Chemical fingerprints of natural origin:

Analyses of secondary metabolite profiles from plants, micro- and macrofungi

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät II

Chemie, Physik und Mathematik

der Martin-Luther-Universität

Halle-Wittenberg

vorgelegt von

Frau Dipl. Leb. Chem. Annegret Laub

geb. am 09.11.1990 in Querfurt

This dissertation has been developed under the supervision of Prof. Dr. Ludger A. Wessjohann and mentorship of Dr. Katrin Franke, Dr. Jürgen Schmidt, and Dr. Norbert Arnold at the Leibniz Institute of Plant Biochemistry (IPB) in cooperation with the Martin Luther University Halle-Wittenberg.

1st Reviewer: Prof. Dr. Ludger A. Wessjohann

2nd Reviewer: Prof. Dr. Georg Pohnert

Date of public defense: 17.05.2022

Die Neugier steht immer an erster Stelle eines Problems,

das gelöst werden will.

Galileo Galilei

Chemical fingerprints of natural origin:

Analyses of secondary metabolite profiles from plants, micro- and macrofungi

Table of Contents

Acknowledgements

List of Abbreviations

Å	Angstrom		**min**	Minute(s)
Ac$_2$O	Acetic anhydride		**MPA**	Malt peptone agar
amu	Atomic mass units		**MS**	Mass spectrometry
calcd	Calculated		**MS^n**	Multistage mass spectrometry
CD	Circular dichroism		**MVDA**	Multivariate data analysis
CH$_3$OH	Methanol		**NP**	Natural product
CH$_3$CN	Acetonitrile		**NCE**	Normalized collision energy
CID	Collision-induced dissociation		**NRPS**	Non-ribosomal peptide synthetase
COSY	Correlation spectroscopy		**NMR**	Nuclear magnetic resonance
d	Doublet in NMR		**P**	Peak number
Da	Dalton		**PC**	Principal component
DAD	Diode array detector		**PCA**	Principal component analysis
DEPT	Distortionless Enhancement by Polarization Transfer		**ppm**	Parts per million
DESI	Desorption electrospray ionization		*q*	Quartett
DMSO	Dimethylsulfoxide		**QC**	Quality control
eq.	Equivalent(s)		**RDBE**	Ring and double bond equivalent
ESI	Electrospray ionization		**r.t.**	Room temperature
FA	Formic acid		**rel. int.**	Relative intensity
Hz	Hertz		*R*$_f$	Retention factor
HCA	Hierarchial cluster analysis		**RP**	Reversed phase
HESI	Heated electrospray ionization		**ROESY**	Rotational frame Overhauser effect
HMBC	Heteronuclear multiple bond correlation		**ROS**	Reactive oxygen species
HMDS	Hexamethyl disiloxane		**SCPC**	Sequential centrifugal partition chromatography
HPLC	High-performance liquid chromatography		**SIM**	Single ion monitoring
HPTLC	High-performance thin-layer chromatography		*s*	Singlet
HRMS	High-resolution mass spectrometry		**S/N**	Signal-to-noise ratio
HSQC	Heteronuclear single quantum Correlation		**spec**.	Species
IC$_{50}$	Concentration of a compound needed to inhibit the growth by half		**spp.**	Species (more than one)
IS	Internal standard		**SPE**	Solid-phase extraction
i.e.	*id est* (that is)		**SPPS**	Solid-phase peptide synthesis
J	coupling constant		**TMS**	Tetramethylsilane
KSH	Kultursammlung Halle		**TOF**	Time-of-flight
LAESI	Laser ablation electrospray ionization		**TLC**	Thin-layer chromatography
L.	Linné		**TOCSY**	Total correlation spectroscopy
LC-MS	Liquid chromatography/mass spectrometry		**TIC**	Total ion chromatogram/current
m/z	Mass-to-charge-ratio		*t*	Triplet
m	Multiplet		**t$_R$**	Retention time
MALDI	Matrix assisted laser desoption/ionization		**UHPLC**	Ultra high-performance liquid chromatography
MeOH	Methanol		**UV/Vis**	Ultraviolet/visible

Summary

Natural products from plants, fungi, and microorganism remain among the most important sources for new bioactive compounds for the development of pharmaceutical products and for their use in other industry sectors. Nowadays, the application and combination of advanced analytical methods in conjunction with sophisticated data mining approaches constitute the foundation for natural product research and discovery. In consequence, the goal of the present thesis was the development of effective workflows for the in-depth investigation of the secondary metabolite profiles of plants (Balsaminaceae), micro- (*Sepedonium*), and macrofungi (*Cortinarius*) using advanced analytical hyphenated techniques, in particular UHPLC-ESI-HRMS and HPTLC-DESI-HRMS. Targeted and untargeted approaches were applied to explore the chemical diversity within different species for the dereplication and identification of their metabolites. Additionally, possible chemophenetic markers for species differentiation, as well as the presence of new bioactive metabolites (assisted by bioactivity screening) were investigated.

A comprehensive untargeted metabolite profiling approach by UHPLC-ESI-HRMS and NMR of 31 *Impatiens* species and *Hydrocera triflora* (including 19 phytochemically uninvestigated species) showed a high interspecific variation as reflected by variability of leaf metabolites and flavonoid glycosides as one of the major compound classes (Chapter 2). The combination of LC-ESI-HRMS and NMR allows for a more complete picture (overview) of the metabolome. Phytochemical investigation of *I. racemosa* DC, *I. ethiopica* Grey-Wilson, and *I. flanaganae* Hemsl lead to the successful isolation and characterization of dihydrochalcone derivatives, astragalin, and 2-methoxynaphthoquinone. Preliminary investigations of the biological activity (anthelmintic, anticancer and antibacterial) indicated the antimicrobial potential of different *Impatiens* species (such as, *I. andringitrensis*, *I. arguta*, *I. balansae*, among others) and *Hydrocera triflora*.

Two new cyclic pentapeptides (named microsporide A and B) were detected and structurally elucidated from the enriched crude extract of *Sepedonium microspermum* by intensive ESI-HRMS^n investigations (Chapter 3). MS-guided isolation yielded microsporide A as isomeric mixture and pure chrysosporide, a known cyclic pentapeptide isolated before from *S. chrysospermum*. Additionally, the synthesis of selected peptides was accomplished by a solid-phase approach in combination with an in-solution cyclization, allowing to confirm the absolute configuration of their constitutive amino acids.

The occurrence of the anthraquinones physcion, emodin, endocrocin, dermolutein, hypericin, and skyrin were investigated by HPTLC-negative ion DESI-HRMS in crude extracts of six Chilean dermocyboid *Cortinarii* (Chapter 4). The implementation of fragmentation experiments (MS^2) for anthraquinones on HPTLC surfaces shows the potential of this technique as a valuable tool for the detection of this compound class. HPTLC reduces sample preparation steps, provides good separation efficiencies and can be performed in an automated and controlled way with respect to

the sample application and the development of the plate. These investigations illustrated the feasibility and capacity of HPTLC-DESI-HRMS as a rapid first screening method for the analysis of anthraquinones in complex mixtures, which may be used in the analysis of anthraquinones in food, plants, fungi, dyes, and cosmetic and pharmaceutical products.

A comprehensive discussion covers several areas and aspects of modern natural product research (Chapter 5). In summary, high-resolution mass spectrometry-based analysis, complemented with other analytic techniques, represents a sophisticated approach for the identification and characterization of natural products from plant and fungal crude extracts.

Zusammenfassung

Naturstoffe aus Pflanzen, Pilzen und Mikroorganismen bleiben eine der wichtigsten Quellen für neue bioaktive Verbindungen, die zum Beispiel für die Entwicklung von pharmazeutischen Produkten von großer Bedeutung sind. Heutzutage stellt die Anwendung und Kombination fortschrittlicher Analysemethoden in Verbindung mit anspruchsvollen Data-Mining-Ansätzen die Grundlage für die Erforschung und Entdeckung von Naturstoffen dar. Im Rahmen dieser Arbeit wurden effektive Arbeitsabläufe für die eingehende Untersuchung der Sekundärmetabolitenprofile von Pflanzen (Balsaminaceae), Mikro- (*Sepedonium*) und Makropilzen (*Cortinarius*) unter Verwendung fortschrittlicher analytischer Kopplungstechniken, im Speziellen UHPLC-ESI-HRMS und HPTLC-DESI-HRMS entwickelt. Zielgerichtete und ungerichtete (umfassende) Metaboliten Analyse-Ansätze wurden verwendet, um die chemische Vielfalt innerhalb verschiedener Spezies zur Dereplikation sowie Identifizierung vorhandener Metaboliten zu untersuchen. Zusätzlich wurden mögliche chemophenetische Marker für die Artendifferenzierung sowie das Vorhandensein neuer bioaktiver Metaboliten (unterstützt durch Bioaktivitätsscreening) untersucht.

Ein nicht-zielgerichteter Ansatz zur Untersuchung der intraspezifischen Varianz basierend auf DC-, UHPLC-ESI-HRMS- und NMR-Metabilotenprofilen von 31 *Impatiens*-Arten und der monophylletischen Schwesterart *Hydrocera triflora* (mit 19 phytochemisch nicht untersuchten Arten) zeigte eine hohe Variabilität der Sekundärmetaboliten in den Blattextrakten (Kapitel 2). Die Kombination von UHPLC-ESI-HRMS und NMR ermöglichte dabei ein vollständigeres Bild des Metaboloms. Erste phytochemische Untersuchungen von *I. racemosa* DC, *I. ethiopica* Grey-Wilson und *I. flanaganae* Hemsl führten zur Isolierung und Charakterisierung von Dihydrochalconderivaten, Astragalin und 2-Methoxynaphthochinon. Untersuchungen zur biologischen Aktivität zeigen das antibakterielle Potenzial verschiedener *Impatiens* Arten (wie z. B. *I. andringitrensis*, *I. arguta*, *I. balansae*) und *Hydrocera triflora*.

Mittels intensiver und zielgerichteter UHPLC-ESI-HRMS[n]-Untersuchungen wurden in dem angereicherten Rohextrakt von *Sepedonium microspermum* zwei neue zyklische Pentapeptide (Microsporide A and B) nachgewiesen und deren Primärstrukturen sequenziert (Kapitel 3). Mit Hilfe einer MS-geleiteten Isolierung konnte Microsporid A als Isomerengemisch und das bereits in der Literatur beschriebenede Chrysosporid gewonnen werden. Zusätzlich wurde die Synthese ausgewählter Peptide mittels Festphasenpeptidsynthese in Kombination mit einer Zyklisierung in Lösung etabliert, wodurch die absolute Konfiguration ihrer konstitutiven Aminosäuren bestätigt werden konnte. Hierbei konnte gezeigt werden, dass die Kombination aus fortschrittlichen Analysetechniken, Isolierung und Synthese ein leistungsfähiger Ansatz für die Identifizierung von Naturstoffen ist.

Rohextrakte von sechs chilenischen dermocyboiden *Cortinarii* wurden mittels HPTLC-Negativ-Ion DESI-HRMS auf das Auftreten der Anthrachinone Physcion, Emodin, Endocrocin, Dermolutein, Hypericin und Skyrin untersucht (Kapitel 4). Fragmentierungsexperimente (MS^2) für Anthrachinone auf HPTLC-Oberflächen zeigen das Potenzial dieser Technik als wertvolles Werkzeug für den Nachweis dieser Verbindungsklasse. HPTLC bietet ohne aufwendige Probenvorbereitung eine gute Trennleistung und kann in Bezug auf die Probenapplikation und die Entwicklung der Platte automatisiert und kontrolliert durchgeführt werden. Die entwickelte HPTLC-DESI-HRMS-Methode bildet ein schnelles erstes Screening-Verfahren für die Analyse von Anthrachinonen in komplexen Gemischen, welches in Zukunft bei der Analyse dieser Substanzklasse in Lebensmitteln, Pflanzen, Pilzen, Farbstoffen und Kosmetika angewendet werden kann.

Abschließend werden verschiedene Bereiche und Aspekte der modernen Naturstoffanalytik umfassend diskutiert (Kapitel 5). Dabei wird der wertvolle Beitrag der hochauflösenden massenspektrometrischen Analyse, ergänzt durch andere Analysetechniken, zur Identifizierung und Charakterisierung von Naturstoffen aus pflanzlichen und pilzlichen Rohextrakten herausgestellt.

Chapter 1 - General Introduction and Objectives

1.1 Nature - a unique source for today's drug discovery

Nature offers a vast, renewable, and highly diverse library of chemical compounds that allows science to move a step forward to a sustainable development and humankind wellbeing. In this sense, the exploration of the high chemical diversity and biochemical specificity underlying natural products (NPs) has contributed significantly to the health care for the treatment of diseases. A remarkable example of this statement arises from the fact that about 56% of all small-molecule drugs approved between 01/1981–09/2019 are NP-based or NP-inspired compounds.[1] This shows that NPs remain a major source of new chemical entities with potential applications for drug discovery, especially as anticancer drug or anti-infective agents to overcome existing resistances.[2-17]

1.2 The origin, structural diversity and function of natural products

When coming to natural sources, it is important to note that organisms produce numerous and diverse assortments of organic compounds. The great majority of them are specialized metabolites, oftentimes referred to as "secondary metabolites", that do not appear to participate directly in growth and development but are necessary for the survival of the producer organism in a given environment. In contrast, "primary metabolites", including amino acids, lipids, and carbohydrates, are essential to sustain life and physiology purposes.[18]

Whereas there is a clear difference regarding functions and roles that primary and secondary metabolites play in organisms, there is not a clear way to differentiate them on the basis of precursor molecules, chemical structures, or biosynthetic origins. Therefore, conventionally it is adopted a functional definition, with primary metabolites associated to nutritional and essential metabolic processes inside organisms, and secondary metabolites mediating ecological interactions between an organism and its environment.

The evolutionary origin of secondary metabolites can be rooted to genetic events, such as gene duplication, mutation, recombination, or even insertion of a transposable element.[19] On the other hand, its evolutionary change in a trait - whether the trait is maintained and regulated, or completely eliminated - can be attributed to biotic agents. While some variability in secondary metabolites composition might not play an adaptive role, much of the diversity of these natural compounds almost certainly results from both biotic and abiotic selection. Thus, ecological and physiological factors seems to have huge implications on secondary metabolites diversity.[20] In plants for example, nutritional factors,[21] light,[22,23] temperature,[24] and humidity[19] influence the production, distribution and structural diversity of phytochemicals.

Consequently, the structural diversity of secondary metabolites is the result of the continuous interaction of the producer organism with the high variability of its biotic and abiotic environment. Some of these interactions may be antagonistic, triggering the synthesis of chemical defenses.[25-27] While others, may be symbiotic or mutualistic, developing specific pools of chemical compounds to tackle specific kinds of interactions. The inherent necessity of nature for using structural diversity as key tool in these interplays should not be analyzed in trivial way, rather on a multifactorial scale: an organism has to struggle with a wide array of other organisms, those "contenders" are differentially adapted against chemical defenses, which mean that certain chemical compounds might be used as host-recognition cues. These "contenders" might play a differential role as agents of mortality in space and time. The presence of toxins in some tissues might be detrimental in mutualistic interactions. Because of the variability that exists in the interaction between organisms and their contenders, differential selection can lead to the evolution of a suite of chemicals within a species, at least in part due to synergistic or additive interactions among chemicals. Also, because of the variability that exists in the interaction between organisms, differential selection can lead to differentiation in chemistry between populations; and because of the differential effectiveness of these combinations of chemicals against individual "contenders", mechanisms allowing independent regulation (e.g., inducibility) of individual components should be operable, to allow for optimal biosynthesis of a suite of chemicals.[19,23] For that reason, at a certain time and population, the selection regime will be unique; in consequence, it is expectable that the associated metabolites profile will be as well unique, but at the same time highly diverse in a multiple zoom out view. Overall, the characteristic diversity of secondary metabolites and their functions can be explained according to the manner in which the pathways initially evolved and have subsequently been exploited.

1.3 Bioactivity of natural products and their pharmacological use

The use of natural sources for treating diseases have been around since the first human communities. Medicines based on what nature could offer for millennia have allowed the treatment of diverse conditions ranging from infections to gastric disorders to pain management to psychiatric syndromes. Most of the ancient therapeutics known relied on mixtures of plant products in different ways, sometimes consumed or applied directly, at times administrated after processing, for example through extraction of active ingredients using water (teas) or alcohol (tinctures), or even fermented. In a sense, one might say that the information compiled for the production of many of these earliest medicines set the ground for modern pharmacological knowledge.[28-30]

The first systematic efforts to isolate, purify and chemically analyze the medicinally active ingredients from plant-derived traditional medicines date from the 19th century. Representative examples of those endeavors are the opiate morphine, the anti-malarial quinine and the analgesic salicylic acid.[31] The 20th Century would complement plant-derived medicines with natural

products derived from microorganisms. The discovery of microbes and the ability to culture them in the lab, significantly expanded the known repertoire of natural products. Microbial natural products stand out for their effectiveness as antibiotics and antifungals, e.g., penicillin, amphotericin, etc., but also as anticancer drugs, e.g. doxorubicin, immune-suppressing agents, e.g. rapamycin, cholesterol lowering medicines, e.g. lovastatin, and antiparasitics, e.g., avermectin.[5,32]

For understanding why NPs are privileged for clinically useful bioactivity, once again it is necessary to spotlight their evolution through natural selection, aiming to provide an evolutive advantage to their producing organisms. This advantage is the result of effective engagement of the compound(s) with their cognate biologically relevant receptors or targets. Therefore, the biological activity of NPs is the result of its intrinsic physico-chemical characteristics. In comparison with chemically synthesized molecules, NPs deploy a set of enriched chemical features an increased number of sp^3-hybridized carbons and of chiral centres, fewer aromatic rings, larger macrocyclic aliphatic rings, lower nitrogen content and increased oxygen content. All these contribute to more complex three-dimensional structures. These characteristics elicit that NPs engage (more) effectively with biological targets then the usually more planar and stereochemically less complex structures prevailing in synthetic compound libraries.[10,17]
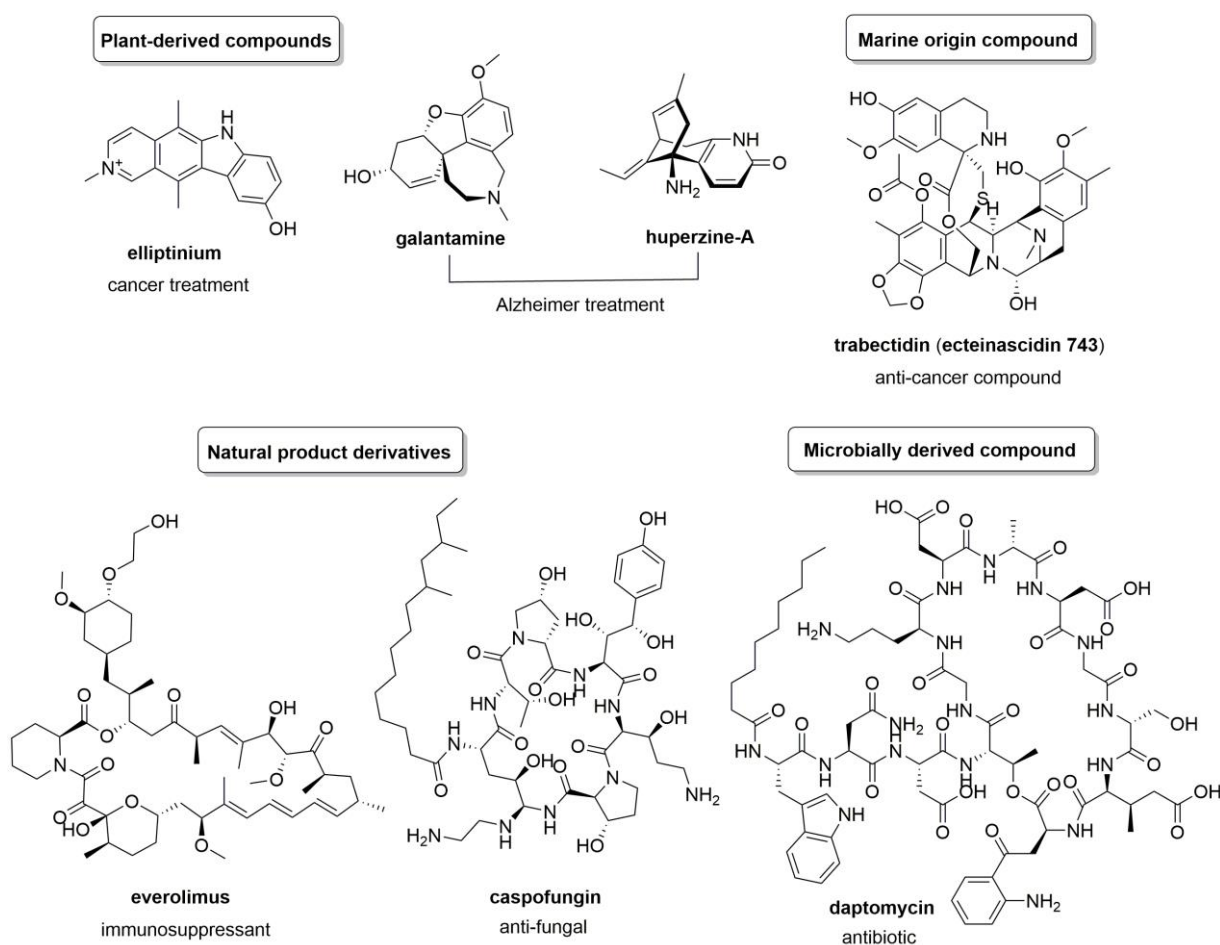
Figure 1.1. Some representative natural products/natural-products-inspired small-molecules and natural-product-based drugs.[1]

In consequence, NPs span a broad spectrum of biological activity covering anti-microbial,[9] anti-inflammatory,[33] anti-helminth,[34] anti-diabetic,[35] and anti-cancer[36] activity, just to mention some. Nowadays, it is well accepted among the scientific community that NPs are modern and effective sources for new drugs, and serve as the most important new drug discovery template reservoir.[13,37-42] In this sense, the NPs have also served as feasible starting materials to generate compounds with improved and new therapeutic efficacy. This includes stability to host biochemistry and targeting of specific tissues, and extension of the receptor range beyond the known biological functions of the starting material itself.[43-47]

The 21st century is witnessing a remarkable and progressive advance in natural-product-based drugs. In this context, there have been described compounds from diverse natural sources such as plants (including elliptinium, galantamine and huperzine), microbes (daptomycin) and animals (exenatide and ziconotide), as well as synthetic or semi-synthetic compounds based on natural products (e.g. tigecycline, everolimus, telithromycin, micafungin and caspofungin) (Figure 1.1).[1,38,48,49]

Recently, an astonishing number of natural-product-derived compounds are undergoing different stages of clinical development (Table 1.1). Although commonly the main focus on natural product projects is directed to cancer or anti-infective studies, there are many other therapeutic areas represented.[4,50] Another interesting approach explored within the last years comprises the use of mixtures of natural compounds from traditionally used medicines for the development of commercial products.[51]

Table 1.1. Natural-product-based drugs at different stages of development.

Origin	Plant	Bacterial	Fungal	Animal	Semi-synthetic	Other	Total[a]
2018	248	51	28	30	57	45	459
2019	266	43	27	35	59	69	499

Source: Pharmaprojects database (February 2019), [a] *Only compounds belonging to the* Top 25 origins of pipeline drugs *for the accounted period.*

So far, plant and microbial sources stand among the main leads in the development from natural products. Considering that relatively little of the world's plant biodiversity has been extensively screened for bioactivity and that very little of the estimated microbial biodiversity has been available for screening,[42,52,53] it is reasonable to forecast that in the upcoming years, researches in these areas will provide many novel chemicals for use in drug discovery assays. In addition, recent advances in genomics,[50,54] along with increasingly sophisticated gene mining[55-57] and manipulation,[58] amplified by synthetic biology strategies[59,60] might expand significantly the use of natural products in drug discovery.

1.4 Analytical techniques in natural product discovery

As discussed above, natural products (NPs) constitute a rich source of bioactive compounds. In consequence, the screening of NPs for the identification of new chemical entities plays a pivotal role in both, drug discovery and development (Figure 1.2). The workflows to obtain well-characterized bioactive NPs as hits and then as lead compounds might be extremely challenging regarding effort, time consumption and costs.[7,61]

The high complexity of the biological matrix (natural crude extracts), which is the starting point of natural product research, makes the work with NPs difficult. Each extract contains typically only a few main constituents and hundreds or thousands of less abundant ones. However, the lower concentrated compounds are often related to the bioactive secondary metabolites. To face these concentration differences and to enable the identification of the NPs, metabolite profiling methods are required, which cover a broad linear dynamic range and generate information rich spectra for their full or partial identification.[62]
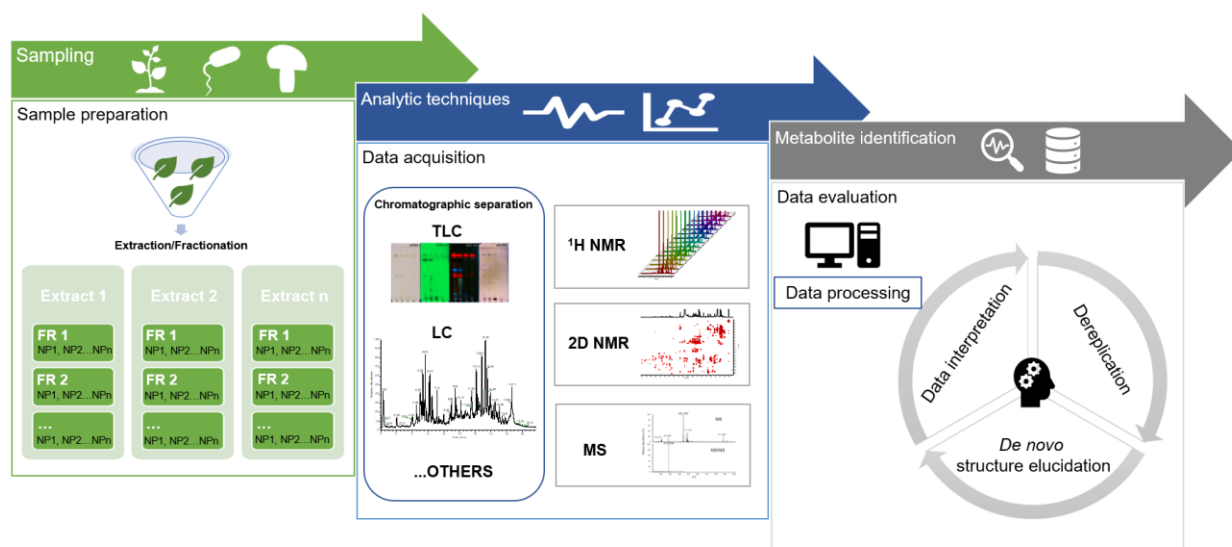
Figure 1.2. Overview of workflow and analytical techniques for natural product isolation and identification.

For the isolation and identification of a bioactive natural product from natural crude extracts (pharmacognosy) different approaches were developed in the last decades. They can be classified in five groups: bioactivity-guided fractionation, synergy-directed fractionation, metabolic profiling strategy, metabolism-directed (biotransformation focused) approach and finally direct phytochemical isolation.[17,28]

For the full characterization of the isolated metabolites a combination of different analytical methods, nuclear magnetic resonance spectroscopy (NMR), high-resolution mass spectrometry (HRMS), and chiroptical techniques (e.g. circular dichroism), are used to obtain their structure and the absolute configuration.[63] This process of natural product isolation and identification is often lengthy and cost-intensive. However, it has led to the major and important findings in natural

product research like the discovery of the antimalarial compound artemisinin by Professor Tu Youyou.[64]

To overcome these drawbacks, a key feature was the development and application of new, more efficient and sensitive analytical techniques, along with more effective high-throughput approaches for the identification of compounds with the desired biological effect within a complex matrix.[1] These metabolic profiling approaches allow dereplication – the differentiation of novel compounds from known ones – prior isolation.[65]

1.4.1 Metabolomics and metabolite profiling in natural product research

Besides metabolic profiling, metabolomics became an important tool in life science since the beginning of the 2000s.[66] In a general approximation, metabolomics might be defined as the scientific field in charge of studying the small molecules in a biological system. In a stricter sense, metabolomics is understood as an all-inclusive analytical approach for the identification and quantification of metabolites in a biological system, in a non-selective and broad scope way. With its universal approach, metabolomics provides a 'snap shot' of the metabolome of a given organism at a given time, aiming to achieve the complete set of metabolites, detect differences between them and generate hypotheses to explain these differences.[67,68]
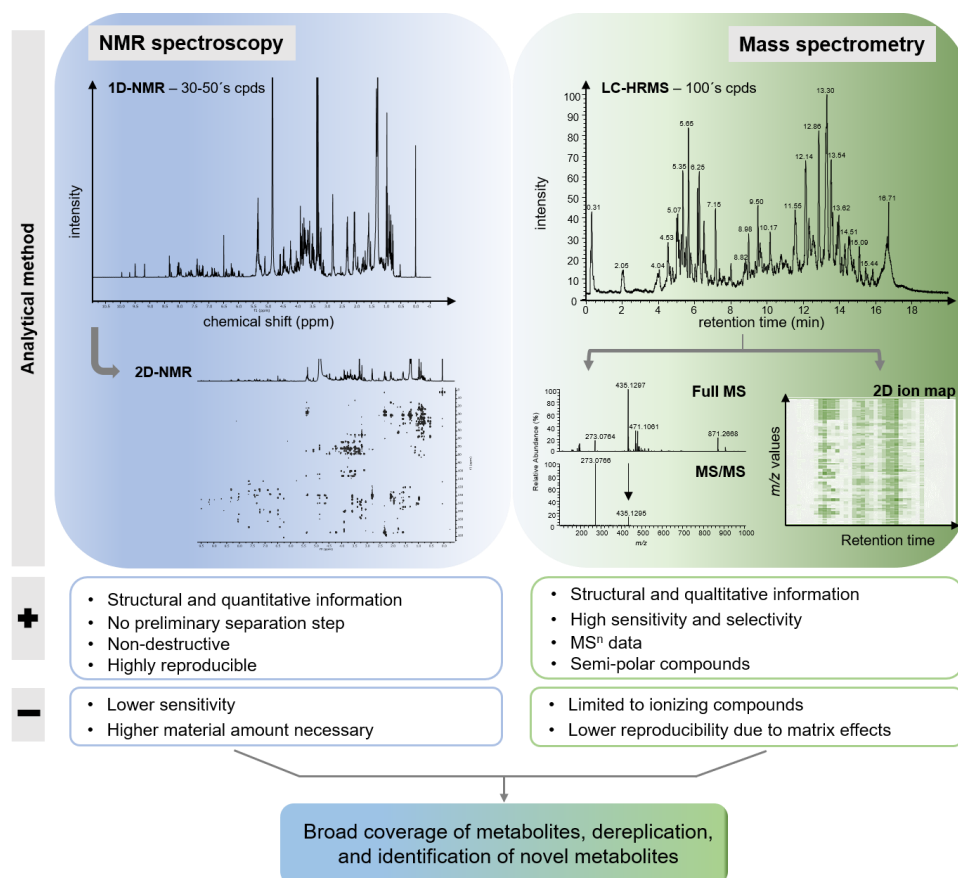
Figure 1.3. Combination of NMR and LC-HRMS for metabolite profiling or fingerprinting of crude extracts. Spectra shown in this figure are based on data acquired along this PhD project.

In metabolomics, the profiling of different constituents in complex biological matrices or extracts can be accomplished on several levels according the scope of the different techniques and concepts.[69] While "metabolite target analysis" directs the investigation to specific metabolites associated, e.g., with a particular pathway, "metabolite profiling" aims to analyze a larger group of metabolites that is either related to a specific metabolic pathway or a class of compounds.[70] On the other hand, "metabolite fingerprinting" is an untargeted approach that allows a rapid classification of samples; the aims of this methodology is, at least initially, not the identification and quantification of metabolites rather than to generate patterns or "finger-prints" of metabolites that change in a given biological system. In this last approach, the use of multivariate data analysis (MVDA) is applied to generate hypothesis based on samples difference and classification.[26]

In metabolite investigations (metabolomics) a diverse range of analytical techniques are used for the study of metabolites in complex biological samples, such as mass spectrometry,[71-75] NMR spectroscopy,[76-82] and infrared spectrophotometry,[83-86] among others.[87-91] Due to the complexity of biological systems, the intrinsic size of a given metabolome, as well as its chemical diversity, a single analytical technique is not sufficient for a comprehensive analysis of all metabolites in a biological model, and the utilization of multiple technologies is necessary. In order to expand the limits of metabolite detection and to assess the dereplication, new methods are being combined.[65,90,92] An example of this approach for metabolite profiling or fingerprinting is highlighted in Figure 1.3.

1.4.2   Mass spectrometry

Mass spectrometry (MS) is one of the most widely used analytical techniques for metabolite detection and identification.[93] Due to high speed, enhanced resolution (high-resolution MS), and greater sensitivity, it has become a method of choice for the rapid and sensitive determination of secondary metabolites, including those that appear at lower concentrations.[94-96] The rapid and continuous development of high-resolution mass spectrometer possessing Orbitrap and Time-of-Flight (TOF) analyzers broadened the applicability of MS towards broader linear dynamic ranges and the possibility to obtain information about elemental compositions of the analytes. The application of different ionization principles like electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), desorption electrospray ionization (DESI),[97] matrix assisted laser desorption/ionization (MALDI) using positive and negative ionization modes; the possibility of fragment generation (through e.g. collision-induced dissociation (CID), higher-energy collision induced dissociation (HCD), electron-transfer dissociation (ETD)); and different ion separation techniques display the versatility of mass spectrometry-based methods.[93,98] Furthermore, fragmentation reactions are an important tool for the structural elucidation of natural products.[99]

Compared with NMR, the sensitivity and resolution of MS is considerably higher, enabling the detection of a huge number of metabolites.[68] However, the main drawback associated to

qualitative and quantitative MS analysis is that detection is extremely dependent on the capability of metabolites to ionize under particular ionization conditions.[75] In order to enhance resolution and diminish the effect of the complex matrix on ionization, a common approach is the use of hyphenated separation technologies such as coupling the MS with chromatographic system like liquid (LC) or gas chromatography (GC).[100-102] The chromatographic separation of metabolites based on differences in their physicochemical properties allows a certain degree of fractionation that often is beneficial for the efficient detection, identification and quantification of the different metabolites in a complex mixture. Therefore, profiling methods should involve suitable chromatographic resolution to meet the standards for data acquisition and dereplication.[26]

1.4.3 Nuclear magnetic resonance spectroscopy (NMR)

Nuclear magnetic resonance spectroscopy (NMR) is commonly employed in metabolomics for catching an insight into the purity and molecular structure of metabolites under investigation.[78,91] Its scope of applications in metabolomics encompasses metabolite fingerprinting, profiling, natural product dereplication and metabolite flux analysis.[91,103-105] NMR allows to obtain spectral data in a quantitative and highly reproducible manner, offering an alternative for those metabolites that are difficult to ionize or would require derivatization for MS analysis.[105] Though, compared to MS it possesses low sensitivity and its capability for the investigation of a large amount of metabolites at low abundance is extremely limited.[106] Therefore, important metabolites such as biomarkers may remain undetected if their concentrations are too low in abundance, requiring larger number of scans and consequently longer measurement time or could also be resolved by employing higher magnetic fields.[91]

Hyphenated HPLC-NMR have been described for the rapid and detailed structural characterization of unknown mixtures. Despite the nowadays existing challenges regarding NMR data interpretation of complex mixtures or crude extracts analysis, various applications of HPLC-NMR and integrated HPLC-NMR-MS have been shown in drug discovery.[102,103,107,108]

1.5 Multivariate data analysis and chemometrics

Multivariate data analysis (MVDA) arises as a key tool for the visualization and interpretation of the large datasets generated with the modern analytical methods. MVDA approaches aim for reducing dataset dimensionality in order to gain an understanding of the dataset trends produced from tested samples and evaluate the significant variations that occur at the metabolome level.[109]

Among the different techniques employed for the conversion of highly multivariate data into compatible and interpretable sized data, principle component analysis (PCA) highlights as the most commonly used one. PCA is an unsupervised method mostly employed for a hypothesis-free explorative analysis.[110] This method allows to define datasets by converting a large number of correlated variables into smaller linear datasets.[111] The outcome of the PCA analysis is depicted

into a scores scatter plot showing patterns that are associated with fundamental arrangements present in the dataset divided into separate clusters by grouping the samples into specific metabolic phenotypes.[112] Overall, PCA allows to find patterns in the data matrix by producing clusters of columns (samples) or rows (features) that have comparable designs, enabling for group differentiation by identifying respective distinguishing metabolic features.[113]

In summary, metabolomics and metabolic profiling in combination with powerful data mining and data analysis strategies has shown great potential as a tool for evaluating new sources of bioactive natural substances using low amounts of material. Still, a large number of plants and fungi are not yet chemically investigated giving space for future research. Furthermore, the problem of matrix effects and method optimization for the optimal detection and characterization of a high variety of secondary metabolites within crude extracts remains challenging. The unambiguous identification of each individual natural product in an extract using generic methods is still challenging and needs further investigation to understand e.g. fragmentation pathways.

1.6 Aim of PhD work

The general objective of the present thesis was thus the in-depth investigation of the secondary metabolite profiles of plants and fungi using advanced analytical hyphenated techniques, in particular UHPLC-ESI-HRMS and HPTLC-DESI-HRMS, for the identification of new bioactive compounds.

In particular, the following aspects were covered within the investigations:

- Development and optimization of efficient workflows for high-resolution mass spectrometry-based metabolic profiling approaches
- Characterization and comparison of secondary metabolite profiles of plant and fungal crude extracts

- The interpretation of fragmentation pathways obtained from high-resolution mass spectrometry experiments for metabolite annotation and its use as template for natural product total synthesis

- Bioactivity screening of extracts and isolated compounds for the evaluation of their potential as possible sources of new bioactive natural products

- Isolation, characterization and structural elucidation of secondary metabolites from phytochemically not investigated species

- Application and evaluation of HPTLC-DESI-HRMS for the screening of anthraquinones from fungi as a tool for chemophenetic comparison

1.7 References

[1]   Newman, D.J.; Cragg, G.M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770-803.

[2]   Cowan, M.M. Plant products as antimicrobial agents. *Clin. Microbiol. Rev.* **1999**, *12*, 564-582.

[3]   Gullo, V.P.; McAlpine, J.; Lam, K.S.; Baker, D.; Petersen, F. Drug discovery from natural products. *J. Ind. Microbiol. Biotechnol.* **2006**, *33*, 523-531.

[4]   Harvey, A.L. Natural products as a screening resource. *Curr. Opin. Chem. Biol.* **2007**, *11*, 480-484.

[5]   Dias, D.A.; Urban, S.; Roessner, U. A historical overview of natural products in drug discovery. *Metabolites* **2012**, *2*, 303-336.

[6]   Savoia, D. Plant-derived antimicrobial compounds: alternatives to antibiotics. *Future Microbiol.* **2012**, *7*, 979-990.

[7]   David, B.; Wolfender, J.-L.; Dias, D.A. The pharmaceutical industry and natural products: historical status and new trends. *Phytochem. Rev.* **2015**, *14*, 299-315.

[8]   Barnes, E.C.; Kumar, R.; Davis, R.A. The use of isolated natural products as scaffolds for the generation of chemically diverse screening libraries for drug discovery. *Nat. Prod. Rep.* **2016**, *33*, 372-381.

[9]   Moloney, M.G. Natural products as a source for novel antibiotics. *Trends Pharmacol. Sci.* **2016**, *37*, 689-701.

[10]   Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on natural products for drug design. *Nat. Chem.* **2016**, *8*, 531-541.

[11]   Rossiter, S.E.; Fletcher, M.H.; Wuest, W.M. Natural products as platforms to overcome antibiotic resistance. *Chem. Rev.* **2017**, *117*, 12415-12474.

[12]   Tacconelli, E.; Carrara, E.; Savoldi, A.; Harbarth, S.; Mendelson, M.; Monnet, D.; Pulcini, C.; Kahlmeter, G.; Kluytmans, J.; Carmeli, Y., et al. Discovery, research, and development of new antibiotics: The WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* **2017**, *18*, 3218-327.

[13]   Thomford, N.E.; Senthebane, D.A.; Rowe, A.; Munro, D.; Seele, P.; Maroyi, A.; Dzobo, K. Natural Products for Drug Discovery in the 21st Century: Innovations for Novel Drug Discovery. *Int. J. Mol. Sci.* **2018**, *19*, 1-29.

[14]   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.

[15]   Wright, G.D. Unlocking the potential of natural products in drug discovery. *Microb. Biotechnol.* **2019**, *12*, 55-57.

[16]   Kuhlborn, J.; Gross, J.; Opatz, T. Making natural products from renewable feedstocks: back to the roots? *Nat. Prod. Rep.* **2020**, *37*, 380-424.

[17]   Lautié, E.; Russo, O.; Ducrot, P.; Boutin, J.A. Unraveling plant natural chemical diversity for drug discovery purposes. **2020**, *11*, 1-37.

[18]   Grotewold, E. Plant metabolic diversity: a regulatory perspective. *Trends Plant Sci.* **2005**, *10*, 57-62.

[19]   Berenbaum, M.R.; Zangerl, A.R. Phytochemical Diversity. In *Phytochemical Diversity and Redundancy in Ecological Interactions*, Romeo, J.T., Saunders, J.A., Barbosa, P., Eds. Springer US: Boston, MA, 1996; pp. 1-24, ISBN: 978-1-4899-1756-0.

[20]   Hathcock, J.N. Nutritional Toxicology. In *Nutritional Toxicology*, Hathcock, J.N., Ed. Academic Press: 1989; ISBN: 978-0-12-332603-4.

[21]   Slansky, F. Chapter 4 - Allelochemical–Nutrient Interactions in Herbivore Nutritional Ecology. In *Herbivores: Their Interactions with Secondary Plant Metabolites (Second Edition)*, Rosenthal, G.A., Berenbaum, M.R., Eds. Academic Press: San Diego, 1992; pp. 135-174, ISBN: 978-0-08-092545-5.

[22]   Downum, K.R. Tansley Review No. 43. Light-Activated Plant Defence. *New Phytol.* **1992**, *122*, 401-420.

[23]   Berenbaum, M. Phototoxicity of plant secondary metabolites: Insect and mammalian perspectives. *Arch. of Insect Biochem. Physiol.* **1995**, *29*, 119-134.

[24]   Wadleigh, R.W.; Koehler, P.G.; Preisler, H.K.; Patterson, R.S.; Robertson, J.L. Effect of temperature on the toxicities of ten pyrethroids to German cockroach (Dictyoptera: Blattellidae). *J. Econ. Entomol.* **1991**, *84*, 1433-1436.

[25] Creek, D.J.; Dunn, W.B.; Fiehn, O.; Griffin, J.L.; Hall, R.D.; Lei, Z.; Mistrik, R.; Neumann, S.; Schymanski, E.L.; Sumner, L.W., et al. Metabolite identification: are you sure? And how do your peers gauge your confidence? *Metabolomics* **2014**, *10*, 350-353.

[26] Wolfender, J.L.; Marti, G.; Thomas, A.; Bertrand, S. Current approaches and challenges for the metabolite profiling of complex natural extracts. *J. Chromatogr. A* **2015**, *1382*, 136-164.

[27] Nair, S.K.; Jez, J.M. Natural product biosynthesis: What's next? An introduction to the JBC Reviews Thematic Series. *J. Biol. Chem.* **2020**, *295*, 335-336.

[28] Atanasov, A.G.; Waltenberger, B.; Pferschy-Wenzig, E.-M.; Linder, T.; Wawrosch, C.; Uhrin, P.; Temml, V.; Wang, L.; Schwaiger, S.; Heiss, E.H., et al. Discovery and resupply of pharmacologically active plant-derived natural products: A review. *Biotechnol. Adv.* **2015**, *33*, 1582-1614.

[29] Heinrich, M. Ethnopharmacy and natural product research—Multidisciplinary opportunities for research in the metabolomic age. *Phytochem. Lett.* **2008**, *1*, 1-5.

[30] Swinney, D.C.; Anthony, J. How were new medicines discovered? *Nat. Rev. Drug Discov.* **2011**, *10*, 507-519.

[31] Desborough, M.J.R.; Keeling, D.M. The aspirin story - from willow to wonder drug. *Br. J. Haematol.* **2017**, *177*, 674-683.

[32] Walsh, C.T.; Tang, Y. *Natural Product Biosynthesis: Chemical Logic and Enzymatic Machinery*; Royal Society of Chemistry: London, 2017, ISBN: 978-1-78801-131-0.

[33] Soares-Bezerra, R.J.; Calheiros, A.S.; da Silva Ferreira, N.C.; da Silva Frutuoso, V.; Alves, L.A. Natural products as a source for new anti-Inflammatory and analgesic compounds through the inhibition of purinergic P2X receptors. *Pharmaceuticals (Basel)* **2013**, *6*, 650-658.

[34] Liu, M.; Panda, S.K.; Luyten, W. Plant-based natural products for the discovery and development of novel anthelmintics against nematodes. *Biomolecules* **2020**, *10*, 426-448.

[35] Zhu, J.; Huang, X.; Gao, H.; Bao, Q.; Zhao, Y.; Hu, J.-F.; Xia, G. A novel glucagon-like peptide 1 peptide identified from Ophisaurus harti. *J.Pept. Sci.* **2013**, *19*, 598-605.

[36] Kinghorn, A.D.; Chin, Y.-W.; Swanson, S.M. Discovery of natural product anticancer agents from biodiverse organisms. *Curr. Opin. Drug Discov. Devel.* **2009**, *12*, 189-196.

[37] von Nussbaum, F.; Brands, M.; Hinzen, B.; Weigand, S.; Häbich, D. Antibacterial natural products in medicinal chemistry--exodus or revival? *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 5072-5129.

[38] Harvey, A.L. Natural products in drug discovery. *Drug Discov. Today* **2008**, *13*, 894-901.

[39] Hagino, N. West meets east-looking for the interphase of Western medicine and traditional oriental medicine in future. *Rinsho Byori* **2000**, *48*, 764-770.

[40] Donia, M.S.; Fricke, W.F.; Ravel, J.; Schmidt, E.W. Variation in tropical reef symbiont metagenomes defined by secondary metabolism. *PLOS ONE* **2011**, *6*, e17897.

[41] Shen, B. A New golden age of natural products drug discovery. *Cell* **2015**, *163*, 1297-1300.

[42] Katz, L.; Baltz, R.H. Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.* **2016**, *43*, 155-176.

[43] DeCorte, B.L. Underexplored opportunities for natural products in drug discovery. *J. Med. Chem.* **2016**, *59*, 9295-9304.

[44] Banwell, M. Research in natural product synthesis: a vital and dynamic global enterprise - Preface. *Tetrahedron* **2008**, *64*, 4669-4670.

[45] Galloway, W.R.J.D.; Bender, A.; Welch, M.; Spring, D.R. The discovery of antibacterial agents using diversity-oriented synthesis. *Chem. Commun.* **2009**, 2446-2462.

[46] Kombarov, R.; Altieri, A.; Genis, D.; Kirpichenok, M.; Kochubey, V.; Rakitina, N.; Titarenko, Z. BioCores: identification of a drug/natural product-based privileged structural motif for small-molecule lead discovery. *Mol. Div.* **2010**, *14*, 193-200.

[47] Tanaka, H.; Yoshizawa, A.; Chijiwa, S.; Ueda, J.-y.; Takagi, M.; Shin-ya, K.; Takahashi, T. Efficient synthesis of the deoxysugar part of versipelostatin by direct and stereoselective glycosylation and revision of the structure of the trisaccharide unit. *Chem. Asian J.* **2009**, *4*, 1114-1125.

[48] Butler, M.S. Natural products to drugs: natural product-derived compounds in clinical trials. *Nat. Prod. Rep.* **2008**, *25*, 475-516.

[49] Lam, K.S. New aspects of natural products in drug discovery. *Trends Microbiol.* **2007**, *15*, 279-289.

[50] Harvey, A.L.; Edrada-Ebel, R.; Quinn, R.J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug. Discov.* **2015**, *14*, 111-129.

[51] Alamgir, A.N.M. *Therapeutic Use of Medicinal Plants and Their Extracts: Volume 1*; 2017, ISBN: 978-3-319-63862-1.

[52] Harvey, A. Strategies for discovering drugs from previously unexplored natural products. *Drug Discov. Today* **2000**, *5*, 294-300.

[53] Baker, D.D.; Chu, M.; Oza, U.; Rajgarhia, V. The value of natural products to future pharmaceutical discovery. *Nat. Prod. Rep.* **2007**, *24*, 1225-1244.

[54] Blin, K.; Shaw, S.; Kautsar, S.A.; Medema, M.H.; Weber, T. The antiSMASH database version 3: increased taxonomic coverage and new query features for modular enzymes. *Nucleic Acids Res.* **2020**, *8*, 639-643.

[55] Pimentel-Elardo, S.M.; Sørensen, D.; Ho, L.; Ziko, M.; Bueler, S.A.; Lu, S.; Tao, J.; Moser, A.; Lee, R.; Agard, D., et al. Activity-independent discovery of secondary metabolites using chemical elicitation and cheminformatic inference. *ACS Chem. Biol.* **2015**, *10*, 2616-2623.

[56] Mohimani, H.; Pevzner, P.A. Dereplication, sequencing and identification of peptidic natural products: from genome mining to peptidogenomics to spectral networks. *Nat. Prod. Rep.* **2016**, *33*, 73-86.

[57] Xu, M.; Wright, G.D. Heterologous expression-facilitated natural products' discovery in actinomycetes. *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 415-431.

[58] Chen, J.; Wu, Q.; Hawas, U.W.; Wang, H. Genetic regulation and manipulation for natural product discovery. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 2953-2965.

[59] Wright, G. Perspective: Synthetic biology revives antibiotics. *Nature* **2014**, *509*, S13.

[60] Smanski, M.J.; Zhou, H.; Claesen, J.; Shen, B.; Fischbach, M.A.; Voigt, C.A. Synthetic biology to access and expand nature's chemical diversity. *Nat. Rev. Microbiol.* **2016**, *14*, 135-149.

[61] Koehn, F.E.; Carter, G.T. The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discov.* **2005**, *4*, 206-220.

[62] van der Hooft, J.J.J.; de Vos, R.C.H.; Ridder, L.; Vervoort, J.; Bino, R.J. Structural elucidation of low abundant metabolites in complex sample matrices. *Metabolomics* **2013**, *9*, 1009-1018.

[63] Hostettmann, K.; Wolfender, J.L.; Terreaux, C. Modern screening techniques for plant extracts. *Pharm. Biol.* **2001**, *39*, 18-32.

[64] Van Voorhis, W.C.; Hooft van Huijsduijnen, R.; Wells, T.N. Profile of William C. Campbell, Satoshi Ōmura, and Youyou Tu, 2015 Nobel Laureates in Physiology or Medicine. *Proc. Nat. Acad. Sci. U.S.A* **2015**, *112*, 15773-15776.

[65] Wolfender, J.L.; Nuzillard, J.M.; van der Hooft, J.J.J.; Renault, J.H.; Bertrand, S. Accelerating metabolite identification in natural product research: toward an ideal combination of liquid chromatography-high-resolution tandem mass spectrometry and NMR profiling, *in silico* databases, and chemometrics. *Anal. Chem.* **2019**, *91*, 704-742.

[66] Guy, C.; Kopka, J.; Moritz, T. Plant metabolomics coming of age. *Physiol. Plant* **2008**, *132*, 113-116.

[67] Wishart, D.S. Applications of metabolomics in drug discovery and development. *Drugs R D* **2008**, *9*, 307-322.

[68] Kosmides, A.K.; Kamisoglu, K.; Calvano, S.E.; Corbett, S.A.; Androulakis, I.P. Metabolomic fingerprinting: challenges and opportunities. *Crit. Rev. Biomed. Eng.* **2013**, *41*, 205-221.

[69] Fiehn, O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp. Funct. Genomics* **2001**, *2*, 155-168.

[70] Wolfender, J.-L.; Glauser, G.; Boccard, J.; Rudaz, S. MS-based Plant Metabolomic Approaches for Biomarker Discovery. *Nat. Prod. Commun.* **2009**, *4*, 1417-1430.

[71] Steinmann, D.; Ganzera, M. Recent advances on HPLC/MS in medicinal plant analysis. *J. Pharm. Biomed.* **2011**, *55*, 744-757.

[72] Krug, D.; Müller, R. Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. *Nat. Prod. Rep.* **2014**, *31*, 768-783.

[73] Ernst, M.; Silva, D.B.; Silva, R.R.; Vêncio, R.Z.N.; Lopes, N.P. Mass spectrometry in plant metabolomics strategies: from analytical platforms to data acquisition and processing. *Nat. Prod. Rep.* **2014**, *31*, 784-806.

[74] Jorge, T.F.; Rodrigues, J.A.; Caldana, C.; Schmidt, R.; van Dongen, J.T.; Thomas-Oates, J.; António, C. Mass spectrometry-based plant metabolomics: Metabolite responses to abiotic stress. *Mass Spectrom. Rev.* **2016**, *35*, 620-649.

[75]   Wang, S.; Blair, I.A.; Mesaros, C. Analytical Methods for Mass Spectrometry-Based Metabolomics Studies. In *Advancements of Mass Spectrometry in Biomedical Research*, Woods, A.G., Darie, C.C., Eds. Springer International Publishing: Cham, 2019; pp. 635-647, ISBN: 978-3-030-15950-4.

[76]   Breton, R.C.; Reynolds, W.F. Using NMR to identify and characterize natural products. *Nat. Prod. Rep.* **2013**, *30*, 501-524.

[77]   Kim, H.K.; Choi, Y.H.; Verpoorte, R. NMR-based metabolomic analysis of plants. *Nat. Protoc.* **2010**, *5*, 536-549.

[78]   Larive, C.K.; Barding, G.A., Jr.; Dinges, M.M. NMR spectroscopy for metabolomics and metabolic profiling. *Anal. Chem.* **2015**, *87*, 133-146.

[79]   Mahrous, E.A.; Farag, M.A. Two dimensional NMR spectroscopic approaches for exploring plant metabolome: A review. *J. Adv. Res.* **2015**, *6*, 3-15.

[80]   Kumar, D. Nuclear magnetic resonance (NMR) spectroscopy for metabolic profiling of medicinal plants and their products. *Crit. Rev. Anal. Chem.* **2016**, *46*, 400-412.

[81]   Deborde, C.; Moing, A.; Roch, L.; Jacob, D.; Rolin, D.; Giraudeau, P. Plant metabolism as studied by NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *102-103*, 61-97.

[82]   Emwas, A.-H.; Roy, R.; McKay, R.T.; Tenori, L.; Saccenti, E.; Gowda, G.A.N.; Raftery, D.; Alahmari, F.; Jaremko, L.; Jaremko, M., et al. NMR spectroscopy for metabolomics research. *Metabolites* **2019**, *9*, 1-39.

[83]   Volpe, M.G.; Costantini, S.; Coccia, E.; Parrillo, L.; Paolucci, M. Evaluation of metabolic changes induced by polyphenols in the crayfish Astacus leptodactylus by metabolomics using Fourier transformed infrared spectroscopy. *J. Biosci.* **2018**, *43*, 585-596.

[84]   Dai, S.; Lin, Z.; Xu, B.; Wang, Y.; Shi, X.; Qiao, Y.; Zhang, J. Metabolomics data fusion between near infrared spectroscopy and high-resolution mass spectrometry: A synergetic approach to boost performance or induce confusion. *Talanta* **2018**, *189*, 641-648.

[85]   Pilatti, F.K.; Ramlov, F.; Schmidt, E.C.; Costa, C.; Oliveira, E.R.d.; Bauer, C.M.; Rocha, M.; Bouzon, Z.L.; Maraschin, M. Metabolomics of *Ulva lactuca* Linnaeus (Chlorophyta) exposed to oil fuels: Fourier transform infrared spectroscopy and multivariate analysis as tools for metabolic fingerprint. *Mar. Pollut. Bull.* **2017**, *114*, 831-836.

[86]   Kuligowski, J.; Pérez-Guaita, D.; Escobar, J.; Lliso, I.; de la Guardia, M.; Lendl, B.; Vento, M.; Quintás, G. Infrared biospectroscopy for a fast qualitative evaluation of sample preparation in metabolomics. *Talanta* **2014**, *127*, 181-190.

[87]   Eugster, P.J.; Guillarme, D.; Rudaz, S.; Veuthey, J.L.; Carrupt, P.A.; Wolfender, J.L. Ultra high pressure liquid chromatography for crude plant extract profiling. *J. AOAC Int.* **2011**, *94*, 51-70.

[88]   Potterat, O.; Hamburger, M. Concepts and technologies for tracking bioactive compounds in natural product extracts: generation of libraries, and hyphenation of analytical processes with bioassays. *Nat. Prod. Rep.* **2013**, *30*, 546-564.

[89]   Alonso, A.; Marsal, S.; Julià, A. Analytical methods in untargeted metabolomics: state of the art in 2015. *Front. Bioeng. Biotechnol.* **2015**, *3*, 23.

[90]   Wolfender, J.L.; Marti, G.; Queiroz, E.F. Advances in techniques for profiling crude extracts and for the rapid identification of natural products: dereplication, quality control and metabolomics. *Curr. Org. Chem.* **2010**, *14*, 1808-1832.

[91]   Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X. Modern analytical techniques in metabolomics analysis. *Analyst* **2012**, *137*, 293-300.

[92]   Marshall, D.D.; Powers, R. Beyond the paradigm: Combining mass spectrometry and nuclear magnetic resonance for metabolomics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *100*, 1-16.

[93]   Bouslimani, A.; Sanchez, L.M.; Garg, N.; Dorrestein, P.C. Mass spectrometry of natural products: current, emerging and future technologies. *Nat. Prod. Rep.* **2014**, *31*, 718-729.

[94]   Gowda, G.A.N.; Djukovic, D. Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges. In *Mass Spectrometry in Metabolomics: Methods and Protocols*, Raftery, D., Ed. Springer New York: New York, NY, 2014; pp. 3-12, ISBN: 978-1-4939-1258-2.

[95]   Kachlicki, P.; Piasecka, A.; Stobiecki, M.; Marczak, Ł. Structural characterization of flavonoid glycoconjugates and their derivatives with mass spectrometric techniques. *Molecules* **2016**, *21*, 1494-1515.

[96]   Alvarez-Rivera, G.; Ballesteros-Vivas, D.; Parada-Alfonso, F.; Ibañez, E.; Cifuentes, A. Recent applications of high resolution mass spectrometry for the characterization of plant natural products. *Trends Anal. Chem.* **2019**, *112*, 87-101.

[97] Takats, Z.; Wiseman, J.M.; Cooks, R.G. Ambient mass spectrometry using desorption electrospray ionization (DESI): instrumentation, mechanisms and applications in forensics, chemistry, and biology. *J. Mass Spectrom.* **2005**, *40*, 1261-1275.

[98] Demarque, D.P.; Crotti, A.E.; Vessecchi, R.; Lopes, J.L.; Lopes, N.P. Fragmentation reactions using electrospray ionization mass spectrometry: an important tool for the structural elucidation and characterization of synthetic and natural products. *Nat. Prod. Rep.* **2016**, *33*, 432-455.

[99] Dührkop, K.; Nothias, L.-F.; Fleischauer, M.; Reher, R.; Ludwig, M.; Hoffmann, M.A.; Petras, D.; Gerwick, W.H.; Rousu, J.; Dorrestein, P.C., et al. Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. *Nat. Biotechnol.* **2020**, *39*, 462-471.

[100] Wolfender, J.-L.; Queiroz, E.F.; Hostettmann, K. The importance of hyphenated techniques in the discovery of new lead compounds from nature. *Expert Opin. Drug Discov.* **2006**, *1*, 237-260.

[101] Sarker, S.D.; Nahar, L. Hyphenated techniques and their applications in natural products analysis. In *Natural Products Isolation*, Sarker, S.D., Nahar, L., Eds. Humana Press: Totowa, NJ, 2012; pp. 301-340, ISBN: 978-1-61779-624-1.

[102] Usman, R.; Badgujar, S.R.; Shaikh, T.Y. Hyphenated techniques of drug analysis. *Sch. Acad. J. Pharm.* **2017**, *6*, 263-272.

[103] Jaroszewski, J.W. Hyphenated NMR methods in natural products research, Part 2: HPLC-SPE-NMR and other new trends in NMR hyphenation. *Planta Med.* **2005**, *71*, 795-802.

[104] Seger, C.; Sturm, S.; Stuppner, H. Mass spectrometry and NMR spectroscopy: modern high-end detectors for high resolution separation techniques – state of the art in natural product HPLC-MS, HPLC-NMR, and CE-MS hyphenations. *Nat. Prod. Rep.* **2013**, *30*, 970-987.

[105] Markley, J.L.; Brüschweiler, R.; Edison, A.S.; Eghbalnia, H.R.; Powers, R.; Raftery, D.; Wishart, D.S. The future of NMR-based metabolomics. *Curr. Opin. Biotechnol.* **2017**, *43*, 34-40.

[106] Klassen, A.; Faccio, A.T.; Canuto, G.A.B.; da Cruz, P.L.R.; Ribeiro, H.C.; Tavares, M.F.M.; Sussulini, A. Metabolomics: Definitions and significance in systems biology. In *Metabolomics: From Fundamentals to Clinical Applications*, Sussulini, A., Ed. Springer International Publishing: Cham, 2017; pp. 3-17, ISBN: 978-3-319-47656-8.

[107] Shockcor, J.P.; Unger, S.E.; Wilson, I.D.; Foxall, P.J.; Nicholson, J.K.; Lindon, J.C. Combined HPLC, NMR spectroscopy, and ion-trap mass spectrometry with application to the detection and characterization of xenobiotic and endogenous metabolites in human urine. *Anal. Chem.* **1996**, *68*, 4431-4435.

[108] Wolfender, J.L.; Rodriguez, S.; Hostettmann, K. Liquid chromatography coupled to mass spectrometry and nuclear magnetic resonance spectroscopy for the screening of plant constituents. *J. Chromat. A* **1998**, *794*, 299-316.

[109] Worley, B.; Powers, R. Multivariate analysis in metabolomics. *Curr. Metabolomics* **2012**, *1*, 92-107.

[110] Ramadan, Z.; Jacobs, D.; Grigorov, M.; Kochhar, S. Metabolic profiling using principal component analysis, discriminant partial least squares, and genetic algorithms. *Talanta* **2006**, *68*, 1683-1691.

[111] Kemsley, E.K.; Le Gall, G.; Dainty, J.R.; Watson, A.D.; Harvey, L.J.; Tapp, H.S.; Colquhoun, I.J. Multivariate techniques and their application in nutrition: a metabolomics case study. *Br. J. Nutr.* **2007**, *98*, 1-14.

[112] Madala, N.E.; Piater, L.A.; Steenkamp, P.A.; Dubery, I.A. Multivariate statistical models of metabolomic data reveals different metabolite distribution patterns in isonitrosoacetophenone-elicited *Nicotiana tabacum* and *Sorghum bicolor* cells. *Springerplus* **2014**, *3*, 254.

[113] Chanana, S.; Thomas, C.S.; Braun, D.R.; Hou, Y.; Wyche, T.P.; Bugni, T.S. Natural product discovery using planes of principal component analysis in R (PoPCAR). *Metabolites* **2017**, *7*, 34-46.

Chapter 2

-

Exploration and Comparison of the Secondary Metabolic Profiles of *Impatiens* species and *Hydocera triflora* (Balsaminaceae) Leaf Extracts using non-targeted Analytical Approaches

*Abstract**

The genus *Impatiens* is one of the largest genera in angiosperms with more than 1000 species. Only a few have been characterized in terms of their phytochemical and pharmacological properties. In the present study, 31 *Impatiens* species and *Hydrocera triflora*, grown in the botanical garden of Bonn, were investigated and compared regarding their metabolic profiles using a non-targeted LC-ESI-HRMS and ^{1}H-NMR profiling approach. The results revealed the presence of a great variety of natural products, especially flavonoid glycosides, hydroxycinnamic acid derivatives, coumarins, naphthalene, and naphthoquinone conjugates within the investigated species. The comparison and chemometric analysis of the data indicated that flavonoid glycosides are the major compound class present in *Impatiens* species and its monotypic sister *Hydrocera triflora*. Possible chemophenetic marker compounds like coumarins, dihydrochalcones, xanthones or napthoquinones detected in this study only appear in some species suggesting a more detailed analysis of the genetic and chemical relationsships towards a subdivision of the genus *Impatiens*. Furthermore, a biological activity screening of the investigated species showed their potential as source for new antibacterial natural products. The detailed phytochemical investigation of *Impatiens racemosa*, *I. ethiopica* and *I. flanaganae* revealed the occurrence of dihydrochalcones in *Impatiens* species and confirmed the presence as well as the antibacterial properties of 2-methoxy-1,4-naphthoquinone. To the best of our knowledge, this is the most comprehensive study on the secondary metabolite profiles and antibacterial properties of the leaf extracts of a high number of *Impatiens* species as well as the first report on the phytochemical composition of *Hydrocera triflora* and 19 *Impatiens* species.

2.1 Introduction

The family of Balsaminaceae consists of two genera, the genus *Impatiens* with more than 1000 species displaying one of the largest genera in angiosperms, and its monotypic sister *Hydrocera*.[1-6] Whereas members of the genus *Impatiens* are geographically widely distributed – mainly in tropical, subtropical and temperate climate zones, in Asia, Africa, Europe, Central and North America[7-11] –, *Hydrocera triflora* Weight-Arn, the single member of *Hydrocera*, its sibling genus, grow in the Indo-Malaysian Tropical region.[1,9] The monotypic *H. triflora* is a semiaquatic plant possessing the distinct characteristic of having dark-red, fleshy, and floatable berries. The genus *Hydrocera* is a perennial herb with rose flowers possessing free petals and sepals, which differentiate it from *Impatiens* species (see Figure 2.1 A). *Hydrocera* has been poorly investigated, and reports on it are merely morphological.[1,4]

Figure 2.1. Pictures of and *Hydrocera triflora* (L.) Wight & Arn. (A), *Impatiens noli-tangere* L. (B), *Impatiens niamniamensis* Gilg. emdend G. M. Schulze (C). (Copyright Dr. Stefan Abrahamczyk)

The scientific name of the genus *Impatiens* (lat. *impatiens* = impatient) is based on the small, green, and juicy capsules, which explode at seed ripeness already at the slightest touch and sometimes can be thrown sometimes far meters. Therefore, these plants are also known as "touch-me-nots".[12] *Impatiens* species are predominantly annual or perennial herbs, rarely single woody plants growing from sea level to 4000 m altitude. They are often found in forest margins, in valleys, roadside troughs, and along rivers on humid soils.[5,13]

The majority of the species – e.g. 91% of the southern Indian species, and almost all species of Madagascar – are locally endemic.[2,14,15] They possess soft, semi-succulent stems and fleshy undivided, often serrated or lobed leaves.[7] Their fragile and zygomorphic flowers exhibit a large diversity in corolla color and morphology (Figure 2.1 B, C). Therefore, the genus *Impatiens* has been regarded as "dicot counterpart of the orchid" and is of interest to horticulturalists.[9] However, its ornamental use sometimes presents detrimental effects since some species become invasive weeds (neophytes), e.g. *I. glandulifera* Royle, *I. capensis* Meerb., *I. parviflora* DC., provoking negative impact on native species diversity and habitats.[16-18]

Interestingly, some *Impatiens* species, as in the case of *I. balsamina* L. and *I. textori* Miq., possess nutritional and pharmaceutical properties, and are used in traditional folk medicine for antimicrobial, antirheumatic, antipruritic, and antitumoural purposes, as well as for the treatment

of difficult labour, superficial infections, fingernail inflammation, detoxification, carbuncles bruises, beri beri, and puerperal pain.[19-25]

The taxonomic differentiation of the genus *Impatiens* is difficult due to the heterogeneity of their morphological characters.[7,9,26,27] However, molecular phylogenetic data has contributed greatly to understand the relationships within the genus *Impatiens*.[3,9,28-33] In 2016, Yu et al.[5] provided the most comprehensive phylogenetic study up to date. Yu´s work describes the phylogenetic investigation of 150 *Impatiens* species, combining forty-six morphological characters and the phylogenetic relationships using three genetic regions, including nuclear ribosomal internal transcribed spacer (ITS) and plastid *atpB-rbcL* and *trnL-F*. The outcome of this study was that *Impatiens* splits into two major clades. Species with three-colpate pollen and four carpels form a monophyletic group (clade I), while clade II shows seven well-supported subclades (A-G) confirming the previously reported data.[5]

Besides their morphological and genetic complexity, *Impatiens* species also exhibit a wide and diverse phytochemical profile. Different approaches to investigate the metabolic profile of different *Impatiens* species have been done, including the targeted profiling towards their flavonoid and phenolic compounds,[34-36] as well as one untargeted metabolic analysis of different leaf and stem extracts of *I. balsamina*.[37] Futhermore, several efforts for the isolation of secondary metabolites from different *Impatiens* species has resulted in the identification of phenolic compounds (e.g. flavonoids[20,38-43], phenols[34,44,45], and coumarins[46,47]), alkaloids[48], quinones (e.g. naphthoquinones[49-55] and dinaphthofuran-7,12-dione derivatives[56,57]), naphtalenes[58,59], as well as triterpenes (e.g. triterpenoids[60,61], steroids[62,63] and saponins[25,64-66]), polysaccharides[67], and fatty acids.[68] The phytochemistry of *Impatiens* species was reviewed in 2018.[69]

Impatiens species exhibit a broad spectrum of biological activity. Several species show a high antioxidant and antimicrobial potential due to their broad spectrum of flavonoid and polyphenol derivatives.[34,40,42,43,70,71] Additionally, antifungal[72,73], antianaphylatic, and anti-tumor activity against HepG2 cells[52] are associated with naphthoquinones isolated from flowers and leaves of *I. balsamina*, e.g. 2-methoxy-1,4-naphthoquinone (2-MNQ) and 2-hydroxy-1,4-naphthoquinone (lawsone).

Relevant attention has to be drawn to the bioactivity investigation of crude extracts from different *Impatiens*.[21,70,74-84] In this direction, extracts of the aerial parts of *I. balfourii*, *I. balsamina*, *I. glandulifera*, and *I. noli-tangere* were tested *in vitro* for their antibacterial and antifungal activities, showing potential activity against gram-positive bacteria.[34]

2.2 Aims and Scope

Due to the lack of information on the phytochemical composition of a great number of species belonging to Balsaminaceae family, and the possible use of specific compounds as chemophenetic markers being useful for species delimitation, there is a necessity for further investigations. Additionally, the search for novel antiinfective natural compounds is still an important research area in the pursuit for new effective drugs. The aim of the present research work was to perform an exploratory investigation for the comparison of the secondary metabolite profiles of leaf extracts from 31 *Impatiens* species and *Hydrocera triflora* based on untargeted LC-ESI-HRMS and ^{1}H-NMR approaches. Additionally, a bioactivity screening should lead to *Impatiens* species that could be used for further phytochemical investigations towards the isolation of new bioactive natural products.

2.3 Results and Discussion

Leaves of 31 *Impatiens* species as well as from *Hydrocera triflora* were collected in 2015 in the Botanical Garden in Bonn by Dr. S. Abrahamczyk including two new species, *I.* species 34628 and *I.* species 34558 from Africa (Table 2.1). The lyophilized material was ground using a ball mill to ensure a homogeneous powder. Different extraction solvents (methanol, ethanol, 80% aqueous methanol, acetone, *n*-hexane) were tested in order to assess the most suitable one for a maximum metabolite extraction. After comparison of the different extracts, by thin-layer chromatography (TLC), LC-MS and ^{1}H-NMR, methanol was chosen as optimal solvent providing the broadest spectrum of metabolite extraction.

In the first stage, TLC was used in order to have an overview of the leaf extracts. For a deeper analysis of the metabolite profiles, an untargeted analysis using UHPLC-ESI-HRMS was performed to obtain specific fingerprints of the selected species. Multivariate data analysis (PCA) was applied for comparison of the obtained profiles and to find unique compounds within the species. The compounds were manually annotated supported by the comparison of known reference spectra using databases and software tools like Reaxys, massbank.eu and Metfrag. The crude extracts were further analyzed by ^{1}H-NMR in order to overcome the limitations regarding non-ionizable metabolites present in the complex matrices.

In addition, a biological screening against the gram-negative bacterium *Aliivibrio fischeri,* the gram-positive *Bacillus subtilis,* the nematode *Caenorhabditis elegans* and the cancer cell lines HT-29 (human colon cancer) and PC3 (human prostate cancer) was conducted. Due to its special metabolite composition, the leaf extract of *I. racemosa* DC., was further chromatographically separated to obtain the main constituent. Furthermore, the methanolic leaf extracts of *Impatiens ethiopica* Grey-Wilson and *I. flanaganae* Hemsl. were phytochemically investigated to characterize those compounds, which caused the detected high antibacterial activity within the performed bioactivity screenings.

Table 2.1. Sample information of *Impatiens* and *Hydrocera* leaf samples.

Species	Code	Herbarium voucher	Clade/ Subclade*	Section	Ancestral origin[⊥]
Impatiens andringitrensis H. Perrier	AND	xx-0-BONN-36655	II / G	Uniflorae	Madagascar
Impatiens arguta Hook. fil. & Thomson	ARG	xx-0-BONN-28423	II / C	Fasciculatae	Asia
Impatiens balansae Hook. fil.	BLN	VN-0-BONN-36588	I / none	Clavicarpa	Asia
Impatiens balfourii Hook. fil.	BLF	xx-0-BONN-36126	II / B	n.d.	Asia
Impatiens balsamina L.	BLS	xx-0-FRT-0000/1958	II / G	Uniflorae	Asia
Impatiens bicolor Royle	BCL	xx-0-BONN-36979	n.d.	n.d.	Asia
Impatiens bisaccata Warb.	BSC	MG-0-BONN-36496	n.d.	n.d.	Madagascar
Impatiens burtonii Hook.fil.	BRT	RW-0-BONN-11528	II / G	Uniflorae	Africa
Impatiens capensis Meerb.	CPN	DE-0-BONN-35960	II / D	Impatiens	North America
Impatiens edgeworthii Hook. f.	EDG	37204 (ID 11060, from Berlin)	II / B	n.d.	Asia
Impatiens ethiopica Grey-Wilson	ETH	xx-0-BONN-36142	n.d.	n.d.	Africa
Impatiens flanaganae Hemsl.	FLN	xx-0-BONN-35967	II / E	Tuberosae	Africa
Impatiens glandulifera Royle	GLN	xx-0-BONN-3774	II / B	n.d.	Asia
Impatiens grandis B. Heyne	GRN	xx-0-BONN-13388	n.d.	n.d.	Asia
Impatiens hochstetteri Warb.	HCH	KE-0-BONN-17872	n.d.	n.d.	Africa
Impatiens marianae Rchb. fil. ex Hook. fil.	MRN	XX-O-Z-20090082	n.d.	n.d.	Asia
Impatiens mengtszeana Hook. fil.	MNG	xx-0-BONN-36134	II / G	Uniflorae	Asia
Impatiens namchabarwensis	NMC	ZW-0-HBG-89-G-1021 (36593)	n.d.	n.d.	Asia
Impatiens niamniamensis Gilg. emdend G.M. Schulze	NMN	RW-0-BONN-11521	II / G	Uniflorae	Africa
Impatiens noli-tangere L.	NLT	DE-0-BONN-33527	II / D	Impatiens	Asia
Impatiens platypetala Lindl.	PLT	xx-0-BONN-36569	II / G	Uniflorae	Asia
Impatiens puberula DC	PBR	NP-0-BONN-36758	n.d.	n.d.	Asia
Impatiens racemosa DC	RCM	xx-0-BONN-36980	II / B	Racemosae	Asia
Impatiens repens Moon	RPN	xx-0-BONN-33522	II / G	Uniflorae	Asia
Impatiens rothii Hook. fil.	RTH	xx-0-BONN-36141	II / E	Tuberosae	Africa
Impatiens scabrida DC	SCB	xx-0-BONN-36760	II / B	Racemosae	Asia
Impatiens sodenii Engl. & Warb.	SDN	xx-0-BONN-28429	II / G	Uniflorae	Africa
Impatiens spec. (Nr. 36248)	362	BI-O-Bonn-36248	n.d.	n.d.	Africa
Impatiens spec. (Nr. 34558)	345	RW-0-BONN-34558	n.d.	n.d.	Africa
Impatiens textori Miq.	TXT	xxx-0-BONN-37372	II / D	Uniflorae	Japan
Impatiens tinctoria var. elegans	TNC	xx-0-BONN-36145	II / E	Tuberosae	Africa
Hydrocera triflora (L.) Wight & Arn.	HYD	xx-0-BONN-35302	Out group	-	Asia

* Genetic information (clade, subclade, section) from [5,9,29,30], [⊥] Oigin information from [8,9,85], n.d. = non defined

2.3.1 Thin-layer chromatography (TLC)

Thin-layer chromatography (TLC) is used as a fast and effective method for the separation of complex mixtures such as plant extracts. This method is widespread in the Pharmacopeias and acts as quality control for plant extract used in pharmaceutical applications. Furthermore, spray reagents enable a derivatization of specific compound classes. In this study, TLC was used as starting point to explore the variability of secondary metabolites present in the leaf extracts of the different species.
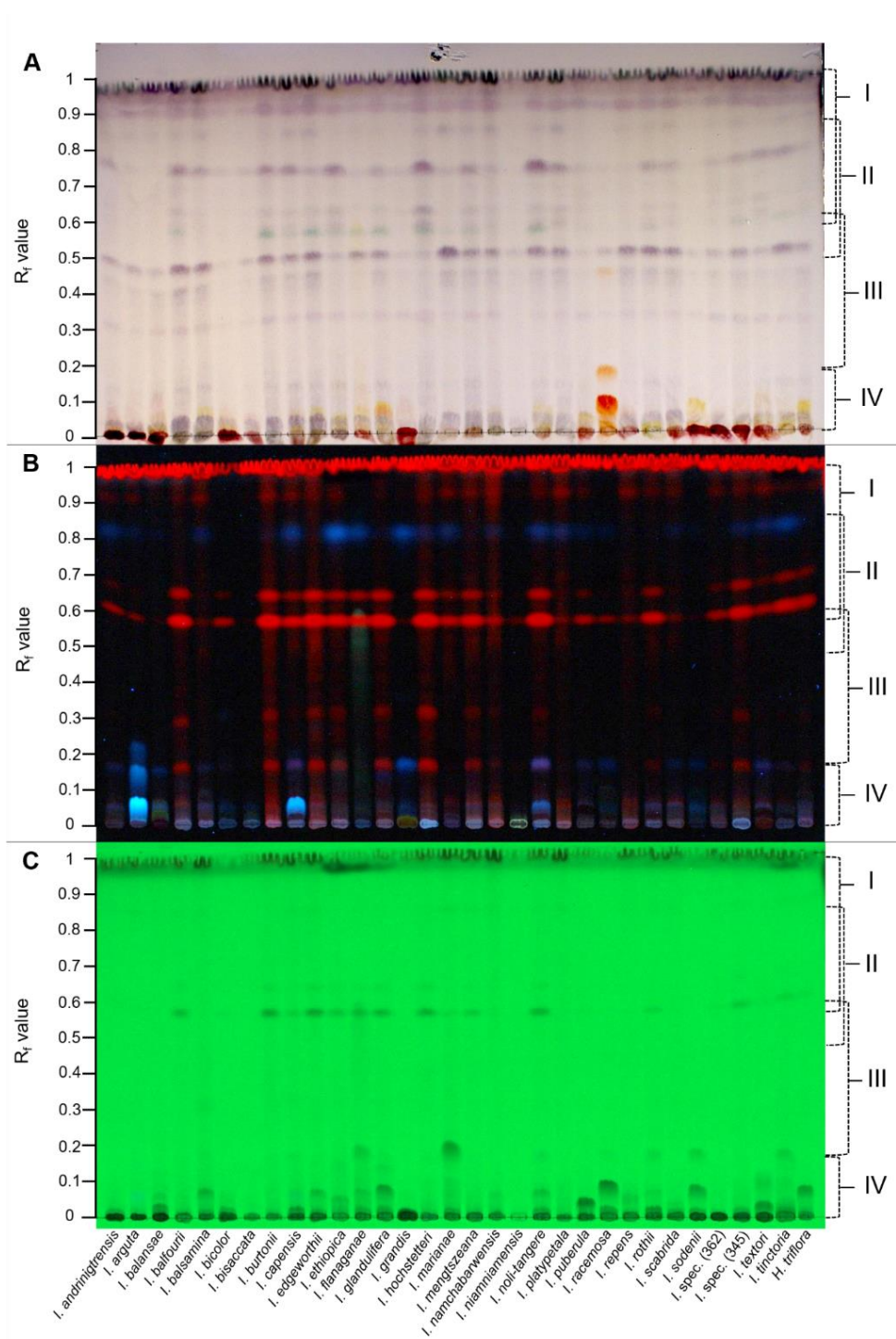
Figure 2.2. Thin-layer chromatograms of crude methanolic leaf extracts (c = 10 mg/ml, mobile phase: chloroform: methanol (9:1, *v/v*)), visualized using derivatization with vanillin-sulfuric acid reagent (A), under UV-light (366 nm (B) and 254 nm (C)). R_f ranges: I (0.55 – 1), II (0.5 – 0.8), III (0.2 – 0.6), IV (0 – 0.2).

In Figure 2.2, the chromatographic separation of the methanolic leaf extracts on normal phase using chloroform/methanol (9:1, *v/v*) as eluents is shown. The used solvent mixture yielded a partial separation of the compounds present in the different extracts. More hydrophobic compounds show bands with higher R_f value, while highly polar compounds are retained in a lower R_f range between 0 - 0.2. At first view, the extracts under investigation show a complex

composition exhibiting a number of similarities but also clear differences in their metabolite patterns. Common bands between a R_f value 0.55 to 1.0 (I, Figure 2.2), which exhibit a strong red fluorescence under 366 nm (Figure 2.2 B, I) correspond to chlorophylls and their degradation products. Further common bands within this R_f range (II, Figure 2.2) are related to more hydrophobic compounds of the extracts like steroids, triterpenoids and related saponins, which are reported for this plant family.[25,63,64,66] This is supported by the blue color change after spraying with vanillin sulfuric acid reagent (II, Figure 2.2 A).

The blue fluorescence band at R_f 0.8 (II, Figure 2.2 B) could refer to coumarins structurally related to 7-hydroxycoumarin used as reference which show a high R_f value using the same eluent system (Appendix, Table A.1, Figure A.1). Flavonoid aglycons, like the reference compound kaempferol, are found in the middle range area of the TLC, while their glycosides are retained at lower R_f (section III of the TLC). Polyhydroxylated organic acids, like caffeic acid, appear in the lowest R_f range between 0 - 0.1 (IV, Figure 2.2). This differentiation is based on the comparison with reference compounds (Appendix Table A.1, Figure A.1), which confirm that the strong absorbance at 254 nm in the mentioned R_f ranges (section III and IV in the TLC) are corresponding with the aromatic rings of these structures (Figure 2.1 C) and the yellow color correspond to the flavonoid glycosides after derivatization (Figure 2.1 A). The dark red spots at the application line after derivatization (Figure 2.1 A) are related to the presence of polymerized catechins, flavan-3-ol glycosides or flavanone glycosides, as spotted by reference compound naringin (Figure 2.2 A, Appendix Figure A.1).

Besides common bands, the extracts exhibit a number of species-specific compounds. The methanolic leaf extract of *I. racemosa* DC. exhibited a distinctive metabolite composition in comparison with the other extracts (Figure 2.2). The main constituents were identified after further phytochemical investigations as phlorizin (**2.5**, Scheme 2.1) and related dihydrochalcone derivatives (see 2.3.5), which cause the accentuated orange colored spots (R_f values 0.09, 0.18, 0.44; Figure 2.2 A). The intense blue fluorescent band at R_f value of 0.07 at 366 nm in the extracts of *I. arguta* and *I. capensis* (Figure 2.2 B) belong to coumarin glycosides, as it was further confirmed by the UHPLC-ESI-HRMS analysis (see 2.3.2). The methanolic extract of *I. grandis* (GRN) exhibit a yellow-green spot at the application point (Figure 2.2 B), which was annotated as the xanthone-*C*-glycoside mangiferin based on its fragmentation behavior, as well as, on its typical absorbance maxima at 239, 258, 319 and 368 nm. (Table 2.2, Appendix Figure A.2).

The investigation of the methanolic leaf extracts by planar TLC gave the first insights into the complexity of the extracts and enabled a first comparison of the different constituents of the investigated species. Furthermore, the properties under UV-light and after derivatization with vanillin-sulfuric acid spray reagent gave hints for the compound classes, like flavonoids, coumarins, dihydrochalcone derivatives and others. However, without reference substances or further characterization by other analytical methods, specific compound identification is rather

difficult. Therefore, a UHPLC-ESI-HRMS based profiling was performed to compare the different species and to annotate selected metabolites based on their fragmentation behavior.

2.3.2 UHPLC-ESI-HRMS based profiling of *Impatiens* species and *Hydrocera triflora*

For more in-depth analysis of the secondary metabolite profiles, their comparison, and the characterization of the chemical constituents, an untargeted UHPLC-ESI-HRMS based profiling was performed. The methanolic extracts were separated on a reversed-phase C18 column using an optimized gradient system. The detection was performed using a high-resolution Orbitrap mass spectrometer allowing the determination of the elemental composition of the detected *m/z* values. Furthermore, MS^2 experiments using data-dependent acquisition were performed for a putative assignment of the metabolites. (see experimental part, 2.5.4)

2.3.2.1 Method development and experimental design

For the optimal extract concentration, a linear dynamic range analysis comprising a dilution series of the quality control sample (QC) was performed (see Appendix, Table A.2, Figure A.3). Two milligram plant material extracted in one milliliter LC-MS grade methanol showed best results for the detection of analytes in an intensity range between 10^4 -10^7. Besides the optimal extract concentration, the UHPLC gradient was optimized allowing a comprehensive elution of plant analytes within 15 minutes and to prevent carry over (Appendix, Figure A.4).

Furthermore, the electrospray ion source conditions (spray voltage, source temperature, capillary temperature, sheath and auxiliary gas) were optimized using the flavonoid glycoside rutin as reference sample to enhance the ionization efficiency and to avoid in-source fragmentation as well as thermal degradation of the analytes during UHPLC-ESI-HRMS analysis. The samples were measured both in positive and negative ionization mode. Due to a higher number of detected peaks, a lower noise and better signal-to-noise ratio (S/N) in the total ion chromatogram (TIC) as well as a higher number of MS/MS spectra acquired in data-dependent mode, the negative ion mass spectra were used for further analysis.

For controlling the performance of the UHPLC with respect to retention time shifts as well as for the external calibration of the instrument, an internal standard (kinetin, $C_{10}H_9ON_5$, 20 µM) was used. This compound being not present in the extracts, did not coelute with any metabolite (Appendix, Figure A.4, t_R 2.04 min) and, showed good ionization properties. In order to monitor the reproducibility of sample preparation and the instrument performance, a pooled QC sample, containing plant powder of all species, was prepared and analyzed alongside individual samples.

