Chemical Investigations of Pacific Marine Organisms

By

Joe Bracegirdle

A thesis submitted to Victoria University of Wellington

For in fulfilment for the requirements of

Doctor of Philosophy in Chemistry

Supervised by Dr. Rob Keyzers

2020

Abstract

This thesis describes a comparative study of three screening methods for the discovery of new marine natural products (MNP) from South Pacific organisms. Over the course of the study, 13 tunicates, one bryozoan and one red alga collected from the waters of the Kingdom of Tonga and New Zealand were investigated. Bioassay, ^{1}H nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) were used to prioritise six organisms for investigation, which resulted in the isolation and characterisation of 16 new and 11 previously reported metabolites.

An NMR–based spectroscopic screening protocol led to the isolation of the new cyclic polyhalogenated monoterpene costatone C (**46**), isolated from a NZ red alga identified as *Plocamium angustum*, which showed moderate antibacterial activity. The absolute configuration of a previously reported and co–isolated *bis*–normonoterpene (**47**) was also deduced using Mosher's method. The NMR screening protocol also prioritised the Tongan bryozoan *Nelliella nelliiformis*, from which the nucleosides nellielloside A and B (**61** and **62**) were isolated and inspired the synthesis of congeners **66–69**. All six compounds showed potent kinase inhibitory activity against a subset of disease–relevant kinases, with IC_{50} values determined to be in the nM range for GSK3A, MAPK14 and RSK2.

An MS screening protocol was also employed utilising molecular networking through the GNPS platform. Of the eight tunicates examined, three were further investigated based on promising constellations in the network, and collectively led to the identification of 13 new compounds. A new aromatic ketone (**28**) was isolated from the NZ tunicate *Distaplia stylifera*, while a comprehensive examination of a NZ *Synoicum kuranui* sample resulted in two new compounds, rubrolides T and U (**110** and **111**). The encrusting tunicate *Didemnum ternerratum* was the most fruitful organism examined in this work, resulting in the isolation of ten new and three previously reported lamellarin sulfates (**147**–**151** and **153**–**156**). The absolute configurations of atropisomers **148**–**151** were determined by comparison of experimental and calculated ECD spectra. Compound **151** showed moderate cytotoxicity against HCT–116 cells.

This work illustrated the power of molecular networking as a screening tool when applied to the MNP field, particularly for the isolation new derivatives within a previously known family of compounds. It also suggested that the more traditional NMR spectroscopy–based screening protocol is still very useful as a standalone method when no comparative standards are

available. These complementary techniques should be used together to maximise new MNP discovery.

Acknowledgements

First, I would like to acknowledge and thank my amazing supervisor Dr. Rob Keyzers for all your help, dedication and commitment. You have inspired me throughout my natural products journey, and I am indebted to you for opening my mind to this field. You provided the perfect amount of direction and support while allowing me to explore and be independent in the lab. You were also the best (and fastest) editor I have ever heard of in science and that is something I appreciated so much. Without you, I wouldn't have been able to achieve anything remotely near what I have, thank you!

Next, I would like to thank everybody past and present in the marine natural products research group, you guys were amazing. In particular, thank you to Sarah Andreassend who I had the pleasure of working beside throughout my whole time at VUW, you were amazing and all the best for your future endeavours. Next, thank you to Helen Woolner, Taitusi Taufa and Jono Singh for being my mentors, you all taught me so much and I appreciate your time and patience in the lab. Thank you very much for proofreading my thesis in such detail Jono, you did such a detailed job! A huge thank you to Jan Vorster who keeps our facilities running smoothly. Your availability and dedication are second to none, without you none of the amazing research at SCPS would be possible, you are a legend.

Thank you to Jeremy Owen for being my secondary supervisor and providing me with exposure to the microbiological world. Even though I didn't get to spend too much time in the lab, I learnt a lot from being around you and your research group and I hope we can continue to work together in the future. I would like to thank Matt Storey and Luke Stevenson for helping me around the lab and carrying out antibacterial assays.

Next, to all of the collaborators that helped out my PhD project in various ways. Thanks for all of the help with chemical synthesis Joanne Harvey, Amira Brackovic and Jordan McCone. To Tony Carroll and Luke Robertson, thank you so much for providing the opportunity to travel to the Gold Coast and running our experiments and computations along with Paul Hume. I would also like to thank Muhammad Ali Hashmi and Zaineb Sohail for your computational work. To the taxonomists and biologists, Joe Zuccarello, Dennis Gordon, Mike Page, Abby Sharrock, Dave Ackerley, Monica Gerth and Mike Fairhurst thank you for your time in contributions to this work.

I am very grateful to the Victoria Doctoral Scholarship and the Curtis–Gordon Research Scholarship in Chemistry for financial support.

Finally, I would like to thank my friends and family for the endless support and helping me get through this journey. Particularly Mum, Dad, Max and Nana for support at home, and Jordan McCone for all of the coffees at lunchtimes. It was full of challenges and hurdles, but our chats and activities helped to keep my mind clear outside of university and allowed me to reach my full potential. Thank you all!

Publications from this work

Chapter 1: Bracegirdle, J.; Keyzers, R. A. Modern Techniques for the Discovery of Marine Natural Products. Chem. N.Z. **2019**, *83* (2), 71–77.

Chapter 3: Bracegirdle, J.; Sohail, Z.; Fairhurst, M. J.; Gerth, M. L.; Zuccarello, G. C.; Hashmi, M. A.; Keyzers, R. A. Costatone C–A New Halogenated Monoterpene from the New Zealand Red Alga *Plocamium angustum*. Mar. Drugs **2019**, *17* (7), 418.

Chapter 4: Bracegirdle, J.; Gordon, D. P.; Harvey, J. E.; Keyzers, R. A. Kinase–Inhibitory Nucleoside Derivatives from the Pacific Bryozoan *Nelliella nelliiformis*. J. Nat. Prod. **2020**, *83* (2), 547–551.

Chapter 5: Bracegirdle, J.; Keyzers, R. A. Marine–derived Polyaromatic Butenolides – Isolation, Synthesis and Biological Evaluations. Curr. Pharm. Des. **2020**, DOI: 10.2174/1381612826666200518110617.

Bracegirdle, J.; Stevenson, L. J.; Page, M. J.; Owen, J. G.; Keyzers, R. A. Targeted Isolation of Rubrolides from the New Zealand Marine Tunicate *Synoicum kuranui*. Mar. Drugs **2020**, *18* (7), 337.

Chapter 6: Bracegirdle, J.; Robertson, L. P.; Hume, P. A.; Page, M. J.; Sharrock, A. V.; Ackerley, D. F.; Carroll, A. R.; Keyzers, R. A. Lamellarin Sulfates from the Pacific Tunicate *Didemnum ternerratum*. J. Nat. Prod. **2019**, *82* (7), 2000–2008.

Other Authors Contributions

Chapter 1: J. Bracegirdle completed this work.

Chapter 2: J. Bracegirdle completed this work. Dr. L. J. Stevenson and Dr. J. G. Owen (VUW) completed the antibacterial bioassay.

Chapter 3: J. Bracegirdle completed all chemical isolation and structure elucidation. Dr. M. A. Hashmi (University of Education, Attock, Pakistan) and Z. Sohail (University of Management and Technology, Lahore, Pakistan) completed the computational structure optimisation and ECD calculations. M. Fairhurst and Dr. M. L. Gerth (VUW) completed the antibacterial bioassays. Prof. G. C. Zuccarello identified the alga and undertook taxonomic analysis.

Chapter 4: J. Bracegirdle completed this work. Assoc. Prof. J. Harvey (VUW) supervised chemical synthesis. Dr. D. P. Gordon (NIWA) identified the bryozoan.

Chapter 5: J. Bracegirdle completed this work. Dr. L. J. Stevenson and Dr. J. G. Owen (VUW) completed the antibacterial bioassays. M. J. Page (NIWA) identified the tunicate.

Chapter 6: J. Bracegirdle completed all chemical isolation and structure elucidation. Dr. L. Robertson and Prof. A. C. Carroll (Griffith University, Australia) completed the computational structure optimisation and ECD calculations. Dr. P. Hume (VUW) completed the free energy of rotation calculations. Dr. A. V. Sharrock and Prof. D. F. Ackerley (VUW) completed the cytotoxicity bioassays. M. J. Page (NIWA) identified the tunicate.

Chapter 7: J. Bracegirdle completed this work.

Table of Contents

List of Figures

List of Schemes

List of Tables

Glossary

δ	Chemical shift (ppm)
^{13}C NMR	Carbon-13 nuclear magnetic resonance
^{1}H NMR	Proton nuclear magnetic resonance
ACN	Acetonitrile
ADP	Adenosine diphosphate
AIDS	Acquired immunodeficiency syndrome
APCI	Atmospheric pressure chemical ionisation
ATP	Adenosine triphosphate
Auto-MS/MS	Untargeted ion fragmentation in MS/MS
br	Broad
C18	Octadecyl-derivatised silica gel
C8	Octyl-derivatised silica
calcd	Calculated *m/z* value
CD_3OD	Deuterated methanol
$CDCl_3$	Deuterated chloroform
CHIKV	Chikungunya virus
CID	Collision induced dissociation
CML	Chronic myelogenous leukaemia
COSY	Homonuclear correlation spectroscopy
d	Doublet
DCC	*N,N'*-Dicyclohexylcarbodiimide
DCM	Dichloromethane
DFT	Density functional theory
DIOL	2,3-Dihydroxypropoxypropyl-derivatised silica gel
DMAP	4-Dimethylaminopyridine
DMSO-d_6	Deuterated dimethyl sulfoxide
ECD	Electronic circular dichroism spectroscopy
EDC.HCl	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride salt
eEF1A2	Eukaryotic elongation factor 1A2
EI	Electron impact ionisation
EtOAc	Ethyl acetate
FA	Formic acid
FDA	United States of America Food and Drug Administration
FRET	Förster resonance energy transfer (aka fluorescence energy transfer)
GC	Gas chromatography
GI_{50}	Growth inhibitory concentration to 50% of test subjects
GLK	Germinal centre kinase-like kinase
GNPS	Global natural products social molecular networking
GSK	Glycogen synthase kinase
HCT-116	Human colorectal carcinoma cell line 116
HIV	Human immunodeficiency virus

HMBC	Heteronuclear multiple bond correlation
HP20/HP20ss	PSDVB stationary support
HPLC	High performance liquid chromatography
HRESIMS	High resolution electrospray ionisation mass spectrometry
HSQC	Heteronuclear single quantum coherence
IC_{50}	Inhibitory concentration to 50% of test subjects
IHD	Index of hydrogen deficiency
IR	Infrared
J	Scalar coupling constant (Hz)
JEV	Japanese encephalitis virus
LC	Liquid chromatography
LC_{50}	Lethal concentration to 50% of test subjects
LH20	Crosslinked dextran-based size exclusion resin
m	Multiplet
m/z	Mass-to-charge ratio
MAPK	Mitogen-activated protein kinase
MCV	Molluscum contagiosum virus
Me_2CO	Acetone
MeMSChem	Meta-mass shift chemical profiling
MeOH	Methanol
MIC	Minimum inhibitory concentration
MNP	Marine natural product(s)
MRSA	Methicillin-resistant *Staphylococcus aureus*
MS	Mass spectrometry; mass spectrometric
MS/MS	Tandem mass spectrometry
mTOR	Mammalian target of rapamycin
MTPA	A-Methoxy-α-trifluoromethylphenylacetic acid
mult	Multiplicity
MW	Molecular weight (Da)
n.d.	Not detected (in NMR spectrum)
NIWA	New Zealand National Institute of Water and Atmospheric Research
NOE	Nuclear Overhauser effect
NZ	New Zealand
PLK	Polo-like kinase
ppm	Parts per million
PSDVB	Poly(styrene-divinylbenzene) copolymer
pTsOH	*para*-Toluenesulfonic acid
quin	Quintet
RA	Rheumatoid arthritis
ROESY	Rotating-frame nuclear Overhauser effect correlation spectroscopy
RSK	Ribosomal S6 kinase
rt	Room temperature
s	Singlet
SAR	Structure-activity relationship

SARS	Severe acute respiratory syndrome virus
SBS	School of Biological Sciences, VUW
SCUBA	Self-contained underwater breathing apparatus
SEM	Scanning electron microscope
t	Triplet
TB	Tuberculosis
TDDFT	Time-dependent DFT
TLC	Thin layer chromatography
UV/vis	Ultraviolet/visible spectroscopy
VUW	Victoria University of Wellington
XRD	X-ray diffraction crystallography

Chapter 1 – Introduction

The medicinal properties of naturally occurring compounds have been exploited by humans for thousands of years. Owing to their ease of access and taxonomical identification, traditional measures mainly focussed on terrestrial plants, which consequently are the topic of the majority of natural products research in the modern era. The terrestrial environment today continues to be a valuable source of new bioactive chemical entities. However, in order to continue to pursue chemical diversity of natural products with potentially novel molecular skeletons, a diverse biota must be interrogated. Specific selection pressures of different habitats lead to organism diversification, which also includes the chemical diversity contained within. Therefore, exploitation of different ecological niches should result in the isolation of new and varied molecular classes essential for drug discovery.

1.1 – Marine Natural Products

The marine environment is one the richest sources of biodiversity on our planet, relatively untouched and unexplored, and therefore represents a unique wealth of unknown chemistry. As 70% of the Earth's surface is oceanic, the volume of habitats is much greater than on terrestrial surfaces, providing much more space for the generation of biodiversity.[1] Before the advent of SCUBA some 70 years ago, it was difficult to collect organisms from subtidal depths, however, today this is common practice. It was fast realised that marine organisms produced a variety of interesting metabolites, with many compound classes not found from terrestrial sources.[2] Such success resulted in the establishment of the marine natural products (MNP) field, which explores the chemical structures and bioactivities of metabolites derived from organisms of the sea. The output began slowly with ~3000 new metabolites reported over the course of a decade (1977–1987).[3-6] However, due to increased worldwide interest and the development of more sensitive techniques/experiments, novel structures became published much more frequently, such as 809 in 1996,[6], 1076 in 2006,[7] and 1277 in 2016,[8] to a total of ~30,000 today.[9] Therefore, marine organisms represent a valuable source for the discovery of new natural products.

The sea is home to a huge range of organisms, all frequently used for the isolation of chemically interesting MNPs. Taken from the numerous annual MNP reviews, the most successful sources are marine algae, microorganisms and invertebrates such as sponges, tunicates, echinoderms and molluscs, each of which produce a variety of different classes of compound.[8,10-18] These

desired bioactive molecules are colloquially known as secondary metabolites. Primary metabolites, such as amino acids, carbohydrates and lipids, are essential to an organism's survival and reproduction, and are found throughout all forms of life, independent of phylogeny. Secondary metabolites are restricted to individual taxonomic groups, however, and provide a selective advantage to the organism that is not strictly essential but increases survival fitness relative to competing organisms. Marine organisms encounter substantially different ecological and physiological challenges compared to their land-dwelling counterparts, which results in the acquisition of bioactive molecules produced by unique pathways with intriguing structures and variations. The loss of biodiversity due to anthropogenic habitat destruction and climate change (acidification and increasing sea temperatures) is cause for concern over the future of novel drug discovery from the sea.[19] Today, marine bioprospection is more important than ever to ensure potentially therapeutic compounds are not lost forever.

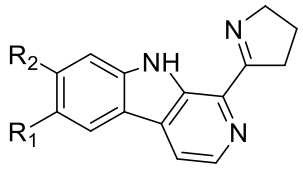
1.2 – Invertebrate Natural Products

Invertebrate animals are found throughout the marine environment and have been a fruitful source of MNPs since the beginning of the field, particularly sponges, cnidarians and echinoderms.[8] In line with being the most successful source of new MNPs globally, the major focus of the Victoria University of Wellington (VUW) MNP research group has been on sponge metabolites sourced from New Zealand (NZ) and Tongan waters. This resulted in the isolation of a plethora of new compounds with varying structures and bioactivities.[20-23] For the work described in this thesis, the marine invertebrate focus was shifted to other phyla in our collections, namely the Chordata and Bryozoa. Chordata (predominantly tunicates) and bryozoans have been the source of many novel molecule classes with ranging bioactivities. A recent statistical study compared various physico-chemical properties of MNPs isolated from different phyla, such as clog*P* (calculated partition coefficient), molecular weight (MW) and number of stereogenic centres, to those of approved drugs. It was deduced that alongside the Actinobacteria and Ascomycota, the Chordata and Bryozoa were the source of MNPs with the most similar structures to approved drugs.[8] However, in recent years the number of new chemical entities published from both phyla is decreasing, reflecting the general shift towards microbial sources. In the most recent review describing MNPs published in 2018, the number of new compounds from tunicates decreased by 45% compared to 2017, as did those from bryozoans (in 2017, 16 metabolites from bryozoans were reported which is unusually high

compared to three in 2018).[10] It is imperative these organisms are not overlooked by the field, and attention should be drawn towards chemical investigations of these organisms.

Tunicates (urochordates) are soft bodied, filter-feeding animals, colloquially referred to as sea squirts. They belong to the phylum Chordata, and phylogenetically represent the closest living relatives to vertebrates.[24] Tunicates are divided into three classes; Thaliacea (72 species), Appendicularia (20 species) and Ascidiacea (>2500 species)[25,26] therefore as the vast majority studied are ascidians; the term is often used interchangeably with tunicate throughout the natural products community. Conversely, bryozoans or "moss animals" are a diverse phylum of over 6000 living species found in marine habitats around the world and are also filter feeding invertebrates.[27,28]

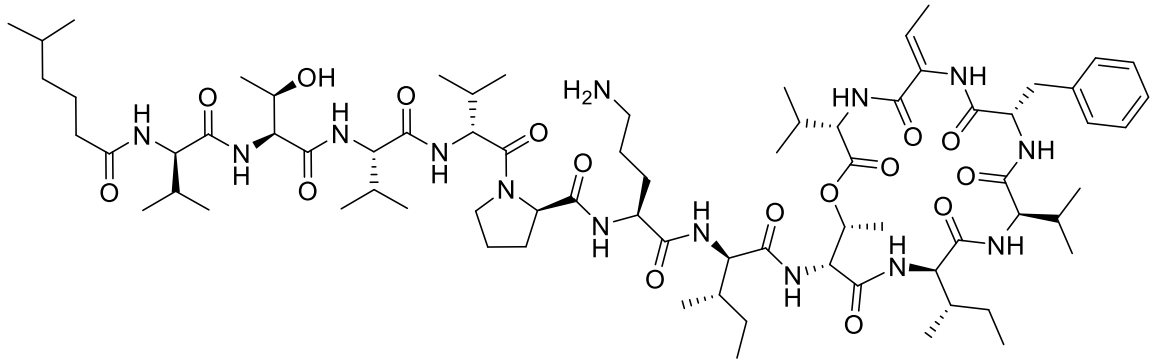
Generally marine invertebrates are sessile organisms, and lack physical protection mechanisms to evade harm. Many tunicates and bryozoans are found in tropical waters where they have a range of predators, particularly grazing fish. It has been demonstrated within our research group[29] and the literature[30,31] that tropical organisms generally have more potent chemical defence mechanisms than those in temperate climates, correlating with increases in both quantity and variation of secondary metabolites. These molecules can act as anti-feedants, produced in larger quantity with the increased predation pressures at lower latitudes.[32] Secondary metabolites also have a host of other functions, such as reducing neighbour growth (anti-fouling) in the competition for resources and even can even provide protection from UV radiation in tropical environments, allowing shallow water habitation that may otherwise be too hazardous.[33] Excess fouling and growth of organisms on the body can be very detrimental to sessile marine organisms,[34] particularly if blocking the siphon for filter-feeding; therefore, metabolites are often produced to control this micro-environment. For example, the alkaloids eudistomin G (**1**) and H (**2**) were isolated from the colonial ascidian *Eudistoma olivaceum*, but were not active as an antifeedant in a fish predation assay.[35] However, they both proved effective in deterring the settlement of larvae, thus providing an effective antifouling chemical defence. This observation also highlights the importance and consequences of selecting the correct bioactivity assay in natural products discovery.



1 R_1 = H R_2 = Br
2 R_1 = Br R_2 = H

1.3 – Algal Natural Products

Similar to their terrestrial counterparts that often contain insecticidal metabolites, seaweeds (macroalgae) also provide bioactive anti-feedant secondary metabolites. Photosynthetic pigments cause the alga to have characteristic red, green, yellow or brown colouring, with the red alga (phylum Rhodophyta) being the largest group.[36] Among these, red algae have been the biggest source of new bioactive chemodiversity,[37] accounting for ~2400 reported structures and have been extensively studied around the world.[9] Metabolites have shown a range of potential bio-applications, particularly as anti-oxidative, anti-inflammatory and anti-cancer agents.[38] Sulfated polysaccharides, such as fucoidan, are isolated from brown algae and have shown promising anti-tumour activity in numerous preclinical *in vivo* studies.[39] The green alga *Bryopsis pennata* is the biogenic source of the only algal natural product to enter phase II clinical trials to date, kahalalide F (**3**), a depsipeptide with potent anti-tumour activity originally isolated from the mollusc *Elysia rufescens*,[40] but produced by an intracellular bacterial symbiont.[41] Kahalalide F passed five phase I clinical trials for advanced solid tumours and psoriasis, however failed phase II for numerous reasons including short half-life and lack of efficacy.[42] Attempts have been made to improve its pharmacokinetics,[43] and derivatives showed potential anti-protozoan activity against leishmaniasis.[44]



3

Griffithsin is the only algal-derived drug currently in the clinical pipeline, first isolated from the red alga *Griffithsia* sp. collected in the waters off of Chatham Island, NZ.[45] Following detection of anti-viral activity of the crude aqueous extract, the 121-amino acid lectin was purified and demonstrated potent (picomolar to nanomolar) activity against a range of enveloped viruses such as human immunodeficiency virus (HIV), hepatitis C and severe acute respiratory syndrome (SARS) coronavirus, with activity attributed to binding the terminal mannose residue of envelope glycoproteins, preventing fusion with cell membranes.[46]

Griffithsin is currently in phase I clinical trials where its ability to act as a antimicrobial and prevent HIV-1 transmission is being accessed.[47] Today, alga of all taxa are a continuing source of novel chemistry.

1.4 – Clinically-Approved Marine Drugs

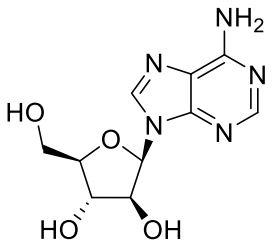
The chemical space occupied by approved drugs mirrors that of natural products,[48,49] with over half of approved drugs being natural products or derived/influenced thereof. Within the field of drug discovery, the most successful result for a candidate compound is the approval for therapeutic use. Drugs are required to proceed through rigorous testing including three phases of clinical trials prior to approval. It is important to recognise the field of MNP discovery is the first step in drug development, as drugs typically require many steps of optimisation before they are ready for clinical use. Although the number of MNP-derived drugs represents a small portion of those used in the clinic, there are many promising leads in clinical/preclinical trials with more being tested every year. As of February 2020, there are 27 marine-derived compounds in Phase I–III trials, and ten approved by the United States Food and Drug Administration (FDA), all of which focus on a range of different molecular targets.[50]

Invertebrates have contributed substantially to the collection, with eight of the ten approved marine drugs derived from molluscs, sponges or tunicates. The foundation was laid by vidarabine (Ara–A (**4**) approved 1976 and since discontinued) and cytarabine (Ara–C (**5**) approved 1969), approved by the FDA for the treatment of *Herpes simplex* virus infections and acute non-lymphocytic leukaemia, respectively.[1] These drugs were synthesised from lead molecules isolated from the sponge *Tectitethya crypta* (originally described as *Cryptotethya crypta*), collected from the Bimini Islands, Bahamas.[51] The other approved sponge-derived drug, eribulin mesylate (**6**, Halaven®, approved 2010) targets microtubules for the treatment of metastatic breast cancer. Although enough halichondrin B (**7**) could be isolated from natural sources for preclinical studies, this was not viable for long term use and synthetic efforts provide costly and complex. However, a synthetic intermediate to the macrolide core was equipotent and far more simple to produce in quantity (although still hugely complicated with 19 chiral centres and MW 731 g mol^{-1}), giving rise to the drug **6**, ultimately the pharmacophore of **7**.[52]

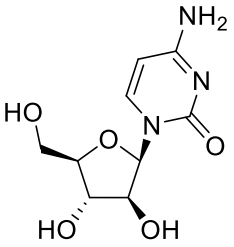
According to the MarinLit database,[9] sponges (phylum Porifera) have been the source of ~11,000 compounds whereas only ~1500 compounds have been described from tunicates

(phylum Chordata). Relatively, the compounds isolated from tunicates are overly successful as drug leads, with two approved for use and one in phase III clinical trials. Trabectedin (**8**) or ET-743 (Yondelis® approved 2015), represents a rare example of a natural product moving to the market unaltered. Trabectedin is the most potent and abundant of the six ecteinascidans isolated from the colonial Caribbean ascidian *Ecteinascida turbinata*.[53] The ecteinascidans are tetrahydroisoquinoline alkaloids produced in very small quantities, with ET-743 only present at 0.0001% yield of dry animal mass. In order to support clinical development, PharmaMar led extremely extensive (and successful) aquaculture farms throughout the Mediterranean, leading to the harvest of 250 tonnes of biomass with a yield of ~1 μg g^{-1}.[54] However, commercialisation of ET-743 required a much more consistent and economical approach. Total synthesis was achieved but was not applicable to industrial scale,[55] with semi-synthesis from cyanosafracin B (**9**) (obtained from the cultured bacterium *Pseudomonas fluorescens* on a kg scale) finally resolving the supply issue.[54] Trabectedin alkylates guanine residues on DNA, bending the helix towards the major groove and causing cell death through disruption of replication events, and was therefore approved as a treatment for soft tissue sarcoma and ovarian cancer.[56] The structurally related lurbinectedin (**10**, Zepsyre®) is currently in phase III clinical trials for small-cell lung cancer and hereditary breast cancers.[57,58] Lurbinectedin is a late-stage synthetic derivative of trabectedin which results in a much higher tolerated dose and efficacy against advanced solid tumours.[59]

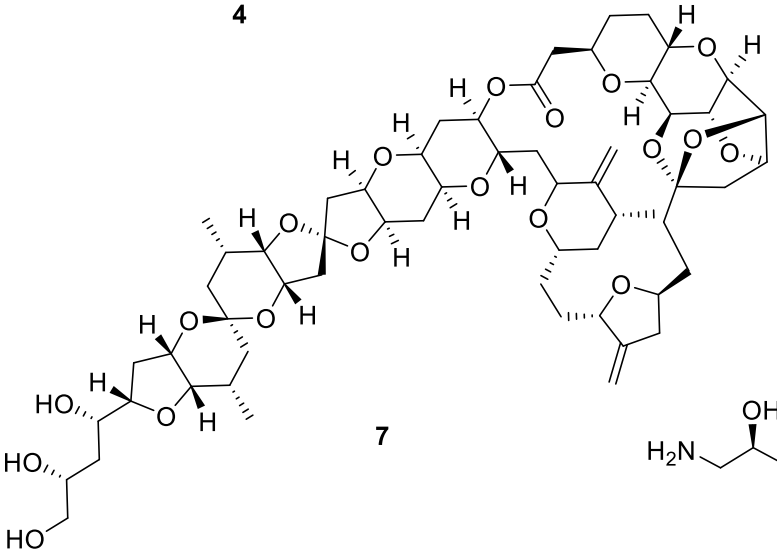
The other approved tunicate-derived drug, the depsipeptide plitidepsin (**11**, Aplidin®),[60] is used to treat resistant multiple myeloma in Australia, and was originally isolated from a Mediterranean *Aplidium albicans* tunicate.[61] Formally known as dehydrodidemnin B, plitidepsin is the most successful congener of the didemnin family of compounds which have shown potent cytotoxic, antiviral and immunosuppressive activities, leading to many clinical trials.[62] Plitidepsin exerts its cytotoxic effects on cancer cells through a number of pathways; however, its primary intracellular target appears to be eukaryotic elongation factor 1A2 (eEF1A2), which is upregulated in multiple myeloma.[63]



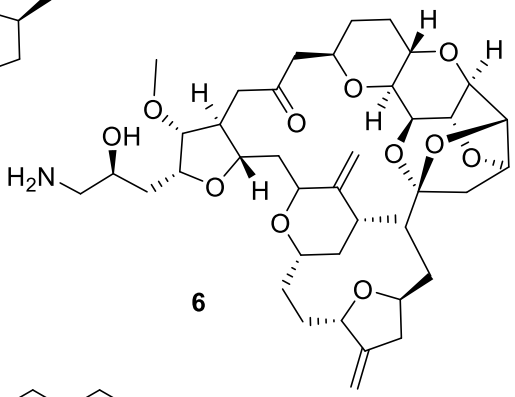
4



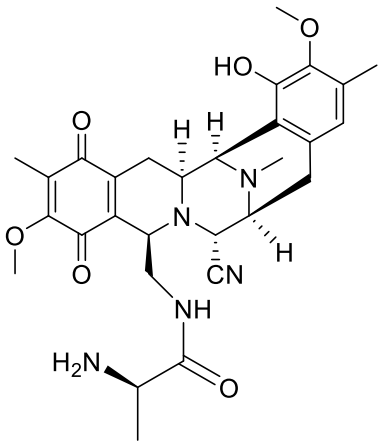
5



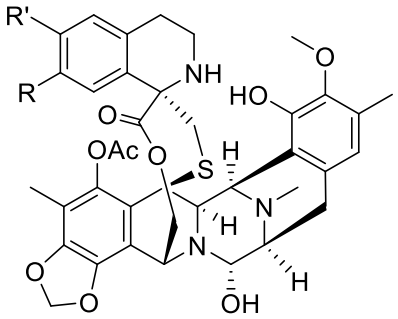
7



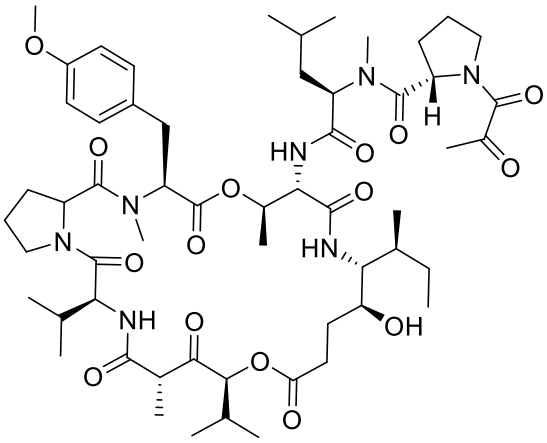
6



9

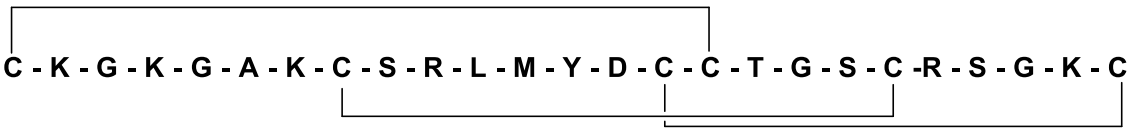


8 R=OCH_3 R'=OH
10 R=H R'=OCH_3

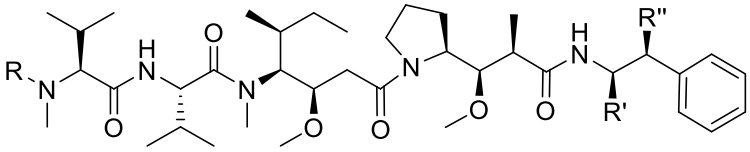


11

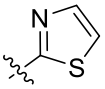
Ziconotide (**12**, Prialt®, approved 2004) is a ribosomally-synthesised 25-residue peptide isolated from the cone snail *Conus magus*, and represents another rare example of a drug approved by the FDA as it is found in nature.[64] The majority of approved drugs derived from natural products are enhanced analogues of the originally isolated compound, many modern examples are antibody-drug conjugates. These utilise the specific targeting ability of antibodies to deliver extremely potent cytotoxic metabolites to their site of action. Monomethyl auristatin E (**13**) is the warhead in the very recently approved enfortumab vedotin (Padcev ™, approved 2019), which is derived from dolastatin 10 (**14**), initially isolated from the mollusc *Dolabella auricularia*.[65,66] The remaining approved MNP drug is a mixture of fish derived omega-3 fatty acids used for the treatment of hypertriglyceridemia.[67]



12



13 R = H R' = CH_3 R'' = OH

14 R = CH_3 R' =  R'' = H

1.5 – Isolation of MNPs at VUW

There is a huge range of specialised metabolites from living organisms produced as a result of various ecological pressures, not all of which are relevant to drug discovery. Common and well-known in the field, Lipinski's rule of five describes the physio-chemical properties of a compound associated with solubility and permeability – an evaluation of druglikeness.[68,69] Drug molecules are typically required to be amphiphilic in nature, to ensure solubility in the relevant setting, the ability to cross membranes but also interact with their target. Although it is very difficult to exploit the number of hydrogen bond donors/acceptors and MW of compounds prior to isolation, the clog*P* can be used as a proxy for polarity. Chromatographic purification tools typically use relevant affinities of polar/non-polar stationary phases to the mobile phase to separate molecules based on their polarity, therefore molecules with both hydrophobic and hydrophilic nature can be prioritised.

In past experience of VUW MNP group members, the crude extracts of marine organisms are typically full of a range of polar (sugars, salts, amino acids) and non-polar compounds (fats and lipids), with secondary metabolites constituting a much smaller amount.[70] This is represented by the mass window (Figure 1.1), which demonstrates the relative distribution and polarities of the constituents in a typical extract. In order to enrich for drug-like molecules, a pre-fractionation using polymeric reversed-phase media (HP20) has been employed over the past 20 years. HP20 beads are a robust stationary phase made from a poly(styrene-divinylbenzene) (PSDVB) copolymer that acts as a reversed-phase resin. The typical protocol involves loading a methanolic extract by sequential dilution with H_2O, washing with neat H_2O and then elution with increasing percentages of an organic modifier in H_2O (e.g. 30%, 75% and 100% Me_2CO in H_2O). By this method, the intermediate polarity fraction is enriched in amphiphilic drug-like molecules, and the extremes are polar and non-polar molecules that either do not adhere or significantly adhere to the resin. These fractions are much easier to handle than a crude extract in terms of solubility and raw mass, and all three fractions are always analysed for novel chemistry in order to prioritise further study.

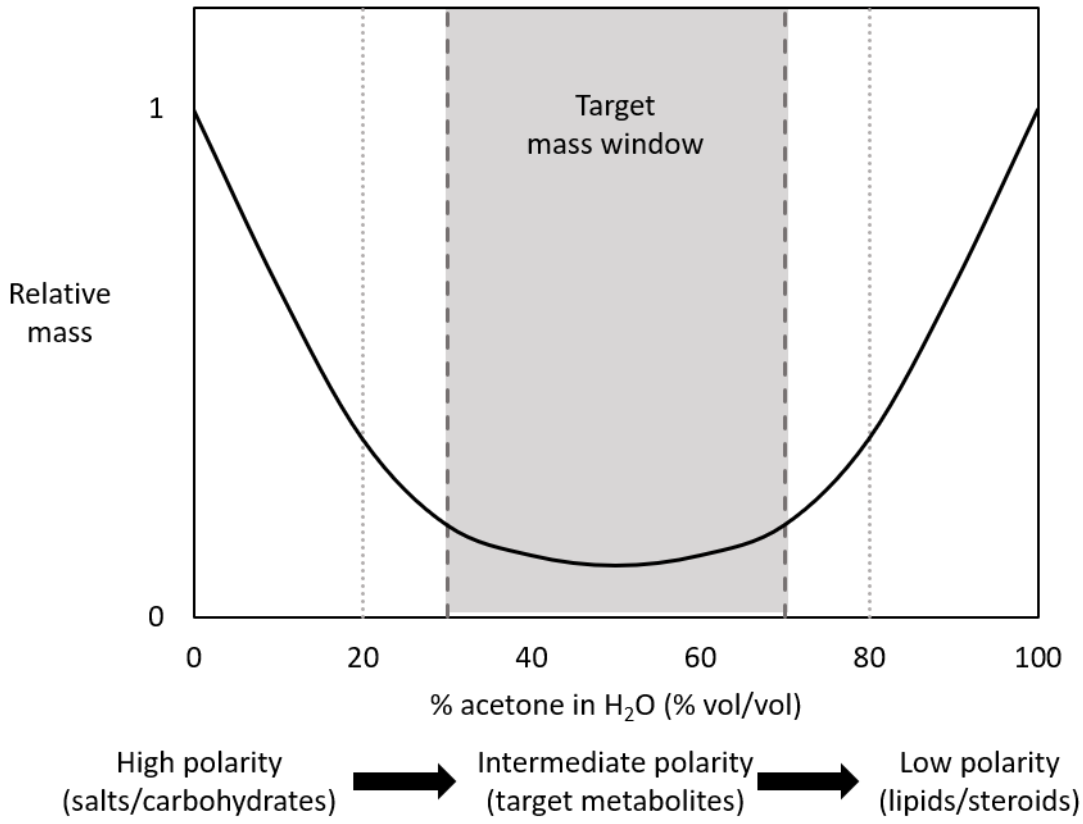


Figure 1.1 – Mass window of a typical marine invertebrate extract and elution over HP20 resin.

1.6 – Dereplication

By far one of the largest challenges in natural products drug discovery is dereplication, the early identification of previously reported metabolites. The probability of re-isolating known compounds is both very high and increasing,[71] which can become extremely costly in terms of both time and money. The faster chemical novelty within an extract can be established, the more optimal the procedure becomes. With this comes the concept of screening (Chapter 2), an assessment of an organism (or extract) for target molecules of interest, which is completed at the beginning of any NP program. Bioassays can be used to probe a desired biological activity within an extract, while spectroscopic techniques profile the types of molecules within. These include 1D ^{1}H nuclear magnetic resonance (NMR) spectroscopy (used in this study) or more advanced 2D NMR (COSY, TOCSY, HSQC, HMBC) which help to increase resolution of chemical shifts by spreading the data over two dimensions, but take longer to acquire,[72] or mass spectrometry (MS; also utilised here). Many spectroscopic and spectrometric tools are utilised for dereplication, and in combination with biological data of the organism (taxonomy, bioactivity etc.), databases such as the Dictionary of Natural Products,[73] DEREP-NP[74] or

MarinLit[9] are searched. These databases are not extensive, however, and do not contain every molecule described in the literature but can give clues as to the structural classes of some 'hits'. They also often rely on precise spectroscopic/metric inputs, which may vary depending on the solvent or concentration used (e.g. ^{1}H NMR chemical shift) and are very difficult to use when mixtures of compounds are present, such as early in the purification procedure, although more robust tools are beginning to be developed capable of analysing these types of data (e.g. SMART NMR).[75] The sharing of raw data between research groups has the power to rapidly speed up the process of dereplication, including studies of organisms that do not result in the isolation of novel chemistry and are therefore not published. A recent communication expressed the need for more data sharing in metabolomics, which faces a similar problem analogous to that of natural products.[76] Therefore although tools exist, dereplication is still a substantial problem today.

Clearly, the earlier dereplication is achieved the better. The main reason why our research group opts to use a spectroscopic-guided isolation procedure as opposed to the more common bioassay-guided is to achieve this. Although the molecules isolated through such guidance do not have guaranteed bioactivity, producing secondary metabolites is very energy demanding and inactive/pointless molecules are removed by evolution, therefore the vast majority are active. Experiments such as 1D (^{1}H and ^{13}C) and 2D (COSY, HSQC and HMBC) NMR spectroscopy and liquid-chromatography mass-spectrometry (LCMS) use very little/none of the material to provide a wealth of data. In contrast, bioassays not only use the sample and can take a long time, they may also give a negative result on a bioactive sample if the wrong assay is used, as was the case with the anti-fouling eudistomins G (**1**) and H (**2**) previously mentioned.[35]

1.7 – Molecular Networking and GNPS

In this work, a recently developed molecular networking tool was also used to screen initial crude fractions by LCMS/MS, alongside the more traditional ^{1}H NMR spectroscopy. MS is well suited as a screening tool; it is easy to run, very sensitive and can be used in a high-throughput manner for rapid generation of large datasets. Through the analytical separation on the LC column, a degree of separation and differentiation is achieved that is not available with NMR-based methods. MS/MS introduces another dimension to data acquisition, providing information on how the molecule dissociates or fragments upon collision with an inert gas.

From an injection of a mixture of compounds, untargeted LCMS/MS selects and fragments the most intense ions above a defined ion intensity threshold, these ions of which are ignored in subsequent scans in preference for less intense ions. This generates a large dataset for each sample with information on many molecules within, including those of lower abundance. This contrasts with classic ^{1}H NMR spectroscopic screening where peaks of minor metabolites may be masked by more concentrated compounds, particularly minor analogues present that can easily be missed.

Global Natural Products Social Molecular Networking (GNPS) is an online tool used for the sharing and processing of MS/MS data.[77] Users upload both annotated and unknown raw data to contribute to the crowd-sourced library. The annotated spectra can then be used to label previous unknowns, and as 'deposited' data are continuously reanalysed, connections can be made in the future if an annotated hit is later uploaded. This tool is extremely powerful for dereplication in the screening process. Molecular networks are generated on the basis that similarly structured molecules will fragment with similar patterns. Therefore, the MS/MS spectra of each precursor ion (node) represents a 'fingerprint' for a molecule, and through spectral alignment these nodes are connected to each other. A proxy for the degree of similarity of the molecules is the cosine score, which is often labelled on the edge (connecting line) or represented for by the edge line thickness. This generates a visual network representing the structural chemical space occupied by the group of metabolites in the sample (Figure 1.2). Within networks, clusters of nodes (referred to as 'constellations') form when analogues with slightly different structures are detected. This is particularly useful when looking at minor derivatives of major metabolites, as structural information can also be ascertained from the difference in precursor ion *m/z*, which is unlikely to be determined from a ^{1}H NMR spectroscopy screen. For example, two nodes with a high cosine score and a difference of $m/z = 2$ is indicative of reduction of a double bond, whereas differences of 14, 28 or 42 suggest differences of methylene groups, possibly part of an alkyl chain or methylation of an alcohol. There are various common functional group changes that can occur between related molecules, all resulting in diagnostic mass shifts and are detailed by meta-mass shift chemical (MeMSChem) profiling.[78] This is a tool initially developed to compare the metabolomes of coral reefs, and showed that although genetically similar, the MeMSChem profiles were unique and thus the metabolites are modified differently depending on the environment the reef inhabits. Therefore, GNPS has two key applications in screening: dereplication of known

metabolites present in the database, and molecular networking of similarly structured molecules – both of which are very valuable for drug discovery.[79]

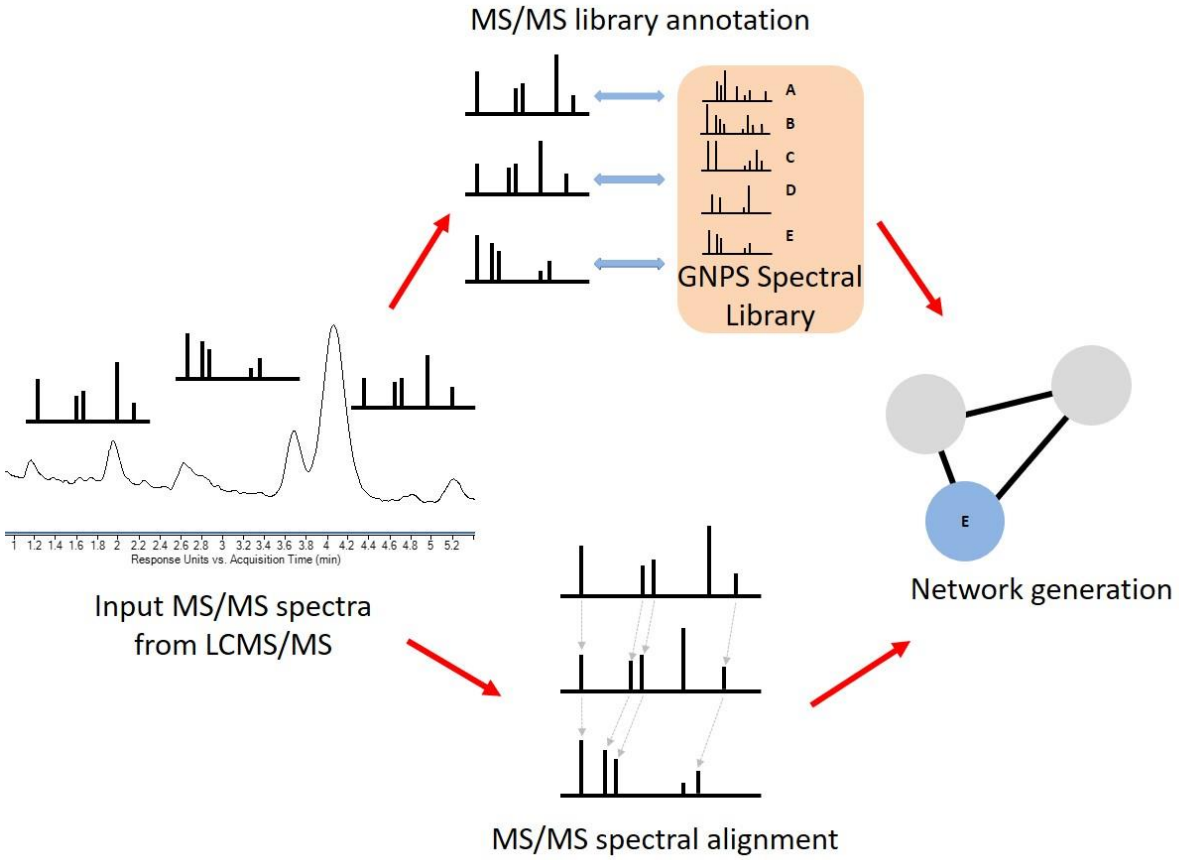
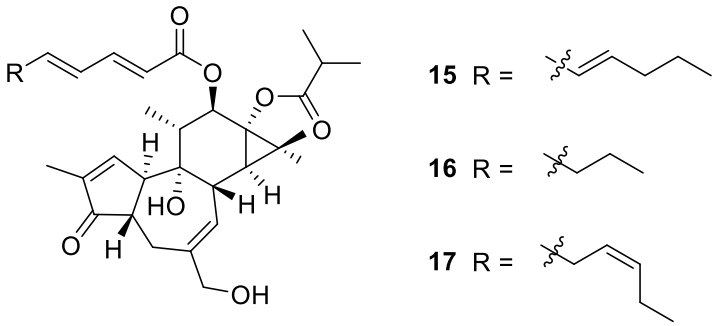


Figure 1.2 – GNPS molecular networking workflow as adapted from Wang et al.[77] In the network; grey nodes are unassigned, the blue node is annotated by database matching as compound E.

In natural products research, a large number of samples are often dealt with, whether it is a variety of extracts or fractionated samples as a result of compound purification. Molecular networking through the GNPS platform allows sample specific data (metadata) to be visualised in the network, represented by features such as the size or colour(s) of the nodes and gives rapid comparisons of features.[80] Aspects such as the samples collection location,[81,82] or culturing conditions for growth,[83] have been used however the most common and relevant to drug discovery is the sample bioactivity. A bioinformatic workflow termed '*bioactive molecular networking*' has been developed to integrate this information from fractions, and predict which nodes (precursor ions) correlate most with the activity of a quantified sample.[84] This can then be used to visualise which nodes should be targeted in a bioactivity guided

isolation procedure. The method was applied to a *Euphorbia densroides* plant extract with anti-chikungunya virus (CHIKV) activity to probe for its bioactive constituents.[84] Previous studies on the extract resulted in the isolation of many new compounds; however, none had antiviral activity representative of that observed for the extract,[85] an event that may occur if the active components are present in small quantities or if the activity is the result of compound synergism. By following the nodes predicted to correlate with bioactivity, three new deoxyphorbol ester derivatives (**15**–**17**) were isolated as minor components of the extract, where **15** and **16** were potent and selective inhibitors of CHIKV.[84]



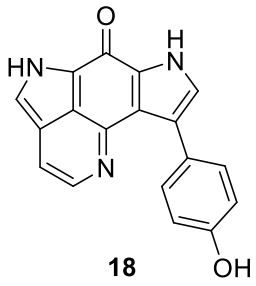
Although it was only recently developed in 2012,[86] the GNPS molecular networking protocol has been readily adopted by the natural products community, and continues to grow in user numbers and database size. Over the past five years, there have been a substantial number of reports utilising the tool in various ways to aid the discovery of new chemical entities, with the majority of publications focusing on microorganisms and terrestrial plants.[85,87-96] In relation to MNPs, marine derived microorganisms, particularly cyanobacteria, have been the focus of the majority of studies,[97-100] which reflects the direction of the current MNP field in general.[10] Currently, all studies on invertebrates have focused on sponges,[82,101-105] with no previous reports on the metabolites of tunicates or bryozoans, aside from a study on *Streptomyces* sp. strain PTY087I2 cultured from a Panamanian *Styela canopus* tunicate.[106]

Thorough reviews of the use of GNPS to discover new natural products have been reported elsewhere.[80,107] Pyrroloiminoquinone alkaloids are produced by a variety of sponge species and are widely studied due to their potential anti-cancer applications.[108] There are many families within the class, such as the makaluvamines, discorhabdins and tsitsikammamines, all with subtle structural variations that makes them perfect candidates for predictable fragmentation and molecular networking studies. The molecular network of an Antarctic *Latrunculia biformis* sponge extract showed clusters dereplicated as discorhabdins and epinardins, with a separate cluster containing the structurally related tsitsikammamine A (*m/z* 304.125) manually

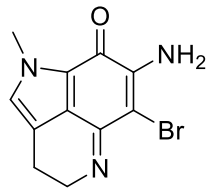
annotated based on its fragmentation pattern. In this cluster, a node at *m/z* 302.111 was linked with a high cosine score, and therefore has a similar chemical structure. Following targeted purification and spectroscopic structure elucidation, the compound was determined to be the previously unknown 16,17-dehydrotsitsikammamine A (**18**).[101] In a separate study, molecular networking of six South African *Tsitsikamma favus* samples showed two distinct chemotypes, revealed by sample specific clusters within a pyrroloiminoquinone constellation.[102] One chemotype of four samples was found to produce mainly discorhabdins and tsitsikammamines, whereas the other had predominantly unbranched and halogenated makaluvamines. Through the network, the previously unreported makaluvamine Q (**19**) was identified and targeted for isolation, the first example of a brominated and *N*–methylated derivative.[102]

Other sponge studies using molecular networking led to the isolation of leucettazine A (**20**) from an Australian *Leucetta* sp.,[103] furofficin (**21**) from *Spongia officinalis*[82] and the dactylocyanines (e.g. dactylocyanine A **22**) from a Polynesian *Dactylospongia metachromia* sample.[104] Sponge metabolites are often produced by symbiotic microbes, however these organisms are often unculturable and therefore biosynthetic origin is difficult to prove.[109] The sponge *T. crypta* is the source of the spongonucleosides, including the anti-inflammatory spongosine (**23**). Comparative studies using molecular networking of sponge extract against six isolated bacterial strains revealed a *Vibrio harveyi* symbiont as the sole producer of **23**, demonstrating the use of GNPS in aspects other than novel compound discovery.[105]

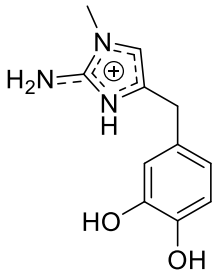
More recently, the chemical constituents of various algae have also been the subject of studies utilising GNPS.[81,110-113] Seasonal variations in the chemical composition of the Baltic seaweed *Fucus vesiculosus* were investigated using a molecular networking approach. A total of 44 compounds were dereplicated through comparison to the GNPS database (mainly primary metabolites), the quantity of which varied based on collection month with phlorotannin highest in summer, and pigments such as chlorophyll and carotenoids highest in the winter and spring.[110] The same research group also investigated 55 fungal strains associated with *F. vesiculosus*, and compared the metabolites produced under various conditions using molecular networking. The cytotoxicity of extracts was mapped onto the network, to reveal clusters of bioactive molecules and the strains specific culturing conditions to which they were produced.[111] A laboratory-grown *Derbesia* sp. chlorophyte collected in American Samoa was the source of the cyclic depsipeptide pagoamide A (**24**). Molecular networking of extract fractions revealed a cluster with no matches in the database that correlated with neuromodulatory activity, and this was traced to the presence of **24**.[113]



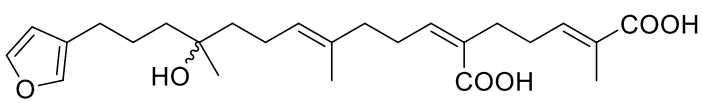
18



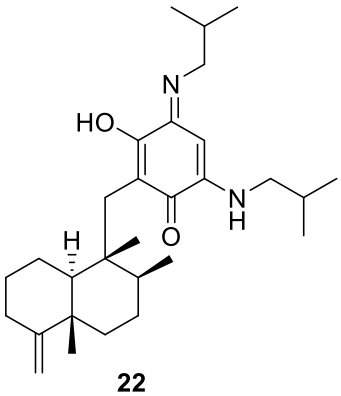
19



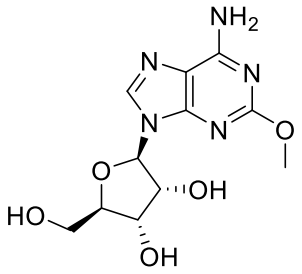
20



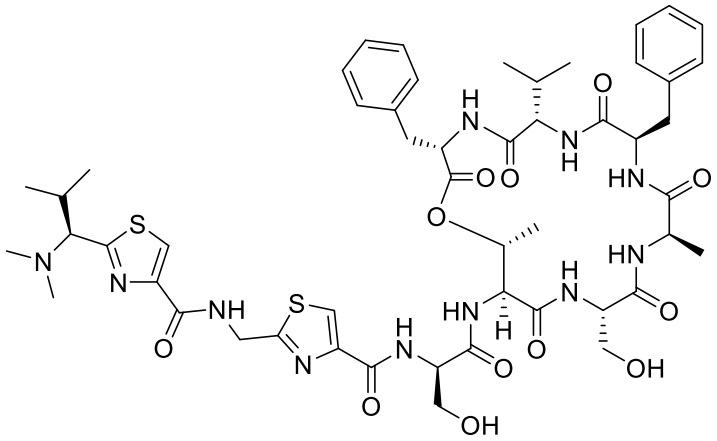
21



22



23



24

1.8 – Research Objectives

The research described in this thesis aimed at investigating the applicability of ^{1}H NMR spectroscopy and MS/molecular network-guided screening methods on a range of marine organisms collected in the South Pacific. Organisms were prioritised based on evidence of interesting secondary metabolites, with the structures of these compounds deduced primarily using NMR and MS techniques. The bioactivity of these compounds was assessed with the overall goal of isolating and identifying new MNPs for the discovery of new drug candidates.

In Chapter 2, I will go through the different approaches used for organism screening, including the benefits and drawbacks associated with each. This is followed by Chapters 3–6, which explore the chemistry of four of the screening hits in detail. Chapter 7 is comprised of concluding remarks which is then followed by experimental details in Chapter 8.

Chapter 2 – Organism Screening

In the context of natural products drug discovery, organism screening is an assessment of the novelty/bioactivity within a secondary metabolome of an organism and is often followed by a decision if it warrants a further, larger scale investigation. Even if the chemical composition looks interesting and the metabolites drug–like, it is essential to ensure that dereplication is the priority, and this may require a combination of analytical techniques.

For microbial natural products research, a bioassay–guided isolation is the preferred method for screening. The consumption of target compounds during an assay is not an issue, as more can be produced by larger scale culture to repeat the process. With non-sample limited microbial natural products, a variety of bioassays can be used, so broad-spectrum activity can be assessed and the chances of picking the wrong assays are reduced. Resources can also be saved by ensuring only positive hits are upscaled after activity is detected, as metabolite bioactivity is often detected at levels lower than that required for NMR spectroscopy-based detection.

However, if samples are limited, such as a marine invertebrate sample or plants from a collection expedition or historical collection, losing mass on a bioassay is undesirable. In a typical sized sample of our marine collection (~100 g), secondary metabolites are very often present in sub-milligram quantities, near the threshold of spectroscopy-based structure elucidation on the instruments at VUW (see Chapter 6). Therefore, over the past 20 years, a non-destructive spectroscopy-guided screen has been typically adopted where samples are returned after analysis. With the development of tools such as GNPS, MS has been adopted by the natural products community as a progressive tool for screening, and offers benefits of both tools. Like spectroscopy, it gives rapid insight into the structure of the metabolites and uses very little compound, and like bioassay, requires only the smallest quantities for analysis as MS can reliably detect molecules on the pico- to nanogram scale. In fact, this may be the biggest downfall with the detection of new and exciting molecules present in vanishingly small amounts that cannot be structurally assigned nor assayed for biological activity post–isolation. In this work, all three screening methods were employed across a range of marine samples with the goal of discovering new molecules for drug development.

2.1 – Bioassay Screening

The ultimate goal of drug discovery is to find novel bioactive molecules; therefore, it is logical to test the biological activity of extracts first. This explains why bioassays are the most commonly used screening technique in traditional natural product investigations. Where extracts show the desired activity, it is likely this is caused by a molecular component which is then targeted through the purification procedure. This method has some downfalls, however, particularly in terms of dereplication. The bioactive entity may also be present in a very small quantity, or the activity could be the result of synergy between multiple metabolites, both of which result in samples lacking the activity observed in the screen when purified. Finally, if the wrong assay is selected for the screen, bioactive molecules may give negative results. For these reasons and the scope of this research, this method was only carried out for one Pacific invertebrate, an undescribed tunicate.

2.1.1 – Purple Tunicate PTN3_40B (558)

Figure 2.1 – Surface photograph of the unidentified tunicate PTN3_40B (558).

In 2010 a range of extracts from organisms collected using SCUBA in Swallow`s Cave, Vava′u, Kingdom of Tonga, were screened for anti–mycobacterial activity, through a *Mycobacterium smegmatus* inhibition assay where a 'hit' was defined by >80% growth inhibition at 100 μg mL^{-1}.[114] *M. smegmatus* is used instead of *M. tuberculosis* for assay, as it is non-pathogenic, much easier to handle and has a faster reproductive rate.[115] Of the hits, an unknown purple ascidian (Figure 2.1, 40 g) was selected for further study in this work, due to its larger biomass than the others. The methanolic extract was purified through HP20 (PSDVB) using mixtures of H_2O and Me_2CO, followed by HP20ss (PSDVB) using mixtures of H_2O and MeOH. Owing to the time consuming and expensive nature of the bioassays, the purification from here on was guided by ^{1}H NMR spectroscopy and MS. Based on the colour eluting from

the column and the fractions analysed by MS, it was clear that two major coloured components were present (Figure 2.2), one orange and one pink. These compounds were then further purified using normal- and reversed-phase column chromatography and reversed-phase HPLC.

The most abundant compound was isolated as an orange solid, and HRESIMS analysis detected a protonated molecular ion at *m/z* 391.1027 ($[M+H]^+$) indicative of the molecular formula $C_{21}H_{19}N_4O_2S$ (calcd. 391.1223). NMR experiments were run in DMSO–d_6 on the compound, showing deshielded signals corresponding to aromatic protons, broad ^1H resonances of exchangeable protons, three methylenes and one methyl from an acetate group (Appendix 1). This data was used to search the MarinLit database, with shermilamine B (**25**) being the clear hit. A comparison of the NMR data from the original isolation confirmed this.[116]

The second, pink coloured compound was subjected to HRESIMS and analysis detected a protonated molecular ion at *m/z* 361.0925 indicative of the molecular formula $C_{20}H_{17}N_4OS$ (calcd. 361.1118). This was rapidly dereplicated as kuanoniamine D (**26**)[117] as it is often co–isolated from tunicates with shermilamine B;[118,119] and the ^1H NMR data matched that of the literature values.

Several minor compounds were detected; however, they were only present in very small amounts and as many other congeners of these compounds have been reported, this tunicate was not investigated further. This example illustrates the downfall of bioassay–guided fractionation, often resulting in the isolation of known compounds after already expending valuable resources. This class of pyridoacridine molecules are of interest as anti-tuberculosis (TB) agents, however both of these molecules have previously been tested and showed only mild activity.[120]

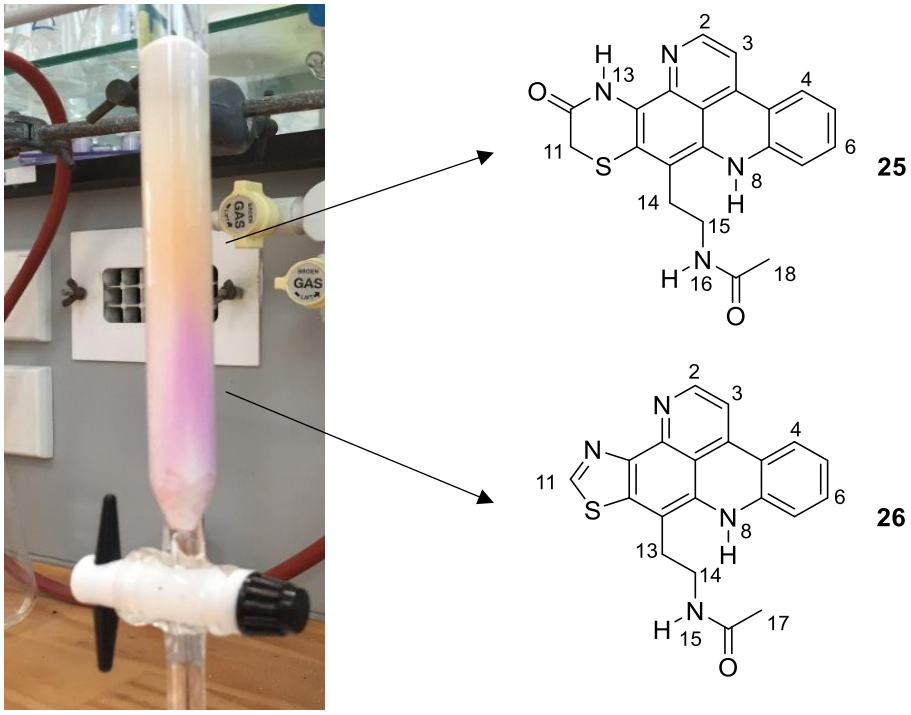


Figure 2.2 – HP20ss column resulting in separation of **25** and **26**.

2.2 – NMR Spectroscopy Screening

Spectroscopy-based methods, i.e. NMR, have formed the main basis of the screening protocols for our research group over the past 20 years. These have included 1D ^{1}H NMR, and more detailed HSQC and HMBC screens, often with computational elements.[29,121,122] The data output gives clear indications of functionalities present for the metabolites, and this, along with taxonomic identification, can be used to postulate the types of molecules within a given extract. The process can be time consuming, as extracts generally require a pre-fractionation step as the initial concentration of compounds in the extract may be too low for detection. Primary metabolites/media components may also interfere with metabolite detection, and solubility of complex extracts can be an issue, therefore this research group has employed a reversed-phase chromatography protocol to enrich extract fractions with amphiphilic drug-like molecules before screening (Chapter 1.5).

In this work, various marine samples were analysed by following the standard HP20 protocol (Experimental: Chapter 2) to generate three fractions for a ^{1}H NMR spectroscopy screen. Resonances corresponding to potentially new secondary metabolites were then used to prioritise samples for bulk extraction and analysis. Very polar compounds such as salts and sugars, are usually removed in the early stages of prefractionation, whereas those non-polar

metabolites are often carried through and eluted in the 75 or 100% fractions. These are easily identified by broad peaks in the ^{1}H NMR spectrum such as the lipid methylene envelope at ~1.4 ppm, or sites of unsaturation of fatty acids at ~5.3 ppm. Aromatic methines in the 6–8 ppm region, oxygenated methines with sharp splitting in the 3–6 ppm region and intense methoxy singlets from 3–4 ppm represents just some examples of diagnostic signals that stand out to the trained eye as a positive hit. These NMR peaks are then followed through the isolation procedure to arrive at a pure compound. Relative to other screens, this method is time consuming due to the larger quantities that must be extracted and then pre-fractionated, to produce enough mass for NMR signals to be attained, particularly for 2D screening experiments. It can also be more costly as deuterated solvents must also be used.

All of the invertebrates screened in this way were collected in the Kingdom of Tonga over the course of several years and multiple collection trips, while the alga was collected in NZ.

2.2.1 – Tunicate PTN4_01A (632)

Figure 2.3 – Surface photograph of the unidentified tunicate PTN4_01A (632).

This tunicate is regularly encountered by the SCUBA divers in Tongan waters, and was collected at Kitahi Beach, Vava′u, Kingdom of Tonga. Following extraction, it was screened by ^{1}H NMR spectroscopy following HP20 fractionation. For all of the fractions, very few resonances of interest were observed in the NMR spectra (Appendix 3), therefore this extract was not taken further.

2.2.2 – Tunicate PTN4_40E (855)

Figure 2.4 – Surface photograph of the unidentified tunicate PTN4_40E (855).

This unidentified tunicate was collected in Pete's Cave, ′Eua, Kingdom of Tonga. The 75% Me_2CO screening fraction ^1H NMR spectrum had many interesting resonances from 6–9 ppm so was bulk extracted and stored at 4 °C (Appendix 3). However, the extract decoloured in storage prior to further purification, and a subsequent ^1H NMR analysis did not show resonances of interest, suggesting the compounds had broken down and therefore this sample was not taken further.

2.2.3 – Tunicate PTN4_03C (650)

Figure 2.5 – Surface photograph of the unidentified tunicate PTN4_03C (650).

This unidentified tunicate was collected from Hunga Island, Vava′u, Kingdom of Tonga. The ^1H NMR spectrum of the 30% Me_2CO/H_2O screening fraction does have many interesting resonances (Appendix 3), however these were of very low intensity with the total fraction weighing only 1.0 mg. This sample was not taken further based on limited mass.

2.2.4 – *Polyandrocarpa polypora*

Figure 2.6 – Surface photograph of the *Polyandrocarpa polypora* PTN4_24F (779).

The tunicate identified as *Polyandrocarpa polypora* was collected all around Vava′u and ′Eua, Kingdom of Tonga. This organism has been the subject of previous investigations within the VUW MNP group,[123,124] which identified an unstable metabolite of interest. In this work, the 75% Me_2CO/H_2O screening fraction showed many interesting deshielded ^1H NMR peaks. The major compound (**27**) was partially purified by reversed-phase HPLC, however degraded before full structure elucidation could be completed. No suitable molecular formula could be identified using HRESIMS however, some substructures could be put together from the acquired NMR data (Appendix 2).

The ^1H NMR spectrum shows two exchangeable resonances at δ_H 11.26 (1H) and 6.10 (2H), three mutually coupled aromatic methines (δ_H 8.13, 6.86 and 6.66) and two aromatic singlet methines (δ_H 7.63 and 7.08). The ^{13}C NMR spectrum suggested the presence of at least 15 aromatic carbon environments, of which five were accounted for by the methines. The coupling pattern of the aromatic methines suggested a 1,2,4-trisubstituted aromatic ring, with HMBC correlations assigning the chemical shifts of the six carbon atoms in the ring. The chemical shift of the least shielded carbon C–4 (δ_C 155.3), suggested this carbon was oxygenated. The putative exchangeable NH shift δ_H 11.26, showed correlations in the HMBC spectrum to C–1 and C–2 (δ_C 114.8 and 141.0 respectively) and a nuclear Overhauser effect NOE (ROESY) through–space correlation to H–3 (δ_H 6.86), thus a secondary amine was likely substituted at C–2. The putative NH also showed an HMBC correlation to δ_C 119.8; however, this carbon

showed no other correlations so is likely more than three bonds from another proton, thus substructure I was deduced (Figure 2.7).

The other two singlet aromatic methines H–3' and H–6' (δ_H 7.63 and 7.08) are positioned *para* to one another on an aromatic ring, sharing HMBC correlations to three deshielded aromatic carbons at δ_C 149.2, 144.6 and 142.1, while H–6 also correlated to δ_C 115.3. This is all that could be deduced for substructure II (Figure 2.7), as no correlations were detected outside of this system, nor to the 2H exchangeable resonance at δ_H 6.10. As the compound degraded before further experiments such as ^{15}N-HSQC/HMBC NMR experiments could be completed, the structure of this compound remains unknown.

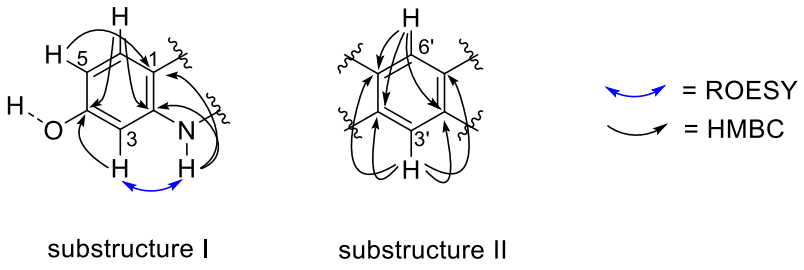


Figure 2.7 – Key HMBC and ROESY correlations supporting substructure I and II for **27**.

2.2.5 – *Plocamium angustum*

The *Plocamium* genus of red algae is a prolific producer of natural products, and are particularly well known for their halogenated monoterpenes.[125] Monoterpenes without ionisable functionalities are often difficult to detect by softer ionisation techniques such as electrospray ionisation (ESI), therefore a 2017 collection of alga identified as *P. angustum* from Moa Point, Wellington, NZ, collected for an undergraduate teaching experiment, was screened by ^{1}H NMR spectroscopy. The 75% Me_2CO in H_2O fraction showed numerous deshielded methylene/methine signals, shielded methyls and peaks in the olefinic range appropriate for a terpenoid compound (Figure 2.8). The results are discussed in Chapter 3.

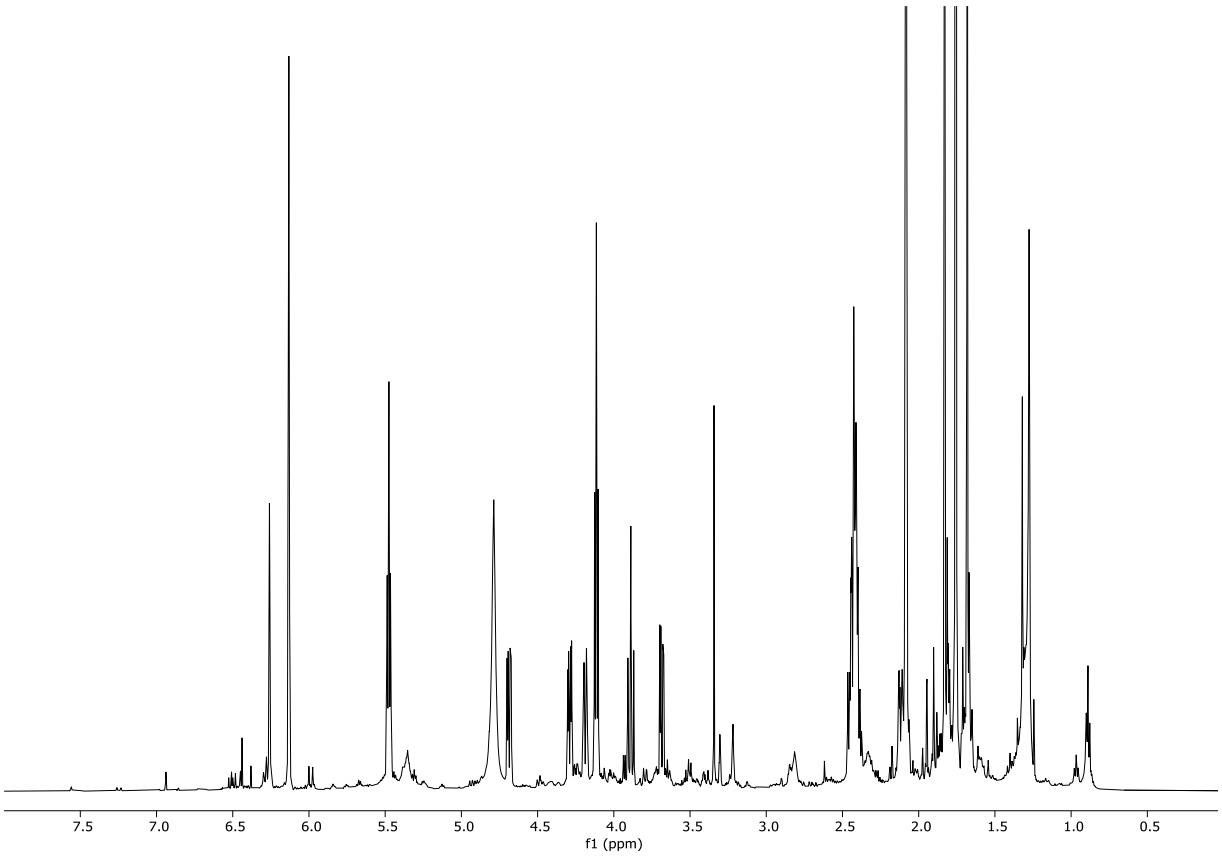


Figure 2.8 – ^{1}H NMR Spectrum of the 75% Me_2CO/H_2O screening fraction of "*P. angustum*" (600 MHz, CD_3OD).

2.2.6 – *Nelliella nelliiformis*

The turfing bryozoan *Nelliella nelliiformis* was collected from Cathedral Cave ʻEua, Kingdom of Tonga. Analysis of the 30% Me_2CO/H_2O screening fraction showed numerous interesting signals, particularly aromatic methines and deshielded methylenes/methines (Figure 2.9). The results of the ensuing chemical investigation are elaborated in Chapter 4.

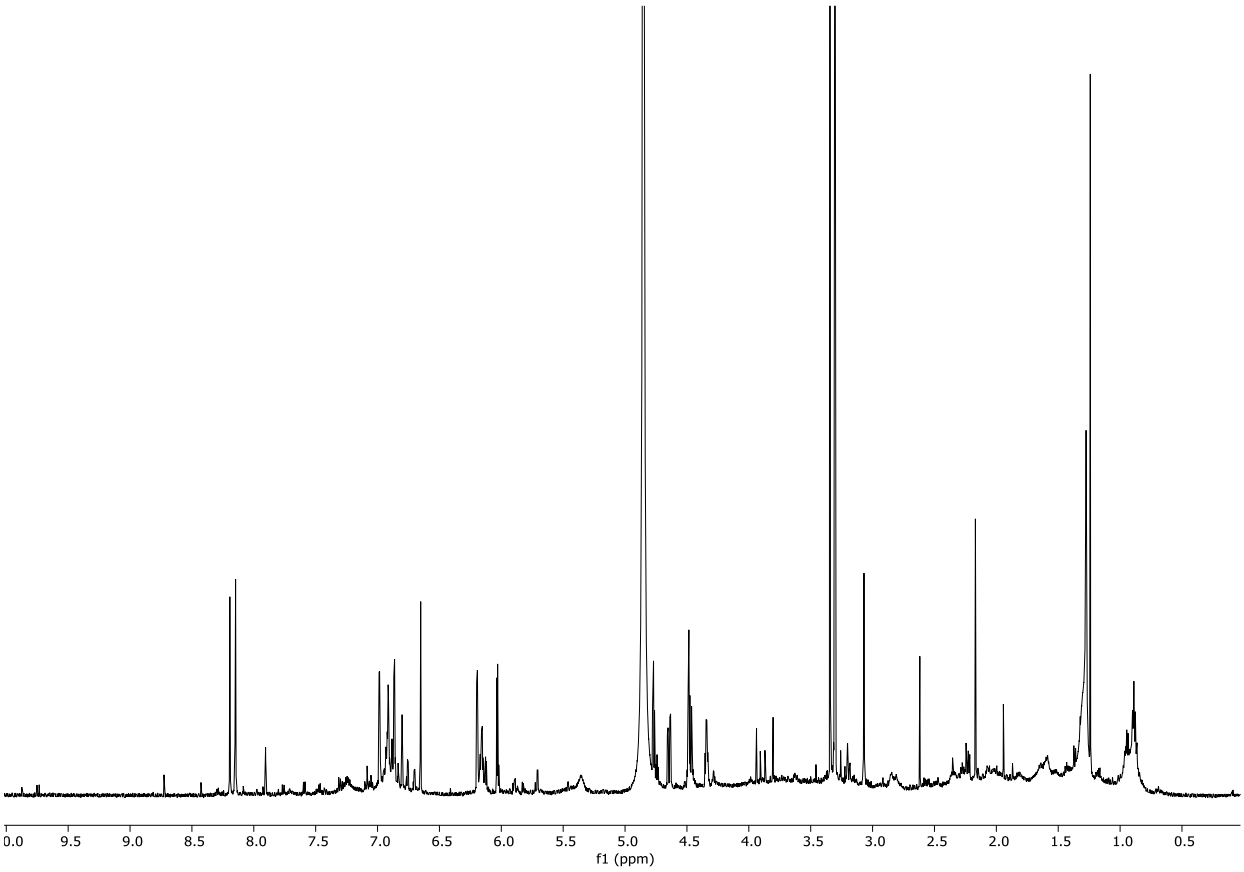


Figure 2.9 – ^1H NMR Spectrum of the 30% Me_2CO/H_2O screening fraction of *N. nelliiformis* (600 MHz, CD_3OD).

2.3 – Mass Spectrometry Screening

As described in Chapter 1, LCMS/MS is well suited for screening complex mixtures of molecules and over the past seven years, GNPS has proved to be an effective tool for the discovery of a huge number of new molecules.[80,107] Although NMR methods provide structural information of the molecules within the fraction, it is often difficult to assign peaks to a particular entity if they are of a similar intensity or chemical shift, providing challenges for dereplication within mixtures of compounds. LCMS/MS provides an initial separation, typically over a polarity gradient, which helps resolve this issue and separate compounds from a crude extraction without any prior purification. Also, by examining a range of *m/z* values, different compounds that were not resolved can still be detected separately from each other. The sensitivity of the technique also allows the detection of molecules that may be missed by NMR. Based on the *m/z* of a parent ion, a molecular formula can be deduced and compared to databases to screen for novelty. Even when the molecular formula is previously described, fragmentation data can be analysed to assess if it is consistent with the compound or if it suggests an isomeric compound. Before complete attention is focused on one ion, it is essential it is present in sufficient intensity relative to the extract for NMR-based structure elucidation. At a screening level, this can be approximated by examining the precursor ion intensity and, if a chromophore is present, manually looking at the ultraviolet (UV) profile of the LCMS/MS trace.

In terms of screening a particular organism, detecting novel molecules by molecular networking through the GNPS platform can be troublesome, as it generates a lot of data and noise as many metabolites form clusters. If the organism is from a group of 'biosynthetically-talented organisms' and data has been previously reported (e.g. classes of compound), then molecular networking is very useful to detect anomalies and differences from prior studies, such as new derivatives or compound modifications. The key strength of molecular networking lays in the general screening of a range of organisms, by displaying numerous datasets in one representation. This allows for a simple visualisation of the chemical space occupied by the extracts relative to each other when nodes are layered (e.g. coloured) for the individual source organism.

When set to display nodes based upon the source organism, constellations of many colours typically depict primary metabolites, as these are common across many species, or otherwise can result from common artefacts of the experiment (e.g. extraction solvent, column impurities), and are therefore ignored. As secondary metabolites are the focus for drug

discovery, and are typically restricted to limited numbers of taxa, constellations made up from a single organism suggest a distinct metabolic class desired for prioritisation. The more nodes in the network, the more different molecules comprise it and are suggestive of a class of structurally similar metabolites. This is of great benefit as structural modifications allow structure activity relationships (SAR) to be performed once bioassay has been completed. Groups made up of only one or two nodes are produced in large numbers and therefore metabolites produced without congeners can be lost in the noise. This is one of the weaknesses of the method, therefore deprioritised extracts should ideally also be screened by complementary or alternative methods.

Following this, the precursor ion of a node can be used to determine if the constellation potentially contains evidence for novel molecules, providing a hypothesis of which metabolites to target down the purification line. From the beginning, the precursor ion can be followed, and this can reduce the need of time–consuming analytical methods such as thin–layer chromatography (TLC) when combining fractions following chromatography. In this work, ^{1}H NMR spectroscopy was also used after fractions had been pooled to ensure the targeted metabolites were present in quantities sufficient for structure elucidation.

In order to screen Pacific tunicates of interest, a methanolic extract of a small amount (~5 g) was analysed by untargeted (+)- and (–)-ion mode LCMS/MS and put through the GNPS workflow (Experimental: Chapter 2). Figure 2.11 shows the full molecular network analysis of the tunicates, where each individual organism is represented by a single colour. Extracts were initially run at 1 mg mL^{-1} where a 10 μL injection (10 μg) showed more nodes with connectivity than 1 μL. As the extracts contain metabolites across a wide polarity range, a gradient from 2.5% to 100% ACN in H_2O over 30 min was used for general separation. The best results were obtained using an active exclusion (i.e. the precursor ion is ignored after it has been fragmented for three sequential scans) for 0.3 min. This is long enough to ensure less abundant metabolites are also fragmented, but not too long so as to miss isomeric compounds eluting after initial selection. The collision induced dissociation (CID) energy for each selected precursor ion was tailored proportionally to the parent mass to charge ratio, with heavier ions bombarded with higher energy to ensure sufficient fragmentation occurred. Every ion was fragmented with a high and low energy (both in proportion to *m/z*), to account for molecules of different fragilities (Figure 2.10).

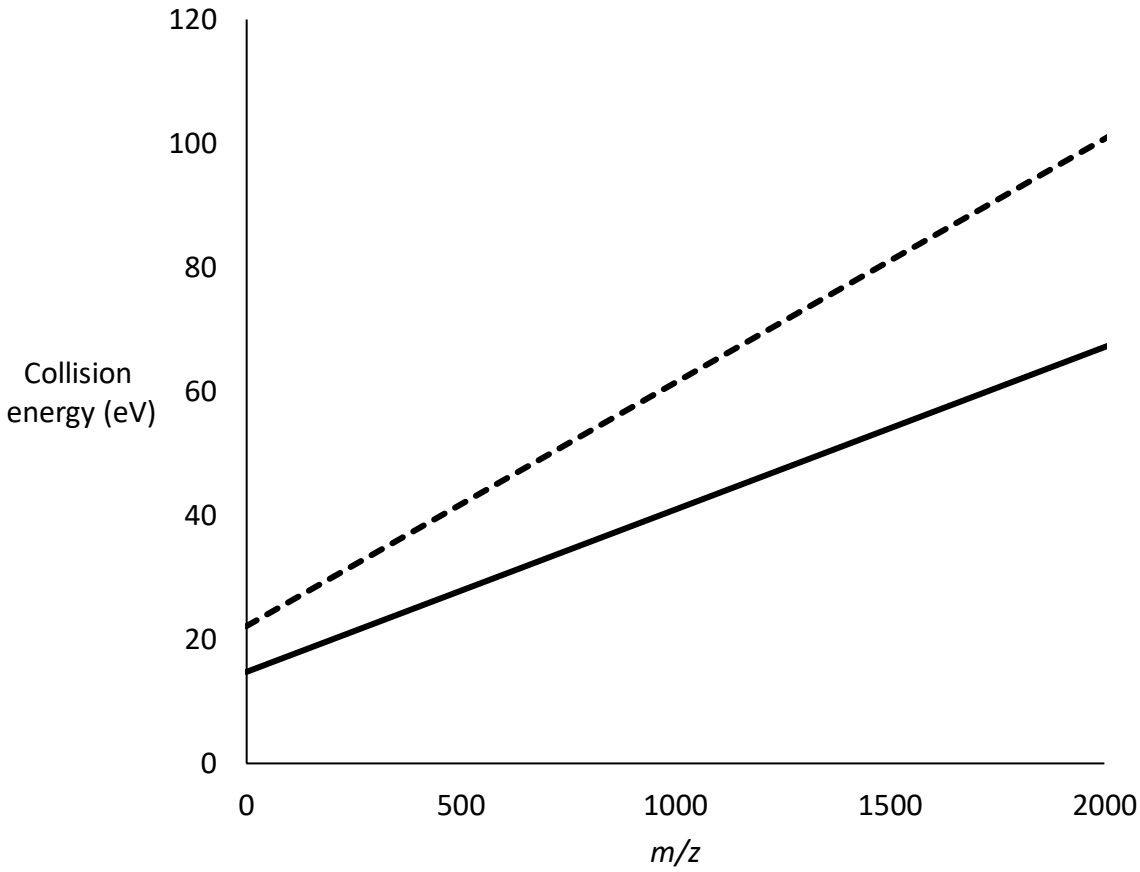


Figure 2.10 – Plot of collision energy tailored to precursor *m/z* for untargeted LCMS/MS. Full line is low energy (gradient 2.62 and offset 14.75) and dotted line is high (gradient 3.93 and offset 22.13).

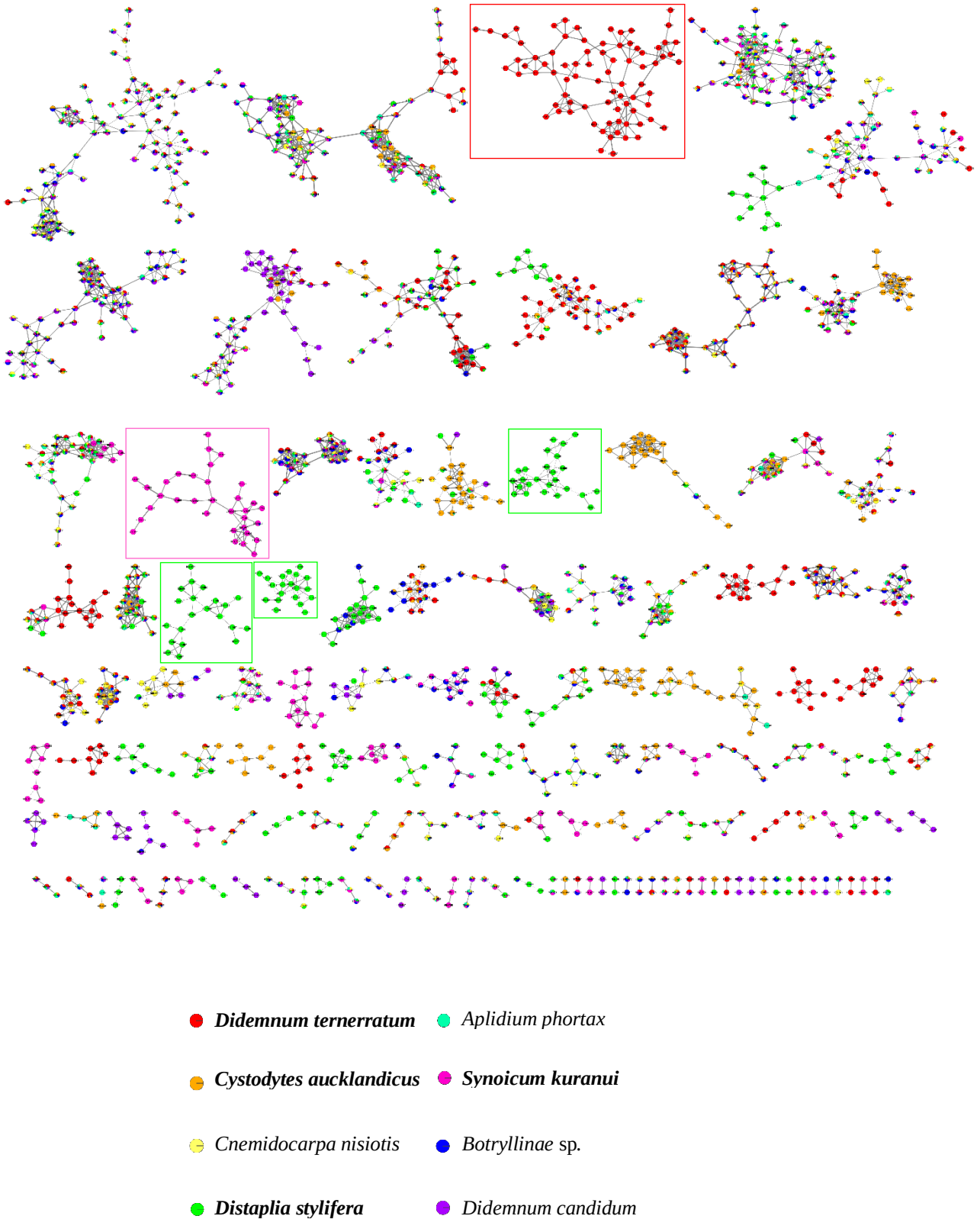


Figure 2.11 – Full molecular network of Pacific tunicates using both positive- and negative-ionisation modes. Red box represents constellation A; pink box represents constellation B; green boxes represent constellations C. Bold species were further studied in this work.

2.3.1 – *Cystodytes aucklandicus*

C. aucklandicus was collected from North Cape, NZ, in 1999 as part of the National Institute of Water and Atmospheric Research (NIWA) collection. In the network, the constellations predominantly made up from this extract also contain a small number of nodes from other organisms (Figure 2.12). This organism was selected for further investigation by ^{1}H NMR, to act as a negative control for the networking screen, as it would typically not be selected for further analysis initially based on the MS molecular networking result. As seen in the ^{1}H NMR spectra of all three screening fractions (Appendix 3), the extract of *C. aucklandicus* lacked resonances typically associated with secondary metabolites, instead showing evidence of unsaturated lipids. This organism was not taken further.

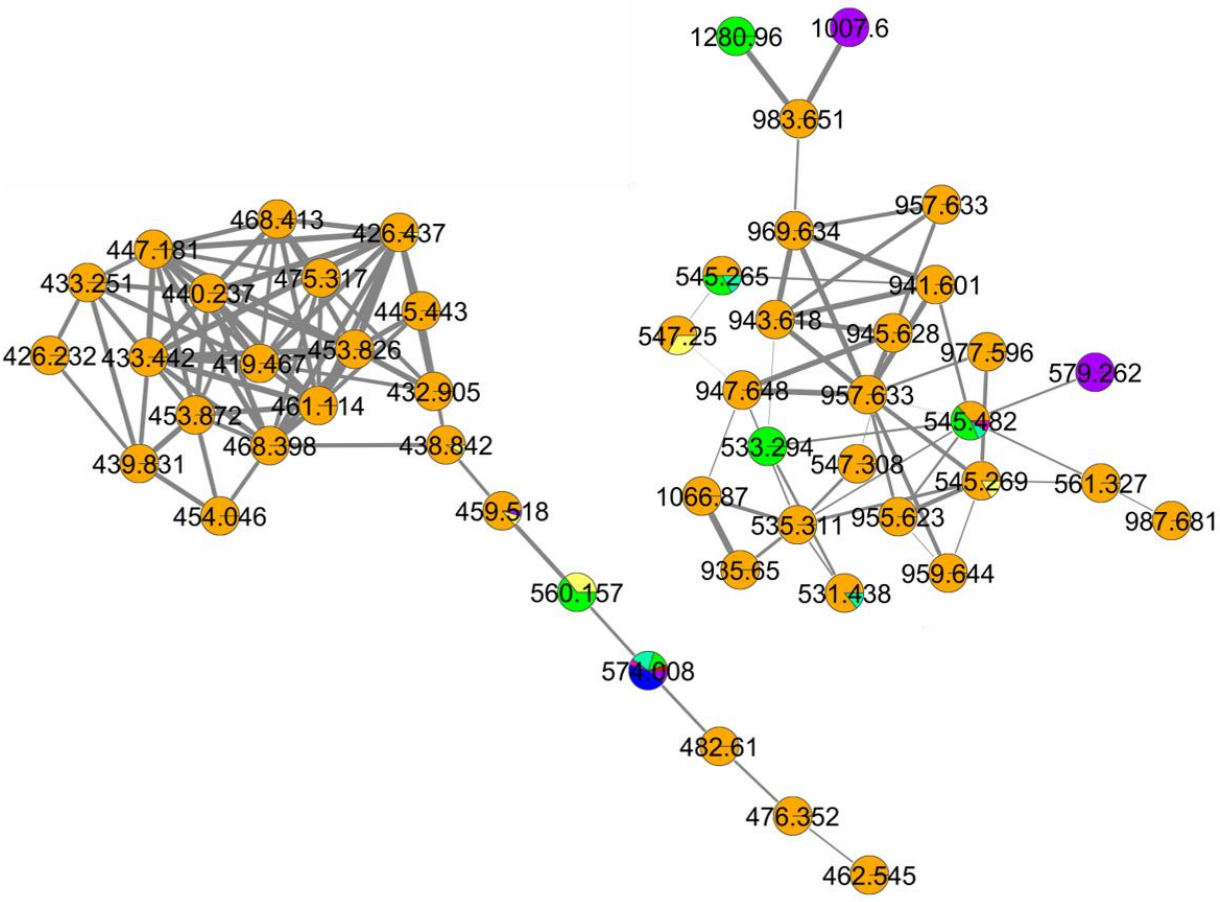


Figure 2.12 – Two constellations with nodes comprised mainly from the *C. aucklandicus* extract. Nodes labelled with precursor ion, thickness of edge (connecting line) is dependent on cosine (similarity) score.

2.3.2 – *Synoicum kuranui*

S. kuranui is a red globular tunicate that was collected from Great Barrier Island, NZ, in 1999 as part of the NIWA collection. Based on the large constellation containing pink-coloured nodes (constellation B, Figure 2.13) this organism was prioritised for bulk extraction and further analysis. These results are presented in Chapter 5.

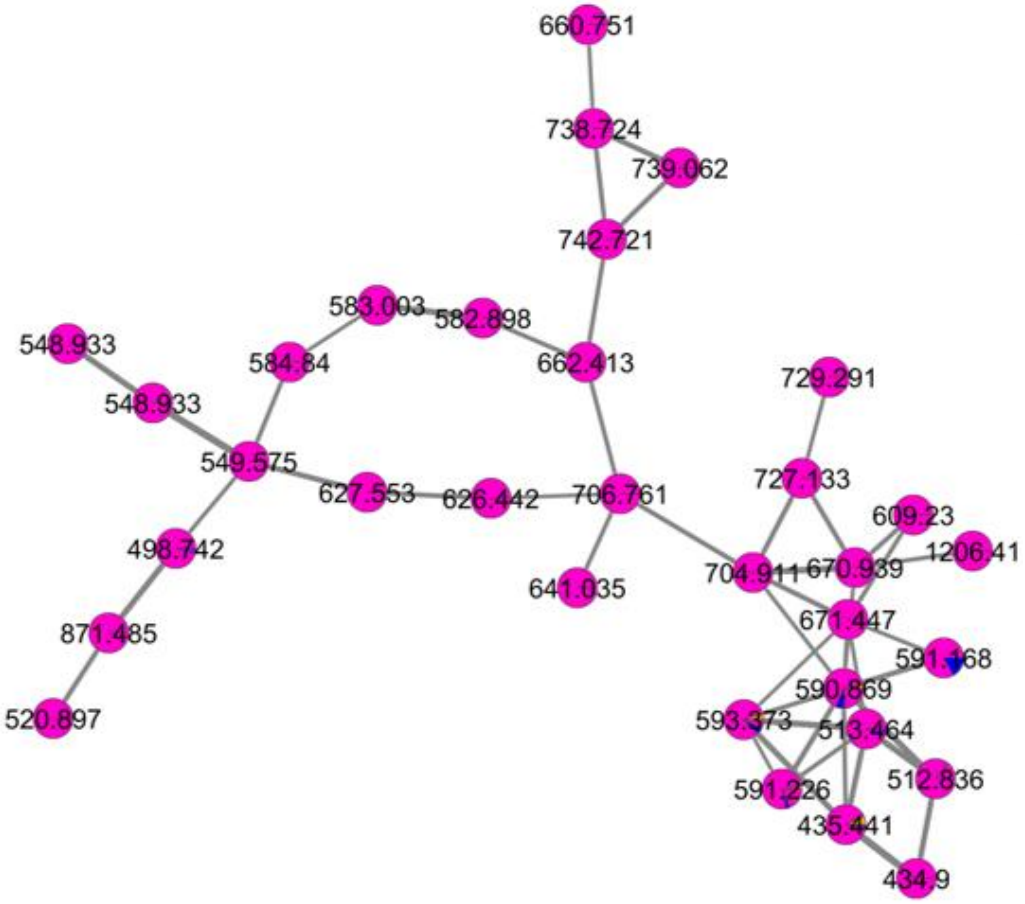


Figure 2.13 – Constellation B with nodes comprised only from the *S. kuranui* extract. Nodes labelled with precursor ion, thickness of edge (connecting line) is dependent on cosine (similarity) score.

2.3.3 – *Didemnum ternerratum*

D. ternerratum is a purple encrusting tunicate that was collected from Pete's Cave, ʻEua, Kingdom of Tonga, in 2016 as part of the VUW MNP research group collection. Based on the large constellation containing red-coloured nodes (constellation A, Figure 2.14) this organism was prioritised for bulk extraction and further analysis. These results are presented in Chapter 6.

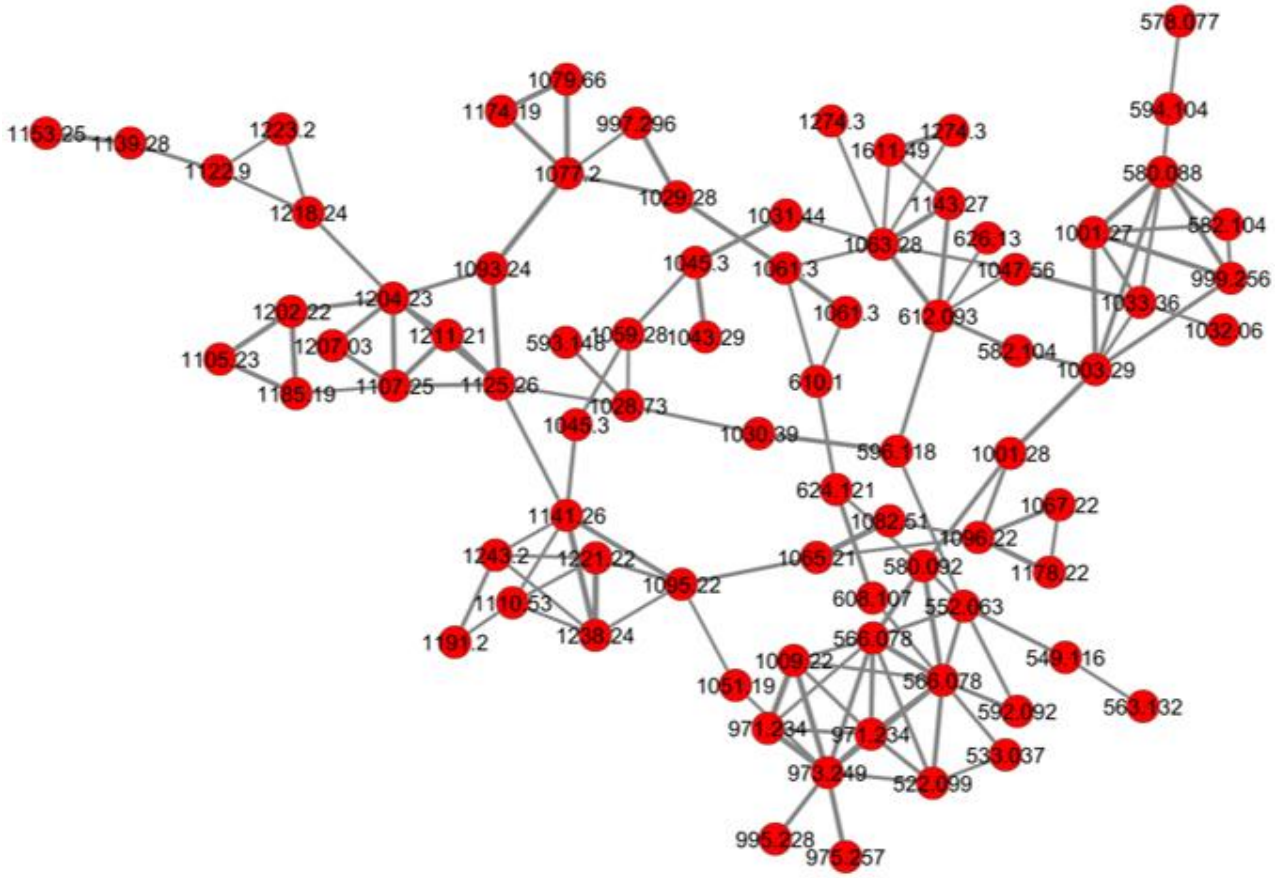


Figure 2.14 – Constellation A with nodes comprised only from the *D. ternerratum* extract. Nodes labelled with precursor ion, thickness of edge (connecting line) is dependent on cosine (similarity) score.

2.3.4 – *Distaplia stylifera*

The metabolites from the methanolic extract of *D. stylifera* are represented by the green nodes in the screening network (Figure 2.11) and based on the many large standalone constellations (constellations C, Figure 2.15), the extract was prioritised for further examination.

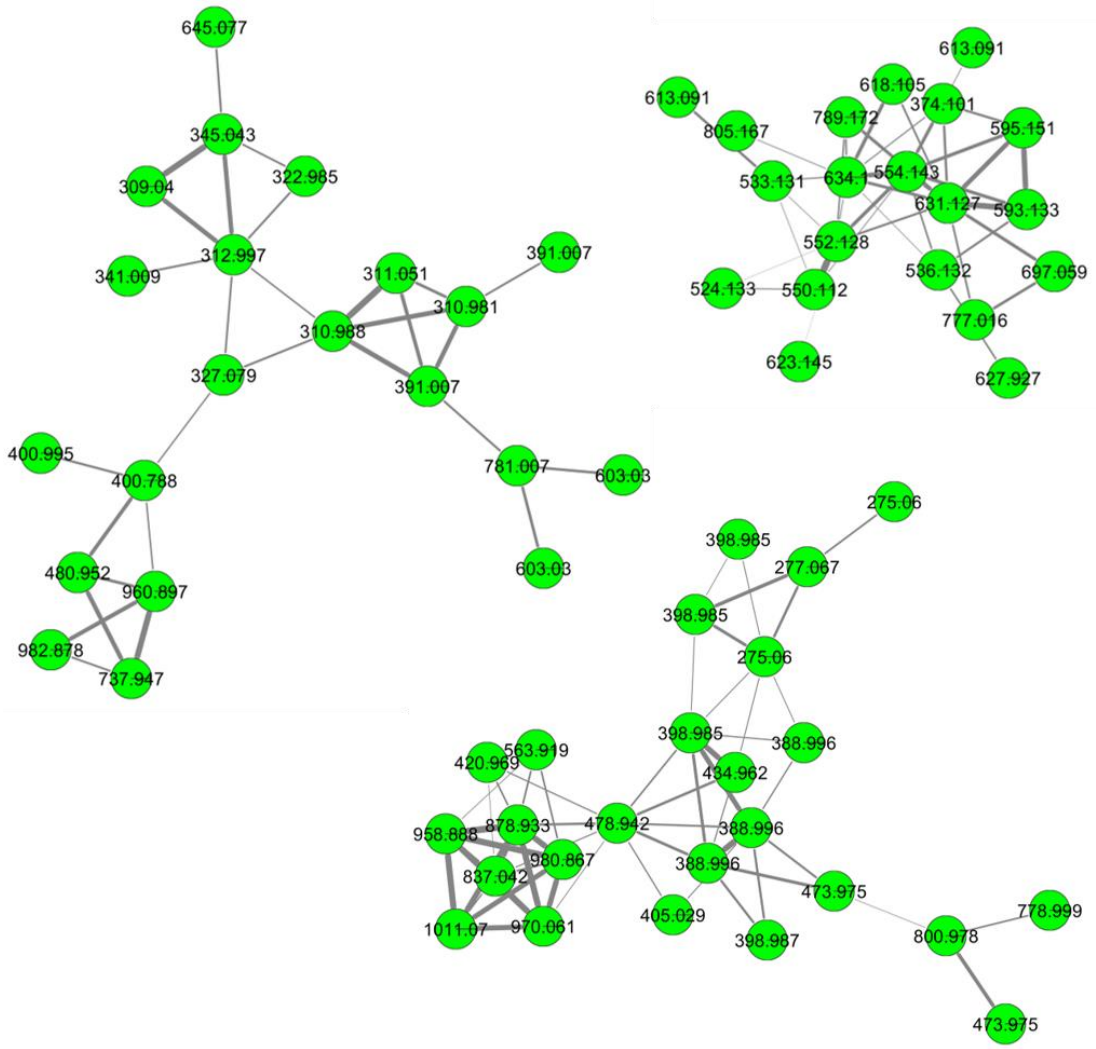


Figure 2.15 – Constellations C with nodes comprised only from the *D. stylifera* extract. Nodes labelled with precursor ion, thickness of edge (connecting line) is dependent on cosine (similarity) score.

Following HP20 prefractionation, ^{1}H NMR spectroscopy was run with the more polar 30% Me_2CO/H_2O screening fraction showing many peaks of interest (Appendix 4), particularly in the aromatic region. *D. stylifera* is a member of the Holozoidae ascidian family and was collected by SCUBA at a depth of 6 m, Cape Reinga, NZ, as part of the NIWA collection in April 1999. Although no compounds have been reported from the species to date, this genus has been the source of nitrogenous secondary metabolites, with 3,6-dibromoindole isolated from *D. regina* collected in Palau,[126] and a South African *D. skoogi* sample the source of both 6-bromo–3-chloroindole and 6-bromo-2-oxindole.[127] The benthic species *D. cylindrica* is widely distributed throughout Antarctica, however in field studies there has been no evidence of predation or fouling from epibionts, and the species was therefore proposed to harbour bioactive deterrents.[128] Thus other species of the genus are worthy of chemical investigation.

Separation by high pressure liquid chromatography (HPLC) was performed on the polar 30% Me_2CO/H_2O screening fraction, with a number of different UV active peaks detected. One major metabolite, compound **28**, was isolated as a yellow oil. The sample was analysed by HRESIMS analysis, detecting a protonated molecule at *m/z* 219.0794, consistent with the molecular formula $C_{11}H_{11}N_2O_3$ (calcd. 219.0764). The ^{13}C NMR spectrum showed 10 deshielded signals (one ketone at δ_C 179.0) and one methoxy resonance, while the ^{1}H NMR spectrum contained signals for the corresponding methoxy singlet (δ_H 3.84) and three coupled aromatic methines (δ_H 6.91 (d, J = 8.3 Hz), 8.11 (d, J = 2.0 Hz) and 8.31 (dd, J = 8.3, 2.0 Hz)), that were indicative of a 1,2,4-trisubstituted benzene ring (Table 2.1 and Appendix 3). It also contained a further two broad, deshielded methines (δ_H 7.26 (br. s) and 7.45 (br. s)) that were coupled to an exchangeable proton at δ_H 13.28 (br. s).

Aside from the carbons of the two broad methines defined by the HSQC (δ_C 121.2 and 130.6 for C–2 and C–4, respectively), only one non-protonated aromatic carbon, C–5 (δ_C 145.2), correlated with both protons in the HMBC spectrum. The only other detected correlation was H–4 (δ_H 7.26) with ketone C–6 (δ_C 179.0), suggesting this substructure is a substituted imidazole ring (Figure 2.16). This ^{13}C chemical shift is relatively low for a ketone (~ 200 ppm), which suggested it was *α,β*-unsaturated. Further evidence for this moiety came from the MS/MS spectrum where the corresponding imidazole acylium ion at *m/z* 95.0302 was detected upon CID at low energy (20.0 eV), suggesting a fragile link to the remainder of the compound.

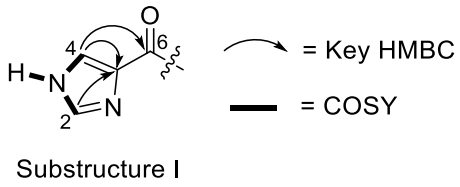


Figure 2.16 – Key HMBC and COSY correlations supporting **28** substructure I.

A second spin system consisted of a 1,2,4-trisubstituted benzene ring, with two of the three substituted carbons C–9 and C–10 being highly deshielded (δ_C 147.2 and 152.0, respectively), which indicated oxygenation. The methoxy group was assigned at C–9, as it shares a through space NOESY correlation in the ROESY spectrum to H–8, and thus based on the molecular formula, a phenol was assigned at C–10. As the ketone C–6 also showed HMBC correlations with all three aromatic methines of this spin system, this must be connected directly to the ring at the other substituted carbon C–7 (δ_C 127.3). The standard HMBC pulse sequence is optimised for three-bond correlations, and the intensity of the $^4J_{H–11,C–6}$ correlation (four bonds) is much weaker than correlations from either H–8 or H–12 (three bonds). Thus, substructure II was deduced as vanillone (Figure 2.17), with further evidence again from the presence of an acylium ion based on an ion in the MS/MS spectrum at *m/z* 151.0480. This completed the structure elucidation of **28**.

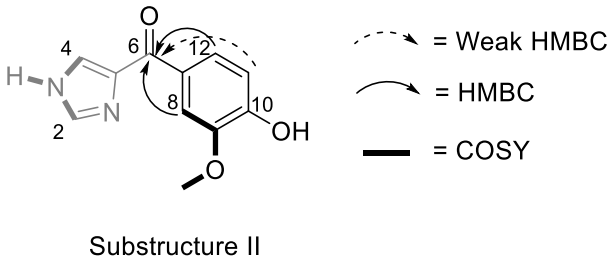
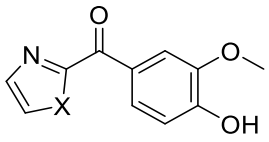
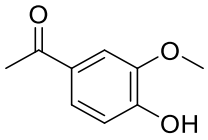


Figure 2.17 – Key HMBC and COSY correlations supporting **28** substructure II.

Compound **28** is (4-hydroxy-3-methoxyphenyl)-1*H*-imidazol-5-yl-methanone. This structure has not been reported in the literature, however the analogous compound **29**, substituted at the 2-position of the imidazole ring, was reported in 1988 from the Australian ascidian *Aplydium pliciferum* alongside thiazole–containing **30**.[129] The ^{1}H NMR spectrum of compound **29** differs in that the imidazole protons only give rise to one peak as they are magnetically equivalent. Although ^{13}C NMR data were not reported for **29**, the ketone carbon in compound **30** showed a resonance at δ_C 182.0 (75 MHz, $CDCl_3$), consistent with the data for compound **28**.



29 X=NH
30 X=S

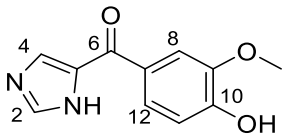


31

Compound **28** was submitted to SBS for antibacterial testing where it did not show any activity against Gram-positive (*Bacillus subtilis*) or Gram-negative (*Escherichia coli*) bacteria at 128 μg mL^{-1}. Although the biological activities of **29** and **30** were not reported, the related compound apocynin (**31**) has been extensively studied for its antioxidant activity.[130,131] It is believed apocynin forms an oxidative dimer in the presence of H_2O_2, leading to the active compound which also inhibits NADPH-oxidase.[132] Compound **28** was not tested for antioxidant activity in this study due to time restrictions, however this assay, alongside the metabolites total synthesis, are part of ongoing studies.

Table 2.1 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for **28** in DMSO-d_6.

position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC
NH–1	–	13.28 (br. s)	2, 4	
2	121.2	7.45 (br. s)	NH–1	4, 5
4	130.6	7.26 (br. s)	NH–1	2, 5, 6
5	145.2	–		
6	179.0	–		
7	127.3	–		
8	113.9	8.11 (d, 2.0)	12, OCH_3–9	6, 9, 10, 12
9	147.2	–		
10	152.0	–		
11	114.9	6.91 (d, 8.3)	12	6, 7, 8, 9, 10
12	126.3	8.31 (dd, 2.0, 8.3)	8, 11	6, 8, 9, 10
OCH_3–9	55.6	3.84 (s)	8	9
OH–10	–	10.09 (br. s)		



28

Chapter 3 – "*Plocamium angustum*"

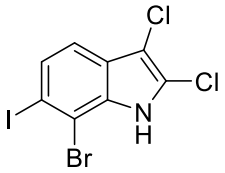
*Chapter 3 describes chemical investigation of a sample identified as the red algae "*Plocamium angustum", *collected from Moa Point, Wellington, NZ. This organism was prioritised based on many interesting signals associated with halogenated monoterpenes in the screening fraction ^{1}H NMR spectrum. This resulted in the isolation of the new cyclic polyhalogenated monoterpene, costatone C (**46**), along with the previously reported (1*E,5Z*)-1,6-dichloro-2-methylhepta-1,5-dien-3-ol (**47**). The structures of both compounds were solved by NMR, with the absolute configuration of **46** determined by comparison of experimental and DFT-calculated ECD spectra, whereas that of **47** was solved using Mosher's method. Compound **46** showed mild antibacterial activity against* Staphylococcus aureus *and* S. epidermidis.

3.1 – *Plocamium* sp. Natural Products

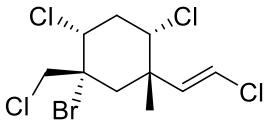
Red algae are seaweeds of the phylum Rhodophyta and have been the source of many new natural products over the past 50 years. Extensive studies have revealed the Rhodophyta as the most prolific producers of halogenated compounds of all macroalgae, with the privileged genera *Laurencia* and *Plocamium* yielding numerous new chemical entities.[133] Red algae have been the focus of recent studies at VUW.[134] A sample of *Callophycus serratus,* collected in Tonga, led to the isolation of six new polyhalogenated meroditerpenoids,[135] while locally collected *Rhodophyllis membranacea* was the source of eleven new tetrahalogenated indoles, including the first example to simultaneously possess the three halogen elements chlorine, bromine and iodine (**32**).[136]

Red algae of the genus *Plocamium* are widely distributed around the world, with chemical studies focusing on species collected from Europe and USA to South Africa, Australia and NZ. Reports of new natural products from the genus date back to 1974, where violacene (**33**) and cartilagineal (**34**) were described from *P. violaceum* and *P. cartilagineum*, respectively.[137,138] Both compounds are polyhalogenated monoterpenes, a class of molecule that the genus would become known as a valuable producing source. At the time of isolation, their structure elucidation was extremely challenging with the techniques available, and correctly assigning the halogen positions was troublesome without X-ray data. These compounds are produced by the alga and act as feeding deterrents to herbivores such as fish. This has been demonstrated by feeding inhibition and reduction in grazing observed in the field (Grenadine Islands and

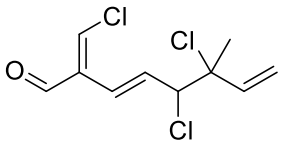
Guam)[139] or in the laboratory (damselfish)[140] caused by cyclised red algal metabolites **35** and **36**.



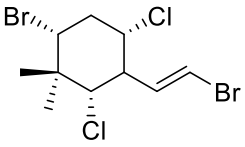
32



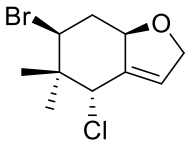
33



34

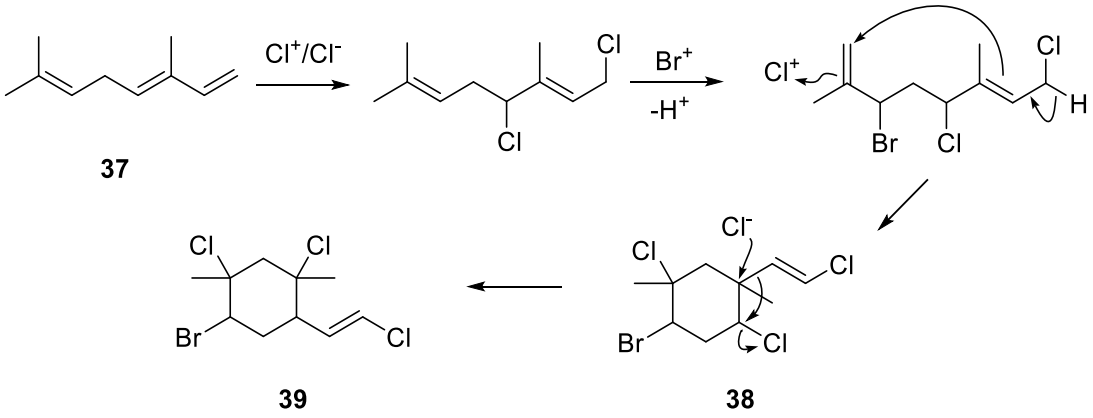


35



36

The monoterpenes of macroalgae are biosynthesised from geranyl pyrophosphate, generated from acetyl CoA via the mevalonate pathway.[125] Algae of the genus *Plocamium* have been shown to localise all enzymes involved in this pathway in an vascular organelle termed the mevalonosome.[141] Owing to the higher concentrations of chlorine and bromine in the ocean than terrestrial environments, they are incorporated into the biosynthesises through vanadium–haloperoxidases, generating halonium ions.[36] These reactive functionalities result in different metabolites formed than terrestrial counterparts, which rarely contain halogens.[142] The halonium ions, alongside those functionalities generated from oxidative processes such as epoxides, are very reactive and often initiate the next biosynthetic step, which includes cyclisation. *Plocamium* algae generate their metabolites from ocimene (**37**),[143] through pathways such as the typical example shown in Scheme 3.1, the only biosynthetic pathway from this genus to be formally chemically investigated so far.[144] This study utilised ^{14}C and ^{3}H labelled mevalonate to observe how the atoms were incorporated into the final products.

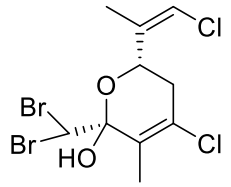


Scheme 3.1 – Biosynthesis of 2,4-dimethyl-1-vinylcyclohexane derivative **39**.

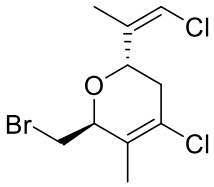
Here, both chloride and ocimene–based halonium ions generate a precursor that undergoes an internal cyclisation to produce 1,3-dimethyl-1-vinylcyclohexane (**38**). This can then undergo further substitution and 1,2-alkyl migration reactions to produce metabolites such as 2,4-dimethyl-1-vinylcyclohexane derivative **39**,[125] giving the alga access to a huge range of cyclisation and halogenation patterns. Costatones A and B (**40** and **41**), are representative of the variation and differentiation of the core terpene scaffold that this process can produce, with both being brominated and chlorinated, centred about a 3,6-dihydropyran ring formed after oxidation. Costatone A (**40**) was isolated from the hexane extract of an alga identified as *P. costatum* collected from Robe, South Australia.[145] The absolute configuration of compound **40** was solved by single–crystal X–ray diffraction (XRD) analysis, and it has subsequently been isolated from an Australian *P. angustum* alga collected from Rocky Point, Victoria,[146] as well as a NZ specimen collected in Doubtless Bay, Northland.[134] Costatone B (**41**) was isolated from a *P. costatum* sample collected from Pandalowie Bay, South Australia, and lacks the hydroxyl and one of the bromine substitutions of **40**.[147]

Isolation protocols for purification and biological activities of previously reported *Plocamium* halogenated monoterpenes have been thoroughly reviewed,[125,145,146] therefore the interested reader is directed to these sources. However, it is relevant to compare the compounds that have been previously reported from algae 'identified' as *P. angustum*. This species is endemic to Australia and NZ, and distributed throughout the Pacific Islands and along the full length of the NZ archipelago.[148] In 1979, Dunlop *et al.* reported the first chemical investigation of the alga from a sample collected from Cape Northumberland, South Australia, where the previously unreported acyclic monoterpene **42** was isolated alongside the previously reported dienone plocamenone (reported as **43**).[149] The structure of **42** was solved by analysis of its NMR and MS data, and comparison to a derivative generated by chemical methods. Compound

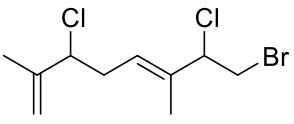
43 was initially isolated from an unidentified *Plocamium* species collected off New South Wales, Australia,[150] however the reported structure misplaced one chlorine atom and was revised following ^{13}C NMR chemical shift calculations.[151] It was not until Timmers *et al.* investigated a cytotoxic and antibacterial extract of *P. angustum* collected off Queenscliff, Victoria, isolating the major plocamenone (**44**) and unstable isomer isoplocamenone (**43**), that the *E* double bond geometry was realised.[152] Purified plocamenone (**44**) showed significant cytotoxicity against P388 murine leukaemia cells (IC_{50} 0.157 μg mL^{-1}) and also potent antimicrobial activity against *B. subtilis*, *C. albicans* and *Cladosporium resinae*, whereas **43** was not stable enough for accurate assay. Recent total syntheses of both **43** and **44** confirmed the proposed structures, and isomerisation of unstable **43** to **44** was observed when a purified sample was irradiated with UV light over several days.[153] The only other report of a study on *P. angustum* re-isolated four previously known compounds. Both **42** and **44** (reported as **43**) were re-isolated from a sample collected from Rocky Point, Victoria,[146] alongside costatone A (**40**) and highly halogenated trienone **45**, both originally isolated from *P. costatum* (now known as *P.cirrhosum*).[154,145] Therefore, even though progress has been made through chemical investigations of the alga, the only completely novel compound isolated is **42**. To date, there have been no reports of studies concerning alga identified as *P. angustum* from NZ waters.



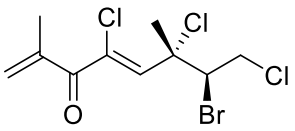
40



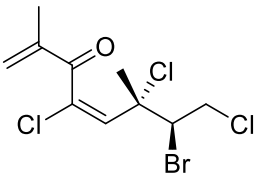
41



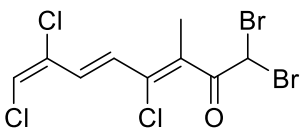
42



43



44



45

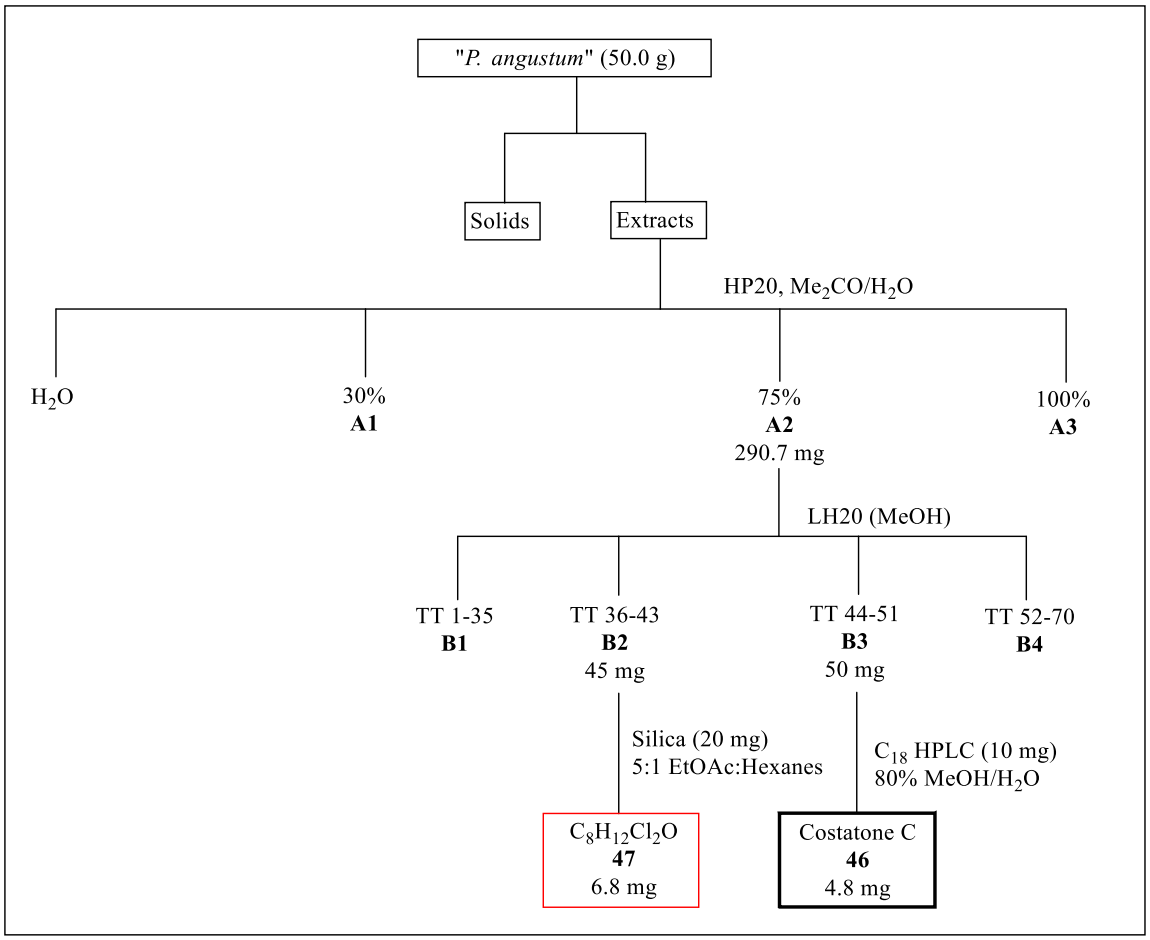
3.2 – Prioritisation and Isolation Procedure

A sample morphologically identified as "*P. angustum*," (Figure 3.1) collected from Moa Point in 2017, was the subject of this chemical investigation. At VUW, an undergraduate teaching laboratory experiment exploring the natural product diversity of local red algae required the annual collection of *P. costatum* (*P. cirrhosum*);[155] however, students seemingly working on the same species across several years isolated different compounds. The "*P. angustum*" sample was prioritised in the current research owing to questions put forward regarding the state of *Plocamium* taxonomy in NZ.[148] It has been recognised that morphological characteristics typically used for identification are not sufficient to distinguish the different species identified by molecular taxonomic methods.[156] European samples of species *P. cartilagineum,* in particular, have been the subject of intense molecular examination, where at least eight divergent species were described.[157] There are currently seven described *Plocamium* species in NZ,[158] however in the first taxonomic investigation to use molecular methods on these samples, at least eleven putative genetic species were identified.[148] As *Plocamium* species are known to be rich in chemical diversity, it is also of interest to inspect how the chemistry across these species varies.

Figure 3.1 – Surface photograph of alga identified as '*P. angustum'*.

In the NMR–spectroscopy guided screening section of Chapter 2, the ^{1}H NMR spectrum of the 75% Me_2CO/H_2O screening fraction of a "*P. angustum*" sample was presented (Figure 2.8). Consistent with previously reported compounds from this genus, the extract showed peaks attributable to halogenated monoterpenes, particularly methyl groups (1.0-2.5 ppm) and deshielded methines and methylenes (3.5-5.0 ppm). It was clear based on the integration of the

key resonances that two major compounds made up this sample. These were consequently separated by LH20 size-exclusion chromatography and individually purified by reversed-phase HPLC and silica–gel chromatography, to a yield new compound costatone C (**46**) and the known (1*E*,3*S*,5*Z*)-1,6-dichloro-2-methylhepta-1,5-dien-3-ol (**47**), respectively. Compound **46** is the first polyhalogenated tetrahydropyran reported from the 'species', sharing the carbon skeleton of **40** and **41** isolated from separate *P. costatum* (now *P.cirrhosum*) samples.[145,147] Compound **47** was previously reported from *P. cruciferum* collected near Kaikoura, NZ;[159,160] however, the absolute configuration of the secondary alcohol had not been established.



Scheme 3.2 – Isolation procedure for compounds from '*P. angustum*', collected at Moa Point, Wellington in 2017. Red boxes are previously reported compounds, black bold boxes denote compounds identified during the current study.

3.3 – Costatone C

Costatone C (**46**) was isolated as a colourless oil. High resolution atmospheric–pressure chemical ionisation mass spectrometry (HRAPCIMS) analysis detected a deprotonated molecule at *m/z* 376.8719, consistent with the molecular formula $C_{10}H_{13}OCl_2Br_2$ (calcd. 376.8716) and giving an index of hydrogen deficiency (IHD) of two. At VUW, ESI is the routine ionisation source used on the high-resolution MS (HRMS) instrument; however, with this ion source no signals were detected as the molecule seemingly does not ionise. This is an example of a situation where a routine MS screening protocol would not be applicable, presumably due to the higher lipophilicity of **46**. With ten carbon atoms, this molecular formula was indicative of a monoterpene structure. Sample **46** was soluble in $CDCl_3$, CD_3OD and DMSO-d_6, however for the purposes of structure elucidation, CD_3OD was used as it provided the best peak resolution (Appendix 5). As evidenced by the two deshielded resonances in the ^{13}C NMR spectrum at δ_C 138.5 and 117.5, one of the two degrees of unsaturation was accounted for by a C=C double bond and with no other evidence of unsaturation, this implied a monocyclic structure. The ^{13}C NMR spectrum, in conjunction with the multiplicity-edited HSQC, also showed signals corresponding to a non-protonated carbon (δ_C 73.6), three methines (δ_C 83.4, 73.8, 54.5), two methylenes (δ_C 39.1, 29.0) and two methyl groups (δ_C 28.9, 13.1).

The olefinic carbon C–1 (δ_C 117.5) was protonated (δ_H 6.28), with the resonance showing quintet splitting (J = 1.4 Hz) and COSY correlations to a vinylic methyl H_3–9 (δ_H 1.84) alongside allylic coupling to H–3 (δ_H 4.22). This spin system was extended by further COSY correlations from H–3 to H_2–4 (δ_H 2.45, 2.15) and finally to deshielded methine H–5 (δ_H 4.75). Correlations in the HMBC spectrum further supported this connectivity, including those of H_3–9 to C–1, C–2 (δ_C 138.5) and likely oxygenated C–3 (δ_C 73.8) (Figure 3.2).

A second spin system was also observed in the COSY spectrum, linking the oxygenated methine H–7 (δ_H 4.29) to methylene H_2–8 (δ_H 3.94, 3.73). The two spin systems were connected to carbon C–6 (δ_C 73.6), through both two and three-bond HMBC correlations from H_2–4, H–5, H–7 and H_2–8 (Figure 3.2). With such a large amount of semi-purified material isolated (50 mg), an acetylation was attempted on a subsample to detect which carbons were attached to free alcohols. Lack of reactivity with Ac_2O suggested **46** did not have a free hydroxyl group, and evidence for a tetrahydropyran moiety was deduced based on the HMBC correlation from H–3 to C–7, supporting the proposed monocyclic structure. The final methyl group H_3–10 (δ_C 28.9) showed HMBC correlations to C–5, C–6 and C–7 thus was attached directly to C–6.

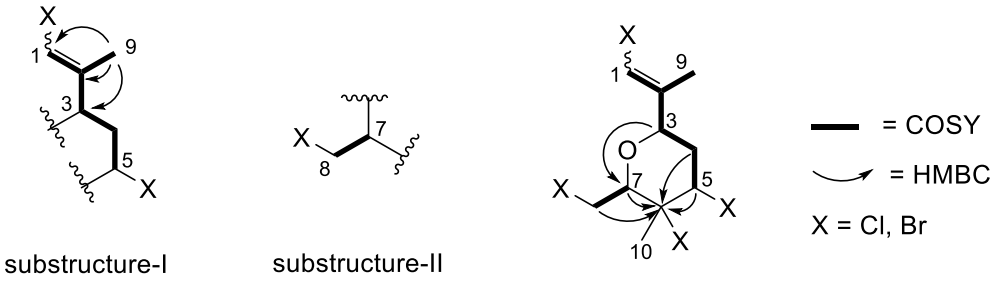


Figure 3.2 – Substructures alongside COSY and key HMBC correlations used to determine the planar structure of **46**.

With the carbon skeleton in hand, the positions of the two chlorine and two bromine atoms (Figure 3.2) were to be determined. As chlorine is more electronegative than bromine, in general the carbon to which it is attached is less shielded than when substituted with bromine. It can also be useful to compare chemical shifts of the carbons and geminal protons to those of previously reported compounds, particularly those of structures solved by X-ray crystallography. Another method for placing chlorine atoms is based on a phenomenon known as the "chlorine isotope effect".[161] Based on the slightly longer bond between the carbon and ^{35}Cl than ^{37}Cl, the carbon experiences less shielding and consequently the NMR resonance is split into an asymmetric doublet with the intensity ratio matching the natural isotopic abundance of chlorine (~3:1).[162] Figure 3.3 shows this characteristic splitting that was observed at δ_C 73.6, and therefore used to place the first chlorine atom at C–6. This effect is only observed for sp^3 carbons and is not observed for bromine as the difference in bond length between ^{79}Br and ^{81}Br is too small.

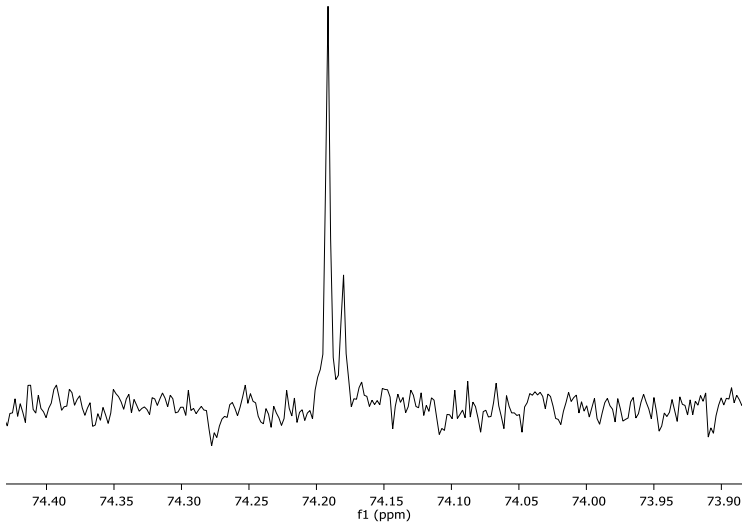
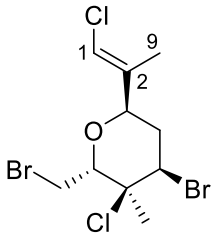


Figure 3.3 – Chlorine isotope effect splitting of the ^{13}C NMR resonance of C–6 of **46** (150 MHz, CD_3OD).

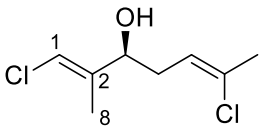
Chemical shift arguments were then used to assign the positions of the remaining halogens. The polyhalogenated monoterpenes costatols C-E (**48–50**) were reported from a previous chemical investigation of *P. costatum*,[147] and each contain a terminal trisubstituted *Z* double bond, with **48** and **49** chlorinated at C–1 and **50** brominated. The structure of compound **49** was solved by XRD analysis, therefore comparison of NMR spectroscopic data is reliable (Table 3.1). The ^{13}C NMR chemical shift of C–1 in compounds **47–49** with chlorovinyl groups is ~ δ_C 116, whereas it was δ_C 105.6 for the bromovinyl containing **50**. With this analysis, a chlorine atom was assigned at C–1, thus **46** contained a chlorovinyl terminal group. As no other ^{13}C NMR resonances showed the characteristic splitting, assigning the second chlorine to a sp^2 hybridised carbon is in agreement with the previous report that chlorine-induced ^{13}C splitting decreases with increased *s* character of the carbon atom.[161] The final two substituted positions at C–5 and C–8 were therefore brominated (δ_C 54.5 and 29.0 respectively). These shielded shifts are further evidenced the comparative shifts to C–5 of **48** (δ_C 64.6) and C–8 of **49** (δ_C 37.4).

Compound	C–1 shift	C–2 shift	C–9 shift
46	117.5	138.5	13.1
47 (Cl, *E*)	115.7	140.3	12.4*
48 (Cl, *Z*)	116.1	138.8	16.4
49 (Cl, *Z*)	116.2	138.7	16.1
50 (Br, *Z*)	105.6	141.7	17.5

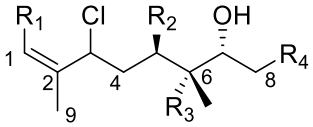
Table 3.1 – Comparison of NMR chemical shifts (**46** and **48–50** in CD_3OD, **47** in $CDCl_3$). *Chemical shift value for C–8.



46



47



48 R_1=Cl R_2=Br R_3=Br R_4=Cl
49 R_1=Cl R_2=Cl R_3=Br R_4=Br
50 R_1=Br R_2=Cl R_3=Cl R_4=Br

The nature of the four chiral centres of compound **46** and its alkene geometry were then deduced by through space NOE correlations and ^{1}H NMR data, particularly coupling constants.

The NOE correlations between H–1 and H–3 implied H–1 and H_3–9 are *trans*, thus the alkene possesses an *E* geometry, with supporting data coming from NMR chemical shift comparison (Tables 3.1 and 3.2). The C–9 methyl resonance of the co-isolated compound **47**, also containing an *E*-chlorovinyl group (δ_C 12.4, Table 3.1) is in much better agreement than **48** and **49**, where the *Z*-chlorovinyl was deduced by NOE correlations between H–1 and H_3–9 and single–crystal XRD.[147]

A *syn* relationship for H–3 and H–5 was proposed based upon their ROESY correlation to each other and also to the same vicinal proton H–4_b (δ_H 2.15). The coupling constants in the ^{1}H NMR spectrum for both interactions were relatively small ($J_{H-3/H-4b}$ = 2.6 Hz and $J_{H-4b/H-45}$ = 4.4 Hz), indicating axial-equatorial type relationships, whereas no detected through-space correlations and large coupling constants with H–4_a ($J_{H-3/H-4a}$ = 11.7 Hz and $J_{H-4a/H-5}$ = 12.0 Hz) suggested axial-axial *anti*-relationships. Shared correlations in the ROESY spectrum between H–5, H–8_a and H_3–10 also put these substituents on the same side of the ring, confirmed by a correlation from H–8_a across to H–3. Thus, the relative configuration was deduced as 3R^*, 5R^*, 6S^*, 7R^* as depicted below (Figure 3.4). The favoured conformation of six-membered pyranose rings minimises 1,3-diaxial interactions by orientating most of the bulky groups in equatorial positions, thus is consistent with the scalar ^{1}H-^{1}H coupling constants observed where all sterically demanding groups, with the exception of bromomethyl C–8, are oriented equatorially. The full NMR data of **46** is presented in Table 3.2.

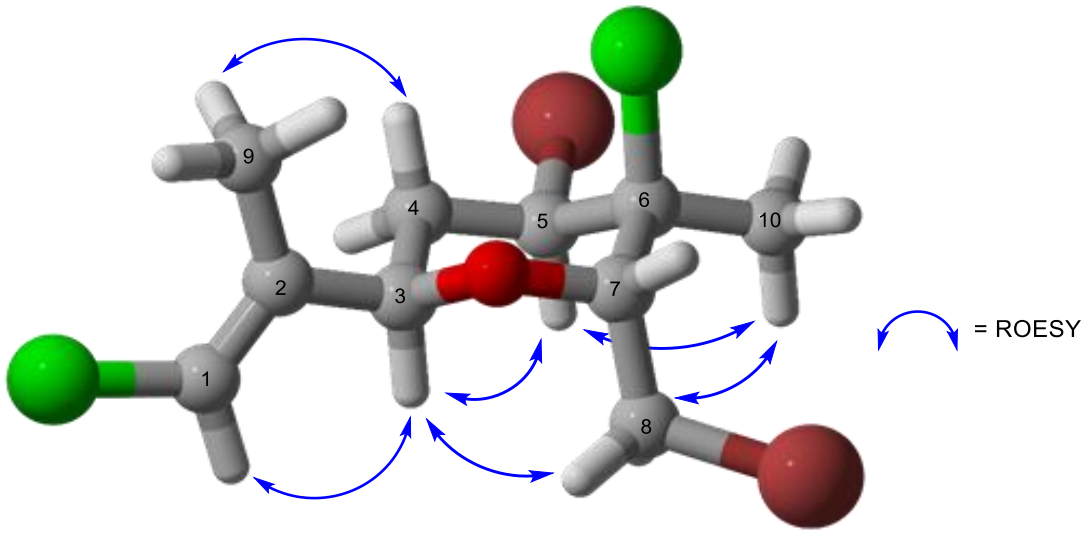
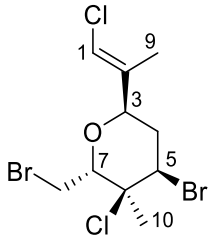


Figure 3.4 – Optimised geometry of **46** at PBE0-D3BJ/def2-TZVP/SMDMeOH level of theory. Key experimental ROESY correlations are shown by arrows.

