UNIVERSITA' DEGLI STUDI DI NAPOLI "FEDERICO II"

DIPARTIMENTO DI FARMACIA

PhD Thesis in Pharmaceutical Sciences

XXXIV cycle

by

Silvia SCARPATO

Date of defense:

Innovative strategies in the identification of new natural lead compounds from marine organisms

Tutor
Prof. MANGONI Alfonso

PhD Coordinator
Prof. D'AURIA Maria Valeria

CONTENTS

Abstract in English

Marine environment represents one of the most interesting sources of inspiration for the development of new drugs. Marine organisms, for example porifera, tunicates, aquatic plants, can synthesize and/or accumulate an important plurality of secondary metabolites with peculiar and unusual molecular structures, implicated in defence mechanisms, in intra- and inter-species communication and in processes of adaptation to extreme conditions. In the spectrum of pharmacological activities boasted by natural products isolated from marine organisms, it is fundamental to mention the antibiotic, antifungal, antiviral and antineoplastic activities. In recent years, about 20 compounds from marine origin have been in phase I, II or III clinical trials, whereas some marine natural product-derived compounds are commercially available. Significant advances in the field of structural elucidation techniques, occurred in recent years, allow for the complete elucidation of sub milligram samples. In addition, the development of innovative techniques combining data obtained from sophisticated analytical techniques and efficient data analysis software, has allowed to stem the difficulty in identifying new natural compounds from complex extracts, facilitating the process of dereplication of the huge number of metabolites. Among these, molecular networking is an innovative computational technique, which is a valuable complement to traditional dereplication techniques. Molecular networking processes the massive amounts of data obtained from LC-MS/MS analysis of complex extracts, allowing the automated identification of the structural similarity between

metabolites based on similarity between the fragmentation spectra of the compounds in the extract.

The research work described in this PhD thesis fits into this scenario. A significant part of the research activity was focused on molecular networking, demonstrating its effectiveness both in the discovery of new natural products and in environmental analysis. In particular, the workflow that integrates the analysis of LC-MS/MS data of extracts of the widely studied sponge *Stylissa caribica* with molecular networking, in its feature-based molecular networking variant, allowed the rapid isolation of a new cyclic heptapeptide, stylissamide L. Moreover, the study with these methods of the extract of the sponge *Clathria faviformis* led to the isolation of a new phospholipid, favilipid A. The validity of the methodology was further verified with the analysis of extracts from the cyanobacterium *Trichodesmium thiebautii*, whose metabolome had shown traits in common with the metabolic profile of the marine sponge *Smenospongia aurea.* Thanks to Feature-Based Molecular Networking, it was possible to focus the analysis on halogenated compounds in the extract, leading to the isolation of 4 new polyketide compounds, included in the class of trichophycins (isotrichophycin C and trichophycins G-I). The cytotoxicity of the newly isolated compounds, along with the previously identified trichophycins was evaluated against the mouse neuroblastoma cell line N2A, as part of a broader structure-activity relationship (SAR) study, aimed at determining the features that modulate bioactivity.

In the context of environmental monitoring of cyanobacteria, the validity of a pipeline that integrates the investigation of the territory thanks to specific

satellites, allowing the rapid identification of potentially toxic cyanobacterial blooms, and the subsequent chemical analysis of collected environmental samples, based on molecular networking, was verified in the analysis of an unusual cyanobacteria bloom that took place in Lake Avernus at the turn of the lockdown period from COVID-19 in Italy. The presence in the extracts of cyanotoxins, particularly microcystins, was verified by genome amplification analysis and by in-depth analysis of molecular networking clusters containing nodes recognized as known toxins by in-silico tools, provided by the freely accessible website GNPS. In addition, the cytotoxic activity of the chloroform extracts was confirmed.

Another important part of the research activity was focused on the detailed structural elucidation of new natural products isolated from the extracts of the seagrass *Zostera marina*. The structure of the new cyclic diarylheptanoids, zosteraphenol A and zosteraphenol B was completely elucidated. They both displayed broad and unresolved signals in 1D and 2D NMR spectra even at low temperature, which were shown to be related to a conformational equilibrium between rotamers with opposite axial chirality with the support of quantum mechanical calculations. Further analysis of extracts from *Z. marina* led to the discovery of two unique diarylheptanoid dimers, zosterabisphenone A and B, each composed of two different cyclic diarylheptanoids. They showed the same conformational equilibrium as zosteraphenols A and B, and their structure elucidation required extensive use of DFT calculations. A possible biosynthesis of zosterabisphenone A was proposed. In cytotoxicity assays performed for these compounds, zosterabisphenone B exhibited time-dependent and concentration-dependent

cytotoxic activity towards HCT116 cells. Finally, the new phenolic acid, 7″,8″-didehydrosalviolanic acid B was isolated and characterized as part of a study investigating seasonal variations of phenolic compounds in *Z. marina*. Its absolute configuration was assigned by comparison of the ECD spectrum of methyl (3,4-dihydroxyphenyl)lactate produced by degradation of the isolated compound, and methyl (3,4-dihydroxyphenyl) lactate produced by degradation of rosmarinic acid, whose absolute configuration is known from previous studies.

Abstract in Italian

L'ambiente marino rappresenta una delle fonti di ispirazione più interessanti per lo sviluppo di nuovi farmaci. Gli organismi marini, come ad esempio poriferi, tunicati, piante acquatiche, possono sintetizzare e/o accumulare una importante pluralità di metaboliti secondari con peculiari ed inusuali strutture molecolari, implicati nei meccanismi di difesa, nella comunicazione intra- e inter-specie e nei meccanismi di adattamento a condizioni estreme. Nello spettro di attività farmacologiche vantate dai prodotti naturali isolati da poriferi è fondamentale citare l'attività antibiotica, antimicotica, antivirale e antineoplastica. Negli ultimi anni, circa 20 composti naturali di origine marina sono stati sottoposti a studi clinici di fase I, II o III, mentre alcuni composti derivanti da prodotti naturali marini sono in commercio. I significativi progressi nel campo delle tecniche di elucidazione strutturale avvenuti negli ultimi anni permettono l'elucidazione completa di una quantità di microgrammi di campione. Inoltre, lo sviluppo di tecniche innovative che combinano i dati ottenuti da sofisticate tecniche analitiche ed efficienti software di analisi e dei dati, ha permesso di arginare la difficoltà nell'identificazione di nuovi composti naturali da estratti complessi, facilitando il processo di dereplicazione dell'ingente numero di metaboliti presenti. Tra questi, il molecular networking è una tecnica computazionale innovativa, che costituisce un valido complemento alle tradizionali tecniche di dereplicazione. Il molecular networking elabora la mole di dati ottenuta dall'analisi LC-MS/MS di estratti complessi, permettendo l'identificazione automatica della similitudine strutturale tra i

metaboliti, che è basata sulla similitudine strutturale tra gli spettri di frammentazione dei composti presenti nell'estratto. In questo scenario si inserisce il lavoro di ricerca descritto in questa tesi di dottorato. Una parte significativa dell'attività di ricerca è stata incentrata sul molecular networking, dimostrando la sua efficacia sia nel campo della ricerca di nuove sostanze naturali sia nell'analisi ambientali. In particolare, il workflow che integra l'analisi dei dati LC-MS/MS degli estratti della spugna *Stylissa caribica*, ampiamente studiata da vari gruppi di ricerca, con il molecular networking, nella sua variante feature-based molecular networking, ha permesso il rapido isolamento di un nuovo eptapeptide ciclico la stylissamide L. Inoltre, lo studio con questi metodi dell'estratto della spugna marina *Clathria faviformis* ha portato all'isolamento di un nuovo fosfolipide, il favilipide A. La validità della metodica è stata ulteriormente verificata con l'analisi degli estratti del cianobatterio *Trichodesmium, thiebautii*, il cui metaboloma ha mostrato tratti in comune con il profilo metabolico della spugna marina *Smenospongia aurea.* Grazie al Feature-Based Molecular Networking è stato possibile focalizzare l'attenzione sui composti alogenati presenti nell'estratto, portando all' isolamento di quattro nuovi polichetidi, inclusi nella classe delle tricoficine (isotricoficina C, tricoficine G-I). La citotossicità dei nuovi composti isolati insieme alle tricoficine precedentemente identificate, è stata valutata nei confronti della linea cellulare N2A di neuroblastoma del topo, nell'ambito di uno studio più ampio di relazione struttura-attività (SAR), volto alla determinazione delle caratteristiche che modulano la bioattività. Nell'ambito del monitoraggio ambientale di cianobatteri, la validità di una

pipeline che integra l'investigazione del territorio grazie a specifici satelliti, permettendo l'identificazione rapida di fioriture di cianobatteri potenzialmente tossiche, e la successiva analisi chimica di campioni ambientali raccolti, basata sul molecular networking, è stata verificata nell'analisi di un insolito fioritura di cianobatteri che ha avuto luogo nel Lago d'Averno a cavallo del periodo di lockdown da COVID-19 in Italia. La presenza di cianotossine, in particolare di microcistine, negli estratti, è stata verificata sia tramite analisi di amplificazione del genoma sia attraverso l'analisi approfondita dei cluster del molecular networking, contenenti nodi riconosciuti da strumenti in-silico come tossine note, sul sito gratuito GNPS. Inoltre, è stata confermata anche l'attività citotossica degli estratti di cloroformio.

Un'altra parte importante dell'attività di ricerca è stata incentrata sull'elucidazione strutturale dettagliata di nuovi prodotti naturali isolati dagli estratti della pianta acquatica *Zostera marina.* La struttura dei nuovi diarileptanoidi ciclici, zosterafenolo A e zosterafenolo B, è stata completamente chiarita. Entrambi hanno mostrato segnali larghi e non risolti negli spettri 1D e 2D NMR anche a bassa temperatura, che hanno dimostrato di essere legati ad un equilibrio conformazionale tra rotameri con chiralità assiale opposta con il supporto di calcoli quantomeccanici. L'ulteriore analisi degli estratti della pianta acquatica *Zostera marina* ha permesso di individuare due dimeri diarileptanoidi unici, zosterabisfenone A e B, ciascuno composto da due diarileptanoidi ciclici. Essi hanno mostrato, lo stesso equilibrio conformazionale degli zosterafenoli A e B e la loro determinazione strutturale ha richiesto un ampio utilizzo dei calcoli DFT. La

possibile biosintesi dello zosterabisfenone A è stata ipotizzata. Nei saggi di tossicità effettuati su questi composti, lo zosterabisfenone B, ha mostrato attività citotossica tempo-dipendente e concentrazione-dipendente nei confronti delle cellule HCT116. Infine, il nuovo acido fenolico,l' acido 7″,8″-dideidrosalvianolico B, è stato isolato e caratterizzato nell'ambito di uno studio sulle variazioni stagionali di composti fenolici della *Zostera marina*. La sua configurazione assoluta è stata assegnata dal confronto dello spettro ECD del (3,4-diidrossifenil)lattato di metile prodotto della degradazione del composto, e del (3,4-diidrossifenil)lattato di metile prodotto dalla degradazione dell'acido rosmarinico, la cui configurazione assoluta è nota da studi precedenti.

INTRODUCTION

Man’s interest in natural substances dates back to the dawn of time, when primitive humans experimented with the effects on the human body of ingesting herbs, roots, berries, and wild fruits. The healing properties of plants have been thoroughly investigated by all ancient civilizations. As an example, a profound knowledge of vegetable drugs was achieved by Egyptian medicine, as witnessed by the papyrus of Ebers (dated around 1500 B.C.), containing a collection of thousands of natural medicinal preparations and the procedures for obtaining them. However, it was not until the second half of the eighteenth century, when the remarkable development of chemistry led to the design of new isolation and purification techniques, that scientists were able to identify the single compounds responsible for the observed effects, later called "active substances".

Moreover, the deeper knowledge of the physiology of animal and vegetal organisms achieved in this century enabled the understanding of their metabolism, with the distinction between primary and secondary metabolites.

Primary metabolites, such as carbohydrates, fatty acids, amino acids, and nucleic acids, obtained with almost universal chemical reactions, represent the key components in maintaining normal physiological processes in the organisms. On the other hand, secondary metabolism consists of the metabolic engineering of non-essential metabolites, which are present only in specific taxa of living organisms. Secondary metabolites are proven to be

involved in the ecological interactions between species, anti-predator strategies and adaptation to marginal conditions. As products of secondary metabolism reactions, these compounds show a peculiar and complex structure.

The plant kingdom has been shown to be a rich source of secondary metabolites, whose pharmacological features have long been investigated to treat a wide range of illnesses. Thus, a good variety of pharmaceutically available drugs is of natural plant origin or is obtained by modifications of natural compounds. Examples include the anticancer drugs vinblastine and vincristine, alkaloids isolated from *Catharanthus roseus*, the anti-inflammatory drug aspirin, derived by the salicylic acid found in the bark of *Salix Alba* and digoxin and digitoxin, indicated in cardiac decompensation, from the leaves of *Digitalis Purpurea*.

Marine organisms, such as sponges, tunicates, molluscs, seaweeds, marine plants, and microorganisms, represent an outstanding and countless storehouse of novel bioactive compounds. Marine natural products (MNs) generally have completely diverse structural and chemical properties from molecules found in terrestrial environment, due to the distinct biochemical reactions that occur in marine and terrestrial organisms which are strongly influenced by different growth environments.[1]

Furthermore, the evident genome heterogeneity observed in a large population of organisms of the same species, leading to the structural diversity among the compounds produced, may likely result in a wide assortment of biological activities. The need for marine organisms to survive in extreme marine ecosystem, characterized by high ion

concentrations, confined living spaces, nutrient scarcity and pH and temperature variability, has led to evolving adaptations to these conditions, with the production of a wide spectrum of secondary metabolites.[2] In particular, the most attractive organisms in the marine environment are Porifera and Tunicates, which, belonging to sessile marine communities, possess the ability to produce bioactive compounds as their only weapon to defend themselves against predators.

However, research into marine organisms faces major problems. Firstly, it is necessary to collect samples from the same site of origin, to avoid quantitative and qualitative variability in the chemical composition of extracts even if the samples belong to the same species.

In addition, pharmaceutical industries often show a lack of interest towards marine natural products, due to the small amounts that can be isolated, insufficient for clinical purpose. While the more logical solution to address this issue seems to be the design of total syntheses of these compounds, their complex structure, rich in chiral centres, complicates both the synthetic workflow and the purification of the intermediates obtained in reaction steps. At the same time, the massive extraction of natural organisms cannot be a viable route, leading to serious consequences for marine biodiversity. Two ways can be envisaged to reconcile the massive production of marine natural products and the need to safeguard marine ecosystem: aquaculture, such as the controlled growth of aquatic organisms in monitored areas, and the production of metabolites in laboratory conditions using cell cultures. In addition, further advances in modern biotechnologies promise to provide the ability to control biochemical mechanisms through gene cloning, resulting

in large-scale production of bioactive compounds by fermentation of heterologous hosts.

At the present times, 30 marine natural products-based compounds are included in the pharmaceutical clinical pipeline: eight of them clinically approved by EMEA, FDA, Japanese Ministry of Health and Australia's Therapeutic Goods Administration.[3]

Among the first approved marine compound-derived drug the ziconotide (Prialt®), a synthetic peptide, derived by the ω-conotoxin MVIIA produced by the tropical cone snail Conus magus, was approved as intrathecal infusion formulation for the treatment of severe and chronic pain. Very interesting are the commercially available anticancer drugs, Ectinascidin-743 or Trabectedin (Yondelis®), a marine alkaloid identified in the extracts of the Caribbean tunicate *Ecteinascidia turbinata,* which finds application in the treatment of refractory soft-tissue sarcoma, whose structure present three fused tetrahydroisoquinoline rings.[4] Another recently approved drug to treat breast cancer is Eribulin mesylate, commercially known as Halaven®, an analogue of halicondrin B, isolated by the extracts of the Japanese sponge *Halicondria okadai*.[5]

Interesting is the approved drug brentuximab vedotin, commercially known as ADCETRIS®, used to treat relapsing or refractory Hodgkin's lymphoma and anaplastic large cell lymphoma (ALCL). This medicament, obtained by recombinant DNA biotechnology, consists in an anti-CD-30 monoclonal antibody linked to monomethylauristatin E (MMAE), a semisynthetic analogue of the depsipeptide dolastatin-10 found in the extracts of the

marine hare *Dolabella auricularia*, but recently identified as a cyanobacterial product.

Indeed, an in-depth study on marine natural compounds reveals that microorganisms living in symbiosis with sponges and tunicates, are often the producers of natural compounds isolated from macro-organisms. Based on the assumption that a proportion of natural products isolated from marine macro-organisms, is actually synthetized by bacteria, the laboratory growth of interesting bacteria could potentially provide a sustainable source of bioactive molecules, for clinical trials or even large-scale drug production.

Although the cultivation of most bacteria is challenging, due to the difficulty in reproducing the actual growth conditions of the microorganism, the identification of the gene cluster responsible for metabolite biosynthesis through metagenomics could allow the transfer of the gene cluster to an engineered bacterial host, producing in a large scale the compound of interest through heterologous expression.

Recent literature reveals an increased wave of attention towards marine natural products in the last years, indeed a huge number of bioactive compounds have been isolated, exhibiting interesting pharmacological activities for example antitumoral, antimicrobial and antiviral activities,[6] which can act as lead compounds for the development of new drugs.

The search for new compounds of marine origin was initially focused on species that were easy to find and available in large quantities, while in the last decades the spotlight was turned on organisms which are endemic to restricted areas or whose habitat consists of exotic environments like polar zones or hydrothermal nozzles or are hard to collect.

Given the considerable progresses made in structural elucidation techniques, which make it possible to fully determine the structure and the stereochemistry of isolated compounds even in very small quantities, the current constraint in natural product research is the fast identification of novel marine products from complex extracts, containing thousands of different molecules, including already known natural substances, primary metabolites, and contaminants. This process, referred to as dereplication, is best achieved by high-resolution liquid chromatography coupled with tandem mass spectrometry. However, despite being the analytical technique of choice for analysis of complex extracts, LC-HRMS/MS provides huge amounts of data matching the complexity of the extract but making manual data analysis challenging. Thus, a lot of bioinformatic tools for natural product dereplication have been devised. Among them, molecular networking is the modern MS-based technique that attracts the greatest interest from researchers. Molecular networking has been proved to be an effective computational tool to process the results of LC-HRMS/MS experiments, allowing an automated identification of structural similarity between metabolites, which is inferred from the relatedness of their MS/MS spectra.

In this scenario, the research carried out during this PhD project was focused on the investigation of extracts of various marine organisms, with the main aim of the identification and the full structural elucidation of new natural compounds as possible lead compounds, complementing traditional dereplication strategies with innovative approaches. In particular, the strength of molecular networking was fully explored both for dereplication

of new chemical entities and for environmental purposes. The study with these methods of the extract of the well-known sponge *Stylissa caribica* led to the straightforward isolation of a new cyclic heptapeptide, stylissamide L (**1**), demonstrating the strength of molecular networking for accurate dereplication of complex mixtures. Furthermore, the workflow that integrates the analysis of LC-MS/MS data of extracts of the sponge *Clathria faviformis* with molecular networking, in its feature-based molecular networking variant, allowed the rapid isolation of a new phospholipid, favilipid A (**2**). The similarity between the metabolomes of the sponge *Smenospongia aurea* and of the cyanobacterium *Trichodesmium* pushed our interest towards the molecules produced by the free-living bacterium *Trichodesmium thiebautii* (this project was pursued in collaboration with the research group of Professor Matthew Bertin at the University of Rhode Island, USA). The visual overview of the chemical repertoire detected by LC-MS/MS, provided by molecular networking, resulted in the characterization of a new group of chlorinated polyketides: isotrichophycin C (**12**) and trichophycins G-I (**13-15**). Moreover, the value of molecular networking was shown in the examination of a cyanobacterial bloom, occurred in Lake Avernus in Campania, providing a comprehensive picture of the toxic peptides accumulated during the harmful bloom. At the same time, the research activity was also dedicated to the in-depth investigation of the extracts of the common eelgrass *Zostera marina*, carried out in collaboration with the research group of Professor Christian Zidorn at the University of Kiel. More precisely, the chemical study of *Zostera marina* led to the isolation of two new tetracyclic diarylheptanoids, zosteraphenols

A (**3**) and B (**4**). Their structure was elucidated through an extensive NMR study, with the support of quantum-mechanical calculations, showing that they experience a slow conformational equilibrium involving the tetracyclic structures between two diastereomeric rotamers with opposite axial chirality. Two unique dimeric diarylheptanoids, zosterabisphenone A (**5**) and B (**6**), each composed of two coupled cyclic diarylheptanoids were also isolated from *Z. marina*. Their cytotoxic activity was evaluated, and zosterabisphenone B (**6**) exhibited cytotoxic activity in a concentration- and time-dependent manner on HCT116 cells up to 48 hours of exposure.

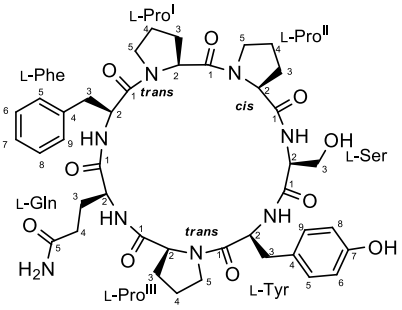
Besides, structural determination and the clarification of the absolute configuration of the new phenolic acid 7″,8″-didehydrosalvianolic acid B (**7**), found in the extracts of *Zostera marina*, was performed. In particular, the absolute configuration of the new compound was thoroughly assigned by the comparison of the ECD spectrum of methyl (3,4-dihydroxyphenyl)lactate (**8**) obtained from the methanolysis of rosmarinic acid (**9**), whose absolute configuration was known from previous studies, and methyl (3,4-dihydroxyphenyl) lactate (**8**) produced by the degradation of compound **7**.

The results obtained can be divided into three different sections according to the organism under investigation.

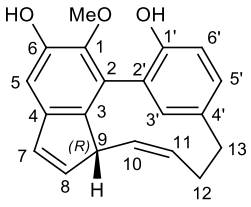
- Exploration of the extracts from marine sponges
- Exploration of the extracts the seagrass *Zostera marina*
- Exploration of the extracts from cyanobacteria

The results obtained during the Ph.D. program have been reported in six published articles.

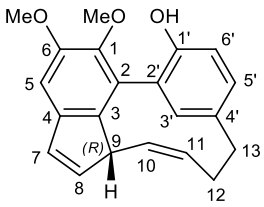
1. Grauso L, Li Y, Scarpato S, Shulha O, Rárová L, Strnad M, Teta R, Mangoni A, Zidorn C. Structure and Conformation of Zosteraphenols, Tetracyclic Diarylheptanoids from the Seagrass *Zostera marina*: An NMR and DFT Study. *Org Lett.* **2020** ;*22(1)*:78-82.
2. Scarpato, S.; Teta, R.; Della Sala, G.; Pawlik, J.R.; Costantino, V.; Mangoni, A. New Tricks with an Old Sponge: Feature-Based Molecular Networking Led to Fast Identification of New Stylissamide L from *Stylissa caribica*. *Mar. Drugs* **2020**, *18*, 443.
3. McManus KM, Kirk RD, Via CW, Lotti JS, Roduit AF, Teta R, Scarpato S, Mangoni A, Bertin MJ. Isolation of Isotrichophycin C and Trichophycins G-I from a Collection of *Trichodesmium thiebautii*. *J Nat Prod.* **2020**, *83(9)*, 2664-2671.
4. Li Y, Grauso L, Scarpato S, Cacciola NA, Borrelli F, Zidorn C, Mangoni A. Stable Catechol Keto Tautomers in Cytotoxic Heterodimeric Cyclic Diarylheptanoids from the Seagrass *Zostera marina*. *Org Lett.* **2021,** *23(18),*7134-7138.
5. Teta, R.; Sala, G.D.; Esposito, G.; Stornaiuolo, M.; Scarpato, S.; Casazza, M.; Anastasio, A.; Lega, M.; Costantino, V. Monitoring Cyanobacterial Blooms during the COVID-19 Pandemic in Campania, Italy: The Case of Lake Avernus. *Toxins* **2021**, *13*, 471.
6. Li, Y.; Rárová, L.; Scarpato, S.; Sezai Çiçeka, S.; Jordheimd, T.; Štenclováb, T.; Strnade, M.; Mangoni, A.; Zidorn, C. Seasonal variation of phenolic compounds in *Zostera marina* (Zosteraceae) from the Baltic Sea. *Phytochemistry,**2022**, 196,113099.*



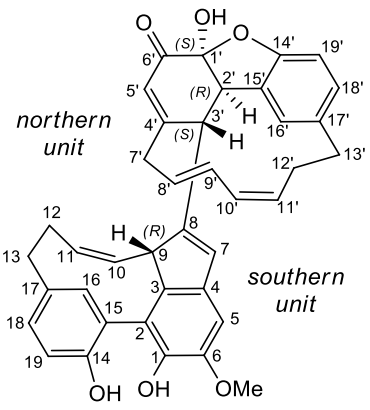
stylissamide L (**1**)



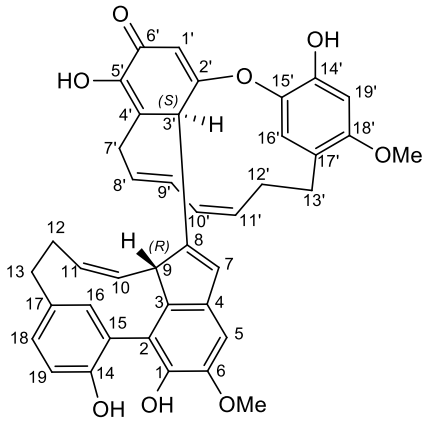
zosteraphenol A (**3**)



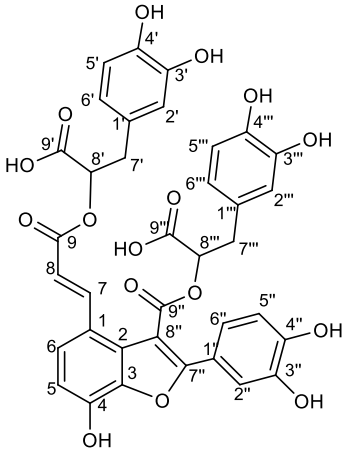
zosteraphenol B (**4**)



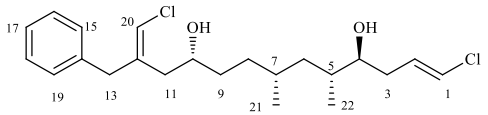
zosterabisphenone A (**5**)



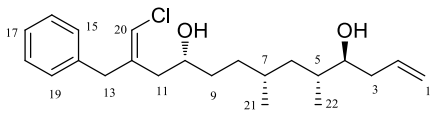
zosterabisphenone B (**6**)



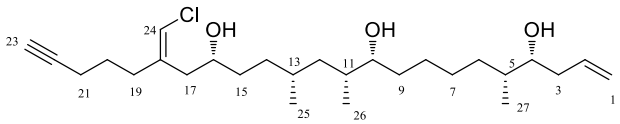
7",8"-didehydrosalvianolic acid B (**7**)



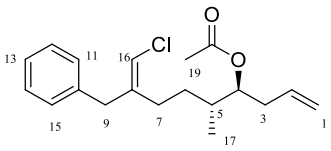
isotrichophycin C (**12**)



trichophicin G (**13**)



trichophicin H (**14**)



trichophicin I (**15**)

Figure 1. Structure of stylissamide L (**1**), favilipid A (**2**), zosteraphenol A (**3**) and B (**4**), zosterabisphenone A (**5**), zosterabisphenone B (**6**), 7″,8″-didehydrosalvianolic acid B (**7**), isotrichophycin C (**12**) and trichophycin G-I (**13**-**15**).

References

[1] Shang, J.; Hu, B.; Wang, J.; Zhu, F.; Kang, Y.; Li, D.; Sun, H.; Kong, D.X.; Hou, T. Cheminformatic Insight into the Differences between Terrestrial and Marine Originated Natural Products. *J. Chem. Inf. Model.* **2018**, *58*, 1182–1193, doi:10.1021/acs.jcim.8b00125.

[2] Conte, M.; Fontana, E.; Nebbioso, A.; Altucci, L. Marine-derived secondary metabolites as promising epigenetic bio-compounds for anticancer therapy. *Mar. Drugs* **2021**, *19*, doi:10.3390/md19010015.

[3] Pereira, F. Have marine natural product drug discovery efforts been productive and how can we improve their efficiency? *Expert Opin. Drug Discov.* **2019**, *14*, 717–722, doi:10.1080/17460441.2019.1604675.

[4] Gordon, E.M.; Sankhala, K.K.; Chawla, N.; Chawla, S.P. Trabectedin for Soft Tissue Sarcoma: Current Status and Future Perspectives. *Adv. Ther.* **2016**, *33*, 1055–1071, doi:10.1007/s12325-016-0344-3.

[5] Molinski, T.F.; Dalisay, D.S.; Lievens, S.L.; Saludes, J.P. Drug development from marine natural products. *Nat. Rev. Drug Discov.* **2009**, *8*, 69–85, doi:10.1038/nrd2487.

[6] Carroll, A.R.; Copp, B.R.; Davis, R.A.; Keyzers, R.A.; Prinsep, M.R. Marine natural products. *Nat. Prod. Rep.* **2021**, *38*, 362–413, doi:10.1039/d0np00089b.

Chapter 1

Structural elucidation methods

Prior to the introduction of spectroscopy, the complete structure of new organic molecules could only be determined using chemical techniques such as degradation or elaboration of functional groups of the compound under study. These techniques were destructive and only suitable for compounds available in tens or hundreds of milligrams. While these methods succeeded in the elucidation of the structure of some surprisingly complex, yet abundant natural products, there are much more natural products that are extremely scarce in the producing organisms and can be isolated only in microgram amounts, and as such cannot be studied using chemical methods. Nowadays, thanks to the introduction and the modern advances of spectroscopic techniques, the structure and the stereochemistry of compounds can be determined from samples as small as a few micrograms. Structural determination of the natural products presented in this thesis is mainly based on mass spectrometry (MS) and nuclear magnetic resonance (NMR). Chiral compounds were characterized by their specific rotation at the sodium D line (589 nm) and CD spectroscopy was used to obtain information about the absolute stereochemistry of the isolated natural products. Yet, the absolute configuration of amino acids was clarified using chemical degradation and functionalization, in a procedure called Marfey's method.

1.1 Mass spectrometry

Structure elucidation of a new natural product begins with the determination of its molecular formula through high-resolution mass spectrometry analysis.

Mass spectrometry is a powerful analytical technique whose output data are needed to completely elucidate the structure of unknown compounds. Specifically, mass spectrometers work by converting analyte molecules into ions and subsequent measurement of their mass to charge ratio (*m/z* ratio).

Three key components of mass spectrometers are the ion source, which produces ions, the mass analyser, which separates ions based on their mass to charge ratio, and the detector, which collects ions and converts them into an electric output. Several technologies are currently in use for ionization of compounds in sources, as well as different types of analysers are available.

The sine qua non for detection processes in the mass spectrometer is the efficient formation of gaseous ions. Depending on the different type of mass spectrometer, samples may be already in the form of ions in solution, or they may be ionized in the source. The natural products presented in this work were analysed by ESI (Electrospray Ionization) mass spectrometry (Figure 1.1) through an Orbitrap system.

Electrospray or ESI is an atmospheric pressure ion source (API), which is considered a “soft” ionization technique, because of the relatively low energy imparted to the analyte, in contrast to other MS ion sources that cause a huge fragmentation. In the ESI source liquid samples, dissolved in a polar and volatile solvent such as H_2O, MeOH, and CH_3CN, are pumped through a metal capillary maintained under vacuum, and, in presence of an

electrostatic field, they are nebulized and charged. The droplets are rapidly evaporated by the application of heated drying gas which increase charge concentration in droplets. The decrease in droplet size causes electrostatic repulsion of the charges. When the Coulombic forces reach the Rayleigh limit, the ions are ejected into the gas phase and forced through a capillary and then into the mass analyser.

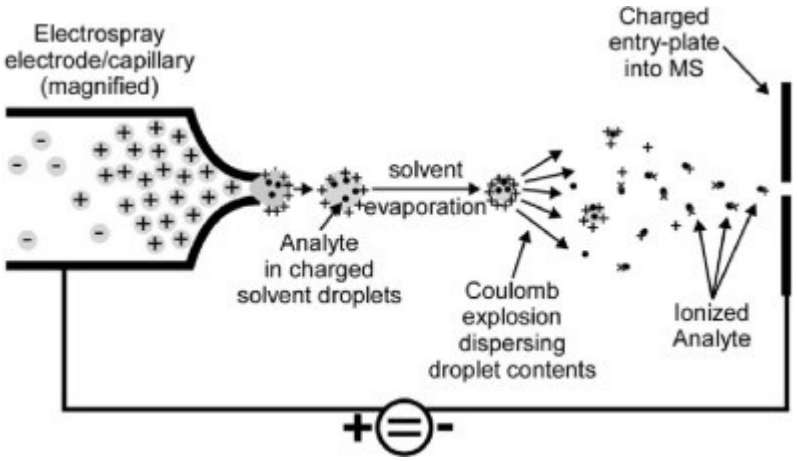


Figure 1.1. Representation of an ESI source [1]

The ESI ionization technique is often combined with high-performance liquid chromatography (HPLC).

In the ESI source molecular ions are not fragmented during ionization. However, after being selected by the analyser, ions can be fragmented by collision with an inert gas like Argon or Nitrogen (Collision Activated Dissociation) in a collision chamber. Product ions are then separated and selected according to their *m/z* ratio in a second analyser. This is known as tandem mass spectrometry or MS/MS.

The Orbitrap analyser, developed by the Russian physic Alexander Alexeyevich Makarov, consists of a central electrode surrounded by a pair

of bell-shaped outer electrodes. Ions which enter the Orbitrap are captured by the electric field and oscillate around the central electrode with stable trajectories. The device obtains the frequencies of these axial oscillations through Fast Fourier Transformation (FFT) and thus the *m/z* relationships of the ions. The Orbitrap analyser provides high-resolution (exact) mass measurement, allowing determination of the molecular formula.

The LTQ-Orbitrap instrument used in this work combines into a single instrument a linear ion trap, used as first analyser and collision chamber in MS/MS experiments, and an Orbitrap analyser, providing excellent mass accuracy, and high MS and MS/MS sensitivity.[2]

1.2 Nuclear Magnetic Resonance

The Nuclear Magnetic Resonance Spectroscopy (NMR) is an essential analytical technique to obtain detailed information about the molecular structure of compounds. This technique works measuring the interaction of an electromagnetic radiation of a molecule subjected to an intense magnetic field. The electromagnetic radiation causes nuclear spin transitions in nuclei possessing a nuclear magnetic moment of spin.

Alongside the standard ^{1}H and ^{13}C NMR spectra, homo and hetero-nuclear 2D NMR experiments, have been extensively recorded. Due to the possibility of finding nuclei resonating in similar regions of the spectra, preventing efficient chemical shift assignment, two-dimensional NMR experiments, provide much more information, as signal overlap is much less likely in two dimensions than in one.

The COSY (Correlation SpectroscopY) experiment, represents one of the most useful 2D experiments, allowing to identify the correlation of protons that are 2 or 3 bonds apart (scalarly coupled). Various versions of COSY experiments exist, depending on the particular pulse sequences used, the use of gradients, the use of multiple-quantum filters, and so on.

The 2D NMR experiment TOCSY (Total Correlation SpectroscopY) is very helpful in the analysis of molecules like peptides and oligosaccharides, which are composed of many separate spin systems. Correlation peaks in TOCSY spectrum are observed not only between coupled protons, but also between all the protons that are in the same spin system.

Among the heteronuclear 2D NMR experiments, it is important to mention the HSQC experiment (Heteronuclear Single Quantum Correlation), which provides information about the carbons directly linked to hydrogens, through the correlation of ^{13}C nuclei with protons coupled with them via a single bond $^{1}J_{CH}$.[3] In addition, the HMBC (Heteronuclear Multi Bond Correlation) experiment provides a spectrum in which correlation peaks show long range ^{1}H-^{13}C heteronuclear couplings, i.e. between ^{13}C nuclei and protons separated through two ($^{2}J_{CH}$) or three bonds ($^{3}J_{CH}$).[4] The information obtained from HMBC spectrum is fundamental for linking the subunits of the molecule, determined by the interpretation of the other two-dimensional spectra. In particular, the HMBC experiment is crucial for the assignment of non-protonated carbons through the correlation peaks with nearby protons, which cannot be achieved through an HSQC experiment. However, in many HMBC spectra not all the correlation peaks that one might expect from the structure of the compound are shown. This is why

$^{2,3}J_{CH}$ can be dissimilar from each other, requiring the experiment to be optimised for each type of coupling.

When several ^{13}C nuclei happen to resonate very close to each other, preventing their assignment by standard HSQC and HMBC experiments, band-selective HSQC and HMBC experiments may be performed. 2D heteronuclear band-selective (bs) NMR experiments have been introduced to improve spectral resolution in the indirect dimension (typically ^{13}C) while keeping the number of experiments at a reasonable value. These experiments use band-selective pulses, i. e. pulses that are effective only for nuclei resonating in a limited range of chemical shifts and allow the spectral width of the heteronuclear chemical shift to be narrowed without causing spectral folding.

1.3 Methods for the determination of relative/absolute configurations

Assessing the relative and absolute stereochemistry of a natural product is a crucial aspect of its structural characterization. Indeed, the stereochemistry of a molecule, linked to its three-dimensional structure, is pivotal in biological studies, in particular drug-receptor interaction, as well as in chemical studies, such as those relating to the total synthesis of complex molecules. Most natural molecules have one or more chiral centres. NMR may provide information on relative configuration of chiral centres through the values of chemical shifts (δ), coupling constants (*J*) and NOE effects.

The chemical shift of a proton is strongly influenced by the chemical surroundings, and typically protons of diastereomers resonate at different

frequencies. However, there are no general rules to correlate chemical shifts with configurations of the surrounding chiral centres.

Coupling constants (J_{HH}) can provide a great deal of information on the relative stereochemistry of a compound. The value of both the homonuclear coupling constants, $^3J_{HH}$, and the heteronuclear coupling constants, $^3J_{CH}$, follows Karplus's law. According to Karplus' law, these values depend on the dihedral angle θ between the coupled atoms.[5] The values of $^3J_{HH}$ are very small, about 0-1.5 Hz, when θ is close to 90°, while the values rise considerably when θ has values between 0° or 180°. In addition, $^3J_{HH}$ can be used to determine the configuration of a double bond, because $^3J_{cis}$ ~ 6-12 Hz, $^3J_{trans}$ ~ 14-20 Hz.

Further structural data about the relative configuration of a molecule can be provided by dipolar coupling, which is experienced by nuclei that are spatially close to each other but not necessarily separated by few bonds (and therefore scalarly coupled). Dipolar coupling is not observed directly in solution NMR, but causes NOE (Nuclear Overhauser Enchancements) effects, which consist of the variation of the signal intensity of a proton following the irradiation of another proton that is close to the observed proton.[6] An increase in signal intensity (positive NOE) is usually observed for small organic molecules, and a decrease for macromolecules (negative NOE). Therefore, NOE can detect the spatial relationship between various parts of a molecule.

Nowadays NOE effect between protons is normally measured using the NOESY experiment (Nuclear Overhauser Effect SpectroscopY), which can provide information about the dipolar couplings between all the pairs of

protons in the molecule in a single two-dimensional experiment. In this homonuclear 2D NMR experiment, the correlation peak indicates that there is a NOE effect between the two protons, and that they are therefore close in space.

The ROESY (Rotating-frame Overhauser SpectroscopY) experiment is a homonuclear correlation experiment based on the detection of ROE (Rotating-frame Overhauser Effect). The ROE effect, similar to the NOE effect, is related to the dipolar coupling between the nuclei, which depends on the geometric distance between the nuclei. Unlike the NOE effect, which is positive for small molecules and negative for macromolecules, the ROE effect is always positive. Therefore, the ROESY experiment is particularly suitable for medium-sized molecules, which would show a NOE effect close to zero.

Being a non-chiral technique, NMR cannot distinguish between enantiomers (i.e. two enantiomers will behave exactly in the same way in any possible NMR experiment) and therefore cannot be used to determine the absolute configuration of a molecule, unless they are turned into diastereomers by chemical derivatization with a chiral reagent (e.g. Mosher's method).

1.3.1 Electronic circular dichroism (ECD)

Electronic circular dichroism spectroscopy (ECD), known until a few years ago as CD (circular dichroism) but now called ECD to distinguish it from vibrational circular dichroism (VCD), is a spectroscopic technique that provides information on the three-dimensional structure of an organic

molecule, and therefore on its configuration (including absolute configuration) and on its conformations.

The basic principle of this technique is the ability of certain substances to absorb to a different extent UV/Vis radiation circularly polarised in one direction and UV/Vis radiation polarised in the opposite direction. The electric field vector associated with the two types of circularly polarised radiation, right-hand and left-hand polarised radiation, describes a helical path, which in the case of right-hand polarised radiation will be a right-hand helix and in the case of left-hand polarised radiation a left-hand helix. The two components are absorbed differently by the chiral medium. At a given wavelength, the circular dichroism of a substance is the difference between the absorbance of left and right circularly polarised light: $\Delta A = A_L - A_R$.

The resulting ECD spectrum is a graph of the circular dichroism ΔA (or the corresponding molar coefficient $\Delta\varepsilon$) as a function of wavelength. Obviously, to be subjected to ECD spectroscopy, the molecule under investigation must absorb UV/Vis light, so it must have at least one chromophore. If the molecule does not have a chromophore, this can in many cases be introduced by subjecting the compound to a derivatisation reaction. Based on the principle that circular dichroism uses asymmetric electromagnetic radiation, enantiomers can also be easily distinguished, as the CD of pure enantiomers are the mirror image of each other.

However, identifying the relationship between the absolute configuration of an unknown compound and its ECD spectrum is not straightforward because the ECD is influenced in many different ways by the electronic and geometrical structure of the molecule. There exist several empirical rules to

do so for specific classes of compounds (e.g. the octant rules for ketones), and in more general terms this can be achieved by comparing the experimental ECD spectrum with the ECD spectrum obtained by quantum mechanical calculations. Currently, ECD is the technique of choice for determining the absolute configuration of a natural product due to its high sensitivity. In fact, a few micrograms of sample are usually sufficient to obtain an analysable ECD spectrum.[7]

1.3.2 Computational methods

Although the analysis of NMR data makes it possible to elucidate the relative stereochemistry of organic molecules based on evaluations of coupling constants in conjunction with analyses of NOESY or ROESY spectra, which is useful for determining the distances between the different protons of molecules, the interpretation of NMR data is not always simple, clear, and unambiguous.

Computational chemistry is a branch of chemistry that, based on the laws of quantum and statistical mechanics, can predict the properties of molecules and compounds through the use of appropriate computer programs. Computational methods are a very useful method for carrying out a qualitative and quantitative simulation of the chemical and physical phenomena of the molecules under investigation. They may be used to predict the geometry of a molecule, including information on bond lengths, bond angles, and torsional angles, the energy of the fundamental electronic states, the frequencies associated with the vibrational motion of the nuclei, and many other properties difficult to obtain experimentally. Thus,

computational chemistry is both a supplementary tool to experimental data and a tool for predicting the relationship between the structure and the properties of a chemical system.

Two main groups of computational methods exist: classical methods, which rely on a model of the molecule based on electrostatics and classical mechanics to determine molecular features, and quantum-mechanical methods, which can in turn be divided into two subgroups: ab initio methods, which study a system using equations and calculations based on chemical first principles, and semi-empirical methods, which include parameters and data obtained experimentally in the calculations.

In both methods, the conformation of a molecule is determined through a process called "minimisation", in which the change in the total energy of the system is evaluated as a result of the displacement of the atoms from their position, in order to identify the geometry of the molecule with the lowest energy. A fundamental problem with minimization is that the geometry (and the corresponding energy) provided by this method refers to a local minimum, i.e. a relative minimum that depends on the starting structure using as input for the calculation, and is not necessarily the absolute minimum, corresponding to the lowest energy geometry. Therefore, a *conformational search* must be performed prior to minimizations to identify a comprehensive set of possible starting conformation to be subjected to minimization. Conformational search is often the most difficult part in a computational study.

One method to perform a conformational search is molecular dynamics. Molecular dynamics simulates the motion of the molecule under the thermal

motion, which provides the energy necessary to cross the energy barrier between conformations. Molecular dynamics is very demanding computationally, and therefore is usually performed using classic methods. Even so, current computers do not allow simulations lasting at most microseconds, while many conformational changes can occur more slowly than this.

To overcome this difficulty, conformational searches can be done using *simulated annealing*. In a simulated annealing calculation, a molecular dynamics simulation is performed molecule starting at a high temperature and gradually lowering the temperature to 0 K. During the high-temperature simulation, the molecule moves quickly between different conformations, whereas during the low temperature simulations the molecule relaxes moving smoothly towards the nearest local minimum, still being able to cross small energy barriers along the way.

The result of conformational search and minimization are a comprehensive or representative set of low-energy conformers, which are significantly populated at room temperature (for rigid molecules, only one significant conformer is often found). These can be used to calculate quantum-mechanically the properties of interest, which can be NMR chemical shift or coupling constants, ECD spectra, or any other measurable quantity. The predicted properties are compared with the experimental values, confirming or disproving a structural hypothesis, or allowing the choice between two possible structures.

1.3.3 Marfey's method

The most widely used techniques to clarify the L or D steric series of the amino acids present in a peptide is Marfey's method.[8] Marfey's method is based on the principle that, while the HPLC retention times of enantiomeric amino acids are obviously the same, retention times are different if enantiomers are converted in diastereomers by reaction with a chiral reagent.

This general principle is implemented through an elegant, effective, and sensitive degradation protocol. The method involves an initial total hydrolysis of the peptide under investigation and subsequent derivatization of the amino acids obtained with the chiral reagent 1-fluoro-2,4-dinitrophenyl-5-L-alaninamide (L-FDAA). The reaction is an aromatic nucleophilic displacement of the fluorine leaving group by the amino group of the amino acid. The steric series of the amino acids present can then be determined by comparing the LC-MS retention times of the derivatized amino acids present in the peptide with the retention times of the corresponding L or D standards derivatized with the same reagent.

The Marfey's method can therefore be schematized in several sequential steps:

- Hydrolysis of the peptide with 6N HCl
- Derivatization of the hydrolysate with L-FDAA
- Preparation of derivatives with L-FDAA of L and D amino acid standards (thus giving L-FDAA-L-amino acid and L-FDAA-D-amino acid). Alternatively, in case the D-series steric standard is not available or is too expensive, the L-amino acid to be analysed can be

derivatized with L and D-FDAA (thus giving L-FDAA-L-amino acid and D-FDAA-L-amino acid; the latter is the enantiomer of L-FDAA-D-amino acid and shows the same retention time).

- HPLC analysis of the derivatized samples and comparison of the respective retention times.

The derivatization reaction is easy and effective, and the overall sensitivity of the method is limited by the sensitivity and selectivity of the HPLC detector. If an LC-MS instrument is used for the analysis, Marfey's method can be used with a sample as small as 1 μg.

In cases where the peptide contains non-proteinogenic amino acids, for which no amino acid standards are available, it is still possible to use a slightly Marfey's method called the advanced Marfey's method, which is based on the order of elution of amino acids derivatized with L-FDAA.[9] In particular, the results of tests carried out on a series of proteinogenic and non-proteinogenic amino acids have shown that, more or less consistently, the L amino acid derivatized with L-FDAA, analysed with a C18 column, has a shorter retention time than its corresponding D enantiomer, also derivatized with L-FDAA.

The different hydrophobicity is caused by the *cis* or *trans* arrangement of the two most hydrophobic substituents on the alpha carbons of L and D amino acids and of L-FDAA, as confirmed by UV and NMR measurements. Specifically, a cis arrangement was observed for the L-FDAA-D-amino acid stereoisomer, which therefore can interact more strongly with the C18 phase and shows a longer retention time. The trans rearrangement of the L-FDAA-

L-amino acid stereoisomer causes lower hydrophobicity and therefore a shorter retention time.

Schematically, to define the configuration of a non-proteinogenic amino acid, it is necessary to

- Derivatise an aliquot of amino acid with L-FDAA
- Derivatise an amino acid portion with D-FDAA
- LC/MS analysis of the derivatives and comparison of their Rt.

References

[1] Grebe SK, Singh RJ. LC-MS/MS in the Clinical Laboratory - Where to From Here? Clin Biochem Rev. 2011 Feb;32(1):5-31. PMID: 21451775; PMCID: PMC3052391.

[2] Hu, Q.; Noll, R.J.; Li, H.; Makarov, A.; Hardman, M.; Cooks, R.G. The Orbitrap: A new mass spectrometer. *J. Mass Spectrom.* **2005**, *40*, 430–443, doi:10.1002/jms.856.

[3] Palmer, A.G.; Cavanagh, J.; Byrd, R.A.; Rance, M. Sensitivity improvement in three-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson.* **1992**, *96*, 416–424, doi:10.1016/0022-2364(92)90097-Q.

[4] Bax, A.; Summers, M.F. 1H and 13C Assignments from Sensitivity-Enhanced Detection of Heteronuclear Multiple-Bond Connectivity by 2D Multiple Quantum NMR. *J. Am. Chem. Soc.* **1986**, *108*, 2093–2094, doi:10.1021/ja00268a061.

[5] Karplus, M. Contact electron-spin coupling of nuclear magnetic moments. *J. Chem. Phys.* **1959**, *30*, 11–15, doi:10.1063/1.1729860.

[6] Sanders, J.K.M.; Mersh, J.D. Nuclear magnetic double resonance; the use of difference spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **1982**, *15*, 353–400, doi:10.1016/0079-6565(82)80011-3.

[7] Grauso, L.; Teta, R.; Esposito, G.; Menna, M.; Mangoni, A. Computational prediction of chiroptical properties in structure elucidation of natural products. *Nat. Prod. Rep.* **2019**, *36*, 1005–1030, doi:10.1039/c9np00018f.

[8] Marfey, P. Determination of o-amino acids . II Use of a bifunctional. *Carlsb. Res. Commun.* **1984**, *49*, 591–596.

[9] Fujii, K.; Ikai, Y.; Mayumi, T.; Oka, H.; Suzuki, M.; Harada, K.I. A nonempirical method using LC/MS for determination of the absolute configuration of constituent amino acids in a peptide: Elucidation of limitations of Marfey's method and of its separation mechanism. *Anal. Chem.* **1997**, *69*, 3346–3352, doi:10.1021/ac9701795.

Chapter 2

Molecular networking

It is well established in the scientific community that marine organisms are an important source of new chemically diverse compounds with potential applications for drug discovery, showing a wide range of pharmacological activities, especially anti-cancer, and antibiotic properties.[1] In contrast to the early years of marine organism research, where only abundant or easy-to-find species were widely analysed, many marine chemists have continued to explore the marine environment, broadening their focus to organisms that live only in limited areas such as polar or tropical seas or are found in hard-to-reach areas. However, identifying new bioactive compounds of interest from complex extracts can be a real challenge, given the complexity of the composition of extracts that can contain thousands of substances including primary metabolites, known secondary metabolites, and contaminants.

The process, defined dereplication, involving the identification of known natural substances from a complex extract is best achieved by high-resolution liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS).[2] Unfortunately, traditional dereplication techniques, including manual analysis of by LC-MS/MS data, are time-consuming and laborious.

Recent significant advances in 21st century 'omics' technologies, e.g. genomics, metabolomics and information technology, are impressively accelerating the pace of discovery and analysis.[3] In particular, efficient

workflows combining sophisticated analytical techniques and powerful data mining software have been developed. One example is molecular networking, a modern computational approach to organising data provided by LC-MS/MS analysis, which is a strong complement to traditional dereplication techniques, and can be performed by any interested researcher using the freely accessible website GNPS, Global Natural Products Social Molecular Networking.[4] Molecular networking is a bio-informatics strategy that facilitates and speeds up the analysis of the overwhelming amount of data obtained from LC-MS/MS analysis of one or more extracts by providing a visual overview of the chemical repertoire detected.

This technique is based on the observation that the architecture of the molecule, the functional groups present, and its stability influence the reactivity of the compound in the dissociation occurring in the collision chamber of an MS/MS instrument, so that similar fractionation patterns can be used as a proxy for chemical similarity.[5] Based on a computational algorithm, the degree of similarity between every MS/MS spectrum is calculated, and metabolites are correlated according to the similarity of MS/MS fragmentation patterns, which is closely related to the structural similarity of the molecules themselves. The MS/MS spectra of the metabolites to be compared are aligned to identify fragment peaks in common between the spectra, the so-called peak matches, defined as fragment peaks that either have the same mass in the two spectra, or are produced by the same neutral loss from the respective molecular ions. A cosine similarity score between every possible couple of consensus MS/MS spectra is then computed, based on the number of matches found and their

relative intensity. The similarity score can be in a range between 0, if the metabolites have no fragmentation peaks in common, and 1, when all peaks in the spectrum coincide (this generally means that the spectra of the same compounds have been compared). In the molecular network, metabolites are represented as nodes, and connected to other similar metabolites by lines (edges). Within the molecular network, a node represents a consensus MS/MS spectrum, given by the mathematical combination of several similar MS/MS spectra, with the same precursor ion and assumed to be from the same compound. The entire molecular network will provide the simultaneous visual representation of compound families and analogues, grouped in clusters of molecules with similar structures. Molecular network visualization can be best approached with the software Cytoscape, in which several additional visual features can be added to aid interpretation of the data, such as the size of the nodes, which can be proportional to the intensity of the precursor ion, or the thickness of the connecting line, in relation to the cosine score.[6]

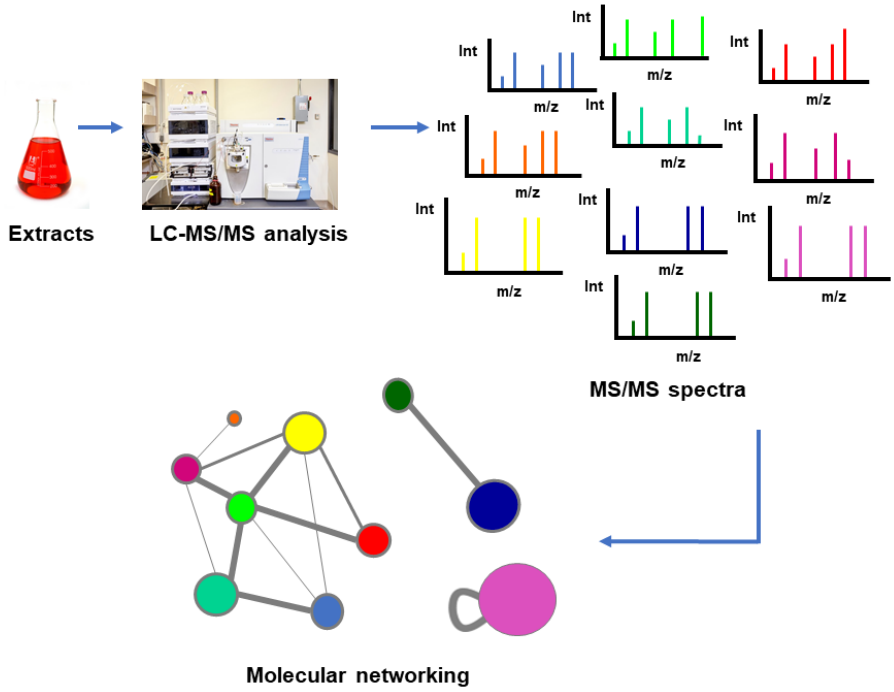


Figure 2.1. Schematic workflow of molecular networking

The considerable advantage of using this technique is its ability to compare the MS/MS spectra of the compounds present within the extracts with the fragmentation patterns of known natural products found in online libraries of MS/MS spectra, greatly facilitating the recognition of new compounds in a very early fractionation step, and avoiding re-isolating any known substances present. Moreover, close analogues of compounds present in the library may be easily identified even if they are not in the library themselves, and even if they are unknown compounds.

Important databases include GNPS or Global Natural Products Social Molecular Networking library, a virtual repository of over 200,000 MS/MS spectra of natural products. All LC-MS/MS data uploaded into GNPS are collected in a virtual platform, MassIVE, to facilitate and encourage the sharing of mass spectrometry data.[7] In addition, nodes can be associated

with additional information provided by the user, called metadata, such as product origin, hydrophobicity, and abundance.[4] Of particular interest is the possibility of including data of the bioactivity of the compounds as metadata, with the possibility of assigning each node, i.e. each metabolite, a score that predicts its potential bioactivity, based on the comparison with the degree of activity reported by other known metabolites that are structurally similar to it.[8] GNPS also provides additional tools, facilitating the recognition of peptide and non-peptide compounds and the chemical classification of compounds grouped in clusters. Indeed, dereplicator and dereplicator+ in silico-tools have recently been developed to annotate known natural products in samples analyzed by the comparison of experimental MS/MS data and the in-silico fragmentation tree of metabolites.[9] Furthermore, MolNetEnhancer workflow, on GNPS platform, provides for the fast automated chemical classification of the compounds in examined samples.[10]

Despite the ability to analyse and compare a large dataset, the main drawback of the MS-Cluster tool commonly used on the Global Natural Product Social Molecular Networking (GNPS) platform is the inability to distinguish between well-resolved isomers eluted at different retention times, as the chromatographic dimension is not considered. Furthermore, in the classical molecular network workflow, isotopic peaks or adduct ions may not be recognised as arising from the same compound, generating different nodes. An approach called Feature Based Molecular Networking,[11] which consists of pre-processing the raw LC-MS/MS data using MS processing software like MZmine2,[12] has been shown to generate

significantly better networks when used with LC-MS/MS data obtained from raw extracts. The Feature Based Molecular Networking pipeline exploits the chromatographic dimension, allowing the distinction of isomeric compounds based on their retention time. In addition, MZmine2 allows filtering of the data obtained, detection of isotopic peaks and quantification of compounds.

In conclusion, molecular Networking appears to be a very useful and versatile tool, showing an excellent compatibility with different mass ionization sources.[5] It is successfully used in many fields of the research, for example it has been recently introduced in the discovery of new lead compounds, with the possibility to simultaneously identify known natural products and new structural analogues from complex extracts, but it also has important applications in clinical diagnostics, in monitoring xenobiotics and their metabolites in body fluids and in the emerging field of precision medicine.[13]

References

[1] Pereira, F. Have marine natural product drug discovery efforts been productive and how can we improve their efficiency? *Expert Opin. Drug Discov.* **2019**, *14*, 717–722,

[2] Ito, T.; Masubuchi, M. Dereplication of microbial extracts and related analytical technologies. *J. Antibiot. (Tokyo).* **2014**, *67*, 353–360, doi:10.1038/ja.2014.12.

[3] Wolfender, J.L.; Litaudon, M.; Touboul, D.; Queiroz, E.F. Innovative omics-based approaches for prioritisation and targeted isolation of natural products-new strategies for drug discovery. *Nat. Prod. Rep.* **2019**, *36*, 855–868, doi:10.1039/c9np00004f.

[4] 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.; et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, 828–837, doi:10.1038/nbt.3597.

[5] Yang, J.Y.; Sanchez, L.M.; Rath, C.M.; Liu, X.; Boudreau, P.D.; Bruns, N.; Glukhov, E.; Wodtke, A.; De Felicio, R.; Fenner, A.; et al. Molecular networking as a dereplication strategy. *J. Nat. Prod.* **2013**, *76*, 1686–1699, doi:10.1021/np400413s.

[6] Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks., doi:10.1101/gr.1239303.

[7] Aron, A.T.; Gentry, E.C.; McPhail, K.L.; Nothias, L.F.; Nothias-Esposito, M.; Bouslimani, A.; Petras, D.; Gauglitz, J.M.; Sikora, N.; Vargas, F.; et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat. Protoc.* **2020**, *15*, 1954–1991, doi:10.1038/s41596-020-0317-5.

[8] Nothias, L.F.; Nothias-Esposito, M.; Da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; et al. Bioactivity-Based Molecular Networking for the Discovery of Drug Leads in Natural Product Bioassay-Guided Fractionation. *J. Nat. Prod.* **2018**, *81*, 758–767, doi:10.1021/acs.jnatprod.7b00737.

[9] Mohimani, H.; Gurevich, A.; Shlemov, A.; Mikheenko, A.; Korobeynikov, A.; Cao, L.; Shcherbin, E.; Nothias, L.F.; Dorrestein, P.C.; Pevzner, P.A. Dereplication of microbial metabolites through database search of mass spectra. *Nat. Commun.* **2018**, *9*, 1–12, doi:10.1038/s41467-018-06082-8.

[10] Ernst, M.; Kang, K. Bin; Caraballo-Rodríguez, A.M.; Nothias, L.F.; Wandy, J.; Chen, C.; Wang, M.; Rogers, S.; Medema, M.H.; Dorrestein, P.C.; et al. Molnetenhancer: Enhanced molecular networks by integrating metabolome mining and annotation tools. *Metabolites* **2019**, *9*, doi:10.3390/metabo9070144.

[11] Nothias, L.F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; et al. Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* **2020**, *17*, 905–908, doi:10.1038/s41592-020-0933-6.

[12] Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010**, *11*, doi:10.1186/1471-2105-11-395.

[13] Quinn, R.A.; Nothias, L.F.; Vining, O.; Meehan, M.; Esquenazi, E.; Dorrestein, P.C. Molecular Networking As a Drug Discovery, Drug Metabolism, and Precision Medicine Strategy. *Trends Pharmacol. Sci.* **2017**, *38*, 143–154, doi:10.1016/j.tips.2016.10.011.

PART 1

Exploration of the extracts from marine sponges

Chapter 3

Porifera

Porifera is the phylum of the so-called sponges, which are the least evolved multicellular invertebrates.[1] Almost all organisms in the phylum are marine sponges. Unlike most organisms belonging to the kingdom Animalia, sponges do not have true differentiated tissues and organs. Furthermore, although they exhibit radial symmetry at first, in the adult stage any form of symmetry is absent as the animal adapts to the surface on which it is fixed, taking on a different shape depending on the substrate.

The body of a sponge can be associated with a sac completely covered with pores, a characteristic from which the term poriferous is derived. The pores covering the surface of the sponge are either inhaling pores with a large opening (oscula) or exhaling pores.

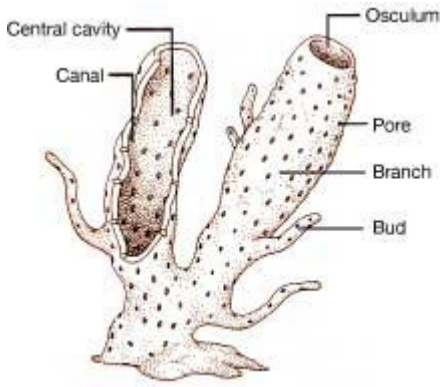


Figure 3.1 Sponge body structure

(Image retrieved from http://angelenamangieri.weebly.com/phylum-porifera.html)

They also have an inner chamber that serves as the gastric cavity, covered by a double wall, formed externally by the dermal lamina and internally by

the gastric lamina. The gastric leaflet is made up of flagellated cells, the choanocytes, through whose movements water enters the inhalation pores, passes through the gastric cavity, and exits into the oscula. (Figure 3.1) The passage of water due to the movement of the choanocytes provides oxygen to the cells, allowing the exchange of respiratory gases, and at the same time the exchange of food particles. Indeed, nutrients are retained and digested by the choanocytes themselves through the emission of pseudopods and distributed to the various tissues by amoeboid cells (amoebocytes). The dermal leaflet consists of flat cells called pinacocytes, and contractile cells, porocytes. In addition, the dermal and gastric lamina are separated by protein-rich gelatinous mesenchyme called mesohyl. Within this mesohyl there are small calcareous and siliceous bodies (spicules), which are produced by specialised cells called scleroblasts together with fibres of a scleroprotein (spongin) produced by specialised cells called spongioblasts.

Porifera can be classified into ascon, sycon and leucon according to their morphological structure. (Figure 3.2)

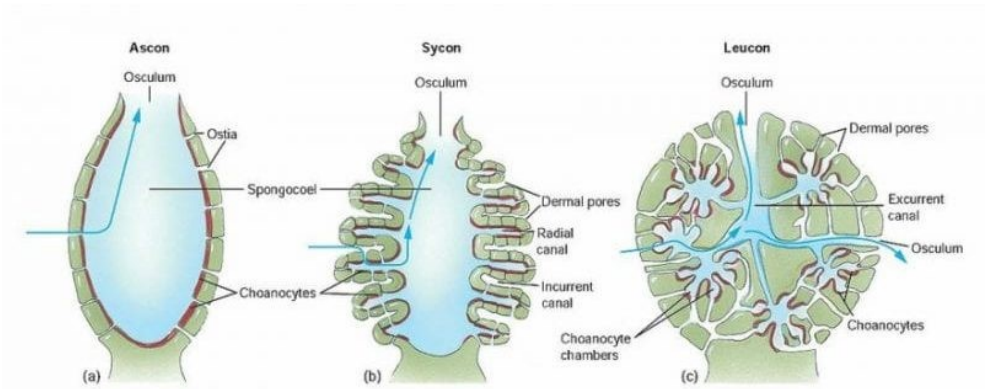


Figure 3.2 The organization of the sponge: a) ascon b) siconoid c) leuconoid.
(Image retrieved from https://biologando206.wordpress.com/2019/07/19/poriferos)

The first type called ascon is characteristic of calcareous sponges and larval stages, while the fundamental characteristic of the second type of

organisation, called sycon, is that the sponges form many extroflexions from the initial gastric cavity, where the choanocytes are located.[2] Finally, in the leucon sponges, the mesenchymal or mesohyl layer is very large, rich in small chambers called flagellated chambers, covered by choanocytes; the flagellated chambers are connected to the gastric cavity through efferent channels and to the inhalant pores through afferent channels.

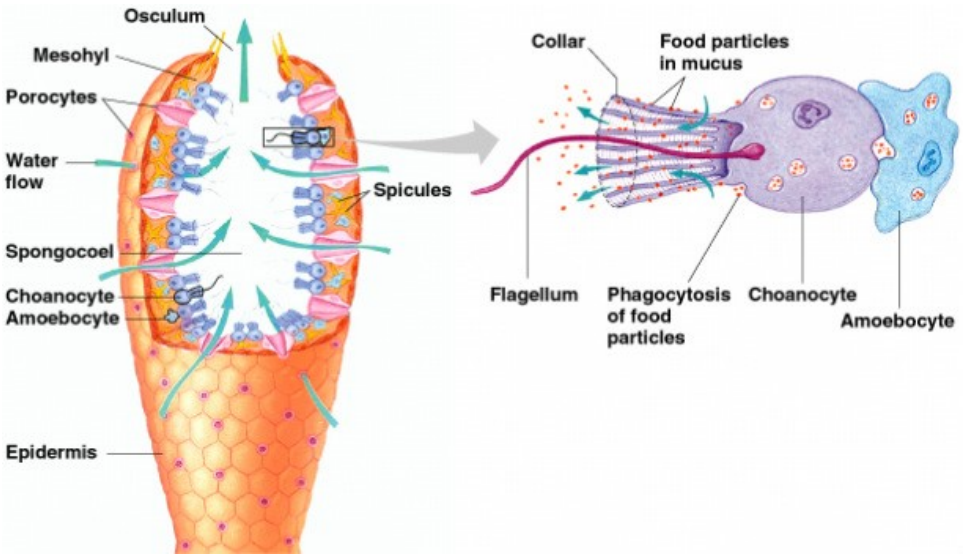


Figure 3.3 The structure of a sponge in detail.
(Image retrieved from https://dsip.weebly.com/porifera-sponges.html)

Generally, Porifera are hermaphroditic, in fact in reproduction each adult organism can act both as female and male.[3] The mesohyl contains the sex cells, sperm and ova. The embryonic cells, from which both sex cells and amoebocytes develop, are called archaeocytes. In most species' fertilization is internal. Released spermatozoa can reach another sponge with the water current and when a spermatozoon is captured by the choanocytes of another sponge the fertilization of an egg by the spermatozoa occurs inside the sponge. After the larva is released, it uses cilia (hairs) to swim, or it may

settle on the sea floor, becoming sessile and growing into an adult. Under certain unfavorable environmental conditions, sponges are also able to reproduce asexually, with the formation of external gems that can then give rise to new sponges when favourable environmental conditions are re-established. In the case of asexual reproduction, the genome of the new sponge will be identical to that of the mother sponge.

Porifera are divided into four classes: Calcarea (or Calcispongiae), Demospongiae, Hexactinellidae and Homoscleromorpha.[4]

3.1 Calcispongiae

Calcispongiae, also called calcareous sponges, are light-coloured organisms. They generally live at shallow depths and are usually do not exceed few centimetres in size. Their outer surface is grainy, covered with needle-shaped calcareous spicules or with three or four rays. Calcispongiae can present all three types of Porifera organisation: ascon, sycon and leucon. Although some species live isolated, in most cases they form colonies of small individuals, always in shallow water.

The Calcispongiae can be divided into two subclasses according to the different cytological and embryological characters.

- Calcinea: with flaccid consistency, covered with scattered and not entangled spicules.
- Calcaronea: with a solid consistency and entangled spicules.

3.2 Demospongiae

Demospongiae sponges constitute the largest class of the phylum Porifera. They are characterized by siliceous spicules (monactins to tetractins) or by a mixture of siliceous spicules and spongy fibres or only spongy fibres.

From a structural point of view the organisation of the Demospongiae is very complex, with numerous flagellates and small chambers communicating with the inhalation and exhalation pores through a network of channels. Three different subclasses are included in the class Demospongiae, which are differentiated according to the latest findings. [5]

- Verongimorpha: these sponges can be with or without a skeleton. If they have a skeleton it is composed of siliceous asters or spongy fibres, with a laminated cortex and a granular or fibrillar pith. [6]
- Keratosa: with spongy fibres, to which external material is often added for reinforcement, which may be homogeneous or fissured and strongly laminated with pith that degrades into bark
- Heteroscleromorpha: this is the group containing the largest number of Demospongiae, comprising about 5000 species. The name comes from the variability in the type of siliceous spicules, which can be monaxon and/or tetraxon spicules.

3.3 Hexactinellidae

This class of sponges is found predominantly in deep water and is distinguished from other sponges by the presence of a well-organised skeletal structure consisting exclusively of triaxonic siliceous spicules. The siliceous spicules, which may be isolated or form a network, are often

cylindrical in shape, with a large atrial cavity, which has an opening with an operculum.

3.4 Homoscleromorphs

Homoscleromorphs are a small group consisting of less than 100 species of sponges that are specifically positioned in shallow waters from 8 to 60 m. Due to their ability to overpower massive sponges, gorgonians and erect bryozoans, homoscleromorphs can be predominant in some areas. Although some specimens have an ascon organisation, these sponges usually have a leuconic organisation.

Their skeletal framework is very small, and they are often mobile and delicate to the touch. Notably, several species have no spicules at all, and their parenchyma consists only of collagen fibres and often spongin. When they do have spicules, they are called calthrops, which are small and form a poorly organised skeleton.

3.5 Sponge -microrganisms association

The symbiotic relationship that exists between bacteria and eukaryotes is widespread in most sponge species belonging to the Demospongiae, Calcarea, and Hexactinellide classes, both marine and freshwater.[7] These classes are often populated by entire clusters of symbiotic bacteria, both generalists and specialists, living in association with sponges. Given the absence of bacteria in the environment surrounding sponges, symbionts are considered "sponge-specific symbionts." It has been shown that very often, symbiont bacteria contribute 38-57% of the total biomass of sponges.[8] The

bacteria, which may be heterotrophic (eubacteria, archaea) or autotrophic, along with fungi, are predominantly located in the mesohyl. Moreover, since some bacteria seem to be specifically and permanently associated with sponges, the existence of a sponge-bacteria symbiosis is confirmed. The mechanism of symbiosis can be defined as obligatory mutualism when the symbiont is essential for the survival of the host, facultative mutualism when the host benefits from the symbiont but the symbiosis is not essential for the survival of the host, or commensalism when there is no obvious benefit to the host.

Benefits derived from the sponge-microbe association are:

- determining a supply of nutrients through intracellular digestion or transport of metabolites:
- production of new products through nitrogen fixation.[9]
- stabilization of the sponge skeleton.[10]
- protection against predators or fouling through the production of secondary metabolites. [11]

References

[1] Taylor, M.W.; Radax, R.; Steger, D.; Wagner, M. Sponge-Associated Microorganisms: Evolution, Ecology, and Biotechnological Potential. *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 295–347, doi:10.1128/mmbr.00040-06.

[2] E.E., R. *Invertebrate Zoology*; 2004.

[3] D.G., S. *Pennak's Freshwater Invertebrates of the United States: Porifera to Crustacea*; 4th Editio.; 2001.

[4] Gazave, E.; Lapébie, P.; Ereskovsky, A. V.; Vacelet, J.; Renard, E.; Cárdenas, P.; Borchiellini, C. No longer Demospongiae: Homoscleromorpha formal nomination as a fourth class of Porifera. *Hydrobiologia* **2012**, *687*, 3–10, doi:10.1007/s10750-011-0842-x.

[5] Morrow, C.; Cárdenas, P. Proposal for a revised classification of the Demospongiae (Porifera). *Front. Zool.* **2015**, *12*, 1–27, doi:10.1186/s12983-015-0099-8.

[6] Cárdenas, P.; Pérez, T.; Boury-Esnault, N. *Sponge Systematics Facing New Challenges*; 2012; Vol. 61; ISBN 9780123877871.

[7] Sarà, G.; Manganaro, A.; Cortese, G.; Pusceddu, A.; Mazzola, A. The relationship between food availability and growth in Mytilus galloprovincialis in the open sea (southern Mediterranean). *Aquaculture* **1998**, *167*, 1–15, doi:10.1016/S0044-8486(98)00281-6.

[8] Hentschel, U.; Horn, M.; Friedrich, A.B.; Wagner, M.; Moore, B.S. Molecular evidence for a uniform microbial community in sponges. **2002**, *68*, 1–10, doi:10.1128/AEM.68.9.4431.

[9] Wilkinson, C.R.; Fay, P. Nitrogen fixation in coral reef sponges with symbiotic cyanobacteria [13]. *Nature* **1979**, *279*, 527–529, doi:10.1038/279527a0.

[10] Wilkinson, C.R.; Nowak, M.; Austin, B.; Colwell, R.R. Specificity of bacterial symbionts in Mediterranean and Great Barrier Reef sponges. *Microb. Ecol.* **1981**, *7*, 13–21, doi:10.1007/BF02010474.

[11] Unson, M.D.; Holland, N.D.; Faulkner, D.J. A brominated secondary metabolite synthesized by the cyanobacterial symbiont of a marine sponge and accumulation of the crystalline metabolite in the sponge tissue. *Mar. Biol.* **1994**, *119*, 1–11, doi:10.1007/BF00350100.

Chapter 4

Feature-Based Molecular networking for the fast detection of a new cyclic peptide from the sponge *Stylissa caribica*

4.1 Synopsis of this chapter [1]

In the frame of the research of new bioactive natural products from porifera, the value of the molecular networking, as complementation of the traditional dereplication strategies, was proved in the study of the organic extracts from the well-known marine sponge *Stylissa caribica*. Although the Bahamian sponge *Stylissa caribica* has been extensively investigated by many research groups, the molecular networking approach allowed the fast identification of a new cyclic heptapeptide, named stylissamide L (**1**).

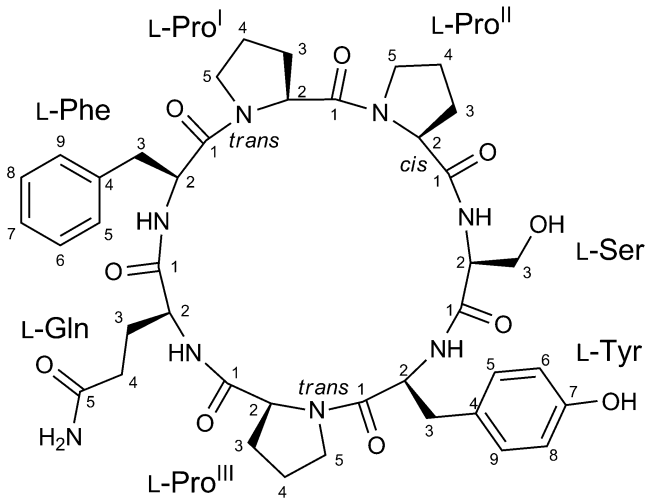


Figure 4.1. Structure of stylissamide L (**1**)

4.2 The marine sponge Stylissa caribica

Sea sponges belonging to the Dictyonellidae and Axinellidae families are a prolific source of new natural bioactive compounds. The marine sponge *Stylissa caribica* was first identified by Lehnert & van Soest in 1998,[2] and it belongs to the order Halichondridae and the family Dictyonellidae. This species has been described as a dark orange sponge with a very irregular surface. Specimens of *S. caribica* are mainly found clinging to rocks in coral reefs.

Figure 4.2. The sponge *Stylissa caribica*
(Image retrieved from https://www.riffgrotte.de/stylissa-sp-oranger-schwamm.html)

The sponge *Stylissa caribica* has been extensively analysed by various research groups, revealing a large pool of secondary metabolites including a wide variety of brominated compounds with a pyrrole-imidazole structure and a few cyclic peptides. A considerable number of brominated alkaloids have been identified over the years from sponges of the order Halichondriae and Agelasidae, whose shared features are the presence of a brominated pyrrole ring and an amino-imidazole ring. The progenitor of the pyrrole-

imidazole alkaloids is oroidin, which can undergo different intra- or intermolecular cyclization reactions, giving rise to a series of derivatives.[3] Twenty brominated pyrrole-imidazole alkaloids have been isolated from the *S. caribica*, and nine of them have been first reported from this sponge, including *N*-methyldibromoisophakellin, oxocyclostylidol, 4-bromopyrrole-2-carboxy-*N*(ε)-lysine, and 4-bromopyrrole-2-carboxyarginine. In addition to the brominated alkaloids, a large pool of cyclic peptides, extracted from both Jamaican and Caribbean *S. caribica* specimens, has been referred to in recent literature. Known cyclic peptides include phakellistatin 2, previously isolated from sponges of the genus *Phakellia* (family Axinillidae), and hymenamides C and F, isolated from sponges of the genus *Hymenidiacion*.[4] The stylissamide group (stylissamide A-H) has enriched the variety of cyclic peptides containing proline residues found in most sponges belonging to the order Halichondrida.[5,6,7] Although extensive screening for bioactivity has been carried out, only the fraction containing stylissamides G and H showed cytotoxic activity against HCT-116 cells (a cell line involved in colon cancer). In addition to the stylissamides, extracts from Jamaican specimens of *S. caribica* showed the presence of two additional cyclic peptides, stylisin 1, whose amino-acid sequence is the reverse of stylissamide D, and stylisin 2.[8]

4.3 Collection and Extraction

Although the Bahamian sponge *Stylissa caribica* has been thoroughly studied, having shown to contain a wide variety of unknown natural products, the molecular networking approach readily provided a fast detection of a novel natural compound, named stylissamide L (**1**).

A frozen sample of the marine sponge *S. caribica*, collected along the coast of the Bahamian island Compass Cay in the Exuma Islands, was extracted with MeOH and $CHCl_3$ at different concentrations. After partitioning the MeOH extract between H_2O and BuOH, the BuOH layer was combined with the other organic extracts. The total organic extract was subjected to reverse phase column chromatography on RP-18 silica gel.

4.4 LC-HRMS/MS Analysis and construction of the molecular network

The RP-18 fractions, rather than the crude organic extract itself, were analysed by LC-HRMS/MS spectrometry. Indeed, the analysis of the RP-18 fractions rather than the organic extract allows better untargeted analysis, reducing the possibility of overlapping peaks in the chromatogram.

Liquid chromatography coupled with high-resolution tandem mass spectrometry (LC-HRMS/MS) was performed using an LTQ Orbitrap instrument equipped with an electrospray source (ESI) and a pentafluorophenyl (PFP) high-performance liquid chromatography (HPLC) column. The four most intense ions from each MS scan of the spectrum were fragmented in subsequent MS/MS scans. A clear molecular network was then obtained after careful pre-processing of the raw data using

MZmine2 and MetGem,[9,10] with the aim to distinguish isomeric compounds and to filtrate adduct and isotopic peaks.

More precisely, raw MS/MS data were uploaded to the MZmine 2.40 program for pre-processing, taking into account the chromatographic dimension: isomeric compounds were distinguished on the basis of their retention time, isotopic peaks were removed, and quantification of the compounds was performed. Following the first standard pre-processing parameters (mass detection, chromatogram construction and chromatogram deconvolution), the MS data obtained from LC-MS/MS runs of individual fractions were grouped into one feature list via the Join aligner module. Adduct peaks ($[M+Na]^+$, $[M+NH_3]^+$, $[M+K]^+$) and ^{13}C and ^{81}Br isotopic peaks were identified and removed, and as a result most compounds gave a unique entry in the feature list. Finally, the converted MS/MS spectra were exported into a .mgf file and its quantitative data into a .csv file. Before uploading the pre-processed data to the GNPS platform, some key parameters of the molecular networking algorithm were optimized with the standalone MetGem program. Indeed, the construction of an informative molecular network is strongly influenced by the nature of the sample, the MS/MS instrument and the methods set up for the LC-MS/MS analysis. In particular, the mass tolerance for peak matching, the minimum required number of matched peaks to calculate a cosine score and the threshold cosine score to connect two nodes were fine-tuned with MetGem, a faster software than GNPS in case of small datasets. A mass tolerance of 0.01 Da for parent and fragment ions, the minimum number of matched peaks of

eight and the minimum cosine score of 0.55 showed to be the most suitable parameters for the generation of the most informative set of clusters.

Lastly, the MS/MS spectra converted into an .mgf file and its quantitative data into a .csv file were submitted on GNPS to build the final molecular network. Contrary to expectations, the feature-based molecular networking, and the old metabolomics workflow on the GNPS platform produced different results, even though we used the same networking parameters from the same .mgf file. In particular, the molecular map obtained with the metabolomics workflow showed smaller clusters and many nodes were absent (including the stylissamide L node). Interestingly, the metabolomics workflow produced the same results as the MetGem programme. The reason for this unexpected result is still unknown.

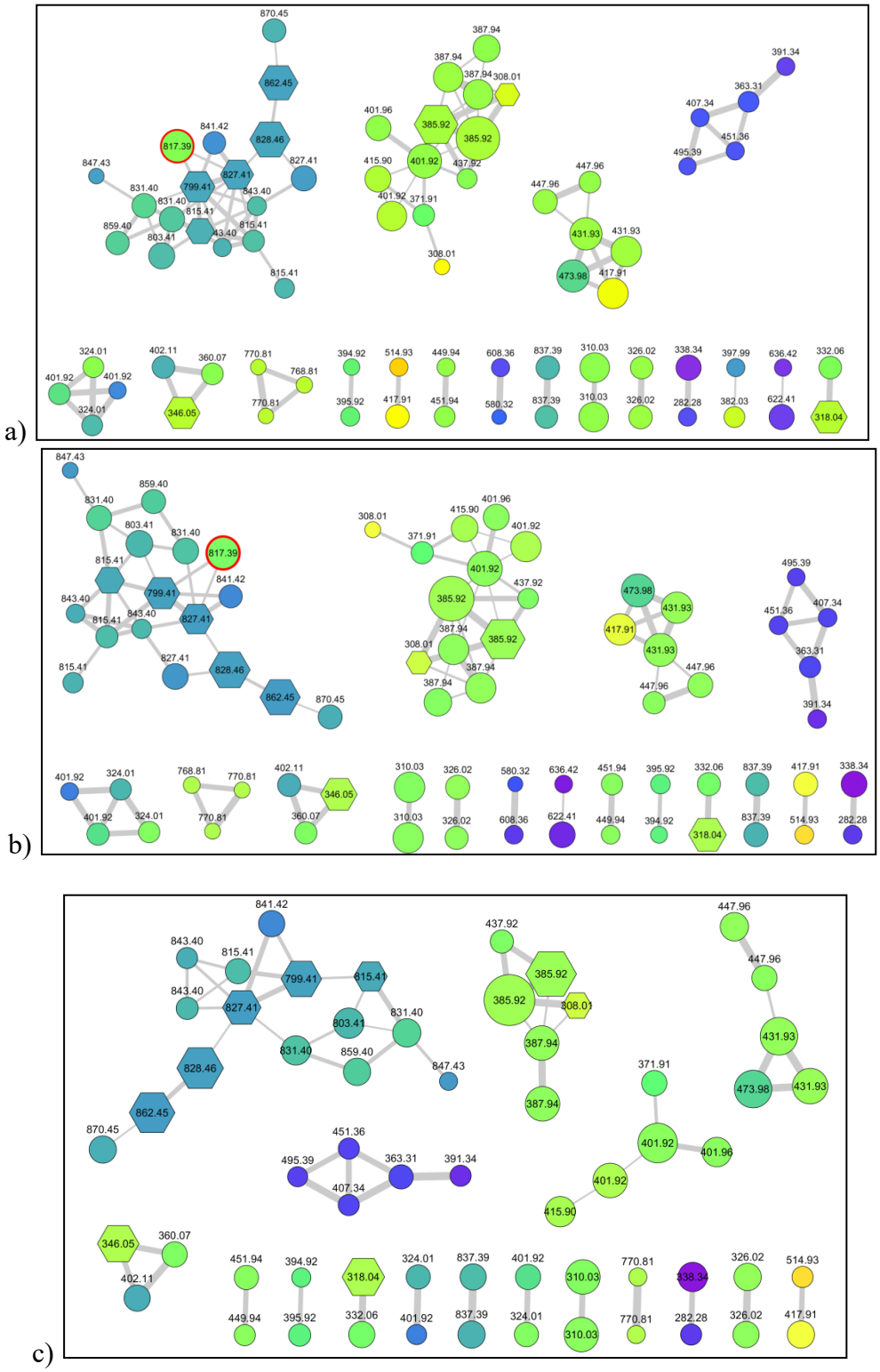


Figure 4.3. Molecular networks of the extract of *S. caribica* obtained with the same data and the same parameters (*m*/*z* tolerance 0.01 Da, cosine score > 0.55, matched peaks > 8, maximum number of neighbor nodes = 10) using (a) the program MetGem (b) the Metabolomics workflow on GNPS, and (c) the Feature-Based Molecular Network workflow on GNPS.

In the output of the MS-based molecular networking, most peptides were grouped into a single cluster. Accurate identification of many known compounds was performed by Dereplicator tool in GNPS,[11] by comparing each spectrum with a wide database of MS/MS spectra, enriched thanks to contributions from all GNPS users. The final visualization of the network was obtained using the Cytoscape software, importing the important features from the quantitation file exported from MZmine2.[12]

In the network, the colour of the node was mapped to the relevant retention time, providing a visual cue to the polarity of the metabolite, and the size of the node was made proportional to the amount of the metabolite. In addition, nodes recognized by Dereplicator as putatively known compounds were represented as hexagons.

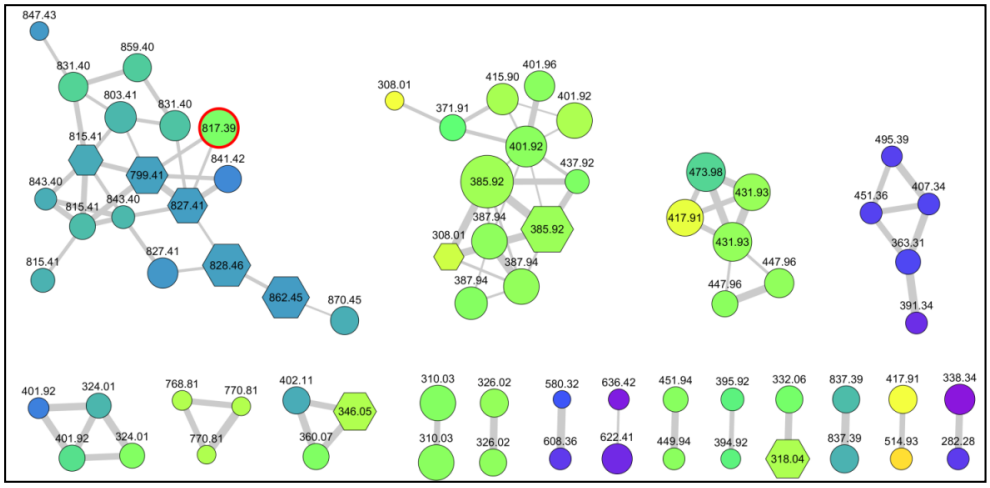


Figure 4.4. The feature-based molecular network of *S. caribica*. The colour of nodes is mapped according to retention times, and their size is related to the amounts of the metabolite. Annotated nodes are shown as hexagons and can be identified using Table 4.1 The node of stylissamide L (**1**) is marked with red borders.

As *S. caribica* is rich in brominated compounds, most of the clusters in the network were related to these compounds. However, the largest cluster in

the network turned out to be related to cyclic peptides. Five of the nodes in the cyclic peptide cluster could be annotated as known peptides, two of which were reported for the first time in this sponge. The other 13 nodes appeared to be new compounds, because their molecular formula did not correspond to any known natural peptide. Surprisingly, the most abundant unknown cyclic peptide (*m/z* 817.39) showed a much shorter retention time than the other peptides in the cluster, so it was eluted in the earlier RP-18 fraction (fraction F3) compared to the fraction where most of the other peptides were eluted (fraction F4).

Table 4.1 Cyclic heptapeptides in the sponge *Stylissa caribica*.

Cyclic peptide	$[M+H]^+$	Exact Mass	Retention time RP-18[a]	Retention time PFP[b]	Citation
Stylissamide A	$C_{44}H_{61}N_8O_9$	845.4556	1.72	26.34	[5,6,7]
Stylissamide B	$C_{44}H_{58}N_7O_8$	812.4341	ND	ND	[5,6,7]
Stylissamide C	$C_{48}H_{60}N_7O_8$	862.4498	20.29	32.66	[5,6,7]
Stylissamide D	$C_{45}H_{62}N_7O_8$	828.4654	19.55	32.48	[5,6,7]
Stylissamide E	$C_{39}H_{59}N_8O_9$	783.4400	13.29	28.76	[6,7]
Stylissamide F	$C_{43}H_{57}N_{10}O_9$	857.4304	13.94	32.21	[6,7]
Stylissamide G	$C_{45}H_{62}N_7O_7$	812.4705	ND	ND	[7]
Stylissamide H	$C_{44}H_{59}N_8O_8$	827.4450	ND	ND	[7]
Stylissamide L	$C_{41}H_{53}N_8O_{10}$	817.3876	6.21	24.95	this work
Hymenamide C	$C_{43}H_{55}N_8O_9$	827.4087	20.71	32.58	[4]
Hymenamide F	$C_{35}H_{61}N_{10}O_8S$	781.4389	1.67	26.40	[4]
Phakellistatin 3	$C_{42}H_{55}N_8O_9$	815.4087	18.50	31.89	*c*
Phakellistatin 13	$C_{42}H_{55}N_8O_8$	799.4137	20.09	32.47	[8]
Stylisin 1	$C_{45}H_{62}N_7O_8$	828.4654	ND	ND	[5,6,7,8]
Stylisin 2	$C_{44}H_{58}N_7O_8$	812.4341	ND	ND	[5,6,7,8]

a. Experiments were performed with a Kinetex 5 µm, 50 mm × 2.10 mm C18 column using a flow rate of 200 µL/min and the following elution gradient: 10% MeOH for 1 min, 10%−100% MeOH over 30 min, and 100% MeOH for 10 min.

b. Experiments were performed with a Kinetex 5 µm, 100 mm × 2.1 mm PFP column using a flow rate of 200 µL/min and the same elution gradient of H_2O and MeOH

c. This compound was not previously detected in specimens of *Stylissa caribica*. It was putatively identified by the Dereplicator tool in GNPS.

*4.5 Isolation of stylissamide L (**1**)*

Isolation of the new compound was performed in a single reverse-phase HPLC chromatography step, yielding a pure compound of 7.2 mg, and named stylissamide L (**1**).

*4.6 Structural elucidation of stylissamide L (**1**)*

The high-resolution ESI mass spectrum of stylissamide L (**1**) showed peaks of $[M+H]^+$ and $[M+Na]^+$ ions at *m/z* 817.3876 and *m/z* 839.3694, which established its molecular formula as $C_{41}H_{53}O_{10}N_8$ with 20 unsaturations.

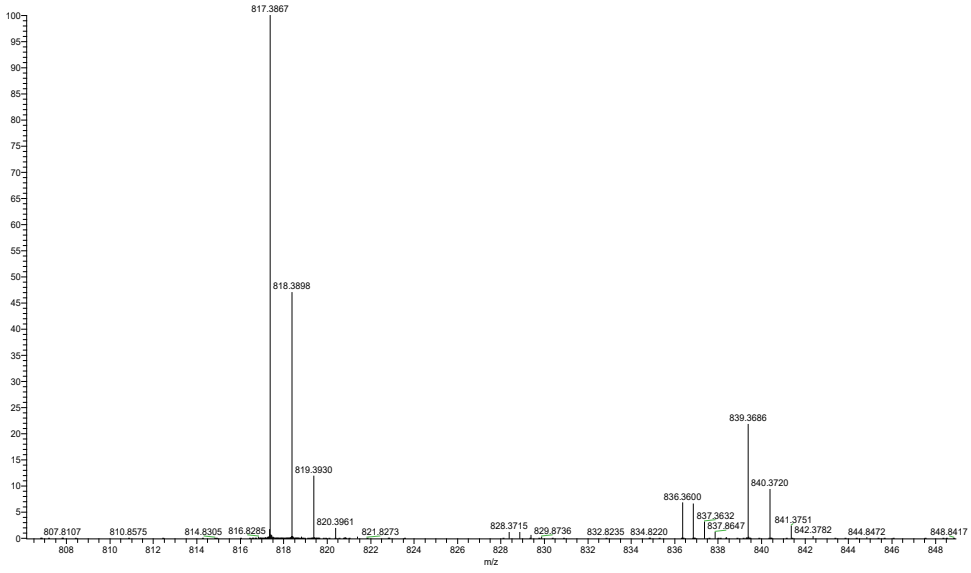


Figure 4.5. Positive ion mode high-resolution ESI mass spectrum of stylissamide L (**1**).

The cyclic heptapeptide structure was confirmed by analysing the fragmentation pattern observed in the MS/MS spectrum of compound **1**.

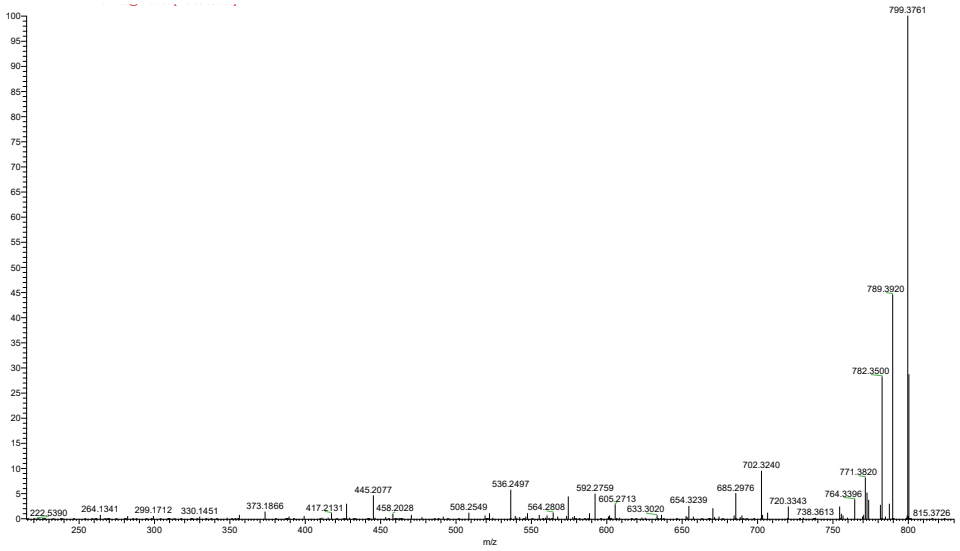


Figure 4.6. Positive ion mode high-resolution ESI MS/MS spectrum of stylissamide L (**1**).

Fragments corresponding to the loss of H_2O and CO and a phenylalanine, glutamine, tyrosine, and proline residue were evident. In addition to the four residues, whose loss from the molecular ion was detected in the MS/MS spectrum, the molecular formula was satisfied with the presence of a serine and two additional proline residues. The amino acid composition of compound **1** was subsequently verified by a full set of NMR analyses. The cyclic structure of compound **1** was confirmed, considering that the seven amino acids accounted for only 19 unsaturation degrees rather than 20 unsaturations as determined by the molecular formula.

A complete array of two-dimensional homonuclear and heteronuclear NMR spectra (COSY, TOCSY, NOESY, HSQC and HMBC) were recorded to confirm the structure of compound **1**. Four NH amide signals and seven α-proton signals were observed in the proton spectrum, as expected for a cyclic heptapeptide with three proline residues. Cross peaks with the corresponding α-proton or amide NH signals in the TOCSY spectrum

allowed the aliphatic proton of each residue to be elucidated, and their assignment was obtained using the COSY and HSQC spectra.