2.3.2.2 Analysis and comparison of Impatiens and Hydrocera leaf metabolites

Chemical constituents of the 31 examined *Impatiens* species and *Hydrocera triflora* leaf extracts were separated via reversed-phase UHPLC resulting in the representative total ion chromatograms shown in Figure 2.3.
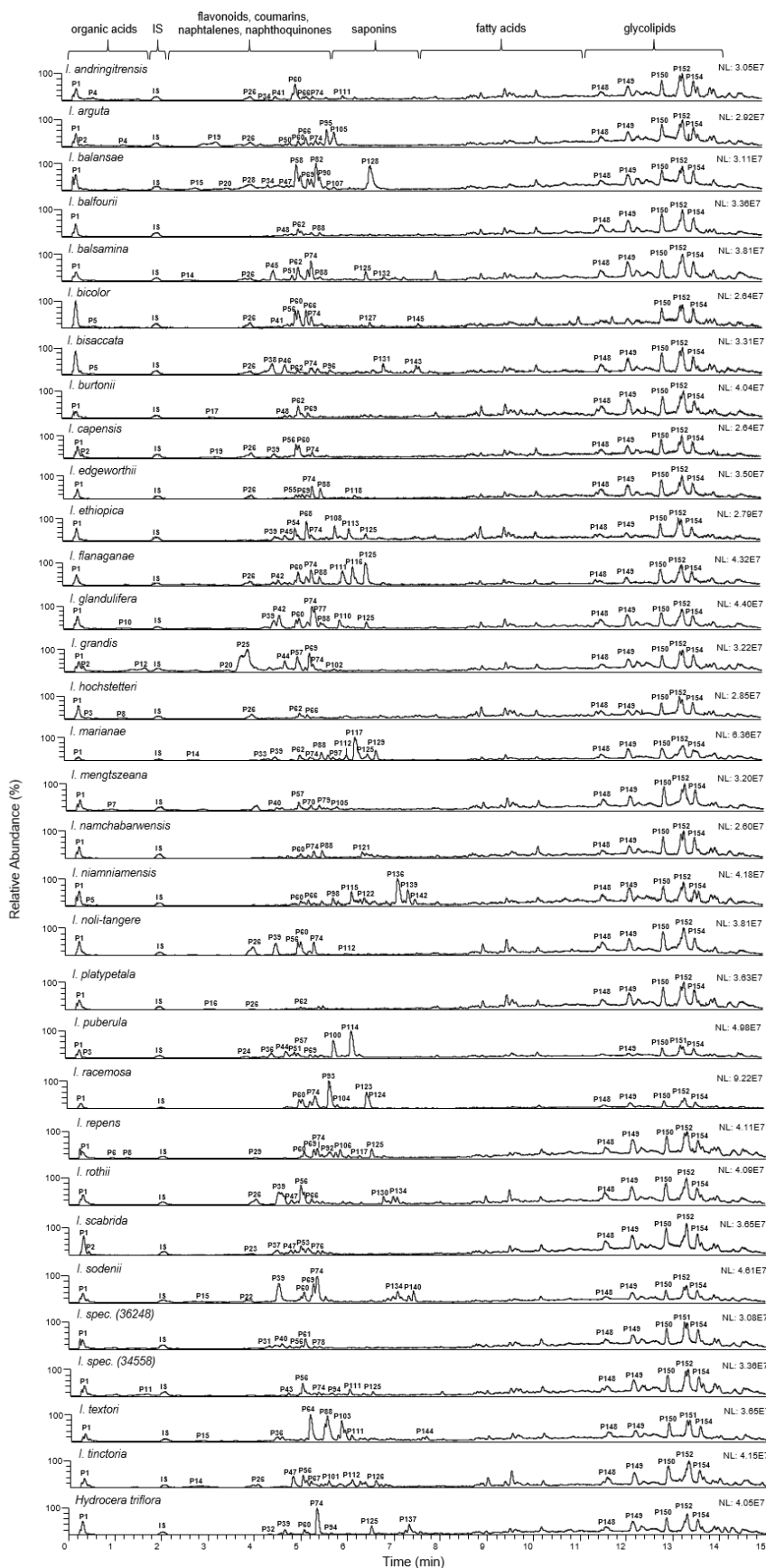
Figure 2.3. Comparison of total ion chromatograms (TICs, *m/z* 100-1500) in negative ion mode of the investigated *Impatiens* species and *Hydrocera triflora*. (IS = internal standard Kinetin, P = peak number, see Table 2.2)

Very polar compounds such as organic acids were detected in the retention time range between 0 to 2 minutes (Figure 2.3, Table 2.2), including the internal standard at t_R 2.05 min. Hydroxycinnamic acid derivatives, coumarin conjugates, flavonoids, and naphthoquninones were found in the retention time range of 3 – 7 min (Figure 2.3) according on their substitution pattern (e.g. level of glycosylation, Table 2.2). Flavonoid aglycones, dihydrochalcone conjugates and some naphthoquinone derivatives showed a stronger interaction with the stationary phase and eluted later from the column. Metabolites, which showa high *m/z* value and a complex glycosylation pattern according to their mass spectral fragmentation, eluted at a retention time of 7 to 8 minutes. Possibly, they belong to the class of saponins (Figure 2.3, Table 2.2). Peaks detected at t_R 8.98 min, 9.50 min, and 10.15 min were identified as fatty acid derivatives based on the comparison with the methanol blank (Figure 2.3, Appendix, Figure A.4 F).

Highly hydrophobic constituents, present in all species, elute with at higher proportion of acetonitrile between 11 to 15 minutes (Figure 2.3). The detection of the metabolites present in the different leaf extracts was performed using electrospray (ESI) high-resolution mass spectrometry (HRMS) to enable the determination of the elemental composition. The negative ion mass spectra are characterized by [M-H]$^-$ ions. In some cases, the formiate adduct [M-H+HCOOH]$^-$ was detected, e.g. in P114 and P146-P156 (Table 2.2). Mass spectral fragmentation studies using data-dependent HR-MS^2 experiments were performed for the peak assignment (Table 2.2). Furthermore, the four-level system for metabolite annotation suggested by Creek et al.[86] was applied. The main compound classes detected during the profiling are shown in Scheme 2.1.
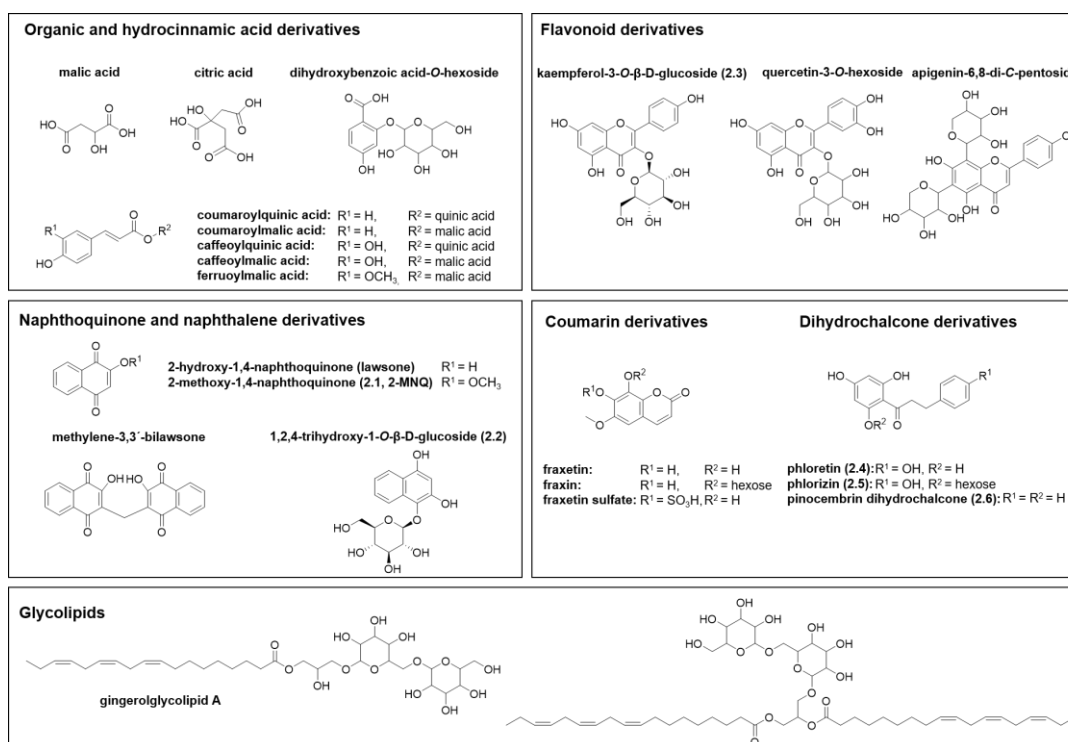
Scheme 2.1. Overview of main compound classes detected within the investigated *Impatiens* species and *Hydrocera triflora* using negative ion mode UHPLC-ESI-HRMS. Identified compounds based on UHPLC-ESI-HRMS and NMR possess a compound number.

Table 2.2. Metabolites putatively annotated in *Impatiens* species and *Hydrocera triflora* methanol leaf extracts via UHPLC-HRMS² in negative ionization mode (CID, NCE 35 %).

Peak no.	Feature	t_R (min)	[M-H]⁻ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P1	133/20	0.33	133.0144	1.4	$C_4H_5O_5^-$	115.0038 ($C_4H_3O_4^-$, **100**), 87.0090 ($C_3H_3O_3^-$, 1), 71.0141 ($C_3H_3O_2^-$, 4)	Malic acid[1]	1	all
P2	191/25	0.41	191.0196	-0.7	$C_6H_7O_7^-$	173.0092 ($C_6H_5O_6^-$, 12), 129.0195 ($C_5H_5O_4^-$, 3), 111.0089 ($C_5H_3O_3^-$, **100**), 87.0090 ($C_3H_3O_3^-$, 2), 85.0298 ($C_4H_5O_2^-$, 1)	Citric acid[2]	1	all
P3	-	0.53	331.0663	0.8	$C_{13}H_{15}O_{10}^-$	313.0560 ($C_{13}H_{13}O_9^-$, 6), 287.0772 ($C_{12}H_{15}O_8^-$, 3), 211.0245 ($C_9H_7O_6^-$, 6), 169.0142 ($C_7H_5O_5^-$, 21), 168.0064 ($C_7H_4O_5^-$, 82), 167.0350 ($C_8H_7O_4^-$, 12), 149.9959 ($C_7H_2O_4^-$, 7), 125.0245 ($C_6H_5O_3^-$, 27)	Glucogallic acid	2	BLS
P4	-	0.67	305.0661	-2.0	$C_{15}H_{13}O_7^-$	287.0558 ($C_{15}H_{11}O_6^-$, 8), 261.0766 ($C_{14}H_{13}O_5^-$, 47), 221.0453 ($C_{11}H_9O_5^-$, 66), 219.0660 ($C_{12}H_{11}O_4^-$, 55), 179.0349 ($C_9H_7O_4^-$, **100**), 167.0349 ($C_8H_7O_4^-$, 10), 139.0401 ($C_7H_7O_3^-$, 10), 137.0245 ($C_7H_5O_3^-$, 23), 125.0245 ($C_6H_5O_3^-$, 34)	Gallocatechin/Epigallocatechin[14]	2	AND
P5	315.1/41	0.70	315.0714	-2.7	$C_{13}H_{15}O_9^-$	225.0402 ($C_{10}H_9O_6^-$, 6), 165.0193 ($C_8H_5O_4^-$, 13), 163.0401 ($C_9H_7O_3^-$, 8), 153.0194 ($C_7H_5O_4^-$, **100**), 152.0116 ($C_7H_4O_4^-$, 42), 109.0296 ($C_6H_5O_2^-$, 9)	Dihydroxybenzoic acid hexoside	2	all
P6	-	0.95	385.0767	-1.7	$C_{16}H_{17}O_{11}^-$	357.0822 ($C_{15}H_{17}O_{10}^-$, 3), 315.0720 ($C_{13}H_{15}O_9^-$, 2), 297.0611 ($C_{13}H_{13}O_8^-$, 97), 223.0245 ($C_{10}H_8O_6^-$, **100**), 205.0140 ($C_{10}H_5O_5^-$, 55), 195.0297 ($C_9H_7O_5^-$, 1), 179.0348 ($C_9H_7O_4^-$, 29), 161.0243 ($C_9H_5O_3^-$, 6), 153.0193 ($C_7H_5O_4^-$, 33)	Unidentified	4	RPN
P7	-	1.00	327.0715	-2.0	$C_{14}H_{15}O_9^-$	309.0611 ($C_{14}H_{13}O_8^-$, 3), 179.0348 ($C_9H_7O_4^-$, 4), 165.0403 ($C_5H_9O_6^-$, **100**), 129.0194 ($C_5H_5O_3^-$, 3)	Unidentified	4	MNG
P8	-	1.29	225.0401	-1.5	$C_{10}H_9O_6^-$	207.0297 ($C_{10}H_7O_5^-$, **100**), 181.0506 ($C_9H_9O_4^-$, 32), 179.0349 ($C_9H_7O_4^-$, 30), 163.0401 ($C_9H_7O_3^-$, 38), 151.0403 ($C_8H_7O_3^-$, 1), 135.0453 ($C_8H_7O_2^-$, 2), 119.0503 ($C_8H_7O_7^-$, 5)	Unidentified	4	HCH, RPN
P9	305.1/78	1.30	305.0662	-1.6	$C_{15}H_{13}O_7^-$	287.0558 ($C_{15}H_{11}O_6^-$, 8), 261.0766 ($C_{14}H_{13}O_5^-$, 45), 221.0453 ($C_{11}H_9O_5^-$, 68), 219.0660 ($C_{12}H_{11}O_4^-$, 56), 179.0349 ($C_9H_7O_4^-$, **100**), 167.0349 ($C_8H_7O_4^-$, 10), 139.0401 ($C_7H_7O_3^-$, 10), 137.0245 ($C_7H_5O_3^-$, 23), 125.0245 ($C_6H_5O_3^-$, 34)	Gallocatechin/Epigallocatechin[14]	2	HYD,AND, ARG, BSC, ETH, GRN, MNG, NLT, RPN
P10	-	1.40	297.0609	-2.0	$C_{13}H_{13}O_8^-$	279.0509 ($C_{13}H_{11}O_7^-$, 2), 179.0350 ($C_9H_7O_4^-$, 20), 135.0300 ($C_8H_7O_2^-$, **100**)	Caffeic acid derivative	3	GLN
P11	-	1.68	337.0921	-2.2	$C_{16}H_{17}O_8^-$	191.0558 ($C_7H_{11}O_6^-$, 7), 173.0453 ($C_7H_9O_5^-$, 5), 163.0399 ($C_9H_7O_3^-$, **100**), 119.0502 ($C_8H_7O^-$, 5)	Coumaroylquinic acid	3	345
P12	-	1.73	407.0976	-2.5	$C_{19}H_{19}O_{10}^-$	389.0873 ($C_{19}H_{17}O_9^-$, 5), 317.0662 ($C_{16}H_{13}O_7^-$, 22), 287.0557 ($C_{15}H_{11}O_6^-$, **100**)	Unidentified	4	GRN
P13	353.1/120	2.00	353.0869	-2.4	$C_{16}H_{17}O_9^-$	191.0560 ($C_7H_{11}O_6^-$, **100**), 179.0350 ($C_9H_7O_4^-$, 45), 173.0455 ($C_7H_9O_5^-$, 3), 135.0453 ($C_8H_7O_2^-$, 8)	Caffeoylquinic acid[16]	1	ARG, ETH, GRN, NLT, PLT, PBR, 345, 362
P14	-	2.67	347.0401	-2.5	$C_{16}H_{11}O_9^-$	303.0508 ($C_{15}H_{11}O_7^-$, 42), 275.0560 ($C_{14}H_{11}O_6^-$, 14), 229.0141 ($C_{12}H_5O_5^-$, **100**), 201.0192 ($C_{11}H_5O_4^-$, 45), 185.0244 ($C_{11}H_5O_3^-$, 16)	Unidentified	4	BLS, ETH, MRN, TNC
P15	-	2.85	577.1342	-1.4	$C_{30}H_{25}O_{11}^-$	559.1238 ($C_{30}H_{23}O_{11}^-$, 7), 451.1027 ($C_{24}H_{19}O_9^-$, 29), 425.0871 ($C_{22}H_{17}O_9^-$, **100**), 407.0766 ($C_{22}H_{15}O_8^-$, 56), 299.0556 ($C_{16}H_{11}O_6^-$, 7), 289.0714 ($C_{15}H_{13}O_6^-$, 21), 287.0558 ($C_{15}H_{11}O_6^-$, 7)	Unidentified	4	BLN, GRN, TXT
P16	321.1/188	3.13	321.0972	-2.3	$C_{16}H_{17}O_7^-$	213.0555 ($C_{13}H_9O_3^-$, 37), 201.0555 ($C_{12}H_9O_3^-$, **100**), 171.0451 ($C_{11}H_7O_2^-$, 13), 159.0451 ($C_{10}H_7O_2^-$, 98), 158.0374 ($C_{10}H_6O_2^-$, 11)	4-Methylumbelliferyl-C-glycoside	2	PLT

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	$[M-H]^-$ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P17	-	3.24	209.0451	-2.0	$C_{10}H_9O_5^-$	191.0349 ($C_{10}H_7O_4^-$, **100**), 147.4503 ($C_8H_7O_2^-$, 2)	Coumarin derivative	3	BRT
P18	-	3.30	289.0713	-1.6	$C_{15}H_{16}O_6^-$	245.0818 ($C_{14}H_{13}O_4^-$, **100**), 205.0506 ($C_{11}H_9O_4^-$, 32), 179.0351 ($C_9H_7O_4^-$, 14), 137.0246 ($C_7H_5O_3^-$, 4), 125.0246 ($C_6H_5O_3^-$, 4)	Unidentified	4	ARG, BLN, ETH, GRN
P19	369.1/199	3.31	369.0817	-2.7	$C_{16}H_{17}O_{10}^-$	354.0587 ($C_{15}H_{14}O_{10}^-$, 1), 207.0296 ($C_{10}H_7O_5^-$, **100**), 192.0062 ($C_9H_4O_5^-$, 11)	Fraxin	2	ARG, BLS, CPN, GLN, GRN, NLT, RCM, RTH, SDN
P20	-	3.53	289.0710		$C_{15}H_{16}O_6^-$	245.0818 ($C_{14}H_{13}O_4^-$, **100**), 205.0506 ($C_{11}H_9O_4^-$, 33), 179.0350 ($C_9H_7O_4^-$, 15), 137.0247 ($C_7H_5O_3^-$, 4), 125.0246 ($C_6H_5O_3^-$, 5)	Unidentified	4	ARG, ETH, GRN
P21	-	3.83	207.0297	-1.3	$C_{10}H_7O_5^-$	192.0063 ($C_9H_4O_5^-$, **100**)	Fraxetin	2	ARG, CPN
P22	-	3.83	449.1080	-1.9	$C_{21}H_{22}O_{11}^-$	287.0560 ($C_{15}H_{11}O_6^-$, **100**), 269.0455 ($C_{15}H_9O_5^-$, 33), 259.0612 ($C_{14}H_{11}O_5^-$, 38)	Unidentified	4	SDN
P23	-	3.95	595.1658	-1.1	$C_{26}H_{27}O_{16}^-$	505.1344 ($C_{24}H_{25}O_{12}^-$, 29), 475.1248 ($C_{23}H_{23}O_{12}^-$, 77), 457.1133 ($C_{23}H_{22}O_{10}^-$, 20), 415.1028 ($C_{21}H_{19}O_9^-$, 31), 385.0923 ($C_{20}H_{17}O_8^-$, 80), 355.0817 ($C_{19}H_{15}O_7^-$, **100**)	Unidentified	4	SCB
P24	-	3.98	401.1446	-1.3	$C_{18}H_{25}O_{10}^-$	293.0874 ($C_{11}H_{17}O_9^-$, 6), 269.1027 ($C_{13}H_{17}O_6^-$, **100**), 233.0664 ($C_9H_{13}O_7^-$, 7), 161.0455 ($C_6H_5O_5^-$, 19), 125.0245 ($C_6H_5O_3^-$, 2)	Unidentified	4	BLS, CPN, MNG, PBR, SDN
P25	421.1/237	3.95	421.0767	-2.2	$C_{19}H_{17}O_{11}^-$	403.0666 ($C_{19}H_{15}O_{10}^-$, 23), 361.0561 ($C_{17}H_{13}O_9^-$, 1), 331.0455 ($C_{16}H_{11}O_8^-$, 89), 301.0350 ($C_{15}H_9O_7^-$, **100**), 259.0246 ($C_{13}H_6O_6^-$, 2)	Mangiferin[90]	1	GRN
P26	295/242	4.03	295.0454	-1.9	$C_{13}H_{11}O_8^-$	179.0349 ($C_9H_7O_4^-$, **100**), 135.045 ($C_8H_7O_2^-$, 5), 133.0143 ($C_4H_5O_5^-$, 84), 115.0038 ($C_4H_3O_4^-$, 5)	Caffeoylmalic acid[91]	1	AND, ARG, BLN, BLF, BLS, BCL, BSC, BRT, CPN, EDG, ETH, FLN, HCH, NMC, NMN, NLT, PLT, PBR, RCM, RPN, RTH, 345, 362, TNC
P27	-	4.03	459.1134	-2.2	$C_{19}H_{23}O_{13}^-$	415.1239 ($C_{18}H_{23}O_{11}^-$, 1), 397.1134 ($C_{18}H_{21}O_{10}^-$, 42), 357.0821 ($C_{15}H_{17}O_{10}^-$, 91), 315.0715 ($C_{15}H_{17}O_{10}^-$, **100**), 153.0193 ($C_7H_5O_4^-$, 1)	Unidentified	4	ARG
P28	-	4.03	593.1500	-1.6	$C_{27}H_{29}O_{15}^-$	467.1185 ($C_{21}H_{23}O_{11}^-$, 6), 327.0502 ($C_{17}H_{11}O_7^-$, 11), 299.0554 ($C_{16}H_{11}O_6^-$, 24), 285.0397 ($C_{15}H_9O_6^-$, **100**), 284.0321 ($C_{15}H_8O_6^-$, 47)	Kaempferol conjugate	3	BLN
P29	-	4.04	207.0296	-1.5	$C_{10}H_7O_5^-$	192.0062 ($C_9H_4O_5^-$, 1), 179.0350 ($C_9H_7O_4^-$, 19), 163.0401 ($C_9H_7O_3^-$, **100**), 145.0296 ($C_9H_5O_2^-$, 7), 135.0453 ($C_8H_7O_2^-$, 24), 119.0504 ($C_8H_7O^-$, 41)	Coumaric acid derivative	3	RPN
P30	-	4.27	387.1649	-3.0	$C_{18}H_{27}O_9^-$	369.1549 ($C_{18}H_{25}O_8^-$, 14), 351.0751 ($C_{16}H_{15}O_6^-$, 12), 225.1123 ($C_{12}H_{17}O_4^-$, 3), 207.1024 ($C_{12}H_{15}O_3^-$, **100**), 163.1128 ($C_{11}H_{15}O^-$, 42), 113.0241 ($C_5H_5O_3^-$, 2)	Hydronaphthalene conjugate	3	ARG
P31	-	4.31	787.1931	-2.7	$C_{33}H_{39}O_{22}^-$	667.1503 ($C_{29}H_{31}O_{18}^-$, 35), 445.0769 ($C_{21}H_{17}O_{11}^-$, 38), 355.0453 ($C_{18}H_{11}O_8^-$, 11), 343.0453 ($C_{17}H_{11}O_8^-$, 11), 301.0347 ($C_{15}H_9O_7^-$, **100**), 300.0271 ($C_{15}H_8O_7^-$, 52), 271.0244 ($C_{14}H_7O_6^-$, 14)	Quercetin conjugate	3	362
P32	-	4.32	431.1915	-2.0	$C_{20}H_{31}O_{10}^-$	385.1861 ($C_{19}H_{29}O_8^-$, **100**), 223.1336 ($C_{13}H_{19}O_3^-$, 4), 153.0921 ($C_9H_{13}O_2^-$, 2)	Unidentified	4	HYD

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	[M-H]− measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P33	–	4.40	173.0245	−0.2	$C_{10}H_5O_3^-$	145.0296 ($C_9H_5O_2^-$, **100**), 111.0816 ($C_8H_5O^-$, 6)	2-Hydroxy-1,4-naphthoquinone (Lawsone)	2	MRN
P34	371.1/264	4.40	371.0976	−2.2	$C_{16}H_{19}O_{10}^-$	353.0874 ($C_{16}H_{17}O_9^-$, 6), 249.0613 ($C_9H_{13}O_8^-$, **100**), 231.0508 ($C_9H_{11}O_7^-$, 7), 121.0296 ($C_7H_5O_2^-$, 2)	Benzoic acid derivative	3	AND, ARG, BLN, BLF, BLS, BCL, BSC, BRT, EDG, ETH, FLN, GLN, GRN,HCH,MRN, MNG,NMC, NMN, NLT, PLT, PBR, RCM, RPN, RTH, SDN, HYD
P35	481.1/265	4.43	481.0979	−1.7	$C_{21}H_{21}O_{13}^-$	319.0454 ($C_{15}H_{11}O_8^-$, 44), 301.0349 ($C_{15}H_9O_7^-$, 25), 257.0452 ($C_{14}H_9O_5^-$, 3), 215.0346 ($C_{12}H_7O_4^-$, 2), 193.0140 ($C_9H_5O_5^-$, **100**), 175.0035 ($C_9H_3O_4^-$, 3)	Coumarin derivative	3	BLN, BCL, BSC
P36	593.2/267	4.45	593.1497	−2.6	$C_{27}H_{29}O_{15}^-$	575.1398 ($C_{27}H_{27}O_{14}^-$, 10), 503.1189 ($C_{24}H_{23}O_{12}^-$, 34), 473.1083 ($C_{23}H_{21}O_{11}^-$, **100**), 455.0979 ($C_{23}H_{19}O_{10}^-$, 4), 413.0871 ($C_{21}H_{17}O_9^-$, 3), 383.0768 ($C_{20}H_{15}O_8^-$, 17), 353.0663 ($C_{19}H_{13}O_7^-$, 30)	Apigenin-6,8-di-C-hexoside[90]	2	GRN, PBR, SCB, 362, TXT
P37	–	4.49	461.1658	−1.5	$C_{20}H_{29}O_{12}^-$	415.1603 ($C_{19}H_{27}O_{10}^-$, **100**)	Unidentified	4	SCB
P38	–	4.51	449.1081	−1.8	$C_{21}H_{21}O_{13}^-$	287.0557 ($C_{15}H_{11}O_6^-$, **100**), 259.0608 ($C_{14}H_{11}O_5^-$, 68), 243.0660 ($C_{14}H_{11}O_4^-$, 6)	Unidentified	4	BSC
P39	337.1/272	4.53	337.0921	−2.2	$C_{16}H_{17}O_8^-$	187.0400 ($C_{11}H_9O_3^-$, 1); 175.0400 ($C_{10}H_7O_3^-$, **100**); 173.0094 (5); 111.0089 ($C_5H_3O_3^-$, 1)	1,2,4-Trihydroxynaphthalene-1-O-β-D-glucoside[58]	1	ARG, BLN, BLS, BCL, CPN, EDG, FLN, GLN, MRN, MNG, NLT, RCM, RTH, SDN, TNC
P40	625.1/277	4.61	625.1397	−1.5	$C_{27}H_{29}O_{17}^-$	607.1295 ($C_{27}H_{27}O_{16}^-$, 2); 505.0977 ($C_{23}H_{21}O_{13}^-$, 32); 463.0873 ($C_{21}H_{19}O_{12}^-$, 22); 445.0767 ($C_{21}H_{17}O_{11}^-$, 22); 343.0452 ($C_{17}H_{11}O_8^-$, 6), 301.0345 ($C_{15}H_9O_7^-$, 74), 300.0271 ($C_{15}H_8O_7^-$, **100**), 273.0400 ($C_{14}H_9O_6^-$, 4), 271.2044 ($C_{14}H_7O_6^-$, 15), 255.0296 ($C_{14}H_7O_5^-$, 7)	Quercetin-di-O-hexoside	2	AND, BLN, BCL, FLN, GLN, MNG, RCM, RTH, SCB, 362, 345, TNC
P41	479.1/278	4.63	479.0822	−1.9	$C_{21}H_{19}O_{13}^-$	461.0718 ($C_{21}H_{17}O_{12}^-$, 2), 359.0403 ($C_{17}H_{11}O_9^-$, 3), 317.0295 ($C_{15}H_9O_8^-$, 56), 316.0220 ($C_{15}H_8O_8^-$, **100**), 178.9984 ($C_8H_3O_5^-$, 3)	Myricetin-O-hexoside[69]	2	AND, ARG, BLS, BCL, BSC, CPN, EDG, ETH, GLN, GRN, HCH, MNG, RPN, SCB, SDN, 345, TXT, HYD

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	[M-H]$^-$ measured	Error (ppm)	Molecular formula	MS^2 product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P42	351.1/279	4.65	351.1076	-1.0	$C_{17}H_{19}O_8^-$	319.0822 ($C_{16}H_{15}O_7^-$, 32), 301.0715 ($C_{16}H_{13}O_6^-$, 3), 259.0611 ($C_{14}H_{11}O_5^-$, 11), 241.0506 ($C_{14}H_9O_4^-$, 9), 231.0662 ($C_{13}H_{11}O_4^-$, 6), 229.0505 ($C_{13}H_9O_4^-$, 25), 211.0400 ($C_{13}H_7O_3^-$, 17), 201.0557 ($C_{12}H_9O_3^-$, 5), 199.0400 ($C_{12}H_7O_3^-$, 8), 189.0556 ($C_{11}H_9O_3^-$, 25), 188.0478 ($C_{11}H_7O_3^-$, 19), 187.0401 ($C_{11}H_7O_3^-$, **100**), 175.0401 ($C_{10}H_7O_3^-$, 27), 173.0245 ($C_{10}H_5O_3^-$, 16), 159.0453 ($C_{10}H_7O_3^-$, 4), 113.0245 ($C_5H_5O_3^-$, 3)	Naphthoquinone derivative	3	BLS, CPN, ETH, FLN, GLN, RTH, TNC
P43		4.70	609.1447	-2.0	$C_{27}H_{29}O_{16}^-$	447.0926 ($C_{21}H_{19}O_{11}^-$, **100**), 285.0400 ($C_{15}H_9O_6^-$, 46), 284.0323 ($C_{15}H_9O_6^-$, 5),	Kaempferol-*O*-hexosyl-*O*-hexoside	2	345
P44	563.1/285	4.75	563.1399	-1.3	$C_{26}H_{27}O_{14}^-$	545.1295 ($C_{26}H_{23}O_{12}^-$, 22), 503.1190 ($C_{24}H_{23}O_{12}^-$, 42), 473.1084 ($C_{23}H_{21}O_{11}^-$, 99), 443.0979 ($C_{22}H_{19}O_{10}^-$, **100**), 425.0973 ($C_{22}H_{17}O_9^-$, 8), 383.0768 ($C_{20}H_{15}O_8^-$, 34), 353.0663 ($C_{19}H_{13}O_7^-$, 36)	Apigenin-*C*-hexoside-*C*-pentoside[92]	2	GRN, PBR, 362
P45	337.1/288	4.80	337.0559	-1.7	$C_{15}H_{13}O_9^-$	191.0196 ($C_6H_6O_7^-$, 11), 173.0091 ($C_6H_5O_6^-$, **100**), 129.0195 ($C_5H_5O_4^-$, 1), 111.0089 ($C_5H_3O_3^-$, 20)	2-Coumaroyl isocitric acid	2	HYD, ETH, GRN, SCB, TNC
P46	527/288	4.80	527.0500	-0.2	$C_{21}H_{19}O_{14}S^-$	447.0927 ($C_{21}H_{19}O_{11}^-$, **100**), 285.0403 ($C_{15}H_9O_6^-$, 2)	Kaempferol-sulfate hexoside	2	BSC
P47	595.1/288	4.80	595.1295	-1.7	$C_{26}H_{27}O_{16}^-$	475.0871 ($C_{22}H_{19}O_{12}^-$, 8), 463.0877 ($C_{21}H_{19}O_{12}^-$, 8), 445.0771 ($C_{21}H_{17}O_{11}^-$, 14), 301.0348 ($C_{15}H_9O_7^-$, 44), 300.0273 ($C_{15}H_8O_7^-$, **100**), 271.0245 ($C_{14}H_7O_6^-$, 10), 179.0033 ($C_8H_3O_5^-$, 2)	Quercetin-*C*-hexoside-*O*-pentoside	2	BLN, BCL, GRN, RTH, SCB, 345, TNC
P48	279.1/293	4.88	279.0505	-2.0	$C_{13}H_{11}O_7^-$	179.0350 ($C_9H_7O_4^-$, 3), 163.0402 ($C_9H_7O_3^-$, **100**), 133.0144 ($C_4H_5O_5^-$, 88), 119.0504 ($C_8H_7O^-$, 10), 115.0038 ($C_4H_3O_4^-$, 3)	Coumaroylmalic acid	1	AND,ARG, BLN, BLF, BLS, BCL, BSC, BRT, CPN, EDG, ETH, FLN, GLN,HCH,MRN, NMC,NMN, NLT, PLT, PBR, RCM, RTH, SDN, 345, 362, TNC, HYD
P49	-	4.89	205.0502	-1.5	$C_{11}H_9O_4^-$	161.0608 ($C_{10}H_9O_2^-$, **100**)	Unidentified	4	MRN
P50	-	4.90	411.0926	-1.6	$C_{18}H_{19}O_{11}^-$	369.0822 ($C_{16}H_{17}O_{10}^-$, 1), 351.0716 ($C_{16}H_{15}O_9^-$, 5), 207.0296 ($C_{10}H_7O_5^-$, **100**), 192.0062 ($C_9H_4O_5^-$, 21)	Acetyl fraxin	2	ARG

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	$[M-H]^-$ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P51	711.2/296	4.93	711.2133	0.3	$C_{32}H_{39}O_{18}^-$	431.0976 ($C_{21}H_{19}O_{10}^-$, 13), 355.1023 ($C_{16}H_{19}O_9^-$, 2), 311.1130 ($C_{15}H_{19}O_7^-$, **100**), 279.0506 ($C_{13}H_{11}O_7^-$, 7), 269.1027 ($C_{13}H_{17}O_6^-$, 6)	Unidentified	4	BLF, BLS, BCL, BSC, EDG, GLN, PBR, SCB
P52	631.1/299	4.98	631.1295	-1.5	$C_{29}H_{27}O_{16}^-$	587.1396 ($C_{28}H_{27}O_{14}^-$, 1), 515.1186 ($C_{25}H_{23}O_{12}^-$, **100**), 471.1288 ($C_{24}H_{23}O_{10}^-$, 7), 337.0923 ($C_{16}H_{17}O_8^-$, 14), 335.0800 ($C_{16}H_{15}O_8^-$, 9), 291.0659 ($C_{18}H_{11}O_4^-$, 7), 277.0561 ($C_{10}H_{13}O_3^-$, 3), 179.0349 ($C_9H_7O_4^-$, 1)	Trihydroxynaphthalene conjugate	3	BLS, CPN, RTH
P53	-	4.99	783.1979	-1.3	$C_{34}H_{39}O_{21}^-$	741.1872 ($C_{32}H_{37}O_{20}^-$, **100**), 723.1770 ($C_{32}H_{35}O_{19}^-$, 17), 619.1296 ($C_{28}H_{27}O_{16}^-$, 3), 489.1034 ($C_{23}H_{21}O_{12}^-$, 4), 301.0342 ($C_{15}H_9O_7^-$, 6), 300.0271 ($C_{15}H_8O_7^-$, 14)	Quercetin derivative	3	SCB
P54	593.2/300	5.00	593.1503	0.3	$C_{27}H_{29}O_{15}^-$	447.0928 ($C_{21}H_{19}O_{11}^-$, **100**), 431.0979 ($C_{21}H_{19}O_{10}^-$, 60), 430.0904 ($C_{21}H_{18}O_{10}^-$, 9), 285.0403 ($C_{15}H_9O_6^-$, 22)	Kaempferol-*O*-rhamno-*O*-hexoside	2	ETH, 362
P55	739.2/300	5.00	739.2081	-1.3	$C_{33}H_{39}O_{19}^-$	593.1504 ($C_{27}H_{29}O_{15}^-$, 23), 575.1399 ($C_{27}H_{27}O_{14}^-$, **100**), 473.1082 ($C_{23}H_{21}O_{11}^-$, 10), 393.0612 ($C_{21}H_{13}O_8^-$, 8), 327.0505 ($C_{17}H_{11}O_7^-$, 15), 285.0395 ($C_{15}H_9O_6^-$, 41), 284.0323 ($C_{15}H_8O_6^-$, 69), 255.0296 ($C_{14}H_7O_5^-$, 21), 227.0348 ($C_{13}H_4O_4^-$, 3)	Kaempferol-di-*O*-rhamoside-*C*-hexoside	2	BLF, EDG, MNG, PBR, SCB, SDN
P56	-	5.01	463.0873	-1.3	$C_{21}H_{19}O_{12}^-$	301.0349 ($C_{15}H_9O_7^-$, **100**), 300.0274 ($C_{15}H_8O_7^-$, 44), 151.0039 ($C_7H_3O_4^-$, 1)	Hyperoside	1	CPN, EDG, ETH, FLN, HCH, NLT, RPN, RTH, 345, TNC
P57	609.1/301	5.01	609.1454	-1.1	$C_{27}H_{29}O_{16}^-$	591.1346 ($C_{27}H_{27}O_{15}^-$, 2), 343.0454 ($C_{17}H_{11}O_8^-$, 7), 301.0347 ($C_{15}H_9O_7^-$, **100**), 300.0273 ($C_{15}H_8O_7^-$, 42), 271.0244 ($C_{14}H_7O_6^-$, 5), 255.0296 ($C_{14}H_7O_5^-$, 3), 179.0030 ($C_8H_5O_5^-$, 2)	Rutin	1	AND, BLN, BLF, BLS, BCL, BSC, BRT, CPN, EDG, FLN, GLN, GRN, HCH, MNG, MN NLT, PBR, RPN, RTH, SCB, SDN, 345, 362, TNC
P58	477.1/302	5.03	477.0667	0.7	$C_{21}H_{17}O_{13}^-$	301.0352 ($C_{15}H_9O_7^-$, **100**)	Quercetin-*O*-glucuronide[87]	2	BLN, BSC
P59	533.1/302	5.03	533.1289	-0.2	$C_{25}H_{25}O_{13}^-$	515.1188 ($C_{25}H_{23}O_{12}^-$, 22), 503.0092 ($C_{24}H_{23}O_{12}^-$, 3), 473.1083 ($C_{23}H_{21}O_{11}^-$, 54), 455.0979 ($C_{23}H_{19}O_{10}^-$, 6), 443.0978 ($C_{22}H_{19}O_{10}^-$, **100**), 425.0870 ($C_{22}H_{17}O_9^-$, 7), 383.0768 ($C_{20}H_{15}O_8^-$, 14), 353.0663 ($C_{19}H_{13}O_7^-$, 13)	Apigenin-di-*C*-pentoside[93]	2	ARG, GRN, PBR, 362

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	[M-H]⁻ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P60	463.1/304	5.07	463.0872	-2.2	$C_{21}H_{19}O_{12}^-$	301.0349 ($C_{15}H_9O_7^-$, **100**), 300.0274 ($C_{15}H_8O_7^-$, 32), 151.0039 ($C_7H_3O_4^-$, 1)	Isoquercitrin[34]	1	AND, ARG, BLN, BLF, BLS, BCL, BSC, BRT, CPN, EDG, ETH, FLN, GLN, GRN, HCH, MRN,MNG,NMC, NMN, NLT, PLT, PBR, RCM, RPN, RTH, SCB, SDN, 345, 362, TXT, TNC, HYD
P61	-	5.09	667.1508	-0.4	$C_{29}H_{31}O_{18}^-$	625.1404 ($C_{27}H_{29}O_{17}^-$, 100), 607.1300 ($C_{27}H_{27}O_{16}^-$, 27), 547.1090 ($C_{25}H_{23}O_{14}^-$, 10), 445.0774 ($C_{21}H_{17}O_{11}^-$, 18), 355.0458 ($C_{18}H_{11}O_8^-$, 5), 301.0347 ($C_{15}H_9O_7^-$, 42), 300.0274 ($C_{15}H_8O_7^-$, 62), 271.0247 ($C_{14}H_7O_6^-$, 11)	Quercetin derivative	3	362
P62	309.1/306	5.10	309.0610	-1.8	$C_{14}H_{13}O_8^-$	193.0505 ($C_{10}H_9O_4^-$, **100**), 149.0610 ($C_9H_9O_2^-$, 1), 133.0144 ($C_4H_5O_5^-$, 12), 115.0038 ($C_4H_3O_4^-$, 2)	Feruloyl malate	2	AND, ARG, BLN, BLF, BLS, BCL, BSC, BRT, CPN, EDG, ETH, FLN, HCH,MRN, NMC, NMN, NLT, PLT, RCM, RPN, RTH, SDN, 345, 362, TNC
P63	579.1/308	5.13	579.1352	-0.6	$C_{26}H_{27}O_{15}^-$	459.0921 ($C_{22}H_{19}O_{11}^-$, 5), 447.0924 ($C_{21}H_{19}O_{11}^-$, 15), 429.0819 ($C_{21}H_{17}O_{10}^-$, 49), 327.0502 ($C_{17}H_{11}O_7^-$, 7), 285.0396 ($C_{15}H_9O_6^-$, 67), 284.0322 ($C_{15}H_8O_6^-$, **100**), 257.0452 ($C_{14}H_9O_5^-$, 6), 255.0295 ($C_{14}H_7O_5^-$, 15), 227.0346 ($C_{13}H_7O_4^-$, 2)	Kaempferol-*O*-xyloside-*C*-hexoside	2	BLN, BLF, BLS, BCL, GLN, GRN, RPN, RTH, SDN, 345, TNC
P64	447.1/309	5.15	447.0928	-1.1	$C_{21}H_{19}O_{11}^-$	327.0507 ($C_{17}H_{11}O_7^-$, 2), 285.0401 ($C_{15}H_9O_6^-$, **100**), 284.0327 ($C_{15}H_8O_6^-$, 1)	Kaempferol-*O*-hexoside[34]	2	AND, BLN, BLF, BLS, BCL, CPN, FLN, GLN, HCH, NLT, PLT, PBR, RCM, RPN, RTH, SCB, SDN, 345, 362, TXT, TNC, HYD
P65	-	5.19	533.1299	-1.6	$C_{25}H_{25}O_{13}^-$	515.1185 ($C_{25}H_{23}O_{12}^-$, 21), 473.1080 ($C_{23}H_{21}O_{11}^-$, 61), 443.0975 ($C_{22}H_{19}O_{10}^-$, **100**), 383.0766 ($C_{20}H_{15}O_9^-$, 11), 353.0661 ($C_{19}H_{13}O_7^-$, 12)	Unidentified	4	GRN

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	[M-H]⁻ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P66	505.1/314	5.23	505.0979	-1.8	$C_{23}H_{21}O_{13}^-$	487.0885 ($C_{23}H_{19}O_{12}^-$, 1), 463.0875 ($C_{23}H_{19}O_{12}^-$, 23), 445.0796 ($C_{21}H_{17}O_{11}^-$, 3), 343.0454 ($C_{17}H_{11}O_8^-$, 3), 301.0348 ($C_{15}H_9O_7^-$, **100**), 300.0273 ($C_{15}H_8O_7^-$, 56), 273.0399 ($C_{14}H_9O_6^-$, 1), 271.2045 ($C_{14}H_7O_6^-$, 1), 255.0295 ($C_{14}H_7O_5^-$, 1)	Quercetin-O-acetyl-hexoside[94]	2	AND, ARG, BLN, BLF, BLS, BCL, BRT, CPN, EDG, FLN, GLN, GRN, HCH, MRN, MNG, NMC, NMN, NLT, PLT, RCM, RTH, SDN, 345, TXT, TNC
P67	-	5.24	433.0759	-3.9	$C_{20}H_{17}O_{11}^-$	301.0349 ($C_{15}H_9O_7^-$, **100**), 300.0271 ($C_{15}H_8O_7^-$, 27), 151.0036 ($C_7H_3O_4^-$, 1)	Quercetin-O-pentoside	2	AND, ARG, HCH, TNC
P68	577.2/315	5.25	577.1546	-2.9	$C_{27}H_{29}O_{14}^-$	431.0979 ($C_{21}H_{19}O_{10}^-$, **100**), 430.0906 ($C_{21}H_{18}O_{11}^-$, 7), 285.0402 ($C_{15}H_9O_6^-$, 30)	Kaempferol-O-pentosyl-pentoside (Kaempferitrin)	2	ETH, PBR, 362
P69	593.2/317	5.28	593.1505	-1.2	$C_{27}H_{29}O_{15}^-$	447.0938 ($C_{21}H_9O_{11}^-$, 1), 327.0507 ($C_{17}H_{11}O_7^-$, 4), 285.0401 ($C_{15}H_9O_6^-$, **100**), 284.0326 ($C_{15}H_8O_6^-$, 8), 257.0453 ($C_{14}H_9O_5^-$, 4), 229.0505 ($C_{13}H_9O_4^-$, 2)	Kaempferol-O-rutinoside[89]	2	AND, BLN, BLF, BLS, BCL, BSC, BRT, EDG, GLN, GRN, HCH, MNG, NLT, PBR, RCM, RPN, RTH, SCB, SDN, 345, TNC
P70	607.1/317	5.28	607.1300	-0.8	$C_{27}H_{27}O_{16}^-$	545.1293 ($C_{26}H_{25}O_{13}^-$, 23), 505.0981 ($C_{23}H_{21}O_{13}^-$, 55), 463.0875 ($C_{23}H_{19}O_{12}^-$, **100**), 301.0350 ($C_{15}H_9O_7^-$, 5)	Quercetin derivative	3	CPN, GLN, GRN, MNG, PLT, RCM, RTH
P71	-	5.28	413.1444	-2.2	$C_{19}H_{25}O_{10}^-$	369.1545 ($C_{18}H_{25}O_8^-$, 12), 351.1440 ($C_{18}H_{23}O_7^-$, 32), 311.1129 ($C_{15}H_{19}O_7^-$, **100**), 269.1025 ($C_{13}H_{17}O_6^-$, 16), 143.0347 ($C_6H_4O_4^-$, 1), 125.0244 ($C_6H_5O_3^-$, 16)	Unidentified	4	CPN, 362
P72	-	5.34	531.1136	-1.6	$C_{25}H_{13}O_{13}^-$	369.0819 ($C_{16}H_{17}O_{10}^-$, **100**), 323.0766 ($C_{15}H_{15}O_8^-$, 4), 207.0295 ($C_{10}H_7O_5^-$, 3), 192.0110 ($C_9H_4O_5^-$, 1)	Fraxetin-di-O-hexoside	2	ARG
P73	-	5.34	677.1708	-2.2	$C_{31}H_{33}O_{17}^-$	531.1347 ($C_{22}H_{27}O_{15}^-$, **100**), 513.1242 ($C_{22}H_{25}O_{14}^-$, 2), 469.1343 ($C_{21}H_{25}O_{12}^-$, 2), 207.0362 (1)	Unidentified	4	ARG
P74/	447.1/321	5.35	447.0929	-0.8	$C_{21}H_{19}O_{11}^-$	327.0506 ($C_{17}H_{11}O_7^-$, 15), 299.0558 ($C_{16}H_{11}O_6^-$, 2), 285.0400 ($C_{15}H_9O_6^-$, **100**), 284.0324 ($C_{15}H_8O_6^-$, 81), 255.0297 ($C_{14}H_7O_5^-$, 8), 227.0348 ($C_{13}H_7O_4^-$, 2), 151.0037 ($C_7H_3O_4^-$, 3)	Kaempferol-O-hexoside	2	all
P75	-	5.35	461.0715	-2.0	$C_{21}H_{17}O_{12}^-$	285.0400 ($C_{15}H_9O_6^-$, **100**), 175.0247 ($C_6H_7O_6^-$, 2)	Kaempferol-O-glucuronide	2	BSC
P76	-	5.36	825.2084	-1.1	$C_{36}H_{41}O_{22}^-$	783.1968 ($C_{34}H_{39}O_{21}^-$, **100**), 741.1861 ($C_{32}H_{37}O_{20}^-$, 10), 661.1395 ($C_{30}H_{29}O_{17}^-$, 3), 609.1450 ($C_{27}H_{29}O_{16}^-$, 2), 489.1030 ($C_{23}H_{21}O_{12}^-$, 3), 343.0454 ($C_{17}H_{11}O_8^-$, 2), 301.0345 ($C_{15}H_9O_7^-$, 4), 300.0269 ($C_{15}H_8O_7^-$, 11)	Quercetin derivative	3	SCB

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	$[M-H]^-$ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P77	449.1/323	5.38	449.1092	0.6	$C_{21}H_{20}O_{11}^-$	287.0558 ($C_{15}H_{11}O_6^-$, **100**), 151.0038 ($C_7H_3O_4^-$, 6)	Eriodictyol-O-hexoside[34]	2	ARG, BLF, BLS, BCL, BSC, BRT, CPN, EDG, ETH, FLN, GRN, HCH, MNG,NMC,NMN, NLT, PLT, RTH, SDN, 345 HYD
P78	-	5.40	559.2025	-1.3	$C_{25}H_{35}O_{14}^-$	515.2133 ($C_{24}H_{35}O_{12}^-$, 1), 497.2019 ($C_{24}H_{33}O_{11}^-$, 9), 457.1709 ($C_{21}H_{29}O_{11}^-$, 45), 415.1605 ($C_{19}H_{27}O_{10}^-$, **100**), 269.1026 ($C_{16}H_{17}O_6^-$, 2)	Apigenin derivative	3	362
P79	-	5.45	297.0974	-2.1	$C_{14}H_{17}O_7^-$	179.0561 ($C_6H_{11}O_6^-$, **100**), 161.0457 ($C_6H_9O_5^-$, 9), 143.0352 ($C_6H_7O_4^-$, 4), 135.0453 ($C_8H_7O_2^-$, 16), 91.0555 (9)	Unidentified	4	MNG, NMC
P80	-	5.46	387.0721	-0.2	$C_{19}H_{15}O_9^-$	225.0402 ($C_{10}H_9O_6^-$, **100**), 207.0298 ($C_{10}H_7O_5^-$, 3), 179.0350 ($C_9H_7O_4^-$, 19), 135.0453 ($C_8H_7O_2^-$, 3),	Unidentified	4	RPN
P81	431.1/328	5.46	431.0976	-1.8	$C_{21}H_{19}O_{10}^-$	341.0662 ($C_{18}H_{13}O_7^-$, 1), 311.0558 ($C_{17}H_{11}O_6^-$, 4), 283.0608 ($C_{16}H_{11}O_5^-$, 1), 269.0452 ($C_{15}H_9O_5^-$, **100**), 268.0377 ($C_{15}H_8O_5^-$, 13)	Apigenin-O-hexoside	2	BLN, PBR, TXT
P82	463.1/328	5.46	463.0874	-1.7	$C_{21}H_{19}O_{12}^-$	301.0350 ($C_{15}H_9O_7^-$, **100**), 300.0272 ($C_{15}H_8O_7^-$,1)	Quercetin-O-hexoside	2	AND, ARG, BLN, FLN, GLN, PBR, RCM, RPN, RTH, 345, TXT
P83	417.1/329	5.48	417.0822	-1.2	$C_{20}H_{17}O_{10}^-$	327.0504 ($C_{17}H_{11}O_7^-$, 9), 285.0397 ($C_{15}H_9O_6^-$, 26), 284.0321 ($C_{15}H_8O_6^-$, **100**), 255.0294 ($C_{14}H_7O_5^-$, 4), 151.0037 ($C_7H_3O_4^-$, 2)	Kaempferol xylopyranoside[69]	2	AND, ARG, BLS, CPN, RTH, FLN, GLN, GRN, HCH, MRN, RCM, RPN, SDN, 345, TNC, HYD
P84	-	5.49	433.1132	-1.5	$C_{21}H_{21}O_{10}^-$	271.0608 ($C_{15}H_{11}O_5^-$, **100**)	Unidentified	4	BSC
P85	-	5.49	465.1028	-2.2	$C_{21}H_{21}O_{12}^-$	303.0502 ($C_{15}H_{11}O_7^-$, **100**), 151.0036 ($C_7H_3O_4^-$, 2)	Unidentified	4	BSC
P86	-	5.50	651.1566	-0.2	$C_{29}H_{31}O_{17}^-$	609.1446 ($C_{27}H_{29}O_{16}^-$, 100), 301.0346 ($C_{15}H_9O_6^-$, 11), 300.0269 ($C_{15}H_8O_7^-$,**100**)	Acetyl-rutin	2	TNC
P87	-	5.55	187.0974	-0.9	$C_9H_{15}O_4^-$	125.0973 ($C_8H_{13}O^-$,**100**)	Unidentified	4	BLS
P88	533.1/332	5.53	533.0923	-2.7	$C_{24}H_{21}O_{14}^-$	447.0926 ($C_{21}H_{19}O_{11}^-$, 1), 429.0829 ($C_{21}H_{17}O_{10}^-$, 1), 327.0506 ($C_{17}H_{11}O_7^-$, 2), 285.0402 ($C_{15}H_9O_6^-$, **100**), 284.0327 ($C_{15}H_8O_6^-$, 7), 257.0454 ($C_{14}H_9O_5^-$, 1)	Kaempferol-O-malonylhexoside[34]	2	ARG, BLF, BLS, BCL, BRT, EDG, ETH, FLN, GLN, HCH,MRN, MNG, NMC, NMN, NLT, PLT, RCM, RPN, RTH, SDN, TXT, TNC
P89	-	5.54	475.0874	-1.7	$C_{22}H_{19}O_{12}^-$	433.0796 ($C_{20}H_{17}O_{11}^-$, 4), 415.0665 ($C_{20}H_{15}O_{10}^-$,50), 301.0348 ($C_{15}H_9O_7^-$, 37), 300.0272 ($C_{15}H_8O_7^-$,**100**), 178.9988 ($C_8H_3O_5^-$, 2)	Quercetin derivative	3	ARG

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	[M-H]⁻ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P90	-	5.54	447.0931	-1.8	$C_{21}H_{19}O_{11}^-$	327.0506 ($C_{17}H_{11}O_7^-$, 3), 285.0400 ($C_{15}H_9O_6^-$, **100**), 284.0324 ($C_{15}H_8O_6^-$, 13), 151.0037 ($C_7H_3O_4^-$, 2)	Kaempferol-*O*-hexoside	2	BLN
P91	461.1/335	5.58	461.1079	-2.3	$C_{22}H_{21}O_{11}^-$	446.0844 ($C_{21}H_{18}O_{11}^-$, 94), 341.0659 ($C_{18}H_{13}O_7^-$, 5), 299.0554 ($C_{16}H_{11}O_6^-$, **100**), 298.0476 ($C_{16}H_{10}O_6^-$, 16), 297.0398 ($C_{16}H_9O_6^-$, 2), 284.0318 ($C_{15}H_8O_6^-$, 7), 283.0246 ($C_{15}H_7O_6^-$, 3)	Chrysoeriol-*O*-hexoside	2	HCH, MRN, PBR, RPN, 362, TXT
P92	-	5.62	637.1561	-0.2	$C_{32}H_{29}O_{14}^-$	491.1191 ($C_{23}H_{23}O_{12}^-$, 2), 329.0663 ($C_{17}H_{13}O_7^-$, **100**), 314.0430 ($C_{16}H_{10}O_7^-$, 9), 299.0198 ($C_{15}H_7O_7^-$, 3)	*O*-methyl-chrysoeriol-*O*-pentosyl-*O*-hexoside	2	RPN
P93	435.1/340	5.65	435.1296	-0.2	$C_{21}H_{23}O_{10}^-$	297.0769 ($C_{17}H_{13}O_5^-$, 1), 273.0768 ($C_{15}H_{13}O_5^-$, 100), 167.0352 ($C_8H_7O_4^-$, 2)	Phloretin-*O*-hexoside (Phlorizin)	1	AND, ARG, BLN, BSC,MRN, NMC, PLT, RCM, RPN, RTH, SCB, SDN, TXT, HYD
P94	-	5.66	415.1965	-2.4	$C_{20}H_{31}O_9^-$	369.1909 ($C_{19}H_{29}O_7^-$, 89), 225.0611 ($C_7H_{13}O_8^-$, 21), 179.0559 ($C_6H_{11}O_6^-$, **100**)	Unidentified	4	BLN, HYD
P95	-	5.68	515.1188	-1.3	$C_{25}H_{23}O_{12}^-$	369.0822 ($C_{16}H_{17}O_{10}^-$, **100**), 207.0297 ($C_{10}H_7O_5^-$, 6), 192.0063 ($C_9H_4O_5^-$, 2)	Coumarin derivative	3	ARG
P96	-	5.76	364.9964	-2.3	$C_{15}H_9O_6S^-$	285.0400 ($C_{15}H_9O_6^-$, **100**)	Kaempferol-sulfate	2	BSC
P97	-	5.77	517.134	-0.5	$C_{25}H_{25}O_{12}^-$	473.1145 ($C_{24}H_{25}O_{10}^-$, **100**), 307.0819 ($C_{15}H_{15}O_7^-$, 3)	Unidentified	4	MRN
P98	-	5.77	563.1032	-1.5	$C_{25}H_{23}O_{15}^-$	519.1132 ($C_{24}H_{23}O_{12}^-$, **100**), 477.1020 ($C_{22}H_{21}O_{12}^-$, 3)	Unidentified	4	NMN
P99	489.1/347	5.78	489.1035	-0.7	$C_{23}H_{21}O_{12}^-$	327.0509 ($C_{17}H_{11}O_7^-$, 1), 285.0400 ($C_{15}H_9O_6^-$, **100**)	Kaempferol-*O*-acetyl-hexoside	2	EDG, FLN, GLN, MRN, NMC, RPN, SDN, 345, TXT, TNC, HYD
P100	799.2/347	5.78	799.2302	-0.1	$C_{35}H_{43}O_{21}^-$	753.2229 ($C_{34}H_{41}O_{19}^-$, **100**), 283.0606 ($C_{16}H_{11}O_5^-$, 1)	Flavonoid-*O*-trisaccharide (2xhexoside,1xpentoside)	3	PBR
P101	-	5.79	693.1657	-1.3	$C_{31}H_{33}O_{18}^-$	651.1554 ($C_{29}H_{31}O_{17}^-$, **100**), 609.1450 ($C_{27}H_{26}O_{16}^-$, 3), 301.0334 ($C_{15}H_9O_7^-$, 3), 300.0271 ($C_{15}H_8O_7^-$, 21)	Rutin derivative	2	TNC
P102	-	5.80	503.1763	-1.1	$C_{22}H_{31}O_{13}^-$	485.1656 ($C_{22}H_{29}O_{12}^-$, 21), 425.1445 ($C_{20}H_{25}O_{10}^-$, 10), 413.1446 ($C_{19}H_{25}O_{10}^-$, 60), 395.1340 ($C_{19}H_{23}O_9^-$, 19), 383.1340 ($C_{18}H_{23}O_9^-$, 65), 365.1235 ($C_{18}H_{21}O_8^-$, 17), 353.1235 ($C_{17}H_{21}O_8^-$, 37), 323.1130 ($C_{16}H_{19}O_7^-$, **100**), 293.1026 ($C_{15}H_{17}O_6^-$, 95)	Unidentified	4	GRN
P103	473.1/350	5.83	473.1077	-2.6	$C_{23}H_{21}O_{11}^-$	413.0874 ($C_{21}H_{17}O_9^-$, 11), 311.0558 ($C_{17}H_{11}O_6^-$, 7), 269.0450 ($C_{15}H_9O_5^-$, **100**), 268.0377 ($C_{15}H_8O_5^-$, 48)	Apigenin-*O*-acetyl-hexoside	2	MRN,RCM, TXT, HYD

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	[M-H]⁻ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P104	581.2/350	5.83	581.1871	-0.8	$C_{27}H_{33}O_{14}^-$	461.1450 ($C_{23}H_{25}O_{10}^-$, 16), 449.1297 ($C_{18}H_{25}O_{13}^-$, **100**), 329.0874 ($C_{14}H_{17}O_9^-$, 36), 257.0818 ($C_{15}H_{13}O_4^-$, 99)	Pinocembrin-dihydrochalcone-O-hexoside-C-hexoside	2	SDN, RCM
P105	607.1/351	5.84	607.1298	-1.1	$C_{27}H_{27}O_{16}^-$	545.1292 ($C_{26}H_{25}O_{13}^-$, 36), 505.0980 ($C_{23}H_{21}O_{13}^-$, 37), 463.0874 ($C_{21}H_{19}O_{12}^-$, **100**), 301.0349 ($C_{15}H_9O_7^-$, 69)	Quercetin-O-derivative	3	ARG, MNG, PBR
P106	-	5.85	401.0870	-1.5	$C_{20}H_{17}O_9^-$	225.0400 ($C_{10}H_9O_6^-$, 4), 207.0295 ($C_{10}H_7O_5^-$, **100**), 193.0503 ($C_{10}H_9O_4^-$, 23). 163.0399 ($C_9H_7O_3^-$, 4)	Unidentified	4	RPN
P107	-	5.85	533.1867*	-1.6	$C_{22}H_{33}O_{14}^-$	487.1811 ($C_{22}H_{31}O_{12}^-$, 9), 293.0872 ($C_{11}H_{17}O_9^-$, **100**), 233.0663 ($C_9H_{13}O_7^-$, 100)	Unidentified	4	BLN
P108	619.2/352	5.86	619.1663	-0.8	$C_{29}H_{31}O_{15}^-$	473.1085 ($C_{23}H_{21}O_{11}^-$, 17), 431.0978 ($C_{21}H_{19}O_{10}^-$, **100**), 430.0903 ($C_{21}H_{18}O_{10}^-$, 17), 285.0403 ($C_{15}H_9O_6^-$, 42)	Kaempferol-(O-acetyl)-rhamnopyranosyl-O-rhamnopyranoside	2	ETH, RCM
P109	299.1/354	5.90	299.0554	-2.3	$C_{16}H_{11}O_6^-$	284.0324 ($C_{15}H_8O_6^-$, **100**)	Chrysoeriol	2	MRN,TXT
P110	-	5.93	515.1553	-1.1	$C_{26}H_{27}O_{11}^-$	471.1649 ($C_{25}H_{27}O_9^-$, 17), 369.1182 ($C_{17}H_{21}O_9^-$, **100**)	Unidentified	4	GLN
P111	301/362	6.03	301.0349	-1.6	$C_{15}H_9O_7^-$	283.0246 ($C_{15}H_7O_6^-$, 1), 273.0402 ($C_{14}H_9O_6^-$, 13), 257.0454 ($C_{14}H_9O_5^-$, 11), 239.0347 ($C_{14}H_7O_4^-$, 3), 229.0505 ($C_{13}H_9O_4^-$, 3), 193.0141 ($C_9H_5O_5^-$, 5), 178.9985 ($C_8H_3O_5^-$, **100**), 151.0037 ($C_7H_3O_4^-$, 84), 107.0140 ($C_6H_3O_2^-$, 4)	Quercetin	1	All
P112	333/365	6.08	333.0399	-1.6	$C_{19}H_9O_6^-$	305.0453 ($C_{18}H_5O_5^-$, 41), 289.0504 ($C_{18}H_9O_4^-$, **100**), 261.0555 ($C_{17}H_9O_3^-$, 31), 245.0606 ($C_{17}H_9O_2^-$, 26), 233.0606 ($C_{16}H_9O_2^-$, 1), 185.0244 ($C_{11}H_5O_3^-$, 1), 157.0296 ($C_{10}H_5O_2^-$, 2), 129.0346 ($C_9H_5O^-$, 1)	Bisumbelliferone	2	BLS, CPN, ETH, FLN, GLN, MRN, NLT, RTH, SDN, TNC
P113	431.1/370	6.16	431.0985	0.3	$C_{21}H_{19}O_{10}^-$	285.0400 ($C_{15}H_9O_6^-$, **100**), 284.0325 ($C_{15}H_8O_6^-$, 24), 257.0453 ($C_{14}H_9O_5^-$, 2), 241.0504 ($C_{14}H_9O_4^-$, 1), 151.0037 ($C_7H_3O_4^-$, 3)	Kaempferol-O-rhamnoside	2	ETH, FLN
P114	637.2/370	6.16	637.1441*	0.2	$C_{29}H_{33}O_{16}^-$	591.1710 ($C_{28}H_{31}O_{14}^-$, **100**), 283.0609 ($C_{16}H_{11}O_5^-$, 57), 268.0373 ($C_{15}H_8O_5^-$, 2)	Flavonoid-O-pentoside-O-hexoside	3	PBR
P115	-	6.16	1119.5592	0.5	$C_{54}H_{87}O_{24}^-$	-	Unidentified	4	NMN
P116	555.2/375	6.25	555.1505	-0.6	$C_{28}H_{27}O_{12}^-$	537.1393 ($C_{28}H_{25}O_{11}^-$, 41), 523.1238 ($C_{27}H_{23}O_{11}^-$, 27), 365.0877 ($C_{17}H_{17}O_9^-$, 7), 351.1079 ($C_{17}H_{19}O_8^-$, **100**), 337.0923 ($C_{16}H_{17}O_8^-$, 8)	Naphthoquinone-naphthalene conjugate	3	FLN, TNC
P117	687.2/376	6.25	687.1929	-0.2	$C_{33}H_{35}O_{16}^-$	525.1397 ($C_{27}H_{25}O_{11}^-$, **100**), 507.1292 ($C_{27}H_{23}O_{10}^-$, 9), 337.0925 ($C_{16}H_{17}O_8^-$, 8), 319.0820 ($C_{16}H_{15}O_7^-$, 1)	Naphthoquinone-naphthalene conjugate	3	MRN, RPN
P118	-	6.26	747.2135	-1.0	$C_{35}H_{39}O_{18}^-$	601.1761 ($C_{26}H_{33}O_{16}^-$, 16), 583.1655 ($C_{26}H_{31}O_{15}^-$, 17), 571.1658 ($C_{25}H_{31}O_{15}^-$, **100**)	Unidentified	4	EDG
P119	541.1/379	6.31	541.1340	-2.0	$C_{27}H_{25}O_{12}^-$	523.1241 ($C_{27}H_{23}O_{11}^-$, 7), 497.1449 ($C_{26}H_{25}O_{10}^-$, 5), 365.0863 ($C_{17}H_{17}O_9^-$, 1), 337.0924 ($C_{16}H_{17}O_8^-$, **100**), 203.0348 ($C_{11}H_7O_4^-$, 1), 175.0453 ($C_{10}H_7O_3^-$, 22)	Naphthoquinone-naphthalene conjugate	3	BLS, FLN, MRN, NLT, RTH, TNC

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	[M-H]⁻ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P120	513.1/381	6.35	513.1399	-0.7	$C_{26}H_{25}O_{11}^-$	477.1182 ($C_{26}H_{21}O_8^-$, 1), 453.1183 ($C_{24}H_{21}O_9^-$, 17), 423.1081 ($C_{23}H_{19}O_8^-$, 4), 351.0871 ($C_{20}H_{15}O_6^-$, 1), 325.0924 ($C_{15}H_{17}O_8^-$, 16), 307.0819 ($C_{15}H_{15}O_7^-$, **100**), 289.0713 ($C_{15}H_{13}O_6^-$, 6), 265.0715 ($C_{13}H_{13}O_6^-$, 41), 235.0609 ($C_{12}H_{11}O_5^-$, 36), 205.0505 ($C_{11}H_9O_4^-$, 4)	Unidentified	4	MRN, RPN
P121	-	6.39	1165.5627	-0.6	$C_{56}H_{89}O_{26}^-$	-	Unidentified	4	NMC
P122	-	6.43	1161.5684	-1.2	$C_{56}H_{89}O_{25}^-$	1119.5568 ($C_{54}H_{87}O_{24}^-$, 16), 1101.5459 ($C_{54}H_{85}O_{23}^-$, 29), 1015.5098 ($C_{50}H_{79}O_{21}^-$, **100**), 999.5157 ($C_{50}H_{79}O_{20}^-$, 22), 817.4371 ($C_{44}H_{65}O_{14}^-$, 17), 703.4422 ($C_{40}H_{63}O_{10}^-$, 26), 541.3892 ($C_{34}H_{53}O_5^-$, 9), 475.3574 ($C_{33}H_{47}O_2^-$, 11)	Acetyl P113	3	NMN
P123	419.1/388	6.46	419.1342	-1.3	$C_{21}H_{23}O_9^-$	299.0922 ($C_{17}H_{15}O_5^-$, 1), 281.0817 ($C_{17}H_{13}O_4^-$, 1), 257.0816 ($C_{15}H_{13}O_4^-$, **100**)	2,4,6-Trihydroxy-dihydrochalcone-*O*-hexoside	1	RCM
P124	273.1/391	6.51	273.0767	-0.5	$C_{15}H_{13}O_5^-$	179.0352 ($C_9H_7O_4^-$, 1), 167.0351 ($C_8H_7O_4^-$, **100**), 125.0247 ($C_6H_5O_3^-$, 4), 123.0454 ($C_7H_7O_2^-$, 3)	Phloretin	1	FLN, RCM,
P125	285/392	6.52	285.0403	-0.7	$C_{15}H_9O_6^-$	285.0403 ($C_{15}H_9O_6^-$, **100**), 257.0454 ($C_{14}H_9O_5^-$, 10), 243.0296 ($C_{13}H_7O_5^-$, 7), 241.0505 ($C_{14}H_9O_4^-$, 7), 229.0505 ($C_{13}H_9O_4^-$, 10), 213.0556 ($C_{13}H_9O_3^-$, 9), 185.0608 ($C_{12}H_9O_2^-$, 6), 169.0659 ($C_{12}H_9O^-$, 8), 151.0038 ($C_7H_3O_4^-$, 17)	Kaempferol	1	AND, ARG, BLN, BLF, BLS, BCL, BSC, BRT, CPN, EDG, ETH, FLN, GLN, GRN, HCH, MRN, NMC, NMN, NLT, PLT, PBR, RCM, RPN, RTH, SDN, 345, 362, TXT, TNC, HYD
P126	-	6.58	735.1774	-0.5	$C_{33}H_{35}O_{19}^-$	693.1661 ($C_{31}H_{33}O_{18}^-$, 100), 651.1555 ($C_{29}H_{31}O_{17}^-$, 8), 505.0977 ($C_{23}H_{21}O_{13}^-$, 3), 301.0355 ($C_{15}H_9O_7^-$, 3), 300.0271 ($C_{15}H_8O_7^-$, 15)	Quercetin derivative	3	TNC
P127	355.1/397	6.60	355.0819	-1.2	$C_{19}H_{15}O_7^-$	207.0295 ($C_{10}H_7O_5^-$, **100**), 163.0399 ($C_9H_7O_3^-$, 1)	Fraxetin conjugate	2	BLF, BCL, BRT, CPN, SDN
P128	-	6.61	1497.6797	2.7	$C_{69}H_{109}O_{35}^-$	1351.6144 ($C_{63}H_{99}O_{31}^-$, 6), 1333.6040 ($C_{63}H_{97}O_{30}^-$, 40), 1289.6281 ($C_{62}H_{97}O_{28}^-$, 1), 1183.5508 ($C_{58}H_{87}O_{25}^-$, 5), 1139.5586 ($C_{57}H_{87}O_{23}^-$, 1), 1071.5227 ($C_{56}H_{79}O_{20}^-$, 1), 925.4789 ($C_{47}H_{73}O_{18}^-$, 11), 779.4205($C_{41}H_{63}O_{14}^-$, 11), 761.4092 ($C_{41}H_{61}O_{13}^-$, 11), 717.4208 ($C_{40}H_{61}O_{11}^-$, **100**), 611.3587 ($C_{36}H_{51}O_8^-$, 7), 499.1662 ($C_{31}H_{47}O_5^-$, 20)	Unidentified	4	BLN
P129	525.1/403	6.71	525.1393	-1.7	$C_{27}H_{25}O_{11}^-$	337.0926 ($C_{16}H_{17}O_8^-$, **100**), 319.0820 ($C_{16}H_{15}O_7^-$, 35), 241.0508 ($C_{14}H_9O_4^-$, 3), 229.0505 ($C_{13}H_9O_4^-$, 4), 217.0504 ($C_{12}H_9O_4^-$, 2), 187.0451 (9)	Naphthoquinone-naphthalene conjugate	3	MRN
P130	-	6.78	1351.6172	-1.1	$C_{63}H_{99}O_{31}^-$	1333.6047 ($C_{63}H_{97}O_{30}^-$, 7), 1219.5717 ($C_{58}H_{91}O_{27}^-$, 4), 703.2286 (28), 647.3787 (**100**), 571.3626 ($C_{34}H_{51}O_7^-$, 4), 471.3473 ($C_{30}H_{47}O_4^-$, 3)	Unidentified	4	RTH, SDN
P131	-	6.89	1335.6228	0.1	$C_{63}H_{99}O_{30}^-$	1317.6082 ($C_{63}H_{97}O_{29}^-$, 7), 1203.5768 ($C_{58}H_{91}O_{26}^-$, 4), 793.4357 ($C_{42}H_{65}O_{14}^-$, 100), 611.3571 ($C_{36}H_{51}O_8^-$, 2), 499.3419 ($C_{31}H_{47}O_5^-$, 3)	Unidentified	4	BSC

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	[M-H]⁻ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P132	-	6.90	1321.6052	-2.0	$C_{62}H_{97}O_{30}^-$	1303.5393 ($C_{62}H_{95}O_{29}^-$, 7), 1189.5618 ($C_{57}H_{89}O_{26}^-$, 6), 647.3793 ($C_{36}H_{55}O_{10}^-$, **100**), 571.3635 ($C_{34}H_{51}O_7^-$, 4), 471.3474 ($C_{30}H_{47}O_4^-$, 4)	Unidentified	4	BLS
P133	455.1/419	6.98	455.1005	-2.7	$C_{20}H_{23}O_{10}S^-$	411.1111 ($C_{19}H_{23}O_8S^-$, 5), 375.1440 ($C_{20}H_{23}O_7^-$, 82), 286.9865 ($C_{10}H_7O_8S^-$, 1), 242.9965 ($C_9H_7O_6S^-$, 8), 207.0295 ($C_{10}H_7O_5^-$, **100**), 163.0399 ($C_9H_7O_3^-$, 35)	Fraxetin-sulfate derivative	2	BLF, BCL
P134	1393.6/426	7.10	1393.6270	-0.9	$C_{65}H_{101}O_{32}^-$	1375.6145 ($C_{65}H_{99}O_{31}^-$, 8), 1351.6143 ($C_{63}H_{99}O_{31}^-$, 17), 1333.6040 ($C_{63}H_{97}O_{30}^-$, 14), 703.2284 (19), 647.3785 ($C_{36}H_{55}O_{10}^-$, **100**), 571.3615 ($C_{34}H_{51}O_7^-$, 3), 471.3474 ($C_{30}H_{47}O_4^-$, 3)	Unidentified	4	BLS, RTH, SDN
P135	-	7.15	1363.6139	-2.7	$C_{64}H_{99}O_{31}^-$	1345.6025 ($C_{64}H_{97}O_{30}^-$, 4), 1321.6041 ($C_{62}H_{97}O_{30}^-$, 27), 1303.5936 ($C_{62}H_{95}O_{29}^-$, 31), 1231.5712 ($C_{59}H_{91}O_{27}^-$, 6), 647.3792 ($C_{36}H_{55}O_{10}^-$, **100**), 571.3633 ($C_{34}H_{51}O_7^-$, 4), 471.3475 ($C_{30}H_{47}O_4^-$, 3)	Unidentified	4	BLS
P136	1057.5/430	7.16	1057.5216	-0.8	$C_{52}H_{81}O_{22}^-$	911.4633 ($C_{46}H_{71}O_{18}^-$, 52), 895.4685 ($C_{46}H_{71}O_{17}^-$, 13), 877.4569 ($C_{46}H_{69}O_{16}^-$, 4), 713.3897 ($C_{40}H_{57}O_{11}^-$, 11), 669.3998 ($C_{39}H_{57}O_9^-$, 11), 641.4051 ($C_{38}H_{57}O_8^-$, **100**), 571.3634 ($C_{34}H_{51}O_7^-$, 10)	Unidentified	4	BLS, BSC, GRN, HCH, NMN, RTH, SDN
P137	-	7.33	985.5000	-1.4	$C_{49}H_{77}O_{20}^-$	925.4786 ($C_{47}H_{73}O_{18}^-$, 12), 853.4577 ($C_{44}H_{69}O_{16}^-$, **100**), 823.4476 ($C_{43}H_{67}O_{15}^-$, 14), 655.3837 ($C_{38}H_{55}O_9^-$, 4), 611.3949 ($C_{37}H_{55}O_7^-$, 17), 583.4000 ($C_{36}H_{55}O_6^-$, 37), 513.3579 ($C_{32}H_{49}O_5^-$, 7)	Unidentified	4	HYD
P138	-	7.33	1435.6367	-2.5	$C_{67}H_{103}O_{33}^-$	1417.6233 ($C_{67}H_{101}O_{31}^-$, 7), 1393.6251 ($C_{65}H_{101}O_{32}^-$, 34), 1375.6144 ($C_{65}H_{99}O_{31}^-$, 31), 1357.6055 ($C_{65}H_{97}O_{30}^-$, 4), 1303.5923 ($C_{62}H_{95}O_{29}^-$, 6), 745.2393 ($C_{34}H_{41}O_{17}^-$, 17), 703.2291 ($C_{34}H_{39}O_{16}^-$, 8), 647.3790 ($C_{36}H_{55}O_{10}^-$, **100**), 629.3690 ($C_{36}H_{55}O_9^-$, 5), 571.3631 ($C_{34}H_{51}O_7^-$, 3), 471.3475 ($C_{30}H_{47}O_4^-$, 3)	Unidentified	4	BLS, SDN
P139	-	7.37	1057.5223	-0.8	$C_{52}H_{81}O_{22}^-$	911.4633 ($C_{46}H_{71}O_{18}^-$, 53), 895.4685 ($C_{46}H_{71}O_{17}^-$, 14), 877.4569 ($C_{46}H_{69}O_{16}^-$, 4), 713.3897 ($C_{40}H_{57}O_{11}^-$, 11), 669.3998 ($C_{39}H_{57}O_9^-$, 11), 641.4051 ($C_{38}H_{57}O_8^-$, **100**), 571.3634 ($C_{34}H_{51}O_7^-$, 7)	Unidentified	4	NMN
P140	1477.6/446	7.43	1477.6460	-2.2	$C_{69}H_{105}O_{34}^-$	1459.6351 ($C_{69}H_{103}O_{33}^-$, 3), 1435.6358 ($C_{67}H_{103}O_{33}^-$, 86), 787.2498 (9), 745.2394 (16), 647.3788 ($C_{36}H_{55}O_{10}^-$, **100**), 585.3786 ($C_{35}H_{53}O_7^-$, 3)	Unidentified	4	BLS, SDN
P141	359.1/451	7.52	359.0555	-1.8	$C_{21}H_{11}O_6^-$	173.0242 ($C_{10}H_5O_3^-$, **100**)	Methylene-3,3'-bilawsone	2	BLS, CPN, FLN, GLN, MRN, NLT, RTH, SDN, TNC
P142	-	7.52	895.4683	-1.6	$C_{46}H_{71}O_{17}^-$	749.4104 ($C_{40}H_{61}O_{13}^-$, 47), 713.3887 ($C_{40}H_{57}O_{11}^-$, 47), 689.3901 ($C_{38}H_{57}O_{11}^-$, 21), 669.3995 ($C_{39}H_{57}O_9^-$, 39), 641.4047 ($C_{38}H_{57}O_8^-$, **100**), 571.3638 ($C_{34}H_{51}O_7^-$, 13), 321.0386 ($C_{12}H_{17}O_{10}^-$, 4)	Unidentified	4	NMN
P143	-	7.59	1011.5154	-2.0	$C_{51}H_{79}O_{20}^-$	865.4681 ($C_{45}H_{69}O_{16}^-$, 14), 785.4467 ($C_{44}H_{65}O_{12}^-$, 34), 746.4189 ($C_{41}H_{62}O_{12}^-$, **100**), 673.4310 ($C_{39}H_{61}O_9^-$, 14)	Unidentified	4	BSC

Table 2.2. (continue)

Peak no.	Feature	t_R (min)	[M-H]⁻ measured	Error (ppm)	Molecular formula	MS² product ions, m/z (elemental composition, rel. intensity [%])	Identification	Annotation level	Species
P144	-	7.63	977.2914	-1.9	$C_{45}H_{53}O_{24}^-$	935.2803 ($C_{43}H_{55}O_{23}^-$, 24), 755.2171 ($C_{37}H_{39}O_{17}^-$, **100**), 737.2054 ($C_{37}H_{37}O_{16}^-$, 10), 713.2071 ($C_{35}H_{37}O_{16}^-$, 14), 683.1947 ($C_{34}H_{35}O_{15}^-$, 11), 435.1288 ($C_{21}H_{23}O_{10}^-$, 91), 415.1238 ($C_{18}H_{23}O_{11}^-$, 45), 393.118 ($C_{19}H_{21}O_9^-$, 11), 301.0729 ($C_{16}H_{13}O_6^-$, 3)	Unidentified	4	TXT, TNC
P145	-	7.64	375.1440	-2.4	$C_{20}H_{23}O_7^-$	207.0294 ($C_{10}H_7O_5^-$, **100**), 163.0399 ($C_9H_7O_3^-$, 2)	Bi-napthoquinone derivative	2	BCL
P146	721.4/539	8.98	721.3641*	-1.5	$C_{34}H_{57}O_{16}^-$	675.3580 ($C_{33}H_{55}O_{14}^-$, **100**), 415.1447 ($C_{15}H_{27}O_{13}^-$, 3), 397.1342 ($C_{15}H_{25}O_{12}^-$, 13)	Gingerolglycolipid A II	2	all
P147	559.3/570	9.5	559.3121*	-0.5	$C_{28}H_{47}O_{11}^-$	513.3062 ($C_{27}H_{45}O_9^-$, **100**), 277.2170 ($C_{18}H_{29}O_2^-$, 61), 253.0927 ($C_9H_{17}O_8^-$, 9), 235.0821 ($C_9H_{15}O_7^-$, 1)	Glycolipid	3	all
P148	997.6/694	11.56	997.5716*	-2.6	$C_{52}H_{85}O_{18}^-$	951.5670 ($C_{51}H_{83}O_{16}^-$, **100**), 933.5535 ($C_{51}H_{81}O_{15}^-$, 2), 823.4256 ($C_{46}H_{63}O_{13}^-$, 41), 673.3433 ($C_{33}H_{53}O_{14}^-$, 4), 515.2233 (1), 397.1361 (1)	Unidentified	4	all
P149	835.5/728	12.12	835.5215*	0.4	$C_{46}H_{75}O_{13}^-$	789.5143 ($C_{45}H_{73}O_{11}^-$, **100**), 771.5038 ($C_{45}H_{71}O_{10}^-$, 13), 731.4725 ($C_{42}H_{67}O_{10}^-$, 1), 667.4050 ($C_{36}H_{59}O_{11}^-$, 1), 553.4255 ($C_{36}H_{57}O_4^-$, 1), 529.3008 ($C_{27}H_{45}O_{10}^-$, 3), 513.3059 ($C_{27}H_{45}O_9^-$, 40), 511.2904 ($C_{27}H_{43}O_9^-$, 16), 293.2118 ($C_{18}H_{29}O_3^-$, 5), 277.2169 ($C_{18}H_{29}O_2^-$, 16), 275.2012 ($C_{18}H_{27}O_2^-$, 5)	Glycolipid	3	all
P150	981.6/772	12.85	981.5799*	-1.1	$C_{52}H_{85}O_{17}^-$	935.5721 ($C_{51}H_{83}O_{15}^-$, **100**), 675.3597 ($C_{33}H_{55}O_{14}^-$, 8), 657.3482 ($C_{33}H_{53}O_{13}^-$, 31), 397.1319 ($C_{22}H_{21}O_7^-$, 2), 379.1237 ($C_{15}H_{23}O_{11}^-$, 4)	Glycolipid	3	all
P151	758.5/795	13.25	758.5413*	2.1	$C_{38}H_{78}O_{14}^-$	712.5358 ($C_{37}H_{76}O_{12}^-$, **100**)	Unidentified	4	all
P152	819.5/798	13.3	819.5257*	-0.7	$C_{46}H_{75}O_{12}^-$	773.5197 ($C_{45}H_{73}O_{10}^-$, 30), 513.3062 ($C_{27}H_{45}O_9^-$, 26), 495.2955 ($C_{27}H_{43}O_8^-$, 1), 277.2170 ($C_{18}H_{29}O_2^-$, **100**)	Glycolipid	3	all
P153	837.5/801	13.35	837.4825	-3.3	$C_{41}H_{73}O_{17}^-$	559.2579 ($C_{23}H_{43}O_{15}^-$, 40)	Unidentified	4	all
P154	959.6/812	13.53	959.5947*	-0.2	$C_{50}H_{87}O_{17}^-$	913.5876 ($C_{49}H_{85}O_{15}^-$, **100**), 675.3579 ($C_{33}H_{55}O_{14}^-$, 4), 657.3479 ($C_{33}H_{55}O_{13}^-$, 12), 653.3741 ($C_{31}H_{57}O_{14}^-$, 4), 635.3636 ($C_{31}H_{55}O_{11}^-$, 14), 397.1330, 379.1244 ($C_{15}H_{23}O_{11}^-$, 2)	Unidentified	4	all
P155	821.5/818	13.63	821.5411*	-1.1	$C_{46}H_{77}O_{12}^-$	775.5355 ($C_{45}H_{75}O_{10}^-$, 66), 515.3219 ($C_{27}H_{47}O_9^-$, 24), 513.3063 ($C_{27}H_{45}O_9^-$, 23), 279.2327 ($C_{18}H_{31}O_2^-$, **100**), 277.2172 ($C_{18}H_{29}O_2^-$, 86)	Unidentified	4	all
P156	815.5/875	14.56	815.4979	0.9	$C_{46}H_{71}O_{12}^-$	815.4983 ($C_{46}H_{71}O_{12}^-$, 39), 559.2579 ($C_{30}H_{39}O_{10}^-$, 39), 537.2736 ($C_{28}H_{41}O_{10}^-$, 11)	Unidentified	4	all

* Formic acid adduct

2.3.2.3 UHPLC-ESI-HRMS annotation of leaf metabolites

The most abundant peaks detected within the methanolic leaf extract of each species, as well as the eighty peaks which caused the main variance within the data set (Table 2.2, section 2.3.2.4), were further analyzed and annotated based on their fragmentation behavior revealing different compound classes. Mostly, they comprise carboxylic and phenolic acids, flavonoids, coumarins, dihydrochalcone derivatives, naphthoquinones and naphthalene conjugates (Table 2.2). A detailed description of the fragmentation behavior of the different compound classes is shown in the following paragraphs.

2.3.2.3.1 Flavonoid conjugates

Flavonoid glycoside conjugates represent a major compound class occurring within the *Impatiens* species, which were also detected in our study.[69] Derivatives of the known flavonoids kaempferol (P43, P46, P54, P55, P63, P64, P68, P69, P74, P75, P88, P90, P96, P99, P108, P113), quercetin (P40, P47, P53, P56, P57, P58, P60, P61, P66, P67, P70, P76, P82, P86, P89, P101, P105, P11), apigenin (P36, P44, P59, P78, P81, P103), chrysoeriol (P91, P92, P109), myricetin (P41), eridictyol (P77), and gallocatechin/epigallocatechin (P5, P9) were detected within the species under investigation (Table 2.2). Kaempferol-*O*-rutinoside (P69), kaempferol-*O*-hexoside (P74), kaempferol-*O*-malonylhexoside (P88), rutin (P57), and quercetin-*O*-hexoside (P60) were detected in a high number or even in all species (Table 2.2).

O-Glycosyl and *C*-glycosyl flavonoids could be directly differentiated due to their striking differences in the fragmentation pathways.[95] Besides, the aglycones, namely chrysoeriol (P109), quercetin (P111), and kaempferol (P125) as well as their glycosides were detected in different leaf extracts, respectively (Table 2.2).

<u>*O*-Glycosides and derivatives</u>

The fragmentation behavior of *O*-glycosides is mainly characterized by losses of the sugar unit leading to intense fragment ions of type [M-162]⁻ for hexose, [M-146]⁻ for deoxyhexose (rhamnose), [M-176]⁻ for glucuronide and [M-132]⁻ for pentose.[96,97] Within the most common peaks mono- and di-*O*-glycosides of quercetin (P40, P57, P58, P60), myricetin (P41), kaempferol (P54, P64, P68, P69, P74, P113), apigenin-*O*-hexoside (P81), and chrysoeriol (P91) were detected. Additionally, a flavonoid-*O*-trisaccharide (P100) and a flavonoid-*O*-pentoside-*O*-hexoside (P114) were detected in the methanolic leaf extract of *I. puberula* showing both an aglycone of *m/z* 283.0606 ($C_{16}H_{11}O_5^-$). Within the ESI-HRMS^2 data of P114 the additional loss of a methyl group was detected (fragment ion at *m/z* 268) suggesting the presence of a methylated flavone (Table 2.2).