Table 3.2 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for **46** in CD_3OD.

position	^{13}C (δ)	^{1}H (δ, mult., *J* in Hz)	COSY	HMBC	ROESY
1	117.5	6.28 (quin, 1.4)	3, 9	2, 3, 9	3
2	138.5				
3	73.8	4.22 (dd, 11.7, 2.6)	1, 4, 9	1, 2, 4, 5, 7, 9	1, 5, 8a
4a	39.1	2.45 (dt, 12.9, 11.9)	3, 4, 5	2, 3, 5, 6	9
4b		2.15 (ddd, 12.9, 4.4, 2.7)	3, 4, 5	3, 5, 6	9
5	54.5	4.75 (dd, 12.0, 4.4)	4	3, 4, 6, 10	3, 8a, 10
6	73.6				
7	83.4	4.29 (dd, 11.7, 3.9)	8	3, 6, 5, 8, 10	10
8a	29.0	3.94 (t, 11.6)	7, 8	6, 7	3, 5, 10
8b		3.73 (dd, 11.6, 3.9)	7, 8	7	10
9	13.1	1.84 (d, 1.4)	1, 3	1, 2, 3	4a, 4b
10	28.9	1.70 (s)		5, 6, 7	5, 7, 8a/b



46

As compound **46** was isolated as an oil, crystallographic data could not be obtained to deduce its absolute configuration, however this was assigned through comparison of experimental and computational electronic circular dichroism (ECD) data. Calculations were carried out by collaborators Dr Muhammad Ali Hashmi (University of Education, Attock, Pakistan) and Zaineb Sohail (University of Management and Technology, Lahore, Pakistan). After an optimised structure and conformational analysis was performed, time-dependent density functional theory (TDDFT) was used to compute a theoretical ECD spectrum which was then compared to that obtained experimentally in MeOH (Figure 3.5). After scaling for peak intensity and wavelength, there is clear agreement that the absolute configuration for **46** is 3*R*, 5*R*, 6*S*, 7*R*.

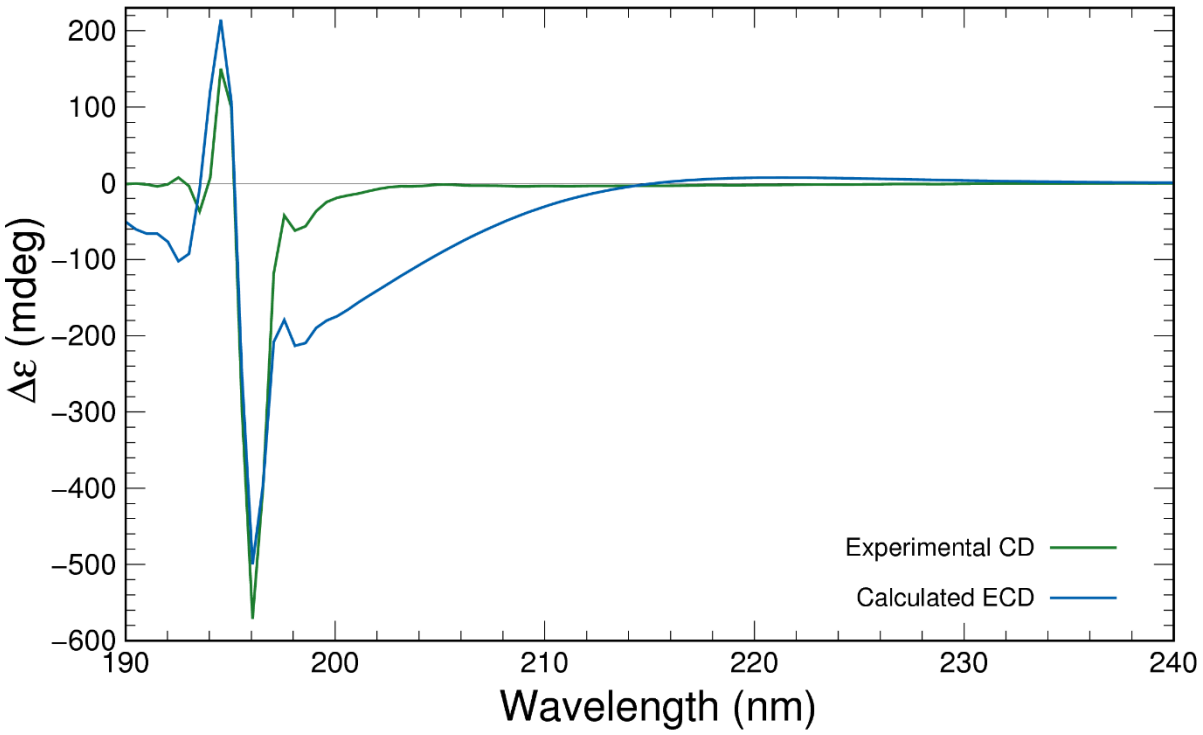
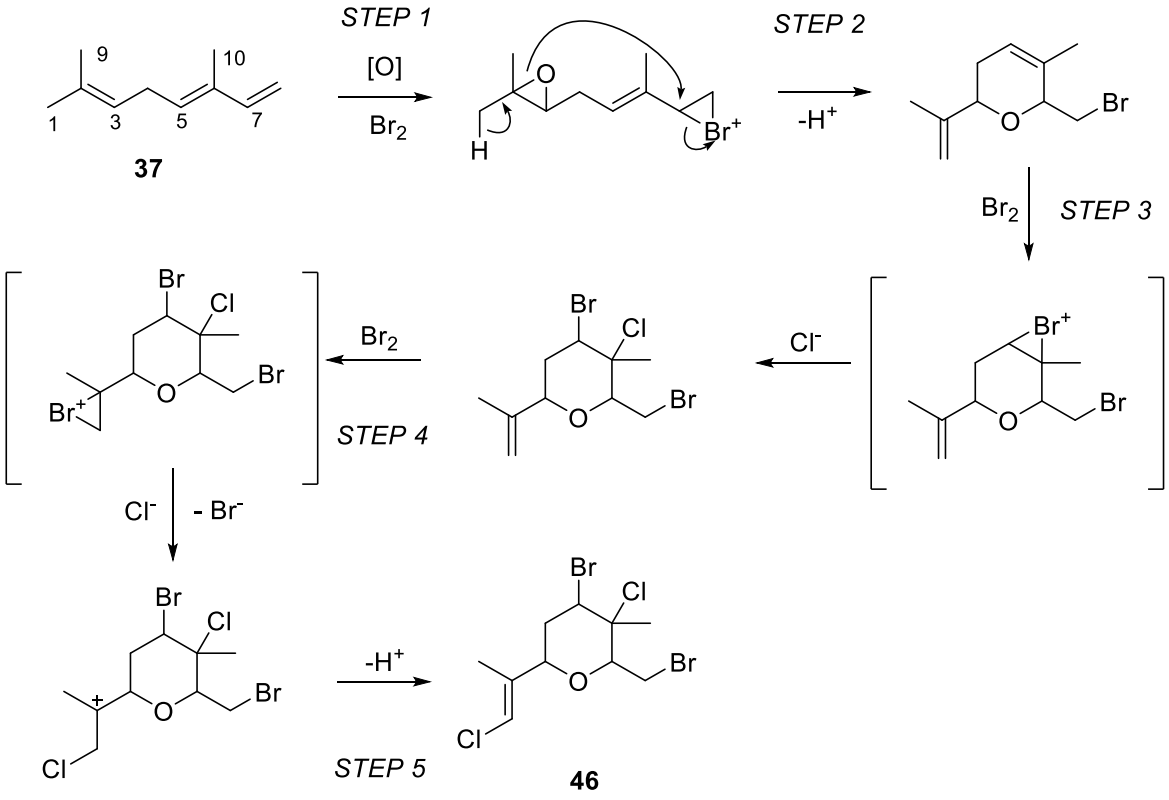


Figure 3.5 – Experimental ECD spectrum of **46** in MeOH compared with the spectrum of (3*R*, 5*R*, 6*S*, 7*R*)-**46** computed at the PBE0-D3BJ/def2-TZVP/SMDMeOH level of theory.

Although there are no methods to verify the biosynthesis of halogenated monoterpenes in the absence of isolating intermediates and/or analysis of the enzymes involved in the biosynthetic gene cluster, a general mechanism can be postulated from the predicted starting material ocimene (**37**). The olefinic double bonds of isoprenoid starting materials are often oxidised by epoxidase enzymes to form reactive epoxides.[163] These can then be the reactive site to catalyse a 6-exo-tet cyclisation, such as the second step in Scheme 3.3 where the epoxide oxygen attacks C–7 of the C–7 to C–8 bromonium ion, resulting in standard Markovnikov addition observed in general for products of bromoperoxidases.[164] A similar process is illustrated in step three, this time a chloride ion attacks the bromonium ion at the more substituted and electropositive C–6. The final vinylic chlorine is substituted at C–1, likely the result of a bromonium ion mediated chloride attack and subsequent elimination of HBr.[165] A more simpler route would be enzyme catalysed chloronium formation and deprotonation, however chloroperoxidases have never been found in marine algae,[165] and bromoperoxidases are only capable of oxidising bromine or iodine.[166]



Scheme 3.3 – A postulated biogenesis of **46**.

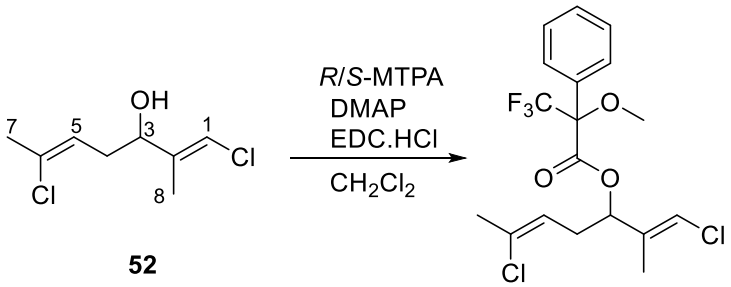
3.4 – (1*E*,3*S*,5*Z*)-1,6-dichloro-2-methylhepta-1,5-dien-3-ol

Compound **47** was isolated as a colourless oil. Its NMR ($CDCl_3$) and MS data were consistent with that previously reported for (1*E*,5*Z*)-1,6-dichloro-2-methylhepta-1,5-dien-3-ol, reported originally from *P. cartilagineum* collected near Kaikoura, NZ.[159,160] Ionisation by ESI again proved ineffective for MS, therefore APCI was used to detect the $[M–H_2O+H]^+$ cluster at *m/z* 176/178/180, a result of in-source dehydration with no ions corresponding to the parent ion detected. Electron impact (EI) ionisation analysis of this compound showed that the base peak was *m/z* 105/107,[159] and our sample showed identical EI fragmentation when using GC-EIMS.

The original isolation publication did not report any configurational analysis of C–3, therefore Mosher ester analysis was used to define the configuration of the secondary alcohol.[167] As acid chlorides of *R*- and *S*-methoxy(trifluoromethyl)phenylacetic acid (MTPA) were not available in our laboratory, the esters were produced under Steglich conditions with their respective carboxylic acids (Scheme 3.4). This also avoids the common issue of changing priority of the Mosher ester when using acid chlorides.[168] The resulting esters were semi-purified, and their ^{1}H NMR resonances assigned based on COSY correlations (Appendix 5). The resulting ^{1}H NMR chemical shifts are reported in Table 3.3.

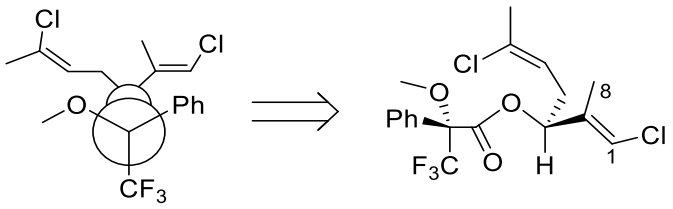
Table 3.3 – ^{1}H NMR chemical shifts of MTPA esters of **47** ($CDCl_3$, 600 MHz).

Position	*R*-MTPA ester	*S*-MTPA ester	$\Delta\delta_{S-R}$
H–1	6.16	6.25	+0.09
H_3–8	1.64	1.78	+0.14
H–3	5.48	5.50	+0.02
H–4a	2.60	2.56	–0.04
H–5	5.32	5.18	–0.14
H_3–7	2.09	2.03	–0.06



Scheme 3.4 – Conditions of MTPA esterification of **47**.

The Mosher method is used to assign the absolute configuration of a secondary alcohol, and relies on the shielding effects of the introduced aryl group on one half of the molecule when the ester sits in its favoured conformation with the CF_3 group eclipsing the proton (Figure 3.6). As illustrated, when the stereocentre has an '*S*' configuration, the phenyl ring of the *R*-MTPA ester (**47b**) shields the eastern side of the molecule as illustrated. This results in more electron density and shielding, thus lowering the chemical shifts, and results in positive $\Delta\delta_{S-R}$ values for H–1 (+0.09) and CH_3–8 (+0.14). As this is consistent with the observed data, the *S* configuration was established. As the observed specific rotation of $[\alpha]_D^{20}$ –22 (c 0.1, $CHCl_3$) for the current isolate agrees with that previously reported of ($[\alpha]_D^{20}$ –9.8 (c 0.1, $CHCl_3$)), this also established the absolute configuration of the *P. cruciferum* metabolite.[159,160] A comparison of experimental and literature NMR data is presented in Table 3.4.



47b **47b**

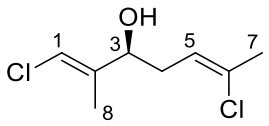
Figure 3.6 – *R*-MTPA ester of **47** (**47b**) conformation.

Table 3.4 – Comparison of isolated and reported NMR chemical shifts for **47**.

Position	Isolated [a] ($CDCl_3$, 600 MHz)		Reported [b] (CCl_4, 270 MHz)	
	δ_C	δ_H	δ_C	δ_H
C–1	115.7	6.15	115.2	6.02
C–2	140.3		139.9	
C–3	74.3	4.19	73.8	4.08
C–4	34.8	2.46	34.4	2.40
C–5	120.9	5.45	120.8	5.40
C–6	133.2		132.3	
C–7	26.5	2.12	26.1	2.13
C–8	12.7	1.80	12.30	1.78

[a] ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data in $CDCl_3$.

[b] ^{13}C (67.89 MHz) and ^{1}H (270 MHz) NMR data in CCl_4.[159]



47

3.5 – Biological Activity

Both isolated metabolites **46** and **47** were submitted to SBS, VUW, for antimicrobial testing against the Gram-negative *Pseudomonas aeruginosa*, and Gram-positive strains *Staphylococcus aureus* and *Staphylococcus epidermidis* (Figure 3.7). No activity was detected at 100 μg mL^{-1} for either compound against *P. aeruginosa*, however **46** showed mild activity against both Gram-positive strains, with minimum inhibitory concentrations (MIC) of 128 and 64 μg mL^{-1} for *S. aureus* and *S. epidermidis*, respectively.

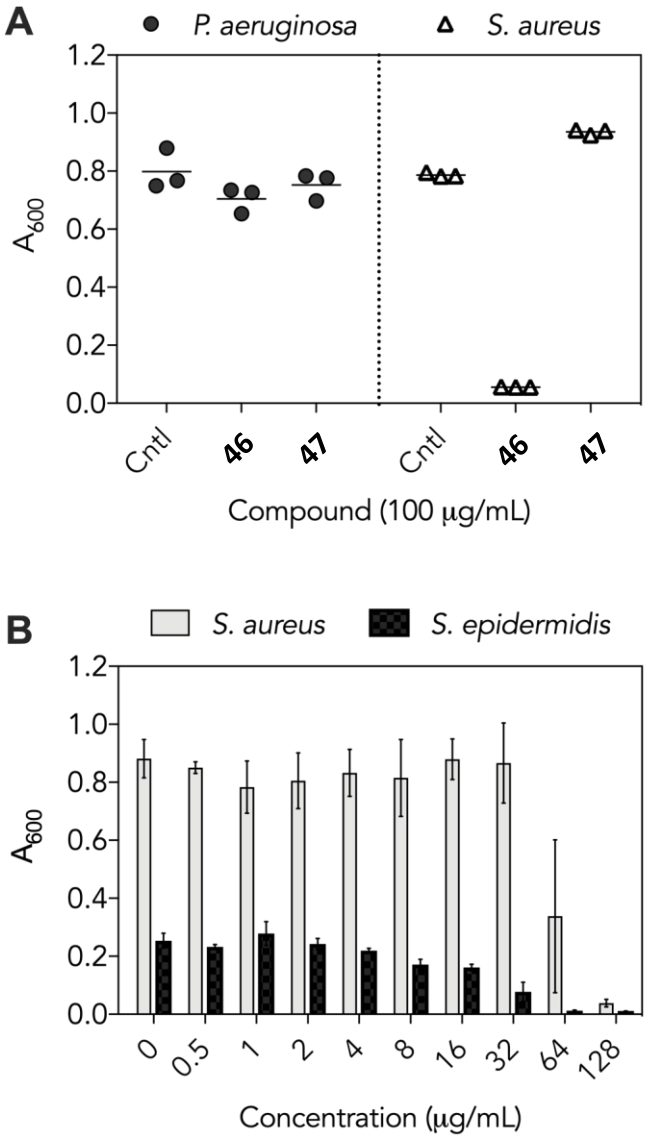


Figure 3.7 – Antibacterial testing of **46** and **47**. (A) Initial screen for growth inhibition of *P. aeruginosa* and *S. aureus* using a 1% DMSO control. (B) Determination of the minimum inhibitory concentration of **46** towards *S. aureus* (grey bars) and *S. epidermidis* (black bars).

3.6 – *Plocamium* taxonomy in NZ and Australia

It is evident that the taxonomy of *Plocamium* species around the world is currently very poorly defined in the scientific literature.[148,169] The routine and simple morphological methods previously used to determine species are clearly inadequate, proven both chemically and phylogenetically. Although chemical constituents are not an absolute descriptor of species as they can vary based on environment, season and other organisms present (e.g. symbiotic microorganisms, epibionts), the huge variation of halogenated monoterpenes observed in seemingly identical species validates the genomic data. There are reports in the literature describing changes in halogenated monoterpene composition for a given species that occurs with varied collection location and season;[170,171] however, it is possible that inaccurate species identification may also contribute to this. An example of this variation was reported in 1985, were two algal samples both identified as *P. cartilagineum* collected near Kaikoura, NZ, had completely different chemical constituents and were so different they were referred in text as 'subtypes.'[172] One of the samples had extremely high levels of dienol **47**, while the other did not contain it at all, and samples collected around Northern America and Europe also do not contain this compound.

Through phylogenetic work carried out as part of a Master of Biology project in SBS, VUW,[148] a range of *Plocamium* samples collected around NZ were phylogenetically analysed by comparison of their cytochrome oxidase subunit 1 gene sequence. The sequence of the algal sample identified as "*P. angustum*" used in this investigation was compared to those results, which suggested this species fell into cryptic species G (Figure 3.8).[148] This sample was collected near Moa Point, Wellington, with the other sample sequences restricted to the Wellington coast and Wairarapa region, indicating cryptic species G to be distinct from all other species from NZ and Australia previously identified as *P. angustum*. The alga used in this work was collected in 2017 and contained large quantities of **46**; however, investigations of the same morphologically identified species collected in 2018 showed no sign of the compound. This was initially attributed to a change in metabolic state of the alga, however with this genomic data, is likely due to a different genotype collected and consequently being analysed. In line with the chemistry detected, this species is most similar to sequences previously identified as *P. cartilagineum* in NZ, however, *P. cartilagineum* is a European species and is therefore unlikely to be the correct identification. Our results indicate the state of NZ *Plocamium* taxonomy is very complex, and there may be several species with additional interesting chemistries and unreported metabolites still to be investigated. In order to fully

understand this phenomenon, a parallel chemical and genomic screen of *Plocamium* species must be carried out in future work.

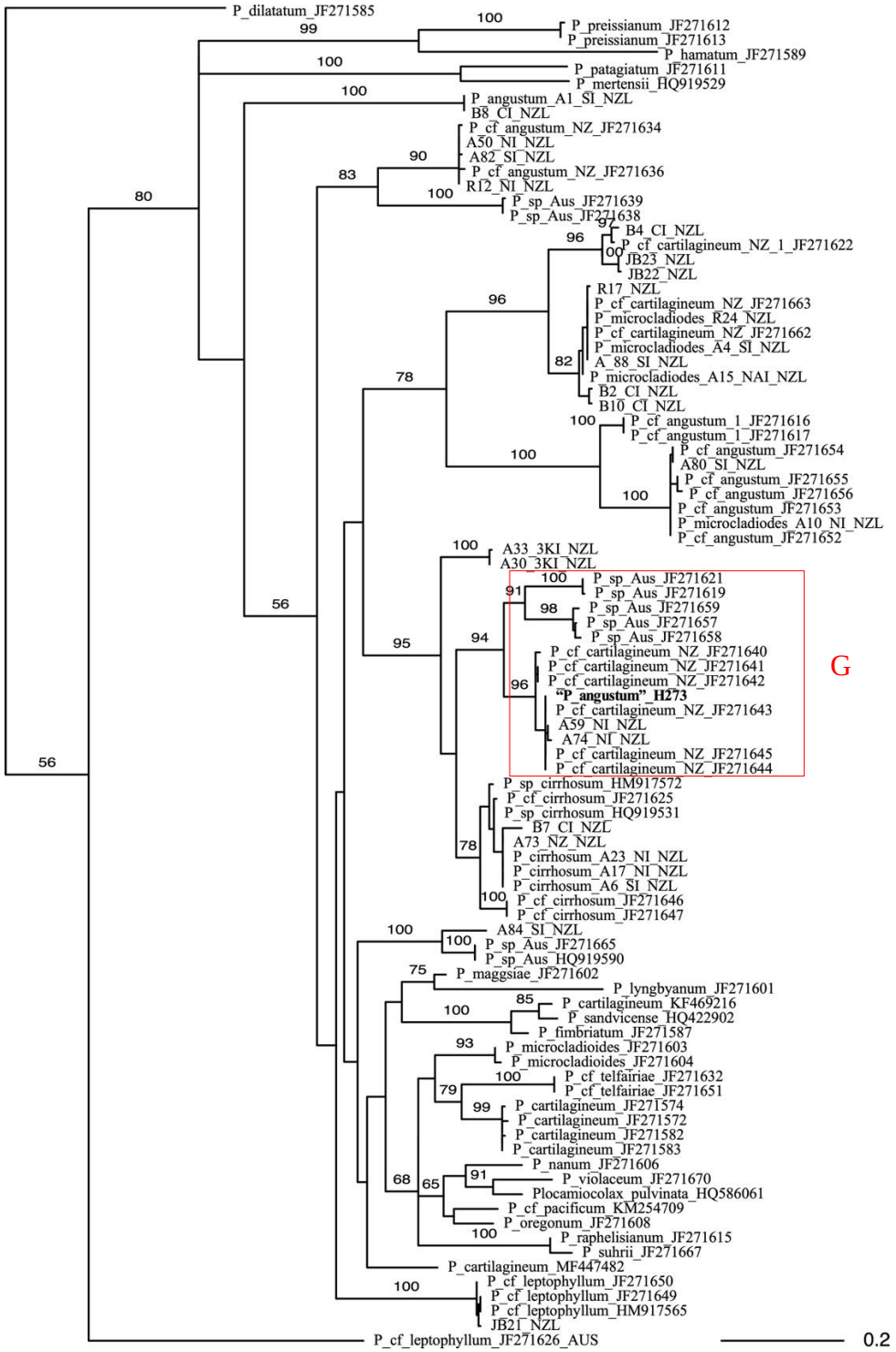


Figure 3.8 – Phylogenetic analysis (maximum–likelihood) of *Plocamium* species based on COI data.

3.7 – Conclusion

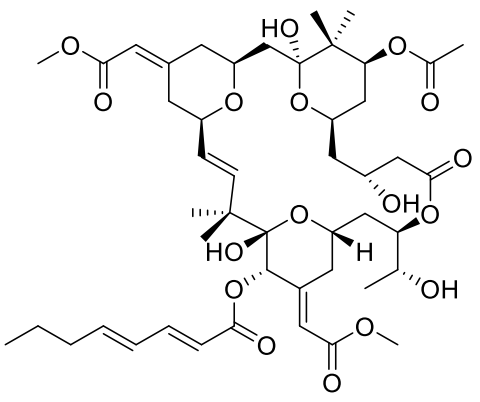
With the poorly defined state of *Plocamium* taxonomy in the scientific literature, there are potentially many different species that have not been chemically investigated to date. This is reflected in the alga identified as "*P. angustum*", which was the source of new compound costatone C (**46**). This metabolite, alongside the previously reported compound **47**, were isolated using a ^{1}H NMR spectroscopy-guided isolation procedure. Neither of these compounds could be ionised by ESI methods, and only **46** showed mild antibacterial activity in the bioassays. Therefore, both a general MS- or bioactivity-guided isolation may have missed this new chemical entity. This study proves the power of ^{1}H NMR-guided isolation for the discovery of new natural products.

Chapter 4 – *Nelliella nelliiformis*

This chapter describes the chemical investigation of the Tongan bryozoan Nelliella nelliiformis. *This organism was prioritised for investigation based on its unique taxonomy within our collection, along with many interesting signals in the screening fraction ^{1}H NMR spectra. This resulted in the isolation of two new nucleosides termed nelliellosides A and B (**61** and **62**), and inspired the synthesis of a further four congeners (**66**-**69**). The bioactivity of these nucleosides was evaluated against a full panel of human disease-relevant kinases, where significant inhibition was observed against three kinase families.*

4.1 – Introduction

Bryozoans are a diverse phylum of filter feeding invertebrates colloquially known as "moss animals". Most lack physical protection from predation and therefore have acquired anti-predatory secondary metabolites with a range of chemical functionalities and bioactivities.[28] The most well-known examples are the bryostatins, first reported in 1982 from a Californian-sourced specimen of *Bugula neritina*.[173] Since the initial report of bryostatin 1 (**51**), 21 different congeners have been reported.[174] Bryostatin 1 is of immense interest owing to its wide range of potent biological applications. These include *in vitro* and *in vivo* anticancer activity for which it has undergone >80 clinical trials,[175,176] modulation of latent HIV–1 infections to increase the efficacy of adjunctive therapies,[177] and has also entered phase II clinical trials for the treatment of Alzheimer's disease.[178] In order to fuel potential drug development, large-scale extractions of *B. neritina* were performed; however, only 18 g of **51** was isolated from 14 tonnes of bryozoan, clearly limiting its utility from natural sources.[179] These compounds have attracted much interest from the synthetic chemistry community, and synthetic procedures have been thoroughly reviewed.[180] Research into bryostatin analogues and improved synthetic routes is ongoing some 40 years after its initial report.

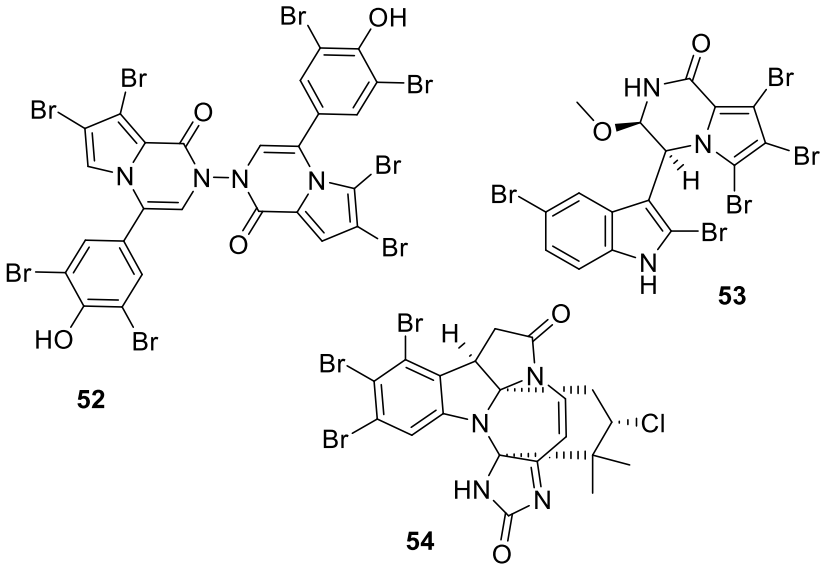


51

Despite the success of the bryostatins, the Bryozoa remain understudied relative to most other investigated marine phyla, with just below 300 reported compounds in MarinLit to date.[9] This is despite a recent study comparing the chemical space occupied by MNPs and approved pharmaceuticals, proving bryozoan-derived secondary metabolites to have one of the most similar profiles to clinically-used drugs of all the phyla to have been investigated.[8] Still, no other bryozoan compounds are currently in clinical trials or the development pipeline, and therefore presents a gaping hole in marine pharmacognosy.

Aside from the polyketide-derived macrocyclic bryostatins, the majority of reported bryozoan metabolites are nitrogenous,[181] including nucleosides previously isolated from terrestrial sources.[182,183] An investigation of a South China Sea *B. neritina* sample led to the isolation of adenosine,[183] while a sample of *Cellaria* spp. contained inosine and the 2'-deoxy derivatives of adenosine and guanosine.[182]

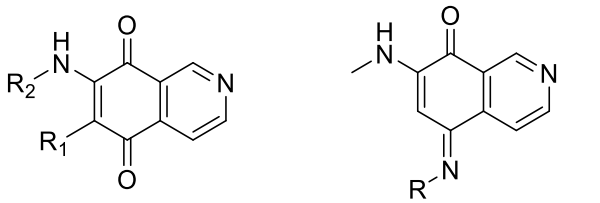
One particularly fruitful order of Bryozoa, to which *B. neritina* belongs, is the Cheilostomatida. According to the MarinLit database, 58 new natural products have been reported from the order.[9] Common motifs include highly condensed aromatics with pyrrole cores, which are often brominated. These are best exemplified by the aspidostomides and aspidazide A, isolated from a Patagonian *Aspidostoma giganteum* sample.[184] These compounds are based on a core of bromotyrosine or bromotryptophan, and share a common bromopyrrole carboxylate moiety within their core. Due to the proton deficient nature of these molecules, with many protons that could be used as spectroscopic "handles" replaced by bromine, structure elucidation can be difficult. Aspidazide A (**52**) is the first marine invertebrate natural product containing a diacylazide moiety, and aspidostomide E (**53**) was the only congener to show cytotoxic activity, with an IC_{50} value of 7.8 μM against the 786-O human renal carcinoma cell line. Similar chemical diversity was observed from the Artic bryozoan *Securiflustra securifrons*, the source of the securidene, securamine and securine alkaloids.[185] One particular extract, containing securamines H–J and securamines C and E, showed broad cytotoxicity against a range of human cancer cell lines. Purified compounds were assayed, and securamine H (**54**) showed the most potent activity, with IC_{50} values of ~ 3 μM against three cancer cell lines and non-malignant MRC-5 lung fibroblast cells.[186] Therefore, organisms of the order Cheilostomatida are a proven source of chemical novelty and diversity, and all currently unexplored genera warrant chemical investigation.



52 **53** **54**

4.2 – Isolation

A VUW MNP expedition to the Kingdom of Tonga in 2016, focused on the collection of sponges, with some tunicates, soft corals, cnidaria and one bryozoan also collected. The lone bryozoan, *N. nelliiformis* (Figure 4.1), was screened as it is from an alternative phylum to the sponges. Additionally, searching through databases such as MarinLit and Scifinder suggested there have been no previous reports of compounds isolated from a Tongan bryozoan, or any others collected in the tropical South Pacific Islands, nor from the genus. There are many reports of bryozoan-derived compounds from NZ and Australia, but the closest to the tropical waters of Tonga was an investigation of the Palauan *Caulibugula intermis*, of the same class as *N. nelliiformis* (Gymnolaemata), which resulted in the isolation of six new alkaloids, caulibugulones A-F (**55–60**).[187] Therefore, with the potential for interesting new chemistry from an understudied taxa, this organism was prioritised for extraction and NMR-based screening.



55 R_1=H R_2=CH_3
56 R_1=Br R_2=CH_3
57 R_1=Cl R_2=CH_3
58 R_1=H R_2=CH_2CH_2OH

59 R_1=H
60 R_1=CH_2CH_2OH

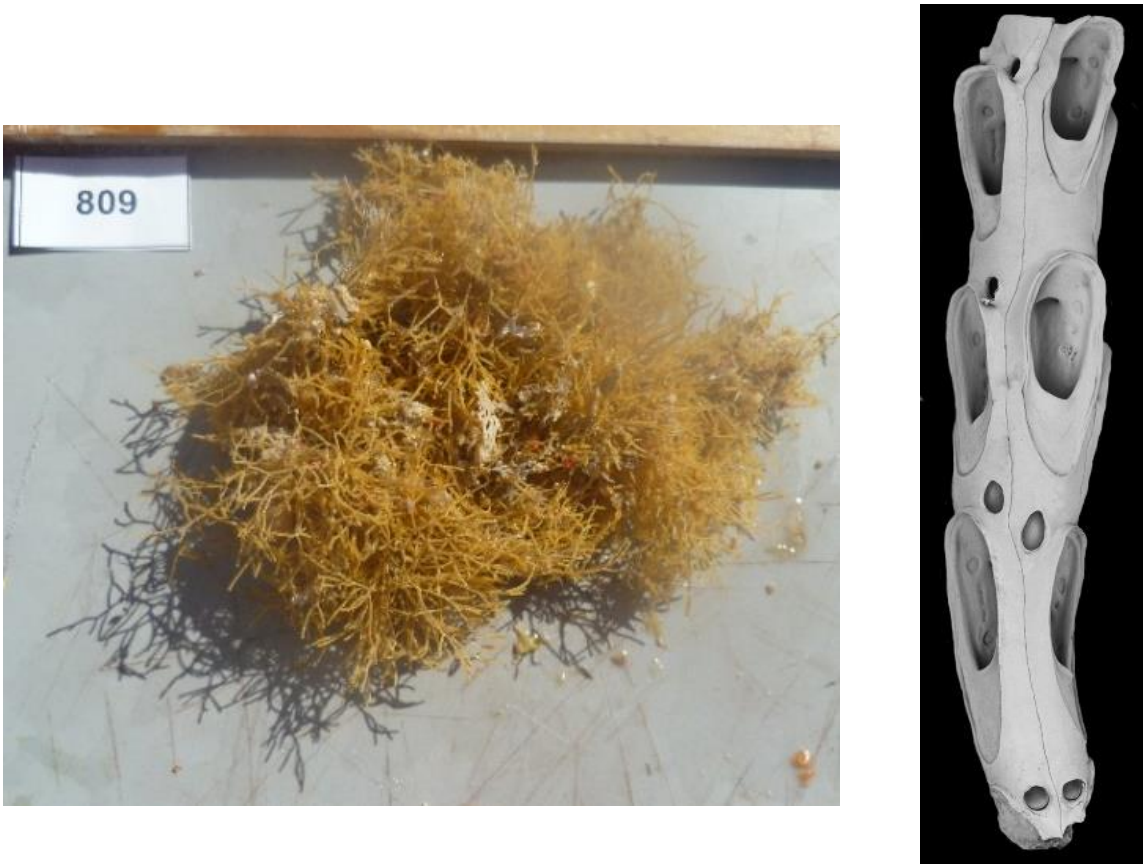
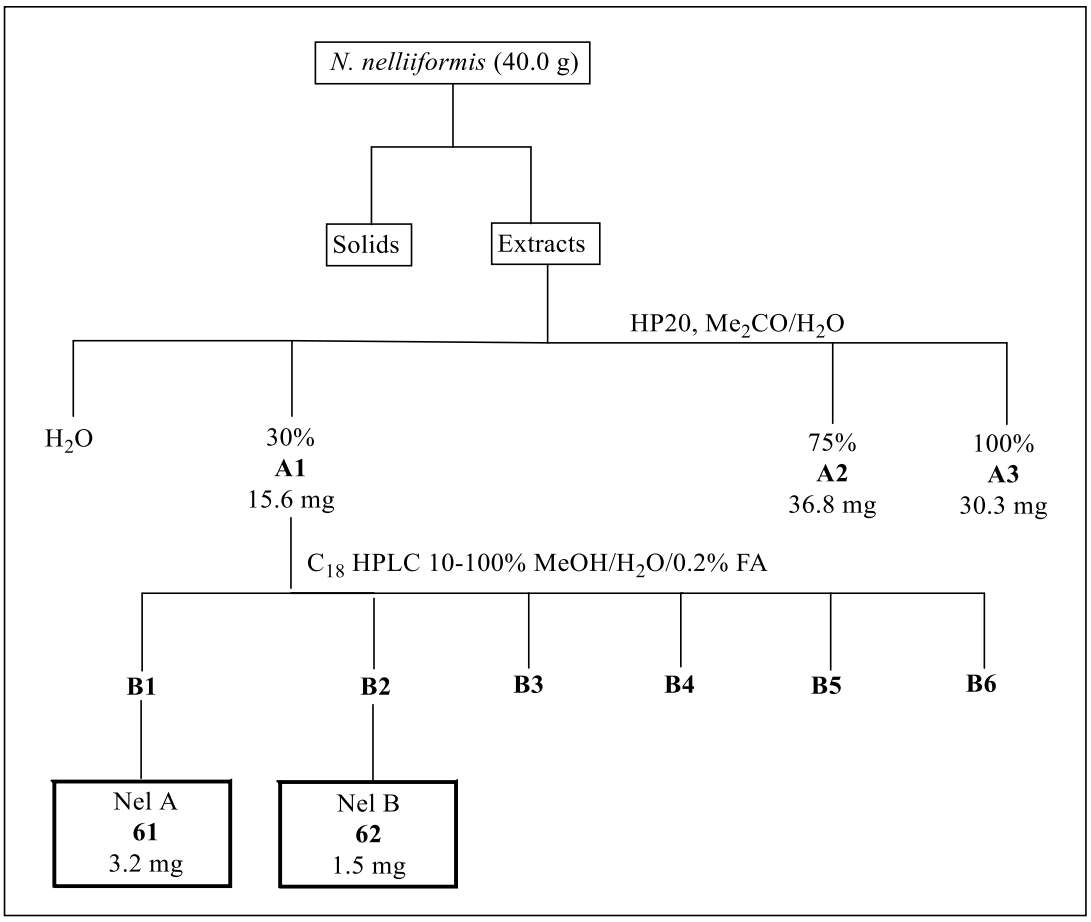


Figure 4.1 – Left: Surface photograph of the bryozoan *N. nelliiformis*; Right: Scanning electron microscopy illustration (SEM) of an internode of *N. nelliiformis*. The stem is 425 μm in diameter mid-length (SEM image reused with permission of Dr. D. Gordon, NIWA).

N. nelliiformis (40.0 g wet weight) was collected by hand using SCUBA, off ʻEua, Kingdom of Tonga, and extracted twice overnight in methanol. The extract was partitioned over HP20 resin using increasing percentages of Me_2CO in H_2O, with the ^1H NMR spectrum of the 30% Me_2CO in H_2O fraction showing many resonances of interest, particularly in the aromatic region. Further purification of the extract by reversed-phase HPLC afforded compounds **61** and **62** as the major components (Scheme 4.1).



Scheme 4.1 – Isolation procedure for compounds from *N. nelliiformis*, collected in ʻEua, Kingdom of Tonga. Black bold boxes denote compounds identified during the current study.

4.3 – Nellielloside A

Nellielloside A (**61**) was isolated as a white solid, with HRESIMS analysis showing a protonated molecular ion at *m/z* 361.1198 $[M + H]^+$, consistent with the molecular formula $C_{15}H_{17}N_6O_5$ (calcd. 361.1255) and an IHD of 11. The infrared (IR) spectrum showed evidence of an ester carbonyl (1674 cm^{-1}), while the UV/vis spectrum suggested the presence of an aromatic compound (Appendix 6). The ^{13}C NMR spectrum showed all 15 resonances, 10 of which were above 100 ppm (one ester at δ_C 160.1) and five between 60 and 90 ppm, suggesting oxygenation. The ^1H NMR spectrum acquired in DMSO-d_6 showed five resonances in the aromatic region, six resonances in the oxygenated methine/methylene region and two exchangeable signals (δ_H 11.94 and 7.28).

From these NMR spectra, alongside the multiplicity-edited HSQC experiment, the presence of four deshielded methines (δ_C 87.6, 81.8, 73.1 and 70.3) and an oxygenated methylene (δ_C 63.4) were observed, attributable to a furanoside. This was supported by COSY correlations around the ring, from the anomeric centre H–1' (δ_H 5.92) to the three contiguous secondary oxygenated methines (δ_H 4.66, 4.32 and 4.18) and terminating at oxygenated methylene H_2–5' (δ_H 4.36 and 4.50). The HMBC correlation between H–4' (δ_H 4.18) and C–1' (δ_C 87.6) confirmed the cyclic nature of the furanoside (Figure 4.2), while the other HMBC correlations were consistent with the COSY data.

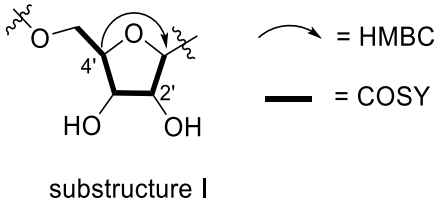


Figure 4.2 – Key HMBC and COSY correlations supporting substructure I of **61**.

Two aromatic fragments were also observed, one containing three non–protonated carbons (δ_C 156.1, 149.4 and 119.1) and two methine carbons (δ_C 152.7 and 139.4), while the other had one non-protonated carbon (δ_C 121.5) and three methine carbons (δ_C 124.5, 115.4 and 109.7).

Regarding the former fragment, an aromatic methine H–2 (δ_H 8.13) correlated to two deshielded carbons (δ_C 149.4 and 156.1), while the other H–8 (δ_H 8.25) correlated to one of these (δ_C 149.4) and to another more shielded resonance (δ_C 119.1). As neither ^1H resonance showed a HMBC correlation to the other's carbon, the two methines likely had more than three bonds between them. The exchangeable NH_2–6 resonance (δ_H 7.28) also gave a strong signal in DMSO-d_6 and showed a $^3J_{CH}$ HMBC correlation to the signal at δ_C 119.1. These data were

compared with the literature to assign an adenine moiety as the substructure II (Figure 4.3). Thus, C–4 was assigned to the signal at δ_C 149.4 as this carbon is three bonds from methine protons, while C–5 and C–6 were assigned to δ_C 119.1 and 156.1 respectively, the higher chemical shift of C–6 justified by the extra electron withdrawal of the two nitrogen atoms. The presence of adenine was corroborated by the (+)–MS/MS fragment ion at *m/z* 136.0622, which is caused by the cleavage of the purine base from the protonated parent ion.

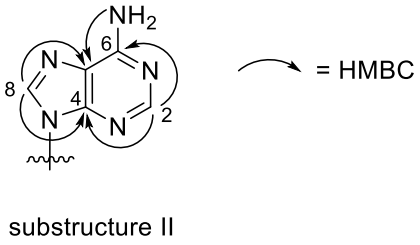


Figure 4.3 – Key HMBC correlations defining substructure II of **61**.

The third spin system showed COSY correlations between the three protons (δ_H 6.17, 6.81 and 7.03) and a broad downfield NH signal (δ_H 11.94). Aside from the correlations to the other protonated carbons as defined by the HSQC experiment, all three resonances showed HMBC correlations to only one other carbon C–2'' (δ_C 121.5), suggestive of a 2-substituted pyrrole moiety. There were, however, no correlations to other atoms outside of this ring system (Figure 4.4).

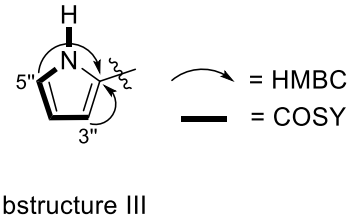
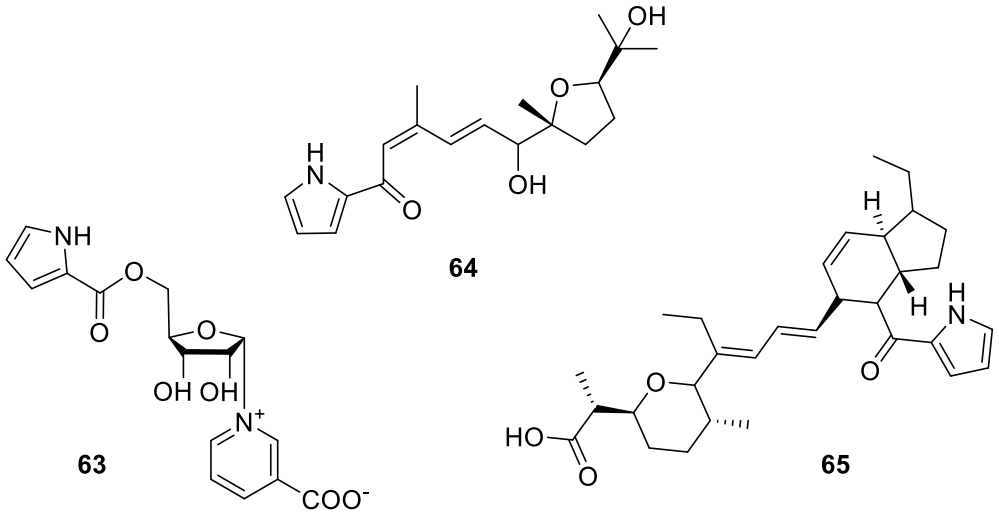


Figure 4.4 – Key COSY and HMBC correlations defining substructure III of **61**.

The anomeric proton H–1' of the furanose moiety (substructure I, Figure 4.2) showed $^3J_{CH}$ HMBC correlations to C–4 and C–8 of adenine, thus connecting substructures I and II. The final unassigned resonance in the ^{13}C NMR spectrum was at δ_C 160.1, suggestive of a *α,β*-unsaturated ester carbonyl. Both protons of the furanose methylene H_2–5' showed HMBC correlations to it, therefore the ester must be attached to the sugar.

The MS/MS spectra for **61** showed a fragment ion at *m/z* 94.0300, attributable to a pyrrole–carboxylate acylium ion (Appendix 6). This suggested that the ester carbonyl was bound directly to substructure III and this indicated a pyrrole–carboxylic acid residue esterified at C–

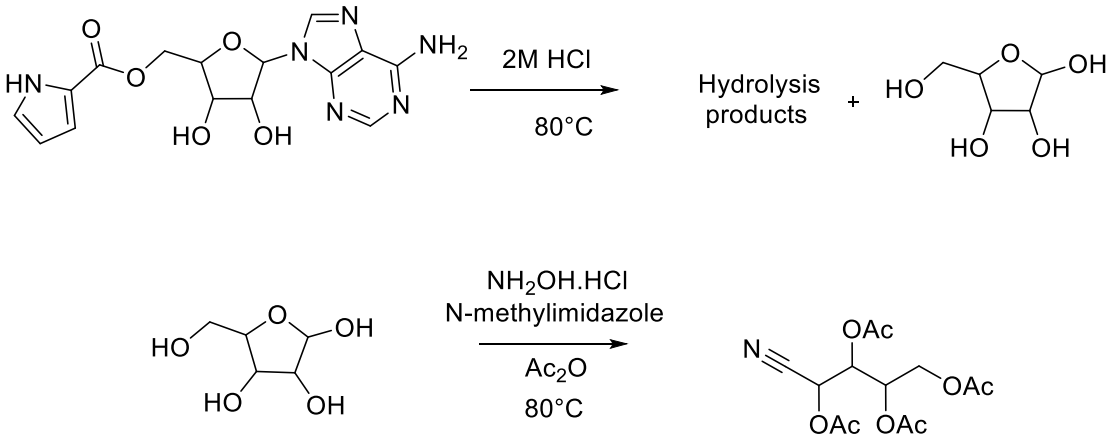
5'. The pyrrole ring was initially speculated to be 2-substituted as opposed to 3-substitutes based on the non-protonated carbon chemical shift being one of the two less shielded, due to the inductive effects of the pyrrole nitrogen causing C-2 and C-5 to resonate at a higher frequency. This was indeed confirmed through the synthesis of both the position-2 and -3 substituted acids (see below). Intriguingly, no HMBC correlations were detected to the carbonyl carbon from any of the protons on the pyrrole ring. The absence of such correlations is noted in other pyrrole–2–carboxyl containing natural products, such as neopetroside B (**63**),[188] the glaciapyrroles (**64**)[189] and the indanomycin-related antibiotics (**65**).[190] Therefore, the planar structure of **61** was deduced.



4.3.1 – Configurational analysis using GCMS

Although the strong NOE (ROESY) correlation between H–1′ and H–4′ suggests a *β*-anomeric configuration, the metabolite **61** has four stereocentres, and their relative configurations could not all be unequivocally deduced using coupling constants and ROESY data. In natural products, D–ribose is the most common furanose sugar, however, this could not be assumed immediately. There are many MNPs with other furanose configurations, including the inaugural FDA approved MNP leads Ara-A and Ara-C (**7** and **8**) that both have D-arabinose sugars. A method for the analysis of a range of carbohydrates, proposed by McGinnis,[191] utilises a conversion for each sugar to its corresponding aldononitrile peracetate, followed by gas chromatography-mass spectrometry (GCMS) analysis. Aldononitrile peracetates are stable, easily synthesised, and volatile enough to be accurately resolved when analysed by GCMS, where comparison of retention time to those of derivatised authentic standards can determine the configuration of the compound. Better yet, the use of chiral stationary phases allows the

absolute configuration of the sugar to be deduced as well. Therefore, a portion of isolated **61** was hydrolysed by heating in 2 M HCl, to produce a mixture of products, including the free furanose. The peracetylated aldononitrile was then synthesised, by treating the mixture with hydroxylamine hydrochloride ($NH_2OH.HCl$) in *N*–methylimidazole and acetic anhydride (Scheme 4.2), and following workup, was analysed using a previously established GCMS method on an achiral column.[192,193]



Scheme 4.2 – Peracetylated aldononitrile synthesis from natural sample of **61**.

The hydrolysed product had a retention time of 15.74 min while authentic samples of the four possible carbohydrate diastereomers were synthesised from commercially available ribose, arabinose, lyxose and xylose, which gave retention times of 15.76, 15.99, 15.88 and 16.15 min, respectively (Figure 4.5). Therefore, the retention time of the hydrolysed carbohydrate of **61** was in very good agreement with that of ribose.

To determine the absolute configuration of the compound, peracetylated aldononitrile derivatives of D- and L-ribose were again synthesised from commercially available standards, and analysed using the previously reported method,[192,193] with a CP-ChiraSil-L-Val capillary chiral column. The retention times of the peracetylated aldononitrile derivatives of D-ribose and L–ribose was 24.91 and 24.80 min, respectively, with the natural sample matching D-ribose with a retention time of 24.92 min (Figure 4.5). This matches the common configuration for furanose sugars in the majority of natural products. As this molecule had not previously been reported prior to this study, it was termed nellielloside A, based on the genus name "*Nelliella*" and the nucleoside structure of the molecule. The structure of this compound was confirmed by total synthesis (see below).

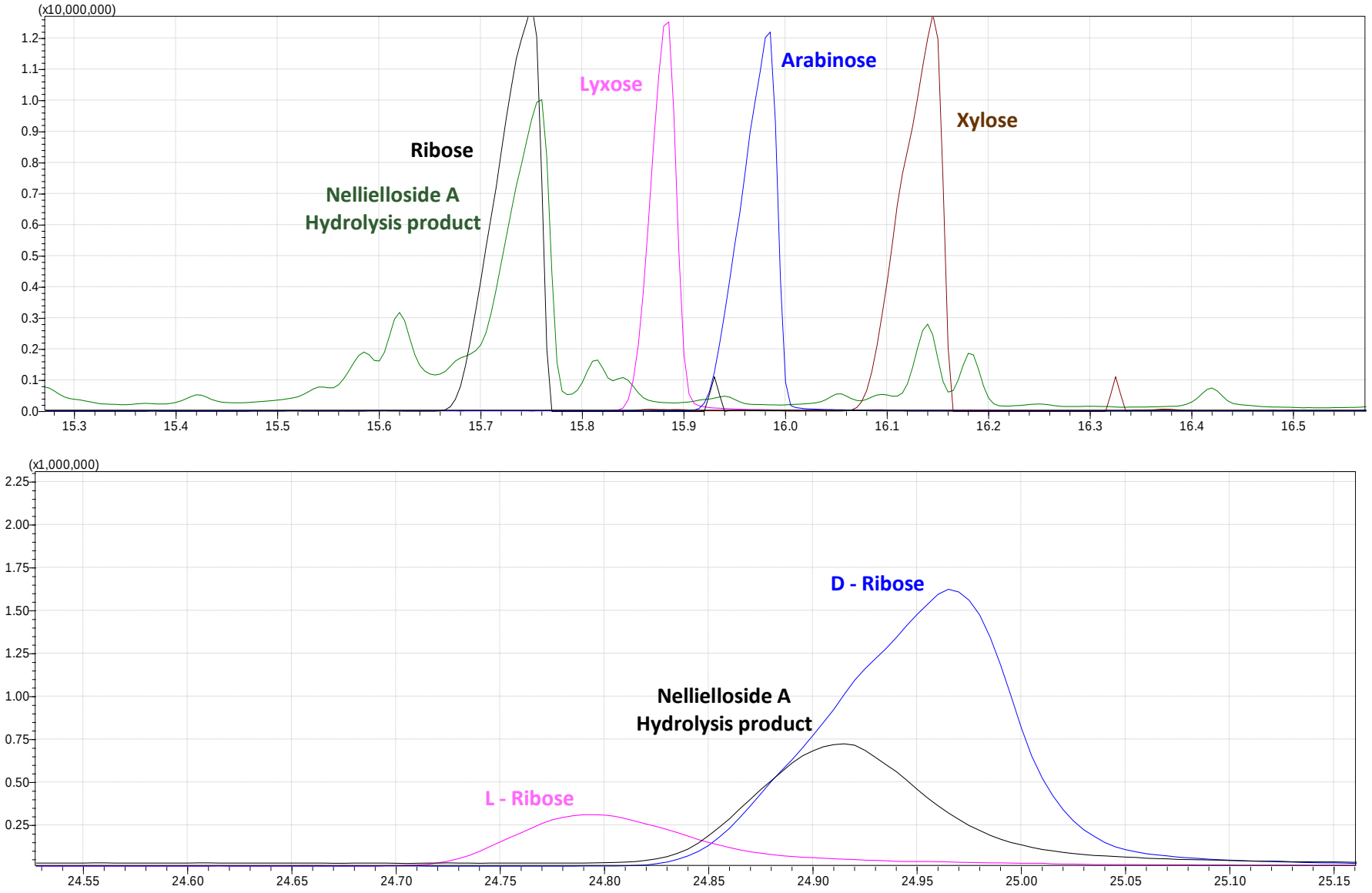
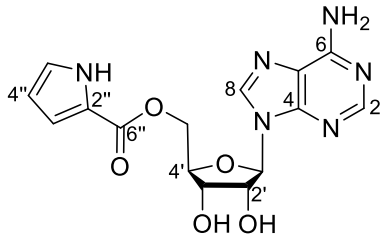


Figure 4.5 – GCMS retention time comparisons of furanose sugars. Top: using an achiral column; bottom: using a chiral column.

Table 4.1 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for Nellielloside A (**61**) in DMSO-d_6.

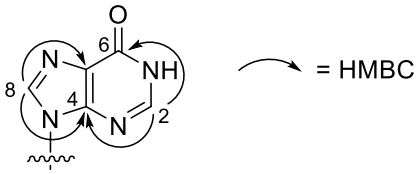
position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC
2	152.7	8.13 (s)		4, 6
4	149.4			
5	119.1			
6	156.1			
NH_2–6		7.28 (s)		5
8	139.4	8.25 (s)		4, 5
1’	87.6	5.92 (d, 5.2)	2’	4, 8, 2’, 3’
2’	73.1	4.66 (t, 5.2)	1’, 3’	3’
3’	70.3	4.32 (t, 4.9)	2’, 4’	1’, 4’
4’	81.8	4.18 (m)	3’, 5’	1’, 5’
5’	63.4	4.50 (dd, 3.7, 12.0)	4’	3’, 6’’
		4.36 (dd, 5.7, 12.0)	4’	3’, 4’, 6’’
NH–1’’		11.94 (s)	3’’, 4’’, 5’’	
2’’	121.5			
3’’	115.4	6.81 (m)	NH–1’’, 4’’, 5’’	
4’’	109.7	6.17 (m)	NH–1", 3", 5"	
5’’	124.5	7.03 (m)	NH–1", 3", 4"	2’’, 3’’, 4’’
6’’	160.1			



61

4.4 – Nellielloside B

Compound **62** was isolated as a white solid, with HRESIMS analysis detecting a protonated molecule at *m/z* 362.1096, consistent with the molecular formula $C_{15}H_{16}N_5O_6$ (calcd. 362.1095), differing from **61** by replacement of an NH equivalent with an oxygen atom. The IR, UV/vis and NMR spectral data were very similar to **61**, particularly for substructures I and III (Figures 4.2 and 4.4). The major differences were shifts in the NMR resonances associated with the adenine base, now with two aromatic methines resonating at δ_H 7.92 and 7.99 compared to δ_H 8.13 and 8.25 for **61** (Appendix 6). The less shielded proton signal showed HMBC correlations to C–4 and C–5 (δ_C 148.8 and 124.5 respectively), while the anomeric resonance of H–1' (δ_H 5.85) also correlated with C–4, thus assigning δ_H 7.99 as H–8. The other proton signal showed HMBC correlations to C–4 and C–6 (δ_C 148.8 and 162.5). C–6 is notably less shielded (~6 ppm) in **62** than **61**, the result of bearing an oxygen atom as opposed to the primary amine in adenine, thus suggesting the purine base was hypoxanthine and the nucleoside core inosine (Figure 4.6). This was confirmed from the MS/MS fragment ion at *m/z* 137.0431 (Appendix 6), caused by fragmentation of the hypoxanthine base from the protonated parent ion.



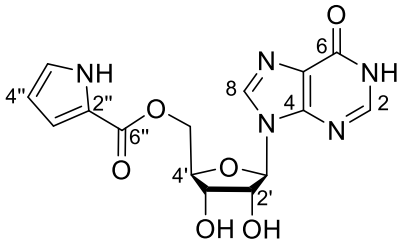
substructure-II

Figure 4.6 – Key HMBC correlations defining substructure II of **62**.

As only 1.5 mg of **62** was isolated, it was not hydrolysed and analysed by GCMS as above. The absolute configuration was proposed to be the same as **61** however, as both molecules have comparable specific rotations ($[\alpha]_D^{20}$ –24 vs –21 (*c* 0.1, MeOH). This was confirmed by total synthesis (see below). As the second molecule in the class, **62** was termed nellielloside B.

Table 4.2 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for Nellielloside B (**62**) in DMSO-d_6.

position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC
2	150.5	7.92 (s)		4, 6
4	148.8			
5	124.5			
6	162.5			
8	136.8	7.99 (s)		4, 5
1’	87.6	5.85 (d, 5.1)	2’	4, 8, 2’
2’	73.2	4.62 (t, 5.1)	1’, 3’	1’, 4’
3’	70.3	4.29 (t, 5.0)	2’, 4’	1’, 5’
4’	81.6	4.16 (m)	3’, 5’	3’
5’	63.5	4.48 (dd, 3.7, 12.0)	4’	3’, 6’’
		4.35 (5.7, 12.0)	4’	3’, 4’, 6’’
NH–1’’		11.96 (s)	3’’, 4’’, 5’’	
2’’	121.4			
3’’	115.4	6.82 (m)	NH–1’’, 4’’, 5’’	2’’, 4’’
4’’	109.7	6.18 (m)	NH–1’’, 3’’, 5’’	
5’’	124.4	7.04 (m)	NH–1’’, 3’’, 4’’	2’’, 3’’, 4’’
6’’	160.1			

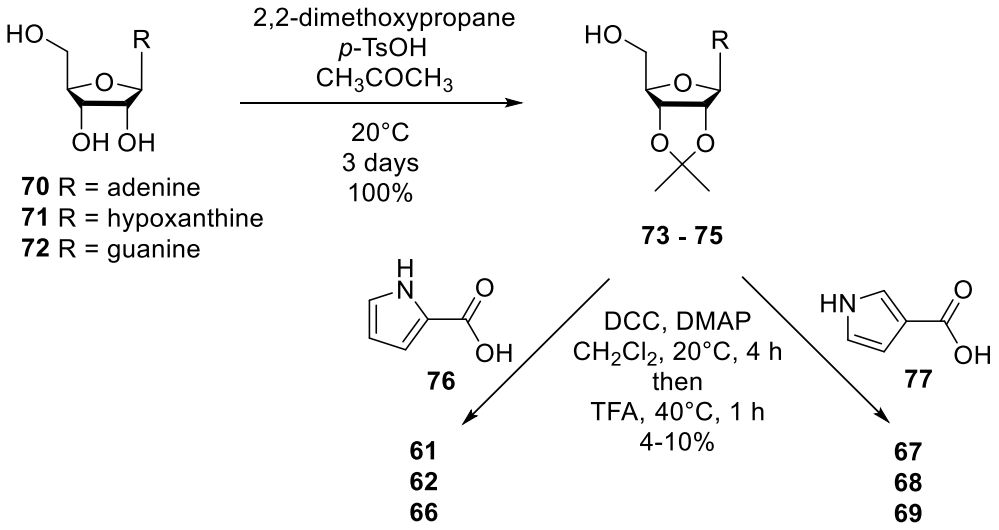


62

4.5 – Total Synthesis of Nelliellosides A and B and Non-natural Analogues

The total synthesis of natural products is desirable for two key reasons; to provide a route for commercial access for supply, and for unequivocal evidence of structure elucidation. Synthesising a molecule using known starting materials and predictable reactions, and comparison of spectroscopic data to that of the natural sample, is the ultimate proof of structure. Total synthesis and comparison of chiro–optical properties is the most robust method to assign absolute configuration, although computational methods are becoming more simple and trustworthy.[10] Total synthesis also provides a route to more compound mass, which is essential for the development of a drug through preclinical and clinical trials, particularly when the natural compound's source cannot be collected easily or is ecologically detrimental in the long term. In the case of the nelliellosides, total synthesis provided enough material to test against a range of biological targets, which would not have been possible using the limited natural samples alone, giving more information on the potential applications of the metabolites.

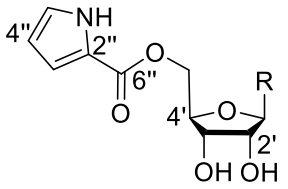
Compounds **61** and **62** are nucleosides, both containing pyrrole-2-carboxylate esters attached at the C-5' alcohol of adenosine and inosine, respectively. As both parent nucleosides (**70** and **71**) are readily available, a method was devised to protect the two free secondary alcohols at C-2' and C-3', leaving the primary alcohol to react with an activated pyrrolic acid. These two alcohols are *cis*, and therefore were protected by acetonide formation using 2,2-dimethoxypropane and catalytic *para*-toluenesulfonic acid (*p*TsOH) (Scheme 4.3). The acetonide was then esterified with pyrrole-2-carboxylic acid (**76**) using Steglich methodology with *N*,*N*′-dicyclohexylcarbodiimide (DCC) and 4-dimethylaminopyridine (DMAP).[194] DMAP acts as a base, deprotonating the pyrrole-2-carboxylic acid. The acid is then activated towards attack from the primary alcohol by DCC, forming the ester and a urea by-product. The esterified acetonide was not purified prior to its deprotection using trifluoroacetic acid (TFA), which afforded the desired products. The NMR data of the synthesised products matched that of those naturally isolated. The recorded specific rotation data ($[\alpha]_D^{20}$ $-24_{(\mathbf{61}\ \text{natural})}$ vs $-25_{(\mathbf{61}\ \text{synthetic})}$; $-21_{(\mathbf{62}\ \text{natural})}$ vs $-24_{(\mathbf{62}\ \text{synthetic})}$) unequivocally confirmed the absolute configurations proposed by the chiral GCMS method.



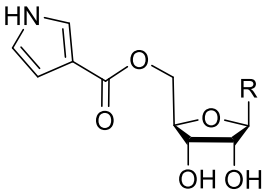
Scheme 4.3 – General procedure for the synthesis of **61**, **62** and **66**–**69**.

When undertaking medicinal chemistry studies, it is beneficial to have more analogues with deviations to a chemical structure for structure-activity relationship (SAR) analyses. Owing to the ease of synthesis of **61** and **62**, a further four non–natural products of the class were synthesised, guanosyl-6′-pyrrole-2-carboxylate (**66**), adenosyl-6′-pyrrole-3-carboxylate (**67**), inosyl-6′-pyrrole-3-carboxylate (**68**) and guanosyl-6′-pyrrole-3-carboxylate (**69**). Compound **66** was synthesised using guanosine (**72**) as the starting nucleoside and was the next logical analogue in the panel owing to the parallel role of guanine in many biological processes to adenine. Guanine is chemically more similar to hypoxanthine, except it bears a NH_2 group at C–2. The product structure was confirmed by NMR and MS analysis (Appendix 5).

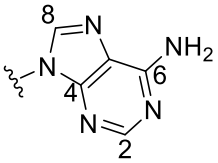
Synthesis of the other three analogues, compounds **67**–**69**, used the same three nucleosides, however coupled to pyrrole-3-carboxylic acid (**77**). This served two key roles: first to test the effect of 2”- vs 3”- substitution on bioactivity (SAR), but also to confirm the attachment of the ester to the position C–2” of the pyrrolic acid. All three were successfully synthesised using the predetermined method (Scheme 4.3), again confirmed using NMR and MS data (Appendix 6).



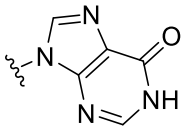
61 R = adenosine
62 R = inosine
66 R = guanosine



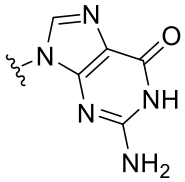
67 R = adenosine
68 R = inosine
69 R = guanosine



Adenosine



Inosine



Guanosine

4.6 – Bioassay

The six compounds were submitted to a range of bioassays within the SBS, VUW, against several different targets. The compounds were first tested for antibacterial activity against *Pseudomonas aeruginosa* and *Staphylococcus aureus*; however, no activity was detected at concentrations below 100 μg mL^{-1}. Next, they were tested against *Saccharomyces cerevisiae* for antifungal activity, where no activity was detected below 200 μM. Finally, the compounds were assayed for cytotoxic activity against the human colon carcinoma cell line HCT-116, where no significant activity was observed for all compounds except weak cytotoxicity at 10 μM for compounds **68** and **69**.

The panel of compounds, as nucleosides, bear a structural similarity to the adenosine phosphates, common cellular metabolites associated with many processes, including being one of the products of adenosine triphosphate (ATP) metabolism. Kinases are enzymes ubiquitous throughout living organisms, responsible for the phosphorylation of substrates, typically using ATP as a starting material and phosphate source. As kinase binding pockets interact with ATP, it was postulated that the nelliellosides and their synthetic congeners may function as kinase inhibitors relevant to the treatment and prevention of a range of diseases.

With no precedence for inhibition against any particular kinase, nellielloside A (**61**) was submitted for a full kinome-wide inhibition assay against 485 human disease-relevant kinases at a commercial testing facility.[195] This screen utilises three main types of assay, depending on

the function of the kinase being tested. The Z'-LYTE assay utilises the disruption of Förster resonance energy transfer (FRET, aka fluorescence energy transfer) to quantify the amount of a phosphorylated synthetic peptide using a development reagent that only cleaves non-phosphorylated peptides and interrupts FRET, thus to get FRET emission, the kinase must be active. The ratio between donor and acceptor emission is then used to determine the extent of phosphorylation and thus kinase inhibition relative to the control. FRET is also used in the LanthaScreen activity assay where a terbium-labelled antibody specific towards only one phosphorylated substrate is used. The substrate has a fluorescein tag, which results in FRET emission when bound by the antibody, thus kinase activity is proportional to relative emission intensity.

Rather than direct substrate phosphorylation, the Adapta assay indirectly observes kinase activity by measuring amounts of the by-product, adenosine diphosphate (ADP). A europium-labelled anti-ADP antibody is added to the reaction mixture along with an ADP tracer, that when bound to the antibody, leads to high FRET emission. Therefore, by measuring the drop in this emission relative to control, the concentration of ADP in the solution is indirectly measured as ADP generated by the kinase reaction outcompetes the ADP tracer and stops FRET.

The results of this screen are presented in Appendix 9. Compound **61** showed significant inhibition (>80% at 10 μM) against 13 kinases of 485 across a relatively narrow range of kinase families including glycogen synthase kinase (GSK), mitogen-activated protein kinase (MAPK), ribosomal s6 kinase (RSK), polo-like kinase (PLK) and germinal centre kinase-like kinase (GLK); many of which are involved in cell cycle progression through the MAPK and mTOR pathways. This was an extremely promising result as it provided targets for further SAR development of the compounds. Broad spectrum activity against a large range of kinases is undesirable and provides no clear developmental leads, which would likely be the result of a conserved element such as the nucleoside portion of the molecule. It was interesting to note that for the assay against IKBKB, a 97% *activation* was observed, thus **61** acted to increase the amount of phosphorylated peptide.

Next, the remaining five nucleosides **62** and **66**–**69** were submitted for single point assays against 10 selected kinases each at the same set concentration of 10 μM. It was decided that IKBKB should be tested due to the unusual initial observation, while the other nine were the kinases most inhibited by **61**. Potent inhibition (and activation) against most of the kinases

tested was observed again (Table 4.3). One kinase from four representative families was then selected to obtain dose-response curves. These results showed potent inhibition of the kinases at sub-micromolar levels, which is promising, considering this is the first stage of drug discovery and development (Table 4.4).

The main objective in producing the panel of compounds was to perform SAR studies once the target was deduced. Comparing the activity against GSK3A, MAPK14 and RSK2, there was a trend in IC_{50} values, with slightly more inhibition of RSK2. At this stage, there is no clear SAR for the purine base or the attachment point of the pyrrolic acid. These target values are extremely important for the future development of the congeners and SAR.

Table 4.3 – Kinase inhibition (%) by **61**, **62** and **66**–**69** at 10 μM against selected targets.

	61	**62**	**66**	**67**	**68**	**69**
GSK3A (GSK3 alpha)	99	94	98	101	101	93
GSK3B (GSK3 beta)	100	94	98	98	99	96
IKBKB (IKK beta)	–97	67	–85	–51	–95	–48
MAPK14 (p38 alpha)	95	95	93	95	95	84
MAPKAPK2	95	42	88	87	89	82
MAPKAPK3	98	83	101	98	100	99
MELK	95	97	96	98	97	87
RPS6KA3 (RSK2)	96	102	96	94	95	93

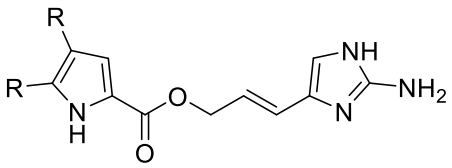
Table 4.4 – IC_{50} values (μM) against selected targets.

	GSK3A	MAPK14	RSK2
61	0.89	1.00	0.80
62	1.11	1.32	0.23
66	0.71	0.95	0.66
67	1.00	1.71	0.97
68	0.61	0.87	0.64
69	0.67	0.99	0.68

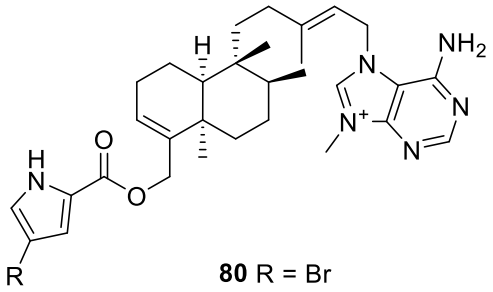
4.7 – Related Compounds

Owing to the widespread involvement of DNA and RNA in cellular processes, nucleoside natural products are of interest for drug discovery,[196] and have demonstrated a range of potential biological applications including antiviral,[197] antibacterial[198] and anticancer[199] activity. Marine sources have contributed a wealth of new structures to the pool of known nucleoside derivatives,[200,201] and two nucleosides hold a vital place in the history of marine natural products; sponge-derived cytarabine and vidarabine (**4** and **5**) were the inaugural marine-derived FDA approved drugs.[202] Only one example of the pyrrole-2-carboxylic acid moiety has been reported on a naturally derived nucleoside, neopetroside B (**63**), isolated from the sponge *Neopetrosia* sp.[188] Too little compound was isolated for biological assay, but it would be intriguing in future to test for kinase inhibition, for comparison against the nelliellosides **61** and **62** and congeners **66**–**69**.

Although these compounds represent the only pyrrole-2-carboxylic acid ester nucleosides reported from any natural source, a range of compounds have been isolated with the functionality bound to other moieties, particularly from sponges. Oroidin (**78**) was one of the first compounds of the class, reported in 1971 from an *Agelas oroides* sample collected in Naples, and is often referred to as the archetypical compound class parent.[203] It has attracted interest from synthetic groups, as have analogues such as the des-bromo derivative, clathroidin (**79**).[204] This may be in part due to their potent antiplasmodial activity with no significant cytotoxicity.[205] When a range of bacteria were challenged with these compounds, **78** reduced the growth of Gram-positive species, whereas **79** was inactive, thus activity requires bromination.[206] The effect of pyrrole bromination on bioactivity was also demonstrated through the purine-containing *A. nakamurai* diterpenes, agelasine G and ageline B (**80** and **81**), differing this time by one bromine substitution. Both compounds showed antimycobacterial activity and lacked cytotoxicity, however, only **78** was active against protein tyrosine phosphatase 1B, a therapeutic target for the treatment of type-2 diabetes and obesity.[207] These examples provide a potential direction for future nellielloside synthetic efforts, particularly to assess the potential antimicrobial activities of the nucleosides with brominated acid derivatives.

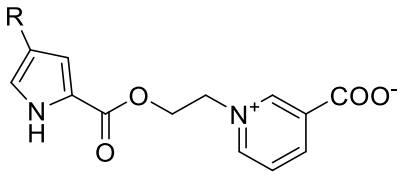


78 R = Br
79 R = H

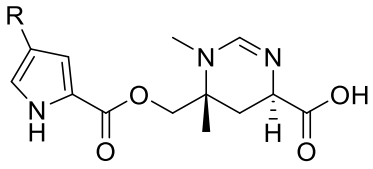


80 R = Br
81 R = H

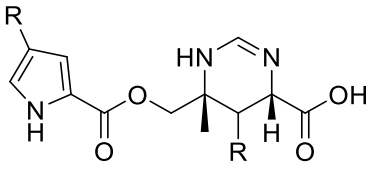
Pyrrole 2-carboxylate-containing marine natural products have shown a wide range of biological activities. A Caribbean sponge, *Agelas longissima*, was the source of agelongine (**82**), a bromopyrrole alkaloid with a nicotinic acid moiety similar to neopetroside B (**63**), which exhibited antiserotonergic activity,[208] while the des-bromo derivative daminin (**83**) was isolated from a Mediterranean *Axinella damicornis* and showed significant neuroprotective effects by reducing a rise in intracellular calcium ions upon neuronal stimulation with no cytotoxic effects.[209] Sponges were also the source of the clathriroles (**84** and **85**) and manzacidins (**86**–**88**),[210,211] which also lack cytotoxic activity and exhibited weak antifungal activity.[212]



82 R = Br
83 R = H



84 R = Br
84 R = H

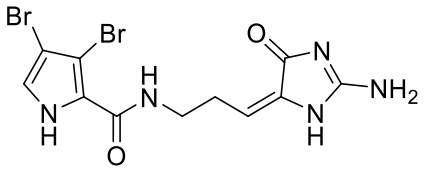


86 R = H
87 R = OH
88 R = H (C-11 epimer)

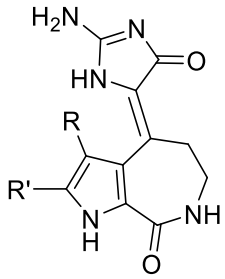
The phosphorylation of proteins by kinases is one of the main drivers of signal transduction and propagation in all cells, and controls many cell cycle processes including metabolism, progression, apoptosis and transcription.[213] Consequently, perturbations to their normal function through mutation are often implicated in the molecular pathology of a huge number of diseases. For example, in nearly all cases of chronic myelogenous leukaemia (CML), the tyrosine kinase ABL is fused to another protein, BCR, causing enhanced activation and cell

cycle proliferation through the MAPK pathway.[214] The ABL kinase therefore makes an attractive target for drug treatments. Imatinib is a selective kinase inhibitor, approved for the treatment of CML. It binds close to the ATP binding site of ABL, rendering it closed and unable to phosphorylate substrates and the kinase is thus switched off.[215] Besides their widespread use in oncology, kinase inhibitors have demonstrated a range of other therapeutic uses, particularly in autoimmune and inflammatory diseases such as asthma, age-related muscular degeneration and rheumatoid arthritis, for which the drug tofacitinib is approved for use.[216] There has even been interest towards the inhibition of kinases of infectious disease-causing pathogens, with kinase targets now recognised for malaria and TB.[217,218]

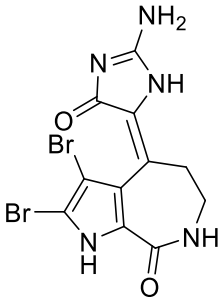
Kinase inhibitors derived from organisms of the sea are a chemically diverse and medicinally relevant family, to which the nelliellosides now belong. Recently isolated examples have been recently thoroughly reviewed.[219] Some compounds with similar structural motifs to the nelliellosides showed kinase inhibitory activity. Extracts of the Indonesian sponges *Stylissa massa* and *Stylissa flabelliformis* resulted in the isolation of 25 brominated pyrrole-alkaloids that were assayed for kinase inhibitory activity.[220] Dispacamide E (**89**) is a dibrominated pyrrole amide, with a similar structure to oroidin (**78**), and showed IC_{50} values of 2–10 μM against a range of kinases, including GSK3 at 2.1 μM, and no significant cytotoxic activity. The most potent inhibitors of GSK3 (10–40 nM) range were cyclised oroidin derivatives, such as hymenialdisine congeners (**90**–**92**) and spongiacidin B (**93**).[220] Methylpenicinoline (**94**) was isolated from a marine *Penicillium* sp., and although it does not have a pyrrole-2-carboxylate ester, it shares an appended pyrrole functionality with the nelliellosides and also inhibits the p38 MAPK pathways.[221] This inhibition stopped the propagation of a LPS-induced inflammatory response *in vitro* and thus may have anti-inflammatory applications.



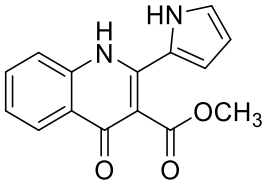
89



90 R = H, R' = H
91 R = H, R' = Br
92 R = R' = Br



93



94

A recent investigation of the Saudi Red Sea tunicate *Symplegma rubra* led to the isolation of a range of purine alkaloids, all of which were tested for kinase inhibition against GSK3.[222] The only nucleoside isolated was inosine (**71**), however, even though the authors state 'moderate kinase inhibitory activity was observed', the IC_{50}'s for all compounds were >10 μM. This served as an excellent control experiment to the nelliellosides and showed the kinase inhibitory activity of the nelliellosides is not a result of the nucleoside core alone.

4.8 – Potential Applications

Owing to the widespread role of kinases in cellular function, particularly in the cell cycle and proliferation, kinase inhibitors are often targeted towards anti-cancer functions. The lack of overt cytotoxic activity observed in HCT-116 cancer cells, yet inhibition of cancer-relevant kinases, suggests applications of the nelliellosides and congeners may lie in alternative diseases. To completely rule out oncological utilisation, however, other cancer cell lines must be assayed to ensure this is not specific to HCT-116; there is a possibility that the mutations involved in making these specific cells cancerous may reduce their sensitivity to these kinases. A broad-spectrum test against the USA National Cancer Institute's (NCI) 60 cancer cell line panel would be the best method to address this.

RSK2 is a serine-threonine kinase, involved in cell proliferation and differentiation through the MAPK pathway.[223] Its main mechanism of action is the phosphorylation of histone H3, controlling chromatin structure and the transcriptional regulation of specific genes, thus it is no

surprise inhibitors typically target cancer.[224] However, inflammatory diseases are often also caused by dysregulation of transcription, such as rheumatoid arthritis (RA) which has been linked to increased serum levels of basic fibroblast growth factor, an upstream regulator of RSK2.[225] Kaempferol, a natural product RSK2 inhibitor,[226] reduced the severity and frequency of RA in a mouse model,[225] therefore proving inhibition may reduce the inflammation and tissue destruction otherwise observed.[227] RSK2 has also been implicated as a target for anti-HIV therapy, owing to its role in the transcriptional regulation of the integrated HIV provirus. The transactivator Tat is responsible for full transcription and viral activity, however in experiments with both RSK2 knockouts and small molecule inhibition, it did not function.[228] Another activator of HIV transcription, ORF45, has also been shown to mediate activation through RSK2,[229] thus selective inhibition may prove a valuable way to control and treat HIV infection.

Although IC_{50} values of **61**, **62** and **66**–**69** were obtained against GSK3A, this enzyme was selected as a representative of the GSK family and in fact, GSK3B was also completely inhibited by all compounds at 10 μM. These two enzymes are paralogues, meaning they are homologous yet derived from different genes, and our result is consistent with all other GSK inhibitors to date, which inhibit both isoforms with similar potency.[230] GSK has gathered much attention due to its unusual cellular activity, being constitutively active and usually inhibited in cell pathways by signals.[231] Its function is widespread and involved in a huge number of pathways, however its main attraction is the direct involvement in the pathology of Alzheimer's disease through interactions with Tau protein, and also its observed inhibition by the mood stabiliser lithium used in the treatment of psychiatric disorders.[232]

Psychiatric disorders such as bipolar and schizophrenia have been linked to abnormally regulated GSK3, through its effect on fundamental neuronal processes and development.[62] It has been shown that two completely different mood stabilisers, lithium and valproic acid, both have varying mechanisms of action, but converge with the inhibition of GSK3.[233,234] Although enzymatic activity is nearly impossible to measure in live brain tissue, the link between GSK3 over-activation and Alzheimer's disease is well established through the pathological hallmarks that all arise from its dysregulation.[235] These include the over-phosphorylation of Tau, increased production of β-amyloid plaques, neuroinflammation and decreased production of acetylcholine, making GSK3 an attractive target for treatment of the disease.[235] The use of GSK3 inhibitors have shown efficacy *in vivo* to reduce β-amyloid-induced neurodegeneration, but caused severe neurological side effects owing to the promiscuous nature of the enzyme.[236]

Therefore, for effective treatment of Alzheimer's disease through GSK3 modulation, a function-specific inhibitor would need to be developed. Over-activated MAPK14 is also involved in Alzheimer's disease pathology,[237] and its inhibition has been shown to mitigate defects in the autophagy-lysosomal system, reducing Alzheimer's pathogenesis in mouse model.[238] As the nucleosides also inhibit MAPK14 and other kinases involved in the MAPK pathway, there are two potential targets and mechanisms to which they could treat Alzheimer's disease. MAPK14 is also involved in the propagation of inflammatory signals and has thus been proposed as a target for inflammatory bowel disease.[239] The potential targets and applications of the nelliellosides are the subject of future studies in our laboratory.

4.9 – Conclusion

A ^{1}H NMR spectroscopy screen of an extract of the Tongan bryozoan *N. nelliiformis* revealed numerous resonances indicative of interesting metabolites, particularly in the aromatic and oxymethine/oxymethylene region of the spectrum. Through a ^{1}H NMR spectroscopy-guided isolation procedure, new compounds nelliellosides A and B (**61** and **62**) were isolated, and they, along with four congeners (**66**–**69**), were synthesised from commercial starting materials. These nucleosides showed potent sub-micromolar activity against a range of kinases including GSK3A, MAPK14 and RSK2. On retrospective assessment of the 30% Me_2CO/H_2O screening fraction ^{1}H NMR spectrum (Figure 2.9), the major peaks are from **61**. This study displays how well ^{1}H NMR spectroscopy can be used as a screening technique, if the metabolite of interest represents a major proportion of the acquired data. This sample would not have been suited to a mass-spectrometric screen as it represented the only bryozoan in the collection, therefore there are no comparative examples to detect chemical differences. The NMR-based screening approach does not require a comparative dataset, therefore works very well as a standalone technique.

Chapter 5 – *Synoicum kuranui*

Chapter 5 describes the chemical investigation of the NZ tunicate Synoicum kuranui. *This organism was prioritised based upon the network generated in the mass spectrometric screen, along with many aromatic signals in the ^{1}H NMR spectrum of its screening fractions. This resulted in the isolation of two new (**110** and **111**) and two known (**108** and **109**) rubrolides present in a complex mixture of previously reported metabolites. All of compounds **108-111** showed strong antibacterial activity against Gram-positive* B. subtilis *bacteria.*

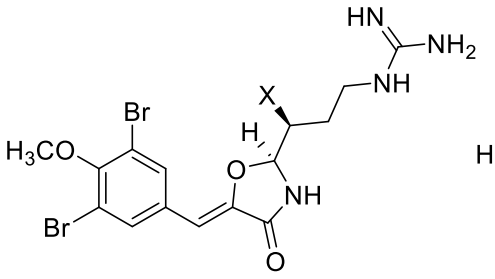
5.1 –*Synoicum* sp.

The genus *Synoicum* is one of nine currently accepted genera of the family Polyclinidae, comprising 82 different species.[240] According to the MarinLit database,[9] these colonial tunicates have been the source of 43 compounds including alkaloids,[241-244] terpenes/steroids,[245,246] macrolides[247,248] and bromotyrosine-derived rubrolides,[249,250] cadiolides,[251] synoilides[251] and prunolides.[252]

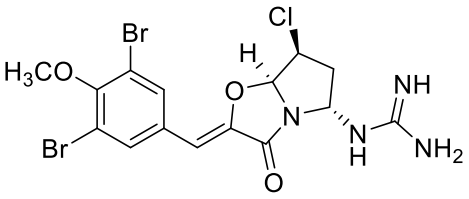
The synoxazolidinones are a family of halogenated alkaloids derived from a brominated tyrosine metabolite, and contain a rare 4-oxazolidinone ring structure.[241,242] Synoxazolidinones A and B (**95** and **96**) were first reported from a sub-Arctic specimen of *S. pulmonaria*, and demonstrated antibacterial and antifungal activities.[241] This was followed by the isolation of synoxazolidinone C (**97**) from a different sample of the same species, which contains an extra pyrrolidine ring as a cyclised derivative of **95**.[242] In contrast to the other congeners, compound **97** showed general cytotoxicity to a range of human cancer cell lines and normal lung fibroblast cells. To determine the absolute configuration of the multiple stereocentres in **95** and **97**, both molecules were subjected to a rigorous comparison of experimental and calculated data for a range of chiroptical techniques, where vibrational circular dichroism proved more reliable and accurate than ECD.[253] The aqueous extract of the *S. pulmonaria* sample contained additional new compounds, the acetylcholinesterase inhibitors pulmonarins A and B (**98** and **99**).[243] Although they share the 4-methoxy-3,5-dibromo-benzyl substructure, it is not believed these metabolites share a common biosynthetic route.

Possibly the most significant *Synoicum*-derived class of compounds to date are the palmerolides, first isolated by Baker and co-workers from an Antarctic *S. adarenum* sample.[247] Palmerolide A (**100**) is an enamide containing polyketide, and showed potent targeted activity against the melanoma cell line UACC-62 (LC_{50} = 18 nM), three orders of magnitude higher than any other cancer cell line assayed in the NCI-60 cell line panel. It was isolated alongside

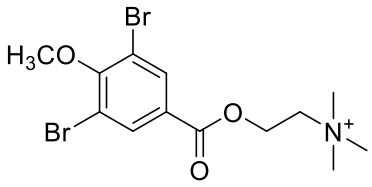
palmerolides B and C (**101** and **102**),[254-256] and further investigation of the minor extract components led to the characterisation of palmerolides D–G (**103**–**106**), typically present in ~100-fold less mass, yet **103** was more potent against melanoma than lead compound **100**.[248] The carboxamide tail of **103** is extended, the only difference from **100**, and was thus an important SAR discovery from the minor component. The naturally isolated palmerolides and synthetic derivatives have been the topic of medicinal chemistry projects,[257] where benzamide analogue **107** is the most potent anti-melanoma palmerolide tested to date with GI_{50} (Mean 50% growth inhibition) of 9 nM against the UACC-62 cell line.[258,259]



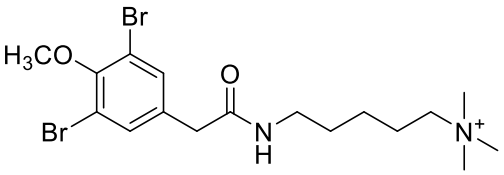
95 R = Cl
96 R = H



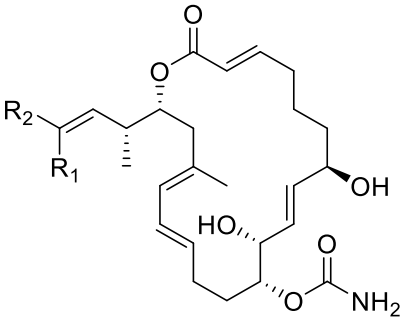
97



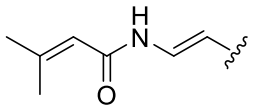
98



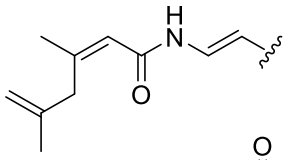
99



100 $R_1 = CH_3$ R_2 =

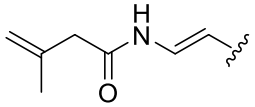


103 $R_1 = CH_3$ R_2 =

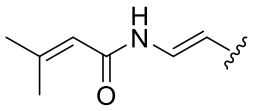


104 $R_1 = CH_3$ R_2 =

105 $R_1 = CH_3$ R_2 =

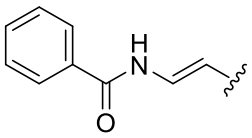


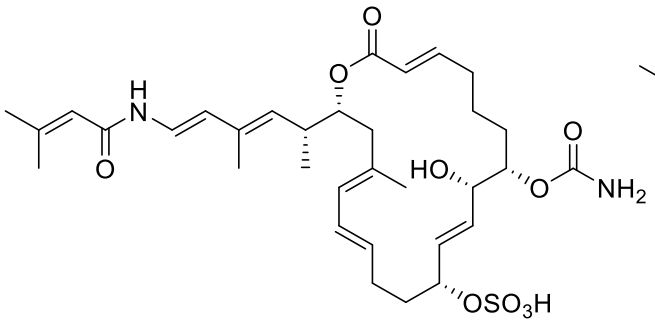
106 R_1 =



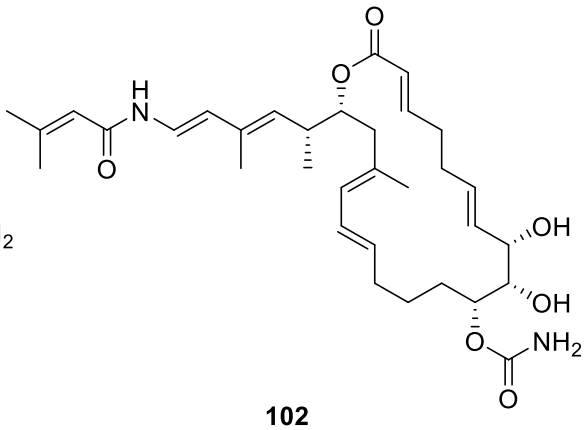
$R_2 = CH_3$

107 $R_1 = CH_3$ R_2 =





101



102

5.2 – Prioritisation and Isolation Procedure

The bright orange tunicate *S. kuranui* (Brewin, 1950),[260,261] was collected from Great Barrier Island in the Hauraki Gulf, NZ, at a depth of 21 m with SCUBA as part of the NIWA collection (Figure 5.1). This tunicate is native to NZ waters, and is commonly encountered up to 40 m deep in the waters of North Cape, Whangarei and Fiordland.[262] With no previously reported investigations of the species, this organism was included in the mass spectrometric screen.

Within the spectrometry-guided Pacific tunicate screen (Figure 2.11), the metabolic extract of *S. kuranui* is represented by the pink nodes. Constellations consisting of nodes made up from only one organism are a clear lead for similarly structured secondary metabolites, and with the second largest number of nodes being pink (Constellation B), *S. kuranui* was highlighted for bulk extraction.

The methanolic extract was fractionated through HP20 (PSDVB) using mixtures of H_2O and Me_2CO, and analysed by ^1H NMR spectroscopy and MS. It was notable that the colour of the extract was maintained in the fractions, most intensely in the 75% Me_2CO/H_2O screening fraction, and when the orange-yellow sample was dissolved in DMSO-d_6, the solution turned a deep red. The ^1H NMR screen spectrum revealed numerous peaks of interest in the aromatic region, and a notable absence of peaks typically associated with fats and primary metabolites in the 0–2 ppm range (Figure 5.2). The sample was then further purified using reversed-phase HPLC (Scheme 5.1) to afford and allow the full characterisation of the known rubrolides A and B (**108** and **109**) and new rubrolides T and U (**110** and **111**).

Figure 5.1 – Underwater photo of the marine tunicate *Synoicum kuranui* (Photo courtesy of Mike Page, NIWA).

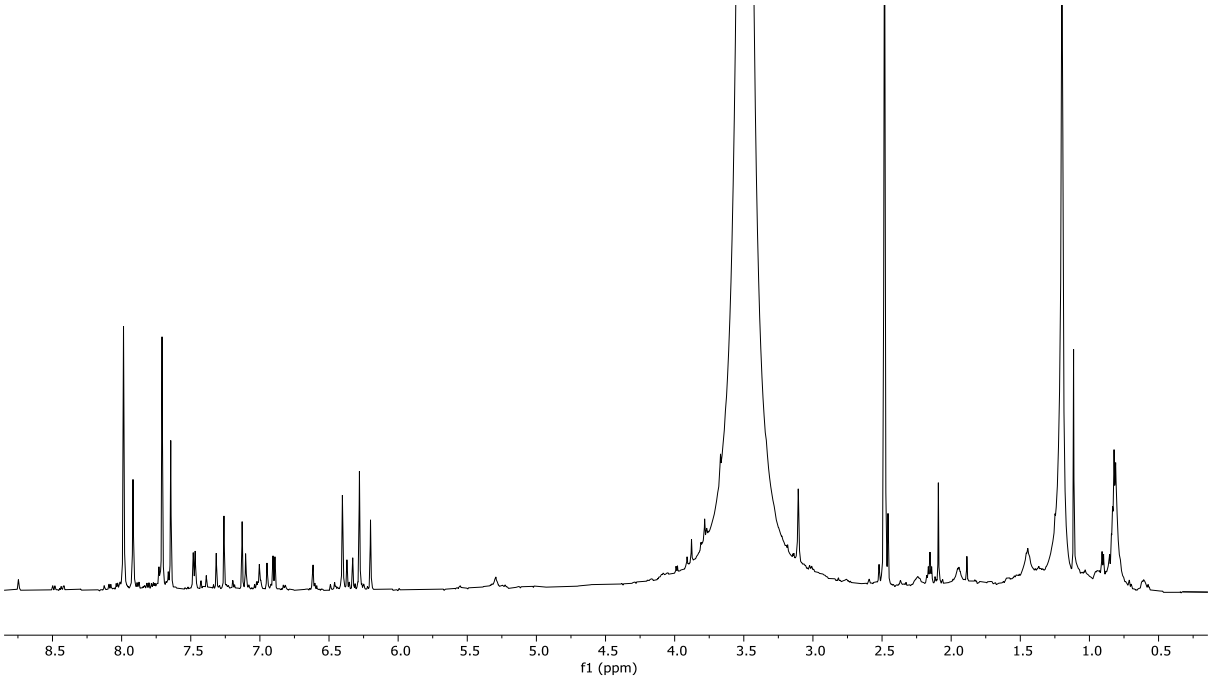
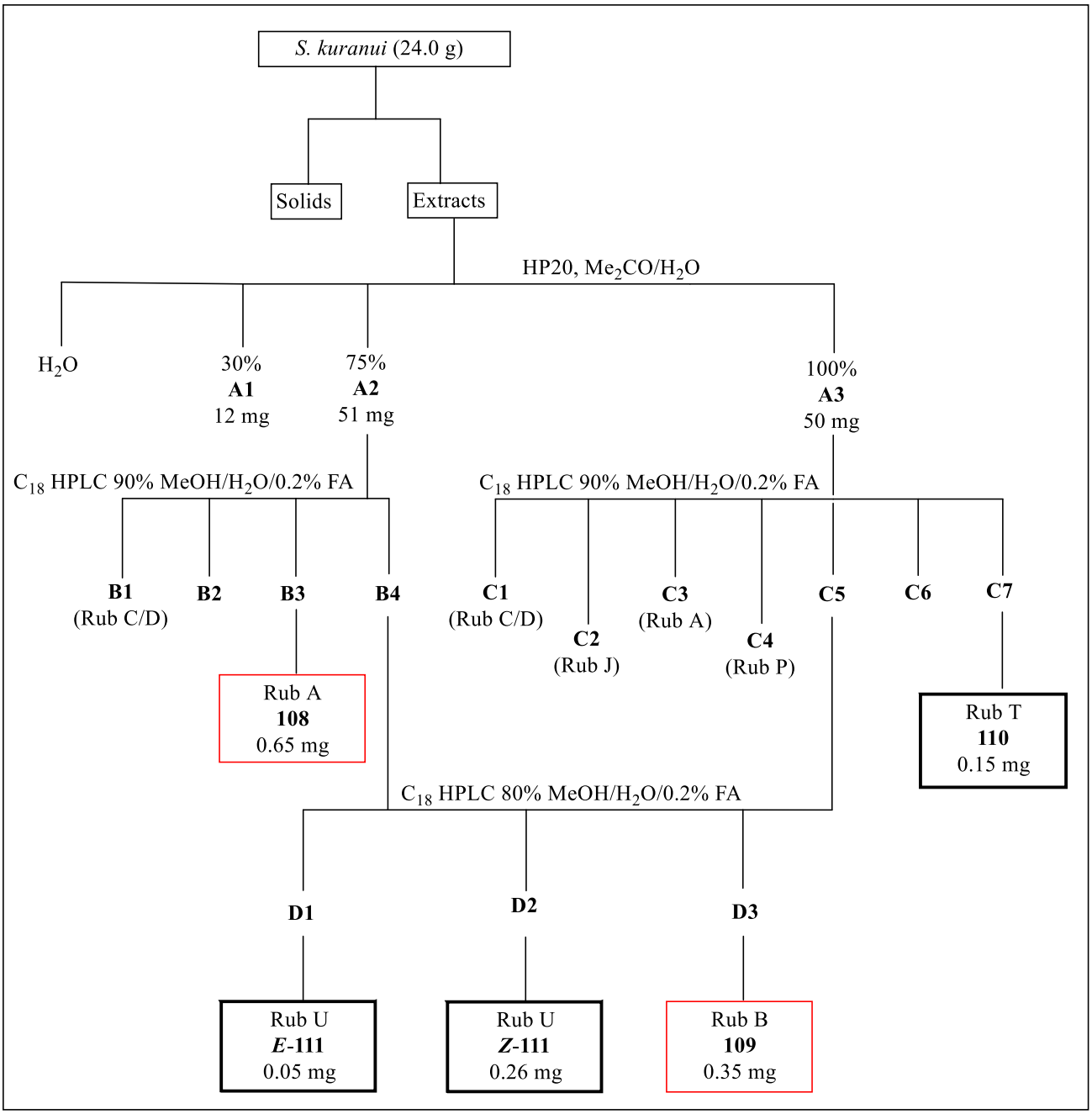


Figure 5.2 – ^{1}H NMR spectrum of *S. kuranui* 75% Me_2CO/H_2O screening fraction (600 MHz, DMSO-d_6).



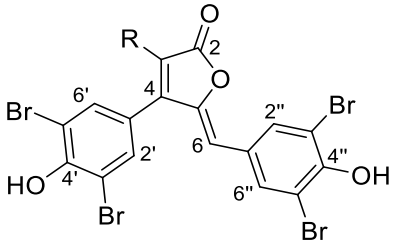
Scheme 5.1 – Isolation procedure for compounds from *S. kuranui*. Red boxes are previously reported compounds, black bold boxes denote compounds identified during the current study. Under each fraction is a tentative assignment of the previously reported rubrolide present based on the *m/z* detected.

5.3 – Previously reported Rubrolides A and B

The major compound present in fractions B3 and C3 was a yellow film that was subjected to HRESIMS analysis to give a deprotonated molecular ion cluster at *m/z* 590.7075, indicative of the molecular formula $C_{17}H_7O_4Br_4$ (calcd. 590.7083). The presence of four bromine atoms was clear from the 1:4:6:4:1 quintet isotopic distribution pattern. The ^{13}C NMR spectrum consisted of only 13 deshielded resonances, including that of an α,β-unsaturated ester carbonyl resonance at δ_C 168.0, and indicating symmetry within the molecule. The ^1H NMR spectrum showed four deshielded singlet methine signals, with two resonances each integrating for two (δ_H 7.77 and 8.05) and one (δ_H 6.35 and 6.55) protons each (Appendix 7).

This data was used to search the MarinLit database, with rubrolide A (**108**) being the clear hit, both spectroscopically and from a taxonomic perspective. A comparison of the NMR data from the original isolation confirmed this.[263] This identification was further verified by the change from orange-yellow in MeOH to deep red in DMSO, which was reported in the original publication.

A minor fraction (D3), also a yellow film, was also subjected to HRESIMS analysis, which detected a deprotonated molecular ion cluster at *m/z* 624.6694, indicative of the molecular formula $C_{17}H_7O_4Br_4Cl$ (calcd. 624.6694). This was dereplicated as rubrolide B (**109**) after matching the MS and ^1H NMR data. As only a small amount was isolated, a ^{13}C NMR spectrum was not obtained, however chemical shift values from HSQC and HMBC correlations corroborate this assignment.[264]



108 R = H
109 R = Cl

5.4 – New Rubrolide Prioritisation by Molecular Networking

With the structural class of the major metabolites deduced from the 75% Me_2CO/H_2O screening fraction, molecular networking was then used to probe the remaining two *S. kuranui* fractions (30% and 100% Me_2CO) for new analogues that have not been reported to date, along with dereplication of known minor metabolites. From the major constellation generated, compounds **108** and **109** and 10 other nodes were putatively annotated, based on the precursor ion matching the *m/z* of a previously reported rubrolide (Figure 5.3). Based on the node connections and precursor ions, two previously unreported target masses were identified, with molecular formulae $C_{18}H_{10}O_4Br_4$ and $C_{18}H_{11}O_4Br_3$. These ions were then used to prioritise fractions for further purification.

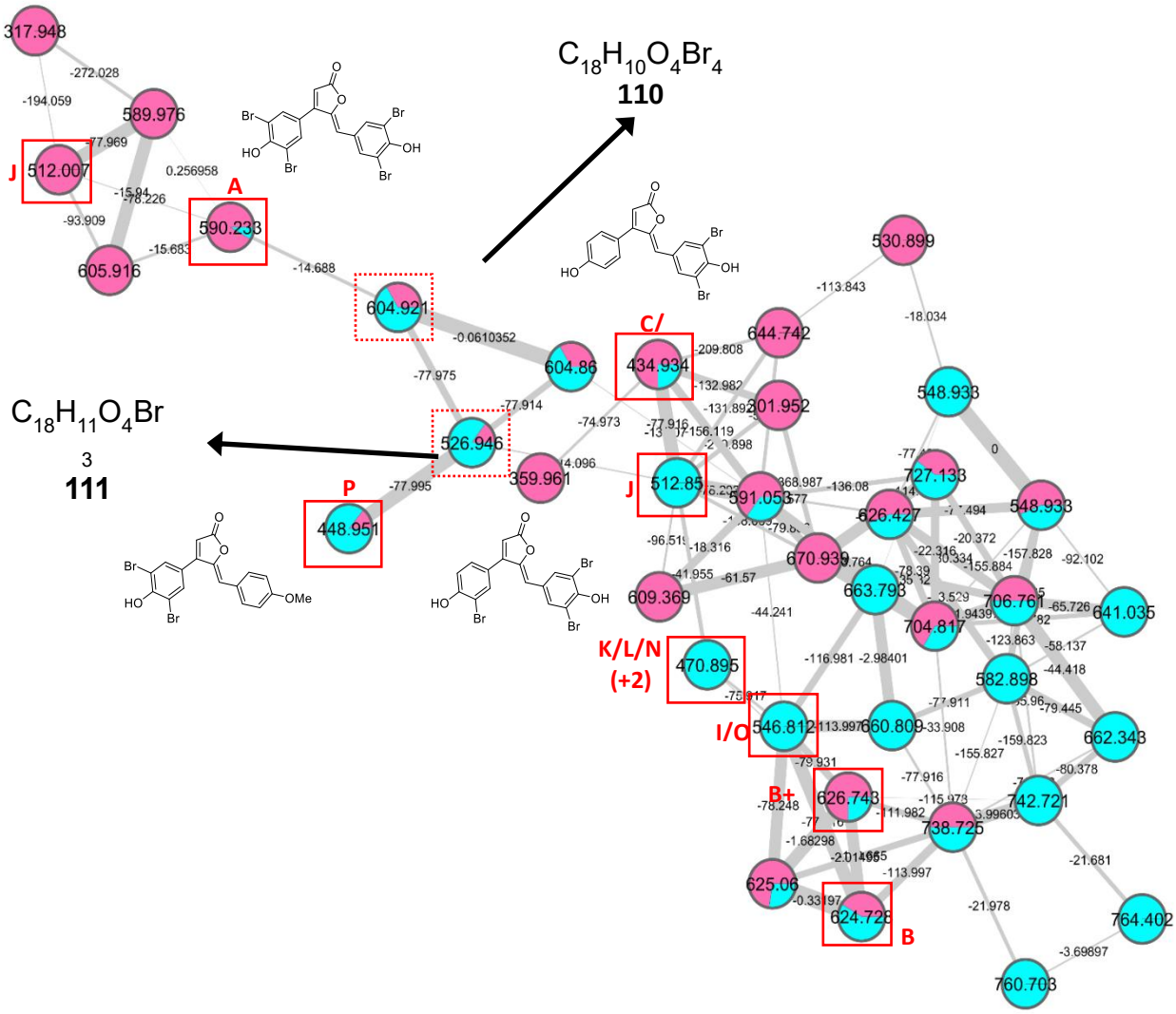


Figure 5.3 – Molecular network of *S. kuranui* screening fractions. Pink is the 30% Me_2CO/H_2O and aqua is the 100% Me_2CO screening fraction. Red letters denote known rubrolide annotations, while +2 denotes the M+2 precursor ion.

5.5 – Rubrolide T

Through a targeted isolation procedure guided by MS, a new analogue named rubrolide T (**110**) was isolated as a yellow film. Analysis of the deprotonated molecule at *m/z* 604.7254 in the HRESIMS, indicated the molecular formula $C_{18}H_9O_4Br_4$ (calcd. 604.7240), consisting of an IHD of 12 and differing from that of **108** by the addition of CH_2. The presence of four bromine atoms was clear from the 1:4:6:4:1 quintet isotopic distribution pattern (Appendix 7). The UV/vis spectrum showed evidence of a highly conjugated chromophore, while the ^{13}C NMR spectrum consisted of one methoxy and 13 deshielded resonances, including an *α,β*-unsaturated ester carbonyl resonance at δ_C 168.5, which indicated an element(s) of symmetry within the molecule similar to **108**. The ^1H NMR spectrum showed a methoxy (δ_H 3.83) and four aromatic singlets, two integrating for two protons (δ_H 7.59 and 8.15) and the other two (δ_H 6.21 and 6.44) for one proton, respectively. The COSY spectrum revealed no correlations; therefore, it was apparent that no protons were vicinal to another.

The aromatic methine signal at δ_H 8.15 showed correlations to a signal at δ_C 134.2 in both the HSQC and HMBC spectra, thus when the integration was also considered, was indicative of two methines symmetrically substituted about a phenyl ring. This proton resonance also showed HMBC correlations to signals at δ_C 153.1, 117.7 and 108.2. The correlation to an oxygenated carbon at δ_C 153.1 was also shared with the methoxy signal at δ_H 3.83 which must be on the same ring. Owing to the singlet methine and the symmetry required, these three substituents must be *meta* to one another. As the carbon signal at δ_C 108.2 is protonated (δ_H 6.44) and this ^1H signal did not show a COSY correlation to δ_H 8.15, it must be outside the aromatic ring, so a pair of brominated carbons were assigned to positions C–3''/C–5'' (δ_C 117.7). The C–3''/C–5'' chemical shift value is typical for a brominated aromatic carbon based on comparison to chemical shifts of previously reported compounds (e.g. **97**).[242] Thus, substructure I was deduced (Figure 5.4). An analogous workflow starting from the other two proton aromatic methine (δ_H 7.59) also derived a 1,3,4,5-tetrasubstituted phenyl ring, however with no methoxy substitution at the C–4' position, gave rise to substructure II.

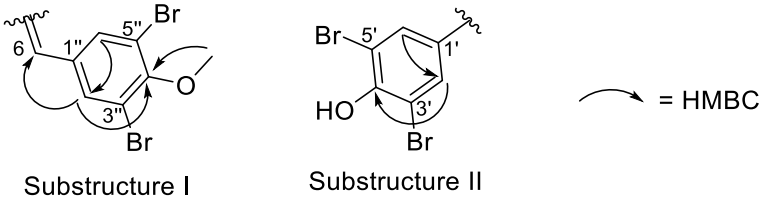
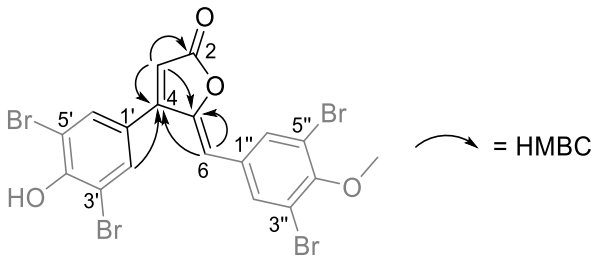


Figure 5.4 – Substructures I and II and key HMBC correlations of **110**.

The remaining substructure of **110** was made up of C_4HO_2, and included an α,β-unsaturated ester carbonyl (C–2, δ_C 168.5). As this component had to account for three IHD units, a butenolide ring was the only plausible way to satisfy the requirements. The remaining proton singlet (δ_H 6.21) was assigned to C–3, as it had the only correlation to carbonyl C–2. H–3 also shared HMBC correlations to quaternary carbon δ_C 156.6 with both H–6 and H–2'/6', which assigned this as C–4, and consequently δ_C 148.8 to oxygenated C–5 (Figure 5.5).



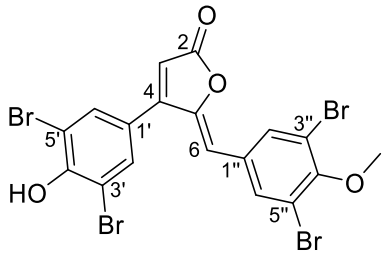
110

Figure 5.5 – Key HMBC correlations completing the structure of **110**.

This molecule differs from the major rubrolide A **108** solely by the presence of a methoxy group at C–4'', and as a new member of the class, is termed rubrolide T. A general trend has been observed that phenol methylation reduces antibacterial and cytotoxic activity of rubrolides in both synthetic[265] and a naturally isolated[266] analogues.

Table 5.1 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for Rubrolide T (**110**) in DMSO–d_6.

position	^{13}C (δ)	^{1}H (δ, int., mult.)	HMBC
2	168.5		
3	105.7	6.21 (1H, s)	2, 4, 5
4	156.7		
5	148.8		
6	108.2	6.44 (1H, s)	4, 5, 2’’/6’’
1’	129.9		
2’/6’	131.9	7.59 (2H, s)	4, 2’/6’, 3’/5’ 4’
3’/5’	115.8		
4'	164.9		
1’’	133.0		
2’’/6’’	134.2	8.15 (2H, s)	6, 2’’/6’’, 3’’/5’’, 4’’
3’’/5’’	117.7		
4''	153.1		
OMe–4’’	60.6	3.83 (3H, s)	4’’



110

5.6 – Rubrolide U

It was noted that during HPLC purification of fractions B4 and C5, containing ions at *m/z* 526.8143 in the HRESIMS spectrum, there was two UV-active peaks with identical *m/z* ions. In ^{1}H NMR analysis, both fractions had identical resonances, however, these were split into two sets of signals of different relative intensities. This suggested both distinct HPLC peaks contained the same two separate compounds. This has been observed before for rubrolides O, P, Q and also cadiolide F,[250,267] all of which exist as two separable but interchangeable isomers, differing by the *E* or *Z* configuration of the C–5 to C–6 double bond. All previous examples are mono-brominated at the benzylidene ring, and are often methylated at the phenolic oxygen.

The major rubrolide U isomer (*Z*-**111**) was isolated as a yellow film, and HRESIMS analysis detected the deprotonated molecule at *m/z* 526.8143, indicative of the molecular formula $C_{18}H_{10}O_4Br_3$ (calcd. 526.8135). The presence of three bromine atoms was clear from the 1:3:3:1 quartet isotopic distribution pattern (Appendix 7). The ^{13}C NMR experiments showed a similar pattern of peaks as **110**, including the single methoxy resonance, however *Z*-**111** had two extra proton signals and with one less bromine in the elemental formula, an element of symmetry was broken. In the ^{1}H NMR spectrum, the coupling pattern of the peaks at δ_H 7.19 (d, 8.7 Hz), 7.81 (dd, 8.7, 2.2 Hz) and 8.11 (d, 2.2 Hz) was characteristic of a 1,2,4-trisubstituted aromatic ring, while the singlet at δ_H 7.56 that integrated for two protons indicated a 1,3,4,5-tetrasubstituted structure, giving rise to four potential isomers (Figure 5.6).

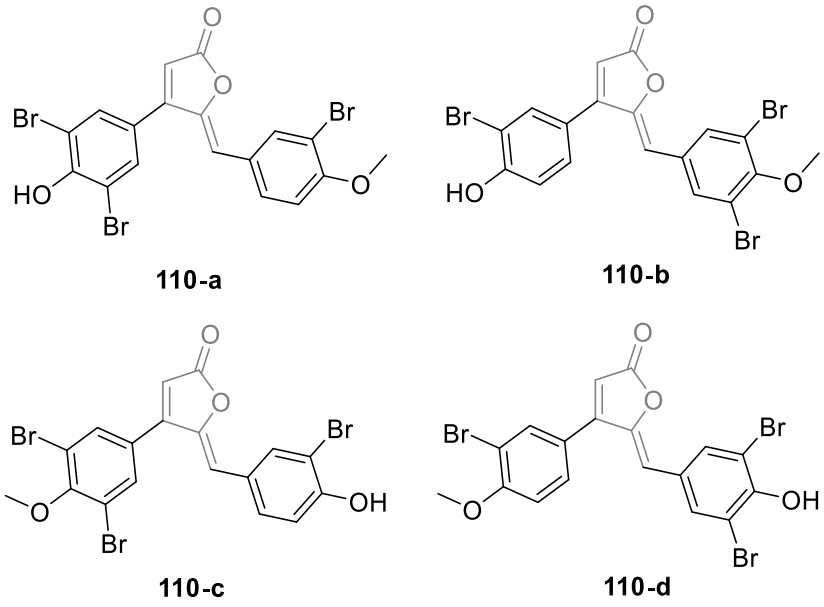


Figure 5.6 – Potential isomers of *Z*-**111**.

The methoxy group at δ_H 3.90 showed an HMBC correlation to C–4'' (δ_C 155.3), which also correlated with the doublet of doublets for H–2'' and the *meta*–coupled doublet for H–6'' (δ_H 7.81 and 8.11, respectively). The methoxy signal also showed a NOE correlation (ROESY experiment) to the *ortho*-coupled doublet H–3'' (δ_H 7.19) only, thus the methoxy group was present on the 1,2,4-trisubstituted aromatic ring (isomers **111-a** or **111-d**). The structure was confirmed as **111-a** by HMBC correlations between H–6 (δ_H 6.37) and both C–6'' and C–2'' (δ_C 134.1 and 131.3). The single bromination of the benzylidene unit is consistent with all other congeners isolated as mixtures of *Z*/*E* C5-C6 isomers. This structure shows a new substitution pattern for the class and was therefore named rubrolide U.

The NMR assignments of the two compounds were made based on relative peak intensity of the resonances in the two separate HPLC fractions, one enriched in *E*-**111** and the other *Z*-**111** (Figure 5.8), in conjunction with chemical shift values and NOE (ROESY) data for stereochemical assignment. For *E*-**111**, the two aromatic rings are closer in space, which resulted in increased shielding of the aromatic signals compared to Z-**111** where they point away from each other (Figure 5.7). In both cases, H–3 showed NOE correlations to H–2'/6', however for the *Z* isomer H–6 correlated to H–2'/6', H–2'' and H–6'' whereas it only correlated to H–2'' and H–6'' for the *E* isomer. A correlation between H–2'/6' and H–2'' was also observed for the *E* isomer but absent for *Z*, consistent with the two aromatic rings in a closer proximity and placing the sterically-demanding C–5'' Br atom away from the other aromatic ring.

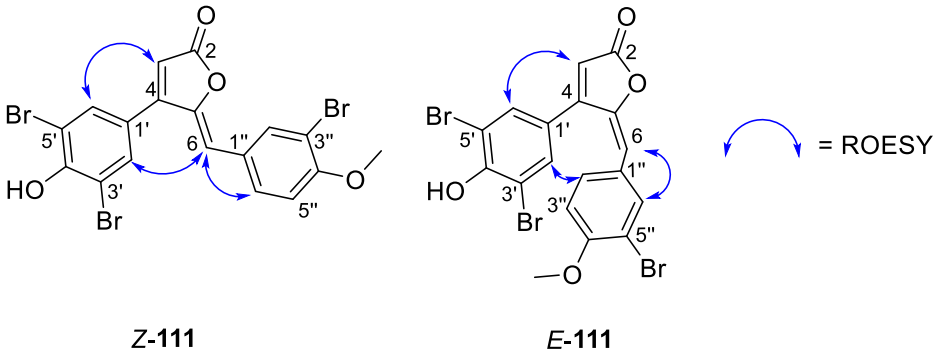
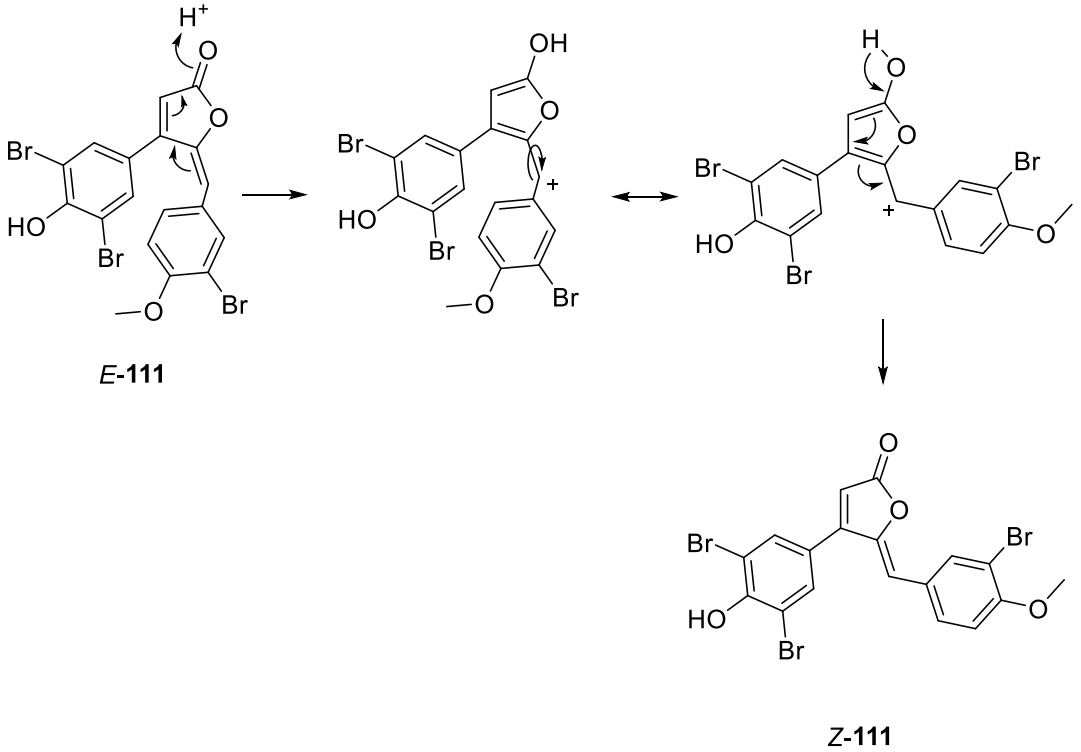


Figure 5.7 – Key ROESY correlations for *E*- and *Z*-**111**.

Although the two HPLC samples were enriched for *E*- or *Z*-**111**, over time the proportion of the resonances corresponding to *Z*-**111** increased, suggesting isomerisation of the *E*-isomer to this compound. In order to obtain better ^{13}C NMR data on the compounds, the two samples

were combined. The *Z* isomer is favoured due to the reduction in steric repulsion of the two aromatic rings that are set co-planar to each other by the extended conjugation across the molecule. When the rubrolides are biosynthesised, a final deprotonation forms the C-5 to C-6 double bond, where both isomers can be formed depending upon the orientation of the benzylidene ring substituent (Scheme 5.2). The electron density of the aromatic ring of the benzylidene group will help stabilise the positive charge formed in the conversion, particularly with the electron-donating methoxy group at the *para* position. Of the four previous compounds reported to display this phenomenon, three also contained a methoxy group on the ring. In fact, *E* isomers of the non–methylated congeners of rubrolide Q and cadiolide F were not detected, therefore it is likely a combination of both factors contributing to carbocation stability. No conversion from the *Z* to the *E* isomer has been observed.



E-**111**

Z-**111**

Scheme 5.2 – Isomerisation from *E*- to *Z*-**111**.

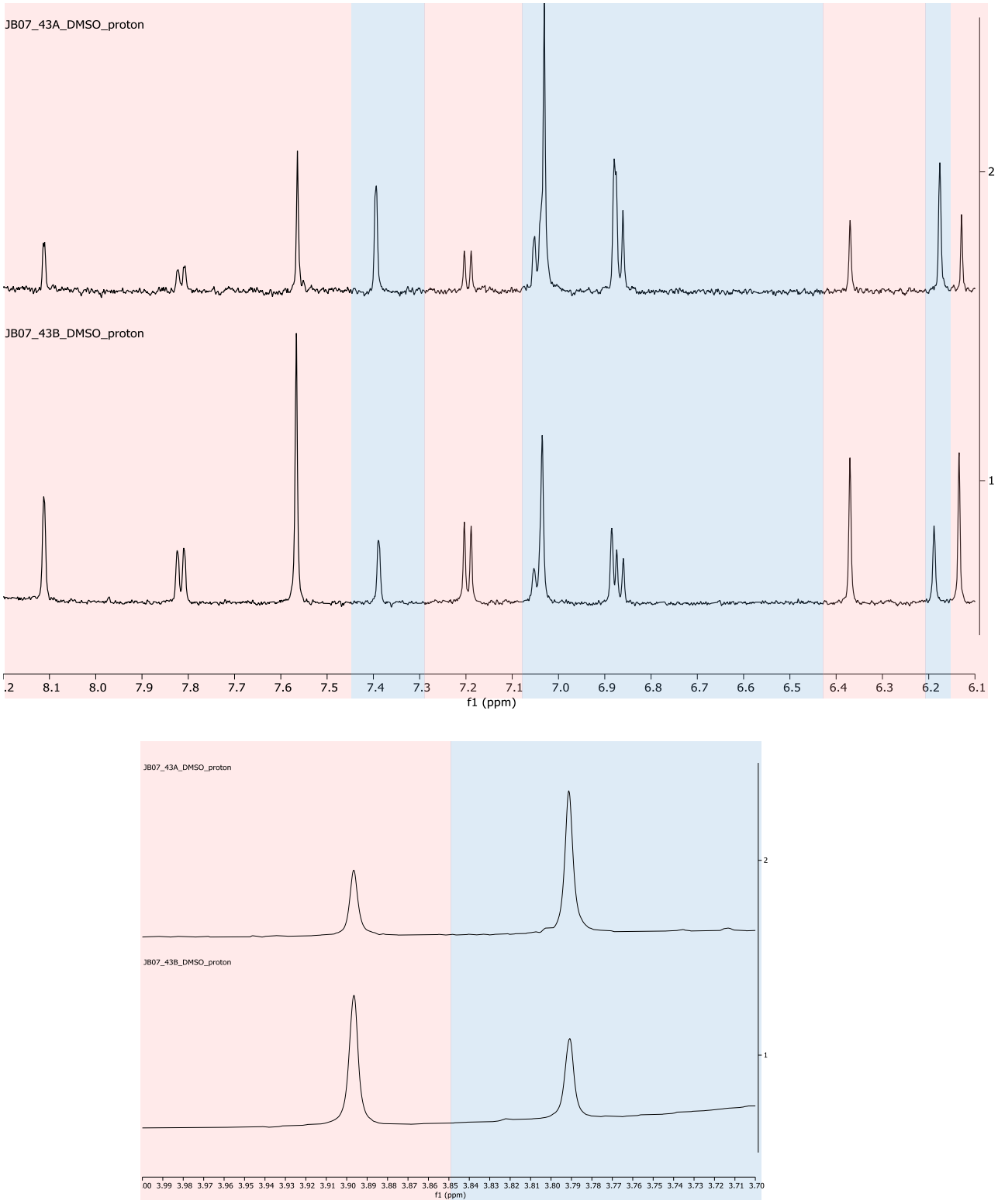
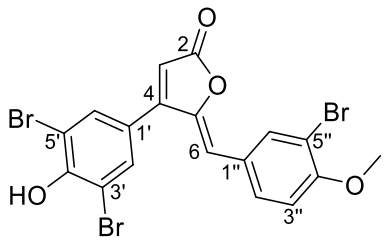


Figure 5.8 – ^{1}H NMR spectra of *E*- (highlighted blue) and *Z*-**111** (highlighted red) (600 MHz, DMSO-d_6).

Table 5.2 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for *Z*-Rubrolide U (*Z*-**111**) in DMSO-d_6.

position	^{13}C (δ)	^{1}H (δ, int., mult., *J* in Hz)	HMBC	ROESY
2	168.6			
3	105.0	6.13 (1H, s)	2, 5	2’/6’
4	156.8			
5	146.7			
6	110.0	6.37 (1H, s)	4, 5, 2’’, 6’’	2’/6’, 2’’, 6’’
1’	131.6			
2’/6’	131.5	7.56 (2H, s)	4, 4’, 3’/5’	3, 6
3’/5’	115.4			
4'	164.6			
1’’	127.4			
2’’	131.1	7.81 (1H, dd, 8.7, 2.2)	4’’, 6’’	6
3’’	112.6	7.19 (1H, d, 8.7)	1’’, 5’’	MeO–4’’
4'’	155.3			
5’’	110.7			
6’’	134.1	8.11 (1H, d, 2.2)	4’’, 5’’	6
MeO–4’’	56.4	3.90 (3H, s)	4’’	3’’

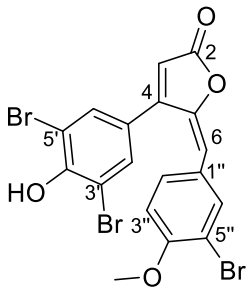


Z-**111**

Table 5.3 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for *E*-Rubrolide U (*E*-**111**) in DMSO-d_6.

position	^{13}C (δ)	^{1}H (δ, int., mult., *J* in Hz)	HMBC	ROESY
2	168.1			
3	110.3	6.18 (1H, s)	2, 5	2'/6'
4	154.7			
5	147.6			
6	113.8	6.88 (1H, s)		2'', 6''
1'	ND			
2'/6'	132.0	7.04 (2H, s)	4, 4', 3'/5'	3, 2'', 6''
3'/5'	114.3			
4'	164.0			
1''	126.5			
2''	130.7	7.05 (1H, dd, 8.7, 2.2)		6, 2'/6'
3''	111.9	6.87 (1H, d, 8.7)	1'', 5''	MeO–4''
4''	155.4			
5''	110.3			
6''	134.1	7.40 (1H, d, 2.2)		6, 2'/6'
MeO–4''	56.2	3.79 (3H, s)	4''	3''

ND = Not detected



***E*-111**

5.7 – Biological Activity

The isolated rubrolides **108**, **109**, **110** and **111** (as a mixture of *E*- and *Z*-isomers) were submitted to SBS, VUW, for antimicrobial testing against Gram-negative (*E. coli*) and Gram-positive (*B. subtilis*) bacteria. No activity was detected at 128 μg mL^{-1} for any compounds against *E. coli*, however all compounds showed strong activity against *B. subtilis*. (Table 5.4). With compounds **110** and **111** methylated at the C–4'' phenol, weak antibacterial activity was expected as this has been observed for other C–4'' methylated rubrolides against the species.[268] However, compound **110** showed roughly ten-fold stronger activity than **108**. All direct comparative studies that observed a decrease in antibacterial activity on compound methylation used other bacterial species such as MRSA, and *S. epidermidis*,[266] therefore this may be a strain specific effect for the new rubrolides upon *B. subtilis*.

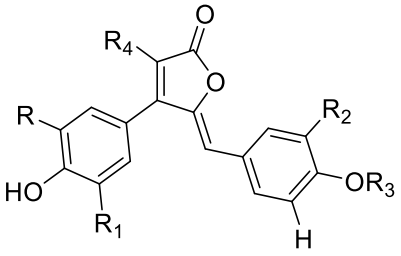
	MIC (μg mL^{-1})	MIC (μM)
108	2	3.36
109	0.5	0.79
110	$\leq$ 0.25	$\leq$ 0.41
111	0.5	0.94
Tetracycline	2	4.5

Table 5.4 – MIC values of rubrolides isolated from *S. kuranui* against *B. subtilis* bacteria. Compound **110** was active at the highest concentration tested. Tetracycline was used as a positive control.

5.8 – Polyaromatic Butenolides

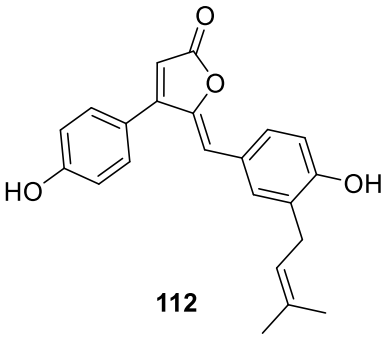
Marine-derived polyaromatic butenolides are a class of often halogenated small molecules, typically isolated from tunicates. Although non-nitrogenous, like many alkaloids they are formed by the condensation of multiple precursors derived from aromatic amino acids such as tryptophan, tyrosine and/or phenylalanine,[269] which can be decorated with halogens to form ring systems that do not contain nitrogen. The class is led by the rubrolides, which are often co-isolated with other butenolide-containing families such as the prunolides and/or the cadiolides. The rubrolides were first isolated from a Canadian (Queen Charlotte Islands, British Columbia) tunicate *Ritterella rubra*,[263] with >20 compounds now known after subsequent isolations from tunicates *S. globosum*,[266] *S. blochmanni*,[249] two unidentified *Synoicum* species,[250,270] *Pseudodistoma antinboja*[268] and the marine-derived fungus *Aspergillus terreus*.[271] Therefore, the discovery in this work of a number of rubrolide derivatives from *S. kuranui* is in line with the taxonomic trend observed for the genus.

The rubrolide core structure consists of β-aryl and γ-benzylidene groups flanking a butenolide core, and are of interest to chemists and biologists alike, owing to their relative structural simplicity yet varied biological activities. The molecules are small, and differ based on their halogen, hydroxyl and methyl substitution patterns,[263] with fungal-derived rubrolides R and S (**112** and **113**) the only prenylated derivatives.[271] Rubrolide E (**114**) and its mono- and dibrominated derivatives (**115** and **116**) all showed moderate antibacterial activity against methicillin-resistant *Staphylococcus aureus* (MRSA) and *Staphylococcus epidermidis*, while their methylated derivatives **117–119** showed a marked decrease in activity.[266] The family has also shown anti-tumour activity, with rubrolide M (**120**) showing the most significant activity against P-388, A-549, HT-29 and MEL-28 cancer cell lines, with an EC_{50} of 1.2 μg mL^{-1} against each.[249] Finally, rubrolide O (**121**) (as a mixture of interconverting *E* and *Z* isomers) showed promise as an anti-inflammatory agent, by inhibiting superoxide production by human neutrophils *in vivo* with an IC_{50} of 35 μM.[250] A range of non–natural rubrolides have been synthesised, showing herbicidal potential,[272,273] and inhibition of biofilm formation against a range of Gram-positive and negative bacteria.[274,275]

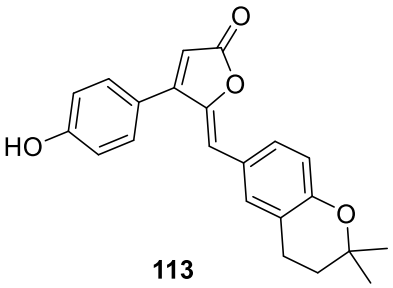


114 R=R_1=R_2=R_3=R_4=H
115 R=R_2=R_3=R_4=H R_1=Br
116 R=R_3=R_4=H R_1=R_2=Br
117 R=R_1=R_2=R_4=H R_3=CH_3
118 R=R_1=R_4=H R_2=Br R_3=CH_3
* **119** R=R_2=R_4=H R_1=Br R_3=CH_3
120 R=R_1=R_3=H R_2=Br R_4=Cl
* **121** R=R_1=R_2=Br R_3=H R_4=Cl

* = isolated as *E*/*Z* isomers

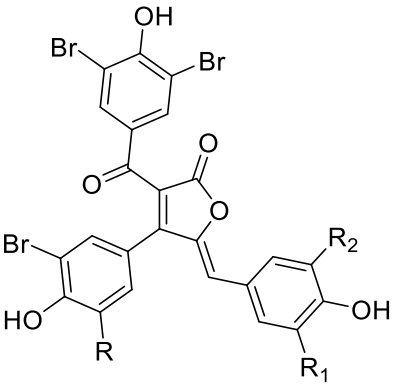


112

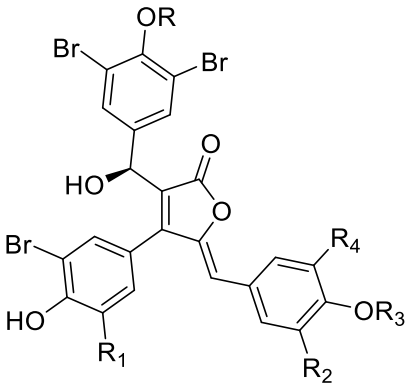


113

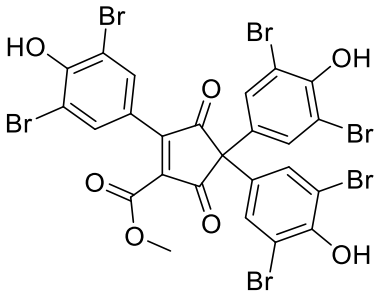
The cadiolides are structurally similar to the rubrolides; however, they bear a third aryl group at the C–2 position of the butenolide core. They were first reported from an Indonesian specimen of *Botryllus* sp.,[276] with cadiolides A and B (**122** and **123**) showing promising anti-viral activity against Japanese encephalitis virus (JEV).[270] Further congeners C–M have been isolated from Korean tunicates *Pseudodistoma antinboja*[267,268] and *Synoicum* sp.,[251] and are often co-isolated alongside rubrolides. The cadiolides have shown more potent antibacterial activity than their rubrolide counterparts, with cadiolides C, K and M (**124**–**126**) having MIC values against Gram-positive MRSA strains comparable to the marketed drug vancomycin,[267,268] and cadiolides E, G and H (**127**–**129**) showed potent activity against Gram-negative *Salmonella enterica* and *Proteus hauseri*.[251] This group of compounds were expanded further this year, with investigation of a Korean *Synoicum* sp. resulting in the isolation of the isocadiolides. These compounds are oxidised and rearranged to different extents, resulting in compounds centred around a cyclopentenedione (e.g. isocadiolide A **130**), dihydrofuran (e.g. isocadiolide F **131**) or a pyranone (isocadiolide H **132**), instead of a butenolide core.[277]



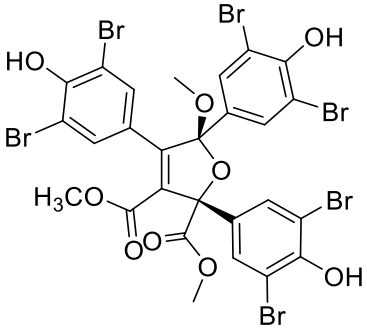
122 R=Br R_1=R_2=H
123 R=R_1=R_2=Br
124 R=R_2=H R_1=Br
127 R=H R_1=R_2=Br



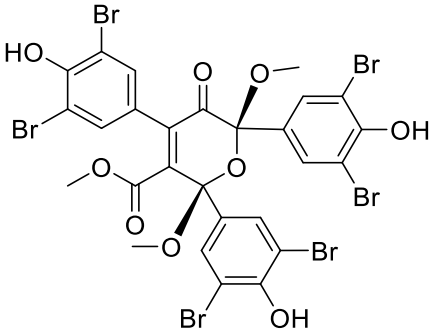
125 R=CH_3 R_1=R_2=Br R_3=R_4=H
126 R=R_3=H R_1=R_2=R_4=Br
128 R=CH_3 R_1=R_3=R_4=H R_2=Br
129 R=R_3=CH_3 R_1=Br R_2=R_4=H



130

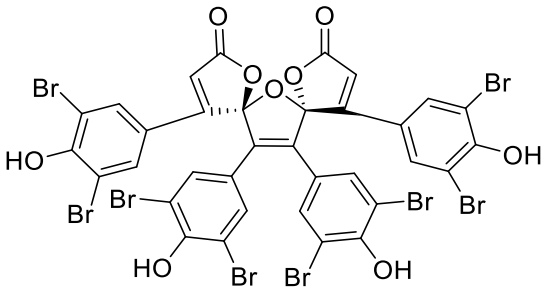


131

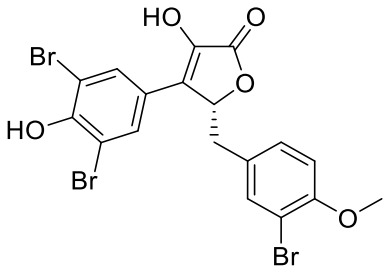


132

The prunolides and procerolides are two minor families within the polyaromatic butenolide class, also isolated from tunicates. The prunolides have been isolated from *S. prunum*, *S. blochmanni* and an unidentified *Synoicum* species,[249,252,270] alongside rubrolides in every case. Owing to their structural similarities, it has been suggested the prunolides form by the oxidative dimerization of a rubrolide precursor.[252] In general, the prunolides are less bioactive than their monomeric counterparts, however, prunolide A (**133**) has shown promising anti-viral activity against JEV at 1 μg mL^{-1}.[270] The procerolides were reported from a *Polycarpa procera* sample collected in Coffs Harbour, Australia, and bear one less degree of unsaturation in the furanone ring and oxygenation at C–2.[278] Procerolide A (**134**) showed potent anti-prion activity comparable to the positive control guanabenz in a yeast-based assay.