Table 4.2. NMR data of stylissamide L (**1**) (^{1}H 700 MHz, ^{13}C 175 MHz, DMSO-d_6)

AA	pos.	δ_C, type		δ_H, mult (J in Hz)	AA	pos.	δ_C, type		δ_H, mult (J in Hz)
Pro^{I}	1	170.3, C			**Pro^{III}**	1	171.9, C		
	2	59.1, CH		4.34, dd (5.1, 8.6)		2	63.1, CH		4.06, t (8.7)
	3	28.1, CH_2	a	2.15, m		3	28.7, CH_2	a	2.22 m
			b	1.75, m				b	1.81, m
	4	24.3, CH_2		1.87, m		4	25.0, CH_2	a	2.11, m
	5	46.7, CH_2	a	3.45, m				b	1.98, m
			b	3.36, m		5	46.9, CH_2	a	3.93, ddd (6.8, 9.8, 9.8)
Pro^{II}	1	171.8, C						b	3.82, m
	2	60.1, CH		4.28, dd (1.5, 8.8)	**Gln**	NH			8.17, d (7,0)
	3	31.8, CH_2	a	2.16, m		1	170.7, C		
			b	2.00, m		2	52.8, CH		4.05, ddd (4.3, 7.0, 10.0)
	4	21.7, CH_2	a	1.77, m		3	25.9, CH_2	a	1.85, m
			b	1.57, m				b	1.73, m
	5	46.8, CH_2	a	3.60, ddd (1.5, 8.4, 10.8)		4	31.5, CH_2	a	2.13, ddd (7.2, 15.7, 7.2)
			b	3.33, ddd (10.8, 10.8, 7.1)				b	2.04, ddd (7.2, 15.7, 7.2)
Ser	NH			7.65, d (5.9)		5	174.5, C		
	1	167.7, C				5-NH_2			6.92, s
	2	60.0, CH		3.85, ddd (3.6, 5.9, 10.2)	**Phe**	NH			7.11, d (7.2)
	3	60.9, CH_2	a	3.46, dd (10.2, 11.9)		1	167.5, C		
			b	3.14, dd (11.9, 3.6)		2	51.5, CH		4.69, ddd (5.8,7.2, 8.0)
Tyr	NH			7.34, d (9.1)		3	36.9, CH_2	a	3.18, dd (8.0, 14.2)
	1	171.5 C						b	2.71, dd (5.8, 14.2)
	2	51.5 CH		4.88 ddd (3.2, 9.1,10.9)		4	138.0, C		
	3	37.0 CH_2	a	3.35, dd (3.2,13.5)		5/9	128.9, CH		7.16, d (7.5)
			b	2.42, dd (10.9, 13.5)		6/8	126.0, CH		7.18, t (7.3)
	4	126.6 C				7	128.0, CH		7.22, t (7.5)
	5/9	130.5 CH		7.08, d (8.5)					
	6/8	114.9 CH		6.66, d (8.5)					
	7	156.0 C							
	7-OH			7.42, s					

Moreover, the HMBC data enabled the amino acid sequence in the peptide to be determined. Additionally, to the standard HMBC experiment, a band-selective HMBC experiment was performed to discriminate the CO signal with very close ^{13}C chemical shift, such as Pro^{II}-C1 and Pro^{III}-C1, increasing the resolution in the ^{13}C dimension (Figure 4.18). Diagnostic HMBC correlations used to elucidate amino acid sequence are shown in Figure 4.7.

The chemical shift assignment of the ^{13}C carbonyl signals of each amino acid was accomplished (except for Ser) based on their HMBC correlations with one or both protons at the respective β-methylene (i.e., at position 3). In addition, the bonds between the amino acid residues were established from the HMBC correlations of the four amide protons (Ser-NH with Pro^{II}-C1, Tyr-NH with Ser-C1, Gln-NH with Pro^{III}-C1, Phe-NH with Gln-C1) and the ε-protons of the proline (Pro^{I}-5b with Phe-C1 and Pro^{II}-5b with Pro^{I}-C1). Thus, the amino acid sequence was defined as cyclo (Pro-Pro-Ser-Tyr-Pro-Gln-Phe).

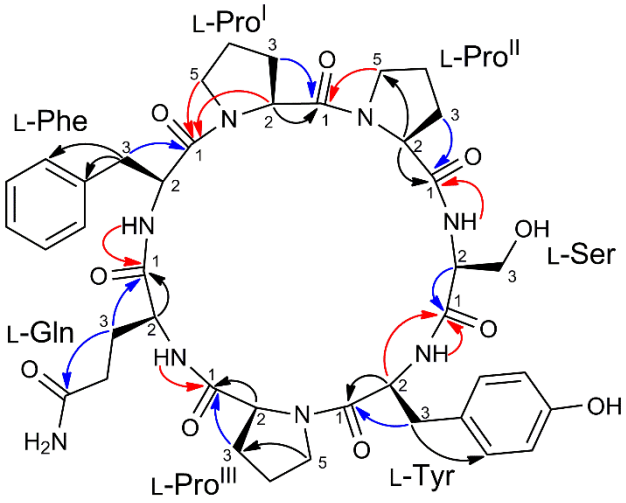


Figure 4.7 Diagnostic HMBC correlation used to elucidate amino acid sequence in stylissamide L (**1**). Blue arrows represent correlations used to assign carbonyl ^{13}C signals, while intra-residual correlations are noted with red arrows.

Table 4.3. Full NMR data of stylissamide L (**1**) (^{1}H 700 MHz, ^{13}C 175 MHz, DMSO-d_6).

AA	pos.	δ_C, type		δ_H, mult (*J* in Hz)	NOESY	HMBC
Pro^{I}	1	170.3, C		-		
	2	59.1, CH		4.34, dd (5.1, 8.6)	Ser-NH, Tyr-NH, Phe-NH	Phe-1
	3	28.1, CH_2	a	2.15, m	Pro^{II}-4b, Pro^{II}-3a	Pro^{I}-1
			b	1.75, m	Pro^{II}-2, Pro^{II}-4b	Pro^{I}-1
	4	24.3, CH_2		1.87, m		
	5	46.7, CH_2	a	3.45, m		
			b	3.36, m		Phe-1,Pro^{I}-2
Pro^{II}	1	171.8, C		-		
	2	60.1, CH		4.28, dd (1.5, 8.8)	Ser-NH, Pro^{I}-2, Pro^{I}-3b	Pro^{II}-1, Pro^{II}-5
	3	31.8, CH2	a	2.16, m		Pro^{II}-1
			b	2.00, m	Ser-NH	
	4	21.7, CH2	a	1.77, m	Pro^{III}-3a	
			b	1.57, m		
	5	46.8, CH2	a	3.60, ddd (1.5, 8.4, 10.8)	Ser-NH	
			b	3.33, ddd (10.8, 10.8, 7.1)	Pro^{III}-4b	Pro^{I}-1
Ser	NH			7.65, d (5.9)	Tyr-NH, Pro^{I}-2, Pro^{II}-2, Pro^{II}-5a, Pro^{II}-4b	Pro^{II} -1
	1	167.7, C		-		
	2	60.0, CH		3.85, ddd (3.6, 5.9, 10.2)	Tyr-NH	Ser -1
	3	60.9, CH_2	a	3.46, dd (10.2, 11.9)	Tyr-NH, Tyr-5/9	
			b	3.14, dd (11.9, 3.6)		
Tyr	NH			7.34, d (9.1)	Ser-NH, Ser-3a	Ser-1
	1	171.5 C		-		
	2	51.5 CH		4.88 ddd (3.2, 9.1,10.9)	Pro^{III}-5a/b	Tyr- 1
	3	37.0 CH2	a	3.35, dd (3.2,13.5)	Pro^{III}-5a/b, Phe-NH, Gln-NH	Tyr-5/9
			b	2.42, dd (10.9, 13.5)	Gln-NH, Phe-NH, Pro^{III}-5a	Tyr-1, Tyr-4
	4	126.6 C		-		
	5/9	130.5 CH		7.08, d (8.5)	Tyr-3a, Tyr-3b	Tyr-7
	6/8	114.9 CH		6.66, d (8.5)		Tyr-7
	7	156.0 C		-		
	7-OH			7.42, s		
Pro^{III}	1	171.9, C		-		
	2	63.1, CH		4.06, t (8.7)	Phe-NH,	Pro^{III}-1
	3	28.7, CH_2	a	2.22 m		
			b	1.81, m	Gln-NH	Pro^{III}- 1
	4	25.0, CH_2	a	2.11, m	Gln-NH, Gln-3a, Gln-3b	
			b	1.98, m		
	5	46.9, CH_2	a	3.93, ddd (6.8, 9.8, 9.8)	Tyr-2, Tyr-3a	
			b	3.82, m	Tyr-2, Gln-NH, Tyr-3a	Pro^{III} -3
Gln	NH			8.17, d (7,0)	Phe-NH, Tyr-3a, Pro^{III}-2, Pro^{III}-3b, Pro^{III}-4a	Pro^{III}-1
	1	170.7, C		-		
	2	52.8, CH		4.05, ddd (4.3, 7.0, 10.0)	Phe-NH, Pro^{III}-4b	Gln-1
	3	25.9, CH_2	a	1.85, m	Pro^{III}-4a	Gln-1, Gln-5
			b	1.73, m	Pro^{III}- 4a	Gln-5
	4	31.5, CH_2	a	2.13, ddd (7.2, 15.7, 7.2)	Pro^{III}-5b, Pro^{III}-3a, Pro^{III}-3b	Gln-5
			b	2.04, ddd (7.2, 15.7, 7.2)	Pro^{III}-4a	Gln-5
	5	174.5, C		-		
	5-NH_2			6.92, s	Gln-4a, Gln-4b	Gln-5
Phe	NH			7.11, d (7.2)	Phe-2, Gln-NH, Pro^{III}-2, Pro^{I}-5b, Tyr-NH	Gln-1
	1	167.5 C				
	2	51.5, CH		4.69, ddd (5.8,7.2, 8.0)	Pro^{I}-5a, Pro^{I}-5b	Phe-1, Phe-4
	3	36.9 CH_2	a	3.18, dd (8.0, 14.2)		Phe-1, Phe-4
			b	2.71 (5.8, 14.2)		Phe-1, Phe-5/9
	4	138.0, C		-		
	5/9	128.9, CH		7.16, d (7.5)		
	6/8	126.0, CH		7.18, t (7.3)		
	7	128.0, CH		7.22, t (7.5)		

*4.7 Stereostructural determination of stylissamide L (**1**)*

An advanced Marfey's method was performed to clarify the absolute configuration of the seven amino acid residues.[13] A high-resolution Orbitrap MS instrument was used as detector to analyse the reaction results, improving sensitivity and specificity, and allowing the use of only a few μg of sample. Firstly, 32 μg of stylissamide L (**1**) was subjected to total hydrolysis by treating it with 6 N HCl/AcOH (1:1) at 120 °C for 18 h. The derivatization of the amino acid residues with the D enantiomer of Marfey's reagent (1-fluoro-2-4-dinitrophenyl-5-D-alanine amide, or D-FDAA) was performed by adding 100 μL of 1% D-FDAA. It is important to note that under the total hydrolysis conditions used, the glutamine residue was converted to glutamic acid. The retention times recorded by high-resolution LC-MS analysis of the obtained D-FDAA derivatives of Pro, Ser, Tyr, Glu and Phe were compared with those of the authentic standards prepared from the reaction of L- and D-FDAA with L-Pro, D-Ser, L-Tyr, L-Glu, L-Phe. LC-MS analysis revealed the L configuration for all amino acids. The presence of only L amino acids was consistent with the other heptacyclopeptides of the stylissamide class.

Like other stylissamides, stylissamide L (**1**) appears to have a highly structured conformation, as deduced from the NOESY spectrum, which showed many correlations between topologically distant protons (e.g., Tyr-NH with Phe-NH or Tyr-NH with Pro^{I}-H2).[7]

The electron circular dichroism (ECD) spectrum showed a fairly complex band structure, with a positive Cotton effect at 236 nm and negative Cotton effects at 219 and 202 nm. Recent literature reveals the presence of

configurational isomerism on proline peptide bonds in strained cyclic peptides such as, for example, in stylissamide H and euryanine A.[7] Based on this assumption, the *cis* or *trans* geometry of the binding of proline residues to the preceding amino acid must be considered a configuration rather than a conformation.

Pro^{II} was determined to be *cis* because of NOESY cross peak between Pro^{II}-H2 and Pro^{I}-H2, and because of the difference between the ^{13}C NMR chemical shifts of Pro^{II}-C3 and Pro^{II}-C4 greater than 8.0 ppm, with Pro^{II}-C4 below 23.3 ppm, in accordance with an empirical rule discussed in ref 5.[5] Similarly, Pro^{I} and Pro^{III} were inferred to be *trans* because the respective chemical shift differences (3.8 and 3.7 ppm) between C-3 and C-4 were well below the 8.0 ppm threshold.

Structurally, stylissamide L (**1**) is similar to the other compounds of the stylissamide class, which are all proline-rich heptapeptides (containing two to four proline residues); however, unlike the other stylissamides, it contains a serine residue. The reason why stylissamide L (**1**) is poorly retained by the RP-18 stationary phase is not immediately clear. The reason cannot be due to the lack of aliphatic amino acids other than proline, because stylissamide F, which exhibits the same characteristics was retained more than compound **1**. On the other hand, compounds appearing to show a similar polarity, such as stylissamide A, have an even shorter retention time than stylissamide L (**1**) (Table 4.1). Probably, peptide conformation can prevent interaction between the hydrophobic regions of the molecule and the chromatographic stationary phase by affecting the RP-18 retention times of stylissamides.

*4.8 Evaluation of cytotoxicity of stylissamide L (**1**)*

The conformational characteristics of stylissamide L (**1**) and the cytotoxic activity exhibited by certain peptides found in *Stylissa* sponges, suggested the evaluation of the growth inhibitory effects of stylissamide L (**1**).[7] The assays were performed on two different cancer cell lines: BxPC-3 from pancreatic cancer and MCF-7 from breast cancer by evaluating dynamic impedance monitoring based on cell proliferation. The cell growth of both cell lines appeared largely unaltered even at the highest dose tested. Furthermore, due to structural similarities between compound **1** and a known cell motility inhibitor stylissamide X,[14] the ability of the new compound to inhibit cell migration was evaluated. Cell migration is a crucial aspect of cancer cell spreading and metastasis formation, making it an attractive target in cancer treatment. Cell motility of MCF-7 and 3AB-OS osteosarcoma steam cells was unaffected or slightly increased at a concentration of 50 μM of compound.

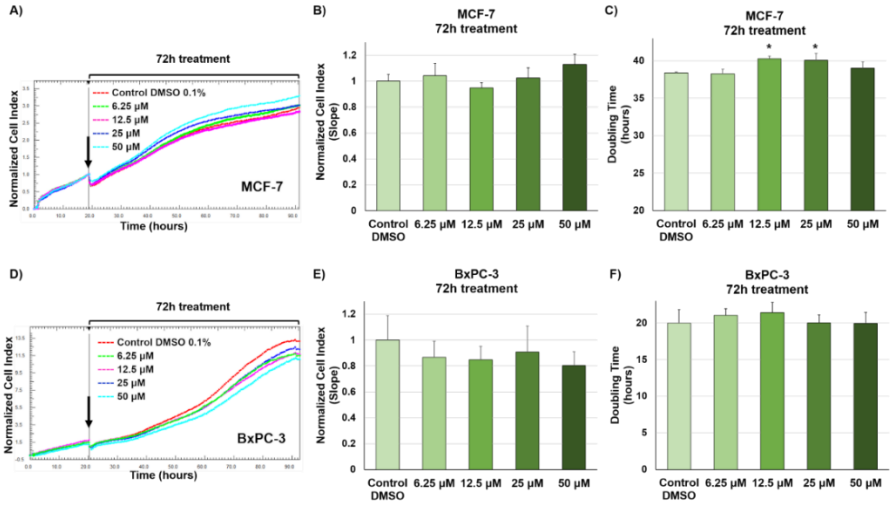


Figure 4.8. Real time monitoring of cancer cell proliferation after exposure to stylissamide L (**1**) and DMSO vehicle, by using the xCELLigence System Real-Time Cell Analyzer. (**A**,**D**) Normalized cell index (NCI) traces of MCF7 (**A**) and BxPC-3 (**D**) cells treated with different concentrations (6.25, 12.5, 25, and 50 µM) of stylissamide L (**1**) and DMSO vehicle (0.1%) for 72 hours. Black arrow shows the start of drug treatment. Each cell index value was normalized at this time. (**B**,**E**) Slope values of growth curves of MCF-7 (**B**) and BxPC-3 (**E**) cells after 72h exposure to different concentrations (6.25, 12.5, 25, and 50 µM) of stylissamide L (**1**) and DMSO vehicle (0.1%). NCI slope values are relative to controls treated with DMSO vehicle. (**C**,**F**) Doubling times of NCI of MCF-7 (**C**) and BxPC-3 (**F**) cells after 72h treatment with different concentrations (6.25, 12.5, 25, and 50 µM) of stylissamide L (**1**) and DMSO (0.1%). Data are presented as mean ± SD; n=3. Statistical significances are referred to the DMSO control. One-way analysis of variance (ANOVA) was applied to compare means of groups and Dunnett's method was used as a post-hoc test to compare multiple groups versus the control group. *p*-values < 0.05 were considered to be statistically significant. Statistical analysis was performed using the GraphPad Prism Software Version 5. * $p < 0.05$.

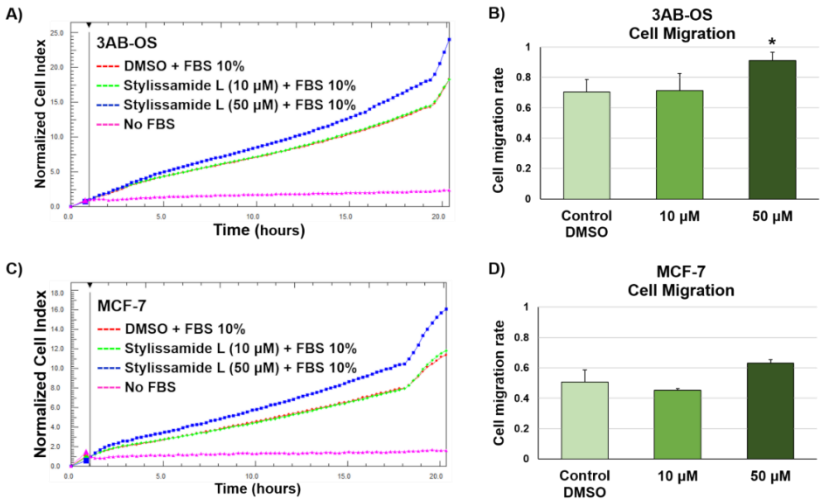


Figure 4.9. Real-time monitoring of 3AB-OS and MCF-7 cell migration after exposure to stylissamide L (**1**). (**A**,**C**) NCI traces of 3AB-OS (**A**) and MCF-7 (**C**) cells seeded with compound **1** or DMSO (0.1%) vehicle, in presence of 10% Fetal Bovine Serum (FBS) as the chemoattractant. Migration was monitored for 20 hours, using the xCELLigence System equipped with specially designed 16-well plates (CIM-plate 16). (**B**,**D**) Migration activity of 3AB-OS (**B**) and MCF-7 (**D**) cells seeded with compound **1** or DMSO (0.1%) vehicle, in presence of 10% Fetal Bovine Serum (FBS) as the chemoattractant. Cell migration rates were recorded for 20 hours and expressed as slope values of NCI curves. Data are presented as mean ± SD; n=3. Statistical significances are referred to the DMSO control. Two-group comparisons were performed using Student's t-test. P-values < 0.05 were considered to be statistically significant. Statistical analysis was performed using the GraphPad Prism Software Version 5.

4.9 Experimental section

4.9.1 General experimental procedures

Optical rotation was measured with a Jasco P-2000 polarimeter (Jasco Europe s.r.l., Cremella, Italy) at the sodium D-line. A Bruker Avance Neo spectrometer (Bruker BioSpin Corporation, Billerica, MA, USA) at 700 MHz with dimethylsulfoxide-d6 (DMSO-d_6) as solvent was used to perform ^{1}H NMR and 2D NMR experiments. All chemical shifts were referenced to the residual solvent signal (δ_H 2.50, δ_C 39.5). The HSQC spectra were optimized for $^1J_{CH}$ = 142 Hz and the HMBC experiments for $^{2,3}J_{CH}$ = 8.3 Hz. A NOESY experiment with a mixing time of 300 ms was analyzed to highlight through space ^{1}H connectivity. A Thermo LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific Inc.,Waltham, MA, USA) combined with a Thermo U3000 HPLC system, was used to record high-resolution ESI-MS and HR-ESI-HPLC experiments. High-performance liquid chromatography (HPLC) separations were achieved on an Agilent 1260 Infinity quaternary LC apparatus (Agilent Technology, Cernusco sul Naviglio, Italy), equipped with a diode-array detector (DAD).

4.9.2 Collection, extraction, isolation

The specimen of *Stylissa caribica*, studied, was found at a depth of 28 m by Scuba along the coast of Compass Cay Island, Exuma Island of the Bahamas (GPS 24° 16.372′ N, 76° 30.141′ W) during a ship-based research expedition in 2010. After collection, the sponge was immediately frozen and kept at a temperature of -20 °C until extraction. The frozen sponge, with a wet weight of 154 g, was first chopped into small pieces and extracted at room temperature with MeOH (4 × 1.5 L), with different mixtures of MeOH

and $CHCl_3$ (2:1, 1:1, 1:2) and finally with $CHCl_3$ (2 × 1.5 L). The orange colored MeOH extract was partitioned between H_2O and *n*-BuOH, the latter was merged with the $CHCl_3$ extracts and dried under vacuum. The total organic extract (8.08 g) was chromatographed using a reversed-phase column chromatography on RP-18 silica gel. LC-MS/MS analysis was performed on RP-18 fractions F3 (eluted with 60% MeOH, 1060 mg), F4 (80% MeOH, 830 mg), F5 (90% MeOH, 410 mg), and F6 (MeOH/$CHCl_3$ (9:1), 460 mg). Fraction F3, which contained the new compound stylissamide L (**1**), was subjected to reversed-phase HPLC separation on a Luna (Phenomenex) C18 column (250 × 10 mm, 10 μm) (Eluent A: 0.1% HCOOH in H_2O; eluent B: MeOH; gradient schedule: 25% B 5 min, 25% → 50% B over 27 min, 50% → 100% B over 3 min, 100% B 7 min; flow rate 5 mL min^{-1}, wavelength 230 nm) to collect the fraction (tR = 24 min) containing 7.2 mg of pure compound **1**.

Stylissamide L (**1**): light yellow powder; $[\alpha]^{20}_{D}$–40 (*c* 0.23, acetonitrile); UV (ACN): λ_{max} (ε) 277 (1250), 232 (5900, shoulder), 195 (50500); ECD (ACN): λ_{max} (Δε) 236 (+4.9), 219 (−15.3), 202 (−18.6); high resolution ESI-MS (positive ion mode, MeOH) *m/z* 817.3876 ($[M + H]^+$, $C_{41}H_{53}O_{10}N_8^+$, calcd. 817.3879), 839.3694 ($[M + Na]^+$, $C_{41}H_{52}O_{10}N_8Na^+$, calcd. 839.3699). ^1H and ^{13}C NMR (DMSO-d6): Table 4.2

4.9.3. LC-HRMS and LC-HRMS/MS

All LC-MS and LC-MS/MS experiments were performed on a Thermo LTQ Orbitrap XL high-resolution equipped with an ESI source, coupled to a

Thermo U3000 HPLC system. Experiments were carried out with a Kinetex 5 μm, 100 mm × 2.1 mm PFP column (Phenomenex, Torrance, CA, USA), at a temperature of at 25 °C, with an elution gradient of H_2O and MeOH running and a flow rate of 200 μL/min. The gradient program was set as follows: 10% MeOH for 1 min, 10−100% MeOH over 30 min, and 100% MeOH for 10 min. Mass spectra were acquired in positive ion detection mode, with resolution set to 60,000 in the range of *m/z* 150–2000. MS parameters used were a spray voltage of 4.80 kV, a capillary temperature of 285 °C, a sheath gas rate of 32 units N_2 (ca. 320 mL/min), and an auxiliary gas rate of 15 units N_2 (ca. 150 mL/min). Data were recorded with data-dependent acquisition (DDA) mode. In particular, the four most intense ions in the full-scan mass spectrum were subjected to high-resolution tandem mass spectrometry (HRMS/MS). HRMS/MS scans were achieved for selected ions with a collision induced dissociation (CID) fragmentation, an isolation width of 3.00 Da, a normalized collision energy of 35 units, an activation Q of 0.250 units, and an activation time of 30 ms.

The analysis of mass data was accomplished with the Thermo Xcalibur software version 2.2 (Thermo Fisher Scientific Inc., Waltham, MA, USA).

4.9.4. LC-HRMS/MS Data Processing and Molecular Networking

Raw LC-HRMS/MS data were processed in batch mode with the software MZmine2 version 2.51.[10] The pre-processed data were presented on MassIVE (accession number: MSV000085867). Molecular networking was performed with both MetGem version 1.2.2 and the GNPS website.[9] The parameters used were *m/z* tolerance 0.01 Da, cosine score > 0.55, matched

peaks > 8, maximum number of neighbouring nodes = 10, maximum number of nodes in a single network = 100. The molecular netwoking output was visualised with Cytoscape version 3.7.1 by importing both the .mgf file and the quantification file generated by MZmine2.[12]

Accurate identification of many known compounds was performed by the Dereplicator V2 tool in GNPS by establishing a precursor ion mass tolerance and fragment ion mass tolerance of 0.02 Da.[11]

4.9.5. Advanced Marfey's Analysis

An amount of 32 μg of stylissamide L (compound **1**) was hydrolysed with 500 μL 6 N HCl/AcOH (1:1) at 120 °C for 18 h. The residual HCl fumes were removed under a direct flow of N_2. The hydrolysate of stylissamide L (compound **1**) was dissolved in TEA/acetone (2:3, 100 μL) and 1% 1-fluoro-2,4-dinitrophenyl-5-D-alaninamide (D-FDAA) in CH_3CN/acetone (1:2) (100 μL). The mixture was heated at 50 °C for 2 h and dried under N_2 flow. It is important to note that in the hydrolysis conditions used, the glutamine residue was transformed into glutamic acid. The resulting D-FDAA derivatives of all amino acids (Pro, Ser, Tyr, Gln, Phe) were dissolved in MeOH (100 μL) for subsequent analysis. Authentic standards of L-Pro, D-Ser, L-Tyr, L-Glu and L-Phe were derivatized with both L-FDAA and D-FDAA following the same steps as described above. The retention times of Marfey's derivatives of compound **1** were compared with those of the standard derivatives by LC-HRMS/MS using a Kinetex C18 (Phenomenex) 150 x 2.1 mm, 5 μm column., setting the gradient conditions as follows: 35 min prerun with 5% ACN, 5% ACN 3 min, 5% → 50% ACN

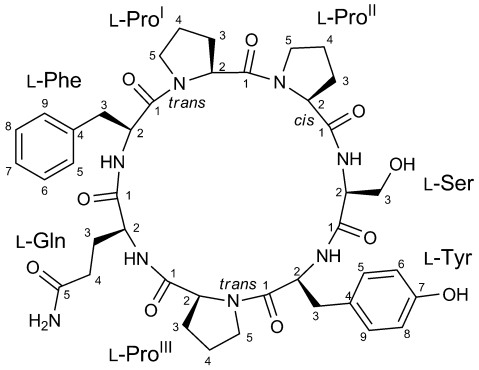
over 30 min, 50% ACN 1 min, 50% → 90% ACN 1 min, 90% ACN 6 min. Spectra acquisition was performed in positive ion detection mode, and raw data were analysed with the Xcalibur program suite.

4.9.6 Cell Proliferation and Migration Assays

Cell proliferation assays were performed using the xCELLigence System real-time cell analyser (ACEA Biosciences, San Diego, CA, USA).

In addition, migration activity of compound **1** was also evaluated using xCELLigence system, but in this case, it was equipped with electronic cell invasion and migration plates (CIM-Plate 16). These plates are composed of upper and lower chambers, separated by a microporous membrane coated with gold microelectrodes, which display decreased electrical conductivity when cells adhere to their surface while moving towards the lower chamber. For the migration assay, 5.0×10^4 cells/well were seeded in the upper chamber with stylissamide L (**1**) or 0.1% DMSO vehicle, in a serum-free growth medium. Growth medium supplemented with the chemoattractant 10% Fetal Bovine Serum (Gibco-Thermo Scientific, Waltham, MA, USA) was added to the lower chambers. Nothing was added to the negative control. Cell migration was monitored every 15 min for 20 h, through real time measurement of electronic impedance variations recorded by the microelectrodes located on the lower side of the microporous membrane. Data were analyzed by the Real-Time Cell Analyzer (RTCA)-integrated software (Version 2.0.0.1301, ACEA Biosciences, San Diego, CA, USA).

4.10 Spectroscopic data



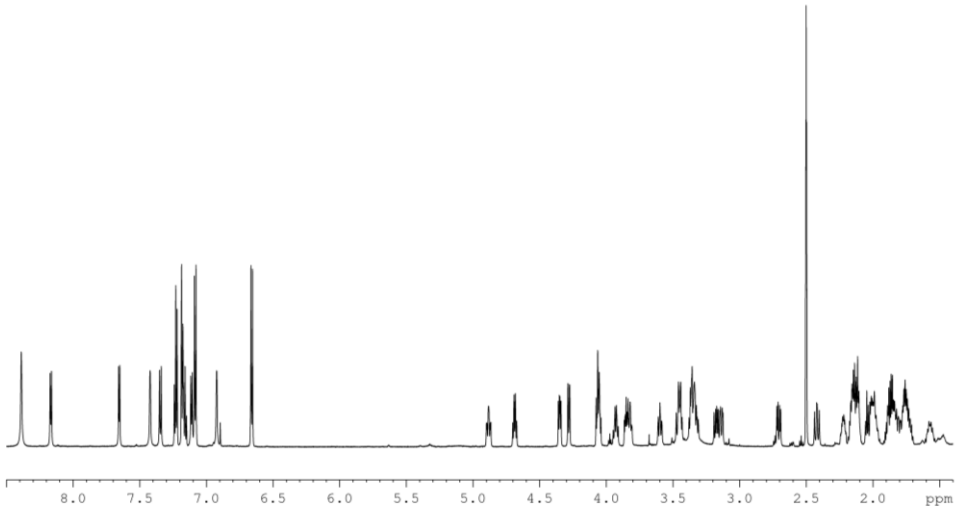


Figure 4.10. ^{1}H-NMR spectrum of stylissamide L (**1**) (700 MHz, DMSO-d_6).

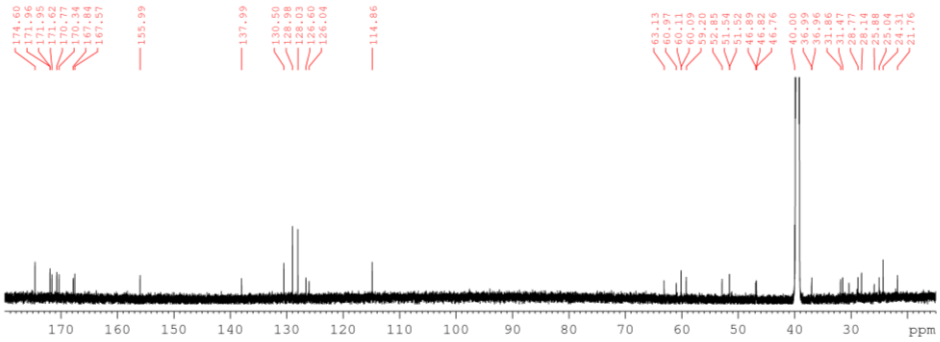


Figure 4.11. ^{13}C-NMR spectrum of stylissamide L (**1**) (175 MHz, DMSO-d_6).

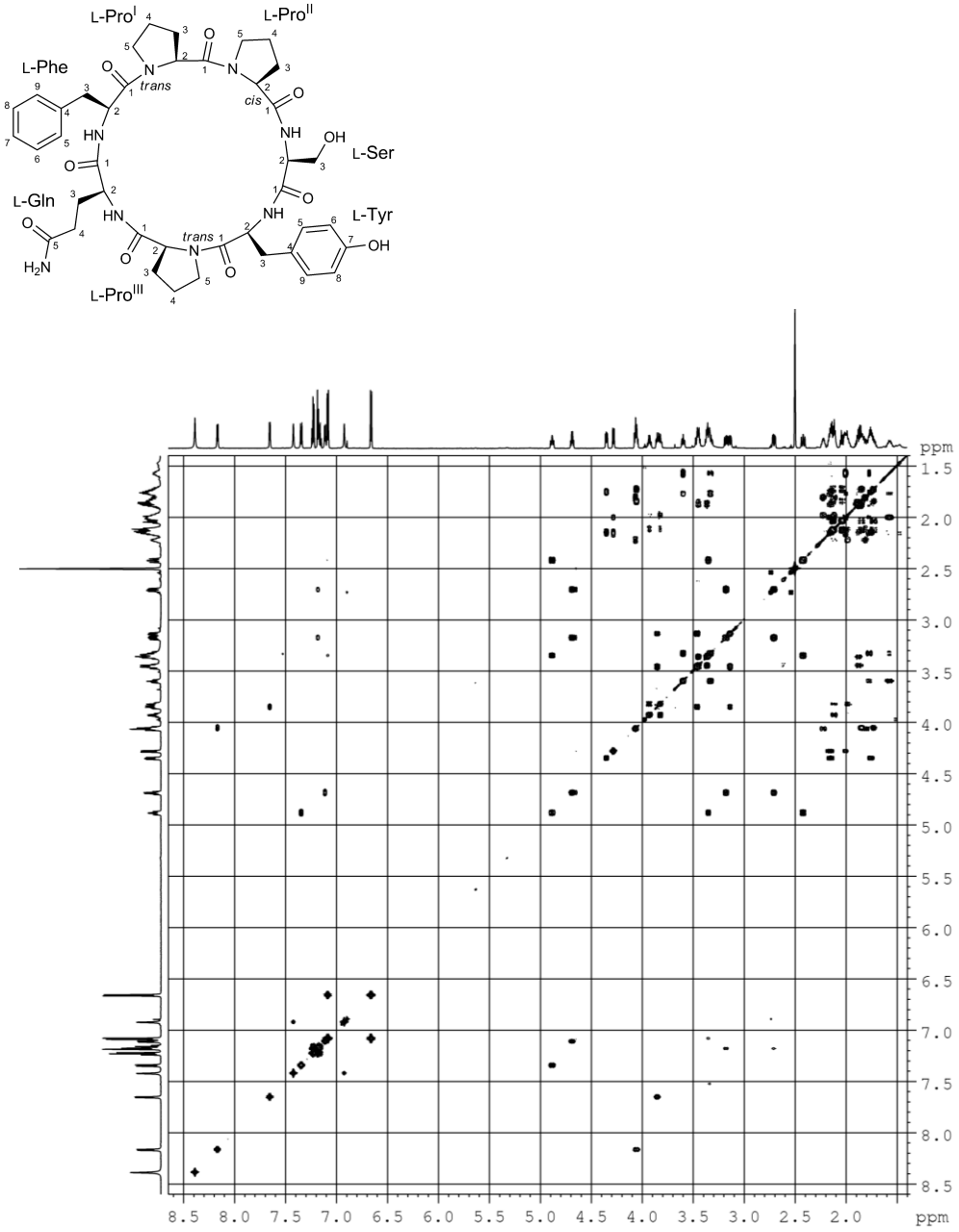


Figure 4.12. COSY spectrum of stylissamide L (**1**) (700 MHz, DMSO-d_6).

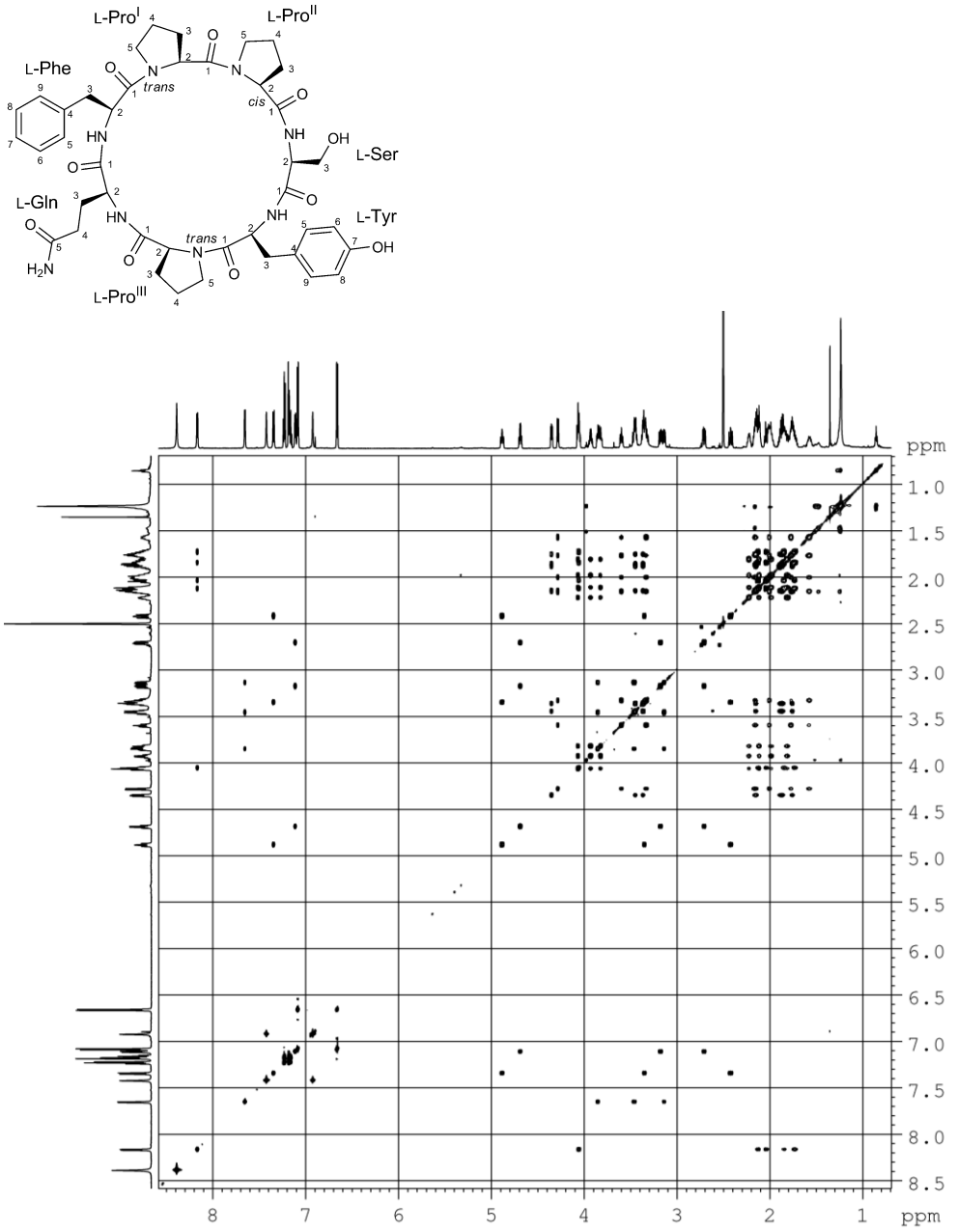
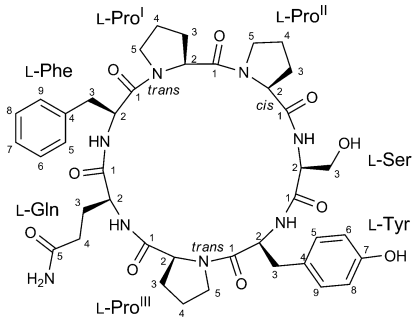


Figure 4.13. TOCSY spectrum of stylissamide L (**1**) (700 MHz, DMSO-d_6).



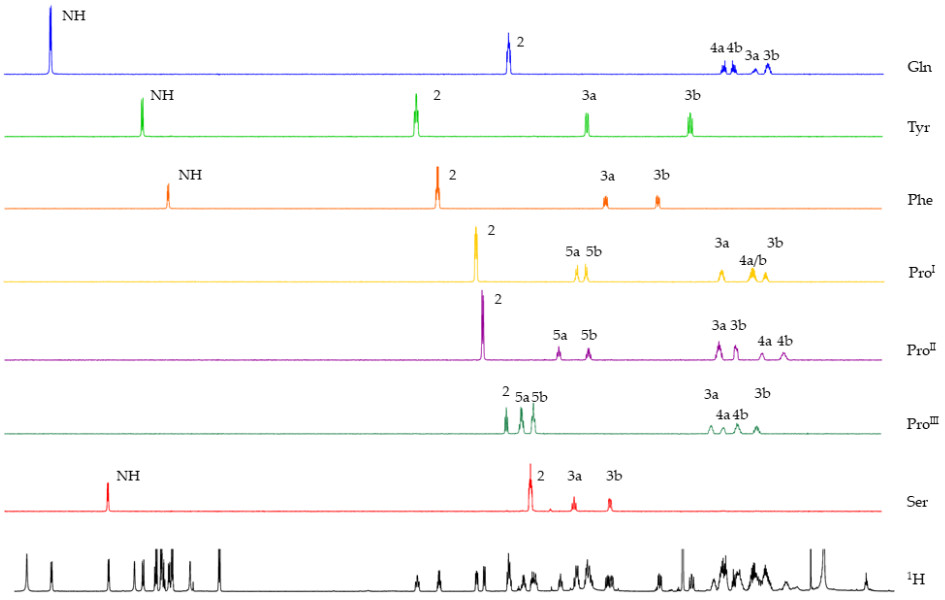
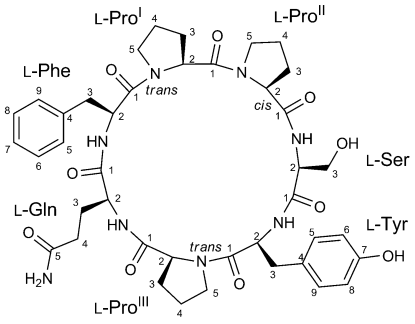


Figure 4.14. Spin systems of stylissamide L (**1**) from the sections of the TOCSY spectrum.



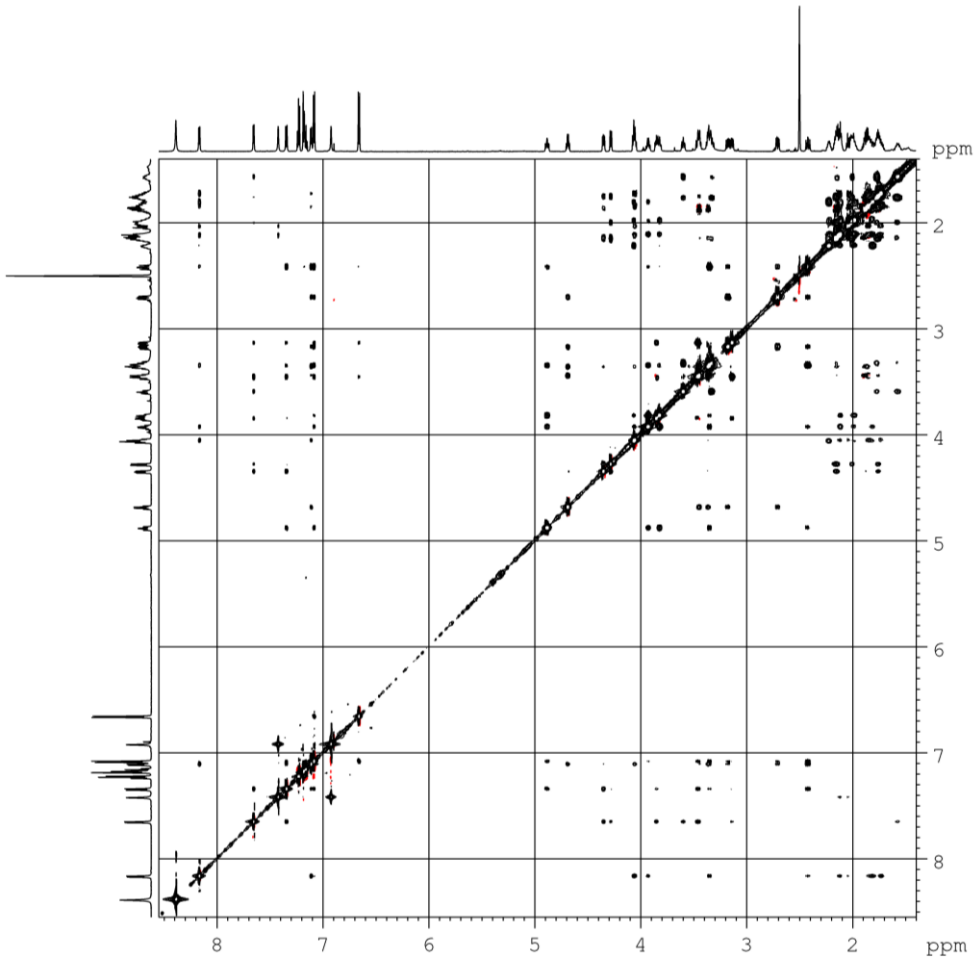


Figure 4.15. NOESY spectrum of stylissamide L (**1**) (700 MHz, DMSO-d_6).

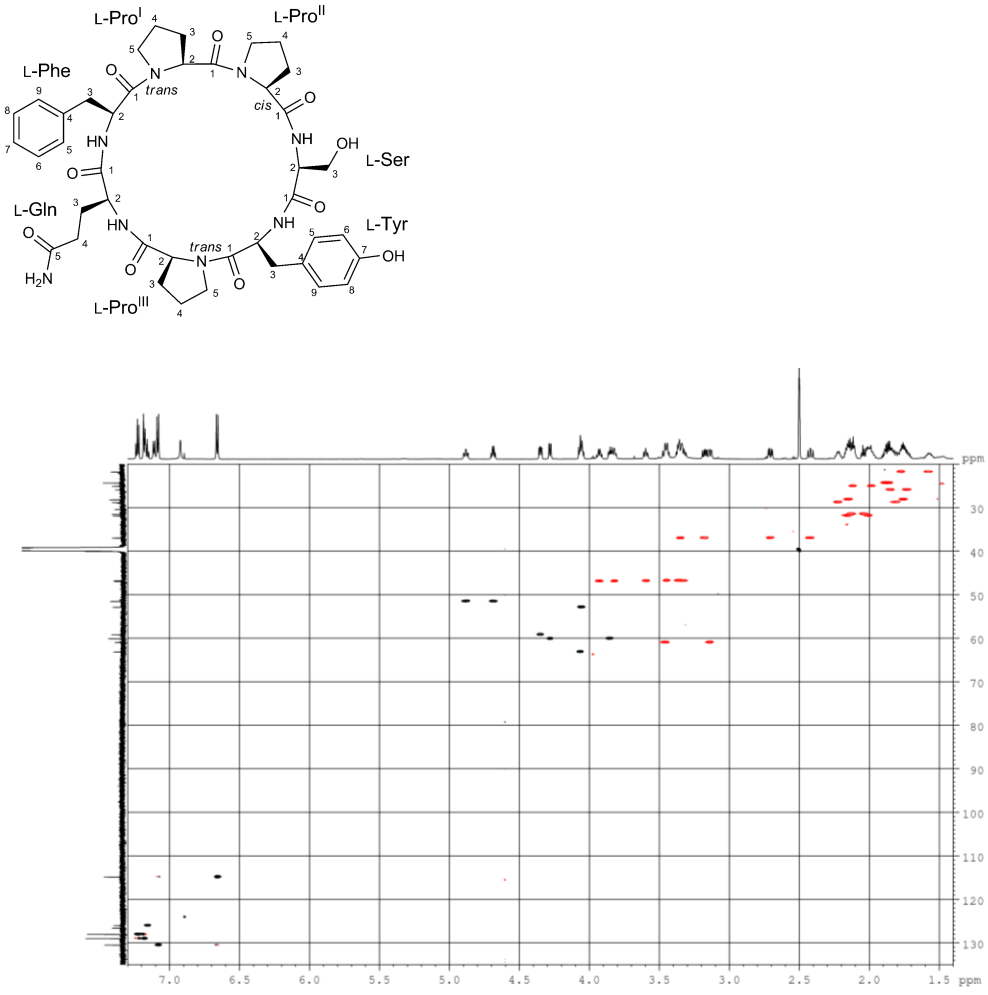


Figure 4.16. HSQC spectrum of stylissamide L (**1**) (700 MHz, DMSO-d_6).

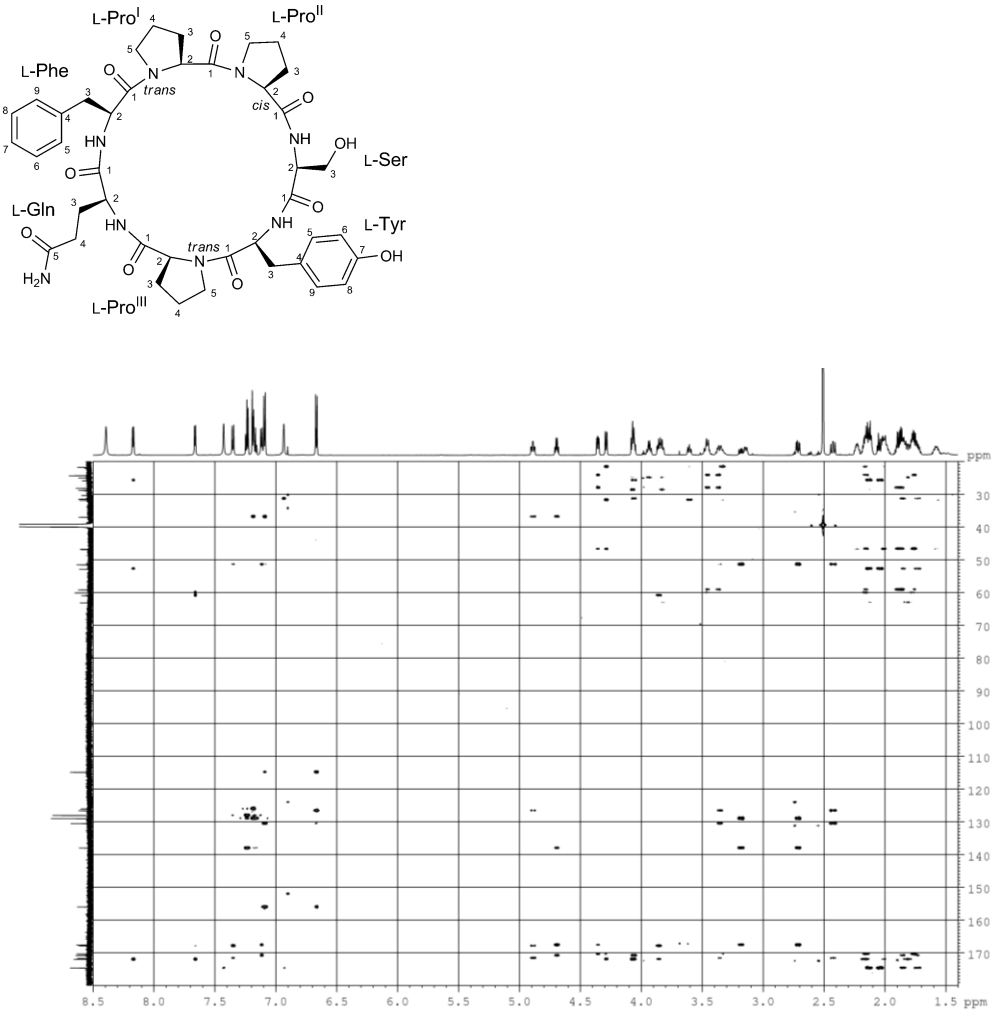
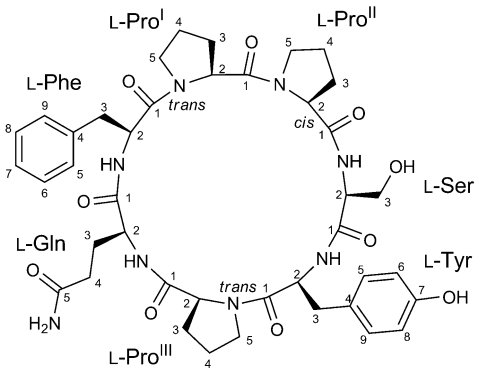


Figure 4.17. HMBC spectrum of stylissamide L (**1**) (700 MHz, DMSO-d_6).



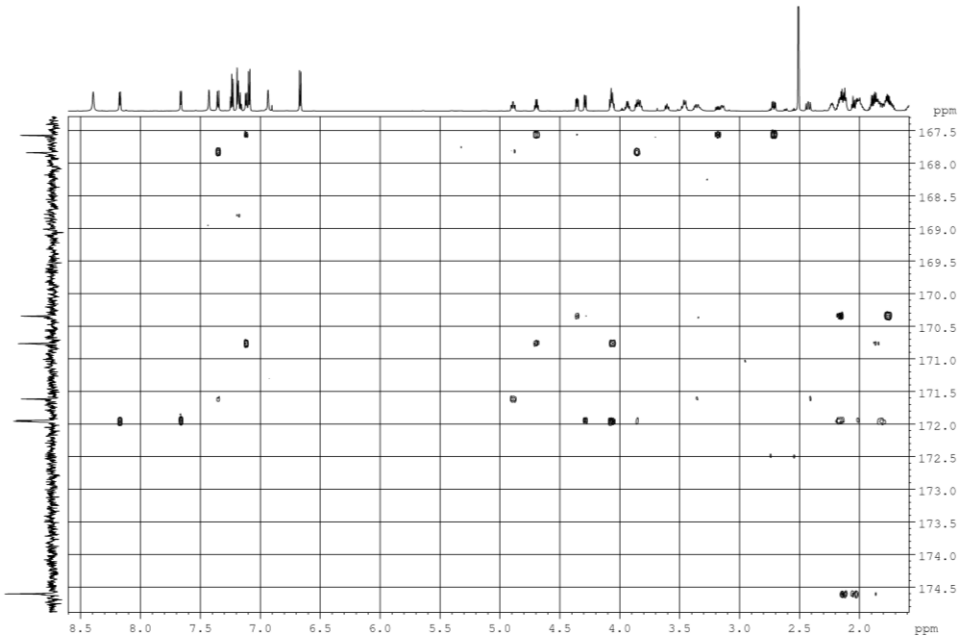


Figure 4.18. Band-selective HMBC spectrum of stylissamide L (**1**) (700 MHz, DMSO-d_6).

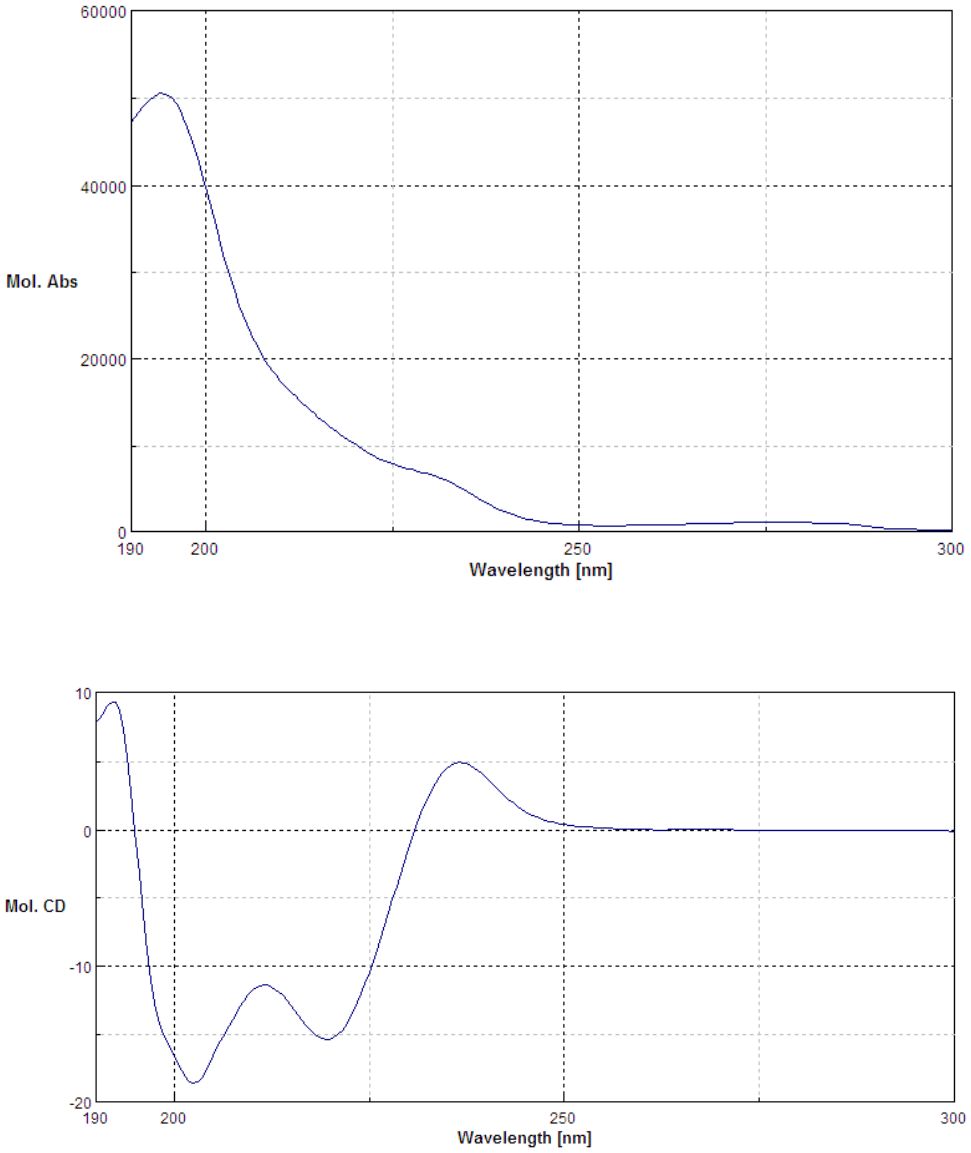


Figure 4.19. UV and ECD spectra of stylissamide L (**1**) in ACN.

References

[1] Scarpato, S.; Teta, R.; Della Sala, G.; Pawlik, J.R.; Costantino, V.; Mangoni, A. New Tricks with an Old Sponge: Feature-Based Molecular Networking Led to Fast Identification of New Stylissamide L from *Stylissa caribica*. *Mar. Drugs* **2020**, *18*, 443. https://doi.org/10.3390/md18090443

[2] Lehnert, H.; Van Soest, R. Shallow water sponges of Jamaica. *Beaufortia* **1998**, *48*, 71–103.

[3] Al Mourabit, A.; Potier, P. Sponge's Molecular Diversity Through the Ambivalent Reactivity of 2-Aminoimidazole: A Universal Chemical Pathway to the Oroidin-Based Pyrrole-Imidazole Alkaloids and Their Palau'amine Congeners. *European J. Org. Chem.* **2001**, *2001*, 237–243, doi:10.1002/1099-0690(200101)2001:2<237::aid-ejoc237>3.0.co;2-v.

[4] Grube, A.; Maier, T.; Kostrzewa, M.; Köck, M. MS-guided fractionation as a fast way to the identification of new natural products - MALDI-TOF-MS screening of the marine sponge Stylissa caribica. *Zeitschrift fur Naturforsch. - Sect. B J. Chem. Sci.* **2007**, *62*, 600–604, doi:10.1515/znb-2007-0420.

[5] Schmidt, G.; Grube, A.; Köck, M. Stylissamides A-D - New proline-containing cyclic heptapeptides from the marine sponge Stylissa caribica. *European J. Org. Chem.* **2007**, *2*, 4103–4110, doi:10.1002/ejoc.200700013.

[6] Cychon C., and Köck,M. Stylissamides E and F, Cyclic Heptapeptides from the Caribbean Sponge Stylissa caribica. *J. Nat. Prod.* **2010,** *73,* 738–742. doi: 10.1021/np900664f.

[7] Wang, X.; Morinaka, B.I.; Molinski, T.F. and H from the Bahamian Sponge Stylissa caribica. **2013**.

[8] Mohammed, R.; Peng, J.; Kelly, M.; Hamann, M.T. Cyclic heptapeptides from the Jamaican sponge Stylissa caribica. *J. Nat. Prod.* **2006**, *69*, 1739–1744, doi:10.1021/np060006n.

[9]Olivon, F.; Elie, N.; Grelier, G.; Roussi, F.; Litaudon, M.; Touboul, D. MetGem Software for the Generation of Molecular Networks Based on the t-SNE Algorithm. *Anal. Chem.* **2018**, *90*, 13900–13908, doi:10.1021/acs.analchem.8b03099.

[10] Olivon, F.; Grelier, G.; Roussi, F.; Litaudon, M.; Touboul, D. MZmine 2 Data-Preprocessing to Enhance Molecular Networking Reliability. *Anal. Chem.* **2017**, *89*, 7836–7840, doi:10.1021/acs.analchem.7b01563.

[11] Mohimani, H.; Gurevich, A.; Shlemov, A.; Mikheenko, A.; Korobeynikov, A.; Cao, L.; Shcherbin, E.; Nothias, L.F.; Dorrestein, P.C.; Pevzner, P.A. Dereplication of microbial metabolites through database search of mass spectra. *Nat. Commun.* **2018**, *9*, 1–12, doi:10.1038/s41467-018-06082-8.

[12] Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks., doi:10.1101/gr.1239303.

[13] Fujii, K.; Ikai, Y.; Mayumi, T.; Oka, H.; Suzuki, M.; Harada, K.I. A nonempirical method using LC/MS for determination of the absolute configuration of constituent amino acids in a peptide: Elucidation of limitations of Marfey's method and of its separation mechanism. *Anal. Chem.* **1997**, *69*, 3346–3352, doi:10.1021/ac9701795.

[14] Arai, M.; Yamano, Y.; Fujita, M.; Setiawan, A.; Kobayashi, M. Stylissamide X, a new proline-rich cyclic octapeptide as an inhibitor of cell migration, from an Indonesian marine sponge of Stylissa sp. *Bioorganic Med. Chem. Lett.* **2012**, *22*, 1818–1821, doi:10.1016/j.bmcl.2011.10.023.

PART 2

Exploration of the extracts of the seagrass *Zostera marina*

Chapter 6

Aquatic plants

Aquatic plants are plant organisms requiring seasonal or stagnant water for their growth and survival, and which have adapted to living submerged or floating on both saltwater and freshwater, such as on lakes, rivers, oceans, and ponds.[1] This group of plants is also defined as hydrophytes or macrophytes, distinguishing them from algae and other microphytes. Included in this group of plants are plants that grow completely below the surface of the water, plants with roots attached to the sediment, or suspended in water, but with leaves floating on the surface, or floating plants with leaves that may be submerged, or partially or completely emerged. The definition of aquatic plants usually includes vascular plants, both angiosperms and ferns, but not algae because the latter have no true conduction tissues. In addition to water, these plants need light and a source of inorganic carbon, such as carbon dioxide, to carry out chlorophyll photosynthesis, oxygen, which is essential for cellular respiration, and nutrients such as nitrogen, phosphorus, and others. Generally, plants with emerging or floating leaves have a ready source of light, carbon dioxide and oxygen and therefore, are more productive than submerged plants, since as they descend deeper into the water, light energy is progressively reduced. Carbon dioxide and oxygen must be taken up by the water and/or stored in the plant's stem. Both freshwater and marine macrophytes are essential in forming important ecosystems for other aquatic organisms, promoting the

diversity of aquatic ecosystem.[2] They supply nutrients, providing both a direct source of food and detrital material after the plant has died and a substrate for other plants and animals for many marine species.

Focusing on marine aquatic plants, it is possible to list about 60 species of marine aquatic plants divided into four families: *Posidoniaceae, Zosteracea*e, *Hydrocharitaceae* and *Cymodoceaceae*. They all belong to the order *Alismatales* in the monocotyledonous class.[3]

Marine aquatic plants are herbaceous plants with a morphology, similar to other flowering plants. Seagrasses have roots, which arise from the lower surface of the rhizomes, capable of living in anoxia, i.e., in the absence of oxygen, which are essential for taking up nutrients from the water and transporting them to various tissues.[4] The basic supply of oxygen is provided by the leaves and rhizomes. Seagrass rhizomes, which can be usually herbaceous, cylindrical, or laterally compressed, have also the fundamental task of keeping the plant anchored to the soil. On the other hand, the leaves are characterised by a thin cuticle, with a species-specific morphology, and an epidermis without stomata, which is the main photosynthetic tissue. [5]

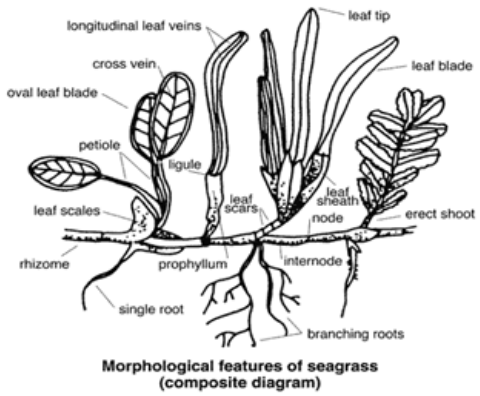


Figure 6.1. Morphology of seagrass
(Image retrieved from https://orienttaking86.blogspot.com/2012/11/karakter-sistem-vegetatif-lamun.html)

Most seagrass species are perennial and generative reproduction takes place with regular flowering. As a result of unique acquired structural adaptations for submerged marine life, seagrasses can efficiently pollinate underwater. From an evolutionary point of view, around 140 million years ago, aquatic plants were derived from the first monocotyledonous land plants. Indeed, through genome sequencing of *Zostera marina*, it was possible to confirm the adaptation of these plants to the marine environment.[6] The loss of genes involved in the biosynthesis of volatile compounds involved in ecological interactions such as ethylene and terpenoids, was verified, as well as genes for protection against ultraviolet radiation and phytochromes used for far-red detection necessary for terrestrial life. At the same time, they have been shown to have introduced into their genome the genes needed for adaptation to the salt concentration and ion homeostasis characteristic of underwater life.[7]

6.1 Zostera marina

Zosteraceae is one of the four families of aquatic plants belonging to the order *Alismatales*. The *Zosteraceae* are perennial flowering plants that complete their life cycle under water. *Zostera* species are widespread in the greatest latitudinal range of all aquatic plant genera, ranging from the tropics to the Arctic and sub-Antarctic waters. There are 22 different species classified in 3 genera: *Heterozostera* (recently included in Zostera), Phyllospadix and Zostera. The latter genus includes at least 15 different species, including *Zostera marina*. [8]

Zostera marina, like all other plants in the Zosteraceae family, is a perennial plant and is the most widespread in the northern hemisphere, living mainly in the colder ocean waters of the North Atlantic and North Pacific. It can also be found in the Arctic Ocean. It is characterised by clusters of up to 7-10 leaves. The single leaf has a rounded or truncated apex and is about 30-40 cm long, up to 70 cm long and about 3-7 mm wide.[3]

This plant can be reproduced by vegetative propagation, i.e., the new plant is formed by separating organs or parts of organs from an existing plant, which will contribute to the formation of a clone, or by sexual reproduction.

Figure 6.2. *Zostera marina*
(Image retrieved from https://oursharedseas.com/blue-solutions/)

Zostera marina is one of the most studied aquatic plants; indeed, it was one of the first aquatic plants whose genome was sequenced. *Zostera marina*, is known to be a rich source of natural compounds, serving defensive purposes in the marine environment. The abundance of natural products from these aquatic plants makes them attractive as a source of pharmaceuticals,

cosmetics, food supplements and antifouling agents. Various secondary metabolites have also been extracted from specimens collected in United States, including: vanillic acid, which has anthelmintic, antimicrobial and, according to some studies, anti-inflammatory properties; gallic acid, an antioxidant used as an anti-hemorrhagic in cases of menorrhagia; gentisic acid, antibacterial and anti-inflammatory; protocatechuic acid, a flavouring for food use and caffeic acid, an inhibitor of the enzyme xanthine oxidase, an excess of which in the tissues indicates that gout is developing.[6,9] .However, it is commonly known that the quantity and quality of the secondary metabolite profile, e.g., the concentration of phenols, is highly dependent on environmental factors such as light, temperature or nutrient availability. The phenolic composition of different species of *Zostera* have been in-depth investigated revealing the presence of zosteric acid, a sulfated phenolic acid, a group of simple phenolic acids, like caffeic acid, coumaric acid, ferulic acid, and rosmarinic acid, a bioactive dimeric phenylpropanoic acid.[10,11,12] The genus *Zostera* also contains flavonoids, which can often contain sulfate moieties distinguishing marine aquatic plants from terrestrial specimen, in which it is mostly absent.

Recently, extracts of the common seagrass *Zostera marina* showed the presence of a cyclic diarylheptanoid previously identified in the extract of *Cymodocea nodosa* (Ucria) and of the new diarylheptanoid, isotedarene A, which showed a structure closely related to tedarene A, a natural product isolated from the extracts of the marine sponge *Tedania ignis*.[13]

During the three years of the PhD program, a fruitful collaboration was established with the research group of Professor Christian Zidorn, full

professor at Kiel University, whose research interests include the search for novel bioactive secondary metabolites from plant extracts and marine organisms, and chemical ecology. The research activity focused on the analysis of extracts of the common eelgrass *Zostera marina*, with the main purpose of exploring the presence of unknown bioactive secondary metabolites.

Specifically, in the laboratories of Kiel University, sample collection, subsequent extraction and isolation of the unknown natural product were performed. In our laboratory, we pinpointed our attention on the structural elucidation, including the investigation of the intriguing conformational and configurational aspects of the new entities.

The collaboration project resulted with the publication of three research papers:

1. Grauso, L.; Li, Y.; Scarpato, S.; Shulha, O.; Rárová, L.; Strnad, M.; Teta, R.; Mangoni, A.; Zidorn, C. Structure and Conformation of Zosteraphenols, Tetracyclic Diarylheptanoids from the Seagrass Zostera marina: An NMR and DFT Study. *Org. Lett.* **2020**, *22*, 78–82, doi:10.1021/acs.orglett.9b03964.
2. Li, Y.; Grauso, L.; Scarpato, S.; Cacciola, N.A.; Borrelli, F.; Zidorn, C.; Mangoni, A. Stable Catechol Keto Tautomers in Cytotoxic Heterodimeric Cyclic Diarylheptanoids from the Seagrass Zostera marina. *Org. Lett.* **2021**, *23*, 7134–7138, doi:10.1021/acs.orglett.1c02537.
3. Li, Y.; Rárová, L.; Scarpato, S.; Çiçek, S.S.; Jordheim, M.; Štenclová, T.; Strnad, M.; Mangoni, A.; Zidorn, C. Seasonal

variation of phenolic compounds in Zostera marina (Zosteraceae) from the Baltic Sea. *Phytochemistry,**2022**, 196,113099,* https://doi.org/10.1016/j.phytochem.2022.113099

References

[1] B, S. Economic Benefits of Aquatic Plants - A Review. *Mod. Approaches Oceanogr. Petrochemical Sci.* **2018**, *1*, 48–49, doi:10.32474/maops.2018.01.000115.

[2] Bakker, E.S.; Wood, K.A.; Pagès, J.F.; Veen, G.F. (Ciska.; Christianen, M.J.A.; Santamaría, L.; Nolet, B.A.; Hilt, S. Herbivory on freshwater and marine macrophytes: A review and perspective. *Aquat. Bot.* **2016**, *135*, 18–36, doi:10.1016/j.aquabot.2016.04.008.

[3] Borum, J.; Duarte, C.; Krause-Jensen, D.; Greve, T.M. *European seagrasses : an introduction to monitoring and management*; 2004; ISBN 8789143213

[4] Larkum, A.W.D.; Orth R, J.; Duarte, C.M. *Seagrasses:Biology,Ecology and Conservation*; 2006; ISBN 9781402029424.

[5] Larkum, A.W.D.; Kendrick, G.A.; Editors, P.J.R. *Seagrasses of Australia*; 2018; ISBN 9783319713526.

[6] Zidorn, C. Secondary metabolites of seagrasses (Alismatales and Potamogetonales; Alismatidae): Chemical diversity, bioactivity, and ecological function. *Phytochemistry* **2016**, *124*, 5–28, doi:10.1016/j.phytochem.2016.02.004

[7] Olsen, J.L.; Rouzé, P.; Verhelst, B.; Lin, Y.C.; Bayer, T.; Collen, J.; Dattolo, E.; De Paoli, E.; Dittami, S.; Maumus, F.; et al. The genome of the seagrass Zostera marina reveals angiosperm adaptation to the sea. *Nature* **2016**, *530*, 331–335, doi:10.1038/nature16548.

[8] Chase, M.W.; Christenhusz, M.J.M.; Fay, M.F.; Byng, J.W.; Judd, W.S.; Soltis, D.E.; Mabberley, D.J.; Sennikov, A.N.; Soltis, P.S.; Stevens, P.F.; et al. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* **2016**, *181*, 1–20, doi:10.1111/boj.12385.

[9] Dybsland, C.S.; Bekkby, T.; Enerstvedt, K.H.; Kvalheim, O.M.; Rinde, E.; Jordheim, M. Variation in phenolic chemistry in zostera marina seagrass along environmental gradients. *Plants* **2021**, *10*, 1–18, doi:10.3390/plants10020334.

[10] Dybsland, C.S.; Bekkby, T.; Enerstvedt, K.H.; Kvalheim, O.M.; Rinde, E.; Jordheim, M. Variation in phenolic chemistry in zostera marina seagrass along environmental gradients. *Plants* **2021**, *10*, 1–18, doi:10.3390/plants10020334.

[11] Grignon-Dubois, M.; Rezzonico, B. First phytochemical evidence of chemotypes for the seagrass zostera noltii. *Plants* **2012**, *1*, 27–38, doi:10.3390/plants1010027.

[12] Papazian, S.; Parrot, D.; Burýšková, B.; Weinberger, F.; Tasdemir, D. Surface chemical defence of the eelgrass Zostera marina against microbial foulers. *Sci. Rep.* **2019**, *9*, 1–12, doi:10.1038/s41598-019-39212-3.

[13] Li, Y.; Mangoni, A.; Shulha, O.; Çiçek, S.S.; Zidorn, C. Cyclic diarylheptanoids deoxycymodienol and isotedarene A from Zostera marina (Zosteraceae). *Tetrahedron Lett.* **2019**, *60*, 150930, doi:10.1016/j.tetlet.2019.07.021.

Chapter 7

An in-depth study of the structural and conformational features of zosteraphenols, tetracyclic diarylheptanoids from *Zostera marina*

7.1 Synopsis of this chapter [1]

Extracts of the marine grass *Zostera marina*, collected from the coast of Kiel, showed the presence of two unknown compounds, named zosteraphenols A (**3**) and B (**4**). The complete structure of zosteraphenols A (**3**) and B (**4**), was fully elucidated revealing a tetracyclic diarylheptanoid structure similar to tedarene B, previously identified from the marine sponge *Tedania igni*s.[2]

Diarylheptanoids represent a class of natural products characterized by two benzene rings, which can usually bear one or more hydroxyl groups, joined by a functionalized seven-carbon chain. Diarylheptanoid compounds can easily be found in plants, the most notable example being curcumin.[3] A subclass of diarylheptanoids comprises cyclic structures. In particular, in the biphenyl type the aromatic rings are directly connected to each other, as opposed to the diphenyl ether type, where the two benzene rings are connected through an oxygen atom. Due to their inherent steric tension, cyclic diarylheptanoids are often found to possess axial and/or planar chirality.[4,5] Indeed, the high energy barrier between conformers, which

exhibit restricted rotation of one or more single bonds, can lead to separable stereoisomers, which are termed atropisomers. For example, tedarene B contains axial, planar and point chirality elements. When the energy barrier between the atropisomers is lower, axially or planarly chiral atropisomers may be in slow equilibrium, this causing coalescent NMR signals.[2] The proton NMR spectrum of zosteraphenol A (**3**) and B (**4**), showing some coalescent signals, agreed with slow equilibrium of both compounds with a minor rotamer with opposite axial chirality. Both 1D and 2D NMR spectra were then recorded at -15° C to obtain sharp NMR signals and determine the complete structure of compound **3** and **4**. The broadness of signals of the elusive minor rotamer, still present at -15°C, prevented the detection of any coupling constants for this rotamer, so ^{1}H chemical shifts were the only NMR parameters available for it. The good agreement between experimental and quantum mechanically calculated NMR chemical shifts of both rotamers, supported the hypothesized equilibrium between conformers. The anti-inflammatory activity of zosteraphenols **3** and **4** was evaluated in a CD62E induction assay (E-selectin, ELAM), but no bioactivity was demonstrated.

*7.2 Isolation of zosteraphenol A (**3**) and B (**4**)*

A quantity of *Zostera marina* (1.25 kg) was collected off the coast of Kiel, Schleswig-Holstein, Germany, during March 2018. Once ground, the plants were extracted first with *n*-hexane and then with acetone at a temperature of 25 °C. The crude *n*-hexane extract was fractionated thanks to two chromatographic steps: through a silica gel column at medium pressure and

thanks to a Sephadex LH-20. Zosteraphenol A (7.7 mg) (**3**) and B (10.5 mg) (**4**) were obtained as pure compounds through a semipreparative RP-HPLC. (Figure 7.1).

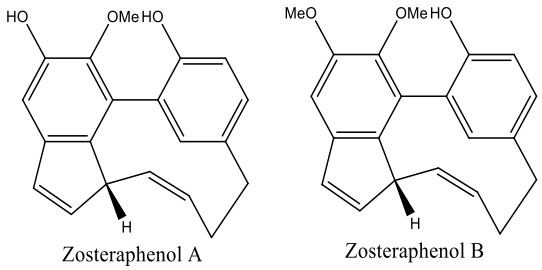


Figure 7.1. Structure of zosteraphenol A (**3**) and B (**4**)

*7.3 Structural elucidation of zosteraphenol A (**3**) and B (**4**)*

The high-resolution ESI mass spectrum of compound **3** showed a pseudomolecular $[M + H]^+$ ion peak at *m/z* 307.1332. The molecular formula $C_{20}H_{18}O_3$, was in good agreement with these data and 12 was the number of unsaturations derived from it.

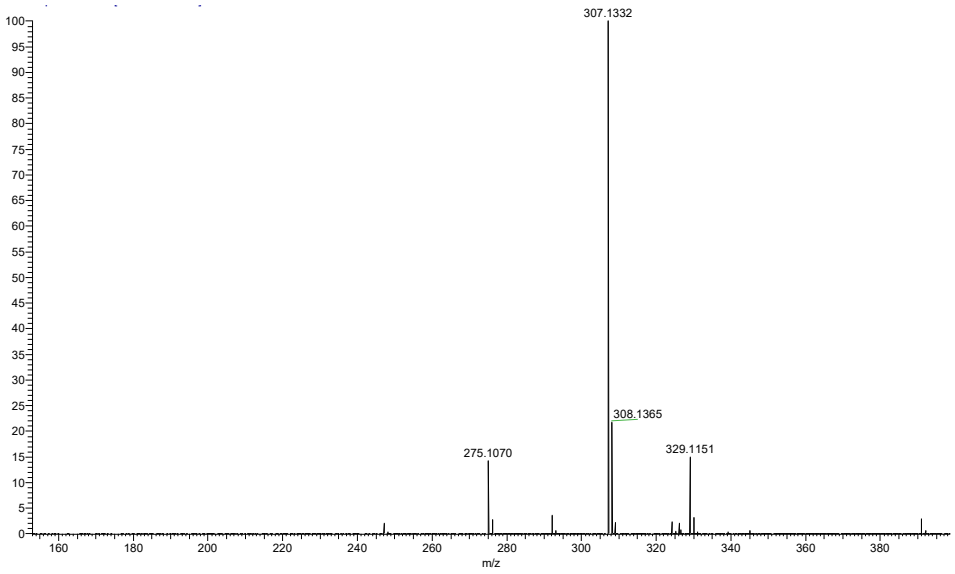


Figure 7.2. High-resolution ESI-MS spectrum of zosteraphenol A (**3**) (positive ion mode).

The proton NMR spectrum of zosteraphenol A (**3**), showing some broad, unresolved signals, suggested the presence of coalescent signals. In addition, the number of signals shown in the ^{13}C NMR spectrum, was lower than those predicted in the molecular formula, suggesting coalescent signals to be present also in this spectrum. The 1D NMR spectra were, therefore, recorded at -15°C (258 K) so that any conformational equilibrium was slowed down, resulting in sharper signals in the ^{1}H NMR spectrum and 20 signals for the 20 carbons present in the molecular formula in the ^{13}C spectrum.

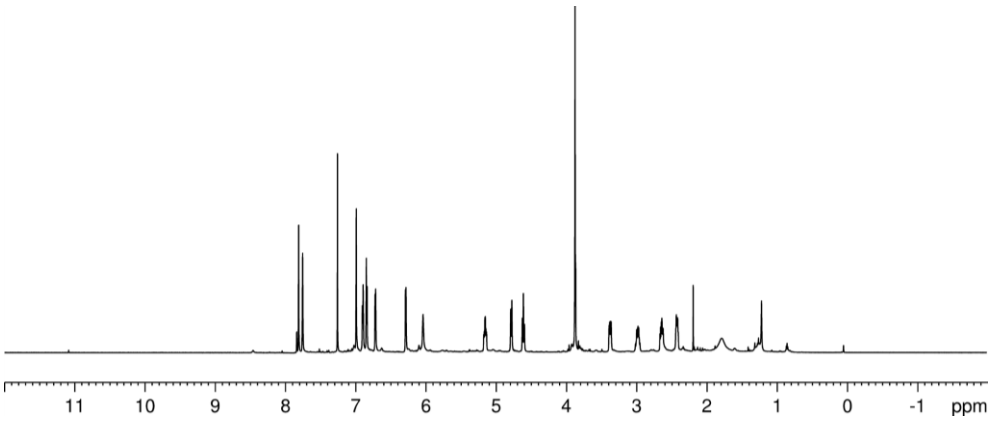


Figure 7.3. Full ^{1}H NMR spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol A (**3**).

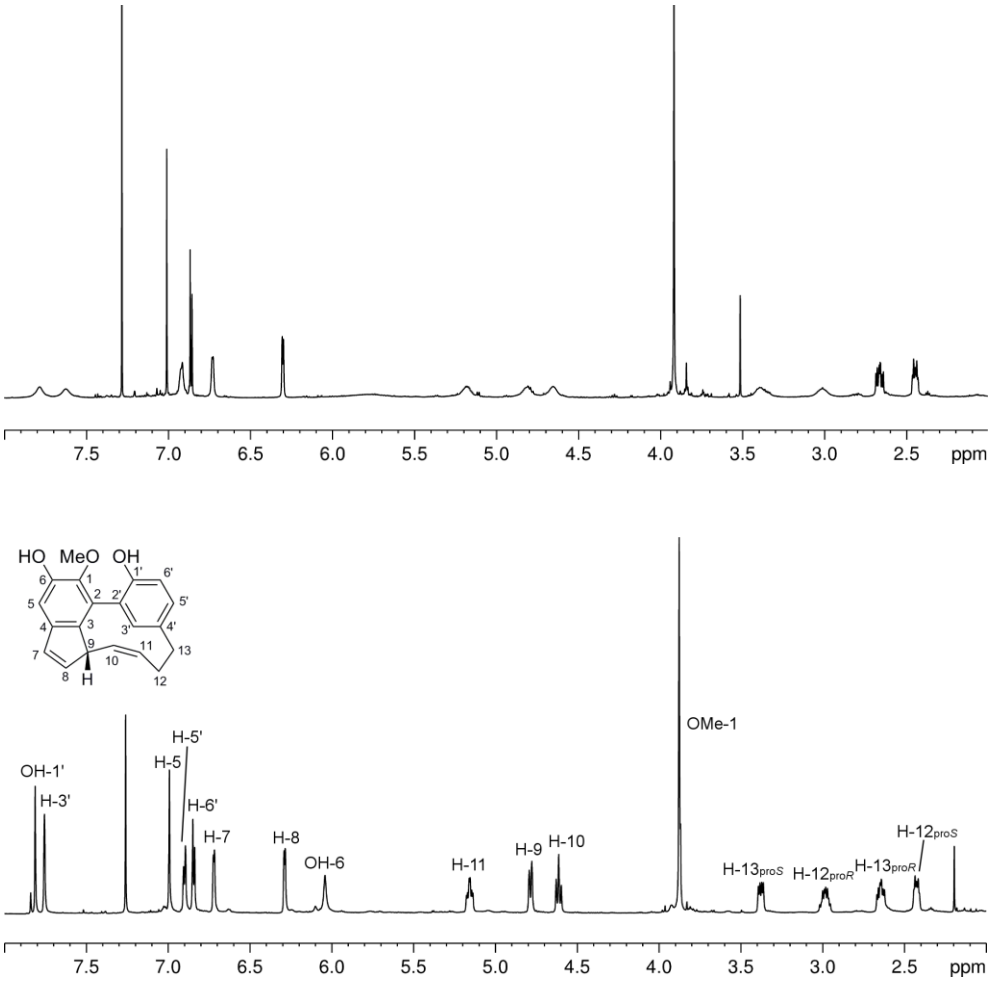


Figure 7.4. Expansion of ^{1}H NMR spectra (700 MHz, $CDCl_3$) of zosteraphenol A (**3**) at 298 K (top spectrum) and 258 K (bottom spectrum).

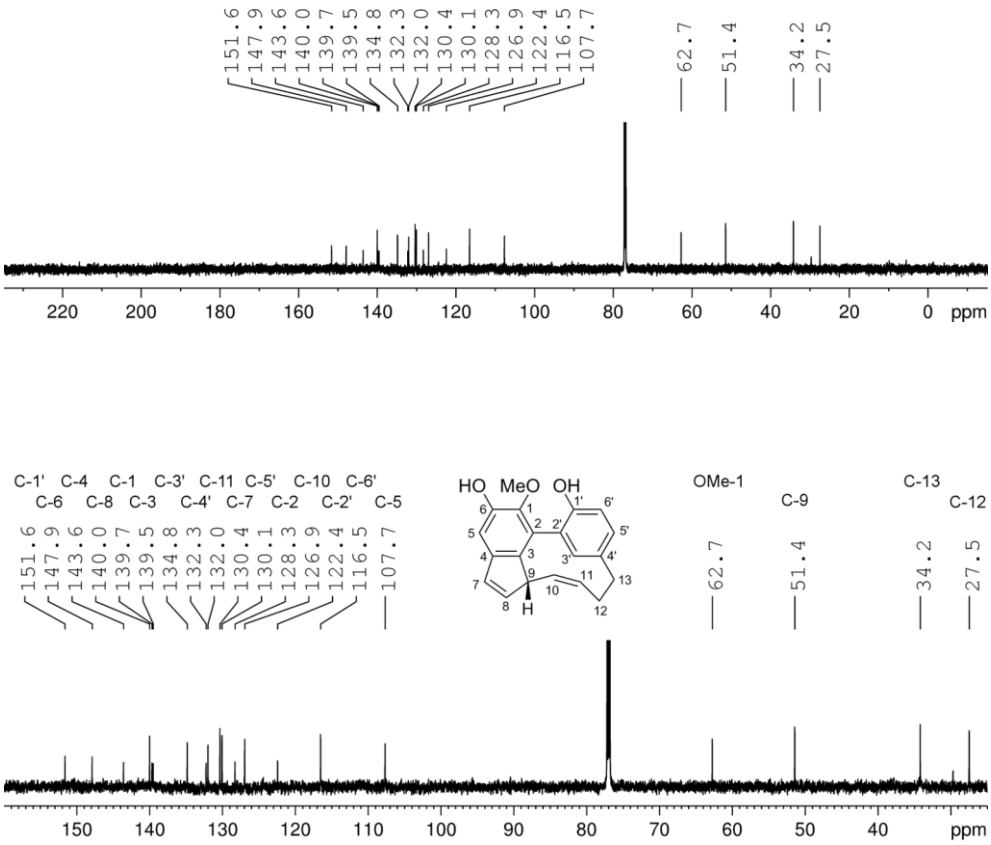


Figure 7.5. Full (top) and expanded (bottom) ^{13}C NMR Spectrum (175 MHz, 258 K, $CDCl_3$) of zosteraphenol A (**3**).

Accordingly, 2D NMR spectra were also performed at -15°C to perform structure elucidation of compound **3**. First, interpretation of the COSY spectrum allowed the identification of two double bonds and no oxygenated functional group in a C_7 chain. The value of the coupling constants helped to assign the configuration of the double bond at position 10 and of the double bond between H-7 and H-8. In detail, a coupling constant of 11.2 Hz, suggested a *Z* alkene, while the unusually small value of the coupling constant between H-7 and H-8 was consistent with a double bond in a five-membered ring, whose presence was confirmed by HMBC correlations of H-8, H-9, and H-10 with C-3 (Figure 7.6).

Analysis of NMR spectra suggested a diarylheptanoid structure, due to the presence of signals representing a 1,2,4-trisubstituted and a pentasubstituted benzene ring, which was in agreement with a diarylheptanoid skeleton of 19 carbon atoms and an additional methoxyl group (δ_H 3.88, δ_C 62.7). In addition to the aromatic signals, the ^{13}C spectrum of compound **3** showed signals for the C_7 chain, i.e. two additional double bonds (δ_C 140.0, 130.4, 132.0, and 126.9) and one methine and two methylene groups. The structure must be tetracyclic, because the π-bonds accounted for only eight of the twelve unsaturations defined by the molecular formula, and this meant that the two aromatic rings must be linked to each other. Full ^{13}C NMR assignment was achieved by HMBC correlation, which also showed a few four-bond correlations. More specifically, the connection between the two rings was elucidated by the H-3' correlation peaks with C-2, and the position of the methoxy group by the correlation of methoxy protons with C-1 (Figure 7.6).

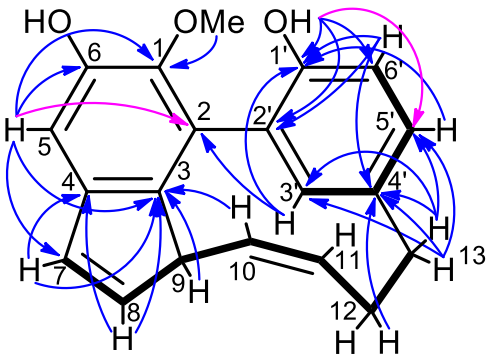
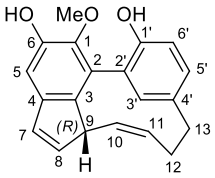


Figure. 7.6. Key connections determined from 2D NMR spectra for zosteraphenol A (**3**). Bold bonds show connections determined from COSY, blue arrows show 2- and 3-bond HMBC correlations, and pink arrows show 4-bond HMBC correlations.

Table 7.1. ^{1}H and ^{13}C NMR Data of zosteraphenol A (**3**) (700 MHz, 258 K, $CDCl_3$)



			Major rotamer	Minor rotamer
Pos.	δ_C		δ_H, mult (*J* in Hz)	δ_H
1	139.7 (C)		-	-
1-OMe	62.7 (CH_3)		3.88, s	3.97
2	128.3 (C)		-	-
3	139.5 (C)		-	-
4	143.6 (C)		-	-
5	107.7 (CH)		6.99, s	6.97
6	147.9 (C)		-	-
6-OH		-	6.04, s	6.04
7	130.4 (CH)		6.72, dd (4.7, 1.7)	6.63
8	140.0 (CH)		6.29, dd (4.7, 1.5)	6.28
9	51.4 (CH)		4.79, br. d (11.2)	3.57
10	126.9 (CH)		4.61, t (11.2)	5.04
11	132.0 (CH)		5.16, ddd (11.8, 11.2, 4.8)	5.77
12	27.5 (CH_2)	pro*R*	2.99, dddd (13.6, 12.6, 11.8, 7.0)	2.15
		pro*S*	2.43 (ddd, 13.6, 6.0, 4.8)	2.43
13	34.2 (CH_2)	pro*R*	2.65, dd (14.4, 12.6, 6.0)	2.65
		pro*S*	3.38, dd (14.4, 7.0)	2.63
1'	151.6 (C)		-	-
1'-OH	-		7.81, s	7.00
2'	122.4 (C)		-	-
3'	134.8 (CH)		7.76, br. s	8.46
4'	132.3 (C)		-	-
5'	130.1 (CH)		6.90, br. d (8.0)	7.02
6'	116.5 (CH)		6.85, d (8.0)	6.85

The ESI-HR-MS spectrum of zosteraphenol B (**4**) revealed a molecular ion peak $[M+NH_4]^+$ at *m/z* 321.1488, a very close value to the exact mass 321.1485 defining the molecular formula of $C_{21}H_{20}O_3$, showing an additional CH_2 compared to zosteraphenol A (**3**).

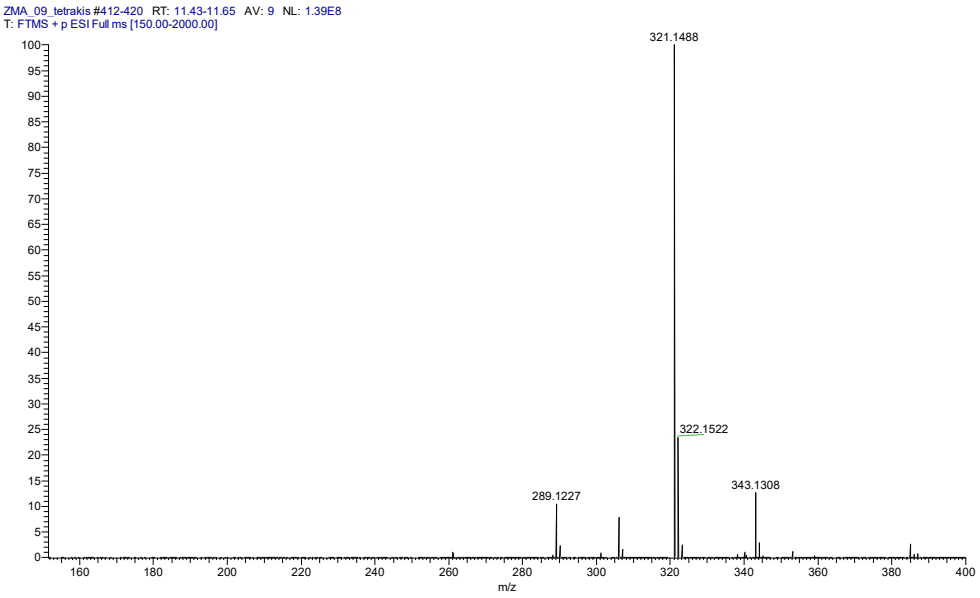


Figure 7.7. High-resolution ESI-MS spectrum of zosteraphenol B (**4**) (positive ion mode).

^{1}H spectrum analysis suggested zosteraphenol B (**4**) as an *O*-methylated derivative of zosteraphenol A (**3**), showing the presence of an additional methyl singlet at δ 3.97 and the absence of the OH singlet at δ 6.04. A complete set of 1D and 2D NMR spectra of compound **4** was recorded, fully supporting the hypothesized structure. Due to the HMBC correlation between C-6 and the methyl protons at δ 3.97, the methyl was located at O-6.

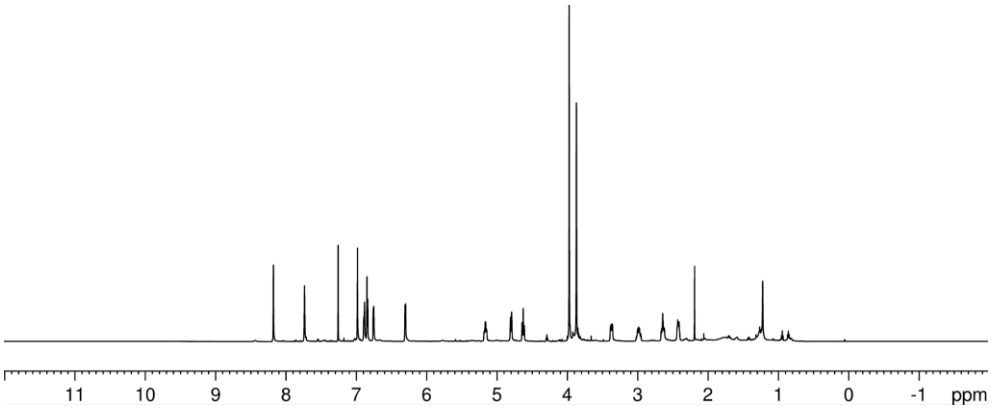


Figure 7.8. Full ^{1}H NMR spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol B (**4**).

