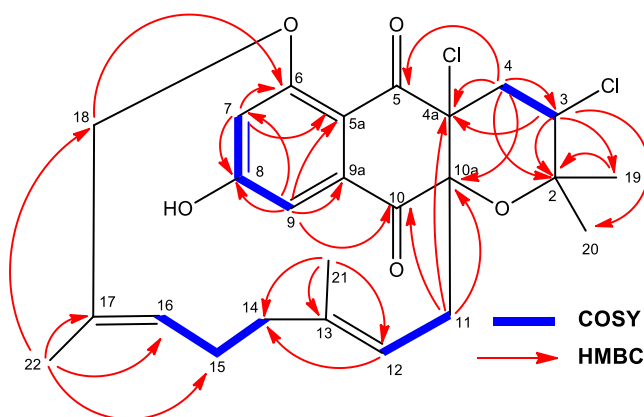
Figure 8. Key COSY and HMBC correlations were observed in the spectra of compound **5**.

The relative configuration of **5** was determined by inspection of the NOESY correlations and multiplet analysis for some key proton signals (Fig. 9). As in napyradiomycin SC (**4**), the values of the coupling constants between H-3 and the pair H_2-4 (J = 11.8, 4.4 Hz) confirmed a chair conformation of the tetrahydropyran ring and the axial orientation of H-3. Although the substitution at C-4a with a chlorine atom in **5** prevents establishing a relative configuration concerning C-10a, it was assumed to be *cis*, as for all tetrahydropyran-containing napyradiomycins described to date. The geometry of the two double bonds in **5** was assigned as *E* based on the existence of NOESY cross-peaks between H_2-11/H_3-21/H_2-14 and H-15/H_3-22/H_2-18, as well as the absence of correlations between H-12/ H_3-21 and H-16/H_3-22 (Fig. 9-a). The absolute configuration of **5** is assumed to be the same as for napyradiomycin C1 considering the comparable specific rotation values and the common biosynthetic origin.[240]
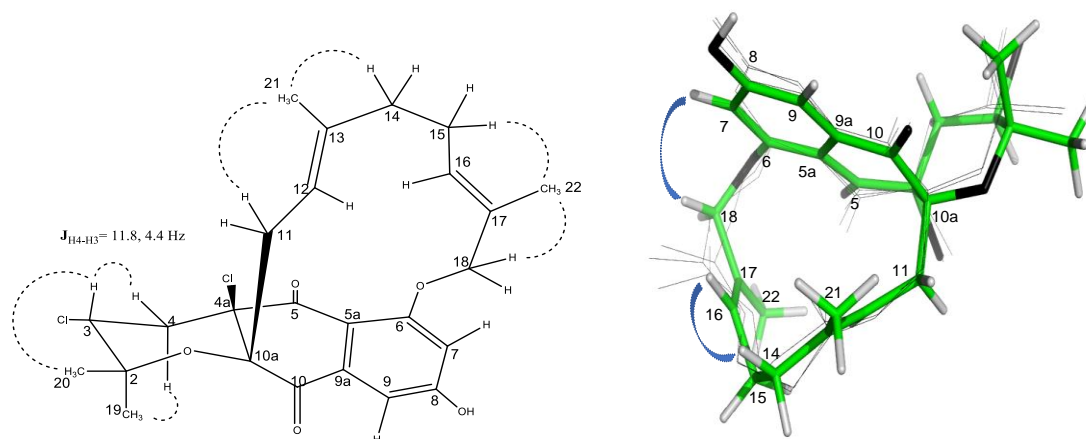
Figure 9. (**a**) Key NOESY correlations (dashed lines) supporting the relative configuration of napyradiomycin D1 (**5**). (**b**) Minimum-energy conformer for **5** and consistent *nOe* cross-peaks.

Apart from the unambiguous NOESY correlations described above, other sets of cross-peaks were observed for proton signals within the linear monoterpenoid chain (Fig. S41). The slight broadening of these signals in the ^{1}H NMR spectrum and the presence of such different sets of *nOe* correlations point to a fast (in NMR time-scale) conformational equilibrium in solution **5**. Not surprisingly, this interconversion might be assisted by the flexibility of the terpenoid chain due to the presence of seven rotatable bonds within it. To better illustrate this, a conformational search using *ConfBuster Web Server*, a recently delivered open-source tool for the conformational analysis of macrocycles was launched.[265] As a result, up to four close-energy conformers were obtained (Fig. S42), differing from each other by about 1.5 kcal/mol, which supports the rapid interconversion between them. Interestingly, the minimum-energy conformer proved to be consistent with the most abundant one, as evidenced by the more intense set of NOESY cross-peaks between H_2-18 (δ_H 4.67 ppm) and the aromatic proton at H-7 (δ_H 7.04 ppm), and between the olefinic proton H-16 (δ_H 4.91 ppm) and H_2-14 (δ_H 1.42 ppm) (Figure 9b).

2.4. Evaluation of Antimicrobial Activity - Antibacterial, Antifungal, and Cytotoxic Activities

Compounds **1–15** were evaluated for their antibacterial and antifungal properties against a clinical isolate of methicillin-resistant *Staphylococcus aureus* (MRSA), *Mycobacterium tuberculosis, Escherichia coli*, *Acinetobacter baumannii,* and *Aspergillus fumigatus* (Table 3). Napyradiomycins **2**, **5**, **7**, **12**, **13**, **14,** and **15** showed antibacterial activities against MRSA with MIC values ranging from 3 to 48 µg/mL. The new napyradiomycin D1 (**5**) was one of the most active compounds and displayed activities comparable to those of napyradiomycin B4 (**13**) and napyradiomycin B5 (**15**). Napyradiomycin A2b (**12**) and napyradiomycin B2 (**14**) exhibited the best antibacterial activities (MIC values of 3–6 µg/mL) among these fifteen napyradiomycins. Except for compounds **3**, **6**, **8,** and **10**, the other ten napyradiomycins isolated showed moderate activity against the Gram-positive bacteria *M. tuberculosis* H37Ra with MIC values fluctuating from 12 to 48 µg/mL. None of the compounds exhibited activity against the Gram-negative bacteria *E. coli* ATCC 25922 or *A. baumannii* MB5973 nor the fungus *A. fumigatus* ATCC 46645.

Eight napyradiomycins, **2**, **5**, **8**, **11**, **12**, **13, 14,** and **15** showed moderate cytotoxic activities with IC_{50} values below 50 µM against the human liver adenocarcinoma cell line (HepG-2), whereas the other seven had reduced cytotoxicity, with IC_{50} values above this concentration.

	MIC (µg/mL)					IC$_{50}$ (µM)
	MRSA	*Mt*	*Ec*	*Ab*	*Af*	**HepG-2**
1	>96	NT[a]	>96	>96	NT[a]	>67.8
2	48	12-24	>96	>96	>96	41.7
3	>64	>64	>64	>64	>64	109.5
4	>96	24-48	>96	>96	>96	263.5
5	12-24	24-48	>96	>96	NT[a]	14.9
6	>96	>96	>96	>96	>96	277.2
7	48-96	12-24	>96	>96	>96	186.9
8	>64	>64	>64	>64	>64	30.2
9	48-96	12-24	>96	>96	>96	71.2
10	48-96	>96	>96	>96	>96	64.4
11	12-24	48-96	>96	>96	>96	30.4
12	12-24	48-96	>96	>96	>96	28.6
13	12-24	12-24	>96	>96	>96	15.6
14	3-6	24-48	>96	>96	>96	27.1
15	12-24	24-48	>96	>96	>96	40.1

Table 3. Antibacterial, antifungal, and cytotoxic activities of compounds 1–15. Note: MRSA, Methicillin-resistant *Staphylococcus aureus* MB5393; Mt, *Mycobacterium tuberculosis* H37Ra; Ec, *Escherichia coli* ATCC 25922; Ab, *Acinetobacter baumannii* MB5973; Af, *Aspergillus fumigatus* ATCC 46645. [a] NT = not tested.

3. Discussion

The napyradiomycins constitute a large class of unique halogenated meroterpenoids, with around fifty members reported to date, produced by marine and terrestrial *Streptomyces* species.[234–238,246–253] The biosynthetic gene cluster (BGC) of napyradiomycins was first described from *Streptomyces aculeolatus* NRRL 18422 and the marine-derived *Streptomyces sp.* CNQ-525, and its analysis established the link between the presence of three vanadium-dependent haloperoxidases (VHPOs), (NapH1, NapH3, and NapH4) and a chloronium-induced meroterpene cyclization pathway.[242] The full biosynthetic route to these metabolites from three precursors (1,3,6,8-tetrahydroxynaphthalene, dimethylallyl pyrophosphate, and geranyl pyrophosphate), has been recently described and highlights the key role of those VHPO enzymes.[243]

The final structures resulting from this biosynthetic pathway, the napyradiomycins, are hybrid terpenoid/polyketide metabolites composed of a semi-naphthoquinone chromophore, a prenyl unit attached to C-4a which is cyclized to produce a tetrahydropyran or dihydropyran ring and a monoterpenoid subunit attached at C-10a, which in turn can be either linear (type A napyradiomycins) or cyclized to 6-membered (type B) or to 14-membered (type C) rings. The different halogenation patterns add complexity and contribute to structural variations of these interesting metabolites.[234–238,246–253] A further proof of the still surprising structural possibilities for these natural products is the herein reported isolation of four new napyradiomycins, A3 (**1**), B7a (**2**), B7b (**3**), and D1 (**5**), showing unusual substitution patterns, the inverted configuration of some chiral centers, or unprecedented cyclic bridging links.

Compound **1** is to the best of our knowledge the first example of a napyradiomycin in the A series bearing a hydroxy group instead of a chlorine atom at position C-3 of the dihydropyran ring. This substitution pattern has been previously reported for napyradiomycins of types B (MDN-0170)[241] and C,[240,263] further displaying the same relative configuration as **1** at this chiral center. The presence of the C-3 hydroxy group with this absolute configuration can be explained considering the precursor of **1** might be the corresponding chlorinated compound at the same position, and that a non-enzymatic S$_N$2 nucleophilic substitution with water on the C-3 chloride would result in the production of **1**.

Compounds **2** and **3** are epimers at C-3 of their dihydropyran ring. The relative configuration at this chlorinated position for compound **3** is reported herein for the first time in the napyradiomycin series since all the natural products of this family found in the literature or databases have the opposite configuration at the same chiral center (i.e., that found for compound **2**) when it has a chlorine substituent. A first tentative explanation for this variant arises from the proposed mechanism of oxidative halogenation and subsequent halonium-induced cyclization in meroterpenoids.[242] Although the vanadium-dependent chloroperoxidase (VCPO) NapH1 has been shown to act in a stereoselective fashion when introducing chlorine atoms in napyradiomycin biosynthetic intermediates, it was also found that the same enzyme catalyzed a non-stereoselective bromination of the same substrate.[266] This result was explained because of the production of a diffusible hypobromous acid, which would depart the active site of the enzyme and then brominate the substrate in a nonspecific manner. Indeed, the involvement of a hypohalous acid (HOX species) in this mechanism is widely accepted for vanadium-dependent haloperoxidases (VHPOs) from algae and fungi, which do not exhibit specificity,[267] while for VCPOs from *Streptomycetes*, it has been postulated that an enzyme-bound chlorine species would make possible the stereoselective halogenation.[268,269]

The isolation herein of both diastereomeric versions of the same chlorinated product (compounds **2** and **3**) could suggest the participation of hypochlorous acid-mediated chlorination along with the enzyme-assisted mechanism. However, the fact that this nonspecific chlorination has not been observed for any other napyradiomycin derivatives isolated in this work, i.e., from the same culture and therefore under the same conditions, did not support this overall hypothesis.

The other possibility is that the chlorine atom at C-3 in compound **3** may come from a S_N2 displacement, therefore with inversion of configuration. The only reasonable way for that to occur is that a chloride ion is displacing the C-3 chlorine in compound **2**. Although it is known that chloride is a poor nucleophile (and a fair leaving group), the activation of this allylic position within the dihydropyran ring could provide enough driving force for the reaction to proceed. Despite this epimerization has not been observed for any other dihydropyran ring-containing napyradiomycin (isolated in this work or not), we consider that the nucleophilic substitution to produce compound **3** may be the most plausible hypothesis.

Finally, compound **5** presents in its structure a 14-membered macrocycle because of a bridging ether link between C-6 of the dihydronaphtoquinone unit and C-18 of the monoterpenoid chain. This kind of ether link is unprecedented within the napyradiomycin family, and therefore compound **5** represents the first member reported of a new series of napyradiomycins that we have designated as napyradiomycin D1. Remarkably, although a similar 15-membered cyclic ether ring has been previously reported for merochlorin C —a compound belonging to another family of meroterpenoid metabolites[270] — the *O*-linked bridging observed in **5** still represents, to the best of our knowledge, one of the largest ether cyclization events observed for a natural product. Regarding its biosynthesis, since NapH1 has been proved responsible for the halogenation and the formation of a 6-membered cyclic ether ring in napyradiomycins,[266] it is tempting to think that the same enzyme could also catalyze this macrocyclization. However, the attachment in napyradiomycin D1 of the ether linkage to a former methyl group without the installation of any chiral chlorinated center vicinal to the linking position is inconsistent with a chloronium-induced cyclization catalyzed by NapH1. The whole genome sequencing of the producer strain CA-271078 and the inspection of its napyradiomycin BGC may further provide valuable data about this, and other biosynthetic questions raised in this work (see Box 1).

Although the specific mechanism of action for this family of meroterpenoids is not clear,[239,246,254] previous studies about the structure-activity relationship (SAR) have shown that structural variations among the napyradiomycin metabolites scaffold known so far, such as the different halogenation patterns or the presence or absence of the methyl group at C-7 among others, can attenuate or enhance their biological activities.[237,254]

Some of the compounds isolated displayed antibacterial activity against MRSA. Notably, the activity of compound **5**, the first in the napyradiomycin D series, was comparable to that of other

known compounds isolated in this work. In line with previous results, none of the compounds isolated were found to be active against Gram-negative bacteria or fungi. Additionally, our cytotoxicity data illustrate that, for the napyradiomycin B series, the C-3 chlorinated derivatives (compounds **2** and **3**) exhibit a higher cytotoxic activity compared to the hydroxylated analogous at that stereocenter (compound **6**). This cytotoxic is in turn significantly influenced by the absolute stereochemistry of the chlorine group at that position (compound **2** vs. **3**). On the other hand, there is a clear correlation between rising levels of cytotoxicity when the substitution pattern at C-4a varies between hydroxy group (compound **4**), hydrogen (compound **9**)**,** or chlorine (compound **13**).

4. Conclusions

In this work, we report the isolation and structural characterization of four new napyradiomycins (**1**–**3, 5**) and the known napyradiomycin SC (**4**), whose structural details had not been previously described. Additionally, another ten known napyradiomycins or related compounds (**6–15**) were also isolated from the same culture broth of the marine-derived *Streptomyces* sp. CA-271078 from MEDINA´s microbial collection. The antibacterial, antifungal, and cytotoxic properties of all the compounds isolated were tested. Napyradiomycins B2 (**14**), B4 (**13**), and B5 (**15**) and the new napyradiomycin D1 (**5**) were the most active compounds, exhibiting similar antibacterial activities against MRSA and *M. tuberculosis* H37Ra, as well as comparable cytotoxic activities against the HepG2 tumoral cell line. On the contrary, none of the compounds tested showed significant activity against *E. coli*, *A. baumannii,* or *A. fumigatus.*

The four newly compounds isolated displayed remarkable structural features. Compound **1** is the first member of napyradiomycins in the A series bearing a hydroxy group rather than a chlorine atom at C-3. The isolation of the new B-type napyradiomycin **3** represents to our knowledge the first example reported with a different relative configuration at the C-3 chlorinated position in napyradiomycins and raises biosynthetic and mechanistic questions to be considered. Compound **5** harbors in its structure an unprecedented 14-membered macrocyclic ether ring between the naphthoquinone moiety and the monoterpenoid chain, thus inaugurating a new class of napyradiomycins, the D series.

The results reported here highlight the wide range of structural possibilities for napyradiomycin metabolites. The further whole genome sequencing of the producer strain CA-271078 and the analysis of the corresponding napyradiomycins gene cluster could bring new data about the biosynthesis of these fascinating natural products.

Box 1. Biosynthesis of Napyradiomycins

To complement the work described in this chapter and try to gain insight into the biosynthesis of the novel napyradiomycin D1 (**5**), newly isolated in this work, whole-genome sequencing was applied to the producing strain CA-271078 at the Center for Biosustainability (DTU-Biosustain) in Copenhagen, Denmark. The BGC putatively responsible for the biosynthesis of napyradiomycins, herein referred to as *nap*, was identified in the genome of the producing strain by using the bacterial version of the antiSMASH software (v. 7.0)[271] with the default 'relaxed strictness' settings (Fig. 10). A total of 30 biosynthetic regions were predicted, one of them (region 30, with an overall size of 42 kbp) showing a high similarity (90%) to a known cluster responsible for producing napyradiomycin analogues in *Streptomyces aculeolatus* NRRL 18422.[242]
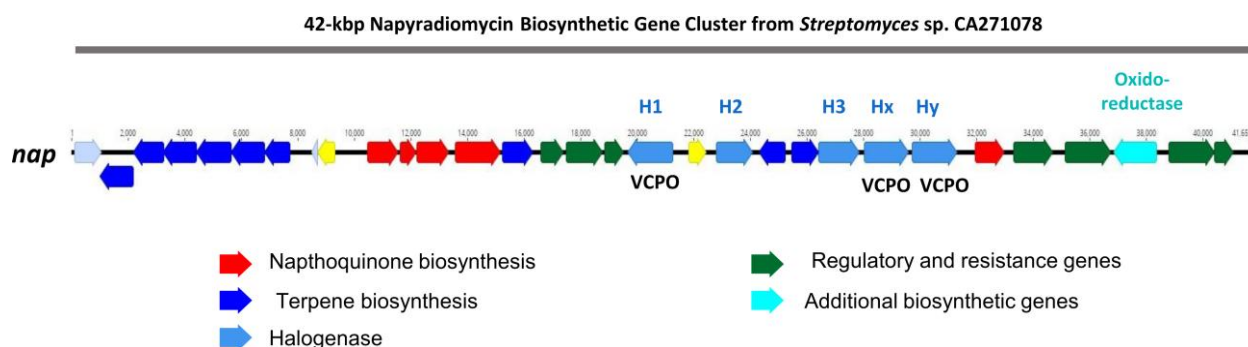
Figure 10. Organization of the *nap* biosynthetic gene cluster in *Streptomyces* sp. CA-271078 and putative gene functions.

The putative functions of genes in the *nap* cluster were determined by sequence comparisons with the only two napyradiomycin BGCs publicly available at NCBI: *Streptomyces aculeolatus* NRRL 18422 and the marine sediment-derived *Streptomyces sp.* CNQ-525.[242] These two available clusters are nearly identical, as they exhibit a similar organization and share 97% nucleotide identity. Hence, the CA-271078 BGC would be the third napyradiomycin-producing biosynthetic gene cluster reported to date.

As mentioned earlier in this chapter, previous comprehensive works on the biosynthetic pathway of these meroterpenoids underscored the pivotal role played by three vanadium-dependent chloroperoxidase enzymes (VCPOs), NapH1, NapH3, and NapH4, which are correlated to a chloronium-induced cyclization pathway. These enzymes are a specific subset of vanadium-dependent haloperoxidases and are instrumental biocatalysts for the enzymatic construction of these natural products (Fig. 11).[243]
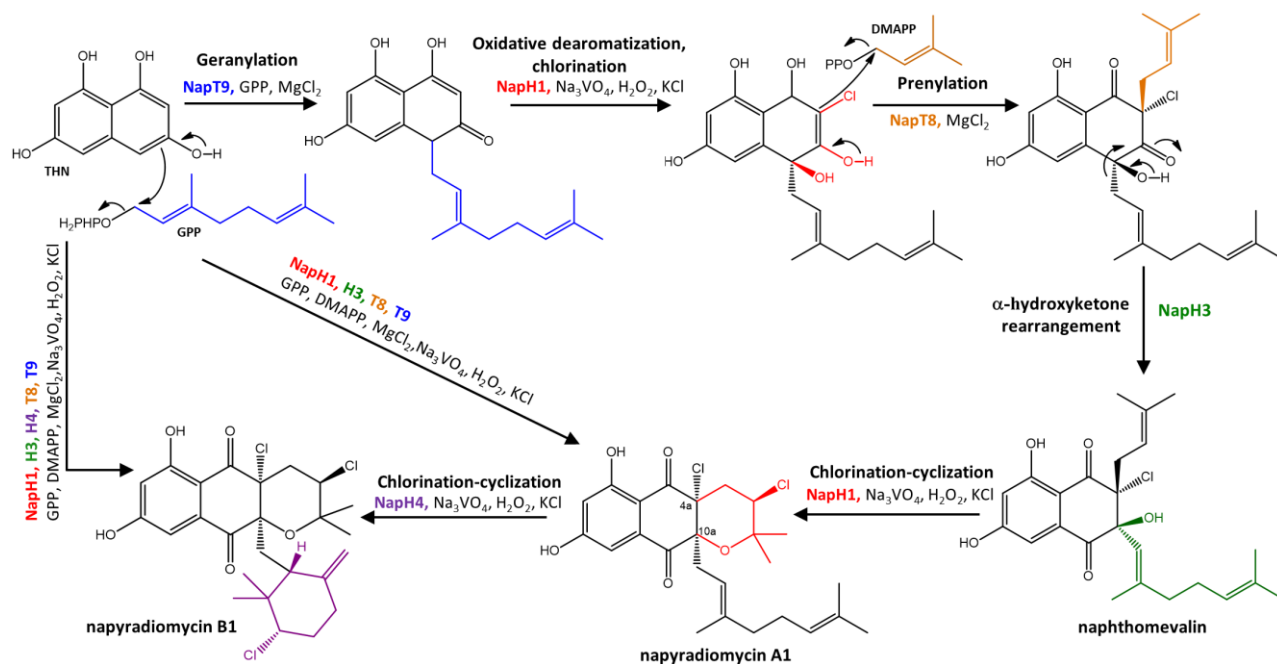
Figure 11. Fully elucidated biosynthetic pathway to napyradiomycin B1 from THN. The napyradiomycins are derived from three precursors: a symmetrical pentaketide 1,3,6,8-tetrahydroxynaphthalene (TNH) and two isoprenyl derivatives, the five-carbon dimethylallyl pyrophosphate (DMAPP) and the ten-carbon geranyl pyrophosphate (GPP). The simplest member of the family is naphthomevalin, which features a naphthoquinone core adorned with prenyl and chlorine substituents at C-4a, and geranyl and hydroxyl groups at C-10a. Figure adapted from McKinnie et al., 2019[243] and Murray et al., 2020.[272]

The majority of marine meroterpenoids are directly derived from THN. Following the initial prenylation of THN, the resulting metabolite undergoes successive processes involving oxidation, halogenation, and cyclization. Some notable examples are the napyradiomycins, described in this chapter, and the merochlorins family, which possesses a methyl substituent at C-4a instead of the prenyl substituent observed in napyradiomycins at the same position. Merochlorins were the first THN-based meroterpenoids to have their biosynthesis fully elucidated[270] and much of the logic established in merochlorin biosynthesis, in terms of reactions and chemical structures, applies to the napyradiomycins.[272] For example, merochlorin BGC (*mcl*) also features genes encoding the rare halogenating enzymes VHPOs (Mcl24 and Mcl40), first discovered in napyradiomycins.[242]

Among the merochlorin family, the particularly unusual 15-membered macrocyclic ether ring of merochlorin C resembles that present in the newly discovered napyradiomycin D1 (**5**) as in both cases the macrocycle ring is formed between a phenol moiety of dihydronaphtoquinone core and the monoterpenoid chain. It has been proposed previously that the VHPO Mcl40 could be responsible for constructing the entropically-demanding 15-membered ring in merochlorin C (Fig. 12, left panel), although the confirmation of it remains pending due to challenges related to protein insolubility and expression issues.[272]

Box 1. Biosynthesis of Napyradiomycins

Interestingly, the genome analysis of the napyradiomycin-producing strain CA-271078 revealed the presence of four putative VCPOs (Fig. 10), rather than the three previously described in the two napyradiomycin BGCs deposited in NCBI. In addition to the identified NapH1 and NapH3 proteins, other two VCPOs, here named NapHx and NapHy, are found in the gene cluster. These two enzymes showed similarity to both NapH1 and NapH4 from the previously described *nap* BGCs, as well as to Mcl40 from the merochlorin BGC. (Table 4)

VHPOs	NapH1 *(S. acuelolatus)*	NapH3 *(S. acuelolatus)*	NapH4 *(S. acuelolatus)*	Mcl40 (*Streptomyces* sp. CHN-189)
NapH1 (CA-271078)	**96% / 98%**			
NapH3 (CA-271078)		**83% / 86%**		
NapHx (CA-271078)	78% / 87%		70% / 82%	64% / 76%
NapHy (CA-271078)	63% / 76%		66% / 78%	61% / 74%

Table 4. Homology degrees (% identity / % similarity) for the VCPOs found in the *nap* BGC of CA-271078, NapH1-H3 from the *nap* BGC in *S. aculeolatus,* and Mcl40 from the merochlorin BGC.

Another unexpected result from the genome analysis was that the *nap* gene cluster in CA-271078 also encodes an additional FAD-dependent oxidoreductase (marked in turquoise blue in Fig. 10) which is not found in the previously described *nap* BGCs. The presence of these additional enzymes in the *nap* BGC of CA-271078 suggests that either NapHx or NapHy, or both, might be responsible for the macrocyclization leading to napyradiomycin D1 (**5**). The involvement of the extra oxidoreductase in this process cannot be ruled out either.

It is worth noting that napyradiomycin D1 (**5**) could be considered as the *O*-linkage analog of the previously reported napyradiomycin C1, both compounds presumably derived from napyradiomycin A1 as a common biosynthetic precursor. (Fig. 12, right panel). To date, however, no enzyme proposal has been described for the macrocyclization in napyradiomycins belonging to the C subfamily. Moreover, the genome of the strains known to produce these macrocyclic napyradiomycins (*S. rubra* MG802-AF1*, Streptomyces* sp. CNPQ-329, and *S. antimicotycus* NT17*)* are not publicly available, which prevented the identification of the napyradiomycins BGC in those strains and their comparison to the *nap* gene cluster identified in CA-271078.

Further insights into the structure and function of NapHx and NapHy, as well as of the additional FAD-dependent oxidoreductase found in the *nap* BGC from C-271078 may expand the knowledge of vanadium-dependent haloperoxidase enzymes and enhance our understanding of the biosynthesis of napyradiomycins.
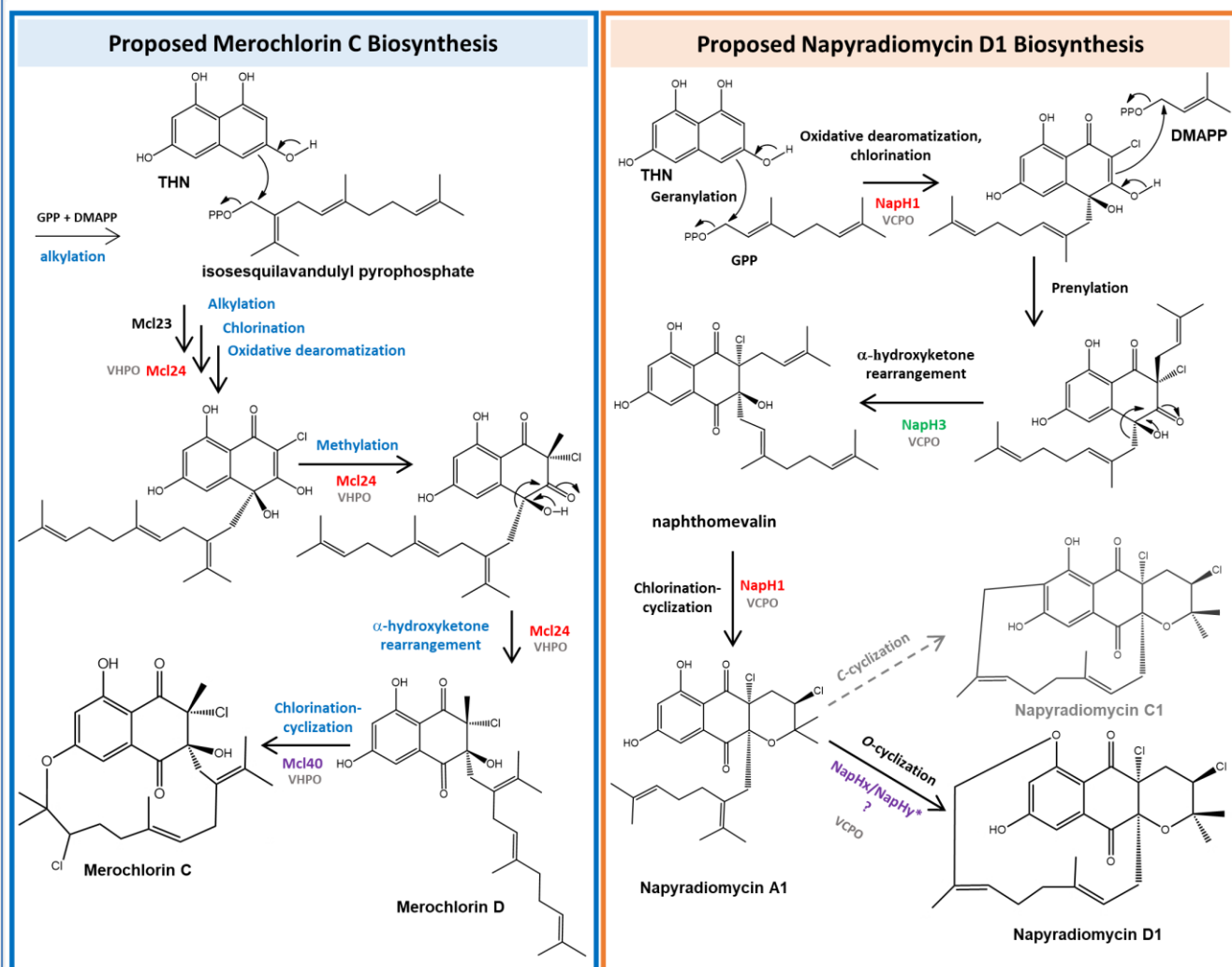
Figure 12. Proposed biosynthetic pathways for merochlorin C[270] (left panel) and napyradiomycin D1 (**5**) (this work, right panel). The proposed merochlorin C biosynthesis involves two VHPOs: Mcl24, responsible for the unexpected α-hydroxyketone rearrangement under basic pH conditions, and Mcl40, which would catalyze the macrocyclic chloro-ether ring formation. The proposed pathway herein to napyradiomycin D1 (**5**) might involve up to four VCPOs NapH1, NapH3, NapHx, and NapHy. Further studies should be conducted to validate this hypothesis.

5. Experimental Section

5.1. General Experimental Procedures

Optical rotations were measured on a Jasco P-2000 polarimeter (JASCO Corporation, Tokyo, Japan). IR spectra were recorded with a JASCO FT/IR-4100 spectrometer (JASCO Corporation) equipped with a PIKE MIRacle™ single reflection ATR accessory. 1D- and 2D-NMR spectra were recorded on a Bruker Avance III spectrometer (500 and 125 MHz for ^{1}H and ^{13}C NMR, respectively) equipped with a 1.7 mm TCI MicroCryoProbe™ (Bruker Biospin, Fällanden, Switzerland). Chemical shifts were reported in ppm using the signals of the residual solvents as internal reference (δ_H 2.51 and δ_C 39.5 ppm for DMSO-d_6). LC-UV-MS analysis was performed on an Agilent 1100 (Agilent Technologies, Santa Clara, CA, USA) single quadrupole LC-MS system as previously described.[261] ESI-TOF spectra were acquired using a Bruker maXis QTOF (Bruker Daltonik GmbH, Bremen, Germany) mass spectrometer coupled to an Agilent 1200 LC (Agilent Technologies, Waldbronn, Germany). Medium-pressure liquid chromatography (MPLC) was performed on semiautomatic flash chromatography (CombiFlash Teledyne ISCO Rf400x) with a precast reversed-phase column. Semi-preparative HPLC separation was performed on Gilson GX-281 322H2 (Gilson Technologies, USA) with a semi-preparative reversed-phase column (Zorbax SB-C18, 250 × 9.4 mm, 5 µm). Preparative HPLC separation was performed on Gilson GX-281 322H2 (Gilson Technologies, USA) with a reversed-phase column (Zorbax SB-C18, 250 × 21.2 mm, 7 µm). Acetone used for extraction was analytical grade. Solvents employed for isolation were HPLC grade. Molecular models were generated using Chem3D Pro 12.0 (CambridgeSoft, PerkinElmer Informatics, Waltham, MA, USA). The structures were energy-minimized by molecular mechanics with the MM2 force field using as gradient convergence criteria an RMS value of 0.001. Molecular modeling figures were generated with PyMol (W. L. DeLano, The PyMOL Molecular Graphics System, DeLano Scientific LLC, Palo Alto, CA, USA, 2002). The conformational search was performed for 5 with ConfBuster Web Server, using as input the corresponding sdf files generated with Chem3D Pro 12.0 (CambridgeSoft, PerkinElmer Informatics, Waltham, MA, USA).

5.2. Taxonomic Identification of the Producing Microorganism

The taxonomic identification of the strain was described in a previous work.[241]

5.3. Fermentation of the Producing Microorganism

A 3 L fermentation of the strain CA-271078 of *Streptomyces* sp. was generated using the conditions described in the reference.[241] In brief, the producing strain was fermented at 28 °C in 1 L of R358 medium for six days. Extraction with an equal volume of acetone and evaporation of the organic solvent afforded an acetone crude extract, which was subsequently subjected to liquid-liquid extraction with ethyl acetate (EtOAc).

5.4. Extraction and Bioassay Guided Isolation

The fermentation broth (3 L) was extracted with acetone (3 L) under continuous shaking at 220 rpm for 1h. The mycelium was separated and discarded by centrifugation at 9000 rpm and filtration and the supernatant (ca. 6L) was concentrated to 3L under reduced pressure of the nitrogen stream. The aqueous crude extract was extracted with EtOAc (3 × 0.9 L) to afford a crude extract of 0.178 g. This EtOAc extract was loaded onto a Reversed-Phase C18 (ODS) column (32 × 100 mm) that was eluted with a gradient of acetonitrile in water (35% to 100% ACN in 50 min + 100% ACN in 10 min, 15 mL/min, 18 mL/fraction) to afford 50 fractions. They were combined into five fractions according to their LC-UV-MS profiles and evaporated to dryness in a centrifugal evaporator: fractions A (6.6 mg), B (5.3 mg), C (9.0 mg), D (12.7 mg), and E (17.8 mg).

Fractions containing the compounds of interest from this chromatography were further purified by semipreparative reversed-phase HPLC.

Fraction A (6.6 mg) was chromatographed by semipreparative reversed-phase HPLC (Zorbax SB-C18, 9.4 × 250 mm, 5 µm; 3.6 mL/min, UV detection at 210 and 280 nm) with an isocratic elution of 33% CH₃CN/ 67% H₂O with 0.1% trifluoroacetic acid over 34 min yielding **1** (0.7 mg, t_R 18 min) and **6** (2.1 mg. t_R 25 min).

Fraction B (5.3 mg) was chromatographed by semipreparative reversed-phase HPLC (Zorbax SB-C18, 9.4 × 250 mm, 5 µm; 3.6 mL/min, UV detection at 210 and 280 nm) with an isocratic elution of 35% CH₃CN/ 65% H₂O with 0.1% trifluoroacetic acid over 34 min yielding **7** (1.1 mg, t_R 21.5 min).

Fraction C (9.0 mg) was subjected to reversed-phase semipreparative HPLC (Zorbax SB-C18, 9.4 × 250 mm, 5 µm; 3.6 mL/min, UV detection at 210 and 280 nm) with and isocratic elution of CH₃CN/ H₂O 55/45 with 0.1% trifluoroacetic acid over 34 min, yielding **3** (0.7 mg, t_R 15 min), **4** (0.5 mg, t_R 22.5 min), **8** (0.5 mg, t_R 20.5 min) and **9** (0.5 mg, t_R 24 min).

Fraction D (12.7 mg) was chromatographed by semipreparative reversed-phase HPLC (Zorbax SB-C18, 9.4 × 250 mm, 5 µm; 3.6 mL/min, UV detection at 210 and 280 nm) with a linear gradient of CH₃CN/ H₂O with 0.1% trifluoroacetic acid, from 55 to 65 % CH₃CN over 34 min, yielding **2** (1.3 mg, t_R 23.5 min), **10** (0.8 mg, t_R 25.2 min), **11** (0.9 mg, t_R 27 min), **12** (1.3 mg, t_R 28.5 min) and **13** (0.9 mg, t_R 31 min).

Fraction E (12.7 mg) was subjected to reversed-phase semipreparative HPLC (Zorbax SB-C18, 9.4 × 250 mm, 5 µm; 3.6 mL/min, UV detection at 210 and 280 nm) with a linear gradient of CH₃CN/ H₂O with 0.1% trifluoroacetic acid, from 65 to 70 % CH₃CN over 34 min, yielding **5** (1.0 mg, t_R 21 min), **14** (1.0 mg, t_R 27.5 min) and **15** (1.0 mg, t_R 32.5 min).

5.5. Characterization Data

Napyradiomycin A3 (1): $[\alpha]^{25}_{D}$ +4.0 (*c* 0.38, MeOH); IR (ATR) cm⁻¹: 3379, 2977, 1702, 1673, 1619, 1267, 1186, 1136, 1024, 837, 720; (+)-ESI-TOF-MS *m/z* 465.1877 [M + Na]⁺ (calcd. for $C_{25}H_{30}NaO_7^+$, 465.1884), 460.2330 [M + NH₄]⁺ (calcd. for $C_{25}H_{34}NO_7^+$, 460.2330), 443,2060 [M + H]⁺ (calcd. for $C_{25}H_{31}O_7^+$, 443.2064), 425.1958 [M + H − H₂O]⁺ (calcd. for $C_{25}H_{29}O_6^+$, 425.1960); ¹H and ¹³C NMR data see Tables 1 and 2.

Napyradiomycin B7a (2): $[\alpha]^{25}_{D}$ −41.0 (*c* 0.4, MeOH); IR (ATR) cm⁻¹: 3374, 2930, 1680, 1615, 1377, 1260, 1202, 1137, 1025, 798, 722; (+)-ESI-TOF-MS *m/z* 1010.3184 [2M + NH₄]⁺ (calcd. for $C_{50}H_{64}{}^{35}Cl_4NO_{12}^+$, 1010.3177),514.1753 [M + NH₄]⁺ (calcd. for $C_{25}H_{34}{}^{35}Cl_2NO_6^+$, 514.1758), 479.1389 [M + H − H₂O]⁺ (calcd. for $C_{25}H_{29}{}^{35}Cl_2O_5^+$, 479.1387), 461.1284 [M + H − 2H₂O]⁺ (calcd. for $C_{25}H_{27}{}^{35}Cl_2O_4^+$, 461.1281); ¹H and ¹³C NMR data see Tables 1 and 2.

Napyradiomycin B7b (3): $[\alpha]^{25}_{D}$ +32.0 (*c* 0.13, MeOH); IR (ATR) cm⁻¹: 3368, 2980, 1680, 1618, 1455, 1378, 1268, 1204, 1137, 803; (+)-ESI-TOF-MS *m/z* 1010.3191 [2M + NH₄]⁺ (calcd. for $C_{50}H_{64}{}^{35}Cl_4NO_{12}^+$, 1010.3177), 514.1752 [M + NH₄]⁺ (calcd. for $C_{25}H_{34}{}^{35}Cl_2NO_6^+$, 514.1758), 479.1394 [M + H − H₂O]⁺ (calcd. for $C_{25}H_{29}{}^{35}Cl_2O_5^+$, 479.1387), 461.1294 [M+H − 2H₂O]⁺ (calcd. for $C_{25}H_{27}{}^{35}Cl_2O_4^+$, 461.1281), 443.1629 [M + H − 3H₂O]⁺ (calcd. for $C_{25}H_{25}{}^{35}Cl_2O_3^+$, 443.1175); ¹H and ¹³C NMR data see Tables 1 and 2.

Napyradiomycin SC (4): $[\alpha]^{25}_{D}$ −43.0 (*c* 0.4, MeOH); IR (ATR) cm⁻¹: 3340, 2976, 1703, 1632, 1615, 1464, 1372, 1264, 1224, 1183, 1082, 1024; (+)-ESI-TOF-MS *m/z* 497.1499 [M + H − H₂O]⁺ (calcd. for $C_{25}H_{31}{}^{35}Cl_2O_6^+$, 497.1492), 479.1389 [M + H − 2H₂O]⁺ (calcd. for $C_{25}H_{28}{}^{35}Cl_2O_5^+$, 479.1387); ¹H and ¹³C NMR data see Tables 1 and 2.

Napyradiomycin D1 (5): $[\alpha]^{25}_{D}$ +14.6 (*c* 0.13, MeOH); IR (ATR) cm⁻¹: 3358, 2978, 2948, 1671, 1608, 1458, 1371, 1298, 1185, 1136, 1079, 1028, 802; (+)-ESI-TOF-MS *m/z* 479.1386 [M+H]⁺ (calcd. for $C_{25}H_{29}{}^{35}Cl_2O_5^+$, 479.1387); ¹H and ¹³C NMR data see Tables 1 and 2.

5.6. Antibacterial and Antifungal Assays Cytotoxic Activities

Compounds **1–15** were tested in antimicrobial assays against the growth of Gram-negative (*E. coli* ATCC 25922 and *A. baumannii* MB5973*)* and Gram-positive bacteria (*M. tuberculosis* H37Ra and MRSA (MB5393), and fungi (*A. fumigatus* ATCC46645) in a cytotoxicity assay against the human liver adenocarcinoma cell line (HepG2) following previously described methodologies.[273–275] Briefly, each compound was serially diluted in DMSO with a dilution factor of 2 to provide 10 concentrations starting at 96 µg/mL for all the antimicrobial assays except for compounds **3** and **8** which started at 64 µg/mL. For the adenocarcinoma cell line, each compound was serially diluted in DMSO with a dilution factor of 2 to provide 10 concentrations starting at 150 µg/mL (compounds **2**, **4**, **6**, **7,** and **9**–**15**), 100 µg/mL (compounds **3** and **8**) or 30.0 µg/mL (compounds **1** and **5**). The MIC was defined as the lowest concentration of compound that inhibited ≥ 95% of the growth of a microorganism after overnight incubation. The Genedata Screener software (Genedata, Inc., Basel, Switzerland) was used to process and analyze the data and to calculate the RZ' factor, which predicts the robustness of an assay.[276] In all experiments performed in this work the RZ' factor obtained was between 0.87 and 0.98.

6. Supplementary Information

The following Supplementary Information is available online at:
http://www.mdpi.com/1660-3397/18/1/22/s1.

List of Supplementary Materials:

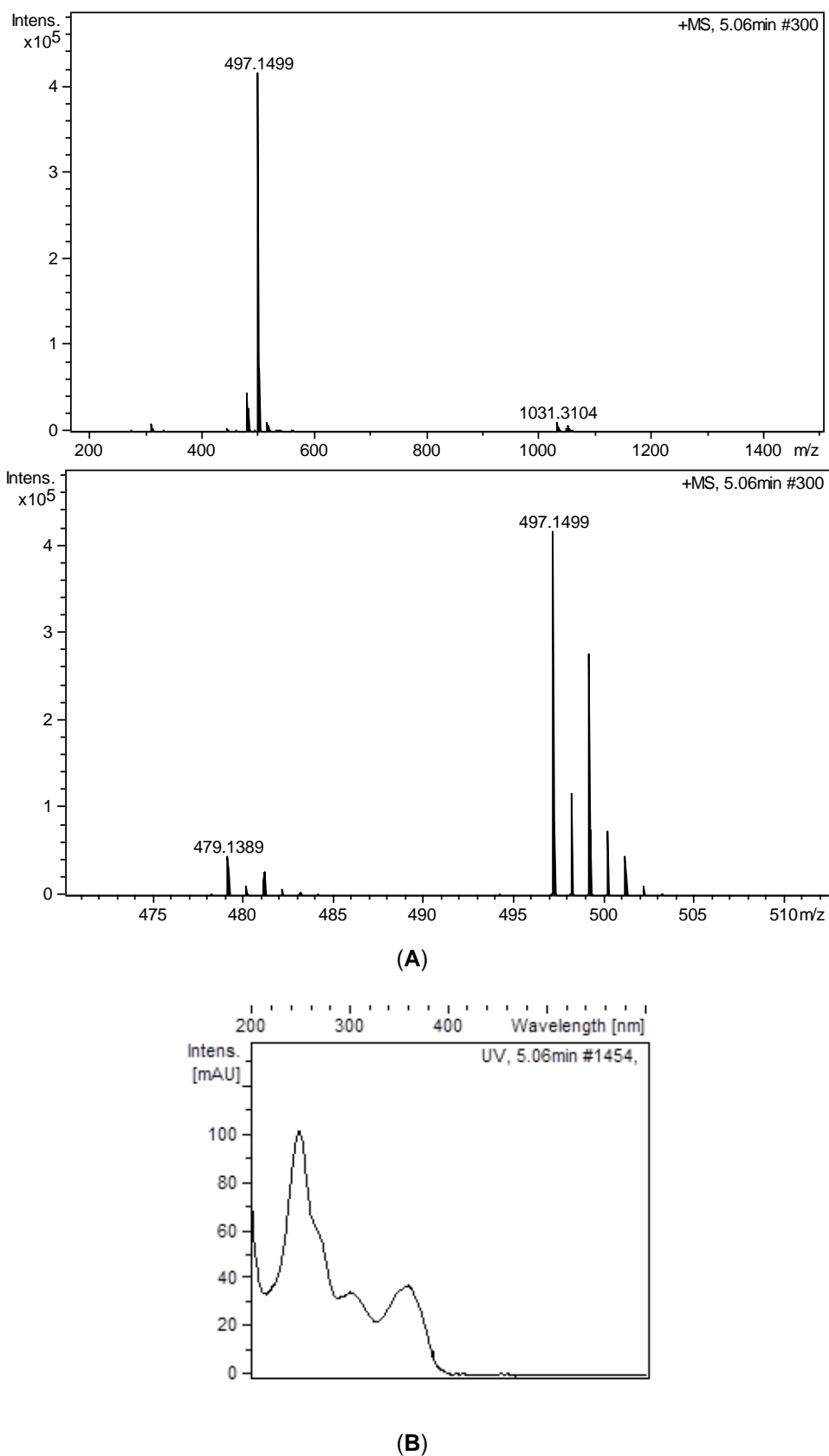(**A**)

(**B**)

Figure S1. Electrospray-time of flight (ESI-TOF) (**A**) and UV (**B**) spectra of compound **1**.

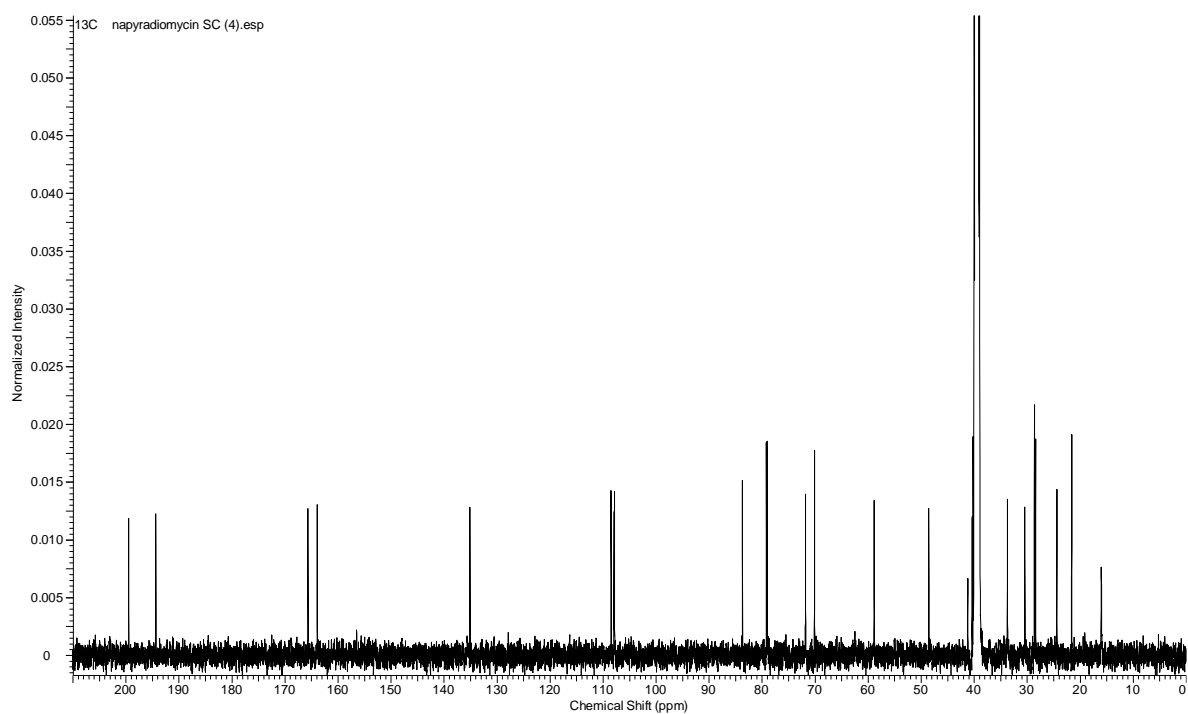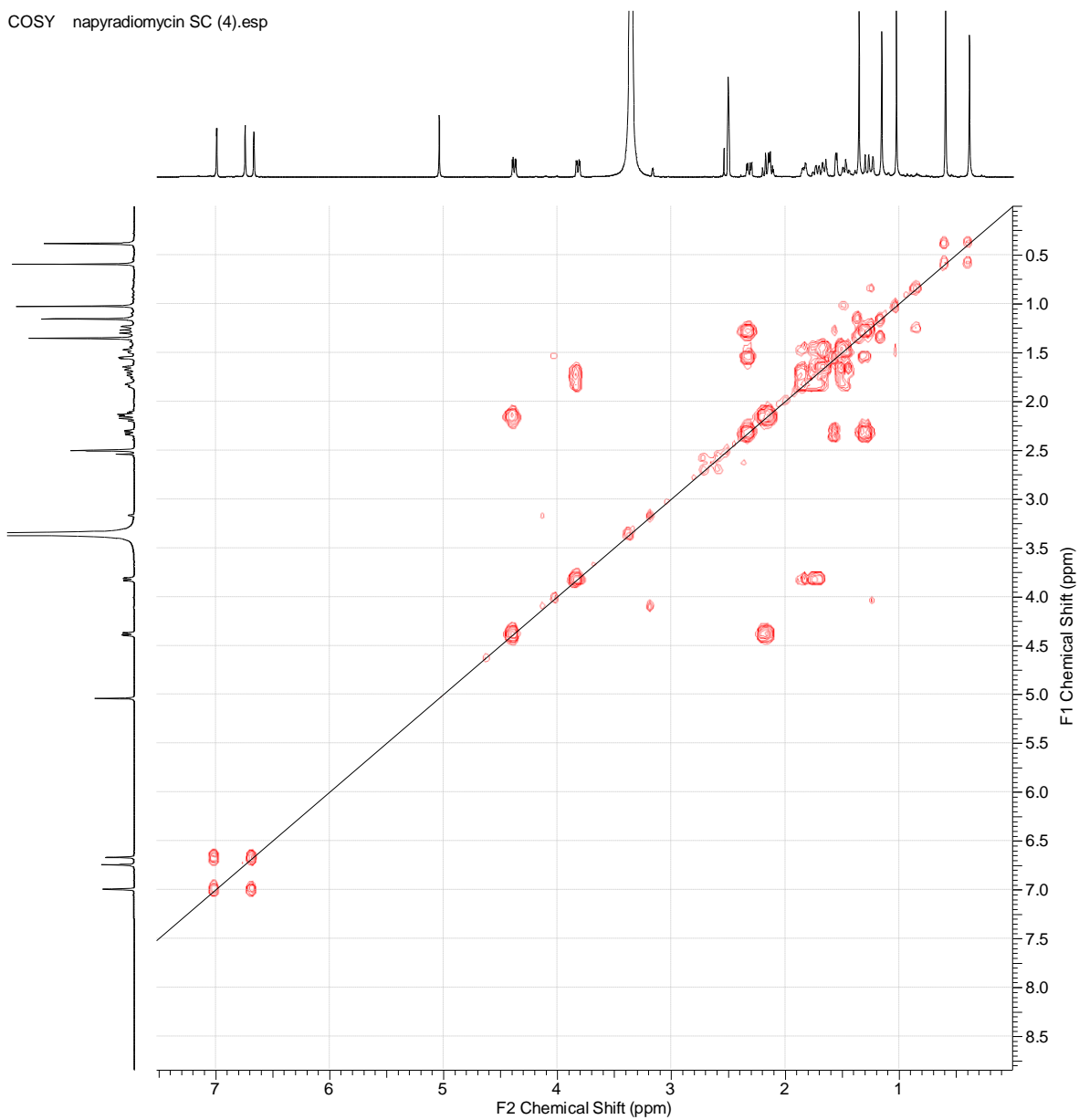**Figure S2.** ^{1}H NMR (DMSO-d_6, 500 MHz) of compound **1**.

Figure S3. ^{13}C NMR (DMSO-d_6, 125 MHz) of compound **1**.

COSY napyradiomycin A3 (1).esp

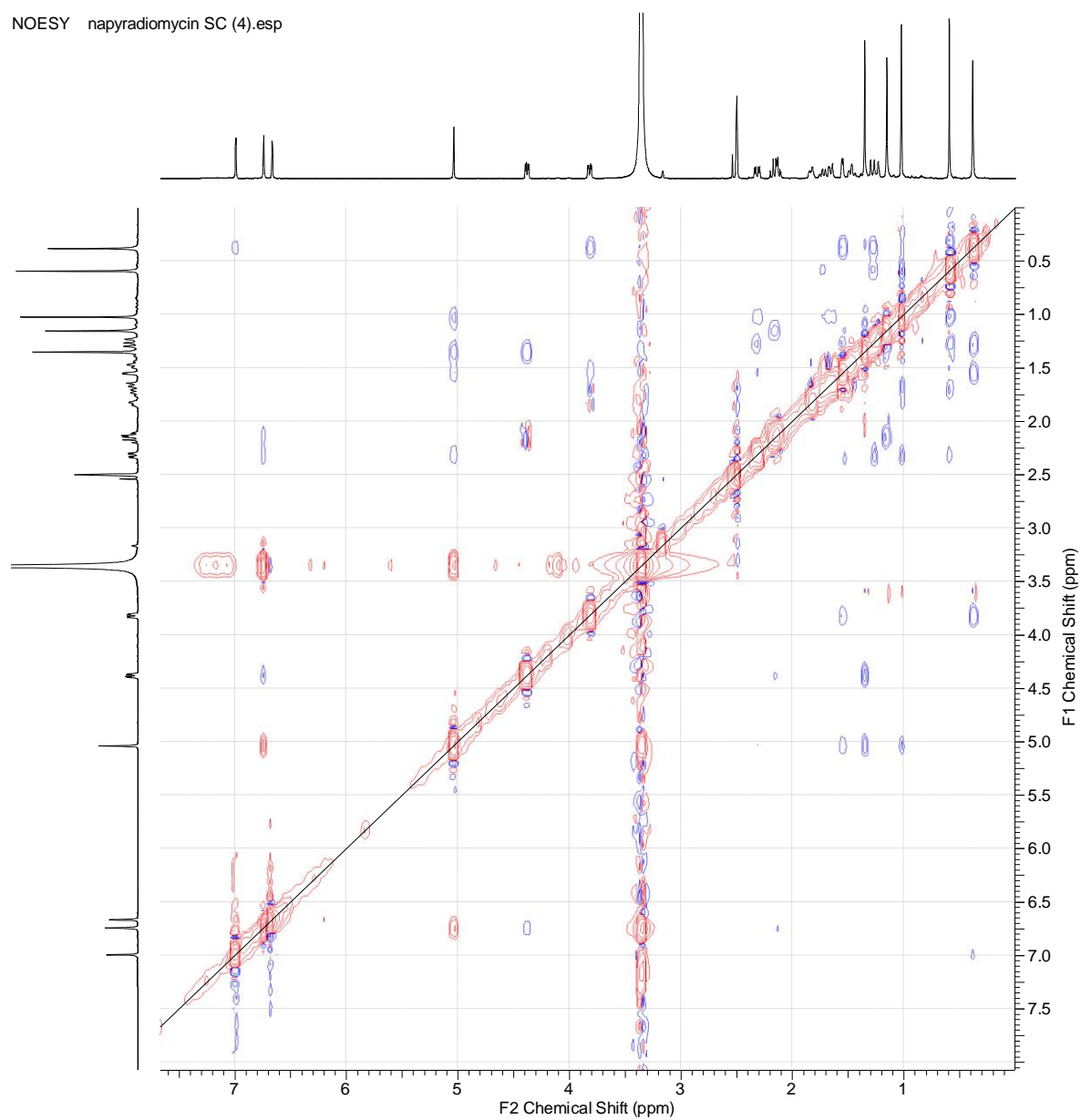**Figure S4.** COSY of compound **1**.

HSQC napyradiomycin A3 (1).esp

Figure S5. HSQC spectrum of compound **1**.

HMBC   napyradiomycin A3 (1).esp

Figure S6. HMBC of compound **1**.

ROESY   napyradiomycin A3 (1).esp
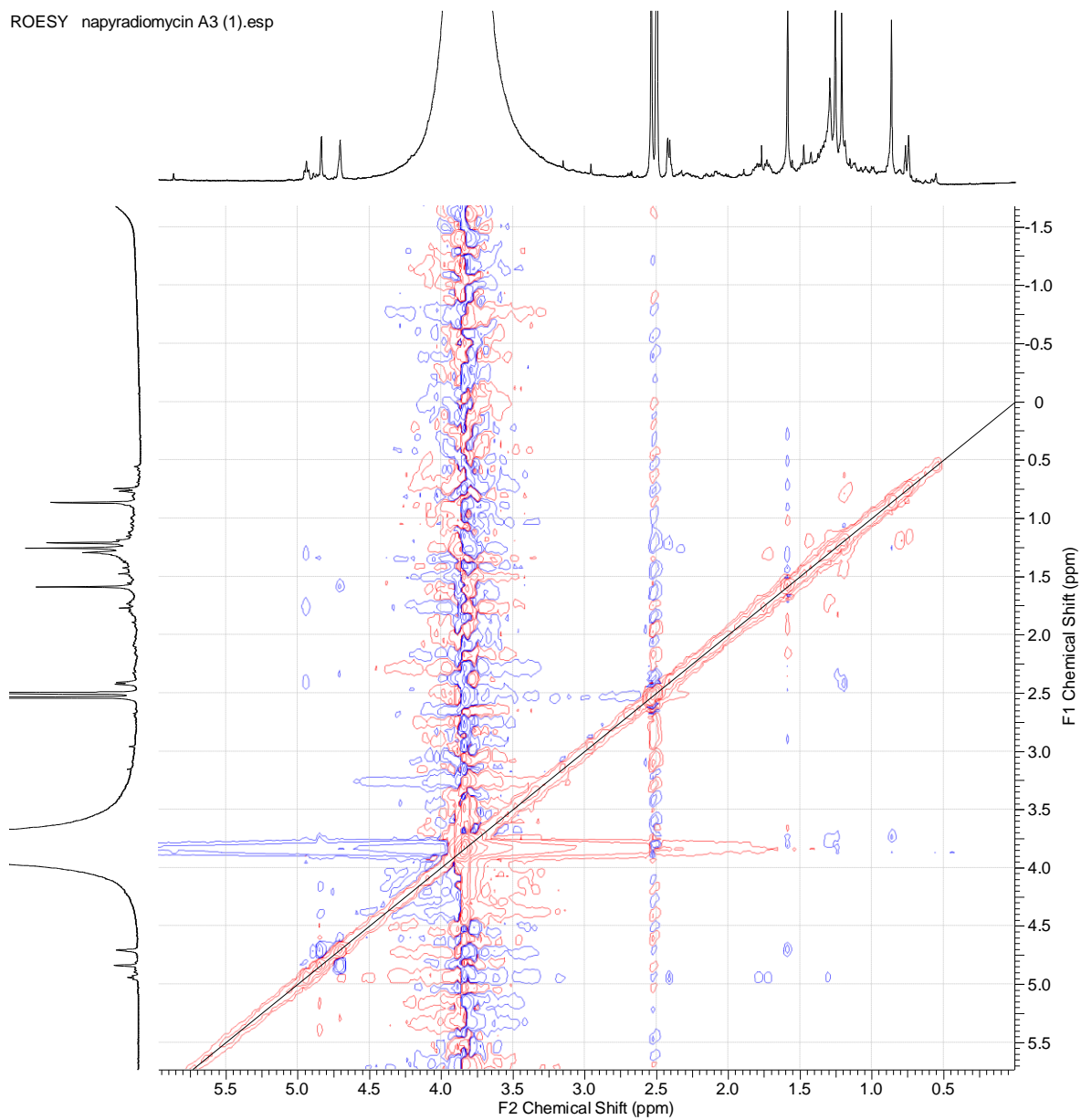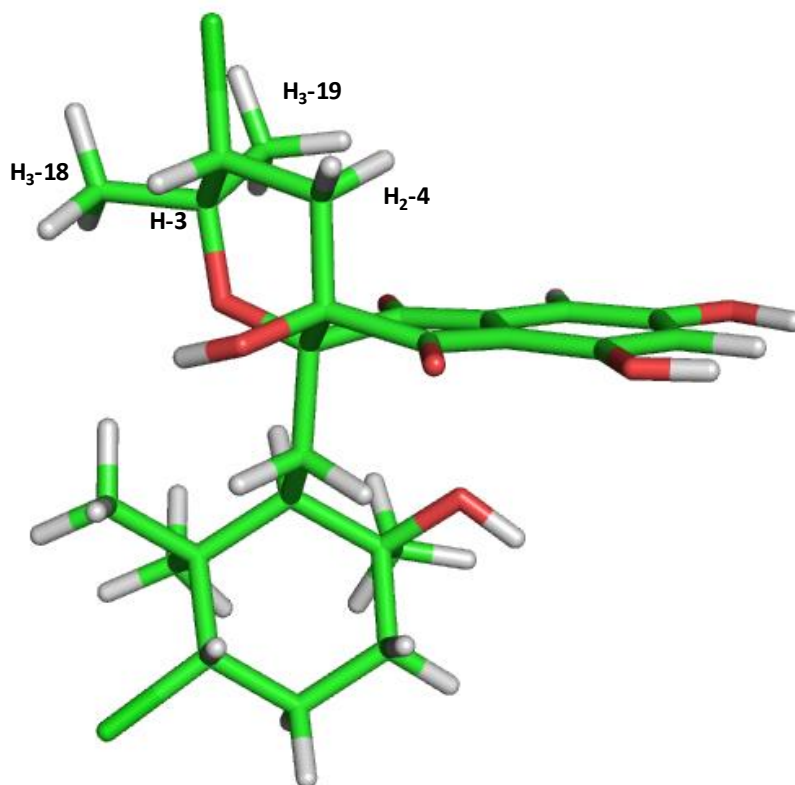
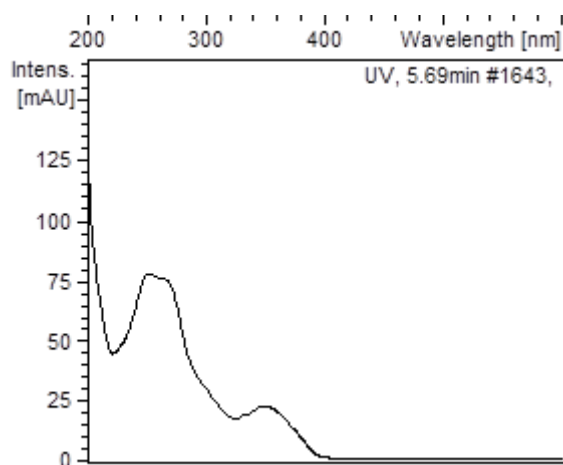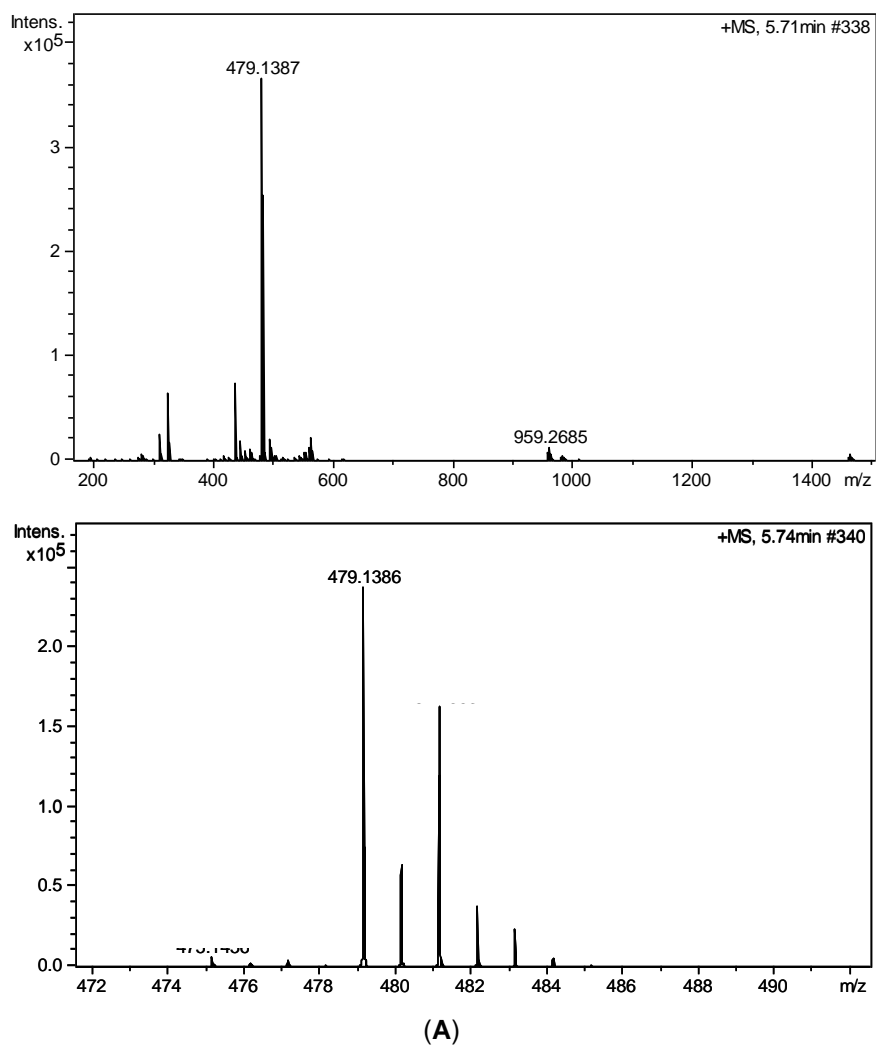

Figure S7. ROESY of compound **1**
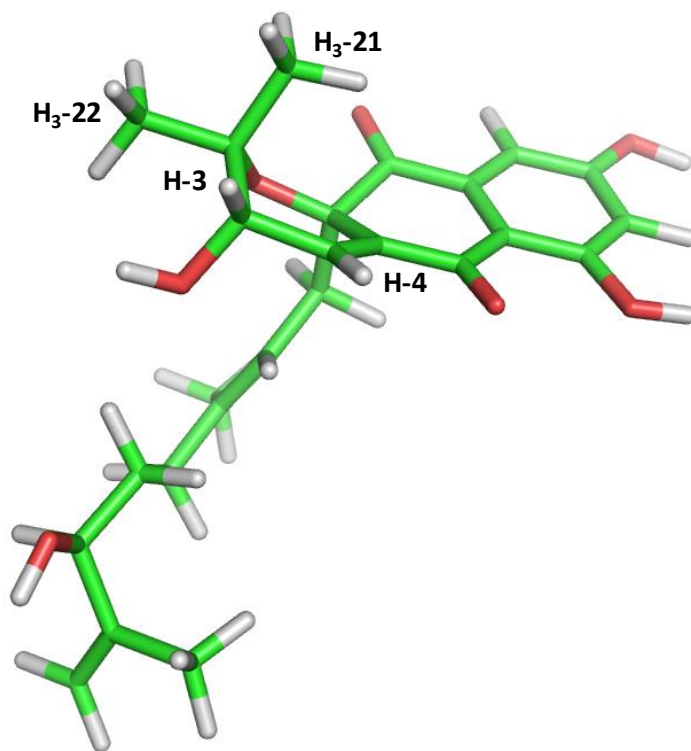
Figure S8. Energy minimized molecular model of compound **1**.

(**A**)

(**B**)
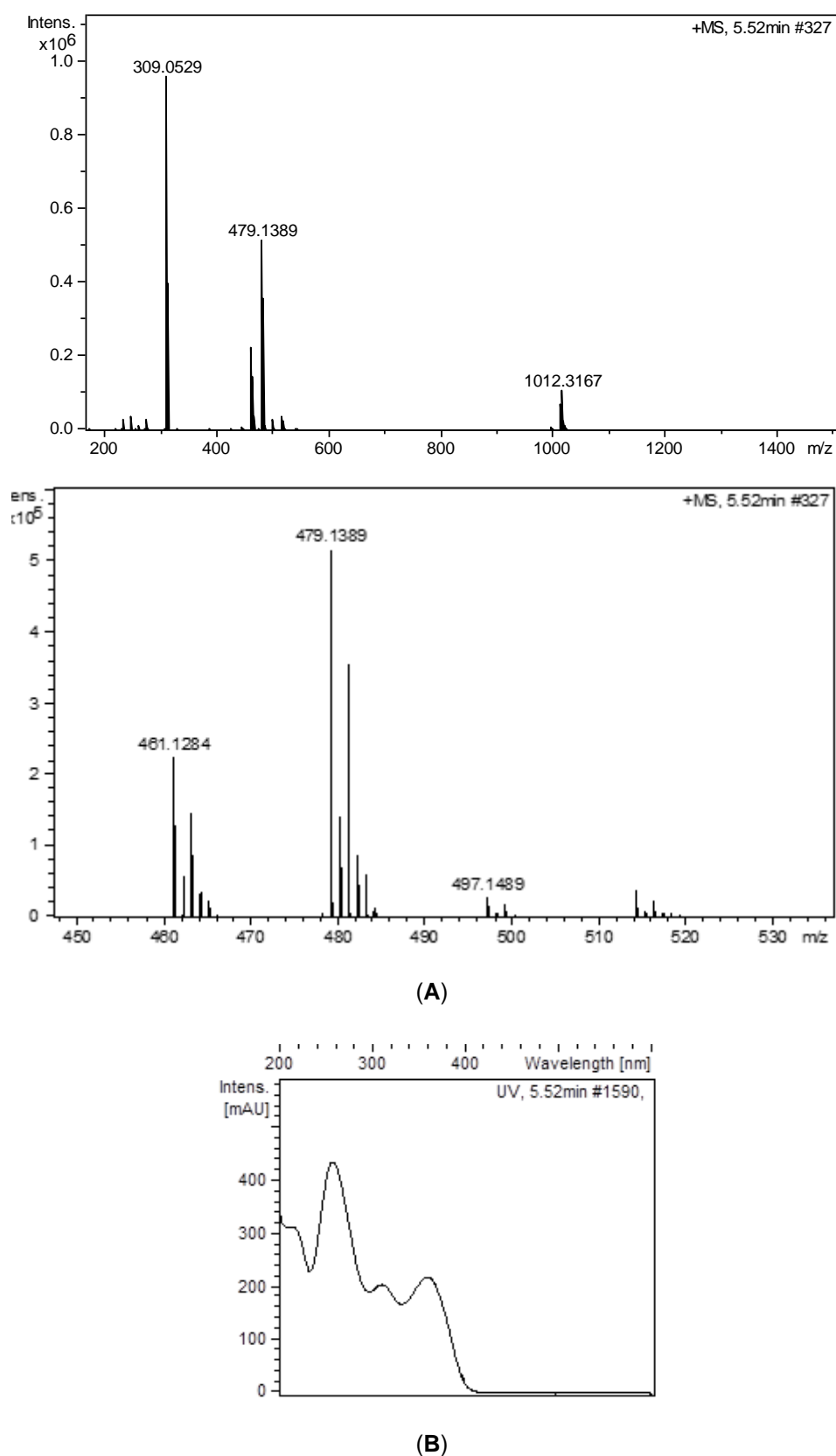
Figure S9. Electrospray-time of flight (ESI-TOF) (**A**) and UV (**B**) spectra of compound **2**.
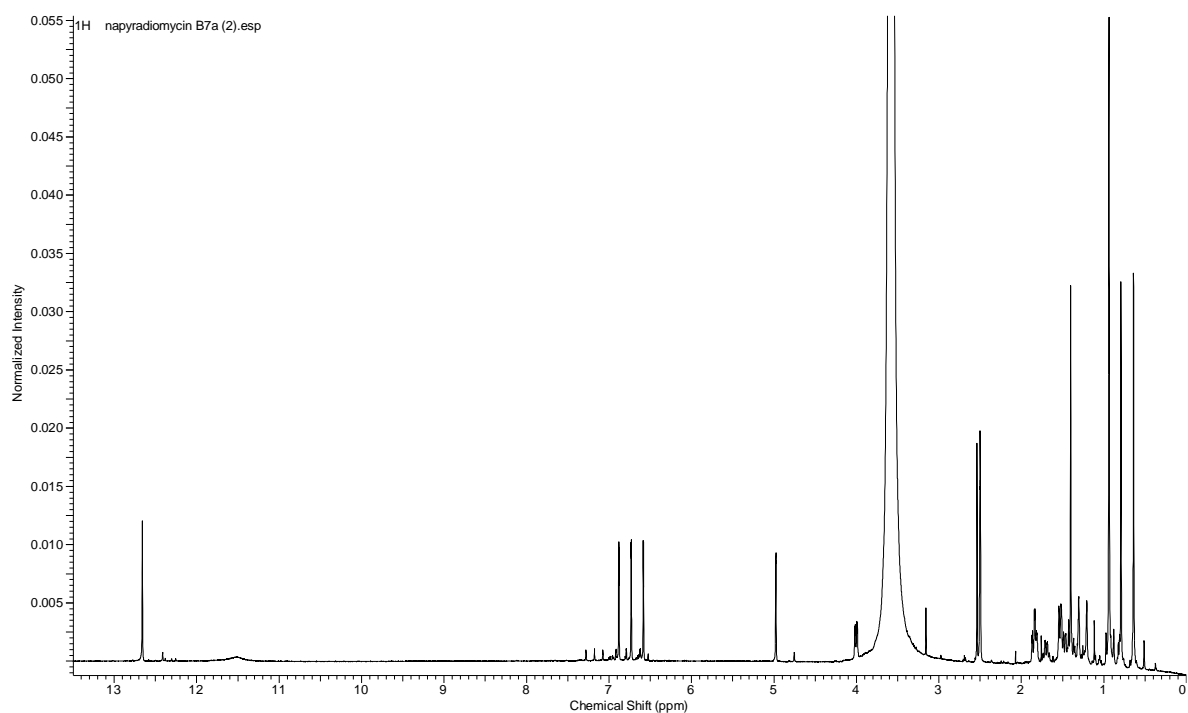
Figure S10. ^1H NMR (DMSO-d_6, 500 MHz) of compound **2**.
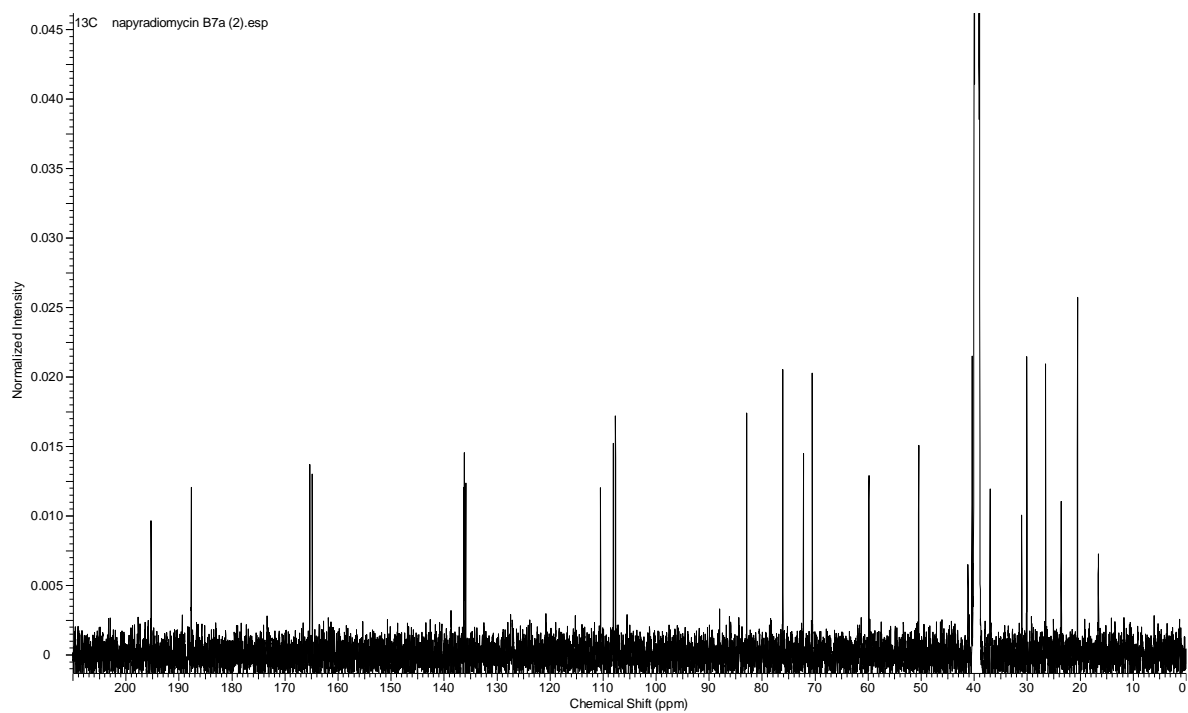
Figure S11. ^{13}C NMR (DMSO-d_6, 125 MHz) of compound **2**.

COSY   napyradiomycin B7a (2).esp
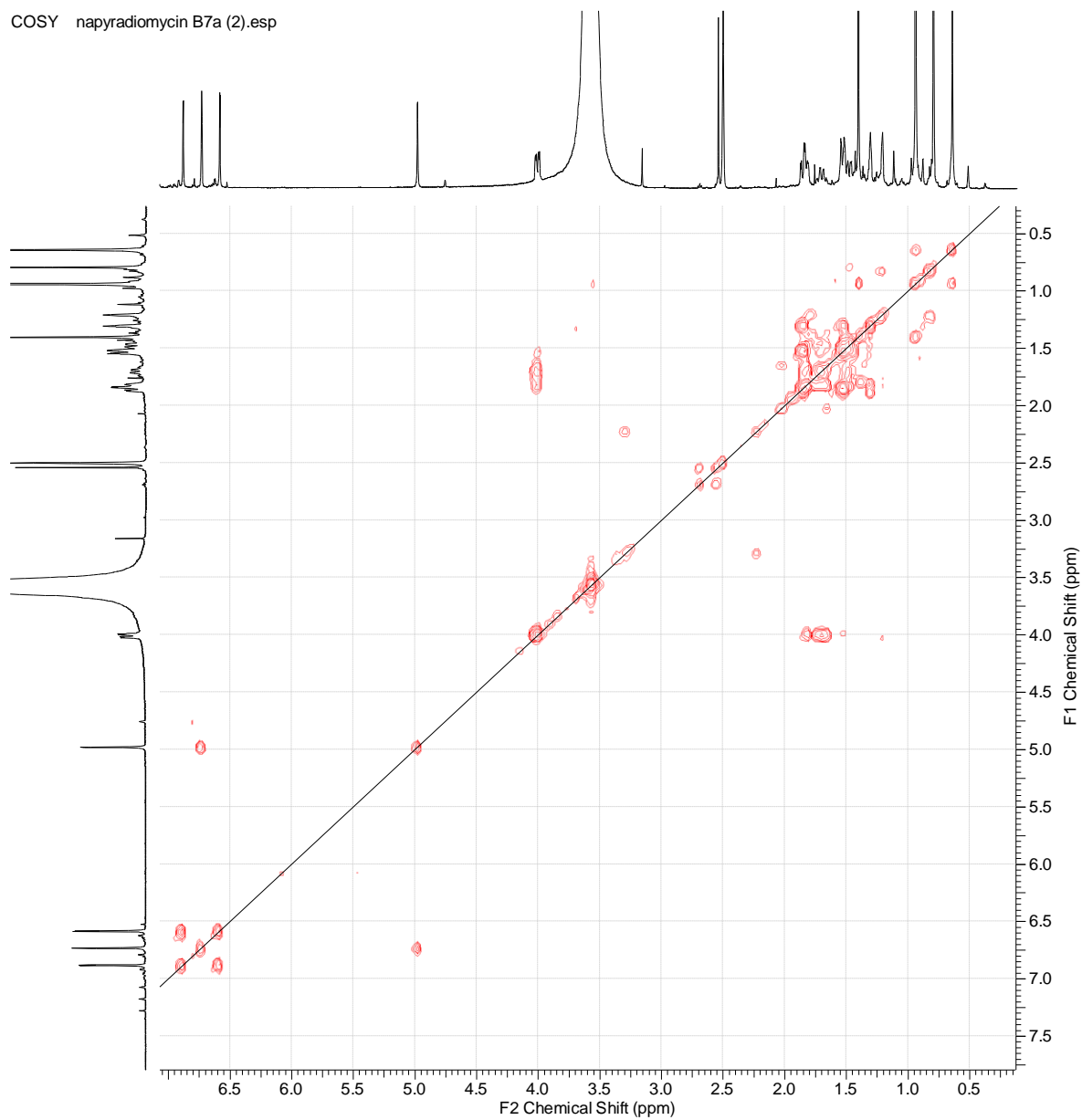
Figure S12. COSY of compound **2**.

HSQC   napyradiomycin B7a (2).esp
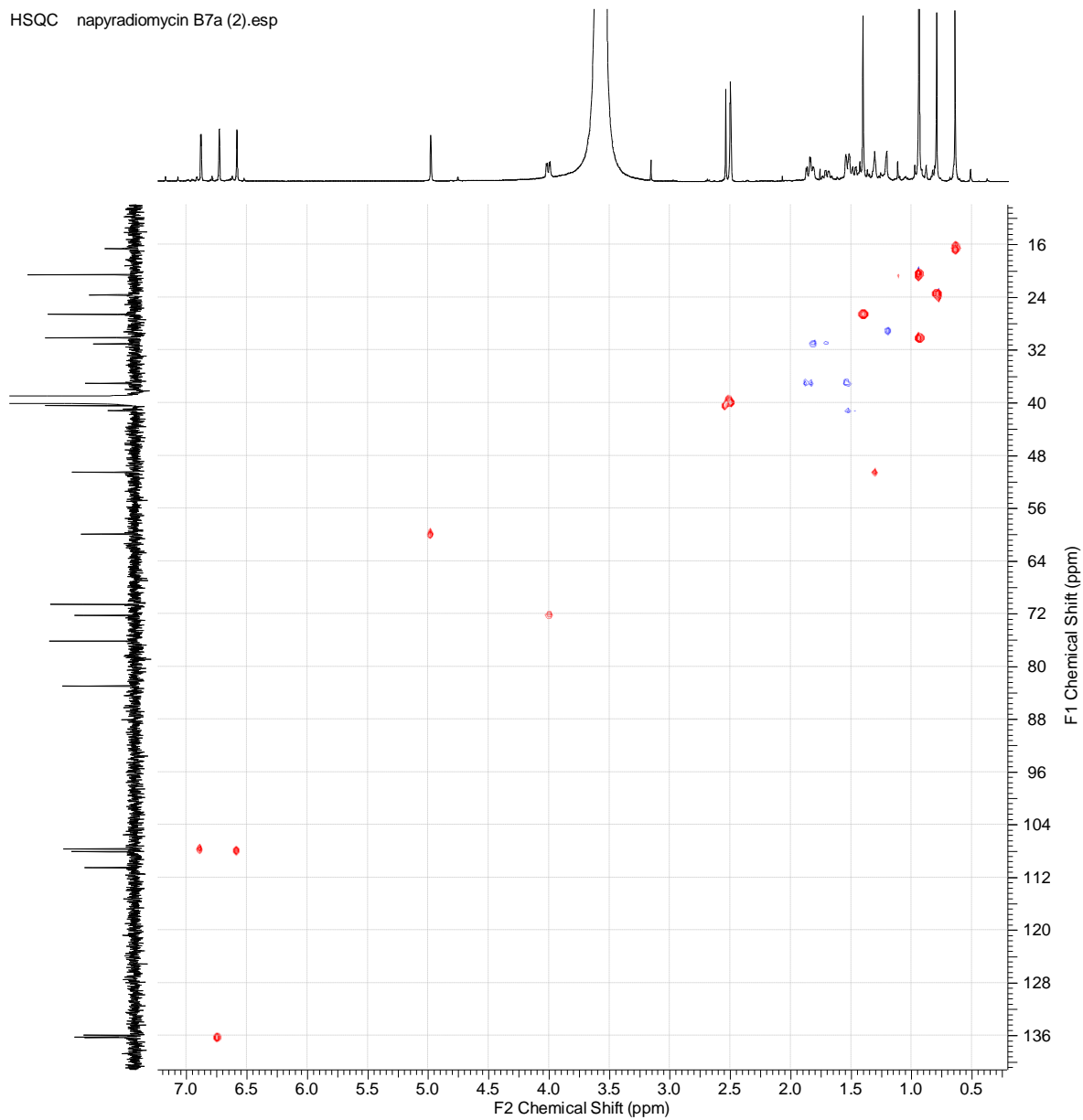
Figure S13. HSQC spectrum of compound **2**.
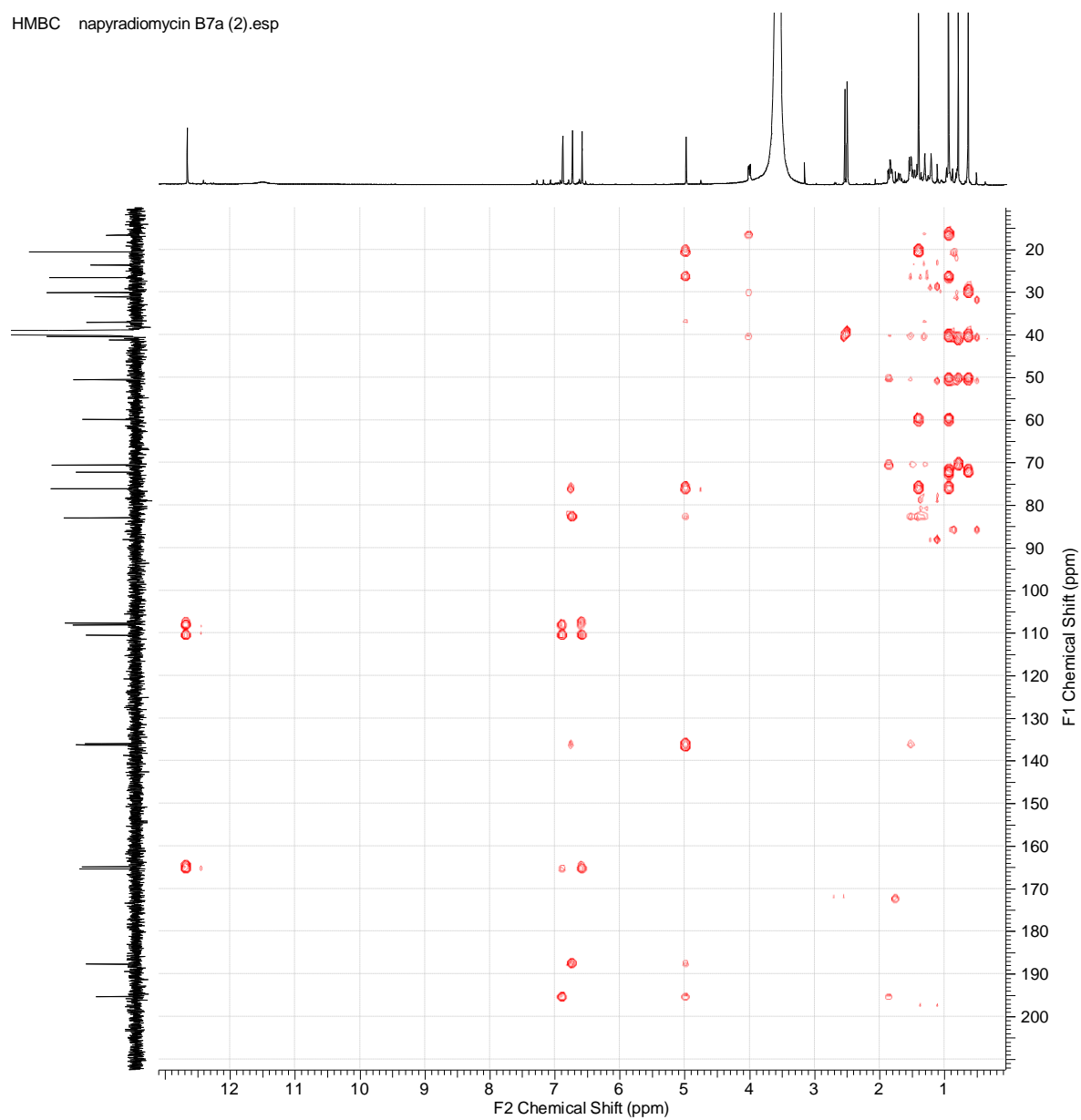
Figure S14. HMBC of compound **2.**

NOESY   napyradiomycin B7a (2).esp
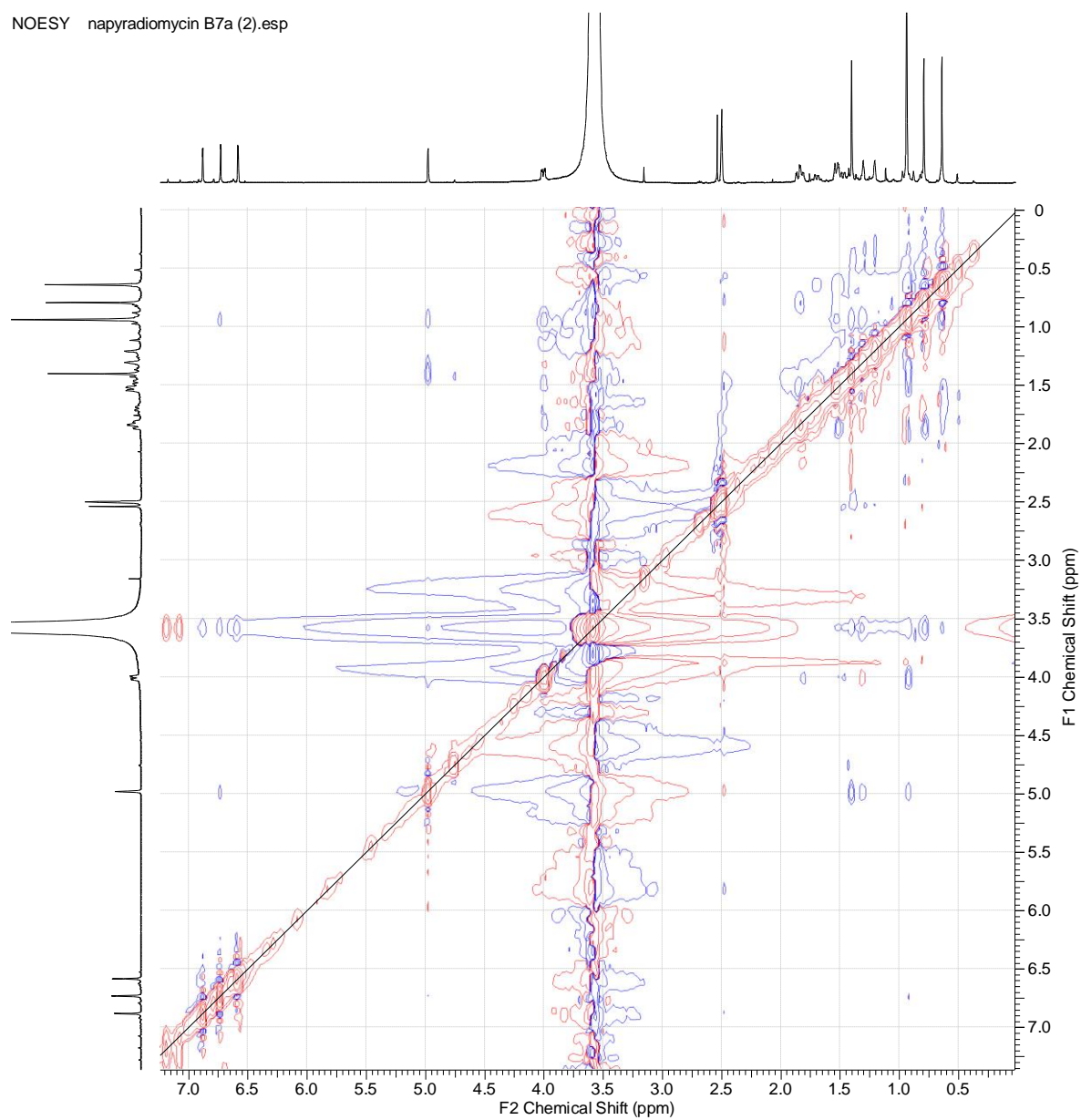
Figure S15. NOESY of compound **2**.