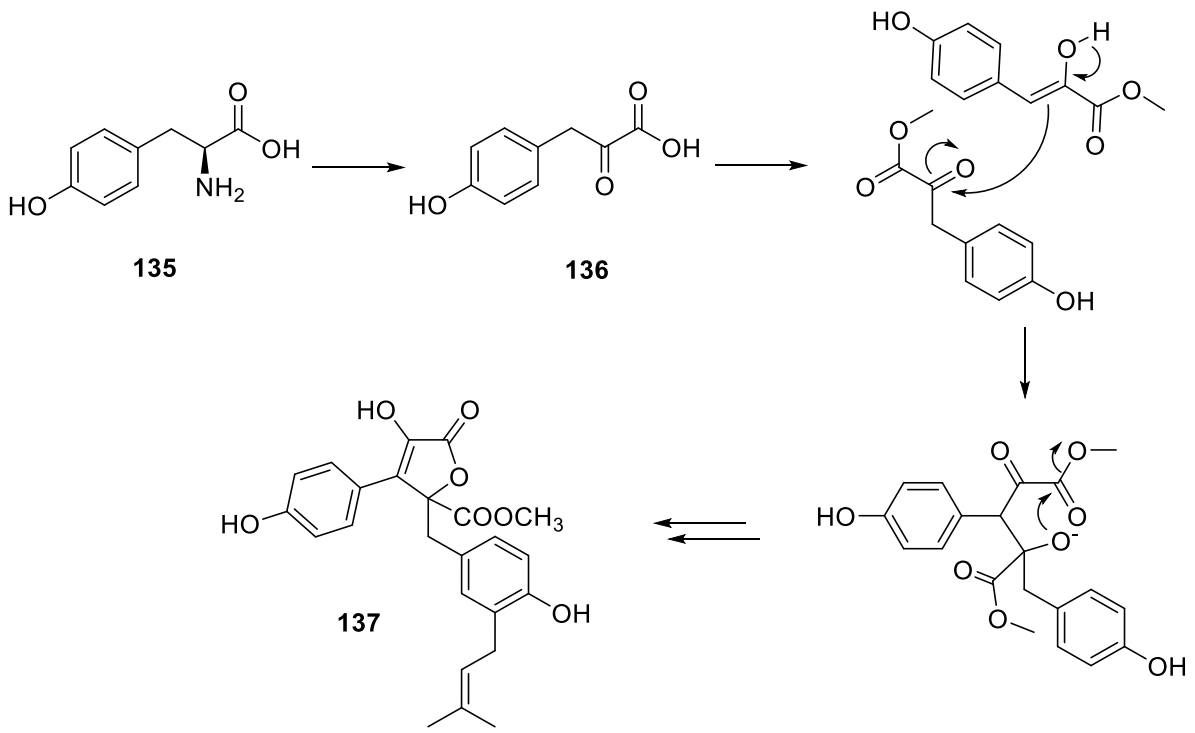
133



134

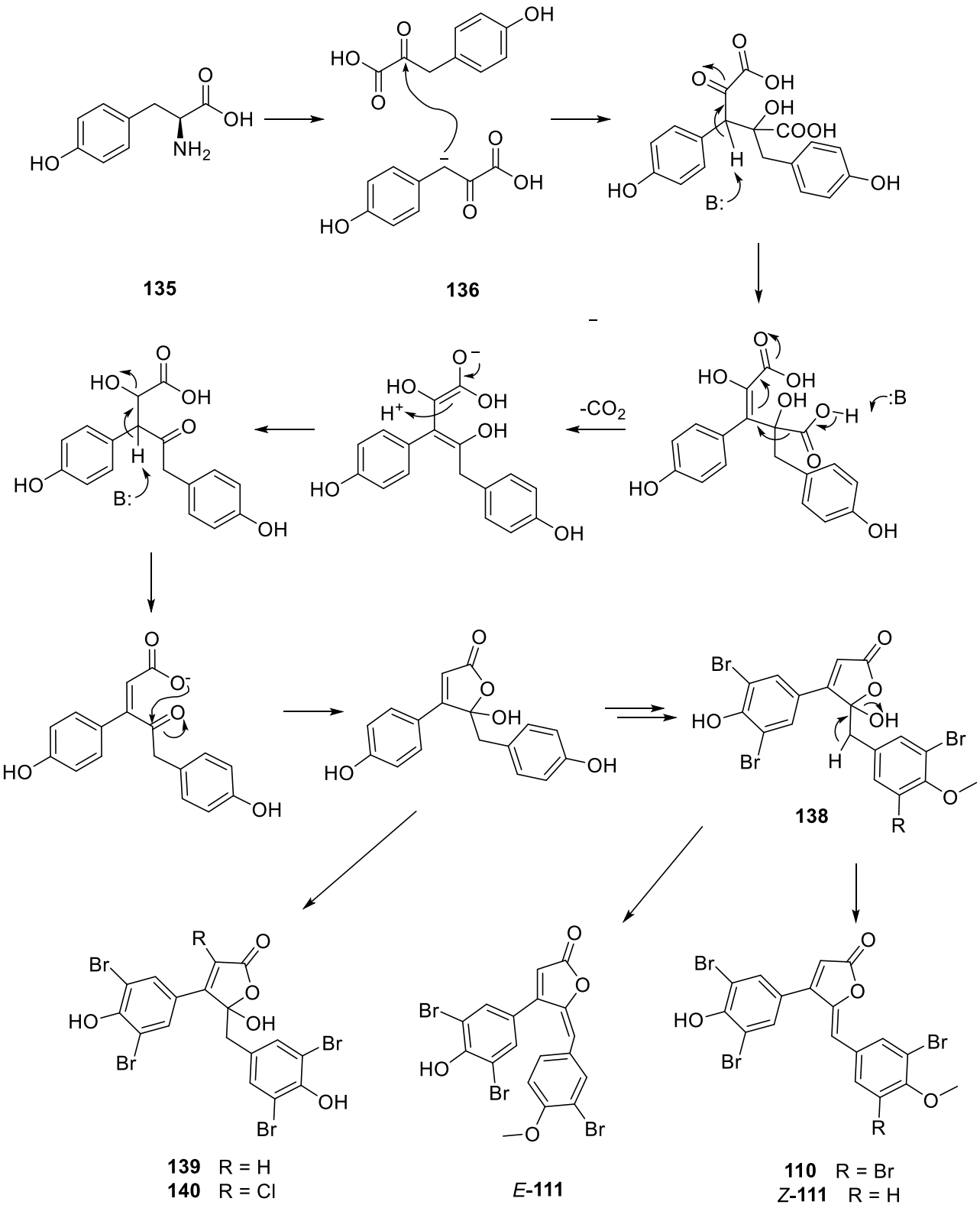
5.9 – Proposed Biogenesis

Although there are no reports formally investigating the biosynthesis of the marine-derived polyaromatic butenolides, there have been studies on structurally related bacterial-derived compounds. It is evident from the two phenol aromatic rings that the rubrolides are derived from the condensation of two molecules of tyrosine (**135**) or phenylalanine and subsequent halogenation/methylation via the α-keto acid **136**. The related butenolide (**137**), isolated from the fermentation broth of *Aspergillus terreus* var. *africanus* IFO 8835, was shown to be biosynthesised via this method (Scheme 5.3) through isotope incorporation studies.[279,280] This species was also the source of prenylated rubrolides R and S (**112** and **113**) which contain the same butenolide core as the tunicate derived rubrolides.[271]



Scheme 5.3 – Biosynthesis of related butenolide (**137**) as deduced by isotope incorporation studies.[279]

These related biosynthetic studies led Miao and Andersen to propose a biogenesis for the rubrolides (Scheme 5.4).[281] Unlike most tunicate amino acid-derived alkaloids, the rubrolides do not contain the amino nitrogen, therefore the first step is the formation of the *α*-keto acid **136** from **135**. Through an aldol-type condensation, two molecules form a reactive intermediate which then undergoes decarboxylation following a keto-enol tautomerisation of the *α*-ketone. This is then followed by a E1CB-dehydration and attack ketone carbonyl by the carboxylic acid oxygen. This mechanism forms the previously reported rubrolides G and H (**139** and **140**) after halogenation.[263] Miao and Andersen suggested this intermediate would then undergo another dehydration reaction to form rubrolide E (**114**), which is halogenated and methylated to form the remaining rubrolides, and by extension **110** and **111**.[281] The presence of interconverting *E* and *Z* isomers with the same halogenation patterns suggested that halogenation occurred before the dehydration reaction.



Scheme 5.4 – Proposed biogenesis of the rubrolides adapted from Miao and Andersen.[281]

5.10 – Conclusion

The tunicate *S. kuranui* was prioritised for investigation based on a molecular network, which revealed a constellation of rubrolide congeners. Although this class of metabolite has been thoroughly investigated, the network revealed two previously unreported precursor ions, which upon a mass–directed isolation led to the elucidation of previously unreported rubrolides T (**110**) and U (**111**). This work reflects how molecular networking can be used to probe for new compounds within a complex extract of previously reported metabolites. Based on the difference in *m/z* of the neighbouring precursor ions, the molecular formula was predicted, which streamlined the following NMR spectroscopy-based structure elucidation. LCMS-based molecular networking is a powerful tool for the identification of previously unreported congeners within a complex and enriched chemical extract.

Chapter 6 – *Didemnum ternerratum*

This chapter describes the chemical investigation of the Tongan tunicate Didemnum ternerratum. *This organism was prioritised based on the large and extensive constellation of nodes of varying masses within the molecular network generated in the mass spectrometric screen, along with many interesting signals detected in the ^{1}H NMR spectrum of its screening fractions. Alongside the known lamellarin K, this study resulted in the isolation of ten new and three known lamellarin sulfates, which is the first expansion of this class of compounds since 1999. The optical activity of four of the new compounds was investigated by ECD, and cytotoxic bioactivity was evaluated against the human colon carcinoma cell line HCT-116.*

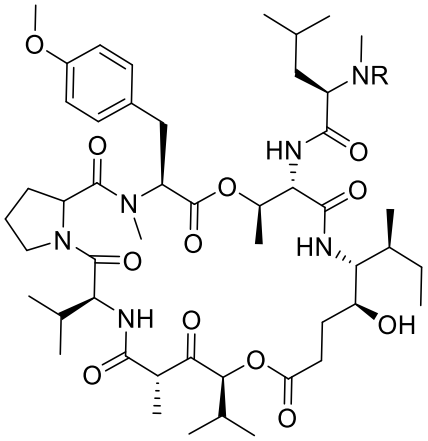
6.1 – *Didemnum* sp. Natural Products

Tunicates have been the source of many novel secondary metabolites. These have been isolated from a range of families including *Perophoridae* (ET-743) and *Polyclinidae* (aplidin, palmerolides), however the *Didemnidae* have been the most prolific source thus far. The *Didemnidae* is a highly diverse family of colonial ascidians, with at least 578 reported species representing 21% of all known ascidian species.[282] With species diversity often comes chemical diversity; ~36% of the known tunicate compounds have come from this family; therefore, didemnids are even more chemically diverse than predicted relative to other ascidians.[9] The high diversity of the family derives from genus *Didemnum*, with more than 200 species described alone,[282] which have been the source of ~215 (45%) of *Didemnidae* natural products.

The colonial nature of the family can be problematic for drug discovery, as often distinct colonies of varying sizes can merge and split, whilst maintaining their genetic/microbiome identity and thus appear morphologically identical yet chemically distinct.[283] This makes it difficult to collect enough material for traditional natural products research, let alone supply of compounds for future clinical development. This very issue is the basis of the chemical diversity present in the family, however, and leads to a 'natural' combinatorial library of metabolites with subtle structural variations between neighbouring colonies.[283] The didemnids host a range of bacterial species,[284] and the relationship with their major cyanobacterial symbiont, *Prochloron didemni*, is well studied.[285-288] Colonial ascidians are more prevalent across the substrate surface in tropical environments compared to solitary ascidians, therefore giving further precedence to study these organisms from the Pacific.[282]

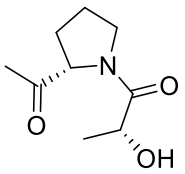
Many secondary metabolites isolated from the *Didemnidae* have significant biological activities. The namesake class, the didemnins, were first isolated from the Caribbean tunicate *Tridemnum solidum*, with didemnin B passing through phase I anti-tumour clinical trials.[289] Didemnins A–C (**141**–**143**) were first reported in 1981 after potent antiviral activity of a *T. solidum* extract was observed on a field expedition.[290] The depsipeptides inhibited both DNA and RNA viruses, showed potent cytotoxic activity against a range of cancer cell lines, and also exhibited immunosuppressive activities.[291] Didemnin B was the first marine metabolite to acquire attention from the synthetic community to supply preclinical and clinical trials, but was abandoned during early Phase II trials owing to significant toxicity.[292] Its dehydro derivative, plitidepsin (**14**), was mentioned in Chapter 1 as the second tunicate-derived drug, approved for treatment of multiple myeloma in 2018.[50] The pyridoacridone alkaloids **144** and **145**, related to ascididemin (**146**), were isolated from the NZ tunicate *Lissoclinum notti* and showed anti-TB activity against *M. tuberculosis* $H_{37}Rv$ but no significant cytotoxicity.[293] Ascididemnin, initially reported from an Okinawan *Didemnum sp.*,[294] has potent antineoplastic/cytotoxic activity, which is undesirable for anti–TB applications.[120] Didemnid compounds have also been the source of anti-fungal,[295] anti-malarial[296,297] and anti-oxidant[298] activities, to name a few.

For this work, the dark purple colonial ascidian *D. ternerratum*, collected in June 2016 using SCUBA from ʻEua, Kingdom of Tonga, was examined. This species was first described by Kott, 2001,[299] from samples sourced in Ellison Bay, South Australia. There are no previous reports of chemistry from this species.

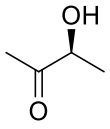


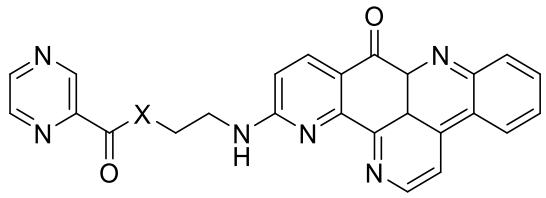
141 A R = H

142 B R =

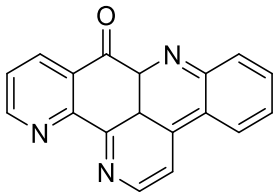


143 C R =





144 X = NH
145 X = O



146

6.2 – Prioritisation and Isolation Procedure

In Chapter 2, a full molecular networking analysis of several Pacific tunicate–based methanolic extracts using the GNPS platform was presented (Figure 2.11). Many individual constellations were observed, including the largest (most nodes) from only one source organism, coming from the *D. ternerratum* extract (coloured in red). This suggests that the extract contains a large number of molecules with different masses or retention times that are structurally similar, something that is often observed chemically from the family *Didemnidae*. It is worth noting that none of the nodes were annotated by the GNPS database, indicating a lack of similarity to known molecular scaffolds. The full tunicate sample (only 2 g) was extracted twice overnight in methanol (50 mL), and fractionated into 30%, 75% and 100% Me_2CO/H_2O samples (fractions A1–A3 respectively) using a PSDVB column. These samples were then analysed by ^1H NMR spectroscopy, with A1 and A2 showing numerous peaks associated with interesting secondary metabolites, such as coupled aromatic methines (6–8 ppm) and methoxy groups (intense 3–4 ppm), yet lacking many signals associated with primary metabolites. These data showed great potential for the discovery of new molecules, therefore *D. ternerratum* was prioritised for chemical investigation.

When the *m/z* values of the nodes were more accurately analysed, it was observed that the majority were in the range of *m/z* 400–700 and detected only in negative-ion mode. To further probe this group of molecules, an optimised molecular networking analysis was performed on fractions A1–A3, selecting only precursor ions in this range. One of the most intense ions observed, *m/*z 610.1024, was analysed over a range of collision energies from 10–100 eV, where 30 eV caused sufficient fragmentation without completely destroying the ion. Therefore, a fixed collision energy of 30 eV was used. The HPLC gradient time was also extended to ensure isomeric compounds with slightly different retention times would generate separate nodes (Figure 6.1). The general untargeted MS/MS method used within our group selects the five most intense ions per MS scan to fragment by CID, which are actively excluded in the

MS/MS ion selection for 0.3 mins once three spectra are obtained. By limiting the precursor ion *m/z* range, the LCMS is stopped from selecting ions not related to the group of molecules and focuses on minor species present that might otherwise be ignored in favour of more intense ions. As fractions A1–A3 crudely separated the extract based on polarity by increasing the solvent strength (increasing organic modifier Me_2CO), the fraction from which the precursor ion is selected from gives some insight into the relative polarity of the molecule.

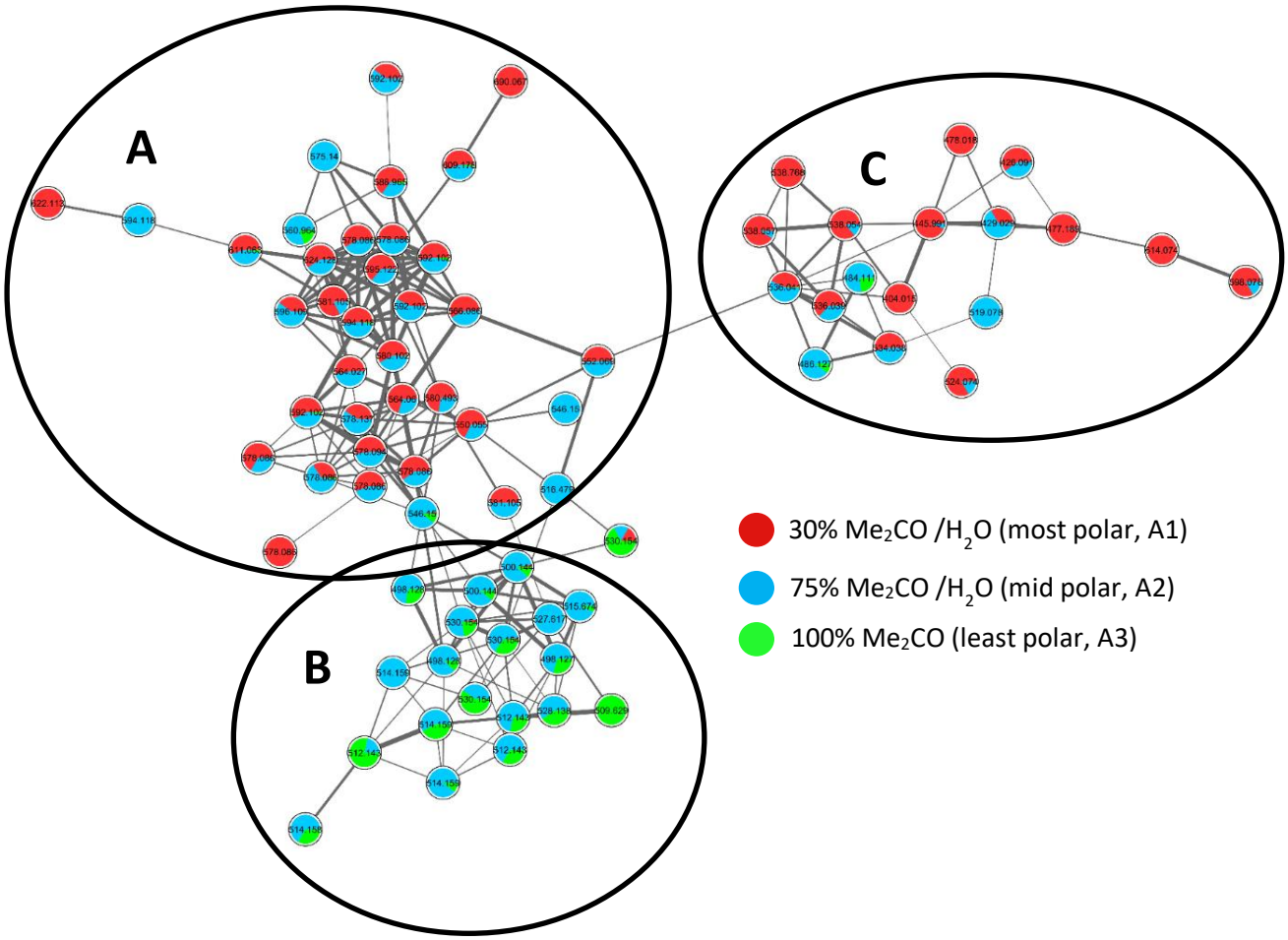
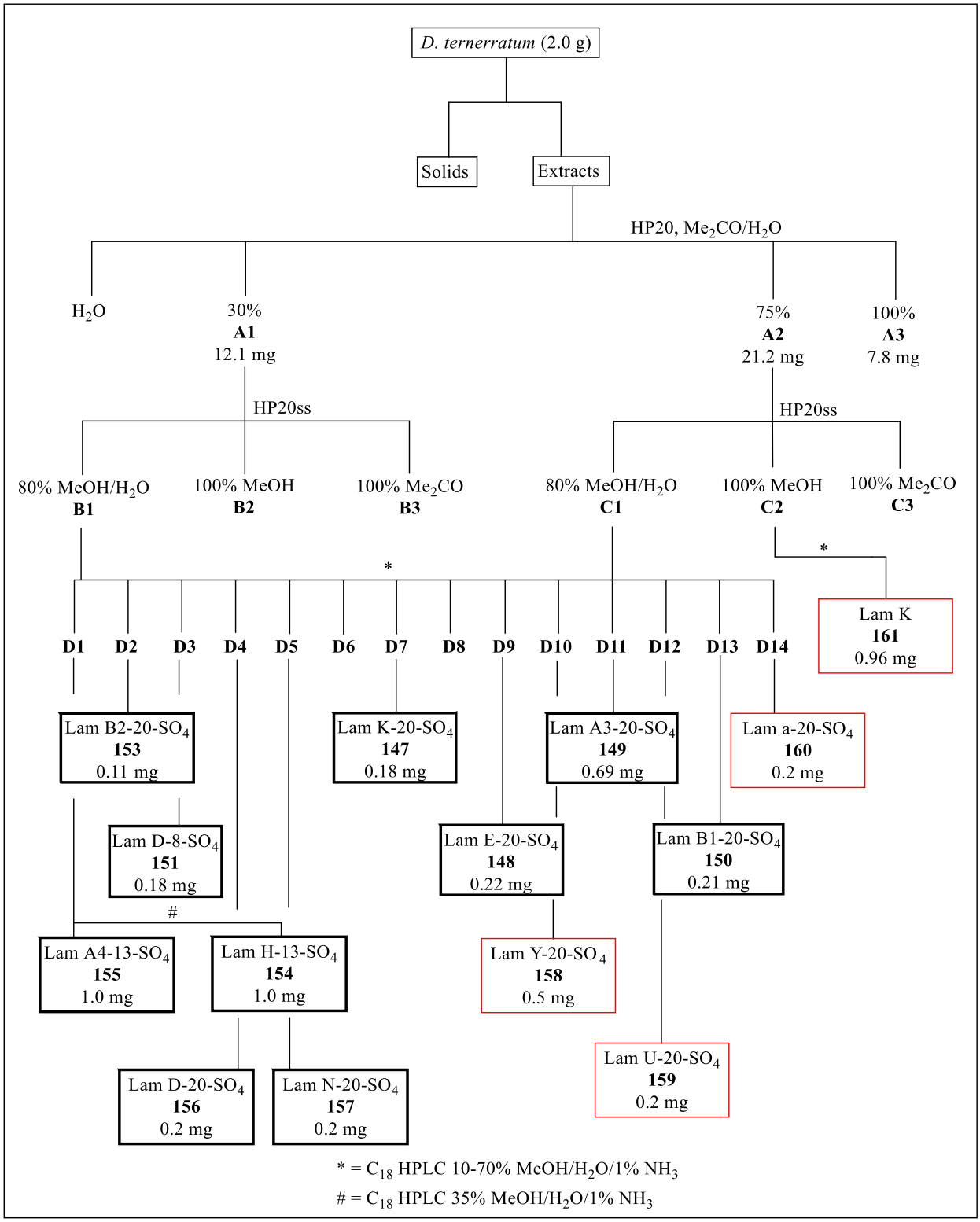


Figure 6.1 – GNPS molecular network of screening fractions of the *D. ternerratum* extract. Cluster A: sulfated lamellarins; Cluster B: non-sulfated lamellarins; Cluster C: non-methylated compounds (see 6.16 – Related compounds).

In the constellation, three key clusters of differing polarity were observed. These were subsequently annotated following the identification of several congeners (see below) as the more polar lamellarin sulfates (A) mainly present in fractions A1/A2, and the less polar methoxylated lamellarins (B) in A2/A3. A third very polar cluster (C) was observed from derivatives containing sulfates and phenolic groups. Clearly, there are a large number of nodes

and potential derivatives in this network, and it is unreasonable to try and isolate every one of these for NMR-based structure elucidation from a 2 g sample of tunicate. The extreme sensitivity of MS is ironically one of the key drawbacks of this type of analysis, as extremely small quantities of easily ionisable compounds are detected and can lead researchers in an unprofitable direction during purification. Therefore, it is pertinent to also monitor fractions by ^{1}H NMR spectroscopy throughout a MS-guided procedure. To increase the likelihood of isolating new compounds, the lamellarin sulfates were targeted for purification, as only nine previous congeners had been reported compared to the >70 known methoxylated-variants. These compounds are easily identifiable by MS/MS, due to the characteristic loss of *m/z* 80 (–SO_3) from collision-induced desulfonation.

Samples A1 and A2 were further purified using HP20 and two rounds of reversed-phase C_{18} HPLC (Scheme 6.1) to yield ten new lamellarin sulfates, and three that were previously reported. Two of the new metabolites isolated had methylation patterns that were not previously reported. As the majority of characters in the English and Greek alphabets have been exhausted in naming previous lamellarins, and the most recently isolated metabolites were termed A1–A6,[300] the two new methylation patterns have been termed B1 and B2.



Scheme 6.1 – Isolation procedure for compounds from *D. ternerratum*, collected in ʻEua, Kingdom of Tonga. Red boxes are previously reported compounds, black bold boxes denote compounds identified during the current study.

6.3 – Lamellarin K-20-sulfate

Lamellarin K-20-sulfate (**147**) was isolated as a brown film. HRESIMS analysis detected a deprotonated molecule at *m/z* 610.1024, consistent with the molecular formula $C_{29}H_{24}NO_{12}S^-$ (calcd. 610.1025). It is clear from this formula the molecule is highly unsaturated with an IHD of 18, and a ratio of hydrogen to "heavy atoms" of 24:43 (0.55). In NMR-based structure elucidation of natural products, "Crews' Rule" states that when the ratio of hydrogen atoms to heavy atoms (C, N, O, S) is < 1, the molecular structure is challenging to determine using conventional 1D and 2D ^{1}H and ^{13}C detected experiments.[301] Long-range correlations between ^{13}C and ^{1}H atoms (HMBC or similar experiment) become increasingly more important when protons are limited in a molecule, however in a molecule violating Crews' rule, these are often insufficient to deduce and assign the full structure. Therefore, based on its formula, this molecule violates Crews' rule and was predicted to be a challenge.

Due to the lack of quantity isolated, an IR spectrum was not obtained for the compound. The UV/vis trace showed maxima at 276, 311 and 339 (shoulder) nm, indicative of a highly conjugated aromatic compound, which is consistent with the large IHD value (Appendix 8). In the MS/MS spectrum (Figure 6.2), the dominant fragment ion at *m/z* 530.1607 (M–80), alongside the presence of sulfur, suggested a sulfate functionality was present. The ^{1}H NMR spectrum contained two mutually coupled methylenes (δ_H 3.03 (t, J = 6.9 Hz) and 4.63 (m)) and four methoxy singlets (δ_H 3.28, 3.32, 3.66, and 3.75) in the shielded region, and six resonances in the deshielded aromatic region. These consisted of three methine singlets (δ_H 6.39, 6.60, and 7.48) and three coupled methines (δ_H 6.89 (dd, J = 7.9, 2.0 Hz), 7.02 (d, J = 7.9 Hz), and 7.04 (d, J = 2.0 Hz)), the latter indicative of a 1,2,4-trisubstituted aromatic ring. Aside from the resonances associated with the shielded methylene (δ_C 21.2 and 41.8) and methoxy (δ_C 54.7, 55.0, 56.0, and 60.3) groups, the ^{13}C NMR spectrum showed 23 signals above 100 ppm in the aromatic region, with the peak at δ_C 154.4 assigned as a α,β-unsaturated ester carbonyl. In the HMBC spectrum, there were no correlations to this carbon from any proton resonances, therefore it was likely to be at least four bonds from any protonated carbon centre.

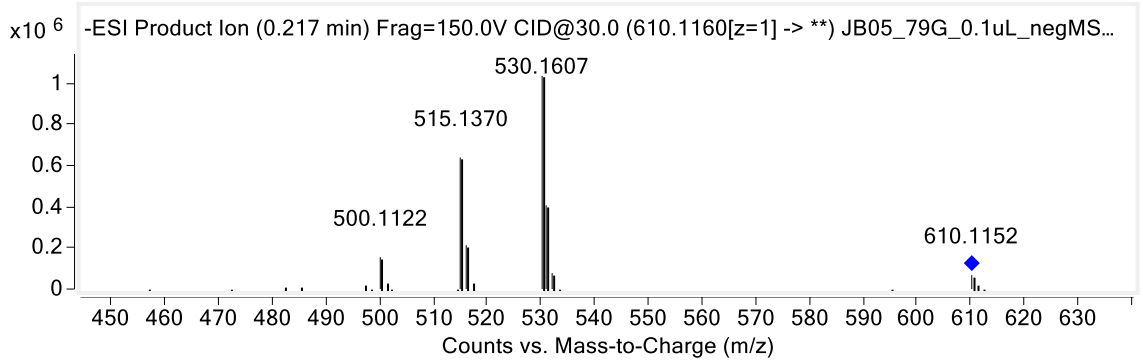


Figure 6.2 – MS/MS spectrum (CID = 30 eV) of **147**.

An aromatic methine resonating at δ_H 7.48 (H–19) showed HMBC correlations to C–17, C–18, C–20, and C–21, while singlet methine δ_H 6.60 (H–22) correlated with C–2, C–18, C–20, and C–21. The shared correlations and lack of splitting positioned these two protons *para* to one another, on a 1,2,4,5-tetrasubstituted benzene ring. The HMBC correlation between OCH_3–21 (δ_H 3.32) and C–21 placed this group on the ring, corroborated by a through-space NOE correlation (ROESY spectrum) to H–22. In similar molecules, the ^{1}H resonance *ortho* to a sulfate group is very deshielded compared to the phenolic homologue, and typically occurs in the δ_H 7.4−7.6 range.[302-304] The resonance of C–19 at δ_C 108.9 suggested it was flanked by two oxygen-bearing carbons, thus a sulfate was assigned at C–20 and substructure I was proposed (Figure 6.3).

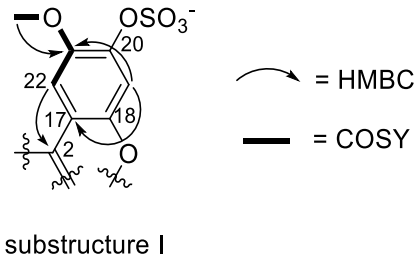


Figure 6.3 – Key COSY and HMBC correlations establishing **147** substructure I.

The COSY experiment typically shows correlations between protons two (geminal) or three (vicinal) bonds away, however in some instances, correlations can be observed up to five bonds away, particularly if both resonances have intense signals. This has been commonly observed with protons *ortho* to methoxy groups on aromatic rings[304] and was observed between singlet methine δ_H 6.39 (H–10) and the methoxy group at δ_H 3.28 (OMe–9). H–10 showed HMBC correlations to C–6a, C–8, C–9, C–10a, and C–10b, with an additional methoxy group δ_H 3.66 (OMe–8) also correlating to C–8. The less shielded methylene δ_H 4.63 (H_2–5) is nitrogen-bearing, based on its ^{13}C chemical shift (δ_C 41.8), and also from shared correlations to C–6a

and C–10b with H–10. The vicinal methylene δ_H 3.03 (H_2–6) also correlated in the HMBC to C–6a, along with the phenolic C–7. These data were suggestive of a trioxygenated-dihydroisoquinoline, substituted with two methoxy groups, thus substructure II was deduced (Figure 6.4).

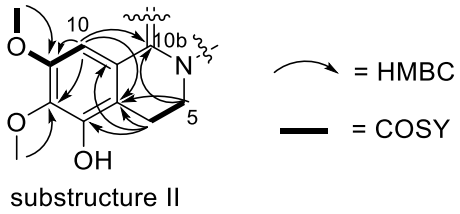


Figure 6.4 – Key COSY and HMBC correlations establishing **147** substructure II.

The final substructure, based around a 1,2,4-trisubstituted benzene ring scaffold, bore the final methoxy group attached to one of two oxygenated carbons C–13 (δ_C 148.6) or C–14 (δ_C 146.6). The methoxy group δ_H 3.28 (OMe–13) had COSY and ROESY correlations to the *meta*-coupled doublet H–12, thus must be on C–13 with a phenolic group at C–14. OCH_3–13 also shared a $^3J_{CH}$ HMBC correlation to C–13 with the H–15 *ortho*-coupled doublet, finalising substructure III (Figure 6.5).

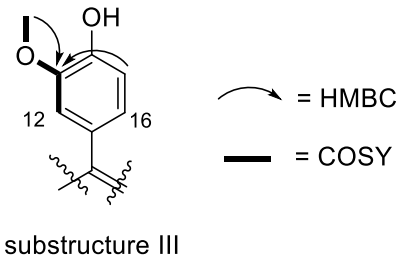


Figure 6.5 – Key COSY and HMBC correlations establishing **147** substructure III.

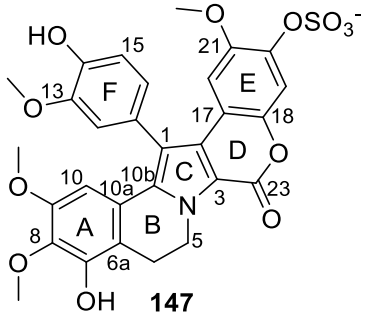
The three defined substructures accounted for $C_{27}H_{24}NO_{11}S$ of the atoms required by the molecular formula ($C_{29}H_{24}NO_{12}S^-$), yet the compounds full structure could not be established from the above analysis, a portrayal of the aforementioned "Crews' rule". The substructures were then used to search the MarinLit database[9] for previously isolated compounds with similar NMR data. Of the similar highly aromatic and methoxylated hits, the lamellarins were a clear match and comparison of the UV/vis and NMR data confirmed this. The C–23 carbonyl could not be directly positioned based on HMBC correlations alone, however, by analogy to reported data, it must be adjacent to the C–18 oxygen as an α,β-unsaturated ester. This particular methylation pattern is the same as lamellarin K, isolated from a *Didemnum* sp.[305] Comparison of the NMR data to the original isolated non-sulfated compound shows that, aside from the

proton shift of H–19 *ortho* to the sulfate, the data is concordant, thus **147** was identified as lamellarin K-20-sulfate.

Table 6.1 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for Lamellarin K-20-sulfate (**147**) in DMSO-d_6.

position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC	ROESY
1	116.0				
2	127.0				
3	113.1				
5	41.8	4.63 (m)	6	6, 6a, 10b	6
6	21.2	3.03 (t, 6.9)	5	5, 6a, 10a, 7	5
6a	114.3				
7	147.3				
8	136.5				
9	150.9				
10	100.9	6.39 (s)	OCH_3–9	6a, 8, 9, 10a,10b	OCH_3–9
10a	122.5				
10b	135.5				
11	125.3				
12	114.5	7.04 (d, 2.0)	15, OCH_3–13	1, 16, 14	OCH_3–13
13	148.6				
14	146.6				
15	116.4	7.02 (d, 7.9)	16	11, 13, 16	
16	123.3	6.89 (dd, 7.9, 2.0)	12, 15	12, 14, 15	
17	112.1				
18	142.3[a]				
19	108.9	7.48 (s)		17, 18, 20, 21	
20	144.6[a]				
21	146.8				
22	105.1	6.60 (s)	OCH_3–21	2, 18, 20, 21	OCH_3–21
23	154.2				
OCH_3–8	60.3	3.66 (s)		8	
OCH_3–9	54.7	3.28 (s)	10	9	10
OCH_3–13	56.0	3.75 (s)	12	13	12
OCH_3–21	55.0	3.32 (s)	22	21	22

[a] interchangeable



147

6.4 – Lamellarin E-20-sulfate

Lamellarin E-20-sulfate (**148**) was isolated as a white film. Again, HRESIMS analysis detected a deprotonated molecule at *m/z* 610.1048, and again consistent with the molecular formula $C_{29}H_{24}NO_{12}S^-$ (calcd. 610.1025), therefore **148** is an isomer of **147**. Unsurprisingly, the ^{1}H NMR spectrum had very similar NMR resonances to **147** including three of the four methoxy groups δ_H 3.28, 3.34, and 3.66 and the three singlet methine resonances δ_H 6.38, 6.67, and 7.47 with similar chemical shifts to those observed for **147**. The major differences in the spectrum were those assigned to *ring F*, with deshielded shifts for both H–15 (δ_H 7.17 (d, J = 8.1 Hz)) and the methoxy group at δ_H 3.83. These two resonances also showed long-range COSY and ROESY correlations to each other, therefore suggesting the methoxy group was at C–14. This was further corroborated by shared $^3J_{CH}$ HMBC correlations from both δ_H 3.83 and δ_H 6.89 (H–12) to the signal for C–14. Therefore, substructure IV of **148** (Figure 6.6) was deduced.

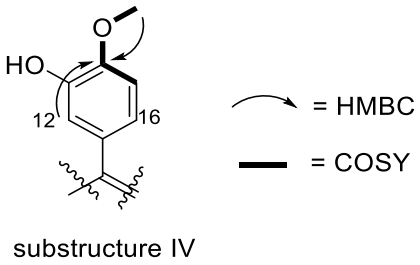


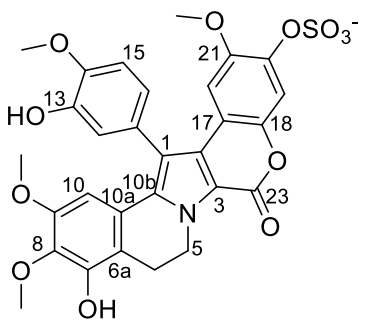
Figure 6.6 – Key COSY and HMBC correlations establishing **148** substructure IV.

The remaining NMR data was consistent with substructures I and II of **147** (Figure 6.3 and 6.4) therefore the methoxylation pattern is the same as parent lamellarin E, and **148** is lamellarin E-20-sulfate. Comparison of the reported NMR data is consistent with this assignment, as the only difference is again the ^{1}H NMR resonance of H–19 *ortho* to the sulfate.[306]

Table 6.2 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for Lamellarin E-20-sulfate (**148**) in DMSO-d_6.

position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC	ROESY
1	n.d				
2	126.7				
3	n.d				
5	41.7	4.64 (m)	6	6a, 10b	
6	21.2	3.03 (t, 7.6)	5	6a	
6a	114.4				
7	147.3				
8	136.4				
9	150.8				
10	100.9	6.38 (s)	OCH_3–9	6a, 8, 9, 10b	OCH_3–9
10a	122.3				
10b	135.1				
11	127.1				
12	117.7	6.89 (d, 2.1)	16	14, 16	
13	147.5				
14	147.7				
15	113.5	7.17 (d, 8.1)	16, OCH_3–14	11, 14	OCH_3–14
16	121.5	6.90 (dd, 8.0, 2.1)	12, 15		
17	n.d				
18	142.2				
19	108.9	7.47 (s)		18, 20, 21	
20	144.5				
21	146.7				
22	104.9	6.67 (s)	OCH_3–21	2, 18, 20, 21	OCH_3–21
23	154.2				
OCH_3–8	60.3	3.66 (s)		8	
OCH_3–9	54.7	3.28 (s)	10	9	10
OCH_3–14	56.0	3.83 (s)	15	14	15
OCH_3–21	55.0	3.34 (s)	22	21	22

n.d. = not detected.



148

From these two isomers, a comparison of the effect of the *ring F* methyl groups on adjacent ^{1}H NMR chemical shifts can be drawn. When going from **147** to **148**, a downfield shift of the methoxy proton resonance was observed when positioned at C–14 rather than C–13 (both without substitution at C–15). When the ^{1}H NMR chemical shifts of all reported closed-hexacyclic lamellarins containing a *ring F* methoxy and one phenol are tabulated (Table 6.3), this trend is also seen.[264,300,302,305-307] This guideline can be used in future structure elucidations simply from ^{1}H NMR data, especially if long-range COSY correlations are not observed or NOESY/ROESY data cannot be obtained; at the very least it provides an extra piece of evidence. Comparison of $^{2}J_{CH}$ vs $^{3}J_{CH}$ correlations in the HMBC experiment of the protons in *ring F* is another method used for this placement, however it is much less reliable than these methods.

Lamellarin	*C–13*	*C–14*
B	3.76	OH
C	3.75	OH
E	OH	3.83
G	OH	3.86
K	3.77	OH
L	OH	3.82
M	3.76	OH
T	OH	3.84
U	OH	3.82
V	OH	3.82/3.83
W	OH	3.85
X	OH	3.85
β	OH	3.88
A1	3.73	OH

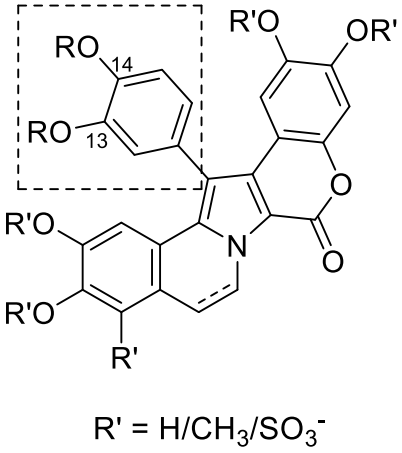


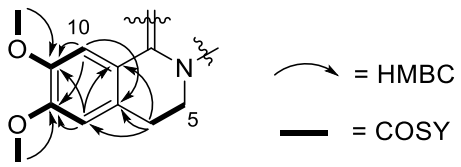
Table 6.3 – Ring F (in blue) methoxy ^{1}H NMR chemical shifts of various lamellarins.

6.5 – Lamellarin A3-20-sulfate

Lamellarin A3-20-sulfate (**149**) was isolated as a white solid. HRESIMS analysis showed a deprotonated molecule at *m/z* 594.1103, consistent with the molecular formula $C_{29}H_{24}NO_{11}S^-$ (calcd. 594.1076) and indicating one less oxygen atom than **147** or **148**. Based on comparison of the MS and UV data, **149** was clearly a lamellarin sulfate. The ^{1}H NMR spectrum showed four methoxy groups, four aromatic methines and signals at δ_H 3.12 (t, J = 6.9 Hz) and 4.66 (m) of methylene groups. The major difference when compared to **147** and **148** was the substitution of one phenolic group replaced with a hydrogen, validated by the presence of an extra aromatic methine at δ_H 6.99.

The ^{1}H NMR data suggested that the sulfate group was yet again attached to C–20, as H–19 was the most deshielded resonance at δ_H 7.48 and the remainder of the data for *ring E* was the same as **147** and **148** (substructure I, Figure 6.3). The characteristic aromatic methine coupling pattern of a 1,2,4-trisubstituted benzene ring was again observed, with long range COSY and ROESY correlations, and OCH_3–13 ^{1}H chemical shift δ_H 3.74 suggesting the same *ring F* substitution pattern as **147** (substructure III, Figure 6.5).

The major difference was observed in the NMR data of *ring A*, where the extra aromatic methine singlet was observed at δ_H 6.99. This resonance showed HMBC correlations to C–6, C–8, C–9 and C–10a, and therefore was assigned to a proton at C–7 in place of the phenol functionality of **147** and **148**. H–10 shared HMBC correlations with C–8 and C–9, and as both are singlets, suggested the two are arranged *para* to each other. The assignment of C–7 was also verified by an HMBC correlation to methylene H_2–6, while H–7 also showed long range COSY and ROESY correlations to OCH_3–8 (δ_H 3.25). This evidence suggested that substructure V of **149** is a dimethoxy-dihydroisoquinoline (Figure 6.7).



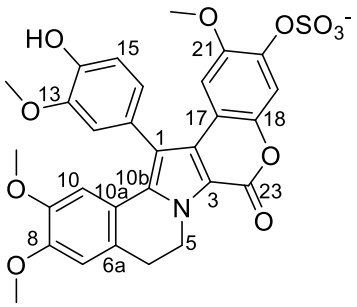
substructure V

Figure 6.7 – Key COSY and HMBC correlations establishing **149** substructure V.

This methylation pattern is the same as lamellarin A3, with all of the remaining ^{13}C and ^{1}H NMR data consistent with that previously reported,[300] aside from the H–19 resonance, giving **149** as lamellarin A3-20-sulfate.

Table 6.4 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for Lamellarin A3-20-sulfate (**149**) in DMSO-d_6.

position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC	ROESY
1	115.1				
2	127.0				
3	109.6				
5	42.0	4.66 (m)	6		6
6	27.7	3.12 (t, 6.9)	5	6a, 7, 10a	5, 7
6a	126.9				
7	111.8	6.99 (s)	OCH_3–8	6, 8, 9, 10a	6, OCH_3–8
8	146.9				
9	148.9				
10	108.6	6.71 (s)	OCH_3–9	6a, 8, 9, 10b	OCH_3–9
10a	119.3				
10b	135.5				
11	125.2				
12	114.4	7.04 (d, 2.0)	16, OCH_3–13	1, 14, 16	OCH_3–13
13	148.5				
14	146.6				
15	116.3	7.02 (d, 8.0)	16	11, 13	
16	123.3	6.91 (dd, 7.9, 2.0)	12, 15	12	
17	112.1				
18	142.3				
19	108.9	7.48 (s)		17, 20, 21	
20	144.6				
21	146.8				
22	105.1	6.65 (s)	OCH_3–21	2, 18, 20, 21	OCH_3–21
23	154.2				
OCH_3–8	54.5	3.25 (s)	7	8	7
OCH_3–9	55.6	3.77 (s)	10	9	10
OCH_3–13	56.0	3.74 (s)	12	13	12
OCH_3–21	55.0	3.34 (s)	22	21	22



149

6.6 – Lamellarin B1-20-sulfate

Lamellarin B1-20-sulfate (**150**) was isolated as a white solid. HRESIMS analysis detected the deprotonated molecule at *m/z* 592.0948, consistent with the molecular formula $C_{29}H_{22}NO_{11}S^-$ (calcd. 592.0919); with two less hydrogens than **149**. There are two core structures of closed hexacyclic lamellarins, those with 18 degrees of unsaturation as found in **147**–**149**, and those with 19 where the C–5 to C–6 bond is oxidised to form an isoquinoline motif. This results in a distinct change in chromophore (Figure 6.8), which was observed in the UV/vis spectrum, particularly the addition of two longer wavelength maxima at 367 and 386 nm. The other distinct alteration is the replacement of the methylene signals with two deshielded, mutually coupled methines (δ_H 7.35 (d, J = 7.3 Hz) and 9.10 (d, J = 7.3 Hz)) in the ^{1}H NMR spectrum.

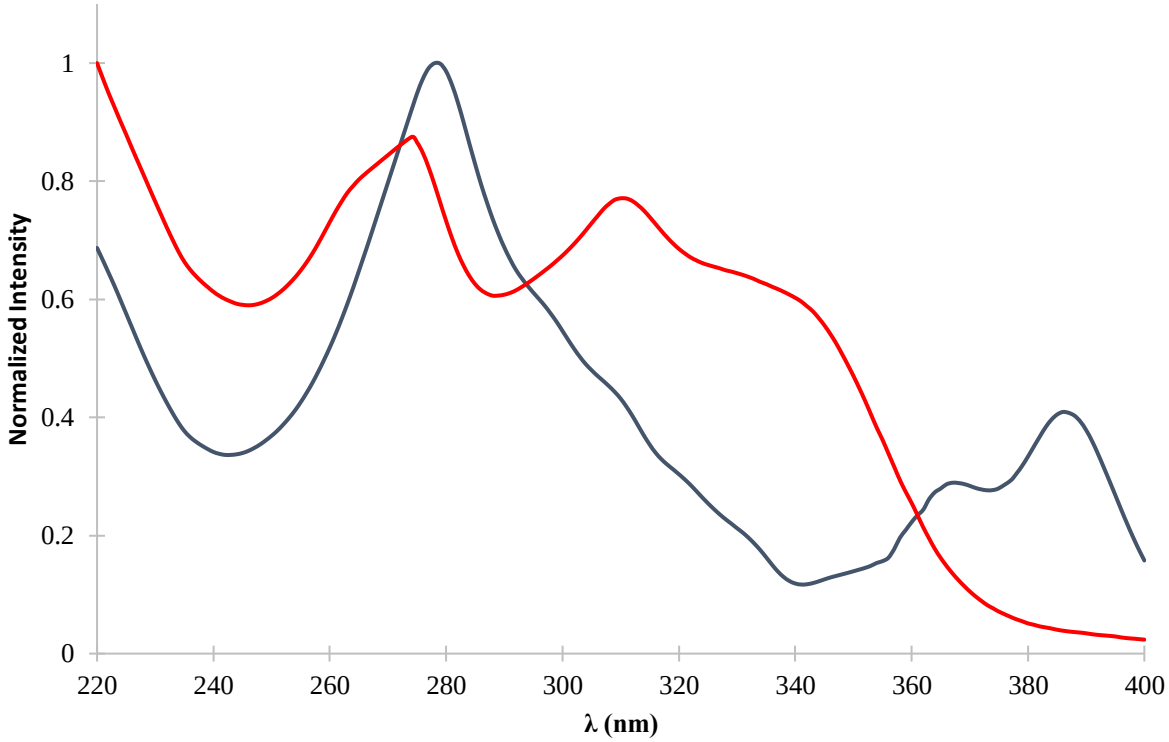
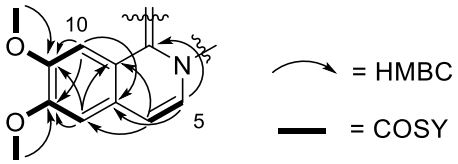


Figure 6.8 – Normalised UV/visible spectra of **149** (red) and **150** (blue) in MeOH.

Aside from the C–5 to C–6 double bond, the NMR data suggested that the remainder of the molecule is identical to **149**. The very deshielded doublet methine at δ_H 9.10 had HMBC correlations to C–6a and C–10b and was therefore assigned to H–5, consistent with a nitrogen-bound alkene causing an extreme deshielding effect. Alkene partner H–6 (δ_H 7.35) had HMBC correlations to C–7 and C–10a, consistent with the structure proposed (Figure 6.9).



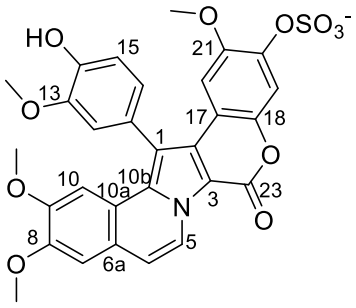
substructure VI

Figure 6.9 – Key COSY and HMBC correlations establishing **150** substructure VI.

This methylation pattern with an isoquinoline core has yet to be reported in the literature, therefore the molecule was termed lamellarin B1-20-sulfate.

Table 6.5 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for Lamellarin B1-20-sulfate (**150**) in DMSO-d_6.

position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC	ROESY
1	111.7				
2	128.4				
3	107.0				
5	122.4	9.1 (d, 7.3)	6	6a, 10b	
6	113.0	7.35 (d, 7.3)	5	7, 10a	
6a	124.4				
7	108.3	7.44 (s)	OCH_3–8	6, 8, 9, 10a	OCH_3–8
8	150.0				
9	149.0				
10	104.9	7.17 (s)	OCH_3–9	6a, 8, 9, 10b	OCH_3–9
10a	118.5				
10b	133.7				
11	125.0				
12	114.9	7.16 (d, 2.0)	16, OCH_3–13	14, 16, 1	OCH_3–13
13	148.7				
14	146.9				
15	116.5	7.11 (d, 8.0)	16	11, 13	
16	123.7	7.03 (dd, 8.0, 2.0)	12, 15		
17	111.7				
18	143.2				
19	108.9	7.57 (s)		17, 18, 20, 21	
20	145.1				
21	146.7				
22	105.8	6.77 (s)	OCH_3–21	2, 18, 20, 21	OCH_3–21
23	154.3				
OCH_3–8	55.7	3.89 (s)	7	8	7
OCH_3–9	54.5	3.37 (s)	10	9	10
OCH_3–13	56.0	3.76 (s)	12	13	12
OCH_3–21	55.0	3.36 (s)	22	21	22

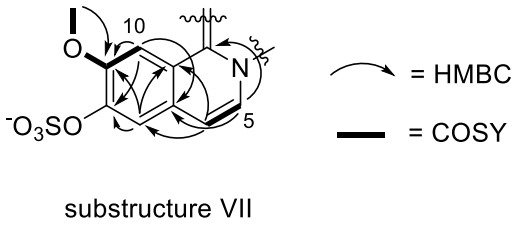


150

6.7 – Lamellarin D-8-sulfate

Lamellarin D-8-sulfate (**151**) was isolated as a yellow solid and HRESIMS analysis indicated a deprotonated molecule at *m/z* 578.0783, suitable for the molecular formula $C_{28}H_{20}NO_{11}S^-$ (calcd. 578.0763). The ^{1}H NMR showed signals for two coupled doublets δ_H 7.28 (1H, d, J = 7.3 Hz) and 9.03 (1H, d, J = 7.3 Hz) which, in conjunction with the UV/vis data, suggested a C–5 to C–6 double bond and isoquinoline structure as with **150**. Also observed were three methoxy signals, four aromatic methine singlets and the three mutually coupled aromatic methines with the characteristic coupling pattern of *ring F*. The *meta*-coupled doublet H–12 (δ_H 7.16) showed long range COSY and ROESY correlations to OCH_3–13 (δ_H 3.77), which, along with the chemical shift, suggested methoxy substitution at C–13 (Table 6.3), thus **151** contains the same substructure III as **147**, **149** and **150** (Figure 6.5).

This molecule had all the same spectral characteristics as the other lamellarin sulfates isolated, such as the prominent M–80 fragment ion in the MS/MS at *m/z* 498.1190 and requirement of sulfur from the molecular formula. However, no ^{1}H NMR resonance was detected in the typical spectral range of δ_H 7.4−7.6 for the proton *ortho* to the sulfate group; instead, a new aromatic methine singlet at δ_H 7.89 was observed. The HSQC spectrum showed a correlation to a carbon resonating at δ_C 117.2, in stark contrast to the C–19 signal that typically resonates at δ_C ~109 in **147**–**149**. The resonance at δ_H 7.89 showed HMBC correlations to carbons assigned as C–6, C–8, C–9 and C–10a, while H–10 (δ_H 7.17) correlated with C–6a, C–8, C–9 and C–10b; therefore, these two singlets must be *para* on *ring A*. This then placed the sulfate group at C–8, similar to the only other reported lamellarin sulfated at a carbon other than C–20, lamellarin G-8-sulfate (**152**).[304] The most deshielded proton resonance (H–7) of **152** was observed in the typical range at 7.40 ppm, however this molecule has a dihydroisoquinoline core and so cannot be directly compared with **151**. Finally, a methoxy group was placed at C–9, based on long range COSY and ROESY correlations to H–8, therefore **151** substructure VII was deduced (Figure 6.10).



substructure VII

Figure 6.10 – Key COSY and HMBC correlations establishing **151** substructure VII.

One methoxy group remained to be assigned, which showed correlations in both the COSY and ROESY spectra to δ_H 6.72. This resonance could feasibly be from H–19 or H–22, however it in turn showed a HMBC correlation to C–2, and therefore must be H–22, thus defining substructure VIII (Figure 6.11).

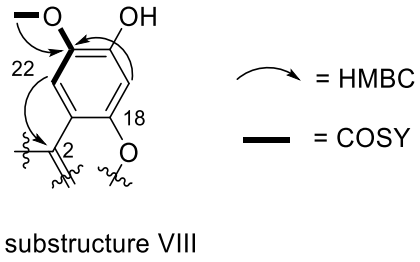


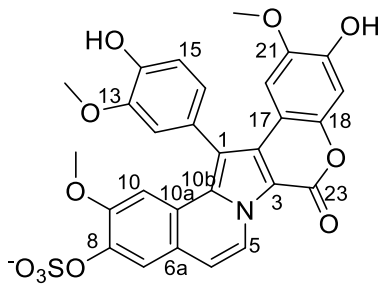
Figure 6.11 – Key COSY and HMBC correlations establishing **151** substructure VIII.

The data for this methylation pattern is consistent with lamellarin D, and therefore **151** is lamellarin D-8-sulfate. As the original isolation publication used Me_2CO-d_6 as the NMR solvent and **151** was not readily soluble in this solvent,[264] direct comparison of the chemical shifts with our data (DMSO-d_6) was not possible.

Table 6.6 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for Lamellarin D-8-sulfate (**151**) in DMSO-d_6.

position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC	ROESY
1	111.5				
2	129.0				
3	n.d				
5	122	9.03 (d, 7.3)	6	6, 6a, 10b	
6	112.8	7.28 (d, 7.3)	5	5, 10a	
6a	123.7				
7	117.2	7.89 (s)		6, 8, 9, 10a	
8	143.8				
9	150.3				
10	105.5	7.17 (s)	OCH_3–9	6a, 8, 9, 10b	OCH_3–9
10a	120.3				
10b	133.5				
11	125.2				
12	114.9	7.16 (d, 2.1)	16, OCH_3–13	1, 11, 13, 14	OCH_3–13
13	148.7				
14	146.9				
15	116.4	7.11 (d, 7.9)	16	11, 13	
16	123.6	7.03 (dd, 7.9, 2.1)	12, 15	12	
17	112.1				
18	144.6				
19	103.6	6.87 (s)		20, 21	
20	146.3				
21	144.7				
22	105.8	6.72 (s)	OCH_3–21	2, 18, 20	OCH_3–21
23	154.5				
OCH_3–9	54.5	3.35 (s)	10	9	10
OCH_3–13	56.0	3.77 (s)	12	13	12
OCH_3–21	55.1	3.38 (s)	22	21	22

n.d. = not detected.

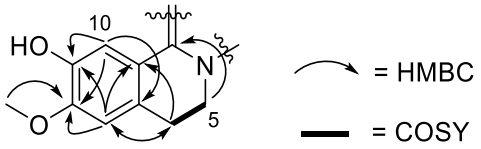


151

6.8 – Lamellarin B2-20-sulfate

Lamellarin B2-20-sulfate (**153**) was isolated as a yellow solid with HRESIMS analysis detecting a deprotonated molecule at *m/z* 566.0758 indicative of the molecular formula $C_{27}H_{20}NO_{11}S^-$ (calcd. 566.0763). The UV/vis spectrum and presence of two methylene resonances (δ_H 3.05 (t, J = 6.7 Hz) and 4.63 (t, J = 6.7 Hz)) in the ^{1}H NMR spectrum implied a dihydroisoquinoline core, with resonances associated with two methoxy groups, four aromatic singlet methines and the three resonances of the *ring F* protons also present in the spectrum. Shared $^3J_{CH}$ HMBC correlations from δ_H 3.88 (OMe–14), 6.75 (H–16), and 6.77 (H–12) to C–14 indicated a methoxy group was attached at C–14 and a phenol group at C–13 (Table 6.3). This was further supported by a ROESY correlation between δ_H 7.09 (H–15) and OCH_3–14, thus sharing the same ring structure with **148** (Figure 6.6).

The other methoxy group, δ_H 3.78, showed a ROESY correlation to δ_H 6.93, which was assigned to H–7 based on its reciprocal $^3J_{CH}$ HMBC correlation to methylene C–6. This placed it at C–8, with further evidence from the shared correlation to C–8 in the HMBC spectrum from OCH_3–8 and H–10 (δ_H 6.49). Therefore, **153** substructure IX was deduced (Figure 6.12).



substructure IX

Figure 6.12 – Key COSY and HMBC correlations establishing **153** substructure IX.

Based on the molecular formula, *ring C* must have phenol and sulfate functionalities only, as both methoxy groups had already been assigned around the molecule. The final two singlet aromatic methines, δ_H 7.21 and 6.50, were assigned to H–19 and H–22, respectively, based on a $^3J_{CH}$ HMBC correlation between H–22 and C–2 (δ_C 125.8). All previously isolated lamellarin sulfate molecules have had a methoxy group at C–21, and in all cases, if the sulfate is at C–20 its *ortho* proton partner is shifted to δ_H ~7.5 (Figure 6.13). This was not observed in the NMR spectrum of **153**; however, as the proton is more shielded by the C–21 phenol. Strong evidence to place the sulfate group at C–20 was also observed when comparing NMR data to related compound lamellarin A1 with two hydroxyls on *ring F*.[300] The chemical shifts of CH–19 reported are less shielded for **153** (δ_H 7.24 and δ_C 109.8 vs δ_H 6.75 and δ_C 103.4), whereas the shifts of CH–22 are much more similar (δ_H 6.50 and δ_C 109.3 vs δ_H 6.65 and δ_C 108.6). Therefore, **153** is the first reported lamellarin with a hydroxyl *ortho* to the sulfate group and

gives insight into how this substitution affects the NMR data. This methylation pattern has not been observed before, therefore **153** was termed lamellarin B2-20-sulfate.

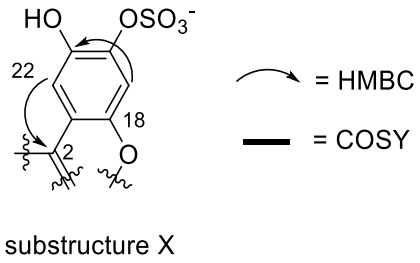
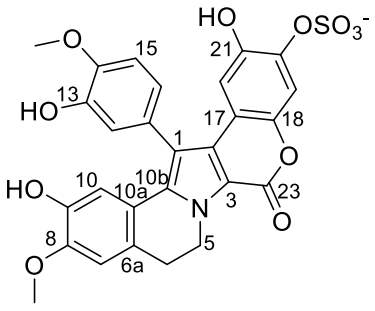


Figure 6.13 – Key COSY and HMBC correlations establishing **153** substructure X.

Table 6.7 – ^{13}C (200 MHz) and ^{1}H (800 MHz) NMR data for Lamellarin B2-20-sulfate (**153**) in DMSO-d_6.

position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC	ROESY
1	115.3				
2	125.8				
3	113.1				
5	41.8	4.63 (t, 6.7)	6	6, 6a, 10b	6
6	27.7	3.05 (t, 6.7)	5	5, 6a, 7, 10a	5, 7
6a	125.5				
7	111.5	6.93 (s)		6, 8, 9, 10a	OCH_3–8
8	148.0				
9	144.7				
10	112.6	6.49 (s)		6a, 8, 10b	
10a	119.6				
10b	135.6				
11	126.7				
12	117.1	6.77 (d, 2.0)	16	12, 16	
13	147.0				
14	147.6				
15	112.7	7.09 (d, 8.2)	16	11, 13	OCH_3–14
16	120.7	6.75 (dd, 8.2, 2.0)	12, 15	12, 14	
17	113.5				
18	144.2				
19	109.8	7.24 (s)		17, 20, 21	
20	140.7				
21	144.4				
22	109.3	6.50 (s)		2, 20, 21	
23	154.2				
OCH_3–8	55.6	3.78 (s)		8	7
OCH_3–14	55.4	3.88 (s)		14	15



153

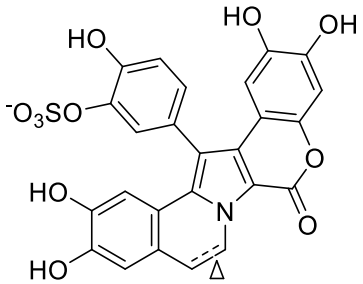
6.9 – Lamellarin H-13-sulfate and Lamellarin A4-13-sulfate

The most polar sulfated lamellarin sample was, based upon NMR and MS analysis, initially isolated as a fraction containing both C–5 to C–6 saturated and unsaturated derivatives with no methyl groups (lamellarin A4 and lamellarin H skeletons, respectively). Upon HPLC purification, lamellarin H-13-sulfate (**154**) was stable enough to acquire ^{1}H NMR, COSY, HSQC and MS data, however it degraded before HMBC and ^{13}C NMR spectra could be acquired on the mass limited sample, whereas dihydro derivative lamellarin A4-13-sulfate (**155**) rapidly degraded upon purification. As a comprehensive NMR analysis of the two-component mixture was performed, it has been used for structural annotation.

Lamellarin H-13-sulfate (**154**) was purified as a yellow solid from the 30% Me_2CO/H_2O PSDVB column elution. HRESIMS analysis detected a deprotonated molecule *m/z* 536.0316 consistent with the molecular formula $C_{25}H_{14}NO_{11}S^-$ (calcd. 536.0293). The ^{1}H NMR spectrum (CD_3OD) showed signals for two mutually coupled doublets δ_H 7.05 (J = 7.3 Hz) and 9.07 (J = 7.3 Hz), which in conjunction with the UV/vis spectrum indicated the presence of a C–5 to C–6 double bond. Additionally, three aromatic methines with the characteristic coupling pattern of *ring F* and four singlet aromatic methines were observed, while significantly no methoxy resonances were detected (Table 6.8). From this data, it was clear that the molecule must be similar to lamellarin H,[306] however the deshielding of the *meta* coupled doublet (H–12) to δ_H 7.52, as opposed to the C–20 or C–9 singlet in all other naturally isolated examples, suggested that the sulfate group was present at C–13. Lamellarin α-13-sulfate has been synthesised,[308] with the corresponding deshielding of H–12 observed at δ_H 7.71 (DMSO-d_6). As the two other aromatic spin systems appeared as 1,2,4,5-tetrasubstituted, this NMR data was consistent with the structure of lamellarin H-13-sulfate (**154**), representing the first lamellarin sulfated at a position other than C–8 or C–20. The mixture of **154** and **155** rapidly degraded to a sample containing only **154**.

Lamellarin A4-13-sulfate (**155**) was semi-purified as a yellow solid from the 30% Me_2CO/H_2O PSDVB column elution, as a two-component mixture deduced by MS and ^{1}H NMR data. HRESIMS analysis detected a deprotonated molecule *m/z* 538.0474 consistent with the molecular formula $C_{25}H_{16}NO_{11}S^-$ (calcd. 538.0450). As NMR data was obtained for purified **154**, the unassigned remaining peaks were used for structure elucidation. The presence of two mutually coupled methylenes at δ_H 3.00 (m) for H_2–6 and 4.40 (m) and 4.89 (obscured) for H_2–5, along with the UV/vis chromatogram (Appendix 8) indicated that the C–5 to C–6 bond was reduced. As the remaining NMR data was analogous to **154**, including the *meta* coupled

doublet (H–12) deshielded to δ_H 7.45 to indicate the sulfate group present at C–13, lamellarin A4-13-sulfate was proposed.



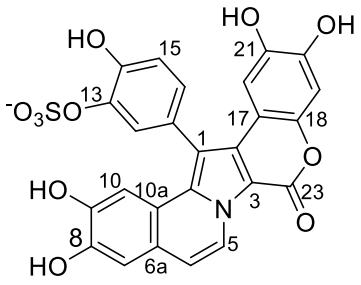
154

155 Δ saturated

Table 6.8 – ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for Lamellarin H-13-sulfate (**154**) in CD_3OD.

position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC
1	112.2			
2	130.2			
3	n.d.			
5	122.8	9.07 (d, 7.3)	6	6, 6a, 10b
6	113.9	7.05 (d, 7.3)	5	5, 7, 10a
6a	126.1			
7	112.4	7.10 (s)		6, 8, 9, 10a
8	149.1			
9	147.7			
10	110.5	7.15 (s)		6a, 8, 9, 10b
10a	119.9			
10b	136.5			
11	128.0			
12	127.2	7.52 (d, 2.1)	16	1, 13, 14, 15
13	142.1			
14	151.2			
15	119.9	7.21 (8.2)	16	11, 13
16	130.4	7.16 (dd, 8.2, 2.1)	12, 15	
17	110.7			
18	143.5			
19	104.4	6.81 (s)		17, 18, 20, 21
20	147.3			
21	148.2			
22	109.7	6.77 (s)		2
23	157.6			

n.d. = not detected



154

6.10 – Lamellarin D-20-sulfate and Lamellarin N-20-sulfate

Over the course of the purification, many metabolites were detected by HRESIMS with masses predicted to correspond to new compounds. However, as these lamellarin sulfates are natively charged, they produce a large signal intensity even when only sparingly small quantities are present. Nevertheless, four peaks were collected from HPLC purification (Figure 6.14, D3–D6) with metabolites showing deprotonated molecule ions at *m/z* 578.0783, consistent with the molecular formula $C_{28}H_{20}NO_{11}S^-$ (calcd. 536.0763), which does not correspond to any previously reported lamellarin sulfates. The first, lamellarin D-8-sulfate (**151**) isolated from D3, is described above. While the mass isolated from D6 was insufficient, D4 and D5 yielded samples with enough mass (>0.1 mg) to acquire ^{1}H detected NMR experiments on the 800 MHz NMR spectrometer at Griffith University, Australia. Based on ROESY correlations, their structures were tentatively assigned as lamellarin D-20-sulfate (**156**) and lamellarin N-20-sulfate (**157**).

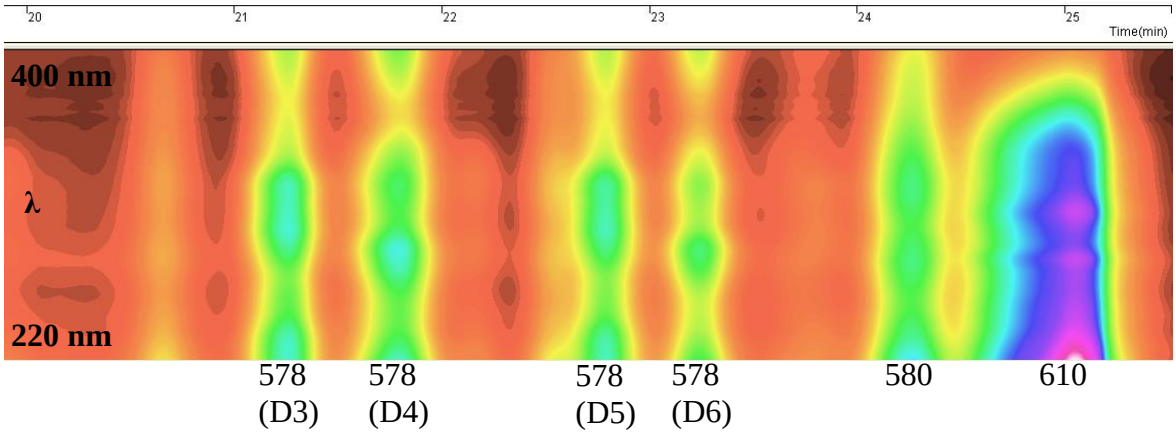


Figure 6.14 – Expanded HPLC trace of lamellarin sulfates (*m/z* and fraction ID of each compound shown below trace).

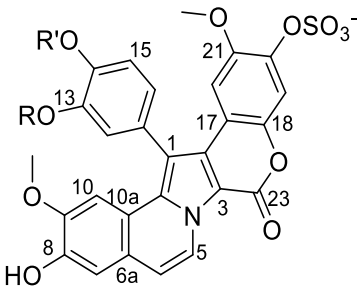
The ^{1}H NMR spectrum of **156** showed the characteristic peaks of a sulfated lamellarin, with a singlet at δ_H 7.54. Owing to the extremely low mass, HMBC data was hard to obtain, however through comparison with the data of the previous compounds and the ROESY spectrum, a tentative structure elucidation was completed. Based on the MS data and UV/vis trace, **156** has an unsaturated lamellarin core, therefore only positioning of the three methoxy and the sulfate group was required.

The ROESY correlation between methoxy δ_H 3.36 and H–22 (δ_H 6.75) placed one methoxy at C–21. An analogous correlation was observed between the *meta* coupled doublet H–15 (δ_H 7.16) and methoxy δ_H 3.77 to place a second methoxy at C–16. Finally, the ROESY correlation

from H–10 to δ_H 3.37 placed the last methoxy at C–9, corroborated by its chemical shift, which is further downfield (~ 3.7 ppm) when at C–8, as in **153**. The sulfate was placed at C–20 based on the deshielding of the C–19 proton (δ_H 7.55) which showed a HSQC correlation to δ_C 109.1. With this data in hand, the chemical structure was assigned as lamellarin D-20-sulfate, matching that observed for **151** with a different sulfate position.

For both **156** and **157**, the peaks of *ring A* were very broad in the ^{1}H NMR spectrum, even when run on the 800 MHz NMR spectrometer. It was postulated that deprotonation of the C–8 phenol leads to shifting of the electron density throughout the conjugated system (Figure 6.15), thus broadening the signals for H–7 and, to a lesser extent, H–10 (Figure 6.16). This was not observed in any of the other compounds isolated, as it is the first example with a phenol at C–8 and unsaturation between C–5 and C–6, with this observation providing further evidence for the phenol position. This phenomenon was also observed for the non-sulfated derivatives isolated from an Australian *Didemnum* sp. sample, which may explain why the structures of these compounds were only obtained after acetylation.[305,309]

The only difference between **156** and **157** was the ROESY correlation between the *ortho* coupled doublet H–15 (δ_H 7.26) and the more deshielded methoxy at δ_H 3.87, both suggesting in this isomer the methoxy is positioned at C–14. Thus, the molecule was assigned as lamellarin N-20-sulfate, however accurate comparisons to literature NMR data of the non-sulfated compounds is not possible for either **156** or **157** as only the triacetate derivatives have been published. Due to a paucity of material, the data obtained for both compounds were of poor quality, therefore full assignment was not possible. If more compound mass was isolated, then peracetylation of the phenol groups would likely be the best option to solve this issue.



156 R=CH_3 R'=H
157 R=H R'=CH_3

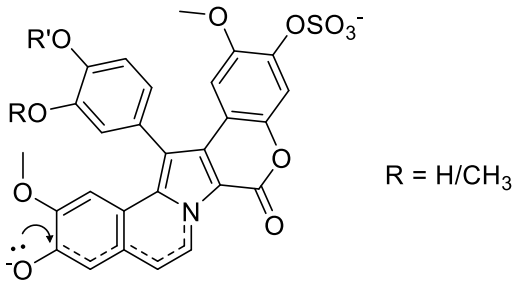


Figure 6.15 – Delocalisation of electron density from phenolic deprotonation resulting in NMR signal broadening for compounds **156** and **157**.

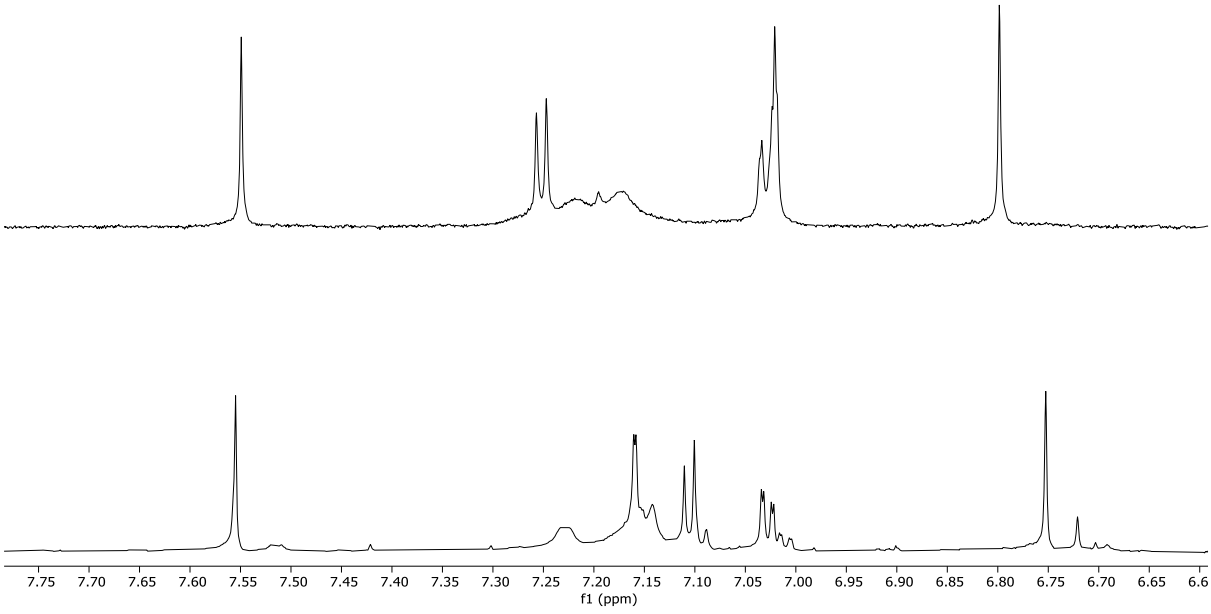


Figure 6.16 – ^{1}H NMR spectra of **156** (bottom) and **157** (top) illustrating broadening of *ring A* peaks.

6.11 – Previously Reported Lamellarin Sulfates

With the interest of discovering new analogues, samples with masses that did not correspond to any previously reported lamellarin sulfate were actively pursued over the course of the purification, therefore only three known examples were isolated, two of which are isomers of new compounds reported above.

Lamellarin Y-20-sulfate (**158**) was isolated as a white film. HRESIMS analysis detected a deprotonated molecule at *m/z* 580.0944 (calcd. 580.0919), consistent with the molecular formula $C_{28}H_{22}NO_{11}S^-$, which is the most common mass of known examples (lamellarin G, Y and L-sufates, all with three methoxy groups and a saturated core). The NMR data for all resonances except those of *ring A* matched that of **148**, which accounted for two methoxy groups (δ_H 3.34, 3.86) and the sulfate group. A long-range COSY correlation between methoxy δ_H 3.80 and H–7 (δ_H 6.94) showed the methoxy group is placed at C–8 and this was corroborated by HMBC data, giving lamellarin Y-20-sulfate. This compound was previously isolated from an unidentified ascidian;[302] all ^{1}H NMR data was consistent with that reported.

Lamellarin U-20-sulfate (**159**) was isolated as a white film. HRESIMS analysis detected a deprotonated molecule at *m/z* 594.1084, consistent with the molecular formula $C_{29}H_{24}NO_{11}S^-$ (calcd. 594.1076), and thus an isomer of **149**. The NMR data for both are very similar, except for the resonances associated with *ring F*. A methoxy group was shifted downfield to δ_H 3.83, caused by C–14 substitution rather than C–13, thus suggesting the structure was lamellarin U-20-sulfate, which was confirmed by comparison with the original published data isolated from the same unidentified tunicate as **158**.[302]

Finally, lamellarin α-20-sulfate (**160**) was isolated as a white film. HRESIMS analysis detected a deprotonated molecule at *m/z* 592.0921, consistent with the molecular formula $C_{29}H_{22}NO_{11}S^-$ (calcd. 592.0919), and thus an isomer of **150**. Akin to **149** and **150**, **160** was the C–5 to C–6 unsaturated analogue of **159**, with the NMR data identifying that all the methoxy groups were on the same carbons. Confirmation of **160** as lamellarin α-20-sulfate was achieved by comparison of spectroscopic data to those previously reported.[303]

6.12 – Lamellarin K

The extract of *D. ternerratum* contained a large amount of methoxylated lamellarins, which were mainly ignored due to the high probability of re-isolating known molecules. When the ^{1}H NMR spectrum of the less polar fraction (C2) was analysed, no peaks indicative of lamellarin sulfates were detected, and instead one major compound was observed with a protonated ion detected at *m/z* 532.1594 using HRESIMS analysis, consistent with the molecular formula $C_{29}H_{26}NO_9^+$ (calcd. 532.1602). Upon HPLC purification following the same method as above, lamellarin K (**161**) was isolated, with its spectral data matching that previously reported.[305] There were many other peaks also detected, however none were purified to a level suitable for NMR-based structure elucidation, as their MWs were the same as previously isolated congeners, however the isolation of **161** was very useful for comparison of bioassay results.

6.13 – Ethanol Extraction

To ensure the methoxylated lamellarins were not artefacts of methanolic extraction, two small samples of *D. ternerratum* were extracted simultaneously with MeOH or EtOH and analysed using the GNPS LCMS protocol. These data were then networked together (Figure 6.17), sharing nodes corresponding to many of the major isolated lamellarin sulfates and also the precursor ion *m/z* 624.119 which corresponds to the mass of two previously reported examples **173** and **177** (see below). As only very small quantities of the tunicate were used for this experiment, the minor metabolites were not present in the network. However, this proves these molecules are legitimately biosynthesised secondary metabolites of the *D. ternerratum*.

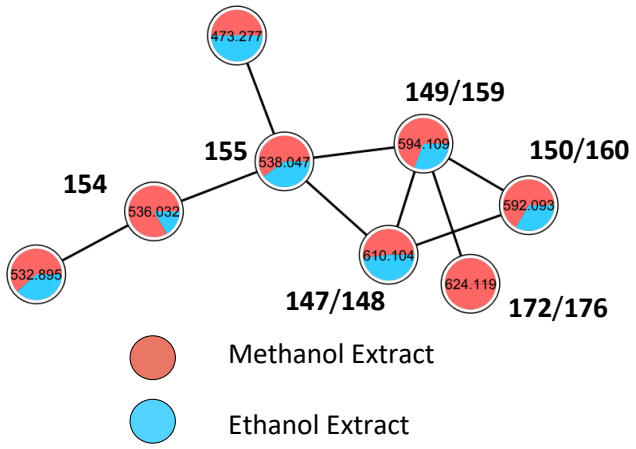
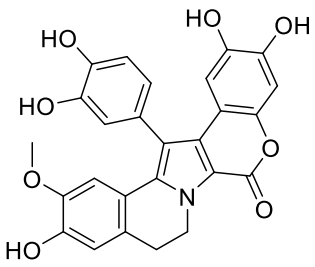


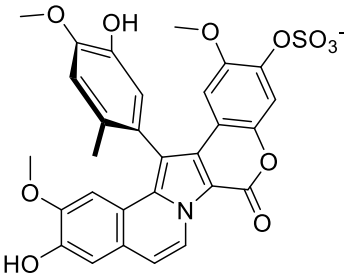
Figure 6.17 – Molecular network comparing MeOH and ethanol extracts of *D. ternerratum*.

6.14 –Atropisomerism of Lamellarins

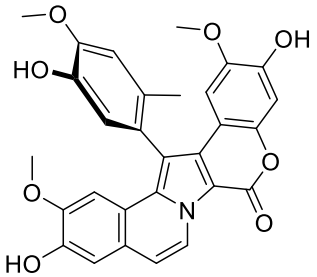
Structurally, the basic lamellarin pentacyclic core is planar, but from single-crystal X-ray diffraction of the first lamellarins isolated (A–D), the appended *ring F* sits 90° out of plane, and thus gives rise to axial chirality (atropisomerism).[264] However, none of these, or the vast majority of naturally isolated lamellarins, have been reported with optical activity. Lamellarin S (**162**) is the only reported optically-active example, where it slowly racemised with a half-life *c.* 90 days.[310] Molecular modelling and variable-temperature NMR calculations suggested the barrier to rotation within various lamellarins is typically on the order of 72–87 kJ mol^{-1}, depending on substitution patterns, thus this semi-restricted rotation about the C–1 to C–11 bond is typically a thermally facile process resulting in racemisation.[311] There have been no reported investigations on what governs atropisomerism in lamellarins, however racemisation would likely be dictated by the extent of substitution about the aromatic appendage and how they affect the rings rotation through steric and electronics (i.e. hydrogen bonds). Bioactivity SAR studies are also difficult as there have been no reports of purification of natural enantiomers, however 16-methyl atropisomers of the potent kinase inhibitor lamellarin N have been synthesised and resolved (*R*- and *S*-**163**).[312] The *R* isomer inhibited nearly all kinases tested, whereas the *S* isomer selectively inhibited GSK3α/β, PIM1 and DYRK1A, therefore semi-synthetic efforts in favour of one atropisomer may be able to reduce the broad-spectrum activity of many other related compounds.



162



S-**163**



R-**163**

The optical activity of **148**–**151** was assessed using ECD spectroscopic analysis. The most routine method for configurational analysis by ECD has become the comparison to data predicted using TDDFT quantum mechanical calculations, as acquiring experimental ECD data is simple and fast, and the calculations can be done on a desktop computer in reasonable time

using both commercial and freely available software.[313] This tool has become increasingly popular over recent years, representing 35% of all configurational assignments of reported MNPs in 2018.[314]

All of **148**–**151** were optically active, producing ECD traces with very similar Cotton effects, and therefore it is likely each has the same absolute configuration, or excess of one enantiomer over the other in the case of these molecules where racemisation is possible (Figure 6.18). This assumption does not hold true for all molecules,[315] particularly when there are slight variations about sources of chirality within a molecule, however between **148**–**151**, the structural differences of hydroxy, methoxy and sulfate substituents are unlikely to affect their ECD spectra.

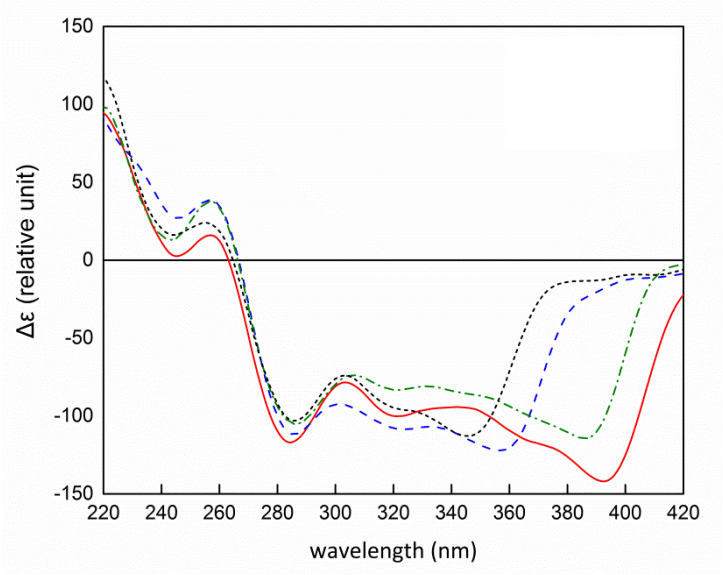


Figure 6.18 – Experimental ECD spectra of **148** (black), **149** (blue), **150** (green) and **151** (red).

(Data acquired by J. Bracegirdle; Figure generated by L. Robertson)

The ECD spectra for the pair of atropisomers, (a*S*)-**150** and (a*R*)-**150** were calculated using TDDFT[316] using the CAM-B3LYP/def2SVP//B3LYP/def2SVP functional/basis set combination (L. Robertson, Griffith University). The calculated spectrum of (a*R*)-**150** matched the experimental data much more closely (Figure 6.19), therefore indicating this to be the most likely configuration. By analogy, this a*R* configuration is also suggested to be favoured for **148**, **149** and **151** although this remains to be verified computationally.

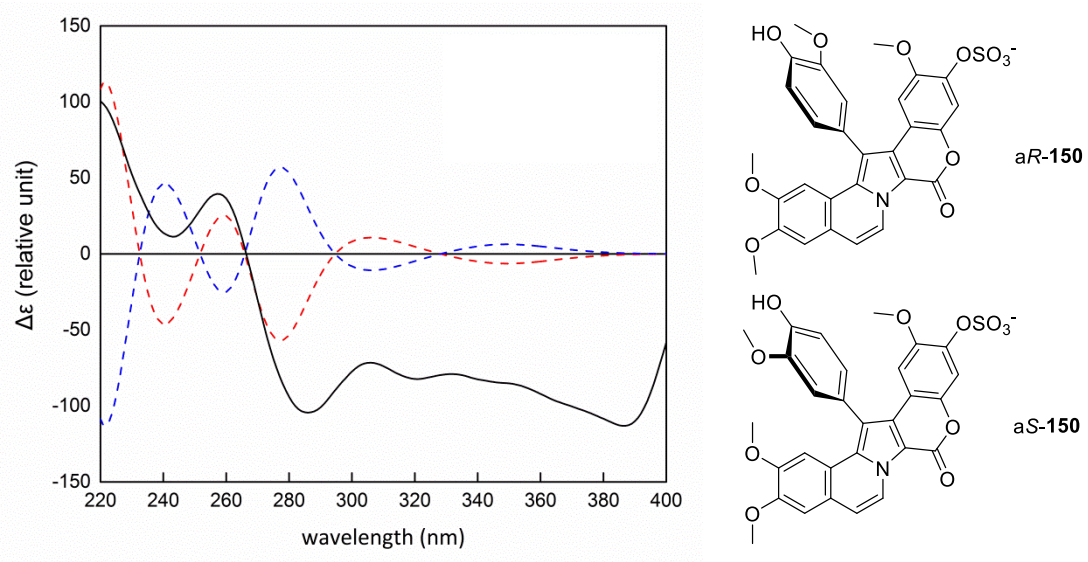


Figure 6.19 – Comparison of the experimental (black) and predicted ECD spectra of (a*R*)-**150** (red) and (a*S*)-**150** (blue).

(Calculations and figure generated by L. Robertson)

In order to better understand the influences upon atropisomerism in this class of compound, specifically the observed atropisomerism in **148**–**151**, the free energy barrier to rotation was computed for each molecule using the B3LYP exchange-correlation functional,[317-319] the 6-311G(d,p) basis set and a polarizable continuum solvent model (MeOH) (P. Hume, VUW), along with lamellarin S , the only previously reported lamellarin to show optical activity, as a control. As *ring F* can rotate in either direction, both resulting in the same change in configuration, two transitions states exist, and both were considered and calculated for each compound. The previous reported experimental results for lamellarin S observed a half-life of ~90 days.[310] Using the Eyring equation, $k = \frac{k_B T}{h} e^{(-\frac{\Delta G^{\ddagger}}{RT})}$, the lower of the two barriers to rotation calculated for lamellarin S (+25.2 kcal mol^{-1}) corresponds to a half-life of ~112 days at 4 °C, in good agreement with the observed racemisation. As lamellarin S has two phenol groups on *ring F* and no methoxy substituents, it was predicted to racemise more readily than **148**–**151**. However, the energies calculated for the new lamellarins were lower, suggesting the opposite.

	$\Delta G_1^{\ddagger}$	$\Delta G_2^{\ddagger}$
162	+25.2 (+25.4)	+25.9 (+25.4)
148	+19.2 (+18.4)	+19.3 (+18.2)
149	+18.5 (+19.6)	+17.4 (+19.2)
150	+21.5 (+22.2)	+21.3 (+21.3)
151	+22.9 (+22.2)	+24.2 (+22.8)

Table 6.9 – DFT-calculated rotational transition state energies for **162** and **148**–**151** using the B3LYP exchange-correlation functional. Energies for protonated versions in parenthesises. All values in kcal mol^{-1}.

To ensure this was not an artefact of the natively charged lamellarin sulfates, protonated versions of **148**–**151** were also used for analogous calculations where this trend was still observed (Table 6.9, values in parenthesises). When the solvent model was removed, these trends were observed again. Finally, to ensure the results were not due to a failure of the B3LYP basis set, rotational barriers for **162** and **150** were calculated using the highly parametrised M06-2X[320] and minimally parametrised PBE1PBE[321-323] exchange-correlation functionals, resulting yet again in comparable values (Table 6.10).

	B3LYP	M06-2X	PBE1PBE
162	+25.2	+26.8	+25.9
150	+21.5	+22.1	+21.7

Table 6.10 – DFT-calculated rotational transition state energies for **162** and **150** with different exchange-correlation functionals. All values in kcal mol^{-1}.

These results are very surprising, considering steric principles, one would assume the methyl groups are bulkier and should thus restrict rotation of the phenyl ring. One clear difference between these molecules is the negatively-charged sulfate group, which may form persistent ion-pair interactions in solution, or interact with solvent molecules, producing an additional steric barrier that is not captured in the calculation. This issue may be able to be resolved with kinetic studies on **148**–**151**, however in the absence of more material, such studies were not possible.

6.15 – Biological Activity

The new lamellarin sulfates (**147**–**153**) and lamellarin K (**161**) were submitted to SBS, VUW, for testing against the human colon carcinoma cell line HCT-116. Compounds **161** and **151** showed moderate cytotoxicity, with IC_{50} values of 4.3 and 9.7 μM respectively, while all other compounds showed only weak activity. This observation is in accordance with a reduction in cytotoxic activity upon sulfation,[303] likely attributable to the increased polarity of the molecule which makes traversing the lipophilic cell membrane more difficult. The full results are presented in Table 6.11.

	Average IC_{50}	Standard Deviation
147	19.5	5.5
148	27.9	2.7
149	>50	–
150	42.0	5.9
151	9.7	2.1
153	40.4	3.5
161	4.3	0.6
Peloruside A	0.04	0.01

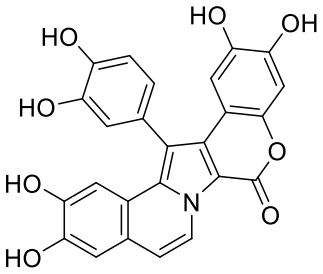
Table 6.11 – IC_{50} values (μM) of compounds isolated from *D. ternerratum* against the HCT-116 cell line. Peloruside A was used as a positive control.

Previous lamellarin sulfates have shown potential as anti-viral agents (see below). Due to sample limitations however, we were unable to test the activity of **147**–**151** or **153**. The lack of overt cytotoxicity observed in our experiments suggests that these compounds may be worthy of further investigation, if the compounds could be re-isolated or synthesised.

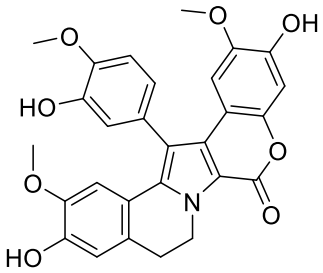
6.16 – Related Compounds

The lamellarins are a family of secondary metabolites that have been isolated from range of marine animals, particularly didemnid ascidians. Andersen *et. al.* first reported the isolation of these DOPA-derived alkaloids in 1985 from the marine prosobranch mollusc *Lamellaria* sp. collected in Palau.[264] Since then, over 70 natural derivatives have been reported, and lamellarin

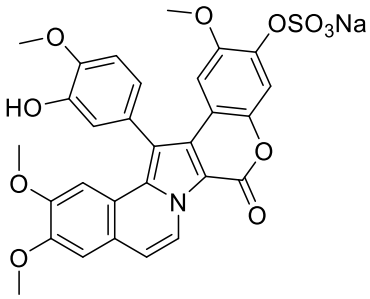
is now the term used to loosely describe the central pyrrole ring structure, with the family mainly consisting of 14-phenyl-*6H*-[1]benzopyrano[4',3':4,5]pyrrolo[2,1-*a*]isoquinolin-6-one hexacyclic ring systems (**164**, **165** and **159**), however some 'open' structures (i.e. lamellarins O (**166**), R (**167**) and Q (**168**)) are also deemed lamellarins. All reported isolations have been thoroughly reviewed many times,[324-328] with the most recent example from the Capon research group isolating lamellarins A1–A6 from two *Didemnum* spp. in 2012.[300]



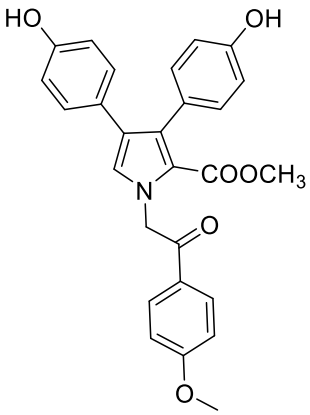
164



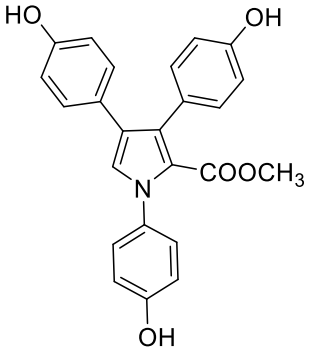
165



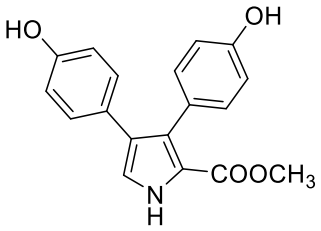
159



166



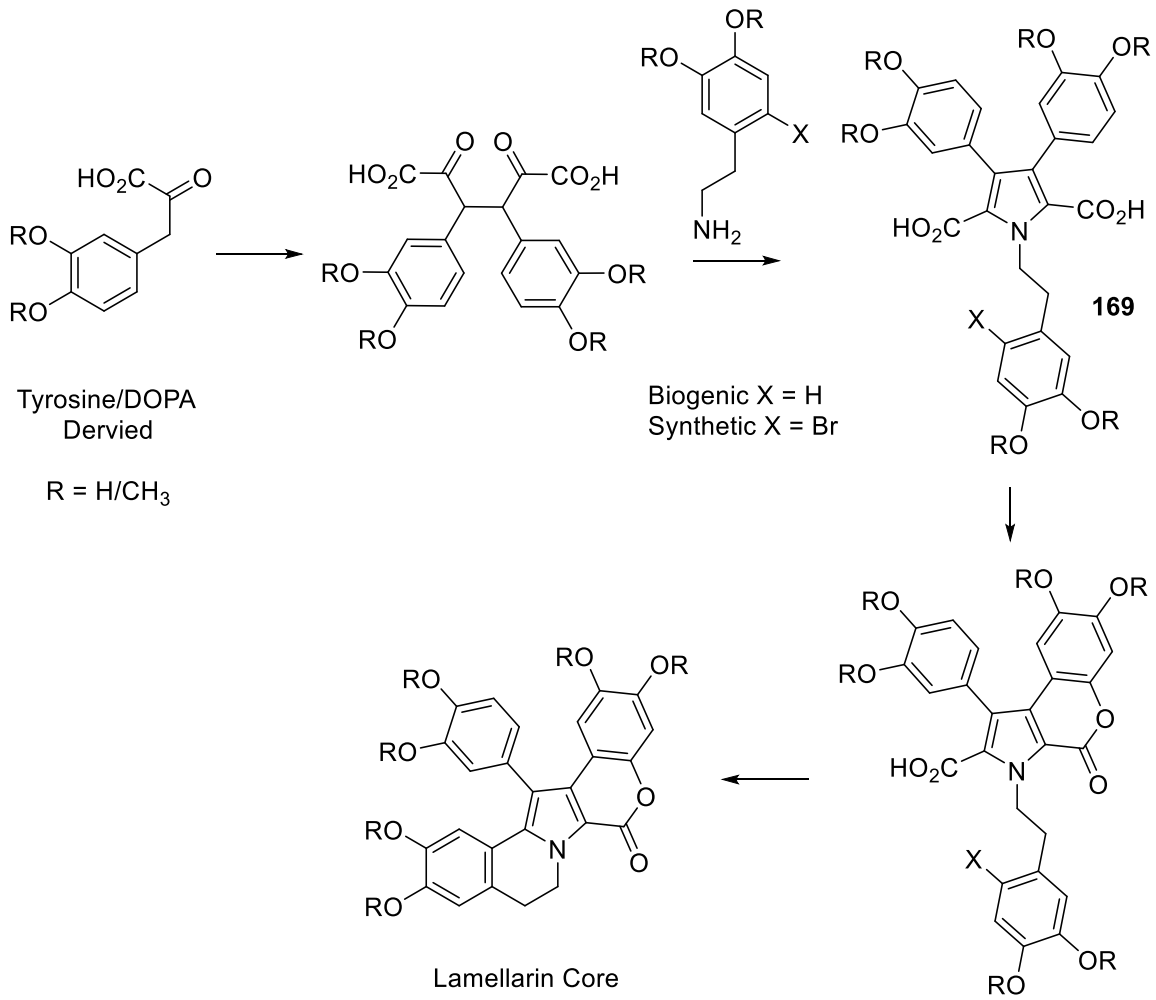
167



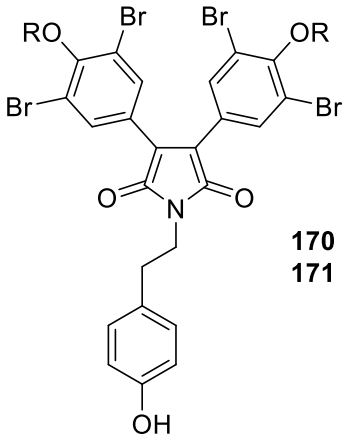
168

Although there have been no biogenic or metabolomic studies reported on how the lamellarins are produced *in vivo*, it is believed they are formed by the condensation of three DOPA or tyrosine amino acid molecules.[326] Two consecutive oxidative cyclisations followed by a pyrrole-forming condensation with the amino group of a third amino acid-derived molecule forms 3,4-diaryl-1-(2-arylethyl)-2,5-pyrrole dicarboxylic acid (**169**). The lamellarin core structure is then formed following several enzymatic ring-closures. Precedence for the biosynthetic route comes from the structurally similar polycitrins A (**170**) and B (**171**), isolated from a *Polycitor* ascidian species.[329] This postulated route has been mimicked synthetically by

Steglich and co-workers to synthesise the lamellarin G trimethyl ether and lamellarin L,[330,331] however in the absence of enzymatic catalysts, the final step utilised a palladium catalysed Heck cyclisation to afford the isoquinoline ring system (Final step in Scheme 6.2). This required a bromine substituent (labelled X) to be included at the position of cyclisation on the third precursor molecule, which undergoes a decarboxylative reaction at the pyrrole double bond.

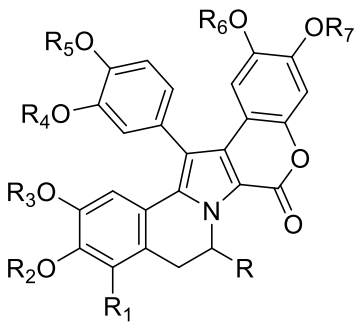


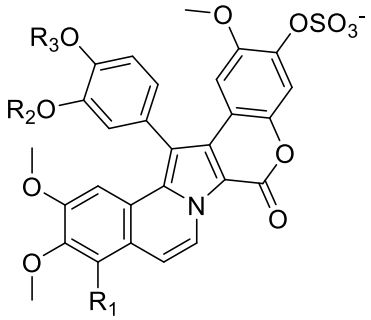
Scheme 6.2 – Proposed biosynthesis and biomimetic synthesis of lamellarins. X = H in enzymatically controlled biosynthetic process, whereas X = Br for chemical synthesis via the Heck cyclisation reaction.



170 R= H
171 R=CH_3

Marine natural products can be sulfated, drastically changing their solubility and bioactivity compared to non-sulfated analogues.[332] In 1997, Reddy *et. al.*[302] reported the first isolation of sulfated lamellarins from an unidentified ascidian collected in the Arabian Sea. Extraction led to the first identification of lamellarins T–X, along with the C–20 sulfate derivatives of T (**172**), U (**159**), V (**173**) and Y (**174**). It was noted that the presence of the sulfate group shifts the *ortho* proton at C–19 downfield from ~6.80 to ~7.50 ppm in the ^{1}H NMR spectrum, which was a key characteristic used in this current study to assist in locating the position of sulfation. The biological activity of these compounds was not analysed as they decomposed before testing.





	R	R_1	R_2	R_3	R_4	R_5	R_6	R_7
152	H	H	SO_3	CH_3	OH	CH_3	OH	CH_3
174	H	H	CH_3	OH	OH	CH_3	CH_3	SO_3
177	H	H	OH	CH_3	OH	CH_3	CH_3	SO_3
159	H	H	CH_3	CH_3	OH	CH_3	CH_3	SO_3
172	H	OCH_3	CH_3	CH_3	OH	CH_3	CH_3	SO_3
176	H	OCH_3	CH_3	CH_3	CH_3	OH	CH_3	SO_3
173	OH	OCH_3	CH_3	CH_3	OH	CH_3	CH_3	SO_3

	R_1	R_2	R_3
160	H	OH	CH_3
175	CH_3	CH_3	OH

Isolation of sulfated lamellarins has only been reported twice since, for a total of nine compounds. First, the Australian ascidian *Didemnum chartaceum* was the source of three new C–20 sulfated derivatives (B (**175**), C (**176**) and L (**177**)) and the first C–8 sulfated derivative (Lamellarin G (**152**)).[304] By measuring the T_1 relaxation by inversion recovery methods, it was observed that the C–19 *ortho* proton of the C–20 sulfated lamellarins shows T_1 relaxation times 3–5 times longer than the derivative lacking the sulfate, presumably due to the lack of relaxation pathways.[304] This is reflected in the ^{1}H NMR signal intensity, integrating for half of the neighbouring signals in standard experiments.

This study was closely followed by the publishing of lamellarin α-20-sulfate (**160**) from the aforementioned unidentified Arabian ascidian,[303] which demonstrated inhibition of HIV-1 integrase with the best therapeutic index of those compounds tested. Integrase is the enzyme responsible for attaching viral cDNA to the host DNA, initially breaking the host DNA and to allow the integrative processes of terminal cleavage and strand transfer to occur. Lamellarin α-20-sulfate reduced the accumulation of products from both processes at a IC_{50} lower than sulfated lamellarins U and V. The isolated enzymes catalytic core domain showed a reduced sensitivity compared to the full enzyme inhibition (IC_{50} 64 μM vs 7 μM), suggesting multisite binding with either/both of the N and/or C terminal domains, which is unlike most other natural product inhibitors.[333] In comparison, lamellarin α showed no inhibition up to 1.6 mM.[308] It can be envisaged that the sulfate group binds competitively to the positively charged DNA binding region typical of the terminal phosphate of DNA.

The baculiferins are a class of similarly structured sulfated pyrrole alkaloids, isolated from the Chinese sponge *Iotrochota baculifera*.[334] Many of these compounds were also found to be potent inhibitors against HIV-1 virus, with a decreased cytotoxicity compared to the non-sulfated derivatives. The most simply structured congener, baculiferin O (**178**), is a sulfated derivative of ningalin A (**179**), originally isolated from a *Didemnum* specimen collected from the Ningaloo Reef, Australia.[335] During the current study, an aqueous extract made by sonicating a sample of *D. ternerratum* in H2O for 20 mins was examined by HRESIMS. Deprotonated molecules for both **178** and **179** were tentatively detected (Figure 6.20); however, neither could be isolated in quantities useful for NMR-based structure elucidation. An ion suitable for baculiferin O (**178**) was also observed in cluster C of the optimised molecular network in Figure 6.1, showing an edge connection to lamellarin A4-13-sulfate (**155**).

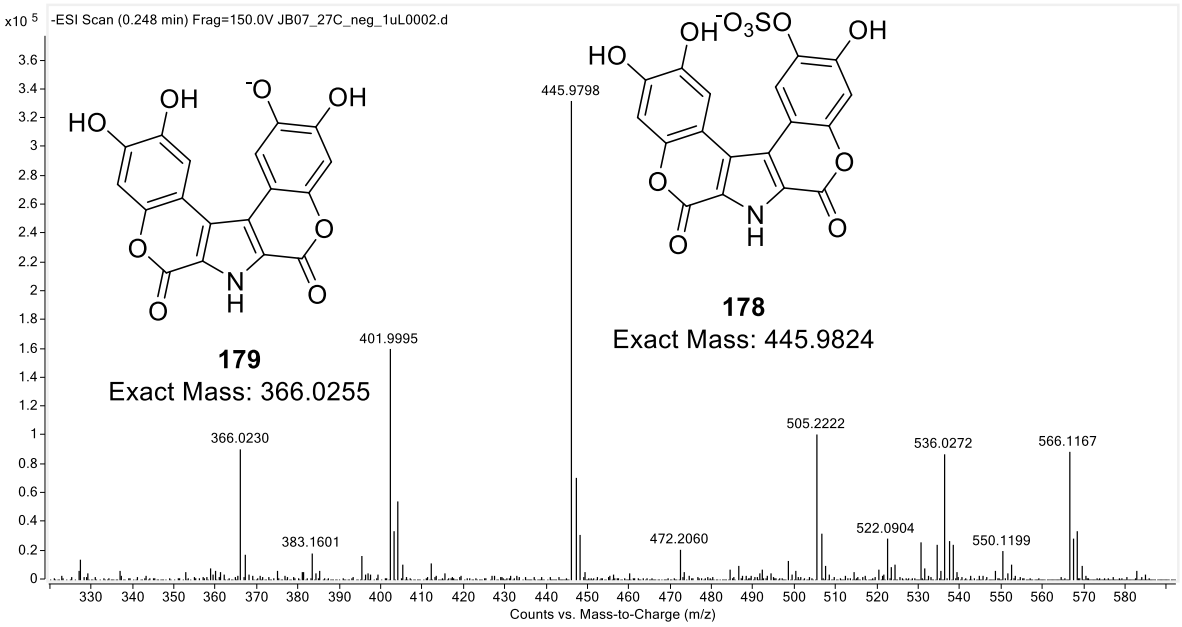
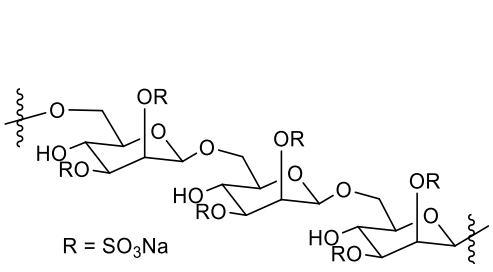
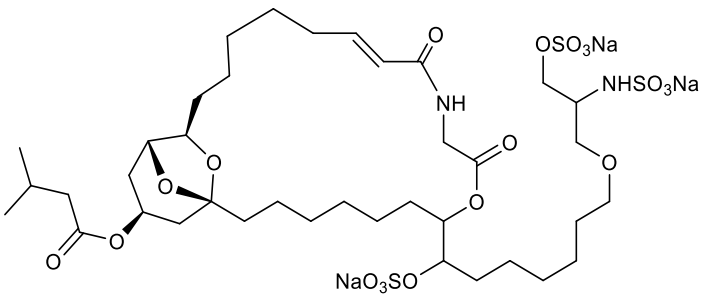


Figure 6.20 – MS/MS spectrum of *D. ternerratum* aqueous extract.

Anti-viral activity resulting from sulfated compounds isolated from ascidians is not unprecedented,[336] particularly those from the *Didemnidae* family. The sulfated mannose homopolysaccharide kakelokelose (**180**) was isolated from the Micronesian tunicate *Didemnum molle* and showed 100% inhibition of HIV infection at 0.3 μg mL^{-1} with no cytotoxicity at 15 μg mL^{-1}.[337] *Didemnum guttatum* collected in Palau was the source of cyclodidemniserinol trisulfate (**181**), which inhibited purified HIV integrase with an IC_{50} of 60 μg mL^{-1},[338] however no anti-viral activity has been reported thus far for the non-sulfated didemniserinolipds.[339]

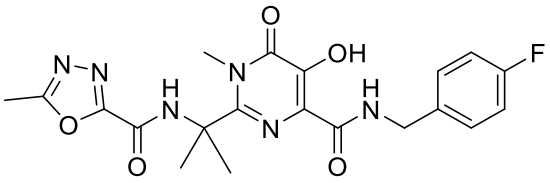


180

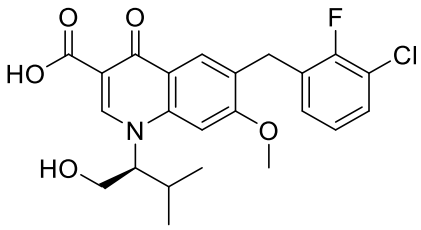


181

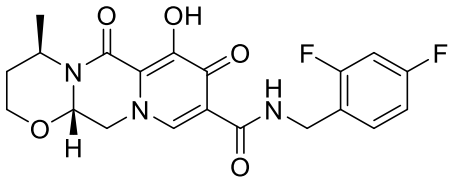
The impact of the acquired immunodeficiency syndrome (AIDS) epidemic needs no introduction, with the cause, HIV-1 retrovirus, encoding three essential enzymes for replication: reverse transcriptase, protease and integrase. The first clinically used anti-HIV therapeutics targeted either reverse transcriptase or protease, however mutations resulting in drug resistance have occurred and now many of these treatments must be taken in combination.[340] Many natural and synthetic compounds target integrase, however it was not until 2007 when raltegravir (**182**), the first anti-integrase drug, was approved by the FDA that has since been followed by elvitegravir (**183**) and dolutegravir (**184**), all of which are synthetic in origin.[341] These drugs bind through the catalytic core domain, to which resistance has been observed and therefore drugs that bind to other/multiple sites are desirable.[342] Integrase represents a superior target for anti-HIV drugs compared with the other enzymes, as no similar proteins exist in human cells, which should result in less off-target activity and thus side effects.



182



183

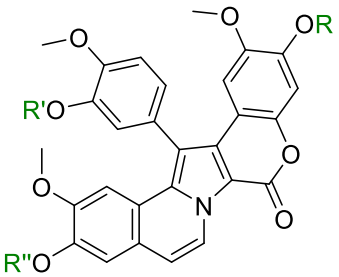


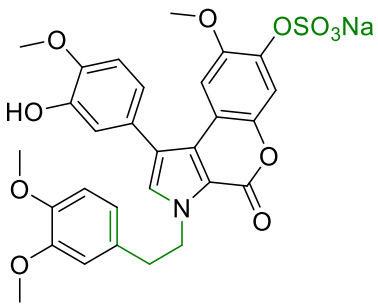
184

Synthetic attempts have been made towards establishing an anti-HIV integrase SAR for lamellarin α-20-sulfate (**160**). Ridley *et. al.*[308] synthesised lamellarin α (**185**), lamellarin α-13,20-disulfate (**186**) and the entirely demethylated derivative lamellarin H (**164**), and compared their inhibition of HIV-1 integrase, Molluscum contagiosum virus (MCV) topoisomerase and cytotoxicity towards HeLa cells to the naturally isolated 20-sulfate. Lamellarin α showed no inhibition of integrase at 1.6 mM compared to **185** (LD_{50} 22 μM) and **186** (LD_{50} 49 μM), proving the sulfate group(s) critical towards the selective activity of methylated derivatives. Lamellarin H (**164**) was the most potent inhibitor tested (IC_{50} 1.3 μM), possibly owing to the three catechol functionalities that have been of interest when designing

synthetic inhibitors,[340] however it was just as potent in the counter-screen against MCV topoisomerase (IC_{50} 0.23 μM) and is therefore a non-selective inhibitor. This non-selectivity was also observed for **186** (LD_{50} 70 μM) but not **185** (LD_{50} >170 μM), postulated to be the result of interacting with targets that typically bind two phosphate groups of DNA resulting in 'nuisance' activity.[308]

Iwao and co-workers were the first group to synthesise lamellarin α-20-sulfate,[343] and using a more recently developed divergent approach,[344] synthesised a panel of several lamellarin α sulfate derivatives (**160** and **185–189**). For the first time, the anti-HIV–1 activity of a ring opened sulfate derivative (**189**) was assayed, showing no growth inhibition. Therefore, the hexacyclic ring system is essential for anti-viral activity. It was determined through confocal laser scanning microscopy that the very hydrophilic sulfated lamellarins have limited cellular uptake opposed to the nonsulfated derivatives, possibly explaining the reduction in cytotoxicity.[344] This suggested that these compounds work through the virus entry step rather than cDNA integration, supported by evidence from HIV-1 envelope glycoprotein mediated cell-cell fusion assays. However, aside from the series in the initial report containing sulfated lamellarins U and V, all the bioactivity testing has focused on lamellarin α (and derivatives thereof).





189

160	R=SO_3Na	R'=H	R"=H
185	R=H	R'=H	R"=H
186	R=SO_3Na	R'=SO_3Na	R"=H
187	R=H	R'=SO_3Na	R"=H
188	R=SO_3Na	R'=SO_3Na	R"=SO_3Na

6.17 – Conclusion

Through a MS-guided molecular networking approach, the Tongan tunicate *D. ternerratum* was prioritised and purification of its chemical components resulted in the isolation and characterisation of three previously reported and 10 new lamellarin sulfates, along with lamellarin K. This study reflects the utility of molecular networking for the discovery of new congeners in a previously studied, but under investigated, chemical class. It was clear from the cluster of nodes corresponding to lamellarin sulfates that there were a variety of different methylation patterns and many of the precursor ion masses had not been previously reported. This allowed rapid detection of new metabolites early on in the project, when other methods such as ^{1}H NMR spectroscopy would not detect new compounds yet. A ^{1}H NMR screen would have also prioritised this extract based on the interesting aromatic peaks, however, dereplication of the thoroughly studied lamellarin class may have discouraged further research. Therefore, the MS-guided approach worked well to direct isolation and characterise the new lamellarin sulfate congeners.

Chapter 7 – Concluding Remarks

The biodiverse marine environment is a unique source of new chemistry, continuing today to provide new compounds of interest for drug discovery. This is exemplified by this work, where the investigation of 15 organisms led to the isolation of 16 previously unreported compounds from five sources, across a range of compound classes, biological activities and potential applications. The tunicates, bryozoan and red alga used in this work were collected from the waters of the Kingdom of Tonga and NZ.

The investigation of these macroorganisms is contrary to the direction the MNP field as a whole has taken in recent times, with a heavier focus on microorganisms.[10,11] Even if the secondary metabolites are produced by symbiotic microorganisms, the scientific community is not in a position to culture every microbe, nor are current metagenomic techniques able to explore the full chemical diversity in a sample. Although there has been significant progress in understanding secondary metabolite biosynthesis, this is still not fully understood to a level where genomic data alone can be used for MNP discovery. Therefore, in order to explore the chemical constituents of macroorganisms (particularly invertebrates), full chemical extraction methods must still be employed.

Three screening methods were used in this study to prioritise organisms for bulk extraction with the ultimate goal of isolating new and bioactive compounds. The unidentified tunicate prioritised based on a positive hit in the bioassay screen, resulted in the re-isolation of two previously reported and commonly encountered metabolites (**25** and **26**).

The ^{1}H NMR spectroscopy-based approach was used to screen six marine organisms, prioritising three for further investigation. Although the structure of the major metabolite (**27**) of *P. polypora* was not determined, bulk extraction of the red alga "*P. angustum*" resulted in the isolation of the new halogenated monoterpene costatone C (**46**), while the bryozoan *N. nelliiformis* afforded the nucleosides nellielloside A and B (**61** and **62**) and inspired the synthesis of four congeners (**66**–**69**).

The strength of this screening approach is the detection of secondary metabolites in quantities applicable to an NMR-based structure elucidation, working very well to eliminate samples such as organisms PTN4_01A and PTN3_03C. It is also a stand-alone technique that does not require any controls or reference samples. The major flaw is the its detection of novelty, which often requires some further degree of purification before dereplication can be accurately completed from NMR spectroscopy alone.

A set of eight tunicates were screened by a mass spectrometry-based approach, utilising the GNPS online platform to analyse LCMS/MS data of the extracts. Organisms that produced unique constellations were prioritised for further investigation. The tunicate *C. aucklandicus* was used as a negative control, where lack of a standalone cluster suggested this organism did not contain unique (and/or ionisable) chemistry in the set and validated by a ^{1}H NMR spectroscopy analysis which showed little evidence of secondary metabolites worth examining. Three tunicates were prioritised, and bulk extracted. Investigation of the major chemical constituent of *D. stylifera* revealed the new aromatic ketone **28**, while thorough examination of an *S. kuranui* extract resulted in the isolation of two new methylated rubrolides (**110** and **111**) which both showed strong antibacterial activity. Molecular networking of the extract of the Tongan tunicate *D. ternerratum* suggested the presence of a number of previously unreported lamellarin sulfates, prompting further investigation that led to the isolation and structure elucidation of ten new derivatives (**147–151** and **153–157**).

Mass spectrometry-based screening using the GNPS platform is an excellent tool for detecting unique compounds against a reference data set. It worked very well when comparing the data sets of a number of related organisms which should have similar primary metabolites. The technique works best when a compound class is known, as precursor ions can be used to predict novelty, the nature of which can be explored by analysing the mass shifts from annotated nodes. As exemplified by the lamellarins and the rubrolides, new derivatives in an extract can be identified and targeted for purification, even when the class of compound has been thoroughly investigated. The main drawbacks lie in *de novo* discovery, such as an organism that has no parallels in a collection (e.g. *N. nelliiformis*) or a compound class that has not been previously reported/annotated. When a full network is produced there are a huge number of constellations that form clusters with only one (self-loop) or two nodes, making it difficult to detect unique metabolites that are present alone in an extract. Molecules that do not ionise or fragment well will not be observed with this technique.

Overall, all three screening techniques described here have a place in the arsenal of an MNP chemist. For best results, they should be used parallel with each other against databases such as MarinLit, in order to detect new bioactive molecules from marine organisms as quickly and effortlessly as possible.

Chapter 8 – Experimental

General Experimental Procedures – UV/vis spectra were obtained on an Agilent 8453 UV/vis spectrophotometer or extracted from HPLC chromatograms. Optical rotations were measured using a Rudolph Autopol II polarimeter. ECD spectra were recorded on a ChiraScan CD spectrometer or a JASCO J-715 spectropolarimeter. A 600 MHz Varian Direct Drive spectrometer equipped with a 5 mm PFG dual broadband probe or a JEOL JNM-ECZ600R with a nitrogen cooled 5 mm SuperCOOL cryogenic probe were used to record the NMR spectra of all compounds (600 MHz for ^{1}H nuclei and 150 MHz for ^{13}C nuclei) except **153**, **156** and **157** which were run on a Brüker Avance III HDX 800 MHz equipped with a triple (TCl) resonance 5 mm cryoprobe (800 MHz for ^{1}H nuclei and 200 MHz for ^{13}C nuclei). The residual solvent peak was used as an internal reference for ^{1}H (δ_H 2.50, DMSO-d_6; 3.31, CD_3OD; 7.26, $CDCl_3$) and ^{13}C (δ_C 39.52, DMSO-d_6; 49.0, CD_3OD; 71.6, $CDCl_3$) chemical shifts.[345] Samples were quantified by ^{1}H NMR spectroscopy using the residual DMSO-d_5 peak, calibrated and acquired according to the parameters described by Pierens and co-workers.[346] Standard pulse sequences supplied by Varian and JEOL were used for NMR analyses. High-resolution (ESI or APCI) mass spectrometric data were obtained with an Agilent 6530 Accurate Mass Q-TOF LCMS equipped with a 1260 Infinity binary pump. IR (film) spectra were recorded using a Bruker Platinum Alpha FTIR spectrometer. EI-GCMS analyses were carried out using a Shimadzu QP2010-Plus gas chromatograph–mass spectrometer equipped with an AOC-20i auto injector and operating a QP2010 MS detector.

Reversed-phase column chromatography was achieved using Supelco Diaion HP20 and HP20ss (PSDVB) chromatographic resin. Size exclusion chromatography was achieved using Sephadex LH20 resin. HPLC purifications were carried out using a Rainin Dynamax SD-200 solvent delivery system with 25 mL pump heads with a Varian Prostar 335 diode array detector, or an Agilent Technologies 1260 Infinity HPLC equipped with a diode array detector and an Agilent 380 evaporative light-scattering detector. Octadecyl-derivatised silica (C_{18}, 5 μm, 100 Å) HPLC columns (Phenomenex) were either analytical (4.6 mm × 250 mm, 1 mL/min) or semi-preparative (10 mm × 250 mm, 4 mL/min), while a Phenomenex Luna octyl-derivatised silica gel (C_8) column (analytical; 4.6 mm × 250 mm, 1 mL/min) was also used. TLC was carried out using Machery-Nagel Polygram Sil G/UV_{254} plates developed using a H_2SO_4 (5% in MeOH) vanillin (0.1% w/v in ethanol) char. The mobile phase for separation is described in each section. All solvents used were of HPLC grade and H_2O was glass distilled. Solvent

mixtures are reported as percent volume/volume. All reagents were of commercial quality, obtained from either Sigma Aldrich or AK Scientific and were used without prior purification.

Chapter 2 – Screening

Organism Material – All specimens sourced from the Kingdom of Tonga were collected over three trips (2006, 2009 and 2016), immediately frozen and transported to NZ. NZ specimens were collected by NIWA. All organisms were stored at –18 °C in the School of Chemical and Physical Sciences, VUW, until extraction.

Bioassay-Guided Screening– The extracts of a range of Tongan organisms were screened by Dr Nathaniel Dasyam, School of Biological Sciences, VUW, for activity against *Mycobacterium smegmatis*, with PTN3_40B showing activity (>80% growth inhibition at 100 μg mL^{-1}).[114] The frozen unknown purple ascidian (40 g) was extracted in MeOH (100 mL) twice overnight. The second extract, followed by the first, were passed through a HP20 column (40 mL), pre-equilibrated in H_2O and combined following elution. The eluent was then diluted with an equal volume of H_2O and passed back through the column twice, followed by a 120 mL H_2O wash. The column was then eluted with 120 mL portions of: (1) 30% Me_2CO/H_2O; (2) 75% Me_2CO/H_2O; and (3) Me_2CO (fractions A1−A3, respectively). Fraction A2 was dry-loaded onto a 10 mL HP20ss column, eluted into 5 mL test tubes with 30 mL portions of 60%, 70%, 80%, 90% MeOH/H_2O, MeOH, and Me_2CO and the resulting fractions were pooled together on the basis of TLC (5% MeOH in DCM), resulting in six major fractions B1–B6. Fractions B5 and B6 were individually purified on analytical C_{18} HPLC, using a linear gradient from 10% MeOH/H_2O (0.1% FA) to 100% MeOH (0.1% FA) over 20 minutes to afford kuanoniamine D (**26**) (1.0 mg, t_R = 16 min) and shermilamine B (**25**) (2.0 mg, t_R = 18 min), respectively.

Kuanoniamine D (**26**); pink film; HRESIMS *m/z* 361.0925 $[M + H]^+$ (calcd. for $C_{20}H_{17}N_4O_2S$, 361.1118); all NMR data matches those previously reported.[117]

Shermilamine B (**25**); orange film; HRESIMS *m/z* 391.1027 $[M + H]^+$ (calcd. for $C_{21}H_{19}N_4O_2S$, 391.1223); all NMR data matches those previously reported.[116]

NMR Spectroscopy-Guided Screening – All organisms were extracted in MeOH twice overnight, and the second extract, followed by the first, were passed through a HP20 column, pre-equilibrated in H_2O and combined following elution. The eluent was then diluted with an

equal volume of H_2O and passed back through the column twice, followed by a H_2O wash. The column was then eluted with portions of: (1) 30% Me_2CO/H_2O; (2) 75% Me_2CO/H_2O; and (3) Me_2CO to generate the three screening fractions. NMR spectra of the 30% and 75% screening fractions were obtained using CD_3OD while $CDCl_3$ was used for the 100% fraction.

***Polyandrocarpa polypora* –** The tunicate *P. polypora* (50 g) was extracted in MeOH (200 mL) twice overnight in the dark (due to concerns regarding possible photochemical instability) and screened as above. The 75% screening fraction was then purified on semi-preparative C_{18} HPLC, using a linear gradient from 10% ACN/H_2O to 100% ACN over 60 minutes, and the major compound (**27**) (2 mg, t_R = 15.9 min) was analysed by NMR before decomposition.

Mass Spectrometry-Guided Screening – Each tunicate (~5 g) was extracted in MeOH (40 mL) overnight, and then filtered, dried and weighed. The extracts were then reconstituted in MeOH to 1 mg mL^{-1}, with 1 mL then passed through a syringe filter prior to LCMS analysis. The fractions were analysed using the mass spectrometer, operating in both (+)- and (–)-polarity at a mass range of *m/z* 50–1500. Instrumental parameters for data acquisition were set as follows: capillary voltage of 3500 V, nebuliser gas (N_2) pressure of 30 psig, ion source temperature of 275 °C, sheath gas temperature of 300 °C and flow of 7 L min^{-1}, and the acquisition rate was three spectra s^{-1}. The ion isolation width was 4 Da. Minutes 0–0.5 were sent to waste and minutes 0.5–25 recorded with untargeted ion fragmentation (auto-MS/MS), where the CID energies were dependent on the precursor mass determined by the pre-set gradient (Figure 2.10). The five most intense ions per MS scan were subjected to CID and were actively excluded after three spectra for 0.3 minutes.

The samples were injected (injection volume: 10 μL) into the system equipped with an Eclipse Plus reversed-phase C_{18} column (30 mm x 2.1 mm, 3.5 μm; Agilent Technologies) at 35 °C. Compound separation was achieved using mobile phase A 99.9% H_2O/0.1% NH_4HCO_2 and B 99.9% ACN/0.1% FA, at 0.4 mL/min with the following gradient method: 0-1 min 2.5% B, 1-20 min 100% B, 20-25 min 100% B and a column re-equilibration period for 3 min.

All LCMS/MS data were converted to the MGF file format using the Agilent Qualitative analysis B.04.00 software and uploaded to the GNPS webserver using FileZilla, along with a meta-data file containing the source organism and ionisation polarity. A molecular network were created using the online workflow at GNPS.[77] The data was filtered by removing all MS/MS peaks within ± 17 Da of the precursor *m/z*. MS/MS spectra were window filtered by choosing only the top six peaks in the ± 50 Da window throughout the spectrum. The data were

then clustered with MS-Cluster with a parent mass tolerance of 2.0 Da and a MS/MS fragment ion tolerance of 0.5 Da to create consensus spectra. Further, consensus spectra that contained less than two spectra were discarded. A network was then created where edges were filtered to have a cosine score above 0.6 and more than three matched peaks. Further edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. The spectra in the network were then searched against GNPS' spectral libraries. The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7 and at least six matched peaks. The data can be found and accessed at: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=89718130592941eab01ef07c0afbd978. The output molecular networking data was visualised using Cytoscape (version 3.7.0).

***Cystodytes aucklandicus* –** The tunicate *C. aucklandicus* (15 g) was extracted in MeOH (80 mL) twice overnight, and in addition to the MS method, the ^{1}H NMR spectroscopy screening method was followed using a 20 mL HP20 column. The three fractions were all analysed by ^{1}H NMR spectroscopy in DMSO-d_6.

***Distaplia stylifera* –** The tunicate *D. stylifera* (20 g) was extracted in MeOH (100 mL) twice overnight, and in addition to the MS method, the ^{1}H NMR spectroscopy screening method was followed using a 20 mL HP20 column. The 30% Me_2CO/H_2O screening fraction was then purified on semi-preparative C_{18} HPLC, using a linear gradient from 10% MeOH/H_2O (0.2% FA) to 100% MeOH (0.2% FA) over 60 minutes, collecting 11 fractions, with the major UV/vis peak affording **28** (0.95 mg, t_R = 15.7 min).

(4-Hydroxy-3-methoxyphenyl)-(1H-imidazol-5-yl) methanone (**28**) white film; UV (MeOH/H_2O) λ_{max} 248, 275, 301, 335 nm; ^{13}C and ^{1}H NMR data, Table 2.1; HRESIMS *m/z* 219.0794 $[M + H]^+$ (calcd. for $C_{11}H_{11}N_2O_3$, 219.0764); HRESIMS/MS (20 eV) *m/z* (%) relative intensity) 151.0480 (100), 123.0518 (11.1), 95.0302 (44.0) and 69.0498 (8.3).

Bioassay – MIC studies were performed on **28** using an established protocol[347] to test isolated compounds against test strains of *E. coli* BL21 and *B. subtilis* 168. All tests were carried out in MH medium at 30 °C, with 16 h growth time in biological triplicates (n = 3 independent experiments). A positive control of tetracycline and negative control DMSO was included in the analysis. All compounds were assayed from 128 μg mL^{-1} in two-fold serial dilution steps across 10 dilution stages to the lowest concentration tested 0.25 μg mL^{-1}. Sterility (media only) and growth (cells only) controls were included in each plate assayed.

Chapter 3 – "*Plocamium Angustum*"

Algal Material – The specimen identified as *Plocamium angustum* were collected by hand using SCUBA at a depth of 3−10 m from Moa Point, Wellington, NZ, in January 2017 and stored at −20 °C until extraction. A voucher specimen (JB06_38) is held at the School of Chemical and Physical Sciences, VUW, NZ.

Extraction and Isolation – A frozen sample (50.0 g wet weight) was extracted in MeOH (200 mL) twice overnight. The second extract, followed by the first, were passed through a HP20 column (30 mL), pre-equilibrated in H_2O and combined following elution. The eluent was then diluted with an equal volume of H_2O and passed back through the column twice, followed by a 100 mL H_2O wash. The column was then eluted with 100 mL portions of: (1) 30% Me_2CO/H_2O; (2) 75% Me_2CO/H_2O; and (3) Me_2CO (fractions A1−A3, respectively). Fraction A2 (300 mg) was then partitioned on LH20 resin with 100% MeOH, and the resulting fractions were pooled together on the basis of TLC (1:3 EtOAc:hexanes), resulting in two major fractions B1 (test tubes 36–43) and B2 (test tubes 44–51). A portion of sample B1 (20 mg) was subjected to silica gel chromatography (5:1 EtOAc:hexanes) to afford **47** (6.8 mg) as a colourless oil. A portion of sample B2 (10 mg) was further purified on a semipreparative C18 HPLC column (80% $MeOH/H_2O$), yielding compound **46** (4.8 mg, t_R = 13.0 min) as a pale–yellow oil.

Costatone C (**46**): yellow oil, $[\alpha]_D^{20}$ −7.2 (c 0.05, $CHCl_3$); IR υ (thin film): 2930, 1642, 1434, 1382, 1102, 1080, 782, 738, 587, 519 cm^{-1}; ^{13}C and ^1H NMR data, Table 3.2; HRAPCIMS m/z 376.8719 $[M - H]^-$ (calcd. for $C_{10}H_{13}OCl_2Br_2$, 376.8716).

(1E,3S,5Z)–1,6–Dichloro–2–methylhepta–1,5–dien–3–ol (**47**): colourless oil, $[\alpha]_D^{20}$ −22 (c 0.1, $CHCl_3$); NMR and MS data consistent with published.[159,160]

Mosher Ester Analysis – A solution of EDC.HCl (12 mg, 64 μmol), *S*-(+)-α-MTPA (15 mg, 64 μmol), and DMAP (14.8 mg, 120 μmol) was stirred in dry DCM (0.5 mL) under Ar at 0 °C for 10 min after which a sample of crude **47** (fraction B1, 4 mg, 12 μmol) in DCM (0.5 mL) was added. The solution was allowed to come to room temperature and stirred under Ar for 48 h. DCM (10 mL) was added and the mixture was washed in turn with 10% HCl (10 mL), H_2O (10 mL), sat. $NaHCO_3$ (10 mL) and H_2O (10 mL) before being dried under reduced pressure. The sample was purified by flash silica gel chromatography (10:1 hexanes:EtOAc) to yield the crude product **47a**, analysed without further purification. The procedure was repeated with *R*-(–)-α-MTPA to yield product **47b**.

Compound **47a**: ^{1}H NMR ($CDCl_3$, 600 MHz) δ_H 7.48–7.40 (5H, m, aromatics); 6.25 (1H, s, H–1); 5.50 (1H, dd, H–3); 5.18 (1H, t, H–5); 3.51 (3H, s, OCH_3); 2.56 (2H, m, H–4); 2.03 (3H, s, H–7); 1.75 (3H, d, H–8).

Compound **47b**: ^{1}H NMR ($CDCl_3$, 600 MHz) δ_H 7.48–7.40 (5H, m, aromatics); 6.16 (1H, s, H–1); 5.48 (1H, dd, H–3); 5.32 (1H, t, H–5); 3.55 (3H, s, OCH_3); 2.60 (2H, m, H–4); 2.09 (3H, s, H–7); 1.64 (3H, s, H–8).

Attempted Acetylation – 15 mg of crude **46** (fraction B3) was stirred in pyridine (0.5 mL) for 10 min, before addition of acetic anhydride (0.5 mL). The reaction was stirred for 2 h, before quenching with H_2O (1 mL). The reaction mixture was passed through a plug of HP20ss (2 mL), diluted with H_2O (2 mL) and passed through again, before stripping the plug with Me_2CO (10 mL). The Me_2CO eluent was dried and analysed by NMR, where only unchanged **46** was observed.

Computational Analysis – All computations were performed by Dr Muhammad Ali Hashmi (University of Education, Attock, Pakistan) and Zaineb Sohail (University of Management and Technology, Lahore, Pakistan) using Gaussian 09 (Revision D.01).[348] Density functional theory (DFT) was used for all the calculations utilising Adamo's hybrid[321] version of Perdew, Burke and Ernzerhof functional (PBE0)[322,323] along with the application of Grimme's empirical dispersion correction (D3) with Becke-Johnston damping (D3BJ).[349-351] All calculations were performed with Ahlrich's triplet ζ basis set def2-TZVP[352] supported by the polarisable continuum model with the integral equation formalism variant (IEFPCM) for solvation modelling.[353-359] The solvent for optimisation and ECD calculation was MeOH which was modelled with the SMD parameter set by Cramer and Truhlar[360] (as implemented in Gaussian 09).[348] Calculated ECD spectra were scaled for both intensity and frequency to the experimental data. Frequency calculations at the same level of theory were used to confirm all the optimised structures to be true minima on the potential energy surface with the absence of imaginary frequencies. The 3D images of optimised molecules were drawn using CYLview program.[361]

Bioassay – *Pseudomonas aeruginosa* (PAO1) or *Staphylococcus aureus* (ATCC 25923) were used to inoculate 100 μL of Mueller Hinton broth amended with 100 μg mL^{-1} of the test compounds in a 96-well plate (control wells contained an equivalent volume of DMSO). Cells were incubated at 37 °C, shaking at 600 RPM for 24 h. The optical density was measured at

600 nm and the absorbance value of the media–only controls were averaged and subtracted from all measurements. Values were calculated from three replicates.

S. aureus and *S. epidermidis* (ATCC 35984) were then tested with **46** to determine the strength of inhibition in Gram-positive bacteria. Similar to the previous experiment, *S. aureus* and *S. epidermidis* were used to inoculate 100 μL of Mueller Hinton broth, amended with a two-fold dilution series of **46** from 0.5 μg mL^{-1} to 128 μg mL^{-1} in a 96–well plate (control wells contained an equivalent volume of DMSO). Cells were incubated at 37 °C, shaking at 600 RPM for 24 h. The optical density was measured at 600 nm and the absorbance value of the media-only controls were averaged and subtracted from all measurements. Values were calculated from three replicates.

Taxonomic analysis – (Completed by Joe Zuccarello) DNA extraction, PCR amplification, and sequencing of the cytochrome oxidase genes followed previously described method.[148] Various sequences of *Plocamium* were downloaded from Genbank or were gained directly from Cooper.[148] Phylogenetic trees were made using RAxML 8[362] to construct maximum-likelihood trees to show the most likely tree from the dataset. RAxML was performed using the GTR+gamma model. The reliability of the maximum-likelihood topologies was evaluated based on 1000 nonparametric bootstrap replicates.[363]

Chapter 4 – *Nelliella nelliiformis*

Animal Material – The marine bryozoan *Nelliella nelliiformis* was hand collected using SCUBA from ʻEua, Kingdom of Tonga (21°23′ S, 174°56′ W), in June 2016. Samples were stored at −18 °C until required for extraction. The sample was identified by an expert bryozoologist (Dennis P. Gordon, NIWA) by examination and comparison of SEM images with original descriptions and illustrations of the type specimen (Figure 4.1). A voucher sample (NIWA 127738) is held at NIWA.

Extraction and Isolation – Frozen bryozoan (40.0 g wet weight) was cut into small pieces and extracted in MeOH (100 mL) twice overnight. The second extract, followed by the first, were passed through a HP20 column (40 mL), pre-equilibrated in H_2O and combined following elution. The eluent was then diluted with an equal volume of H_2O and passed back through the column twice, followed by a 30 mL H_2O wash. The column was then eluted with 120 mL portions of (1) 30% Me_2CO/H_2O, (2) 75% Me_2CO/H_2O, and (3) Me_2CO (fractions A1−A3, respectively). Fraction A1 was then purified on a semi-preparative C18 HPLC column, using

a linear gradient from 10% MeOH/H_2O (0.2% FA) to 100% MeOH/H_2O (0.2% FA) over 60 min, generating fractions B1–B6. Fractions B1 and B2 were further purified by analytical C18 HPLC (same method) to afford **61** (3.2 mg, t_R = 16.0 min) and **62** (1.5 mg t_R = 18.2 min), respectively.

Nellielloside A (**61**): white solid; $[\alpha]_D^{20}$ –24 (*c* 0.2, MeOH); UV (MeOH) λ_{max} (log ε) 263 (3.87) nm; IR (film) v_{max} 3484, 3182, 1686, 1674, 1408, 1336, 1123, 1077 cm^{-1}; ^{13}C and ^{1}H NMR data, Table 4.1; HRESIMS *m/z* 361.1250 $[M + H]^+$ (calcd. for $C_{15}H_{17}N_6O_5$, 361.1255); HRESIMS/MS (20 eV) *m/z* (% relative intensity) 136.0617 (100), 119.0349 (20), 94.0298 (6.8).

Nellielloside B (**62**): white solid; $[\alpha]_D^{20}$ –21 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε) 264 (3.19) nm; IR (film) v_{max} 3238, 1680, 1587, 1554, 1451, 1315, 1126, 1076 cm^{-1}; ^{13}C and ^{1}H NMR data, Table 4.2; HRESIMS *m/z* 362.1096 $[M + H]^+$ (calcd. for $C_{15}H_{16}N_5O_6$, 362.1095); HRESIMS/MS (20 eV) *m/z* (% relative intensity) 137.0456 (85.2), 94.0299 (100).