Furthermore, also an acetylation of the sugar unit represented by a neutral loss of 42 amu or 60 amu for additional acetylation at the C6-position of the sugar unit could be detected during the analysis (P66, P99, P103, P108, Table 2.2). Besides acetylation, also a functionalization of the

sugar with a malonyl group (neutral loss 86 ($C_3H_2O_3$), P83) and another di-acid (P70, P105) was observed (Scheme 2.2).
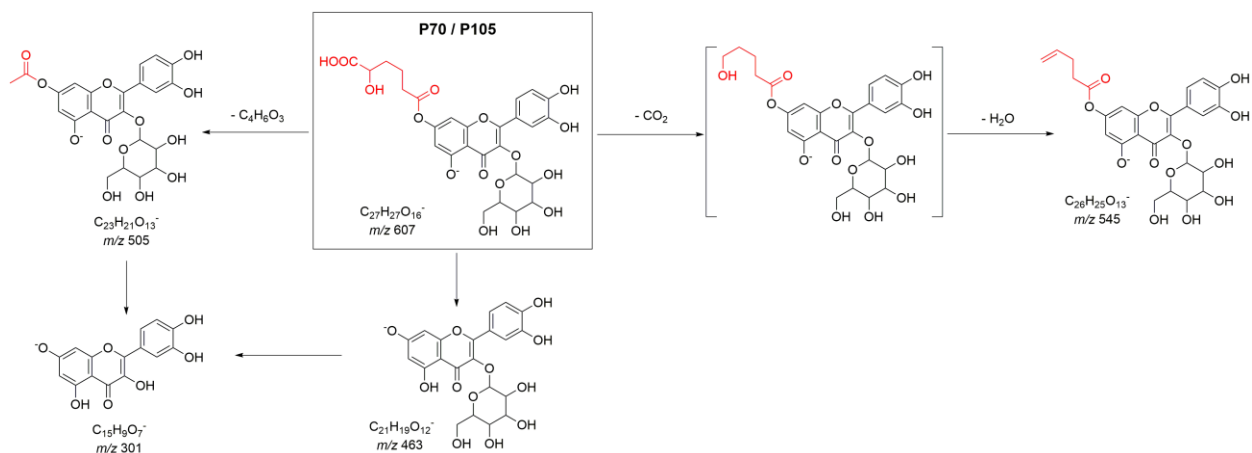
Scheme 2.2. Proposed mass spectral fragmentation of the [M-H]⁻ -ion of the non-identified quercetin derivatives P70 and P105.

The elemental composition of peak P46 was determined as $C_{21}H_{20}O_{14}S$ ([M-H]⁻ at *m/z* 527.0500, calculated for $C_{21}H_{19}O_{14}S^-$ 527.0501, see Table 2.2) based on the ESI-HRMS data. The presence of sulfur is indicated by the isotopic pattern, as well as, the loss of 80 Da (SO_3) observed in the ESI-MS^2 spectrum (Table 2.2). Sulfated flavonoids are known from different species like e.g. *Flaveria*.[98-100] Due to the presence of the fragment ions at *m/z* 447 and *m/z* 284 and the detection of the aglycone moiety within the extract of *Impatiens bisaccata* (P96, *m/z* 396.9964, calculated for $C_{15}H_{19}O_6S^-$ 396.9964), the compound was assigned as kaempferol-sulfate-*O*-hexoside.[101]

C-Glycosides and derivatives

The fragmentation pathway of *C*-glycosides is dominated by a cross-ring cleavage [(*O*-C1 and C2-C3) or (*O*-C1 and C3-C4)] of the sugar unit resulting in [M-H-120/90]⁻ for *C*-hexosides, [M-H-90/60]⁻ *C*-pentosides, and [M-H-104/74]⁻ for *C*-desoxyhexosides. Furthermore, fragment ions [aglycone + 41/71]⁻ in mono-*C*- and [aglycone + 83/113]⁻ for di-*C*-glycosides can be useful for the identification of the type of aglycone, which is linked to the remaining sugar fragment. In case of apigenin the corresponding fragment peaks appear at *m/z* 311/341 and *m/z* 353/383, respectively.[95,102,103]

The ESI-MS^2 spectra of peaks P36 (*m/z* 593.1497, $C_{27}H_{29}O_{15}^-$), P44 (*m/z* 563.1399, $C_{26}H_{27}O_{14}^-$), and P59 (*m/z* 533.1289, $C_{25}H_{25}O_{13}^-$) shared all the fragmentation patterns of apigenin-di-*C*-glycosides displaying typical fragment ions at *m/z* 353, 383, and 473 (Scheme 2.3). However, the linkage position of the sugar moiety could not be determined.
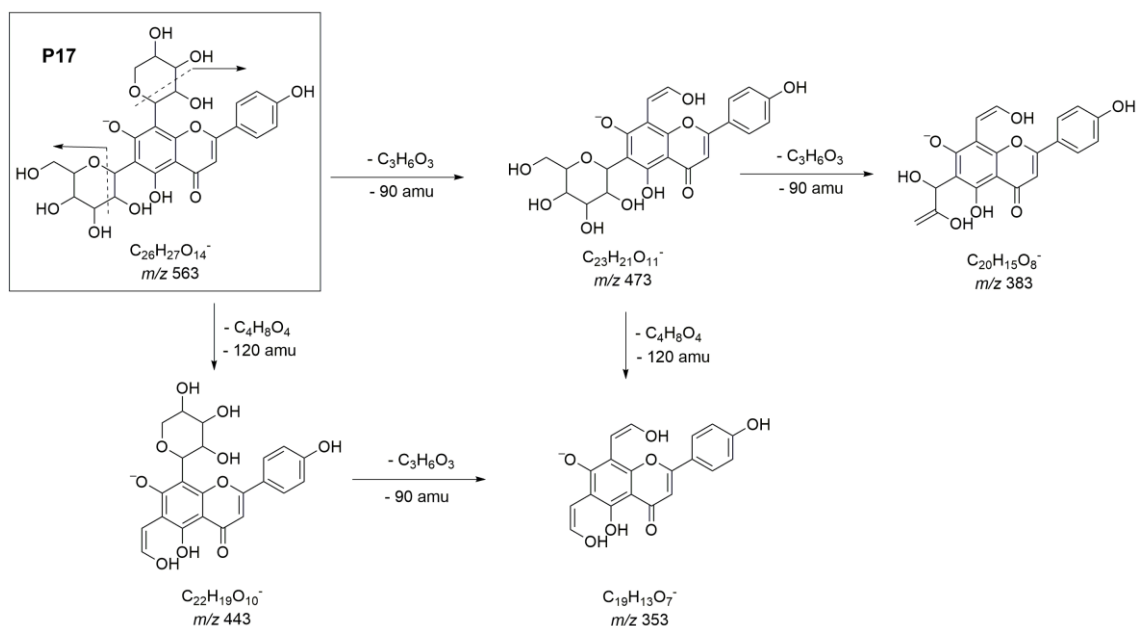
Scheme 2.3. Mass spectral fragmentation of the [M-H]⁻-ion of apigenin-*C*-hexoside-*C*-pentoside (P17).

C, O-Glycosides and derivatives

In addition, three mixed *C, O*-glycosylated flavonoids were detected within this investigation (P47, P55, P63). Based on the relative abundance of the major key fragments in these conjugates, it was possible to distinguish between an *O*-linked or a *C*-linked sugar position at the aglycone. In case of P63 showing a [M-H]⁻ ion at *m/z* 579.1352 (calculated for $C_{26}H_{27}O_{15}^-$, 579.1355, see Table 2.2) the appearance of an abundant ion at *m/z* 447 ([M-H-132]⁻) characterizes the loss of a pentose moiety which is *O*-linked to the aglycone. On the other hand, the characteristic fragment ion at *m/z* 459 ([M-H-120]⁻) being typically for the cross-ring cleavage of a *C*-hexoside, is further decomposed by the loss of $C_5H_8O_4$ indicating an *O*-connected pentose unit (*m/z* 327). The aglycone kaempferol shows an ion at *m/z* 284 (see Figure 2.4).
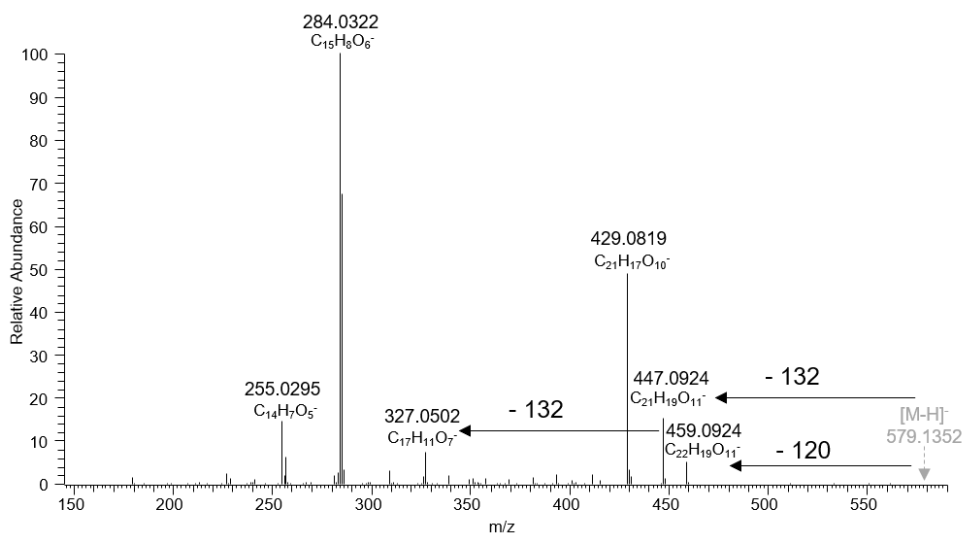
Figure 2.4. (-)-ESI-HRMS² spectra of *m/z* [M-H]⁻ 579 (P63), a kaempferol-*O*-pentoside-*C*-hexoside ($C_{26}H_{27}O_{15}^-$).

2.3.2.3.2 Coumarins

Fraxidin and Isofraxidin were previously detected in the leaf extract of *Impatiens balsamina* using LC-MS.[37] However, in our investigation only the glycosylated conjugate fraxin (P19) as well as other coumarin conjugates (P16, P35, P127, P133) and a biscoumarin (P112) were detected (Table 2.2). P16 and P35 both show a linkage to a hexose. In the (-)-ESI-HRMS^2 of P16, characteristic fragments of a *C*-glycoside were detected (Scheme 2.4 A), which are also known from flavonoids (see, *C-glycosides and derivatives*). Based on the fragment ions at *m/z* 159 and *m/z* 213, which loses a CH$_2$CO unit (*m/z* 171), a methyl group at position 4 might be suggested. Therefore, the compound was putatively assigned as 4-methylumbelliferyl-*C*-hexoside (P16, Scheme 2.4).
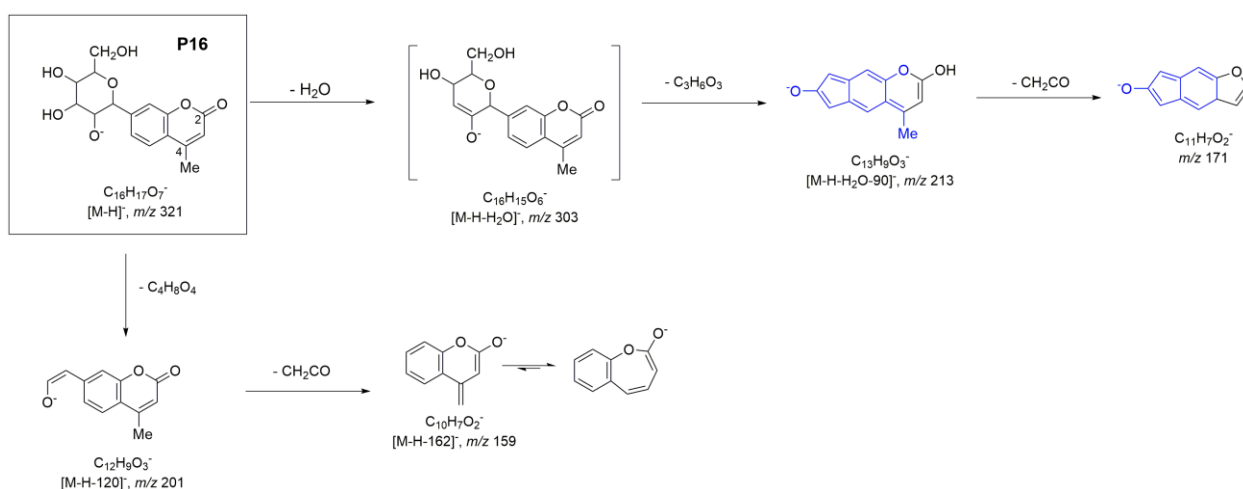
Scheme 2.4. Proposed mass spectral fragmentation of the [M-H]$^-$ ion of the non-identified peaks P16.

MS/MS fragmentation of the annotated peak P127, whose elemental composition could be determined as C$_{19}$H$_{16}$O$_7$ ([M-H]$^-$ at *m/z* 355.0819, calculated for C$_{19}$H$_{15}$O$_7^-$ 355.0823, see Table 2.2) leads to a base peak ion at *m/z* 207 (C$_{10}$H$_7$O$_5^-$, Scheme 2.5). This ion corresponds to a fraxetin moiety and forms an ion at *m/z* 163 (loss of CO$_2$). Probably, P127 represents an ester comprising fraxetin and phloretic acid (dihydrocoumaric acid). It can be assumed that the ester bond is easily splitted during the fragmentation. However, in that case more detailed MS^n investigations are necessary, e.g. by a positive ion MS^2 to detect the assumed phloretyl cation.
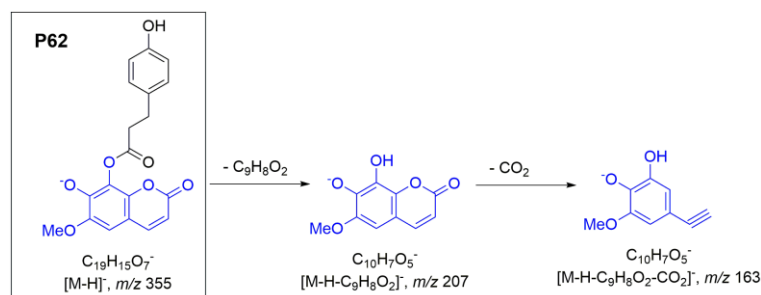
Scheme 2.5. Mass spectral fragmentation of the [M-H]$^-$ ion (*m/z* 355) of the proposed fraxetin derivative P62.

45

According to the high-resolution negative ion ESI measurements, the elemental composition of the peak P133 was determined as $C_{20}H_{24}O_{10}S$ ([M-H]$^-$ at m/z 455.1002, calculated for $C_{20}H_{23}O_{10}S^-$ 455.1017, see Table 2.2). The presence of sulfur is also shown by the isotopic pattern. In agreement to peak P127, fraxetin seems to be the core unit as indicated by the base peak ion at m/z 207 (Scheme 2.6., Table 2.2).
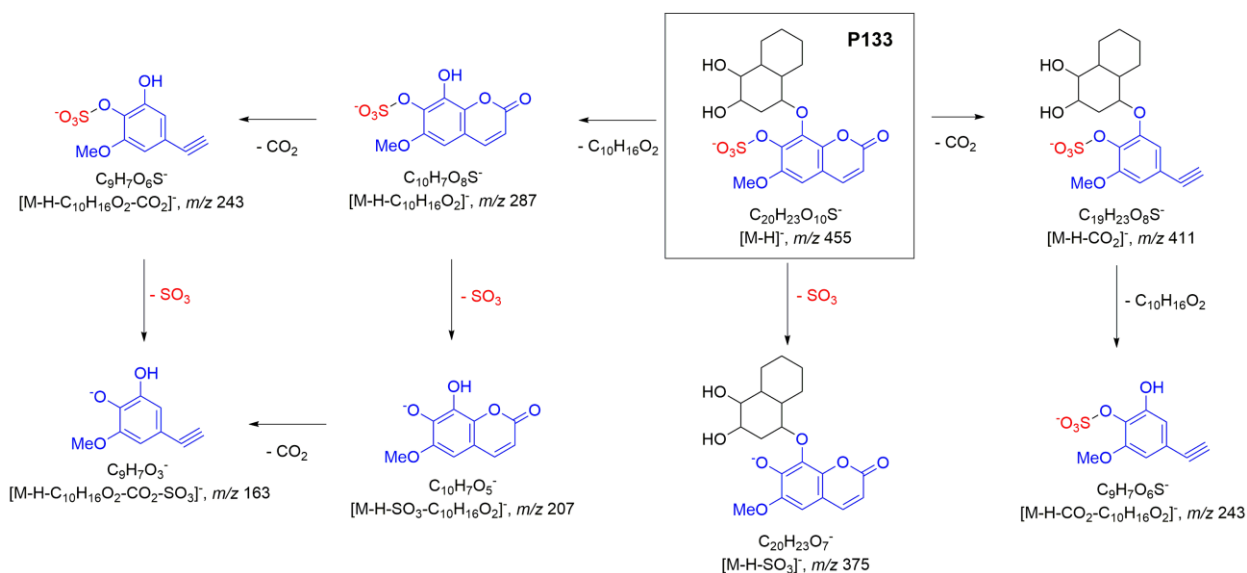
Scheme 2.6. Mass spectral fragmentation of the [M-H]$^-$ ion (m/z 455) of the proposed sulfated fraxetin derivative (P133).

While the expulsion of SO_3 leads to the abundant ion at m/z 375, the loss of 168 amu ($C_{10}H_{16}O_2$) yield an ion at m/z 287 indicating that the fraxetin moiety is sulfated. A trihydroxydecalin unit could be a possible component, which is attached to fraxetin. This is also supported by the ion at m/z 243 ($C_9H_7O_6S^-$) formed by loss of CO_2 from m/z 287. Further fragment ions being in agreement with the proposed structure are observed at m/z 411 ([M-CO$_2$]$^-$) and m/z 163 ([M-H-SO$_3$-C$_{10}$H$_{16}$O$_2$-CO$_2$]$^-$) (Scheme 2.6). However, a final proof can be given by a comparison of a MS^3 experiment of m/z 207 from compounds P127 and P133 with an MS^2 one of the [M-H]$^-$ ion (m/z 207) of a reference compound of fraxetin.

The elemental composition of the bisumbelliferone (P112) was determined as $C_{19}H_{10}O_6$ as deduced from the [M-H]$^-$ ion at m/z 333.0399 (calculated for $C_{19}H_9O_6$ 333.0405, see Table 2.2). This would correspond to a biscoumarin derivative consisting of two umbelliferones linked via a CH_2-unit. This would also be in agreement both with its MS^2 spectrum and the obtained UV data. The proposed mass spectral fragmentation of the deprotonated ion at m/z 333 is mainly characterized by primary losses of carbon monoxide and carbon dioxide leading to fragment ions at m/z 305 and m/z 289, respectively. Further consecutive fragmentation leads to key ions at m/z 245 ([M-2CO$_2$]$^-$), ion **a** at m/z 185 and m/z 157 ([**a**-CO]) (Scheme 2.7). However, the present structure is a hypothetical structure and the exact linkage of the two umbelliferones remains unknown.
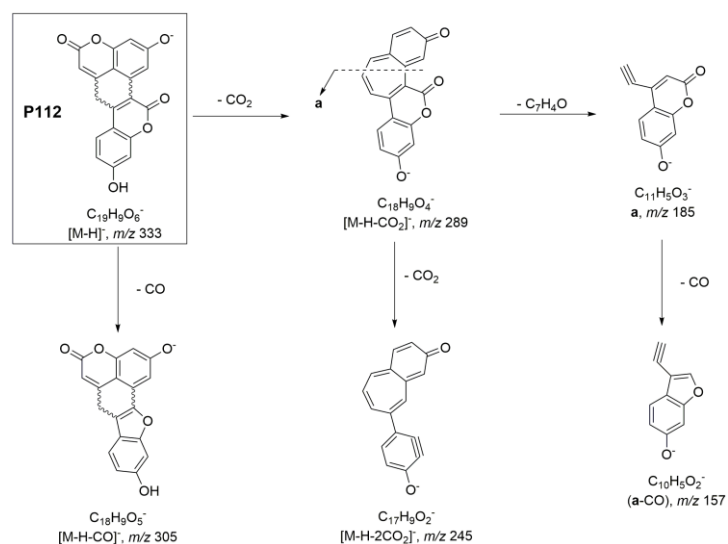
Scheme 2.7. Proposed mass spectral fragmentation of the [M-H]$^-$ ion of the bisumbelliferone methylene (P112).

2.3.2.3.3 Naphthoquinones and Naphthalene Conjugates

Peak P116 possess an elemental composition of $C_{28}H_{28}O_{12}$ ([M-H]$^-$ at *m/z* 555.1505, calculated 555.1508 for $C_{28}H_{27}O_{12}^-$, see Table 2.2). Its structure consists of three parts: 1) a dihydronaphthoquinone with a methoxy function, 2) a glucose and 3) balsaminolate, a 1,4-naphthoqunione derivative known from the corolla of *I. balsamina.*[104]
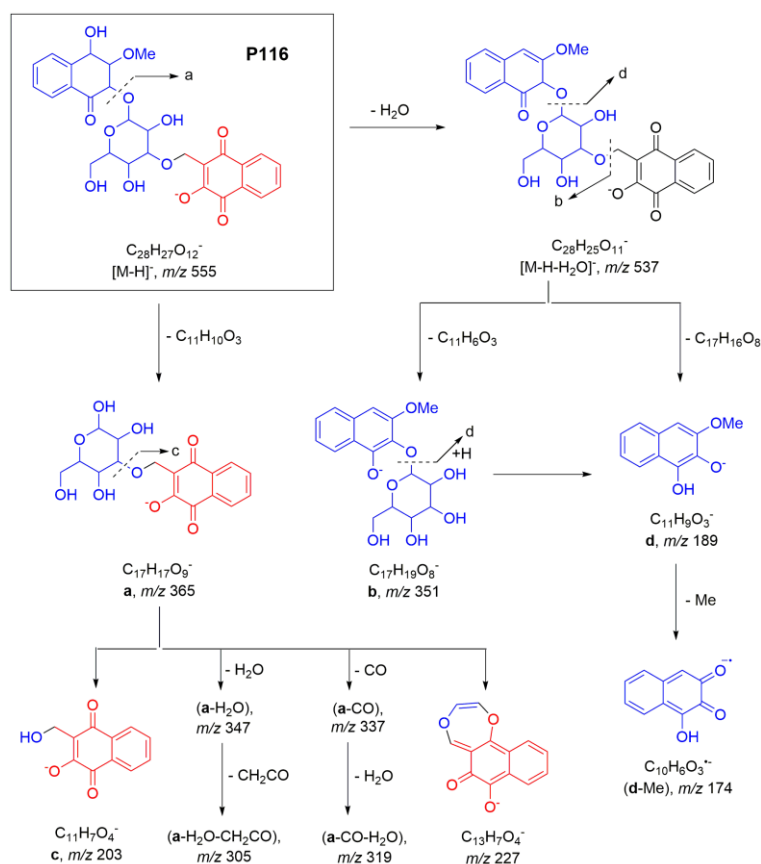
Scheme 2.8. Mass spectral fragmentation of the [M-H]- ion (*m/z* 555) of peak P116.

These three units are connected as given in Scheme 2.8. Ion **a** at *m/z* 365 comprising the glucose attached to the balsaminolate scaffold and is further decomposed to form a series of ions representing an evidence for the indicated partial structure (*m/z* 347, 337, 319, 305 and 227). Especially, the appearance of ion **c** at *m/z* 203 reflects one naphthoquinone moiety (balsaminolate). Another fragmentation pathway of the (M-H)⁻ ion (*m/z* 555) via a primary loss of H_2O (*m/z* 537) leads to the key ions **b** (*m/z* 351) and **d** (*m/z* 189) reflecting the connection of the other naphthoquinone core unit with the glucose. Ion **d** (*m/z* 189) can lose a methyl forming the ion at *m/z* 174 indicating the presence of a methoxy function at the naphthoquinone part.

Peak P119 (Scheme 2.9) possessing an elemental composition of $C_{27}H_{26}O_{12}$ ([M-H]⁻ at *m/z* 541.1340, calculated 541.1351 for $C_{27}H_{25}O_{12}^-$, see Table 2.2) represents a structurally related component compared with P116 (Scheme 2.8). It also consists of two naphthoquinone units and a glucose. However, in contrast to P116 in compound P119 the two naphthoquinones are linked. Both naphthoquinone units are reflected in the key ions at *m/z* 175 and *m/z* 203 (**d**). While ion **d** appears in agreement to compound P119, the fragment at *m/z* 175 indicates the core unit corresponding to 1,2,4-trihydroxynaphthalene. Therefore, based on the obtained (-)-ESI-HRMS[2] data, it can be assumed that compound P119 represents a 1,2,4-trihydroxynaphthalene-*O*-glucoside connected with the 1,4-naphthoquinone balsaminolate. This is supported by the fact, that the key ion at *m/z* 337 originated by loss of the core unit of balsaminolate would correspond to the [M-H]⁻ ion of 1,2,4-trihydroxynaphthalene-*O*-glucoside (P39).
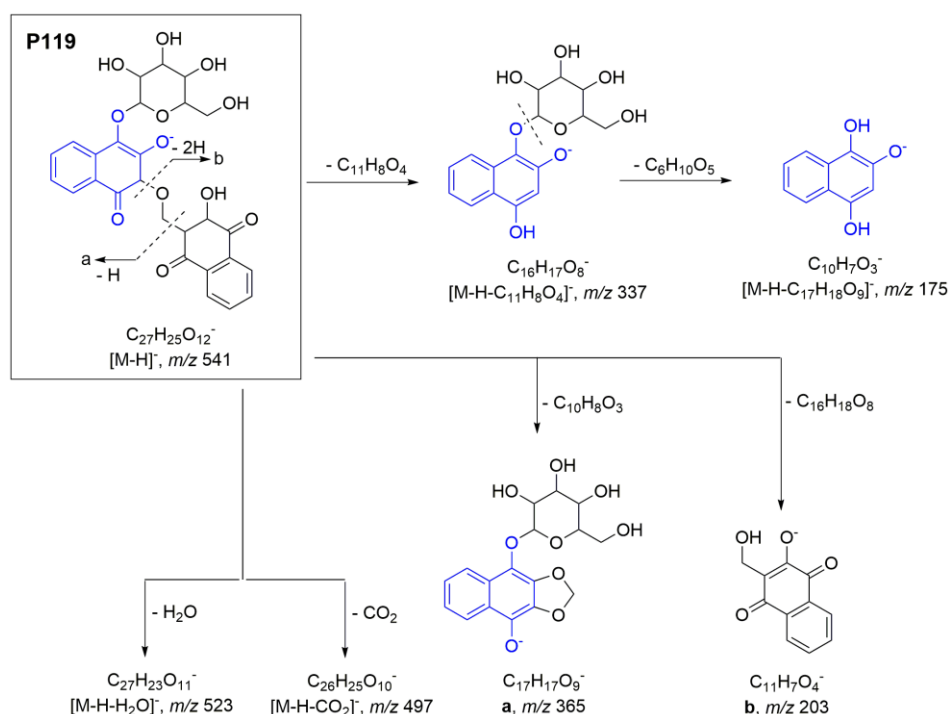
Scheme 2.9. Mass spectral fragmentation of the [M-H]- ion (*m/z* 541) of the trihydroxynaphthalene derivative P119.

2.3.2.3.4 Dihydrochalcones

The elemental composition of the dihydrochalcone glycoside phlorizin (P93, **2.5**, Scheme 2.1) was determined as $C_{21}H_{24}O_{10}$ as deduces from the [M-H]$^-$ ion at *m/z* 435.1296 (calculated for $C_{21}H_{23}O_{10}^-$ 435.1296, see Table 2.2). During collision induced fragmentation the loss of the sugar unit [M-H-162]$^-$ resulting in the fragment ion at *m/z* 273 (phloretin, [aglycone-H]$^-$) is observed. Key ions showing the phloroglucinol unit are found at *m/z* 167, 125 and 123. This was further supported by the fragmentation pattern of the aglycone phloretin (P124, **2.4**, Scheme 2.1) found in in the methanolic leaf extract of *I. racemosa*.
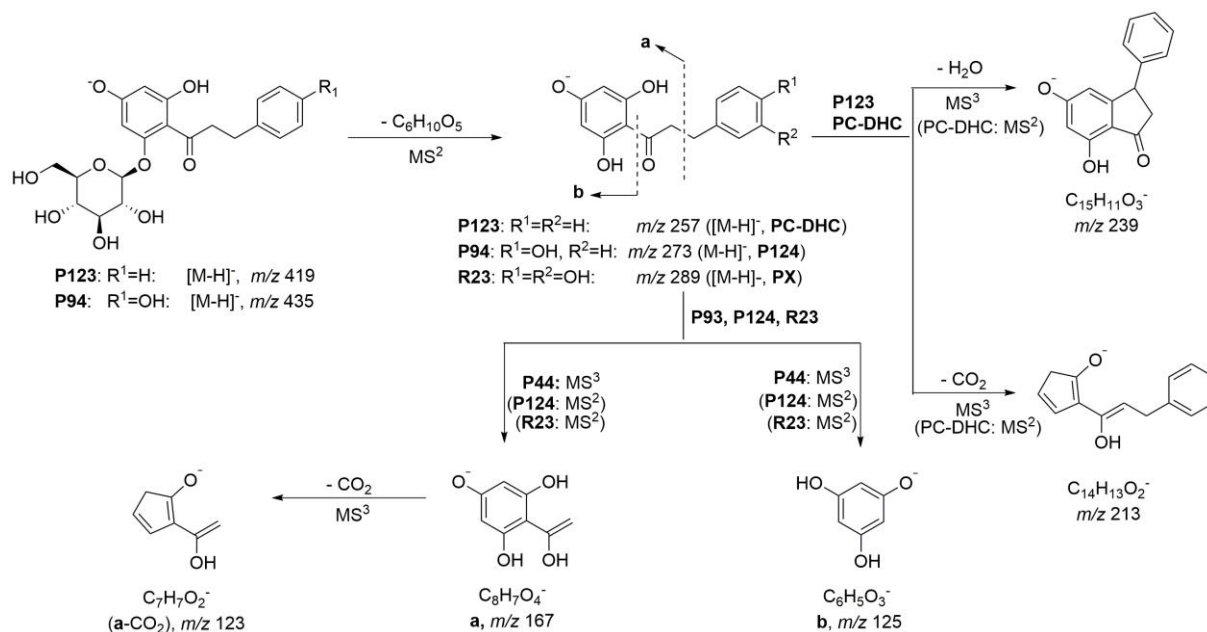
Scheme 2.10. Negative ion mass spectral fragmentation of the dihydrochalcone derivatives phlorizin (P94, **2.5**), pinocembrin dihydrochalcone glycoside (P123), phloretin (P124, **2.4**) and caffeoyl dihydrochalcone (R23, see section 2.3.5).

Peak P123 showed a similar fragmentation behavior as phlorizin. However, the base peak ion was detected at *m/z* 257.0816 (calculated for $C_{15}H_{13}O_4^-$ 257.0819, loss of sugar unit, see Table 2.2). Compared with phloretin, a mass difference of 16 amu was observed suggesting 2,4,6-trihydroxy-dihydrochalcone as aglycone. This was further supported by MS^3 data (Appendix, Figure A.49, A.50, A.56, A.59).

In case of P104 showing a [M-H]$^-$ ion at *m/z* 581.1871 (calculated for $C_{27}H_{33}O_{14}^-$, 581.1876, see Table 2.2) the appearance of an abundant ion at *m/z* 449 ([M-H-132]$^-$) characterizes the α-cleavage within the aglycone. On the other hand, the characteristic fragment ion at *m/z* 461 ([M-H-120]$^-$) being typical for the cross-ring cleavage of a *C*-hexoside, is further decomposed by the loss of C_9H_8O (*m/z* 329). The aglycone shows an ion at *m/z* 257, which is similar to P123. Therefore, the peak P104 was putatively assigned as pinocembrin-dihydrochalcone-*O*-hexoside-*C*-hexoside.

2.3.2.4 Chemometrics for the comparison of secondary metabolite profiles of methanolic leaf extracts

For the further chemical composition-based classification between *Impatiens* species and *Hydrocera triflora* chemometric tools were used defining both, similarities and differences among the samples. Although differences in the metabolic pattern could be observed by visual inspection of the total ion chromatograms of the investigated species (Figure 2.3), principal component analysis (PCA) and hierarchical cluster analysis (HCA) were used as unsupervised clustering methods to investigate the possible heterogeneities among the species and to cluster the UHPLC-ESI-HRMS profiles based on their similarities. These approaches reduce the dimensionality of a data set while preserving most of the variance within.[105-108]

The UHPLC-ESI-HRMS raw data were processed using xcms 3 (see section 3.5.3) to obtain a feature table including 7151 features and their intensities for further data analysis. A feature represents the combination of a mass-to-charge ratio (*m/z*) with a retention time. For feature determination, the parameters for peak picking and grouping were optimized using an intensity threshold of 1000.

2.3.2.4.1 Hierarchical cluster analysis (HCA)

Hierarchical clustering analysis reveals differences between classes. Applying this approach to the total ion chromatogram (TIC) data from the UHPLC-ESI-HRMS analysis a dendrogram with two very distinct clusters emerged (Figure 2.5). Within the two clusters, the technical replicates of each species clustered in general well together. However, one technical replicate of *I. andringitrensis*, *I. balfourii*, and *I. edgeworthii* clustered outside the group due to lower intensity in the TIC. In case of one replicate of *I. andringitrensis* the vial position was incorrect.

Within cluster B it is noticeable that *Impatiens racemosa* DC shows an independent cluster differentiating it from the other 12 species. This is in accordance with TLC observations, where its profile shows major differences to the other species (see section 2.3.1 Thin-layer chromatography (TLC)). The neighboring species *I. flanganae* showed low intensities of phloretin (P124, **2.4**, Table 2.2), which cause the close clustering to *I. racemosa*. Surprisingly, *Hydrocera triflora* does not form an individual cluster, although it is genetically different from *Impatiens* species. This indicates that both genera are closely related regarding their secondary metabolite profile. The two new species, *Impatiens* spec. (34558) and *Impatiens* spec. (36248) occur both in cluster A. Only species present in cluster B show the detected naphthoquinone and naphthalene conjugates (P42, P116, P117, P119, P129, P141; Table 2.2), which is one explanation for the distinction between the two groups and suggest this compound class as possible biomarkers for a subdivision. Further comparison of the metabolite profiles acquired using negative ion UHPLC-ESI-HRMS showed that the species occuring in cluster B show similiarities and high abundances in their occurring flavonoids (e.g. isoquercitrin (P60), kaempferol-*O*-hexosides (P74), and kaempferol (P125), Table 2.2), which was confirmed during PCA analysis (see section 2.3.2.4.2).
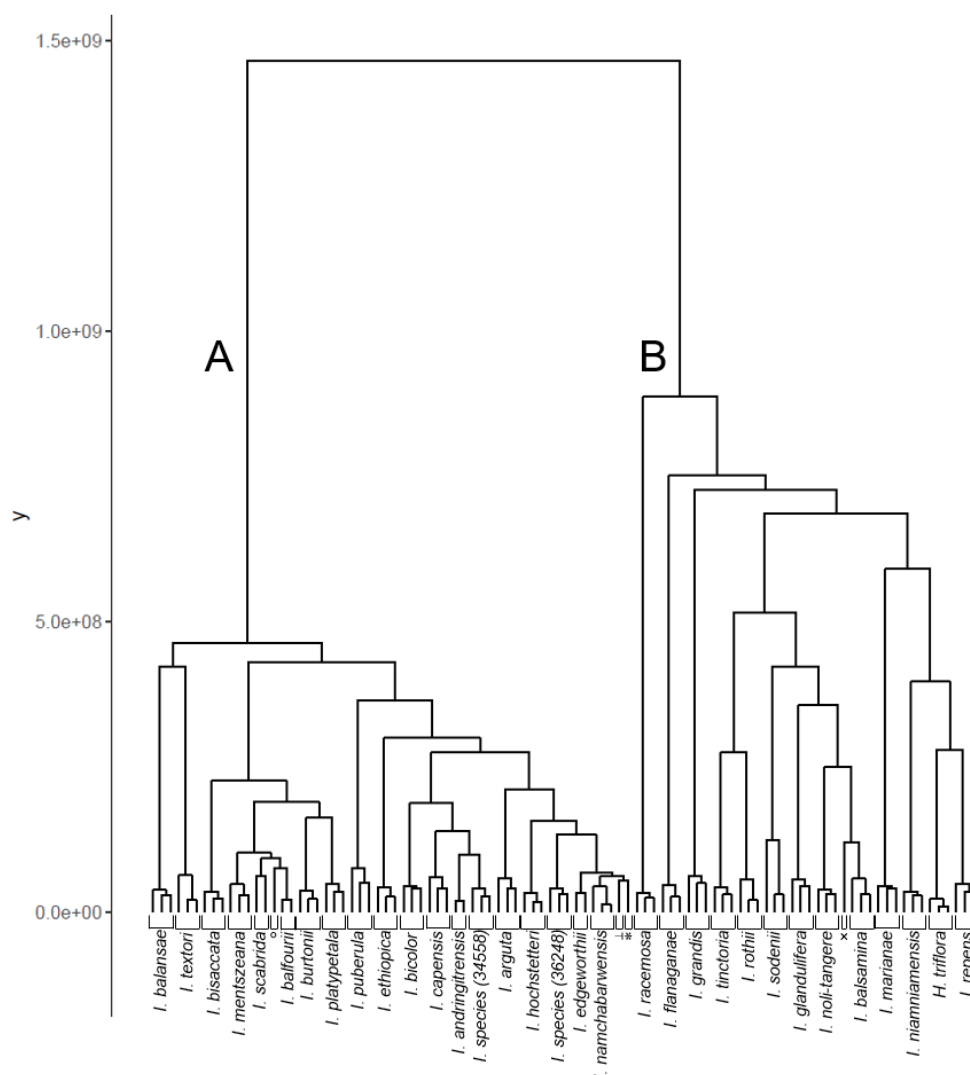
Figure 2.5. Hierarchical cluster analysis (HCA) of *Impatiens* species and *Hydrocera triflora* based on ward algorithm cluster of mass spectrometric profiles (n = 3, technical replicate of ⊥ = *I. andringitrensis*, ∗ = *I. balfourii*, ° = *I. edgeworthii*, and ˣ = *I. scabrida*).

In addition, meta data comprising the phylogenetic background, the ancestral origin (see Table 2.1), and biological activity information (based on our screening results, for details see section 2.3.4) was used to assess correlations within the recorded data. The hierarchical clusters do not reflect the phylogenetic background, and origin (Appendix, Figure A.5).

Staining of the species based on their tested antibiotic activity revealed difference between the effect against the gram-negative bacterium *Aliivibrio fischeri* and the gram-positive *Bacillus subtilis* at the higher concentration (Figure 2.6 II, IV). A clear concentration dependence effect appears to occur, since at the lower applied concentration there is a loss of activity along the tested species (Figure 2.6 I, II). Interestingly, at that lower concentration the methanolic leaf extracts of certain species, i.e. *I. ethiopica*, *I. flanaganae*, *I. glandulifera*, *I. textori*, and *I. tinctoria*, remain active against both tested microorganisms. Further discussion and considerations regarding biological activity is presented in section 2.3.4.
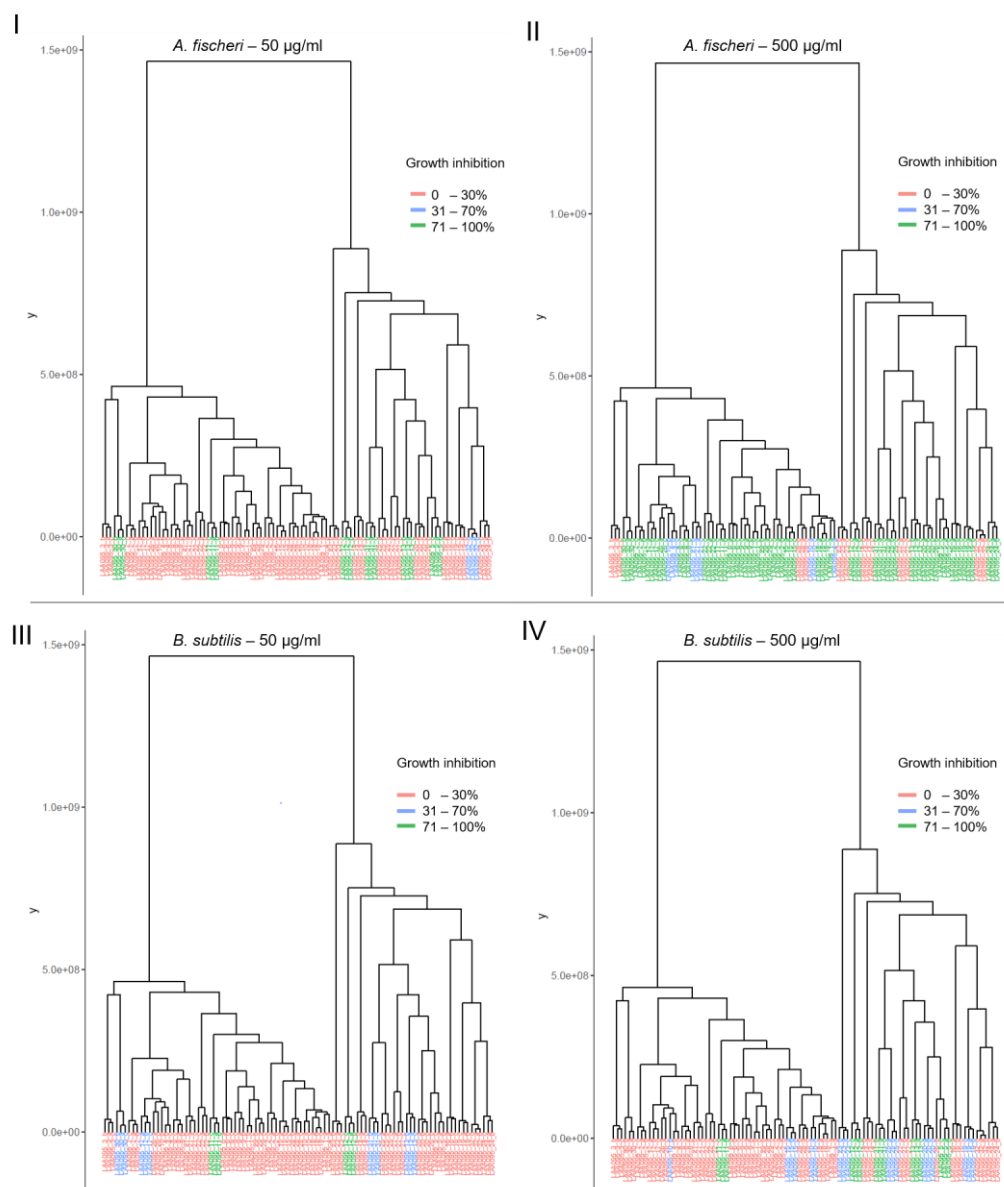
Figure 2.6. HCA colored by metadata based on the antibacterial activities of the methanolic leaf extracts against *Aliivibrio fischeri* (I – extract concentration 50 µg/mL, II – extract concentration 500 µg/mL) and *Bacillus subtilis* (III – extract concentration 50 µg/mL, IV– extract concentration 500 µg/mL).

The analyzed data set, acquired using negative ion UHPLC-ESI-HRMS, suggest that there is no direct correlation between the biological activity and the observed clusters within the HCA. This might indicate that the bioactivity is independent from the differentiating metabolites. However, the exclusive consideration of negative ion UHPLC-ESI-HRMS data is not sufficient and is limited to ionizing substances using this ionization mode. For example, 2-methoxy-1,4-naphthoquinone (**2.1**, Scheme 2.1) that shown to be responsible for antibacterial effects (section 2.3.6) does not ionize well in negative ion mode. A different clustering might occur when using positive ion mode or NMR data, which might lead to a well distinguished correlation of the active metabolites and the observed bioactivity. Therefore, NMR data and data from positive ion should be considered to overcome this limitation.

2.3.2.4.2 Principal component analysis (PCA)

In comparison to the HCA, the PCA is one of the most commonly used unsupervised dimensionality reduction tool in metabolic profiling to investigate the main variance and detect grouping trends and outliers in a data set. These variations of the dataset are visualized along the principal components (PCs), wherein the first PC constitutes the highest explained variation.[105] Comparable information responsible for clustering or differentiation of samples were obtained from the loadings plots of the PCA analysis. During data processing, different scaling methods were considered for the PCA analysis. Due to the large variance in the data set, scaling was omitted, which enormously increased the complexity of the loadings plot. The quality control sample measured within the batch nicely represents the different extracts by clustering in the middle of the samples (Appendix, Figure A.6).

For the analysis of the features being responsible for the main variance within this UHPLC-ESI-HRMS data set, the first ten principal components (PCs) were evaluated, explaining together 79.47% of the variance. The identified features (m/z / t_R (s)) and tentatively assigned compounds (numbered according to increasing retention time, for assignment and fragmentation schemes see section 2.3.2.3), as well as their occurrence within the investigated species are listed in Table 2.2. The assignment of a peak representing a metabolite in the MS profile of the leaf extract was based on the following criteria: intensity $> 1xE^4$, presence of the isotopic pattern and detection in all three replicates (threshold for processing 1000). Due to the complexity of the matrix, co-elution of some compounds happened, showing the same nominal mass but exhibiting differences in their decimals leading to different elemental compositions. Misleading data interpretation might occur during the data pre-processing regarding grouping steps and retention time correction, whereby isobaric peaks were partly picked together. Therefore, the data were checked manually to ensure the right annotation and differentiation of the features.

The untargeted UHPLC-ESI-HRMS analysis showed that the metabolite profiles of the investigated *Impatiens* are highly variable with little in common. The observed great chemical diversity within the different species is also underlined by a large number of features in the data set (7151 features).

The metabolome clusters were located at different points in the 2D space described by two vectors, principal component 1 (PC1), accounting for 20.3% of variance, and PC2 which explained 13.8% of the variance (Figure 2.7) covering only 34.1% of the main variance within the data set. This shows that the PCA analysis might not be the most suitable method for evaluating this data set due to the highly diverse composition of the methanolic leaf extracts of the investigated species. Nevertheless, the PCA based on the complete feature list shows a species differentiation by specific constituents (Figure 2.7, Table 2.2). 11 out of 32 species are separated from the others using PC1 and PC2. Generally, all three technical replicates of each species clustered together showing the reproducibility of sample preparation and measurement. Examination of the loadings

plot refers to the metabolites present along the investigated species, mainly caused by differences in their abundances within the extracts (Figure 2.7 B).
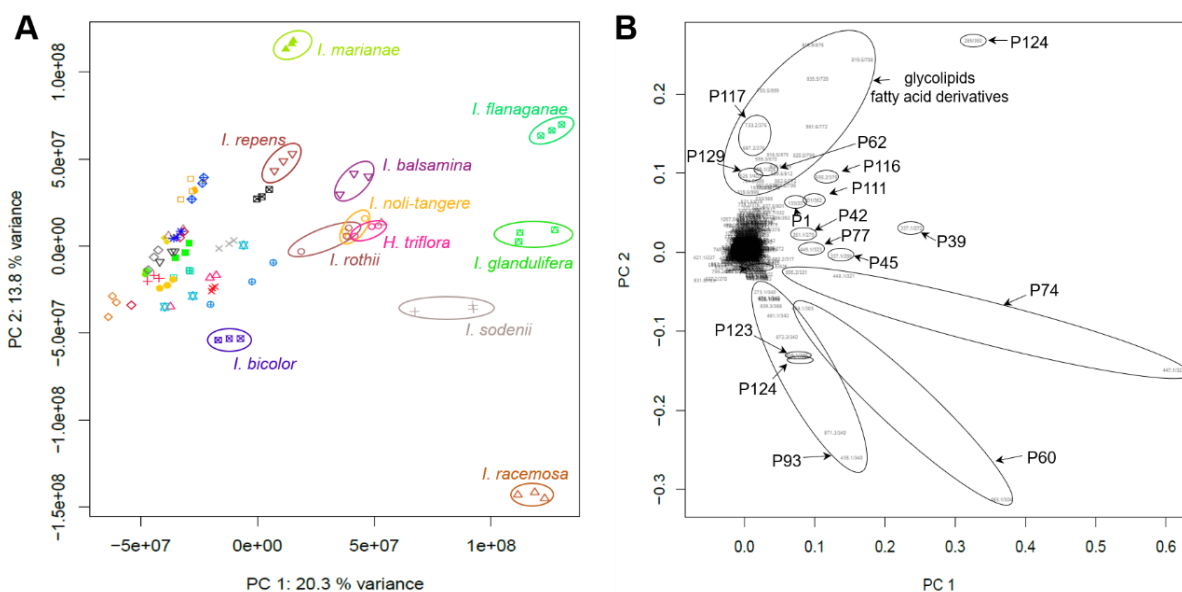
Figure 2.7. A) PCA scores plot (components 1 and 2). Classes are defined according to the species and B) PCA loadings plot (components 1 and 2). Shown features correspond to measured mass-over-charge ratios (*m/z*) and retention times in seconds and are numbered according to Table 2.2.

Impatiens racemosa as well as *I. sodenii*, *I. glandulifera*, *I. flanaganae*, *I. rothii*, *H. triflora*, *I. noli-tangere, and I. balsamina* can be differentiated from the other examined species in principal component 1 (PC1, Figure 2.7 A). Likewise, all these species also occur in the same cluster (cluster B) in the performed hierarchical cluster analysis (Figure 2.5). This distinctive differentiation in PCA and HCA is related to higher abundances of flavonoids and the occurrence of naphthoquinone and naphthalene derivatives.

In detail, *I. racemosa* is clearly separated due to its content of dihydrochalcone derivatives (435.1/340 (P93), 419/388 (P123), 273/39 (P124), Table 2.2), which were found to be unique metabolites for this species. These compounds were isolated from the crude methanolic extract and their structures confirmed by NMR and ESI-HRMS[n] (see section 2.3.5). To our knowledge, this is the first report on the occurrence of this substance class within *Impatiens* species. Furthermore, *I. racemosa* showed the highest abundance of quercetin-*O*-hexoside (P60, Table 2.2), although this metabolite is present in all extracts. Kaempferol-*O*-hexoside (P74) is distributed over nearly all species showing increased abundancies within *I. sodenii*, *I. racemosa*, *I. glandulifera*, *I. flanaganae* and *H. triflora*. *I. sodenii*. Furthermore, *I. sodenii* showed an increased level of 1,2,4-trihydroxynaphthalene-1-*O*-β-D-glucoside (P39, **2.2**, Scheme 2.1), which was isolated before from *I. glandulifera*,[58] as well as of peak P45, which was putatively assigned as 2-coumaroyl isocitric acid (Table 2.2). The distant clustering of *I. glandulifera* is mostly attributed to an abundance of a naphthoquinone derivative (P42, Table 2.2). The methanolic leaf extract of *I. flanaganae* exhibited high intensities of the flavonoids quercetin (P111) and kaempferol (P125).

Additionally, the unique feature 555.2/375 (P116), suggesting a new naphthoquinone-naphthalene conjugates, was detected for this species (Scheme 2.8, Table 2.2). Along PC2 *I. marianae*, *I. repens* and *I. bicolor* show distinct clusters. P117 and P129, which were also putatively annotated as new naphthoquinone-naphthalene conjugate were detected exclusively in the leaf extracts of *I. marianae* and *I. repens*. The present glycolipids and fatty acid derivatives (P146-P156) as well as feruloyl malate (P62) are present in comparable levels in all extracts of *Impatiens* and *Hydrocera triflora* (Table 2.2), and thus do not contribute to any differentiation.

In PC3 (Figure 2.8 A/B), *I. grandis* shows a distinct cluster due to its unique metabolite mangiferin (P25) and the high abundance of kaempferol-*O*-rutinoside (P69), as well as apigenin-*C*-hexoside-*C*-pentoside (P44). Metabolite P119, which showed a fragmentation behavior of a naphthoquinone-naphthalene conjugate (Scheme 2.9.), is responsible for the clustering of *I. tinctoria*, *I. noli-tangere*, and *I. rothii*. Along PC5/6 no new clusters and responsible feature were observed (Appendix, Figure A.7).
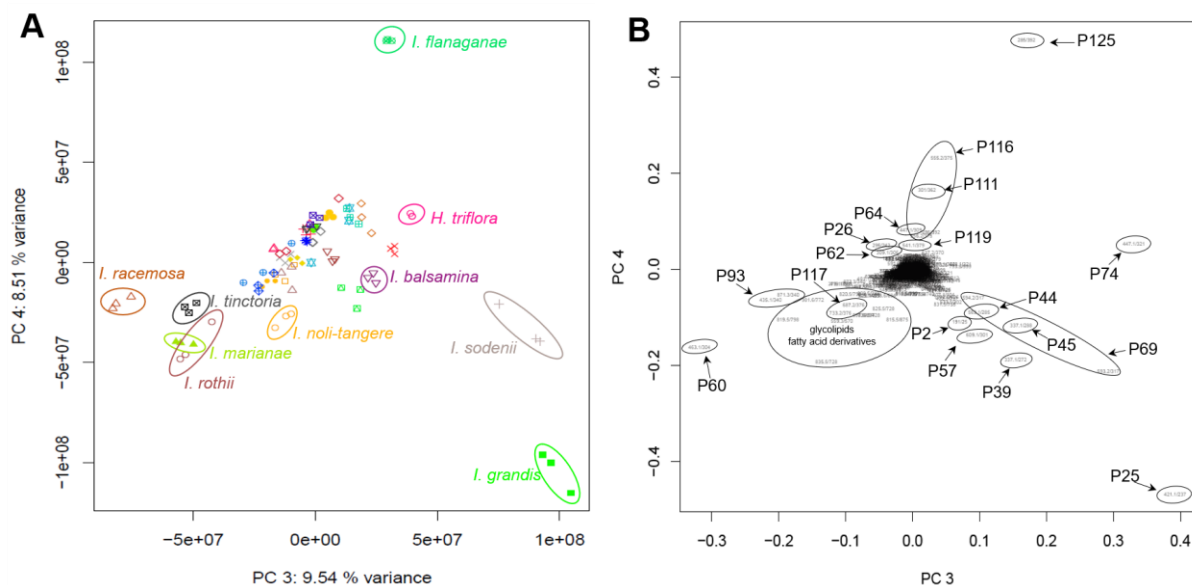
Figure 2.8. PCA scores (A) and loadings plots (B) of PC3/4 (A, B). Classes are defined according to the species.

In PC7 (Appendix, Figure A.8) *I. textori*, *I. marianae* and *I. puberula* show delimited clusters mainly because of their increased content of apigenin glycosides (P103, P81, P36), kaempferol-*O*-glycosides (P64, P88, P99), and chrysoeriol conjugates (P109, P91). *I. niamniamensis* clusters because of high abundance of the unknown metabolite P136 (1057.5/430) in its leaf extract. Additionally, compounds found during the analysis of PC1/2, *I. glandulifera* showed high intensity of eriodictyol-*O*-hexoside (P77). Besides a high number of flavonoid derivatives also coumarins are detected during the UHPLC-ESI-HRMS analysis. Peak P112, putatively annotated as a biscoumarin (Scheme 2.7, Table 2.2), was highly abundant in the methanolic extract of *I. marianae*, *I. flanaganae* and *I. tinctoria*. Further peaks related to coumarin-based compounds were observed in PC9 and 10 (Appendix, Figure A.8). Fraxin (P7, Table 2.2) was detected with a high abundance in the leaf extract of *I. arguta*. Fraxetin, the aglycone of fraxin, was found as core

structure of peaks P127 and P133 during the analysis of their (-)-ESI-HRMS[2] spectra (Scheme 2.5, Scheme 2.6., Table 2.2). Furthermore, the sulfated derivative (P133) was identified as unique metabolite of *I. bafourii* and *I. bicolor*.

In PC9 *I. balansae* showed a distinct cluster (Appendix, Figure A.8). The methanolic leaf extract showed peaks related with quercetin-*O*-glucuronide (P58), which is further present only in *I. bissacata* with a low abundancy (Table 2.2), as well as a high content of quercetin-*O*-hexoside (P82). P35, a possible coumarin derivative (Scheme 2.4), is contributing to the clustering of this species. The kaempferol-*C*/*O*-glycoside (P63, Figure 2.4) was found in high abundance for *I. balansae* and *I. bicolor*.

Furthermore *I. ethiopica* and *I. puberula* showed a clear differentiation to the other species. *I. ethiopica* clustered due to its content and high abundances of different kaempferol derivatives (P54, P68, P108, P113). For *I. puberula*, P44 (apigenin-*C*-hexoside-*C*-pentoside, Table 2.2), P100 (a flavonoid-*O*-trisaccharide, Table 2.2) and P114 (a flavonoid-*O*-pentoside-*O*-hexoside, Table 2.2) were identified as responsible for the cluster.

Multivariate data analysis using PCA led to a differentiation between the species caused mainly by differences in the relative abundance of peaks corresponding to flavonoid derivatives. However, species that contain unique/specific metabolites like *Impatiens racemosa* (e.g. dihydrochalcone derivatives) or *I. grandis* (e.g mangiferin), separate as distinct clusters in PC1-PC4. However, the clustering does not show any specific patterns associated to the applied metadata (Table 2.1), which is a reflection of the big differences along the investigated species (Appendix, Figure A.9--A.13). In order to complement the outcome of this study more specified analysis can be conducted, e.g. analysis of individual substance classes such as flavonoids and naphthoquinones within the species. Further insights for a comprehensive view on the investigated species can be achieved by performing PCA and HCA on positive ion mode UHPLC-ESI-HRMS and NMR data.

2.3.3. ¹H-NMR profiling

¹H-NMR was used as a complementary analytic technique, seeking for further information not supplied by UHPLC-ESI-HRMS. Beside quantitative information of the crude extracts under investigation, ¹H-NMR profiling might lead to the detection of compounds that barely ionize in MS negative mode. As depicted in Figure 2.9., the ¹H-NMR spectra show a very diverse profile. This confirms the results of the UHPLC-ESI-HRMS screening, pointing out the high variability of metabolites within the different species under investigation.

The spectra show characteristic signals for primary and secondary metabolites present in the methanolic leaf extracts (Figure 2.9). Due to the complexity of the extracts, signals are overlapping especially in the aromatic (yellow) and sugar (blue) region. Fatty acids are characterized by the two signals at δ 1.31-1.33 ppm and δ 0.88-0.91 ppm related to the methylenes and the terminal methyl group. Based on the proton multiplicity and chemical shift, ¹H-NMR also confirmed the

presence of unsaturated fatty acids appearing at a chemical shift range of olefinic protons at δ 5.3-5.36 ppm (Figure 2.9).

Sugars were detected showing hydroxylated CH and CH$_2$ in the range of δ 3.4-4.6 ppm (blue). Interestingly, all extracts show a doublet at δ 4.47 ppm (J = 7.82 Hz), which corresponds to the anomeric proton of free *β*-glucose. The signals at a chemical shift of δ 5.2 ppm (J = 3.70 Hz) can be assigned to the anomeric proton of free α-glucose. Signals related with *O*-glycosidic bonds are shifted depending on their linkage.[109] For *C*-glycosidic compounds the anomeric proton appears upfield at δ 4.90 ppm,[110] which is overlapped with the solvent signal of water.
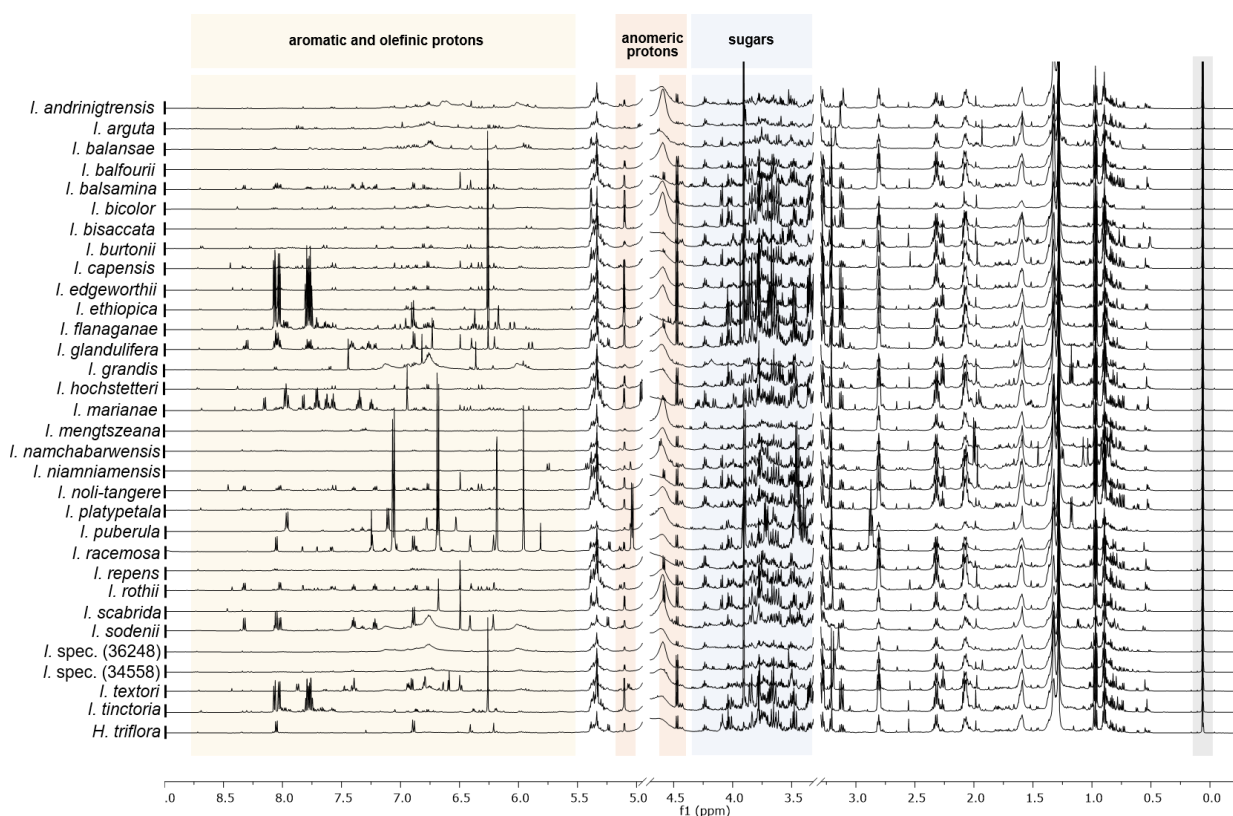
Figure 2.9. Stacked ^{1}H-NMR spectra (600 MHz, CD$_3$OD-d_4, 25°C) of methanolic leaf extracts of *Impatiens* species and *Hydrocera triflora*. Solvent signals are cut off (H$_2$O (4.70 – 4.95 ppm), CD$_3$OD (3.2-3.3 ppm)) and the internal standard HMDS (δ=0.062 ppm) is highlighted in grey. Minor differences in the chemical shifts are due to small differences in pH or concentration of the NMR samples.

Within the above described areas, the investigated extracts show a high number of common signals, suggesting similarities in their sugar composition and sugars incorporated in glycosidic compounds, as well as, similarities in their fatty acid constituents.

In contrast, the region of olefinic and aromatic protons (δ 5.5-8.5 ppm) is exceptionally diverse, suggesting a high diversity in the secondary metabolite patterns. During the UHPLC-ESI-HRMS based screening, flavonoids were found as one of the major compound classes detected in all *Impatiens* species and *Hydrocera triflora*. Characteristic chemical shifts of the A- and B-rings of the aglycones are detected within the aromatic region. For *H. triflora*, *I. flanaganae*, *I. racemosa*,

and *I. sodenii* intensive signals corresponding to the characteristic ¹H-NMR signals of kaempferol were detected (para substituted B-Ring, doublets H-6 (δ 6.21 ppm) and H-8 (δ 6.41 ppm), showing that kaempferol conjugates represent the main consistuent within these extracts (Appendix Figure A.14). This is in agreement with the results observed during UHPLC-ESI-HRMS (section 2.3.2.3).

I. puberula exhibited a different pattern in the aromatic region, confirming its uniqueness regarding the flavonoid constiuents as observed by UHPLC-ESI-HRMS² and multivariate data analysis (see Table 2.2, section 2.3.2.3.1, Appendix, Figure A.8). 1D- and 2D-NMR data indicate 5,7-dihydroxyl-4'-methoxyflavone as flavonoid skeleton in *I. puberula* (Appendix, Figure A.15-A.17).

In the extracts of *I. ethiopica*, *I. flanaganae*, *I. glandulifera* and *I. tinctoria* a sharp singulet at δ 6.29 ppm and signals of aromatic protons at chemical shifts of δ 7.76, 7.79, 8.03, 8.07 ppm reveal the presence of a 1,4-naphthoquinone skeleton. Structure elucidation was supported by 2D-NMR experiments (HSQC, HMBC, Appendix, Figure A.18-A.26) confirming the presence of 2-MNQ (**2.1**, Scheme 2.1, details see 2.3.6) in the methanolic leaf extracts. Although in lower amount, this compound was also detected in the leaf extract of *I. balsamina*, as previously described in the literature.[111]

In *I. sodenii*, *I. rothii*, *I. noli-tangere*, *I. glandulifera* and *I. balsamina* characteristic signals for the trihydroxynaphthalene skeleton were detected (annotation exemplified in the crude extract of *I. sodenii*, Appendix, Figure A.27). HMBC correlations confirmed the connection with a sugar moiety. Due to the coupling constant of the peak at 4.58 ppm (J = 7.9 Hz) corresponding with the anomeric proton, the sugar unit was thus identified as β-glycoside (Appendix Figure A.28-A.29). These observations, supported with our UHPLC-ESI-HRMS based profiling (Table 2.2), and compared with a literature report[58] allows to confirm the occurence of 1,2,4-trihydroxynaphthalene-1-*O*-β-D-glucoside (lawsone glycoside, **2.2**, Scheme 2.1) in the above mentioned *Impatiens* species.

2.3.4 Biological activity screening of methanolic leaf extracts of Balsaminaceae

In addition to the phytochemical analysis of the methanolic leaf extracts, their biological activity was evaluated against the gram-negative bacterium *Aliivibrio fischeri* and the gram-positive bacterium *Bacillus subtilis*. Furthermore, the anthelmintic properties of the crude extracts were tested against the model organism *Caenorhabditis elegans*, using a concentration of 500 µg plant extract/mL (see Appendix, Figure A.30). None of the extracts showed significant effect against the latter model organism. Previously, adverse effects on *C. elegans* were reported in literature for the 55% hydroethanolic stem extract of *I. balsamina*, as well as for the naphthoquinones lawsone and 2-methoxy-1,4-naphthoquinone,[82] compounds that appear as major constituents in the leaf extracts of some *Impatiens* species described in this work (i.e. *I. balsamina, I. ethiopica, I. flanaganae, I. glandulifera* and *I. tinctoria*, see section 2.3.3). Differences in the biological activity

of the tested extracts/compounds report within the scope of this thesis and those presented in the literature,[82] might be related to the different incubation times of the aforementioned nematode in the presence of the test substances. In our experiments, an incubation time of 30 min is used to investigate the acute toxicity of the samples, however in the article of Jiang et al. the incubation period lasted for 48 h.[82] Longer exposure of the nematodes with the tested extracts/compounds might lead to higher uptake provoking long-term toxicity through different metabolic pathways. In addition, Jiang et al. showed that kaempferol and quercetin increase the lifespan and improve the locomotion of larvae of *C. elegans*.[82] These results suggest that the presence of flavonoids have a positive effect on the organism survival. The methanolic leaf extracts investigated in this thesis possess a complex and high abundance of flavonoids (see sections 2.3.2, 2.3.3), which might also explain the lack of activity for the short-time incubation period. In our anthelmintic assay the presence of flavonoids within the tested extracts could attenuate the toxicity of the active compounds.

Although the research frame of the present work is focused on the discovery of new sources of antiinfective compounds, a preliminary cytotoxicity screening against the cancer cell lines HT-29 (human colon cancer) and PC3 (human prostate cancer) was conducted. Two different concentrations were tested, 0.05 and 50 µg/mL, in order to assess the effectivity of the extracts against the mentioned cell lines. Results show high growth inhibition at the high concentration and no effect at the low concentration, except for the methanolic leaf extracts of *I. niamniamensis*, *I. noli-tangere*, *I. sodenii*, and *I.* spec. (36248) which show moderate activity, for the investigated extracts (see Appendix, Figure A.31-A.32). Further investigations in order to determine the IC$_{50}$ and the mode of action of the tested extracts/isolated compounds will be conducted in the future.

2.3.4.1 Aliivibrio fischeri

The antibacterial activity of the leaf extract was examined against gram-negative *A. fischeri* using a luminescence based 96 well microtiter plate assay (experimental see section 2.5.9). At first, a fast screening was performed by applying two different concentrations of the corresponding leaf extracts, 500 µg/mL and 50 µg/mL (Figure 2.10). The commercial antibiotic chloramphenicol was used as positive control. At the higher concentration 26 *Impatiens* species display 100% growth inhibition, while *I. balansae*, *I. grandis*, *I. racemosa*, *I. sodenii*, *I.* spec. (36248) and *Hydrocera triflora* extracts acts as growth promoters (Figure 2.10).

A clear concentration dependence effect appears to occur, since at the lower applied concentration there is a loss of activity along the tested species. This means that the concentration of active metabolites is no longer sufficient to inhibit the growth of the bacteria. Furthermore, synergistic effects that previously enhanced the effect may be lost due to dilution. *Impatiens balsamina*, *I. ethiopica*, *I. flanaganae*, *I. glandulifera*, *I. textori,* and *I. tinctoria* still showed 100% growth inhibition at the lower concentration of the fast screening (50 µg/mL). Accordingly, *I. ethiopica*, *I. flanaganae*, and *I. glandulifera* were chosen as prime exploratory examples for further

concentration screening. Results show that all three *Impatiens* species inhibit the bacterial growth by 100% ratio at concentration of 15 µg/mL. *I. flanaganae* and *I. glandulifera* show 100% growth inhibition also at 6.25 µg/mL and 12.5 µg/mL, respectively. An outstanding outcome was found for *I. ethiopica*, displaying full inhibition at a concentration as low as 3.13 µg/mL (Appendix, Figure A.33).
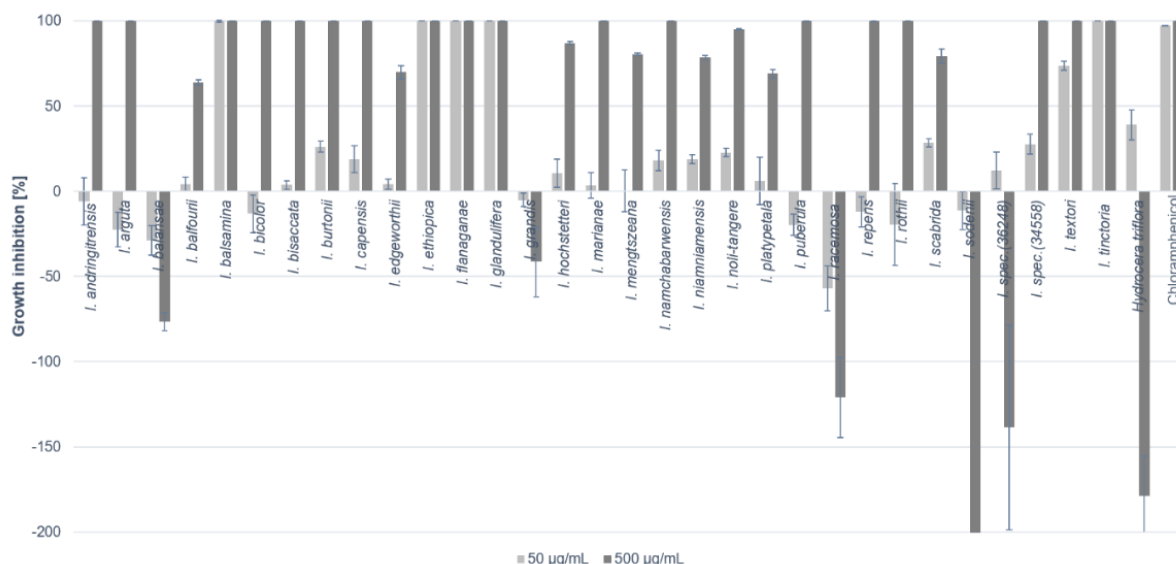
Figure 2.10. Growth inhibition of methanolic crude leaf extracts of 31 *Impatiens* species and *Hydocera triflora* against *A. fischeri* (concentrations 50 µg/mL and 500 µg/mL in DMSO, positive control: chloramphenicol, negative control: DMSO), n = 2 (technical replicates).

In comparison to gram-positive bacteria, the outer membrane of gram-negative microorganisms, containing lipopolysaccharides, constitute an additional barrier that compounds have to overcome for displaying its biological activity. This physiological feature of gram-negative bacteria is one of the main factors contributing to their intrinsic antibiotic resistances. In consequence, the remarkable result shown by the methanolic extract of *I. ethiopica* point out a certain degree of specialization. Interestingly, the most active species described above, showed in the previous analytical investigation using negative ion UHPLC-ESI-HRMS, similarities in their metabolite profiles (see sections 2.3.2, 2.3.3, 2.3.6). Notably, all of them present higher abundances of flavonoids and contain naphthoquinone and naphthalene derivatives (see section 2.3.2.4). Therefore, these classes of metabolites might have an influence on the observed antibacterial activity (*vide infra*).

2.3.4.2 Bacillus subtilis

For the evaluation of the antibacterial activity against gram-positive *B. subtilis*, the methanolic leaf extracts were tested in a turbidimetric assay using chloramphenicol as positive control. Initially, a fast screening was performed by applying two different concentrations of the corresponding crude extracts, 500 µg/mL and 50 µg/mL. At the higher concentration 5 *Impatiens* species display over

80% growth inhibition, while the rest of the species did not show a significant bioactivity. At the lower test concentration of 50 µg/mL the two species, *I. ethiopica* and *I. flanaganae*, completely inhibited the growth of the gram-positive test organism. (Figure 2.11).
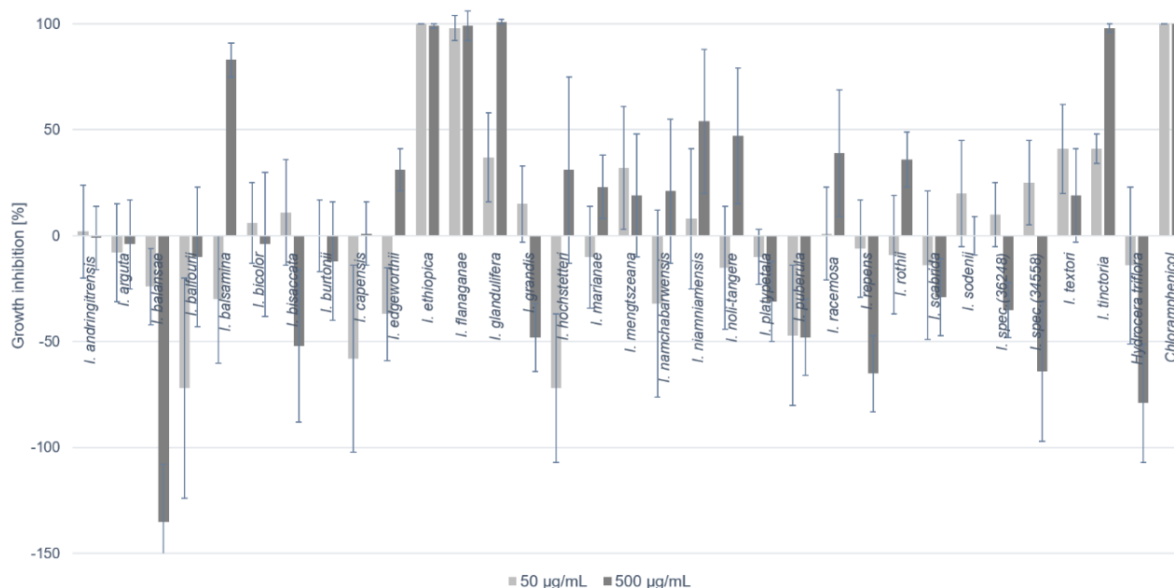
Figure 2.11. Biological activity of methanolic crude leaf extracts of 31 *Impatiens* species and *Hydocera triflora* against *B. subtilis* (concentration 50 µg/mL and 500 µg/mL in DMSO, positive control Chloramphenicol, negative control DMSO), n = 3 (technical replicates).

The extracts active against *B. subtilis* showed similarities in their NMR profiles. The high intensity signals of the aromatic protons suggest that 1,4-naphthoquinones are the main constituents of these extracts (see section 2.5.5), rising the hypothesis that this compound class might be related with the observed bioactivity. As previously described, the 1,4-naphthoquinone moieties are e.g. present in many naturally occurring alkaloids and have been identified as potential antibacterial candidates.[112-114] The mechanism of action of these compounds and its analogues involves enhanced ROS generation leading to cell apoptosis.[115-117] Various compounds presenting the 1,4-naphthoquinone molecular scaffold display relevant biological toxicity acting as antimicrobial,[118] anticancer,[119] antitubercular,[120] antimalarial,[121] and trypanocidal agents.[122]

In summary, *I. balsamina*, *I. flanaganae*, *I. ethiopica*, *I. glandulifera*, and *I. tinctoria*, are active against gram-negative and gram-positive bacteria, suggesting that the bioactive constituents of the corresponding extracts do not present a specific selectivity against the tested microorganisms, or they posses different active metabolite classes. On the other hand, *I. burtonii*, *I. capensis*, *I. namchabarwensis*, *I. niamniamensis*, *I. noli-tangere*, *I. scabrida*, and *I. textori* showed selective activity against the gram-negative *Aliivibrio fischeri* at the tested concentrations, suggesting that the metabolites present in their corresponding leaf extracts present a higher degree of specialization against gram-negative bacteria. Based on these results, *Impatiens* species are suitable sources for isolating new drugs to combat the increasing resistance to gram-negative bacteria, which is a major issue in current research and medicine.[123]

2.3.5 Phytochemical investigation of *Impatiens racemosa* DC*.*

The methanolic leaf extract of *I. racemosa*, compared to the other species, showed significant difference in its metabolite composition during analysis with thin-layer chromatography (see section 2.3.1), LC-ESI-HRMS (see section 2.3.2), and NMR analysis (see section 2.3.3). In literature, no information on the phytochemical composition of *I. racemosa* are described so far. Within the biological activity screening (see section 2.3.4), the extract exhibits strong growth promoting effects for the gram-negative bacterium *Aliivibrio fischeri* and the gram-positive *Bacillus subtilis*. Anthelmintic properties against *C. elegans* were not observed (Appendix, Figure A.30). Additionally, the antifungal properties of the extract were tested against *Botrytis cinerea*, *Septoria tritici* and *Phytophtora infestans* showing moderate inhibiting properties against *P. infestans* (Appendix, Figure A.34).
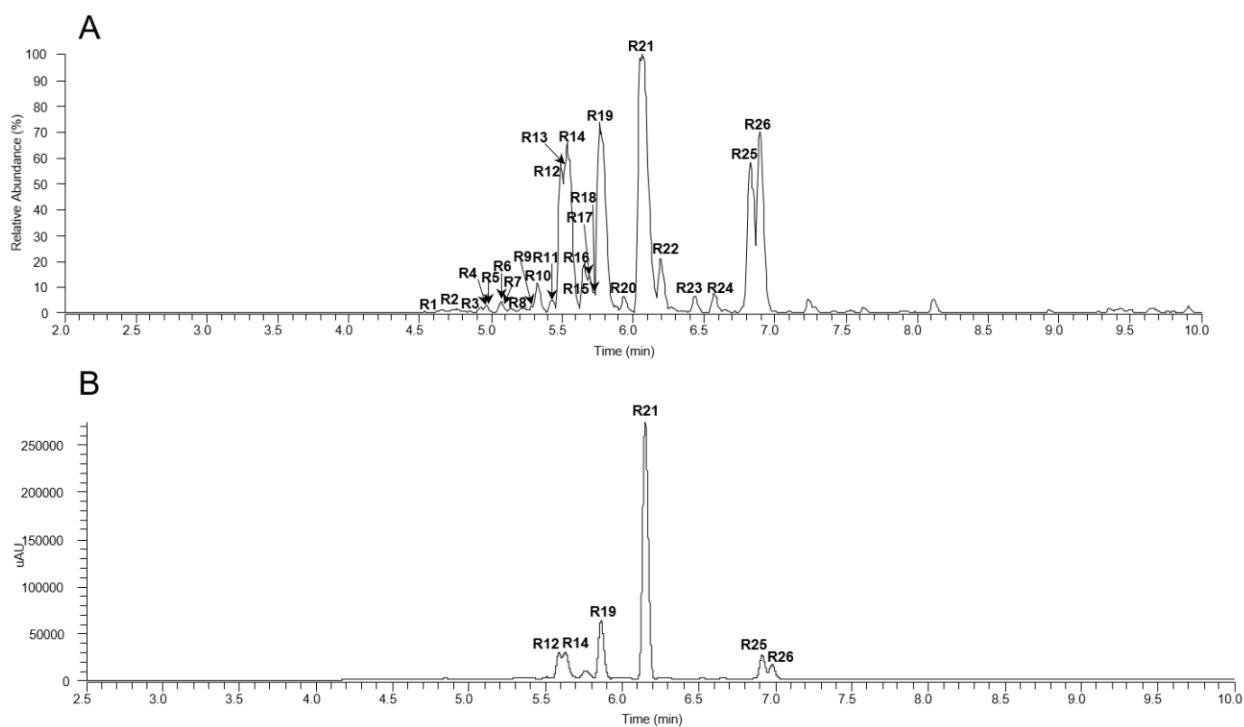
Figure 2.12. Basepeak chromatogram (A, *m/z* 150-2000) and PDA spectrum (λ = 190-400 nm) obtained from the methanolic crude extract of *I. racemosa,* R = Peak number.

Including the annotated metabolites reported in section 2.3.2 (Table 2.2), 26 metabolites were annotated within the crude extracts (Figure 2.12) and summarized in Appendix, Table A.3. Besides various flavonoid conjugates (R5, R12, R14, R18, R19, R20), hydrocinnamic acid derivatives (R1, R7, R11, R13, R15, R17), coumarins, including scopolin (R2) and coniferin (R3) were detected in the leaf extract of *I. racemosa*. As already known from the UHPLC-ESI-HRMS profiling*, I. racemosa* showed high abundance of peaks which are correlated to dihydrochalcone derivatives. In the base peak chromatogram as well as in the PDA spectrum the highest abundant peak (R21) was identified as phlorizin (**2.6**, Scheme 2.11). Beside its aglycone phloretin (**2.4**, R26), the higher

glycosylated derivatives phloretin-*O*-triglycoside (R9) and phloretin-*O*-diglycoside (R16) were detected. (Appendix, Table A.3, Figure 2.12) These compounds are known e.g. from *Malus* species and *Lithocarpus litseifolius*.[124-126] Furthermore, pinocembrin dihydrochalcone (**2.6**, R25, Scheme 2.1) was found to be one of the main constituents within the methanolic leaf extract of *I. racemosa* (Figure 2.12, Scheme 2.10). Based on their fragmentation behavior pinocembrin dihydrochalcone-*C*-hexoside-*O*-hexoside (R22), pinocembrin-*C*-hexosyl-*O*-pentoside (R24) were also detected (Appendix, Table A.3, Figure 2.12).

For detailed NMR investigations to confirm the proposed structures, the chromatographic separation of the methanolic crude extract leaf extract using preparative RP-HPLC-MS was performed. This resulted in seven fractions (Appendix, Figure A.35-A.36).

Fraction 1 and 2 mainly contained glycosylated coumarins as well as phenolic acid derivatives (Appendix, Figure A.37-Figure A.40). Major compounds in fraction 3 were identified as quercetin-*O*-glycosides (R12, R14; Appendix, Table A.3, Figure A.41-Figure A.44). Furthermore, the phloretin-tri-*O*-hexoside was found based on its fragmentation pattern and signals of the aglycone phloretin in the ^{1}H NMR and confirmed the presence of this compound (Appendix, Figure A.41). Fraction 4 showed two abundant peaks at a retention time of 5.69 min (R18, Appendix, Table A.3) and 5.77 min (R19, Appendix, Table A.3). The peaks showed the same *m/z* at 447.0927 ([M-H]⁻, $C_{21}H_{19}O_{11}^-$, calcd. 447.0933). In the ESI-HRMS^2 spectra the loss of 162 amu resulting in a fragment peak of *m/z* 285.0401 ($C_{15}H_9O_6^-$,) suggested the presence of kaempferol-*O*-hexosides. Based on 1D- and 2D-NMR experiments and in comparison with literature[127] the main constituent in fraction 4 was identified as astragalin (**2.3**, Scheme 2.1, kaempferol-3-*O*-β-D-glucoside, Appendix, Figure A.45-Figure A.47).

The chromatographic separation yielded two pure compounds phlorizin (**2.5**, fraction 5) and pinocembrin dihydrochalcone (**2.6**, fraction 6). The obtained NMR and ESI-HRMS data (Appendix, Figure A.51-Figure A.57) were in agreement with the previous reported ones.[125,126,128] Interestingly, phlorizin occurs in really high amount of 33% in the used 42 mg of dried material (see section 2.5.6). Fraction 7 contains the phloretin, the aglycone of phlorizin (Appendix, Figure A.58-Figure A.60). This is the first report on the chemical composition of *Impatiens racemosa* and the occurrence of dihydrochalcone derivatives within *Impatiens* species.

2.3.6 Isolation of bioactive compounds from *Impatiens ethiopica* Grey-Wilson and *Impatiens flanaganae* Hemsl.

Within the biological screening of the 31 Impatiens *species*, both extracts exhibited high activities against the bacteria tested (see section 2.3.4). The results against *A. fischeri* should be highlighted where the extract of *I. ethiopica* at a concentration of 3.13 µg/mL and the extract of *I. flanaganae* at 6.25 µg/mL still caused 100% growth inhibition of the bacteria (Appendix, Figure A.33). Furthermore, these extracts were also tested against resistant bacterial strains, showing significant

inhibition zones in a disc diffusion assay against *Stahpylococcus aureus* ATCC 43300 and *S. aureus* ATCC 6538P (Appendix, Figure A.61-A.62).

Thus, the aim was to perform a first separation using sequential centrifugal partition chromatography (SCPC) towards the isolation of the bioactive compounds. After in-depth literature revieweing and to the best of our knowledge these species are not yet phytochemically investigated. The dried leaf powders of *I. ethiopica* and *I. flanaganae* were extracted with 100% MeOH to obtain the crude leaf extracts These extracts were characterized by HPLC and NMR. Figure 2.13 shows representative HPLC chromatograms of the crude methanol leaf extracts of *I. ethiopica* (A) and *I. flanganae* (B) that share the peak with the highest absorbance at 6.8 minutes and the peak at 4.3 min. In total, the extract of *I. flanganae* contains more compounds with an absorbance at 254 nm than the leaf extract of *I. ethiopica*.
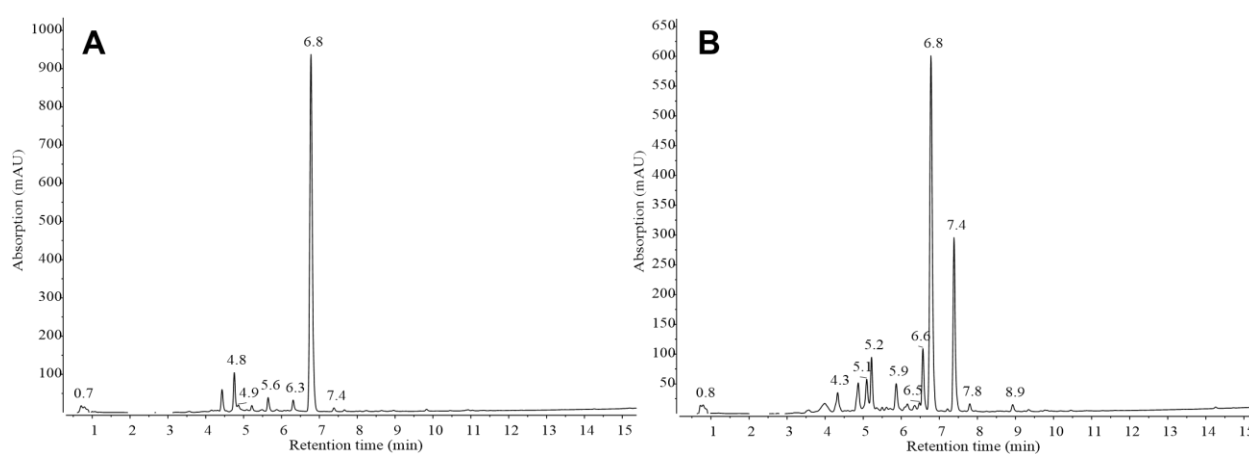
Figure 2.13. Representative HPLC chromatograms (λ= 254 nm) of crude methanolic leaf extracts from (A) *I. ethiopica* and (B) *I. flanaganae*.