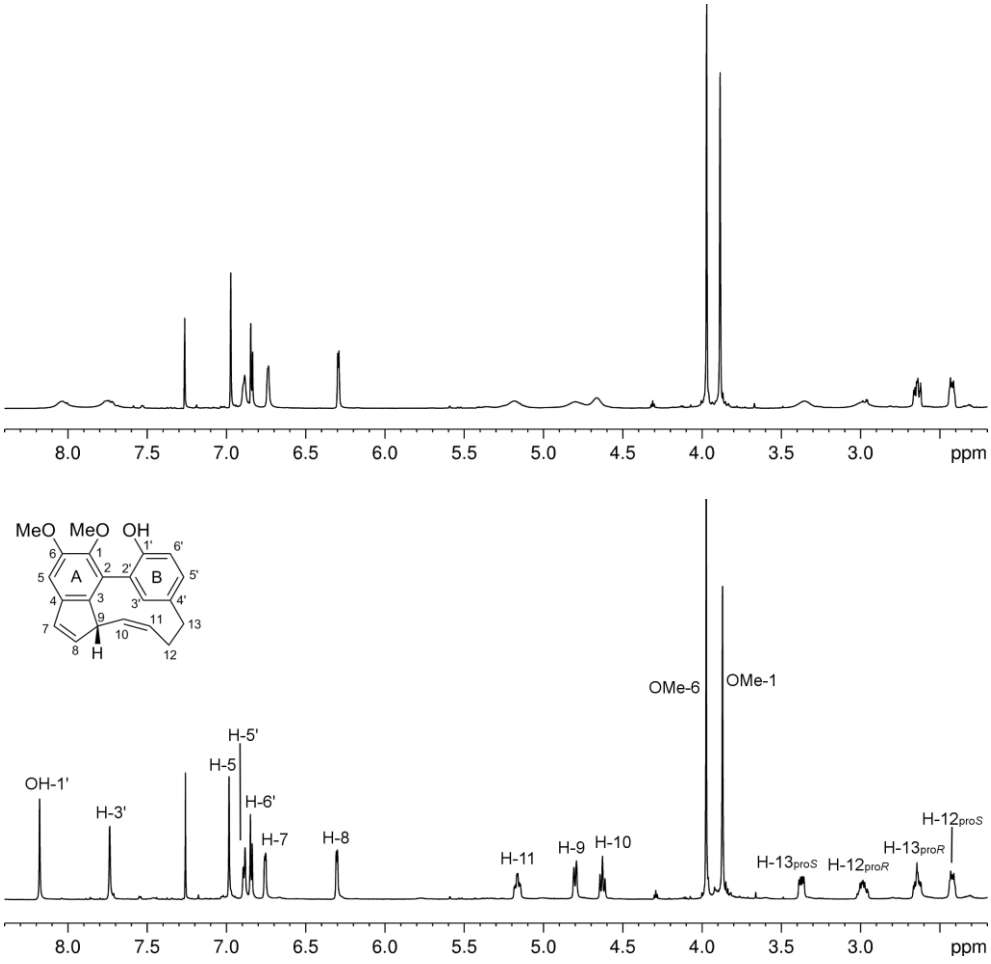


Figure 7.9. Expansion of ^{1}H NMR spectra (700 MHz, $CDCl_3$) of zosteraphenol B (**4**) at 298 K (top spectrum) and 258 K (bottom spectrum).

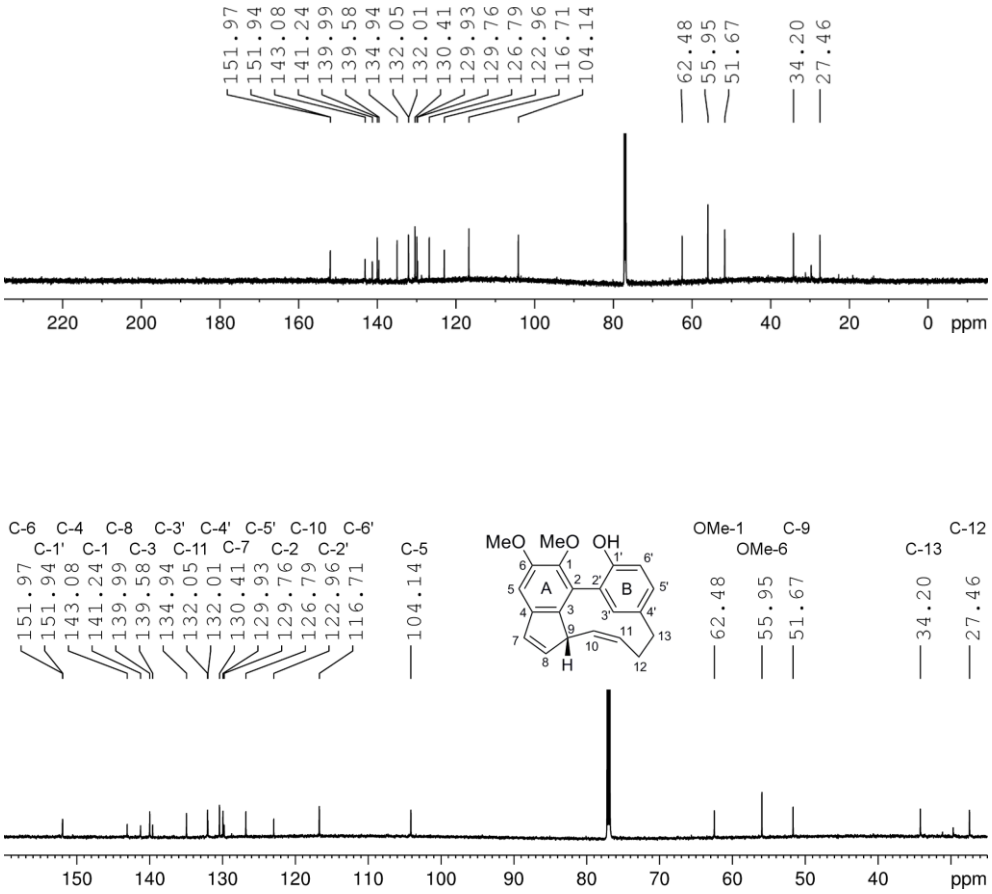
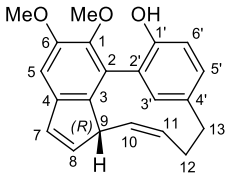


Figure 7.10. Full (top) and expanded (bottom) ^{13}C NMR Spectrum (175 MHz, 258 K, $CDCl_3$) of zosteraphenol B (**4**).

Table 7.2 ^{1}H and ^{13}C NMR Data of zosteraphenol B (**4**) (700 MHz, 258 K, $CDCl_3$)



			Major rotamer	Minor rotamer
Pos.	δ_C		δ_H, mult (*J* in Hz)	δ_H
1	141.2 (C)		-	-
1-OMe	62.5 (CH_3)		3.87, s	3.92
2	129.8 (C)		-	-
3	139.6 (C)		-	-
4	143.1 (C)		-	-
5	104.1 (CH)		6.98, s	6.98
6	152.0 (C)		-	-
6-OMe	56.0 (CH_3)		3.97, s	3.97
7	130.4 (CH)		6.75, dd (4.7, 1.7)	6.67
8	140.0 (CH)		6.30, dd (4.7, 1.5)	6.28
9	51.7 (CH)		4.80, br. d (11.2)	3.60
10	126.8 (CH)		4.63, t (11.2)	5.01
11	132.1 (CH)		5.16, ddd (11.8, 11.2, 4.8)	5.78
12	27.5 (CH_2)	pro*R*	2.99, dddd (13.6, 12.6, 11.8, 7.0)	2.15
		pro*S*	2.43 (ddd, 13.6, 6.0, 4.8)	2.43
13	34.2 (CH_2)	pro*R*	2.65, dd (14.4, 12.6, 6.0)	2.61
		pro*S*	3.38, dd (14.4, 7.0)	2.63
1'	151.9 (C)		-	-
1'-OH	-		8.18, s	7.46
2'	123.0 (C)		-	-
3'	134.9 (CH)		7.74, br. s	8.44
4'	132.0 (C)		-	-
5'	129.9 (CH)		6.89, br. d (8.0)	7.01
6'	116.7 (CH)		6.85, d (8.0)	6.85

*7.4 Conformational study of zosteraphenol A (**3**) and B (**4**)*

The explanation for the coalescence of NMR signals was not clear at first glance. The stereochemistry, i.e., three-dimensional structure of a compound, including configuration and conformation, is a fundamental aspect, which influences chemical, spectroscopic, and biological properties. However, the distinction between configuration and conformation is very subtle. In fact, the high energy barrier between conformers, due to restricted rotation of one or more single bonds, can lead to the clear separation of stereoisomers, which in this case are called atropisomers.

An example of a natural product exhibiting slow conformational changes or atropisomerism is tedarene B, a tetracyclic diarylheptanoid isolated from the marine sponge *Tedania ignis*. This compound shows central, planar, and axial chirality: a stereogenic carbon, a chiral axis involving the biphenyl system, fixed in the S_a configuration, and a chiral plane of the alkene, showing slow conformational exchange between the R_p and S_p configurations.

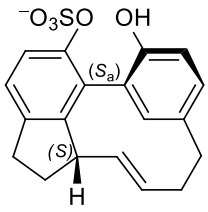


Figure 7.11. Chemical structures of tedarene B from *T. ignis.*

In zosteraphenol A (**3**), because of the impossibility of flipping the *cis* double bond at C-10, the hypothesis was that the conformational equilibrium involved the axial chirality of the biphenyl system.

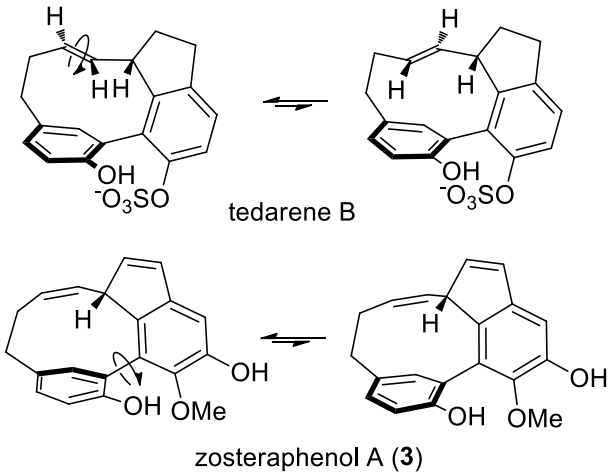


Figure 7.12. The rotameric equilibria of tedarene B and zosteraphenol A (**3**).

Although, the two rotamers were expected to be diastereomeric due to the presence of a chiral C-9 carbon in the compound, in NMR spectra recorded at -15 °C, only one set of signals was evident. Nevertheless, close examination of the NOESY spectrum, which is referred to as EXSY in these cases,[6] dispelled any perplexity. The NOESY/EXSY spectrum of zosteraphenol A (**3**) performed at -15°C, generating correlation peaks between exchanged protons, revealed many exchange peaks correlating ^{1}H signals of the major with ^{1}H signals of a minor rotamer. The exchange peaks can be recognized in that they have opposite sign compared to NOE correlation peaks. It could be seen that some protons appeared significantly shielded in the minor rotameter (H-9, proR-H-12, and proS-H-13), while others were strongly deshielded (H-10, H-11, and H-3′) (Figure 7.14)

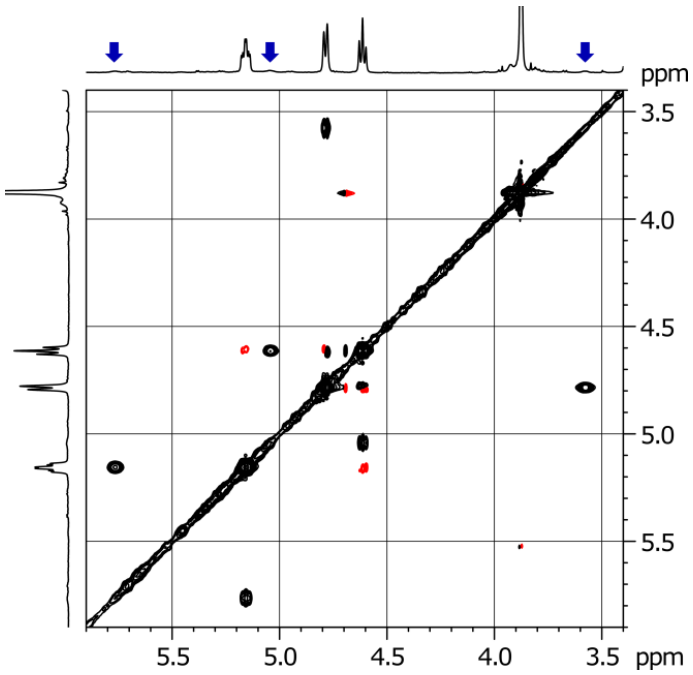


Figure 7.13. Illustrative portion of the NOESY/EXSY spectrum of compound **3** recorded at −15 °C, showing the exchange peaks (in black) between signals of the major rotamer and signals of the minor rotamer (indicated by the blue arrows). Red cross peaks have opposite sign and are NOE cross peaks.

The reason for this effect may be the anisotropic effects of the trisubstituted catechol ring. Integration of the proton signals aided in estimating the population of the minor rotamer as 6% of the total. As a consequence of the shorter lifetime of the minor conformer, its ^{1}H NMR signals appeared broad and not resolved even at -20 °C, the lowest temperature available to the NMR spectrometer in use; therefore, only ^{1}H chemical shift could be determined, while neither ^{1}H coupling constants nor ^{13}C chemical shifts could be measured. Equal signal coalescence was observed for zosteraphenol B (**4**), suggesting a similar conformational equilibrium. Indeed, the ^{1}H chemical shifts of the minor rotamer of zosteraphenol B (**4**) were also assigned using the EXSY spectrum and were very similar to those of the minor rotamer of zosteraphenol A (**3**).

7.5 Molecular modeling and Quantum mechanical calculations (DFT)

The nature of the conformational equilibrium of the zosteraphenol A (**3**) was confirmed by molecular modeling and quantum mechanical calculations (DFT).

In brief, the two possible rotamers (**3a** and **3b**), differing in the chirality of the biphenyl system, were subjected to a series of high-temperature molecular dynamics (MD) simulation in the CFF91 force field at increasing temperatures. In line with the slow conformational equilibrium, up to 1500 K both rotamers showed a conformational rigidity in the 10 ns duration of the simulation. Only in an MD simulation performed at 2000K, inversion of the axial chirality was observed.

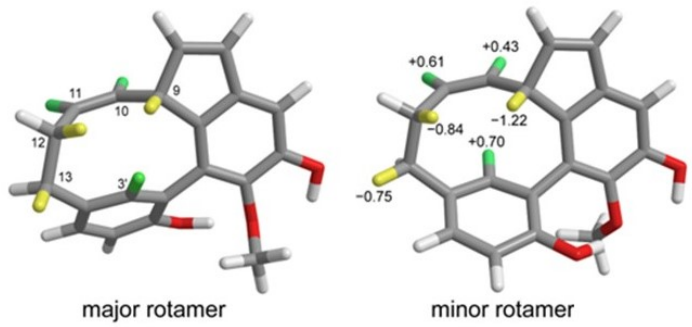


Figure 7.14. Most prominent ^{1}H chemical shift differences between the minor and major rotamers of zosteraphenol A (**3b** and **3a**, respectively). Protons that are deshielded (shielded) in the minor rotamer are marked in green (yellow).

One of the protocols and scaling factors suggested by Tantillo et al.[7] was used to perform the prediction of NMR chemical shift, to be compared with the experimental ones. After the quantum mechanical optimization of the selected rotamers with the Gaussian 16 suite at the B3LYP/6-31+G(d,p) level, chemical shifts were calculated at the PBE0/6-311+G(2d,p) level, using the PCM model for the chloroform, used as solvent of NMR spectra.[8] DFT calculation results were fully in line with the hypothesized conformational equilibrium for zosteraphenol A (**3**). Results of the DFT calculation revealed an excellent match between the predicted and experimental ^{1}H and ^{13}C chemical shifts for the major rotamer **3a**, with higher accuracy than that expected for the level of theory used (RMSD 1.67 ppm for ^{13}C NMR and 0.09 ppm for ^{1}H NMR, compared to expected RMSD 2.45 and 0.15 ppm, respectively). In addition, the ^{1}H chemical shifts predicted for minor rotamer **3b**, were also in good agreement with the experimental values assigned through the EXSY spectrum (RMSD 0.10 ppm). Moreover, the different shielding and deshielding effects experienced by each proton in the two rotamers were predicted even better than this,

because the predicted differences between the ^{1}H chemical shifts of the rotamers matched the experimental ones within 0.08 ppm (Figure 7.15).

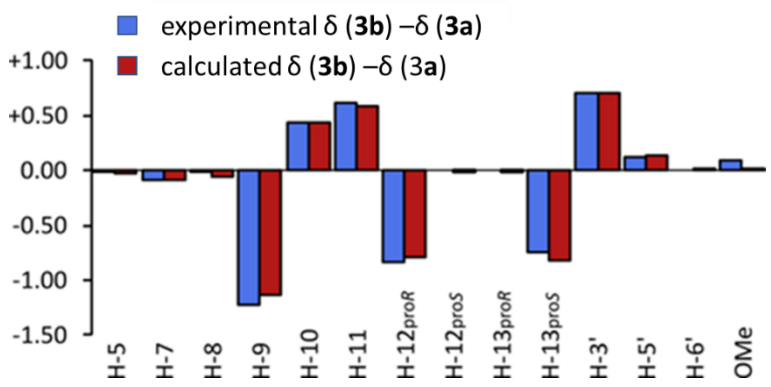


Figure 7.15. Experimental and DFT calculated differences between chemical shifts of the minor (**3b**) and the major (**3a**) rotamers of zosteraphenol A (**3**).

7.6 Absolute configuration of zosteraphenol A (3) and B (4)

Once the structure and conformation of the zosteraphenol A (**3**) and B (**4**) were completely elucidated, we were able to determine their absolute configuration using quantum mechanical prediction of its ECD spectrum.[9] The absolute configuration of the stereogenic carbon of both compounds was determined as *R* by comparing the predicted electronic circular dichroism (ECD) spectra and the experimental ECD spectra. Specifically, calculation of the theoretical spectrum of the major rotameter of the 9*R* enantiomer of **2** was done at the B3LYP/6-311+G(2d,p) level using the same geometry as the NMR calculations, and an ECD spectrum was obtained using the program SpecDis v. 1.71.[10] The same absolute configuration was assigned to zosteraphenol B (**4**), showing a similar ECD spectrum to that of zosteraphenol A (**3**).

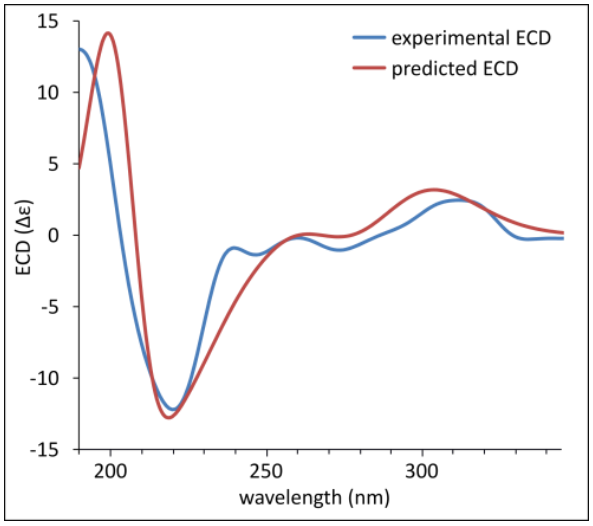


Figure 7.16. Predicted (red line) and experimental (blue line) ECD spectra of 9*R*-zosteraphenol A (**3**).

*7.7 Bioactivity evaluation of zosteraphenol A (**3**) and B (**4**)*

The diarylheptanoid zosteraphenols A (**3**) and B (**4**) anti-inflammatory activity was explored in a CD62E induction assay (Eselectin, ELAM), showing no activity at the concentration of 10 μM tested.[11] They were also tested for possible cytotoxic activity against human acute lymphoblastic leukemia (CEM), human mammary adenocarcinoma (MCF7), human cervical carcinoma (HeLa), and normal fibroblasts (BJ), showing no interesting results.[12]

Table 7.3. Cytotoxicity (μM) in human acute lymphoblastic leukemia (CEM), human breast adenocarcinoma (MCF7), human cervical carcinoma (HeLa) cells, and normal fibroblasts (BJ) after 72 h of treatment of zosteraphenols A (**3**) and B (**4**). Experiment was repeated three times in triplicates.

	IC_{50} (μM)			
Compound	CEM	MCF7	HeLa	BJ
3	>50	>50	>50	>50
4	>50	>50	>50	>50

7.8 Experimental section

7.8.1 General methods

A Jasco P-200 polarimeter at 589 nm was used to evaluate optical rotations using a 10 cm cell. UV spectra and ECD spectra were respectively measured on a Jasco V-530 spectrophotometer and on a Jasco 715 spectropolarimeter. NMR spectra were performed on Varian Unity Inova spectrometers at 700 MHz and chemical shifts were referenced to the residual solvent signal ($CDCl_3$: δ_H 7.26, δ_C 77.0). The 2D NMR experiments (HSQC, HMBC, COSY, NOESY) were recorded with standard Varian pulse sequences. A Thermo LTQ Orbitrap XL mass spectrometer was used to perform high-resolution ESI mass spectra and Column chromatography separations were performed with Kieselgel Si 60 (Merck, Darmstadt, Germany).

7.8.2 Collection, extraction, isolation

Freshly beached unrooted plants of *Zostera marina L.* (Zosteraceae) were collected along the coast of close to the Olympiazentrum Schilksee, Kiel, Schleswig-Holstein, Germany in March 2018 at selected coordinates: N 54°25'39.0", E 10°10'17.5"; alt.: 0 m. A voucher specimen is deposited in the herbarium of the Institut für Botanik, University of Kiel (voucher code: LP-20180326A-1; KIEL0005003). First, whole *Z. marina* plants (1.25 kg) were ground and extracted with *n*-hexane and acetone at room temperature five times (each time with 3.5 L). The extracts were dried under vacuum obtaining 6.75 g and 9.86 g residue, respectively.

Fractionation of the crude hexane extract was performed using medium pressure silica gel column chromatography, performed using a Büchi PrepChrom C-700 system with a silica gel column (Sepacore Silica 40–63

μm, 80 g, 194 × 31 mm) at a flow rate of 10.0 ml/min. The employed linear gradients started with 100% hexane to 100% CH_2Cl_2, and then continued from 100% CH_2Cl_2 to 100% MeOH; run time 2 hours, yielding a total of 22 fractions.

Fraction A19 eluting with 45 % CH_2Cl_2 and 55% MeOH contained zosteraphenol A (**3**) while fraction A17 eluting with 52 % CH_2Cl_2 and 48% MeOH contained zosteraphenol B (**4**). Further fractionation of both fractions was performed by Sephadex LH-20 column chromatography (Sephadex LH-20, 2 cm × 100 cm, CH_2Cl_2/acetone (85:15 v: v) as eluent). Finally, partially purified purification compounds **3** (16.5 mg) and **4** (23.5 mg) were subjected to semi-preparative reverse-phase HPLC, rendering zosteraphenol A (**3**) (7.7 mg), collected at 22.0-25.0 min, and B (**4**) (10.5 mg), eluted at 40.0-43.5 min as pure compounds. Chromatographic purification was performed using Waters e2695 instrument with a Nucleodur C_8, cm × 25 cm column and isocratic flow at 2.00ml/min in 70% MeOH, 30% water. Wavelengths detected were 210 nm and 254 nm.

Zosteraphenol A (**3**): colorless oil; $[\alpha]_D^{20}$ −18 (c 0.15, MeOH); UV (MeOH) λ_{max} (ε): 206 (20600), 254 (7700), 307 (3000) nm; ^{1}H NMR data, see Table 7.1; ECD (MeOH) λ_{max} (Δε): 195 (+13.0), 225 (–12.2), 252 (–1.4), 278 (–1.0), 317 (+2.5); ^{13}C NMR data, see Table 7.1; HR-ESI-MS *m/z* 307.1332 $[M+H]^+$ (calcd. for $C_{20}H_{19}O_3$ 307.1329).

Zosteraphenol B (**4**): colorless oil; $[\alpha]_D^{20}$ −16 (c 0.14, MeOH); UV (MeOH) λ_{max} (ε): 206 (20000), 254 (7400), 306 (3000) nm; ECD (MeOH) λ_{max} (Δε):

197 (+22.2), 225 (–17.4), 276 (–1.0), 317 (+3.9); ^{1}H NMR data, see Table 7.2; ^{13}C NMR data, see Table 7.2; HR-ESI-MS *m/z* 321.1488 $[M+H]^+$ (calcd. for $C_{21}H_{21}O_3$ 321.1485).

7.8.3 Molecular Dynamics Simulations.

MD calculations were performed using the CFF91 force field in the INSIGHT II/Discover package. The effect of the solvent was approximated by using a dielectric constant of 4.81 (chloroform). A series of 10-ns MD simulations were performed. The coordinates were saved every 50 ps, giving 200 structures for each simulation, which were minimized in the same force field. Because MD simulations must necessarily be very short, it is unlikely that a conformational change causing coalescence (i.e. when the average life of conformers is in the order of many milliseconds) can be observed during the duration of the simulation. If temperature is raised, however, the equilibrium is faster and the conformational change may eventually be observed. Therefore, simulations were performed at increasingly high temperature, namely 800 K, 1000 K, 1500 K, and 2000 K. The double bond at position 10 was restrained to cis during the simulations to avoid isomerization, and the restraint was removed before minimization. [2] The 9*R* enantiomer of zosteraphenol A (**3**) (9*R*-**3**) was arbitrarily chosen for calculations. No remarkable conformational change was observed for **3** in simulations carried out up to 1500 K. At 2000 K, inversion of the axial chirality of the molecule occurred several times during the simulation. After minimization, the 200 structures from MD simulation converged into a set of 11 conformations for each axial rotamer. The rotamers were called **3a**

and **3b**, respectively, all within 3 kcal/mol from the lowest energy conformer (Figure 7.17). The conformers within each set showed the same conformation of the carbon skeleton, but differed in the dihedral angle about the C—O bonds. These were used as starting structures for DFT calculations.

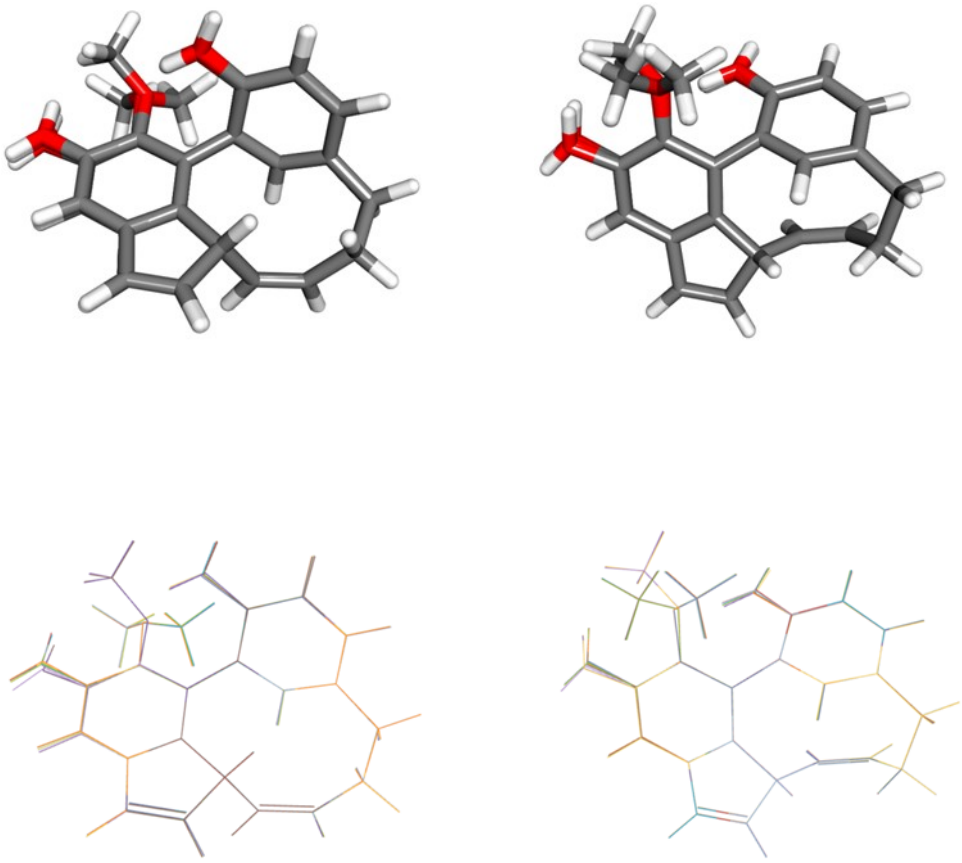


Figure 7.17. Overlay of the conformers of the two rotamers **3a** and **3b** of zosteraphenol A (**3**) generated by a 2000 K molecular dynamics simulation, represented as stick (top) and wireframe (bottom) models.

7.8.4 Quantum-Mechanical Calculation of ^{1}H and ^{13}C NMR Chemical Shifts and ECD Spectrum of compounds **3** and **4**.

Quantum-mechanical calculations were performed using density functional theory (DFT) in the program Gaussian 16.[8] The geometry of the two sets of conformers from MD was first optimized using the B3LYP/6-31G(d) level of theory to roughly estimate their energies and filter out high energy conformers. Optimization caused the 11 conformers of the rotamer **3a**, to converge into 6 conformers; in addition, all the conformers other than the lowest-energy conformer showed relative energies higher than 3 kcal/mol,

and were therefore excluded from subsequent calculations. Likewise, the 11 conformers of rotamer **3b** converged into 6 conformers, and only the lowest energy conformer was retained for subsequent calculations.

The lowest energy conformers of rotamers **3a** and **3b** were re-optimized at the B3LYP/6-31+G(d,p) level; vibrational frequency analysis revealed no imaginary frequencies, confirming **3a** and **3b** to be in true energy minima. Their Cartesian coordinates are reported in Table 7.4. Finally, the ^{1}H and ^{13}C NMR chemical shifts of **3a** and **3b** were calculated using the Gauge Invariant Atomic Orbitals (GIAO) method at the PBE0/6-311+G(2d,p) level of theory, using the PCM model for the solvent (chloroform). The scaling factors needed to convert the isotropic shielding constants calculated by Gaussian at this level of theory into chemical shifts have been pre-calculated and are available in the literature (^{1}H: slope -1.0958, intercept 31.7532; ^{13}C: slope -1.0533, intercept: 187.3123).[7] Isotropic shielding constants and calculated chemical shift for **3a** and **3b** are reported in Table 7.5.

The theoretical ECD spectrum of zosteraphenol A (**3**) was calculated using time-dependent DFT (TDDFT), the same geometry as for NMR calculations, the B3LYP/6-311+G(2d,p) level of theory, and the PCM model for the solvent (methanol). The calculated rotational strengths of the rotamers **3a** and 3**b** are listed in Table 7.6. The Boltzmann-weighted mean of the two spectra was calculated using Boltzmann statistics (T = 298 K) on the basis of their population as estimated from NMR data. From this, the ECD curve of **3** was then obtained using the program SpecDis v. 1.71[10] and the adjustable parameters were optimized for the best fit between the

theoretical and experimental spectra (the exponential half-width of the ECD bands, σ, was empirically set to 0.28 eV and the UV shift was set to –5 nm).

7.8.5 Bioactivity assays

Cytotoxicity of zosteraphenol A (**3**) and B (**4**) in human acute lymphoblastic leukaemia (CEM), human mammary adenocarcinoma (MCF7), human cervical carcinoma (HeLa) and normal fibroblasts (BJ) was determined by resazurin assay following the manufacturer's protocol (Sigma Aldrich, St. Louis, MO, USA). The procedure was described earlier, but it was performed using another dye (resazurin). The data shown are means ± standard deviation (SD) obtained from at least three independent experiments performed in triplicates.[12]

The Cell-Surface ELISA CD62E (E-Selectin, ELAM) assay was performed using enzyme-linked activity assay (ELISA) to verify the levels of cell adhesion molecule ELAM on HUVECs after 30 min of incubation with tested compounds and 4 h of stimulation with TNFα.[11] Experiments were repeated three times in triplicate. In addition, Calcein AM (Molecular Probes, Invitrogen, Karlsruhe, Germany) cytotoxicity assays after 4 h of treatment in the HUVECs were used to measure the cytotoxicity of compounds for ELAM expression assay.[11] Triplicates of at least three independent experiments were used.

7.9 Computational and spectroscopic data

Table 7.4. Cartesian coordinates of the lowest-energy conformations of the two rotamers of zosteraphenol A (**3**), optimized at the B3LYP/6-31+G(d,p) level.

Major rotamer (**3a**), E = −998.647341 hartree

C	3.236618	0.356271	-0.006001
C	3.015391	1.726151	-0.123069
C	1.695605	2.178117	-0.099970
C	0.594072	1.298467	0.024385
C	0.807505	-0.097126	0.113641
C	2.150448	-0.523292	0.125953
C	-0.344583	-1.047037	0.112611
C	-1.487889	-0.658370	0.821677
C	-2.765597	-1.142192	0.552309
C	-2.868599	-2.218191	-0.338061
C	-1.729068	-2.753441	-0.936797
C	-0.471923	-2.160004	-0.756942
C	1.201577	3.552355	-0.207346
C	-0.141434	3.556776	-0.128978
C	-0.691307	2.154195	0.066221
C	-3.957178	-0.308655	0.982060
H	-3.092028	1.624854	1.516427
C	-3.710210	1.216422	0.710590
O	4.513145	-0.125858	-0.021606
H	3.857806	2.401128	-0.233227
O	2.490656	-1.881024	0.202404
H	-1.381820	0.161560	1.519434
H	-3.845117	-2.624721	-0.591135
H	-1.797056	-3.595669	-1.618106
O	0.548540	-2.662697	-1.508964
H	1.400701	-2.467225	-1.074002
H	1.841929	4.419980	-0.324828
H	-0.789581	4.424570	-0.173553
C	-1.809475	1.840712	-0.911368
H	-4.195557	-0.430112	2.047244
H	-4.840988	-0.642026	0.425821
C	-3.079788	1.497538	-0.635022
H	-4.674955	1.735555	0.779551
C	2.356846	-2.477232	1.513631
H	4.457426	-1.094617	-0.047548
H	-3.734854	1.333125	-1.491277
H	-1.528146	1.944685	-1.958624
H	3.001122	-1.954727	2.228219
H	1.316853	-2.437091	1.848018
H	2.675542	-3.515557	1.408805
H	-1.096066	2.124134	1.086967

Minor rotamer (**3b**) , E = −998.646161 hartree

C	3.202564	0.488294	-0.130926
C	2.911747	1.820330	0.160599
C	1.572160	2.181746	0.312732
C	0.517172	1.255764	0.140377
C	0.809109	-0.106952	-0.089326
C	2.163964	-0.449861	-0.242605
C	-0.288321	-1.114071	-0.024017
C	-1.475124	-0.785139	-0.686466
C	-2.721692	-1.287048	-0.316426
C	-2.739896	-2.324996	0.626550
C	-1.549190	-2.816978	1.164818
C	-0.321153	-2.205424	0.875223
C	1.013417	3.448483	0.793877

C	-0.318467	3.325876	0.948146
C	-0.812788	1.971915	0.460242
C	-3.954032	-0.515705	-0.738216
H	-3.612504	0.897615	0.883342
C	-3.904420	0.947393	-0.171390
O	4.502137	0.094851	-0.266095
H	3.722575	2.526194	0.308652
O	2.564059	-1.781235	-0.409483
H	-1.424949	-0.006582	-1.431725
H	-3.687115	-2.748681	0.952485
H	-1.553003	-3.642012	1.870261
O	0.778107	-2.668296	1.539550
H	1.580221	-2.444393	1.032470
H	1.607325	4.327357	1.020968
H	-1.000523	4.098854	1.283524
C	-1.732611	2.257686	-0.725534
H	-4.853518	-1.010576	-0.354949
H	-4.054019	-0.473984	-1.830610
C	-3.001051	1.874729	-0.953135
H	-4.921158	1.359915	-0.199918
C	2.324948	-2.342025	-1.720349
H	4.508381	-0.874496	-0.317507
H	-3.438857	2.234501	-1.885891
H	-1.267943	2.896667	-1.476687
H	2.710096	-3.362513	-1.688679
H	1.254819	-2.353637	-1.943554
H	2.859143	-1.760919	-2.479431
H	-1.361402	1.430022	1.240411

Table 7.5. Experimental chemical shifts, calculated isotropic shielding constants, and predicted chemical shifts of the two rotamers of zosteraphenol A (**3**). Calculations were performed at the PBE0/6-311+G(2d,p)/PCM($CHCl_3$)//B3LYP/6-31+G(d,p) level.

	major rotamer (**3a**)				minor rotamer (**3b**)			
	exp. δ_H	shielding	calcd. δ_H[a]	$\Delta\delta_H$	exp. δ_H	shielding	calcd. δ_H[a]	$\Delta\delta_H$
H-5	6.99	24.34	6.88	−0.11	6.97	24.37	6.85	−0.12
H-7	6.72	24.52	6.72	0.00	6.63	24.61	6.63	0.00
H-8	6.29	24.94	6.32	0.03	6.28	25.01	6.26	−0.02
H-9	4.79	26.58	4.81	0.02	3.57	27.82	3.67	0.10
H-10	4.61	26.69	4.71	0.10	5.04	26.22	5.14	0.10
H-11	5.16	26.02	5.33	0.17	5.77	25.38	5.92	0.15
H-12$_{pro\mathit{R}}$	2.99	28.52	3.02	0.03	2.15	29.37	2.24	0.09
H-12$_{pro\mathit{S}}$	2.43	29.16	2.43	0.00	2.43	29.17	2.43	0.00
H-13$_{pro\mathit{R}}$	2.65	28.90	2.67	0.02	2.65	28.91	2.67	0.02
H-13$_{pro\mathit{S}}$	3.38	28.12	3.40	0.02	2.63	29.00	2.58	−0.05
H-3'	7.76	23.45	7.70	−0.06	8.46	22.68	8.41	−0.05
H-5'	6.90	24.36	6.86	−0.04	7.02	24.22	6.99	−0.03
H-6'	6.85	24.47	6.76	−0.09	6.85	24.46	6.77	−0.08
		28.07				27.69		
1-OMe	3.88	27.66	3.68[b]	−0.20	3.97	27.55	3.75[b]	−0.22
		27.71				27.98		
			RMSD ^{1}H	**0.09**			**RMSD ^{1}H**	**0.10**

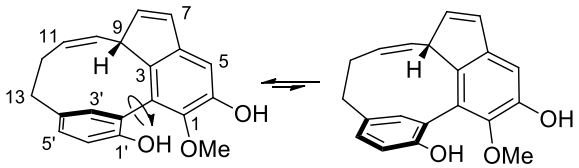
	exp. δ_C	shielding	calcd. δ_C[c]	$\Delta\delta_C$
C-1	139.7	42.52	138.2	−1.5
C-2	128.3	51.27	129.8	1.5
C-3	139.5	41.35	139.3	−0.2
C-4	143.6	36.76	143.7	0.1
C-5	107.7	76.37	105.8	−1.9
C-6	147.9	31.74	148.5	0.6
C-7	130.4	49.70	131.3	0.9
C-8	140.0	38.05	142.4	2.4
C-9	51.4	130.63	53.8	2.4
C-10	127.0	51.42	129.6	2.6
C-11	132.0	46.00	134.8	2.8
C-12	27.5	156.77	28.8	1.3
C-13	34.2	148.96	36.3	2.1
C-1'	151.6	27.15	152.9	1.3
C-2'	122.4	59.15	122.2	−0.2
C-3'	134.8	46.01	134.8	0.0
C-4'	132.3	49.53	131.4	−0.9
C-5'	130.1	51.42	129.6	−0.5
C-6'	116.5	66.06	115.6	−0.9
C-OMe	62.8	124.63	59.5	−3.3
			RMSD ^{13}C	**1.7**

[a] ^{1}H chemical shifts were obtained from the computed isotropic shielding constants according to ref. 7 using the equation: $\delta = (31.7532 - \text{shielding})/1.0958$

[b] The average chemical shift of the three methyl protons is reported

[c] ^{13}C chemical shifts were obtained from the computed isotropic shielding constants according to ref. 7 using the equation: $\delta = (187.3123 - \text{shielding})/1.0533$

Table 7.6. Experimental and calculated chemical shift differences between the minor (**3b**) and the major (**3a**) rotamers of zosteraphenol A (**3**).



major rotamer (**3a**) minor rotamer (**3b**)

	δ(**3b**) − δ(**3a**)		
	exp.	calcd.	error
H-5	−0.02	−0.03	−0.01
H-7	−0.09	−0.09	+0.00
H-8	−0.01	−0.06	−0.05
H-9	**−1.22**	**−1.14**	**+0.08**
H-10	**+0.43**	**+0.43**	**+0.00**
H-11	**+0.61**	**+0.59**	**−0.02**
H-12$_{proR}$	**−0.84**	**−0.78**	**+0.06**
H-12$_{proS}$	+0.00	−0.01	−0.01
H-13$_{proR}$	+0.00	−0.01	−0.01
H-13$_{proS}$	**−0.75**	**−0.82**	**−0.07**
H-3'	**+0.70**	**+0.71**	**+0.01**
H-5'	+0.12	+0.13	+0.01
H-6'	+0.00	+0.02	+0.02
OMe	+0.09	+0.02	−0.07

Table 7.7. Rotatory strengths (length formalism) calculated for rotamers **3a** and **3b**.

major rotamer (**3a**)		minor rotamer (**3b**)	
Wavelength (nm)	Rotatory strength (10^{-40} erg esu cm Gauss^{-1})	Wavelength (nm)	Rotatory strength (10^{-40} erg esu cm Gauss^{-1})
174.3	2.526000	175.2	5.845800
174.9	11.100900	175.4	10.798600
175.3	-8.707300	175.7	2.925100
175.9	3.713700	175.9	-37.875900
176.1	9.015100	176.2	-5.918100
176.4	-2.029500	176.9	4.093400
176.8	-2.784600	177.1	7.116600
177.2	16.889600	177.8	9.999500
177.4	-22.305500	178.3	5.861500
178.0	9.300700	178.7	-5.385900
178.2	12.834100	179.0	-9.374400
179.0	-56.526700	179.3	50.183300
179.5	45.076900	179.6	17.229700
180.2	-19.352000	180.0	8.109500
180.7	34.096000	180.4	2.476300
181.6	6.834000	180.8	-6.145500
182.0	12.124300	181.6	1.643900
182.7	-62.245700	181.8	-38.750300
183.4	-0.942800	182.2	3.609900
184.3	16.294400	183.2	-19.427600
184.9	2.420100	183.8	-10.055800
185.3	41.504500	184.9	0.666200
185.6	10.734900	185.5	-3.247700
186.1	-2.384000	185.7	1.013500
186.9	-12.282200	186.0	1.455800
187.3	-9.672800	186.8	-0.874000
187.9	-16.267400	187.2	-26.098200
188.3	3.147600	187.4	11.149400

188.8	6.762600	189.0	21.834800
189.0	-6.593100	189.4	-70.338600
190.1	-44.721300	190.7	-22.213300
190.5	40.798200	190.9	11.796100
191.5	-25.100700	191.9	-4.294700
192.1	10.890600	193.0	-32.158300
193.0	27.666200	193.6	-24.222000
193.6	-8.476600	193.8	-4.030400
194.1	-18.017100	194.3	-40.140300
194.6	19.724300	194.8	-66.379000
195.4	-15.080700	195.1	-8.110000
196.4	26.524000	196.6	70.683700
197.4	-13.998700	197.3	-12.739000
197.8	-24.407200	197.9	25.611000
198.3	30.724700	198.0	-17.786600
198.9	7.236700	198.8	11.932400
199.4	-14.311500	200.6	35.600800
200.8	-1.007300	200.9	3.176700
201.3	-16.694500	201.4	-0.576200
201.9	13.868700	202.9	-74.876300
202.7	2.659500	203.7	-17.119700
203.9	7.713500	205.1	-0.385200
205.1	14.131400	206.2	-29.953000
206.5	32.822500	207.6	-14.451600
208.1	-36.978400	208.4	3.164300
210.7	13.240300	208.7	22.589700
211.2	20.696600	211.4	-23.361000
212.4	-18.249200	212.2	70.924700
212.6	154.360500	212.5	-106.583700
213.6	0.499900	212.9	16.831200
215.8	60.793700	215.1	17.916600
218.5	-46.097100	217.4	-28.223600
220.0	-50.096900	217.6	18.293300
221.0	-24.959500	221.4	-23.792000
222.9	-45.006600	221.9	-25.812500
225.3	-63.064600	223.9	5.690600
227.0	35.844500	226.1	136.279200
230.6	-6.871200	232.0	6.096300
233.0	-16.674900	232.5	44.209100
235.4	8.636100	235.0	14.964800
236.3	-64.752600	238.4	29.264800
238.0	9.117700	239.6	19.702400
240.8	-22.538700	242.0	-1.432600
245.1	-14.510100	246.5	-87.201300
247.6	10.544300	249.1	35.901300
250.4	-7.120400	250.9	38.139200
256.0	-18.868600	259.2	11.015400
268.8	11.983200	264.7	-45.935000
275.3	16.270700	273.4	-24.125700
280.4	-19.832600	279.9	24.051100
296.4	-15.278000	294.4	46.060000
311.6	58.923700	316.7	-46.777300

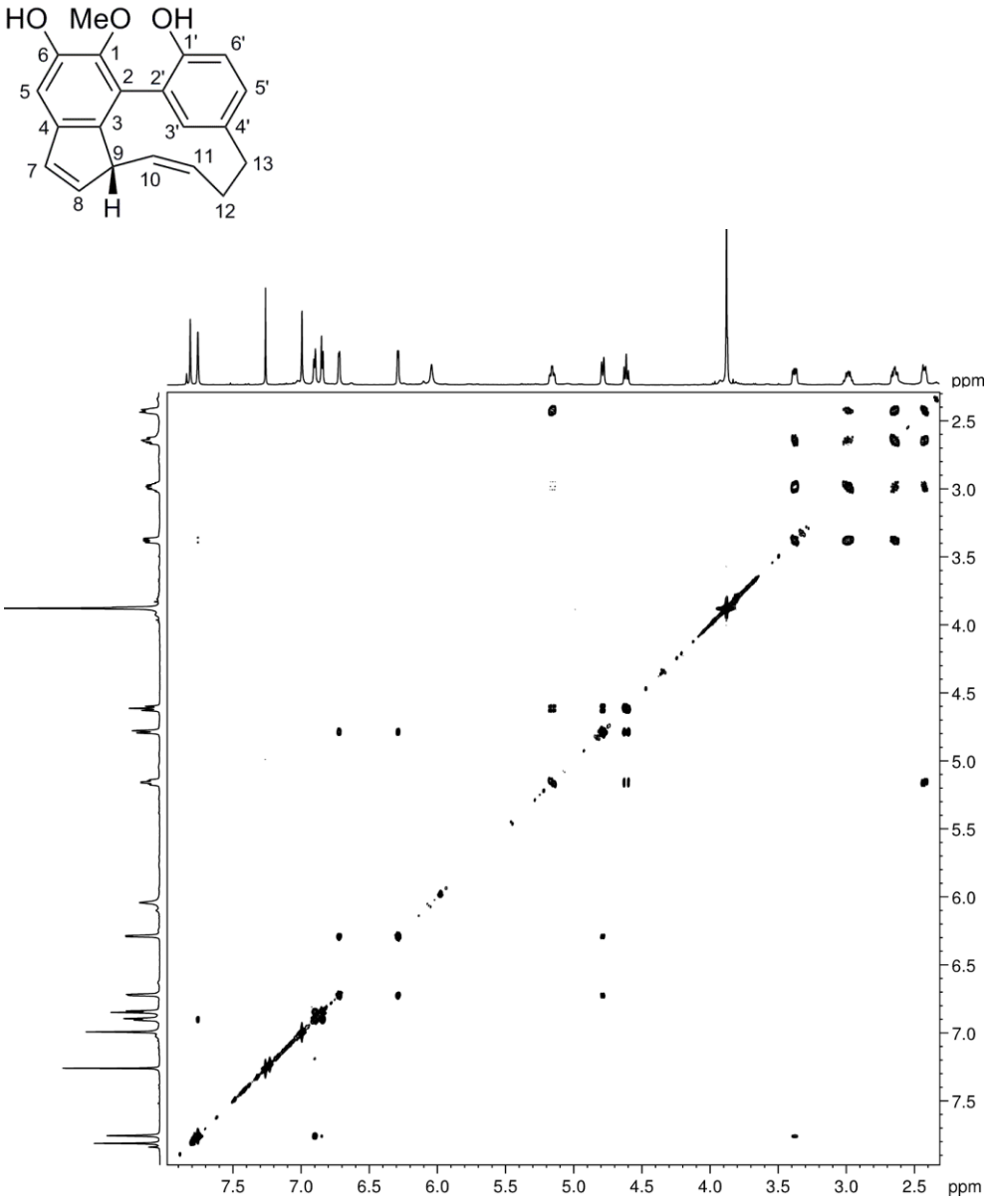


Figure 7.18. COSY spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol A (**3**).

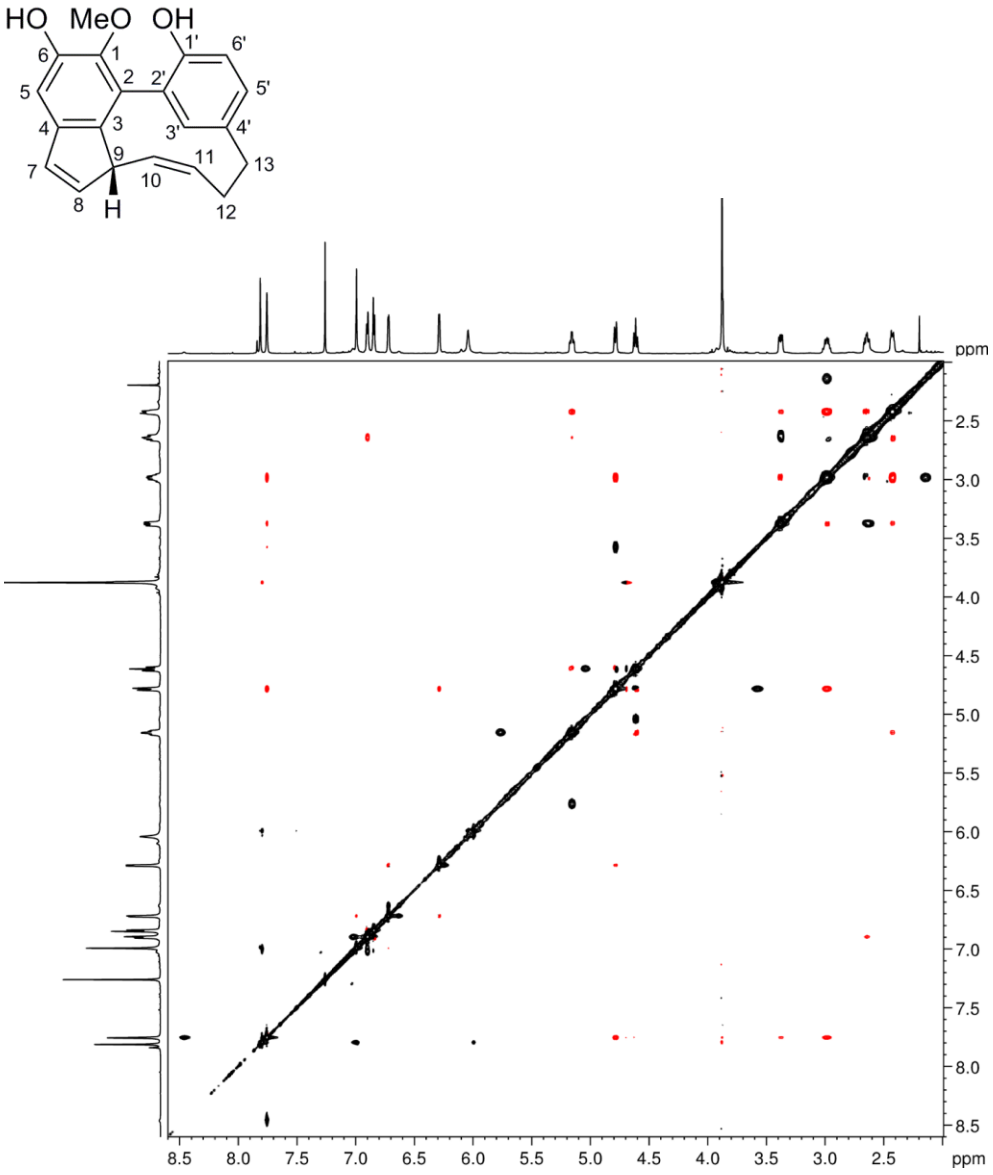


Figure 7.19. NOESY/EXSY spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol A (**3**). In the positive NOE regime, exchange peaks (black) have the same phase as diagonal peaks and NOE peaks (red) have opposite phase.

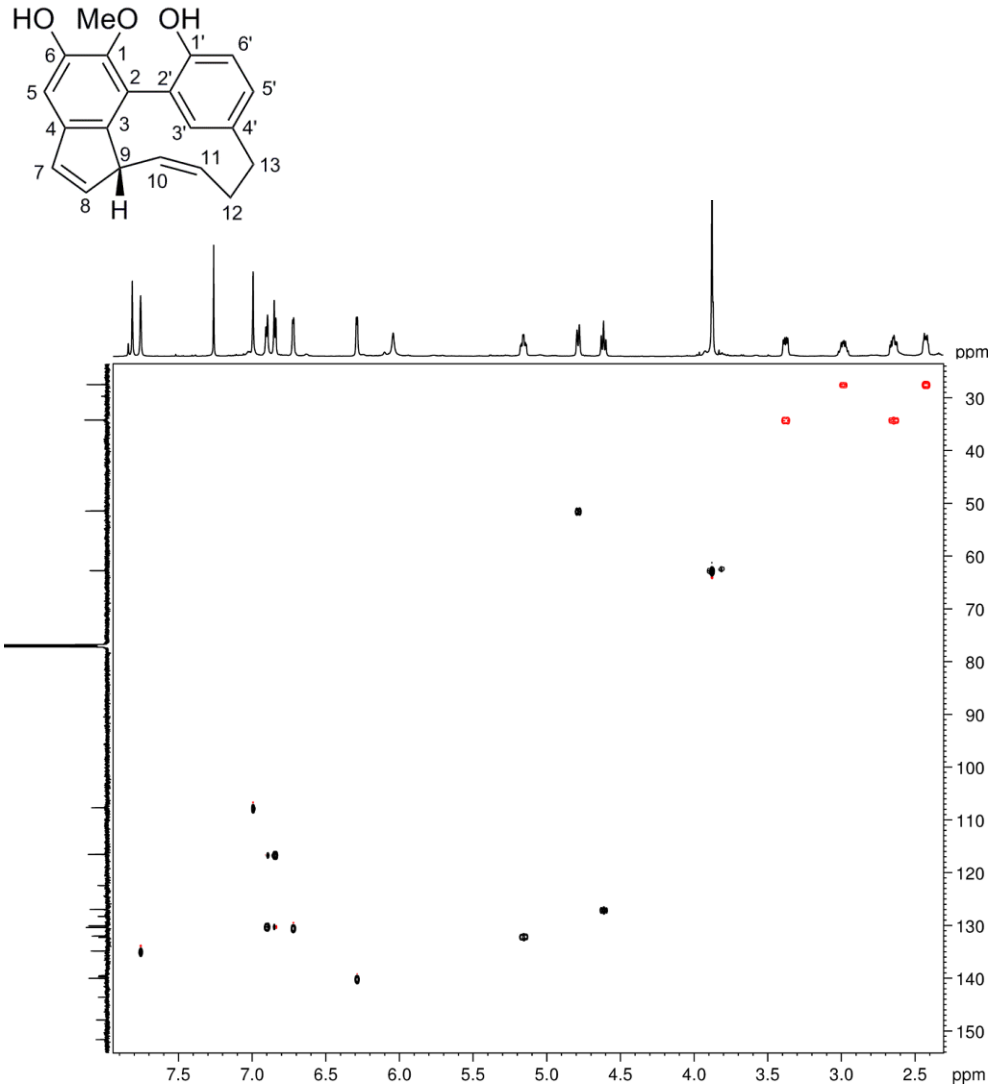


Figure 7.20. HSQC spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol A (**3**).

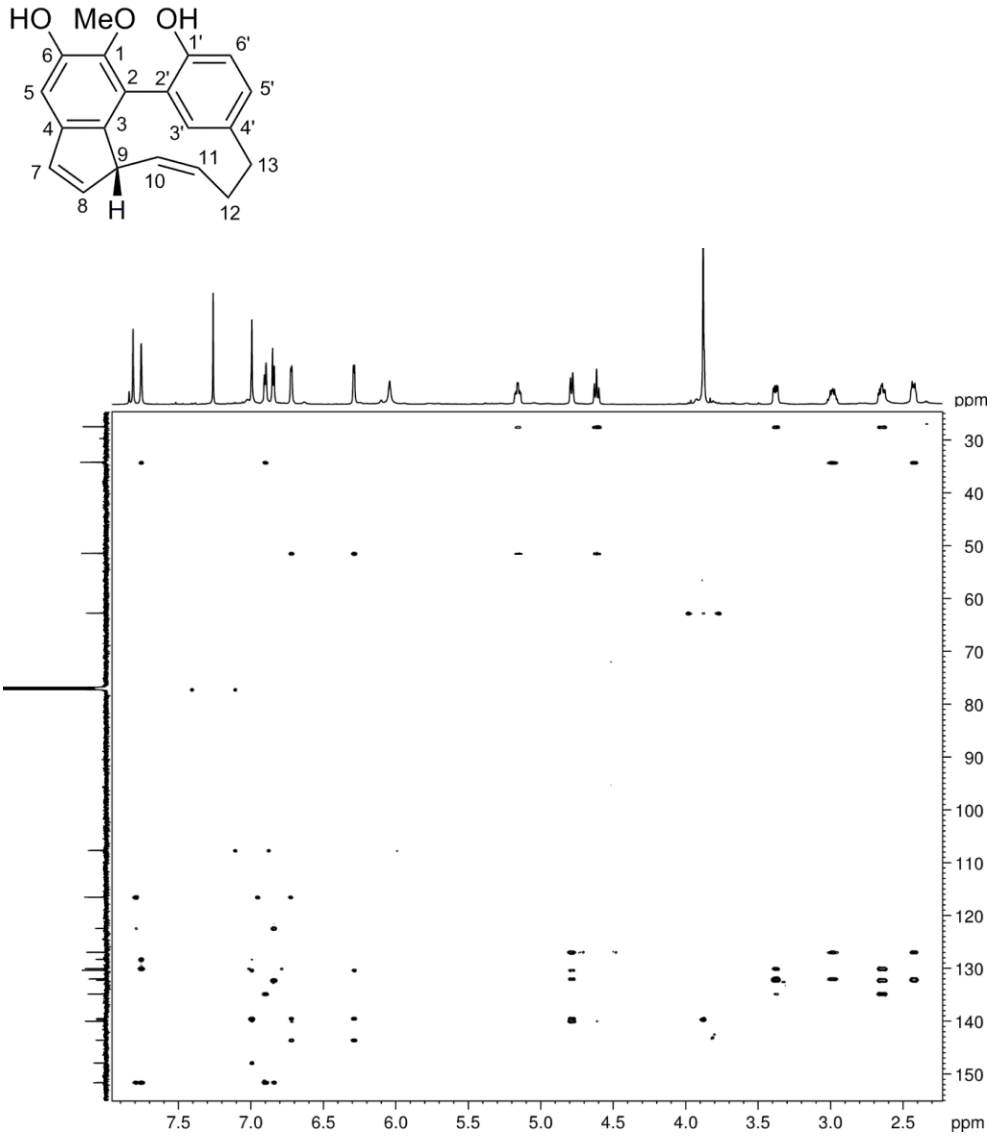


Figure 7.21. HMBC spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol A (**3**).

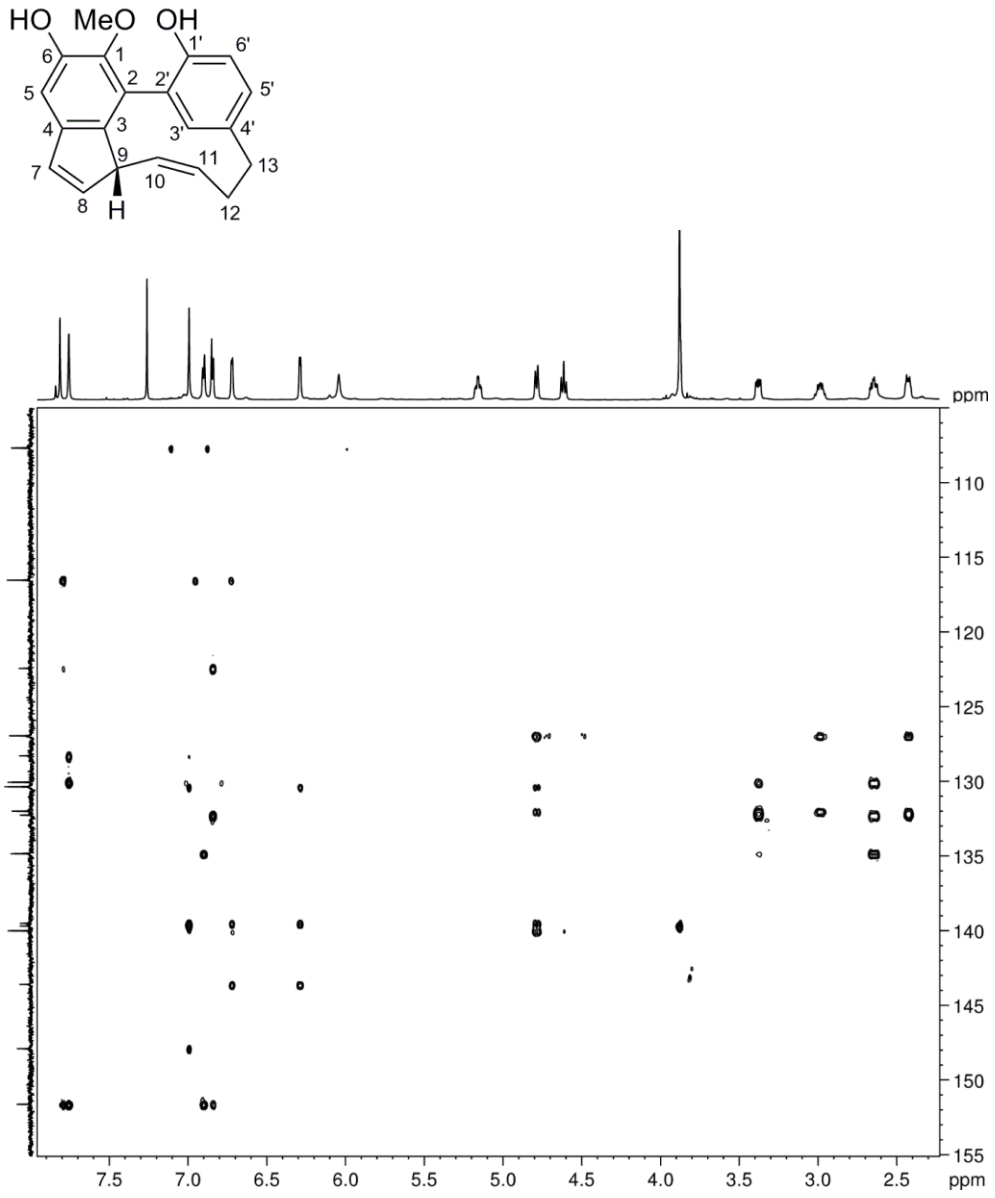


Figure 7.22. Expansion of HMBC spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol A (**3**).

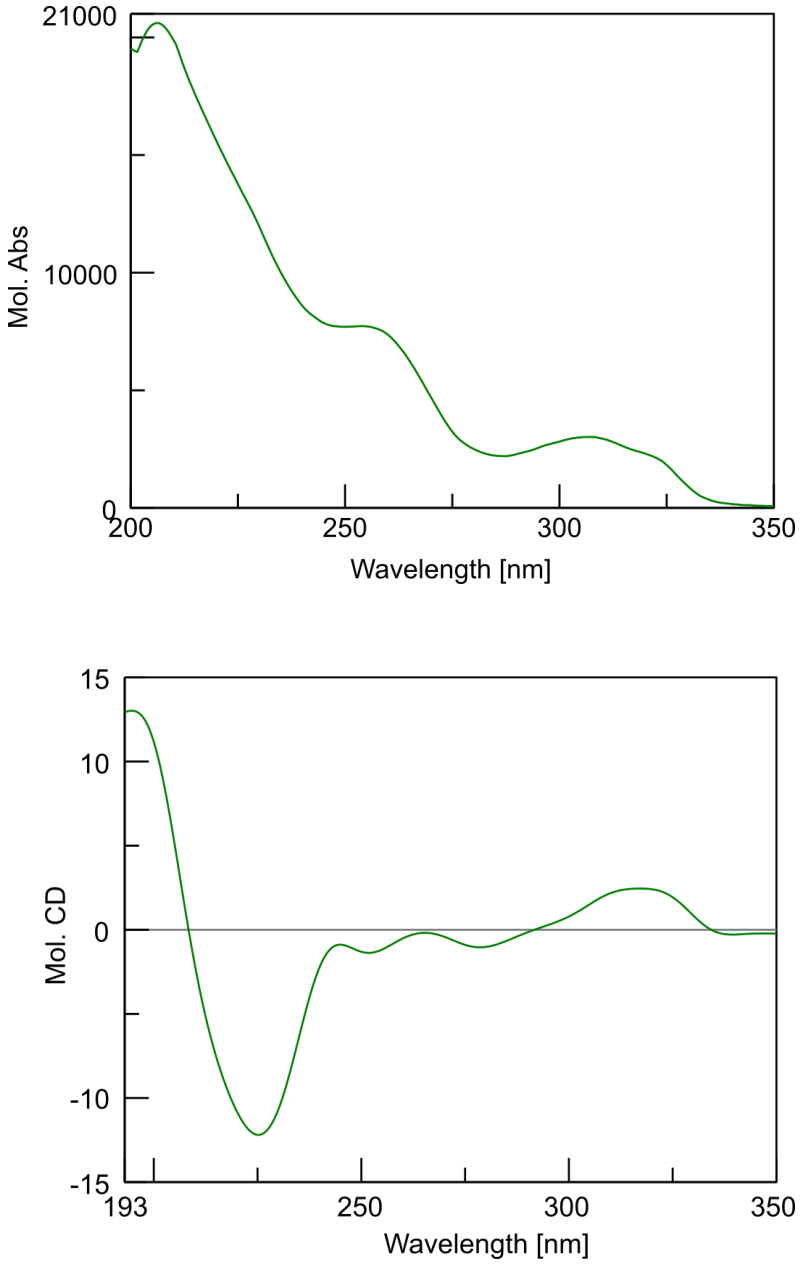


Figure 7.23. UV and ECD spectra (CH_3OH) of zosteraphenol A (**3**).

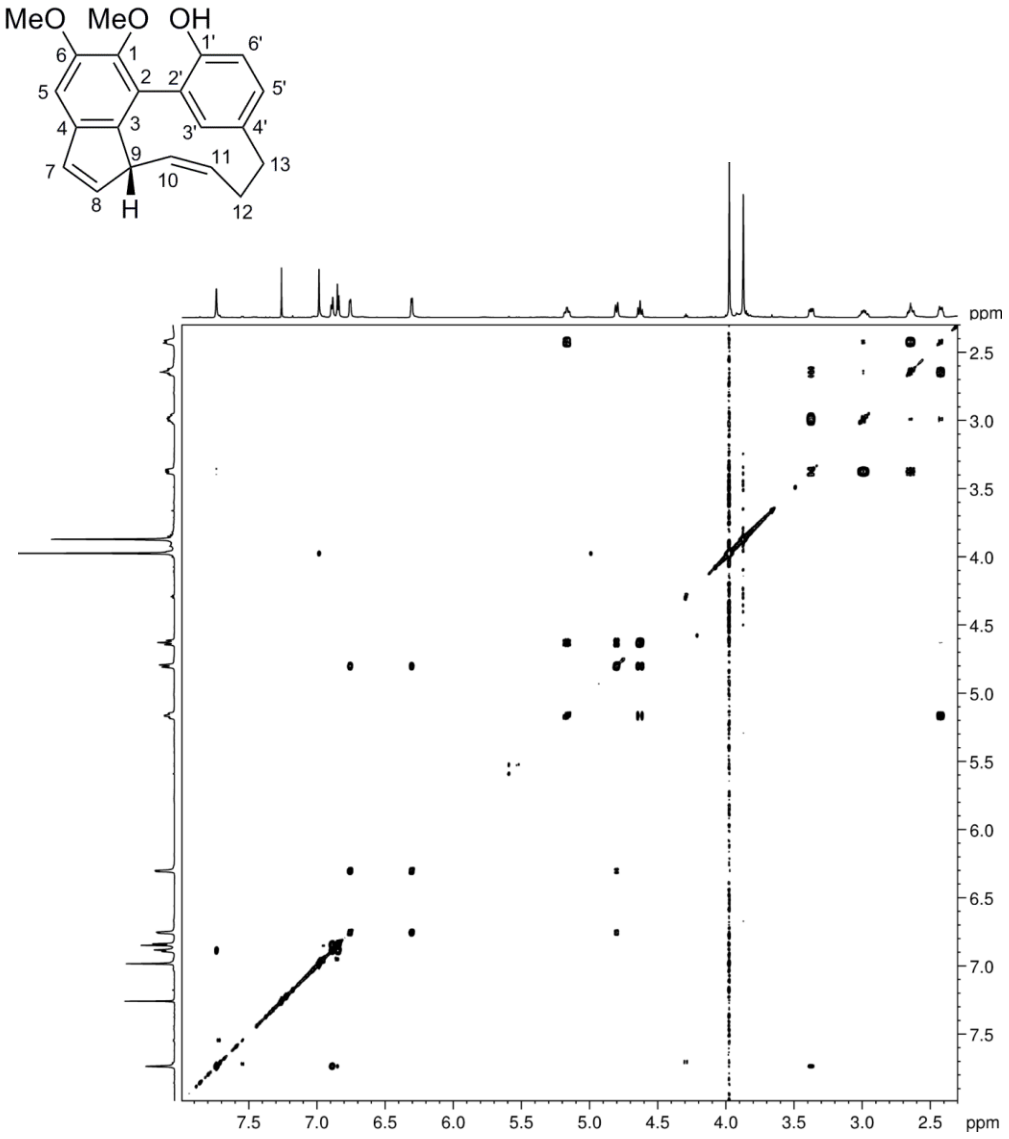


Figure 7.24. COSY spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol B (**4**).

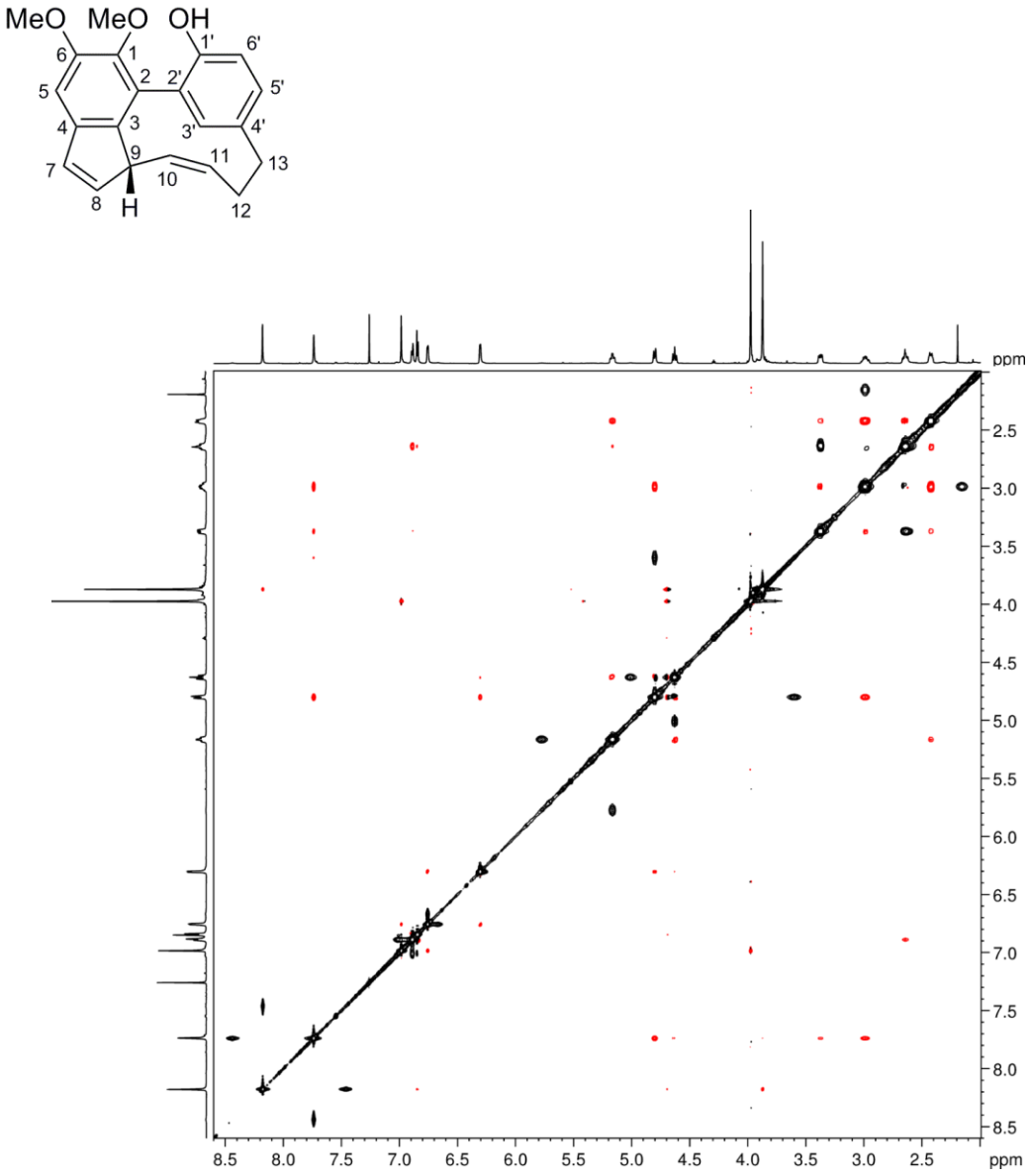


Figure 7.25. NOESY/EXSY spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol B (**4**). In the positive NOE regime, exchange peaks (black) have the same phase as diagonal peaks and NOE peaks (red) have opposite phase.

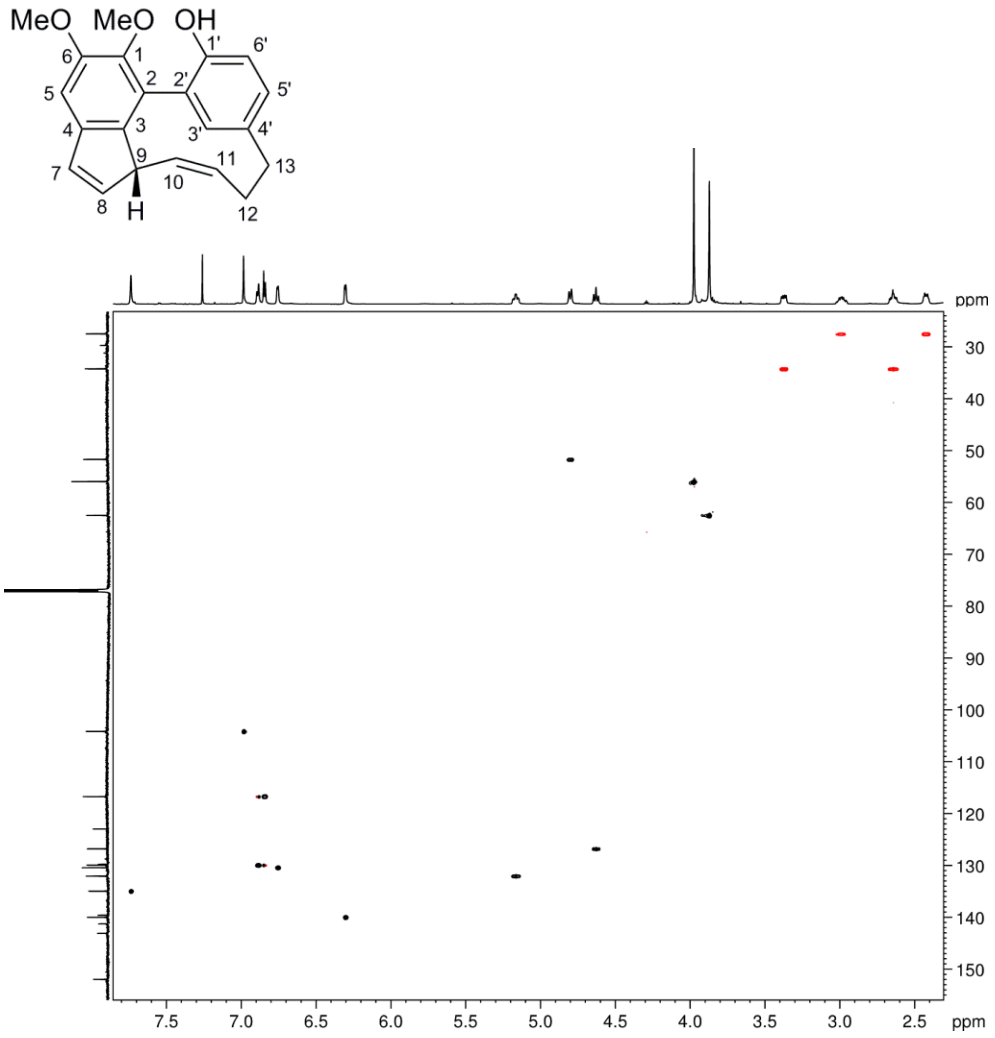


Figure 7.26. HSQC spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol B (**4**).

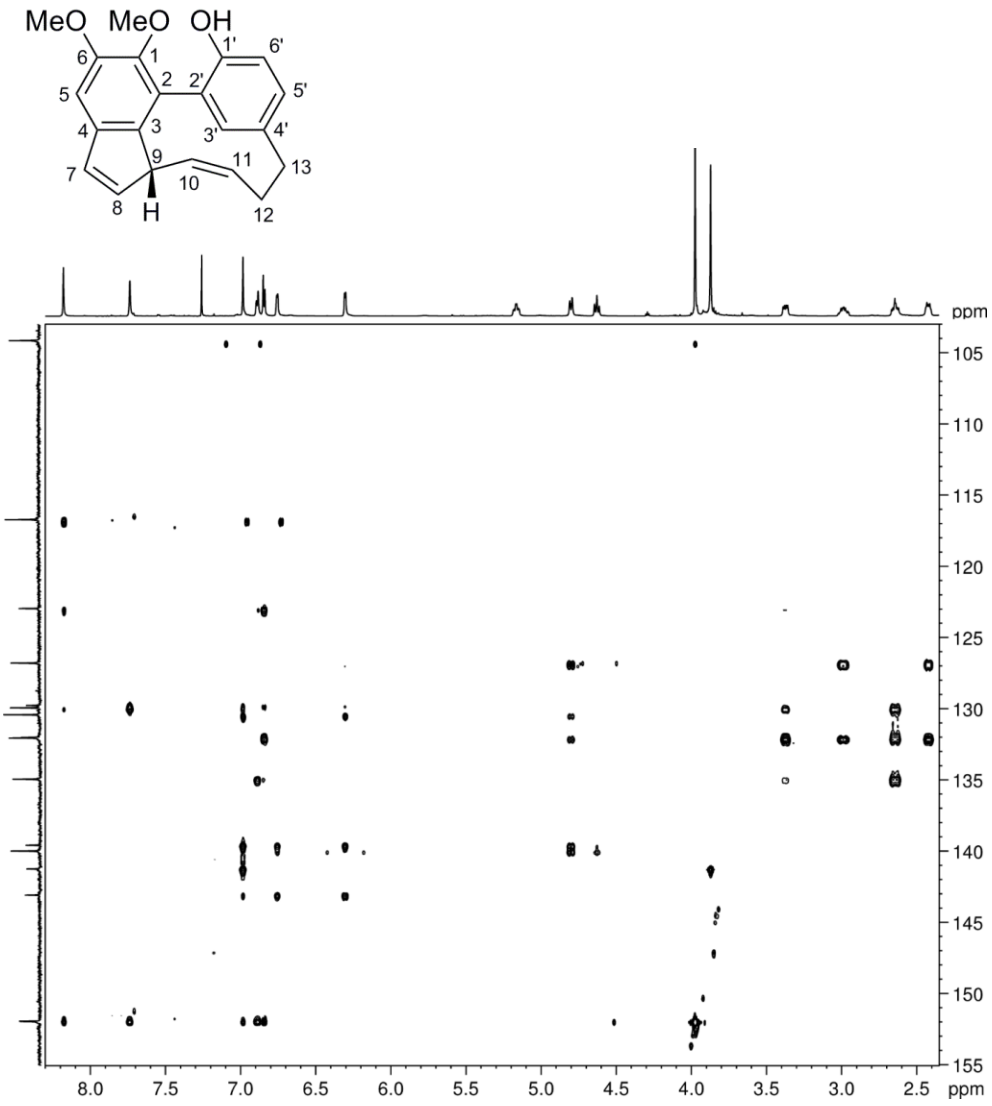


Figure 7.27 HMBC spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol B (**4**).

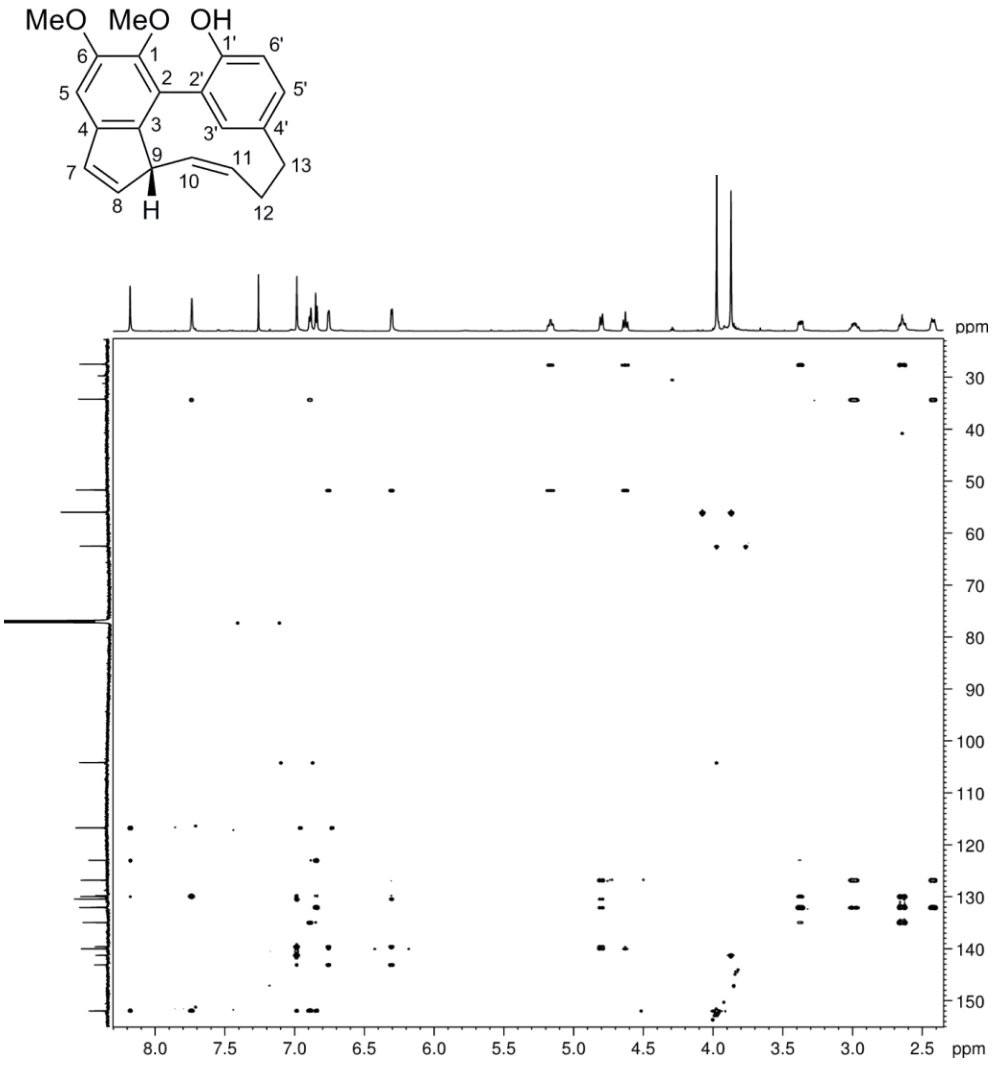


Figure 7.28. Expansion of the HMBC spectrum (700 MHz, 258 K, $CDCl_3$) of zosteraphenol B (**4**).

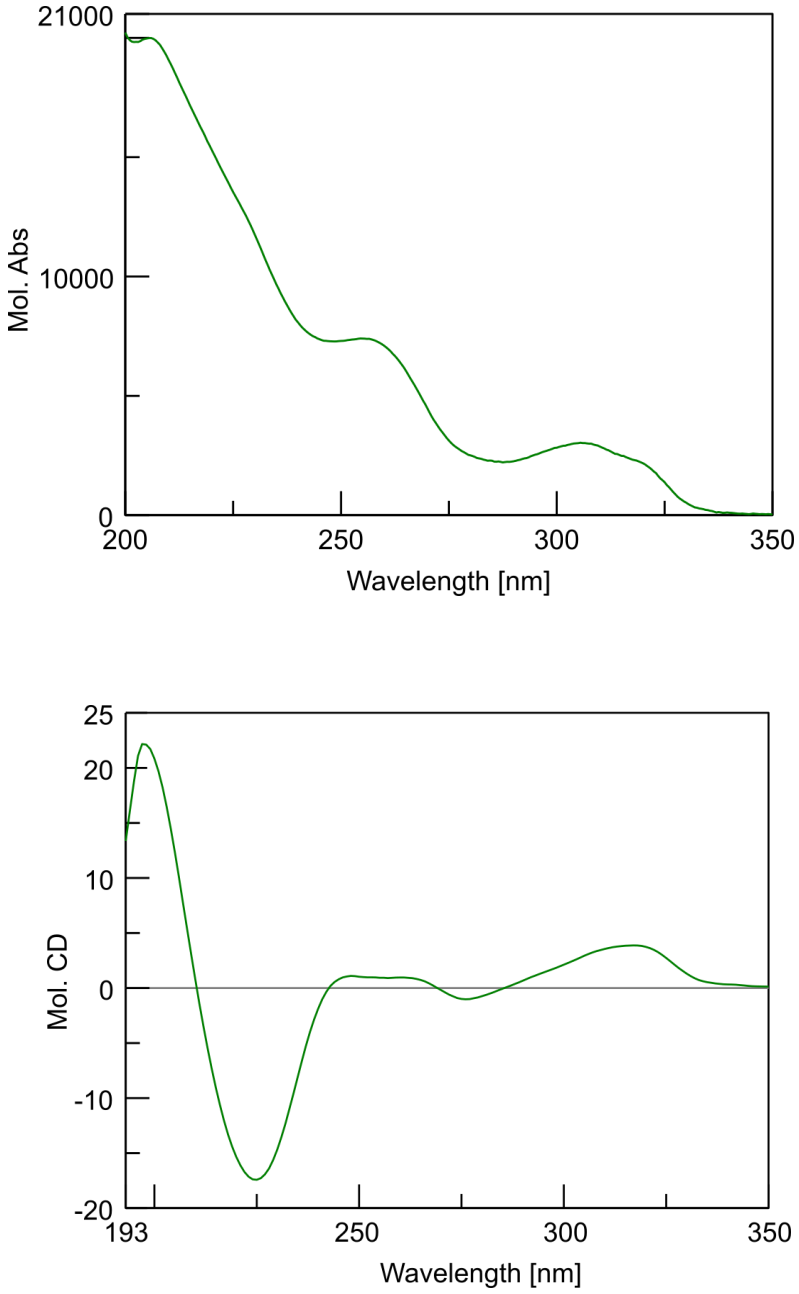


Figure 7.29. UV and ECD spectra (CH_3OH) of zosteraphenol B (**4**).

References

[1] Grauso, L.; Li, Y.; Scarpato, S.; Shulha, O.; Rárová, L.; Strnad, M.; Teta, R.; Mangoni, A.; Zidorn, C. Structure and Conformation of Zosteraphenols, Tetracyclic Diarylheptanoids from the Seagrass Zostera marina: An NMR and DFT Study. *Org. Lett.* **2020**, *22*, 78–82, doi:10.1021/acs.orglett.9b03964.

[2] Costantino, V.; Fattorusso, E.; Mangoni, A.; Perinu, C.; Teta, R.; Panza, E.; Ianaro, A. Tedarenes A and B: Structural and Stereochemical Analysis of Two New Strained Cyclic Diarylheptanoids from the Marine Sponge Tedania ignis. *J. Org. Chem.* **2012**, *77*, 6377–6383, doi:10.1021/jo300295j.

[3] Alberti, Á.; Riethmüller, E.; Béni, S. Characterization of diarylheptanoids: An emerging class of bioactive natural products. *J. Pharm. Biomed. Anal.* **2018**, *147*, 13–34, doi:10.1016/j.jpba.2017.08.051.

[4] Smyth, J.E.; Butler, N.M.; Keller, P.A. A twist of nature-the significance of atropisomers in biological systems. *Nat. Prod. Rep.* **2015**, *32*, 1562–1583, doi:10.1039/c4np00121d.

[5] Jahng, Y.; Park, J.G. Recent studies on cyclic 1,7-diarylheptanoids: Their isolation, structures, biological activities, and chemical synthesis. *Molecules* **2018**, *23*, doi:10.3390/molecules23123107.

[6] Claridge, T.D.W. *High-Resolution NMR Techniques in Organic Chemistry*; Oxford, E., Ed.; 3rd ed.;Elsevier:Oxford **2016**.

[7] Lodewyk, M.W.; Siebert, M.R.; Tantillo, D.J. Computational prediction of 1H and 13C chemical shifts: A useful tool for natural product, mechanistic, and synthetic organic chemistry. *Chem. Rev.* **2012**, *112*, 1839–1862.

[8] 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. et al. Gaussian 16, Revision A.03 **2016**.

[9] Grauso, L.; Teta, R.; Esposito, G.; Menna, M.; Mangoni, A. Computational prediction of chiroptical properties in structure elucidation of natural products. *Nat. Prod. Rep.* **2019**, *36*, 1005–1030, doi:10.1039/c9np00018f.

[10] Bruhn, T.; Schaumlöffel, A.; Hemberger, Y.; Pescitelli, G. SpecDis ver-sion 1.71, Berlin, Germany, **2017**, http://specdis-software.jimdo.com.

[11] Morrogh-Bernard, H.C.; Foitová, I.; Yeen, Z.; Wilkin, P.; De Martin, R.; Rárová, L.; Doležal, K.; Nurcahyo, W.; Olšanský, M. Self-medication by orang-utans (Pongo pygmaeus) using bioactive properties of Dracaena cantleyi. *Sci. Rep.* **2017**, *7*, 1–7, doi:10.1038/s41598-017-16621-w.

[12] Rárová, L.; Steigerová, J.; Kvasnica, M.; Bartůněk, P.; Křížová, K.; Chodounská, H.; Kolář, Z.; Sedlák, D.; Oklestkova, J.; Strnad, M. Structure activity relationship studies on cytotoxicity and the effects on steroid receptors of AB-functionalized cholestanes. *J. Steroid Biochem. Mol. Biol.* **2016**, *159*, 154–169, doi:10.1016/j.jsbmb.2016.03.017.

Chapter 8

Novel heterodimeric cyclic diarylheptanoids showing a stable catechol keto tautomerism from the Seagrass *Zostera marina*

8.1 Synopsis of this chapter [1]

The in-depth analysis of the secondary metabolites of *Zostera marina* led to the isolation of two new dimeric diarylheptanoids, zosterabisphenone A (**5**) and B (**6**), enriching the family of diarylheptanoid found from the extracts of *Zostera marina*.[2,3] Structure elucidation of both compounds revealed several structural peculiarities. The "southern" diarylheptanoid unit was closely related to zosteraphenol A (**3**), except that the methoxy group was located at C-6 and not C-1 and showed the same conformational equilibrium leading to coalescent signals in NMR spectra recorded at room temperature. All NMR spectra were, therefore, recorded at low temperature so that the conformational equilibrium was slower, and all signals were sharp enough for structure elucidation. The "northern" diarylheptanoid unit of both zosterabisphenone, whose structure was determined mainly on the basis of HMBC data, is characterized by a large modification of one of the benzene rings. In zosterabisphenone A (**5**), this ring appears to be the diketo tautomer of a catechol (1,2-dihydroxybenzene), in which the carbonyl group at position 1' forms a cyclic hemiacetal with the nearby hydroxyl group at position 14'. This structure, which shows the first

occurrence of the stable catechol tautomer in an organic molecule, has been validated by DFT prediction of the ^{1}H and ^{13}C which resulted in very good agreement with the experimental values (RMSD of 1.65 ppm for ^{13}C and 0.097 ppm for ^{1}H). The prevalence of the keto tautomer over the usual aromatic form of catechol can be explained by steric strain, because in a planar aromatic structure the bond C-8/C-3' and the C_7 bridge would necessarily collide. Additionally, HMBC data of the northern unit of zosterabisphenone B (**6**), revealed the presence of a 1,2,4,5-tetrasubstituted, trioxygenated benzene ring, while the second ring in the northern diarylheptanoid unit was determined to be a cross-conjugated cyclohexadienone. Finally, the absolute configuration of zosterabisphenone A (**5**) and B (**6**) was determined by prediction of the ECD spectrum. The cytotoxic activity of zosterabisphenone A (**5**) and B (**6**) was also evaluated, resulting in zosterabisphenone B (**6**) exhibiting cytotoxic activity in a concentration- and time-dependent manner on HCT116 cells up to 48 hours of exposure

*8.2 Isolation of zosterabisphenone A (**5**)*

The unrooted leaves of the seagrass *Zostera marina* were collected from the shore and washed fresh. After allowing them to air dry, they were extracted with acetone, and the extract was subjected to several purification steps, sequentially, by SiO_2 column chromatography, Sephadex LH-20 chromatography, and reversed-phase HPLC to obtain pure zosterabisphenone A (**5**) (4.6 mg) and B (**6**) (4.2 mg).

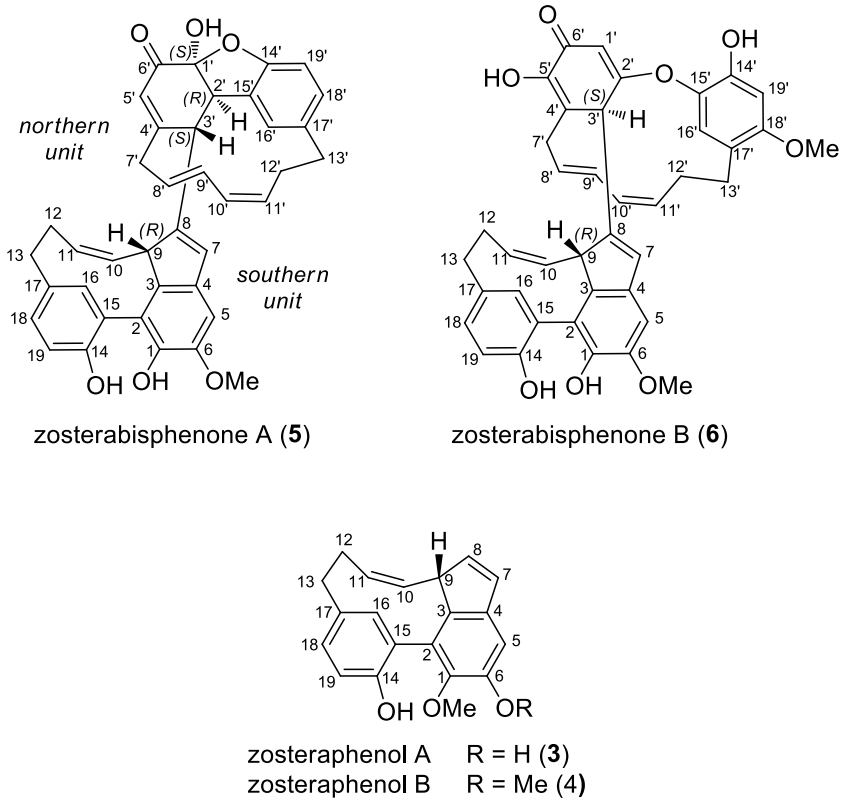


Figure 8.1. Structure of zosteraphenol A, B (**3**,**4**) and zosterabisphenone A and B (**5**,**6**)

*8.3 Structural elucidation of zosterabisphenone A (**5**)*

The high-resolution ESI mass spectrum of zosterabisphenone A (**5**) showed a $[M+Na]^+$ ion at *m/z* 621.2235 establishing a molecular formula $C_{39}H_{34}O_6$ with 23 unsaturations, suggesting a dimeric diarylheptanoid structure.