ROESY   napyradiomycin B7a (2).esp
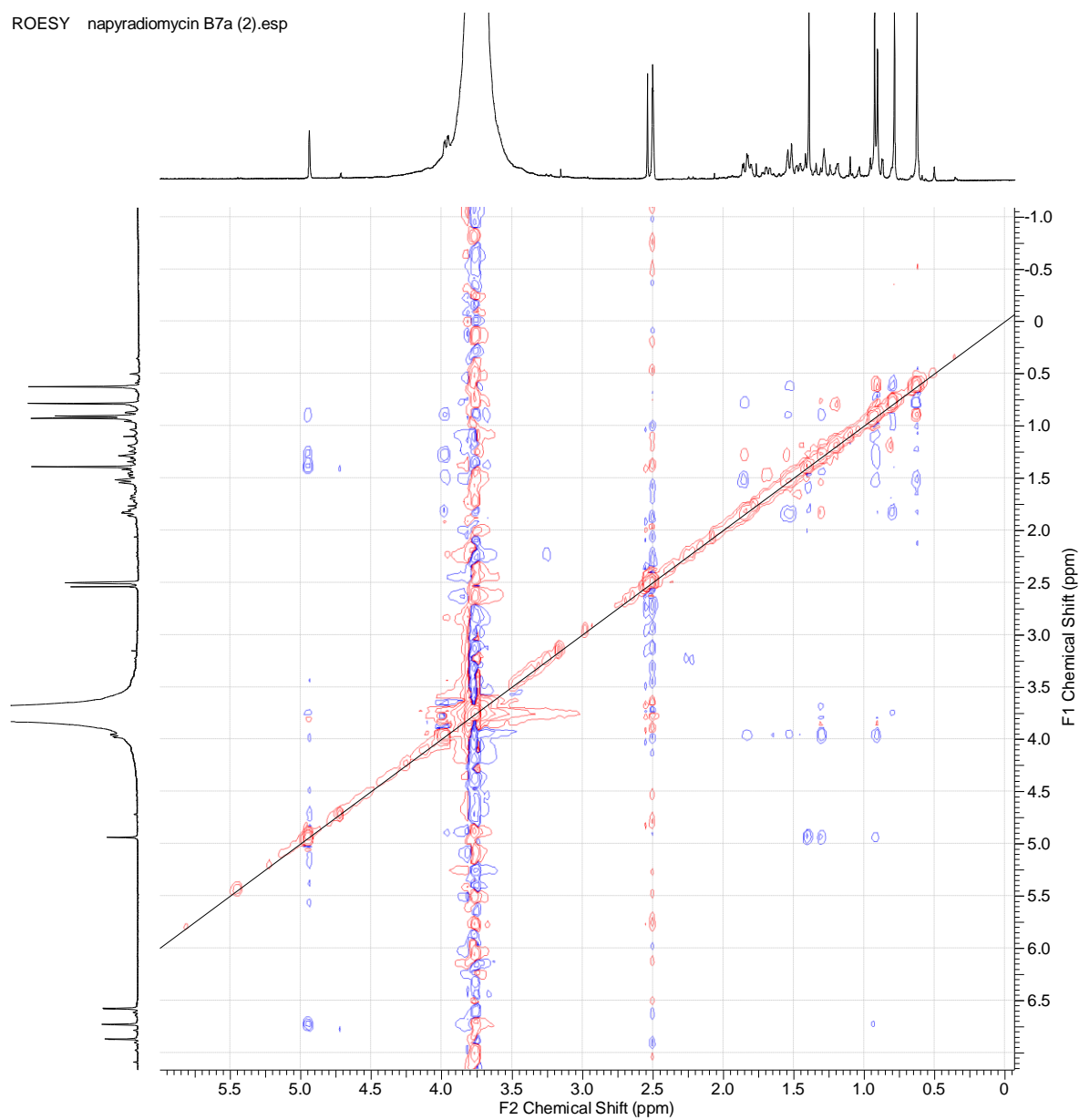
Figure S16. ROESY of compound **2**.
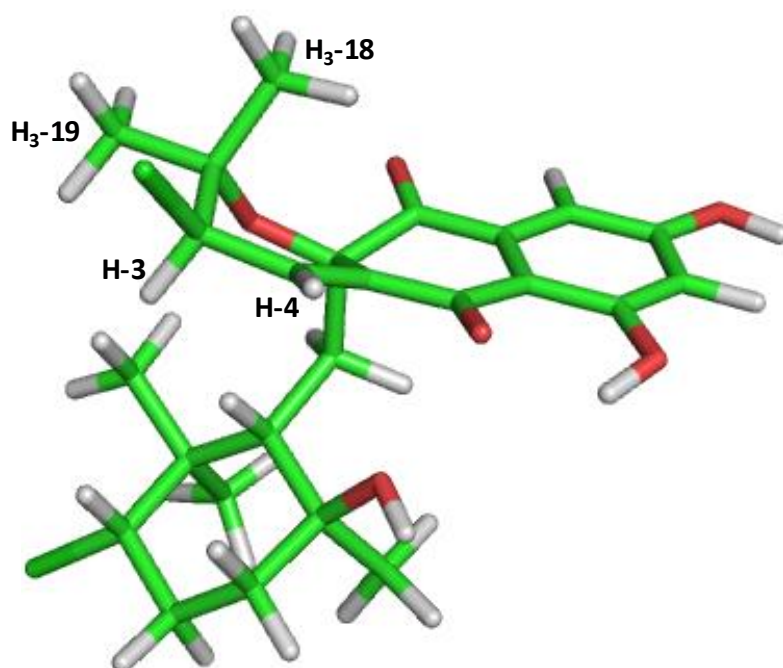
Figure S17. Energy minimized molecular model of compound **2**.
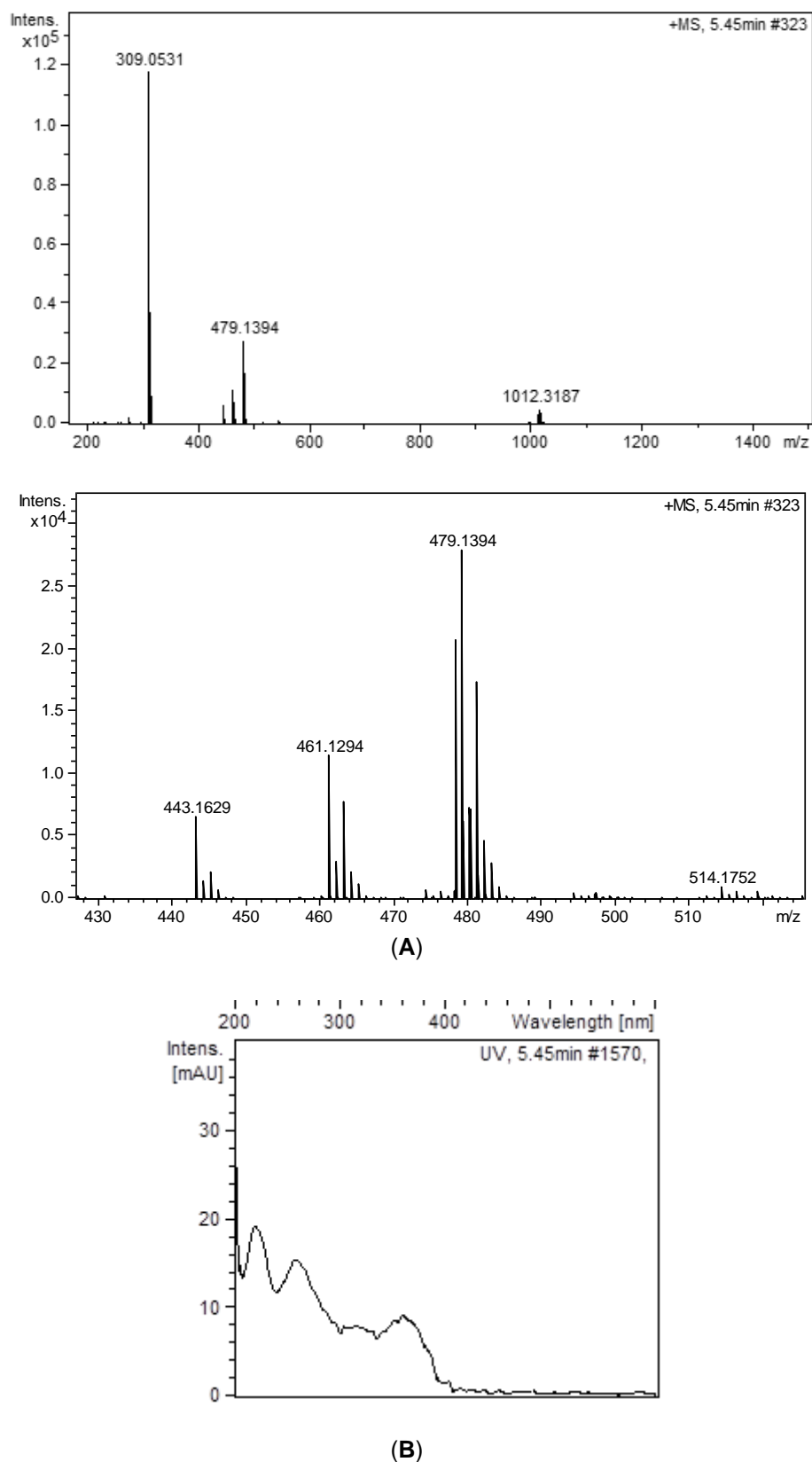
Figure S18. Electrospray-time of flight (ESI-TOF) (**A**) and UV (**B**) spectra of compound **3**.
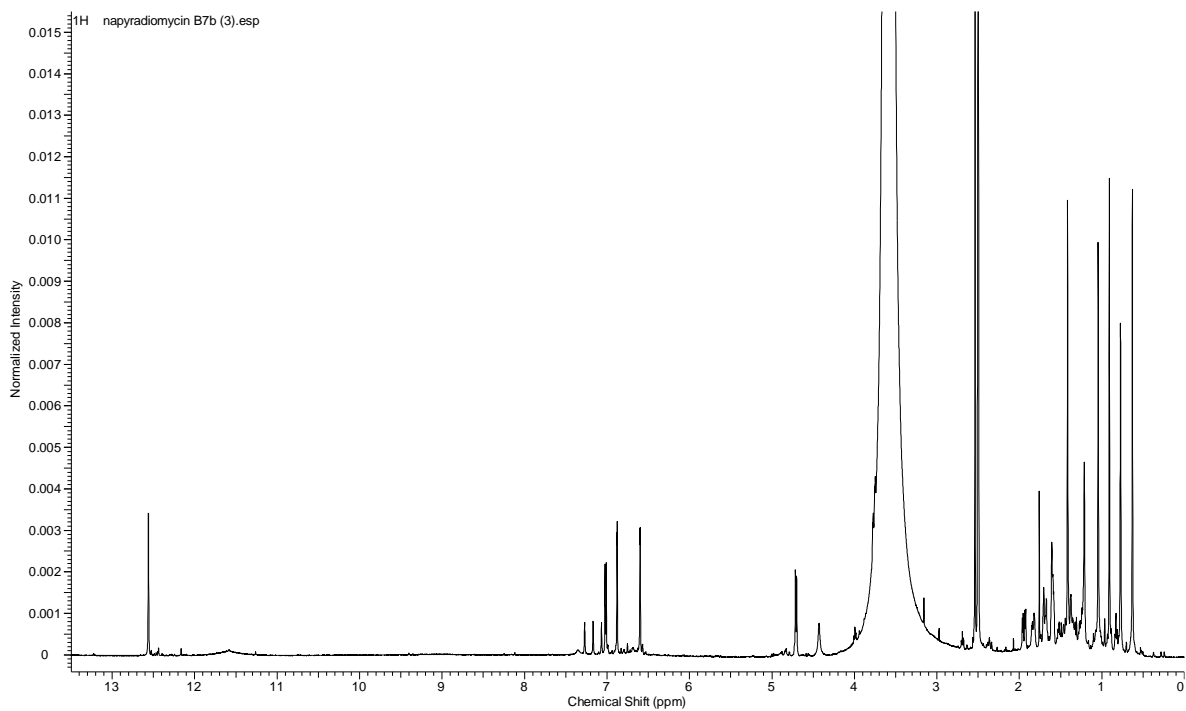
Figure S19. ^{1}H NMR (DMSO-d_6, 500 MHz) of compound **3**.
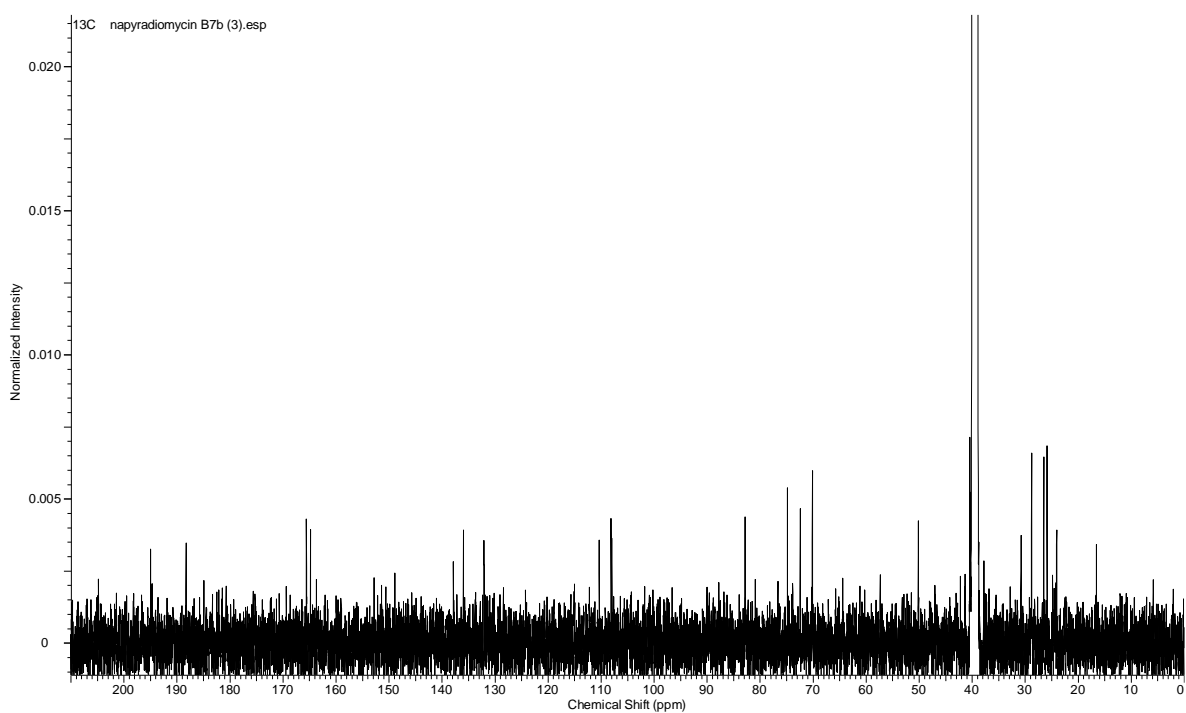
Figure S20. ^{13}C NMR (DMSO-d_6, 125 MHz) of compound **3**.

COSY   napyradiomycin B7b (3).esp
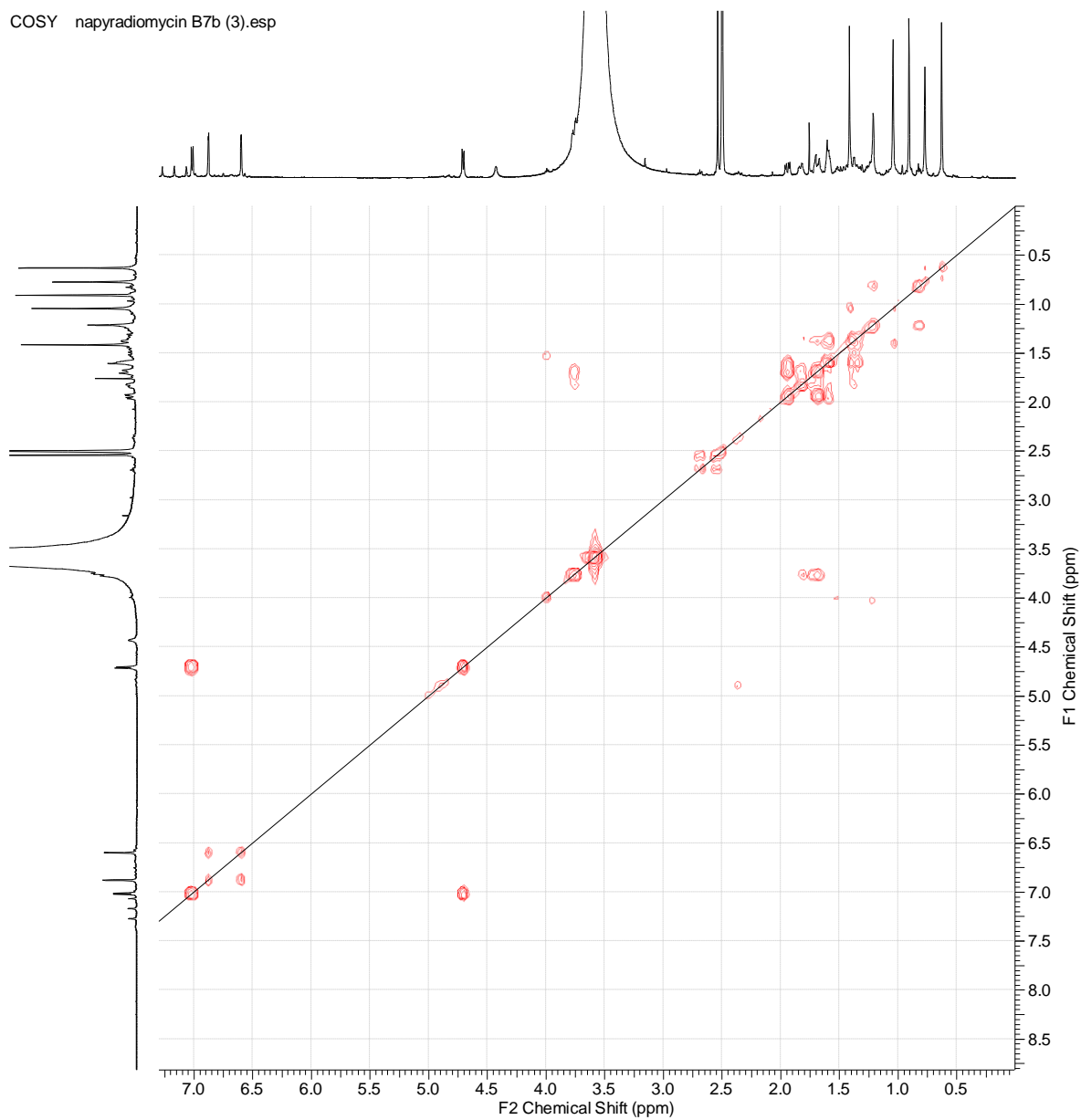
Figure S21. COSY of compound **3**.

HSQC   napyradiomycin B7b (3).esp
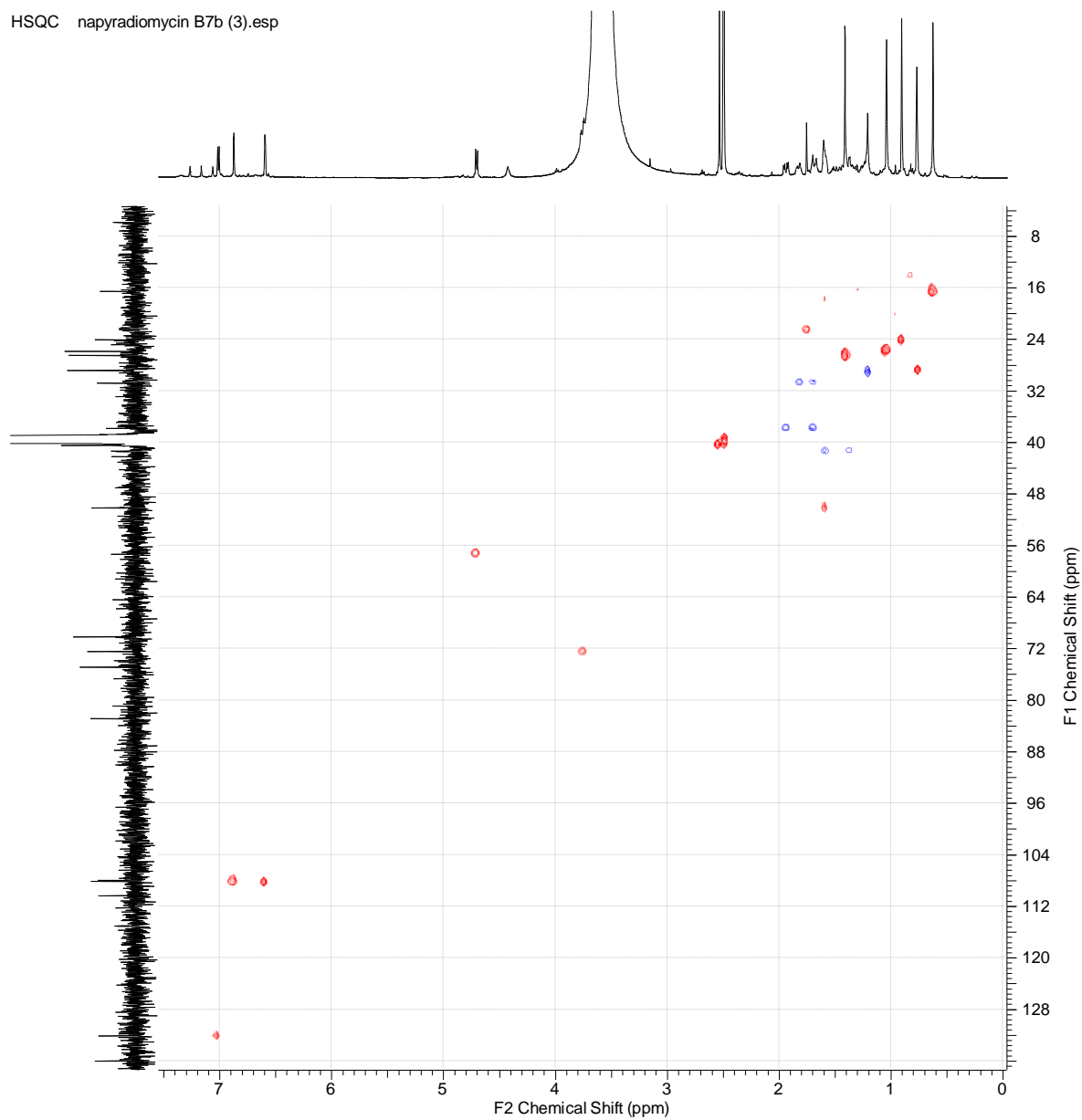
Figure S22. HSQC spectrum of compound **3**.
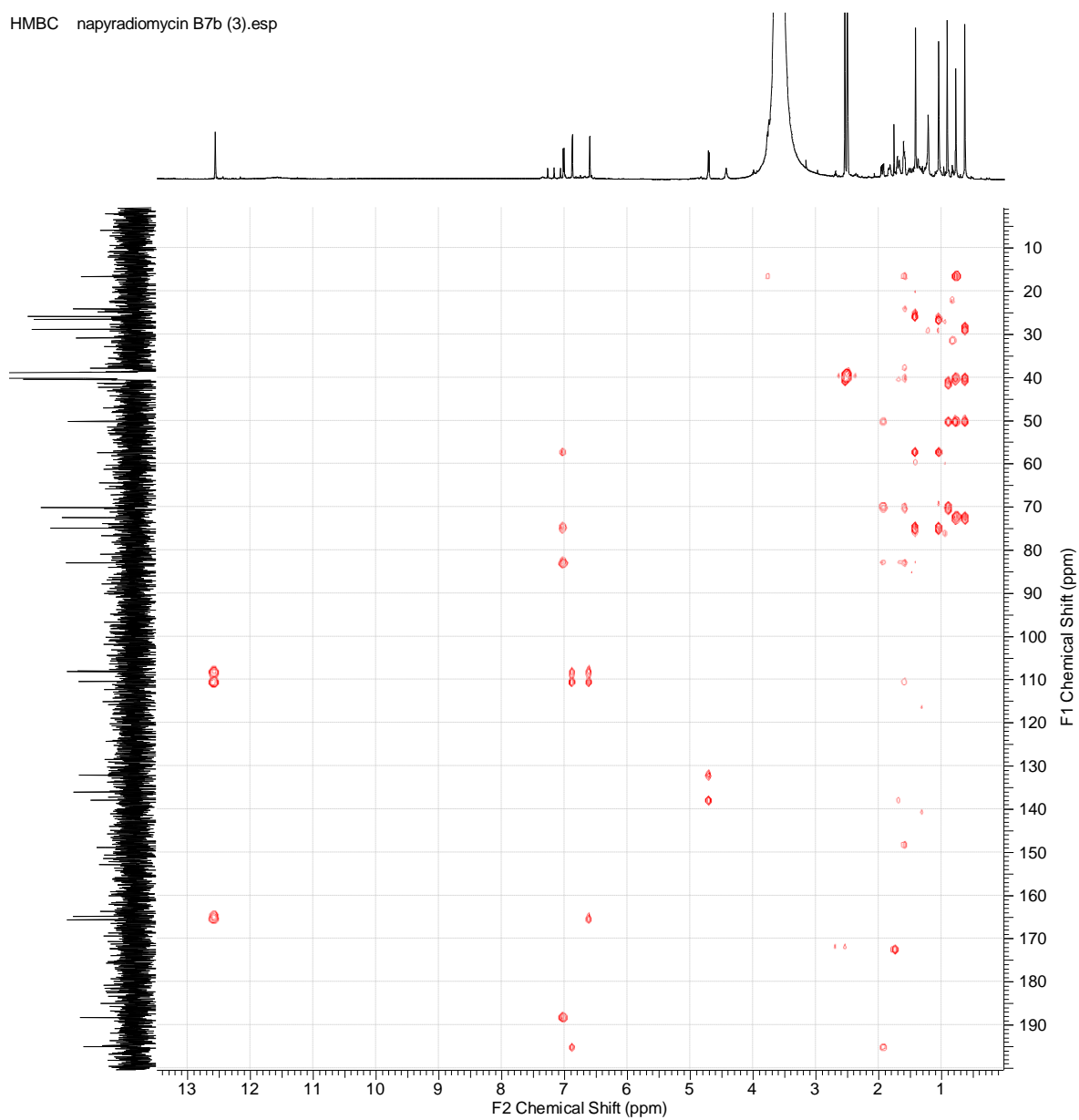
Figure S23. HMBC of compound **3**.

NOESY   napyradiomycin B7b (3).esp
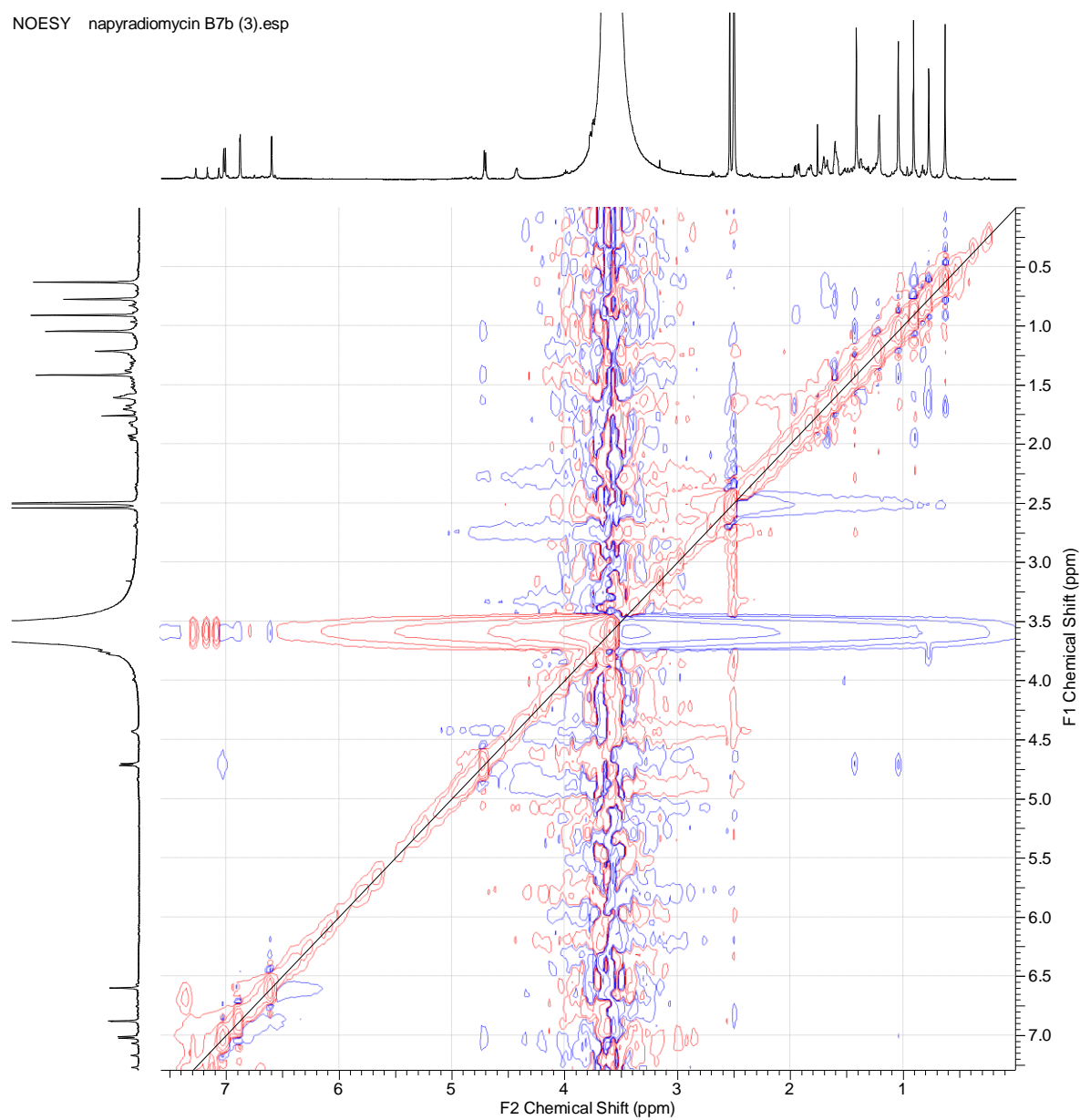
Figure S24. NOESY of compound **3**.

ROESY napyradiomycin B7b (3).esp
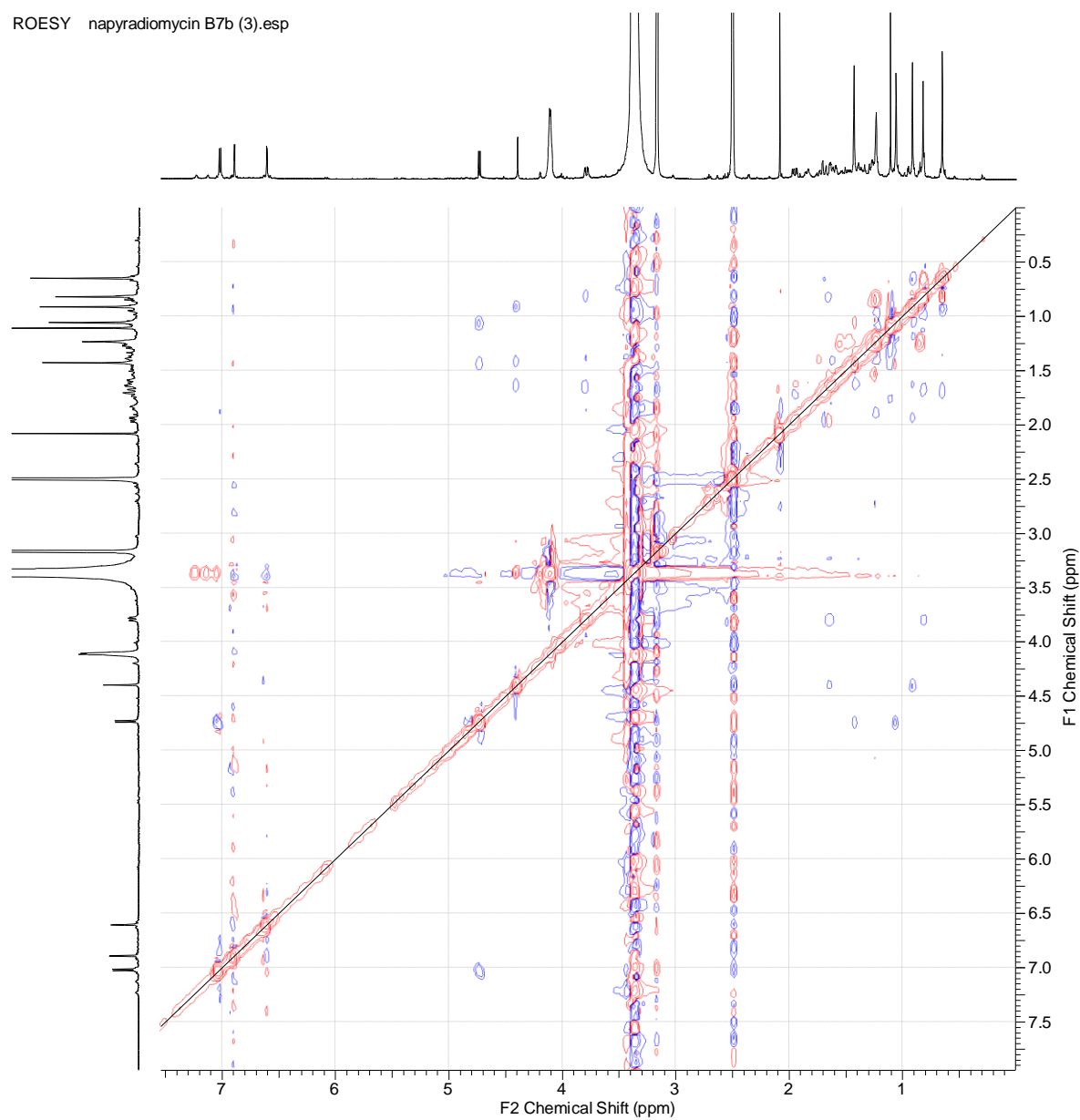
Figure S25. ROESY of compound **3**.
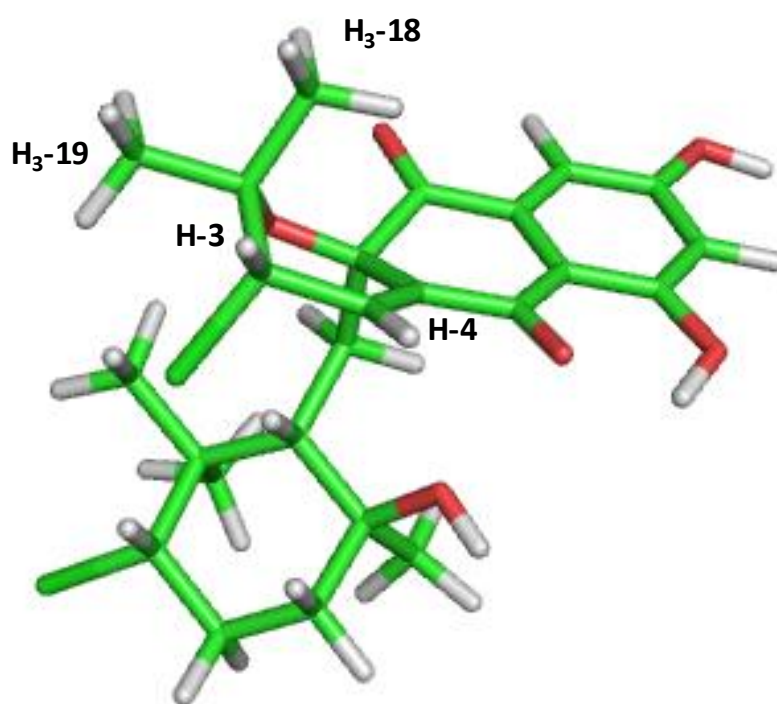
Figure S26. Energy minimized molecular model of compound **3**.

(**A**)

(**B**)

Figure S27. Electrospray-time of flight (ESI-TOF) (**A**) and UV (**B**) spectra of compound **4**.
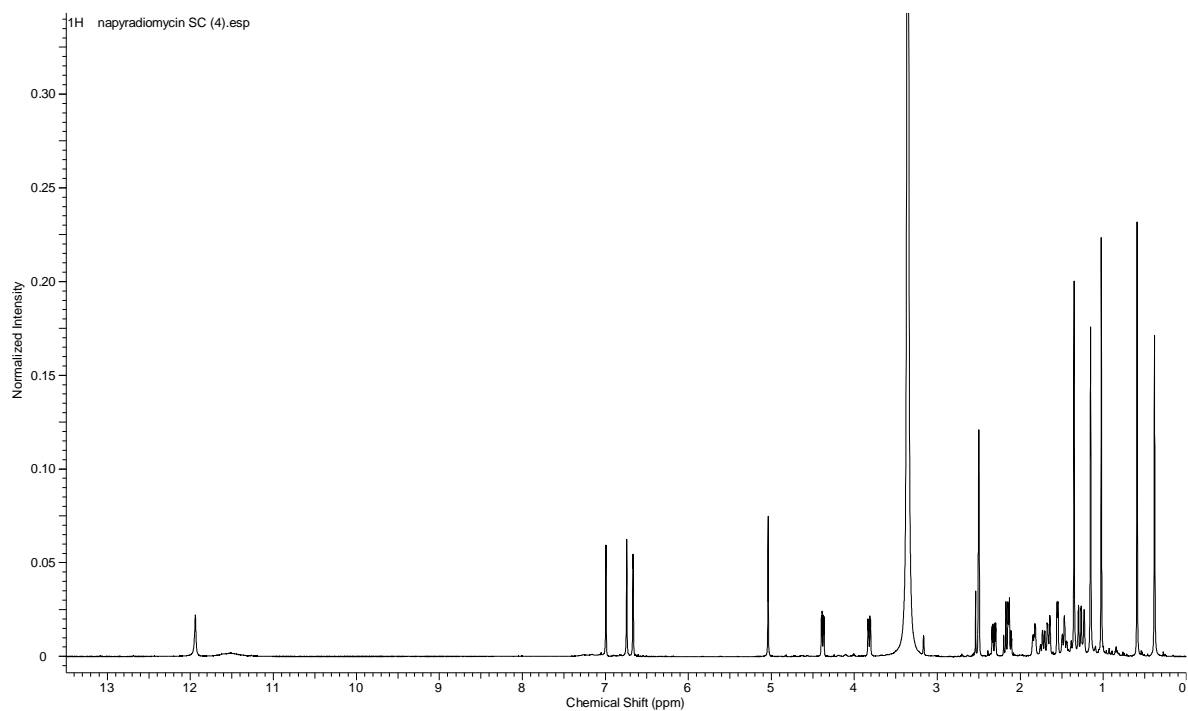
Figure S28. ^{1}H NMR (DMSO-d_6, 500 MHz) of compound **4**.

Figure S29. ^{13}C NMR (DMSO-d_6, 125 MHz) of compound **4**.

COSY   napyradiomycin SC (4).esp

Figure S30. COSY of compound **4**.

HSQC   napyradiomycin SC (4).esp
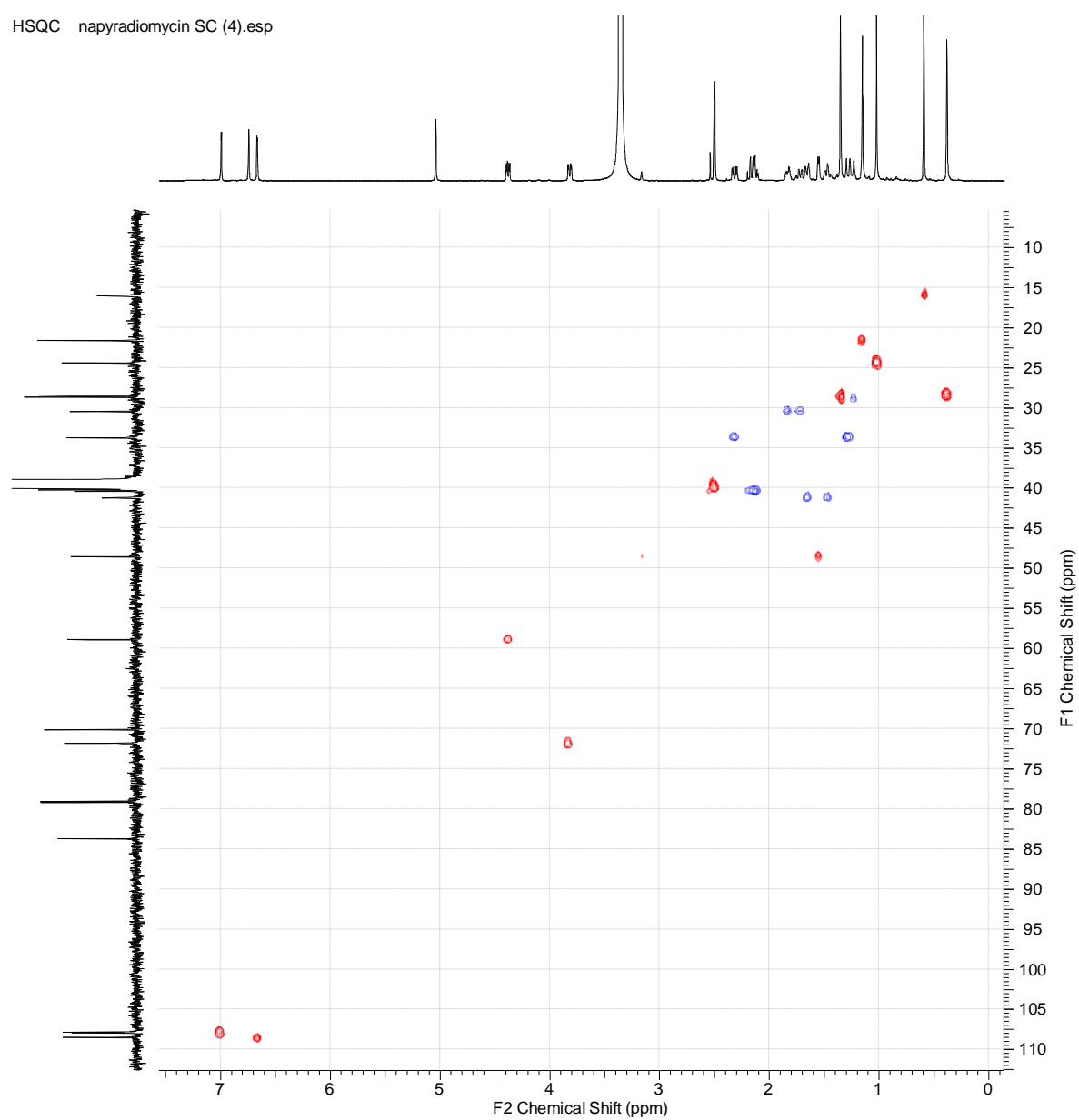
**Figure S31.** HSQC spectrum of compound **4**.

HMBC   napyradiomycin SC (4).esp
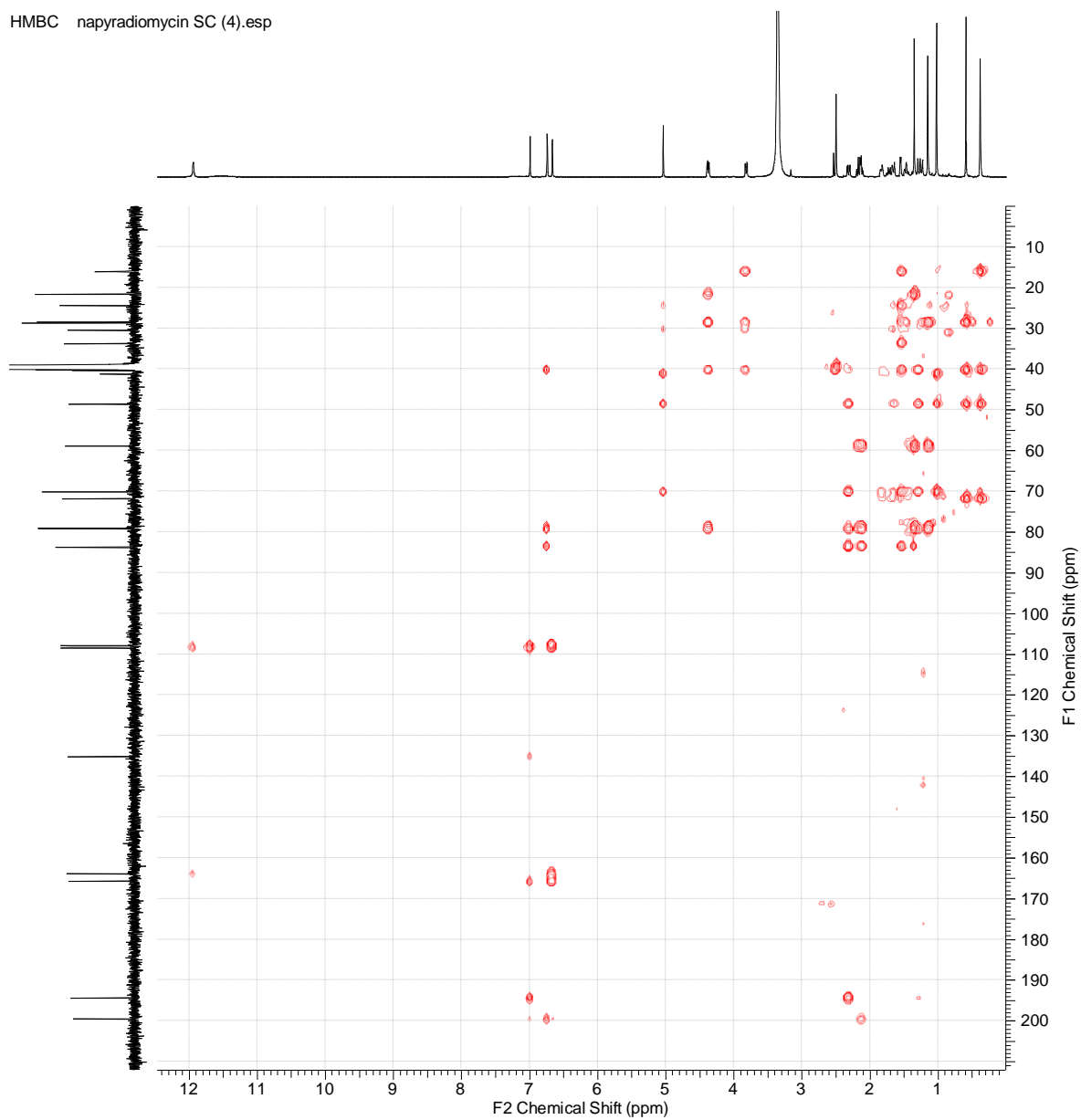
Figure S32. HMBC of compound **4**.

NOESY   napyradiomycin SC (4).esp

F1 Chemical Shift (ppm)

F2 Chemical Shift (ppm)

Figure S33. NOESY of compound **4**.

Figure S34. Energy minimized molecular model of compound **4**.

(**A**)

(**B**)

Figure S35. Electrospray-time of flight (ESI-TOF) (**A**) and UV (**B**) spectra of compound **5**.
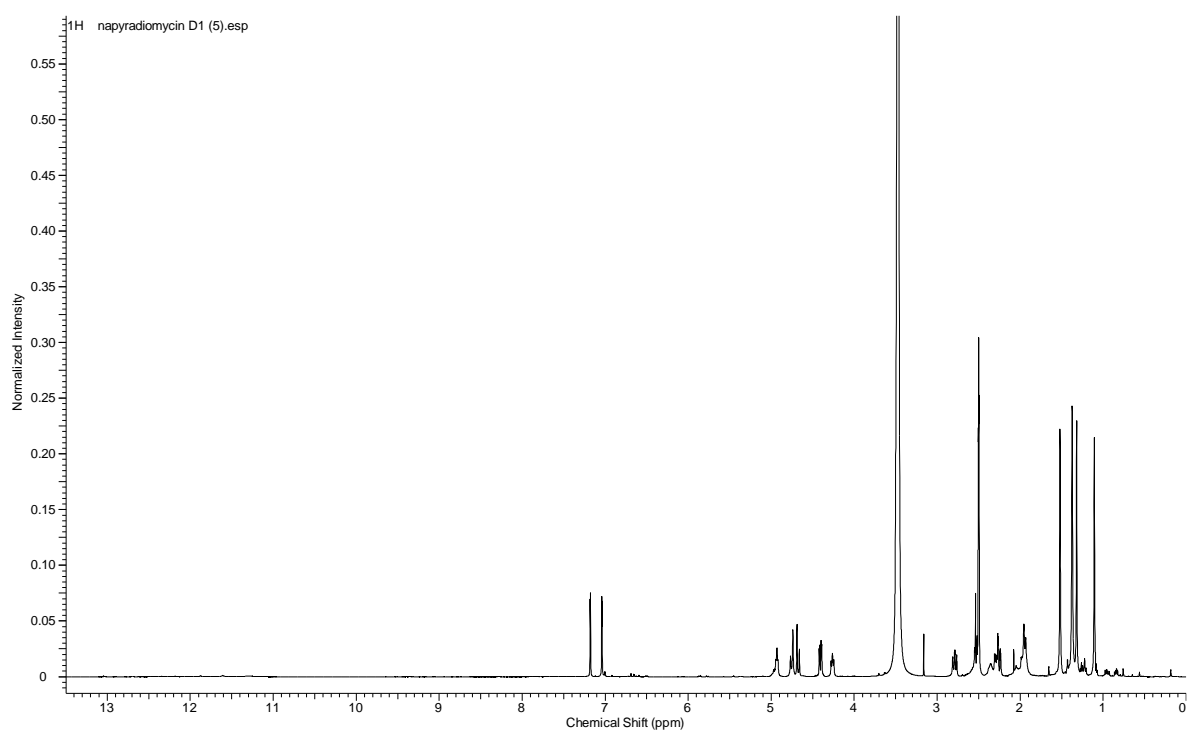
Figure S36. ^{1}H NMR (DMSO-d_6, 500 MHz) of compound **5**.
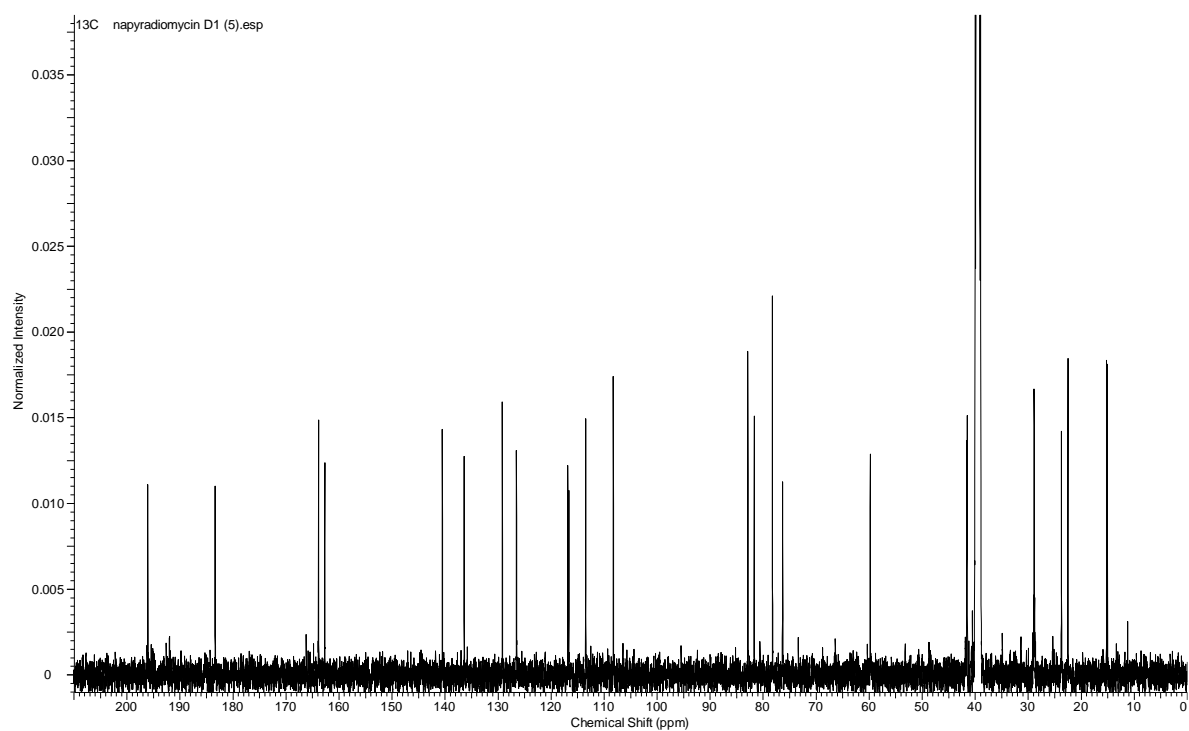
Figure S37. ^{13}C NMR (DMSO-d_6, 125 MHz) of compound **5**.

COSY   napyradiomycin D1 (5).esp
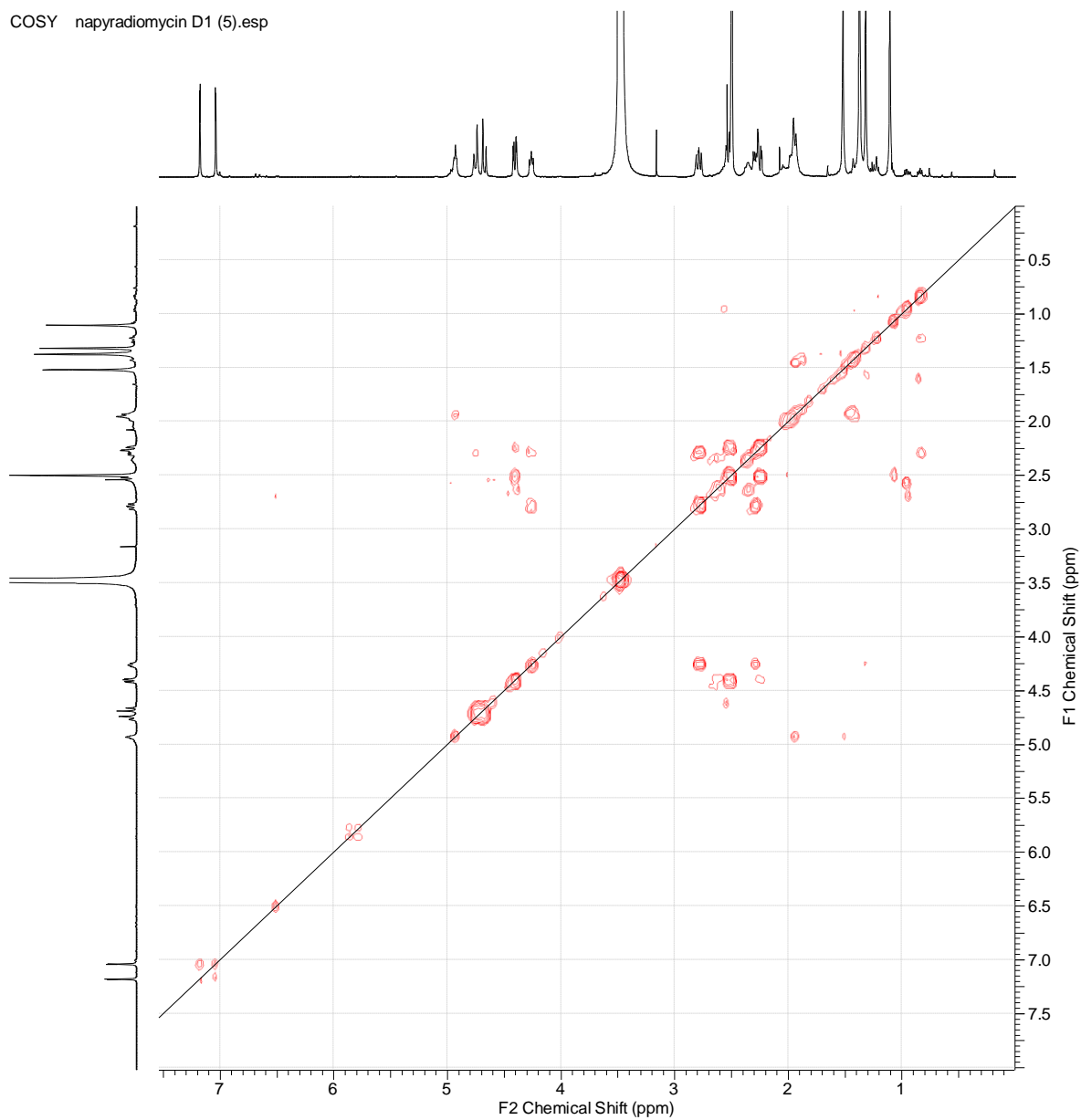
Figure S38. COSY of compound **5**.

HSQC   napyradiomycin D1 (5).esp
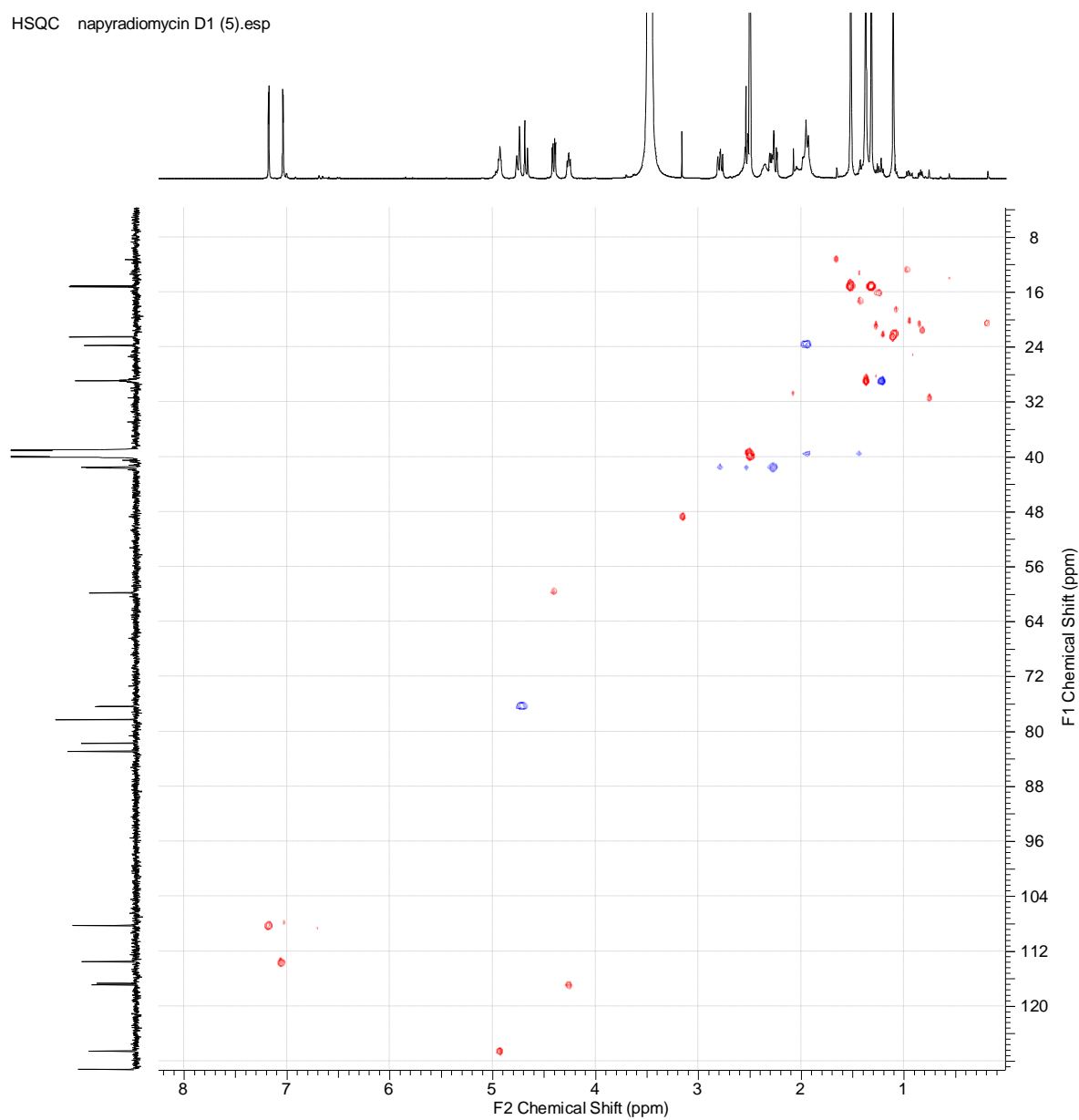
Figure S39. HSQC spectrum of compound **5**.

HMBC   napyradiomycin D1 (5).esp

(a)
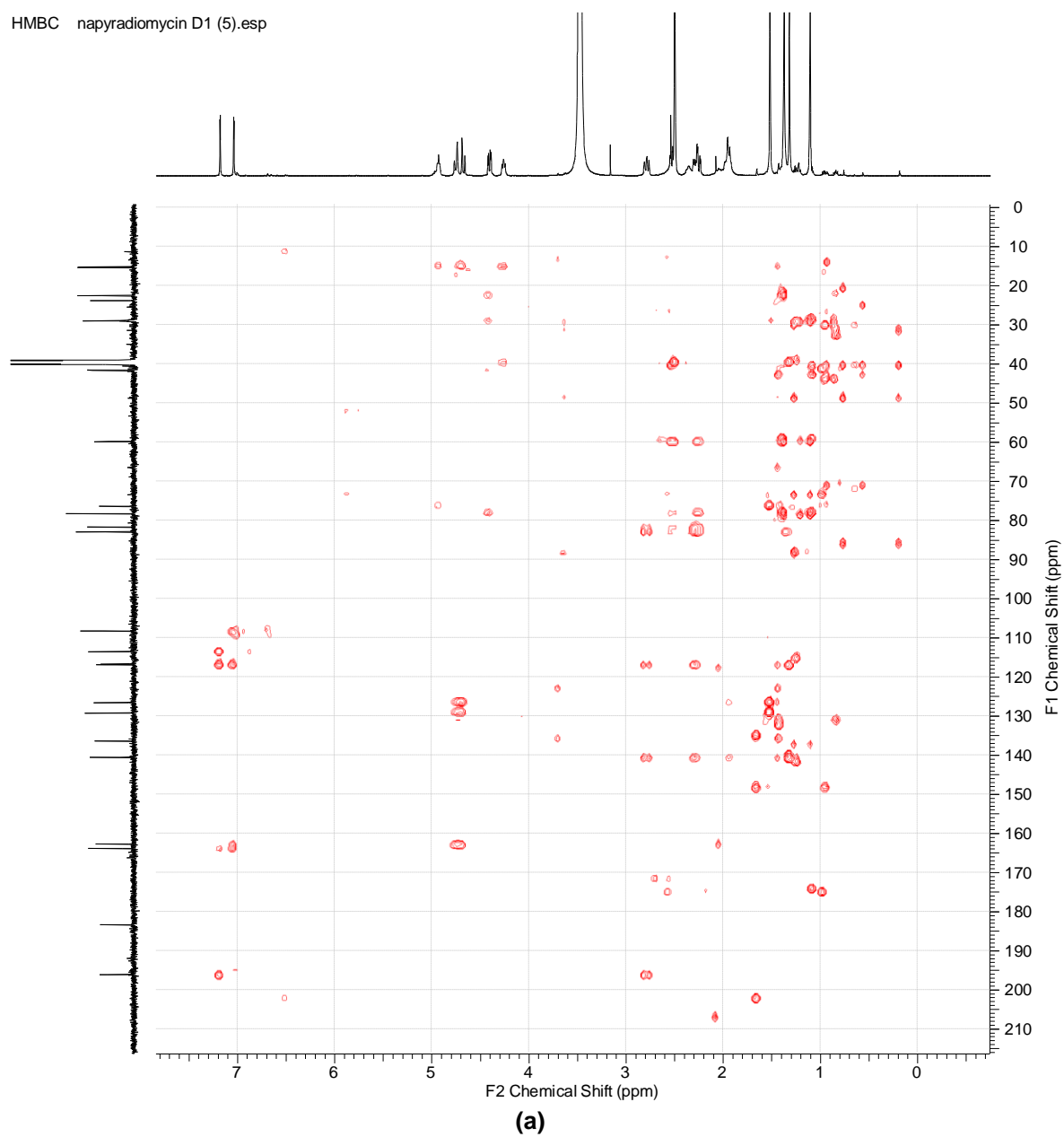
HMBC   napyradiomycin D1 (5).esp
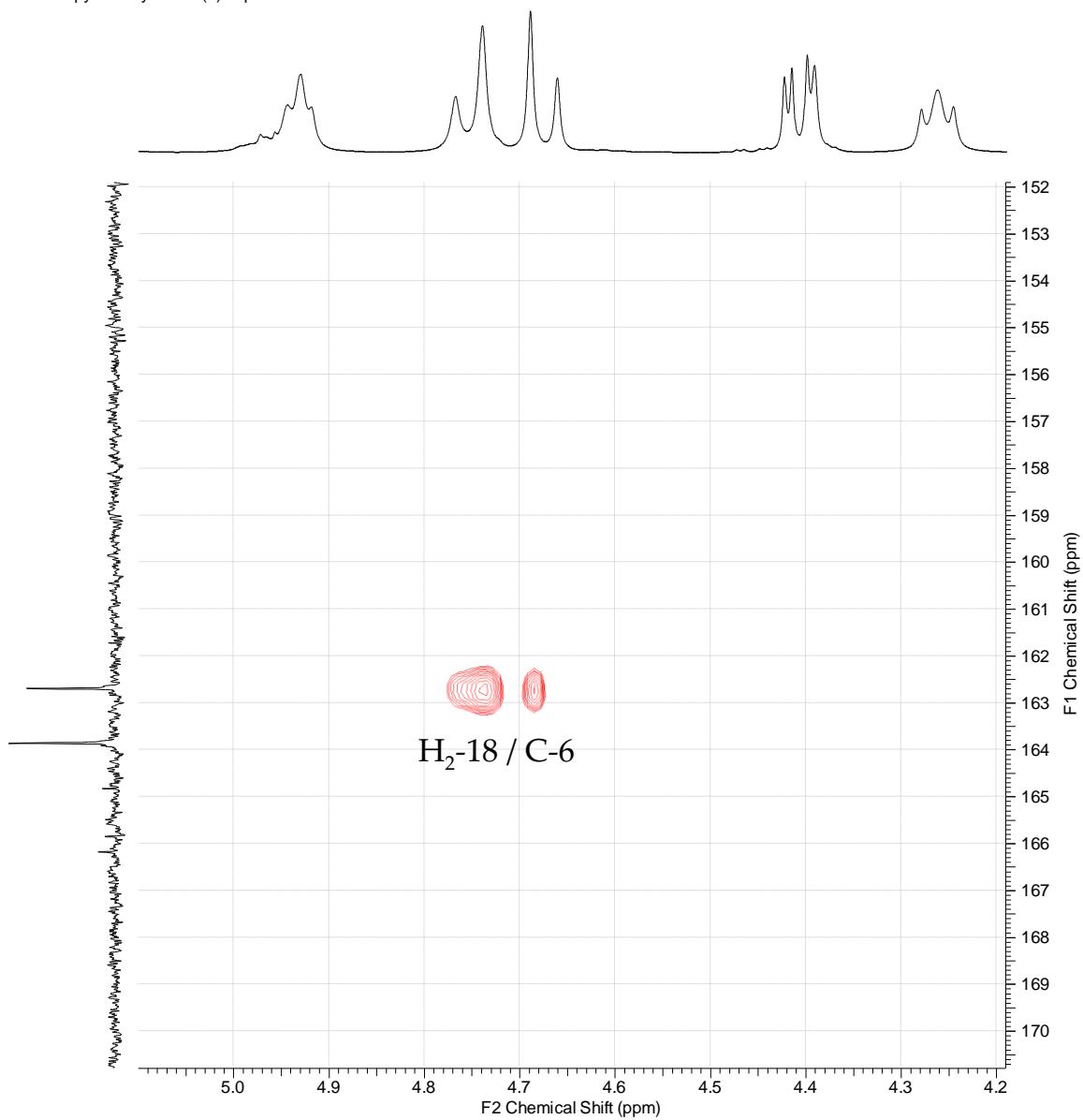
(b)

Figure S40. (a) HMBC of compound **5**. (b) Key HMBC correlation supporting the the etherification of the 18-hydroxy group in **5.**

NOESY napyradiomycin D1 (5).esp
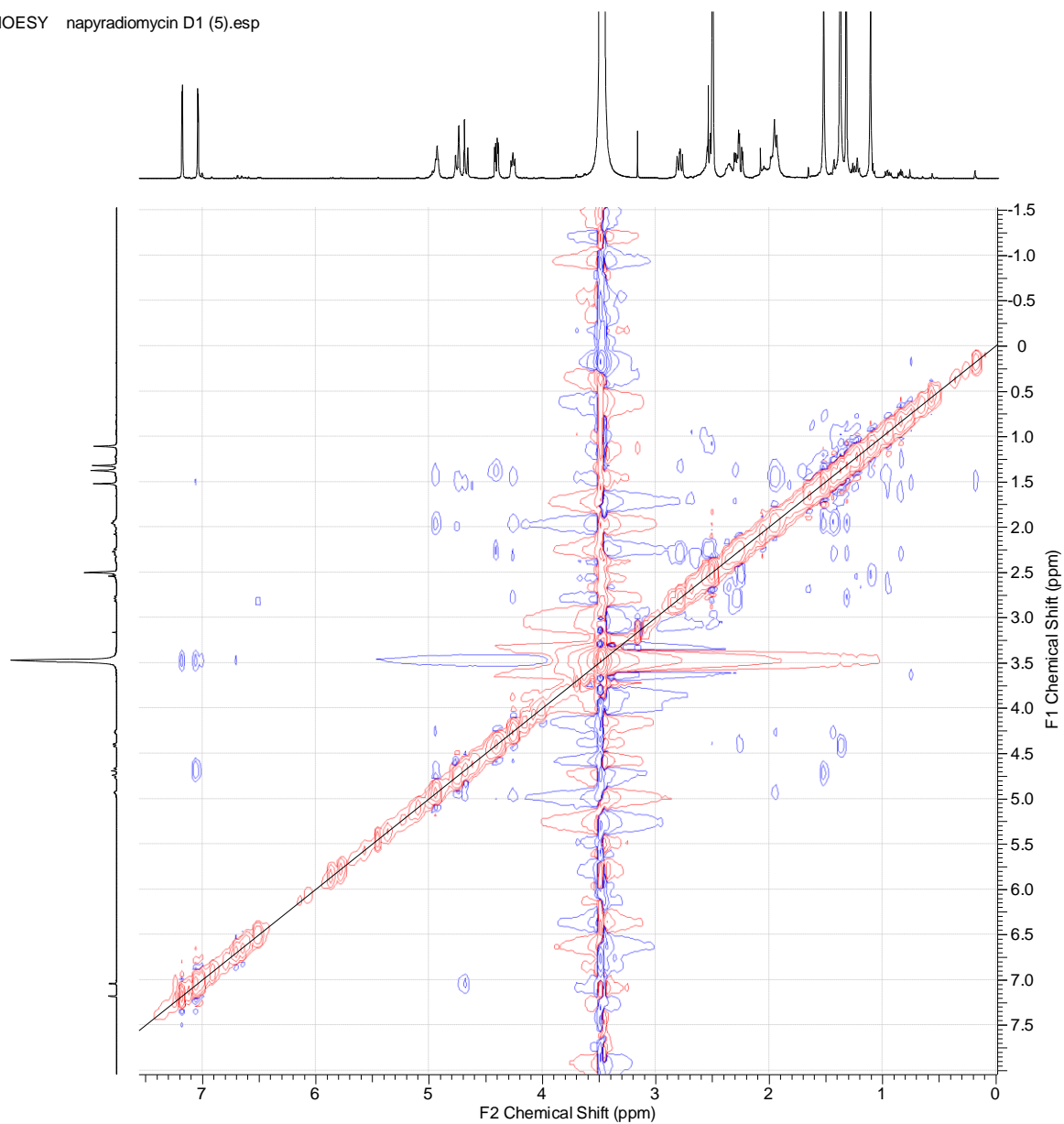
Figure S41. NOESY of compound **5**.
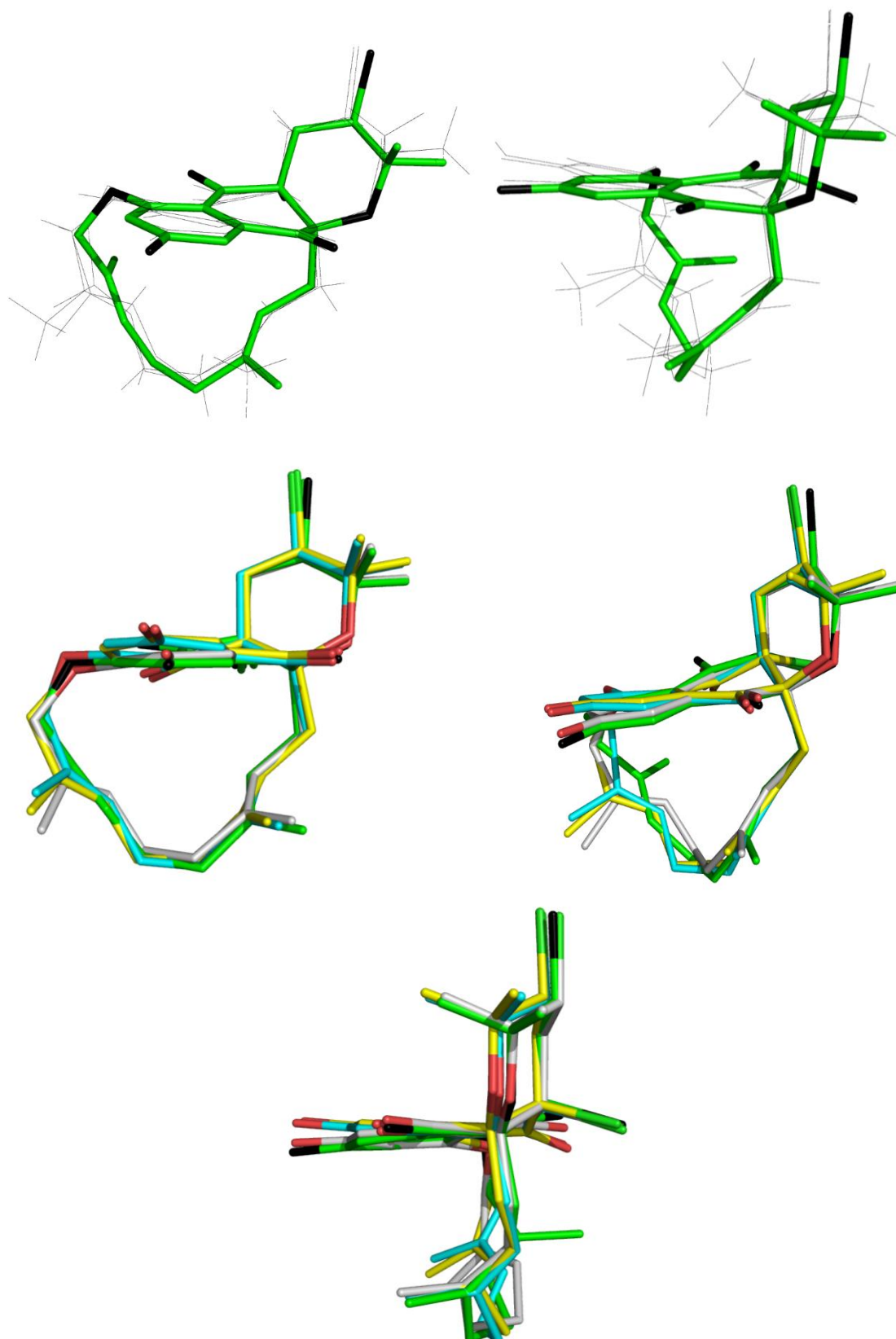
Figure S42. Overlay of different conformers for compound **5.**
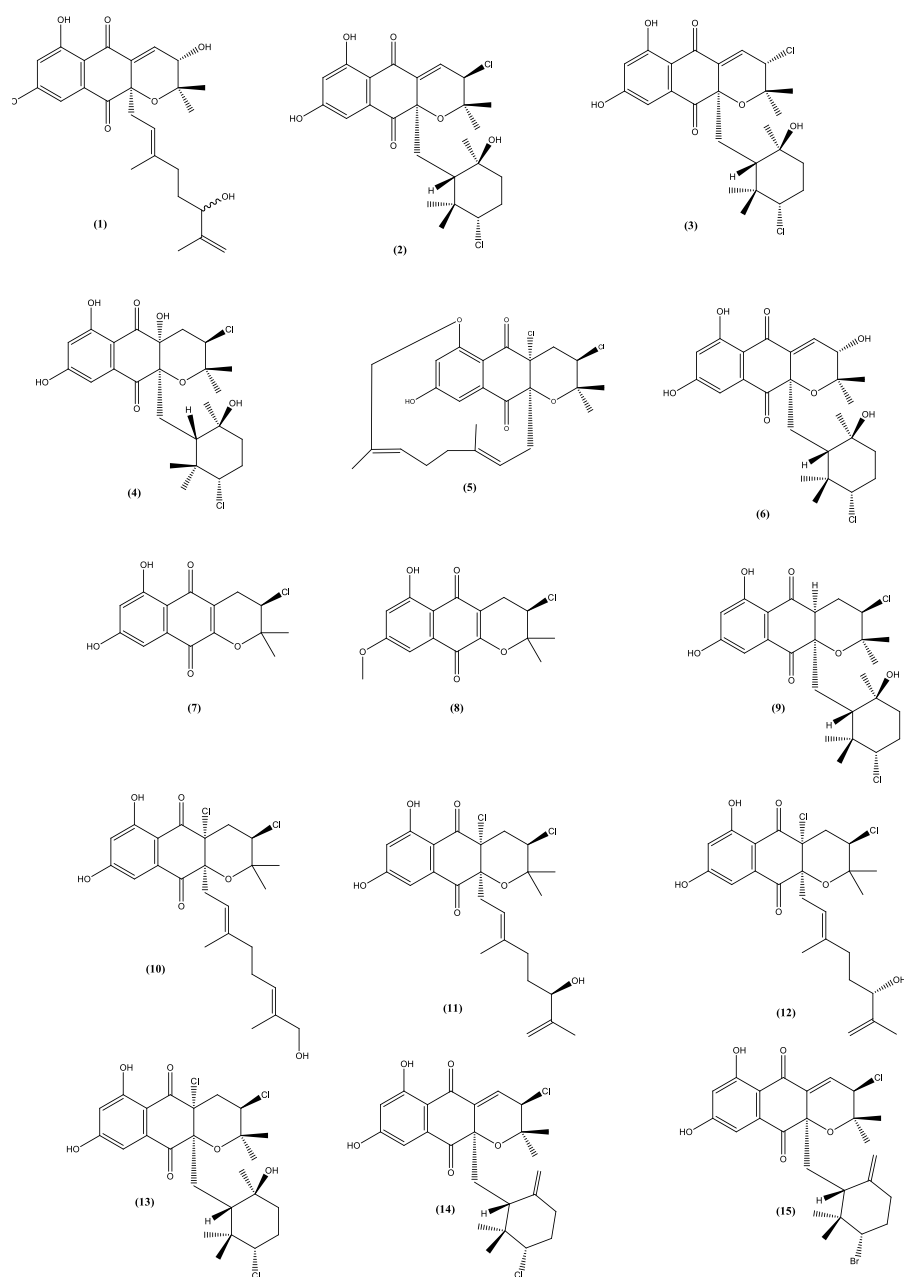
Figure S43. Napyradiomycin metabolites isolated from CA-271078 (**1-15**).

Acknowledgments

Note on structures

All structures were drawn using the program ChemDraw (v.19.1) and then inserted into the text.

Disclaimer

Box 1 presented in this Chapter is complementary work and it is not part of the original peer-reviewed article.
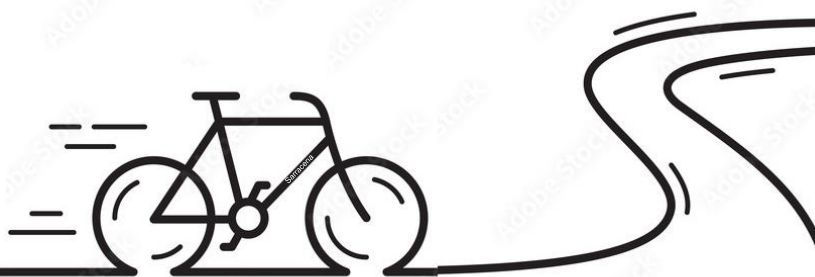
CHAPTER 2

Chapter 2. Pentaminomycins F and G, the first non-ribosomal peptides containing 2-pyridylalanine

Daniel Carretero-Molina,[1,#] Francisco Javier Ortiz-López,[1,#,*] Jesús Martín,[1] Ignacio González,[1] Marina Sánchez-Hidalgo,[1] Fernando Román-Hurtado,[1] Caridad Díaz,[1] Mercedes de la Cruz,[1] Olga Genilloud[1] and Fernando Reyes[1,*]

Resumen: Pentaminomicinas F-H (**1-3**), un grupo de tres nuevos pentapéptidos cíclicos que contienen hidroxiarginina, fueron aisladas de cultivos de *Streptomyces cacaoi* subsp. *cacaoi* junto con las pentaminomicinas A-E conocidas. Las estructuras de estos nuevos péptidos se determinaron mediante una combinación de espectrometría de masas y RMN y análisis de Marfey. Entre ellos, se demostró que las pentaminomicinas F (**1**) y G (**2**) contienen en sus estructuras el aminoácido singular 3-(2-piridil)-alanina (L-2-PAL). Este hallazgo representa el primer ejemplo reportado de péptidos no ribosomales que contienen este residuo. La secuencia quiral LDLLD determinada para los tres nuevos compuestos estaba de acuerdo con la descrita para las pentaminomicinas previamente aisladas y era consistente con los dominios de epimerización presentes en el posible grupo de genes biosintéticos de la péptido sintetasa no ribosomal (NRPS).

Affiliations

[1] Fundación MEDINA, Centro de Excelencia en Investigación de Medicamentos Innovadores en Andalucía, Avda. del Conocimiento 34, 18016, Armilla (Granada), Spain; daniel.carretero@medinaandalucia.es (D.C.-M.); jesus.martin@medinaandalucia.es (J.M.); ignacio.gonzalez@medinaandalucia.es (I.G.); marina.sanchez@medinaandalucia.es (M.S.-H.); fernando.roman@medinaandalucia.es (F.R.-H); caridad.diaz@medinaandalucia.es (C.D.); mercedes.delacruz@medinaandalucia.es (M.d.l.C.); olga.genilloud@medinaandalucia.es (O.G.)

[#] D.C.-M. and F.J.O.-L contributed equally to this work.

[*] Correspondence: javier.ortiz@medinaandalucia.es (F.J.O.-L.); fernando.reyes@medinaandalucia.es (F.R.); Tel.: +34-958-993-965 (F.J.O.-L. and F.R.)

Abstract: Pentaminomycins F-H (**1-3**), a group of three new hydroxyarginine-containing cyclic pentapeptides, were isolated from cultures of *Streptomyces cacaoi* subsp. cacaoi strain along with the known pentaminomycins A-E. The structures of the new peptides were determined by a combination of mass spectrometry and NMR and Marfey´s analyses. Pentaminomycins F (**1**) and G (**2**) were shown to contain the rare amino acid 3-(2-pyridyl)-alanine (L-2-PAL). This finding represents the first reported examples of non-ribosomal peptides containing this residue. The LDLLD chiral sequence found for the three compounds was in agreement with that reported for previously isolated pentaminomycins and consistent with the epimerization domains present in the putative non-ribosomal peptide synthetase (NRPS) biosynthetic gene cluster.

Keywords: hydroxyarginine-containing cyclic pentapeptides, *Streptomyces cacaoi*, pentaminomycins, nonribosomal peptide synthetase (NRPS) biosynthetic gene cluster

1. Introduction

Microbial secondary metabolites are a major source of novel chemical entities with privileged structures, many of which have become lead compounds for clinical use as antibiotics, immunosuppressive, or anticancer agents. Among them, cyclic peptides are of great importance due to their structural variety and multiple functions, ranging from therapeutic to ecological roles.[277] Recently, a family of novel cyclic pentapeptides characterized by the presence of N^5-hydroxy-arginine, pentaminomycins A-E (**4-8**), has been uncovered.[278–280] The putative biosynthetic gene cluster (BGC) for both pentaminomycins and the previously known BE-18257 cyclic peptides,[281] was identified in two closely related *Streptomyces cacaoi* strains. This BGC was shown to harbor two different NRPSs, a P450 enzyme accounting for the hydroxylation of the arginine residue and a putative hydrolase presumably responsible for the release and cyclization of the peptides, along with regulatory and transport-related genes. The number and location of the epimerization (E) domains within the two different NRPS proteins were shown to be fully consistent with the DLDDL or the enantiomeric LDLLD chiral sequences in BE-18257 peptides (Box 1, Fig. 4) and pentaminomycins (Box 1, Fig. 5), respectively.[279,280]

Interestingly, a recent work on a *Streptomyces cacaoi* subsp. *cacaoi* strain CA-170360 developed by our group led to the discovery of cacaoidin, a novel glycosylated lanthipeptide with unprecedented structural features[162] which represents the first member of class V lanthipeptides (lanthidins).[195] The access to this strain and its whole genome sequence prompted us to search the BGC of pentaminomycins within the genome of the strain CA-170360 to confirm its presence, compare it with that previously reported, and explore the possibility of producing new cyclic pentapeptides of this interesting family. The gene cluster of pentaminomycins was found in the genome of CA-170360,[290] and the subsequent culture of the strain under different growth conditions allowed the detection of all the pentaminomycins previously described together with additional unreported members of the family. Targeted isolation of these new analogues led to the identification of pentaminomycins F-H (**1-3**) (Fig. 1). Among them, **1** and **2** were revealed as unprecedented non-ribosomal peptides containing the unusual amino acid 3-(2-pyridylalanine) (2-Pal). The isolation and structural elucidation of these three new cyclic pentapeptides are described herein.

2. Results

2.1. Isolation and Taxonomy of the Producing Microorganism

The strain CA-170360 was identified as *Streptomyces cacaoi* subsp. *cacaoi* NBRC 12748[T] based on its 16S rDNA sequence (GenBank accession number MW131875), as described previously.[162] The draft genome sequence of CA-170360 was analyzed with antiSMASH 5.1.2[190], and the putative BGC of pentaminomycins and BE-18257 peptides were searched within it. As expected, a nearly identical BGC containing two NRPS genes (*cppB, cppM*) with high homology to those previously reported, was identified in the genome sequence of CA-170360.[162] These NRPS enzymes were individually linked to the biosynthesis of BE-18257 cyclic peptides (CppB) and pentaminomycins (CppM) based on their E domains, as previously reported.[279,280]

2.2. Extraction, Dereplication, and Isolation

The strain CA-170360 was cultured in three different growth media (MPG, YEME, and R2YE) for 21 days at 28ºC and the culture broths and mycelia were extracted with acetone and processed as described previously.[282] These three extracts were analyzed by LC-DAD-MS and the chromatograms were interrogated for the presence of pentaminomycins and/or BE-18257 peptides. Different components with molecular formulae matching those of some known members of both

families were detected in the three extracts. Thus, cyclic peptides BE-18257A, BE-18257B/C, and pentaminomycin D (**7**) were detected in the acetone extract of YEME (Fig. S1), while that of the MPG culture contained mainly the BE-18257 peptides (Fig. S2). On the other hand, the culture in R2YE medium showed the richest profile in pentaminomycin content (Fig. S3). Interestingly, the analysis of the R2YE extract also revealed the presence of two components with $[M+H]^+$ ions at 719.3980 and 753.3830 *m/z*, which were indicative of compounds having molecular formulae of $C_{36}H_{50}N_{10}O_6$ (Δ -1.11 ppm) and $C_{39}H_{48}N_{10}O_6$ (Δ -0.20 ppm), respectively (Fig. S4). These previously unreported molecular formulae indicated the presence of compounds with a close structural relationship with pentaminomycins but containing an additional nitrogen atom that suggested the presence of a basic amino acid residue not present in previously isolated members of the family. This finding prompted us to isolate these new pentaminomycins and determine their structures. For this purpose, the producing strain was cultivated in a 3L-scale (R2YE, 21 days, 28 ºC) and the broth and mycelia were extracted with acetone as described in the experimental section. After evaporation of the organic solvent, the mycelial debris was discarded by filtration. Considering the basicity of previously known pentaminomycins (pKa was calculated as 9.6 by using the Marvin suite, ChemAxon)[283], the pH of the aqueous broth was therefore adjusted to 10 and extracted with MEK. The organic extract was purified by MPLC (reversed phase C18), and the resulting fractions were analyzed by LC-MS. The more polar fractions containing the two new compounds mentioned above were pooled and repurified by semipreparative RP-HPLC (Fig. S5) to yield pentaminomycins F (**1**) and G (**2**). Pentaminomycins A-E (**4-8**) were also found in different less polar adjacent fractions of this chromatography.
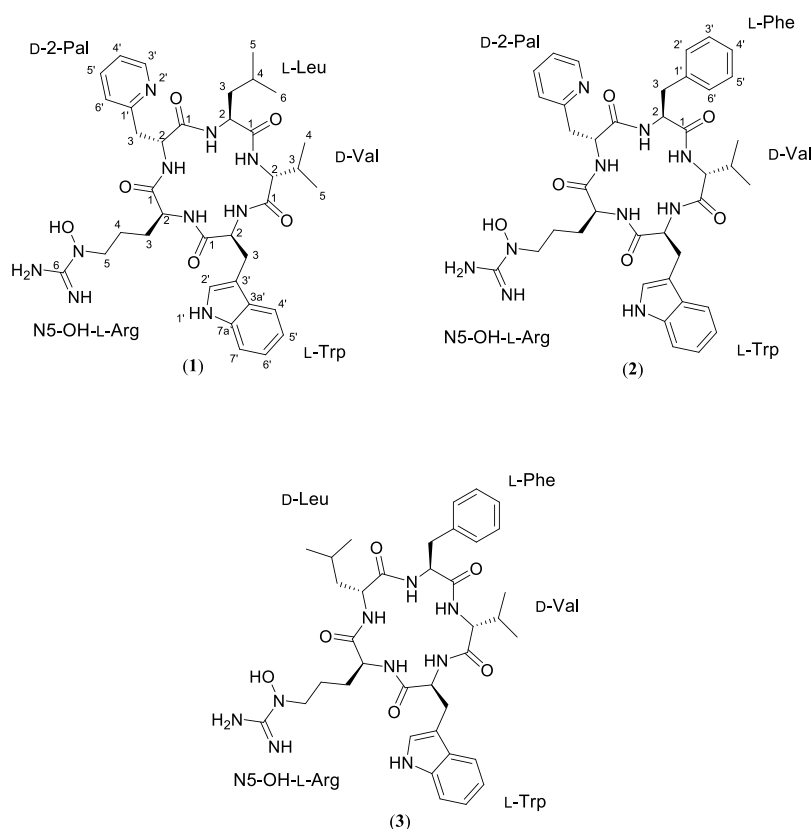

Figure 1. Compounds **1**–**3** isolated from culture broths of *Streptomyces cacaoi* subsp. *cacaoi* CA-170360.

2.3. Structural Elucidation

Compound **1** was isolated as an off-white amorphous solid (1.5 mg) and the molecular formula $C_{36}H_{50}N_{10}O_6$ previously identified was confirmed based on ESI-TOF MS data (Fig. S6). The peptidic nature of **1** was evidenced by the presence of different exchangeable amide NH and α-amino acid hydrogen signals in the ^{1}H-NMR spectrum (Table 1; Fig. S7), as well as of characteristic amide-type carbonyl signals in the ^{13}C-NMR spectrum (Table 1; Fig. S8). Analysis of 2D-NMR spectroscopic data (COSY, HSQC, HMBC, and TOCSY) (Figs. S9-S12) of **1** revealed the presence of valine (Val), tryptophan (Trp), N^5-hydroxyarginine (N5-OH-Arg) and leucine (Leu), as well as of an additional non-proteinogenic aromatic amino acid residue. Along with the similar chemical shift values reported for N5-OH-Arg in the previously reported pentaminomycins A-E,[278–280] the hydroxylation at N-5 of the arginine residue in **1** was supported by two different experimental evidences: the downfield shift of C-5 (δ_C: 50.5 ppm; see Table 1), which strongly suggested the hydroxy group to be located at the adjacent atom (N-5), and the detection of key fragment ions in the HRMS/MS spectrum of **1** (Fig. S13).

Pentaminomycin F (**1**)			Pentaminomycin G (**2**)			Pentaminomycin H (**3**)		
position	δ_C	δ_H mult (J in Hz)	position	δ_C	δ_H mult (J in Hz)	position	δ_C	δ_H mult (J in Hz)
L-Leu 1	172.4, C		**L-Phe** 1	171.2		**L-Phe** 1	171.1, C	
2	50.4, CH	4.41, *dd* (16.0, 8.2)	2	53.0, CH	4.66, *dd* (7.6, 7.3)	2	53.2, CH	4.61, *m*
3	41.4, CH$_2$	1.34, *m*	3	38.4, CH$_2$	2.83, *m*	3	37.4, CH2	2.82, *m*
4	24.3, CH	1.44, *m*			2.77, *m*			
5	22.7, CH$_3$	0.85, *d* (6.5)	1'	137.1, C		1'	137.5, C	
6	22.0, CH$_3$	0.81, *d* (6.5)	2'	129.1, CH	7.16, *m*	2'	129.2, CH	7.21, *m*
NH		7.51, *d* (7.6)	3'	128.0, CH	7.22, *m*	3'	128.0, CH	7.22, *m*
			4'	126.2, CH	7.16, *m*	4'	126.2, CH	7.15, *m*
			5'	128.0, CH	7.22, *m*	5'	128.0, CH	7.22, *m*
			6'	129.1, CH	7.16, *m*	6'	129.2, CH	7.21, *m*
			NH		7.66, *m*	NH		7.93, *d* (9.1)
D-Val 1	171.4, C		**D-Val** 1	171.3		**D-Val** 1	171.2, C	
2	60.1, CH	3.68, *dd* (9.9, 7.4)	2	60.1, CH	3.61, *dd* (9.8, 7.7)	2	59.6, CH	3.72, *dd* (9.3, 7.6)
3	28.2, CH	1.66, m	3	28.2, CH	1.56, *m*	3	29.0, CH	1.57, *m*
4	19.1, CH$_3$	0.75, *d* (6.6)	4	19.3, CH$_3$	0.54, *d* (6.5)	4	18.9, CH$_3$	0.56, *d* (6.5)
5	18.5, CH$_3$	0.33, *d* (6.6)	5	18.8, CH$_3$	0.29, *d* (6.5)	5	18.5, CH$_3$	0.32, *d* (6.5)
NH		8.51, *d* (7.4)	NH		8.38, *d* (7.7)	NH		8.08, *d* (7.6)
L-Trp 1	171.6		**L-Trp** 1	171.6		**L-Trp** 1	171.7, C	
2	55.3, CH	4.30, *m*	2	55.3, CH	4.28, *m*	2	55.5, CH	4.25, *m*
3	27.0, CH$_2$	3.20, *dd* (14.8, 3.3)	3	26.9, CH$_2$	3.18, *m*	3	26.9, CH$_2$	3.17, br *d* (4.3)
		2.92, *m*			2.90, *m*			2.91, *dd* (14.2, 11.8)

	Position	Compound 1 δC	Compound 1 δH	Compound 2 δC	Compound 2 δH	Compound 3 δC	Compound 3 δH
	2-NH		8.60, d (7.9)		8.59, d (8.3)		8.56, d (7.7)
	1´ (NH)		10.76, s		10.76, br s		10.77, s
	2'	123.9, CH	7.17, m	123.9, CH	7.16, m	123.9, CH	7.16, br d (2.2)
	3'	110.2, C		110.2, C		110.2, C	
	3a'	126.9, C		126.8, C		126.9, C	
	4'	117.9, CH	7.51, br d	117.8, CH	7.51, br d	117.9, CH	7.51, br d (7.9)
	5'	118.3, CH	6.98, dd (7.2, 7.2)	118.2, CH	6.97, dd (7.4, 7.3)	118.3, CH	6.98, dd (7.9, 7.3)
	6'	120.8, CH	7.05, br dd (7.6, 7.2)	120.8, CH	7.04, m	120.9, CH	7.04, m
	7'	111.3, CH	7.30, m	111.3, CH	7.29, m	111.4, CH	7.30, br d (8.1)
	7a'	126.9, C		136.1, C		136.2, C	
N5-OH-L-Arg	1	170.3		170.4		170.6, C	
	2	52.7, CH	4.16, dd (14.2, 7.2)	52.7, CH	4.16, m	52.6, CH	4.20, dd (14.4, 7.3)
	3	28.3, CH2	1.54. m	28.2, CH2	1.54, m	28.5, CH2	1.61, m
	4	22.0, CH2	1.33, m	22.0, CH2	1.32, m	22.3, CH2	1.48, m
			1.24, m		1.19, m		
	5	50.5, CH2	3.44, m	50.5, CH2	3.44, m	50.8, CH2	3.52, m
	6	157.5, C		157.4, C		157.6, C	
	2-NH		7.31, m		7.28, m		7.25, m
	N-OH NH2		10.47, s		10.51, s		10.52, s
D-2-Pal / D-2-Pal / D-Leu	1	170.5		170.3		171.1, C	
	2	51.8, CH	4.81, dd (15.1, 8.0)	51.7, CH	4.8, dd (8.0, 7.8)	50.8, CH	4.24, m
	3	36.1, CH2	3.14. m	36.2, CH2	3.09, dd (14.2, 6.6)	37.4, CH2	1.37, m
			2.94, m		2.91, m		
	4					24.1, CH	1.33, m
	5					22.0, CH3	0.73, d (6.2)
	6					22.5, CH3	0.82, d (6.0)
	NH		8.91, d (8.2)		8.91, d (7.6)		8.58, d (7.9)
	1´	157.8, C		157.4, C			
	2´ (N)						
	3´	148.4, CH	8.43, br d (4.2)	148.7, CH	8.42, br d (4.2)		
	4´	121.8, CH	7.19, m	121.6, CH	7.19, m		
	5´	136.1, CH	7.65, ddd (7.6, 7.6, 1.6)	136.3, CH	7.64, m		
	6´	123.9, CH	7.25, br d (7.9)	123.7, CH	7.22, m		

Table 1. ^{1}H NMR (500 MHz in DMSO-d_6) and ^{13}C NMR (125 MHz in DMSO-d_6) data for compounds 1-3.

As mentioned above, compound **1** showed signals accounting for a second aromatic amino acid in the 2D-NMR spectra. (Figs. S9-S12). A spin system in the Total Correlation Spectroscopy (TOCSY) spectrum (Fig. S12) comprising four aromatic proton signals at δ_H 7.25 (H-6'; δ_C: 123.9 according to the HSQC spectrum), 7.65 (H-5'; δ_C: 136.1), 7.19 (H-4'; δ_C: 121.7) and the low-field signal at δ_H 8.43 (H-3'; δ_C: 148.4), along with a quaternary carbon signal at δ_C 157.8 ppm (C-1') in the ^{13}C-NMR spectrum (Table 1), jointly suggested the presence of a 2-substituted pyridine ring, which was further supported by a set of key HMBC correlations (Fig. 2; Fig. S11). Another spin system was identified in the TOCSY spectrum (Fig. S12) consisting of an NH amide hydrogen at δ_H 8.91, along with an α-amino acid hydrogen at δ_H 4.81 (H-2) and one diastereotopic methylene group at δ_H 3.14 and 2.94 (H-3) coupled to carbon signals at δ_C 51.8 (C-2) and 36.1 (C-3), respectively, in the HSQC spectrum (Fig. S10). HMBC correlations from H-2 and H-3 to C-1' at δ_C 157.8, and from H-3 to C-6' at δ_C 123.9, revealed the connection between this spin system and the pyridine ring, thus establishing the presence of the rare 3-(2-pyridylalanine) amino acid (2-Pal) residue in **1**.
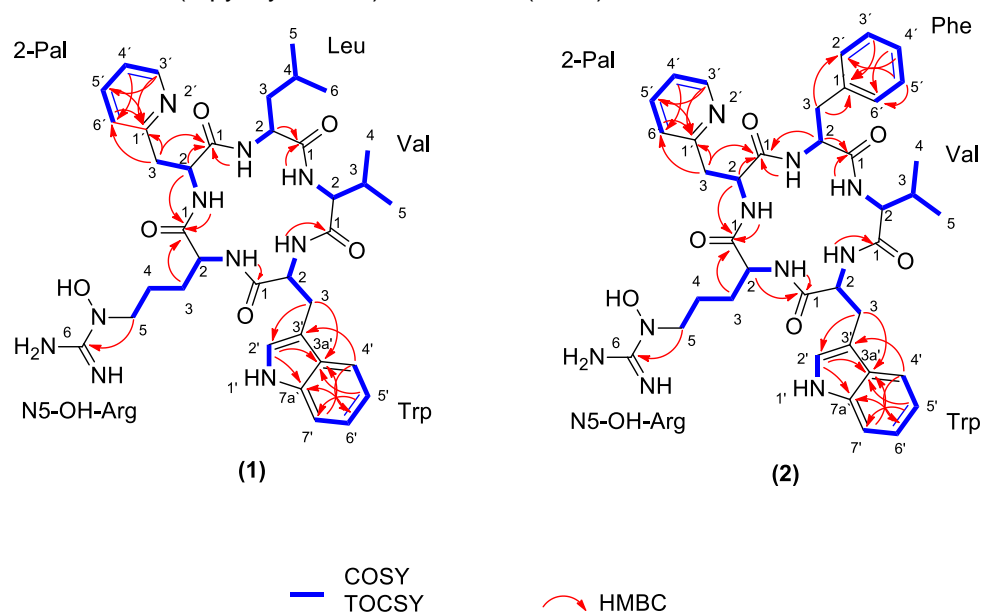
Figure 2. Key COSY/TOCSY and HMBC correlations were observed in the spectra of 1 and 2.

The planar structure of **1** was elucidated based on the key HMBC correlations from NH- to -CO (C-1) within each amide bond (Fig. 2; Fig. S11). HMBC correlations from Trp NH (δ_H 8.60) to Val C-1(δ_C 171.4), from Val NH (δ_H 8.51) to Leu C-1(δ_C 172.4), from Leu NH (δ_H 7.51) to 2-Pal C-1 (δ_C 170.5), from 2-Pal NH (δ_H 8.91) and H-2 (δ_H 4.81) to N5-OH-Arg C-1 (δ_C 170.3) and from N5-OH-Arg NH (δ_H 7.31) to Trp C-1 (δ_C 171.6) allowed to determine the amino acid sequence of **1** as cyclo-(Leu-Val-Trp-N5-OH-Arg-2-Pal). This sequence was in agreement with the molecular formula of **1** and was furtherly supported by HRMS/MS fragmentation (Fig. 3)
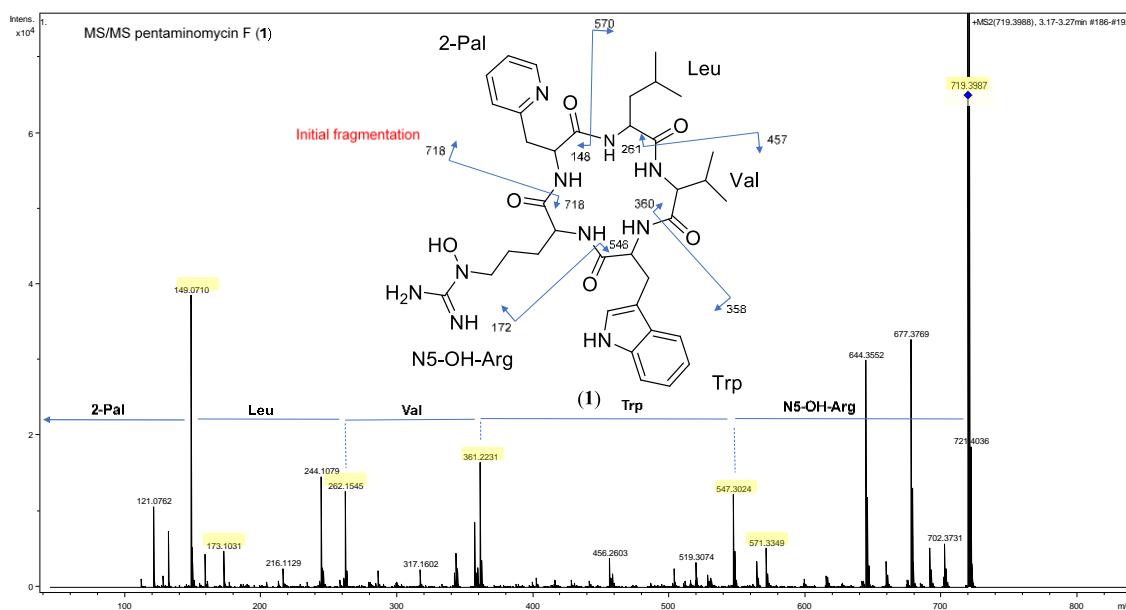
Figure 3. HRMS/MS spectrum and amino acid sequencing of **1**.

The absolute configuration of the amino acid residues in **1** was determined by advanced Marfey´s analysis[284] under two different conditions: A sample of **1** was hydrolyzed (HCl 6 N, 110 ºC) in the presence of 5% thioglycolic acid to prevent Trp degradation.[285] In parallel, the hydroxy group of N5-OH-Arg was removed by using hydriodic acid (HI) for the hydrolysis, as previously reported.[286] Each hydrolysate was separately derivatized with L-FDVA, analyzed by HPLC, and the retention times were compared to those of standard amino acid derivatives. The results jointly showed that the absolute configurations of the amino acids in **1** were L-Leu, D-Val, L-Trp, N5-OH-L-Arg (detected as L-Arg in the second hydrolysate with HI), and D-2-Pal (Figs. S32-S38). Compound **1** was therefore elucidated as cyclo-(L-Leu-D-Val-L-Trp-N5-OH-L-Arg-D-2-Pal) and named pentaminomycin F. The LDLLD chiral sequence found in **1** is, not surprisingly, the same as that reported for pentaminomycins A-E, some of which were also detected in the culture broth.