Preparation of Peracetylated Aldononitrile Derivatives – A portion of isolated **61** (0.2 mg) was hydrolysed by heating at 90 °C for 8 h in HCl (1 M, 1 mL), and the reaction was then dried under reduced pressure. The hydrolysis products, D/L-ribose, D-arabinose, D-lyxose and D-xylose were each separately dissolved in H_2O (0.2 mL), to which a solution of $NH_2OH.HCl$ (10 mg) dissolved in *N*–methylimidazole (0.4 mL) was added. All the solutions were stirred at 80 °C for 10 min, then Ac_2O (4 x 0.25 mL) was added over 1 min after which the reaction was stirred for 2 h (80 °C). Each reaction mixture was extracted with DCM (5 mL) and the extract washed sequentially with H_2O (3 mL), sat. $NaHCO_3$(*aq*) (3 mL), 10% H_2SO_4(*aq*) (3 mL), and finally H_2O (2 × 3 mL). The organic layers were dried and taken up in $CHCl_3$ (0.5 mL) for GC analysis.

GC-MS analyses used helium as the carrier gas. Mass spectra were obtained at 70 eV in positive ion mode, scanning from *m/z* 40–600 every 0.3 s. The ion source was held at 200 °C while the MS-transfer line was at 305 °C. Compound identity was determined using both retention time and mass spectral fragmentation pattern. Samples were introduced (1 μL) into a glass split/splitless line at 270 °C.

Non-chiral-phase gas chromatographic separations were performed using a Restek RXI-5Sil-MS column (30 m × 0.25 mm × 0.25 μm) with splitless injection using constant carrier gas flow (linear velocity 43.4 cm s^{-1}; 1.38 mL min^{-1}). The initial oven temperature was 50 °C, held

for 2 min after which a temperature ramp of 10 °C min^{-1} to 300 °C was used, with a final hold of 5 min (total analysis time 32 min). The retention time of derivatised authentic standards (arabinose: 15.99 min; lyxose: 15.88 min; ribose: 15.76 min; xylose: 16.15 min) was used to compare to the nellielloside A hydrolysis product (15.74 min), establishing the furanose as ribose.

For chiral-phase gas chromatographic separations, the method from Keyzers was followed.[192,193] Chiral-phase GC analyses of the derivatised D- and L-ribose and hydrolysate prepared from **61** were performed on a CP-ChiraSil-L-Val capillary column (25 m × 0.25 mm × 0.12 μm; Agilent)[193] with a 5:1 split injection of 1 μL of derivatised sample with constant flow control (linear velocity 48.2 cm s^{-1}; 1.40 mL min^{-1}). The initial oven temperature was 100 °C, held for 1 min, after which it was ramped at 2 °C min^{-1} to 180 °C, at which point it was held for a further 2 min. The peracetylated aldononitrile derivatives of the hydrolysed natural product, D- and L-ribose, were eluted with retention times of 24.91 and 24.80 min, respectively. The hydrolysed ribose from **61** eluted with a retention time of 24.92 min.

Synthetic Procedures – To a solution of the appropriate nucleoside (**70**–**72**) (500 mg, 1.87 mmol) in distilled Me_2CO (20 mL) under argon were added *p*-toluenesulfonic acid (388 mg, 2.26 mmol) and 2,2–dimethoxypropane (0.916 mL, 7.48 mmol). The solution was stirred for three days at room temperature (rt), upon which TLC analysis (monitored by UV, eluent 15% MeOH/CH_2Cl_2) indicated no starting material remained. The reaction was then quenched with sat. $NaHCO_3$ (*aq*) (30 mL) and extracted with $CHCl_3$ (3 × 20 mL) to afford the corresponding protected acetonides **73**–**75** (quantitative yield).[364]

To a solution of the acetonide-protected alcohols **73**–**75** (50 mg, 0.163 mmol) and pyrrole-2-carboxylic acid (**76**) or pyrrole-3-carboxylic acid (**77**) (18.1 mg, 0.163 mmol) in dry DCM (10 mL) under Ar, were added DCC (40.4 mg, 0.196 mmol) and DMAP (6.0 mg, 0.049 mmol). The reaction was stirred at rt for 4 h, after which it was quenched with sat. $NaHCO_3$ (*aq*) (20 mL). The solution was extracted with DCM (2 × 20 mL) and concentrated under reduced pressure to afford the corresponding pyrrole-coupled acetonides. These protected products were immediately reacted with 3 mL of TFA for 1 h at 40 °C, then concentrated to dryness under reduced pressure. The residues were loaded onto HP20ss plugs (2 mL), washed with 6 mL of H_2O and then eluted with 6 mL of 50% MeOH/H_2O. The crude material was then purified by semi-preparative HPLC (using the same method as for isolation-derived samples), to afford pyrrole nucleosides **61**, **62**, **66–69** (4–10% yield).

Nellielloside A (**61**): white solid; $[\alpha]_D^{20}$ –25 (*c* 0.1, MeOH); spectroscopic data matched those of the natural material, as reported above.

Nellielloside B (**62**): white solid; $[\alpha]_D^{20}$ –24 (*c* 0.1, MeOH); spectroscopic data matched those of the natural material, as reported above.

Guanosyl-6′-pyrrole-2-carboxylate (**66**): white solid; $[\alpha]_D^{20}$ –15 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε) 262 (3.91) nm; IR (film) ν_{max} 3324, 3113, 1676, 1555, 1409, 1311, 1123, 1025 cm^{-1}; ^{13}C and ^{1}H NMR data, Appendix 6 – Table A1; HRESIMS *m/z* 377.1200 $[M + H]^+$ (calcd. for $C_{15}H_{17}N_6O_6$, 377.1204); HRESIMS/MS (20 eV) *m/z* (% relative intensity) 152.0558 (100), 135.0302 (20), 94.0288 (13.3).

Adenosyl–6′-pyrrole-3-carboxylate (**67**): white solid; $[\alpha]_D^{20}$ –25 (*c* 0.2, MeOH); UV (MeOH) λ_{max} (log ε) 258 (3.82 nm); IR (film) ν_{max} 3362, 3138, 1671, 1604, 1332, 1149, 1048 cm^{-1}; ^{13}C and ^{1}H NMR data, Appendix 6 – Table A2; HRESIMS *m/z* 361.1253 $[M + H]^+$ (calcd. for $C_{15}H_{17}N_6O_5$, 361.1255); HRESIMS/MS (20 eV) *m/z* (% relative intensity) 136.0618 (100), 94.0297 (25.7).

Inosyl-6′-pyrrole-3-carboxylate (**68**): white solid; $[\alpha]_D^{20}$ –7 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε) 249 (3.30) nm; IR (film) ν_{max} 3336, 2933, 1696, 1592, 1416, 1331, 1178, 1091 cm^{-1}; ^{13}C and ^{1}H NMR data, Appendix 6 – Table A3; HRESIMS *m/z* 362.1088 $[M + H]^+$ (calcd. for $C_{15}H_{16}N_5O_6$, 362.1095); HRESIMS/MS (20 eV) *m/z* (% relative intensity) 137.0442 (100), 94.0275 (19.2).

Guanosyl-6′-pyrrole-3-carboxylate (**69**): white solid; $[\alpha]_D^{20}$ –27 (*c* 0.1, MeOH); UV (MeOH) λ_{max} (log ε) 254 (4.12) nm; IR (film) ν_{max} 3308, 3116, 1688, 1603, 1534, 1329, 1169, 1083 cm^{-1}; ^{13}C and ^{1}H NMR data, Appendix 6 – Table A4; HRESIMS *m/z* 377.1207 $[M + H]^+$ (calcd. for $C_{15}H_{17}N_6O_6$, 377.1204); HRESIMS/MS (20 eV) *m/z* (% relative intensity) 152.0571 (100), 135.0306 (17.3), 94.0302 (30.8).

Bioassay – The kinase inhibitory activity of nellielloside A (**61**) was assayed against 485 human-disease relevant kinases at 10 μM, using a commercial service (ThermoFisher Life Technologies SelectScreen® Whole Panel ACCESS Program, Wisconsin, USA. (https://www.thermofisher.com/nz/en/home/products–and–services/services/customservices/screening–and–profiling–services/selectscreen–profiling–service/selectscreen–kinaseprofiling–service.html). Of the successful hits, representatives of

kinase families were selected for single point assay (at 10 μM) using compounds **62** and **66**–**69**. Finally, IC_{50} values were generated for all six compounds against GSK3A, MAPK14 and RSK2 using 10 concentration points at the same commercial service. All measurements were made using solutions of the compounds in 100% DMSO.

Chapter 5 – *Synoicum kuranui*

Animal Material – The tunicate *Synoicum kuranui* was hand collected using SCUBA from Great Barrier Island, NZ in 1999 as part of the NIWA collection. The tunicate was identified by Michael J. Page (NIWA). Samples were stored at −18 °C until required.

LCMS/MS Analysis and Molecular Networking – The HP20 screening fractions (A1 and A3, 1 mg mL^{-1} in MeOH, see below) were analysed using the aforementioned screening method, however only operating with a negative polarity at a mass range of *m/z* 50–1500.

All LCMS/MS data was converted to the MGF file format using the Agilent Qualitative analysis B.04.00 software and uploaded to the GNPS webserver. A molecular network was created using the online workflow at GNPS.[77] The data was filtered by removing all MS/MS peaks within ± 17 Da of the precursor *m/z*. MS/MS spectra were window filtered by choosing only the top six peaks in the ± 50 Da window throughout the spectrum. The data was then clustered with MS-Cluster with a parent mass tolerance of 2.0 Da and a MS/MS fragment ion tolerance of 0.5 Da to create consensus spectra. Further, consensus spectra that contained less than two spectra were discarded. A network was then created where edges were filtered to have a cosine score above 0.6 and more than four matched peaks. Further edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. The spectra in the network were then searched against GNPS' spectral libraries. The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7 and at least six matched peaks. The data can be found and accessed at: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=0faf2a77b21b457c9a047343e9b813e0#. The output molecular networking data was visualised using Cytoscape (version 3.7.0).

Extraction and Isolation – The tunicate (24 g wet weight) was extracted in MeOH (100 mL) twice overnight. The second extract, followed by the first, were passed through a HP20 column (20 mL), pre-equilibrated in H_2O and combined following elution. The eluent was then diluted with an equal volume of H_2O and passed back through the column twice, followed by a 60 mL

H_2O wash. The column was then eluted with 60 mL portions of (1) 30% Me_2CO/H_2O, (2) 75% Me_2CO/H_2O, and (3) Me_2CO (fractions A1−A3, respectively). Fractions A2 and A3 were then purified on a semi-preparative C_{18} HPLC column, using isocratic 90% MeOH/H_2O (0.2% FA), generating fractions B1–B4 and C1–C7 respectively. Fractions B3 and C7 solely afforded rubrolides **108** (0.65 mg, t_R = 6.3 min) and **110** (0.15 mg, t_R = 9.9 min) respectively. Fractions B4 and C5 were further purified by on an analytical C_{18} HPLC column using 80% MeOH/H_2O (0.2% FA) to afford rubrolide *E*-**111** (0.05 mg, t_R = 6.3 min), *Z*-**111** (0.26 mg, t_R = 9.9 min) and **109** (0.35 mg, t_R = 13.1 min). The two separate **111** samples were combined for NMR analysis.

Rubrolide A (**108**) yellow film; HRESIMS *m/z* 590.7075 $[M - H]^-$ (calcd. for $C_{17}H_7Br_4O_4$, 590.7083); all NMR data matches those previously reported.[263]

Rubrolide B (**109**) yellow film; HRESIMS *m/z* 624.6694 $[M - H]^-$ (calcd. for $C_{17}H_6Br_4ClO_4$, 624.6694); all NMR data matches those previously reported.[263]

Rubrolide T (**110**); yellow film; UV (MeOH/H_2O) λ_{max} 230, 250, 257 (sh), 340 nm; ^{13}C and ^1H NMR data, Table 5.1; HRESIMS *m/z* 604.7254 $[M - H]^-$ (calcd. for $C_{18}H_9Br_4O_4$, 604.7240); HRESIMS/MS (50 eV) *m/z* (% relative intensity) 589.6992 (5.4), 545.7086 (3.1), 510.7802 (20.8), 482.7853 (14.5), 454.7901 (5.0), 402.8585 (8.3), 287.8473 (18.8), 272.8544 (75), 78.9188 (100).

E/Z–Rubrolide U (**111**); yellow film; UV (MeOH/H_2O) λ_{max} 232, 254, 361 nm; ^{13}C and ^1H NMR data, Table 5.2/5.3; HRESIMS *m/z* 526.8143 $[M - H]^-$ (calcd. for $C_{18}H_{10}Br_3O_4$, 526.8135); HRESIMS/MS (50 eV) *m/z* (% relative intensity) 511.7898 (3.4), 483.7952 (3.0), 467.8005 (3.3), 432.8721 (12.5), 404.8767 (5.0), 375.8743 (20.0), 295.9485 (22.5), 272.8561 (50), 78.9194 (100).

Bioassay – MIC studies were performed on **108**–**111** using an established protocol[347] to test isolated compounds against test strains of *E. coli* BL21 and *B. subtilis* 168. All tests were carried out in MH medium at 30 °C, with 16 h growth time in biological triplicates (n = 3 independent experiments). A positive control of tetracycline and negative control DMSO was included in the analysis. All compounds were assayed from 128 μg mL^{-1} in two-fold serial dilution steps across 10 dilution stages to the lowest concentration tested 0.25 μg mL^{-1}. Sterility (media only) and growth (cells only) controls were included in each plate assayed.

Chapter 6 – *Didemnum ternerratum*

Animal Material – The marine tunicate *Didemnum ternerratum* was hand collected using SCUBA from ʻEua, Kingdom of Tonga, in June 2016. The tunicate was identified by Michael J. Page (NIWA). Samples were stored at −18 °C until required. A voucher sample (PTN4_36E) is held at NIWA, Nelson.

LCMS/MS Analysis and Molecular Networking – The HP20 screening fractions (A1–A3, 1 mg mL^{-1} in MeOH) were analysed using the aforementioned mass spectrometer, operating with a negative polarity at a mass range of *m/z* 50–1500. Instrumental parameters for data acquisition were set as follows: capillary voltage of 3500 V, nebuliser gas (N_2) pressure of 30 psig, ion source temperature of 275 °C, sheath gas temperature of 300 °C and flow of 7 L min^{-1} and the acquisition rate was three spectra s^{-1}. The isolation width was 4 Da. Minutes 0–0.5 were sent to waste and minutes 0.5–25 recorded with auto-MS/MS, where the CID energy was set at 30 eV. The precursor selection window was set to the optimised *m/z* 400–700, and the five most intense ions per MS scan were subjected to CID and were actively excluded after three spectra for 0.3 minutes.

The samples were injected (injection volume: 10 μL) into the system equipped with an Eclipse Plus reversed-phase C_{18} column (30 mm × 2.1 mm, 3.5 μm; Agilent Technologies) at 35 °C. Metabolite separation was achieved using mobile phase A 99.9% H_2O/0.1% NH_4HCO_2 and B 99.9% ACN/0.1% FA, pumped at 0.4 mL min^{-1} for the following gradient method: 0–1 min 2.5% B, 1–20 min 50% B, 20–25 min 100% B and a column re-equilibration period for 3 min.

All LCMS/MS data was converted to the MGF file format using the Agilent Qualitative analysis B.04.00 software, and uploaded to the Global Natural Products Social Molecular Networking webserver. A molecular network was created using the online workflow.[77] The data was filtered by removing all MS/MS peaks within ± 17 Da of the precursor *m/z*. MS/MS spectra were window filtered by choosing only the top six peaks in the ± 50 Da window throughout the spectrum. The data was then clustered with MS-Cluster with a parent mass tolerance of 2.0 Da and a MS/MS fragment ion tolerance of 0.5 Da to create consensus spectra. Further, consensus spectra that contained less than two spectra were discarded. A network was then created where edges were filtered to have a cosine score above 0.6 and more than five matched peaks. Further edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. The spectra in the network were then searched against GNPS spectral libraries. The library spectra were filtered

in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7 and at least six matched peaks. The data can be found and accessed at: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f7150c51146f481da039d8dbcd7cb706. The output molecular networking data was visualised using Cytoscape (version 3.7.0).[33]

Extraction and Isolation – The tunicate (2 g wet weight) was extracted in MeOH (50 mL) twice overnight. The second extract, followed by the first, were passed through a HP20 column (10 mL), pre-equilibrated in H_2O and combined following elution. The eluent was then diluted with an equal volume of H_2O and passed back through the column twice, followed by a 30 mL H_2O wash. The column was then eluted with 30 mL portions of (1) 30% Me_2CO/H_2O, (2) 75% Me_2CO/H_2O, and (3) Me_2CO (fractions A1−A3, respectively). Fractions A1 and A2 were each dry-loaded onto a 10 mL HP20ss column, eluted with 30 mL portions of (1) 80% MeOH/H_2O, (2) MeOH, and (3) Me_2CO (fractions B1−B3, and C1–C3, respectively). Fraction B1 was then purified on analytical C_{18} HPLC, using a linear gradient from 10% MeOH/H_2O (1% NH_3) to 70% MeOH/H_2O (1% NH_3) over 40 minutes. Fraction C1 was purified by the same method, and fractions were combined based on retention time and HRESIMS analysis, resulting in 12 fractions (D1–D14).

Further purification of fractions D2-D14 using the same analytical C_{18} HPLC method, afforded eight new lamellarin sulfates, lamellarin B2-20-sulfate (**153**) (0.11 mg, t_R = 20.6 min), lamellarin D-8-sulfate (**151**) (0.18 mg, t_R = 21.2 min), lamellarin D-20-sulfate (**156**) (0.2 mg, t_R = 21.8 min), lamellarin N-20-sulfate (**157**) (0.2 mg, t_R = 22.8 min), lamellarin K-20-sulfate (**147**) (0.18 mg, t_R = 25.0 min), lamellarin E-20-sulfate (**148**) (0.22 mg, t_R = 26.2 min), lamellarin A3-20-sulfate (**149**) (0.69 mg, t_R = 29.5 min), and lamellarin B1-20-sulfate (**150**) (0.21 mg, t_R = 33.0 min). During this HPLC procedure, three previously reported lamellarin sulfates were also purified, lamellarin Y-20-sulfate (**158**) (0.5 mg, t_R = 27.6 min), lamellarin U-20-sulfate (**159**) (0.2 mg, t_R = 32.0 min) lamellarin α-20-sulfate (**160**) (0.2 mg, t_R = 34.0 min).

Fraction D1 was further purified by isocratic 35% MeOH/H_2O (1% NH_3) using an analytical C_8 column and afforded two new compounds, lamellarin A4-13-sulfate (**155**) (1.0 mg, t_R = 17.1 min) and lamellarin H-13-sulfate (**154**) (1.0 mg, t_R = 20.8 min).

Fraction C2 was purified on analytical C_{18} HPLC, using a linear gradient from 10% MeOH/H_2O (1% NH_3) to 70% MeOH/H_2O (1% NH_3) over 40 minutes to afford compound lamellarin K (**161**) (0.96 mg, t_R = 32.0 min) as the major component.

Lamellarin B2-20-sulfate (**153**); white film; UV (MeOH/H_2O) λ_{max} 274, 313, 339 (sh) nm; ^{13}C and ^{1}H NMR data, Table 6.7; HRESIMS *m/z* 566.0758 $[M - H]^-$ (calcd. for $C_{27}H_{20}NO_{11}S^-$, 566.0763); HRESIMS/MS (30 eV) *m/z* (% relative intensity) 486.1185 (58.0), 471.0958 (100), 456.0719 (67.7).

Lamellarin D-8-sulfate (**151**); yellow film; UV (MeOH/H_2O) λ_{max} 228, 256 (sh), 294, 315, 399 nm; ^{13}C and ^{1}H NMR data, Table 6.6; HRESIMS *m/z* 578.0783 $[M - H]^-$ (calcd. for $C_{28}H_{21}NO_{11}S^-$, 578.0763); HRESIMS/MS (30 eV) *m/z* (% relative intensity) 498.1190 (100), 483.0950 (50.0), 468.0723 (7.4).

Lamellarin D-20-sulfate (**156**); yellow film; UV (MeOH/H_2O) λ_{max} 283, 318, 336 (sh), 399 nm; HRESIMS *m/z* 578.0782 $[M - H]^-$ (calcd. for $C_{28}H_{21}NO_{11}S^-$, 578.0763).

Lamellarin N-20-sulfate (**157**); yellow film; UV (MeOH/H_2O) λ_{max} 280, 321, 336 (sh), 399 nm; HRESIMS *m/z* 578.0785 $[M - H]^-$ (calcd. for $C_{28}H_{21}NO_{11}S^-$, 578.0763).

Lamellarin K-20-sulfate (**147**); brown film; UV (MeOH/H_2O) λ_{max} 274, 311, 339 (sh) nm; ^{13}C and ^{1}H NMR data, Table 6.1; HRESIMS *m/z* 610.1024 $[M - H]^-$ (calcd. for $C_{29}H_{24}NO_{12}S^-$, 610.1025); HRESIMS/MS (30 eV) *m/z* (% relative intensity) 530.1607 (100), 515.1370 (59.0), 500.1122 (14.3).

Lamellarin E-20-sulfate (**148**); white film; UV (MeOH/H_2O) λ_{max} 276, 303, 335 (sh) nm; ^{13}C and ^{1}H NMR data, Table 6.2; HRESIMS *m/z* 610.1048 $[M - H]^-$ (calcd. for $C_{29}H_{24}NO_{12}S^-$, 610.1025); HRESIMS/MS (30 eV) *m/z* (% relative intensity) 530.1500 (100), 515.1265 (71.4), 500.1019 (18.1).

Lamellarin A3-20-sulfate (**149**); white film; UV (MeOH/H_2O) λ_{max} 274, 310, 340 (sh) nm; ^{13}C and ^{1}H NMR data, Table 6.4; HRESIMS *m/z* 594.1103 $[M - H]^-$ (calcd. for $C_{29}H_{25}NO_{11}S^-$, 594.1076); HRESIMS/MS (30 eV) *m/z* (% relative intensity) 514.1504 (100), 499.1271 (40.0), 484.1025 (6.8).

Lamellarin B1-20-sulfate (**150**); white film; UV (MeOH/H_2O) λ_{max} 278, 311 (sh), 367, 386 (sh) nm; ^{13}C and ^{1}H NMR data, Table 6.5; HRESIMS *m/z* 592.0948 $[M - H]^-$ (calcd. for

$C_{29}H_{23}NO_{11}S^-$, 592.0919); HRESIMS/MS (30 eV) *m/z* (% relative intensity) 512.1366 (100), 497.1128 (39.3), 482.0883 (8.9).

Lamellarin Y-20-sulfate (**158**); white film; HRESIMS *m/z* 580.0944 $[M - H]^-$ (calcd. for $C_{28}H_{23}NO_{11}S^-$, 580.0919); all NMR data matches those previously reported.[302]

Lamellarin U-20-sulfate (**159**); white film; HRESIMS *m/z* 592.1084 $[M - H]^-$ (calcd. for $C_{29}H_{25}NO_{11}S^-$, 594.1076); all NMR data matches those previously reported.[302]

Lamellarin α-20-sulfate (**160**); white film; HRESIMS *m/z* 592.0921 $[M - H]^-$ (calcd. for $C_{29}H_{23}NO_{11}S^-$, 592.0919); all NMR data matches those previously reported.[303]

Lamellarin A4-13-sulfate (**155**); white film; UV (MeOH/H_2O) λ_{max} 277, 323, 341 (sh) nm; HRESIMS *m/z* 538.0474 $[M - H]^-$ (calcd. for $C_{25}H_{16}NO_{11}S^-$, 538.0450).

Lamellarin H-13-sulfate (**154**); white film; UV (MeOH/H_2O) λ_{max} 284, 315 (sh), 399 nm; ^{13}C and ^1H NMR data, Table 6.8. HRESIMS *m/z* 536.0317 $[M - H]^-$ (calcd. for $C_{25}H_{14}NO_{11}S^-$, 536.0293); HRESIMS/MS (30 eV) *m/z* (% relative intensity) 456.0715 (100), 428.0778 (8.0).

Lamellarin K (**161**); white film; HRESIMS *m/z* 532.1594 $[M + H]^+$ (calcd. for $C_{29}H_{26}NO_9^+$, 532.1602); all NMR data matches those previously reported.[305]

Ethanol Extraction – Two small samples of *D. ternerratum* (~0.1 g) were extracted with either MeOH or EtOH (20 mL). The samples were dried, reconstituted in 1 mL of the extraction solvent and analysed by LCMS/MS as above through the GNPS platform. The data can be found and accessed at: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=689bee3957bd4cf987f875e85cabfac2.

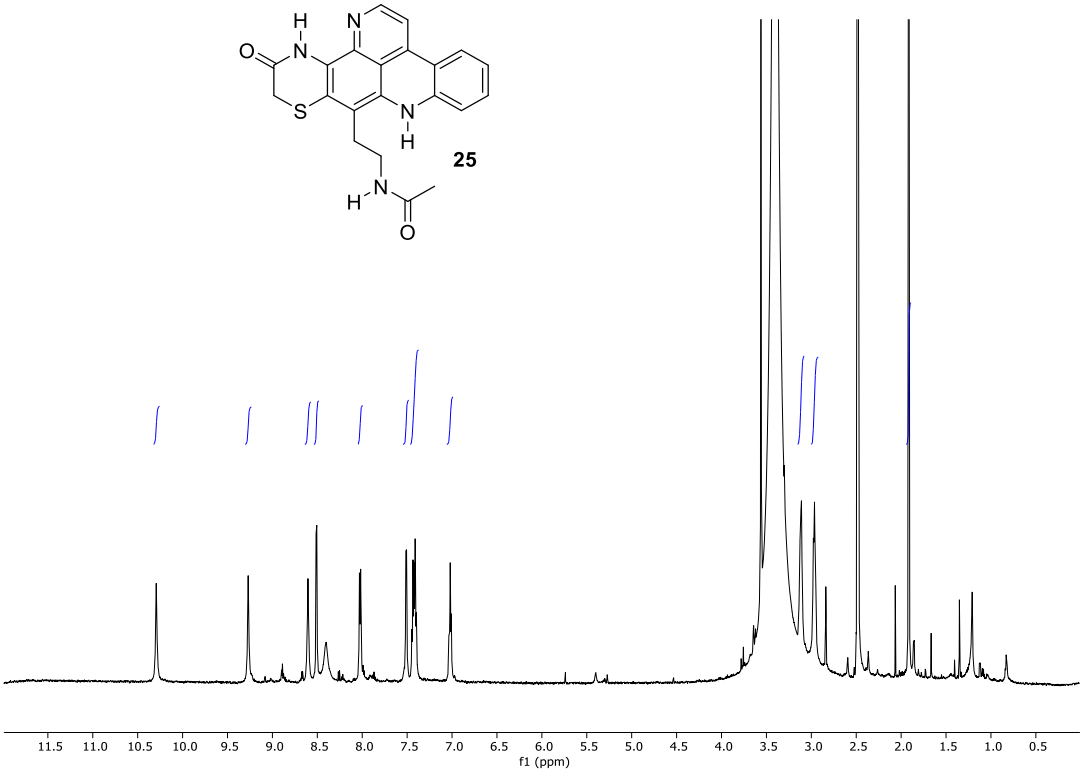
ECD Calculations – The computational procedure used by Dr Luke Robertson (Griffith University, Australia) was based on the method described by Pescitelli and Bruhn.[313] Conformational analysis of **150** was performed with Schrödinger MacroModel 2016 by following the method of Willoughby et al.[365] Initial geometry optimisations were then carried out on each of the 10 generated conformers using density functional theory calculations with the B3LYP/6-31G(d) functional/basis set combination and Grimme's empirical dispersion corrections (D3).[350] Each conformer was then re–optimised at the B3LYP/def2SVP level using empirical dispersion corrections (D3) with the addition of the Polarisable Continuum Model.[358] Single point energy calculations were carried out at the same level. Electronic transition and rational strength were calculated using TDDFT at the CAM-B3LYP/def2SVP level with

consideration of the methanol solvent effect using the Polarisable Continuum Model. Boltzmann-weighting of UV and ECD spectra performed using SpecDis[366] with a half-bandwidth of 0.25 eV. Experimental ECD spectra of **148**–**151** were processed using SDAR.[367] All DFT calculations were performed using GAUSSIAN 16.[368]

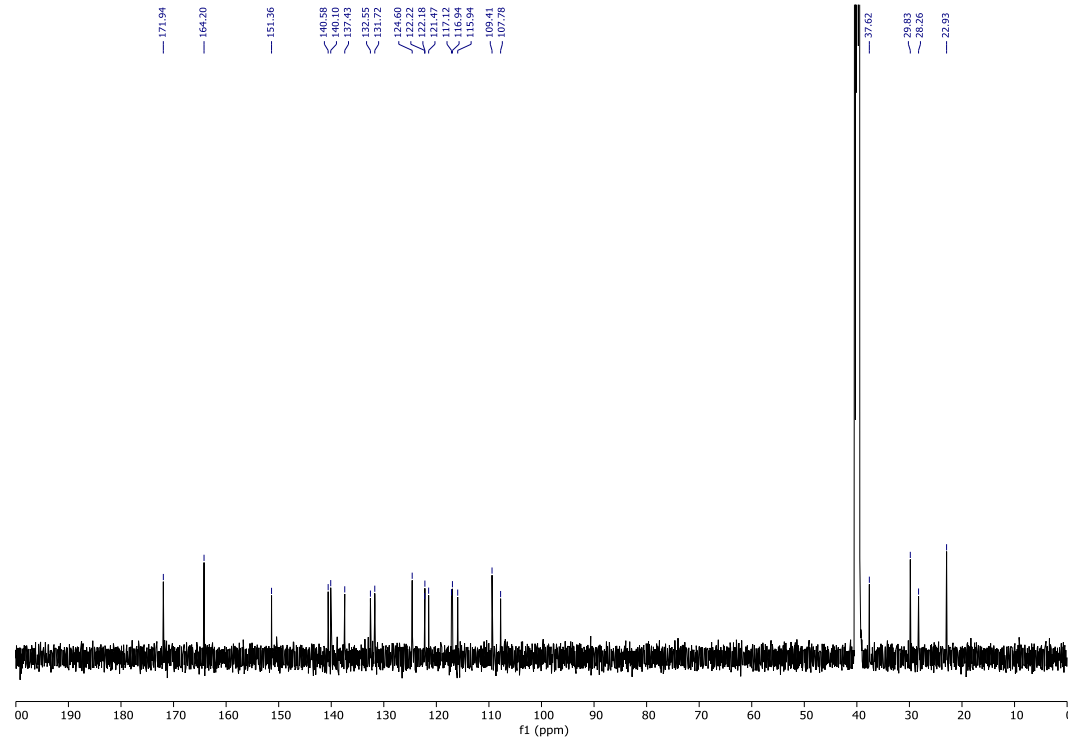
Rotational Transition State Modelling – DFT calculations were performed by Dr Paul Hume (VUW) using the Gaussian 09 software package.[348] All geometry optimisations were performed using the B3LYP hybrid functional[358] and the 6-311G(d,p) basis set, employing very tight convergence criteria, a superfine integration grid and two–electron integral accuracy of 10^{-12}. The self-consistent field convergence threshold was lowered to 10^{-9}. Dispersion effects were incorporated by the use of Grimme's empirical dispersion correction (D3).[350] Solvent effects were incorporated using a Polarisable Continuum Model with MeOH as the solvent. Stationary points were characterised by normal mode vibrational frequency calculations. Transition states were further validated by performing intrinsic reaction coordinate calculations. Predicted free energy values were calculated as the sum of the electronic energy and the thermal correction to the Gibbs energy at 298.15 K. Gibbs free energy calculations were chosen in order to incorporate the effects of entropy on both barrier heights and reaction energies.

Bioassay – A standard 48 h MTS cell proliferation assay was used to evaluate cytotoxic activity against the human colon carcinoma cell line HCT-116 ($n \geq 3$ independent experiments with duplicate wells per experiment). Cells were treated with compound at various concentrations, and a dose-response was generated relative to a control of untreated HCT-116 cells.

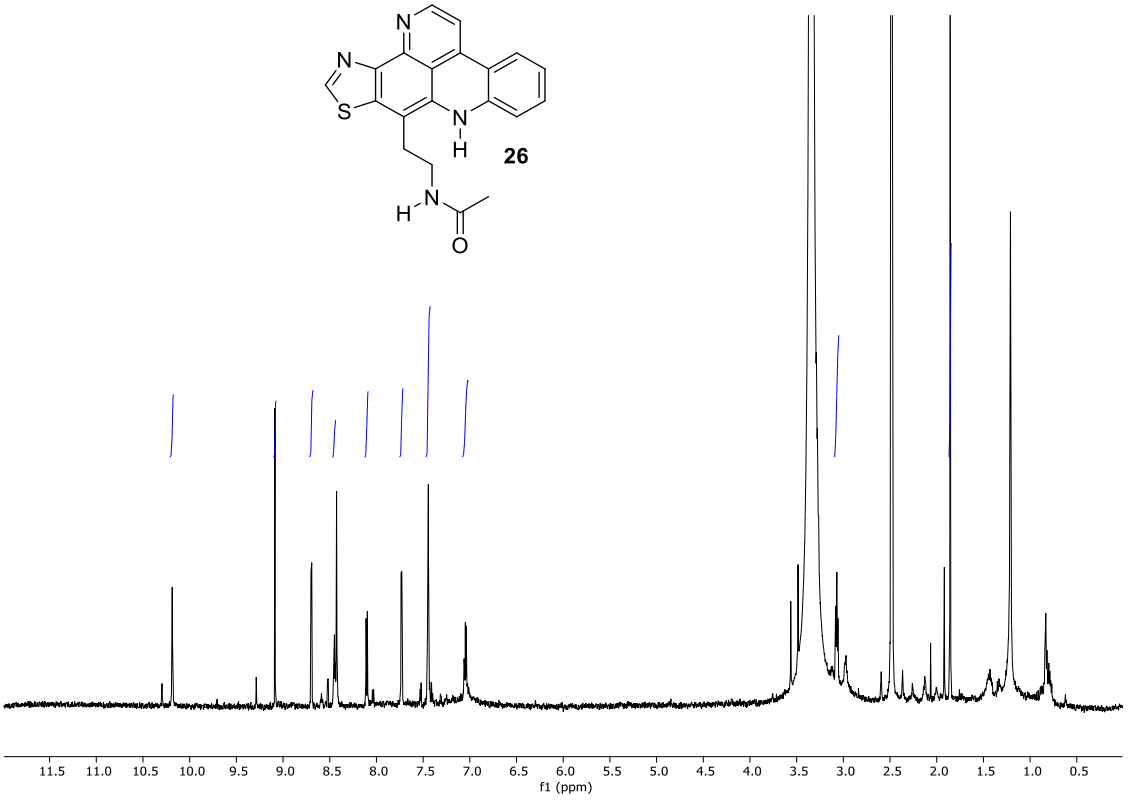
Appendix 1 – Ascidian PTN3_40B NMR Spectra



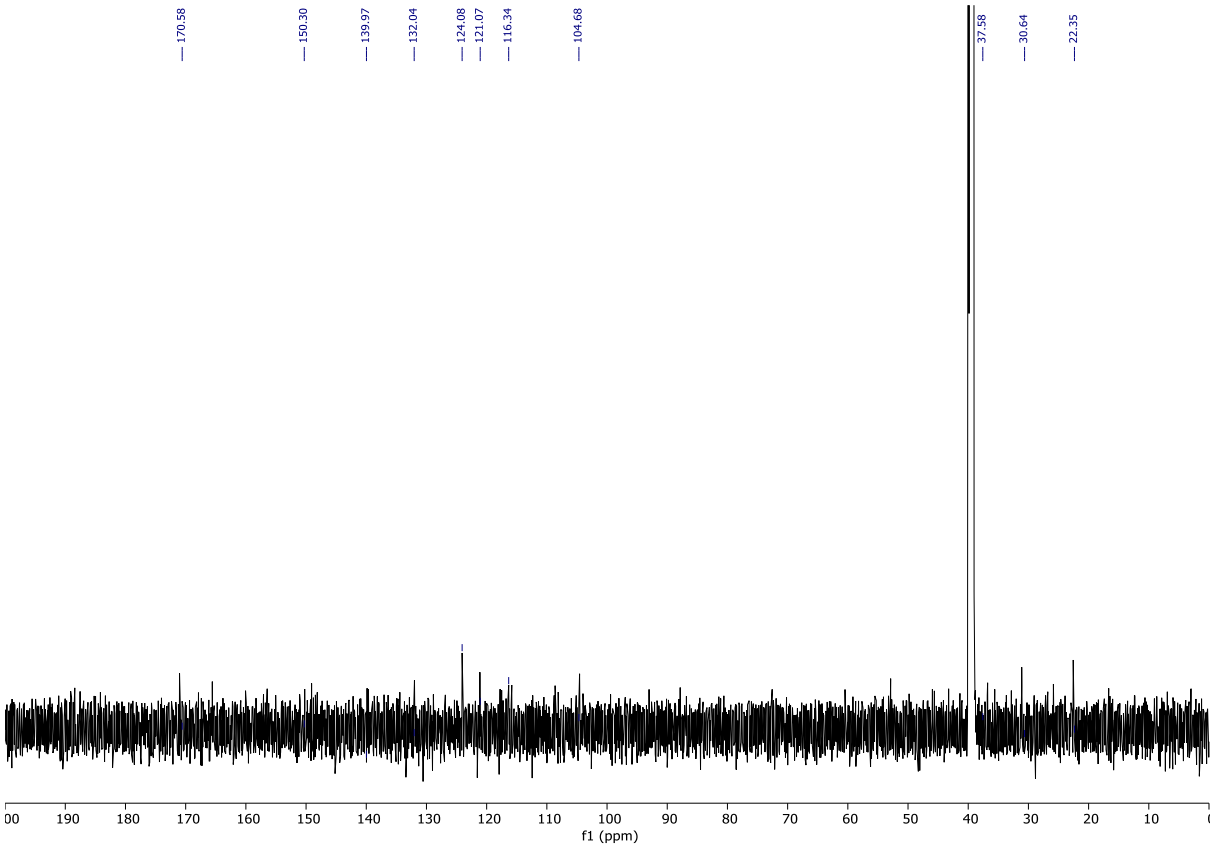
^{1}H NMR spectrum (600 MHz, DMSO-d_6) of **25**



^{13}C NMR spectrum (150 MHz, DMSO-d_6) of **25**

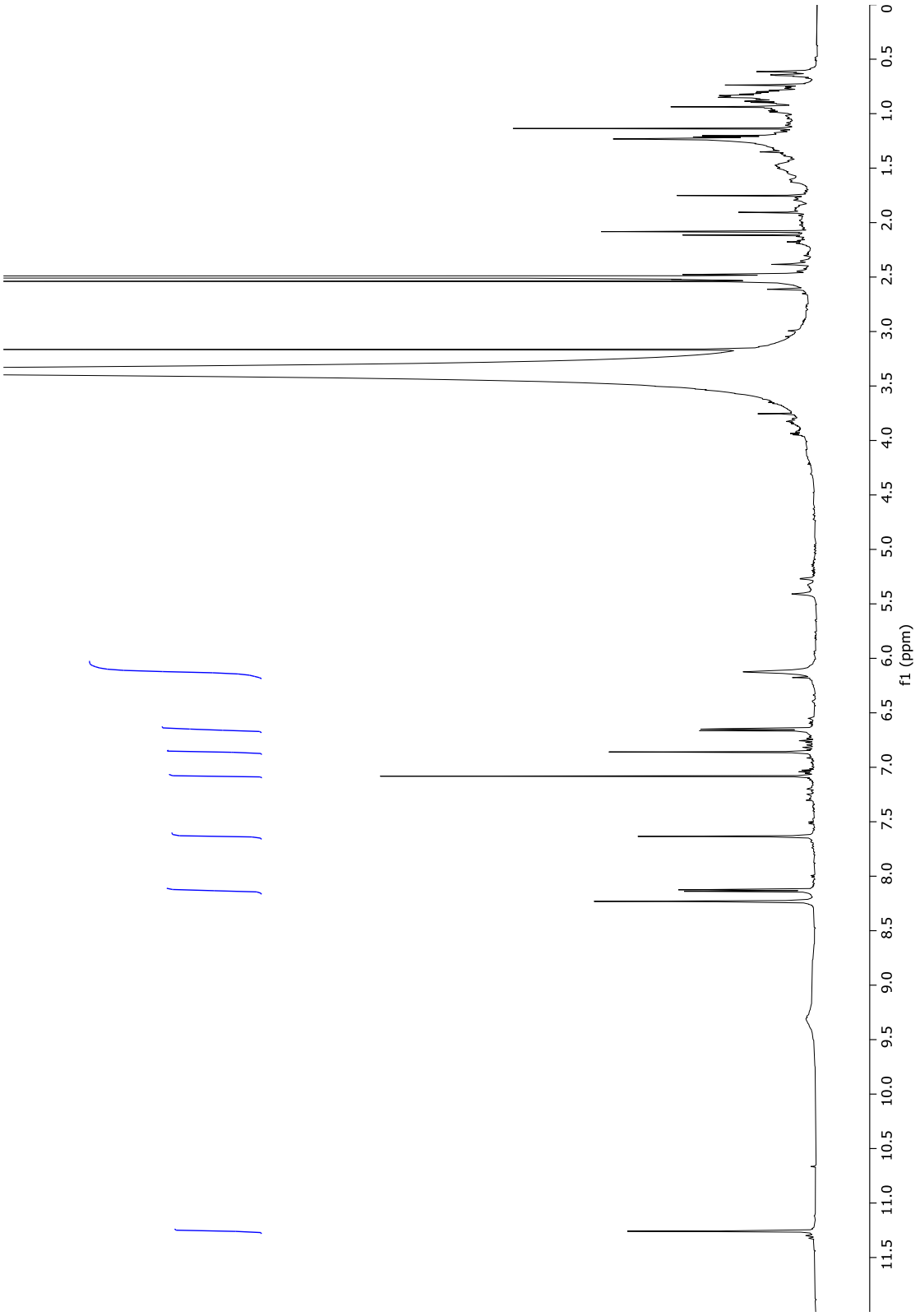


^{1}H NMR spectrum (600 MHz, DMSO-d_6) of **26**

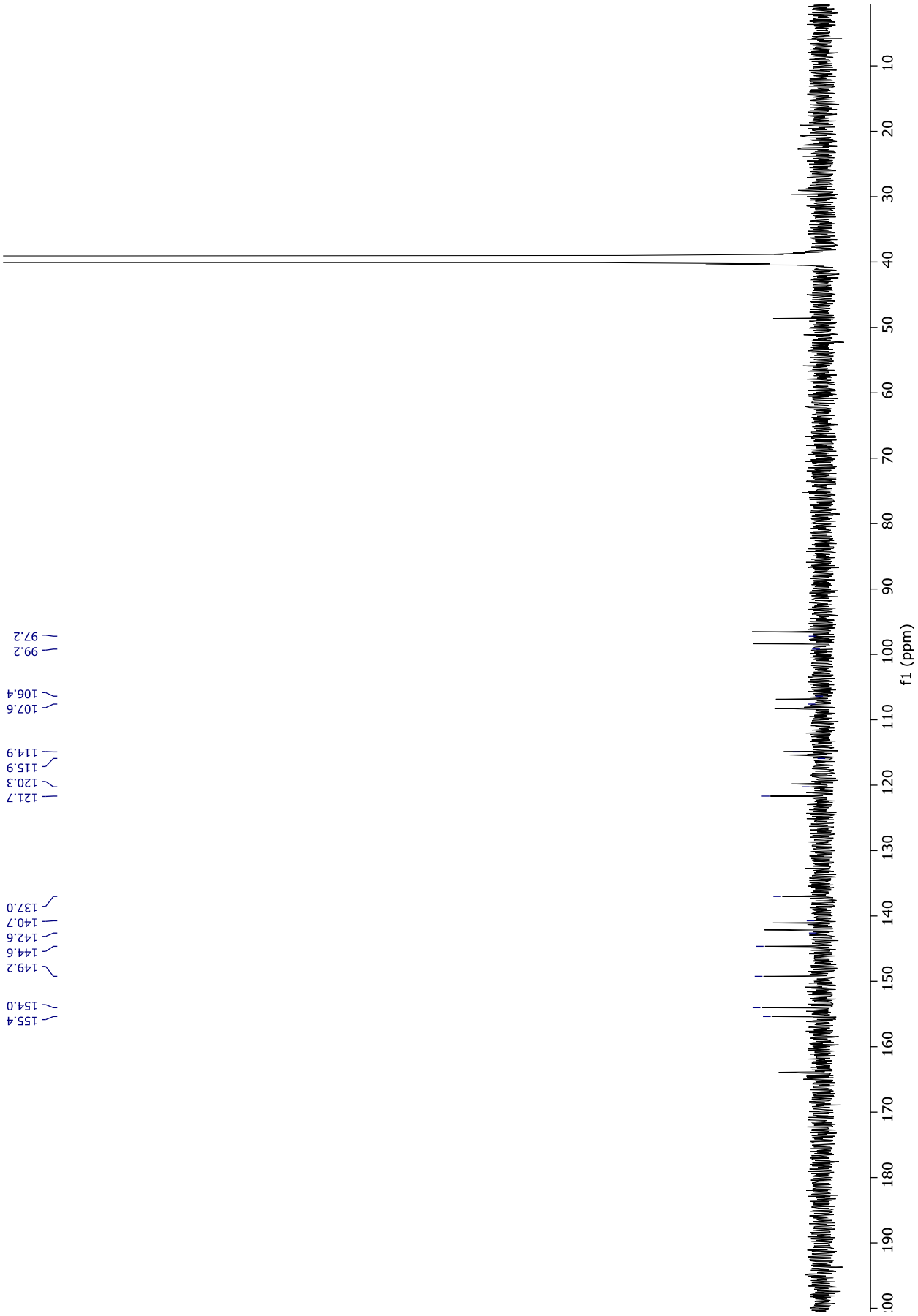


^{13}C NMR spectrum (150 MHz, DMSO-d_6) of **26**

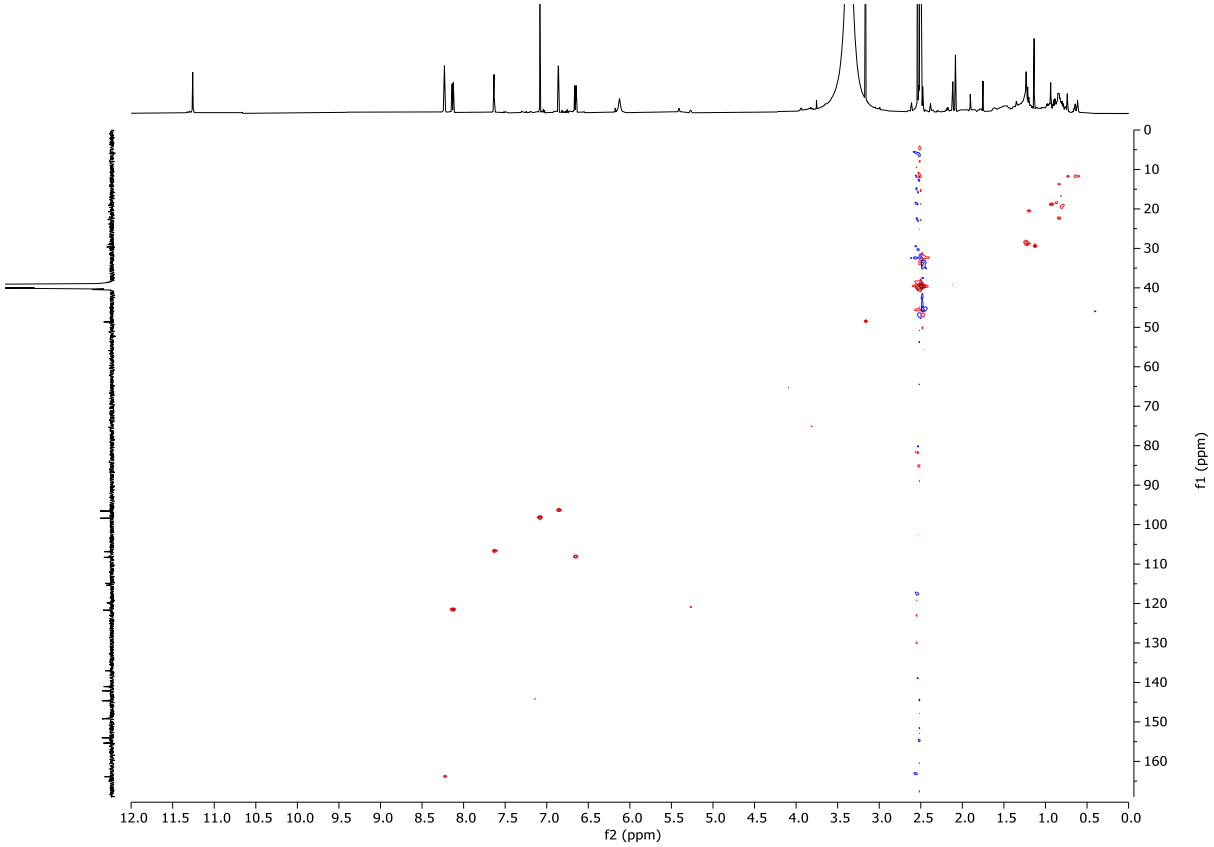
Appendix 2 – *Polyandrocarpa polypora* NMR Spectra



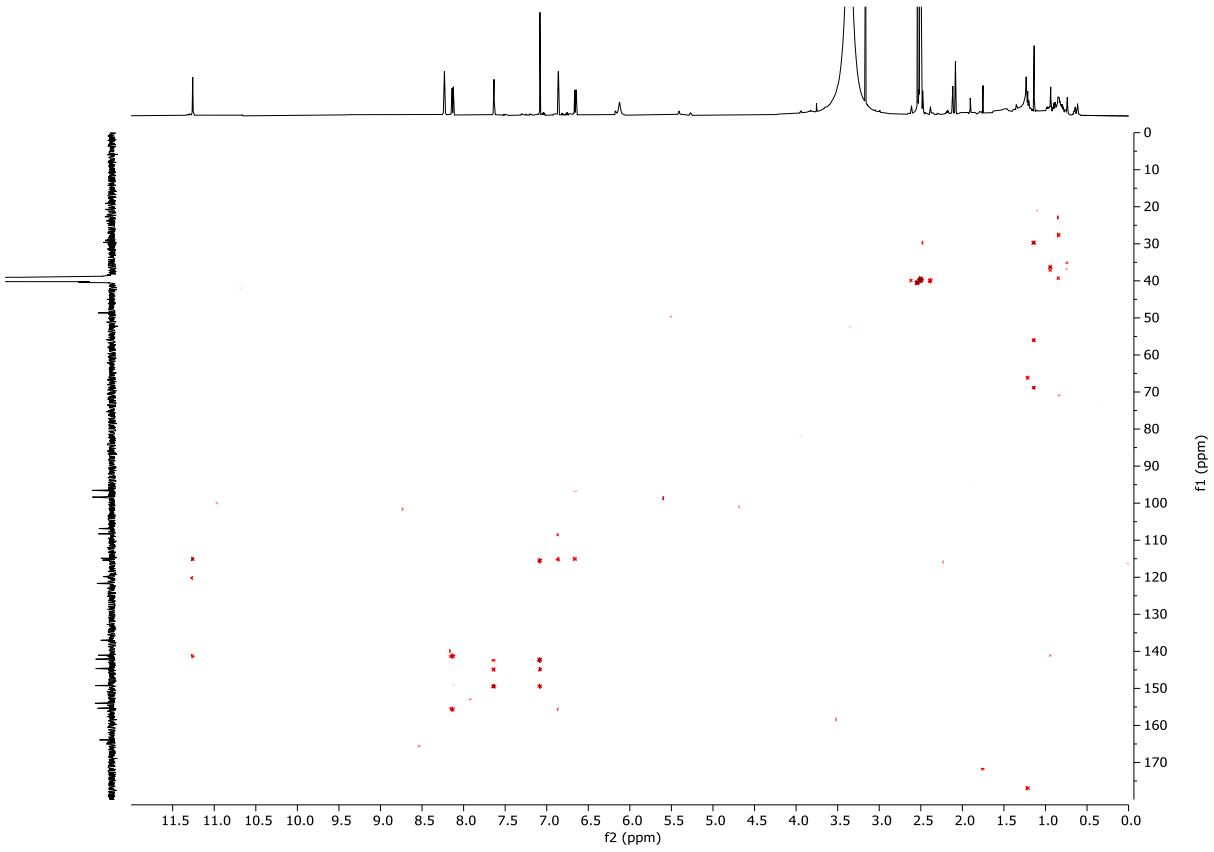
^{1}H NMR spectrum (600 MHz, DMSO-d_6) of **27**



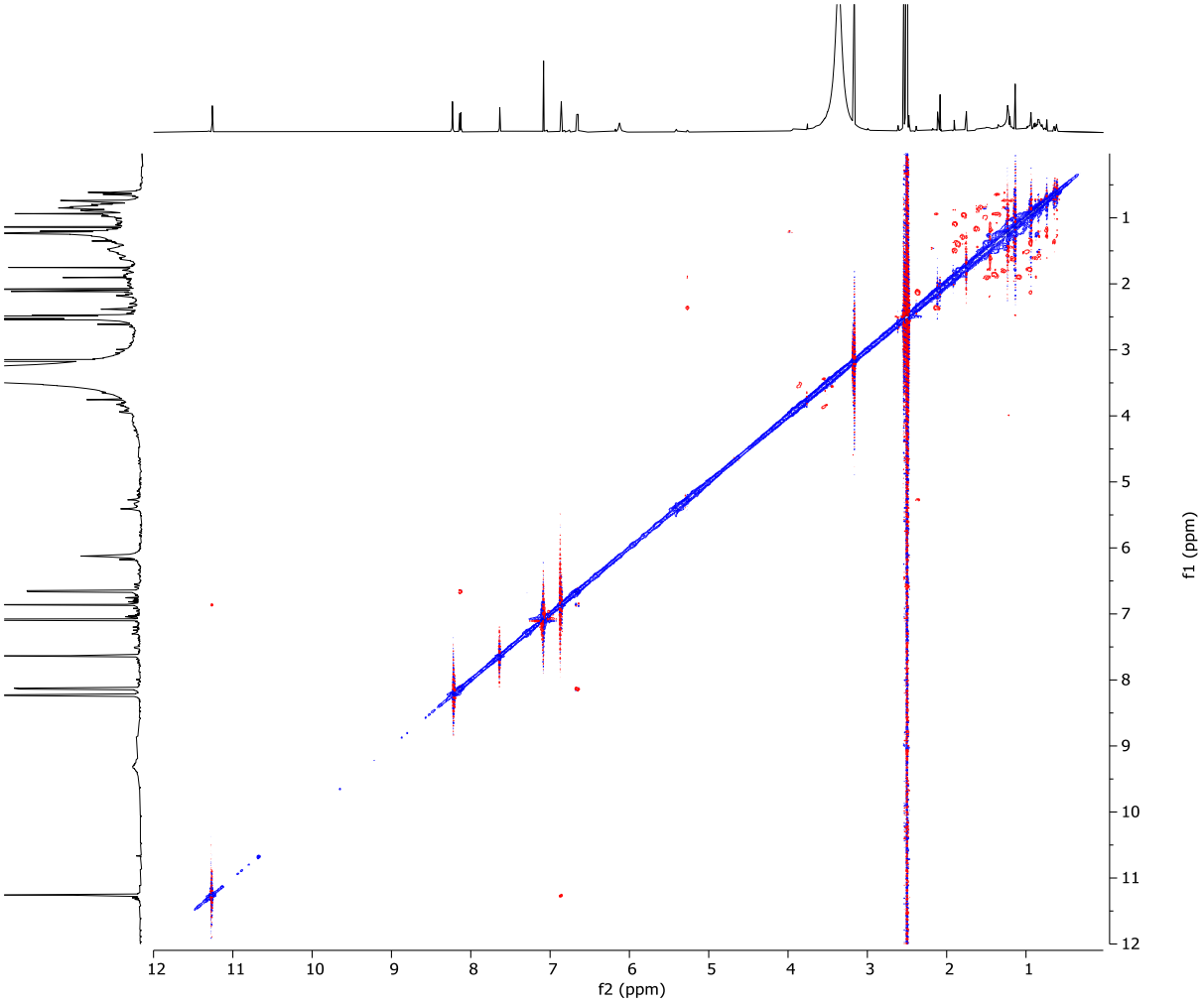
^{13}C NMR spectrum (150 MHz, DMSO-d_6) of **27**



HSQC NMR spectrum (600 MHz, DMSO-d_6) of **27**

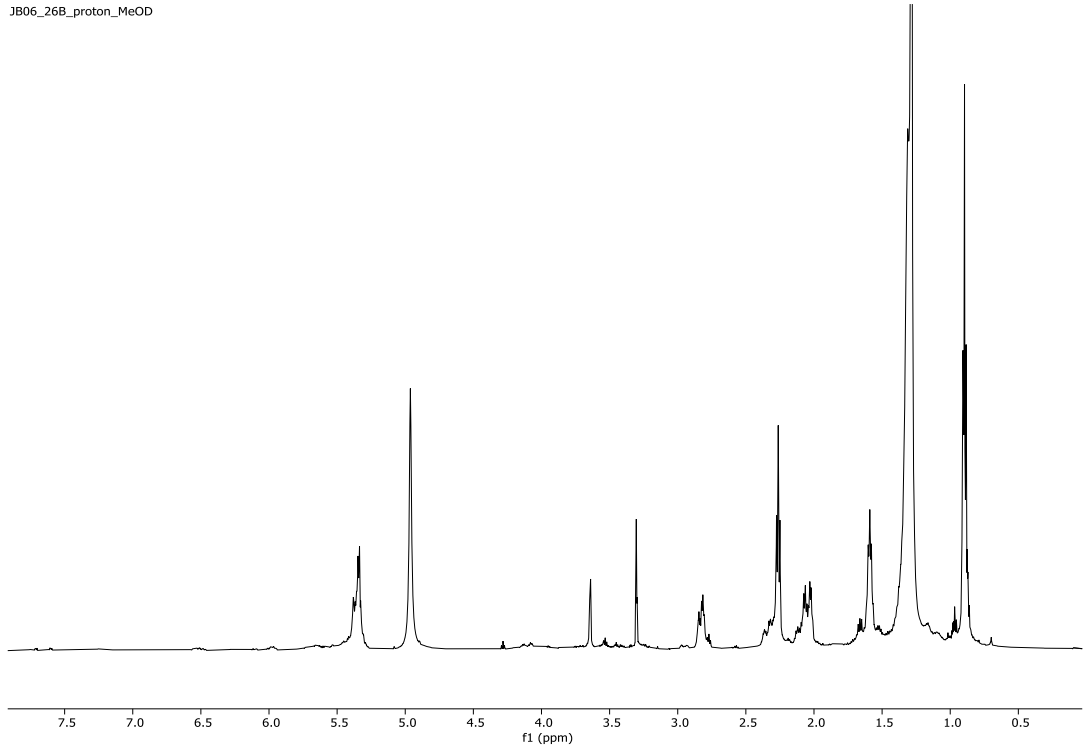


HMBC NMR spectrum (600 MHz, DMSO-d_6) of **27**

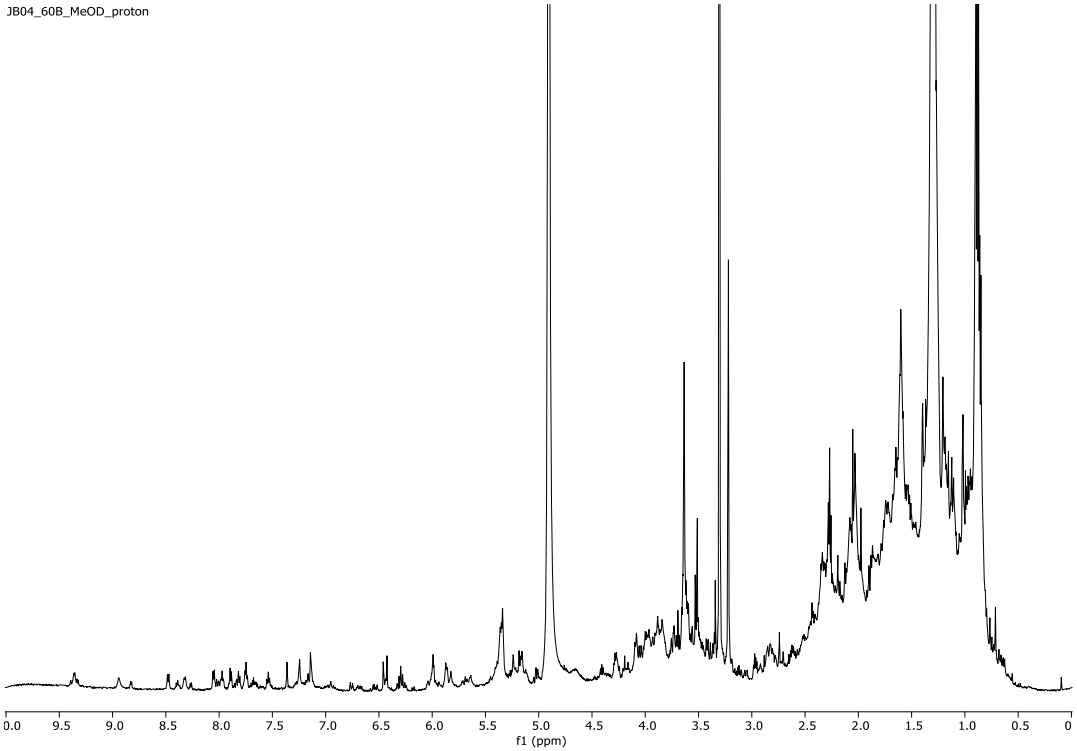


ROESY NMR spectrum (600 MHz, DMSO-d_6) of **27**

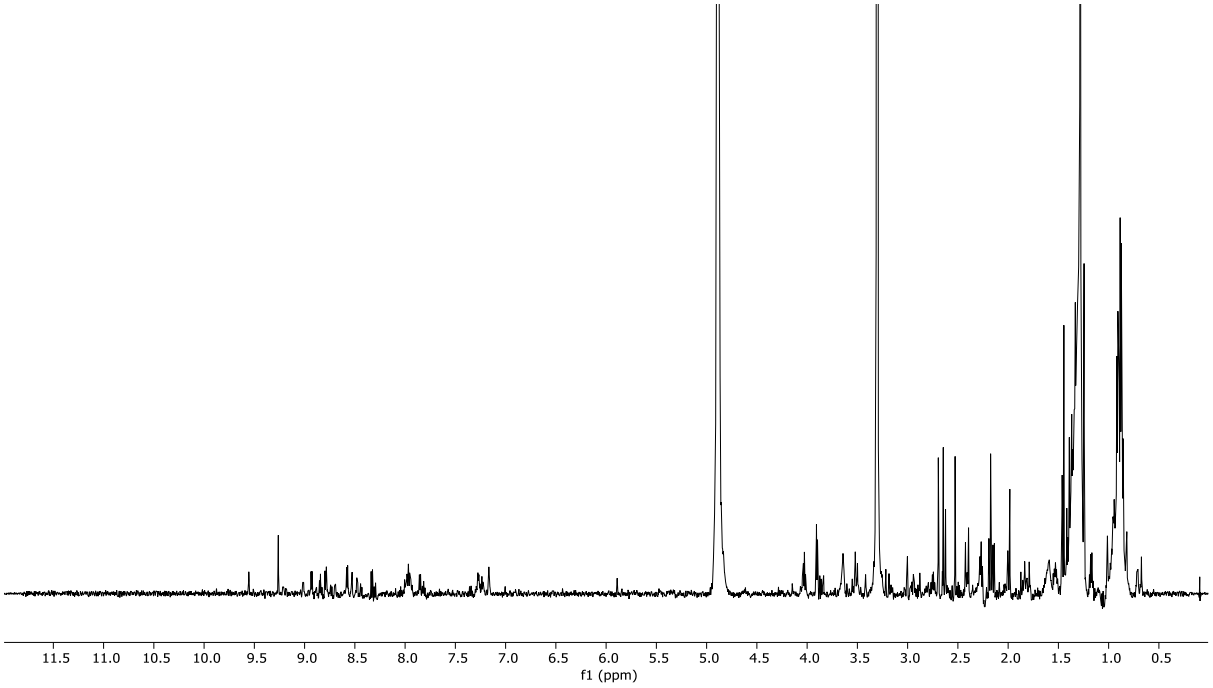
Appendix 3 – Organisms Not Further Investigated



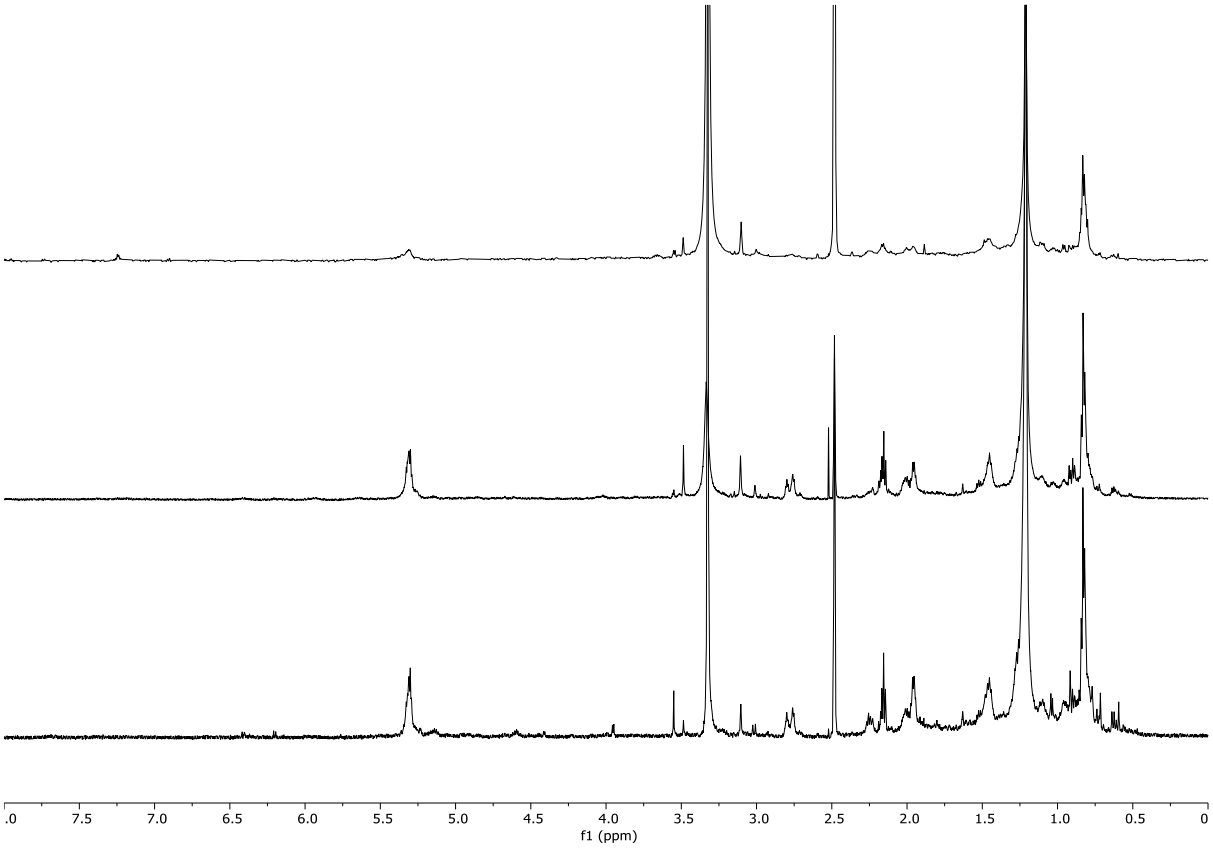
^{1}H NMR spectrum of the 75% Me_2CO/H_2O screening fraction of PTN4_01A (632) (600 MHz, CD_3OD).



^{1}H NMR spectrum of the 75% Me_2CO/H_2O screening fraction of PTN4_40E (855) (600 MHz, CD_3OD).

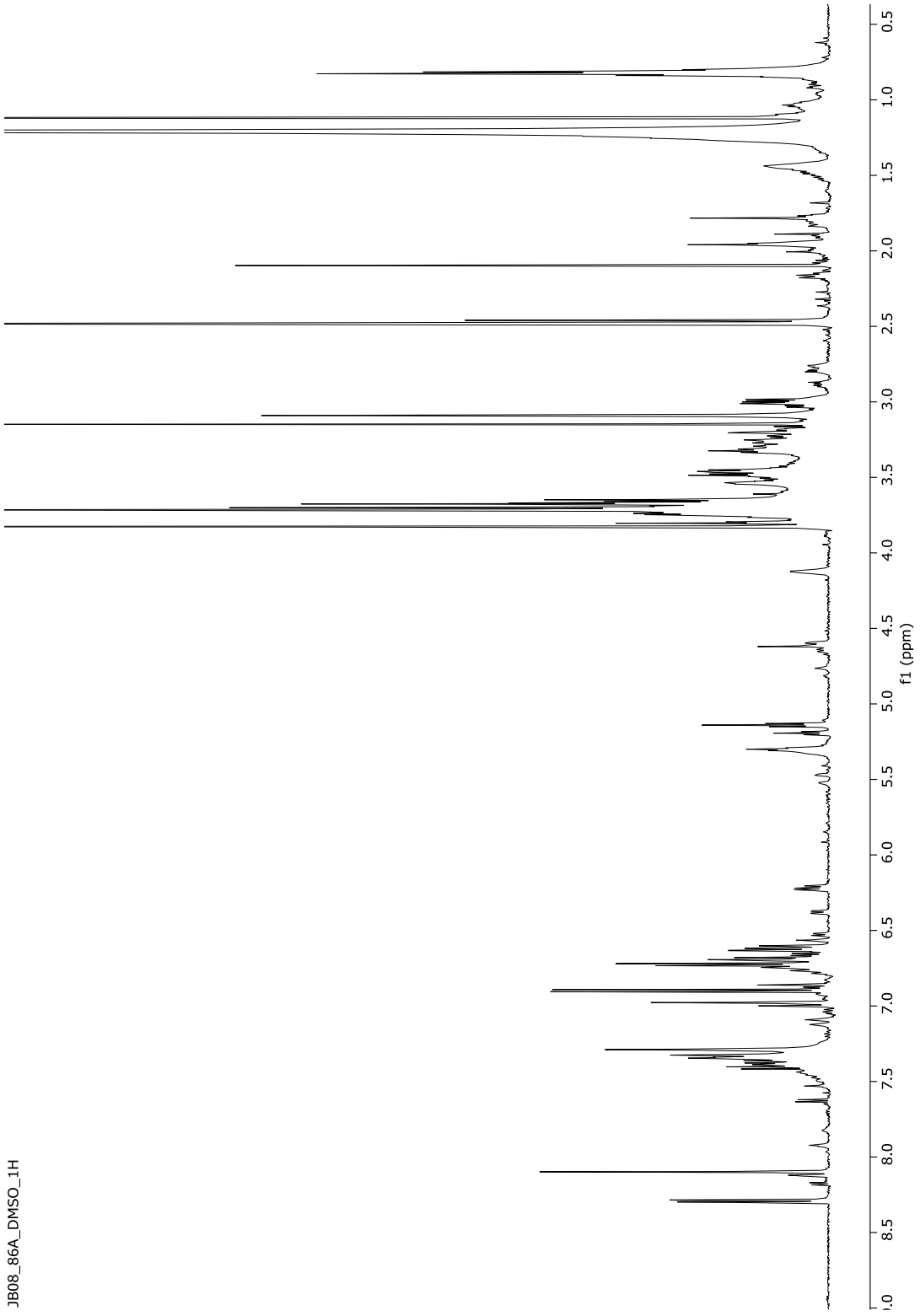


^{1}H NMR spectrum of the 30% Me_2CO/H_2O screening fraction of PTN3_03C (650) (600 MHz, CD_3OD).

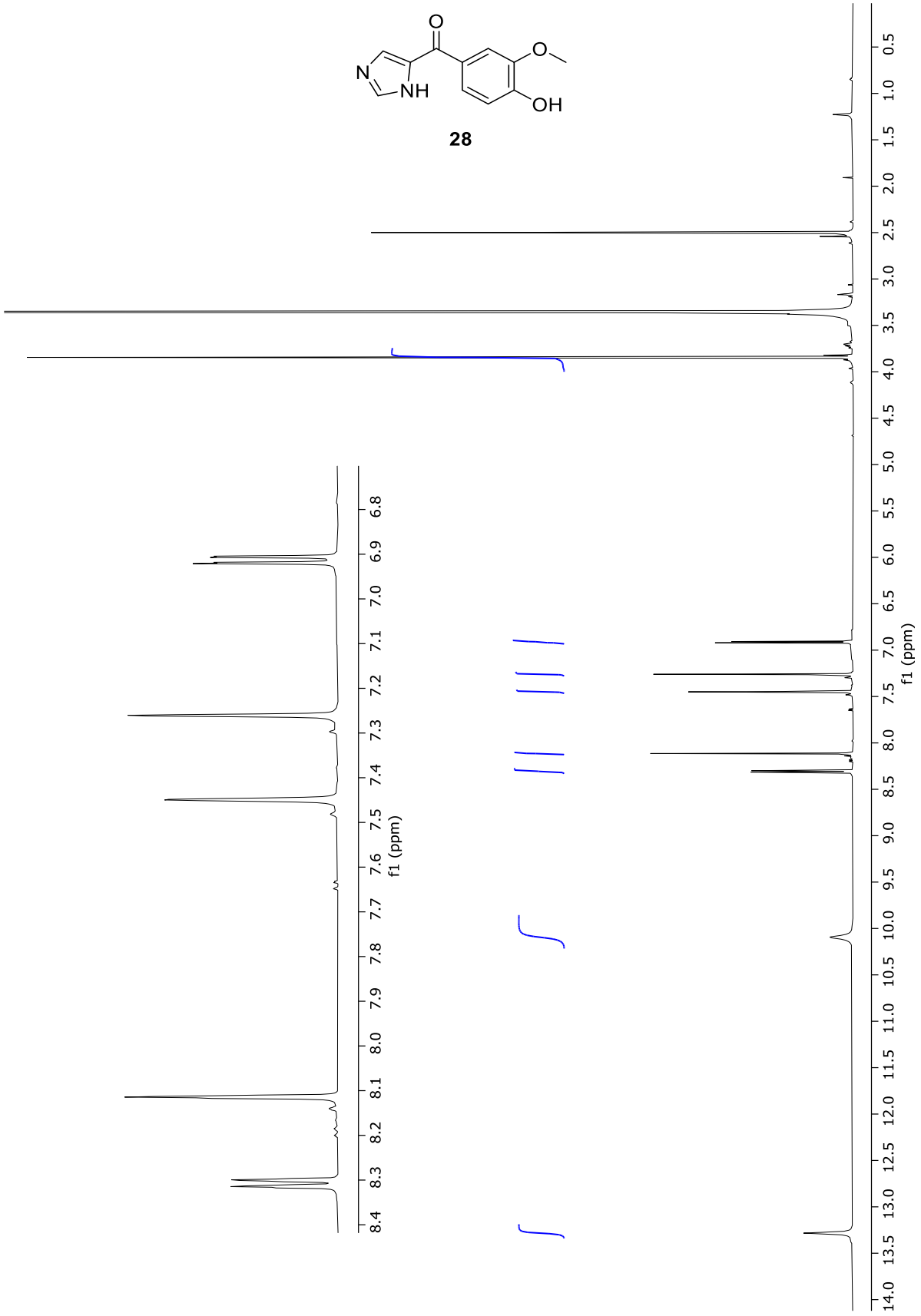


^{1}H NMR spectra of 30%, 75% and 100% (top to bottom respectively) Me_2CO/H_2O screening fractions of *C. aucklandicus* (600 MHz, CD_3OD).

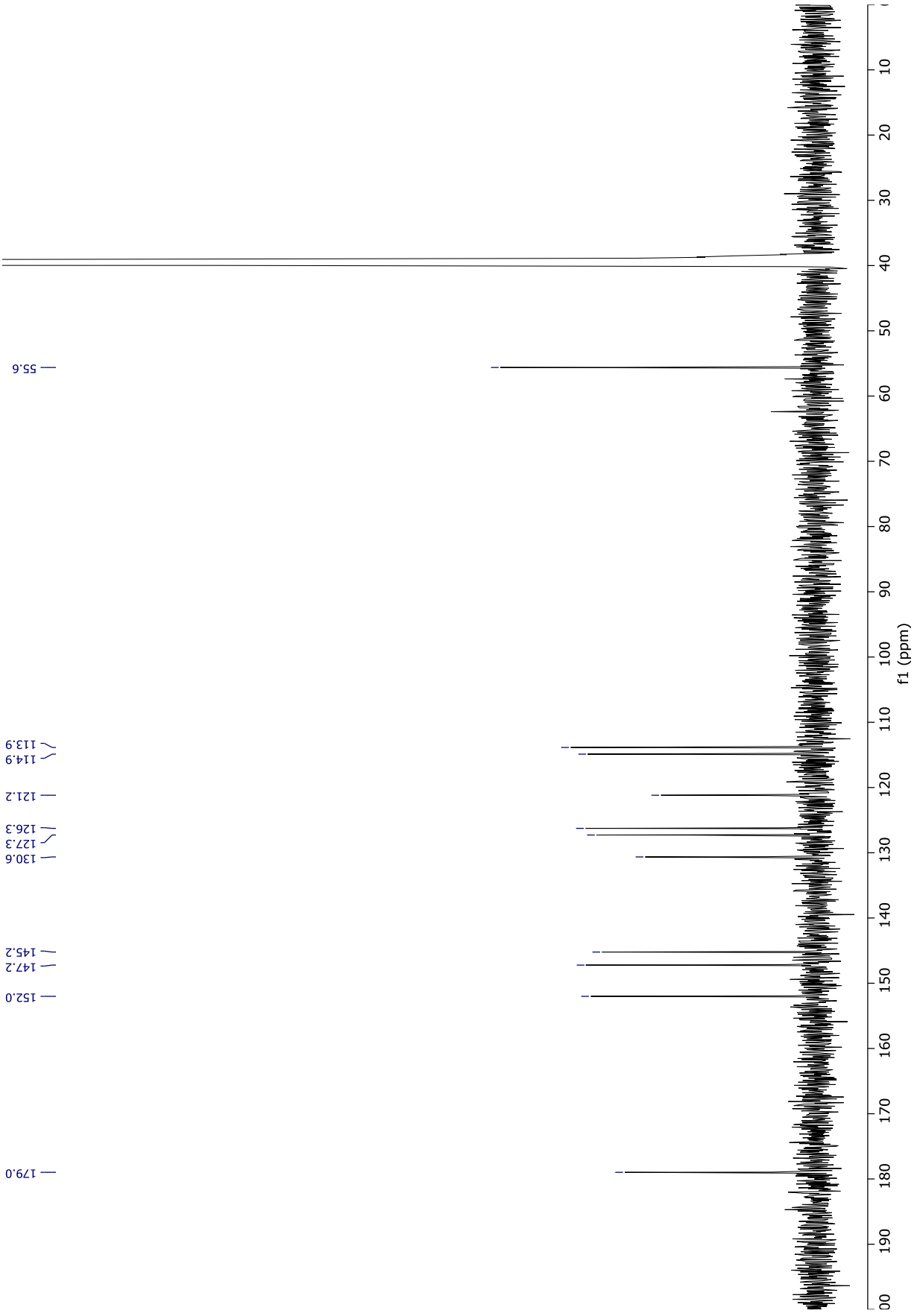
Appendix 4 – *Distaplia stylifera* Spectra



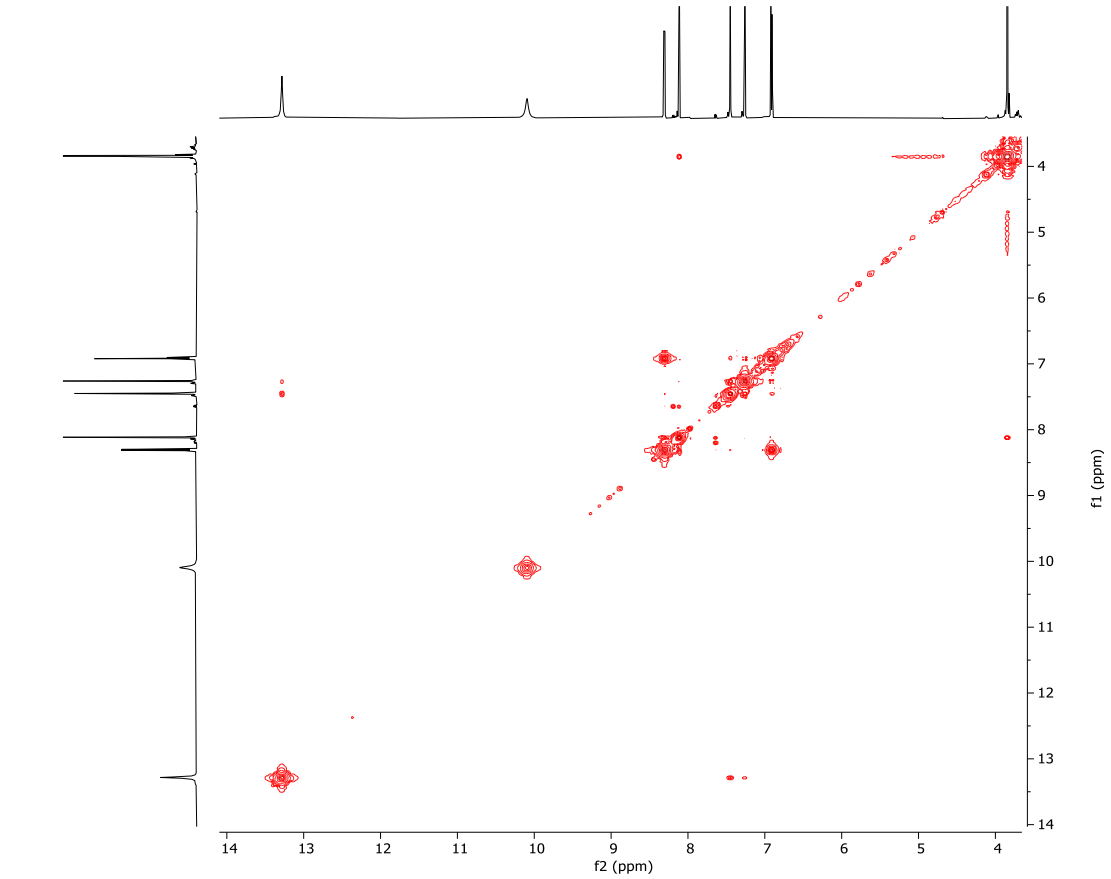
^{1}H NMR spectrum of the 30% Me_2CO/H_2O screening fraction of *D. stylifera* (600 MHz, DMSO-d_6).



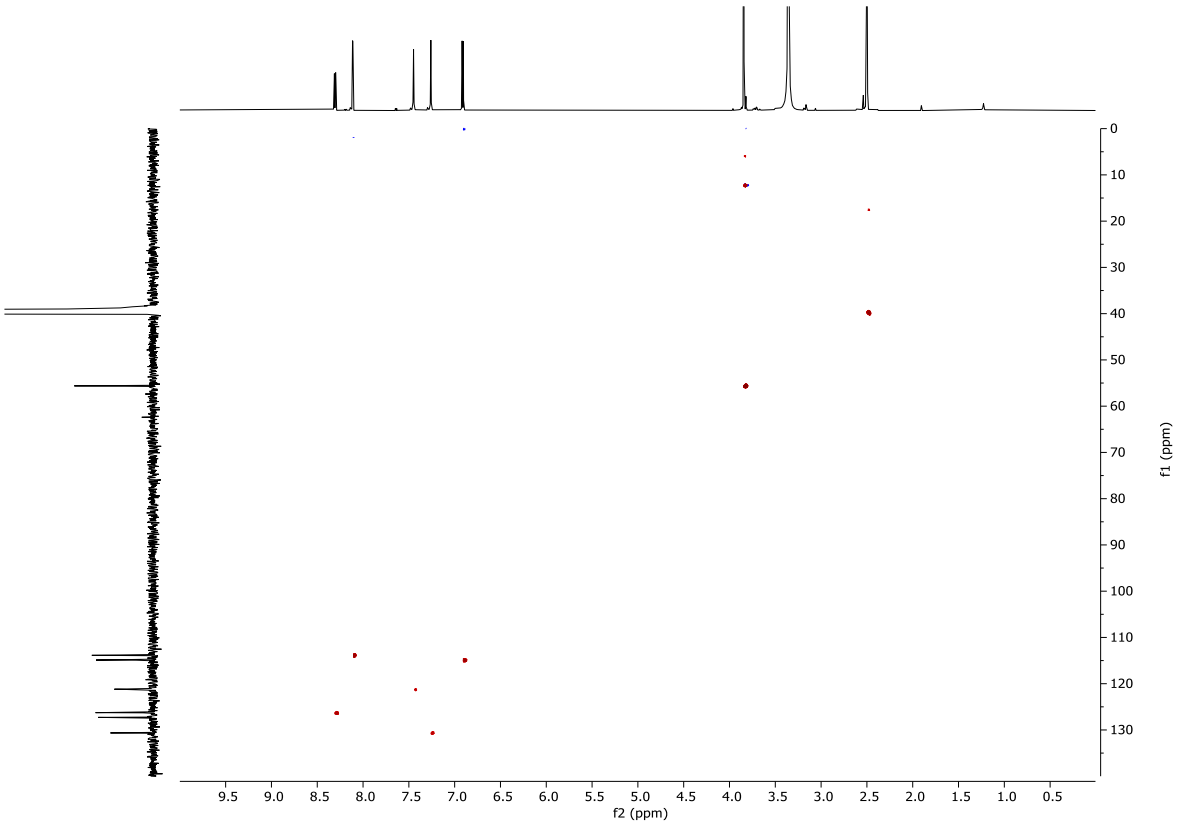
^{1}H NMR spectrum (600 MHz, DMSO-d_6) of **28**



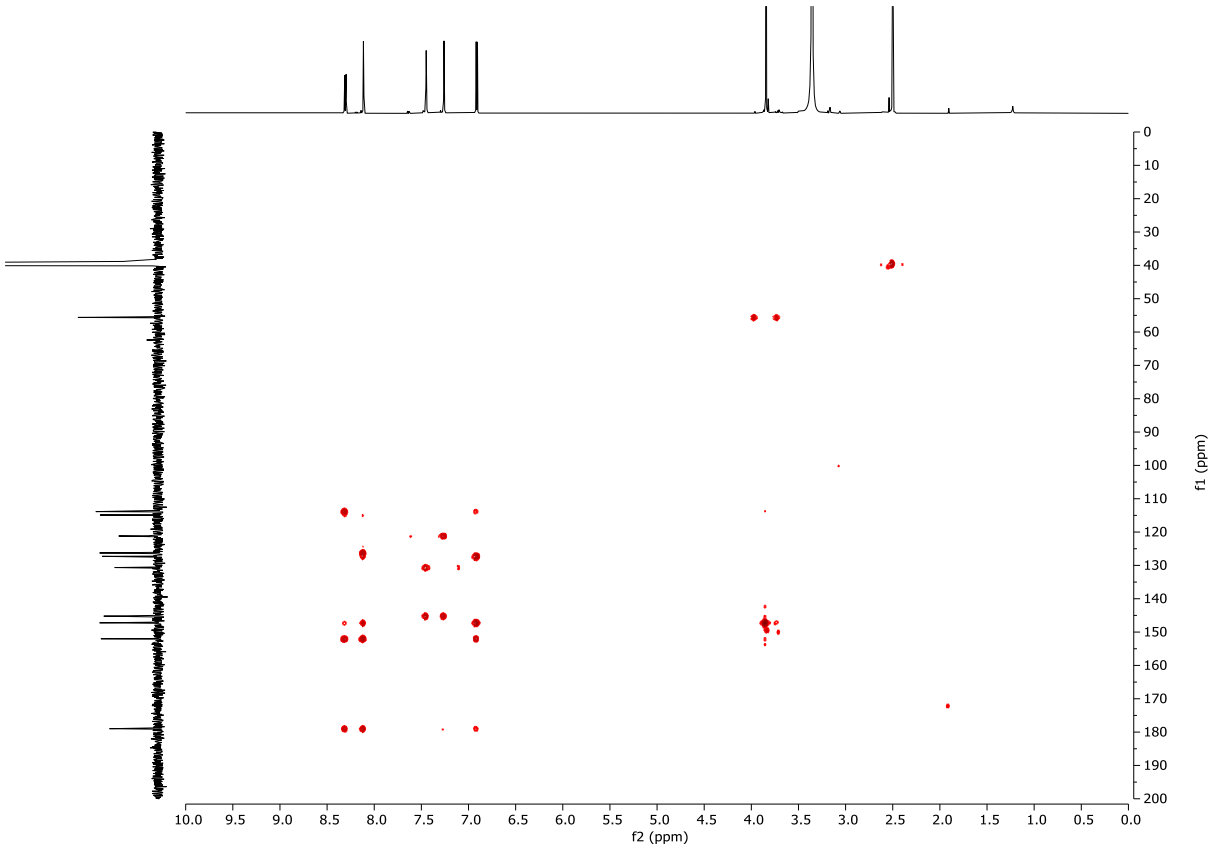
^{13}C NMR spectrum (150 MHz, DMSO-d_6) of **28**



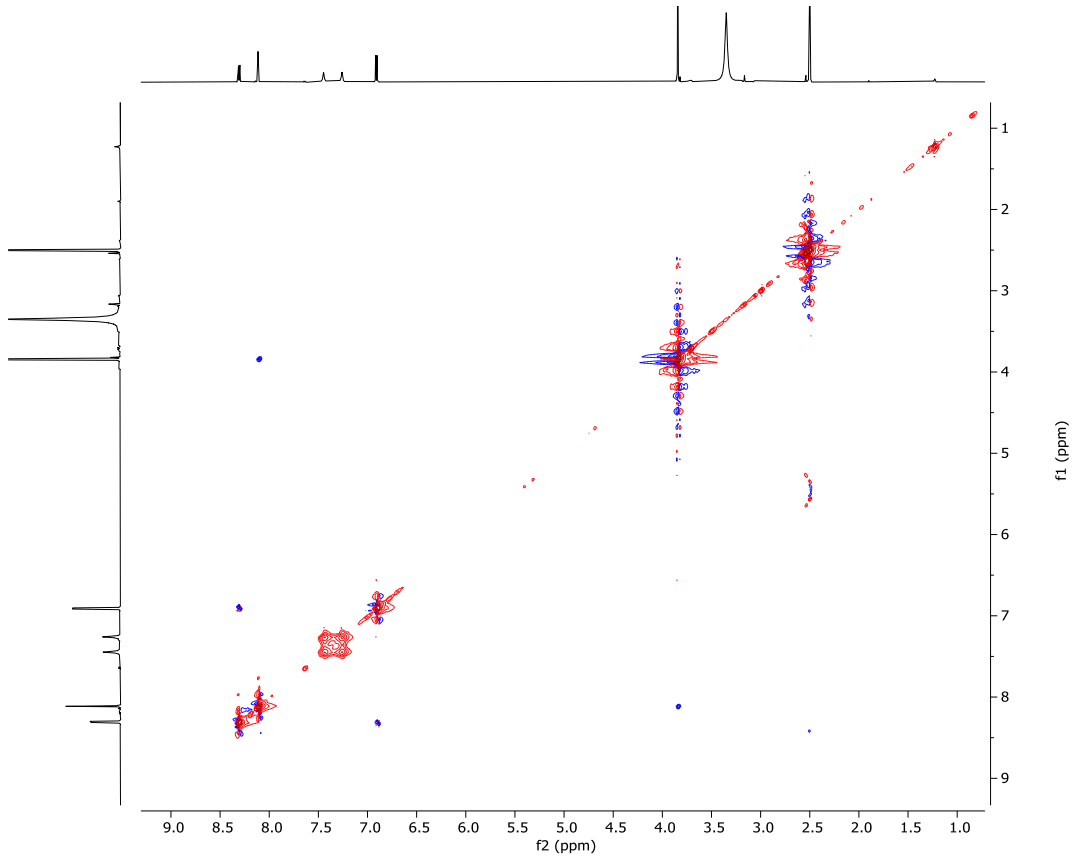
COSY NMR spectrum (600 MHz, DMSO-d_6) of **28**



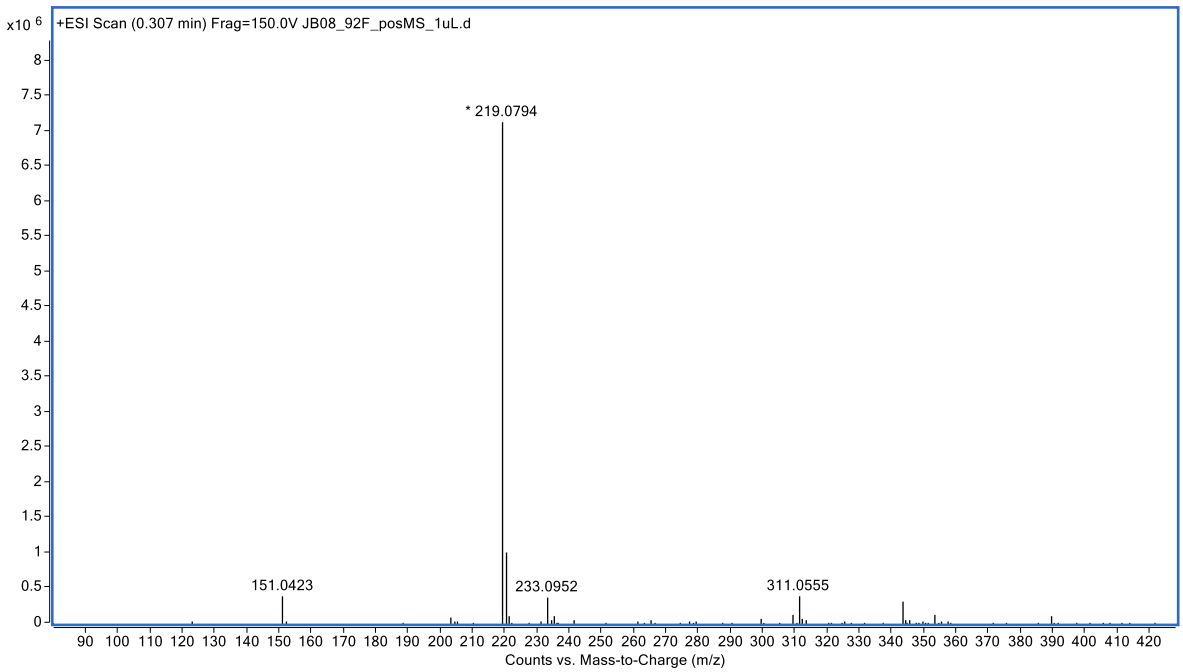
HSQC NMR spectrum (600 MHz, DMSO-d_6) of **28**



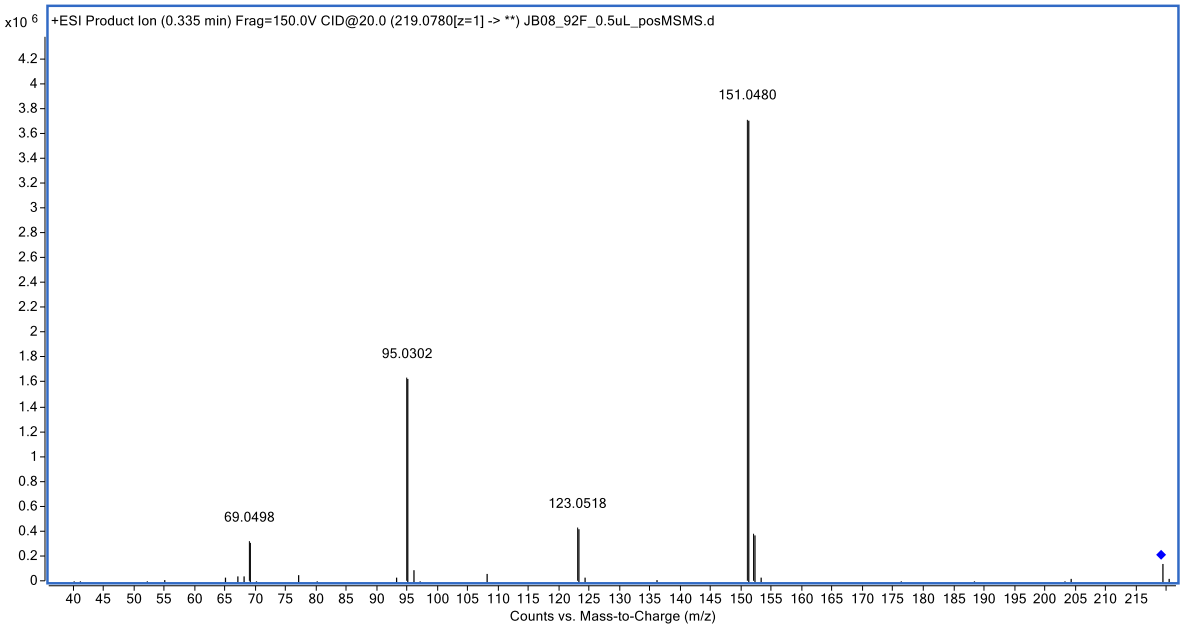
HMBC NMR spectrum (600 MHz, DMSO-d_6) of **28**



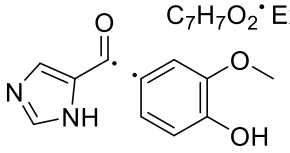
ROESY NMR spectrum (600 MHz, DMSO-d_6) of **28**



(+)-HRESIMS Spectrum of **28**

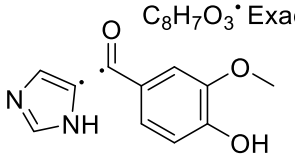


(+)-HRESIMS/MS spectrum of **28** using CID of 20 eV



$C_7H_7O_2^{\bullet}$ Exact Mass: 123.0446

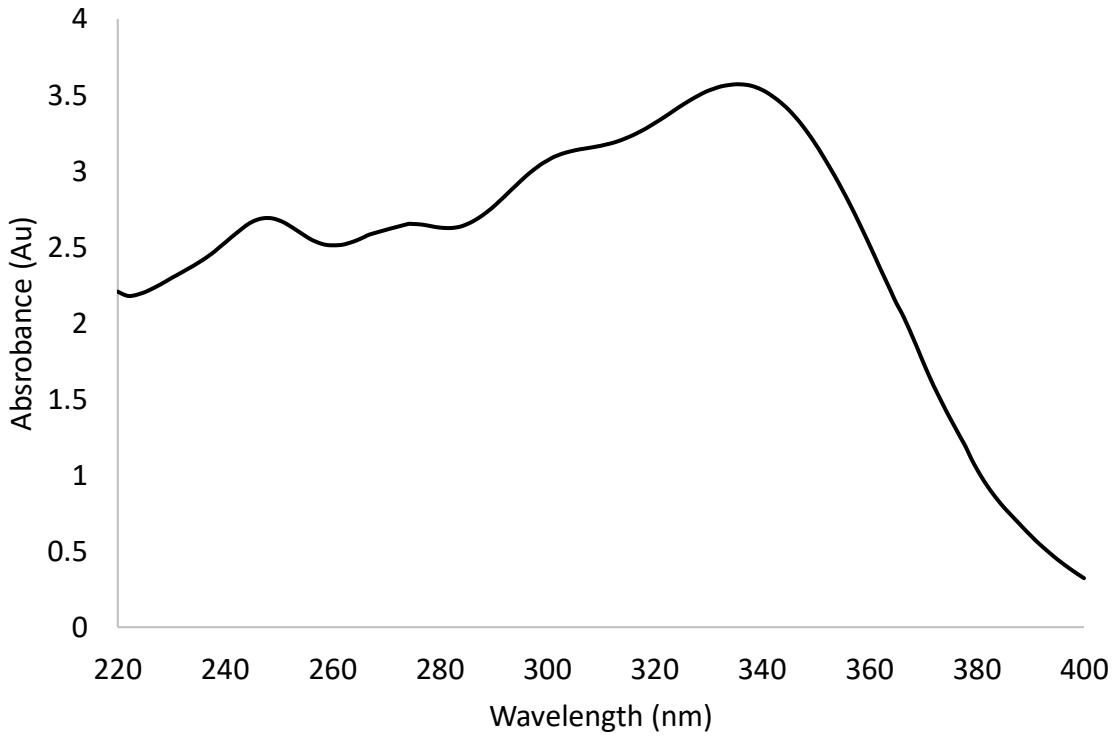
$C_4H_3N_2O^{\bullet}$ Exact Mass: 95.0245



$C_8H_7O_3^{\bullet}$ Exact Mass: 151.0395

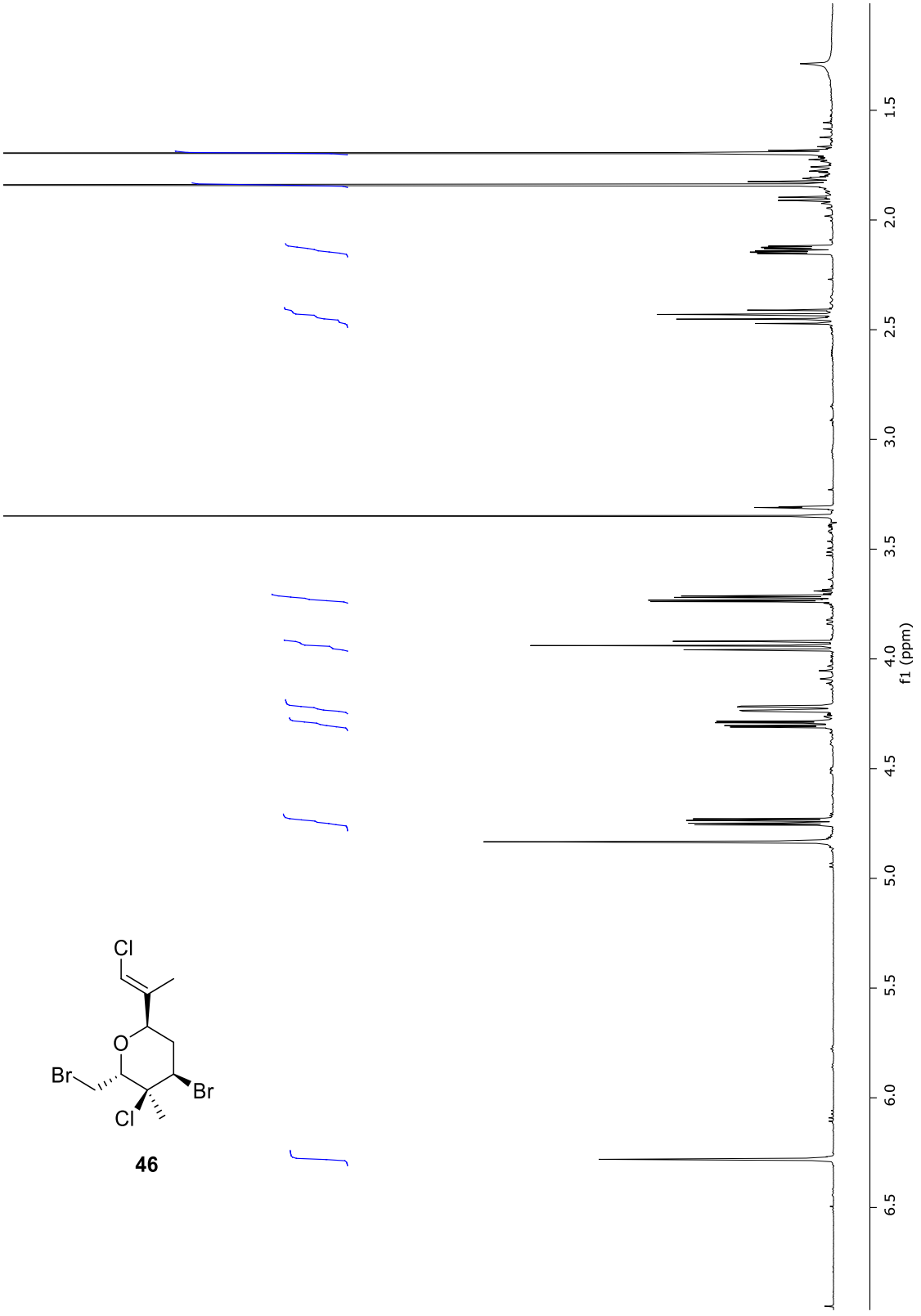
$C_3H_3N_2^{\bullet}$ Exact Mass: 67.0296

28 mass fragments shown in MS/MS

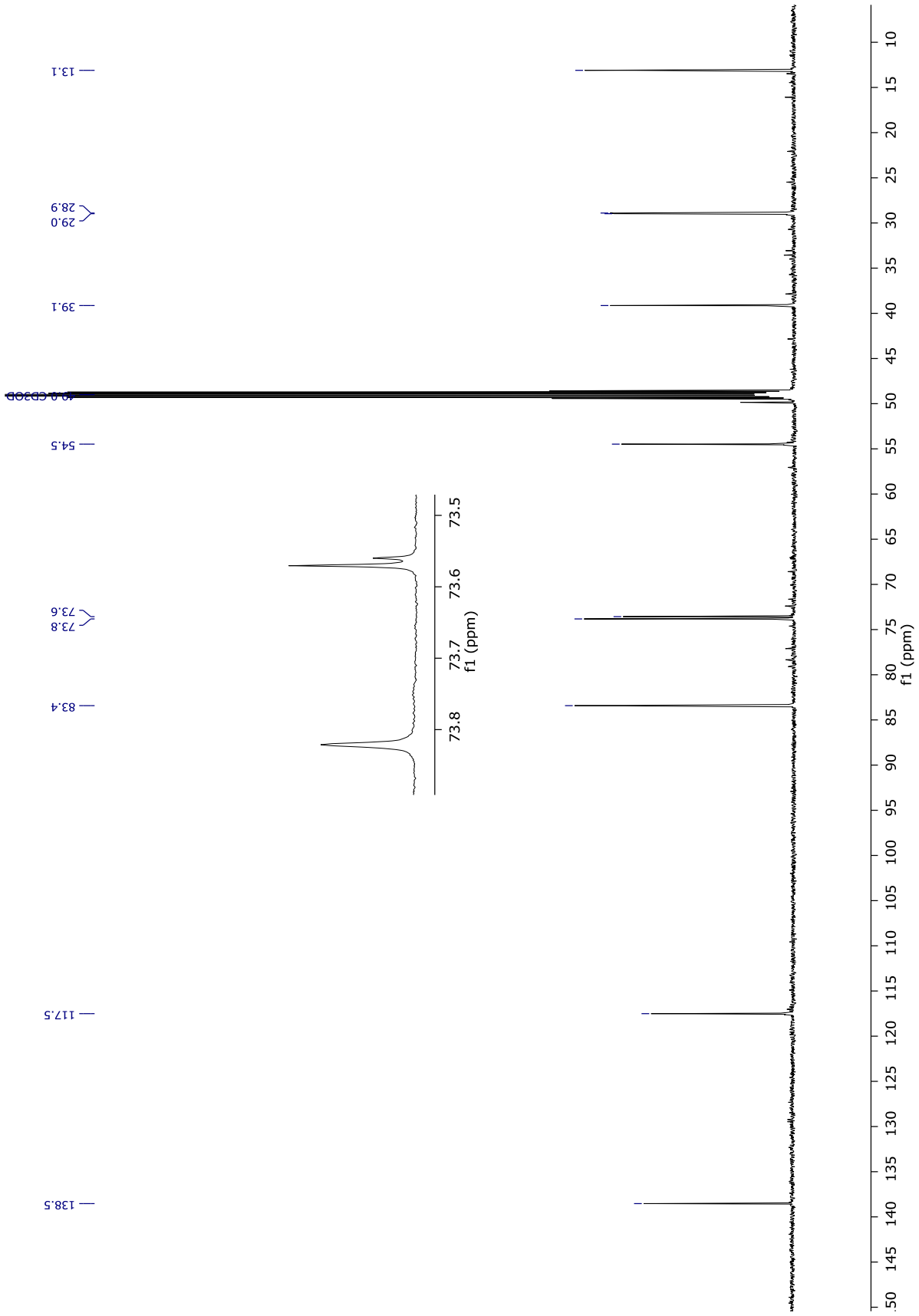


UV/vis Spectrum of **28**

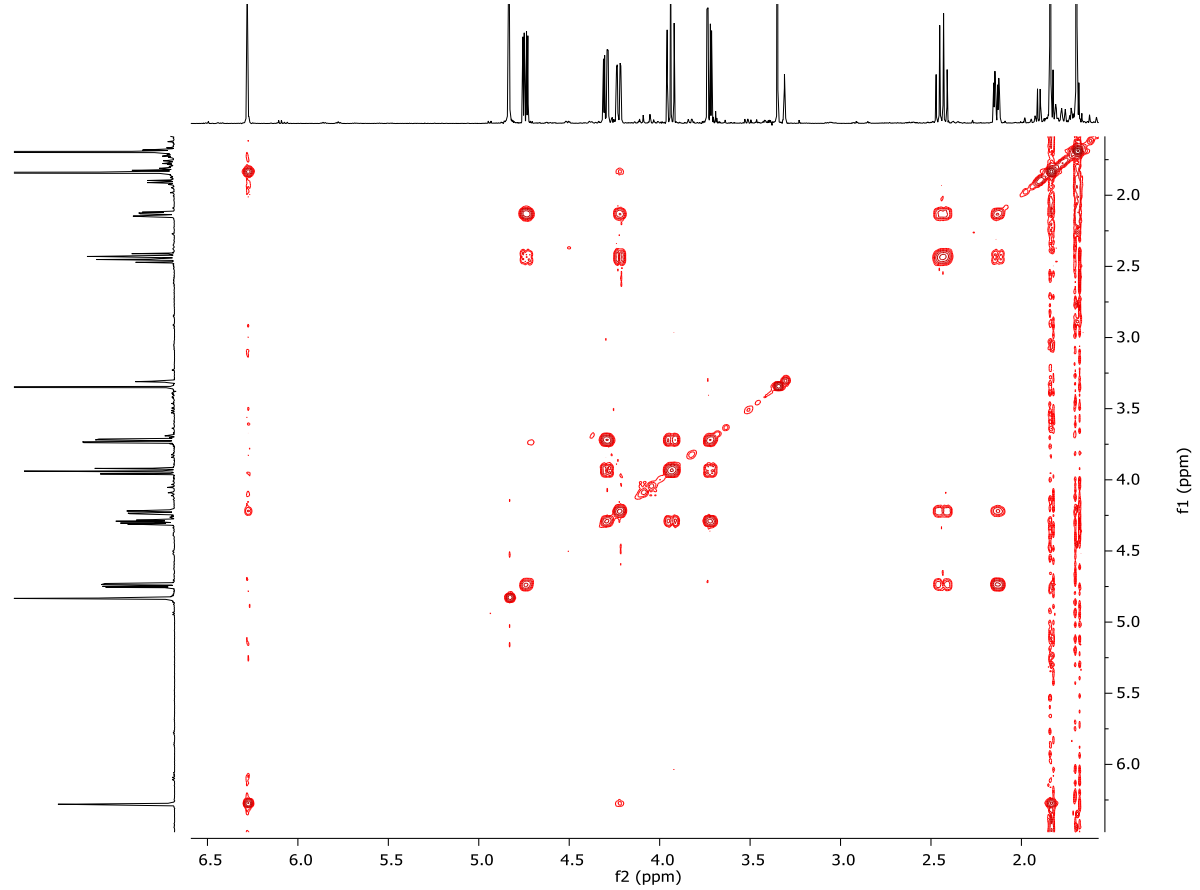
Appendix 5 – '*Plocamium angustum*' Spectra



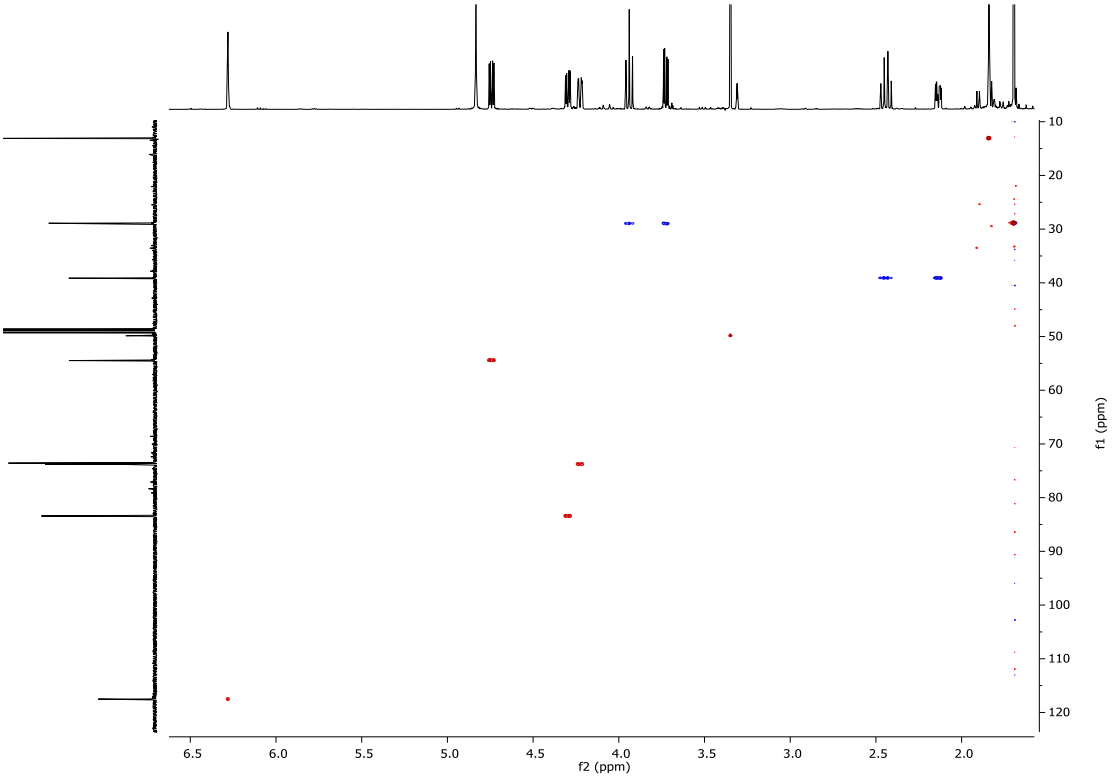
^{1}H Spectrum (600 MHz, CD_3OD) of **46**



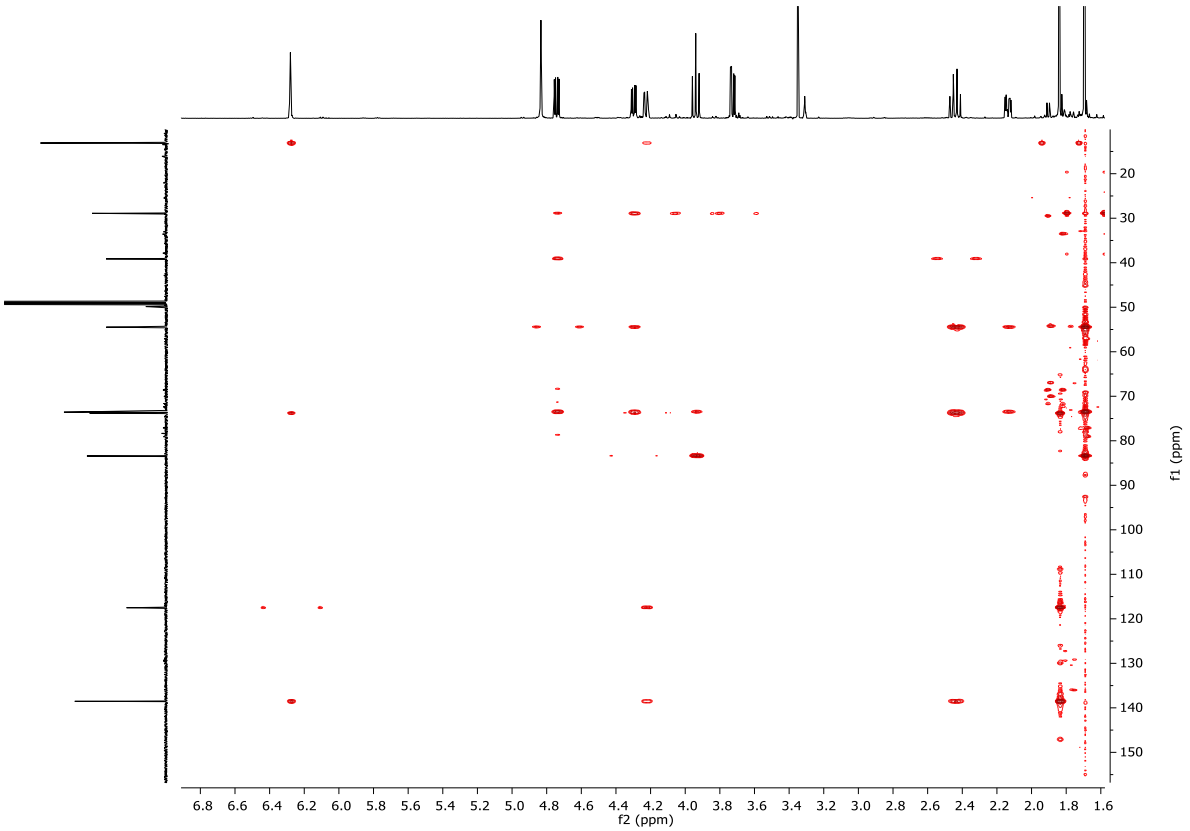
^{13}C Spectrum (150 MHz, CD_3OD) of **46**



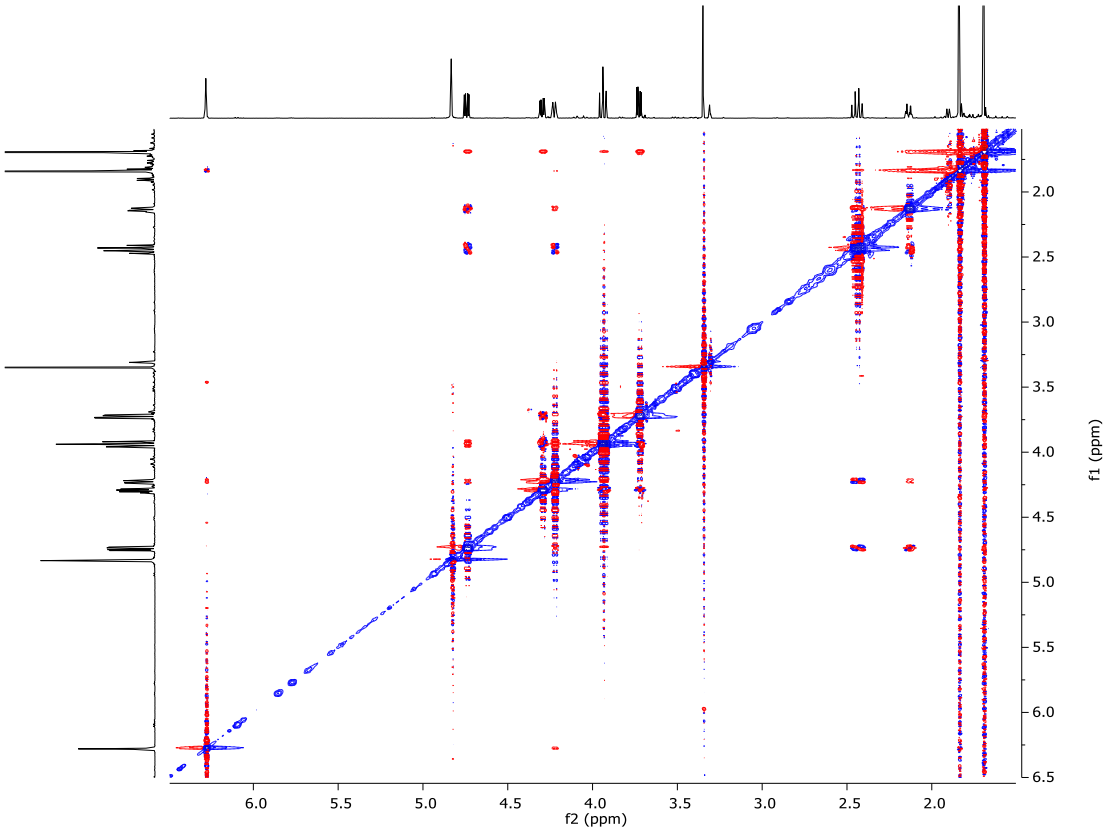
COSY Spectrum (600 MHz, CD_3OD) of **46**



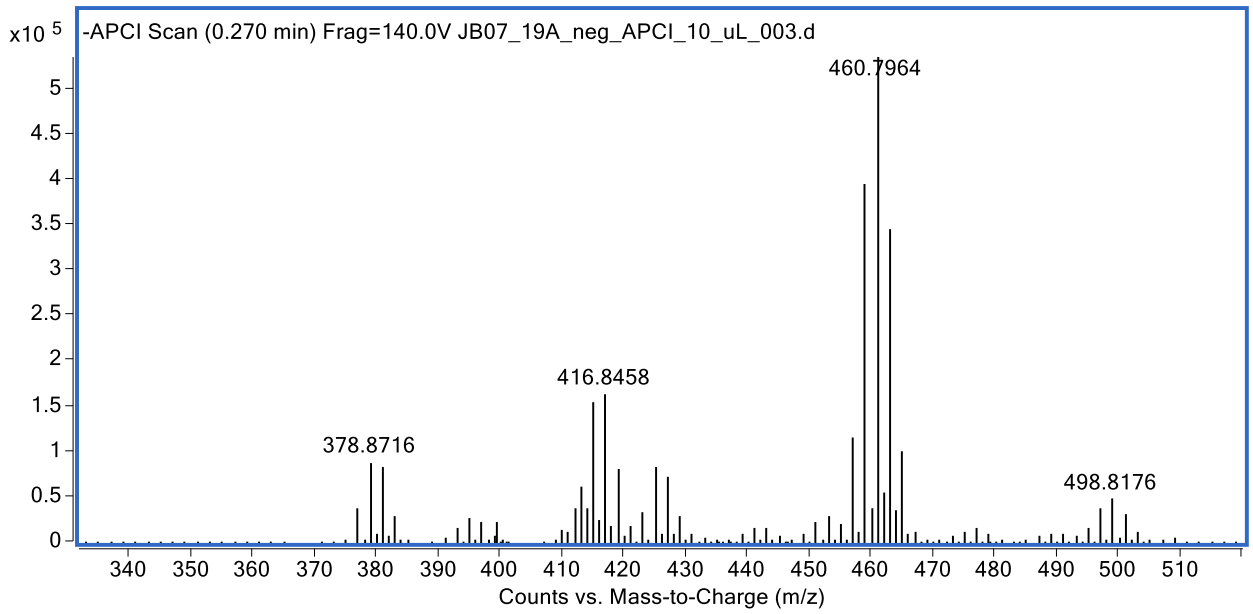
HSQC Spectrum (600 MHz, CD_3OD) of **46**



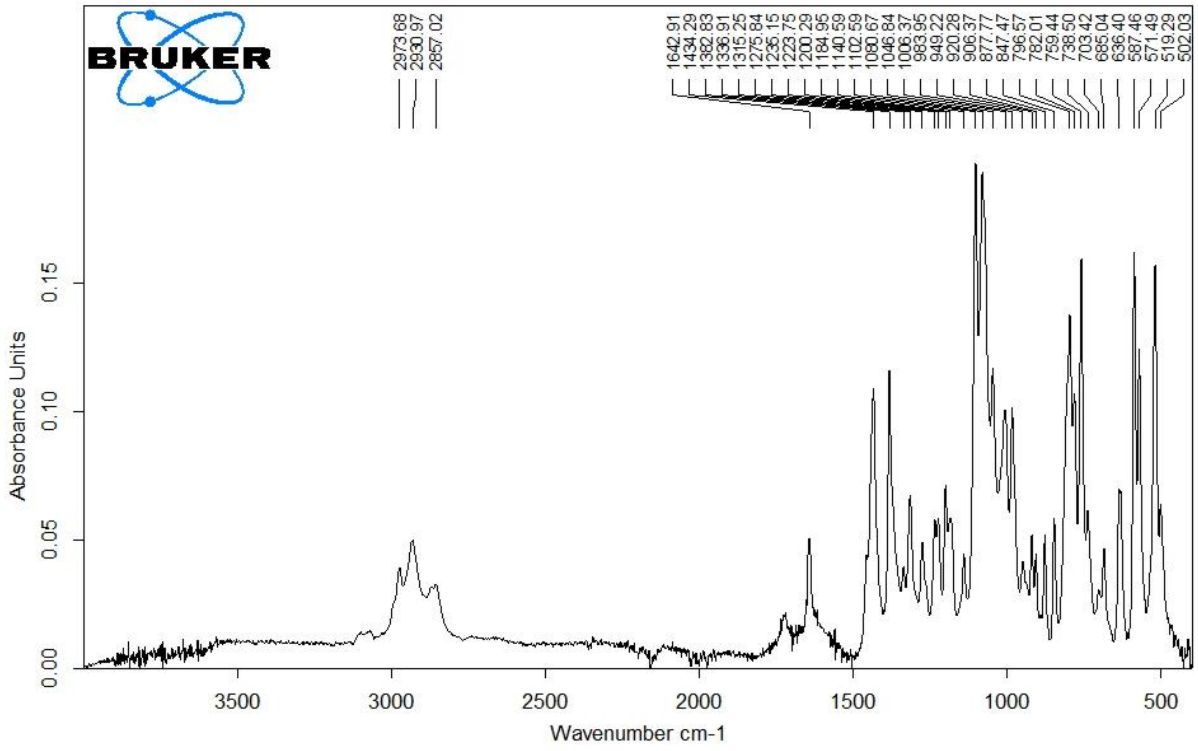
HMBC Spectrum (600 MHz, CD_3OD) of **46**



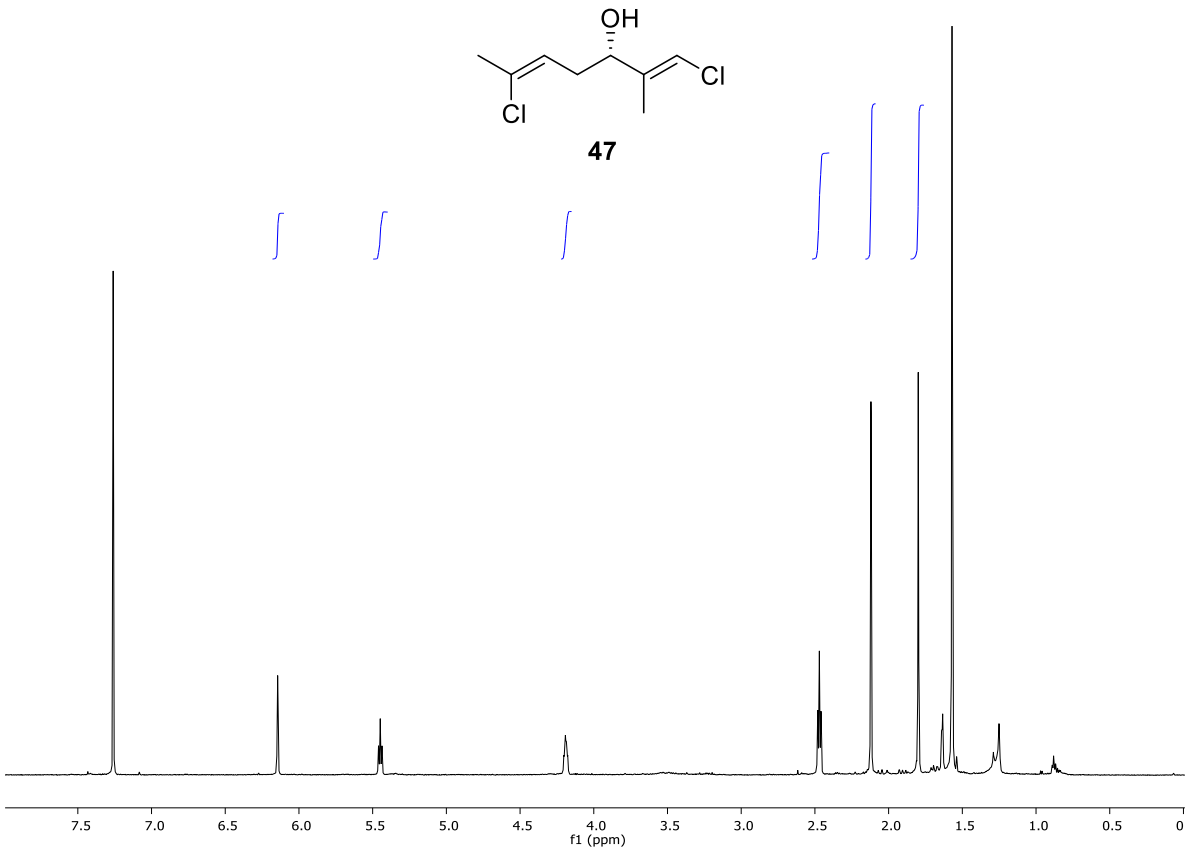
ROESY Spectrum (600 MHz, CD_3OD) of **46**



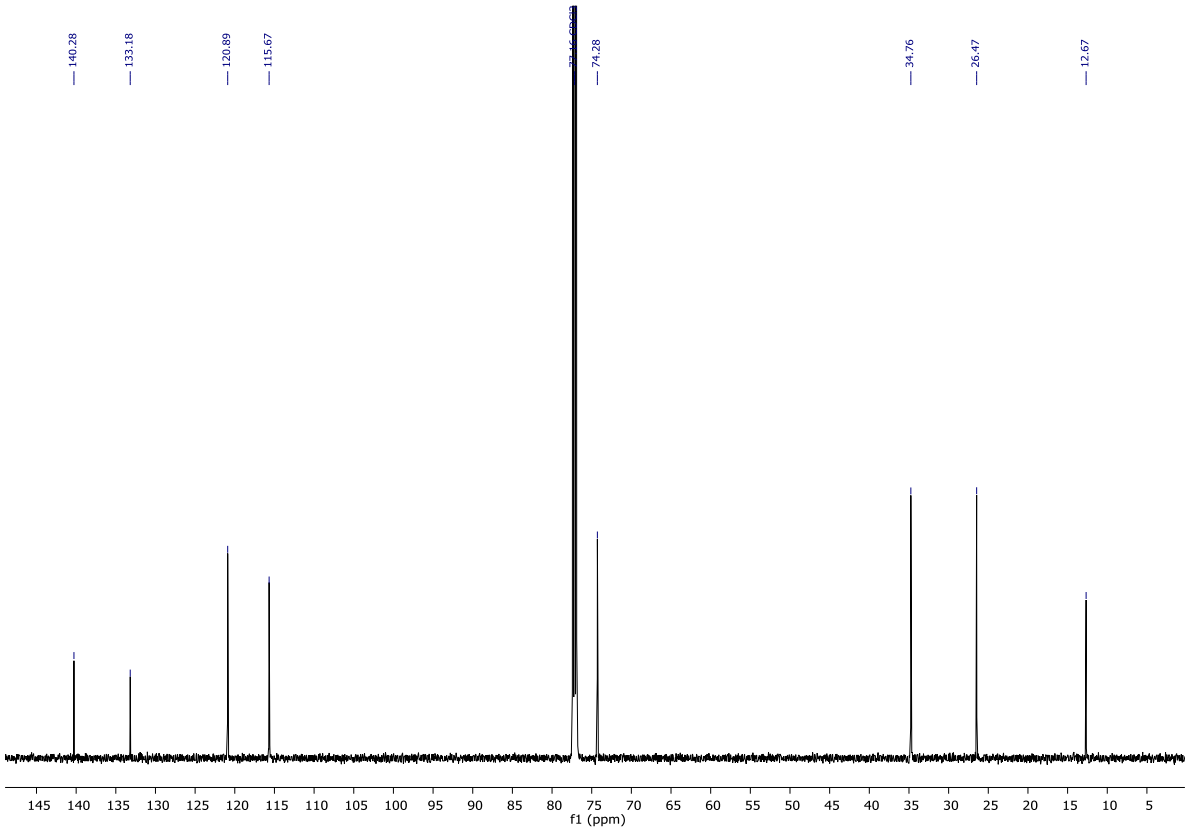
(–)-HRAPCIMS Spectrum of **46**



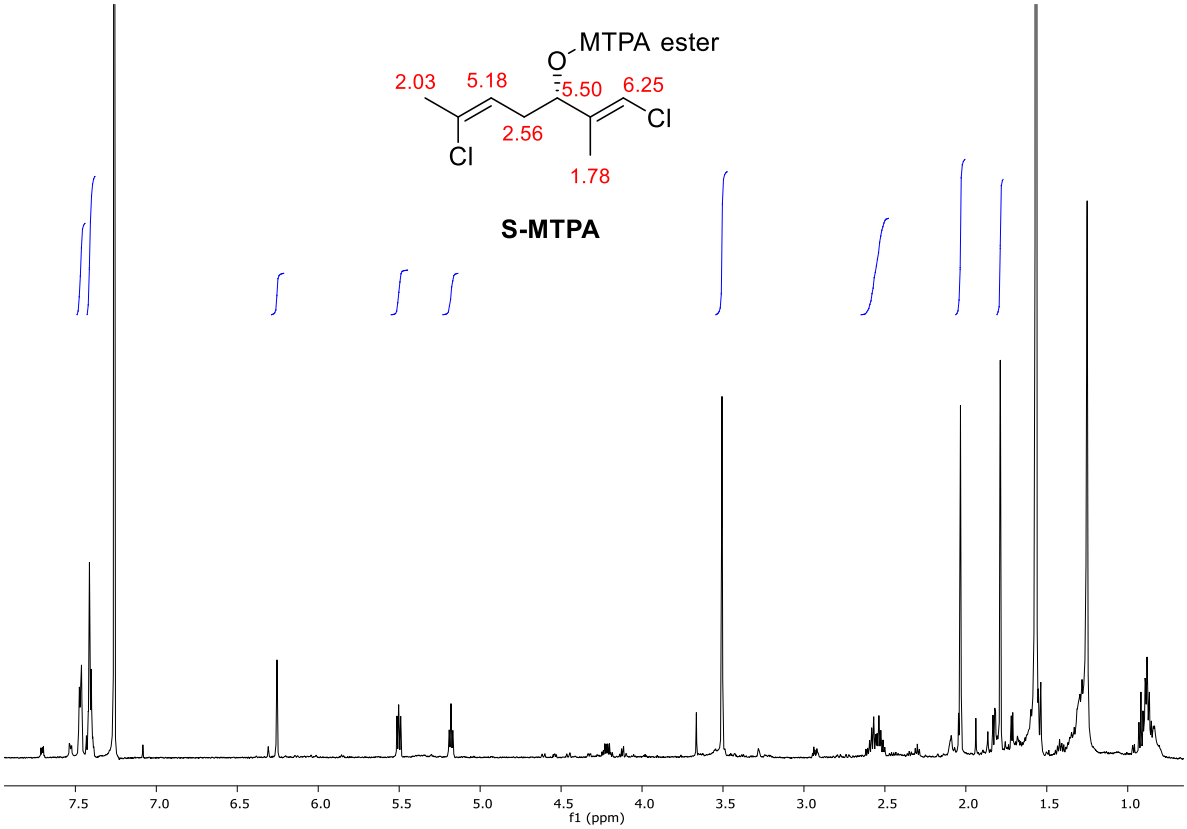
IR Spectrum (film) of **46**



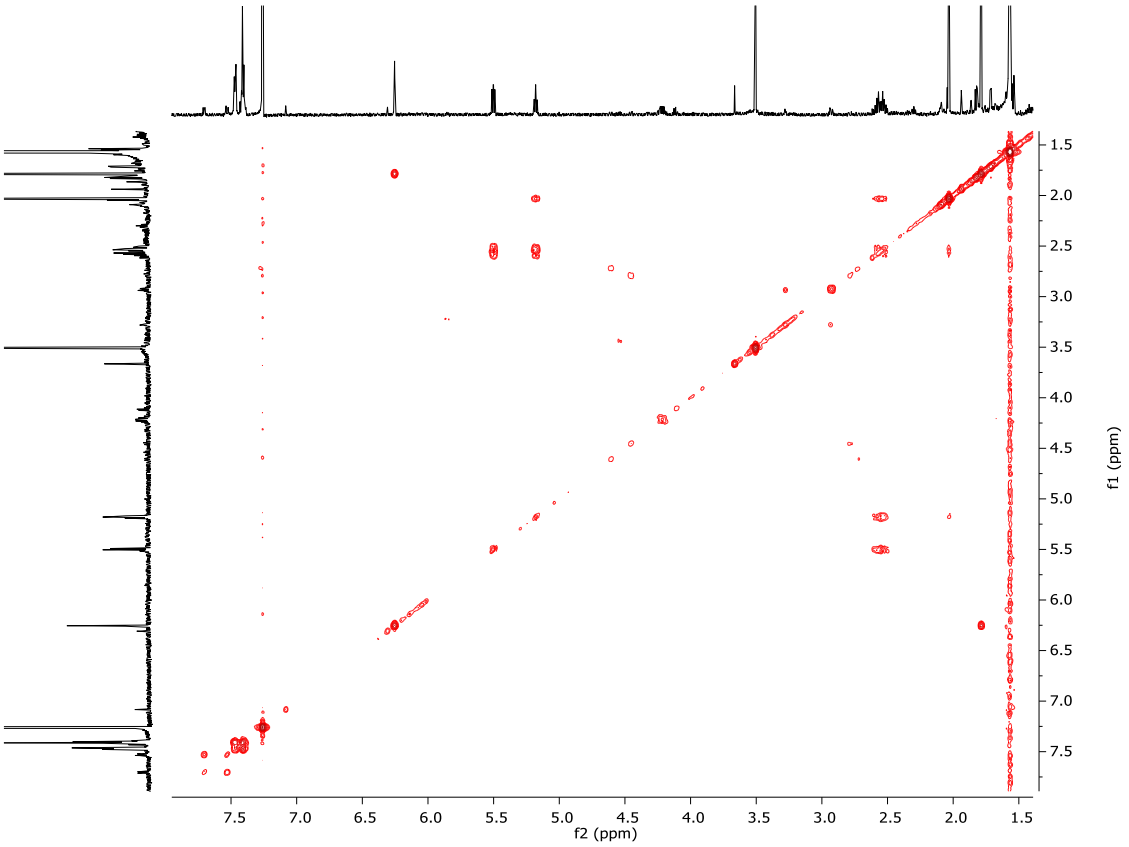
^{1}H Spectrum (600 MHz, $CDCl_3$) of **47**



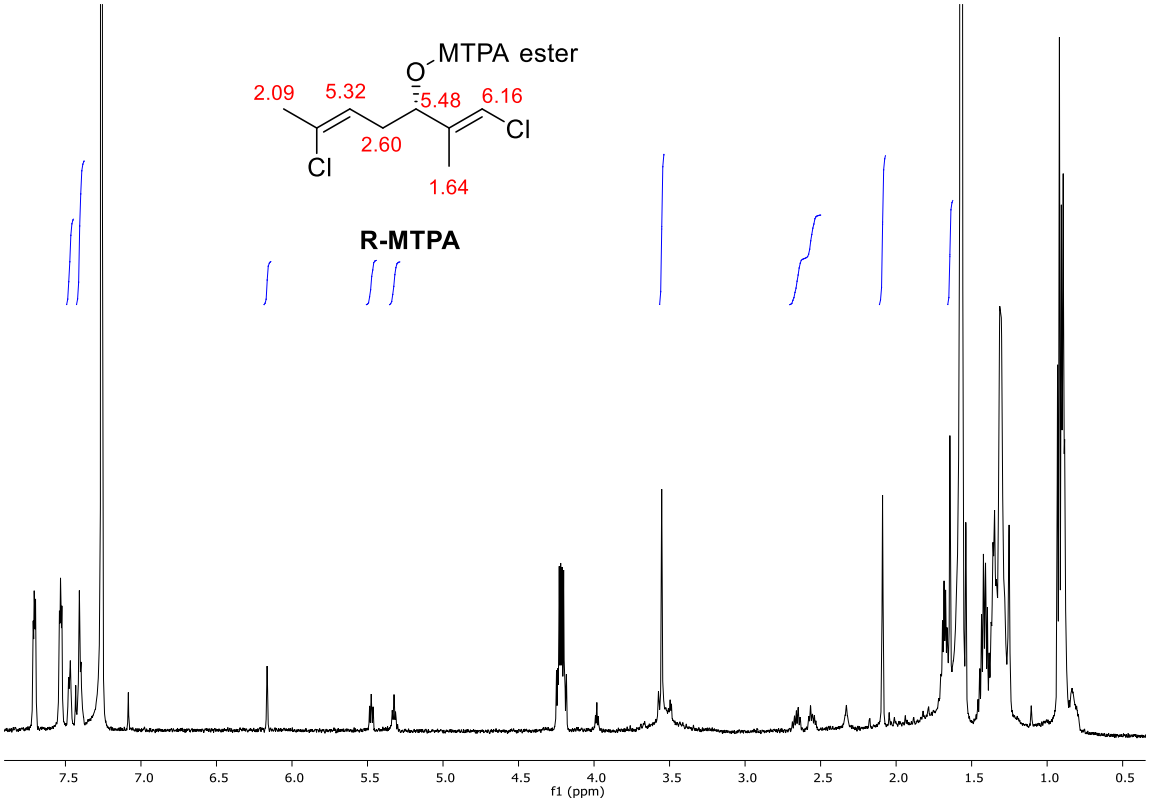
^{13}C Spectrum (150 MHz, $CDCl_3$) of **47**