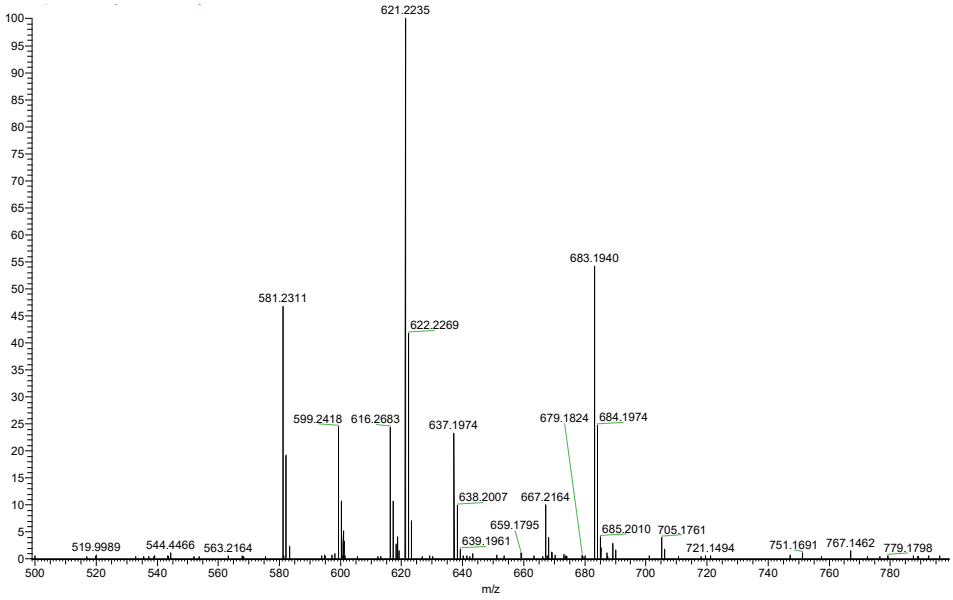


Figure 8.2. High-resolution ESI mass spectrum of zosterabisphenone A (**5**) (positive ion mode, MeOH/0.1% aqueous HCOOH).

NMR spectra of zosterabisphenone A (**5**) showed recorded at room temperature many coalescent signals, similarly to the monomeric diarylheptanoids zosteraphenols A (**3**) and B (**4**) described in Chapter 7. Therefore, all NMR spectra of zosterabisphenone A (**5**) were recorded at low temperature (253 K) where the conformational equilibrium was slower, and the signals appeared to be sharp enough to perform structural elucidation.

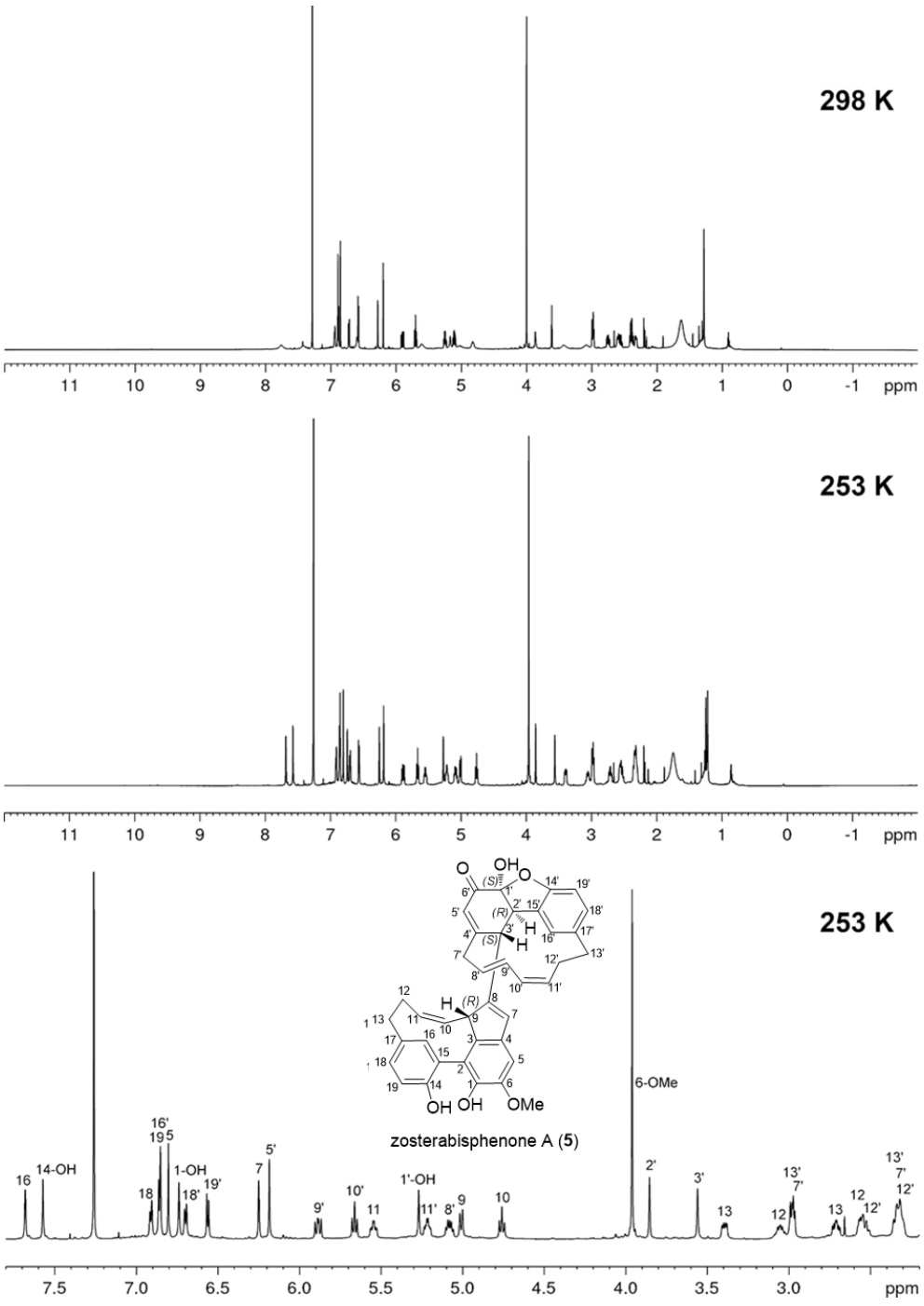


Figure 8.3. ^{1}H NMR spectrum of zosterabisphenone A (**5**) recorded at 298 K and 253 K (700 MHz, $CDCl_3$)

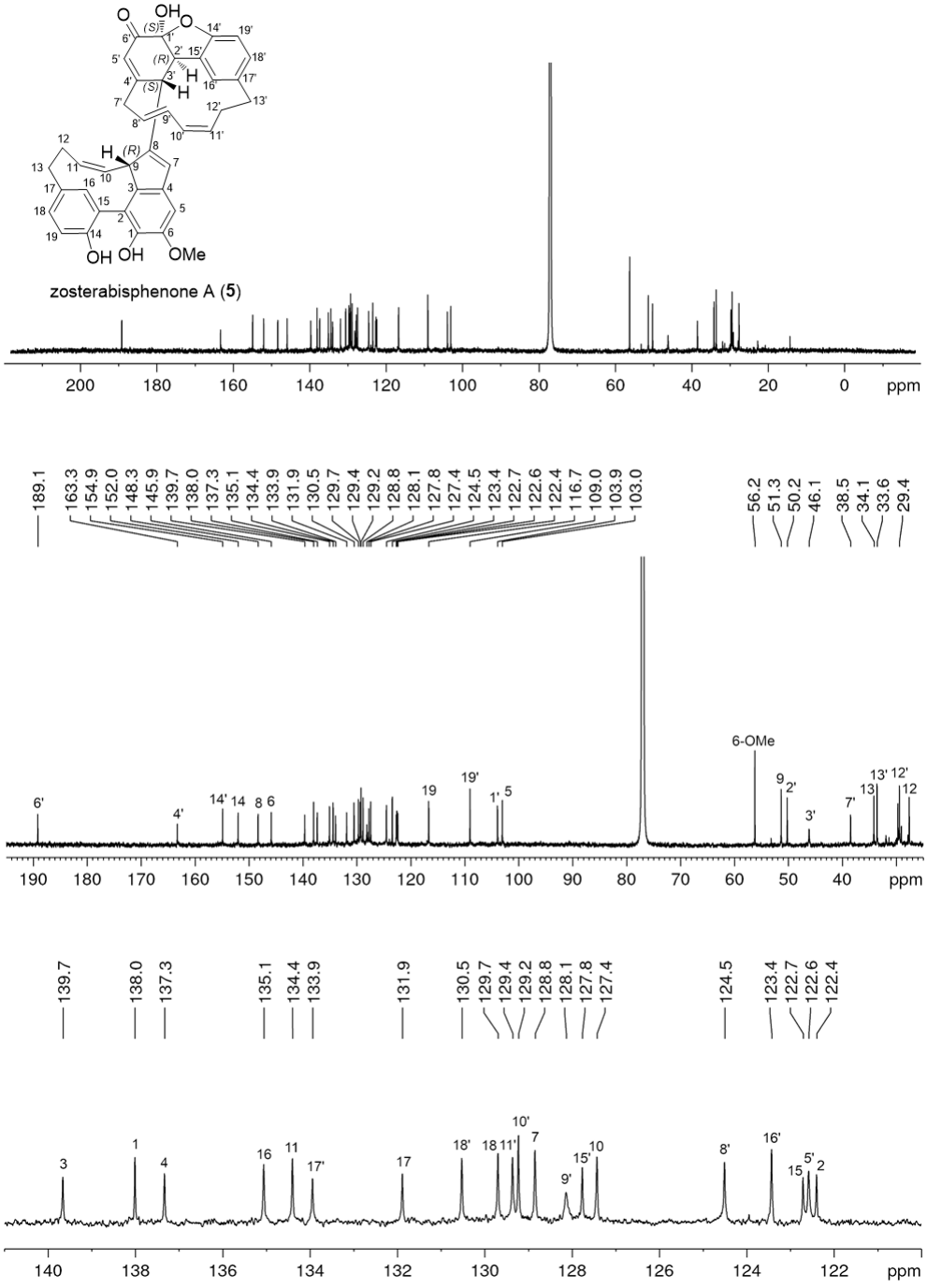


Figure 8.4. ^{13}C NMR spectrum of zosterabisphenone A (**5**) recorded at 253 K (175 MHz, $CDCl_3$)

In-depth analysis of the NMR data revealed a similar structure between one of the diarylheptanoid units ('southern unit') and zosteraphenol A (**3**). It was assumed that C-8 was involved in the linkage of the other diarylheptanoid unit because, unlike zosteraphenol A (**3**), it was not protonated. A further difference was that the methoxy group of compound **5** was located at C-6, and not at C-1 as in zosteraphenol A (**3**).

NMR signals from the second diarylheptanoid unit ('northern' unit) were predictive of a 1,2,4-trisubstituted benzene ring C-14' to C-19') and a hepta-2,4-diene-1,7-diyl chain (C-7' to C-13'). HMBC correlations of H-16' and H-18' with C-13' were indicative of the linkage between C-13' and C-17'. The remaining six carbons in the molecule, which included two sp^3 methine carbons and one carbonyl carbon atom, were indicative of a modification of the second benzene ring of the northern unit. In particular, one of the sp^3 methine carbons at δ 50.2 (C-2') was bound to C-15', as suggested by the HMBC correlation of H-2' with C-14', C-15', and C-16'. In addition, the other sp^3 methine carbon at δ 46.1 (C-3') was shown to be bound to C-2', from the HMBC correlations of H-3' with C-2' and C-15'. The multiplicity of the vicinal protons H-2' and H-3', which resonated as singles, will be rationalized below by the dihedral angle close to 90° between them. HMBC correlations of H-2' with C-4' and of H-3' with C-4' and C-5' suggested that the remaining three sp^2 carbon atoms (δ 189.1, C, C-6'; δ 122.6, CH, C-5'; δ 163.3, C, C-4') were involved in an α,β-unsaturated ketone, linked to C-3' through its β carbon atom. (Figure 8.5). The last carbon atom to be assigned resonating at δ 103.9 (C-1') was a non-protonated carbon. Due to the HMBC correlation of the OH proton at 1' and C-1', C-2', and C-6', C-1’ was

assigned as a hemiacetal carbon atom connecting C-2' with C-6' and defining a cyclohexenone ring. Therefore, zosterabisphenone A (**5**) contains a novel cyclohexenone tautomer of the catechol, in which one carbonyl is involved in a cyclic hemiacetal function with the OH group at the 14' position.

The HMBC data, in particular the correlations of H-7, H-2', H-3 and H-5', helped to define the link between C-4' and C-7' connecting the cyclohexenone ring with the seven-carbon chain, and the C-8/C-3' link connecting the two diarylheptanoid units. Finally, the only possibility of satisfying the 23 unsaturations implied by the molecular formula was the presence of an ether bridge between C-1' and C-14'.

The interpretation of the ROESY correlation allowed the assignment of the relative configuration of the three stereocenters on the northern unit (C-1', C-2', and C-3'). In particular, the ROESY correlation between OH-1' and H-2', indicated their *cis*-relationship, and the ROESY correlations of H-3' with H-9' and H-16', showed H-3' pointing inwards into the macrocycle and thus establishing the relative configuration at C-3'.

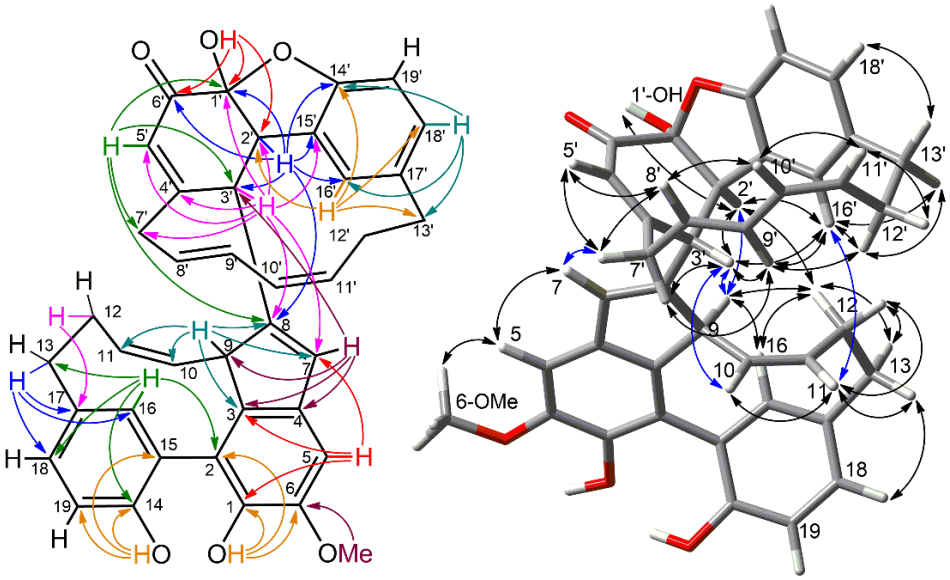


Figure 8.5. Most significant 2D NMR data of zosterabisphenone A (**5**). Left: HMBC correlations. Right: ROESY correlations.

However, the stereochemistry of the last stereocenter, which defines the stereochemical relationship between the northern and southern units, could not be clarified from the NMR data. Consequently, a detailed DFT study was performed using the Gaussian 16 program to provide additional information to the structural features determined from the NMR data. DFT calculations made it possible to clarify the relative configuration between the two diarylheptanoid units and the absolute configuration of the entire molecule.

Table 8.1. ^{1}H and ^{13}C NMR Data of zosterabisphenone A (**5**) (700 MHz, 253 K, $CDCl_3$)

Position	δ_C		δ_H, mult (*J* in Hz)	HMBC	ROESY
1	138.0 (C)		-		
1-OH			6.74, s	1,2,6	-
2	122.4 (C)		-		
3	139.7 (C)		-		
4	137.3 (C)		-		
5	103.0 (CH)		6.80, s	1,3,6,7	7, 6-OMe
6	145.9 (C)		-		
6-OMe	56.2 (CH_3)		3.96, s	6	5
7	128.8 (CH)		6.25, s	3,4,9,3'	5,7'$_{proR}$
8	148.3 (C)		-		
9	51.3 (CH)		5.01, br. d (11.2)	3,7,8,10,11	12$_{proR}$,16,2',3'
10	127.4 (CH)		4.76, t (11.2)	9,12	11,3'
11	134.4 (CH)		5.55, ddd (11.8, 11.2, 4.8)		10,12$_{proS}$,13$_{proR}$,16'
12	27.6 (CH_2)	pro*R*	3.05, dddd (13.6, 12.6, 11.8, 7.0)		9,12$_{proS}$,13$_{proS}$,16,2'
		pro*S*	2.56, overlapped	10,11,17	11,12$_{proR}$
13	34.1 (CH_2)	pro*R*	2.71, ddd (14.4, 12.6, 6.0)	16,17,18	11,18
		pro*S*	3.39, dd (14.4, 7.0)	12,16,17,18	12$_{proR}$,16
14	152.0 (C)		-		
14-OH	-		7.57, s	14,15,19	-
15	122.7 (C)		-		
16	135.1 (CH)		7.68, br. s	2,13,14, 18	9,12$_{proR}$,13$_{proS}$
17	131.9 (C)		-		
18	129.7 (CH)		6.91, br. d (8.0)	13,14,16	13$_{proR}$
19	116.7 (CH)		6.86, d (8.0)	14,15,17	-
1'	103.9 (C)				
1'-OH			5.27	1', 2', 6'	2'
2'	50.2 (CH)		3.85, s	8,1',3',4',6',14',15',16'	9,12$_{proR}$,1'-OH,3',16'
3'	46.1 (CH)		3.56, s	7,8,1',2',4',5',7',15'	7,9,10,2',7'$_{proS}$,9',16'
4'	163.3 (C)		-		
5'	122.6 (CH)		6.18, s	8,1',3',7'	7'$_{proR}$,8'
6'	189.1 (C)		-		
7'	38.5 (CH_2)	pro*R*	2.97, dd (12.4, 7.4)	3',4',5',8',9'	7,5',7'$_{proS}$,8'
		pro*S*	2.33, overlapped		3',7'$_{proR}$,9'
8'	124.5 (CH)		5.08, ddd (15.3, 8.0, 8.0)	10'	5',7'$_{proR}$,10'
9'	128.1 (CH)		5.89, dd (15.3, 10.6)	7',11'	3',7'$_{proS}$,12'$_{proS}$,16'
10'	129.2 (CH)		5.65, t (10.6)	8',9',12'	8',11'
11'	129.4 (CH)		5.22, ddd (11.5, 10.6, 5.3)		10',12'$_{proR}$
12'	29.4 (CH_2)	pro*R*	2.31, overlapped		11',12'$_{proS}$
		pro*S*	2.53, dq (3.3,12.2)		9',12'$_{proR}$,13'$_{proR}$,16'
13'	33.6 (CH_2)	pro*R*	2.98, overlapped		12'$_{proS}$,13'$_{proS}$,16'
		pro*S*	2.34, overlapped		13'$_{proR}$,18'
14'	154.9 (C)		-		
15'	127.8 (C)		-		
16'	123.4 (CH)		6.85, br. s	2',13',14',18'	11,2',3',9', 12'$_{proS}$,13'$_{proR}$
17'	133.9 (C)		-		
18'	130.5 (CH)		6.70, br. d (8.0)	13',14',16'	13'$_{proS}$
19'	109.0 (CH)		6.56, d (8.0)	14',15',17'	-

*8.4 Molecular modeling and Quantum mechanical calculations (DFT) of zosterabisphenone A (**5**)*

The structure of the southern unit was closely related to zosteraphenols A (**3**) and B (**4**), whose favored conformation was known (Chapter 7).

Therefore, a model of the southern unit was produced starting from the lowest energy conformer zosteraphenol B (**4**) and removing the methyl group at O-1, which was followed by DFT optimization.

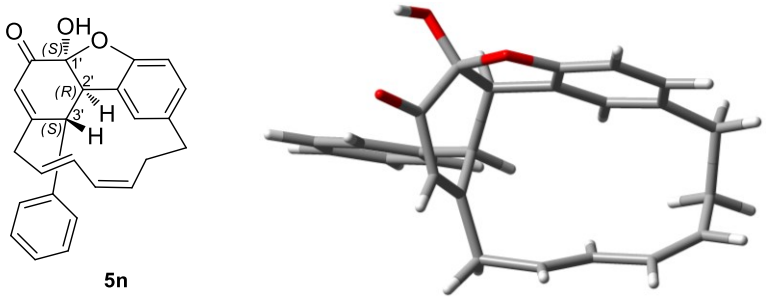


Figure 8.6 The model compound **5n** for the northern diarylheptanoid unit of unit of zosterabisphenone A (**5**) and its lowest-energy conformation at the B3LYP/6-31+G(d,p) level.

A model of the northern unit of zosterabisphenone A (**5**) was then created, replacing the southern unit with a phenyl ring, using the relative configurations at C-1', C-2', and C-3' as established by NMR studies (**5n**, Figure 8.6). The model obtained was subjected to molecular dynamics (MD) conformational search, using the configurations 1'*S*,2'*R*,3'*S*. To perform the MD simulation, the same protocol as for tedarene B was used (Chapter 7), which could reproduce conformational changes with $k < 1\ s^{-1}$ and thus could appropriately explore the entire conformational space of the model compound. The MD simulation was performed at 2000 K, allowing any slow conformational changes to occur within the 10-ns duration of the simulation. The conformational search protocol yielded 200 minimized structures that identified 16 unique conformers for **5n** in a 7.76 kcal/mol range, whose geometries were refined by DFT optimization at the

B3LYP/6-31+G(d,p) level. During DFT optimization, many conformers converged to the same structure, leaving only 7 unique conformers. The geometry of the lowest-energy conformer described well the conformation of the northern unit of zosterabisphenone A (**5**), because it was in perfect agreement with the ROESY data (Table 8.1), while the geometry the second lowest-energy conformer (ΔE = 1.42 kcal/mol) was not. The other 5 conformers were discarded due to high DFT energies (ΔE > 5 kcal/mol). Based on the established conformation of the northern unit and the conformation of the southern unit determined for zosteraphenols, models were generated for two diastereomers of compound **5**, which differ in the relative configuration between the diarylheptanoid units, namely (9*R*,1'*S*,2'*R*,3'*S*)-**5** (called just **5** in the following text) and (9*S*,1'*S*,2'*R*,3'*S*)-**5** (*epi*-**5** in the following text). To clarify the conformation around the rotable bond C-8/C-3', which connected the two diarylheptanoid units, the dihedral angle C-9/C-8/C-3'/C-2' was scanned in steps of 10°, optimizing the produced structures at the B3LYP/6-31G(d) level.

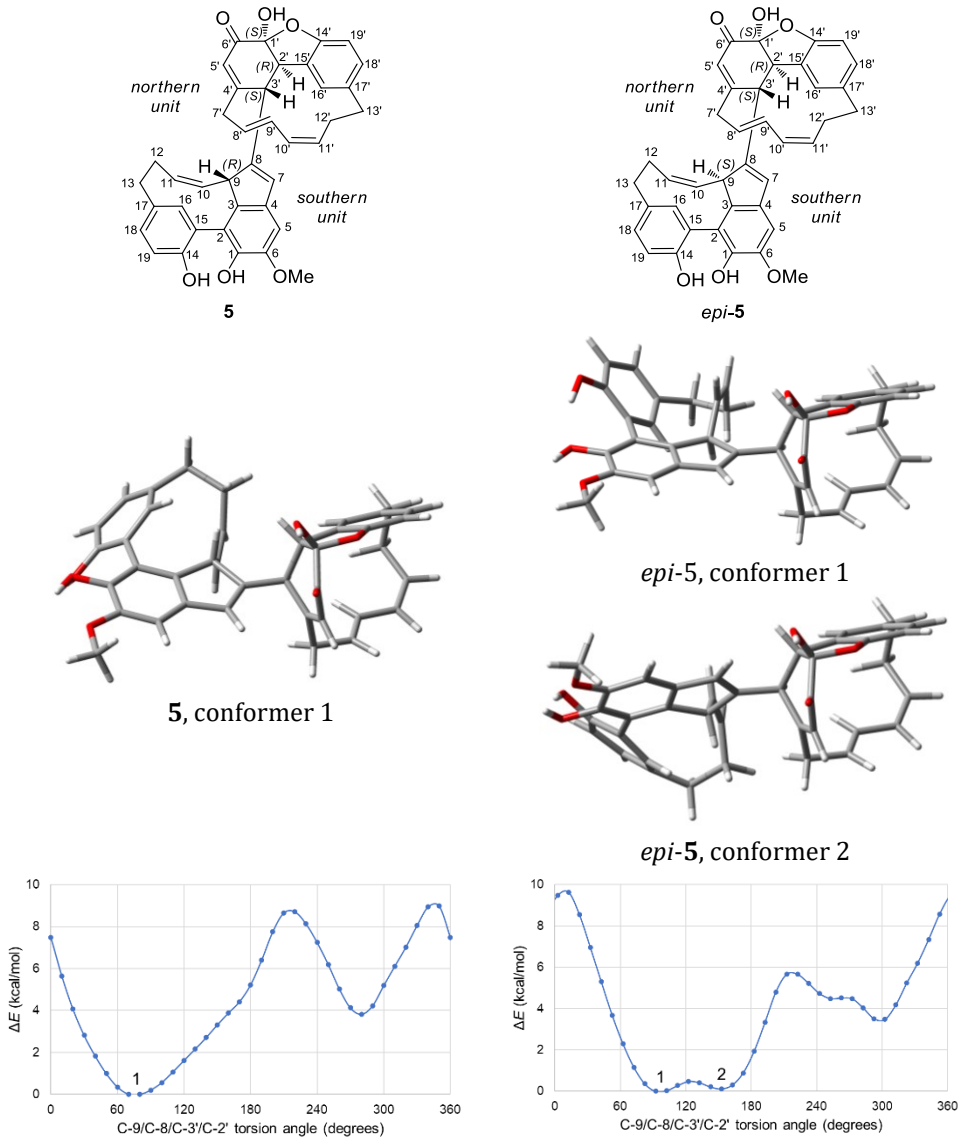


Figure 8.7. DFT energy of zosterabisphenone A (**5**) and its C-9 epimer epi-**5** as a function of the torsion angle about the C-9/C-8/C-3'/C-2' bond, i.e. the bond that connects the two cyclic diarylheptanoid units. The significantly populated minimum energy conformations are shown above the graphs.

Results revealed the existence of two energy minima at 80° and 280° (Figure 8.7). The second one was not significantly populated, being 3.8 kcal/mol higher in energy. In contrast, results obtained for *epi*-**5**, showed three minima, two of them with neighboring low energy minima at 93° (ΔE = 0) and 153° (ΔE = 0.12 kcal/mol), divided by an almost flat potential profile, and one high energy minimum (ΔE = 3.5 kcal/mol). The identified

low energy conformer of **5** and the two low energy conformers of *epi*-**5** were re-optimized at the B3LYP/6-31+G(d,p) level, giving the resulting structures that were used for subsequent NMR and ECD prediction. Prediction of ^{1}H and ^{13}C NMR chemical shifts allowed a safe selection between the alternative diastereomers **5** and *epi*-**5**. The theory level PBE0/6-311+G(2d,p) was used to calculate NMR isotropic shieldings, including the PCM continuous solvent model for chloroform. It is important to note that for *epi*-**5**, Boltzmann-averaged isotropic shieldings over the two conformers were considered. Calculated isotropic shieldings were converted into chemical shifts with the conversion factors proposed by the Tantillo group for this level of theory.[4]

The matching of the experimental values was remarkably better for diastereomer **5** (RMSD of 1.66 ppm for ^{13}C and 0.113 ppm for ^{1}H) than for *epi*-**5** (RMSD of 1.92 ppm for ^{13}C and 0.148 ppm for ^{1}H). Furthermore, some predicted chemical shifts of *epi*-**5** (C-9, H-5', and H-7') exhibited large deviations from the experimental ones. (Figure 8.8). Finally, DP4+ analysis of the predicted chemical shifts showed a 100.00% probability for **5** to be the correct stereoisomer.[5]

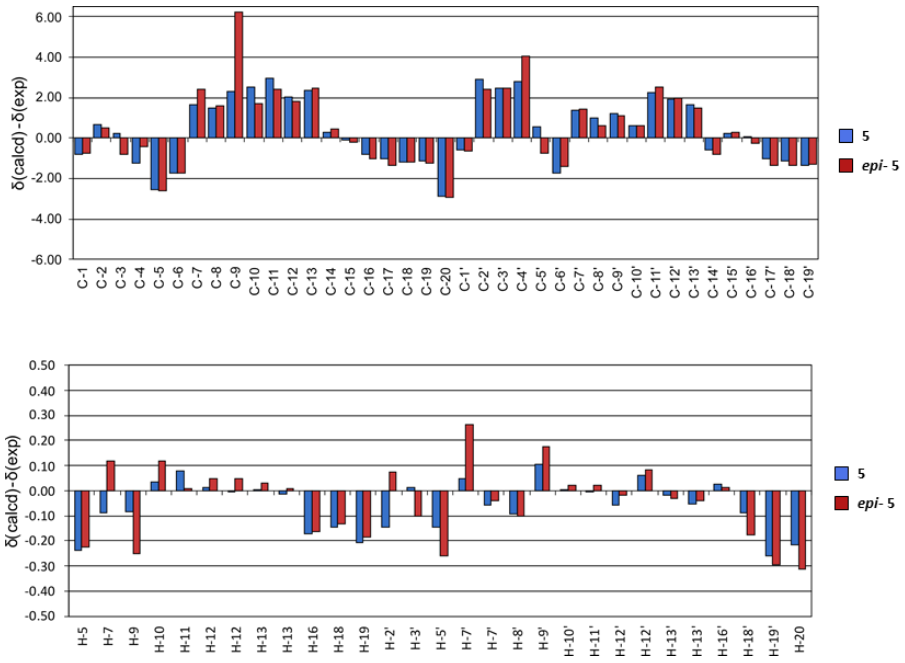


Figure 8.8. Deviations from experimental values of calculated ^{13}C (top panel) and ^{1}H (bottom panel) NMR chemical shifts of zosterabisphenone A (**5**) and its C-9 epimer *epi*-**5**.

Because the accuracy of the method (reported as 2.45 for ^{13}C and 0.15 ppm for ^{1}H) was considerably better than expected, DFT prediction of NMR parameters also provided a strong validation to the unique structure of zosterabisphenone A (**5**). An additional confirmation of the structure, configuration, and conformation of the two diarylheptanoid units of zosterabisphenone A (**5**) was obtained with the DFT prediction of ^{1}H-^{1}H scalar coupling, which were calculated following the suggestions of Bally and Rablen.[6] The experimentally recorded multiplicity of ^{1}H NMR signals were in excellent agreement with the predicted coupling constants. In particular, the predicted coupling between the vicinal protons H-2' and H-3' (1.1 Hz), in turn linked to the –82.5° torsion angle between them, was in accordance with the singlet resonance recorded for the two protons. Finally,

all the observed ROESY cross peaks were assignable to protons which were close in the DFT minimum energy conformation.

The confirmation of absolute configuration was achieved by the comparison of the ECD spectrum of zosterabisphenone A (**5**) calculated at the ωB97XD/6-31+G(d,p) level, using the PCM continuous model for the acetonitrile, solvent used for the experimental measurement.[7] The calculated spectrum was visualized using the SpecDis program,[8] which was also used for the optimization of the half-band width σ (0.51 eV) and the UV correction (+27 nm). The good agreement between the predicted spectrum and the experimental one, defined the (9*R*,1'*S*,2'*R*,3'*S*) configuration for zosterabisphenone A (**5**).

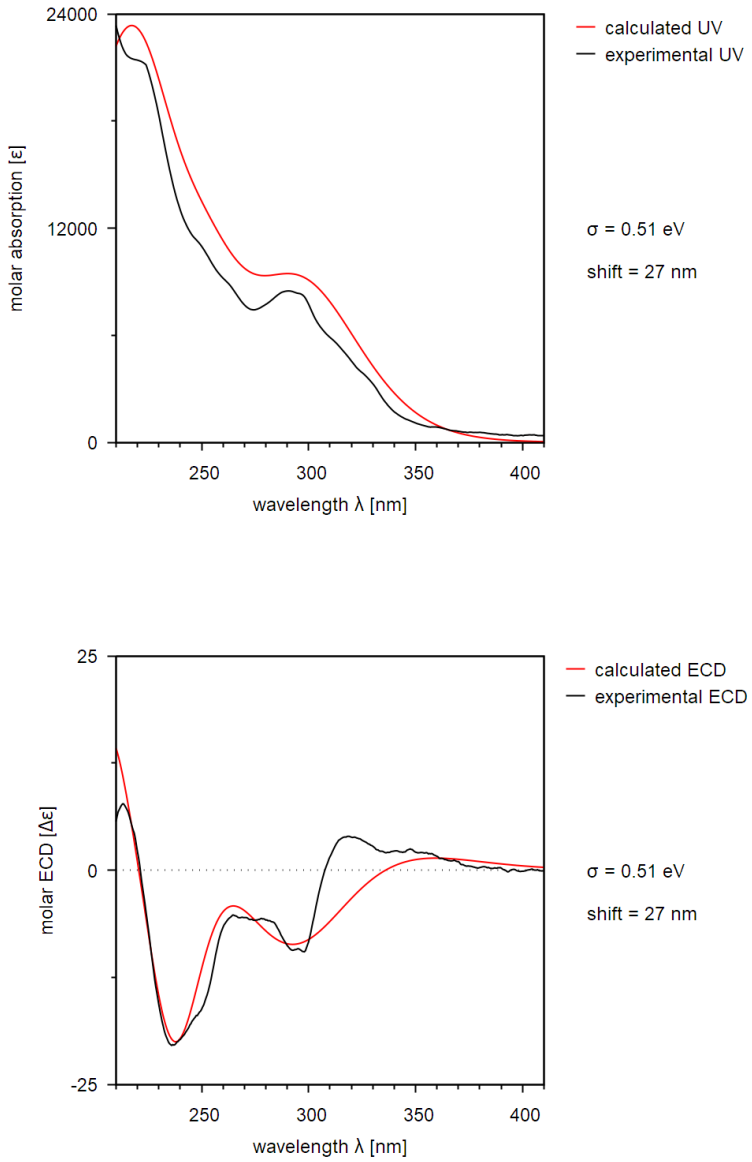


Figure 8.9. Top panel: calculated and experimental UV spectra of zosterabisphenone A (**5**) in acetonitrile; bottom panel: calculated and experimental ECD spectra of zosterabisphenone A (**5**) in acetonitrile.

8.5 Hypothetic biosynthetic pathways of zosterabisphenone A (5)

The presence of the stable enone tautomer rather than the aromatic tautomer **5a**, which has a higher energy than tautomer **5** of 7.95 kcal/mol, is clearly due to steric reasons. In fact, in the catechol tautomer **5a**, a clash between the bulky southern unit at C-4' and the seven-carbon atom bridge results

unless the aromatic rings deviate widely from planarity. As an evidence for this fact, in the DFT optimized structure of **5a** the dihedral angle observed between the ortho bonds at C-2' and C-3' (C-15'/C-2'/C-3'/C-8) was 40°.

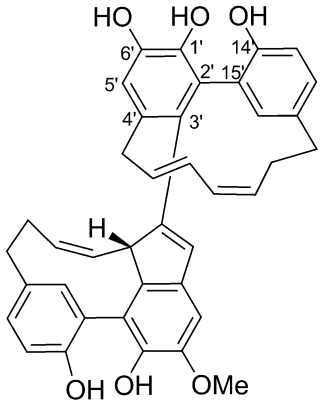


Figure 8.10. Hypothetical aromatic tautomer of zosterabisphenone A **5a**

In contrast, in the enone tautomer **5** the single bond C-3'/C-4' does not place constraints on the C-15'/C-2'/C-3'/C-8 dihedral angle, which is in fact 90°, resulting in a much lower steric strain. Steric strain at the level of the northern diaryleptanoid unit is further reduced by the presence of two atoms with sp^3 hybridization (C-1' and C-2') at the ring junction allowing the planes of cyclohexenone and benzoxolane to be nearly perpendicular.

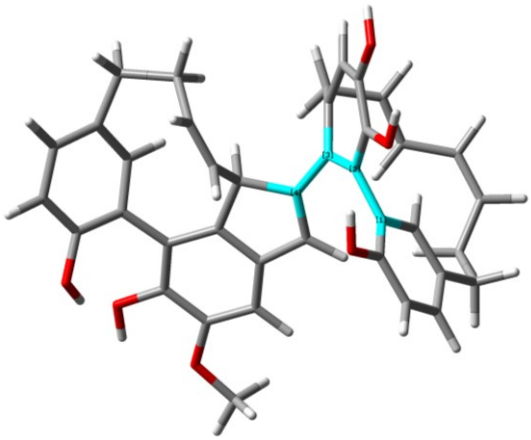


Figure 8.11. Minimum energy conformation of the aromatic tautomer **5a** of zosterabisphenone A. The energy of the aromatic tautomer calculated at the B3LYP/6-31+G(d,p) level was higher by 7.95 kcal/mol than that of the keto tautomer. The two ortho bonds at C-2' and C-3' are pushed out of planarity by steric hindrance up to 40.8° (the value of the dihedral angle C-15'/C-2'/C-3'/C-8, highlighted in the figure).

From a biogenetic point of view, zosterabisphenone A (**5**) could be derived from the reaction of a zosteraphenol A analogue with a free 1-OH and a hydroxylated deoxymodienol analogue. A plausible hypothesis about the mechanism for its biosynthesis involves radical oxidative coupling. This step is followed by tautomerization to enedione and hemiacetal formation, requiring a free OH group at the C-1 level. It is interesting to note that the two monomeric zosteraphenols A and B (**3** and **4**) isolated previously from *Z. marina* extracts both have a methoxy group at the C-1 position, thus making this type of reaction not possible.

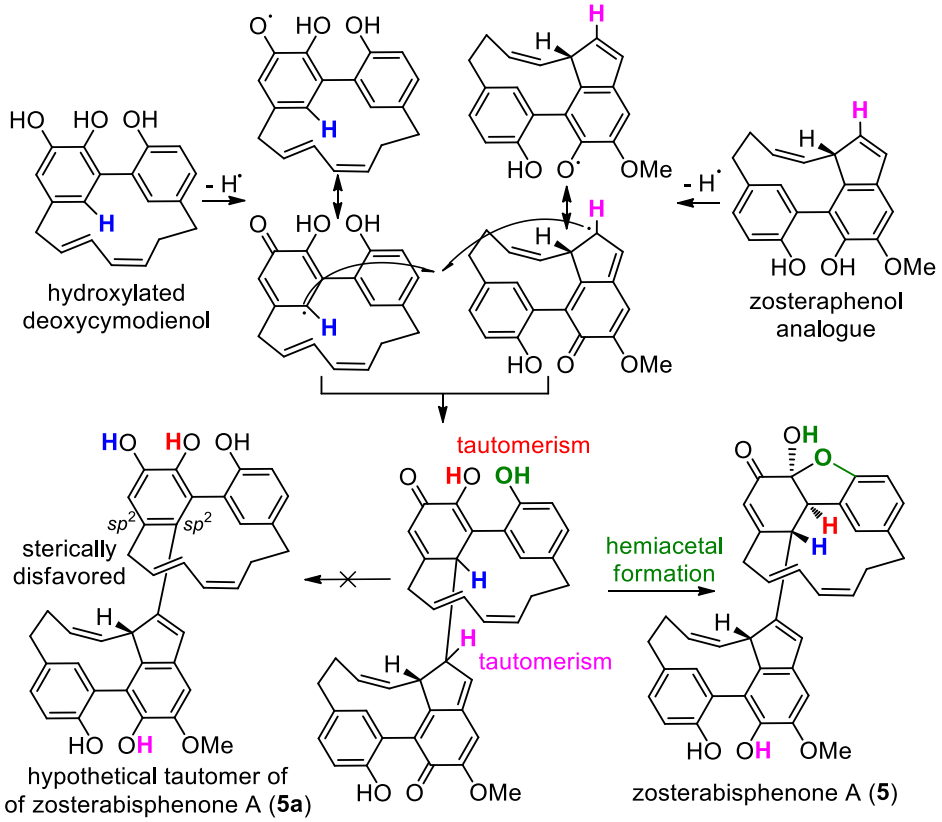


Figure 8.12. Scheme of the putative biosynthetic pathways of zosterabisphenone A (**5**).

*8.6 Structural elucidation of zosterabisphenone B (**6**)*

The molecular formula of zosterabisphenone B was determined to be $C_{40}H_{36}O_8$ (23 unsaturations), established by the $[M+H]^+$ ion at *m*/*z* 645.2471, detected in high-resolution ESI mass spectrum, i. e. with one more carbon atom than zosterabisphenones A (**5**). This data was in perfect accordance with the presence of an additional methoxy group, as confirmed by the ^1H NMR spectrum.

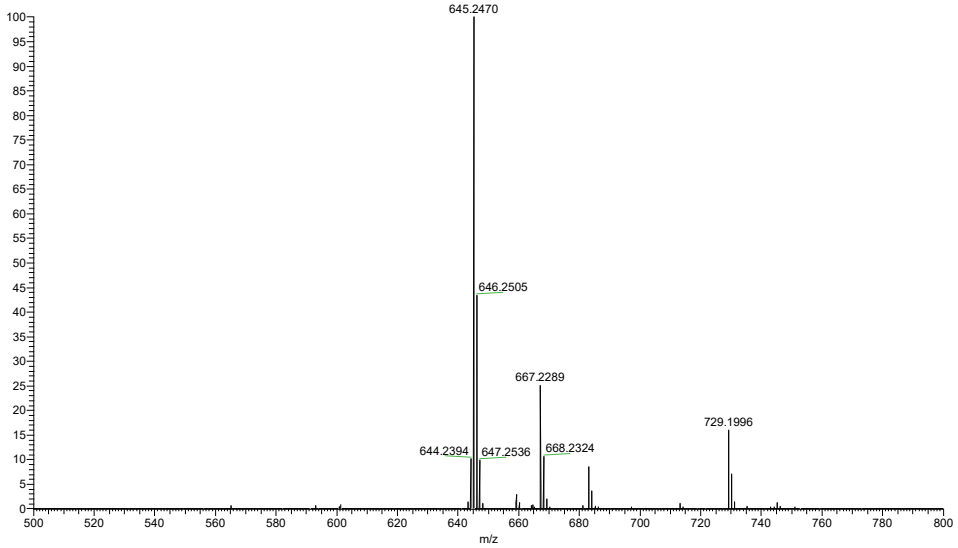


Figure 8.13. High-resolution ESI mass spectrum of zosterabisphenone B (**6**) (positive ion mode, MeOH/0.1% aqueous HCOOH).

The presence of coalescent signals in the ^{1}H and ^{13}C NMR spectra was still evident when the spectra were recorded at 253 K, then NMR spectra were recorded at even lower temperature, 238 K, resulting in sharper signals suitable for the structure elucidation work. The southern diarylheptanoid unit was identical to that present in zosterabisphenone A (**5**), showing similar ^{1}H and ^{13}C chemical shifts; the northern unit contained the same C_7 chain between the benzene rings as in zosterabisphenone A (**5**), as determined from the COSY spectrum.

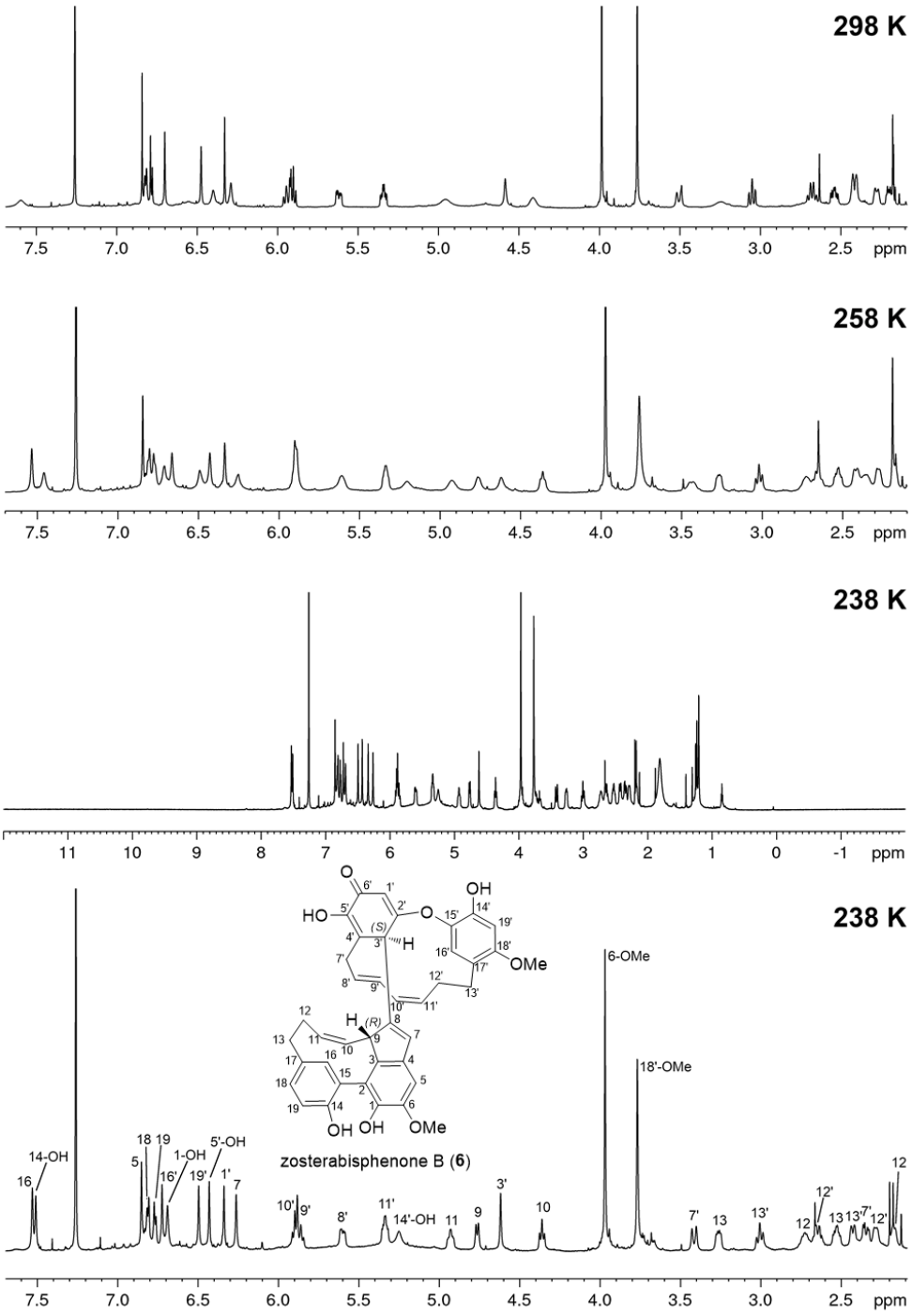


Figure 8.14. ^{1}H NMR spectrum of zosterabisphenone B (**6**) recorded at 298 K, 258 K and 238 K (700 MHz, $CDCl_3$)

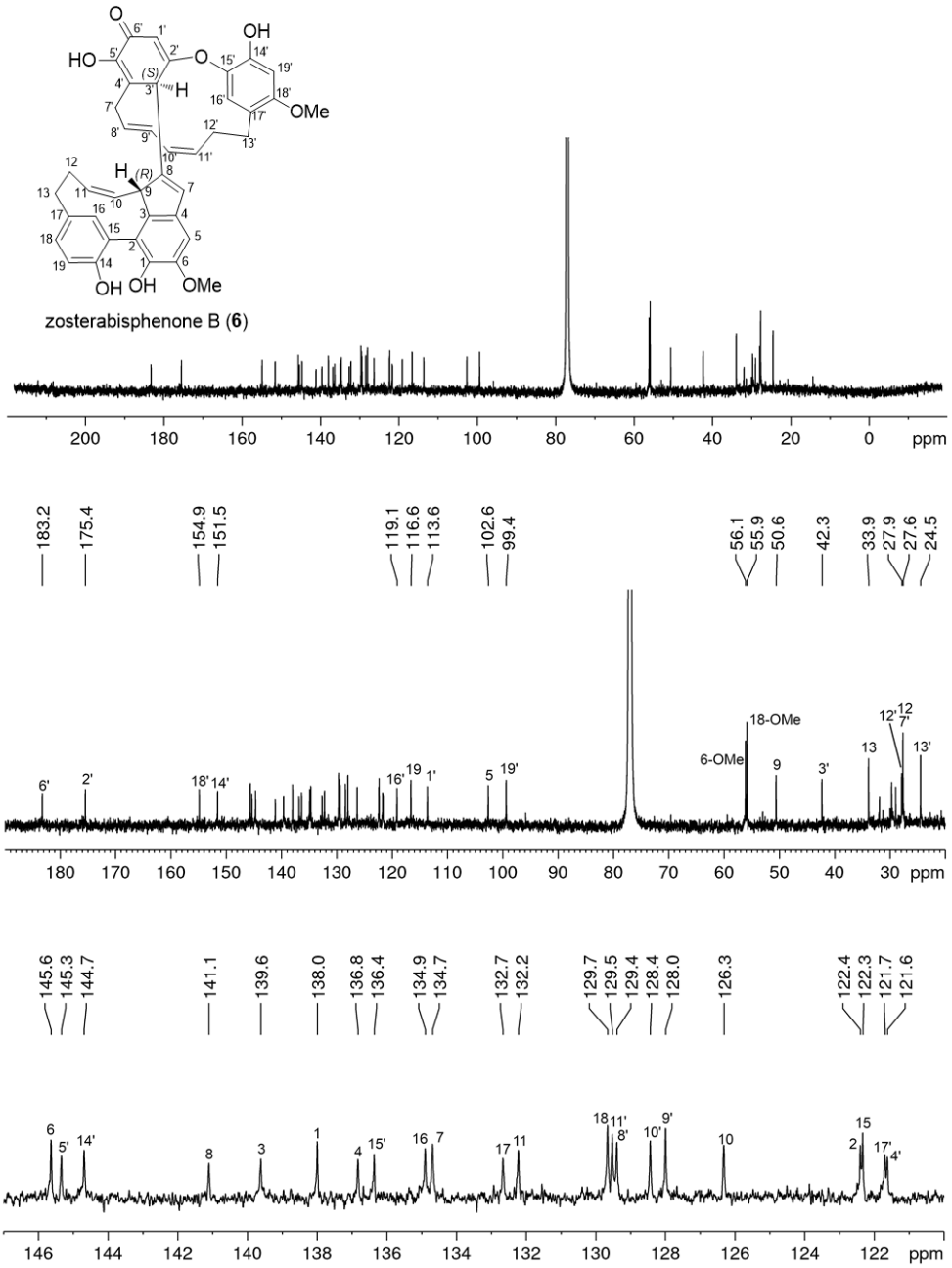


Figure 8.15. ^{13}C NMR spectrum of zosterabisphenone B (**6**) recorded at 298 K, 258 K and 238 K (700 MHz, $CDCl_3$)

However, HMBC correlation peaks analysis revealed that the 1,2,4-trisubstituted benzene ring of zosterabisphenone A (**5**) was replaced by a 1,2,4,5-tetrasubstituted trioxygenated benzene ring. Diagnostic were the HMBC correlations of the two para protons H-16' and H-19' and of protons at C-13' (Figure 8.17). The different magnitude between $^{2}J_{CH}$ (1.0-3.0 Hz) and $^{3}J_{CH}$ (6.5-8.5 Hz) typically observed in aromatic rings was exploited in the assignment of the non-protonated carbon atoms of the ring. The ^{1}H-^{13}C coupling constants of H-16 and H-19, evaluated from the rows of the HMBC spectrum at δ 154.9, 144.7, 136.4, and 121.7 allowed the assignments of C-18', C-14', C-15', and C-17' ^{13}C signals.

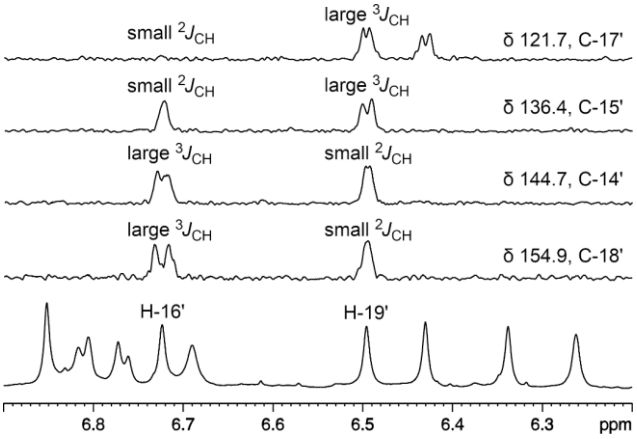


Figure 8.16. Sections of the ^{13}C-coupled HMBC spectrum of zosterabisphenone B (**6**) were used to estimate the magnitude of ^{1}H-^{13}C coupling constants of H-16' and H-19', and therefore to distinguish between the smaller $^{2}J_{CH}$ and the larger $^{3}J_{CH}$.

The second ring in the northern diarylheptanoid unit was determined to be a cross-conjugated cyclohexadienone. The structure was clarified based on HMBC data.

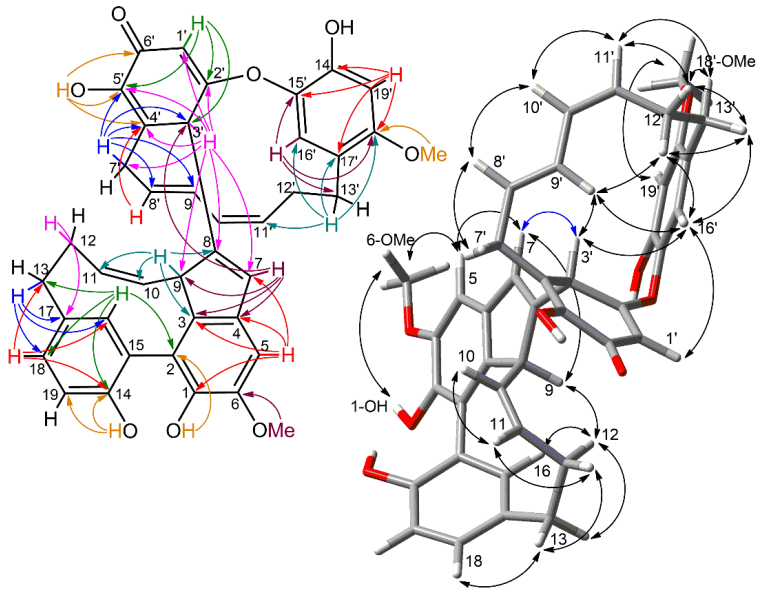


Figure 8.17. Most significant 2D-NMR data of zosterabisphenone B (**6**). Left: HMBC correlations. Right: ROESY correlations.

Thanks to the cross peaks of H-8' with C-4', of H-7' with C-3', C-4', and C-5', C-4' was assigned as the carbon linked to the C_7 chain and both C-5' and the sp^3 carbon atom C-3' (δ 42.3) respectively. Furthermore, the cross peaks of OH-5' with C4', C-5', and C-6' assigned the carbonyl carbon atom C-6' (δ 183.2) defining an α-hydroxyenone system. The second double bond of the molecule was defined by the cross peaks of H-3' with C-1', C-2', C-4', and C-5, with the assignment of C-1' and C-2'. In particular C-2' was strongly deshielded (δ 175.4) for being β to the carbonyl group and linked to O. The closure of the six-membered ring was confirmed by the HMBC correlations of H-1' with C-2', C-3', and C-5'. Finally, the C-8/-C-3' connection between the two diarylheptanoid units was demonstrated by the cross peak of H-3' with C-7, C-8, and C-9. At this point, the three oxygenated carbon atoms C-2', C14', and C-15' remained to be assigned. However the NMR spectrum showed only an additional OH signal, with no HMBC or ROESY

correlation, suggesting the presence of an ether bridge between C-2' and C-14' or C-15'. Due to the presence of a weak H-1' and H-16' NOESY peak, it was possible to localize the ether bridge between C-2' and C-15' because of the weak but clear NOESY peak between H-1' and H-16'. Confirmation of correct assignment was obtained by DFT chemical shift prediction and ^{1}H-^{1}H scalar coupling prediction.

Table 8.2. ^{1}H and ^{13}C NMR Data of zosterabisphenone B (**6**) (700 MHz, 238 K, $CDCl_3$).

Position	δ_C		δ_H, mult (*J* in Hz)	HMBC	ROESY
1	138.0 (C)		-		
1-OH			6.69, s	2	6-OMe
2	122.4 (C)		-		
3	139.6 (C)		-		
4	136.8 (C)		-		
5	102.6 (CH)		6.85, s	1,3,6,7	6-OMe,7
6	145.6 (C)		-		
6-OMe	56.1 (CH_3)		3.97, s	6	1-OH,5
7	134.7 (CH)		6.26, s	3,4,9,3'	5,3'
8	141.1 (C)		-		
9	50.6 (CH)		4.76, br. d (11)	3,8,10,11	$12_{pro\mathit{R}}$,16
10	126.3 (CH)		4.36, t (11)	12	11
11	132.2 (CH)		4.93, ddd (11, 11, 5)	-	$10,12_{pro\mathit{S}}$,
12	27.6 (CH_2)	pro*R*	2.72, m	-	$9,12_{pro\mathit{S}},13_{pro\mathit{S}},16$
		pro*S*	2.17, m	17	$11,12_{pro\mathit{R}},13_{pro\mathit{R}}$
13	33.9 (CH_2)	pro*R*	2.53, ddd (13, 13, 6)	16,17,18	$12_{pro\mathit{S}},13_{pro\mathit{S}},18$
		pro*S*	3.26, dd (13, 6)	11,17	$12_{pro\mathit{R}},13_{pro\mathit{R}}$
14	151.5 (C)		-		
14-OH	-		7.51, s	14,19	-
15	122.3 (C)		-		
16	134.9 (CH)		7.53, br. s	2,13,14,18	$9,12_{pro\mathit{R}}$
17	132.7 (C)		-		
18	129.7 (CH)		6.81, br. d (8)	13,14,16	$13_{pro\mathit{R}}$
19	116.6 (CH)		6.77, d (8)	14,15,17	
1'	113.6 (CH)		6.34, s	2',3',5'	16'
2'	175.4 (C)		-		
3'	42.3 (CH)		4.62, s	7,8,9,1',2',4',5',7'	7,9',16'
4'	121.6 (C)		-		
5'	145.3 (C)		-		
5'-OH			6.43, s	4',5',6'	-
6'	183.2 (C)		-		
7'	27.6(CH_2)	pro*R*	2.35, dd (19, 6)	4',5',8',9'	$7'_{pro\mathit{S}}$,8'
		pro*S*	3.41, br. d (19)	3',4',5',8',9'	$7'_{pro\mathit{R}}$
8'	129.4 (CH)		5.60, br. dd (15, 6)	4',10'	$7'_{pro\mathit{R}}$,10'
9'	128.0 (CH)		5.86, m	-	$3',12'_{pro\mathit{S}}$,16'
10'	128.4 (CH)		5.90, m	8',12'	8',11'
11'	129.5 (CH)		5.34, ddd (11, 11, 5)	-	$10',12'_{pro\mathit{R}},13'_{pro\mathit{S}}$
12'	27.9 (CH_2)	pro*R*	2.28, m	-	$11',13'_{pro\mathit{R}},12'_{pro\mathit{S}}$
		pro*S*	2.65, br. quartet (13)	-	$9',12'_{pro\mathit{R}},13'_{pro\mathit{R}},16'$
13'	24.5 (CH_2)	pro*R*	2.43, br. d (14)	11',17'	$12'_{pro\mathit{R}},12'_{pro\mathit{S}},13'_{pro\mathit{S}},16'$
		pro*S*	3.01, br. t (14)	11',16',17',18'	$11',12'_{pro\mathit{R}},13'_{pro\mathit{R}}$
14'	144.7 (C)		-		
14'-OH			5.25, s	-	
15'	136.4 (C)		-		
16'	119.1 (CH)		6.72, s	13',14',15',18'	$1',3',9',12'_{pro\mathit{S}},13'_{pro\mathit{R}}$
17'	121.7 (C)		-		
18'	154.9 (C)		-		
18'-OMe			3.77, s	18'	19'
19'	99.4 (CH)		6.50, s	14',15',17',18'	18'-OMe

*8.7 Molecular modeling and quantum mechanical DFT calculations of zosterabisphenone B (**6**)*

Similar to zosterabisphenol A (**5**), to confirm the relative configuration of the two diarylheptanoid units of zosterabisphenone B (**6**), molecular modeling and a detailed DFT study were performed with a protocol similar

to that used for zosterabisphenone A (**5**). First, a simplified model of the northern diarylheptanoid unit was made (**6n**), replacing the southern unit with a phenyl ring. The model was subjected to high-temperature MD conformational search, using the 3'*S* enantiomer, which produced 29 conformers in a range of 11.54 kcal/mol. Their geometry was then optimized with DFT at the B3LYP/6-31+G(d,p) level. The conformer with the lowest energy was in perfect agreement with the data obtained from the ROESY spectrum, but the second lowest energy conformer was not, and was discarded. The remaining conformers were discarded due to their very high DFT energy.

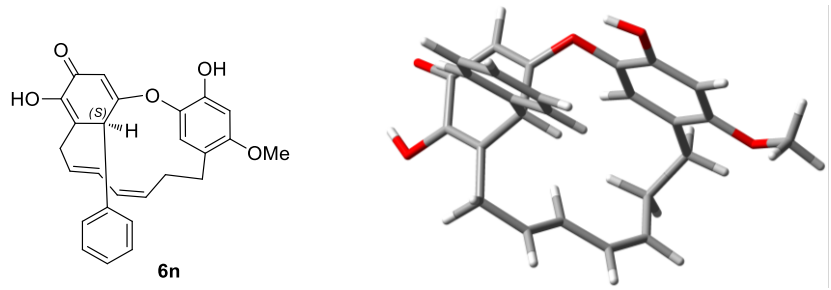


Figure 8.18. The simplified model compound **6n** for the northern diarylheptanoid unit of unit of zosterabisphenone B (**6**) and its lowest-energy conformation at the B3LYP/6-31+G(d,p) level.

The generation of the two possible diastereomers of zosterabisphenone B (**6**) was achieved by replacing the phenyl group of **6n** with both enantiomers of the southern unit, to give **6** (the 9*R*,3'*S* stereoisomer) and *epi*-**6** (the 9*S*,3'*S* stereoisomer). At this point, the torsion angle around the C-8/C-3' bond was scanned in the same manner used for **5**/*epi*-**5**. As result, the graph of potential energy vs. dihedral angle (Figure 8.19) showed two low energy conformers for **6** and two low energy conformers for *epi*-**6**.

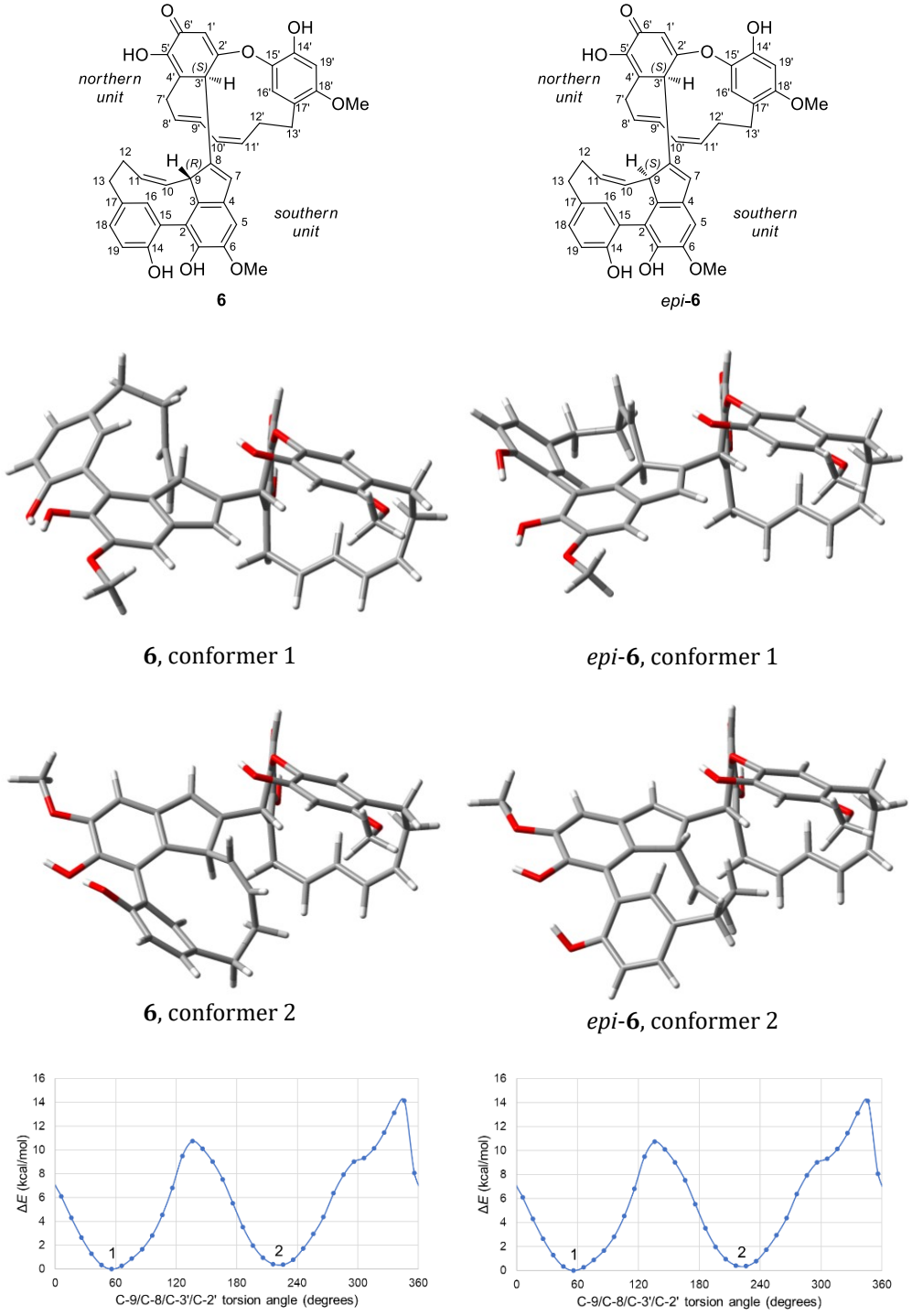


Figure 8.19. DFT energy of zosterabisphenone B (**6**) and its C-9 epimer epi-**6** as a function of the torsion angle about the C-9/C-8/C-3'/C-2' bond, i.e. the bond that connects the two cyclic diarylheptanoid units. The significantly populated minimum energy conformations are shown above the graphs**.**

Once re-optimized at the B3LYP/6-31+G(d,p) level, the resulting geometries were used to perform NMR data prediction and ECD prediction.

The PBE0/6-311+G(2d,p)/PCM level was used to calculate the isotropic NMR shielding, which were averaged over the two conformers of each compound using the Boltzmann statistics. The obtained data were converted into chemical shifts using the conversion factors discussed above. Although the calculated ^{13}C chemical shifts were accurately similar for both conformers (RMSD of 2.05 ppm for **6** and 2.08 for *epi*-**6**), the accuracy of ^{1}H chemical shifts was better for **6** (RMSD of 0.126 ppm for **6** and 0.149 for epi-**6**) (Figure 8.20). In addition, further confirmation was provided by DP4+ analysis, providing 100.00% probability of **6** being the correct stereoisomer.

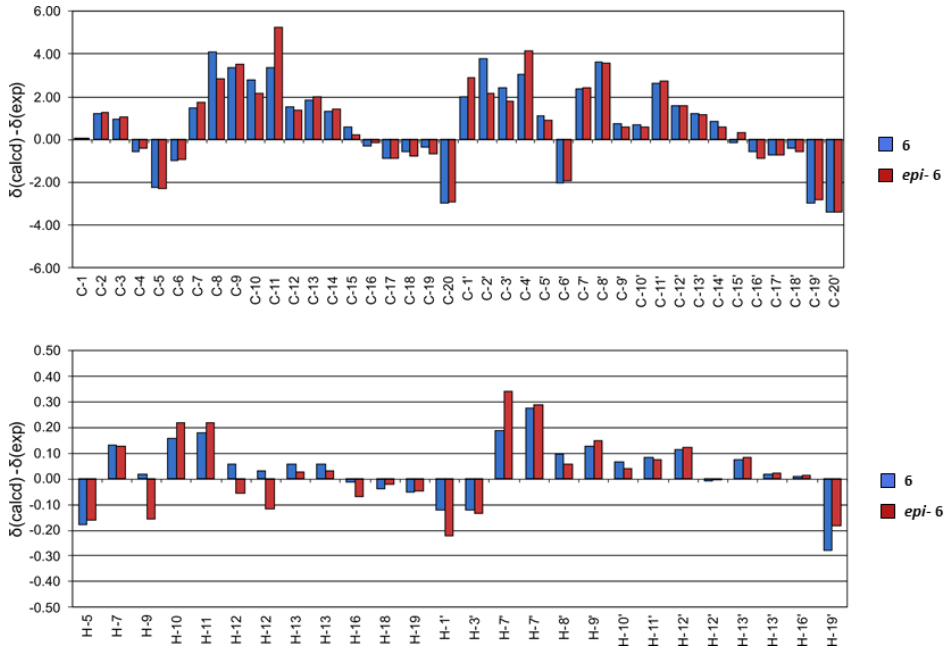


Figure 8.20. Deviations from experimental values of calculated ^{13}C (top panel) and ^{1}H (bottom panel) NMR chemical shifts of zosterabisphenone B (**6**) and its C-9 epimer epi-**6**.

To confirm the absolute configuration of zosterabisphenone B (**6**), the ECD spectrum was generated at the ωB97XD/6-31+G(d,p) level using the

continuous PCM model for acetonitrile, and the mid-band width σ = 0.30 eV and UV correction = +9 nm in the SpecDis program.[9] The results obtained showed good coincidence between the predicted ECD spectrum and the experimental spectrum (Figure 8.21), thus defining the (9R,3'S) configuration for zosterabisphenone B (**6**).

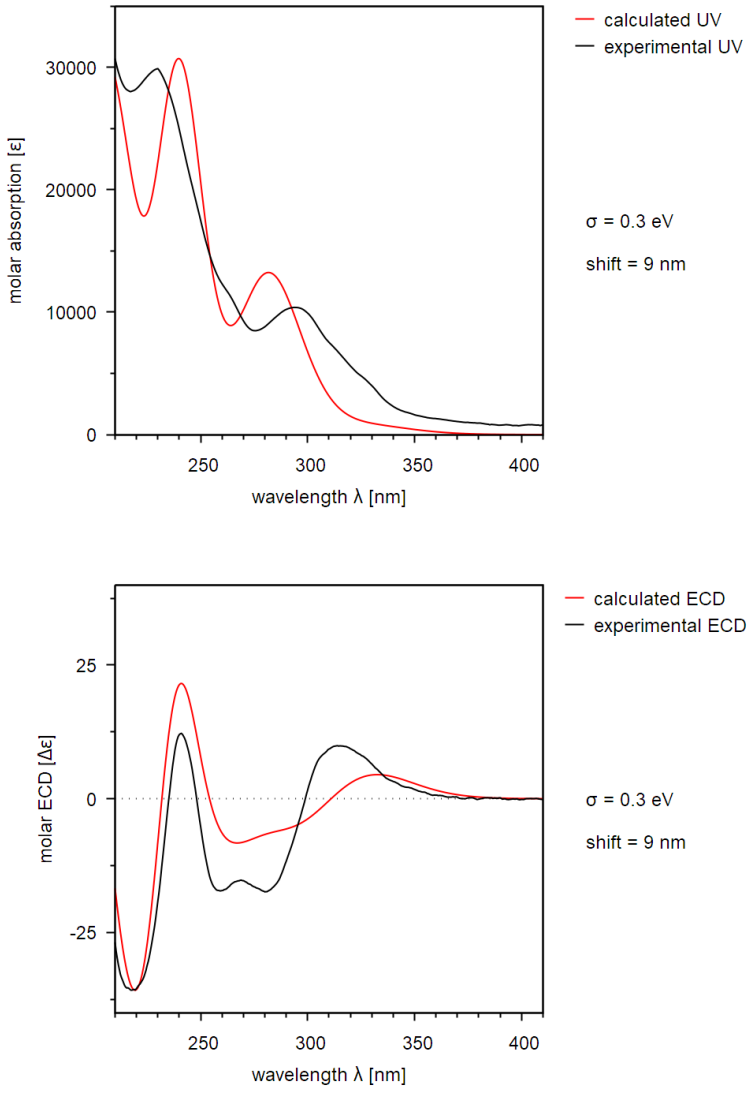


Figure 8.21. Top panel: calculated and experimental UV spectra of zosterabisphenone B (**6**) in acetonitrile; bottom panel: calculated and experimental ECD spectra of zosterabisphenone B (**6**) in acetonitrile. Zosterabisphenone B (**6**) constitutes a stable keto tautomer of a catechol and, similarly to zosterabisphenone A, its stability can be attributed to steric interactions of the hypothetical aromatic tautomer 6a (Table S1).

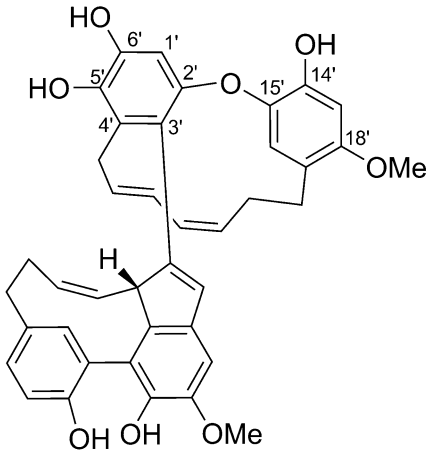


Figure 8.22. Hypothetical aromatic tautomer of zosterabisphenone B (**6**).

Particularly interesting is the presence of an unusual *meta,meta* ether bridge in the northern unit, which deviates from the *meta,para* coupling prevalent in almost all known cyclic ether-bridged diarylheptanoids, including the same isotedarene A also isolated in *Z. marina*.[2]

In literature, it constitutes the second example of a *meta,meta* ether-bridged cyclic diarylheptanoid.[10] Given the absence of an OH substituent at C-1', unlike the other diarylheptanoids, in which the para carbon at the C7 chain is oxygenated, it is very difficult to hypothesize a biosynthetic mechanism for zosterabisphenone B (**6**) or the structure of the monomeric precursor of its northern unit.

*8.8 Evaluation of cytotoxicity of zosterabisphenone A (**5**) and B (**6**)*

The ability of zosterabisphenones to act as antitumor agents was tested by evaluating their effects on cell viability of two cell lines (HCT116 and Hep G2 cells) by MTT assay.[11] It was possible to verify that zosterabisphenone B (**6**) showed a cytotoxic activity. More precisely, the ability to reduce in a concentration- and time-dependent manner the viability rate of HCT116

cells up to 48 hours of exposure, was verified reaching 97.4% inhibition at 10 µM (IC50 3.9±1.2 µM at 24 hours, 3.6±1.1 µM at 48 hours). Given the ability to inhibit the viability of Hep G2 cells only at the highest concentration tested, zosterabisphenone cytotoxicity resulted to be selective on HCT116 cells compared with Hep G2 cells (Figure 8.23, Table S17). On the other hand, zosterabisphenone A (**5**) exhibited a weak cytotoxicity on HCT116 only at the highest concentration and period tested. No effect was recorded on Hep G2.

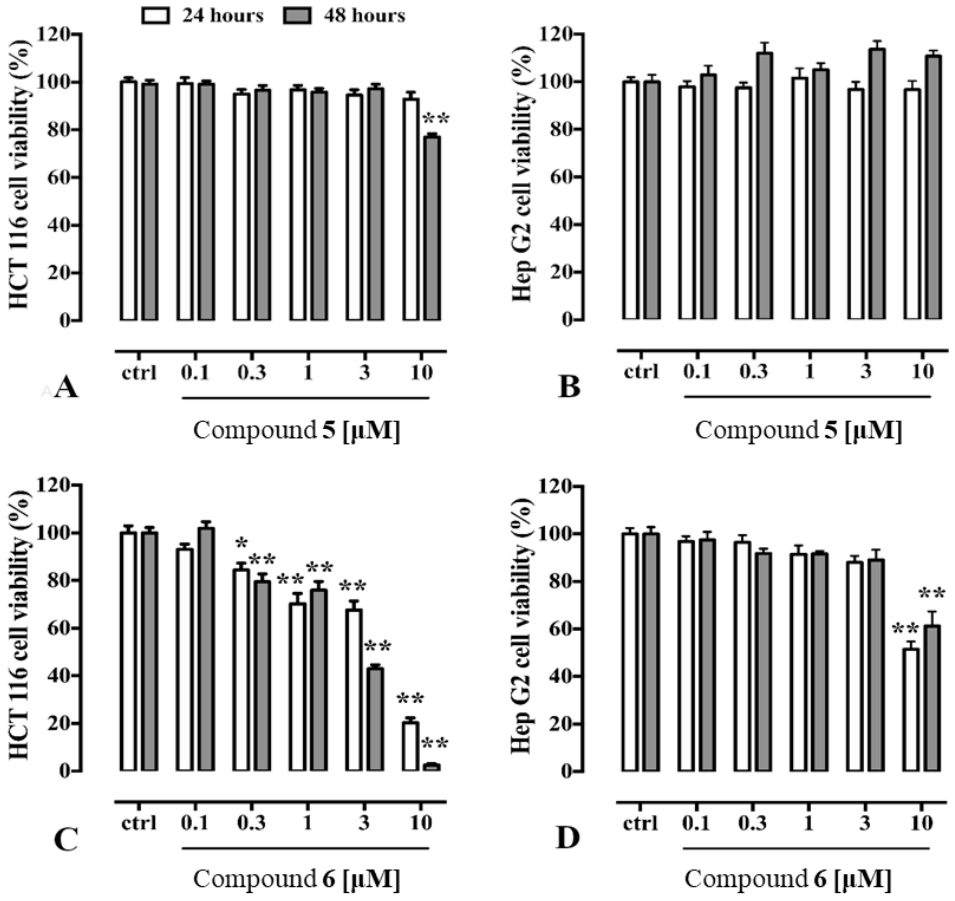


Figure 8.23. Cytotoxic effects of zosterabisphenone A (**5**) (panels A, B) and B (**6**) (panels C, D) on HCT116 and Hep G2 cells (0.1–10 µg/mL at 24-h and 48-h exposure). Cell viability rate is expressed as percentage. Each bar represents the mean ± SEM of three independent experiments (including 5–6 replicates for each treatment). * $p < 0.01$ and ** $p < 0.0001$ vs. control (ctrl, i.e. untreated cells).

Table 8.3. Effects of zosterabisphenone A (**5**) on the viability of HCT116 and Hep G2 cells after 24 hours and 48 hours of treatment. The results are expressed as percentage of cell viability. Each value represents the mean ± SEM of 3 experiments including 5–6 replicates for each treatment. **P<0.0001.

Concentration of zosterabisphenone A (**5**) [μM]	HCT116		Hep G2	
	24 h	48 h	24 h	48h
Ctrl	100.0±1.53	100.0±1.53	100.0±1.92	100.0±2.95
0.1	102.0±2.10	97.6±1.47	97.9±2.43	103.0±4.00
0.3	98.6±2.08	95.2±1.83	97.5±2.13	112.0±4.49
1	98.9±2.07	96.0±1.59	102.0±4.16	105.0±2.75
3	98.5±2.25	95.3±2.06	96.7±3.25	114.0±3.56
10	97.6±2.27	77.3±1.41 **	96.7±3.68	111.0±2.35

Table 8.4. Effects of zosterabisphenone B (**6**) on the viability of HCT116 and Hep G2 cells after 24 hours and 48 hours of treatment. The results are expressed as percentage of cell viability. Each value represents the mean ± SEM of 3 experiments including 5–6 replicates for each treatment. *P<0.01; **P<0.0001.

Concentration of zosterabisphenone B (**6**) [μM]	HCT116		Hep G2	
	24 h	48 h	24 h	48h
Ctrl	100.0±2.90	100.0±2.96	100.0±2.52	100.0±2.87
0.1	93.1±2.17	102.0±2.75	96.8±2.27	97.4±3.38
0.3	84.4±2.94 *	79.4±3.30 **	96.5±2.96	91.7±2.04
1	70.2±4.36 **	76.0±3.69 **	91.4±3.79	91.6±3.69
3	67.6±3.75 **	42.9±1.65 **	88.0±2.82	89.0±4.33
10	20.3±2.10 **	2.6±0.65 **	51.4±3.31 **	61.2±6.23 **

8.9 Experimental section

8.9.1 General methods

A Jasco P-200 polarimeter at 589 nm, was used to perform optical rotations using a 10 cm cell, while a Jasco V-530 spectrophotometer and a Jasco 715 spectropolarimeter, were used to record UV and ECD spectra, respectively. High-resolution ESI mass spectra were measured on a Thermo LTQ Orbitrap XL mass spectrometer coupled to a Thermo Ultimate 3000 UPLC system. NMR spectra were recorded on a 700 MHz Bruker Avance Neo spectrometer equipped with a cold probe and a BCU-II variable temperature

unit. Chemical shifts were referenced to solvent peaks at δ_H 7.26 and δ_C 77.0 for $CDCl_3$.

8.9.2 Extraction and Isolation

Whole plants of *Zostera marina L.*, were sampled along the shore close to the Olympiazentrum Schilksee, Kiel, Schleswig-Holstein, Germany in December 2018 at the following coordinates N 54°25'39.0", E 10°10'17.5"; alt.: 0 m. The collected plants were unrooted plants freshly washed ashore. A voucher specimen is preserved in the Herbarium of the Institute of Botany, Kiel University (voucher code: YL-20181222A-1; KIEL0005004). Air-dried, ground whole plants of *Z. marina* (1.75 kg) were extracted five times with 5.5 L acetone each at room temperature. Following evaporation of the solvent in vacuum, the crude extract weighing 18.3 g was subjected to silica gel column chromatography, using a 7.5 × 50-cm column, and eluted in sequence with hexane/dichloromethane (7:3), hexane/dichloromethane (5:5), dichloromethane/acetone (7:3), dichloromethane/acetone (5:5), and acetone (1 L of each mixture) to yield 24 subfractions. A medium-pressure silica gel column was used to purify the N fraction weighing 1.25 g, containing zosterabisphenones A (**5**) and B (**6**). A PrepChrom C-700 system (Büchi, Essen, Germany) equipped with a silica gel column (Sepacore Silica 40-63 μm, 80 g, 194 × 31 mm) and a flow rate of 20.0 ml/min was used. The employed linear gradients were: 0 min, hexane; 30 min, hexane/CH_2Cl_2 (8:2); 80 min, hexane/CH_2Cl_2 (1:1); 81 min, CH_2Cl_2; 120 min, CH_2Cl_2/MeOH (8:2), with a run time set to 2 hours, yielding 15 fractions. From the 15 fractions, Fraction N10 (107 mg, eluted with 70% hexane and 30% CH_2Cl_2) containing compound **5** and Fraction N11 (114 mg, eluted

with 68% hexane and 32% CH_2Cl_2) containing compound **6** were selected. Fractions N10 and N11 were further purified by Sephadex LH-20 column chromatography (2 × 100 cm) using CH_2Cl_2 /acetone (85:15) as eluents. Partially purified compounds 5 (23.5 mg) and 6 (21.2 mg) were finally purified by semi-preparative reversed-phase HPLC using a Waters e2695 instrument and a Nucleodur C8 column (1 × 25 cm), flow rate 2 ml/min, UV detection at 210 and 254 nm and mixtures of MeOH and 0.025% formic acid in water as the mobile phase in isocratic separation. Zosterabisphenone A **5** (4.6 mg) was eluted isocratically with 75% MeOH (Rt 40-46 min) while pure zosterabisphenone B **6** (4.2 mg) was eluted with 70% MeOH (Rt 40-46 min).

Zosterabisphenone A (**5**): $[\alpha]_D^{25}$ –140 (*c* 0.01, ACN); UV/Vis (ACN): λ_{max} (ε) 291 (8500), 196 nm (29300); ECD (ACN): λ_{max} (Δε) 347 (+2.3), 319 (+3.9), 296 (–9.5), 237 (–20.4), 214 (+7.6), 202 nm (–8.1); HRMS (ESI/Orbitrap) *m/z*: $[M+Na]^+$ calcd for $C_{39}H_{34}O_6Na^+$ 621.2248, found 621.2235; $[M+H-H_2O]^+$ calcd for $C_{39}H_{33}O_5^+$ 581.2323, found 581.2311; $[M+H]^+$ calcd for $C_{39}H_{35}O_6^+$ 599.2428, found 599.2418; $[M+NH_4]^+$ calcd for $C_{39}H_{38}O_6N^+$ 616.2694, found 616.2683; $[M+K]^+$ calcd for $C_{39}H_{34}O_6K^+$ 637.1987, found 637.1974; $[M+HCOO+Ca]^+$ calcd for $C_{40}H_{35}O_8Ca^+$ 683.1952, found 683.1940.

*Zosterabisphenone B (**6**):* $[\alpha]_D^{25}$ –55 (*c* 0.01, ACN); UV/Vis (ACN): λ_{max} (ε) 294 (10400), 230 (30000), 200 nm (40700); ECD (ACN): λ_{max} (Δε) 314 (+9.9), 280 (–17.4), 259 (–17.2), 240 (+12.2), 218 (–35.8), 192 nm (+26.2);;

HRMS (ESI/Orbitrap) *m/z*: $[M+H]^+$calcd for $C_{40}H_{37}O_8^+$ 645.2483, found 645.2470; $[M+Na]^+$ calcd for $C_{40}H_{36}O_8Na^+$ 667.2302, found 667.2289; $[M+K]^+$ calcd for $C_{40}H_{36}O_8K^+$ 683.2042, found 683.2029; $[M+HCOO+Ca]^+$ calcd for $C_{41}H_{37}O_{10}Ca^+$ 729.2007, found 729.1996.

8.9.3 General computational methods

Conformational search was performed using molecular dynamics (MD) with the INSIGHT II/Discover package (BIOVIA, 5005 Wateridge Vista Drive, San Diego, CA 92121, USA). All the MD simulations were performed at 2000 K to allow possible slow conformational changes to occur in the short duration of the simulation, constraining the geometry of double bonds to prevent *cis/trans* isomerization. The effect of the solvent (chloroform) was approximated by using a dielectric constant of 4.81. The search protocol involved a 10-ns MD simulation in the CFF91 force field. The coordinates were saved every 50 ps and subsequently minimized in the same force field, giving 200 minimized structures, which were used as input for the subsequent quantum-mechanical calculations. Density functional theory (DFT) calculations were performed using the program Gaussian 16 (Revision C.01, Gaussian Inc., Wallingford CT, USA), using the B3LYP/6-31+G(d,p) level of theory for structure optimization, the Gauge Invariant Atomic Orbitals (GIAO[12] method at the PBE0/6-311+G(2d,p) level of theory and the PCM solvent model for chemical NMR shift prediction, and the time-dependent DFT (TDDFT) method at the ωB97XD/6-31+G(d,p) level of theory and the PCM solvent model for ECD prediction. Proton-proton NMR scalar couplings were calculated according to the suggestions

of Bally and Rablen: calculations were performed at the B3LYP/6-31G(d,p) level of theory in vacuum.[6] Only the Fermi contact terms were calculated, which were then scaled by a factor of 0.9117. The NMR isotopic shielding computed for each H or C nucleus was converted into a chemical shift using the scaling factors proposed by the Tantillo group for the level of theory used (^{1}H: slope -1.0958, intercept 31.7532; ^{13}C: slope -1.0533, intercept: 187.3123).[4] When more than one conformer were significantly populated, the weighted mean of the chemical shifts and scalar couplings of individual conformers was calculated using Boltzmann statistics (T = 253 KThe predicted ECD curves were obtained using the program SpecDis v. 1.71,[8] adjusting the parameters σ and UV shift for the best fit between the predicted and experimental spectra. When more than one conformer were significantly populated, the weighted mean of the ECD curves of individual conformers was calculated using Boltzmann statistics (T = 298 K).

8.9.4 Bioactivity assays

Cytotoxic activity of zosterabisphenone A (**5**) and B (**6**) was evaluated on the human adenocarcinoma colon cancer cell line (HCT116) and the human liver colon cancer cell line (Hep G2).[11] Both cell lines were purchased from ATCC (Manassas, VA, USA). HCT116 and Hep G2 cells were cultured in McCoy's 5A medium and Minimal Essential Medium (Euroclone, MI, Italy), respectively. Both cell lines were supplemented with 10% Fetal Bovine Serum (FBS) plus 2mM glutamine and 100UI penicillin and 100 μg/ml streptomycin. Cells were maintained under standard cell culture conditions at 37 °C in a humidified atmosphere of 5% CO_2 in air. To perform the cell viability assay (MTT assay), compounds **5** and **6** were

solubilized in pure EtOH (Sigma Aldrich, St Louis, MO, USAFinal EtOH concentration in cell culture medium never exceeded 0.1% (v/v) and equal amounts of the solvent were added to control cells (untreated cells). Cell viability was assessed by using the tetrazolium dye [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide] (MTT). Briefly, HCT116 (5 × 103) cells or Hep G2 cells (1 × 104) cells were seeded in culture media which contained 10% FBS on 96-well plates and left to adhere for 24 h. After that, the cells were incubated with increasing concentrations (0-10 μM) of zosterabisphenones in cell culture medium containing 1% FBS for 24 or 48 h. Control cells received only the vehicle (EtOH). After the indicated incubation period times, MTT (0.25 mg/mL) was added into each well according to the manufacturer's instructions for 1 h at 37 °C. Thereafter, 100 μL of DMSO was added into each well and the absorbance (OD) was measured at 570 nm using a microplate reader (BioTek™ Cytation™ 3, Winooski, VT, USA). Statistical analysis, performed using GraphPad Prism 7.0 (GraphPad software, San Diego, CA, USA), was determined by two-way analysis of variance (ANOVA) followed by a Turkey-Kramer multiple comparisons test. The concentration of compounds that produced 50% inhibition of cell viability (EC_{50}) was calculated by nonlinear regression analysis using the equation for a sigmoid concentration-response curve. $P < 0.05$ was considered significant. Data were expressed as the mean ± mean standard error (SEM) of n experiments.

8.10 Computational and spectroscopic data

Table 8.4. Cartesian coordinates of the lowest-energy conformer of zosterabisphenone A (**5**) optimized at the B3LYP/6-31+G(d,p) level.

conformer 1, *E* = -1957.9933839 Harthrees >99% population			
C	-0.8302212481	3.3360981514	-4.0016888195
C	-1.30728812	2.3620698545	-3.1098754281
C	-0.3618412221	1.8354251534	-2.1939006135
C	0.9610354032	2.3153920842	-2.1761052344
C	1.4121077453	3.2964924499	-3.0675897145
C	0.4995897203	3.7974194683	-3.9885354095
C	1.7134152178	1.6338681012	-1.1269619797
C	0.9333525279	0.7229967886	-0.5072900982
C	-0.4656212651	0.713373843	-1.1394505371
C	-1.5753078221	0.7976196798	-0.1089810056
C	-2.5567786924	-0.0919583558	0.1260603994
C	-2.8885096942	-1.3543031146	-0.6383490534
C	-4.1625601176	-1.2210392268	-1.5449773523
C	-3.8401633089	2.8223617612	-2.9937235948
C	-2.736669641	1.9363635752	-3.0673465709
C	-2.9872567506	0.583605587	-2.8116820278
C	-4.1797096687	0.1114639926	-2.2676540562
C	-5.2441042999	1.0154903715	-2.1688286537
C	-5.0890663156	2.3376310687	-2.5836657951
C	2.0823108716	5.2741341821	-5.0412678128
C	2.946134361	-1.9941237492	-0.4071699148
C	1.5051443859	-1.7045399236	0.0970207019
C	1.2952693647	-0.2423477928	0.6047965125
C	2.4830301555	0.2727400028	1.3979511879
C	3.7436745419	-0.0054586338	0.9992634367
C	3.9911190801	-0.8859778053	-0.1411397972
C	2.1854246081	1.1336156518	2.6150424402
C	2.1969723958	0.2628233564	3.8566655382
C	1.1383008362	-0.414390669	4.3386391115
C	1.2177943048	-1.3388888613	5.4659006745
C	0.3372329933	-2.3125228365	5.7757646621
C	-0.8936445204	-2.7095044925	5.0074279518
C	-0.6897527985	-4.0040769546	4.1613939109
C	2.4998347364	-3.4592012754	1.3259220207
C	1.3270451728	-2.7188328012	1.2011596401
C	0.2702117571	-2.9357029457	2.0748731967
C	0.3979322434	-3.8751209944	3.1139284869
C	1.5865112449	-4.6158526375	3.1998326232
C	2.6514638823	-4.431489465	2.3069806892
O	-1.6919054212	3.8746946588	-4.9365905276
O	0.7595419817	4.750893764	-4.9453879211
O	-3.7506490712	4.165541115	-3.2171110521
O	4.9901828143	-0.8443826084	-0.8519116253
O	2.9503572847	-2.3566526499	-1.7371651263
O	3.4609866563	-3.0999164414	0.4203216506
H	2.4381559585	3.6439308274	-3.0355852994
H	2.755062404	1.8368220509	-0.9058994971
H	-0.5514133344	-0.2419445131	-1.6733012747
H	-1.5399681635	1.6896308119	0.5163755884
H	-3.2565307977	0.1585337791	0.924466513
H	-2.042028508	-1.6589700344	-1.2615978843
H	-3.0554682815	-2.1792928764	0.0663867405
H	-4.1783167953	-2.076940719	-2.2325256861
H	-5.0635590565	-1.3013994265	-0.9259906435
H	-2.1576148591	-0.1035855516	-2.9164146366
H	-6.1919518943	0.6972120027	-1.7407096868
H	-5.9118876879	3.0432751988	-2.5281387976
H	2.0615193184	5.9915192353	-5.8619892799
H	2.3697442783	5.784544589	-4.1147050787
H	2.8024537129	4.4787818343	-5.265385414
H	0.8236967727	-1.8930627774	-0.7384771892
H	0.4302761673	-0.2836530893	1.276701934
H	4.6107583148	0.4542516191	1.4676009932
H	1.2108703414	1.6181047781	2.4934142515
H	2.9386709973	1.9231786449	2.7084614586
H	3.1721186672	0.1135058287	4.3200267237
H	0.1739149639	-0.2965593386	3.8473241025
H	2.1082505346	-1.2451086534	6.0870002137
H	0.5707409838	-2.9375167349	6.6377431528
H	-1.7199375441	-2.8920880032	5.7071869209
H	-1.2202158907	-1.9015790444	4.3447519334
H	-1.6502482764	-4.2490322736	3.6886467182
H	-0.4521586737	-4.837128577	4.8331993118
H	-0.643234342	-2.3520478342	1.9809984836
H	1.696035624	-5.345768879	3.9979206706
H	3.5692810868	-5.003273964	2.3916858959
H	-1.1903493486	4.5149287443	-5.4680558512
H	-3.0523380667	4.3341753813	-3.8766830573
H	3.8607410849	-2.2192694623	-2.0590284434

Table 8.5. Cartesian coordinates of the lowest-energy conformers of the C-9 epimer of zosterabisphenone A (*epi*-**5**) optimized at the B3LYP/6-31+G(d,p) level.