By comparing both ^{1}H NMR spectra (Figure 2.14) it is striking that they have a high number of NMR signals in common. The leaf extract of *I. flanaganae* contains more sugars represented by the ^{1}H signals in the range of δ 3.5 to 4 ppm and exhibits more signals in the aromatic proton region between δ 6.75 to 7.25 ppm, which could be related to a higher content of flavonoids. Both extracts show significant signals with a chemical shift in the region from δ 7.75 ppm to 7.78 and δ 8.05 to 8.10 ppm. For further investigations both extracts were separated using SCPC.
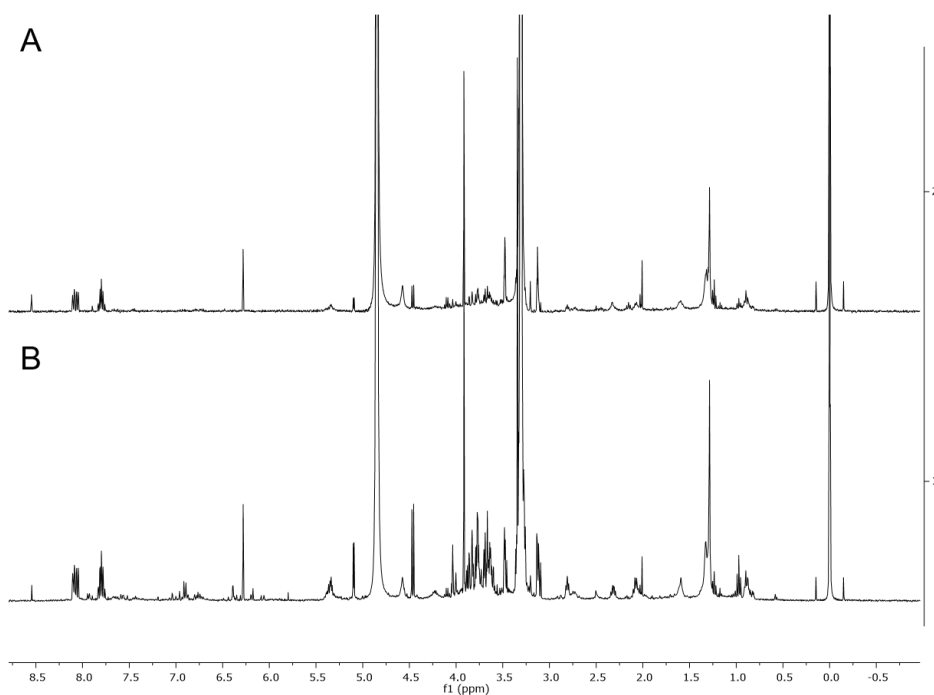
Figure 2.14. ¹H-NMR spectra (400 MHz, CD₃OD-*d4*, 25°C, 256 scans) of crude methanolic leaf extracts from (A) *I. ethiopica* and (B) *I. flanaganae*.

Fractionation using sequential centrifugal partition chromatography (SCPC)

The SCPC separation of crude extracts (Appendix, Figure A.63-A.64) monitored by ESI-MS, TLC and HPLC resulted in seven leaf extract fractions from *I. ethiopica* (AI-AVII, Appendix, Figure A.65) and six fractions from *I. flanaganae* (BI-BVI, Appendix, Figure A.66). Fraction I eluted in both cases as a sharp peak at a retention time of circa 7 to 11 min and 15 min, respectively. These fractions (AI, BI) contain mainly sugars, represented in the NMR spectra by a high number of proton signals in the region of δ 3.5 to 4 ppm (Figure 2.15). Fractions AII and BII eluted early during SCPC fractionation using the polar phase as mobile phase. In the corresponding ¹H-NMR spectra signals appear in the aromatic proton region which are common for e.g. flavonoids (Figure 2.15).
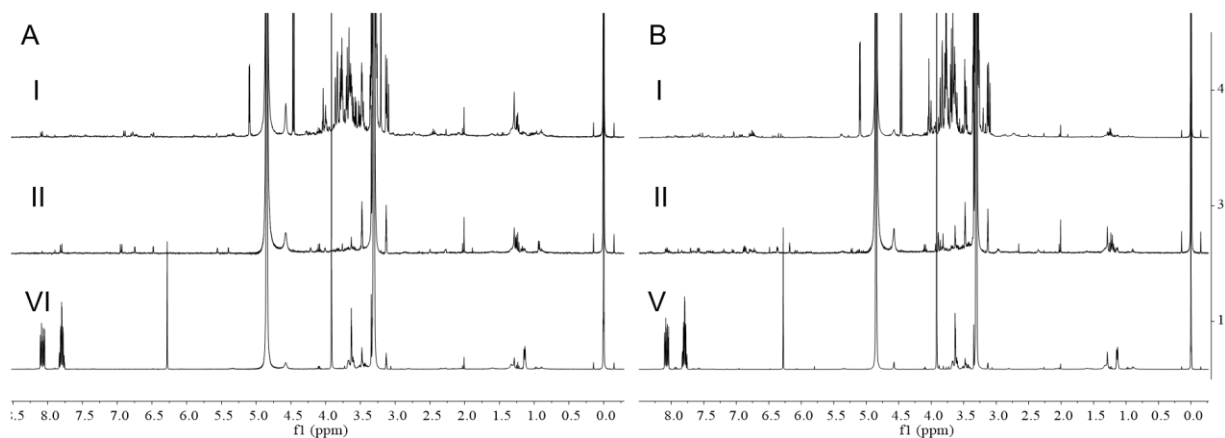
Figure 2.15. ¹H-NMR spectra (400 MHz, CD₃OD-*d4*, 25°C) of SCPC fractions of methanolic leave extracts from (A) *I. ethiopica* and (B) *I. flanaganae*

By switching the elution mode, fractions AV and AVI and BIV and BV eluted as sharp peaks within two to three minutes. Hereby, fractions BV and AVI were obtained with high purity and show an identical ¹H-NMR spectra. All obtained fractions where tested in the antibacterial assay against *A. fischeri* and *B. subtilis* (Appendix, Figure A.67-A.68). Fraction I of both extracts showed due to the sugar content growth promoting effects whereas fraction AV and AVI from *I. ethiopica* and fractions BIII, BIV and BV from *I. flanaganae* exhibited 100% growth inhibition against *A. fischerii* for both concentrations (500 and 50 µg/mL, Appendix, Figure A.67-A.68). Fraction AVI and BV exhibited 100% growth inhibition also against *B. subtilis* (Appendix, Figure A.67-A.68).

Based on these results, it is clear that the biological activity is not due to only one class of substance. In fraction AV signals were found belonging to Lawsone and flavonoid derivatives. The isolated 2-methoxy-1,4-naphthoquinone (AVI, BV) showed strong antibacterial activity consistent with those described in the literature.[20,53]

Structure elucidation of fraction BV from I. flanaganae

Fraction BV contains 2-methoxy-1,4-naphthoquinone, which was elucidated based on its ¹H-(Figure 2.16) and ^{13}C-NMR spectra (600 MHz for ¹H and 150.84 MHz for ^{13}C spectra in CD₃OD-d_4, including 2D HMBC and HSQC spectra, Appendix, Figure A.69-A.70). The molecular formula was deduced as $C_{11}H_9O_3$ from high-resolution ESI-MS in agreement with its $[M+H]^+$ ion at *m/z* 189.0546 (Appendix, Figure A.71). This compound also known from other *Impatiens* species represents the major constituent of both investigated species and might be responsible for the observed antibacterial activity.
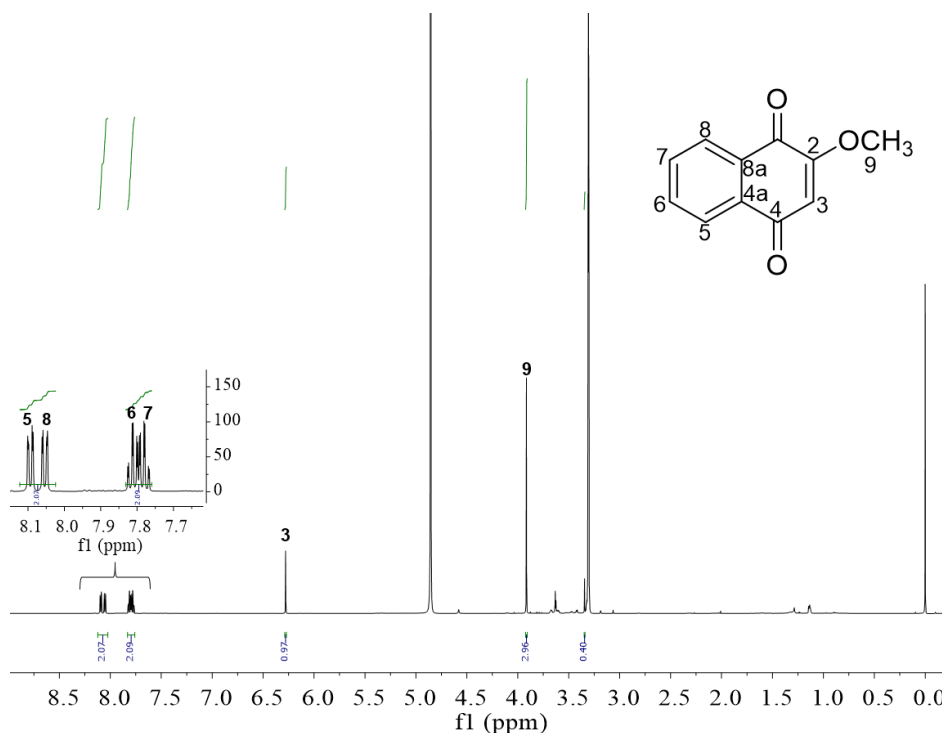
Figure 2.16. ¹H-NMR spectra (600 MHz, CD₃OD-*d4*, 25°C) of SCPC fraction BV obtained from the methanolic leaf extracts of *I. flanaganae.*

2.4. Conclusions

This work represents the first attempt to explore interspecific variation in secondary metabolites from a wide range of *Impatiens* species and *Hydrocera triflora*. Herein, we report on the first phytochemical investigation and provide preliminary information on the antibiotic activity of *I. andringitrensis, I. arguta, I. balansae, I. bisaccata, I. burtonii, I. ethiopica, I. flanaganae, I. grandis, I. hochstetteri, I. marianae, I. marianae, I. platypetala, I. puberula, I. racemosa, I. repens, I. rothii, I. scabrida, I. sodenii,* and *Hydrocera triflora.* Using an optimized protocol for the sample preparation, as well as optimizing parameters for the measurements, a broad coverage of secondary metabolites within the analysis was achieved. The investigation of the methanolic leaf extracts by planar TLC gave the first insights into the complexity of the extracts and enabled a first comparison of the different constituents of the investigated species. UHPLC-ESI-HRMS and NMR profiling showed a high variability regarding the metabolites present in the leaf extracts of the investigated species. Herein, we have further demonstrated that combining UHPLC-ESI-HRMS and NMR allows for a more complete picture (overview) of the metabolome.

Detailed compound composition analysis using UHPLC-ESI-HRMS/MS lead to the annotation of 156 peaks, revealing that flavonoid conjugates are broadly distributed along all investigated Balsaminaceae species and has allowed for the first report on *C*-glycosides within the genus *Impatiens*. Furthermore, known carboxylic acids, hydroxycinnamic acid derivatives, coumarins, naphthoquinones, and naphthalenes were identified and core structures of new naphthalene derivatives were suggested based on their fragmentation behavior. Due to the large differences in secondary metabolite profiles, no specific biomarkers for the genus *Impatiens* could be defined in this study. It should be mentioned that the monotypic species *Hydrocera* generally has a very similar metabolite composition compared to the *Impatiens* species studied. This suggests that the current taxonomic classification of the family Balsaminaceae in two genera might be insufficient, launching the possibility for new classification paths. Although samples cluster in the hierarchical clustering analysis as well as in the principal component analysis, this segregation does not correlate well to the phylogenetic background, geographic origin or the antibacterial activity. This is due to the great diversity within the chemical profiles of the investigated species. The high interspecific variance of the acquired data causes low principle components within the multivariate data analysis. Based on the evaluated data, it was concluded that a full features PCA analysis is not the most appropriate method for multispecies comparison. In order to complement the outcome of this study more specified analysis can be conducted, e.g. analysis of individual substance classes such as flavonoids and naphthoquinones within the species. Furthermore, the exclusive consideration of negative ion UHPLC-ESI-HRMS data is not sufficient and is limited to ionizing substances using this ionization mode. 2-Methoxy-1,4-naphthoquinone, which was found to be one of the antibacterial metabolites, does not show well ionizing properties in negative ion mode. Therefore, PCA and HCA on positive ion mode UHPLC-ESI-HRMS and NMR data should be performed in the future to overcome ionization limitations.

Based on the results of metabolic profiling, the methanolic extract of *I. racemosa* was further fractionated chromatographically. Thereby, three flavonoid glycosides and three dihydrochalcone derivatives were successfully isolated. To the best of our knowledge, the occurrence of dihydrochalcone derivatives within *Impatiens* species is described for the first time.

As shown in the manuscript, the methanolic leaf extracts of *I. ethiopica* and *I. flanaganae* display the highest antibacterial activity against the gram-negative bacterium *Aliivibrio fischeri* amongst all investigated species. After separation of the extracts using SCPC, 2-methoxy-1,4-naphthoquinone was isolated as one of likely several active compounds.

In spite of the outstanding bioactivity associated to some of the leaf extracts, at this point it was not possible to establish a direct correlation between chemical composition and antibacterial activity. However, through these initial studies we were able to show that the antibiotic activity of some species against the gram-positive bacterium *B. subtilis* might be related to metabolites containing a 1,4-naphthoquinone moiety, which was confirmed through the isolation of 2-methoxy-1,4-naphthoquinone from *I. ethiopica* and *I. flanaganae*. The outstanding antibacterial activity against the gram-negative bacterium *Aliivibrio fischeri* seems to be related to different compound classes including flavonoids, naphthoquinone and naphthalene derivates. However, this has to be further investigated and confirmed through isolation of the bioactive constituents. In summary, an efficient workflow based on HRMS and NMR was implemented for the successful isolation and characterization of new bioactive compounds from different genera of the Balsaminaceae family. This represents the basis for future detailed studies towards the isolation of bioactive compounds from more *Impatiens* species for deeper elucidation on its bioactive principles.

2.5 Experimental Section

2.5.1 Plant Material

Whole plants of 31 *Impatiens* species and *Hydrocera triflora* were collected in 2015 by Dr. Abrahamczyk (Botanical Garden Bonn/ Nees-Institute for Biodiversity Bonn). Herbal vouchers are stored in the Nees Institute for Biodiversity Bonn. The plants were separated into the single organs and directly frozen (-50 °C). Afterwards the material was stored at –80 °C and later lyophilized for analysis. An overview of the plant material available is listed in Table 2.1. The perennial species are clones of the species of wild origin. Therefore, the material, which was harvested from different plants of the perennial species, always belong to one genetic individual. The annual species have been propagated by seeds. Correspondingly, the material for annual species comes from several individuals.

2.5.2 Sample preparation

Dried plant material was ground using a liquid nitrogen cooled vibrating mill (MM400, Retsch GmbH, Germany, applying a frequency of 30 Hz for 30 seconds). Afterwards the powder was dried by lyophilisation, and stored at room temperature in darkness for further investigations. A quality control (QC) sample was prepared, which is composed of equal parts of all individual samples. These pooled sample is considered as data set representative for quality control and method development.

2.5.3 Thin-layer chromatography (TLC)

For TLC analysis the plant material (approximately 25 mg) was weighed in 2 mL Eppendorf tubes. 1 mL methanol was used for extraction under ultrasound for 15 minutes. After 10 min centrifugation at 14000xg (room temperature) the supernatant was separated and dried under nitrogen flow. The obtained dried extracts were redissolved and adjusted to a concentration of 10 mg extract/ 1 mL methanol. Chromatographic separation was performed on silica gel 60 F_{254} TLC plates (with fluorescence indicator, 20×20 cm, layer thickness 150–200 µm, Merck, Germany) using a mixture of chloroform: methanol (9:1; *v/v*) as mobile phase. After the application of 2 µL of the extract on the plate, the plate was conditioned for 10 min within the TLC chamber containing the solvent mixture. The plate was developed for 10 cm, taken out and air-dried. Further, the developed plate was photographed in daylight, 254 nm and 366 nm UV-light. For further information, a derivatizion using vanillin sulfuric acid spray reagent [1.2 g vanillin (Fluka), 25 mL glacial acetic acid (Roth), 11 mL sulfuric acid (Roth), 212 mL methanol (distilled)] was performed. For documentation and R_f-value determination, a CAMAG TLC visualizer (CAMAG, Muttenz, Switzerland) was used with the software winCATS (version 1.4.9.2001, CAMAG, Switzerland).

2.5.4 UHPLC-ESI-HRMS analysis

Non-targeted UHPLC-ESI-HRMS based profiling

Approximately 2 mg of dried plant powder of each species (three technical replicates per species) were weighed into a 1.5 mL Eppendorf tube and stored at 25 °C until measurements. For extraction, the volume of methanol was added to end up with a concentration of 2 mg plant powder/mL. Extraction was performed in the ultrasonic bath for 15 min followed by centrifugation at 25 °C and 14000xg (5415 C Eppendorf, Eppendorf AG, Hamburg, Germany). The negative ion high-resolution ESI mass spectra (*m/z* range 100-1500) were obtained from an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Germany) equipped with a heated electrospray ion source (negative spray voltage 4 kV, capillary temperature 325 °C, source heater temperature 300 °C, FTMS resolution 15.000, Nitrogen was used as a sheath and auxiliary gas. The MS system was coupled to an ultra-high-performance liquid chromatography (UHPLC) system (Dionex UltiMate 3000, Thermo Fisher Scientific), equipped with a RP-18 column (particle size 1.9 µm, 50 × 2.1 mm ID, Hypersil Gold, Thermo Fisher Scientific; column temperature 45 °C). The mobile phases were H_2O (A; MilliQ-biocel apparatus from Millipore (Billerica, USA)) and CH_3CN (B; Chromasolv™, for LC-MS, Honeywell Riedel de Haën™) with 0.1% formic acid (additive for LC-MS, LiChropur®, Merck). Chromatographic separation was realized using a gradient system starting from 5% B (isocratic for 1 min) increasing to 98% B within 10 min, followed by further 3 min at 98% B (flow rate 0.4 mL/min, injection volume 2 µL). The re-equilibration time of the column was set to 5 min at 5% B. The CID mass spectra (buffer gas: helium) using data dependent acquisition were recorded using a normalized collision energy (NCE) of 35%. The instrument was externally calibrated by the Pierce ESI negative ion calibration solution (product no. 88324) from Thermo Fisher Scientific. The data were evaluated using the software Xcalibur 2.2 SP1.

A complete randomized design was applied for the batch design. The batch started with three methanol blanks followed by four QC samples for the equilibration of the column. After 6 study samples the QC was injected. In addition, an external standard mix (MM8) was injected every 16th sample for the detection of shifts and effects in mass-to-charge ratios (*m/z*) and retention times (RT). The following substances were used in the MM8: 2-phenylglycine (Fluka), kinetin (Roth), rutin (Acros Organics), *O*-methylsalicylic acid (Sigma), phlorizin dihydrate (Sigma), N-(3-indolyacetyl)-*L*-valine (Sigma), 3-indolylacetonitrile (Fluka) and biochanin A (Sigma).

UHPLC-ESI-HRMS data processing

The raw LC/MS data were converted to the open data format mzML[129] format using ProteoWizard (3.0.20115) to get centroided data. Chromatographic peak picking was performed in R 3.6.2 (available at https://cran.r-project.org) in Rstudio (1.2.5033) with the package xcms 3.9.1[130] using the centWave algorithm and the following parameters: ppm =7, peakwidth=7,25, snthresh=10, prefilter=3,1000, fitgauss=FALSE, verbose.columns = FALSE. Grouping of chromatographic peaks was performed using the peakDensity algorithm, without log

transformation, and the following parameters: minfrac=1, bw =3. Retention time correction was performed using the function retcor in XCMS using the parameters method =loess, family =gaussian, span=0.2. The resulting feature lists (*m/z*, t_R) of all leaf samples comprised 7151 features.

Chemometrics

The chemometric analyses were performed based on the generated feature table (see data processing) under R environment using the following packages: pcaMethods (1.78.0), Corrplot (0.84) and ggdendro (0.1.22).

2.5.5 Nuclear magnetic resonance (NMR)

Non-targeted NMR profiling

For NMR analysis, 80 mg of the dried and powdered plant material was ultrasonically extracted for 15 minutes with deuterated methanol MeOH-*d4* (99.8 % D, Deutero GmbH, Kastellaun, Germany), containing 0.935 mM hexamethyldisiloxane (HMDS) as internal standard. Afterwards the samples were centrifuged for 10 minutes at 14000*xg* (5415 C Eppendorf, Eppendorf AG, Hamburg, Germany) to separate plant residues and 750 µL of the resulting supernatant was transferred to quantitative Deu-Quant-5-7 NMR tubes (Deutero GmbH, Kastellaun, Germany).

The NMR spectra of the methanolic crude extracts were recorded on an Agilent VNMRS 600 NMR spectrometer at 599.83 MHz (^{1}H) and 150.84 MHz (^{13}C) operating at 25 °C equipped with a 5-mm inverse detection cryoprobe using standard CHEMPACK 7.1 pulse sequences (s2pul) implemented in a Varian VNMRJ 4.2 spectrometer software. The signals were referenced to internal HMDS at 0.062 ppm. Quantitative ^{1}H-NMR spectra were measured with the following parameters: pulse angle = 90°, relaxation delay (d1) + acquisition time (at) = 30.0 s, number of scans (nt) = 64, and digital resolution = 128 K, spectral width -1 – 14 ppm. Data analysis was carried out with MestreNova (12.0.4-220023, Mestrelab Research, S.L. (Santiago de Compostela, Spain).

NMR analysis of reference and isolated compounds

For NMR analysis the compounds were dissolved in 800 µL deuterated methanol MeOH-*d4* (99.8 % D, Deutero GmbH, Kastellaun, Germany), containing 0.03 % TMS (euriso-top, Saarbrücken, Germany) as internal standard and transferred to qualitative NMR tubes (Deu-500, Deutero GmbH, Kastellaun, Germany). The NMR spectra were recorded on a Varian MERCURY-VX 400 system or Varian 600 VNMRS system at +25 °C operating at a proton NMR frequency of 399.82 MHz or 599.83 MHz and a carbon NMR frequency of 150.84 MHz and 100 MHz, respectively. 2D NMR spectra (HSQC, HMBC, COSY, NOESY, TOCSY) were obtained from a Varian 600 VNMRS system. The HSQC experiment was optimized for 1JCH = 146 Hz with DEPT-like editing and ^{13}C-decoupling during acquisition time. The HMBC experiment was optimized for a long-range coupling of 8 Hz.

2.5.6 Fractionation of Impatiens racemosa DC using semipreparative HPLC-MS

Fractionation of the crude methanolic extract of *Impatiens racemosa* DC (42 mg) was performed using a semipreparative RP-18 HPLC-MS using a custom set system of Agilent 1200 Infinity I series consisting of a preparative binary pump (G1361A), an analytical quaternary pump (G1311B) used as make-up pump, an auto sampler (G2260A), a column select valve (G1159A), a fraction collector (G1364B), a multiple wavelength detector (G1365D), a splitter (G1968D) and single quadrupole MS detector with an electrospray ionization source (ESI, 6120). Electrospray mass spectrometry measurements were performed in negative and positive ionization mode (fragmentor, 5 V; threshold spectral abundance, 150; capillary voltage, 3000 V) using a mass range of m/z 200–1200 Da. Nitrogen was used as the nebulizer gas (nebulizer pressure, 35 psig). The drying gas flow was set to 12.0 L/min and the drying gas temperature to 250°C. The chromatographic separation was conducted using a Nucleodur C18 ec column (5 μm, 100 Å, 150 x 10 mm internal diameter, Macherey-Nagel, Germany). For the separation the mobile phases were H_2O (A; bidistilled, Millipore) and CH_3CN (B; gradient grade, LiChrosolv, Merck) with formic acid (0.1%, Merck) and a gradient system was used (0–30 min, 5–100% B; flow rate 3.7 mL/min). Automated fractionation was carried based on scan mode in negative ion mode. The whole system was controlled by the Agilent Chemstation (Rev. 8.04.03 SP1). The fractionation yielded seven fractions: fraction 1 (t_R: 2.1-3.0 min, 10.5 mg), fraction 2 (t_R: 6.0-7.6 min, 3 mg), fraction 3 (t_R: 8.2-8.9 min, 3.7 mg), fraction 4 (t_R: 9.0-9.7 min, 3.7 mg), fraction 5 (t_R: 10.0-11.1 min, 13.8 mg), fraction 6 (t_R: 12.7-13.4 min, 3.7 mg), and fraction 7 (t_R: 14.0-14.9 min, 3.8 mg). Fraction 4 contains kaempferol-*O*-hexoside as main constituent. Fraction 5 was identified as the known dihydrochalcone glycoside phlorizin and fraction 7 as the aglycone phloretin. Fraction 6 was identified as pinocembrin dihydrochalcone-*O*-glycoside (configuration as phlorizin).

Fraction 5 (Phlorizin): ^{1}H NMR (600 MHz, Methanol-*d₄*) δ 7.08 – 7.04 (m, 2H, H-2, H-6), 6.70 – 6.66 (m, 2H, H-3, H-5), 6.17 (d, *J* = 2.3 Hz, 1H, H-3′), 5.94 (d, *J* = 2.3 Hz, 1H, H-5′), 5.03 (d, *J* = 7.2 Hz, 1H, anomeric proton of *β*-glucopyranoside (H-1″)), 3.90 (dd, *J* = 12.1, 2.2 Hz, 1H, H-6″), 3.71 (dd, *J* = 12.1, 5.5 Hz, 1H, H-6″), 3.49 – 3.43 (m, 4H, H-2″, H-3″, H-4″, H-5″), 3.42 – 3.36 (m, 2H, H-α), 2.87 (ddd, J = 8.8, 6.9, 2.4 Hz, 2H, H-β). ^{13}C NMR (Information extracted from HMBC/HSQC, 151 MHz, Methanol-*d₄*) δ 130.4 (*C*H, C-2, C-6), 116.1 (*C*H, C-3, C-5), 98.4 (*C*H, C-5′), 95.6 (*C*H, C-3′), 102.1 (*C*H, anomeric carbon of *b*-glucopyranoside (C-1″)), 78.5, 74.8, 71.1, 71.1 (*C*H, glucopyranoside), 62.5 (*C*H₂, C-6″), 47.0 (*C*H₂, C-α), 30.9 (*C*H₂, C-β), which was in accordance with previous reports on phlorizin[113,114] and the reference compound. UHPLC-ESI-HRMS: m/z at 435.1295 ([M-H]-, $C_{21}H_{23}O_{10}^-$, calcd. 435.1297)

Fraction 6 (Pinocembrin dihydrochalcone glucoside): ^{1}H NMR (600 MHz, Methanol-*d₄*) δ 7.27 – 7.20 (m, 4H, H-2, H-3, H-5, H-6), 7.15 – 7.11 (m, 1H, H-4), 6.13 (d, *J* = 2.3 Hz, 1H, H-3′), 5.90 (d, *J* = 2.3 Hz, 1H, H-5′), 5.06 – 5.02 (m, 1H, anomeric proton of *b*-glucopyranoside (H-1″)), 3.90 (dd, *J* = 12.1, 2.2 Hz, 1H, H-6″), 3.71 (dd, *J* = 12.2, 5.4 Hz, 1H, H-6″), 3.53 – 3.43 (m, 4H, H-2″,

H-3″, H-4″, H-5″), 3.43 – 3.35 (m, 2H, H-α), 2.97 (ddd, *J* = 8.1, 6.9, 2.7 Hz, 2H, H-β). UHPLC-ESI-HRMS: *m/z* at 419.1345 ([M-H]⁻, $C_{21}H_{23}O_9^-$, calcd. 419.1348), which was in accordance with previous reports on the aglycone pinocembrin.[118]

2.5.7 Extraction and SCPC separation

The plant powders (*I. ethiopica* m= 267.4 mg; *I. flanaganae* m= 266.0 mg) were extracted using methanol for 15 minutes in an ultrasonic bath. After 10 minutes centrifugation at 14000xg the supernatants were transferred to 6 mL glass vials and dried under nitrogen flow. Sequential centrifugal partition chromatography (SCPC) was performed using a Gilson Glider CPC system with a 250 mL rotor, adjustable rotation of 200–2000 rpm and a PLC purification system (Gilson) pump. Detection was carried out using a UV-VIS detector (Gilson PLC,). A manual sample injection valve was used to introduce the samples into the column. The methanolic crude leaf extracts of *I. ethiopica* (28.7 mg) and *I. flanaganae* (48.3 mg) were subjected to CPC separation using a mixture of *n*-hexane: EtOAc:MeOH:H₂O (2:5:2:6, *v/v/v/v*) as biphasic system. The separation was run at a revolution speed of 2200 rpm. Initially, the lower layer (water based) was used as the mobile phase, while the upper layer of the biphasic system was used as the stationary phase in a tail to head or descending mode (flow rate 5 mL/min). The effluent of the column was automatically collected in 8 mL aliquots and this procedure resulted in two fractions. Then after 50 min, the organic layer was used as mobile phase in a head to tail or ascending mode. Total number of 46 fractions for both extracts were collected. Based on the comparison of the ESI-MS and HPLC chromatograms the obtained CPC fractions were combined as following for *Impatiens ethiopica*: fraction AI (1-4, t_R: 7.0-13.3 min, 32 mL, 8 mg), fraction AII (5-7, t_R: 13.3-18.2 min, 24 mL,1 mg), fraction AIII (8-14, t_R: 18.2-29.4 min, 56 mL, 0.5 mg), fraction AIV (15-27, t_R: 29.4-50.4 min, 104 mL, 1.2 mg), fraction AV (28-29, t_R: 50.4-53.5 min, 16 mL, 12.5 mg), fraction AVI (30-33, t_R: 53.5-60.0 min, 32 mL, 14.7 mg), fraction AVII (34-46, t_R: 60.0-80.0 min, 104 mL, 9.9 mg); and for *Impatiens flanaganae*: fraction BI (1-7, t_R: 7.0-18.0 min, 56 mL, 25.1 mg), fraction BII (8-13, t_R: 18.2-28.0 min, 48 mL, 1.7 mg), fraction BIII (14-28, t_R: 28.0-52.2 min, 120 mL, 2.8 mg), fraction BIV (29-30, t_R: 52.2-55.3 min, 16 mL, 10.1 mg), fraction BV (31-38, t_R: 55.3-68.3 min, 64 mL, 6.9 mg), fraction BVI (39-46, t_R: 68.3-80.0 min, 64 mL, 0.9 mg)

2-Methoxy-1,4-naphthoquinone (**2.1**): Yellow amorphous powder. ¹H NMR (600 MHz, Methanol-*d₄*) δ 8.094 (1H, dd, *J* = 7.4, 1.5 Hz, *H*-8), 8.054 (1H, dd, *J* = 7.3, 1.5 Hz, *H*-5), 7.812 (1H, ddd, *J* = 7.4,7.4, 1.6 Hz, *H*-6), 7.780 (1H, ddd, *J* = 7.4, 7.4, 1.5 Hz, *H*-7), 6.279 (1H, s, *H*-3), 3.915 (3H, s, *H*-9). ¹³C NMR (151 MHz, CD₃OD) δ 186.4 (*C*-8a), 181.4 (*C*-1), 162.2 (*C*-2), 135.5 (*C*-6), 134.5 (*C*-7), 133.4 (*C*-4a), 132.6 (*C*-4), 127.4 (*C*-8), 126.9 (*C*-5), 110.67 (*C*-3), 57.10 (*C*-9), UHPLC-ESI-HRMS: *m/z* at 189.0546 ([M+H]⁺, $C_{11}H_9O_3^-$, calcd. 189.0546), which was in accordance with previous reports on *2*-methoxy-1,4-naphthoquinone[82,131] and the reference compound.

2.5.8 Analytical high-performance liquid chromatography (HPLC)

Analytical RP18 HPLC was carried out with an Agilent 1260/1290 system equipped with a quarternary pump and a diodenarray detector VL+ using a Poroshell 120 C18ec column (length: 50 mm, internal diameter: 4.6 mm, pore size: 100 Å, particle size 2.7 µM, Agilent). For the separation the mobile phases were H_2O (A; bidistilled, Millipore) and CH_3CN (B; gradient grade, LiChrosolv, Merck) with formic acid (0.1%, Merck) and a gradient system was used (0–15 min, 5–100% B (5 min); flow rate 0.8 mL/min).

2.5.9 Biological Assays

2.5.9.1 Antibacterial Bioassay

Aliivibrio fischeri

The assay was performed according to a procedure described by Stark (phD thesis 2016)[132] using the gram-negative *A. fischeri* test strain DSM507 (batch no. 1209). Briefly, for each test run a fresh glycerol stock was incubated in 25 mL Boss medium at 100 rpm and 23 °C for 16‑18 h and was afterward diluted with fresh BOSS medium to an appropriate cell number (luminescence value between 30.000 and 50.000 RLU). The assay was conducted on black flat bottom 96 well plates (Brand cellGrade™premium, STERILE R) in a final volume of 200 µl of Boss medium containing 1% DMSO in each well (100 µl diluted bacterial solution and 100 µl test solution). The respective extracts were applied in two concentrations (50 µg/ml and 500 µg/ml) in DMSO/medium (2%, *v/v*). The plates were incubated in the dark at 23 °C and 100% humidity without lid and without shaking for 24 h. The bioluminescence (obtained in relative luminescence units, RLU) is dependent on the cell density and was determined after 24 h using the microplate reader TecanGeniosPro. Therefore, the whole wavelength range was detected for 1000 ms without preliminary shaking to avoid secondary oxygen effects. The results (mean value ± standard deviation, n=6) are given as relative values (% inhibition) in comparison to the negative control (bacterial growth, 1% DMSO, without test compound). Negative values indicate an elevation of luminescence/increase of bacterial growth. Chloramphenicol was used as positive control.

Bacillus subtilis

The activity against the gram-positive *Bacillus subtilis* 168 (DSM 10) is determined in a turbidimetric assay, which was described in a previous work.[133] A preculture in 100 mL tryptone yeast medium (TY) was incubated at 37 °C for 24 h as agitated culture (120 rpm). Afterward, the microbial density was adjusted to the absorption of the medium at λ=612 nm. The assay was performed on clear flat bottom 96 well plates (Greiner Bio-one) in a final volume of 200 µl TY medium containing 1% DMSO in each well (100 µl diluted bacterial suspension and 100 µl test solution). To eliminate the influence of colored compounds for each test sample, color controls without bacteria were treated in the same manner. The plate with lid was incubated for 16 h at 37 °C under shaking (80 rpm) and the absorption determined at 612 nm using the microplate reader TecanGeniosPro. The results (mean value ± standard deviation, n=6) are given as relative values

(% inhibition) in comparison to the negative control (bacterial growth, 1% DMSO, without test compound) after subtracting background absorption of color control samples. Chloramphenicol was used as positive control.

2.5.9.2 Anthelmintic Bioassay

The Bristol N2 wild type strain of *Caenorhabditis elegans* was used in the anthelmintic assay. The nematodes were cultured on NGM (Nematode Growth Media) petri plates using the uracil auxotroph *E. coli* strain OP50 as food source according to the methods described by Stiernagle.[134] The anthelmintic bioassay was carried out following the method developed by Thomsen et al..[135] In all the assays, the solvent DMSO (2%) and the standard anthelmintic drug ivermectin (10 µg/mL) were used as negative and positive controls, respectively. All the assays were carried out in triplicate.

2.5.9.3 Anticancer Bioassay

Cell lines and cultivation

Colon adenocarcinoma cells (HT-29) were kindly provided by Prof. B. Seliger, Immunology Department, Martin-Luther-University Halle-Wittenberg, Germany. Human refractory prostate cancer cells (PC3) was purchased from the German Collection of Microorganisms and Cell Cultures (Leibniz-DSMZ, Germany). Cells are regularly tested for mycoplasma infections and were cultured as follows. B16F10 and PC3: RPMI 1640 medium with L-glutamine and sodium hydrogen carbonate (Sigma-Aldrich, R8758), 10% fetal bovine serum (Sigma-Aldrich, F2442), penicillin-streptomycin solution (Sigma-Aldrich, P4333, 10 000 units/mL penicillin and 100 mg/mL streptomycin). HT-29: Dulbecco's Modified Eagle's Medium with 1000 mg/L glucose, L-glutamine, and sodium hydrogen carbonate (Sigma-Aldrich D6046), 10% fetal bovine serum (Sigma-Aldrich, F2442), penicillin-streptomycin solution (Sigma-Aldrich, P4333, 10 000 units/mL penicillin and 100 mg/mL streptomycin). All cells were cultured at 37 °C in a humidified incubator with a 5% CO_2 atmosphere.

MTT and CV assays

For HT-29 some 2000 cells/well were seeded. 6000 cells/well have been used for PC3 cell line. After 24 h from incubation, cells were treated with the tested compounds in various concentrations. Stock solutions of the investigated compounds were prepared at 20 mM concentration in DMSO. According to the experiments, suitable dilutions in the medium were prepared from the stock solution. MTT and CV assays were performed as described in the literature.[136] For the fast screening with MTT and CV assays, 2 concentrations have been used (0.05 and 50 µM) in quadruplicate. Digitonin (125 µM) was used as control. In this study, the incubation time for all compounds with different cell lines was 72 h. The absorbance of the dissolved dyes was measured in an automated microplate reader at 540 nm with a reference wavelength of 670 nm. The results are presented as a percentage of control values obtained from untreated cultures.

2.6. References

[1]   Grey-Wilson, C. *Hydrocera triflora*, Its floral morphology and relationship with *Impatiens*: Studies in Balsaminaceae: V. *Kew Bulletin* **1980**, *35*, 213-219.

[2]   Fischer, E.; Rahelivololona, M.E. New taxa of *Impatiens* (Balsaminaceae) from Madagascar. III. *Adansonia* **2004**, *26*, 37-52.

[3]   Janssens, S.B.; Smets, E.F.; Vrijdaghs, A. Floral development of *Hydrocera* and *Impatiens* reveals evolutionary trends in the most early diverged lineages of the Balsaminaceae. *Ann. Bot.* **2012**, *109*, 1285-1296.

[4]   Abrahamczyk, S.; Reinken, B.; Neumann, K.M. *Hydrocera triflora*, die weitgehend unbekannte Schwesterart der Springkräuter (*Impatiens*). **2015**, *79*, 16-18.

[5]   Yu, S.-X.; Janssens, S.B.; Zhu, X.-Y.; Lidén, M.; Gao, T.-G.; Wang, W. Phylogeny of *Impatiens* (Balsaminaceae): integrating molecular and morphological evidence into a new classification. **2016**, *32*, 179-197.

[6]   Li, Z.Z.; Saina, J.K.; Gichira, A.W.; Kyalo, C.M.; Wang, Q.F.; Chen, J.M. Comparative genomics of the Balsaminaceae sister genera *Hydrocera triflora* and *Impatiens pinfanensis*. *Int. J. Mol. Sci.* **2018**, *19*, 319-336.

[7]   Grey-Wilson, C. *Impatiens* of Africa; Morphology, Pollination and Pollinators, Ecology, Phytogeography, Hybridization, Keys and a Systematics of All African Species with a Note on Collecting and Cultivation; A A Balkema Publishers: The Netherlands, 1980, ISBN 9789061910411.

[8]   Janssens, S.B.; Knox, E.B.; Huysmans, S.; Smets, E.F.; Merckx, V.S. Rapid radiation of *Impatiens* (Balsaminaceae) during Pliocene and Pleistocene: result of a global climate change. *Mol. Phylogenet. Evol.* **2009**, *52*, 806-824.

[9]   Yuan, Y.M.; Song, Y.; Geuten, K.; Rahelivololona, E.; Wohlhauser, S.; Fischer, E.; Smets, E.; Küpfer, P. Phylogeny and biogeography of Balsaminaceae inferred from ITS sequences. *TAXON* **2004**, *53*, 391-404.

[10]   Tan, Y.-H.; Liu, Y.-N.; Jiang, H.; Zhu, X.-X.; Zhang, W.; Yu, S.-X. *Impatiens pandurata* (Balsaminaceae), a new species from Yunnan, China. *Bot. Stud.* **2015**, *56*, 29-35.

[11]   Gan, Q.; Li, X. *Impatiens zhuxiensis* (Balsaminaceae), a new species of Hubei, China. *Nord. J. Bot.* **2020**, *38*, 1-5.

[12]   Abrahamczyk, S.; Reinken, B.; Neumann, M.; Lobin, W. Die Ungeduldigen: Fleißige Lieschen, Balsaminen und Springkräuter (*Impatiens*, Balsaminaceae). *Der Palmengarten* **2016**, *80*, 31-39.

[13]   Vrchotova, N.; Sera, B.; Krejčová, N. Allelopathic activity of extracts from *Impatiens* species. *Plant Soil Environ.* **2011**, *57*, 57-60.

[14]   Rao, R.V.S.; Ayyangar, K.R.; Sampathkumar, R. On the karyological characteristics of some members of Balsaminaceae. *CYTOLOGIA* **1986**, *51*, 251-260.

[15]   Fischer, E.; Rahelivololona, M.E. New taxa of *Impatiens* (Balsaminaceae) from Madagascar. I. *Adansonia* **2002**, *24*, 271-294.

[16]   Bartomeus, I.; Vilà, M.; Steffan-Dewenter, I. Combined effects of *Impatiens glandulifera* invasion and landscape structure on native plant pollination. *J. Ecol.* **2010**, *98*, 440-450.

[17]   Skálová, H.; Jarošík, V.; Dvořáčková, Ś.; Pyšek, P. Effect of intra- and interspecific competition on the performance of native and invasive species of *Impatiens* under varying levels of shade and moisture. *PLoS One* **2013**, *8*, e62842.

[18]   Čuda, J.; Skálová, H.; Pyšek, P. Spread of *Impatiens glandulifera* from riparian habitats to forests and its associated impacts: insights from a new invasion. *Weed Res.* **2020**, *60*, 8-15.

[19]   Ishiguro, K.; Fukumoto, H.; Murashima, T.; Kuriyama, M.; Semma, M.; Isoi, K. Antianaphylactic effects of the ethanolic extract from the petals of *Impatiens balsamina* L. in mice. *Phytother. Res.* **1992**, *6*, 112-113.

[20]   Yang, X.; Summerhurst, D.K.; Koval, S.F.; Ficker, C.; Smith, M.L.; Bernards, M.A. Isolation of an antimicrobial compound from *Impatiens balsamina* L. using bioassay-guided fractionation. *Phytother. Res.* **2001**, *15*, 676-680.

[21]   Ueda, Y.; Oku, H.; Iinuma, M.; Ishiguro, K. Antianaphylactic and antipruritic effects of the flowers of *Impatiens textori* MIQ. *Biol. Pharm. Bull.* **2005**, *28*, 1786-1790.

[22]   Wang, Y.C.; Wu, D.C.; Liao, J.J.; Wu, C.H.; Li, W.Y.; Weng, B.C. In vitro activity of *Impatiens balsamina* L. against multiple antibiotic-resistant *Helicobacter pylori*. *Am. J. Chin. Med.* **2009**, *37*, 713-722.

[23]   Kumar, M.; Paul, Y.; Anand, V. An ethnobotanical study of medicinal plants used by the locals in Kishtwar, Jammu and Kashmir, India. *Ethnobot. Leafl.* **2009**, *13*, 1240-1256.

[24]   Imam, M.Z.; Nahar, N.; Akter, S.; Rana, M.S. Antinociceptive activity of methanol extract of flowers of *Impatiens balsamina*. *J. Ethnopharmacol.* **2012**, *142*, 804-810.

[25]   Li, Q.; Cao, J.; Yuan, W.; Li, M.; Yang, L.; Sun, Y.; Wang, X.; Zhao, Y. New triterpene saponins from flowers of *Impatiens balsamina* L. and their anti-hepatic fibrosis activity. *J. Funct. Foods* **2017**, *33*, 188-193.

[26]   Warburg, O.; Reiche, K. Balsaminaceae. In *Die Natürlichen Pflanzenfamilien*, Engler, H.G.A., Prantl, K.A.E., Eds. Wilhelm: 1895; Vol. 3, pp. 383-392.

[27]   Hooker, J.D. Les espe`ces du genre "*Impatiens*" dans l'herbier du Museum de Paris. *Nov. Arch. Mus. Nat. Hist. Paris.* **1908**, *10*, 233-272.

[28]   Fujihashi, H.; Akiyama, S.; Ohba, H. Origin and relationships of the Sino-Himalayan *Impatiens* (Balsaminaceae) based on molecular phylogenetic analysis, chromosome numbers and gross morphology. *J. Jap. Bot.* **2002**, *77*, 284-295.

[29]   Janssens, S.; Geuten, K.; Yuan, Y.-M.; Song, Y.; Küpfer, P.; Smets, E. Phylogenetics of *Impatiens* and *Hydrocera* (Balsaminaceae) using chloroplast atpB-rbcL spacer sequences. *Syst. Bot.* **2006**, *31*, 171-180.

[30]   Janssens, S.; Geuten, K.; Viaene, T.; Yuan, Y.M.; Song, Y.; Smets, E. Phylogenetic utility of the AP3/DEF K-domain and its molecular evolution in *Impatiens* (Balsaminaceae). *Mol. Phylogenet. Evol.* **2007**, *43*, 225-239.

[31]   Shajitha, P.P.; Dhanesh, N.R.; Ebin, P.J.; Laly, J.; Aneesha, D.; Reshma, J.; Augustine, J.; Linu, M. A combined chloroplast atpB–rbcL and trnL-F phylogeny unveils the ancestry of balsams (*Impatiens* spp.) in the Western Ghats of India. *3 Biotech* **2016**, *6*, 258.

[32]   Rahelivololona, E.M.; Fischer, E.; Janssens, S.B.; Razafimandimbison, Sylvain G. Phylogeny, infrageneric classification and species delimitation in the Malagasy *Impatiens* (Balsaminaceae). *PhytoKeys* **2018**, *110*, 51-67.

[33]   Peng, S. Taxonomic revision and phylogenetic study of the genus *Impatiens* in Kenya. Hunan Normal University, 2019.

[34]   Szewczyk, K.; Zidorn, C.; Biernasiuk, A.; Komsta, Ł.; Granica, S. Polyphenols from *Impatiens* (Balsaminaceae) and their antioxidant and antimicrobial activities. *Ind. Crops Prod.* **2016**, *86*, 262-272.

[35]   Szewczyk, K.; Olech, M. Optimization of extraction method for LC–MS based determination of phenolic acid profiles in different *Impatiens* species. *Phytochem. Lett.* **2017**, *20*, 322-330.

[36]   Szewczyk, K.; Heise, E.M.; Piwowarski, J.P. Preliminary characterization and bioactivities of some *Impatiens* L. water-soluble polysaccharides. *Molecules* **2018**, *23*, 631-644.

[37]   Chua, L.S. Untargeted MS-based small metabolite identification from the plant leaves and stems of *Impatiens balsamina*. *Plant Physiol. Biochem.* **2016**, *106*, 16-22.

[38]   Clevenger, S. The flavonols of *Impatiens balsamina* L. *Arch. Biochem. Biophys.* **1958**, *76*, 131-138.

[39]   Fukumoto, H.; Ishiguro, K.; Murashima, T.; Yamaki, M.; Isoi, K. Structure determination of a kaempferol 3-rhamnosyldiglucoside from *Impatiens balsamina*. *Phytochemistry* **1994**, *37*, 1486-1488.

[40]   Hasan, A.; Tahir, M.N. Flavonoids from the leaves of *Impatiens bicolor*. *Turk. J. Chem.* **2005**, *29*, 65-70.

[41]   Lei, J.; Qian, S.H.; Jiang, J.Q. Two new flavone glycosides from the seeds of *Impatiens balsamina* L. *J Asian Nat. Prod. Res.* **2010**, *12*, 1033-1037.

[42]   Vieira, M.N.; Winterhalter, P.; Jerz, G. Flavonoids from the flowers of *Impatiens glandulifera* Royle isolated by high performance countercurrent chromatography. *Phytochem. Anal.* **2016**, *27*, 116-125.

[43]   Szewczyk, K.; Sezai Cicek, S.; Zidorn, C.; Granica, S. Phenolic constituents of the aerial parts of *Impatiens glandulifera* Royle (Balsaminaceae) and their antioxidant activities. *Nat. Prod. Res.* **2019**, *33*, 2851-2855.

[44]   Shahwar, D.; Rehman, S.u.; Raza, M. Acetyl Cholinesterase inhibition potential and antioxidant activities of ferulic acid isolated from *Impatiens bicolor* Linn. *J. Med. Plant Res.* **2010**, *4*, 260-266.

[45]   Orzelska-Górka, J.; Szewczyk, K.; Gawrońska-Grzywacz, M.; Kędzierska, E.; Głowacka, E.; Herbet, M.; Dudka, J.; Biała, G. Monoaminergic system is implicated in the antidepressant-like effect of hyperoside and protocatechuic acid isolated from *Impatiens glandulifera* Royle in mice. *Neurochem. Int.* **2019**, *128*, 206-214.

[46]   Panjchayupakaranant, P.; Noguchi, H.; De-Eknamkul, W.; Sankawa, U. Naphthoquinones and coumarins from *Impatiens balsamina* root cultures. *Phytochemistry* **1995**, *40*, 1141-1143.

[47] Panichayupakaranant, P.; Noguchi, H.; De-Eknamkul, W. A New Biscoumarin from *Impatiens balsamina* root cultures. *Planta Med.* **1998**, *64*, 774-775.

[48] Chand, K.; Akanksha; Rahuja, N.; Mishra, D.P.; Srivastava, A.K.; Maurya, R. Major alkaloidal constituent from *Impatiens niamniamensis* seeds as antihyperglycemic agent. *Med. Chem. Res.* **2011**, *20*, 1505-1508.

[49] Glennie, C.W.; Bohm, B.A. The isolation of 2-hydroxy-1,4-naphthoquinone from *Impatiens balsamina* L. (Balsaminaceae). *Can. J. Biochem.* **1965**, *43*, 293-295.

[50] Fukumoto, H.; Yamaki, M.; Isoi, K.; Ishiguro, K. Antianaphylactic effects of the principal compounds from the white petals of *Impatiens balsamina* L. *Phytother. Res.* **1996**, *10*, 202-206.

[51] Panichayupakaranant, P. Naphthoquinone formation in cell cultures of *Impatiens balsamina*. *Pharm. Biol.* **2001**, *39*, 293-296.

[52] Ding, Z.S.; Jiang, F.S.; Chen, N.P.; Lv, G.Y.; Zhu, C.G. Isolation and identification of an anti-tumor component from leaves of *Impatiens balsamina*. *Molecules* **2008**, *13*, 220-229.

[53] Sakunphueak, A.; Panichayupakaranant, P. Comparison of antimicrobial activities of naphthoquinones from *Impatiens balsamina*. *Nat. Prod. Res.* **2012**, *26*, 1119-1124.

[54] Ruckli, R.; Hesse, K.; Glauser, G.; Rusterholz, H.P.; Baur, B. Inhibitory potential of naphthoquinones leached from leaves and exuded from roots of the invasive plant *Impatiens glandulifera*. *J. Chem. Ecol.* **2014**, *40*, 371-378.

[55] Li, Q.; Guo, Z.; Wang, K.; Zhang, X.; Lou, Y.; Zhao, Y.-q. Two new 1,4-naphthoquinone derivatives from *Impatiens balsamina* L. flowers. *Phytochem. Lett.* **2015**, *14*, 8-11.

[56] Ishiguro, K.; Ohira, Y.; Oku, H. Antipruritic dinaphthofuran-7,12-dione derivatives from the pericarp of *Impatiens balsamina*. *J. Nat. Prod.* **1998**, *61*, 1126-1129.

[57] Pei, H.; Lei, J.; Qian, S.H. A new cytotoxic dinaphthofuran-7,12-dione derivatives from the seeds of *Impatiens balsamina*. *Zhong yao cai = Zhongyaocai = J. Chin. med. mater.* **2012**, *35*, 407-410.

[58] Triska, J.; Vrchotova, N.; Sykora, J.; Moos, M. Separation and identification of 1,2,4-trihydroxynaphthalene-1-O-glucoside in *Impatiens glandulifera* Royle. *Molecules* **2013**, *18*, 8429-8439.

[59] Kim, D.H.; Lee, T.H.; Subedi, L.; Kim, S.Y.; Lee, K.R. Chemical constituents of *Impatiens balsamina* stems and their biological activities. *Nat. Prod. Sci.* **2019**, *25*, 130-135.

[60] Li, W.; Bi, X.; Wang, K.; Li, D.; Satou, T.; Koike, K. Triterpenoid saponins from *Impatiens siculifer*. *Phytochemistry* **2009**, *70*, 816-821.

[61] Motz, V.A.; Bowers, C.P.; Kneubehl, A.R.; Lendrum, E.C.; Young, L.M.; Kinder, D.H. Efficacy of the saponin component of *Impatiens capensis* Meerb. in preventing urushiol-induced contact dermatitis. *J. Ethnopharmacol.* **2015**, *162*, 163-167.

[62] Wang, Y.-C.; Li, W.-Y.; Wu, D.-C.; Wang, J.-J.; Wu, C.-H.; Liao, J.-J.; Lin, C.-K. *In vitro* activity of 2-methoxy-1,4-naphthoquinone and stigmasta-7,22-diene-3β-ol from *Impatiens balsamina* L. against multiple antibiotic-resistant *Helicobacter pylori*. *Evid. Based Complementary Altern. Med.* **2011**, *2011*, 704721.

[63] Cimmino, A.; Mathieu, V.; Evidente, M.; Ferderin, M.; Moreno Y Banuls, L.; Masi, M.; De Carvalho, A.; Kiss, R.; Evidente, A. Glanduliferins A and B, two new glucosylated steroids from *Impatiens glandulifera*, with *in vitro* growth inhibitory activity in human cancer cells. *Fitoterapia* **2016**, *109*, 138-145.

[64] Zhou, X.F.; Zhao, X.Y.; Tang, L.; Ruan, H.L.; Zhang, Y.H.; Pi, H.F.; Xiao, W.L.; Sun, H.D.; Wu, J.Z. Three new triterpenoid saponins from the rhizomes of *Impatiens pritzellii* var. *hupehensis*. *J. Asian Nat. Prod. Res.* **2007**, *9*, 379-385.

[65] Grabowska, K.; Podolak, I.; Galanty, A.; Żmudzki, P.; Koczurkiewicz, P.; Piska, K.; Pękala, E.; Janeczko, Z. Two new triterpenoid saponins from the leaves of *Impatiens parviflora* DC. and their cytotoxic activity. *Ind. Crops Prod.* **2017**, *96*, 71-79.

[66] Grabowska, K.; Wróbel, D.; Żmudzki, P.; Podolak, I. Anti-inflammatory activity of saponins from roots of *Impatiens parviflora* DC. *Nat. Prod. Res.* **2020**, *34*, 1581-1585.

[67] Hromádková, Z.; Košťálová, Z.; Vrchotová, N.; Ebringerová, A. Non-cellulosic polysaccharides from the leaves of small balsam (*Impatiens parviflora* DC.). *Carbohydr. Res.* **2014**, *389*, 147-153.

[68] Ortin, Y.; Evans, P. *trans*-Tetradec-2-enoic acid in *Impatiens glandulifera*. *Synth. Commun.* **2013**, *43*, 1404-1412.

[69] Szewczyk, K. Phytochemistry of the genus *Impatiens* (Balsaminaceae): A review. *Biochem. Syst. Ecol.* **2018**, *80*, 94-121.

[70] Anwer, N.; Waqar, M.; Mushtaq, M.; Sobia, A. Phytochemical analysis, free radical scavenging capacity and antimicrobial properties of *Impatiens bicolor* plant. *Int. Food Res. J.* **2012**, *20*, 99-103.

[71]   Tsushiro, K.; Kurizaki, A.; Watanabe, T.; Devkota, H.P. Chemical constituents from the aerial parts of *Impatiens hypophylla* Makino var. *hypophylla*. *Biochem. Syst. Ecol.* **2019**, *83*, 10-12.

[72]   Little, J.E.; Sproston, T.J.; Foote, M.W. Isolation and antifungal action of naturally occurring 2-methoxy-1,4-naphthoquinone. *J. Biol. Chem.* **1948**, *174*, 335-342.

[73]   Rodriguez, S.; Wolfender, J.-L.; Hakizamungu, E.; Hostettmann, K. An antifungal naphthoquinone, xanthones and secoiridoids from *Swertia calycina*. *Planta Med.* **1995**, *61*, 362-364.

[74]   Ueda, Y.; Oku, H.; Iinuma, M.; Ishiguro, K. Effects on blood pressure decrease in response to PAF of *Impatiens textori* MIQ. *Biol. Pharm. Bull.* **2003**, *26*, 1505-1507.

[75]   Nisar, M.; Qayum, M.; Shah, M.R.; Kaleem, W.A.; Ali, I.; Zia-Ul-Haq, M. Antimicrobial screening of *Impatiens bicolor* Royle. *Pak. J. Bot.* **2010**, *42*, 523-526.

[76]   Qayum, M.; Nisar, M.; Zia-Ul-Haq, M.; kaleem, W.; Inamullah, A.; Marwat, K. Biological screening of *Impatiens bicolor* Royle. *Pak. J. Bot.* **2010**, *42*, 355-359.

[77]   Qayum, M.; Nisar, M.; Khan, B.; Akhtar, M.; kaleem, W.; Khan, M.I.; Mahmood, T.; Shahid, S.; Zia-Ul-Haq, M. Nematicidal potential of *Impatiens bicolor* Royle. *J. Med. Plant Res.* **2011**, *5*, 5138-5141.

[78]   Nisar, M.; Qayum, M.; Cavar, S.; Shah, M.R.; Zia-Ul-Haq, M.; Khan, I.; Ahmad, K.W.; Qayum, Z.A. Chemical constituents and antioxidant activity of *n*-hexane extract of *Impatiens bicolor*. *Chem. Nat. Compd.* **2012**, *48*, 143-146.

[79]   Jinfeng, Y. Antioxidant, antiproliferative, and α-glucosidase inhibitory activities of extracts from *Impatiens textori* Miq. *J. Med. Plants Res.* **2012**, *6*, 391-397.

[80]   Kang, S.N.; Goo, Y.M.; Yang, M.R.; Ibrahim, R.I.; Cho, J.H.; Kim, I.S.; Lee, O.H. Antioxidant and antimicrobial activities of ethanol extract from the stem and leaf of *Impatiens balsamina* L. (Balsaminaceae) at different harvest times. *Molecules* **2013**, *18*, 6356-6365.

[81]   Mushtaq, M.; Anwer, N.; Waqar, M.A.; Latif, S.; Shahid, S.A.; Azam, A. Antioxidant, antimicrobial potential and phytochemical attributes of *Impatiens edgeworthii*. *Asian J. Chem.* **2013**, *25*, 9800-9804.

[82]   Jiang, H.F.; Zhuang, Z.H.; Hou, B.W.; Shi, B.J.; Shu, C.J.; Chen, L.; Shi, G.X.; Zhang, W.M. Adverse effects of hydroalcoholic extracts and the major components in the stems of *Impatiens balsamina* L. on *Caenorhabditis elegans*. *Evid. Based Complement. Alternat. Med.* **2017**, *2017*, 4245830.

[83]   Szewczyk, K.; Orzelska-Górka, J.; Polakowska, M.; Biała, G. Antinociceptive and antianxiety activity of hydroethanolic extracts of three *Impatiens* species in mice. *Acta Pol. Pharm.* **2018**, *75*, 989-1001.

[84]   Paun, G.; Neagu, E.; Moroeanu, V.; Albu, C.; Ursu, T.-M.; Zanfirescu, A.; Negres, S.; Chirita, C.; Radu, G.L. Anti-inflammatory and antioxidant activities of the *Impatiens noli-tangere* and *Stachys officinalis* polyphenolic-rich extracts. *Rev. Bras. Farmacogn.* **2018**, *28*, 57-64.

[85]   Ruchisansakun, S.; van der Niet, T.; Janssens, S.B.; Triboun, P.; Techaprasan, J.; Jenjittikul, T.; Suksathan, P. Phylogenetic analyses of molecular data and reconstruction of morphological character evolution in asian *Impatiens* section Semeiocardium (Balsaminaceae). *Syst. Bot.* **2016**, *40*, 1063-1074.

[86]   Creek, D.J.; Dunn, W.B.; Fiehn, O.; Griffin, J.L.; Hall, R.D.; Lei, Z.; Mistrik, R.; Neumann, S.; Schymanski, E.L.; Sumner, L.W., et al. Metabolite identification: are you sure? And how do your peers gauge your confidence? *Metabolomics* **2014**, *10*, 350-353.

[87]   Llorent-Martínez, E.J.; Zengin, G.; Lobine, D.; Molina-García, L.; Mollica, A.; Mahomoodally, M.F. Phytochemical characterization, *in vitro* and in silico approaches for three *Hypericum* species. *New J. Chem.* **2018**, *42*, 5204-5214.

[88]   Dou, J.; Lee, V.S.Y.; Tzen, J.T.C.; Lee, M.-R. Identification and comparison of phenolic compounds in the preparation of Oolong tea manufactured by semifermentation and drying processes. *J. Agric. Food Chem.* **2007**, *55*, 7462-7468.

[89]   Jiménez-López, J.; Ruiz-Medina, A.; Ortega-Barrales, P.; Llorent-Martínez, E.J. *Rosa rubiginosa* and *Fraxinus oxycarpa* herbal teas: characterization of phytochemical profiles by liquid chromatography-mass spectrometry, and evaluation of the antioxidant activity. *New J. Chem.* **2017**, *41*, 7681-7688.

[90]   De Beer, D.; Schulze, A.E.; Joubert, E.; De Villiers, A.; Malherbe, C.J.; Stander, M.A. Food ingredient extracts of *Cyclopia subternata* (Honeybush): variation in phenolic composition and antioxidant capacity. *Molecules* **2012**, *17*, 14602-14624.

[91]   Szajwaj, B.; Moldoch, J.; Masullo, M.; Piacente, S.; Oleszek, W.; Stochmal, A. Amides and esters of phenylpropenoic acids from the aerial parts of *Trifolium pallidum*. *Nat. Prod. Commun.* **2011**, *6*, 1293-1296.

[92]   Li, S.-S.; Wu, J.; Chen, L.-G.; Du, H.; Xu, Y.-J.; Wang, L.-J.; Zhang, H.-J.; Zheng, X.-C.; Wang, L.-S. Biogenesis of *C*-glycosyl flavones and profiling of flavonoid glycosides in Lotus (*Nelumbo nucifera*). *PLOS ONE* **2014**, *9*, e108860.

[93]   Llorent-Martínez, E.J.; Ortega-Barrales, P.; Zengin, G.; Mocan, A.; Simirgiotis, M.J.; Ceylan, R.; Uysal, S.; Aktumsek, A. Evaluation of antioxidant potential, enzyme inhibition activity and phenolic profile of *Lathyrus cicera* and *Lathyrus digitatus*: Potential sources of bioactive compounds for the food industry. *Food Chem. Toxicol.* **2017**, *107*, 609-619.

[94]   Zhang, Y.; Xiong, H.; Xu, X.; Xue, X.; Liu, M.; Xu, S.; Liu, H.; Gao, Y.; Zhang, H.; Li, X. Compounds identification in Semen Cuscutae by ultra-high-performance liquid chromatography (UHPLC) coupled to electrospray ionization mass spectrometry. *Molecules* **2018**, *23*, 1199-1215.

[95]   Kachlicki, P.; Piasecka, A.; Stobiecki, M.; Marczak, Ł. Structural characterization of flavonoid glycoconjugates and their derivatives with mass spectrometric techniques. *Molecules* **2016**, *21*, 1494-1515.

[96]   Cuyckens, F.; Claeys, M. Mass spectrometry in the structural analysis of flavonoids. *J. Mass Spectrom.* **2004**, *39*, 1-15.

[97]   Yang, W.-Z.; Qiao, X.; Bo, T.; Wang, Q.; Guo, D.-A.; Ye, M. Low energy induced homolytic fragmentation of flavonol 3-*O*-glycosides by negative electrospray ionization tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2014**, *28*, 385-395.

[98]   Barron, D.; Varin, L.; Ibrahim, R.K.; Harborne, J.B.; Williams, C.A. Sulphated flavonoids—an update. *Phytochemistry* **1988**, *27*, 2375-2395.

[99]   Teles, Y.C.F.; Souza, M.S.R.; Souza, M.d.F.V.d. Sulphated flavonoids: biosynthesis, structures, and biological activities. *Molecules* **2018**, *23*, 480-491.

[100]   Kleinenkuhnen, N.; Büchel, F.; Gerlich, S.C.; Kopriva, S.; Metzger, S. A novel method for identification and quantification of sulfated flavonoids in plants by neutral loss scan mass spectrometry. *Front. Plant Sci.* **2019**, *10*, 885.

[101]   Bylka, W.; Stobiecki, M.; Frański, R. Sulphated flavonoid glycosides from leaves of *Atriplex hortensis*. *Acta Physiol. Plant.* **2001**, *23*, 285-290.

[102]   Davis, B.D.; Brodbelt, J.S. Determination of the glycosylation site of flavonoid monoglucosides by metal complexation and tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1287-1299.

[103]   Figueirinha, A.; Paranhos, A.; Pérez-Alonso, J.J.; Santos-Buelga, C.; Batista, M.T. *Cymbopogon citratus* leaves: Characterization of flavonoids by HPLC–PDA–ESI/MS/MS and an approach to their potential as a source of bioactive polyphenols. *Food Chem.* **2008**, *110*, 718-728.

[104]   Oku, H.; Ishiguro, K. Cyclooxygenase-2 inhibitory 1,4-naphthoquinones from *Impatiens balsamina* L. *Biol. Pharm. Bull.* **2002**, *25*, 658-660.

[105]   Worley, B.; Powers, R. Multivariate analysis in metabolomics. *Curr. Metabolomics* **2013**, *1*, 92-107.

[106]   Boccard, J.; Rudaz, S. Harnessing the complexity of metabolomic data with chemometrics. *J. Chemom.* **2014**, *28*, 1-9.

[107]   Gad, H.A.; El-Ahmady, S.H.; Abou-Shoer, M.I.; Al-Azizi, M.M. Application of chemometrics in authentication of herbal medicines: A review. *Phytochem. Anal.* **2013**, *24*, 1-24.

[108]   Salem, M.A.; Perez de Souza, L.; Serag, A.; Fernie, A.R.; Farag, M.A.; Ezzat, S.M.; Alseekh, S. Metabolomics in the context of plant natural products research: From sample preparation to metabolite analysis. *Metabolites* **2020**, *10*, 37-67.

[109]   Porzel, A.; Farag, M.A.; Mülbradt, J.; Wessjohann, L.A. Metabolite profiling and fingerprinting of *Hypericum* species: a comparison of MS and NMR metabolomics. *Metabolomics* **2014**, *10*, 574-588.

[110]   Farag, M.A.; Otify, A.; Porzel, A.; Michel, C.G.; Elsayed, A.; Wessjohann, L.A. Comparative metabolite profiling and fingerprinting of genus *Passiflora* leaves using a multiplex approach of UPLC-MS and NMR analyzed by chemometric tools. *Anal. Bioanal. Chem.* **2016**, *408*, 3125-3143.

[111]   Lobstein, A.; Brenne, X.; Feist, E.; Metz, N.; Weniger, B.; Anton, R. Quantitative determination of naphthoquinones of *Impatiens* species. *Phytochem. Anal.* **2001**, *12*, 202-205.

[112]   Fernando de Carvalho da, S.; Vitor Francisco, F. Natural naphthoquinones with great importance in medicinal chemistry. *Curr. Org. Synth.* **2016**, *13*, 334-371.

[113]   Rho, Y.S.; Kim, S.Y.; Kim, W.J.; Yun, Y.K.; Sin, H.S.; Yoo, D.J. Convenient syntheses of daunomycinone-7-*D*-glucuronides and doxorubicinone-7-*D*-glucuronides. *Synth. Commun.* **2004**, *34*, 3497-3511.

[114]   Ravichandiran, P.; Sheet, S.; Premnath, D.; Kim, A.R.; Yoo, D.J. 1,4-Naphthoquinone analogues: Potent antibacterial agents and mode of action evaluation. *Molecules* **2019**, *24*, 1437-1451.

[115] Ma, W.-D.; Zou, Y.-P.; Wang, P.; Yao, X.-H.; Sun, Y.; Duan, M.-H.; Fu, Y.-J.; Yu, B. Chimaphilin induces apoptosis in human breast cancer MCF-7 cells through a ROS-mediated mitochondrial pathway. *Food Chem. Toxicol.* **2014**, *70*, 1-8.

[116] Wellington, K.W. Understanding cancer and the anticancer activities of naphthoquinones – a review. *RSC Adv.* **2015**, *5*, 20309-20338.

[117] Widhalm, J.R.; Rhodes, D. Biosynthesis and molecular actions of specialized 1,4-naphthoquinone natural products produced by horticultural plants. *Hortic. Res.* **2016**, *3*, 16046-16063.

[118] Novais, J.S.; Campos, V.R.; Silva, A.C.J.A.; de Souza, M.C.B.V.; Ferreira, V.F.; Keller, V.G.L.; Ferreira, M.O.; Dias, F.R.F.; Vitorino, M.I.; Sathler, P.C., et al. Synthesis and antimicrobial evaluation of promising 7-arylamino-5,8-dioxo-5,8-dihydroisoquinoline-4-carboxylates and their halogenated amino compounds for treating gram-negative bacterial infections. *RSC Adv.* **2017**, *7*, 18311-18320.

[119] Manickam, M.; Boggu, P.R.; Cho, J.; Nam, Y.J.; Lee, S.J.; Jung, S.-H. Investigation of chemical reactivity of 2-alkoxy-1,4-naphthoquinones and their anticancer activity. *Bioorg. Med. Chem. Lett.* **2018**, *28*, 2023-2028.

[120] Pullella, G.A.; Wild, D.A.; Nealon, G.L.; Elyashberg, M.; Piggott, M.J. What Is the structure of the antitubercular natural product eucapsitrione? *J. Org. Chem.* **2017**, *82*, 7287-7299.

[121] Lanfranchi, D.A.; Cesar-Rodo, E.; Bertrand, B.; Huang, H.-H.; Day, L.; Johann, L.; Elhabiri, M.; Becker, K.; Williams, D.L.; Davioud-Charvet, E. Synthesis and biological evaluation of 1,4-naphthoquinones and quinoline-5,8-diones as antimalarial and schistosomicidal agents. *Org. Biomol. Chem.* **2012**, *10*, 6375-6387.

[122] Lara, L.S.; Moreira, C.S.; Calvet, C.M.; Lechuga, G.C.; Souza, R.S.; Bourguignon, S.C.; Ferreira, V.F.; Rocha, D.; Pereira, M.C.S. Efficacy of 2-hydroxy-3-phenylsulfanylmethyl-1,4-naphthoquinone derivatives against different *Trypanosoma cruzi* discrete type units: Identification of a promising hit compound. *Eur. J. Med. Chem.* **2018**, *144*, 572-581.

[123] Exner, M.; Bhattacharya, S.; Christiansen, B.; Gebel, J.; Goroncy-Bermes, P.; Hartemann, P.; Heeg, P.; Ilschner, C.; Kramer, A.; Larson, E., et al. Antibiotic resistance: What is so special about multidrug-resistant gram-negative bacteria? *GMS Hyg. Infect. Control* **2017**, *12*, 1-24.

[124] Williams, A.H. Dibenzoylmethanes and flavones of *Malus*. *Phytochemistry* **1979**, *18*, 1897-1898.

[125] Rui-Lin, N.; Tanaka, T.; Zhou, J.; Tanaka, O. Phlorizin and trilobatin, sweet dihydrochalcone-glucosides from leaves of *Lithocarpus litseifolius* (Hance) Rehd. (Fagaceae). *Agr. Biol. Chem.* **1982**, *46*, 1933-1934.

[126] Qin, X.; Xing, Y.; Zhou, Z.; Yao, Y. Dihydrochalcone Compounds isolated from crabapple leaves showed anticancer effects on human cancer cell lines. *Molecules* **2015**, *20*, 21193-21203.

[127] Ghavam-Haghi, F.; Sadeghi Dinani, M. Isolation and identification of astragalin and 2-methoxy tyrosol from the bulbs of *Allium paradoxum*. *J. Herbmed. Pharmacol.* **2017**, *6*, 114-118.

[128] Hegde, V.R.; Pu, H.; Patel, M.; Das, P.R.; Butkiewicz, N.; Arreaza, G.; Gullo, V.P.; Chan, T.-M. Two antiviral compounds from the plant *Stylogne cauliflora* as inhibitors of HCV NS3 protease. *Bioorganic Med. Chem. Lett.* **2003**, *13*, 2925-2928.

[129] Martens, L.; Chambers, M.; Sturm, M.; Kessner, D.; Levander, F.; Shofstahl, J.; Tang, W.H.; Rompp, A.; Neumann, S.; Pizarro, A.D., et al. mzML--a community standard for mass spectrometry data. *Mol. Cell Proteomics* **2011**, *10*, R110.000133.

[130] Tautenhahn, R.; Patti, G.J.; Rinehart, D.; Siuzdak, G. XCMS Online: a web-based platform to process untargeted metabolomic data. *Anal. Chem.* **2012**, *84*, 5035-5039.

[131] Mori, N.; Toume, K.; Arai, M.A.; Koyano, T.; Kowithayakorn, T.; Ishibashi, M. 2-Methoxy-1,4-naphthoquinone isolated from *Impatiens balsamina* in a screening program for activity to inhibit Wnt signaling. *J. Nat. Med.* **2011**, *65*, 234-236.

[132] Stark, S. Utilization of the Ugi four-component reaction for the synthesis of lipophilic peptidomimetics as potential antimicrobials. Martin-Luther-University Halle-Wittenberg, Halle (Saale), **2016**.

[133] Dos Santos, C.H.C.; de Carvalho, M.G.; Franke, K.; Wessjohann, L. Dammarane-type triterpenoids from the stem of *Ziziphus glaziovii* Warm. (Rhamnaceae). *Phytochemistry* **2019**, *162*, 250-259.

[134] Stiernagle, T. Maintenance of *C. elegans*. In Wormbook WormBook; Editor: The *C. elegans* Research Community. Available online: http://dx.doi.org/10.1895/wormbook.1.101.1 (accessed on 17.06.2020).

[135] Thomsen, H.; Reider, K.; Franke, K.; Wessjohann, L.A.; Keiser, J.; Dagne, E.; Arnold, N. Characterization of constituents and anthelmintic properties of *Hagenia abyssinica*. *Sci. Pharm.* **2012**, *80*, 433-446.

[136] Krajnović T:, Kaluđerović G.N.; Wessjohann L.A.; Mijatović S.; Maksimović-Ivanić D. Versatile antitumor potential of isoxanthohumol: Enhancement of paclitaxel activity *in vivo*. *Pharmacol. Res.* **2016**, *105*, 62-73.

Chapter 3

-

Isolation, Identification and Total Synthesis of Two New Cyclic Pentapeptides from *Sepedonium microspermum* Besl

*Abstract**

Two new cyclic pentapeptides (**3.1**, **3.2**), named microsporide A and B, were detected within the extract of the fungi *Sepedonium microspermum* Besl (strain KSH 584) besides the known cyclic pentapeptide chrysosporide (**3.3**). Their primary structures were elucidated by extensive (+)-ESI-HRMS^{n} analysis. The structural elucidation with NMR was supported by a comparison with reference compounds obtained via solid-phase peptide synthesis followed by in-solution cyclization. A mass spectrometry-guided isolation yielded cyclic peptide **3.1** as an isomeric mixture and **3.3** as pure compound. The absolute configuration of all chiral amino acids in the natural cyclic pentapeptides **3.2** and **3.3** and for one stereoisomer of **3.1** could be unambiguously determined. The amino acid sequence of the stereoisomer **3.1a** was thus identified as cyclo-(D-Ala-L-Leu-L-Leu-D-Val-L-Val), while **3.2** differs in an exchange of D-Val for D-Leu and L-Val for L-Ala, respectively. The structure of cyclic peptide **3.3** was confirmed as cyclo-(D-Ala-L-Leu-L-Leu-D-Leu-L-Val). Furthermore, combined NMR and CD studies suggest the existence of β-turn conformations.

3.1 Introduction

Mycoparasitic species of the genus *Sepedonium* (Hypocreaceae, Ascomycota) are well-known to produce a notable variety of structurally diverse secondary metabolites. Besides tropolones,[1,2] mono- and bisanthraquinones,[3] an isoquinoline alkaloid[4] and azaphilone,[5] various peptide-based compounds like peptaibols,[6-9] and also one cyclic pentapeptide[10] were described in recent years.

Homodetic cyclic peptides are an interesting group of macrocyclic oligopeptides present in different organisms from plants, fungi, bacteria, sponges, algae and mammals.[11-12] In particular, filamentous fungi produce this class of natural products in remarkable structural diversity, ranging from tri- to octadeca(cyclic)peptides.[11] A large number of fungal cyclic peptides as well as peptaibols are biosynthesized by non-ribosomal peptide synthetases (NRPS),[13,14] enabling the incorporation of non-proteogenic amino acids – including D-configured homologues during their biosynthesis.[15]

Fungal cyclic pentapeptides have been recognized in the genera *Aspergillus*, *Fusarium*, *Hamigera*, *Penicillum*, *Pseudallescheria,* and *Xylaria*.[11] They exhibit a high structural diversity containing e.g. *N*-methylation,[16] anthranilic acid moieties,[17,18] and chlorinated compounds.[19] From a New Zealand sample of *Sepedonium chrysospermum* (strain CANU E609) the cyclic pentapeptide chrysosporide (**3.3**) was described only comprising aliphatic amino acids in the primary sequences.[10] Similar compounds were also isolated from different endophytic and marine fungi.[20-24] Furthermore, fungal cyclic pentapeptides show a wide range of bioactivities like antiviral,[25] antibacterial,[26] insecticidal,[27] antiplasmodial,[28] cytotoxic,[29] and chitinase inhibiting properties.[30,31]

3.2 Aims and Scope

This chapter reports on the recognition, the structural elucidation and total synthesis of two new cyclic pentapeptides, named microsporide A (**3.1**) and B (**3.2**), besides the known chrysosporide (**3.3**) from *Sepedonium microspermum* Besl (strain KSH 584). Their primary sequences were established using intensive (+)-ESI-HRMS^n analysis. The absolute configuration of microsporide A (**3.1**) and B (**3.2**) were determined by comparison with the corresponding synthetic products using cicular dichroism (CD) spectroscopy. Furthermore, results of biological evaluation results are discussed.

3.3 Results and Discussion

3.3.1. Structural elucidation and isolation of cyclic pentapeptides 3.1-3.3

As part of ongoing studies for new peptidogenic metabolites based on our highly diverse *Sepedonium* strain collection, the culture extract of *Sepedonium mircospermum* was investigated using an UHPLC-(+)-ESI-HRMS screening approach.[32] From an enriched fraction of the crude extract of *S. microspermum*, the unknown cyclic pentapeptides (**3.1**, **3.2**, Figure 3.1E) as well as

the known chrysosporide[10] (**3.3**, Figure 3.1 E) could be detected. They occur as traces in comparison to the main component (t_R 16.78 min), which was annotated as the peptaibol ampullosporin A[6] (Figure 3.1 A, Appendix, Figure B.1-B.2).
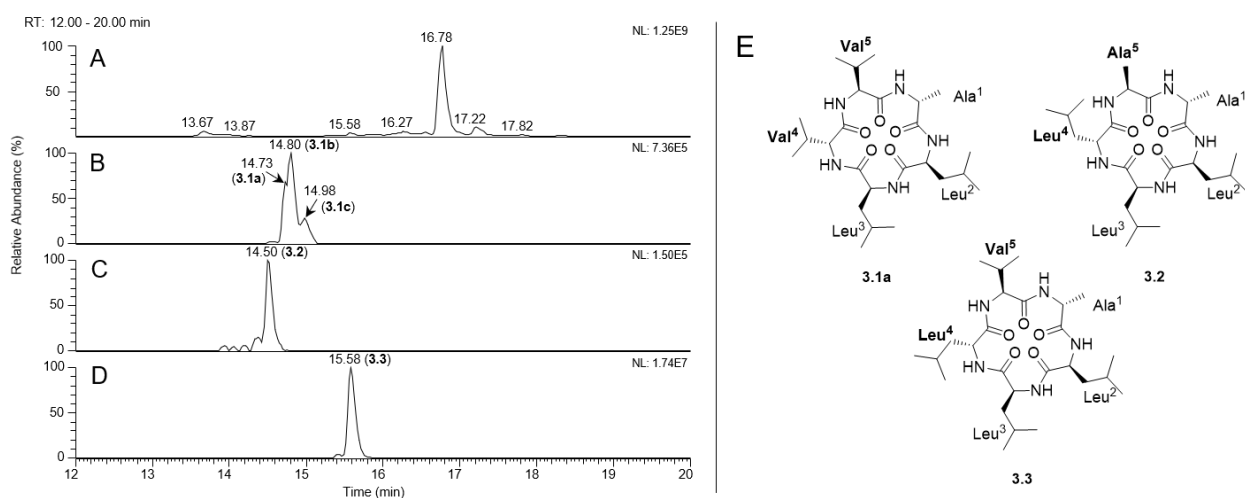
Figure 3.1. Total ion chromatogram (TIC, A) and extracted ion chromatograms (EIC) of cyclic pentapeptides **3.1** (C, *m/z* 496), **3.2** (B, *m/z* 482) and **3.3** (D, *m/z* 510) obtained from UHPLC-(+)-ESI-HRMS measurements of enriched extracts of *S. microspermum* (concentration 1 mg/mL). (E) Structures of new cyclic pentapeptides (**3.1a**, **3.2**), and chrysosporide (**3.3**).

The amino acid sequences of the cyclic pentapeptides **3.1**-**3.3** were determined by extensive (+)-ESI-HRMS^n studies (Appendix, Table B.1). The obtained complex HRMS^2 spectra display several series of both **a-** and **b**-ions. Furthermore, ions of type [M+H-H$_2$O]$^+$, [M+H-CO]$^+$ and [M+H-CO-NH$_3$]$^+$ also gave hints to the presence of a cyclic peptide structure. Cyclic peptides undergo multiple and indiscriminate ring-opening pathways during collision-induced dissociation (CID, as exemplary shown for **3.3** in Figure 3.2).[33,34] In analogy to the classical peptide sequencing a series of **a-** and **b**-ions are formed. It should be noted, that according to the literature not all theoretical ion series are detected during MS fragmentation experiments.[35] The fragment ion nomenclature was adapted from Ngoka and Gross, whereby the first letter is representing the *C*-terminal amino acid and the second the *N*-terminus.[36]

The [M+H]$^+$ ion of the known cyclic pentapeptide chrysosporide (**3.3**) at *m/z* 510.3650 (calcd. for [C$_{26}$H$_{48}$N$_5$O$_5$]$^+$ *m/z* 510.3650) was detected at a UHPLC retention time of 15.58 min (Figure 3.1 D). The primary sequence of the cyclic pentapeptide **3.3** as cyclo-(L-Val-D-Ala-L-Leu-L-Leu-D-Leu) was originally determined by comparison of molecular mechanic calculations with observed NOEs.[10] We could now confirm the primary sequence of **3.3** by fragmentation experiments in the MS^2/MS^3 mode.

From chrysosporide (**3.3**) five linear peptide acylium ions can be formed (Figure 3.2). The exact *m/z* values and the corresponding elemental composition detected in the Orbitrap analyzer enabled the assignment of the fragment ions unambiguously. The ESI-HRMS^2 spectrum of **3.3** with *m/z* 510 ([M+H]$^+$) shows a base peak at *m/z* 482.3699 (calcd. for [C$_{25}$H$_{48}$N$_5$O$_4$]$^+$ *m/z* 482.3701,

Figure 3.2, Appendix, Table B.1), representing the loss of carbon monoxide (CO) after cleavage of the ring structure.
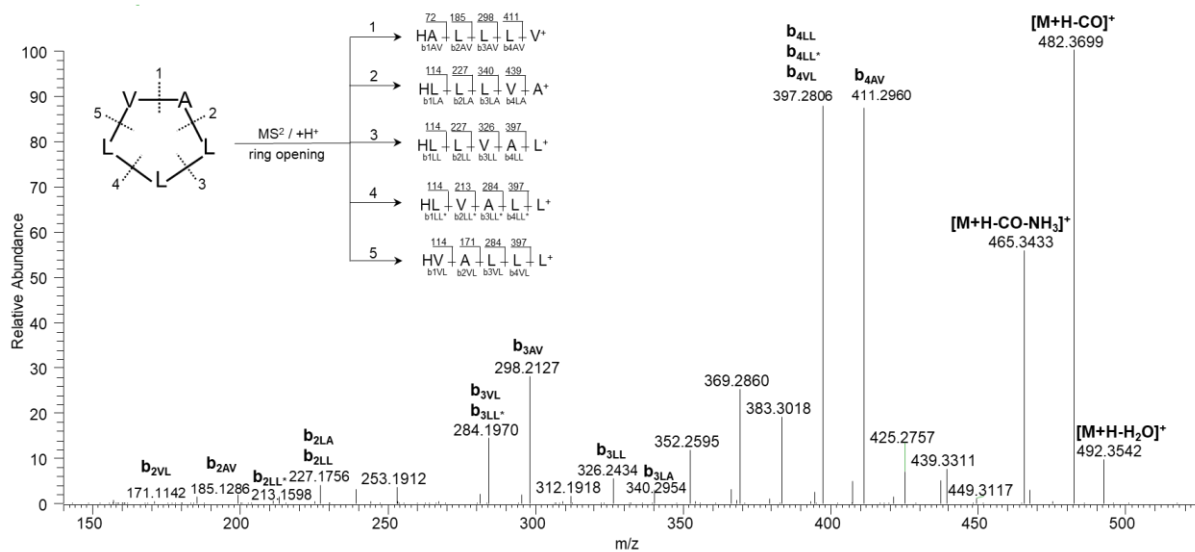
Figure 3.2. (+)-ESI-HRMS^2-spectrum of the [M+H]$^+$ ion at m/z 510 of the cyclopentapeptide chrysosporide (**3.3**, cyclo-(D-Ala-L-Leu-L-Leu-D-Leu-L-Val)) obtained from the UHPLC-(+)-ESI-HRMS measurements of the *S. microspermum* (KSH 584) extract.

Besides a series of **b**-type ions (**b2**, **b3**, and **b4**), the **a**-type ion series was detected (Appendix, Table B.1). Furthermore, the **b**-ions can lose both NH$_3$ and CO (see Appendix, Table B.1). However, the linear sequence of the **b4LA**-ion (HL-L-L-V$^+$) is not detectable, which suggests that the direct neutral loss of the "dipeptide V-A" results in the **b3LA**-ion at m/z 340.2599 (calcd. for [C$_{18}$H$_{34}$N$_3$O$_3$]$^+$ m/z 340.2595, Appendix, Table B.1). For further information, MS^3 experiments of the **b4** ions were carried out. So, the MS^3 spectrum of m/z 411 (**b4AV**) shows a loss of alanine from the *N*-terminus forming an ion at m/z 340 ([**b4AV**-Ala]$^+$) besides the ions of type **b3AV** (m/z 298) and **b2AV** (m/z 285). Based on the neutral losses from the **b**-type ions, the primary structure of compound **3.3** was verified as cyclo-(Ala^1-Leu^2-Leu^3-Leu^4-Val^5). Based on this results, the detected **b**-type fragment ions of **3.3** were used for annotation of the new compounds **3.1** and **3.2** in their corresponding ESI-HRMS^2 spectra.