Compound **2** was isolated as an amorphous solid (1.3 mg) and its ESI-TOF MS data confirmed the molecular formula $C_{39}H_{48}N_{10}O_6$ previously assigned (Fig. S14). Combined analysis of 1D and 2D NMR set data (Table 1; Figs. S15-S20) identified Val, Trp, N5-OH-Arg, Phe, and 2-Pal as the constituent amino acid residues in **2**, thus advancing the close relatedness to **1.**

Detailed analysis of HMBC correlations (Fig 1; Fig. S19) confirmed the replacement of Leu by Phe (concerning **1**) and established its planar structure as cyclo-(Phe-Val-Trp-N5-OH-Arg-2-Pal), which was in concordance with the HRMS/MS data (Fig. S21) The absolute configuration of the amino acid residues in **2** was determined by advanced Marfey's analysis (Figs. S39-S43) and the same LDLLD chiral sequence was found. The structure of **2** was established as cyclo-(L-Phe-D-Val-L-Trp-N5-OH-L-Arg-D-2-Pal) and the compound named pentaminomycin G.

Remarkably, although the non-proteinogenic amino acid 3-(2-pyridyl)-L-alanine (also named 2´aza-L-phenylalanine) has been previously isolated from cultures of *Streptomyces sp.* SF2538,[287] the presence of 2-Pal (either in its L or D form) as an amino acid residue within a peptide is reported herein for the first time, thus highlighting the structural novelty of pentaminomycins F and G.

Further, LC-HRMS-ESI and tandem mass spectrometry analysis of the MPLC fractions containing the known pentaminomycins A-E (**4**-**8**) revealed the presence of a new compound (**3**) isobaric to pentaminomycin C (**6**). Detection of different key fragments ions in the HRMS/MS

spectrum of compound **3** in comparison with those of pentaminomycin C (**6**) suggested a different order in the amino acids sequence (Fig. S22). The new peptide (**3**) was isolated by semipreparative HPLC along with pentaminomycins A-E (**4-8**) (Figs. S23, S59). Interpretation of 2D NMR data of compound **3** (Table 1; Figs. S25-S31) determined **3** as a sequence isomer of pentaminomycin C (**6**), in which the Leu and Phe residues were interchanged in the peptide backbone. The planar structure of **3** was then established as cyclo-(Phe-Val-Trp-N5-OH-Arg-Leu). The absolute configurations of the amino acid residues in **3** were determined by advanced Marfey´s analysis (Figs. S44-S48), and **3**, named pentaminomycin H, was elucidated as cyclo-(L-Phe-D-Val-L-Trp-N5-OH-L-Arg-D-Leu)

2.4. Evaluation of antimicrobial activity

Finally, compounds **1**–**8** were evaluated for their antibacterial properties against a panel of Gram-negative pathogens such as *Escherichia coli* ATCC 25922, *Acinetobacter. baumannii* MB5973, *Klebsiella pneumoniae* ATCC700603 or *Pseudomonas aeruginosa* MB5919. Two of them, compounds **5** and **6**, showed modest bioactivity against a clinical isolate of *A. baumannii*, presenting MIC values of 16-32 μg/Ml but were inactive against *Escherichia coli*, *Klebsiella pneumoniae,* and *Pseudomonas aeruginosa*.

3. Discussion

The co-isolation during this work of three new pentaminomycins (**1-3**) along with the known pentaminomycins A-E (**4-8**) suggests a common biosynthetic origin for all these compounds. According to previous results,[279,280] bioinformatic analysis[288,289] of the adenylation (A) domains of the CppM protein[290] predicted the incorporation of Val and Arg for the second and fourth A domains, while suggesting a polar aromatic amino acid in the third A domain, the latter being consistent with a Trp residue at this position in all pentaminomycins A-H. For the first and fifth A domains, the substrates predicted were in both cases hydrophobic amino acids such as Val, Leu, or Ile, thus advancing the promiscuity of these domains. The variety of residues found at these positions in the backbone of pentaminomycins A-H is experimental evidence of this promiscuity, and the acceptance of 2-Pal as a substrate of the fifth A domain (**1** and **2**) further highlights it.

Regarding the production of pentaminonycins F (**1**) and G (**2**), it may be assumed that the amino acid 2-Pal is biosynthesized, i.e., is not present in the medium, and then recruited by the biosynthetic pathway of pentaminomycins. As mentioned above, this amino acid (in its L form) has been previously isolated from *Streptomyces* sp. SF253,[287] although, nothing is known about its biosynthesis. It has been postulated a type II PKS locus is responsible for the biosynthesis of 2´-aza-L-phenylalanine (i.e., 2-L-Pal) and the related amino acid 2´-aza-L-tyrosine (which would ultimately result in the formation of the aza-β-tyrosine substructure present in the kedarcidin chromophore), but this hypothesis could not be proved.[291] Apart from that set of genes, which were not found anywhere within the genome of CA-170360, there are no other obvious candidate genes. Also, none of the putative proteins encoded in the BGC of pentaminomycins seems to be related to the synthesis of 2-Pal, and its biosynthetic origin remains unclear.

4. Conclusions

In summary, the use of different culture conditions for the strain *Streptomyces cacaoi* subsp. *cacaoi* CA-170360 led to the isolation of the three new pentaminomycins F-H (**1-3**). Of major significance, pentaminomycins F (**1**) and G (**2**) reported herein are the first non-ribosomal peptides incorporating in their structures the rare 3-(2-pyridylalanine) amino acid residue. This work highlights the importance of exploring different conditions to extend the microorganism´s biosynthetic potential and thus have access to novel and privileged structures.

Box 1. Biosynthesis of BE-18257 A–C and pentaminomycins

As demonstrated in this chapter, the strain *Streptomyces cacaoi* CA-170360 can produce two different groups of cyclic pentapeptides, the BE-18257 peptides and the pentaminomycins. These two families of cyclopeptides are synthesized by a pair of non-ribosomal peptide synthetases (NRPSs) adjacently found within the same BGC.[279,280] By applying heterologous expression, our research team has successfully linked these NRPSs to their respective groups of cyclopeptides.[290] Additionally, the role of a standalone enzyme belonging to the Penicillin Binding Protein (PBP) family in the macrocyclization-mediated release structure of the peptides was uncovered.

Despite no other similar BGCs have been found in genome sequence databases beyond *Streptomyces cacaoi* species, this work has paved the way to identify additional tandem biosynthetic genes organized within a single BGC encoding the biosynthesis of related families of compounds and/or new analogs of both BE-18257 antibiotics and pentaminomycins.
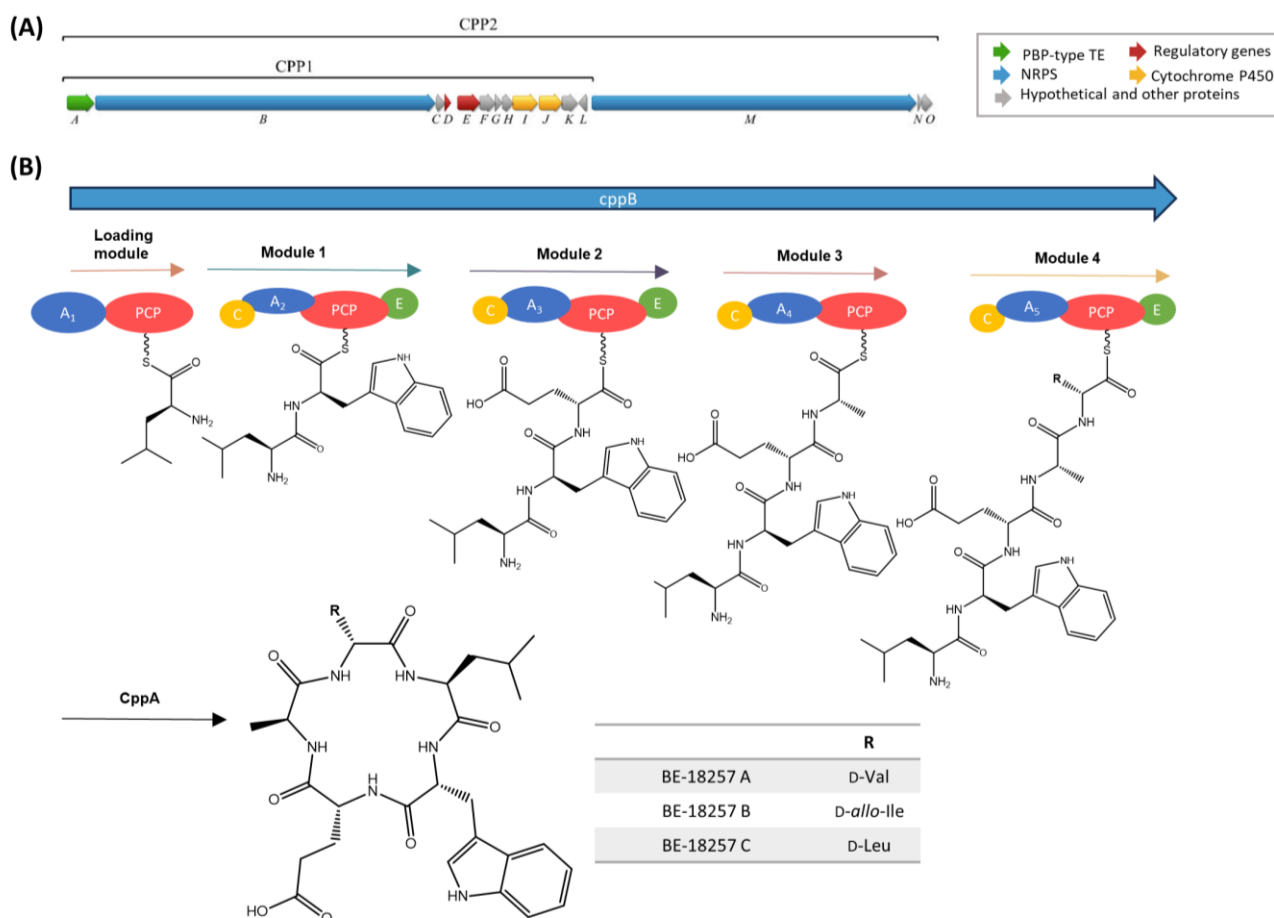
Figure 4. (**A**) The *cpp* biosynthetic gene cluster encodes two NRPS genes (blue), Penicillin Binding Protein (PBP)-type thioesterase (TE) gene (green), two cytochrome P450 (yellow), two regulatory genes (red) and other genes with unknown functions (grey). (**B**) Proposed biosynthetic pathway for the BE-18257 A-C cyclic peptides by the non-ribosomal peptide synthetase CppB modular organization. A1-A5, adenylation domains; PCP, peptidyl carrier protein; C, condensation domain; E, epimerase domain; CppA, PBP-type TE.

Box 1. Biosynthesis of BE-18257 A–C and pentaminomycins
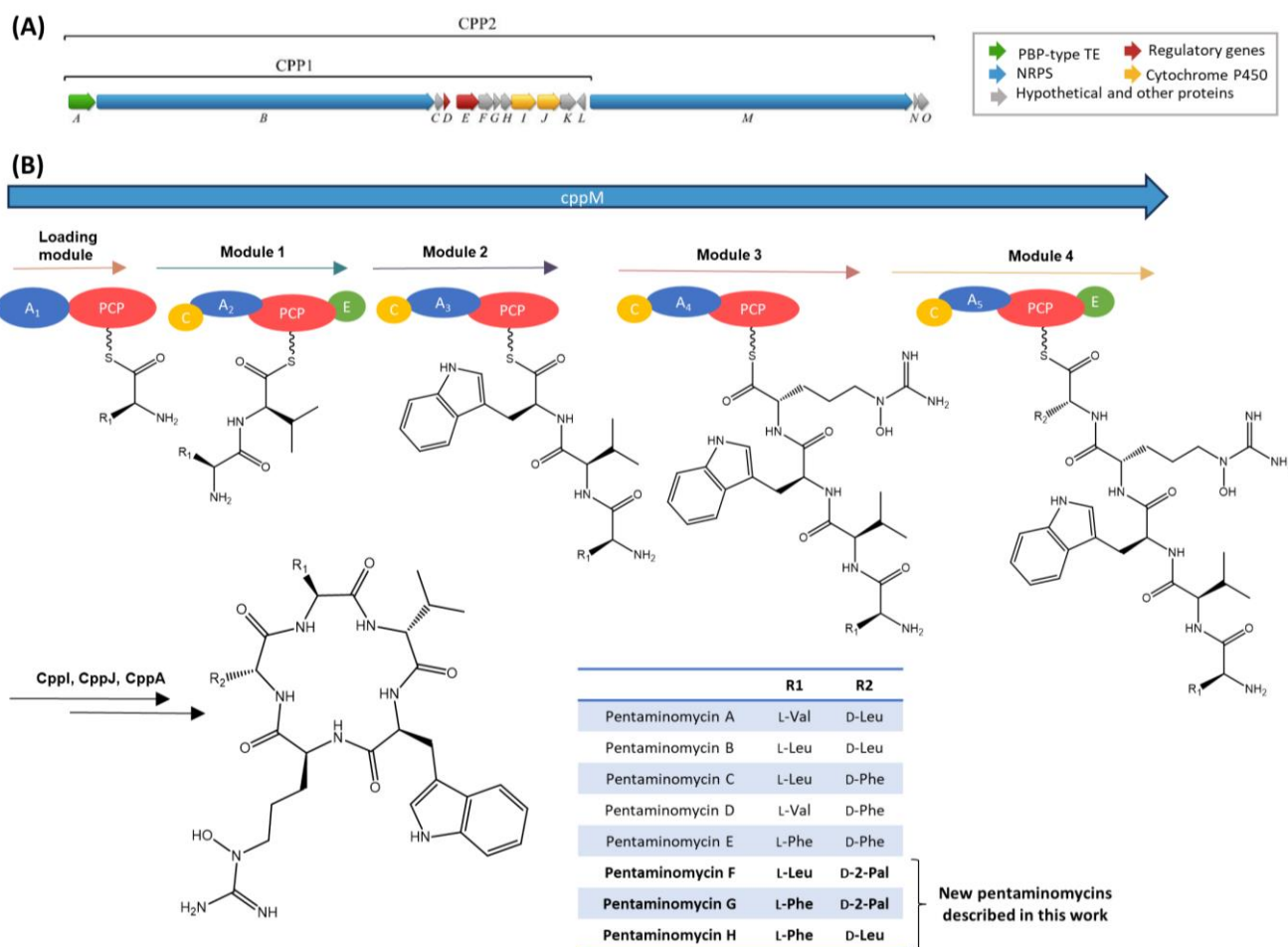
Figure 5. (A) The *cpp* biosynthetic gene cluster encodes two NRPS genes (blue), Penicillin Binding Protein (PBP)-type thioesterase (TE) gene (green), two cytochrome P450 (yellow), two regulatory genes (red), and other genes with unknown functions (grey). **(B)** Proposed biosynthetic pathway for the pentaminomycins A–H by the non-ribosomal peptide synthetase CppM modular organization. A1-A5, adenylation domains; PCP, peptidyl carrier protein; C, condensation domain; E, epimerase domain; CppI and CppJ, cytochromes P450; CppA, PBP-type TE.

5. Experimental Section

5.1. General Experimental Procedures

Optical rotations were measured on a Jasco P-2000 polarimeter. IR spectra were recorded with a JASCO FT/IR-4100 spectrometer equipped with a PIKE MIRacle single reflection ATR accessory. NMR spectra were recorded on a Bruker Avance III spectrometer (500 and 125 MHz for ^{1}H and ^{13}C NMR, respectively) equipped with a 1.7 mm TCI MicroCryoProbe. Chemical shifts were reported in ppm using the signals of the residual solvents as internal reference (δ_H 2.51 and δ_C 39.5 for DMSO-d_6). LC-UV-LRMS analysis was performed on an Agilent 1100 single quadrupole LC-MS system as previously described.[261] ESI-TOF and MS/MS spectra were acquired using a Bruker maXis QTOF mass spectrometer coupled to an Agilent Rapid Resolution 1200 LC. The mass spectrometer was operated in positive ESI mode. The instrumental parameters were 4 kV capillary voltage, drying gas flow of 11 L min^{-1} at 200 °C, and nebulizer pressure of 2.8 bar. TFA-Na cluster ions were used for mass calibration of the instrument before sample injection. Pre-run calibration was done by infusion with the same TFA-Na calibrant. Medium-pressure liquid chromatography (MPLC) was performed on semiautomatic flash chromatography (CombiFlash Teledyne ISCO Rf400×) with a precast reversed-phase column. Semi-preparative HPLC separation was performed on Gilson GX-281 322H2 with a semi-preparative reversed-phase column (Zorbax SB-C18, 250 × 9.4 mm, 5 μm). Acetone used for extraction was analytical grade. The solvents employed for isolation were all HPLC grade. Chemical reagents and standards were purchased from Sigma-Aldrich.

5.2. Strain identification

The taxonomic identification of the strain was reported previously.[162] According to the EzTaxon server analysis[292] (http://www.ezbiocloud.net/eztaxon), the strain was identified as *Streptomyces cacaoi* subsp. *cacaoi* NBRC 12748T (similarity of 99.93%) based on its nearly complete 16S rDNA sequence (1410 bp) (GenBank accession number MW131875).

5.3. Culture conditions

The strain CA-170360 was cultivated in EPA vials (10 mL of medium/ 40 mL vial) in three different media: MPG,[162] YEME[293] and modified R2YE,[293] the latter consisting of yeast extract (5 g/L), sucrose (103 g/L), K_2SO_4 (0.25 g/L), $MgCl_2$-$6H_2O$ (10.12 g/L), glucose (10 g/L), casamino acids (0.1 g/L), KH_2PO_4 (0.5%, 1 mL), $CaCl_2$-$2H_2O$ (3.68%, 8 mL), L-proline (20%, 1.5 mL), TES buffer 5.73% adjusted to pH 7.2 (10 mL), trace element solution 0.2 mL, NaOH 1N (0.5mL), trace element solution: $ZnCl_2$ (40 mg/L), $FeCl_3$-$6H_2O$ (200 mg/L), $CuCl_2$-$2H_2O$ (10 mg/L), $MnCl_2$-$4H_2O$ (10 mg/L), $Na_2B_4O_7$-$10H_2O$ (10 mg/L), $(NH_4)_6Mo_7O_{24}$-$4H_2O$ (10 mg/L); The EPA vials were incubated at 28 °C for 21 days in a rotary shaker at 220 rpm and 70% humidity before harvesting. The whole broth of each one was extracted with an equal volume of acetone under shaking for 1h. After evaporation of the organic solvent, the mycelial debris was discarded by centrifugation and filtration. The aqueous broth was then extracted with MEK (2 × 5 mL). The organic extract was evaporated to dryness and a portion of the residue was reconstituted in 20% DMSO: H_2O (2 × WBE, Whole Broth Equivalent units).

5.4. Large-scale fermentation, extraction, and isolation

The producing strain CA-170360 was cultured in a 3L-scale as follows: A first seed culture of the producing organism was prepared by inoculating 10 mL of seed medium ATCC-2 [soluble starch (20 g/L), dextrose (10 g/L), NZ amine EKC (Sigma) (5 g/L), Difco beef extract (3 g/L), Bacto peptone (5

g/L), yeast extract (5 g/L), and CaCO3 (1 g/L), adjusted to pH 7.0 with NaOH before the addition of CaCO$_3$], in a 40 mL tube with 0.5 mL of a frozen inoculum stock of the producing strain and incubating the tube at 28 °C with shaking at 220 rpm for about 48 h. The second seed culture was prepared by inoculating 50 mL of seed medium in two 250 mL flasks with 2.5 mL of the first seed. A 5% aliquot of the second seed culture was transferred to each of the 20 × 500 mL Erlenmeyer flasks containing 150 mL of the production medium R2YE. The 3 L whole broth was then extracted with an equal volume of acetone under continuous shaking for 1h. After evaporating most of the organic solvent (concentration to initial volume) under an N$_2$ stream, the mycelial debris was discarded by centrifugation and vacuum filtration. The resulting aqueous residue was adjusted to pH 10.0 with NaOH 4 N, extracted with MEK (3 × 800 mL), and the extract was rotary evaporated to dryness to afford a crude extract of 0.910 g.

A portion of this MEK extract (250 mg) was loaded onto a Reversed-Phase C18 (Phenomenex Luna) column (32 × 100 mm) that was eluted with an H$_2$O-CH$_3$CN gradient, with both solvents containing 0.1% trifluoroacetic acid (TFA) (25% to 55% ACN in 36 min + 100% ACN in 10 min, 10 mL/min, 18 mL/fraction) to afford 14 fractions. Fractions containing the target compounds were pooled into two different groups according to their LC-UV-MS profiles and evaporated to dryness in a centrifugal evaporator to yield fractions A (18.5 mg) and B (35.2 mg).

Fraction A (18.5 mg) was further chromatographed by semipreparative reverse-phase HPLC (Zorbax SB-C18, 9.4 × 250 mm, 5 µm; 3.6 mL/min, UV detection at 210 and 280 nm) applying a linear H$_2$O-CH$_3$CN gradient (3.6 mL/min; 20-32% CH$_3$CN in 30 min; UV detection at 210 nm), both solvents containing 0.1% trifluoroacetic acid (TFA) yielding **1** (1.5 mg, rt 23 min) and **2** (1.3 mg, rt 27.5 min).

Fraction B (35.2 mg) was chromatographed by semipreparative reversed-phase HPLC (Zorbax SB-C18, 9.4 × 250 mm, 5 µm; 3.6 mL/min, UV detection at 210 and 280 nm) applying a linear H$_2$O-CH$_3$CN gradient (3.6 mL/min; 30-45% CH$_3$CN in 30 min; UV detection at 210 nm), both solvents containing 0.1% trifluoroacetic acid (TFA) yielding **3** (1.6 mg, rt 26.5 min), **4** (0.6 mg, rt 19 min), **5** (0.8 mg, rt 25.5 min), **6** (2.1 mg, rt 27.5 min), **7** (3.8 mg, rt 22 min), and **8** (0.8 mg, rt 28 min).

5.5. Characterization Data

Pentaminomycin F (**1**): $[\alpha]^{25}_D$ -8.7 (*c* 0.30, MeOH); UV (MeOH) λ_{max} 219, 267; IR (ATR) ν_{max} 3286, 2964, 2875, 1667, 1546, 1439, 1203, 1136 cm^{-1}; ^{1}H and ^{13}C NMR data, Table 1; (+)-ESI-TOF-MS *m/z* 719.3992 [M + H]$^+$ (calcd. for C$_{36}$H$_{50}$N$_{10}$O$_6^+$, 719.3993), 360.2031 [M + 2H]$^{2+}$ (calcd. for C$_{36}$H$_{51}$N$_{10}$O$_6^{2+}$, 360.2008).

Pentaminomycin G (**2**): $[\alpha]^{25}_D$ -5.7 (*c* 0.24, MeOH); UV (MeOH) λ_{max} 219, 267; IR (ATR) ν_{max} 3276, 2971, 2936 1668, 1639, 1545, 1438, 1202, 1134 cm^{-1}; ^{1}H and ^{13}C NMR data, Table 1; (+)-ESI-TOF-MS *m/z* 753.3829 [M + H]$^+$ (calcd. for C$_{39}$H$_{48}$N$_{10}$O$_6^+$, 753.3837), 377.1947 [M + 2H]$^{2+}$ (calcd. for C$_{39}$H$_{49}$N$_{10}$O$_6^{2+}$, 377.1930).

Pentaminomycin H (**3**): $[\alpha]^{25}_D$ -0.6 (*c* 0.24, MeOH); UV (MeOH) λ_{max} 219, 279; IR (ATR) ν_{max} 3274, 2959, 2872, 1659, 1635, 1543, 1439, 1202, 1135 cm^{-1}; ^{1}H and ^{13}C NMR data, Table 1; (+)-ESI-TOF-MS *m/z* 718.4042 [M + H]$^+$ (calcd. for C$_{37}$H$_{52}$N$_9$O$_6^+$, 718.4035), 359.7048 [M + 2H]$^{2+}$ (calcd. for C$_{37}$H$_{53}$N$_9$O$_6^{2+}$, 359.7054)

Pentaminomycin A (**4**): UV (MeOH) λ_{max} 219, 279; ^{1}H and ^{13}C NMR data (Figs. S49, S50); (+)-ESI-TOF-MS *m/z* 670.4036 [M + H]$^+$ (calcd. for C$_{33}$H$_{52}$N$_9$O$_6^+$, 670.4035), 335.7051 [M + 2H]$^{2+}$ (calcd. for C$_{33}$H$_{53}$N$_9$O$_6^{2+}$, 335.7054)

Pentaminomycin B (**5**): UV (MeOH) λ_{max} 219, 279; ^{1}H and ^{13}C NMR data (Figs. S51, S52); (+)-ESI-TOF-MS *m/z* 684.4199 [M + H]$^+$ (calcd. for $C_{34}H_{54}N_9O_6^+$, 684.4192), 342.7129 [M + 2H]$^{2+}$ (calcd. for $C_{34}H_{55}N_9O_6^{2+}$, 342.7132)

Pentaminomycin C (**6**): UV (MeOH) λ_{max} 219, 279; ^{1}H and ^{13}C NMR data (Figs. S53, S54); (+)-ESI-TOF-MS *m/z* 1436.8028 [2M + H]$^+$ (calcd. for $C_{74}H_{103}N_{18}O_{12}^+$, 1436.8031), 718.4035 [M + H] $^+$ (calcd. for $C_{37}H_{52}N_9O_6^+$, 718.4036),

Pentaminomycin D (**7**): UV (MeOH) λ_{max} 219, 279; ^{1}H and ^{13}C NMR data (Figs. S55, S56); (+)-ESI-TOF-MS *m/z* 736.3615 [M + CH$_3$OH + H]$^+$ (calcd. for $C_{37}H_{54}N_9O_7^+$, 736.4141), 704.3880 [M + H]$^+$ (calcd. for $C_{36}H_{50}N_9O_6^+$, 704.3879), 352.6968 [M + 2H]$^{2+}$ (calcd. for $C_{36}H_{50}N_9O_6^{2+}$, 352.6954).

Pentaminomycin E (**8**): UV (MeOH) λ_{max} 219, 279; ^{1}H and ^{13}C NMR data (Figs. S57, S58); (+)-ESI-TOF-MS *m/z* 752.3883 [M + H]$^+$ (calcd. for $C_{40}H_{50}N_9O_6^+$, 752.3884), 376.6970 [M + 2H]$^{2+}$ (calcd. for $C_{40}H_{50}N_9O_6^{2+}$, 376.6954).

5.6. Marfey´s Analysis of Compounds 1-3

Samples (300 µg) of compounds **1-3** were separately dissolved in 0.6 mL of 6 N HCl containing 5% (v/v) of thioglycolic acid and heated at 110ºC for 16 h in a sealed vial. Additionally, a second 300 µg batch of compound **1** was dissolved in 0.6 mL of 4 N HI and heated at 150ºC for 20 h in a sealed vial. The crude hydrolysates were evaporated to dryness under a nitrogen stream and each residue was dissolved in 100 µL of water. A 1% (w/v) solution (100 µL) of L-FDVA (Marfey's reagent,N-(2,4-dinitro-5-fluorophenyl)-L-valinamide) in acetone was added to an aliquot (50µL) of a 50 mM solution of each amino acid standard (leucine, phenylalanine, valine, tryptophan, arginine and 3-(2-Pyridyl)-L-alanine, D, L, or DL mixture) and the aqueous solution of each compound hydrolysate. After the addition of 20 µL of 1 M NaHCO$_3$ solution, each mixture was incubated for 60 min at 40 ºC. The reactions were quenched by the addition of 10 µL of 1 N HCl and the crude mixtures were diluted with 700 µL of acetonitrile and analyzed by ESI LC/MS on an Agilent 1100 single Quadrupole LC/MS.

Separations were carried out on an Agilent Zorbax SB-C8 column (2.1 × 30 mm, 5µm), maintained at 40 ºC. A mixture of two solvents, A (10% CH$_3$CN, 90% H$_2$O) and B (90% CH$_3$CN, 10% H$_2$O), both containing 1.3 mM trifluoroacetic acid and 1.3 mM ammonium formate, was used as the mobile phase under a linear gradient elution mode (10-26% B in 6 min, isocratic 26% B for 2 min, 26-28% B in 0.1 min, isocratic 28% B for 2 min, 28-40% B in 1 min, isocratic 40% B for 2 min, 40-100% B in 0.1 min and then isocratic 100% B for 2 min) at a flow rate of 0.3 mL/min. For D and L arginine, the separation was carried out on Waters XBridge C$_{18}$ column (4.6 × 150 mm, 5µm), maintained at 40 ºC. A mixture of two solvents, A (100% H$_2$O) and B (100% CH$_3$CN), both containing 0.1% of trifluoroacetic acid, was used as the mobile phase under a linear gradient elution mode (isocratic 32% B for 8 min, 32-40% B in 27 min, 40-100% B in 9 min and then isocratic 100% B for 2 min) at a flow rate of 1 mL/min. Retention times (min) for the derivatized (L-FDVA) amino acid standards and the observed peaks in the HPLC trace of each L-FDVA-derivatized hydrolysis product under the reported conditions were as follows:

o   Retention times (min) for the derivatized (L-FDVA) standards amino acids (both L- and D- forms) present in **1** were: L-Trp: 9.18, D-Trp: 11.36, L-Val: 7.04, D-Val: 10.37, L-Leu: 8.70, D-Leu: 12.83, L-2-Pal: 3.03, D-2-Pal: 4.59, L-Arg: 5.18, D-Arg: 4.21.

o   Retention times for the observed peaks in the HPLC trace of the L-FDVA derivatized hydrolysis product of **1** were: L-Trp: 9.29, D-Val: 10.50, L-Leu: 8.83, D-2-Pal: 4.64, L-Arg: 5.37. Retention

times (min) for the derivatized (L-FDVA) standards amino acids (both L- and D- forms) present in **2** were: L-Trp: 9.19, D-Trp: 11.34, L-Val: 7.06, D-Val: 10.33, L-Phe: 8.84, D-Phe: 12.12, L-2-Pal: 2.93, D-2-Pal: 4.65.

o   Retention times for the observed peaks in the HPLC trace of the L-FDVA derivatized hydrolysis product of **2** were: L-Trp: 9.31, D-Val: 10.52, L-Phe: 8.95, D-2-Pal: 4.57.

o   Retention times (min) for the derivatized (L-FDVA) standards amino acids (both L- and D- forms) present in **3** were: L-Trp: 9.54, D-Trp: 11.86, L-Val: 7.33, D-Val: 10.82, L-Phe: 9.52, D-Phe: 12.58, L-Leu: 9.17, D-Leu: 13.04.

o   Retention times for the observed peaks in the HPLC trace of the L-FDVA derivatized hydrolysis product of **3** were: L-Trp: 9.64, D-Val: 10.96, L-Phe: 9.42, D-Leu: 13.23.

5.7. Antibacterial Assays

Compounds **1**–**8** were tested in antimicrobial assays for the growth inhibition of Gram-negative pathogens (*E. coli* ATCC 25922, *A. baumannii* MB5973, *K. pneumoniae* ATCC700603, or *P. aeruginosa* MB5919) following previously described methodologies.[273,294] Briefly, each compound was serially diluted in DMSO with a dilution factor of 2 to provide 10 concentrations starting at 96 µg/mL for all the antimicrobial assays. The MIC was defined as the lowest concentration of compound that inhibited ≥ 95% of the growth of a microorganism after overnight incubation. The Genedata Screener software (Genedata, Inc., Basel, Switzerland) was used to process and analyze the data and also to calculate the RZ' factor, which predicts the robustness of an assay.[276] In all experiments performed in this work the RZ' factor obtained was between 0.87 and 0.98.

Note on structures

All structures were drawn using the program ChemDraw (v.19.1) and then inserted into the text.

Disclaimer

Box 1 presented in this Chapter is complementary work and it is not part of the original peer-reviewed article.

6. Supplementary Information

The following Supplementary Information is available online at:
https://pubs.acs.org/doi/10.1021/acs.jnatprod.0c01199.

List of Supplementary Materials:

Figure S1. LC-DAD-MS analysis of the YEME culture broth extract of CA-170360. Detection of pentaminomycins and BE-18257 peptides.

Figure S2. LC-DAD-MS analysis of the MPG culture broth extract of CA-170360. Detection of BE-18257 peptides.
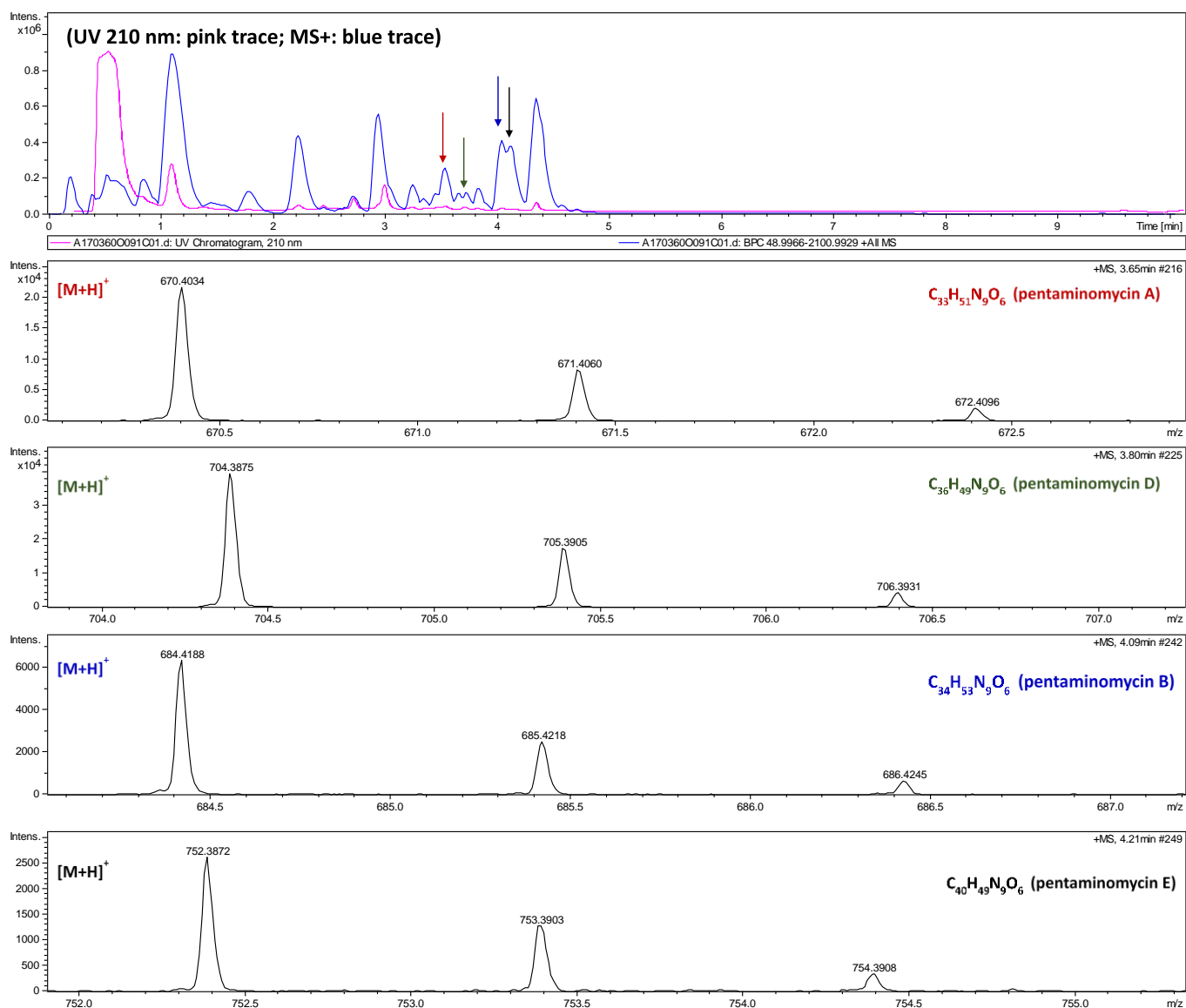
Figure S3. LC-DAD-MS analysis of the R2YE culture broth extract of CA-170260. Detection of pentaminomycins.
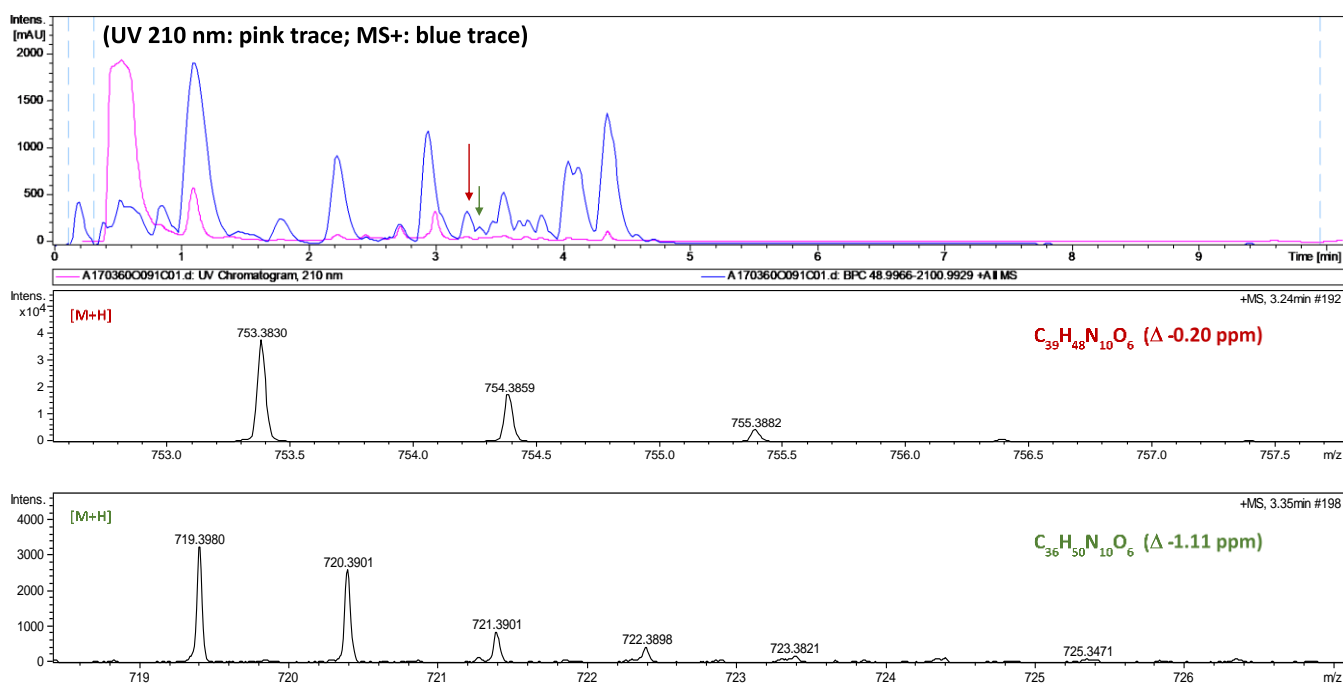
Figure S4. LC-DAD-MS analysis of the R2YE culture broth extract of CA-170360. Detection of unreported molecular formulas related to pentaminomycins.

Figure S5. Targeted isolation of compounds **1** and **2**. LC-UV semipreparative chromatogram was recorded at 210 (blue trace) and 280 nm (orange trace).

136

(A)

(B)
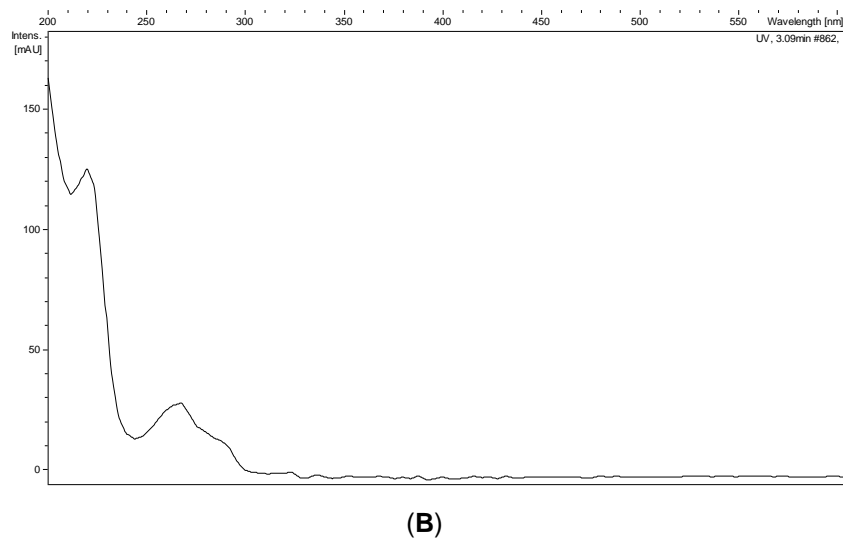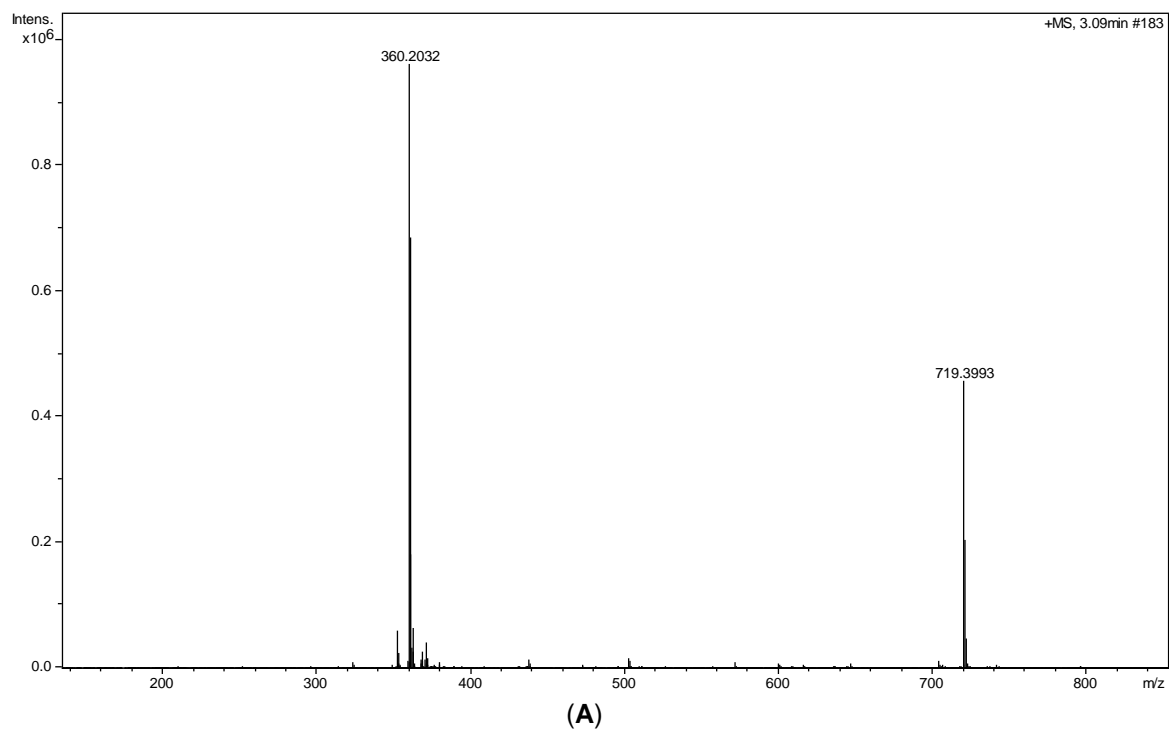
Figure S6. (+)-ESI-TOF (**A**) and UV (**B**) spectra of **1.**
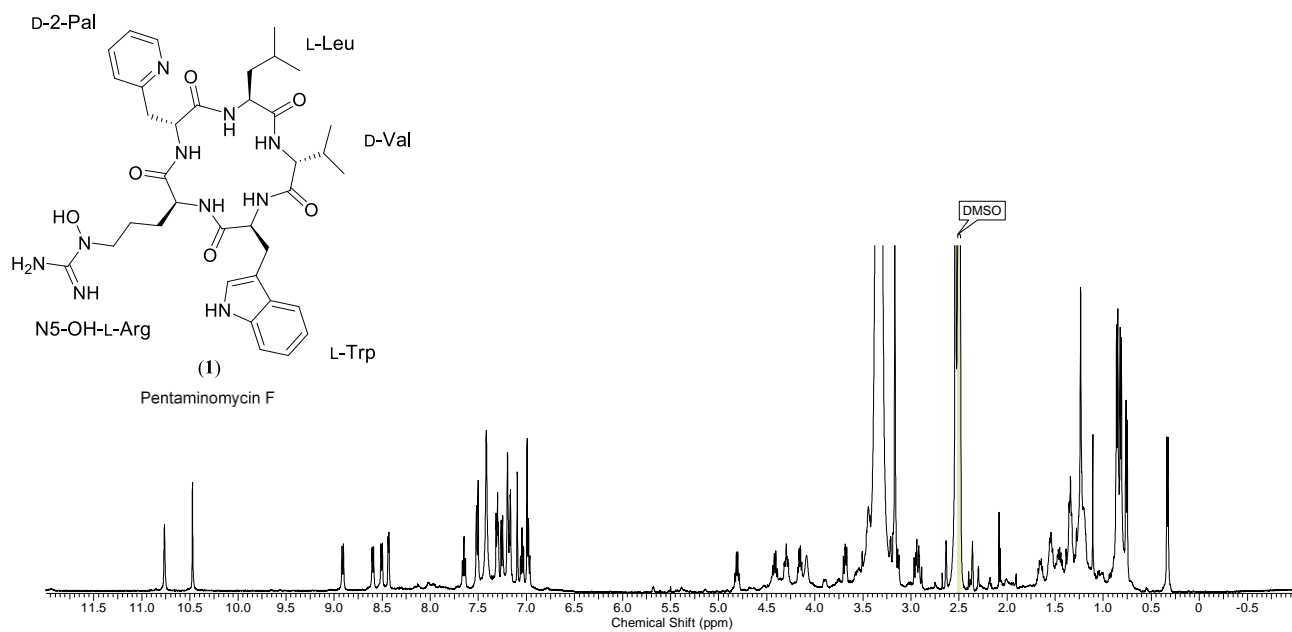
Figure S7. ^{1}H NMR (DMSO-d_6, 500 MHz) of **1.**
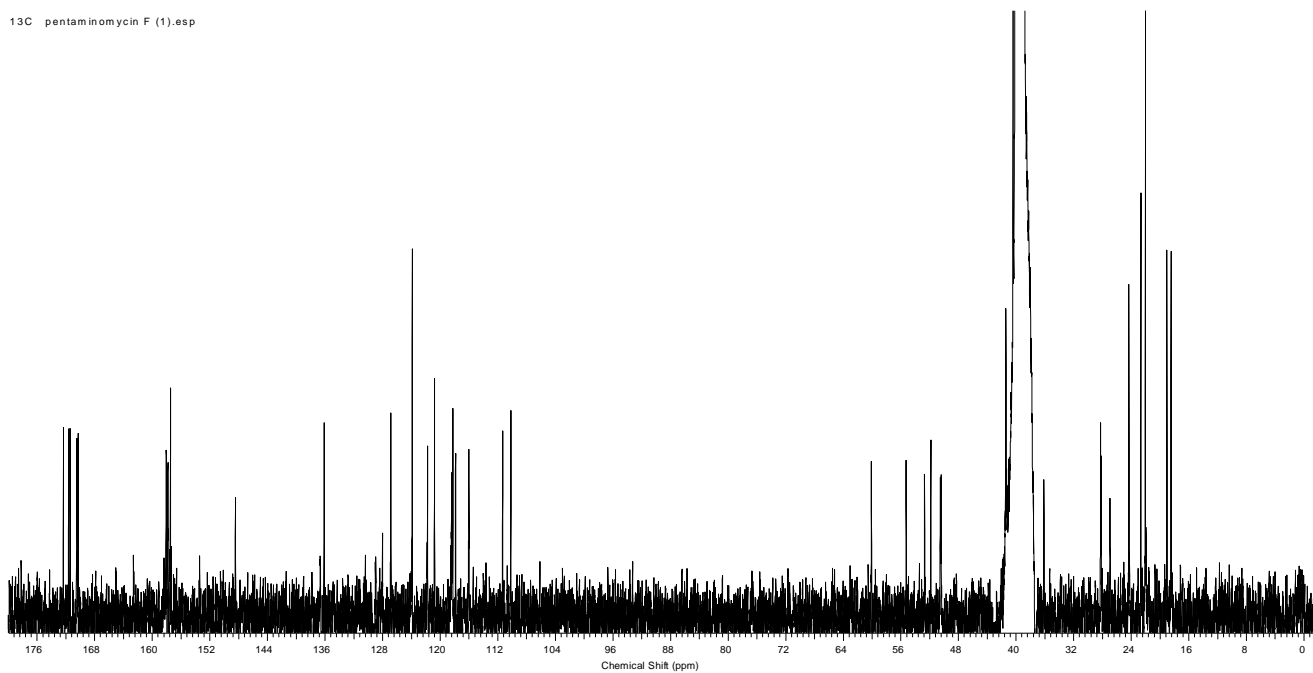
Figure S8. ^{13}C NMR (DMSO-d_6, 125 MHz) of **1.**

COSY pentaminomycin F (1).esp
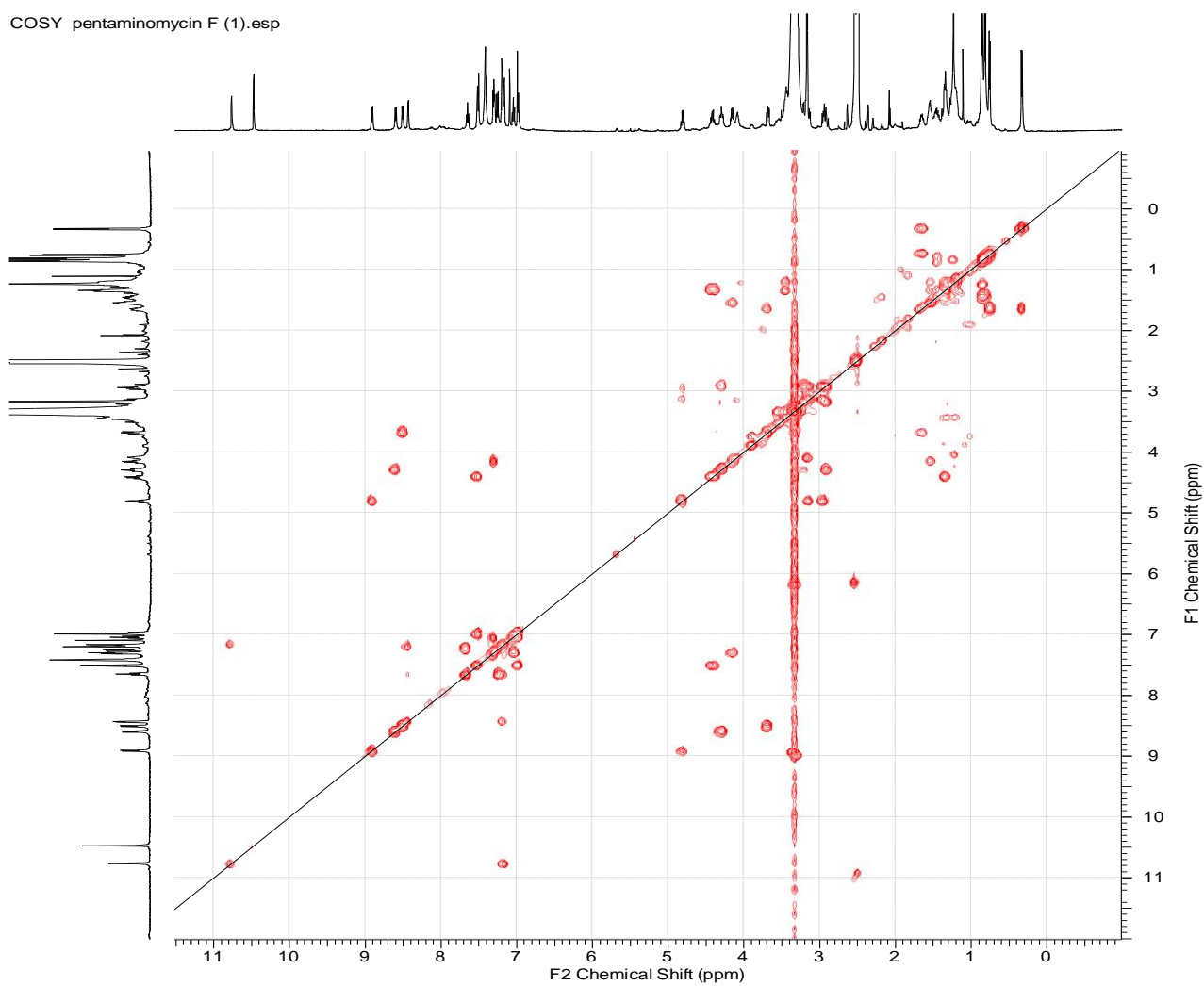
Figure S9. COSY spectrum of **1.**

HSQC pentaminomycin F (1).esp
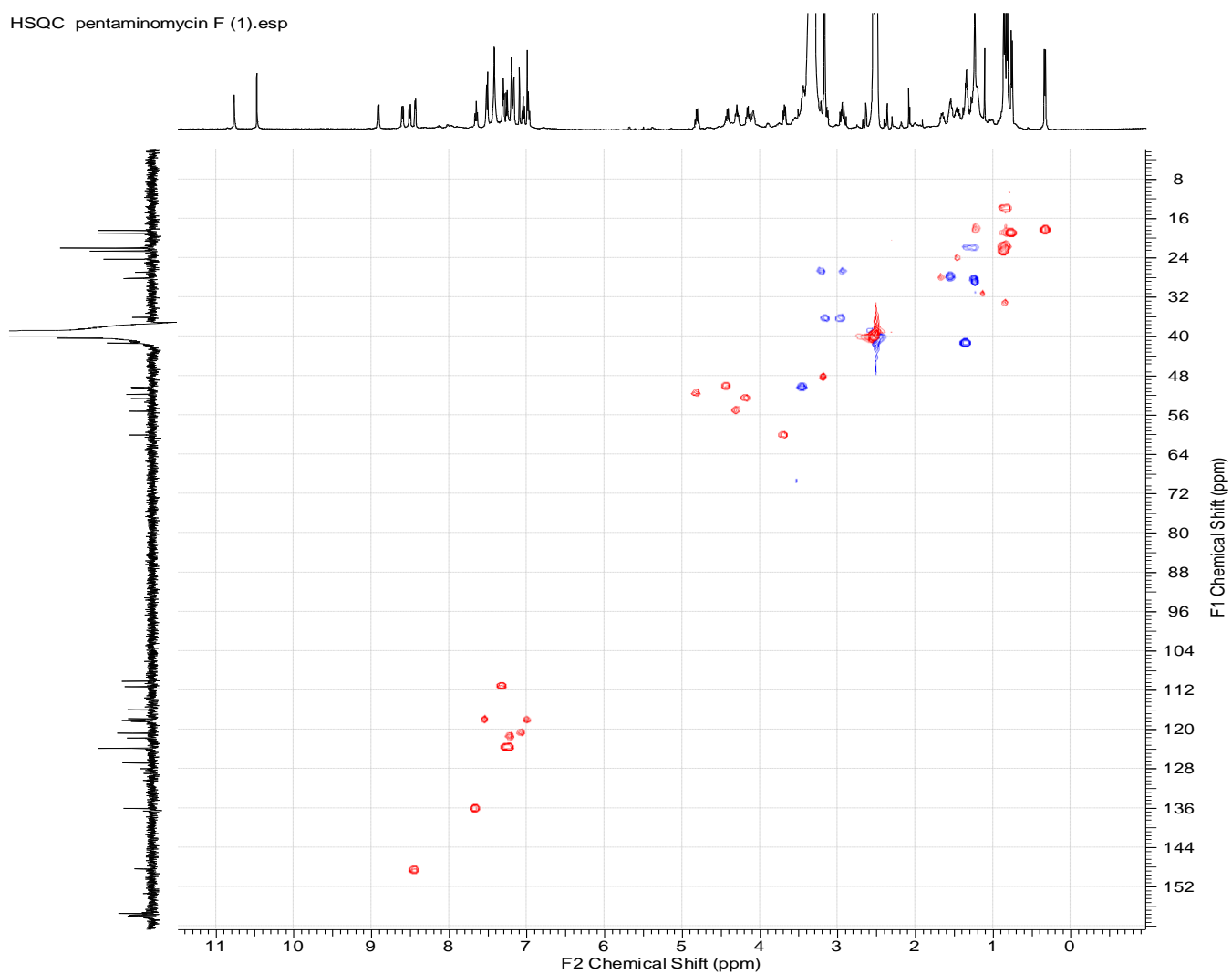
Figure S10. HSQC spectrum of **1.**

HMBC pentaminomycin F (1).esp
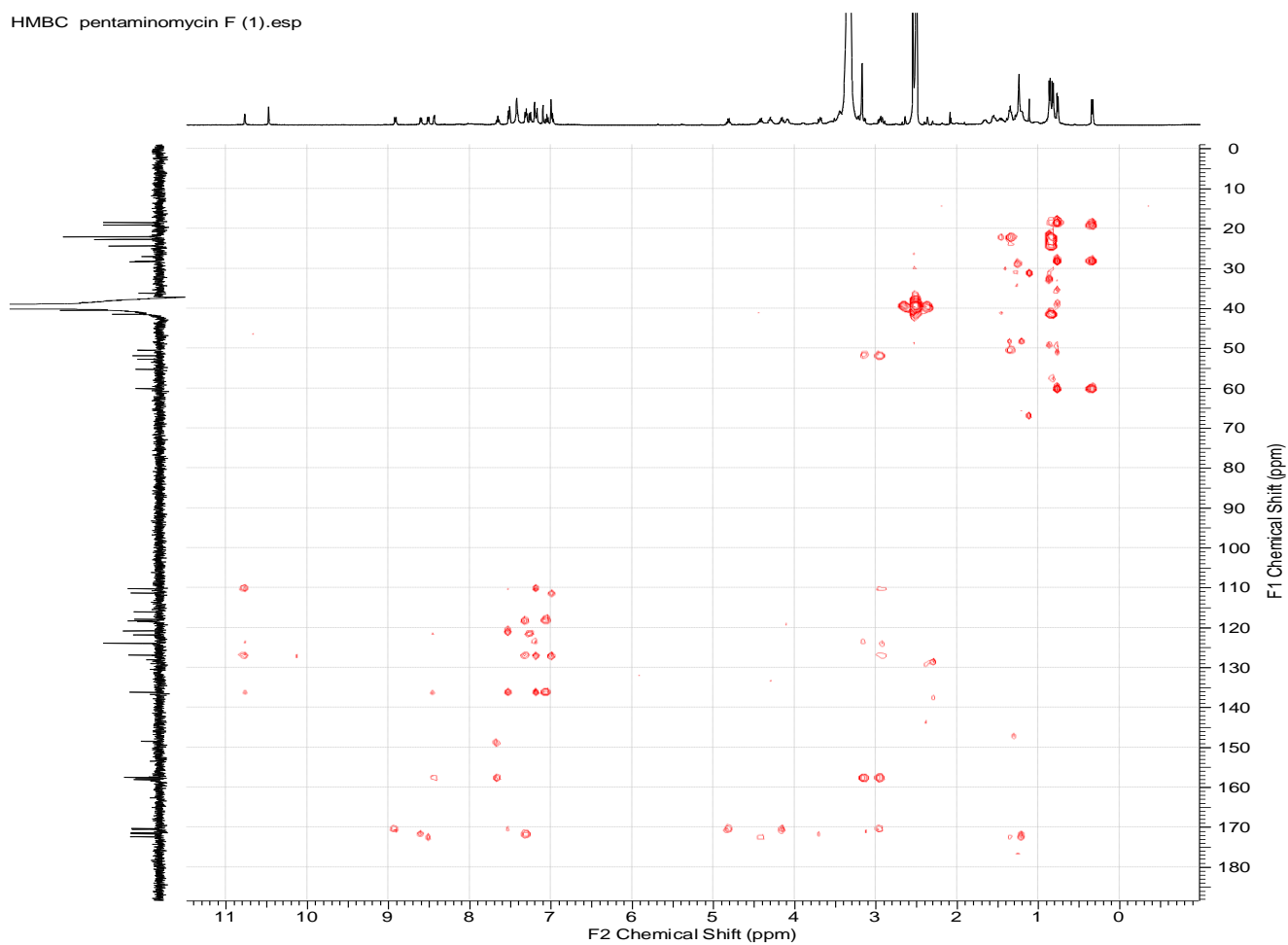
Figure S11. HMBC spectrum of **1.**

TOCSY pentaminomycin F (1).esp
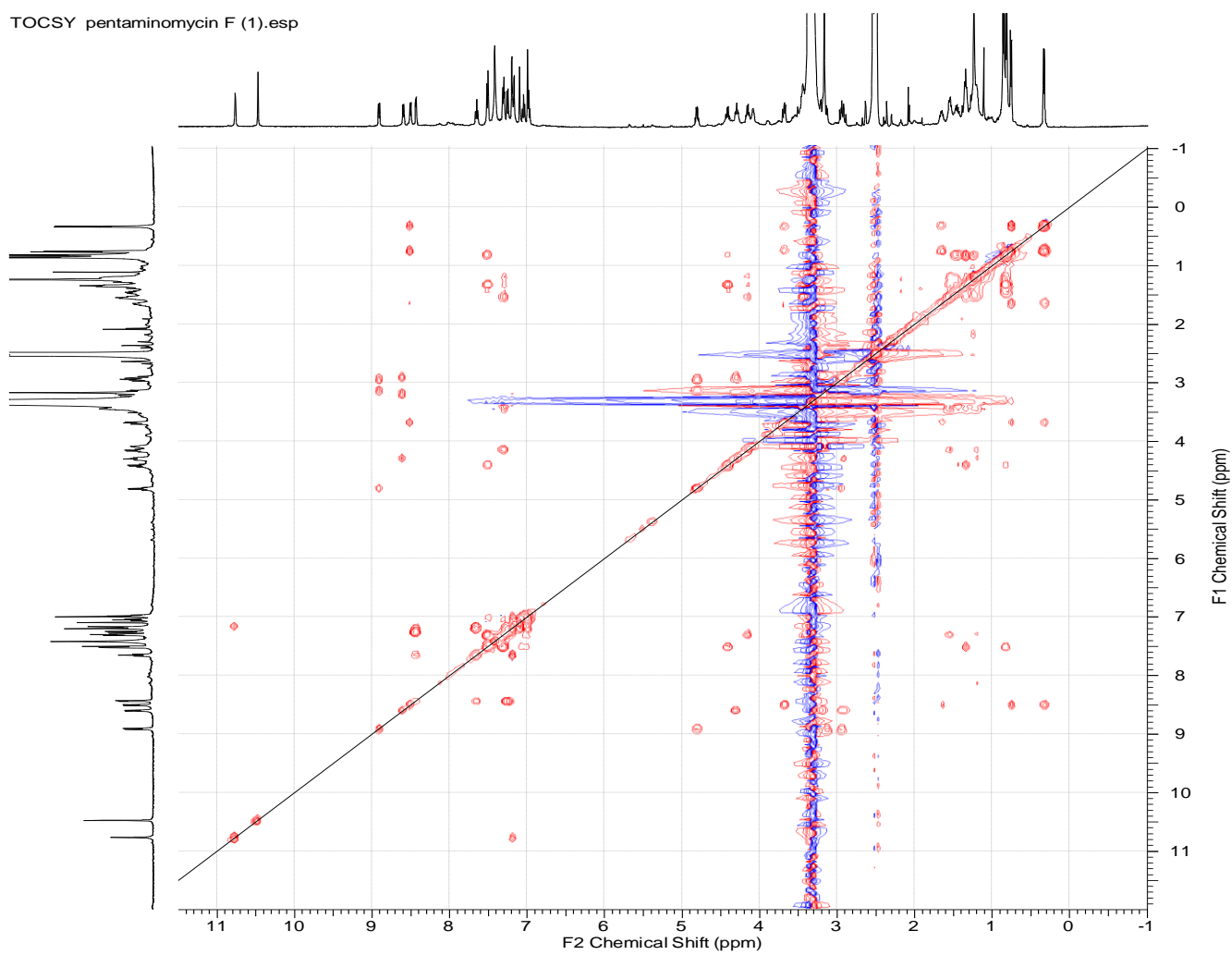
Figure S12. TOCSY spectrum of 1.
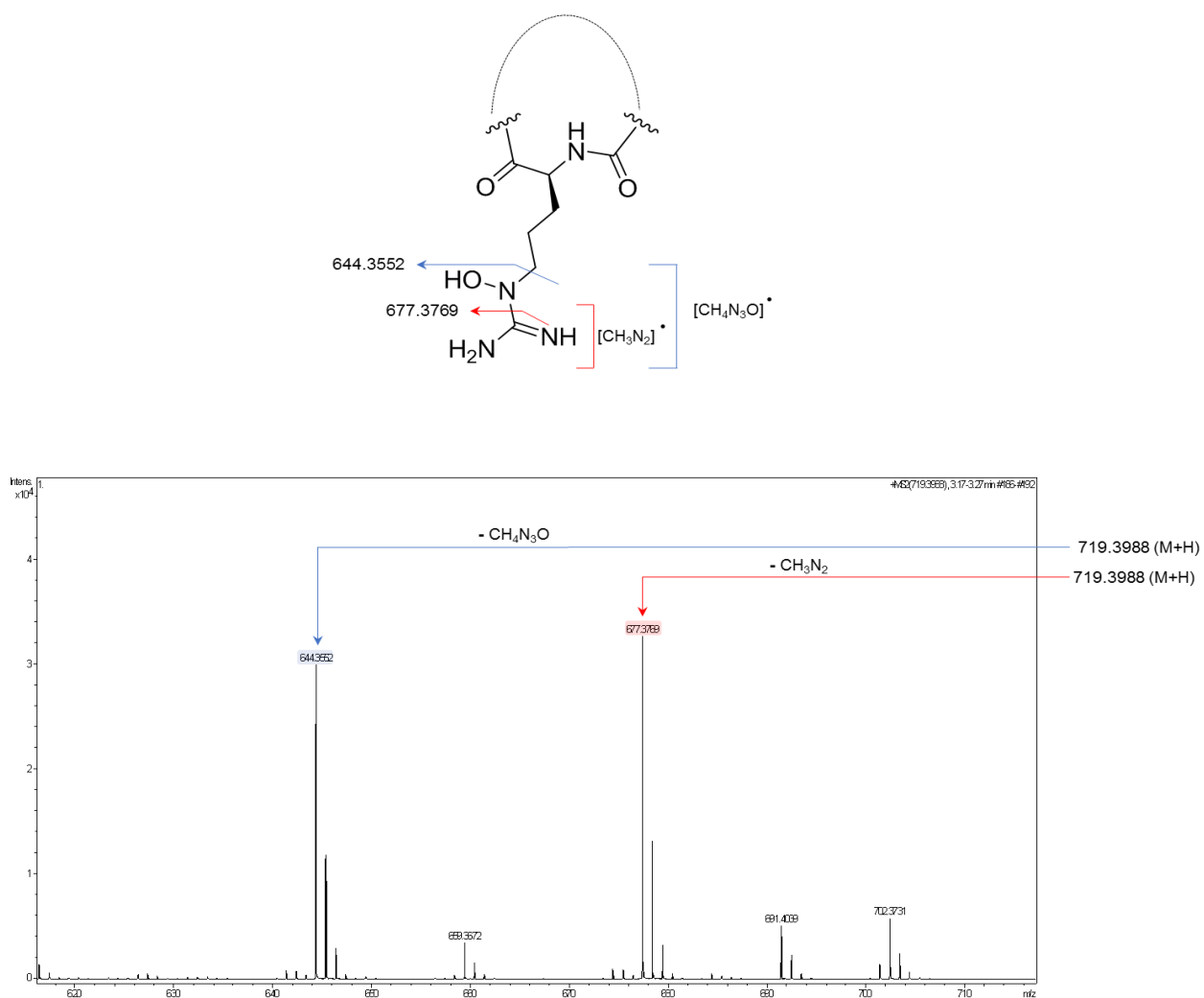
Figure S13. Confirmation of hydroxylation at N-5 of the arginine residue by HRMSMS for compound **1.**

(A)

(B)
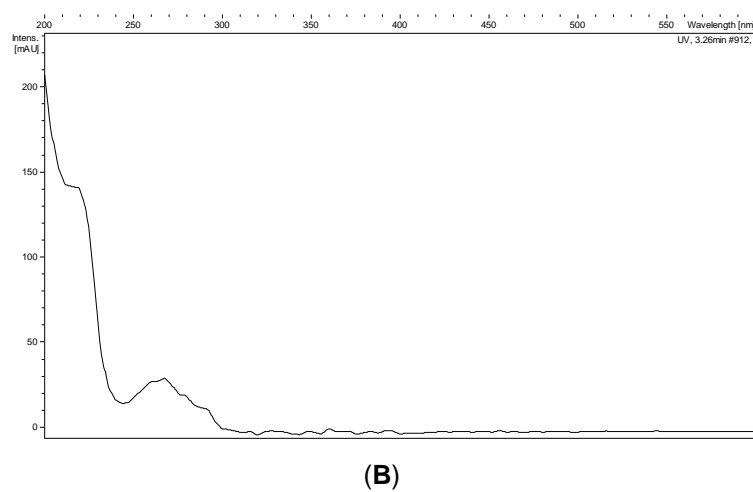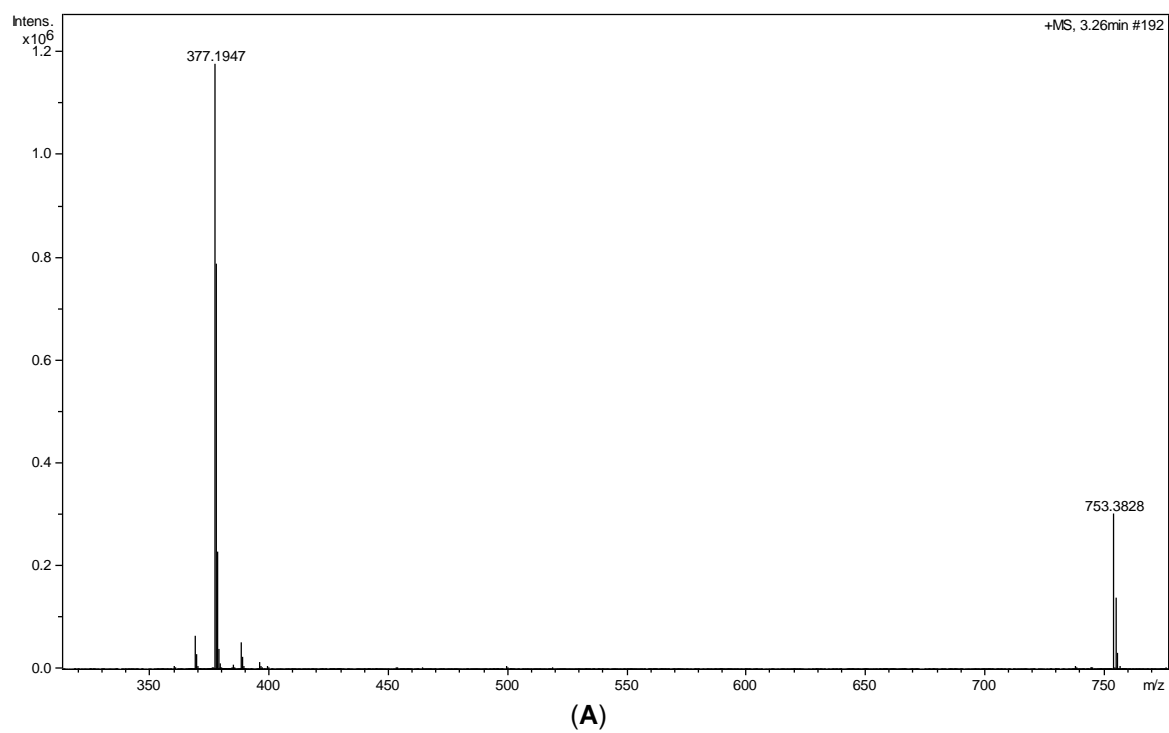
Figure S14. (+)-ESI-TOF (**A**) and UV (**B**) spectra of **2.**
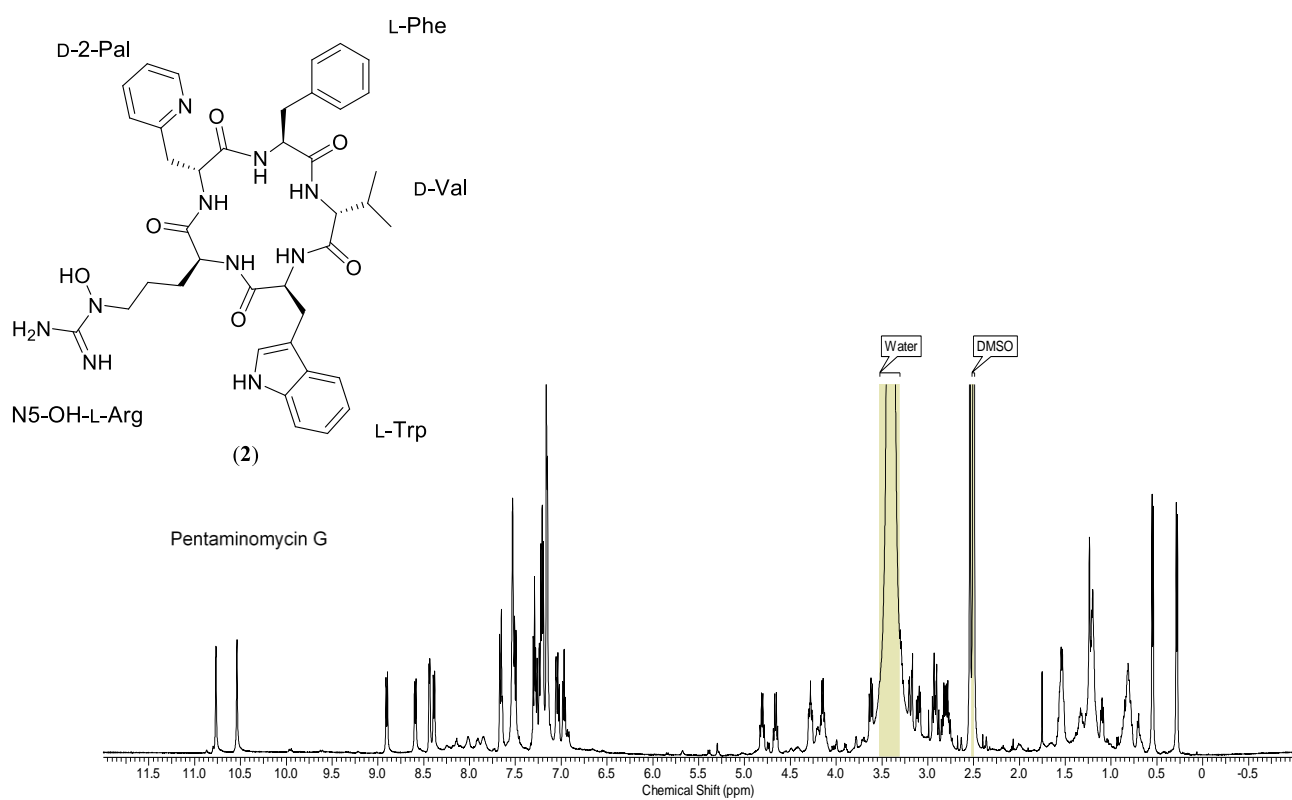
Figure S15. ^{1}H NMR (DMSO-d_6, 500 MHz) of **2.**
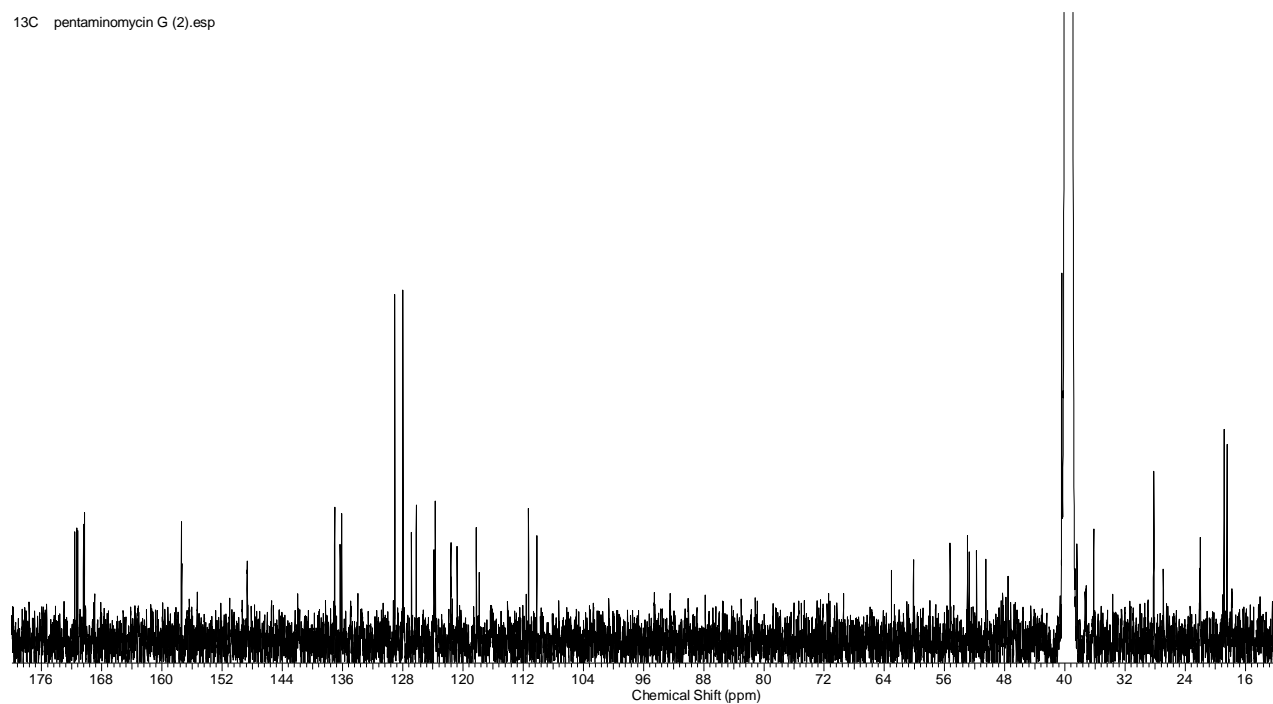
Figure S16. ^{13}C NMR (DMSO-d_6, 125 MHz) of **2.**

COSY   pentaminomycin G (2).esp
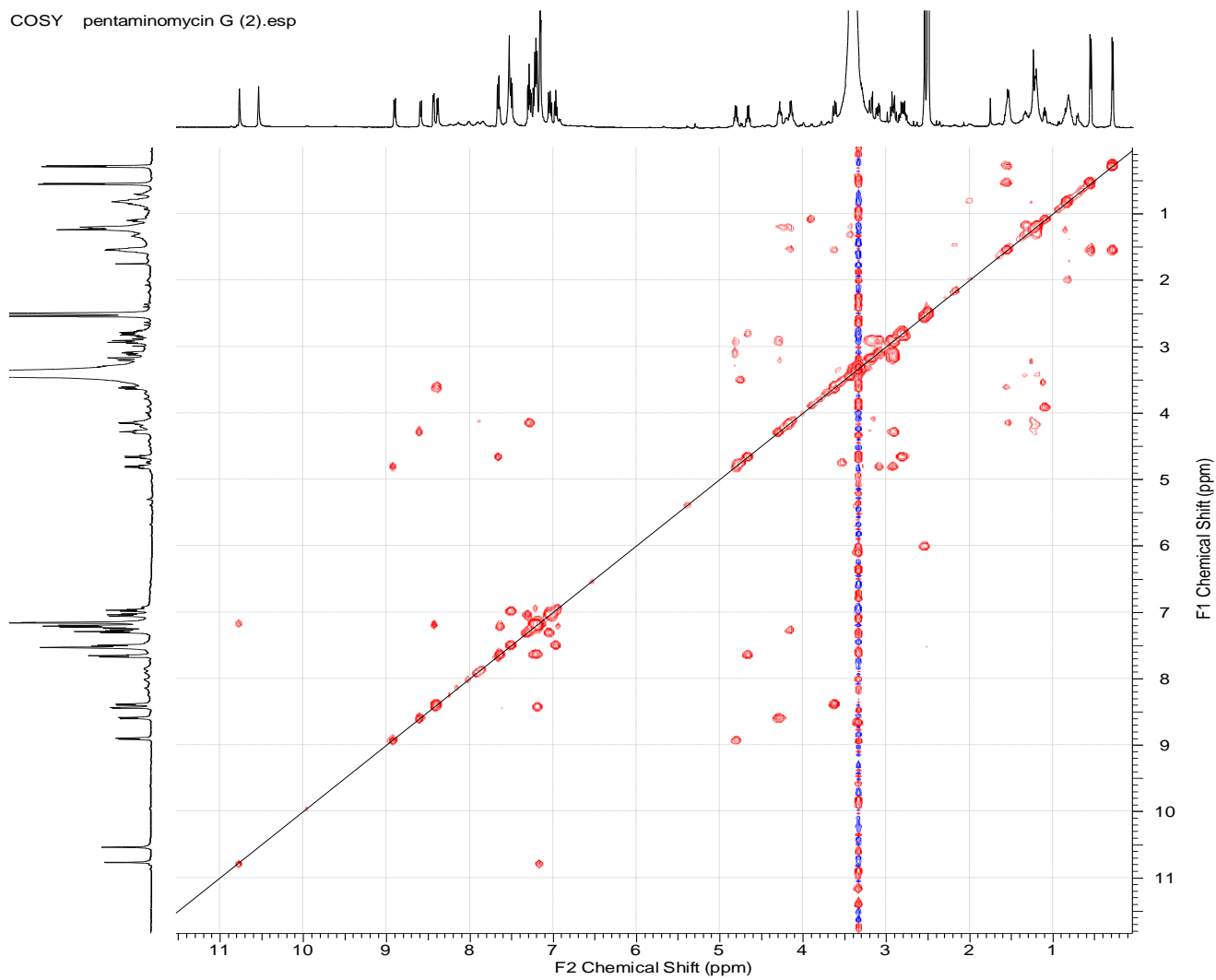
Figure S17. COSY spectrum of **2.**

HSQC   pentaminomycin G (2).esp
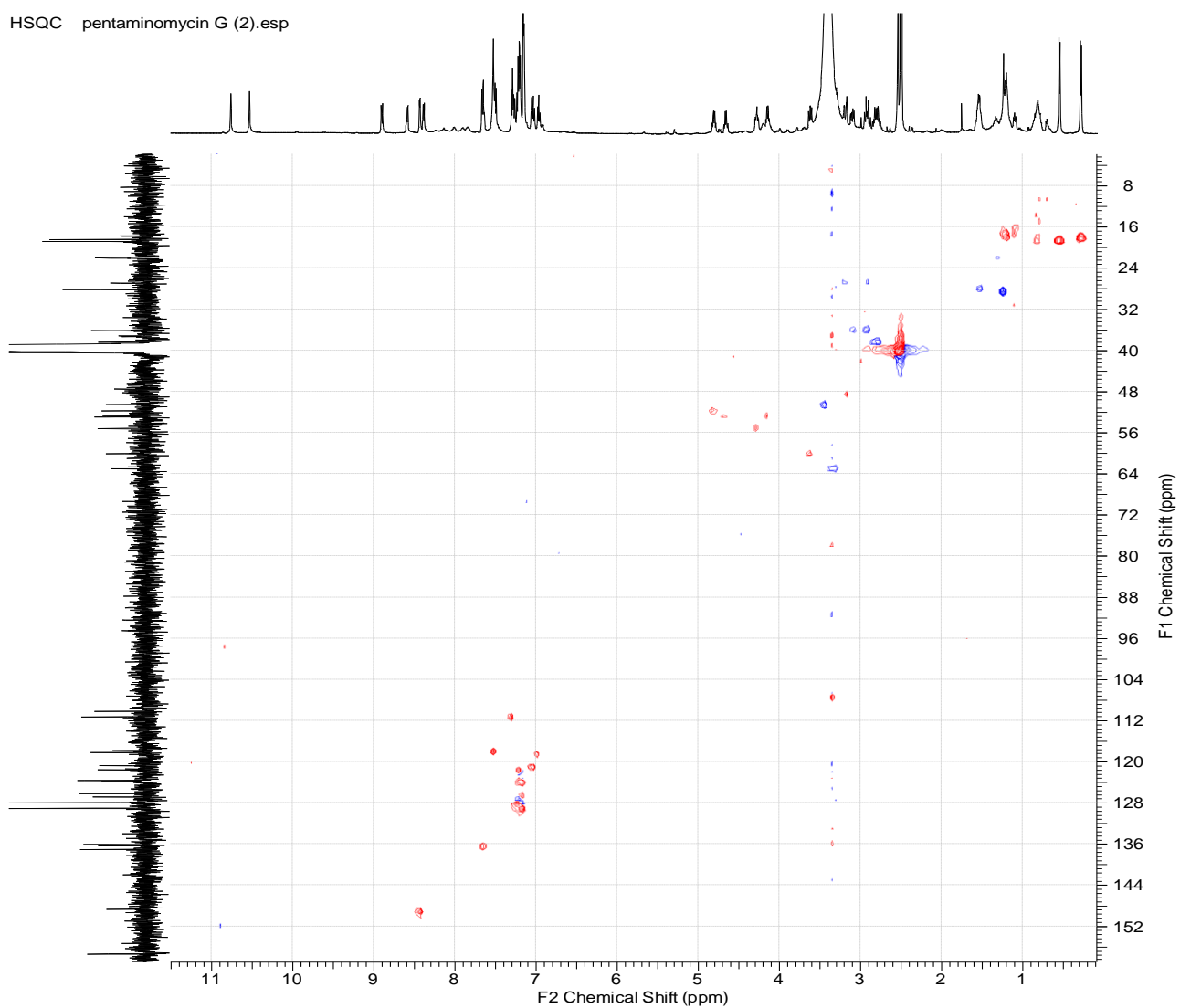
Figure S18. HSQC spectrum of **2.**

HMBC   pentaminomycin G (2).esp
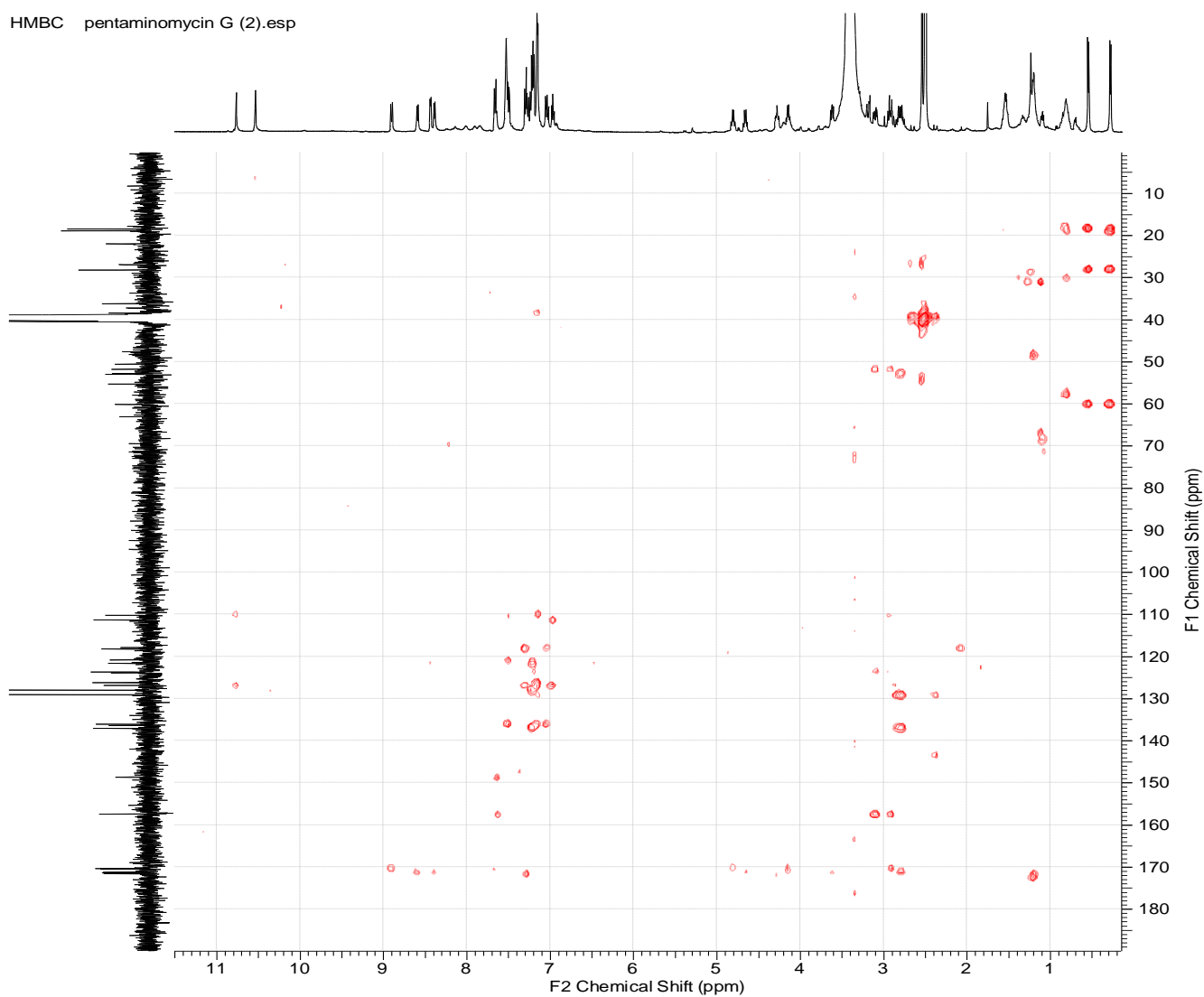
Figure S19. HMBC spectrum of **2.**

TOCSY   pentaminomycin G (2).esp
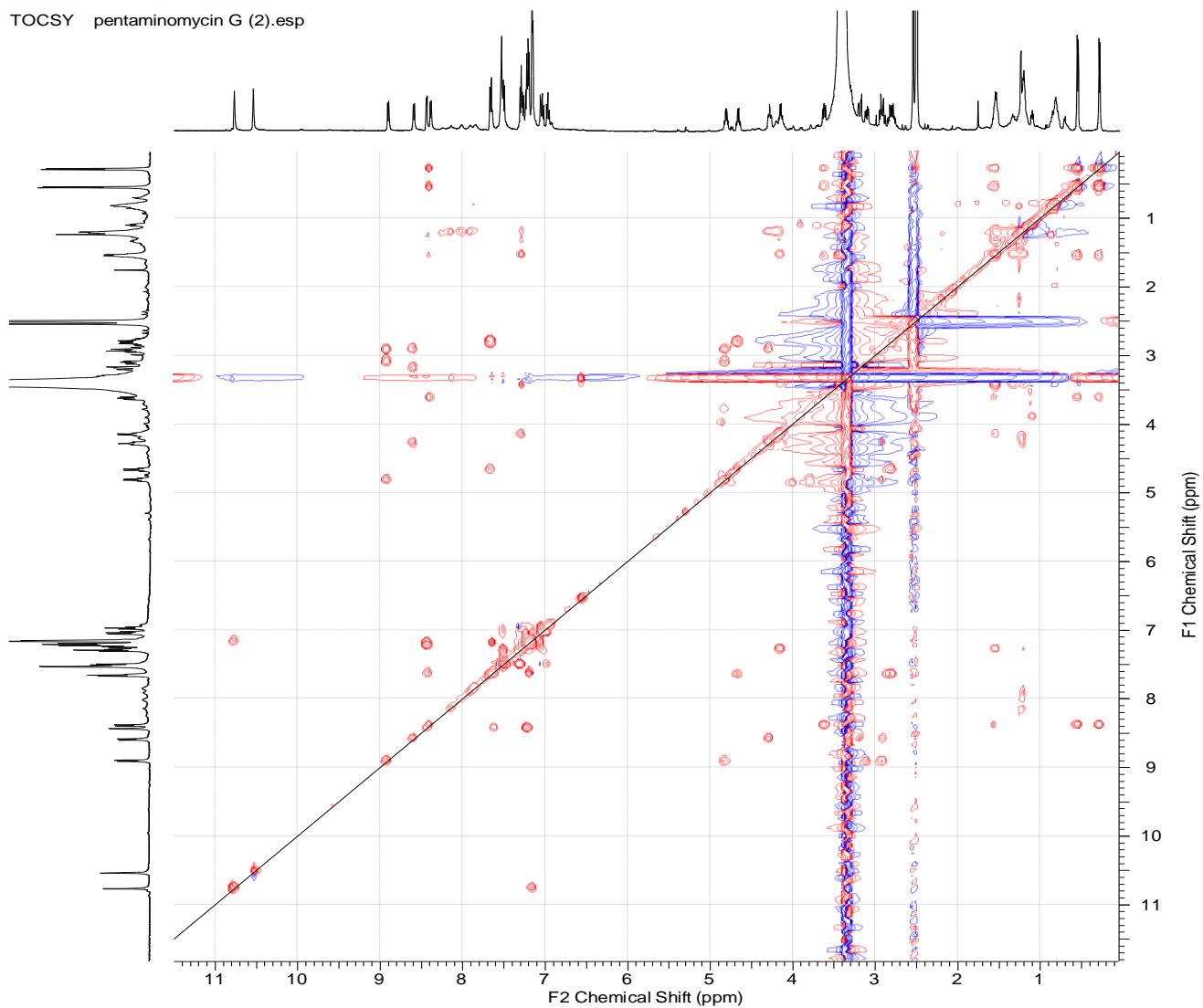
Figure S20. TOCSY spectrum of **2.**
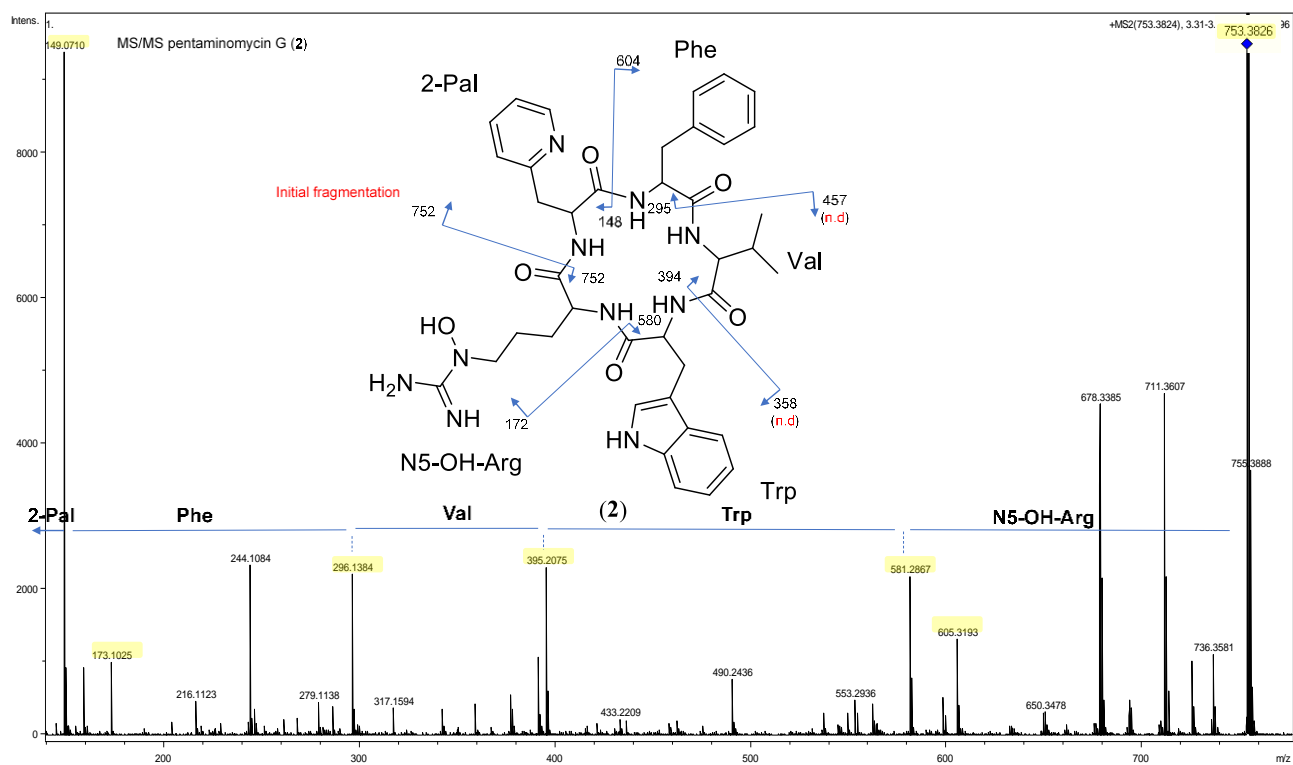
Figure S21. HRMS/MS spectrum and amino acid sequencing of **2**. (n.d.: not detected).
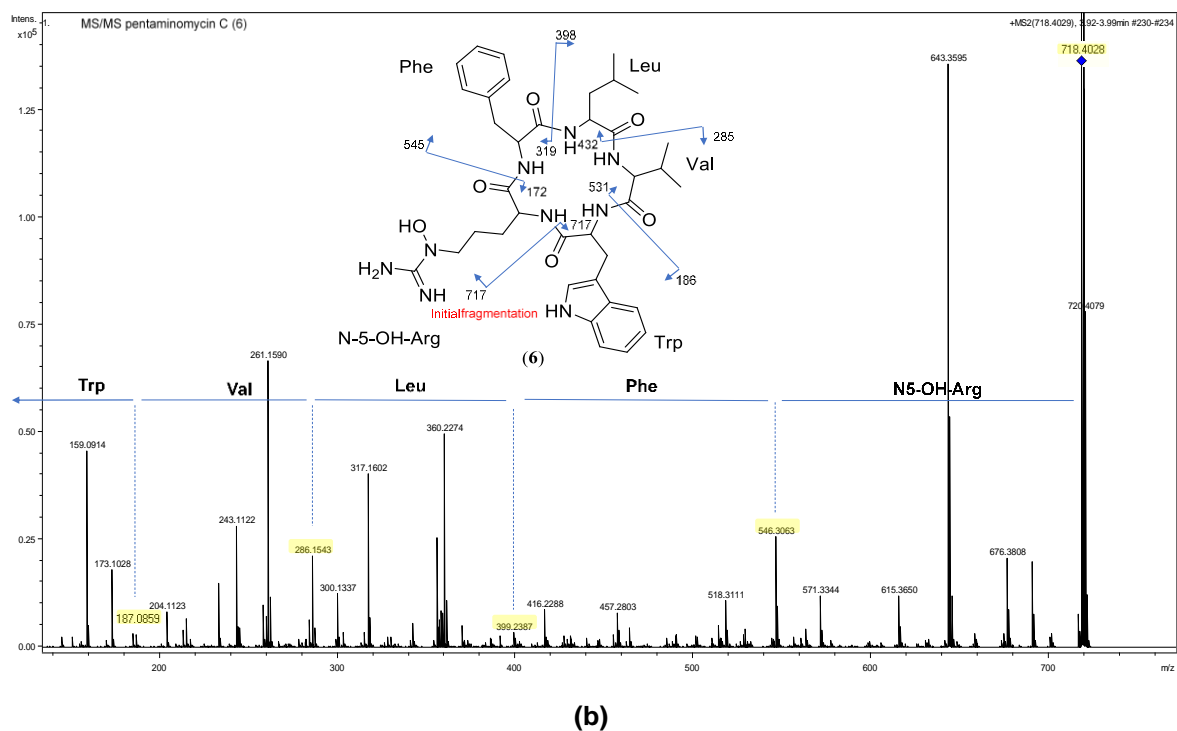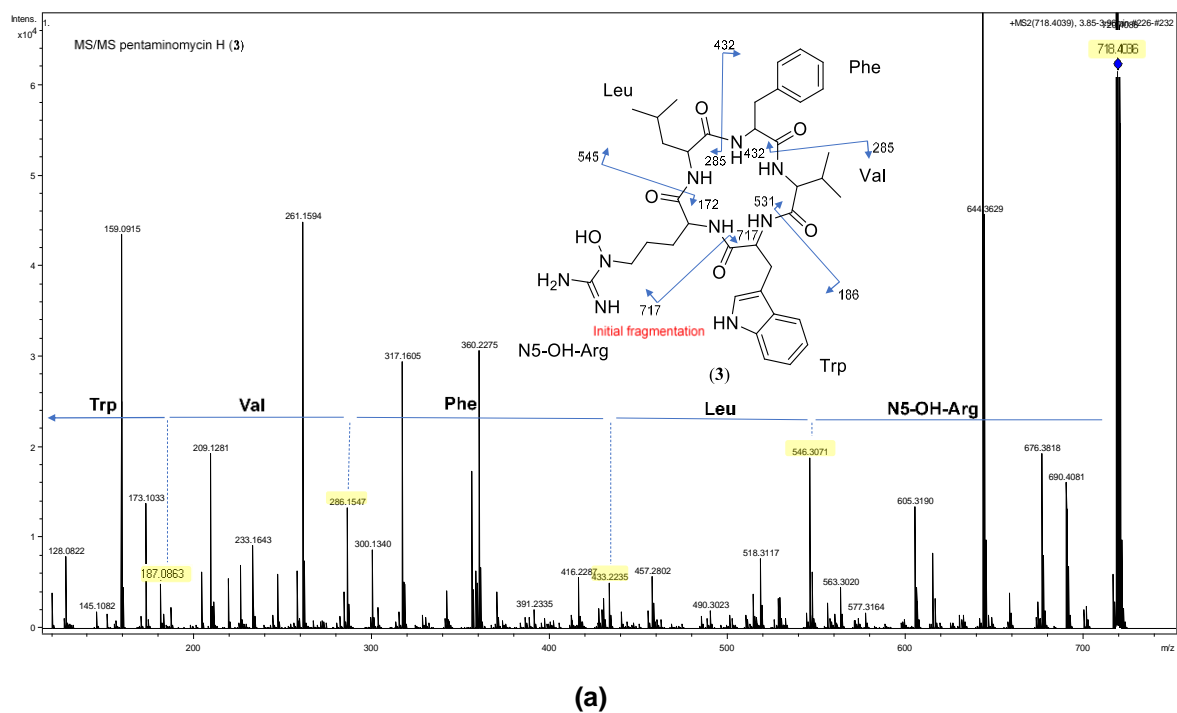
Figure S22. Comparison of different key fragment ions in the HRMS/MS spectrum of compound **3** (**a**) with those for compound **6** (**b**).
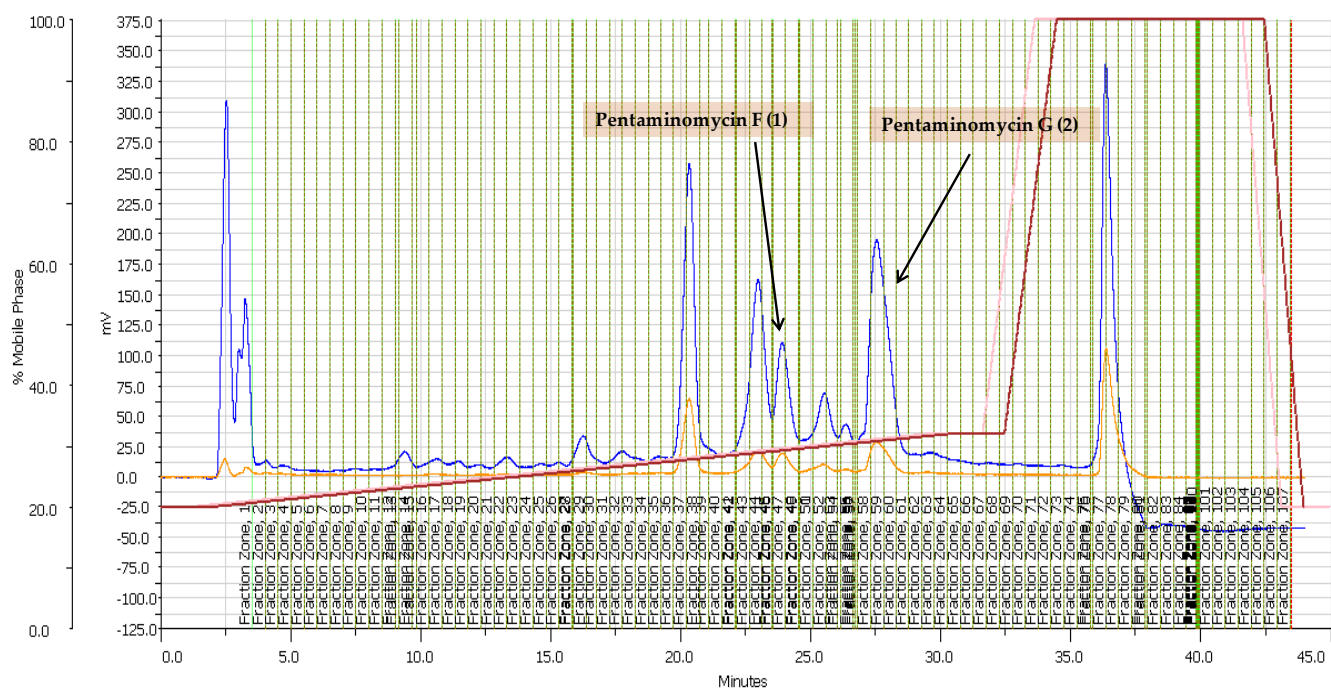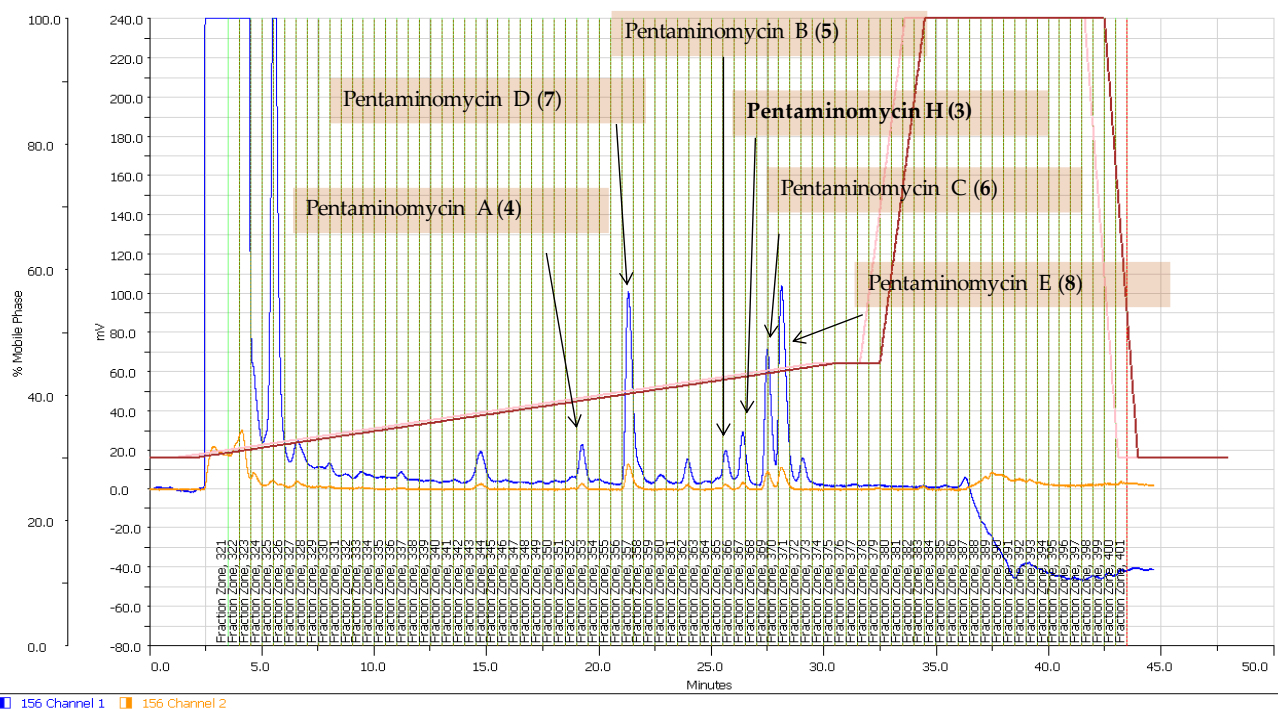
Figure S23. Isolation of compounds **3**-**8**. LC-UV semipreparative chromatogram was recorded at 210 (blue trace) and 280 nm (orange trace).