	Conformer 1, *E* = -1957.9916158 Harthrees Δ*E* = 0.00 kcal/mol, population 77.6%				Conformer 2, -1957.9906194 Harthrees Δ*E* = 0.63 kcal/mol, population 22.4%		
C	4.9624260509	0.7058031051	-0.4374890323	C	-1.0095253365	3.7796838506	-3.5480806728
C	3.8963971201	-0.1947465935	-0.2829552745	C	-1.3934263001	2.815341615	-2.60191899
C	2.593449951	0.3624903501	-0.3325663153	C	-0.3463694195	2.0502604632	-2.025276921
C	2.4143216762	1.7490545181	-0.4985771317	C	0.9829917579	2.2323734233	-2.4509851966
C	3.4934424062	2.6301440795	-0.641855566	C	1.3405118692	3.1870869888	-3.4107991062
C	4.7735121886	2.0899528009	-0.6122171115	C	0.3261901685	3.9696772882	-3.9480587334
C	0.9888394319	2.0609576603	-0.5010117424	C	1.854783574	1.2968557893	-1.750215629
C	0.262275536	0.9340082806	-0.3442100087	C	1.1437430853	0.5615779061	-0.8701547533
C	1.1943764366	-0.2843106243	-0.2810820754	C	-0.3343514984	0.9834487657	-0.9050538771
C	0.8797996476	-1.231224747	0.8611096513	C	-1.2714248933	-0.2026613055	-1.0404664438
C	0.5133587421	-2.5233828913	0.7884443674	C	-2.2639847535	-0.5845885434	-0.2160068959
C	0.4292361209	-3.4144423104	-0.4308398234	C	-2.8015108512	0.1150833687	1.013206563
C	1.6040481278	-4.4512478797	-0.527937033	C	-4.1978119198	0.7912013342	0.7900047499
C	4.9453054435	-2.1864517517	0.9752982777	C	-3.8469620083	2.3235952992	-3.2189986194
C	4.0990243601	-1.6505870741	-0.0270049084	C	-2.8224773197	2.5495187657	-2.2655326164
C	3.1972043646	-2.5300408272	-0.6365051801	C	-3.1137994276	2.2391234511	-0.9331027158
C	2.9269593368	-3.8100257335	-0.1579004567	C	-4.2396059573	1.5196489811	-0.5383724202
C	3.7518916495	-4.303126344	0.8601345127	C	-5.2210585045	1.2768841991	-1.5064745375
C	4.7838728581	-3.5196783015	1.3750425834	C	-5.0483202965	1.7267425346	-2.8145475414
C	5.880359173	4.2065600737	-0.9099709287	C	1.7976386718	5.2444332122	-5.3631019697
C	-2.1744149319	2.3541624378	1.5830834071	C	4.2840665151	-0.5668957281	-0.197617462
C	-1.8437805229	0.8981714754	1.1503998457	C	2.9242472543	-1.2553055847	-0.5134817848
C	-1.2481602779	0.7783579743	-0.289454259	C	1.6769199235	-0.5192778238	0.0565903171
C	-1.9051141727	1.7332116623	-1.2712674103	C	1.943210195	0.0504762739	1.452420721
C	-2.2319505624	2.9903166759	-0.8973632364	C	3.1268913511	0.6331635075	1.7442741323
C	-2.0570536682	3.4333584736	0.4835387445	C	4.2264882259	0.6266303294	0.7847616924
C	-2.1367349924	1.220960384	-2.6836817546	C	0.8310275157	-0.0575198373	2.4826724551
C	-3.5517673764	0.6854815001	-2.7916375904	C	1.0524980644	-1.296712844	3.328751489
C	-3.9327737697	-0.5693212502	-2.4877102568	C	0.5775754359	-2.522472912	3.0391677546
C	-5.3239301031	-1.0123945016	-2.4850358774	C	0.8939403877	-3.7173433208	3.816426595
C	-5.8254472073	-2.0981128402	-1.8610867195	C	0.7817440303	-4.9977476467	3.4072540234
C	-5.0878192884	-3.0701276375	-0.9813233465	C	0.371865262	-5.4867723044	2.0449825563
C	-5.3507116313	-2.8351852368	0.538481465	C	1.5846244009	-5.9443619709	1.1762349376
C	-4.1395042203	1.1364054609	1.6597901138	C	4.3588338778	-2.6951562725	0.6953516187
C	-3.1828250537	0.2074134934	1.258830867	C	3.1017513898	-2.6245082292	0.1013717611
C	-3.5602318027	-1.0946444613	0.9602896861	C	2.2275458183	-3.6986975627	0.1961778179
C	-4.9156955694	-1.4655999593	1.0193935391	C	2.594231424	-4.8418785083	0.9293451014
C	-5.8506252233	-0.5095096163	1.4439681683	C	3.8754050899	-4.8856686118	1.4993727583
C	-5.4822671349	0.8002643584	1.7816798695	C	4.7829446851	-3.8234057447	1.3858225096
H	3.3264369012	3.6938024732	-0.7653353705	H	2.3740561775	3.3042758827	-3.7147740885
H	0.5955784963	3.0627473909	-0.6316739566	H	2.9212241563	1.2203381574	-1.9326987543
H	1.0316810724	-0.8157646775	-1.2293510075	H	-0.5404031842	1.4940444555	0.0453881561
H	0.9548075912	-0.7826696757	1.8508586566	H	-1.0816137495	-0.8199603019	-1.918479028
H	0.3247141429	-3.0266439467	1.7373277011	H	-2.81047061	-1.4838256385	-0.5032471874
H	0.4129369926	-2.8133590489	-1.3458248442	H	-2.0978183915	0.8780867872	1.3598110224
H	-0.5166770886	-3.971880426	-0.4171774266	H	-2.8973485949	-0.6061185913	1.8355087362
H	1.6039403097	-4.8689729742	-1.5435006711	H	-4.3936777549	1.4523756475	1.6446725657
H	1.4113859702	-5.2878151227	0.1534338013	H	-4.9829037515	0.0263902903	0.8015164141
H	2.583290063	-2.1355172451	-1.436093346	H	-2.3503327547	2.4543109759	-0.1967162677
H	3.5733559647	-5.2922105577	1.2758252513	H	-6.1125481037	0.709402189	-1.2491479141
H	5.4402368976	-3.8971120222	2.1526749162	H	-5.8084707371	1.5588115438	-3.570842054
H	6.9125923412	4.5491051949	-0.9859687659	H	1.68757068	6.0573188815	-6.081265937
H	5.3382416926	4.4664202923	-1.8265708704	H	2.452886256	5.5628227671	-4.5441571143
H	5.3981274346	4.6797439365	-0.0469209614	H	2.2270331852	4.3682616013	-5.8624734522
H	-1.1208029971	0.4983266991	1.8659564713	H	2.8268709735	-1.3133299708	-1.6022088578
H	-1.4680098095	-0.2444684162	-0.6137180869	H	0.8869553806	-1.2717818058	0.1566778868
H	-2.5754879382	3.7315131799	-1.6150476498	H	3.2915127269	1.158838395	2.6816571503
H	-1.4090215385	0.435626943	-2.9155088993	H	-0.1380502532	-0.1188204128	1.9814819072
H	-1.9897457799	2.0310586758	-3.4056936352	H	0.8253195756	0.832508127	3.1214558649
H	-4.3165017083	1.4204656603	-3.0420980058	H	1.7235228804	-1.1768589473	4.1791002851
H	-3.1721104203	-1.2912264323	-2.1950177228	H	-0.0553874678	-2.642587091	2.1615945714
H	-6.0228240347	-0.3639322598	-3.0127300824	H	1.2967061946	-3.5367012343	4.8127141912
H	-6.9007224625	-2.2605622478	-1.9355655973	H	1.0929294524	-5.7727172085	4.10775918
H	-5.402242276	-4.0942387155	-1.2226667859	H	-0.3088166482	-6.3419161727	2.1499277707
H	-4.0094009699	-3.0244636425	-1.164628881	H	-0.180339232	-4.713911685	1.5006301108
H	-4.8241804663	-3.6225632204	1.0941574206	H	1.1901285146	-6.3248879139	0.224820033
H	-6.4192291618	-2.9716007817	0.7420123039	H	2.0796662038	-6.7890989375	1.6692073722
H	-2.811974296	-1.819441875	0.6468427756	H	1.2426599704	-3.644577948	-0.2634070636
H	-6.900862113	-0.7854532866	1.4953002429	H	4.1702297118	-5.7649705285	2.0664035085
H	-6.2164394239	1.5332134054	2.0983827598	H	5.7670286267	-3.8653532484	1.8400066913
H	6.8546104466	0.9840938731	-0.5440728502	H	-1.5276584683	5.1871090414	-4.7394400464
H	6.2133775435	-0.7488178973	1.0747992003	H	-3.1150769121	3.33533524	-4.6761953503
H	-1.4764888114	3.6944728704	2.7370241989	H	5.621594446	0.4495017396	-1.0883303618
O	6.2576055177	0.2277553589	-0.4194167542	O	-1.9751559184	4.582151378	-4.124780172
O	5.9489444331	2.7921208775	-0.7426761868	O	0.4838712052	4.964889689	-4.8843047888
O	5.8732731553	-1.4542520469	1.6558571857	O	-3.7114694567	2.5733003773	-4.5532769625
O	-1.8856380328	4.5995361391	0.8259051712	O	5.1327658232	1.4526548337	0.7469046267
O	-1.4584122855	2.7200610654	2.7026089687	O	4.9454338694	-0.2010089278	-1.3556298475
O	-3.6193434649	2.3832043001	1.8764442731	O	5.0872464913	-1.5450160329	0.5481465362

Table 8.6. Cartesian coordinates of the lowest-energy conformers of zosterabisphenone B (**6**) optimized at the B3LYP/6-31+G(d,p) level.

	Conformer 1, E = -2147.7357258 Harthrees ΔE = 0.00 kcal/mol, population 85.4 %				Conformer 2, E = -2147.7342802 Harthrees ΔE = 0.91 kcal/mol, population 14.6%		
C	-3.9540693072	2.6117500189	0.2511530279	C	-5.033172493	0.2404203709	0.1983802419
C	-3.6353556242	1.266071911	0.0043509724	C	-3.8404870028	-0.355976338	-0.2420827363
C	-2.2767255582	0.8977939612	0.1870112582	C	-2.6591969717	0.4142309946	-0.0862159231
C	-1.3409017107	1.8495062528	0.6360815255	C	-2.7118674573	1.6827239284	0.5231104436
C	-1.6834389561	3.1831886585	0.889383472	C	-3.9114502339	2.2518543687	0.9689011338
C	-3.0054838653	3.5567138814	0.6841050472	C	-5.0767822216	1.5163488956	0.7910963514
C	-0.0326698056	1.218827088	0.7589083499	C	-1.3687746722	2.2460906078	0.5901169628
C	-0.0961813855	-0.0786462313	0.3961713963	C	-0.479576379	1.4066562571	0.0219714577
C	-1.5199101726	-0.432613426	-0.0592798011	C	-1.1967036079	0.1598769134	-0.515260854
C	-2.0854758944	-1.6942931868	0.5693751753	C	-0.5368181998	-1.1372087024	-0.0883162738
C	-2.5238103348	-2.8036354712	-0.0497521513	C	-0.0231783457	-2.1002091645	-0.8736284191
C	-2.643932661	-3.0827006113	-1.5310454336	C	-0.0556130288	-2.2164822866	-2.3802793979
C	-4.1240919156	-3.0880295762	-2.048285	C	-1.0647786776	-3.2968007895	-2.90388704
C	-5.8869510287	0.05021895	0.3121134042	C	-4.3578135285	-2.8799413517	-0.1594608411
C	-4.6676354053	0.2599858415	-0.3788386347	C	-3.7898403346	-1.743620291	-0.7877502266
C	-4.2907564523	-0.7080121008	-1.3153155575	C	-2.9002602916	-1.9769750632	-1.8419794063
C	-4.9121462479	-1.9493333294	-1.4318907553	C	-2.3742591663	-3.2312225062	-2.1429389943
C	-6.0972303375	-2.1450259145	-0.7131819451	C	-2.9259646922	-4.3350060505	-1.4819725143
C	-6.60493922	-1.1338372289	0.1017341645	C	-3.9451966595	-4.1615797085	-0.5465751101
C	-2.6500376864	5.8597450591	1.28489233	C	-6.5096835196	3.1930530546	1.7549331863
C	1.1437435716	-2.8060339479	-1.3996249638	C	1.8762420702	3.5196631452	1.298659068
C	1.3308830707	-1.5208108534	-1.0462102185	C	1.5852029556	2.2156815704	1.1254313085
C	1.1211173931	-1.002562904	0.3557989751	C	1.0257452733	1.6523020575	-0.1610993934
C	1.0940561649	-2.0894749489	1.4163383037	C	1.2941232882	2.541935601	-1.3700127282
C	0.9369654533	-3.3784140171	1.0516602903	C	1.6004865917	3.8437040344	-1.18480813
C	0.874215763	-3.805253237	-0.3702694564	C	1.8408494796	4.4286206768	0.1591574363
C	1.2311869847	-1.6544321181	2.8555917736	C	1.168522317	1.9393455368	-2.7483929198
C	2.5351777716	-0.9478181256	3.1774561422	C	2.1458768398	0.8165205904	-3.0365879085
C	3.6712824664	-1.0312551907	2.4594434845	C	3.2890418863	0.5772192343	-2.3678627711
C	4.8771201313	-0.2593507791	2.7276170381	C	4.1530139692	-0.5709412403	-2.6048672072
C	5.9925825935	-0.2210499891	1.9683199719	C	5.2548096344	-0.9090659842	-1.9029775816
C	6.2843612801	-1.0031384164	0.7151722813	C	5.8994610901	-0.1568384289	-0.7690054204
C	6.4012169839	-0.1371944494	-0.5691225939	C	5.8440488715	-0.8885868688	0.5979122569
C	2.6512868051	1.5842821612	-1.9606881625	C	1.9899371225	-0.9559056497	2.6222660367
C	2.7767932193	0.2114418728	-1.7487294973	C	2.4728995431	0.2160622018	2.0426604229
C	3.9903592125	-0.3267629804	-1.3260412178	C	3.7108086966	0.2299911274	1.4052553658
C	5.0988145068	0.4762493276	-1.0504936055	C	4.4833062082	-0.9259613814	1.2725676923
C	4.9581256525	1.8668120614	-1.2635693532	C	3.9755668871	-2.1122608188	1.8511353532
C	3.7528643177	2.4114178611	-1.7238373007	C	2.7483422153	-2.1245090601	2.5260857892
C	5.98651697	4.0393992805	-1.1908093355	C	4.3040151474	-4.4565953184	2.2770303992
H	-0.9361176537	3.8911850353	1.2280801373	H	-3.9209486145	3.2318560598	1.4315810479
H	0.8668274053	1.7394046368	1.072521979	H	-1.1408828103	3.2128061016	1.027357854
H	-1.4580931874	-0.5861724082	-1.1451834926	H	-1.1524556188	0.2310462576	-1.6098339595
H	-2.1170922841	-1.6772969227	1.6577957025	H	-0.4641630826	-1.2742440273	0.9889282776
H	-2.8875015972	-3.6031408319	0.5963749059	H	0.4244638842	-2.9511111632	-0.3595491237
H	-2.0697178006	-2.3500014763	-2.1069653599	H	-0.3036588953	-1.2521502744	-2.8342557269
H	-2.1950023144	-4.0587830685	-1.7536676134	H	0.9460918989	-2.4734170703	-2.7471104927
H	-4.6036666997	-4.0371736126	-1.7829516323	H	-0.631290814	-4.2958316288	-2.7788712678
H	-4.1024538666	-3.0450888179	-3.1453147563	H	-1.1938987459	-3.1458806393	-3.9840317146
H	-3.3816103492	-0.529863161	-1.8752068741	H	-2.5012605456	-1.1113812802	-2.3547005437
H	-6.6123242657	-3.101691127	-0.7619375168	H	-2.5407754179	-5.3339157537	-1.6739611123
H	-7.5302600548	-1.2686470764	0.6528010107	H	-4.391695254	-5.0108205214	-0.0391590452
H	1.2590817712	-3.1404957281	-2.4250366326	H	2.2118624365	3.9045353405	2.2555382402
H	1.9699133261	-0.3512902509	0.5893000785	H	1.487163085	0.6751967765	-0.3421561597
H	1.1040831051	-2.5257042221	3.5092020973	H	1.2813673132	2.733926512	-3.4956159233
H	0.4084815726	-0.9685546077	3.1056639023	H	0.1494750872	1.5501889386	-2.8908695048
H	2.5151158797	-0.3049154557	4.0581545091	H	1.8638111315	0.1478039709	-3.8500863751
H	3.683313631	-1.6852600292	1.5904309086	H	3.5700488177	1.2556676728	-1.5659981948
H	4.8452805421	0.3749914739	3.6136049648	H	3.8429160767	-1.2359282624	-3.4110272315
H	6.7894863495	0.4512926736	2.2863757973	H	5.757479955	-1.835376135	-2.1816849302
H	5.5440946895	-1.7945472874	0.5633449636	H	5.4850226736	0.8515728731	-0.6773790731
H	7.2482403475	-1.5147669656	0.8461735148	H	6.9618331617	-0.0195449853	-1.0145450745
H	7.1327245202	0.6569113921	-0.4035428311	H	6.2072134952	-1.9113604486	0.4723827375
H	6.8099936574	-0.7742460002	-1.3653189136	H	6.5567510805	-0.3938315331	1.2728252967
H	4.0611496846	-1.4032305047	-1.1985772571	H	4.0748588933	1.1709697689	1.004730419
H	3.6305692673	3.4730661736	-1.8958774716	H	2.351289716	-3.022559524	2.9815960916
H	-3.2640998943	6.7581360207	1.3524835547	H	-7.5779137655	3.2978669835	1.9461258898
H	-1.8481027165	6.0160429298	0.5540097148	H	-6.1764801414	3.9907426903	1.0809645142
H	-2.2180363904	5.6355920178	2.2672067311	H	-5.9595174718	3.25297168	2.7011666177
H	6.9693492744	4.4279943266	-0.920993719	H	5.0790793308	-5.18933281	2.0486226641
H	5.2259392773	4.4904160838	-0.5411540631	H	3.3571157131	-4.7746768108	1.8232690614
H	5.770617056	4.291306073	-2.2365872834	H	4.1819027208	-4.3837669681	3.3648746507
H	0.8129975693	1.4515113904	-2.4464740127	H	0.4097867003	-0.1004470213	3.2633376714
H	0.7291335951	-5.2064020383	1.3957362611	H	1.9918258733	5.575269254	-1.7806857775
H	-5.2822813566	3.9917375672	0.2657786348	H	-6.9343401819	0.1202823323	0.3996822692
H	-6.1271325537	1.8156582614	1.0249019344	H	-5.7769679352	-2.0044123583	0.790998388
O	1.4854882273	2.1473885623	-2.3879379782	O	0.7918896748	-0.9919146782	3.2753733618
O	1.6549105936	-0.592773672	-2.014946948	O	1.6877380662	1.3680115649	2.2045337314
O	0.6856765281	-5.0042684064	-0.6157675292	O	2.092990391	5.6386649668	0.2378160757
O	0.8450345329	-4.4012734269	1.9395109771	O	1.7665543277	4.7239583268	-2.2071039205
O	6.0624556933	2.6326059191	-1.0011322985	O	4.7588562016	-3.228606718	1.7230089937
O	-5.2523467302	3.0421409561	0.0629707604	O	-6.2238127275	-0.4436696564	0.0518651778
O	-3.5262012041	4.817433062	0.8639553394	O	-6.3442060512	1.9129468728	1.1497320879
O	-6.3807579198	0.9035433886	1.2569624039	O	-5.2378272357	-2.8118508403	0.8812029228

Table 8.7. Cartesian coordinates of the lowest-energy conformers of the C-9 epimer of zosterabisphenone B (*epi*-**6**) optimized at the B3LYP/6-31+G(d,p) level.

Conformer 1, E = -2147.7361081 Harthrees ΔE = 0.00 kcal/mol, population 84.5 %				Conformer 2, E = -2147.7347249 Harthrees ΔE = 0.87 kcal/mol, population 15.5%			
C	3.9724629296	2.5575843289	-0.7100062783	C	4.982891869	0.2122387687	-0.550297574
C	3.6730965553	1.2178066469	-0.4118908415	C	3.7955927151	-0.4256339482	-0.1550636816
C	2.2968198955	0.87254825	-0.3843826474	C	2.632785082	0.3864937199	-0.1135190943
C	1.3155062298	1.8579365735	-0.6056127588	C	2.714617661	1.7631210419	-0.4006376931
C	1.6360712011	3.1918150195	-0.8853034029	C	3.9147982779	2.3824592292	-0.7716964212
C	2.9822580289	3.5303189551	-0.9414456416	C	5.0499623553	1.5855068405	-0.8526660304
C	-0.0069219136	1.2525225872	-0.505637238	C	1.3955362569	2.3673974477	-0.2598485982
C	0.0954195658	-0.0664802075	-0.2417929965	C	0.4856209964	1.4316233931	0.0771243184
C	1.5718411185	-0.4813164691	-0.1804587988	C	1.1532428113	0.0542640491	0.1795732468
C	1.9408899276	-1.3009376563	1.0433984246	C	0.8438608792	-0.6664127568	1.4785675677
C	2.4627964895	-2.5386218225	1.0840520435	C	0.2432229379	-1.8575222723	1.6439645917
C	2.8952534436	-3.4250821492	-0.0619314205	C	-0.1982202161	-2.8467760654	0.5905186939
C	4.452179333	-3.5529296373	-0.2012387556	C	0.7244227366	-4.112157516	0.4971435055
C	5.756751965	0.4065245797	0.8698528564	C	4.6263868699	-2.4416216371	1.2163587886
C	4.7295321835	0.2159392902	-0.0876306958	C	3.7527728933	-1.8567334506	0.2653657227
C	4.535819215	-1.0846345385	-0.5647693904	C	2.6166250337	-2.5912226559	-0.0928602061
C	5.1231855036	-2.2101514338	0.0087737116	C	2.1871405235	-3.7256063293	0.5929976095
C	6.111605951	-1.9969716542	0.9767467266	C	3.0522608257	-4.2643951888	1.5525141833
C	6.4614680027	-0.7017464688	1.356578511	C	4.2810513476	-3.660245172	1.8146285774
C	2.5762901643	5.8460059898	-1.4556928964	C	6.4963805817	3.3837383645	-1.5366781067
C	-1.2264699118	-2.6956699544	1.7131348673	C	-1.8915076841	3.4133195993	-1.348525453
C	-1.3861846089	-1.4376077365	1.2641839879	C	-1.5672966554	2.1314999607	-1.0938523647
C	-1.1042004505	-1.0099607698	-0.1584569301	C	-1.0275836559	1.6543157875	0.2364950707
C	-0.9905925923	-2.1650964208	-1.1355309544	C	-1.3447049908	2.6022710466	1.3830286256
C	-0.823915766	-3.4213244594	-0.6717393962	C	-1.6814863341	3.8826844903	1.1191043389
C	-0.8660252636	-3.7538696859	0.7781903687	C	-1.910157132	4.3846281726	-0.2603968668
C	-1.0593477211	-1.8347765434	-2.6084546632	C	-1.2390729295	2.0675754997	2.7898344018
C	-2.3556941806	-1.1764232261	-3.0441327495	C	-2.2140652755	0.9485526494	3.1018610233
C	-3.5285060551	-1.2372985604	-2.3854995166	C	-3.352377119	0.6904420806	2.4318191737
C	-4.7281426059	-0.5056084218	-2.7698776254	C	-4.2097494762	-0.4595756116	2.6842756225
C	-5.8844053119	-0.4346988947	-2.0770114931	C	-5.295085688	-0.8294067972	1.9724766447
C	-6.2370413037	-1.1342566281	-0.7912805356	C	-5.9232808685	-0.1182512777	0.8030653253
C	-6.4434527516	-0.183242964	0.4192810284	C	-5.8525208345	-0.9019492915	-0.5340150045
C	-2.8122207644	1.6922626411	1.9226941273	C	-1.9426547104	-1.1192224619	-2.4342618091
C	-2.897407903	0.3078210958	1.7749063537	C	-2.4254471678	0.0869508861	-1.9299883405
C	-4.0740420407	-0.2786711871	1.3139025998	C	-3.6826101871	0.1480566553	-1.3337136939
C	-5.181377531	0.4835426274	0.9363708271	C	-4.4775136532	-0.9892245617	-1.1741835973
C	-5.0813753214	1.8858862225	1.0849725504	C	-3.9745977275	-2.2074281931	-1.6866433642
C	-3.9150126836	2.4808677712	1.5821187525	C	-2.7251294801	-2.2696218168	-2.3170362857
C	-6.148792374	4.0275425567	0.8453806516	C	-4.3395267621	-4.5592802352	-2.0368847599
H	0.8553583605	3.9258163844	-1.047068571	H	3.9465581758	3.4440651312	-0.987322163
H	-0.9359244716	1.7951920669	-0.6485475463	H	1.191300497	3.4191710896	-0.4331452891
H	1.7456148883	-1.0958331362	-1.0733690196	H	0.7579682238	-0.5403132716	-0.6545433291
H	1.7356526402	-0.8075358806	1.9927086877	H	1.1539863984	-0.1277895963	2.3728231688
H	2.6475418257	-2.9516884221	2.0761341141	H	0.1146722625	-2.1942212518	2.6732784091
H	2.4818453438	-3.0601848637	-1.0071117253	H	-0.2407956015	-2.369355113	-0.3915690076
H	2.4783584221	-4.4299960407	0.0797042169	H	-1.2211116522	-3.1792094508	0.8081065979
H	4.6711175899	-3.9881821304	-1.1852366708	H	0.4859687554	-4.635795	-0.4382760013
H	4.8320345896	-4.2606820833	0.5444583639	H	0.4867015341	-4.8048321318	1.3129324186
H	3.7808416005	-1.2268231895	-1.3273931531	H	1.9706217363	-2.1716820909	-0.8532363283
H	6.5905530588	-2.8443162027	1.4620748292	H	2.7560793597	-5.1426873756	2.1217097703
H	7.2361148061	-0.5234890226	2.0956473522	H	4.9672932317	-4.0769125974	2.5450492747
H	-1.4053580572	-2.9606994539	2.7494549234	H	-2.2063499824	3.7360889233	-2.3348644021
H	-1.9575716806	-0.4011906084	-0.4758136352	H	-1.4733134698	0.6809440338	0.467628377
H	-0.8860048593	-2.747501998	-3.1910836776	H	-1.371989258	2.8923681621	3.4999231696
H	-0.2376436968	-1.151068559	-2.8688302609	H	-0.2217959822	1.6880051364	2.9592454058
H	-2.2970686761	-0.58848763	-3.9606178016	H	-1.9294839098	0.2952507332	3.9267914922
H	-3.5801295974	-1.8355162476	-1.4785424105	H	-3.6312812532	1.3516577297	1.6145906581
H	-4.6550330708	0.0692742785	-3.6931716652	H	-3.9069265147	-1.099511559	3.5131659577
H	-6.6703201089	0.2020636438	-2.4834418071	H	-5.7948015596	-1.7518020321	2.2690841723
H	-5.4951758194	-1.8989102709	-0.5419474441	H	-5.5034528516	0.8843294026	0.6781262835
H	-7.1835177076	-1.6723191442	-0.9417945235	H	-6.9885740142	0.0322173999	1.0276635738
H	-7.1725334755	0.5855474941	0.153563018	H	-6.2366978953	-1.9127440108	-0.3780410974
H	-6.8942727725	-0.7694928842	1.2317270157	H	-6.5407475241	-0.4210526986	-1.2434125925
H	-4.1155163888	-1.3614985942	1.2389954871	H	-4.0432704148	1.1104943551	-0.9850715024
H	-3.8246557827	3.5523752417	1.7062803353	H	-2.3298622153	-3.1931964609	-2.7200749242
H	3.1884855698	6.7283954304	-1.6437190976	H	7.5514166414	3.4901303026	-1.7900979916
H	1.9376076174	6.0209003447	-0.5821111542	H	6.2580785165	4.0230521492	-0.6787649131
H	1.9542352134	5.6362652966	-2.3337180244	H	5.8806661461	3.6721251047	-2.3965263891
H	-7.1228584474	4.3792119947	0.5028770065	H	-5.1376321535	-5.266488552	-1.8075743161
H	-5.3626581937	4.4637179038	0.2162834358	H	-3.4143940289	-4.8828040413	-1.5438056338
H	-5.9968277567	4.3358434349	1.8872951501	H	-4.1836023724	-4.5264964407	-3.1223537525
H	-0.9999555254	1.6248579077	2.5118138762	H	-0.3459900435	-0.3105154777	-3.0985050386
H	-0.5774132661	-5.2647159163	-0.8719810386	H	-2.1387505771	5.6313466451	1.611200759
H	5.3078354783	3.8989910418	-1.0038867872	H	6.8531666616	0.0728865273	-0.9397852991
H	5.8920524679	2.3185527642	0.7760491903	H	6.1492722964	-1.3031718417	0.951840271
O	-1.682966549	2.3015936823	2.3830437677	O	-0.7134676975	-1.205418838	-3.0278497346
O	-1.7763950022	-0.4521599593	2.1467855461	O	-1.6146091326	1.2155128667	-2.1228226152
O	-0.6846360669	-4.9328709278	1.1106580117	O	-2.1965870154	5.5798193073	-0.4126555616
O	-0.6554038501	-4.5004542106	-1.4788090916	O	-1.8968337924	4.8119826363	2.0876890936
O	-6.1849094274	2.6117518561	0.7233230325	O	-4.7824029998	-3.3032697635	-1.5387129843
O	5.293339071	2.9533085443	-0.7819844765	O	6.1457373272	-0.526228984	-0.6490098468
O	3.4904157508	4.7798428981	-1.2128966222	O	6.3081947621	2.0078421214	-1.21421995
O	6.0580573073	1.6148815213	1.4299320815	O	5.7682462799	-1.8451206214	1.6673427238

Table 8.8. Cartesian coordinates of the lowest-energy conformer of the aromatic tautomer of zosterabisphenone A (**5a**) optimized at the B3LYP/6-31+G(d,p) level.

conformer 1, *E* = -1957.9807093 Harthrees >99% population			
C	-2.0142010139	3.6080249593	-1.9358689038
C	-0.612673093	3.5605796004	-1.8570357437
C	-0.0654679587	2.5066778337	-1.0795596923
C	-0.905409037	1.5447366036	-0.4897970639
C	-2.3008227248	1.5975790582	-0.6038205196
C	-2.846593394	2.650415709	-1.3273863811
C	-0.1012171351	0.5592668163	0.2216686948
C	1.2129938351	0.8739630653	0.1600628463
C	1.3826065846	2.1646060579	-0.6791331967
C	2.3898040927	1.9894490711	-1.8016258685
C	3.5260754522	2.6771755137	-2.0050749496
C	4.0648758602	3.8578806553	-1.2284650023
C	3.929798698	5.2297165808	-1.9776368622
C	0.1457650662	4.8000324126	-3.9858327803
C	0.2688502226	4.5054951149	-2.6042662332
C	1.4521175629	4.9054488715	-1.9731610534
C	2.581089529	5.347558007	-2.6597003571
C	2.4434927164	5.5969058089	-4.0300393431
C	1.2226113418	5.3813714476	-4.6676474398
C	-5.1251719046	1.9584841683	-0.9790820353
C	4.1046079283	-1.5327431228	0.2860883813
C	2.7634069843	-1.1357782251	0.2725763148
C	2.4453463761	0.2000059715	0.6478236555
C	3.3920147991	0.9278168008	1.4085577528
C	4.7145022057	0.4737618226	1.4779064453
C	5.088558133	-0.7025380963	0.8320686331
C	2.8886998613	1.9499799992	2.4321305208
C	2.450286971	1.13775316	3.6456506762
C	1.2552538049	0.5215578391	3.7527364966
C	0.9697223108	-0.5302542061	4.7249549999
C	0.0125117321	-1.4819283719	4.6462854274
C	-1.0816684325	-1.6662252536	3.6200620067
C	-1.0454868664	-3.0130396575	2.8310573573
C	1.3822398652	-3.0643911255	-0.8013811742
C	1.6662107478	-2.1456331246	0.2207812888
C	0.8879023886	-2.1970622511	1.3934703474
C	-0.2133373423	-3.0332043708	1.5545116145
C	-0.5205910757	-3.8846554149	0.4789389595
C	0.271844848	-3.9111493103	-0.6696416232
O	-2.6189579828	4.6296416886	-2.6441760515
O	-4.1888546231	2.8779525904	-1.5338382712
O	-0.94292587	4.4715331783	-4.7397959393
O	6.3744106916	-1.2017426474	0.7885378117
O	4.431733161	-2.7783917652	-0.1989883422
O	2.1204481034	-3.1519609089	-1.9499091569
H	-2.923122394	0.8395750452	-0.1425061293
H	-0.5292286782	-0.2921835885	0.7313764335
H	1.7503016615	2.9426552123	-0.0016219511
H	2.1470283518	1.1888937932	-2.4989790017
H	4.1180440707	2.3894136266	-2.8745985169
H	3.5642453902	3.9346344013	-0.2582541449
H	5.1272618599	3.6929559563	-1.0051306572
H	4.1089865328	6.0330116163	-1.2502217239
H	4.7176946044	5.31433919	-2.7350959829
H	1.5325005442	4.7249583041	-0.9093580803
H	3.299340113	5.9310269757	-4.6122501843
H	1.0987141784	5.584903066	-5.7265744739
H	-6.111190285	2.3255652111	-1.2653385987
H	-4.9748051007	0.952003391	-1.3866989222
H	-5.0464670668	1.9306225314	0.1141028236
H	5.4501138507	1.0275110575	2.0591085927
H	2.0442338835	2.5248961227	2.0456237855
H	3.6803275955	2.6568198233	2.7035887735
H	3.2293079124	0.9057754451	4.373261549
H	0.5002672818	0.730870397	2.9982718991
H	1.6735045375	-0.604298606	5.5543035947
H	0.0135532063	-2.2372226845	5.4332825343
H	-2.0365054427	-1.6428785358	4.1641181245
H	-1.1147742487	-0.8277682273	2.9178145736
H	-2.0754163644	-3.2803660348	2.5652064152
H	-0.7015699782	-3.8061851506	3.5108565184
H	1.1736066277	-1.5360404554	2.2000032198
H	-1.3781428316	-4.5512554944	0.5419655769
H	0.0485707144	-4.5937996537	-1.4835356252
H	-3.579901611	4.4987543078	-2.5870893294
H	-1.7396107839	4.496575544	-4.1777564362
H	6.9829910979	-0.6138832029	1.2534212979
H	5.3801590807	-2.9284256136	-0.0663513899
H	3.0421052701	-2.9206093243	-1.7495348111

Table 8.9. Experimental chemical shifts, calculated isotropic shielding constants, and predicted chemical shifts of zosterabisphenone A (**5**). Calculations were performed at the PBE0/6-311+G(2d,p)/PCM($CHCl_3$)//B3LYP/6-31+G(d,p) level.

	exp. δ_C	shielding constants	calcd. δ_C[a]	$\Delta\delta_C$
C-1	138.0	42.81	137.19	-0.81
C-2	122.4	57.71	123.04	0.64
C-3	139.7	39.94	139.92	0.22
C-4	137.3	44.03	136.03	-1.27
C-5	103.0	81.52	100.44	-2.56
C-6	145.9	35.51	144.12	-1.78
C-7	128.8	49.92	130.44	1.64
C-8	148.3	29.57	149.76	1.46
C-9	51.3	130.85	53.61	2.31
C-10	127.4	50.49	129.90	2.50
C-11	134.4	42.66	137.33	2.93
C-12	27.6	156.11	29.62	2.02
C-13	34.1	148.96	36.41	2.31
C-14	152.0	26.82	152.37	0.37
C-15	122.7	58.19	122.59	-0.11
C-16	135.1	45.91	134.25	-0.85
C-17	131.9	49.50	130.84	-1.06
C-18	129.7	51.97	128.50	-1.20
C-19	116.7	65.62	115.54	-1.16
6-OMe	56.2	131.18	53.29	-2.91
C-1'	103.9	78.54	103.27	-0.63
C-2'	50.2	131.30	53.18	2.98
C-3'	46.1	136.09	48.63	2.53
C-4'	163.3	12.38	166.08	2.78
C-5'	122.6	57.60	123.15	0.55
C-6'	189.1	-10.01	187.34	-1.76
C-7'	38.5	145.33	39.86	1.36
C-8'	124.5	55.17	125.45	0.95
C-9'	128.1	51.10	129.32	1.22
C-10'	129.2	50.61	129.79	0.59
C-11'	129.4	48.65	131.64	2.24
C-12'	29.4	154.32	31.32	1.92
C-13'	33.6	150.20	35.24	1.64
C-14'	154.9	24.80	154.29	-0.61
C-15'	127.8	52.50	127.99	0.19
C-16'	123.5	57.20	123.53	0.03
C-17'	133.9	47.39	132.84	-1.06
C-18'	130.5	51.10	129.32	-1.18
C-19'	109.0	73.95	107.62	-1.38
			RMSD ^{13}C	**1.66**

	exp. δ_H	shielding constants	calcd. δ_H[b]	$\Delta\delta_H$
H-5	6.80	24.56	6.56	-0.24
H-7	6.25	25.00	6.16	-0.09
H-9	5.01	26.35	4.93	-0.08
H-10	4.76	26.50	4.80	0.04
H-11	5.55	25.58	5.63	0.08
H-12	3.05	28.39	3.07	0.02
H-12	2.56	28.95	2.56	0.00
H-13	3.39	28.03	3.40	0.01
H-13	2.71	28.79	2.70	-0.01
H-16	7.68	23.52	7.51	-0.17
H-18	6.91	24.34	6.77	-0.14
H-19	6.86	24.46	6.66	-0.20
H-2'	3.85	27.69	3.71	-0.14
H-3'	3.56	27.84	3.57	0.01
H-5'	6.18	25.14	6.04	-0.14
H-7'	2.33	29.17	2.36	0.03
H-7'	2.97	28.56	2.91	-0.06
H-8'	5.08	26.31	4.97	-0.11
H-9'	5.89	25.19	5.99	0.10
H-10'	5.65	25.56	5.65	0.00
H-11'	5.22	26.05	5.21	-0.01
H-12'	2.31	29.28	2.25	-0.06
H-12'	2.53	28.91	2.59	0.06
H-13'	2.98	28.51	2.96	-0.02
H-13'	2.34	29.26	2.28	-0.06
H-16'	6.85	24.22	6.88	0.03
H-18'	6.70	24.50	6.61	-0.09
H-19'	6.56	24.85	6.30	-0.26
		27.45		
6-OMe	3.96	27.73	3.75	-0.21
		27.76		
			RMSD ^{1}H	**0.113**

[a] ^{13}C chemical shifts were obtained from the isotropic shielding constants according to ref. 4,[4] using the equation: $\delta = (187.3123 - \text{shielding})/1.0533$

[b] ^{1}H chemical shifts were obtained from the isotropic shielding constants according to ref. 4,[4] using the equation: $\delta = (31.7532 - \text{shielding})/1.0958$

Table 8.10 Experimental chemical shifts, calculated isotropic shielding constants, and predicted chemical shifts of the two conformers of the C-9 epimer of zosterabisphenone A (*epi*-**5**). Calculations were performed at the PBE0/6-311+G(2d,p)/PCM($CHCl_3$)//B3LYP/6-31+G(d,p) level.

		isotropic shieldings				
	exp. δ_C	Conformer 1 ΔE = 0.00 kcal/mol population 77.6%	Conformer 2 ΔE = 0.63 kcal/mol population 22.4%	average	calcd. δ_C[a]	$\Delta\delta_C$
C-1	138.0	42.70	43.13	42.79	137.21	-0.79
C-2	122.4	57.93	57.73	57.89	122.87	0.47
C-3	139.7	40.93	41.34	41.02	138.89	-0.81
C-4	137.3	43.27	42.77	43.16	136.86	-0.44
C-5	103.0	81.57	81.64	81.58	100.38	-2.62
C-6	145.9	35.45	35.50	35.46	144.17	-1.73
C-7	128.8	48.72	50.62	49.15	131.17	2.37
C-8	148.3	30.86	24.60	29.46	149.87	1.57
C-9	51.3	126.00	129.25	126.72	57.52	6.22
C-10	127.4	51.65	50.26	51.34	129.09	1.69
C-11	134.4	43.14	43.49	43.22	136.80	2.40
C-12	27.6	156.58	155.52	156.35	29.40	1.80
C-13	34.1	148.73	149.14	148.82	36.54	2.44
C-14	152.0	26.61	26.71	26.63	152.55	0.55
C-15	122.7	58.34	58.20	58.31	122.48	-0.22
C-16	135.1	46.10	46.22	46.13	134.04	-1.06
C-17	131.9	49.74	50.11	49.82	130.53	-1.37
C-18	129.7	51.88	52.26	51.97	128.50	-1.20
C-19	116.7	65.77	65.46	65.70	115.46	-1.24
6-OMe	56.2	131.21	131.18	131.20	53.27	-2.93
C-1'	103.9	78.70	78.26	78.60	103.21	-0.69
C-2'	50.2	131.84	131.82	131.83	52.67	2.47
C-3'	46.1	135.75	137.29	136.09	48.63	2.53
C-4'	163.3	11.98	7.98	11.08	167.31	4.01
C-5'	122.6	58.91	59.33	59.00	121.82	-0.78
C-6'	189.1	-10.55	-9.59	-10.34	187.65	-1.45
C-7'	38.5	145.32	145.23	145.30	39.89	1.39
C-8'	124.5	55.40	56.01	55.54	125.11	0.61
C-9'	128.1	51.39	50.87	51.27	129.15	1.05
C-10'	129.2	50.57	50.60	50.58	129.81	0.61
C-11'	129.4	48.41	48.38	48.40	131.88	2.48
C-12'	29.4	154.27	154.25	154.26	31.38	1.98
C-13'	33.6	150.39	150.37	150.38	35.06	1.46
C-14'	154.9	25.01	25.21	25.05	154.05	-0.85
C-15'	127.8	52.44	52.43	52.44	128.05	0.25
C-16'	123.5	57.55	57.40	57.52	123.23	-0.27
C-17'	133.9	47.81	47.51	47.74	132.51	-1.39
C-18'	130.5	51.38	51.08	51.31	129.12	-1.38
C-19'	109.0	73.84	74.04	73.88	107.69	-1.31
					RMSD ^{13}C	**1.93**

	exp. δ_H	Conformer 1	Conformer 2	average	calcd. δ_H[b]	$\Delta\delta_H$
H-5	6.80	24.56	24.50	24.54	6.58	-0.22
H-7	6.25	24.92	24.27	24.77	6.37	0.12
H-9	5.01	26.51	26.64	26.54	4.76	-0.25
H-10	4.76	26.39	26.45	26.40	4.88	0.12
H-11	5.55	25.66	25.65	25.66	5.56	0.01
H-12	3.05	28.32	28.49	28.36	3.10	0.05
H-12	2.56	28.84	29.07	28.89	2.61	0.05
H-13	3.39	27.98	28.07	28.00	3.42	0.03
H-13	2.71	28.75	28.83	28.77	2.72	0.01
H-16	7.68	23.49	23.60	23.51	7.52	-0.16
H-18	6.91	24.31	24.36	24.33	6.78	-0.13
H-19	6.86	24.43	24.45	24.44	6.68	-0.18
H-2'	3.85	27.44	27.50	27.45	3.93	0.08
H-3'	3.56	27.98	27.90	27.96	3.46	-0.10
H-5'	6.18	25.20	25.46	25.26	5.92	-0.26
H-7'	2.33	29.02	28.63	28.93	2.58	0.25
H-7'	2.97	28.55	28.51	28.54	2.93	-0.04
H-8'	5.08	26.30	26.38	26.32	4.96	-0.12
H-9'	5.89	25.09	25.20	25.11	6.06	0.17
H-10'	5.65	25.53	25.56	25.54	5.67	0.02
H-11'	5.22	26.02	26.03	26.02	5.23	0.01
H-12'	2.31	29.23	29.29	29.24	2.29	-0.02
H-12'	2.53	28.87	28.96	28.89	2.61	0.08
H-13'	2.98	28.51	28.55	28.52	2.95	-0.03
H-13'	2.34	29.24	29.26	29.24	2.29	-0.05
H-16'	6.85	24.21	24.31	24.23	6.87	0.02
H-18'	6.70	24.61	24.56	24.60	6.53	-0.17
H-19'	6.56	24.90	24.85	24.89	6.27	-0.29
		27.77	27.44			
6-OMe	3.96	27.76	27.74	27.76	3.65	-0.31
		27.85	27.71			
					RMSD ^{1}H	**0.148**

[a] ^{13}C chemical shifts were obtained from the isotropic shielding constants according to ref. 4,[4] using the equation: δ = (187.3123 − shielding)/1.0533

[b] ^{1}H chemical shifts were obtained from the isotropic shielding constants according to ref. 4,[4] using the equation: δ = (31.7532 − shielding)/1.0958

Table 8.11. Experimental chemical shifts, calculated isotropic shielding constants, and predicted chemical shifts of the two conformers of zosterabisphenone B (**6**). Calculations were performed at the PBE0/6-311+G(2d,p)/PCM($CHCl_3$)//B3LYP/6-31+G(d,p) level.

		isotropic shieldings				
	exp. δ_C	Conformer 1 ΔE = 0.00 kcal/mol population 85.4%	Conformer 2 ΔE = 0.91 kcal/mol Population 14.6%	average	calcd. δ_C[a]	$\Delta\delta_C$
C-1	138.0	42.58	42.69	42.59	138.05	0.06
C-2	122.4	57.65	57.78	57.67	123.63	1.23
C-3	139.6	39.92	40.39	39.99	140.55	0.94
C-4	136.8	44.59	43.62	44.45	136.28	-0.54
C-5	102.6	81.95	81.79	81.93	100.40	-2.22
C-6	145.6	35.70	35.60	35.68	144.67	-0.96
C-7	134.7	43.40	51.53	44.58	136.15	1.47
C-8	141.1	35.82	31.12	35.14	145.19	4.10
C-9	50.6	130.93	127.76	130.47	53.94	3.37
C-10	126.3	51.74	53.12	51.94	129.11	2.79
C-11	132.2	45.70	42.25	45.20	135.56	3.35
C-12	27.6	156.40	156.11	156.36	29.16	1.53
C-13	33.9	149.54	149.34	149.51	35.72	1.87
C-14	151.5	27.19	26.83	27.14	152.85	1.31
C-15	122.3	58.41	58.39	58.41	122.92	0.59
C-16	134.9	46.22	46.35	46.24	134.57	-0.32
C-17	132.7	49.04	49.86	49.16	131.77	-0.89
C-18	129.7	52.01	51.86	51.99	129.06	-0.59
C-19	116.6	65.36	65.61	65.40	116.22	-0.36
6-OMe	56.1	131.31	131.36	131.32	53.13	-2.98
C-1'	113.6	66.09	65.62	66.02	115.63	2.01
C-2'	175.4	-0.41	-0.48	-0.42	179.23	3.79
C-3'	42.3	139.66	142.95	140.14	44.69	2.40
C-4'	121.6	57.38	51.99	56.59	124.66	3.03
C-5'	145.3	33.57	35.47	33.84	146.43	1.10
C-6'	183.2	-2.43	-2.62	-2.46	181.18	-2.03
C-7'	27.6	155.60	154.82	155.49	29.99	2.36
C-8'	129.4	47.76	48.62	47.88	132.99	3.61
C-9'	128.0	52.30	52.71	52.36	128.70	0.72
C-10'	128.4	51.90	52.05	51.92	129.13	0.70
C-11'	129.5	48.79	48.57	48.75	132.16	2.64
C-12'	27.9	155.80	157.50	156.05	29.45	1.59
C-13'	24.5	160.00	160.07	160.01	25.66	1.20
C-14'	144.7	34.67	35.42	34.78	145.54	0.85
C-15'	136.4	44.52	44.45	44.51	136.22	-0.13
C-16'	119.1	62.65	64.85	62.97	118.55	-0.55
C-17'	121.7	60.34	61.04	60.44	120.97	-0.72
C-18'	154.9	25.43	25.82	25.49	154.43	-0.43
C-19'	99.4	86.14	86.07	86.13	96.38	-2.98
18'-OMe	55.8	132.04	132.09	132.05	52.43	-3.41
					RMSD ^{13}C	**2.05**

	exp. δ_H	Conformer 1	Conformer 2	average	calcd. δ_H[b]	$\Delta\delta_H$
H-5	6.85	24.57	24.56	24.56	6.67	-0.18
H-7	6.26	24.92	24.57	24.87	6.39	0.13
H-9	4.76	26.61	26.68	26.62	4.78	0.02
H-10	4.36	26.93	26.74	26.90	4.52	0.16
H-11	4.93	26.24	26.37	26.26	5.11	0.18
H-12	2.72	28.77	28.84	28.78	2.79	0.07
H-12	2.17	29.41	29.44	29.41	2.20	0.03
H-13	2.53	29.00	29.00	29.00	2.59	0.06
H-13	3.26	28.20	28.20	28.20	3.32	0.06
H-16	7.53	23.65	23.64	23.65	7.52	-0.01
H-18	6.81	24.46	24.44	24.46	6.77	-0.04
H-19	6.77	24.52	24.48	24.52	6.72	-0.05
H-1'	6.34	25.06	25.03	25.06	6.22	-0.12
H-3'	4.62	26.88	27.18	26.92	4.50	-0.12
H-7'	3.41	27.93	27.64	27.89	3.61	0.20
H-7'	2.35	29.00	28.76	28.96	2.62	0.27
H-8'	5.60	25.64	25.53	25.62	5.70	0.10
H-9'	5.86	25.31	25.23	25.30	6.00	0.14
H-10'	5.90	25.35	25.32	25.34	5.96	0.06
H-11'	5.34	25.92	25.92	25.92	5.42	0.08
H-12'	2.65	28.81	28.77	28.80	2.76	0.11
H-12'	2.28	29.32	29.37	29.32	2.28	0.00
H-13'	3.01	28.45	28.49	28.45	3.09	0.08
H-13'	2.43	29.14	29.18	29.14	2.45	0.02
H-16'	6.72	24.50	24.56	24.50	6.73	0.01
H-19'	6.50	25.03	25.19	25.05	6.22	-0.28
		27.43	27.48			
6-OMe	3.97	27.74	27.79	27.64	3.83	-0.14
		27.74	27.75			
		27.62	27.69			
18'-OMe	3.77	28.01	28.14	27.89	3.60	-0.17
		28.01	28.09			
					RMSD ^{1}H	**0.126**

[a] ^{13}C chemical shifts were obtained from the isotropic shielding constants according to ref. 4,[4] using the equation: δ = (187.3123 − shielding)/1.0533

[b] ^{1}H chemical shifts were obtained from the isotropic shielding constants according to ref. 4,[4] using the equation: δ = (31.7532 − shielding)/1.0958

Table 8.12. Experimental chemical shifts, calculated isotropic shielding constants, and predicted chemical shifts of the two conformers of the C-9 epimer of zosterabisphenone B (*epi*-**6**). Calculations were performed at the PBE0/6-311+G(2d,p)/PCM($CHCl_3$)//B3LYP/6-31+G(d,p) level.

	isotropic shieldings					
	exp. δ_C	Conformer 1 ΔE = 0.00 kcal/mol population 84.5%	Conformer 2 ΔE = 0.87 kcal/mol population 25.5%	average	calcd. δ_C[a]	$\Delta\delta_C$
C-1	138.0	42.55	42.74	42.58	138.07	0.08
C-2	122.4	57.63	57.54	57.62	123.67	1.27
C-3	139.6	39.74	40.52	39.86	140.67	1.06
C-4	136.8	44.47	43.37	44.30	136.42	-0.40
C-5	102.6	82.04	81.82	82.01	100.33	-2.29
C-6	145.6	35.67	35.48	35.64	144.71	-0.92
C-7	134.7	43.34	49.63	44.31	136.41	1.73
C-8	141.1	37.19	32.38	36.44	143.94	2.85
C-9	50.6	130.85	127.55	130.33	54.07	3.50
C-10	126.3	52.34	54.12	52.61	128.46	2.14
C-11	132.2	43.51	41.57	43.21	137.47	5.26
C-12	27.6	156.53	156.48	156.53	29.00	1.37
C-13	33.9	149.39	149.14	149.35	35.87	2.02
C-14	151.5	27.08	26.86	27.04	152.94	1.40
C-15	122.3	58.86	58.52	58.81	122.53	0.20
C-16	134.9	46.09	45.78	46.04	134.76	-0.13
C-17	132.7	49.13	49.32	49.16	131.77	-0.89
C-18	129.7	52.17	52.13	52.16	128.90	-0.75
C-19	116.6	65.74	65.74	65.74	115.90	-0.68
6-OMe	56.1	131.26	131.21	131.25	53.19	-2.92
C-1'	113.6	64.77	67.02	65.12	116.49	2.87
C-2'	175.4	1.95	-2.25	1.30	177.58	2.14
C-3'	42.3	140.42	142.81	140.79	44.06	1.77
C-4'	121.6	56.05	51.84	55.39	125.80	4.17
C-5'	145.3	33.79	35.34	34.03	146.25	0.92
C-6'	183.2	-2.45	-3.02	-2.54	181.26	-1.95
C-7'	27.6	155.45	155.25	155.42	30.06	2.43
C-8'	129.4	48.06	47.17	47.92	132.95	3.57
C-9'	128.0	52.32	53.52	52.51	128.57	0.59
C-10'	128.4	52.07	51.69	52.01	129.04	0.61
C-11'	129.5	48.52	49.45	48.67	132.24	2.72
C-12'	27.9	155.80	157.49	156.06	29.44	1.58
C-13'	24.5	160.07	159.89	160.05	25.63	1.17
C-14'	144.7	34.94	35.72	35.06	145.26	0.57
C-15'	136.4	44.04	44.13	44.05	136.66	0.31
C-16'	119.1	63.02	64.91	63.32	118.22	-0.88
C-17'	121.7	60.44	60.32	60.42	120.99	-0.70
C-18'	154.9	25.59	25.84	25.63	154.29	-0.57
C-19'	99.4	86.06	85.54	85.98	96.53	-2.83
18'-OMe	55.8	132.02	131.91	132.01	52.47	-3.37
					RMSD ^{13}C	**2.08**

	exp. δ_H	Conformer 1	Conformer 2	average	calcd. δ_H[b]	$\Delta\delta_H$
H-5	6.85	24.55	24.55	24.55	6.69	-0.16
H-7	6.26	24.91	24.66	24.87	6.39	0.13
H-9	4.76	26.82	26.72	26.81	4.61	-0.15
H-10	4.36	26.81	26.94	26.83	4.58	0.22
H-11	4.93	26.23	26.14	26.22	5.15	0.22
H-12	2.72	28.90	28.91	28.90	2.67	-0.05
H-12	2.17	29.52	29.83	29.57	2.05	-0.12
H-13	2.53	28.25	28.15	28.24	3.29	0.03
H-13	3.26	29.02	29.03	29.02	2.56	0.03
H-16	7.53	23.74	23.57	23.71	7.46	-0.07
H-18	6.81	24.44	24.43	24.44	6.79	-0.02
H-19	6.77	24.51	24.49	24.51	6.72	-0.05
H-1'	6.34	25.17	25.18	25.17	6.12	-0.22
H-3'	4.62	26.86	27.37	26.94	4.49	-0.13
H-7'	3.41	27.69	27.87	27.72	3.76	0.35
H-7'	2.35	28.92	29.10	28.95	2.63	0.28
H-8'	5.60	25.65	25.77	25.67	5.66	0.06
H-9'	5.86	25.28	25.25	25.27	6.02	0.16
H-10'	5.90	25.36	25.41	25.37	5.93	0.03
H-11'	5.34	25.92	25.98	25.93	5.42	0.08
H-12'	2.65	28.80	28.76	28.80	2.77	0.12
H-12'	2.28	29.31	29.37	29.32	2.29	0.01
H-13'	3.01	28.44	28.45	28.44	3.10	0.09
H-13'	2.43	29.14	29.14	29.14	2.45	0.02
H-16'	6.72	24.49	24.53	24.50	6.74	0.02
H-19'	6.50	24.97	24.83	24.95	6.32	-0.18
		27.43	27.43			
6-OMe	3.97	27.74	27.77	27.65	3.83	-0.14
		27.77	27.76			
		27.62	27.58			
18'-OMe	3.77	27.99	27.96	27.86	3.63	-0.14
		27.99	27.95			
					RMSD ^{1}H	**0.145**

[a] ^{13}C chemical shifts were obtained from the isotropic shielding constants according to ref. 4,4 using the equation: δ = (187.3123 − shielding)/1.0533

[b] ^{1}H chemical shifts were obtained from the isotropic shielding constants according to ref. 4,[4] using the equation: δ = (31.7532 − shielding)/1.0958

Table 8.13. Experimental multiplicity of ^{1}H NMR signals and predicted ^{1}H-^{1}H *J* couplings (Hz) of zosterabisphenone A (**5**). Predicted couplings whose magnitude is smaller than 0.2 Hz are not reported.

		predicted ^{1}H-^{1}H *J* couplings (Hz)											
	southern unit	H-5	H-7	H-9	H-10	H-11	H-12 proR	H-12 proS	H-13 proR	H-13 proS	H-16	H-18	H-19	H-3'
Position	δ_H, mult (*J* in Hz)													
H-5	6.80, s		-0.3											-0.2
H-7	6.25, s	-0.3		-2.1										-1.4
H-9	5.01, br. d (11.2)		-2.1		10.8	-0.8	-0.7				-0.4			-0.5
H-10	4.76, t (11.2)			10.8		11.4	-0.9	-1.7		0.5				
H-11	5.55, ddd (11.8, 11.2, 4.8)			-0.8	11.4		12.1	5.1		-0.4				
H-12 proR	3.05, dddd (13.6, 12.6, 11.8, 7.0)			-0.7	-0.9	12.1		-13.4	11.8	7.2				
H-12 proS	2.56, overlapped				-1.7	5.1	-13.4		6.0	1.0				
H-13 proR	2.71, ddd (14.4, 12.6, 6.0)						11.8	6.0		-14.4		-0.4		
H-13 proS	3.39, dd (14.4, 7.0)				0.5	-0.4	7.2	1.0	-14.4		-1.2	-1.1	0.6	
H-16	7.68, br. s			-0.4						-1.2		2.0	0.4	
H-18	6.91, br. d (8.0)								-0.4	-1.1	2.0		8.1	
H-19	6.86, d (8.0)									0.6	0.4	8.1		

		predicted ^{1}H-^{1}H *J* couplings (Hz)														
	northern unit	H-2'	H-3'	H-5'	H-7' proR	H-7' proS	H-8'	H-9'	H-10'	H-11'	H-12' proR	H-12' proS	H-13' proR	H-13' proS	H-16'	H-18'	H-19'
Position	δ_H, mult (*J* in Hz)																
H-2'	3.85, s		1.1										-0.4		-1.2	-0.9	0.5
H-3'	3.56, s	1.1			0.4	-0.4									-0.2		
H-5'	6.18, s				-1.1	-0.5											
H-7' proR	2.97, dd (12.4, 7.4)		0.4	-1.1		-11.8	7.5	-1.0	0.5	-0.5	0.5						
H-7' proS	2.33, overlapped		-0.4	-0.5	-11.8		9.1	-1.1	0.3	-0.4	0.4						
H-8'	5.08, ddd (15.3, 8.0, 8.0)				7.5	9.1		15.6	-1.0	1.0	-0.8						
H-9'	5.89, dd (15.3, 10.6)				-1.0	-1.1	15.6		10.7	-1.4	0.6	-0.4					
H-10'	5.65, t (10.6)				0.5	0.3	-1.0	10.7		11.5	-1.6	-0.9	0.3				
H-11'	5.22, ddd (11.5, 10.6, 5.3)				-0.5	-0.4	1.0	-1.4	11.5		5.7	11.8	-0.2	-0.2			
H-12' proR	2.31, overlapped				0.5	0.4	-0.8	0.6	-1.6	5.7		-12.8	3.5	3.4			
H-12' proS	2.53, dq (3.3,12.2)							-0.4	-0.9	11.8	-12.8		3.9	13.0			
H-13' proR	2.98, overlapped	-0.4							0.3	-0.2	3.5	3.9		-13.5	-0.9	-0.8	0.5
H-13' proR	2.34, overlapped									-0.2	3.4	13.0	-13.5			-0.4	
H-16'	6.85, br. s	-1.2	-0.2										-0.9	-		1.7	0.5
H-18'	6.70, br. d (8.0)	-0.9											-0.8	-0.4	1.7		8.1
H-19'	6.56, d (8.0)	0.5											0.5		0.5	8.1	

Table 8.14. Experimental multiplicity of ^{1}H NMR signals and predicted ^{1}H-^{1}H *J* couplings of zosterabisphenone B (**6**). Predicted couplings whose magnitude is smaller than 0.2 Hz are not reported.

		predicted ^{1}H-^{1}H *J* couplings (Hz)												
	southern unit	H-5	H-7	H-9	H-10	H-11	H-12 proR	H-12 proS	H-13 proR	H-13 proS	H-16	H-18	H-19	H-3'
Position	δ_H, mult (*J* in Hz)													
H-5	6.85, s		-0.3											
H-7	6.26, s	-0.3		-2.2										-0.5
H-9	4.76, br. d (11)		-2.2		10.4	-0.9	-0.7				-0.4			-0.5
H-10	4.36, t (11)			10.4		11.5	-0.9	-1.8		0.4				
H-11	4.93, ddd (11, 11, 5)			-0.9	11.5		12.2	5.1	-0.2	-0.5				
H-12 proR	2.72, m			-0.7	-0.9	12.2		-13.5	11.8	7.2				
H-12 proS	2.17, m				-1.8	5.1	-13.5		6.0	1.0				
H-13 proR	2.53, ddd (13, 13, 6)					-0.2	11.8	6.0		-14.4		-0.4		
H-13 proS	3.26, dd (13, 6)				0.4	-0.5	7.2	1.0	-14.4		-1.2	-1.1	0.6	
H-16	7.53, br. s			-0.4						-1.2		2.0	0.4	
H-18	6.81, br. d (8)								-0.4	-1.1	2.0		8.1	
H-19	6.77, d (8)									0.6	0.4	8.1		

		predicted ^{1}H-^{1}H *J* couplings (Hz)												
	northern unit	H-1'	H-3'	H-7' proR	H-7' proS	H-8'	H-9'	H-10'	H-11'	H-12' proR	H-12' proS	H-13' proR	H-13' proS	H-16'	H-19'
Position	δ_H, mult (*J* in Hz)														
H-1'	6.34, s														
H-3'	4.62, s			-1.6	-0.3										
H-7' proR	2.35, dd (19, 6)		-1.6		-19.7	7.0	-1.6	0.4	-0.8	1.0					
H-7' proS	3.41, br. d (19)		-0.3	-19.7		3.1	-3.4	1.1	-2.2	2.4		0.2			
H-8'	5.60, br. dd (15, 6)			7.0	3.1		15.8	-0.8	1.1	-1.2					
H-9'	5.86, m			-1.6	-3.4	15.8		11.6	-1.2	0.6	-0.5				
H-10'	5.90, m			0.4	1.1	-0.8	11.6		11.2	-2.1	-0.7	0.3			
H-11'	5.34, ddd (11, 11, 5)			-0.8	-2.2	1.1	-1.2	11.2		5.6	11.8	-0.2	-0.2		
H-12' proR	2.28, m			1.0	2.4	-1.2	0.6	-2.1	5.6		-13.9	5.0	2.5		
H-12' proS	2.65, br. quartet (13)						-0.5	-0.7	11.8	-13.9		2.7	13.4	-0.3	
H-13' proR	2.43, br. d (14)				0.2			0.3	-0.2	5.0	2.7		-14.8	-1.2	0.6
H-13' proS	3.01, br. t (14)								-0.2	2.5	13.4	-14.8			
H-16'	6.72, s										-0.3	-1.2			0.3
H-19'	6.50, s											0.6		0.3	

Table 8.15. Rotatory strengths (length formalism) calculated for zosterabisphenone A (**5**). Calculations were performed at the ωB97XD/6-31+G(d,p)//B3LYP/6-31+G(d,p) level.

Wavelength (nm)	Rotatory strength (10^{-40} erg esu cm Gauss^{-1})	Wavelength (nm)	Rotatory strength (10^{-40} erg esu cm Gauss^{-1})
154.2	-5.772300	172.4	12.977700
154.3	1.541200	173.3	28.706500
154.5	3.893400	174.0	-4.519900
154.6	-10.440700	174.4	-27.216800
154.7	-7.697300	175.2	-15.858000
154.8	7.970900	175.8	24.097800
155.2	-31.635200	176.1	59.054500
155.3	12.115700	176.5	67.683100
155.6	-5.555600	176.7	13.526000
156.1	-4.629800	177.1	-0.621700
156.3	18.611000	177.4	-21.616900
156.6	-13.818100	177.7	35.119500
156.7	5.712400	178.0	-36.729400
156.9	1.595000	178.6	28.256100
157.0	-3.321700	179.0	-13.432700
157.3	27.079500	179.6	-3.024600
157.5	22.680600	179.7	-30.905200
157.6	28.708000	180.0	34.459000
158.1	5.242200	180.2	41.971900
158.3	-10.164400	180.9	-5.846900
158.5	-0.567000	181.1	37.356200
159.0	-52.780000	181.4	-4.326600
159.1	-3.250000	181.5	-85.923100
159.2	13.062600	182.9	-6.718500
159.5	12.140200	183.8	-44.258000
159.7	34.046700	184.6	147.774000
159.8	-1.469900	184.8	-40.092300
160.0	10.957300	185.5	-71.196700
160.1	-0.695600	186.0	-19.141000
160.3	8.561600	187.4	92.584100
160.4	-16.036500	187.8	-5.689900
160.8	-55.483000	188.5	12.489100
160.9	-1.154500	188.8	124.362100
161.2	-0.585200	189.3	15.474100
161.5	-27.228300	190.1	32.238800
161.6	4.859500	191.2	-100.354400
161.9	2.391500	192.2	-106.711400
162.2	-52.850900	192.3	-92.401300
162.4	-1.034100	193.8	-45.835500
162.7	5.868300	194.2	7.917400
163.0	7.152700	195.3	-19.843200
163.2	1.120300	195.5	-19.959400
163.5	-0.615500	196.1	45.606000
163.7	-26.574100	197.1	42.327800
164.1	-67.908300	198.2	69.862700
164.4	14.053000	198.4	18.801200
164.6	-23.579100	200.5	52.275300
165.0	15.012400	200.9	86.801900
165.3	-14.487900	202.1	-78.328000
165.7	36.021900	202.7	-5.918400
165.9	19.796600	203.9	223.407200
166.2	56.541300	206.0	-263.886900
166.3	-46.183900	206.4	12.806700
166.6	-1.189200	206.8	-73.070300
166.7	-54.545500	209.0	-26.396900
166.9	-5.216300	210.4	-59.320200
167.2	33.239600	215.1	-64.030700
167.5	8.547700	216.5	14.225400
167.8	-6.360400	217.2	-159.651700
168.2	8.413500	218.5	-24.128500
168.6	-3.269600	219.7	36.298000
168.7	2.910900	225.1	218.617000
169.1	-2.461300	226.8	189.465800
169.4	-39.800000	231.6	-291.876800
169.5	-24.449300	239.3	35.972300
169.7	2.466500	253.4	40.881000
170.0	-8.438600	264.0	-188.955400
170.6	-1.489600	264.7	47.877300
171.3	11.400800	270.4	109.088900
171.4	33.352500	272.9	-100.810700
171.5	-0.981200	275.0	-8.082200
172.0	-3.698200	285.8	3.111100
172.2	-1.949900	315.4	30.210900

Table 8.16. Rotatory strengths (length formalism) calculated for the two conformers of zosterabisphenone B (**6**). Calculations were performed at the ωB97XD/6-31+G(d,p)//B3LYP/6-31+G(d,p) level.

Conformer 1, Δ*E* = 0.00 kcal/mol, population 85.4%		Conformer 2, Δ*E* = 0.91 kcal/mol, population 14.6%	
Wavelength (nm)	Rotatory strength (10^{-40} erg esu cm Gauss^{-1})	Wavelength (nm)	Rotatory strength (10^{-40} erg esu cm Gauss^{-1})
157.4	10.695000	157.5	7.422900
157.5	7.878900	157.7	23.052500
157.7	30.015100	157.8	-7.352800
157.9	-3.903700	158.0	-8.935200
158.1	11.341500	158.2	20.811200
158.2	-6.628300	158.5	-20.022900
158.3	1.738600	158.8	-2.355600
158.4	-2.726700	159.0	13.069900
158.6	6.436900	159.1	-36.318800
158.9	4.174900	159.4	10.940500
159.2	-29.048700	159.5	-20.836100
159.4	-9.213800	159.8	-1.502500
159.7	27.055500	160.1	-18.692300
159.8	26.716500	160.4	-6.776300
160.0	-12.860000	160.5	22.978000
160.3	-0.881700	160.6	7.936000
160.4	-8.400100	160.8	-10.933100
160.5	1.422000	161.0	-1.208600
160.8	0.814900	161.3	-3.707200
161.1	-25.898000	161.6	-9.683800
161.4	-3.591100	162.0	-1.269500
161.6	-23.514400	162.3	13.831800
161.8	15.463200	162.4	-4.985100
161.9	-0.227900	162.5	5.799800
162.0	-9.241400	162.9	-6.869100
162.3	-6.218800	163.0	23.561200
162.5	-9.932300	163.2	-12.109200
162.7	-43.780200	163.3	10.420800
163.1	1.234600	163.5	-20.024700
163.3	-4.189000	163.7	46.345400
163.5	19.068100	164.0	-32.850400
163.7	18.930500	164.2	0.458100
163.8	-6.920200	164.4	-10.645800
164.0	13.454300	164.6	-15.024300
164.4	2.465900	165.0	-7.340500
164.8	8.656900	165.3	5.918500
164.9	-18.239300	165.6	6.680500
165.4	6.317600	165.9	-4.114300
165.6	0.834500	166.1	-0.437000
165.9	9.111500	166.2	-1.458000
166.2	3.212300	166.8	-7.150100
166.4	-13.012500	167.2	-24.239600
166.6	7.802100	167.3	-7.370600
166.8	7.808700	167.5	11.572800
167.0	-44.886300	167.6	-9.866600
167.2	-6.793700	168.1	10.924400
167.5	-8.949700	168.4	-8.839600
167.7	34.076400	168.5	9.612900
168.1	17.703500	168.8	-9.111300
168.3	-24.049300	169.2	-21.623000
168.4	-4.713900	169.3	11.535900
168.9	-14.873200	169.5	17.727300
169.2	-7.410100	169.7	-2.586700
169.8	-2.649200	169.8	2.877600
170.0	34.952300	170.0	-50.819800
170.4	4.109300	170.5	-22.541100
170.6	-33.338600	170.7	25.009300
170.9	27.664900	171.0	-9.545700
171.3	-5.784700	171.4	12.855000
171.4	-15.723300	171.9	-20.441500
171.6	1.438100	172.3	14.726900
171.8	-8.947900	172.4	4.285000
172.2	-0.131700	172.7	3.055100
172.6	-22.696000	172.8	6.293900
173.0	3.991700	173.4	36.684300
173.4	-7.924000	174.1	-11.563700
173.8	0.438800	174.3	-19.232200
174.2	9.996100	174.4	4.563300
174.6	29.170500	174.8	11.591200
174.8	-8.373000	175.1	21.195200
175.2	58.104100	175.7	-13.861400
175.5	-8.174000	175.9	-21.134200
176.4	15.202300	176.1	25.709900
176.8	-19.474500	176.2	43.046500
177.3	0.428600	177.2	-8.758200
177.4	3.756200	177.5	28.682400

177.6	13.496000	177.8	-40.761700
178.2	-12.263900	178.2	-40.775300
179.0	105.205200	178.7	40.186100
179.4	-65.625200	179.0	25.232700
179.8	0.148300	179.2	20.465300
180.4	-53.543800	179.5	-3.996400
180.6	-16.170300	179.7	-29.328800
181.0	-16.126800	179.8	13.125100
181.3	50.062300	179.9	-0.688800
181.5	-34.422100	180.3	47.758900
181.8	16.247400	180.4	-47.865300
182.9	145.109900	181.0	24.699500
183.4	70.388500	181.2	21.442700
183.6	-6.940200	181.4	-4.414000
184.2	-34.444600	182.4	8.346900
184.7	-66.281200	183.0	-9.676700
185.2	-175.027600	183.6	-7.660700
185.5	-2.387500	184.3	59.696200
186.3	76.099000	185.2	-94.917300
186.6	46.920000	185.4	-4.499400
186.8	13.026400	186.6	241.948800
187.2	31.558300	187.2	-155.501700
187.5	71.671100	187.8	-22.002300
188.8	123.496000	188.4	12.366100
189.1	-117.271600	188.7	123.844800
189.5	126.382600	189.0	-52.067100
190.4	-11.592400	189.6	67.812700
192.3	-29.985900	190.1	-61.410100
192.9	64.312900	192.1	28.899900
194.0	-44.472300	193.1	-14.705800
194.5	34.931600	193.5	-122.324700
195.1	-9.721300	194.7	-38.745800
196.0	25.222300	195.0	-67.400900
196.2	-2.360700	195.4	150.982300
197.0	43.157100	197.0	-39.845500
198.1	0.926000	197.8	-39.257900
198.8	-111.100200	198.6	-54.483800
199.8	19.194400	199.6	115.692700
200.9	-9.548400	200.0	188.851600
201.9	-11.150300	201.0	28.754800
202.1	0.291400	201.5	22.303200
203.8	-54.910800	202.4	-28.152500
204.0	-45.061800	204.3	29.945900
205.3	-4.543600	204.9	-8.621600
206.4	9.851500	205.3	-66.458900
207.7	-32.310600	205.6	-45.324000
208.0	-130.152500	207.0	-15.682100
211.0	-21.341500	208.4	26.425000
212.0	-8.066200	208.5	14.180200
212.4	2.229500	210.8	31.262700
214.9	-109.322300	211.7	-248.795100
217.0	-0.993200	211.8	-36.374900
218.2	-74.338500	213.8	-42.555200
219.2	-22.882900	215.7	-123.511200
222.5	-46.927700	217.6	-185.555200
225.8	195.291500	218.1	49.057400
226.5	-136.384100	219.3	-9.291700
230.6	20.305800	224.2	30.399500
232.0	439.674400	225.9	-20.845800
233.7	-126.587000	228.3	2.441500
239.4	-2.924500	232.4	723.136500
244.4	-94.636100	237.3	-363.145000
257.5	-42.293900	239.5	-13.833600
263.9	-10.720800	258.3	-67.448600
268.8	-136.052500	265.6	-11.359900
270.6	109.171600	269.8	-32.713900
277.0	19.384500	273.5	48.933000
278.1	1.933700	276.2	-64.725700
285.2	-119.194000	279.6	50.076700
292.5	18.452100	286.2	-2.397900
319.5	86.494600	294.7	-32.153400
		332.2	81.472400

Table 8.17. Rotatory strengths (length formalism) calculated for the two conformers of zosterabisphenone B (**6**). Calculations were performed at the ωB97XD/6-31+G(d,p)//B3LYP/6-31+G(d,p) level.

Conformer 1, Δ*E* = 0.00 kcal/mol, population 85.4%		Conformer 2, Δ*E* = 0.91 kcal/mol, population 14.6%	
Wavelength (nm)	Rotatory strength (10^{-40} erg esu cm Gauss^{-1})	Wavelength (nm)	Rotatory strength (10^{-40} erg esu cm Gauss^{-1})
157.4	10.695000	157.5	7.422900
157.5	7.878900	157.7	23.052500
157.7	30.015100	157.8	-7.352800
157.9	-3.903700	158.0	-8.935200
158.1	11.341500	158.2	20.811200
158.2	-6.628300	158.5	-20.022900
158.3	1.738600	158.8	-2.355600
158.4	-2.726700	159.0	13.069900
158.6	6.436900	159.1	-36.318800
158.9	4.174900	159.4	10.940500
159.2	-29.048700	159.5	-20.836100
159.4	-9.213800	159.8	-1.502500
159.7	27.055500	160.1	-18.692300
159.8	26.716500	160.4	-6.776300
160.0	-12.860000	160.5	22.978000
160.3	-0.881700	160.6	7.936000
160.4	-8.400100	160.8	-10.933100
160.5	1.422000	161.0	-1.208600
160.8	0.814900	161.3	-3.707200
161.1	-25.898000	161.6	-9.683800
161.4	-3.591100	162.0	-1.269500
161.6	-23.514400	162.3	13.831800
161.8	15.463200	162.4	-4.985100
161.9	-0.227900	162.5	5.799800
162.0	-9.241400	162.9	-6.869100
162.3	-6.218800	163.0	23.561200
162.5	-9.932300	163.2	-12.109200
162.7	-43.780200	163.3	10.420800
163.1	1.234600	163.5	-20.024700
163.3	-4.189000	163.7	46.345400
163.5	19.068100	164.0	-32.850400
163.7	18.930500	164.2	0.458100
163.8	-6.920200	164.4	-10.645800
164.0	13.454300	164.6	-15.024300
164.4	2.465900	165.0	-7.340500
164.8	8.656900	165.3	5.918500
164.9	-18.239300	165.6	6.680500
165.4	6.317600	165.9	-4.114300
165.6	0.834500	166.1	-0.437000
165.9	9.111500	166.2	-1.458000
166.2	3.212300	166.8	-7.150100
166.4	-13.012500	167.2	-24.239600
166.6	7.802100	167.3	-7.370600
166.8	7.808700	167.5	11.572800
167.0	-44.886300	167.6	-9.866600
167.2	-6.793700	168.1	10.924400
167.5	-8.949700	168.4	-8.839600
167.7	34.076400	168.5	9.612900
168.1	17.703500	168.8	-9.111300
168.3	-24.049300	169.2	-21.623000
168.4	-4.713900	169.3	11.535900
168.9	-14.873200	169.5	17.727300
169.2	-7.410100	169.7	-2.586700
169.8	-2.649200	169.8	2.877600
170.0	34.952300	170.0	-50.819800
170.4	4.109300	170.5	-22.541100
170.6	-33.338600	170.7	25.009300
170.9	27.664900	171.0	-9.545700
171.3	-5.784700	171.4	12.855000
171.4	-15.723300	171.9	-20.441500
171.6	1.438100	172.3	14.726900
171.8	-8.947900	172.4	4.285000
172.2	-0.131700	172.7	3.055100
172.6	-22.696000	172.8	6.293900
173.0	3.991700	173.4	36.684300
173.4	-7.924000	174.1	-11.563700
173.8	0.438800	174.3	-19.232200
174.2	9.996100	174.4	4.563300
174.6	29.170500	174.8	11.591200
174.8	-8.373000	175.1	21.195200
175.2	58.104100	175.7	-13.861400
175.5	-8.174000	175.9	-21.134200
176.4	15.202300	176.1	25.709900
176.8	-19.474500	176.2	43.046500
177.3	0.428600	177.2	-8.758200
177.4	3.756200	177.5	28.682400
177.6	13.496000	177.8	-40.761700
178.2	-12.263900	178.2	-40.775300

179.0	105.205200
179.4	-65.625200
179.8	0.148300
180.4	-53.543800
180.6	-16.170300
181.0	-16.126800
181.3	50.062300
181.5	-34.422100
181.8	16.247400
182.9	145.109900
183.4	70.388500
183.6	-6.940200
184.2	-34.444600
184.7	-66.281200
185.2	-175.027600
185.5	-2.387500
186.3	76.099000
186.6	46.920000
186.8	13.026400
187.2	31.558300
187.5	71.671100
188.8	123.496000
189.1	-117.271600
189.5	126.382600
190.4	-11.592400
192.3	-29.985900
192.9	64.312900
194.0	-44.472300
194.5	34.931600
195.1	-9.721300
196.0	25.222300
196.2	-2.360700
197.0	43.157100
198.1	0.926000
198.8	-111.100200
199.8	19.194400
200.9	-9.548400
201.9	-11.150300
202.1	0.291400
203.8	-54.910800
204.0	-45.061800
205.3	-4.543600
206.4	9.851500
207.7	-32.310600
208.0	-130.152500
211.0	-21.341500
212.0	-8.066200
212.4	2.229500
214.9	-109.322300
217.0	-0.993200
218.2	-74.338500
219.2	-22.882900
222.5	-46.927700
225.8	195.291500
226.5	-136.384100
230.6	20.305800
232.0	439.674400
233.7	-126.587000
239.4	-2.924500
244.4	-94.636100
257.5	-42.293900
263.9	-10.720800
268.8	-136.052500
270.6	109.171600
277.0	19.384500
278.1	1.933700
285.2	-119.194000
292.5	18.452100
319.5	86.494600

178.7	40.186100
179.0	25.232700
179.2	20.465300
179.5	-3.996400
179.7	-29.328800
179.8	13.125100
179.9	-0.688800
180.3	47.758900
180.4	-47.865300
181.0	24.699500
181.2	21.442700
181.4	-4.414000
182.4	8.346900
183.0	-9.676700
183.6	-7.660700
184.3	59.696200
185.2	-94.917300
185.4	-4.499400
186.6	241.948800
187.2	-155.501700
187.8	-22.002300
188.4	12.366100
188.7	123.844800
189.0	-52.067100
189.6	67.812700
190.1	-61.410100
192.1	28.899900
193.1	-14.705800
193.5	-122.324700
194.7	-38.745800
195.0	-67.400900
195.4	150.982300
197.0	-39.845500
197.8	-39.257900
198.6	-54.483800
199.6	115.692700
200.0	188.851600
201.0	28.754800
201.5	22.303200
202.4	-28.152500
204.3	29.945900
204.9	-8.621600
205.3	-66.458900
205.6	-45.324000
207.0	-15.682100
208.4	26.425000
208.5	14.180200
210.8	31.262700
211.7	-248.795100
211.8	-36.374900
213.8	-42.555200
215.7	-123.511200
217.6	-185.555200
218.1	49.057400
219.3	-9.291700
224.2	30.399500
225.9	-20.845800
228.3	2.441500
232.4	723.136500
237.3	-363.145000
239.5	-13.833600
258.3	-67.448600
265.6	-11.359900
269.8	-32.713900
273.5	48.933000
276.2	-64.725700
279.6	50.076700
286.2	-2.397900
294.7	-32.153400
332.2	81.472400

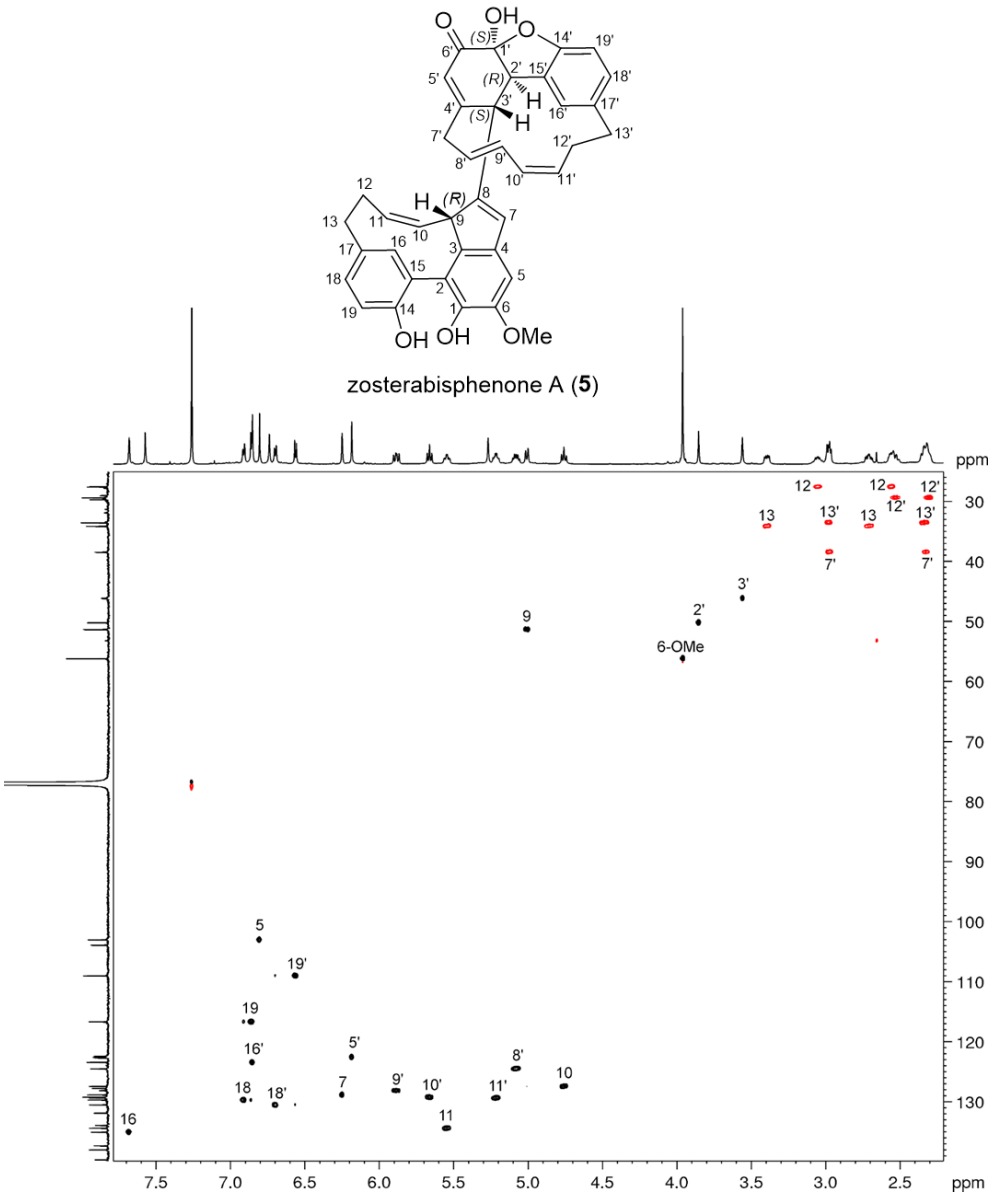


Figure 8.24. HSQC spectrum of zosterabisphenone A (**5**) recorded at 253 K (700 MHz, $CDCl_3$)

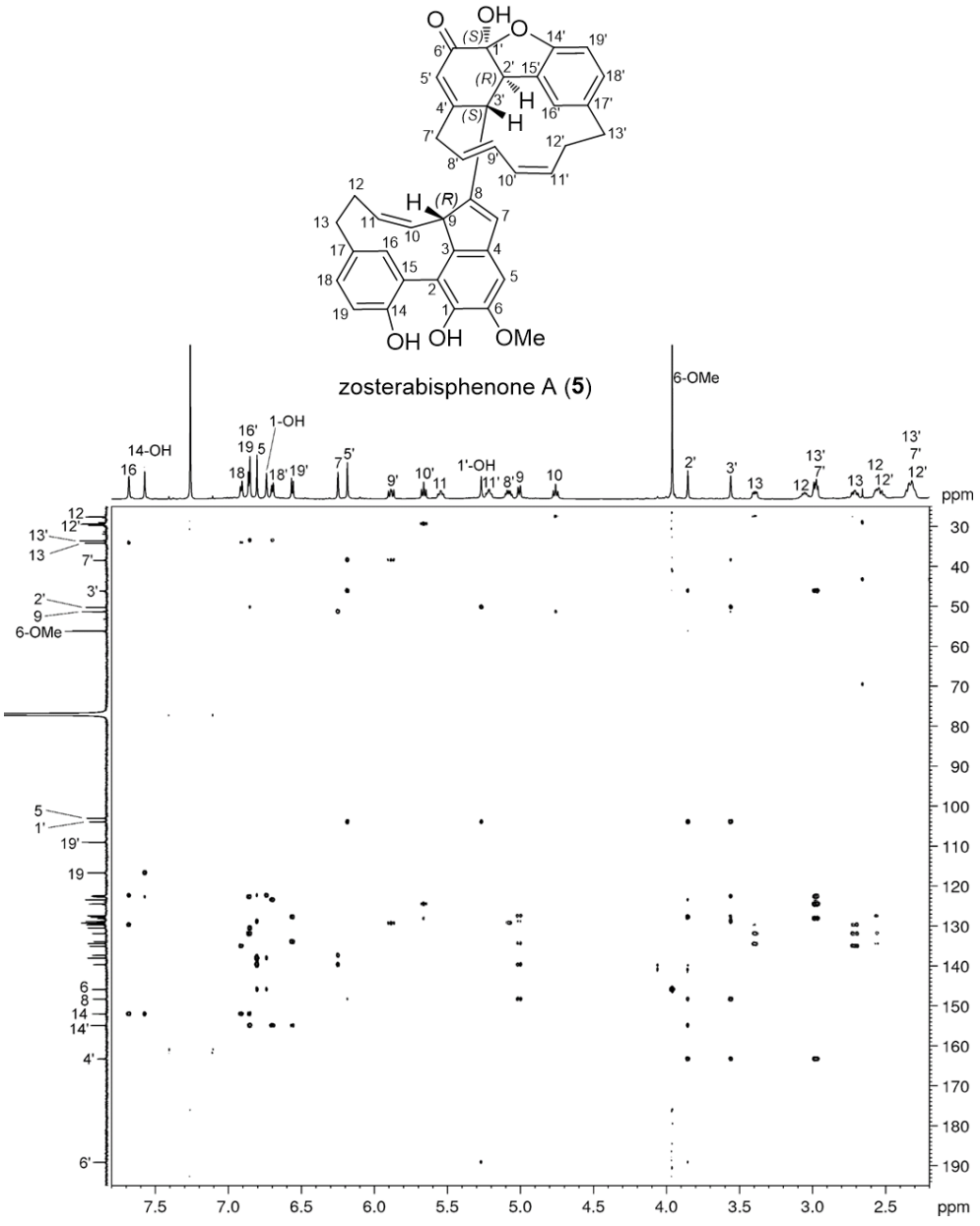


Figure 8.25. HMBC spectrum of zosterabisphenone A (**5**) recorded at 253 K (700 MHz, $CDCl_3$)

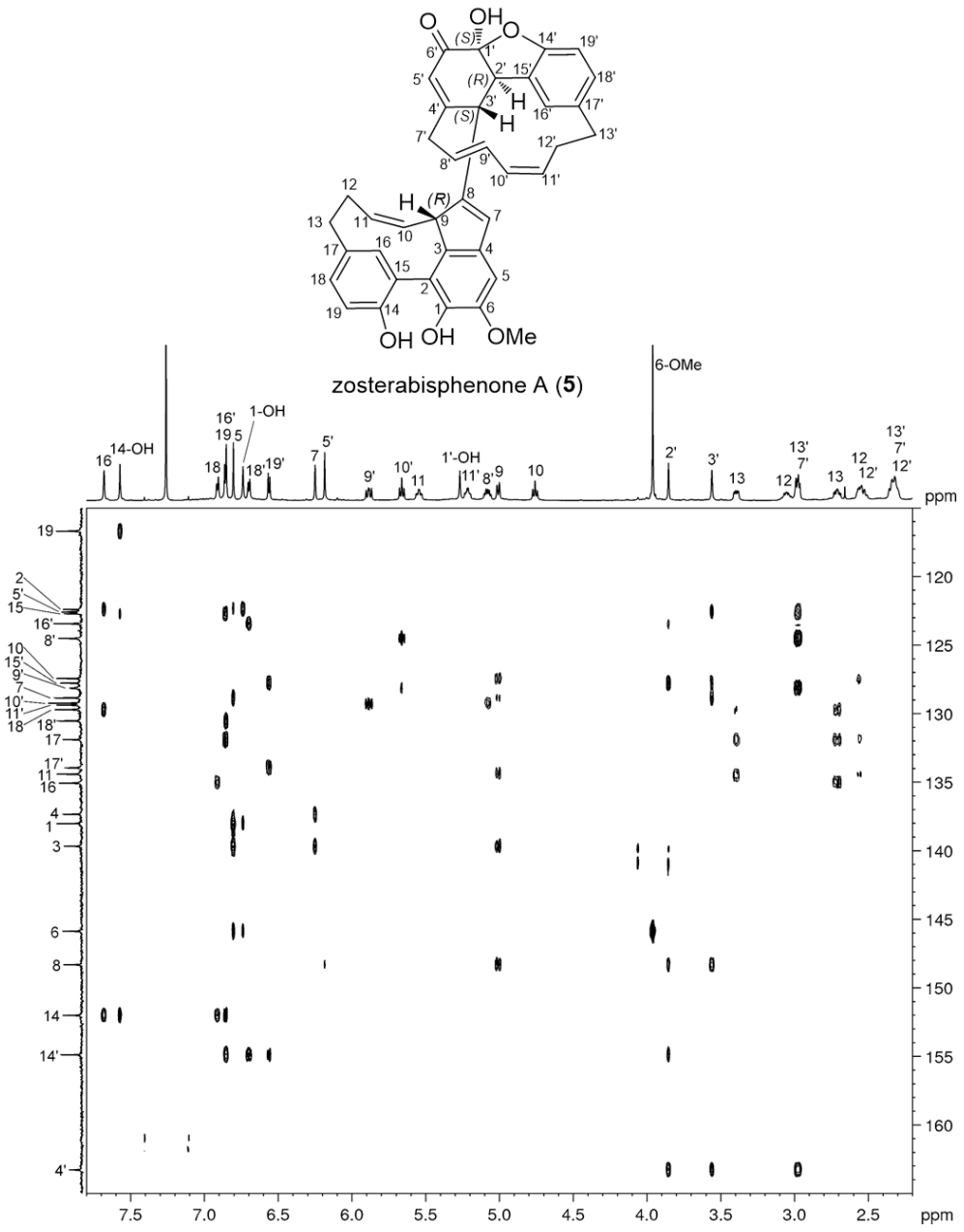


Figure 8.26. Expansion of the HMBC spectrum of zosterabisphenone A (**5**) recorded at 253 K (700 MHz, $CDCl_3$)

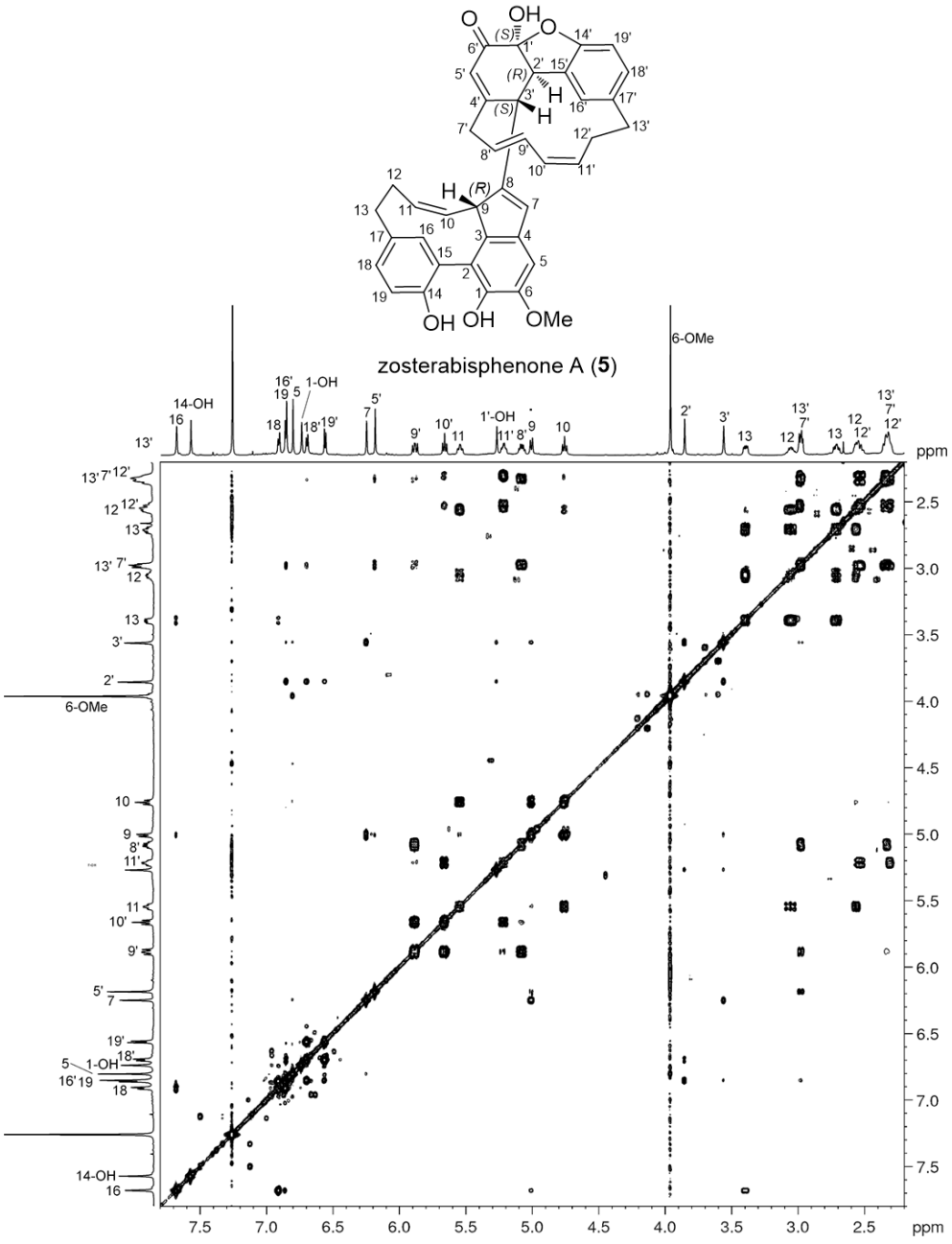


Figure 8.27. COSY spectrum of zosterabisphenone A (**5**) recorded at 253 K (700 MHz, $CDCl_3$)

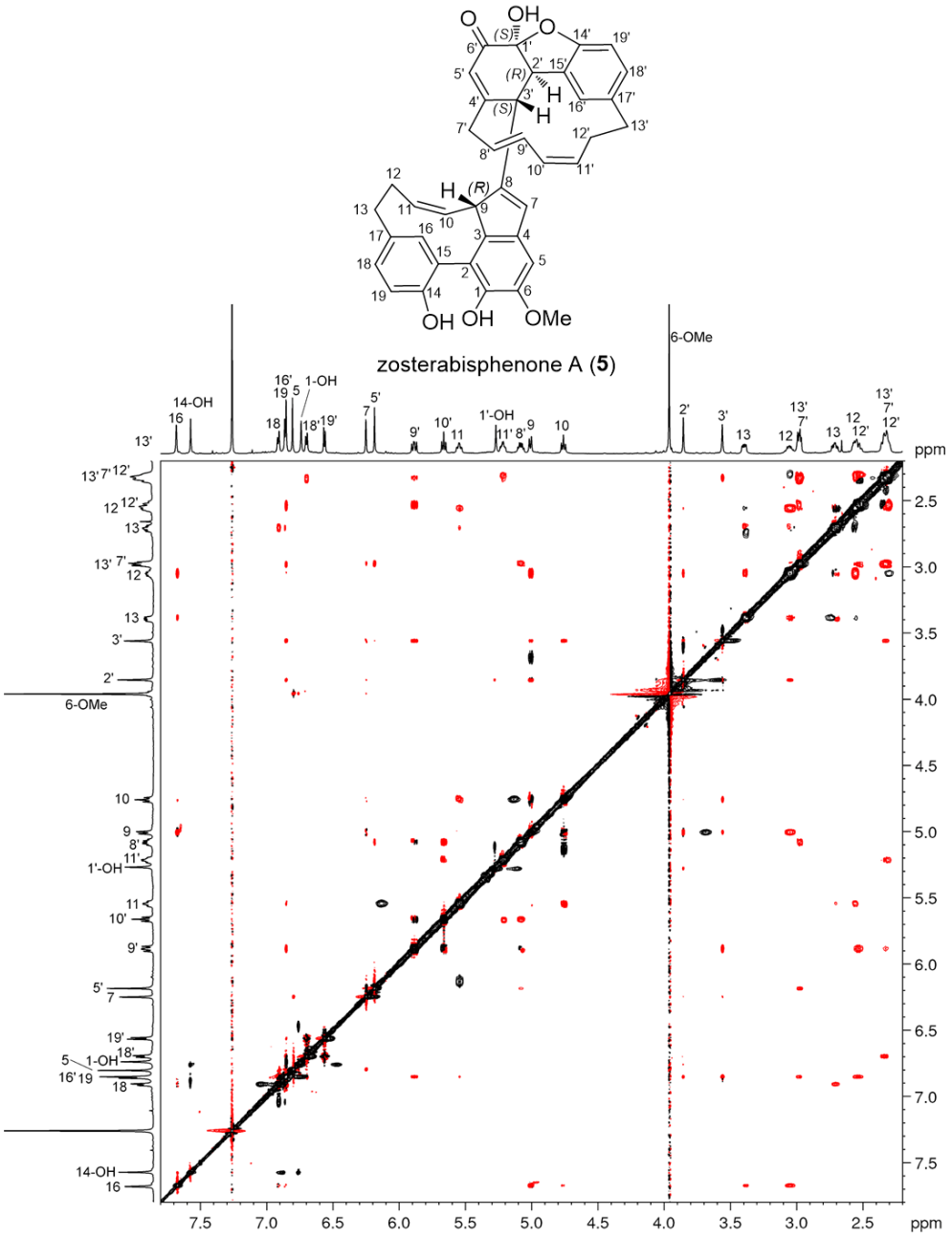


Figure 8.28. ROESY spectrum of zosterabisphenone A (**5**) recorded at 253 K (700 MHz, $CDCl_3$)

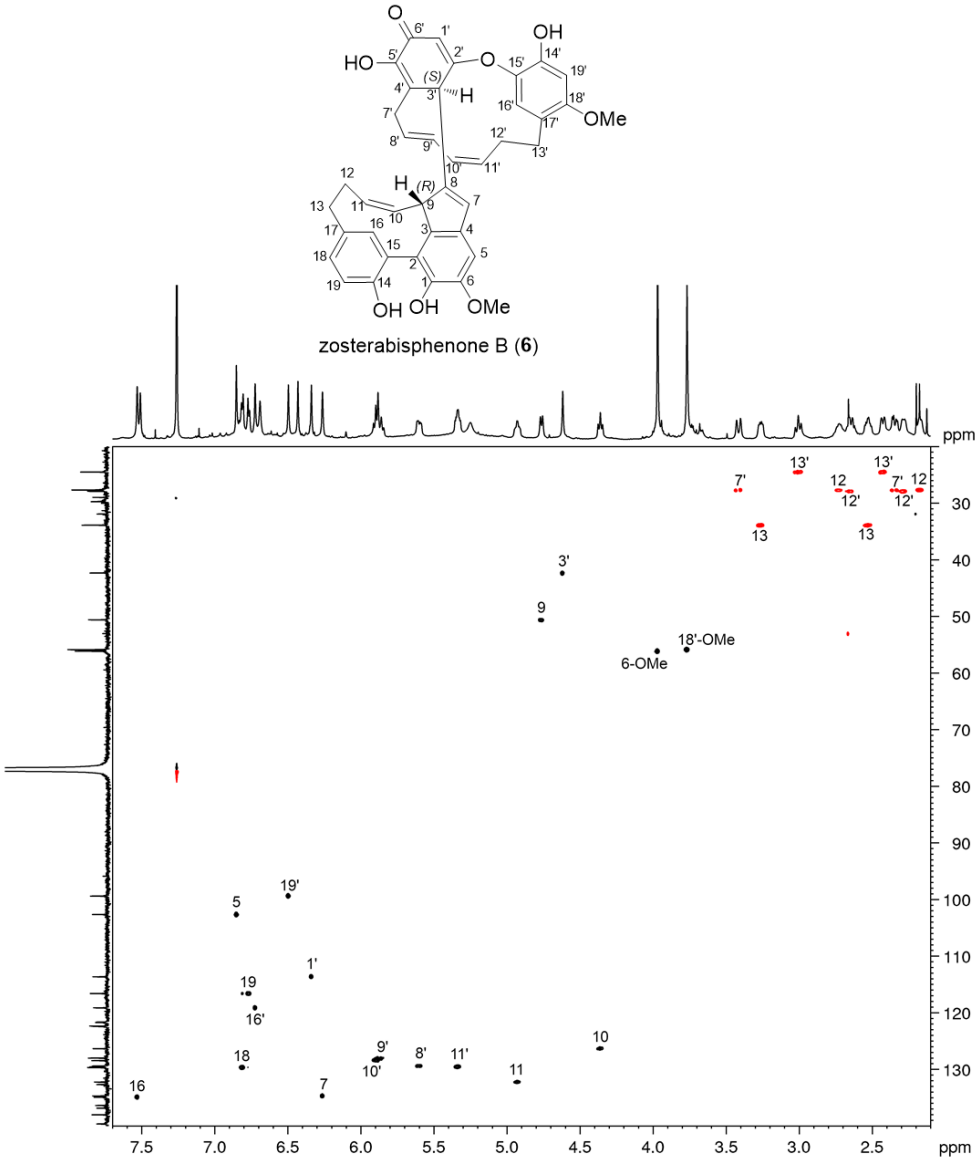


Figure 8.29. HSQC spectrum of zosterabisphenone B (**6**) recorded at 238 K (700 MHz, $CDCl_3$)

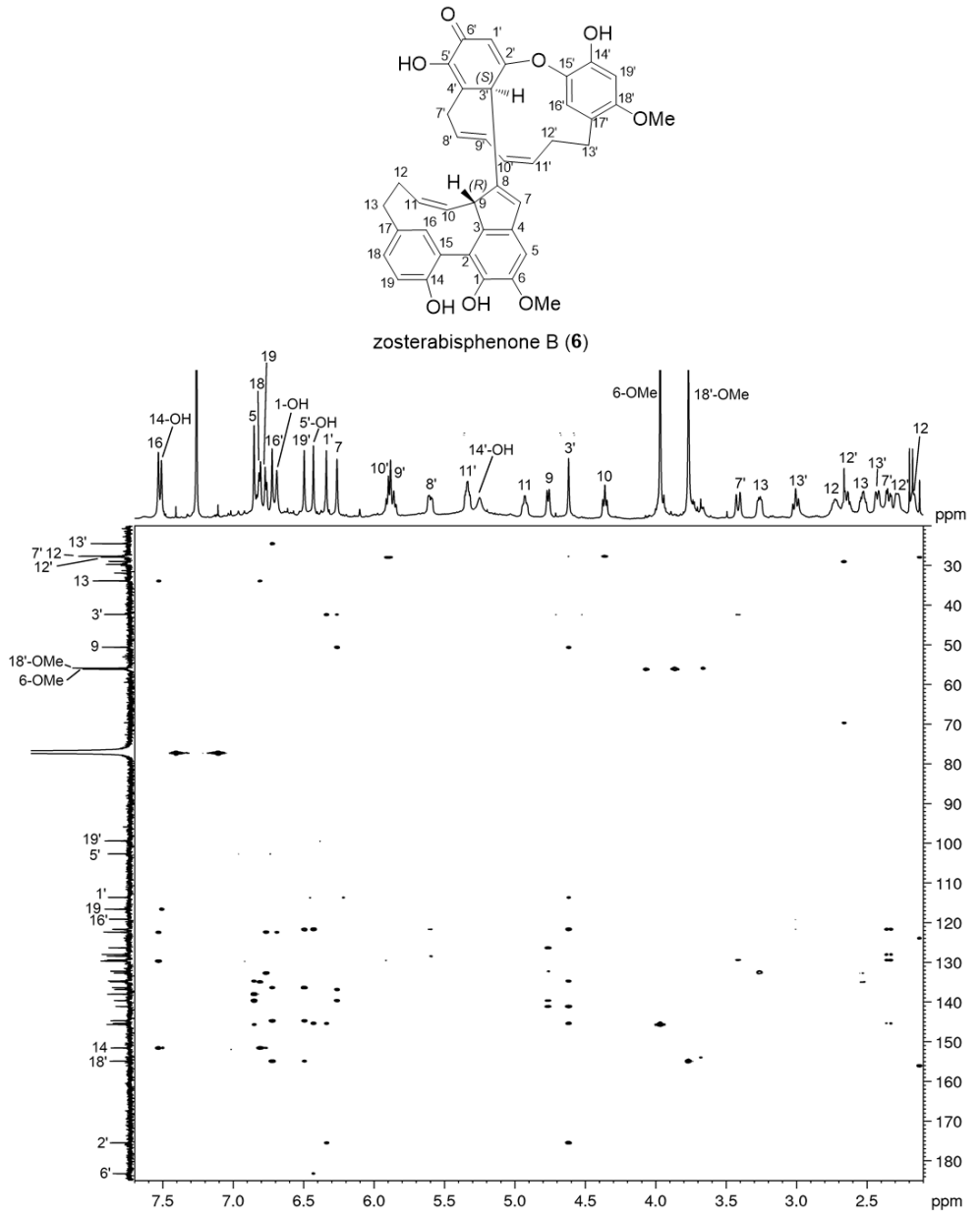


Figure 8.30. HMBC spectrum of zosterabisphenone B (**6**) recorded at 238 K (700 MHz, $CDCl_3$).