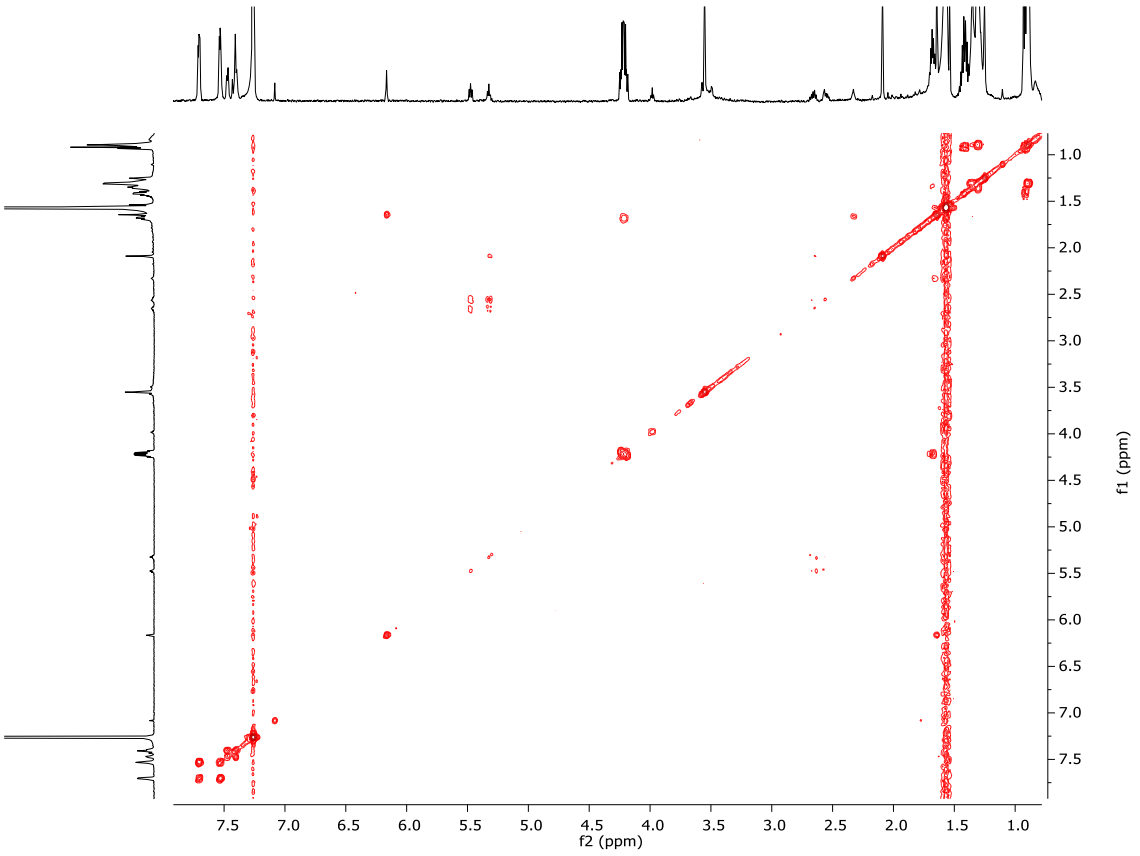
^{1}H Spectrum (600 MHz, $CDCl_3$) of *S*-MTPA ester of **47** (**47a**)



COSY Spectrum (600 MHz, $CDCl_3$) of *S*-MTPA ester of **47** (**47a**)

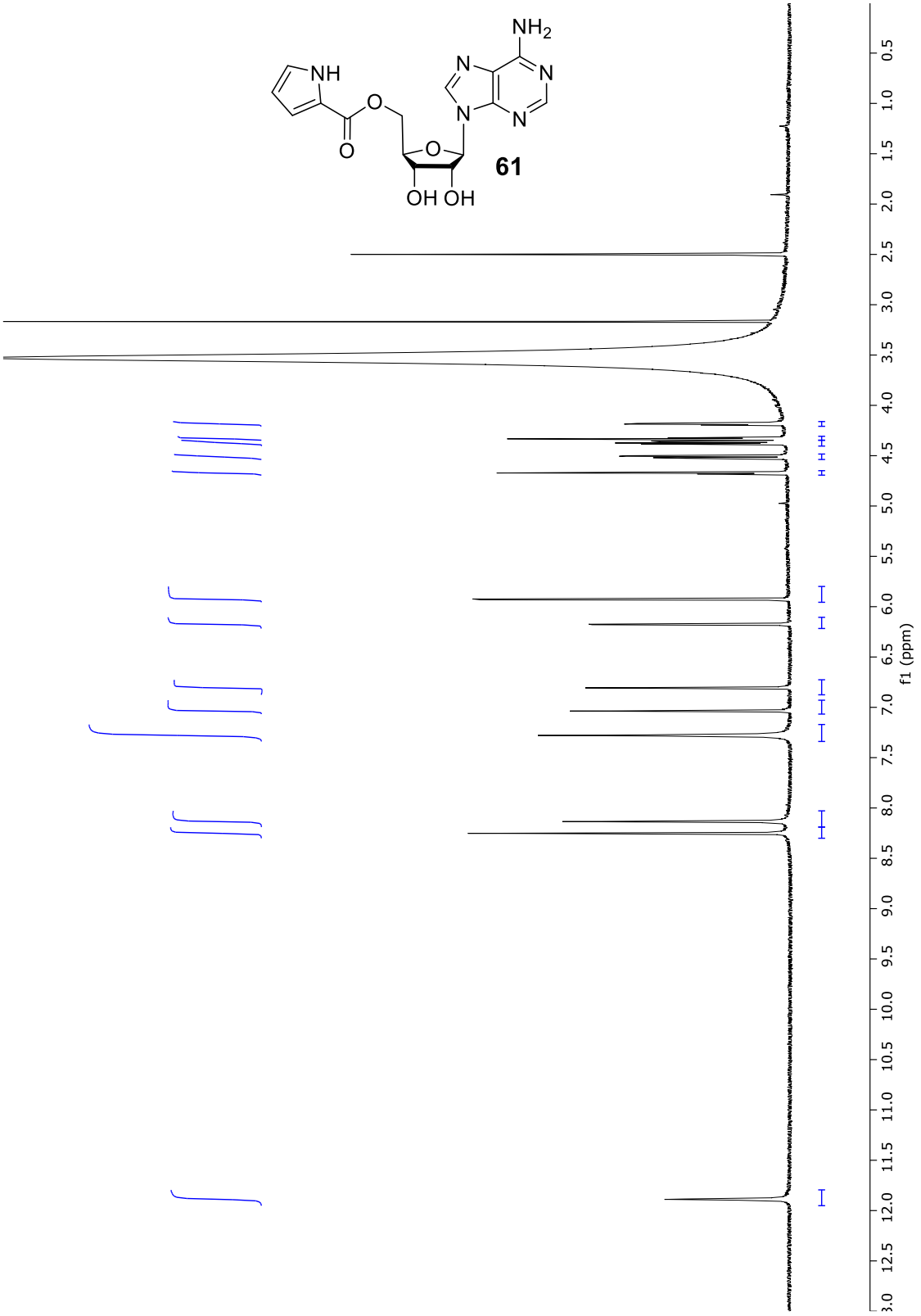


^{1}H Spectrum (600 MHz, $CDCl_3$) of *R*-MTPA ester of **47** (**47b)**

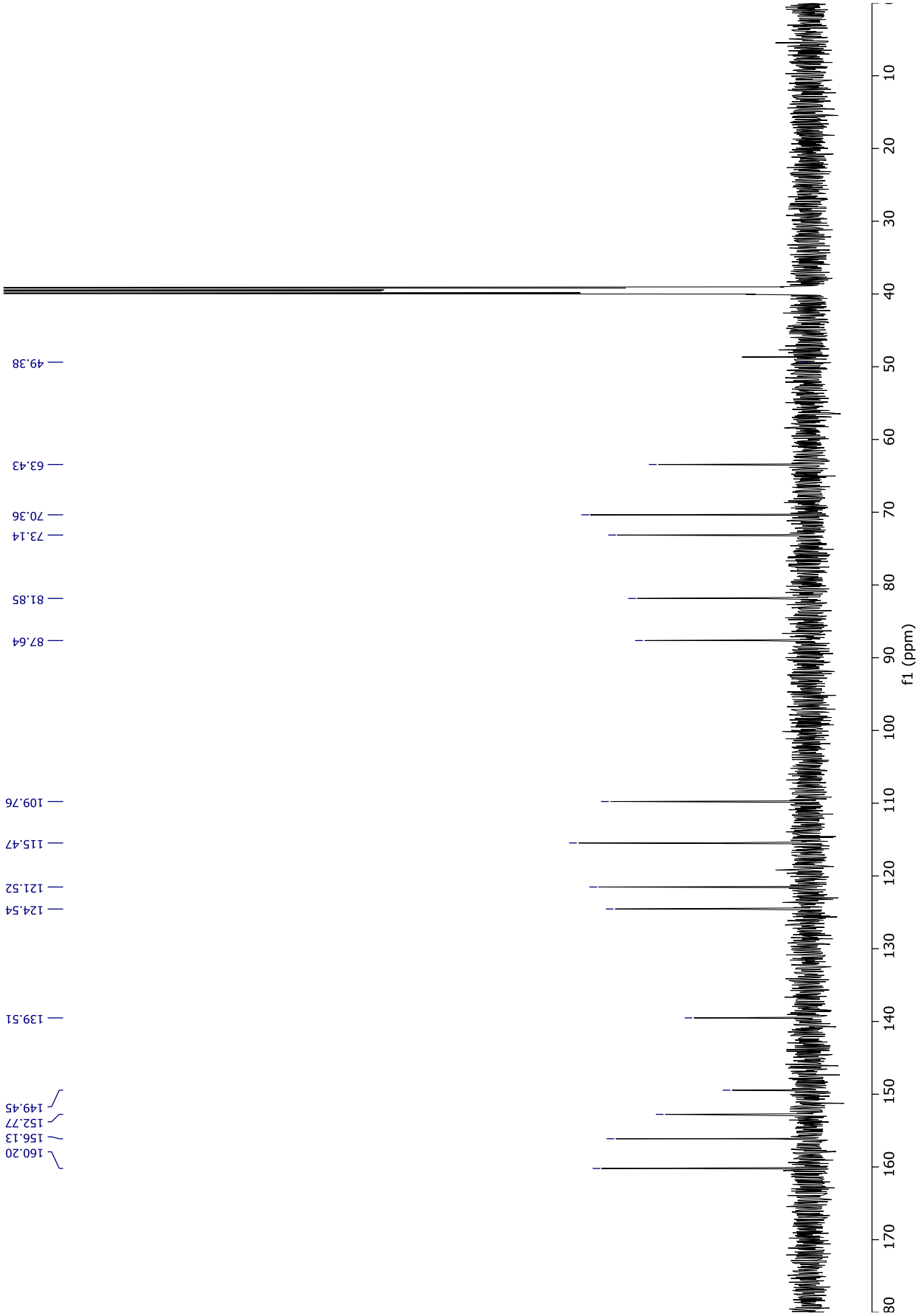


COSY Spectrum (600 MHz, $CDCl_3$) of *R*-MTPA ester of **47** (**47a**)

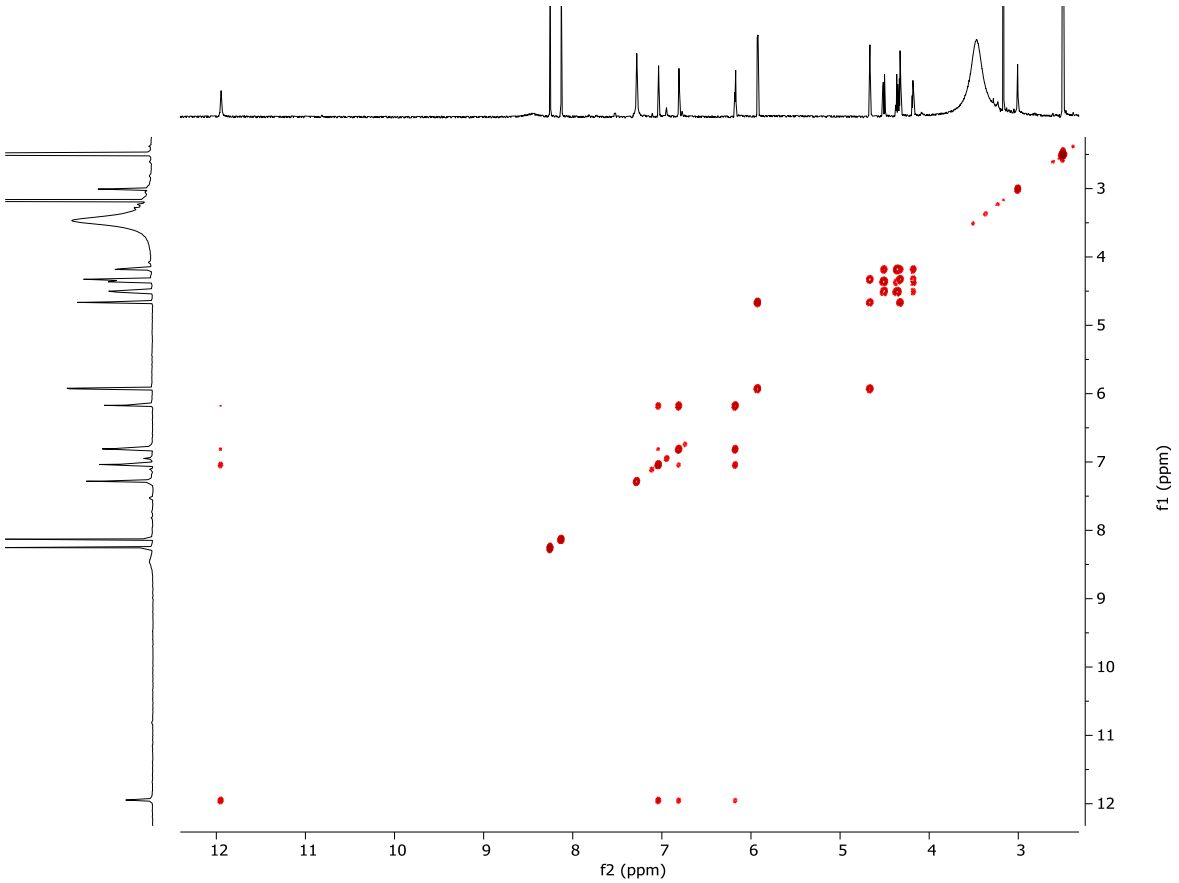
Appendix 6 – *Nelliella nelliiformis* Spectra



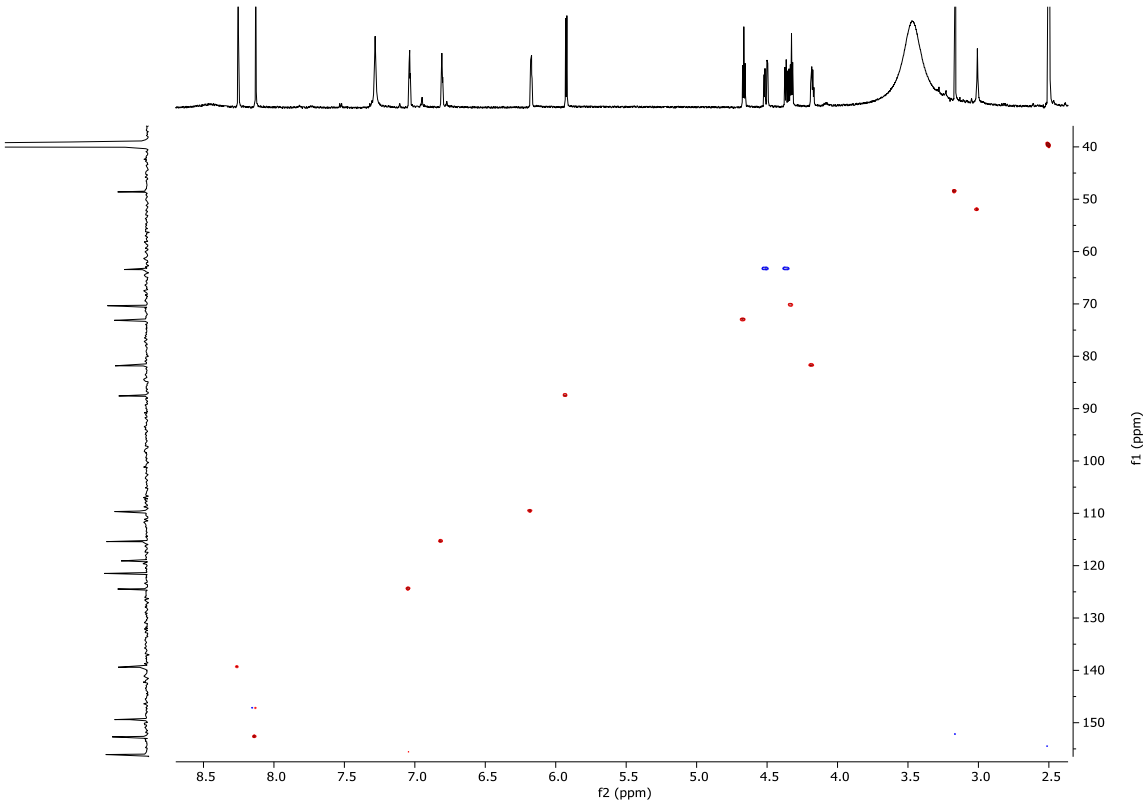
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **61**



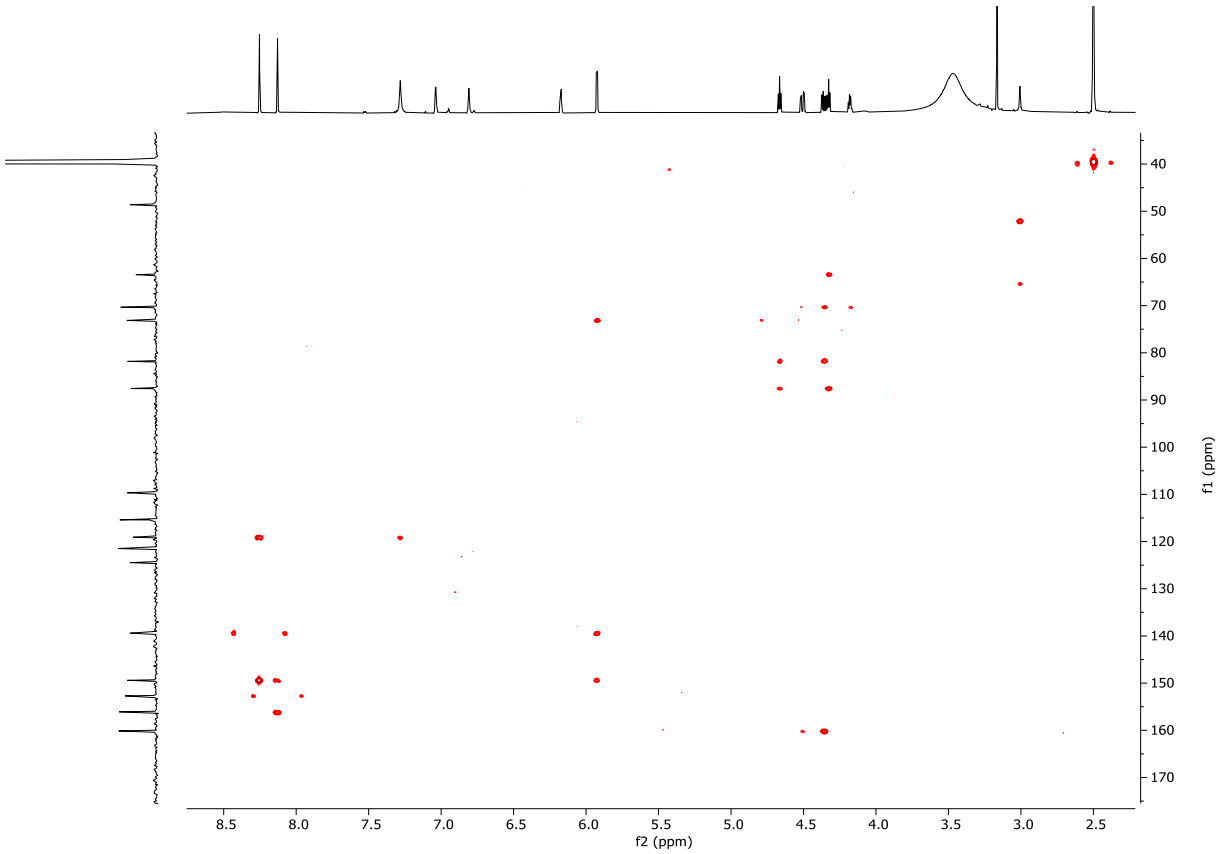
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **61**



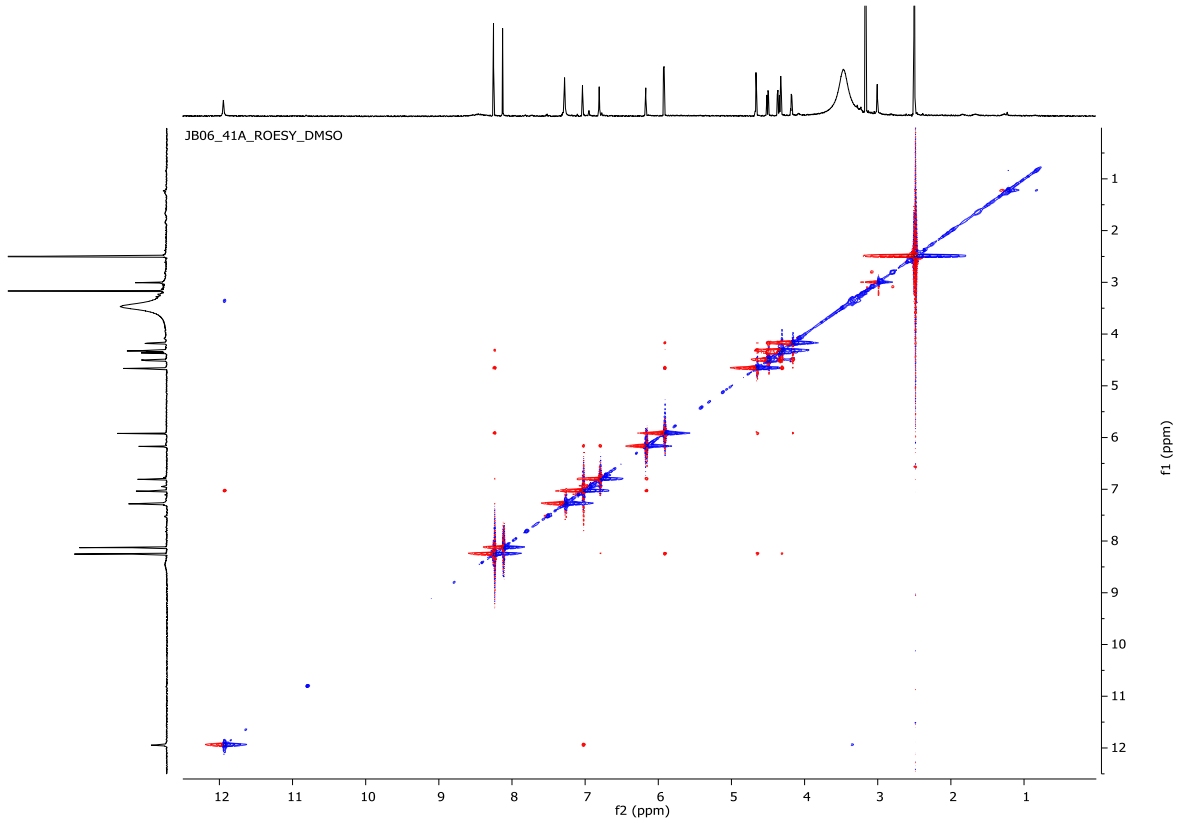
COSY Spectrum (600 MHz, DMSO-d_6) of **61**



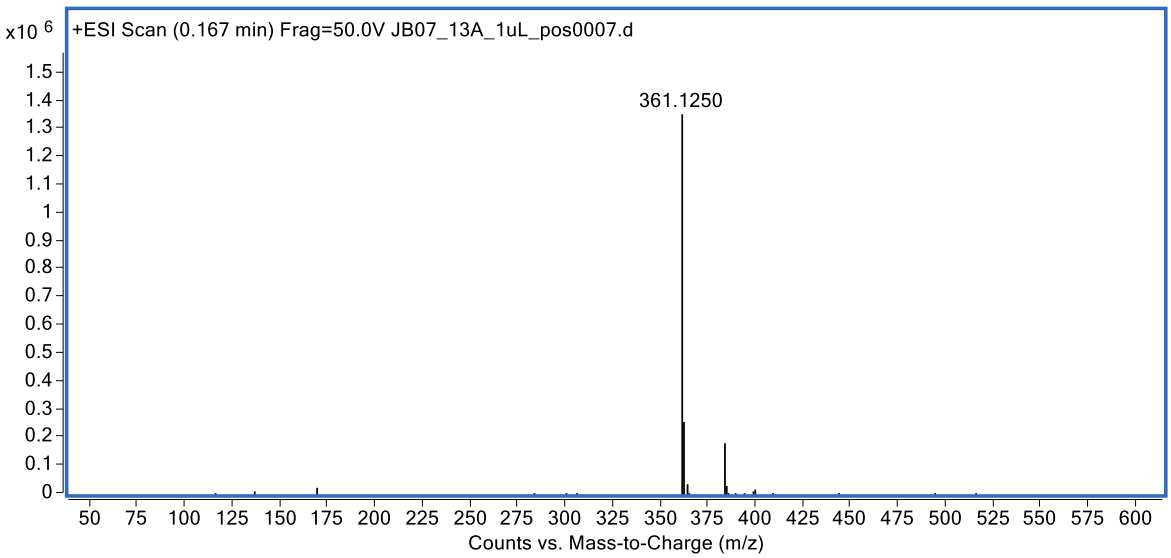
HSQC Spectrum (600 MHz, DMSO-d_6) of **61**



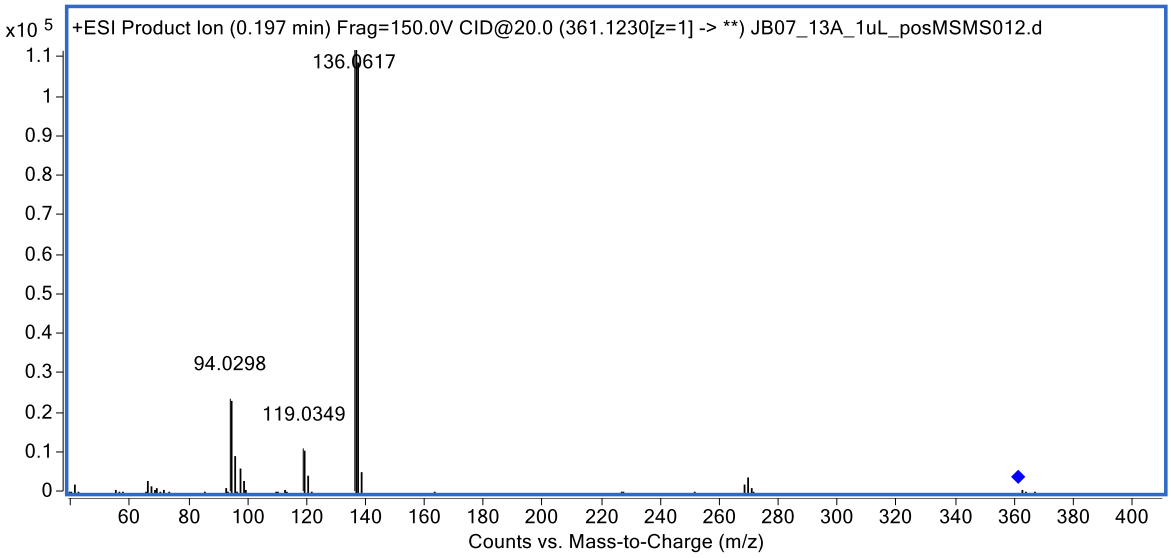
HMBC Spectrum (600 MHz, DMSO-d_6) of **61**



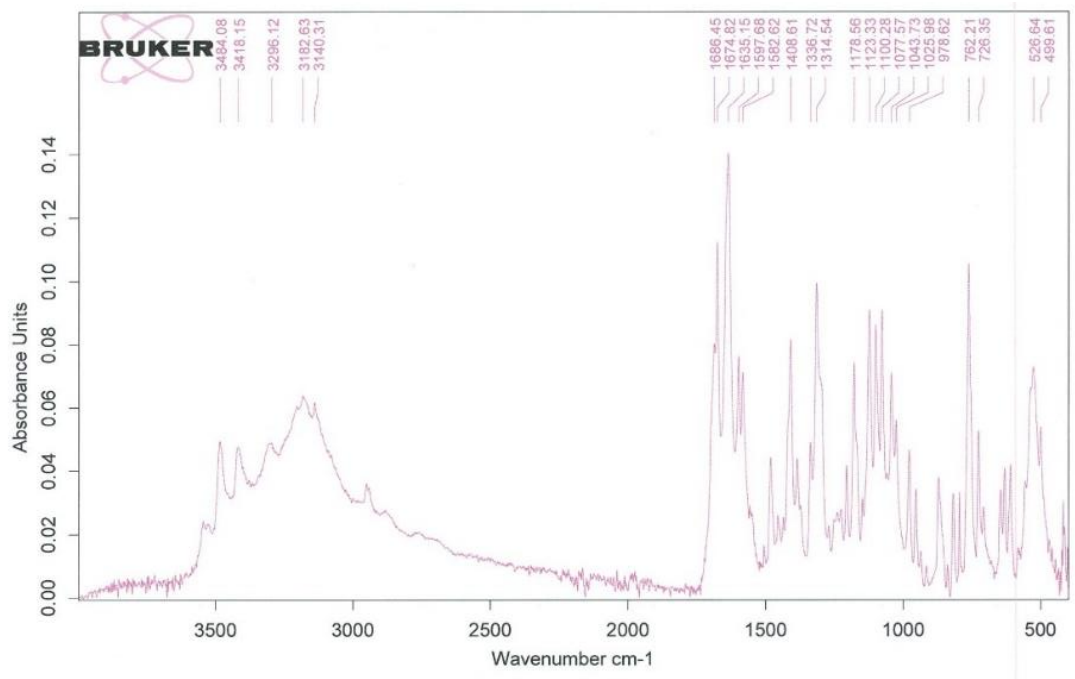
ROESY Spectrum (600 MHz, DMSO-d_6) of **61**



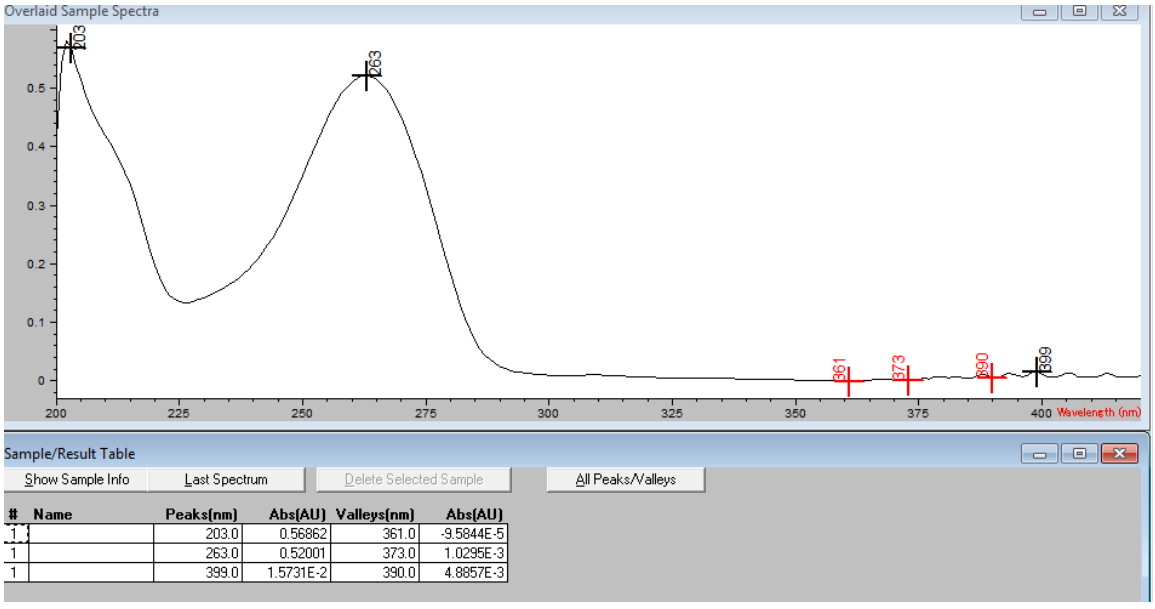
(+)-HRESIMS spectrum of **61**



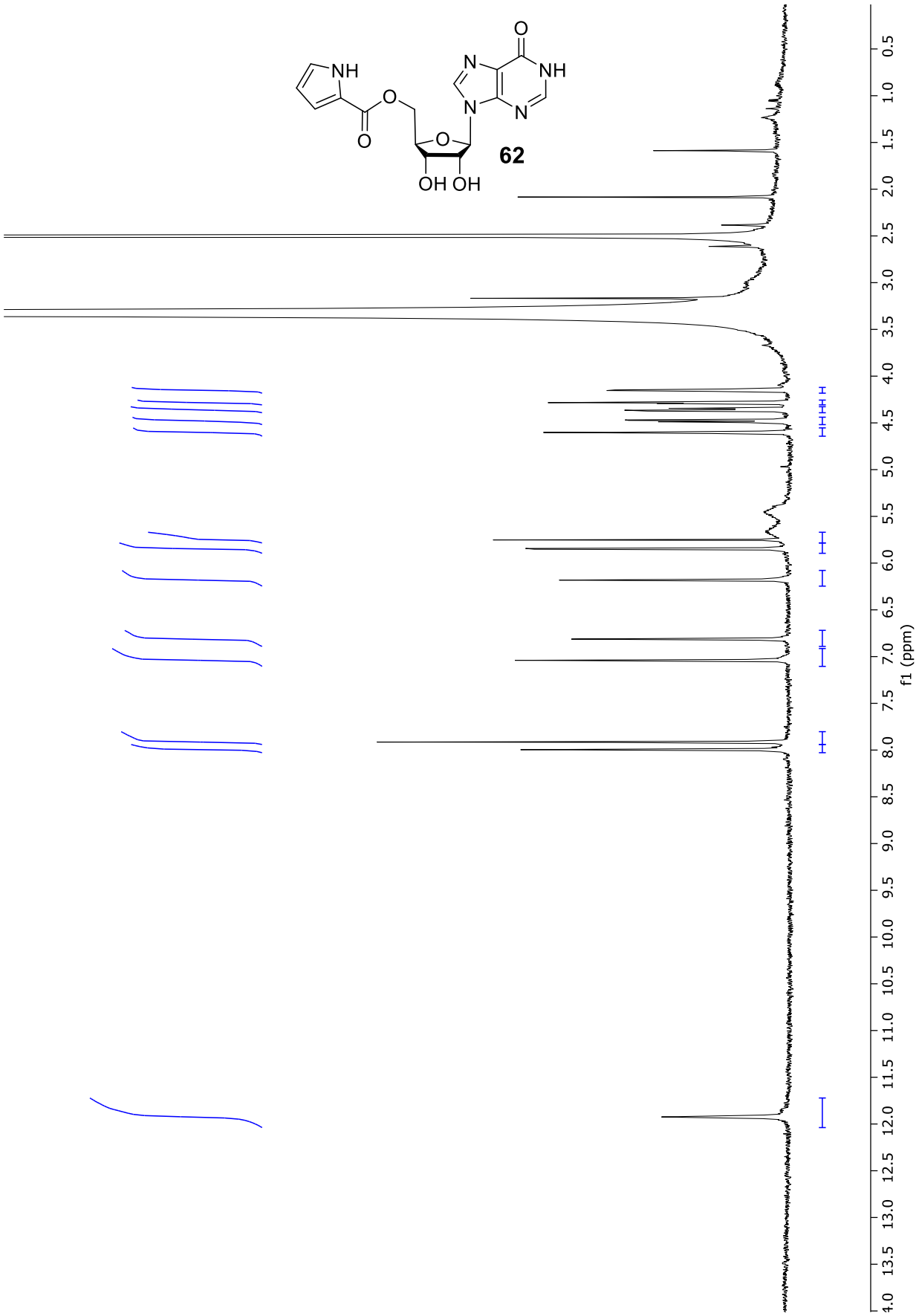
(+)-HRESIMS/MS spectrum of **61**



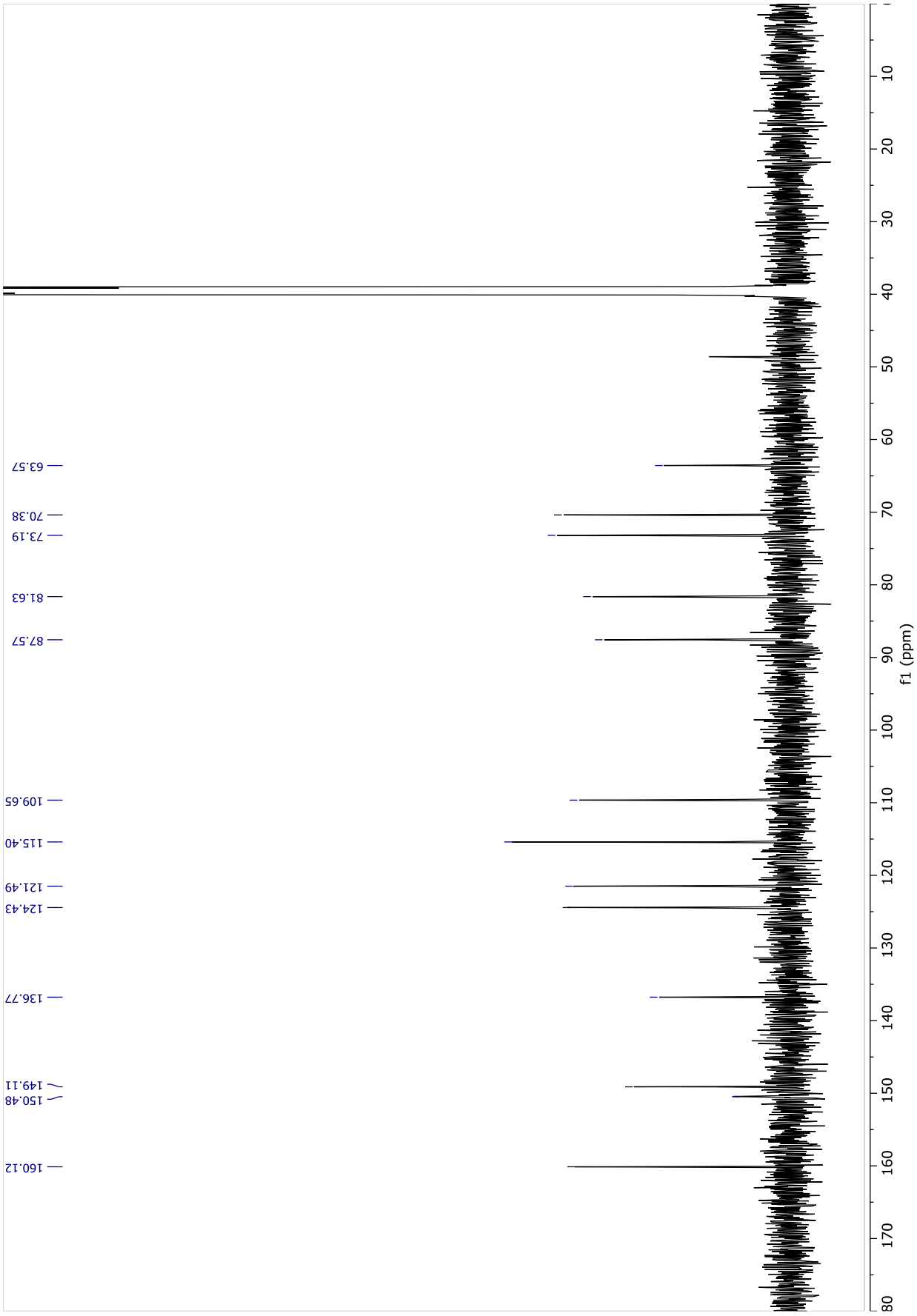
IR spectrum of **61**



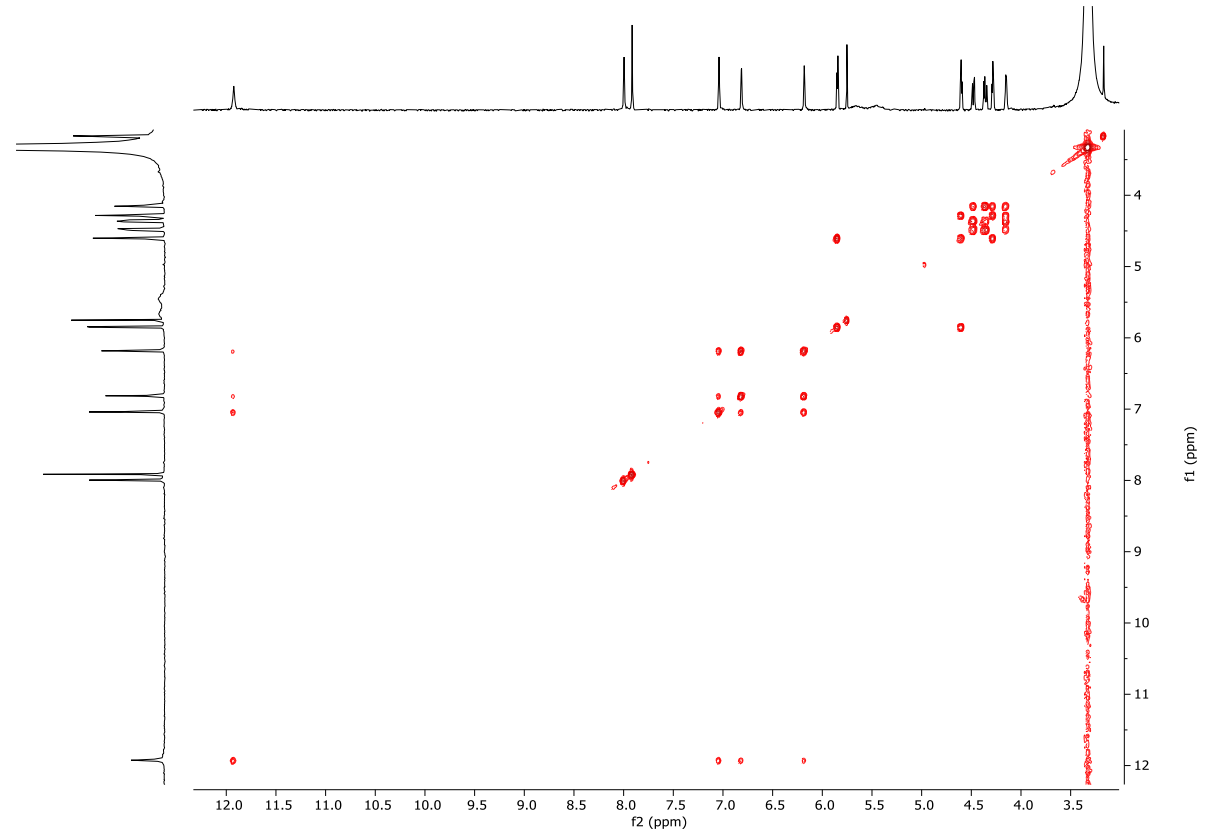
UV/vis spectrum of **61**



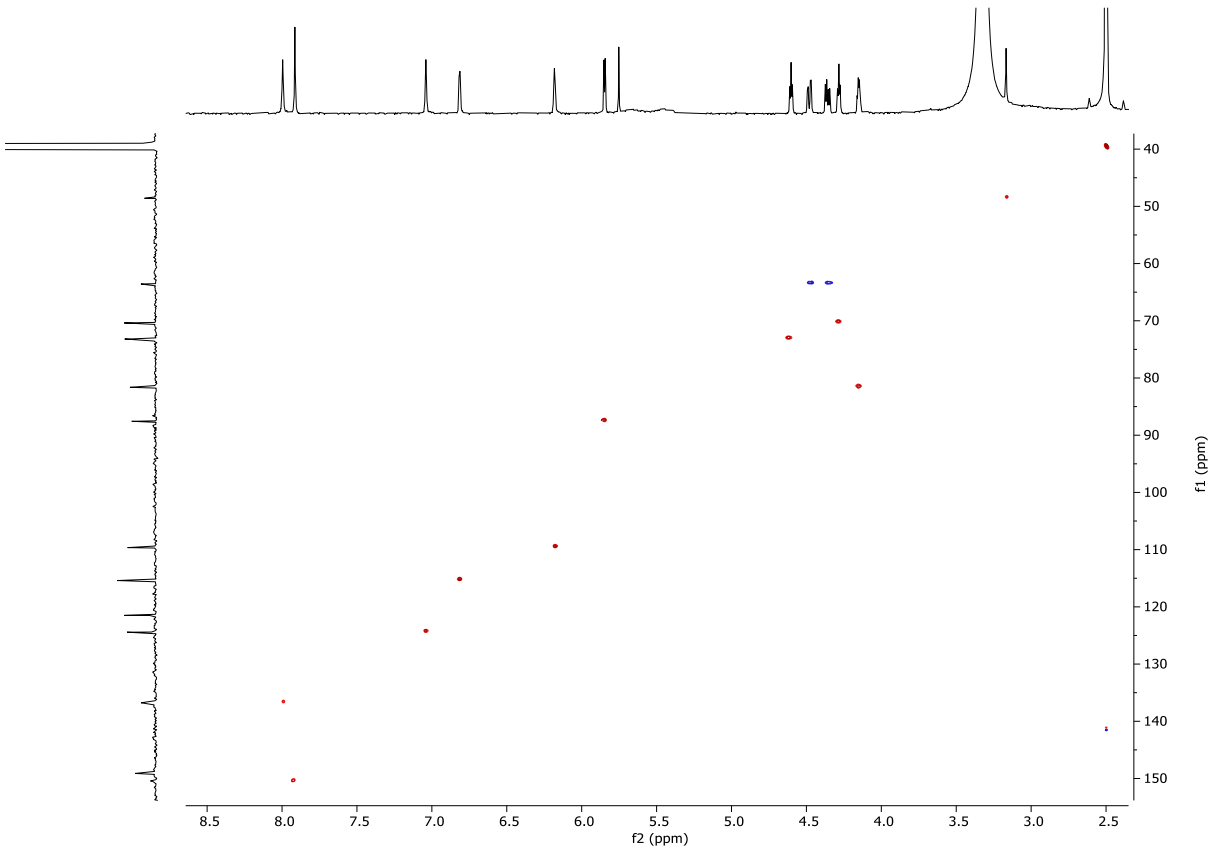
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **62**



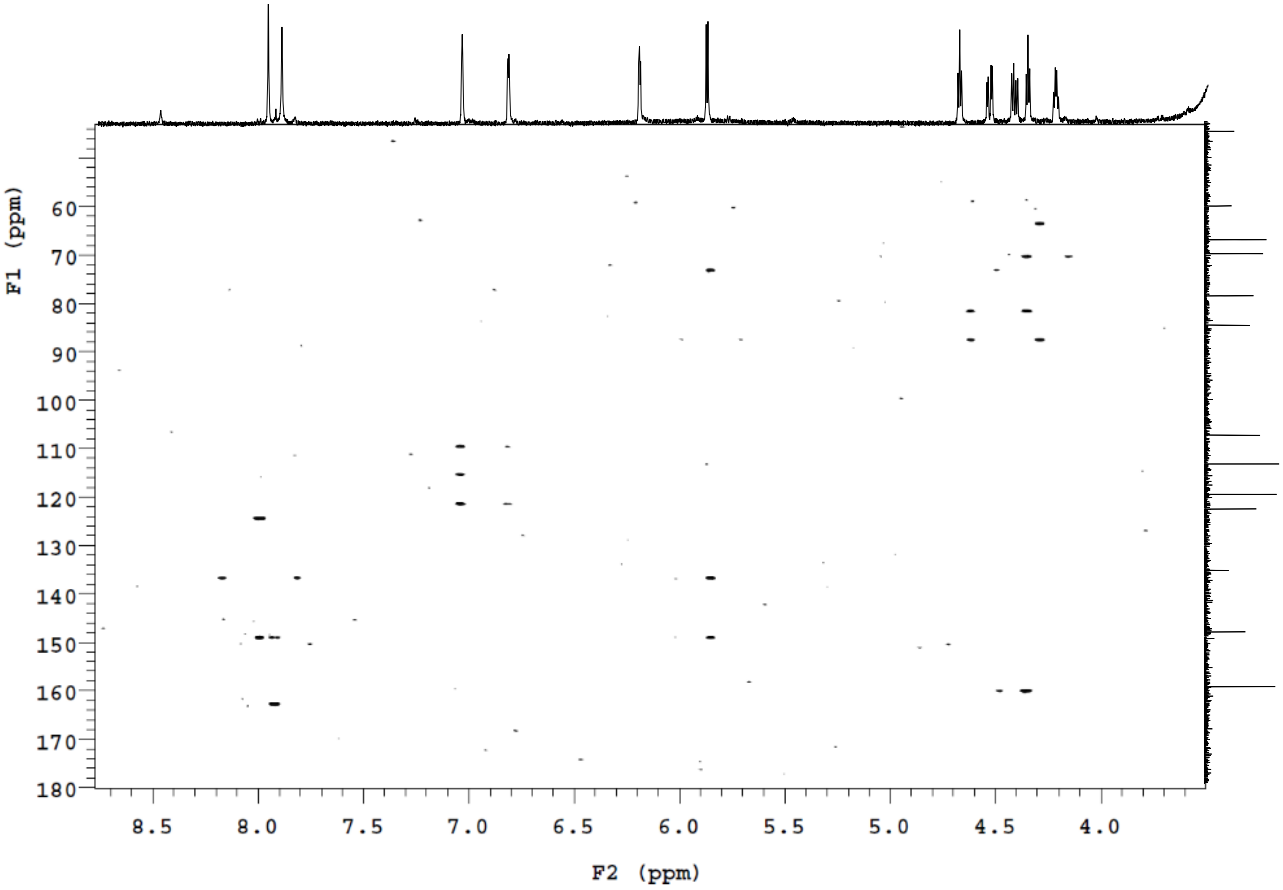
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **62**



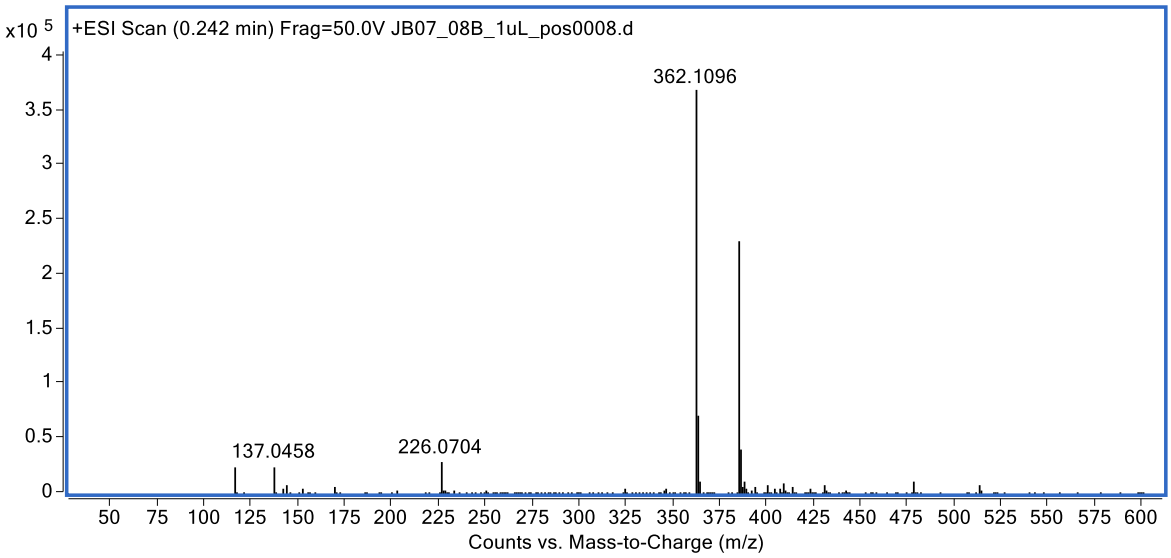
COSY Spectrum (600 MHz, DMSO-d_6) of **62**



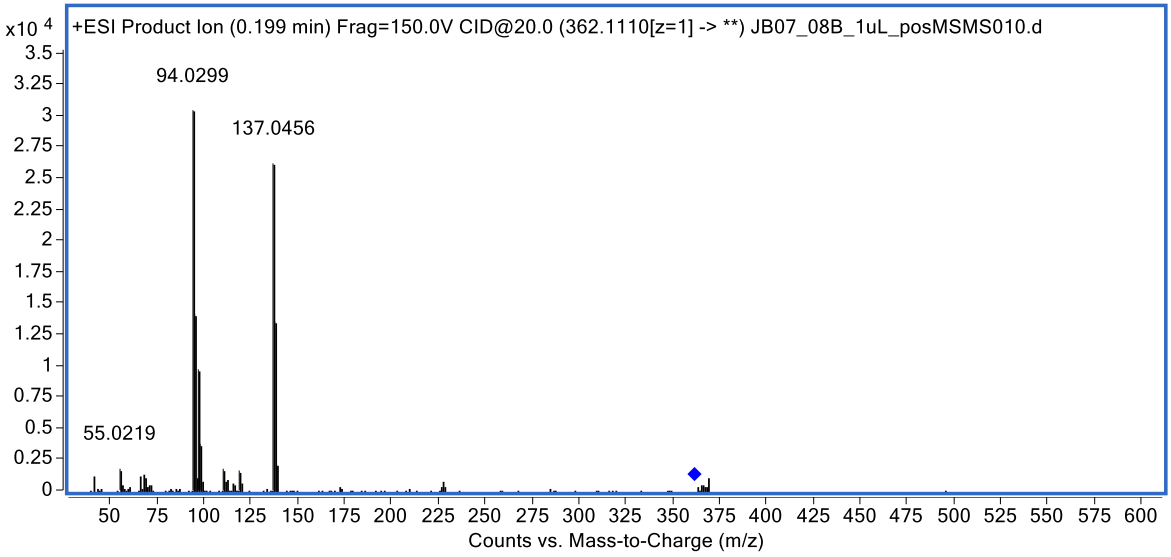
HSQC Spectrum (600 MHz, DMSO-d_6) of **62**



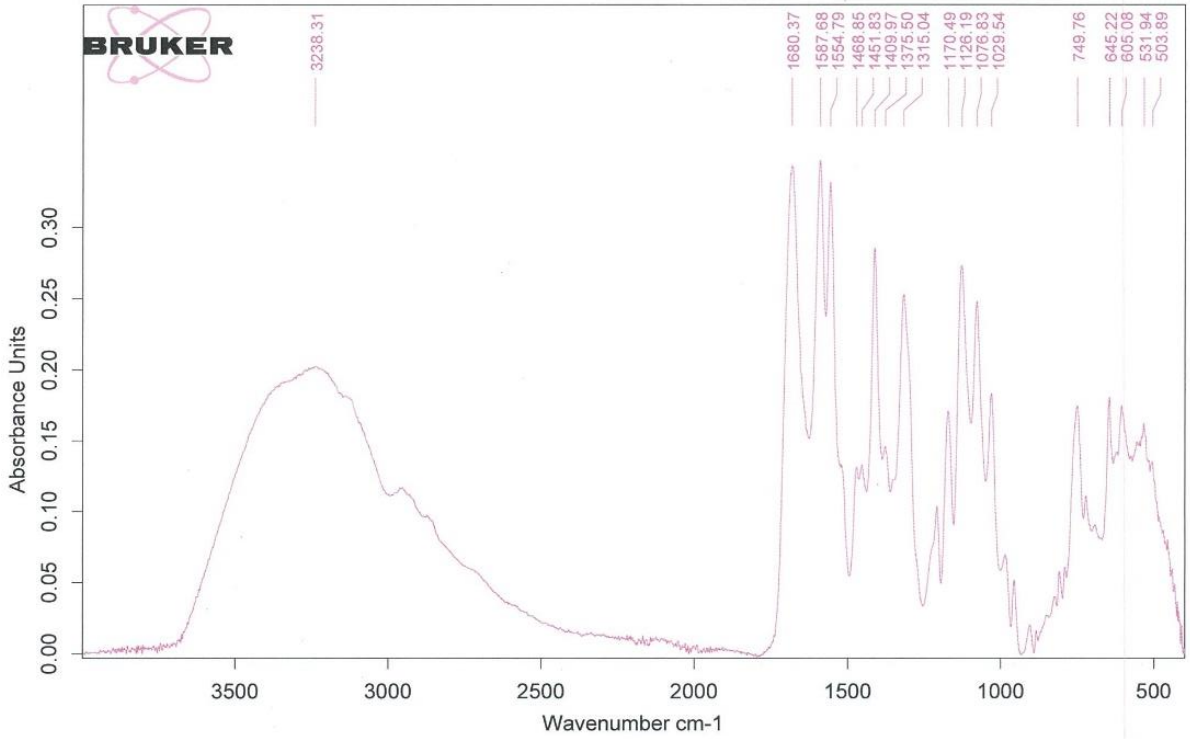
HMBC Spectrum (600 MHz, DMSO-d_6) of **62**



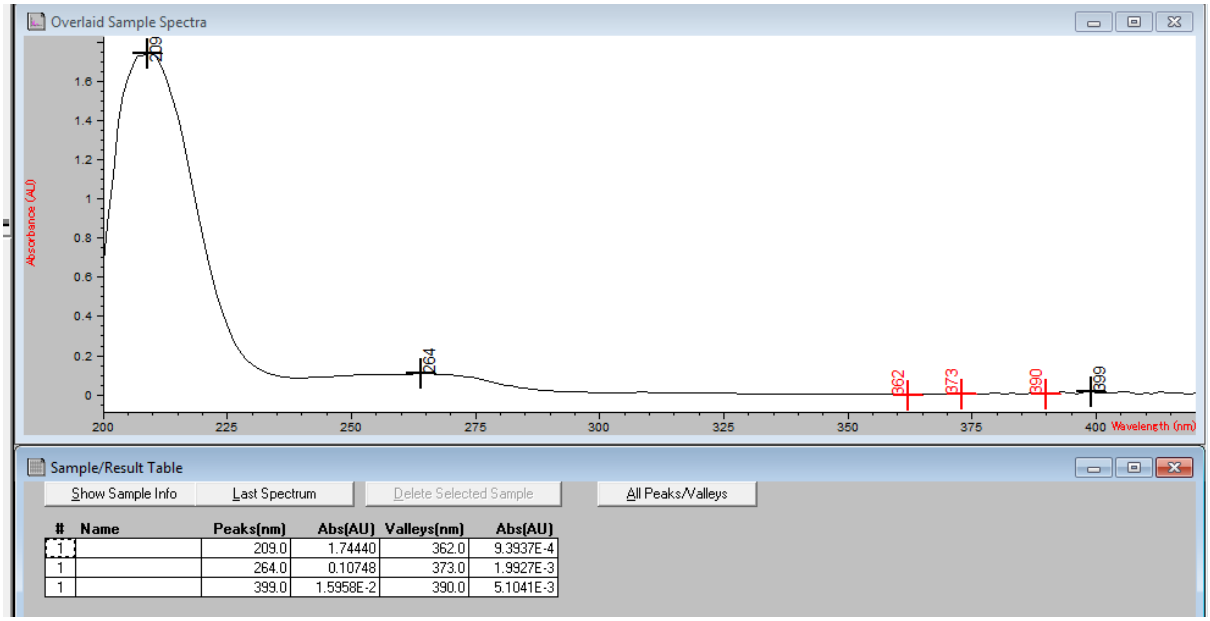
(+)-HRESIMS spectrum of **62**



(+)-HRESIMS/MS spectrum of **62**



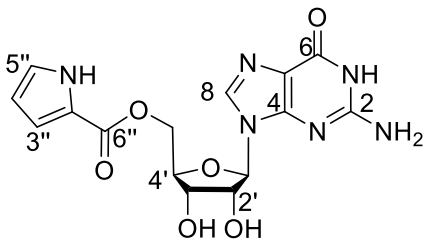
IR spectrum of **62**



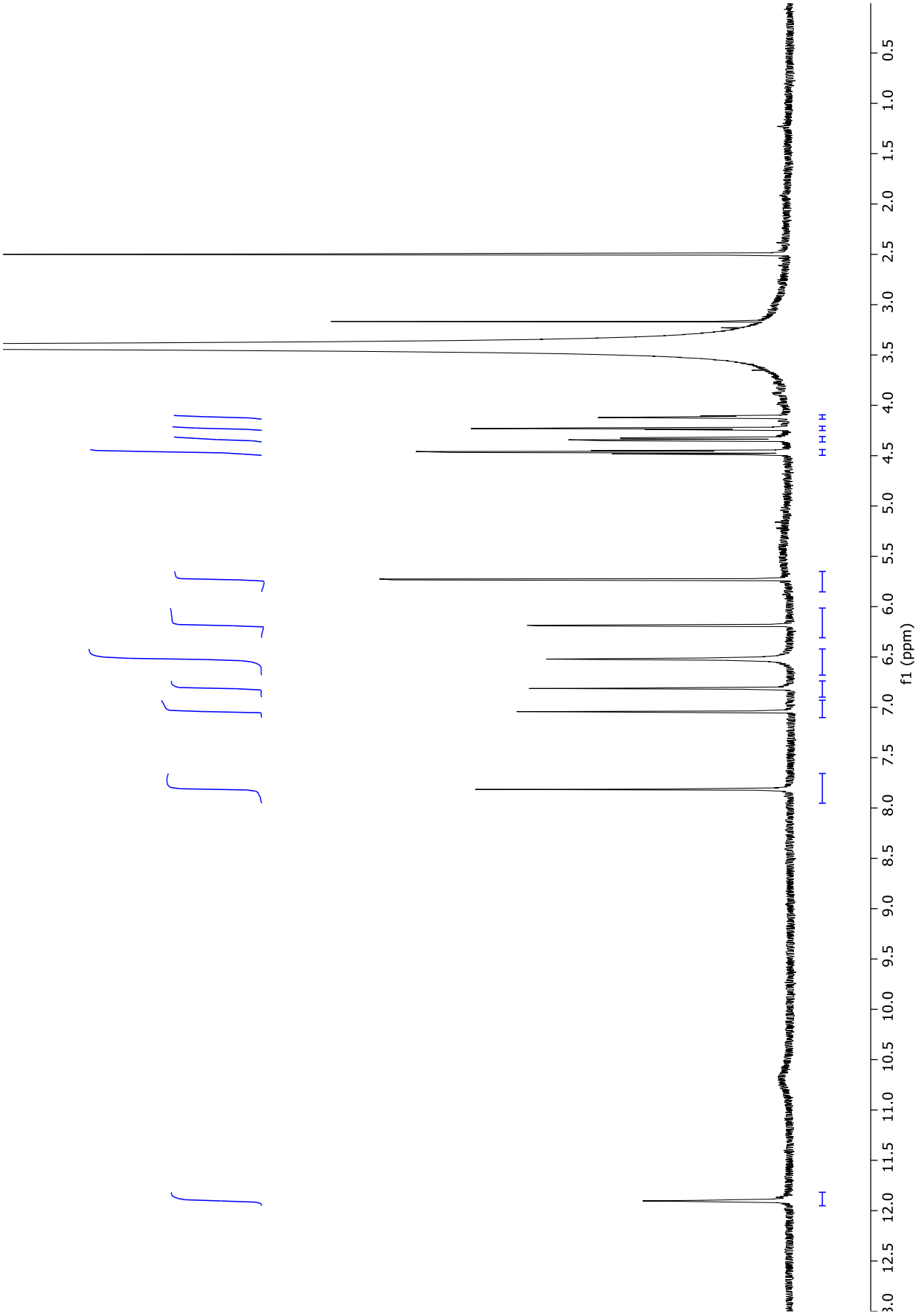
UV/vis spectrum of **62**

Table A1: ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for guanosyl-6'-pyrrole-2-carboxylate (**66**) in DMSO-d_6.

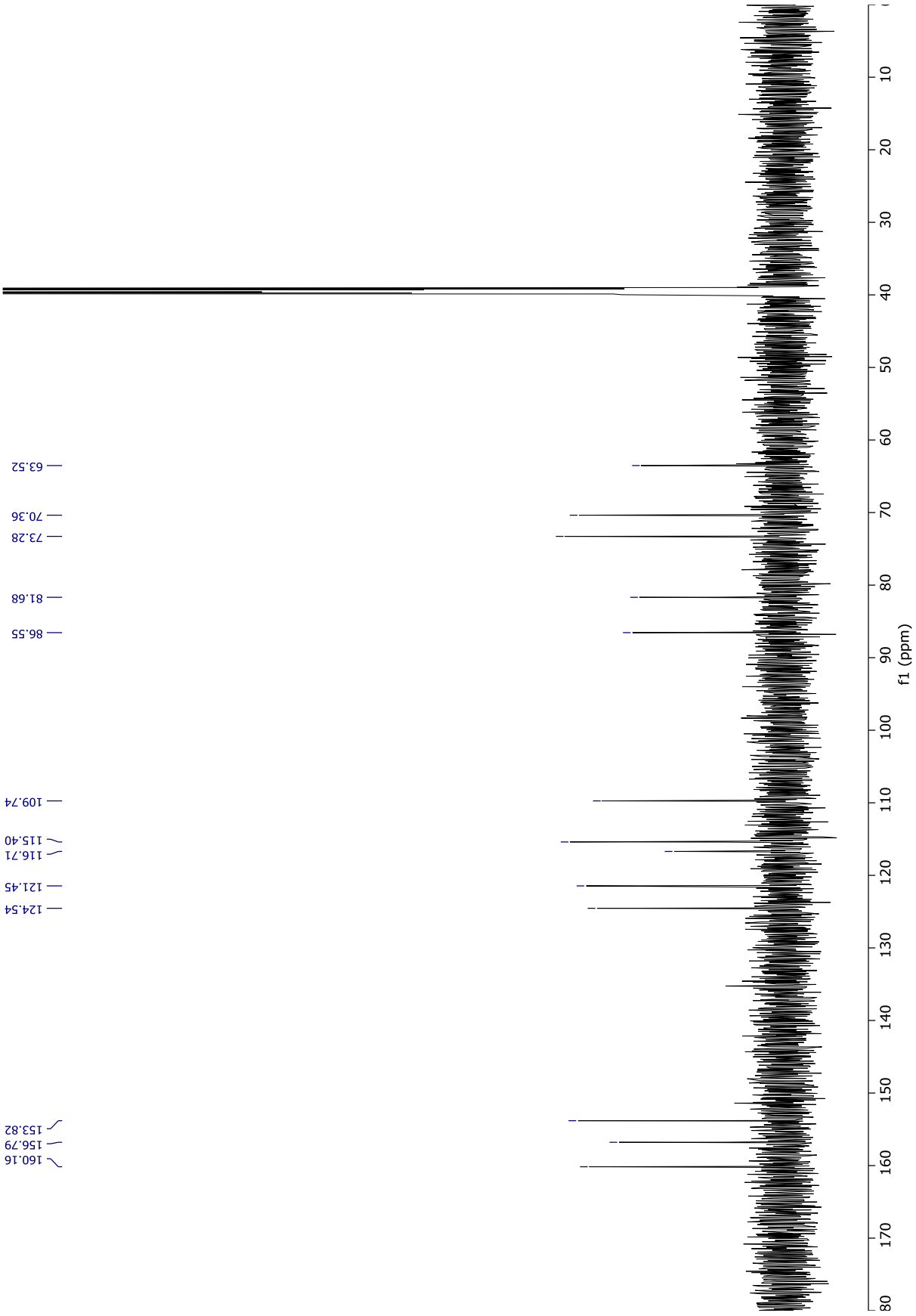
position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC
2	153.8			
NH_2–2		6.52 (s)		
4	151.3			
5	116.7			
6	156.8			
8	135.6	7.82 (s)		4, 5
1'	86.5	5.73 (d, 5.4)	2'	4, 8, 2'
2'	73.3	4.46 (t, 5.2)	1', 3'	1', 3', 4'
3'	70.4	4.23 (t, 4.8)	2', 4'	1', 5'
4'	81.7	4.12 (dt, 4.1, 5.6)	3', 5'	1', 5'
5'	63.5	4.48 (dd, 12.0, 3.8)	4'	3', 6"
		4.34 (dd, 12.0, 5.7)	4'	3', 4', 6"
NH–1"		11.9 (s)	3", 4", 5"	
2"	121.4			
3"	115.4	6.81 (m)	NH–1", 4", 5"	
4"	109.7	6.19 (m)	NH–1", 3", 5"	
5"	124.5	7.04 (m)	NH–1", 3", 4"	
6"	160.1			



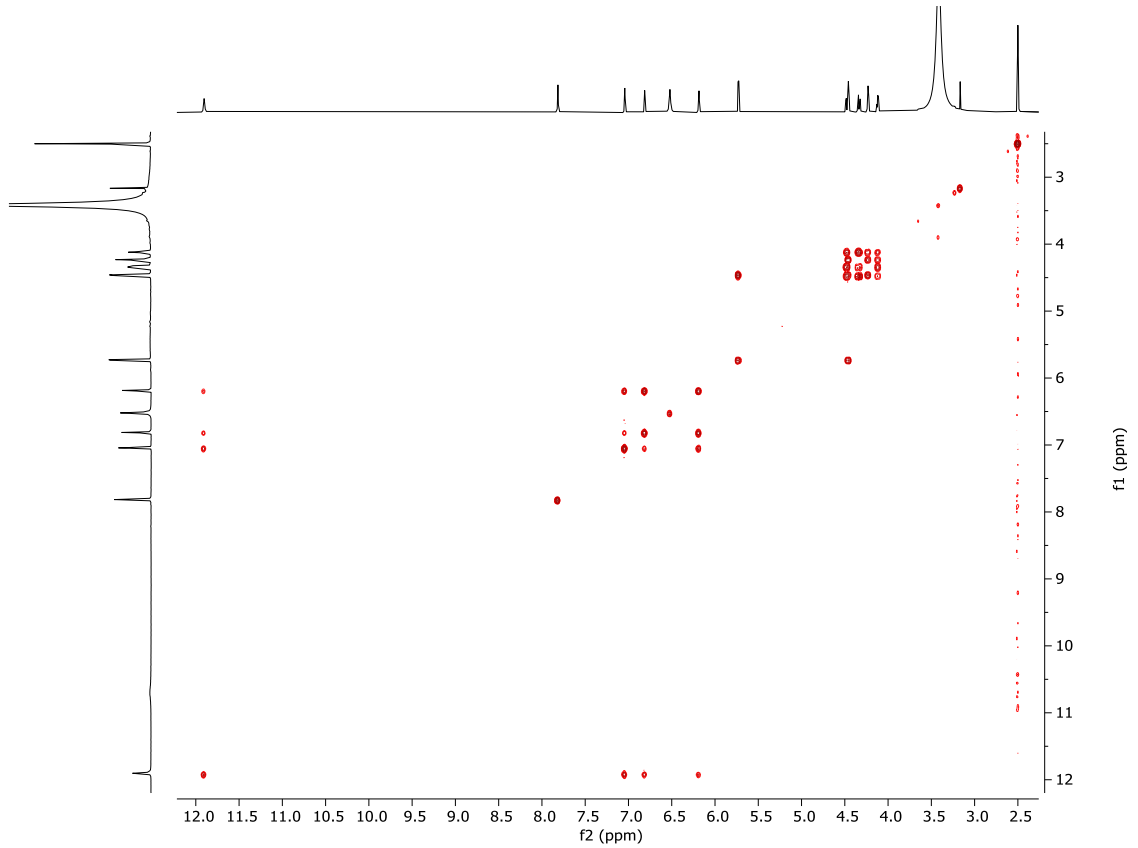
66



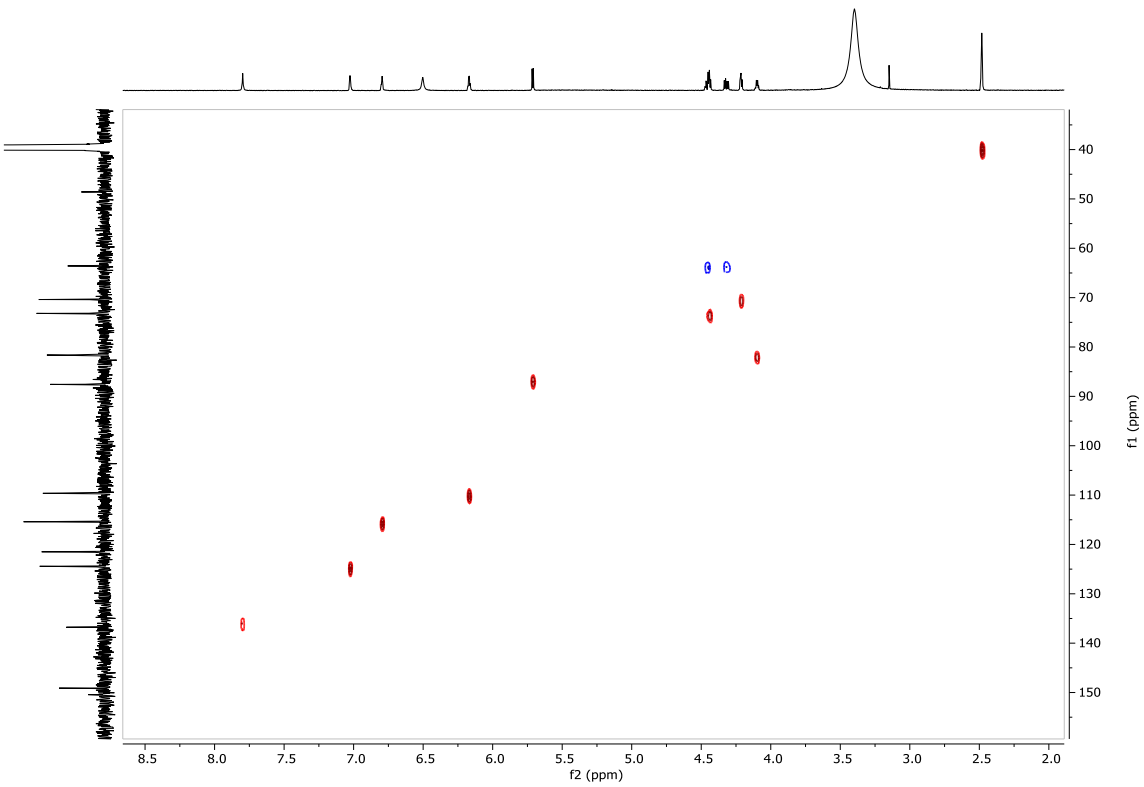
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **66**



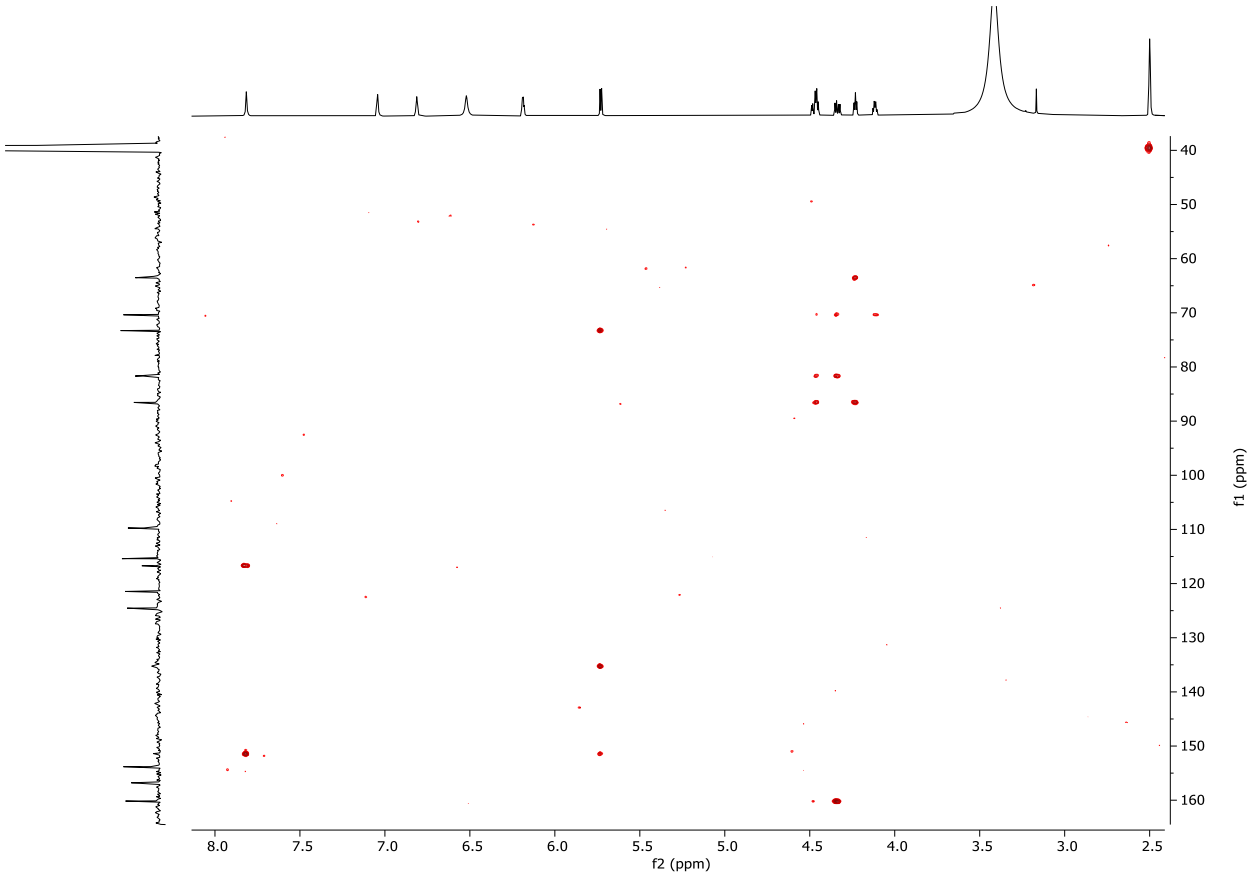
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **66**



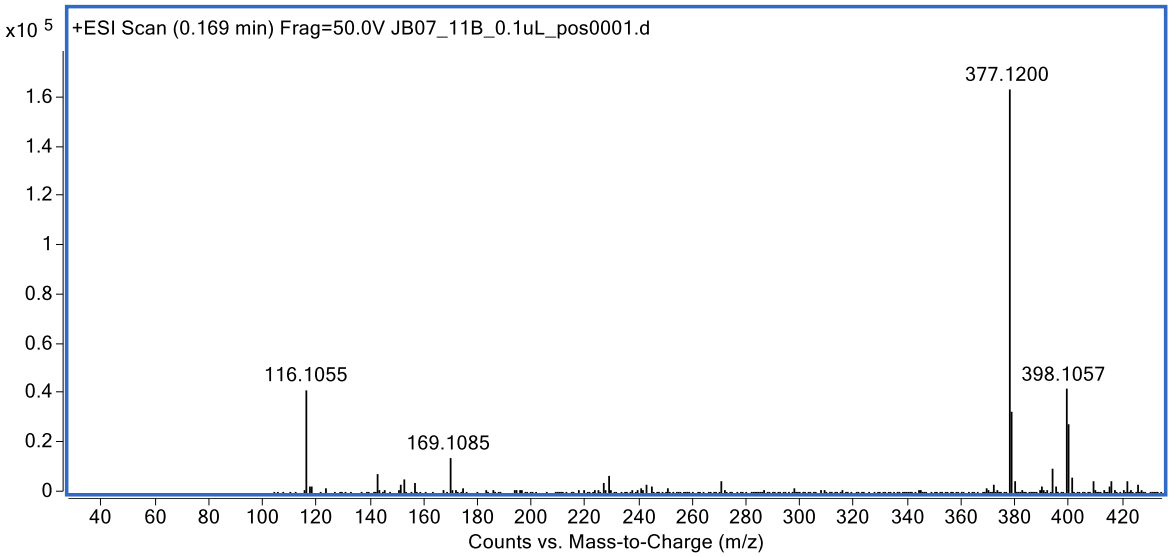
COSY Spectrum (600 MHz, DMSO-d_6) of **66**



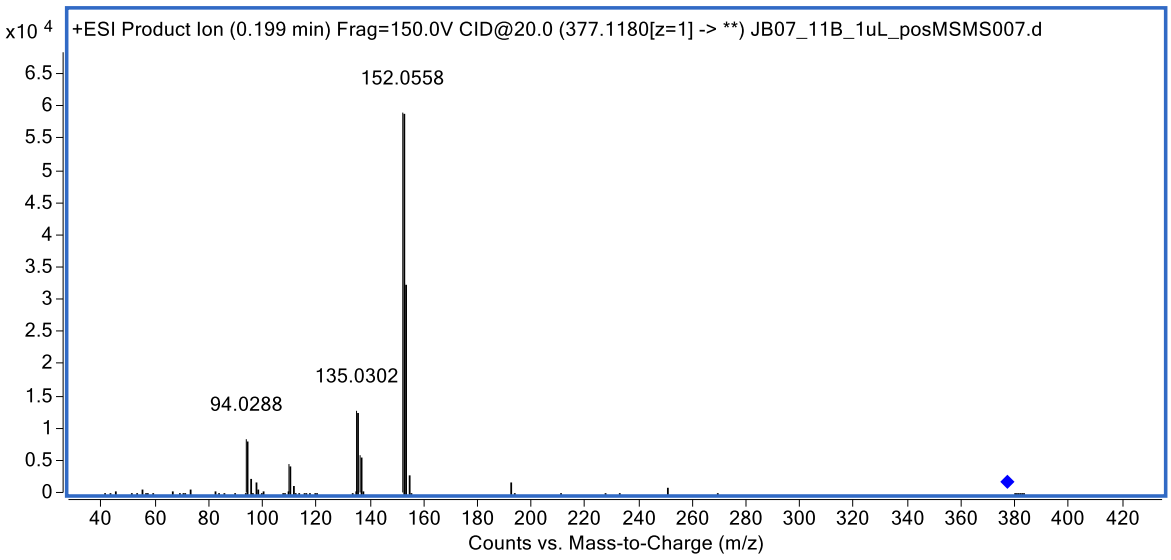
HSQC Spectrum (600 MHz, DMSO-d_6) of **66**



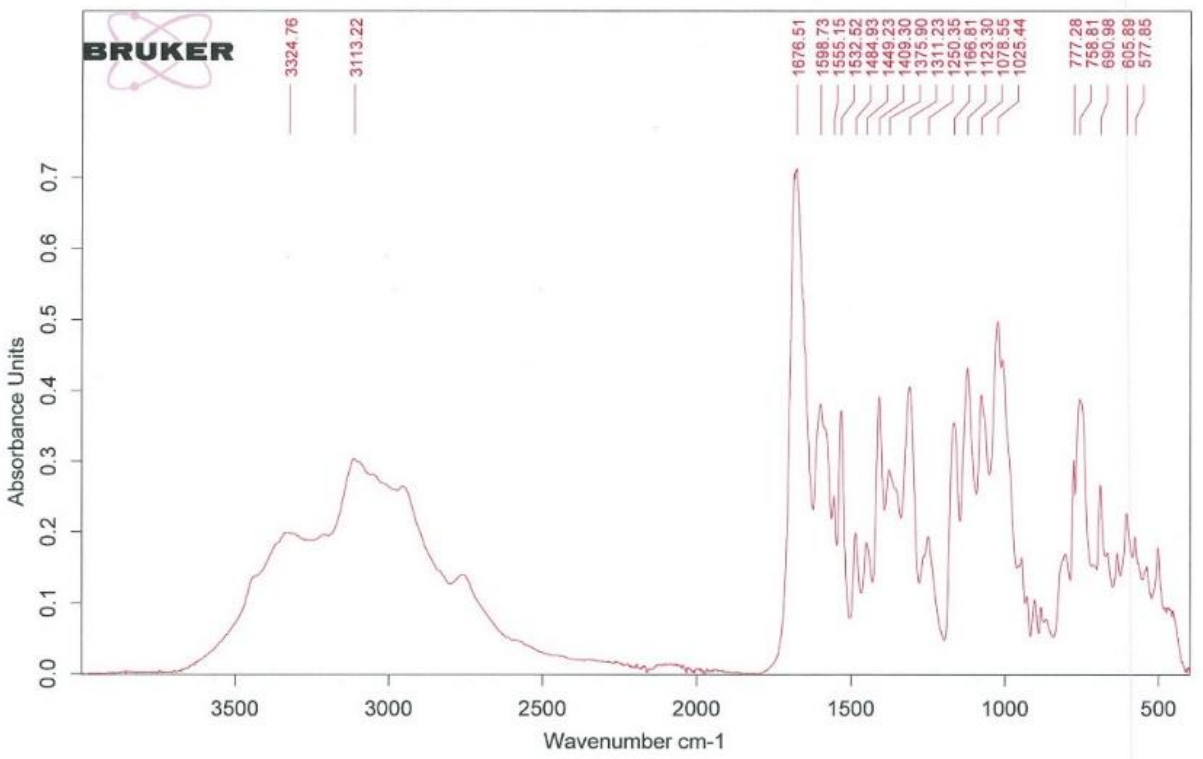
HMBC Spectrum (600 MHz, DMSO-d_6) of **66**



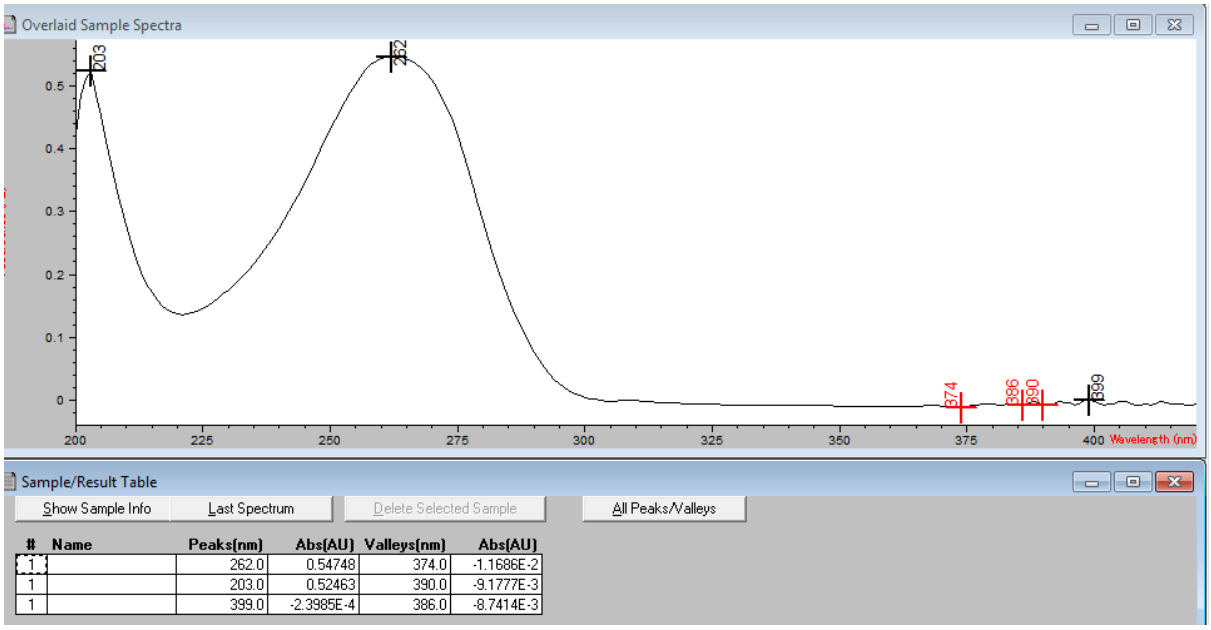
(+)-HRESIMS spectrum of **66**



(+)-HRESIMS/MS spectrum of **66**



IR spectrum of **66**



Sample/Result Table

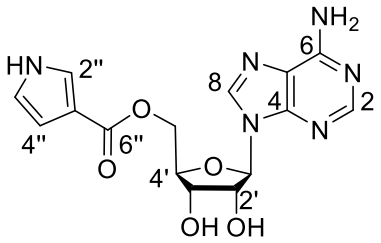
Show Sample Info | Last Spectrum | Delete Selected Sample | All Peaks/Valleys

#	Name	Peaks(nm)	Abs(AU)	Valleys(nm)	Abs(AU)
1		262.0	0.54748	374.0	-1.1686E-2
1		203.0	0.52463	390.0	-9.1777E-3
1		399.0	-2.3985E-4	386.0	-8.7414E-3

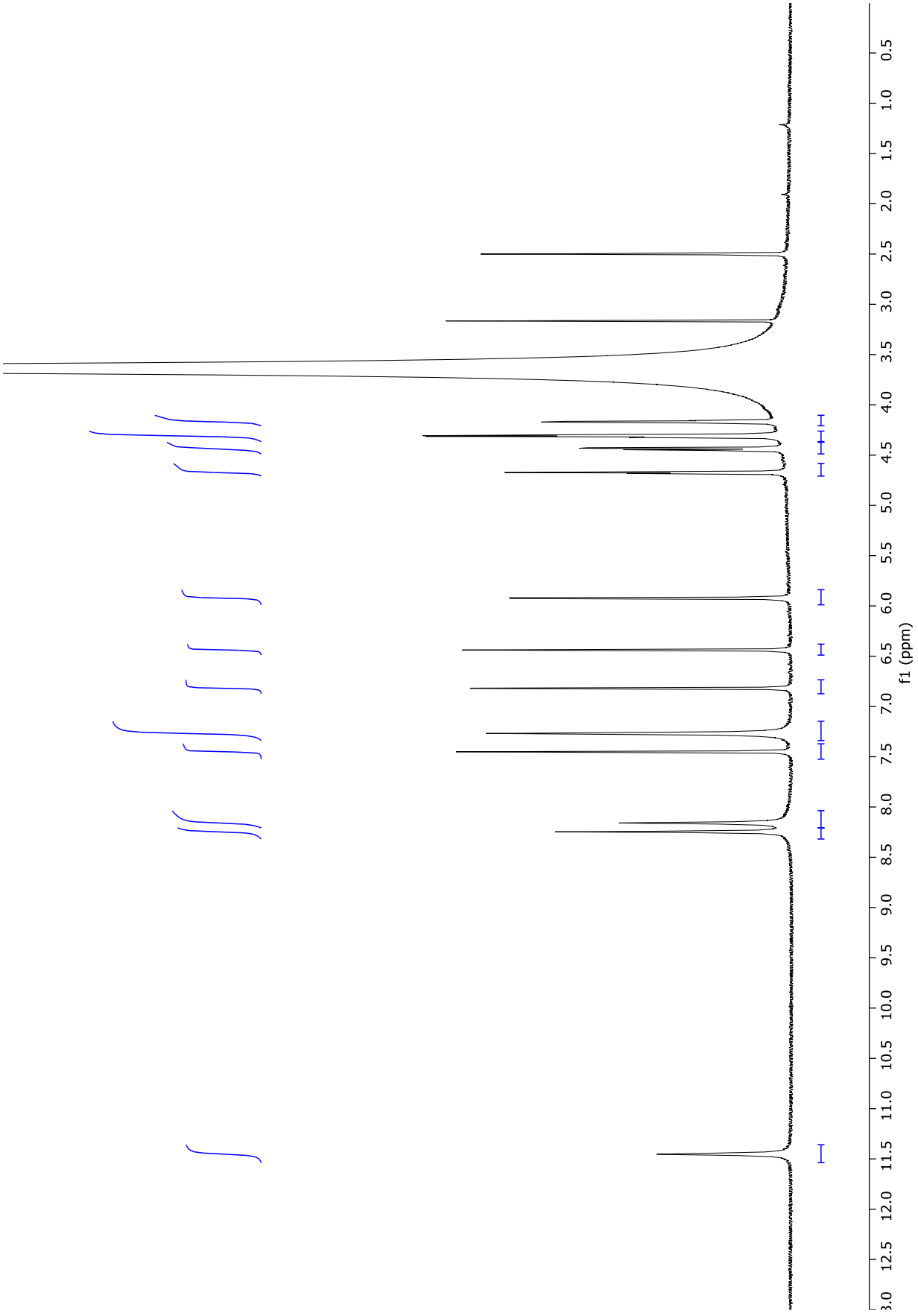
UV/vis spectrum of **66**

Table A2: ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for adenosyl-6'-pyrrole-3-carboxylate (**67**) in DMSO-d_6.

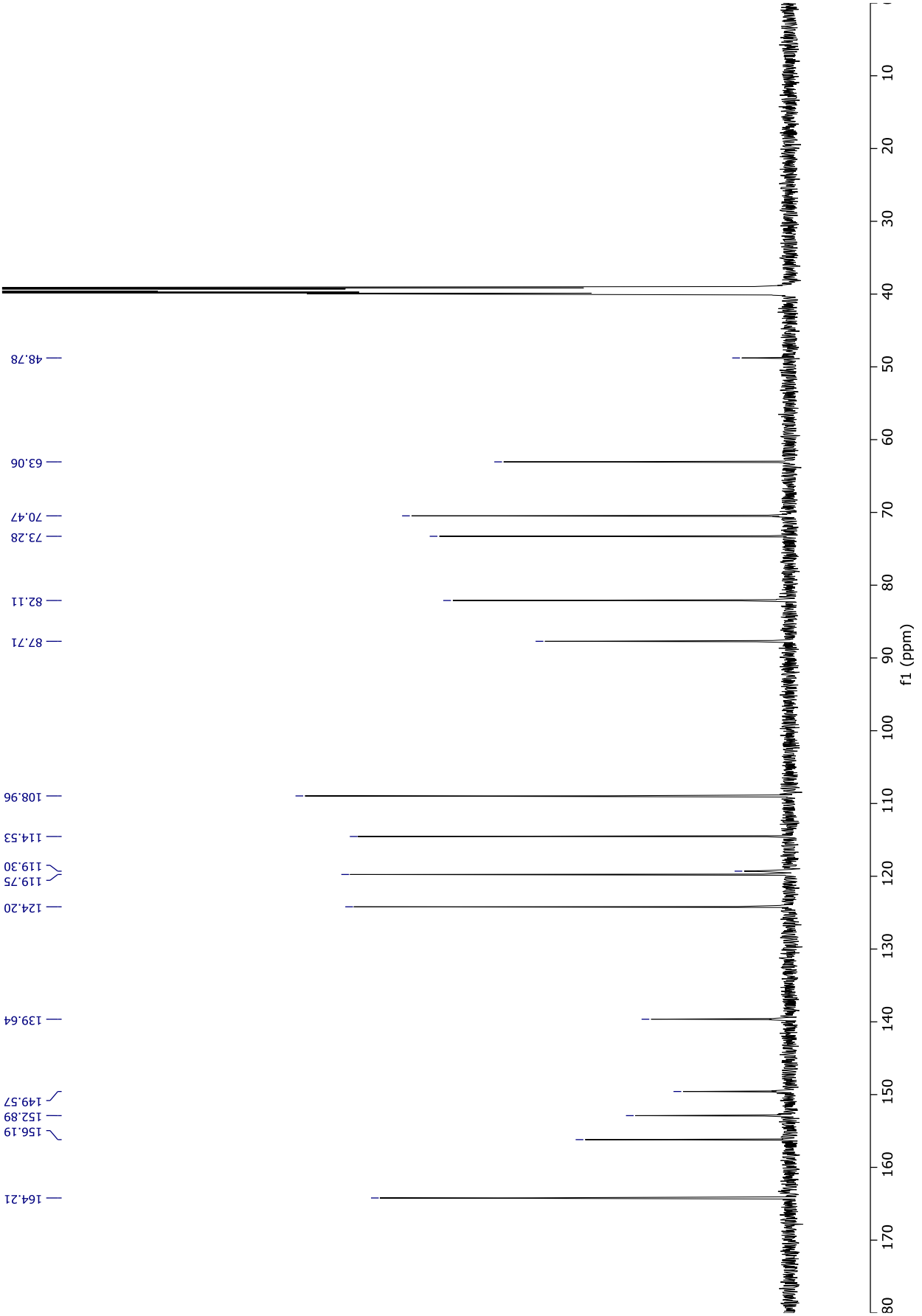
position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC
2	152.9	8.16 (s)		4, 6
4	149.6			
5	119.3			
6	156.2			
NH_2–6		7.27 (s)		5
8	139.6	8.24 (s)		4, 5
1'	87.7	5.92 (d, 5.1)	2'	4, 8, 2', 3'
2'	73.3	4.67 (t, 5.2)	1', 3'	1', 4'
3'	70.5	4.31 (m)	2', 4'	1', 4'
4'	82.1	4.17 (m)	3', 5'	3', 5'
5'	63.1	4.43 (dd, 12.1, 3.6)	4'	3', 4", 6"
		4.32 (m)	4'	3', 4", 6"
NH–1"		11.45 (s)	2", 4", 5"	
2"	124.2	7.45 (m)	NH–1", 4", 5"	3", 4", 5"
3"	114.5			
4"	109	6.44 (m)	NH–1", 2", 5"	2", 3", 5"
5"	119.7	6.82 (m)	NH–1", 2", 4"	2", 3", 4"
6"	164.2			



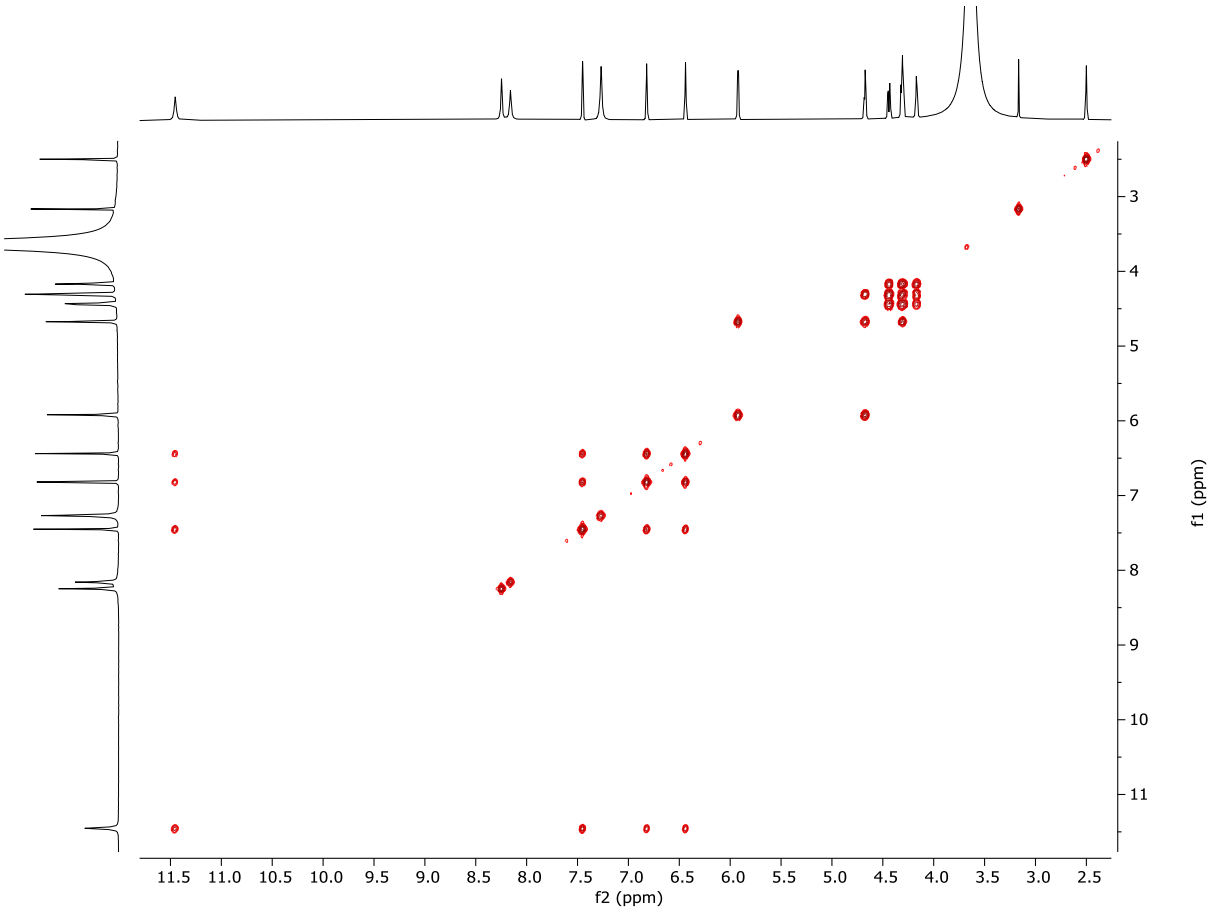
67



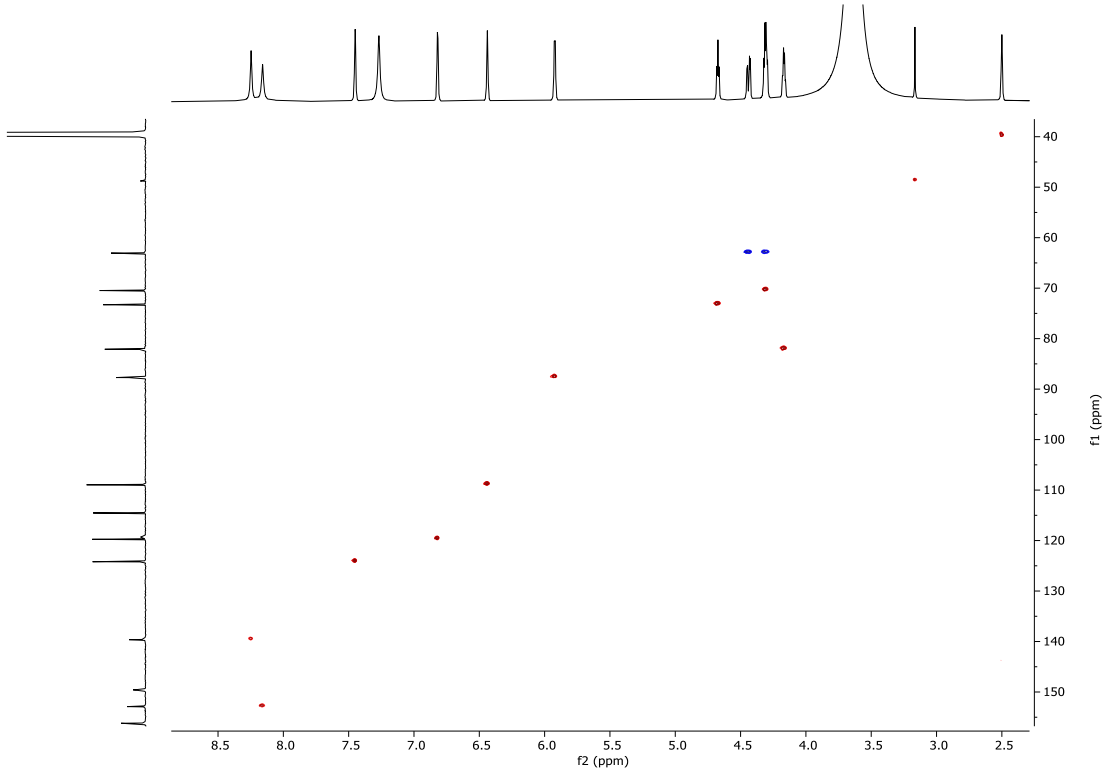
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **67**



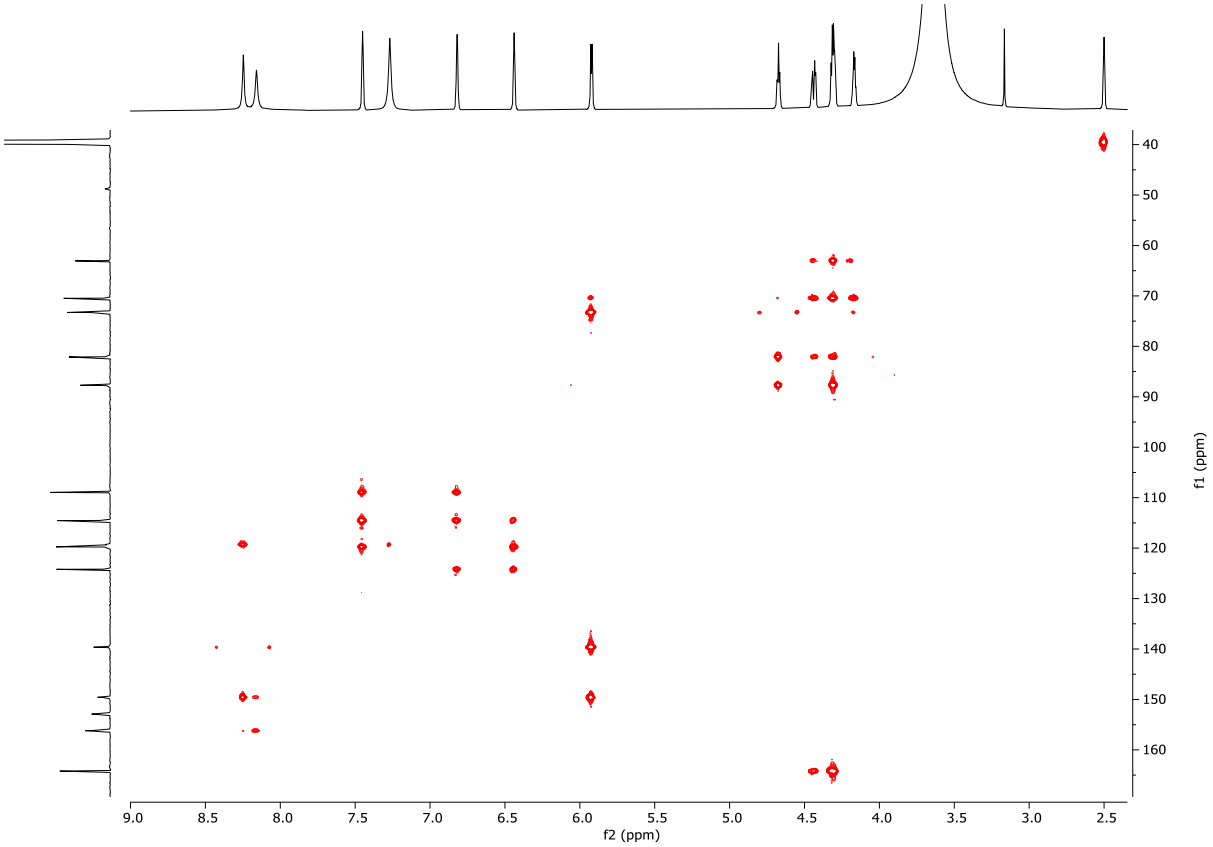
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **67**



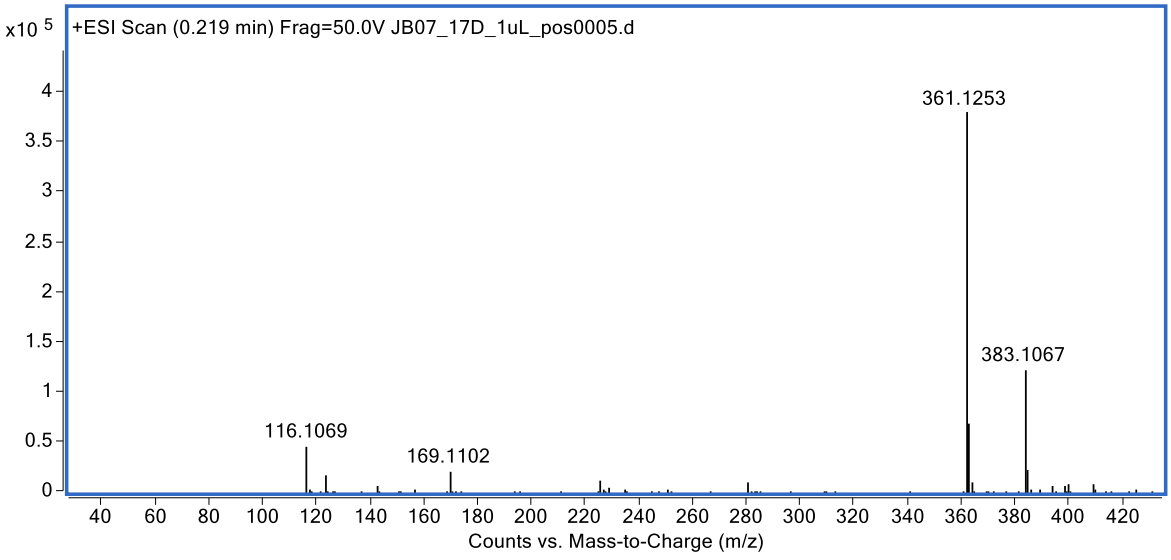
COSY Spectrum (600 MHz, DMSO-d_6) of **67**



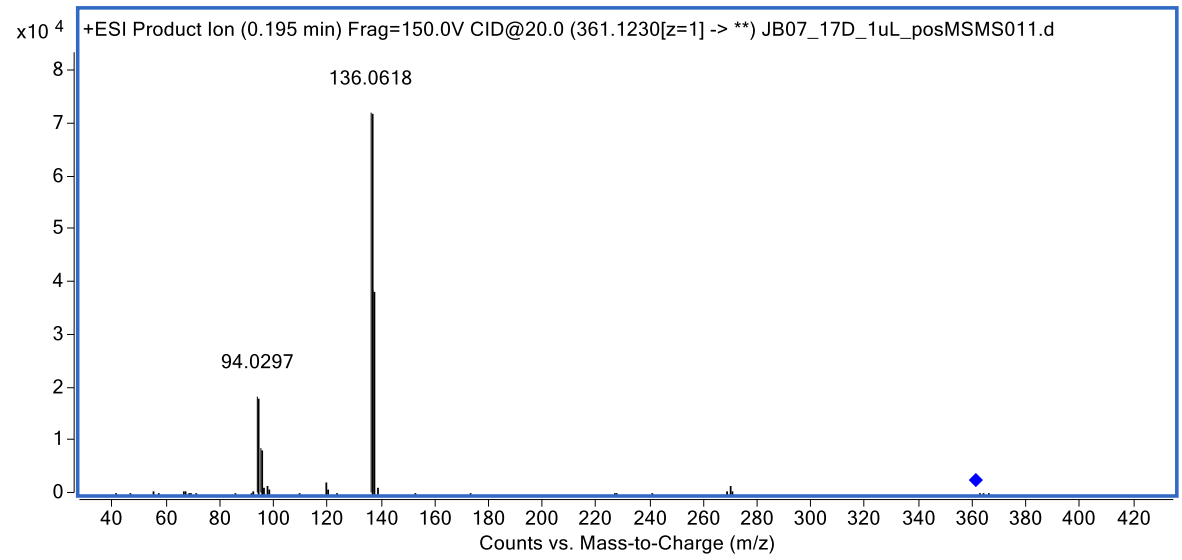
HSQC Spectrum (600 MHz, DMSO-d_6) of **67**



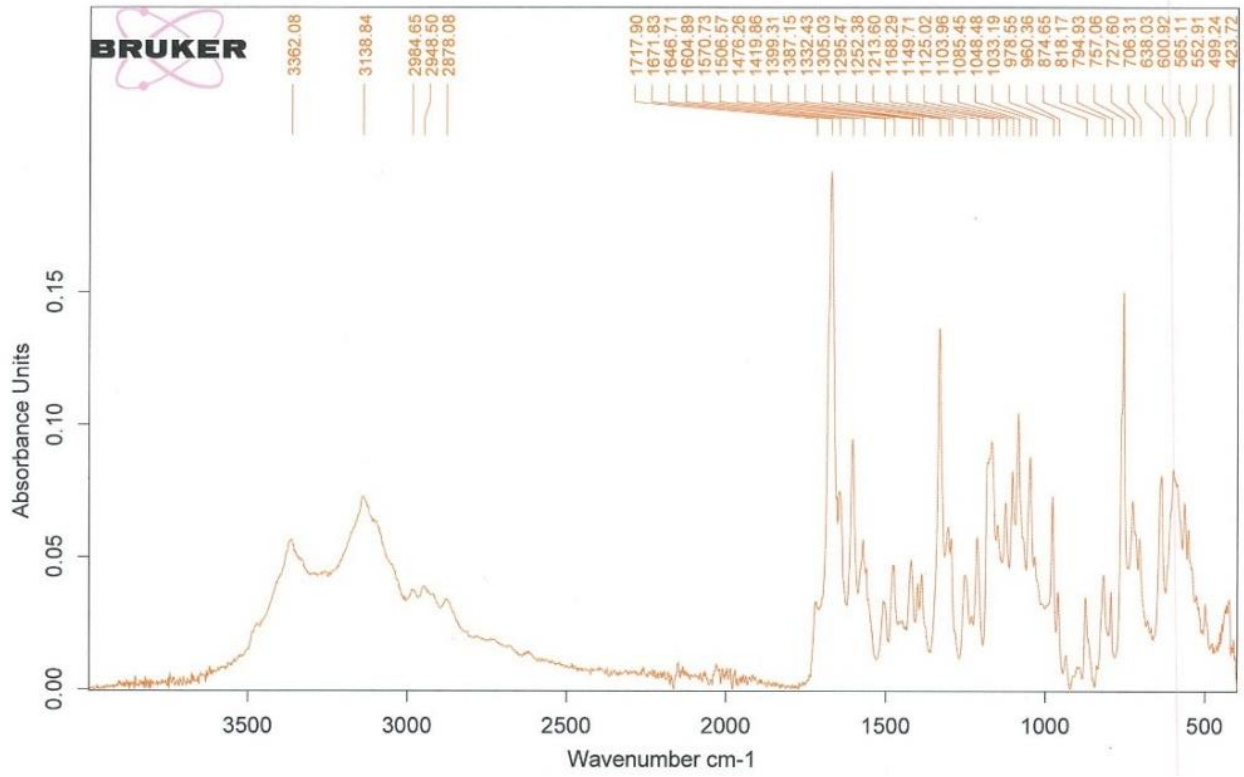
HMBC Spectrum (600 MHz, DMSO-d_6) of **67**



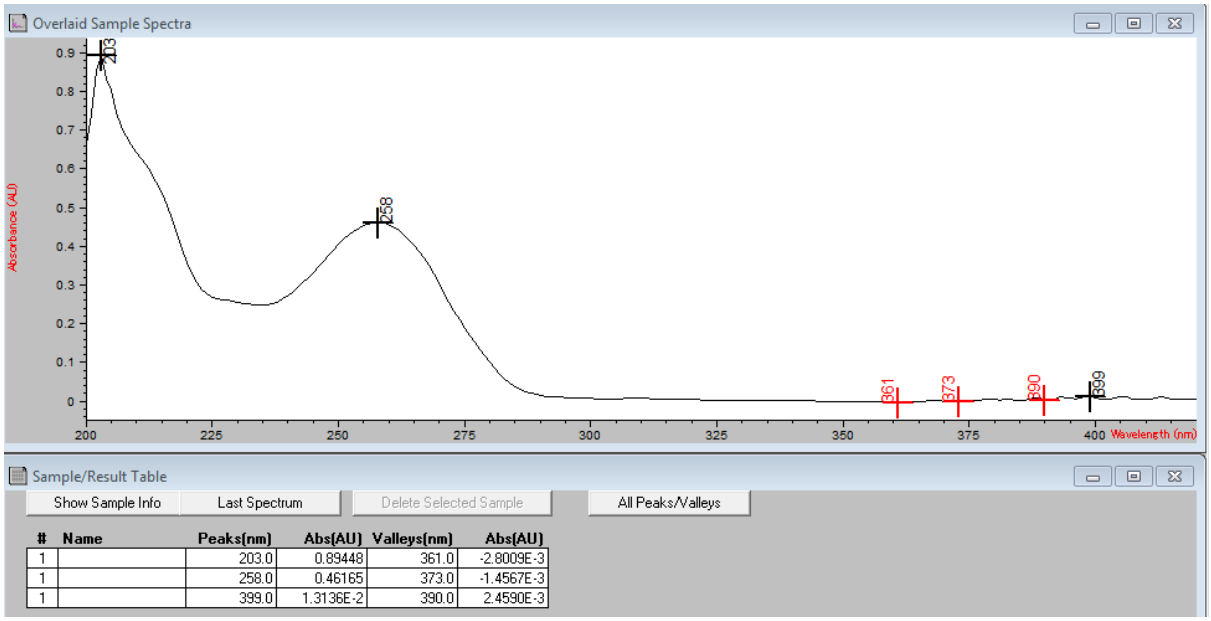
(+)-HRESIMS spectrum of **67**



(+)-HRESIMS/MS spectrum of **67**



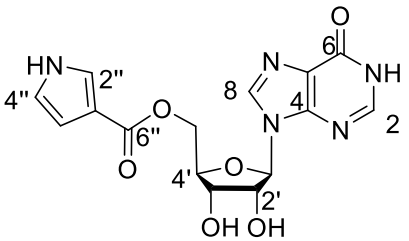
IR spectrum of **67**



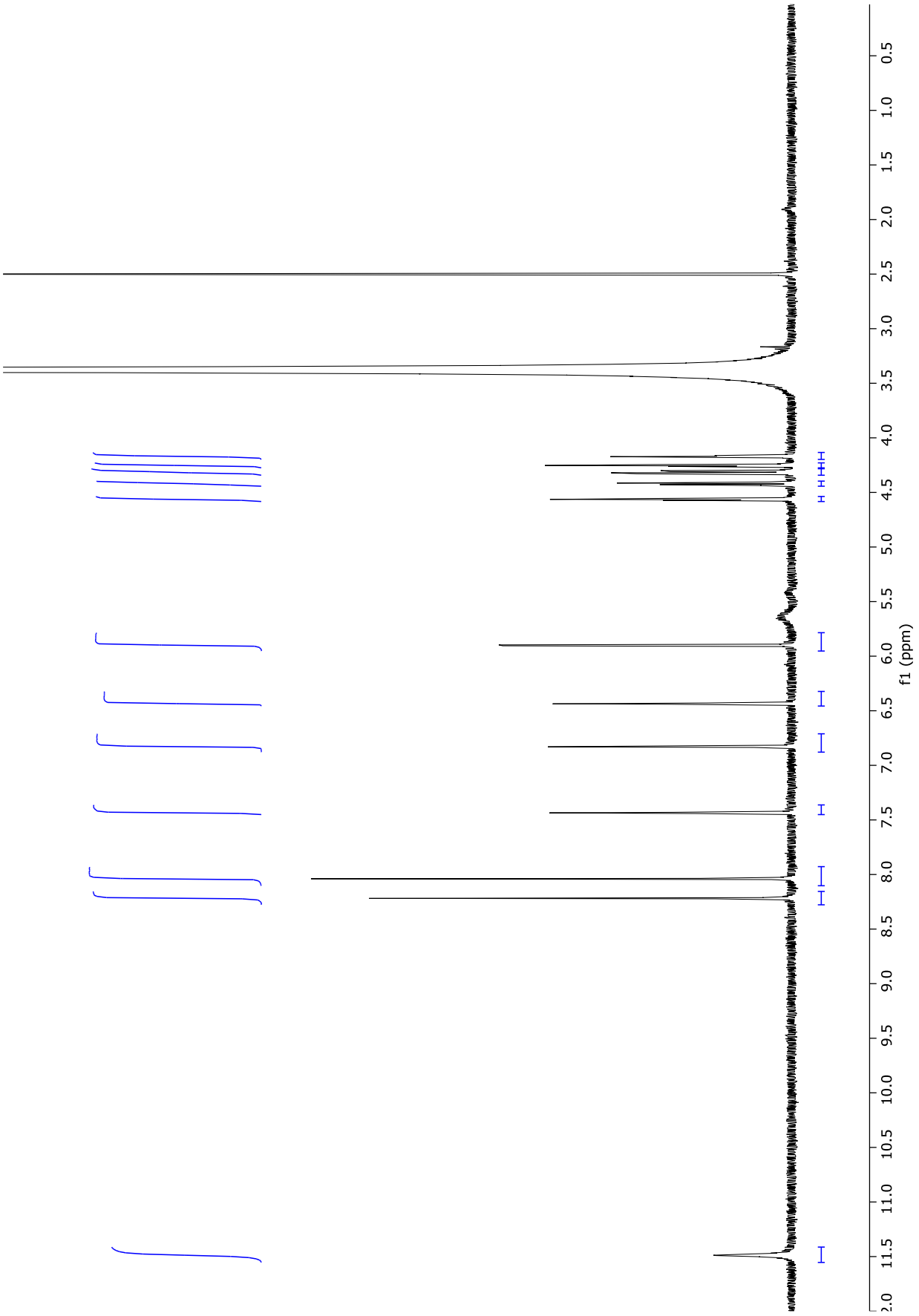
UV/vis spectrum of **67**

Table A3: ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for inosyl-6'-pyrrole-3-carboxylate (**68**) in DMSO-d_6.

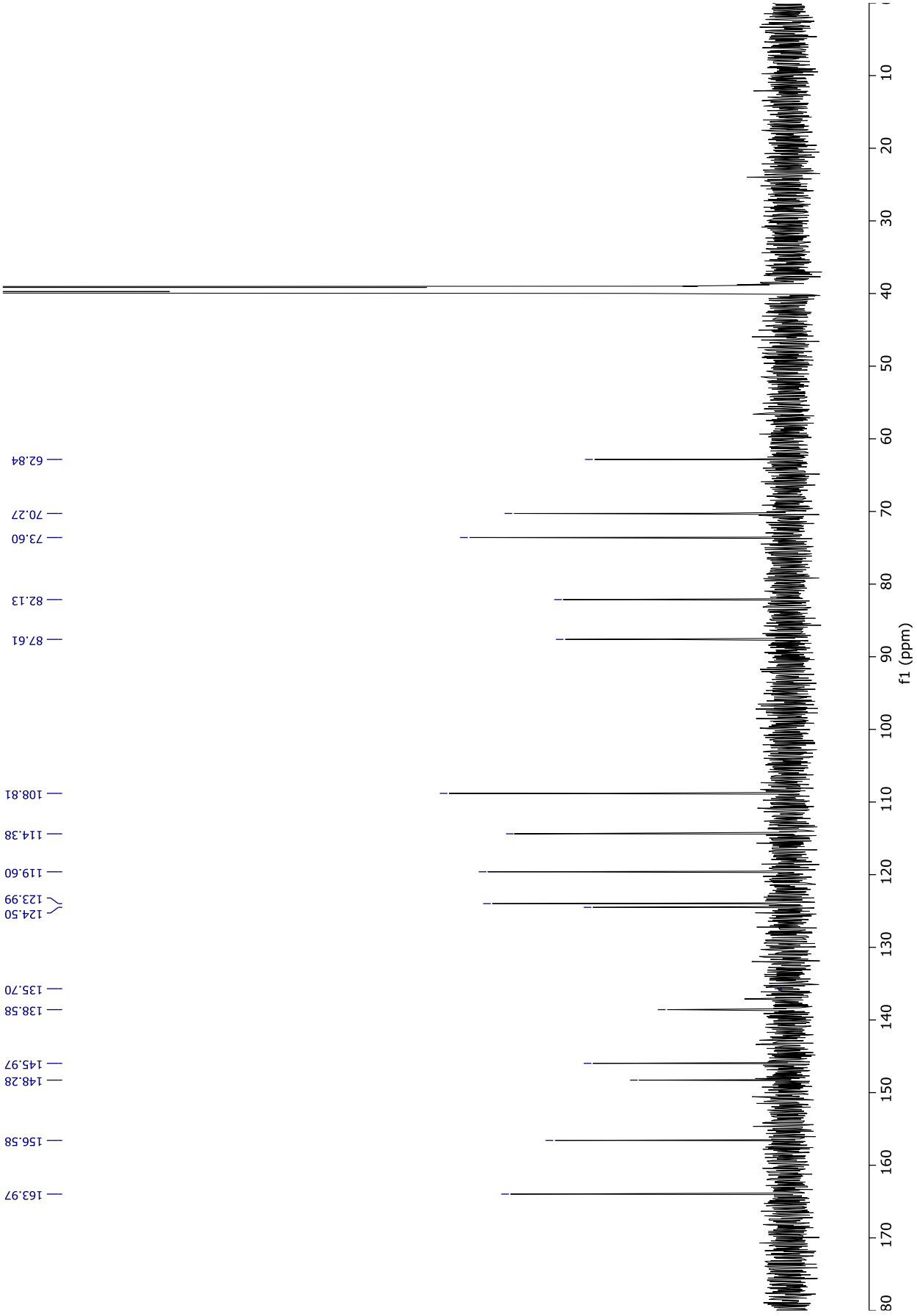
position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC
2	146	8.04 (s)		4, 6
4	148.3			
5	124.5			
6	156.6			
8	138.6	8.22 (s)		4, 5
1'	87.6	5.90 (d, 5.1)	2'	4, 8, 2'
2'	73.6	4.56 (t, 5.1)	1', 3'	1', 4'
3'	70.3	4.25 (t, 5.1)	2', 4'	1', 5'
4'	82.1	4.17 (m)	3', 5'	3'
5'	62.8	4.41 (dd, 12.1, 3.7)	4'	3', 6"
		4.32 (dd, 12.1, 5.4)	4'	3', 4', 6"
NH–1"		11.49 (s)	2", 4", 5"	
2"	124.0	7.43 (m)	NH–1", 4", 5"	3", 4", 5"
3"	114.4			
4"	108.8	6.43 (m)	NH–1", 2", 5"	2", 3", 5"
5"	119.6	6.83 (m)	NH–1", 2", 4"	
6"	164.0			



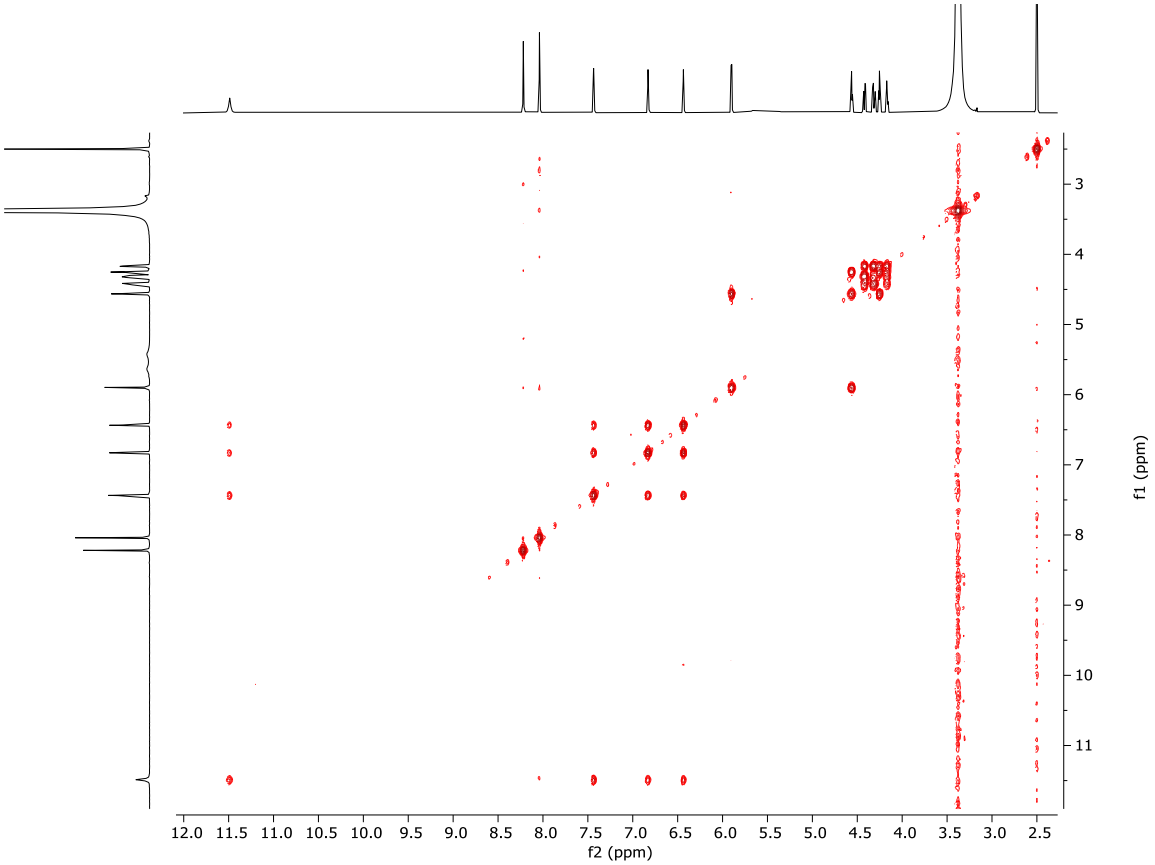
68



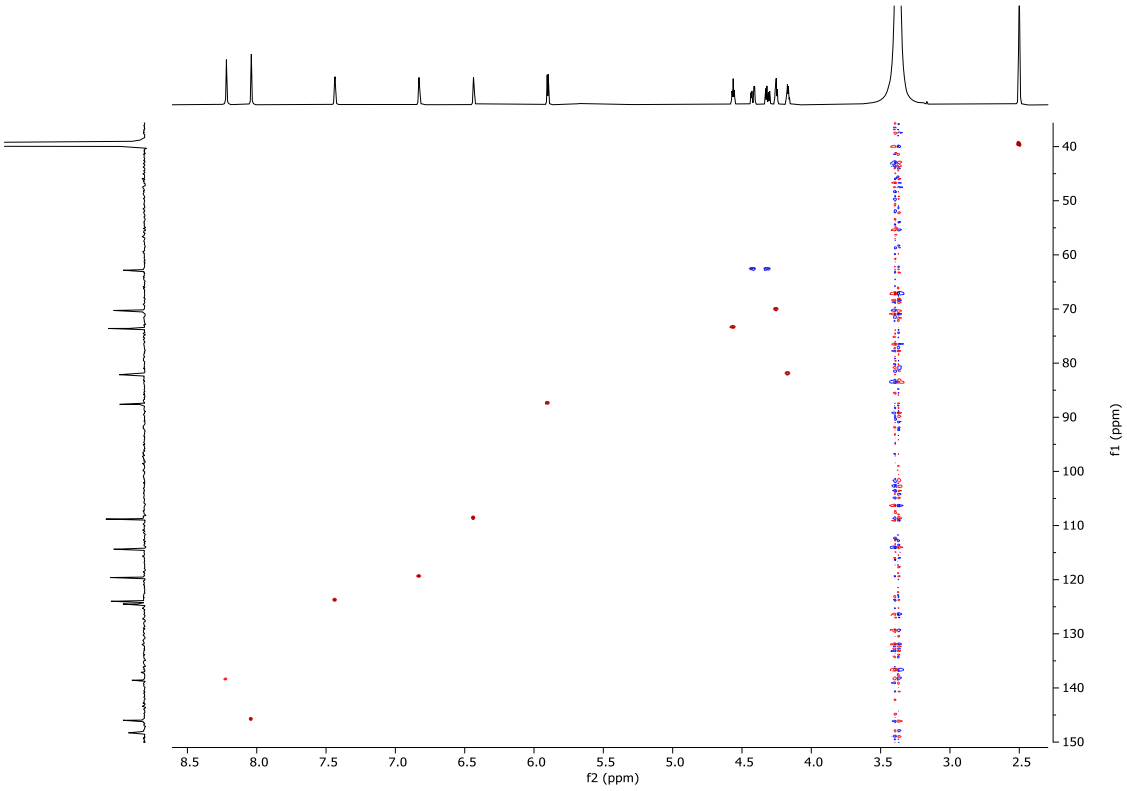
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **68**



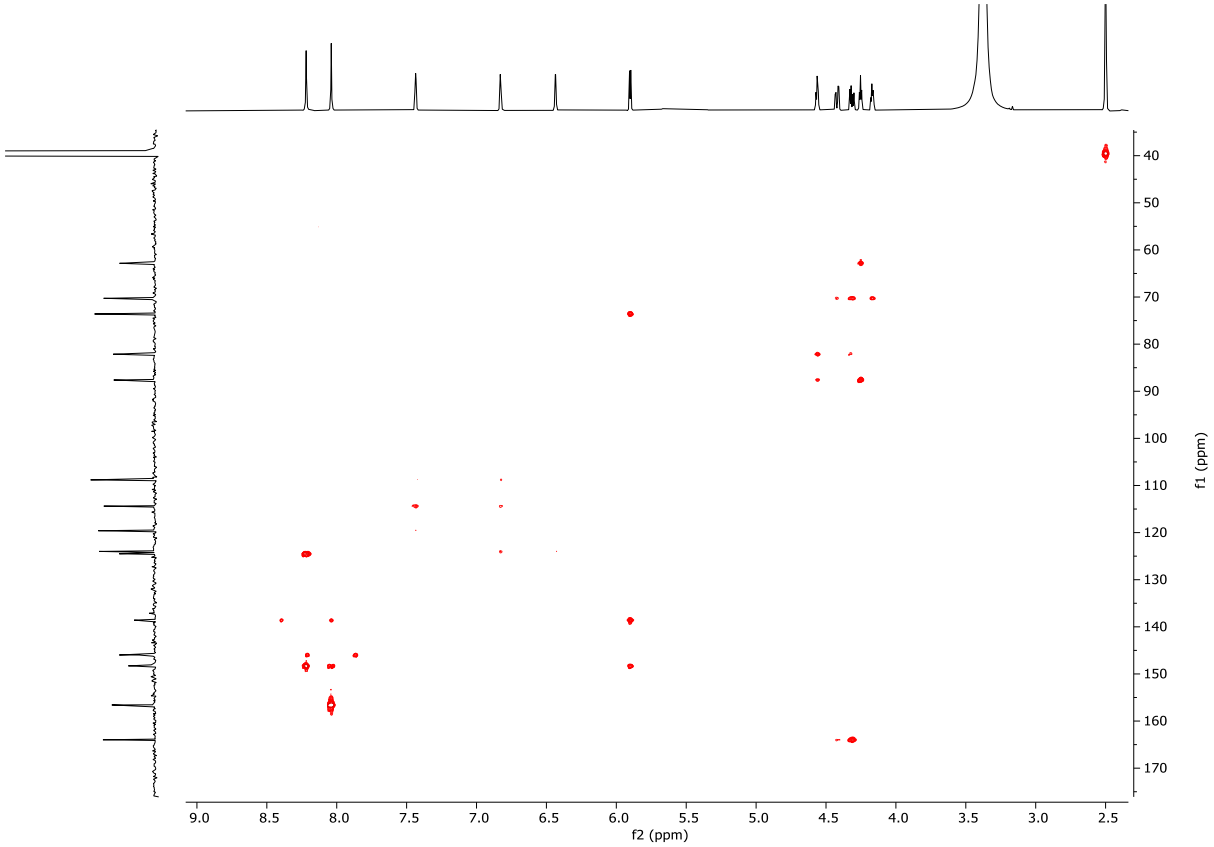
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **68**



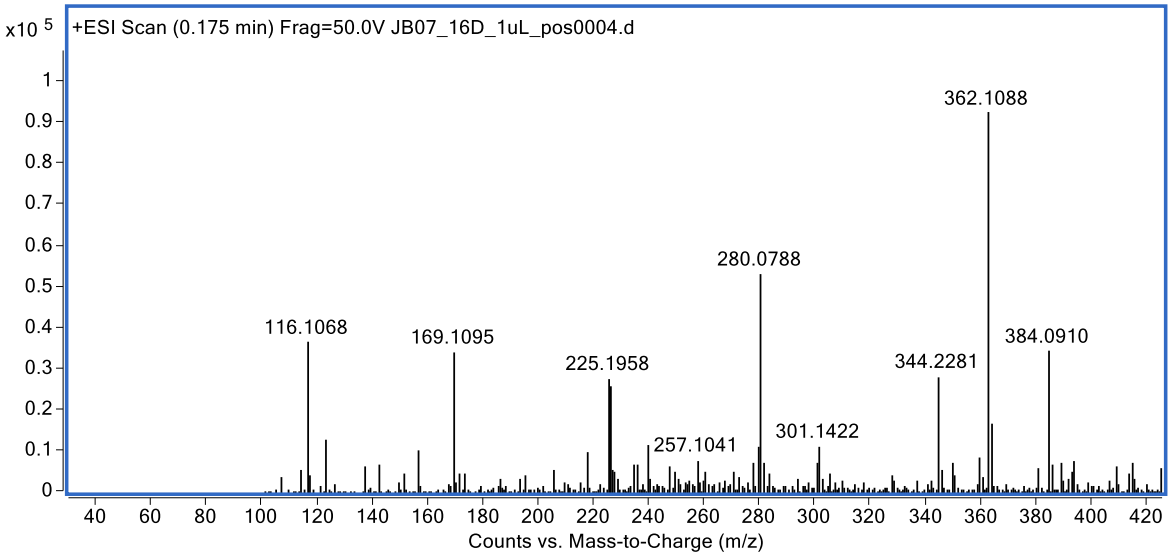
COSY Spectrum (600 MHz, DMSO-d_6) of **68**



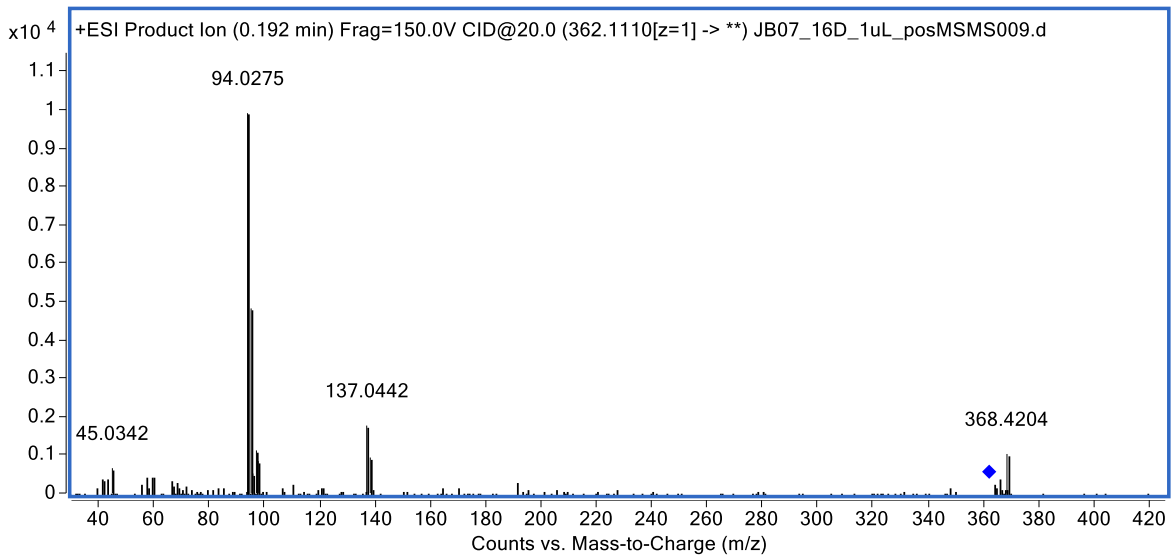
HSQC Spectrum (600 MHz, DMSO-d_6) of **68**



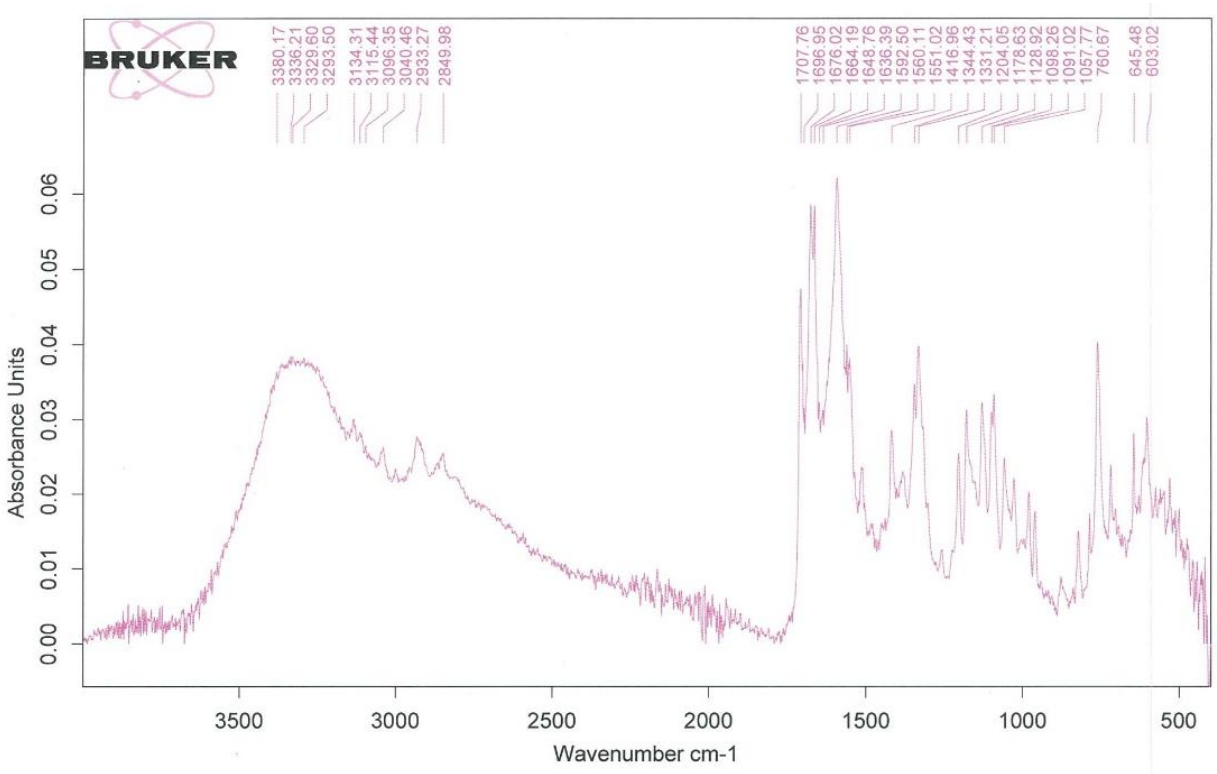
HMBC Spectrum (600 MHz, DMSO-d_6) of **68**



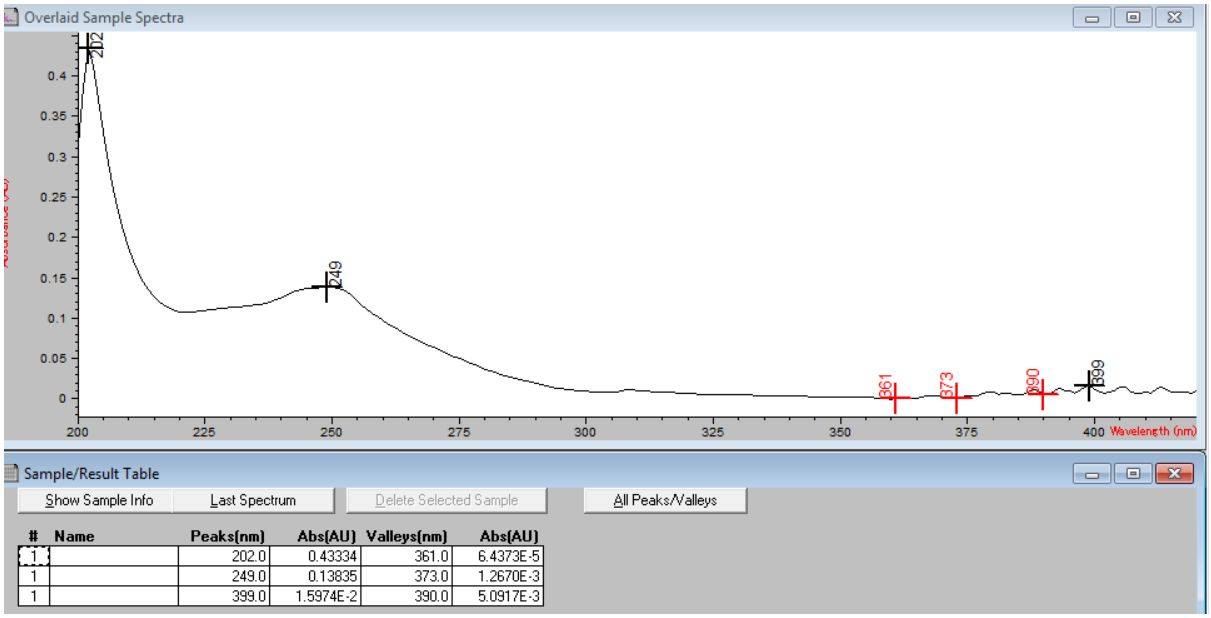
(+)-HRESIMS spectrum of **68**



(+)-HRESIMS/MS spectrum **68**



IR spectrum of **68**

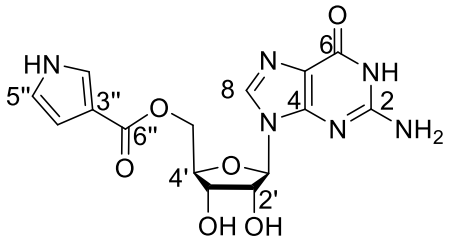


#	Name	Peaks(nm)	Abs(AU)	Valleys(nm)	Abs(AU)
1		202.0	0.43334	361.0	6.4373E-5
1		249.0	0.13835	373.0	1.2670E-3
1		399.0	1.5974E-2	390.0	5.0917E-3

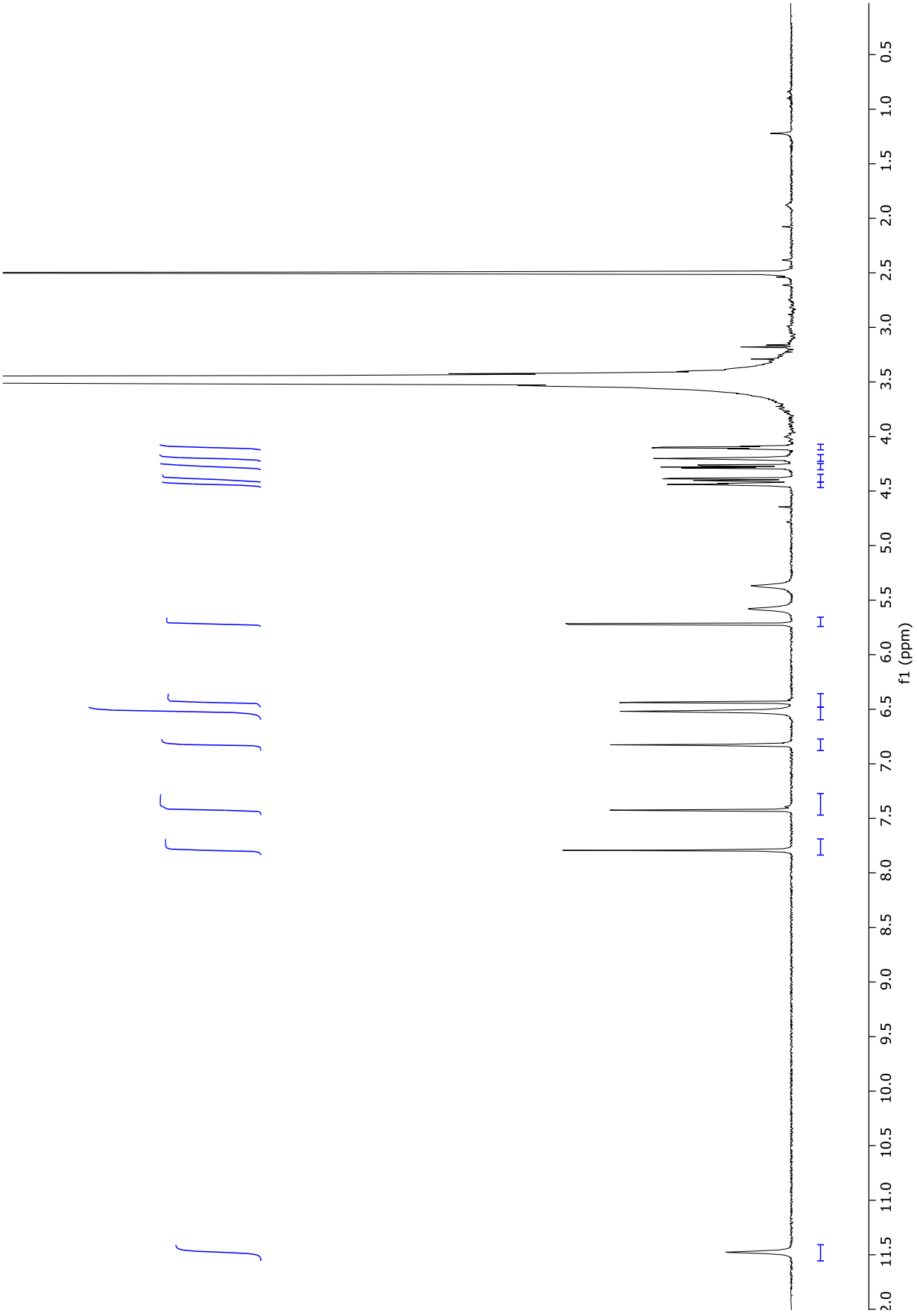
UV/vis spectrum of **68**

Table A4: ^{13}C (150 MHz) and ^{1}H (600 MHz) NMR data for guanosyl-6'-pyrrole-3-carboxylate (**69**) in DMSO-d_6.