(A)
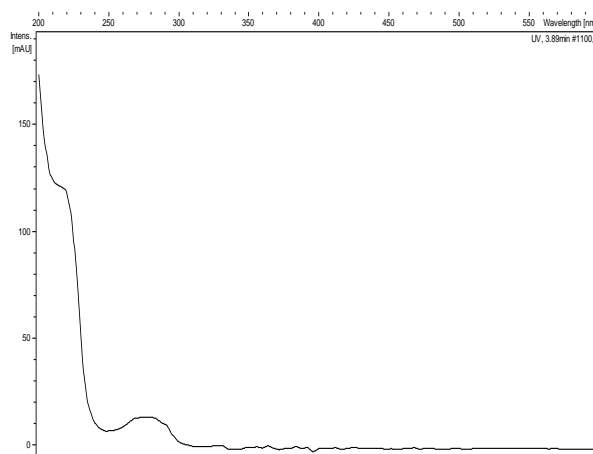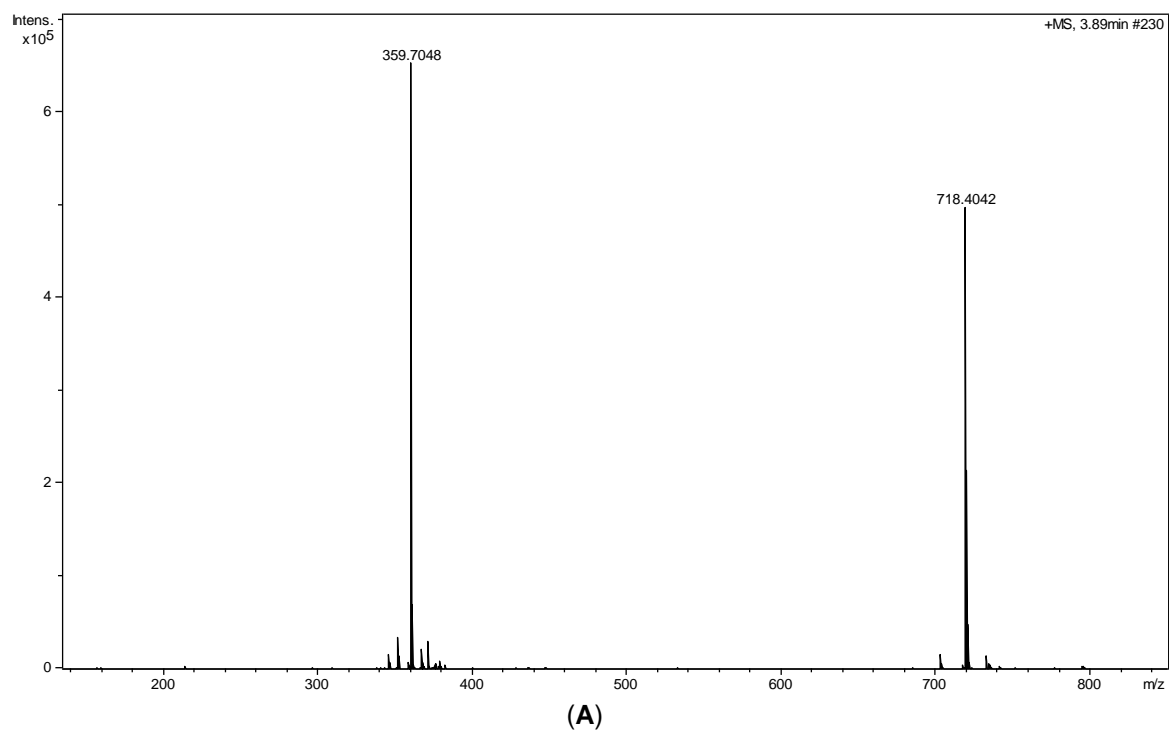
(B)
Figure S24. (+)-ESI-TOF (**A**) and UV (**B**) spectra of **3.**
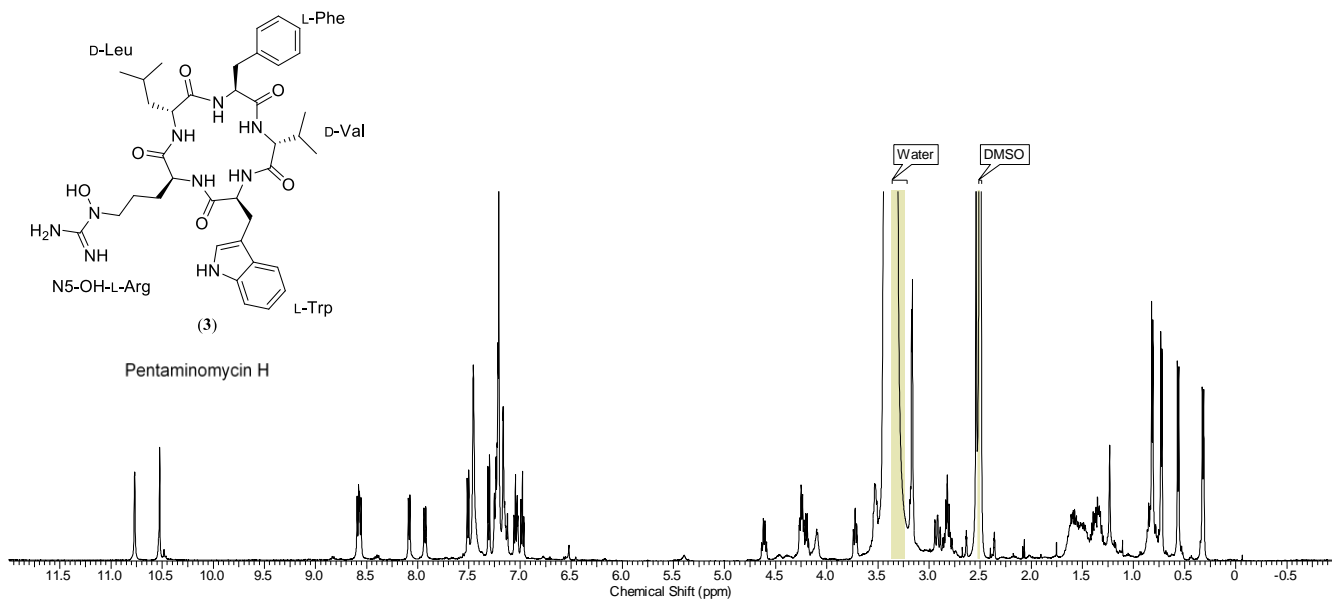
Figure S25. ^{1}H NMR (DMSO-d_6, 500 MHz) of **3.**
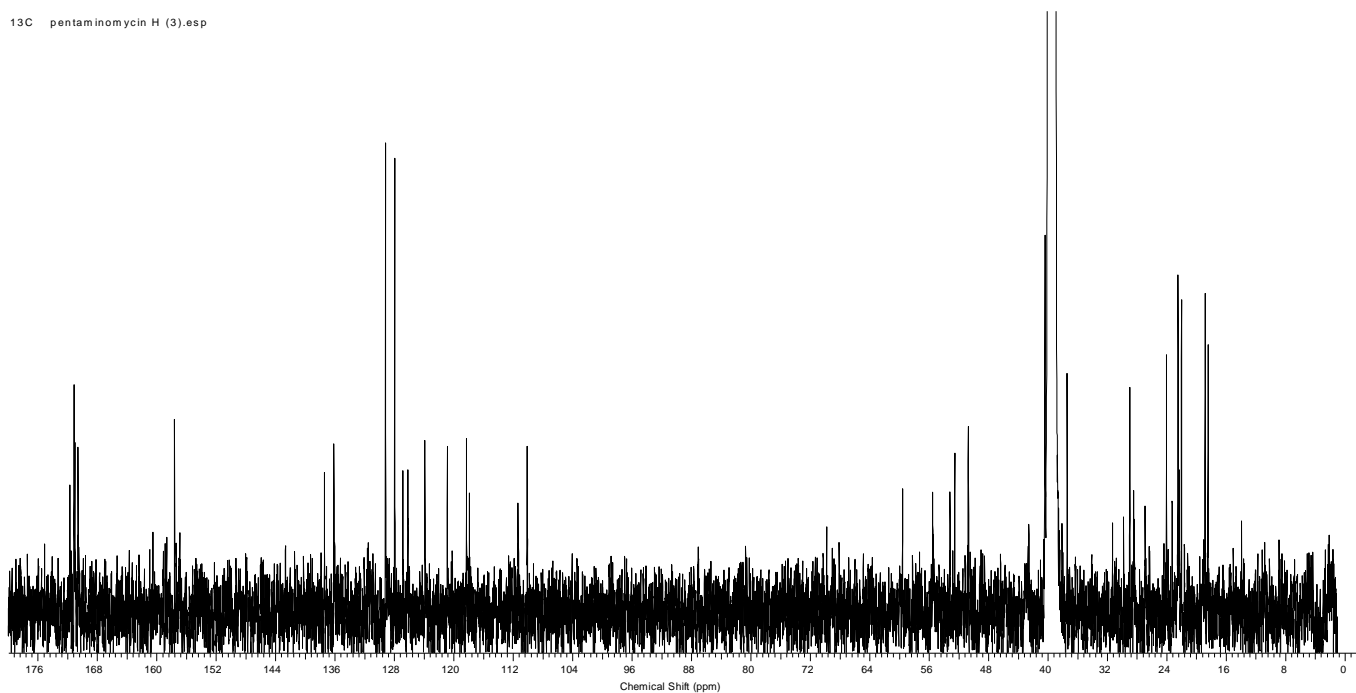
Figure S26. ^{13}C NMR (DMSO-d_6, 500 MHz) of **3.**

COSY   pentaminomycin H (3).esp
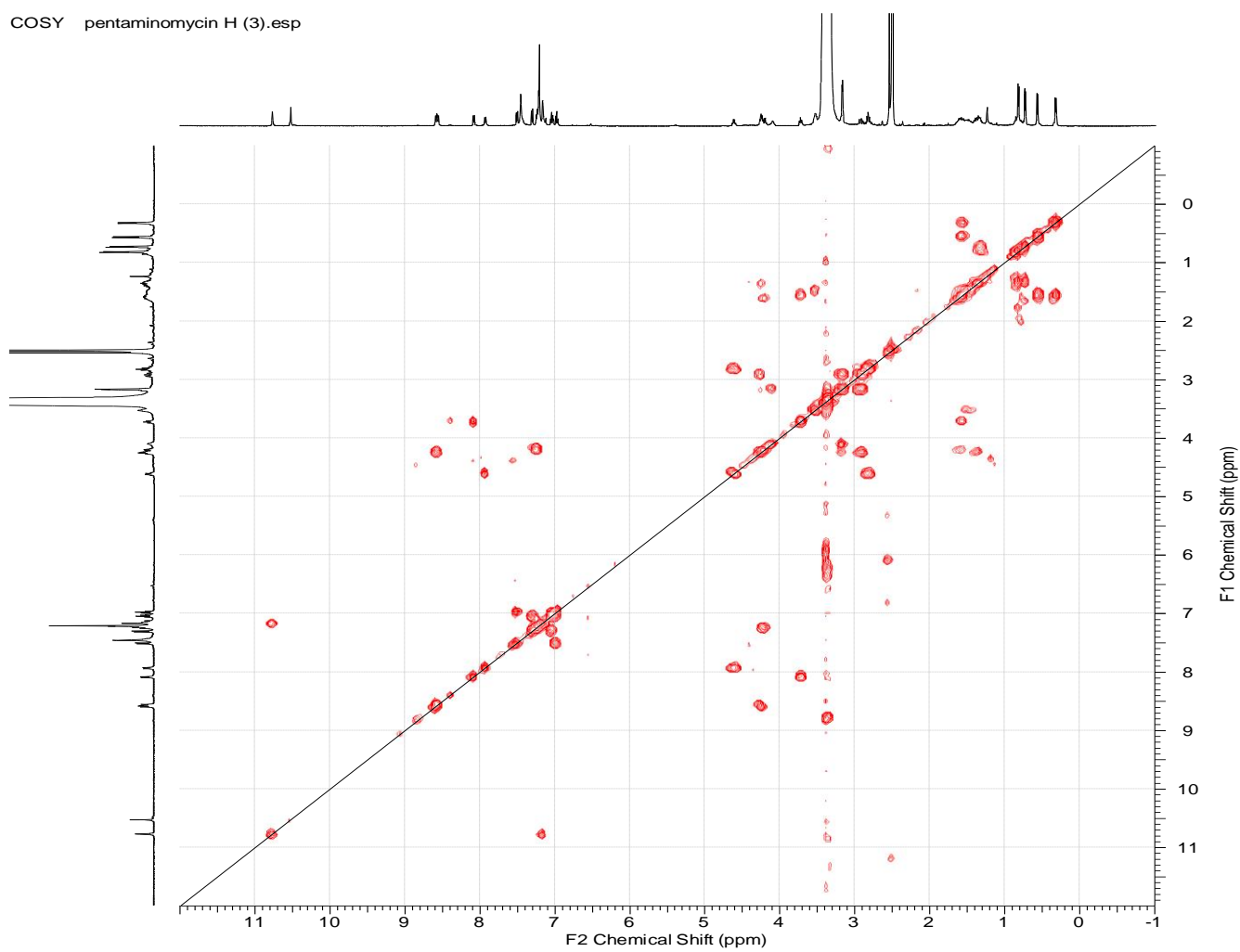
Figure S27. COSY spectrum of **3.**

HSQC   pentaminomycin H (3).esp
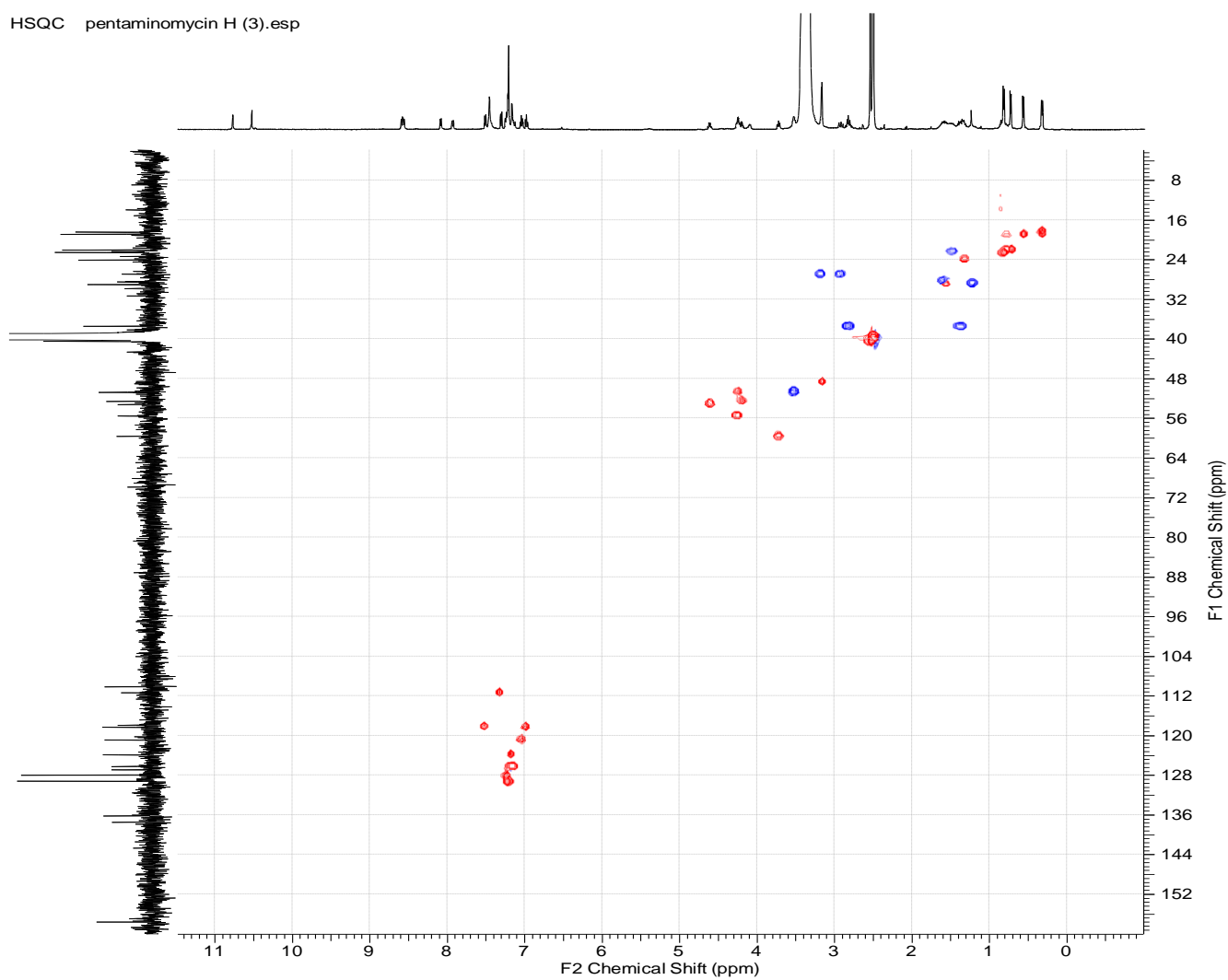
Figure S28. HSQC spectrum of **3.**

HMBC   pentaminomycin H (3).esp
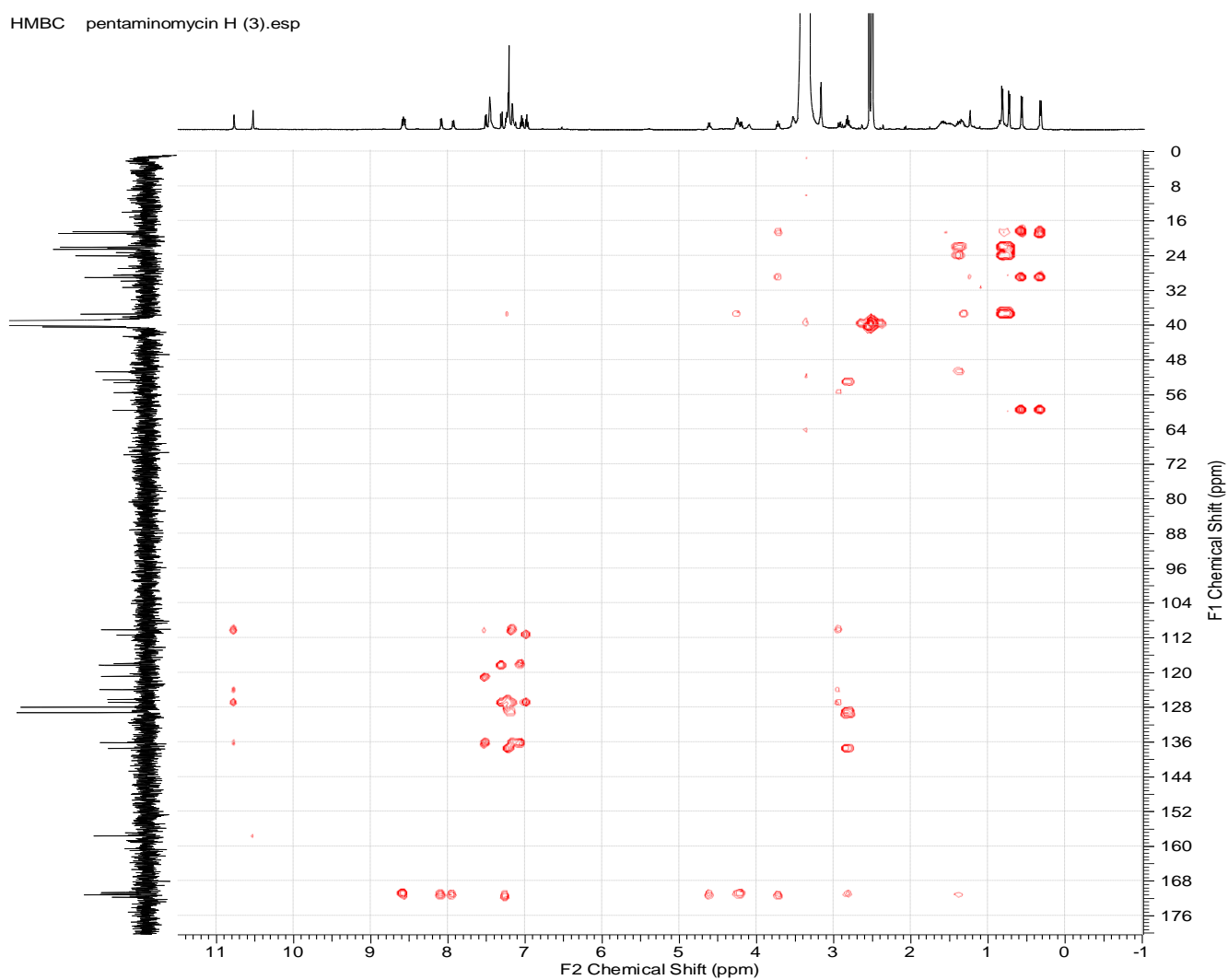
Figure S29. HMBC spectrum of **3.**

NOESY   pentaminomycin H (3).esp
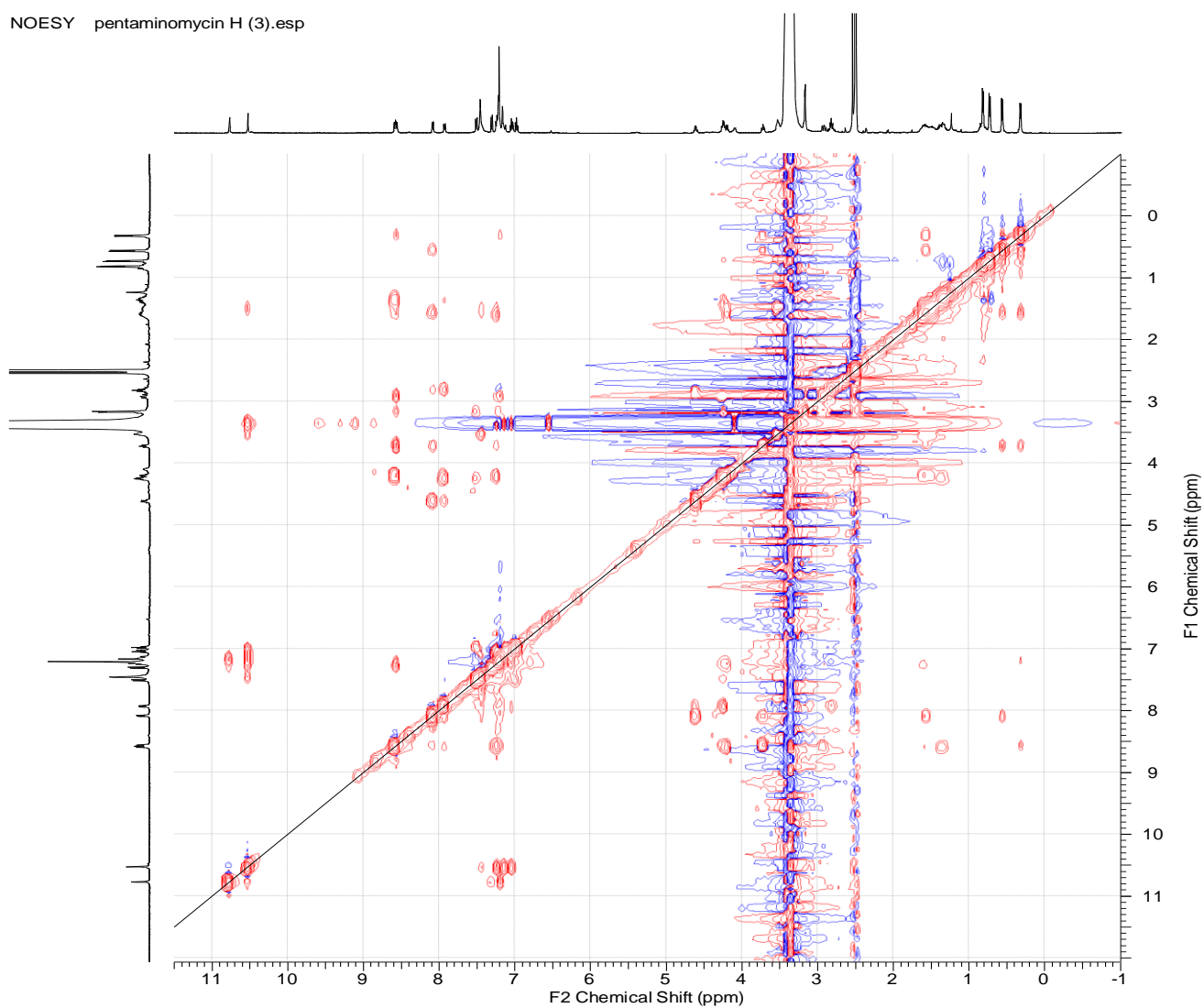
Figure S30. NOESY spectrum of **3.**
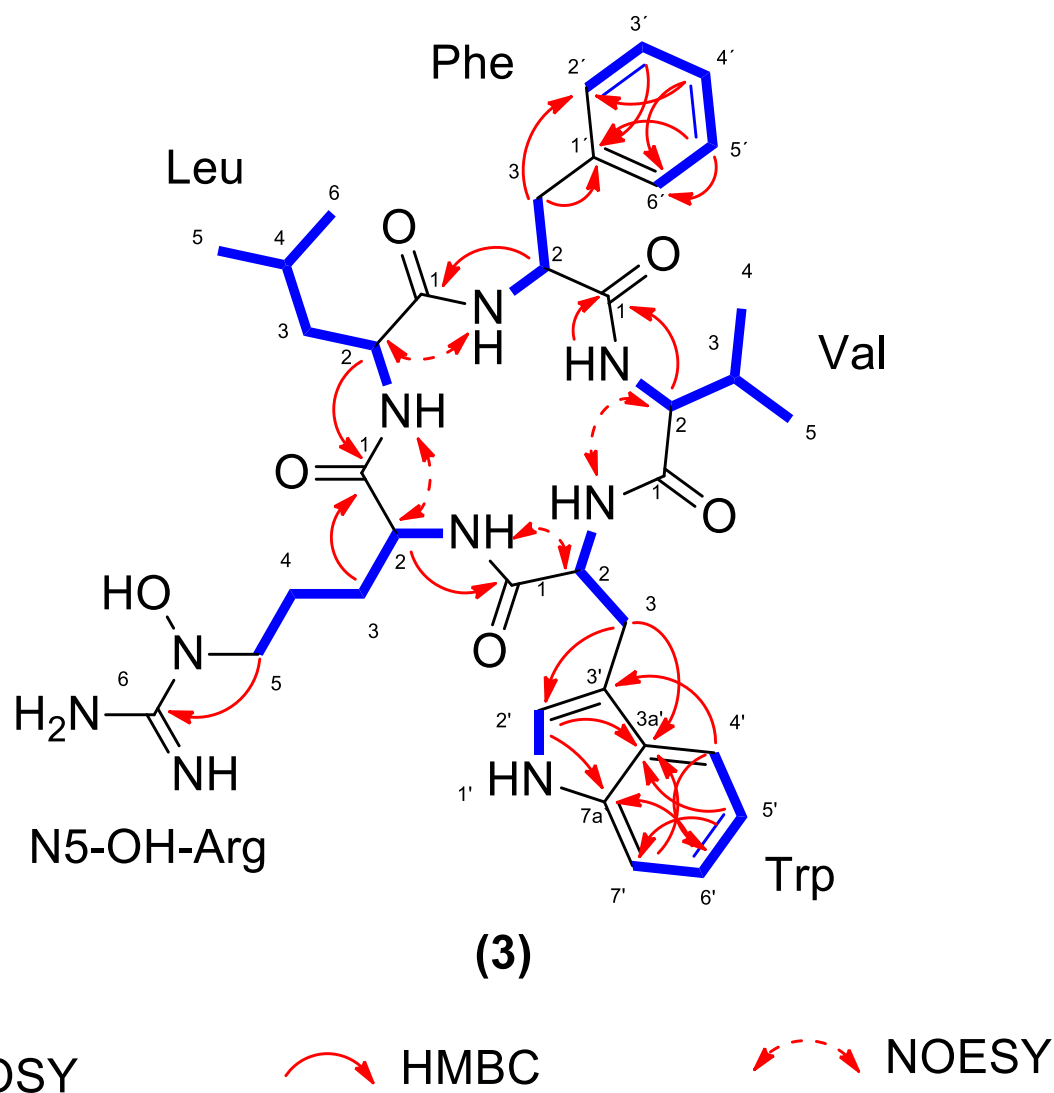
Figure S31. Key COSY, HMBC, and NOESY correlations observed for **3**.
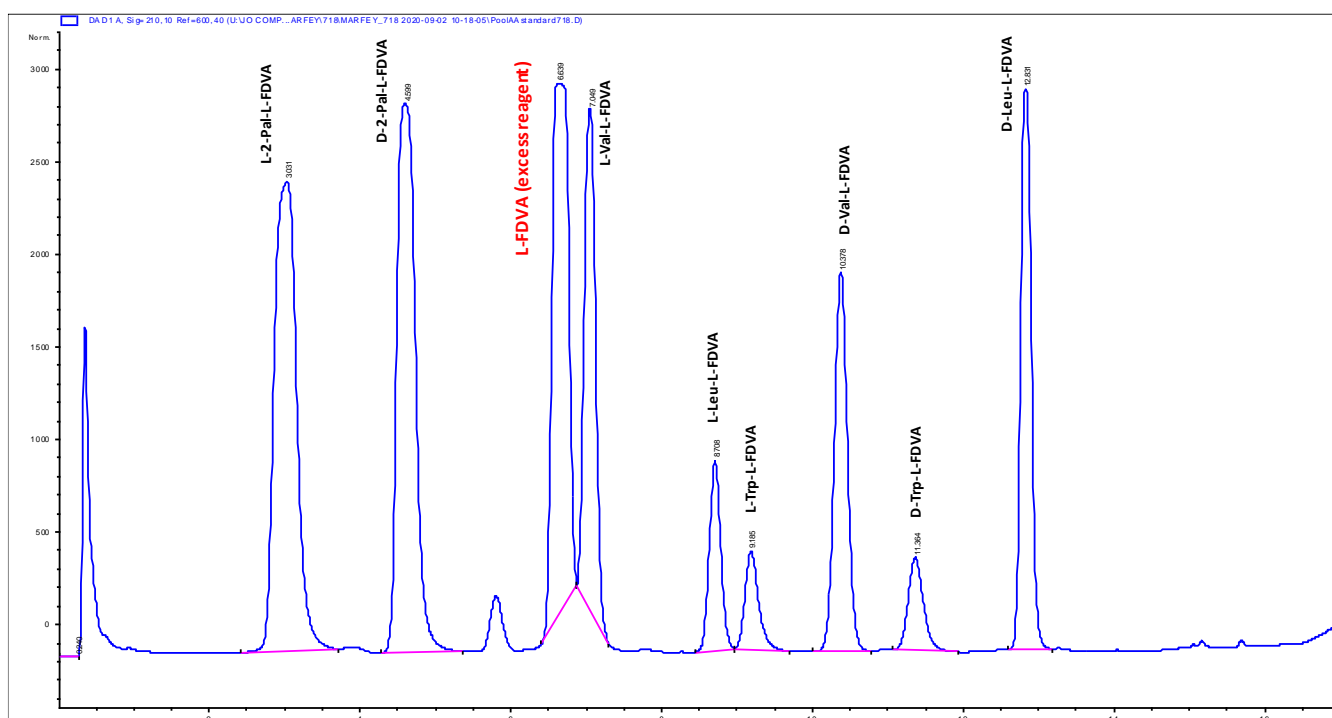
Figure S32. LC-UV chromatogram of a mixture of L-FDVA derivatization reactions of L-Val / D-Val / L-Trp / D-Trp / L-2-Pal / D-2-Pal / L-Leu / D-Leu standards amino acids present in **1.**
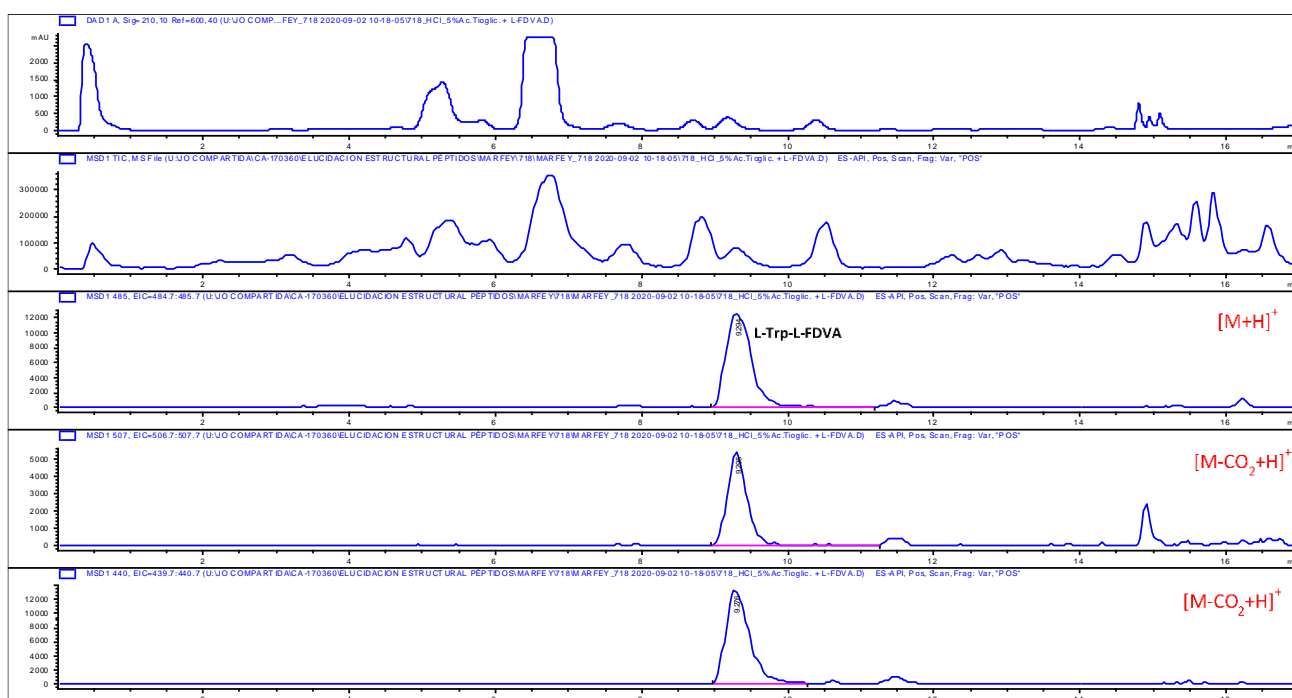
Figure S33. LC-UV and extracted-ion chromatograms (EIC) of Trp-L-FDVA derivatized hydrolysis of **1** with HCl (5% Thioglycolic acid). L-Trp was found in **1.**
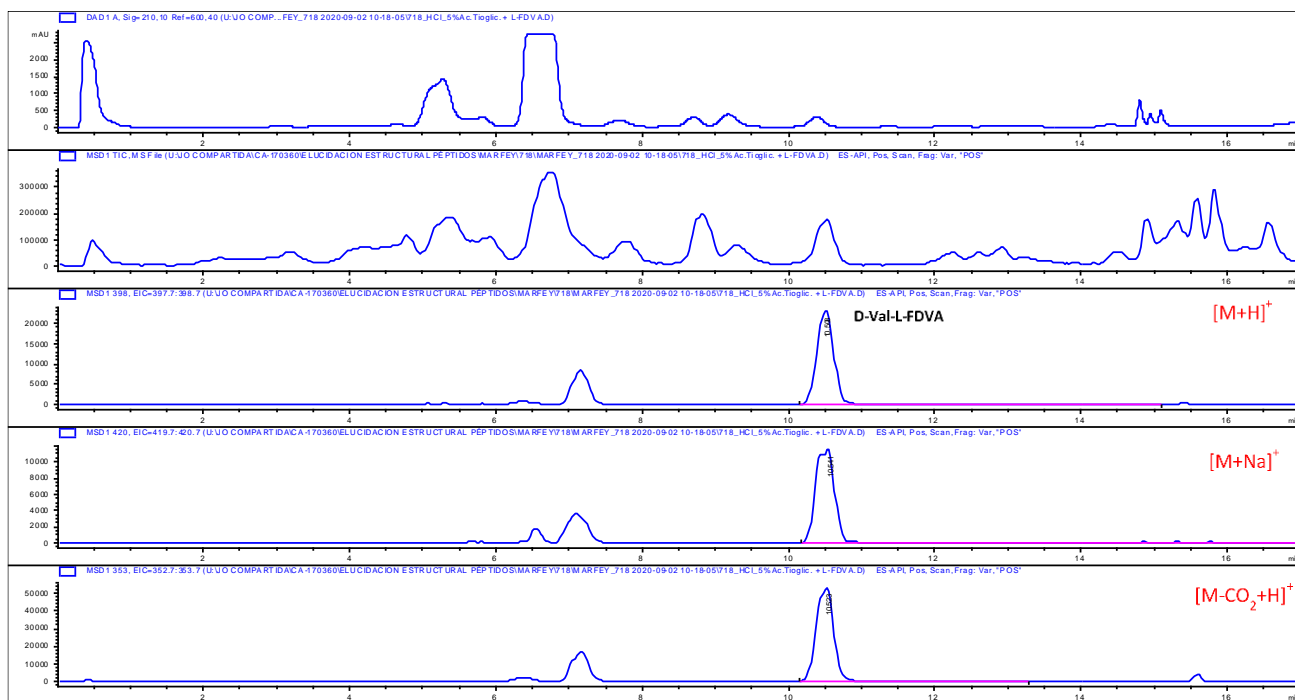
Figure S34. LC-UV and extracted-ion chromatograms (EIC) of Val-L-FDVA derivatized hydrolysis of **1** with HCl (5% Thioglycolic acid). D-Val was found in **1.**
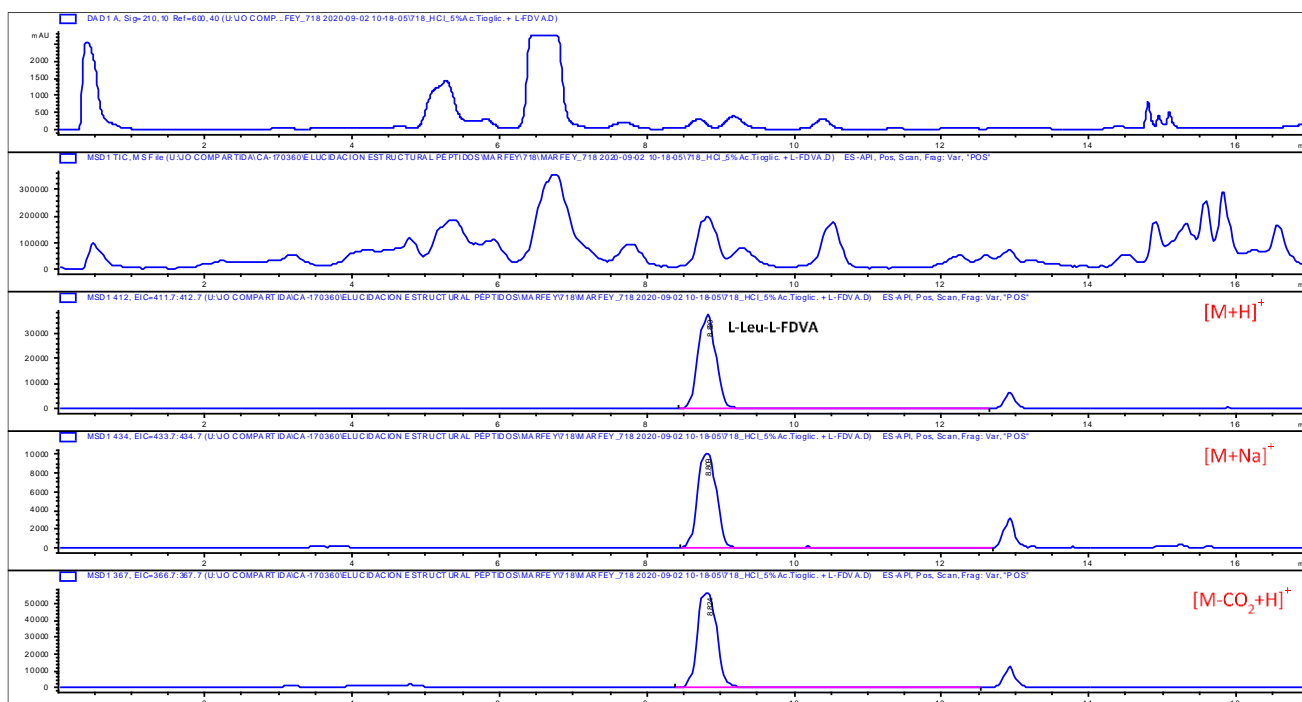
Figure S35. LC-UV and extracted-ion chromatograms (EIC) of Leu-L-FDVA derivatized hydrolysis of **1** with HCl (5% Thioglycolic acid). L-Leu was found in **1.**
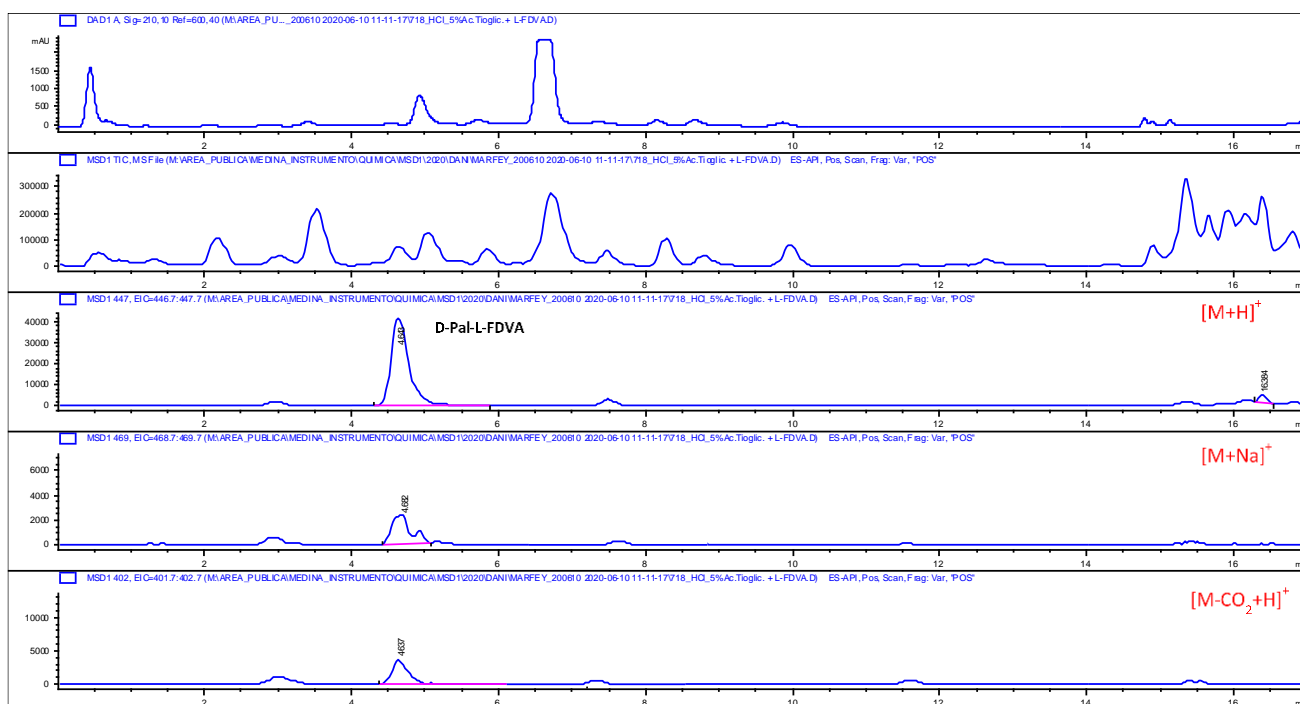
Figure S36. LC-UV and extracted-ion chromatograms (EIC) of 2-Pal-L-FDVA derivatized hydrolysis of **1** with HCl (5% Thioglycolic acid). D-2-Pal was found in **1.**
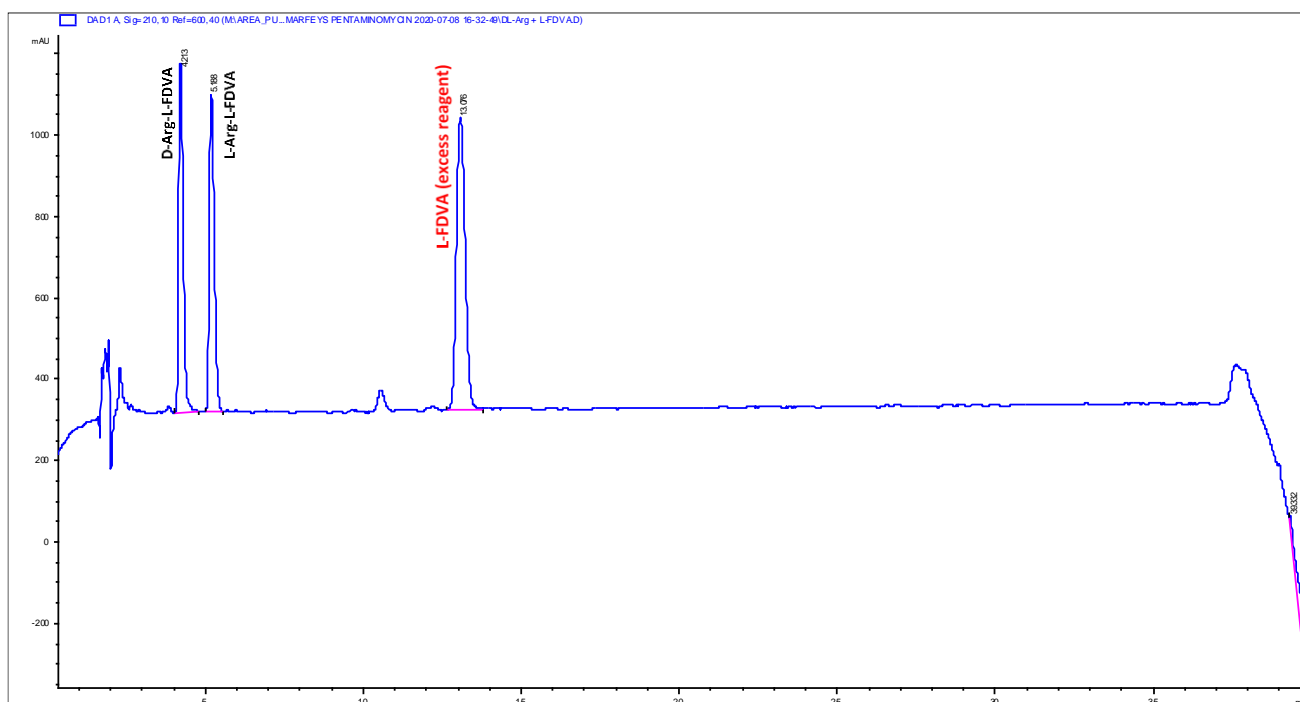
Figure S37. LC-UV chromatogram of a mixture of L-FDVA derivatization reactions of L-Arg / D-Arg standard amino acids present in **1.**
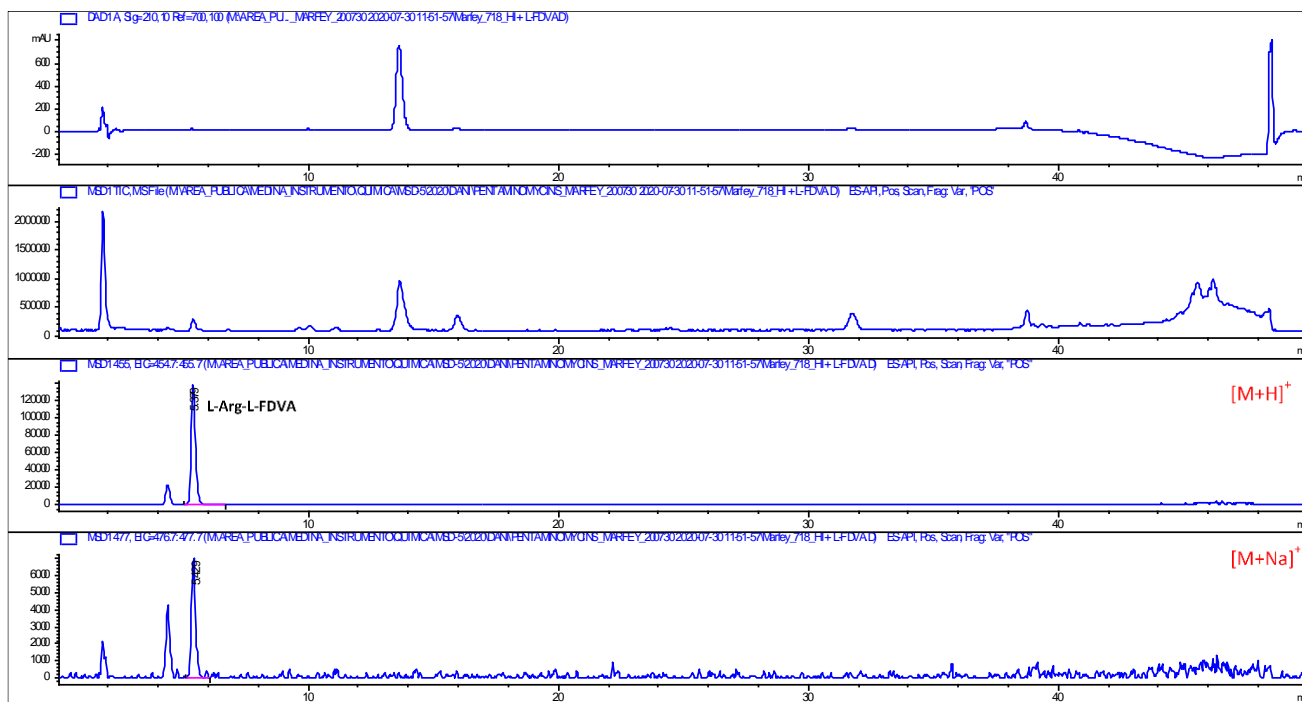
Figure S38. LC-UV and extracted-ion chromatograms (EIC) of Arg-L-FDVA derivatized hydrolysis of **1** with HI. L-Arg was found in **1.**
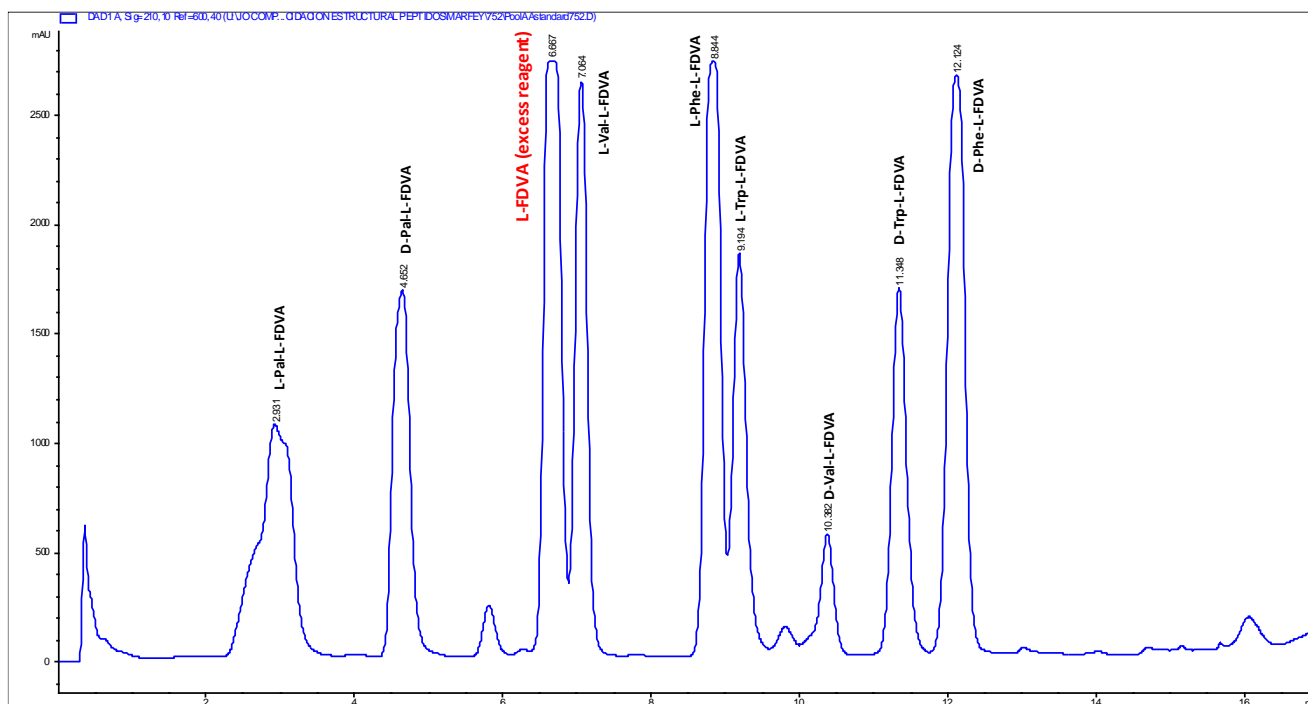
Figure S39. LC-UV chromatogram of a mixture of L-FDVA derivatization reactions of L-Val / D-Val / L-Trp / D-Trp / L-2-Pal / D-2-Pal / L-Phe / D-Phe standards amino acids present in **2.**
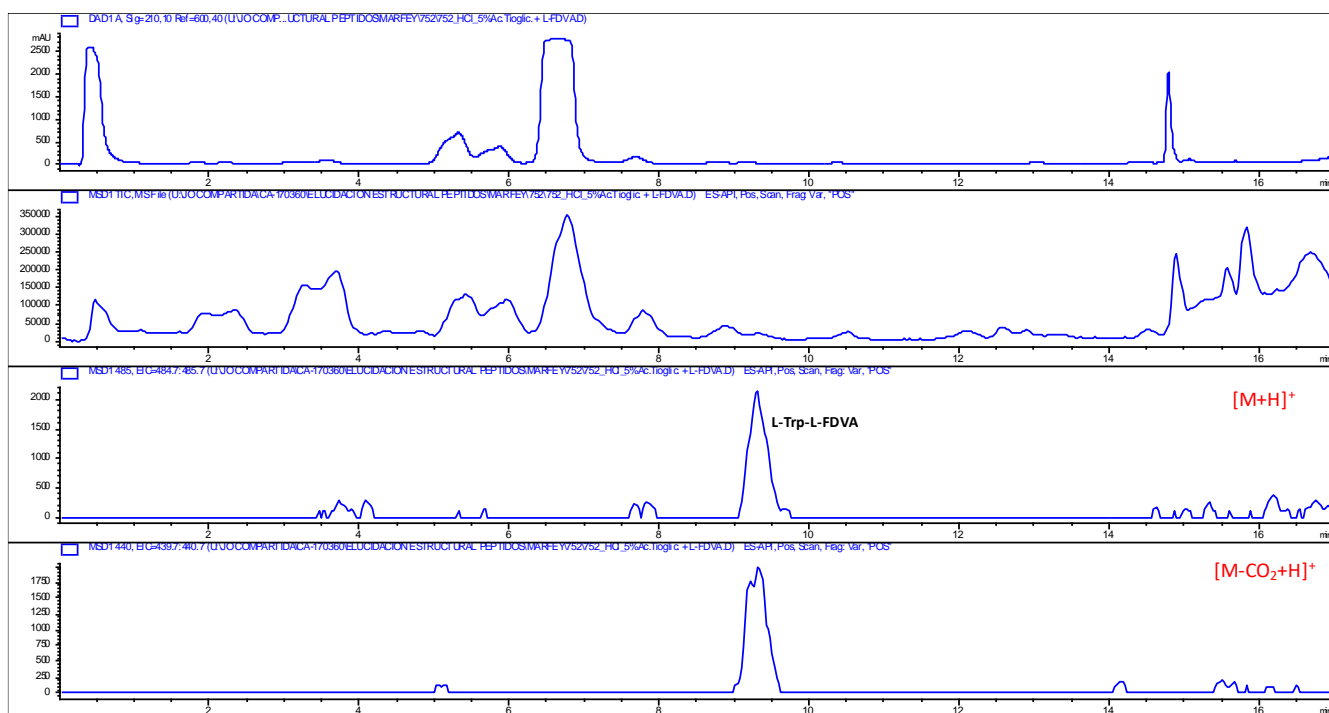
Figure S40. LC-UV and extracted-ion chromatograms (EIC) of Trp-L-FDVA derivatized hydrolysis of **2** with HCl (5% Thioglycolic acid). L-Trp was found in **2.**
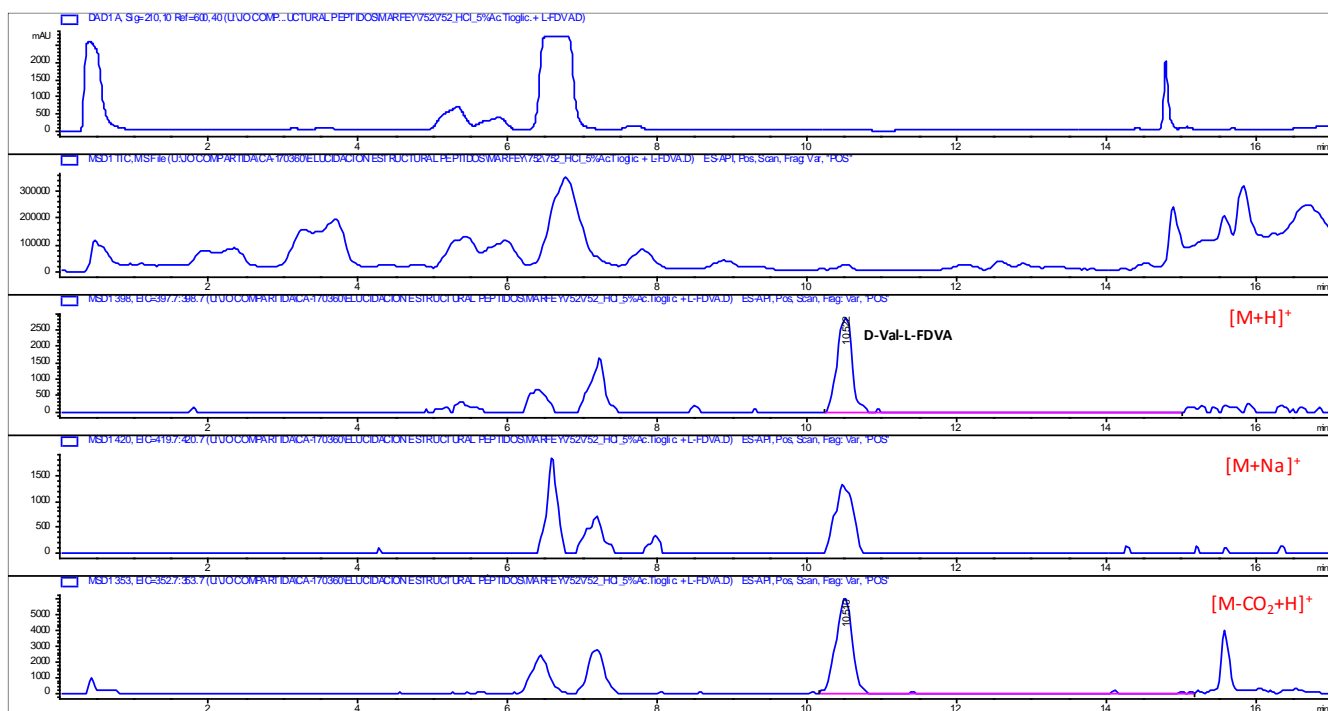
Figure S41. LC-UV and extracted-ion chromatograms (EIC) of Val-L-FDVA derivatized hydrolysis of **2** with HCl (5% Thioglycolic acid). D-Val was found in **2.**
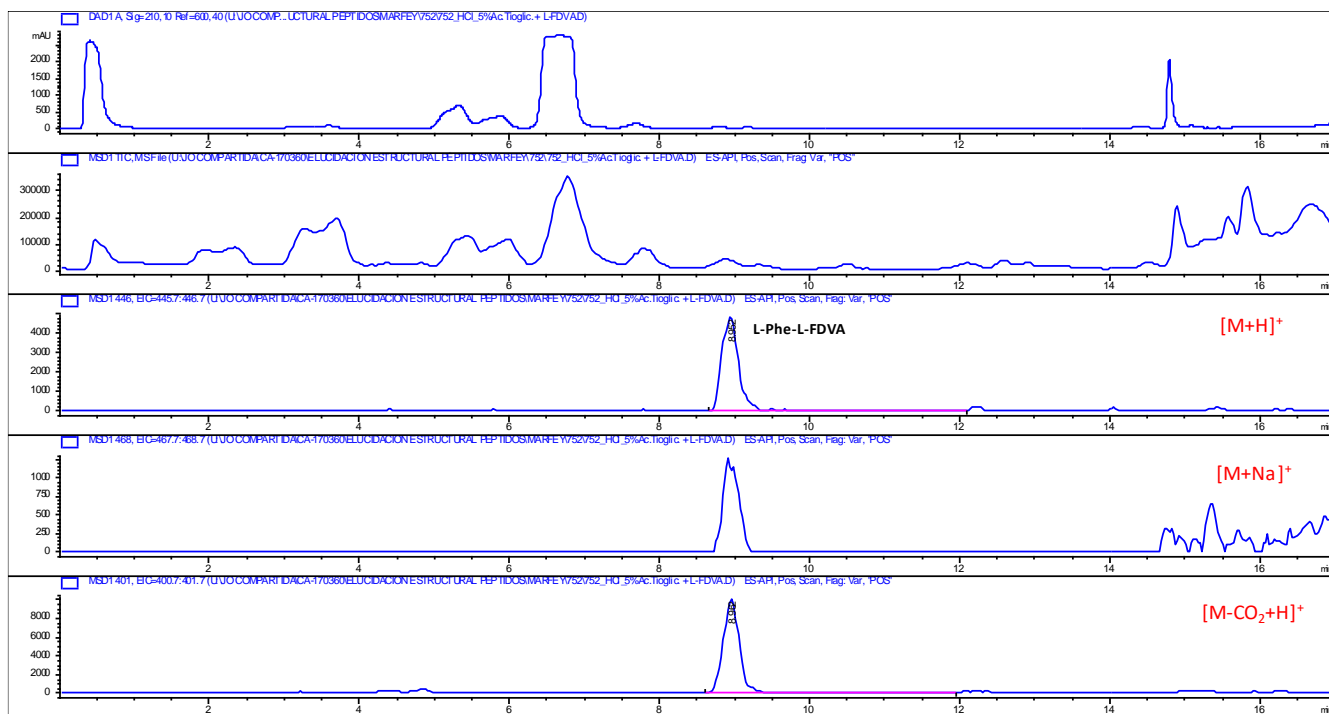
Figure S42. LC-UV and extracted-ion chromatograms (EIC) of Phe-L-FDVA derivatized hydrolysis of **2** with HCl (5% Thioglycolic acid). L-Phe was found in **2.**
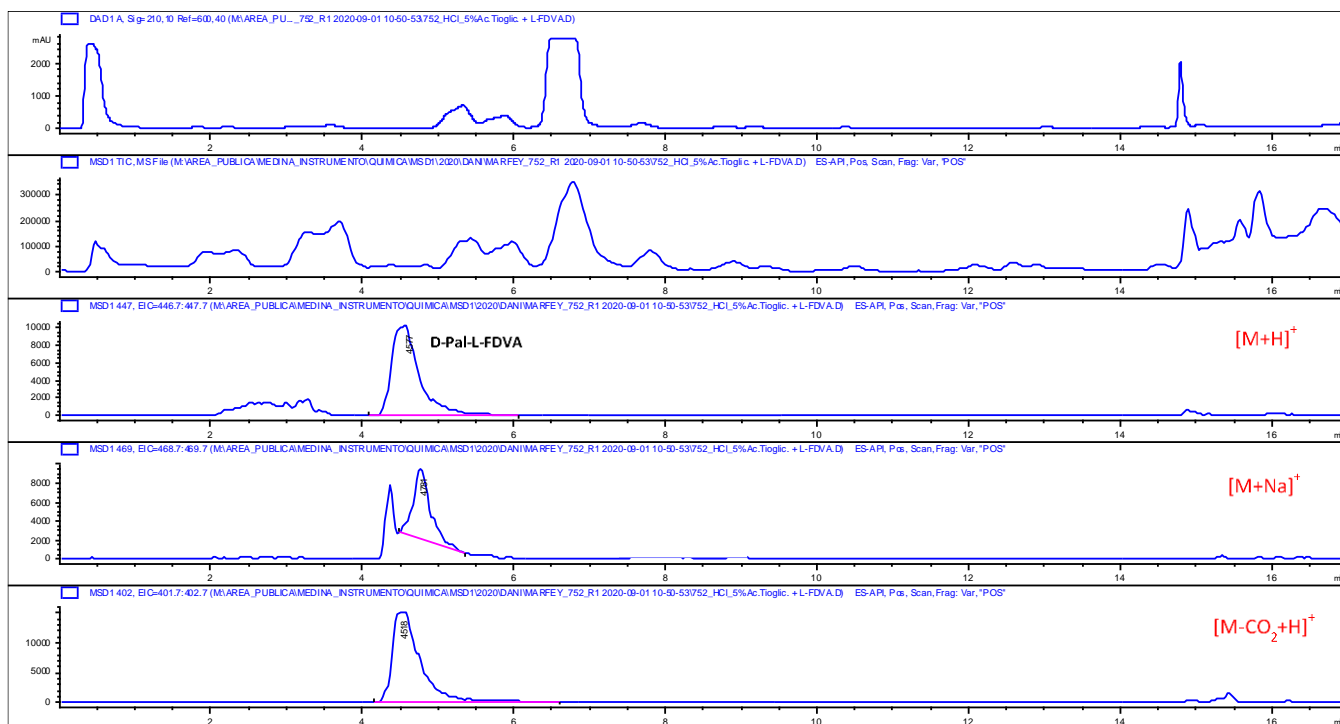
Figure S43. LC-UV and extracted-ion chromatograms (EIC) of 2-Pal-L-FDVA derivatized hydrolysis of **2** with HCl (5% Thioglycolic acid). D-2-Pal was found in **2.**
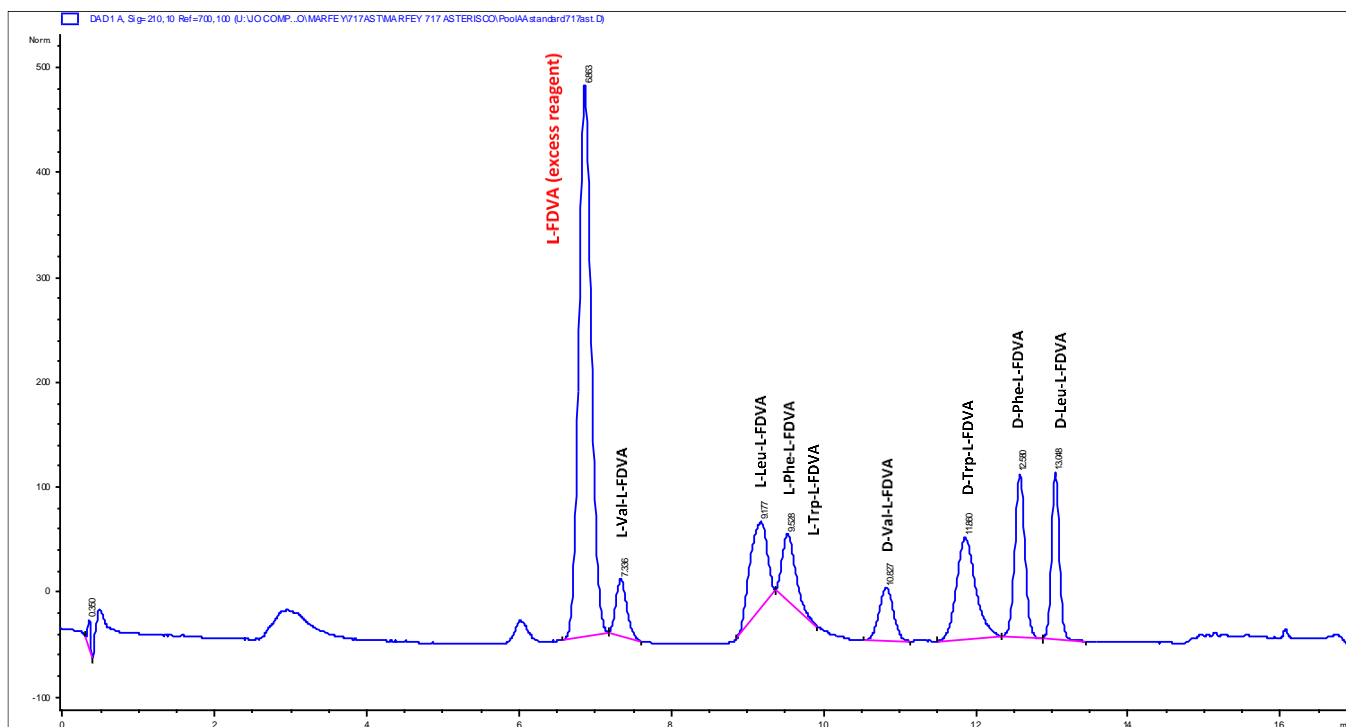
Figure S44. LC-UV chromatogram of a mixture of L-FDVA derivatization reactions of L-Val / D-Val / L-Trp / D-Trp / L-Leul / D-Leu / L-Phe / D-Phe standard amino acids present in **3.**
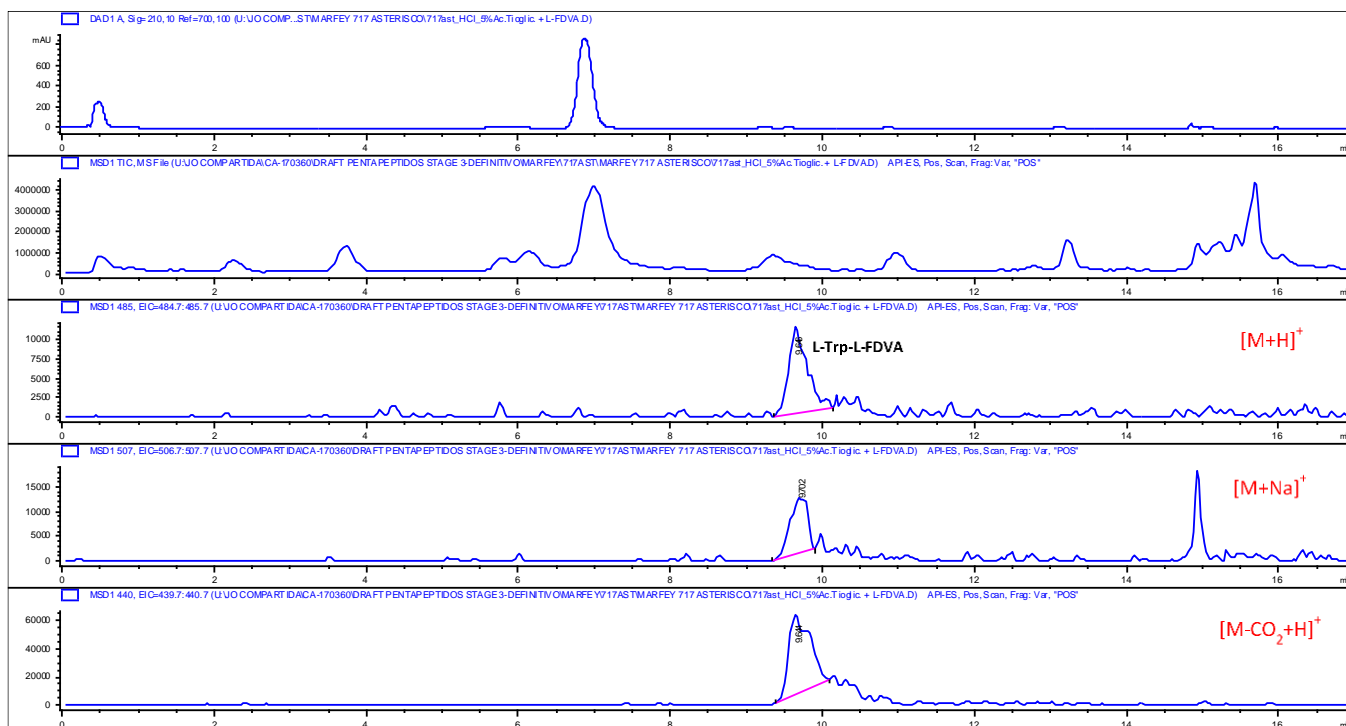
Figure S45. LC-UV and extracted-ion chromatograms (EIC) of Trp-L-FDVA derivatized hydrolysis of **3** with HCl (5% Thioglycolic acid). L-Trp was found in **3.**
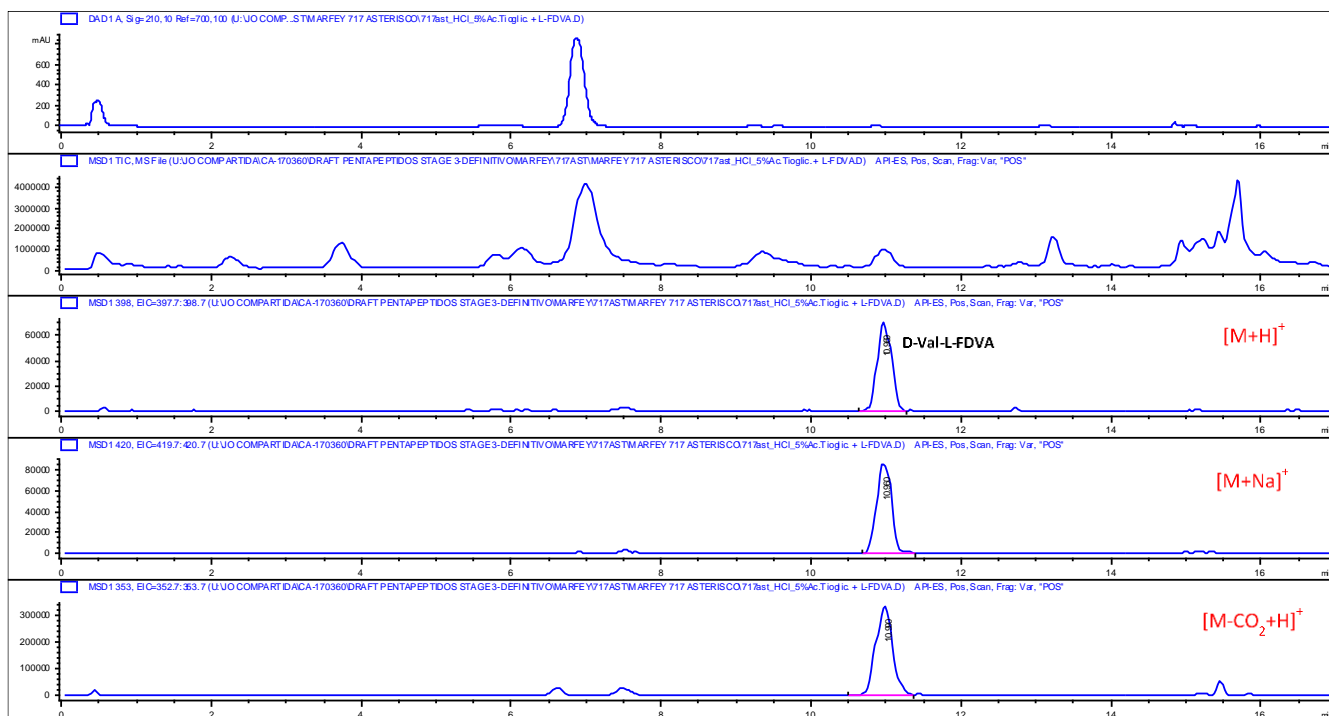
Figure S46. LC-UV and extracted-ion chromatograms (EIC) of Val-L-FDVA derivatized hydrolysis of **3** with HCl (5% Thioglycolic acid). D-Val was found in **3**.
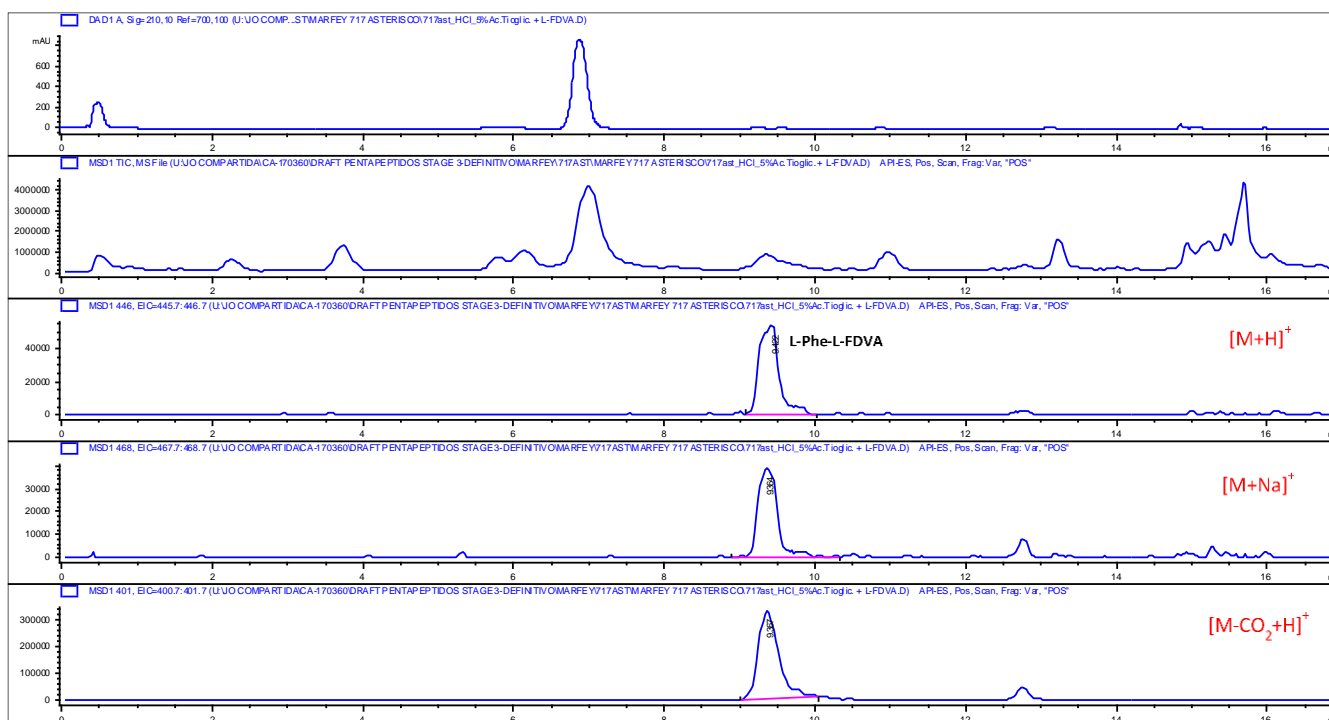
Figure S47. LC-UV and extracted-ion chromatograms (EIC) of Phe-L-FDVA derivatized hydrolysis of **3** with HCl (5% Thioglycolic acid). L-Phe was found in **3**.
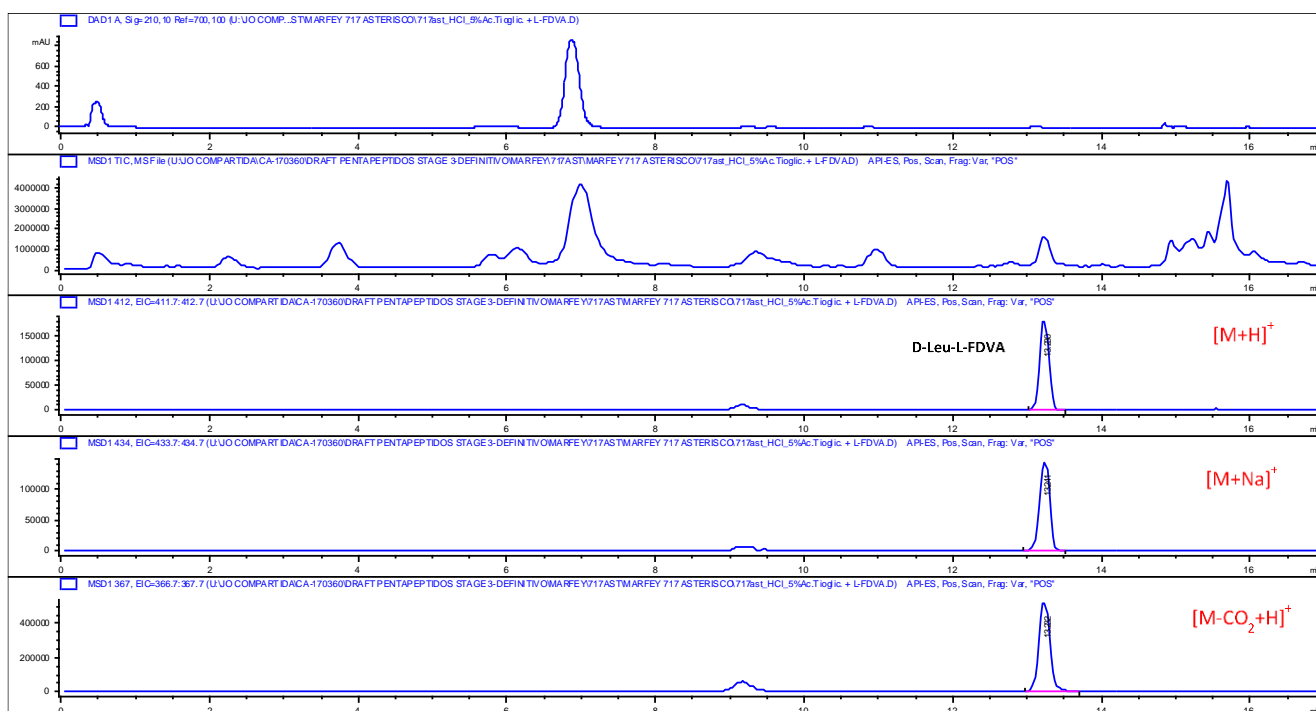
Figure S48. LC-UV and extracted-ion chromatograms (EIC) of Leu-L-FDVA derivatized hydrolysis of **3** with HCl (5% Thioglycolic acid). D-Leu was found in **3.**
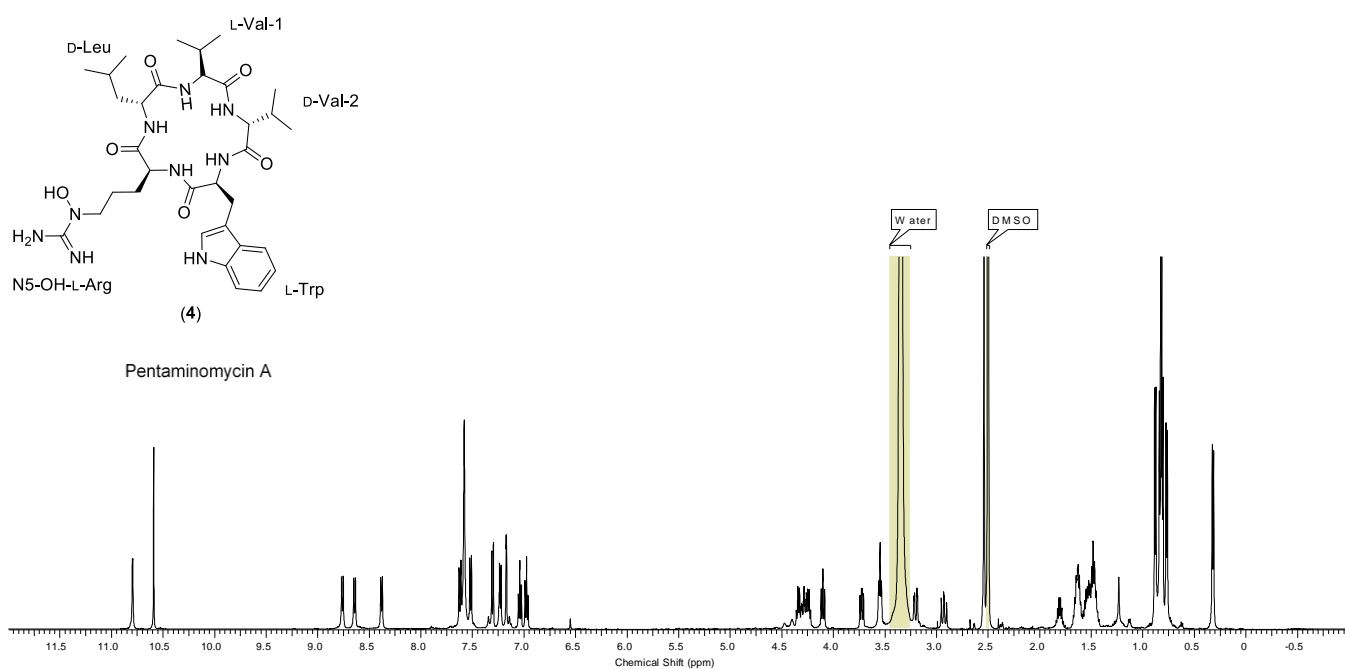
Figure S49. ^{1}H NMR spectrum (DMSO-d_6, 500 MHz) of **4.**
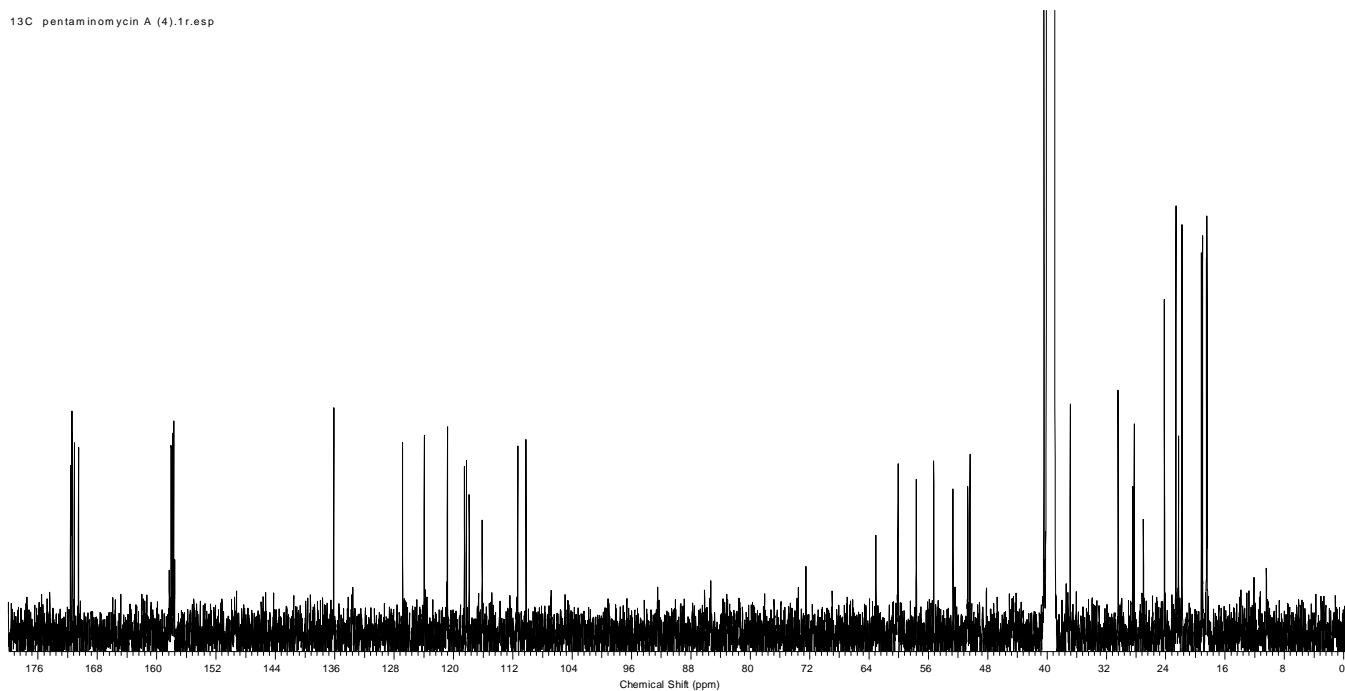
Figure S50. ^{13}C NMR spectrum (DMSO-d_6, 125 MHz) of **4.**
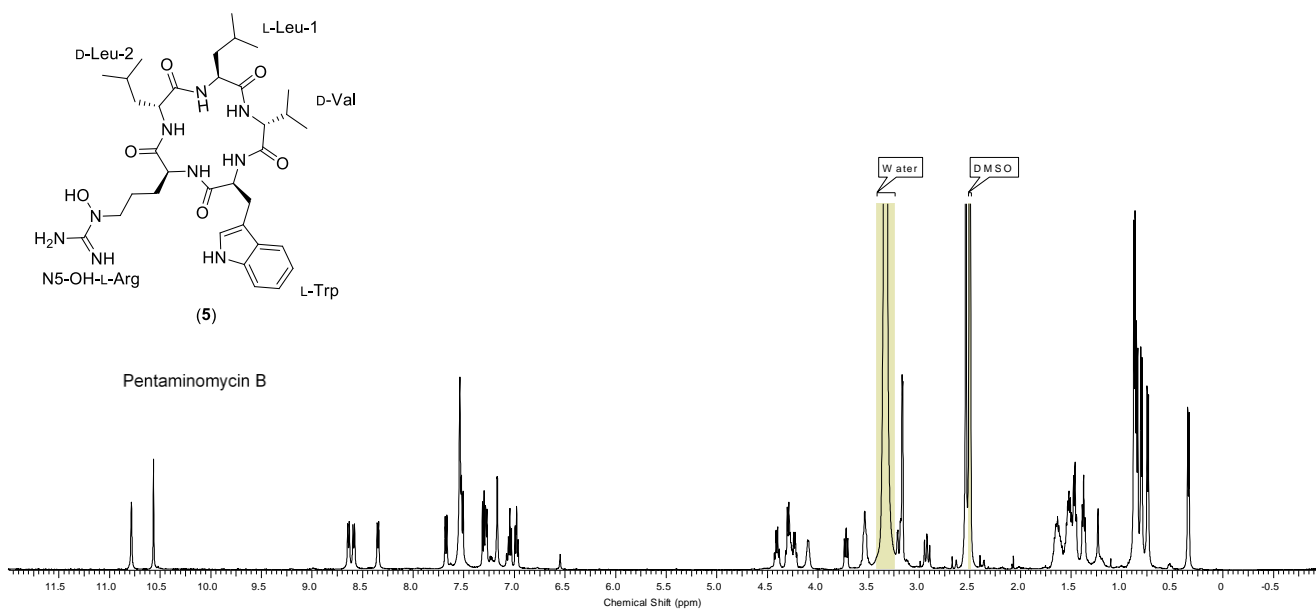
Figure S51. ^{1}H NMR spectrum (DMSO-d_6, 500 MHz) of **5.**
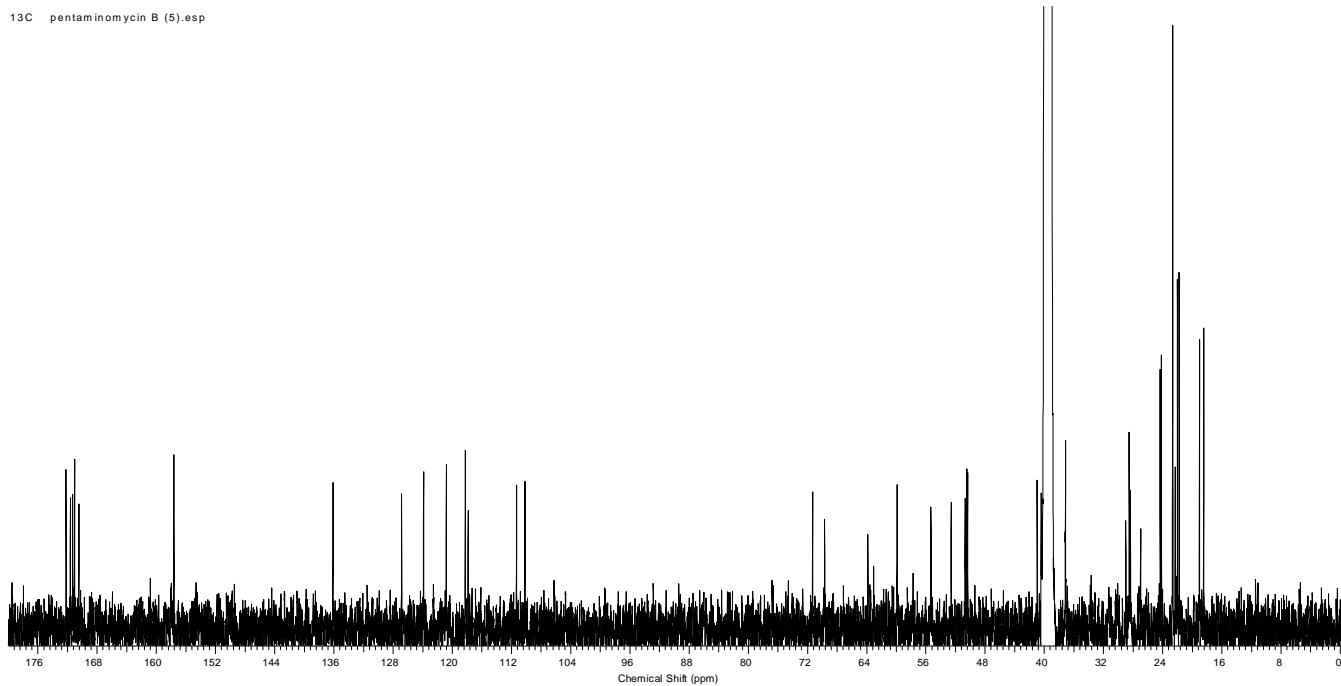
Figure S52. ^{13}C NMR spectrum (DMSO-d_6, 125 MHz) of **5.**
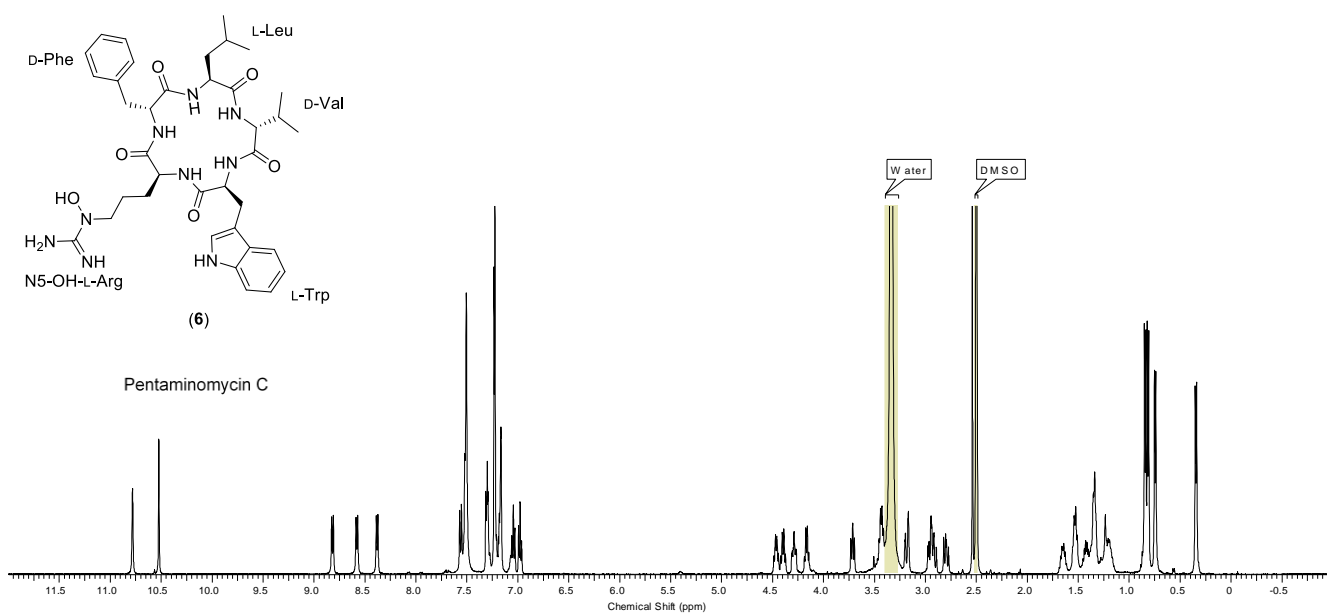
Figure S53. ¹H NMR spectrum (DMSO-d_6, 500 MHz) of **6.**
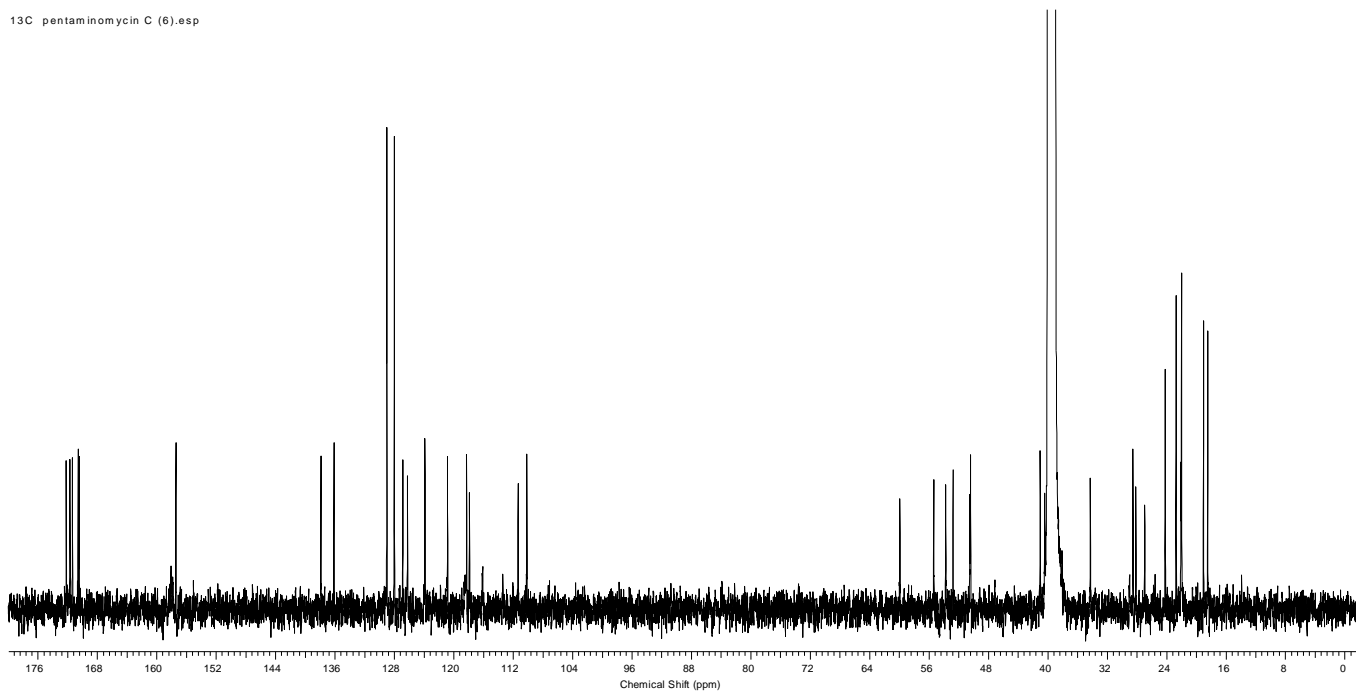
Figure S54. ¹³C NMR spectrum (DMSO-d_6, 125 MHz) of **6.**
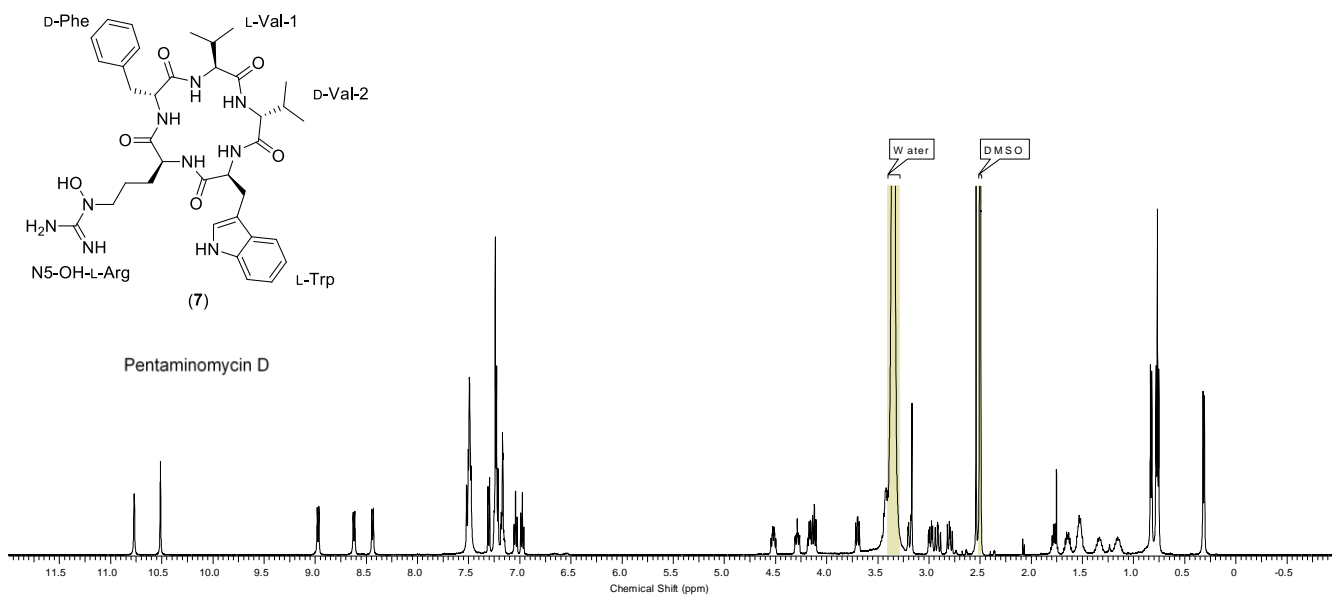
Figure S55. ¹H NMR spectrum (DMSO-d_6, 500 MHz) of **7.**
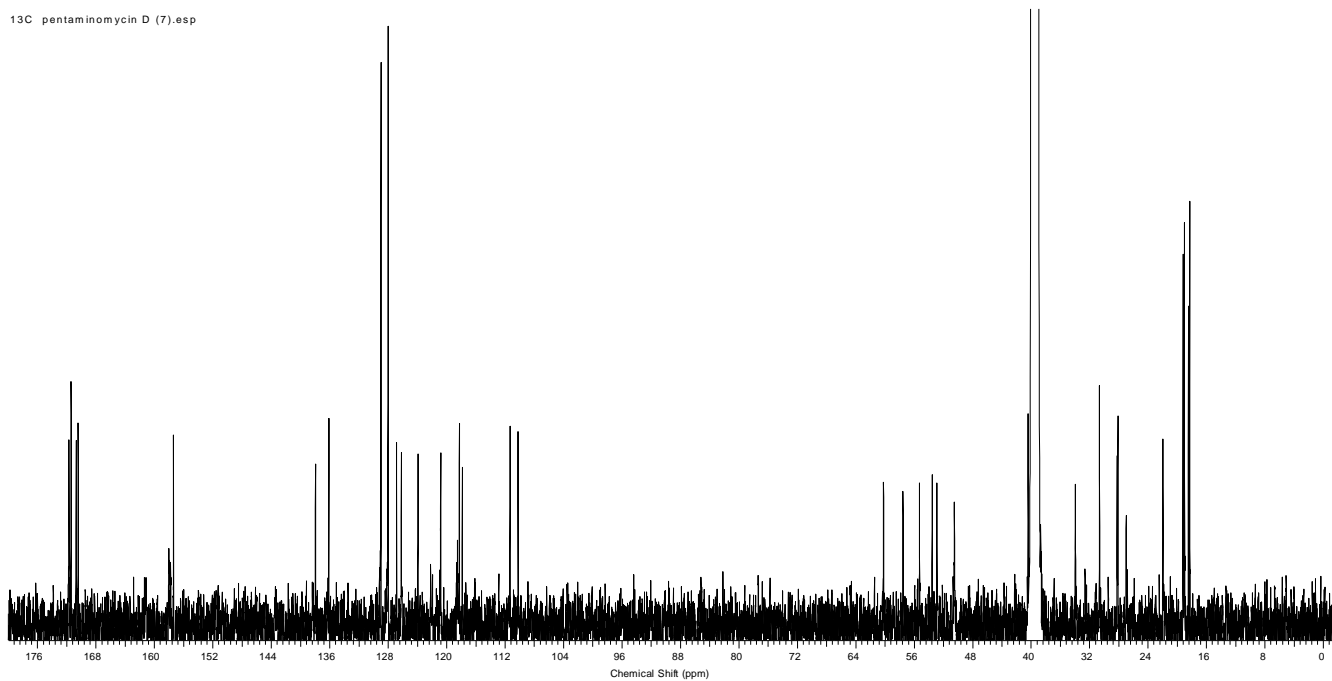
Figure S56. ¹³C NMR spectrum (DMSO-d_6, 125 MHz) of **7.**
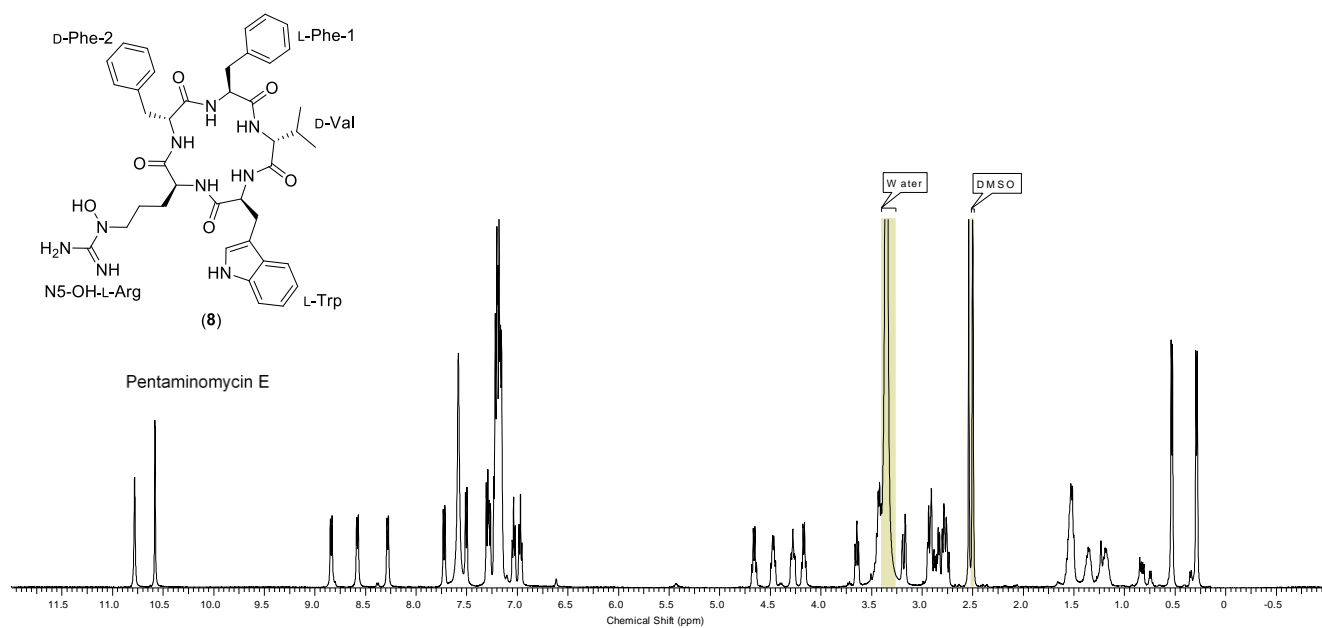
Figure S57. ¹H NMR spectrum (DMSO-d_6, 500 MHz) of **8**.
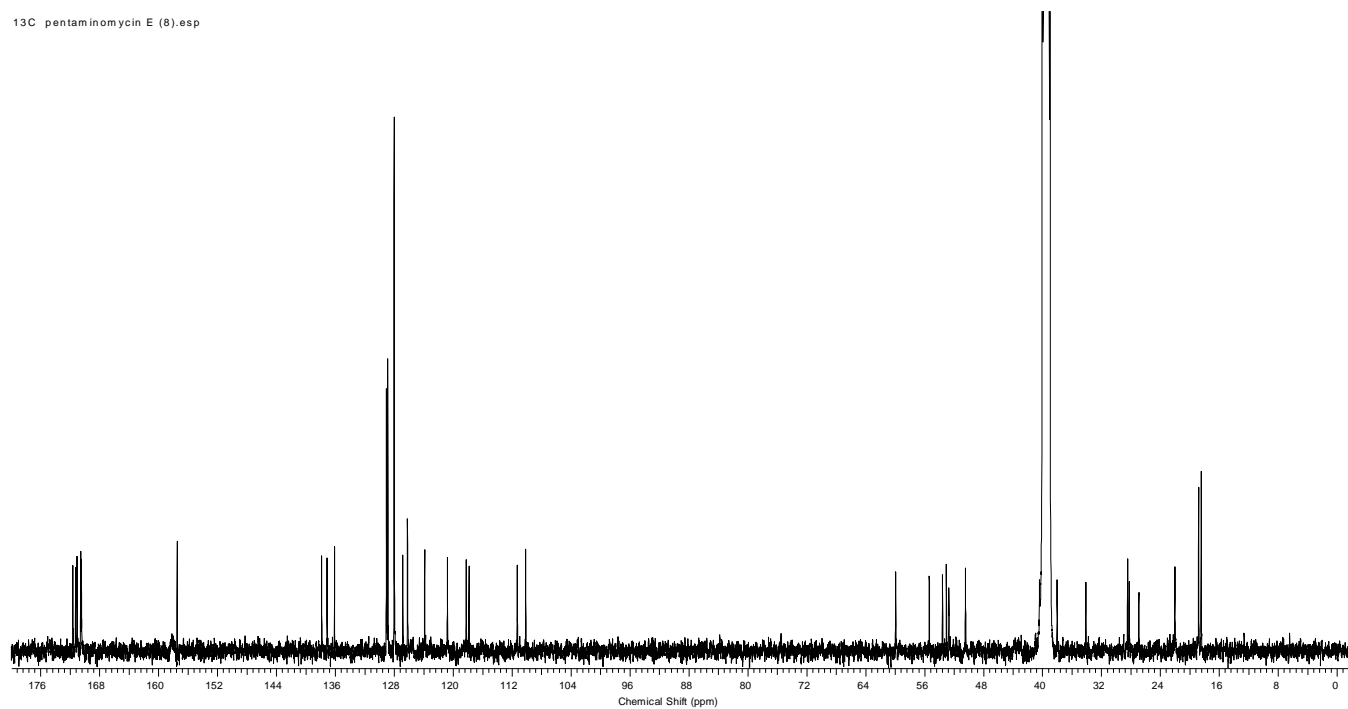
Figure S58. ¹³C NMR spectrum (DMSO-d_6, 125 MHz) of **8.**
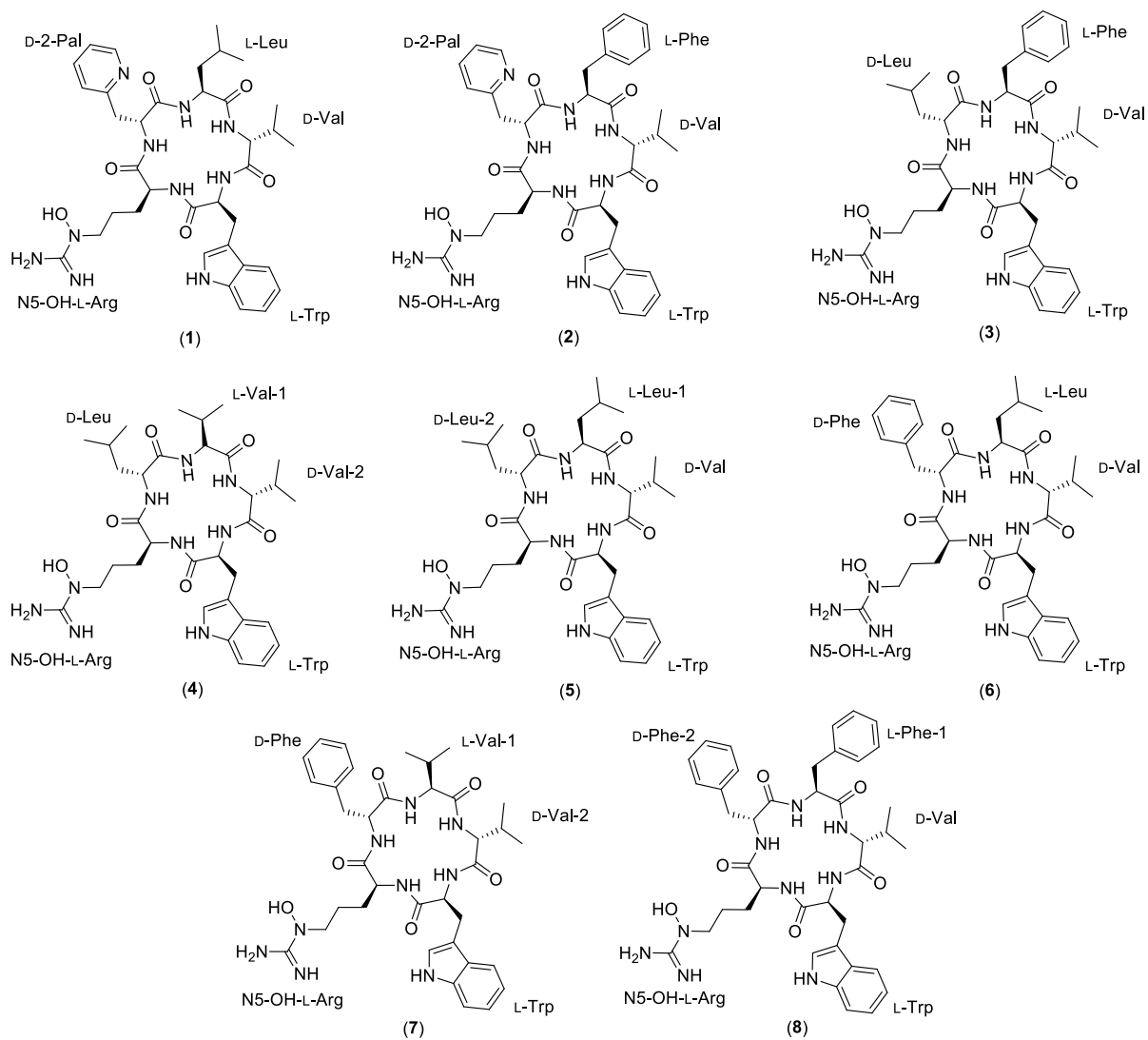
Figure S59. Pentaminomycins A-H isolated from strain CA-170360 (**1-8**).
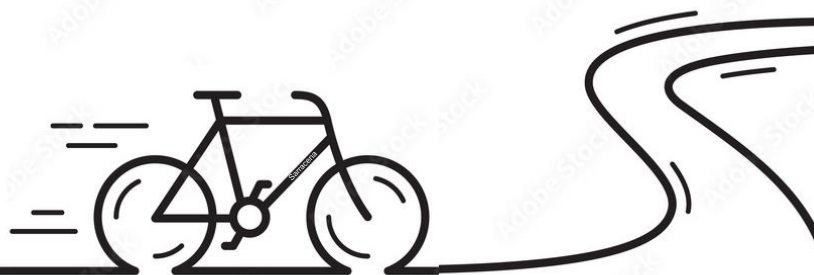
CHAPTER 3