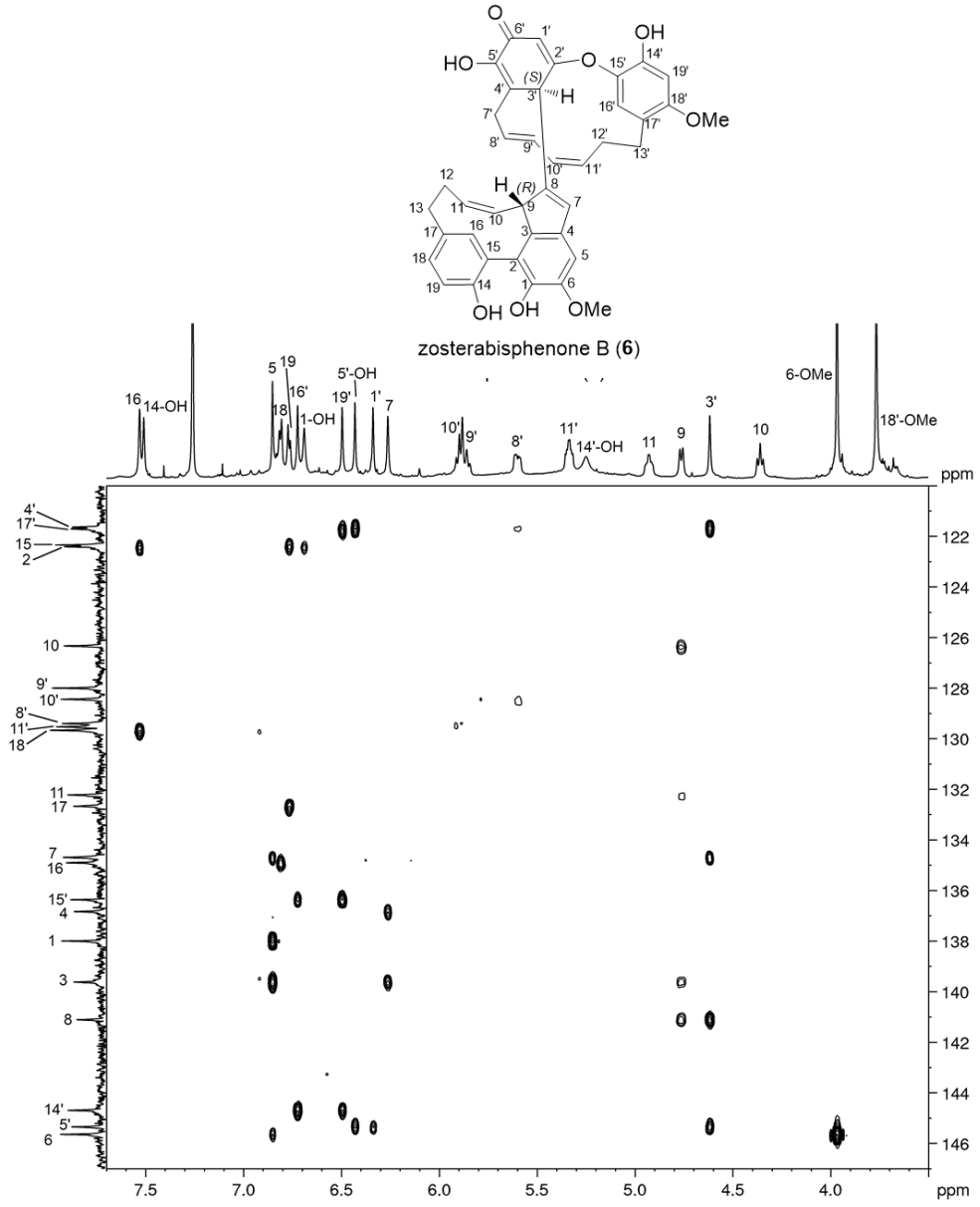


Figure 8.31. Expansion of the HMBC spectrum of zosterabisphenone B (**6**) recorded at 238 K (700 MHz, $CDCl_3$)

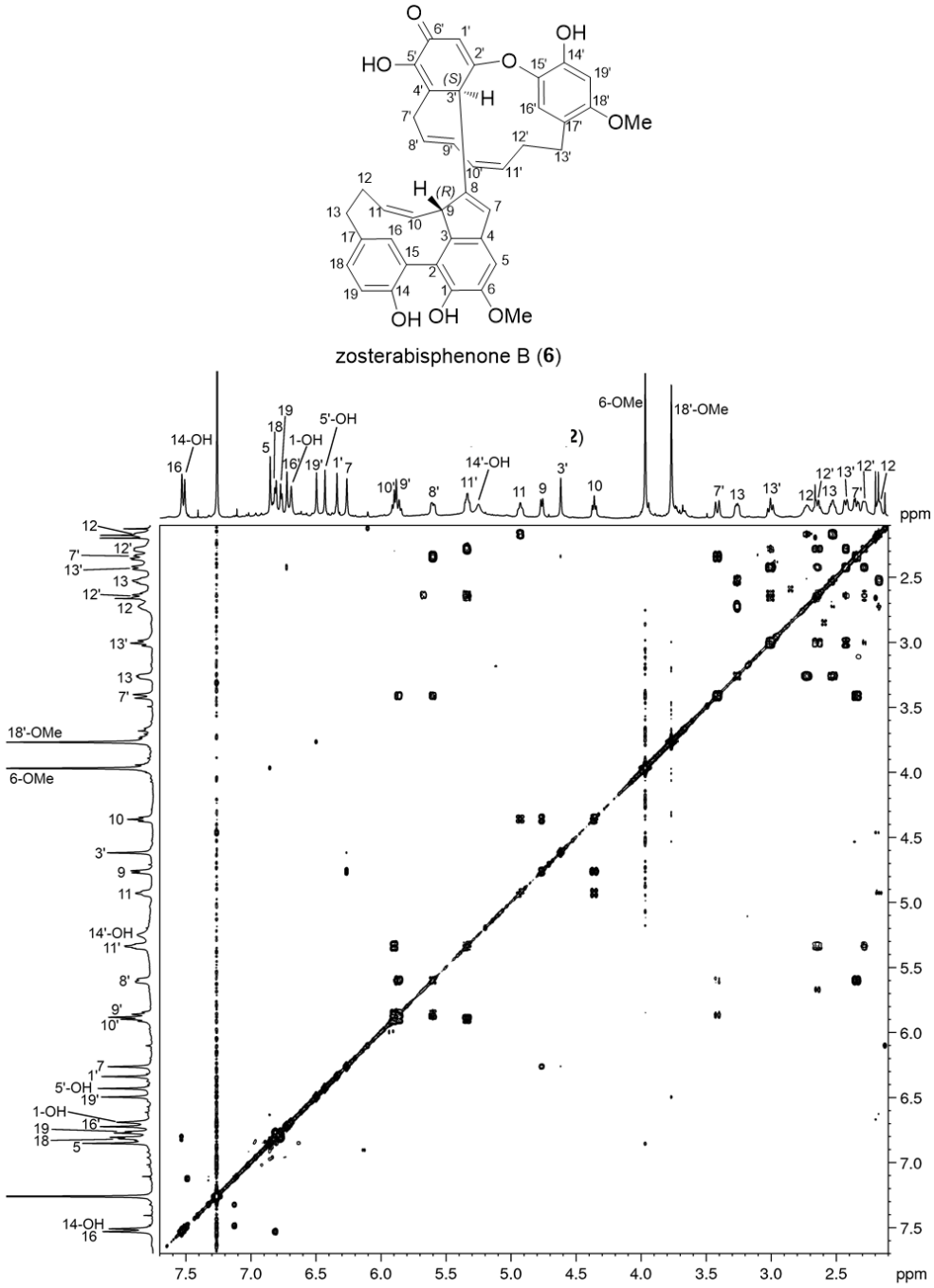


Figure 8.32 COSY spectrum of zosterabisphenone B (**6**) recorded at 238 K (700 MHz, $CDCl_3$)

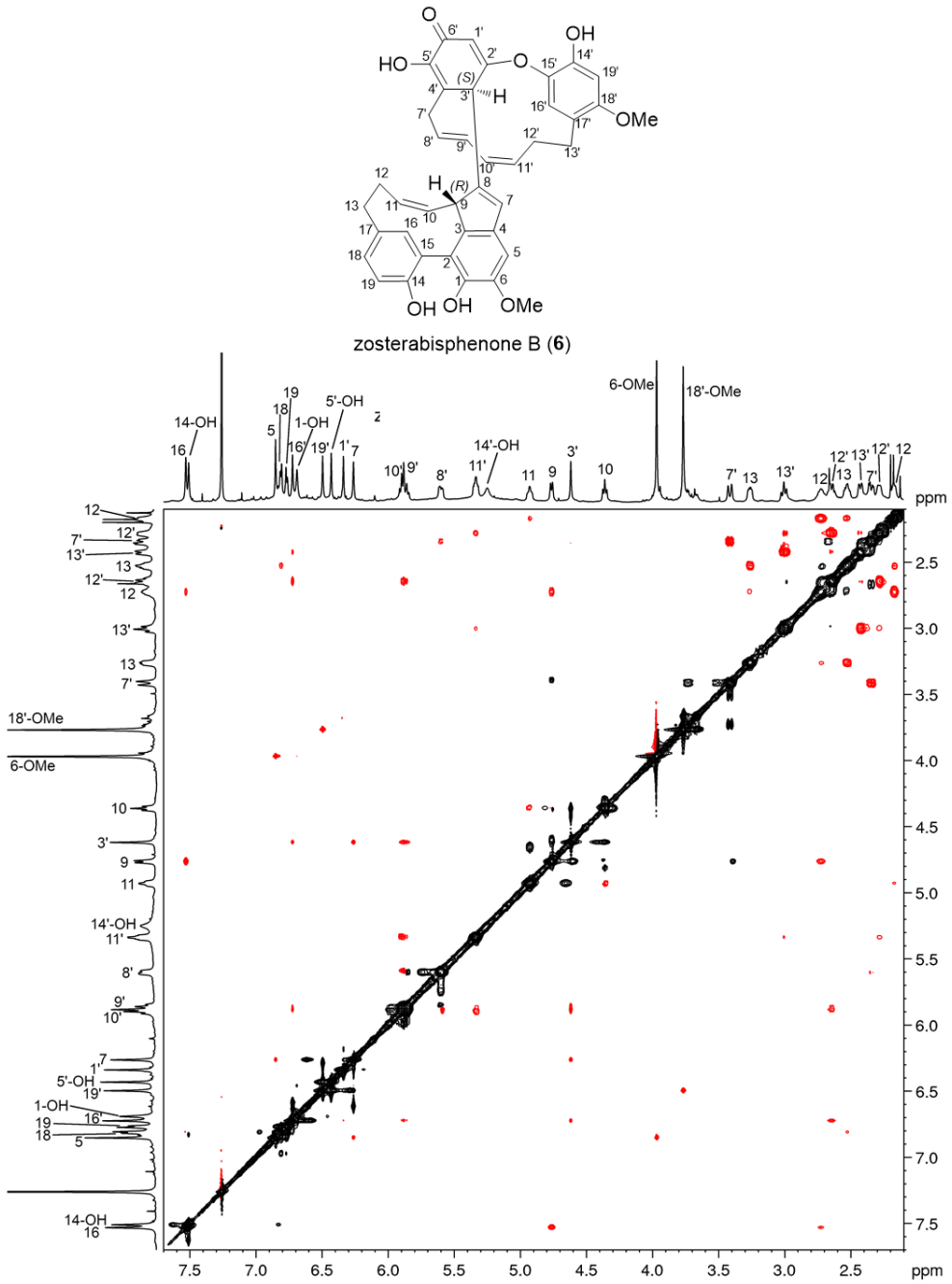


Figure 8.33. ROESY spectrum of zosterabisphenone B (**6**) recorded at 238 K (700 MHz, $CDCl_3$)

References

[1] Li, Y.; Grauso, L.; Scarpato, S.; Cacciola, N.A.; Borrelli, F.; Zidorn, C.; Mangoni, A. Stable Catechol Keto Tautomers in Cytotoxic Heterodimeric Cyclic Diarylheptanoids from the Seagrass Zostera marina. *Org. Lett.* **2021**, *23*, 7134–7138, doi:10.1021/acs.orglett.1c02537.

[2] Li, Y.; Mangoni, A.; Shulha, O.; Çiçek, S.S.; Zidorn, C. Cyclic diarylheptanoids deoxycymodienol and isotedarene A from Zostera marina (Zosteraceae). *Tetrahedron Lett.* **2019**, *60*, 150930, doi:10.1016/j.tetlet.2019.07.021.

[3] Grauso, L.; Li, Y.; Scarpato, S.; Shulha, O.; Rárová, L.; Strnad, M.; Teta, R.; Mangoni, A.; Zidorn, C. Structure and Conformation of Zosteraphenols, Tetracyclic Diarylheptanoids from the Seagrass Zostera marina: An NMR and DFT Study. *Org. Lett.* **2020**, *22*, 78–82, doi:10.1021/acs.orglett.9b03964.

[4] Lodewyk, M.W.; Siebert, M.R.; Tantillo, D.J. Computational prediction of 1H and 13C chemical shifts: A useful tool for natural product, mechanistic, and synthetic organic chemistry. *Chem. Rev.* 2012, *112*, 1839–1862.

[5] Grimblat, N.; Zanardi, M.M.; Sarotti, A.M. Beyond DP4: An Improved Probability for the Stereochemical Assignment of Isomeric Compounds using Quantum Chemical Calculations of NMR Shifts. *J. Org. Chem.* **2015**, *80*, 12526–12534, doi:10.1021/acs.joc.5b02396.

[6] Bally, T.; Rablen, P.R. Quantum-chemical simulation of 1H NMR spectra. 2. Comparison of DFT-based procedures for computing proton-proton coupling constants in organic molecules. *J. Org. Chem.* **2011**, *76*, 4818–4830, doi:10.1021/jo200513q.

[7] Grauso, L.; Teta, R.; Esposito, G.; Menna, M.; Mangoni, A. Computational prediction of chiroptical properties in structure elucidation of natural products. *Nat. Prod. Rep.* **2019**, *36*, 1005–1030, doi:10.1039/c9np00018f.

[8] Pescitelli, G.; Bruhn, T. Good Computational Practice in the Assignment of Absolute Configurations by TDDFT Calculations of ECD Spectra. *Chirality* **2016**, *28*, 466–474, doi:10.1002/chir.22600.

[9] 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. et al. Gaussian 16, Revision A.03 2016.

[10] Singldinger, B.; Dunkel, A.; Bahmann, D.; Bahmann, C.; Kadow, D.; Bisping, B.; Hofmann, T. New Taste-Active 3-(O-β- d -Glucosyl)-2-oxoindole-3-acetic Acids and Diarylheptanoids in Cimiciato-Infected Hazelnuts. *J. Agric. Food Chem.* **2018**, *66*, 4662–4673, doi:10.1021/acs.jafc.8b01216.

[11] Cacciola, N.A.; Squillaci, G.; D'Apolito, M.; Petillo, O.; Veraldi, F.; Cara, F. La; Peluso, G.; Margarucci, S.; Morana, A. Castanea sativa Mill. Shells aqueous extract exhibits anticancer properties inducing cytotoxic and pro-apoptotic effects. *Molecules* **2019**, *24*, 1–12, doi:10.3390/molecules24183401.

[12] Marcarino, M.O.; Zanardi, M.M.; Cicetti, S.; Sarotti, A.M. NMR calculations with quantum methods: Development of new tools for structural elucidation and

beyond. *Acc. Chem. Res.* **2020**, *53*, 1922–1932, doi:10.1021/acs.accounts.0c00365

Chapter 9

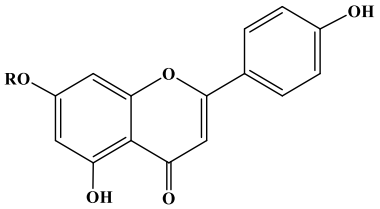
Structure and stereochemistry elucidation of a new phenolic acid from the seagrass *Zostera marina*

9.1 Synopsis of this chapter [1]

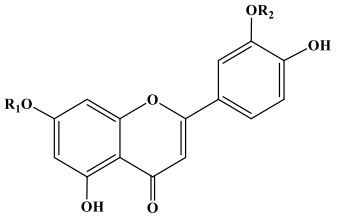
In the framework of a study concerning the full investigation of the phenolic profile of leaves of *Zostera marina* collected along the Baltic Sea coast, and the evaluation of its seasonal variability, a new secondary metabolite was identified. The extraction and the isolation steps of the 7″,8″-didehydrosalvianolic acid B (**7**), were performed in the laboratories of Kiel University under the supervision of Professor Christian Zidorn, while the complete elucidation of its planar structure and stereochemistry were carried out in our laboratory. In particular, the absolute configuration of the new compound was thoroughly assigned by the comparison of the ECD spectrum of methyl (3,4-dihydroxyphenyl)lactate (**8**) obtained from the methanolysis of rosmarinic acid (**9**), whose absolute configuration was known from previous studies, and methyl (3,4-dihydroxyphenyl) lactate (**8**) produced by the degradation of compound **7**.

9.2 Phenolic composition of Zostera marina

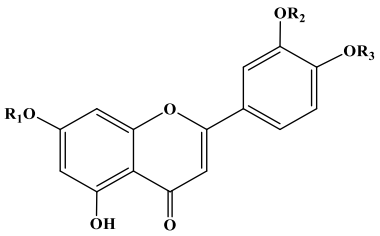
In the perspective to characterize the complete range of phenolics of *Z. marina* from the Baltic Sea and to clarify the variability in flavonoid and phenolic acid composition, a quantity of leaves was extracted with a mixture of H_2O: MeOH (1:1). The data recorded by UHPLC-DAD-MS of the extract demonstrated the presence of a large pool of compounds (four phenolic acids and eight flavonoids) already known from the literature,[2,3] previously isolated in other *Zostera* species. The composition of flavonoids in *Z. marina* collected along the German coast was less variable than the composition of leaves collected in the North Sea off Norway.[4] Moreover, both contained specific disulfated flavonoids, not known from other areas of the range of *Z. marina*. (luteolin 7,3′-*O*-disulfate, diosmetin 7,3′-O-disulfate), while three phenolic acids were previously reported in *Zostera marina* (zosteric acid; caffeic acid; rosmarinic acid) [11,12]



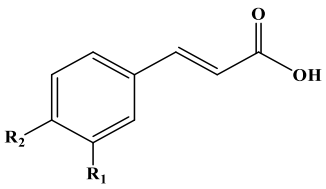
apigenin	R = H	
apigenin 7-*O*-sulfate	R = SO_3H	
apigenin 7-*O*-glucoside	R = GLC	
apigenin 7-*O*-(6″-malonyl) glucoside	R = MAL-6-*O*-GLC	



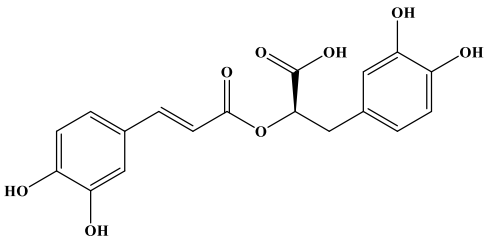
luteolin	R_1 = H	R_2 = H
luteolin 3′-*O*-sulfate	R_1 = H	R_2 = SO_3H
luteolin 7-*O*-glucoside	R_1 = GLC	R_2 = H
luteolin 7-*O*-(6″-malonyl) glucoside	R_1 = MAL-6-*O*-GLC	R_2 = H
luteolin 7-*O*-sulfate	R_1 = SO_3H	R_2 = H



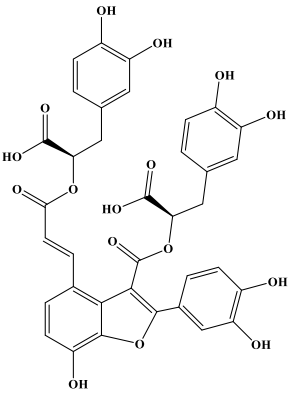
diosmetin	R_1 = H	R_2 = H	R_3 = CH_3
chrysoeriol	R_1 = H	R_2 = CH_3	R_3 = H
diosmetin 3′-*O*-sulfate	R_1 = H	R_2 = SO_3H	R_3 = CH_3
diosmetin 7-*O*-(6″-malonyl) glucoside	R_1 = MAL-6-*O*-GLC	R_2 = H	R_3 = CH_3
diosmetin 7-*O*-sulfate	R_1 = SO_3H	R_2 = H	R_3 = CH_3
chrysoeriol 7-*O*-sulfate	R_1 = SO_3H	R_2 = CH_3	R_3 = H
diosmetin 7,3′-*O*-disulfate	R_1 = SO_3H	R_2 = SO_3H	R_3 = CH_3



zosteric acid	R_1 =H	R_2 = OSO_3H
caffeic acid	R_1 =OH	R_2 = OH



rosmarinic acid (**9**)



7″,8″-didehydrosalvianolic acid B. (**7**)

Figure. 9.1. Structures of the flavonoids found in *Zostera* species.

9.3 Seasonal variation of phenolic compounds

Seasonal variability of phenolic compounds was investigated from August 2019 to July 2020. In addition, the percentage of each compound out of its total class of compounds (flavonoids and phenylpropanoic acids) is displayed in Table 9.1, Table 9.2 and Figure 9.2. It was possible to observe that the content of phenolic acids in the extract, with a variation from 0.07 to 16.8 mg/g dry weight (DW), showed an increase starting in September reaching a peak in November and decreasing until the end of the winter season, rising progressively until April reaching the maximum concentration peak in June. The results are in perfect conformity with the data reported for extracts of *Zostera marina* from the North Sea, showing a peak of phenolic acid concentration in the months of May/June and in sharp contrast to the results reported for *Zostera noltei,* which showed a peak of concentration in the coldest months of the year.[4] However, it must be kept in mind that the seagrass sampled in January-March were not as freshly washed ashore as in the other months, so sampling of *Zostera marina* in these months cannot be considered representative. Analysis of phenolic acid content shows a clear predominance of rosmarinic acid **9** (12.9 - 95.7%, but > 90% in almost half of the year) in all months, even reaching a concentration of 15.7 mg/g in June, followed by zosteric acid (0.04 - 0.15 mg/g), with a proportion of 0.9 - 50.8%, caffeic acid (0.03 - 0.39 mg/g) (1.5 - 36.3%), and 7″,8″-didehydrosalvianolic acid B (0.02 - 0.51 mg/g) (1.7 - 11.7%). (Table 9.1). As for the concentration of phenylpropanoid acids, it was about 2-5 times lower and ranged from 0.05 to 3.57 mg/g (dry leaves). In the Baltic Sea samples, eight flavonoids, five of which had one or two sulfate groups,

accounted for 82.2% to 100% of the total flavonoids, including two disulfated flavones (luteolin 7,3′-O-disulfate and diosmetin 7,3′-O-disulfate) and three mono-sulfated flavones (luteolin 7-O-sulfate, diosmetin 7-O-sulfate, and chrysoeriol 7-O-sulfate). Certainly, a clear predominance in all months of the year was represented by luteolin 7,3′-O-disulfate and diosmetin 7-O-sulfate, which affected the total flavonoid content by 24.7 to 58.7% and 17.4 to 41.3%, respectively, with a concentration ranging from 0.03 mg/g to 1.30 mg/g and from 0.02 to 1.24 mg/g dry weight, respectively. It was possible to verify two significant increases in November or June in the concentration of sulfated flavones, in contrast to the flavonoid aglycones apigenin and luteolin, which showed very low concentrations throughout the year. In conclusion, compared to the data reported for Norwegian North Sea *Zostera marina* extracts, the quantity of phenolic compounds was much lower, with a maximum peak in June.[4] Interestingly, the concentrations of the dominant compounds in the extract, luteolin 7, 3′-O-disulfate, and diosmetin 7-O-sulfate, showed an inverse trend over the months, reaching the same value in June. From the analysis of flavonoid content, it was possible to define the presence of different chemotypes in *Z. marina*, the chemotype 1, Baltic Sea, or luteolin 7,3′-O-disulfate/diosmetin 7-O-sulfate type and a chemotype 2, North Sea, diosmetin 7-O-sulfate/luteolin 7,3′-O-disulfate type.

Table 9.1 Phenolic amounts in plants collected in different months. Data are expressed as mg/g (dry weight, DW) of plant material (mean values ± s.d.).

Sample	Aug. 2019	Sept. 2019	Oct. 2019	Nov. 2019	Dec. 2019	Jan. 2020	Feb. 2020	Mar. 2020	Apr. 2020	May 2020	Jun. 2020	Jul. 2020
Flavonoids												
apigenin	nq	nq	0.03 ± 0.00^{d,e,f}	0.03 ± 0.00^{c,d,e}	0.02 ± 0.00^{e,f}	0.02 ± 0.00^{g}	nq	0.04 ± 0.01^{c}	0.03 ± 0.00^{cd}	0.04 ± 0.01^{b}	0.08 ± 0.02^{a}	var
luteolin	var[5]	var	0.03 ± 0.00^{e}	0.03 ±0.00^{d,e}	0.02 ± 0.00^{f}	0.02 ± 0.00^{f}	nq	0.04 ± 0.00^{d}	0.03 ±0.00^{de}	0.05 ± 0.01^{c}	0.14 ± 0.02^{a}	
luteolin 7-*O*-glucoside	nq[6]	nq	nq	nq	nq	nq	nq	nq	0.04 ± 0.00^{a}	0.02 ± 0.00^{b}	0.05 ± 0.01^{a}	0.06 ± 0.00^{b}
luteolin 7-*O*-sulfate	0.02 ± 0.00^{g}	nq	0.07 ± 0.01^{d}	0.12 ± 0.01^{c}	0.06 ± 0.00^{d}	0.03 ±0.00^{f,g}	nq	0.04 ± 0.00^{f}	0.20 ± 0.01^{b}	0.06 ± 0.01^{d}	0.32 ± 0.03^{a}	0.02 ± 0.00^{b}
luteolin 7,3′-*O*-disulfate	0.05 ± 0.00^{e,f}	0.03 ± 0.00^{g}	0.53 ± 0.06^{d}	1.25 ± 0.05^{b}	0.76 ± 0.02^{c}	0.13 ± 0.01^{f}	var	0.27 ± 0.02^{e}	0.79 ± 0.03^{c}	0.15 ± 0.02^{f}	1.30 ± 0.12^{a}	0.05 ± 0.00^{e}
diosmetin 7-O-sulfate	0.03 ± 0.00^{f}	0.02 ± 0.00^{f}	0.18 ± 0.01^{g}	0.44 ± 0.03^{c}	0.23 ± 0.00^{f}	0.07 ± 0.01^{h}	nq	0.33 ± 0.01^{e}	0.72 ± 0.03b	0.19 ± 0.00^{g}	1.24 ± 0.09^{a}	0.51 ± 0.03^{d}
chrysoeriol 7-O-sulfate	0.02 ± 0.00^{h}	nq	0.05 ± 0.01^{e}	0.08 ± 0.01^{b}	0.07 ± 0.00^{c}	0.03 ± 0.01^{g}	nq	0.04 ± 0.00^{f}	0.07 ± 0.01d	0.04 ± 0.00^{f}	0.14 ± 0.02^{a}	0.39 ± 0.03^{d}
diosmetin 7,3′-*O*-disulfate	0.02 ± 0.00^{h}	nq	0.09 ± 0.01^{f}	0.22 ± 0.01^{c}	0.16 ± 0.01^{d}	0.05 ± 0.01^{g}	nq	0.10 ± 0.01^{e}	0.27 ± 0.01^{b}	0.06 ± 0.01^{g}	0.31 ± 0.03^{a}	0.04 ± 0.00^{f}
Phenolic acids												0.10 ± 0.01^{e}
zosteric acid	0.10 ± 0.00^{e}	0.05 ± 0.00^{h}	0.14 ± 0.01^{b}	0.13 ± 0.02^{c}	0.10 ± 0.01^{e}	0.06 ± 0.02^{g}	0.04 ± 0.01[1]	0.04 ± 0.00^{i}	0.10 ± 0.01^{e}	0.12 ± 0.01^{d}	0.15 ± 0.01^{a}	0.08 ± 0.01^{f}
caffeic acid	0.14 ± 0.03^{b}	0.04 ± 0.01^{f,g}	0.10 ± 0.01^{d}	0.10 ± 0.01^{d}	0.07 ± 0.00^{e}	0.04 ± 0.00^{f}	0.03 ± 0.00^{g}	0.05 ± 0.03^{f}	0.12 ± 0.01^{c}	0.11 ± 0.02^{c,d}	0.39 ± 0.04^{a}	0.07 ± 0.01^{e}
rosmarinic acid	0.25 ± 0.07^{g}	nq	1.92 ± 0.41^{d}	3.87 ± 0.28^{c}	1.63 ± 0.16^{d}	0.14 ± 0.02^{g}	nq	0.78 ±0.16^{e,f}	7.80 ± 0.49^{b}	0.85 ± 0.16^{e}	15.7 ± 1.5^{a}	0.32 ±0.01^{f,g}
7	0.02 ± 0.00^{e}	nq	0.07 ± 0.02^{d}	0.14 ± 0.02^{b}	0.06 ± 0.01^{d}	0.03 ± 0.01^{e}	nq	nq	0.13 ± 0.02b	0.10 ± 0.01^{c}	0.51 ± 0.05^{a}	var
TDS[1]	0.07 ± 0.00^{h}	0.03 ± 0.00^{h}	0.62 ± 0.06^{e}	1.47 ± 0.05^{b}	0.92 ± 0.02^{d}	0.18 ± 0.01^{g}		0.37 ± 0.01^{f}	1.06 ± 0.03^{c}	0.21 ± 0.02^{g}	1.61 ± 0.12^{a}	0.61 ± 0.03^{e}
TMS[2]	0.07 ± 0.00^{f,g}	0.02 ± 0.00^{g}	0.30 ± 0.01^{e}	0.64 ± 0.03^{c}	0.37 ± 0.01^{e}	0.12 ± 0.01^{f}		0.40 ± 0.01^{e}	0.99 ± 0.03^{b}	0.30 ± 0.03^{e}	1.70 ± 0.09^{a}	0.48 ± 0.03^{d}
TF[3]	0.14 ± 0.02^{g}	0.05 ± 0.01^{h}	0.98 ± 0.11^{d}	2.16 ± 0.08^{b}	1.34 ± 0.04^{c}	0.34 ± 0.01^{f}		0.86 ±0.04^{d,e}	2.14 ± 0.10 ^{b}	0.61 ± 0.04^{e}	3.57 ± 0.35^{a}	1.17 ± 0.04^{c}
TPA[4]	0.51 ± 0.07 ^{f,g}	0.09 ± 0.03^{g}	2.23 ± 0.41^{d}	4.25 ± 0.28^{c}	1.86 ± 0.16^{d,e}	0.27±0.02 ^{f,g}	0.07 ± 0.00^{g}	0.87 ±0.16^{f,g}	8.15 ± 0.49^{b}	1.18 ± 0.17^{f}	16.8 ± 1.5^{a}	0.48±0.06 ^{f,g}

Total disulfate flavonoids; 2 total mono-sulfate flavonoids; 3 total flavonoids; 4 total phenolic acids; 5 variable: only detected in two samples, but absent from the remaining samples; 6 nq: could not be quantified due to lack of baseline and very low concentrations; 7 not real 0.00, means standard deviation < 0.005.

Table 9.2 Relative content (%) of each compound in respect to its compound class (phenolic acids or flavonoids).

Sample	Aug. 2019	Sept. 2019	Oct. 2019	Nov. 2019	Dec. 2019	Jan. 2020	Feb. 2020	Mar. 2020	Apr. 2020	May 2020	Jun. 2020	Jul. 2020
Flavonoids %												
apigenin	0.0	0.0	2.7	1.3	1.6	5.7	0.0	4.2	1.4	7.2	2.3	4.2
luteolin	0.0	0.0	2.9	1.6	1.5	5.6	0.0	4.6	1.5	7.5	4.0	5.1
luteolin 7-*O*-glucoside	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.9	3.2	1.3	1.9
luteolin 7-*O*-sulfate	15.1	0.0	7.0	5.6	4.8	8.0	0.0	4.4	9.4	10.1	9.1	4.1
luteolin 7,3′-*O*-disulfate	37.5	58.7	54.4	57.6	56.8	38.3	0.0	32.0	36.7	24.7	36.4	41.3
diosmetin 7-*O*-sulfate	19.9	41.3	18.5	20.2	17.4	19.5	0.0	38.7	33.6	31.6	34.7	31.9
chrysoeriol 7-*O*-sulfate	12.2	0.0	5.6	3.8	5.6	7.8	0.0	4.2	3.1	6.3	3.8	2.9
diosmetin 7,3′-*O*-disulfate	15.2	0.0	8.9	10.0	12.3	15.2	0.0	12.0	12.4	9.5	8.6	8.5
Phenolic acids %												
zosteric acid	19.6	50.8	6.2	3.0	5.4	21.5	61.7	4.6	1.2	9.9	0.9	15.5
caffeic acid	27.8	36.3	4.5	2.3	3.8	15.9	38.3	5.3	1.5	9.1	2.4	13.5
rosmarinic acid (**9**)	48.8	12.9	86.3	91.3	87.4	52.4	0.0	90.1	95.7	72.5	93.8	59.3
7″,8″-didehydrosalvianolic acid B (**7**)	3.9	0.0	3.0	3.4	3.4	10.2	0.0	0.0	1.7	8.4	3.0	11.7
TDS[1]	52.7	58.7	63.3	67.6	69.1	53.4	0.0	44.0	49.2	34.2	44.9	49.8
TMS[2]	47.3	41.3	31.1	29.6	27.8	35.3	0.0	47.3	46.1	48.0	47.5	38.9
TSF[3]	100	100	94.4	97.1	96.9	88.7	0.0	91.3	95.3	82.2	92.5	88.7

[1] Total disulfated flavonoids; [2] Total mono-sulfated flavonoids; [3] Total sulfated flavonoids.

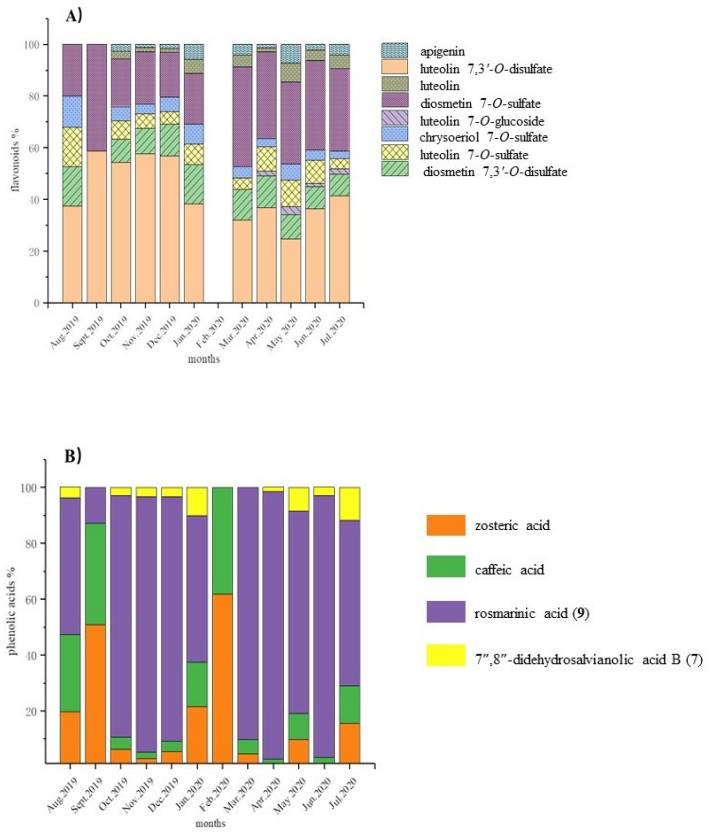


Figure 9.2. Relative content (%) of flavonoids A) and phenolic acids B) detected for each sampling period.

9.4 Isolation of 7″,8″-didehydrosalvianolic acid B (7)

In addition to known phenolic compounds, the study of UHPLC-DAD-MS analysis of extracts of the eelgrass *Zostera marina* from the Baltic Sea, enabled the identification of a new secondary metabolite, 7″,8″-didehydrosalvianolic acid B (**7**). To isolate 7″,8″-didehydrosalvianolic acid B, (**7**) the water-methanol extract (1:1) was first subjected to a preparative C_{18}-MPLC column with H_2O-MeOH gradient and the fraction containing the compound of interest, was purified with a LH-20 Sephadex methanolic column, obtaining a pure compound with the appearance of an amorphous brownish powder.

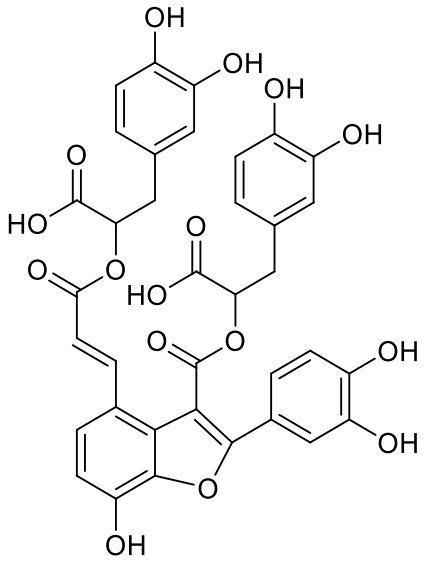


Figure 9.3. Structure of 7″,8″-didehydrosalvianolic acid B (**7**).

9.5 Elucidation of the planar structure of 7″,8″-didehydrosalvianolic acid B (7)

The structural determination of 7″,8″-didehydrosalvianolic acid B (**7**) was carried out using mass spectrometry (MS) and nuclear magnetic resonance (NMR) both one- and two-dimensional homonuclear and heteronuclear (COSY, NOESY, TOCSY, HMBC, HSQC) techniques.

The HR-ESI-MS spectrum of 7″,8″-didehydrosalvianolic acid B (**7**) revealed the presence of the molecular ion peak $[M+NH_4]^+$ at *m/z* 734.1710, a value very close to the exact mass 734.1716 calculated for the ion of formula $C_{36}H_{32}NO_{16}^+$. This allowed the molecular formula of compound **7** to be defined as $C_{36}H_{28}O_{16}$.

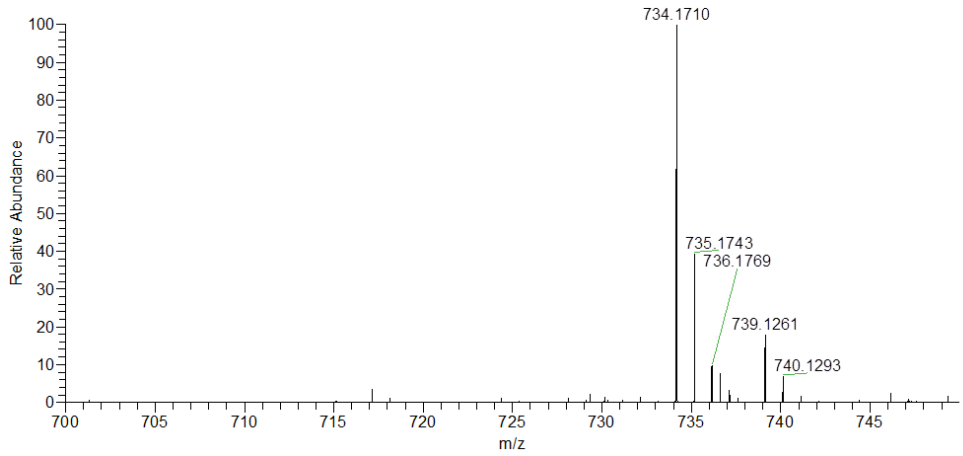


Figure 9.4 . HS-ESI-MS spectrum of 7″,8″-didehydrosalvianolic acid B (**7**)

The ^{1}H-NMR spectrum showed the presence of 13 olefinic protons resonating at a chemical shift between δ_H 6.33 and 8.30, the presence of 2 oxygenated methynes at δ_H 5.33 and 5.20 and, finally, 4 methylene protons resonating at δ_H 3.00 and 3.09.

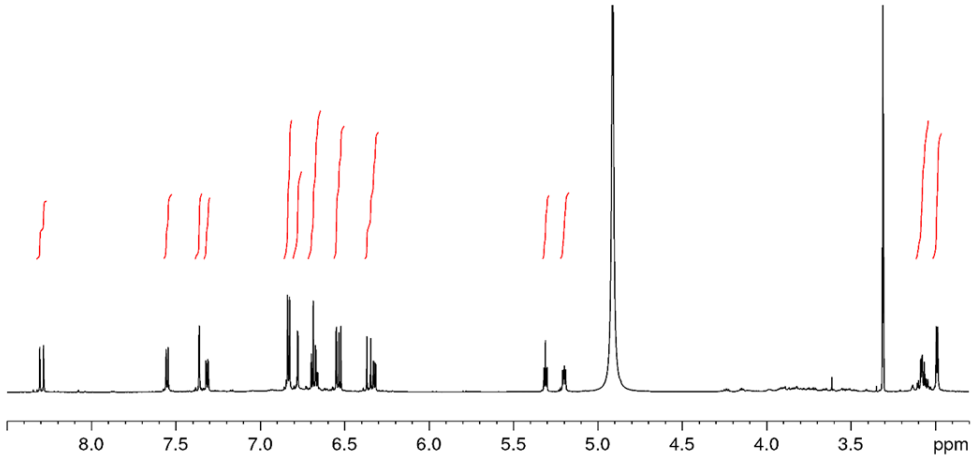


Figure 9.5. ^{1}H-NMR spectrum of 7″,8″-didehydrosalvianolic acid B (**7**)

The ^{13}C NMR spectrum showed signals relating to four carboxylic carbons with chemical shifts between δ_C 167.2 and 174.5; signals relating to three 1,3,4-trisubstituted benzene rings between δ_C 116.4 and 149.1; four olefinic carbons between δ_C 109.1 and 160.2; two oxygenated methynes at δ_C 76.5 and 75.1 and finally two methylenes at δ_C 38.0 and 37.5. It was possible to perform a complete interpretation of ^{13}C NMR spectra with the analysis of data from HSQC and HMBC experiments. The complete set of ^{1}H and ^{13}C chemical shift is reported in Table 9.3. It was possible to verify that that signals were very similar to data of salvianolic acid B and its derivatives.[5,6]

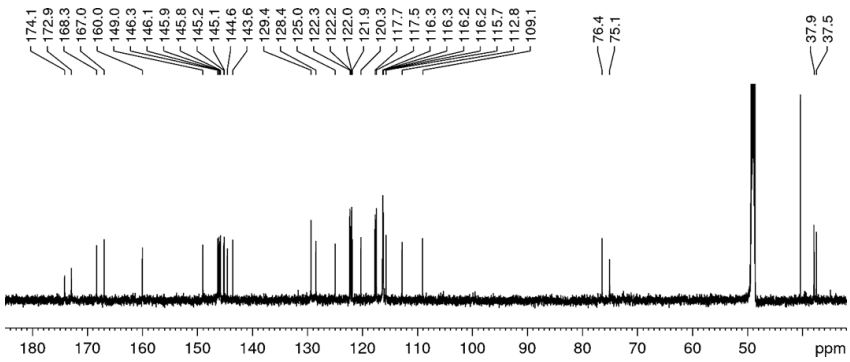
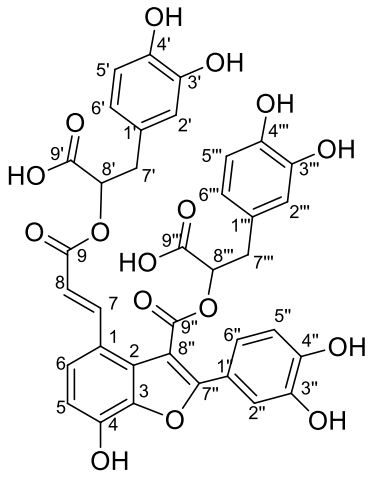


Figure 9.6. ^{13}C-NMR spectrum of 7″,8″-didehydrosalvianolic acid B (**7**)

Table 9.3. ^{1}H and ^{13}C NMR (^{1}H 700 MHz, ^{13}C 175 MHz, CD_3OD) of 7″,8″-didehydrosalvianolic acid B (**7**)



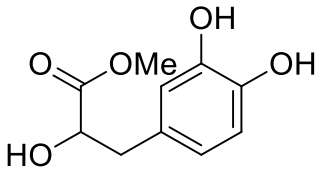
No.	δ_H	δ_C	No.	δ_H	δ_C
1		120.3	1″		121.9
2		129.4	2″	7.36 1H, *d* (2.1)	116.2
3		143.6	3″		146.3
4		145.9	4″		149.0
5	6.83 1H, *m*	112.8	5″	6.83 1H, *m*	116.3
6	7.55 1H, *d* (8.5)	125.0	6″	7.31 1H, *dd* (8.1, 2.1)	122.2
7	8.30 1H, *d* (15.7)	144.6	7″		160.0
8	6.36 1H, *d* (15.7)	115.7	8″		109.1
9		168.3	9″		167.0
1′		129.4	1‴		128.4
2′	6.78 1H, *d* (1.9)	117.5	2‴	6.55 1H, *d* (2.0)	117.7
3′		146.1	3‴		145.8
4′		145.2	4‴		145.1
5′	6.69 1H, *d* (8.0)	116.3	5‴	6.53 1H, *d* (8.1)	116.2
6′	6.67 1H, *dd* (8.0, 1.9)	122.0	6‴	6.32 1H, *dd* (8.1, 2.0)	122.3
7′	3.08 2H, *m*	37.9	7‴	2.99 2H, *d* (6.1)	37.5
8′	5.20 1H, *dd* (8.1, 4.7)	75.1	8‴	5.31 1H, *t* (6.1)	76.4
9′		174.1	9‴		172.9

From the interpretation of the COSY spectrum, it was possible to identify three ABX systems of aromatic protons, i.e., spin systems formed by three protons, two of which had chemical shifts close to each other: H2′/H5′/H6′, H2″/H5″/H6″, and H2‴/H5‴/H6‴. This demonstrated the presence of three 1,3,4-trisubstituted aromatic rings. Two AB systems of aromatic or olefinic

protons were also present: H5/H6, indicative of a 1,2,3,4-tetrasubstituted aromatic ring, and H7/H-8, which, due to the high coupling constant (15.7 Hz) was attributable to disubstituted alkene with trans configuration. Finally, signals from two -CH(O-)-CH_2- systems were present: H8'-H27' and H8"-H27". The HMBC spectrum allowed the fragments identified by the COSY spectrum to be linked together. In particular, the HMBC spectrum showed a diagnostic correlation from H-2", H-6" to C-7", indicating the presence of an aromatic ring linked to an oxygenated olefinic carbon at position C-7". Finally, thanks to the HMBC spectrum, it was also possible to assign the chemical shift of the carbon at position 8", which resonates at a low chemical shift (δ 109.1) given its position in α to a carbonyl and its proximity to the carbon at position 7".

9.6 Absolute configuration of 7",8"-didehydrosalvianolic acid B (7)

The absolute configuration of 7",8"- didehydrosalviolanic acid B (**7**) was elucidated by comparing the ECD spectra of methyl (3,4-dihydroxyphenyl) lactate (**8**) produced by the degradation of compound **7**, and methyl (3,4-dihydroxyphenyl) lactate produced by the degradation of rosmarinic acid (**9**), the absolute configuration of which is known from previous studies.

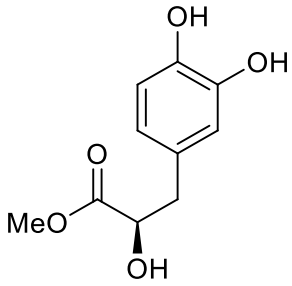


methyl (3,4-dihydroxyphenyl)lactate

Figure 9.7. Structure of methyl (3,4-dihydroxyphenyl)lactate (**8**)

Rosmarinic acid (**9**) isolated from *Zostera marina* is dextrorotatory, showing the same specific positive rotation of the rosmarinic acid found in rosemary (*Rosmarinus officinalis*), whose absolute configuration has been determined. Consequently, it has the same absolute configuration *R* in the only stereogenic center of the molecule.[7] Rosamarinic acid (**9**) (2 mg) was subjected to an acid methanolysis reaction by treating the sample with 1M HCl in 92% MeOH for 72 hours at room temperature and its reaction product was separated by HPLC. Methyl (*R*)-(3,4-dihydroxyphenyl)lactate (**10**) was identified by proton NMR spectrum analysis. In addition, UV and ECD spectra were also recorded to be used as reference standards for comparison with the ECD spectra of the degradation product of compound **7**.

The same procedure was also carried out on 7',8''-dehydroxalviolanic acid B (**7**). However, rather than the complete methanolysis product of sample **7**, the proton NMR spectrum of the reaction mixture showed only the cleavage of the ester bond between C-9 and C-8'. Consequently, the methanolysis product was purified using HPLC on an RP-18 column, which resulted in the isolation of the partially methanolized derivative (**11**) and methyl (3,4-dihydroxyphenyl)lactate (**8**). The ^{1}H-NMR, UV and ECD spectra of the latter compound were identical to those of methyl (*R*)-(3,4-dihydroxyphenyl)lactate (**10**) from rosmarinic acid, and this allowed the configuration of the chiral carbon C-8' to be assigned as *R*.



methyl (*R*)-(3,4-dihydroxyphenyl) lactate

Figure 9.8. Structure of methyl (*R*) -(3,4-dihydroxyphenyl)lactate (**10**)

Once isolated in pure form, compound **11,** obtained from partial methanolysis, was fully characterized by mass spectrometry and one- and two-dimensional NMR experiments. (Table 9.4).

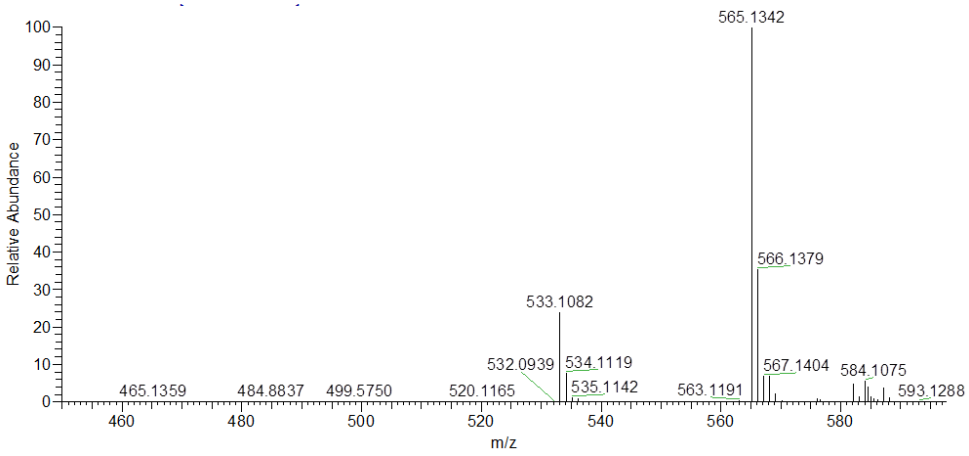


Figure 9.9 HR-ESI-MS spectrum of partially methanolized derivative (**11**)

Particularly diagnostic for understanding the position of the ester bond broken by the methanolysis reaction was the HMBC spectrum, which showed correlations between the methoxy group and the C-9 of the ester carbonyl and the H-7 trans proton in the double bond. (Figure 9.10).

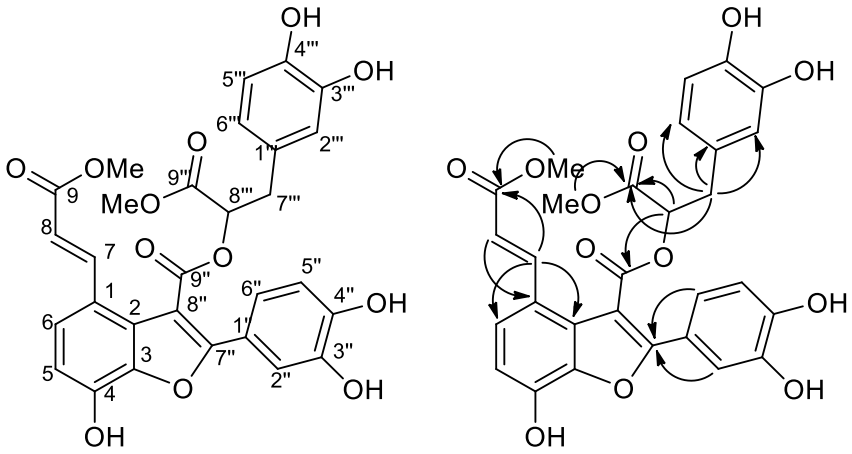
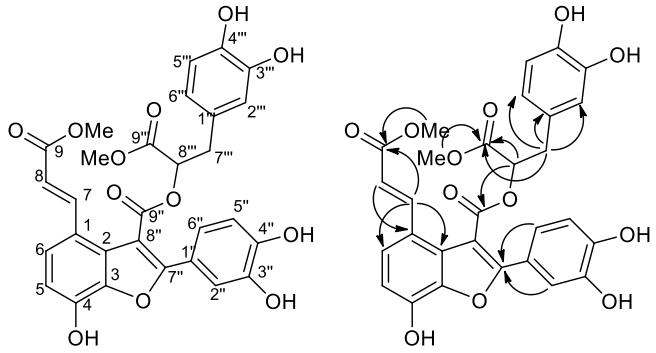


Figure 9.10. Chemical structure and HMBC key correlations of the methanolysis derivative of 7″,8″-didehydrosalvianolic acid B.(**11**)

Table 9.4 ^{1}H and ^{13}C NMR (^{1}H 700 MHz, ^{13}C 175 MHz, CD_3OD) of partially methanolized derivative (**11**)



No.	δ_H	δ_C	No.	δ_H	δ_C
1		118.6	6″	7.30 1H, *dd* (8.3, 2.1)	122.0
2		129.1	7″		160.0
3		144.7	8″		108.7
4		144.1	9″		166.7
5	6.77 1H, *d* (8.4)	113.9	1‴		128.1
6	7.50 1H, *d* (8.4)	125.1	2‴	6.55 1H, *d* (2.0)	117.6
7	8.15 1H, *d* (15.6)	144.0	3‴		146.2
8	6.26 1H, *d* (15.6)	114.9	4‴		145.5
9		169.8	5‴	6.55 1H, *d* (8.1)	116.2
9-OMe	3.76 1H, *s*	52.0	6‴	6.35 1H, *dd* (8.1, 2.0)	122.1
1″		122.2	7‴	3.00 2H, *m*	37.6
2″	7.36 1H, *d* (2.1)	116.3	8‴	5.33 1H, *t* (6.3)	76.4
3″		146.5	9‴		171.7
4″		148.9	9‴-OMe	3.70 1H, *s*	52.7
5″	6.85 1H, *d* (8.3)	116.3			

To elucidate the configuration of the second stereogenic center, the partially methanolised derivative (**11**) was subjected to a further methanolysis reaction using more drastic conditions, in the attempt to methanolise the other dihydroxyphenyllactate unit as well. The sample, again treated with 1M HCl in 92% MeOH, was left to react for 120 h at a temperature of 70 °C. At the end of 120 h, a new sample of methyl (3,4-dihydroxyphenyl)lactate (**8**) was obtained from the reaction mixture by

HPLC on an analytical RP18 column, and was subjected to UV and ECD spectroscopy. The ECD spectra of the two samples of methyl (3,4-dihydroxyphenyl)lactate (**8**), obtained respectively from the methanolysis of compound **7** and from the methanolysis of compound **11**, were identical. This showed that both units of methyl-(3,4-dihydroxyphenyl)lactate (**8**) derived from 7″,8″-didehydrosalvianolic acid B (**7**) have absolute *R* configuration (Figure 9.10).

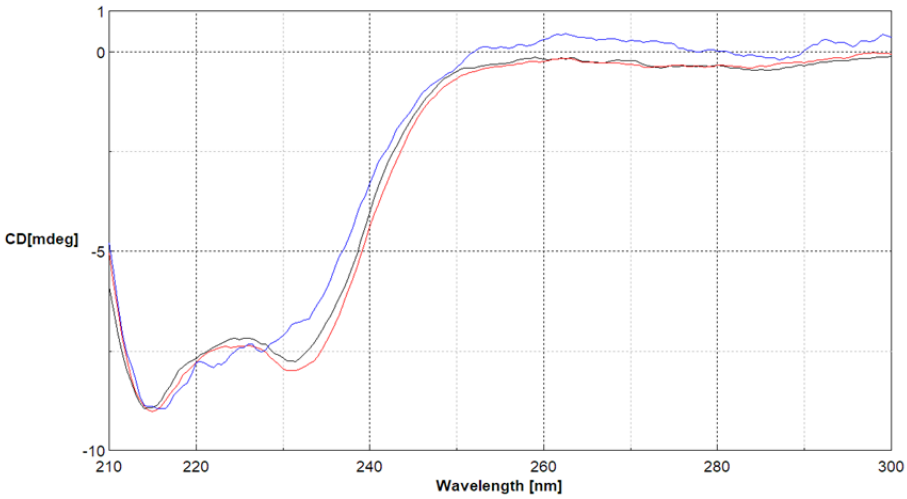


Figure 9.11. ECD spectrum of methyl (*R*)-(3,4-dihydyroxyphenyl) lactate (**10**) obtained from the methanolides of rosmarinic acid **9** (in black), partial methanolysis of compound **7** (in red), and methanolysis of compound compound **11** (in blue).

9.7 Pharmacological essays of 7″,8″-didehydrosalvianolic acid B (7)

The cytotoxic activity of 7″,8″-didehydrosalvianolic acid B (**7**) was evaluated on three human cancer cell lines and against normal human fibroblasts (breast adenocarcinoma, cervical carcinoma, malignant melanoma, and skin fibroblasts). However, cells resulted to be unaffected even after 72h of treatment. Moreover, the anti-inflammatory activity, angiogenesis, or ATPase inhibitory activity of 7″,8″-didehydrosalvianolic

acid B (**7**) were also evaluated. It was evident that compound **7**, did not influence angiogenesis or inflammation. However, compound **7** showed a selective inhibition of the isolated enzyme Na^+ / K^+ -ATPase. This bioactivity, evaluated measuring activity of ATPase after incubation with **7**, makes 7″,8″-didehydrosalvianolic acid B (**7**) an interesting lead compound for the development of new drugs for prevention of cardiac hypertrophy and heart failure.

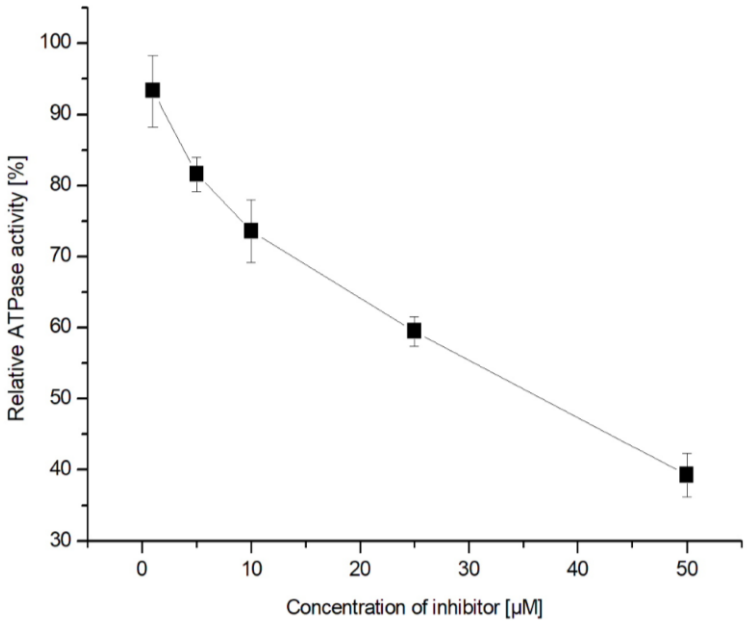


Figure. 9.12. Relative ATPase activity treated with **7** in micromolar range.

9.8 Experimental section

9.8.1 General methods

Distilled water (lab-made), methanol and acetonitrile (both gradient grade, VWR, Darmstadt, Germany) were used for extraction, MPLC, and HPLC analyses, respectively. LC-MS grade water and acetonitrile (VWR, Darmstadt, Germany) were used for UHPLC-DAD-MS analysis. Phenolic

standards were purchased from Sigma-Aldrich (Steinheim, Germany). Optical rotation was measured with a Jasco P-2000 polarimeter at sodium D-line. High-performance liquid chromatography (HPLC) separations carried out by the research group of Prof. Dr. Christian Zidorn were performed on a Thermo Scientific Dionex Ultimate 3000 apparatus, with a Phenomenex Luna Omega C18 column, 1.6 µm, 100×2. 1 mm, using the following elution scheme eluent A: 0.5 % trifluoroacetic acid in water; eluent B: 100 % acetonitrile; gradient: 0 min 10 % B, 25 min 20 % B, 40 min 50 % B, 40.1 min 10 % B, 55 min 10 % B, with a flow rate of 0.200 ml/min; injection volume 5 µl; column compartment temperature 30 °C. UHPLC-DAD-MS experiments were performed using a Shimadzu Nexera X2 system with LC-30AD binary pump connected to the SIL-30AC autosampler, CTO-20AC column heating system, SPD-M30A diode array detector and a Shimadzu LCMS 8030 triple quadrupole analyser. MS spectra were recorded in electron spray ionization (ESI) mode, scanning in a range of m/z = 200-1500 (Q1) and 100-1000 (Q3) in positive and negative mode. The interface voltage was 4.5 kV, the desolvation line temperature 250 °C, the heat block temperature 400 °C, the nebulizer gas flow 3 l/min, the drying gas flow 15 l/min, and the collision induced dissociation 230 kPa. The separation of compounds was achieved using a Phenomenex Luna Omega C18, 1.6 µm, 100 × 2.1 mm column, with an elution flow rate of 0.20 ml/min using the binary gradient (v/v) 1 % formic acid in water and acetonitrile (B). The linear gradient used was: 0 min 5 % B, 20 min 20 % B, 40 min 30 % B, 50 min 50 % B, 50.1 min stop, post time 10 min. The injection volume was 2 µl. In our laboratory, the compounds were isolated

using an Agilent 1260 Infinity Quaternary LC system with a diode array detector (DAD). NMR spectra were acquired using a Bruker spectrometer at 700 MHz with CD_3OD (MeOH-d_4) as solvent; chemical shifts refer to the CD_3OD solvent signal: δ_H 3.31, δ_C 49.0. The HSQC spectra were optimised for a $^{1}J_{CH}$ of 142 Hz while the HMBC experiments were optimised for a $^{2,3}J_{CH}$ of 8 Hz. High resolution ESI-MS experiments were performed on the Thermo LTQ Orbitrap XL mass spectrometer combined with a Thermo U3000 HPLC system.

9.8.2 Collection, extraction, isolation

Zostera marina was collected monthly for a whole year, from 12 August 2019 to 12 July 2020 along the coast of Kiel, Schleswig-Holstein, Germany near the Olympiazentrum Schilksee, (coordinates: N 54°25'39.0", E 10°10'17.5"; alt.: 0 m). A voucher specimen has been deposited in the Herbarium of the Institut für Botanik, Kiel University (voucher codes: YL-20181222A-1; KIEL0005004). The collected plant material were unrooted plants, freshly washed ashore. After collection, the samples were quickly transferred into the lab, washed in fresh water to remove sand, and then dried at room temperature in the dark. The dried leaves of *Zostera marina*, weighing 1 kg, were cut, ground, and and extracted with 50% aqueous methanol for 3 times (each time 24 h). The extracts were filtered through glass wool, and the methanol (also partial water) was removed using a rotary evaporator under reduced pressure at 27 ◦C. The remaining water (600 mL) was partitioned with either ethyl acetate (300 mL) or *n*-butanol (300 mL) three times. Once the butanol phase was evaporated under

vacuum, 1.52 g of residue was obtained and subjected to preparative MPLC. A PrepChrom C-700 system with a C18 PrepChrom HPLC column (15 µm, 250×30 mm) and eluted with a water-methanol gradient elution (flow rate: 20 ml/min; 0 min - 10% MeOH, 10 min - 20 % MeOH, 60 min - 50% MeOH, 100 min - 65% MeOH, 120 min - 80 % MeOH). Nine subfractions, L1 to L9, were collected. Fraction L8, eluted with a percentage of 60% MeOH and 40% H_2O, at 90 min weighing 90.4 mg, was again purified with an LH-20 Sephadex column (MeOH, 1 m × 2 cm). With this last step, 70.5 mg of 7″,8″-didehydrosalvianolic acid B was obtained (**7**).

7″,8″-didehydrosalvianolic acid B (7)

7″,8″-didehydrosalvianolic acid B (**7**) appearing as a colourless oil, exhibits a $[\alpha]^{20}_{D}$ -4.1 (*c* 0.5, MeOH); ECD (MeOH) λ_{max} ($\Delta\varepsilon$) 355 (+2.5), 301 (+1.6), 279 (-1.6), 226 (-11. 5) nm; UV (MeOH) λ_{max} (ε) 343 (38000), 335 (28000), 233 (45000, shoulder) nm; IR (thin solid film) ν_{max} 3263, 2938, 1704, 1607, 1520, 1444, 1367, 1261, 1196, 1165, 1023, 816 cm^{-1}. The ^{1}H and ^{13}C NMR data are described in Table 9.3; HR-ESI-MS acquired in positive, *m/z* 739.1261 ($[M+NH_4]^+$, with molecular formula $C_{36}H_{28}O_{16}Na$: 739.1270), 734.1710 ($[M+NH_4]^+$, calcd for $C_{36}H_{32}NO_{16}$: 734.1716).

9.8.3 Quantitative analysis

The quantitative detection procedure was based on the method described by Enerstvedt et al. (2016) with a few modifications.[2] Three replicates of 1.0 g of grounded leaf-samples were extracted with 20 ml of 50% aqueous methanol for 60 min at room temperature and sonicated for 5 min at the

beginning of each extraction. Extraction was repeated twice, and the total crude extracts were stored in a round bottle flask. After transferring of the extracts in a 10 ml volumetric flask, HPLC analysis were performed. Data was recorded at 360 ± 10 nm and 280 ± 10 nm using a calibration curve of luteolin, rosmarinic acid (**9**) as analytical standard [analytical standard, 97% (HPLC)]. The results are presented as mg luteolin/RA equivalents ± one standard deviation (SD) per gram of dry weight (DW) plant material. Compound identification was performed using seagrasses collected in Bergen, Norway. Kiel and Norwegian plants were subjected to the same extraction procedure and analysed by a similar HPLC method. Moreover, qualitative analysis was performed using as standard reference compounds (apigenin, luteolin, and rosmarinic acid).

9.8.4 Procedures for determining absolute configuration

Rosmarinic acid methanolysis

Rosmarinic acid (**9**) (2 mg) isolated from *Zostera marina* was subjected to acidic methanolysis (1 M HCl in 92% MeOH) at room temperature for 72 h. The reaction product was fractionated by HPLC using a Phenomenex Luna C18 (2), 5 µm, 100 × 4.6 mm column, flow rate of 1.0 ml/ min, detector UV at 280 nm, and the following water/MeOH binary linear gradient: 0 min 20% MeOH, 1 min 20% MeOH, 31 min 100% MeOH, 36 min 100% MeOH. The fraction containing methyl (*R*)-(3,4-dihydroxyphenyl)lactate (**10**) was collected and the compound was identified by ^{1}H-NMR spectrum. Subsequently, the UV and ECD spectra were also recorded to use them as a reference.

Partial methanolysis of 7″,8″-dideohydroxyalviolanic acid B (7)

The same procedure, (acid methanolysis under the same conditions and HPLC following the same elution scheme), was also repeated on 7″,8″-didehydrosalvianolic acid B (1.0 mg) (**7**). Separation by HPLC allowed us to obtain methyl (*R*)-(3,4-dihydroxyphenyl)lactate (**10**) and the partially methanolized compound (**11**). The UV and ECD spectra of methyl (*R*)-(3,4-dihydroxyphenyl)lactate (**10**) were recorded and found to be comparable to those used as reference.

7″,8″-didehydrosalvianolic acid B partially methanolised (**11**)

Colourless oil, $[\alpha]^{20}_{D}$ +14 (*c* 0.01, MeOH); UV (MeOH) λ_{max} (ε) 341 (37000) nm; ^{1}H and ^{13}C NMR data are described in Table 9.4; HR-ESI-MS *m/z* 565.1342 ($[M+H]^{+}$, calcd for $C_{29}H_{25}O_{12}Na$: 565.1341).

*Methanolysis of partially methanolized 7″,8″-didehydrosalvianolic acid B (**10**)*

The partially methanolized compound **11** (100 μg) was subjected to a new acid methanolysis, this time under more drastic conditions such as 70 °C for 120 h. Through an HPLC separation, a fraction containing pure methyl (*R*)-(3,4-dihydroxyphenyl)lactate (**10**) was obtained. The UV and ECD spectra of this fraction were recorded and showed the same sign and comparable magnitude compared to the above reference spectra.

9.8.5 Bioassays

Cytotoxicity assay

The cytotoxicity of compounds was evaluated on three different cancer cell lines: malignant melanoma G-361, cervix epithelial carcinoma HeLa, breast carcinoma MCF7 and human foreskin fibroblasts (BJ), after 72 h of treatment stained by resazurin.[8]

Inhibition of Na^+/K^+-ATPase activity

Isolation of Na+/K+-ATPase from porcine kidney and its activity measurement method were described previously.[9,10] Microwell plates using microplate reader Tecan were used to calculate ATPase activity, based on the colorimetric detection. In this assay, ammonium molybdate interacts with inorganic phosphate (product of ATP hydrolysis), which leads to a colour change measured at 710 nm.[11,12.] DMSO was used to prepare a stock solution of 7″ ,8″ -didehydrosalvianolic acid B with a concentration of 10 mM. This resulted in a maximal final concentration of DMSO in the assay of 0.5%. Prior tests had shown that this concentration was not influencing the activity of ATPase.

All data are presented as the specific activity of Na+/K+-ATPase that was standardly estimated using treatment by 500 mM ouabain, which served as a highly specific inhibitor of Na+/K+-ATPase. The residual activity in the presence of ouabain was subtracted from the total ATPase activity in the absence of ouabain, yielding the specific activity of Na+/ K+-ATPase. The

values represent the means ± S.D. of four independent replicates in three repetitions.

9.9 Spectroscopic data

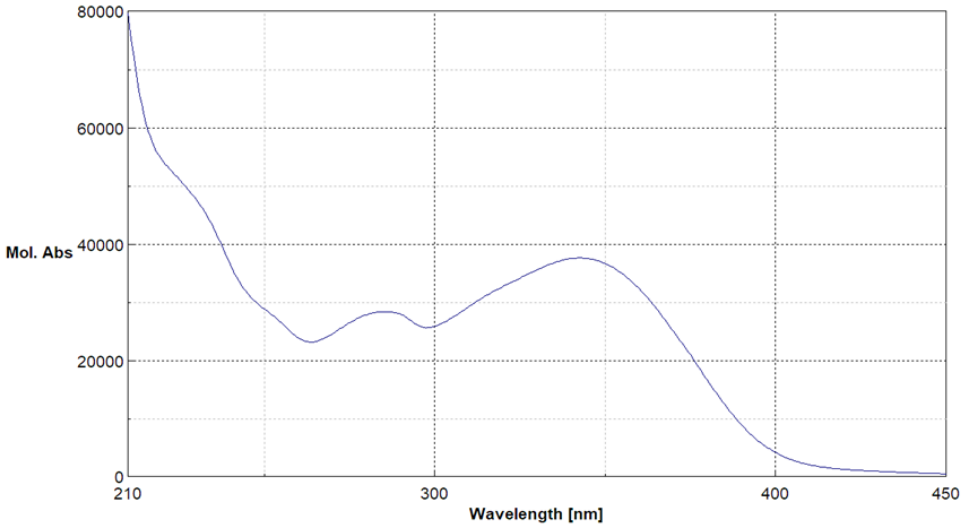


Figure. 9.13. UV spectrum of pure compounds **7** isolated from *Zostera marina*.

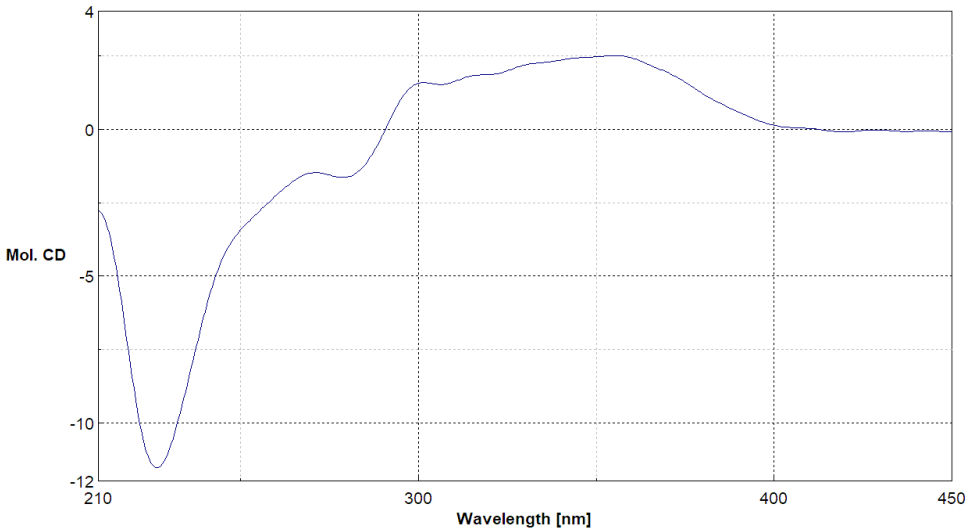


Figure. 9.14. ECD spectrum of 7",8"-didehydrosalvianolic acid B (**7**) from *Zostera marina.*

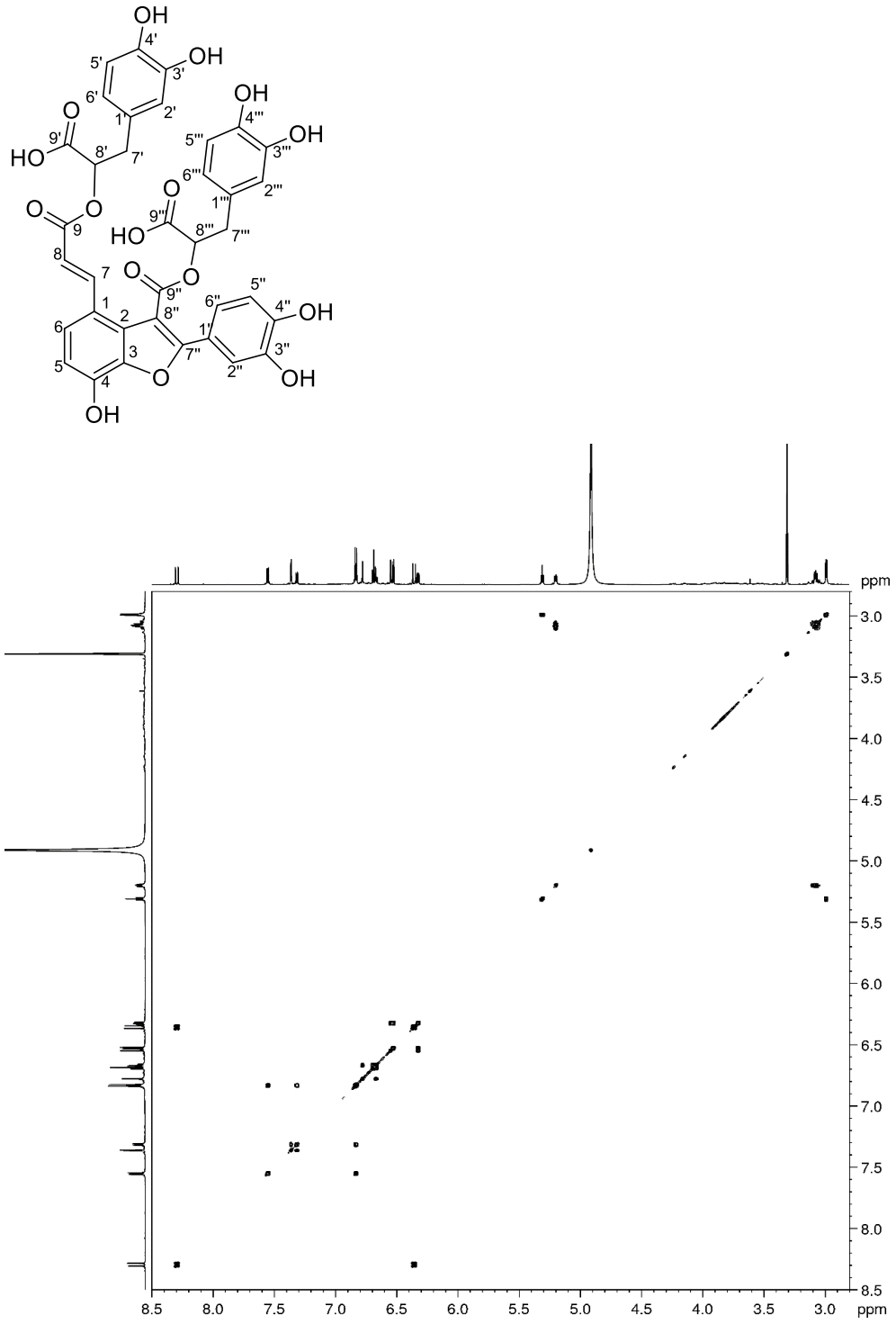


Figure 9.15 H, H-COSY spectrum of 7",8"-didehydrosalvianolic acid B (**7**) (CD_3OD, 700 MHz).

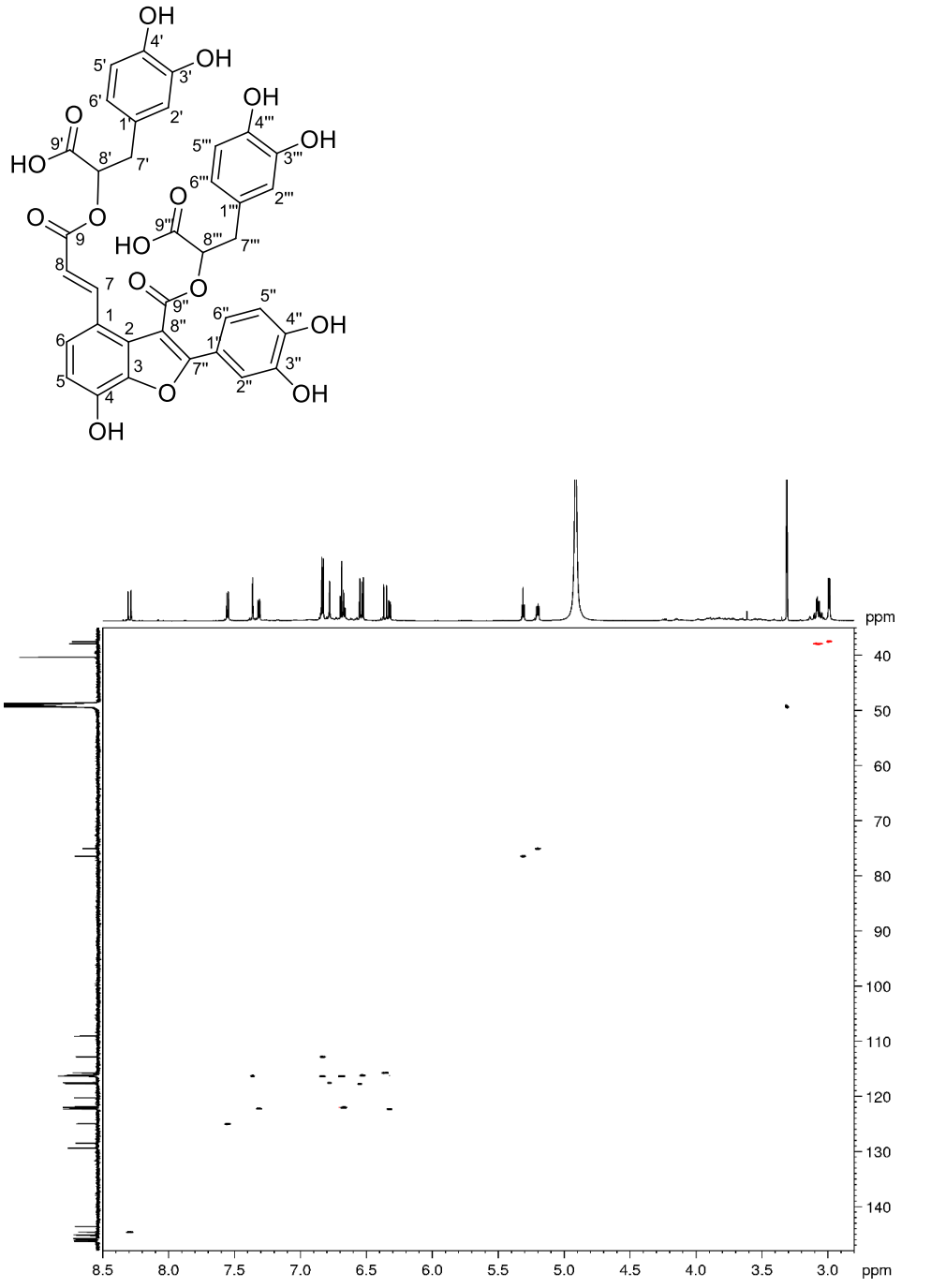


Figure 9.16. HSQC spectrum of 7",8"-didehydrosalvianolic acid B (**7**) (CD_3OD, 700 MHz).

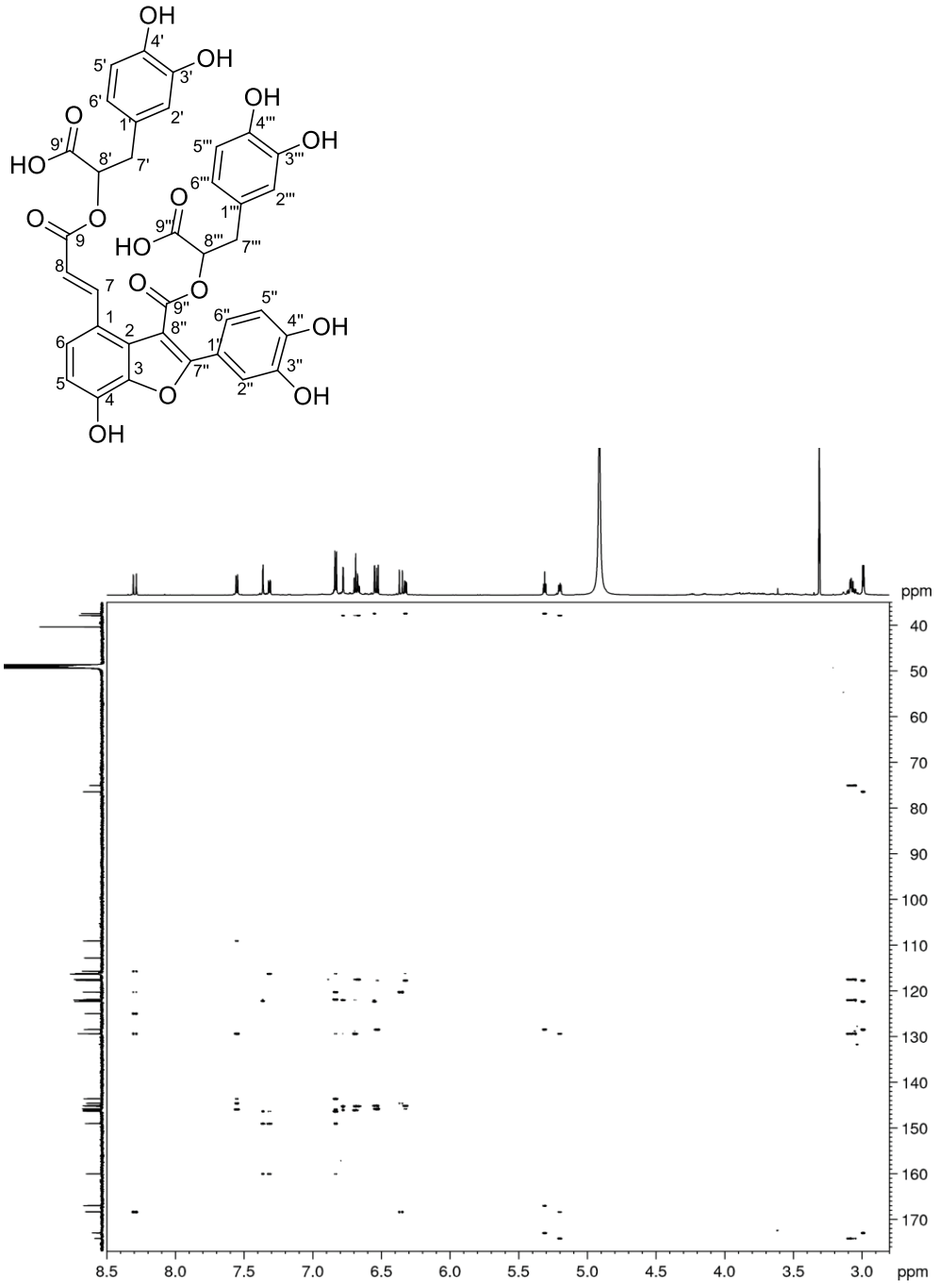


Figure 9.17 HMBC spectrum of 7",8"-didehydrosalvianolic acid B (**7**) (CD_3OD, 700 MHz).

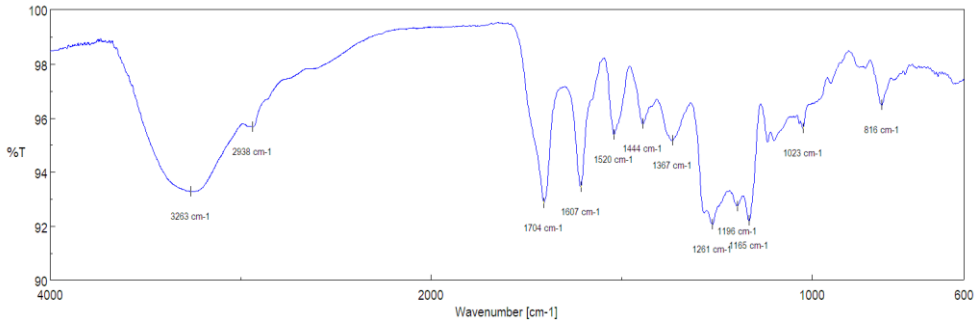


Figure 9.18. IR spectrum (thin solid film) of 7",8"-didehydrosalvianolic acid B (**7**).