The peaks with t$_R$ = 14.73 min, t$_R$ = 14.80 min, t$_R$ = 14.98 exhibit the identical elemental composition (C$_{25}$H$_{45}$N$_5$O$_5$) as deduced from (+)-ESI-HRMS measurements of the [M+H]$^+$ ions at m/z 496.3486 (calcd. for [C$_{25}$H$_{46}$N$_5$O$_5$]$^+$ m/z 496.3480). Surprisingly, the MS^2 and MS^3 studies demonstrate identical fragmentation patterns (Appendix, Table B.1). The different retention times of the compounds between t$_R$ = 14.73 – 14.80 min suggested different absolute configurations of underlying amino acids. Therefore, the peaks were assigned to **3.1a** (t$_R$ = 14.73 min), **3.1b** (t$_R$ = 14.80 min), and **3.1c** (t$_R$ = 14.98 min). The molecular ion of **3.1a-c** revealed a mass difference of 14 amu in comparison to **3.3** corresponding to a lack of a CH$_2$-group in an amino acid side chain suggesting a replacement of a leucine by valine (Figure 3.3).
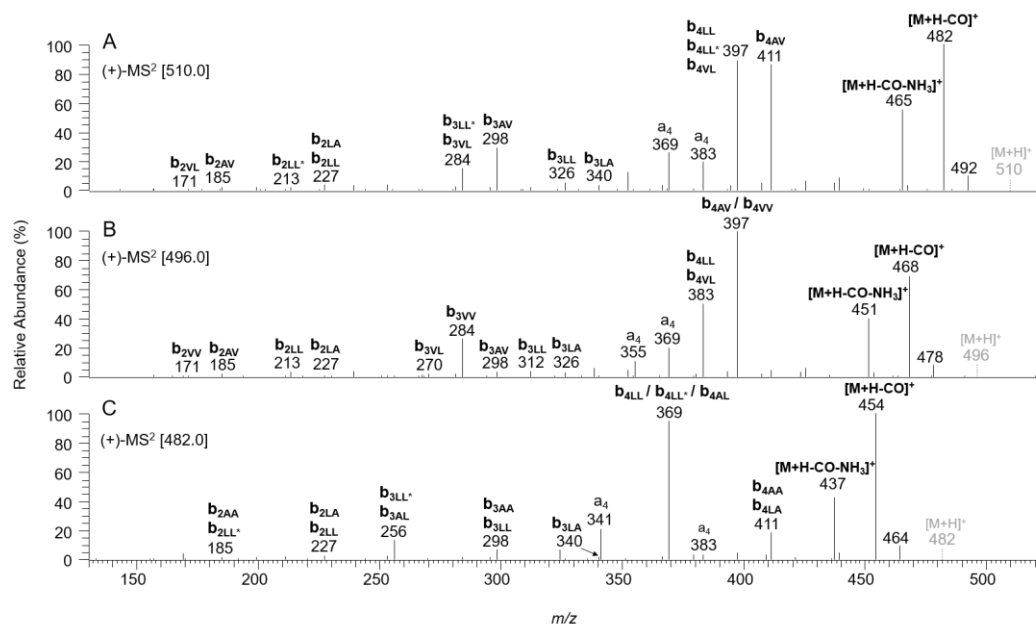
Figure 3.3. Comparison of the (+)-ESI-HRMS^2 spectra of the known chrysosporide (**3.3**, A, *m/z* 510) to the new cyclic pentapeptides **3.1** (B, *m/z* 496, **3.1a**, t_R = 14.8 min), and **3.2** (C, *m/z* 482) detected in *S. microspermum* (fragment ions are shown as nominal *m/z* values for better clarity).

High-resolution CID experiments in the MS^2/MS^3 mode (Figure 3.3, Appendix, Table B.1, Scheme B.1) show a series of **b**-type ions displaying the four different linear peptide fragments originated by scissions of the amide bonds. As in compound **3.3**, a **b$_4$**-ion, in particular **b$_{4LA}$** ion (HL-L-V-V$^+$), was not detectable. In accordance with **3.3**, the cyclic pentapeptides **3.1a-c** also show a **b$_4$**-fragment ion at *m/z* 397 (Figure 3.3). This ion corresponds to both linear acylium ions HV-A-L-L$^+$ (**b$_{4VV}$**) and HA-L-L-V$^+$ (**b$_{4AV}$**) formed by the loss of a valine from the protonated *C*-terminus. The linear sequences of these **b$_4$** ions were confirmed by a MS^3 experiment (Appendix, Table B.1). A series of **b$_3$** and **b$_2$** ions as well as related fragment **a** could be detected. Additionally, the loss of valine from the *C*-terminus of the **b$_{4VV}$** ion at *m/z* 298.2128 (calcd. for [C$_{15}$H$_{28}$N$_3$O$_3$]$^+$ *m/z* 298.2125, Table B.1) and the loss of alanine from the **b$_{4AV}$** ion at *m/z* 326.2433 (calcd. for [C$_{17}$H$_{32}$N$_3$O$_3$]$^+$ *m/z* 326.2438, Appendix, Table B.1) were observed. Based on the obtained fragment ion series, the primary sequence of the cyclic pentapeptide **3.1** was established as cyclo-(Ala^1-Leu^2-Leu^3-Val^4-Val^5) and named microsporide A.

Compound **3.2** (t_R 14.5 min) shows a [M+H]$^+$ ion at *m/z* 482.3330 (calcd. for [C$_{24}$H$_{44}$N$_5$O$_5$]$^+$ *m/z* 482.3337, Appendix, Table B.1). No isomers of **3.2** as compared to **3.1a-c** could be observed, it differs by -28 amu towards lower masses compared with chrysosporide (**3.3**). This fact indicates two CH$_2$-units less in the amino acid side chains. In case of compound **3.2,** all five linear acylium ions, represented by the **b$_4$** ions at *m/z* 411.2967 (calcd. for [C$_{21}$H$_{39}$N$_4$O$_4$]$^+$ *m/z* 411.2966; **b$_{4AA}$**, **b$_{4LA}$**) and *m/z* 369.2493 (calcd. for [C$_{18}$H$_{33}$N$_4$O$_4$]$^+$ *m/z* 369.2496; **b$_{4LL}$**, **b$_{4LL*}$**, **b$_{4AL}$**) were detectable (Appendix, Table B.1, Scheme B.2). In analogy to **3.3**, the cyclic pentapeptide **3.2** shows the **b$_4$**-fragment ion at *m/z* 411 (Figure 3.3). This ion corresponds to the linear acylium ion HA-L-L-L$^+$ (**b$_{4AA}$/b$_{4LA}$**), formed by the loss of alanine from the *C*-terminus.

87

The ESI-HRMS[3] spectrum of the **b4** ion at *m/z* 411 (**b4AA/b4LA**) shows dominant ions at *m/z* 383 (**a4AL/a4LA**) corresponding to a loss of CO and a base peak at *m/z* 298 (**b3AA**/[**b4LA**-Leu]$^+$). The [**b4AA**-Ala]$^+$ ion at *m/z* 340.2593 (calcd. for $[C_{18}H_{34}N_3O_3]^+$ *m/z* 340.2595) confirms the sequence HA-L-L-L-A$^+$ (Appendix, Table B.1, Scheme B.2). The sequence is also supported by the HRMS[3] data of the ion at *m/z* 369 (**b4LL/b4LL*/b4AL**) displaying besides a loss of CO (*m/z* 341) both a loss of alanine (*m/z* 298) and leucine (*m/z* 256). Accordingly, (+)-ESI-HRMS measurements yield the evidence for the occurrence of two alanine residues as well as three leucine residues within the primary sequence of compound **3.2**. Therefore, the sequence of the cyclic pentapeptide **3.2** was determined to cyclo-(Ala^1-Leu^2-Leu^3-Leu^4-Ala^5) and named microsporide B.

For the isolation of the target compounds **3.1**-**3.3**, *S. microspermum* was grown on semi-solid media for 21 days. Chromatographic separation of the culture broth and mycelial crude extract using Diaion HP 20 in combination with preparative RP-C18 HPLC resulted in the isolation of compounds **3.1** as isomeric mixture (Appendix, Figure B.3-B.5) and compound **3.3** (Appendix, Figure B.6-B.13). Compound **3.2,** recognized by a UHPLC-(+)-ESI-HRMS screening of the enriched crude extract, could not be obtained through isolation in a quantity sufficient for a detailed spectroscopic analysis.

3.3.2. Total synthesis and absolute configuration of cyclic pentapeptides 3.1-3.3

The total synthesis of the natural products was performed to support the sequence of cyclic peptides **3.1-3.3** and to establish the absolute configurations of the cyclic pentapeptides **3.1** and **3.2.** To develop the synthesis protocol and for comparison purposes, the cyclic peptide **3.3** was synthesized from a linear peptide bearing the following primary sequence: D-Ala-L-Leu-L-Leu-D-Leu-L-Val. The synthetic strategy relied on a solid-phase peptide synthesis (SPPS) approach to afford the linear precursors (**3.4**-**3.6**), followed by a head-to-tail cyclization under high dilution conditions after cleavage from the resin (Scheme 3.1).

The linear peptides (**3.4**-**3.6**, spectral data see Appendix, Figure B.48-B.56) were synthesized using a standard Fmoc (fluorenylmethoxycarbonyl) protocol with a hydroxymethylphenoxy (Wang-type, HMP) resin and a combination of hydroxybenzotriazole (HOBt) and diisopropylcarbodiimide (DIC) in dimethylformamide (DMF) as coupling reagents. The completion of each coupling was checked by a Kaiser test. Owing to the low polarity of the linear unprotected peptide precursors, precipitation from cold ethyl ether was not possible. Instead, the mixture, previously dried under reduced pressure, was dissolved in 10% aqueous acetic acid and washed with $CHCl_3$ in order to extract the remains of protecting groups and non-volatile by-products.
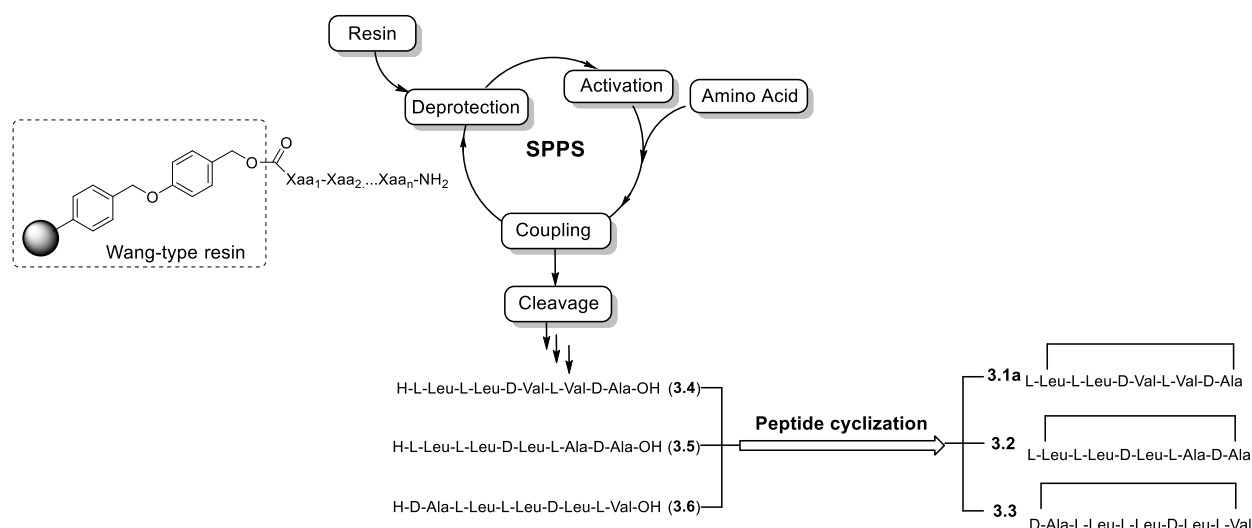
Scheme 3.1 Solid-phase peptide synthesis (SPPS) of linear precursors **3.4**-**3.6** followed by head-to-tail cyclization towards synthetic compounds **3.1a**-**3.3**. Reagents and conditions: *resin*: **3.4**-**3.5** non-preloaded HMP Wang-type resin (L-Ala was manually preloaded as first amino acid, see 3.5.9 in experimental section), **3.6**: L-Val-preloaded HMP Wang-type resin; *activation*: 4 eq. Fmoc-aa$^{1-4}$, 4 eq. HOBt, 4 eq. DIC in DMF, 3 mL, 4 min; *coupling*: after addition activated Fmoc-aa$^{1-4}$ to resin, 1 h; *deprotection*: piperidine (20% in DMF), 6 mL, 2 x 10 min; *cleavage*: TFA/H$_2$O/TIS (95:2.5:2.5, *v/v/v*), 5 mL, 2 h; *peptide cyclization*: 1 eq. linear peptide **3.4**-**3.6**, 4 eq. PyBOP, 8 eq. DIPEA in DMF (details see 3.5.10 in experimental section).

Cyclization was performed under *pseudo*-high dilution conditions (0.001 M) to avoid dimerization. Benzotriazol-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) was used for activating the *C*-terminus and *N,N*-diisopropylethylamine (DIPEA) as base. The reaction was monitored by HPLC and completion was observed after three days. Careful analysis revealed, however, that cyclopeptide **3.3** was formed as a mixture of two isomers in a 4:1 ratio, containing the natural occuring chrysosporide (**3.3**) as the major isomer (Appendix, Figure B.31). The minor diastereomer (epi-**3.3**) is the result of a racemization at the *C*-terminus, mainly owed to the presence of a hindered amino acid (L-Val) at this position of activation for cyclization (see Appendix, Figures B.40-B.47). For further improvements of the synthesis of chrysosporide (**3.3**) (not discussed here), the linear peptide sequence was changed to D-Leu-L-Val-D-Ala-L-Leu-L-Leu which allowed obtaining the cyclic peptide **3.3** with a diastereomeric ratio of 7:1. The observed isomerization occurred alpha to the activated coupling position and not within the primary sequence, a known problem in peptide coupling.

The synthesis protocols and the peptide designs for **3.1a** and **3.2** were established based on the information offered by the primary sequence and stereochemistry of peptide **3.3**. Thus, for cyclic peptides **3.1a** and **3.2** Ala^1 and Val^4/Leu^4 were set as D-configured amino acids (Figure 3.1). In order to favor the cyclization process, linear peptides were designed bearing opposite stereochemistry of the amino acids at either terminus. Therefore, for the synthesis of cyclic peptides **3.1a** and **3.2**, the sequences L-Leu-L-Leu-D-Val-L-Val-D-Ala and L-Leu-L-Leu-D-Leu-L-Ala-D-Ala, respectively, were selected as the linear peptide precursors. The synthesis of these two cyclic peptides took place with similar reaction conditions as those employed for cyclic peptide

3.3, which allowed obtaining synthetic compounds **3.1a** and **3.2** in diastereomeric ratios of 5:1 and 5.6:1, respectively (see Appendix, Figure B.14 and B.22).

Table 3.1 ^{1}H and ^{13}C NMR data of synthetic cyclic pentapeptides (**3.1a**-**3.3**) (600/150 MHz, DMSO-*d6*, 40 °C, δ in ppm).

Pos.	3.1a		3.2		3.3	
	δ_H [ppm], mult. *J* [Hz]	δ_C [ppm]	δ_H [ppm], mult. *J* [Hz]	δ_C [ppm]	δ_H [ppm], mult. *J* [Hz]	δ_C [ppm]
	D-Ala^1		**D-Ala^1**		**D-Ala^1**	
NH	8.197, d, 6.7		8.057, d, 6.6		8.386, d, 6.4	
C=O		172.2		172.0		172.2
α	4.26^b	48.4	4.23^b	48.3	4.212, p, 6.8/6.8/6.9/6.9	48.5
β	1.142, d, 6.9	16.9	1.149, dd, 6.9/2.1	16.8	1.144, d, 6.9	16.3
	L-Leu^2		**L-Leu^2**		**L-Leu^2**	
NH	8.478, d, 7.4		8.344, d, 7.3		8.552, d, 7.6	
C=O		171.4		171.5		171.5
α	3.963, ddd, 10.4/7.4/4.8	53.0	3.976, ddd, 10.1/7.3/5.2	52.9	3.982, ddd, 10.6/7.6/4.6	52.7
β	1.53 m, 1.45 m	39.8	1.50 m, 1.46 m	39.7	1.54^a, 1.48^a m	39.8
γ	1.634, dt, 15.4/13.1/6.6	24.4	1.632, dh, 8.5/6.6	24.1	1.637 m	24.2
δ	0.88^b	22.6	0.89^b	22.6	0.89^b	22.6
δ´	0.81^b	21.2	0.82^b	21.2	0.82^b	20.9
	L-Leu^3		**L-Leu^3**		**L-Leu^3**	
NH	7.023, d, 7.3		7.126, d, 7.6		7.141, d, 7.2	
C=O		171.0		171.0		170.8
α	4.28^b	51.4	4.28^b	50.8	4.271, ddd, 9.0/7.2/5.4	51.3
β	1.53 m, 1.33 m	40.7	1.55 m, 1.33 m	40.5	1.58^a m, 1.35^a m	40.4
γ	1.38	24.6	1.41 m	24.3	1.40^a m	24.6
δ	0.88^b	21.8	0.87^b	21.9	0.85^b	22.7
δ´	0.84^b	22.8	0.84^b	21.4	0.83^b	21.7
	D-Val^4		**D-Leu^4**		**D-Leu^4**	
NH	8.664, d, 8.4		8.451, d, 7.9		8.658, d, 7.9	
C=O		171.0		171.4		171.25
α	3.886, dd, 10.4/8.4	58.8	4.26^b	50.3	4.324 m	50.1
β	1.992, dt, 6.7/3.3	26.6	1.49 m, 1.45 m	37.4	1.49^a	36.8
γ	0.86^b	19.3	1.54 m	23.8	1.56^a m	23.9
γ´	0.83^b	18.7	-	-	-	-
δ	-	-	0.88^b	21.9	0.89^b	21.7
δ´	-	-	0.81^b	21.2	0.81^b	21.4
	L-Val^5		**L-Ala^5**		**L-Val^5**	
NH	7.644, d, 9.2		7.925, d, 8.6		7.647, d, 9.1	
C=O		170.8		171.7		171.1
α	4.091, dd, 9.2/7.9	57.3	4.380, dq, 8.6/7.0	46.8	4.121, dd, 9.1/7.1	57.0
β	1.899 m	29.3	1.149, dq, 6.9/2.1	16.8	1.864 m	30.0
γ	0.85^b	19.2	-	-	0.87^b	19.1
γ´	0.83^b	18.0	-	-	0.84^b	17.8

^{a} chemical shifts are from HSQC and HMBC experiments, ^{b} overlapping signals, chemical shifts are from TOCSY experiments

To verify the proposed structures, the UHPLC-(+)-ESI-HRMS chromatograms, the ^{1}H-NMR as well as CD spectra of the natural and synthetic products **3.1a**-**3.3** (Figure 3.7) were compared. Furthermore, detailed 2D-NMR experiments of the synthetic peptides **3.1a**-**3.3** were performed (Table 3.1, Appendix, Figures B.17-B.21, B.25-B.30, B.34-B.38) for the confirmation of the absolute configuration of the synthetic cyclic pentapeptides **3.1a**-**3.3**. Both, the natural and synthetic chrysosporide **3.3** (cyclo-(D-Ala-L-Leu-L-Leu-D-Leu-L-Val) possess identical NMR signals, CD values as well as the same HPLC-retention time and fragmentation pattern (see Figure 3.7, Appendix, Figures B.58-B.59). Additionally, the data are in agreement with those reported in literature.[10] During cyclization, an isomer of chrysosporide (epi-**3.3**), whose primary sequence and absolute configuration was determined as cyclo-(D-Ala-L-Leu-L-Leu-D-Leu-D-Val) was

formed. The epimerization of L-Valine to D-Valine during cyclization causes changes of all chemical shifts within the ^{1}H-NMR (Appendix, Figure B.41) and a complete change of the CD spectrum (Figure 3.7 A). Furthermore, the retention time of the isomer is different during UHPLC-ESI-HRMS analysis and shows differences in the relative abundances of the fragment ions obtained from the HRMS[n] investigations. This phenomenon is described for the analysis of linear peptides, where ESI-MS/MS measurements have shown the capacity to discriminate between peptides that differ in the configuration at their α-carbons.[35] All detected **b4** ions of the synthetic isomer, which possess a higher number of D-configured amino acids in its primary sequence (**b4LL**/**b4LL***/**b4VL**), show higher abundances in comparison to chrysosporide (**3.3**, Appendix, Table B.2). This information was used for the comparison of the data obtained from the synthetic and natural products of the new cyclic peptides **3.1** and **3.2**.

The synthetic cyclic peptide **3.1a**, with the configuration cyclo-(D-Ala-L-Leu-L-Leu-D-Val-L-Val) has identical retention time (t_R 8.12 min) in comparison with one of the isomers occuring in the natural mixture **3.1** (Figure 3.4 A). Based on the Val-epimer of chrysosporide (epi-**3.3**) discrimination (*vide supra*), it can be suggested that the isomer from the natural mixture at t_R 8.12 min has the same configuration of the underlying amino acids than the synthetic cyclic peptide **3.1a**.
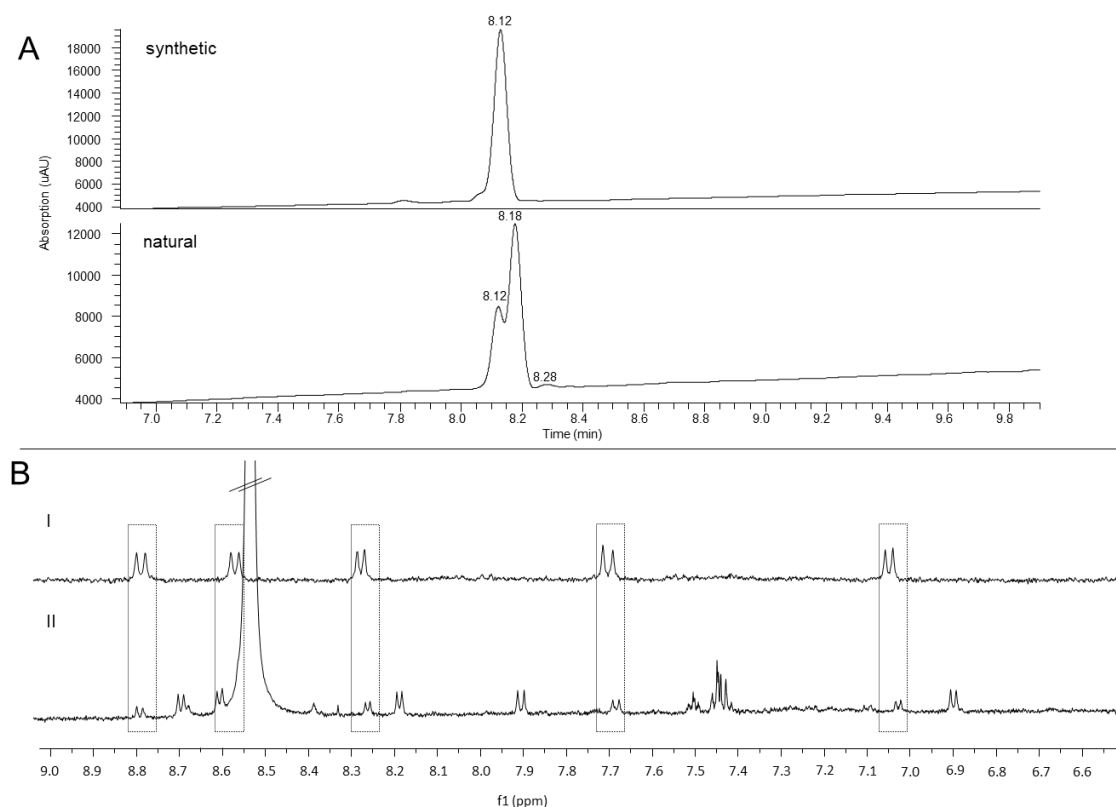
Figure 3.4 A) Comparison of UV spectra ($\lambda = 190\text{-}600$ nm) acquired during UHPLC-(+)-ESI-HRMS; B) Comparison of NH region in the ^{1}H spectrum of I) synthetic cyclic pentapepide **3.1a** and II) isolated isomeric mixture of peptide **3.1** (600 MHz, DMSO-*d6*, 25 °C). Minor differences in the chemical shifts of the two NH protons are due to small differences in pH and/or concentration of the two NMR samples.

This was further supported by comparing the fragment ion abundances of the natural and synthetic cyclic peptide **3.1a** obtained during MS-based sequencing (Appendix, Table B.3). Within the ^{1}H-NMR spectrum of the natural isomeric mixture of **3.1**, the five NH resonances corresponding to the amide protons (Figure 3.4, B/II) exhibit the same chemical shifts like the NH signals of the synthetic peptide **3.1a** (Figure 3.4, B/I). In comparison to natural compound **3.1a**, compounds **3.1b** and **3.1c** exhibit differences in retention times as well as in their fragment abundancies obtained during HRMS[2] experiments (see Appendix, Table B.3).

By analyzing the CD spectra of the natural isomeric mixture of compound **3.1** with that of the synthetic product **3.1a** (Figure 3.7 B), the natural mixture containing **3.1** exhibit additional minima at 213, 219 and 225 nm in addition to the significant maximum at 196 nm and the minimum at 206 nm. This indicates that in addition to isomer **3.1a**, whose CD spectrum coincides with that of the isomer mixture, there are also other isomers in the mixture. These data are consistent with the ^{1}H-NMR and UHPLC data (Figure 3.4). The synthesis of the isomers **3.1b** and **3.1c** could not be achieved. Therefore, the primary sequence of **3.1b** and **3.1c** could be determined as cyclo-(Ala-Leu-Leu-Val-Val) without knowledge of the absolute configuration of the amino acids.

For the synthesized cyclic pentapeptide **3.2** the total ion chromatogram (TIC) as well as the fragmentation pattern were compared to natural **3.2** with t_R = 7.87 min (Appendix, Figure B.57). Both the UHPLC retention time and the ESI-HRMS[2] spectra were in excellent agreement.

Full assignment of the resonance signals for the cyclic peptides was achieved by the analysis of TOCSY, HSQC and HMBC spectra. Figure 3.5 summarizes the TOCSY information as well as key HMBC correlations confirming the connection between the different spin systems. For cyclic peptide **3.3**, the data are in accordance with the previously reported cyclo-(L-Val-D-Ala-L-Leu-L-Leu-D-Leu).[10]
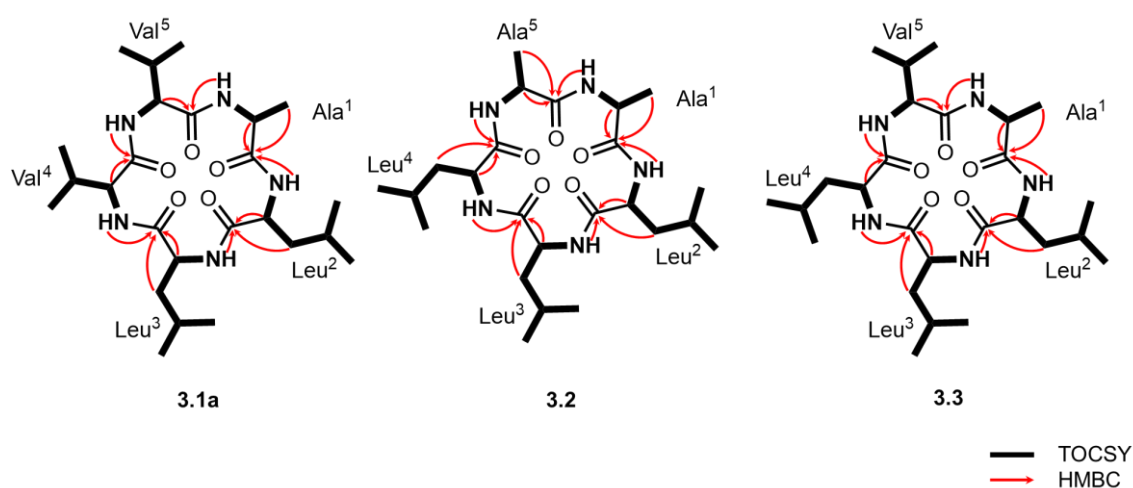
Figure 3.5 Key HMBC (H to C) and TOCSY correlations of synthetic cyclic pentapeptides **3.1a**-**3.3**.

The analysis of the ROESY spectrum and the amide proton coupling constants ($^3J_{HN-H\alpha}$), allows conclusions of structural features of compounds **3.1a-3.3**. The observation of $^3J_{HN-H\alpha}$ higher than 8 Hz is a known characteristic of skeletons with relative fixed conformations around the phi dihedral angles (ϕ), commonly found in native beta sheet structures. For L-amino acids this means $\phi = -120 \pm 30^0$ while for D-amino acids $\phi = +120 \pm 30^0$. Additionally, the ROESY spectra of these molecules were not rich in interresidual crosspeaks, but mostly were confined to sequential correlations (i.e. between amino acids consecutive in the sequence). The most relevant ROEs found are also illustrated in Figure 3.6.
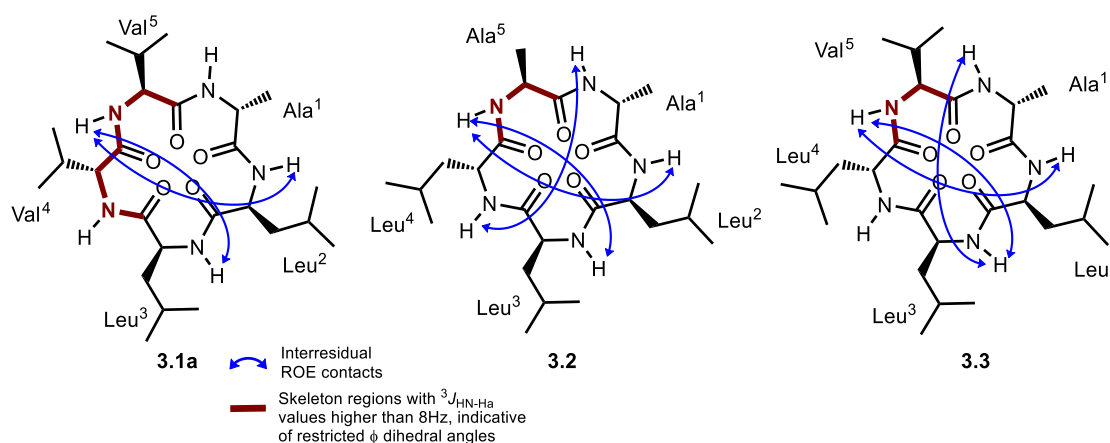
Figure 3.6 Key Structural information deduced from NMR data for synthetic cyclic pentapeptides **3.1a**-**3.3**.

When compared, it seems that cyclic peptides **3.2** and **3.3** share structural similarities, with a fixed phi dihedral angle at residue 5. Apparently, a bulkier side chain informs of a Val residue at position 5 in peptide **3.3** forces the backbone into a slightly different arrangement if compared with compound **3.2**. This latter is evidenced by the different ROE contacts of the amide proton of Ala[1], indicating spatial proximity with the amide proton of Leu[4] for peptide **3.2** or Leu[3] for peptide **3.3**, respectively. In contrast, in cyclic peptide **3.1a** the $^3J_{HN-H\alpha}$ values are higher than 8 Hz for residues Val[4] and Val[5], indicative of a different conformation governed by fixed phi angles at these residues. Interestingly, in the three compounds the amide protons from residue 5 show a correlation with those of residues 2 and 3, suggesting that these atoms point to the inside of the ring.

From the CD spectra of the synthesized compounds it can also be noticed that cyclic peptides **3.2** and **3.3** share some structural similarities. Even when there are some resemblances with the CD spectra for random coil peptides, the deep minimum around 204 nm (rather than 195 nm in random coil) and the positive absorption band around 190 nm are evidences of partial 3_{10}-helicity, which could be correlated to the existence of β-turn conformations. This latter could be expected from a short cyclic peptide sequence such as the one studied. In contrast, the synthetic cyclic peptide **3.1a** shows a clear maximum at 196 nm while keeping a minimum around 206 nm (Figure 3.7).
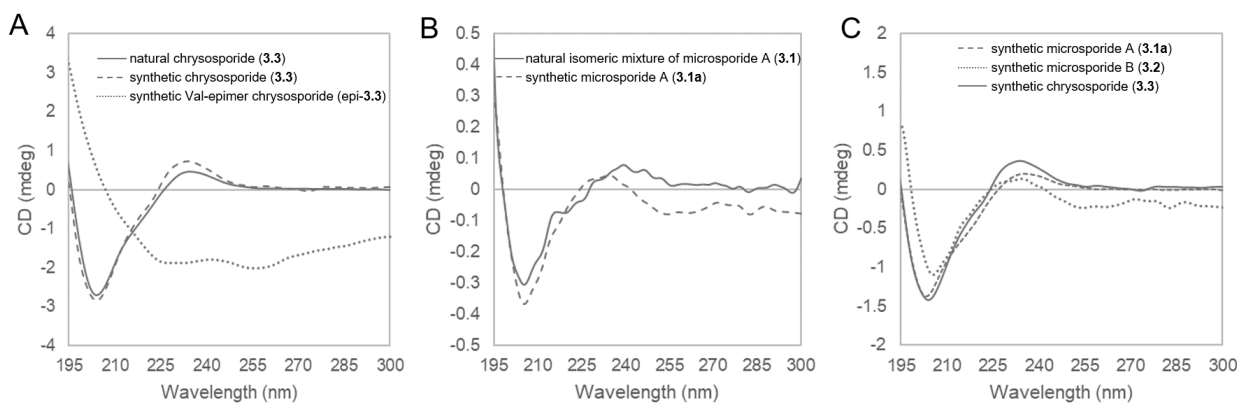
Figure 3.7 Comparison of circular dichroism spectra of A) natural and synthetic chrysosporide (**3.3**, epi-**3.3**); B) natural and synthetic microsporide A (**3.1a**) and, C) the synthetic cyclic pentapeptides microsporide A (**3.1a**), B (**3.2**) and chrysosporide (**3.3**). (Intensities of the spectra were normalized due to concentration differences. Original data see Appendix, Figure B.7, B.16, B.24, B.33, B.40)

3.3.3. Antibacterial and anthelmintic activity

Chrysosporide (**3.3**) was evaluated for its antibacterial and anthelmintic activities. For the antibacterial activity screening the gram-negative strain *Aliivibrio fischeri* (Appendix Figure B.61)*,* and standard pathogenic control strains, namely *Staphylococcus aureus* (ATCC 6538P), *Staphylococcus aureus* MRSA (ATCC 43300), *Enterococcus faecalis* (ATCC 29212), *Enterococcus* spp. (ATCC 25922), *Acinetobacter baumannii* (ATCC 19606), and *Pseudomonas aeruginosa* (ATCC 27853), were selected for testing (Appendix, Table B.4). Chrysosporide (**3.3**) did not exhibit inhibitory activities against all these mentioned strains. Furthermore, compound **3.3** was tested concerning its anthelmintic properties against *Caenorhabditis elegans* (Appendix, Figure B.60), but it did not show any activity. These results are in agreement with observations of Li et *al.*[20]. In that case structurally similar cyclic pentapeptides (cyclo-(L-Leu-L-Leu-D-Leu-L-Leu-L-Ile/L-Val/L-Leu) also did not indicate antibacterial activity and cyclo-(L-Leu-L-Leu-L-Leu-L-Leu-L-Ile) exhibited only moderate inhibitory effects against the three plant pathogenic fungi *Aphanomyces cochlioides*, *Pythium ultimum*, and *Rhizoctonia solani*.[19] Furthermore, their possible biological functoin as signal molecule could not be demonstrated.[20]

3.4 Conclusions

The two new cyclic pentapeptides **3.1a** (cyclo-(D-Ala-L-Leu-L-Leu-D-Val-L-Val)) and **3.2** (cyclo-(L-Leu-L-Leu-D-Leu-L-Ala-D-Ala)) as well as chrysosporide (**3.3**) (cyclo-(D-Ala-L-Leu-L-Leu-D-Leu-L-Val)) were detected and structurally elucidated from the enriched crude extract of *Sepedonium microspermum* by intensive (+)-ESI-HRMS^n investigations. This is the first detection of cyclic peptides as metabolites from the mycoparasitic fungus *S. microspermum*. MS-guided isolation yielded compound **3.1** as isomeric mixture (**3.1a-c**) and pure chrysosporide (**3.3**). It is known from peptaibols as characteristic constituents of *Sepedonium* spp. that they usually occur as complex mixtures of homologues Furthermore, one major compound often dominates compared

with other minor components.[14] The structures of the herewith described cyclic pentapeptides **3.1-3.3** represent minor constituents of yet unknown biological function. Their structures were elucidated by extensive (+)-ESI-HRMS^n experiments as a powerful method for peptide sequencing. Furthermore, the relative and absolute configuration and the structural confirmation of the compounds **3.1a**-**3.3** could be elucidated by a direct comparison with reference compounds obtained from total synthesis. The use of a combination of advanced analytical techniques, isolation and synthesis efforts again was confirmed as the most powerful approach for unequivocal natural product identification.

3.5 Experimental Section

3.5.1. General Procedures

Adsorption chromatography was performed using a batch approach with Diaion HP 20 purchased from Supelco (USA). CD spectra were obtained from a Jasco J-815 CD spectrometer.

3.5.2. NMR

NMR spectra of the cyclic peptides were recorded with an Agilent VNMRS 600 NMR spectrometer operating at 40 °C. The compounds were dissolved in DMSO-*d6* (99.96% D) and spectra were recorded at 599.83 MHz (^{1}H) and 150.84 MHz (^{13}C) using a 5-mm inverse detection cryoprobe. 1D (^{1}H) and 2D NMR spectra (^{1}H,^{13}C gHSQCAD, ^{1}H,^{13}C gHMBCAD, ^{1}H,^{1}H gDQCOSY, ^{1}H,^{1}H zTOCSY, ^{1}H,^{1}H ROESY) were recorded using standard CHEMPACK 8.1 pulse sequences (s2pul) implemented in Varian VNMRJ 4.2 spectrometer software. The mixing time for the TOCSY experiments was set to 80 msec, for the ROESY experiments was set to 300 msec. The HSQC experiment was optimized for ^{1}J$_{CH}$ = 146 Hz with DEPT-like editing and ^{13}C-decoupling during acquisition time. The HMBC experiment was optimized for a long-range coupling of 8 Hz and the ^{13}C band selective HMBC of 4 Hz. ^{1}H chemical shifts are referenced to internal TMS (δ_H = 0 ppm) and ^{13}C chemical shifts to internal DMSO-*d6* (δ_C = 39.5 ppm). The NMR spectra of the linear peptides were recorded on a Varian DD2 spectrometer at 399.82 MHz (^{1}H) and 100.54 MHz (^{13}C), respectively.

3.5.3. UHPLC-ESI-HRMS

The positive ion high-resolution ESI mass spectra and collision-induced dissociation (CID) MS^n spectra were obtained from an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Germany) equipped with a HESI electrospray ion source (spray voltage 4.0 kV/4.5 kV, capillary temperature 275 °C/325 °C, source heater temperature 250 °C/300 °C, FTMS resolution 15.000). Nitrogen was used as sheath and auxiliary gas. The MS system was coupled to an ultra-high-performance liquid chromatography (UHPLC) system (Dionex UltiMate 3000, Thermo Fisher Scientific) equipped with an RP18 column (particle size 1.9 µm, pore size 175 Å, 50 x 2.1 mm internal diameter, Hypersil GOLD, Thermo Fisher Scientific, column temperature 30°C or 45 °C) and a photodiode array detector (190-400 nm, Thermo Fisher Scientific). The mobile phases consisted of H_2O (A, Merck Millipore Milli-Q equipment) and CH_3CN (B, Fluka Analytical, LC-MS Chromasolv) with acetic acid (0.2%, Fluka, solvent system I) or formic acid (0.1%, Fluka, solvent system II). Chromatographic separation during the UHPLC-(+)-ESI-HRMS based screening and for the analysis of the linear peptide precursors I/II was established using a gradient system with solvent system I starting with 5 % B increasing within 15 min to 100 %, holding isocratic for further 10 min (flow rate 0.15 mL/min). For the analysis of the pure natural and synthetic cyclic compounds the following gradient system (solvent system II) was used: 0-1 min, 5 % B, 1-11 min 5-98 % B, 11-14 min 98 % B, flow rate 0.4 ml/min. CID mass spectra (buffer gas: helium) were recorded using normalized collision energies (NCE) of 35-45 % (see Appendix).

The instrument was externally calibrated using Pierce ESI positive-ion calibration solution (product No. 88323) from Thermo Fisher Scientific. The data were evaluated using the software Xcalibur 2.2 SP1.

3.5.4. General SPPS

The synthesis of the linear peptides was performed in a reaction vessel equipped with a sintered glass bottom using the general Fmoc/*t*-Bu strategy for solid-phase peptide synthesis (SPPS). The L-configured Fmoc-Val-OH and Fmoc-Leu-OH were supplied from Novabiochem (Germany) and the D-configured Fmoc-Ala-OH and Fmoc-Leu-OH from Carbolution Chemicals GmbH (Germany). The Fmoc-L-Val preloaded as well as the non-preloaded Wang-type resin were purchased from Iris Biotech GmbH (Germany). Piperidine, pyridine, diisopropylcarbodiimide (DIC), dimethylaminopyridine (DMAP), triisopropyl silane (TIS), hydroxybenzotriazole (HOBt), acetic anhydride (Ac$_2$O), *N*,*N*-Diisopropylethylamine (DIPEA), trifluoroacetic acid (TFA), benzotriaz-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP), and dimethylformamide (DMF) were purchased from Sigma-Aldrich (Germany). Dichlormethane (DCM) was used in distilled quality.

3.5.5. Preparative HPLC

Preparative HPLC was carried out with a Shimadzu prominence system equipped with a CBM-20A communications bus module, a SPD-M20A diode array detector, a DGU-20A5R degassing unit, a LC-20AT liquid chromatograph, and a SIL-20A HT auto sampler using either column 1 (YMC Pack Pro C18 column (5 µm, 120 Å, 150 x 10 mm internal diameter, YMC, USA) or column 2 (LiChrospher C18, 10 µm, 100 Å, 250 x 10 mm i. d., Merck, Germany). The system was controlled using the LabSolutions software (version 5.57).

3.5.6. Semipreparative HPLC-MS

Semipreparative RP-18 HPLC-MS was performed using a custom set of Agilent 1200 Infinity I series consisting of a preparative binary pump (G1361A), an analytical quaternary pump (G1311B) used as make-up pump, an auto sampler (G2260A), a column select valve (G1159A), a fraction collector (G1364B), a multiple wavelength detector (G1365D), a splitter (G1968D) and single quadrupole MS detector with an electrospray ionization source (ESI, 6120). Electrospray mass spectrometry measurements were performed in positive ionization mode (fragmentor, 5 V; threshold spectral abundance, 150; capillary voltage, 3000 V) using a mass range of 100–600 Da. Automated fractionation was carried out using SIM mode. Nitrogen was used as the nebulizer gas (nebulizer pressure, 35 psig). The drying gas flow was set to 12.0 L/min and the drying gas temperature to 250°C. For chromatographic separation an ODS-A C18 ec column (5 µm, 120 Å, 150 x 20 mm internal diameter, YMC, USA) was used. The whole system was controlled by the Agilent Chemstation (Rev. 8.04.03 SP1).

3.5.7. Fungal strain and cultivation

Sepedonium microspermum Besl (strain KSH 584) was isolated in September 2001 from *Xerocomus chrysenteron* (Bull.) Quél. (leg./det. I. Wagner) in Italy (Monte Caloria Fagnano Castello (Cosenza) 1100 m). *S. microspermum* (strain KSH 584) is preserved at the Leibniz Institute of Plant Biochemistry, Halle (Saale). The culture of *S. microspermum* was stored on malt peptone agar (MPA) plates (10 g of malt, 2.5 g peptone and 15 g of agar in 1000 mL deionized water) and transferred periodically.

For the UHPLC-ESI-HRMS screening of the cyclic peptides, fungal cultures were grown for 14 days on malt-peptone agar in triplicate at room temperature and then stored at -65 °C until extraction.

The up-scaled semisolid cultures, used for isolation of cyclic peptides, were grown in 20 Erlenmeyer flasks (size 1 L), each containing 1.5 g of cotton wool and 250 mL of malt peptone medium (2.5 g malt and 0.625 g peptone in 250 mL deionized water), resulting in a total volume of 2.5 L. Each culture flask was inoculated with a 10 x 10 mm agar plug of colonized fungus and incubated for 21 days at room temperature without agitation.

3.5.8. Extraction, sample preparation and isolation

For the preparation of the enriched extracts for the UHPLC-ESI-HRMS-based screening, three stored deep frozen agar plate cultures of *S. microspermum* were crushed in small pieces and extracted with DCM:MeOH (1:1, *v/v*, 3 x 500 mL) in an ultrasonic bath at room temperature. The combined extracts were filtered, and evaporated to dryness *in vacuo*. The dried crude extract was redissolved in MeOH/H_2O (1:2, *v/v*) to a concentration of 50 mg/mL. The resulting solution was separated on SPE cartridges Chromabond® C 18 (loading 200 mg/ 3 mL, particle size 45 µm, Macherey-Nagel), targeted peptides **3.1-3.3** were eluted with MeOH (100%). After evaporation to dryness *in vacuo*, the enriched fraction was redissolved in MeOH to a concentration of 1 mg/mL. Aliquots of 2 µL were used for further analysis by UHPLC-(+)-ESI-HRMS.

For isolation of the natural cyclic pentapeptides, the mycelia were separated from the culture broth by vacuum filtration, frozen with liquid nitrogen, and extracted with MeOH (3 × 1.5 L). The yellow solution was evaporated to dryness, redissolved in H_2O, and combined with the culture broth. Activated Diaion HP 20 (20 g) was added and agitated for 12 hours at room temperature. Diaion HP 20 was removed by vacuum filtration, washed with H_2O, eluted with MeOH, and the extract was evaporated *in vacuo* to dryness. The resulting methanolic crude extract (3.25 g) was resuspended in a MeOH/H_2O mixture (1.4:1, *v/v*). The insoluble part (285.2 mg), which contains the targeted cyclic peptides, was split by centrifugation and further separated using preparative HPLC using column 1 with H_2O (A) and CH_3CN (B), both containing 0.1 % FA, as eluents (0-10 min, 30-80% B, 10-20 min, 80-90% B, 20-21 min, 90-100% B, flow rate 4.5 mL/min), obtaining chrysosporide **3.3** (Rt = 12.0 min, 5.31 mg) and an isomeric mixture of **3.1** (Rt= 11.0

min, 1.7 mg). Compound **3.2** was only detected by UHPLC-(+)-ESI-HRMS in traces, making isolation impossible.

Natural cyclic pentapeptide 3.3: White amorphous solid. UV (PDA signal LC-HRMS) λ_{max} = 222 nm. (+)-ESI-HRMS: calcd. for [M+H]$^+$ 510.3650 , found 510.3642. NMR: see Appendix, Figure B.8-B.13, ESI-HRMS^n: see Appendix, Table B.1.

3.5.9. Solid-phase peptide synthesis

General procedure. The synthesis of the linear intermediate peptides was carried out on a Wang-type resin. In some indicated instances, it was required to load the the first amino acid onto the resin using 10 equivalents (eq.) relative to the resin capacity of the corresponding Fmoc derivatives. The Fmoc amino acid in DCM (3 mL/mmol) was placed in a round bottom flask with a magnetic stirrer with the aid of some drops of DMF at 0 °C. At this point 5 eq. of DIC were added and the mixture was stirred for 10 min at 0 °C. The solution was added to the resin (previously swollen) and 1 eq. DMAP in DMF was also added. After 1 h stirring, the resin was washed three times with DCM (3x3 mL) and three times with DMF (3x3 mL). Acetic anhydride (5 eq.) and pyridine (1 eq.) in DMF were added to the resin and the mixture was stirred for 30 min in order to cap the non-reactive groups on the resin. The pre-loaded resin was swollen in DCM (3 mL) for 20 min. Afterwards, it was treated twice with piperidine (20% in DMF, 2x3 mL) for 10 min each in order to deprotect the first amino acid. For the sequential peptide coupling, 4 eq. of the corresponding amino acid were activated using HOBt (4 eq.) and DIC (4 eq.) in 1 mL DMF for 4 min, and then this mixture was added to the resin. The completion of each coupling was checked after 20 min with the Kaiser test. Peptides were cleavage from the resin using the mixture TFA/H$_2$O/TIS (95:2.5:2.5, *v/v/v*; 5 mL) for 2 h. Then, the mixture was poured onto cold ether and the volatiles were removed under reduced pressure. After re-dissolving the crude in 10% acetic acid, the mixture was washed three times with chloroform. The aqueous phase was dried under reduced pressure, the residue re-dissolved in glacial acetic acid and lyophilized.

Linear peptide 3.4: Synthesis was carried out in 0.2 mmol scale using a non-preloaded Wang-resin (loading 0.6-1.2 mmol/g). Loading of the first amino acid (Fmoc-D-Ala-OH) on the resin, as well as, the sequential peptide coupling was performed as described in the general procedure to yield 195 mg of peptide **3.4** (85%, 76% HPLC purity). (+)-ESI-MS: [M+H]$^+$ *m/z* 514, [M+Na]$^+$ *m/z* 536; NMR Data: ^{1}H NMR (DMSO-*d6*, 400 MHz) δ 8.55 (1H, d, *J* = 8.4 Hz, N*H*), 8.17 (1H, d, *J* = 7.3 Hz, N*H*), 8.06 (1H, d, *J* = 8.7 Hz, N*H*), 7.92 (1H, d, *J* = 9.2 Hz, N*H*), 4.56–4.50 (1H, m), 4.31 (1H, dd, *J* = 8.5, 6.3), 4.27–4.17 (m, 3H), 2.04–1.93 (2H, m), 1.87–1.80(1H, m) 1.69–1.43 (6H, m), 1.25 (3H, d, *J* = 7.2 Hz), 0.94–0.77 (24H, m); ^{13}C NMR (DMSO-*d6*, 100 MHz) δ 173.8, 171.6, 170.8, 170.5; 168.7 (C, *C*=O), 57.7, 57.1, 51.4, 50.7, 47.4 (*C*H), 41.5, 40.3 (*C*H$_2$), 30.7, 24.1, 23.4 (*C*H), 23.0, 22.8, 21.7, 21.6, 19.3, 19.2, 17.8, 17.5 (*C*H$_3$).

Linear peptide 3.5 Synthesis was carried out in 0.2 mmol scale using a non-preloaded Wang-resin (loading 0.6-1.2 mmol/g). Loading of the first amino acid (Fmoc-D-Ala-OH) on the resin, as well

as, the sequential peptide coupling was performed as described in the general procedure to yield 86 mg of peptide **3.5** (77%, 95% HPLC purity). (+)-ESI-MS [M+H]$^+$ *m/z* 500, [M+Na]$^+$ *m/z* 522; NMR: ^{1}H NMR (DMSO-*d6*, 400 MHz) δ 8.37 (1H, d, *J* = 7.8 Hz, N*H*), 8.07 (1H, d, *J* = 7.8 Hz, N*H*), 7.89 (1H, d, *J* = 7.5 Hz, N*H*), 7.83 (1H, d, *J* = 7.8 Hz, N*H*), 4.21 (1H, dd, *J* = 14.8, 7.7 Hz), 4.16–4.05 (2H, m), 3.63 (1H, s), 1.49–1.20 (8H, m); 1.08 (3H, d, *J* = 7.2 Hz), 1.01 (3H, d, *J* = 7.0 Hz), 0.86 (6H, d, *J* = 6.2 Hz), 0.76–0.61 (12H, m); ^{13}C NMR (DMSO-d6, 100 MHz) δ 173.9, 171.7, 171.5, 171.4, 168.8 (C, *C*=O), 51.6, 51.2, 50.7, 47.8, 47.4 (CH), 41.2, 40.6 (CH$_2$), 24.2, 24.1, 23.4 (CH), 23.1, 22.9, 22.8, 21.8, 21.6, 21.2, 18.7, 17.4 (CH$_3$).

*Linear peptide **3.6***: Synthesis was carried out in 0.2 mmol scale using commercial, preloaded Fmoc-L-Val-Wang resin (loading 0.53 mmol/g). Peptide coupling was performed as described in the general procedure to yield 80 mg of peptide **3.6** (68%, 97% HPLC purity). (+)-ESI-MS [M+H]$^+$ *m/z* 528, [M+Na]$^+$ *m/z* 550; NMR Data: ^{1}H NMR (DMSO-*d6*, 400 MHz) δ 8.54 (1H, d, *J* = 8.3 Hz, N*H*), 8.11 (1H, d, *J* = 4.5 Hz, N*H*), 7.99 (1H, d, *J* = 8.5 Hz, N*H*), 7.92 (1H, d, *J* = 6.3 Hz, N*H*), 4.47–4.34 (2H, m), 4.28 (1H, dd, *J* = 14.6, 7.3 Hz), 4.12 (1H, dd, *J* = 8.7, 5.8 Hz), 3.87 (1H, dd, *J* = 13.9, 6.9 Hz), 2.09–1.99 (1H, m), 1.61–1.38 (9H, m), 1.33 (3H, d, *J* = 7.3 Hz), 0.93–0.95 (24H, m); ^{13}C NMR (DMSO-*d6*, 100 MHz) δ 173.0 (*C*=O), 172.1 (*C*=O), 171.6 (*C*=O), 171.5 (*C*=O), 57.1 (*C*H), 51.5 (*C*H), 51.0 (*C*H), 50.8 (*C*H), 48.3 (*C*H), 41.2(*C*H), 40.8 (*C*H), 30.2, 24.3, 24.2, 23.2, 22.9, 21.8, 21.3, 21.1, 19.2, 18.0, 17.6.

3.5.10. In-solution peptide cyclization

General procedure. The linear peptide dissolved in 10 mL DMF was added dropwise (0.5 mL/h) into a three necked flask containing a mixture of PyBOP (4 eq.) and DIPEA (8 eq.) dissolved in 90 mL DMF. The invase addition reaction was run for 3 days under nitrogen atmosphere and high dilution conditions (0.001 M). After monitoring the reaction completion by ESI-MS and HPLC, the reaction mixture was concentrated under reduced pressure.

*Synthetic cyclic pentapeptide **3.1a** (microsporide A)*: Peptide **3.4** (100 mg, 0.174 mmol) dissolved in 10 mL DMF and a mixture of PyBOP (406 mg in 180 mL DMF) and 270 µL of DIPEA was added. The cyclization was performed as described in the general procedure. Precipitation of the crude material with methanol at room temperature yielded 12.3 mg of cyclic pentapeptide **3.1a** (purified yield 14.3%) as a white amorphous solid. Analytical UHPLC (λ= 205-210 nm): t$_R$ = 8.14 min, 91% purity. UV (PDA signal LC-HRMS) λ_{max} = 222 nm; (+)-ESI-HRMS: calcd. for [M+H]$^+$ 496.3493, found 496.3491; NMR Data: see Table 3.1.

*Synthetic cyclic pentapeptide **3.2** (mycrosporide B)*: Peptide **3.5** (60 mg, 0,107 mmol) dissolved in 10 mL DMF and a mixture of PyBOP (249.8 mg in 120 mL DMF) and 166 µL of DIPEA was added The cyclization was performed as described in the general procedure. Purification of the crude material using semi-preparative HPLC and column 2 with H$_2$O + 0.1% FA (A) and CH$_3$CN + 0.1% FA (B) as eluents (0-5 min, 5-40% B; 5-10 min, 40-60%; 10-12 min, 60-100%; flow rate 5 mL/min) yielded 1.3 mg of cyclic pentapeptide **3.2** (purified yield 2.5%) as a white amorphous

solid. Analytical UHPLC (λ = 205-210 nm): t_R 7.95 min, 84% purity. UV (PDA signal LC-HRMS) λ_{max} = 222 nm. (+)-ESI-HRMS: calcd. for $[M+H]^+$ 482.3337, found 482.3334. NMR Data: see Table 3.1.

*Synthetic cyclic pentapeptide **3.3** (chrysosporide)*: Peptide **3.6** (60 mg, 0.102 mmol) dissolved in 10 mL DMF and a mixture of PyBOP (237.3 mg in 90 mL DMF) and 158 µL of DIPEA was added. The cyclization was performed as described in the general procedure. Purification of the crude material using semi-preparative HPLC-MS with H_2O + 0.1% FA (A) and MeOH + 0.1% FA (B) as eluents (0-15 min, 70-90% B; flow rate 9.5 mL/min) yielded 14.5 mg of cyclic pentapeptide **3.3** (isolated yield, t_R 11.6 min,) and 11.3 mg of its synthetic isomer epi-**3.3** (isolated yield 22%, t_R 9.2 min), both as white amorphous solids. Synthetic cyclic pentapeptide **3.3**: Analytical UHPLC (λ = 205-210 nm): Rt 8.66 min, 92% purity. UV (PDA signal LC-HRMS) λ_{max} = 222 nm. ESI-HRMS: calcd. for $[M + H]^+$ 510.3650 , found 510.3651. NMR Data: see Table 3.1. Synthetic cyclic pentapeptide epi-**3.3**: Analytical UHPLC (λ = 205-210 nm): t_R 8.13 min, 90% purity; UV (PDA signal LC-HRMS) λ_{max} = 222 nm. (+)-ESI-HRMS: calcd. for $[M+H]^+$ 510.3650 , found 510.3647. NMR Data: Table 3.1.

3.5.11 Biological Assays

3.5.11.1 Antibacterial Bioassay

Aliivibrio fischeri

The assay was performed according to a procedure described by Stark (phD thesis 2016)[37] using the gram-negative *A. fischeri* test strain DSM507 (batch no. 1209). Briefly, for each test run a fresh glycerol stock was incubated in 25 mL Boss medium at 100 rpm and 23 °C for 16‑18 h and was afterward diluted with fresh BOSS medium to an appropriate cell number (luminescence value between 30.000 and 50.000 RLU). The assay was conducted on black flat bottom 96 well plates (Brand cellGrade™premium, STERILE R) in a final volume of 200 µl of Boss medium containing 1% DMSO in each well (100 µl diluted bacterial solution and 100 µl test solution). The respective extracts were applied in two concentrations (50 µg/ml and 500 µg/ml) in DMSO/medium (2%, *v/v*). The plates were incubated in the dark at 23 °C and 100% humidity without lid and without shaking for 24 h. The bioluminescence (obtained in relative luminescence units, RLU) is dependent on the cell density and was determined after 24 h using the microplate reader TecanGeniosPro. Therefore, the whole wavelength range was detected for 1000 ms without preliminary shaking to avoid secondary oxygen effects. The results (mean value ± standard deviation, n=6) are given as relative values (% inhibition) in comparison to the negative control (bacterial growth, 1% DMSO, without test compound). Negative values indicate an elevation of luminescence/increase of bacterial growth. Chloramphenicol was used as positive control.

Broth microdilution method

In vitro antibacterial activity was investigated determining the minimum inhibitory concentration (MIC) and minimum bactericidal concentration (MBC). A broth microdilution method using sterile 96-well microdilution plate was used to determine the MIC, bottom wells with Muller Hinton broth according CLSI 2016. The methodology also included a positive control for growth (bacterial strains without compound) and negative control (compound without bacterial strains). The microplates were incubated at 37°C for 24 h. The MBC was determined pouring 100 µL of culture from well representing the MIC and at least three of the more concentrated test product dilutions in 5 mL of tryptic soy broth and cultivated at 37°C for 24 h. Evaluated concentration range: 2-128 µg/mL, methanol or DMSO concetration in microplate: 6%. The following bacterial strains were used: *Staphylococcus aureus* (ATCC 6538P), *Staphylococcus aureus* MRSA (ATCC 43300), *Enterococcus faecalis* (ATCC 29212), *Enterococcus* spp. (ATCC 25922), *Acinetobacter baumannii* (ATCC 19606), and *Pseudomonas aeruginosa* (ATCC 27853).

3.5.11.2 Anthelmintic Bioassay

The Bristol N2 wild type strain of *Caenorhabditis elegans* was used in the anthelmintic assay. The nematodes were cultured on NGM (Nematode Growth Media) petri plates using the uracil auxotroph *E. coli* strain OP50 as food source according to the methods described by Stiernagle.[38] The anthelmintic bioassay was carried out following the method developed by Thomsen et al..[39] In all the assays, the solvent DMSO (2%) and the standard anthelmintic drug ivermectin (10 µg/mL) were used as negative and positive controls, respectively. All the assays were carried out in triplicate.

3.6 References

[1]   Divekar, P.V.; Vining, C. Reaction of anhydrosepedonin with alkali. *Can. J. Chem.* **1964**, *42*, 63-68.

[2]   Divekar, P.V.; Raistrick, H. Sepedonin, a tropolone metabolite of *Sepedonium chrysospermum* Fries. *Can. J. Chem.* **1965**, *43*, 1835-1848.

[3]   Shibata, S.; Shoji, J.; Ohta, A.; Watanabe, M. Metabolic products of fungi. XI. Some observation on the occurrence of skyrin and rugulosin in mold metabolites, with a reference to structural relationship between penicilliopsion and skyrin. *Pharm. Bull.* **1957**, *5*, 380-382.

[4]   Quang, D.N.; Schmidt, J.; Porzel, A.; Wessjohann, L.A.; Haid, M.; Arnold, N. Ampullosine, a new isoquinoline alkaloid from *Sepedonium ampullosporum* (Ascomycetes). *Nat. Prod. Commun.* **2010**, *5*, 869-872.

[5]   Closse, A.; Hauser, D. Isolierung und Konstitutionsermitllung von Chrysodin. *Helv. Chim. Acta.* **1973**, *56*, 2694-2698.

[6]   Ritzau, M.; Heinze, S.; Dornberger, K.; Berg, A.; Fleck, W.; Schlegel, B.; Hartl, A.; Gräfe, U. Ampullosporin, a new peptaibol-type antibiotic from *Sepedonium ampullosporum* HKI-0053 with neuroleptic activity in mice. *J. Antibiot. (Tokyo)* **1997**, *50*, 722-728.

[7]   Neuhof, T.; Berg, A.; Besl, H.; Schwecke, T.; Dieckmann, R.; Von Döhren, H. Peptaibol production by *Sepedonium* strains parasiting *Boletales*. *Chem. Biodiv.* **2007**, *4*, 1103-1115.

[8]   Otto, A.; Laub, A.; Wendt, L.; Porzel, A.; Schmidt, J.; Palfner, G.; Becerra, J.; Kruger, D.; Stadler, M.; Wessjohann, L., et al. Chilenopeptins A and B, Peptaibols from the Chilean *Sepedonium* aff. *chalcipori* KSH 883. *J. Nat. Prod.* **2016**, *79*, 929-938.

[9]   Otto, A.; Laub, A.; Haid, M.; Porzel, A.; Schmidt, J.; Wessjohann, L.; Arnold, N. Tulasporins A-d, 19-Residue peptaibols from the mycoparasitic fungus *Sepedonium tulasneanum*. *Nat. Prod. Commun.* **2016**, *11*, 1821-1824.

[10]   Mitova, M.I.; Stuart, B.G.; Cao, G.H.; Blunt, J.W.; Cole, A.L.J.; Munro, M.H.G. Chrysosporide, a cyclic pentapeptide from a New Zealand sample of the fungus *Sepedonium chrysospermum*. *J. Nat. Prod.* **2006**, *69*, 1481-1484.

[11]   Wang, X.; Lin, M.; Xu, D.; Lai, D.; Zhou, L. Structural diversity and biological activities of fungal cyclic peptides, excluding cyclodipeptides. *Molecules* **2017**, *22*, 2069-2116.

[12]   Wessjohann, L.A.; Ruijter, E.; Garcia-Rivera, D.; Brandt, W. What can a chemist learn from nature's macrocycles? – A brief, conceptual view. *Mol. Divers.* **2005**, *9*, 171-186.

[13]   Marahiel, M.A. Working outside the protein-synthesis rules: insights into non-ribosomal peptide synthesis. *J. Pept. Sci.* **2009**, *15*, 799-807.

[14]   Reiber, K.; Neuhof, T.; Ozegowski, J.H.; von Dohrend, H.; Schwecke, T. A nonribosomal peptide synthetase involved in the biosynthesis of ampullosporins in *Sepedonium ampullosporum*. *J. Pept. Sci.* **2003**, *9*, 701-713.

[15]   Jegorov, A.; Hajduch, M.; Sulc, M.; Havlicek, V. Nonribosomal cyclic peptides: specific markers of fungal infections. *J. Mass Spectrom.* **2006**, *41*, 563-576.

[16]   Igarashi, Y.; Hanafusa, T.; Gohda, F.; Peterson, S.; Bills, G. Species-level assessment of secondary metabolite diversity among *Hamigera* species and a taxonomic note on the genus. *Mycology* **2014**, *5*, 102-109.

[17]   Brauer, M.C.N.; Neves Filho, R.A.W.; Westermann, B.; Heinke, R.; Wessjohann, L.A. Synthesis of antibacterial 1,3-diyne-linked peptoids from an Ugi-4CR/Glaser coupling approach. *Beilstein J. Org. Chem.* **2015**, *11*, 25-30.

[18]   Kobayashi, R.; Samejima, Y.; Nakajima, S.; Kawai, K.-I.; Udagawa, S.-I. Studies on fungal products. XI. Isolation and Sstructures of novel cyclic pentapeptides from *Aspergillus* sp. NE-45. *Chem. Pharm. Bull.* **1987**, *35*, 1347-1352.

[19] Mizutani, K.; Hirasawa, Y.; Sugita-Konishi, Y.; Mochizuki, N.; Morita, H. Structural and conformational analysis of hydroxycyclochlorotine and cyclochlorotine, chlorinated cyclic peptides from *Penicillium islandicum*. *J. Nat. Prod.* **2008**, *71*, 1297-1300.

[20] Li, H.J.; Lin, Y.C.; Yao, J.H.; Vrijmoed, L.L.; Jones, G.E. Two new metabolites from the mangrove endophytic fungus no. 2524. *J. Asian Nat. Prod. Res.* **2004**, *6*, 185-191.

[21] Talontsi, F.M.; Facey, P.; Tatong, M.D.; Tofazzal Islam, M.; Frauendorf, H.; Draeger, S.; Tiedemann, A.; Laatsch, H. Zoosporicidal metabolites from an endophytic fungus *Cryptosporiopsis* sp. of *Zanthoxylum leprieurii*. *Phytochemistry* **2012**, *83*, 87-94.

[22] Gulder, T.; Hong, H.; Correa, J.; Egereva, E.; Wiese, J.; Imhoff, J.; Gross, H. Isolation, Structure elucidation and total synthesis of lajollamide A from the marine fungus *Asteromyces cruciatus*. *Mar. Drugs* **2012**, *10*, 2912-2935.

[23] Ye, P.; Shen, L.; Jiang, W.; Ye, Y.; Chen, C.T.; Wu, X.; Wang, K.; Wu, B. Zn-driven discovery of a hydrothermal vent fungal metabolite clavatustide C, and an experimental study of the anti-cancer mechanism of clavatustide B. *Mar. Drugs* **2014**, *12*, 3203-3217.

[24] Li, G.; Kusari, S.; Golz, C.; Strohmann, C.; Spiteller, M. Three cyclic pentapeptides and a cyclic lipopeptide produced by endophytic *Fusarium decemcellulare* LG53. *RSC Adv.* **2016**, *6*, 54092-54098.

[25] Ma, X.; Nong, X.-H.; Ren, Z.; Wang, J.; Liang, X.; Wang, L.; Qi, S.-H. Antiviral peptides from marine gorgonian-derived fungus *Aspergillus* sp. SCSIO 41501. *Tetrahedron Lett.* **2017**, *58*, 1151-1155.

[26] Chen, M.; Shao, C.L.; Fu, X.M.; Kong, C.J.; She, Z.G.; Wang, C.Y. Lumazine peptides penilumamides B-D and the cyclic pentapeptide asperpeptide A from a gorgonian-derived *Aspergillus* sp. fungus. *J. Nat. Prod.* **2014**, *77*, 1601-1606.

[27] Lewer, P.; Graupner, P.R.; Hahn, D.R.; Karr, L.L.; Duebelbeis, D.O.; Lira, J.M.; Anzeveno, P.B.; Fields, S.C.; Gilbert, J.R.; Pearce, C. Discovery, synthesis, and insecticidal activity of cycloaspeptide E. *J. Nat. Prod.* **2006**, *69*, 1506-1510.

[28] Dalsgaard, P.W.; Larsen, T.O.; Christophersen, C. Bioactive cyclic peptides from the psychrotolerant fungus *Penicillium algidum*. *J. Antibiot. (Tokyo)* **2005**, *58*, 141-144.

[29] Zhang, Y.; Liu, S.; Liu, H.; Liu, X.; Che, Y. Cycloaspeptides F and G, cyclic pentapeptides from a *Cordyceps*-colonizing isolate of *Isaria farinosa*. *J. Nat. Prod.* **2009**, *72*, 1364-1367.

[30] Arai, N.; Shiomi, K.; Yamaguchi, Y.; Masuma, R.; Iwai, Y.; Turberg, A.; Kolbl, H.; Omura, S. Argadin, a new chitinase inhibitor, produced by *Clonostachys* sp. FO-7314. *Chem. Pharm. Bull.* **2000**, *48*, 1442-1446.

[31] Shiomi, K.; Arai, N.; Iwai, Y.; Turberg, A.; Kölbl, H.; Ōmura, S. Structure of argifin, a new chitinase inhibitor produced by *Gliocladium* sp. *Tetrahedron Lett.* **2000**, *41*, 2141-2143.

[32] Krause, C.; Kirschbaum, J.; Bruckner, H. Peptaibiomics: an advanced, rapid and selective analysis of peptaibiotics/peptaibols by SPE/LC-ES-MS. *Amino Acids* **2006**, *30*, 435-443.

[33] Eckart, K. Mass spectrometry of cyclic peptides. *Mass Spectrom. Rev.* **1993**, *13*, 23-55.

[34] Ngoka, L.C.M.; Gross, M.L. Multistep tandem mass spectrometry for sequencing cyclic peptides in an ion-trap mass spectrometer. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 732-746.

[35] Ngoka, L.C.M.; Gross, M.L. A Nomenclature system for labeling cyclic peptide fragments. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 360-363.

[36] Serafin, S.S.; Maranan, R.; Zhang, K.; Morton, T.H. Mass spectrometric differentiation of linear peptides composed of L-amino acids from isomers containing one D-amino acid residue. *Anal. Chem.* **2005**, *77*, 5480-5487.

[37] Stark, S. Utilization of the Ugi four-component reaction for the synthesis of lipophilic peptidomimetics as potential antimicrobials. Martin-Luther-University Halle-Wittenberg, Halle (Saale), **2016**.

[38] Stiernagle, T. Maintenance of *C. elegans*. In Wormbook WormBook; Editor: The *C. elegans* Research Community. Available online: http://dx.doi.org/10.1895/wormbook.1.101.1 (accessed on 17.06.2020).

[39] Thomsen, H.; Reider, K.; Franke, K.; Wessjohann, L.A.; Keiser, J.; Dagne, E.; Arnold, N. Characterization of constituents and anthelmintic properties of *Hagenia abyssinica*. *Sci. Pharm.* **2012**, *80*, 433-446.

Chapter 4

-

HPTLC-DESI-HRMS based Profiling of Anthraquinones in Complex Mixtures –

A Proof-Of-Concept Study using Crude Extracts of Chilean Mushrooms

*Abstract**

High-performance thin-layer chromatography (HPTLC) coupled with negative ion desorption electrospray ionization high-resolution mass spectrometry (DESI-HRMS) was used for the analysis of anthraquinones in complex crude extracts of Chilean dermocyboid Cortinarii. For this proof-of-concept study, the known anthraquinones emodin, physcion, endocrocin, dermolutein, hypericin, and skyrin were identified by their elemental composition. HRMS also allowed the differentiation of the investigated anthraquinones from accompanying compounds with the same nominal mass in the crude extracts. An investigation of the characteristic fragmentation pattern of skyrin in comparison with a reference compound showed, exemplarily, the feasibility of the method for the determination of these coloring, bioactive and chemotaxonomically important marker compounds. Accordingly, we demonstrate that the coupling of HPTLC with DESI-HRMS represents an advanced and efficient technique for the detection of anthraquinones in complex matrices. This analytical approach may be applied in the field of anthraquinone-containing food and plants such as *Rheum* spp. (rhubarb), *Aloe* spp., *Morinda* spp., *Cassia* spp. and others. Furthermore, the described method can be suitable for the analysis of anthraquinone-based colorants and dyes, which are used in the food, cosmetic, and pharmaceutical industry.

4.1 Introduction

Anthraquinones represent a large family of naturally occurring pigments, which are produced by plants, microbes, lichens, insects, and fungi.[1] Besides their coloring properties, these natural products exhibit a broad range of bioactivities such as antibacterial, antiparasitic, anti-inflammatory, fungicidal, insecticidal, laxative, antiviral, and anticancer but also DNA intercalating properties.[2–7] The chemical structure of anthraquinones is based on an anthracene skeleton with two keto groups in position 9 and 10. The basic core unit can be further substituted at various positions and connected with sugar molecules, forming the corresponding glycosides.[8,9]

In the literature, about 700 anthraquinone derivatives are described, in which emodin, physcion, catenarin, and rhein are the most frequently reported.[9–11] Two hundred of these are described for flowering plants, which also occur in edible plants and vegetables such as *Rheum*, *Aloe* and *Cassia* species, while the remaining ones are produced by lichens and fungi.[7,8,12]

The genus *Cortinarius* (including Dermocybe) is one of the most diverse genera of basidiomycetous fungi, containing a large variety of anthraquinones.[13–15] The occurrence and distribution of these pigments is closely linked to species diversity and allows their use as chemotaxonomic marker compounds in species delimination.[16–22]

The analysis of anthraquinones is of interest due to their wide range of applications. A continuous improvement of the analytical techniques is needed to overcome difficulties with respect to interference with various types of matrices and low abundance of the analytes within complex mixtures.[7]

Thin-layer chromatography is an effective method for the chromatographic separation of anthraquinones.[23–26] Furthermore, several mass spectrometry-based methods have been developed for a deep analysis of anthraquinones providing characteristic [M-H]⁻ ions in negative ion mode.[27–29]

Desorption electrospray ionization mass spectrometry (DESI-MS) represents a powerful ambient ionization mass spectrometric technique which enables a direct ionization of analytes from surfaces with subsequent mass spectrometric detection.[30–32] The coupling of DESI-MS with high-performance thin-layer chromatography (HPTLC) provides a robust methodological approach for the separation and highly sensitive detection of secondary metabolites in plants and fungi.[33–35] Furthermore, this method is suitable for the fingerprint analysis of crude extracts in natural product research.[36,37] Recently, the detection of excreted polyhydroxyanthraquinones from the surface of fungal culture agar plates using DESI-MS in a negative ion mode was reported.[38]

4.2 Aims and Scope

In the present paper, we report the development of a rapid profiling method of anthraquinones, exemplified with the analysis of different crude extracts from Chilean dermocyboid Cortinarii concerning their anthraquinone pattern based on the combination of HPTLC with negative ion DESI-HRMS. For this proof-of-concept study, extracts from fruiting bodies of six dermocyboid Cortinarii were investigated for the occurrence of the known anthraquinones emodin, physcion, endocrocin, dermolutein, hypericin, and skyrin. Furthermore, the possibility of performing MS/MS experiments on the desorbed analytes directly from the HPTLC plate was exemplarily shown for the bisanthraquinone skyrin in comparison with data obtained from direct-infusion MS experiments.

4.3 Results and Discussion

4.3.1 Method Development

The normal-phase HPTLC plates were developed for 55 mm in one dimension to enable the separation of anthraquinones according to their polarity (Figure 4.1). The geometry of the source, the composition of the spray solvent, the flow rate as well as the scanning rate were optimized for the analysis.
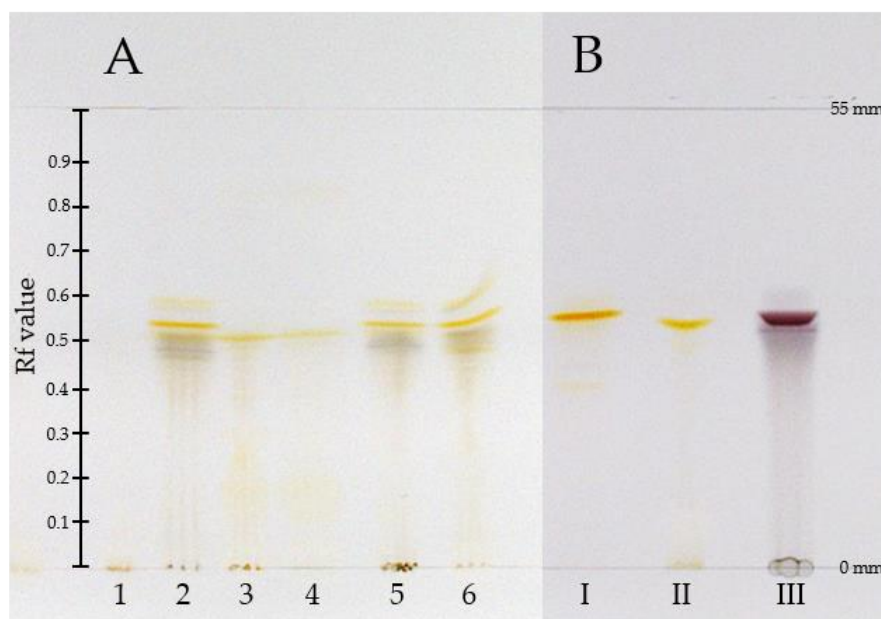
Figure 4.1 High-performance thin-layer chromatograph (HPTLC) of (**A**) methanolic crude extracts of *Cortinarius* (Dermocybe) spec. (1), *C.* (D.) *austronanceiensis* (2), *C.* (D.) *icterina* (3), *C.* (D.) *icterinula* (4), *C. obscuro-olivea* (5), *C. viridulifolius* (6) and (**B**) reference compound skyrin (**4.6**, I), endocrocin (**4.3**, II), and hypericin (**4.5**, III) (mobile phase: toluene, ethyl formate, and formic acid (10:5:3; *v/v/v*) distance from application line to solvent front: 55 mm).

To enhance the ionization and desorption efficiency, different mixtures of methanol and water (with and without formic acid as additive) were tested as spray solvents. A mixture of methanol

109

and water of 1:1 (*v/v*) yielded the best results. During optimization, a flow rate of 2 µL/min showed good results to obtain adequate signal intensities. On the other hand, higher flow rates led to a partial detachment of silica gel particles. Additionally, different velocities of the DESI spray head were tested to obtain an efficient number of precursor ions for the MS^2 experiments. Lower scan rates led to better signal intensities due to the better desorption of the analytes from the surface of the HPTLC plates. Therefore, we used a lower velocity for the MS^2 experiments in the final measurements than in the full scan runs. Each band was recorded by scanning the surface in the y-direction (R_f 0 to 1.0) with an automated DESI source coupled to an Orbitrap Elite mass spectrometer within a total run time of 4.6 min. Before starting the experiment, the spray head was positioned on the application line of the HPTLC followed by the manual start of the MS measurement.

4.3.2 Profiling of Anthraquinones in Crude Extracts

The pigment pattern of the methanolic extracts of dermocyboid Cortinarii *Cortinarius* (Dermocybe) *austronanceiensis*, *C*. (D.) *icterina*, *C*. (D.) *icterinula*, *C*. (D.) *obscuro-olivea*, *C*. (D.) spec., and *C*. (D.) *viridulifolius* (Figure 4.1) was analyzed by high-performance thin-layer chromatography (HPTLC) coupled to desorption electrospray ionization (DESI) mass spectrometry in the negative ion mode. An unspotted HPTLC band was scanned to assign background related peaks (Appendix, Figure C1) and to ensure the absence of the target compounds before applying the crude extracts on the HPTLC plate. No anthraquinone-related peaks could be detected by scanning the empty band on the HPTLC plate after running with the solvent system. This is demonstrated by the extracted ion chromatograms based on the theoretical calculated *m/z* value of the [M-H]$^-$ ions using an 25 ppm window (four decimals) (Appendix, Figure C2).
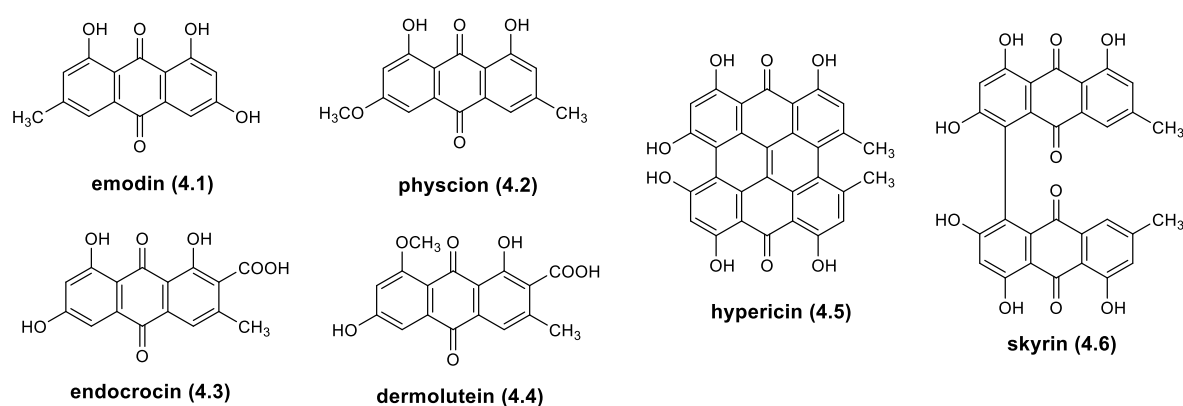
Figure 4.2 Structures of investigated anthraquinones **4.1**-**4.6**.

The established analytical method was applied to identify anthraquinones **4.1**–**4.6** (Figure 4.2). These anthraquinones were chosen for this proof-of-concept study because their occurrence in different *Cortinarius* and *Dermoybe* species is described in the literature.[15] The assignment of the

structures is based on their elemental composition determined by high-resolution mass spectrometry (HRMS) (Table 4.1 and Appendix, Table C2).

Table 4.1 Detected anthraquinones (**4.1**-**4.6**) using HPTLC-desorption electrospray ionization (DESI)-high-resolution mass spectroscopy (HRMS).

No.	Elemental Composition [M-H]$^-$	Theoretical *m/z* [M-H]$^-$	*C.* (D.) *austronanceiensis*	*C.* (D.) *icterina*	*C.* (D.) *icterinula*	*C.* (D.) *obscuro-olivea*	*C.* (D.) spec.	*C.* (D.) *viridulifolius*
4.1	$C_{15}H_9O_5^-$	269.0455	+	+	+	+	+	+
4.2	$C_{16}H_{11}O_5^-$	283.0612	+	+	n.d.	+	+	+
4.3	$C_{16}H_9O_7^-$	313.0354	+	+	+	+	n.d.	+
4.4	$C_{17}H_{11}O_7^-$	327.0510	+	+	+	+	+	+
4.5	$C_{30}H_{15}O_8^-$	503.0772	+	n.d.	n.d.	+	n.d.	+
4.6	$C_{30}H_{17}O_{10}^-$	537.0827	+	n.d.	n.d.	+	n.d.	+

n.d. = not detected; + = detected.

The negative ion DESI mass spectra of the methanolic extract of *C.* (D.) *austronanceiensis* afforded characteristic deprotonated ions of emodin (**4.1**), [M-H]$^-$ at *m/z* 269.0450 calcd for $C_{15}H_9O_5^-$ 269.0455), physcion (**4.2**), [M-H]$^-$ at *m/z* 283.0611 calcd for $C_{16}H_{11}O_5^-$ 283.0612), endocrocin (**4.3**), [M-H]- at *m/z* 313.0349 calcd for $C_{16}H_9O_7^-$ 313.0354), dermolutein (4.4), [M-H]$^-$ at *m/z* 327.0505 calcd for $C_{17}H_{11}O_7^-$ 327.0510), hypericin (**4.5**, [M-H]$^-$ at *m/z* 503.0763 calcd for $C_{30}H_{15}O_8^-$ 503.0772) and skyrin (**4.6**, [M-H]$^-$ at *m/z* 537.0817 calcd for $C_{30}H_{17}O_{10}^-$ 537.0827) (Figure 4.3A).
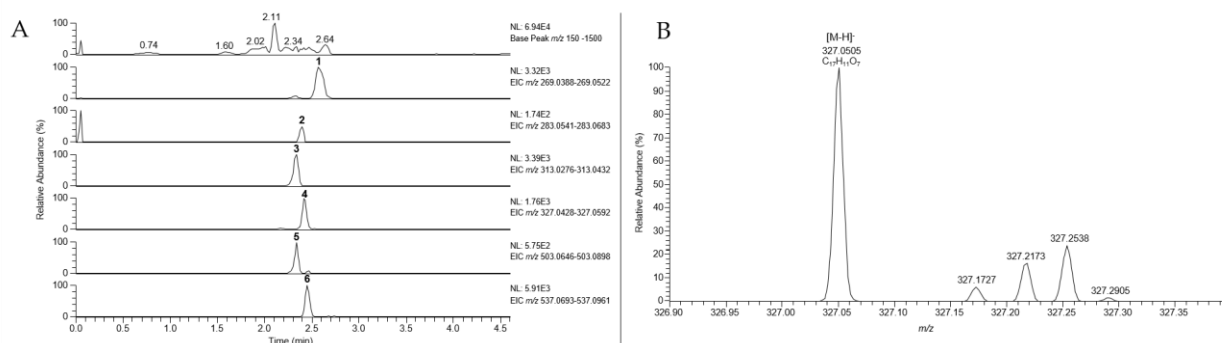
Figure 4.3 (**A**) Extracted ion chromatograms (EIC, mass window: 25 ppm) of anthraquinones **4.1**-**4.6** from crude extract of *Cortinarius* (D.) *austronanceiensis*, (**B**) Extracted ion chromatogram (EIC) of dermolutein (**4.3**, *m/z* 327) acquired during DESI-HR-MS measurement of methanolic extract from *C.* (*D.*) *austronanceiensis*.

For the data evaluation, the target *m/z* values were extracted from the total ion chromatogram using a 25 ppm mass window with a mass accuracy of four decimals to obtain the corresponding extracted ion chromatograms (EICs) for each analyte. The EICs for the methanolic extracts of *C.* (D.) *icterina*, *C.* (D.) *icterinula*, *C.* (D.) *obscuro-olivea*, *C.* (D.) spec., and *C.* (D.) *viridulifolius* are shown in the Appendix Figures C3–C7, and the presences of the different target analytes within the extracts are visualized in Table 4.1. Due to the resolving power of the orbitrap detector, a

differentiation of isobaric ions was possible as shown in the EIC of dermolutein (**4.4**, Figure 4.3 B). The anthraquinone peak *m/z* 327.0505 is clearly separated from other accompanying ions at the same nominal mass using a resolution of 30,000.

Comparing the pigment patterns of the different fungal extracts (Table 4.1), all targets could be detected in the methanolic extracts of *Cortinarius* (D.) *austronanceiensis*, *C*. (D.) *obscuro-olivea* and *C*. (D.) *viridulifolius*. The naphthodianthrone hypericin (**4.5**) and the bisanthraquinone skyrin (**4.6**) were not detectable along the HPTLC bands of *C*. (D.) *icterina*, *C*. (D.) *icterinula* and *C*. (D.) spec.

Based on the retention time and the velocity of scanning the HPTLC bands (see equation 1), R_f values can be calculated and compared with the R_f values determined directly from the HPTLC plate (Table 4.2). The results of the developed HPTLC plates of the extracts and the reference compounds (see Appendix, Figure C7–C10) were reproducible and comparable, exemplified based on the extracted ion chromatograms of endocrocin (Appendix, Figure C11). Therefore, the determination of R_f values based on the retention time of the HPTLC-DESI-HRMS measurements of UV/VIS inactive analytes becomes possible.

$$R_f = t_R \text{ (min)} \times \text{velocity (mm/s)} \times 60 \times 1/\text{distance from application line to solvent front (mm)}$$

$$R_f = t_R \times 0.200 \text{ mm/s} \times 60 \times 1/55 \text{ mm}$$

(1)

Table 4.2 R_f value and calculation from crude extract of *C*. (D.) *austronanciensis* (Figure 4.1, band 2).

Compound	R_f (experimental)	t_R DESI (min)	R_f (calculated)	Spot Color Visible Light	Spot Color UV Light (254 nm)	Spot Color UV Light (366 nm)
4.1	0.58	2.57	0.56	yellow	dark	orange
4.2	0.54	2.40	0.52	yellow	dark	orange
4.3 *	0.5	2.34	0.51	yellow	dark	orange
4.4	0.49	2.42	0.53	yellow	dark	red
4.5 *	0.5	2.34	0.51	black	dark	red
4.6*	0.53	2.45	0.53	yellow-orange	dark	red brown

* Confirmed with reference compound.