Chapter 3. Discovery of gargantulides B and C, new 52-membered macrolactones from *Amycolatopsis* sp. The complete absolute stereochemistry of the gargantulide family

Daniel Carretero-Molina,[1,#] Francisco Javier Ortiz-López,[1,#,*] Tetiana Gren,[2,#] Daniel Oves-Costales,[1] Jesús Martín,[1] Fernando Román-Hurtado,[1] Tue Sparholt Jørgensen,[2] Mercedes de la Cruz,[1] Caridad Díaz,[1] Francisca Vicente,[1] Kai Blin,[2] Fernando Reyes,[1] Tilmann Weber[2,*] and Olga Genilloud[1,*]

Resumen: Las gargantulidas B y C, dos nuevas y altamente complejas macrolactonas glicosiladas de 52 miembros, fueron aisladas de la cepa *Amycolatopsis* sp. CA-230715 durante una campaña de evaluación antibacteriana. Las estructuras de estos enormes macrólidos se elucidaron mediante espectroscopia de RMN 2D y demostraron estar relacionadas con gargantulida A, aunque conteniendo monosacáridos adicionales de β-glucopiranosa y/o α-arabinofuranosa unidos por separado a sus estructuras principales. La secuenciación del genoma permitió la identificación de un grupo de genes biosintéticos sorprendentemente grande de 216 kbp, uno de los grupos de PKS de tipo I más grandes descritos hasta la fecha, y la propuesta de una ruta biosintética, previamente no reportada, para gargantulidas A-C. Las configuraciones absolutas de gargantulida B y C se asignaron en base a una combinación de RMN y análisis bioinformático de los dominios de cetoreductasa y enoilreductasa dentro de la PKS de tipo I multimodular. Además, se ha revisado y completado la estereoquímica absoluta para gargantulida A. Las gargantulidas B y C presentan una potente actividad antibacteriana frente a un set de bacterias Gram-positivas multiresistentes y una actividad moderada contra el patógeno Gram-negativo de gran relevancia clínica *Acinetobacter baumannii*.

Affiliations

[1] Fundación MEDINA, Centro de Excelencia en Investigación de Medicamentos Innovadores en Andalucía, Avda. del Conocimiento 34, 18016, Armilla (Granada), Spain; daniel.carretero@medinaandalucia.es (D.C.-M.); jesus.martin@medinaandalucia.es (J.M.); ignacio.gonzalez@medinaandalucia.es (I.G.); marina.sanchez@medinaandalucia.es (M.S.-H.); fernando.roman@medinaandalucia.es (F.R.-H); caridad.diaz@medinaandalucia.es (C.D.); mercedes.delacruz@medinaandalucia.es (M.d.l.C.);

[2] The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kemitorvet, building 220, 2800 Kgs. Lyngby, Denmark; tetgre@biosustain.dtu.dk (T.G); tuspjo@biosustain.dtu.dk (T.S.J.); kblin@biosustain.dtu.dk (K.B)

[#] D.C.-M., F.J.O.-L, and T. G contributed equally to this work.

[*] Correspondence: javier.ortiz@medinaandalucia.es (F.J.O.-L.); olga.genilloud@medinaandalucia.es (O.G.) and tiwe@biosustain.dtu.dk (T.W). Tel.: +34 958 993 965 (F.J.O.-L. and O.G.); +45 24 89 61 32 (T.W.)

Abstract: Gargantulides B and C, two new and highly complex 52-membered glycosylated macrolactones, were isolated from *Amycolatopsis* sp. strain CA-230715 during an antibacterial screening campaign. The structures of these giant macrolides were elucidated by 2D NMR spectroscopy and shown to be related to gargantulide A, although containing additional β-glucopyranose and/or α-arabinofuranose monosaccharides separately attached to their backbones. Genome sequencing allowed the identification of a strikingly large 216 kbp biosynthetic gene cluster, among the largest type I PKS clusters described so far, and the proposal of a previously unreported biosynthetic pathway for gargantulides A-C. The absolute configurations of gargantulides B and C were assigned based on a combination of NMR and bioinformatics analysis of ketoreductase and enoylreductase domains within the multimodular type I PKS. In addition, the absolute stereochemistry of gargantulide A has now been revised and completed. Gargantulides B and C display potent antibacterial activity against a set of drug-resistant Gram-positive bacteria and moderate activity against the clinically relevant Gram-negative pathogen *Acinetobacter baumannii*.

1. Introduction

Natural products remain a source of privileged chemical scaffolds with high structural diversity and prominent biological activities which can serve to address urgent clinical needs.[54] Among them, macrolides[295] is a particularly large and widespread structural class of secondary metabolites mainly produced by actinomycetes[124] and typically encoded by multimodular type I polyketide synthase (T1PKS) biosynthetic gene clusters (BGC). The best-known macrolide is erythromycin, not only relevant as an antibacterial agent but also as a model to study the biosynthesis of macrolide antibiotics.[296] Other clinically relevant members of this class are, for example, amphotericin B, avermectin, or rapamycin, which cover antimicrobial, insecticide, and immunosuppressant activities.[295,297] In the last years, some of the largest and more complex macrocyclic polyketides ever isolated from actinomycetes have been reported, including quinolidomicin A_1 (60-membered macrolactone)[298] and stambomycins A-D (51-membered).[299] Establishing the structures of such complex and stereochemically rich large polyol macrolides has usually been extremely complicated and implied tedious chemical approaches.[300–303] Fortunately, the advances in genome mining tools and the combination of NMR and bioinformatics analysis of T1PKS gene clusters have arisen as a powerful approach to assign the stereochemistry of these complex polyol macrolides,[159,160,304–306] greatly facilitating this task.

Recently, the discovery of gargantulide A, a complex 52-membered macrolactone with antibacterial activity against Gram-positive pathogens isolated from a *Streptomyces* sp. strain was reported.[307] However, no genetic information about the biosynthesis of this macrolide has been provided to date. Based only on NMR analysis and given its structural complexity, the stereochemistry assignment of gargantulide A could not be fully accomplished in the original work. During a screening program targeting a panel of Gram-negative pathogens, culture extracts of the strain *Amycolatopsis* sp. CA-230715 displayed selective antimicrobial activity against *Acinetobacter baumannii*. A bioassay-guided purification process led to the isolation of gargantulides B (**1**) and C (**2**) (Fig. 1), two new giant glycosylated macrolactones structurally related to their formerly isolated congener gargantulide A (**3**).

The planar structures of the new bioactive macrolides were fully elucidated by 2D NMR spectroscopy and were shown to contain the same 52-membered polyol ring and 22-membered side chain as that reported for gargantulide A, although showing additional glucose and/or arabinofuranose glycosylation at their polyketide backbones. The BGC putatively responsible for the biosynthesis of gargantulides B and C (*gar*) was identified in the genome of the producing strain and was reasoned to be responsible for the production of gargantulide A as well. The 216 kbp *gar* gene cluster codes for a giant T1PKS comprising 10 subunits and 35 modules, thus representing, - to our best knowledge-, the largest macrolide-encoding BGC discovered so far. Interestingly, it also encodes genes presumably responsible for the biosynthesis of the rare amino sugar β-3,6-deoxy-3-methylamino glucose (maG), which is present exclusively in gargantulides A-C. Integration of genomic data from bioinformatics analysis of the T1PKS enzymes and detailed NMR-based analysis allowed us to confidently assign the full stereochemistry of the polyketide chain of gargantulides B and C and to revise and complete that of gargantulide A. In addition, we determined the absolute configuration of the additional monosaccharide residues in gargantulides B and C, thus completing the full stereostructure of these giant macrolides (Fig. 1).
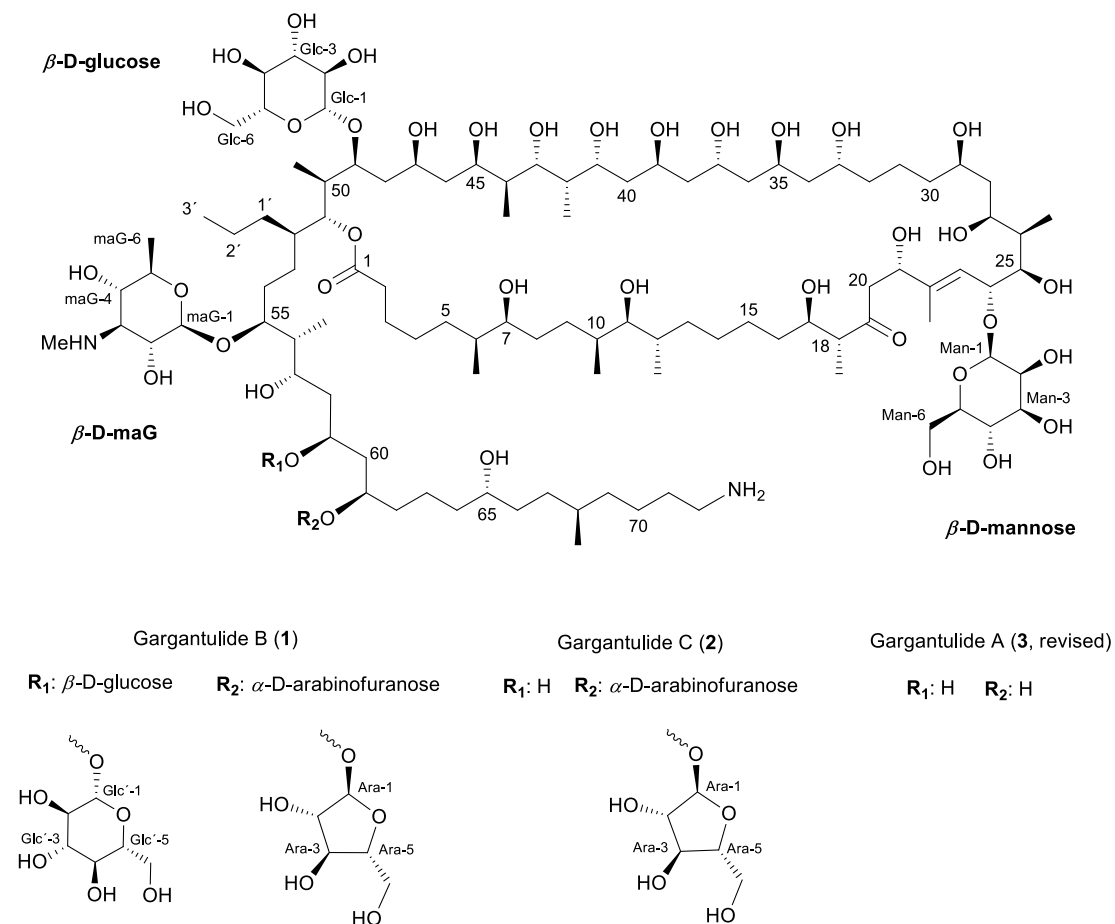
Figure 1. Structures of gargantulide B (**1**) and C (**2**) isolated from *Amycolatopsis sp.* CA-230715 and that of gargantulide A (**3**, revised in this work).

2. Results and Discussion

2.1. Isolation and planar structures of gargantulides B and C

The acetone extract of a fermentation broth of the strain *Amycolatopsis sp.* CA-230715 showed selective antibacterial activity against *Acinetobacter baumannii* MB5973. Bioactivity-guided fractionation (Fig. S1) and subsequent HRMS-ESI analyses of the active fractions traced this activity to gargantulides B (**1**) and C (**2**), two peaks with [M+H]$^+$ ions at *m/z* 2392.4804 and 2230.4286 (Fig. S2 and S3), which were indicative of the previously unreported molecular formulae $C_{116}H_{218}N_2O_{47}$ (Δ -0.13 ppm) and $C_{110}H_{208}N_2O_{42}$ (Δ +0.22 ppm) respectively. These formulae assignments were further supported by the presence of [M+2H]$^{2+}$ and [M+3H]$^{3+}$ ions for each compound (Fig. S2 and S3) and advanced the close relationship between them. Searches in natural products databases failed to identify known compounds with the observed exact masses for **1** and **2**, suggesting that they were new secondary metabolites. The structures of gargantulides B (**1**) and C (**2**) were therefore fully elucidated by combining extensive NMR spectroscopy, genome-based bioinformatics, and chemical approaches.

The planar structure of **1** was determined by 1D (^{1}H and ^{13}C) and 2D NMR (COSY, TOCSY, HSQC, HSQC-TOCSY, and HMBC) (Table S1, Fig. S4). Detailed analyses of the ^{13}C NMR and HSQC spectra revealed the presence of 3 quaternary carbons (including one ketone at δ_C 214.6 and one ester at δ_C 175.2), 1 olefinic methine (δ_C 123.1), 46 oxygenated methines (including five anomeric carbons at δ_C 108.2, 104.4, 103.9, 102.4 and 97.0 ppm, representing five sugar moieties) 12 aliphatic methines, 40 methylenes (one of them being attached to an amino group, at δ_C 41.0) and 14 methyl carbons (including a *N*HMe group at δ_C 31.1). From 9

double-bond equivalents deduced from the molecular formula, one was assigned to an olefinic double bond and two to carbonyl moieties, suggesting that **1** contained six rings in its structure. All these data jointly advanced the glycosylated polyketide macrolactone nature of **1**. Furthermore, many of these NMR features and the high molecular weight of **1** resembled those of the previously reported macrolide gargantulide A (**3**).[307] Indeed, a comparison of the HSQC and ^{13}C NMR spectra of **1** (Fig. S4) with those reported for **3** evidenced the close relationship between both compounds. From the analysis of COSY and TOCSY spectra (Fig. S4) three spin systems comprising C-2 to C-18, C-20 to C-21, and C-23 to C-72, could be constructed (Fig. 2). For the assignment of contiguous methylenes within each segment, we relied on key HSQC-TOCSY and HMBC correlations (Fig. S4). Further ^{1}H–^{13}C-HMBC correlations observed from the methylene hydrogens H-2 and H-3 to carbonyl C-1 (δ_C 175.2) and from H-17/H-18 and H-20/H-21 to the carbonyl ketone C-19 (δ_C 214.6), enabled the extension of the linear polyketide chain from C-1 to C-21 (Fig. 2). Additionally, intense HMBC cross-peaks from H-21 to the olefinic carbons C-22/C-23 linked the last independent spin systems and allowed the construction of the whole polyketide backbone from C1 to C-72. Finally, one long-range correlation from H-51 to C-1 alongside the downfield-shifted resonance frequency of the oxymethine H-51 (δ_H 4.88, Table S1) formally assembled a 52-membered macrolide ring system (Fig. 2).
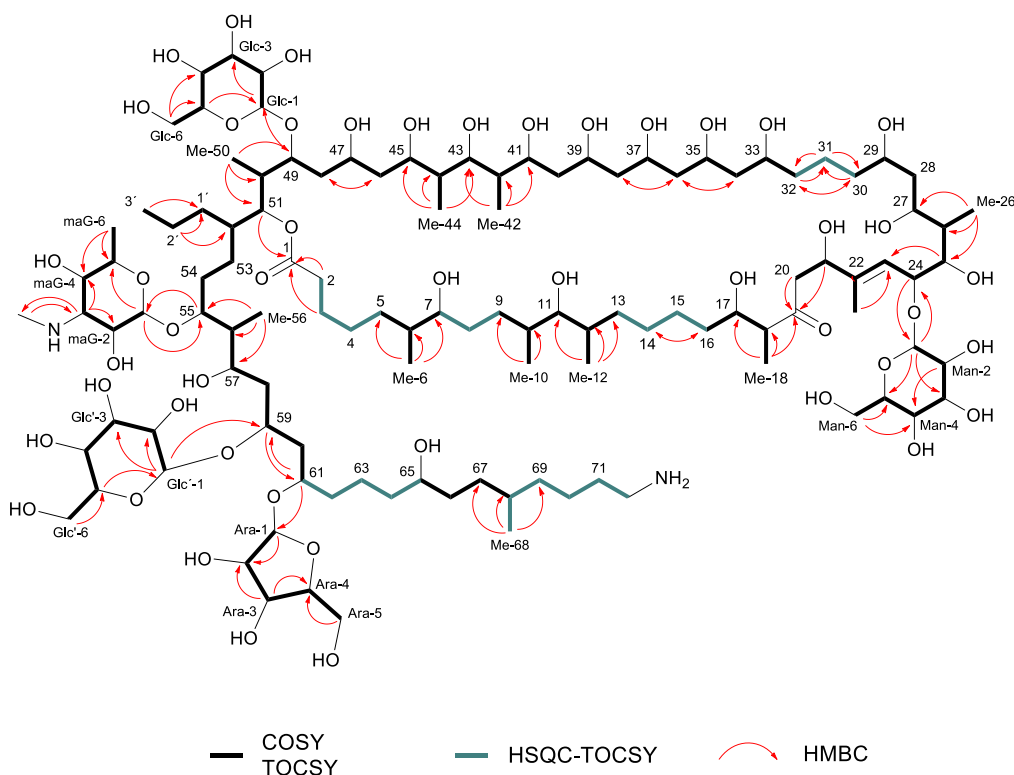
Figure 2. Key COSY, TOCSY, HSQC-TOCSY, and HMBC correlations observed for gargantulide B (**1**).

Gargantulide B (**1**) (Fig. 1) has only one double bond between C22/C-23 whose *E* geometry was determined based on NOESY correlations between H-21/H-23 and Me-22/H-24. The presence of five monosaccharide residues in **1** was revealed by the analysis of the remaining 2D NMR data (Fig. S4). Starting from the anomeric protons at δ_H 4.58, 4.20, and 4.46, three hexose units were constructed by analysis of COSY and TOCSY spectra. Based on $^{3}J_{H,H}$ coupling constants, and NOESY correlations, we identified the three pyranose sugar components as β-mannose, β-glucose and the rare β-3,6-deoxy-3-methylamino glucose (maG), already reported as components of gargantulide A. Key HMBC correlations between

their anomeric protons and the corresponding positions in the backbone of **1** at C-24, C-49 and C-55, respectively, allowed us to establish the independent *O*-glycosidic linkages of these residues at positions equivalent to those described for **3**. An additional hexose residue was found in **1** and determined to be an extra β-glucose unit based on the $^3J_{H,H}$ coupling constants, and *nOe* cross-peaks. HMBC correlations indicated that this sugar residue was *O*-attached to C-59 in the backbone of **1**. Finally, the remaining 2D NMR data (Fig. S4) accounted for the presence of a pentofuranose moiety in **1**. The α-arabinofuranosyl configuration was assigned from ^{13}C-NMR fingerprint chemical shift values,[308,309] as well as from small coupling constants $^3J_{1,2}$ (1.6 Hz) and $^3J_{2,3}$ (3.9 Hz) (Table S1). Based on HMBC correlations, this arabinose unit was found to be *O*-linked to C-61 in gargantulide B (**1**).

As advanced above, gargantulide C (**2**) (Fig. 1) was assigned the molecular formula $C_{110}H_{208}N_2O_{42}$ by HRMS-ESI(+)-TOF ([M+H]$^+$ *m/z* 2230.4281, calculated for $[C_{110}H_{209}N_2O_{42}]^+$ 2230.4274) (Fig. S3), accounting for eight degrees of unsaturation, one less than gargantulide B (**1**). The ^{1}H, ^{13}C, and HSQC NMR data of **2** (Table S1, Fig. S5) showed high similarity to those of **1**, except for the absence of the signals corresponding to the β-glucose moiety *O*-attached to C-59. Further analysis of COSY and HSQC-TOCSY spectra confirmed the lack of this glucose residue in **2**, also reflected in the upfield-shifted resonance of C-59 compared to **1** (Table S1) and consistent with the relationship between molecular formulae of both compounds. The connectivity of the polyketide backbone of **2** was established based on 2D NMR correlations (Fig. S5 and S6) and found to be identical to that of **1**. The four sugar moieties present in **2** were also identified as β-mannose, β-glucose, β-3,6-deoxy-3-methylamino glucose (maG), and α-arabinofuranose (analogously to **1**), and their respective glycosidic linkages positions to the backbone were found to be the same as in **1** (C-24, C-49, C-55, and C-61, respectively).

Interestingly, although the bioassay-guided isolation process led to the isolation of gargantulides B (**1**) and C (**2**), species with an exact mass and UV spectrum identical to those of the formerly reported congener gargantulide A was co-detected together with **1** and **2** in a parallel small-scale fermentation (Fig. S7). Unfortunately, the low production titers for this compound prevented its isolation and confirmation as gargantulide A. Nevertheless, this finding, along with the high similarity between the three macrolides, strongly supports a common biosynthetic origin for all of them. Since no genetic or biochemical information on this class of macrolides has been reported to date, the genome of the strain *Amycolatopsis* sp. CA-230715 was fully sequenced (GenBank CP059997.1-CP059998.1)[310] and the putative T1PKS gene cluster was identified. This allowed us to gain insight into the biosynthetic machinery of these complex macrolactones and complete the determination of their absolute configurations, as discussed below.

2.2. Identification of the 216 Kbp type I PKS gene cluster responsible for the biosynthesis of gargantulides

To identify the BGC of gargantulides B (**1**) and C (**2**) in the genome of the producing strain, we used the bacterial version of the antiSMASH software (v. 6.0.0) with the default 'relaxed strictness' settings.[190] A total of 45 biosynthetic regions were predicted, all located on the circular chromosome (Table S2). One of the predicted T1PKS regions (region 11) fits very well with the size and structure of gargantulides. This gene cluster (further referred to as *gar*), has an overall size of 216 kbp and encodes 10 Type I PKS genes (*garP1* to *garP10*) harbouring 35 PKS modules and spanning approximately 180 kbp (Fig. 3).
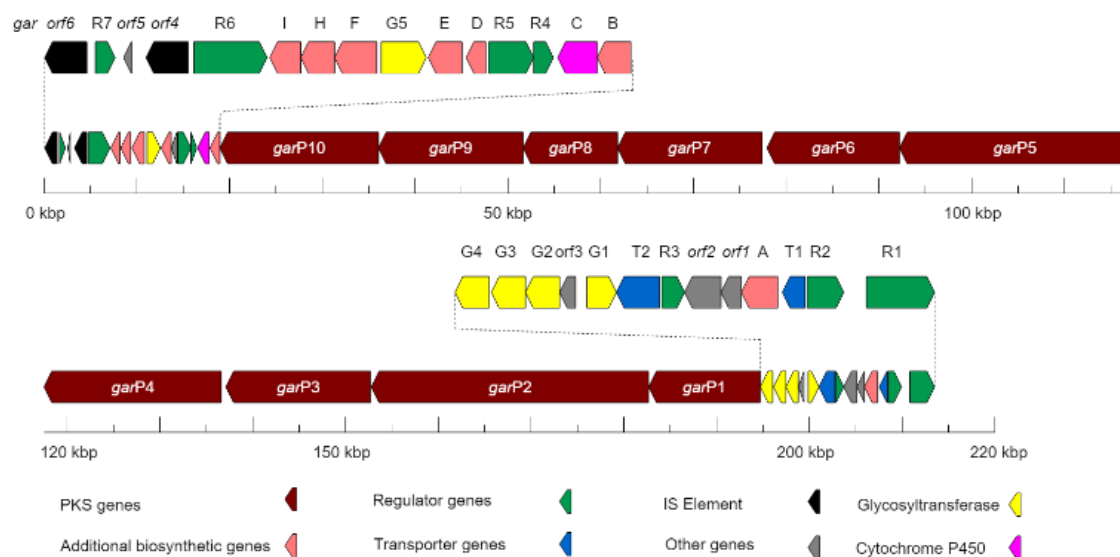
Figure 3. The *gar* biosynthetic gene cluster with putative functions of all genes (see Table S3).

These features place it among the largest uninterrupted BGCs ever discovered and, - to our knowledge -, as the largest T1PKS cluster by the number of encoded modules. The putative functions of all genes in the *gar* cluster were determined by sequence comparisons (Table S3). The genome of *Streptomyces* sp. A42983, the original producer strain of gargantulide A[307] is not publicly available, which prevented the direct comparison. Nevertheless, the structure of the polyketide backbone common to gargantulides A-C is in excellent agreement with the product that can be predicted from the PKS domain organization (Table S4). Based on NMR data, the starter unit of GarP1 is proposed to be 4-aminobutanoyl-CoA, which might be derived from ornithine via putative oxidative deamination and decarboxylation reactions[311,312] and introduced as a starter unit through the function of the putative acyl-CoA ligase GarA.

To predict the substrates of all AT domains as well as the activity and stereochemistry of KR, DH, and ER domains, sequence alignments were performed and key residues involved in substrate selection or stereochemical outcome were extracted and compared with references (Table S4, Fig. S8).[313–316] In the case of AT domains, all but one (module 9), have amino acid signatures indicating malonyl-CoA (modules 2-6, 8, 11, 12, 15-21, 23, 25, 27, 28, 31, 33, 34) or methylmalonyl-CoA (modules 1, 7, 10, 13, 14, 22, 24, 26, 29, 30, 32) specificity, which are common extender units for the biosynthesis of polyketides.[316] AT domain in module 9 is different and is proposed to be responsible for the introduction of the unusual branched extender unit propylmalonyl-CoA (Fig. S8-A), giving rise to the *n*-propyl chain at C-52. Consistent with the latter, the *gar* BGC encodes a putative crotonyl-CoA carboxylase/reductase (GarF) and a putative FabH protein (ketoacyl-ACP synthase III) GarH, which are well known to be involved in the biosynthesis of unusual branched extender units.[316] The cluster contains one inactive KR domain (C1 type) at module 26, which is in agreement with the presence of a keto function instead of a hydroxy group at C-19 in the structures. All other KR domains belong to the A1, B1, and B2 types (Table S4). NMR data established the presence of a single double bond between C-22/C-23 introduced by the DH domain at module 24, thus suggesting that the DH domains at modules 8 and 11 are inactive. Indeed, a close inspection of the DH domain at module 11 revealed the absence of a key catalytic histidine residue. In the case of the DH domain at module 8, although the key residues in the catalytic domain are conserved, several supporting active site residues are missing (Fig. S8-D).[313]

Taking these findings together and assuming the collinearity rule, we can propose a biosynthetic pathway for the common backbone of gargantulides A-C (Fig. 4). After the

181

polyketide chain is released from the PKS complex by TE-domain mediated cyclization, hydroxylation at C-24 and several different glycosylation occur.
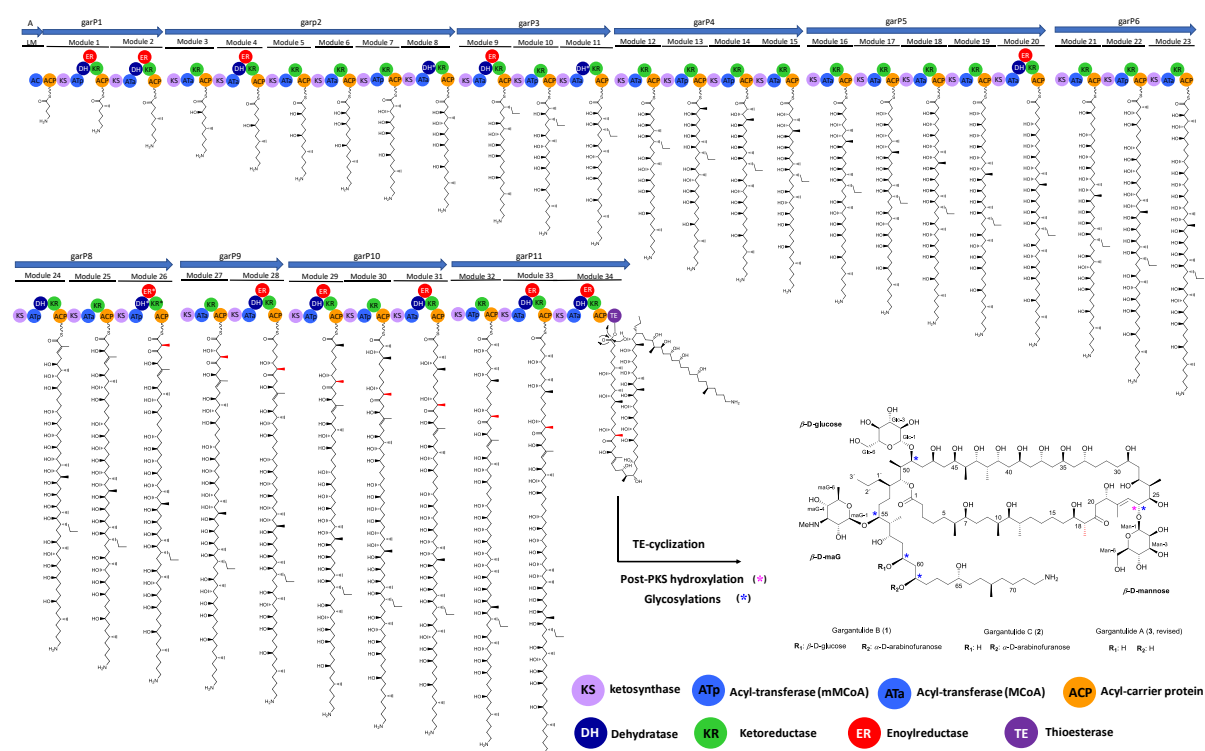
Figure 4. Proposed biosynthesis of the polyketide core of gargantulides A-C. Inactive domains (KR, DH, or ER) are marked with asterisks. Methyl-epimerization at C-18 (redox-inactive KR domains in module 26) is marked in red.

Despite the high degree of glycosylation of gargantulides A-C, only a few genes encoding for proteins putatively involved in sugar biosynthesis are found in the *gar* BGC. Of note, all three macrolides contain the unusual amino sugar 3,6-deoxy-3-methylamino glucose (maG), which, to the best of our knowledge, is present exclusively in this family of compounds. We hypothesized that this amino sugar could share a common biosynthetic origin with mycaminose (3,6-deoxy-3-dimethylamino-D-glucose), an amino sugar found in other glycosylated macrolides (e.g. tylosin), whose biosynthesis has been extensively studied.[208,317,318] Interestingly, a set of genes encoding proteins homologous to those described for mycaminose biosynthesis were also found in the genome of the producer strain. Among them, GarD and GarE within the *gar* cluster (Fig. 3) were identified as homologues of Tyl1a (TDP-4-keto-6-deoxy-D-glucose 3,4-isomerase)[319,320] and TylB (aminotransferase)[321] respectively, from the mycaminose biosynthetic pathway in the tylosin BGC (Supplementary note, Tables S5-S6, and Fig. S9).

Five glycosyltransferases are encoded in the *gar* BGC and assumed to be responsible for the attachment of the different monosaccharide residues to the gargantulide aglycon (Fig. 3, Table S3). GarG1-GarG4 are glycosyltransferases of the GT-1 family, whereas GarG5 belongs to the GT-39 family.[322] Unlike described for the mycaminose attachment to the tylosin aglycon,[323] no glycosyltransferase requiring a P450-like activator is found in the *gar* BGC. Moreover, GarC, the only P450 enzyme encoded in the cluster, contains the conserved cysteine residue within the Cys-pocket found in canonical P450 proteins, which is reported to be missing in the P450-like activators of these glycosyltransferases.[324] Thus, GarC is likely to function as a canonical P450 enzyme catalyzing the hydroxylation at position C-24.

2.3. Correlations between NMR-based and gene cluster analyses. Proposal of absolute stereochemistry of gargantulides A-C

Gargantulides A (**3**), B (**1**), and C (**2**) contain 34 stereocenters in the shared common polyketide backbone, along with 15, 24, or 19 additional chiral centers respectively, from their monosaccharide residues. Despite the complexity of these huge and stereochemically rich polyol macrolides, the relative configurations for most of the chiral carbons within the polyketide chain of gargantulide A were originally established by a combination of NOESY NMR, *J*-based configuration analysis (JBCA)[325] and Kishi´s universal NMR database method.[326,327] Absolute configurations for these centers were also proposed by using the stereochemically defined sugar residues as internal probes, employing the corresponding *nOe* correlations between the monosaccharide units and the aglycon.[307] However, the configurations of eight stereocenters (C-6, C-7, C-10, C-11, C-12, C-17, C-18 and C-68) remained unassigned. The likely common biosynthetic origin of gargantulides A-C (**1-3**) ensures the same absolute configuration for the three congeners. Therefore, to establish or complete the stereochemistry of gargantulides A-C, the bioinformatic-derived assignments for each stereocenter within the polyketide backbone (except that of C-24) were compared and validated with those obtained from NMR data. Unexpectedly, some discrepancies were found between the *in silico* assigned stereochemistry for gargantulides B and C and that determined experimentally for gargantulide A (Table S4, Fig. S10). In terms of relative configuration, the NMR-based determinations for gargantulide A are in excellent agreement with the bioinformatics prediction for all the stereocenters reported, except for C-55.

Strikingly, bioinformatics analysis of the large central segment comprising C-25 to C-52, predicted the same relative stereochemistry but opposite absolute configuration to that reported for gargantulide A. On the other hand, configurations of stereocenters within the segment C-6 to C-18, as well as that of C-68, were not assigned for gargantulide A but they have now been predicted *in silico* (Table S4, Fig. S10). To make a definitive proposal of the absolute stereochemistry for gargantulides A-C, we set out to clarify contradictions between the *in-silico* assignments and the experimental determinations for gargantulide A, as well as to experimentally verify the predicted absolute configurations for those stereocenters not previously determined. For this purpose, we employed a combination of Kishi´s universal NMR database,[326,327] ^{1}H-^{1}H coupling constant, and NOE analyses [328,329] and qualitative $^3J_{C,H}$-based configurational analysis. For the latter, we classified the magnitude of the heteronuclear coupling constants as "large" or "small", based on the intensity of their HMBC cross-peaks, as reported previously.[160,330–332] Focusing on **1**, we first confirmed that the relative stereochemistry of gargantulide B (and thus also that of C) was consistent with the prediction and that it was essentially the same as that partially reported for gargantulide A. Of note, a single comparison of the nearly identical ^{1}H and ^{13}C chemical shift values (except for the differentially substituted positions) as well as the ^{1}H-^{1}H coupling constants (Table S1) for gargantulides A, B, and C, provided a clear indication that we would obtain the same results. Thus, similar to what was reported previously,[307] relative configurations within segments C-27 to C-49 and C-57–C61 were assigned by applying Kishi´s NMR databases corresponding to the 1,3-diol (C-27–C-29), 1,3,5,7-pentol (C-33–C-41), 1,3,5-triol (C-45–C-49 and C-57–C61) and the hydroxy/methyl/hydroxy/methyl (C-41–C-45) sets.[328,333,334] Additionally, the Matsunaga empirical rule for 1,5-diols [335] allowed us to establish *anti* and *syn* configurations for the pairs C-29/C-33 and C-61/C-65, respectively. In parallel, relative configurations within C-24–C-27 and C-49-C-52 segments were also confirmed to be consistent with those reported for gargantulide A based on NOESY, $^3J_{H,H,}$ and $^3J_{C,\ H}$-based configuration analyses.

However, the absolute configuration of the full segment from C-24 to C-52 in gargantulide A was assigned, through *nOe* correlations between key hydrogens of the polyketide backbone and the mannose and glucose units at C-24 and C-49 respectively, as the opposite of that deduced from our BGC analysis.

The absolute configuration of C-55 in gargantulide A was originally defined as *R*, and the C-55/C-56/C-57 substereocluster was assigned a *syn/syn* relative configuration.[307] However, the *in silico* analysis predicted a 55*S* configuration and the *anti/syn* relationship for those centers (Table S4, Fig. S10). From the NMR-based analysis, we found the following evidence supporting this bioinformatics assignment: First, in spite that glycosylation at C-55-OH formally precludes the application of the Kishi´s data set I (Fig. S11-a) due to a possible shielding γ-gauche effect by the glycosyl (maG) moiety, the chemical shift value of Me-56 (δ_C 10.7) is very close to that for the *anti/syn* configuration. In contrast, a *syn/syn* configuration for the fragment C-55/C-56/C-57 would impose a chemical shift value for Me-56 lower than 7.1 ppm (according to Kishi´s data set I and to possible additional shielding by the maG sugar), too far away from the experimental value. Second, NOESY correlations between Me-56 and H-55, H-54 (one of the methylene hydrogens at δ_H 1.89), and H-58 strongly suggested the *anti/syn* configuration for C-55/C-56/C-57 (Fig. S11-b and S11-c). Third, although $^3J_{H,H}$ constants could not be accurately measured for H-55/H-56/H-57, the higher multiplicity of H-55 compared to that of H-57 as well as the appearance of H-56 as a "dqd" with only one large coupling constant ($^3J_{H-55,H-56}$ = ca. 8 Hz; $^3J_{H-56,Me-56}$ = 6.9 Hz; $^3J_{H-56,H-57}$ = ca. 2 Hz), jointly suggested the *anti/syn* configuration as well (Fig. S11-d). Finally, the small estimated values (based on the low intensity of their HMBC cross-peaks) for the coupling constants $^3J_{Me-56–H55}$, $^3J_{C-54–H56}$, $^3J_{C-58–H56}$ and $^3J_{C-55–H57}$ and the large one for $^3J_{Me-56–H57}$, assigned Me-56/H-55, C-54/H-56 and Me-56/57-OH as *gauche* and Me-56/H-57 as *anti* (Fig. S11-b), being consistent with the bioinformatics prediction. Taken together, the 55*S* absolute configuration was proposed.

In view of these results, we might conclude that disagreements between the absolute configurations determined within the segment C24-C55 in gargantulide A (**3**) and those predicted from the gene cluster analysis, may have arisen from misinterpretations of some *nOe* correlations between the sugar units and the polyketide backbone of **3** in the previous work.[307]

We then focused on chiral centers whose stereochemistry had not previously been determined for **3**. The *syn* configuration for C-6/C-7 was deduced from Kishi´s NMR database for the 4-methylnonan-5-ol isomers (Fig. S12),[336] based on the chemical shift of the branched methyl Me-6 at δ_C 14.5 ppm. This configuration assignment matched with that of the bioinformatics assignment.

Regarding the stereocluster C-10–C-12, although ^{1}H-^{1}H coupling constants of H-10 or H-12 could not be measured due to signal overlapping, the appearance of H-11 as double-doublet (dd) with one large (8.2 Hz) and one small (3.2 Hz) $^3J_{H,H}$ coupling constant indicated either a *syn/anti* or *anti/syn* configuration for C10/C-11/C-12. The different chemical shifts for the methyl carbons Me-10 (δ_C 13.4 ppm) and Me-12 (δ_C 16.7 ppm) (Table S1) also referred to one of these two configurations. The greater intensity of the COSY cross-peak for H-11/H-12 compared to that of H-10/H-11 ($^3J_{H10-H11}$ < $^3J_{H11-H12}$) definitely suggested a *syn/anti* configuration (Fig. S13-a), which was further supported by NOESY correlations between H-10/ H-11 and H-9/ H-11 (Fig. S13-c). The large (estimated) value for the heteronuclear coupling constant $^3J_{Me-10–H11}$ and the small one for $^3J_{Me-12–H11}$ were consistent with this configuration assignment (Fig. S13-b and S13-d), which in turn, agreed with the *in-silico* analysis.

Moving to the C-17/C-18 stereocluster, the absolute configuration 17*R*, 18*S* (*syn* relative configuration) was predicted from the gene cluster analysis (Fig. S10, Table S4). However, the NMR analysis revealed a large coupling constant $^3J_{H17-H18}$ (8.6 Hz) (Table S1, Fig. S14). Furthermore, key NOESY correlations Me-18/H-17 and Me-18/H-16 were also consistent with the *anti*-relative configuration (Fig. S14). This result did not match with the stereochemical outcome deduced from the predicted type C1 KR domain at module 26. However, redox-inactive KR domains lacking the conserved active site Tyr residue characteristic of short-chain dehydrogenase/reductase (therefore classified as type C1 KR domains according to Keatinge-Clay guidelines),[314] have recently been shown to catalyze α-methyl epimerization.[337] Moreover, the involvement of the paired DH domain from module 26 in the epimerization of the α-methyl

group by itself or in cooperation with the type C1 KR domain can also be hypothesized.[338] Altogether, the 17*R*, 18*R* absolute configuration was eventually assigned.

Finally, since C-68 is an isolated stereocenter, it was not possible to relate it to any other stereocluster and we assumed the 68*R* configuration deduced from the *in-silico* analysis.

To complete the stereochemical assignments of gargantulides B and C, absolute configurations of the non-amino sugars present in **1** and **2** were determined after hydrolysis and derivatization with L or D-Cys and *o*-tolyl isothiocyanate, as reported previously (Experimental section and Fig. S15).[339] This way, the β-glucose units (two in **1**, only one in **2**) as well as the β-mannose and α-arabinose residues were all determined as D sugars, thus matching the previous determinations for gargantulide A. Regarding the rare amino sugar maG, although its absolute configuration in gargantulides B and C could not be determined experimentally, we propose a D configuration assuming a common biosynthetic origin with mycaminose, as discussed above. The absolute configuration of maG was originally proposed as L in gargantulide A based on NOESY correlations between the monosaccharide and the aglycon.[307] However, it should be noted that maG is *O*-glycosidically bound to C-55 in all three macrolides and that this stereocenter has been assigned in this work as the opposite absolute configuration to that originally reported for gargantulide A. Therefore, it is reasonable to think that a possible original misassignment of the C-55 stereochemistry in gargantulide A may have led to the previous proposal of maG as an L sugar.

2.4. Bioactivity of gargantulides B and C

Compounds **1** and **2** were evaluated for their antimicrobial properties against a panel of different pathogens, including Gram-positive and Gram-negative bacteria and fungi. Thus, gargantulides B (**1**) and C (**2**) both showed potent activity against the Gram-positive bacteria methicillin-resistant and susceptible *Staphylococcus aureus* (MRSA, MSSA) and vancomycin-resistant *Enterococcus* (VRE), with MIC values in the range from 1 to 8 µg/mL (Table S7). In contrast, but similar to gargantulide A, **1** and **2** showed low activity against the fungus *Aspergillus fumigatus* ATCC46645 (MIC 32-64 µg/mL) and were inactive against *Candida albicans* ATCC64124 (MIC > 128 µg/mL). Remarkably and consistent with the bioassay-guided isolation of **1** and **2**, both compounds showed moderate and selective activity against a clinical isolate of *Acinetobacter baumannii* (MIC values of 16-32 µg/mL), but no activity against any other Gram-negative bacteria tested (Table S7). To our knowledge, the activity of large macrolactones against Gram-negative bacteria such as *A. baumannii* has been rarely reported, indicating that gargantulide macrolides may represent and alternative in the fight against the increasing spread of this clinically relevant pathogen.

3. Conclusions

Two new glycosylated 52-membered macrolactones with 22-membered linear side chains, gargantulides B (**1**) and C (**2**), were isolated and structurally characterized from the strain *Amycolatopsis* sp. CA-230715. The extraordinarily large 216 kbp biosynthetic gene cluster (*gar*) encoding a 35 modules T1PKS was identified and analyzed, allowing us to establish the first genetic evidence for the biosynthesis of this interesting class of compounds. Additionally, the *gar* gene cluster includes a set of genes putatively responsible for the biosynthesis of the rare sugar 3,6-deoxy-3-methyl amino D-glucose (maG), which, as far as we know, is present exclusively in this macrolide family. The combination of NMR spectroscopy and bioinformatics analysis of the gene cluster allowed the assignments of the absolute configuration for all chiral centers in gargantulides B (**1**) and C (**2**). Based on the high structural and spectroscopic similarities with the previously reported congener gargantulide A (**3**), we propose a common biosynthetic origin and the same stereochemistry for all gargantulide macrolides. This allowed us to revise the stereochemical assignments originally proposed for gargantulide A, including that of the amino sugar maG. The new macrolactones **1** and **2** showed potent antibacterial activity against Gram-positive bacteria including MRSA and VRE, and moderate but unusual

activity against the Gram-negative pathogen *A. baumannii*. The results presented in this work illustrate the powerful and reliable combination of bioactivity screenings along with NMR and bioinformatic analyses in the discovery and structural elucidation of bioactive natural products.

4. Experimental Section

4.1. General experimental procedures

Optical rotations were measured on a Jasco P-2000 polarimeter. IR spectra were recorded with a JASCO FT/IR-4100 spectrometer equipped with a PIKE MIRacle single reflection ATR accessory. NMR spectra were recorded on a Bruker Avance III spectrometer (500 and 125 MHz for ^{1}H and ^{13}C NMR, respectively) equipped with a 1.7 mm TCI MicroCryoProbe. Chemical shifts were reported in ppm using the signals of the residual solvents as internal reference (δH 3.31 and δC 49.1 for CD$_3$OD).

LC-UV-LRMS analysis was performed on an Agilent 1100 single quadrupole LC-MS system as previously described.[261] ESI-TOF and MS/MS spectra were acquired using a Bruker maXis QTOF mass spectrometer coupled to an Agilent Rapid Resolution 1200 LC. The mass spectrometer was operated in positive ESI mode. The instrumental parameters were 4 kV capillary voltage, drying gas flow of 11 L/min at 200 °C, and nebulizer pressure of 2.8 bar. TFA-Na cluster ions were used for mass calibration of the instrument before sample injection. Pre-run calibration was done by infusion with the same TFA-Na calibrant. Semipreparative HPLC separation was performed on a Gilson GX-281 322H2 with a semi-preparative reversed-phase column (Waters XBridge Phenyl, 10 x 150 mm, 5 μm). Acetone used for extraction was analytical grade. The solvents employed for isolation were all HPLC grade. Chemical reagents and standards were purchased from Sigma-Aldrich.

4.2. Strain isolation, identification, and fermentation

The strain CA-230715 was isolated from a soil sample collected in a waterlogged forest in the Central African Republic. Initial similarity-based search with the 16S rDNA sequence (1318nt) against the EzBioCloud database indicated that the strain is closely related to Amycolatopsis minnesotensis 32U-2(T) (98.71% identity, 91.6% completeness).

A 250 mL fermentation of the producing microorganism was obtained as follows: a seed culture of the strain was obtained by inoculating one 150 x 25 mm tube containing 14 mL of seed-medium (soluble starch 20 g/L, glucose 10 g/L, NZ Amine Type E 5 g/L, meat extract 3 g/L, peptone 5 g/L, yeast extract 5 g/L, calcium carbonate 1 g/L, pH 7) with 0.7 mL of freshly thawed inoculum stock of CA-230715. The tube was incubated at 28 ºC, 70% relative humidity, and 220 rpm for 5 days. The fresh inoculum thus generated was employed to inoculate two 250 mL conical flasks each containing 125 mL of APM9 medium (glucose 50 g/L, soluble starch 12 g/L, soy flour 30 g/L, CoCl2·6H2O 2 mg/L, calcium carbonate 7 g/L, pH 7). The flasks were incubated at 28 °C, 70% relative humidity, and 220 rpm for 13 days before harvesting.

4.3. Isolation of gargantulides B (1) and C (2)

The 250 mL culture broth was extracted with an equal volume of acetone under continuous shaking at 220 rpm for 1h. The mycelial debris was discarded by centrifugation at 9000 rpm and the filtered supernatant/acetone mixture (ca. 0.5 L) was concentrated to 0.25 L under a nitrogen stream. The aqueous crude extract was loaded onto a Sepabeads® SP207ssresin column and eluted using acetone/methanol (1:1). The resulting organic extract was chromatographed by semipreparative reverse-phase HPLC (Waters XBridge Phenyl, 10 x 150 mm, 5 μm; 3.6 mL/min, UV detection at 210 and 280 nm) with a linear gradient of CH$_3$CN/H$_2$O/0.1% trifluoroacetic acid (TFA), from 15 to 28% CH$_3$CN (0.1% TFA) over 15 min

followed by an isocratic step of 28% CH$_3$CN (0.1% TFA) over 19 min, to yield gargantulides B (1, 2.1 mg, rt 22 min) and C (2, 1.7 mg, rt 24.5min) as white amorphous powders.

4.4. Characterization data

Gargantulide B (1): [α]$^{25}_D$ -28.3 (*c* 0.33, MeOH); IR (ATR) cm^{-1}: 3343, 2934, 1674, 1456, 1428, 1379, 1201, 1134, 1068, 1023; (+)-ESI-TOF-MS *m/z* 2392.4804 [M+H]$^+$ (calcd. for C$_{116}$H$_{219}$N$_2$O$_{47}$$^+$, 2392.4808), 1197.2455 [M+2H]$^{2+}$ (calcd. for C$_{116}$H$_{220}$N$_2$O$_{47}$$^{2+}$, 1197.2455), 798.4995 [M+3H]$^{3+}$ (calcd. for C$_{116}$H$_{221}$N$_2$O$_{47}$$^{3+}$, 798.4994); ^{1}H and ^{13}C NMR data see Table S1.

Gargantulide C (2): [α]$^{25}_D$ -20.5 (*c* 0.33, MeOH); IR (ATR) cm^{-1}: 3342, 2935, 1674, 1456, 1425, 1379, 1201, 1135, 1067, 1024; (+)-ESI-TOF-MS *m/z* 2230.4286 [M+H]$^+$ (calcd. for C$_{110}$H$_{209}$N$_2$O$_{42}$$^+$, 2230.4280), 1116.2170 [M+2H]$^{2+}$ (calcd. for C$_{110}$H$_{210}$N$_2$O$_{42}$$^{2+}$, 1116.2191), 744.4808 [M+3H]$^{3+}$ (calcd. for C$_{110}$H$_{211}$N$_2$O$_{42}$$^{3+}$, 744.4818); ^{1}H and ^{13}C NMR data see Table S1.

4.5. Determination of the absolute configurations of the non-amino sugars present in gargantulides B and C

Compounds 1 and 2 (400 µg each) were separately dissolved in 0.1 mL of 1 N HCl and heated at 85 °C for 3.5 h in a sealed vial. The crude hydrolysates were evaporated to dryness under a nitrogen stream and each residue was dissolved in 100 µL of pyridine containing 0.5 mg of L-cysteine methyl ester hydrochloride and heated at 60 °C for 1 h. A 100 µL solution of *O*-tolyl isothiocyanate (0.5 mg) in pyridine was added to the mixture, which was heated at 60 °C for 1 h. Standard monosaccharides (D-glucose, D-mannose, and D-arabinose) were treated in the same manner with both L- and D- forms of cysteine methyl ester hydrochloride separately.

The reaction mixtures (20 µL) were diluted with 50 µL of methanol and analyzed by ESI LC/MS on an Agilent 1100 single Quadrupole LC/MS. Separations were carried out on a Waters XBridge C18 column (4.6 × 150 mm, 5µm), maintained at 40 °C. A mixture of two solvents, A (10% CH$_3$CN, 90% H$_2$O) and B (90% CH$_3$CN, 10% H$_2$O), both containing 1.3 mM trifluoroacetic acid and 1.3 mM ammonium formate, was used as the mobile phase under a linear gradient elution mode (isocratic 15% B for 25 min, 15-100% B in 0.1 min and then isocratic 100% B for 5 min) at a flow rate of 1.0 mL/min.

Retention times (min) for the derivatized standard monosaccharides (with both L- and D- forms of cysteine methyl ester hydrochloride) present in 1 and 2 were: D-glucose + L-Cys/o-TolylNCS: 17.05, D-glucose + D-Cys/o-TolylNCS: 15.70, D-mannose + L-Cys/o-TolylNCS: 10.14, D-mannose + D-Cys/o-TolylNCS: 17.22, D-arabinose + L-Cys/o-TolylNCS: 20.41, D-arabinose + D-Cys/o-TolylNCS: 18.99. Retention times (min) for the observed peaks in the HPLC trace of the L-Cys/o-TolylNCS derivatized hydrolysis product of 1 were: D-glucose + L-Cys/o-TolylNCS: 17.11, D-mannose + LCys/o-TolylNCS: 10.10, D-arabinose + L-Cys/o-TolylNCS: 20.48. Retention times (min) for the observed peaks in the HPLC trace of the L-Cys/o-TolylNCS derivatized hydrolysis product of 2 were: D-glucose + L-Cys/o-TolylNCS: 17.10, D-mannose + L-Cys/o-TolylNCS: 10.10, D-arabinose + L-Cys/o-TolylNCS: 20.56.

4.6. Genome sequence alignments

The sequences of AT, KR, DH, and ER domains were manually extracted from the genome sequence of CA-230715 and analyzed using Geneious 9.1.8 software platform.[340] The catalytic regions and regions valuable for the prediction of stereochemistry were identified as described in previous works.[313–316] Multiple sequence alignments were run using MUSCLE Alignment function[341] of Geneious, set to 8 maximum iterations. As a comparison, domains manually extracted from the erythromycin cluster of *S. erythraea* NRRL2338 were used.

4.7. Antimicrobial bioassays

Compounds 1 and 2 were tested in antimicrobial assays against the growth of the gram-negative bacteria *E. coli* ATCC 25922, *A. baumannii* MB5973, *K. pneumoniae* ATCC 700603 and *P. aeruginosa* MB5919 and against the gram-positive bacteria methicillin-resistant *S. aureus* (MRSA) MB5393, methicillin-susceptible *S. aureus* (MSSA) and Van-A vancomycin-resistant *Enterococcus* (VRE) following previously described methodologies.[273–275] Compounds 1 and 2 were also evaluated for their antifungal activity and spectrum against the filamentous fungus *Aspergillus fumigatus* ATCC46645 following previously described methodologies[274] and the yeast *Candida albicans* ATCC64124 as follows: a thawed stock inoculum suspension from cryovials was streaked onto Sabouraud dextrose agar (SDA) plates and incubated at 37 °C overnight to obtain isolated colonies. The single colonies were inoculated into 20 mL of RPMI 1640 medium (Sigma) in 250 mL Erlenmeyer flasks and were adjusted at 0.25-0.28 ODU (612 nm) and then diluted 1:10 to obtain the inoculum for assay. The variation of the growth was quantified by measuring absorbance (OD612nm) using the plate reader EnVision® Multilabel (Perkin Elmer) with two readings, one at the initial time and the other after the plate was incubated at 37 °C for 20 hours.

Each compound was serially diluted in DMSO with a dilution factor of 2 to provide 10 concentrations starting at 128 µg/mL for all the assays. The MIC was defined as the lowest concentration of compound that inhibited ≥ 95% of the growth of a microorganism after overnight incubation. The Genedata Screener software (Genedata, Inc., Basel, Switzerland) was used to process and analyze the data and also to calculate the RZ' factor, which predicts the robustness of an assay.[276] In all experiments performed in this work the RZ' factor obtained was between 0.76 and 0.93.

5. Acknowledgments

Note on structures

All structures were drawn using the program ChemDraw (v.19.1) and then inserted into the text.

6. Supplementary Information

The following Supplementary Information is available online at:
https://www.rsc.org/suppdata/d1/qo/d1qo01480c/d1qo01480c1.

List of Figures, Supplementary Materials

List of Tables, Supplementary Materials

Figure S1. Bioassay-guided isolation of gargantulides B (**1**) and C (**2**). Semipreparative HPLC-UV chromatogram (blue trace: 210 nm; (orange trace: 280 nm).

Figure S2. UV (DAD) and ESI(+)-TOF spectrum of gargantulide B (1):
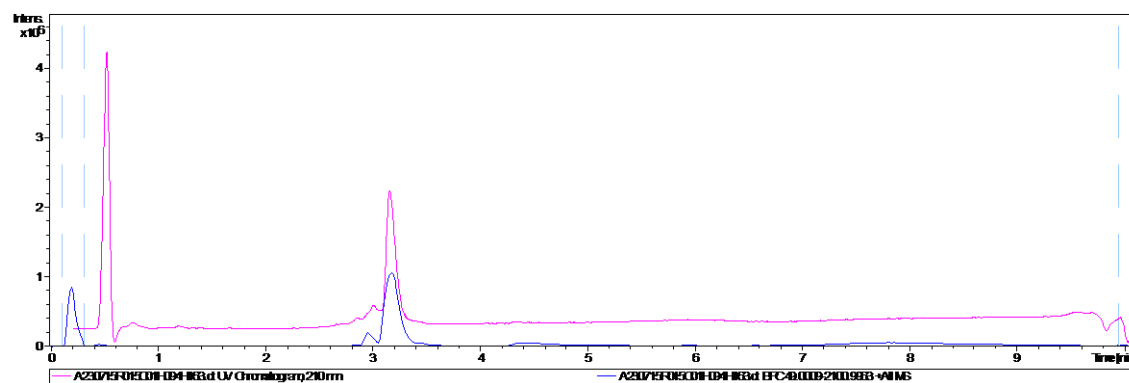
a) LC-UV-HRMS chromatogram (UV 210 nm: pink trace; MS+: blue trace) for **1.**
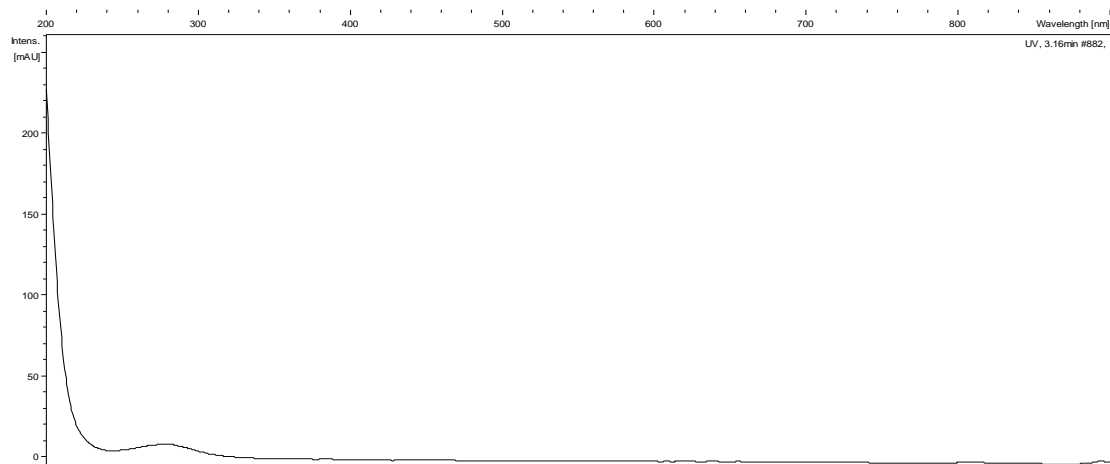
b) UV spectrum of 1.
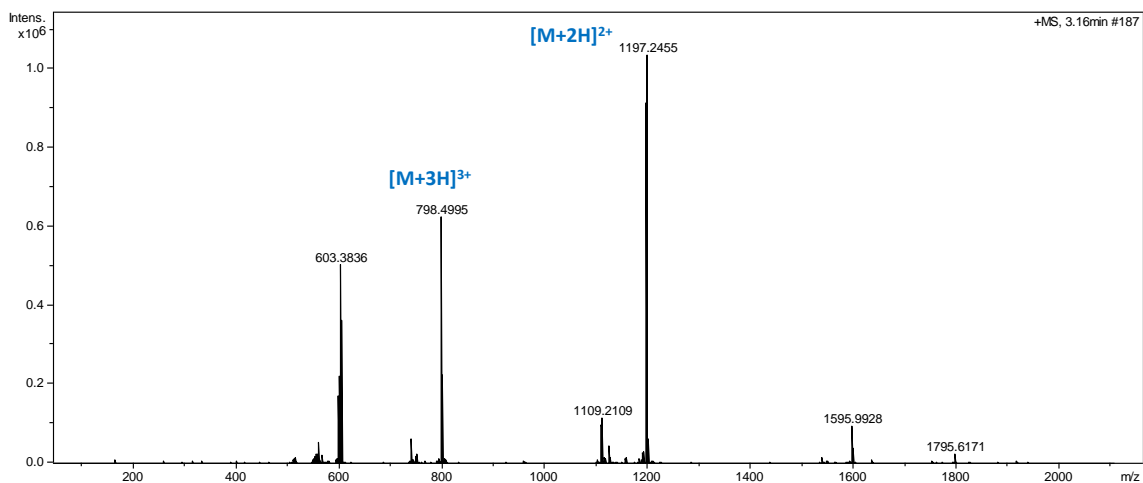
c) HRESIMS(+)-TOF spectra (ISCID 0 eV) of **1** (overview).
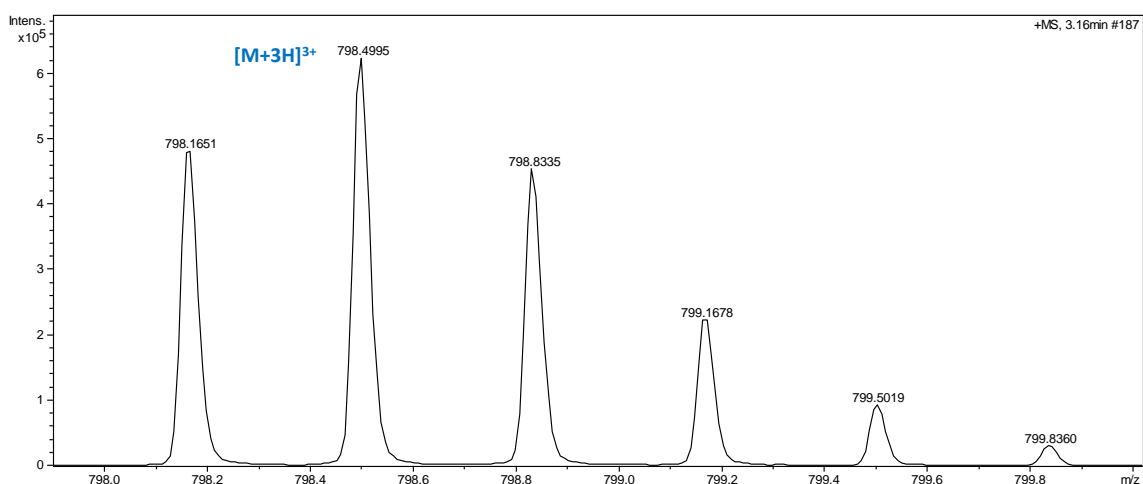
d) HRESIMS(+)-TOF spectra (ISCID 0 eV) of **1**. Triply-charged adducts (zoom in region).
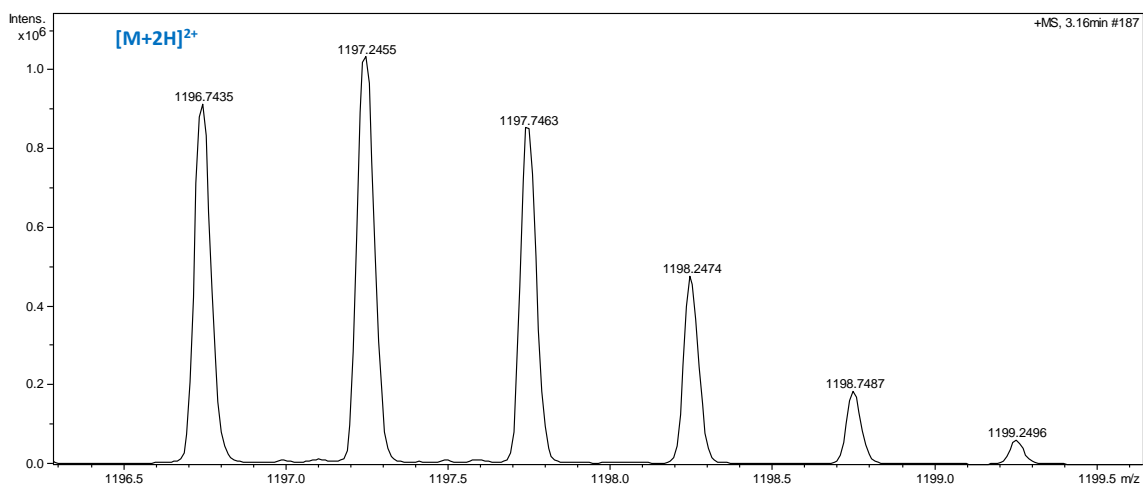
e) HRESIMS(+)-TOF spectra (ISCID 0 eV) of **1**. Doubly-charged adducts (zoom in region).
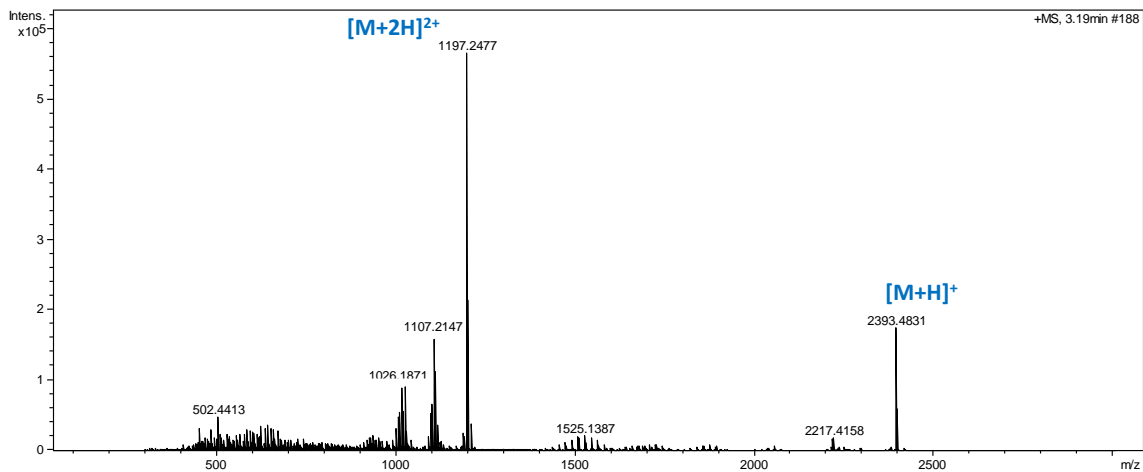
f) . HRESIMS(+)-TOF spectra (ISCID 75 eV) of **1** (overview).
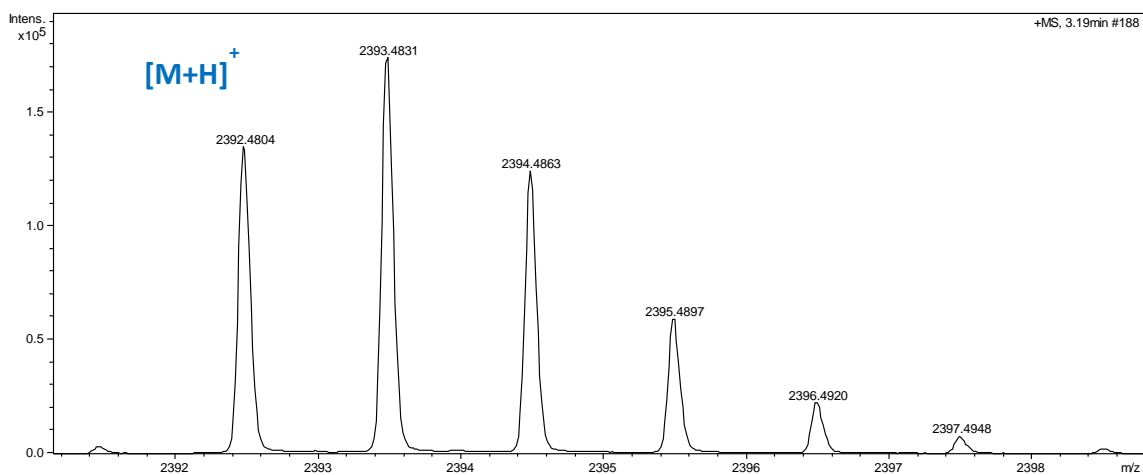
g) HRESIMS(+)-TOF spectra (ISCID 75 eV) of **1**. [M+H]⁺ adduct (zoom in region).
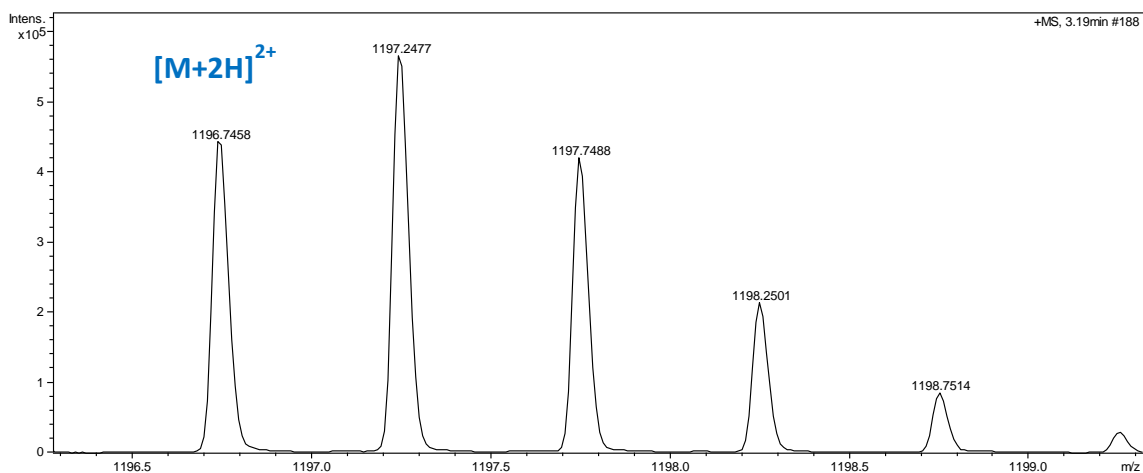
h) HRESIMS(+)-TOF spectra (ISCID 75 eV) of **1**. [M+2H]²⁺ adduct (zoom in region).