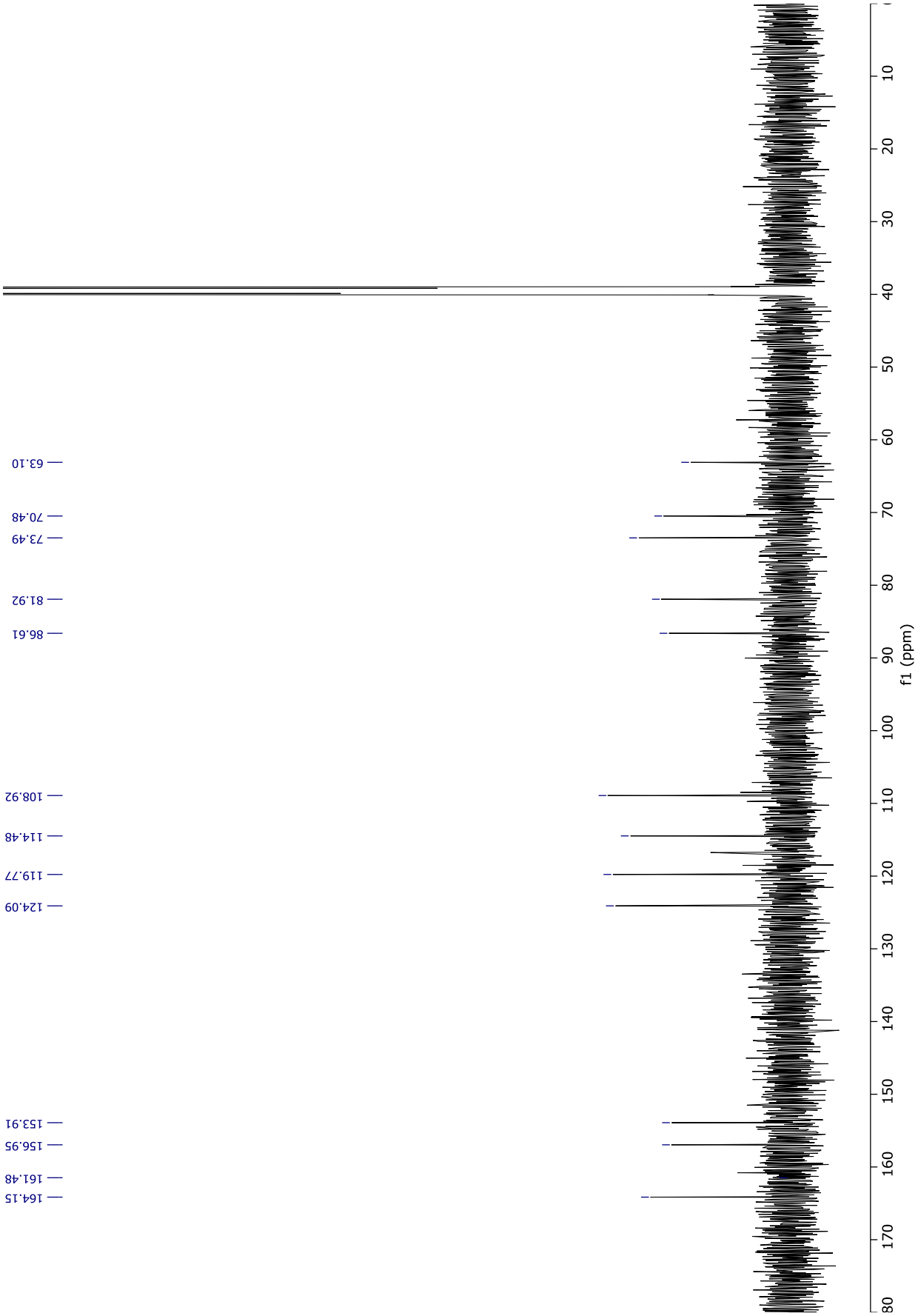
position	^{13}C (δ)	^{1}H (δ, mult., J in Hz)	COSY	HMBC
2	153.9			
NH_2–2		6.52 (s)		
4	151.5			
5	116.8			
6	156.9			
8	135.6	7.77 (s)		4, 5
1'	86.6	5.7 (d, 5.4)	2'	4, 8, 2'
2'	73.5	4.42 (t, 5.3)	1', 3'	
3'	70.5	4.18 (t, 4.8)	2', 4'	1'
4'	81.9	4.08 (m)	3', 5'	
5'	63.1	4.37 (dd, 12.0, 3.7)	4'	3', 6"
		4.26 (dd, 12.0, 5.3)	4'	4', 6"
NH–1"		11.46 (s)	2", 4", 5"	
2"	124.1	7.4 (m)	NH–1", 4", 5"	
3"	114.5	–		
4"	108.9	6.42 (m)	NH–1", 2", 5"	
5"	119.8	6.81 (m)	NH–1", 2", 4"	
6"	164.1			



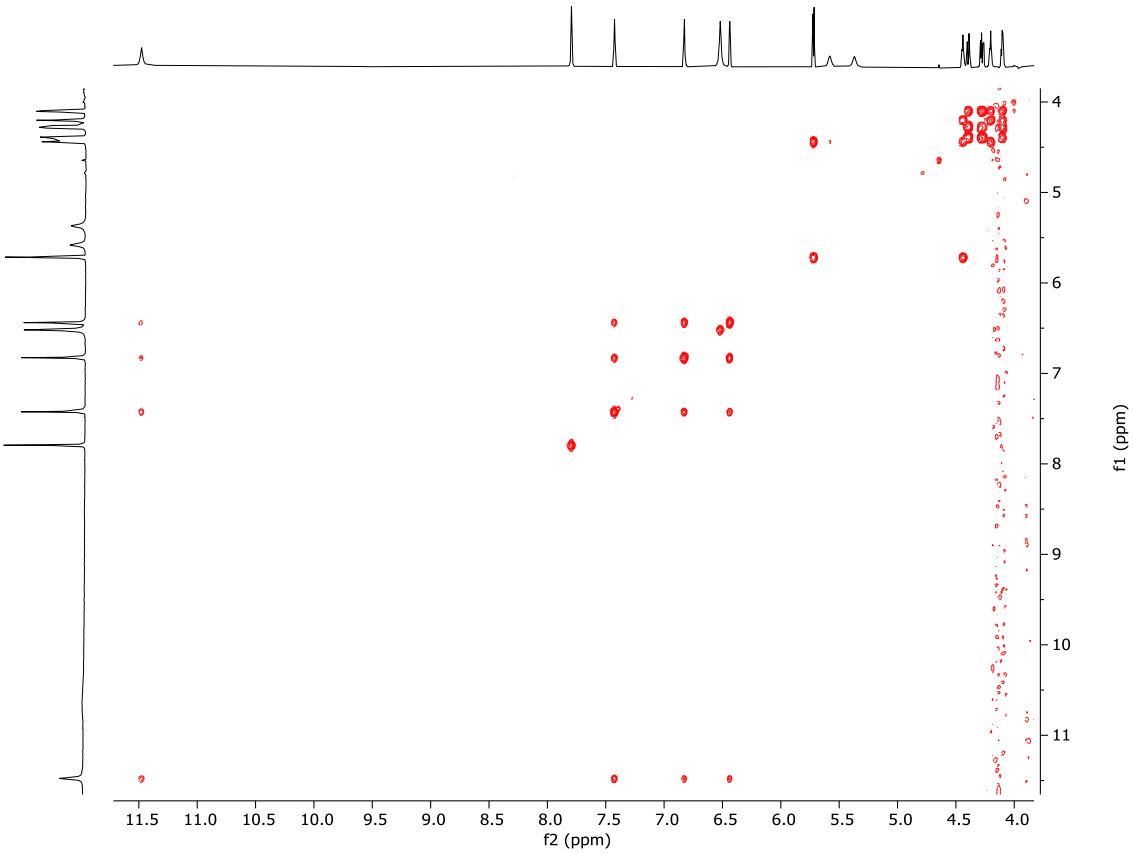
69



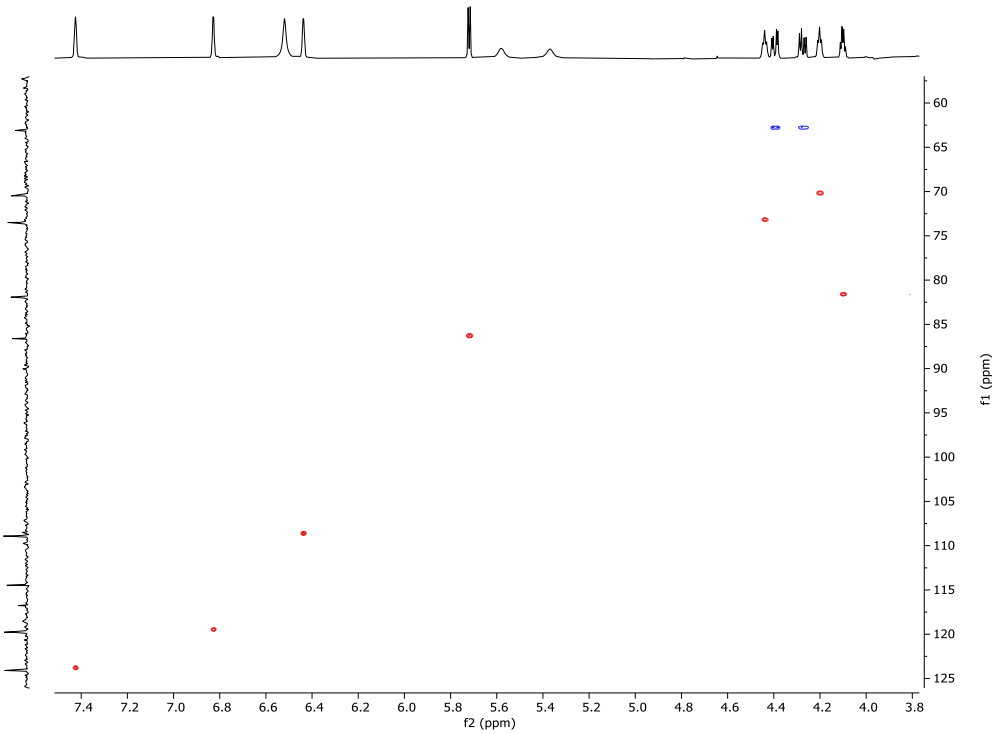
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **69**



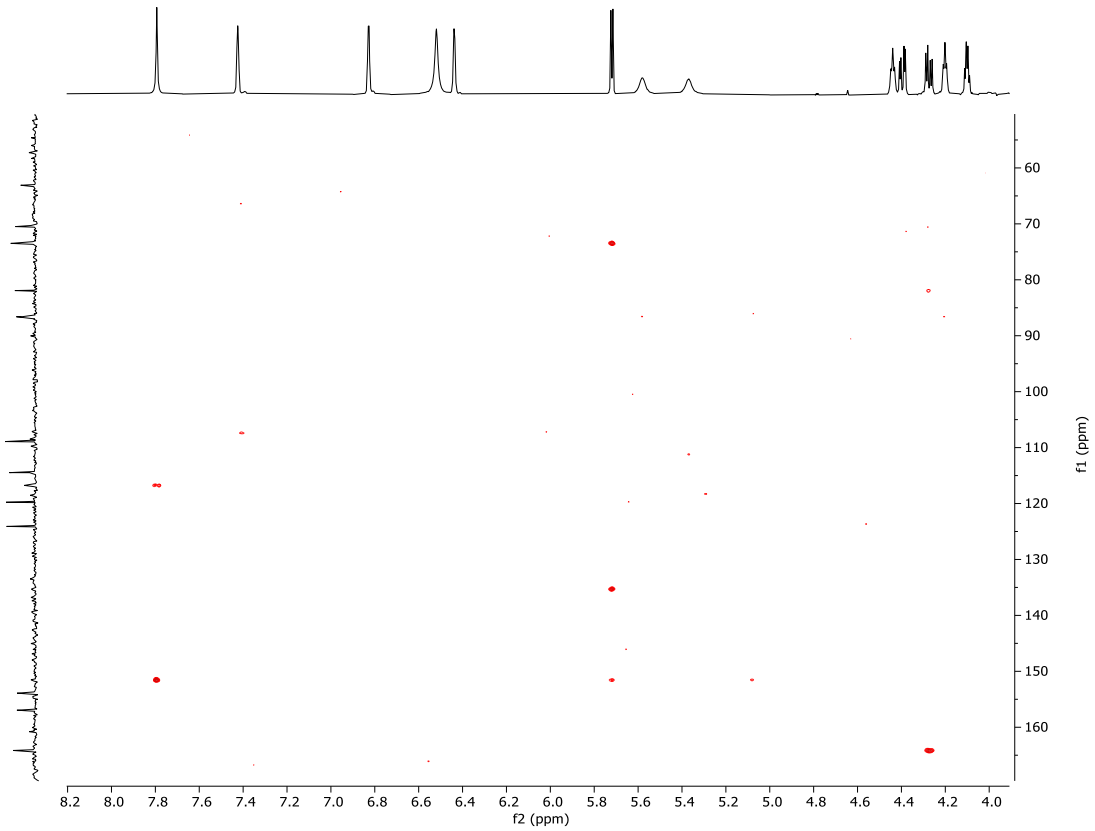
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **69**



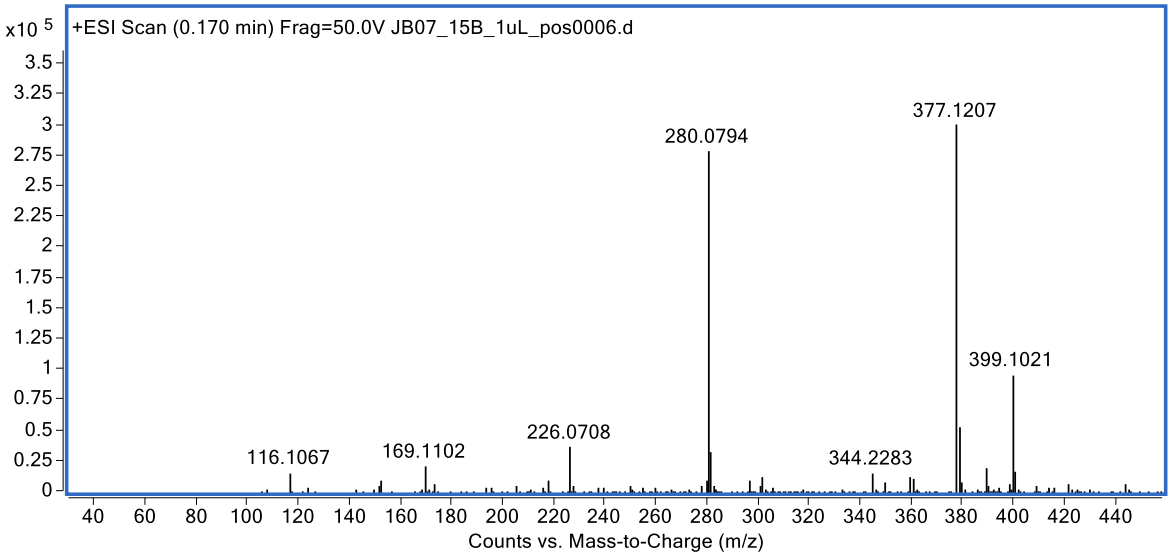
COSY Spectrum (600 MHz, DMSO-d_6) of **69**



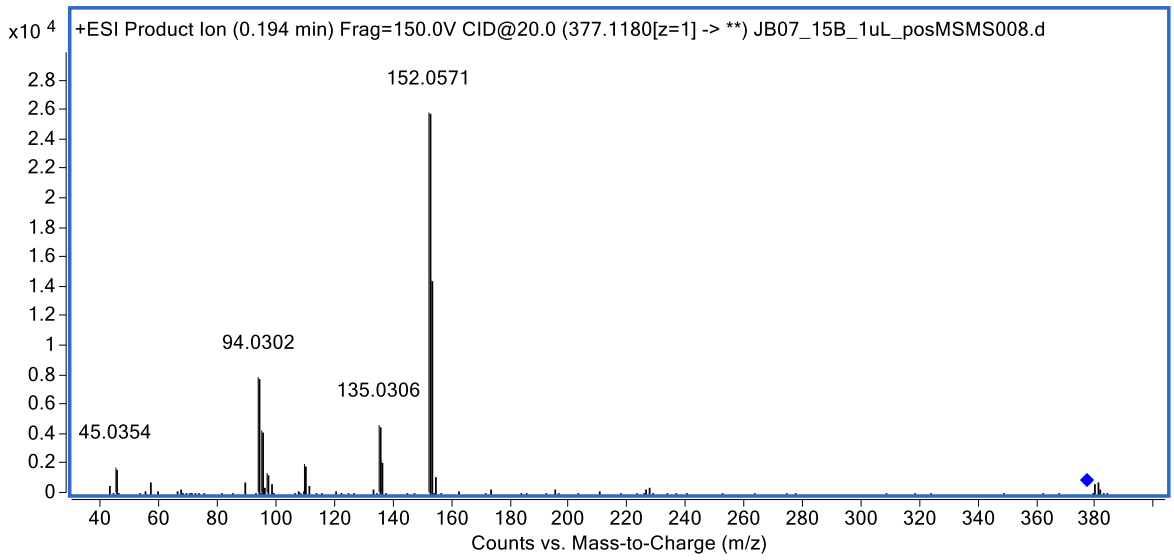
HSQC Spectrum (600 MHz, DMSO-d_6) of **69**



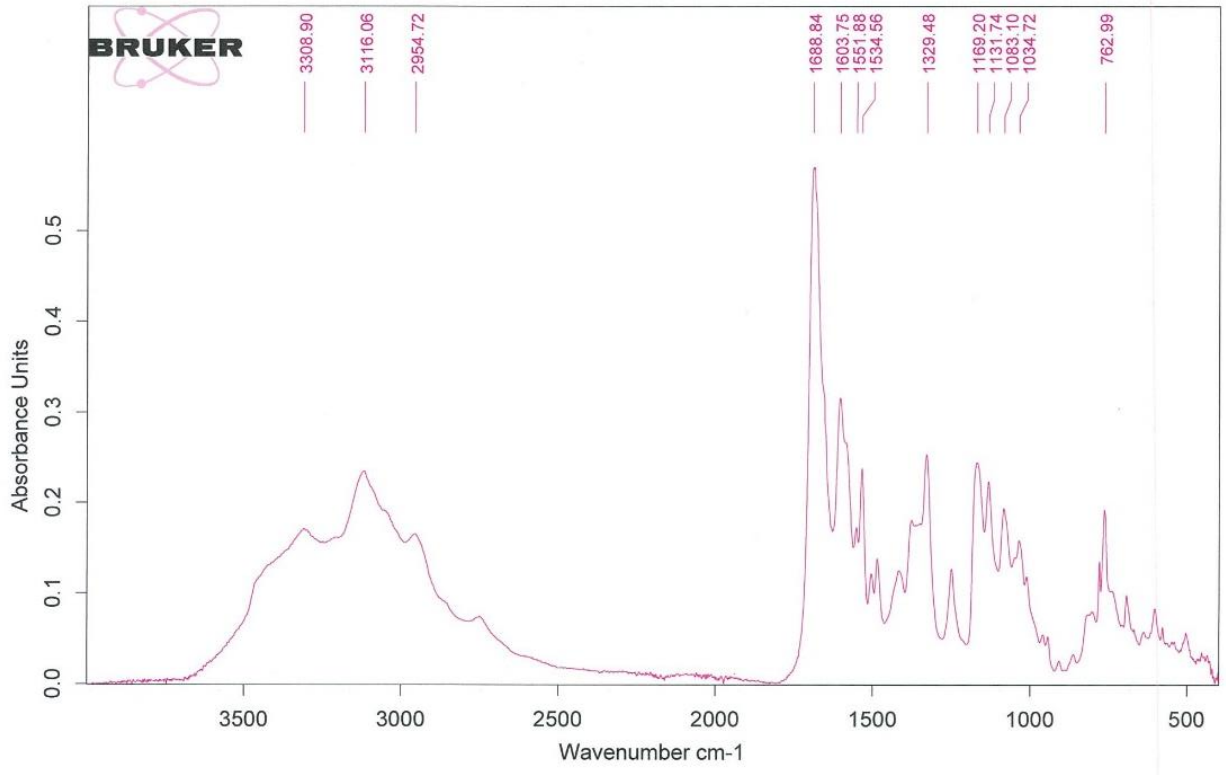
HMBC Spectrum (600 MHz, DMSO-d_6) of **69**



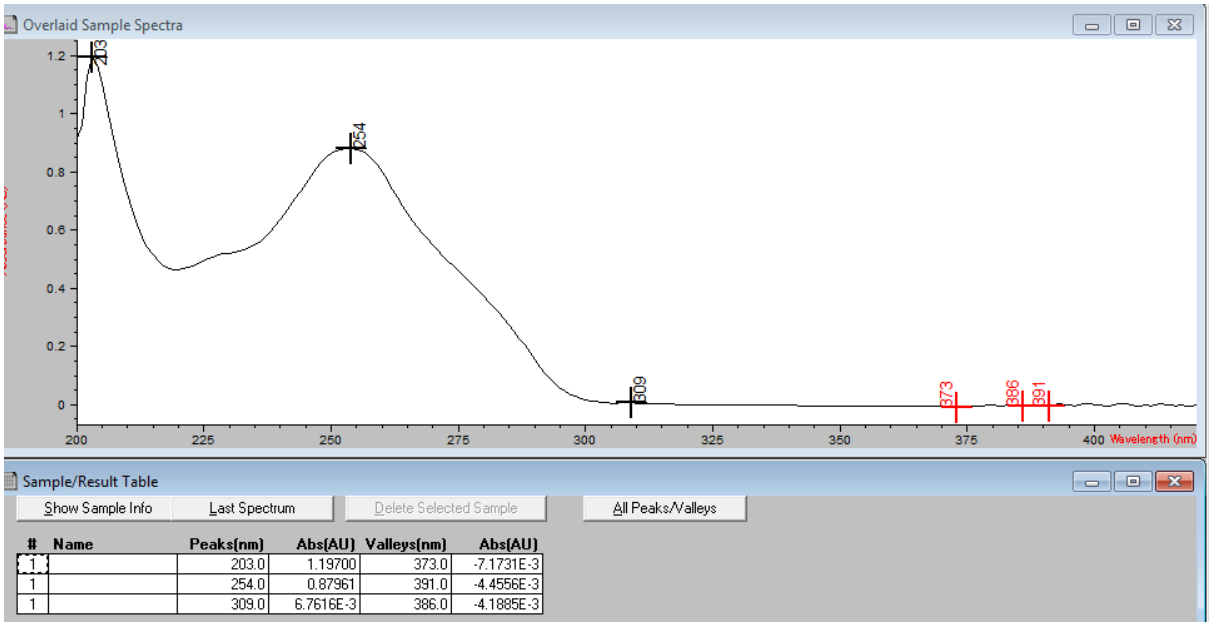
(+)-HRESIMS spectrum of **69**



(+)-HRESIMS/MS spectrum of **69**

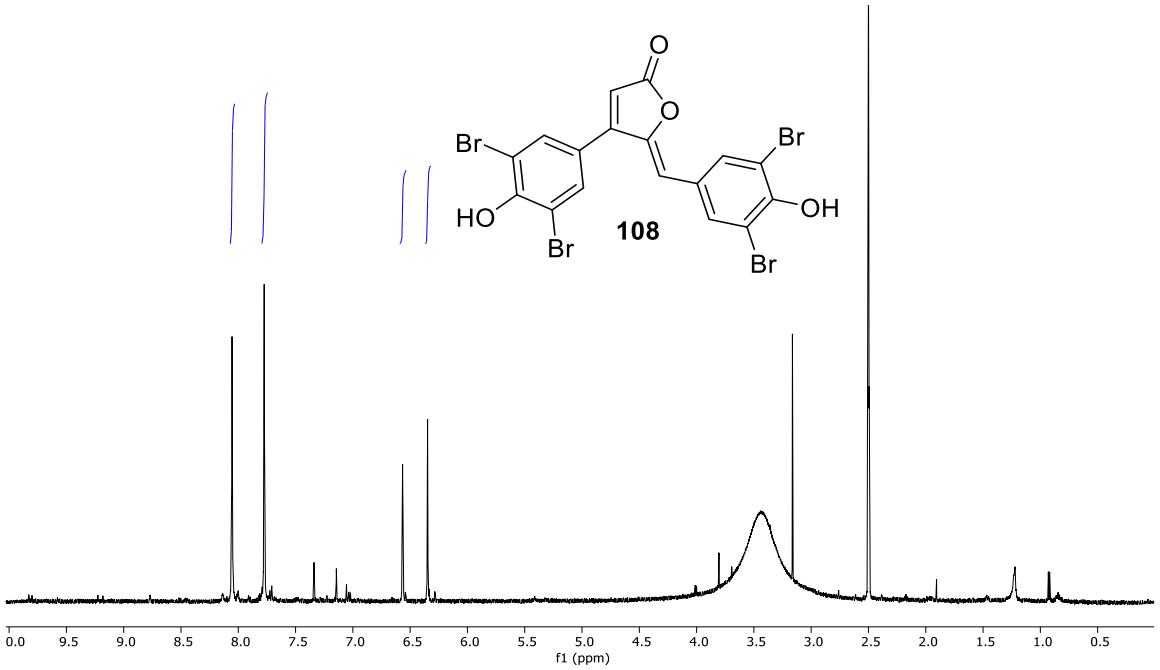


IR spectrum of **69**

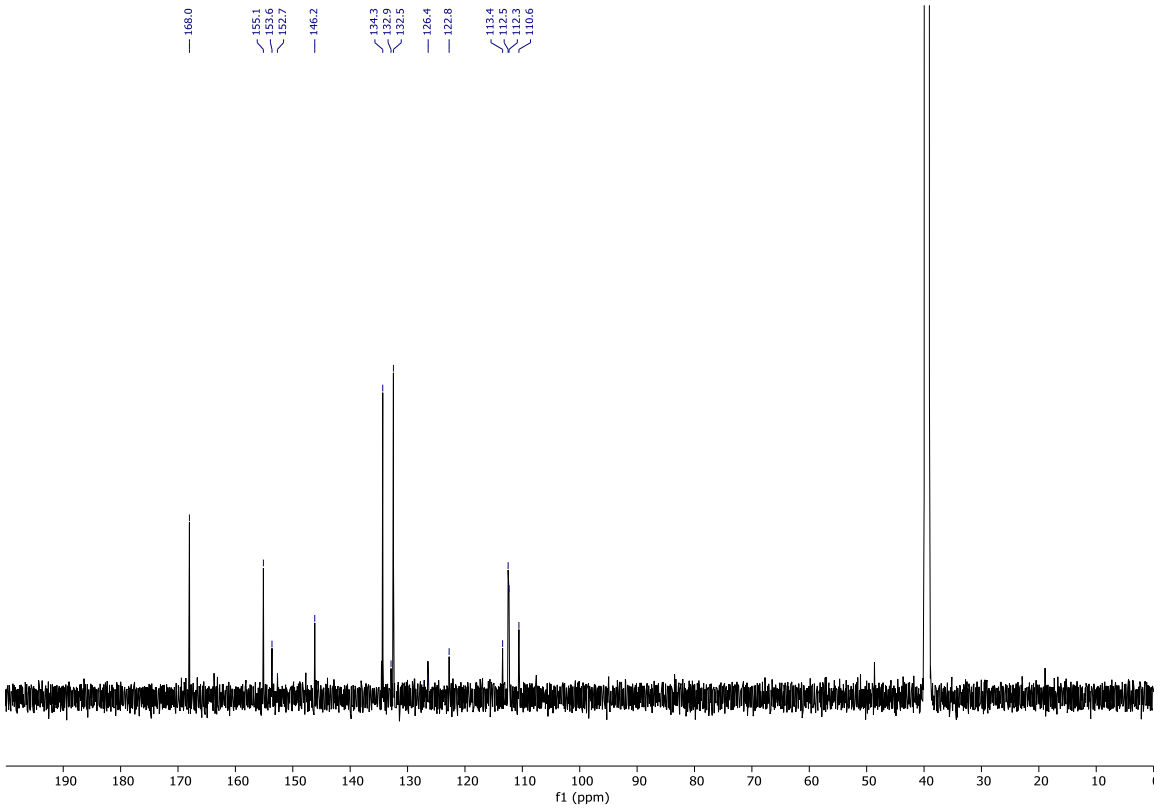


UV/vis spectrum of **69**

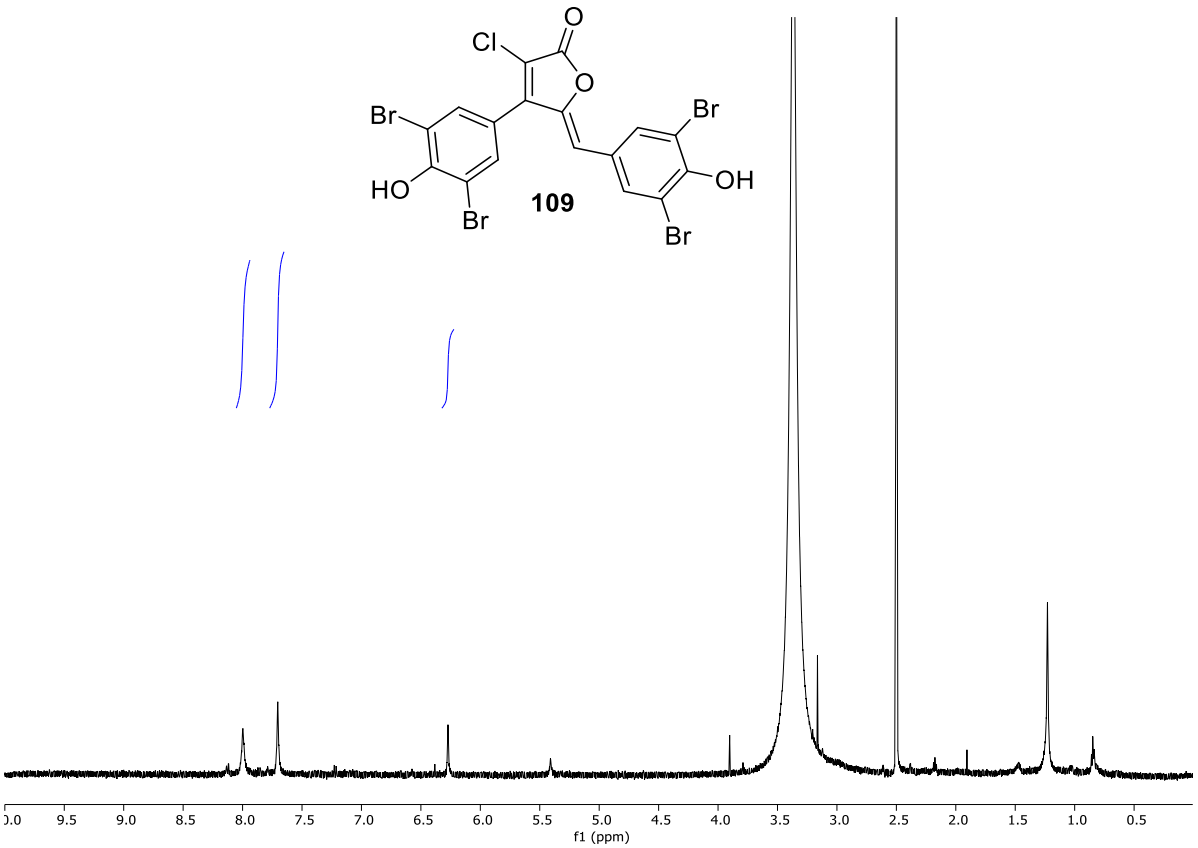
Appendix 7 – *Synoicum kuranui* Spectra



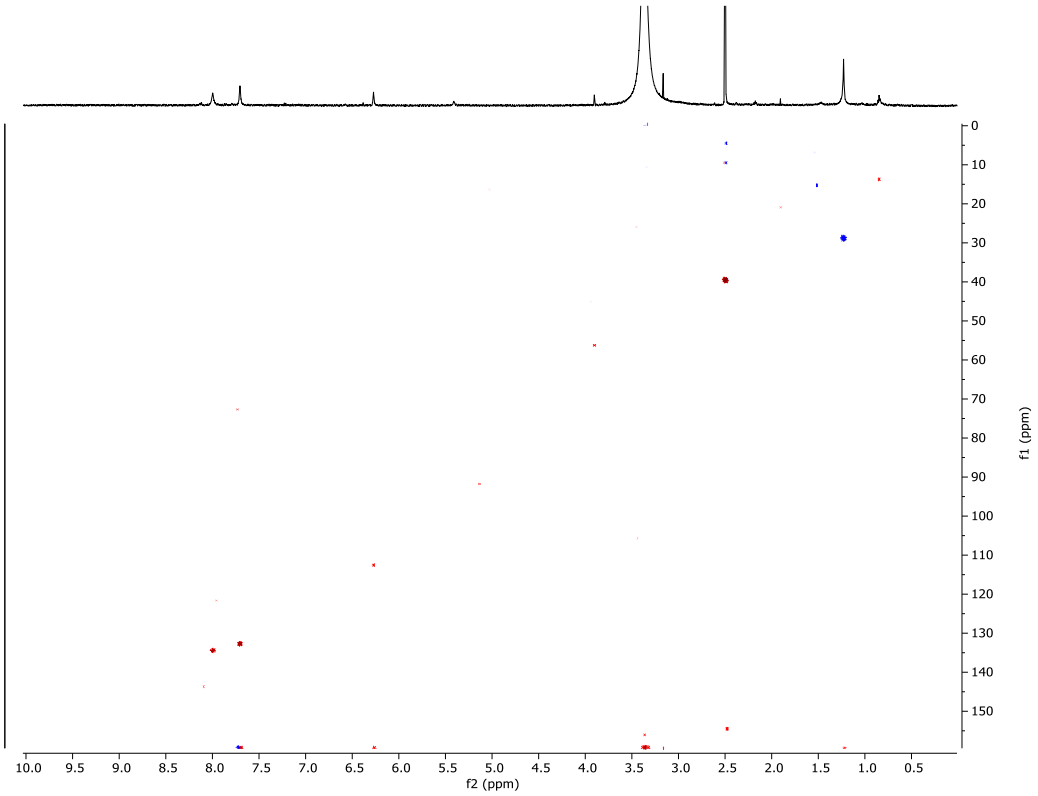
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **108**



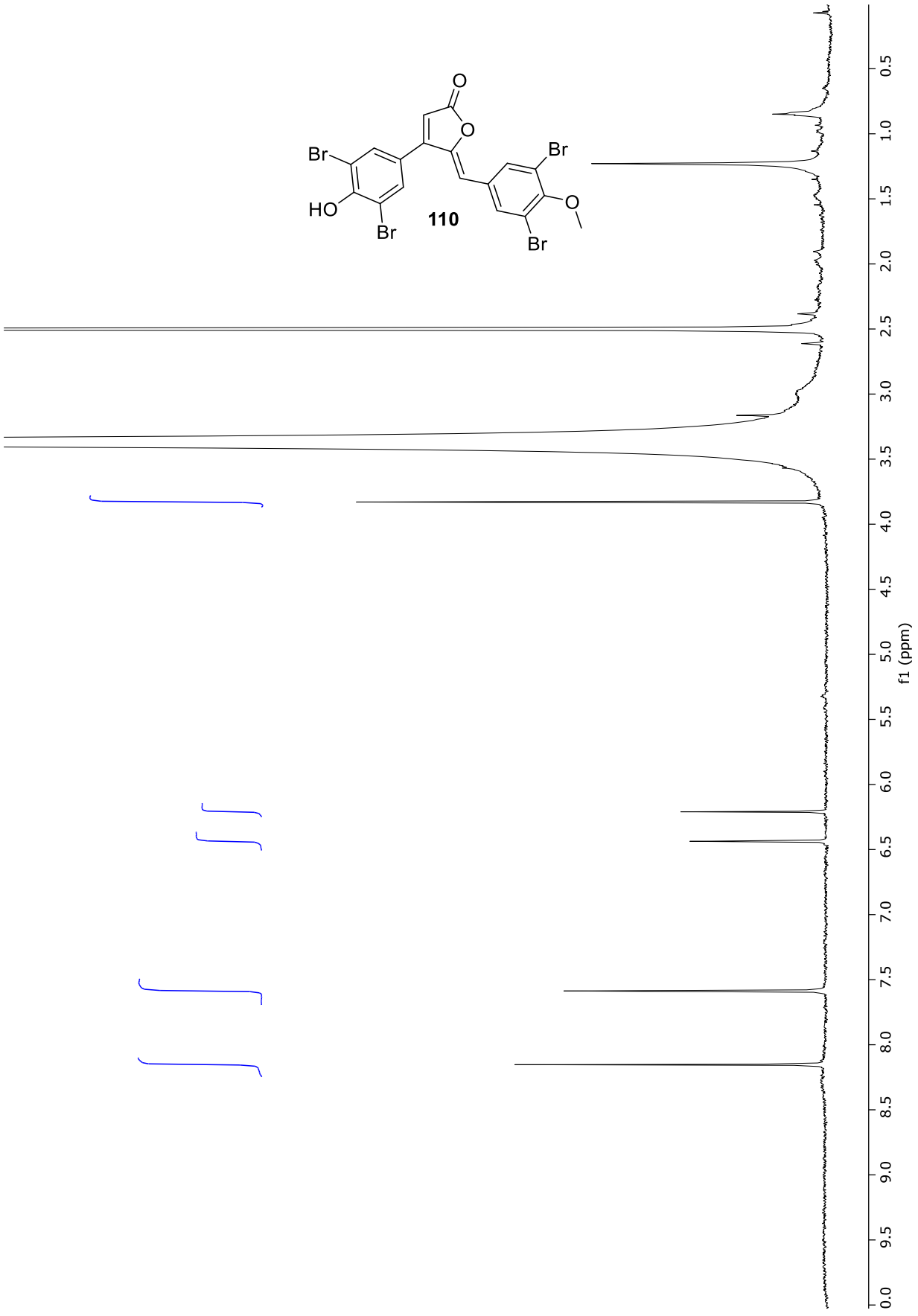
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **108**



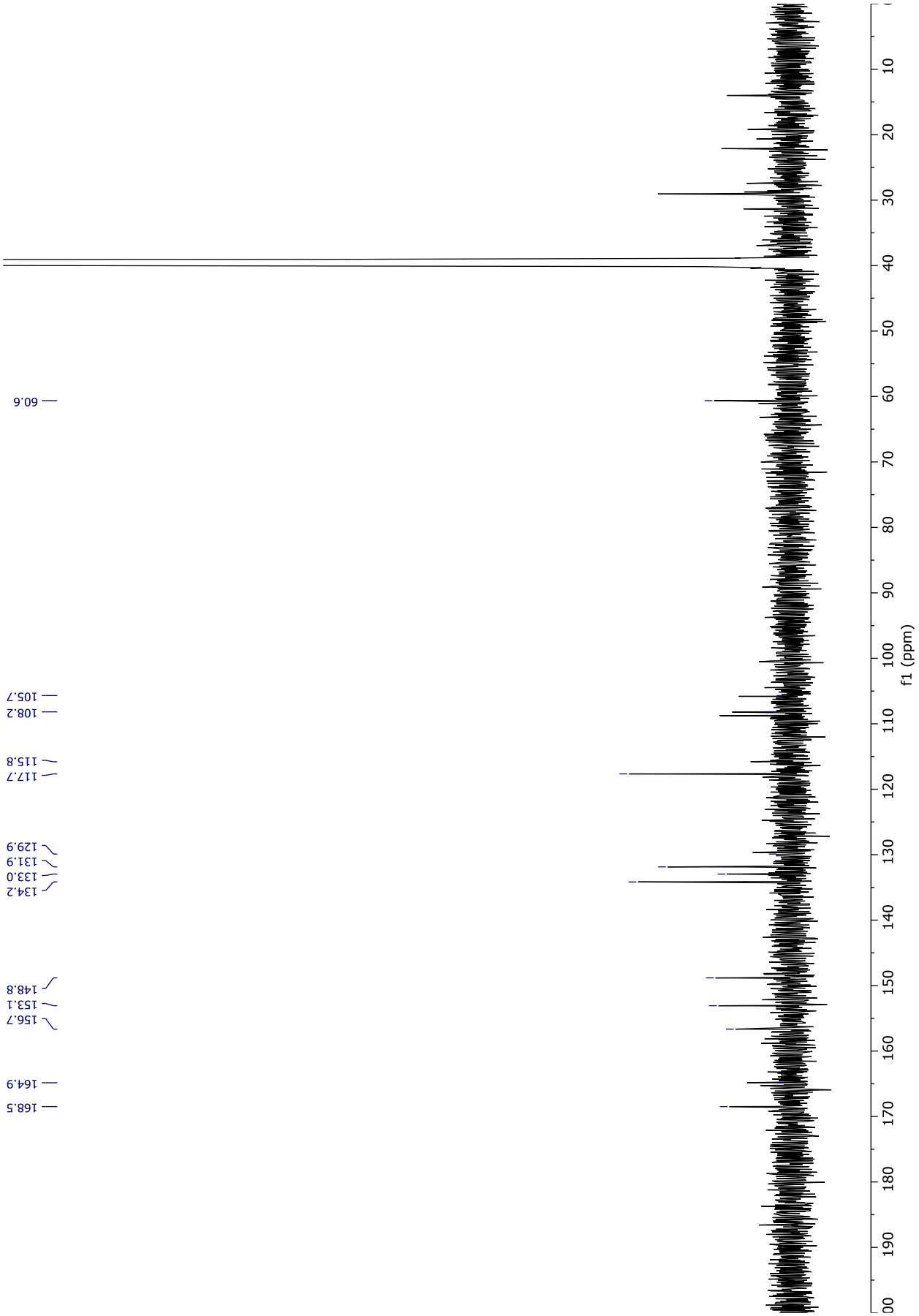
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **109**



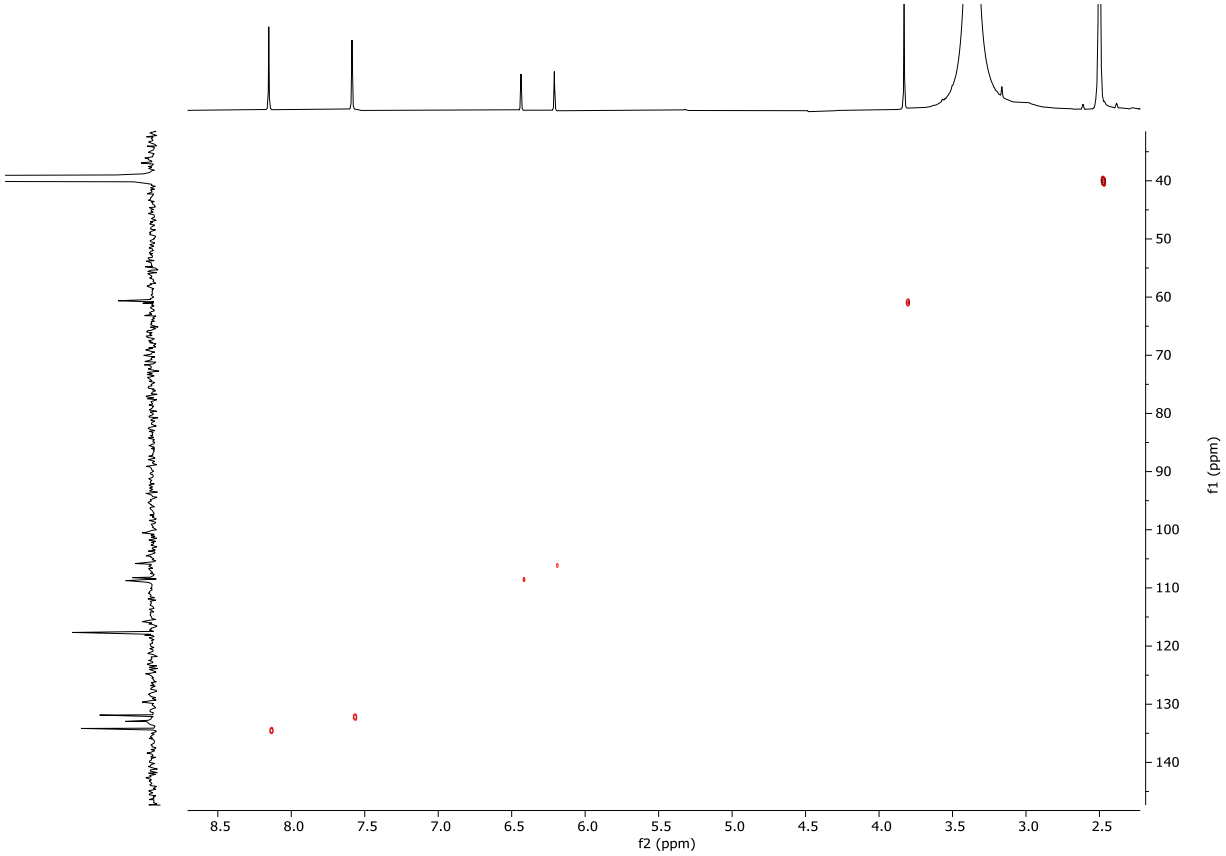
HSQC NMR Spectrum (600 MHz, DMSO-d_6) of **109**



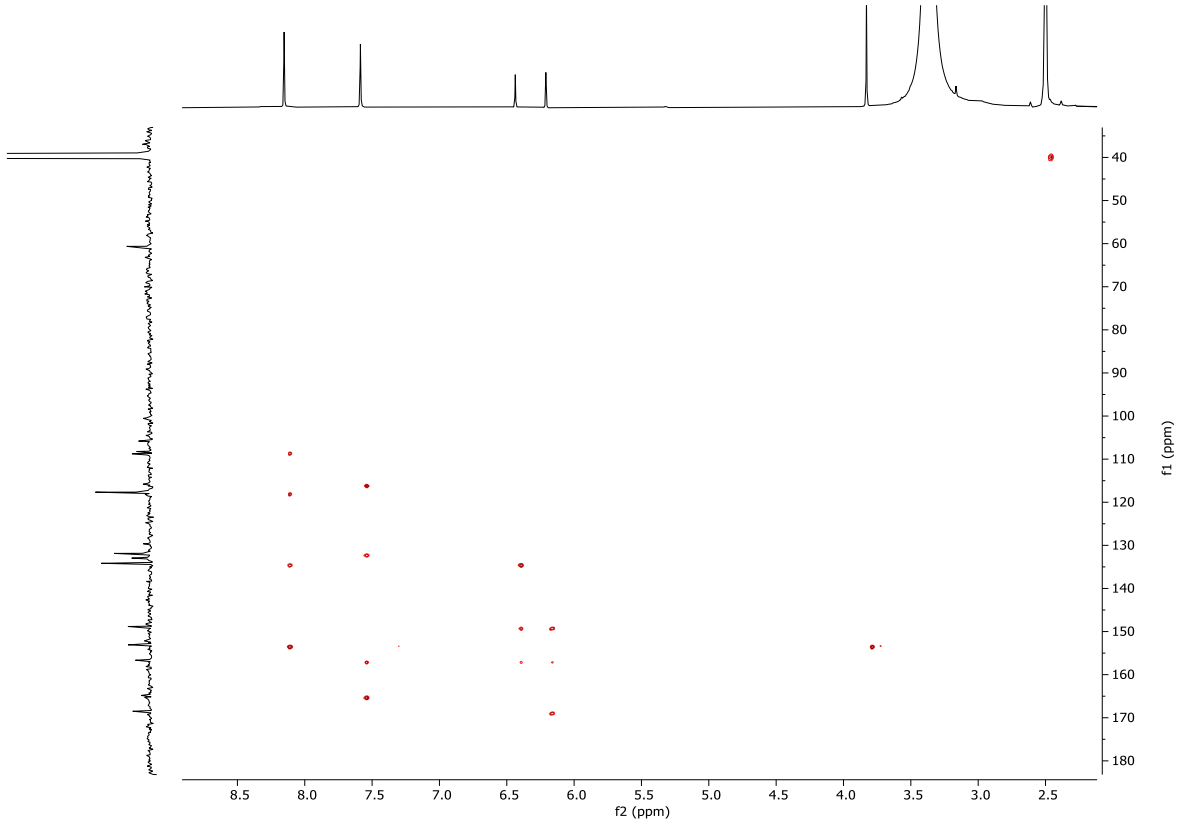
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **110**



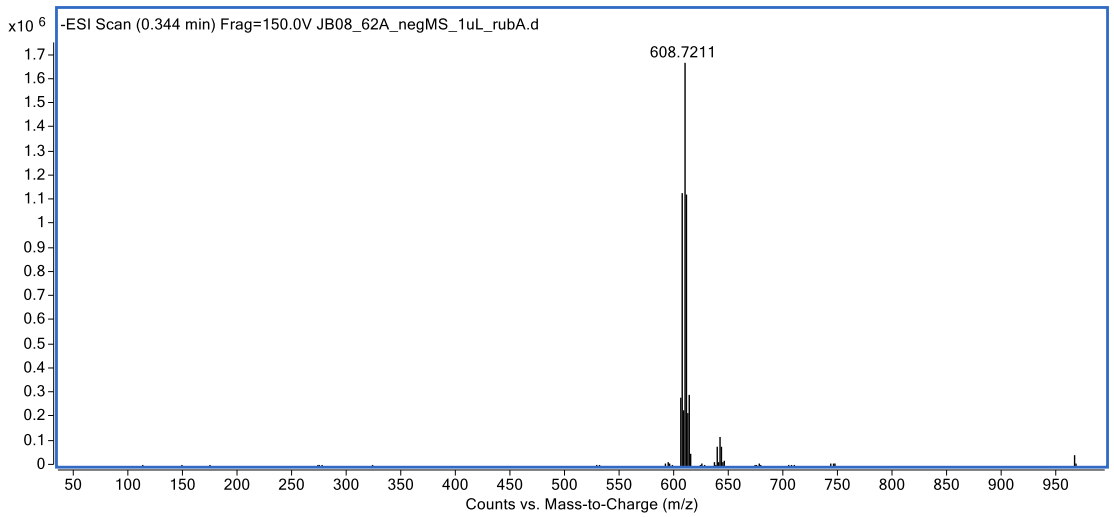
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **110**

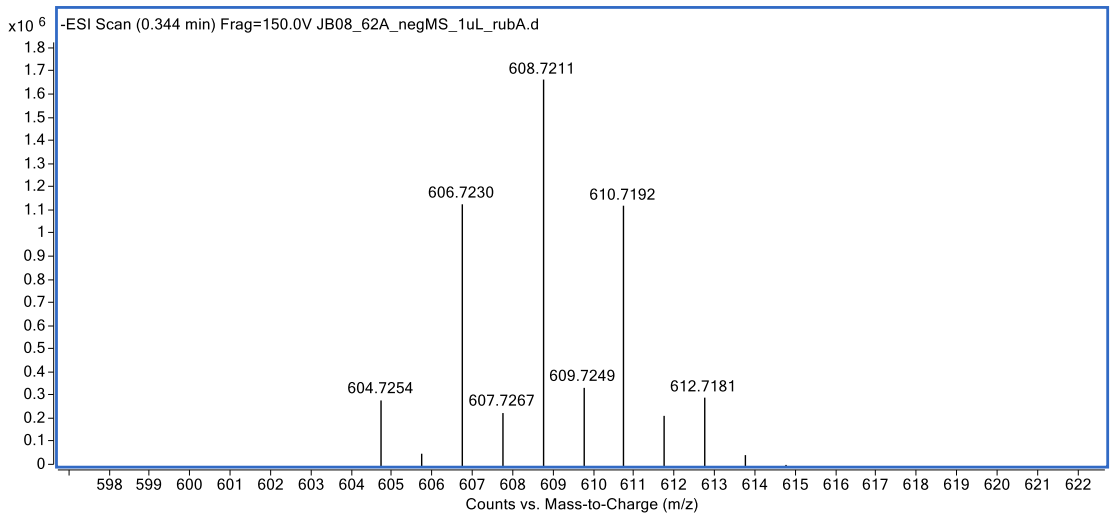


HSQC NMR Spectrum (600 MHz, DMSO-d_6) of **110**

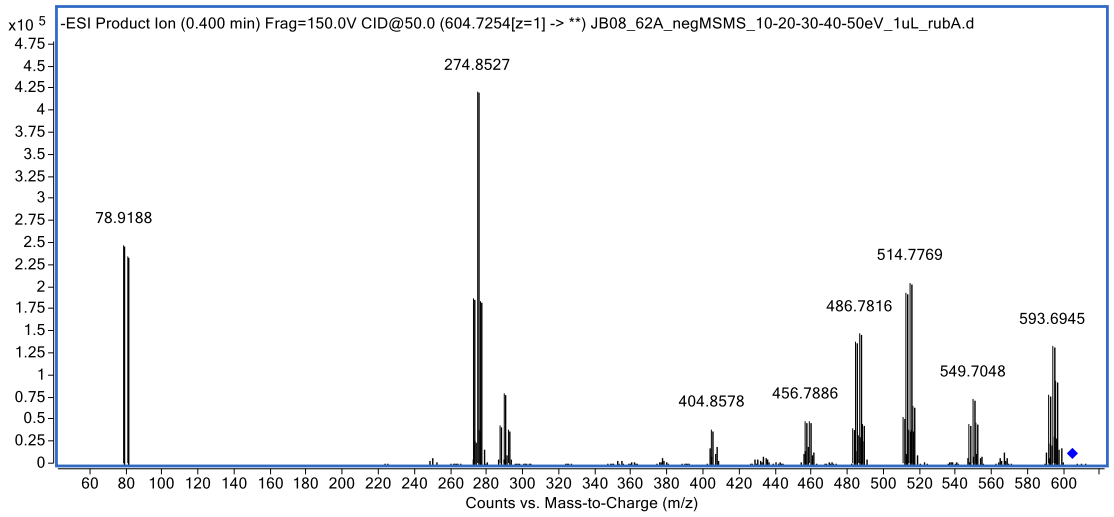


HMBC NMR Spectrum (600 MHz, DMSO-d_6) of **110**

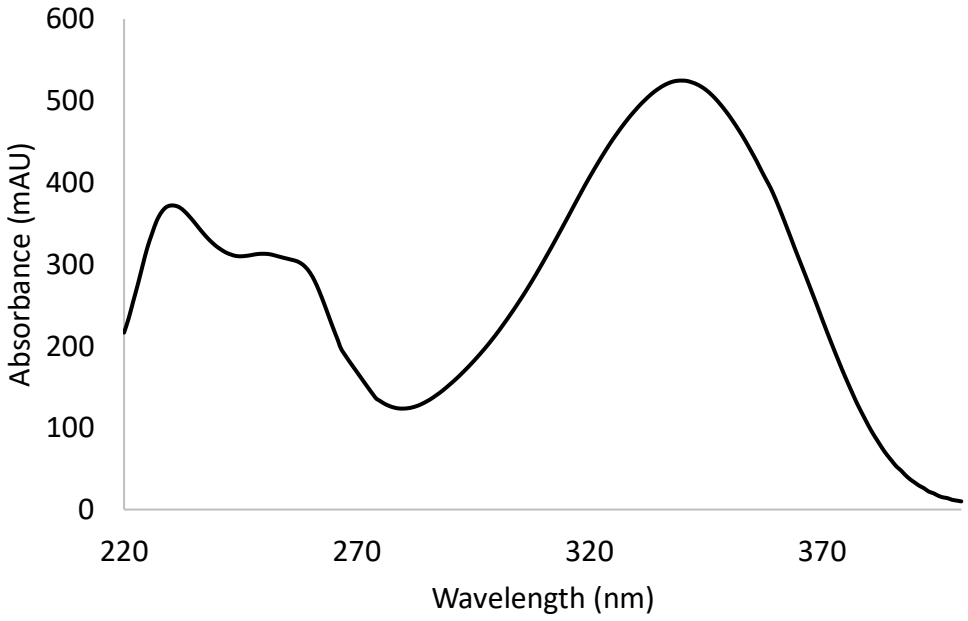




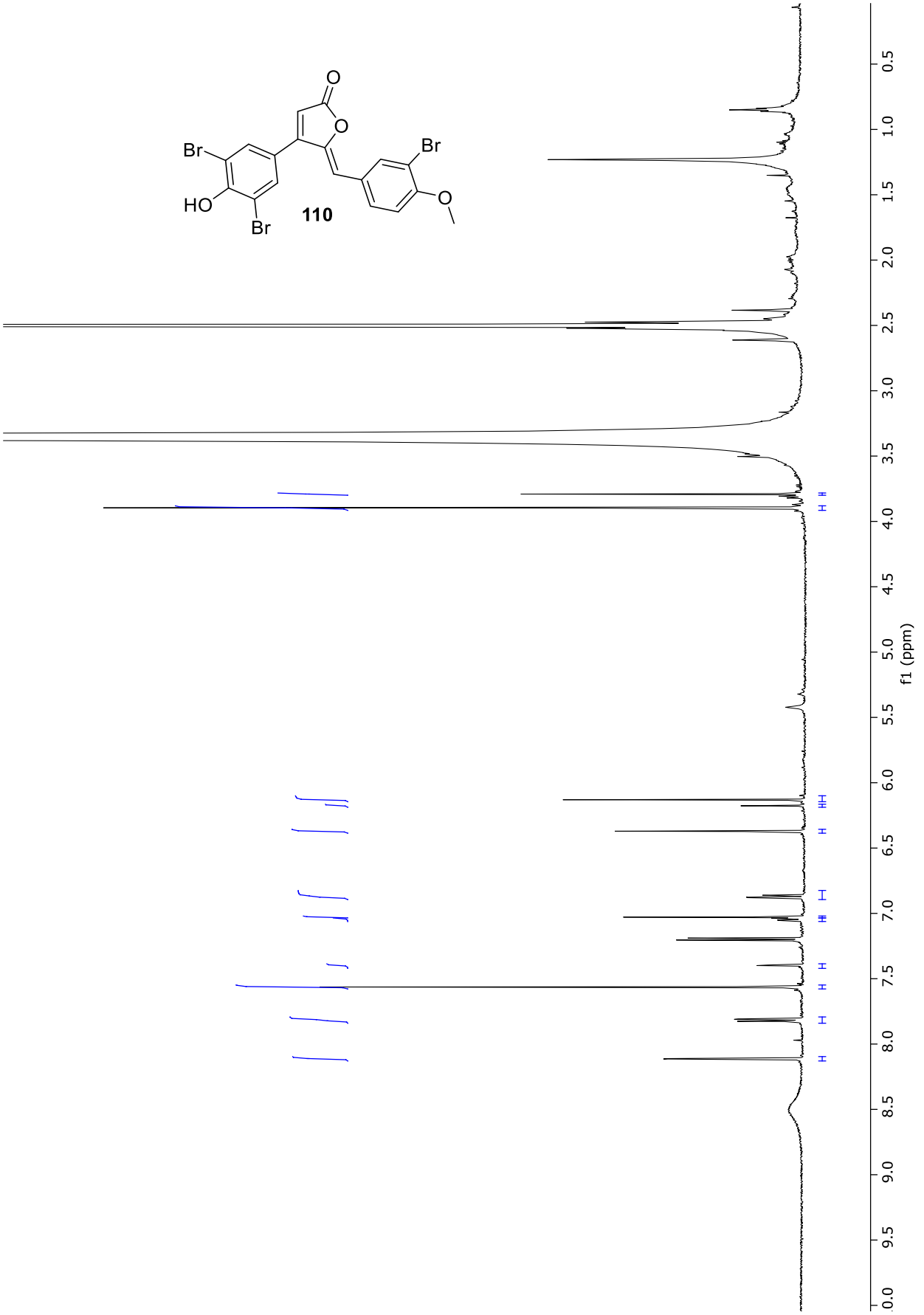
(–)-HRESIMS spectra of **110**



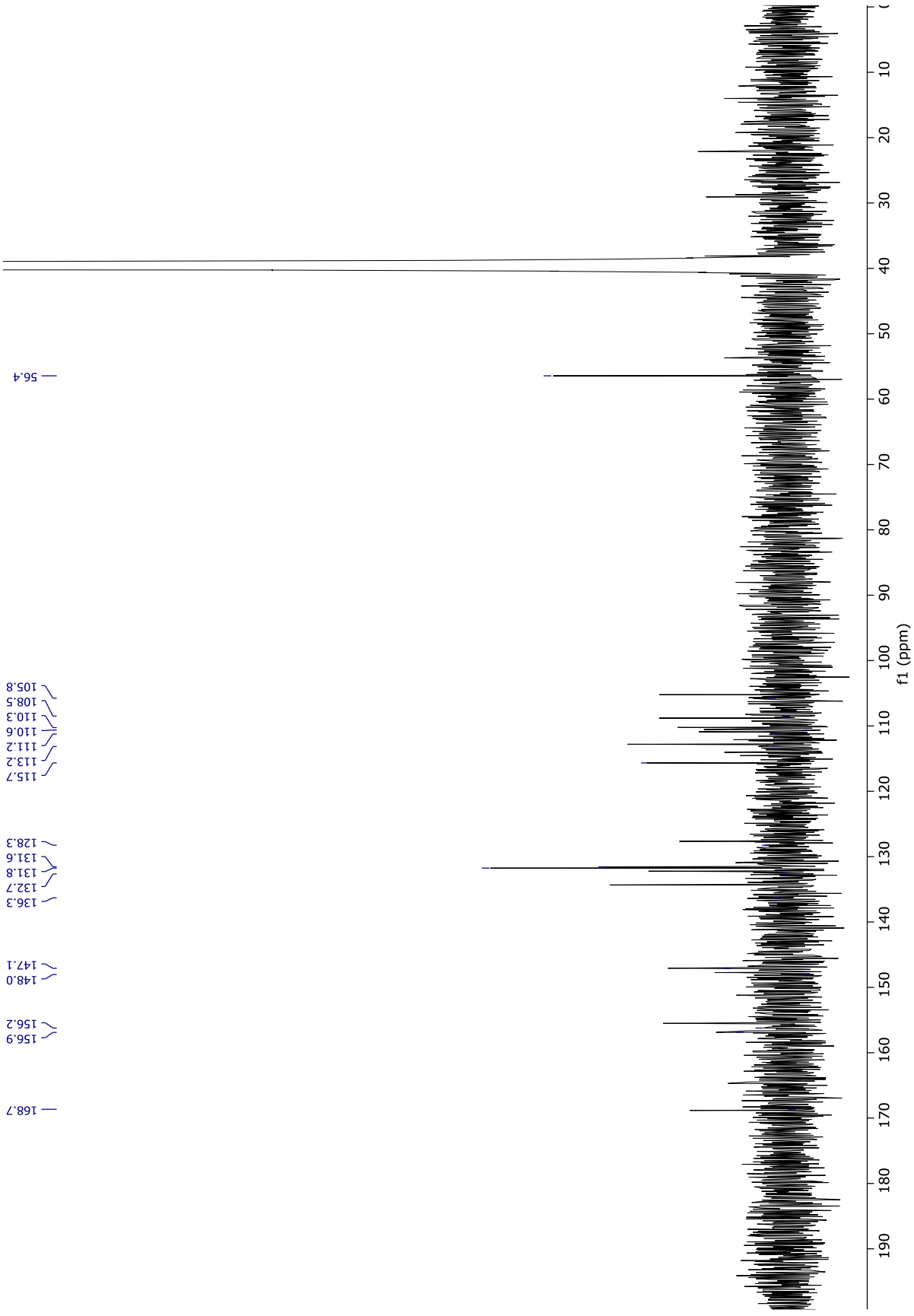
(–)-HRESIMS/MS spectrum of **110**



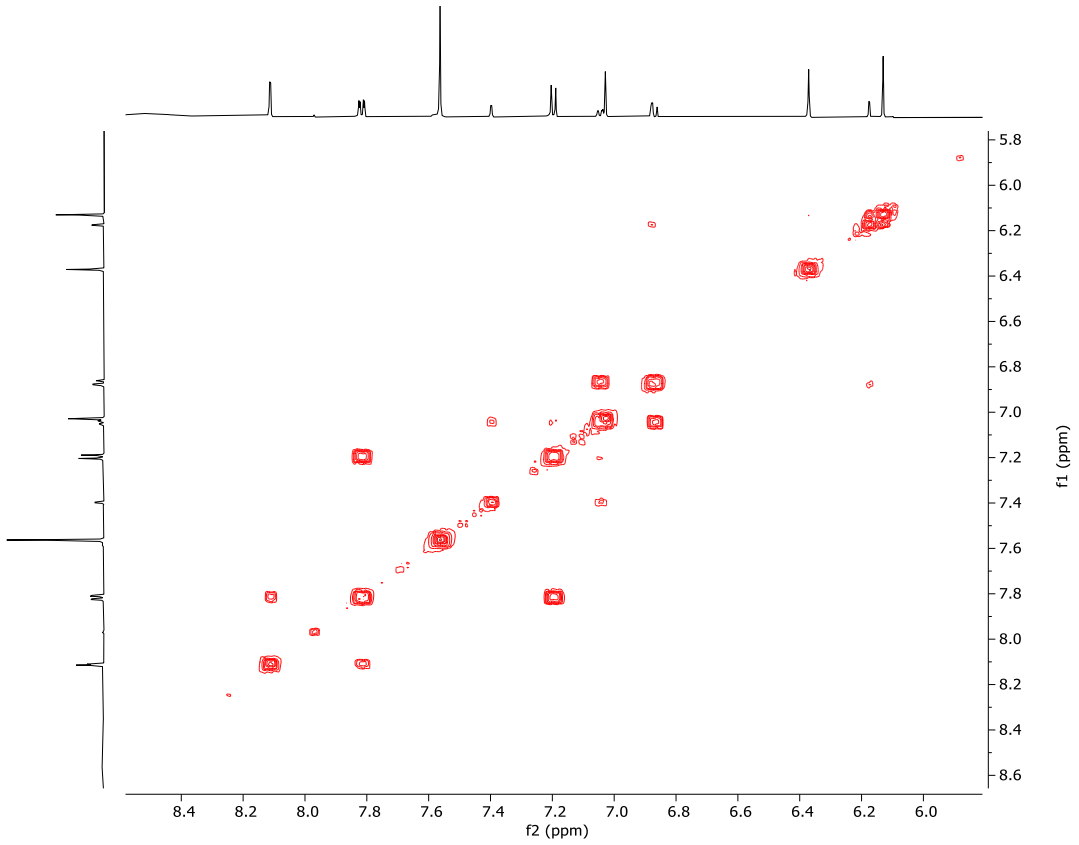
UV/vis spectrum of **110**



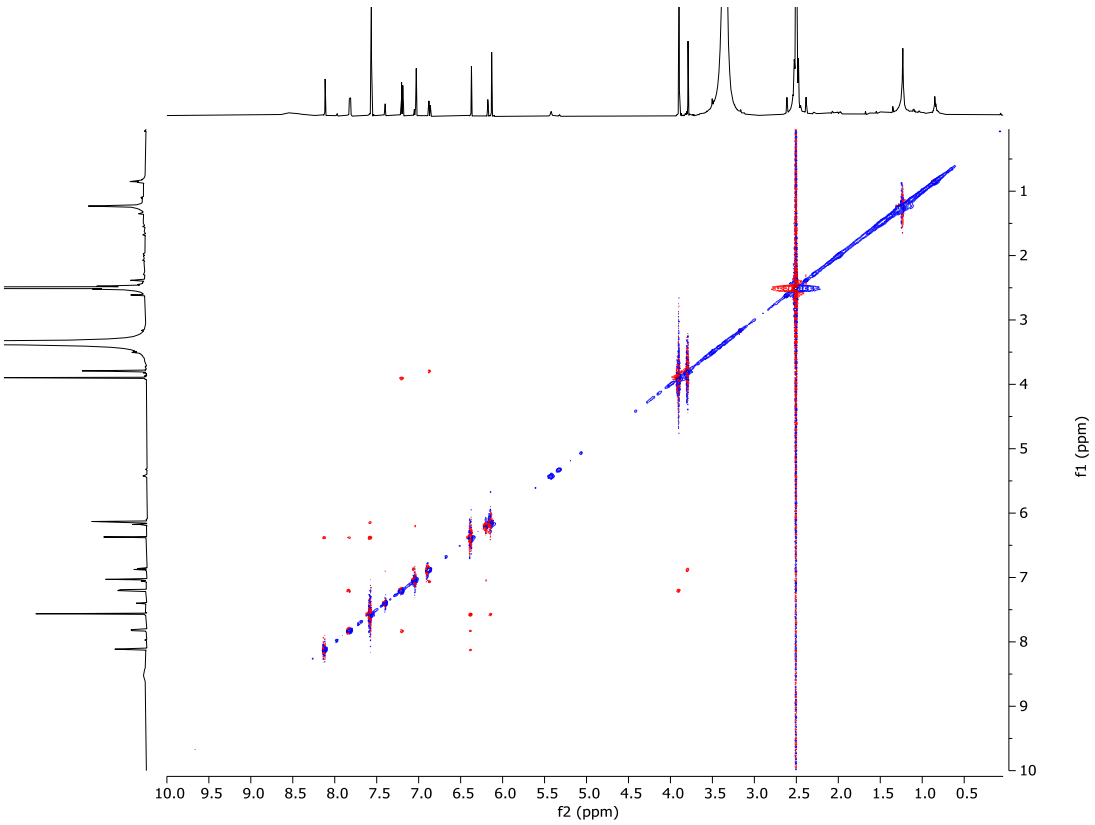
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **111**



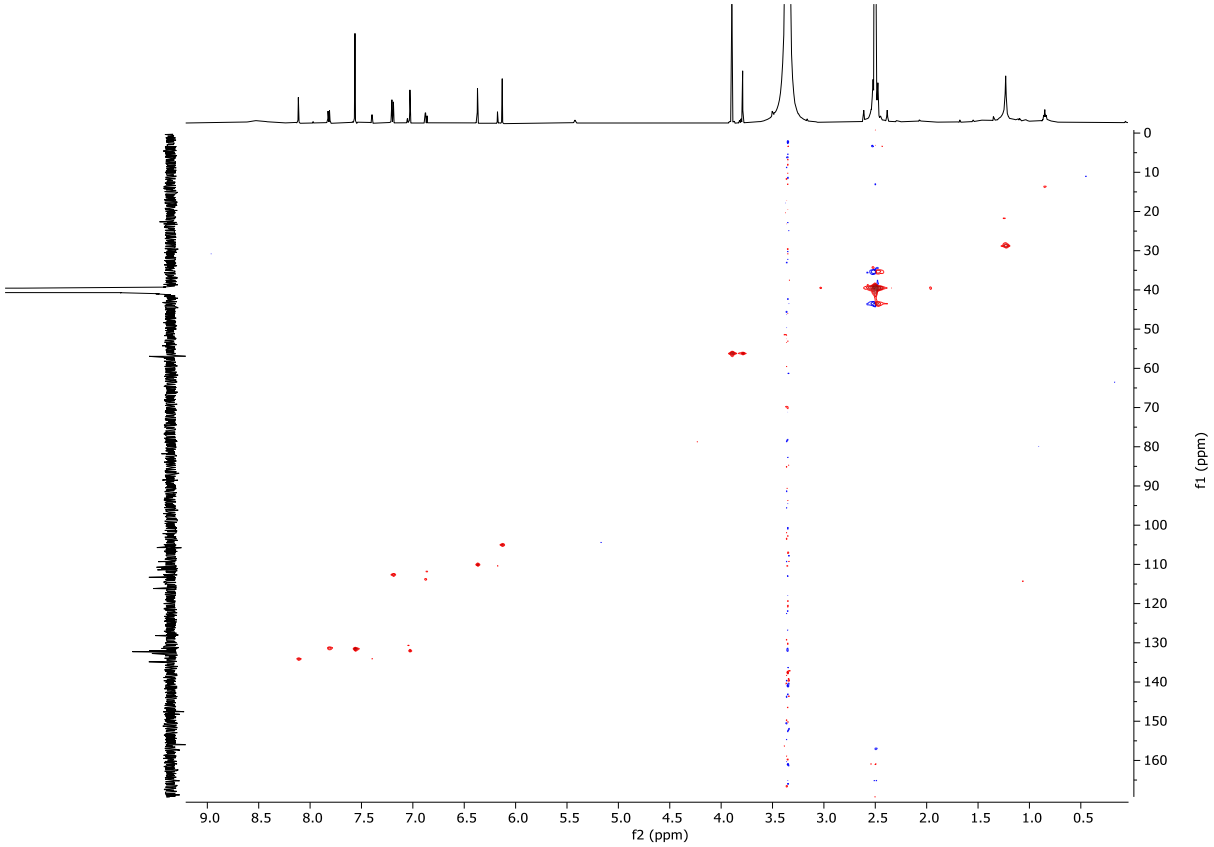
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **111**



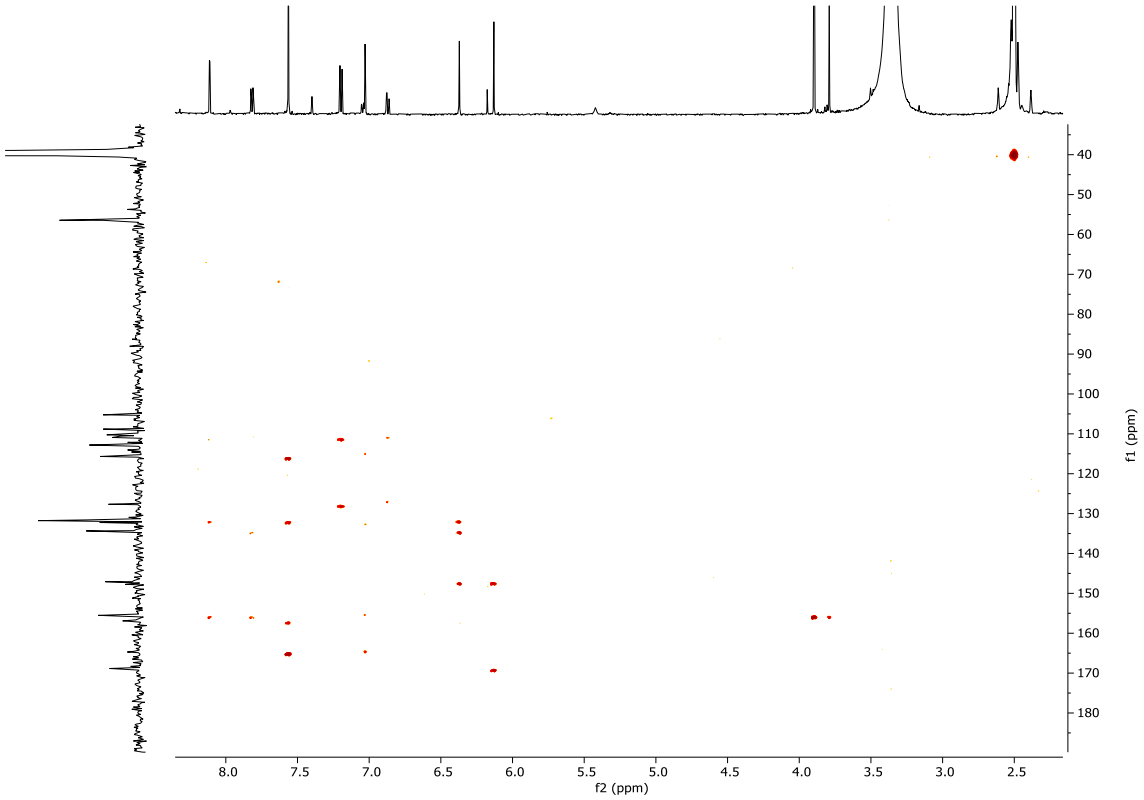
COSY NMR Spectrum (600 MHz, DMSO-d_6) of **111**



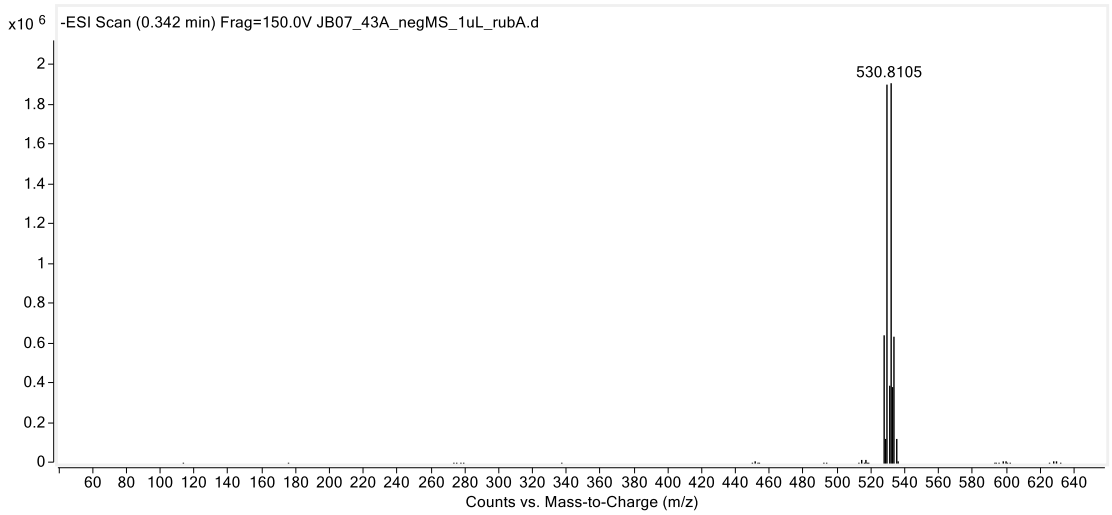
ROESY NMR Spectrum (600 MHz, DMSO-d_6) of **111**

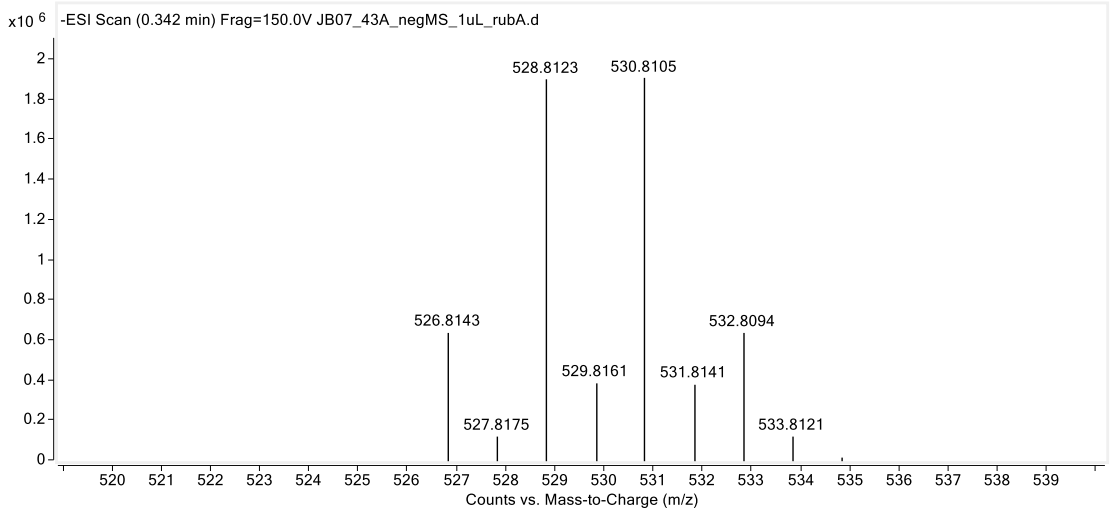


HSQC NMR Spectrum (600 MHz, DMSO-d_6) of **111**

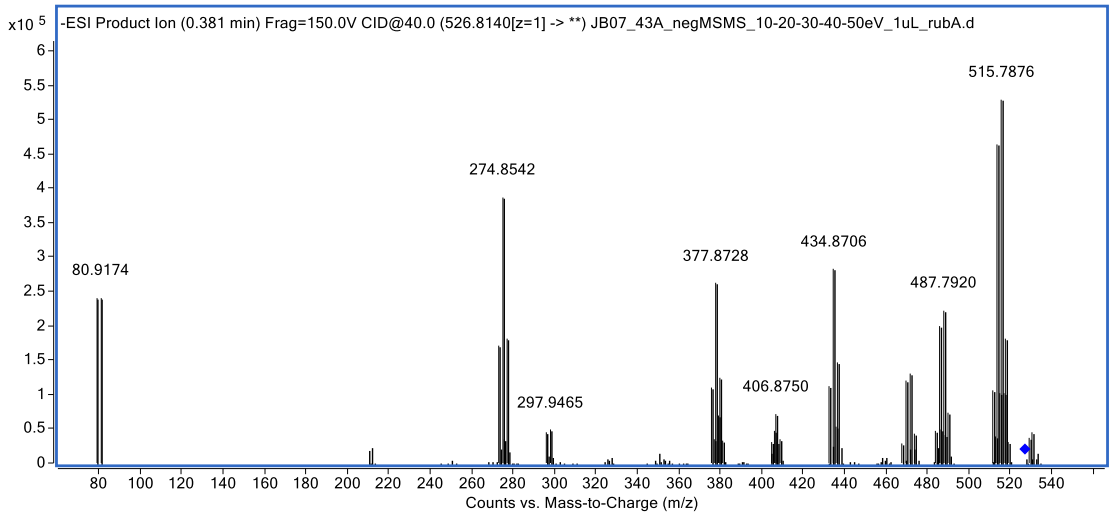


HMBC NMR Spectrum (600 MHz, DMSO-d_6) of **111**

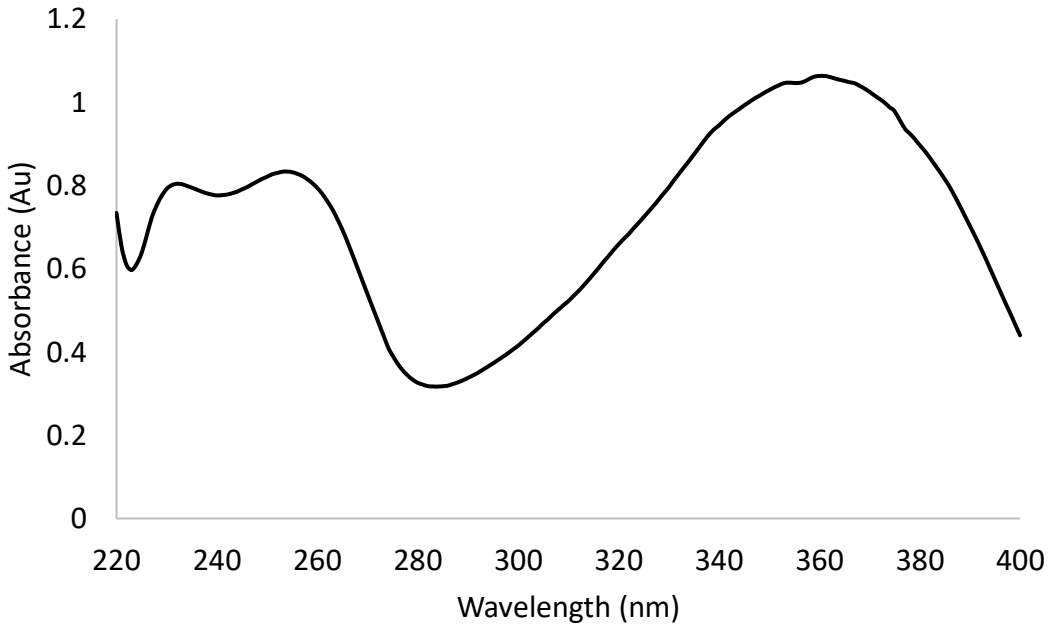




(–)-HRESIMS spectra of **111**

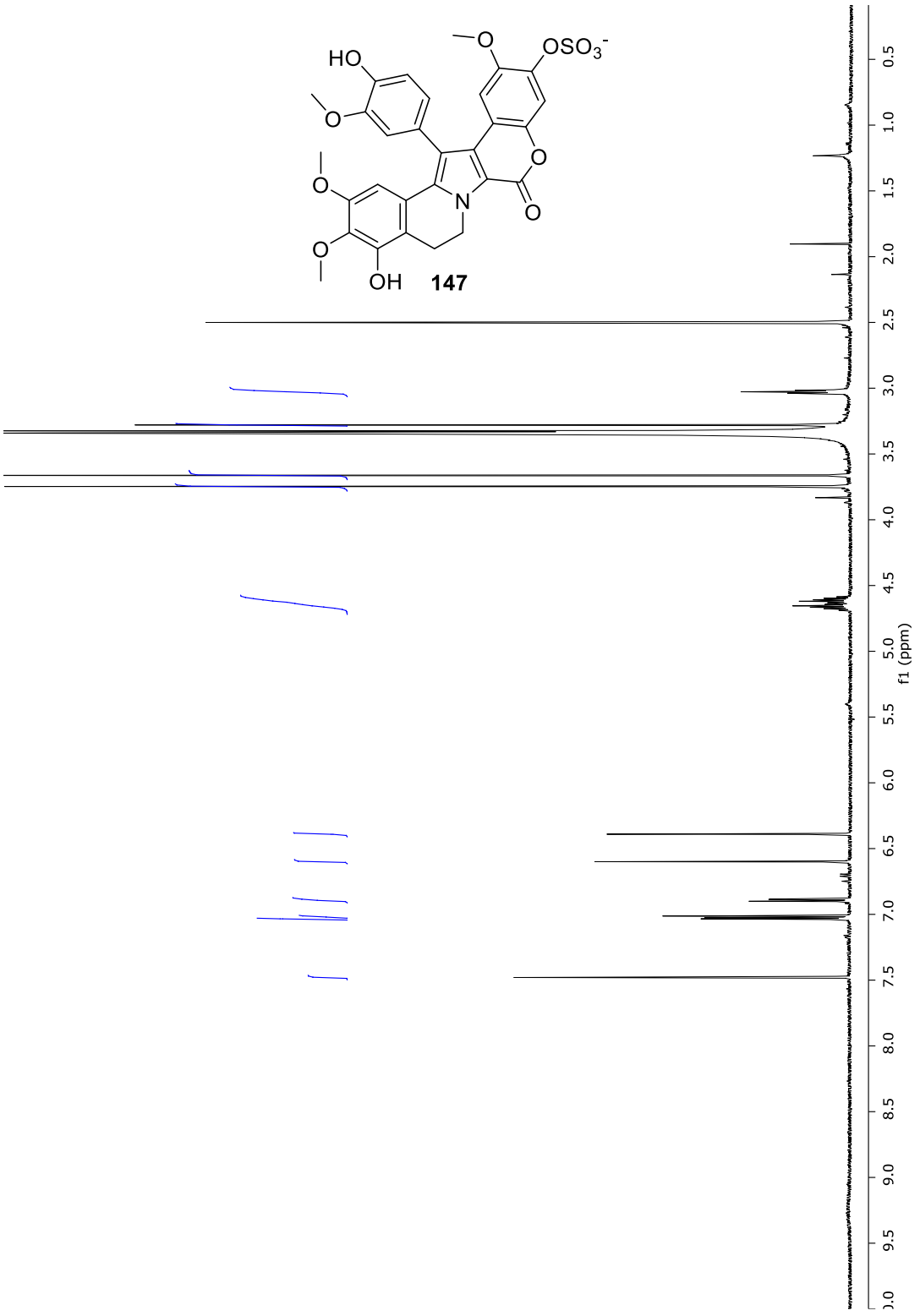


(–)-HRESIMS/MS spectrum of **111**

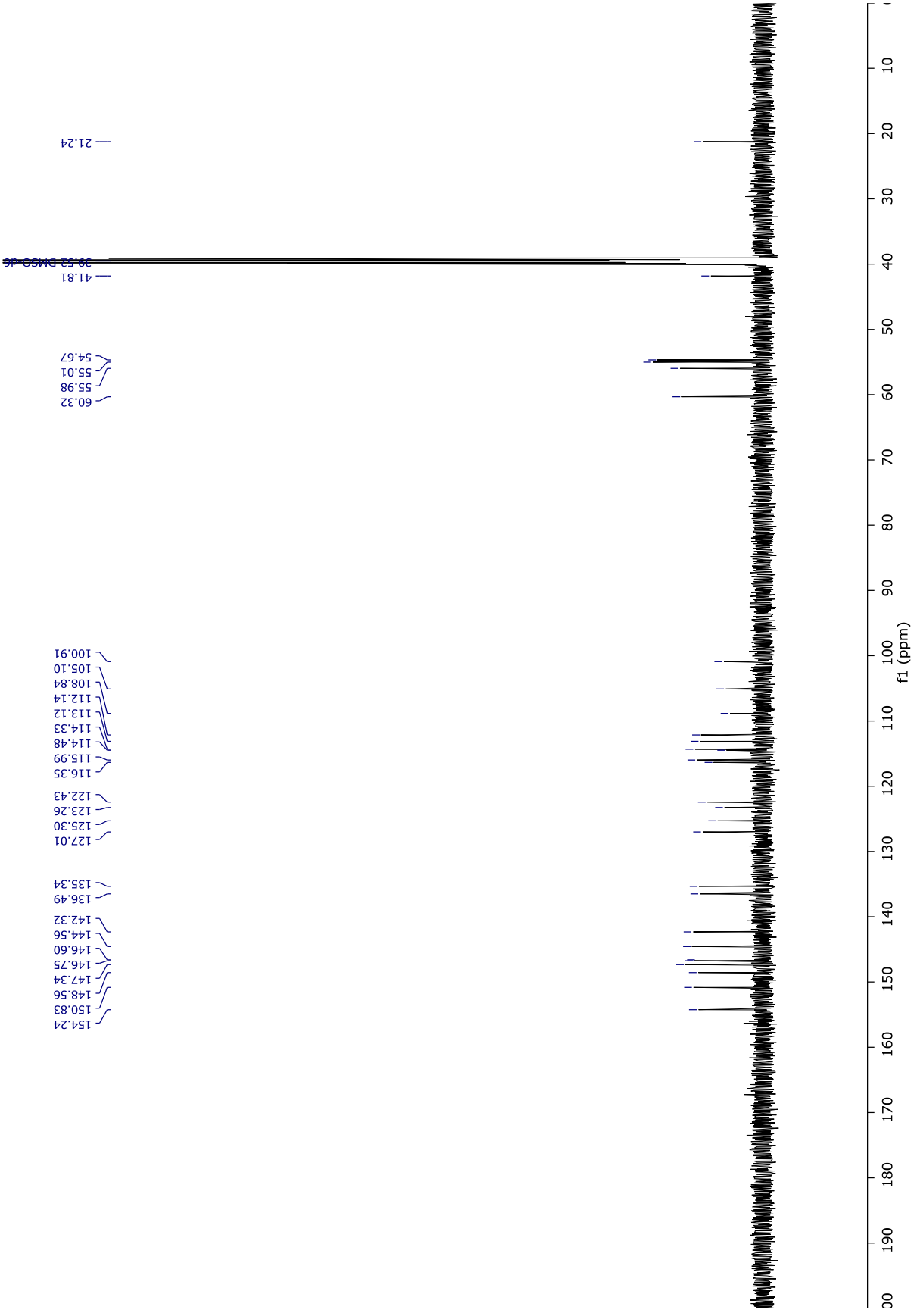


UV/vis spectrum of **111**

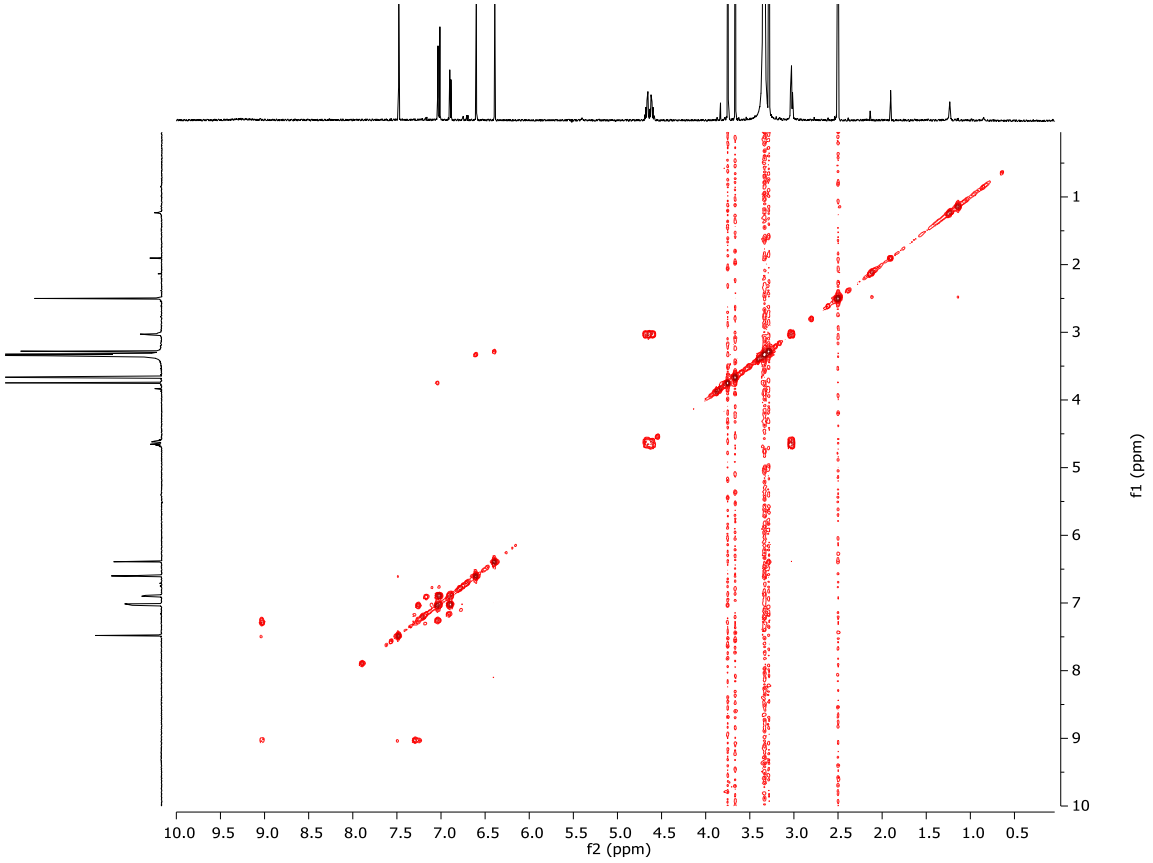
Appendix 8 – *Didemnum ternerratum* Spectra



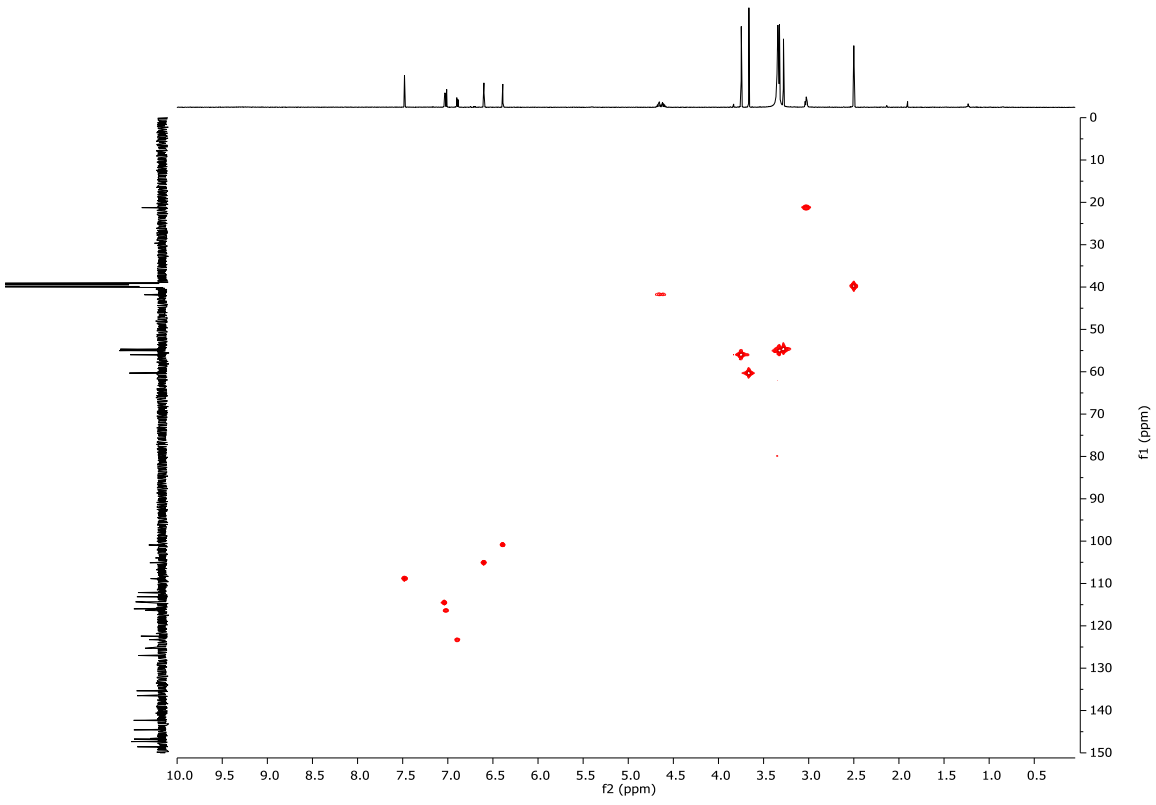
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **147**



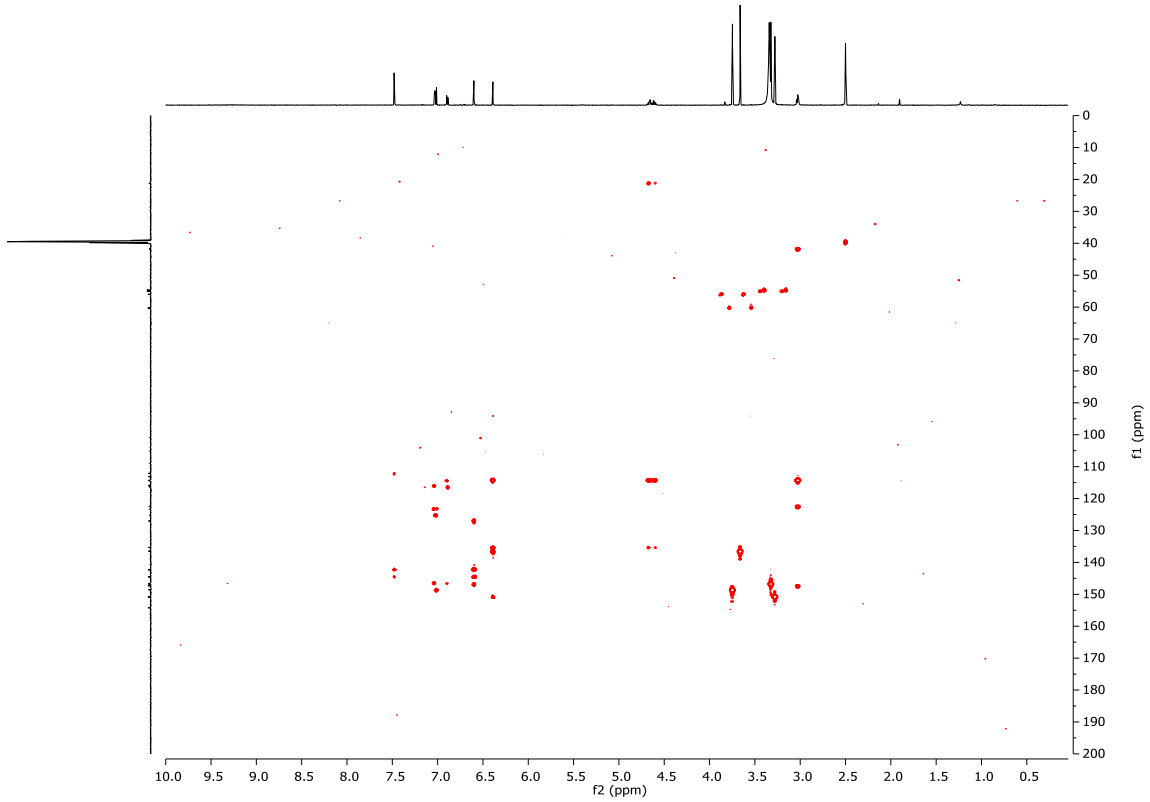
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **147**



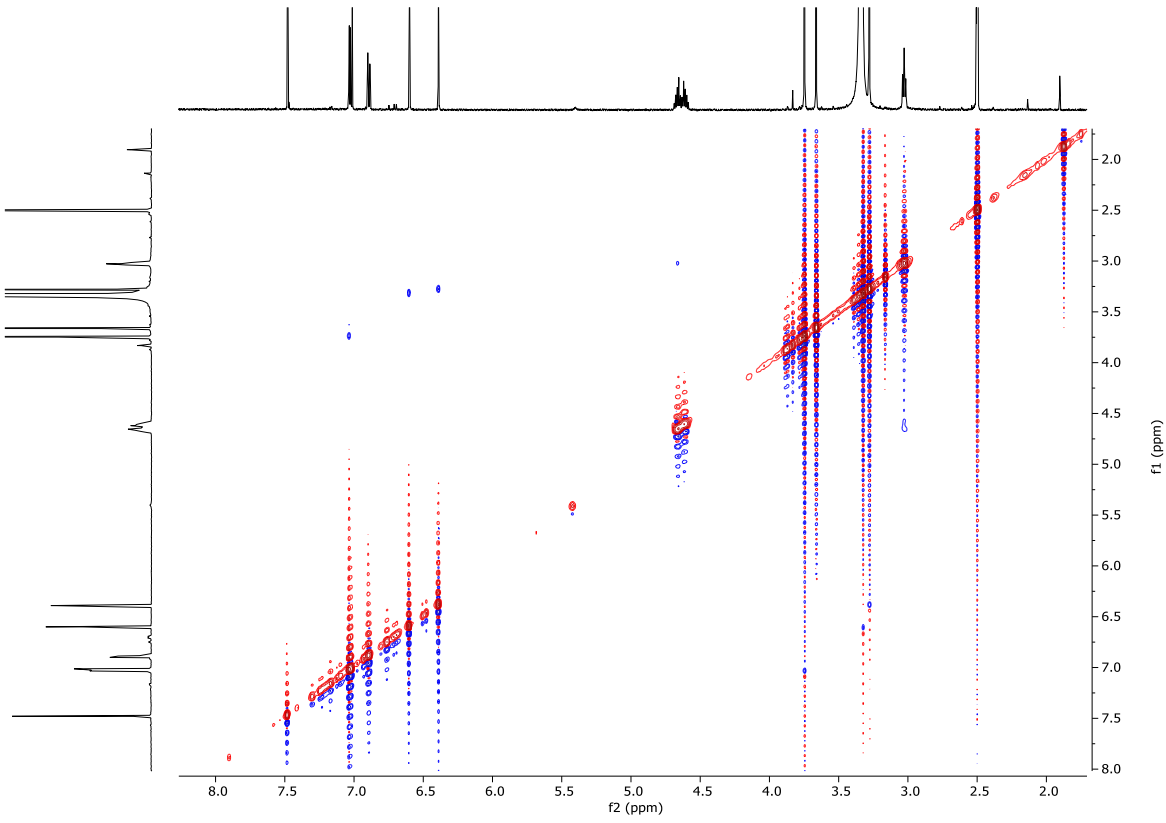
COSY Spectrum (600 MHz, DMSO-d_6) of **147**



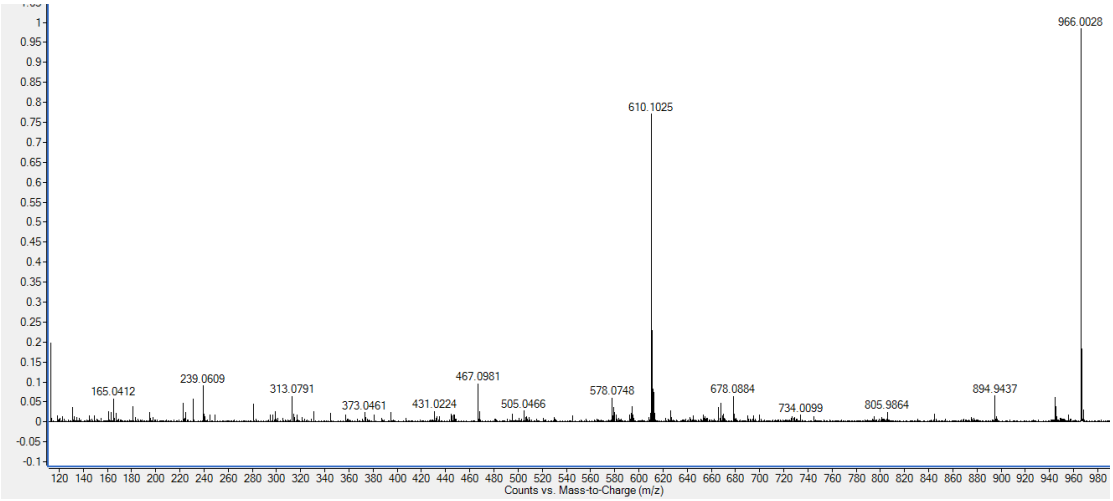
HSQC Spectrum (600 MHz, DMSO-d_6) of **147**



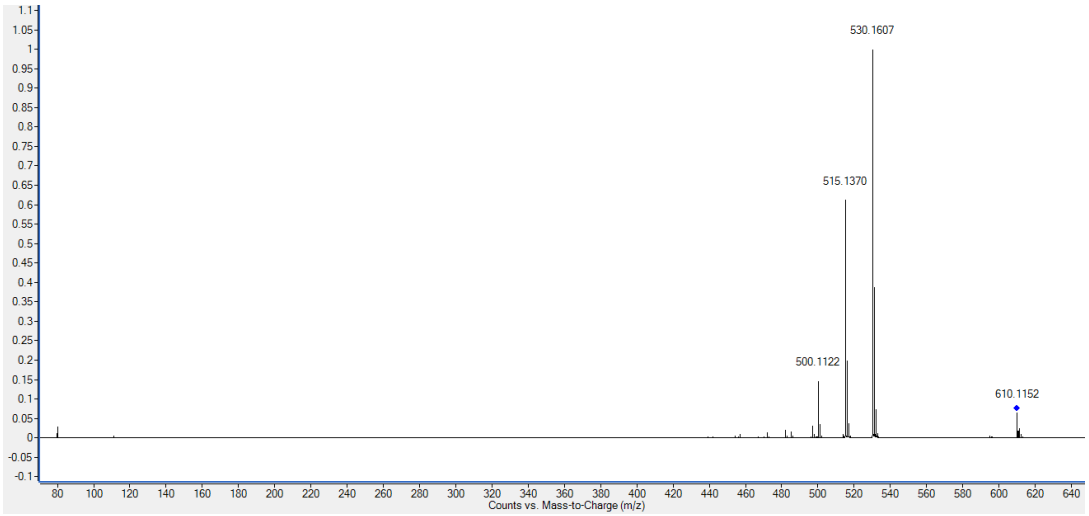
HMBC Spectrum (600 MHz, DMSO-d_6) of **147**



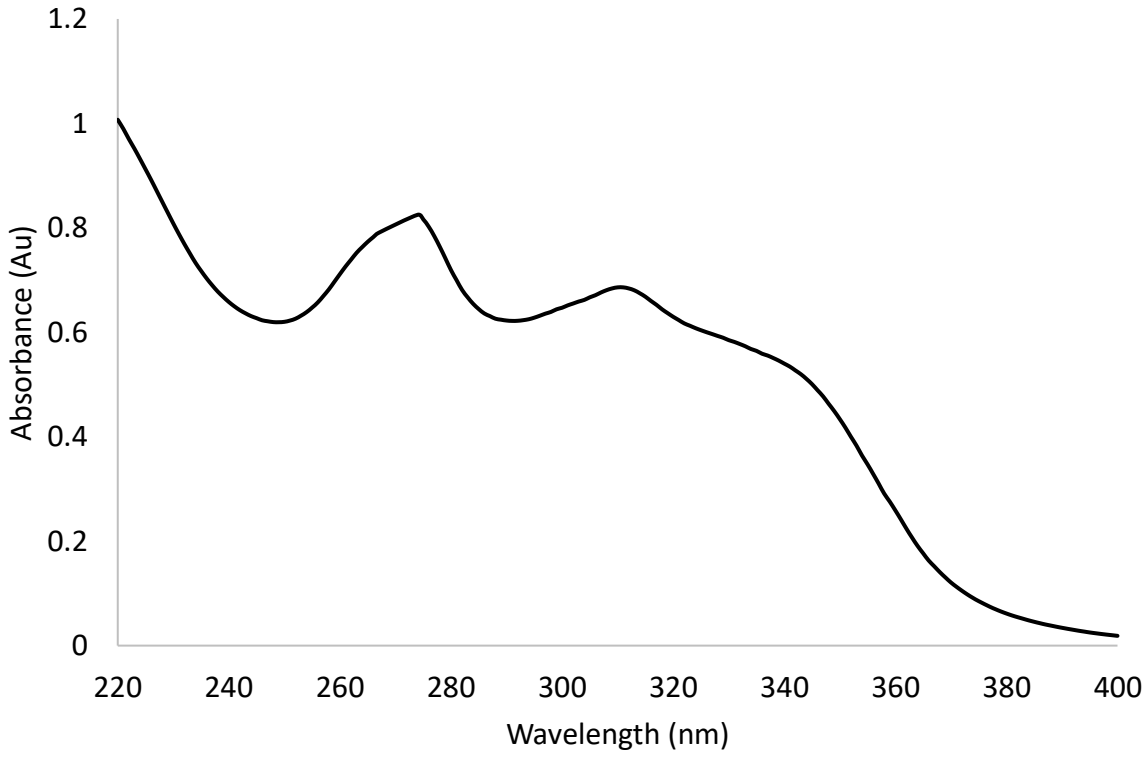
ROESY NMR Spectrum (600 MHz, DMSO-d_6) of **147**



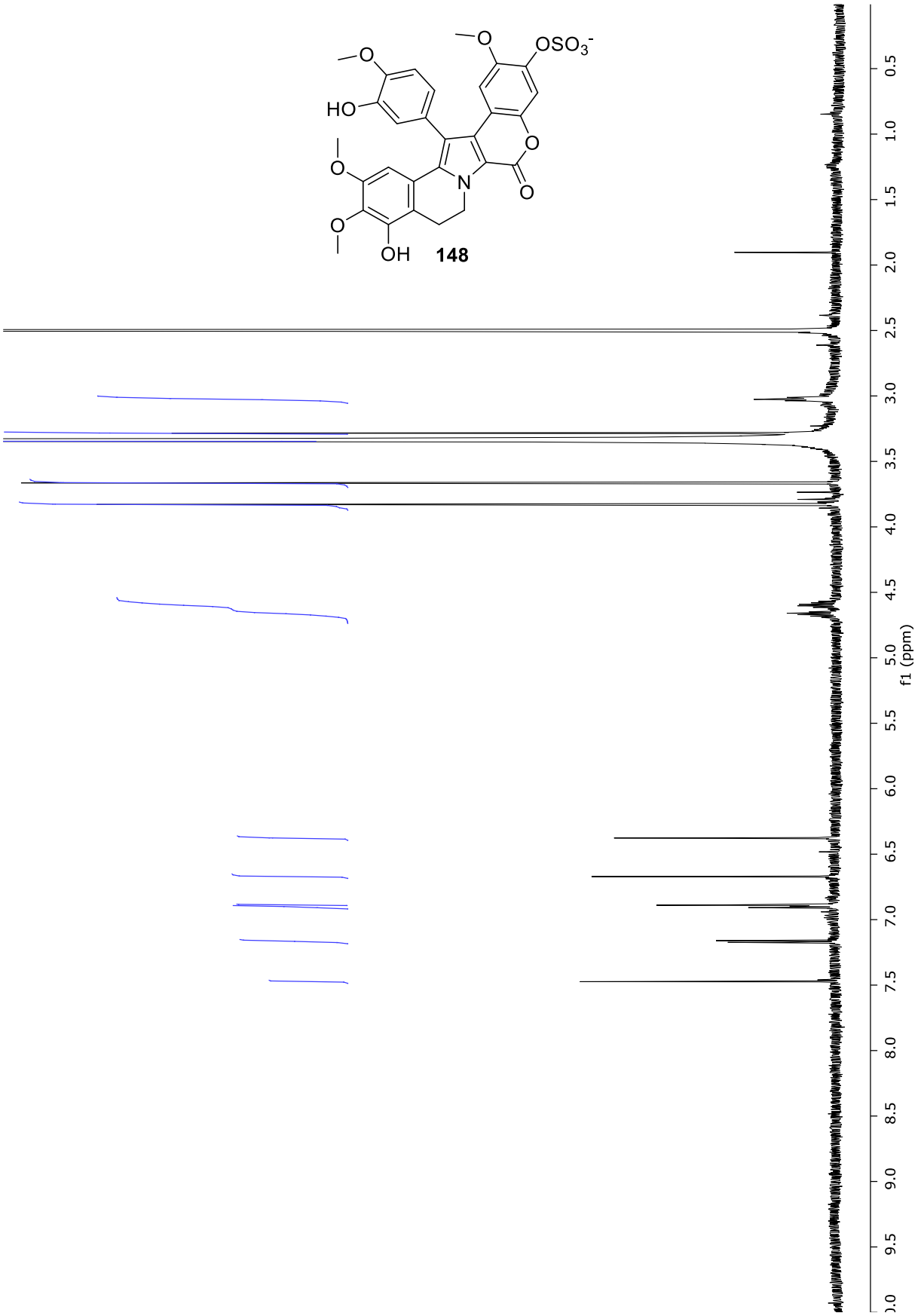
(–)-HRESIMS Spectrum of **147**



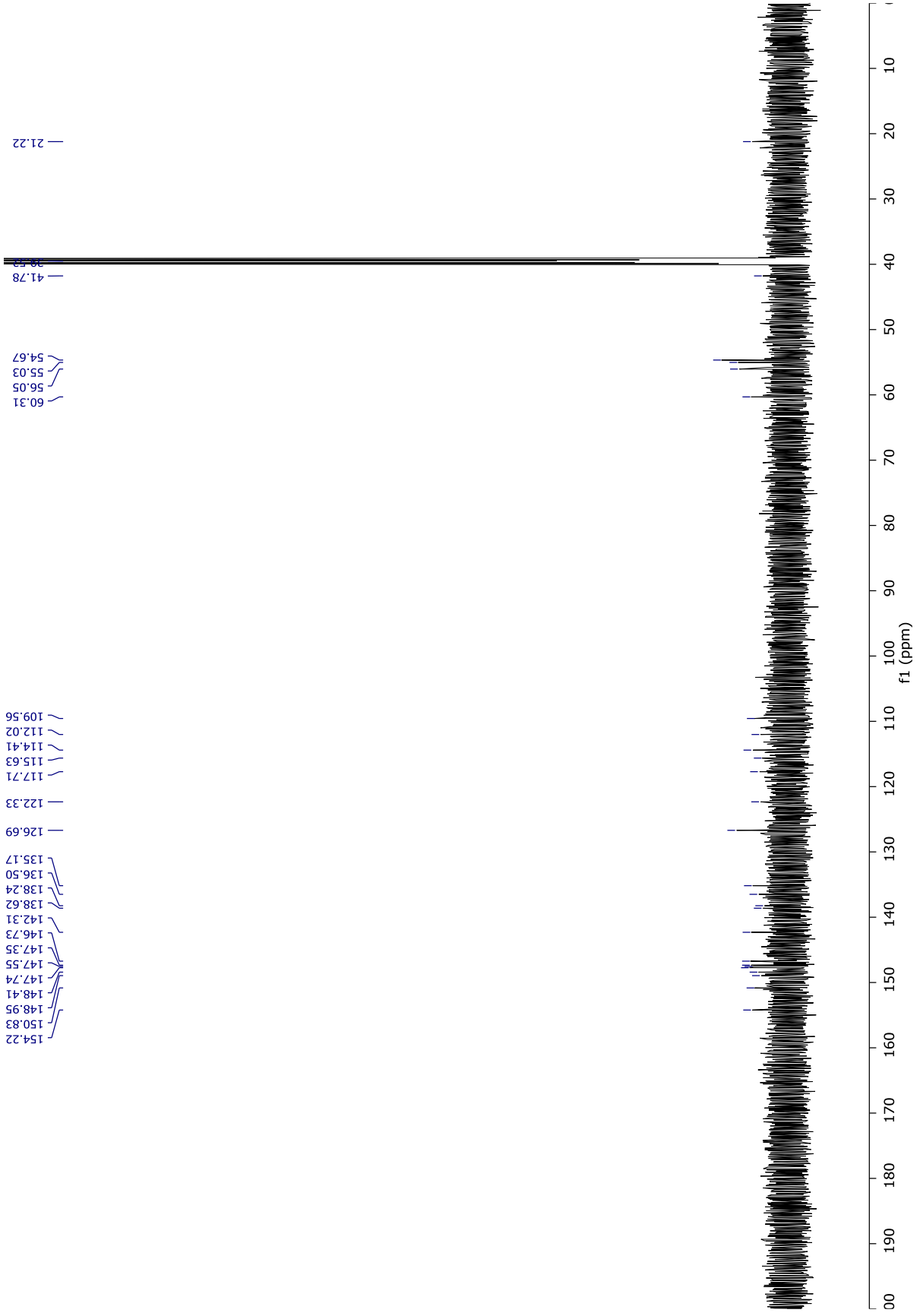
(–)-HRESIMS/MS Spectrum of **147**



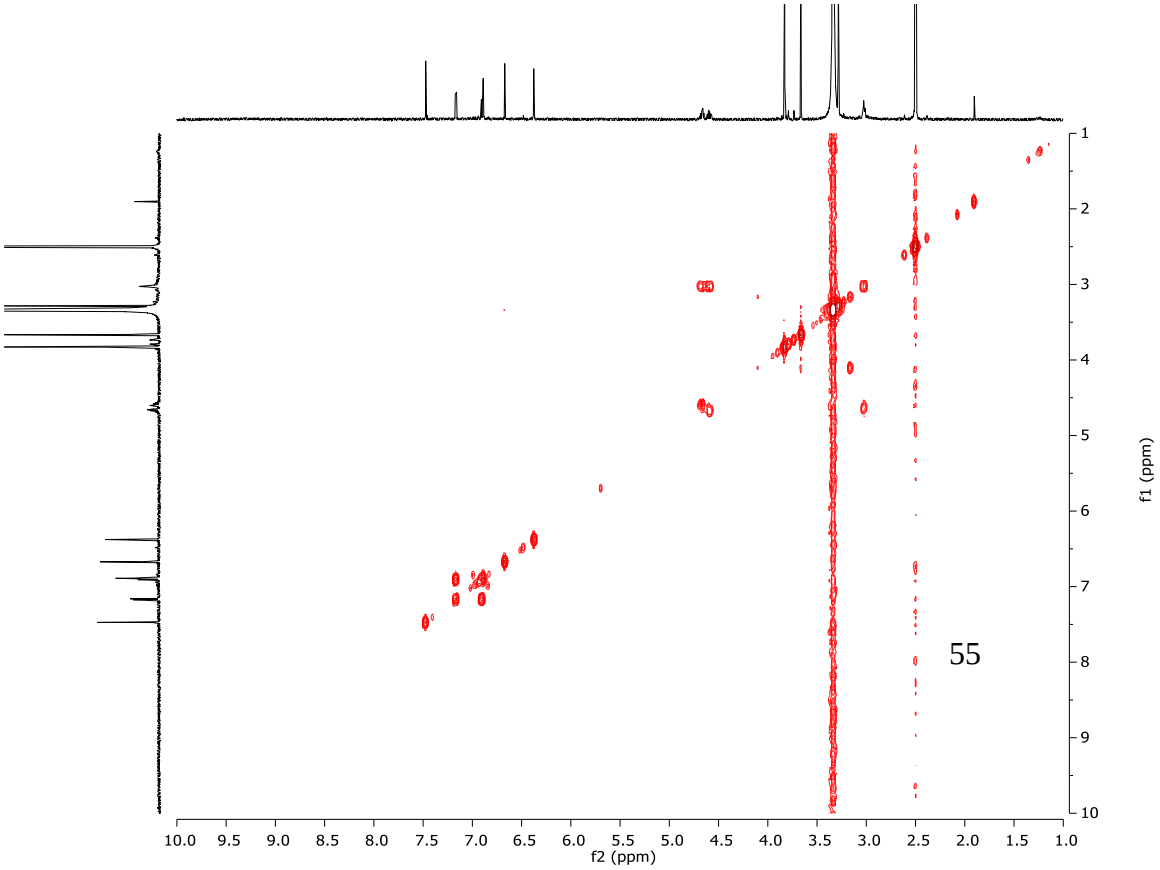
UV/vis Spectrum of **147**



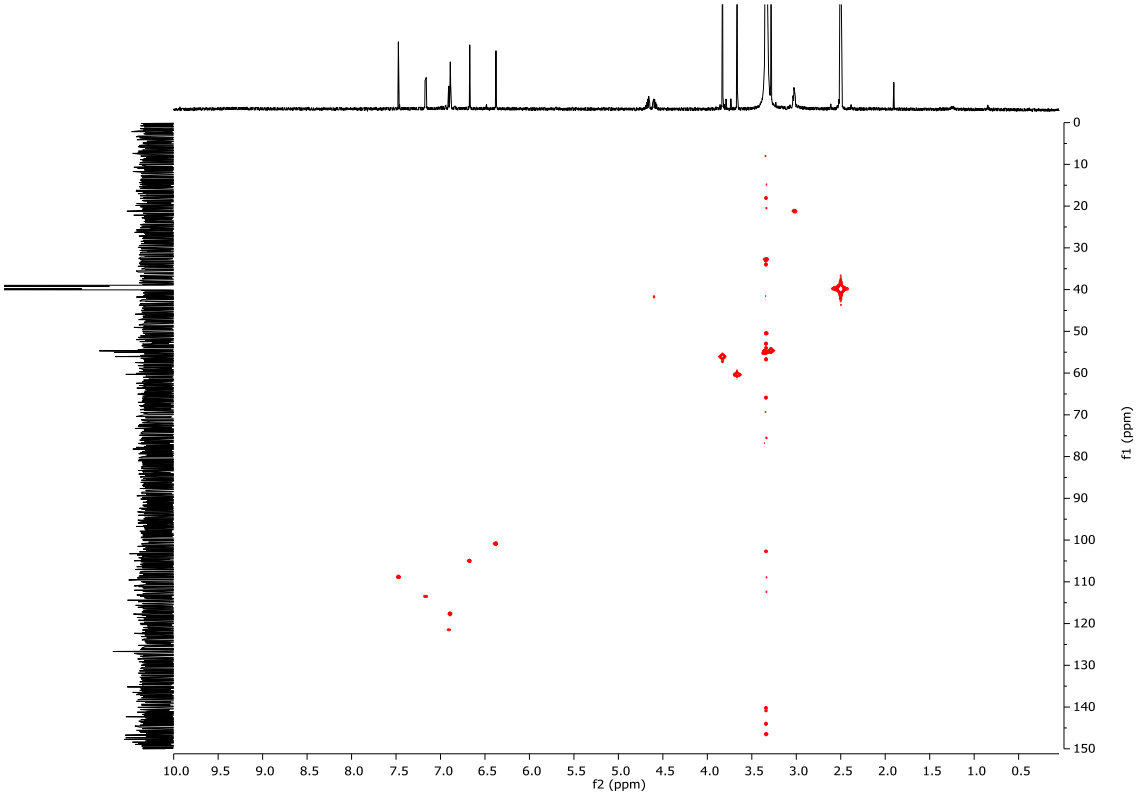
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) **148**



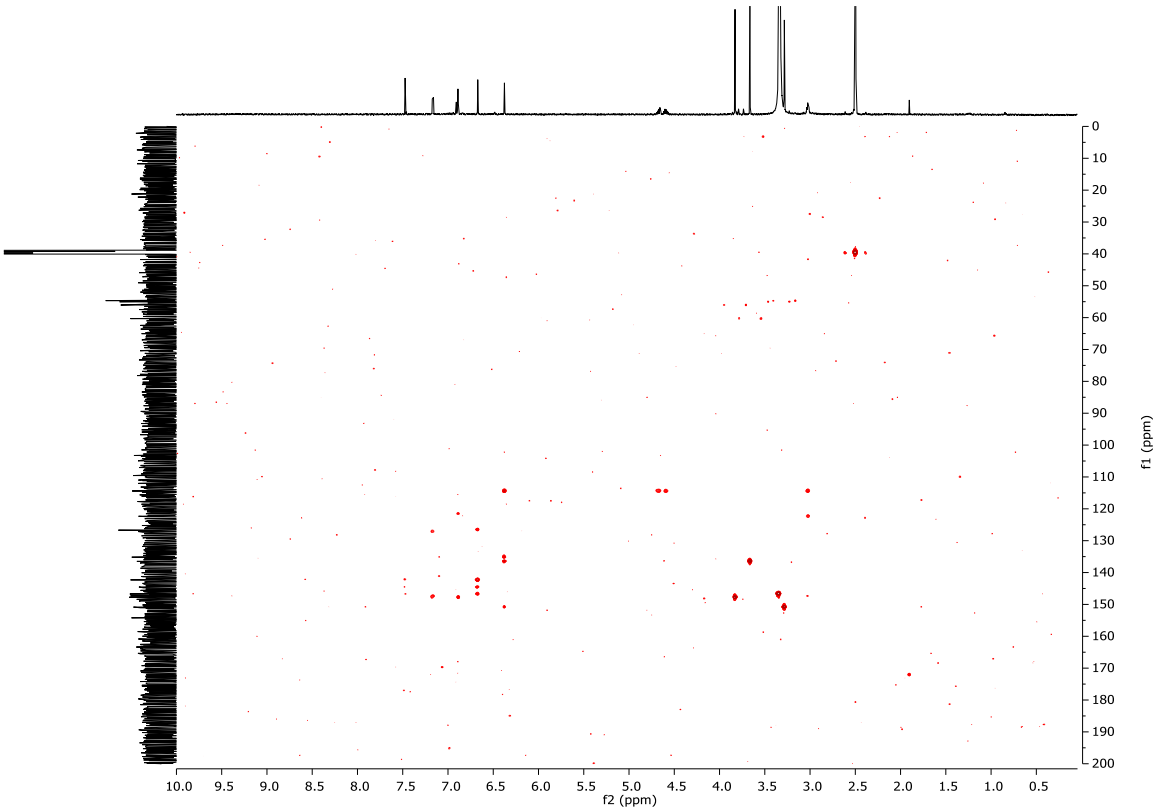
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) **148**



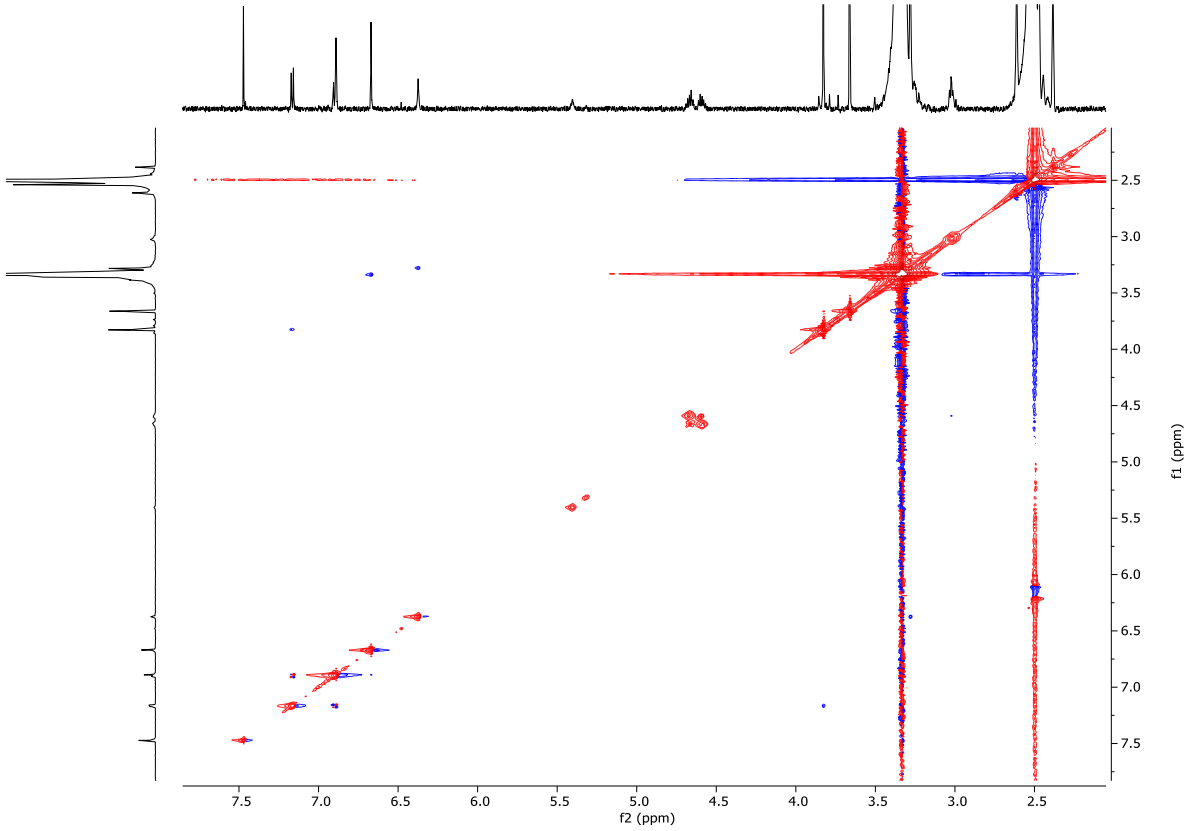
COSY Spectrum (600 MHz, DMSO-d_6) of **148**