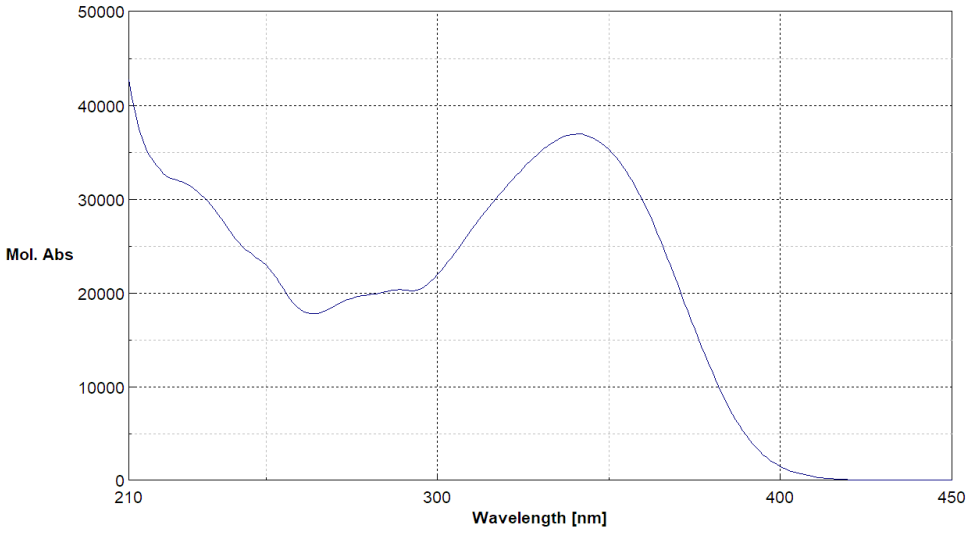


Figure 9.19 UV spectrum of the partially methanolized compound **11.**

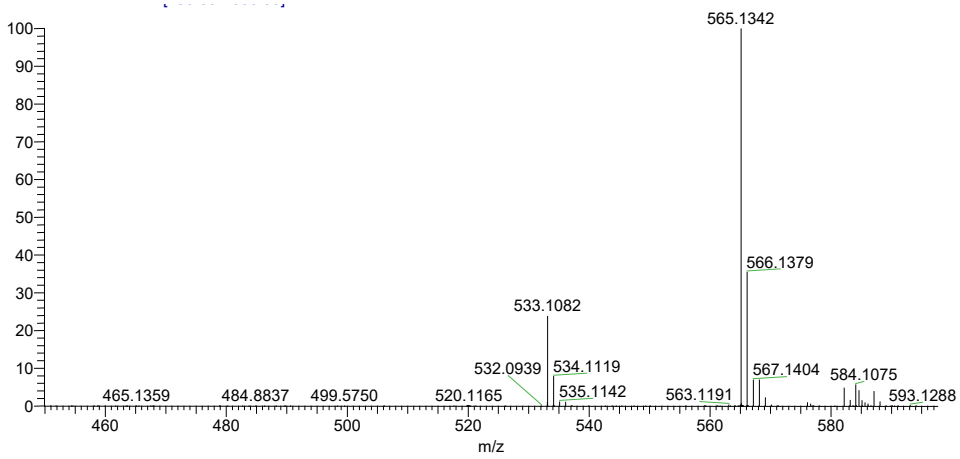


Figure 9.20. High-resolution ESI MS spectrum of the partially methanolized compound **11.**

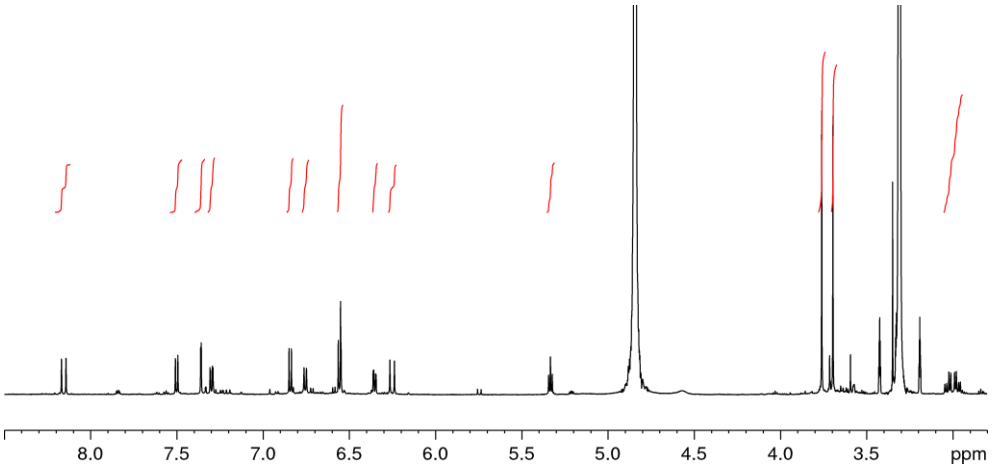


Figure 9.21 ^{1}H NMR spectrum of the partially methanolized compound **11** (CD_3OD, 600 MHz).

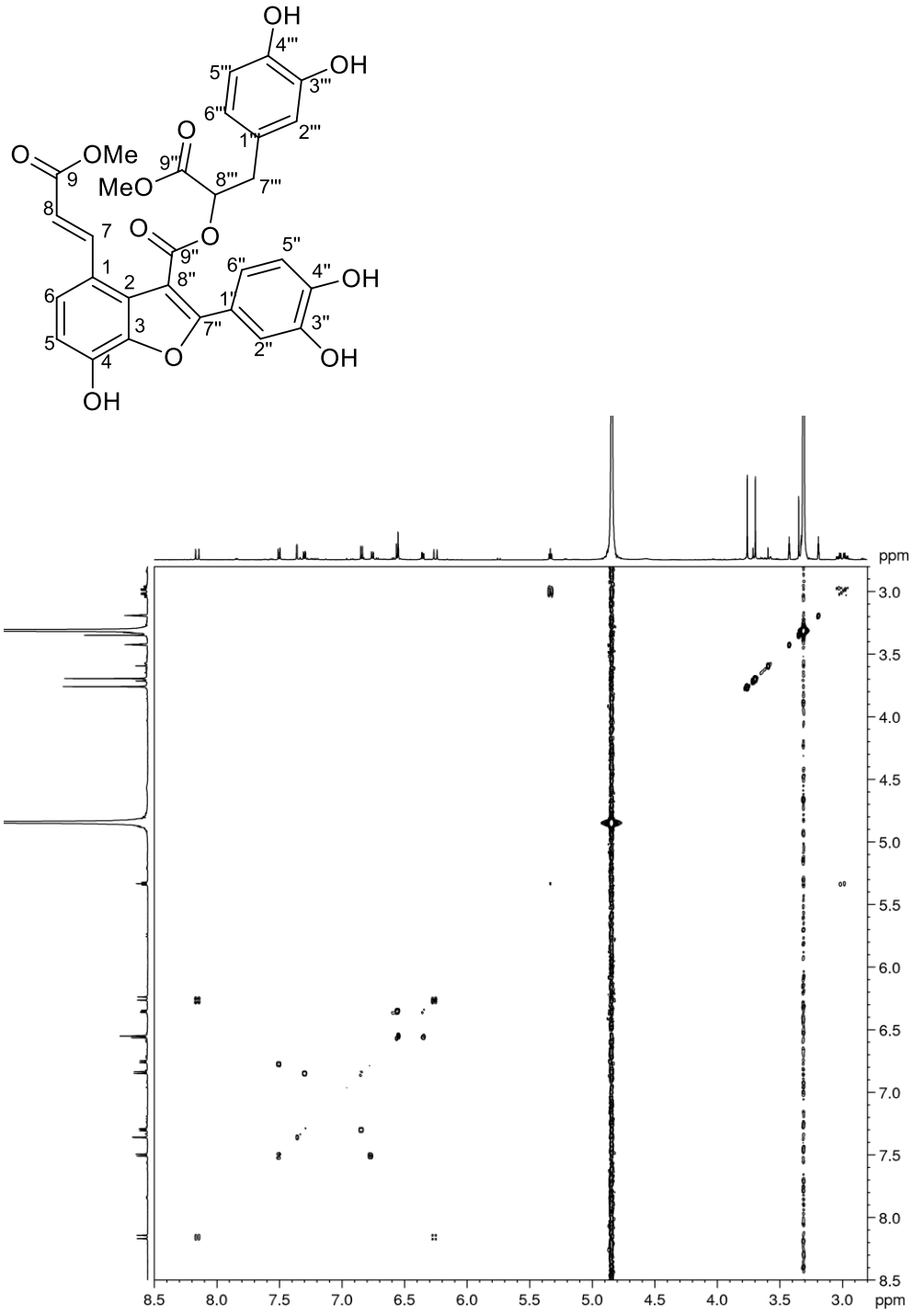


Figure 9.22 H,H-COSY spectrum of the partially methanolized compound **11** (CD_3OD, 600 MHz).

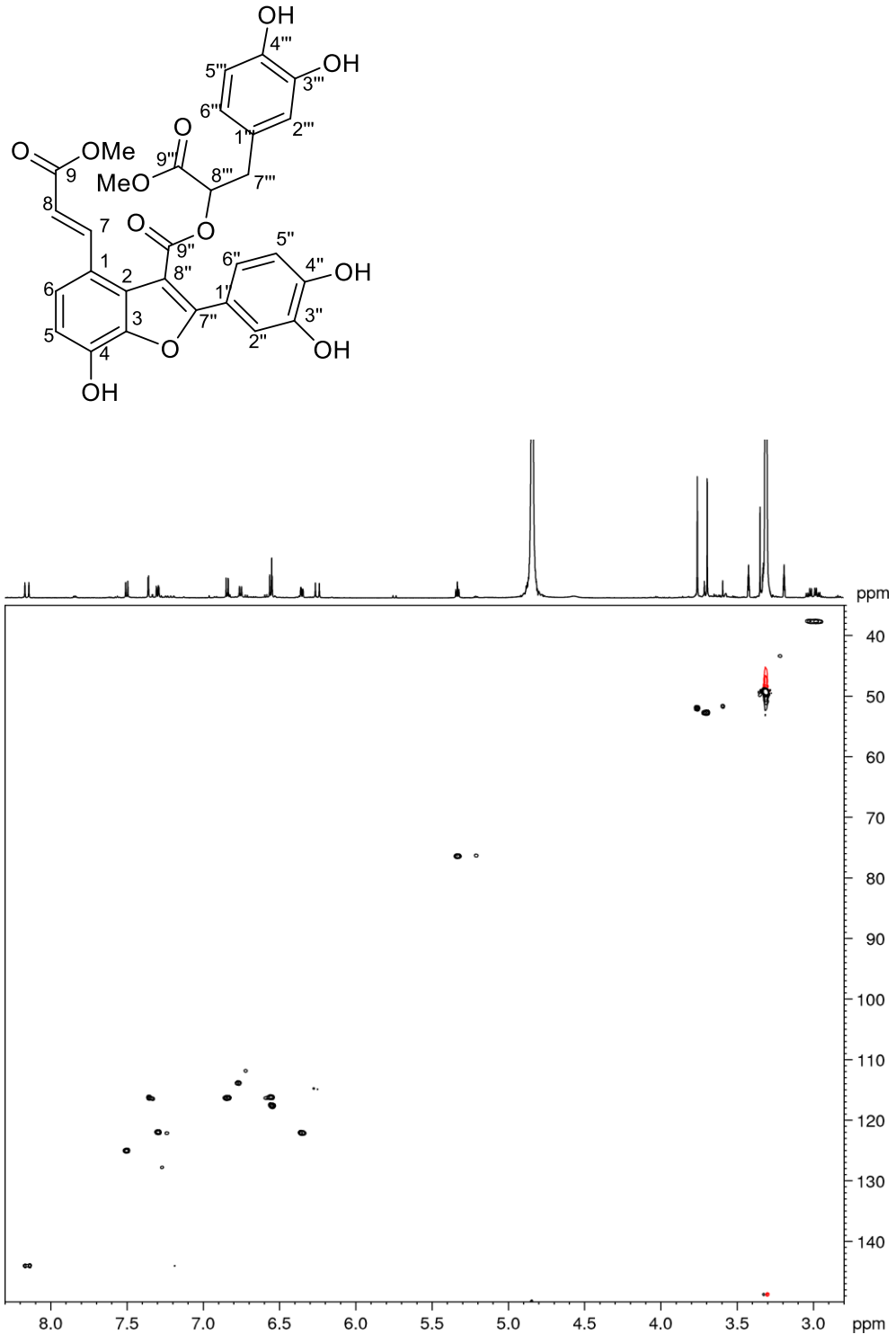


Figure 9.23 HSQC spectrum of the partially methanolized compound **11** (CD_3OD, 600 MHz).

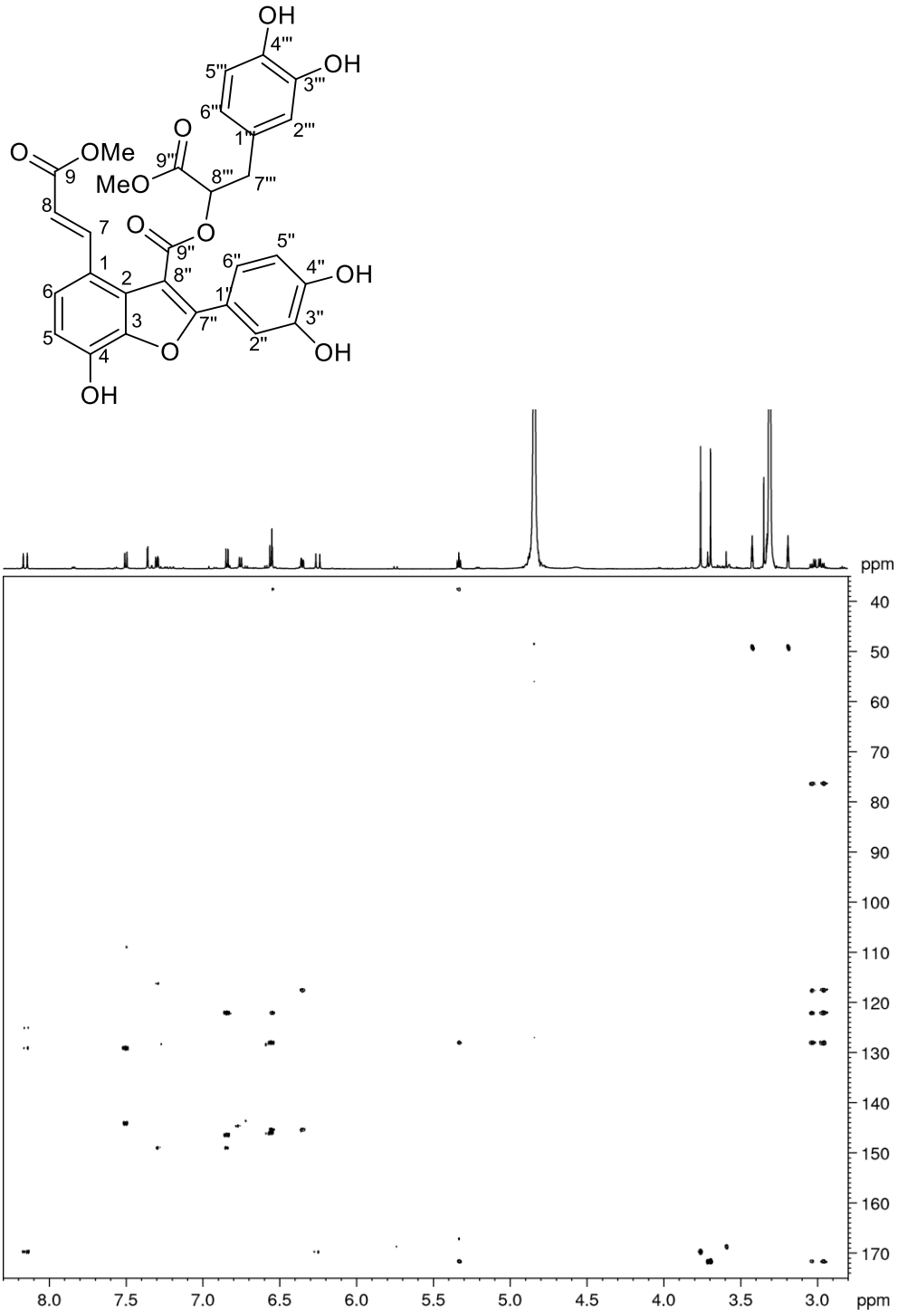
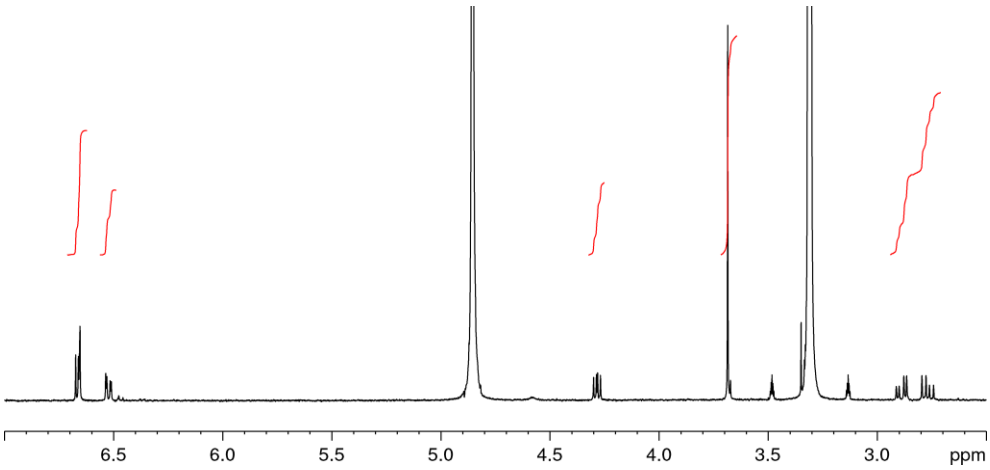


Figure 9.24 HMBC spectrum of the partially methanolized compound **11** (CD_3OD, 600 MHz).



F**igure 9.25** ^{1}H NMR spectrum of methyl (3,4-dihydyroxyphenyl)lactate (**8**) from methanolysis of 7",8"-didehydrosalvianolic acid B (**7**) (700 MHz, CD_3OD).

References

[1] Li, Y.; Rárová, L.; Scarpato, S.; Sezai Çiçeka, S.; Jordheimd, T.; Štenclováb, T.; Strnade, M.; Mangoni, A.; Zidorn, C. Seasonal variation of phenolic compounds in *Zostera marina* (Zosteraceae) from the Baltic Sea. *Phytochemistry,**2022**, 196,113099,* https://doi.org/10.1016/j.phytochem.2022.113099

[2] Enerstvedt, K.H.; Jordheim, M.; Andersen, Ø.M. Isolation and Identification of Flavonoids Found in <i>Zostera marina</i> Collected in Norwegian Coastal Waters. *Am. J. Plant Sci.* **2016**, *07*, 1163–1172, doi:10.4236/ajps.2016.77111.

[3] Grignon-Dubois, M.; Rezzonico, B. Phenolic chemistry of the seagrass Zostera noltei Hornem. Part 1: First evidence of three infraspecific flavonoid chemotypes in three distinctive geographical regions. *Phytochemistry* **2018**, *146*, 91–101, doi:10.1016/j.phytochem.2017.12.006.

[4] Hasle Enerstvedt, K.; Lundberg, A.; Jordheim, M. Characterization of Polyphenolic Content in the Aquatic Plants Ruppia cirrhosa and Ruppia maritima - A Source of Nutritional Natural Products. *Molecules* **2017**, *23*, doi:10.3390/molecules23010016.

[5] Gong, J.; Ju, A.; Zhou, D.; Li, D.; Zhou, W.; Geng, W.; Li, B.; Li, L.; Liu, Y.; He, Y.; et al. Salvianolic acid Y: A new protector of PC12 cells against hydrogen peroxide-induced injury from Salvia officinalis. *Molecules* **2015**, *20*, 683–692,

[6] Wang, J.; Pan, X.; Han, Y.; Guo, D.; Guo, Q.; Li, R. Rosmarinic acid from eelgrass shows nematicidal and antibacterial activities against pine wood nematode and its carrying bacteria. *Mar. Drugs* **2012**, *10*, 2729–2740, doi:10.3390/md10122729.

[7] Trute, A.; Nahrstedt, A. Separation of rosmarinic acid enantiomers by three different chromatographic methods (HPLC, CE, GC) and the determination of rosmarinic acid in Hedera helix L. *Phytochem. Anal.* **1996**, *7*, 204–208, doi:10.1002/(SICI)1099-

[8] Rárová, L.; Steigerová, J.; Kvasnica, M.; Bartůněk, P.; Křížová, K.; Chodounská, H.; Kolář, Z.; Sedlák, D.; Oklestkova, J.; Strnad, M. Structure activity relationship studies on cytotoxicity and the effects on steroid receptors of AB-functionalized cholestanes. *J. Steroid Biochem. Mol. Biol.* **2016**, *159*, 154–169, doi:10.1016/j.jsbmb.2016.03.017.

[9] Seflova, J., Cechova, P., Biler, M., Hradil, P., Kubala, M. Inhibition of Na+/K+-ATPase by 5,6,7,8-tetrafluoro-3-hydroxy-2-phenylquinolin-4(1H)-one. *Biochimie*, **2017**, *138,* 56e61. https://doi.org/10.1016/j.biochi.2017.04.009.

[10] Kubala, M., Cechova, P., Geleticova, J., Biler, M., Stenclova, T., Trouillas, P., Biedermann, D., Flavonolignans as a novel class of sodium pump inhibitors. *Front. Physiol.* **2016**, *7*, 115. https://doi.org/10.3389/fphys.2016.00115.

[11] Baginski, E.; Weiner, L.M.; Zak, B. The simple determination of nucleotide phosphorus. *Clin. Chim. Acta* **1964**, *10*, 378–379, doi:10.1016/0009-8981(64)90071-3.

[12] Cariani, L.; Thomas, L.; Brito, J.; Del Castillo, J.R. Bismuth citrate in the

quantification of inorganic phosphate and its utility in the determination of membrane-bound phosphatases. *Anal. Biochem.* **2004**, *324*, 79–83, doi:10.1016/j.ab.2003.09.008.

PART 3

Exploration of the extracts from cyanobacteria

Chapter 10

Cyanobacteria

Among the first life forms found on the planet it is possible to include cyanobacteria (Cyanophyta) or blue-green algae, also known as *Cyanophyceae*. Cyanobacteria are often mistakenly confused with algae, thanks to the ability of some genera to organize themselves in filamentous structures. Indeed, they can live as solitary, free-living cells or as colonies or filaments.[1]

Figure 10.1. Cyanobacterial cells.
(Image retrieved from https://www.deq.ok.gov/state-environmental-laboratory-services/environmental-public-health-information/harmful-algal-blooms/what-are-cyanobacteria/)

Cyanobacteria have made a significant contribution to the development of life on the planet, given their ability to grow in conditions of anoxia and at the same time to release oxygen into the atmosphere, thanks to their photosynthetic activity, whose slow accumulation over billions of years, has contributed to the development of animal life.[2] Cyanobacteria proliferate significantly in stagnant fresh water, and in the presence of nutrients, mainly nitrogen and phosphorus, at a temperature that reaches more than 20 ° C,

illuminated by sunlight. However, they can easily adapt to extreme conditions and waters with different degrees of salinity and temperature.[3] It is also possible to count some species of thermophilic cyanobacteria which can tolerate temperatures of 50-60° C.[4]

Cyanobacteria can live freely but a certain number of them are able to form symbiotic associations with other organisms. They are unicellular prokaryotic organisms, showing a gram-negative type of wall and a characteristic blue-green coloration, due to the simultaneous presence of accessory pigments and chlorophyll A. The presence of accessory pigments such as phycobilins, which include c-phycocyanin, blue in color containing copper; c-phycoerythrin, red in color with iron; allophycocyanin and phycoerythrocyanin, enables to broaden the spectrum of usable light, allowing cyanobacteria to grow even at low intensity of light radiation.[5] They are also rich in cyanophycin (polymer of arginine and asparagine) which is a nitrogen reserve and a source of energy. Interesting is their ability to transform atmospheric nitrogen (N_2) into ammoniacal nitrogen (NH_3 and NH_4^+), process known as azofixation, making it available to plant life forms. From a morphological point of view, cyanobacteria exhibit a high dimensional variability, ranging range from unicellular organisms of 0.2 μm to filamentous forms with length up to 200 μm. They exhibit a system of equidistant photosynthetic membranes called thylakoids, which they store during the light phase of photosynthesis. Differing from the other photosynthetic bacteria, they show the dual photosystem (PS I and PS II), which uses water as photoreductant producing chemical energy with the consequent liberation of oxygen. The central portion of the cytoplasm is

called centrosome and is the area where DNA forms chromosome-like structures.[6] Due to the presence of gaseous vacuoles, cyanobacteria can move actively along the water column, modulating their buoyancy according to changes in light, water turbulence and nutrient availability. In this way they can also use the nutrients "trapped" in the coldest and deepest layers.[3]

It has been verified that the activities of man, like urbanization and agricultural activities, but also particular environmental conditions, such as increasing global temperatures, contribute to recreate conditions suitable for harmful cyanobacterial blooms. In particular, the presence of an excessive amount of nutrients (nitrogen, phosphorus, sulfur) in the water, caused by the release of industrial effluents contributing to the formation of cyanobacteria blooms.

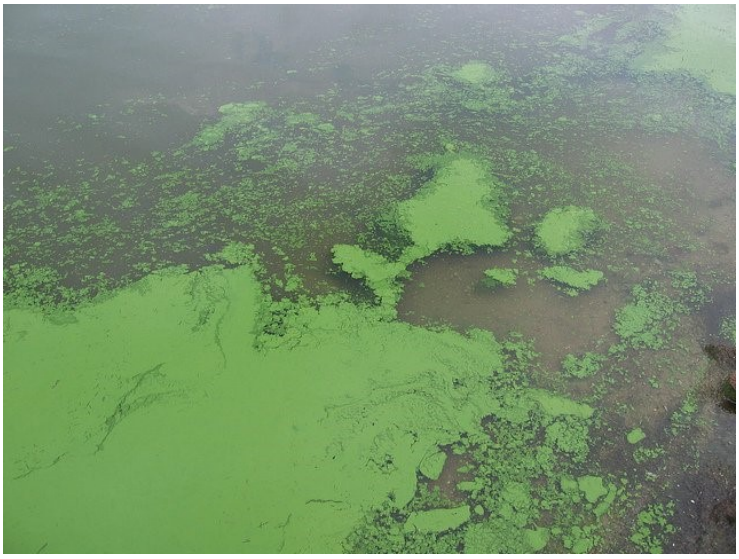


Figure. 10.2 A cyanobacterial bloom
(Image retrieved from
https://www.flickr.com/photos/48722974@N07/5120831456/)

The uncontrolled proliferation of cyanobacteria causes devastating effects on the ecosystem, given the ability of some cyanobacteria (such as

Oscillatoria, *Nodularia*, *Microcystis* and *Anabaena*) to produce a large pool of toxins as secondary metabolites, including neuro and hepatotoxins.[7] It has been shown that some of these can be released into the aquatic environment surrounding the cell by active transport processes to defend themselves and/or attack competing organisms. The massive production of cyanotoxins, mainly peptide molecules that exhibit toxicity to both animals and humans, makes cyanobacterial blooms one of the world's biggest problems. Among cyanotoxins, it is worth mentioning the extensively studied microcystins (MCs), with about 100 congeners, which easily accumulate in fish, mussels, representing a major food safety concern.

The research carried out during this PhD project was focused on members of the cyanobacterial genus *Trichodesmium*, which are known to play a predominant ecological role given their nitrogen-fixing capacity. The specific metabolome of *Trichodesmium* has been extensively investigated in recent years, showing the presence of several natural products, mainly chlorinated polyketides, along with hybrid polyketide-peptide metabolites, some of which have shown moderately potent cytotoxicity against neuronal cell lines. The extensive search for new secondary metabolites from *Trichodesmium* extracts, led to the isolation to four new compounds counted in the trichophycin family.[8]

In addition, great attention has been directed toward the study of an unusual out-of-season toxic bloom of cyanobacteria in Lake Avernus, detected in the months prior to the COVID-19 lockdown, through Fast detection strategies (FDS), integrating satellite imagery and biomolecular investigation. The

bioactivity of the organic extracts of this winter bloom was evaluated revealing a high cytotoxic activity.[9]

References

[1] Catherine, Q.; Susanna, W.; Isidora, E.S.; Mark, H.; Aurélie, V.; Jean-François, H. A review of current knowledge on toxic benthic freshwater cyanobacteria - Ecology, toxin production and risk management. *Water Res.* **2013**, *47*, 5464–5479, doi:10.1016/j.watres.2013.06.042.

[2] Schirrmeister, B.E.; Gugger, M.; Donoghue, P.C.J. Cyanobacteria and the Great Oxidation Event: Evidence from genes and fossils. *Palaeontology* **2015**, *58*, 769–785, doi:10.1111/pala.12178.

[3] Huisman, J.; Codd, G.A.; Paerl, H.W.; Ibelings, B.W.; Verspagen, J.M.H.; Visser, P.M. Cyanobacterial blooms. *Nat. Rev. Microbiol.* **2018**, *16*, 471–483, doi:10.1038/s41579-018-0040-1.

[4] Patel, A.; Matsakas, L.; Rova, U.; Christakopoulos, P. A perspective on biotechnological applications of thermophilic microalgae and cyanobacteria. *Bioresour. Technol.* **2019**, *278*, 424–434, doi:10.1016/j.biortech.2019.01.063.

[5] Pagels, F.; Guedes, A.C.; Amaro, H.M.; Kijjoa, A.; Vasconcelos, V. Phycobiliproteins from cyanobacteria: Chemistry and biotechnological applications. *Biotechnol. Adv.* **2019**, *37*, 422–443, doi:10.1016/j.biotechadv.2019.02.010.

[6] Castenholz, R.W. General Characteristics of the Cyanobacteria . *Bergey's Man. Syst. Archaea Bact.* **2015**, 1–23, doi:10.1002/9781118960608.cbm00019.

[7] Huang, I.S.; Zimba, P. V. Cyanobacterial bioactive metabolites—A review of their chemistry and biology. *Harmful Algae* **2019**, *83*, 42–94, doi:10.1016/j.hal.2018.11.008.

[8] McManus, K.M.; Kirk, R.D.; Via, C.W.; Lotti, J.S.; Roduit, A.F.; Teta, R.; Scarpato, S.; Mangoni, A.; Bertin, M.J. Isolation of isotrichophycin C and trichophycins G–I from a collection of trichodesmium thiebautii. *J. Nat. Prod.* **2020**, *83*, 2664–2671, doi:10.1021/acs.jnatprod.0c00550.

[9] Teta, R.; Sala, G. Della; Esposito, G.; Stornaiuolo, M.; Scarpato, S.; Casazza, M.; Anastasio, A.; Lega, M.; Costantino, V. Monitoring cyanobacterial blooms during the COVID-19 pandemic in Campania, Italy: The case of lake avernus. *Toxins (Basel).* **2021**, *13*, doi:10.3390/toxins13070471.

Chapter 11

Identification of new members of the trichophycin family from a collection of *Trichodesmium thiebautii*

11.1 Synopsis of this chapter [1]

Environmental collections of *Trichodesmium sp.* bloom revealed the presence of many chlorinated polyketides, included in the trichophycin family. With the main aim to in-depth explore the variety of trichophycin compounds, extracts of the cyanobacterium *Trichodesmium thiebautii*, were analyzed by MS/MS-based molecular networking, obtaining an overall visualization of halogenated compounds present. Among many unknown chlorinated metabolites, isotrichophycin C (**12**) and trichophycin G-I (**13-15**) were isolated. The variability in the structure from previously characterized trichophycins, suggested a SAR study to explore the structural features associated with toxicity to murine neuroblastoma cells.

11.2 Previously identified compounds from the extracts of the sponge Smenospongia and trichodesmium

Organic extracts of the cyanobacterium *Trichodesmium sp*, represent a remarkable source of new compounds with polyketide and hybrid polyketide-peptide structures. [2,3,4,5] In particular, it is important to mention the interesting discovery of some common compounds such as smenamide A and smenothiazole A, three structural analogues of smenamide A (smenamides C-E), together with a group of chlorinated polyketides with a

smenamide-like structure (trichophycin A-F) in both extracts of *Trichodesmium sp* and the sponge *Smenospongia aurea*.[6] This finding led to a comparative study of the sponge and cyanobacteria extracts, through MS-based molecular networking, to assess the traits in common of the two metabolomes. The results showed the presence of four new polyketides, smenolactones A-D, in the *Smenospongia aurea* extracts and two new conulothiazole analogs, isoconulothiazole B and conulothiazole C from the *Trichodesmium* extracts.

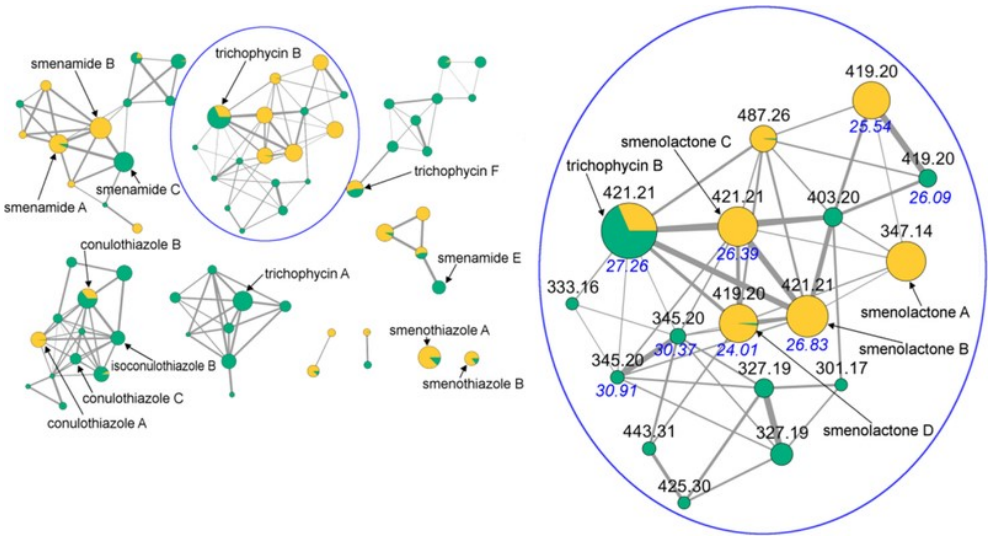


Figure 11.1. The molecular network generated by combining *S. aurea* (yellow) and *Trichodesmium sp* (blue) extracts. Node size is relative to ion abundance while edge thickness is a function of cosine similarity score. Nodes are labeled with m /z values while retention times of isomeric compounds are indicated in blue.

The antiproliferative activity shown by smenolactones and the moderately potent cytotoxicity against neuronal cell lines shown by trichophycins prompted to a further analysis of *Trichodesmium* extracts in search of new therapeutically relevant compounds.

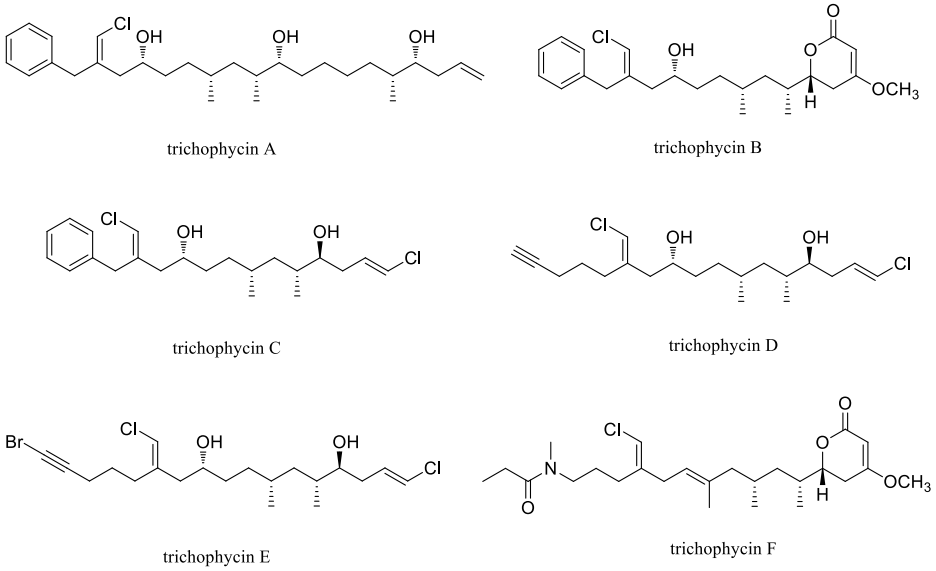


Figure 11.2. Structure of the previously isolated Trichophycin A-F

11.3 Collection and extraction of samples

A cyanobacterial sample was collected at the water's surface in the Gulf of Mexico. The evaluation of some morphological characters suggested a *Trichodesmium* species, whose identification as a strain of *Trichodesmium thiebautii*, was clarified by phylogenetic analysis. Its lipophilic extracts were in-depth analyzed.

11.4 Molecular Networking of extracts of Trichodesmium thiebatii

The analysis of the environmental collection of bloom material, represent a challenging activity due to the heterogeneity of organisms, which can be present during a bloom event, such as diatoms, dinoflagellates, and heterotrophic bacteria. [7,8,9] To simplify and to visualize the overall composition of metabolites, MS/MS-based molecular networking was performed. Since the interest was mainly in trichophycins, which are characterized by a chlorovinylidene functionality, a specific workflow pinpointing on geometric isomers and chlorinated metabolites was

performed on Mzmine2 software. [10] In the resulting molecular map, nodes were color-mapped according to the isotopic patterns of the compounds, clearly showing the presence of over 100 metabolites containing at least one chlorine atom.

A substantial number of halogenated compounds, in particular metabolites containing at least one chlorine atom, including some previously isolated compounds, could be verified in the molecular network. In addition, the molecular network showed some nodes, representing already known metabolites, isolated from extracts of the marine sponge *Smenospongia aurea* (smenothiazole B, conulothiazole A and B, and smenamide F). [11,12,13] From the analysis of molecular clusters showing some previously characterized trichophycins, new compounds, isotrichophycin C (**12**) and trichophycin H (**14**), were detected. The isolation of compounds **12** and **14** was performed. Also reported in this study was the isolation of trichophycin G (**13**), which appears as a single node in the network, probably due to low abundance in the extract or a lack of informative MS/MS fragments, and the isolation of trichophycin I (**15**), represented by no node in the network, but isolated during the chromatographic steps of isolation of the other compounds.

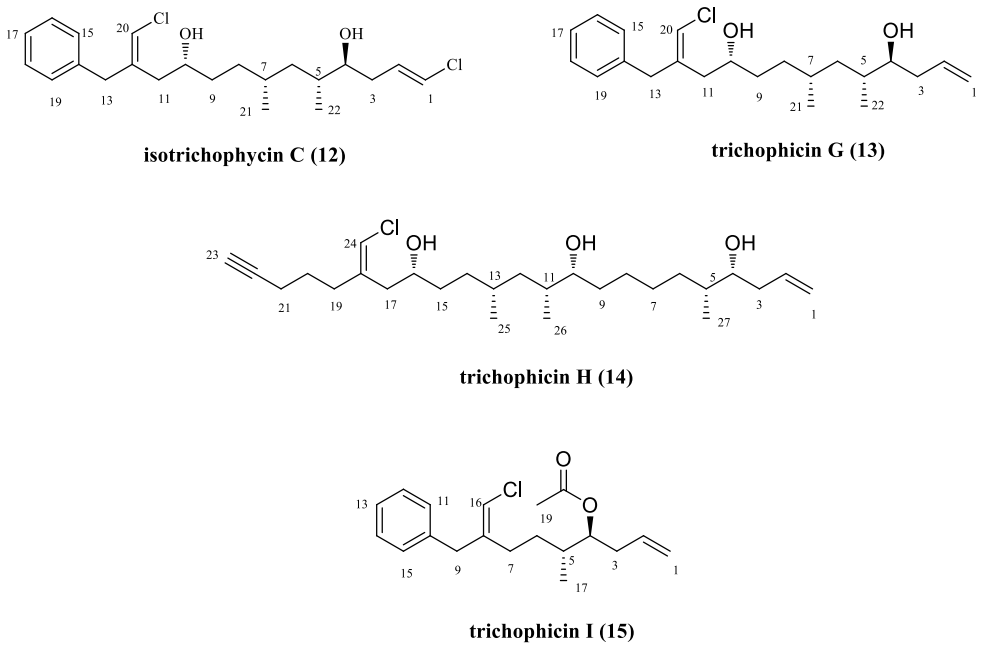


Figure 11.3. Structure of isotrichophycin C **(12)** and trichophycin G-I (**13-15**)

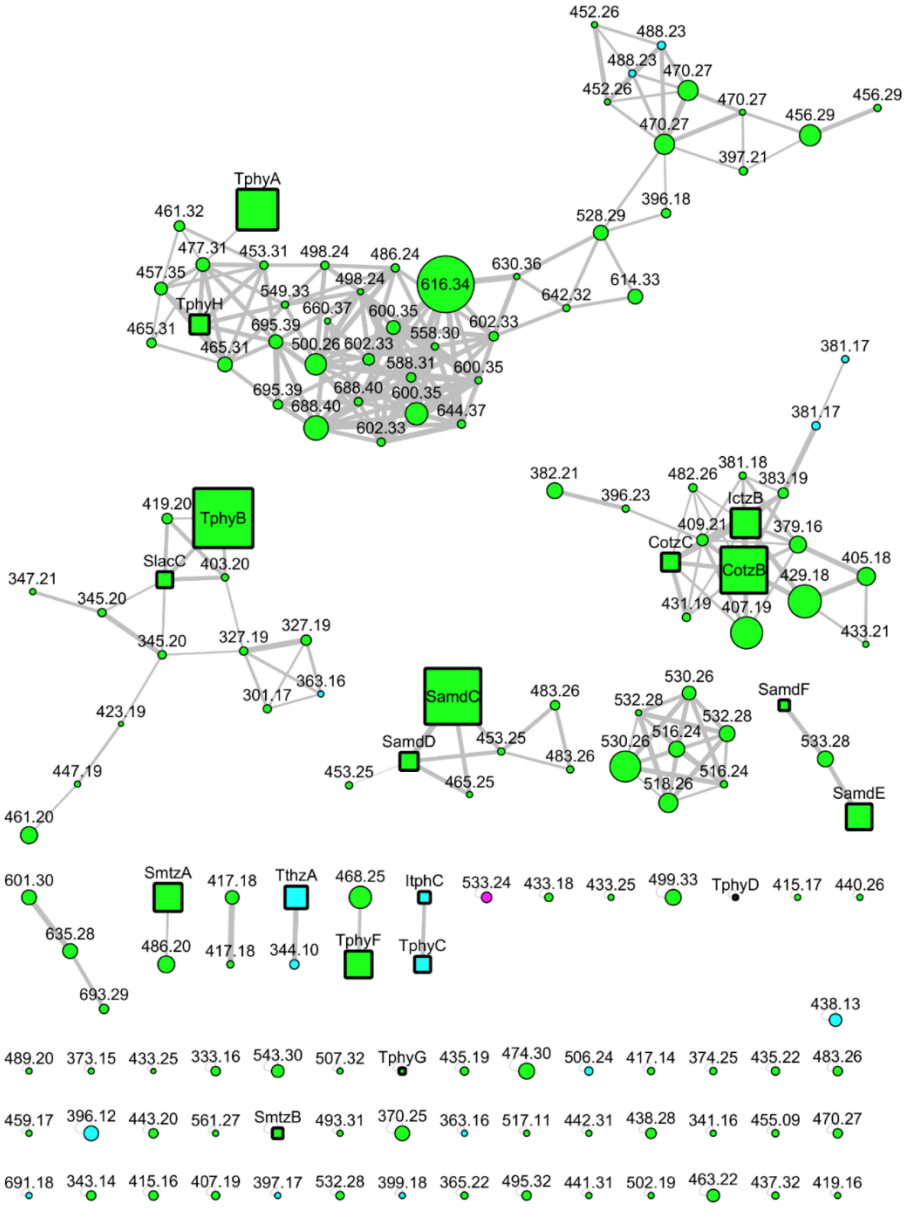


Figure 11.4. Molecular network of the *Trichodesmium* bloom extract. Only halogenated molecules are shown in the network. Color-coding of halogenation pattern is as follows: green = Cl, cyan = Cl2, and pink = ClBr. Node size is proportional to LC-MS peak area. Previously reported compounds and newly reported molecules (**12-14**) are shown as square nodes and abbreviated by name (Cotz, conulothiazole; Ictz, isoconulothiazole; Samd, smenamide; Tphy, trichophycin; Itph, isotrichophycin; Slac, smenolactone; Smtz, smenothiazole). Circular nodes are currently uncharacterized metabolites.

*11.5 Structural elucidation of isotrichophycin C (**12**) and trichophycin G-I (**13-15**)*

The high-resolution ESI mass spectrum of isotrichophycin C (**12**), showed an $[M+H]^+$ of *m/z* 399.1855, suggesting an identical molecular formula of the previously characterized trichophycin C (Figure 11.2), $C_{22}H_{32}Cl_2O_2$.

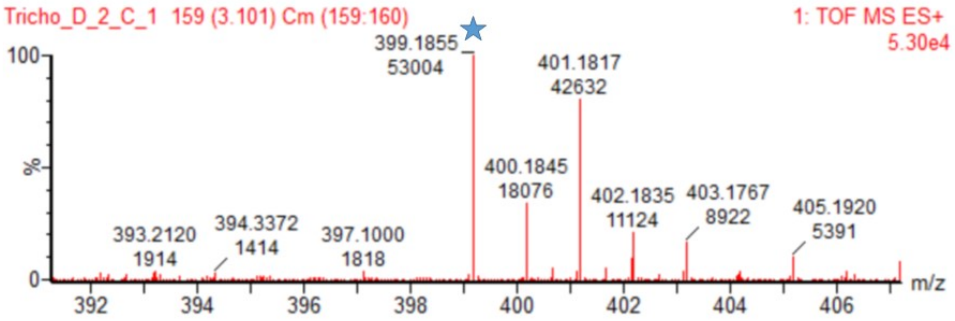


Figure 11.5. HR-ESI-MS of isotrichophycin C (**12**)

However, the analysis of ^{1}H NMR and ^{13}C NMR spectra was indicative of a similar but not identical structure to trichophycin C (Figure 11.2). Moreover, the presence of a node referred to trichophycin C in the same cluster of isotrichophycin C (**12**) in the molecular network confirmed it was a new chemical entity (Figure 11.4, TphyC-ItphC cluster).

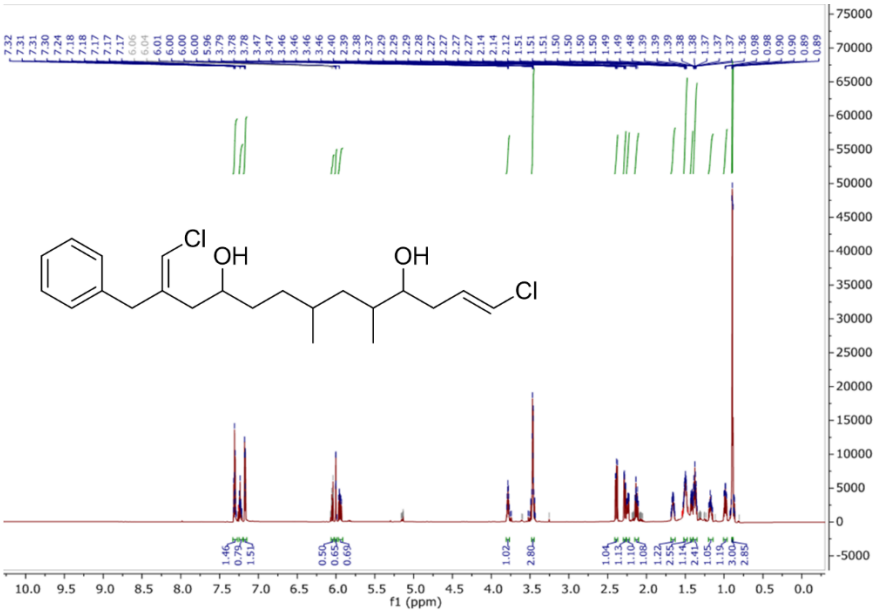


Figure 11.6. ^{1}H NMR spectrum of isotrichophycin C (**12**) (800 MHz, $CDCl_3$).

The main differences between isotrichophycin C (**12**) and trichophicin C were clarified by NMR signals, which showed diverse chemical shift from C-9 to C-13. The analysis of NOESY spectrum revealed correlations between H-20 (δ_H 6.00) and H-15 (δ_H 7.17) and H_2-13 (δ_H 3.46) supporting a *Z* geometry of the chlorovinylidene group at C-12, in contrast to the *E*-geometry of the C-12 chlorovinylidene group of trichophycin C. Furthermore, the study of coupling constant of C-1 and C-2, showing a large vicinal ^{1}H-^{1}H coupling constant (13.2 Hz) determined a *E*-geometry.

The presence of no deviations in the chemical shifts of ^{13}C spectrum at the C-4, C-5 and C-7 stereogenic centers supported the hypothesis that the new compound was a *cis-trans* isomer of trichophycin C.

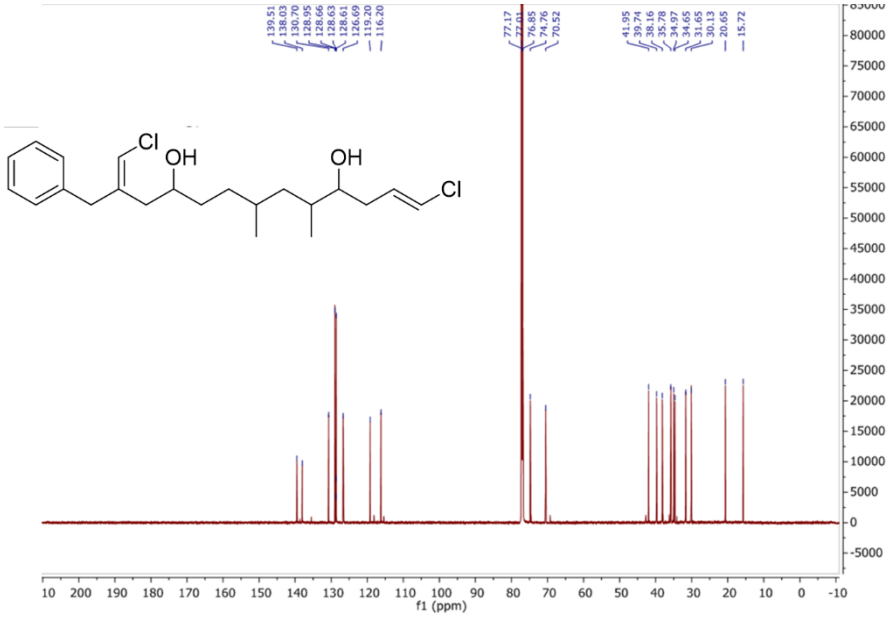
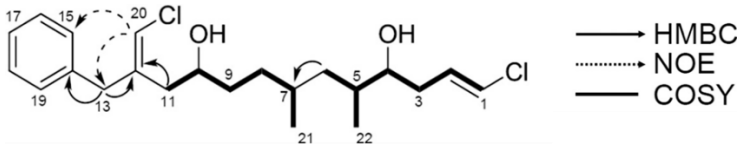


Figure 11.7. ^{13}C NMR spectrum of isotrichophycin C (**12**) (200 MHz, $CDCl_3$).

Table 11.1 NMR data for isotrichophycin C (**12**).

pos	δ_C, type	δ_H (*J* in Hz)	HMBC	COSY
1	119.2, CH	6.05, d (13.2)	2, 3	2
2	130.7, CH	5.95, m	1, 3, 4	1, 3a, 3b
3a	35.0, CH_2	2.24, dddd (14.4, 7.1, 3.3, 1.5)	1, 2, 4, 5	2, 3b, 4
3b		2.13, dddd (14.4, 9.2, 8.6, 1.2)	1, 2, 4, 5	2, 3a, 4
4	74.8, CH	3.46, ddd (8.7, 5.3, 3.4)	2, 3, 5, 6, 22	3a, 3b, 5
5	35.8, CH	1.66, m	3, 4, 6, 7, 22	4, 6a, 6b, 22
6a	39.8, CH_2	1.38, ddd (13.4, 8.5, 4.6)	4, 5, 7, 8, 21, 22	5, 6b
6b		0.98, ddd (13.6, 9.1, 5.5)	4, 5, 7, 8, 21, 22	5, 6a
7	30.1, CH	1.51, ovlp[a]	6, 8, 9, 21	8a, 8b, 21
8a	31.7, CH_2	1.37, ovlp	6, 7, 9, 10, 21	7, 8b
8b		1.17, m	6, 7, 9, 10, 21	7, 8a
9a	34.7, CH_2	1.50, ovlp	7, 8, 10, 11	8b, 9b
9b		1.41, m	7, 8, 10, 11	8b
10	70.5, CH	3.78, m	8, 9, 11, 12	9a, 9b, 11a, 11b
11a	38.2, CH_2	2.39, dd (13.5, 8.8)	9, 10, 12, 13, 20	10, 11b
11b		2.28, dd (13.5, 4.2)	9, 10, 12, 13, 20	10, 11a
12	139.5, C			
13	42.0, CH_2	3.46, ovlp	11, 12, 14, 20	
14	138.0, C			
15	129.0, CH	7.17, d (7.6)	13, 17	16
16	128.6, CH	7.30, t (7.6)	14	17
17	126.7, CH	7.24, t (7.6)	15/19	
18	128.6, CH	7.30, t (7.6)	14	
19	129.0, CH	7.17, d (7.6)	13, 17	18
20	116.2, CH	6.00, s	11, 12, 13	
21	20.7, CH_3	0.90, d (6.6)	6, 7, 8	7
22	15.7, CH_3	0.89, d (6.8)	4, 5, 6	5



12

Figure 11.8. Key 2D NMR correlations for isotrichophycin C (**12**)

However, isotrichophycin C (**12**), showed the opposite sign to that of trichophycin C (+26.8) by the examination of the optical rotation (-9.5). Further information was obtained by performing Mosher methodology, with the synthesis of diastereomeric bis-MTPA esters of isotrichophycin C (**12**). Detailed examination of ^{1}H chemical shifts of diastereomeric bis-MTPA esters allowed for assignment of absolute *R* and *S* configurations of C-10 and C-4 respectively, resulting in an identical configuration of trichophycin C.

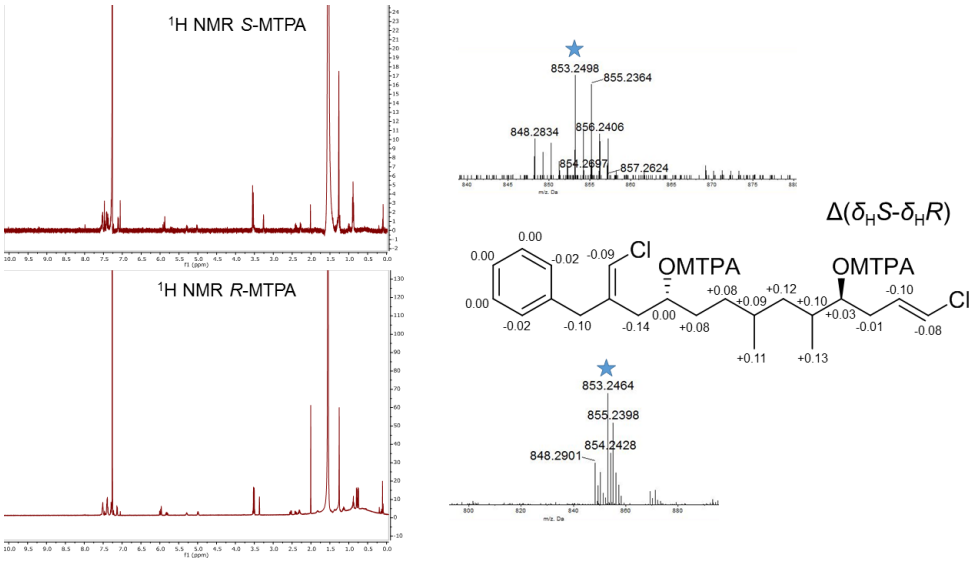


Figure 11.9 Comparison of ^{1}H NMR spectra of *S*-MTPA and *R*-MTPA esters of isotrichophycin C (**12**) with Δ *δS-δR* values noted (500 MHz, $CDCl_3$) as well as HRESIMS values.

Results were further supported by ECD analysis, which showed nearly overlapping spectra (Figure 11.10). A *syn* relationship between the two methyl groups was confirmed by the difference in chemical shift of the diastereotopic protons on C-6. [14] Moreover, the same coupling constant of 5.3 Hz measured for H-4 and H-5, in tricophycin C, was indicative of an *anti*-relationship between the two substituents. The *anti*-relationship between H-4 and H-5 in trichophycin C and isotrichophycin C (**12**) was, indeed, strongly supported by DFT calculations (M06-2X/6-31G(d,p) level of theory), revealing a calculated *J*-coupling constants between H-4 and H-5 of 4.2 Hz for 4R5S (*syn*) and 6.3 Hz for 4R5R (*anti*) for trichophycin C. [5] The in-depth analysis of NMR data confirmed isotricophycin C (**12**) as a *cis-trans* isomer of tricophycin C.

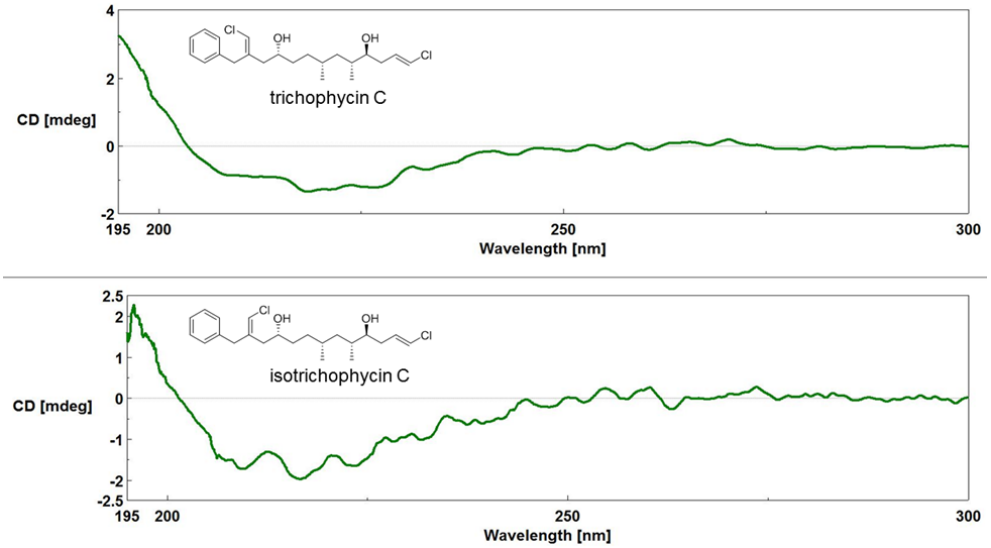


Figure 11.10. Comparison of trichophycin C and isotrichophycin C (**12**) ECD spectra (0.4 mg/mL, CH_3CN).

Showing the HR-ESI-MS spectrum of the compound **13** a $[M+Na]^+$ of *m/z* 387.2071, suggested a molecular formula of $C_{22}H_{33}ClO_2$. The presence of a single chlorine atom was confirmed by the intensity ratio in the trichophycin G (**13**) isotope pattern.

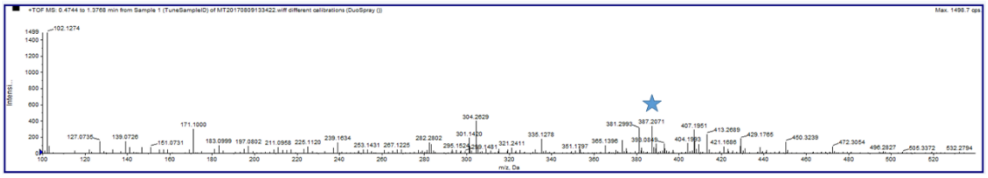


Figure 11.11. HR-ESI-MS spectrum of trichophycin G **(13)**.

By comparing the ^1H NMR spectra of isotrichophycin C (**12**) and trichophycin G (**13**), it was possible to see that, although they were very similar, the resonances at δ_H 5.17 and 5.14 (H-1a and H-1b, respectively) were correlated with H-2 (δ_H 5.84) in trichophycin G (**13**) in contrast to isotrichophycin C (**12**).

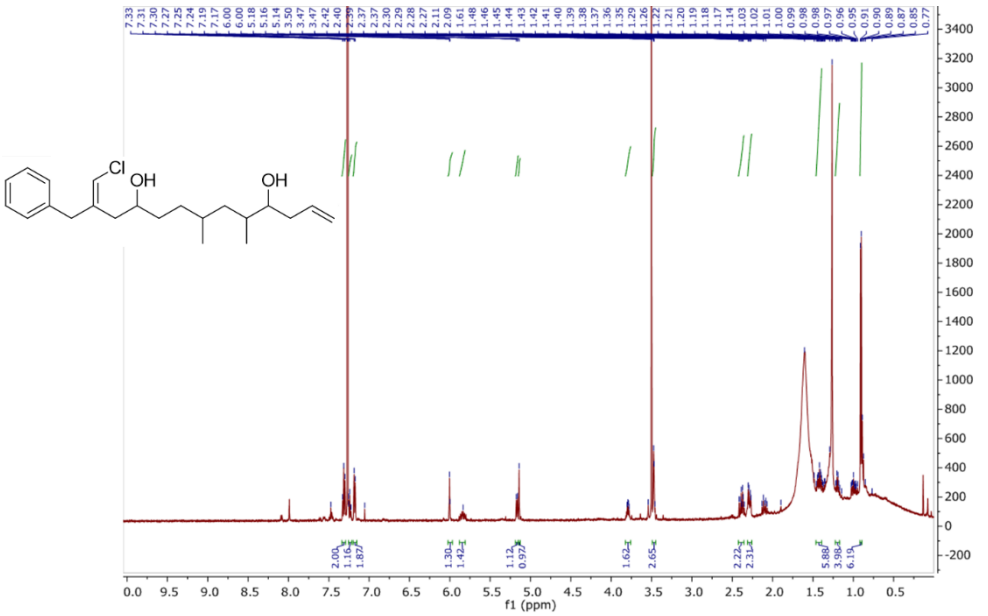
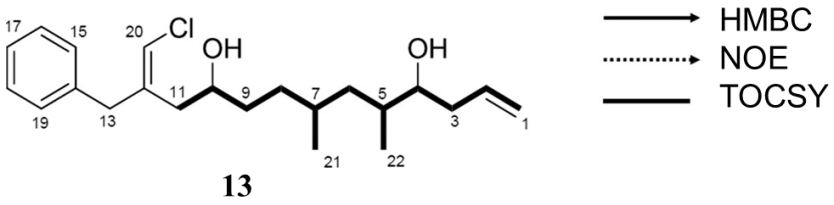


Figure 11.12. ^1H NMR spectrum of trichophycin G (**13**) (500 MHz, $CDCl_3$).

From the ^{1}H-^{1}H COSY cross peaks, it was evident that trichophycin G (**13**), presents a terminal alkene rather than the terminal vinyl chloride of isotrichophycin C (**12**). Inspection of the NMR data and optical rotation values allowed the absolute configuration to be defined identical to that of isotrichophycin C (**12**).



13

Figure 11.13. Key 2D NMR correlations for trichophycin G (**13**)

Table 11.2 NMR data for trichophycin G (**13**) (500 MHz for ^{1}H NMR, $CDCl_3$).

pos	δ_H (*J* in Hz)	TOCSY
1a	5.17, m	2
1b	5.14, m	2
2	5.84, m	1, 3b
3a	2.32, m	2, 3b, 4
3b	2.10, m	2, 3a, 4
4	3.47, ovlp[a]	3a, 3b, 5
5	1.63, m	4, 6a, 6b, 22
6a	1.42, ddd (13.5, 8.6, 4.6)	5, 6b
6b	1.00, ddd (14.2, 9.2, 5.6)	5, 6a
7	1.50, ovlp	21
8a	1.35, ovlp	8b
8b	1.19, m	8a
9a	1.50, ovlp	
9b	1.45, m	
10	3.79, m	9a, 9b, 11a, 11b
11a	2.39, dd (13.7, 8.8)	10, 11b
11b	2.28, dd (13.5, 4.2)	10, 11a
12		
13	3.47, ovlp	20
14		
15	7.18, d (7.4)	16
16	7.31, t (7.4)	17
17	7.24, m	
18	7.31, t (7.4)	
19	7.18, d (7.4)	18
20	6.00, s	13
21	0.91, ovlp	7
22	0.91, ovlp	5

[a]overlapping signals

High resolution ESI-MS analysis of compound (**14**) gave an $[M+Na]^+$ of *m/z* 477.3107, suggesting a molecular formula of $C_{27}H_{47}ClO_3$, which counted for 4 degrees of unsaturation.

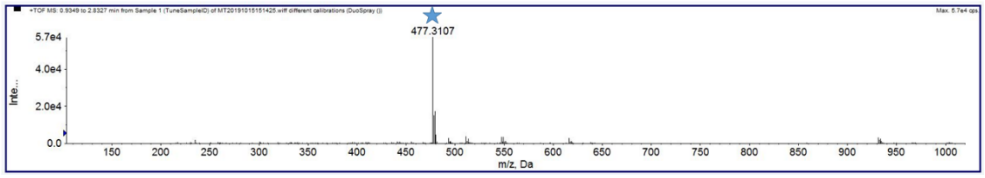


Figure 11.14. HR-ESI-MS spectrum of trichophycin H **(14)**.

It was possible to visualize the node of compound **14**, referred to its $[M+H]^+$ (*m/z* 455, TphyH) in the same cluster of trichophycin A (TphyA) in the molecular network (Figure 11.4). The similarity between this compound and the previously reported trichophycin A, was evaluated by the comparison of their ^1H NMR spectra. The ^1H NMR spectrum of trichophycin H (**14**) showed no coherent resonances with the presence of a benzene ring (Figure 11.15).

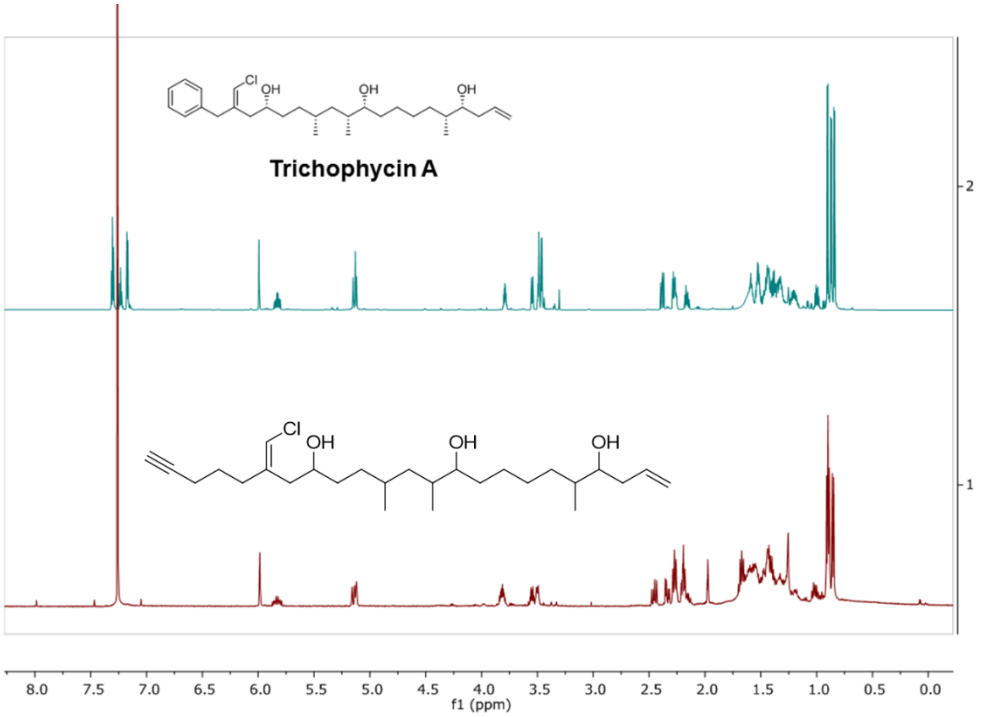


Figure 11.15. Comparison of ^{1}H NMR spectra: trichophycin A (800 MHz, $CDCl_3$) and trichophycin H (**14**) (500 MHz, $CDCl_3$).

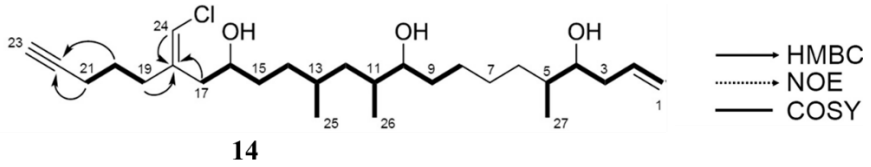
The in-depth examination of ^{1}H NMR spectrum showed a resonance δ_H 1.98, H-23, absent in the ^{1}H NMR spectrum trichophycin A, which showed a correlation to H_2-21 in the ^{1}H-^{1}H COSY spectrum. (Figure 11.31) COSY spectrum investigation exhibited a correlation between that H_2-21 and H_2-20 (δ_H 1.68), the latter itself was correlated to the deshielded methylene H_2-19 (δ_H 2.28)., which itself was correlated to C-18 (δ_C 139.0) by the following interpretation of the HMBC spectrum of trichophycin H (**14**). The presence of the alkyne group was established thanks to the correlations between H_2-20, H_2-21 and C-22 (δ_C 83.5), which were evident in the HMBC spectrum. Moreover, it was possible to establish an identical absolute configuration to that of trichophycin A, from the comparison of the relative configuration and ^{13}C NMR chemical shifts. (Table 11.4). By analyzing the chemical shift

values of a methylene group adjacent to chlorovinylidene in both *Z* and *E*-configurations shown in trichophycins and other metabolites, it was possible to assign the configuration of chlorovinylidene in trichophycin H (**14**). More specifically, the methylene carbon adjacent to chlorovinylidene (C-17 in **14**) is shielded in molecules with *Z* configuration, differing from the same carbon in compounds with *E*- configuration (Table 11.3), so it was possible to assign a *Z* configuration to chlorovinylidene in **14** because the chemical shift of C-17 was δ_C 38.2.

Table 11.3 NMR data for trichophycin H (**14**) (500 MHz for ^{1}H NMR, $CDCl_3$).

pos	δ_C, type	δ_H (*J* in Hz)	HMBC	COSY	$\Delta\delta$ ^{13}C[a]
1a	117.9, CH_2	5.15, m		2	0.0
1b		5.13, m		2	
2	135.6, CH	5.87, m		1, 3a, 3b	0.0
3a	39.0, CH_2	2.26, m		4	0.1
3b		2.17, m		4	
4	73.9, CH	3.56, dt (8.5, 4.0)		3a, 3b, 5	0.0
5	37.8, CH	1.54, ovlp[b]		4, 27	0.0
6a	33.0, CH_2	1.45, m		6b	0.0
6b		1.26, ovlp		6a	
7a	27.4, CH_2	1.41, ovlp		7b	0.0
7b		1.35, ovlp		7a	
8a	26.7	1.42, ovlp			0.0
8b		1.32, m			
9	34.6, CH_2	1.42, ovlp	10, 11	8b	0.0
10	74.5, CH	3.50 dt (7.6, 4.0)		9, 11	0.0
11	35.2, CH	1.61, m		10, 26	0.0
12a	40.8, CH_2	1.41, ovlp	10, 25, 26	11, 13	0.0
12b		1.02, dt (14.1, 7.4)	10, 25, 26	11, 12a	
13	29.8, CH	1.57, ovlp		14a, 25	0.0
14a	32.6, CH_2	1.42, ovlp		14b	0.1
14b		1.26, ovlp		14a	
15a	34.8, CH_2	1.55, ovlp		15b	0.0
15b		1.44, ovlp		15a	
16	70.6, CH	3.82, m		15a, 17a, 17b	0.1
17a	38.2, CH_2	2.46, dd (13.6, 8.7)	16, 18, 19, 24	16, 17b	0.0
17b		2.36, dd (13.6, 4.4)	16, 18, 24	16, 17a	
18	139.0, C				
19	34.2, CH_2	2.28, m	18, 20, 21, 24	20	
20	26.3, CH_2	1.68, m	18, 19, 21, 22	19a, 19b	
21	17.8, CH_2	2.20	19, 22	20	
22	83.5, C				

[a]$\Delta\delta$ ^{13}C (trichophycin A-**3**)
[b]overlapping signals



14

Figure 11.16. Key 2D NMR correlations for trichophycin H (**14**)

Thanks to the results of the analysis of the tris-MTPA esters of trichophycin A and trichophycin H (**14**), it was possible to conclude that they had the same absolute configuration (4*R*,10*R*,16*R*). Furthermore, the configuration assignment was confirmed due to the coupling constants between H-4 and H-5 and between H-10 and H-11, which were found to have the same values of trichophycin A and due to the large difference in the chemical shifts of the diastereotopic protons attached to C-12 (Δ = 0.39 ppm).

High-resolution mass spectrum analysis of trichophycin I (**15**) revealed a $[M+Na]^+$ of *m/z* 343.1454, suggesting a molecular formula of $C_{19}H_{25}ClO_2$ with 7 degrees of unsaturation.

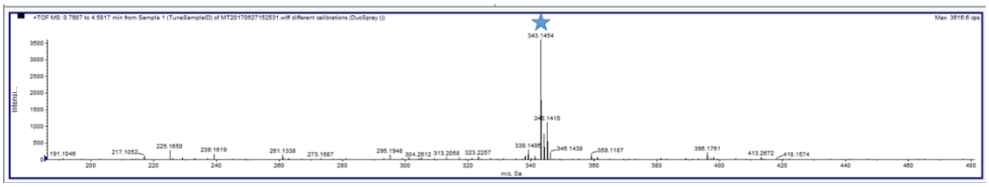


Figure 11.17. HR-ESI-MS spectrum of trichophycin I **(15)**.

Careful investigation of the NMR data revealed the presence of the monosubstituted benzene ring, deshielded methylene, and chlorovinylidene (C-8 to C-16), which agreed with the trichophycin family (Table 11.4). To clarify the geometry of the vinyl chloride group, NOE correlations were

studied. The cross-peak between H-16 (δ_H 5.84) and H_2-9 (δ_H 3.36) supported a *Z*-geometry of the vinyl chloride group.

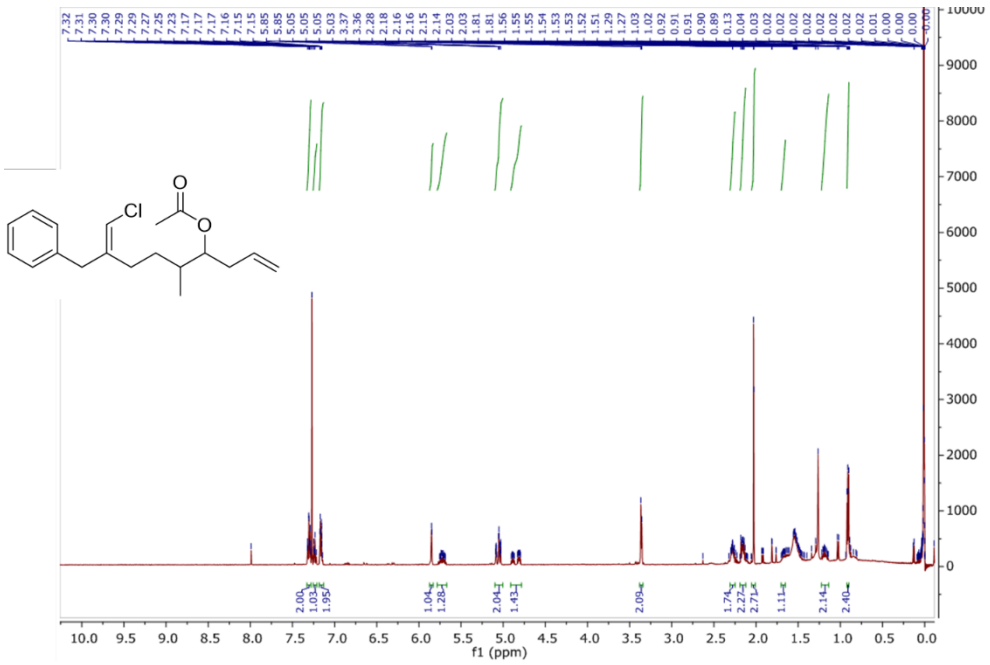
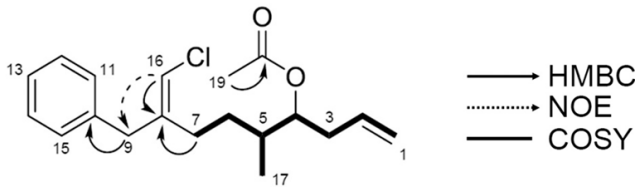


Figure 11.18. ^{1}H NMR of trichophycin I (**15**) (500 MHz, $CDCl_3$).

Moderately deshielded methylene protons (H_2-7, δ_H 2.14) were correlated with C-8 (δ_C 141.9), in HMBC spectrum, and with H_2-6 (H-6a, δ_H 1.50; H-6b, δ_H 1.17) in COSY spectrum. The COSY spectrum showed a cross-peak between the methylene group H-6 with the methine H-5 (δ_H 1.67). The latter was correlated with H_3-17 (δ_H 0.90) and oxymethine H-4 (δ_H 4.80). In addition, correlations between the methylene group H_2-3 at oxymethine and H-2 (δ_H 5.71), and correlations between H-2 and H_2-1 (H-1a, δ_H 5.07; H-1b, δ_H 5.03) were evident in the COSY spectrum, defining a terminal alkene group in trichophycin I (**15**). Compared to the other members of the trichophycin family, the H-4 oxymetine group in trichophycin I (**15**) was significantly deshielded (δ_H 4.80). The HMBC correlation from a deshielded methyl group (H_3-19, δ_H 2.02) to a carbonyl (C-18, δ_C 170.7) agreed with

the presence of an acetyl group, which accounted for the remaining degree of unsaturation. Due to the small amount isolated, the absolute configuration of trichophycin I (**15**) could not be determined.



15

Figure 11.19. Key 2D NMR correlations for trichophycin I (**15**)

However, the *anti*-configuration between H-4 and H-5, was supported by the coupling constant of 5.9 Hz, regarding these stereocenters.

Table 11.4 NMR data of trichophycin I (**15**) (500 MHz for ^{1}H NMR, $CDCl_3$)

pos	δC, type	δH (*J* in Hz)	HMBC	COSY
1a	117.3, CH_2	5.07, m	3	2, 3
1b		5.03, m	3	2, 3
2	134.1, CH	5.71, m		1a, 1b, 3
3	35.8, CH_2	2.27, m	1, 2, 4	2, 4
4	76.6, CH	4.80, m	2	3, 5
5	36.0, CH	1.67, m		4, 17
6a	29.2, CH_2	1.50, m		6b, 7
6b		1.17, m		5, 6a, 7
7	27.6, CH_2	2.14, m	6, 8, 16	6a, 6b
8	141.9, C			
9	41.2, CH_2	3.36, m	7, 8, 10, 11/15, 16	16
10	138.2, C			
11/15	128.9, CH	7.15, d (7.4)	9, 13	9, 12/14
12/14	128.5, CH	7.30, t (7.4)	10	11/15
13	126.6, CH	7.22, t (7.4)		
16	114.2, CH	5.84, s	7, 8, 9	9
17	15.2, CH_3	0.90, d (6.8)	4, 5, 6	5
18	170.7, C			
19	21.2, CH_3	2.02, s	18	

11.6 Biactivity evaluation

The possible cytotoxic activity of the newly isolated compounds was evaluated against neuro-2A cells and compared to the bioactivity of the

previously isolated trichophycins. Isotrichophycin C (**12**) and trichophycin G (**13**), showed potent cytotoxic activity. In detail, isotrichophycin C (**12**) exhibited more potent cytotoxicity (EC_{50}: 13.4 ± 0.35 μM) against neuro-2A cells than trichophycin C (EC_{50}: 23.8 ± 4.2 μM). Furthermore, trichophycin G (**13**) shows to be a more potent cytotoxic agent than both metabolites (EC_{50}: 8.4 ± 3.0 μM) with similar potency to trichophycin A (EC_{50}: 6.5 ± 1.4 μM). Moreover, results of the bioactivity assay revealed that trichophycin H (**14**), which contains an alkene group, was approximately 50% less potent than trichophycin A (EC_{50}: 12.1 ± 1.1 μM). With the data obtained from cytotoxicity assays for 12 of the 15 isolated trichophycin/trichotoxin metabolites, it was possible to reach some conclusions regarding structure-activity relationships (SARs) against neuro-2A cells. (Figure 11.20)

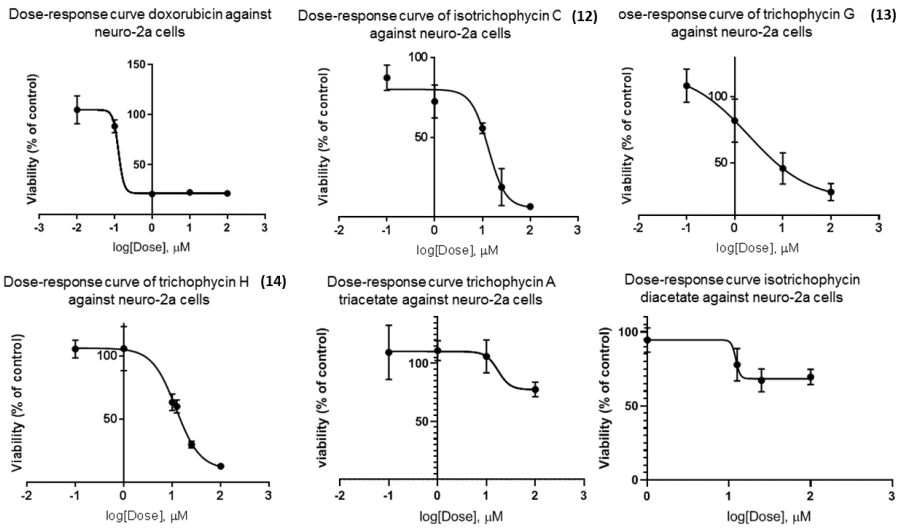
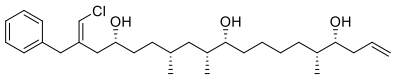
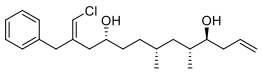
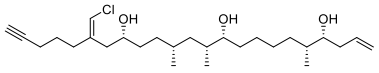
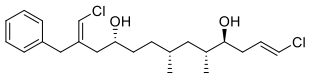
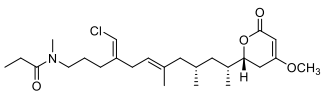
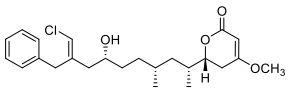
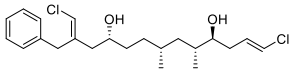
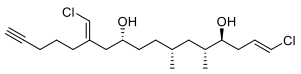
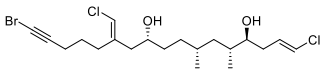
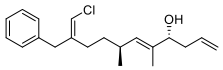
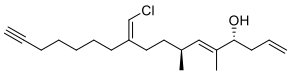
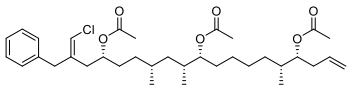
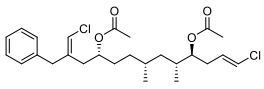
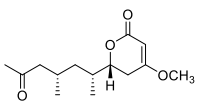


Figure 11.20. EC_{50} curves of doxorubicin, **12-14**, trichophycin A triacetate and isotrichophycin C diacetate.

In particular, it was verified that compounds containing longer polyketide chains are more active, as well as compounds containing phenyls are more

toxic than those containing alkenes. Interestingly, the toxicity of the compounds is negatively affected by the presence of halogens, in fact increased halogenation seems to decrease cytotoxicity. Furthermore, it was verified that the four most potent compounds all have *Z* configurations, in fact isotrichophycin C (**12**) is significantly more potent than trichophycin C. The ability of the compounds to donate hydrogen bonds, could contribute to the activity of the compounds, in fact, the acetylated versions of isotrichophycin C and trichophycin A, showed no activity. The acetylated version of trichophycin A, was previously obtained to explore the absolute configuration of its chiral centers. Inactivity could also be due to the added steric mass of acetyl groups.

Table 11.5 Cytotoxicity comparison of trichophycins, trichotoxins, and tricholactone against neuro-2a cells.

Compound Name	Structure	EC_{50} [mean ± SE (μM)]
Trichophycin A		6.5 ± 1.4
Trichophycin G (**13**)		8.4 ± 3.0
Trichophycin H (**14**)		12.1 ± 1.1
Isotrichophycin C (**12**)		13.4 ± 0.4
Trichophycin F		14.3 ± 2.3
Trichophycin B		14.8 ± 2.4
Trichophycin C		23.8 ± 4.2[8]
Trichophycin D		39.8 ± 3.8[8]
Trichophycin E		>100[8]
Trichotoxin A		>50[10]
Trichotoxin B		>50[10]
Trichophycin A tri Ac		>100
Isotrichophycin C di Ac		>100
Tricholactone		>100

11.7 Experimental section

11.7.1 General methods

A Jasco P-2000 polarimeter was used to perform optical rotations while UV spectra were recorded using a Beckman Coulter DU-800 spectrophotometer. ECD spectra were recorded using a Jasco J-1100 CD spectrometer, and IR spectra were recorded using a Thermo Scientific Nicolet 380 FT-IR spectrometer. A Bruker 800 MHz NMR instrument equipped with cryoprobe, and a Varian 500 MHz instrument were used to perform NMR spectra. The chemical shifts reported were referenced to the residual solvent peaks of $CDCl_3$ (δ_H 7.26 and δ_C 77.2). An AB SCIEX Triple TOF 4600 mass spectrometer was used to perform HRESIMS analysis and the mass spectra were analyzed by Analyst TF software. LC-HRESIMS data for molecular networking were recorded using a Thermo LTQ Orbitrap XL high-resolution ESI mass spectrometer coupled to Thermo U3000 HPLC system, equipped with a solvent reservoir, a degasser, binary pump and refrigerated autosampler. A Thermo Fisher Scientific ISQ mass spectrometer with an electrospray ionization (ESI) source was used to perform low resolution LC-MS analysis. A Dionex UltiMate 3000 HPLC system equipped with a micro vacuum degasser, an autosampler, and a diode-array detector, were used to perform semi-preparative HPLC separations.

11.7.2 Collection of biological material

Biomass from the *Trichodesmium sp*. bloom was collected at the water surface in the Gulf of Mexico.[15] The individual filaments of the bloom were

identified as *Trichodesmium* following the examination of morphological characteristics using light microscopy. The biomass was then frozen for subsequent chemical analysis.

11.7.3 DNA Extraction, Amplification, Sequencing, and Phylogenetic Analysis.

The Qiagen DNeasy® Plant Mini Kit, following the manufacturer's specifications, was used to extract DNA from filaments stored in RNA*later*®. A Mini-Beadbeater-96 (BioSpec Products Inc.) was used for two minutes with 0.5 mm and 0.1 mm beads to facilitate cell lysis. A NanoDrop 2000c (Thermo Scientific) was used to measure the integrity and purity of the extracted DNA. Lineage- specific primers were designed using the IDT primer quest tool to amplify the SSU rRNA gene. The primer set used was as follows: forward 5'- GTA GCG GTG AAA TGC GTA GA -3' and reverse 5'- CTC CCT TTC GGG TTA GAG TAA TG -3'. The PCR reaction components consisted of 3 µl of DNA (36 ng), 7 µl of PCR grade water (ThermoFisher), 1.25 µl of forward and reverse primers (10 µM), and 12.5 µl of 2X Platinum SuperFi PCR Master Mix (ThermoFisher) for a total of 25 µl. The PCR reaction was performed using an Eppendorf Mastercycler X50a®. First, denaturation occurred for 30 seconds at 98 ºC followed by 33 cycles of 10 seconds at 98 ºC, 30 seconds at 47 ºC, 30 seconds at 72 ºC, and a final elongation for five minutes at 72 ºC. A QIAquick PCR purification kit was used to perform the following purification of the PCR product following the manufacturer's specifications. Sanger sequencing of the PCR product was performed with the same primers used for amplification. Partial 16S rRNA sequences from the relevant cyanobacterial strains were collected

from Genbank, aligned with ClustalW, and trimmed. After generation of a phylogenetic tree using the maximum likelihood method and the Tamura-Nei model in MEGA X.22,[16] the final tree used for phylogenetic inference was made from 1000 bootstrap replicates. The partial 16S sequence has been deposited in GenBank under the accession number MT478931.

11.7.4 Data-dependent LC-HRMS/MS analysis

LC-HRMS/MS analysis sample of the crude CH_2Cl_2:CH_3OH extract of *Trichodesmium sp*., was prepared dissolving the sample in MeOH at a concentration of (10 mg/mL). LC-HRMS/MS analysis were carried out with a 5 μm Kinetex C18 column (50 × 2.10 mm), a flow rate of 200 $\mu L \cdot min^{-1}$, using 0.1% HCOOH in H_2O (eluent A) and CH_3CN (eluent B) as eluents. The gradient program was set as follows: 45% B for 1 min, 30%->80% B over 30 min, 80%->100% B over 1 min, 100% B for 9 min. A data-dependent acquisition mode of the Orbitrap spectrometer software was used to select halogenated compounds from fragmentation, avoiding uninteresting fragmentation of non-halogenated compounds. The halogenation patterns Cl, Cl_2, BrCl, and $BrCl_2$ were considered of interest and a separate run was performed for each pattern. Fragmentation was triggered by an M+2 isotope peak with intensity 34±10% to select Cl ions, an M+2 isotope peak with intensity 60±10% to select Cl_2 ions, an M+4 isotope peak with intensity 32±10% to select BrCl ions, and an M+4 isotope with intensity 72±10% to select $BrCl_2$ ions. For each full scan, the five most intense ions, meeting the specifications, were fragmented. Based on the incapacity of data-dependent acquisition of the spectrometer software to be

triggered by an accurate mass difference, some non-halogenated ions were subjected to fragmentation, showing the required M+2 or M+4 isotope peak ratios due to other co-eluting compounds. These compounds were filtered out by the subsequent data processing.

11.7.5 MZmine processing and molecular networking

Mzmine 2.51 software was used to perform preprocessing of LC-HRMS/MS data. LC-MS .raw files generated by the Orbitrap mass spectrometer were uploaded without conversion to MZmine2. After standard preliminary data treatment, data crop between 18-37 min and *m/z* 70-700, mass detection, ADAP chromatogram construction, smoothing, and deconvolution using the local minimum search algorithm, the Adduct Search module was used to identify, Na^+, NH_4^+, and K^+ adducts and ^{13}C isotopic peaks (mass_difference = 1.0033) and annotate them for subsequent filtering. The different .raw files from the four runs selecting the different halogenation patterns were aligned using the Join Aligner module. Peaks annotated as adducts were filtered out using the Feature List Rows Filter module. Known compounds were identified using the Custom Database Search module, and a list of mass and retention times in our standard chromatographic conditions. Automatic isotopic pattern identification is not implemented by MZmine, but could be achieved with the Adduct Search module repeatedly with the following parameters: (in this order): for $BrCl_2$, mass_difference = -3.995, Max relative height of adduct peak < 160%; for BrCl, mass_difference = -3.995, Max relative height of adduct peak < 400%; for Cl_2, mass_difference = -1. 997, Max relative adduct peak height <

180%; for Cl, mass_difference = -1.997, Max relative adduct peak height < 400% (because negative mass differences are used, peak M is seen by the software as an "adduct" of peaks M+2 pr M+4). Consequently, M peaks that showed M+2 or M+4 isotopic peaks with the required mass and intensity were annotated as $BrCl_2$, BrCl, Cl_2, or Cl. In addition, the Feature List Rows Filter module was used to remove unannotated peaks, peaks without associated MS^2 data, and very minor compounds (peak area < 2.0×10^7). Finally, the MS data were exported to an MGF file. Finally, the MS data were exported to an MGF file, while quantitative data, along with the annotated isotope pattern and identification, were exported to a CSV file.

The GNPS platform was used to generate molecular network. Due to the impossibility of the FBMN workflow on GNPS to propagate the isotope our isotope pattern annotations, the Metabolomics workflow on GNPS was used.[17] A parent mass tolerance 0.02 Da, MS/MS fragment ion tolerance 0.02 Da, the cosine score > 0.6, matched peaks > 6, and maximum neighbor number (topK) = 12 parameters were set. Moreover, the network visualization and analysis were performed using Cytoscape 3.7.2, mapping quantitative data and annotations in the CSV file to the relevant nodes of the network (https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=416bebdcdeb1468cb39610db0aab871f).[18]

11.7.6 Isolation of isotrichophycin C (**12**), trichophycin G-I (**13**-**15**)

After thawing of the frozen biomass, it was repeatedly extracted with 2:1 CH_2Cl_2:CH_3OH. The crude extract obtained was fractionated by vacuum

liquid chromatography obtaining 9 VLC fractions (A-I).[4] Fraction C eluted with 80% hexanes in EtOAc (144.2 mg), and Fraction D, eluted with 60% hexanes in EtOAc (293.4 mg), were chromatographed using a 2 g C18 SPE column, eluting with 100% CH_3OH to generate HPLC pre-fractions C-1 and D-1 (53.3 mg), respectively. Whereas fraction I, (100% CH_3OH, 2,090.0 mg) was separated over a 10 g C18 SPE by eluting with 50% CH_3CN in H_2O, 100% CH_3CN, 100% CH_3OH and 100% EtOAC. Fraction I-2, eluted with 100% CH_3CN (273.3 mg), was subsequently purified. Fraction D-1, which from HPLC-DAD analytical analysis appeared not to contain many metabolites, was subjected to semi-preparative RP-HPLC using a YMC 5 µm ODS column (250 x 10 mm) with a mobile phase of 75% CH_3CN in H_2O with 0.05% formic acid added, and a flow rate of 3 mL/min. From this fraction, it was possible to isolate isotrichophycin C (**12**) (5.0 mg, tR: 13.65 min). Purification of the C-1 fraction was also performed by RP-HPLC using the same YMC column and flow rate with a mobile phase of 65% CH_3CN in water with 0.05% formic acid added, resulting in the isolation of Trichophycin I (**15**) (2 mg, t_R: 20.25 min). The evident complexity of Fraction I-2 in metabolite composition, prompted to a different purification protocol. It was subjected to RP-HPLC using the YMC column, a 3 mL/min flow rate, and a gradient HPLC method and time-based collection (5 minutes increment). The mobile phase consisted of H_2O and CH_3CN and the gradient was 50% B for 5 min, 50% → 100% B over 25 min, 100% B over 5 min, and a return to initial conditions for 10 min, Time-based collection of peaks from 20 to 25 min was designated as Fraction E, solvent was concentrated under vacuum leaving a clear oil residue (57.9 mg), which was

further processed using a Kinetex 5 μm C18 column (250 x 10 mm). The analysis was carried out using mobile phase: 75% CH_3CN in water with 0.05% formic acid added to each solvent, flow rate 3 mL/min. Two fractions of 2.2 mg and 4.7 mg were collected (t_R, 10-11 min for fraction E-3, and 11-12 min for fraction E-4), for further purification. Final purification of fraction E-3 was achieved using the same Kinetex column and flow rate with a mobile phase of 65% CH_3CN in H_2O with 0.05% formic acid added to each solvent, leading to the isolation of trichophycin G (**13**) (1.4 mg, tR: 23.75 min). Fraction E-4 was also subjected to a final purification using the Kinetex 5 μm C18 column (250 x 10 mm), a flow rate of 3 mL/min and a mobile phase of 75% CH_3CN in water with 0.05% formic acid added to each solvent, and tricophycin H (**14**) was isolated (1.1 mg, tR: 9.30 min).

Isotrichophycin C (**12**): colorless oil; $[\alpha]^{23}_D$ -9.5 (*c* 0.20, MeOH); UV (MeOH) λ_{max} (log ε) 208 (4.2) nm; ECD (c 1.0 mM, CH_3CN) λmax (Δε) 228 (0.68) nm; IR (ZnSe) νmax 3400 (br), 2917, 2849, 1743, 1733, 1183 cm^{-1};^{1}H NMR (800 MHz, $CDCl_3$), and ^{13}C NMR (200 MHz, $CDCl_3$), **see Table 11.1**; HRESIMS *m/z* 399.1855 $[M+H]^+$ (calcd for $C_{22}H_{33}Cl_2O_2$, 399.1858).

Trichophycin G (**13**): colorless oil; $[\alpha]^{23}_D$ -20.8 (*c* 0.06, MeOH); UV (MeOH) λ_{max} (log ε) 203 (3.6); ^{1}H NMR (500 MHz, $CDCl_3$) see Table 11.2; HRESIMS *m/z* 387.2071 $[M+Na]^+$ (calcd for $C_{22}H_{33}ClO_2Na$, 387.2067).

Trichophycin H (**14**): colorless oil; $[\alpha]^{23}_{D}$ +22.9 (*c* 0.07, MeOH); UV (MeOH) λ_{max} (log ε) 201 (3.6) nm; IR (ZnSe) νmax 3420 (br), 2920, 2313, 1646 cm−1; ^{1}H NMR (500 MHz, $CDCl_3$), and ^{13}C NMR (125 MHz, $CDCl_3$), see Table 11.3; HRESIMS *m/z* 477.3107 $[M+Na]^+$ (calcd for $C_{27}H_{47}ClO_3Na$, 477.3111).

Trichophycin I (**15**) colorless oil; α^{22}_{D} -6.2; UV (MeOH) λ_{max} (log ε) 202 (3.8) nm; ^{1}H NMR (500 MHz, $CDCl_3$) and ^{13}C NMR (125 MHz, $CDCl_3$), see Table 11.4; HRESIMS *m/z* 343.1454 $[M+Na]^+$ (calcd for $C_{19}H_{25}ClO_2$, 343.1441).