Figure S3. UV (DAD) and ESI(+)-TOF spectrum of gargantulide C (**2**):
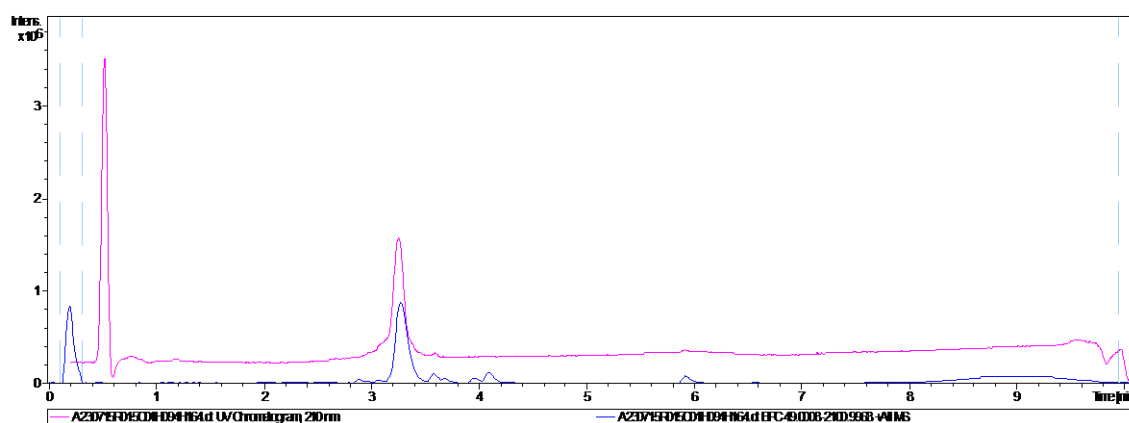
a) LC-UV-HRMS chromatogram (UV 210 nm: pink trace; MS+: blue trace) for **2.**
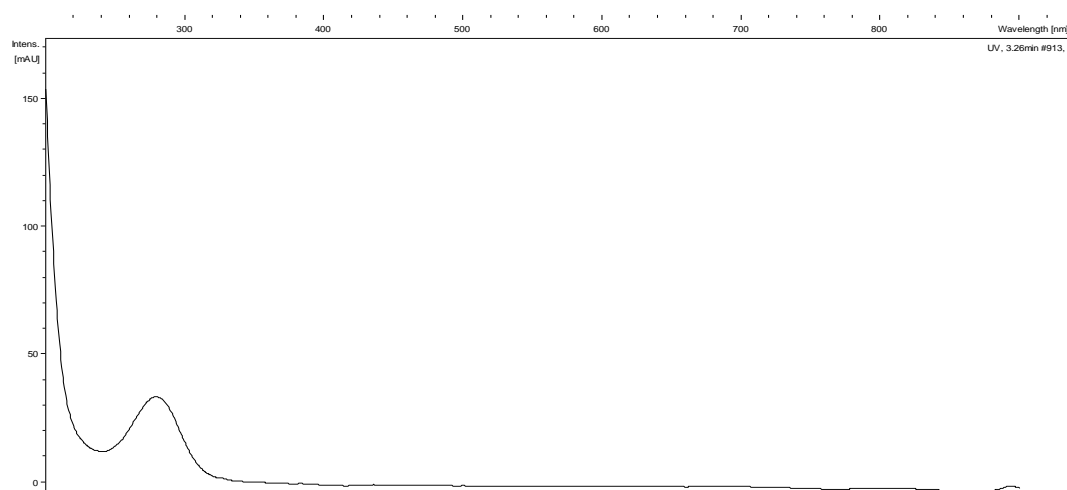
b) UV spectrum of **2.**
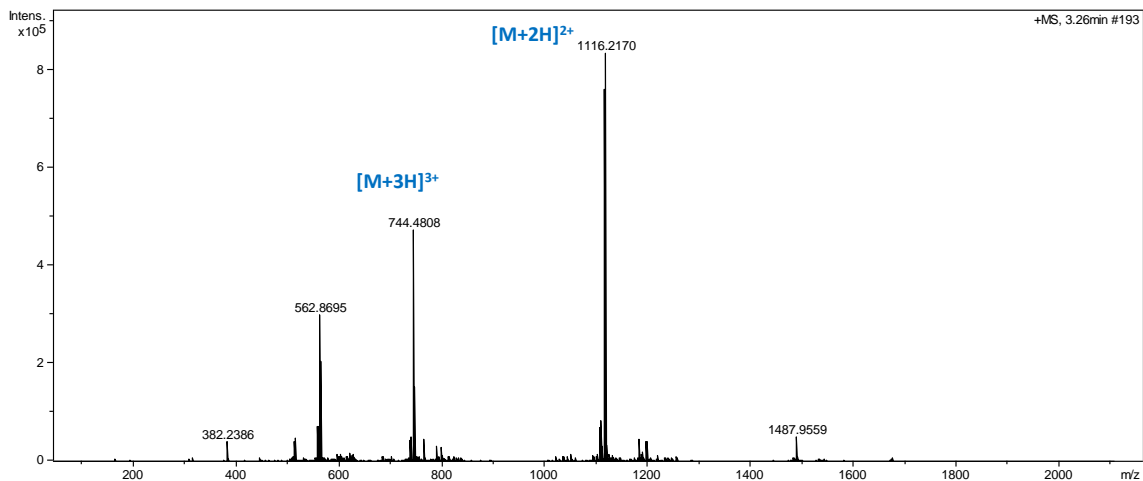
c) HRESIMS(+)-TOF spectra (ISCID 0 eV) of **2** (overview).
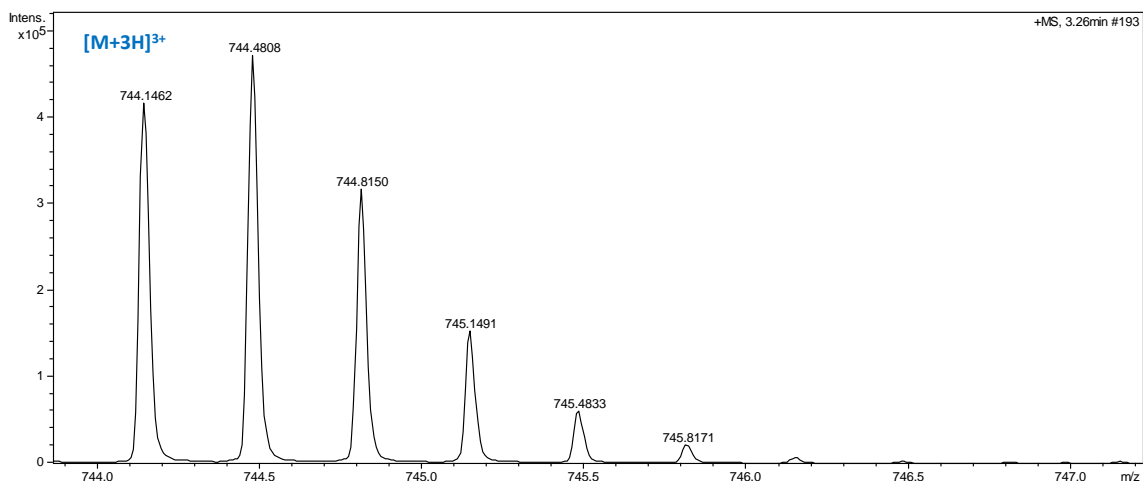
d) HRESIMS(+)-TOF spectra (ISCID 0 eV) of **1**. Triply-charged adducts (zoom in region).
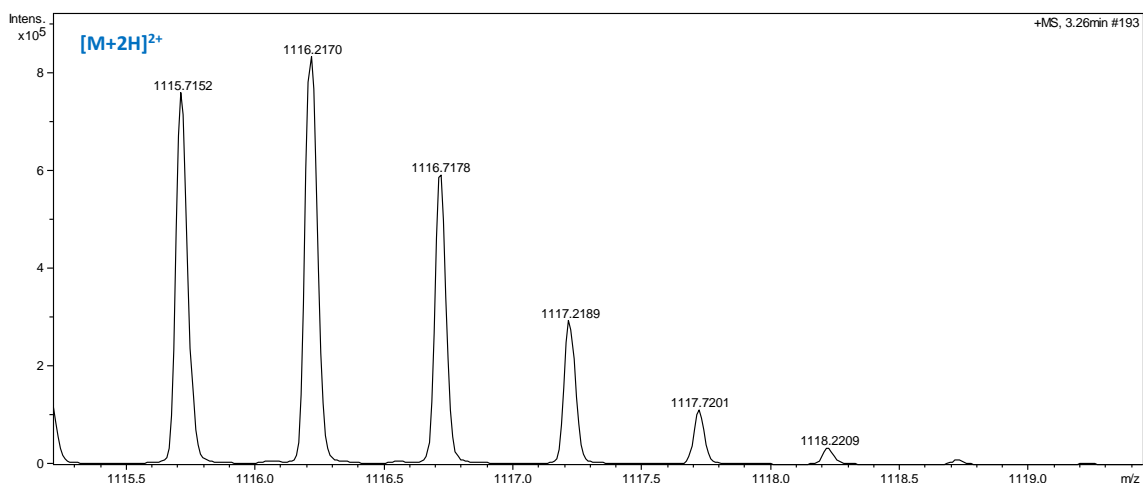
e) HRESIMS(+)-TOF spectra (ISCID 0 eV) of **1**. Doubly-charged adducts (zoom in region).
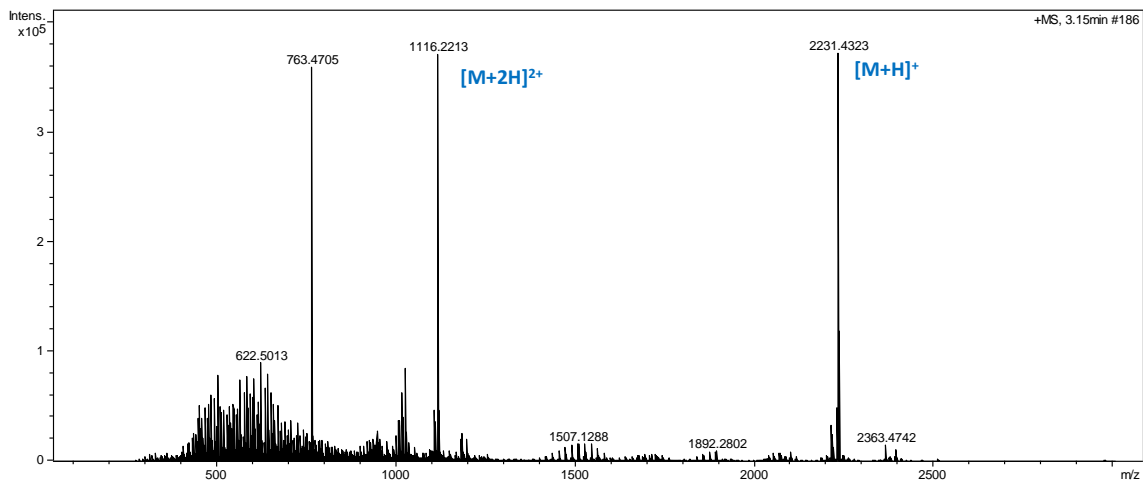
f) HRESIMS(+)-TOF spectra (ISCID 75 eV) of **2** (overview).
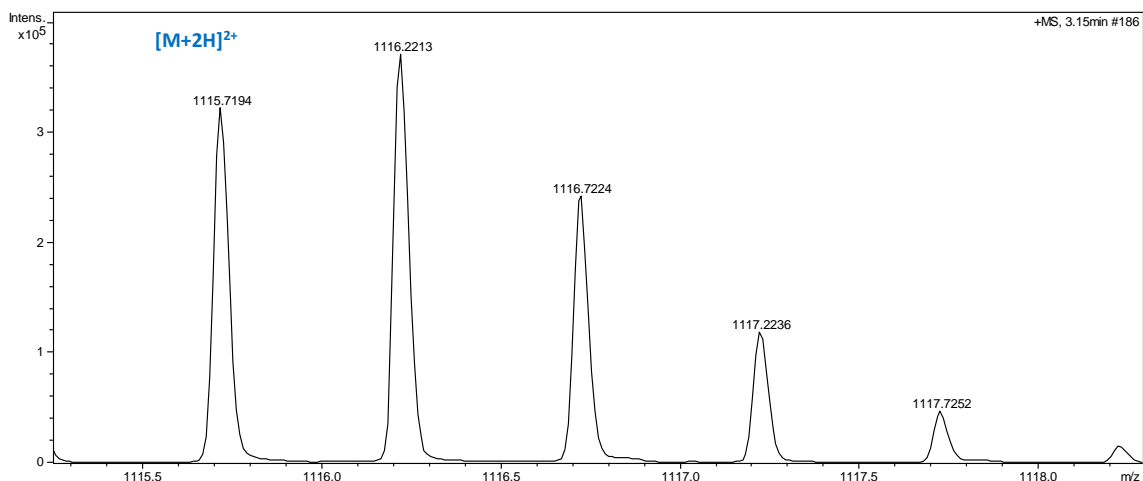
g) HRESIMS(+)-TOF spectra (ISCID 75 eV) of **2**. [M+H]$^+$ adduct (zoom in region).
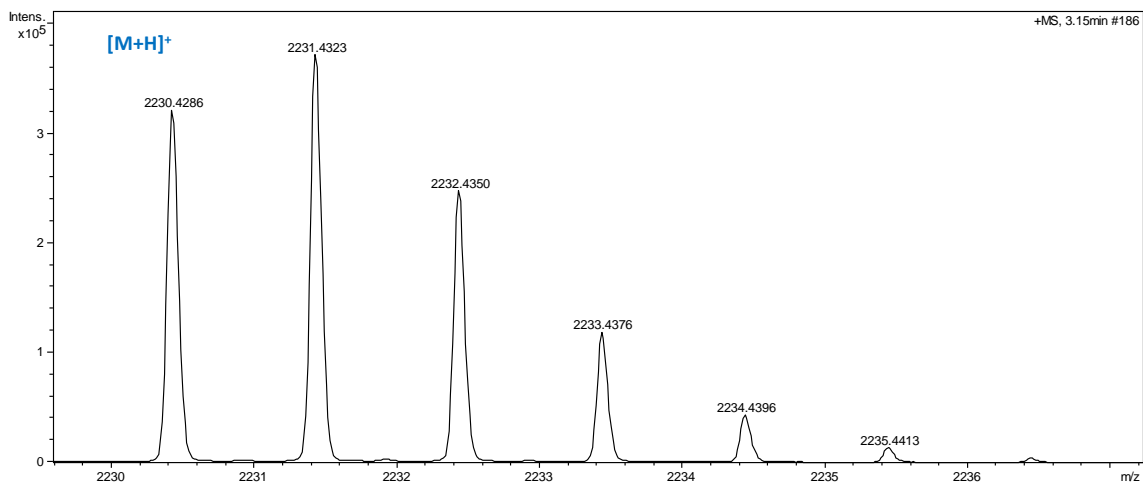
h) HRESIMS(+)-TOF spectra (ISCID 75 eV) of **2**. [M+2H]$^{2+}$ adduct (zoom in region).

Table S1. NMR spectroscopic data (CD$_3$OD) of **1** and **2**.

	Gargantulide B (**1**)			Gargantulide C (**2**)	
nº	δ ^{13}C	δ ^{1}H (mult, *J*, Hz)	nº	δ ^{13}C	δ ^{1}H (mult, *J*, Hz)
1	175.2, C		1	175.2, C	
2	35.5, CH$_2$	2.36, *m*	2	35.6, CH$_2$	2.36, *m*
3	26.9, CH$_2$	1.61 [a], *m*	3	26.9, CH$_2$	1.62 [a], *m*
4	28.5 [a], CH$_2$	1.32 [a], *m*	4	28.5 [a], CH$_2$	1.33 [a], *m*
		1.43 [a], *m*			1.43 [a], *m*
5	34.4 [b], CH$_2$	1.19, *m*	5	34.4 [a], CH$_2$	1.19, *m*
		1.50 [a], *m*			1.50 [a], *m*
6	39.9, CH	1.48 [a], *m*	6	39.9, CH	1.48 [a], *m*
Me-6	14.6, CH$_3$	0.88 [a], *d* (6.7)	Me-6	14.6, CH$_3$	0.89 [a], *d* (6.7)
7	76.2, CH	3.42, *m*	7	76.2, CH	3.42, *m*
8	32.9, CH$_2$	1.46 [a], *m*	8	32.9, CH$_2$	1.47 [a], *m*
9	32.5, CH$_2$	1.38 [a], *m*	9	32.5, CH$_2$	1.38 [a], *m*
		1.46 [a], *m*			1.47 [a], *m*
10	36.2, CH	1.65 [a], *m*	10	36.2, CH	1.65 [a], *m*
Me-10	13.4, CH$_3$	0.86 [a], *d* (6.7)	Me-10	13.3, CH$_3$	0.86 [a], *d* (6.7)
11	79.9, CH	3.12, *dd* (8.2, 3.2)	11	79.9, CH	3.12, *dd* (8.3, 3.2)
12	37.6, CH	1.55 [a], *m*	12	37.6, CH	1.55, *m*
Me-12	16.8, CH$_3$	0.86 [a], *d* (6.7)	Me-12	16.8, CH$_3$	0.86 [a], *d* (6.8)
13	34.0, CH$_2$	1.10 [a], *m*	13	34.1, CH$_2$	1.10 [a], *m*
		1.75, *m*			1.75, *m*
14	28.5 [a], CH$_2$	1.48 [a], *m*	14	28.5 [a], CH$_2$	1.48 [a], *m*
		1.22, *m*			1.22, *m*
15	27.2, CH$_2$	1.33 [a], *m*	15	27.1, CH$_2$	1.32 [a], *m*
		1.57 [a], *m*			1.57 [a], *m*
16	35.6, CH$_2$	1.35 [a], *m*	16	35.9, CH$_2$	1.35 [a], *m*
		1.57 [a], *m*			1.57 [a], *m*
17	74.8, CH	3.67 [a], *m*	17	74.8, CH	3.60 [a], *m*
18	54.1, CH	2.69, *dq* (8.6, 6.9)	18	53.9, CH	2.68, *dq* (8.6, 7.0)
Me-18	13.5, CH$_3$	1.00, *d* (6.9)	Me-18	13.7, CH$_3$	0.99, *d* (6.8)
19	214.6, C		19	214.6, C	
20	49.9, CH$_2$	2.72, *dd* (17.2, 5.1)	20	49.9, CH$_2$	2.80, *dd* (17.6, 6.9)
		2.89, *dd* (17.2, 7.8)			2.87, *dd* (17.6, 6.8)
21	74.1, CH	4.49, *dd* (7.8, 5.1)	21	74.1, CH	4.47, *dd* (6.9, 6.8)
22	146.8, C		22	146.7, C	
Me-22	12.6, CH$_3$	1.81 [a], *s*	Me-22	11.9, CH$_3$	1.78 [a], *s*
23	123.2, CH	5.21, *d* (10.3)	23	123.5, CH	5.22, *d* (10.3)
24	76.1, CH	4.76, *dd* (10.3, 8.1)	24	75.8, CH	4.81, *dd* (10.3, 8.0)
25	80.6, CH	3.56, *dd* (8.1, 3.5)	25	80.4, CH	3.64, *dd* (8.0, 2.8)
26	39.3, CH	1.69 [a], *m*	26	39.4, CH	1.74 [a], *m*
Me-26	11.9, CH$_3$	1.06, *d* (7.0)	Me-26	11.8, CH$_3$	1.06, *d* (7.2)
27	68.6, CH	4.22 [a], *m*	27	68.8 [a], CH	4.20 [a], *m*
28	43.9, CH$_2$	1.27, *m*	28	43.9, CH$_2$	1.27, *m*
		1.59 [a], *m*			1.59 [a], *m*
29	68.6 [a], CH	3.64 [a], *m*	29	68.8 [a], CH	3.61 [a], *m*
30	39.2 [a], CH$_2$	1.40 [a], *m*	30	39.2 [a], CH$_2$	1.41 [a], *m*
		1.44 [a], *m*			1.44 [a], *m*
31	23.1, CH$_2$	1.53 [a], *m*	31	23.2, CH$_2$	1.51 [a], *m*
32	39.2 [a], CH$_2$	1.48 [a], *m*	32	39.2 [a], CH$_2$	1.48 [a], *m*
		1.52 [a], *m*			1.52 [a], *m*
33	69.4, CH	3.84 [a], *m*	33	69.4, CH	3.84 [a], *m*
34	46.0 [a], CH$_2$	1.51 [a], *m*	34	46.3 [a], CH$_2$	1.51 [a], *m*
		1.59 [a], *m*			1.59 [a], *m*
35	66.6, CH	4.10 [a], *m*	35	66.6, CH	4.10 [a], *m*
36	46.9 [a], CH$_2$	1.57 [a], *m*	36	46.9 [a], CH$_2$	1.58 [a], *m*

Gargantulide B (**1**)			Gargantulide C (**2**)		
nº	δ ^{13}C	δ ^{1}H (mult, *J*, Hz)	nº	δ ^{13}C	δ ^{1}H (mult, *J*, Hz)
37	66.6 [a], CH	4.10 [a], *m*	37	66.6, CH	4.10 [a], *m*
38	46.9 [a], CH$_2$	1.58 [a], *m*	38	46.9 [a], CH$_2$	1.58 [a], *m*
39	66.6 [a], CH	4.08 [a], *m*	39	66.6 [a], CH	4.08 [a], *m*
40	44.1, CH$_2$	1.57 [a], *m*	40	44.2, CH$_2$	1.57 [a], *m*
		1.61 [a], *m*			1.61 [a], *m*
41	73.8, CH	4.04, *m*	41	73.8, CH	4.03, *m*
42	41.1, CH	1.63 [a], *m*	42	40.9, CH	1.63 [a], *m*
Me-42	6.9, CH$_3$	0.95 [a], *d* (6.8)	Me-42	6.8, CH$_3$	0.95 [a], *d* (6.9)
43	77.5, CH	3.80 [a], *m*	43	77.5, CH	3.80 [a], *m*
44	41.9, CH	1.67 [a], *m*	44	41.9, CH	1.67 [a], *m*
Me-44	10.8, CH$_3$	0.84, *d* (6.9)	Me-44	10.8, CH$_3$	0.84, *d* (6.8)
45	72.4, CH	4.23 [a], *m*	45	72.4, CH	4.22 [a], *m*
46	40.8, CH$_2$	1.58 [a], *m*	46	40.7, CH$_2$	1.57 [a], *m*
47	71.6, CH	4.06, *m*	47	71.6, CH	4.05, *m*
48	38.2, CH$_2$	1.58 [a], *m*	48	38.3, CH$_2$	1.57 [a], *m*
		1.67 [a], *m*			1.66 [a], *m*
49	77.1, CH	3.69 [a], *m*	49	77.1, CH	3.69 [y], *m*
50	38.5, CH	2.30, *m*	50	38.5, CH	2.30, *m*
Me-50	10.8, CH$_3$	0.93 [a], *d* (6.9)	Me-50	10.7, CH$_3$	0.93 [a], *d* (6.9)
51	76.9, CH	4.88, *m*	51	77.0, CH	4.88, *m*
52	41.1, CH	1.52 [a], *m*	52	41.1, CH	1.52 [a], *m*
53	24.7, CH$_2$	1.20 [a], *m*	53	24.6, CH$_2$	1.20 [a], *m*
		1.78 [a], *m*			1.78 [a], *m*
54	31.9, CH$_2$	1.44 [a], *m*	54	31.8, CH$_2$	1.45 [a], *m*
		1.89, *m*			1.89, *m*
55	83.9, CH	3.68 [a], *m*	55	83.8, CH	3.67 [a], *m*
56	43.4, CH	1.73, *m*	56	43.3, CH	1.73, *m*
Me-56	10.7, CH$_3$	0.94 [a], *d* (6.9)	Me-56	10.8, CH$_3$	0.94 [a], *d* (7.0)
57	68.8, CH	4.19 [a], *m*	57	68.8, CH	4.19 [a], *m*
58	44.2, CH$_2$	1.39 [a], *m*	58	44.6, CH$_2$	1.41 [a], *m*
		1.67 [a], *m*			1.67 [a], *m*
59	77.0 [a], CH	3.84 [a], *m*	59	68.8, CH	3.61, *m*
60	46.0 [a], CH$_2$	1.58 [a], *m*	60	46.0 [a], CH$_2$	1.58 [a], *m*
		1.63 [a], *m*			1.63 [a], *m*
61	77.0 [a], CH	3.84 [a], *m*	61	76.9, CH	3.83, *m*
62	38.7, CH$_2$	1.53, *m*	62	38.7, CH$_2$	1.53, *m*
		1.41 [a], *m*			1.41 [a], *m*
63	22.6, CH$_2$	1.41 [a], *m*	63	22.5, CH$_2$	1.41 [a], *m*
		1.61 [a], *m*			1.61 [a], *m*
64	38.6, CH$_2$	1.40 [a], *m*	64	38.5, CH$_2$	1.40 [a], *m*
		1.49 [a], *m*			1.49 [a], *m*
65	73.1 [a], CH	3.50 [a], *m*	65	72.9 [a], CH	3.50 [a], *m*
66	36.1, CH$_2$	1.36, *m*	66	36.1, CH$_2$	1.35 [a], *m*
		1.51 [a], *m*			1.51 [a], *m*
67	34.2 [a], CH$_2$	1.10 [a], *m*	67	34.2 [a], CH$_2$	1.11 [a], *m*
		1.50 [a], *m*			1.50 [a], *m*
68	34.2 [a], CH	1.42 [a], *m*	68	34.2 [a], CH	1.43 [a], *m*
Me-68	20.2, CH$_3$	0.92 [a], *d* (6.5)	Me-68	20.2, CH$_3$	0.92 [a], *d* (6.5)
69	37.7, CH$_2$	1.20 [a], *m*	69	37.7, CH$_2$	1.19 [a], *m*
		1.39 [a], *m*			1.38 [a], *m*
70	25.1, CH$_2$	1.39 [a], *m*	70	25.1, CH$_2$	1.39 [a], *m*
		1.45 [a], *m*			1.44 [a], *m*
71	29.1, CH$_2$	1.64 [a], *m*	71	29.1, CH$_2$	1.64 [a], *m*
72	40.9, CH$_2$	2.92, *t* (7.7)	72	40.9, CH$_2$	2.92, *t* (7.7)

Gargantulide B (**1**)			Gargantulide C (**2**)		
nº	**δ ¹³C**	**δ ¹H (mult, _J_, Hz)**	**nº**	**δ ¹³C**	**δ ¹H (mult, _J_, Hz)**
1'	34.4 [b], CH₂	1.09, _m_	1'	34.4 [a], CH₂	1.10, _m_
		1.36 [a], _m_			1.35 [a], _m_
2'	21.6, CH₂	1.29, _m_	2'	21.6, CH₂	1.28, _m_
		1.48 [a], _m_			1.48 [a], _m_
3'	14.9, CH₃	0.89 [a], _t_ (7.2)	3'	14.9, CH₃	0.89 [a], _t_ (7.2)
Man-1	97.1, CH	4.58, br _s_	Man-1	96.9, CH	4.58, br _s_
Man-2	73.1 [a], CH	3.81 [a], _m_	Man-2	73.1 [a], CH	3.81 [a], _m_
Man-3	75.4, CH	3.47, _dd_ (9.5, 3.2)	Man-3	75.4, CH	3.47, _dd_ (9.5, 3.2)
Man-4	68.6 [a], CH	3.67 [a], _t_ (9.5)	Man-4	68.8 [a], CH	3.61[a], _t_ (9.5)
Man-5	78.5, CH	3.17, _m_	Man-5	78.6, CH	3.15, _m_
Man-6	62.8 [a], CH₂	3.75 [m'], _m_	Man-6	62.9 [a], CH₂	3.74 [m'], _m_
		3.87 [a], _m_			3.87[a], _m_
Glc-1	102.4, CH	4.20 [a], _d_ (7.8)	Glc-1	102.4, CH	4.22 [a], _d_ (7.8)
Glc-2	74.9, CH	3.20 [a], _m_	Glc-2	74.9, CH	3.20 [a], _m_
Glc-3	78.2, CH	3.33 [a], _m_	Glc-3	78.2, CH	3.33 [a], _m_
Glc-4	71.8, CH	3.32 [a], _m_	Glc-4	71.7, CH	3.33 [a], _m_
Glc-5	78.3 [a], CH	3.21 [a], _m_	Glc-5	78.3 [a], CH	3.21 [a], _m_
Glc-6	62.8 [a], CH₂	3.68 [a], _m_	Glc-6	62.9 [a], CH₂	3.68 [a], _m_
		3.84 [a], _m_			3.85 [a], _m_
maG-1	104.4, CH	4.46, _d_ (7.5)	maG-1	104.7, CH	4.45, _d_ (7.4)
maG-2	70.5, CH	3.50[a], _dd_ (10.4, 7.5)	maG-2	70.6, CH	3.49[a], _dd_ (10.6, 7.4)
maG-3	66.1, CH	3.06, _t_ (10.4)	maG-3	66.0, CH	3.05, _t_ (10.6)
maG-4	71.5, CH	3.39[a], _dd_ (10.4, 8.7)	maG-4	71.5, CH	3.38 [a], _dd_ (10.6, 8.8)
maG-5	74.4, CH	3.40, _dq_ (8.7, 6.2)	maG-5	74.3, CH	3.41, _dq_ (8.8, 6.1)
maG-6	18.4, CH₃	1.34 [a], _d_ (6.2)	maG-6	18.4, CH₃	1.33[a], _d_ (6.1)
N-Me	31.1, CH₃	2.81, br _s_	N-Me	31.1, CH₃	2.81, br _s_
Ara-1	108.3, CH	5.01, _d_ (1.6)	Ara-1	108.2, CH	5.01, _d_ (1.5)
Ara-2	83.9, CH	3.96, _dd_ (3.9, 1.6)	Ara-2	83.9, CH	3.96, _dd_ (3.9, 1.5)
Ara-3	78.7, CH	3.86 [a], _m_	Ara-3	78.6, CH	3.85 [a], _m_
Ara-4	85.4, CH	3.98, _m_	Ara-4	85.4, CH	3.98, _m_
Ara-5	63.2, CH₂	3.64 [a], _m_	Ara-5	63.2, CH₂	3.64 [a], _m_
		3.74 [a], _m_			3.75 [a], _m_
Glc´-1	103.9, CH	4.33, _d_ (7.6)			
Glc´-2	75.2, CH	3.21 [a], _m_			
Glc´-3	78.3 [a], CH	3.37 [a], _m_			
Glc´-4	72.5, CH	3.23 [a], _m_			
Glc´-5	78.3 [a], CH	3.21 [a], _m_			
Glc´-6	63.4, CH₂	3.93, _m_			
		3.70 [a], _m_			

_a_ Overlapping with other isochronous ¹³C-NMR / ¹H-NMR signals.

* No multiplicity is given (m) for ¹H-NMR signals for which coupling constants (_Js_) could not be measured directly in ¹H or JRES spectra nor determined from HSQC traces. The assignments were supported by HSQC, HMBC, TOCSY, and HSQC-TOCSY.

Figure S4. 1D/2D NMR spectra of gargantulide B (**1**):

1H   gargantulide B (1)..esp
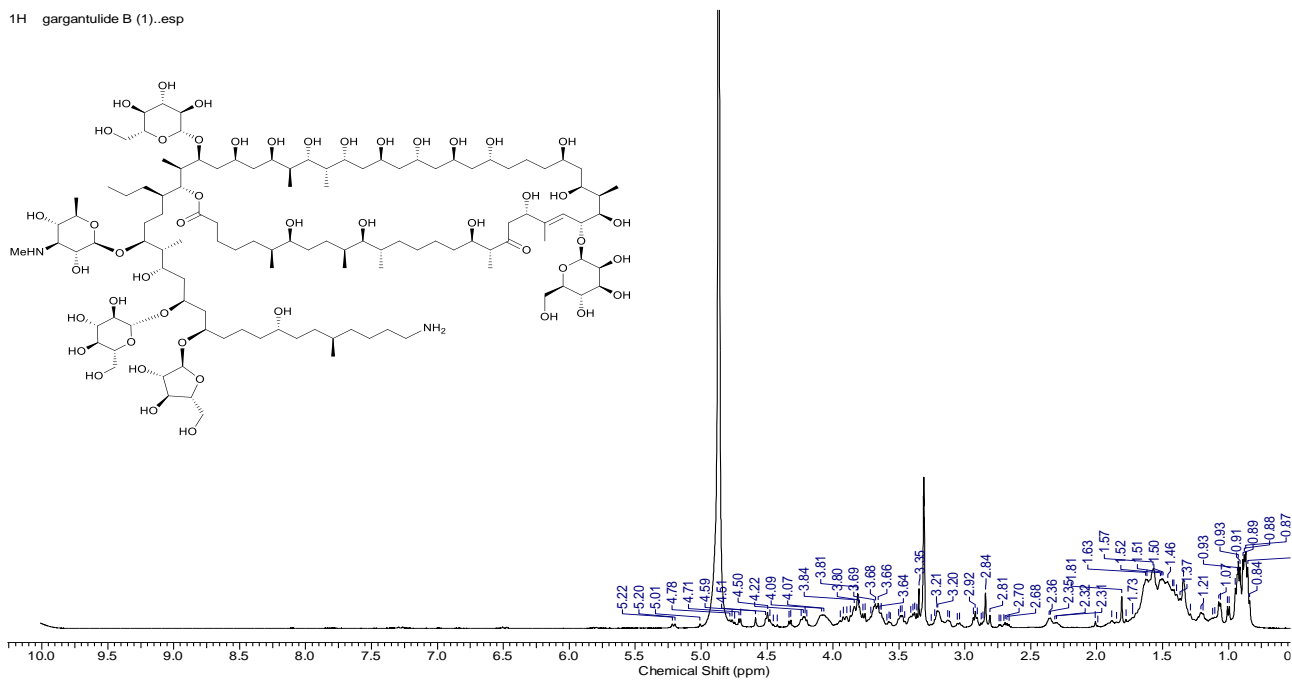
a)   ¹H NMR spectrum (CD₃OD, 500 MHz) of **1**. Full scale (0-10 ppm).

1H   gargantulide B (1)..esp
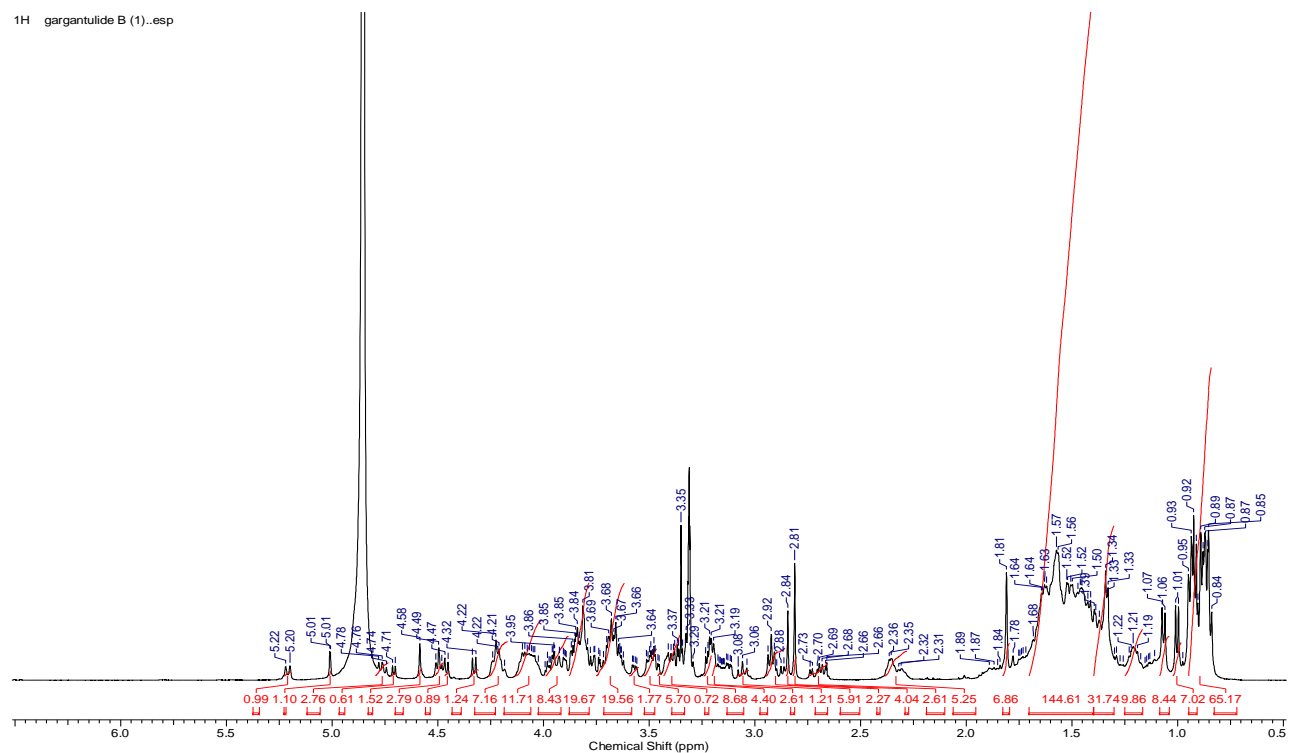
b) ¹H NMR spectrum (CD₃OD, 500 MHz) of **1**. Expansion between 0 and 7 ppm showing integral values.

c) ^{13}C NMR spectrum (CD$_3$OD, 125 MHz) of **1.**

d) ^{13}C NMR spectrum (CD$_3$OD, 125 MHz) of **1**. Expansion between 0 and 65 ppm.

13C gargantulide B (1).esp

e) ^{13}C NMR spectrum (CD$_3$OD, 125 MHz) of **1**. Expansion between 65 and 125 ppm.

COSY   gargantulide B (1).esp
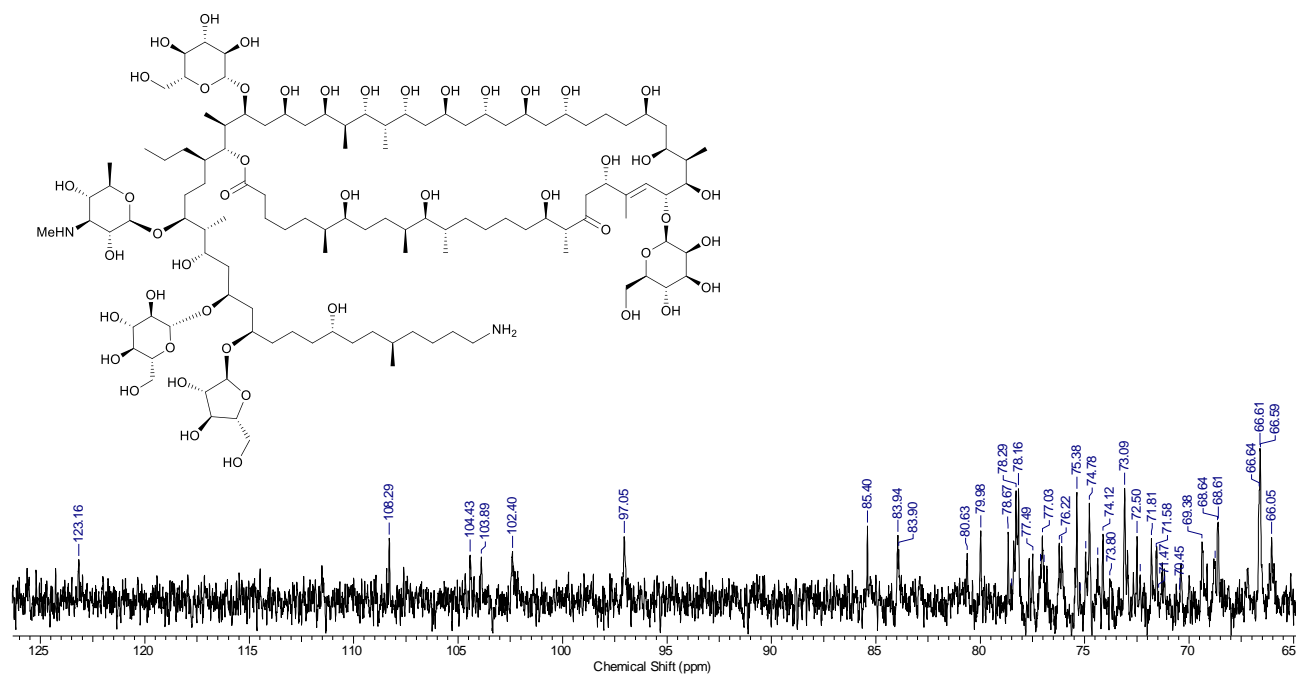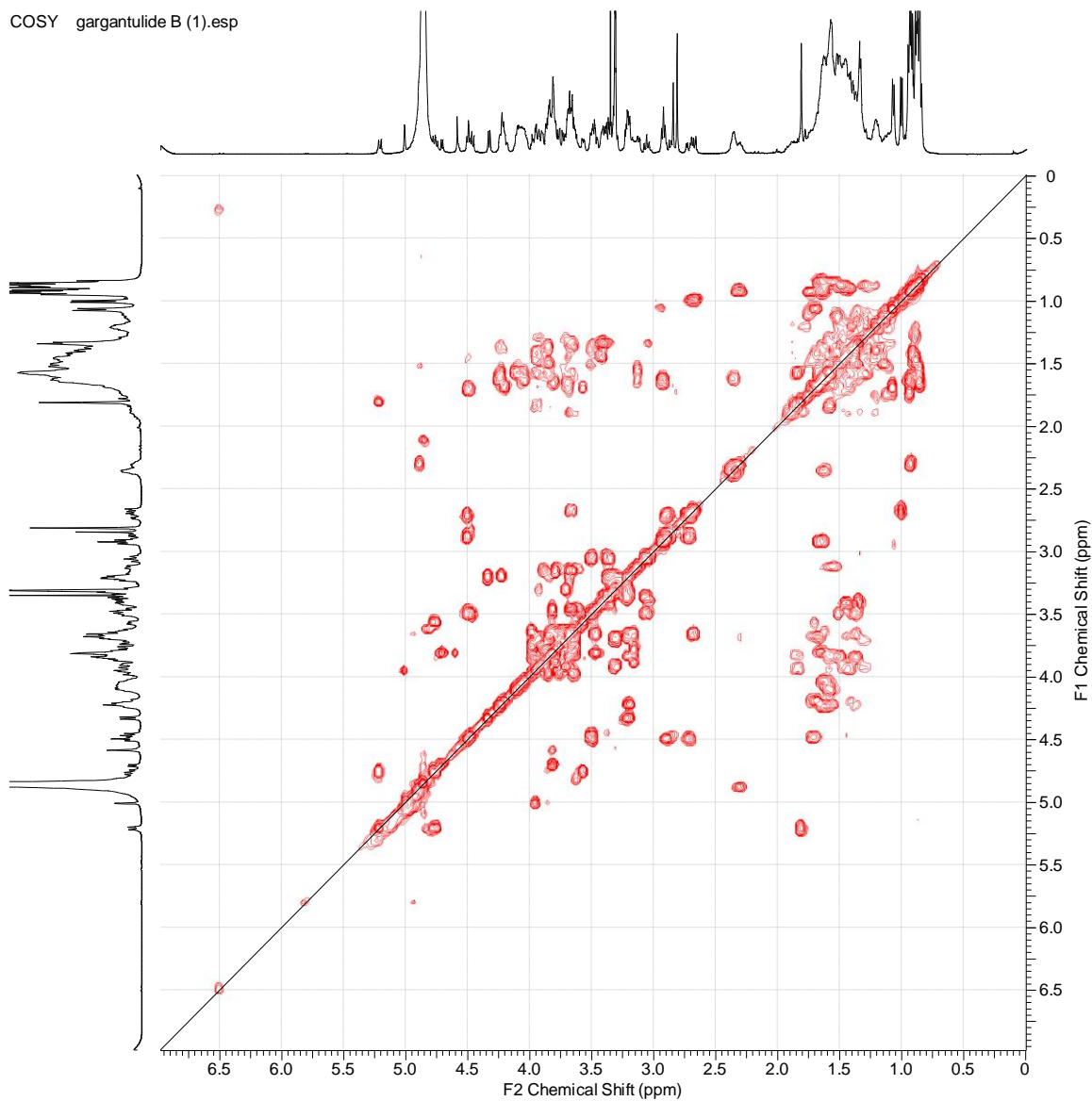
f) COSY spectrum of **1**.

g) HSQC spectrum of **1.**

HMBC   gargantulide B (1).esp

h) HMBC spectrum of **1.**

TOCSY   gargantulide B (1).esp

i) TOCSY spectrum of **1.**

NOESY   gargantulide B (1).esp
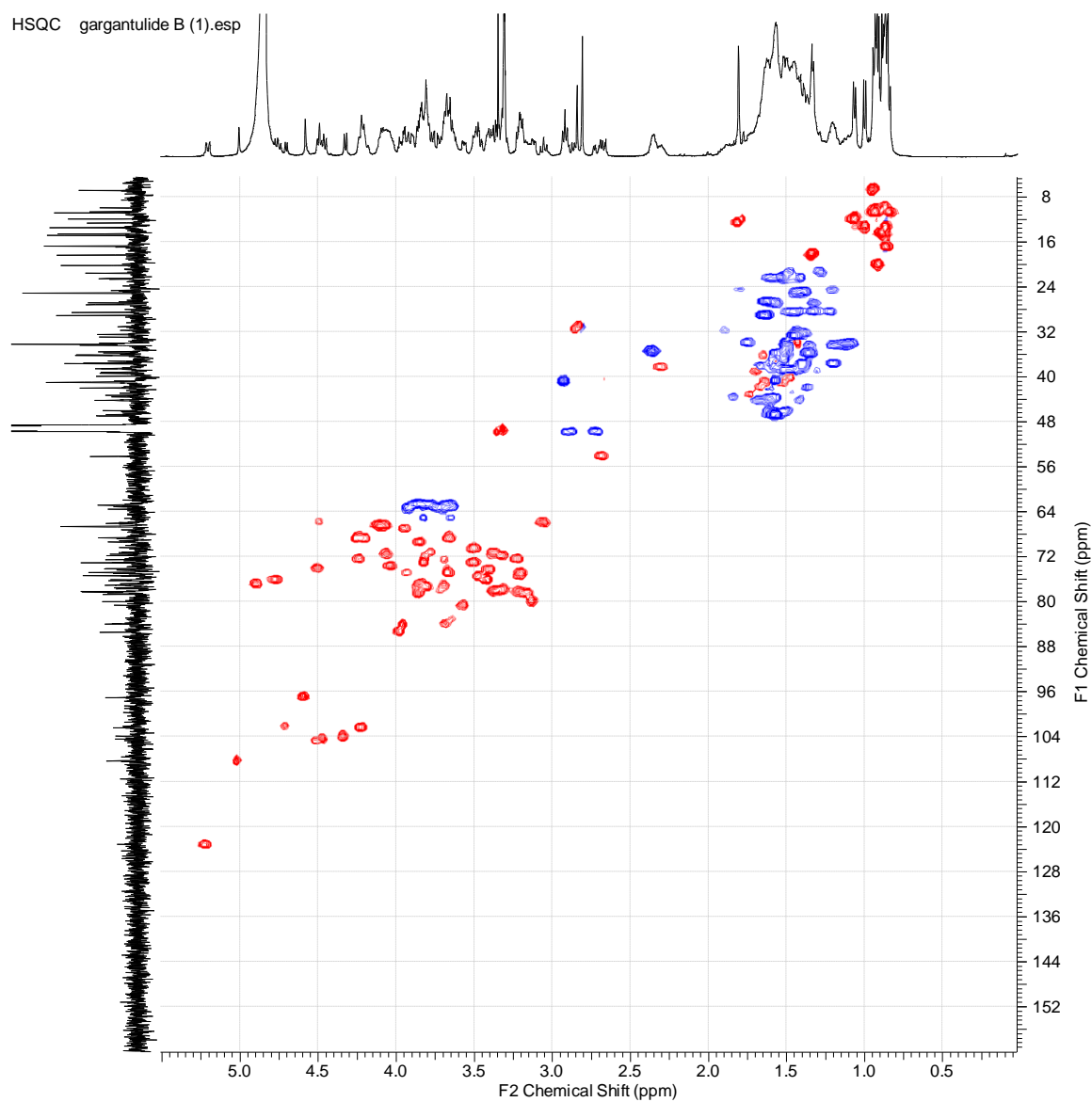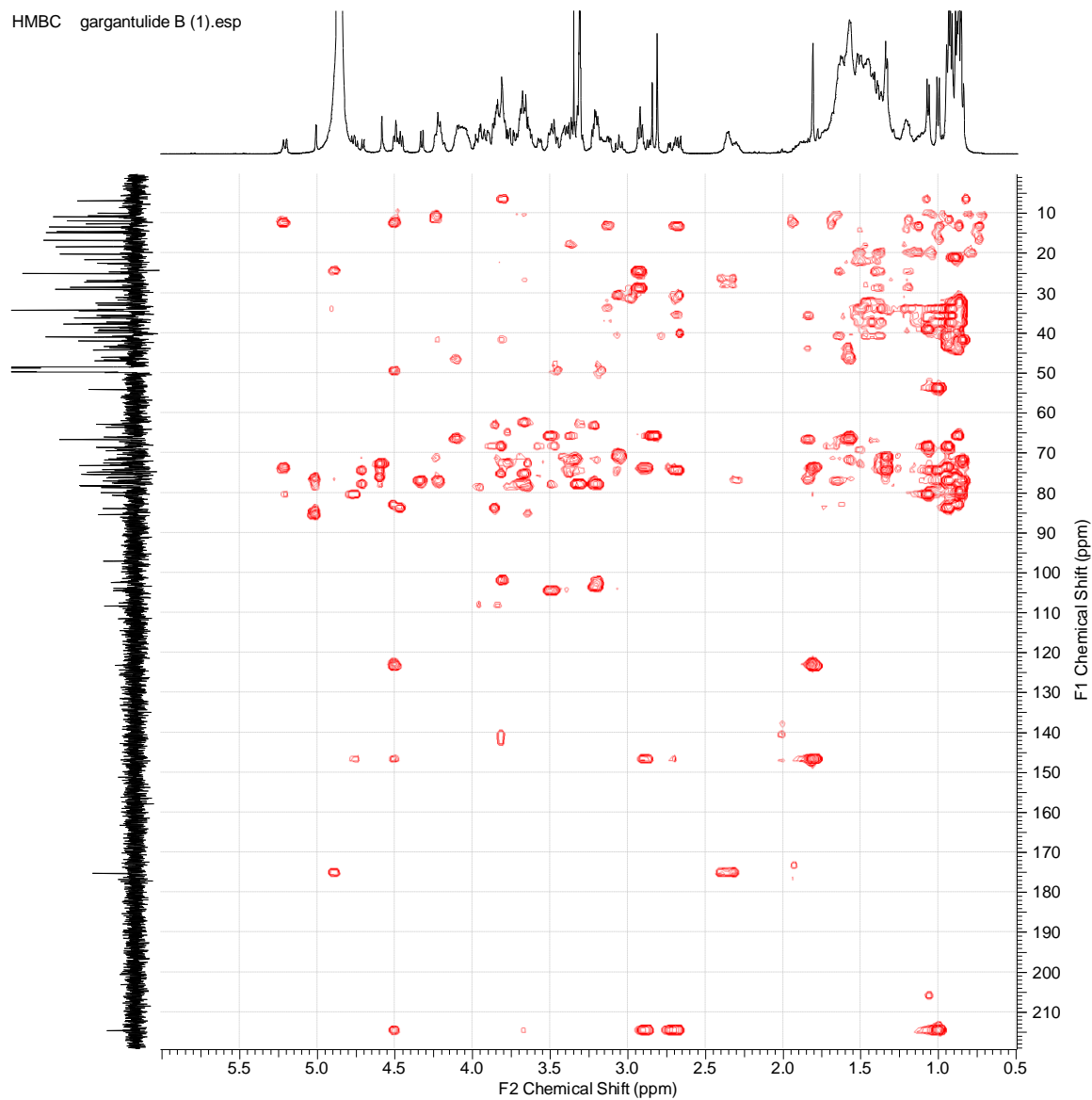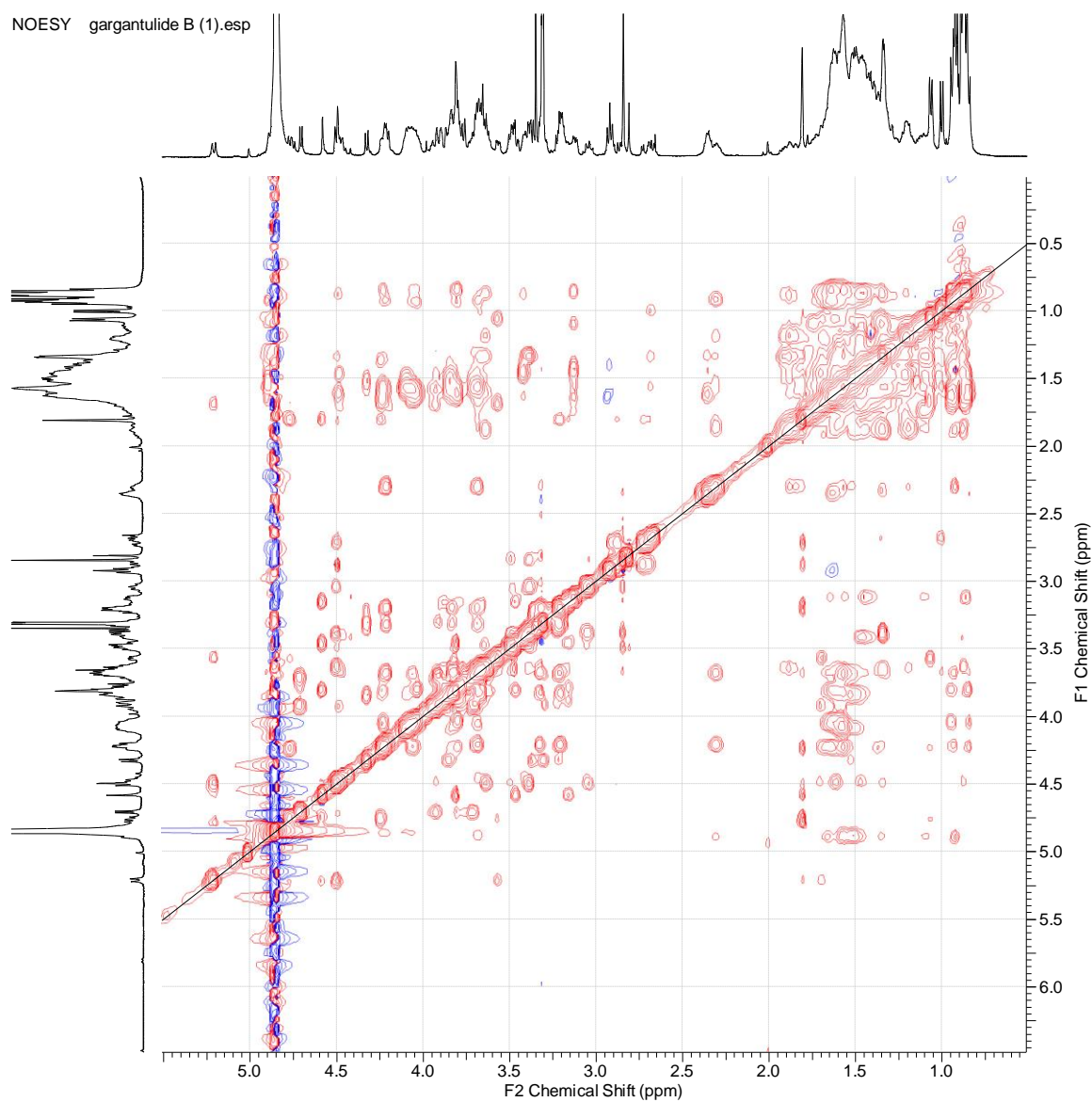
j) NOESY spectrum of **1.**

k) HSQC-TOCSY spectrum of **1.**

JRES gargantulide B (1).esp
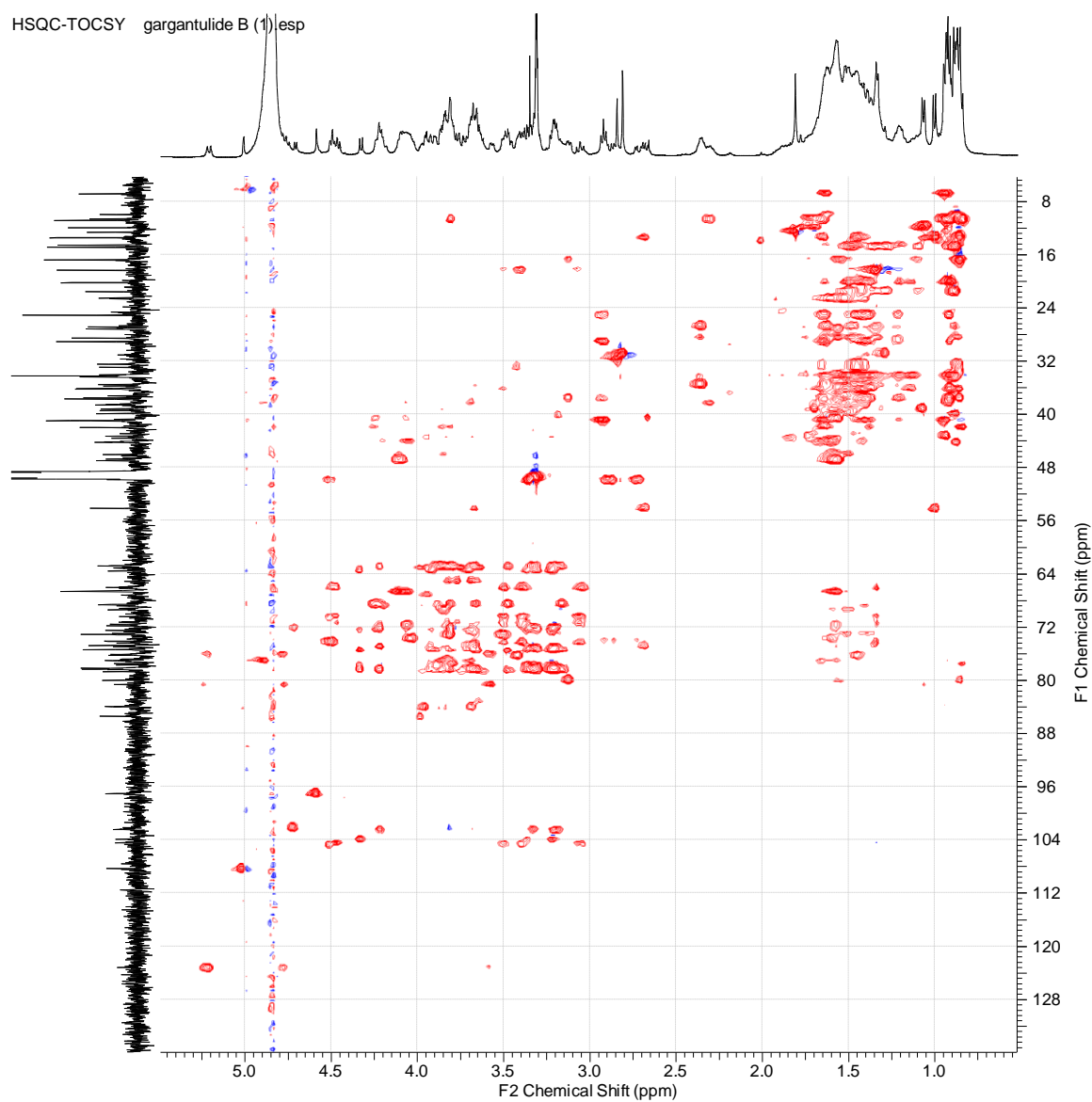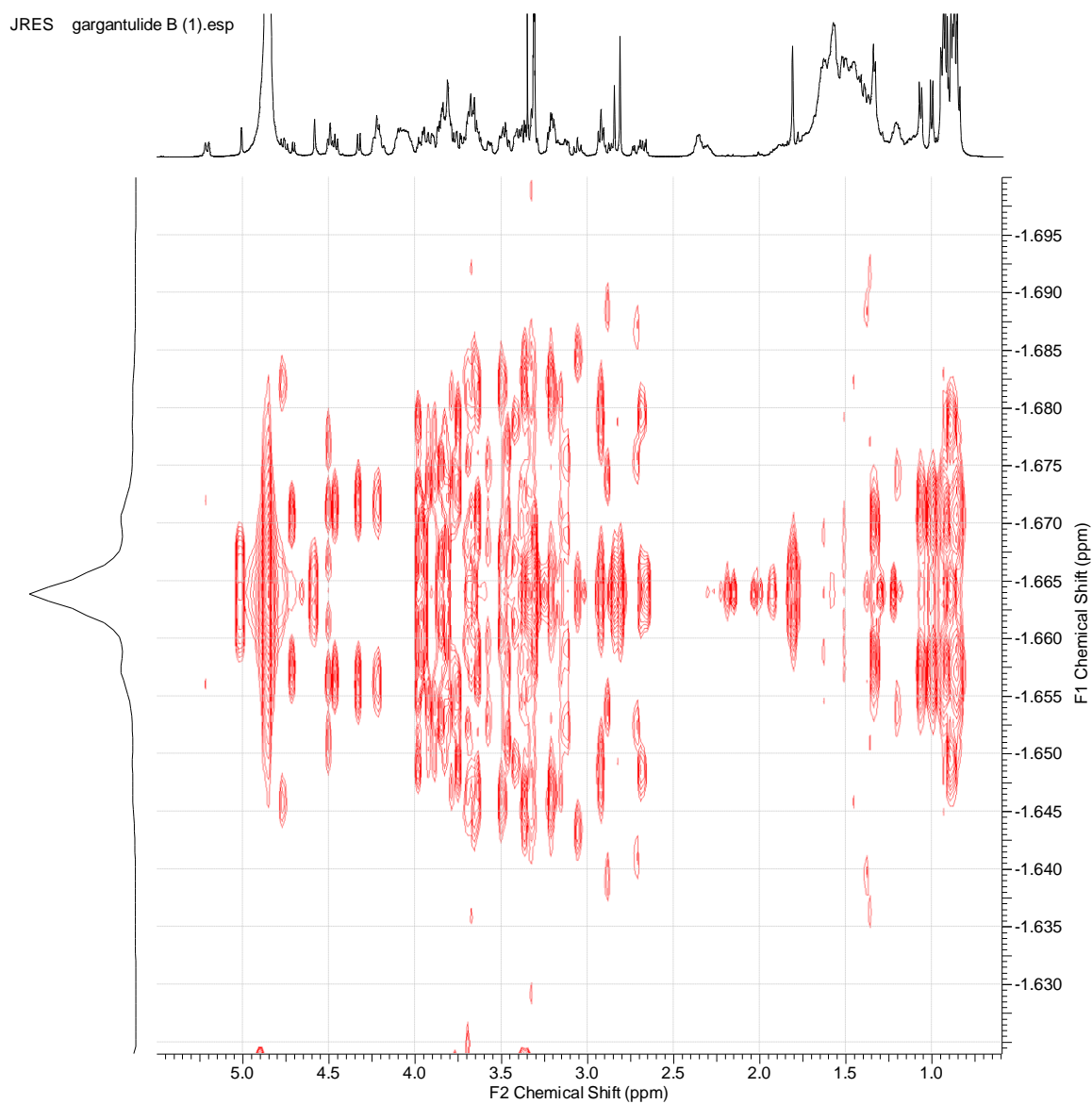
l) JRES spectrum of **1.**

Figure S5. 1D/2D NMR spectra of gargantulide C (**2**)
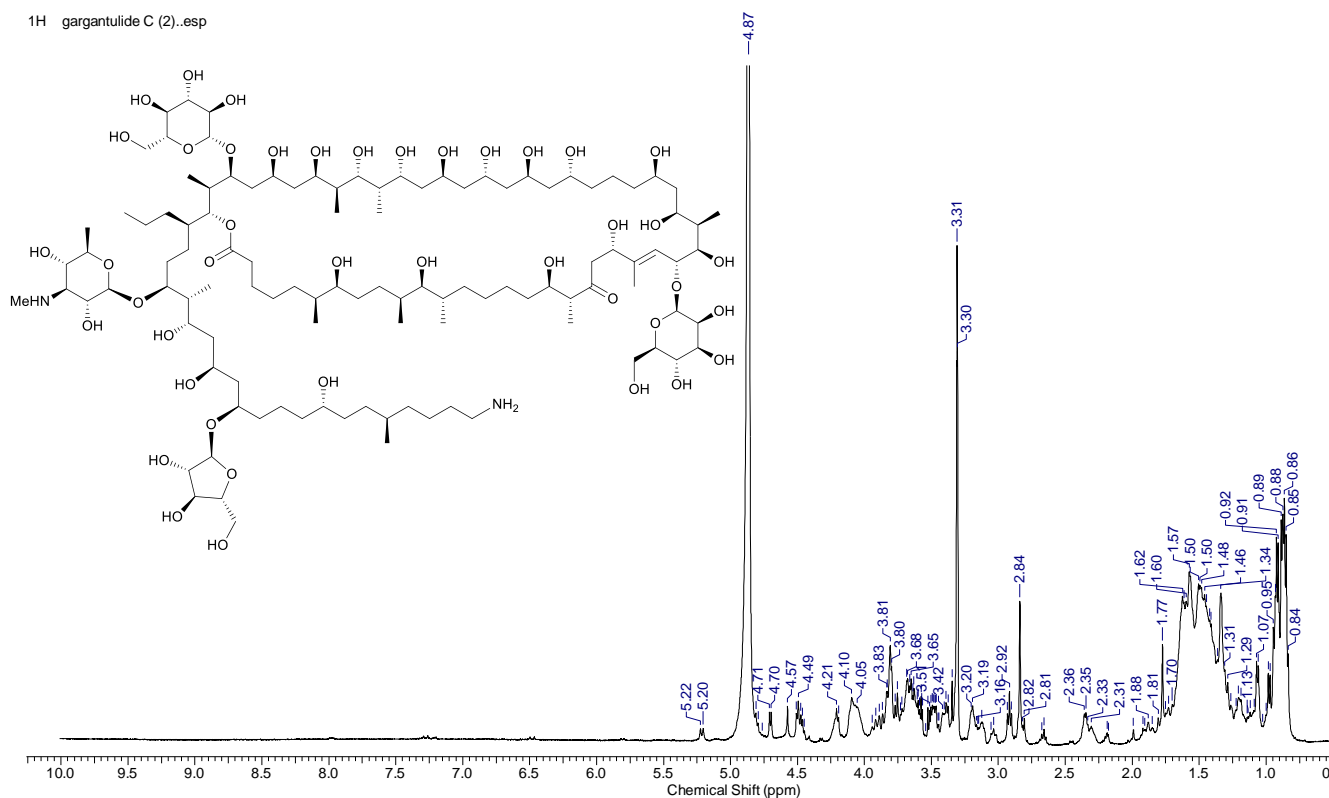
a) ¹H NMR spectrum (CD₃OD, 500 MHz) of **2**. Full scale (0-10 ppm).
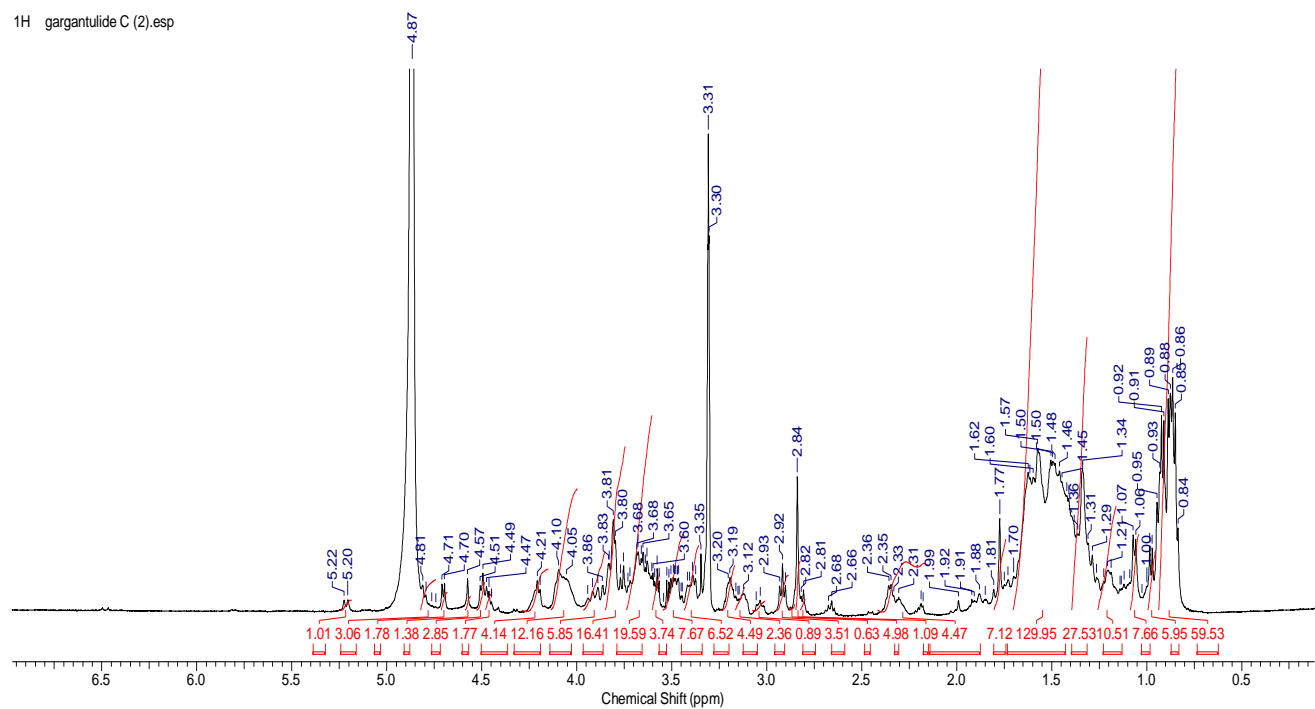
b) ¹H NMR spectrum (CD₃OD, 500 MHz) of **2**. Expansion between 0 and 7 ppm showing integral values.

13C gargantulide C (2).esp
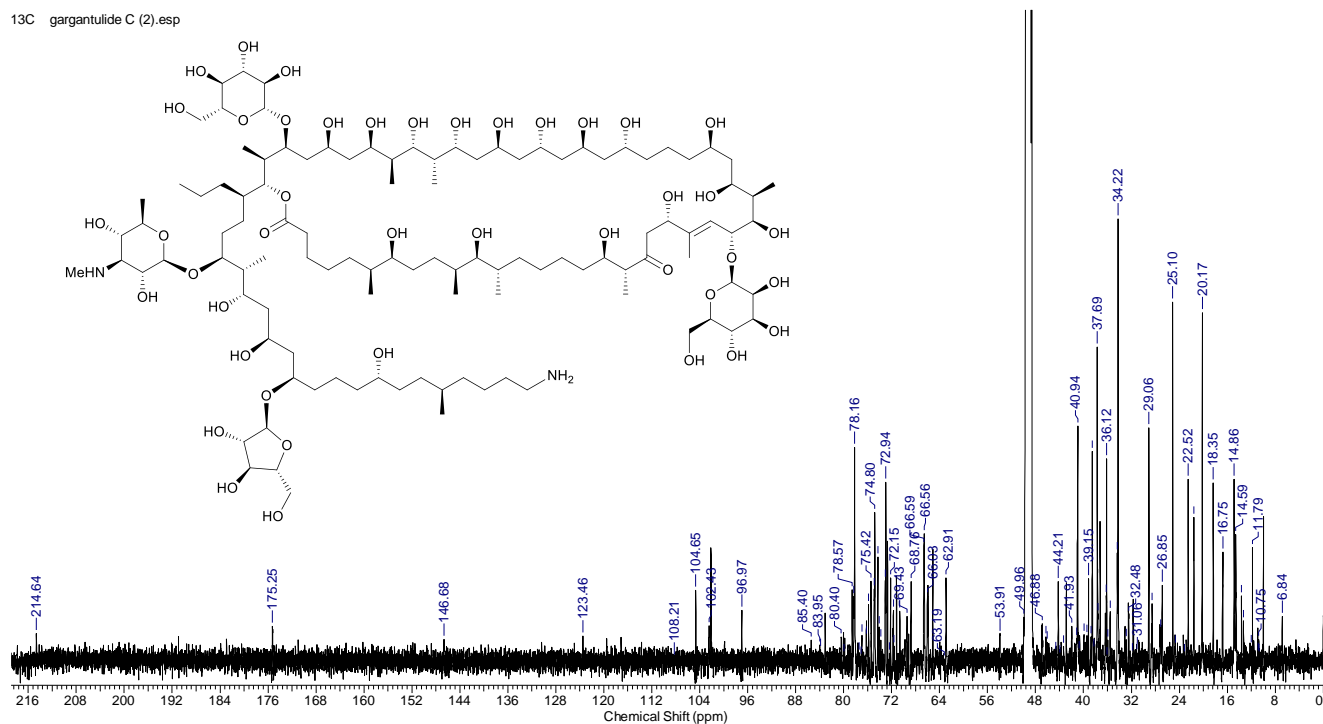
c) ^{13}C NMR spectrum (CD$_3$OD, 125 MHz) of **2**.

13C gargantulide C (2).esp
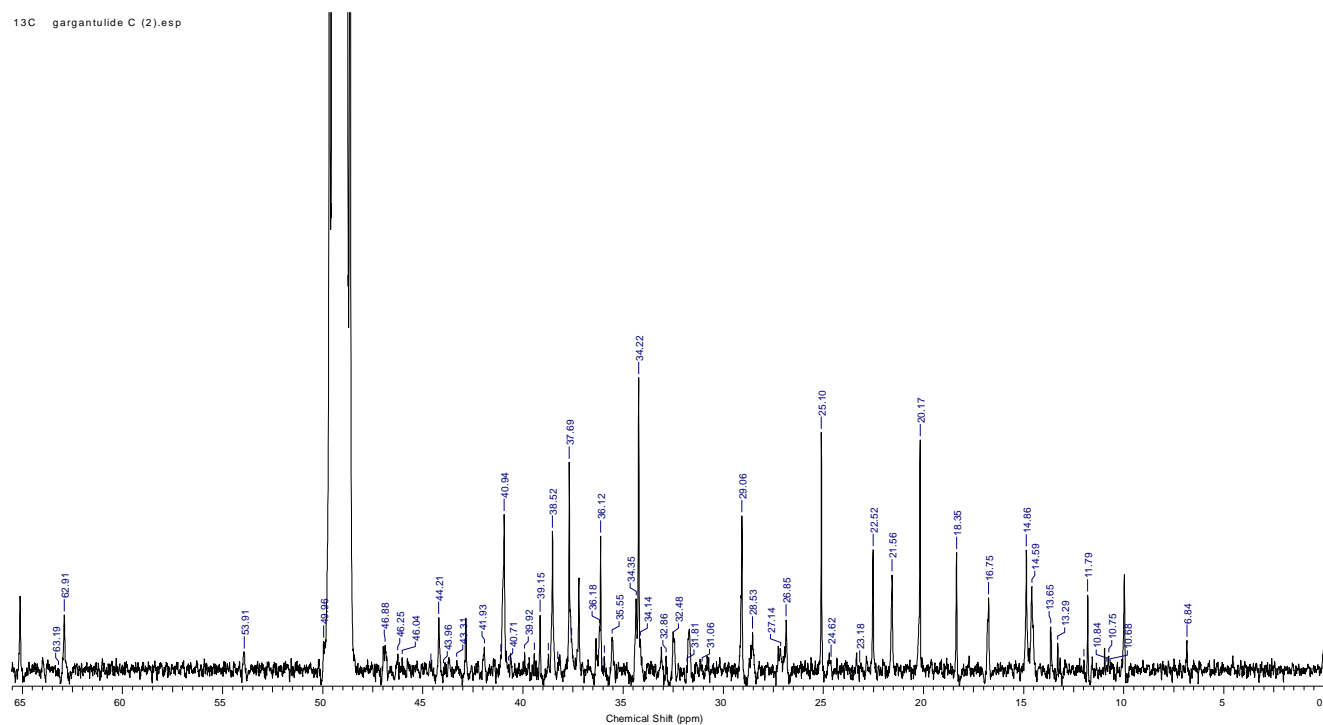
d) ^{13}C NMR spectrum (CD$_3$OD, 125 MHz) of **2**. Expansion between 0 and 65 ppm.

13C gargantulide C (2).esp
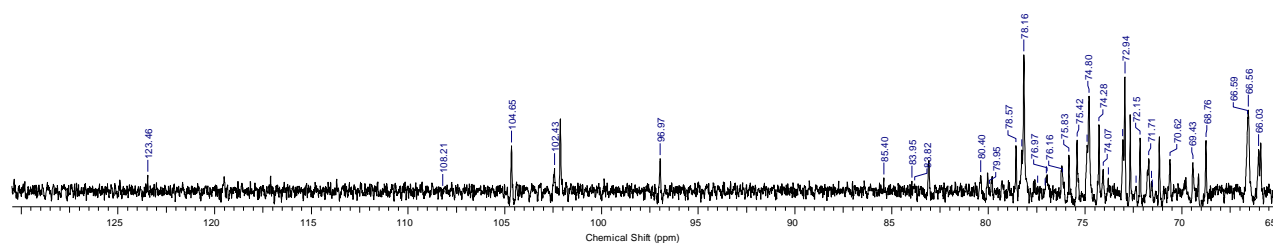
e) ^{13}C NMR spectrum (CD$_3$OD, 125 MHz) of **2**. Expansion between 65 and 130 ppm.

COSY gargantulide C (2).esp

f) COSY spectrum of **2.**

HSQC   gargantulide C (2).esp

g) HSQC spectrum of **2.**

HMBC   gargantulide C (2).esp

h) HMBC spectrum of **2.**

TOCSY   gargantulide C (2).esp
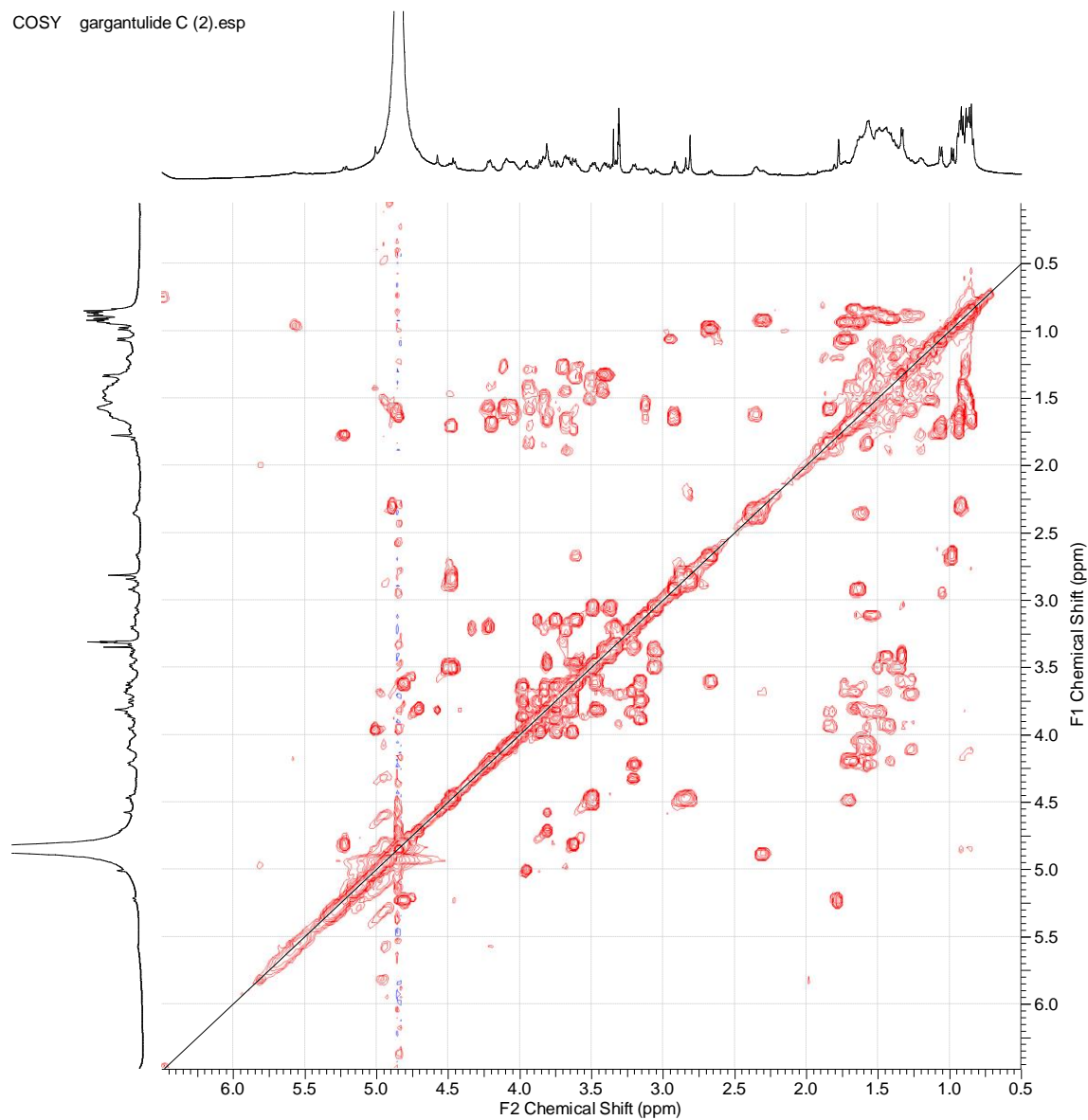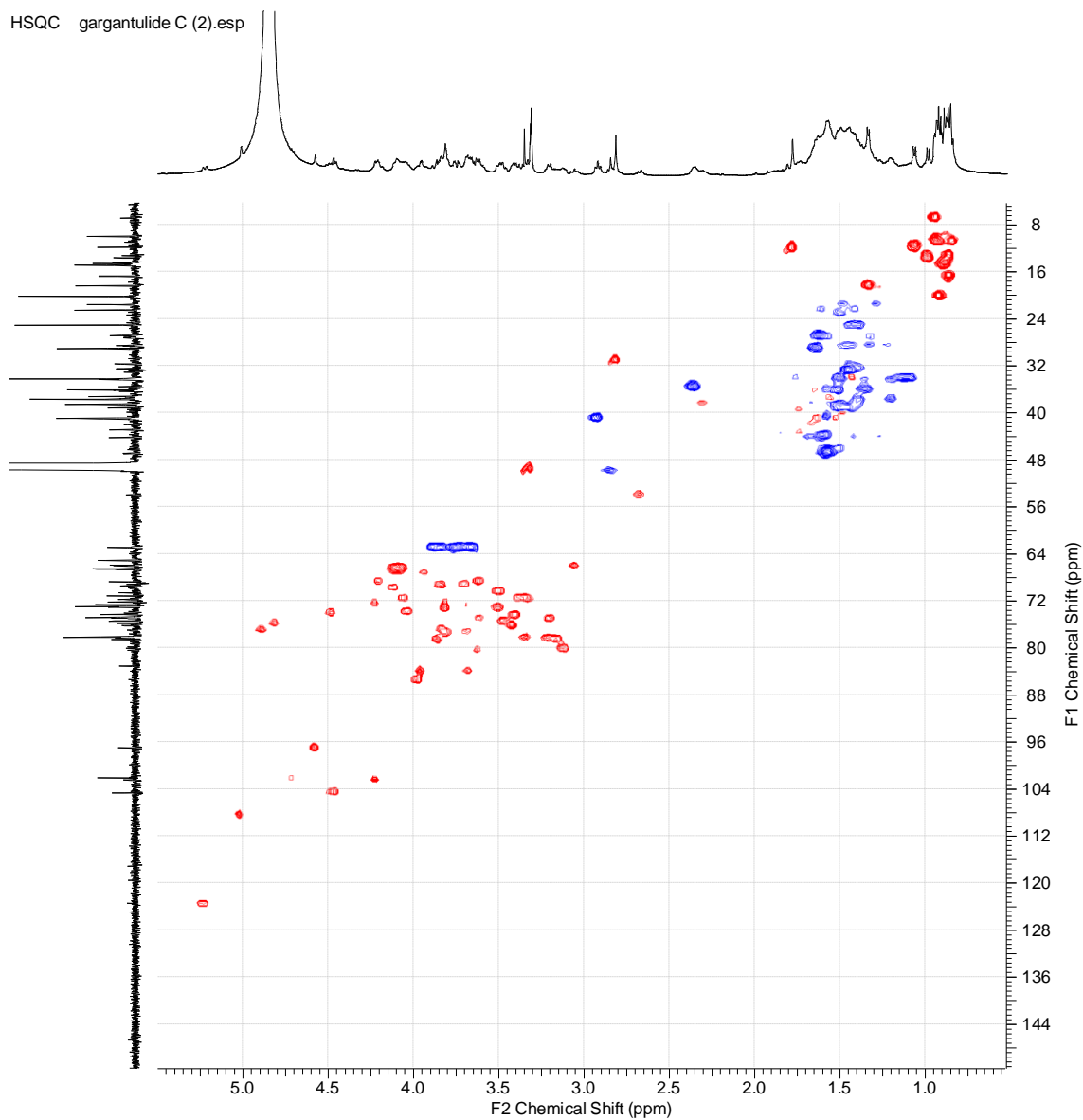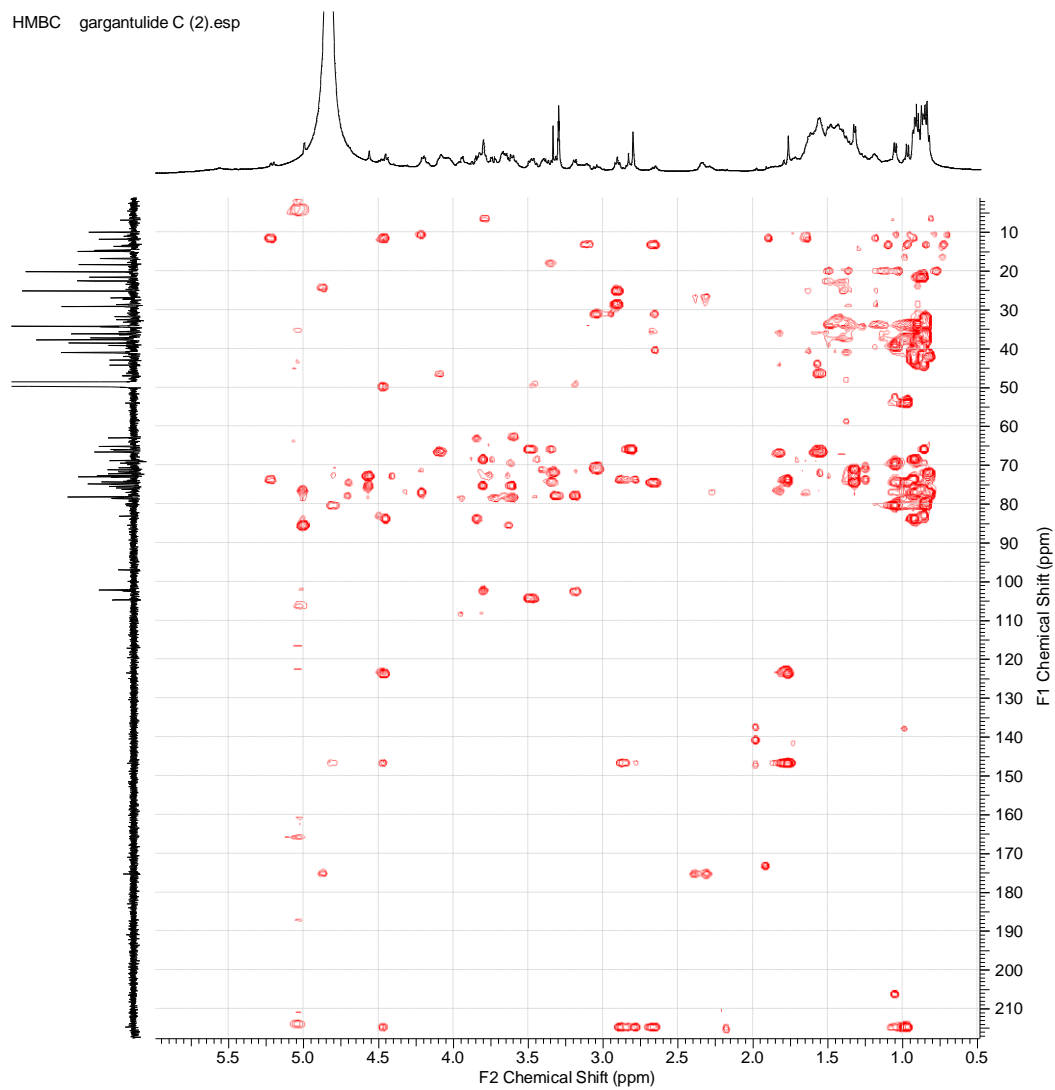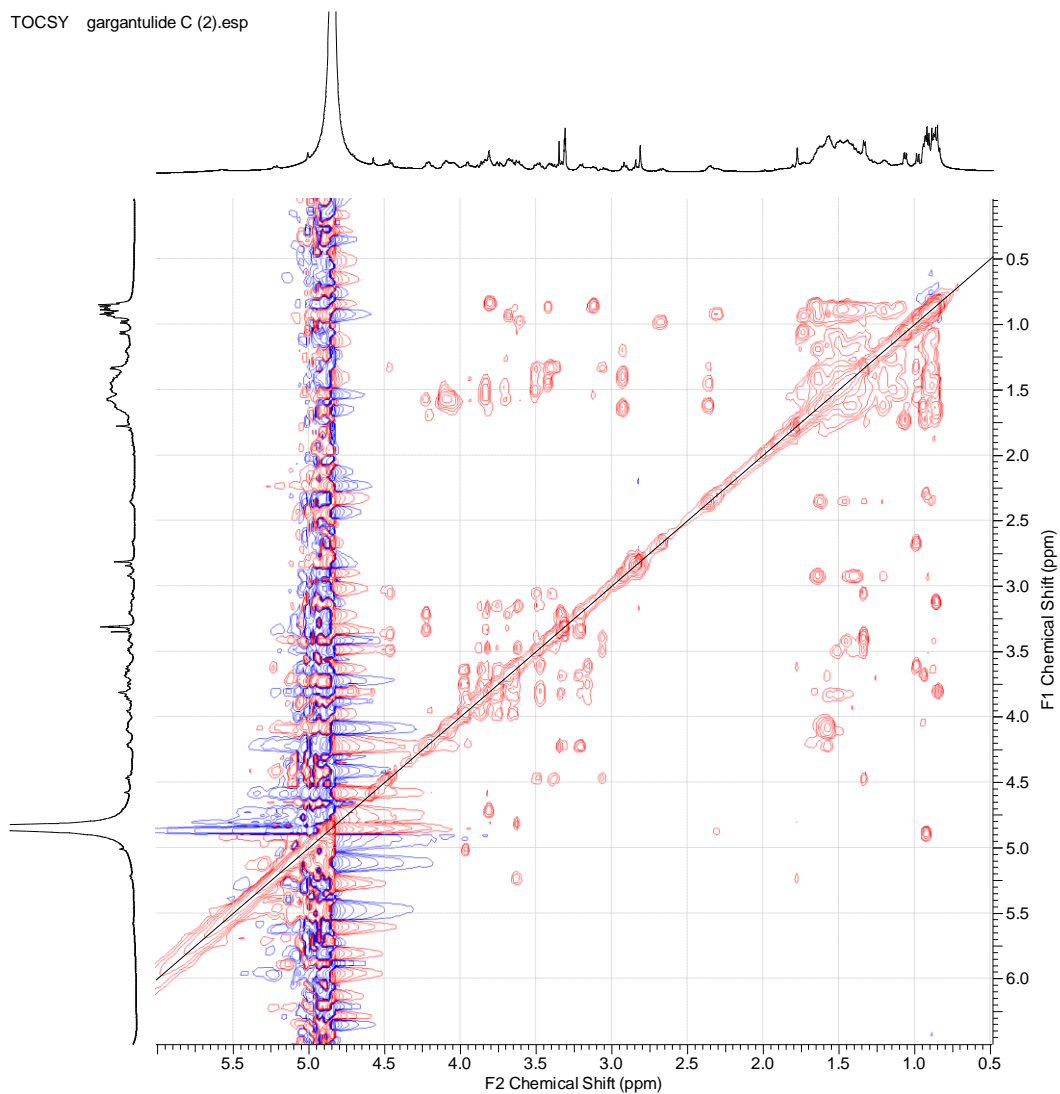
i) TOCSY spectrum of **2.**
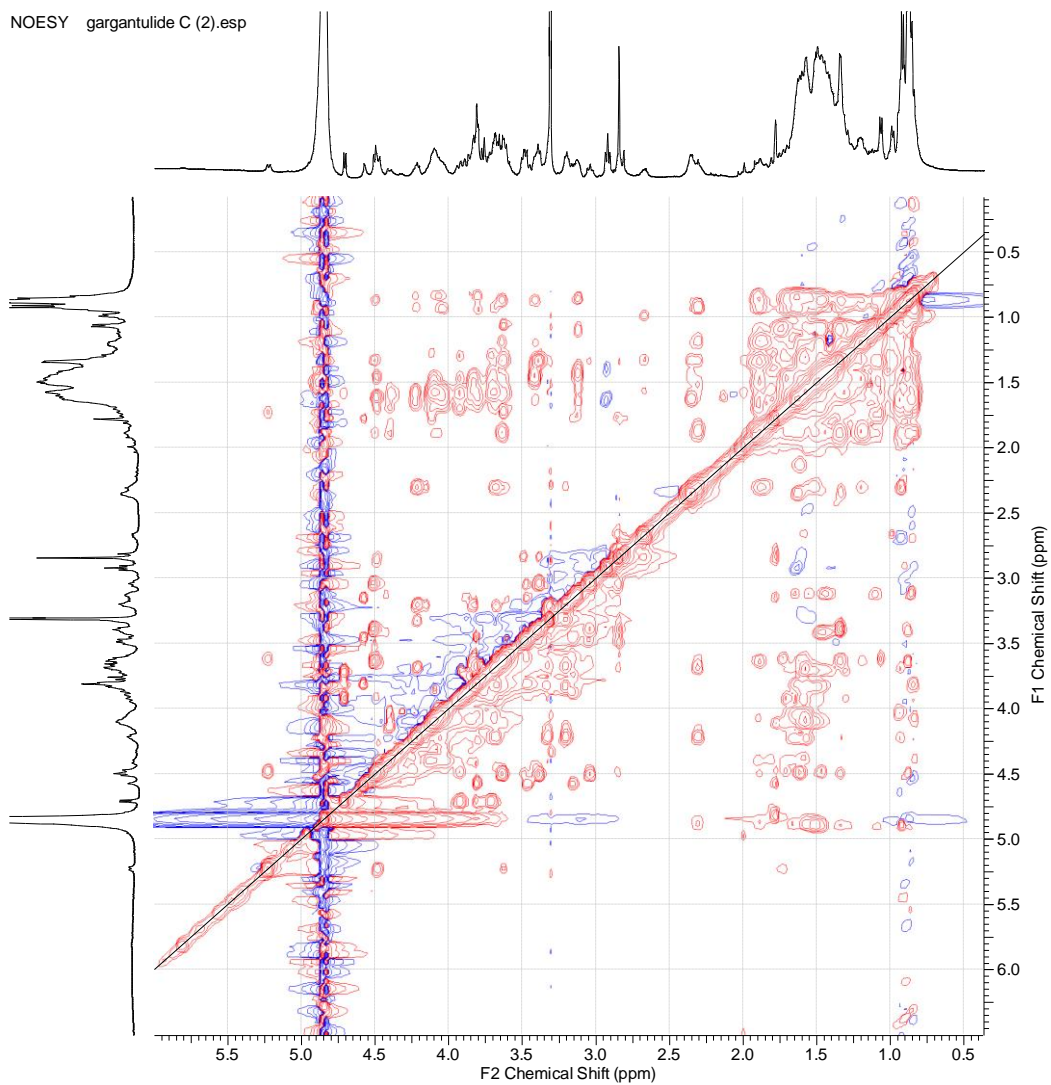
j) NOESY spectrum of **2.**

k) HSQC-TOCSY spectrum of **2.**

JRES gargantulide C (2).esp
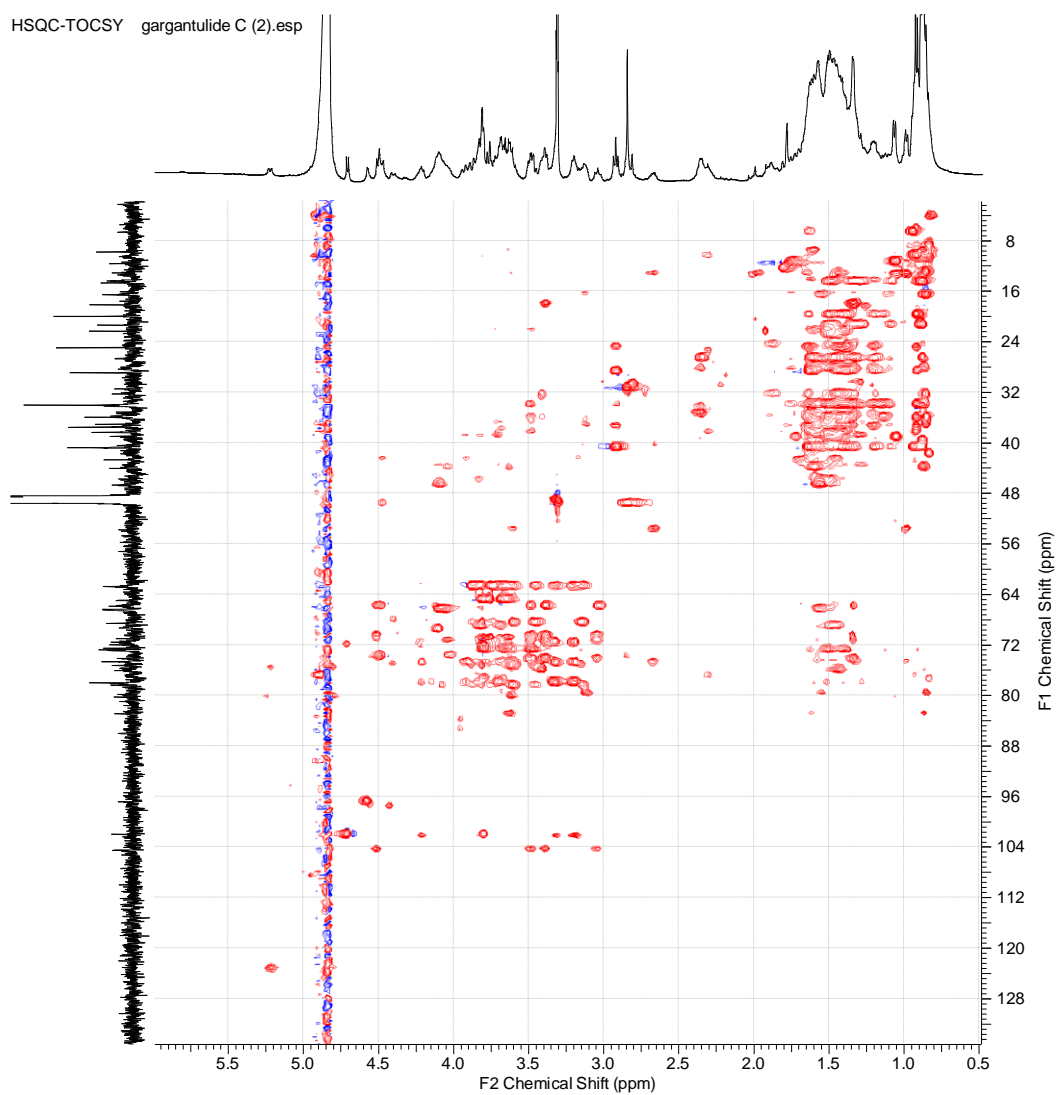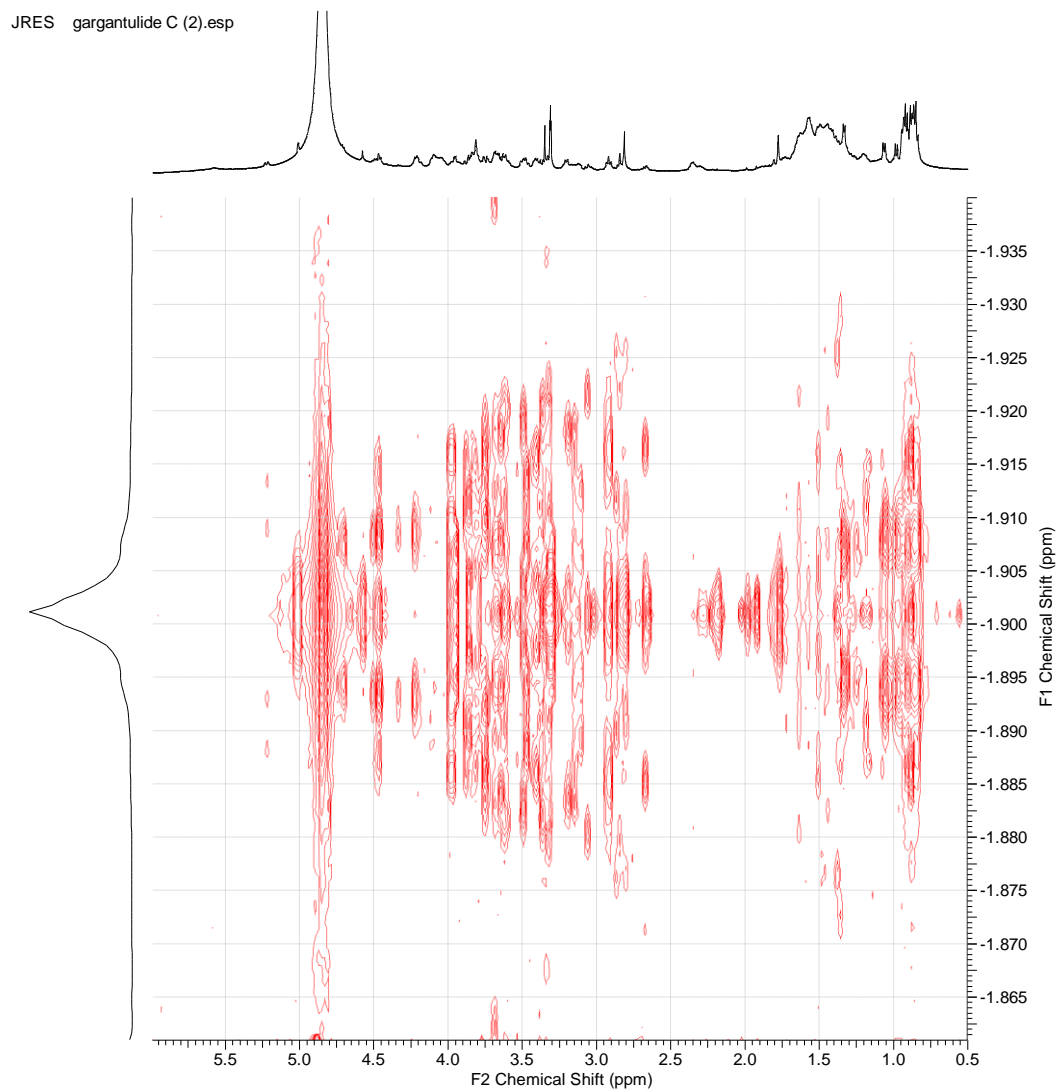
l) JRES spectrum of **2.**

Figure S6. Key COSY, TOCSY, HSQC-TOCSY and HMBC correlations observed for **2**
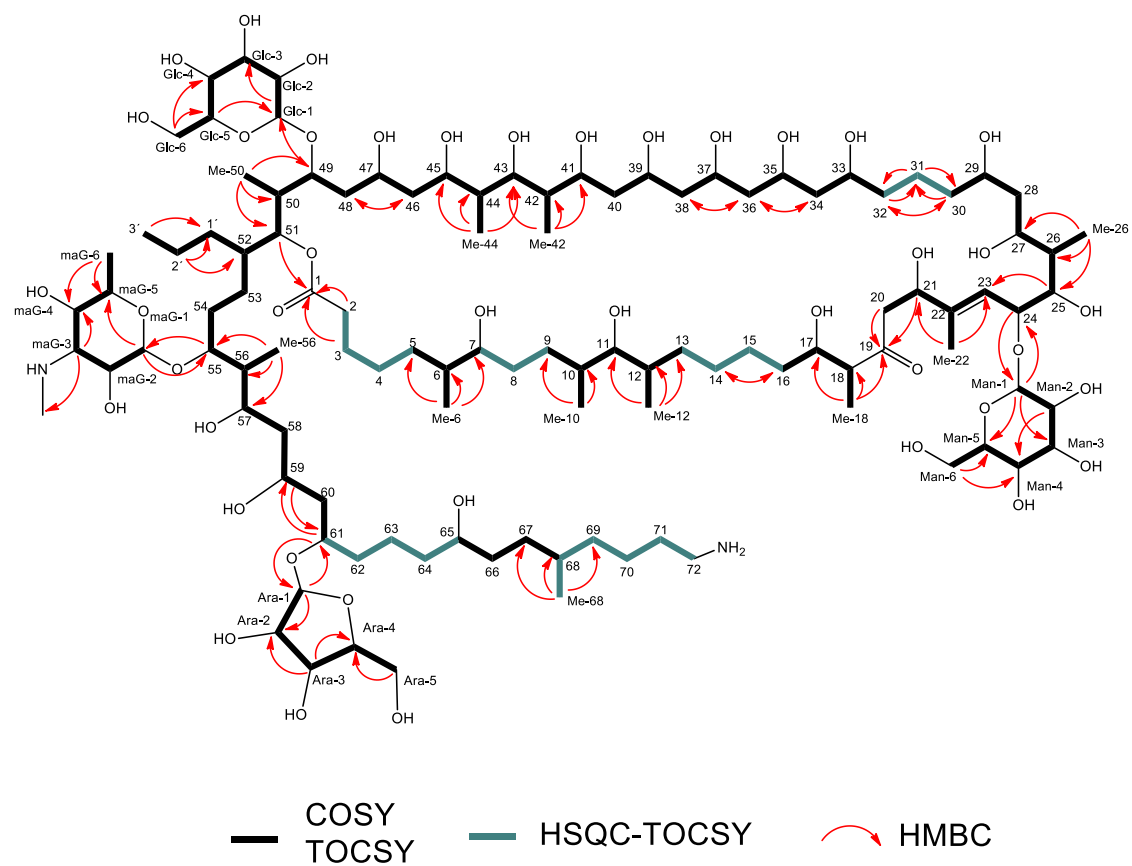
Figure S7. LC-HRMS co-detection of gargantulides A (**3**), B (**1**) and C (**2**) in the acetone crude extract of a micro-scale fermentation. (UV 210 nm: pink trace; MS+: blue trace) HRESIMS(+)-TOF spectra (ISCID 0 eV) of gargantulides A (triply and doubly-charged adducts) B and C (doubly-charged ions).
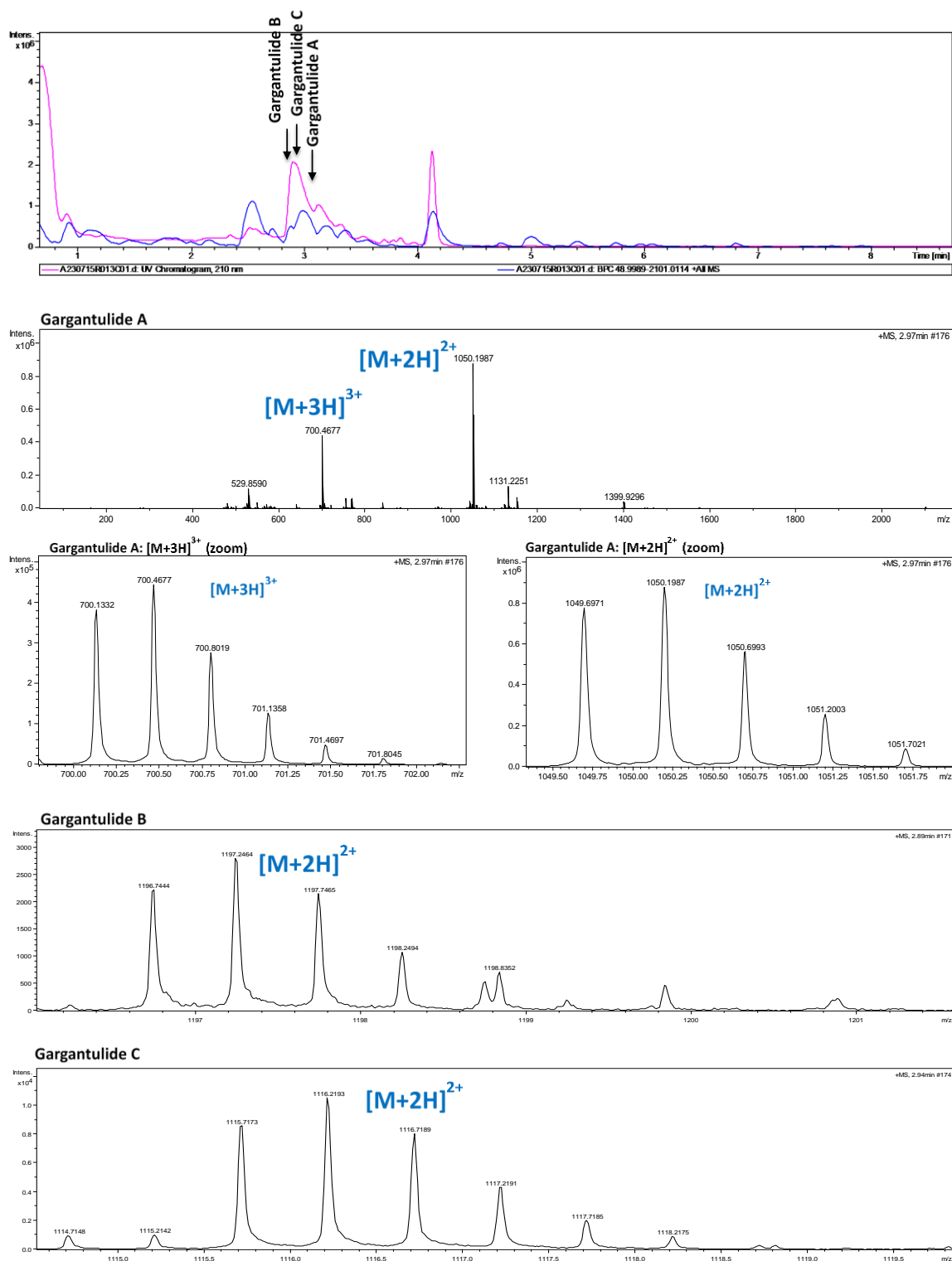
Table S2. antiSMASH results (bacterial version, relaxed mode). All types of BGCs are listed as detected by antiSMASH.