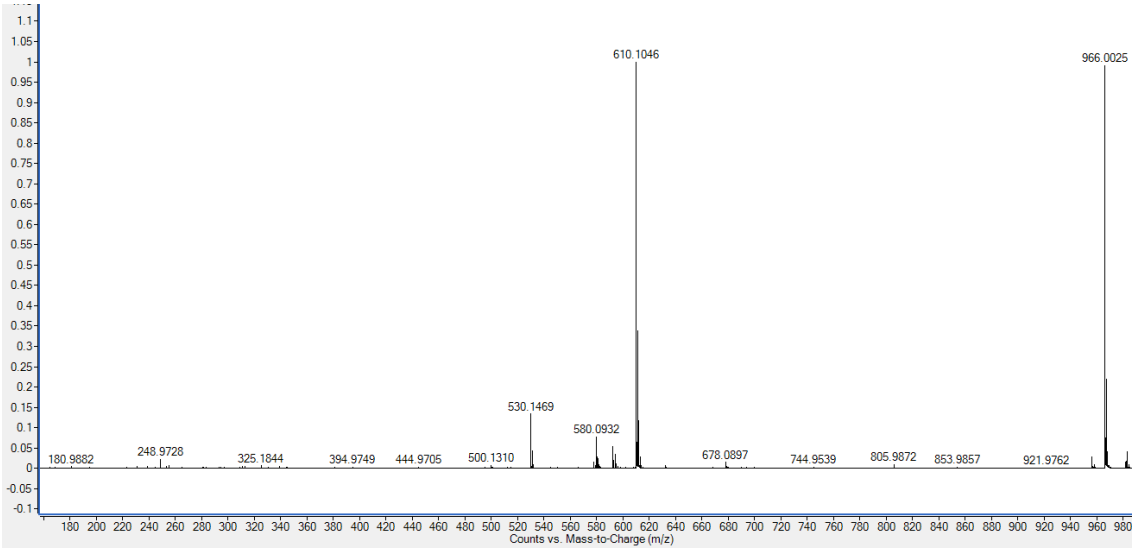
HSQC Spectrum (600 MHz, DMSO-d_6) of **148**



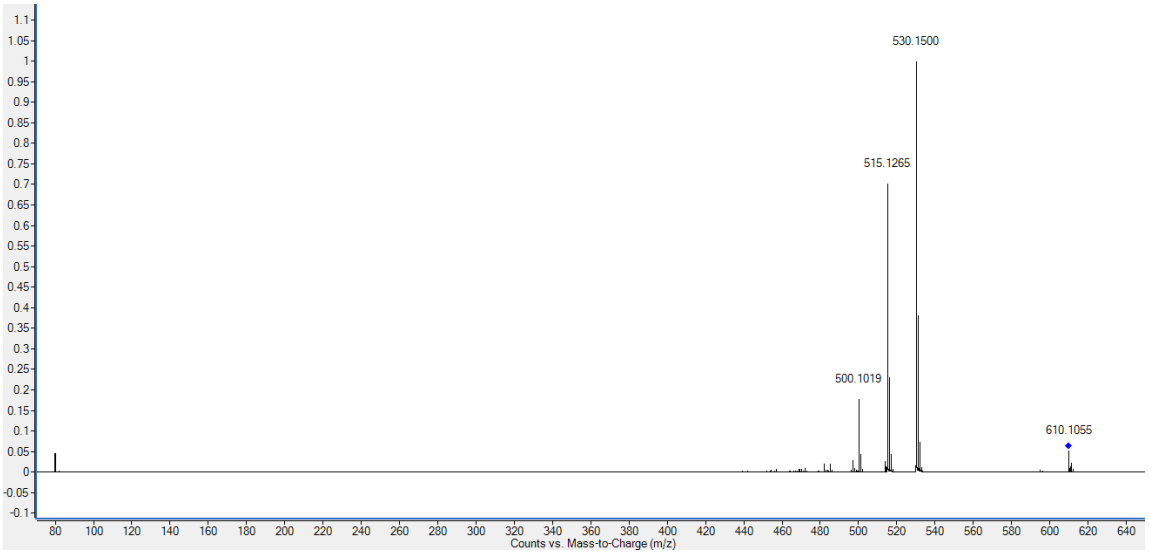
HMBC Spectrum (600 MHz, DMSO-d_6) of **148**



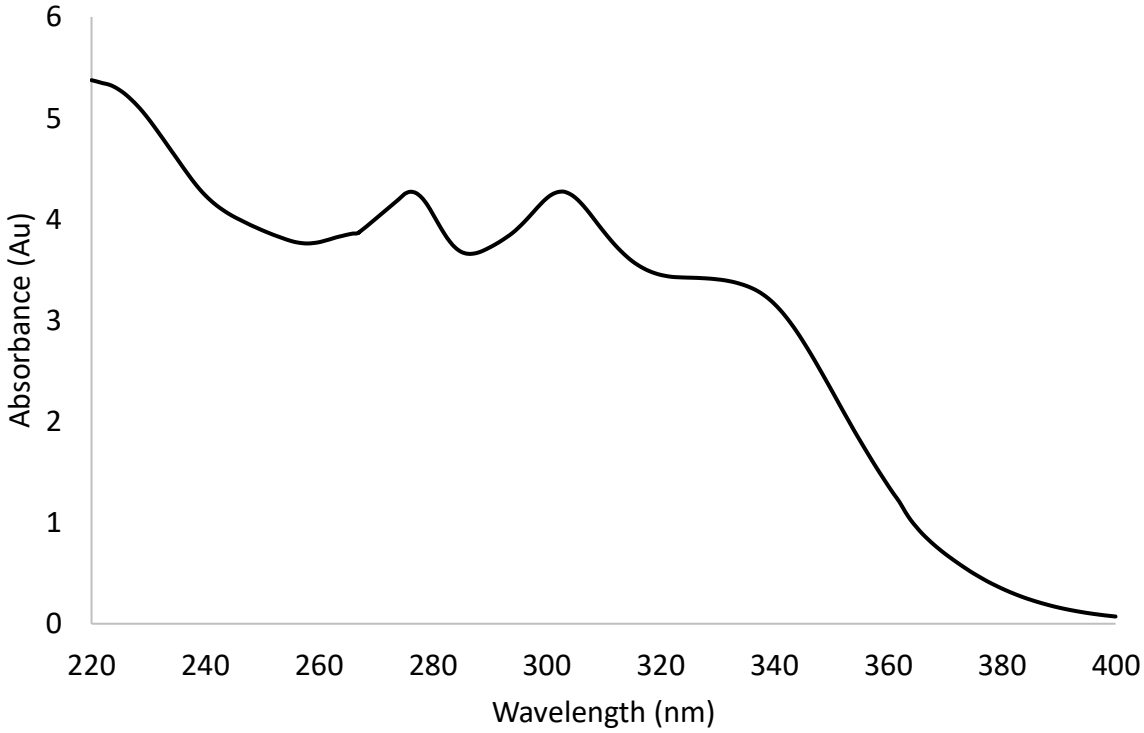
ROESY Spectrum (600 MHz, DMSO-d_6) of **148**



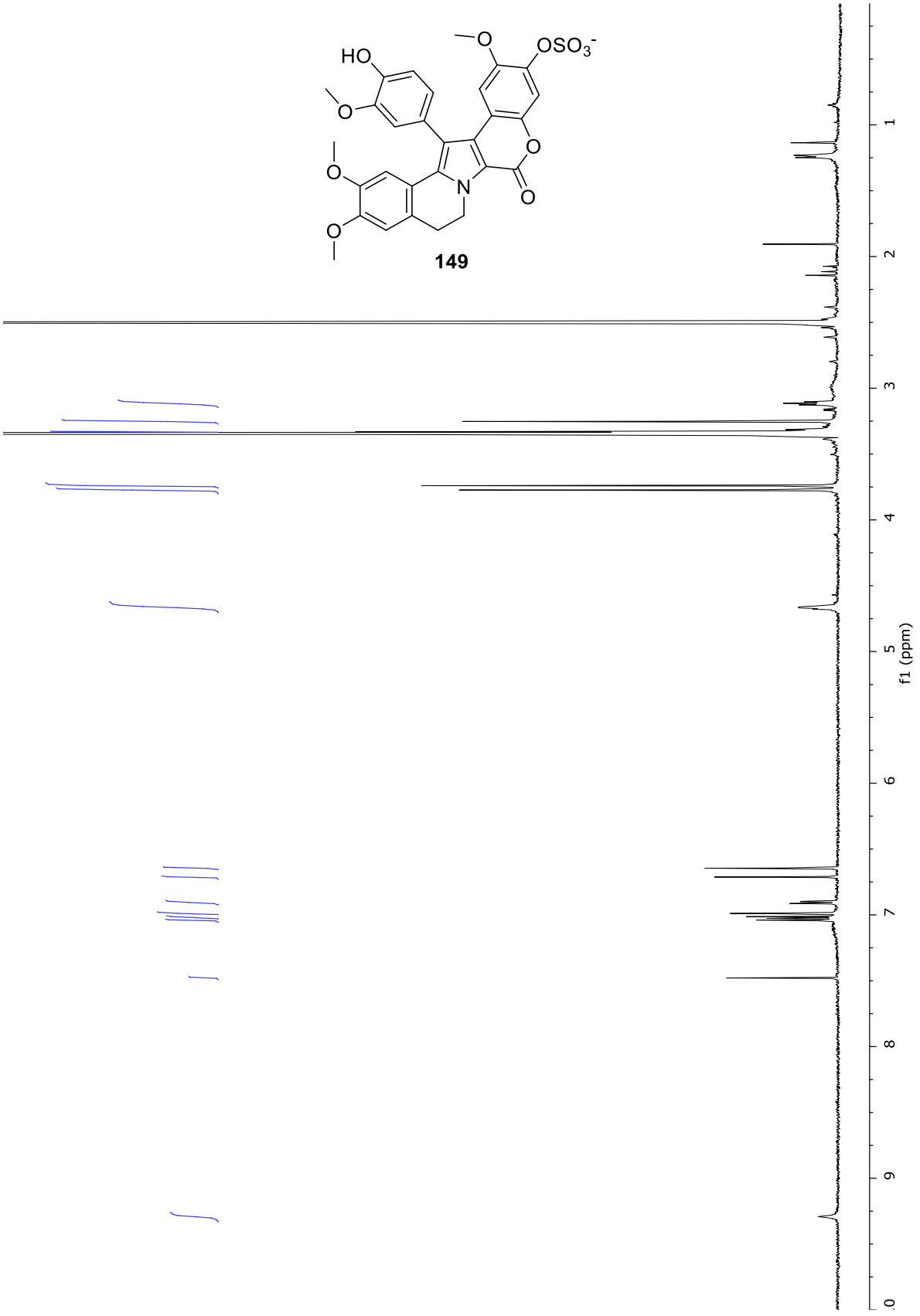
(–)-HRESIMS Spectrum of **148**



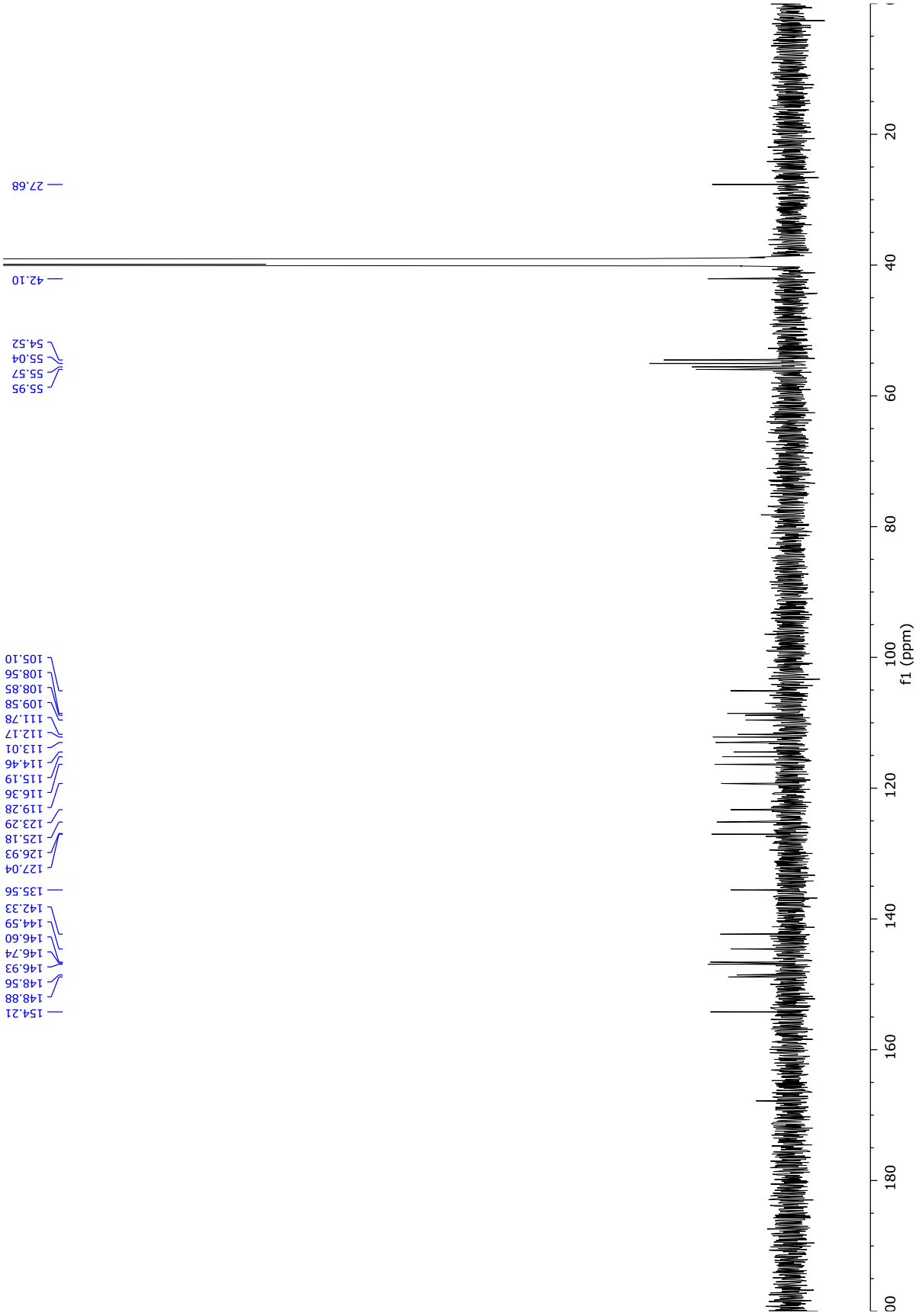
(–)-HRESIMS/MS Spectrum of **148**



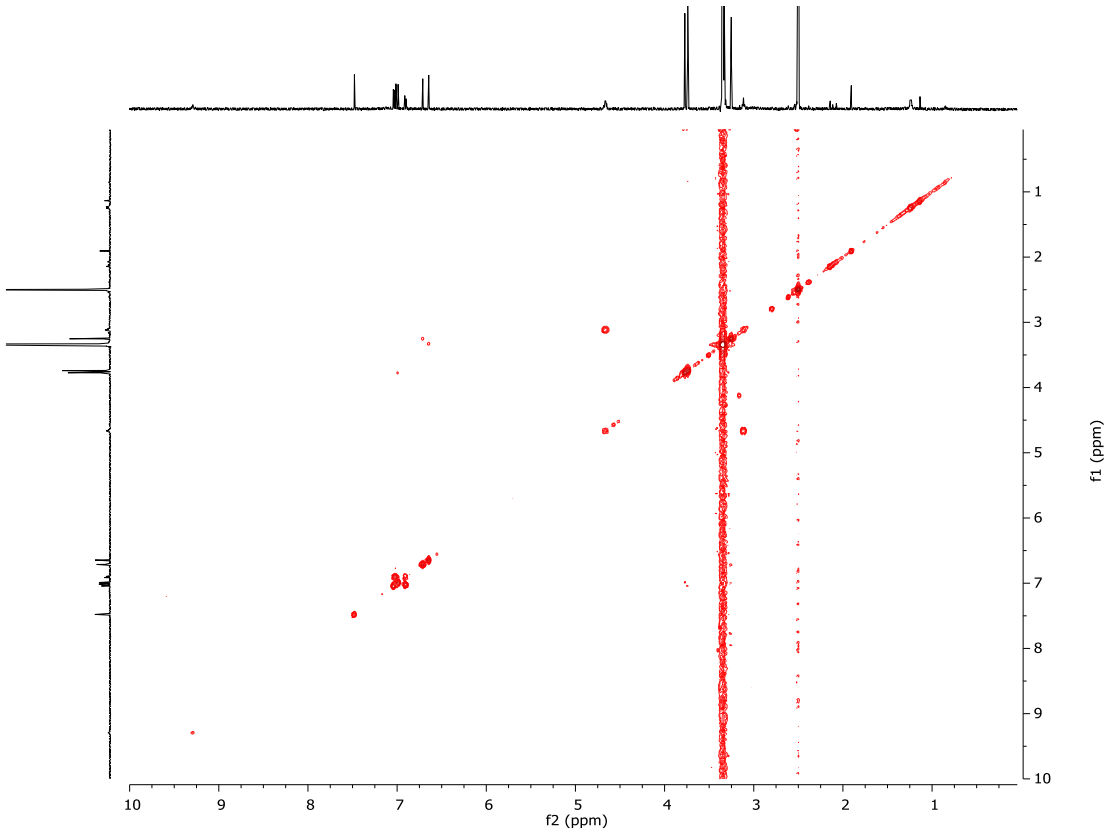
UV/vis Spectrum of **148**



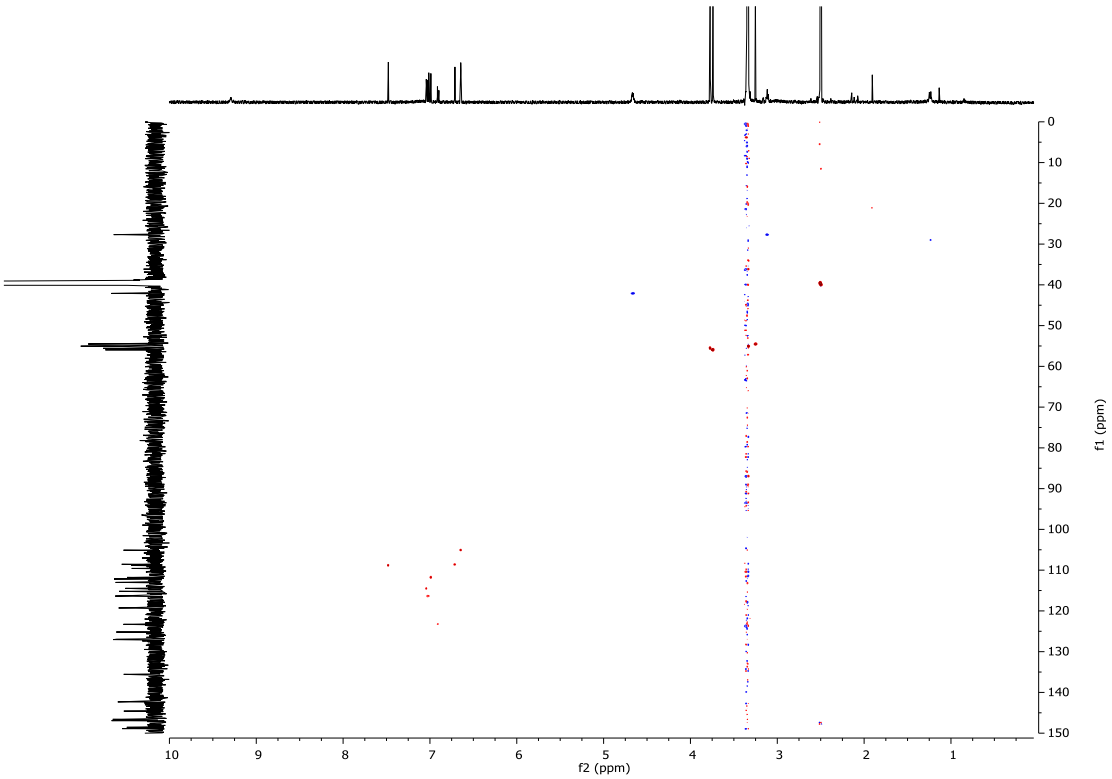
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **149**



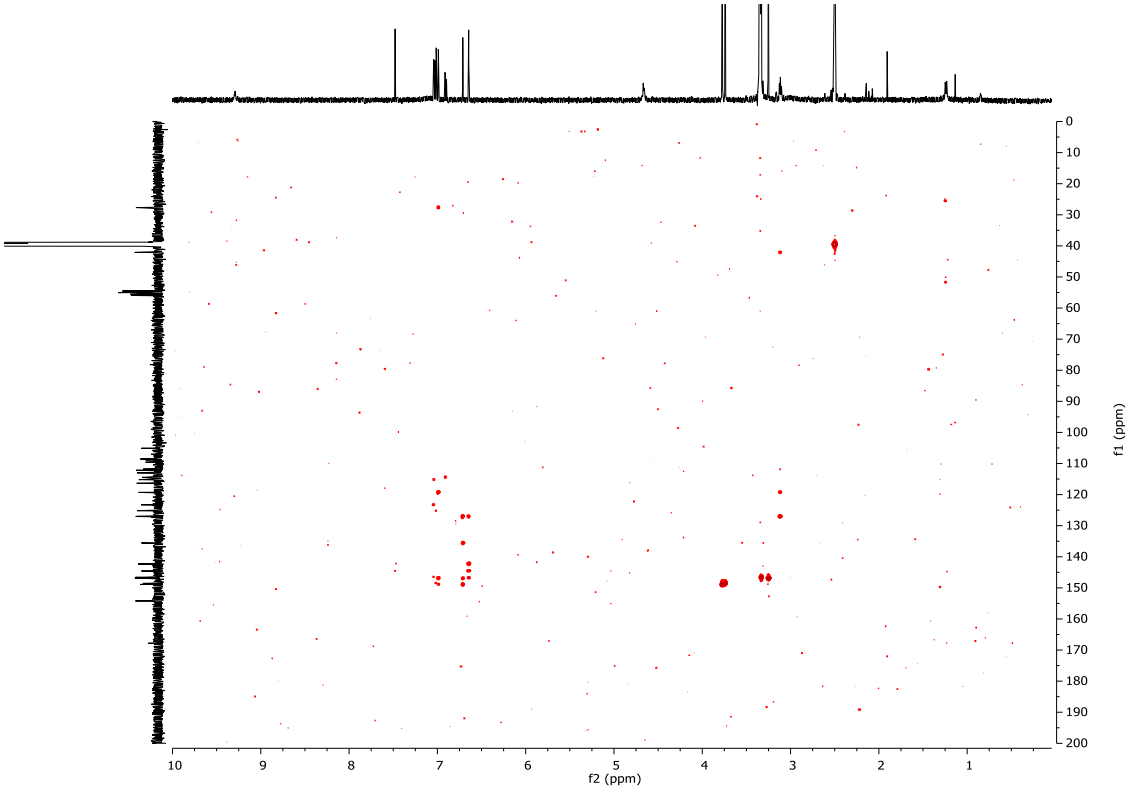
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **149**



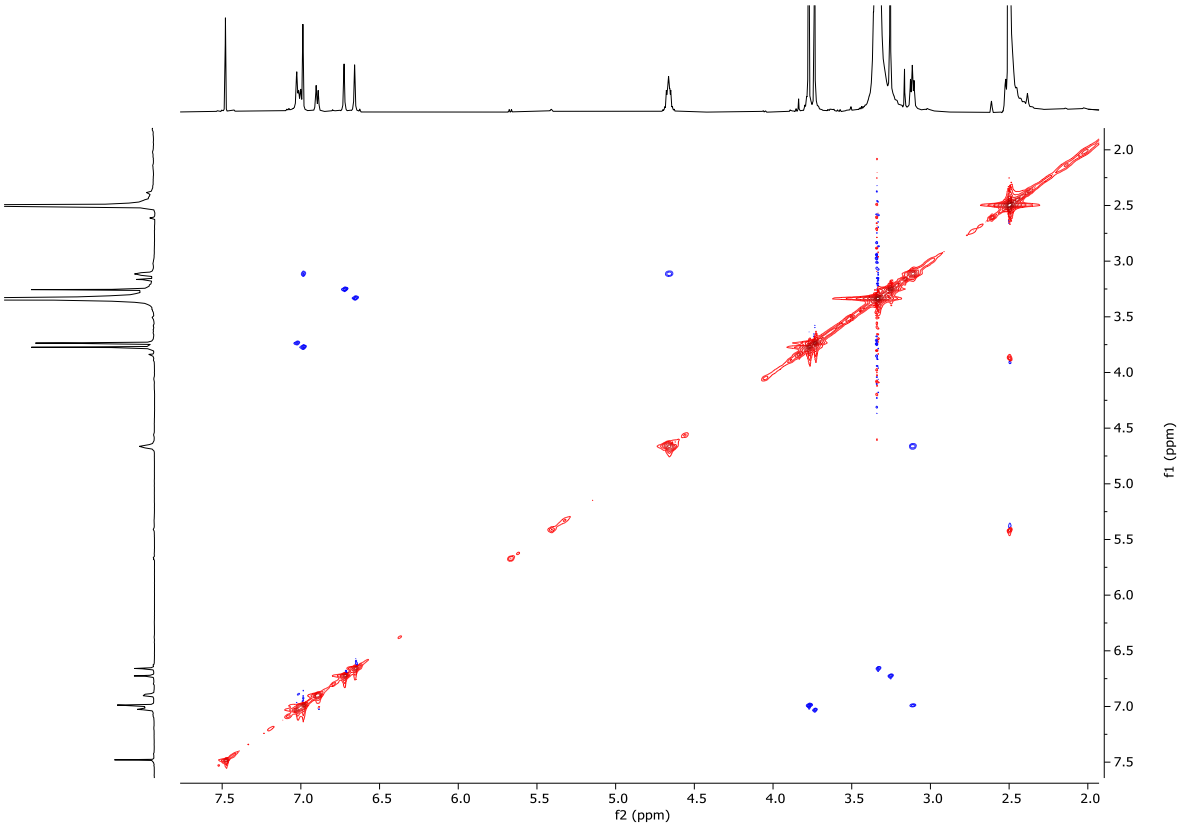
COSY Spectrum (600 MHz, DMSO-d_6) of **149**



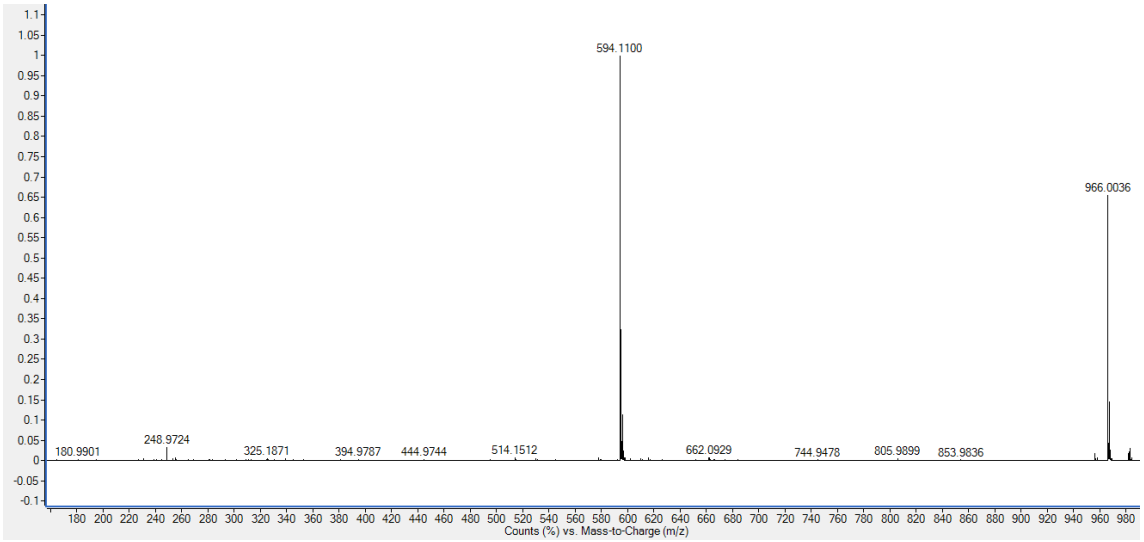
HSQC Spectrum (600 MHz, DMSO-d_6) of **149**



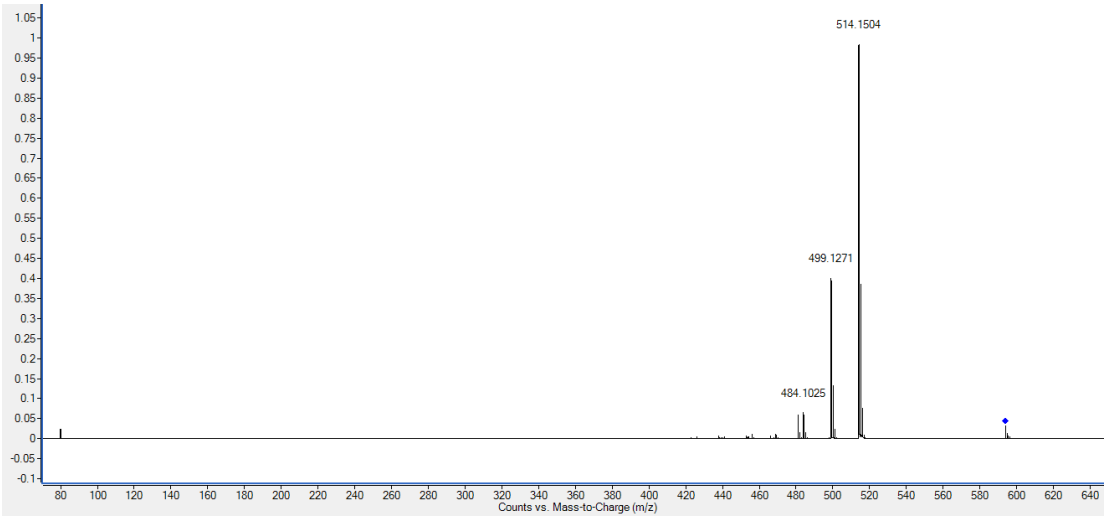
HMBC Spectrum (600 MHz, DMSO-d_6) of **149**



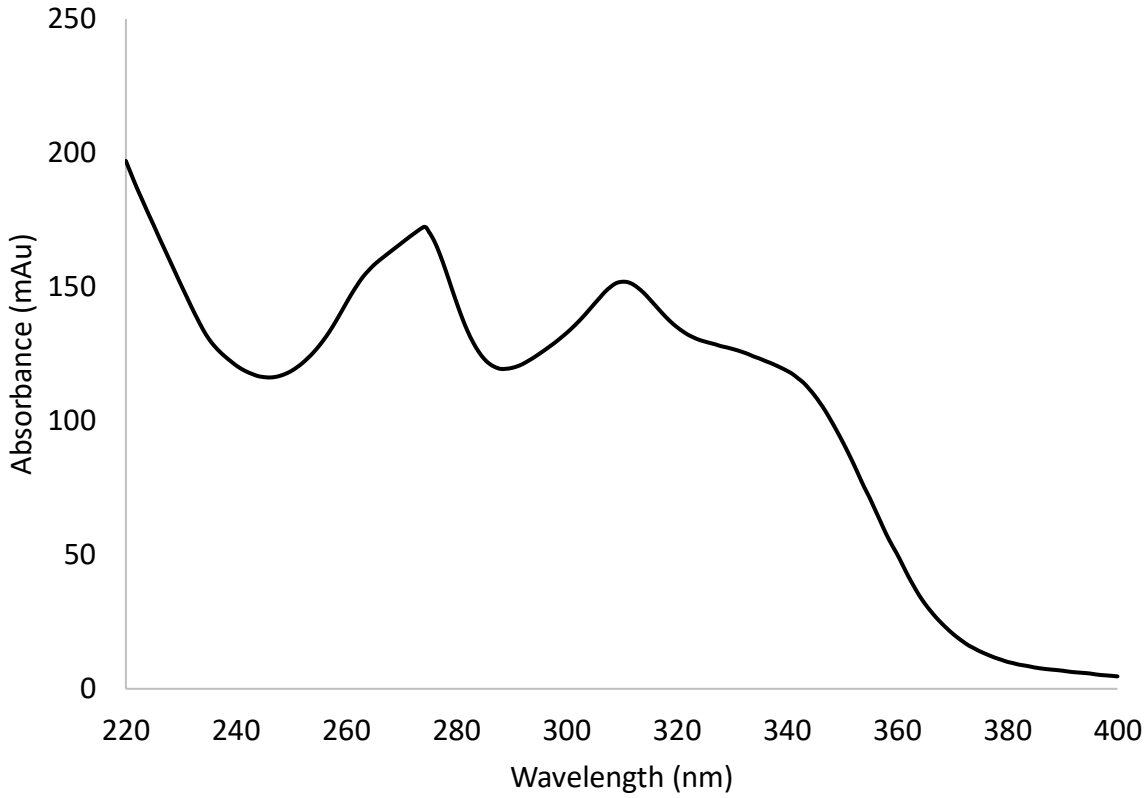
ROESY Spectrum (600 MHz, DMSO-d_6) of **149**



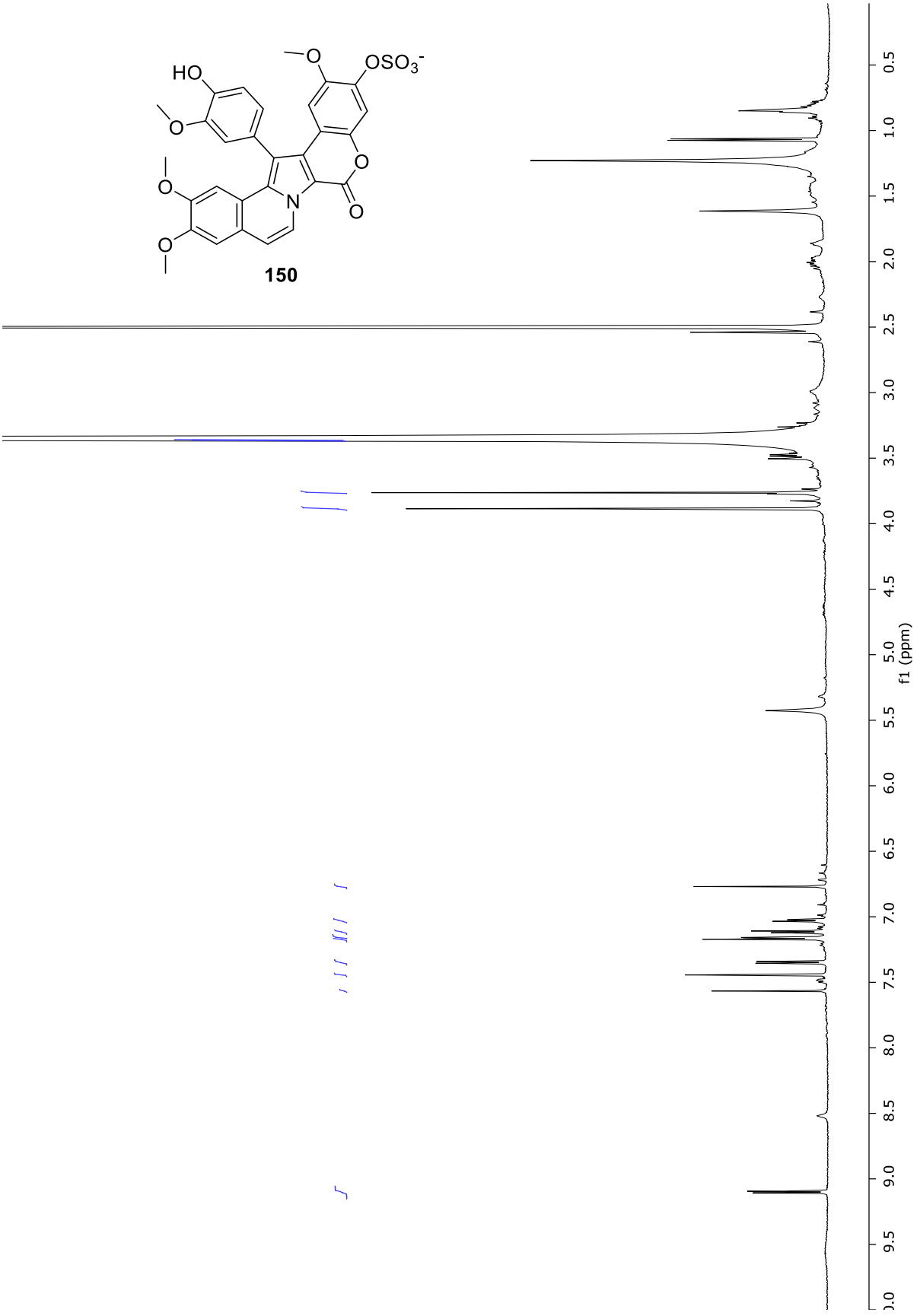
(–)-HRESIMS Spectrum of **149**



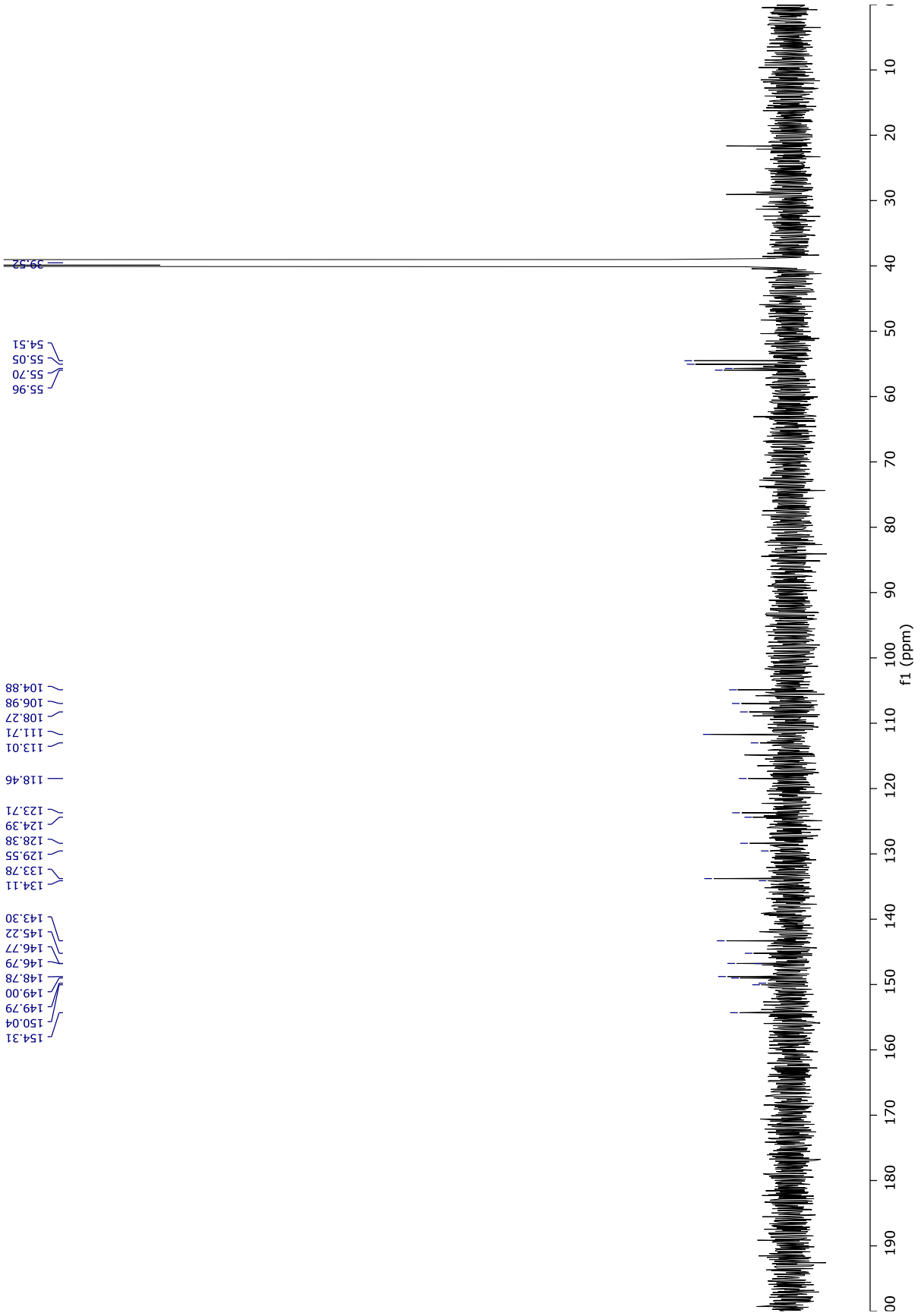
(–)-HRESIMS/MS Spectrum of **149**



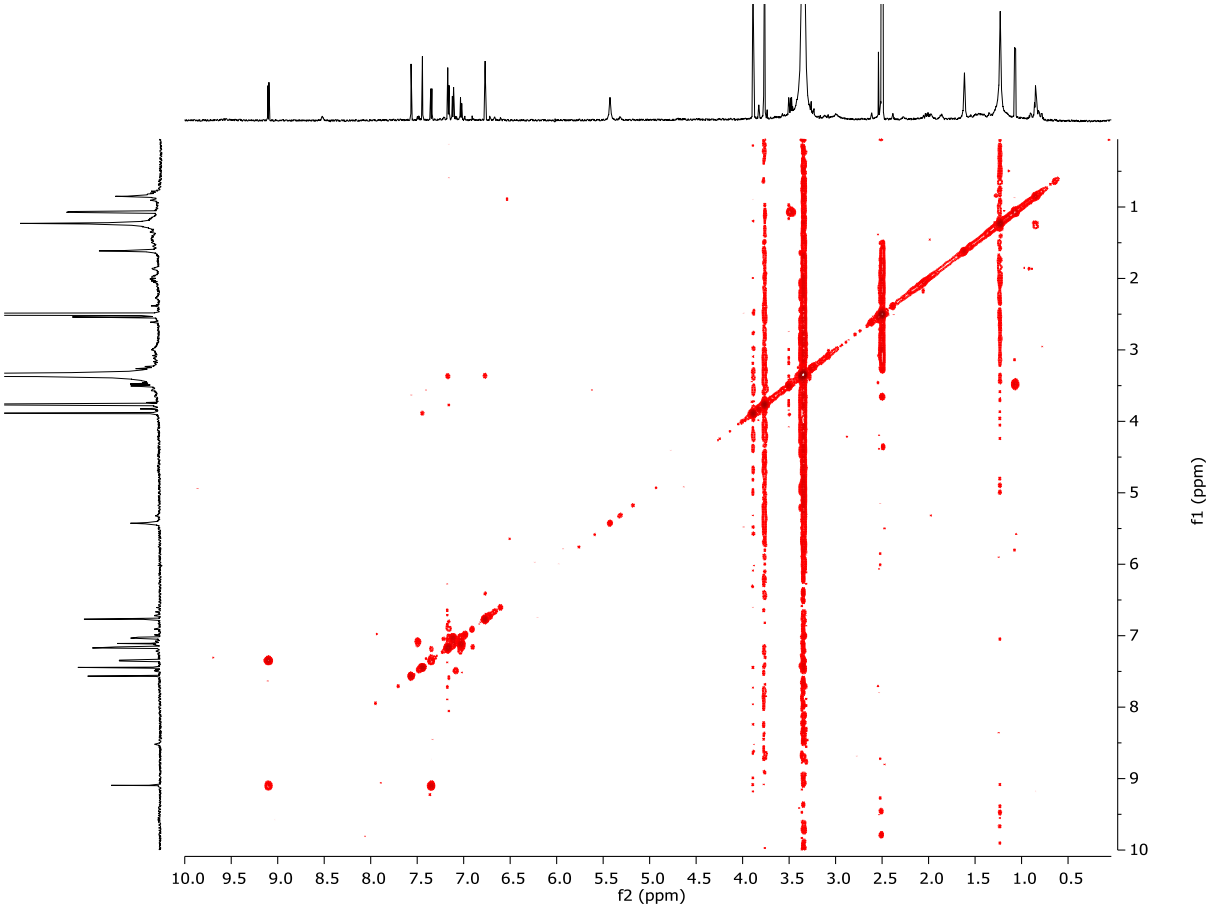
UV/vis Spectrum of **149**



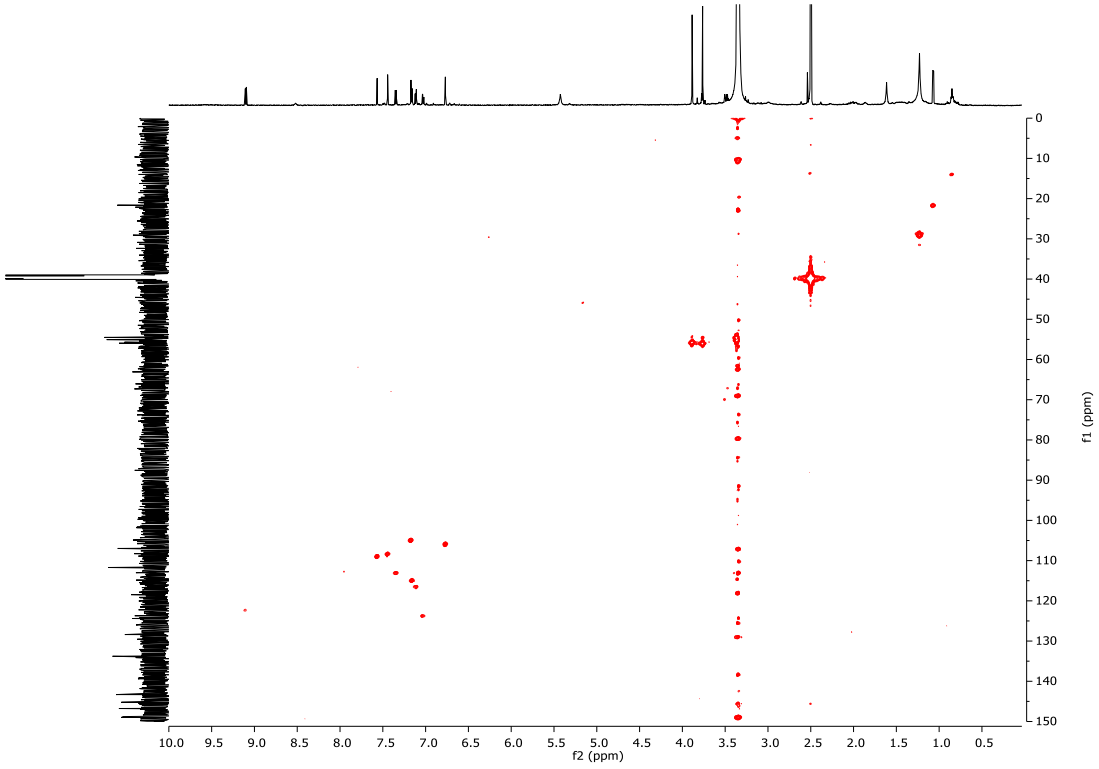
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **150**



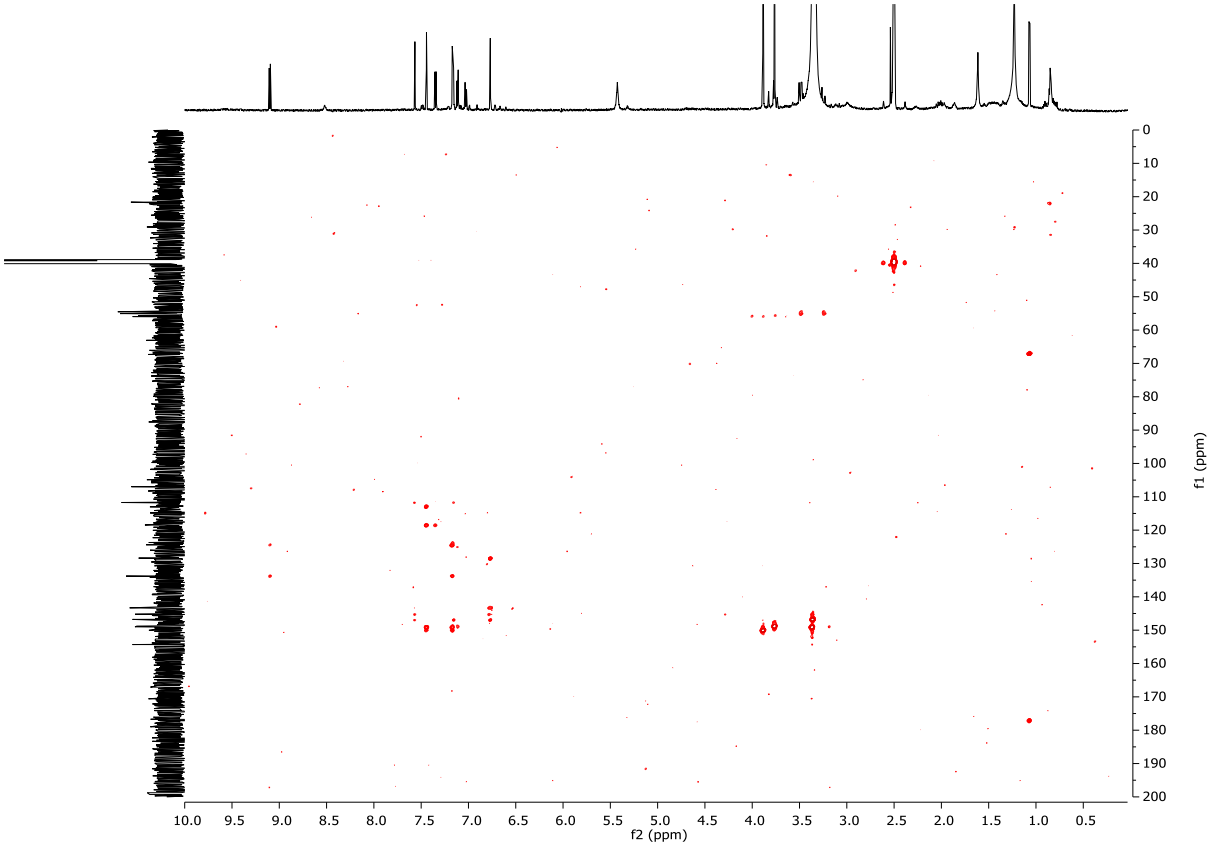
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **150**



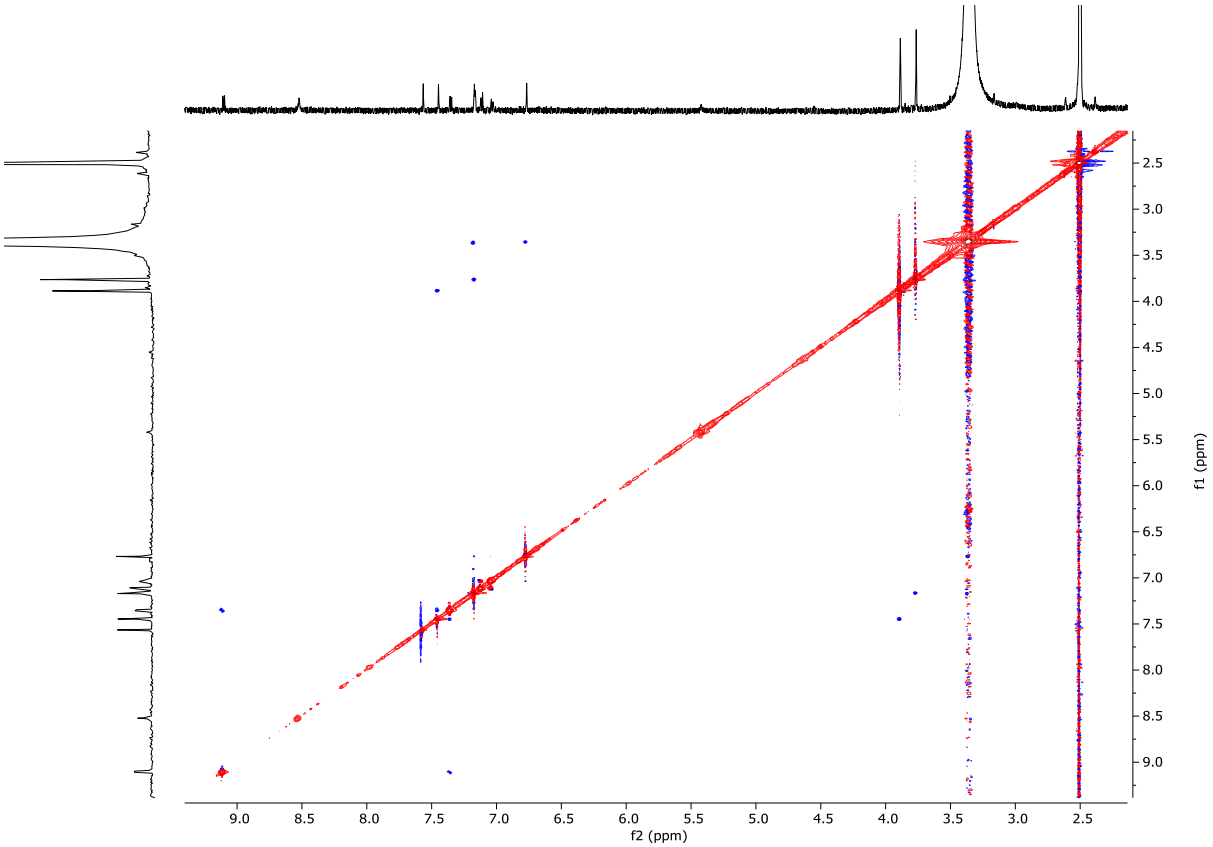
COSY Spectrum (600 MHz, DMSO-d_6) of **150**



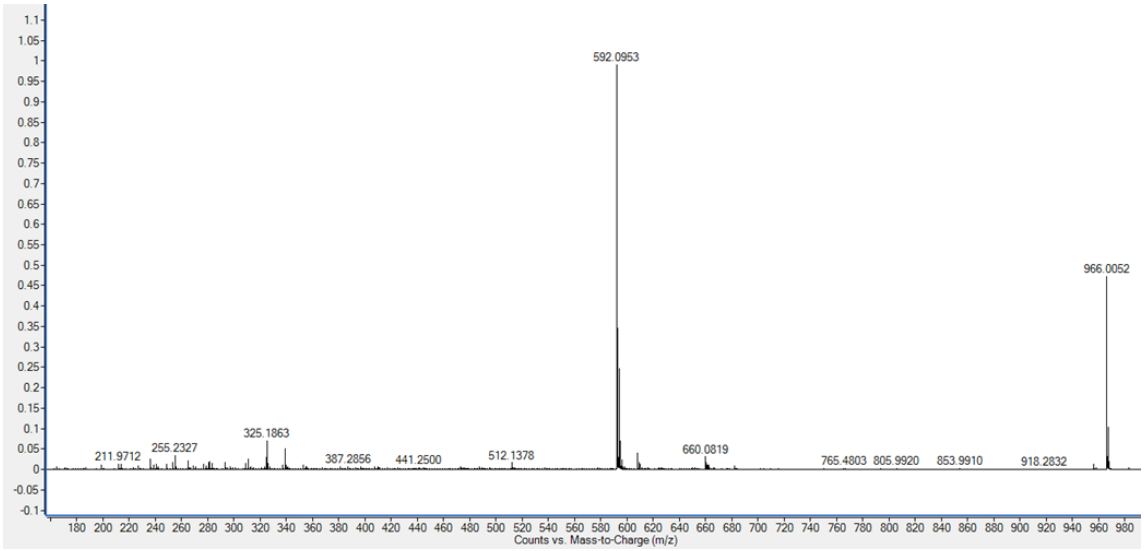
HSQC Spectrum (600 MHz, DMSO-d_6) of **150**



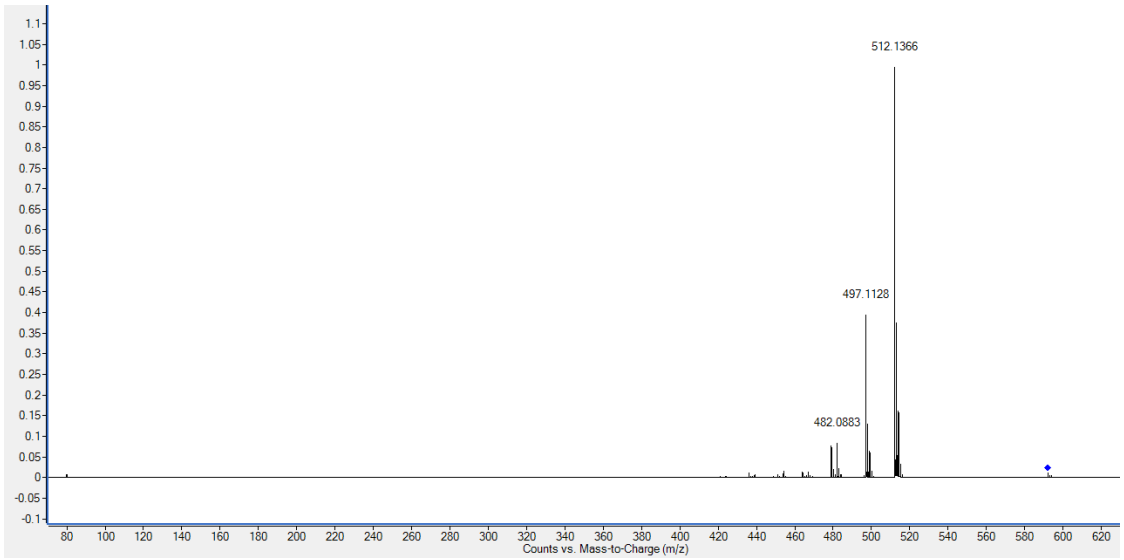
HMBC Spectrum (600 MHz, DMSO-d_6) of **150**



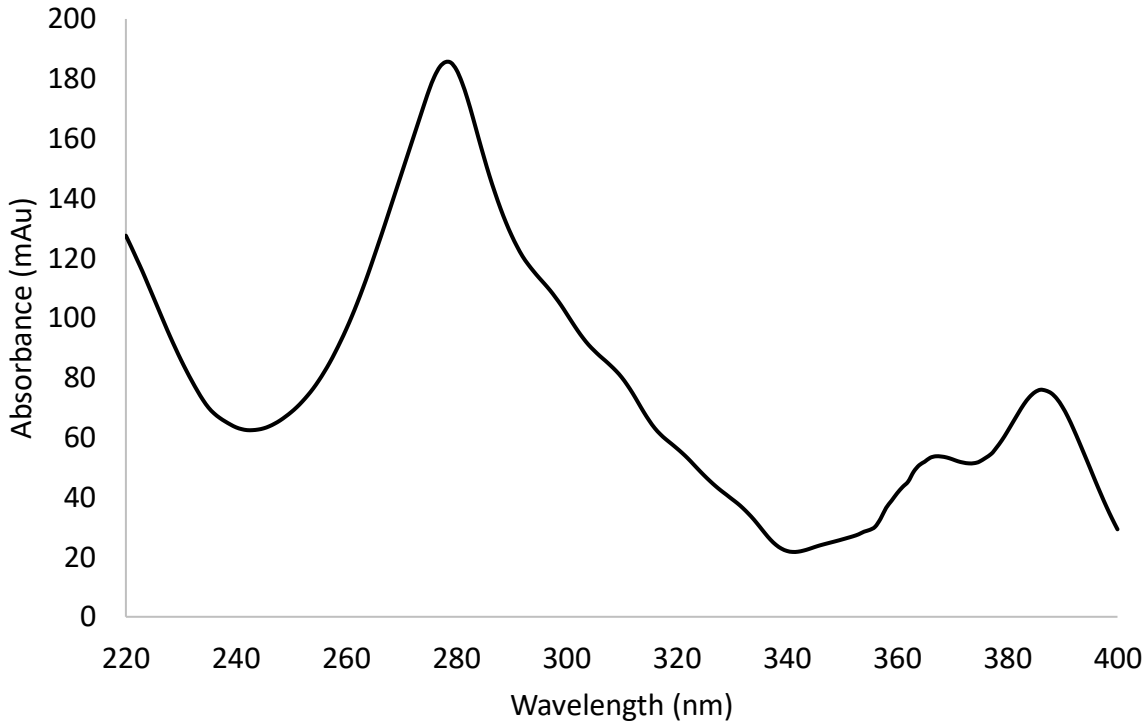
ROESY Spectrum (600 MHz, DMSO-d_6) of **150**



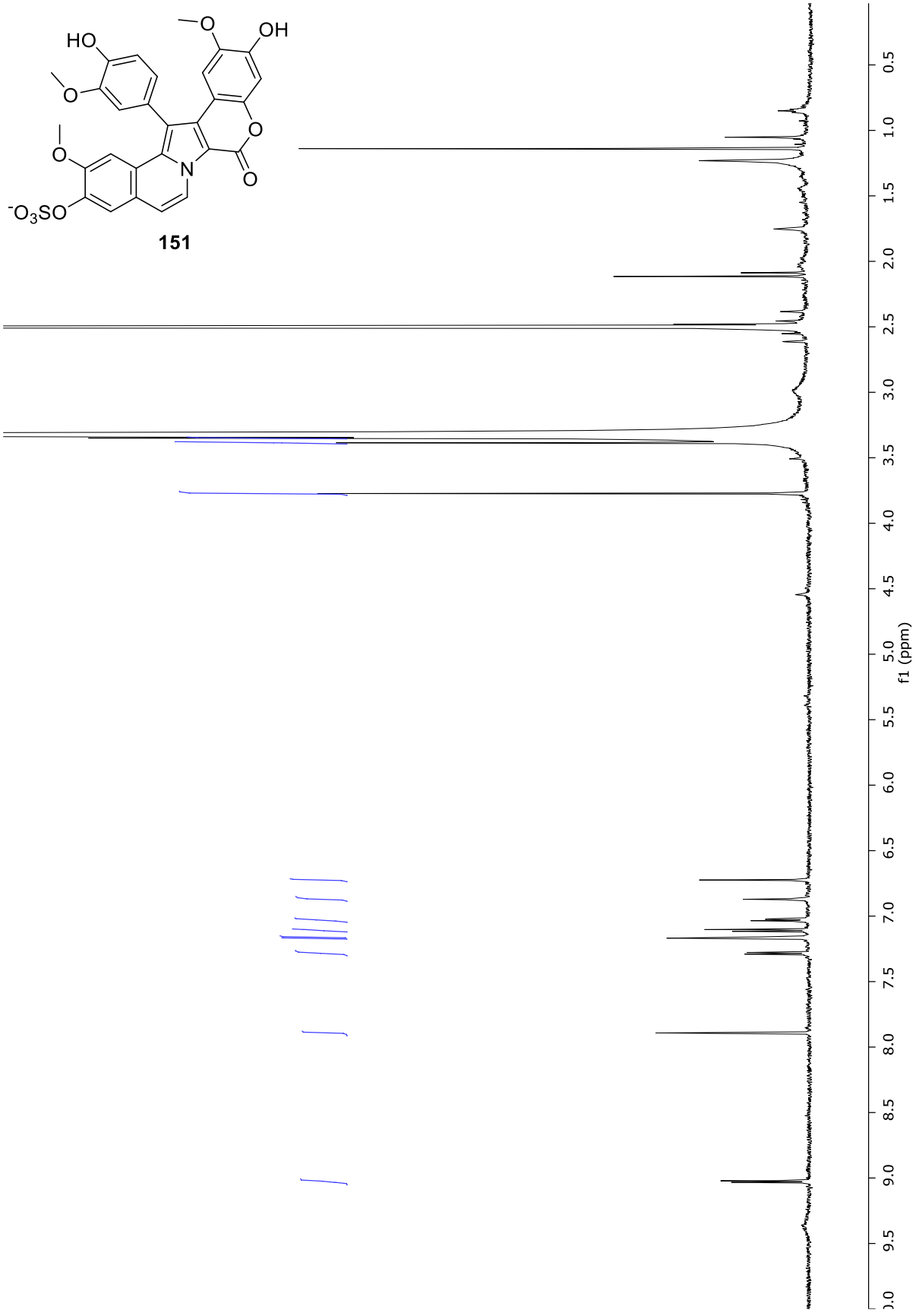
(–)-HRESIMS Spectrum of **150**



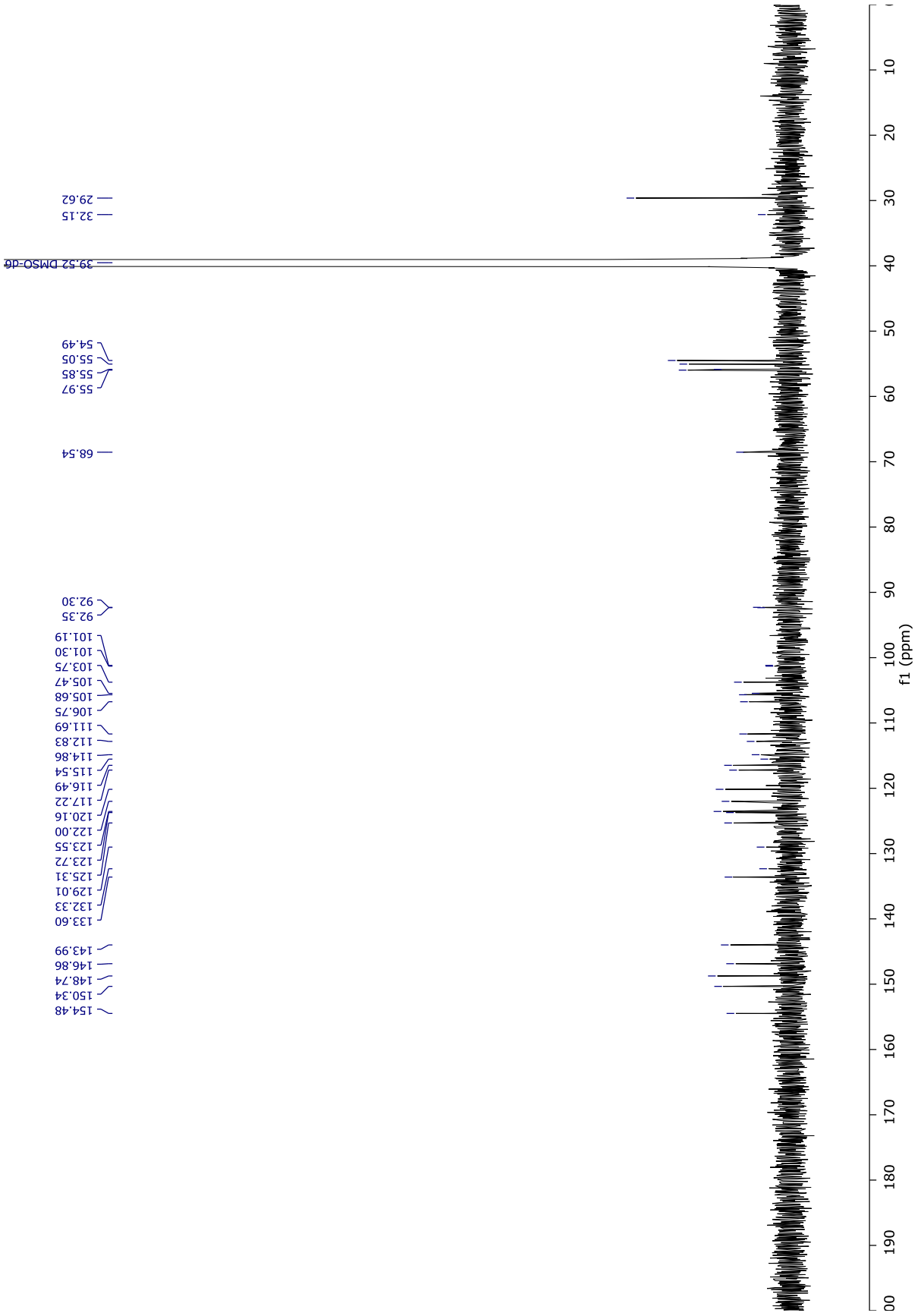
(–)-HRESIMS/MS Spectrum of **150**



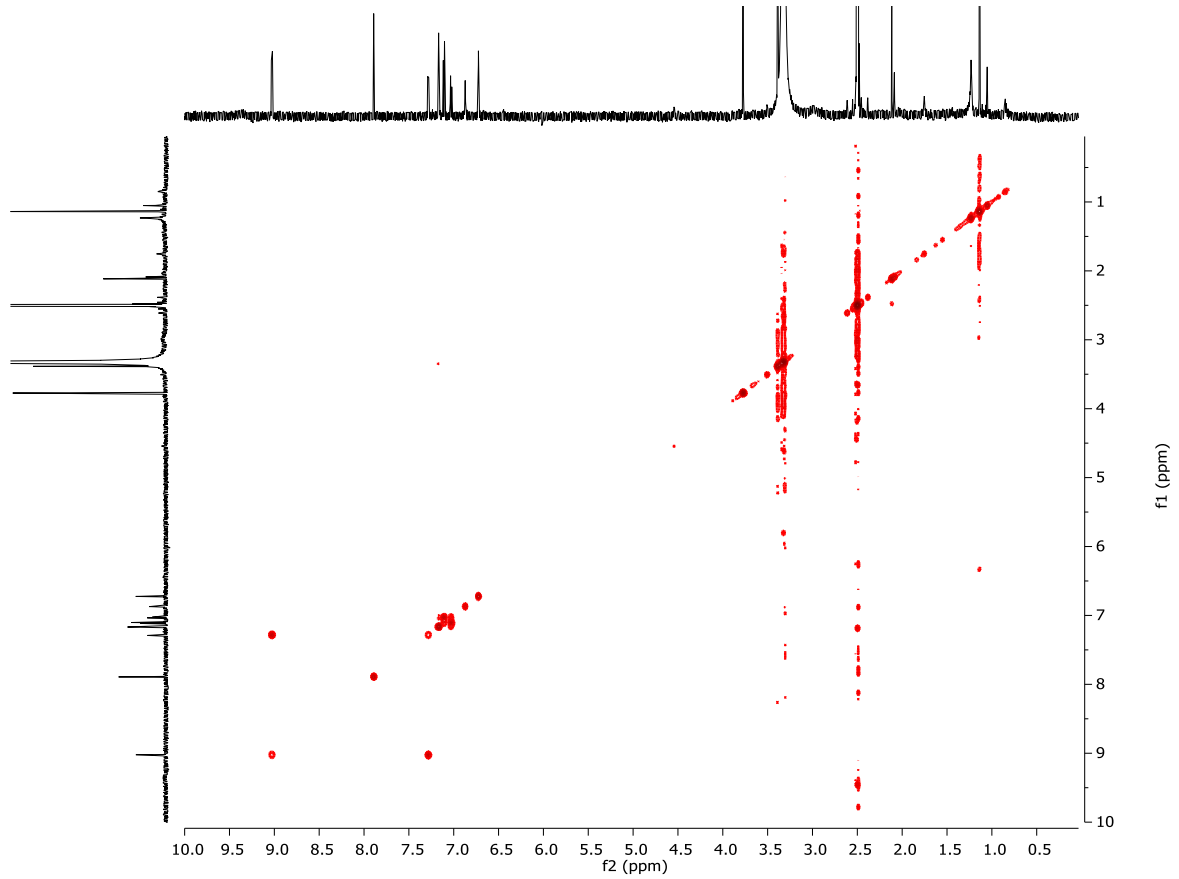
UV/vis Spectrum of **150**



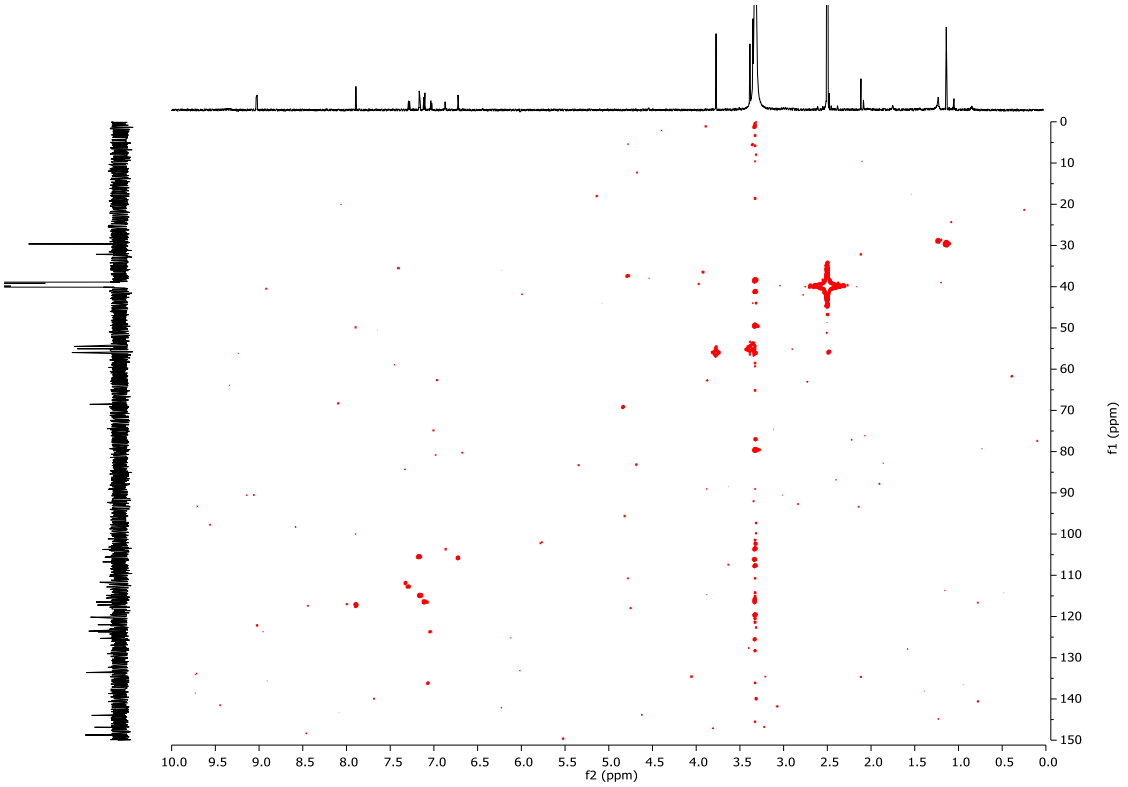
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **151**



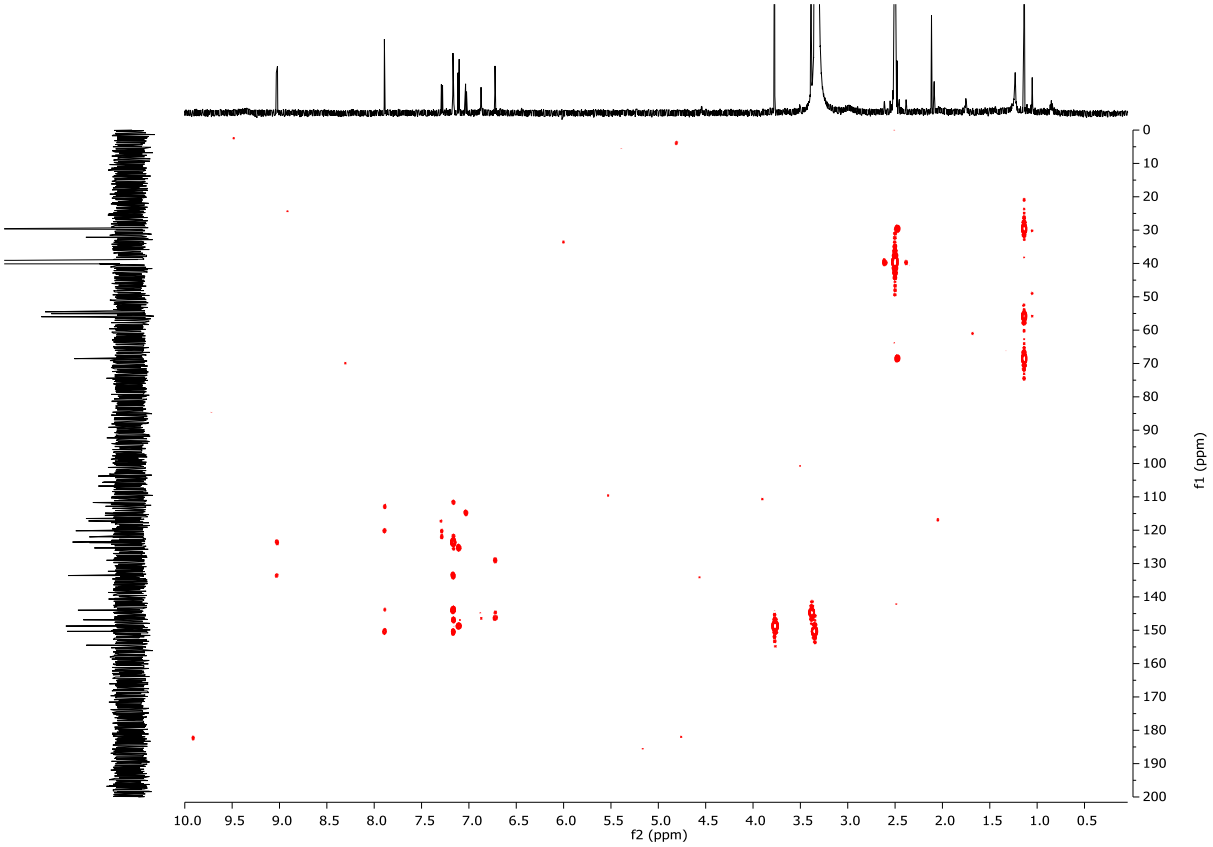
^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **151**



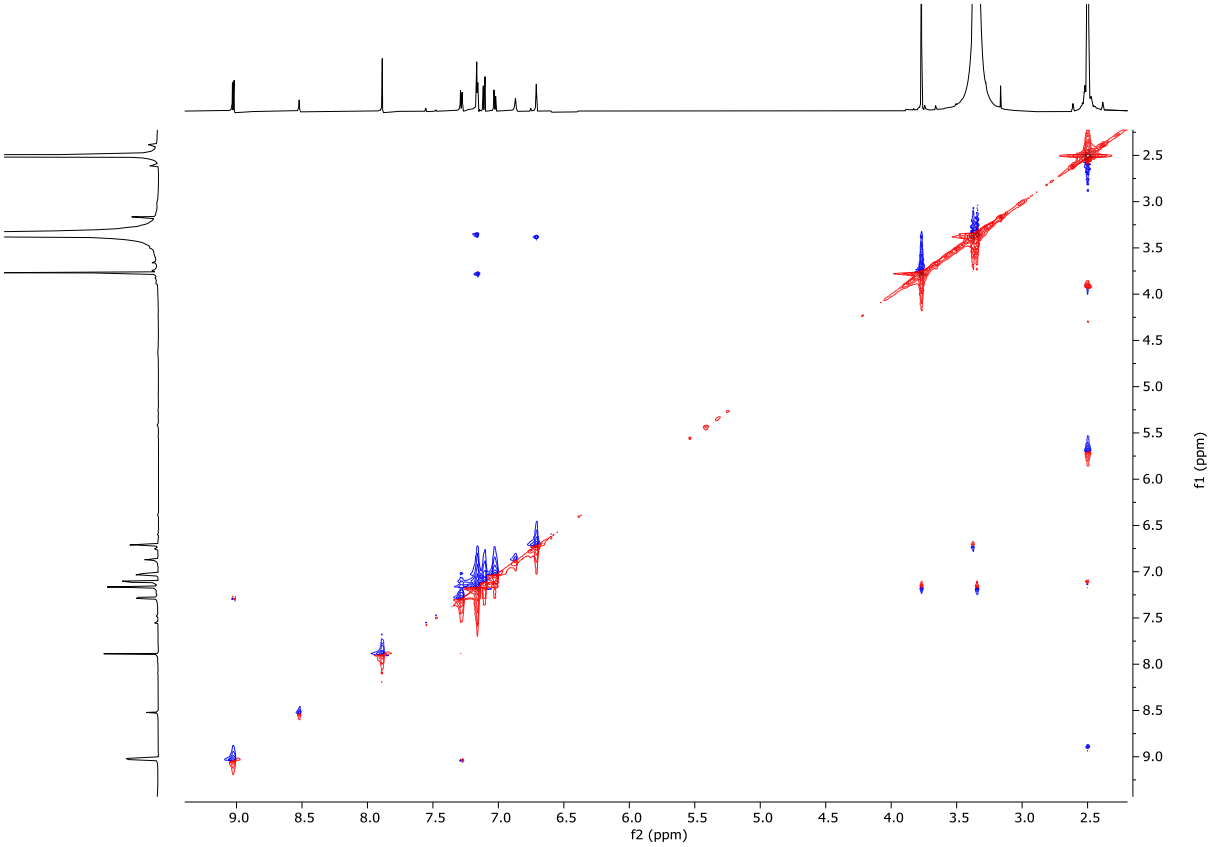
COSY Spectrum (600 MHz, DMSO-d_6) of **151**



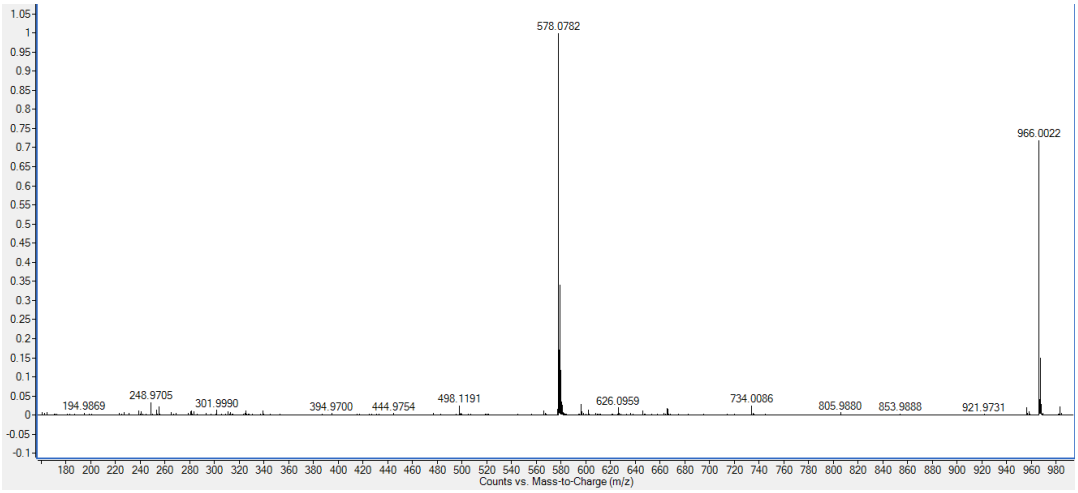
HSQC Spectrum (600 MHz, DMSO-d_6) of **151**



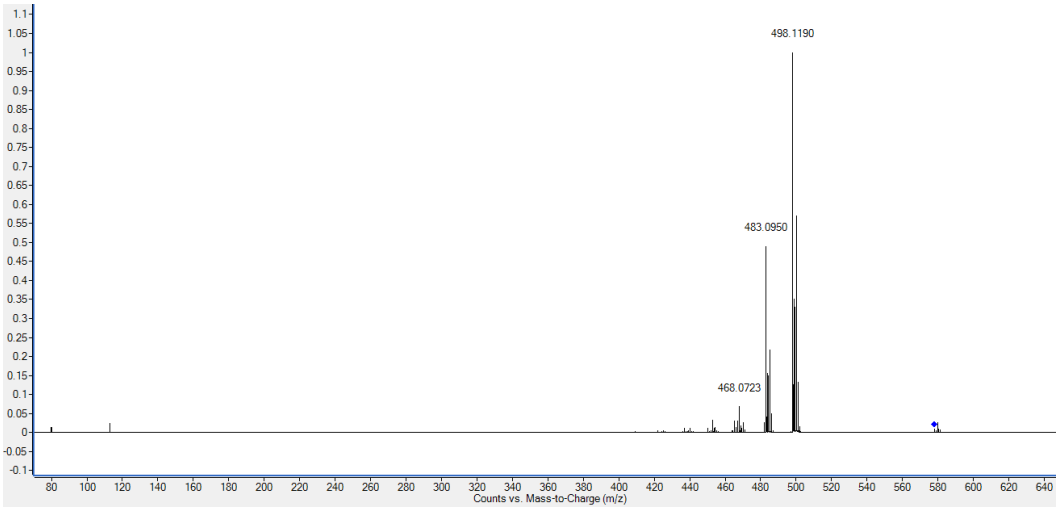
HMBC Spectrum (600 MHz, DMSO-d_6) of **151**



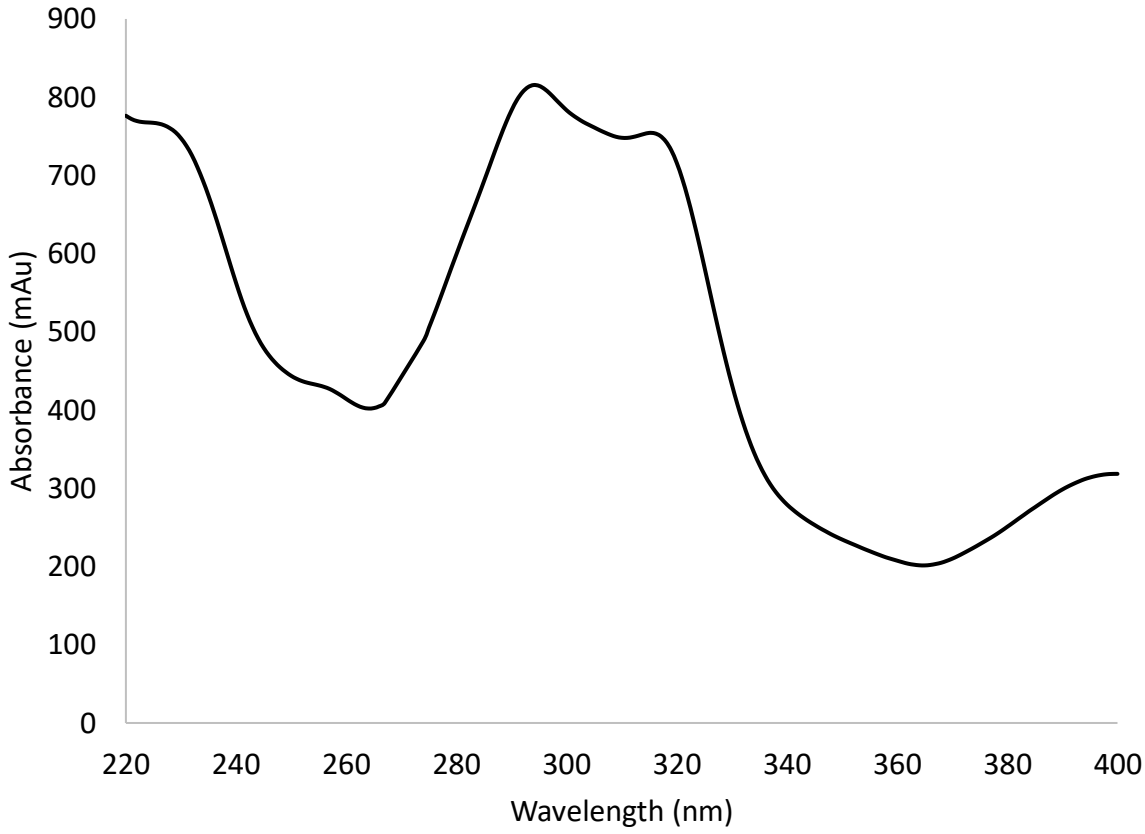
ROESY Spectrum (600 MHz, DMSO-d_6) of **151**



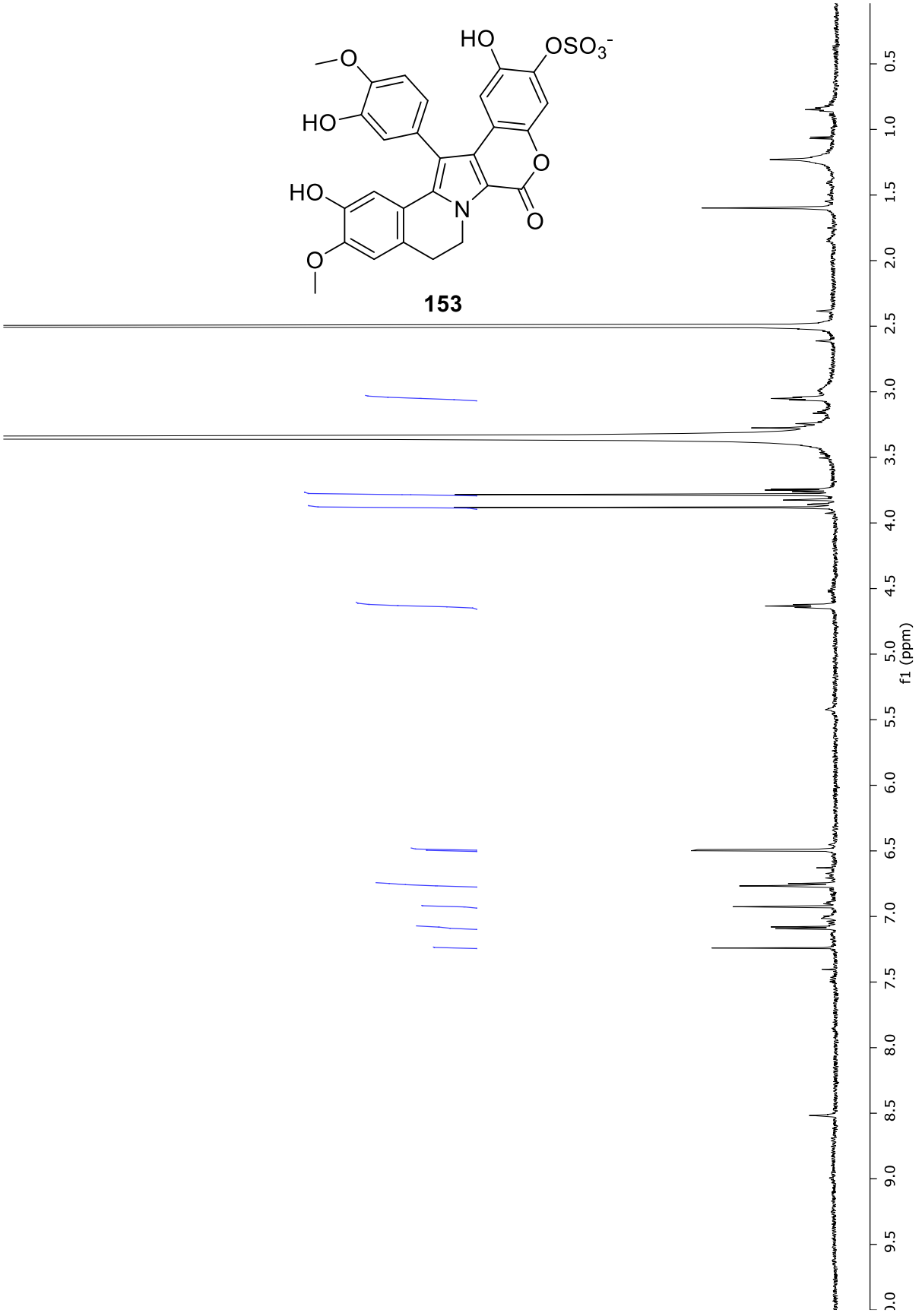
(–)-HRESIMS Spectrum of **151**



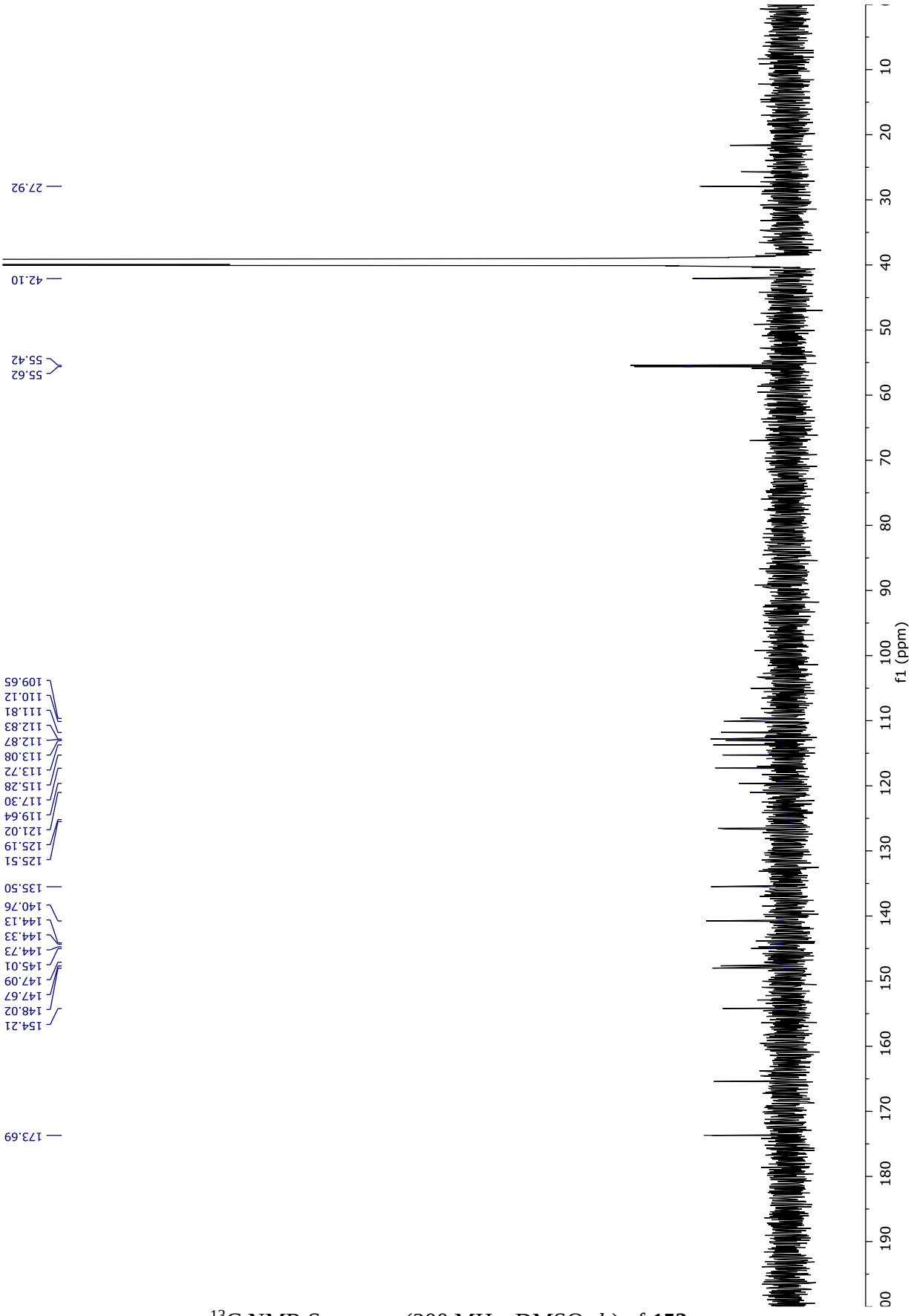
(–)-HRESIMS/MS Spectrum of **151**



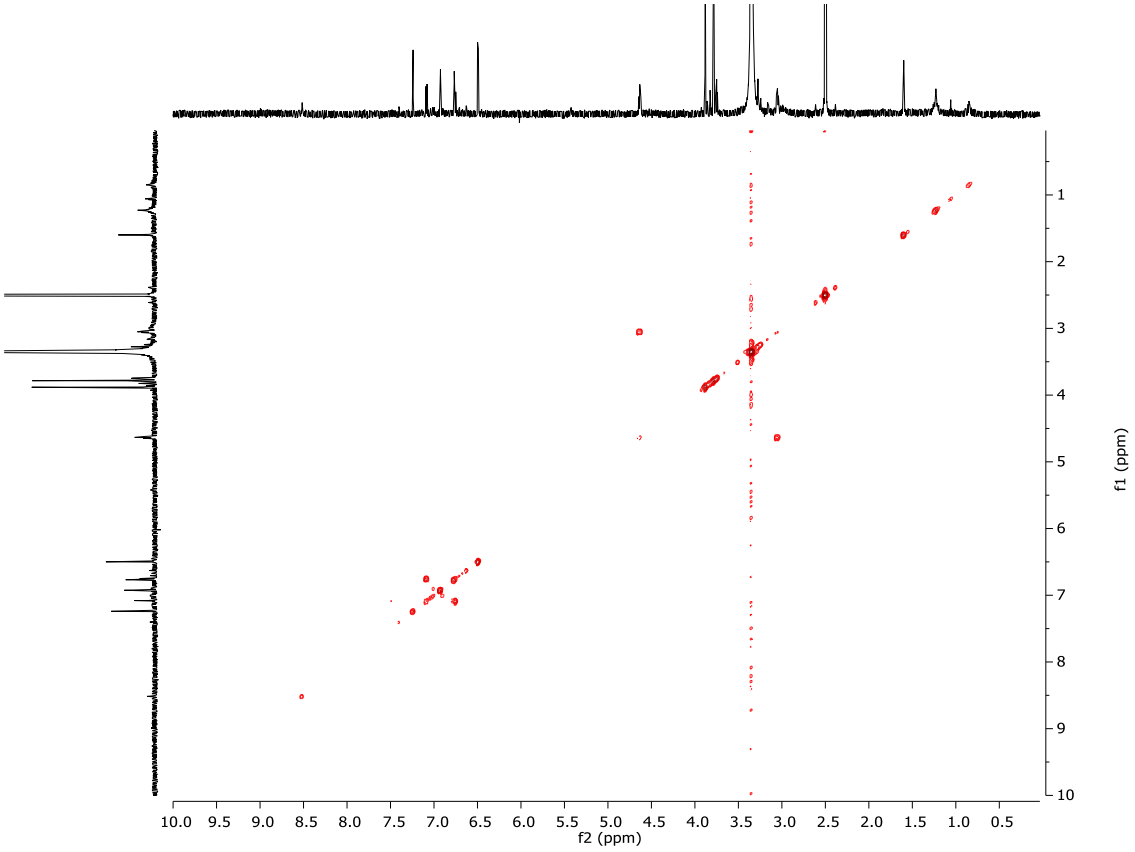
UV/vis Spectrum of **151**



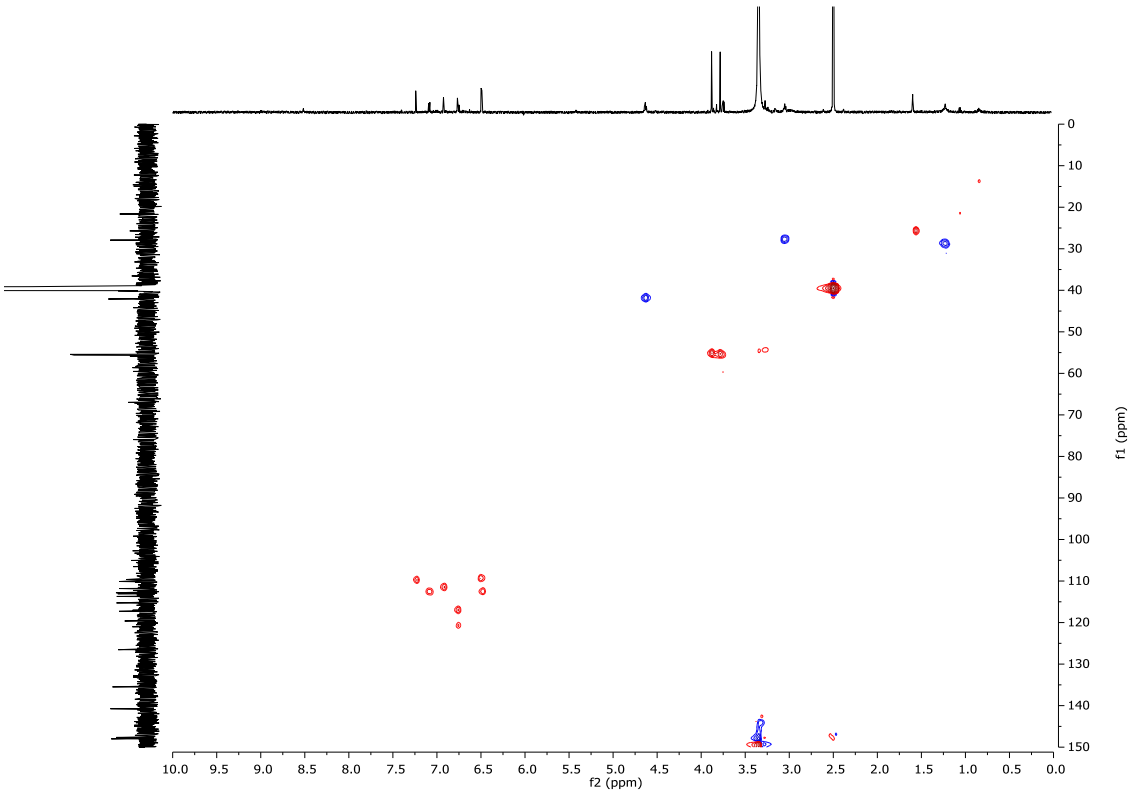
^{1}H NMR Spectrum (800 MHz, DMSO-d_6) of **153**



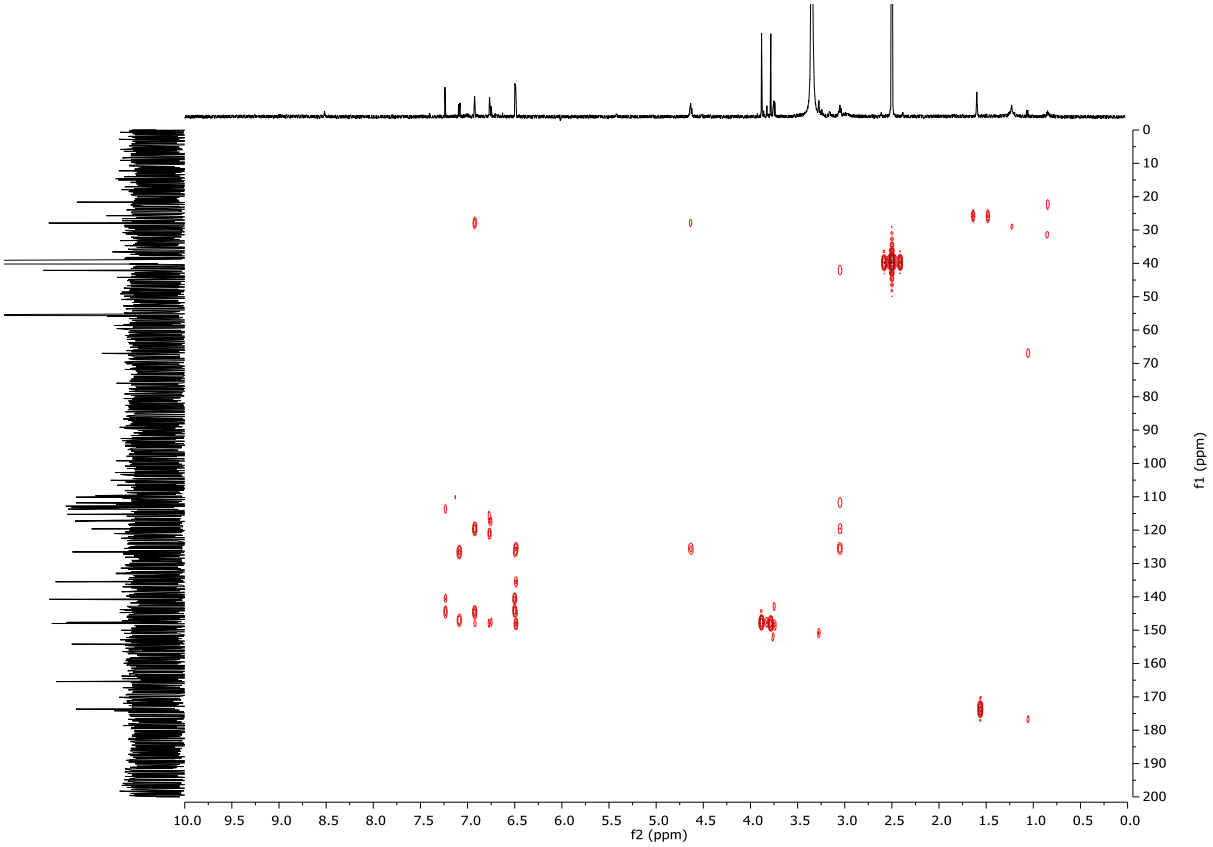
^{13}C NMR Spectrum (200 MHz, DMSO-d_6) of **153**



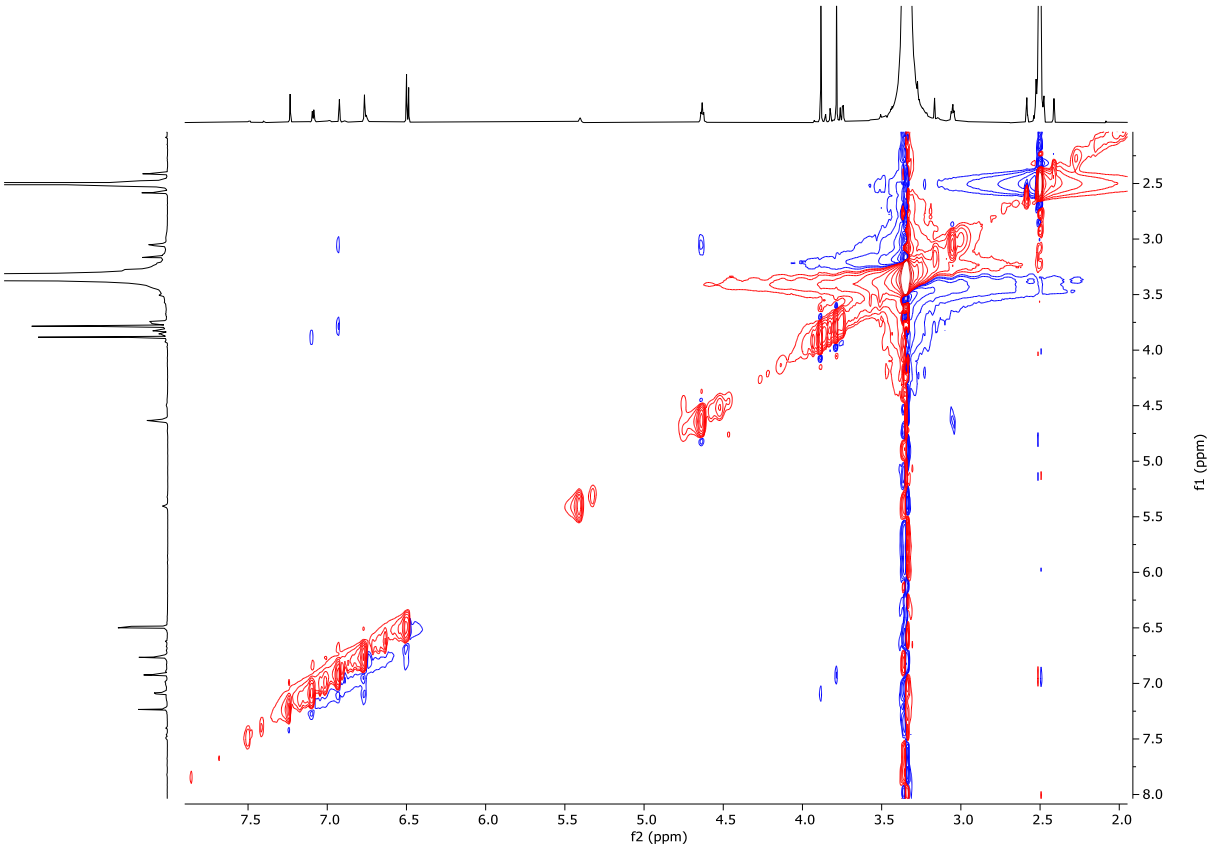
COSY Spectrum (800 MHz, DMSO-d_6) of **153**



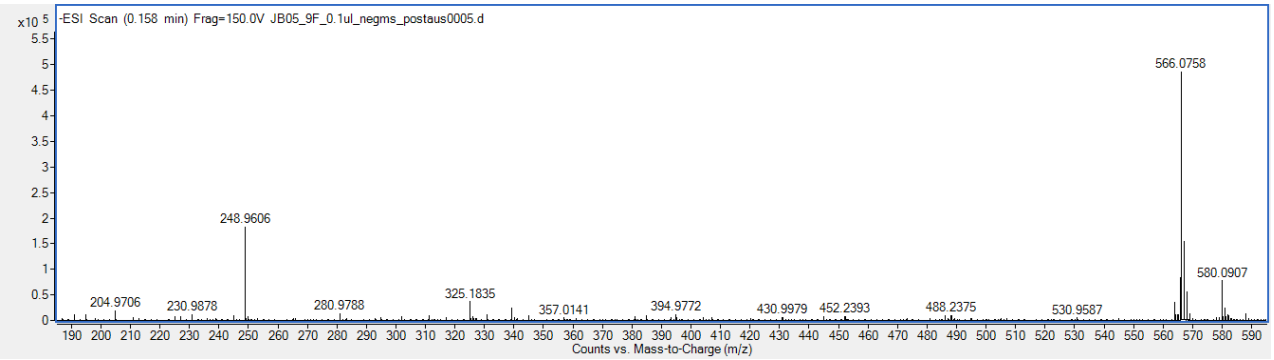
HSQC Spectrum (800 MHz, DMSO-d_6) of **153**



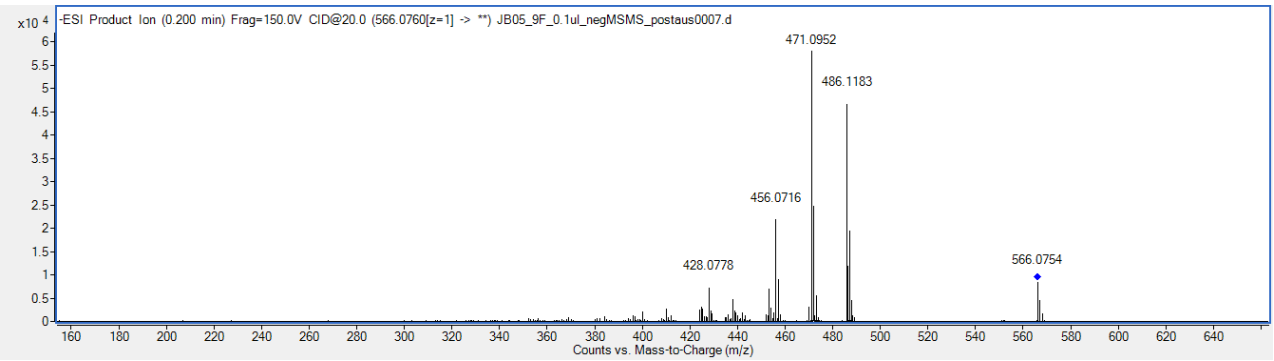
HMBC Spectrum (800 MHz, DMSO-d_6) of **153**



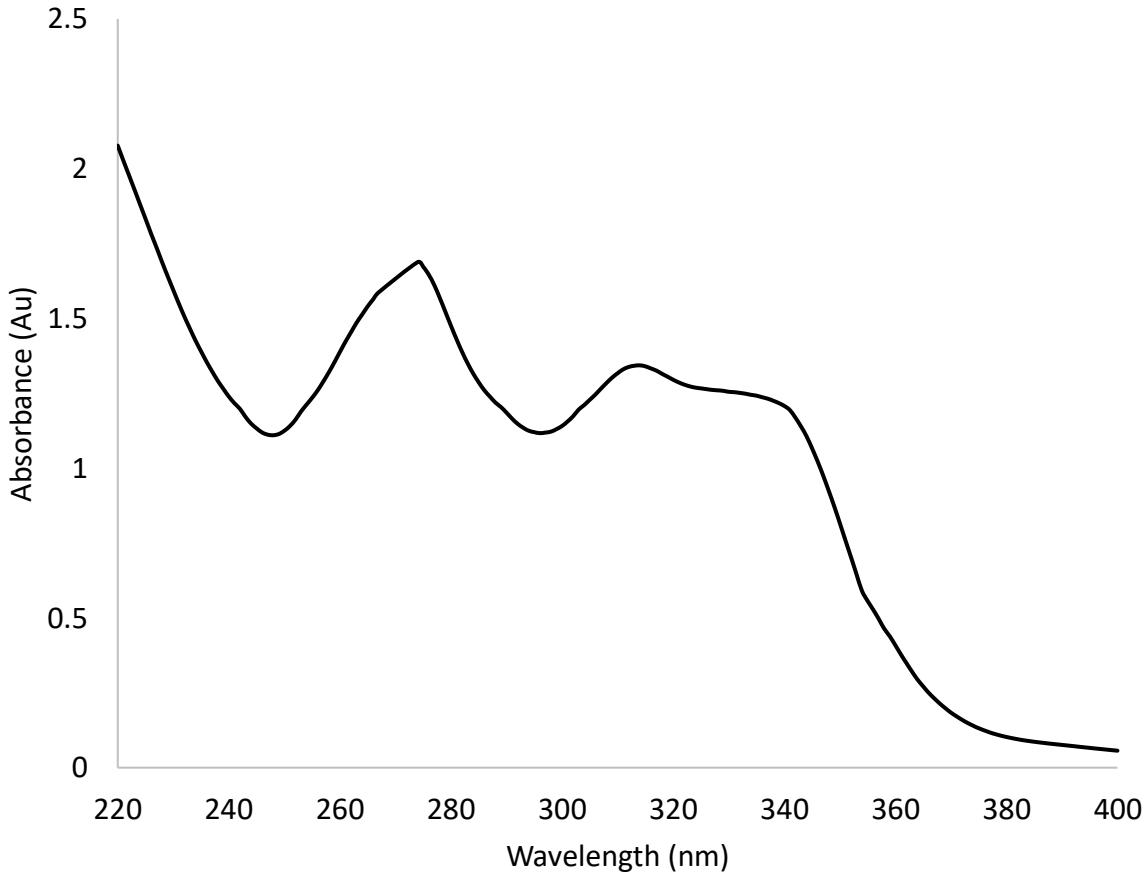
ROESY Spectrum (800 MHz, DMSO-d_6) of **153**



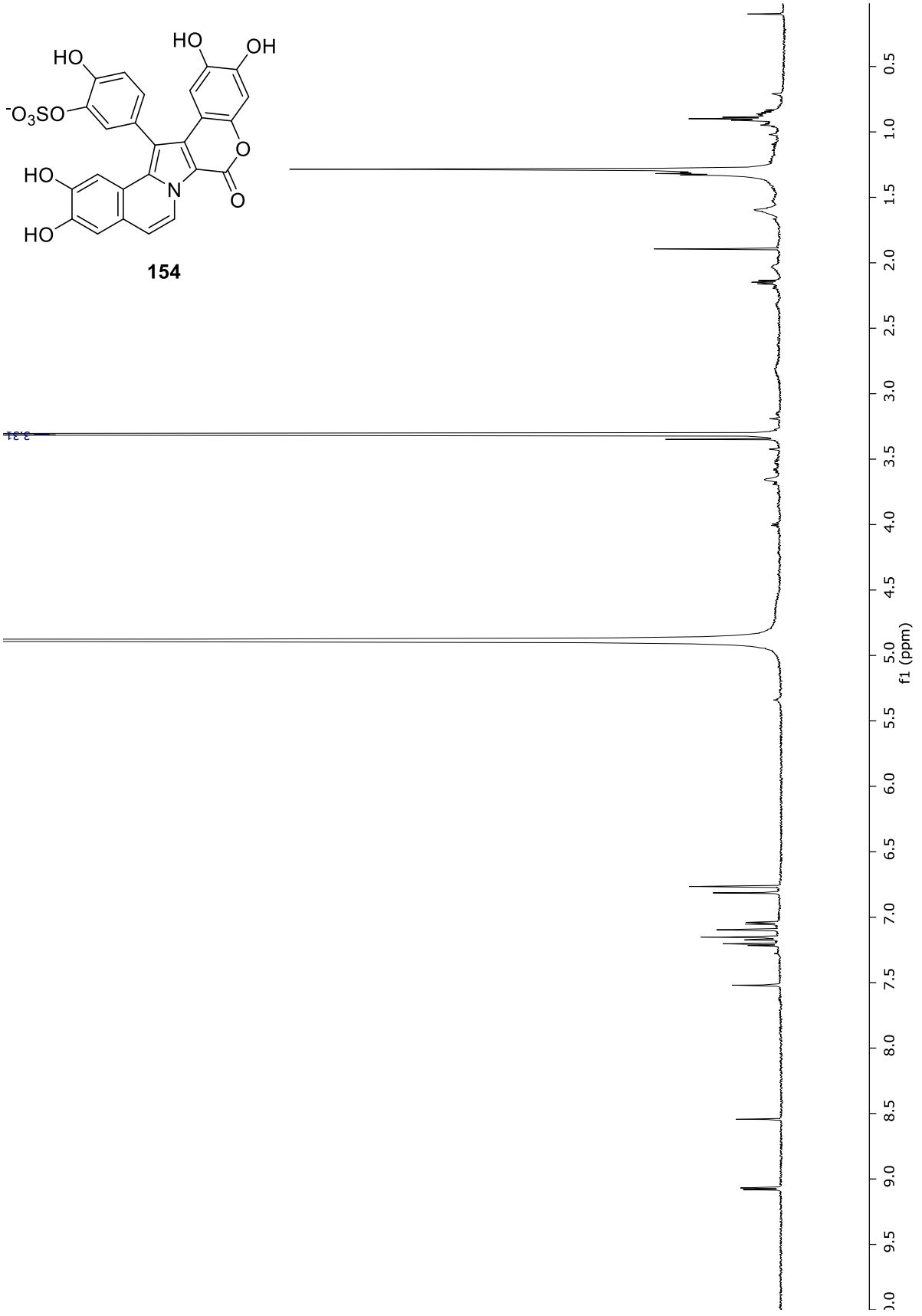
(–)-HRESIMS Spectrum of **153**



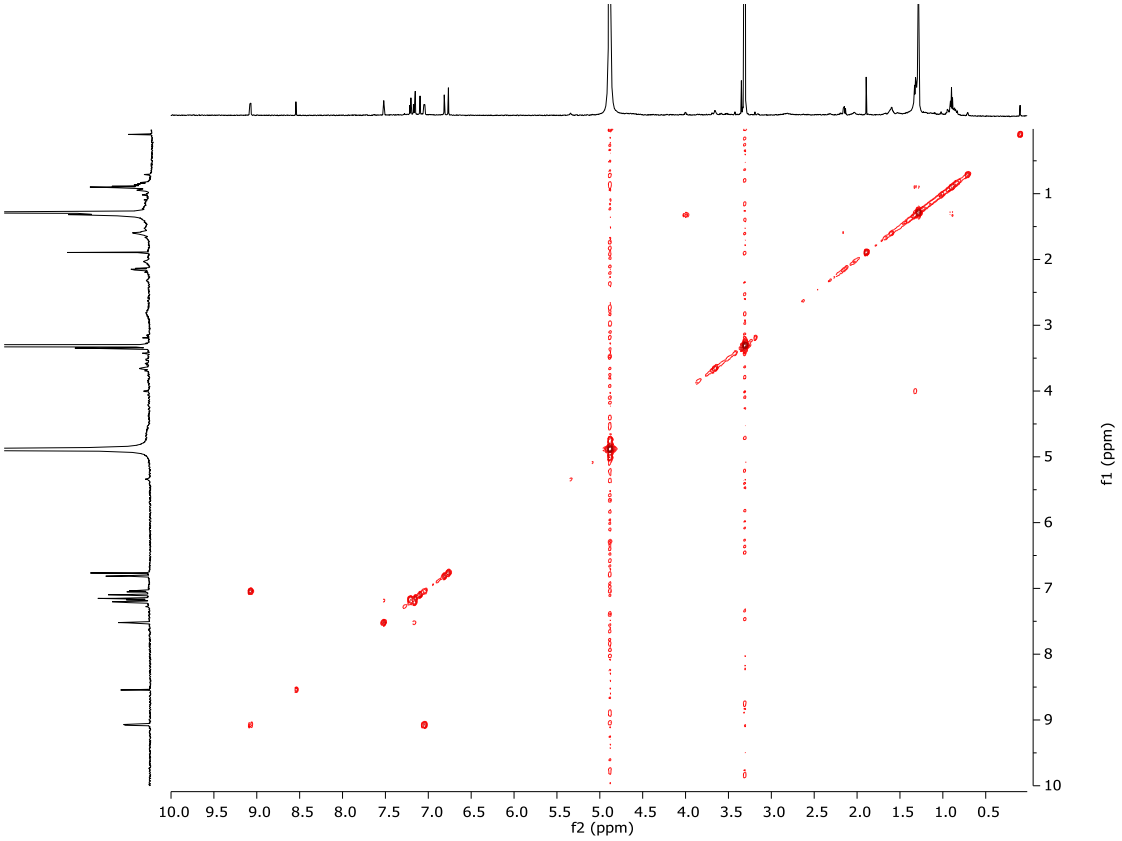
(–)-HRESIMS/MS Spectrum of **153**



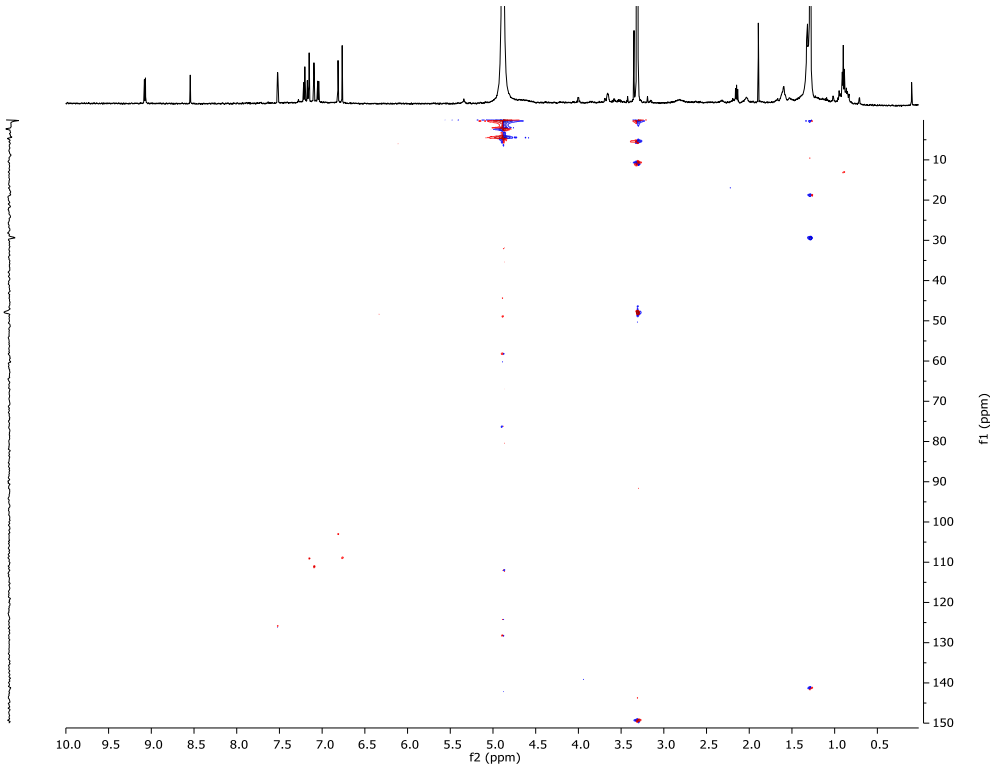
UV/vis Spectrum of **153**



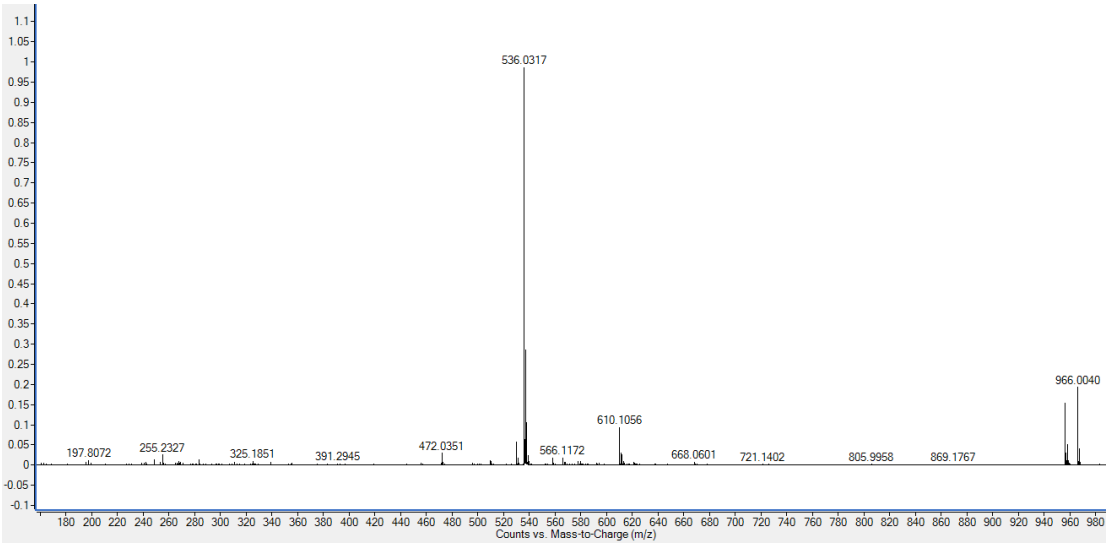
^{1}H NMR Spectrum (600 MHz, CD_3OD) of **154**



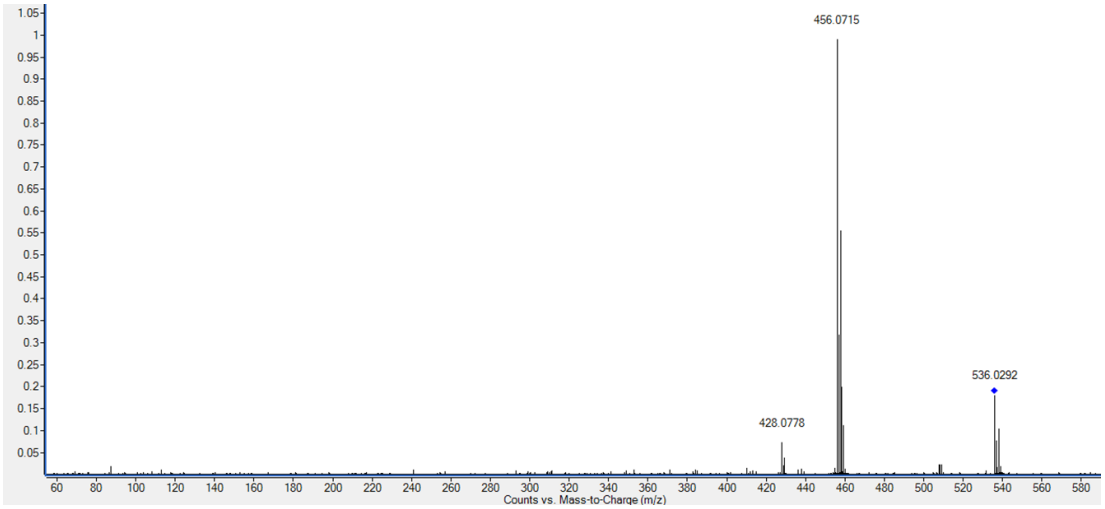
COSY Spectrum (600 MHz, CD_3OD) of **154**



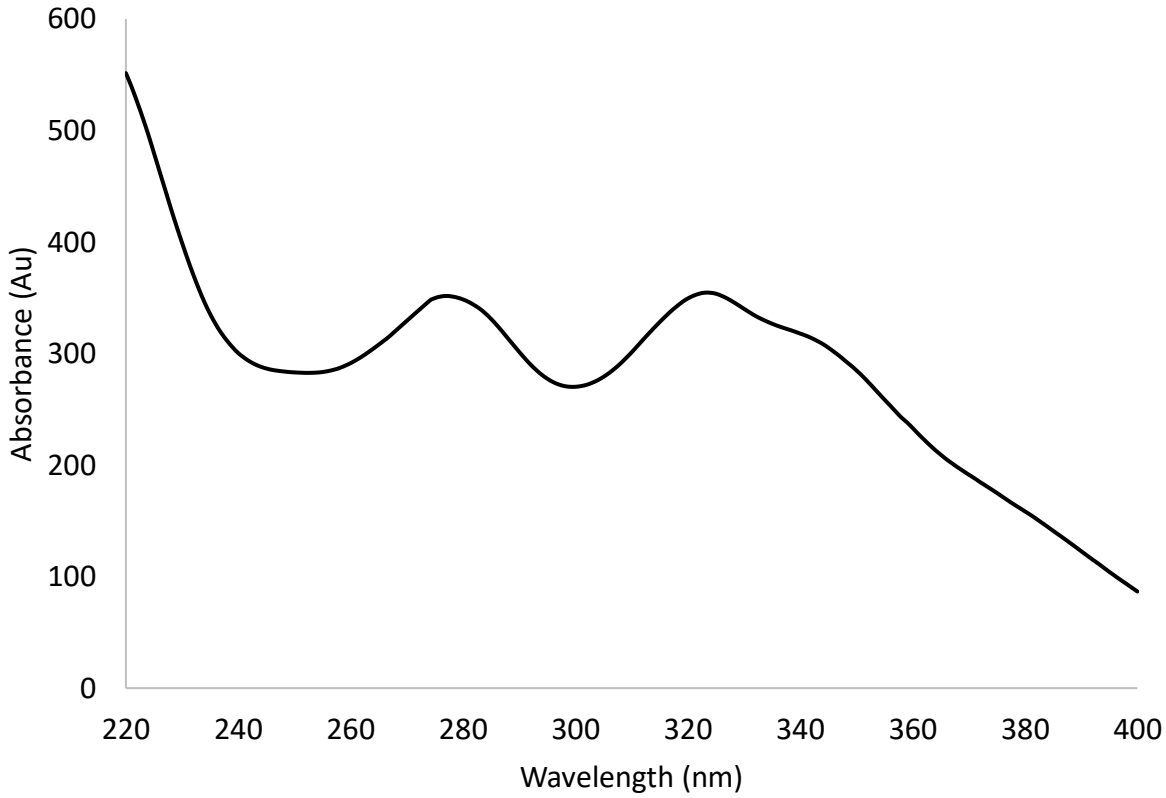
HSQC Spectrum (600 MHz, CD_3OD) of **154**



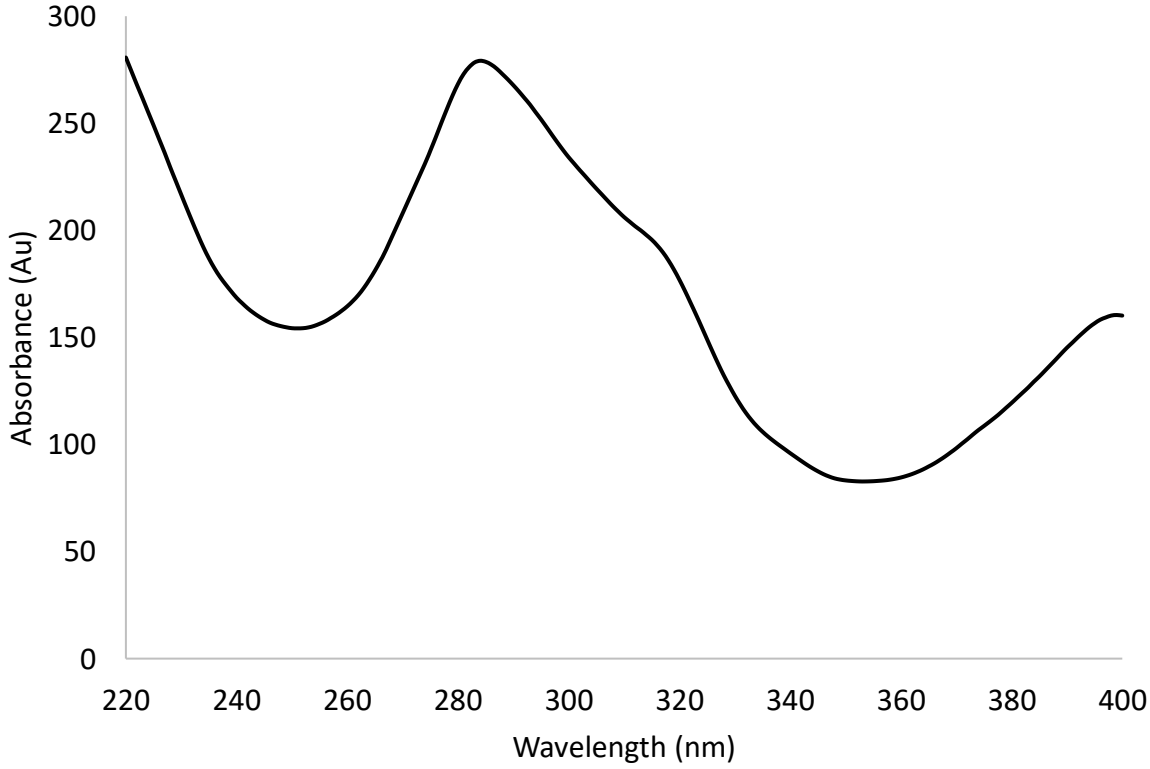
(–)-HRESIMS Spectrum of **154**



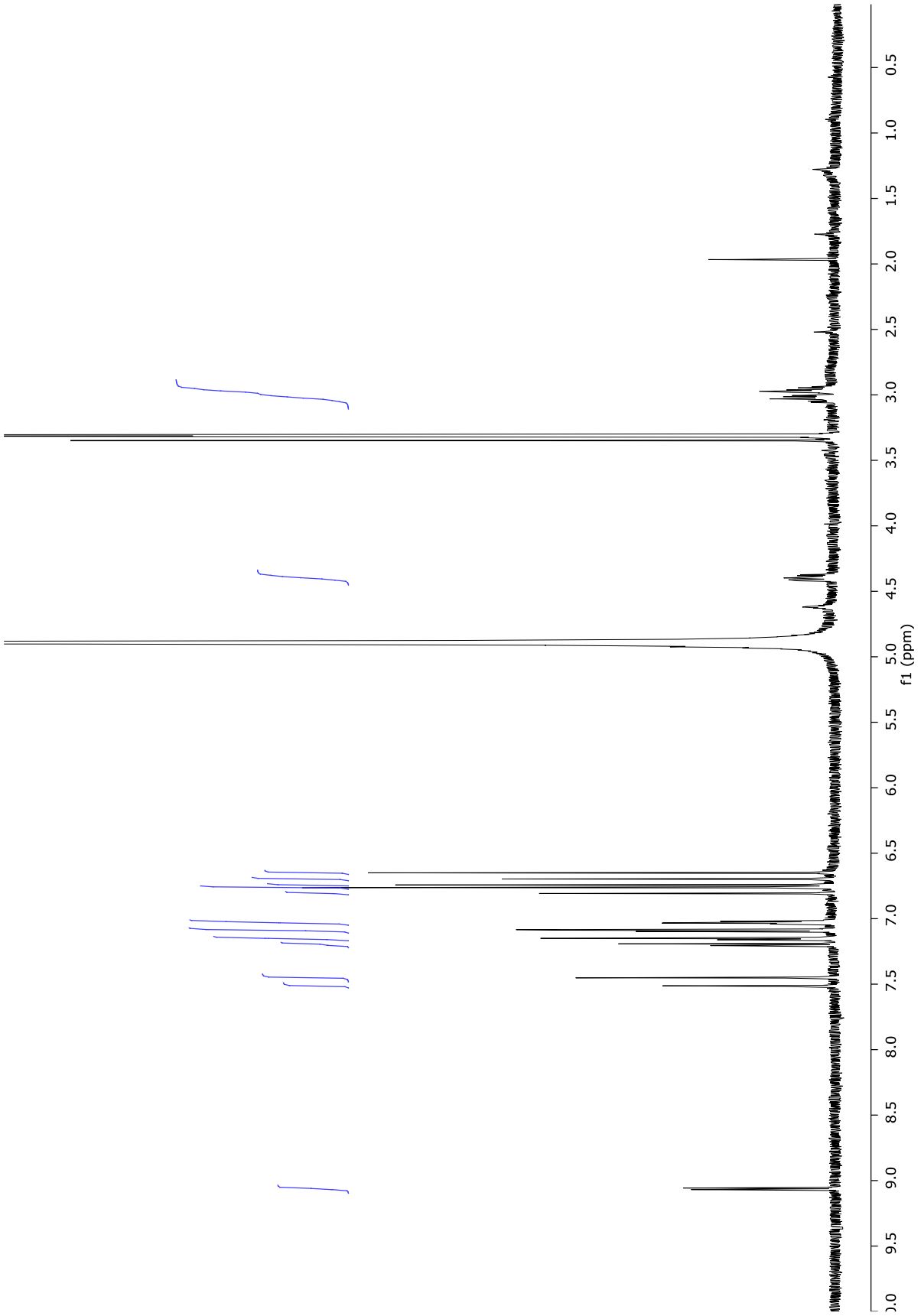
(–)-HRESIMS/MS Spectrum of **154**



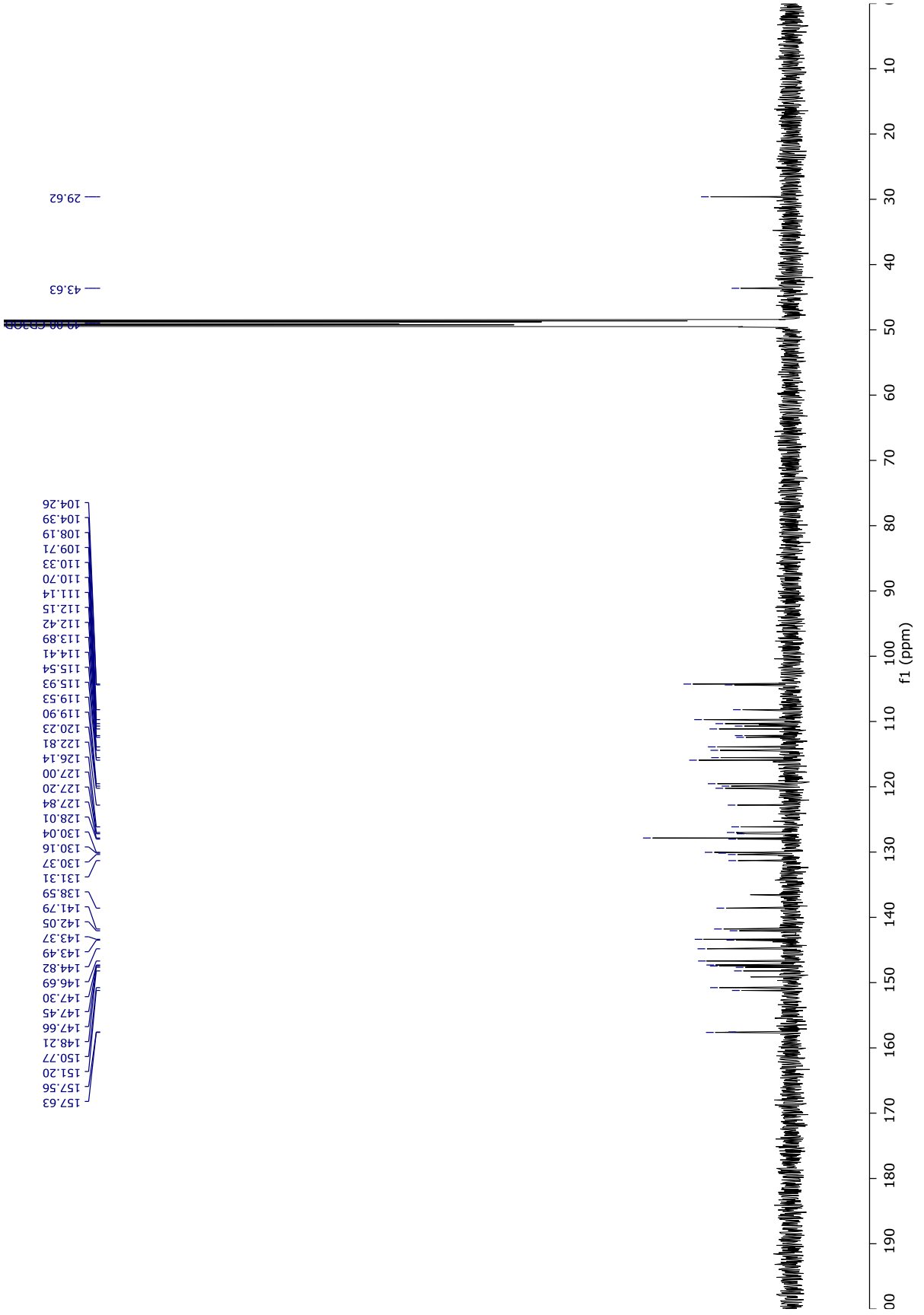
UV/vis Spectrum of **154**



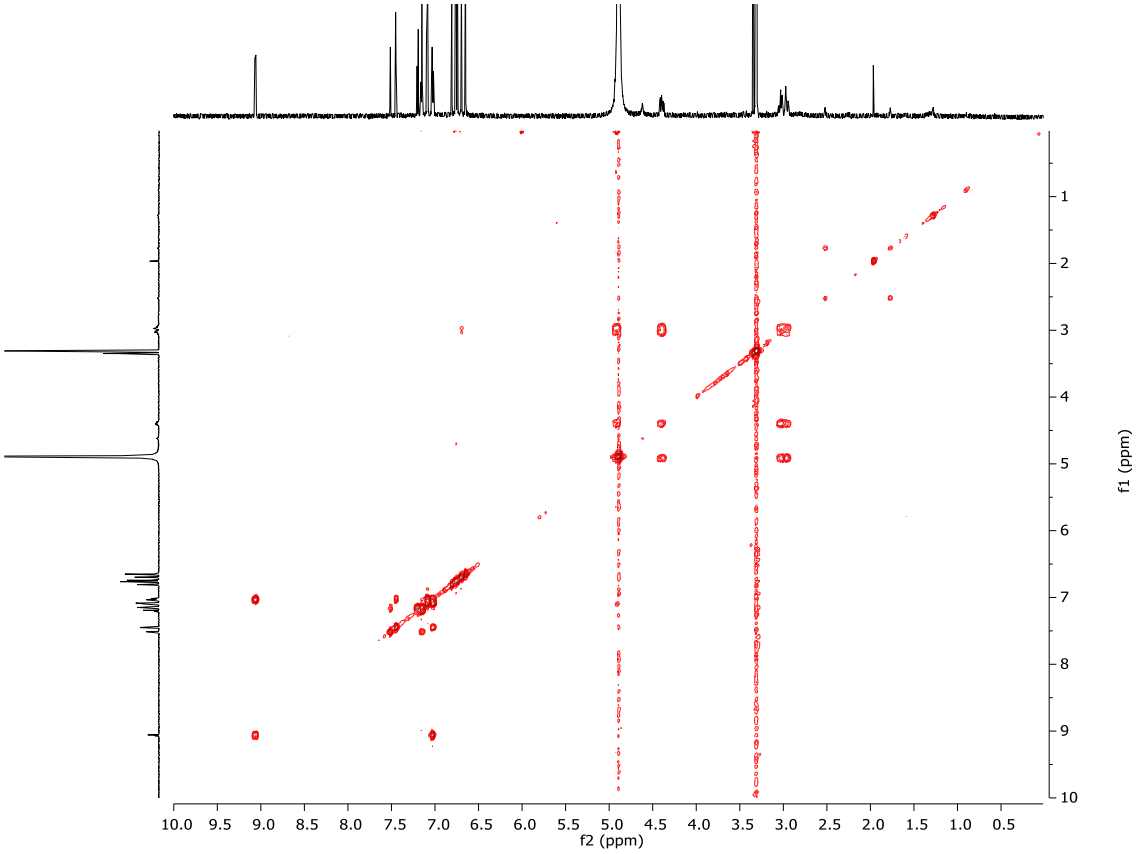
UV/vis Spectrum of **155**



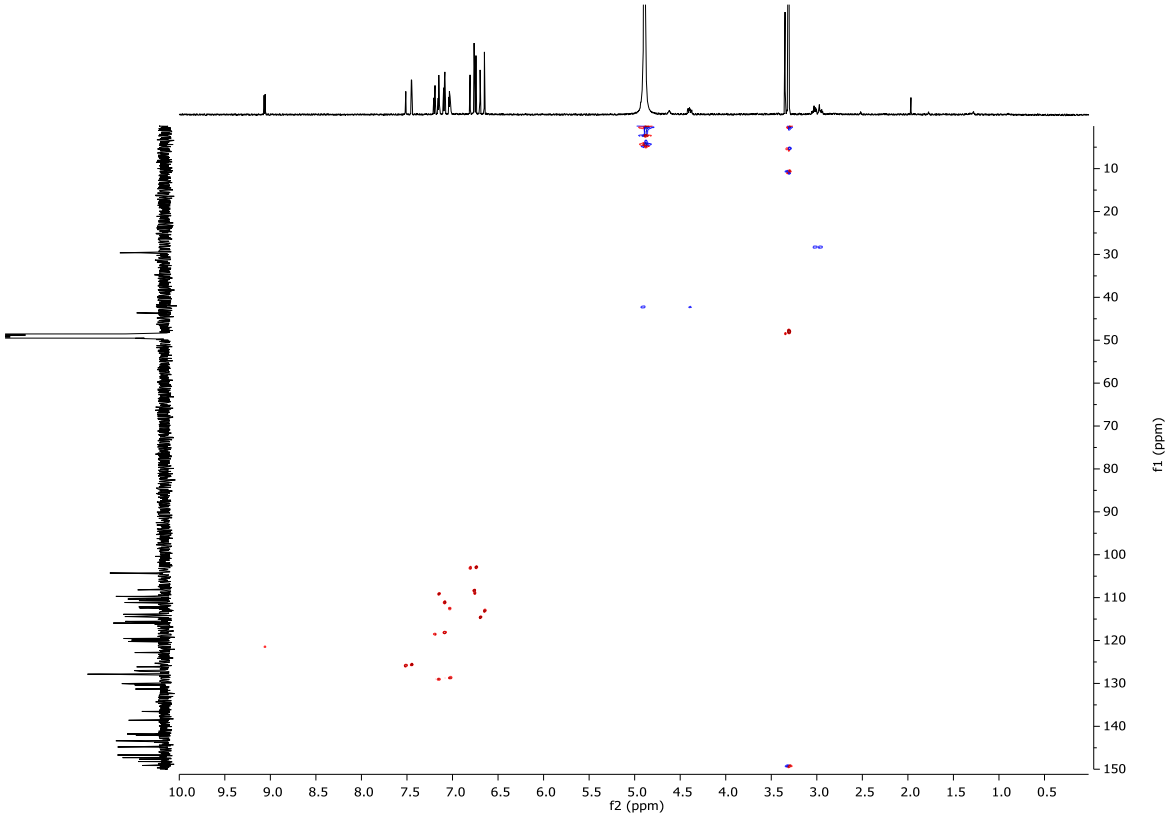
^{1}H NMR Spectrum (600 MHz, CD_3OD) of **154** and **155** (mixture)