11.7.7 Preparation of isotrichophycin C (**12**) peracetylation products

The acetylation reaction was performed by adding a 1:1 mixture of pyridine and acetic anhydride to 1 mg of isotricophycin C (**12**) stirring the samples for 24 hours. After drying of the mixtures under a flow of N_2, the residues were subjected to RP-HPLC using a Kinetex 5 µm C18 column (250 x 10 mm), a flow rate of 3 mL/min, and a mobile phase of 95% CH_3CN in water with 0. 05% formic acid added to each solvent, isolating Isotrichophycin C diacetate (1.0 mg, tR: 7.25 min) .

Isotrichophycin C diacetate: ^{1}HNMR (500 MHz, $CDCl_3$) δ 7.31 (2H, t, J = 7.4 Hz, H-16, H-18), 7.23 (1H, m, H-17), 7.17 (2H, d, J = 7.4 Hz, H-15, H-19), 6.01 (1H, d, J = 13.4 Hz, H-1), 5.93 (1H, s, H-20), 5.84 (1H, m, H-2), 5.09 (1H, m, H-10), 4.78 (1H, dt, J = 7.6, 5.1 Hz, H-4), 3.43 (2H, d, J = 3.7 Hz, H-13), 2.54 (1H, dd, J = 13.7, 8.5 Hz, H-11a), 2.28 (1H, m, H-11b), 2.27 (2H, m, H-3), 2.05 (6H, s, 2-Ac groups), 1.79 (1H, m, H-5), 1.61 (2H,

m, H-9), 1.45 (1H, m, H-7), 1.31 (1H, m, H-8a), 1.25 (1H, m, H-6a), 1.05 (1H, m, H-8b), 0.97 (1H, m, H-6b), 0.89 (3H, d, *J* = 6.7 Hz, H-21), 0.87 (3H, d, *J* = 6.8 Hz, H-22); HRESIMS *m/z* 505.1889 $[M+Na]^+$ (calcd for $C_{26}H_{36}Cl_2O_4$, 505.1888).

11.7.8 Preparation and analysis of MTPA esters

The methodology described by Hoye et al., (2007)[19] with minor modification was performed. 1.5 mg of isotrichophycin C (**12**) was dissolved in dry $CDCl_3$ and separated into two equal aliquots in 4 mL vials. In the first vial Dry pyridine (10 µL) and (*S*)-(+)-α-methoxy-α-(trifluoromethyl) phenylacetyl chloride (15 µL) were added, whereas an equal amount of pyridine and (*R*)-(–)-a-methoxy-a-(trifluoromethyl)phenylacetyl chloride, were added in the second vial. The vials were capped, and the reaction mixtures were stirred for 24h. After 24h, the reaction mixtures were first by ^{1}H NMR to determine the progress of the esterification by examining the presence of deshielded oxymethine signals at 5.29 and 4.99 ppm. Following NMR analysis, each sample was separated between DCM and H_2O of each sample, the DCM layers were dried under a stream of N_2. The residues were subjected to RP-HPLC using a Kinetex 5 µm C18 column (250 x 10 mm) with a mobile phase of 100% ACN and a flow rate of 3 mL/min and both bis-MTPA esters were isolated (t_R, 7.80 min). The same procedure was performed using 1.5 mg of trichophycin A. The tris-MTPA esters were isolated using the same HPLC method (t_R,12.0).

Isotrichophycin C bis-MTPA: ***S*-ester** (partial) ^{1}HNMR (500 MHz, $CDCl_3$) δ 7.28 (2H, t, J = 7.4 Hz, H-16, H-18), 7.23 (1H, m, H-17), 7.11 (2H, d, J = 7.4 Hz, H-15, H-19), 5.87 (1H, s, H-20), 5.91 (1H, s, H-1), 5.71 (1H, m, H-2), 5.29 (1H, m, H-10), 5.02 (1H, m, H-4), 3.26 (1H, d, m, H-13), 2.40 (2H, m, H-11), 2.27 (2H, m, H-3), 1.93 (1H, m, H-5), 1.64 (1H, m, H-9), 1.55 (2H, m, H-8), 1.47 (1H, m, H-7), 1.25 (1H, m, H-6), 0.90 (3H, d, J = 6.6 Hz, H-21), 0.88 (3H, d, J = 6.6 Hz, H-22); HRESIMS *m/z* 853.2498 $[M+Na]^+$ (calcd for $C_{42}H_{46}Cl_2F_6O_6Na$, 853.2473); ***R*-ester** (partial) ^{1}HNMR (500 MHz, $CDCl_3$) δ 7.28 (2H, d, J = 7.4 Hz, H-16, H-18), 7.23 (1H, m, H-17), 7.13 (2H, d, J = 7.4 Hz, H-15, H-19), 5.96 (1H, s, H-20), 5.99 (1H, s, H-1), 5.81 (1H, m, H-2), 5.29 (1H, m, H-10), 4.99 (1H, m, H-4), 3.36 (1H, m, H-13), 2.54 (2H, m, H-11), 2.28 (2H, m, H-3), 1.83 (1H, m, H-5), 1.56 (1H, m, H-9), 1.47 (1H, m, H-8), 1.38 (1H, m, H-7), 1.13 (1H, m, H-6), 0.79 (3H, d, J = 6.6 Hz, H-21), 0.75 (3H, d, J = 6.8 Hz, H-22); HRESIMS *m/z* 853.2464 $[M+Na]^+$ (calcd for $C_{42}H_{46}Cl_2F_6O_6Na$, 853.2473).

Trichophycin A tris-MTPA: ***S*-ester** (partial) ^{1}HNMR (500 MHz, $CDCl_3$) δ 7.29 (2H, d, J = 7.4 Hz, H-22, H-24), 7.23 (1H, m, H-23), 7.10 (2H, d, J = 7.4 Hz, H-21, H-25) 5.85 (1H, s, H-26), 5.73 (1H, m, H-2), 5.28 (1H, m, H-16), 5.10 (1H, m, H-4), 5.11 (1H, m, H-1a), 5.08 (1H, m, H-1b) 5.00 (1H, m, H-10), 3.24 (2H, d, J = 3.9 Hz), 2.41 (2H, m, H-3), 2.34 (2H, m, H-17), 1.72 (1H, m, H-11), 1.67 (2H, ovlp, H-15), 1.67 (1H, m, H-5), 1.63 (2H, m, H-9), 1.52 (2H, ovlp, H-14), 1.51 (2H, ovlp, H-8), 1.40 (1H, m, H-13), 1.23 (2H, m, H-6), 1.19 (2H, m, H-7), 1.14 (2H, m, H-12), 0.85 (3H, d, J = 6.8, H-29), 0.81 (3H, d, J = 6.8 Hz, H-28), 0.78 (3H, d, J = 6.6 Hz, H-27); HRESIMS *m/z* 1149.4317 $[M+Na]^+$ (calcd for $C_{59}H_{68}ClF_9O_9Na$,

1149.4306); ***R*-ester** (partial) ^{1}HNMR (500 MHz, $CDCl_3$) δ 7.28, (2H, d, J = 7.4 Hz, H-22, H-24), 7.23 (1H, m, H-23), 7.13 (2H, d, J = 7.4 Hz, H-21, H-25), 5.95 (1H, s, H-26), 5.65 (1H, m, H-2), 5.29 (1H, m, H-16), 5.09 (1H, m, H-4), 5.03 (1H, m, H-1a), 5.00 (1H, m, H-1b), 5.00 (1H, m, H-10), 3.37 (2H, s, H-19), 2.53 (2H, dd, J = 14.0, 7.9, H-17), 2.35 (2H, m, H-3), 1.71 (1H, m, H-11), 1.70 (1H, m, H-5), 1.59 (2H, m, H-9), 1.53 (2H, m, H-15), 1.47 (2H, m, H-14), 1.45 (2H, m, H-8), 1.39 (1H, m, H-13), 1.27 (2H, m, H-6), 1.16 (2H, ovlp, H-12), 1.15 (2H, ovlp, H-7), 0.89 (3H, d, J = 6.7 Hz, H-29), 0.83 (3H, d, J = 6.7 Hz, H-28), 0.74 (3H, d, J = 6.7 Hz, H-27); HRESIMS *m/z* 1149.4316 $[M+Na]^+$ (calcd for $C_{59}H_{68}ClF_9O_9Na$, 1149.4306).

11.7.9 Biological assays

Murine neuroblastoma cells (neuro-2A) were used to perform cytotoxicity assays. For each concentration of compounds tested **12**-**14**, [4] four technical replicates were prepared. To determine EC_{50} values, assays were performed in triplicates, using doxorubicin as positive control (EC_{50}: 126.8 μM). Four technical replicates were prepared for each concentration of isotrichophycin C diacetate and trichophycin A triacetate and viability values were determined. GraphPad Prism software was used to generated EC_{50} curves and to perform statistical analysis.

11.8 Spectroscopic data

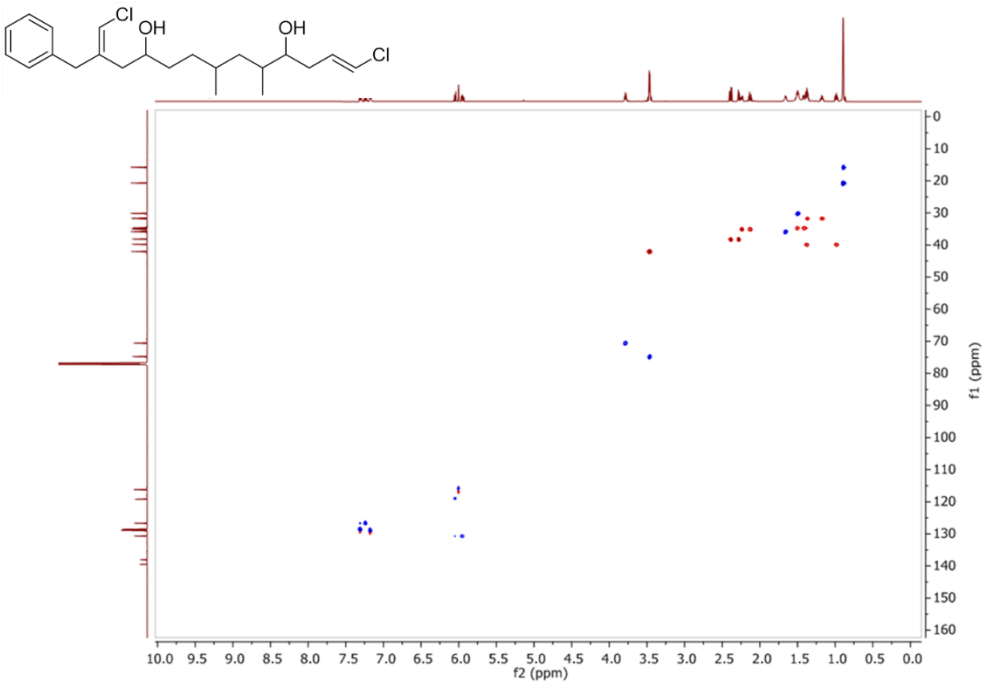


Figure 11.21. HSQC spectrum of isotrichophycin C (**12**)

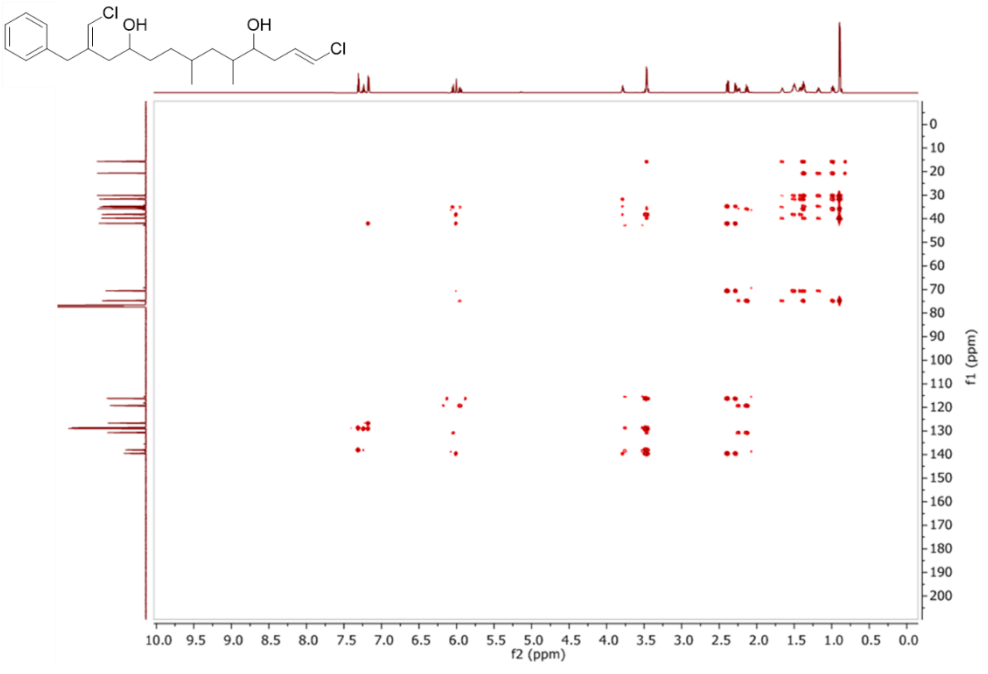


Figure 11.22. HMBC spectrum of isotrichophycin C (**12**)

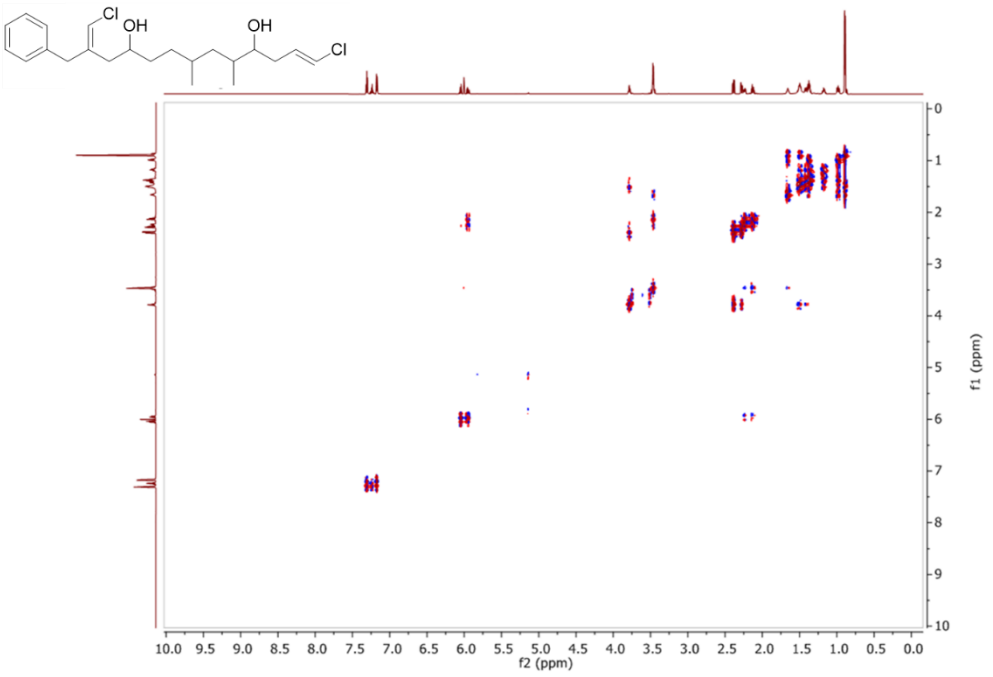


Figure 11.23. DQF-COSY spectrum of isotrichophycin C (**12**)

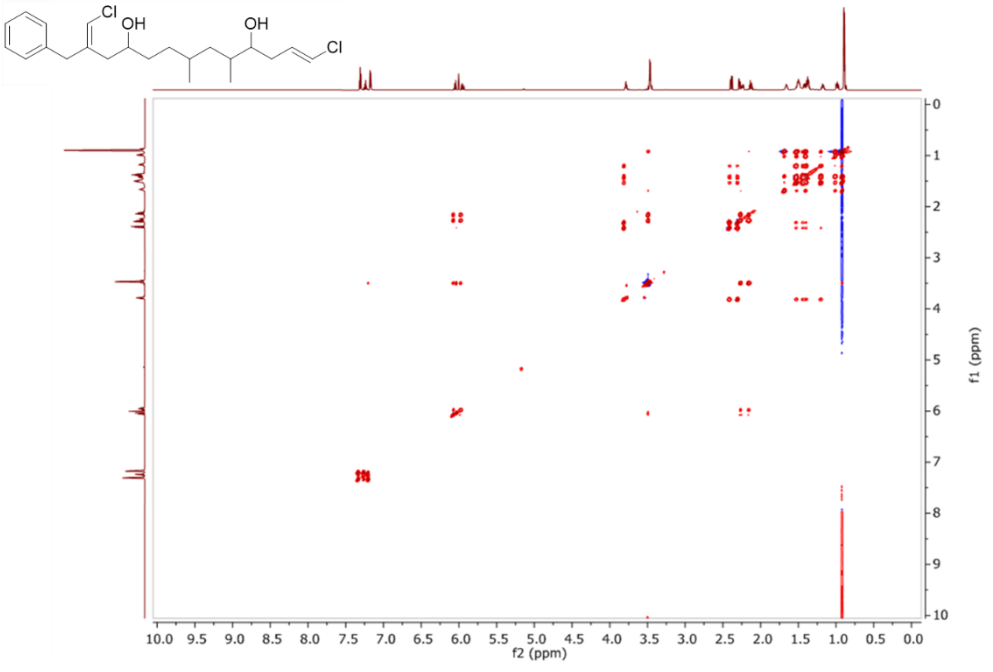


Figure 11.24. TOCSY spectrum of isotrichophycin C (**12**)

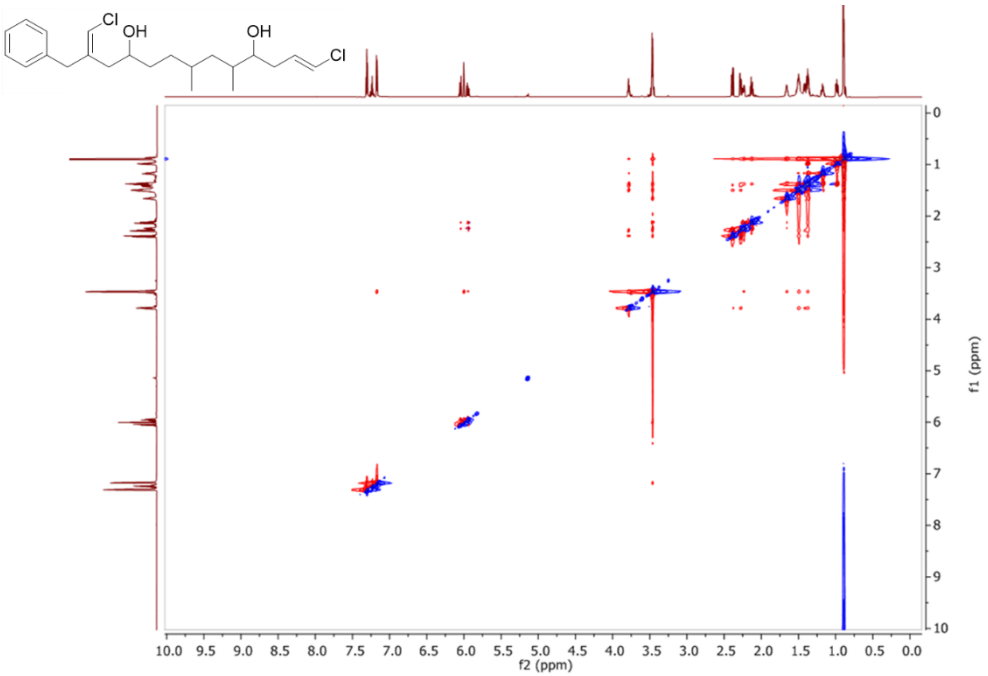


Figure 11.25. NOESY spectrum of isotrichophycin C (**12**)

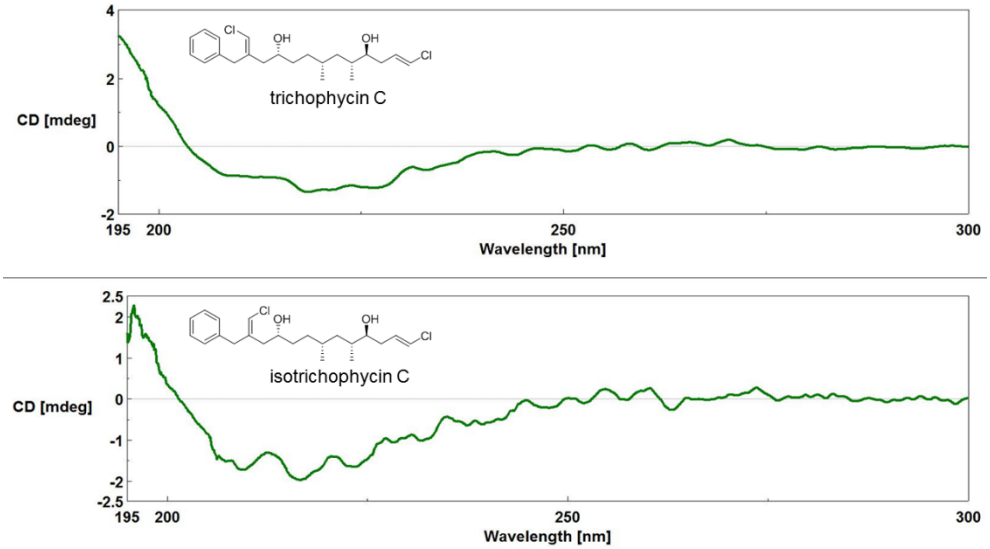


Figure 11.26. Comparison of trichophycin C and isotrichophycin C (**12**) ECD spectra (0.4 mg/mL, CH_3CN).

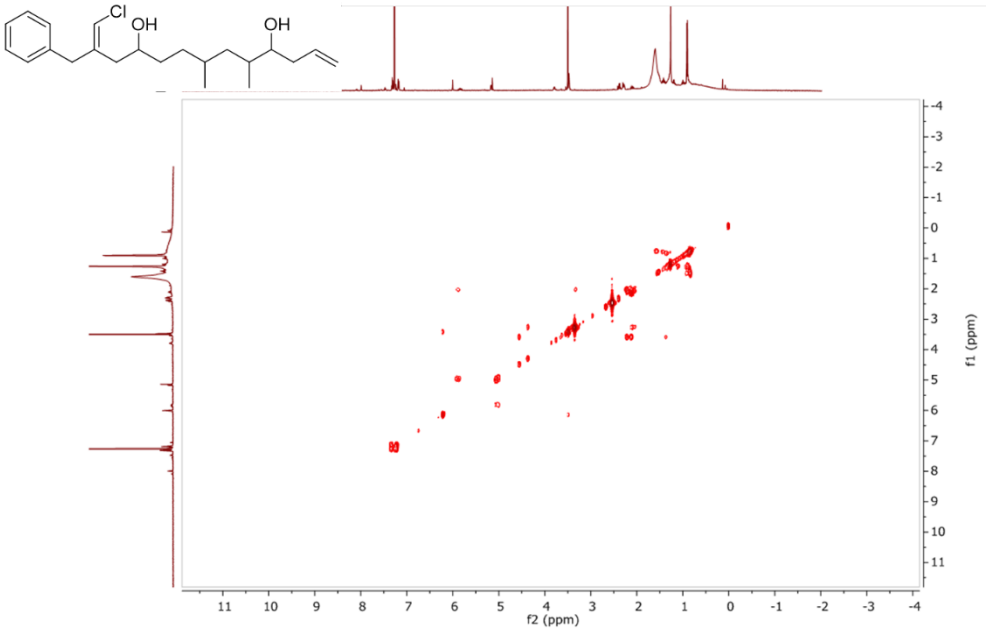


Figure 11.27. COSY spectrum of trichophycin G (**13**)

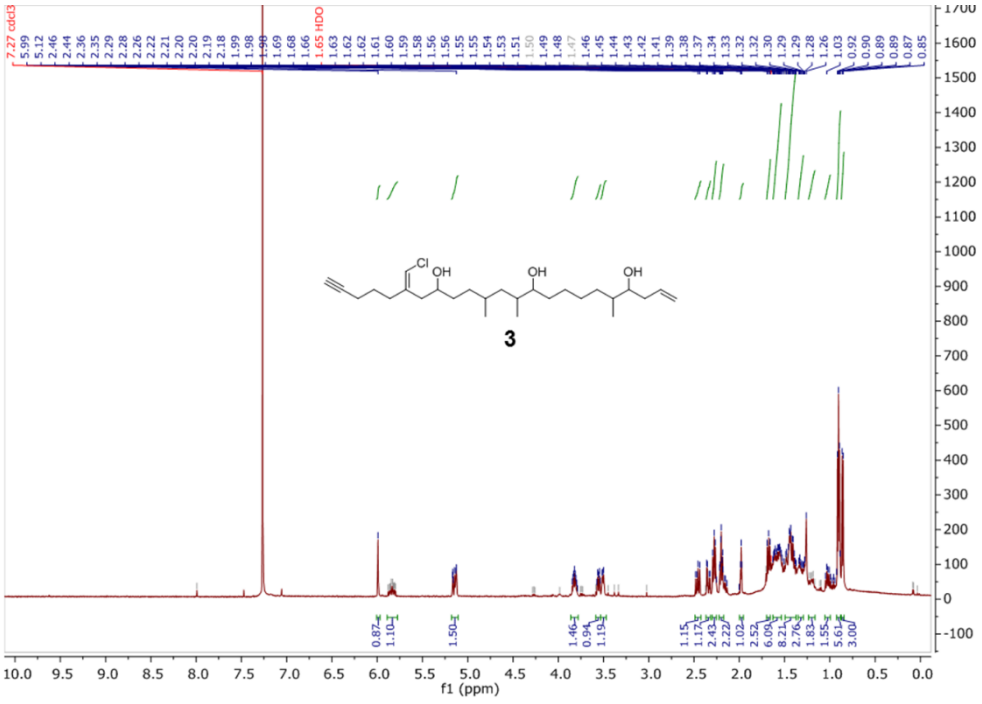


Figure 11.28. ^{1}H NMR spectra of trichophycin H (**14**) (500 MHz, $CDCl_3$).

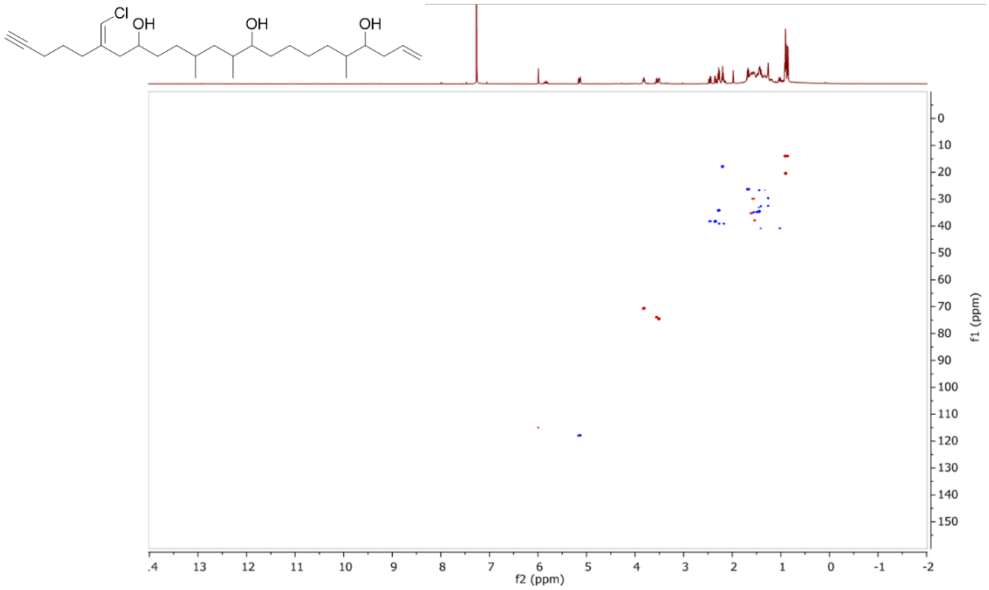


Figure 11.29 HSQC spectrum of trichophycin H (**14**)

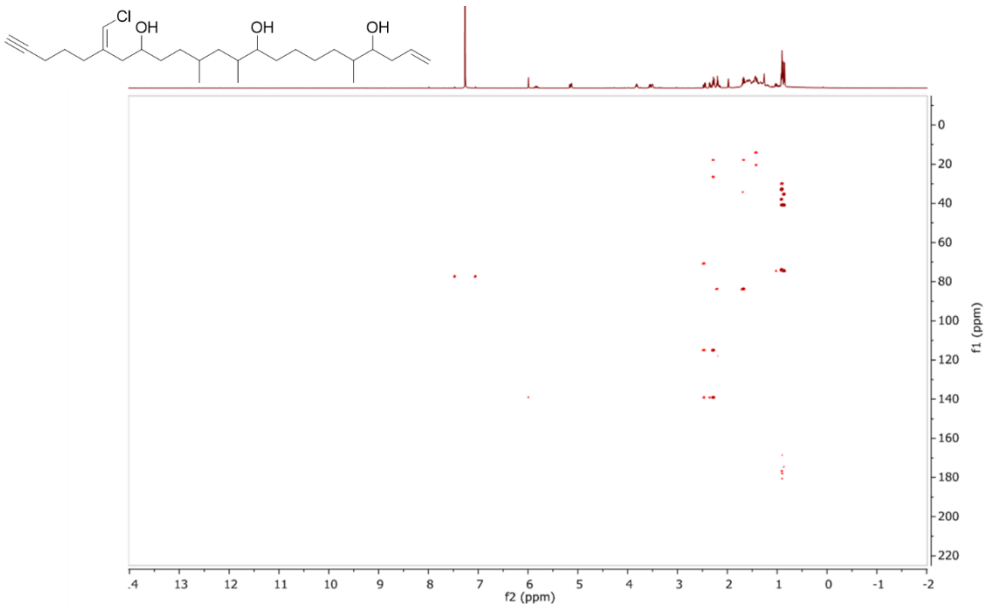


Figure 11.30. HMBC spectrum of trichophycin H (**14**)

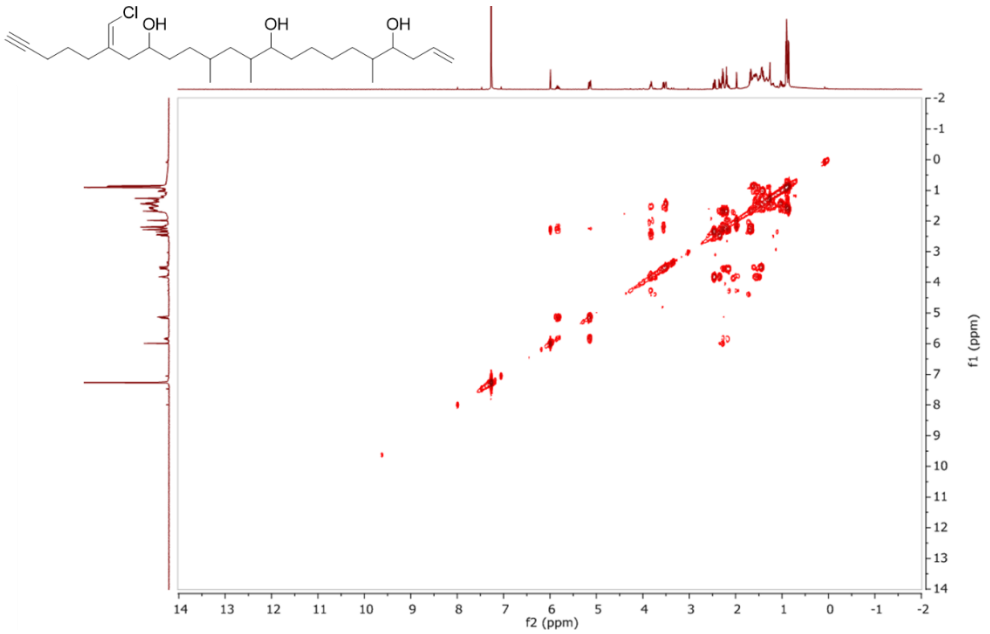


Figure 11.31. COSY spectrum of trichophycin H (**14**)

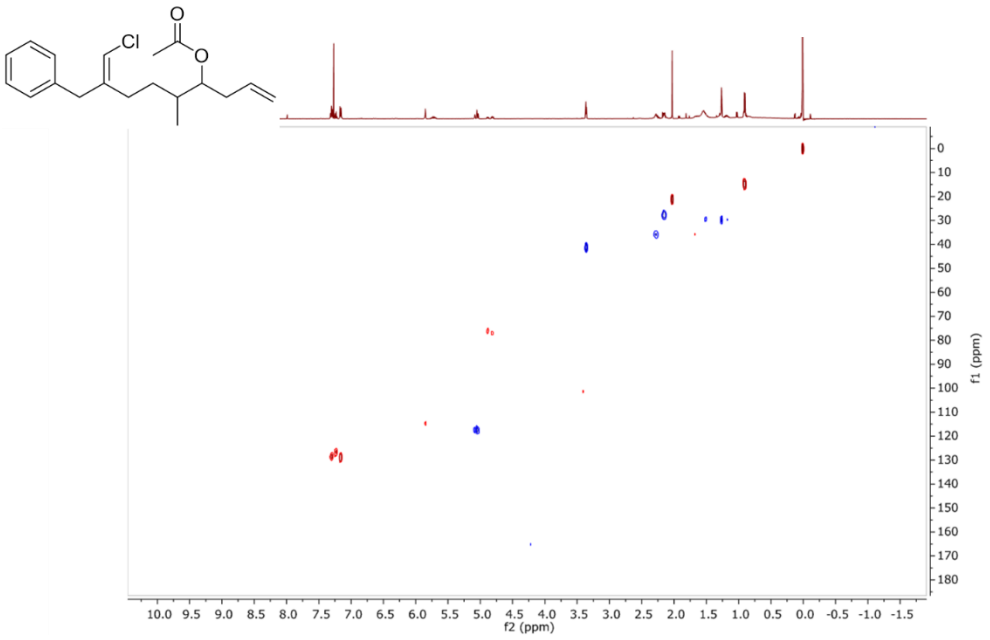


Figure 11.32. HSQC spectrum of trichophycin I (**15**)

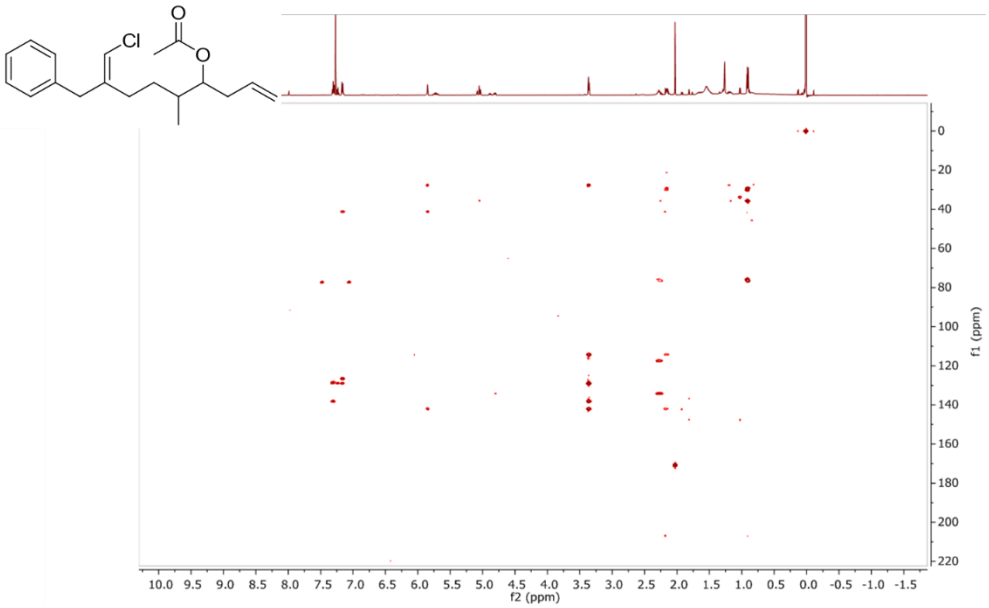


Figure 11.33. HMBC spectrum of trichophycin I (**15**)

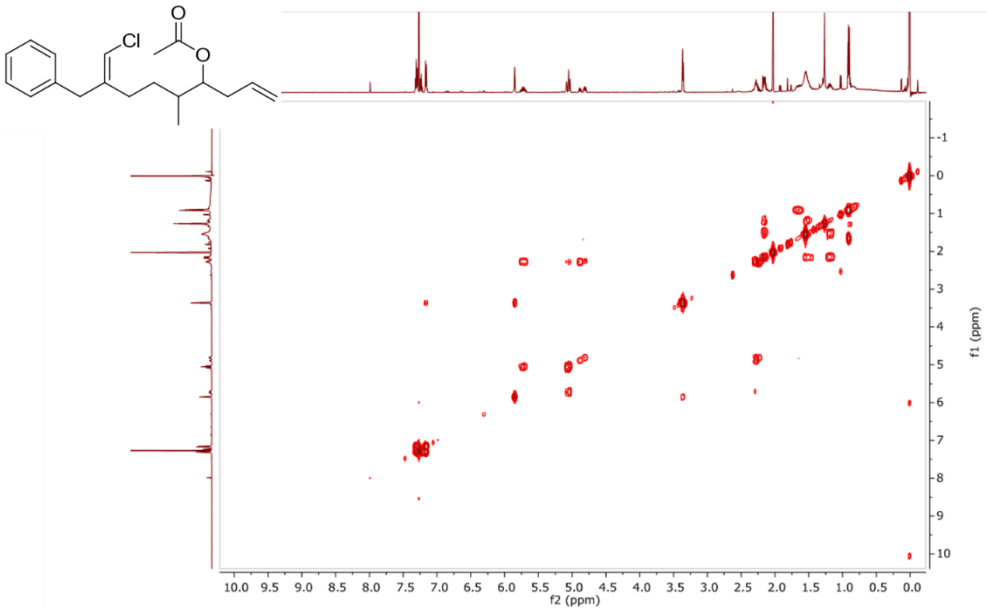


Figure 11.34. COSY spectrum of trichophycin I (**15**)

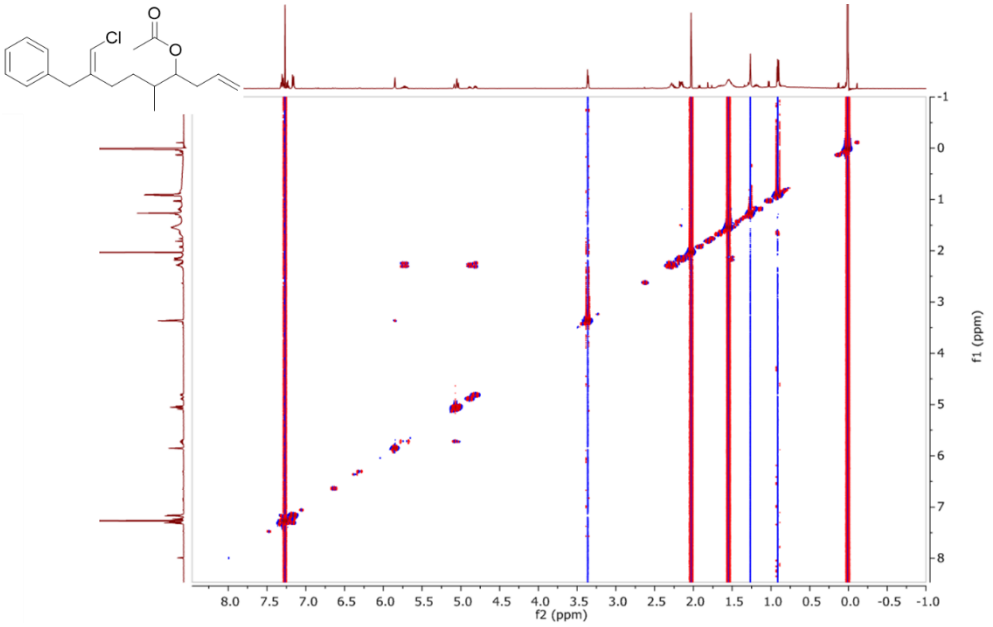


Figure 11.35. NOESY spectrum of trichophycin I (**15**)

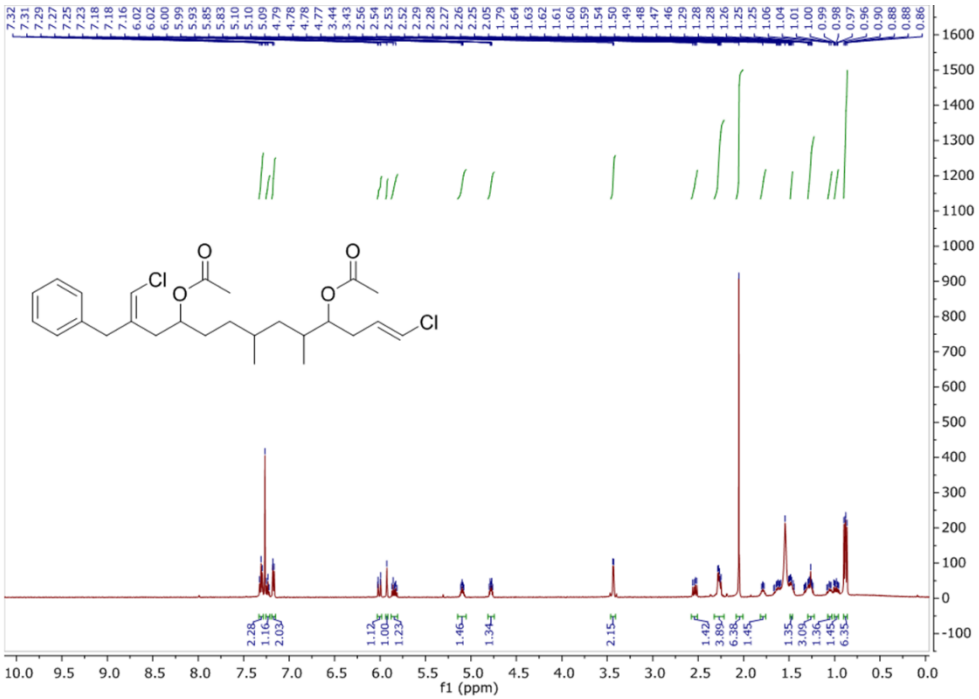


Figure 11.36. ^{1}H NMR of isotrichophycin C diacetate (500 MHz, $CDCl_3$).

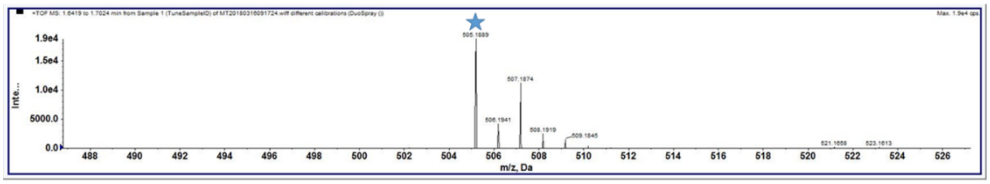


Figure 11.37. HRESIMS of isotrichophycin C diacetate

References

[1] McManus, K.M.; Kirk, R.D.; Via, C.W.; Lotti, J.S.; Roduit, A.F.; Teta, R.; Scarpato, S.; Mangoni, A.; Bertin, M.J. Isolation of isotrichophycin C and trichophycins G−I from a collection of trichodesmium thiebautii. *J. Nat. Prod.* **2020**, *83*, 2664–2671, doi:10.1021/acs.jnatprod.0c00550.

[2] Schock, T.B.; Huncik, K.; Beauchesne, K.R.; Villareal, T.A.; Moeller, P.D.R. Identification of trichotoxin, a novel chlorinated compound associated with the bloom forming cyanobacterium, trichodesmium thiebautii. *Environ. Sci. Technol.* **2011**, *45*, 7503–7509, doi:10.1021/es201034r.

[3] Malloy, K.L.; Suyama, T.L.; Engene, N.; Debonsi, H.; Cao, Z.; Matainaho, T.; Spadafora, C.; Murray, T.F.; Gerwick, W.H. Credneramides A and B: Neuromodulatory phenethylamine and isopentylamine derivatives of a vinyl chloride-containing fatty acid from cf. Trichodesmium sp. nov. *J. Nat. Prod.* **2012**, *75*, 60–66, doi:10.1021/np200611f.

[4] Bertin, M.J.; Wahome, P.G.; Zimba, P. V.; He, H.; Moeller, P.D.R. Trichophycin A, a cytotoxic linear polyketide isolated from a Trichodesmium thiebautii bloom. *Mar. Drugs* **2017**, *15*, 1–8, doi:10.3390/md15010010.

[5] Bertin, M.J.; Saurí, J.; Liu, Y.; Via, C.W.; Roduit, A.F.; Williamson, R.T. Trichophycins B-F, Chlorovinylidene-Containing Polyketides Isolated from a Cyanobacterial Bloom. *J. Org. Chem.* **2018**, *83*, 13256–13266, doi:10.1021/acs.joc.8b02070.

[6] Teta, R.; Della Sala, G.; Esposito, G.; Via, C.W.; Mazzoccoli, C.; Piccoli, C.; Bertin, M.J.; Costantino, V.; Mangoni, A. A joint molecular networking study of a: Smenospongia sponge and a cyanobacterial bloom revealed new antiproliferative chlorinated polyketides. *Org. Chem. Front.* **2019**, *6*, 1762–1774, doi:10.1039/c9qo00074g.

[7] Sheridan, C.C.; Steinberg, D.K.; Kling, G.W. The microbial and metazoan community associated with colonies of Trichodesmium spp.: A quantitative survey. *J. Plankton Res.* **2002**, *24*, 913–922, doi:10.1093/plankt/24.9.913.

[8] Frischkorn, K.R.; Rouco, M.; Van Mooy, B.A.S.; Dyhrman, S.T. Epibionts dominate metabolic functional potential of Trichodesmium colonies from the oligotrophic ocean. *ISME J.* **2017**, *11*, 2090–2101, doi:10.1038/ismej.2017.74.

[9] Rouco, M.; Haley, S.T.; Dyhrman, S.T. Microbial diversity within the Trichodesmium holobiont. *Environ. Microbiol.* **2016**, *18*, 5151–5160, doi:10.1111/1462-2920.13513.

[10] Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010**, *11*, doi:10.1186/1471-2105-11-395.

[11]Esposito, G.; Teta, R.; Miceli, R.; Ceccarelli, L.S.; Sala, G. Della; Camerlingo, R.; Irollo, E.; Mangoni, A.; Pirozzi, G.; Costantino, V. Isolation and assessment of the in vitro anti-tumor activity of smenothiazole a and B, chlorinated thiazole-containing peptide/polyketides from the Caribbean sponge, Smenospongia aurea. *Mar. Drugs* **2015**, *13*, 444–459, doi:10.3390/md13010444.

[12] Caso, A.; Esposito, G.; Sala, G. Della; Pawlik, J.R.; Teta, R.; Mangoni, A.; Costantino, V. Fast detection of two smenamide family members using molecular networking. *Mar. Drugs* **2019**, *17*, 1–12, doi:10.3390/md17110618.

[13] Esposito, G.; Della Sala, G.; Teta, R.; Caso, A.; Bourguet-Kondracki, M.L.; Pawlik, J.R.; Mangoni, A.; Costantino, V. Chlorinated Thiazole-Containing Polyketide-Peptides from the Caribbean Sponge Smenospongia conulosa: Structure Elucidation on Microgram Scale. *European J. Org. Chem.* **2016**, *2016*, 2871–2875, doi:10.1002/ejoc.201600370

[14] Schmidt, Y.; Lehr, K.; Colas, L.; Breit, B. Assignment of relative configuration of desoxypropionates by 1hnmr spectroscopy: Method development, proof of principle by asymmetric total synthesis of xylarinic acid a and applications. *Chem. - A Eur. J.* **2012**, *18*, 7071–7081, doi:10.1002/chem.201103988.

[15] Bertin, M.J.; Zimba, P. V.; He, H.; Moeller, P.D.R. Structure revision of trichotoxin, a chlorinated polyketide isolated from a Trichodesmium thiebautii bloom. *Tetrahedron Lett.* **2016**, *57*, 5864–5867, doi:10.1016/j.tetlet.2016.11.062.

[16] Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549, doi:10.1093/molbev/msy096.

[17] Nothias, L.F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; et al. Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* **2020**, *17*, 905–908, doi:10.1038/s41592-020-0933-6.

[18] Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks., doi:10.1101/gr.1239303.

[19] Hoye TR, Jeffrey CS, Shao F. Mosher ester analysis for the determination of absolute configuration of stereogenic (chiral) carbinol carbons. Nat Protoc. 2007;2(10):2451-8. doi: 10.1038/nprot.2007.354. PMID: 17947986.

Chapter 12

Monitoring of a bloom of cyanobacteria occurred in the Lake Avernus in Campania, Italy during the COVID-19 pandemic

12.1 Synopsis of this chapter [1]

The skills acquired during the PhD study allowed to collaborate in the analysis of an unusual off-season cyanobacterial bloom occurred in Lake Avernus (Lago d'Averno), a volcanic lake located in the Phlegrean Fields (Campi Flegrei) in Campania. A particular time-efficient protocol, integrating satellite imaging to address in situ sampling in selected spots and laboratory analysis of collected samples, including microscopy observations and chemical analysis, was used revealing a widespread bloom of *Microcystis sp.* straddling the COVID-19 pandemic lockdown in Italy (March 6 to June 15, 2020). Monitoring of Lake Avernus at the end of the COVID-19 block, when almost all urban activities were stopped, showed the disappearance of the bloom. Knowing that cyanoHABs have a harmful activity, the cytotoxicity of the crude extracts was evaluated in two different human cell lines, normal dermal fibroblasts (NHDF) and breast adenocarcinoma cells (MCF-7). The results of cytotoxicity assays showed that chloroform extract exerted the highest activity. MS-based molecular network performed on the crude extracts revealed a large group of cyanotoxins, namely microcystins, micropeptins, anabaenopeptins, and

aeruginopeptins in the chloroform extract, whose presence was verified to be related to bioactivity. Moreover, it was confirmed by metagenomic detection of the mcyB gene the production of MCs confirming the presence of *Microcystis sp*.

12.2 Remote and proximal sensing analysis

Multispectral data were recorded by the Sentinel 2 satellite and downloaded monthly. Despite conditions not favorable for a cyanobacterial bloom, due to low winter temperatures, the satellite data- processing sequence (NDVI index rendering in false color maps), an unexpected cyanobacterial bloom was verified in Lake Avernus which was confirmed to have begun in February 2020. In fact, starting this month a gradual increase in chlorophyll was recorded by remote sensing analysis. Lake Avernus (Lago d'Averno) is a volcanic Lake located in Phlegrean Fields, in the surroundings of which are spread several agricultural activities, housing, etc. It was hypothesized that the cyanobacterial bloom occurred due to an unusual heat wave recorded in February 2020. Abundant weather-driven vegetation blooms, with precipitation transporting nutrients, pesticides, and fertilizers into the lake, recreated favorable conditions for the resulting cyanobacterial bloom. Monitoring of recorded data revealed the highest peak of the bloom straddling the Covid-19 lockdown in Italy (April 2020) and a halt at the end of the lockdown, revealing a strong connection to anthropogenic activities.

Both the evolution of the bloom and the distribution of cyanobacteria on the shores were defined thanks to the analysis of satellite data.

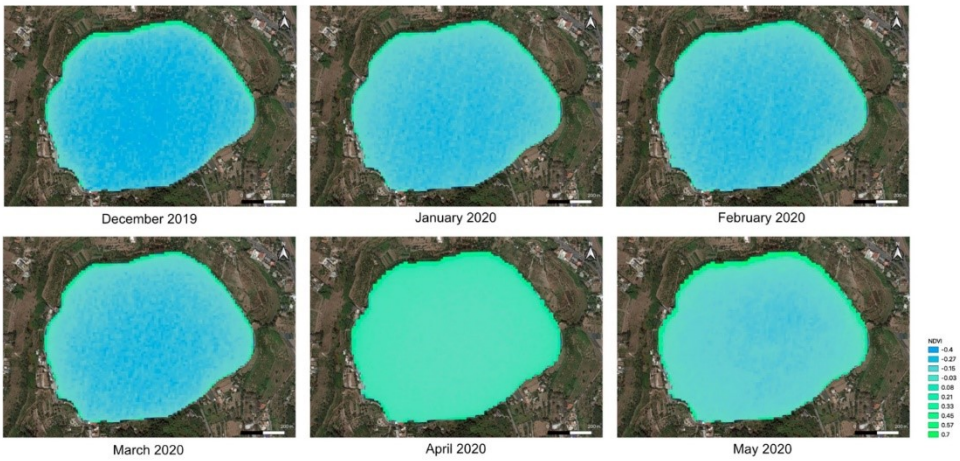


Figure 12.1. Temporal sequence from December 2019 to May 2020 of the processing of satellite multispectral data; in particular, the rendering of the NDVI index in false color maps is reported. The dataset used is that of Sentinel 2 and the spatial resolution is 10m/px.

On the NDVI map, it was not possible to clearly highlight the points of origin and of greatest biomass accumulation, so the increased resolution surveys of the area recorded by specific drones equipped with specific payloads, allowed a better localization of the sampling sites.[2]

Figure 12.2. Image acquired during the drone flight performed on 11th February 2020. The scanned area was centered on lake shores where the blooming was most widespread as indicated by NDVI data and selected as the sampling target.

Figure 12.3. Images acquired during sampling carried out on 11th February 2020. All the sampling points were defined within the hierarchical process started using satellites and drones.

12.3. Sampling and morphological identification

From the analysis of the samples collected, it was possible to detect a concentration of nitrates (0.5 mg/L) and phosphates (0.82 mg/L) higher than the limits imposed by the EEA, the European Environment Agency. Cyanobacterial cells were recognized as *Microcystis sp*. by the presence of typical morphological features observed in the floating suspension of cells under the microscope. The cells appeared spherical (2-8 um), rich in dark-colored vesicles, filled with gas, and covered with a colorless sheath.

Moreover, quantitative PCR on environmental DNA was performed to verify the presence of the biosynthetic gene cluster, to confirm the production of MCs.

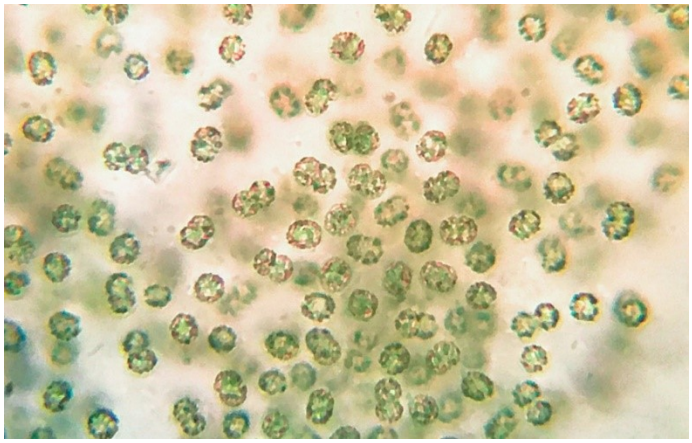


Figure 12.4. *Microcystis sp*. from cyanobacterial bloom at Lake Avernus.

12.4. Chemical analysis of the bloom

Centrifugation of the samples resulted in separation of pellet and supernatant, the latter showing a blue coloration indicative of the presence of phycocyanin. UV/Vis (-λ_MAX= 260nm) analysis of a filtered quantity of the sample confirmed the presence of phycocyanin.

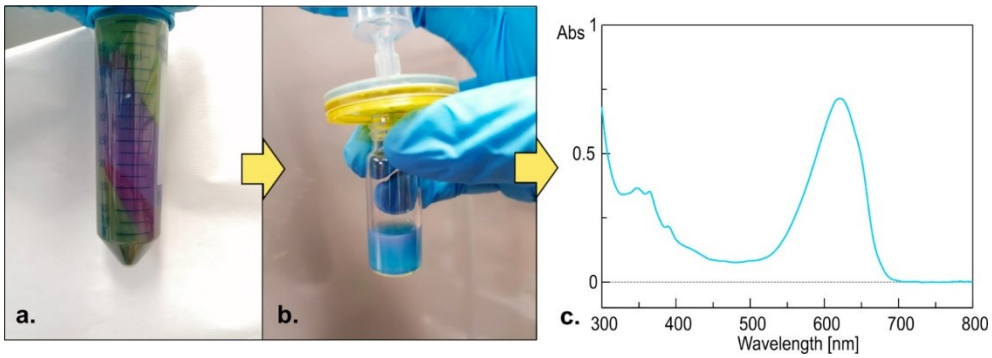


Figure 12.5. a) Centrifugation of the cellular suspension, b) filtration and c) UV/Vis analysis of the supernatant indicated the presence of phycocyanin in the sample, a typical cyanobacterial pigment.

The supernatant and the pellet were treated separately. The first one was partitioned with BuOH, whereas the pellet was extracted with MeOH and

$CHCl_3$ mixtures. The cytotoxicity activity of organic extracts was then evaluated.

MS-based molecular networking was carried out to obtain a visual representation of the metabolites present in the extracts. LC-HR-MS/MS (Liquid Chromatography High Resolution Tandem Mass Spectrometry) data were recorded using an LTQ Orbitrap instrument equipped with an electrospray (ESI) source. Data dependent acquisition was performed, selecting metabolites to be fragmented in a mass range of 700-1500 amu, in which most of the masses of known peptide toxins fall.

After the preprocessing of raw data with the MZmine2 program, [3] the .mgf file generated, was submitted on the GNPS (Global Natural Products Social Molecular Networking) website, performing Feature based molecular networking.[4] To easily dereplicate both peptide and non-peptide known natural products in the extracts, Dereplicator+ tool, was performed.[5] Moreover, MolNetEnhancer workflow, providing chemical structural information from GNPS library, was performed for the fast recognition of most abundant chemical classes.[6] Cytoscape 3.8.2 was used to visualize the enhanced network produced, in which nodes were mapped based on the classification of compounds obtained from MolNetEnhancer and metabolites putatively identified by Dereplicator + tool, were represented as hexagons.[7] The network consists of 107 molecular clusters, 41 of which belong to 7 recognized chemical classes, for a total of 435 nodes.

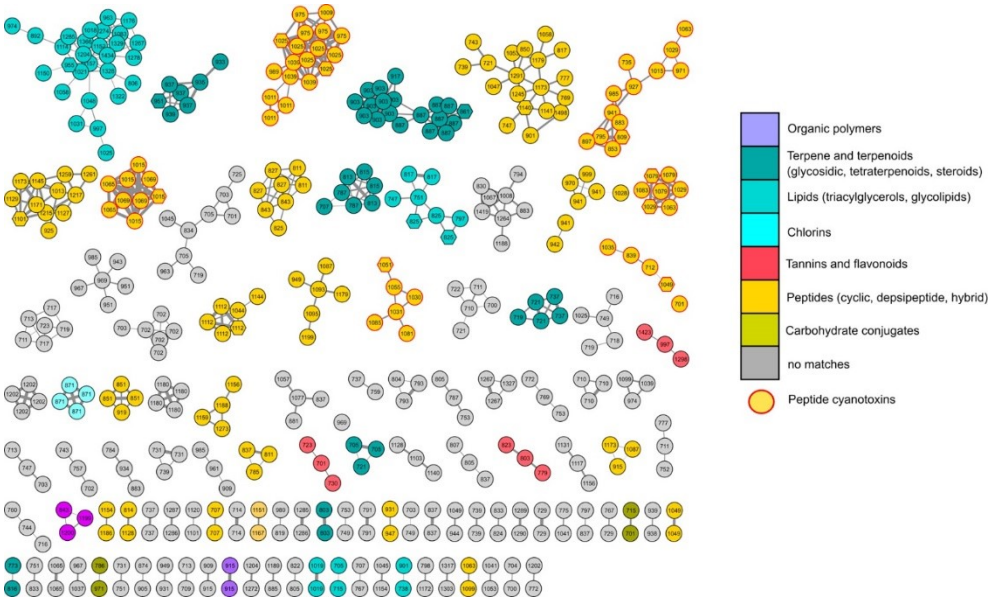


Figure 12.6. Enhanced molecular network of Lake Avernus' cytotoxic extract. Nodes are labeled with parent mass. The color of each node is related to the compound classification derived from MolNetEnhancer and nodes annotated by Dereplicator+ with a putatively identified metabolite are represented as hexagons. Nodes belonging to cyanotoxins are bordered in red. Edge thickness is related to cosine score similarity.

As demonstrated by the molecular network, the most abundant class of compounds in the extracts was represented by peptides, comprising cyclic peptides, hybrids, and depsipeptides. Nodes referred to peptide cyanotoxins were shown in six clusters, demonstrating the toxicity of the extract. More precisely, three nodes were annotated by the Dereplicator+ tool as microcystins, (MC-LW, MC-M(O)R, [D-Asp]^{3}MC-FR),[8] two nodes as micropeptin (MP-88B, MP-T1), [9,10] and two others were annotated as anabaenopeptin C and aeruginopeptin 228B.[11,12] The cluster containing cyanotoxins, revealed the presence of novel natural products, with a similar structure to the known cyanotoxins.

Table 12.1 Peptide cyanotoxins as annotated by Dereplicator+ and molecular networking.

t_R (min)	*m/z*	Molecular Formula $[M+H]^+$	GNPS' Annotation[1]
24.5	1015.5304	$C_{51}H_{71}O_{12}N_{10}^+$	[D-Asp[3]] MC-FR
25.8	1025.5350	$C_{54}H_{73}O_{12}N_8^+$	MC-LW
25.2	1029.5084	$C_{48}H_{73}O_{13}N_{10}S^+$	MC-M(O)R
29.5	1051.5120	$C_{55}H_{71}O_{13}N_8^+$	MP-T1
23.8	1079.5288	$C_{53}H_{75}O_{16}N_8^+$	MP-88B
33.6	809.4596	$C_{41}H_{61}O_9N_8^+$	AnaP-C
23.7	1049.5165	$C_{52}H_{73}O_{15}N_8^+$	AerP-228B

[1] Abbreviations: MC - microcystin; MP - micropeptin; AnaP - anabaenopeptin; AerP – aeruginopeptin

12.5 Metagenomic detection of the mcy gene

The further confirmation of the presence of *Microcystis sp.* in the extracts was obtained by metagenomic DNA analysis. Environmental DNA was tested for the presence of the *mcy*B gene from MC biosynthetic gene cluster, confirming the production of MCs by the collected cyanobacteria. Metagenomic DNA extraction and analysis by quantitative PCR using SYBR Green I technology and *mcy*B-specific primers was performed. [13] At the 27th cycle of PCR amplification, the cyanobacterial 16S-ITS (internal transcribed spacer) rRNA gene, used as housekeeping gene, was evident, giving confirmation of the quality of the extracted DNA. In addition, it was possible to verify the presence of the gene, in the sample because at the 32nd PCR amplification cycle, the *mcy*B amplicon appeared.

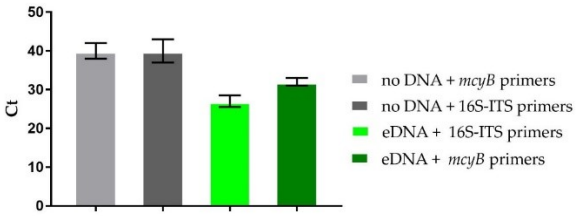


Figure 12.7. Presence of the mcyB gene in the cyanobacterial sample collected in Lake Avernus. qPCR results showing the primer specific amplification of the mcyB gene as function of the threshold PCR cycle (Ct). The amplification of 16S-ITS is shown as positive control of amplification. No amplification could be detected in the absence of eDNA (Ct > 38).

12.6 Cytotoxicity evaluation of organic extract

The effective production of harmful microcystins was also confirmed through cytotoxicity assays. The cytotoxic activity of methanol, chloroform, and *n*-butanol extracts of the cyanobacterial bloom was evaluated on two different human cancer cell lines. Normal dermal fibroblasts (NHDF) were used as a normal cell model, while MCF-7 mammary adenocarcinoma cells were used as a tumor cell model. A real-time monitoring of cell proliferation was assessed through the xCELLigence cell analyzer, which provides a normalized cellular index (NCI), from alterations in electronic impedance, reflecting cell viability and morphology. Three different concentrations of extracts (50, 100, 200 μg/mL) were evaluated for 48 hours. The most significant cytotoxicity toward both NHDF and MCF-7 cells was shown by the chloroform extract within the first 24 hours of exposure. More precisely, a significant reduction (approximately 60%) in the slope of the growth curve at the highest dose tested was exhibited by the chloroform fraction. From the comparison of relative controls treated with 0.5% DMSO vehicle, it was possible to note the increase in cell doubling time, in cells treated with the chloroform fraction. These results led to the confirmation of the richness of cyanobacterial cytotoxic metabolites in this fraction.

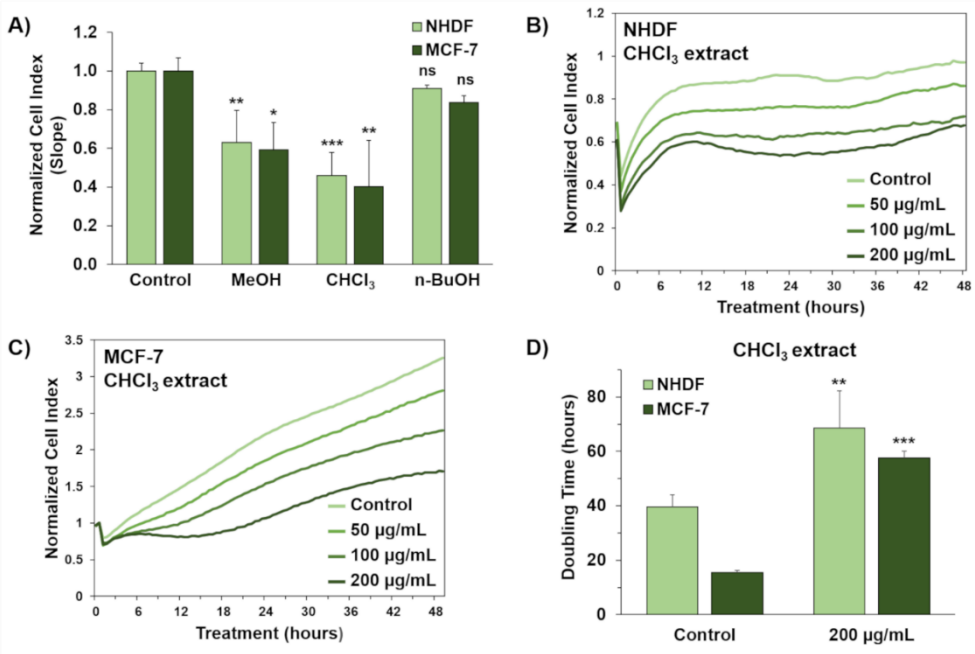


Figure 12.8 In vitro cytotoxicity of organic extracts from the cyanobacterial bloom in Lake Avernus. (**A**) Evaluation of cytotoxic activity of organic extracts at the highest concentration used in this study (200 µg/mL). Cytotoxic effects are reported as slope of normalized cell index (NCI) to describe the changing rate of growth curves after 24h exposure to the crude extracts. NCI slope values are relative to controls treated with 0.5% DMSO vehicle. (**B,C**) Representative NCI traces of NHDF (**B**) and MCF-7 (**C**) cells after exposure to 0.5% DMSO vehicle (control) and different concentrations (50, 100, 200 µg/mL) of the $CHCl_3$ extract from the cyanobacterial bloom. (**D**) NCI doubling times of NHDF and MCF-7 cells after 24 h exposure to 0.5% DMSO vehicle (control) and 200 µg/mL of the $CHCl_3$ extract from the cyanobacterial bloom. Doubling time is the time required for cell index to double. Data are presented as mean ± SD; all experiments have been performed at least in triplicate. The one-way analysis of variance (ANOVA) method was used to compare means of groups and Dunnett's method was used as post-hoc test. Statistical significances are referred to the DMSO control. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns, not significant.

12.7 Conclusion

The verified out-of-season blooming of microcystins in Lake Avernus represent a noteworthy environmental issue for living beings but also for ecosystems, given the connection of the lake with the sea, where more extensive coastal marine phenomena can occur. Indeed, the spread of contamination can also affect the food chain, due to a multitude of shellfish farms in the lake and on the coast, causing serious consequences.

12.8 Experimental section

12.8.1 Remote and proximal sensing analysis

To calculate the biophysical parameters of the observed areas, satellite data are used. In detail, the availability of multispectral data has allowed the calculation of specific indices such as NDVI, which provides values between -1 and 1 being calculated through a combination of red bands and Near Infra-Red (NIR). [14] The expression to calculate NDVI is:

$$NDVI = \frac{NIR - red}{NIR + red}$$

where NIR and red refer to the reflectance values for the specified satellite bands. Sentinel-2, operated by the European Space Agency (ESA), is constituted by a group of two polar satellites operating in the same orbit, phased at 180_ to each other. Its swath width is 290 km and its revisit time is 10 days (at the equator, 1 satellite) or 5 days (2 satellites, cloud-free conditions, equivalent to 2–3 days at mid-latitudes). Sentinel-2 carries an optical instrument payload, the multispectral instrument (MSI), that with a push-broom concept samples 13 spectral bands: 4 bands at 10 m, 6 bands at 20 m, and 3 bands at 60 m spatial resolution. In the case of Sentinel-2, the red band corresponds to the fourth channel (bandwidth: 640–690 nm), while the NIR band corresponds to the eighth channel (bandwidth: 780–910 nm), with a pixel spatial resolution of 10 m. The same index can be calculated using band 5 (NIR) and band 4 (red). However, the pixel spatial resolution is of 30 m. This is why, in order to obtain a higher spatial resolution, Sentinel-2 data were used instead of Landsat 8. In addition to the lower spatial resolution, the rendering of the Landsat 8 data presented several

problems due to the high cloud cover in the acquisition dates and a high "banding" that did not allow us to correctly examine the phenomenon. Several other indices (CI, MCI, CARI, Chlgreen, Chlred-edge, RBD, and CVI) could also be calculated on the basis of available satellite data having specific available spectral bands to formulate specialized indices that allow researchers to carry out not only qualitative, but also quantitative analyses.

Among the others, the MODerate resolution imaging spectroradiometer (MODIS), installed on both the Terra and Aqua satellites, and providing imagery of 36 bands at three spatial resolutions (from 250 m to 1000 m), allows researchers to develop an algorithm specifically designed to quantify the concentration of Chlorophyll-a. Chlorophyll-a near-surface concentrations, expressed in mg/m^3, are obtained with a pixel spatial resolution of 1 km x 1 km with a daily temporal resolution. [15]

The limited resolution of most of these products of these platforms reduce the study of only large-scale phenomena. The combined use of remote (satellites and airplanes) and proximal (drones) sensing techniques has exceeded the previous limits allowing researchers to reach new goals in earth observation. The purpose of remote/proximal sensing is related to the early detection of blooms; thus, this study focuses on qualitative detection methods, which provide a reliable first screening of bloom phenomena. In particular, the application of these observation techniques was to detect the presence of anomalies and to circumscribe the areas on which to perform further investigations and samplings. This is why the NDVI index mapping was sufficient to meet the planned objective and, above all, to maximize the efficiency of the process in terms of time and costs. NDVI maps were

generated, on the basis of elaborated Sentinel-2 data, classifying the results on 10 distinct classes and representing them in a false-color scale, both customized for the specific problem and calibrated with field experiments. In this regard, ground truth points were defined, where the certainty of the presence of bloom was attributed. The entire analysis process was standardized in all steps, so that it can be replicated for each download of data associated with different dates. This approach made it possible to compare each processing and clearly visualize the trends. Regarding the proximal sensing techniques, a mini-drone was used for the specific purpose. It carried two electro-optical sensors: a 12-megapixel multispectral camera able to acquire three bands (Orange + Cyan + NIR) and a 20 megapixels RGB camera with a one-inch sensor.

12.8.2 Sampling, extraction, LC-MS and molecular networking analysis

The cyanobacterial bloom was sampled along the coast of Lake Avernus (40°50′20.4″N 14°04′33.6″E) on 11 February. An optical microscope combined with an OMAX 18 MP CMOS cooled camera was used to visualize the cell suspension and the software ToupView was used to analyze data. The taxonomic classification and the species identification were accomplished thanks to Komarek et al.,[16] evaluating environmental parameters temperature, salinity, pH, nitrate and phosphate content (Hanna Instrument). First, sample (0.3 L) was centrifuged (10,000 rpm for 5 min) and pellets were sonicated for 5 minutes and extracted with MeOH (100%, 0.3 L), MeOH/$CHCl_3$ (1:1, 0.3 L) and $CHCl_3$ (100%, 0.3 L). All extracts were filtered with paper and concentrated in vacuo. All MeOH and all

chloroform extracts were separately combined giving 851.8 mg of MeOH and 302.8 mg of chloroform extracts. In contrast, supernatant was extracted with BuOH and this latter was concentrated in vacuo, yielding to 109.6 mg of BuOH extract. Cell viability assays were performed with all the obtained extracts.

The active chloroform extract was re-suspended in MeOH (100%) at a concentration of 5 mg/mL and subjected to LC-HRMS and LC-HRMS/MS analysis. A Thermo LTQ Orbitrap XL high-resolution ESI mass spectrometer coupled to an Agilent model 1100 LC system, equipped with a solvent reservoir, in-line degasser, binary pump, and refrigerated autosampler, was used to perform LC-HRMS and LC-HRMS/MS analysis. A Kinetex 5 μm, 100 mm × 2.1 mm C18 column (Phenomenex, Torrance, CA, USA), at a temperature of at 25 °C, was eluted at 200 μL/min with H_2O (supplemented with 0.1% HCOOH) and MeOH using gradient elution. The gradient program was set as follows: 10% MeOH for 3 min, 10%→100% MeOH for 30 min, 100% MeOH for 7 min. Mass spectra were acquired in positive ion detection mode using the following MS parameters: a spray voltage of 5 kV, a capillary temperature of 285 °C, a sheath gas rate of 32 units N_2 (ca. 150 mL/min), and an auxiliary gas rate of 15 units N_2 (ca. 50 mL/min).

Data-dependent acquisition mode was used to collect the data. Specifically, the ninth most intense ions from a full-scan mass spectrum were analyzed. In addition, only ions with an *m/z* in the range of 700 to 1500 amu were selected and subjected to CID fragmentation using the following parameters: isolation width of 2.0, normalized collision energy of 35, an

activation Q of 0.250, and activation time of 30 ms. Thermo Xcalibur software was used to analyze the raw data. The raw files were directly imported into MZmine 2.53.[3] Detection of exact masses with mass level 1 and centered masses with mass level 2 was performed, setting the noise level to 1,000. After chromatogram building with the ADAP module (with a minimum height of 1,000, and *m/z* tolerance of 0.01 (or 20 ppm), chromatographic deconvolution was performed. A chromatographic threshold = 5%, minimum retention time interval = 0.20 min, minimum relative height = 30%, minimum absolute height = 1,000, minimum peak top/edge ratio = 1.3, and peak duration range = 0.0-6.0 min were set. Join aligner algorithm was performed for peak alignment, using the following parameters: *m/z* tolerance at 0.02 (or 10 ppm), absolute RT tolerance at 0.5 min. Final filtering of the adduct ([M+Na-H], [M+K-H], [M+Mg-2H], [M+NH_3], [M-Na+NH_4], [M+1, ^{13}C]) was performed by setting the maximum relative height to 100% and peaks without MS/MS spectra were filtered out. Clustered data were exported into a .mgf file and the quantitative of data into a .csv file. Both files were submitted to GNPS where a molecular network was generated. The Metabolomics workflow was performed by setting the parent mass tolerance and MS/MS fragment ion tolerance to 0.05 Da and 0.5 Da, respectively, the cosine score to above 0.6, and matched peaks to above 5. Spectra were retained only if the nodes appeared in each other's respective top 10 most similar nodes. Dereplication of the compounds in the network was performed by comparing their spectra with GNPS spectral libraries using a cosine score above 0.6 and at least 5 matched peaks. Dereplicator+ tool was used to annotate peptide and non-

peptide NPs in the network using the following parameters: precursor and fragment ion mass tolerance were set to 0.05 Da and 0.2 Da, respectively; minimum score for significance was set to 4 and fragmentation mode to 2-1-3.[5] Results obtained from molecular networking and Dereplicator+ were combined through MolNetEnhancer [6] workflow. the final network was visualized by Cytoscape software.[7]

12.8.3 Cell viability assays

xCELLigence System Real-Time Cell Analyzer (ACEA Biosciences, San Diego, CA, USA) was used to perform cell viability assays. NHDF cells were seeded at a density of 5000 cells/well, while MCF-7 cells were seeded at a concentration of 3000 cells/well and both were cultured in DMEM high glucose (4.5 g/L) medium, supplemented with 10% fetal bovine serum, penicillin–streptomycin (100 U/ml), and 2 mM L-glutamine. Approximately 24h after seeding, the medium was removed and cells were treated with a medium containing 50, 100, 200 μg/mL of organic extracts from the cyanobacterial bloom sample. Organic extracts were dissolved in DMSO to prepare 40 mg/mL stock solutions, which were diluted in culture medium to perform antiproliferative assays. All antiproliferative assays were performed with a maximum final concentration of DMSO of 0.5%, shown to be well tolerated with no observable toxic effects to cells. The cytotoxicity was represented as a) slope values of cell growth curves or as b) cell doubling times. Cell index variations were measured to generate growth curves. Moreover, just before treatment cell index was normalized, obtaining a normalized cell index (NCI). RTCA-integrated software (version

2.0.0.1301, ACEA Biosciences, San Diego, CA, USA) was used to calculate NCI slopes, doubling times and real-time NCI traces.

12.8.4 RT-PCR experiments

Amplification of 250 ng of environmental DNA was performed by PCR in an Applied Byo-Sistem StepOne Plus Real-Time Detector System (Thermo Fisher Scientific) equipped with SYBR Green double-stranded DNA fluorescent dye (Thermo Fisher Scientific). Cyanobacteria-specific primers 16S-ITS rRNA CYA359F (5′-GGGGAATYTTCCGCAATGG-3′) and 373R (5′-CTAACCTGAGCTAAT-3′), and *mcyB*-specific primers FAA (5′-CTATGTTATTTATACATCAGG-3′) and RAA (5′-CTCAGCTTAACTTGATTATC -3′) were used to perform PCR screening of metagenomic DNA. Primers were used at a final concentration of 200 nM. The cycling conditions used were 95°C for 10 min followed by 40 cycles at 95°C for 15 s and 60°C for 1 min, running the samples in triplicate. The sample was considered positive for the presence of the *mcyB* gene, when the amplification product appeared before the end of the exponential phase of the PCR.

12.9 Spectroscopic data

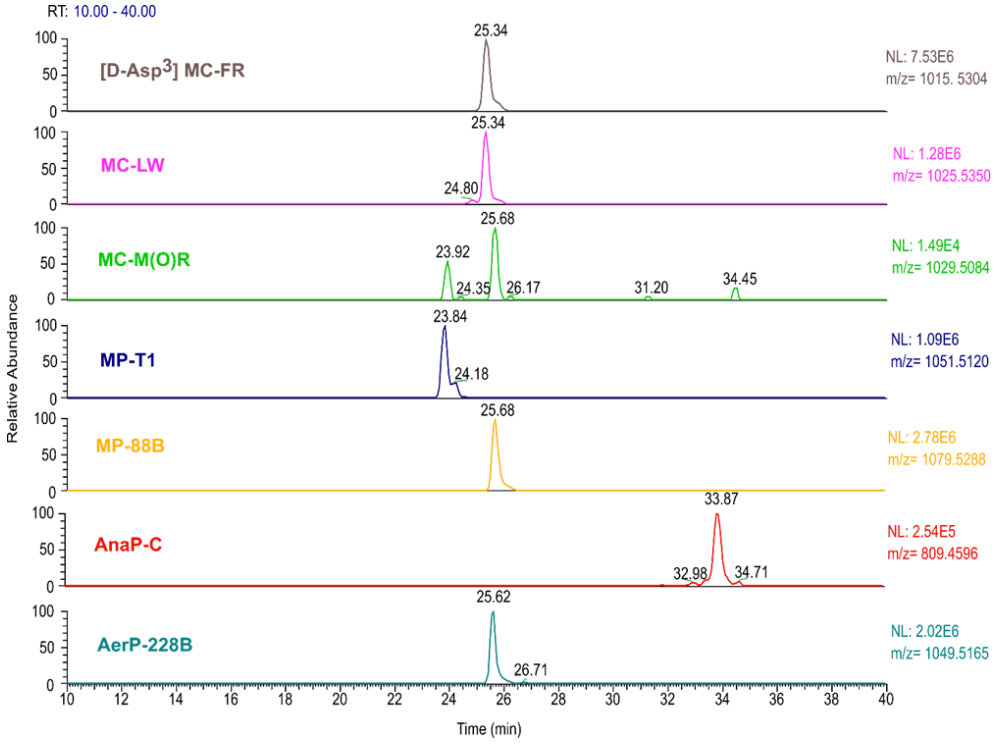


Figure 12.9. Extracted ion chromatograms of the putatively identified peptide cyanotoxins at *m/z* 1015.5304 ([D-Asp3] MC-FR), *m/z* 1025.5350 (MC-LW), *m/z* 1029.5084 (MC-M(O)R), *m/z* 1051.5120 (MP-T1), *m/z* 1079.5288 (MP-88B), *m/z* 809.4596 (AnaP-C), *m/z* 1049.5165 (AerP-228B).

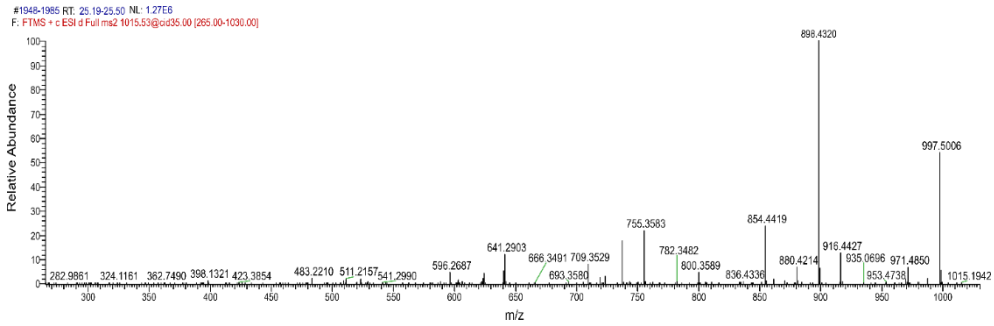


Figure 12.10. High resolution MS/MS spectrum of the $[M+H]^+$ ion at *m/z* 1015.5304, putatively identified as [D-Asp^3] MC-FR.

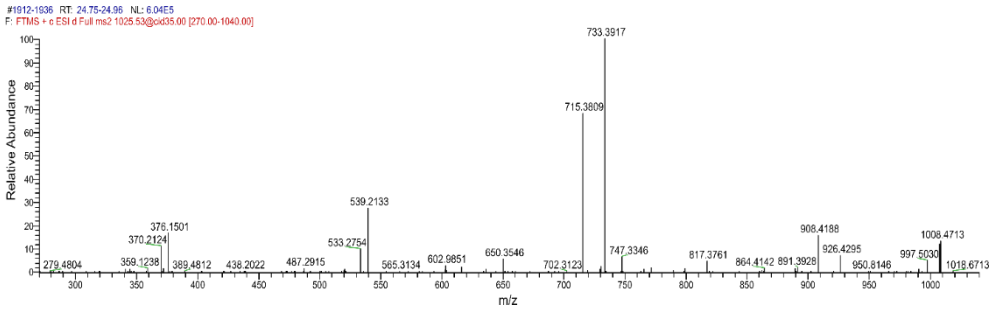


Figure 12.11. High resolution MS/MS spectrum of the $[M+H]^+$ ion at *m/z* 1025.5350, putatively identified as MC-LW.

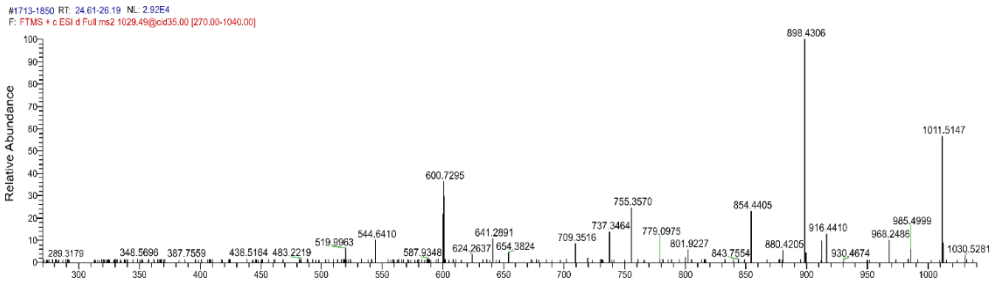


Figure 12.12. High resolution MS/MS spectrum of the $[M+H]^+$ ion at *m/z* 1029.5084, putatively identified as MC-M(O)R.

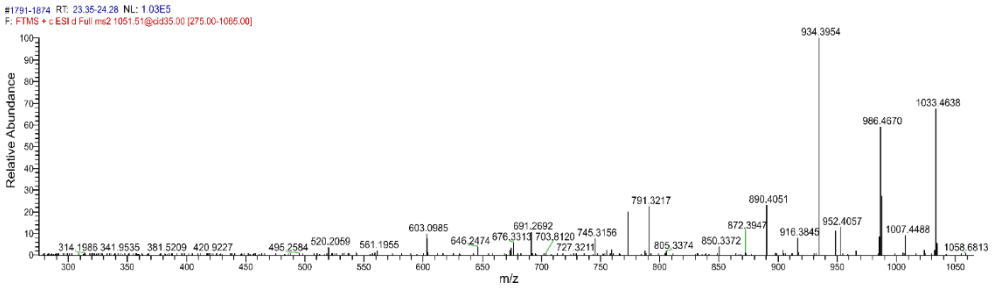


Figure 12.13. High resolution MS/MS spectrum of the $[M+H]^+$ ion at *m/z* 1051.5120, putatively identified as MP-T1.

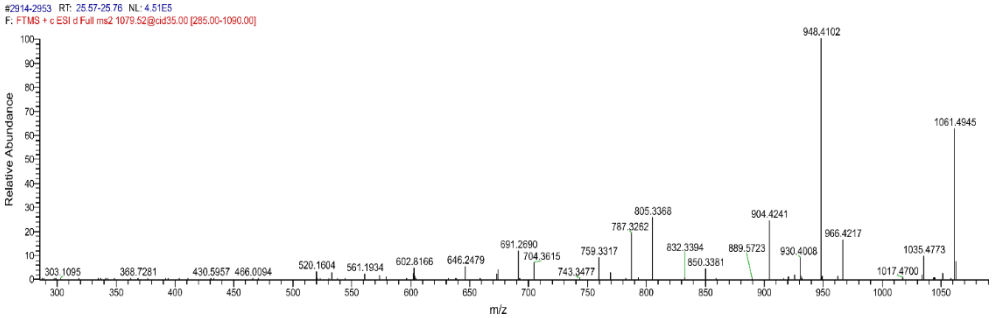


Figure 12.14. High resolution MS/MS spectrum of the $[M+H]^+$ ion at *m/z* 1079.5288, putatively identified as MP-88B.

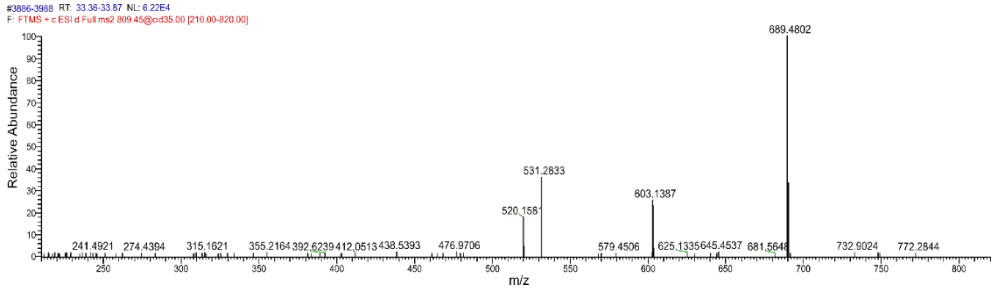


Figure 12.15. High resolution MS/MS spectrum of the $[M+H]^+$ ion at *m/z* 809.4596, putatively identified as AnaP-C.

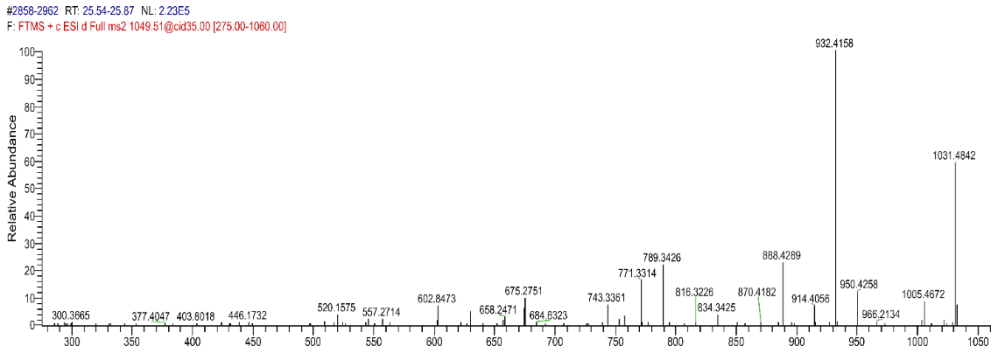


Figure 12.16. High resolution MS/MS spectrum of the $[M+H]^+$ ion at *m/z* 1049.5165, putatively identified as AerP-228B.

References

[1] Teta, R.; Sala, G. Della; Esposito, G.; Stornaiuolo, M.; Scarpato, S.; Casazza, M.; Anastasio, A.; Lega, M.; Costantino, V. Monitoring cyanobacterial blooms during the COVID-19 pandemic in Campania, Italy: The case of lake avernus. *Toxins (Basel).* **2021**, *13*, doi:10.3390/toxins13070471.

[2] Di Fiore, V.; Cavuoto, G.; Punzo, M.; Tarallo, D.; Casazza, M.; Guarriello, S.M.; Lega, M. Integrated hierarchical geo-environmental survey strategy applied to the detection and investigation of an illegal landfill: A case study in the Campania Region (Southern Italy). *Forensic Sci. Int.* **2017**, *279*, 96–105, doi:10.1016/j.forsciint.2017.08.016.

[3] Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010**, *11*, doi:10.1186/1471-2105-11-395.

[4] Aron, A.T.; Gentry, E.C.; McPhail, K.L.; Nothias, L.F.; Nothias-Esposito, M.; Bouslimani, A.; Petras, D.; Gauglitz, J.M.; Sikora, N.; Vargas, F.; et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nat. Protoc.* **2020**, *15*, 1954–1991, doi:10.1038/s41596-020-0317-5.

[5] Mohimani, H.; Gurevich, A.; Shlemov, A.; Mikheenko, A.; Korobeynikov, A.; Cao, L.; Shcherbin, E.; Nothias, L.F.; Dorrestein, P.C.; Pevzner, P.A. Dereplication of microbial metabolites through database search of mass spectra. *Nat. Commun.* **2018**, *9*, 1–12, doi:10.1038/s41467-018-06082-8.

[6] Ernst, M.; Kang, K. Bin; Caraballo-Rodríguez, A.M.; Nothias, L.F.; Wandy, J.; Chen, C.; Wang, M.; Rogers, S.; Medema, M.H.; Dorrestein, P.C.; et al. Molnetenhancer: Enhanced molecular networks by integrating metabolome mining and annotation tools. *Metabolites* **2019**, *9*, doi:10.3390/metabo9070144.

[7] Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks., doi:10.1101/gr.1239303.

[8] Puddick, J. Spectroscopic Investigations of Oligopeptides from Aquatic Cyanobacteria Characterisation of New Oligopeptides , Development of Microcystin Quantification Tools and Investigations into Mi, The University of Waikato, Hamilton, New Zealand, 2013.

[9] Ishida, K.; Matsuda, H.; Murakami, M. Micropeptins 88-A to 88-F, chymotrypsin inhibitors from the cyanobacterium Microcystis aeruginosa (NIES-88). *Tetrahedron* **1998**, *54*, 5545–5556, doi:10.1016/S0040-4020(98)00242-7.

[10] Kodani, S.; Suzuki, S.; Ishida, K.; Murakami, M. Five new cyanobacterial peptides from water bloom materials of lake Teganuma (Japan). *FEMS Microbiol. Lett.* **1999**, *178*, 343–348, doi:10.1016/S0378-1097(99)00379-1.

[11] Namikoshi, M.; Rinehart, K.L. Bioactive compounds produced by cyanobacteria. *J. Ind. Microbiol. Biotechnol.* **1996**, *17*, 373–384, doi:10.1007/bf01574768.

[12] Harada, K. ichi; Mayumi, T.; Shimada, T.; Suzuki, M.; Kondo, F.; Watanabe, M.F. Occurrence of four depsipeptides, aeruginopeptins, together with

microcystins from toxic cyanobacteria. *Tetrahedron Lett.* **1993**, *34*, 6091–6094, doi:10.1016/S0040-4039(00)61736-7.

[13] Teta, R.; Della Sala, G.; Glukhov, E.; Gerwick, L.; Gerwick, W.H.; Mangoni, A.; Costantino, V. Combined LC-MS/MS and Molecular Networking Approach Reveals New Cyanotoxins from the 2014 Cyanobacterial Bloom in Green Lake, Seattle. *Environ. Sci. Technol.* **2015**, *49*, 14301–14310, doi:10.1021/acs.est.5b04415.

[14] Vermote, E.; Justice, C.; Claverie, M.; Franch, B. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sens. Environ.* **2016**, *185*, 46–56, doi:10.1016/j.rse.2016.04.008.

[15] Hooker, S.B.; Firestone, E.R.; O'reilly, J.E.; Maritorena, S.; O'brien, M.C.; Siegel, D.A.; Toole, D.; Menzies, D.; Smith, R.C.; Mueller, J.L.; et al. *SeaWiFS Postlaunch Technical Report Series Volume 11, SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3*; 2000; Vol. 11;.

[16] Hu, C.; Lee, Z.; Franz, B. Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *J. Geophys. Res. Ocean.* **2012**, *117*, doi:10.1029/2011JC007395.

Conclusion

The research carried out during this PhD project was aimed on the investigation of extracts from different marine organisms, with the main goal of identifying and structurally elucidating novel natural products as possible lead compounds. A significant part of the research activity was focused on molecular networking, an innovative bio-informatic strategy that provides a visual overview of the chemical repertoire of complex extracts, processing the massive amount of data obtained by LC-MS/MS analysis of one or more extracts. The effectiveness of integrating traditional dereplication approaches with molecular networking, was demonstrated both in the discovery of new natural compounds and for environmental purposes. In particular, in the frame of the research of new bioactive natural products from porifera, the workflow that integrates the analysis of LC-MS/MS data of extracts with Feature-Based Molecular Networking allowed to the straightforward isolation of a new cyclic heptapeptide, stylissamide L (**1**) from the extract of the widely studied sponge *Stylissa caribica*. The molecular networking-based isolation of a new phospholipid from the extracts of the sponge *Clathria faviformis*, further demonstrated the usefulness of the methodology. However, the small amount isolated of favilipid A (**2**) highlights an important issue with MS based identification of new natural products. Even though the node at *m/z* 533.33 is one of the largest in the network, the amounts of of corresponding metabolite favilipid A (**2**) are not as large as expected. In fact, despite the intensity of an LC-MS peak is certainly proportional to the amounts of the compounds in the

extract, the ionization efficiency in the ESI source can be dramatically different from compound to compound. Favilipid A, (**2**) a zwitterionic compound, is apparently ionized very efficiently, and give rise to large LC-MS peak even if present in a low concentration in the extract. A further confirmation of the strength of the methodology was obtained from the analysis of extracts of the free-living cyanobacterium *Trichodesmium thiebautii*, resulting in the isolation of 4 new polyketide compounds, included in the class of trichophycins (isotrichophycin C (**12**) and trichophycins G-I (**13**-**15**), through a Feature-Based Molecular Networking approach, focused on halogenated compounds in the extract. In the context of environmental monitoring, aimed at the identification of potentially toxic cyanobacterial blooms, the validity of the use of molecular networking for the analysis of extracts was verified in the study of an unusual cyanobacteria bloom occurred in Lake Avernus at the turn of the lockdown period from COVID-19 in Italy. Indeed, this method provided a comprehensive picture of the compounds accumulated during the harmful bloom, allowing the identification of toxic peptides by in-silico dereplication, performed on the freely accessible website GNPS, and confirmed by genome amplification analysis. At the same time, the in-depth investigation of extracts from the common seagrass *Zostera marina*, allowed to strengthen the expertise on the structural elucidation of unknown secondary metabolites from marine extracts, leading to the structural determination and the clarification of the absolute configuration of two new tetracyclic diarylheptanoids, zosteraphenol A (**3**) and zosteraphenol B (**4**) and two unique dimeric diarylheptanoid, zosterabisphenone A (**5**) and B (**6**) through an extensive

NMR study, with the support of quantum-mechanical calculations and of the new phenolic acid 7″,8″-didehydrosalvianolic acid (**7**).