4.3.3 Structural Characterization Using MS² Experiments

As an example, the fragmentation behavior of the bisanthraquinone skyrin (**4.6**) was investigated by a MS^2 measurement compared with the results obtained directly from the extract, data of a reference compound measured by HPTLC-DESI-HR-MS and with direct infusion DESI-HR-MS (Figure 4.4 A–C, Appendix, Table C3). Skyrin (**4.6**), in its MS^2 spectrum, shows a base peak ion at *m/z* 493.0923 ([M-H-CO$_2$]$^-$, calcd for C$_{29}$H$_{17}$O$_8^-$ 493.0929 Figure 4.4 A, Appendix, Table C3). Furthermore, a loss of carbon suboxide (C$_3$O$_2$) is observed at *m/z* 469.0926 (calcd for C$_{27}$H$_{17}$O$_8^-$ 469.0929), indicating a 1,3-dihydroxybenzene feature, which is also described for flavones and

other polyphenols.[39,40] The obtained data are in good agreement with the reported MS^2 data of skyrin.[41]
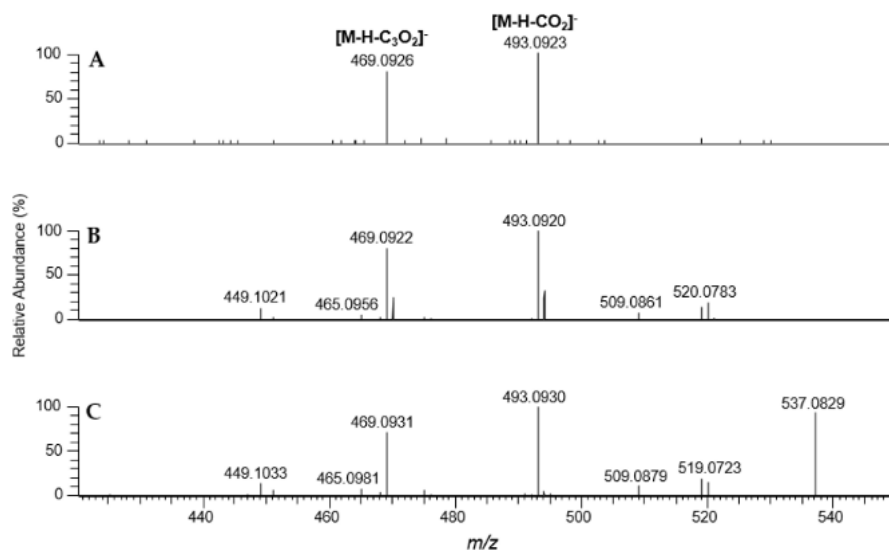
Figure 4.4 HPTLC-DESI-HRMS^2 data of skyrin (**4.6**); (**A**) from fungal extract of *Cortinarius* (D.) *austronanceiensis* (NCE: 50%); (**B**) skyrin standard (NCE 35%); (**C**) direct infusion ESI-HRMS^2 (NCE 30%).

4.4 Conclusions

Crude extracts of six Chilean dermocyboid Cortinarii were investigated by HPTLC-negative ion DESI-HRMS concerning the occurrence of the anthraquinones physcion (**4.1**), emodin (**4.2**), endocrocin (**4.3**), dermolutein (**4.4**), hypericin (**4.5**), and skyrin (**4.6**). The compounds were identified by their elemental composition. It should be pointed out that the high-resolution mass spectrometry (HRMS) approach also allows a mass spectral distinction of isobaric ions as demonstrated for the detection of dermolutein (**4.4**) whose nominal mass is accompanied by other compounds in the crude extract. Furthermore, the implementation of fragmentation experiments (MS^2) for anthraquinones on HPTLC surfaces is possible, as exemplarily shown for the detection of skyrin (**4.6**) in the extract of *C.* (D) *austronanceiensis*, and could be a valuable tool for the presence of these compound classes. The corresponding results are in good agreement with the data obtained by direct infusion and in comparison, with the LC-MS data reported in literature.

HPTLC provides good separation efficiencies and can be performed in an automated and controlled way with respect to the sample application and the development of the plate. In classical approaches, a derivatization of the HPTLC plate is needed; however, combined with DESI-MS, this step is not required. Although the separation power of HPTLC is lower than in (U)HPLC, several analyses can be performed on plate and within a short analysis time. In case of the presented approach, a HPTLC plate (total length 100 mm) with a developing length of 55 mm, and a total scanning time of only 4.6 min was sufficient to obtain the presented results. After the extraction of the material, no further sample preparation steps are necessary, and the crude extracts

can be directly applied to the plate, representing an advantage compared with other analytical techniques.

In summary, the obtained results illustrate the feasibility and capacity of HPTLC-DESI-HRMS to provide a rapid first screening method for the analysis of anthraquinones in complex mixtures, which may be used in the analysis of anthraquinones in food, plants, fungi, dyes, and cosmetic and pharmaceutical products.

4.5 Experimental Section

4.5.1 Reagents and Chemicals

The authentic reference compounds endocrocin (**3**), hypericin (**5**) and skyrin (**6**) were available from the in-house compound library of the Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry (IPB), Halle (Saale), Germany. Methanol and toluene were used at analytical grade. Ethyl formate was purchased from Merck (Darmstadt, Germany) and formic acid from Roth (Karlsruhe, Germany). LC-MS grade methanol was obtained from Merck (Darmstadt, Germany), and purified water was prepared by Merck Millipore Milli-Q equipment (Darmstadt, Germany).

4.5.2 Sampling Sites and Extraction

Fruiting bodies of *Cortinarius* (D.) *austronanceiensis*, *C.* (D.) *icterina*, *C.* (D.) *icterinula*, *C.* (D.) *obscuro-olivea*, *C.* (D.) spec., and *C.* (D.) *viridulifolius* were collected in Chile (detailed information see Table S1). Voucher specimens are deposited in the Fungarium of Concepción University (CONC-F). A duplicate is deposited at the Leibniz Institute of Plant Biochemistry.

Air-dried fruiting bodies (2 g) were homogenized using 15 mL of acetone in a blender followed by an ultrasonic extraction for 15 min to remove interfering compounds such as fatty acids from the material. After vacuum-supported filtration, the fungal material residue was further extracted twice with 15 mL methanol each. The resulting extracts were filtrated and dried under reduced pressure using a rotary evaporator. The crude methanolic extracts were redissolved in methanol and directly spotted on the HPTLC plate for chromatographic separation.

4.5.3 HPTLC

HPTLC was performed on glass HPTLC silica gel 60 F_{254} plates (10×10 cm, layer thickness 150–200 µm, Merck) using a mixture of toluene, ethyl formate, and formic acid (10:5:3; *v/v/v*) as a mobile phase. After air drying, the developed plates were parted by a glass cutter and subjected to the mass spectral analysis. For documentation and R_f-value determination, a CAMAG TLC visualizer (CAMAG, Muttenz, Switzerland) was used with the software winCATS (version 1.4.9.2001, CAMAG, Switzerland).

4.5.4 DESI-Orbitrap-MS and MS[2]

All experiments were performed using a 2D-DESI source (Omnispray System OS-3201, Prosolia, Indianapolis, IN, USA) coupled to an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) operated in the negative ion mode. The DESI source settings were as follows: spray voltage, 3 kV; solvent flow rate, 2 µL/min; nebulizing gas (nitrogen), pressure, 7 bar; tip-to-surface distance, 2–2.5 mm; tip-to-inlet distance, 3.5 mm; incident angle (relative to the surface plane), 55°. The DESI spray solvent was 50:50 (*v/v*) methanol/water. MS experiments were performed by continuously scanning every HPTLC band in the x-direction at a surface velocity of 200 µm/s while acquiring mass spectra in full scan mode (*m/z* 150–1500; resolution 30,000) and

150 µm/s in MS^2 mode. Collision-induced dissociation was performed using normalized collision energies (NCE) of 35 and 50 (arbitrary units) and an isolation width of ± 2 Da. The data were evaluated using the software Xcalibur 2.2 SP1 (Thermo Fisher Scientific).

4.5.5 Supplementary Materials

The following data are available see Appendix C and online at www.mdpi.com/2304-8158/9/2/156/s1, The spectroscopic data are available at the public repository RADAR.[42]

4.6 References

[1]   Thomson, R.H. Naturally Occurring Quinones IV, 4th ed.; Springer: Dordrecht, The Netherlands, 1997; pp. 309–483, ISBN: 978-0-7514-0248-3.

[2]   Wakulinski, W.; Kachlicki, P.; Sobiczewski, P.; Schollenberger, M.; Zamorski, C.; Lotocka, B.; Sarova, J. Catenarin production by isolates of *Pyrenophora tritici-repentis* (Died.) Drechsler and its antimicrobial activity. *J. Phytopathol.* 2003, *151*, 74-79.

[3]   Reynolds, T. Aloes—The Genus Aloe, 1st ed.; CRC Press: Boca Raton, FL, USA, 2004, ISBN: 978-8183293075.

[4]   Srinivas, G.; Babykutty, S.; Sathiadevan, P.P.; Srinivas, P. Molecular Mechanism of emodin sction: Transition from laxative ingredient to an antitumor agent. *Med. Res. Rev.* **2007**, *27*, 591-608.

[5]   Locatelli, M. Anthraquinones: Analytical techniques as a novel tool to investigate on the triggering of biological targets. *Curr. Drug Targets* **2011**, *12*, 366-380.

[6]   Chien, S.C.; Wu, Y.C.; Chen, Z.W.; Yang, W.C. Naturally occurring anthraquinones: chemistry and therapeutic potential in autoimmune diabetes. *Evid. Based Complement. Alternat. Med.* **2015**, *2015*, 1-13.

[7]   Duval, J.; Pecher, V.; Poujol, M.; Lesellier, E. Research advances for the extraction, analysis and uses of anthraquinones: A review. *Ind. Crops Prod.* **2016**, *94*, 812-833.

[8]   Seigler, D.S. Plant Secondary Metabolism; Springer: New York, NY, USA, 1998; pp. 85, ISBN: 978-0-412-01981-4.

[9]   Gessler, N.N.; Egorova, A.S.; Belozerskaya, T.A. Fungal anthraquinones. *Appl. Biochem. Microbiol.* **2013**, *49*, 85-99.

[10]   Caro, Y.; Anamale, L.; Fouillaud, M.; Laurent, P.; Petit, T.; Dufosse, L. Natural hydroxyanthraquinoid pigments as potent food grade colorants: an overview. *Nat. Prod. Bioprospect.* **2012**, *2*, 174-193.

[11]   Fouillaud, M.; Venkatachalam, M.; Girard-Valenciennes, E.; Caro, Y.; Dufosse, L. Anthraquinones and derivatives from marine-derived fungi: structural diversity and selected biological activities. *Mar. Drugs* **2016**, *14*, 1-64.

[12]   Fouillaud, M.; Caro, Y.; Venkatachalam, M.; Grondin, I.; Laurent, D. Anthraquinones. In Phenolic Compounds in Food—Characterization and Analysis; Nollet, L.M.L., Gutierrez-Uribe, J.A., Eds.; CRC Press: Boca Raton, FL, USA, 2018; pp. 130–170, ISBN: 78-1-4987-2296-4.

[13]   Gill, M.; Steglich, W. Pigments of fungi (Macromycetes). *Progr. Chem. Org. Chem. Nat. Prod.* **1987**, *51*, 1-317.

[14]   Gill, M. New pigments of *Cortinarius* Fr. and *Dermocybe* (Fr.) Wünsche (Agaricales) from Australia and New Zealand. *Beih. Sydowia* **1995**, *1*, 73-87.

[15]   Keller, G.; Moser, M.; Horak, E.; Steglich, W. Chemotaxonomic investigations of species of *Dermocybe* (Fr. Wünsche (Agaricales) from New Zealand, Papua New Guinea and Argentina. *Beih. Sydowia* **1988**, *10*, 101-126.

[16]   Steglich, W.; Austel, V. Die Struktur des Dermocybins und des Dermoglaucins. *Tetrahedron Lett.* **1966**, *26*, 3077-3079.

[17]   Gruber, I. Anthrachinonfarbstoffe in der Gattung *Dermocybe* und Versuch ihrer Auswertung für die Systematik. *Zeitschr. f. Pilzk.* **1970**, *36*, 95-112.

[18]   Arnold, N.; Besl, A.; Bresinsky, A.; Kemmer, H. Notizen zur Chemotaxonomie der Gattung *Dermocybe* (Agaricales) und zu ihrem Vorkommen in Bayern. *Z. Mykol.* **1987**, *53*, 187-194.

[19]   Arnold, N. Morphologisch-Anatomische und Chemische Untersuchungen an der Untergattung Telamonia (Cortinarius, Agaricales); IHW-Verlag: Eching, München, Germany, **1993**.

[20]   Jones, R.H.; May, T.W. Pigment chemistry and morphology support recognition of *Cortinarius austrocinnabarinus* sp. nov. (Fungi: Cortinariaceae) from Australia. *Muelleria* **2008**, *26*, 77-87.

[21]   Stefani, F.O.; Jones, R.H.; May, T.W. Concordance of seven gene genealogies compared to phenotypic data reveals multiple cryptic species in Australian dermocyboid *Cortinarius* (Agaricales). *Mol. Phylogenet. Evol.* **2014**, *71*, 249-260.

[22]   Greff, A.; Porzel, A.; Schmidt, J.; Palfner, G.; Arnold, N. Pigment pattern of the Chilean mushroom *Dermocybe nahuelbutensis* Garrido & E. Horak. *Rec. Nat. Prod.* **2017**, *11*, 547-551.

[23]   Shibata, S.; Takito, M.; Tanaka, O. Paper chromatography of anthraquinone pigments. *J. Am. Chem. Soc.* **1950**, *72*, 2789-2790.

117

[24]   Kidd, C.B.M.; Caddy, B.; Robertson, J.; Tebbett, I.R. Thin-layer chromatography as an aid for identification of *Dermocybe* species of *Cortinarius*. *Trans. Br. mycol. Soc.* **1985**, *85*, 213-221.

[25]   Ma, X.; Chen, Y.; Hui, R. Analysis of Anthraquinones in *Rheum Franzenbachii* Münt (Rhubarb) by Thin-Layer Chromatography. *Chromatographia* **1989**, *27*, 465-466.

[26]   Räisänen, R.; Björk, H.; Hynninen, H. Two-Dimensional TLC separation and mass spectrometric identification of anthraquinones isolated from the fungus *Dermocybe sanguinea*. *Z. Naturforsch.* **2000**, *55c*, 195-202.

[27]   Lin, C.-C.; Wu, C.-I.; Lin, T.-C.; Sheu, S.-J. Determination of 19 rhubarb constituents by high-performance liquid chromatography–ultraviolet–mass spectrometry. *J. Sep. Sci.* **2006**, *29*, 2584-2593.

[28]   Ye, M.; Han, J.; Chen, H.; Zheng, J.; Guo, D. Analysis of phenolic compounds in rhubarbs using liquid chromatography coupled with electrospray ionization mass spectrometry. *J. Am. Soc. Mass. Spectrom.* **2007**, *18*, 82-91.

[29]   Derksen, G.C.H.; Niederländer, H.A.G.; Van Beek, T.A. Analysis of anthraquinones in *Rubia tinctorium* L. by liquid chromatography coupled with diode-array UV and mass spectrometric detection. *J. Chromatogr. A* **2002**, *978*, 119-127.

[30]   Takats, Z.; Wiseman, J.M.; Gologan, B.; Cooks, R.G. Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science* **2004**, *306*, 471-473.

[31]   Takats, Z.; Wiseman, J.M.; Cooks, R.G. Ambient mass spectrometry using desorption electrospray ionization (DESI): instrumentation, mechanisms and applications in forensics, chemistry, and biology. *J. Mass Spectrom.* **2005**, *40*, 1261-1275.

[32]   Cooks, R.G.; Ouyang, Z.; Takats, Z.; Wiseman, J.M. Detection Technologies. Ambient Mass Spectrometry. *Science* **2006**, *311*, 1566-1570.

[33]   Van Berkel, G.J.; Tomkins, B.A.; Kertesz, V. Thin-layer chromatography/desorption electrospray ionization mass spectrometry: investigation of goldenseal alkaloids. *Anal. Chem.* **2007**, *79*, 2778-2789.

[34]   Lane, A.L.; Nyadong, L.; Galhena, A.S.; Shearer, T.L.; Stout, E.P.; Parry, R.M.; Kwasnik, M.; Wang, M.D.; Fernandez, F.M.; Kubanek, J. Desorption electrospray ionization mass spectrometry reveals surface-mediated antifungal chemical defense of a tropical seaweed. *PNAS* **2009**, *106*, 7314-7319.

[35]   Nyadong, L.; Hohenstein, E.G.; Galhena, A.; Lane, A.L.; Kubanek, J.; Sherrill, C.D.; Fernandez, F.M. Reactive desorption electrospray ionization mass spectrometry (DESI-MS) of natural products of a marine alga. *Anal. Bioanal. Chem.* **2009**, *394*, 245-254.

[36]   Bagatela, B.S.; Lopes, A.P.; Cabral, E.C.; Perazzo, F.F.; Ifa, D.R. High-performance thin-layer chromatography/desorption electrospray ionization mass spectrometry imaging of the crude extract from the peels of *Citrus aurantium* L. (Rutaceae). *Rapid Commun. Mass Spectrom.* **2015**, *29*, 1530-1534.

[37]   Kennedy, J.H.; Wiseman, J.M. Direct analysis of *Salvia divinorum* leaves for salvinorin A by thin layer chromatography and desorption electrospray ionization multi-stage tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2010**, *24*, 1305-1311.

[38]   Figueroa, M.; Jarmusch, A.K.; Raja, H.A.; El-Elimat, T.; Kavanaugh, J.S.; Horswill, A.R.; Cooks, R.G.; Cech, N.B.; Oberlies, N.H. Polyhydroxyanthraquinones as quorum sensing inhibitors from the guttates of *Penicillium restrictum* and their analysis by desorption electrospray ionization mass spectrometry. *J. Nat. Prod.* **2014**, *77*, 1351-1358.

[39]   Fabre, N.; Rustan, I.; De Hoffmann, E.; Quetin-Leclercq, J. Determination of flavone, flavonol, and flavanone aglycones by negative ion liquid chromatography electrospray ion trap mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 707-715.

[40]   Schmidt, J. Negative ion electrospray high-resolution tandem mass spectrometry of polyphenols. *J. Mass Spectrom.* **2016**, *51*, 33-43.

[41]   Jahn, L.; Schafhauser, T.; Wibberg, D.; Ruckert, C.; Winkler, A.; Kulik, A.; Weber, T.; Flor, L.; van Pee, K.H.; Kalinowski, J., et al. Linking secondary metabolites to biosynthesis genes in the fungal endophyte *Cyanodermella asteris*: The anti-cancer bisanthraquinone skyrin. *J. Biotechnol.* **2017**, *257*, 233-239.

[42]   Laub, A.; Sendatzki, A.-K.; Schmidt, J.; Arnold, N. Dataset: Desi-ms data for "HPTLC-DESI-HRMS based profiling of anthraquinones in complex mixtures—A proof-of-concept study using crude extracts of chilean mushrooms". RADAR (Reasearch Data Repository) v1.33. 2015.

Chapter 5 - General Discussion and Conclusions

The demand for new active active principles as a basis for drugs against serious diseases such as cancer or neurodegenerative disorders is high. In addition, due to the increasing resistance to antibiotics, new sources of effective bioactive compounds are required. Natural products from plants, fungi and microorganisms are the most important sources in drug discovery since the start of any medicine treatment. They exhibit fantastic chemical diversity and specificity, which has been continuously refined during evolution. The work with crude extracts as starting point for natural product isolation and identification is challenging due to matrix complexity and the meager bioactive compounds quantity within it. Therefore, sensitive and advanced analytical methods constitute the foundation of any novel advances in natural product research.

New analytical approaches allowing for higher performance in secondary metabolite identification are of utmost relevance. In this regard, targeted and untargeted metabolomic approaches, in combination with data mining workflows, provide high reliability for metabolite annotation and a bonus amenable to a certain degree of automation. [1-7] Furthermore, the exploration of new biological resources opens the scope towards the discovery of new metabolites and substance classes, pushing the frontiers to a better understanding of MS fragmentation patterns and NMR profiles. This knowledge is the basis for comprehensive databases,[8] as well as for the development and optimization of automated annotation tools, e.g. *in-silico* fragmentation tools, useful for dereplication processes and natural product identification.[9-11]

The present dissertation shows the versatility of high-resolution mass spectrometry for natural product analysis using untargeted (Chapter 2) and targeted (Chapter 3 and 4) metabolic profiling approaches. The combination of efficient chromatographic separation techniques with different ionization methods (e.g. UHPLC-ESI or HPTLC-DESI) were implemented for the exploration of complex plant and fungal extracts with regard to their secondary metabolite profiles. The investigated material comprises Balsaminaceae as representatives of higher plants, and the fungal genera *Sepedonium* and *Cortinarius* belonging to ascomycetes and basidiomycetes, respectively. Thus, the material covers organism form different kingdoms and phyla, characterized by completely different metabolites and matrices.

Extension of MS data analysis in synergy with comprehensive 1D- and 2D-NMR experiments has led to the identification of a pool of natural compounds, some of which are proposed to have novel structures (e.g. fraxetin derivatives or naphthalene-naphthoquinone derivatives, see Chapter 2) and others are for the first time described in the scientific literature (e.g. microsporide A and B, see Chapter 3). Besides, a successful dereplication process based on HRMS and NMR for the different organisms investigated was attained (see Chapters 2, 3, and 4). Herein, the first extensive and systematic phytochemical research for chemically non-investigated *Impatiens* species, *Hydrocera triflora*, and *Sepedonium microspermum* was described. Additionally, antiinfective activity

screenings were conducted for natural crude extracts, as well as for isolated natural products, seeking potential sources of new bioactive natural substances (see Chapter 2 and 3).
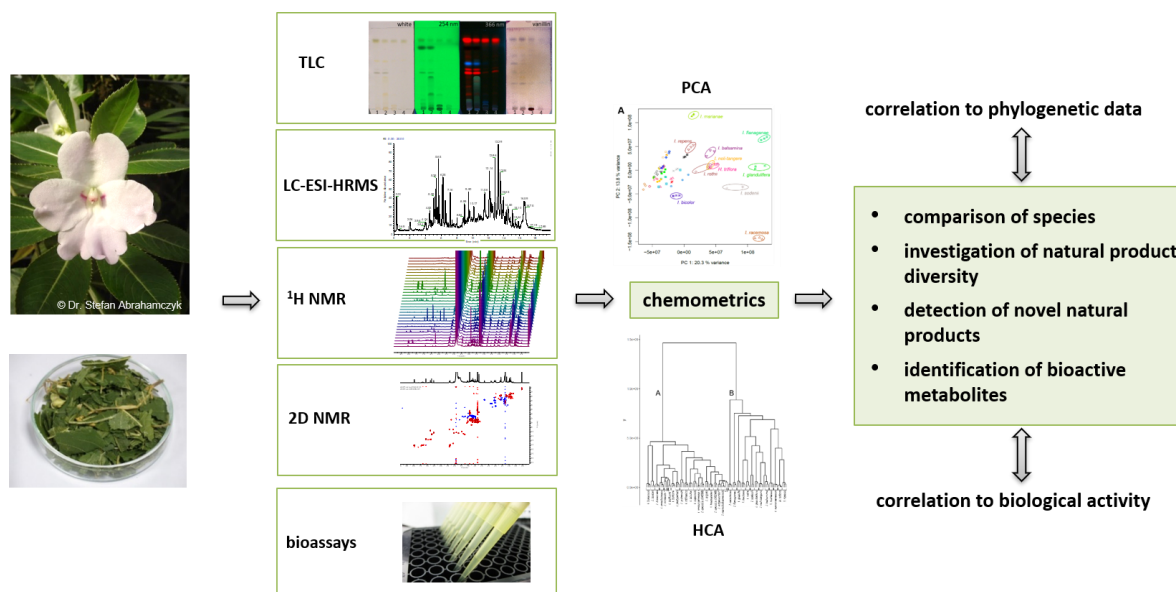
Figure 5.1. Overview of the developed untargeted metabolite profiling workflow in combination with bioactivity screening for the investigation of the methanolic leaf extracts of 31 *Impatiens* species and *Hydrocera triflora* (Balsaminaceae).

Key motivations for implementing untargeted analysis strategies are derived from its potential for the rapid identification of previously known natural products in complex, crude extracts avoiding the unnecessary re-isolation (dereplication), while simultaneously allowing for the detection of potentially new compounds.[12] The use of chemometrics to investigate chemosystematic and chemophenetic relationships based on untargeted analysis can contribute to species differentiation within a genus or taxa, to phylogenetic data correlation and the identification of biomarkers.[13-15] In addition, the acquired analytical information can be used in the frame of molecular network analyses to link different levels of information such as bioactivity data and phylogenetic information to obtain insights in, e.g., biosynthetic pathways. [4,16-18] In spite bioinformatic tools can provide an interconnected view of different features within the analysed samples the use of e.g., principal component analysis, is limited if the samples under investigation show too large differences in their metabolite spectra, as described in Chapter 2.Results shown in this thesis indicate that, for the acquired data set in negative ion LC-ESI-HRMS,the PCA analysis might not be the most suitable method for evaluation. Nevertheless, the PCA based on the complete feature list shows a species differentiation by specific constituents and 11 out of 32 species are separated from the others using PC1 and PC2. To apply multivariate data analysis in the future, the data set should be limited to features that belong for example to a specific substance class (flavonoids, naphthoquinones, and others), to investigate targeted differences between species. This approach might improve correlation with additional data such as phylogenetic information or bioactivity data. Furthermore, in addition to the negative ion LC-ESI-HRMS data, positive ion mode and

NMR data have to be considered for multivariate data analysis in order to enhance the overview of the samples set. In addition, feeding the PCA analysis with data from positive ion mode and NMR might lead to a better metabolites differentiation, while overcoming limitations regarding ionizability.Despite the outstanding benefits of untargeted metabolic profiling approaches, there are some challenges and limitations within these methodologies. The complexity and size of the metabolome, which include a great structural diversity, cannot be easily deciphered using a single analytical technique. On the other hand, the extraction processes constrain the scope of untargeted analyses since compounds that are not extracted or appear in very low abundance might be neglected during detection and identification. Furthermore, metabolites can undergo transformations induced by the used physical and chemical extraction processes resulting in extraction artifacts or degradation.

As extensively described in the literature, the main platforms used for untargeted metabolite profiling are based on LC-HRMS and NMR.[12,19] While NMR allows for quantitative analysis, HRMS can broaden the secondary metabolite detection due to its higher sensitivity. Nevertheless, HRMS faces limitations regarding non-ionizable metabolites and ion suppression effects as highlighted in Chapter 2. Although a wide range of different classes of secondary metabolites of Balsaminaceae were detected in negative ion mode, there were other compound classes, including prevalent naphthoquinones (e.g. 2-methoxy-1,4-naphthoquinone), displaying bad or no ionization efficiencies, and then were better detected by NMR analysis. Beyond detection assistance, the use of NMR as complementary technique can also be exploited to validate MS data interpretation, as shown for phlorizin and 1,2,4-trihydroxynaphthalene-1-*O*-β-glycoside in Chapter 2.

In MS investigations, measuring in positive and negative ion mode can also serves as self-complementing techniques regarding analyte detection, i.e. depending on analyte ionization behavior, it can be detected preferentially either in negative (e.g. flavonoids, hydroxycinnamic acids) or positive (e.g. alkaloids, naphthoquinones) ion mode, which permits to widen the scope of detection. However, this is just effective when an efficient and resolutive chromatographic separation technique is used. In the implementation of the research project described in Chapter 2, it was presented that the acquired data in positive ion mode exhibited low quality due to the accumulation of interfering components on the UHPLC column, provoking low efficiency during the automated recording of the data-dependent MS^2 spectra. In this regard, further investigations have to be conducted to repeat the recording of the data in the positive ion mode in order to assess to which extent this information is complementary with the already processed data.

Processing and interpretation of big data sets requires the use of different statistical analysis and software tools. The processing pipeline presented in Chapter 2 for the evaluation of metabolite profiles from *Impatiens* is based on xcms -an open-source package for LC-MS data processing-and other statistical and data handling tools available from the free platform R. Further enhancement of this platform might include new correlations and network analyzes in order to

combine the available information in a broad context and to understand relationships in the data set. This workflow can be used as a basis for future projects that analyze different complex crude extracts from various organism, and for analysis reproduction to guarantee traceability within the scientific community.

The comprehensive implementation of the optimized workflow described in Chapter 2 allowed for the detection, identification and characterization of known and unknown metabolites from different species of the Balsaminaceae family and constitutes the first report of a systematic untargeted metabolic profiling approach for the exploration of the chemical diversity for a very large set of different species from this family. As previously described in literature, the combination of HRMS and NMR for profiling and fingerprinting is a powerful tool.[20,21] In this regard, in Chapter 2 the first complementary profiling approach based on UHPLC-ESI-HRMS and ¹H-NMR for the analysis of *Impatiens* species and *Hydrocera triflora* is presented. The results show a high diversity of metabolites within the different species. However, and likely because of the chemical diversity, the clustering in the chemometric analysis did not show any specific patterns to associate with phylogenetic data. This suggests that the current taxonomic classification of the family Balsaminaceae in two genera might be insufficient, launching the possibility for new classification paths.

The development and application of targeted analytical approaches is of high interest in different research fields due to their high specificity and reproducibility. The targeted analyses of a defined group or groups of chemically characterized and biochemically annotated metabolites within different matrices can be performed in a qualitative and even quantitative or semi-quantitative manner (if an internal standard is used). Since the analytes of interest are defined, the target-based optimization of sample preparation steps and method parameters reduces the dominance of highly abundant molecules and furthermore prevents analytical artifacts.[22,23] In addition, in natural product discovery it is possible to carry out a targeted, analytically-based isolation of metabolites present in lower abundance in complex mixtures, supported by classical phytochemical isolation procedures.[12] Targeted metabolic profiling is also used in applications for quality control purposes, e.g. species-specific compounds are used as fingerprints of phytoextracts, and for chemosystematic approaches to support species differentiation.

Due to their specificity, targeted analytical methods have a wide range of uses, but there are some challenges and limitations that arise during method development. The high structural diversity and different physicochemical properties of the target molecules require an intensive optimization and adaptation of the sample processing as well as the well-thought-out choice of a suitable analysis method. For example, in the particular case of anthraquinones, unspecific or irreversible interactions with stationary phases commonly used in HPLC applications demand the use of alternative separation techniques, e.g. thin-layer chromatography. In general, for the development

of robust analytical methods, which allow to ensure the accuracy and legitimacy of the produced data, the establishment of reproducible protocols is of paramount importance.
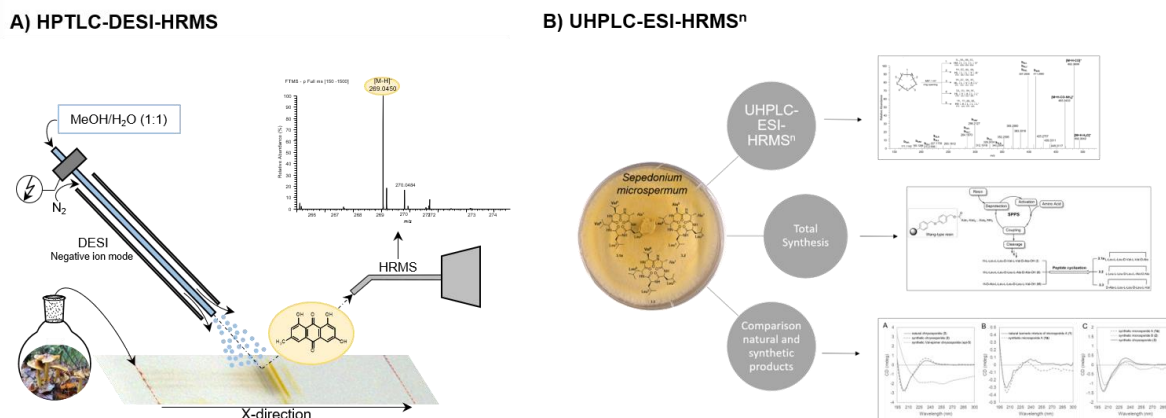
Figure 5.2. Overview of the developed targeted metabolite profiling approaches for the investigation of A) anthraquinones in Chilean dermocyboid Cortinarii[24] and B) cyclic pentapeptides from Sepedonium microspermum Besl.

Analytical techniques mostly used in targeted metabolic profiling include MS, NMR, FT-IR (Fourier-transform infrared spectroscopy), gas and liquid chromatography (GC, HPLC, UHPLC), with LC-ESI-MS and NMR showing the broadest range of applications. LC-ESI-MS can detect a large number of ionizable analytes. However, this technique is limited due to the restricted compatibility with e.g. nonpolar solvents that are necessary to separate more hydrophobic compounds. The development of other ambient ionization principles allowed the detection of surfaces without direct coupling to the MS system (e.g. DESI, MALDI, LAESI) expanding the application of targeted analysis approaches. As highlighted in Chapter 4, the first application of HPLTC-DESI-HRMS for the separation, within a comparably short run time of 4 minutes, and detection of anthraquinone patterns allowed the comparison of different species of the *Cortinarius* fungi as alternative to LC-MS based approaches. However, higher amounts of the extracts are necessary to obtain a sufficient number of ions for the MS detection and further efforts for the optimization of the chromatographic conditions have to be made.

The use of HRMS provides additional information necessary for the determination of the elemental composition of the analyte, as well as for the fragment ions. Misleading information can arise from a superficial interpretation of intensities in MS spectra, suggesting the identification of majority components in the mixture, while in fact quantitative NMR investigations show that intensive ions in MS sometimes could be associated with minority secondary metabolites in the crude samples. This kind of behavior is frequently associated to peptide substrates, which in general present good ionization performance, although they might be present as minor components in a sample (see Chapter 3, Figure 3.1)

Q-TOF, LTQ-Orbitrap, and FTICR analyzers are suitable either to perform global or targeted profiling studies, contributing with their high sensitivity, mass resolving power and mass accuracy.

123

From all of them, LTQ-Orbitrap is the only type of analyzer that allows for MS^n experiments. This characteristic is extremely useful for metabolite identification because it provides a step-wise fragmentation pattern in a multilevel scale for a chosen fragment ion.

Therefore, LTQ-Orbitrap was the system selected for the detailed structural investigation of different compound classes handled along the projects developed in the thesis. UHPLC-ESI-HRMS[2] experiments in Chapter 2 allowed the annotation and partial identification of metabolites belonging to different compound classes (carboxylic acids, hydroxycinnamic acid derivatives, flavonoid glycosides, coumarins, naphthoquinones, naphthalenes, and glycolipids) as well as the proposal of core structures of new naphthalene derivatives. However, some peaks remain unknown and need to be investigated further. UHPLC-ESI-HRMS[3] experiments were used to confirm the structure of the dihydrochalcone derivatives isolated from *Impatiens racemosa* DC. In a similar way, UHPLC-ESI-HRMS[n] (MS^2-MS^4) investigations were applied in Chapter 3 for the determination of the primary sequence of two new cyclic pentapeptides (see example in Figure 5.3.) isolated from *Sepedonium micospermum* as well as for structure verification of the known chrysosporide. In Chapter 4 HPTLC-DESI-HRMS was used to identify anthraquinones by their elemental composition and to demonstrate the implementation of fragmentation experiments ($HRMS^2$) for anthraquinones on HPTLC surfaces.

In conclusion, valuable contributions to the method development and application of high-resolution mass spectrometry-based analysis for the characterization of natural products from plant and fungal crude extracts were achieved. The use of sensitive analytical methods allows resource-saving work and enable a precise assessment of existing substance classes in the investigated matrices and their potential as new active ingredients in combination with bioactivity studies. This was further supported by other modern analytical techniques like NMR and CD spectroscopy as well as organic-chemical total synthesis in combination with classical natural product isolation. Thus, several areas and aspects of modern natural product research are covered in this thesis. The results additionally suggest that *Impatiens* species are a promising source of novel bioactive natural products, although further investigation should be performed in order to decipher all the structure novelty of secondary metabolites still not assessed in this family. Furthermore, it was shown, that Desorption Electrospray Ionization coupled to planar chromatography in combination with high-resolution mass spectrometry is a powerful tool for the comparison of different complex matrices and can be used in the future for chemophenetic studies of different species. The achievements obtained in this thesis are expected to build the base for future studies using mass spectrometry as a key technique in natural product discovery.

5.1 References

[1] Wolfender, J.L.; Marti, G.; Queiroz, E.F. Advances in techniques for profiling crude extracts and for the rapid identification of natural products: dereplication, quality control and metabolomics. *Curr. Org. Chem.* **2010**, *14*, 1808-1832.

[2] Kleigrewe, K.; Almaliti, J.; Tian, I.Y.; Kinnel, R.B.; Korobeynikov, A.; Monroe, E.A.; Duggan, B.M.; Di Marzo, V.; Sherman, D.H.; Dorrestein, P.C., et al. Combining mass spectrometric metabolic profiling with genomic analysis: a powerful approach for discovering natural products from cyanobacteria. *J. Nat. Prod.* **2015**, *78*, 1671-1682.

[3] Wolfender, J.L.; Marti, G.; Thomas, A.; Bertrand, S. Current approaches and challenges for the metabolite profiling of complex natural extracts. *J. Chromatogr. A* **2015**, *1382*, 136-164.

[4] Olivon, F.; Allard, P.-M.; Koval, A.; Righi, D.; Genta-Jouve, G.; Neyts, J.; Apel, C.; Pannecouque, C.; Nothias, L.-F.; Cachet, X., et al. Bioactive natural products prioritization using massive multi-informational molecular metworks. *ACS Chem. Biol.* **2017**, *12*, 2644-2651.

[5] Allard, P.-M.; Bisson, J.; Azzollini, A.; Pauli, G.F.; Cordell, G.A.; Wolfender, J.-L. Pharmacognosy in the digital era: shifting to contextualized metabolomics. *Curr. Opin. Biotechnol.* **2018**, *54*, 57-64.

[6] Wolfender, J.L.; Nuzillard, J.M.; van der Hooft, J.J.J.; Renault, J.H.; Bertrand, S. Accelerating metabolite identification in natural product research: toward an ideal combination of liquid chromatography-high-resolution tandem mass spectrometry and NMR profiling, *in silico* databases, and chemometrics. *Anal. Chem.* **2019**, *91*, 704-742.

[7] Lautié, E.; Russo, O.; Ducrot, P.; Boutin, J.A. Unraveling plant natural chemical diversity for drug discovery purposes. **2020**, *11*, 1-37

[8] Sorokina, M.; Steinbeck, C. Review on natural products databases: where to find data in 2020. *J. Cheminformatics* **2020**, *12*, 1-51.

[9] Schüler, J.-A.; Neumann, S.; Müller-Hannemann, M.; Brandt, W. ChemFrag: Chemically meaningful annotation of fragment ion mass spectra. *J. Mass Spectrom.* **2018**, *53*, 1104-1115.

[10] Kind, T.; Tsugawa, H.; Cajka, T.; Ma, Y.; Lai, Z.; Mehta, S.S.; Wohlgemuth, G.; Barupal, D.K.; Showalter, M.R.; Arita, M., et al. Identification of small molecules using accurate mass MS/MS search. *Mass Spectrom. Rev.* **2018**, *37*, 513-532.

[11] Dührkop, K.; Nothias, L.-F.; Fleischauer, M.; Reher, R.; Ludwig, M.; Hoffmann, M.A.; Petras, D.; Gerwick, W.H.; Rousu, J.; Dorrestein, P.C., et al. Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. *Nat. Biotechnol.* **2020**, *39*, 462-471.

[12] 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.

[13] Wolfender, J.-L.; Glauser, G.; Boccard, J.; Rudaz, S. MS-based Plant Metabolomic Approaches for Biomarker Discovery. *Nat. Prod. Commun.* **2009**, *4*, 1417-1430.

[14] Rosato, A.; Tenori, L.; Cascante, M.; De Atauri Carulla, P.R.; Martins Dos Santos, V.A.P.; Saccenti, E. From correlation to causation: analysis of metabolomics data using systems biology approaches. *Metabolomics* **2018**, *14*, 1-20.

[15] Zidorn, C. Plant chemophenetics − A new term for plant chemosystematics/plant chemotaxonomy in the macro-molecular era. *Phytochemistry* **2019**, *163*, 147-148.

[16] Boccard, J.; Rudaz, S. Harnessing the complexity of metabolomic data with chemometrics. *J. Chemom.* **2014**, *28*, 1-9.

[17] Fox Ramos, A.E.; Evanno, L.; Poupon, E.; Champy, P.; Beniddir, M.A. Natural products targeting strategies involving molecular networking: different manners, one goal. *Nat. Prod. Rep.* **2019**, *36*, 960-980.

[18] Nair, S.K.; Jez, J.M. Natural product biosynthesis: What's next? An introduction to the JBC Reviews Thematic Series. *J. Biol. Chem.* **2020**, *295*, 335-336.

[19] Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X. Modern analytical techniques in metabolomics analysis. *Analyst* **2012**, *137*, 293-300.

[20] Porzel, A.; Farag, M.A.; Mülbradt, J.; Wessjohann, L.A. Metabolite profiling and fingerprinting of *Hypericum* species: a comparison of MS and NMR metabolomics. *Metabolomics* **2014**, *10*, 574-588.

[21] Marshall, D.D.; Powers, R. Beyond the paradigm: Combining mass spectrometry and nuclear magnetic resonance for metabolomics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *100*, 1-16.

[22]   Griffiths, W.J.; Koal, T.; Wang, Y.; Kohl, M.; Enot, D.P.; Deigner, H.-P. Targeted Metabolomics for Biomarker Discovery. *Angew. Chem. Int. Ed.* **2010**, *49*, 5426-5445.

[23]   Roberts, L.D.; Souza, A.L.; Gerszten, R.E.; Clish, C.B. Targeted Metabolomics. *Curr. Protoc. Mol. Biol.* **2012**, *98*, 30.2.1-30.2.24.

[24]   **Laub, A.**; Sendatzki, A.-K.; Palfner, G.; Wessjohann, L.A.; Schmidt, J.; Arnold, N. HPTLC-DESI-HRMS-Based Profiling of Anthraquinones in Complex Mixtures—A Proof-of-Concept Study Using Crude Extracts of Chilean Mushrooms. *Foods* **2020**, *9*, 156.

Appendix

A. Appendix Chapter 2

Table A.1. Overview of reference materials used for TLC analysis (Figure A1).

No.	R_f value	Compound	Compound class	Color reaction with vanillin-sulfuric acid	Molecular weight [g/moL]	Company
1	0.76	7-hydroxycoumarin	coumarin	n.d.	162.14	Alfa Aesar (98%)
2	0.40	6, 7-dihydroxycoumarin	coumarin	n.d.	178.14	Aldrich (98%)
3	0.09	catechin	flavan-3-ol	red	290.26	Sigma
4	0.00	naringin	flavanone	red	580.541	Sigma (>90%)
5	0.66	naringenin	flavanone	pink	272.257	Aldrich
6	0.74	kaempferide	flavonoid	yellow	300.26	Extrasynthese(>99%)
7	0.45	kaempferol	flavonoid	yellow	286.23	Sigma (>97%)
8	0.25	quercetin	flavonoid	yellow	302.236	Sigma (>95%)
9	0.00	rutin	flavonoid	yellow	610.52	Sigma
10	-	*I. flanaganae*	methanolic leaf extract	-	-	
11	-	*I. ethiopica*	methanolic leaf extract	-	-	
12	-	*I. racemosa*	methanolic leaf extract	-	-	
13	1.00	2-methoxy-1,4-naphthoquinone	naphthoquinone	n.d.	188.18	Aldrich (98%)
14	0.20	2-hydroxy-1,4-naphthoquinone	naphthoquinone	yellow	174.15	
15	0.22	phloroglucinol	phenol	orange-red	126.11	Fluka (>95%)
16	0.36	ferulic acid	phenolic acid	light blue	194.18	Sigma Aldrich (99%)
17	0.12	caffeic acid	phenolic acid	light blue	180.16	Sigma
18	0.00	chlorogenic acid	phenolic acid	n.d.	354.31	Aldrich (>95%)
19	0.03	phlorizin	dihydrochalcone	orange-red	436.4	TCI (>97%)
20	0.86	stigmasterin	sterol	dark blue-purple	412.69	Merck [>95%]
21	1.00	xanthone	xanthone	n.d.	196.19	Aldrich (97%)
22	-	Pooled quality control sample containing leaf material of all species (LAA095_QC)	methanolic leaf extract	-	-	-

Table A.2. Overview of ions used for determination of the linear dynamic range in negative ion mode LC-ESI-HRMS based on the measurement of one replicate of a quality control sample containing leaf material of all species (LAA093_LDR_QC_L_n_3b).

[M-H]⁻	retention time [min]	intensity
749.3265	6.7	1,89*e3
447.1283	7	6.9*e3
303.0499	4.82	3.77*e3
475.1235	6.92	2.31*e4
510.0881	5.37	7.73*e4
529.3004	8.03	4.64*e4
301.0347	6.05	4.03*e5
593.1501	5.29	9.77*e5
721.3633	9	9.43*e5
435.1291	5.68	2.97*e6
981.5786	12.9	3.8*e6
1057.5208	7.16	1.85*e6

Table A.3. Metabolites annotated in the methanolic crude leaf extract of *Impatiens racemosa* DC. by UHPLC-ESI-HRMS^2 in negative ion mode (CID, NCE 35%)

No.	t_R (min)	Compound	[M-H]- measured	Error (ppm)	Elemental Composition	UV (nm)	MS^2 product ions *m/z* (rel. intensity [%])	Fraction
R1	4.61	Coumaric acid glycoside	325.0919	-3.1	$C_{15}H_{17}O_8^-$	216, 286	163.0398 ($C_9H_7O_3^-$, **100**), 119.0501($C_8H_7O^-$, 8)	1
R2	4.63	Scopolin$^\perp$	399.0916	-2.7	$C_{17}H_{19}O_{11}^-$	205, 221, 290, 338	353.0871 ($C_{16}H_{17}O_9^-$, 3), 191.0347 ($C_{10}H_7O_4^-$, **100**), 176.0113 ($C_9H_4O_4^-$, 8)	2
R3	4.88	Coniferin$^\perp$	429.1022	-3.8	$C_{18}H_{21}O_{12}^-$	215, 294	383.0970 ($C_{17}H_{19}O_{10}^-$, 5), 221.0450 ($C_{11}H_9O_5^-$, **100**), 206.0215 ($C_{10}H_6O_5^-$, 6)	2
R4	4.97	Coumarin-*O*-glycoside$^\perp$	447.1494	-3.0	$C_{19}H_{27}O_{12}^-$	215	401.1440 ($C_{18}H_{25}O_{10}^-$, 5)	2
R5	4.99	Kaempferol-*O*-di-hexoside	609.1443	-3.0	$C_{27}H_{29}O_{16}^-$	-	447.0923 ($C_{21}H_{19}O_{11}^-$, 47), 285.0400 ($C_{15}H_9O_6^-$, **100**), 284.0320 ($C_{15}H_8O_6^-$, 4)	2
R6	5.07	unidentified	351.1288	-2.4	$C_{14}H_{23}O_{10}^-$	218, 325	291.1079 ($C_{12}H_{19}O_8^-$, 4), 249.0611 ($C_9H_{13}O_8^-$, **100**), 231.0505 ($C_9H_{11}O_7^-$, 12), 113.0243 ($C_5H_5O_3^-$, 11)	1
R7	5.09	Caffeoylmalic acid	295.0452	-2.6	$C_{13}H_{11}O_8^-$	-	179.0348 ($C_9H_7O_4^-$, **100**), 135.0452 ($C_8H_7O_2^-$, 6), 133.0142 ($C_4H_5O_5^-$, 81), 115.0037 ($C_4H_3O_4^-$, 3)	1
R8	5.21	unidentified	505.1702	-2.7	$C_{25}H_{29}O_{11}^-$	219, 273	487.1605 ($C_{25}H_{27}O_{10}^-$, 52), 475.1605 ($C_{24}H_{27}O_{10}^-$, 31), 343.1183 ($C_{19}H_{19}O_6^-$, 24), 325.1078 ($C_{19}H_{17}O_5^-$, **100**), 313.1079 ($C_{18}H_{17}O_5^-$, 9), 295.0973 ($C_{18}H_{15}O_4^-$, 2)	2
R9	5.29	Phloretin-*O*-tri-glycoside$^\perp$	805.2391	-2.1	$C_{34}H_{45}O_{22}^-$	-	759.2335 ($C_{33}H_{43}O_{20}^-$, **100**), 597.1823 ($C_{27}H_{33}O_{15}^-$, 53)	2
R10	5.33	unidentified$^\perp$	461.1656	-1.3	$C_{20}H_{19}O_{12}^-$	-	415.1607 ($C_{19}H_{27}O_{10}^-$, **100**)	2
R11	5.42	Coumaroyl-malic acid	279.0504	-2.2	$C_{13}H_{11}O_7^-$	221, 314	163.0399 ($C_9H_7O_3^-$, **100**)	1
R12	5.49	Quercetin-*O*-hexoside	463.0873	-1.8	$C_{21}H_{19}O_{12}^-$	203, 256, 354	301.0351 ($C_{15}H_9O_7^-$, **100**), 300.0277 ($C_{15}H_8O_7^-$, 45), 151.0039 ($C_7H_3O_4^-$, 1)	3
R13	5.51	Coumaroyl-malic acid	279.0503	-2.5	$C_{13}H_{11}O_7^-$	218, 310	163.0402 ($C_9H_7O_3^-$, **100**), 133.0144 ($C_4H_5O_5^-$, 61), 119.0504 ($C_8H_7O^-$, 6), 115.0038 ($C_4H_3O_4^-$, 2)	1
R14	5.54	Quercetin-*O*-hexoside	463.0874	-1.7	$C_{21}H_{19}O_{12}^-$	203, 256, 354	301.0353 ($C_{15}H_9O_7^-$, **100**), 300.0279 ($C_{15}H_8O_7^-$, 34), 151.0039 ($C_4H_3O_4^-$, 1)	3
R15	5.62	Feruloyl malate	309.0606	-3.1	$C_{14}H_{13}O_8^-$	219, 327	193.0502 ($C_{10}H_9O_4^-$, **100**)	1
R16	5.66	Phloretin-*O*-di-glycoside	597.1824	-0.1	$C_{27}H_{33}O_{15}^-$	221, 285	449.1299 ($C_{18}H_{25}O_{13}^-$, **100**), 329.0877 ($C_{14}H_{17}O_9^-$, 32), 273.0768 ($C_{15}H_{13}O_5^-$, 95)	3
R17	5.68	Feruloyl malate	309.0606	-2.9	$C_{14}H_{13}O_8^-$	219, 322	193.0502 ($C_{10}H_9O_4^-$, **100**), 133.0142 ($C_4H_5O_5^-$, 12), 115.0036 ($C_4H_3O_4^-$, 2)	1
R18	5.69	Kaempferol-*O*-glycoside	447.0927	-1.3	$C_{21}H_{19}O_{11}^-$	-	327.0507 ($C_{17}H_{11}O_7^-$, 14), 285.0401 ($C_{15}H_9O_6^-$, 54), 284.0327 ($C_{15}H_8O_6^-$, **100**)	4
R19	5.77	Kaempferol-*O*-glycoside	447.0935	-0.7	$C_{21}H_{19}O_{11}^-$	-	327.0507 ($C_{17}H_{11}O_7^-$, 18), 285.0401 ($C_{15}H_9O_6^-$, 72), 284.0327 ($C_{15}H_8O_6^-$, **100**)	4
R20	5.94	Kaempferol-*O*-acetyl-hexoside	489.1030	-1.7	$C_{23}H_{21}O_{12}^-$	-	285.0400 ($C_{15}H_9O_6^-$, **100**)	extract
R21	6.07	Phlorizin*	435.1295	-0.3	$C_{21}H_{23}O_{10}^-$	195, 224, 287	273.0766 ($C_{15}H_{13}O_5^-$, **100**), 167.0350 ($C_8H_7O_4^-$, 2)	5
R22	6.20	Pinocembrin dihydrochalcone-*C*-hexoside-*O*-hexoside	581.1871	-0.9	$C_{27}H_{33}O_{14}^-$		449.1300 ($C_{18}H_{25}O_{13}^-$, **100**), 329.0877 ($C_{14}H_{17}O_9^-$, 34), 257.0820 ($C_{15}H_{13}O_4^-$, 95)	3

Table A.3. (continue)

R23	6.43	Caffeoyl dihydrochalcone	289.0715	-0.8	$C_{15}H_{13}O_6^-$	-	271.0613 ($C_{15}H_{11}O_5^-$, 45), 245.0820 ($C_{14}H_{13}O_4^-$, 15), 167.0351 ($C_8H_7O_4^-$, **100**), 125.0247 ($C_6H_5O_3^-$, 34)	extract
R24	6.57	Pinocembrin-*C*-hexosyl-*O*-pentoside	549.1605	-1.5	$C_{26}H_{29}O_{13}^-$	-	429.1191 ($C_{22}H_{21}O_9^-$, 21), 387.1086 ($C_{20}H_{19}O_8^-$, 10), 357.0980 ($C_{19}H_{17}O_7^-$, 4), 297.0768 ($C_{17}H_{13}O_5^-$, 11), 279.0664 ($C_{17}H_{11}O_4^-$, 4), 255.0663 ($C_{15}H_{11}O_4^-$, **100**)	extract
R25	6.83	Pinocembrin-dihydrochalcone-*O*-glycoside	419.1345	-0.6	$C_{21}H_{23}O_9^-$	192, 207, 287	257.0820 ($C_{15}H_{13}O_4^-$, **100**)	6
R26	6.89	Phloretin*	273.0770	0.4	$C_{15}H_{13}O_5^-$	192, 223, 286	179.0352 ($C_9H_7O_4^-$, 1), 167.0351 ($C_8H_7O_4^-$, **100**), 125.0247 ($C_6H_5O_3^-$, 4), 123.0454 ($C_7H_7O_2^-$, 3)	7

$^\perp$ Formic acid adduct, * confirmed with reference compound
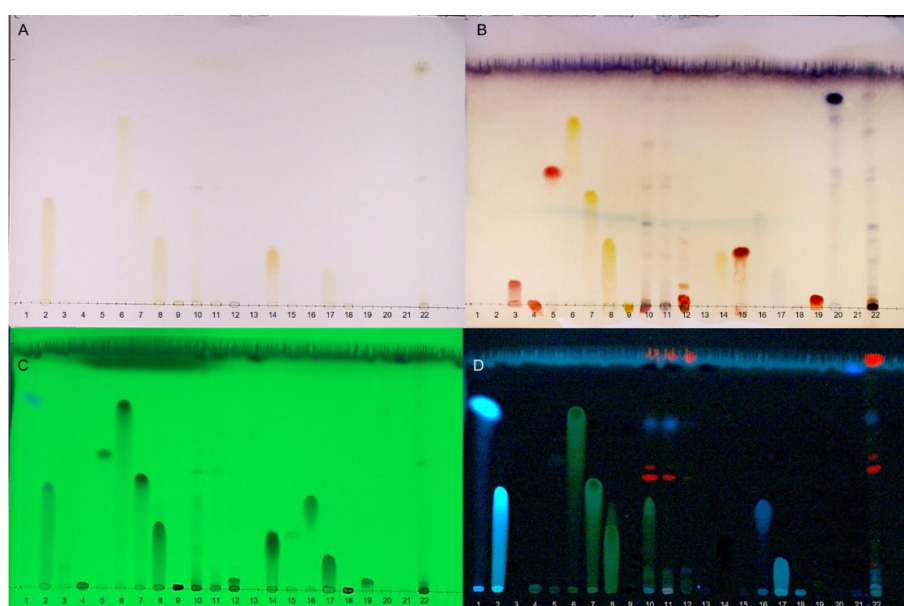
Figure A.1. Thin-layer chromatograms of reference compounds (listed in Table A1) (concentration reference: 10 mM, extracts: 10 mg/mL, mobile phase: chloroform: methanol (9:1, *v/v*)), visualized under day light (A), derivatized using vanillin-sulfuric acid reagent (B), and under UV-light (254 nm (B), 366 nm (C)).
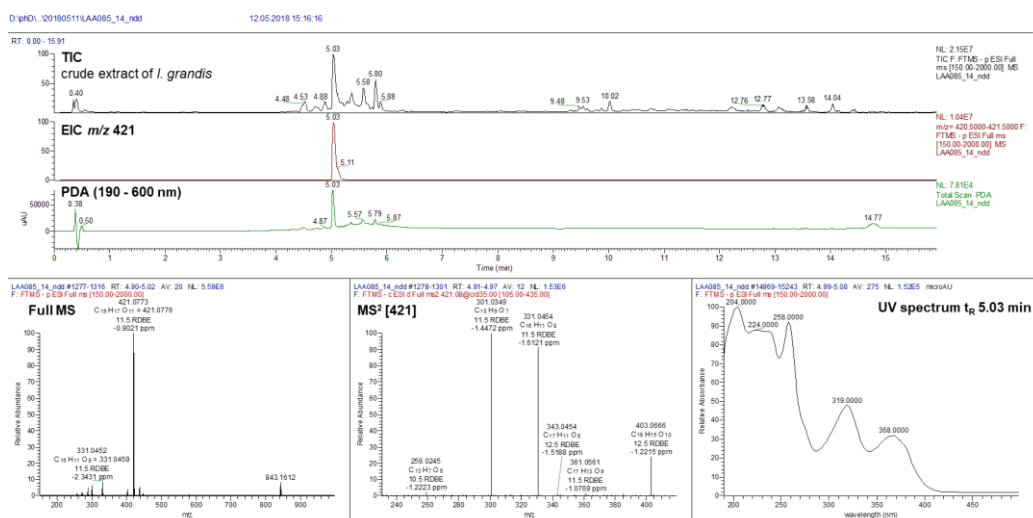
Figure A.2. Total ion chromatogram (TIC) in negative ion mode of crude methanolic leaf extract of *I. grandis* and the (-)-ESI-HRMS^2 and UV spectrum of its constituent mangiferin ([M-H]$^-$ *m/z* 421).
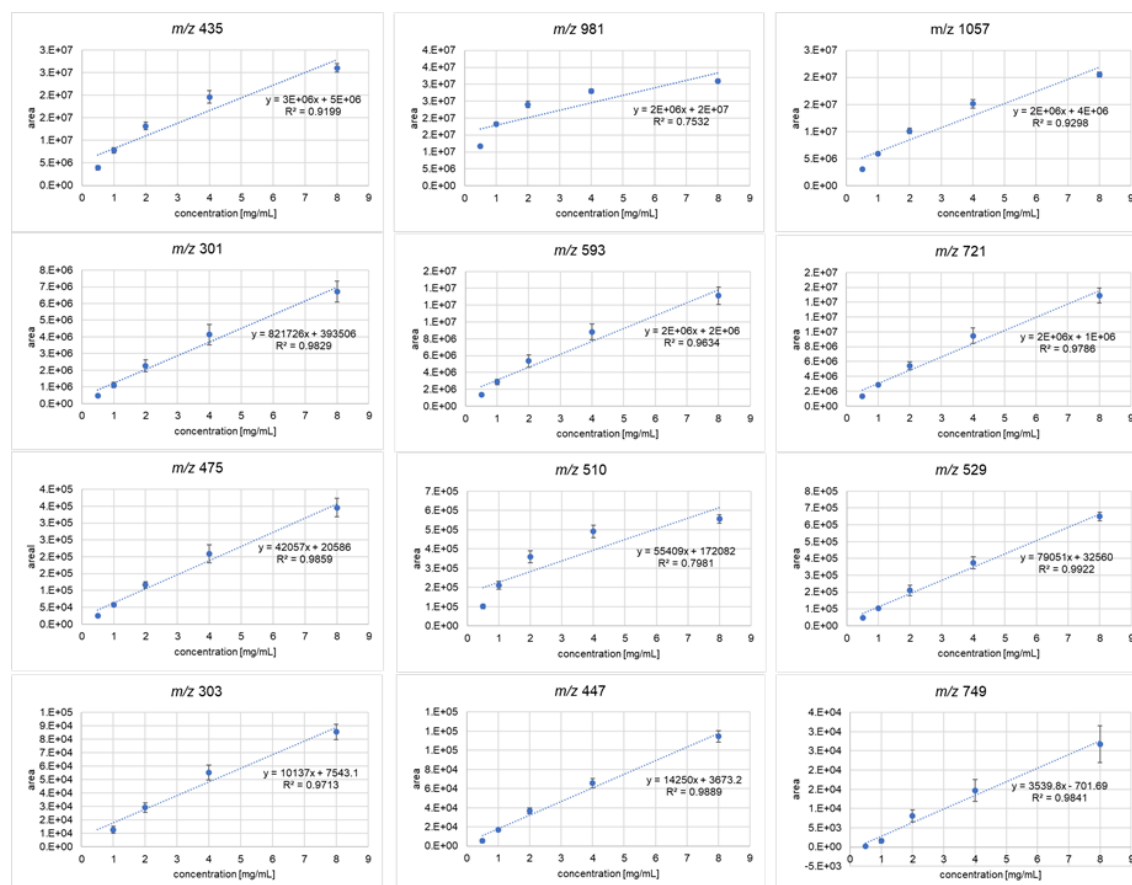
Figure A.3. Evaluation of linear dynamic range showing the concentration (mg dried plant powder/1 mL solvent, x-axis) and the corresponding peak area of the peaks listed in Table A.2 (y-axis, integration performed with Quanbrowser implemented in Thermo Xcalibur 2.2 SP1.48, Thermo Fisher Scientific).
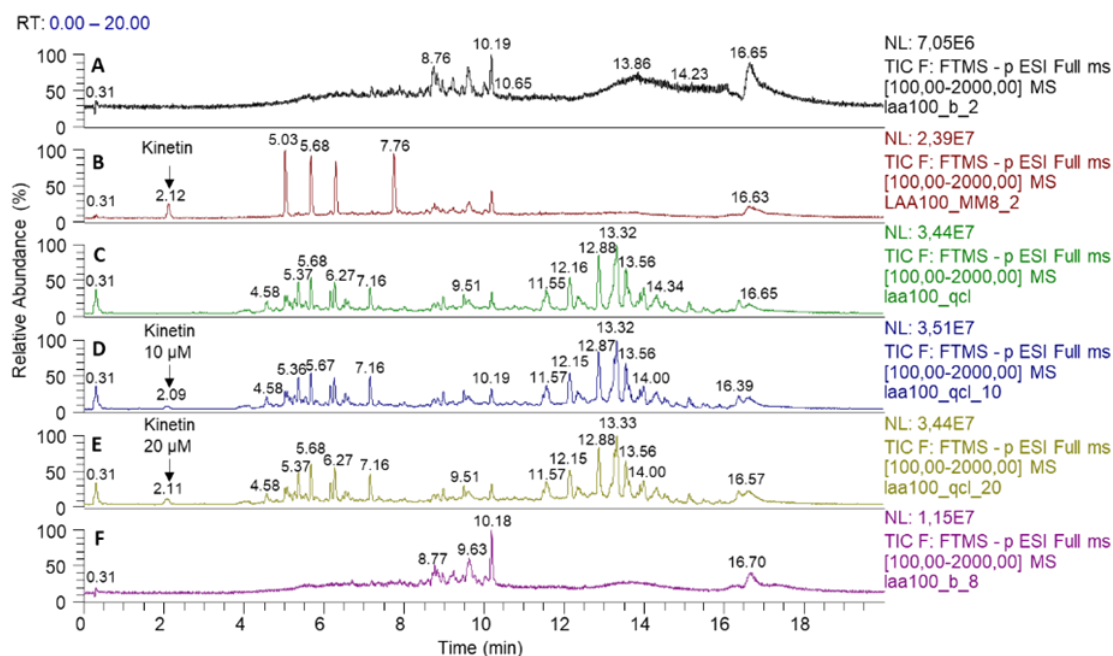
Figure A.4. Comparison of total ion chromatograms (TIC) of A) methanol blank before measurements; B) reference mix MM8 (concentration kinetin: 10 µM); C) Leaf QC sample extracted with methanol; D) Leaf QC sample extracted with methanol containing 10 µM kinetin, E) Leaf QC sample extracted with methanol containing 20 µM kinetin, F) TIC methanol blank after measurements.