Region number	Type	Approx. size,kbp	Most similar known cluster	Remarks
1	PKS-like	40,855	No hits	
2	NRPS	44,211	Daptomycin (4%)	
3	Siderophore	11,832	Macrotetrolide (33%)	
4	RIPP-like	9,971	Enduracidin (4%)	
5	LAP, thiopeptide	27,977	No hits	
6	Terpene	21,678	Geosmin (100%)	
7	NRPS-like, terpene	42,088	Isorenieratene (42%)	Possibly two clusters
8	T1PKS	45,457	EDHA (33%)	
9	Terpene	19,827	Lipopetide 8D1-1 (6%)	
10	T1PKS	45,036	Sch-47554/Sch-47555 (3%)	
11	T1PKS, oligosaccharide	216,516	Funisamine (30%)	Putative cluster, coding for gargantulide compounds
12	LAP, ladderane, thioamide-NRP, NRPS	76,620	Ishigamide (72%)	Possibly several clusters
13	Other, thiopeptide, terpene, lantipeptide clas II	72,243	2-methylisoborneol (100%)	
14	NAPAA	31,640	Himastatin (8%)	
15	NRPS	62,803	Albachelin (70%)	
16	CDPS	20,683	No hits	
17	PKS-like	40,596	No hits	
18	Lantipeptide class I	25,334	No hits	
19	Lantipeptide class III	22,579	Ery-9 / Ery-6 / Ery-8 / Ery-7 / Ery-5 / Ery-4 / Ery-3 (100%)	
20	Arylpolyene	40,340	Ibomycin (5%)	
21	Ectoine	8,532	Ectoine (100%)	
22	NRPS, T1PKS	73,041	Ecumicin (10%)	
23	thiopeptide, LAP	30,258	No hits	
24	redox-cofactor	22,025	Lankacidin C (20%)	
25	cyanobactin, PKS-like, NRPS, T1PKS, lantipeptide class V, linaridin	85,858	prenylagaramide B / prenylagaramide C (8%)	
26	Lantipeptide class V	40,587	Arginomycin (13%)	
27	Linaridin	20,557	Legonaridin (25%)	
28	Lantipeptide class V	41,924	Ashimides (8%)	
29	Lantipeptide class II	23,053	No hits	
30	NAPAA	32,569	No hits	
31	hglE-KS, T1PKS	49,901	Rifamorpholine A/B/C/D/E (4%)	
32	NRPS	57,957	Telomycin (5%)	
33	NRPS, ectoine, T1PKS	80,296	Kosinostatin (16%)	
34	NRPS, NRPS-like	74,191	WS9326 (10%)	
35	NAPAA	32,384	No hits	
36	Terpene	21,115	SF2575 (6%)	
37	Ladderane, NRPS, T1PKS	173,101	Spiramycin (22%)	Possibly three clusters
38	NRPS, arylpolyene	130,298	Kedarcidin (14%)	
39	T1PKS	67,639	Meilingmycin (5%)	
40	Lassopeptide	24,502	No hits	
41	ladderane,NRPS,NRPS-like,transAT-PKS,T1PKS,NAPAA	245,836	Incednine (13%)	Possibly several clusters
42	NRPS, RIPP-like	49,672	Capreomycin IA/IB (6%)	
43	Furan	20,998	No hits	
44	NRPS, thioamitides	80,729	Mannopeptimycin(44%)	
45	Betalactone	25,210	No hits	

Table S3. Putative functions of genes in *gar* BGC.

Protein (NCBI accession)	a.a.	Proposed function	BLAST hit protein	Query coverage/ Per. identity/ E value	Accession number of BLAST hit
GarR1 (HUW46_01962)	938	Putative MalT family HTH-containing regulator	Helix-turn-helix transcriptional regulator (*Kutzneria albida*)	99% / 49.58% / 0.0	WP_025354 824.1
GarR2 (HUW46_01961)	414	Putative sensor histidine kinase	Sensor histidine kinase (*Asanoa ferruginea*)	95% / 51.23% / 3e-110	WP_116067 603.1
GarT1 (HUW46_01960)	240	Putative ABC-type multidrug transport system, ATPase component	ATP-binding cassette domain-containing protein (*Micromonospora pattaloongensis*)	96% / 51.95% / 1e-62	WP_091558 576.1
GarA (HUW46_01959)	467	Acyl-CoA ligase	Acyl-CoA ligase (*Kibdelosporangium aridum*)	99% / 73.18% /0.0	WP_084424 151.1
Orf1 (HUW46_01958)	180	PH domain-containing protein	PH domain-containing protein (Pseudonocardiaceae bacterium)	91% /59.76%/ 3e-56	MPZ66135.1
Orf2 (HUW46_01957)	524	PH domain-containing protein	PH domain-containing protein (*Kutzneria albida*)	95% /52.08 / 6e-169	WP_025356 057.1
GarR3 (HUW46_01956)	228	TetR family transcriptional regulator	TetR/AcrR family transcriptional regulator [*Kutzneria albida*]	81% / 48.92% / 4e- 55	WP_025359 812.1
GarT2 (HUW46_01955)	615	Putative ABC-type multidrug transport system, ATPase component	ATP-binding cassette domain-containing protein [*Amycolatopsis* sp. EGI 650086]	95% / 63.27 / 0.0	WP_158888 982.1
GarG1 (HUW46_01954)	396	Putative glycosyltransferase of MGT family	glycosyltransferase MGT family protein [*Mycolicibacterium thermoresistibile*]	98% / 57.11% / 6e-142	WP_003926 533.1
Orf3 (HUW46_01953)	84	Hypothetical protein	hypothetical protein [*Sciscionella* sp. SE31]	91% / 53.25% / 7e-17	WP_031470 089.1
GarG2 (HUW46_01952)	457	Putative glycosyltransferase of MGT family	glycosyltransferase family 1 protein [*Amycolatopsis* sp. SID8362]	92% / 43.88% / 5e-115	WP_160698 909.1
GarG3 (HUW46_01951)	421	Putative glycosyltransferase of MGT family	glycosyltransferase family 1 protein [*Allokutzneria albata*]	97% / 46.12% / 1e-115	WP_030432 743.1

Protein (NCBI accession)	a.a.	Proposed function	BLAST hit protein	Query coverage/ Per. identity/ E value	Accession number of BLAST hit
GarG4 (HUW46_01950)	419	Putative glycosyltransferase of MGT family	glycosyltransferase family 1 protein [*Allokutzneria albata*]	98% / 43.06% / 1e-104	WP_030432 743.1
GarP1 (HUW46_01949)	4568	Modular polyketide synthase	type I polyketide synthase [*Actinophytocola oryzae*]	97% / 55.09% / 0.0	WP_133901 401.1
GarP2 (HUW46_01948)	9791	Modular polyketide synthase	modular polyketide synthase [*Streptomyces* sp. RK95-74]	96% / 56.64% /0.0	BAW35608. 1
GarP3 (HUW46_01947)	5411	Modular polyketide synthase	PKS I [*Kutzneria albida* DSM 43870]	99% / 56.34% / 0.0	AHH99926.1
GarP4 (HUW46_01946)	6168	Modular polyketide synthase	type I polyketide synthase [*Streptomyces* sp. LamerLS-31b]	98% / 55.92% / 0.0	WP_093893 008.1
GarP5 (HUW46_01945)	8075	Modular polyketide synthase	beta-ketoacyl synthase [*Streptomyces hygroscopicus*]	99% / 54.13% / 0.0	AQW47576. 1
GarP6 (HUW46_01944)	4542	Modular polyketide synthase	type I polyketide synthase [*Streptomyces cinnamoneus*]	99% / 62.72% / 0.0	WP_099199 035.1
GarP7 (HUW46_01943)	5347	Modular polyketide synthase	type I polyketide synthase [*Amycolatopsis orientalis*]	100% / 54.65% / 0.0	WP_051173 827.1
GarP8 (HUW46_01942)	3617	Modular polyketide synthase	type I polyketide synthase [*Streptomyces cinnamoneus*]	99% / 59.56% / 0.0	WP_099199 033.1
GarP9 (HUW46_01941)	5662	Modular polyketide synthase	PKS I [*Kutzneria albida* DSM 43870]	99% / 58.65% / 0.0	AHH99926.1
GarP10 (HUW46_01940)	5910	Modular polyketide synthase	type I polyketide synthase [*Streptomyces cinnamoneus*]	95% / 61.75% / 0.0	WP_104650 952.1
GarB (HUW46_01939)	314	AT domain containing protein	ACP S-malonyltransferase [unclassified *Kitasatospora*]	96% / 53.77% / 5e-111	WP_057230 119.1
GarC (HUW46_01938)	421	Putative cytochrome p450	cytochrome P450 [*Streptomyces cinnamoneus*]	99% / 56.87 / 1e-165	WP_099199 030.1
GarR4 (HUW46_01937)	222	Putative response regulator	response regulator [*Propionibacteriales bacterium*]	92% / 69.42% / 1e-94	MPZ63410.1
GarR5 (HUW46_01936)	452	Sensor histidine kinase	sensor histidine kinase [*Propionibacteriales bacterium*]	87% / 51.87% / 1e-103	MPZ96507.1
GarD (HUW46_01935)	156	Putative TDP-4-keto-6-deoxy-D-glucose 3,4-isomerase	WxcM-like domain-containing protein [*Amycolatopsis antarctica*]	88% / 74.64% / 4e-71	WP_094863 117.1

Protein (NCBI accession)	a.a.	Proposed function	BLAST hit protein	Query coverage/ Per. identity/ E value	Accession number of BLAST hit
GarE (HUW46_01934)	367	Putative aminotransferase	DegT/DnrJ/EryC1/StrS family aminotransferase [Actinoalloteichus]	100% / 68.39% / 0.0	WP_075741 202.1
GarG5 (HUW46_01933)	513	Putative glycosyltransferase	glycosyltransferase, family 39 [*Streptomyces cinnamoneus*]	100% / 60.23% / 0.0	WP_099199 029.1
GarF (HUW46_01932)	450	Putative crotonyl-CoA reductase	crotonyl-CoA carboxylase/reductase [*Actinomadura macra*]	99% / 79.15% / 0.0	WP_067467 652.1
GarH (HUW46_01931)	338	Putative *fabH*	ketoacyl-ACP synthase III [*Micromonospora pisi*]	97% / 77.2 %/ 2e-175	WP_121159 999.1
GarI (HUW46_01930)	291	Putative 3-hydroxybutyryl-CoA dehydrogenase	3-hydroxybutyryl-CoA dehydrogenase [*Micromonospora pisi*]	92% / 74.72% / 3e-138	WP_121160 000.1
GarR6 (HUW46_01929)	956	Putative HTH transcriptional regulator	helix-turn-helix transcriptional regulator [*Kutzneria albida*]	97% / 49.89% / 0.0	WP_025354 824.1
Orf4 (HUW46_01928)	369	Putative IS4-like element	IS4-like element ISMfl1 family transposase [Mycobacteriaceae]	98% / 36.22% / 3e-52	WP_011891 513.1
Orf5 (HUW46_01927)	73	Hypothetical protein	hypothetical protein [*Streptomyces kasugaensis*]	100% / 60.27% /	WP_131126 254.1
GarR7 (HUW46_01926)	222	DNA-binding response regulator	DNA-binding response regulator, NarL/FixJ family, contains REC and HTH domains [*Sinosporangium album*]	98% / 62.39 / 1e-86	SDG85138.1
Orf6 (HUW46_01925)	370	Putative IS256 family element	IS256 family transposase [*Actinopolymorpha alba*]	66% / 30.26% / 5e-36	WP_040420 563.1

Table S4. Prediction of activity and stereochemistry of AT, KR, ER and DH domains from bioinformatics analysis. Stereochemical outcomes for gargantulides B and C.

Protein	Module	Substrate of AT domain	KR domain	DH domain	ER domain	Predicted stereochemistry (Keatinge-Clay)[a]	Predicted stereochemistry (fully assembled linear polyketides)[b]	Stereochemical outcome for gargantulides B and C[c]
GarP1	1	Mmal-CoA	B1	Active	R	"*R*" (C-68)	*R* (C-68)	*R* (C-68)
	2	Mal-CoA	B1	Active	R	-	-	-
GarP2	3	Mal-CoA	B1	-	-	"*R*" (C-65)	*S* (C-65)	*S* (C-65)
	4	Mal-CoA	B1	Active	S	-	-	-
	5	Mal-CoA	B1	-	-	"*R*" (C-61)	*R* (C-61)	*R* (C-61)
	6	Mal-CoA	B1	-	-	"*R*" (C-59)	*R* (C-59)	*S* (C-59)[d]
	7	Mmal-coA	A1	-	-	"*R,S*" (C-56,57)	*S,S* (C-56,57)	*R,S* (C-56,57)[e]
	8	Mal-CoA	B1	n.d.	-	"*R*" (C-55)	*S* (C-55)	*S* (C-55)
GarP3	9	Pmal-CoA[f]	B1	Active	R	"*R*" (C-52)	*R* (C-52)	*R* (C-52)
	10	Mmal-CoA	B1	-	-	"*R,R*" (C-50,51)	*S,R* (C-50,51)	*S,R* (C-50,51)
	11	Mal-CoA	B1	Inactive	-	"*R*" (C-49)	*S* (C-49)	*S* (C-49)
GarP4	12	Mal-CoA	B1	-	-	"*R*" (C-47)	*R* (C-47)	*S* (C-47)[g]
	13	Mmal-CoA	B2	-	-	"*S,R*" (C-44,45)	*S,R* (C-44,45)	*S,R* (C-44,45)
	14	Mmal-CoA	A1	-	-	"*R,S*" (C-42,43)	*S,R* (C-42,43)	*S,R* (C-42,43)
	15	Mal-CoA	A1	-	-	"*S*" (C-41)	*R* (C-41)	*R* (C-41)
GarP5	16	Mal-CoA	B1	-	-	"*R*" (C-39)	*S* (C-39)	*S* (C-39)
	17	Mal-CoA	A1	-	-	"*S*" (C-37)	*R* (C-37)	*R* (C-37)
	18	Mal-CoA	B1	-	-	"*R*" (C-35)	*S* (C-35)	*S* (C-35)
	19	Mal-CoA	A1	-	-	"*S*" (C-33)	*R* (C-33)	*R* (C-33)
	20	Mal-CoA	B1	Active	S	-	-	-
GarP6	21	Mal-CoA	B1	-	-	"*R*" (C-29)	*R* (C-29)	*R* (C-29)
	22	Mmal-CoA	A1	-	-	"*R,S*" (C-26,27)	*S,S* (C-26,27)	*R,S* (C-26,27)[h]
	23	Mal-CoA	B1	-	-	"*R*" (C-25)	*S* (C-25)	*R* (C-25)[h]
GarP7	24	Mmal-CoA	B1	Active	-	"*E*" double bond	*E* double bond	*E* double bond
	25	Mal-CoA	B1	-	-	"*R*" (C-21)	*S* (C-21)	*S* (C-21)
	26	Mmal-CoA	C1[i]	n.d.	n.d.	"*R*" (C-18)	*S* (C-18)	*R* (C-18)[i]
GarP8	27	Mal-CoA	A1	-	-	"*S*" (C-17)	*R* (C-17)	*R* (C-17)
	28	Mal-CoA	B1	Active	S	-	-	-
GarP9	29	Mmal-CoA	B1	Active	S	"*S*" (C-12)	*S* (C-12)	*S* (C-12)
	30	Mmal-CoA	A1	-	-	"*R,S*" (C-10,11)	*S,S* (C-10,11)	*S,S* (C-10,11)
	31	Mal-CoA	B1	Active	S	-	-	-
GarP10	32	Mmal-CoA	A1	-	-	"*R,S*" (C-6,7)	*S,S* (C-6,7)	*S,S* (C-6,7)
	33	Mal-CoA	B1	Active	S	-	-	-
	34	Mal-CoA	B1	Active	S	-	-	-

[a] Bioinformatics prediction for each separate module.

[b] See Fig. S10 (fully assembled linear polyketides).

[c] Stereochemistry determined by a combination of NMR and bioinformatics gene cluster analysis for gargantulides B and C (and extrapolated to gargantulide A, revised in this work).

[d] The Cahn-Ingold-Prelog descriptor in gargantulides B and C for C-59 is inverted to *S* due to the glycosylation at C-61.

[e] The Cahn-Ingold-Prelog descriptor in gargantulides A-C for C-56 is inverted to *R* due to the glycosylation at C-55.

[f] Propyl malonyl-CoA established by NMR analysis

[g] The Cahn-Ingold-Prelog descriptor in gargantulides A-C for C-47 is inverted to *S* due to macrolactone cyclization and glycosylation at C-49.

[h] The Cahn-Ingold-Prelog descriptor in gargantulides A-C for both C-25 and C-26 is inverted to *R* due to the post-PKS hydroxylation at C-24.

[i] Redox-inactive, methyl-epimerizing KR; (module 26).

Figure S8. Extracted amino acid sequence alignments of AT, KR, DH and ER domains.
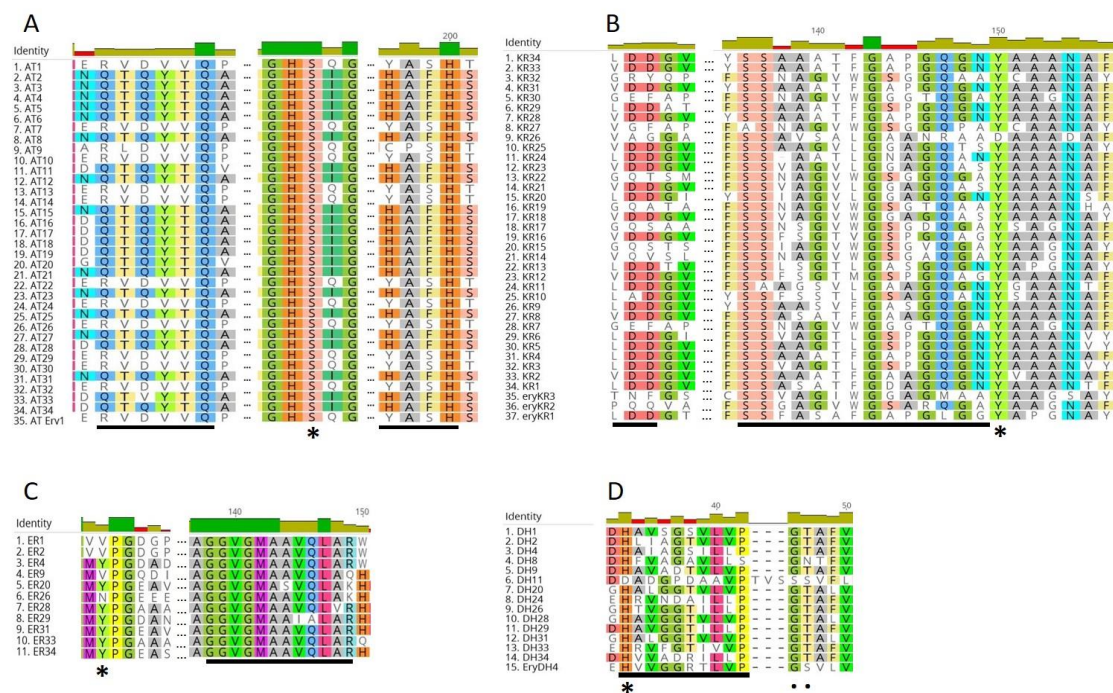
A - amino acid sequence alignment of AT domains. Characteristic residues for substrate recognition are underlined with black bars. The catalytic residue is identified with asterisk; **B** - amino acid sequence alignment of KR domains. Two motifs, responsible for determination of A/B/C type of KRs are underlined with black bars; the catalytic tyrosine residue is identified with an asterisk; **C** - amino acid sequence alignment of ER domains. The amino acid related to the stereochemistry of reduced residue is identified with an asterisk. NADPH binding site is identified with a black bar; **D** - amino acid sequence alignment of DH domains. The conserved motif "HXXXGXXXXP" is identified with a black bar. The catalytic residue is identified with an asterisk. Two supporting catalytic residues are identified with filled black spots. The assignment of characteristic residues for substrate recognition and catalytic residues for ATs, KRs, DHs and ERs was done as described in previous works.[313–316]

Supplementary note: Biosynthesis of amino sugar maG

Within the *gar* BGC, *garD* codes for a protein with a high level of similarity (51% identity, 64% similarity) to Tyl1a from the mycaminose biosynthetic pathway in the tylosin BGC,[319,320] suggesting it could function as a TDP-4-keto-6-deoxy-D-glucose 3,4-isomerase. *garE* encodes a PLP-dependent aminotransferase with high homology (62% identity, 73% similarity) to the biochemically characterized aminotransferase TylB from mycaminose biosynthesis.[321] Thus, GarE is proposed to catalyze a C-3 transamination to afford TDP-3-amino-3,6-dideoxy-D-glucose. Surprisingly, the *gar* cluster lacks a gene coding for a putative *N*-methyltransferase to convert TDP-3-amino-3,6-dideoxy-D-glucose into *N*-demethyl-mycaminose (i.e., maG). To look for a candidate gene responsible for the *N*-methylation, we employed the amino acid sequences of some known methyltransferases found in the mycaminose, desosamine, L-ossamine, and D-forosamine pathways from different microorganisms. The best candidate was HUW46_03188, located just outside of the boundaries of the BGC 15, which encodes for a putative S-adenosylmethionine-dependent methyltransferase with high similarity (51% / 65% amino acid identity / similarity) to DesVI, the biochemically characterized *N,N*-dimethyltransferase from desosamine biosynthesis[342] (Table S5). Interestingly, other DesVI homologous proteins, such as TylM1 (tylosin gene cluster) or SpnS (spinosyns BGC), have been shown to catalyze *N*-methylation in a stepwise manner, involving the release of the monomethylated product from the active site.[342,343] This could explain the formation of the monomethylated amino sugar maG in gargantulides A-C. Additionally, genes encoding for a putative glucose-1-phosphate thymidyltransferase and dTDP-glucose 4,6-dehydratase, required in the first two steps to afford the intermediate TDP-4-keto-2,6-dideoxy-D-glucose, are also found elsewhere in the genome (Table S6). Based on these findings, a biosynthetic pathway to maG from glucose-1-phosphate can be proposed (Fig. S9).

Table S5. Levels of identity and similarity of the putative *N*-Methyltransferase HUW46_03188 from CA-230715 with other known *N*-Methyltransferases.

Protein	Microorganism	% Identity to HUW46_03188	% Similarity to HUW46_03188
TylM1	*Streptomyces fradiae*	43	61
DesVI	*Streptomyces venezuelae*	51	65
SpnS	*Saccharopolyspora spinosa*	42	58
OssI	*Streptomyces ossamyceticus*	43	59
Orf1C	*Streptomyces ambofaciens*	44	62
Orf9c	*Streptomyces ambofaciens*	44	57

Table S6. Identified genes encoding for putative glucose-1-phosphate thymidyltransferase and dTDP-glucose 4,6-dehydratase in the genome of CA-230715.

CA-230715 ORF	Protein a. a.	Putative function	Best match accession number	% Identity / Similarity
HUW46_03193	293	glucose-1-phosphate thymidylyltransferase	WP_020668504.1	87 / 93
HUW46_02674	330	dTDP-glucose 4,6-dehydratase	WP_144591396.1	88 / 92
HUW46_09296	335	dTDP-glucose 4,6-dehydratase	WP_125677294.1	72 / 83

Figure S9. Proposed biosynthetic pathway for the amino sugar 3,6-deoxy-3-methylamino D-glucose (maG).
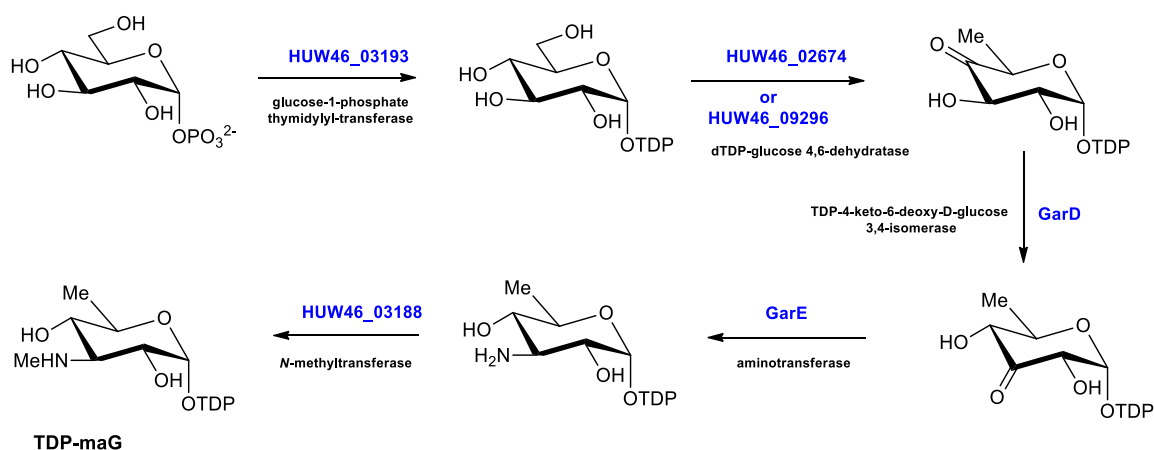
Figure S10. Bioinformatics assignments (*gar* BGC analysis) of the absolute configurations for the gargantulides polyketide aglycon (**a**). Comparison with the absolute configurations previously reported for gargantulide A (**b**).[307] KR domains determining the stereochemical outcome is indicated in blue color. Chiral centers showing disagreement between the *in-silico* prediction and the previously reported NMR-based structure for gargantulide A, are highlighted in red color. Undetermined chiral centers for gargantulide A in the original work, which has been now assigned by bioinformatic analysis, are highlighted in pink color.

(a)

(b)

Org. Chem. Front., **2022**, 9, 462-470
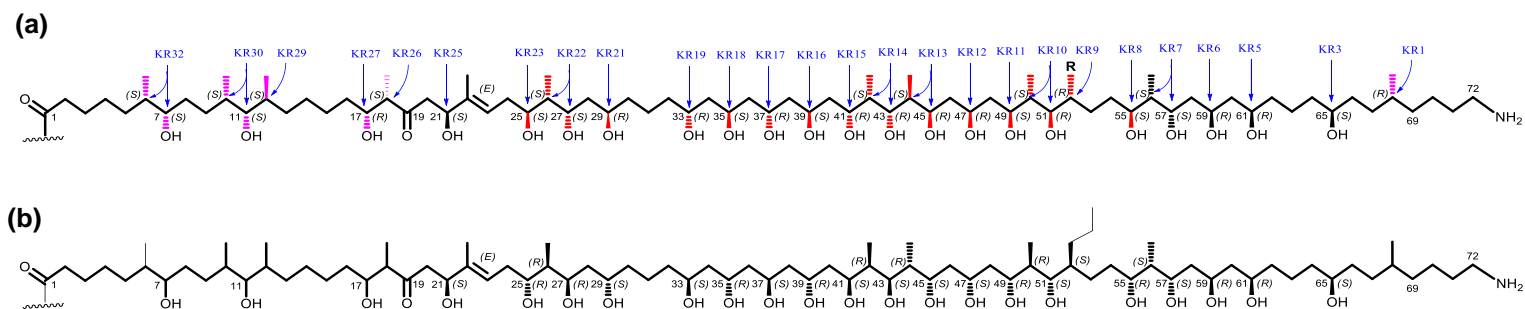
Figure S11. Determination of the relative configuration of the C-55 to C-57 stereocluster for gargantulide B (**1**).

NMR data set I

syn/syn   7.1
syn/anti   10.7
anti/syn   10.7
anti/anti   11.6

} ppm

10.7 ppm

R= maG
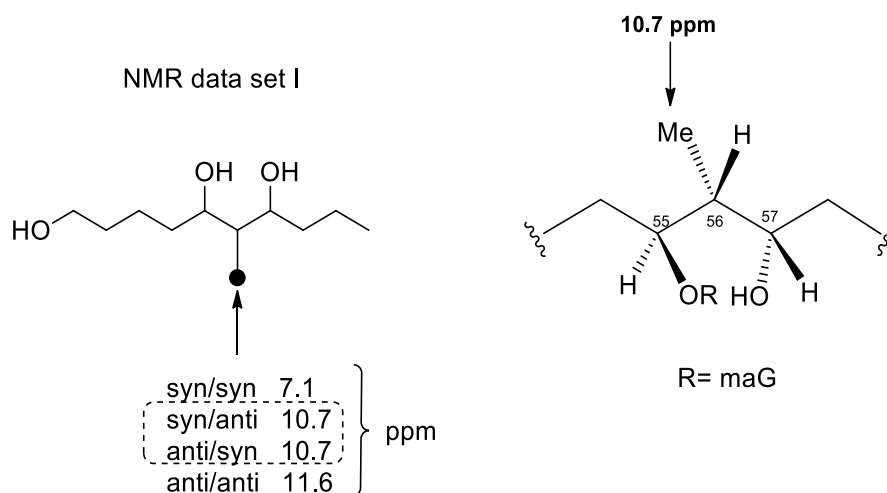
a) Kishi´s NMR data set I for 2-methyl-1,3-diols[327] (left) and chemical shift value of Me-56 (right).

$^3J_{H55-H56}$ = ca. 9 Hz

$^3J_{C54-H56}$ = ca. small

$^3J_{C57-H55}$ = ca. small

$^3J_{Me56-H55}$ = ca. small

$^2J_{C56-H55}$ = ca. small

$^2J_{C55-H56}$ = ca. large

$^3J_{H56-H57}$ = ca. 2 Hz

$^3J_{C55-H57}$ = ca. small

$^3J_{C58-H56}$ = ca. small

$^3J_{Me56-H57}$ = ca. large
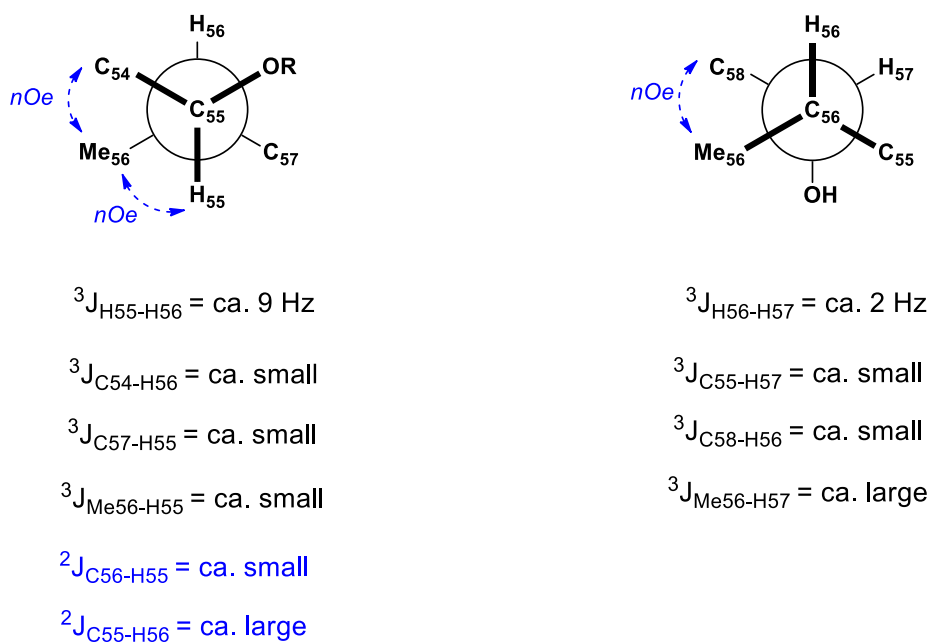
b) Qualitative $^3J_{C,H}$-based configuration analysis of the C-55 to C-57 stereocluster. Other $^2J_{C,H}$ constants are classified as "small" or large" according to Murata´s method.[325] Key NOESY correlations are also indicated (NOESY correlation Me-56/H-58 partially overlapped with other *nOe* cross-peaks).
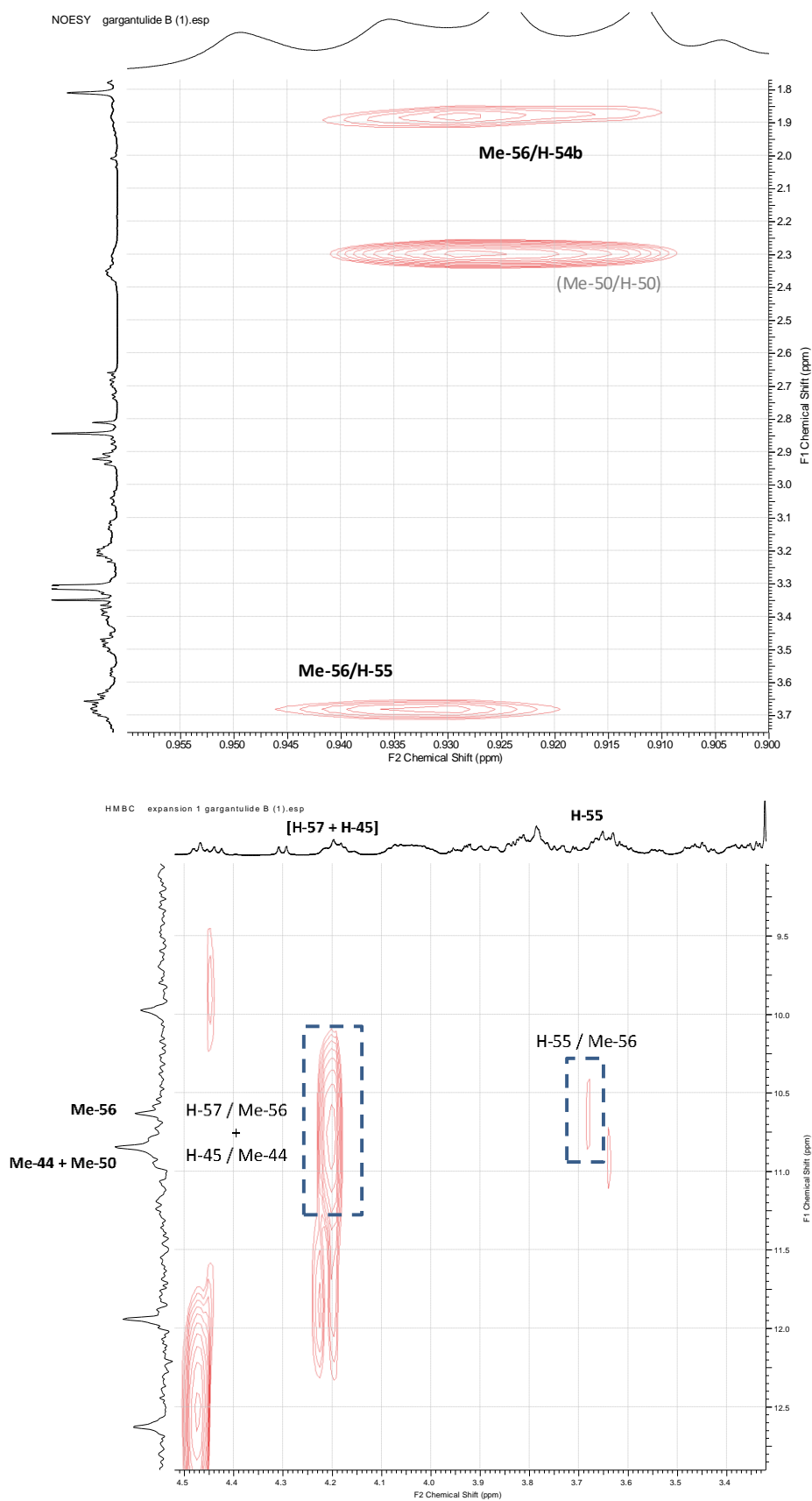
c) NOESY (top) and HMBC (bottom) expansion supporting the *anti/syn* configuration for C-55/C-56/C-57.
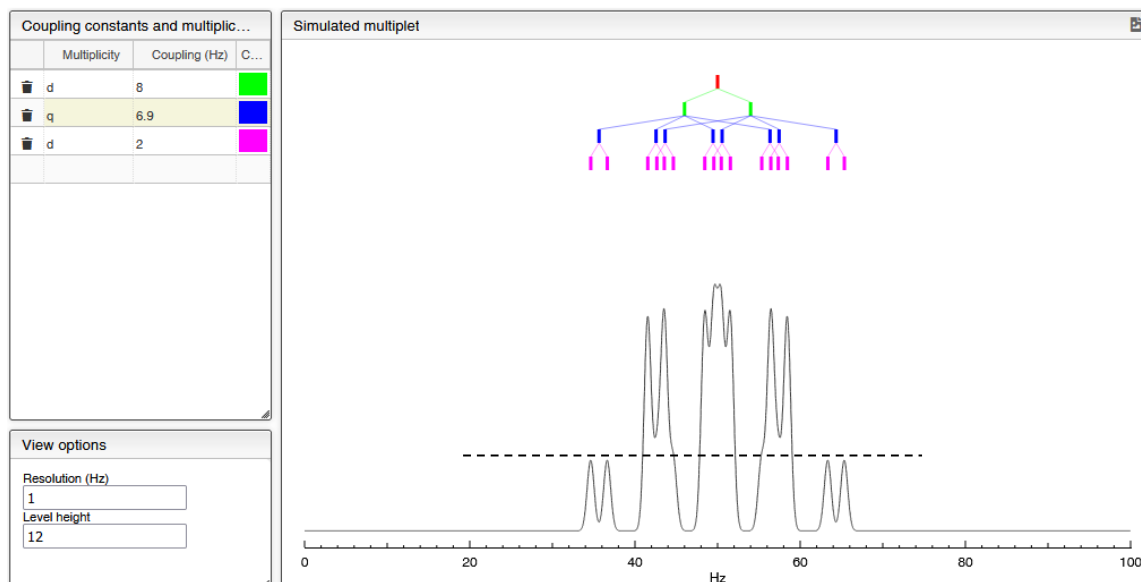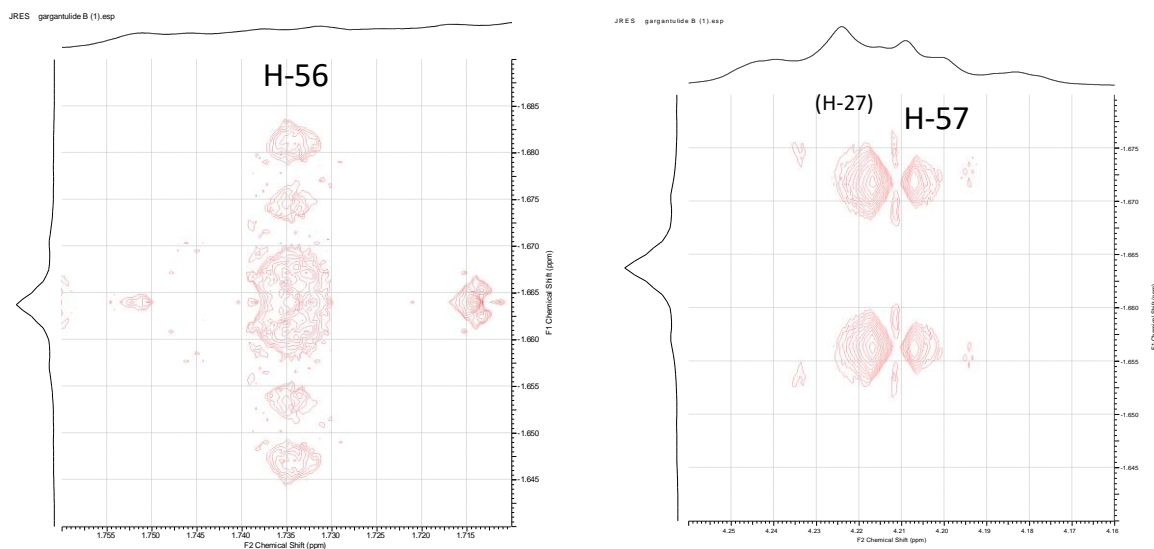
d) Expansions of the JRES spectrum of **1** showing the higher multiplicity of H-56 compared to that of H-57 (top). Multiplet simulation[344] for H-56 (applied values for ^{1}H-^{1}H coupling constants are indicated) showing good fitting with the multiplicity experimentally observed (bottom).

Figure S12. Determination of the relative configuration of the C-6–C-7 stereocluster for gargantulide B (**1**).
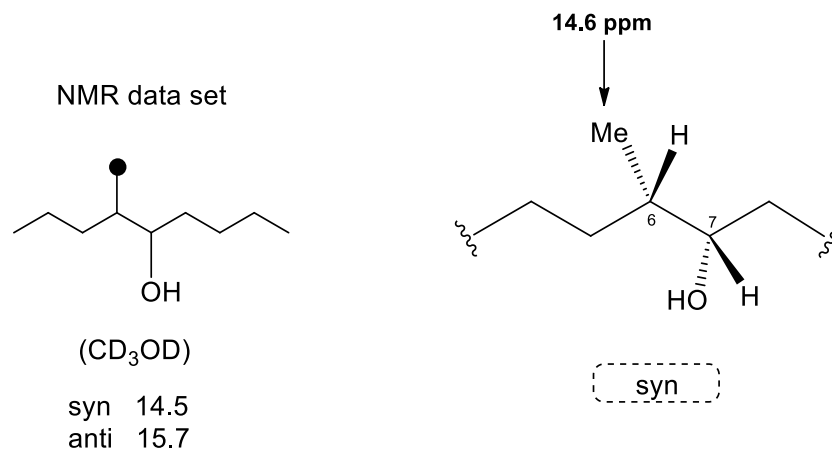
Application of Kishi´s NMR data set for 4-methylnonan-5-ol isomers.[336]

Figure S13. Determination of the relative configuration of the C-10–C-12 stereocluster for gargantulide B (**1**).
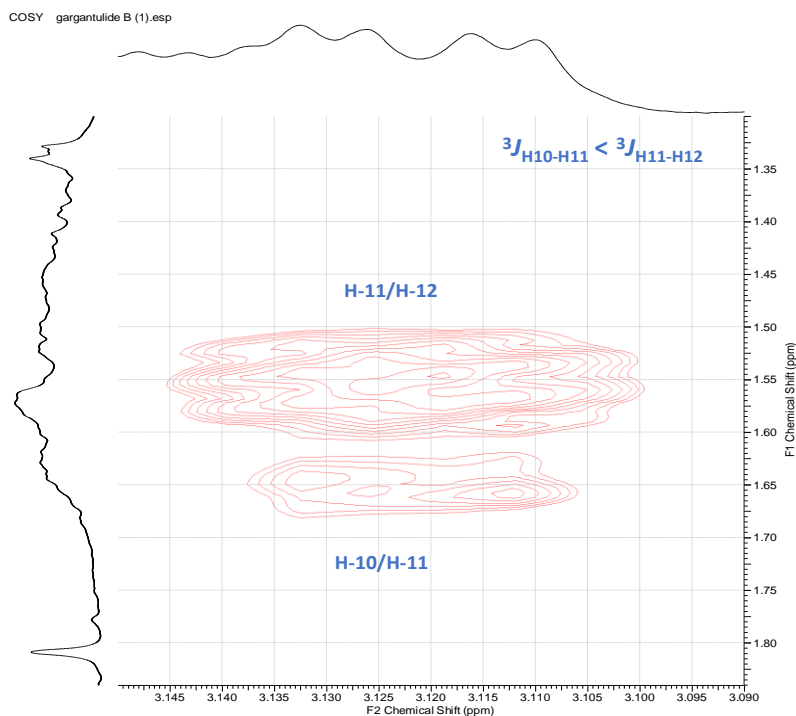
a) Key COSY correlations supporting the *syn/anti* configuration for C-10/C-11/C-12.

$^3J_{H10-H11}$ = 3.2 Hz

$^3J_{Me10-H11}$ = ca. large

$^3J_{C9-H11}$ = ca. small

$^3J_{C12-H10}$ = ca. small

$^3J_{H11-H12}$ = 8.2 Hz

$^3J_{Me12-H11}$ = ca. small

$^3J_{C10-H12}$ = ca. small

$^3J_{C13-H11}$ = ca. small
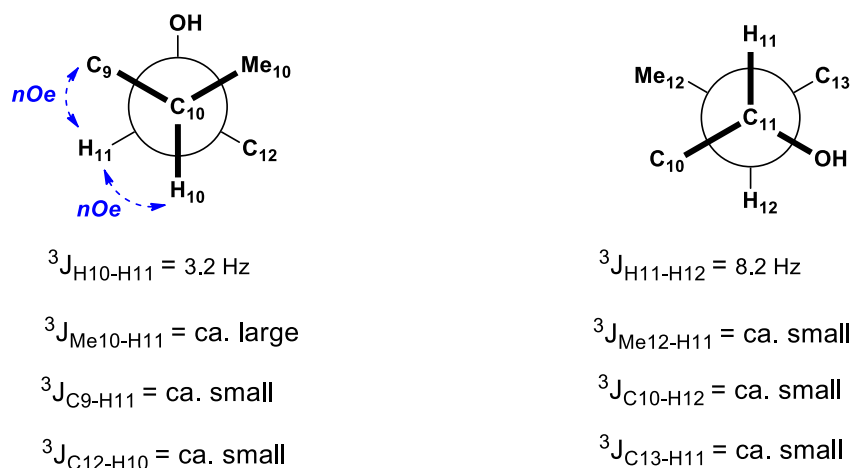
b) Qualitative $^3J_{C,H}$-based configuration analysis of the C-10 to C-12 stereocluster. Key NOESY correlations are also indicated (although critical overlapping of proton NMR signals corresponding to Me-10 and Me-12 make impossible to distinguish the key NOESY correlation H-10/Me-12 from the genuine cross-peak H-10/Me-10, the clear absence of NOESY correlations between H-10 and both H-13 further supports the configurational proposal).
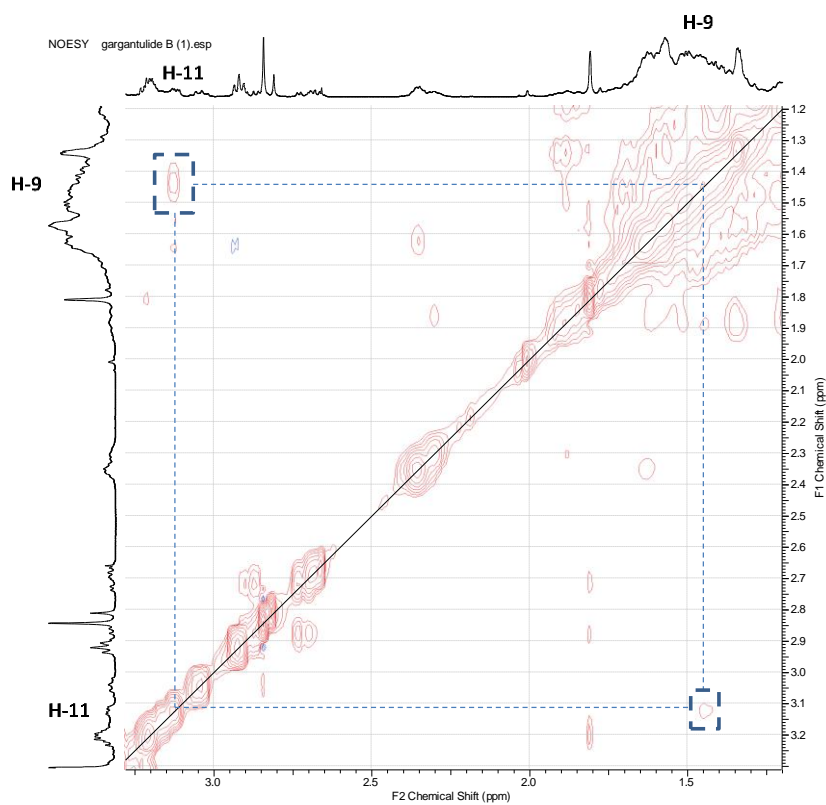
c) Key NOESY correlation supporting the *anti* configuration for C-10–C-12 stereocluster.
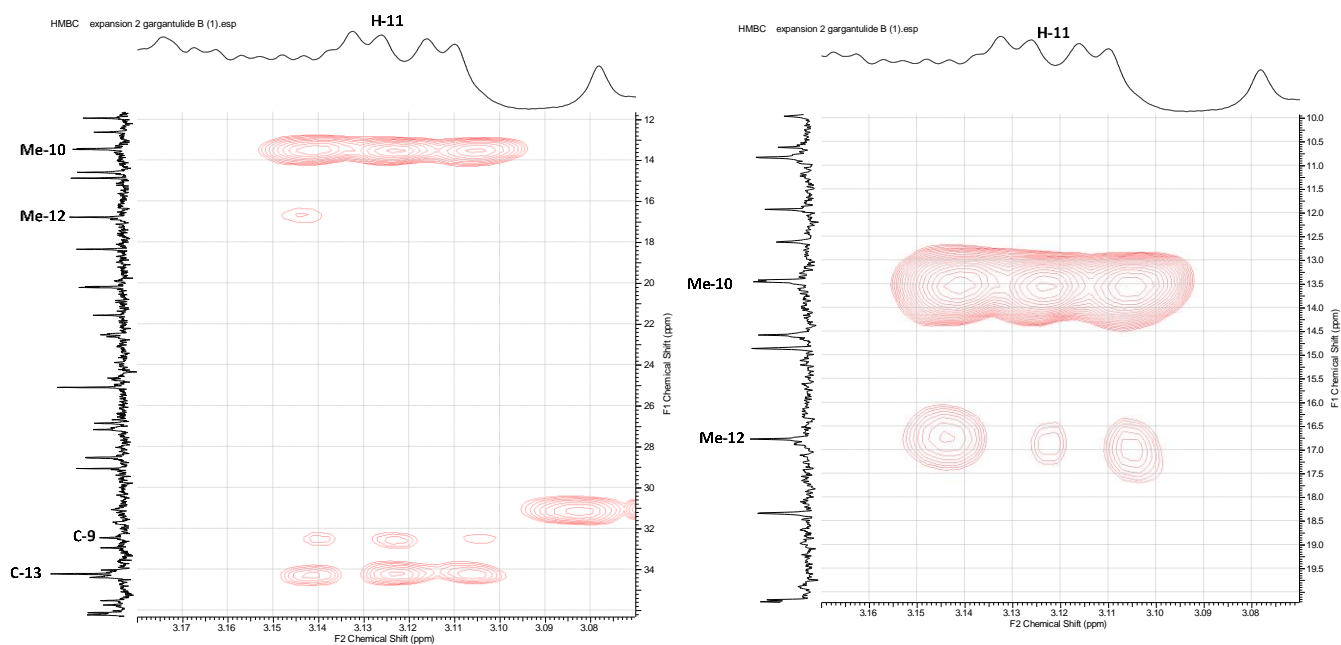
d) HMBC correlations supporting the *anti* configuration for C-10–C-12 stereocluster.

Figure S14. Determination of the relative configuration of the C-17(R)–C-18(R) stereocluster for gargantulide B (**1**).

anti relative configuration

17(R), 18(R) abs. configuration

$^3J_{H17-H18}$ = 8.6 Hz

$^3J_{Me18-H17}$ = ca. small

$^3J_{C16-H18}$ = ca. small

$^3J_{C19-H17}$ = ca. small
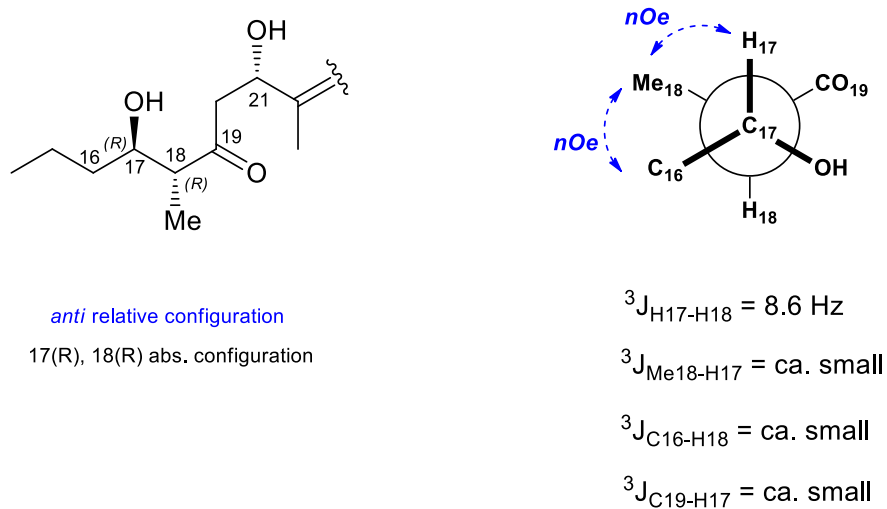
a) Qualitative $^3J_{C,H}$-based configuration analysis of the C-17–C-18 stereocluster.
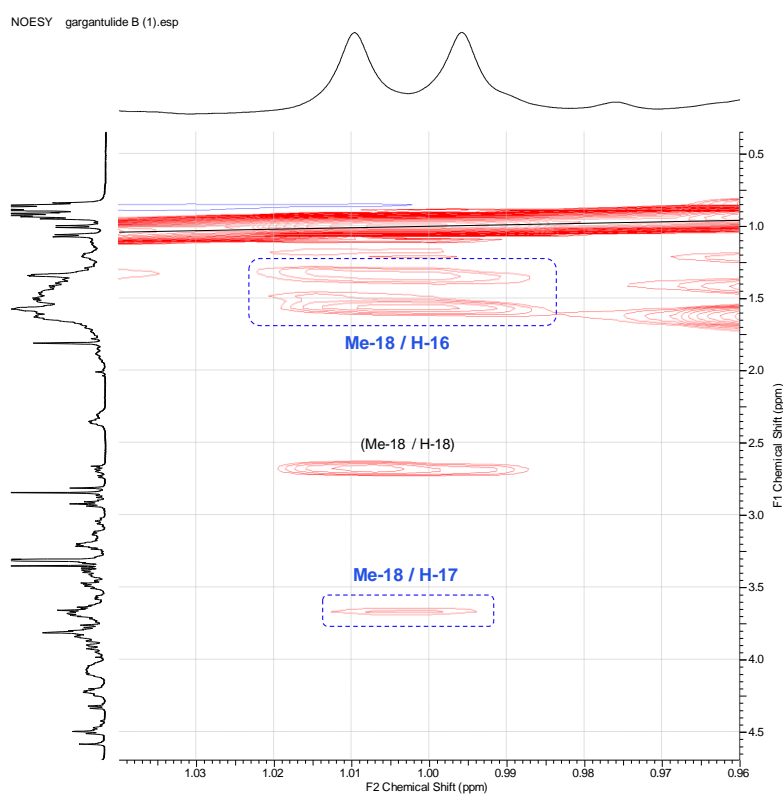
b) Key NOESY correlations supporting the anti-configuration for C-17–C-18 stereocluster.

Figure S15. Determination of the absolute configuration of the non-amino sugars present in gargantulides B (**1**) and C (**2**).
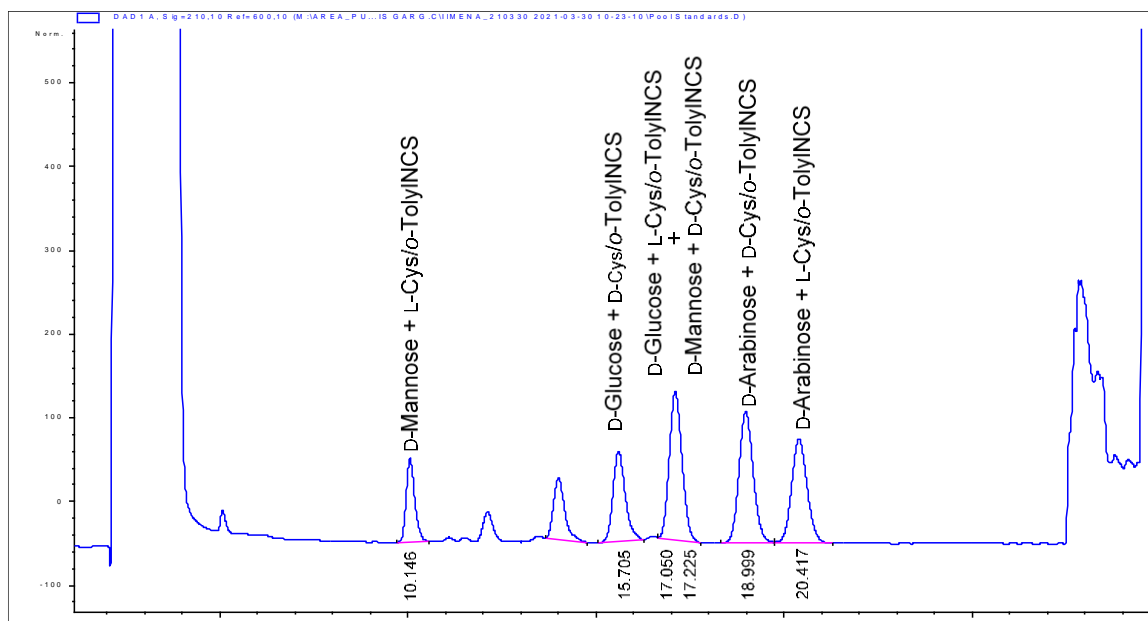
a) LC-UV chromatogram of a mixture of both ʟ- and ᴅ-cysteine methyl ester hydrochloride / *o*-tolyl isothiocyanate derivatization reactions of ᴅ-Mannose / ᴅ-Glucose / ᴅ-arabinose standard monosaccharides.
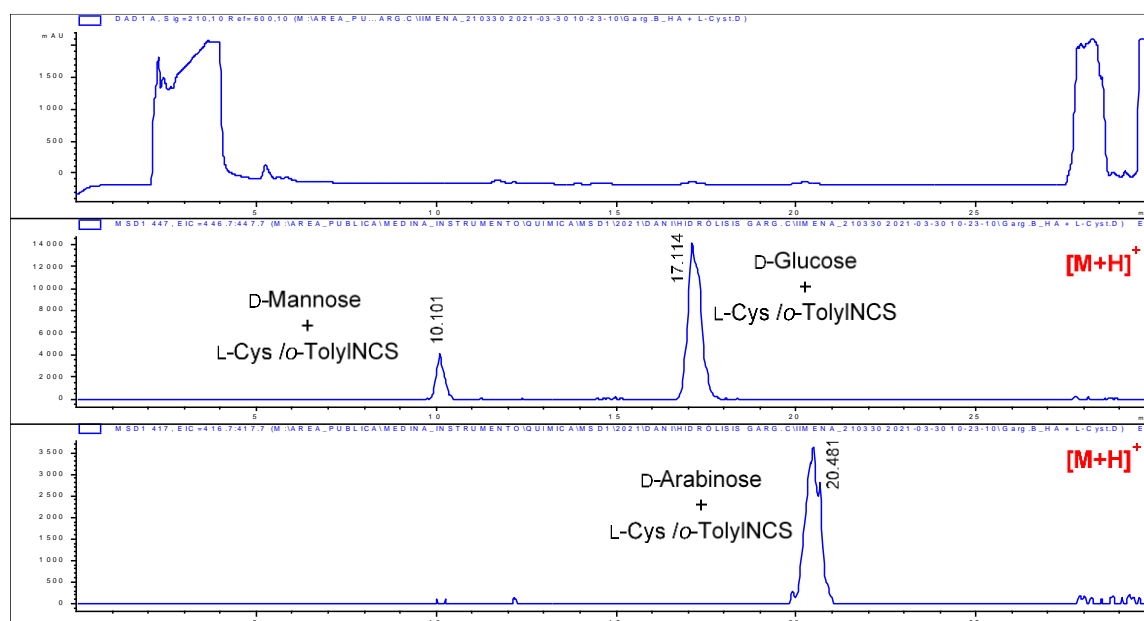
b) LC-UV and extracted-ion chromatograms (EIC) of ʟ-cysteine methyl ester hydrochloride / *o*-tolyl isothiocyanate derivatized hydrolysate of **1** with HCl. ᴅ-Mannose, ᴅ-Glucose and ᴅ-Arabinose were found in **1**.
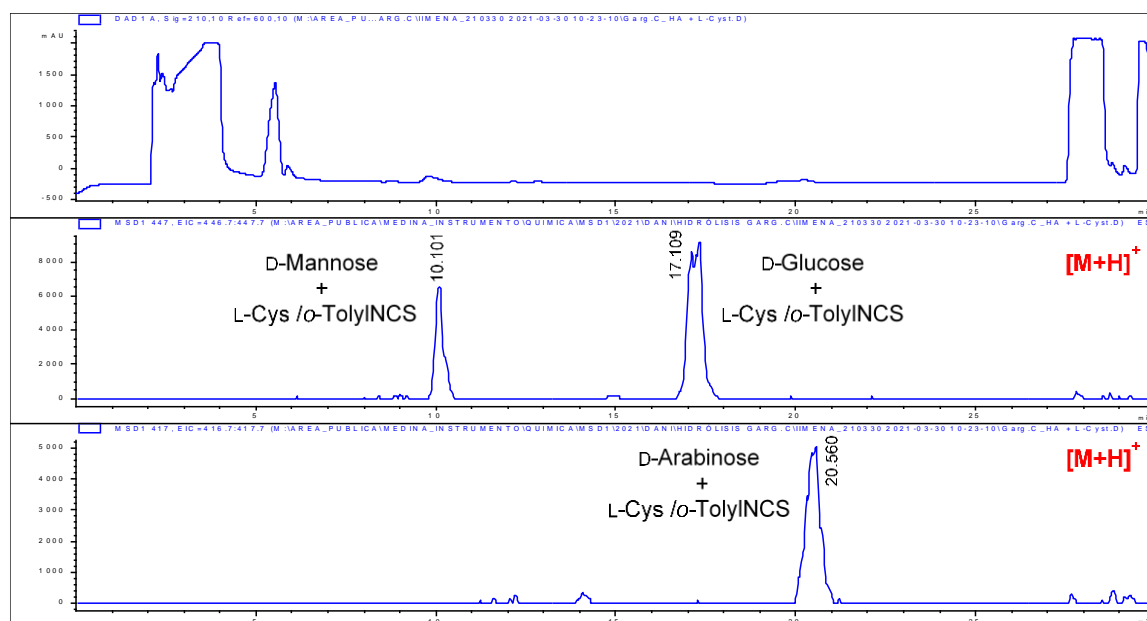
c) LC-UV and extracted-ion chromatograms (EIC) of L-cysteine methyl ester hydrochloride / *o*-tolyl isothiocyanate derivatized hydrolysate of **2** with HCl. D-Mannose, D-Glucose and D-Arabinose were found in **2.**

Table S7. Antibacterial and antifungal activities of compounds **1** and **2**.

Microbial strain	Strain number	MIC (µg/mL)					
		(1)	(2)	R	A	Am	V
A. baumannii	MB5973	16-32	16-32	2-4			
P. aeruginosa	MB5919	>128	>128		1		
E. coli	ATCC 25922	>128	>128		0.5-0.25		
K. pneumoniae	ATCC 700603	>128	>128		>16		
MRSA		4-8	2-4				2-4
MSSA		2-4	2-4				1
VRE		2-4	1-2				>128
A. fumigatus	ATCC46645	32-64	32-64			4	
C. albicans	ATCC64124	>128	>128			4	

*Positive controls: R Rifampicin, A Aztreonam, Am Amphotericin B, V Vancomycin.

CONCLUSIONS

Conclusions

I.   As a result of this research work, we have isolated and characterized the structures of nine new secondary metabolites from the MEDINA´s actinomycetes collection: four meroterpenoids, three nonribosomal peptides and two glycosylated macrolactones (type I polyketides). These compounds expand the chemical space of the three NP classes they belong to, since they incorporate in their structures novel features like unprecedented cyclization patterns, rare amino acids, or amino sugars, or an exceptionally complex polyol scaffold harboring multiple stereocenters.

II. We have reported the isolation and structural characterization of four new napyradiomycins (napyradiomycin A3, napyradiomycin B7a, napyradiomycin B7b and napyradiomycin D1) along with the non-previously described structural details of the known napyradiomycin SC. Additionally, ten other known analogue compounds were also isolated from the same culture broth of the marine-derived *Streptomyces* sp. CA-271078 from MEDINA's microbial collection. The new napyradiomycins isolated displayed novel structural features:

   a. Napyradiomycin A3 is the first member of napyradiomycins in the A series bearing a hydroxy group rather than a chlorine atom at C-3.

   b. Napyradiomycin B7b represents the first example of a B-type napyradiomycin reported with a different relative configuration at the C-3 chlorinated position. A plausible pathway to this compound involving nucleophilic displacement from the congener B7a or MDN-0170 is proposed. This alternative stereochemistry at C-3 seems to impair the antibacterial activity against MRSA and *M. tuberculosis* and the cytotoxicity against the HpeG2 tumoral cell line.

   c. Remarkably, napyradiomycin D1 harbors in its structure an unprecedented 14-membered macrocyclic ether ring between the naphthoquinone moiety and the monoterpenoid chain, thus inaugurating a new subfamily of napyradiomycins, the D series.

III. The bioactivity profiling of all the napyradiomycins isolated in this work revealed new napyradiomycin D1 and napyradiomycins B2, B4, and B5 as the most active compounds, exhibiting similar antibacterial activities against MRSA and *M. tuberculosis* H37Ra, as well as comparable cytotoxic activities against the HepG2 tumoral cell line.

IV. The whole genome sequencing of the producing strain CA-271078 revealed the presence of four putative VCPOs, instead of the three found in other previously described napyradiomycin BGCs, along with an additional FAD-dependent oxidoreductase also absent in the latter. This finding suggests the possible involvement of such enzymes in the macrocyclization of the monoterpenoid chain that would lead to napyradiomycin D1, although further biochemical characterization is needed to validate this hypothesis.

V.   Applying different culture conditions to the strain *Streptomyces cacaoi* subsp. *cacaoi* CA-170360 we have elicited the production of three new pentaminomycins F-H, along with the known pentaminomycins A-E. These compounds are novel *N*-hydroxyarginine-containing cyclic pentapeptides with a LDLLD chiral sequence.

VI.   The analysis of the available genome sequence identified the BGC responsible for producing both the pentaminomycins and BE-18257 cyclic peptides in the genome of the producing strain, *Streptomyces cacaoi* CA-170360. In a parallel work carried out by the MEDINA team, this BGC was heterologously expressed and unambiguously linked to the biosynthesis of both subfamilies of nonribosomal peptides.

VII.   Pentaminomycins F and G, newly discovered in this work, are the first nonribosomal peptides incorporating in their structures the rare amino acid 3-(2-pyridyl)-alanine (2-L-Pal). The production of these two novel compounds indicates that the amino acid 2-L-Pal is recruited by the adenylation domain of the fifth NRPS module and incorporated into the growing peptide chain. This flexibility or substrate tolerance is further reflected by the different amino acid residues found at this position in the different pentaminomycins and opens the door to obtaining new-to-nature analogue compounds by feeding with synthetic amino acids.

VIII.   The antimicrobial activity for this family of pentapeptides was evaluated against a panel of Gram-negative pathogens for the first time, with pentaminomycin B and C showing low but selective activity against a clinical isolate of *Acinetobacter baumannii*.

IX.   Gargantulides B and C, two exceptionally complex 52-membered macrolactones, were isolated from the strain *Amycolatopsis* sp. CA-230715 in the context of an antibacterial screening campaign. The structures of these huge macrolides were fully characterized by combining 2D NMR, genome analysis, and chemical approaches.

X.   The extraordinarily large 216 kbp biosynthetic gene cluster (*gar* BGC) encodes 10 T1PKS enzymes comprising 35 modules, thus representing the largest macrolide-encoding BGC discovered so far. The analysis of the *gar* gene cluster allowed us to establish the first genetic evidence for the biosynthesis of gargantulides.

XI.   Gargantulides A-C contain the rare amino sugar 3,6-deoxy-3-methyl aminoglucose (maG), which, as far as we know, is present exclusively in this family of macrolactones. The analysis of the *gar* gene cluster allowed us to establish for this sugar a common biosynthetic origin with mycaminose, and therefore to propose its absolute configuration as D.

XII.   The combination of NMR spectroscopy and *in silico* analysis of the *gar* gene cluster allowed to assign the absolute configuration for all chiral centers in gargantulides B and C. Based on the co-detection of the previously reported congener gargantulide A and the high structural similarity between the three macrolides, we propose a common biosynthetic origin and the same stereochemistry for all of them. This allowed us to revise the stereochemical assignments originally proposed for gargantulide A, including that of the amino sugar maG, mainly based on NMR studies.

XIII.   The new macrolactones, gargantulides B and C, exhibited potent antibacterial activity against the Gram-positive bacteria MRSA and VRE, and moderate but unusual activity against the Gram-negative pathogen *A. baumannii*.

Conclusiones

I. Como resultado de este trabajo de investigación, hemos aislado y caracterizado las estructuras de nueve nuevos metabolitos secundarios a partir de la colección de actinomicetos de MEDINA: cuatro meroterpenoides, tres péptidos no ribosomales y dos macrolactonas glicosiladas (policétidos tipo I). Estos compuestos amplían el espacio químico de las tres clases de productos naturales a las que pertenecen, ya que presentan características estructurales novedosas, como patrones de ciclización sin precedentes, aminoácidos inusuales o aminoazúcares, o una estructura de poliol excepcionalmente compleja que incorpora múltiples centros quirales.

II. Hemos reportado el aislamiento y la caracterización estructural de cuatro nuevas napiradiomicinas (napiradiomicina A3, napiradiomicina B7a, napiradiomicina B7b y napiradiomicina D1), junto con los detalles estructurales de la napiradiomicina SC previamente conocida. Además, se aislaron también otros diez compuestos análogos conocidos del mismo caldo de cultivo de *Streptomyces* sp. CA-271078 de la colección microbiana de MEDINA. Las nuevas napiradiomicinas aisladas mostraron características estructurales novedosas:

   a. Napiradiomicina A3 es el primer miembro de napiradiomicinas en la serie A con un grupo hidroxilo en lugar de un átomo de cloro en C-3.

   b. Napiradiomicina B7b representa el primer ejemplo de una napiradiomicina de tipo B con una configuración relativa diferente en la posición clorada C-3. Se propone una vía plausible para este compuesto que involucra una sustitución nucleofílica a partir del congénere B7a o MDN-0170. Esta estereoquímica alternativa en C-3 parece afectar la actividad antibacteriana contra MRSA y *M. tuberculosis* y la citotoxicidad contra la línea celular tumoral HpeG2.

   c. A destacar, la napiradiomicina D1 alberga en su estructura un anillo éter macrocíclico de 14 miembros sin precedentes entre el residuo naftoquinona y la cadena monoterpenoide, inaugurando así una nueva subfamilia de napiradiomicinas, la serie D.

III. La secuenciación completa del genoma de la cepa productora CA-271078 reveló la presencia de cuatro posibles enzimas VCPOs, en lugar de las tres encontradas en los clústeres de genes biosintéticos de napiradiomicinas descritos previamente. Este hallazgo sugiere la posible implicación de dichas enzimas en la macrociclación de la cadena monoterpenoide que conduciría a napiradiomicina D1, aunque es necesaria su caracterización bioquímica para validar esta hipótesis.

IV. El perfil de bioactividad de todas las napiradiomicinas aisladas en este trabajo reveló que la napiradiomicina D1 y las napiradiomicinas B2, B4 y B5 son los compuestos más activos, mostrando actividades antibacterianas similares contra MRSA y M. tuberculosis H37Ra, así como actividades citotóxicas comparables frente a la línea celular tumoral HepG2.

V. Aplicando diferentes condiciones de cultivo a la cepa *Streptomyces cacaoi* subsp. *cacaoi* CA-170360 se ha inducido la producción de tres nuevas pentaminomicinas F-H, junto con las pentaminomicinas A-E conocidas. Estos compuestos son nuevos pentapéptidos cíclicos que contienen N-hidroxiarginina con una secuencia quiral LDLLD.

VI.   El análisis de la secuencia genómica disponible identificó el grupo de genes responsables de producir tanto las pentaminomicinas como los péptidos cíclicos BE-18257 en el genoma de la cepa productora, *Streptomyces cacaoi* CA-170360. En un trabajo paralelo realizado por el equipo de MEDINA, este grupo de genes fue expresado heterólogamente y vinculado de manera inequívoca a la biosíntesis de ambas subfamilias de péptidos no ribosomales.

VII.   Las pentaminomicinas F y G, descubiertas recientemente en este trabajo de tesis, son los primeros péptidos no ribosomales que incorporan en sus estructuras el inusual aminoácido 3-(2-piridil)-alanina (2-L-Pal). La producción de estos dos nuevos compuestos indica que el aminoácido 2-L-Pal es reclutado por el dominio de adenilación del quinto módulo NRPS e incorporado a la cadena peptídica en crecimiento. Esta flexibilidad o tolerancia a sustratos se refleja además en los diferentes residuos de aminoácidos encontrados en esta posición en las diferentes pentaminomicinas y abre la puerta para obtener nuevos compuestos análogos mediante la adición de aminoácidos sintéticos.

VIII. La actividad antimicrobiana de esta familia de pentapéptidos se evaluó contra un panel de patógenos Gram-negativos por primera vez, mostrando para pentaminomicina B y C una actividad baja pero selectiva frente a un aislado clínico de *A. baumannii*.

IX.   Las gargantulidas B y C, dos macrolactonas de 52 miembros excepcionalmente complejas, se aislaron de la cepa *Amycolatopsis* sp. CA-230715 en el contexto de una campaña de cribado antibacteriano. Las estructuras de estos enormes macrólidos fueron totalmente caracterizadas mediante la combinación de RMN en 2D, análisis genómico y enfoques químicos.

X.   El extraordinariamente grande grupo de genes biosintético de 216 kpb (*gar* BGC) codifica 10 enzimas T1PKS que comprenden 35 módulos, lo que representa el grupo de genes de codificación de macrólidos más grande descubierto hasta la fecha. El análisis del grupo de genes *gar* nos permitió establecer la primera evidencia genética de la biosíntesis de las gargantulidas.

XI.   Las gargantulidas A-C contienen el inusual aminoazúcar 3,6-desoxi-3-metilaminoglucosa (maG), que, hasta donde sabemos, está presente exclusivamente en esta familia de macrolactonas. El análisis del grupo de genes *gar* nos permitió establecer para este azúcar un origen biosintético común con la micaminosa y, por lo tanto, proponer su configuración absoluta como D.

XII.   La combinación de espectroscopia de RMN y análisis *in silico* del grupo de genes *gar* permitió asignar la configuración absoluta para todos los centros quirales en las gargantulidas B y C. Basándonos en la co-detección del congénere previamente reportado, gargantulida A, y en la alta similitud estructural entre los tres macrólidos, proponemos un origen biosintético común y la misma estereoquímica para todos ellos. Esto nos permitió revisar las asignaciones esterequímicas originalmente propuestas para gargantulida A, incluido el aminoazúcar maG, principalmente basadas en estudios de RMN.

XIII. Las nuevas macrolactonas gargantulidas B y C, mostraron una potente actividad antibacteriana contra las bacterias Gram-positivas MRSA y VRE, y una actividad moderada pero inusual contra el patógeno Gram-negativo *A. baumannii*.

EPILOGUE

Epilogue

The extensive range of chemical scaffolds found in natural products remains highly relevant for the discovery of new drugs. Methodological approaches applied to drug discovery are nowadays more efficient, sophisticated, and precise than those used in the golden era of NP, largely due to advances in analytical chemistry and the development of methods to capture their inherent value. Progress in technologies for isolating and identifying bioactive compounds, instruments facilitating genetic data analysis, bioinformatics prediction of chemical structures, and fusion of datasets encompassing varied information (like multi-omics tools) have greatly simplified the analysis of complex compound mixtures. This has led to a deeper understanding of the specificity ingrained in secondary metabolites.

Nevertheless, challenges related to the isolation of minor abundance metabolites, technical barriers to screening, and lack of reproducibility during microbial culturing remain. These limitations increase the complexity of the process, escalating the usage of solvents, time, and cost for the discovery of new bioactive compounds. Efforts to expand the natural chemical diversity process still include analytical and computational developments, synthetic biology, high throughput screening improvements, and a better integration of omics techniques. For instance, the creation of fractionated libraries of natural products could streamline complex mixtures and amplify the influence of minor components in assays, allowing for a more focused analysis of their properties. The convergence of these technologies opens up exciting possibilities for the development of new drugs.

Furthermore, our knowledge of natural product biosynthesis is continuously growing due to the advances in our understanding of molecular genetics. Although experimental verification and *de novo* structure elucidation remain bottlenecks, synthetic biology has introduced new approaches to increase the throughput and accelerate the identification process of NPs. Natural product genome mining is a powerful tool for discovering new interesting chemical entities, and its impact on drug discovery and other fields is expected to continue growing in the future. Analysis of microbial genetics has revealed a vast number of gene clusters responsible for producing untapped natural products, including "silent genes" that remain unexpressed under normal conditions. The structures of the cryptic metabolites encoded by these genes and their potential roles introduce additional fascinating questions. The availability of genome sequence data, combined with a deeper understanding of molecular genetics underlying natural product biosynthesis and the accessibility of online computational tools, provides unprecedented opportunities to investigate the range and distribution of biosynthetic gene clusters responsible for natural product synthesis. A standardized framework, such as MIBiG (Minimum Information about a Biosynthetic Gene cluster), is essential to facilitate the deposition and retrieval of BGCs and associated metadata, streamlining the sharing and analysis of data from different sources.

However, the inability to recognize unfamiliar biosynthetic gene clusters and forecast the biological properties of identified natural products still obstructs its regular application in the drug discovery process. Achieving a comprehensive understanding of the complex biosynthesis of natural products at the molecular level is crucial for rational pathway engineering. There are still many cases where this level of understanding has not been reached, leading to ineffective manipulations of biosynthetic pathways or low yields of desired products. Therefore, further extensive biochemical investigations are warranted. Given the intricate structural complexity of natural products, this approach has the potential to become a cornerstone in the development of pharmaceutical leads in the years to come.

To advance the search for novel bioactive compounds, we must continue exploring the untapped biosynthetic capabilities of actinobacteria, especially underexplored minor taxa and overlooked bacteria such as marine actinomycetes. These microorganisms may hold immense potential for the evolution of unique secondary metabolic pathways influenced by their distinct local environments. As culture-independent techniques continue to advance, a more productive approach to strain isolation could involve leveraging the distribution patterns of deep-sea and minor taxa actinomycetes or directly cloning and expressing functional genes. This progress holds great promise for the future exploration and exploitation of the immense potential offered by these lesser-known microorganisms.

The utilization of antimicrobial substances in line with evolutionary biology inevitably exerts selection pressure, leading to the emergence of drug resistance over time. This proliferation of antibiotic resistance surpasses our current capacity to develop novel antimicrobials and countermeasures,

emphasizing the urgent need for innovative drugs. In today's interconnected world, the risk of infectious diseases is shared globally, as demonstrated by the COVID-19 pandemic and the rapid spread of new strains.

To address this challenge, advancements in bioinformatics, heterologous expression, synthetic biology, and metabolomics are indispensable. Analytical techniques like mass spectrometry (MS) and nuclear magnetic resonance (NMR), as well as innovative approaches to explore microbial and chemical diversities, serve as crucial starting points for predicting and discovering novel compound families. The integration of natural product chemistry and molecular biology opens up opportunities for the next biotechnological revolution in drug discovery.

Careful observation, deep thinking, and innovation are key factors for potential success. Collaboration among scientists and researchers from diverse disciplines will further fuel the exploration and harnessing of natural products' immense potential. Embracing cutting-edge tools such as artificial intelligence and advanced analytical techniques promises to unlock new insights and accelerate the development of groundbreaking solutions. By combining traditional wisdom with modern approaches, we can forge a path toward a brighter future where the power of nature is fully harnessed for the benefit of all.

BIBLIOGRAPHY

Bibliography

(1) Verpoorte, R.; Choi, Y. H.; Kim, H. K. Metabolomics: Will It Stay? *Phytochemical Analysis* **2010**, *21* (1), 2–3. https://doi.org/10.1002/pca.1191.

(2) Lubbe, A.; Verpoorte, R. Cultivation of Medicinal and Aromatic Plants for Specialty Industrial Materials. *Industrial Crops and Products* **2011**, *34* (1), 785–801. https://doi.org/10.1016/j.indcrop.2011.01.019.

(3) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *Journal of Natural Products* **2020**, *83* (3), 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285.

(4) Newman, D. J.; Cragg, G. M.; Snader, K. M. The Influence of Natural Products upon Drug Discovery. *Natural Product Reports* **2000**, *17* (3), 215–234. https://doi.org/10.1039/a902202c.

(5) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products* **2016**, *79* (3), 629–661. https://doi.org/10.1021/acs.jnatprod.5b01055.

(6) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Last 25 Years. *Journal of Natural Products* **2007**, *70* (3), 461–477. https://doi.org/10.1021/np068054v.

(7) Newman, D. J.; Cragg, G. M. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *J. Nat. Prod.* **2012**, *75* (3), 311–335. https://doi.org/10.1021/np200906s.

(8) Cragg, G. M.; Newman, D. J.; Snader, K. M. Natural Products in Drug Discovery and Development. *Journal of Natural Products* **1997**, *60* (1), 52–60. https://doi.org/10.1021/np9604893.

(9) Newman, D. J.; Cragg, G. M.; Snader, K. M. Natural Products as Sources of New Drugs over the Period 1981−2002. *Journal of Natural Products* **2003**, *66* (7), 1022–1037. https://doi.org/10.1021/np030096l.

(10) Newman, D. J.; Giddings, L.-A. Natural Products as Leads to Antitumor Drugs. *Phytochemistry Reviews* **2014**, *13* (1), 123–137. https://doi.org/10.1007/s11101-013-9292-6.

(11) Koehn, F. E.; Carter, G. T. The Evolving Role of Natural Products in Drug Discovery. *Nature Reviews Drug Discovery* **2005**, *4* (3), 206–220. https://doi.org/10.1038/nrd1657.

(12) Chin, Y.; Balunas, M. J.; Chai, H. B.; Kinghorn, A. D. Drug Discovery from Natural Sources. *The AAPS Journal* **2006**, *8* (2), 239–253. https://doi.org/10.1007/bf02854894.

(13) Li, J. W. H.; Vederas, J. C. Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science* **2009**, *325*, 161–165. https://doi.org/10.1126/science.1168243.

(14) Pham, J. V.; Yilma, M. A.; Feliz, A.; Majid, M. T.; Maffetone, N.; Walker, J. R.; Kim, E.; Cho, H. J.; Reynolds, J. M.; Song, M. C.; Park, S. R.; Yoon, Y. J. A Review of the Microbial Production of Bioactive Natural Products and Biologics. *Frontiers in Microbiology* **2019**, *10*, 1–27. https://doi.org/10.3389/fmicb.2019.01404.

(15) Fouillaud, M.; Dufossé, L. Microbial Secondary Metabolism and Biotechnology. *Microorganisms* **2022**, *10* (1), 123. https://doi.org/10.3390/microorganisms10010123.

(16) Abdel-Razek, A. S.; El-Naggar, M. E.; Allam, A.; Morsy, O. M.; Othman, S. I. Microbial Natural Products in Drug Discovery. *Processes* **2020**, *8* (4), 470. https://doi.org/10.3390/pr8040470.

(17) Cavalier-Smith, T. Origins of Secondary Metabolism. *Ciba Foundation symposium* **1992**, 64–87. https://doi.org/10.1002/9780470514344.ch5.

(18) Heinemann, M.; Sauer, U. Systems Biology of Microbial Metabolism. *Current Opinion in Microbiology* **2010**, *13* (3), 337–343. https://doi.org/10.1016/j.mib.2010.02.005.

(19) Reaves, M. L.; Rabinowitz, J. D. Metabolomics in Systems Microbiology. *Current Opinion in Biotechnology* **2011**, *22* (1), 17–25. https://doi.org/10.1016/j.copbio.2010.10.001.

(20) Seyedsayamdost, M. R. Toward a Global Picture of Bacterial Secondary Metabolism. *Journal of Industrial Microbiology and Biotechnology* **2019**, *46* (3–4), 301–311. https://doi.org/10.1007/s10295-019-02136-y.

(21) Santamaria, G.; Liao, C.; Lindberg, C.; Chen, Y.; Wang, Z.; Rhee, K.; Pinto, F. R.; Yan, J.; Xavier, J. B. Evolution and Regulation of Microbial Secondary Metabolism. *eLife* **2022**, *11*, 1–27. https://doi.org/10.7554/eLife.76119.

(22) Fischbach, M. A.; Clardy, J. One Pathway, Many Products. *Nature Chemical Biology* **2007**, *3* (7), 353–355. https://doi.org/10.1038/nchembio0707-353.

(23) Sanchez, S.; Demain, A. L. Secondary Metabolites. In *Comprehensive Biotechnology*; Elsevier, **2011**; *1*, 155–167. https://doi.org/10.1016/B978-0-08-088504-9.00018-0.

(24) Scherlach, K.; Hertweck, C. Mining and Unearthing Hidden Biosynthetic Potential. *Nature Communications* **2021**, *12* (1). https://doi.org/10.1038/s41467-021-24133-5.

(25) Covington, B. C.; Xu, F.; Seyedsayamdost, M. R. *A* Natural Product Chemist's Guide to Unlocking Silent Biosynthetic Gene Clusters. *Annual Review of Biochemistry* **2021**, *90* (1), 763–788. https://doi.org/10.1146/annurev-biochem-081420-102432.

(26) Baltz, R. H. Molecular Beacons to Identify Gifted Microbes for Genome Mining. *Journal of Antibiotics* **2017**, *70* (5), 639–646. https://doi.org/10.1038/ja.2017.1.

(27) Fischbach, M. A.; Walsh, C. T. Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *Chemical Reviews* **2006**, *106* (8), 3468–3496. https://doi.org/10.1021/cr0503097.

(28) WHO global report on traditional and complementary medicine 2019. Geneva: *World Health Organization*; **2019**. ISBN: 9789241515436

(29) Zhong, G.; Wan, F. An outline on the early pharmacetical development before Galen. *Chinese Journal of Medical History* **1999**, *29* (3), 178–182. PMID: 11624108

(30) Zhang, L.; Song, J.; Kong, L.; Yuan, T.; Li, W.; Zhang, W.; Hou, B.; Lu, Y.; Du, G. The Strategies and Techniques of Drug Discovery from Natural Products. *Pharmacology & Therapeutics* **2020**, *216*. https://doi.org/10.1016/j.pharmthera.2020.107686.

(31) Cragg, G. M.; Newman, D. J. Biodiversity: A Continuing Source of Novel Drug Leads. *Pure and Applied Chemistry* **2005**, *77* (1), 7–24. https://doi.org/10.1351/pac200577010007.

(32) Montinari, M. R.; Minelli, S.; De Caterina, R. The First 3500 years of Aspirin History from Its Roots – A Concise Summary. *Vascular Pharmacology* **2019**, *113*, 1–8. https://doi.org/10.1016/j.vph.2018.10.008.

(33) Přichystal, J.; Schug, K. A.; Lemr, K.; Novák, J.; Havlíček, V. Structural Analysis of Natural Products. *Analytical Chemistry* **2016**, *88* (21), 10338–10346. https://doi.org/10.1021/acs.analchem.6b02386.

(34) Fleming, A. On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to Their Use in the Isolation of *B. Influenzæ*. *British Journal of Experimental Pathology* **1929**, 10 (3), 226–236. PMCID: PMC2048009

(35) Gaynes, R. The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use. *Emerging Infectious Diseases* **2017**, *23* (5), 849–853. https://doi.org/10.3201/eid2305.161556.

(36) Katz, L.; Baltz, R. H. Natural Product Discovery: Past, Present, and Future. *Journal of Industrial Microbiology and Biotechnolog* **2016**, *43* (2-3), 155–176. https://doi.org/10.1007/s10295-015-1723-5.

(37) Waksman, S. A.; Woodruff, H. B. Bacteriostatic and Bactericidal Substances Produced by a Soil Actinomyces. *Experimental Biology and Medicine* **1940**, *45* (2), 609–614. https://doi.org/10.3181/00379727-45-11768.

(38) Waksman, S. A. Streptomycin: Background, Isolation, Properties, and Utilization. *Science* **1953**, *118* (3062), 259–266. https://doi.org/10.1126/science.118.3062.259.

(39) Conly, J.; Johnston, B. Where Are All the New Antibiotics? The New Antibiotic Paradox. *Canadian Journal of Infectious Diseases and Medical Microbiology* **2005**, *16* (3), 159–160. https://doi.org/10.1155/2005/892058.

(40) Shen, B. A New Golden Age of Natural Products Drug Discovery. *Cell* **2015**, *163* (6), 1297–1300. https://doi.org/10.1016/j.cell.2015.11.031.

(41) Baker, D. D.; Chu, M.; Oza, U.; Rajgarhia, V. The Value of Natural Products to Future Pharmaceutical Discovery. *Natural Products Reports* **2007**, *24* (6), 1225–1244. https://doi.org/10.1039/B602241N.

(42) Butler, M. S. Natural Products to Drugs: Natural Product-Derived Compounds in Clinical Trials. *Natural Products Reports* **2008**, *25* (3), 475–516. https://doi.org/10.1039/B514294F.

(43) Spellberg, B. The Future of Antibiotics. *Critical Care* **2014**, *18* (3), 228. https://doi.org/10.1186/cc13948.

(44) Murray, C. J. L.; Ikuta, K. S.; Sharara, F.; Swetschinski, L.; Robles Aguilar, G.; Gray, A.; Han, C.; Bisignano, C.; Rao, P.; Wool, E.; Johnson, S. C.; Browne, A. J.; Chipeta, M. G.; Fell, F.; Hackett, S.; Haines-Woodhouse, G.; Kashef Hamadani, B. H.; Kumaran, E. A. P.; McManigal, B.; Achalapong, S.; Agarwal, R.; Akech, S.; Albertson, S.; Amuasi, J.; Andrews, J.; Aravkin, A.; Ashley, E.; Babin, F.-X.; Bailey, F.; Baker, S.; Basnyat, B.; Bekker, A.; Bender, R.; Berkley, J. A.; Bethou, A.; Bielicki, J.; Boonkasidecha, S.; Bukosia, J.; Carvalheiro, C.; Castañeda-Orjuela, C.; Chansamouth, V.; Chaurasia, S.; Chiurchiù, S.; Chowdhury, F.; Clotaire Donatien, R.; Cook, A. J.; Cooper, B.; Cressey, T. R.; Criollo-Mora, E.; Cunningham, M.; Darboe, S.; Day, N. P. J.; De Luca, M.; Dokova, K.; Dramowski, A.; Dunachie, S. J.; Duong Bich, T.; Eckmanns, T.; Eibach, D.; Emami, A.; Feasey, N.; Fisher-Pearson, N.; Forrest, K.; Garcia, C.; Garrett, D.; Gastmeier, P.; Giref, A. Z.; Greer, R. C.; Gupta, V.; Haller, S.; Haselbeck, A.; Hay, S. I.; Holm, M.; Hopkins, S.; Hsia, Y.; Iregbu, K. C.; Jacobs, J.; Jarovsky, D.; Javanmardi, F.; Jenney, A. W. J.; Khorana,

M.; Khusuwan, S.; Kissoon, N.; Kobeissi, E.; Kostyanev, T.; Krapp, F.; Krumkamp, R.; Kumar, A.; Kyu, H. H.; Lim, C.; Lim, K.; Limmathurotsakul, D.; Loftus, M. J.; Lunn, M.; Ma, J.; Manoharan, A.; Marks, F.; May, J.; Mayxay, M.; Mturi, N.; Munera-Huertas, T.; Musicha, P.; Musila, L. A.; Mussi-Pinhata, M. M.; Naidu, R. N.; Nakamura, T.; Nanavati, R.; Nangia, S.; Newton, P.; Ngoun, C.; Novotney, A.; Nwakanma, D.; Obiero, C. W.; Ochoa, T. J.; Olivas-Martinez, A.; Olliaro, P.; Ooko, E.; Ortiz-Brizuela, E.; Ounchanum, P.; Pak, G. D.; Paredes, J. L.; Peleg, A. Y.; Perrone, C.; Phe, T.; Phommasone, K.; Plakkal, N.; Ponce-de-Leon, A.; Raad, M.; Ramdin, T.; Rattanavong, S.; Riddell, A.; Roberts, T.; Robotham, J. V.; Roca, A.; Rosenthal, V. D.; Rudd, K. E.; Russell, N.; Sader, H. S.; Saengchan, W.; Schnall, J.; Scott, J. A. G.; Seekaew, S.; Sharland, M.; Shivamallappa, M.; Sifuentes-Osornio, J.; Simpson, A. J.; Steenkeste, N.; Stewardson, A. J.; Stoeva, T.; Tasak, N.; Thaiprakong, A.; Thwaites, G.; Tigoi, C.; Turner, C.; Turner, P.; van Doorn, H. R.; Velaphi, S.; Vongpradith, A.; Vongsouvath, M.; Vu, H.; Walsh, T.; Walson, J. L.; Waner, S.; Wangrangsimakul, T.; Wannapinij, P.; Wozniak, T.; Young Sharma, T. E. M. W.; Yu, K. C.; Zheng, P.; Sartorius, B.; Lopez, A. D.; Stergachis, A.; Moore, C.; Dolecek, C.; Naghavi, M. Global Burden of Bacterial Antimicrobial Resistance in 2019: A Systematic Analysis. *The Lancet* **2022**, *399* (10325), 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0.

(45) Global antimicrobial resistance and use surveillance system (GLASS) report 2022. Geneva: *World Health Organization*; **2022**. ISBN: 9789240062702.

(46) WHO Regional Office for Europe/European Centre for Disease Prevention and Control. Antimicrobial Resistance Surveillance in Europe 2022 – 2020 Data. Copenhagen: *WHO Regional Office for Europe*; **2022**. ISBN: 9789289056687.

(47) Urban-Chmiel, R.; Marek, A.; Stępień-Pyśniak, D.; Wieczorek, K.; Dec, M.; Nowaczek, A.; Osek, J. Antibiotic Resistance in Bacteria—A Review. *Antibiotics* **2022**, *11* (8), 1079. https://doi.org/10.3390/antibiotics11081079.

(48) Baker, R. E.; Mahmud, A. S.; Miller, I. F.; Rajeev, M.; Rasambainarivo, F.; Rice, B. L.; Takahashi, S.; Tatem, A. J.; Wagner, C. E.; Wang, L. F.; Wesolowski, A.; Metcalf, C. J. E. Infectious Disease in an Era of Global Change. *Nature Reviews Microbiology* **2022**, *20*, 193–205. https://doi.org/10.1038/s41579-021-00639-z.

(49) Cândido, E. de S.; de Barros, E.; Cardoso, M. H.; Franco, O. L. Bacterial Cross-Resistance to Anti-Infective Compounds. Is It a Real Problem? *Current Opinion in Pharmacology* **2019**, *48*, 76–81. https://doi.org/10.1016/j.coph.2019.05.004.

(50) Langford, B. J.; So, M.; Raybardhan, S.; Leung, V.; Soucy, J.-P. R.; Westwood, D.; Daneman, N.; MacFadden, D. R. Antibiotic Prescribing in Patients with COVID-19: Rapid Review and Meta-Analysis. *Clinical Microbiology and Infection* **2021**, *27* (4), 520–531. https://doi.org/10.1016/j.cmi.2020.12.018.

(51) Rawson, T. M.; Moore, L. S. P.; Zhu, N.; Ranganathan, N.; Skolimowska, K.; Gilchrist, M.; Satta, G.; Cooke, G.; Holmes, A. Bacterial and Fungal Coinfection in Individuals With Coronavirus: A Rapid Review To Support COVID-19 Antimicrobial Prescribing. *Clinical Infectious Diseases* **2020**, *71* (9), 2459–2468. https://doi.org/10.1093/cid/ciaa530.