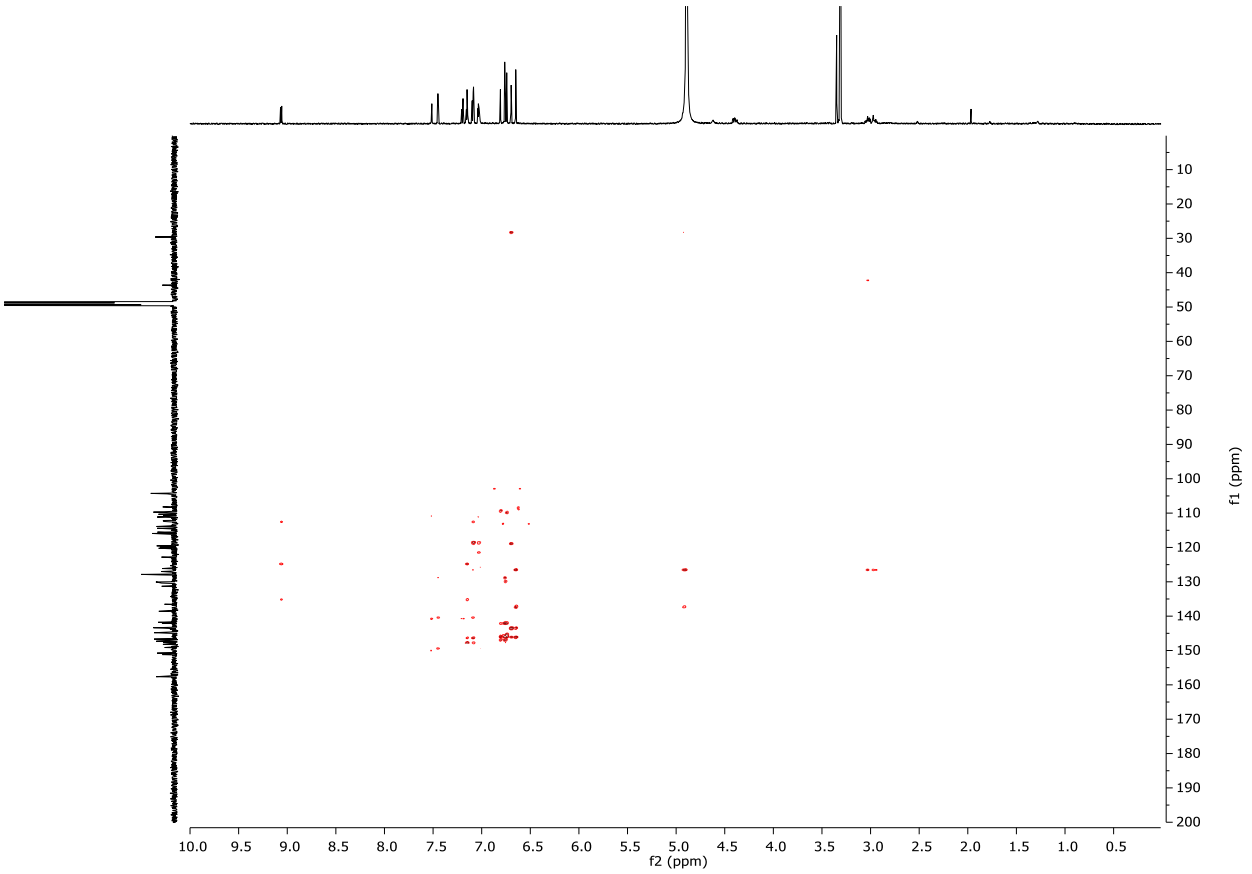
^{13}C NMR Spectrum (150 MHz, CD_3OD) of **154** and **155** (mixture)



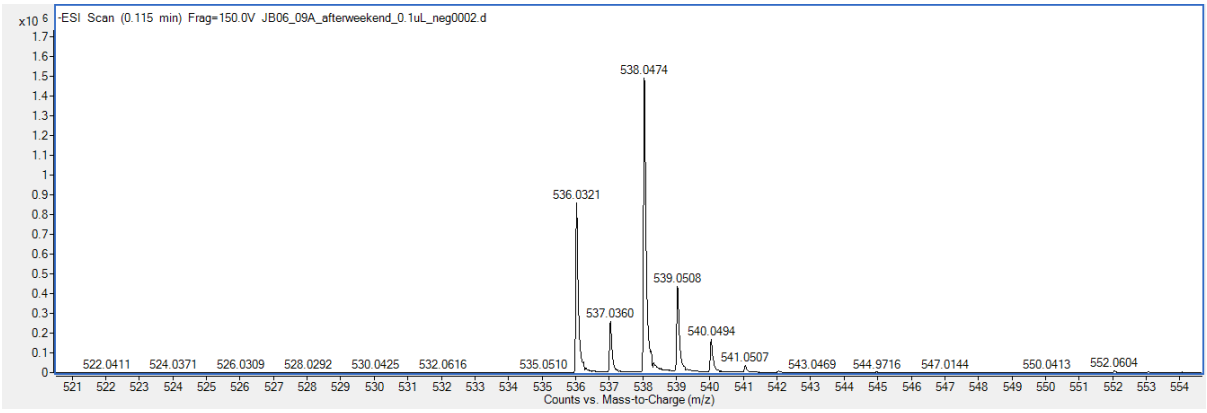
COSY Spectrum (600 MHz, CD_3OD) of **154** and **155** (mixture)



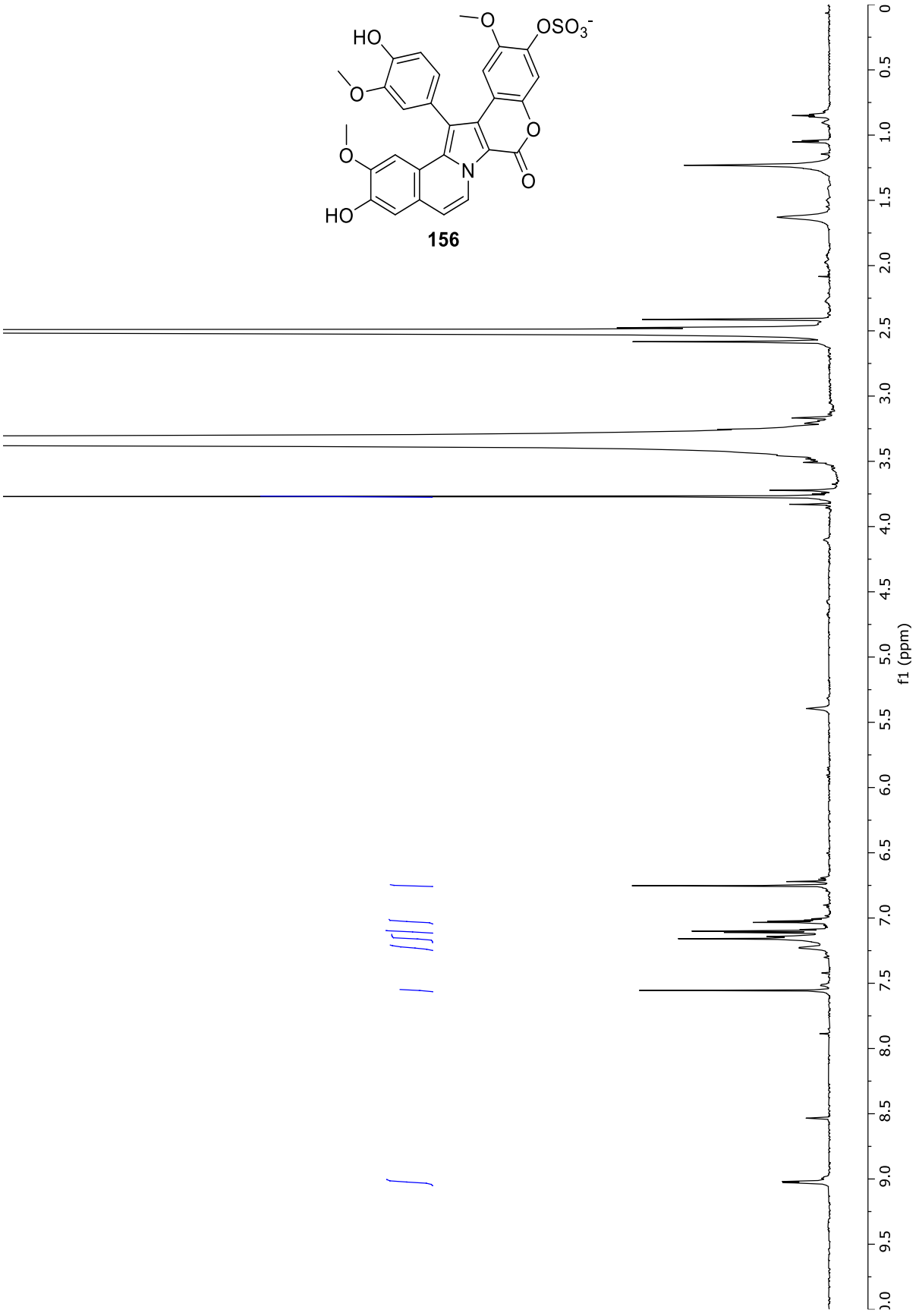
HSQC Spectrum (600 MHz, CD_3OD) of **154** and **155** (mixture)



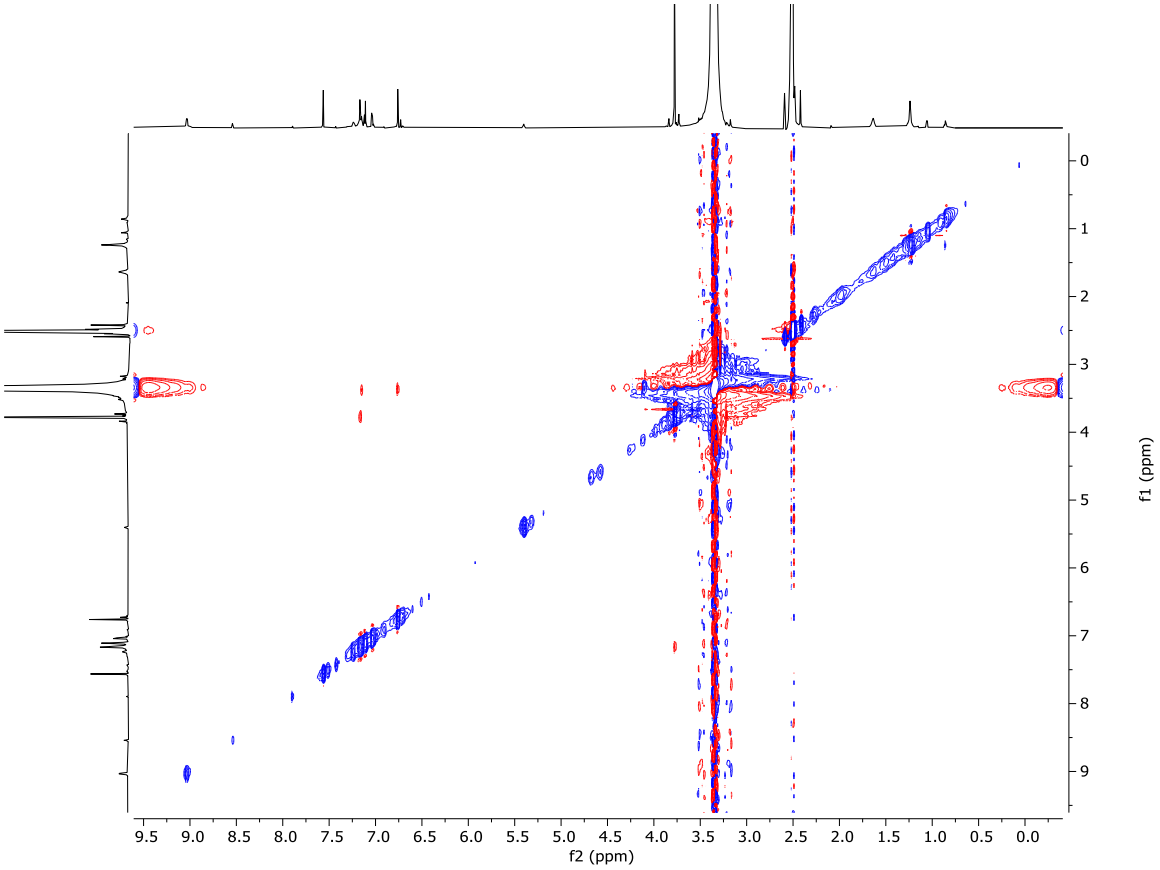
HMBC Spectrum (600 MHz, CD_3OD) of **154** and **155** (mixture)



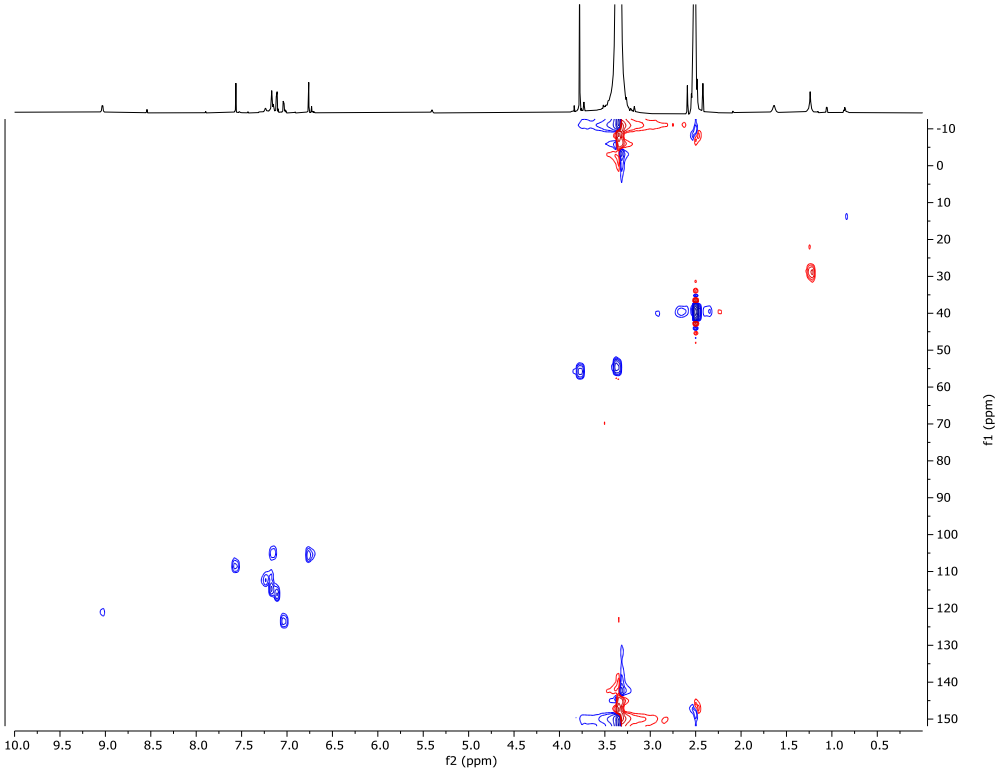
(–)-HRESIMS Spectrum of **154** and **155** (mixture)



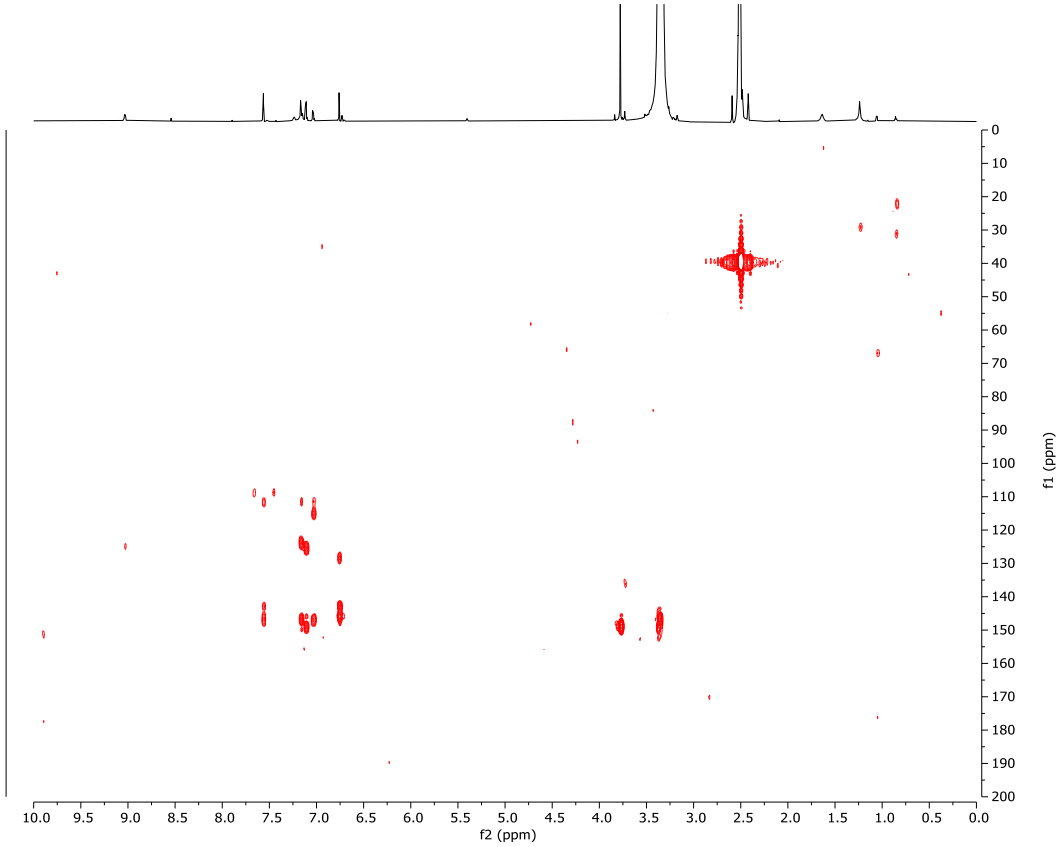
^{1}H NMR Spectrum (800 MHz, DMSO-d_6) of **156**



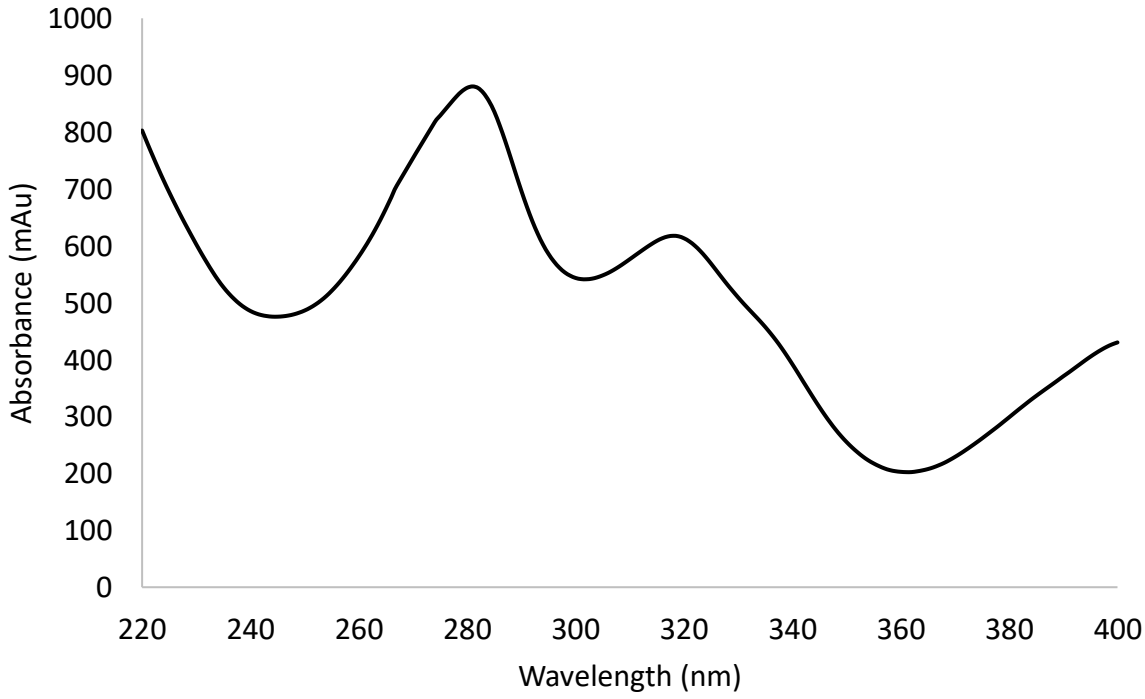
ROESY NMR Spectrum (800 MHz, DMSO-d_6) of **156**



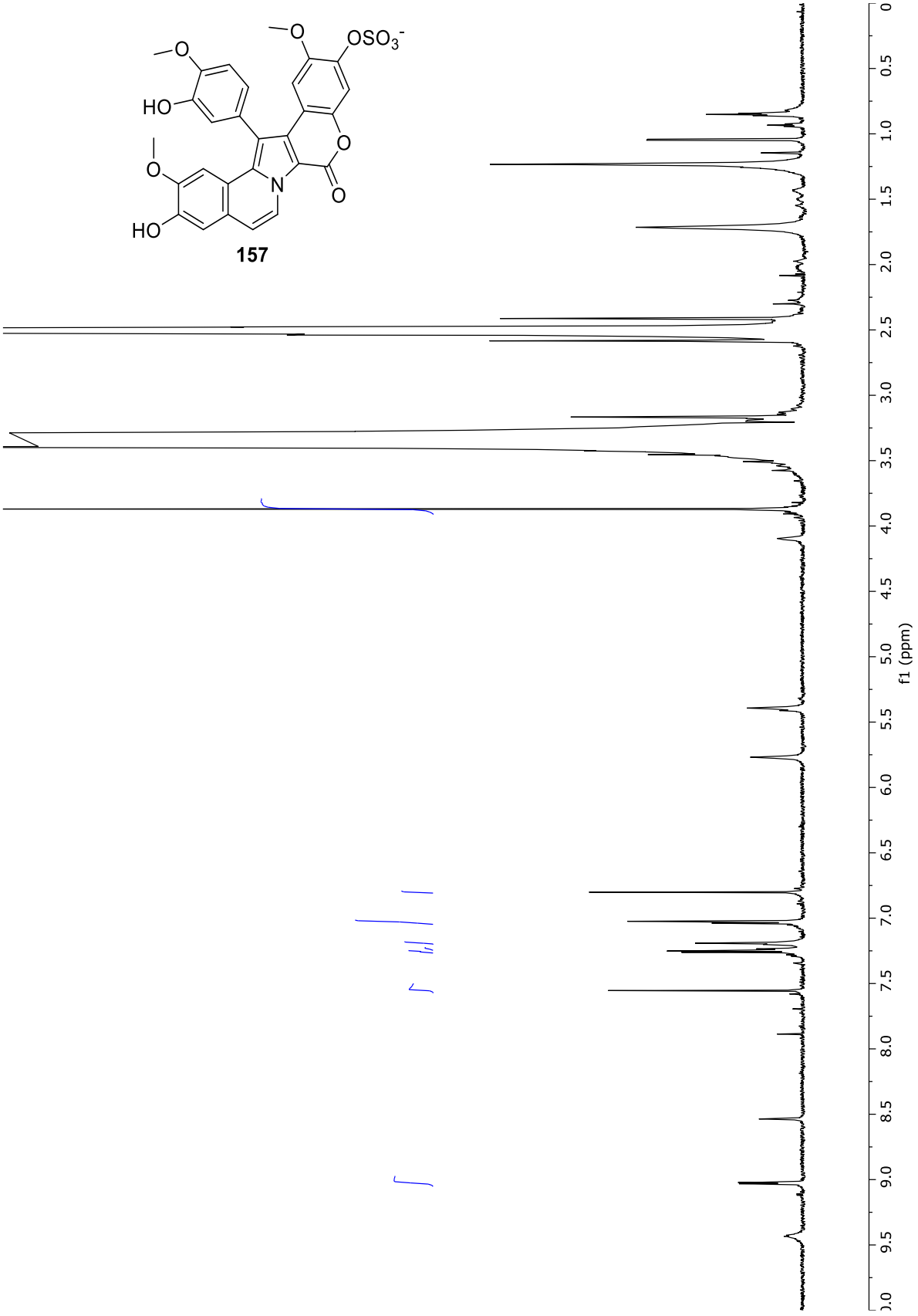
HSQC Spectrum (800 MHz, DMSO-d_6) of **156**



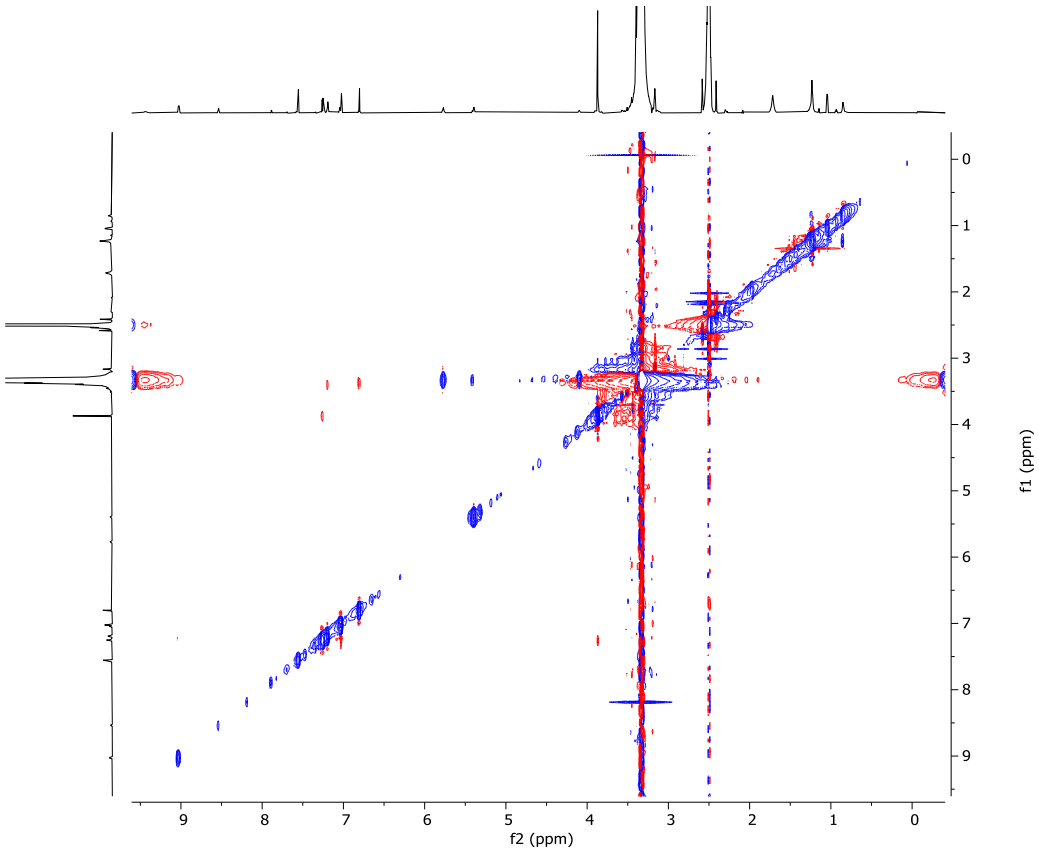
HMBC Spectrum (800 MHz, DMSO-d_6) of **156**



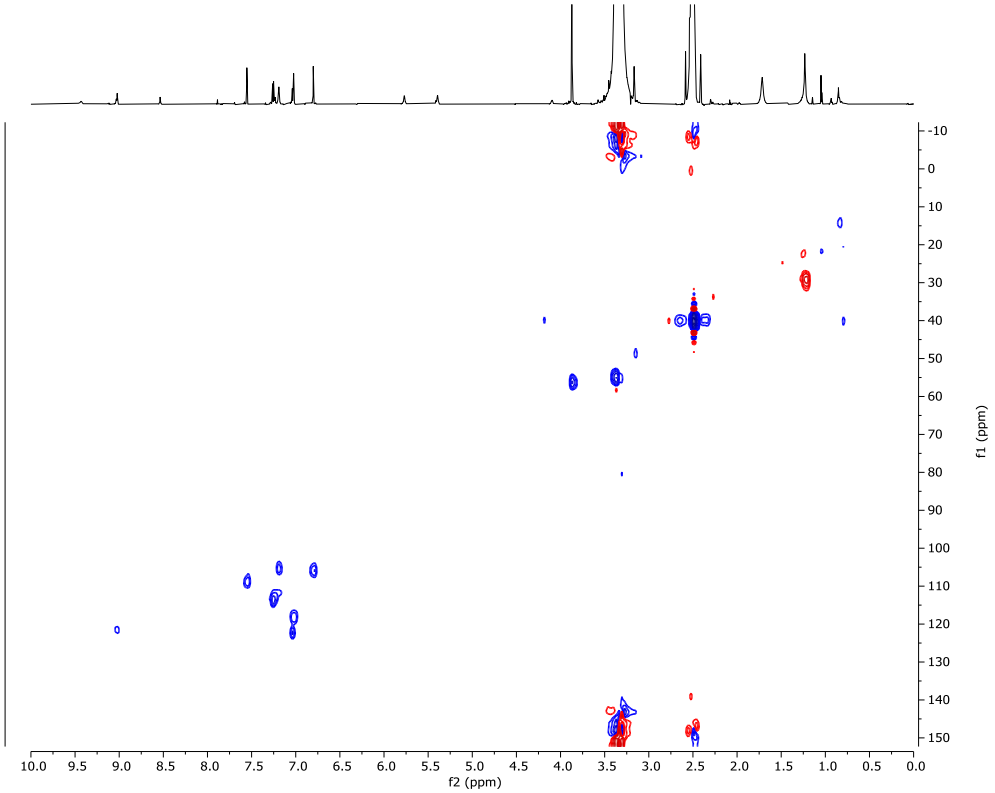
UV/vis Spectrum of **156**



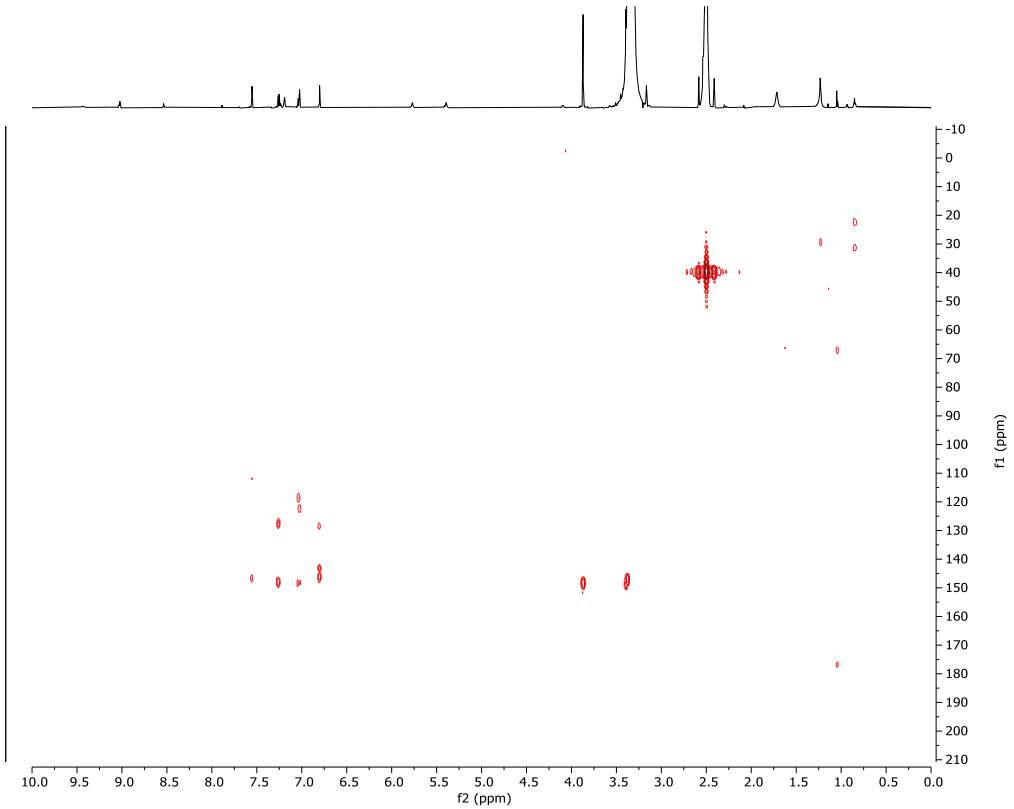
^{1}H NMR Spectrum (800 MHz, DMSO-d_6) of **157**



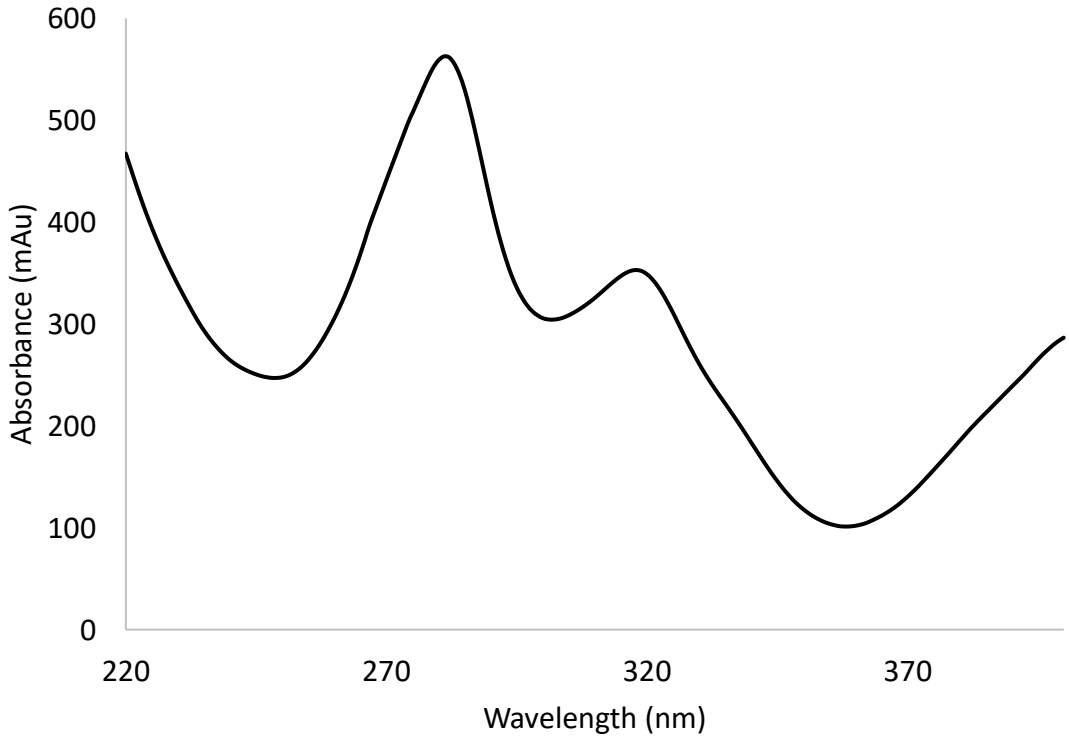
ROESY NMR Spectrum (800 MHz, DMSO-d_6) of **157**



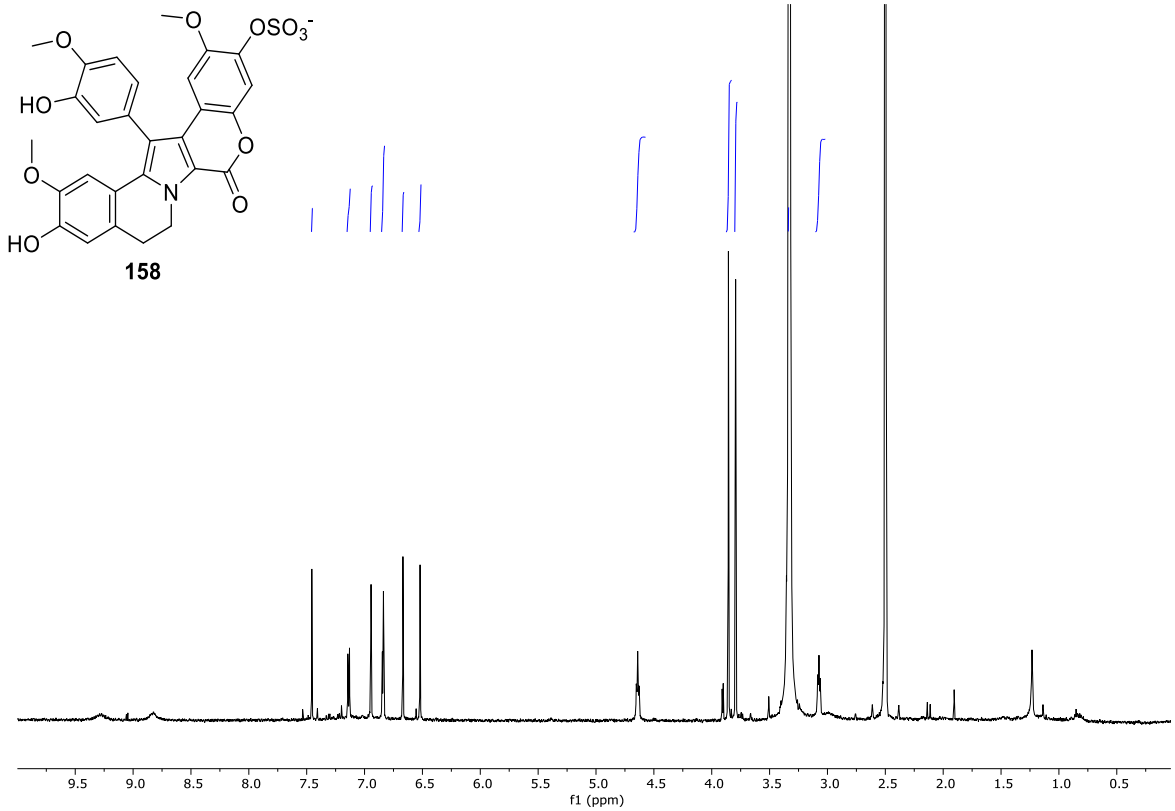
HSQC Spectrum (800 MHz, DMSO-d_6) of **157**



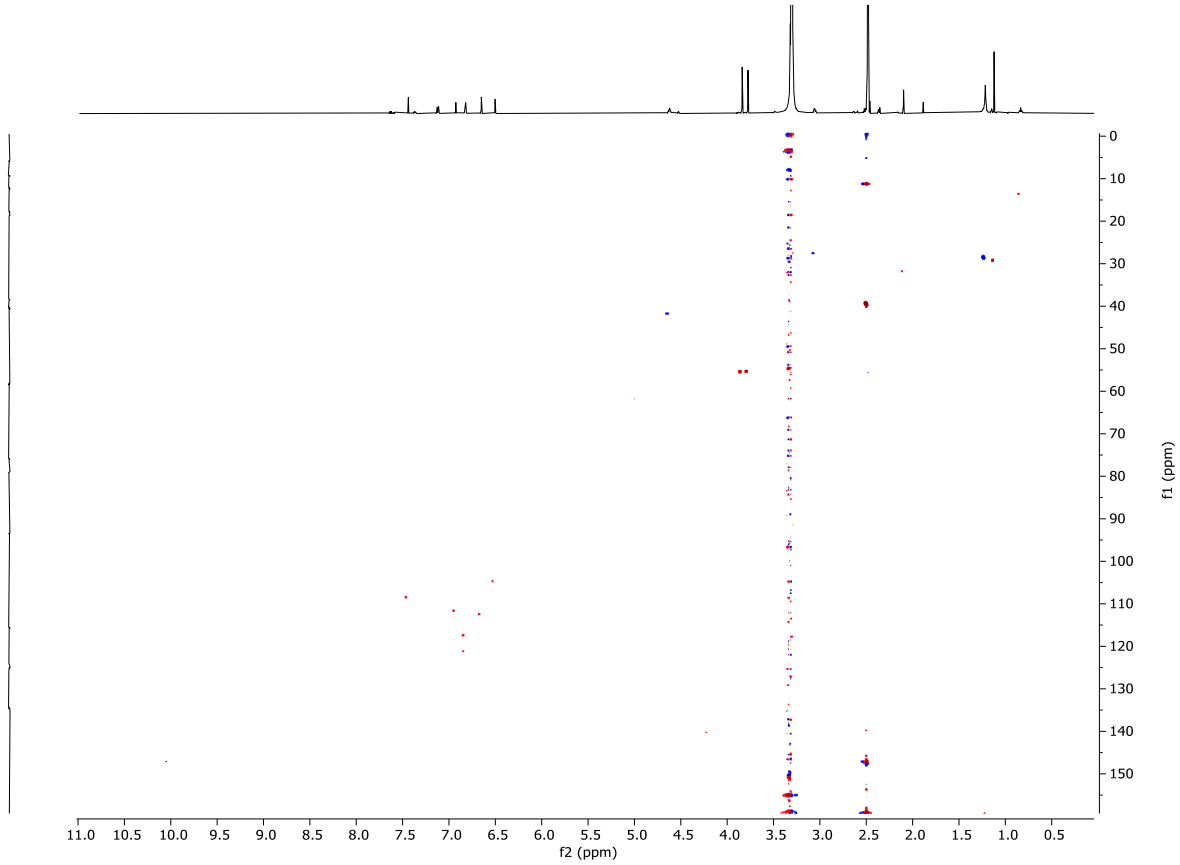
HMBC Spectrum (800 MHz, DMSO-d_6) of **157**



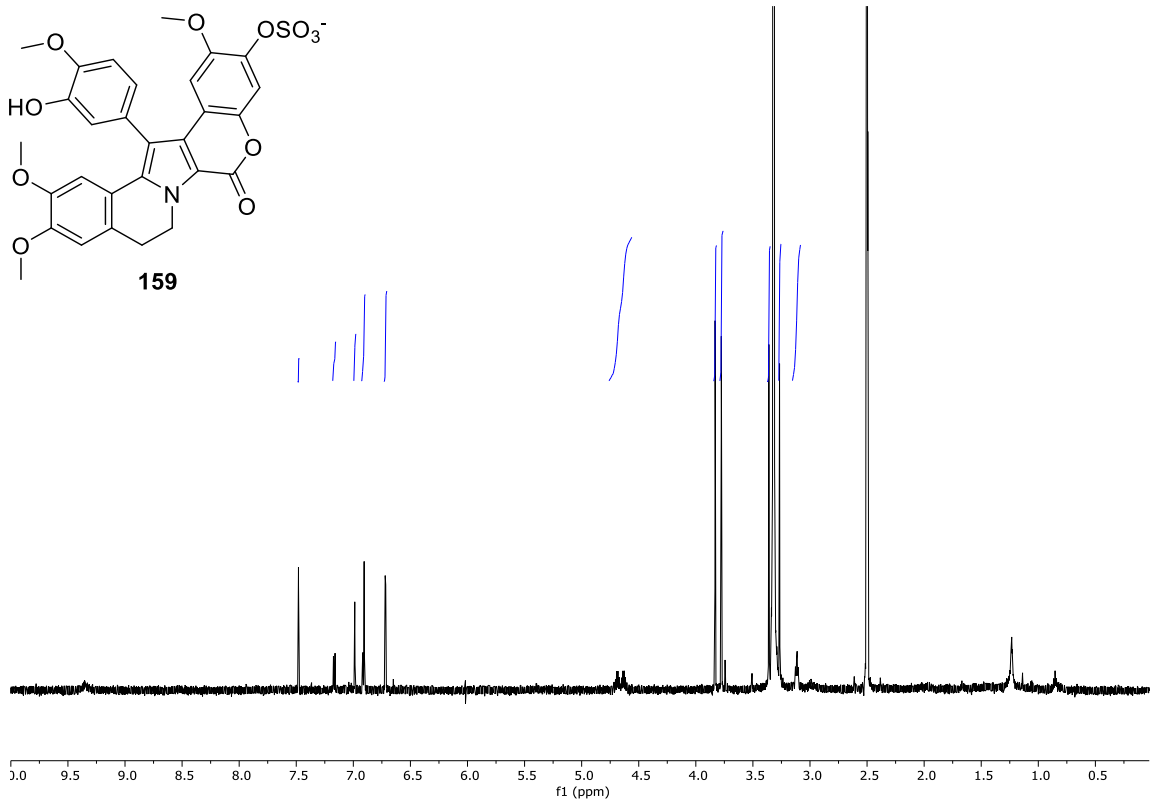
UV/vis Spectrum of **157**



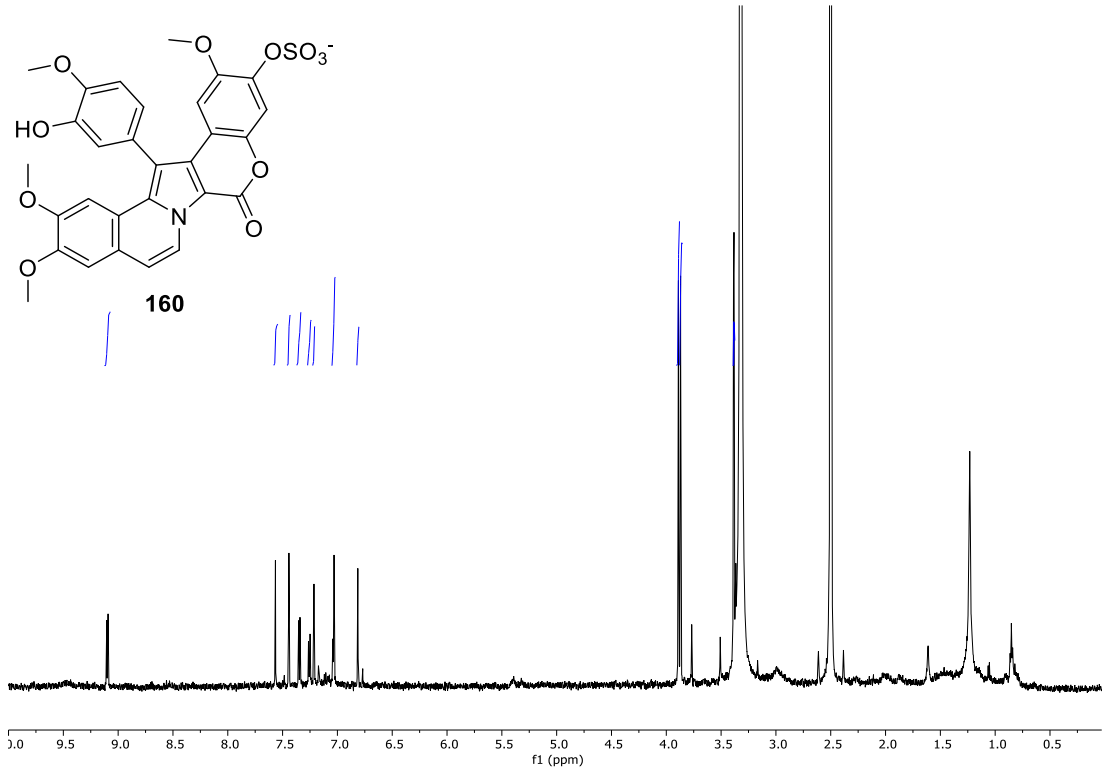
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **158**



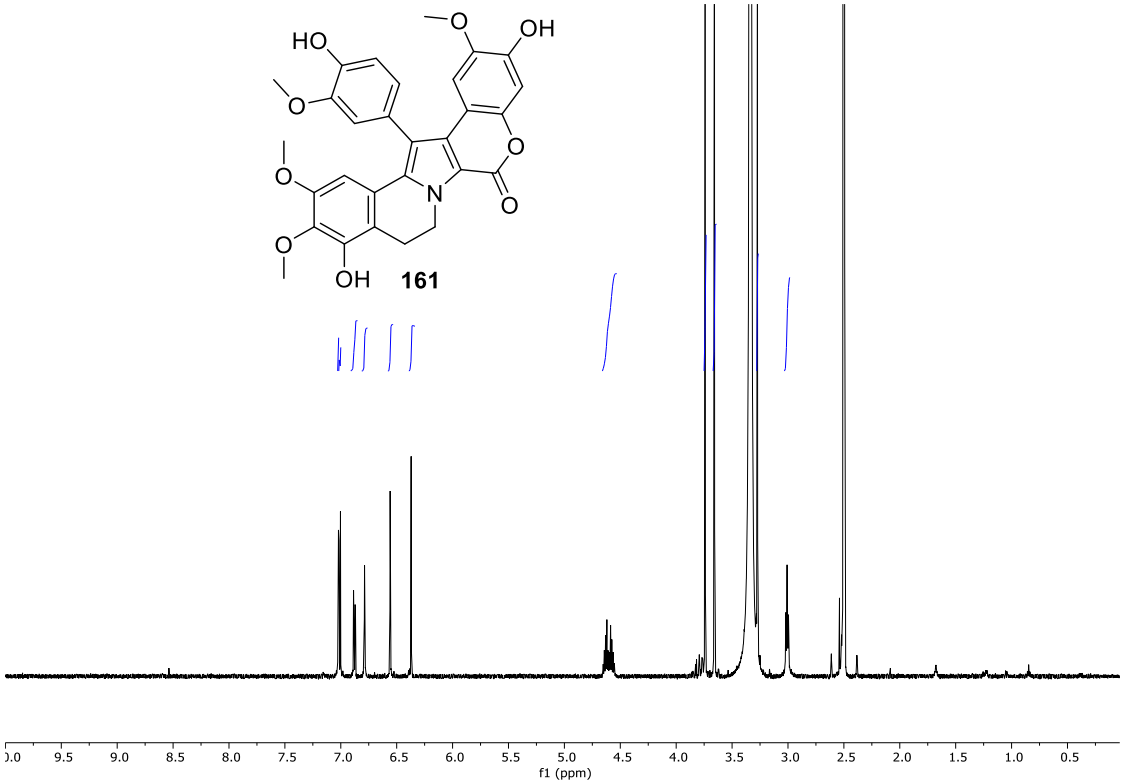
HSQC NMR Spectrum (600 MHz, DMSO-d_6) of **158**



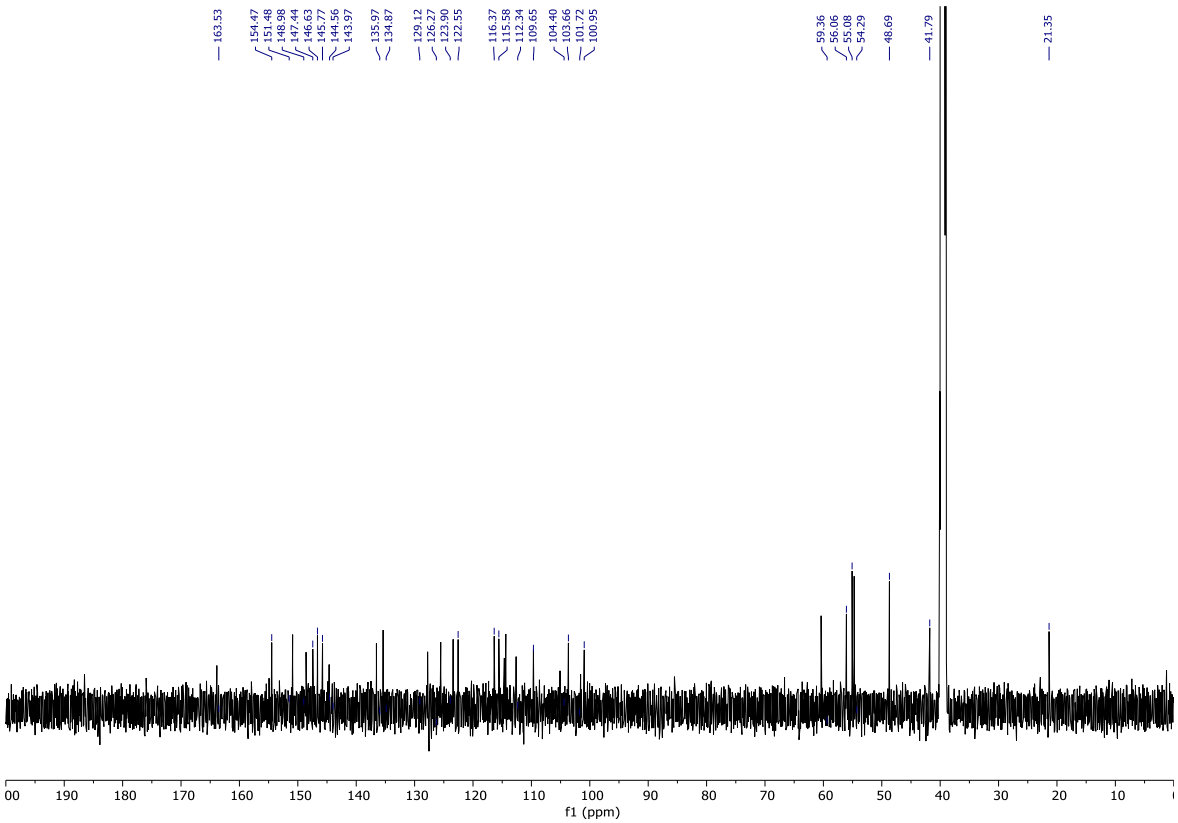
^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **159**



^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **160**



^{1}H NMR Spectrum (600 MHz, DMSO-d_6) of **161**



^{13}C NMR Spectrum (150 MHz, DMSO-d_6) of **161**

Appendix 9 – Kinase Inhibition Data

Inhibition of 485 human–disease relevant kinases by nellielloside A (**61**) at 10 μM (ThermoFisher Life Technologies SelectScreen® Whole Panel ACCESS Program, Wisconsin, USA. All measurements were made using solutions of the compounds in 100% DMSO; results are quoted as percent inhibition. The Z'–LYTE®, LanthaScreen and Adapta Screening Protocols were used to obtain the results for the respective kinases as reported below. Red = >80% inhibition, blue <40% inhibition.

Kinase	Mean	Kinase	Mean
GSK3B (GSK3 beta)	100	GSK3A (GSK3 alpha)	99
MAPKAPK3	98	RPS6KA3 (RSK2)	96
MELK	95	MAPKAPK2	95
MAPK14 (p38 alpha)	95	GSG2 (Haspin)	95
CAMK1 (CaMK1)	94	PLK3	93
GRK4	91	RPS6KA6 (RSK4)	88
ROCK2	80	BMPR1B (ALK6)	78
MAP4K5 (KHS1)	77	BMPR2	76
AMPK A2/B1/G1	75	LRRK2 G2019S	74
GRK6	72	MAPK8 (JNK1)	72
SGK (SGK1)	69	MAP2K4 (MEK4)	71
CLK4	68	AURKA (Aurora A)	68
MAP4K4 (HGK)	67	LRRK2 G2019S FL	67
BLK	66	RPS6KA5 (MSK1)	66
MAP2K1 (MEK1) S218D S222D	64	PRKACA (PKA)	64
CAMK2A (CaMKII alpha)	62	LRRK2 FL	63
NLK	60	DYRK2	61
LRRK2 R1441C	60	SGK2	60
TGFBR2	59	PRKD2 (PKD2)	59
BMX	58	MARK4	58
AMPK (A2/B1/G2)	58	LRRK2 I2020T	58
AMPK A1/B1/G1	57	FGR	58
ACVR1 (ALK2) R206H	57	LRRK2	57
FLT3 D835Y	57	SRMS (Srm)	57
FLT3	56	CAMK2B (CaMKII beta)	57
ACVR2B	54	STK22D (TSSK1)	55
EGFR (ErbB1) T790M L858R	53	MAPKAPK5 (PRAK)	53
NUAK2	53	RET Y791F	53
ROS1	52	ADRBK2 (GRK3)	53
RPS6KA4 (MSK2)	50	STK17A (DRAK1)	51

ABL1 E255K	49	VRK2	49
BRSK1 (SAD1)	48	ACVRL1 (ALK1)	48
RET V804L	48	ACVR1 (ALK2)	48
NEK2	47	TXK	47
DDR2 T654M	46	CAMK2D (CaMKII delta)	47
RET	46	MLCK (MLCK2)	46
CAMK4 (CaMKIV)	45	MARK3	46
CSNK1G3 (CK1 gamma 3)	43	DYRK3	44
BMPR1A (ALK3)	43	ABL1 Q252H	43
TNK2 (ACK)	42	LATS2	42
LYN B	42	CLK2	42
MAP2K6 (MKK6) S207E T211E	41	AMPK (A2/B1/G3)	41
STK38L (NDR2)	40	ABL1 Y253F	41
LCK	40	CHUK (IKK alpha)	40
MAP2K6 (MKK6)	39	MAPK9 (JNK2)	40
JAK3	39	ABL1 M351T	39
STK23 (MSSK1)	39	EPHA4	39
KDR (VEGFR2)	39	FER	39
TGFBR1 (ALK5)	38	ULK3	38
CDK8/cyclin C	37	CDK5/p35	38
BTK	37	MAPK10 (JNK3)	37
MYLK2 (skMLCK)	36	PRKD1 (PKC mu)	36
AMPK (A1/B2/G3)	36	PRKG2 (PKG2)	36
TTK	36	CDK11 (Inactive)	36
MAP3K19 (YSK4)	36	CDK11/cyclin C	36
FGFR1 V561M	36	KIT D816V	36
NTRK3 (TRKC)	35	RET V804E	35
STK38 (NDR)	35	ALK R1275Q	35
PKN2 (PRK2)	34	ALK C1156Y	34
ERN2	33	TEK (TIE2) Y897S	34
NEK1	33	MAP4K1 (HPK1)	33
MAP3K8 (COT)	32	ADCK3	33
MINK1	32	PAK4	32
EPHB3	31	ALK T1151_L1152insT	31
CAMK1G (CAMKI gamma)	31	MKNK2 (MNK2)	31
ABL1 H396P	31	FGFR3 K650M	31
NTRK1 (TRKA)	31	TYRO3 (RSE)	31
PRKACB (PRKAC beta)	31	RPS6KB1 (p70S6K)	31
ALK F1174L	30	IRAK1	30
MAP2K6 (MKK6)	30	FYN	30
DDR2 N456S	30	RET S891A	30
STK24 (MST3)	29	ABL1	30
FGFR1	29	NTRK2 (TRKB)	29
PRKCQ (PKC theta)	29	ABL1 G250E	29
MAPK1 (ERK2)	29	ABL1 T315I	29
MAP3K14 (NIK)	28	GAK	28

CAMK2G (CaMKII gamma)	27	PRKCG (PKC gamma)	28
CHEK1 (CHK1)	27	PLK1	27
KSR2	27	SNF1LK2	27
BRAF	27	MAP2K1 (MEK1)	27
YES1	27	CAMKK2 (CaMKK beta)	27
ALK	26	EPHA1	26
NUAK1 (ARK5)	26	PRKG1	26
EPHB4	25	MAP3K10 (MLK2)	25
LYN A	25	ADRBK1 (GRK2)	25
CHEK2 (CHK2)	25	CSNK1G2 (CK1 gamma 2)	25
ROCK1	25	ALK L1196M	25
CSNK1D (CK1 delta)	24	PIK3CD/PIK3R1 (p110 delta/p85 alpha)	25
MAP4K2 (GCK)	24	ACVR2A	24
CDK2/cyclin E1	24	BRAF V599E	24
AMPK (A1/B2/G2)	24	WEE1	24
EGFR (ErbB1) d747–749 A750P	23	STK22B (TSSK2)	23
MAP2K1 (MEK1)	23	FES (FPS)	23
MAP3K9 (MLK1)	23	EPHA7	23
JAK2	23	RPS6KA1 (RSK1)	23
AURKB (Aurora B)	23	RAF1 (cRAF) Y340D Y341D	23
AMPK (A1/B1/G2)	22	MYLK4	22
CDK5/p25	22	ABL2 (Arg)	22
RET A883F	22	MET (cMet)	22
MAPK8 (JNK1)	22	SPHK1	22
MUSK	21	CDK5 (Inactive)	21
STK32C (YANK3)	21	PLK4	21
PAK7 (KIAA1264)	21	MAP2K2 (MEK2)	21
CSNK1G1 (CK1 gamma 1)	21	RPS6KB2 (p70S6Kb)	21
EGFR (ErbB1) T790M	20	PRKCH (PKC eta)	20
PRKCA (PKC alpha)	20	KIT A829P	20
EIF2AK2 (PKR)	20	PIK3CA E542K/PIK3R1 (p110 alpha E542K/p85 alpha)	20
TEK (Tie2)	20	AMPK (A1/B1/G3)	20
HCK	20	PAK6	20
PRKCB2 (PKC beta II)	19	KIT D816H	20
CDK7/cyclin H/MNAT1	19	CAMK1D (CaMKI delta)	19
RPS6KA2 (RSK3)	19	PAK2 (PAK65)	19
STK32B (YANK2)	19	SLK	19
MET M1250T	19	FGFR2	19
PRKX	19	ANKK1	19
MARK2	18	CDK2/cyclin A1	18
CDK2/cyclin A	18	PTK2 (FAK)	18
NEK6	18	TYK2	18
RET M918T	18	TEC	18
PEAK1	18	JAK2 JH1 JH2	18
MAP2K2 (MEK2)	17	STK17B (DRAK2)	18
SPHK2	17	PIK3CA/PIK3R1 (p110 alpha/p85 alpha)	17

MARK1 (MARK)	17	MAP3K11 (MLK3)	17
EPHA2	17	DYRK1B	17
STK39 (STLK3)	17	SGKL (SGK3)	17
FGFR3 G697C	16	PRKCN (PKD3)	16
KIT V559D T670I	16	CDK3/cyclin E1	16
AAK1	16	CDC7/DBF4	16
MST4	16	DAPK2	16
PRKACG (PRKAC gamma)	15	CSNK1E (CK1 epsilon)	16
TLK2	15	MAPK7 (ERK5)	15
IRAK3	15	RET G691S	15
MYO3B (MYO3 beta)	15	HIPK4	15
MYLK (MLCK)	15	MAP3K7/MAP3K7IP1 (TAK1–TAB1)	15
MERTK (cMER) A708S	15	CSNK1E (CK1 epsilon) R178C	15
TAOK3 (JIK)	14	ULK1	14
ZAK	14	CDK13/cyclin K	14
CLK3	14	TNIK	14
MAP4K3 (GLK)	14	AURKC (Aurora C)	14
PI4KB (PI4K beta)	14	MAP3K5 (ASK1)	14
CDK16 (PCTK1)/cyclin Y	14	RIPK3	14
MYO3A (MYO3 alpha)	14	AMPK (A1/B2/G1)	14
MASTL	13	SBK1	13
FGFR3 V555M	13	ERBB4 (HER4)	13
MAPK10 (JNK3)	13	MET D1228H	13
EGFR (ErbB1) d746–750	13	PLK2	13
MAPK3 (ERK1)	13	MLK4	13
KIT Y823D	13	STK3 (MST2)	13
FGFR3 K650E	13	SIK1	13
AXL R499C	12	IRAK4	12
STK4 (MST1)	12	CASK	12
NEK8	12	CDK9/cyclin T1	12
INSR	12	PKMYT1	12
IGF1R	12	BRAF V599E	12
PASK	12	EPHB1	12
PRKCB1 (PKC beta I)	12	PIK3CA/PIK3R3 (p110 alpha/p55 gamma)	12
PIP5K1B	12	FGFR2 N549H	12
EGFR (ErbB1) G719C	12	TAOK1	12
PTK2B (FAK2)	11	EPHA6	12
CLK1	11	MAP2K5 (MEK5)	11
KIT N822K	11	CDK2/cyclin O	11
FLT3 ITD	11	RET V804M	11
FRK (PTK5)	11	RAF1 (cRAF) Y340D Y341D	11
ABL1 F317I	11	JAK1	11
PIK3CB/PIK3R1 (p110 beta/p85 alpha)	10	PRKCD (PKC delta)	10
AMPK (A2/B2/G1)	10	CSK	10
EPHB2	10	CDK14 (PFTK1)/cyclin Y	10
AXL	10	KIT D820E	10

ULK2	10	EPHA8	10
PI4K2B (PI4K2 beta)	10	DDR2	10
STK16 (PKL12)	10	PIK3C2G (PI3K–C2 gamma)	10
SYK	10	ERN1	10
EGFR (ErbB1) G719S	10	PHKG1	10
GRK1	9	MERTK (cMER)	10
ITK	9	FLT4 (VEGFR3)	9
PTK6 (Brk)	9	LIMK2	9
PDK1	9	CSNK1A1L	9
PIK3C2B (PI3K–C2 beta)	9	PIK3CA E545K/PIK3R1 (p110 alpha E545K/p85 alpha)	9
FGFR3	9	ABL1 F317L	9
TNK1	8	TESK1	8
BRAF	8	EGFR (ErbB1) T790M C797S L858R	8
EGFR (ErbB1)	8	HIPK2	8
TLK1	8	MAPK15 (ERK7)	8
AMPK (A2/B2/G2)	8	FGFR4	8
PI4K2A (PI4K2 alpha)	7	SIK3	7
PDGFRA D842V	7	MST1R (RON)	7
PHKG2	7	MAPK9 (JNK2)	7
AMPK (A2/B2/G3)	7	MAP3K2 (MEKK2)	7
TBK1	6	CDK1/cyclin B	6
STK25 (YSK1)	6	EEF2K	6
SRPK2	6	DAPK1	6
CDK9/cyclin K	6	TEK (TIE2) R849W	6
STK33	6	EGFR (ErbB1) C797S	6
DCAMKL2 (DCK2)	6	CAMKK1 (CAMKKA)	6
CDK17/cyclin Y	6	CSNK1A1 (CK1 alpha 1)	6
CSNK2A1 (CK2 alpha 1)	6	MAPK13 (p38 delta)	6
MAPK11 (p38 beta)	5	DYRK1A	6
TAOK2 (TAO1)	5	HUNK	5
ICK	5	KIT T670E	5
PDGFRB (PDGFR beta)	5	JAK2 JH1 JH2 V617F	5
EPHA5	5	TESK2	5
PDGFRA (PDGFR alpha)	5	PIK3C2A (PI3K–C2 alpha)	5
KIT V654A	4	PIK3CB/PIK3R2 (p110 beta/p85 beta)	4
FRAP1 (mTOR)	4	GRK7	4
NEK4	3	KIT T670I	4
PRKCI (PKC iota)	3	WNK3	3
BRSK2	3	CDC42 BPA (MRCKA)	3
SRPK1	3	PI4KA (PI4K alpha)	3
DMPK	3	ERBB2 (HER2)	3
FYN A	2	KIT V559D V654A	2
PDGFRA V561D	2	ZAP70	2
PAK1	2	INSRR (IRR)	2
PIK3C3 (hVPS34)	2	CDC42 BPB (MRCKB)	2
KIT V560G	2	PIP4K2A	2

TEK (TIE2) Y1108F	2	PDGFRA T674I	2
DYRK4	2	DDR1	2
KIT V559D	1	NEK9	2
CSNK2A2 (CK2 alpha 2)	1	KIT	1
EGFR (ErbB1) L858R	1	FLT1 (VEGFR1)	1
NIM1K	1	DAPK3 (ZIPK)	1
ACVR1B (ALK4)	1	CDC42 BPG (MRCKG)	1
EGFR (ErbB1) L861Q	1	HIPK3 (YAK1)	1
MAPK14 (p38 alpha) Direct	–1	CDK4/cyclin D3	1
CDKL5	–1	CSF1R (FMS)	–1
DNA–PK	–2	MKNK1 (MNK1)	–2
CDK18/cyclin Y	–2	PIM1	–2
LTK (TYK1)	–2	PAK3	–2
EPHA3	–3	PIP5K1C	–2
PIK3CG (p110 gamma)	–4	MAPK12 (p38 gamma)	–3
CDK6/cyclin D1	–5	DCAMKL1 (DCLK1)	–4
PIP5K1A	–5	PKN1 (PRK1)	–5
PIM3	–7	RIPK2	–6
LIMK1	–7	PRKCE (PKC epsilon)	–7
CDK9 (Inactive)	–8	WNK2	–7
PRKCZ (PKC zeta)	–10	IKBKE (IKK epsilon)	–10
GRK5	–13	PIM2	–12
SRC N1	–22	MATK (HYL)	–19
SRC	–26	IKBKB (IKK beta)	–97

References

(1) Andersen, R. J.; Williams, D. E. In *Chemistry in the Marine Environment*; Hester, R. E., Harrison, R. M., Eds.; The Royal Society of Chemistry: 2000; Vol. 13, p 55-80.
(2) Cragg, G. M.; Newman, D. J. *BBA Gen. Subjects* **2013**, *1830*, 3670-3695.
(3) Faulkner, D. J. *Nat. Prod. Rep.* **1984**, *1*, 551-598.
(4) Faulkner, D. J. *Nat. Prod. Rep.* **1986**, *3*, 1-33.
(5) Faulkner, D. J. *Nat. Prod. Rep.* **1987**, *4*, 539-576.
(6) Faulkner, D. J. *Nat. Prod. Rep.* **1988**, *5*, 613-663.
(7) Blunt, J. W.; Copp, B. R.; Hu, W.-P.; Munro, M. H. G.; Northcote, P. T.; Prinsep, M. R. *Nat. Prod. Rep.* **2008**, *25*, 35-94.
(8) Blunt, J. W.; Carroll, A. R.; Copp, B. R.; Davis, R. A.; Keyzers, R. A.; Prinsep, M. R. *Nat. Prod. Rep.* **2018**, *35*, 8-53.
(9) MarinLit Database http://pubs.rsc.org/marinlit/ (Accessed December, 2019)
(10) Carroll, A. R.; Copp, B. R.; Davis, R. A.; Keyzers, R. A.; Prinsep, M. R. *Nat. Prod. Rep.* **2020**, *37*, 175-223.
(11) Carroll, A. R.; Copp, B. R.; Davis, R. A.; Keyzers, R. A.; Prinsep, M. R. *Nat. Prod. Rep.* **2019**, *36*, 122-173.
(12) Blunt, J. W.; Copp, B. R.; Keyzers, R. A.; Munro, M. H. G.; Prinsep, M. R. *Nat. Prod. Rep.* **2017**, *34*, 235-294.
(13) Blunt, J. W.; Copp, B. R.; Keyzers, R. A.; Munro, M. H. G.; Prinsep, M. R. *Nat. Prod. Rep.* **2016**, *33*, 382-431.
(14) Blunt, J. W.; Copp, B. R.; Keyzers, R. A.; Munro, M. H. G.; Prinsep, M. R. *Nat. Prod. Rep.* **2015**, *32*, 116-211.
(15) Blunt, J. W.; Copp, B. R.; Keyzers, R. A.; Munro, M. H. G.; Prinsep, M. R. *Nat. Prod. Rep.* **2014**, *31*, 160-258.
(16) Blunt, J. W.; Copp, B. R.; Munro, M. H. G.; Northcote, P. T.; Prinsep, M. R. *Nat. Prod. Rep.* **2010**, *27*, 165-237.
(17) Blunt, J. W.; Copp, B. R.; Munro, M. H. G.; Northcote, P. T.; Prinsep, M. R. *Nat. Prod. Rep.* **2006**, *23*, 26-78.
(18) Blunt, J. W.; Copp, B. R.; Munro, M. H. G.; Northcote, P. T.; Prinsep, M. R. *Nat. Prod. Rep.* **2004**, *21*, 1-49.
(19) Newman, D. J. *Expert Opin. Drug Dis.* **2019**, *14*, 521-525.
(20) Taufa, T.; Gordon, R. M. A.; Hashmi, M. A.; Hira, K.; Miller, J. H.; Lein, M.; Fromont, J.; Northcote, P. T.; Keyzers, R. A. *Tetrahedron Lett.* **2019**, *60*, 1825-1829.
(21) Taufa, T.; Singh, A. J.; Harland, C. R.; Patel, V.; Jones, B.; Halafihi, T. i.; Miller, J. H.; Keyzers, R. A.; Northcote, P. T. *J. Nat. Prod.* **2018**, *81*, 2539-2544.
(22) Woolly, E. F.; Singh, A. J.; Russell, E. R.; Miller, J. H.; Northcote, P. T. *J. Nat. Prod.* **2018**, *81*, 387-393.
(23) Damodaran, V.; Ryan, J. L.; Keyzers, R. A. *J. Nat. Prod.* **2013**, *76*, 1997-2001.
(24) Delsuc, F.; Brinkmann, H.; Chourrout, D.; Philippe, H. *Nature* **2006**, *439*, 965.
(25) Tsagkogeorga, G.; Turon, X.; Hopcroft, R. R.; Tilak, M.-K.; Feldstein, T.; Shenkar, N.; Loya, Y.; Huchon, D.; Douzery, E. J.; Delsuc, F. *BMC Evol. Biol.* **2009**, *9*, 187.
(26) Holland, L. Z. *Curr. Biol.* **2016**, *26*, R146-R152.
(27) Figuerola, B.; Avila, C. *Mar. Drugs* **2019**, *17*, 477.
(28) Fortunato, H. *Reg. Stud. Mar. Sci.* **2015**, *2*, 32-44.
(29) Singh, A. J., The Structure-directed Isolation of New Secondary Metabolites from South Pacific Marine Sponges, PhD Thesis, University of Wellington, New Zealand, **2012**.
(30) Ruzicka, R.; Gleason, D. F. *Oecologia* **2008**, *154*, 785-794.
(31) Bolser, R. C.; Hay, M. E. *Ecology* **1996**, *77*, 2269-2286.

(32) Freestone, A. L.; Osman, R. W.; Ruiz, G. M.; Torchin, M. E. *Ecology* **2011**, *92*, 983-993.
(33) Palanisamy, S. K.; Rajendran, N. M.; Marino, A. *Nat. Prod. Bioprospect.* **2017**, *7*, 1-111.
(34) Bryan, P.; McClintock, J.; Slattery, M.; Rittschof, D. *Biofouling* **2003**, *19*, 235-245.
(35) Davis, A. R. *Mar. Biol.* **1991**, *111*, 375-379.
(36) Woolner, V. H., New Halogenated Secondary Metabolites from Red Algae of the South Pacific, PhD Thesis, Victoria Univesity of Wellington, New Zealand, **2017**.
(37) Harizani, M.; Ioannou, E.; Roussis, V. In *Progress in the Chemistry of Organic Natural Products*; Kinghorn, A. D., Falk, H., Gibbons, S., Kobayashi, J. i., Eds.; Springer: Cham, Switzerland, 2016; Vol. 102, p 91-252.
(38) Lee, J.-C.; Hou, M.-F.; Huang, H.-W.; Chang, F.-R.; Yeh, C.-C.; Tang, J.-Y.; Chang, H.-W. *Cancer Cell Int.* **2013**, *13*, 55.
(39) Kwak, J.-Y. *Mar. Drugs* **2014**, *12*, 851-870.
(40) Hamann, M. T.; Scheuer, P. J. *J. Am. Chem. Soc.* **1993**, *115*, 5825-5826.
(41) Zan, J.; Li, Z.; Tianero, M. D.; Davis, J.; Hill, R. T.; Donia, M. S. *Science* **2019**, *364*, eaaw6732.
(42) Alves, C.; Silva, J.; Pinteus, S.; Gaspar, H.; Alpoim, M. C.; Botana, L. M.; Pedrosa, R. *Front. Pharmacol.* **2018**, *9*.
(43) Wang, B.; Waters, A. L.; Valeriote, F. A.; Hamann, M. T. *Biochim. Biophys. Acta* **2015**, *1850*, 1849-1854.
(44) Cruz, L. J.; Luque-Ortega, J. R.; Rivas, L.; Albericio, F. *Mol. Pharm.* **2009**, *6*, 813-824.
(45) Mori, T.; O'Keefe, B. R.; Sowder, R. C.; Bringans, S.; Gardella, R.; Berg, S.; Cochran, P.; Turpin, J. A.; Buckheit, R. W.; McMahon, J. B.; Boyd, M. R. *J. Biol. Chem.* **2005**, *280*, 9345-9353.
(46) Lusvarghi, S.; Bewley, C. A. *Viruses* **2016**, *8*, 296.
(47) Griffithsin-based Rectal Microbicide for PREvention of Viral ENTry (PREVENT) https://ClinicalTrials.gov/show/NCT04032717 (Accessed February, 2020)
(48) Stratton, C. F.; Newman, D. J.; Tan, D. S. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 4802-4807.
(49) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2020**, *83*, 770-803.
(50) Marine Clincal Pipeline http://marinepharmacology.midwestern.edu/clinical_pipeline.html (Accessed February, 2020)
(51) Bergmann, W.; Burke, D. C. *J. Org. Chem.* **1956**, *21*, 226-228.
(52) Yeung, B. K. S. *Curr. Opin. Chem. Biol.* **2011**, *15*, 523-528.
(53) Rinehart, K. L.; Holt, T. G.; Fregeau, N. L.; Stroh, J. G.; Keifer, P. A.; Sun, F.; Li, L. H.; Martin, D. G. *J. Org. Chem.* **1990**, *55*, 4512-4515.
(54) Cuevas, C.; Pérez, M.; Martín, M. J.; Chicharro, J. L.; Fernández-Rivas, C.; Flores, M.; Francesch, A.; Gallego, P.; Zarzuelo, M.; de la Calle, F.; García, J.; Polanco, C.; Rodríguez, I.; Manzanares, I. *Org. Lett.* **2000**, *2*, 2545-2548.
(55) Corey, E. J.; Gin, D. Y.; Kania, R. S. *J. Am. Chem. Soc.* **1996**, *118*, 9202-9203.
(56) Barone, A.; Chi, D.-C.; Theoret, M. R.; Chen, H.; He, K.; Kufrin, D.; Helms, W. S.; Subramaniam, S.; Zhao, H.; Patel, A.; Goldberg, K. B.; Keegan, P.; Pazdur, R. *Clin. Cancer Res.* **2017**, *23*, 7448-7453.
(57) He, W.; Zhang, Z.; Ma, D. *Angew. Chem. Int. Ed.* **2019**, *58*, 3972-3975.
(58) Belgiovine, C.; Bello, E.; Liguori, M.; Craparotta, I.; Mannarino, L.; Paracchini, L.; Beltrame, L.; Marchini, S.; Galmarini, C. M.; Mantovani, A.; Frapolli, R.; Allavena, P.; D'Incalci, M. *Brit. J. Cancer* **2017**, *117*, 628.
(59) Pereira, R. B.; Evdokimov, N. M.; Lefranc, F.; Valentão, P.; Kornienko, A.; Pereira, D. M.; Andrade, P. B.; Gomes, N. G. M. *Mar. Drugs* **2019**, *17*, 329.

(60) Alonso-Álvarez, S.; Pardal, E.; Sánchez-Nieto, D.; Navarro, M.; Caballero, M. D.; Mateos, M. V.; Martín, A. *Drug Des. Devel. Ther.* **2017**, *11*, 253-264.
(61) Rinehart, K.; Lithgow-Berelloni, A. *PCT Int. Pat. Appl. WO* **1991**, *91*, 18.
(62) Sakai, R.; Rinehart, K. L.; Kishore, V.; Kundu, B.; Faircloth, G.; Gloer, J. B.; Carney, J. R.; Namikoshi, M.; Sun, F.; Hughes, R. G.; Grávalos, D. G.; de Quesada, T. G.; Wilson, G. R.; Heid, R. M. *J. Med. Chem.* **1996**, *39*, 2819-2834.
(63) Leisch, M.; Egle, A.; Greil, R. *Future Oncol.* **2019**, *15*, 109-120.
(64) McIntosh, M.; Cruz, L. J.; Hunkapiller, M. W.; Gray, W. R.; Olivera, B. M. *Arch. Biochem. Biophys.* **1982**, *218*, 329-334.
(65) McGregor, B. A.; Sonpavde, G. *Expert Opin. Inv. Drug* **2019**, *28*, 821-826.
(66) Pettit, G. R.; Kamano, Y.; Herald, C. L.; Fujii, Y.; Kizu, H.; Boyd, M. R.; Boettner, F. E.; Doubek, D. L.; Schmidt, J. M.; Chapuis, J.-C.; Michel, C. *Tetrahedron* **1993**, *49*, 9151-9170.
(67) Koski, R. R. *Pharm. Ther.* **2008**, *33*, 271-303.
(68) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3-26.
(69) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3-25.
(70) Taufa, T., Natural Products Analysis of South Pacific Marine Sponges, PhD Thesis, Victoria Univesity of Wellington, New Zealand, **2018**.
(71) Blunt, J. W.; Munro, M. H. G. *Phytochem. Rev.* **2013**, *12*, 435-447.
(72) Buedenbender, L.; Habener, L. J.; Grkovic, T.; Kurtböke, D. İ.; Duffy, S.; Avery, V. M.; Carroll, A. R. *J. Nat. Prod.* **2018**, *81*, 957-965.
(73) Dictionary of Natural Products http://dnp.chemnetbase.com/ (Accessed February, 2020)
(74) Zani, C. L.; Carroll, A. R. *J. Nat. Prod.* **2017**, *80*, 1758-1766.
(75) Zhang, C.; Idelbayev, Y.; Roberts, N.; Tao, Y.; Nannapaneni, Y.; Duggan, B. M.; Min, J.; Lin, E. C.; Gerwick, E. C.; Cottrell, G. W.; Gerwick, W. H. *Sci. Rep.* **2017**, *7*, 14243.
(76) Spicer, R. A.; Steinbeck, C. *Metabolomics* **2017**, *14*, 16.
(77) 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.; Crusemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderon, 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 P, C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodriguez, 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.; Muller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. O.; Pogliano, K.; Linington, R. G.; Gutierrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. *Nat. Biotech.* **2016**, *34*, 828-837.

(78) Hartmann, A. C.; Petras, D.; Quinn, R. A.; Protsyuk, I.; Archer, F. I.; Ransome, E.; Williams, G. J.; Bailey, B. A.; Vermeij, M. J. A.; Alexandrov, T.; Dorrestein, P. C.; Rohwer, F. L. *Proc. Natl. Acad. Sci.* **2017**, *114*, 11685-11690.
(79) Quinn, R. A.; Nothias, L.-F.; Vining, O.; Meehan, M.; Esquenazi, E.; Dorrestein, P. C. *Trends Pharmacol. Sci.* **2017**, *38*, 143-154.
(80) Fox Ramos, A. E.; Evanno, L.; Poupon, E.; Champy, P.; Beniddir, M. A. *Nat. Prod. Rep.* **2019**, *36*, 960-980.
(81) Philippus, A. C.; Zatelli, G. A.; Wanke, T.; Gabriela de A. Barros, M.; Kami, S. A.; Lhullier, C.; Armstrong, L.; Sandjo, L. P.; Falkenberg, M. *RSC Adv.* **2018**, *8*, 29654-29661.
(82) Bauvais, C.; Bonneau, N.; Blond, A.; Pérez, T.; Bourguet-Kondracki, M.-L.; Zirah, S. *Metabolites* **2017**, *7*, 27.
(83) Crüsemann, M.; O'Neill, E. C.; Larson, C. B.; Melnik, A. V.; Floros, D. J.; da Silva, R. R.; Jensen, P. R.; Dorrestein, P. C.; Moore, B. S. *J. Nat. Prod.* **2017**, *80*, 588-597.
(84) Nothias, L.-F.; Nothias-Esposito, M.; da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; Paolini, J.; Alexandrov, T.; Litaudon, M.; Dorrestein, P. C. *J. Nat. Prod.* **2018**, *81*, 758-767.
(85) Esposito, M.; Nothias, L.-F.; Nedev, H.; Gallard, J.-F.; Leyssen, P.; Retailleau, P.; Costa, J.; Roussi, F.; Iorga, B. I.; Paolini, J.; Litaudon, M. *J. Nat. Prod.* **2016**, *79*, 2873-2882.
(86) Watrous, J.; Roach, P.; Alexandrov, T.; Heath, B. S.; Yang, J. Y.; Kersten, R. D.; van der Voort, M.; Pogliano, K.; Gross, H.; Raaijmakers, J. M.; Moore, B. S.; Laskin, J.; Bandeira, N.; Dorrestein, P. C. *Proc. Natl. Acad. Sci.* **2012**, *109*, E1743–E1752.
(87) Nguyen, D. D.; Melnik, A. V.; Koyama, N.; Lu, X.; Schorn, M.; Fang, J.; Aguinaldo, K.; Lincecum Jr, T. L.; Ghequire, M. G. K.; Carrion, V. J.; Cheng, T. L.; Duggan, B. M.; Malone, J. G.; Mauchline, T. H.; Sanchez, L. M.; Kilpatrick, A. M.; Raaijmakers, J. M.; De Mot, R.; Moore, B. S.; Medema, M. H.; Dorrestein, P. C. *Nat. Microbiol.* **2016**, *2*, 16197.
(88) von Eckardstein, L.; Petras, D.; Dang, T.; Cociancich, S.; Sabri, S.; Grätz, S.; Kerwat, D.; Seidel, M.; Pesic, A.; Dorrestein, P. C.; Royer, M.; Weston, J. B.; Süssmuth, R. D. *Chem. Eur. J.* **2017**, *23*, 15316-15321.
(89) Lorig-Roach, N.; Still, P. C.; Coppage, D.; Compton, J. E.; Crews, M. S.; Navarro, G.; Tenney, K.; Crews, P. *J. Nat. Prod.* **2017**, *80*, 2304-2310.
(90) Olivon, F.; Apel, C.; Retailleau, P.; Allard, P. M.; Wolfender, J. L.; Touboul, D.; Roussi, F.; Litaudon, M.; Desrat, S. *Org. Chem. Front.* **2018**, *5*, 2171-2178.
(91) Otogo N'Nang, E.; Bernadat, G.; Mouray, E.; Kumulungui, B.; Grellier, P.; Poupon, E.; Champy, P.; Beniddir, M. A. *Org. Lett.* **2018**, *20*, 6596-6600.
(92) Nothias-Esposito, M.; Nothias, L. F.; Da Silva, R. R.; Retailleau, P.; Zhang, Z.; Leyssen, P.; Roussi, F.; Touboul, D.; Paolini, J.; Dorrestein, P. C.; Litaudon, M. *J. Nat. Prod.* **2019**, *82*, 1459-1470.
(93) Nothias, L.-F.; Nothias-Esposito, M.; da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; Paolini, J.; Alexandrov, T.; Litaudon, M.; Dorrestein, P. C. *J. Nat. Prod.* **2018**.
(94) Remy, S.; Olivon, F.; Desrat, S.; Blanchard, F.; Eparvier, V.; Leyssen, P.; Neyts, J.; Roussi, F.; Touboul, D.; Litaudon, M. *J. Nat. Prod.* **2018**.
(95) Nothias, L.-F.; Boutet-Mercey, S.; Cachet, X.; De La Torre, E.; Laboureur, L.; Gallard, J.-F.; Retailleau, P.; Brunelle, A.; Dorrestein, P. C.; Costa, J.; Bedoya, L. M.; Roussi, F.; Leyssen, P.; Alcami, J.; Paolini, J.; Litaudon, M.; Touboul, D. *J. Nat. Prod.* **2017**, *80*, 2620-2629.
(96) Remy, S.; Olivon, F.; Desrat, S.; Blanchard, F.; Eparvier, V.; Leyssen, P.; Neyts, J.; Roussi, F.; Touboul, D.; Litaudon, M. *J. Nat. Prod.* **2018**, *81*, 901-912.

(97) Naman, C. B.; Rattan, R.; Nikoulina, S. E.; Lee, J.; Miller, B. W.; Moss, N. A.; Armstrong, L.; Boudreau, P. D.; Debonsi, H. M.; Valeriote, F. A.; Dorrestein, P. C.; Gerwick, W. H. *J. Nat. Prod.* **2017**, *80*, 625-633.
(98) Zhang, C.; Naman, B. C.; Engene, N.; Gerwick, H. W. *Mar. Drugs* **2017**, *15*.
(99) Tao, Y.; Li, P.; Zhang, D.; Glukhov, E.; Gerwick, L.; Zhang, C.; Murray, T. F.; Gerwick, W. H. *J. Org. Chem.* **2018**, *83*, 3034-3046.
(100) Kristoffersen, V.; Rämä, T.; Isaksson, J.; Andersen, J.; Gerwick, W.; Hansen, E. *Mar. Drugs* **2018**, *16*, 163.
(101) Li, F.; Janussen, D.; Peifer, C.; Pérez-Victoria, I.; Tasdemir, D. *Mar. Drugs* **2018**, *16*, 268.
(102) Kalinski, J.-C. J.; Waterworth, S. C.; Siwe Noundou, X.; Jiwaji, M.; Parker-Nance, S.; Krause, R. W. M.; McPhail, K. L.; Dorrington, R. A. *Mar. Drugs* **2019**, *17*, 60.
(103) Prasad, P.; Salim, A. A.; Khushi, S.; Khalil, Z. G.; Quezada, M.; Capon, R. J. *Mar. Drugs* **2019**, *17*, 106.
(104) Bonneau, N.; Chen, G.; Lachkar, D.; Boufridi, A.; Gallard, J.-F.; Retailleau, P.; Petek, S.; Debitus, C.; Evanno, L.; Beniddir, M. A.; Poupon, E. *Chem. Eur. J.* **2017**, *23*, 14454-14461.
(105) Bertin, M. J.; Schwartz, S. L.; Lee, J.; Korobeynikov, A.; Dorrestein, P. C.; Gerwick, L.; Gerwick, W. H. *J. Nat. Prod.* **2015**, *78*, 493-499.
(106) Sung, A. A.; Gromek, S. M.; Balunas, M. J. *Mar. Drugs* **2017**, *15*, 250.
(107) Bracegirdle, J.; Keyzers, R. A. *Chem. N.Z.* **2019**, *83*, 71-77.
(108) Antunes, E. M.; Copp, B. R.; Davies-Coleman, M. T.; Samaai, T. *Nat. Prod. Rep.* **2005**, *22*, 62-72.
(109) Brinkmann, C. M.; Marker, A.; Kurtböke, D. İ. *Diversity* **2017**, *9*, 40.
(110) Heavisides, E.; Rouger, C.; Reichel, A. F.; Ulrich, C.; Wenzel-Storjohann, A.; Sebens, S.; Tasdemir, D. *Mar. Drugs* **2018**, *16*, 503.
(111) Fan, B.; Parrot, D.; Blümel, M.; Labes, A.; Tasdemir, D. *Mar. Drugs* **2019**, *17*, 67.
(112) Kersten, R. D.; Lee, S.; Fujita, D.; Pluskal, T.; Kram, S.; Smith, J. E.; Iwai, T.; Noel, J. P.; Fujita, M.; Weng, J.-K. *J. Am. Chem. Soc.* **2017**, *139*, 16838-16844.
(113) Li, Y.; Yu, H.-B.; Zhang, Y.; Leao, T.; Glukhov, E.; Pierce, M. L.; Zhang, C.; Kim, H.; Mao, H. H.; Fang, F.; Cottrell, G. W.; Murray, T. F.; Gerwick, L.; Guan, H.; Gerwick, W. H. *J. Nat. Prod.* **2020**, *83*, 617-625.
(114) Dasyam, N., Identification of Anti-tubercular Compounds in Marine Organisms from Aotearoa, PhD Thesis, Victoria University of Wellington, New Zealand, **2014**.
(115) Husson, R. N. In *Mycobacteria Protocols*; Parish, T., Stoker, N. G., Eds.; Humana Press: Totowa, NJ, 1998, p 199-206.
(116) Carroll, A. R.; Cooray, N. M.; Poiner, A.; Scheuer, P. J. *J. Org. Chem.* **1989**, *54*, 4231-4232.
(117) Carroll, A. R.; Scheuer, P. J. *J. Org. Chem.* **1990**, *55*, 4426-4431.
(118) Bontemps, N.; Bry, D.; López-Legentil, S.; Simon-Levert, A.; Long, C.; Banaigs, B. *J. Nat. Prod.* **2010**, *73*, 1044-1048.
(119) Bry, D.; Banaigs, B.; Long, C.; Bontemps, N. *Tetrahedron Lett.* **2011**, *52*, 3041-3044.
(120) Appleton, D. R.; Pearce, A. N.; Copp, B. R. *Tetrahedron* **2010**, *66*, 4977-4986.
(121) Keyzers, R. A., The Isolation of Biologically Active Secondary Metabolites from New Zealand Marine Organisms, PhD Thesis, Victoria University of Wellington, New Zealand, **2003**.
(122) Ryan, J. M., Novel Secondary Metabolites from New Zealand Marine Sponges, PhD Thesis, Victoria University of Wellington, New Zealand, **2007**.
(123) Aitken, E. J., Research Project Thesis, Victoria University of Wellington, **2013**.
(124) Harland, C. R., Chemical investigations of Pacific Marine invertebrates, Master's Thesis, Victoria University of Wellington, New Zealand, **2017**.

(125) Cikoš, A.-M.; Jurin, M.; Čož-Rakovac, R.; Jokić, S.; Jerković, I. *Mar. Drugs* **2019**, *17*, 537.
(126) Qureshi, A.; Faulkner, D. J. *Nat. Prod. Lett.* **1999**, *13*, 59-62.
(127) Bromley, C. L.; Parker-Nance, S.; de la Mare, J.-A.; Edkins, A. L.; Beukes, D. R.; Davies-Coleman, M. T. *S. Afr. J. Chem.* **2013**, *66*, 00-00.
(128) McClintock, J. B.; Amsler, M. O.; Amsler, C. D.; Southworth, K. J.; Petrie, C.; Baker, B. J. *Mar. Biol.* **2004**, *145*, 885-894.
(129) Arabshahi, L.; J. Schmitz, F. *Tetrahedron Lett.* **1988**, *29*, 1099-1102.
(130) Heumüller, S.; Wind, S.; Barbosa-Sicard, E.; Schmidt, H. H. H. W.; Busse, R.; Schröder, K.; Brandes, R. P. *Hypertension* **2008**, *51*, 211-217.
(131) Dunaway, S.; Odin, R.; Zhou, L.; Ji, L.; Zhang, Y.; Kadekaro, A. L. *Front. Pharmacol.* **2018**, *9*.
(132) Stefanska, J.; Pawliczak, R. *Mediators Inflamm.* **2008**, *2008*, 106507-106507.
(133) Cabrita, M. T.; Vale, C.; Rauter, A. P. *Mar. Drugs* **2010**, *8*, 2301-2317.
(134) Popplewell, W., L., Isolation and Structure Elucidation of New Secondary Metabolites from New Zealand Marine Red Algae, PhD Thesis, Victoria University of Wellington, New Zealand, **2008**.
(135) Woolner, V. H.; Gordon, R. M. A.; Miller, J. H.; Lein, M.; Northcote, P. T.; Keyzers, R. A. *J. Nat. Prod.* **2018**, *81*, 2446-2454.
(136) Woolner, V. H.; Jones, C. M.; Field, J. J.; Fadzilah, N. H.; Munkacsi, A. B.; Miller, J. H.; Keyzers, R. A.; Northcote, P. T. *J. Nat. Prod.* **2016**, *79*, 463-469.
(137) Mynderse, J. S.; Faulkner, D. J. *J. Am. Chem. Soc.* **1974**, *96*, 6771-6772.
(138) Crews, P.; Kho, E. *J. Org. Chem.* **1974**, *39*, 3303-3304.
(139) Paul, V. J.; Hay, M. E.; Duffy, J. E.; Fenical, W.; Gustafson, K. *J. Exp. Mar. Biol. Ecol.* **1988**, *114*, 249-260.
(140) Paul, V. J.; McConnell, O. J.; Fenical, W. *J. Org. Chem.* **1980**, *45*, 3401-3407.
(141) Paradas, W. C.; Crespo, T. M.; Salgado, L. T.; de Andrade, L. R.; Soares, A. R.; Hellio, C.; Paranhos, R. R.; Hill, L. J.; de Souza, G. M.; Kelecom, A. G. A. C.; Da Gama, B. A. P.; Pereira, R. C.; Amado-Filho, G. M. *J. Phycol.* **2015**, *51*, 225-235.
(142) Carter-Franklin, J. N.; Parrish, J. D.; Tschirret-Guth, R. A.; Little, R. D.; Butler, A. *J. Am. Chem. Soc.* **2003**, *125*, 3688-3689.
(143) Naylor, S.; Hanke, F. J.; Manes, L. V.; Crews, P. In *Progress in the Chemistry of Organic Natural Products*; Buchanan, J. G., Crews, P., Epe, B., Evans, F. J., Hanke, F. J., Manes, L. V., Mondon, A., Naylor, S., Taylor, S. E., Eds.; Springer Vienna: Vienna, 1983, p 189-241.
(144) Barrow, K. D.; Temple, C. A. *Phytochem.* **1985**, *24*, 1697-1704.
(145) Stierle, D. B.; Wing, R. M.; Sims, J. J. *Tetrahedron Lett.* **1976**, *17*, 4455-4458.
(146) Brownlee, R. T. C.; Hall, J. G.; Reiss, J. A. *Org. Magn. Reson.* **1983**, *21*, 544-547.
(147) Motti, C. A.; Thomas-Hall, P.; Hagiwara, K. A.; Simmons, C. J.; Willis, R.; Wright, A. D. *J. Nat. Prod.* **2014**, *77*, 1193-1200.
(148) Cooper, M. W., A Taxonomic Investigation into the Red Alga Plocamium within New Zealand, Master's Thesis, Victoria University of Wellington, New Zealand, **2017**.
(149) Dunlop, R. W.; Murphy, P. T.; Wells, R. J. *Aust. J. Chem.* **1979**, *32*, 2735-2739.
(150) Stierle, D. B.; Sims, J. J. *Tetrahedron Lett.* **1984**, *25*, 153-156.
(151) Naylor, S.; Manes, L. V.; Crews, P. *J. Nat. Prod.* **1985**, *48*, 72-75.
(152) Timmers, M. A.; Dias, D. A.; Urban, S. *Mar. Drugs* **2012**, *10*, 2089-2102.
(153) Bucher, C.; Deans, R. M.; Burns, N. Z. *J. Am. Chem. Soc.* **2015**, *137*, 12784-12787.
(154) Wynne, M. J. *N. Z. J. Bot.* **2002**, *40*, 137-142.
(155) Harvey, J. E. *Chem 305 Laboratory Manual*; Victoria University of Wellington, 2019.
(156) Cianciola, E. N. *Diversity* **2010**, *v. 2*, pp. 946-958-2010 v.2012 no.2016.

(157) Saunders, G. W.; Virginia Lehmkuhl, K. *Eur. J. Phycol.* **2005**, *40*, 293-312.
(158) Adams, N. M. *Seaweeds of New Zealand*; Canterbury University Press: Christchurch, New Zealand, 1994.
(159) Bates, P.; Blunt, J. W.; Hartshorn, M. P.; Jones, A. J.; Munro, M. H. G.; Robinson, W. T.; Yorke, S. C. *Aust. J. Chem.* **1979**, *32*, 2545-2554.
(160) Blunt, J. W.; Hartshorn, M. P.; Munro, M. H. G.; Yorke, S. C. *Tetrahedron Lett.* **1978**, *19*, 4417-4418.
(161) Aliev, A. E.; Harris, K. D. M. *Magn. Reson. Chem.* **1993**, *31*, 54-57.
(162) Sergeyev, N. M.; Sandor, P.; Sergeyeva, N. D.; Raynes, W. T. *J. Magn. Res., Ser. A* **1995**, *115*, 174-182.
(163) Araki, Y.; Awakawa, T.; Matsuzaki, M.; Cho, R.; Matsuda, Y.; Hoshino, S.; Shinohara, Y.; Yamamoto, M.; Kido, Y.; Inaoka, D. K.; Nagamune, K.; Ito, K.; Abe, I.; Kita, K. *Proc. Natl. Acad. Sci.* **2019**, *116*, 8269-8274.
(164) Polzin, J. P.; Rorrer, G. L. *Biotechnol. Bioeng.* **2003**, *82*, 415-428.
(165) Polzin, J. J.; Rorrer, G. L.; Cheney, D. P. *Biomol. Eng.* **2003**, *20*, 205-215.
(166) Winter, J. M.; Moore, B. S. *J. Biol. Chem.* **2009**, *284*, 18577-18581.
(167) Dale, J. A.; Dull, D. L.; Mosher, H. S. *J. Org. Chem.* **1969**, *34*, 2543-2549.
(168) Seco, J. M.; Quiñoá, E.; Riguera, R. *Chem. Rev.* **2004**, *104*, 17-118.
(169) Bracegirdle, J.; Sohail, Z.; Fairhurst, M. J.; Gerth, M. L.; Zuccarello, G. C.; Hashmi, M. A.; Keyzers, R. A. *Mar. Drugs* **2019**, *17*, 418.
(170) Sabry, O. M. M.; Goeger, D. E.; Valeriote, F. A.; Gerwick, W. H. *Nat. Prod. Res.* **2017**, *31*, 261-267.
(171) Fuller, R. W.; Cardellina, J. H.; Kato, Y.; Brinen, L. S.; Clardy, J.; Snader, K. M.; Boyd, M. R. *J. Med. Chem.* **1992**, *35*, 3007-3011.
(172) Blunt, J. W.; Bowman, N. J.; Munro, M. H. G.; Parsons, M. J.; Wright, G. J.; Yeow, K. K. *Aust. J. Chem.* **1985**, *38*, 519-525.
(173) Pettit, G. R.; Herald, C. L.; Doubek, D. L.; Herald, D. L.; Arnold, E.; Clardy, J. *J. Am. Chem. Soc.* **1982**, *104*, 6846-6848.
(174) Yu, H.-B.; Yang, F.; Li, Y.-Y.; Gan, J.-H.; Jiao, W.-H.; Lin, H.-W. *J. Nat. Prod.* **2015**, *78*, 1169-1173.
(175) Hornung, R. L.; Pearson, J. W.; Beckwith, M.; Longo, D. L. *Cancer Res.* **1992**, *52*, 101-107.
(176) Mayer, A. M. S.; Glaser, K. B.; Cuevas, C.; Jacobs, R. S.; Kem, W.; Little, R. D.; McIntosh, J. M.; Newman, D. J.; Potts, B. C.; Shuster, D. E. *Trends Pharmacol. Sci.* **2010**, *31*, 255-265.
(177) Mehla, R.; Bivalkar-Mehla, S.; Zhang, R.; Handy, I.; Albrecht, H.; Giri, S.; Nagarkatti, P.; Nagarkatti, M.; Chauhan, A. *PLoS One* **2010**, *5*, e11160.
(178) Nelson, T. J.; Sun, M.-K.; Lim, C.; Sen, A.; Khan, T.; Chirila, F. V.; Alkon, D. L. *J. Alzheimers Dis.* **2017**, *58*, 521-535.
(179) Schaufelberger, D. E.; Koleck, M. P.; Beutler, J. A.; Vatakis, A. M.; Alvarado, A. B.; Andrews, P.; Marzo, L. V.; Muschik, G. M.; Roach, J.; Ross, J. T.; Lebherz, W. B.; Reeves, M. P.; Eberwein, R. M.; Rodgers, L. L.; Testerman, R. P.; Snader, K. M.; Forenza, S. *J. Nat. Prod.* **1991**, *54*, 1265-1270.
(180) Wu, R.; Chen, H.; Chang, N.; Xu, Y.; Jiao, J.; Zhang, H. *Chem. Eur. J.* **2019**, *0*.
(181) Sharp, J. H.; Winson, M. K.; Porter, J. S. *Nat. Prod. Rep.* **2007**, *24*, 659-673.
(182) Anthoni, U.; Larsen, C.; Nielsen, P. H.; Christophersen, C.; Lidgren, G. *Comp. Biochem. Physiol. B* **1989**, *92*, 711-713.
(183) Tian, X.-R.; Tang, H.-F.; Feng, J.-T.; Li, Y.-S.; Lin, H.-W.; Fan, X.-P.; Zhang, X. *Mar. Drugs* **2014**, *12*, 1987-2003.

(184) Patiño C, L. P.; Muniain, C.; Knott, M. E.; Puricelli, L.; Palermo, J. A. *J. Nat. Prod.* **2014**, *77*, 1170-1178.
(185) Michael, P.; Hansen, K. Ø.; Isaksson, J.; Andersen, J. H.; Hansen, E. *Molecules* **2017**, *22*, 1236.
(186) Hansen, K. Ø.; Isaksson, J.; Bayer, A.; Johansen, J. A.; Andersen, J. H.; Hansen, E. *J. Nat. Prod.* **2017**, *80*, 3276-3283.
(187) Milanowski, D. J.; Gustafson, K. R.; Kelley, J. A.; McMahon, J. B. *J. Nat. Prod.* **2004**, *67*, 70-73.
(188) Shubina, L. K.; Makarieva, T. N.; Yashunsky, D. V.; Nifantiev, N. E.; Denisenko, V. A.; Dmitrenok, P. S.; Dyshlovoy, S. A.; Fedorov, S. N.; Krasokhin, V. B.; Jeong, S. H.; Han, J.; Stonik, V. A. *J. Nat. Prod.* **2015**, *78*, 1383-1389.
(189) Macherla, V. R.; Liu, J.; Bellows, C.; Teisan, S.; Nicholson, B.; Lam, K. S.; Potts, B. C. M. *J. Nat. Prod.* **2005**, *68*, 780-783.
(190) Lian, X.-Y.; Zhang, Z. *Nat. Prod. Res.* **2013**, *27*, 2161-2167.
(191) McGinnis, G. D. *Carbohydr. Res.* **1982**, *108*, 284-292.
(192) Keyzers, R. A.; Arendse, C. E.; Hendricks, D. T.; Samaai, T.; Davies-Coleman, M. T. *J. Nat. Prod.* **2005**, *68*, 506-510.
(193) Keyzers, R. A.; Samaai, T.; Davies-Coleman, M. T. *Tetrahedron Lett.* **2004**, *45*, 9415-9418.
(194) Neises, B.; Steglich, W. *Angew. Chem.* **1978**, *17*, 522-524.
(195) Thermo SelectScreens https://www.thermofisher.com/nz/en/home/industrial/pharma-biopharma/drug-discovery-development/target-and-lead-identification-and-validation/kinasebiology/kinases-disease-research.html (Accessed November, 2019)
(196) Newman, D. J. *J. Braz. Chem. Soc.* **2016**, *27*, 1320-1333.
(197) Seley-Radtke, K. L.; Yates, M. K. *Antivir. Res.* **2018**, *154*, 66-86.
(198) Niu, G.; Tan, H. *Trends Microbiol.* **2015**, *23*, 110-119.
(199) Shelton, J.; Lu, X.; Hollenbaugh, J. A.; Cho, J. H.; Amblard, F.; Schinazi, R. F. *Chem. Rev.* **2016**, *116*, 14379-14455.
(200) Bhakuni, D. S.; Rawat, D. S. In *Bioactive Marine Natural Products*; Springer Netherlands: Dordrecht, 2005; Vol. 1, p 208-234.
(201) Huang, R.-M.; Chen, Y.-N.; Zeng, Z.; Gao, C.-H.; Su, X.; Peng, Y. *Mar. Drugs* **2014**, *12*, 5817-5838.
(202) Martins, A.; Vieira, H.; Gaspar, H.; Santos, S. *Mar. Drugs* **2014**, *12*, 1066-1101.
(203) Forenza, S.; Minale, L.; Riccio, R.; Fattorusso, E. *J. Chem. Soc. D* **1971**, 1129-1130.
(204) Morales, J. J.; Rodríguez, A. D. *J. Nat. Prod.* **1991**, *54*, 629-631.
(205) Tasdemir, D.; Topaloglu, B.; Perozzo, R.; Brun, R.; O'Neill, R.; Carballeira, N. M.; Zhang, X.; Tonge, P. J.; Linden, A.; Rüedi, P. *Bioorgan. Med. Chem.* **2007**, *15*, 6834-6845.
(206) Zidar, N.; Montalvão, S.; Hodnik, Ž.; Nawrot, D. A.; Žula, A.; Ilaš, J.; Kikelj, D.; Tammela, P.; Mašič, L. P. *Mar. Drugs* **2014**, *12*, 940-963.
(207) Abdjul, D. B.; Yamazaki, H.; Kanno, S.-i.; Takahashi, O.; Kirikoshi, R.; Ukai, K.; Namikoshi, M. *J. Nat. Prod.* **2015**, *78*, 1428-1433.
(208) Cafieri, F.; Fattorusso, E.; Mangoni, A.; Taglialatela-Scafati, O.; Carnuccio, R. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 799-804.
(209) Aiello, A.; D'Esposito, M.; Fattorusso, E.; Menna, M.; Müller, W. E. G.; Perović-Ottstadt, S.; Tsuruta, H.; Gulder, T. A. M.; Bringmann, G. *Tetrahedron* **2005**, *61*, 7266-7270.
(210) Woo, S.-Y.; Win, N. N.; Wong, C. P.; Ito, T.; Hoshino, S.; Ngwe, H.; Aye, A. A.; Han, N. M.; Zhang, H.; Hayashi, F.; Abe, I.; Morita, H. *J. Nat. Med.* **2018**, *72*, 803-807.
(211) Kobayashi, J.; Kanda, F.; Ishibashi, M.; Shigemori, H. *J. Org. Chem.* **1991**, *56*, 4574-4576.

(212) Tsukamoto, S.; Tane, K.; Ohta, T.; Matsunaga, S.; Fusetani, N.; van Soest, R. W. M. *J. Nat. Prod.* **2001**, *64*, 1576-1578.
(213) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science* **2002**, *298*, 1912-1934.
(214) Cohen, P. *Nat. Rev. Drug Discov.* **2002**, *1*, 309-315.
(215) Iqbal, N.; Iqbal, N. *Chemother. Res. Pract.* **2014**, *2014*, 357027-357027.
(216) Ferguson, F. M.; Gray, N. S. *Nat. Rev. Drug Discov.* **2018**, *17*, 353.
(217) Lucet, I. S.; Tobin, A.; Drewry, D.; Wilks, A. F.; Doerig, C. *Future Med. Chem.* **2012**, *4*, 2295-2310.
(218) Gordon, S.; Simithy, J.; Goodwin, D. C.; Calderón, A. I. *Perspect. Medicin. Chem.* **2015**, *7*, PMC.S13212.
(219) Li, T.; Wang, N.; Zhang, T.; Zhang, B.; Sajeevan, T. P.; Joseph, V.; Armstrong, L.; He, S.; Yan, X.; Naman, C. B. *Mar. Drugs* **2019**, *17*, 493.
(220) Ebada, S. S.; Linh, M. H.; Longeon, A.; de Voogd, N. J.; Durieu, E.; Meijer, L.; Bourguet-Kondracki, M.-L.; Singab, A. N. B.; Müller, W. E. G.; Proksch, P. *Nat. Prod. Res.* **2015**, *29*, 231-238.
(221) Kim, D.-C.; Lee, H.-S.; Ko, W.; Lee, D.-S.; Sohn, J. H.; Yim, J. H.; Kim, Y.-C.; Oh, H. *Molecules* **2014**, *19*, 18073-18089.
(222) Youssef, D. T. A.; Mohamed, G. A.; Shaala, L. A.; Badr, J. M.; Bamanie, F. H.; Ibrahim, S. R. M. *Phytochem. Lett.* **2015**, *13*, 212-217.
(223) Sassone-Corsi, P.; Mizzen, C. A.; Cheung, P.; Crosio, C.; Monaco, L.; Jacquot, S.; Hanauer, A.; Allis, C. D. *Science* **1999**, *285*, 886-891.
(224) Ludwik, K. A.; Campbell, J. P.; Li, M.; Li, Y.; Sandusky, Z. M.; Pasic, L.; Sowder, M. E.; Brenin, D. R.; Pietenpol, J. A.; O'Doherty, G. A.; Lannigan, D. A. *Mol. Cancer Ther.* **2016**, *15*, 2598-2608.
(225) Lee, C.-J.; Moon, S.-J.; Jeong, J.-H.; Lee, S.; Lee, M.-H.; Yoo, S.-M.; Lee, H. S.; Kang, H. C.; Lee, J. Y.; Lee, W. S.; Lee, H.-J.; Kim, E.-K.; Jhun, J.-Y.; Cho, M.-L.; Min, J.-K.; Cho, Y.-Y. *Cell Death Dis.* **2018**, *9*, 401.
(226) Yao, K.; Chen, H.; Liu, K.; Langfald, A.; Yang, G.; Zhang, Y.; Yu, D. H.; Kim, M. O.; Lee, M.-H.; Li, H.; Bae, K. B.; Kim, H.-G.; Ma, W.-Y.; Bode, A. M.; Dong, Z.; Dong, Z. *Cancer Prev. Res.* **2014**, *7*, 958-967.
(227) Derer, A.; Böhm, C.; Grötsch, B.; Grün, J. R.; Grützkau, A.; Stock, M.; Böhm, S.; Sehnert, B.; Gaipl, U.; Schett, G.; Hueber, A. J.; David, J.-P. *Ann. Rheum. Dis.* **2016**, *75*, 413-421.
(228) Hetzer, C.; Bisgrove, D.; Cohen, M. S.; Pedal, A.; Kaehlcke, K.; Speyerer, A.; Bartscherer, K.; Taunton, J.; Ott, M. *PLoS One* **2007**, *2*, e151.
(229) Karijolich, J.; Zhao, Y.; Peterson, B.; Zhou, Q.; Glaunsinger, B. *J. Virol.* **2014**, *88*, 7024-7035.
(230) Lo Monte, F.; Kramer, T.; Gu, J.; Anumala, U. R.; Marinelli, L.; La Pietra, V.; Novellino, E.; Franco, B.; Demedts, D.; Van Leuven, F.; Fuertes, A.; Dominguez, J. M.; Plotkin, B.; Eldar-Finkelman, H.; Schmidt, B. *J. Med. Chem.* **2012**, *55*, 4407-4424.
(231) Beurel, E.; Grieco, S. F.; Jope, R. S. *Pharm. Ther.* **2015**, *148*, 114-131.
(232) Nguyen, T.; Fan, T.; George, S. R.; Perreault, M. L. *Front. Aging Neurosci.* **2018**, *9*.
(233) Jope, R. S.; Roh, M.-S. *Curr. Drug Targets* **2006**, *7*, 1421-1434.
(234) Muneer, A. *Clin. Psychopharmacol. Neurosci.* **2017**, *15*, 100-114.
(235) Hooper, C.; Killick, R.; Lovestone, S. *J. Neurochem.* **2008**, *104*, 1433-1439.
(236) Hu, S.; Begum, A. N.; Jones, M. R.; Oh, M. S.; Beech, W. K.; Beech, B. H.; Yang, F.; Chen, P.; Ubeda, O. J.; Kim, P. C.; Davies, P.; Ma, Q.; Cole, G. M.; Frautschy, S. A. *Neurobiol. Dis.* **2009**, *33*, 193-206.

(237) Zhu, X.; Rottkamp, C. A.; Hartzler, A.; Sun, Z.; Takeda, A.; Boux, H.; Shimohama, S.; Perry, G.; Smith, M. A. *J. Neurochem.* **2001**, *79*, 311-318.
(238) Alam, J.; Scheper, W. *Autophagy* **2016**, *12*, 2516-2520.
(239) FENG, Y. J.; LI, Y. Y. *J. Dig. Dis.* **2011**, *12*, 327-332.
(240) World Register of Marine Species http://www.marinespecies.org/ (Accessed February, 2020)
(241) Tadesse, M.; Strøm, M. B.; Svenson, J.; Jaspars, M.; Milne, B. F.; Tørfoss, V.; Andersen, J. H.; Hansen, E.; Stensvåg, K.; Haug, T. *Org. Lett.* **2010**, *12*, 4752-4755.
(242) Tadesse, M.; Svenson, J.; Jaspars, M.; Strøm, M. B.; Abdelrahman, M. H.; Andersen, J. H.; Hansen, E.; Kristiansen, P. E.; Stensvåg, K.; Haug, T. *Tetrahedron Lett.* **2011**, *52*, 1804-1806.
(243) Tadesse, M.; Svenson, J.; Sepčić, K.; Trembleau, L.; Engqvist, M.; Andersen, J. H.; Jaspars, M.; Stensvåg, K.; Haug, T. *J. Nat. Prod.* **2014**, *77*, 364-369.
(244) Venkateswarlu, Y.; Ravinder, K.; Yadav, J. S.; Sarathkumar, Y.; Ramakrishna, S.; Diwan, P. V.; Rao, J. V.; Ramesh, R.; Laatsch, H. *PCT Int. Appl.* **2005**, 0-0.
(245) Miyata, Y.; Diyabalanage, T.; Amsler, C. D.; McClintock, J. B.; Valeriote, F. A.; Baker, B. J. *J. Nat. Prod.* **2007**, *70*, 1859-1864.
(246) Carroll, A. R.; Bowden, B. F.; Coll, J. C. *Aust. J. Chem.* **1993**, *46*, 1079-1083.
(247) Diyabalanage, T.; Amsler, C. D.; McClintock, J. B.; Baker, B. J. *J. Am. Chem. Soc.* **2006**, *128*, 5630-5631.
(248) Noguez, J. H.; Diyabalanage, T. K. K.; Miyata, Y.; Xie, X.-S.; Valeriote, F. A.; Amsler, C. D.; McClintock, J. B.; Baker, B. J. *Bioorgan. Med. Chem.* **2011**, *19*, 6608-6614.
(249) Ortega, M. a. J.; Zubía, E.; Ocaña, J. M.; Naranjo, S.; Salvá, J. *Tetrahedron* **2000**, *56*, 3963-3967.
(250) Pearce, A. N.; Chia, E. W.; Berridge, M. V.; Maas, E. W.; Page, M. J.; Webb, V. L.; Harper, J. L.; Copp, B. R. *J. Nat. Prod.* **2007**, *70*, 111-113.
(251) Won, T. H.; Jeon, J.-e.; Kim, S.-H.; Lee, S.-H.; Rho, B. J.; Oh, D.-C.; Oh, K.-B.; Shin, J. *J. Nat. Prod.* **2012**, *75*, 2055-2061.
(252) Carroll, A. R.; Healy, P. C.; Quinn, R. J.; Tranter, C. J. *J. Org. Chem.* **1999**, *64*, 2680-2682.
(253) Hopmann, K. H.; Šebestík, J.; Novotná, J.; Stensen, W.; Urbanová, M.; Svenson, J.; Svendsen, J. S.; Bouř, P.; Ruud, K. *J. Org. Chem.* **2012**, *77*, 858-869.
(254) Baker, B.; Diyabalanage, T.; McClintock, J.; Amsler, C. (University of South Florida and University of Alabama at Birmingham Research Foundation) US20050187286A1, **2005**.
(255) Noguez, J. H., Chemical Investigation of the Antarctic Marine Invertebrates Synoicum adareanum and Artemisina plumosa, PhD Thesis, University of South Florida, **2010**.
(256) Florence, G. J.; Wlochal, J. *Chem. Eur. J.* **2012**, *18*, 14250-14254.
(257) Lisboa, M. P.; Dudley, G. B. *Chem. Eur. J.* **2013**, *19*, 16146-16168.
(258) Nicolaou, K. C.; Leung, G. Y. C.; Dethe, D. H.; Guduru, R.; Sun, Y.-P.; Lim, C. S.; Chen, D. Y. K. *J. Am. Chem. Soc.* **2008**, *130*, 10019-10023.
(259) Ravu, V. R.; Leung, G. Y. C.; Lim, C. S.; Ng, S. Y.; Sum, R. J.; Chen, D. Y.-K. *Eur. J. Org. Chem.* **2011**, *2011*, 463-468.
(260) Brewin, B. J. *Trans. R. Soc. N.Z.* **1950**, *78*, 354–362.
(261) Millar, R. H. *The Marine Fauna of New Zealand: Ascidiacea*; New Zealand Dept. of Scientific and Industrial Research: Wellington, New Zealand, 1982.
(262) Awesome Ascidians - a guide to the sea squirts of New Zealand https://niwa.co.nz/coasts-and-oceans/marine-identification-guides-and-fact-sheets/seasquirt-id-guide (Accessed January 2020,
(263) Miao, S.; Andersen, R. J. *J. Org. Chem.* **1991**, *56*, 6275-6280.

(264) Andersen, R. J.; Faulkner, D. J.; He, C. H.; Van Duyne, G. D.; Clardy, J. *J. Am. Chem. Soc.* **1985**, *107*, 5492-5495.
(265) Bellina, F.; Anselmi, C.; Martina, F.; Rossi, R. *Eur. J. Org. Chem.* **2003**, *2003*, 2290-2302.
(266) Sikorska, J.; Parker-Nance, S.; Davies-Coleman, M. T.; Vining, O. B.; Sikora, A. E.; McPhail, K. L. *J. Nat. Prod.* **2012**, *75*, 1824-1827.
(267) Wang, W.; Kim, H.; Patil, R. S.; Giri, A. G.; Won, D. H.; Hahn, D.; Sung, Y.; Lee, J.; Choi, H.; Nam, S.-J.; Kang, H. *Bioorg. Med. Chem. Lett.* **2017**, *27*, 574-577.
(268) Wang, W.; Kim, H.; Nam, S.-J.; Rho, B. J.; Kang, H. *J. Nat. Prod.* **2012**, *75*, 2049-2054.
(269) Shen, G. Q.; Baker, B. J. *Tetrahedron Lett.* **1994**, *35*, 4923-4926.
(270) Smitha, D.; Kumar, M. M. K.; Ramana, H.; Rao, D. V. *Nat. Prod. Res.* **2014**, *28*, 12-17.
(271) Zhu, T.; Chen, Z.; Liu, P.; Wang, Y.; Xin, Z.; Zhu, W. *J. Antibiot.* **2013**, *67*, 315.
(272) Barbosa, L. C. A.; Maltha, C. R. A.; Lage, M. R.; Barcelos, R. C.; Donà, A.; Carneiro, J. W. M.; Forlani, G. *J. Agric. Food Chem.* **2012**, *60*, 10555-10563.
(273) Pereira, U. A.; Barbosa, L. C. A.; Demuner, A. J.; Silva, A. A.; Bertazzini, M.; Forlani, G. *Chem. Biodivers.* **2015**, *12*, 987-1006.
(274) de Nys, R.; Wright, A. D.; König, G. M.; Sticher, O. *Tetrahedron* **1993**, *49*, 11213-11220.
(275) Yin, Q.; Liang, J.; Zhang, W.; Zhang, L.; Hu, Z.-L.; Zhang, Y.; Xu, Y. *Mar. Biotechnol.* **2019**, *21*, 88-98.
(276) Smith, C. J.; Hettich, R. L.; Jompa, J.; Tahir, A.; Buchanan, M. V.; Ireland, C. M. *J. Org. Chem.* **1998**, *63*, 4147-4150.
(277) Bae, J.; Cho, E.; Park, J. S.; Won, T. H.; Seo, S.-Y.; Oh, D.-C.; Oh, K.-B.; Shin, J. *J. Nat. Prod.* **2020**, *83*, 429-437.
(278) Jennings, L. K.; Robertson, L. P.; Rudolph, K. E.; Munn, A. L.; Carroll, A. R. *J. Nat. Prod.* **2019**, *82*, 2620-2626.
(279) Kiriyama, N.; Nitta, K.; Sakaguchi, Y.; Taguchi, Y.; Yamamoto, Y. *Chem. Pharm. Bull.* **1977**, *25*, 2593-2601.
(280) Nitta, K.; Fujita, N.; Yoshimura, T.; Arai, K.; Yamamoto, Y. *Chem. Pharm. Bull.* **1983**, *31*, 1528-1533.
(281) Miao, S., Novel Secondary Metabolites from Selected Marine Invertebrates, PhD Thesis, University of British Columbia, Canada, **1985**.
(282) Shenkar, N.; Swalla, B. J. *PLoS One* **2011**, *6*, e20657.
(283) Schmidt, E. W.; Donia, M. S.; McIntosh, J. A.; Fricke, W. F.; Ravel, J. *J. Nat. Prod.* **2012**, *75*, 295-304.
(284) Hirose, E.; Oka, A. T.; Hirose, M. *Zoological Science* **2009**, *26*, 362-368.
(285) Lewin, R. A. *Nature* **1976**, *261*, 697.
(286) Martins, J.; Vasconcelos, V. *Mar. Drugs* **2015**, *13*, 6910.
(287) Schmidt, E. W.; Nelson, J. T.; Rasko, D. A.; Sudek, S.; Eisen, J. A.; Haygood, M. G.; Ravel, J. *Proc. Natl. Acad. Sci.* **2005**, *102*, 7315-7320.
(288) Donia, M. S.; Ravel, J.; Schmidt, E. W. *Nat. Chem. Biol.* **2008**, *4*, 341-343.
(289) Menna, M. *Phytochem. Rev.* **2009**, *8*, 461-472.
(290) Rinehart, K. L.; Kishore, V.; Bible, K. C.; Sakai, R.; Sullins, D. W.; Li, K.-M. *J. Nat. Prod.* **1988**, *51*, 1-21.
(291) Rinehart, K. L.; Gloer, J. B.; Cook, J. C.; Mizsak, S. A.; Scahill, T. A. *J. Am. Chem. Soc.* **1981**, *103*, 1857-1859.
(292) Newman, D. J. *Front. Mar. Sci.* **2018**, *5*.
(293) Appleton, D. R.; Pearce, A. N.; Copp, B. R. *Tetrahedron* **2010**, *66*, 4977-4986.

(294) Kobayashi, J.; Cheng, J.-f.; Nakamura, H.; Ohizumi, Y.; Hirata, Y.; Sasaki, T.; Ohta, T.; Nozoe, S. *Tetrahedron Lett.* **1988**, *29*, 1177-1180.
(295) Searle, P. A.; Molinski, T. F. *J. Org. Chem.* **1993**, *58*, 7578-7580.
(296) Liberio, M. S.; Sooraj, D.; Williams, E. D.; Feng, Y.; Davis, R. A. *Tetrahedron Lett.* **2011**, *52*, 6729-6731.
(297) Carroll, A. R.; Nash, B. D.; Duffy, S.; Avery, V. M. *J. Nat. Prod.* **2012**, *75*, 1206-1209.
(298) Krishnaiah, P.; Reddy, V. L. N.; Venkataramana, G.; Ravinder, K.; Srinivasulu, M.; Raju, T. V.; Ravikumar, K.; Chandrasekar, D.; Ramakrishna, S.; Venkateswarlu, Y. *J. Nat. Prod.* **2004**, *67*, 1168-1171.
(299) Lefranc, F.; Koutsaviti, A.; Ioannou, E.; Kornienko, A.; Roussis, V.; Kiss, R.; Newman, D. *Nat. Prod. Rep.* **2019**, *36*, 810-841.
(300) Plisson, F.; Huang, X.-C.; Zhang, H.; Khalil, Z.; Capon, R. J. *Chem. Asian J.* **2012**, *7*, 1616-1623.
(301) Molinski, T. F.; Morinaka, B. I. *Tetrahedron* **2012**, *68*, 9307-9343.
(302) Reddy, M. V. R.; Faulkner, D. J.; Venkateswarlu, Y.; Rao, M. R. *Tetrahedron* **1997**, *53*, 3457-3466.
(303) Reddy, M. V. R.; Rao, M. R.; Rhodes, D.; Hansen, M. S. T.; Rubins, K.; Bushman, F. D.; Venkateswarlu, Y.; Faulkner, D. J. *J. Med. Chem.* **1999**, *42*, 1901-1907.
(304) Davis, R. A.; Carroll, A. R.; Pierens, G. K.; Quinn, R. J. *J. Nat. Prod.* **1999**, *62*, 419-424.
(305) Carroll, A. R.; Bowden, B. F.; Coll, J. C. *Aust. J. Chem.* **1993**, *46*, 489-501.
(306) Lindquist, N.; Fenical, W.; Van Duyne, G. D.; Clardy, J. *J. Org. Chem.* **1988**, *53*, 4570-4574.
(307) Ham, J.; Kang, H. *B. Korean Chem. Soc.* **2002**, 163-166.
(308) Ridley, C. P.; Reddy, M. V. R.; Rocha, G.; Bushman, F. D.; Faulkner, D. J. *Bioorgan. Med. Chem.* **2002**, *10*, 3285-3290.
(309) Carroll, A., Personal Correspondance. Griffith University; July 2018.
(310) Urban, S.; Capon, R. J. *Aust. J. Chem.* **1996**, *49*, 711-713.
(311) Fukuda, T.; Itoyama, R.; Minagawa, T.; Iwao, M. *Heterocycles* **2014**, *88*, 1121-1133.
(312) Yoshida, K.; Itoyama, R.; Yamahira, M.; Tanaka, J.; Loaëc, N.; Lozach, O.; Durieu, E.; Fukuda, T.; Ishibashi, F.; Meijer, L.; Iwao, M. *J. Med. Chem.* **2013**, *56*, 7289-7301.
(313) Pescitelli, G.; Bruhn, T. *Chirality* **2016**, *28*, 466-474.
(314) Carroll, A.; Copp, B. R.; Davis, A. R.; Keyzers, R. A.; Prinsep, M. R. **2020**.
(315) Zhou, H.; Li, L.; Wu, C.; Kurtán, T.; Mándi, A.; Liu, Y.; Gu, Q.; Zhu, T.; Guo, P.; Li, D. *J. Nat. Prod.* **2016**, *79*, 1783-1790.
(316) Li, X.-C.; Ferreira, D.; Ding, Y. *Curr. Org. Chem.* **2010**, *14*, 1678-1697.
(317) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098-3100.
(318) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785-789.
(319) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.
(320) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
(321) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158-6170.
(322) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396-1396.
(323) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865-3868.
(324) Bailly, C. *Curr. Med. Chem. Anticancer Agents* **2004**, *4*, 363-378.
(325) Fan, H.; Peng, J.; Hamann, M. T.; Hu, J.-F. *Chem. Rev.* **2008**, *108*, 264-287.
(326) Pla, D.; Albericio, F.; Alvarez, M. *Anti-Cancer Agents Med. Chem.* **2008**, *8*, 746-760.
(327) Bailly, C. In *Outstanding Marine Molecules*; Wiley-VCH Verlag GmbH & Co. KGaA: 2014, p 377-386.
(328) Bailly, C. *Mar. Drugs* **2015**, *13*, 1105.

(329) Rudi, A.; Goldberg, I.; Stein, Z.; Frolow, F.; Benayahu, Y.; Schleyer, M.; Kashman, Y. *J. Org. Chem.* **1994**, *59*, 999-1003.
(330) Heim, A.; Terpin, A.; Steglich, W. *Angew. Chem. Int. Ed.* **1997**, *36*, 155-156.
(331) Peschko, C.; Winklhofer, C.; Steglich, W. *Chem. Eur. J.* **2000**, *6*, 1147-1152.
(332) Kornprobst, J.-M.; Sallenave, C.; Barnathan, G. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **1998**, *119*, 1-51.
(333) Handy, S. T.; Zhang, Y.; Bregman, H. *J. Org. Chem.* **2004**, *69*, 2362-2366.
(334) Fan, G.; Li, Z.; Shen, S.; Zeng, Y.; Yang, Y.; Xu, M.; Bruhn, T.; Bruhn, H.; Morschhäuser, J.; Bringmann, G.; Lin, W. *Bioorgan. Med. Chem.* **2010**, *18*, 5466-5474.
(335) Kang, H.; Fenical, W. *J. Org. Chem.* **1997**, *62*, 3254-3262.
(336) Uzair, B.; Mahmood, Z.; Tabassum, S. *BioImpacts* **2011**, *1*, 203-211.
(337) Riccio, R.; Kinnel, R. B.; Bifulco, G.; Scheuer, P. J. *Tetrahedron Lett.* **1996**, *37*, 1979-1982.
(338) Mitchell, S. S.; Rhodes, D.; Bushman, F. D.; Faulkner, D. J. *Org. Lett.* **2000**, *2*, 1605-1607.
(339) González, N.; Rodríguez, J.; Jiménez, C. *J. Org. Chem.* **1999**, *64*, 5705-5707.
(340) Hajimahdi, Z.; Zarghi, A. *Iran J. Pharm. Res.* **2016**, *15*, 595-628.
(341) Summa, V.; Petrocchi, A.; Bonelli, F.; Crescenzi, B.; Donghi, M.; Ferrara, M.; Fiore, F.; Gardelli, C.; Gonzalez Paz, O.; Hazuda, D. J.; Jones, P.; Kinzel, O.; Laufer, R.; Monteagudo, E.; Muraglia, E.; Nizi, E.; Orvieto, F.; Pace, P.; Pescatore, G.; Scarpelli, R.; Stillmock, K.; Witmer, M. V.; Rowley, M. *J. Med. Chem.* **2008**, *51*, 5843-5855.
(342) Mbisa, J. L.; Martin, S. A.; Cane, P. A. *Infect. Drug Resist.* **2011**, *4*, 65-76.
(343) Yamaguchi, T.; Fukuda, T.; Ishibashi, F.; Iwao, M. *Tetrahedron Lett.* **2006**, *47*, 3755-3757.
(344) Kamiyama, H.; Kubo, Y.; Sato, H.; Yamamoto, N.; Fukuda, T.; Ishibashi, F.; Iwao, M. *Bioorgan. Med. Chem.* **2011**, *19*, 7541-7550.
(345) Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. *Organometallics* **2010**, *29*, 2176-2179.
(346) Pierens, G. K.; Carroll, A. R.; Davis, R. A.; Palframan, M. E.; Quinn, R. J. *J. Nat. Prod.* **2008**, *71*, 810-813.
(347) Wiegand, I.; Hilpert, K.; Hancock, R. E. W. *Nat. Protoc.* **2008**, *3*, 163-175.
(348) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 09 Rev. D.01*; Wallingford, CT, 2016.
(349) Grimme, S. *J. Comp. Chem.* **2006**, *27*, 1787-1799.
(350) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
(351) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comp. Chem.* **2011**, *32*, 1456-1465.
(352) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297-3305.
(353) Cammi, R.; Mennucci, B.; Tomasi, J. *J. Phys. Chem. A* **2000**, *104*, 5631-5637.
(354) Cossi, M.; Barone, V. *J. Chem. Phys.* **2000**, *112*, 2427-2435.
(355) Cossi, M.; Barone, V. *J. Chem. Phys.* **2001**, *115*, 4708-4717.

(356) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Chem. Phys.* **2001**, *114*, 5691-5701.
(357) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43-54.
(358) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999-3094.
(359) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comp. Chem.* **2003**, *24*, 669-681.
(360) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378-6396.
(361) CYLview, 1.0b http://www.cylview.org (Accessed April 2019,
(362) Stamatakis, A. *Bioinformatics* **2014**, *30*, 1312-1313.
(363) Felsenstein, J. *Evolution* **1985**, *39*, 783-791.
(364) Mujumdar, P.; Bua, S.; Supuran, C. T.; Peat, T. S.; Poulsen, S.-A. *Bioorg. Med. Chem. Lett.* **2018**, *28*, 3009-3013.
(365) Willoughby, P. H.; Jansma, M. J.; Hoye, T. R. *Nat. Protoc.* **2014**, *9*, 643-660.
(366) Bruhn, T.; Schaumlöffel, A.; Hemberger, Y.; Bringmann, G. *Chirality* **2013**, *25*, 243-249.
(367) Weeratunga, S.; Hu, N.-J.; Simon, A.; Hofmann, A. *BMC Bioinformatics* **2012**, *13*, 201.
(368) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. C.01*; Wallingford, CT, 2016.