v
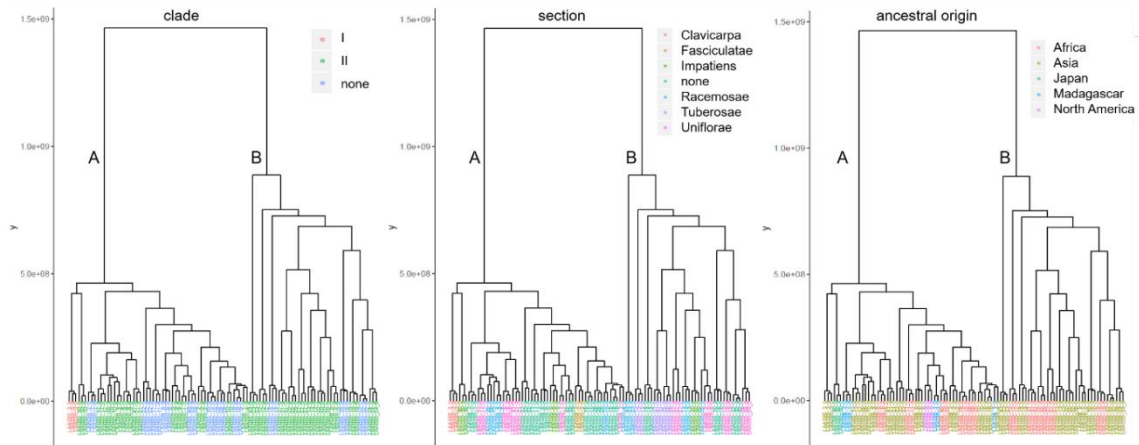
Figure A.5. HCA colored by metadata based on the phylogenetic background, section and ancestral origin of the different *Impatiens* species (data used see Table 2.1).
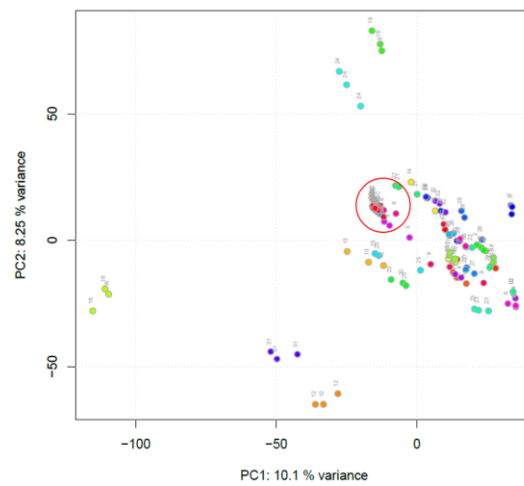
Figure A.6. PCA scores plot (components 1 and 2) including quality control samples.
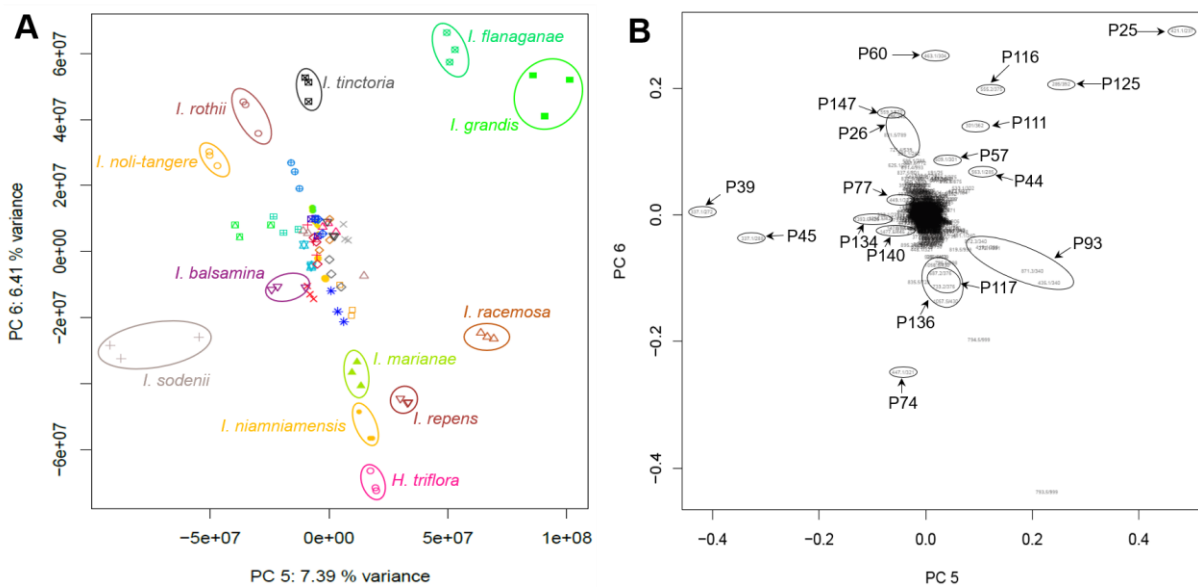
Figure A.7. A) PCA scores plot (components 5 and 6). Classes are defined according to the species and B) PCA loadings plot (components 5 and 6). Shown features correspond to measured mass-over-charge ratios (*m/z*) and retention times in seconds and are numbered according to Table 2.2.
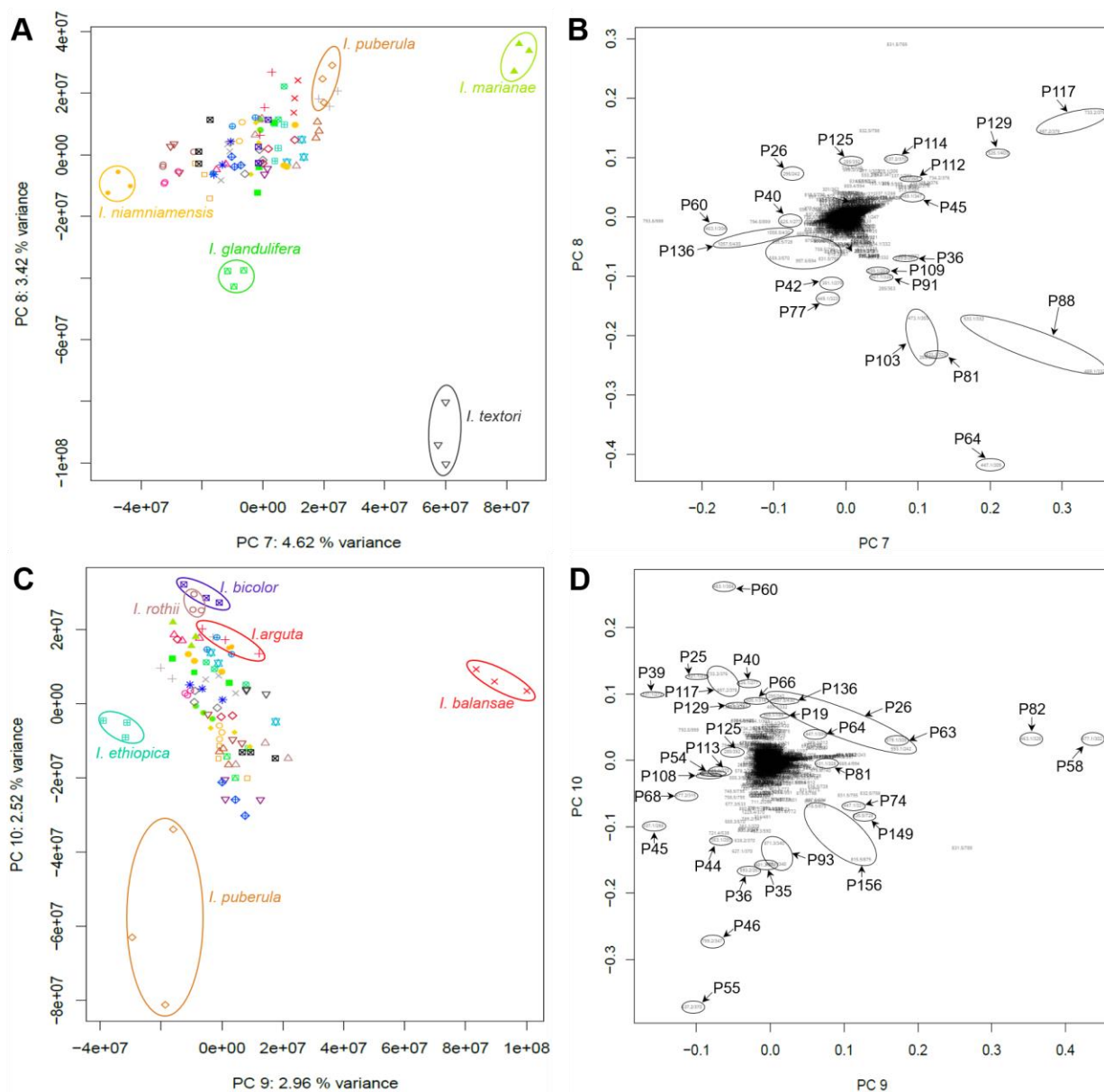
Figure A.8. A) PCA scores plot components 7 and 8, B) PCA loadings plot components 7 and 8, C) PCA scores plot components 9 and 10, D) PCA loadings plot components 9 and 10. Classes are defined according to the species and. Shown features correspond to measured mass-over-charge ratios (*m/z*) and retention times in seconds and are numbered according to Table 2.2.
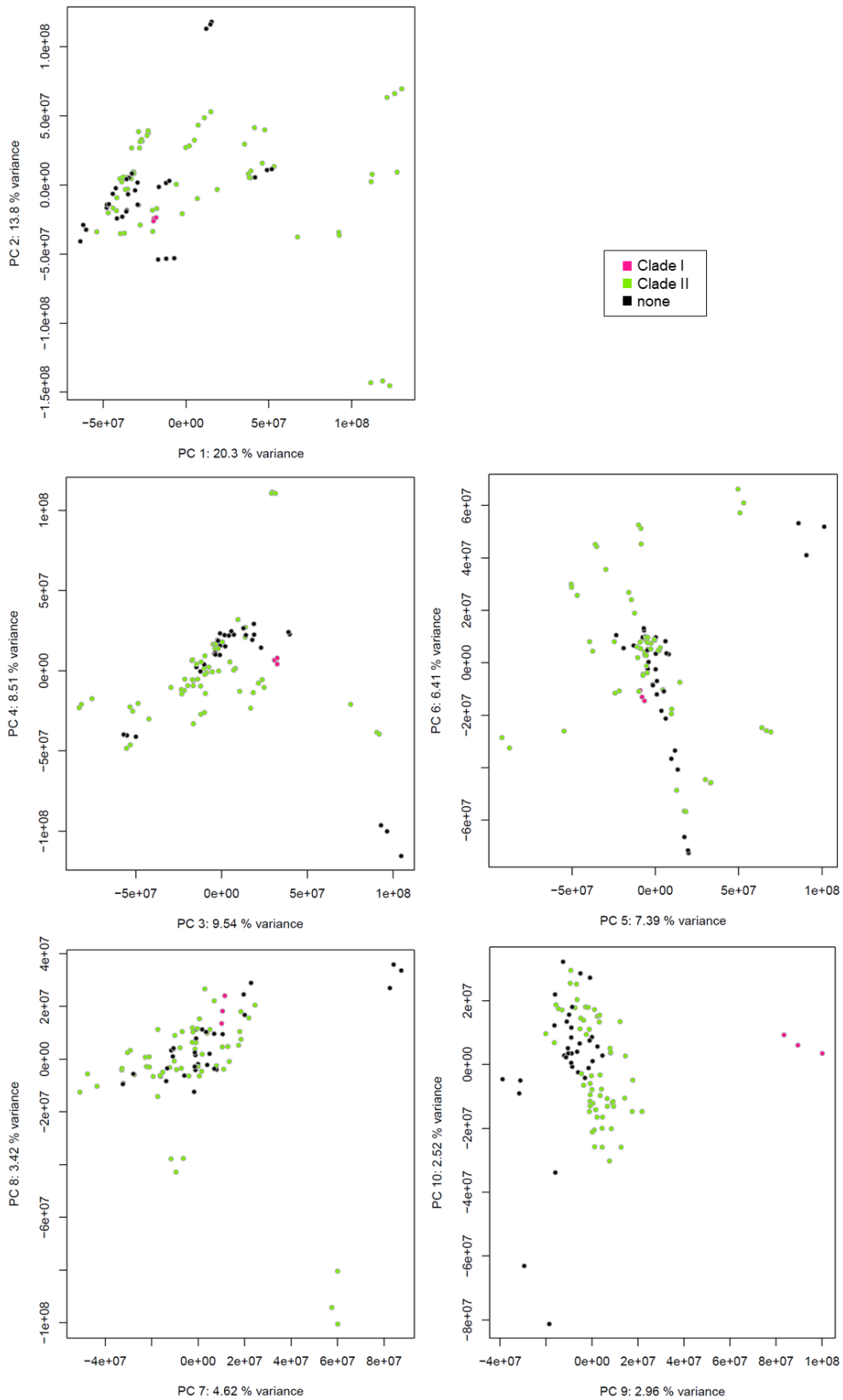
Figure A.9. PCA colored by metadata based on the phylogenetic background (clade) of the different *Impatiens* species (Table 2.1).
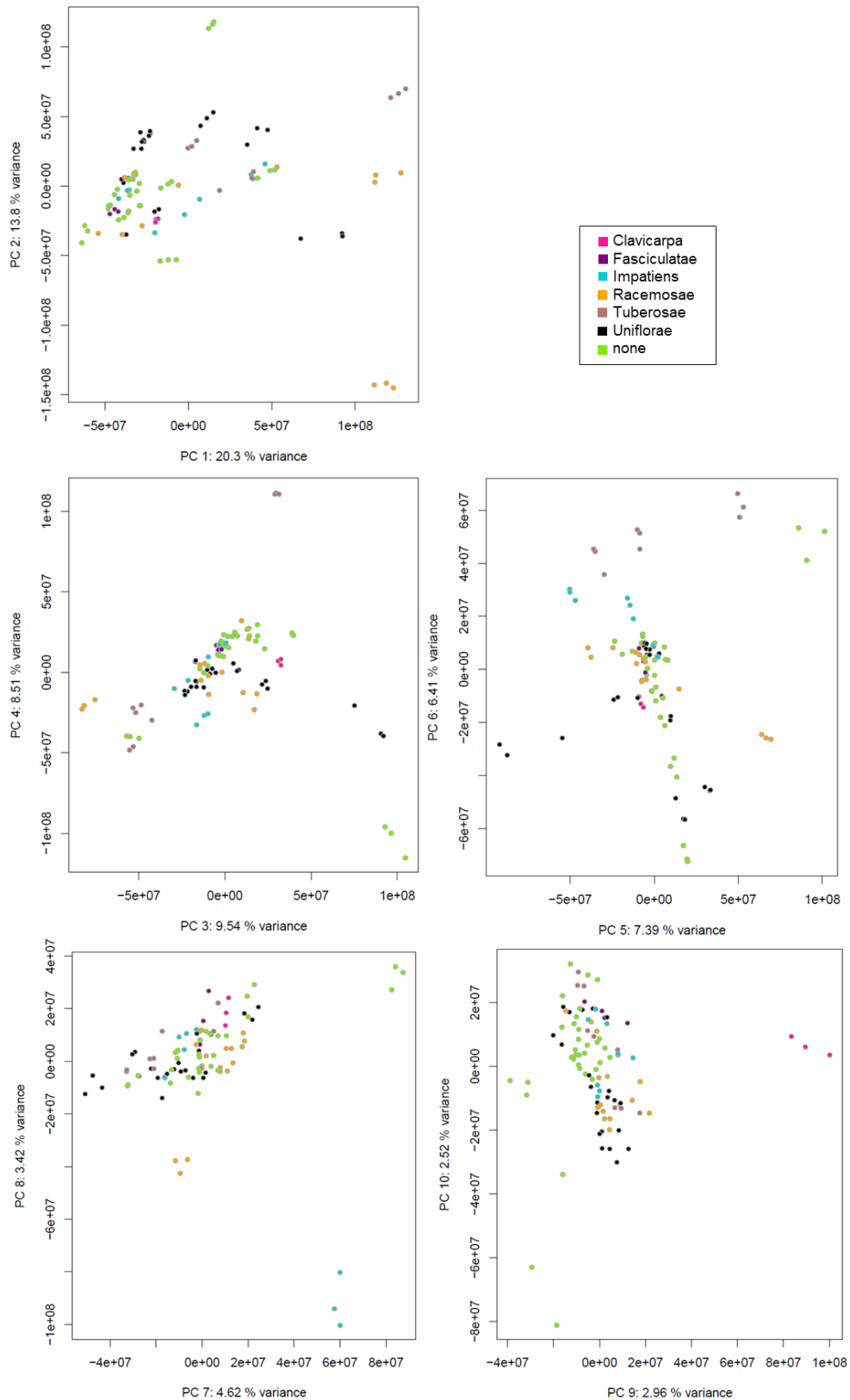
Figure A.10. PCA colored by metadata based on the sections of the different *Impatiens* species (Table 2.1).
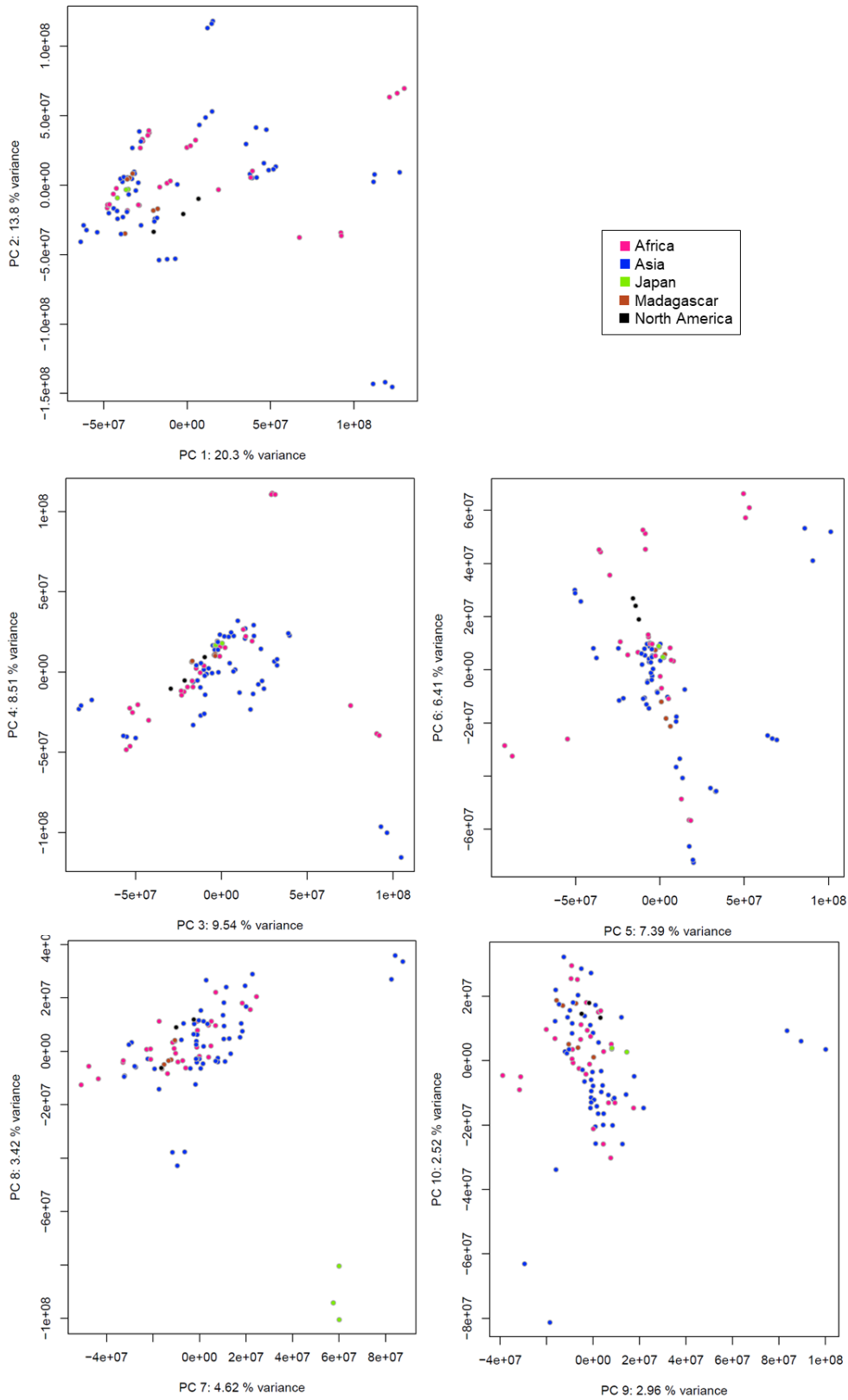
Figure A.11. PCA colored by metadata based on the ancestral origin of the different *Impatiens* species (Table 2.1).
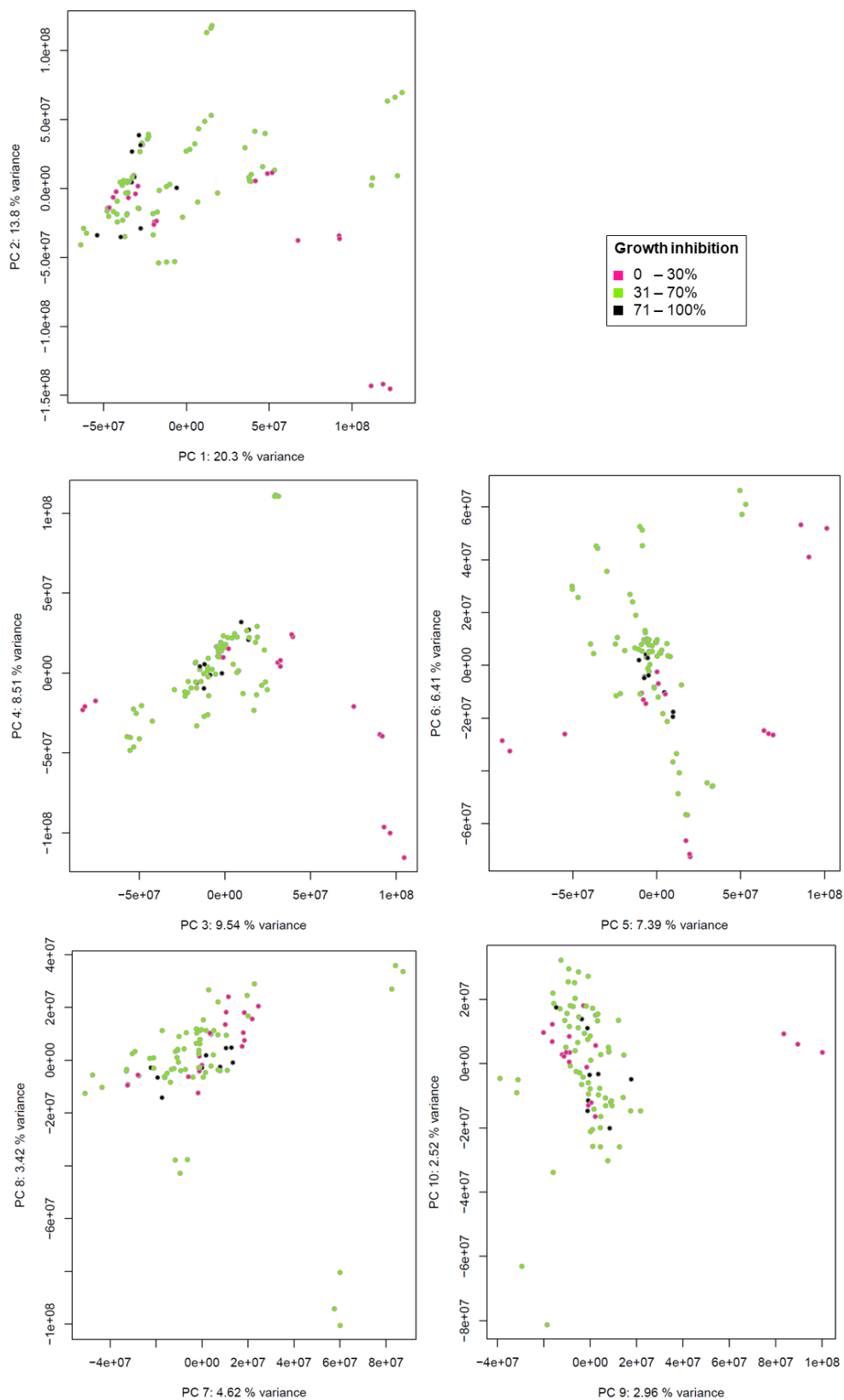
Figure A.12. PCA colored by metadata based on the biological activities of the methanolic leaf extracts against *A. fischeri*.
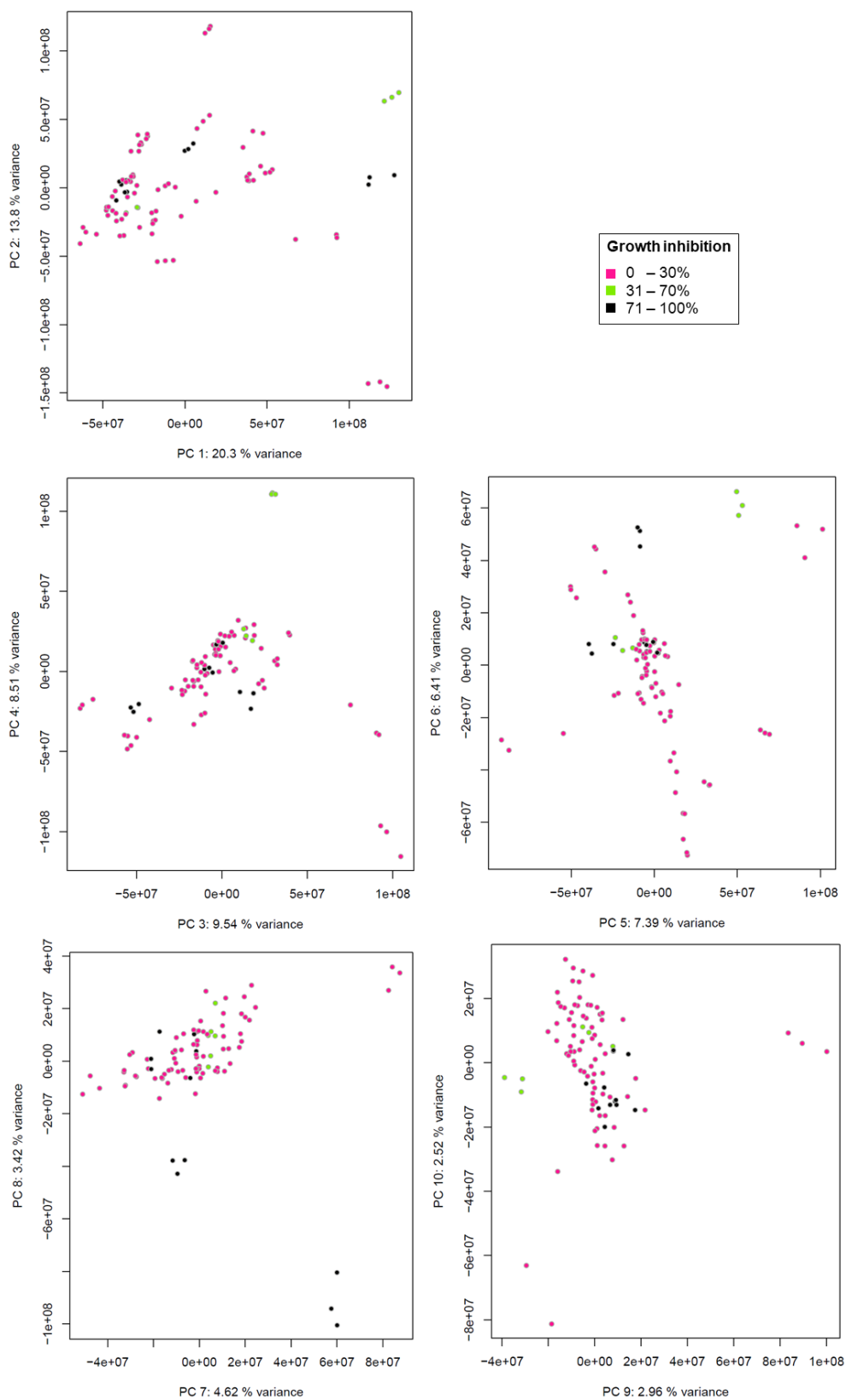
Figure A.13. PCA colored by metadata based on the biological activities of the methanolic leaf extracts against *B. subtilis*.
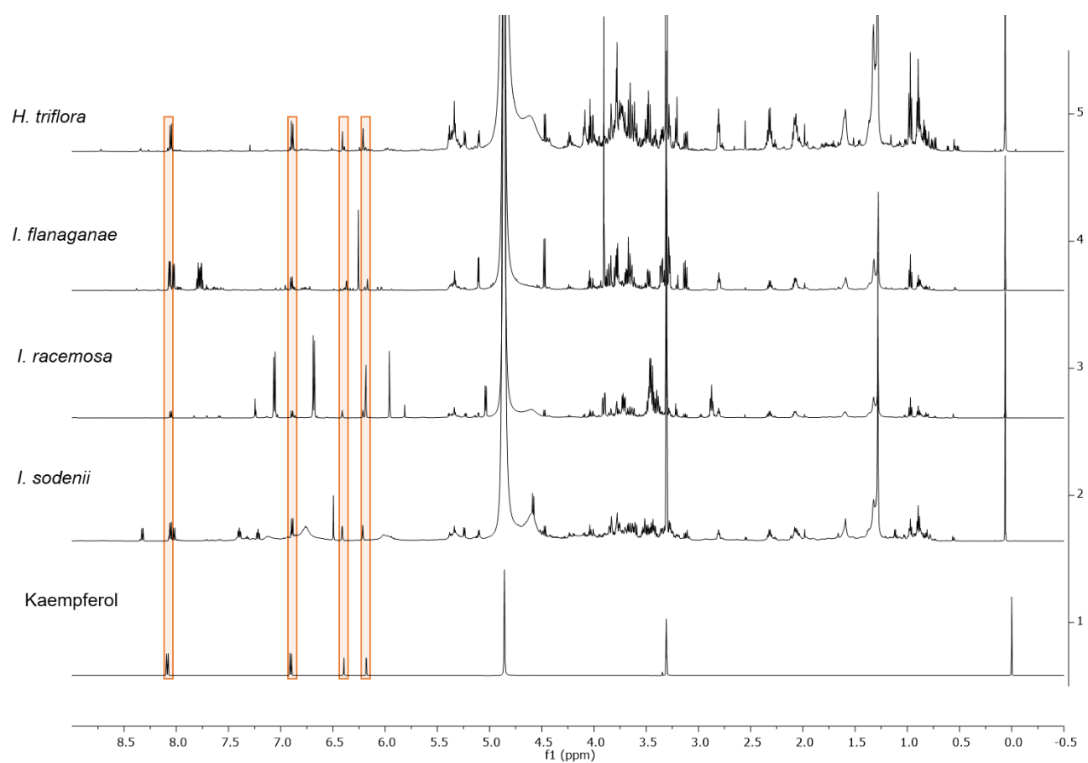
Figure A.14. Comparison of ^{1}H-NMR spectra of the methanolic crude leaf extracts of *H. triflora*, *I. flanaganae*, *I. racemosa*, *I. sodenii* and the reference compound kaempferol (common signals are highlighted in orange, 600 MHz, CD$_3$OD-*d4*, 25°C).).
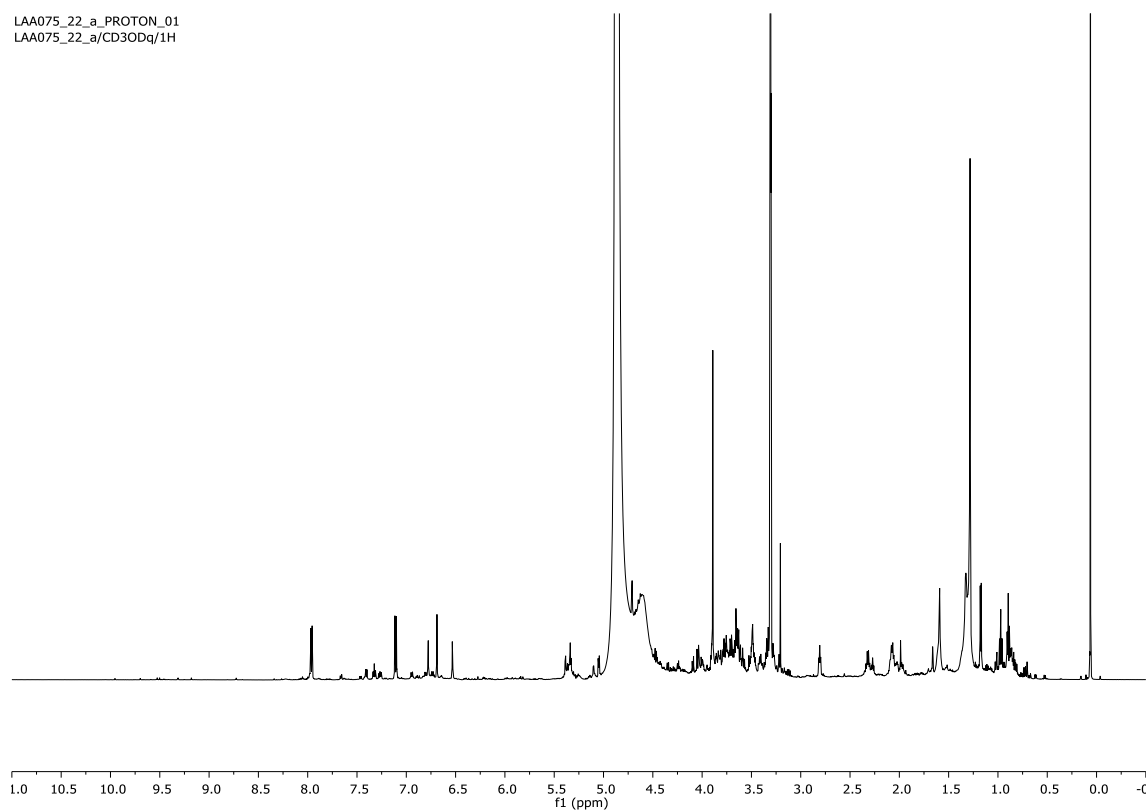
Figure A.15. ^{1}H-NMR spectrum of the methanolic crude leaf extracts of *Impatiens puberula* (600 MHz, CD$_3$OD-*d4*, 25°C).
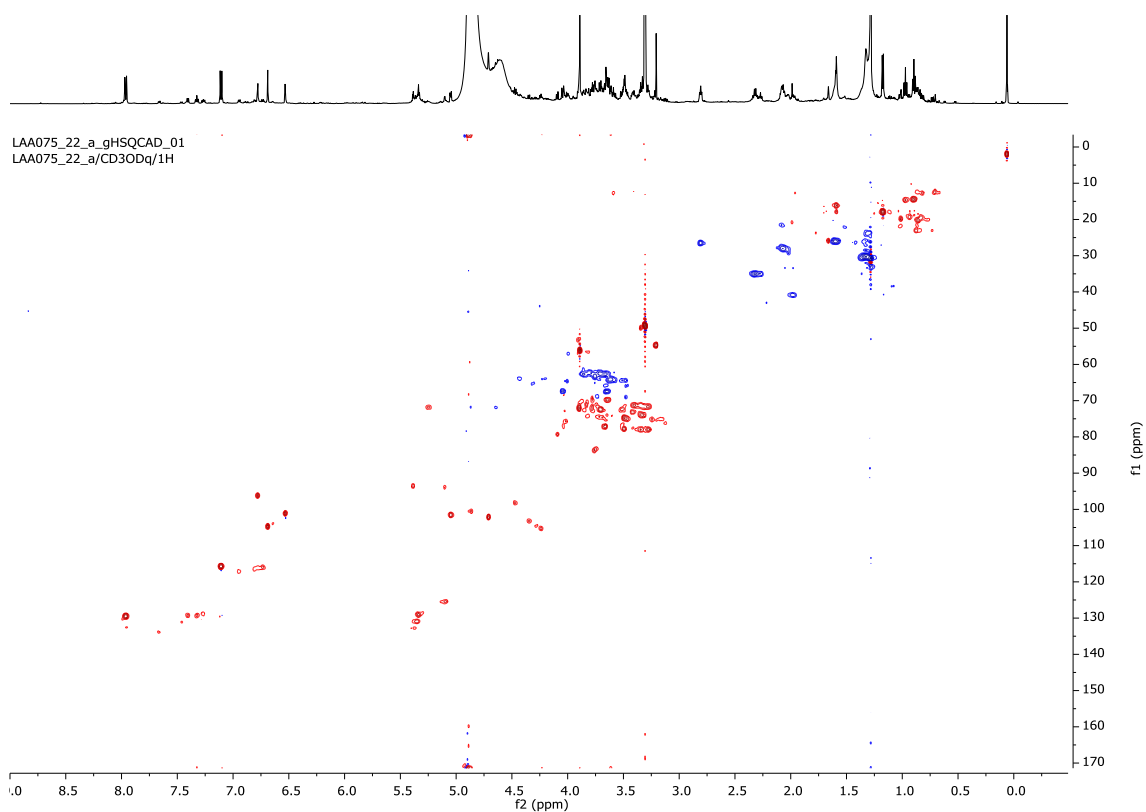
Figure A.16. ^{1}H, ^{13}C HSQC spectrum of the methanolic crude leaf extracts of *Impatiens puberula* (600 MHz, CD$_3$OD-*d4*, 25°C).
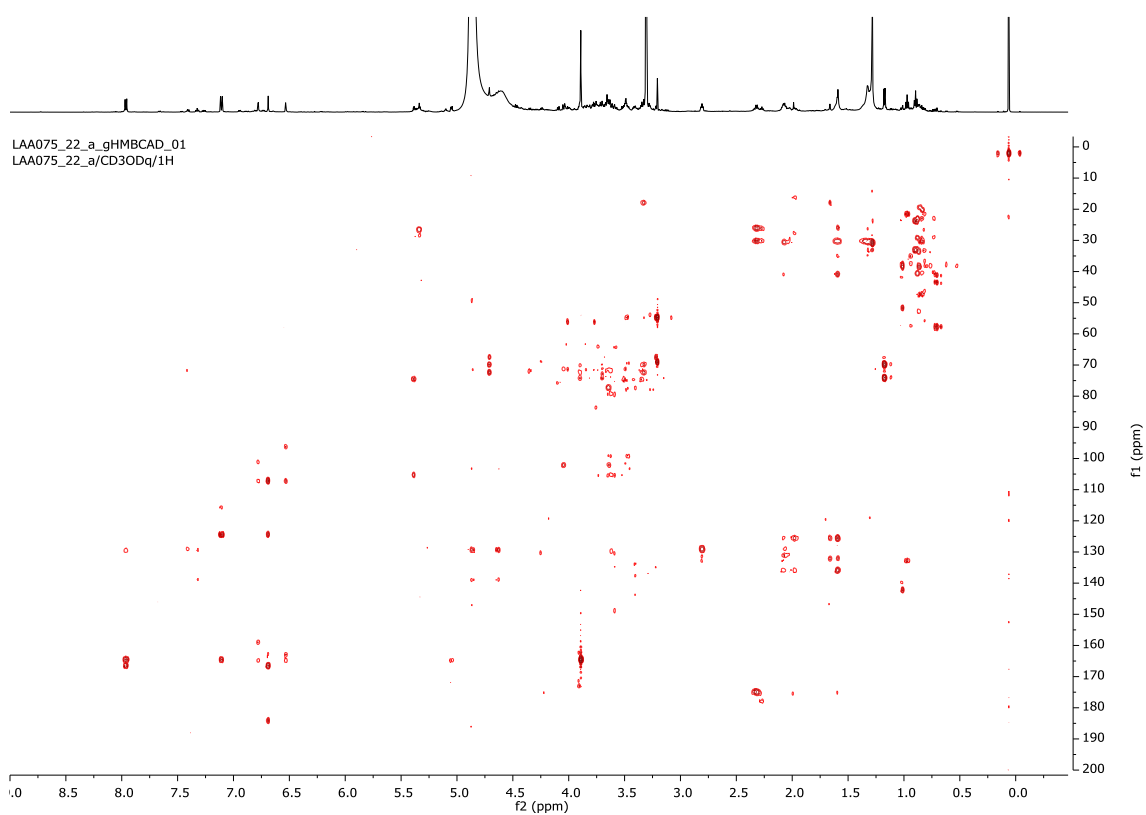
Figure A.17. ^{1}H, ^{13}C HMBC spectrum of the methanolic crude leaf extracts of *Impatiens puberula* (600 MHz, CD$_3$OD-*d4*, 25°C).
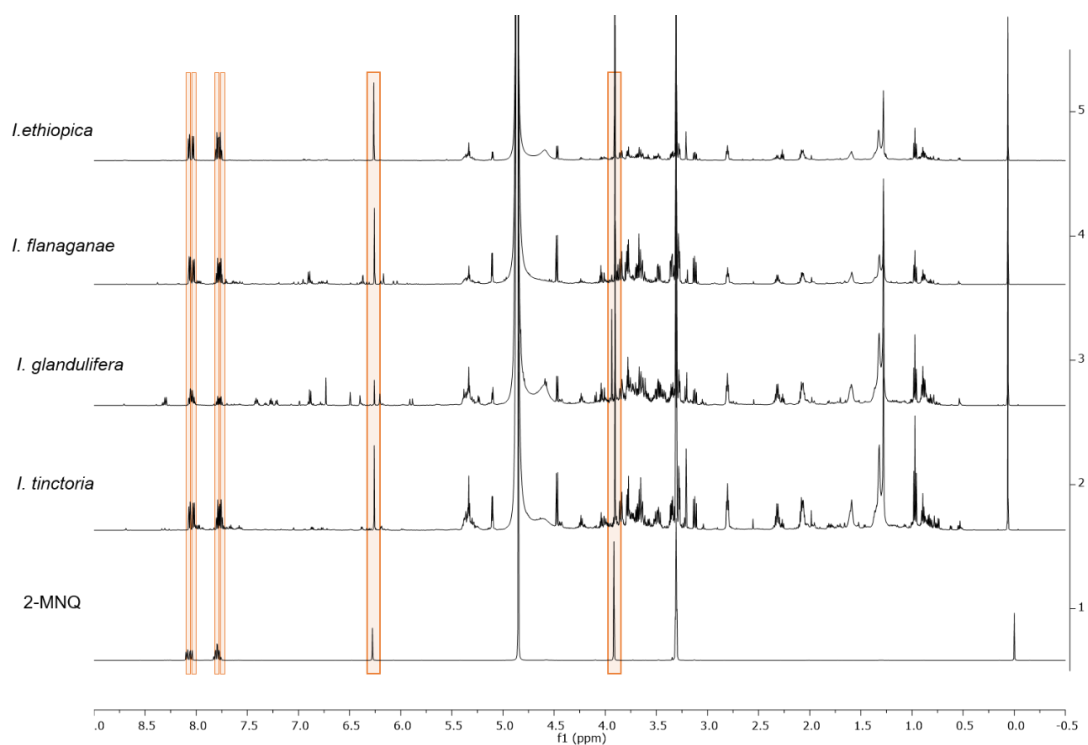
Figure A.18. Comparison of ^{1}H-NMR spectra of the methanolic crude leaf extracts of *I. ethiopica*, *I. flanaganae*, *I. glandulifera*, *I. tinctoria* and the reference compound 2-methoxy-1,4-naphthoquinone (2-MNQ, common signals are highlighted in orange, 600 MHz, CD$_3$OD-*d4*, 25°C).
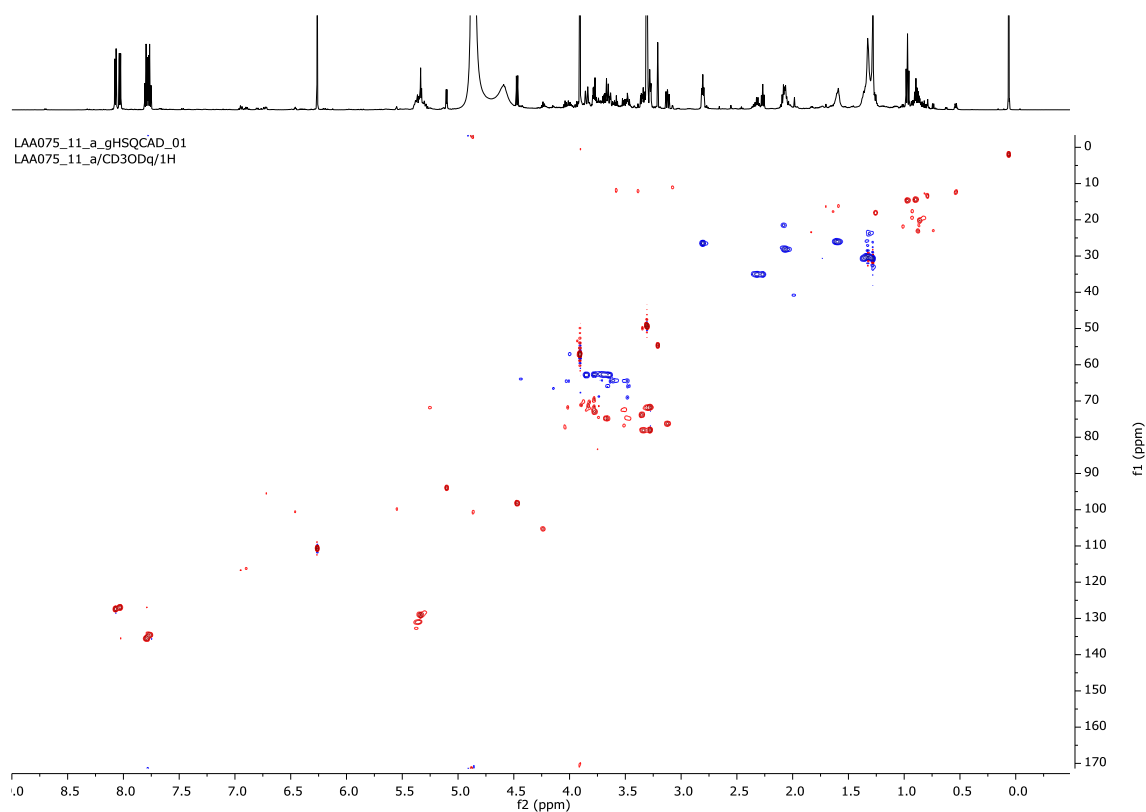
Figure A.19. ^{1}H, ^{13}C HSQC spectrum of the methanolic crude leaf extracts of *Impatiens ethiopica* (600 MHz, CD$_3$OD-*d4*, 25°C).
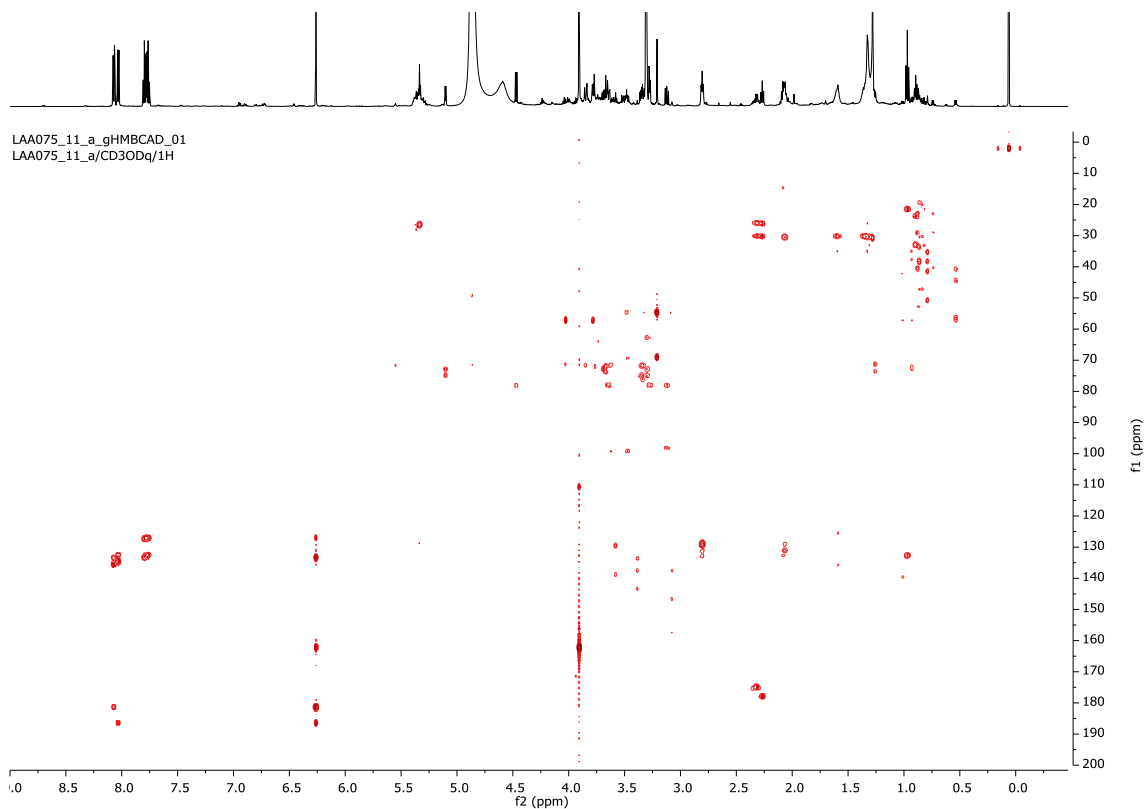
Figure A.20. ¹H, ¹³C HMBC spectrum of the methanolic crude leaf extracts of *Impatiens ethiopica* (600 MHz, CD₃OD-*d4*, 25°C).
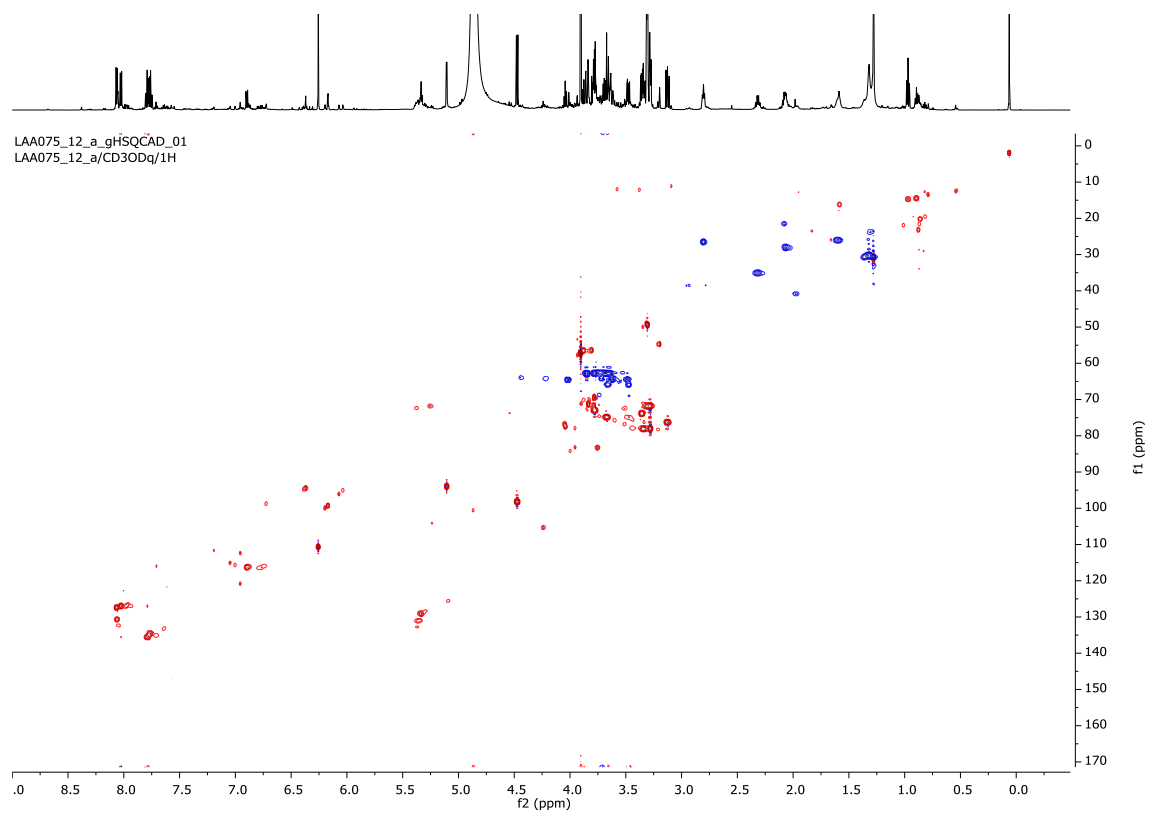
Figure A.21. ¹H, ¹³C HSQC spectrum of the methanolic crude leaf extracts of *Impatiens flanaganae* (600 MHz, CD₃OD-*d4*, 25°C).
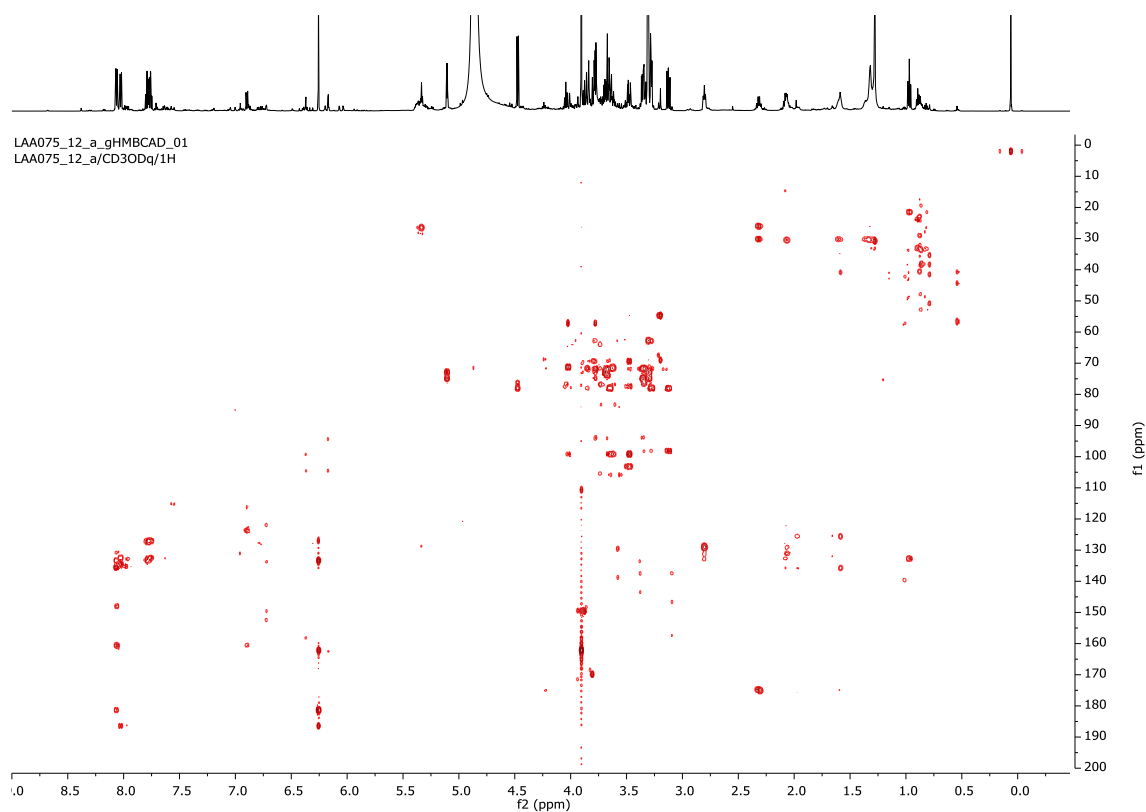
Figure A.22. ^{1}H, ^{13}C HMBC spectrum of the methanolic crude leaf extracts of *Impatiens flanaganae* (600 MHz, CD$_3$OD-*d4*, 25°C).
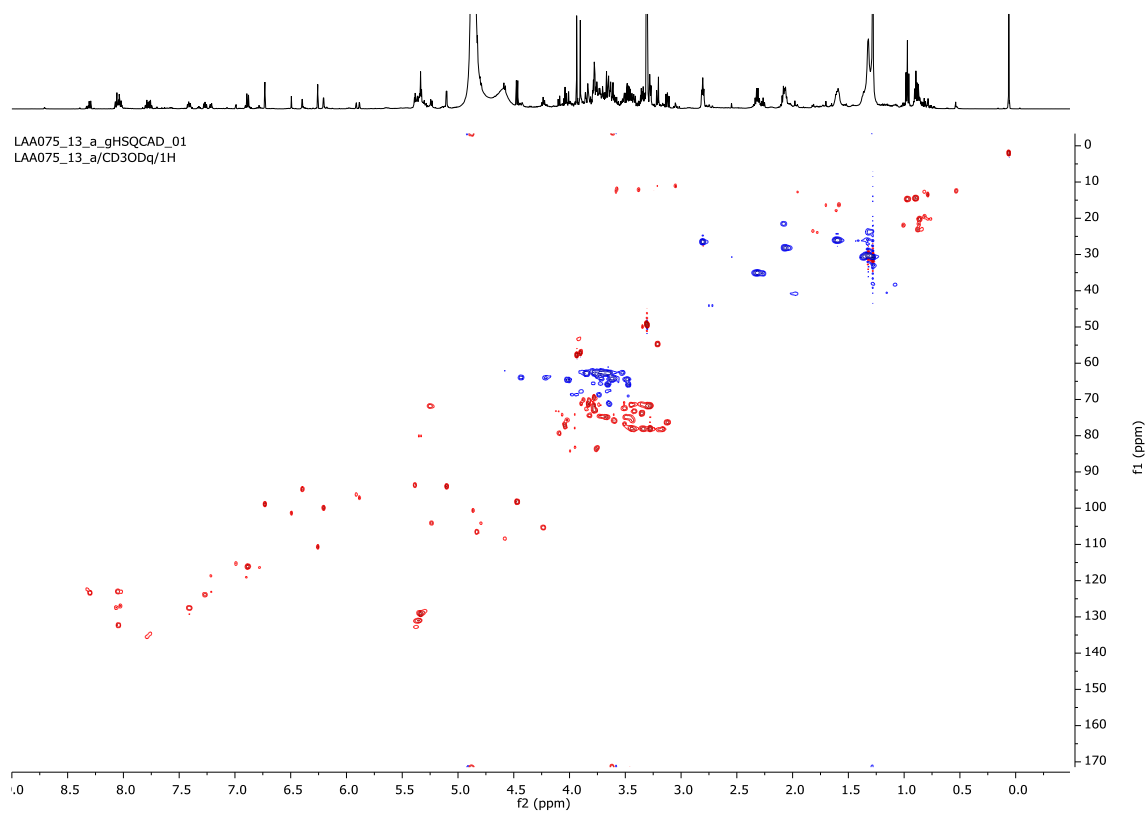
Figure A.23. ^{1}H, ^{13}C HSQC spectrum of the methanolic crude leaf extracts of *Impatiens glandulifera* (600 MHz, CD$_3$OD-*d4*, 25°C).
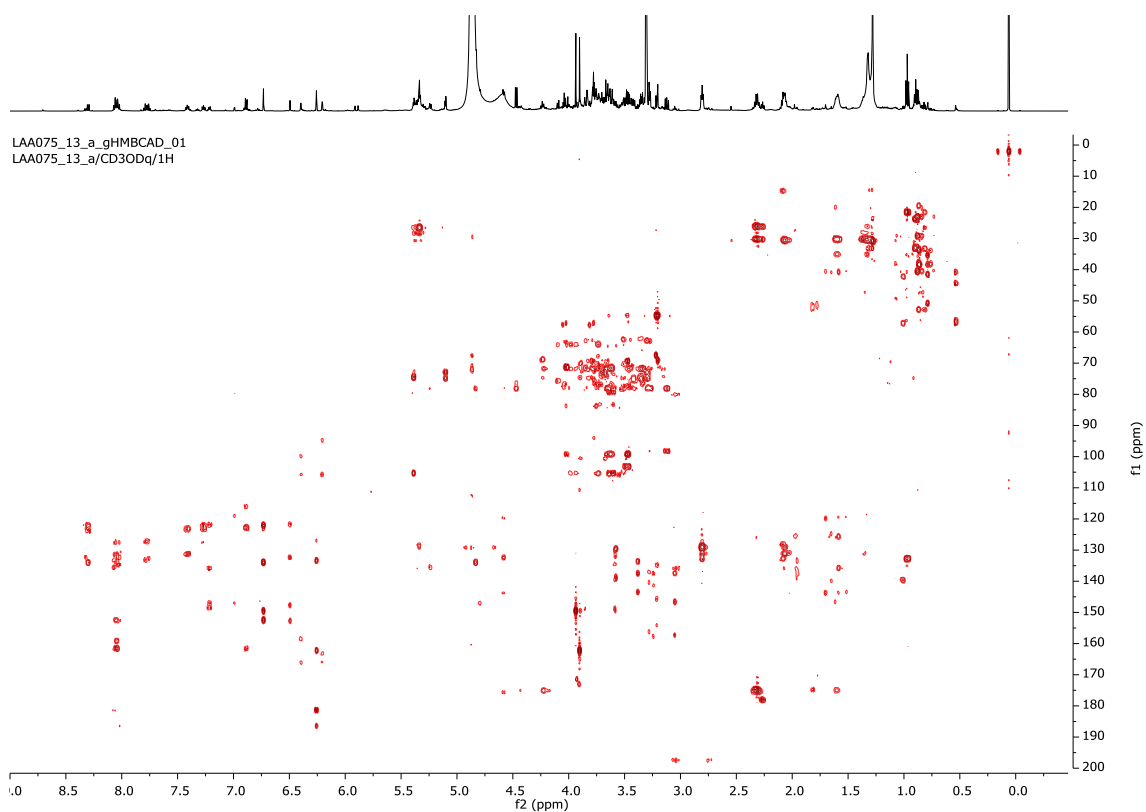
Figure A.24. ^{1}H, ^{13}C HMBC spectrum of the methanolic crude leaf extracts of *Impatiens glandulifera* (600 MHz, CD₃OD-*d4*, 25°C).
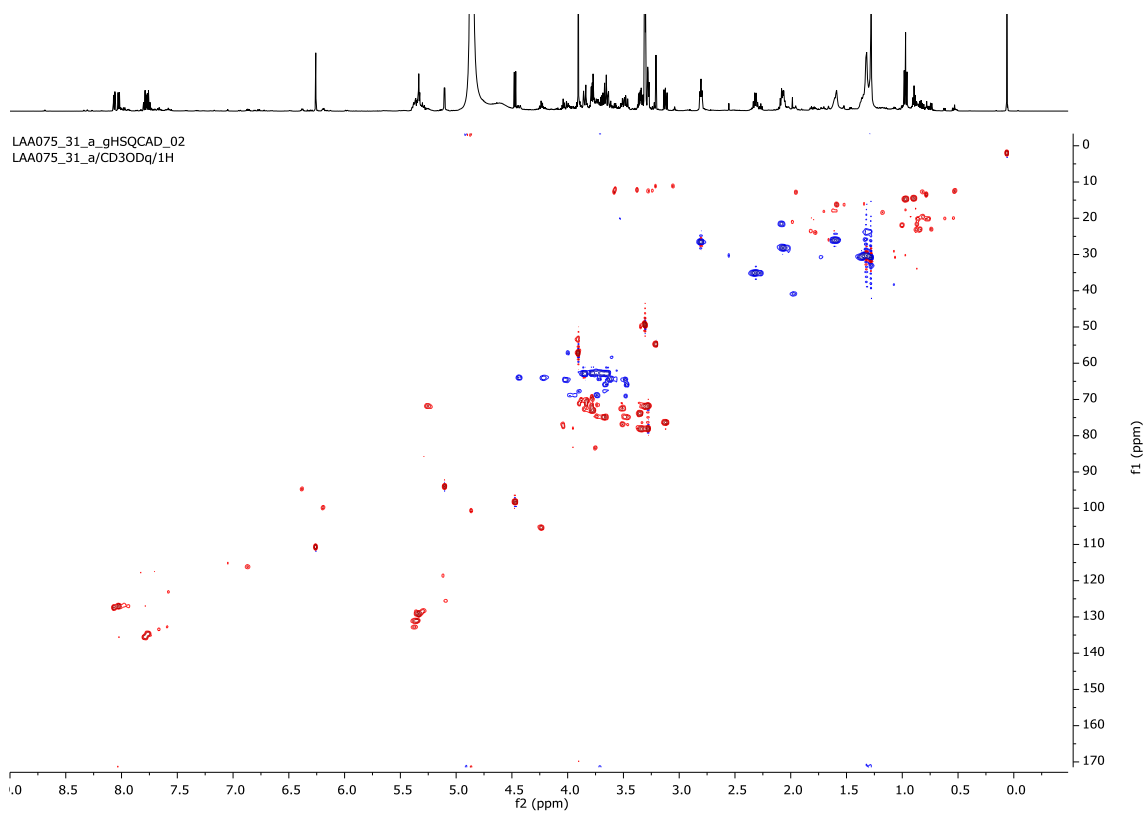
Figure A.25. ^{1}H, ^{13}C HSQC spectrum of the methanolic crude leaf extracts of *Impatiens textori* (600 MHz, CD₃OD-*d4*, 25°C).
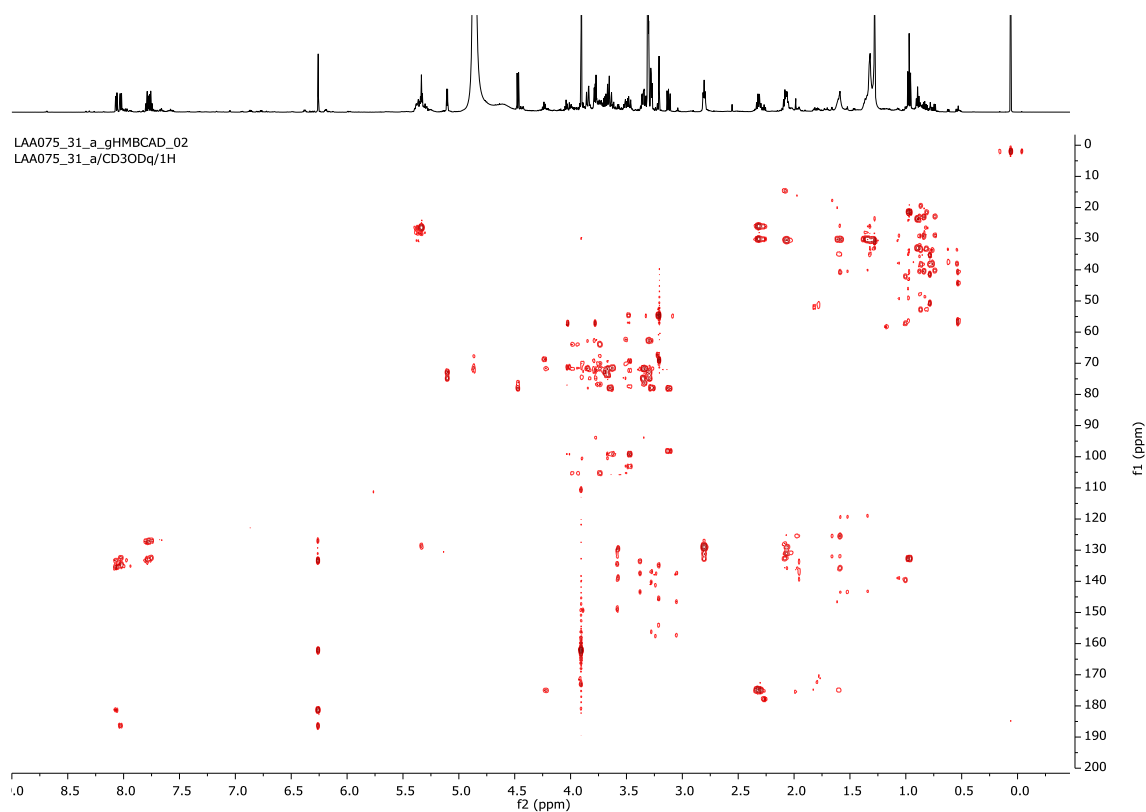
Figure A.26. ^{1}H, ^{13}C HMBC spectrum of the methanolic crude leaf extracts of *Impatiens textori* (600 MHz, CD₃OD-*d4*, 25°C).
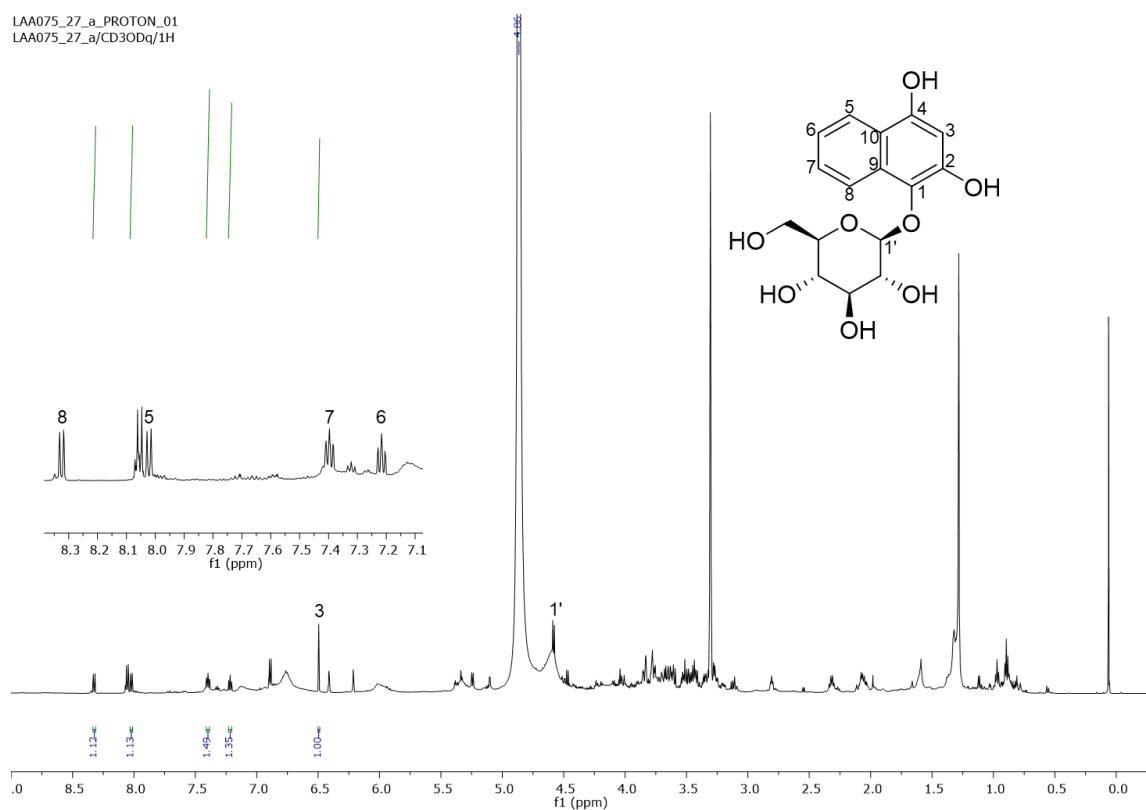
Figure A.27. ^{1}H-NMR spectrum of the methanolic crude leaf extracts of *Impatiens sodenii* including the assignment of 1,2,4-trihydroxynaphthalene-1-*O*-β-D-glucoside (600 MHz, CD₃OD-*d4*, 25°C).
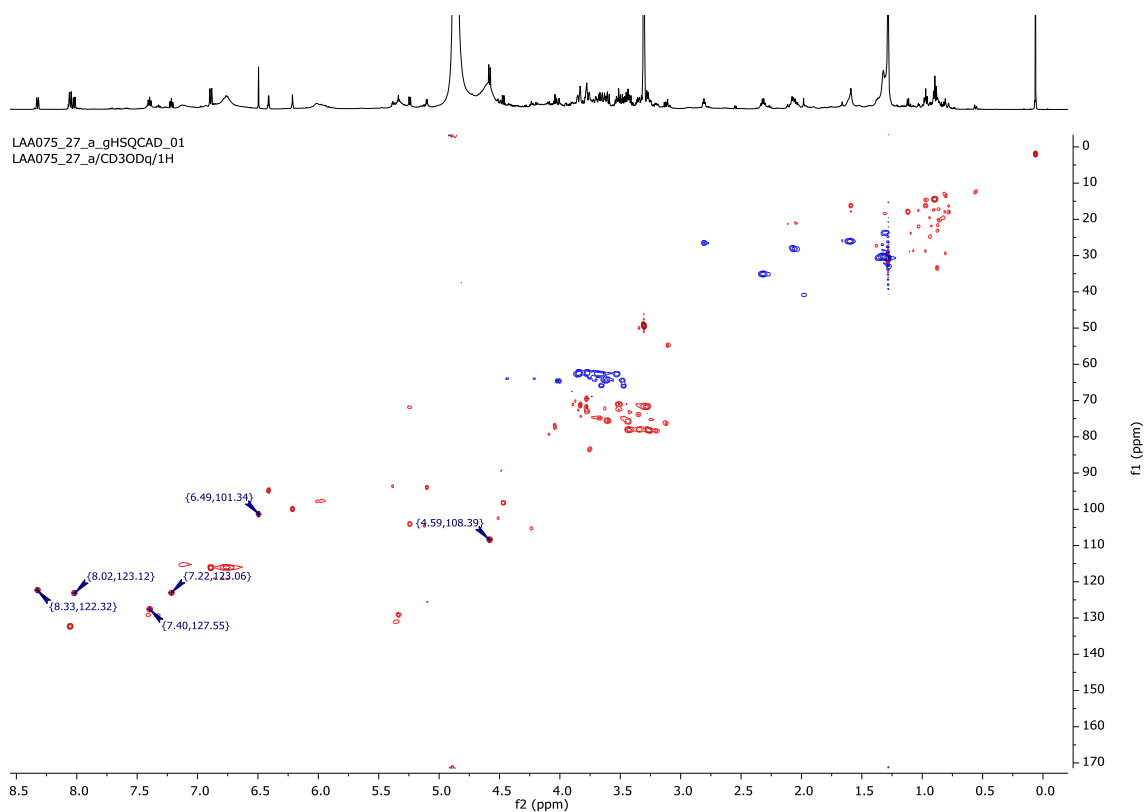
Figure A.28. ^{1}H, ^{13}C HSQC spectrum of the methanolic crude leaf extracts of *Impatiens sodenii* including picked peaks corresponding to 1,2,4-trihydroxynaphthalene-1-*O*-β-D-glucoside (600 MHz, CD$_3$OD- *d4*, 25°C).
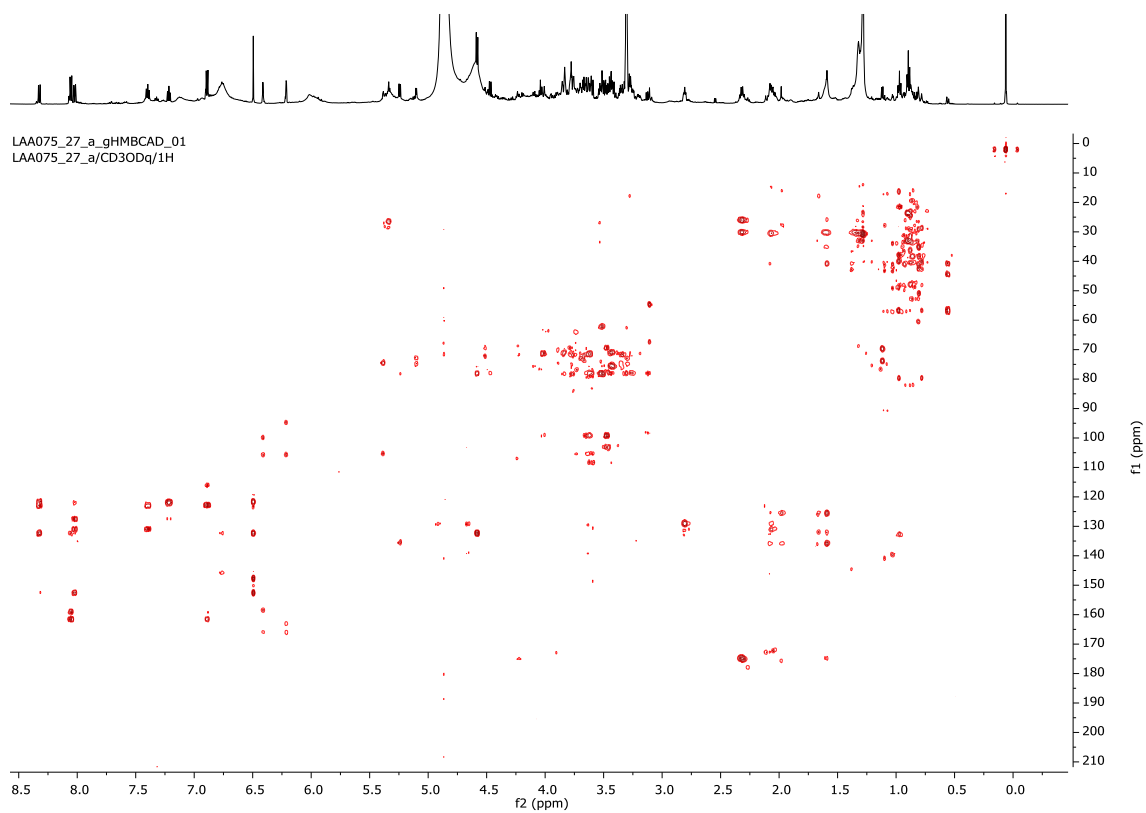
Figure A.29. ^{1}H, ^{13}C HMBC spectrum of the methanolic crude leaf extracts of *Impatiens sodenii* (600 MHz, CD$_3$OD-*d4*, 25°C).
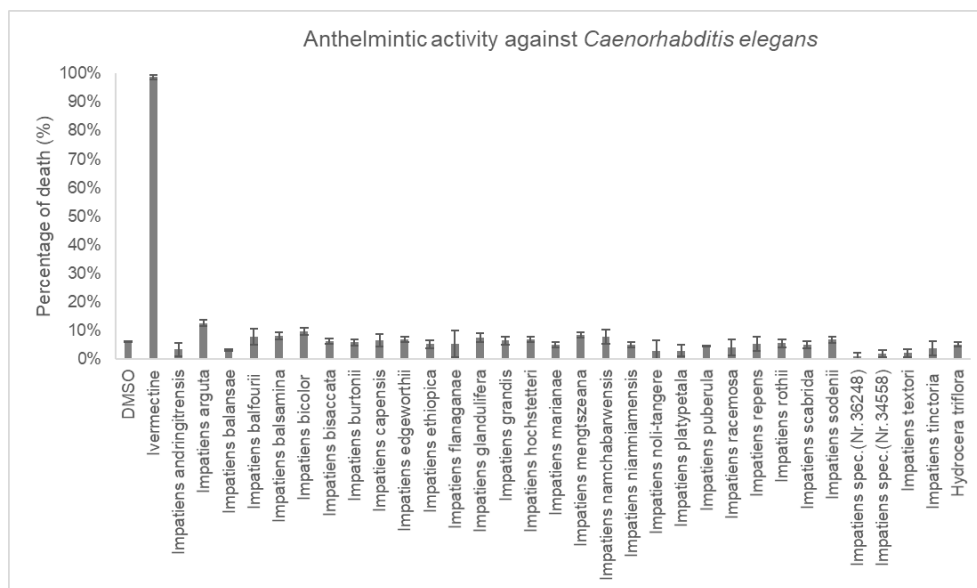
Figure A.30. Anthelmintic activity of methanolic leaf extracts (500 µg/mL) against *Caenorhabditis elegans*. In all the assays, the solvent DMSO (2%) and the standard anthelmintic drug ivermectin (10 µg/mL) were used as negative and positive controls, respectively. All the assays were carried out in triplicate.
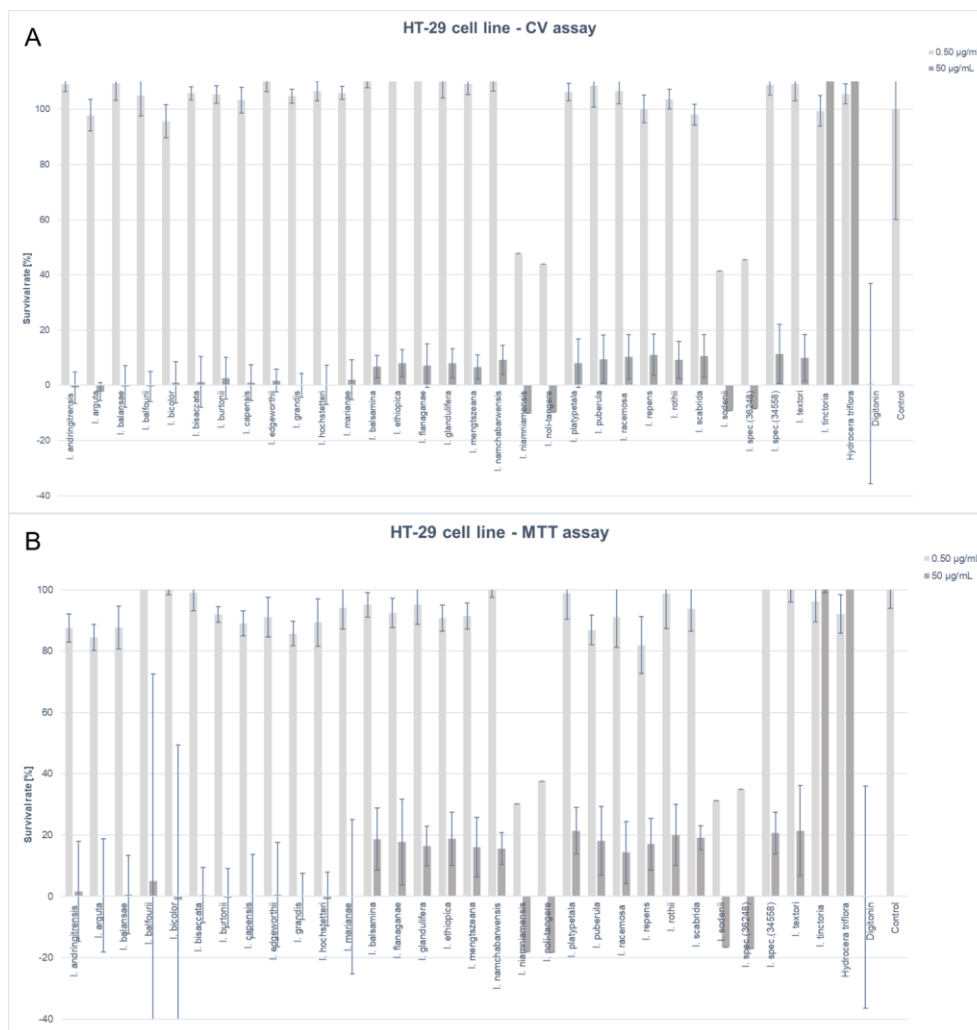
Figure A.31. Anticancer activity of methanolic crude leaf extracts of *Impatiens* species and *Hydrocera triflora* against HT-29 cancer cell line using A) CV assay and B) MTT assay.
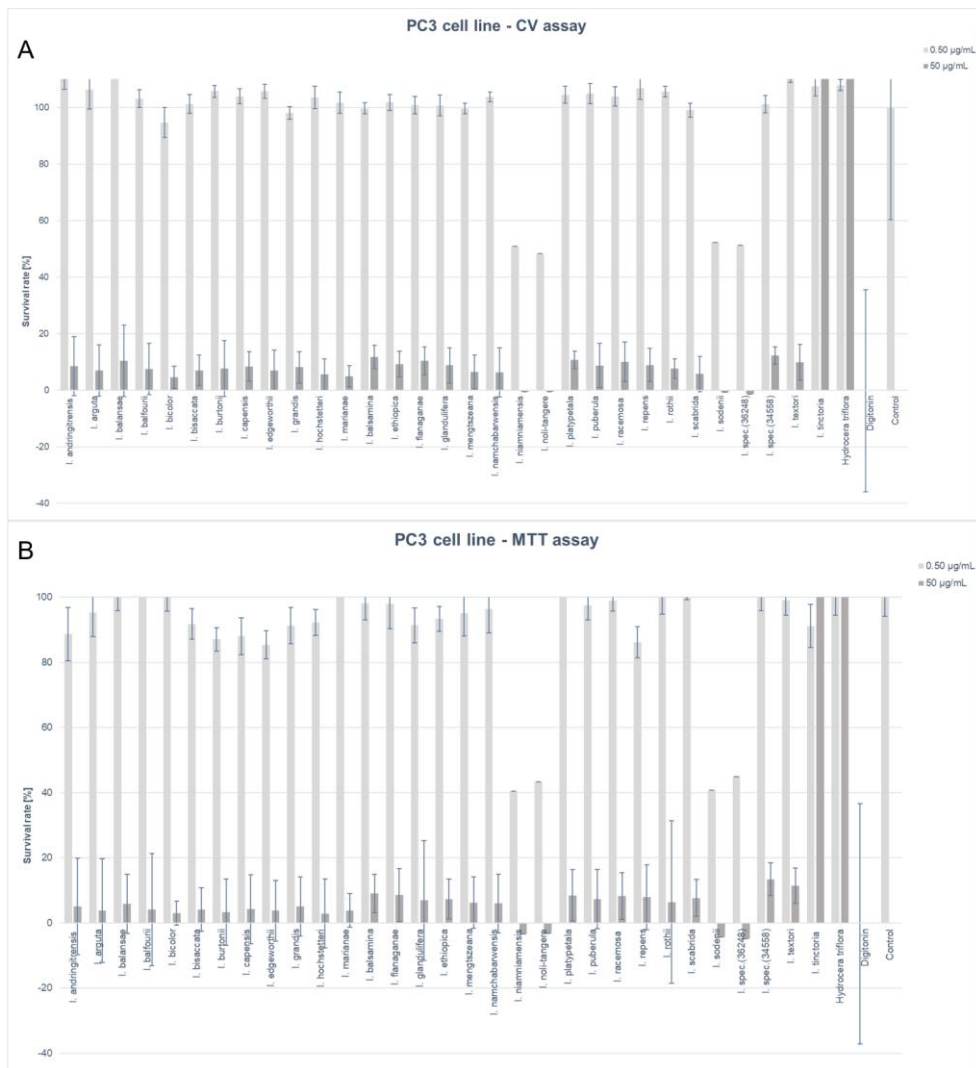
Figure A.32. Anticancer activity of methanolic crude leaf extracts of *Impatiens* species and *Hydrocera triflora* against against PC3 cancer cell line using A) CV assay and B) MTT assay.
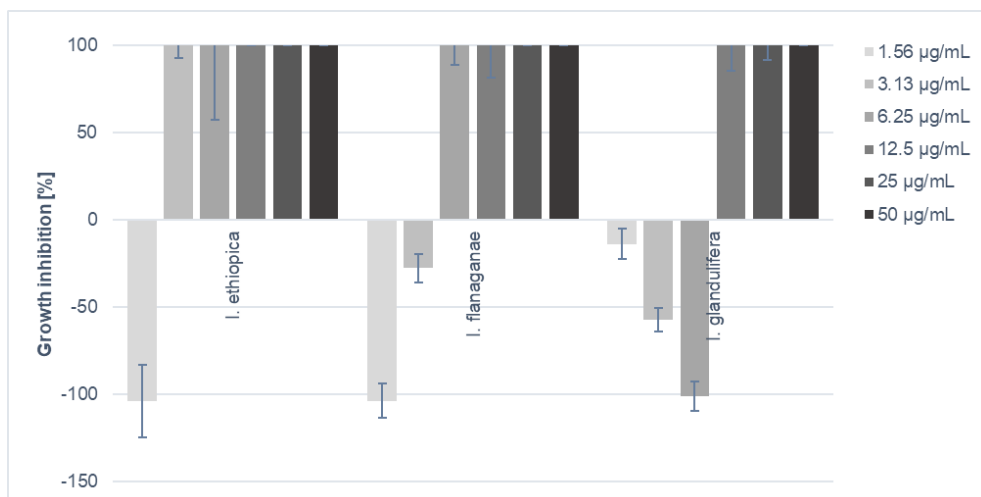
Figure A.33. Antibacterial activity of methanolic crude leaf extracts *I. ethiopica*, *I. flanaganae*, and *I. glandulifera* against gram-negative *A. fischeri*.
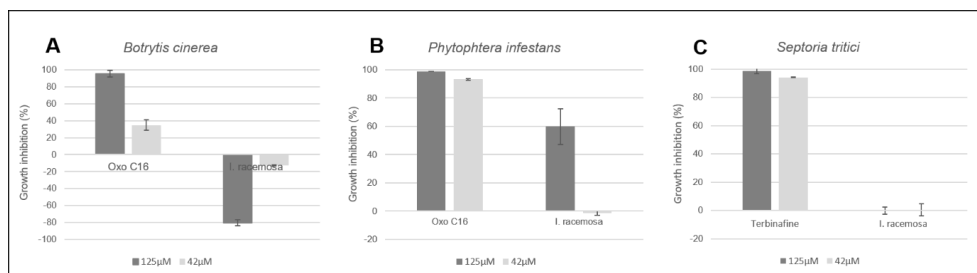
Figure A.34. Antifungal activity of methanolic leaf extract of *Impatiens racemosa* DC against A: *Botrytis cinerea*, B: the oomycete *P. infestans* and the C: *S. tritici*. Positive controls were the known fungicides Oxocrotonat C 16 (A, B) and Terbinafine (C).
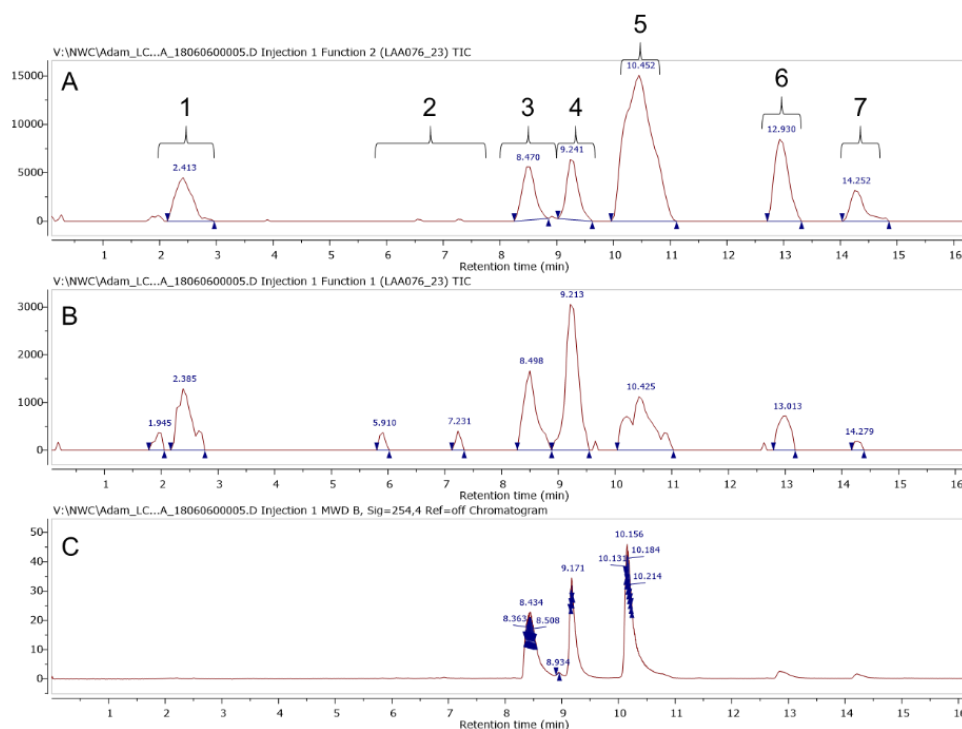
Figure A.35. Preparative LC-MS chromatogram of fractioned methanolic extract of *I. racemosa*. (A: TIC negative ion mode, B: TIC positive ion mode, C: MWD, λ = 254 nm)
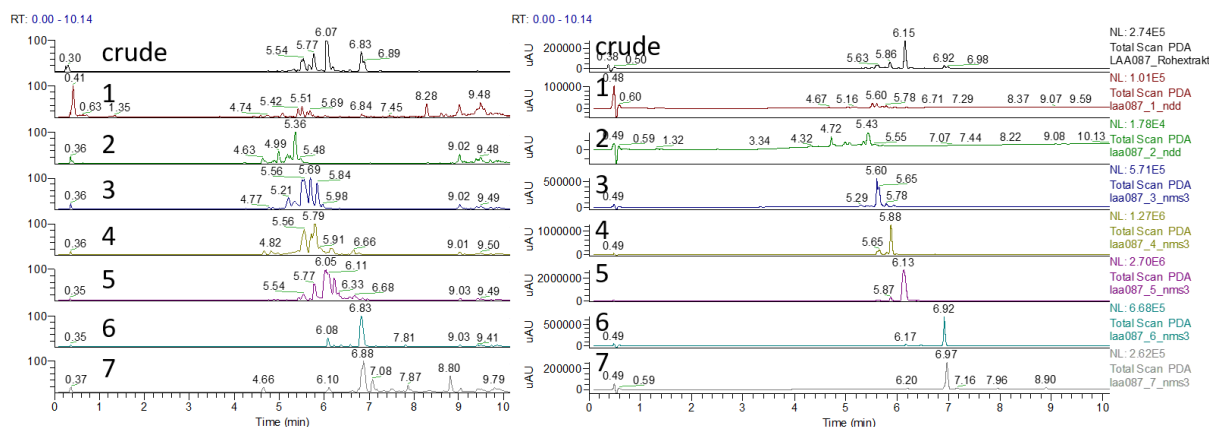
Figure A.36. Comparison of the total ion chromatograms (left) and PDA spectra (right) of the crude methanolic extract of *Impatiens racemosa* and the obtained fractions after preparative HPLC-MS separation by UHPLC-PDA-ESI-HRMS in negative ion mode.
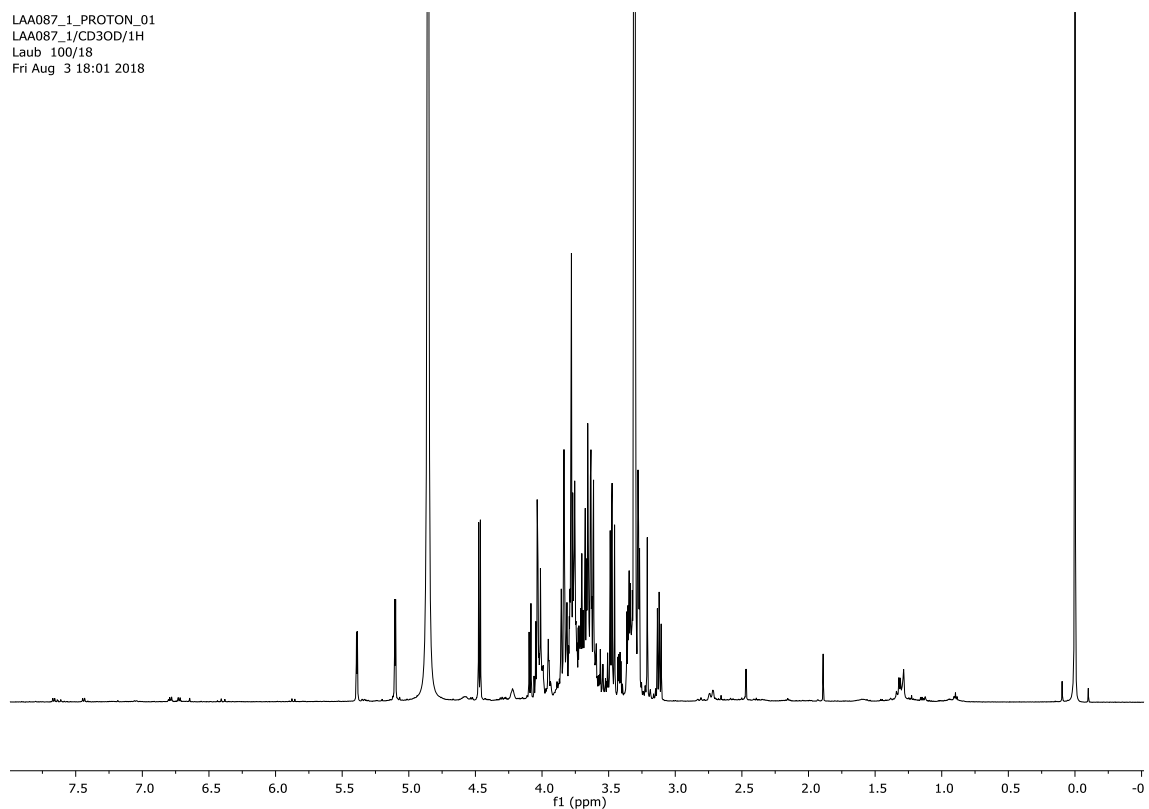
Figure A.37. ^{1}H-NMR spectrum of fraction 1 (t_R (preparative HPLC-MS): 2.1-3.0 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* (600 MHz, CD$_3$OD-*d4*, 25°C).
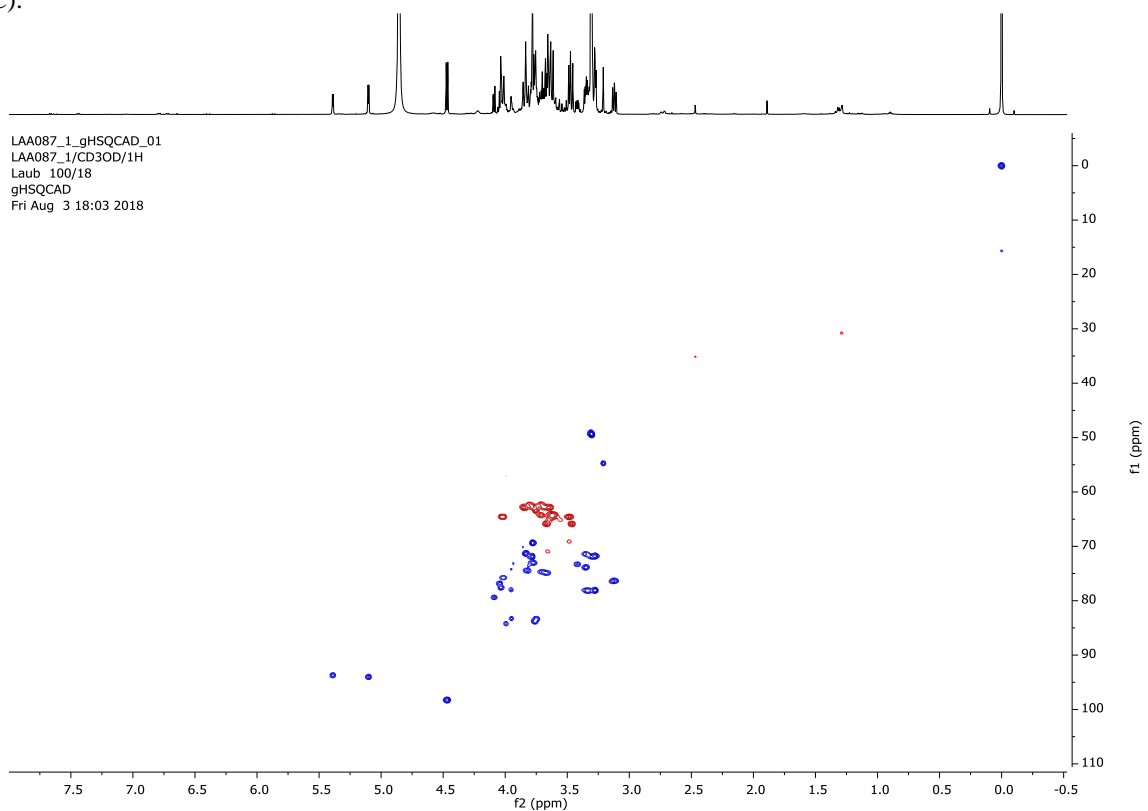
Figure A.38. ^{1}H, ^{13}C HSQC spectrum of fraction 1 (t_R (preparative HPLC-MS): 2.1-3.0 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* (600 MHz, CD$_3$OD-*d4*, 25°C).
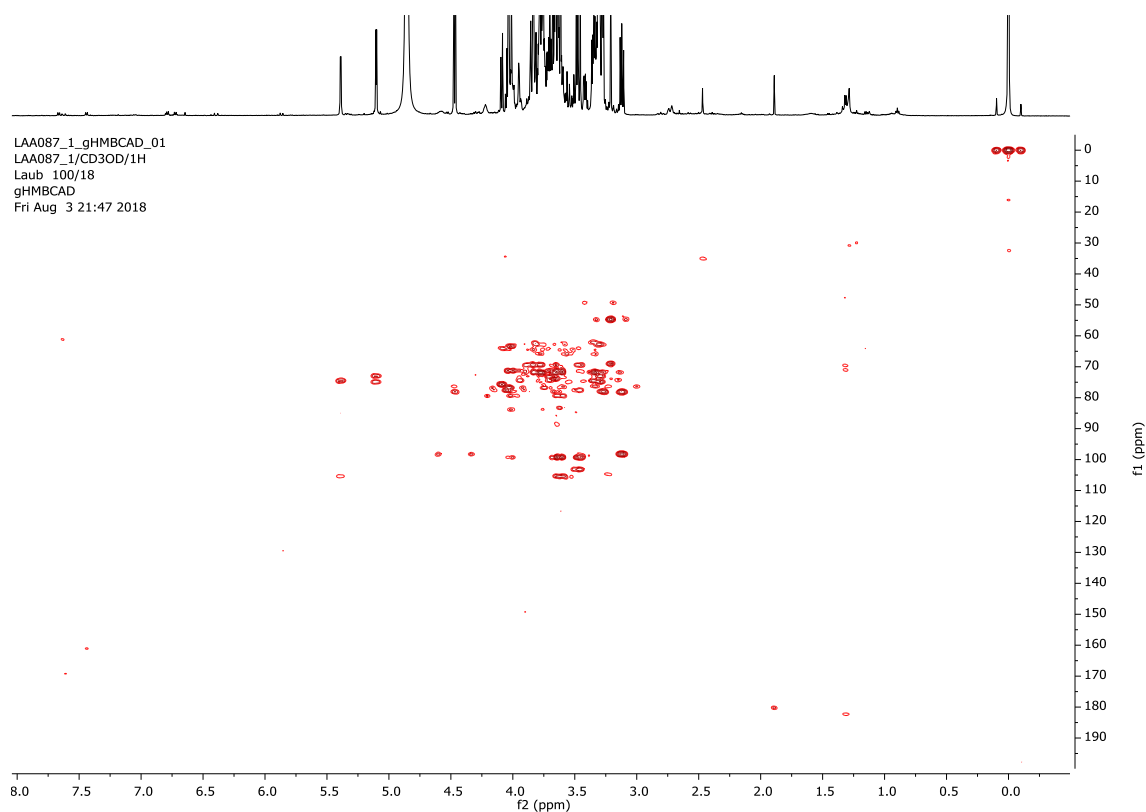
Figure A.39. ^{1}H, ^{13}C HSQC spectrum of fraction 1 (t_R (preparative HPLC-MS): 2.1-3.0 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* (600 MHz, CD$_3$OD-*d4*, 25°C).
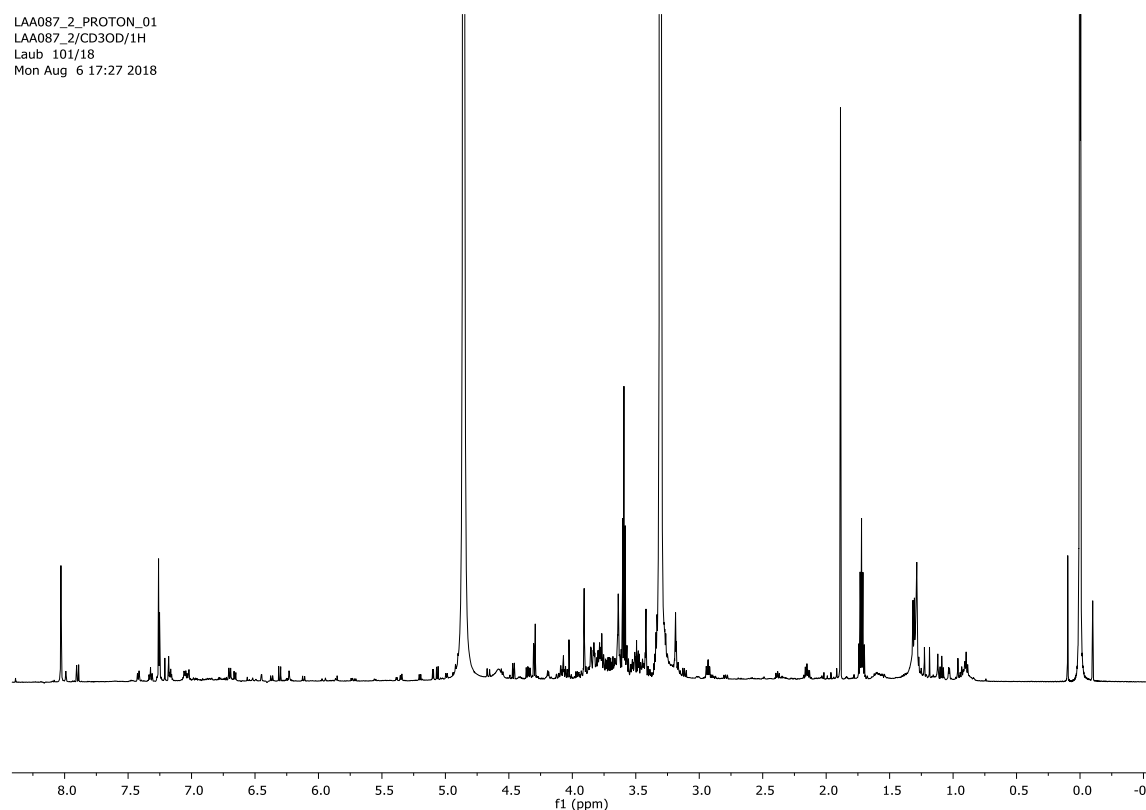
Figure A.40. ^{1}H-NMR spectrum of fraction 2 (t_R (preparative HPLC-MS): 6.0-7.6 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* (600 MHz, CD$_3$OD-*d4*, 25°C).
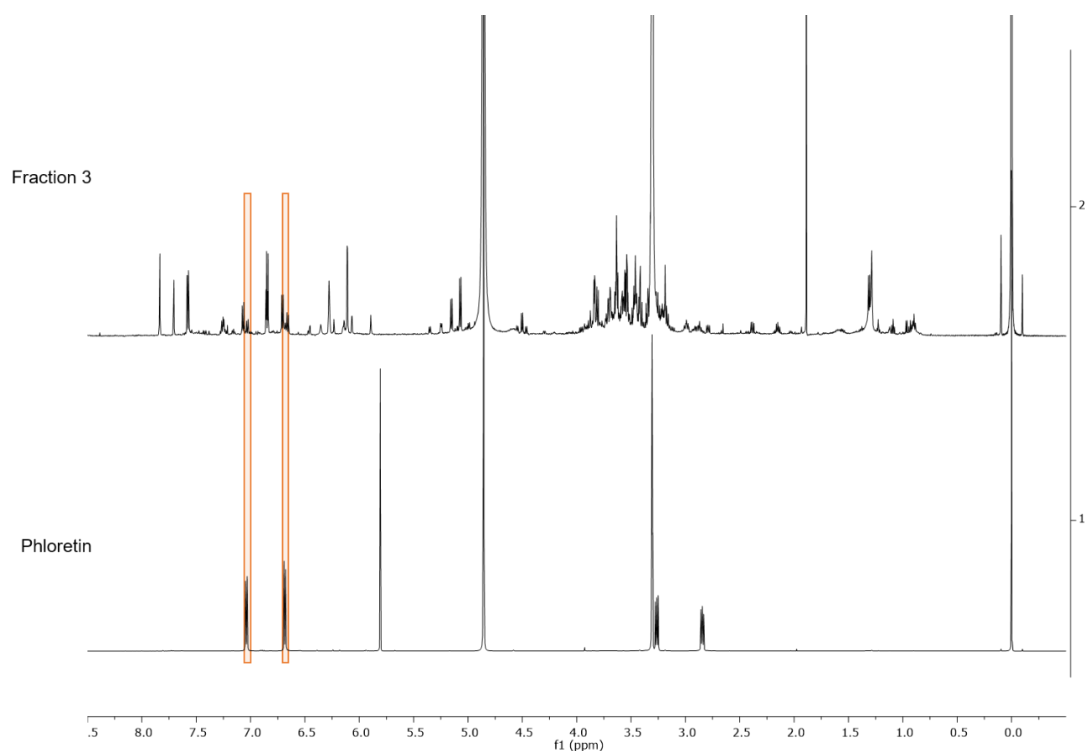
Figure A.41. Comparison of ^{1}H-NMR spectra of fraction 3 (t_R (preparative HPLC-MS): 8.2-8.9 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* and the reference compound phloretin (common signals are highlighted in orange, 600 MHz, CD$_3$OD-*d4*, 25°C).
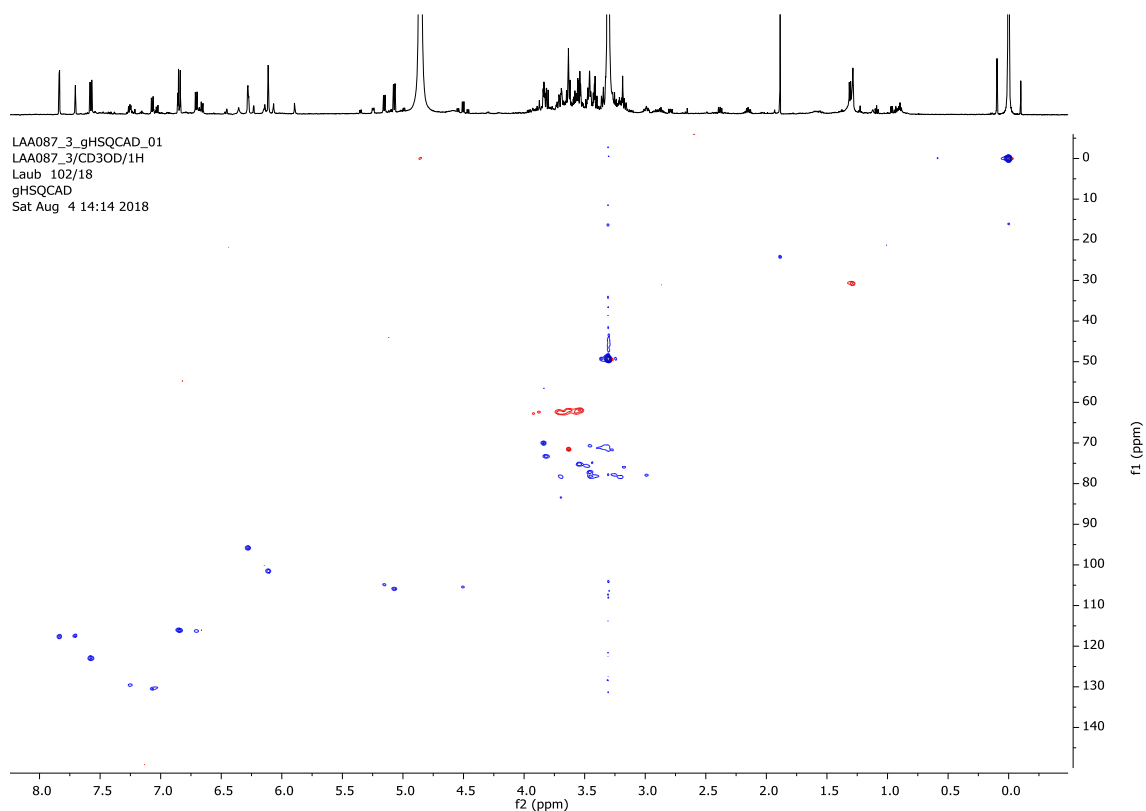
Figure A.42. ^{1}H, ^{13}C HSQC spectrum of fraction 3 (t_R (preparative HPLC-MS): 8.2-8.9 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* (600 MHz, CD$_3$OD-*d4*, 25°C).
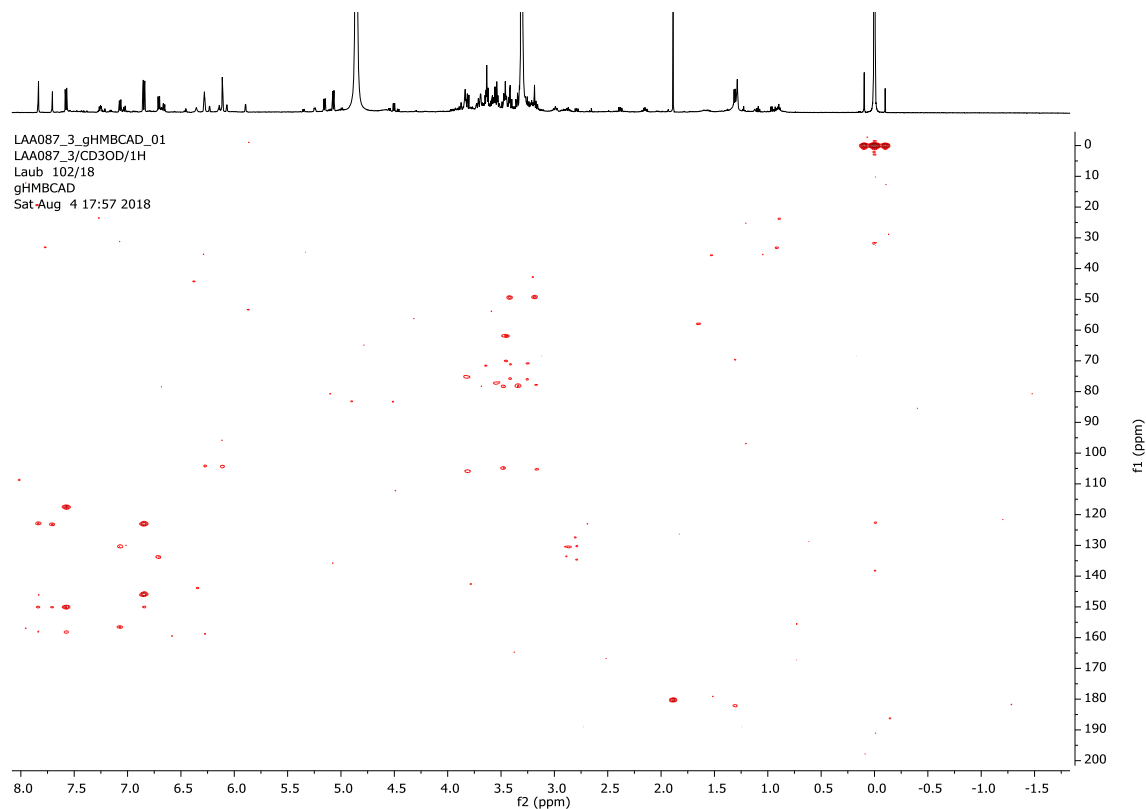
Figure A.43. ^{1}H, ^{13}C HMBC spectrum of fraction 3 (t_R (preparative HPLC-MS): 8.2-8.9 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* (600 MHz, CD$_3$OD-*d4*, 25°C).
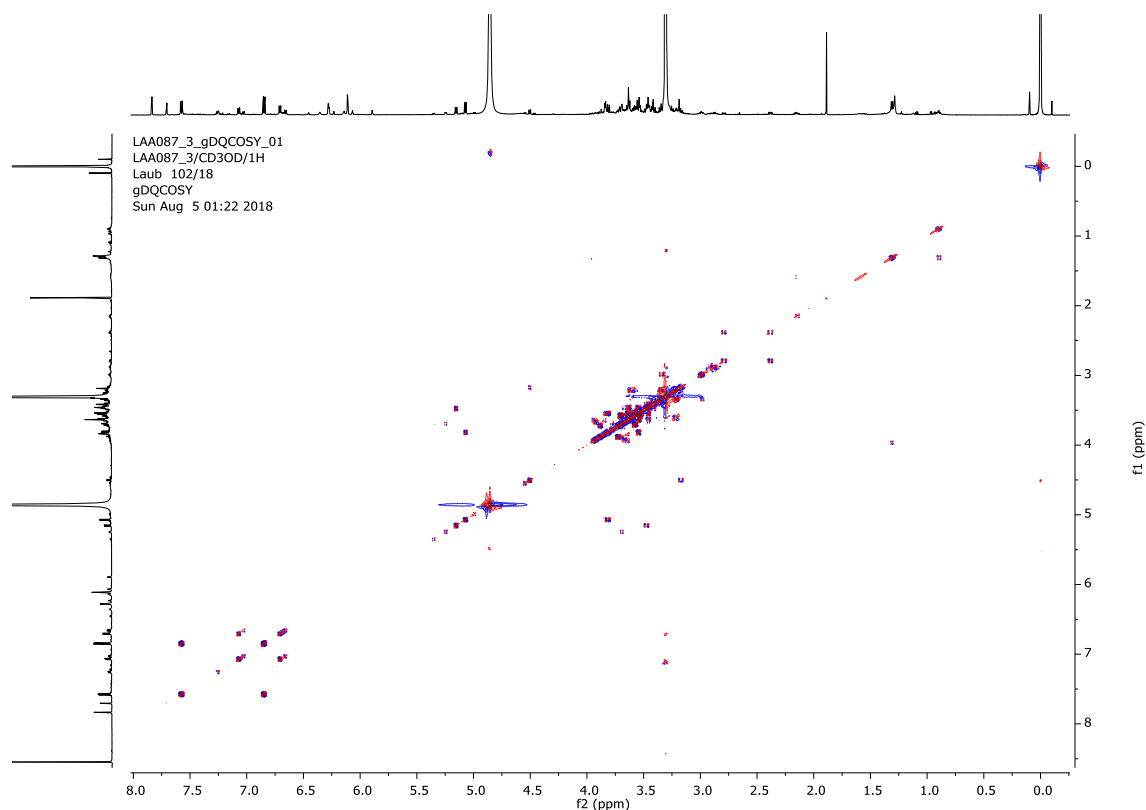
Figure A.44. ^{1}H, ^{1}H COSY spectrum of fraction 3 (t_R (preparative HPLC-MS): 8.2-8.9 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* (600 MHz, CD$_3$OD-*d4*, 25°C).
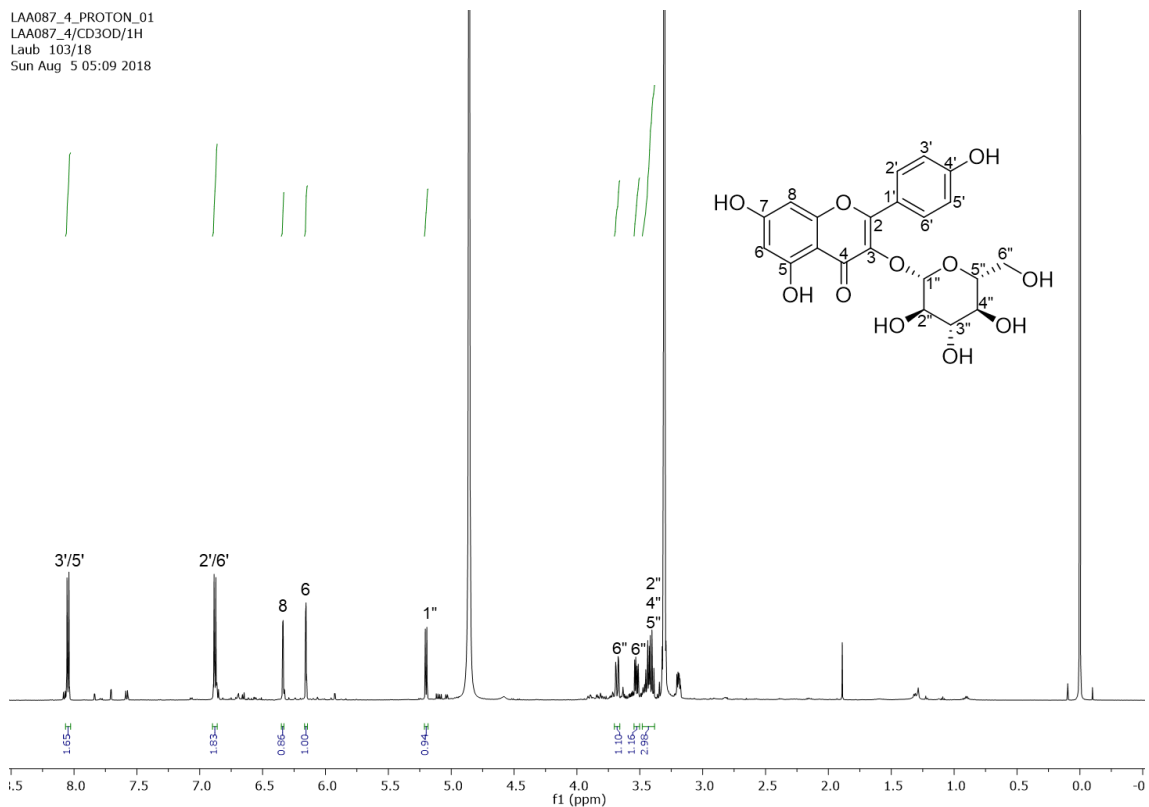
Figure A.45. ¹H-NMR spectrum of fraction 4 (t_R (preparative HPLC-MS): 9.0-9.7 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* including the assignment of astragalin (**2.3**, 600 MHz, CD₃OD-*d4*, 25°C).
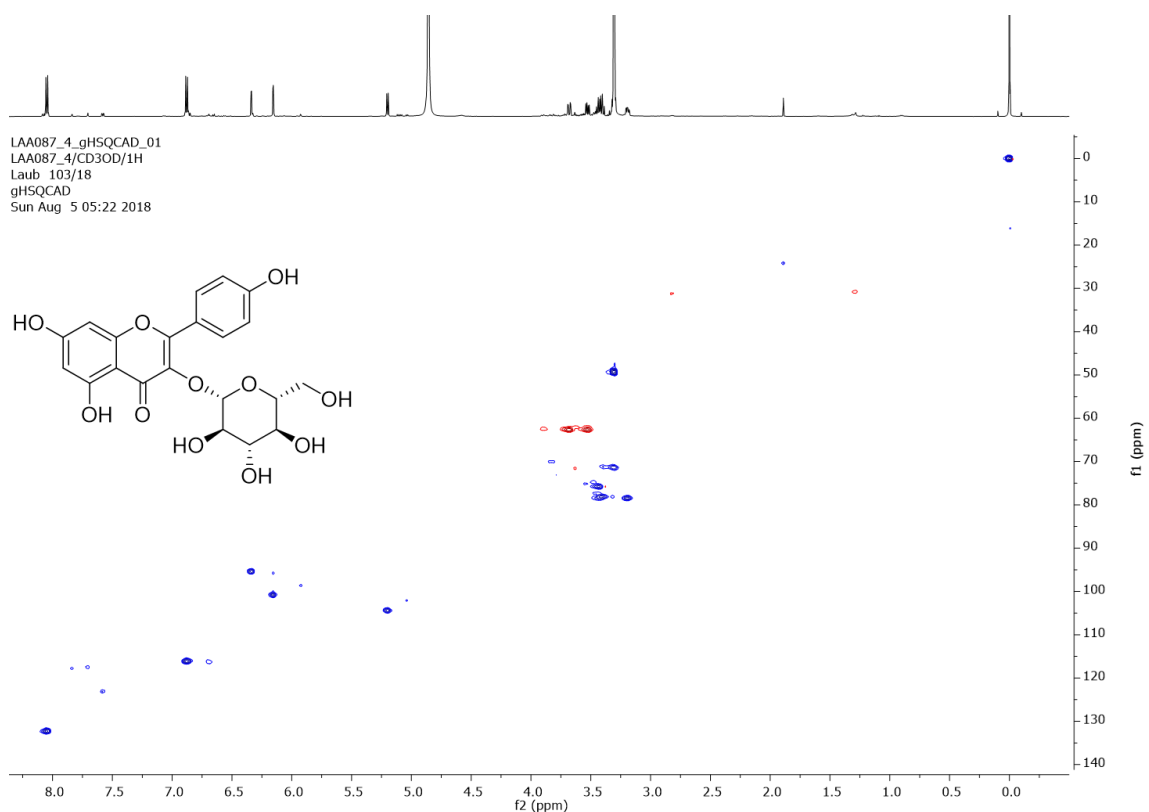
Figure A.46. ¹H, ¹³C HSQC spectrum of fraction 4 (**2.3**, t_R (preparative HPLC-MS): 9.0-9.7 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* (600 MHz, CD₃OD-*d4*, 25°C).
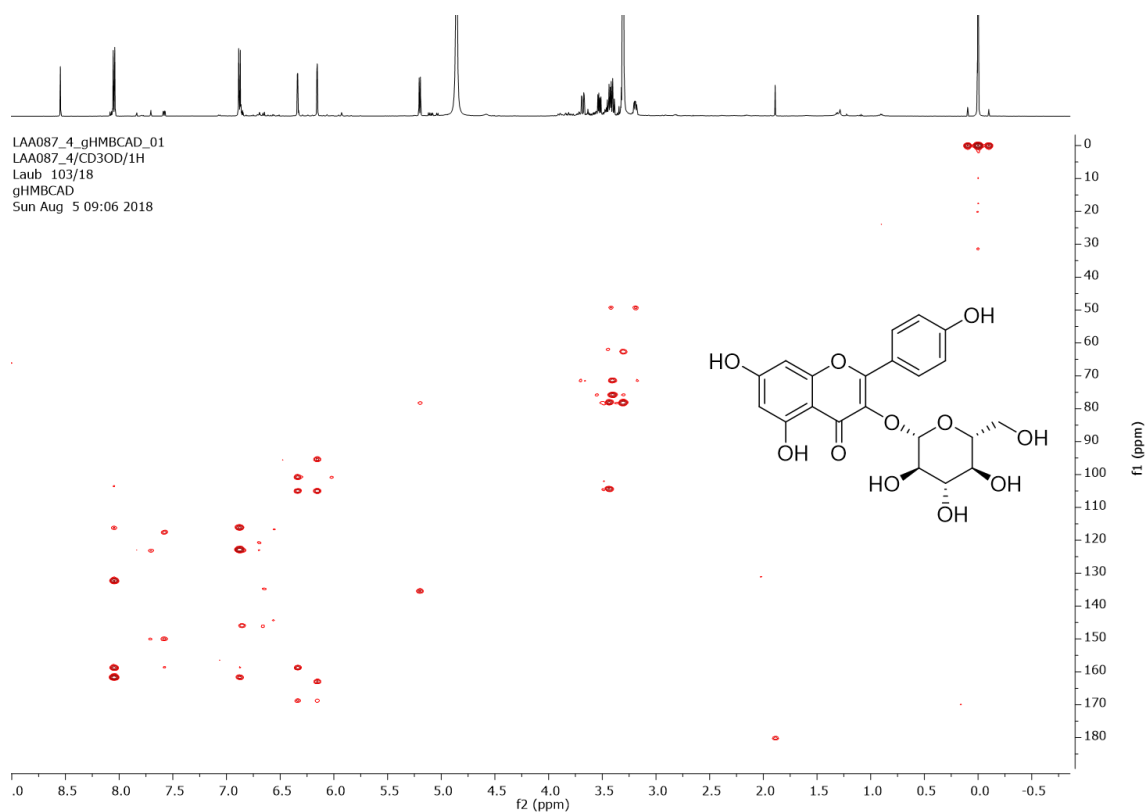
Figure A.47. ^{1}H, ^{13}C HMBC spectrum of fraction 4 (**2.3**, t_R (preparative HPLC-MS): 9.0-9.7 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* (600 MHz, CD$_3$OD-*d4*, 25°C).
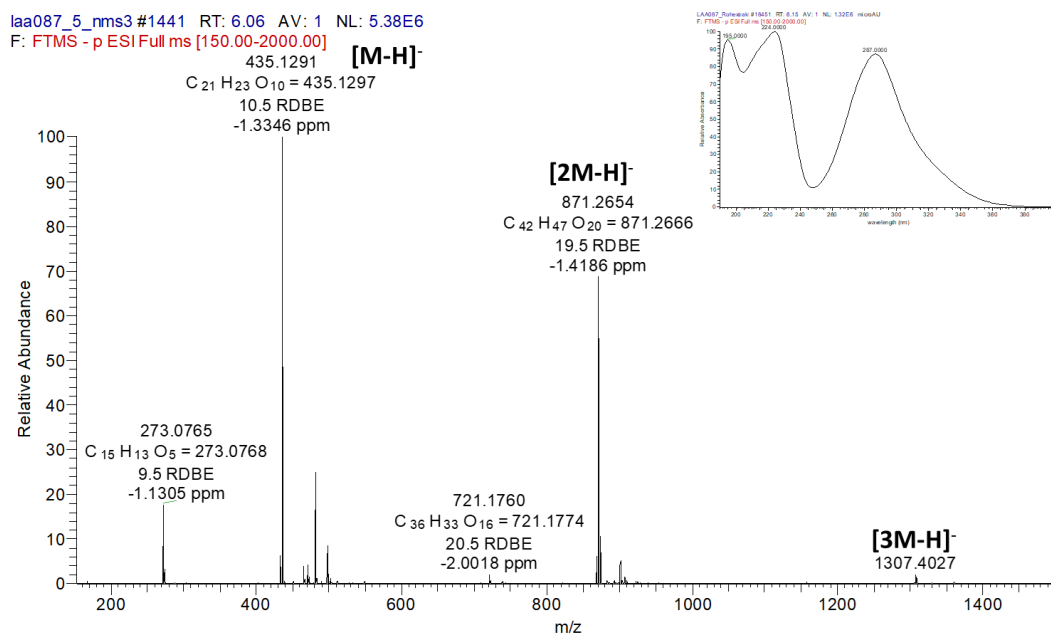
Figure A.48. Full negative ion ESI-HRMS spectrum of phlorizin (fraction 5, t_R (preparative HPLC-MS): 10.0-11.1 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa*.