(52) Adebisi, Y. A.; Alaran, A. J.; Okereke, M.; Oke, G. I.; Amos, O. A.; Olaoye, O. C.; Oladunjoye, I.; Olanrewaju, A. Y.; Ukor, N. A.; Lucero-Prisno, D. E. COVID-19 and Antimicrobial Resistance: A Review. *Infectious Diseases: Research and Treatment* **2021**, *14*. https://doi.org/10.1177/11786337211033870.

(53) Knight, G. M.; Glover, R. E.; McQuaid, C. F.; Olaru, I. D.; Gallandat, K.; Leclerc, Q. J.; Fuller, N. M.; Willcocks, S. J.; Hasan, R.; van Kleef, E.; Chandler, C. I. R. Antimicrobial Resistance and COVID-19: Intersections and Implications. *eLife* **2021**, *10*. https://doi.org/10.7554/eLife.64139.

(54) Atanasov, A. G.; Zotchev, S. B.; Dirsch, V. M.; Supuran, C. T. Natural Products in Drug Discovery: Advances and Opportunities. *Nature Reviews Drug Discovery* **2021**, *20* (3), 200–216. https://doi.org/10.1038/s41573-020-00114-z.

(55) Newman, D. J. Natural Products and Drug Discovery. *National Science Review* **2022**, *9* (11). https://doi.org/10.1093/nsr/nwac206.

(56) Pettit, R. K. Mixed Fermentation for Natural Product Drug Discovery. *Applied Microbiology and Biotechnology* **2009**, *83* (1), 19–25. https://doi.org/10.1007/s00253-009-1916-9.

(57) Laatsch, H. Antibase 2012 - *The Natural Compound Identifier*. Wiley-VCH Verlag GmbH & Co., **2012**.

(58) Chiang, Y. M.; Szewczyk, E.; Nayak, T.; Davidson, A. D.; Sanchez, J. F.; Lo, H. C.; Ho, W. Y.; Simityan, H.; Kuo, E.; Praseuth, A.; Watanabe, K.; Oakley, B. R.; Wang, C. C. C. Molecular Genetic Mining of the Aspergillus Secondary Metabolome: Discovery of the Emericellamide Biosynthetic Pathway. *Chemistry and Biology* **2008**, *15* (6), 527–532. https://doi.org/10.1016/j.chembiol.2008.05.010.

(59) Schroeckh, V.; Scherlach, K.; Nutzmann, H. W.; Shelest, E.; Schmidt-Heck, W.; Schuemann, J.; Martin, K.; Hertweck, C.; Brakhage, A. A. Intimate Bacterial-Fungal Interaction Triggers Biosynthesis of Archetypal Polyketides in Aspergillus Nidulans. *Proceedings of the National Academy of Sciences* **2009**, *106* (34), 14558–14563. https://doi.org/10.1073/pnas.0901870106.

(60) Bergmann, S.; Schümann, J.; Scherlach, K.; Lange, C.; Brakhage, A. A.; Hertweck, C. Genomics-Driven Discovery of PKS-NRPS Hybrid Metabolites from Aspergillus Nidulans. *Nature Chemical Biology* **2007**, *3* (4), 213–217. https://doi.org/10.1038/nchembio869.

(61) Hibbing, M. E.; Fuqua, C.; Parsek, M. R.; Peterson, S. B. Bacterial Competition: Surviving and Thriving in the Microbial Jungle. *Nature Reviews Microbiology* **2010**, *8* (1), 15–25. https://doi.org/10.1038/nrmicro2259.

(62) Galagan, J. E.; Calvo, S. E.; Cuomo, C.; Ma, L. J.; Wortman, J. R.; Batzoglou, S.; Lee, S. I.; Baştürkmen, M.; Spevak, C. C.; Clutterbuck, J.; Kapitonov, V.; Jurka, J.; Scazzocchio, C.; Farman, M.; Butler, J.; Purcell, S.; Harris, S.; Braus, G. H.; Draht, O.; Busch, S.; D'Enfert, C.; Bouchier, C.; Goldman, G. H.; Bell-Pedersen, D.; Griffiths-Jones, S.; Doonan, J. H.; Yu, J.; Vienken, K.; Pain, A.; Freitag, M.; Selker, E. U.; Archer, D. B.; Peñalva, M. Á.; Oakley, B. R.; Momany, M.; Tanaka, T.; Kumagai, T.; Asai, K.; Machida, M.; Nierman, W. C.; Denning, D. W.; Caddick, M.; Hynes, M.; Paoletti, M.; Fischer, R.; Miller, B.; Dyer, P.; Sachs, M. S.; Osmani, S. A.; Birren, B. W. Sequencing of Aspergillus Nidulans and Comparative Analysis with A. Fumigatus and A. Oryzae. *Nature* **2005**, *438* (7071), 1105–1115. https://doi.org/10.1038/nature04341.

(63) Moody, S. C. Microbial Co-Culture: Harnessing Intermicrobial Signaling for the Production of Novel Antimicrobials. *Future Microbiology* **2014**, *9* (5), 575–578. https://doi.org/10.2217/fmb.14.25.

(64) Barka, E. A.; Vatsa, P.; Sanchez, L.; Gaveau-Vaillant, N.; Jacquard, C.; Klenk, H.-P.; Clément, C.; Ouhdouch, Y.; van Wezel, G. P. Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiology and Molecular Biology Reviews* **2016**, *80* (1), 1–43. https://doi.org/10.1128/MMBR.00019-15.

(65) Ventura, M.; Canchaya, C.; Tauch, A.; Chandra, G.; Fitzgerald, G. F.; Chater, K. F.; van Sinderen, D. Genomics of Actinobacteria : Tracing the Evolutionary History of an Ancient Phylum . *Microbiology and Molecular Biology Reviews* **2007**, *71* (3), 495–548. https://doi.org/10.1128/mmbr.00005-07.

(66) Yagüe, P.; López-García, M. T.; Rioseras, B.; Sánchez, J.; Manteca, Á. Pre-Sporulation Stages of Streptomyces Differentiation: State-of-the-Art and Future Perspectives. *FEMS Microbiology Letters* **2013**, *342* (2), 79–88. https://doi.org/10.1111/1574-6968.12128.

(67) Nett, M.; Ikeda, H.; Moore, B. S. Genomic Basis for Natural Product Biosynthetic Diversity in the Actinomycetes. *Natural Product Reports* **2009**, *26* (11), 1362–1384. https://doi.org/10.1039/b817069j.

(68) Parte, A. C.; Carbasse, J. S.; Meier-Kolthoff, J. P.; Reimer, L. C.; Göker, M. List of Prokaryotic Names with Standing in Nomenclature (LPSN) Moves to the DSMZ. *International Journal of Systematic and Evolutionary Microbiology* **2020**, *70* (11), 5607–5612. https://doi.org/10.1099/ijsem.0.004332.

(69) Alam, K.; Mazumder, A.; Sikdar, S.; Zhao, Y.-M.; Hao, J.; Song, C.; Wang, Y.; Sarkar, R.; Islam, S.; Zhang, Y.; Li, A. Streptomyces: The Biofactory of Secondary Metabolites. *Frontiers in Microbiology* **2022**, *13*. https://doi.org/10.3389/fmicb.2022.968053.

(70) Hopwood, D. A. Soil To Genomics: The Streptomyces Chromosome. *Annual Review of Genetics* **2006**, *40* (1), 1–23. https://doi.org/10.1146/annurev.genet.40.110405.090639.

(71) van der Meij, A.; Worsley, S. F.; Hutchings, M. I.; van Wezel, G. P. Chemical Ecology of Antibiotic Production by Actinomycetes. *FEMS Microbiology Reviews* **2017**, *41* (3), 392–416. https://doi.org/10.1093/femsre/fux005.

(72) Waksman, S. A.; Lechevalier, H. A. Neomycin, a New Antibiotic Active against Streptomycin-Resistant Bacteria, Including Tuberculosis Organisms. *Science* **1949**, *109* (2830), 305–307. https://doi.org/10.1126/science.109.2830.305.

(73) Umezawa, H. Kanamycin: Its Discovery. *Annals of the New York Academy of Sciences* **1958**, *76* (2), 20–26. https://doi.org/10.1111/j.1749-6632.1958.tb54688.x.

(74) Ehrlich, J.; Bartz, Q. R.; Smith, R. M.; Joslyn, D. A.; Burkholder, P. R. Chloromycetin, a New Antibiotic From a Soil Actinomycete. *Science* **1947**, *106* (2757), 417–417. https://doi.org/10.1126/science.106.2757.417.

(75) Kahan, J. S.; Kahan, F. M.; Goegelman, R.; Currie, S. A.; Jackson, M.; Stapley, E. O.; Miller, T. W.; Miller, A. K.; Hendlin, D.; Woodruff, H. B.; Birnbaum, J.; Mochales, S.; Hernandez, S. Thienamycin, a New β-Lactam Antibiotic i. Discovery, Taxonomy, Isolation and Physical Properties. *The Journal of Antibiotics* **1979**, *32* (1), 1–12. https://doi.org/10.7164/antibiotics.32.1.

(76) Ding, T.; Yang, L.-J.; Zhang, W.-D.; Shen, Y.-H. The Secondary Metabolites of Rare Actinomycetes: Chemistry and Bioactivity. *RSC Advances* **2019**, *9* (38), 21964–21988. https://doi.org/10.1039/C9RA03579F.

(77) Subramani, R.; Aalbersberg, W. Culturable Rare Actinomycetes: Diversity, Isolation and Marine Natural Product Discovery. *Applied Microbiology and Biotechnology* **2013**, *97* (21), 9291–9321. https://doi.org/10.1007/s00253-013-5229-7.

(78) Tiwari, K.; Gupta, R. K. Rare Actinomycetes: A Potential Storehouse for Novel Antibiotics. *Critical Reviews in Biotechnology* **2012**, *32* (2), 108–132. https://doi.org/10.3109/07388551.2011.562482.

(79) McGuire, J. M.; Bunch, R. L.; Anderson, R. C.; Boaz, H. E.; Flynn, E. H.; Powell, H. M.; Smith, J. W. Ilotycin, a New Antibiotic. *Antibiotics & chemotherapy (Northfield, Ill.)* **1952**, *2* (6), 281–283. PMID: 24541924.

(80) Somma, S.; Gastaldo, L.; Corti, A. Teicoplanin, a New Antibiotic from Actinoplanes Teichomyceticus Nov. Sp. *Antimicrobial Agents and Chemotherapy* **1984**, *26* (6), 917–923. https://doi.org/10.1128/aac.26.6.917.

(81) Zhanel, G. G.; Walkty, A. J.; Karlowsky, J. A. Fidaxomicin: A Novel Agent for the Treatment of Clostridium Difficile Infection. *Canadian Journal of Infectious Diseases and Medical Microbiology* **1900**, *26*. https://doi.org/10.1155/2015/934594.

(82) 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,1 a New Antibiotic Complex from Micromonospora. *Journal of Medicinal Chemistry* **1963**, *6* (4), 463–464. https://doi.org/10.1021/jm00340a034.

(83) Malabarba, A.; Goldstein, B. P. Origin, Structure, and Activity in Vitro and in Vivo of Dalbavancin. *Journal of Antimicrobial Chemotherapy* **2005**, *55* (suppl_2), ii15–ii20. https://doi.org/10.1093/jac/dki005.

(84) Castiglione, F.; Lazzarini, A.; Carrano, L.; Corti, E.; Ciciliato, I.; Gastaldo, L.; Candiani, P.; Losi, D.; Marinelli, F.; Selva, E.; Parenti, F. Determining the Structure and Mode of Action of Microbisporicin, a Potent Lantibiotic Active Against Multiresistant Pathogens. *Chemistry & Biology* **2008**, *15* (1), 22–31. https://doi.org/10.1016/j.chembiol.2007.11.009.

(85) Ortiz-López, F. J.; Oves-Costales, D.; Carretero-Molina, D.; Martín, J.; Díaz, C.; de la Cruz, M.; Román-Hurtado, F.; Álvarez-Arévalo, M.; Jørgensen, T. S.; Reyes, F.; Weber, T.; Genilloud, O. Crossiellidines A–F, Unprecedented Pyrazine-Alkylguanidine Metabolites with Broad-Spectrum Antibacterial Activity from Crossiella Sp. *Organic Letters* **2023**. https://doi.org/10.1021/acs.orglett.3c01088.

(86) Barna, J. C. J.; Williams, D. H. The Structure and Mode of Action of Glycopeptide Antibiotics of the Vancomycin Group. *Annual Review of Microbiology* **1984**, *38* (1), 339–357. https://doi.org/10.1146/annurev.mi.38.100184.002011.

(87) Sensi, P.; Greco, A. M.; Ballota, R. Rifomycin. I. Isolation and Properties of Rifomycin B and Rifomycin Complex. *Antibiotics Annual* **1960**, *7*, 262–270. PMID: 14444943.

(88) Kunimoto, S.; Lu, J.; Esumi, H.; Yamazaki, Y.; Kinoshita, N.; Honma, Y.; Hamada, M.; Ohsono, M.; Ishizuka, M.; Takeuchi, T. Kigamicins, Novel Antitumor Antibiotics: I. Taxonomy, Isolation, Physico-Chemical Properties and Biological Activities. *The Journal of Antibiotics* **2003**, *56* (12), 1004–1011. https://doi.org/10.7164/antibiotics.56.1004.

(89) Bauermeister, A.; Calil, F. A.; das C. L. Pinto, F.; Medeiros, T. C. T.; Almeida, L. C.; Silva, L. J.; de Melo, I. S.; Zucchi, T. D.; Costa-Lotufo, L. V.; Moraes, L. A. B. Pradimicin-IRD from Amycolatopsis Sp. IRD-009 and Its Antimicrobial and Cytotoxic Activities. *Natural Product Research* **2019**, *33* (12), 1713–1720. https://doi.org/10.1080/14786419.2018.1434639.

(90) Banskota, A. H.; McAlpine, J. B.; Sørensen, D.; Ibrahim, A.; Aouidate, M.; Piraee, M.; Alarco, A.-M.; Farnet, C. M.; Zazopoulos, E. Genomic Analyses Lead to Novel Secondary Metabolites. *The Journal of Antibiotics* **2006**, *59* (9), 533–542. https://doi.org/10.1038/ja.2006.74.

(91) Hopmann, C.; Kurz, M.; Brönstrup, M.; Wink, J.; LeBeller, D. Isolation and Structure Elucidation of Vancoresmycin—a New Antibiotic from Amycolatopsis Sp. ST 101170. *Tetrahedron Letters* **2002**, *43* (3), 435–438. https://doi.org/10.1016/S0040-4039(01)02171-2.

(92)   Khalil, Z. G.; Salim, A. A.; Vuong, D.; Crombie, A.; Lacey, E.; Blumenthal, A.; Capon, R. J. Amycolatopsins A–C: Antimycobacterial Glycosylated Polyketide Macrolides from the Australian Soil Amycolatopsis Sp. MST-108494. *The Journal of Antibiotics* **2017**, *70* (12), 1097–1103. https://doi.org/10.1038/ja.2017.119.

(93)   Shi, Y.; Zhang, J.; Tian, X.; Wu, X.; Li, T.; Lu, C.; Shen, Y. Isolation of 11,12- Seco -Rifamycin W Derivatives Reveals a Cleavage Pattern of the Rifamycin Ansa Chain. *Organic Letters* **2019**, *21* (4), 900–903. https://doi.org/10.1021/acs.orglett.8b03792.

(94)   Xiao, Y. S.; Zhang, B.; Zhang, M.; Guo, Z. K.; Deng, X. Z.; Shi, J.; Li, W.; Jiao, R. H.; Tan, R. X.; Ge, H. M. Rifamorpholines A–E, Potential Antibiotics from Locust-Associated Actinobacteria Amycolatopsis Sp. Hca4. *Organic & Biomolecular Chemistry* **2017**, *15* (18), 3909–3916. https://doi.org/10.1039/C7OB00614D.

(95)   Pan, C.; Kuranaga, T.; Liu, C.; Lu, S.; Shinzato, N.; Kakeya, H. Thioamycolamides A-E, Sulfur-Containing Cycliclipopeptides Produced by the Rare Actinomycete Amycolatopsis Sp. *Organic Letters* **2020**, *22* (8), 3014–3017. https://doi.org/10.1021/acs.orglett.0c00776.

(96)   Zhang, C.; Herath, K.; Jayasuriya, H.; Ondeyka, J. G.; Zink, D. L.; Occi, J.; Birdsall, G.; Venugopal, J.; Ushio, M.; Burgess, B.; Masurekar, P.; Barrett, J. F.; Singh, S. B. Thiazomycins, Thiazolyl Peptide Antibiotics from Amycolatopsis Fastidiosa. *Journal of Natural Products* **2009**, *72* (5), 841–847. https://doi.org/10.1021/np800783b.

(97)   Hashizume, H.; Sawa, R.; Yamashita, K.; Nishimura, Y.; Igarashi, M. Structure and Antibacterial Activities of New Cyclic Peptide Antibiotics, Pargamicins B, C and D, from Amycolatopsis Sp. ML1-hF4. *The Journal of Antibiotics* **2017**, *70* (5), 699–704. https://doi.org/10.1038/ja.2017.34.

(98)   Hashizume, H.; Iijima, K.; Yamashita, K.; Kimura, T.; Wada, S.; Sawa, R.; Igarashi, M. Valgamicin C, a Novel Cyclic Depsipeptide Containing the Unusual Amino Acid Cleonine, and Related Valgamicins A, T and V Produced by Amycolatopsis Sp. ML1-hF4. *The Journal of Antibiotics* **2018**, *71* (1), 129–134. https://doi.org/10.1038/ja.2017.135.

(99)   Tsuji, N.; Kamigauchi, T.; Kobayashi, M.; Terui, Y. New Glycopeptide Antibiotics: II. The Isolation and Structures of Chloroorienticins. *The Journal of Antibiotics* **1988**, *41* (10), 1506–1510. https://doi.org/10.7164/antibiotics.41.1506.

(100)   Berdnikova, T. F.; Shashkov, A. S.; Katrukha, G. S.; Lapchinskaya, O. A.; Yurkevich, N. V.; Grachev, A. A.; Nifant'ev, N. E. The Structure of Antibiotic Eremomycin B. *Russian Journal of Bioorganic Chemistry* **2009**, *35* (4), 497–503. https://doi.org/10.1134/S1068162009040128.

(101)   Xu, X.; Han, L.; Zhao, L.; Chen, X.; Miao, C.; Hu, L.; Huang, X.; Chen, Y.; Li, Y. Echinosporin Antibiotics Isolated from Amycolatopsis Strain and Their Antifungal Activity against Root-Rot Pathogens of the Panax Notoginseng. *Folia Microbiologica* **2019**, *64* (2), 171–175. https://doi.org/10.1007/s12223-018-0642-z.

(102)   Breinholt, J.; Kulik, A.; Gürtler, H.; Fiedler, H.-P.; Wahlberg, I.; Juliá, L.; Persson, O. Tigloside: A New Tigloylated Tetrasaccharide from Amycolatopsis Sp. *Acta Chemica Scandinavica* **1998**, *52*, 1239–1242. https://doi.org/10.3891/acta.chem.scand.52-1239.

(103)   Guo, Z.-K.; Jiao, R.-H.; Dai, H.-F.; Mei, W.-L.; Tan, R.-X.; Ge, H.-M. Actinotetraoses I-K: Tetrasaccharide Metabolites Produced by an Insect-Derived Actinobacteria, Amycolatopsis Sp. HCa1. *Chemistry & Biodiversity* **2013**, *10* (2), 296–302. https://doi.org/10.1002/cbdv.201200224.

(104)   Ma, S. Y.; Xiao, Y. S.; Zhang, B.; Shao, F. L.; Guo, Z. K.; Zhang, J. J.; Jiao, R. H.; Sun, Y.; Xu, Q.; Tan, R. X.; Ge, H. M. Amycolamycins A and B, Two Enediyne-Derived Compounds from a Locust-Associated Actinomycete. *Organic Letters* **2017**, *19* (22), 6208–6211. https://doi.org/10.1021/acs.orglett.7b03113.

(105)   Li, X.-M.; Li, X.-M.; Lu, C.-H. Abscisic Acid-Type Sesquiterpenes and Ansamycins from Amycolatopsis Alba DSM 44262. *Journal of Asian Natural Products Research* **2017**, *19* (10), 946–953. https://doi.org/10.1080/10286020.2017.1285909.

(106)   Song, Z.; Xu, T.; Wang, J.; Hou, Y.; Liu, C.; Liu, S.; Wu, S. Secondary Metabolites of the Genus Amycolatopsis: Structures, Bioactivities and Biosynthesis. *Molecules* **2021**, *26* (7), 1884. https://doi.org/10.3390/molecules26071884.

(107)   Cario, A.; Oliver, G. C.; Rogers, K. L. Exploring the Deep Marine Biosphere: Challenges, Innovations, and Opportunities. *Frontiers in Earth Science* **2019**, *7*. https://doi.org/10.3389/feart.2019.00225.

(108)   Carroll, A. R.; Copp, B. R.; Davis, R. A.; Keyzers, R. A.; Prinsep, M. R. Marine Natural Products. *Natural Product Reports* **2023**, *40* (2), 275–325. https://doi.org/10.1039/D2NP00083K.

(109)   Schwartsmann, G.; da Rocha, A. B.; Berlinck, R. G.; Jimeno, J. Marine Organisms as a Source of New Anticancer Agents. *The Lancet Oncology* **2001**, *2* (4), 221–225. https://doi.org/10.1016/S1470-2045(00)00292-8.

(110) Green, S. D.; Konig, H. Treatment of Acute Myeloid Leukemia in the Era of Genomics—Achievements and Persisting Challenges. *Frontiers in Genetics* **2020**, *11*. https://doi.org/10.3389/fgene.2020.00480.

(111) Kamjam, M.; Sivalingam, P.; Deng, Z.; Hong, K. Deep Sea Actinomycetes and Their Secondary Metabolites. *Frontiers in Microbiology* **2017**, *8*. https://doi.org/10.3389/fmicb.2017.00760.

(112) McCauley, E. P.; Piña, I. C.; Thompson, A. D.; Bashir, K.; Weinberg, M.; Kurz, S. L.; Crews, P. Highlights of Marine Natural Products Having Parallel Scaffolds Found from Marine-Derived Bacteria, Sponges, and Tunicates. *The Journal of Antibiotics* **2020**, *73* (8), 504–525. https://doi.org/10.1038/s41429-020-0330-5.

(113) Malve, H. Exploring the Ocean for New Drug Developments: Marine Pharmacology. *Journal of Pharmacy Bioallied Sciences* **2016**, *8* (2), 83–91. https://doi.org/10.4103/0975-7406.171700.

(114) Urdiales, JoséL.; Morata, P.; De Castro, I. N.; Sánchez-Jiménez, F. Antiproliferative Effect of Dehydrodidemnin B (DDB), a Depsipeptide Isolated from Mediterranean Tunicates. *Cancer Letters* **1996**, *102* (1–2), 31–37. https://doi.org/10.1016/0304-3835(96)04151-1.

(115) Cuevas, C.; Francesch, A. Development of Yondelis® (Trabectedin, ET-743). A Semisynthetic Process Solves the Supply Problem. *Natural Product Reports* **2009**, *26* (3), 322. https://doi.org/10.1039/b808331m.

(116) Gordon, E. M.; Sankhala, K. K.; Chawla, N.; Chawla, S. P. Trabectedin for Soft Tissue Sarcoma: Current Status and Future Perspectives. *Advances in Therapy* **2016**, *33* (7), 1055–1071. https://doi.org/10.1007/s12325-016-0344-3.

(117) Barreca, M.; Spanò, V.; Montalbano, A.; Cueto, M.; Díaz Marrero, A. R.; Deniz, I.; Erdoğan, A.; Lukić Bilela, L.; Moulin, C.; Taffin-de-Givenchy, E.; Spriano, F.; Perale, G.; Mehiri, M.; Rotter, A.; P. Thomas, O.; Barraja, P.; Gaudêncio, S. P.; Bertoni, F. Marine Anticancer Agents: An Overview with a Particular Focus on Their Chemical Classes. *Marine Drugs* **2020**, *18* (12), 619–646. https://doi.org/10.3390/md18120619.

(118) Jensen, P. R.; Moore, B. S.; Fenical, W. The Marine Actinomycete Genus Salinispora: A Model Organism for Secondary Metabolite Discovery. *Natural Product Reports* **2015**, *32* (5), 738–751. https://doi.org/10.1039/c4np00167b.

(119) US National Institutes of Health. *A phase III trial of marizomib in patients with newly diagnosed glioblastoma (NCT03345095).* https://clinicaltrials.gov/ct2/show/NCT03345095. (accessed 2023-07-26).

(120) Jang, K. H.; Nam, S.-J.; Locke, J. B.; Kauffman, C. A.; Beatty, D. S.; Paul, L. A.; Fenical, W. Anthracimycin, a Potent Anthrax Antibiotic from a Marine-Derived Actinomycete. *Angewandte Chemie International Edition* **2013**, *52* (30), 7822–7824. https://doi.org/10.1002/anie.201302749.

(121) Subramani, R.; Sipkema, D. Marine Rare Actinomycetes: A Promising Source of Structurally Diverse and Unique Novel Natural Products. *Marine Drugs* **2019**, *17* (5), 249–288. https://doi.org/10.3390/md17050249.

(122) Locey, K. J.; Lennon, J. T. Scaling Laws Predict Global Microbial Diversity. *Proceedings of the National Academy of Sciences* **2016**, *113* (21), 5970–5975. https://doi.org/10.1073/pnas.1521291113.

(123) Jensen, P. R. Natural Products and the Gene Cluster Revolution. *Trends in Microbiology* **2016**, *24* (12), 968–977. https://doi.org/10.1016/j.tim.2016.07.006.

(124) Genilloud, O. Actinomycetes: Still a Source of Novel Antibiotics. *Natural Product Reports* **2017**, *34* (10), 1203–1232. https://doi.org/10.1039/C7NP00026J.

(125) Jorge, T. F.; Rodrigues, J. A.; Caldana, C.; Schmidt, R.; van Dongen, J. T.; Thomas-Oates, J.; António, C. Mass Spectrometry-Based Plant Metabolomics: Metabolite Responses to Abiotic Stress. *Mass Spectrometry Reviews* **2016**, *35*, 620–649. https://doi.org/10.1002/mas.21449.

(126) Hong, J.; Yang, L.; Zhang, D.; Shi, J. Plant Metabolomics: An Indispensable System Biology Tool for Plant Science. *International Journal of Molecular Sciences* **2016**, *17* (6), 767–782. https://doi.org/10.3390/ijms17060767.

(127) Aksenov, A. A.; da Silva, R.; Knight, R.; Lopes, N. P.; Dorrestein, P. C. Global Chemical Analysis of Biology by Mass Spectrometry. *Nature Reviews Chemistry* **2017**, *1* (54). https://doi.org/10.1038/s41570-017-0054.

(128) da Silva, R. R.; Wang, M.; Nothias, L.-F.; van der Hooft, J. J. J.; Caraballo-Rodríguez, A. M.; Fox, E.; Balunas, M. J.; Klassen, J. L.; Lopes, N. P.; Dorrestein, P. C. Propagating Annotations of Molecular Networks Using in Silico Fragmentation. *PLoS Computational Biologyl* **2018**, *14* (4). https://doi.org/10.1371/journal.pcbi.1006089.

(129) Wolfender, J.-L.; Glauser, G.; Boccard, J.; Rudaz, S. MS-Based Plant Metabolomic Approaches for Biomarker Discovery. *Natural Product Communications* **2009**, *4* (10), 1417–1430. https://doi.org/10.1177/1934578X0900401019.

(130) Ng, J.; Bandeira, N.; Liu, W. T.; Ghassemian, M.; Simmons, T. L.; Gerwick, W. H.; Linington, R.; Dorrestein, P. C.; Pevzner, P. A. Dereplication and de Novo Sequencing of Nonribosomal Peptides. *Nature Methods Methods* **2009**, *6* (8), 569–599. https://doi.org/10.1038/nmeth.1350.

(131) Gabrielson, S. W. Scifinder, 2018. https://scifinder-n.cas.org.

(132) Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E. PubChem 2023 Update. *Nucleic Acids Research* **2023**, *51* (D1), D1373–D1380. https://doi.org/10.1093/nar/gkac956.

(133) Pence, H. E.; Williams, A. ChemSpider: An Online Chemical Information Resource. *Journal of Chemical Education* **2010**, *87* (11), 1123–1124. https://doi.org/10.1021/ed100697w.

(134) Goodman, J. Computer Software Review: Reaxys. *Journal of Chemical Information and Modeling* **2009**, *49* (12), 2897–2898. https://doi.org/10.1021/ci900437n.

(135) van Santen, J. A.; Poynton, E. F.; Iskakova, D.; McMann, E.; Alsup, T. A.; Clark, T. N.; Fergusson, C. H.; Fewer, D. P.; Hughes, A. H.; McCadden, C. A.; Parra, J.; Soldatou, S.; Rudolf, J. D.; Janssen, E. M.-L.; Duncan, K. R.; Linington, R. G. The Natural Products Atlas 2.0: A Database of Microbially-Derived Natural Products. *Nucleic Acids Research* **2022**, *50* (D1), D1317–D1323. https://doi.org/10.1093/nar/gkab941.

(136) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, 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.; Kleigrewe, 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, C. A. P.; 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.; Pogliano, K.; Linington, 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. *Nature Biotechnology* **2016**, *34* (8), 828–837. https://doi.org/10.1038/nbt.3597.

(137) Allard, P.-M.; Péresse, T.; Bisson, J.; Gindro, K.; Marcourt, L.; Pham, V. C.; Roussi, F.; Litaudon, M.; Wolfender, J.-L. Integration of Molecular Networking and *In-Silico* MS/MS Fragmentation for Natural Products Dereplication. *Analytical Chemistry* **2016**, *88* (6), 3317–3323. https://doi.org/10.1021/acs.analchem.5b04804.

(138) Moumbock, A. F. A.; Gao, M.; Qaseem, A.; Li, J.; Kirchner, P. A.; Ndingkokhar, B.; Bekono, B. D.; Simoben, C. V.; Babiaka, S. B.; Malange, Y. I.; Sauter, F.; Zierep, P.; Ntie-Kang, F.; Günther, S. StreptomeDB 3.0: An Updated Compendium of Streptomycetes Natural Products. *Nucleic Acids Research* **2021**, *49* (D1), D600–D604. https://doi.org/10.1093/nar/gkaa868.

(139) Sorokina, M.; Merseburger, P.; Rajan, K.; Yirik, M. A.; Steinbeck, C. COCONUT Online: Collection of Open Natural Products Database. *Journal of Cheminformatics* **2021**, *13* (1). https://doi.org/10.1186/s13321-020-00478-9.

(140) Rutz, A.; Sorokina, M.; Galgonek, J.; Mietchen, D.; Willighagen, E.; Gaudry, A.; Graham, J. G.; Stephan, R.; Page, R.; Vondrášek, J.; Steinbeck, C.; Pauli, G. F.; Wolfender, J.-L.; Bisson, J.; Allard, P.-M. The LOTUS Initiative for Open Knowledge Management in Natural Products Research. *eLife* **2022**, *11*. https://doi.org/10.7554/eLife.70780.

(141) Zhao, H.; Yang, Y.; Wang, S.; Yang, X.; Zhou, K.; Xu, C.; Zhang, X.; Fan, J.; Hou, D.; Li, X.; Lin, H.; Tan, Y.; Wang, S.; Chu, X.-Y.; Zhuoma, D.; Zhang, F.; Ju, D.; Zeng, X.; Chen, Y. Z. NPASS Database Update 2023: Quantitative Natural Product Activity and Species Source Database for Biomedical Research. *Nucleic Acids Research* **2023**, *51* (D1), D621–D628. https://doi.org/10.1093/nar/gkac1069.

(142) Wishart, D. S.; Sayeeda, Z.; Budinski, Z.; Guo, A.; Lee, B. L.; Berjanskii, M.; Rout, M.; Peters, H.; Dizon, R.; Mah, R.; Torres-Calzada, C.; Hiebert-Giesbrecht, M.; Varshavi, D.; Varshavi, D.; Oler, E.; Allen, D.; Cao, X.; Gautam, V.; Maras, A.; Poynton, E. F.; Tavangar, P.; Yang, V.; van Santen, J. A.; Ghosh, R.; Sarma, S.; Knutson, E.; Sullivan, V.; Jystad, A. M.; Renslow, R.; Sumner, L. W.; Linington, R. G.; Cort, J. R. NP-MRD: The Natural Products Magnetic Resonance Database. *Nucleic Acids Research* **2022**, *50* (D1), D665–D677. https://doi.org/10.1093/nar/gkab1052.

(143) Lyu, C.; Chen, T.; Qiang, B.; Liu, N.; Wang, H.; Zhang, L.; Liu, Z. CMNPD: A Comprehensive Marine Natural Products Database towards Facilitating Drug Discovery from the Ocean. *Nucleic Acids Research* **2021**, *49* (D1), D509–D515. https://doi.org/10.1093/nar/gkaa763.

(144) Jones, M. R.; Pinto, E.; Torres, M. A.; Dörr, F.; Mazur-Marzec, H.; Szubert, K.; Tartaglione, L.; Dell'Aversano, C.; Miles, C. O.; Beach, D. G.; McCarron, P.; Sivonen, K.; Fewer, D. P.; Jokela, J.; Janssen, E. M.-L. CyanoMetDB, a Comprehensive Public Database of Secondary Metabolites from Cyanobacteria. *Water Research* **2021**, *196.* https://doi.org/10.1016/j.watres.2021.117017.

(145) Wishart, D. S.; Oler, E.; Peters, H.; Guo, A.; Girod, S.; Han, S.; Saha, S.; Lui, V. W.; LeVatte, M.; Gautam, V.; Kaddurah-Daouk, R.; Karu, N. MiMeDB: The Human Microbial Metabolome Database. *Nucleic Acids Research* **2023**, *51* (D1), D611–D620. https://doi.org/10.1093/nar/gkac868.

(146) Zani, C. L.; Carroll, A. R. Database for Rapid Dereplication of Known Natural Products Using Data from MS and Fast NMR Experiments. *Journal of Natural Products* **2017**, *80* (6), 1758–1766. https://doi.org/10.1021/acs.jnatprod.6b01093.

(147) Hilton, B. D.; Martin, G. E. Investigation of the Experimental Limits of Small-Sample Heteronuclear 2D NMR. *Journal of Natural Products* **2010**, *73* (9), 1465–1469. https://doi.org/10.1021/np100481m.

(148) Pin, M.; Poynton, E. F.; Jordan, T.; Kim, J.; Ledingham, B.; van Santen, J. A.; Yang, V.; Maras, A.; Tavangar, P.; Gautam, V.; Peters, H.; Sajed, T.; Lee, B. L.; Shreffler, H. A.; Koller, J. T.; Tretter, Z. M.; Cort, J. R.; Sumner, L. W.; Wishart, D. S.; Linington, R. G. A Data Deposition Platform for Sharing Nuclear Magnetic Resonance Data. *Journal of Natural Products* **2023**. https://doi.org/10.1021/acs.jnatprod.3c00795.

(149) Halabalaki, M.; Vougogiannopoulou, K.; Mikros, E.; Skaltsounis, A. L. Recent Advances and New Strategies in the NMR-Based Identification of Natural Products. *Current Opinion in Biotechnology* **2014**, *25.* https://doi.org/10.1016/j.copbio.2013.08.005.

(150) Sandvoss, M.; Preiss, A.; Levsen, K.; Weisemann, R.; Spraul, M. Two New Asterosaponins from the starfishAsterias Rubens: Application of a Cryogenic NMR Probe Head. *Magnetic Resonance in Chemistry* **2003**, *41* (11), 949–954. https://doi.org/10.1002/mrc.1277.

(151) Olson, D. L.; Norcross, J. A.; O'Neil-Johnson, M.; Molitor, P. F.; Detlefsen, D. J.; Wilson, A. G.; Peck, T. L. Microflow NMR: Concepts and Capabilities. *Analytical Chemistry* **2004**, *76* (10), 2966–2974. https://doi.org/10.1021/ac035426l.

(152) Berlinck, R. G. S.; Monteiro, A. F.; Bertonha, A. F.; Bernardi, D. I.; Gubiani, J. R.; Slivinski, J.; Michaliski, L. F.; Tonon, L. A. C.; Venancio, V. A.; Freire, V. F. Approaches for the Isolation and Identification of Hydrophilic, Light-Sensitive, Volatile and Minor Natural Products. *Natural Product Reports* **2019**, *36* (7), 981–1004. https://doi.org/10.1039/C9NP00009G.

(153) Molinski, T. F. NMR of Natural Products at the 'Nanomole-Scale.' *Natural Product Reports* **2010**, *27* (3), 321–329. https://doi.org/10.1039/B920545B.

(154) Serber, Z.; Richter, C.; Moskau, D.; Böhlen, J.-M.; Gerfin, T.; Marek, D.; Häberli, M.; Baselgia, L.; Laukien, F.; Stern, A. S.; Hoch, J. C.; Dötsch, V. New Carbon-Detected Protein NMR Experiments Using CryoProbes. *Journal of the American Chemical Society* **2000**, *122* (14), 3554–3555. https://doi.org/10.1021/ja991371m.

(155) Grimblat, N.; Zanardi, M. M.; Sarotti, A. M. Beyond DP4: An Improved Probability for the Stereochemical Assignment of Isomeric Compounds Using Quantum Chemical Calculations of NMR Shifts. *Journal of Organic Chemistry* **2015**, *80* (24), 12526–12534. https://doi.org/10.1021/acs.joc.5b02396.

(156) Grimblat, N.; Gavín, J. A.; Hernández Daranas, A.; Sarotti, A. M. Combining the Power of J Coupling and DP4 Analysis on Stereochemical Assignments: The J-DP4 Methods. *Organic Letters* **2019**, *21* (11), 4003–4007. https://doi.org/10.1021/acs.orglett.9b01193.

(157) Tsai, Y.-H.; Amichetti, M.; Zanardi, M. M.; Grimson, R.; Daranas, A. H.; Sarotti, A. M. ML-J-DP4: An Integrated Quantum Mechanics-Machine Learning Approach for Ultrafast NMR Structural Elucidation. *Organic Letters* **2022**, *24* (41), 7487–7491. https://doi.org/10.1021/acs.orglett.2c01251.

(158) Marcarino, M. O.; Passaglia, L.; Zanardi, M. M.; Sarotti, A. M. Breaking the DFT Energy Bias Caused by Intramolecular Hydrogen-Bonding Interactions with MESSI, A Structural Elucidation Method Inspired by Wisdom of Crowd Theory. *Chemistry – A European Journal* **2023**, *29* (35). https://doi.org/10.1002/chem.202300420.

(159) Kim, M. C.; Machado, H.; Jang, K. H.; Trzoss, L.; Jensen, P. R.; Fenical, W. Integration of Genomic Data with NMR Analysis Enables Assignment of the Full Stereostructure of Neaumycin B, a Potent Inhibitor of Glioblastoma from a Marine-Derived Micromonospora. *Journal of the American Chemical Society* **2018**, *140* (34), 10775–10784. https://doi.org/10.1021/jacs.8b04848.

(160) Pérez-Victoria, I.; Oves-Costales, D.; Lacret, R.; Martín, J.; Sánchez-Hidalgo, M.; Díaz, C.; Cautain, B.; Vicente, F.; Genilloud, O.; Reyes, F. Structure Elucidation and Biosynthetic Gene Cluster Analysis of Caniferolides A-D, New Bioactive 36-Membered Macrolides from the Marine-Derived: Streptomyces Caniferus CA-271066. *Organic and Biomolecular Chemistry* **2019**, *17* (11), 2954–2971. https://doi.org/10.1039/c8ob03115k.

(161) Pérez-Bonilla, M.; Oves-Costales, D.; González, I.; de la Cruz, M.; Martín, J.; Vicente, F.; Genilloud, O.; Reyes, F. Krisynomycins, Imipenem Potentiators against Methicillin-Resistant Staphylococcus Aureus, Produced by Streptomyces Canus. *Journal of Natural Products* **2020**, *83* (9), 2597–2606. https://doi.org/10.1021/acs.jnatprod.0c00294.

(162) Ortiz-López, F. J.; Carretero-Molina, D.; Sánchez-Hidalgo, M.; Martín, J.; González, I.; Román-Hurtado, F.; de la Cruz, M.; García-Fernández, S.; Reyes, F.; Deisinger, J. P.; Müller, A.; Schneider, T.; Genilloud, O. Cacaoidin, First Member of the New Lanthidin RiPP Family. *Angewandte Chemie International Edition* **2020**, *59* (31), 12654–12658. https://doi.org/10.1002/anie.202005187.

(163) Onaka, H.; Mori, Y.; Igarashi, Y.; Furumai, T. Mycolic Acid-Containing Bacteria Induce Natural-Product Biosynthesis in Streptomyces Species. *Applied and Environmental Microbiology* **2011**, *77* (2), 400–406. https://doi.org/10.1128/AEM.01337-10.

(164) Hong, K.; Gao, A. H.; Xie, Q. Y.; Gao, H.; Zhuang, L.; Lin, H. P.; Yu, H. P.; Li, J.; Yao, X. S.; Goodfellow, M.; Ruan, J. S. Actinomycetes for Marine Drug Discovery Isolated from Mangrove Soils and Plants in China. *Marine Drugs* **2009**, *7* (1), 24–44. https://doi.org/10.3390/md7010024.

(165) Marmann, A.; Aly, A. H.; Lin, W.; Wang, B.; Proksch, P. Co-Cultivation - A Powerful Emerging Tool for Enhancing the Chemical Diversity of Microorganisms. *Marine Drugs* **2014**, *12* (2), 1043–1065. https://doi.org/10.3390/md12021043.

(166) Lewis, W. H.; Tahon, G.; Geesink, P.; Sousa, D. Z.; Ettema, T. J. G. Innovations to Culturing the Uncultured Microbial Majority. *Nature Reviews Microbiology* **2021**, *19* (4), 225–240. https://doi.org/10.1038/s41579-020-00458-8.

(167) Wakefield, J.; Hassan, H. M.; Jaspars, M.; Ebel, R.; Rateb, M. E. Dual Induction of New Microbial Secondary Metabolites by Fungal Bacterial Co-Cultivation. *Frontiers in Microbiology* **2017**, *8*, 1284–1294. https://doi.org/10.3389/fmicb.2017.01284.

(168) Hertweck, C. Hidden Biosynthetictreasures Brought to Light. *Nature Chemical Biology* **2009**, *5* (7), 450–452. https://doi.org/10.1038/nchembio0709-450.

(169) Bertrand, S.; Bohni, N.; Schnee, S.; Schumpp, O.; Gindro, K.; Wolfender, J.-L. Metabolite Induction via Microorganism Co-Culture-A Potential Way to Enhance Chemical Diversity for Drug Discovery. *Biotechnology Advances Journal* **2014**, *32*, 1180–1204. https://doi.org/10.1016/j.biotechadv.2014.03.001.

(170) Li, L.; MacIntyre, L. W.; Brady, S. F. Refactoring Biosynthetic Gene Clusters for Heterologous Production of Microbial Natural Products. *Current Opinion in Biotechnology* **2021**, *69*, 145–152. https://doi.org/10.1016/j.copbio.2020.12.011.

(171) Huo, L.; Hug, J. J.; Fu, C.; Bian, X.; Zhang, Y.; Müller, R. Heterologous Expression of Bacterial Natural Product Biosynthetic Pathways. *Natural Products Reports* **2019**, *36* (10), 1412–1436. https://doi.org/10.1039/C8NP00091C.

(172) De Boer, A. L.; Schmidt-Dannert, C. Recent Efforts in Engineering Microbial Cells to Produce New Chemical Compounds. *Current Opinion in Chemical Biology* **2003**, *7* (2), 273–278. https://doi.org/10.1016/S1367-5931(03)00023-1.

(173) Selegato, D. M.; Castro-Gamboa, I. Enhancing Chemical and Biological Diversity by Co-Cultivation. *Frontiers in Microbiology* **2023**, *14*. https://doi.org/10.3389/fmicb.2023.1117559.

(174) Moon, K.; Xu, F.; Zhang, C.; Seyedsayamdost, M. R. Bioactivity-HiTES Unveils Cryptic Antibiotics Encoded in Actinomycete Bacteria. *ACS Chemical Biology* **2019**, *14* (4), 767–774. https://doi.org/10.1021/acschembio.9b00049.

(175) Bode, H. B.; Bethe, B.; Höfs, R.; Zeeck, A. Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity. *Chembiochem* **2002**, *3* (7), 619–627. https://doi.org/10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9.

(176) Rateb, M. E.; Houssen, W. E.; Harrison, W. T. A.; Deng, H.; Okoro, C. K.; Asenjo, J. A.; Andrews, B. A.; Bull, A. T.; Goodfellow, M.; Ebel, R.; Jaspars, M. Diverse Metabolic Profiles of a Streptomyces Strain Isolated from a Hyper-Arid Environment. *Journal of Natural Products* **2011**, *74* (9), 1965–1971. https://doi.org/10.1021/np200470u.

(177) Romano, S.; Jackson, S. A.; Patry, S.; Dobson, A. D. W. Extending the "One Strain Many Compounds" (OSMAC) Principle to Marine Microorganisms. *Marine Drugs* **2018**, *16* (7), 1–29. https://doi.org/10.3390/md16070244.

(178) Wei, H.; Lin, Z.; Li, D.; Gu, Q.; Zhu, T. OSMAC (One Strain Many Compounds) Approach in the Research of Microbial Metabolites--a Review. *Wei Sheng Wu Xue Bao* **2010**, *50* (6), 701–709. PMID: 20687332.

(179) Liou, G. F.; Khosla, C. Building-Block Selectivity of Polyketide Synthases. *Current Opinion in Chemical Biology* **2003**, *7* (2), 279–284. https://doi.org/10.1016/S1367-5931(03)00016-4.

(180) Lawrence, J. G.; Roth, J. R. Selfish Operons: Horizontal Transfer May Drive the Evolution of Gene Clusters. *Genetics* **1996**, *143* (4), 1843–1860. https://doi.org/10.1093/genetics/143.4.1843.

(181) Dobrindt, U.; Hochhut, B.; Hentschel, U.; Hacker, J. Genomic Islands in Pathogenic and Environmental Microorganisms. *Nature Reviews Microbiology* **2004**, *2* (5), 414–424. https://doi.org/10.1038/nrmicro884.

(182) Machado, H.; Sonnenschein, E. C.; Melchiorsen, J.; Gram, L. Genome Mining Reveals Unlocked Bioactive Potential of Marine Gram-Negative Bacteria. *BMC Genomics* **2015**, *16* (1). https://doi.org/10.1186/s12864-015-1365-z.

(183) Weber, T.; Kim, H. U. The Secondary Metabolite Bioinformatics Portal: Computational Tools to Facilitate Synthetic Biology of Secondary Metabolite Production. *Synthetic and Systems Biotechnology* **2016**, *1* (2), 69–79. https://doi.org/10.1016/j.synbio.2015.12.002.

(184) Albarano, L.; Esposito, R.; Ruocco, N.; Costantini, M. Genome Mining as New Challenge in Natural Products Discovery. *Marine Drugs* **2020**, *18* (4). https://doi.org/10.3390/md18040199.

(185) Singh, T. A.; Passari, A. K.; Jajoo, A.; Bhasin, S.; Gupta, V. K.; Hashem, A.; Alqarawi, A. A.; Abd_Allah, E. F. Tapping Into Actinobacterial Genomes for Natural Product Discovery. *Frontiers in Microbiology* **2021**, *12*. https://doi.org/10.3389/fmicb.2021.655620.

(186) Zhang, M. M.; Wong, F. T.; Wang, Y.; Luo, S.; Lim, Y. H.; Heng, E.; Yeo, W. L.; Cobb, R. E.; Enghiad, B.; Ang, F. L.; Zhao, H. CRISPR–Cas9 Strategy for Activation of Silent Streptomyces Biosynthetic Gene Clusters. *Nature Chemical Biology* **2017**, *13* (6), 607–609. https://doi.org/10.1038/nchembio.2341.

(187) Lee, N.; Hwang, S.; Kim, J.; Cho, S.; Palsson, B.; Cho, B.-K. Mini Review: Genome Mining Approaches for the Identification of Secondary Metabolite Biosynthetic Gene Clusters in Streptomyces. *Computational and Structural Biotechnology Journal* **2020**, *18*, 1548–1556. https://doi.org/10.1016/j.csbj.2020.06.024.

(188) 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. https://doi.org/10.1038/417141a.

(189) Scott, T. A.; Piel, J. The Hidden Enzymology of Bacterial Natural Product Biosynthesis. *Nature Reviews Chemistry* **2019**, *3* (7), 404–425. https://doi.org/10.1038/s41570-019-0107-1.

(190) Blin, K.; Shaw, S.; Kloosterman, A. M.; Charlop-Powers, Z.; van Wezel, G. P.; Medema, M. H.; Weber, T. antiSMASH 6.0: Improving Cluster Detection and Comparison Capabilities. *Nucleic Acids Research* **2021**, *49* (W1), W29–W35. https://doi.org/10.1093/nar/gkab335.

(191) 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 research* **2017**, *45* (W1), W49–W54. https://doi.org/10.1093/nar/gkx320.

(192) Starcevic, A.; Zucko, J.; Simunkovic, J.; Long, P. F.; Cullum, J.; Hranueli, D. ClustScan: An Integrated Program Package for the Semi-Automatic Annotation of Modular Biosynthetic Gene Clusters and in Silico Prediction of Novel Chemical Structures. *Nucleic Acids Research* **2008**, *36* (21), 6882–6892. https://doi.org/10.1093/nar/gkn685.

(193) Weber, T.; Rausch, C.; Lopez, P.; Hoof, I.; Gaykova, V.; Huson, D. H.; Wohlleben, W. CLUSEAN: A Computer-Based Framework for the Automated Analysis of Bacterial Secondary Metabolite Biosynthetic Gene Clusters. *Journal of biotechnology* **2009**, *140* (1–2), 13–17. https://doi.org/10.1016/j.jbiotec.2009.01.007.

(194) Vm, L. Major Classes of Natural Product Scaffolds and Enzymatic Biosynthetic Machinery. In *Natural Product Biosynthesis*; The Royal Society of Chemistry, **2022**; *1*, 1–21. https://doi.org/10.1039/BK9781839165641-00001.

(195) Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Grande Burgos, M. J.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; van Wezel, G. P.; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A. New Developments in RiPP Discovery, Enzymology and Engineering. *Natural Product Reports* **2020**. https://doi.org/10.1039/D0NP00027B.

(196) Risdian, C.; Mozef, T.; Wink, J. Biosynthesis of Polyketides in Streptomyces. *Microorganisms* **2019**, *7* (5). https://doi.org/10.3390/microorganisms7050124.

(197) Smith, S.; Tsai, S. The Type I Fatty Acid and Polyketide Synthases: A Tale of Two Megasynthases. *Natural Product Reports* **2007**, *24* (5), 1041–1072. https://doi.org/10.1039/b603600g.

(198) Rawlings, B. J. Type I Polyketide Biosynthesis in Bacteria (Part B). *Natural Product Reports* **2001**, *18* (3), 231–281. https://doi.org/10.1039/b100191o.

(199) Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A. Type II Polyketide Synthases: Gaining a Deeper Insight into Enzymatic Teamwork. *Natural Product Reports* **2007**, *24* (1), 162–190. https://doi.org/10.1039/b507395m.

(200) Austin, M. B.; Noel, J. P. The Chalcone Synthase Superfamily of Type III Polyketide Synthases. *Natural Product Reports* **2003**, *20* (1), 79–110. https://doi.org/10.1039/b100917f.

(201) Shen, B. Polyketide Biosynthesis beyond the Type I, II and III Polyketide Synthase Paradigms. *Current Opinion in Chemical Biology* **2003**, *7* (2), 285–295. https://doi.org/10.1016/S1367-5931(03)00020-6.

(202) Sherman, D. H. The Lego-Ization of Polyketide Biosynthesis. *Nature Biotechnology* **2005**, *23* (9), 1083–1084. https://doi.org/10.1038/nbt0905-1083.

(203) El-Sayed, A. K.; Hothersall, J.; Cooper, S. M.; Stephens, E.; Simpson, T. J.; Thomas, C. M. Characterization of the Mupirocin Biosynthesis Gene Cluster from Pseudomonas Fluorescens NCIMB 10586. *Chemistry & Biology* **2003**, *10* (5), 419–430. https://doi.org/10.1016/S1074-5521(03)00091-7.

(204) Mast, Y.; Weber, T.; Gölz, M.; Ort-Winklbauer, R.; Gondran, A.; Wohlleben, W.; Schinko, E. Characterization of the 'Pristinamycin Supercluster' of Streptomyces Pristinaespiralis. *Microbial Biotechnology* **2011**, *4* (2), 192–206. https://doi.org/10.1111/j.1751-7915.2010.00213.x.

(205) Motamedi, H.; Hutchinson, C. R. Cloning and Heterologous Expression of a Gene Cluster for the Biosynthesis of Tetracenomycin C, the Anthracycline Antitumor Antibiotic of Streptomyces Glaucescens. *Proceedings of the National Academy of Sciences* **1987**, *84* (13), 4445–4449. https://doi.org/10.1073/pnas.84.13.4445.

(206) Ridley, C. P.; Ho, Y. L.; Khosla, C. Evolution of Polyketide Synthases in Bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105* (12), 4595–4600. https://doi.org/10.1073/pnas.0710107105.

(207) Moore, B. S.; Hopke, J. N. Discovery of a New Bacterial Polyketide Biosynthetic Pathway. *ChemBioChem* **2001**, *2* (1), 35–38. https://doi.org/10.1002/1439-7633(20010105)2:1<35::AID-CBIC35>3.3.CO;2-T.

(208) Baltz, R. H.; Seno, E. T. Genetics of Streptomyces Fradiae and Tylosin Biosynthesis. *Annual Review of Microbiology* **1988**, *42*, 547–574. https://doi.org/10.1146/annurev.mi.42.100188.002555.

(209) Zhang, H.; Wang, Y.; Wu, J.; Skalina, K.; Pfeifer, B. A. Complete Biosynthesis of Erythromycin A and Designed Analogs Using E. Coli as a Heterologous Host. *Chemistry & Biology* **2010**, *17* (11), 1232–1240. https://doi.org/10.1016/j.chembiol.2010.09.013.

(210) Choi, S.-S.; Hur, Y.-A.; Sherman, D. H.; Kim, E.-S. Isolation of the Biosynthetic Gene Cluster for Tautomycetin, a Linear Polyketide T Cell-Specific Immunomodulator from Streptomyces Sp. CK4412. *Microbiology* **2007**, *153* (4), 1095–1102. https://doi.org/10.1099/mic.0.2006/003194-0.

(211) Hopwood, D. A. Genetic Contributions to Understanding Polyketide Synthases. *Chemical Reviews* **1997**, *97* (7), 2465–2498. https://doi.org/10.1021/cr960034i.

(212) Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angewandte Chemie International Edition* **2009**, *48* (26), 4688–4716. https://doi.org/10.1002/anie.200806121.

(213) Rix, U.; Fischer, C.; Remsing, L. L.; Rohr, J. Modification of Post-PKS Tailoring Steps through Combinatorial Biosynthesis. *Natural Product Reports* **2002**, *19* (5), 542–580. https://doi.org/10.1039/b103920m.

(214) Kohli, R. M.; Walsh, C. T. Enzymology of Acyl Chain Macrocyclization in Natural Product Biosynthesis. *Chemical Communications* **2003**, *3* (3), 297–307. https://doi.org/10.1039/b208333g.

(215) Hong, H.; Fill, T.; Leadlay, P. F. A Common Origin for Guanidinobutanoate Starter Units in Antifungal Natural Products. *Angewandte Chemie International Edition* **2013**, *52* (49), 13096–13099. https://doi.org/10.1002/anie.201308136.

(216) Moore, B. S.; Hertweck, C. Biosynthesis and Attachment of Novel Bacterial Polyketide Synthase Starter Units. *Natural Product Reports* **2002**, *19* (1), 70–99. https://doi.org/10.1039/b003939j.

(217) Kornfuehrer, T.; Eustáquio, A. S. Diversification of Polyketide Structures via Synthase Engineering. *Medicinal Chemistry Communications* **2019**, *10* (8), 1256–1272. https://doi.org/10.1039/C9MD00141G.

(218) Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R. H.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structure of a Modular Polyketide Synthase. *Nature* **2014**, *510* (7506), 512–517. https://doi.org/10.1038/nature13423.

(219) Schwarzer, D.; Finking, R.; Marahiel, M. A. Nonribosomal Peptides: From Genes to Products. *Natural Product Reports* **2003**, *20* (3), 275-287. https://doi.org/10.1039/b111145k.

(220) Süssmuth, R. D.; Mainz, A. Nonribosomal Peptide Synthesis-Principles and Prospects. *Angewandte Chemie International Edition* **2017**, *56* (14), 3770–3821. https://doi.org/10.1002/anie.201609079.

(221) Sieber, S. A.; Marahiel, M. A. Molecular Mechanisms Underlying Nonribosomal Peptide Synthesis: Approaches to New Antibiotics. *Chemical Reviews* **2005**, *105* (2), 715–738. https://doi.org/10.1021/cr0301191.

(222) McErlean, M.; Overbay, J.; Van Lanen, S. Refining and Expanding Nonribosomal Peptide Synthetase Function and Mechanism. *Journal of Industrial Microbiology and Biotechnology* **2019**, *46* (3–4), 493–513. https://doi.org/10.1007/s10295-018-02130-w.

(223) Winn, M.; Fyans, J. K.; Zhuo, Y.; Micklefield, J. Recent Advances in Engineering Nonribosomal Peptide Assembly Lines. *Natural Product Reports* **2016**, *33* (2), 317–347. https://doi.org/10.1039/C5NP00099H.

(224) Duban, M.; Cociancich, S.; Leclère, V. Nonribosomal Peptide Synthesis Definitely Working Out of the Rules. *Microorganisms* **2022**, *10* (3). https://doi.org/10.3390/microorganisms10030577.

(225) Rudolf, J. D.; Alsup, T. A.; Xu, B.; Li, Z. Bacterial Terpenome. *Natural Product Reports* **2021**, *38* (5), 905–980. https://doi.org/10.1039/D0NP00066C.

(226) Helfrich, E. J. N.; Lin, G.-M.; Voigt, C. A.; Clardy, J. Bacterial Terpene Biosynthesis: Challenges and Opportunities for Pathway Engineering. *Beilstein Journal of Organic Chemistry* **2019**, *15*, 2889–2906. https://doi.org/10.3762/bjoc.15.283.

(227) Karunanithi, P. S.; Zerbe, P. Terpene Synthases as Metabolic Gatekeepers in the Evolution of Plant Terpenoid Chemical Diversity. *Frontiers in Plant Science* **2019**, *10*. https://doi.org/10.3389/fpls.2019.01166.

(228) Kirby, J.; Keasling, J. D. Biosynthesis of Plant Isoprenoids: Perspectives for Microbial Engineering. *Annual Review of Plant Biology* **2009**, *60* (1), 335–355. https://doi.org/10.1146/annurev.arplant.043008.091955.

(229) Murphy, A. C.; Gao, S.-S.; Han, L.-C.; Carobene, S.; Fukuda, D.; Song, Z.; Hothersall, J.; Cox, R. J.; Crosby, J.; Crump, M. P.; Thomas, C. M.; Willis, C. L.; Simpson, T. J. Biosynthesis of Thiomarinol A and Related Metabolites of Pseudoalteromonas Sp. SANK 73390. *Chemical Science* **2014**, *5* (1), 397–402. https://doi.org/10.1039/C3SC52281D.

(230) Qu, Y.; Safonova, O.; De Luca, V. Completion of the Canonical Pathway for Assembly of Anticancer Drugs Vincristine/Vinblastine in Catharanthus Roseus. *The Plant Journal* **2019**, *97* (2), 257–266. https://doi.org/10.1111/tpj.14111.

(231) Grigalunas, M.; Brakmann, S.; Waldmann, H. Chemical Evolution of Natural Product Structure. *Journal of the American Chemical Society* **2022**, *144* (8), 3314–3329. https://doi.org/10.1021/jacs.1c11270.

(232) Park, S. R.; Yoo, Y. J.; Ban, Y.-H.; Yoon, Y. J. Biosynthesis of Rapamycin and Its Regulation: Past Achievements and Recent Progress. *The Journal of Antibiotics* **2010**, *63* (8), 434–441. https://doi.org/10.1038/ja.2010.71.

(233) Schor, R.; Schotte, C.; Wibberg, D.; Kalinowski, J.; Cox, R. J. Three Previously Unrecognised Classes of Biosynthetic Enzymes Revealed during the Production of Xenovulene A. *Nature Communications* **2018**, *9* (1), 1963–1971. https://doi.org/10.1038/s41467-018-04364-9.

(234) Shiomi, K.; Iinuma, H.; Hamada, M.; Naganawa, H.; Manabe, M.; Matsuki, C.; Takeuchi, T.; Umezawa, H. Novel Antibiotics Napyradiomycins. Production, Isolation, Physico-Chemical Properties and Biological Activity. *The Journal of Antibiotics* **1986**, *39* (4), 487–493. https://doi.org/10.7164/antibiotics.39.487.

(235) Fukuda, D. S.; Mynderse, J. S.; Baker, P. J.; Berry, D. M.; BOECK, L. D.; Yao, R. C.; Mertz, F. P.; Nakatsukasa, W. M.; Mabe, J.; Ott, J.; Counter, F. T.; Ensminger, P. W.; Allen, N. E.; Alborn, W. E.; Hobbs, J. N. A80915, a New Antibiotic Complex Produced by Streptomyces Aculeolatus. Discovery, Taxonomy, Fermentation, Isolation, Characterization, and Antibacterial Evaluation. *The Journal of Antibiotics* **1990**, *43* (6), 623–633. https://doi.org/10.7164/antibiotics.43.623.

(236) Soria-Mercado, I. E.; Prieto-Davo, A.; Jensen, P. R.; Fenical, W. Antibiotic Terpenoid Chloro-Dihydroquinones from a New Marine Actinomycete. *Journal of Natural Products* **2005**, *68* (6), 904–910. https://doi.org/10.1021/np058011z.

(237) Wu, Z.; Li, S.; Li, J.; Chen, Y.; Saurav, K.; Zhang, Q.; Zhang, H.; Zhang, W.; Zhang, W.; Zhang, S.; Zhang, C. Antibacterial and Cytotoxic New Napyradiomycins from the Marine-Derived Streptomyces Sp. SCSIO 10428. *Marine Drugs* **2013**, *11* (6), 2113–2125. https://doi.org/10.3390/md11062113.

(238) Shiomi, K.; Nakamura, H.; Iinuma, H.; Naganawa, H.; Takeuchi, T.; Umezawa, H.; Iitaka, Y. New Antibiotic Napyradiomycins A2 and B4 and Stereochemistry of Napyradiomycins. *The Journal of Antibiotics* **1987**, *40* (9), 1213–1219. https://doi.org/10.7164/antibiotics.40.1213.

(239) Farnaes, L.; La Clair, J. J.; Fenical, W. Napyradiomycins CNQ525.510B and A80915C Target the Hsp90 Paralogue Grp94. *Organic & Biomolecular Chemistry* **2014**, *12* (3), 418–423. https://doi.org/10.1039/C3OB41355A.

(240) Shiomi, K.; Nakamura, H.; Iinuma, H.; Naganawa, H.; ISSHIKI, K.; Takeuchi, T.; Umezawa, H.; Iitaka, Y. Structures of New Antibiotics Napyradiomycins. *The Journal of Antibiotics* **1986**, *39* (4), 494–501. https://doi.org/10.7164/antibiotics.39.494.

(241) Lacret, R.; Pérez-Victoria, I.; Oves-Costales, D.; De la Cruz, M.; Domingo, E.; Martín, J.; Díaz, C.; Vicente, F.; Genilloud, O.; Reyes, F. MDN-0170, a New Napyradiomycin from Streptomyces Sp. Strain CA-271078. *Marine Drugs* **2016**, *14* (10). https://doi.org/10.3390/md14100188.

(242) Winter, J. M.; Moffitt, M. C.; Zazopoulos, E.; McAlpine, J. B.; Dorrestein, P. C.; Moore, B. S. Molecular Basis For Chloronium-Mediated Meroterpene Cyclization: Cloning, Sequencing, And Heterologous Expression Of The Napyradiomycin Biosynthetic Gene Cluster. *Journal of Biological Chemistry* **2007**, *282* (22), 16362–16368. https://doi.org/10.1074/jbc.M611046200.

(243) McKinnie, S. M. K.; Miles, Z. D.; Jordan, P. A.; Awakawa, T.; Pepper, H. P.; Murray, L. A. M.; George, J. H.; Moore, B. S. Total Enzyme Syntheses of Napyradiomycins A1 and B1. *Journal of the American Chemical Society* **2018**, *140* (51), 17840–17845. https://doi.org/10.1021/jacs.8b10134.

(244) Luo, Y.; Cobb, R. E.; Zhao, H. Recent Advances in Natural Product Discovery. *Current Opinion in Biotechnology* **2014**, *30*, 230–237. https://doi.org/10.1016/j.copbio.2014.09.002.

(245) Subko, K. Natural Product Discovery: Top-down vs. Bottom-up Approach, DTU Bioengineering, Kgs. Lyngby, Denmark, **2020**. https://orbit.dtu.dk/en/publications/natural-product-discovery-top-down-vs-bottom-up-approach.

(246) Haste, N. M.; Farnaes, L.; Perera, V. R.; Fenical, W.; Nizet, V.; Hensler, M. E. Bactericidal Kinetics of Marine-Derived Napyradiomycins against Contemporary Methicillin-Resistant Staphylococcus Aureus. *Marine Drugs* **2011**, *9* (4), 680–689. https://doi.org/10.3390/md9040680.

(247) Motohashi, K.; Sue, M.; Furihata, K.; Ito, S.; Seto, H. Terpenoids Produced by Actinomycetes: Napyradiomycins from Streptomyces Antimycoticus NT17. *Journal of Natural Products* **2008**, *71* (4), 595–601. https://doi.org/10.1021/np070575a.

(248) Henkel, T.; Zeeck, A. Secondary Metabolites by Chemical Screening. 15. Structure and Absolute Configuration of Naphthomevalin, a New Dihydro-Naphthoquinone Antibiotic from Streptomyces Sp. *The Journal of Antibiotics* **1991**, *44* (6), 665–669. https://doi.org/10.7164/antibiotics.44.665.

(249) Soria-Mercado, I. E.; Jensen, P. R.; Fenical, W.; Kassel, S.; Golen, J. 3,4a-Di-chloro-10a-(3-Chloro-6-Hydroxy-2,2,6-Tri-methyl-cyclo-hexyl-methyl)-6,8-Di-hydroxy-2,2,7-Tri-methyl-3,4,4a,10a-Tetra-hydro-2H-Benzo-[g]-chromene-5,10-Dione. *Acta Crystallographica Section E* **2004**, *60* (9), o1627–o1629. https://doi.org/10.1107/S1600536804020094.

(250) Gomi, S.; Ohuchi, S.; Sasaki, T.; Itoh, J.; Sezaki, M. Studies on New Antibiotics SF2415. II. The Structural Elucidation. *The Journal of Antibiotics* **1987**, *40* (6), 740–749. https://doi.org/10.7164/antibiotics.40.740.

(251) Umezawa, K.; Masuoka, S.; Ohse, T.; Naganawa, H.; Kondo, S.; Ikeda, Y.; Kinoshita, N.; Hamada, M.; Sawa, T.; Takeuchi, T. Isolation from Streptomyces of a Novel Naphthoquinone Compound, Naphthablin, That Inhibits Abl Oncogene Functions. *The Journal of Antibiotics* **1995**, *48* (7), 604–607. https://doi.org/10.7164/antibiotics.48.604.

(252) Cho, J. Y.; Kwon, H. C.; Williams, P. G.; Jensen, P. R.; Fenical, W. Azamerone, a Terpenoid Phthalazinone from a Marine-Derived Bacterium Related to the Genus Streptomyces (Actinomycetales). *Orgic Letters* **2006**, *8* (12), 2471–2474. https://doi.org/10.1021/ol060630r.

(253) Motohashi, K.; Irie, K.; Toda, T.; Matsuo, Y.; Kasai, H.; Sue, M.; Furihata, K.; Seto, H. Studies on Terpenoids Produced by Actinomycetes. *The Journal of Antibiotics* **2008**, *61* (2), 75–80. https://doi.org/10.1038/ja.2008.113.

(254) Farnaes, L.; Coufal, N. G.; Kauffman, C. A.; Rheingold, A. L.; DiPasquale, A. G.; Jensen, P. R.; Fenical, W. Napyradiomycin Derivatives, Produced by a Marine-Derived Actinomycete, Illustrate Cytotoxicity by Induction of Apoptosis. *Journal of Natural Products* **2014**, *77* (1), 15–21. https://doi.org/10.1021/np400466j.

(255) Hori, Y.; Abe, Y.; Shigematsu, N.; Goto, T.; Okuhara, M.; Kohsaka, M. Napyradiomycins A and B1: Non-Steroidal Estrogen-Receptor Antagonists Produced by a Streptomyces. *The Journal of Antibiotics* **1993**, *46* (12), 1890–1893. https://doi.org/10.7164/antibiotics.46.1890.

(256) Dantzig, A. H.; Minor, P. L.; Garrigus, J. L.; Fukuda, D. S.; Mynderse, J. S. Studies on the Mechanism of Action of A80915A, a Semi-Naphthoquinone Natural Product, as an Inhibitor of Gastric (H+-K+)-ATPase. *Biochemical Pharmacology* **1991**, *42* (10), 2019–2026. https://doi.org/10.1016/0006-2952(91)90603-3.

(257) Prestinaci, F.; Pezzotti, P.; Pantosti, A. Antimicrobial Resistance: A Global Multifaceted Phenomenon. *Pathogens and Global Health* **2015**, *109* (7), 309–318. https://doi.org/10.1179/2047773215Y.0000000030.

(258) Carroll, A. R.; Copp, B. R.; Davis, R. A.; Keyzers, R. A.; Prinsep, M. R. Marine Natural Products. *Natural Product Reports* **2019**, *36* (1), 122–173. https://doi.org/10.1039/C8NP00092A.

(259) Hassan, S. S. ul; Shaikh, A. L. Marine Actinobacteria as a Drug Treasure House. *Biomedicine & Pharmacotherapy* **2017**, *87*, 46–57. https://doi.org/10.1016/j.biopha.2016.12.086.

(260) Lacret, R.; Oves-Costales, D.; Gómez, C.; Díaz, C.; De la Cruz, M.; Pérez-Victoria, I.; Vicente, F.; Genilloud, O.; Reyes, F. New Ikarugamycin Derivatives with Antifungal and Antibacterial Properties from Streptomyces Zhaozhouensis. *Marine Drugs* **2015**, *13* (1), 128–140. https://doi.org/10.3390/md13010128.

(261) Perez-Victoria, I.; Martin, J.; Reyes, F. Combined LC/UV/MS and NMR Strategies for the Dereplication of Marine Natural Products. *Planta Medica* **2016**, *82* (9–10), 857–871. https://doi.org/10.1055/s-0035-1568644.

(262) Chapman & Hall/CRC. *Dictionary of Natural Products (DNP)*. CHEMnetBASE. https://dnp.chemnetbase.com/chemical/ChemicalSearch. (accessed 2019-11-26).

(263) Cheng, Y.-B.; Jensen, P. R.; Fenical, W. Cytotoxic and Antimicrobial Napyradiomycins from Two Marine-Derived Streptomyces Strains. *European Journal of Organic Chemistry* **2013**, *2013* (18), 3751–3757. https://doi.org/10.1002/ejoc.201300349.

(264) Kamimura, D.; Yamada, K.; Tsuji, T. Agent for Inhibiting Production of Venous Cell-Adhering Molecule-1 and Napyradiomycin SC. 09110689. Japanese Patent. **1997** Apr 28

(265) Barbeau, X.; Vincent, A. T.; Lagüe, P. ConfBuster: Open-Source Tools for Macrocycle Conformational Search and Analysis. *Journal of Open Research Software* **2018**, *6* (1). https://doi.org/10.5334/jors.189.

(266) Bernhardt, P.; Okino, T.; Winter, J. M.; Miyanaga, A.; Moore, B. S. A Stereoselective Vanadium-Dependent Chloroperoxidase in Bacterial Antibiotic Biosynthesis. *Journal of the American Chemical Society* **2011**, *133* (12), 4268–4270. https://doi.org/10.1021/ja201088k.

(267) Wever, R.; Krenn, B. E.; Renirie, R. Chapter Six - Marine Vanadium-Dependent Haloperoxidases, Their Isolation, Characterization, and Application. In *Methods in Enzymology*; Moore, B. S., Ed.; Academic Press, **2018**; *605*, 141–201. https://doi.org/10.1016/bs.mie.2018.02.026.

(268) Diethelm, S.; Teufel, R.; Kaysser, L.; Moore, B. S. A Multitasking Vanadium-Dependent Chloroperoxidase as an Inspiration for the Chemical Synthesis of the Merochlorins. *Angewandte Chemie International Edition* **2014**, *53* (41), 11023–11026. https://doi.org/10.1002/anie.201405696.

(269) McKinnie, S. M. K.; Miles, Z. D.; Moore, B. S. Chapter Thirteen - Characterization and Biochemical Assays of Streptomyces Vanadium-Dependent Chloroperoxidases. In *Methods in Enzymology*; Moore, B. S., Ed.; Academic Press, **2018**; *604*, 405–424. https://doi.org/10.1016/bs.mie.2018.02.016.

(270) Kaysser, L.; Bernhardt, P.; Nam, S.-J.; Loesgen, S.; Ruby, J. G.; Skewes-Cox, P.; Jensen, P. R.; Fenical, W.; Moore, B. S. Merochlorins A–D, Cyclic Meroterpenoid Antibiotics Biosynthesized in Divergent Pathways with Vanadium-Dependent Chloroperoxidases. *Journal of the American Chemical Society* **2012**, *134* (29), 11988–11991. https://doi.org/10.1021/ja305665f.

(271) Blin, K.; Shaw, S.; Augustijn, H. E.; Reitz, Z. L.; Biermann, F.; Alanjary, M.; Fetter, A.; Terlouw, B. R.; Metcalf, W. W.; Helfrich, E. J. N.; van Wezel, G. P.; Medema, M. H.; Weber, T. antiSMASH 7.0: New and Improved Predictions for Detection, Regulation, Chemical Structures and Visualisation. *Nucleic Acids Research* **2023**, *51* (W1), W46–W50. https://doi.org/10.1093/nar/gkad344.

(272) Murray, L. A. M.; McKinnie, S. M. K.; Moore, B. S.; George, J. H. Meroterpenoid Natural Products from Streptomyces Bacteria – the Evolution of Chemoenzymatic Syntheses. *Natural Product Reports* **2020**, *37* (10), 1334–1366. https://doi.org/10.1039/D0NP00018C.

(273) Martín, J.; Da Sousa, T. S.; Crespo, G.; Palomo, S.; González, I.; Tormo, J. R.; De La Cruz, M.; Anderson, M.; Hill, R. T.; Vicente, F.; Genilloud, O.; Reyes, F. Kocurin, the True Structure of PM181104, an Anti-Methicillin-Resistant Staphylococcus Aureus (MRSA) Thiazolyl Peptide from the Marine-Derived Bacterium Kocuria Palustris. *Marine Drugs* **2013**, *11* (2), 387–398. https://doi.org/10.3390/md11020387.

(274) Monteiro, M. C.; De La Cruz, M.; Cantizani, J.; Moreno, C.; Tormo, J. R.; Mellado, E.; De Lucas, J. R.; Asensio, F.; Valiante, V.; Brakhage, A. A.; Latgé, J. P.; Genilloud, O.; Vicente, F. A New Approach to Drug Discovery: High-Throughput Screening of Microbial Natural Extracts against Aspergillus Fumigatus Using Resazurin. *Journal of Biomolecular Screening* **2012**, *17* (4), 542–549. https://doi.org/10.1177/1087057111433459.

(275) Zhang, L.; Ravipati, A. S.; Koyyalamudi, S. R.; Jeong, S. C.; Reddy, N.; Bartlett, J.; Smith, P. T.; de la Cruz, M.; Monteiro, M. C.; Melguizo, Á.; Jiménez, E.; Vicente, F. Anti-Fungal and Anti-Bacterial Activities of Ethanol Extracts of Selected Traditional Chinese Medicinal Herbs. *Asian Pacific Journal of Tropical Medicine* **2013**, *6* (9), 673–681. https://doi.org/10.1016/S1995-7645(13)60117-0.

(276) Zhang, J. H.; Chung, T. D. Y.; Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular Screening* **1999**, *4* (2), 67–73. https://doi.org/10.1177/108705719900400206.

(277) Abdalla, M. A.; McGaw, L. J. Natural Cyclic Peptides as an Attractive Modality for Therapeutics: A Mini Review. *Molecules* **2018**, *23* (8). https://doi.org/10.3390/molecules23082080.

(278) Jang, J. P.; Hwang, G. J.; Kwon, M. C.; Ryoo, I. J.; Jang, M.; Takahashi, S.; Ko, S. K.; Osada, H.; Jang, J. H.; Ahn, J. S. Pentaminomycins A and B, Hydroxyarginine-Containing Cyclic Pentapeptides from Streptomyces Sp. RK88-1441. *Journal of Natural Products* **2018**, *81* (4), 806–810. https://doi.org/10.1021/acs.jnatprod.7b00882.

(279) Kaweewan, I.; Hemmi, H.; Komaki, H.; Kodani, S. Isolation and Structure Determination of a New Antibacterial Peptide Pentaminomycin C from Streptomyces Cacaoi Subsp. Cacaoi. *Journal of Antibiotics* **2020**, *73* (4), 224–229. https://doi.org/10.1038/s41429-019-0272-y.

(280) Hwang, S.; Luu Le, L. T. H.; Jo, S. II; Shin, J.; Lee, M. J.; Oh, D. C. Pentaminomycins c–e: Cyclic Pentapeptides as Autophagy Inducers from a Mealworm Beetle Gut Bacterium. *Microorganisms* **2020**, *8* (9). https://doi.org/10.3390/microorganisms8091390.

(281) Mlyata, S.; Hashimoto, M.; Fujie, K.; Nlshikawa, M.; Klyoto, S.; Okuhara, M.; Kohsaka, M. Ws-7338, New Endothelin Receptor Antagonists Isolated from Streptomyces Sp. No. 7338: II. Biological Characterization and Pharmacological Characterization of Ws-7338 b. *The Journal of Antibiotics* **1992**, *45* (1), 83–87. https://doi.org/10.7164/antibiotics.45.83.

(282) de la Cruz, M.; González, I.; Parish, C. A.; Onishi, R.; Tormo, J. R.; Martín, J.; Peláez, F.; Zink, D.; El Aouad, N.; Reyes, F.; Genilloud, O.; Vicente, F. Production of Ramoplanin and Ramoplanin Analogs by Actinomycetes. *Frontiers in Microbiology* **2017**, *8*. https://doi.org/10.3389/fmicb.2017.00343.

(283) Marvin 17.21.0, ChemAxon (https://www.chemaxon.com).

(284) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. A Nonempirical Method Using LC/MS for Determination of the Absolute Configuration of Constituent Amino Acids in a Peptide: Combination of Marfey's Method with Mass Spectrometry and Its Practical Application. *Analytical Chemistry* **1997**, *69* (24), 5146–5151. https://doi.org/10.1021/ac970289b.

(285) Joergensen, L.; Thestrup, H. N. Determination of Amino Acids in Biomass and Protein Samples by Microwave Hydrolysis and Ion-Exchange Chromatography. *Journal of Chromatography A* **1995**, *706* (1–2), 421–428. https://doi.org/10.1016/0021-9673(94)01107-P.

(286) Inglis, A.; Nicholls, P.; Roxburgh, C. Hydrolysis of the Peptide Bond and Amino Acid Modification with Hydriodic Acid. *Australian Journal of Biological Sciences* **1971**, *24* (4), 1235. https://doi.org/10.1071/BI9711235.

(287) Iwata, M. β-(2-Pyridyl)-L-α-Alanine from Streptomyces. *Sci Rep Meiji Seika Kaisha* **1988**, *27:63–66.*

(288) Prieto, C.; García-estrada, C.; Lorenzana, D.; Martín, J. F. NRPSSP: Non-Ribosomal Peptide Synthase Substrate Predictor. *Bioinformatics* **2012**, *28* (3), 426–*427.* https://doi.org/10.1093/bioinformatics/btr659.

(289) 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 Research* **2019**, *47* (W1), W81–W87. https://doi.org/10.1093/nar/gkz310.

(290) Román-Hurtado, F.; Sánchez-Hidalgo, M.; Martín, J.; Ortiz-López, F. J.; Carretero-Molina, D.; Reyes, F.; Genilloud, O. One Pathway, Two Cyclic Non-Ribosomal Pentapeptides: Heterologous Expression of BE-18257 Antibiotics and Pentaminomycins from Streptomyces Cacaoi CA-170360. *Microorganisms* **2021**, *9* (1). https://doi.org/10.3390/microorganisms9010135.

(291) Lohman, J. R.; Huang, S. X.; Horsman, G. P.; Dilfer, P. E.; Huang, T.; Chen, Y.; Wendt-Pienkowski, E.; Shen, B. Cloning and Sequencing of the Kedarcidin Biosynthetic Gene Cluster from Streptoalloteichu*s* Sp. ATCC 53650 Revealing New Insights into Biosynthesis of the Enediyne Family of Antitumor Antibiotics. *Molecular BioSystems* **2013**, *9* (3), 478–491. https://doi.org/10.1039/c3mb25523a.

(292) Kim, O. S.; Cho, Y. J.; Lee, K.; Yoon, S. H.; Kim, M.; Na, H.; Park, S. C.; Jeon, Y. S.; Lee, J. H.; Yi, H.; Won, S.; Chun, J. Introducing EzTaxon-e: A Prokaryotic 16s rRNA Gene Sequence Database with Phylotypes That Represent Uncultured Species. *International Journal of Systematic and Evolutionary Microbiology* **2012**, *62* (Pt_3), 716–721. https://doi.org/10.1099/ijs.0.038075-0.

(293) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Practical Streptomyces Genetics.*; John Innes Foundation, Norwich, **2000**; *3.*

(294) Carretero-Molina, D.; Ortiz-López, F. J.; Martín, J.; Oves-Costales, D.; Díaz, C.; De La Cruz, M.; Cautain, B.; Vicente, F.; Genilloud, O.; Reyes, F. New Napyradiomycin Analogues from Streptomyces Sp. Strain CA-271078. *Marine Drugs* **2020**, *18* (1). https://doi.org/10.3390/md18010022.

(295) Omura, S. *Macrolide Antibiotics*; Elsevier, **2003**. https://doi.org/10.1016/B978-0-12-526451-8.X5000-0.

(296) Robertsen, H. L.; Musiol-Kroll, E. M. Actinomycete-Derived Polyketides as a Source of Antibiotics and Lead Structures for the Development of New Antimicrobial Drugs. *Antibiotics* **2019**, *8* (4). https://doi.org/10.3390/antibiotics8040157.

(297) Dinos, G. P. The Macrolide Antibiotic Renaissance. *British Journal of Pharmacology* **2017**, *174* (18), 2967–2983. https://doi.org/10.1111/bph.13936.

(298) Hashimoto, T.; Hashimoto, J.; Kozone, I.; Amagai, K.; Kawahara, T.; Takahashi, S.; Ikeda, H.; Shin-Ya, K. Biosynthesis of Quinolidomicin, the Largest Known Macrolide of Terrestrial Origin: Identification and Heterologous Expression of a Biosynthetic Gene Cluster over 200 Kb. *Organic Letters* **2018**, *20* (24), 7996–7999. https://doi.org/10.1021/acs.orglett.8b03570.

(299) Laureti, L.; Song, L.; Huang, S.; Corre, C.; Leblond, P.; Challis, G. L.; Aigle, B. Identification of a Bioactive 51-Membered Macrolide Complex by Activation of a Silent Polyketide Synthase in Streptomyces Ambofaciens. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108* (15), 6258–6263. https://doi.org/10.1073/pnas.1019077108.

(300) Komatsu, K.; Tsuda, M.; Tanaka, Y.; Mikami, Y.; Kobayashi, J. Absolute Stereochemistry of Immunosuppressive Macrolide Brasilinolide A and Its New Congener Brasilinolide C. *Journal of Organic Chemistry* **2004**, *69* (5), 1535–1541. https://doi.org/10.1021/jo035773v.

(301) Essig, S.; Menche, D. Stereochemistry and Total Synthesis of Complex Myxobacterial Macrolides. *Pure and Applied Chemistry* **2013**, *85* (6), 1103–1120. https://doi.org/10.1351/PAC-CON-12-09-12.

(302) Paterson, I.; Housden, M. P.; Cordier, C. J.; Burton, P. M.; Mühlthau, F. A.; Loiseleur, O. Synthetic Studies toward the Brasilinolides: Controlled Assembly of a Protected C1-C38 Polyol Based on Fragment Union by Complex Aldol Reactions. *Organic and Biomolecular Chemistry* **2015**, *13* (20), 5716–5733. https://doi.org/10.1039/c5ob00498e.

(303) Takeuchi, T.; Hatano, M.; Umekita, M.; Hayashi, C.; Wada, S. I.; Nagayoshi, M.; Sawa, R.; Kubota, Y.; Kawada, M.; Igarashi, M.; Shibasaki, M. ATP Depletion Assay Led to the Isolation of New 36-Membered Polyol Macrolides Deplelides A and B from Streptomyces Sp. MM581-NF15. *Organic Letters* **2017**, *19* (16), 4207–4210. https://doi.org/10.1021/acs.orglett.7b01807.

(304) Hu, Y.; Wang, M.; Wu, C.; Tan, Y.; Li, J.; Hao, X.; Duan, Y.; Guan, Y.; Shang, X.; Wang, Y.; Xiao, C.; Gan, M. Identification and Proposed Relative and Absolute Configurations of Niphimycins C-E from the Marine-Derived Streptomyces Sp. IMB7-145 by Genomic Analysis. *Journal of Natural Products* **2018**, *81* (1), 178–187. https://doi.org/10.1021/acs.jnatprod.7b00859.

(305) Kim, M. C.; Winter, J. M.; Cullum, R.; Li, Z.; Fenical, W. Complementary Genomic, Bioinformatics, and Chemical Approaches Facilitate the Absolute Structure Assignment of Ionostatin, a Linear Polyketide from a Rare Marine-Derived Actinomycete. *ACS Chemical Biology* **2020**, *15* (9), 2507–2515. https://doi.org/10.1021/acschembio.0c00526.

(306) Lee, S. R.; Guo, H.; Yu, J. S.; Park, M.; Dahse, H.-M.; Jung, W. H.; Beemelmanns, C.; Kim, K. H. Revised Structural Assignment of Azalomycins Based on Genomic and Chemical Analysis. *Organic Chemistry Frontiers* **2021**, *8* (17), 4791–4798. https://doi.org/10.1039/d1qo00610j.

(307) Rho, J. R.; Subramaniam, G.; Choi, H.; Kim, E. H.; Ng, S. P.; Yoganathan, K.; Ng, S.; Buss, A. D.; Butler, M. S.; Gerwick, W. H. Gargantulide A, a Complex 52-Membered Macrolactone Showing Antibacterial Activity from Streptomyces Sp. *Organic Letters* **2015**, *17* (6), 1377–1380. https://doi.org/10.1021/acs.orglett.5b00068.

(308) Gorin, P. A. J.; Mazurek, M. Further Studies on the Assignment of Signals in 13 C Magnetic Resonance Spectra of Aldoses and Derived Methyl Glycosides . *Canadian Journal of Chemistry* **1975**, *53* (8), 1212–1223. https://doi.org/10.1139/v75-168.

(309) Ritchie, R. G. S.; Cyr, N.; Korsch, B.; Koch, H. J.; Perlin, A. S. Carbon-13 Chemical Shifts of Furanosides and Cyclopentanols. Configurational and Conformational Influences. *Canadian Journal of Chemistry* **1975**, *53* (10), 1424–1433. https://doi.org/10.1139/v75-197.

(310) Jørgensen, T. S.; Gren, T.; Oves-Costales, D.; Ortiz-López, F. J.; Carretero-Molina, D.; González, I.; Blin, K.; Genilloud, O.; Weber, T. Complete Genome Sequence of Amycolatopsis Sp. CA-230715, Encoding a 35-Module Type I Polyketide Synthase. *Microbiology Resource Announcements* **2021**, *10* (38). https://doi.org/10.1128/MRA.00805-21.

(311) Mayer, M.; Thiericke, R. Biosynthetic Relationships in the Desertomycin Family. *Journal of the Chemical Society, Perkin Transactions 1* **1993**, (21), 2525–2531. https://doi.org/10.1039/p19930002525.

(312) Zerlin, M.; Thiericke, R. Common Principles in Macrolactone (Marginolactone) Biosynthesis. Studies on the Desertomycin Family. *Journal of Organic Chemistry* **1994**, *59* (23), 6986–6993. https://doi.org/10.1021/jo00102a023.

(313) Keatinge-Clay, A. Crystal Structure of the Erythromycin Polyketide Synthase Dehydratase. *Journal of Molecular Biology* **2008**, *384* (4), 941–953. https://doi.org/10.1016/j.jmb.2008.09.084.

(314) Keatinge-Clay, A. T. A Tylosin Ketoreductase Reveals How Chirality Is Determined in Polyketides. *Chemistry & Biology* **2007**, *14* (8), 898–908. https://doi.org/10.1016/j.chembiol.2007.07.009.

(315) Kwan, D. H.; Schulz, F. The Stereochemistry of Complex Polyketide Biosynthesis by Modular Polyketide Synthases. *Molecules* **2011**, *16* (7), 6092–6115. https://doi.org/10.3390/molecules16076092.

(316) Musiol-Kroll, E. M.; Wohlleben, W. Acyltransferases as Tools for Polyketide Synthase Engineering. *Antibiotics* **2018**, *7* (3). https://doi.org/10.3390/antibiotics7030062.

(317) Melançon, C. E.; Yu, W.; Liu, H. TDP-Mycaminose Biosynthetic Pathway Revised and Conversion of Desosamine Pathway to Mycaminose Pathway with One Gene. *Journal of the American Chemical Society* **2005**, *127* (35), 12240–12241. https://doi.org/10.1021/ja053835o.

(318) Cundliffe, E.; Bate, N.; Butler, A.; Fish, S.; Gandecha, A.; Merson-Davies, L. The Tylosin-Biosynthetic Genes of Streptomyces Fradiae. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology* **2001**, *79* (3–4), 229–234. https://doi.org/10.1023/A:1012065300116.

(319) Melançon, C. E.; Hong, L.; White, J. A.; Liu, Y. N.; Liu, H. W. Characterization of TDP-4-Keto-6-Deoxy-D-Glucose-3,4-Ketoisomerase from the D-Mycaminose Biosynthetic Pathway of Streptomyces Fradiae: In Vitro Activity and Substrate Specificity Studies. *Biochemistry* **2007**, *46* (2), 577–590. https://doi.org/10.1021/bi061907y.

(320) Tello, M.; Rejzek, M.; Wilkinson, B.; Lawson, D. M.; Field, R. A. Tyl1a, a TDP-6-Deoxy-D-Xylo-4-Hexulose 3, 4-Isomerase from Streptomyces Fradiae: Structure Prediction, Mutagenesis and Solvent Isotope Incorporation Experiments to Investigate Reaction Mechanism. *ChemBioChem* **2008**, *9* (8), 1295–1302. https://doi.org/10.1002/cbic.200800021.

(321) Chen, H.; Yeung, S.-M; Que, N. L. S.; Müller, T.; Schmidt, R. R.; Liu, H. Expression, Purification, and Characterization of TylB, an Aminotransferase Involved in the Biosynthesis of Mycaminose. *Journal of the American Chemical Society* **1999**, *121* (30), 7166–7167. https://doi.org/10.1021/ja991213v.

(322) Liang, D.-M.; Liu, J.-H.; Wu, H.; Wang, B.-B.; Zhu, H.-J.; Qiao, J.-J. Glycosyltransferases: Mechanisms and Applications in Natural Product Development. *Chemical Society Reviews* **2015**, *44* (22), 8350–8374. https://doi.org/10.1039/C5CS00600G.

(323) Melançon, C. E.; Takahashi, H.; Liu, H. W. Characterization of tylM3/tylM2 and mydC/mycB Pairs Required for Efficient Glycosyltransfer in Macrolide Antibiotic Biosynthesis. *Journal of the American Chemical Society* **2004**, *126* (51), 16726–16727. https://doi.org/10.1021/ja043900e.

(324) Moncrieffe, M. C.; Fernandez, M. J.; Spiteller, D.; Matsumura, H.; Gay, N. J.; Luisi, B. F.; Leadlay, P. F. Structure of the Glycosyltransferase EryCIII in Complex with Its Activating P450 Homologue EryCII. *Journal of Molecular Biology* **2012**, *415* (1), 92–101. https://doi.org/10.1016/j.jmb.2011.10.036.

(325) Matsumori, N.; Kaneno, D.; Murata, M.; Nakamura, H.; Tachibana, K. Stereochemical Determination of Acyclic Structures Based on Carbon−Proton Spin-Coupling Constants. A Method of Configuration Analysis for Natural Products. *The Journal of Organic Chemistry* **1999**, *64* (3), 866–876. https://doi.org/10.1021/jo981810k.

(326) Kobayashi, Y.; Tan, C.-H.; Kishi, Y. Stereochemical Assignment of the C21–C38 Portion of the Desertomycin/Oasomycin Class of Natural Products by Using Universal NMR Databases: Prediction. *Angewandte Chemie International Edition* **2000**, *39* (23), 4279–4281. https://doi.org/10.1002/1521-3773(20001201)39:23<4279::AID-ANIE4279>3.0.CO;2-R.

(327) Kobayashi, Y.; Tan, C.-H.; Kishi, Y. Toward Creation of a Universal NMR Database for Stereochemical Assignment: Complete Structure of the Desertomycin/Oasomycin Class of Natural Products. *Journal of the American Chemical Society* **2001**, *123* (9), 2076–2078. https://doi.org/10.1021/ja004154q.

(328) Bifulco, G.; Dambruoso, P.; Gomez-Paloma, L.; Riccio, R. Determination of Relative Configuration in Organic Compounds by NMR Spectroscopy and Computational Methods. *Chemical Reviews* **2007**, *107* (9), 3744–3779. https://doi.org/10.1021/cr030733c.

(329) Menche, D. New Methods for Stereochemical Determination of Complex Polyketides: Configurational Assignment of Novel Metabolites from Myxobacteria. *Natural Product Reports* **2008**, *25* (5), 905–918. https://doi.org/10.1039/b707989n.

(330) Bermel, W.; Wagner, K.; Griesinger, C. Proton-Detected C,H Correlation via Long-Range Couplings with Soft Pulses; Determination of Coupling Constants. *Journal of Magnetic Resonance (1969)* **1989**, *83* (2), 223–232. https://doi.org/10.1016/0022-2364(89)90186-8.

(331) Hansen, P. E. Assignment of the Natural Abundance Carbon-13 Spectrum of Proteins Using Carbon-13-Proton Detected Heteronuclear Multiple-Bond Correlation NMR Spectroscopy: Structural Information and Stereospecific Assignments from Two- and Three-Bond Carbon-Hydrogen Coupling Constants. *Biochemistry* **1991**, *30* (43), 10457–10466. https://doi.org/10.1021/bi00107a014.

(332) Kontou, E. E.; Gren, T.; Ortiz-López, F. J.; Thomsen, E.; Oves-Costales, D.; Díaz, C.; de la Cruz, M.; Jiang, X.; Jørgensen, T. S.; Blin, K.; Charusanti, P.; Reyes, F.; Genilloud, O.; Weber, T. Discovery and Characterization of Epemicins A and B, New 30-Membered Macrolides from Kutzneria Sp. CA-103260. *ACS Chemical Biology* **2021**, *16* (8), 1456–1468. https://doi.org/10.1021/acschembio.1c00318.

(333) Kobayashi, Y.; Tan, C. H.; Kishi, Y. Toward Creation of a Universal NMR Database for Stereochemical Assignment: The Case of 1,3,5-Trisubstituted Acyclic Systems. *Helvetica Chimica Acta* **2000**, *83* (9), 2562–2571. https://doi.org/10.1002/1522-2675(20000906)83:9<2562::AID-HLCA2562>3.0.CO;2-Z.

(334) Hayashi, N.; Kobayashi, Y.; Kishi, Y. Toward the Creation of NMR Databases in Chiral Solvents for Assignments of Relative and Absolute Stereochemistry: Scope and Limitation. *Organic Letters* **2001**, *3* (14), 2249–2252. https://doi.org/10.1021/ol010109r.

(335) Miyata, Y.; Matsunaga, S. Structure Elucidation of 21,22-Dihydroxyonnamides A1-A4 from the Marine Sponge Theonella Swinhoei: An Empirical Rule to Assign the Relative Stereochemistry of Linear 1,5-Diols. *Tetrahedron Letters* **2008**, *49* (44), 6334–6336. https://doi.org/10.1016/j.tetlet.2008.08.057.

(336) Takada, K.; Imae, Y.; Ise, Y.; Ohtsuka, S.; Ito, A.; Okada, S.; Yoshida, M.; Matsunaga, S. Yakushinamides, Polyoxygenated Fatty Acid Amides That Inhibit HDACs and SIRTs, from the Marine Sponge Theonella Swinhoei. *Journal of Natural Products* **2016**, *79* (9), 2384–2390. https://doi.org/10.1021/acs.jnatprod.6b00588.

(337) Xie, X.; Garg, A.; Khosla, C.; Cane, D. E. Mechanism and Stereochemistry of Polyketide Chain Elongation and Methyl Group Epimerization in Polyether Biosynthesis. *Journal of the American Chemical Society* **2017**, *139* (8), 3283–3292. https://doi.org/10.1021/jacs.7b00278.

(338) Xie, X.; Garg, A.; Khosla, C.; Cane, D. E. Elucidation of the Cryptic Methyl Group Epimerase Activity of Dehydratase Domains from Modular Polyketide Synthases Using a Tandem Modules Epimerase Assay. *Journal of the American Chemical Society* **2017**, *139* (28), 9507–9510. https://doi.org/10.1021/jacs.7b05502.

(339) Tanaka, T.; Nakashima, T.; Ueda, T.; Tomii, K.; Kouno, I. Facile Discrimination of Aldose Enantiomers by Reversed-Phase HPLC. *Chemical and Pharmaceutical Bulletin* **2007**, *55* (6), 899–901. https://doi.org/10.1248/cpb.55.899.

(340) Genius 9.1.8. https://www.geneious.com.

(341) Edgar, R. C. MUSCLE: A Multiple Sequence Alignment Method with Reduced Time and Space Complexity. *BMC Bioinformatics* **2004**, *5*. https://doi.org/10.1186/1471-2105-5-113.

(342) Chen, H.; Yamase, H.; Murakami, K.; Chang, C. wei; Zhao, L.; Zhao, Z.; Liu, H. wen. Expression, Purification, and Characterization of Two N,N-Dimethyltransferases, Ty1M1 and DesVI, Involved in the Biosynthesis of Mycaminose and Desosamine. *Biochemistry* **2002**, *41* (29), 9165–9183. https://doi.org/10.1021/bi020245j.

(343) Hong, L.; Zhao, Z.; Melançon, C. E.; Zhang, H.; Liu, H. W. In Vitro Characterization of the Enzymes Involved in TDP-D-Forosamine Biosynthesis in the Spinosyn Pathway of Saccharopolyspora Spinosa. *Journal of the American Chemical Society* **2008**, *130* (14), 4954–4967. https://doi.org/10.1021/ja0771383.

(344) Cheminfo.org. Multiplet Simulator. http://www.cheminfo.org/Spectra/NMR/Tools.