laa087_5_nms3 #1423 RT: 6.17 AV: 1 NL: 2.14E6
F: FTMS - c ESI Full ms3 435.00@cid35.00 273.00@

167.0349
$C_8 H_7 O_4$ = 167.0350
5.5 RDBE
-0.6925 ppm

125.0245
$C_6 H_5 O_3$ = 125.0244
4.5 RDBE
0.2788 ppm

m/z
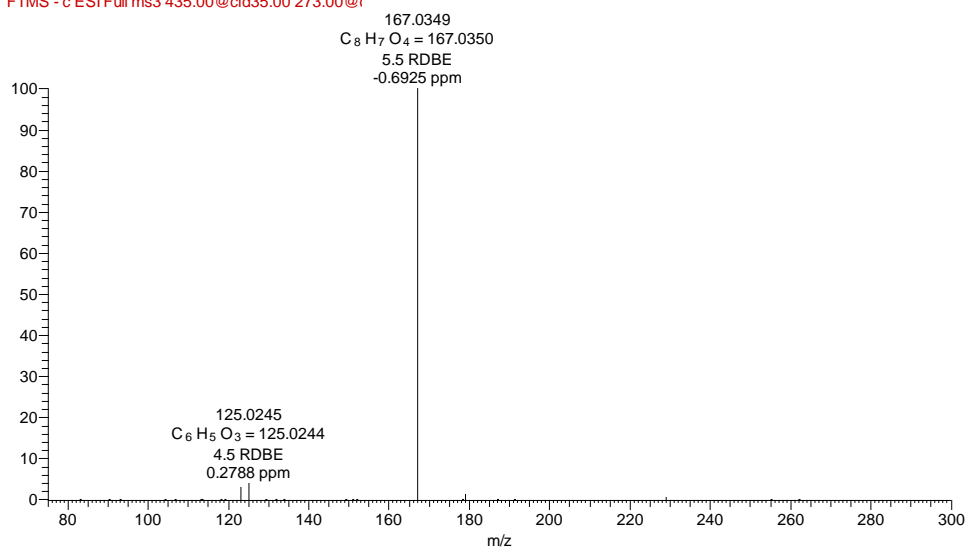
Figure A.49. Negative ion ESI-HRMS³ spectrum of *m/z* 273 (aglycone of phlorizin, fraction 5 obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa*, t$_R$ (preparative HPLC-MS): 10.0-11.1 min).

laa087_5_nms3 #1405-1434 RT: 5.97-6.04 AV: 6 NL: 2.68E4
F: FTMS - c ESI Full ms3 435.00@cid35.00 167.00@

123.0452
$C_7 H_7 O_2$
4.5 RDBE
0.0431 ppm

125.0244
$C_6 H_5 O_3$
4.5 RDBE
0.1226 ppm

151.0025
$C_7 H_3 O_4$
6.5 RDBE
-8.1218 ppm

m/z
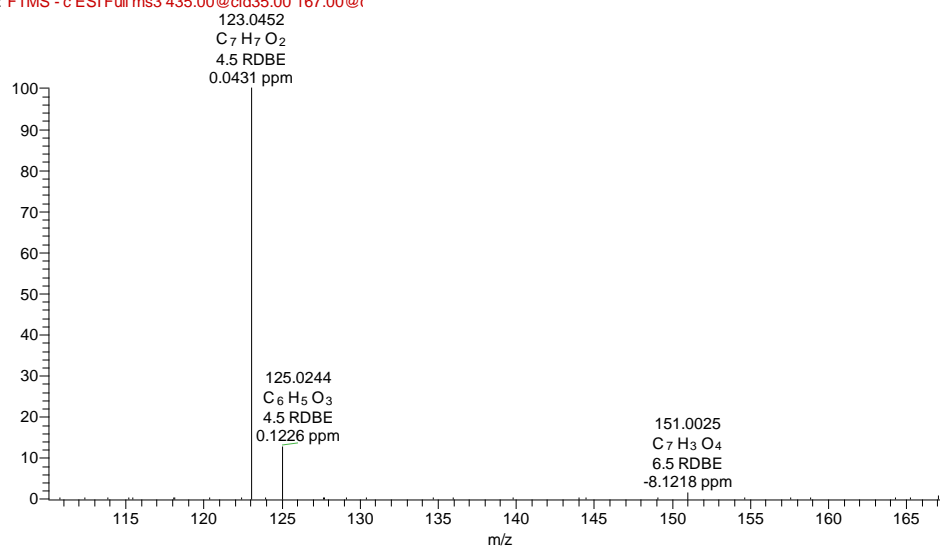
Figure A.50. Negative ion ESI-HRMS³ spectrum of *m/z* 167 (fragment ion of phlorizin, fraction 5 obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa*, t$_R$ (preparative HPLC-MS): 10.0-11.1 min).
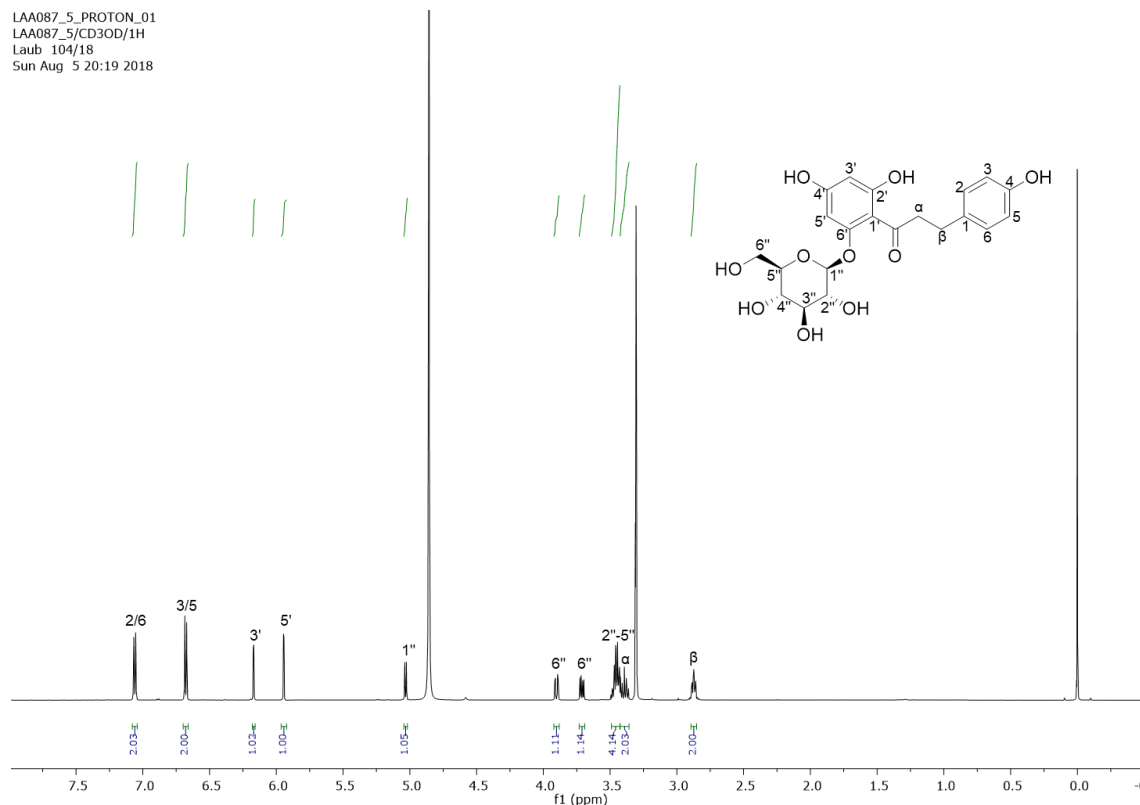
Figure A.51. ^{1}H-NMR spectrum of phlorizin (fraction 5 obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa*, t_R (preparative HPLC-MS): 10.0-11.1 min, 600 MHz, CD$_3$OD-*d4*, 25 °C).
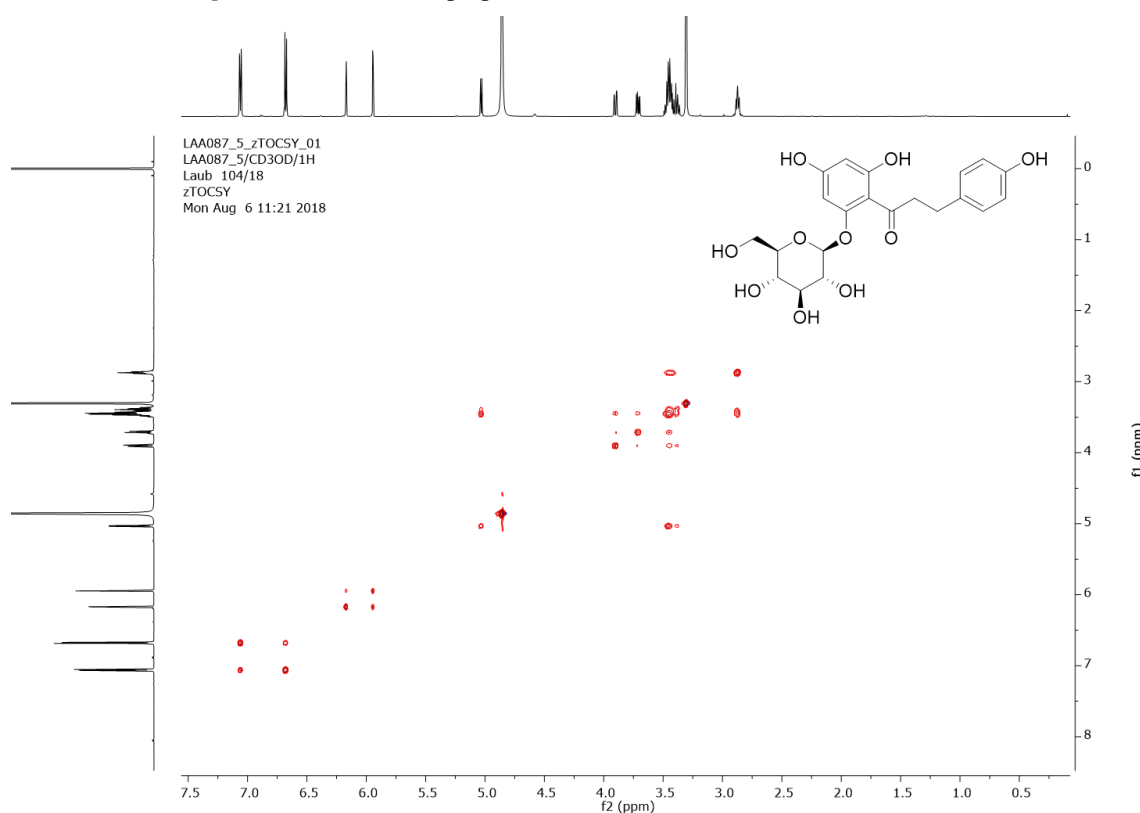
Figure A.52. [1]H,[1]H-TOCSY spectrum of phlorizin (fraction 5 obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa,* t_R (preparative HPLC-MS): 10.0-11.1 min, 600 MHz, CD$_3$OD-*d4*, 25 °C).
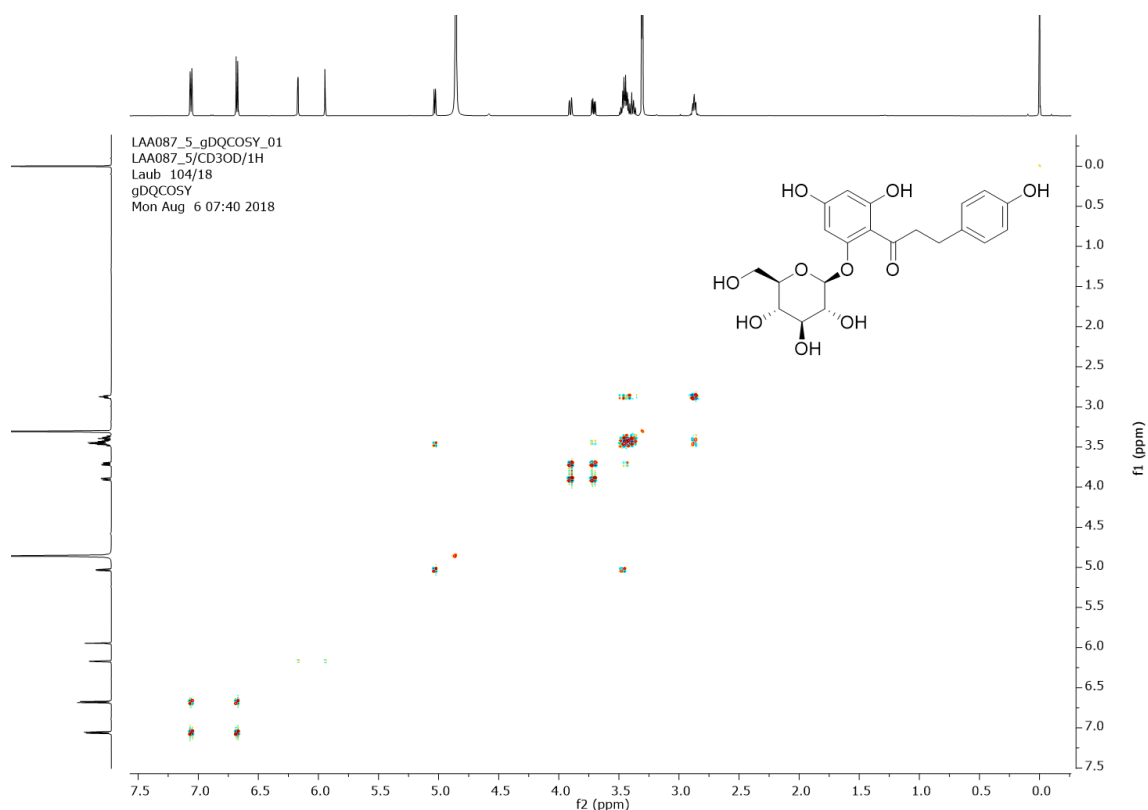
Figure A.53. ^{1}H,^{1}H COSY spectrum of phlorizin (fraction 5 obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa*, t_R (preparative HPLC-MS): 10.0-11.1 min, 600 MHz, CD$_3$OD-*d4*, 25 °C).
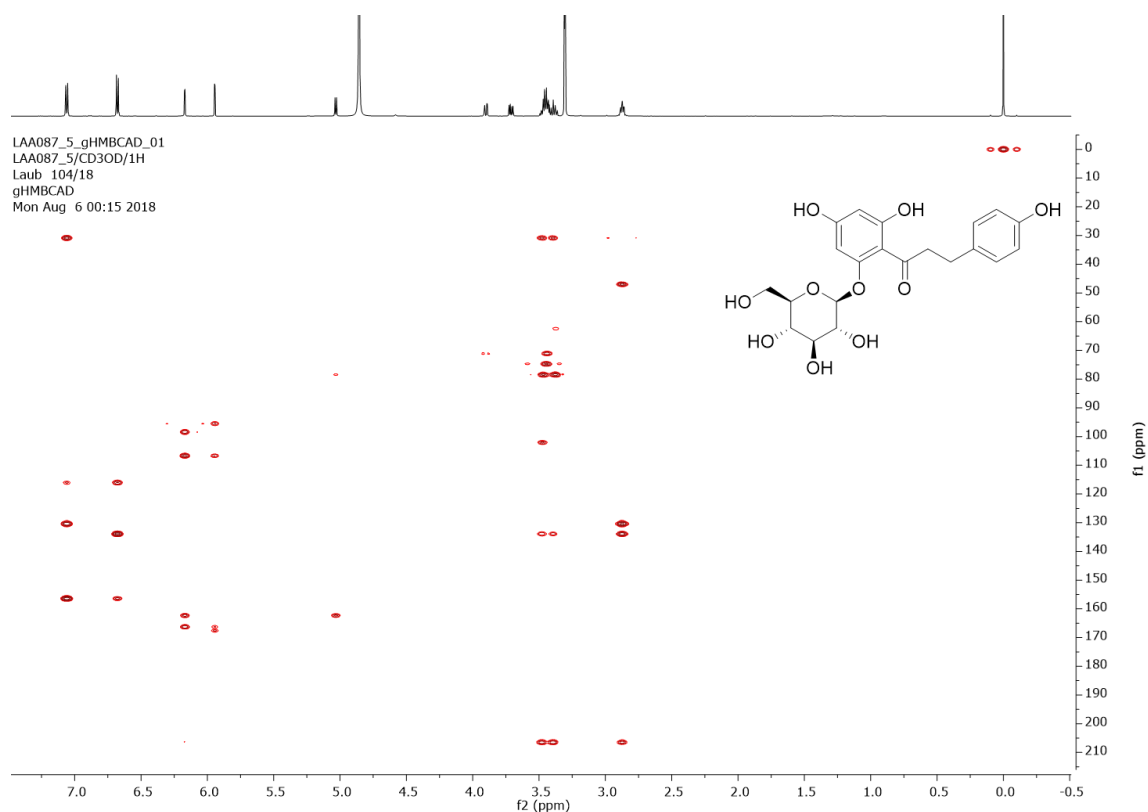
Figure A.54. ^{1}H, ^{13}C HSQC spectrum of phlorizin (fraction 5 obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa*, (t_R (preparative HPLC-MS): 10.0-11.1 min, 600 MHz, CD$_3$OD-*d4*, 25 °C).
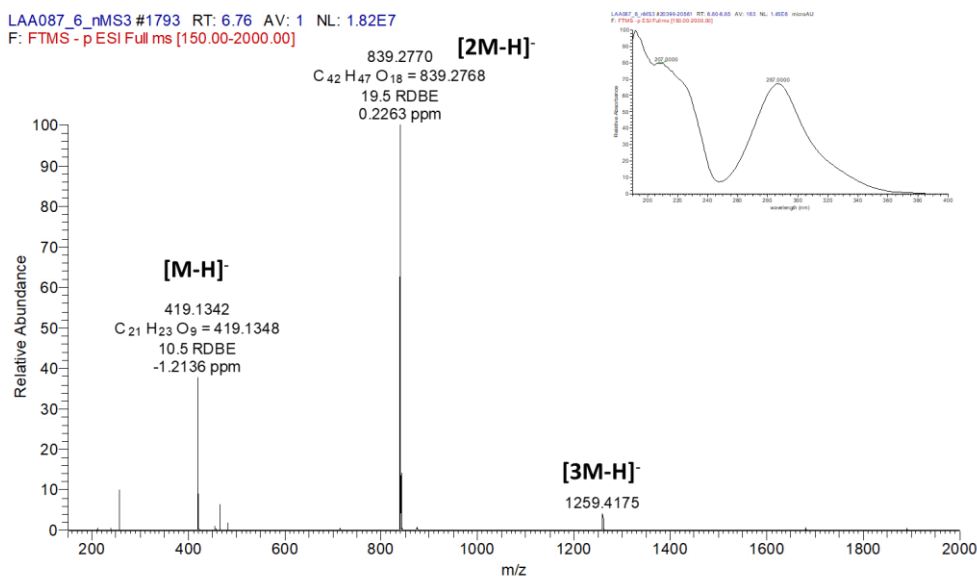
Figure A.55. Full negative ion ESI-HRMS spectrum of pinocembrin dihydrochalcone glycoside (fraction 6, t_R (preparative HPLC-MS): 12.7-13.4 min) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa*.
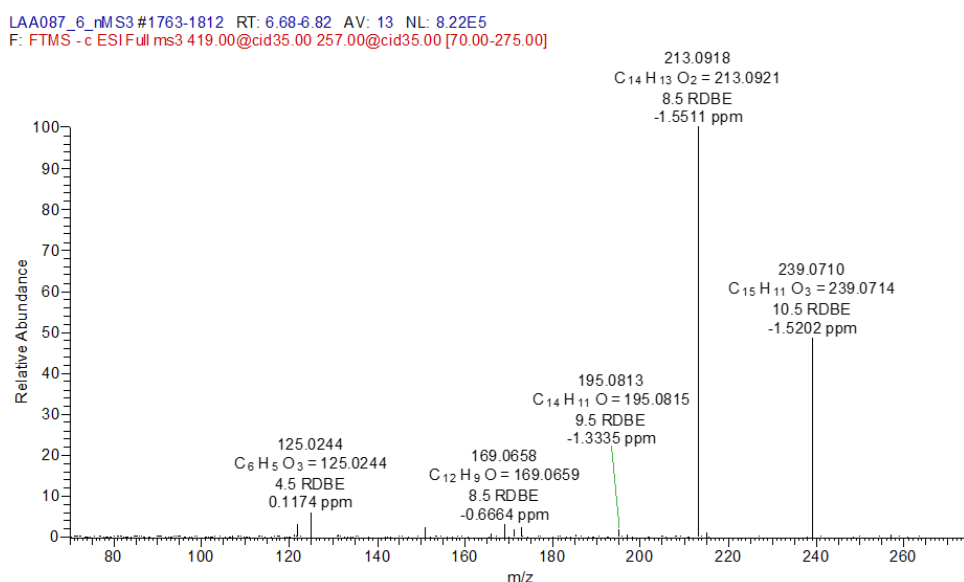
Figure A.56. Negative ion ESI-HRMS[3] spectrum of *m/z* 257 (aglycone of pinocembrin dihydrochalcone glycoside, fraction 6 obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa*, t_R (preparative HPLC-MS): 12.7-13.4 min).
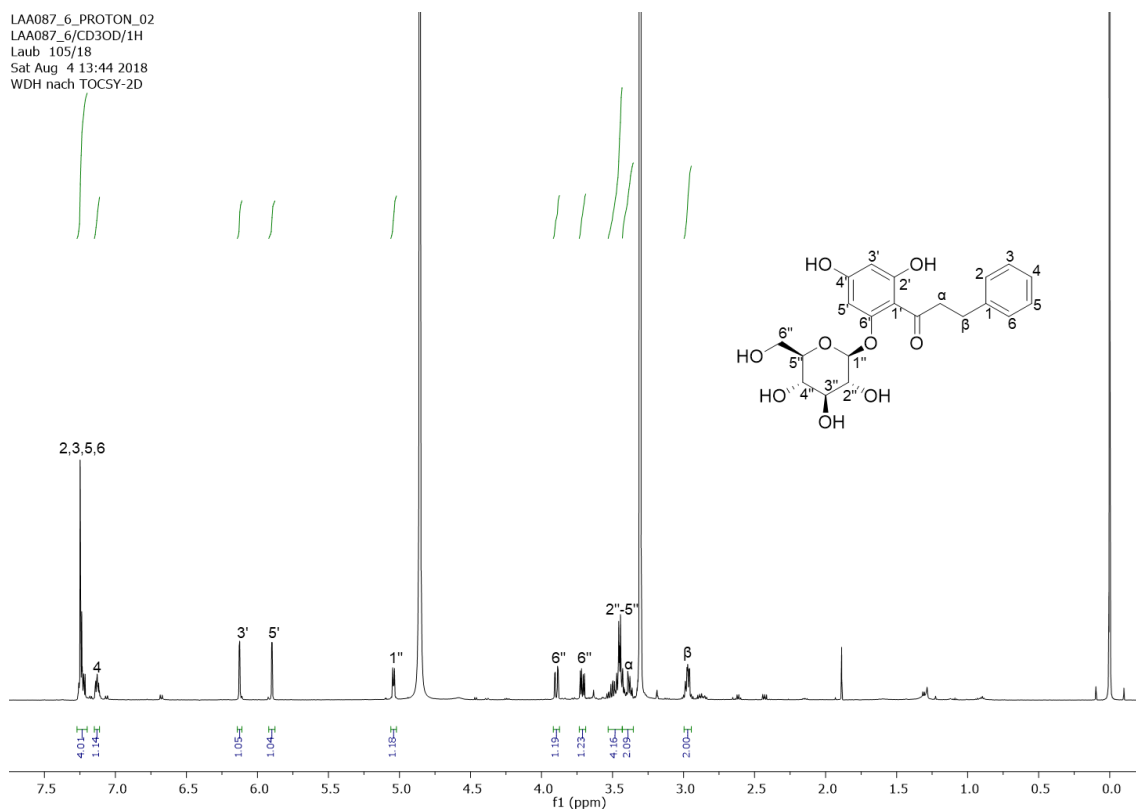
Figure A.57. ¹H-NMR spectrum of pinocembrin dihydrochalcone glycoside (fraction 6 obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa*, t_R (preparative HPLC-MS): 12.7-13.4 min, 600 MHz, CD₃OD-*d4*, 25°C).
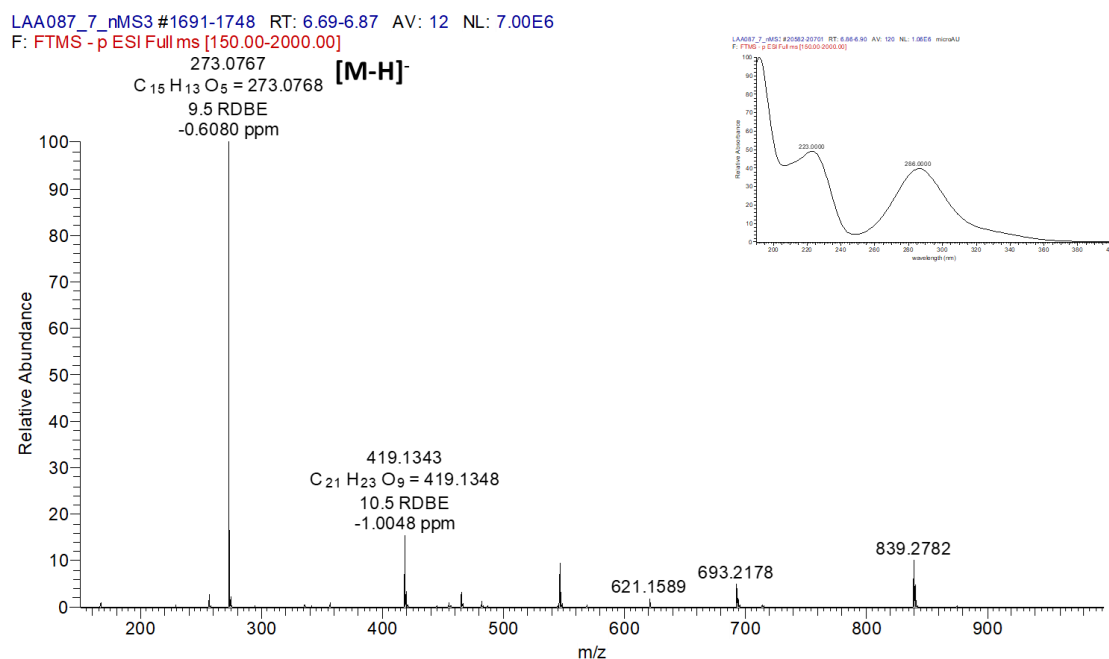
Figure A.58. Full negative ion ESI-HRMS spectrum of phloretin (fraction 7) obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa*.

LAA087_7_nMS3 #1712-1732 RT: 6.77-6.81 AV: 4 NL: 2.84E6
F: FTMS - c ESI Full ms3 273.00@cid35.00 167.00@cid35.00 [50.00-175.00]

123.0452
$C_7H_7O_2 = 123.0452$
4.5 RDBE
0.3465 ppm

151.0037
$C_7H_3O_4 = 151.0037$
6.5 RDBE
0.2487 ppm
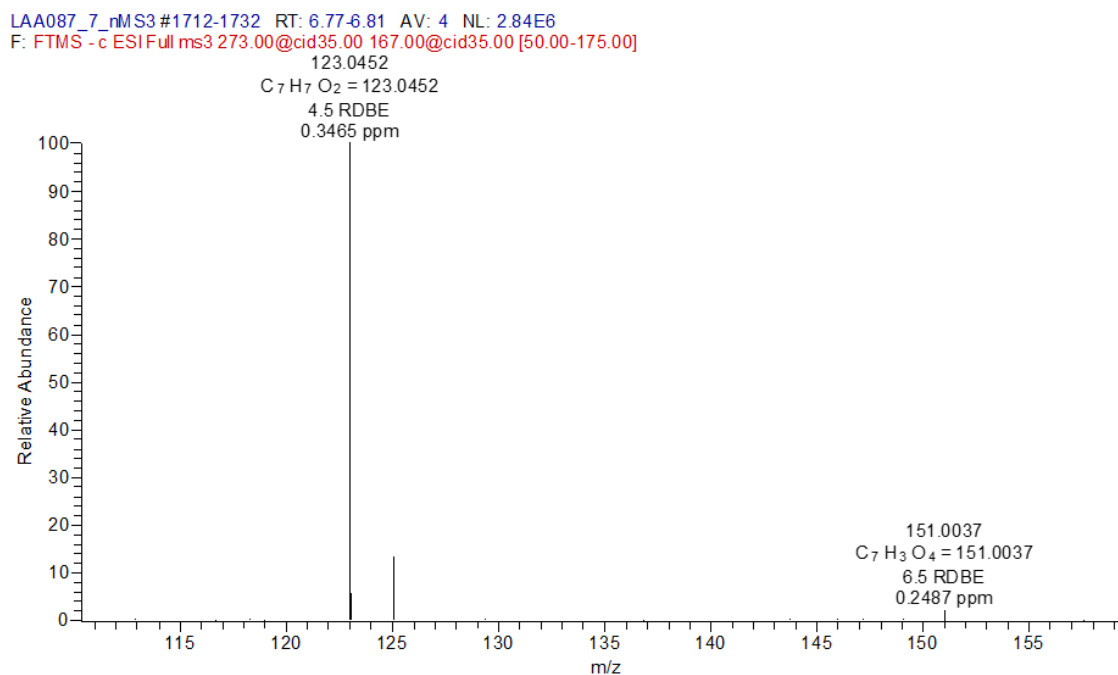
Figure A.59. Negative ion ESI-HRMS[3] spectrum of *m/z* 167 (fragment ion of phloretin, fraction 7, t_R (preparative HPLC-MS): 14.0-14.9 min).

LAA087_7_PROTON_01
LAA087_7/CD3OD/1H
Laub 106/18
Mon Aug 6 12:38 2018
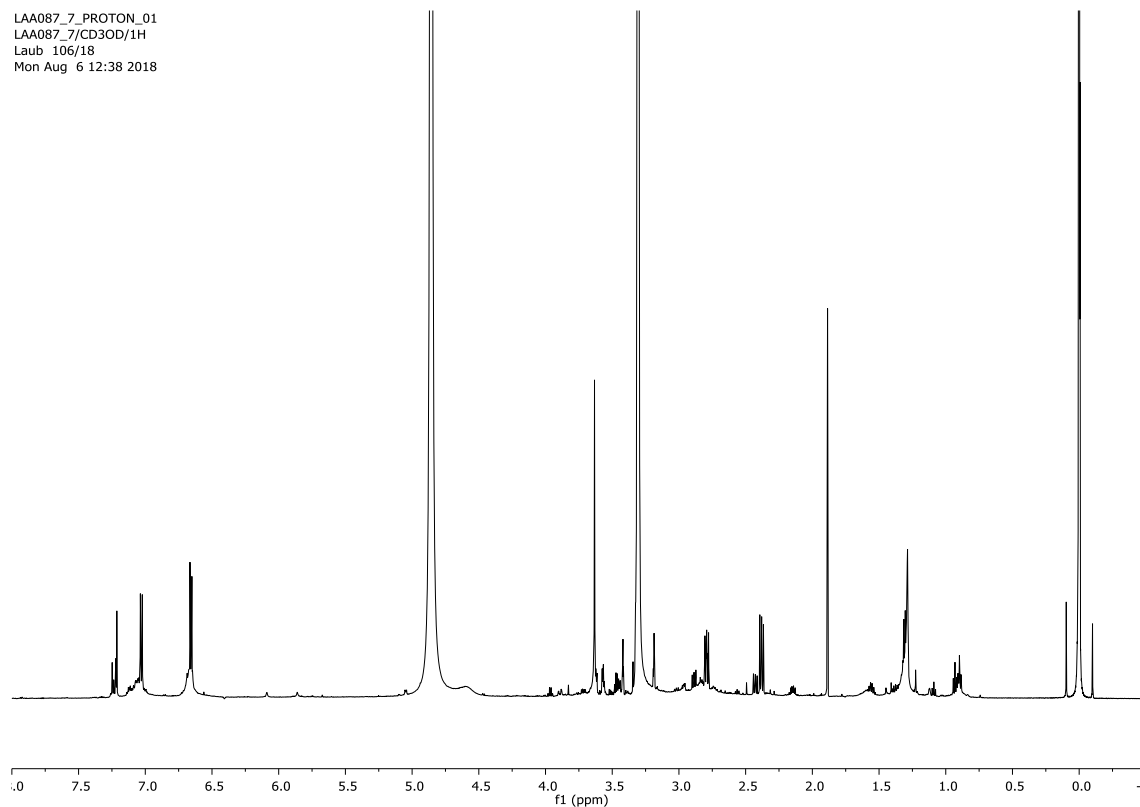
Figure A.60. ^{1}H-NMR spectrum of fraction 7 obtained after chromatographic separation of the methanolic crude leaf extracts of *Impatiens racemosa* (t_R (preparative HPLC-MS): 14.0-14.9 min, 600 MHz, CD₃OD-*d4*, 25°C).
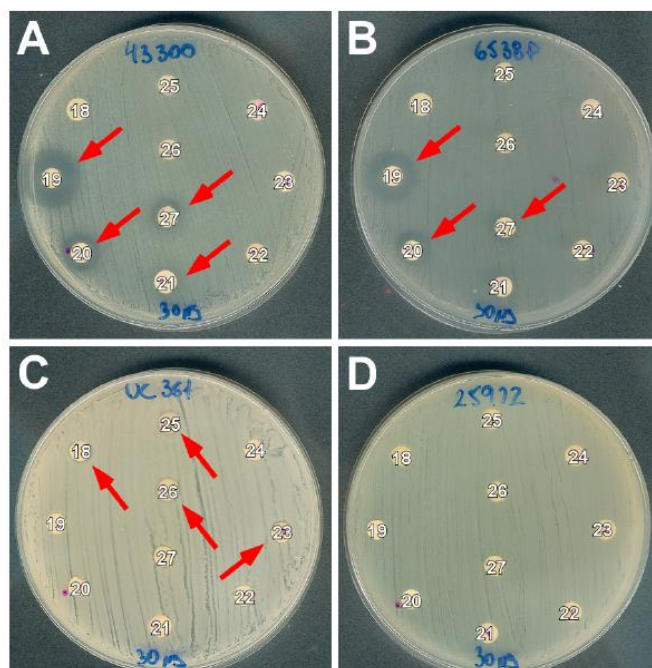
Figure A.61. Antibacterial activity assay by agar disc-diffusion method against gram-positive and gram-negative strains (A) *Staphylococcus aureus* ATCC 43300, B) *Staphylococcus aureus* ATCC 6538P, C) *Klebsiella pneumoniae* UC361, D) *Escherichia coli* ATCC 25922) for the methanolic leaf extracts of *I. balsamina* (18), *I. ethiopica* (19), *I. flanaganae* (20), *I. glandulifera* (21), *I. niamniamensis* (22), *I. noli-tangere* (23), *I. sodenii* (24), *I. species* (No. 36248) (25), *I. textori* (26), *I. tinctoria* (27). The discs were loaded with 30 µg of each extract. The active extracts are indicated with a red arrow.
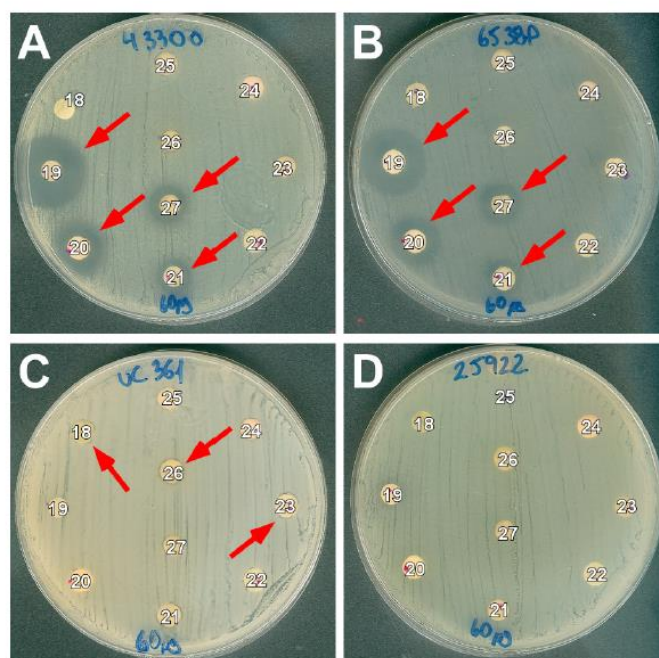
Figure A.62. Antibacterial activity assay by agar disc-diffusion method against gram-positive and gram-negative strains (A) *Staphylococcus aureus* ATCC 43300, B) *Staphylococcus aureus* ATCC 6538P, C) *Klebsiella pneumoniae* UC361, D) Escherichia coli ATCC 25922) for the methanolic leaf extracts of *I. balsamina* (18), *I. ethiopica* (19), *I. flanaganae* (20), *I. glandulifera* (21), *I. niamniamensis* (22), *I. noli-tangere* (23), *I. sodenii* (24), *I. species* (No. 36248) (25), *I. textori* (26), *I. tinctoria* (27). The discs were loaded with 60 µg of each extract. The active extracts are indicated with a red arrow.
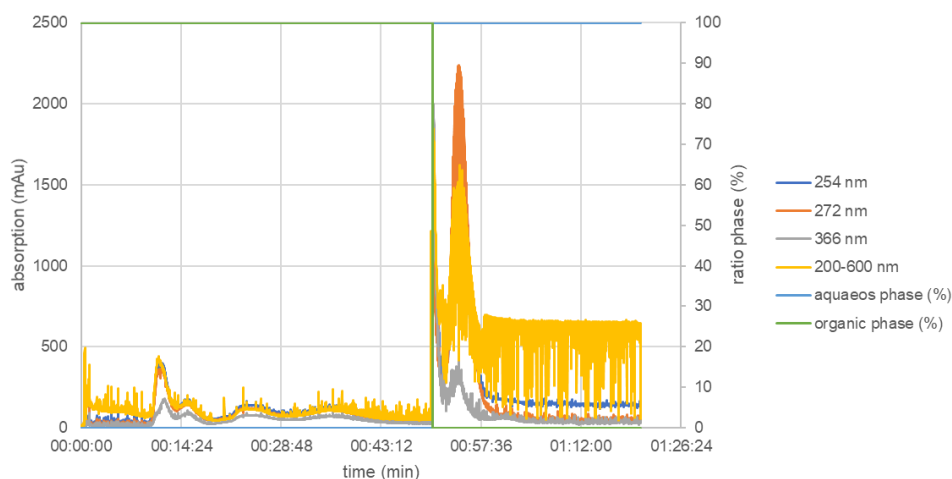
Figure A.63. SCPC chromatogram of crude methanolic leaf extracts of *Impatiens ethiopica*.
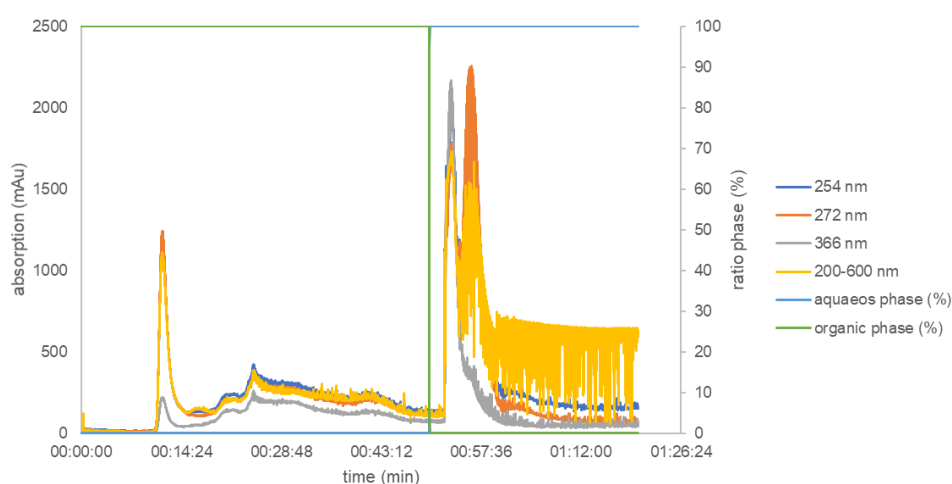
Figure A.64. SCPC chromatogram of crude methanolic leaf extracts of *Impatiens flanaganae*.
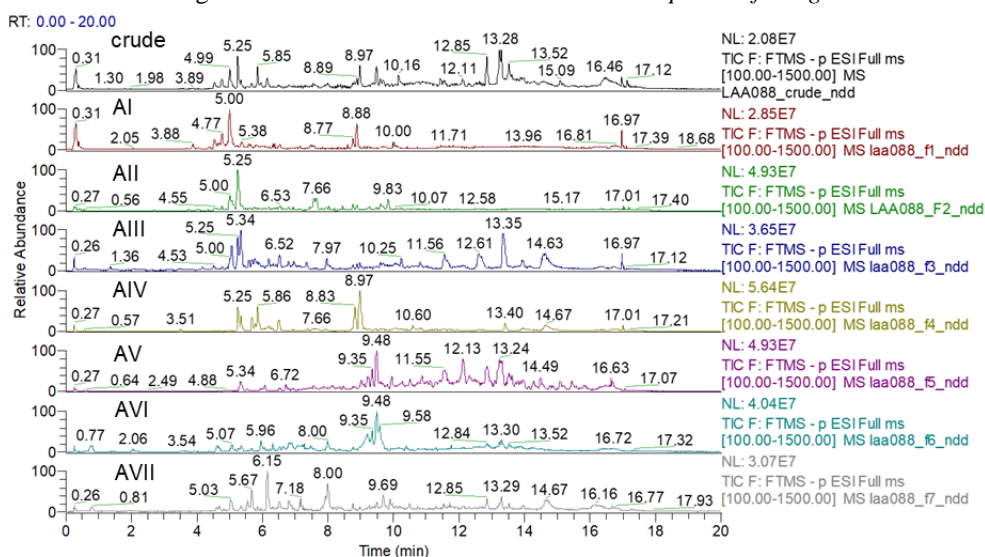
Figure A.65. Comparison of the total ion chromatograms of the crude methanolic extract of *Impatiens ethiopica* and the obtained fractions (AI-AVII) after SCPC separation recorded via UHPLC-PDA-ESI-HRMS in negative ion mode.
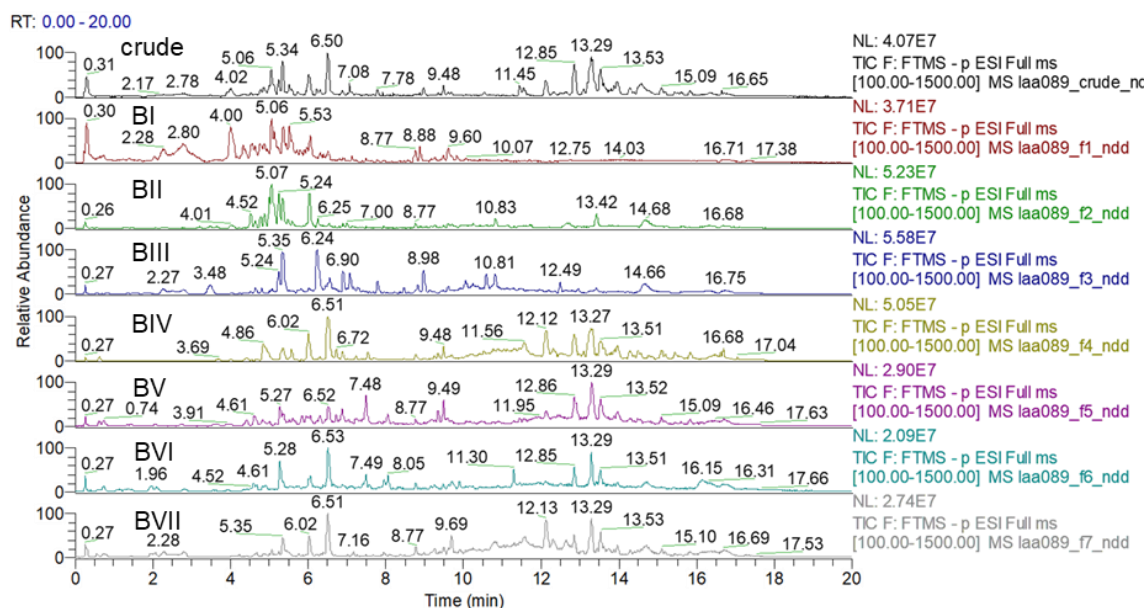
Figure A.66. Comparison of the total ion chromatograms of the crude methanolic extract of *Impatiens flanaganae* and the obtained fractions (BI-BVII) after SCPC separation recorded via UHPLC-PDA-ESI-HRMS in negative ion mode.
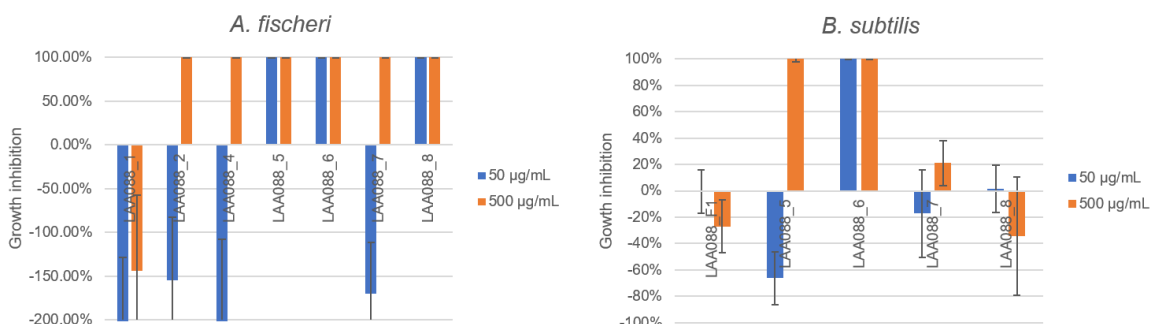
Figure A.67. Antibacterial activity of SCPC fractions obtained from *Impatiens ethiopica*. Negative values (=growth) are promoted by nutritional compounds in crude extracts.
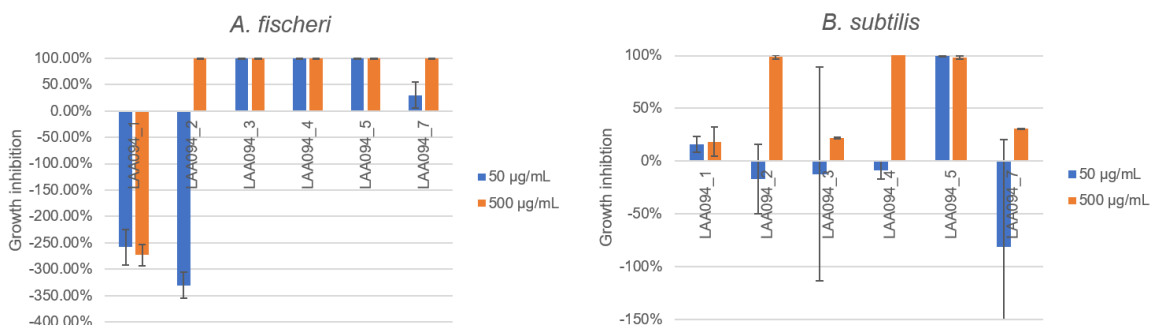
Figure A.68. Antibacterial activity of SCPC fractions obtained from *Impatiens flanaganae*. Negative values (=growth) are promoted by nutritional compounds in crude extracts.
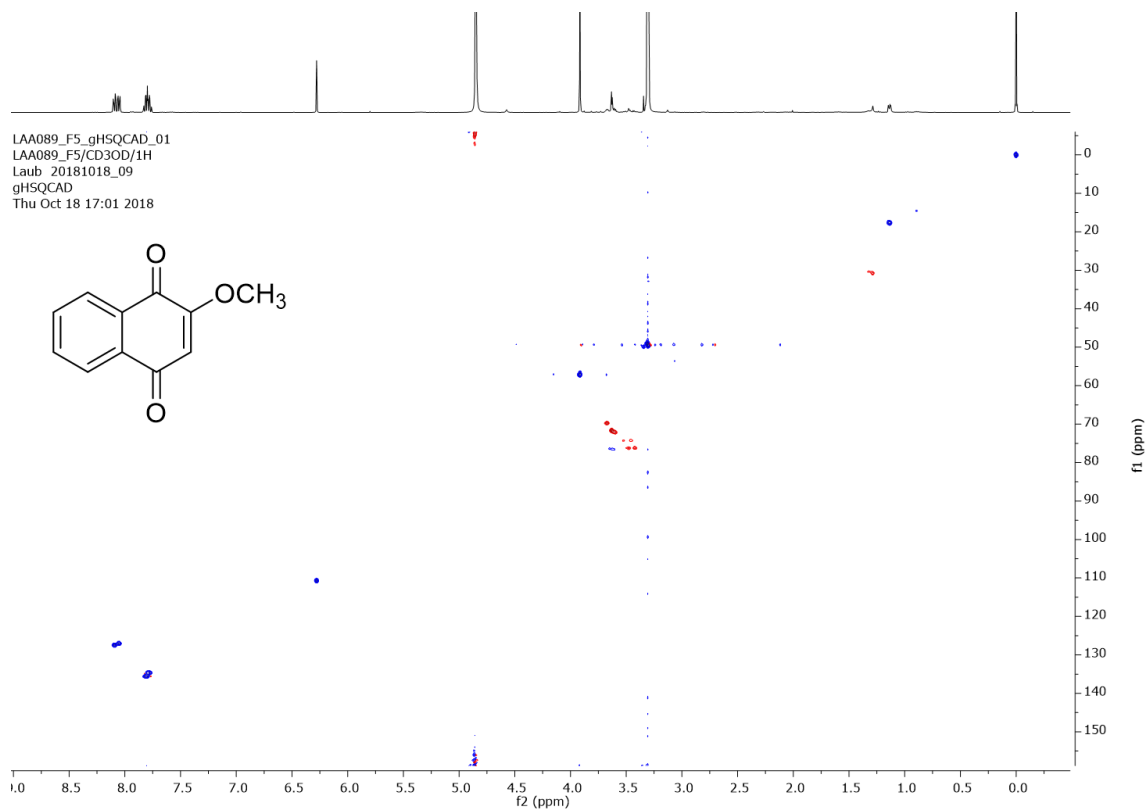
Figure A.69. ¹H, ¹³C HSQC spectrum of 2-methoxy-1,4-napthoquinone (fraction BV obtained after SCPC separation of the methanolic crude leaf extracts of *Impatiens flanaganae* (600 MHz, CD₃OD-*d4*, 25 °C).
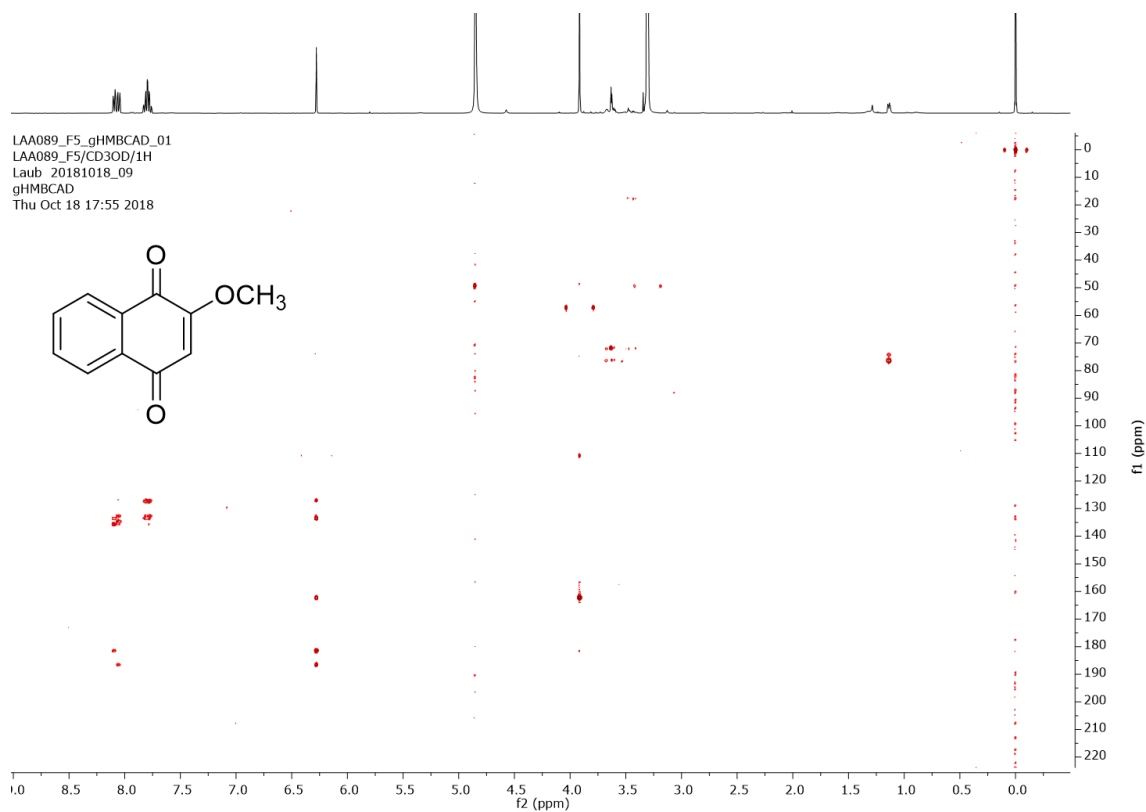
Figure A.70. ¹H, ¹³C HMBC spectrum of 2-methoxy-1,4-napthoquinone (fraction BV obtained after SCPC separation of the methanolic crude leaf extracts of *Impatiens flanaganae* (600 MHz, CD₃OD-*d4*, 25 °C).
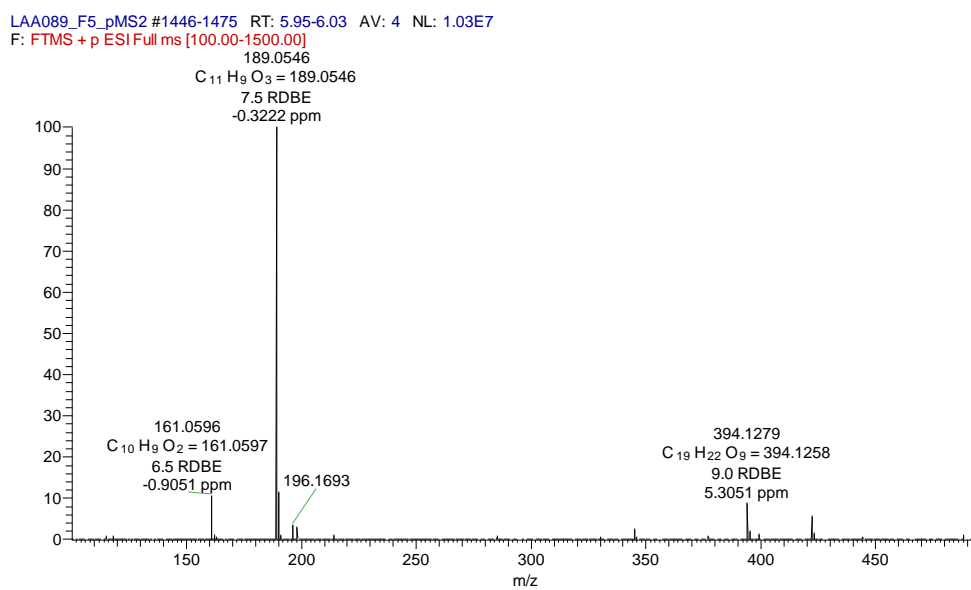
Figure A.71. Positive ion ESI-HRMS spectrum of 2-methoxy-1,4-naphthoquinone (*m/z* 189).

B.   Appendix Chapter 3

Table B.1. (+)-ESI-HRMS[n] data of natural cyclic pentapeptides (**3.1**-**3.3**) detected within the enriched extract of *S. microspermum*.

cpd	t_R [min]	scan mode [*m/z* (NCE in %)]	*m/z* (relative intensity in %, error in ppm)
3.1a	14.73	(+) full MS [a]	496.3491 (**[M+H]**[+], $C_{25}H_{46}O_5N_5^+$, -0.6)
3.1a	14.73	(+) MS[2] **[496.0 (45)]**[a]	478.3395 ([M+H-H₂O]⁺, 8, 1.5), 468.3542 ([M+H-CO]⁺, 66, -0.4), 451.3276 ([M+H-CO-NH₃]⁺, 39, -0.5), 397.2805 (**b₄ₐᵥ/b₄ᵥᵥ**, **100**, -1.0), 383.2654 (**b₄ₗₗ/b₄ᵥₗ**, 44, 0.3), 369.2860 (**a₄ₐᵥ/a₄ᵥᵥ**, 22, 0.1), 355.2707 (**a₄ₗₗ/a₄ᵥₗ**, 13, 1.0), 326.2433 (**b₃ₗₐ**, 3, -1.6), 312.2274 (**b₃ₗₗ**, 4, -2.3), 298.2126 (**b₃ₐᵥ**, 13, 0.1), 284.1970 (**b₃ᵥᵥ**, 11, 0.4), 270.1813 (**b₃ᵥₗ**, 4, 0.2), 227.1760 (**b₂ₗₐ**, 1, 2.5), 213.1600 (**b₂ₗₗ**, 2, 1.0), 199.1804 (**a₂ₗₐ**, 1, -0.4), 199.1441 (**b₂ᵥₗ**, 1, 0.1), 185.1654 (**a₂ₗₗ**, 1, 2.9), 185.1288 (**b₂ₐᵥ**, 1, 1.9), 171.1504 (**a₂ᵥₗ**, 0.4, 5), 171.1132 (**b₂ᵥᵥ**, 1, 2.5), 157.1335 (**a₂ₐᵥ**, 1, -0.3)
		(+) MS[3] [496.0 (45) -> **397.0 (45)**][b]	369.2863 (**a₄ₐᵥ/a₄ᵥᵥ**, **100**, 0.9), 352.2598 ([**a₄ₐᵥ/a₄ᵥᵥ**-NH₃]⁺, 7, 1.0), 326.2433 ([**b₄ₐᵥ**-Ala]⁺, 8, -1.6), 298.2128 (**b₃ₐᵥ**/[**b₄ᵥᵥ**-Val]⁺, 93, 1.0), 284.1972 (**b₃ᵥᵥ**, 52, 1.2), 185.1288 (**b₂ₐᵥ**, 4, 2.0), 157.1337 (**a₂ₐᵥ**, 1, 0.8)
		(+) MS[3] [496.0 (45) -> **383.0 (45)**][b]	355.2710 (**a₄ₗₗ/a₄ᵥₗ**, **100**, 1.8), 312.2302 (**b₃ₗₗ**, 15, 6.0), 284.1978 ([**b₄ᵥₗ**-Val]⁺, 84, 3.2), 270.1823 (**b₃ᵥₗ**/[**b₄ₗₗ**-Leu]⁺, 31, 3.8)
3.1b	14.80	(+) full MS [a]	496.3493 (**[M+H]**[+], $C_{25}H_{46}O_5N_5^+$, -0.2)
3.1b	14.80	(+) MS[2] **[496.0 (45)]**[a]	478.3398 ([M+H-H₂O]⁺, 8, 2.0), 468.3545 ([M+H-CO]⁺, 69, 0.2), 451.3279 ([M+H-CO-NH₃]⁺, 40, 0.1), 397.2807 (**b₄ₐᵥ/b₄ᵥᵥ**, **100**, -0.5), 383.2656 (**b₄ₗₗ/b₄ᵥₗ**, 50, 0.7), 369.2863 (**a₄ₐᵥ/a₄ᵥᵥ**, 20, 0.7), 355.2706 (**a₄ₗₗ/a₄ᵥₗ**, 10, 0.7), 326.2433 (**b₃ₗₐ**, 3, -1.6), 312.2285 (**b₃ₗₗ**, 4, 1.2), 298.2126 (**b₃ₐᵥ**, 13, 0.1), 284.1971 (**b₃ᵥᵥ**, 26, 1.0), 270.1817 (**b₃ᵥₗ**, 2, 1.9), 227.1755 (**b₂ₗₐ**, 0.3, 0.3), 213.1600 (**b₂ₗₗ**, 3, 1.4), 185.1650 (**a₂ₗₗ**, 2, 0.7), 185.1290 (**b₂ₐᵥ**, 2, 3.1), 171.1143 (**b₂ᵥᵥ**, 0.4, 9.0), 157.1339 (**a₂ₐᵥ**, 1, 2.3)
		(+) MS[3] [496.0 (45) -> **397.0 (45)**][b]	369.2861 (**a₄ₐᵥ/a₄ᵥᵥ**, 93, 0.2), 352.2604 ([**a₄ₐᵥ/a₄ᵥᵥ**-NH₃]⁺, 7, 2.7), 326.2436 ([**b₄ₐᵥ**-Ala]⁺, 9, -0.6), 298.2126 (**b₃ₐᵥ**/[**b₄ᵥᵥ**-Val]⁺, 42, 0.4), 284.1970 (**b₃ᵥᵥ**, **100**, 0.4), 185.1285 (**b₂ₐᵥ**, 5, 0.5), 157.1334 (**a₂ₐᵥ**, 1, 5.0)
		(+) MS[3] [496.0 (45) -> **383.0 (45)**][b]	355.2706 (**a₄ₗₗ/a₄ᵥₗ**, **100**, 0.7), 312.2285 (**b₃ₗₗ**, 21, 0.9), 284.1967 ([**b₄ᵥₗ**-Val]⁺, 69, -0.4), 270.1815 (**b₃ᵥₗ**/[**b₄ₗₗ**-Leu]⁺, 23, 1.1)
3.1c	14.98	(+) full MS [a]	496.3486 (**[M+H]**[+], $C_{25}H_{46}O_5N_5^+$, -1.5)
3.1c	14.98	(+) MS[2] **[496.0 (45)]**[a]	478.3400 ([M+H-H₂O]⁺, 11, 2.2), 468.3542 ([M+H-CO]⁺, 84, -0.5), 451.3276 ([M+H-CO-NH₃]⁺, 41, -0.6), 397.2805 (**b₄ₐᵥ/b₄ᵥᵥ**, **100**, -1.1), 383.2654 (**b₄ₗₗ/b₄ᵥₗ**, 50, 0.2), 369.2862 (**a₄ₐᵥ/a₄ᵥᵥ**, 16, 0.5), 355.2705 (**a₄ₗₗ/a₄ᵥₗ**, 12, 0.3), 326.2438 (**b₃ₗₐ**, 5, -0.1), 312.2291 (**b₃ₗₗ**, 1, 3.0), 298.2124 (**b₃ₐᵥ**, 1, -0.5), 284.1969 (**b₃ᵥᵥ**, 18, 0.2), 270.1814 (**b₃ᵥₗ**, 5, 0.7), 227.1753 (**b₂ₗₐ**, 2.3, -0.6), 213.1602 (**b₂ₗₗ**, 2, 2.2), 199.1803 (**a₂ₗₐ**, 1, -1.0), 185.1661 (**a₂ₗₗ**, 2, 0.7), 171.1149 (**b₂ᵥᵥ**, 0.8)
		(+) MS[3] [496.0 (45) -> **397.0 (45)**][b]	369.2862 (**a₄ₐᵥ/a₄ᵥᵥ**, 84, 0.5), 352.2584 ([**a₄ₐᵥ/a₄ᵥᵥ**-NH₃]⁺, 6, -3.0), 326.2444 ([**b₄ₐᵥ**-Ala]⁺, 15, 1.8), 298.2134 (**b₃ₐᵥ**/[**b₄ᵥᵥ**-Val]⁺, 22, 2.9), 284.1971 (**b₃ᵥᵥ**, **100**, 0.9)
		(+) MS[3] [496.0 (45) -> **383.0 (45)**][b]	355.2703 (**a₄ₗₗ/a₄ᵥₗ**, **100**, -0.3), 312.2287 (**b₃ₗₗ**, 5, 1.7), 284.1970 ([**b₄ᵥₗ**-Val]⁺, 44, 0.3), 270.1815 (**b₃ᵥₗ**/[**b₄ₗₗ**-Leu]⁺, 34, 0.9)
3.2	14.50	(+) full MS	482.3335 (**[M+H]**[+], $C_{24}H_{44}O_5N_5^+$, -0.5)
3.2	14.50	(+) MS[2] **[482.0 (45)]**[a]	464.3237 ([M+H-H₂O]⁺, 9, 0.6), 454.3380 ([M+H-CO]⁺, **100**, -1.6), 437.3116 ([M+H-CO-NH₃]⁺, 42, -0.7), 411.2967 (**b₄ₐₐ/b₄ₗₐ**, 18, -2.3), 383.3026 (**a₄ₐₐ/a₄ₗₐ**, 3, 2.3), 369.2493 (**b₄ₗₗ/b₄ₗₗ*/b₄ₐₗ**, 95, -0.3), 341.2544 (**a₄ₗₗ/a₄ₗₗ*/a₄ₐₗ**, 21, -0.3), 340.2592 (**b₃ₗₐ**, 1, -0.2), 324.2279 ([**a₄ₗₗ/a₄ₗₗ*/a₄ₐₗ**-NH₃]⁺, 7, -0.2), 298.2126 (**b₃ₐₐ/b₃ₗₗ**, 6, 0.1), 256.1655 (**b₃ₗₗ*/b₃ₐₗ**, 13, -0.1), 227.1753 (**b₂ₗₐ/b₂ₗₗ**, 2, -0.1), 199.1818 (**a₂ₗₐ/a₂ₗₗ**, 1, 1.3), 185.1284 (**b₂ₐₐ/b₂ₗₗ***, 1, -0.1), 157.1356 (**a₂ₐₗ/a₂ₗₗ**, 1, 2.1)
3.2		(+) MS[3] [482.0 (45) -> **411.0 (45)**][b]	383.3012 (**a₄ₐₐ/a₄ₗₐ**, 75, -1.2), 340.2593 (**b₃ₗₐ**/[**b₄ₐₐ**-Ala]⁺, 6, -0.5), 298.2128 (**b₃ₐₐ**/[**b₄ₗₐ**-Leu]⁺, **100**, 1.1)
3.2		(+) MS[3] [482.0 (45) -> **369.0 (45)**][b]	341.2544 (**a₄ₐₗ/a₄ₗₗ/a₄ₗₗ***, **100**, -1.0), 324.2283 ([**a₄ₐₗ/a₄ₗₗ/a₄ₗₗ***-NH₃]⁺, 9, 0.5), 298.2125 (**b₃ₗₗ**/[**b₄ₐₗ**-Ala]⁺, 19, -0.1), 256.1655 (**b₃ₐₗ/b₃ₗₗ***/[**b₄ₗₗ/b₄ₗₗ***-Leu]⁺, 72, -0.4), 211.1440 ([**a₃ₐₗ/a₃ₗₗ***-NH₃]⁺, 6, 3.0)

Table B.1 (continue)

cpd	t_R [min]	scan mode [*m/z* (NCE in %)]	*m/z* (relative intensity in %, error in ppm)
3.3	15.58	(+) full MS	510.3650 ([**M+H**]⁺, $C_{26}H_{48}O_5N_5^+$, 0.1)
3.3	15.58	(+) MS² **[510.0 (45)]**	492.3546 ([M+H-H₂O]⁺, 10, 0.2), 482.3702 ([M+H-CO]⁺, **100**, 0.1), 465.3436 ([M+H-CO-NH₃]⁺, 61, 0.1), 411.2964 (**b₄ₐᵥ**, 92, -0.2), 397.2809 (**b₄ₗₗ/b₄ₗₗ*/b₄ᵥₗ**, 96, 0.1), 383.3021 (**a₄ₐᵥ**, 22, 0.4), 369.2863 (**a₄ₗₗ/a₄ₗₗ*/a₄ᵥₗ**, 29, 0.3), 366.2755 ([**a₄ₐᵥ**-NH₃]⁺, 4, 0.4), 352.2597 ([**a₄ₗₗ/a₄ₗₗ*/a₄ᵥₗ**-NH₃]⁺, 16, 0.3), 340.2599 (**b₃ₗₐ**, 3, 0.4), 326.2437 (**b₃ₗₗ**, 7, -0.2), 298.2128 (**b₃ₐᵥ**, 33, 0.3), 284.1971 (**b₃ₗₗ*/b₃ᵥₗ**, 18, 0.2), 227.1758 (**b₂ₗₐ/b₂ₗₗ**, 5, 0.4), 213.1601 (**b₂ₗₗ***, 2, 0.3), 199.1809 (**a₂ₗₗ/a₂ₗₐ**, 2, 0.4), 185.1650 (**a₂ₗₗ***, 1, 0.1), 185.1287 (**b₂ₐᵥ**, 2, 0.3), 171.1129 (**b₂ᵥₗ**, 0.7, 0.1), 157.1336 (**a₂ₐᵥ**, 0.9, 0.1)
3.3		(+) MS³ [510.0 (45) -> 411.0 (45)]	383.3018 (**a₄ₐᵥ**, 55, 0.2), 366.2753 ([**a₄ₐᵥ**-NH₃]⁺, 5, 0.2), 340.2587 ([**b₄ₐᵥ**-Ala]⁺, 7, 0.1), 298.2127 (**b₃ₐᵥ**, **100**, 0.2), 253.1913 ([**b₃ₐᵥ**-Ala]⁺, 8, 0.2), 185.1282 (**b₂ₐᵥ**, 5, 0.2), 157.1334 (**a₂ₐᵥ**, 1.5, -0.1)
3.3		(+) MS³ [510.0 (45) -> 397.0 (45)]	369.2861 (**a₄ₗₗ/a₄ₗₗ*/a₄ᵥₗ**, **100**, 0.1), 352.2598 ([**a₄ₗₗ/a₄ₗₗ*/a₄ᵥₗ**-NH₃]⁺, 7, 0.3), 326.2438 (**b₃ₗₗ**, 17, -0.1), 298.2126 ([**b₄ᵥₗ**-Val]⁺, 39, 0.1), 284.1970 (**b₃ᵥₗ/b₃ₗₗ***, 69, 0.1), 227.1755 (**b₂ₗₗ**, 3, 0.1), 213.1595 (**b₂ₗₗ***, 3, -0.2), 185.1286 ([**b₃ᵥₗ**-Val]⁺, 1, 0.2), 171.1123 (**b₂ᵥₗ**, 2, -0.5)

Table B.2. Fragment abundances of the most different fragment ions of the natural, synthetic and synthetic isomer of chrysosporide (**3.3**).

	Relative intensities (%)		
	Natural (3.3)	**Synthetic (3.3)**	**Synthetic Valine isomer epi 3.3***
t_R [min]	8.55	8.55	8.06
[M+H-H₂O]⁺	10	10	6
b₄ₐᵥ/b₄ₗₗ/b₄ₗₗ*	78	80	98
b₄ᵥₗ	80	80	92

*cyclo-(D-Ala-L-Leu-L-Leu-D-Leu-D-Val)

Table B.3. Fragment abundances of most different fragment ions natural, synthetic and synthetic isomers of microsporide A (**3.1**).

	Relative intensities (%)					
	Natural (3.1a)	**Natural (3.1b)**	**Natural (3.1c)**	**Synthetic (3.1a)**	**Synthetic isomer I***	**Synthetic isomer II****
t_R [min]	8.04	8.10	8.19	8.04	8.10	8.05
b₃ₐᵥ	8	1	1	7	8	6
b₃ᵥᵥ	6	15	15	5	5	5
b₄ₐᵥ/b₄ᵥᵥ	100	100	95	100	100	100
b₄ₗₗ/ b₄ᵥₗ	42	49	53	42	58	41
a₅.NH₃	32	34	41	32	34	30
a₅	74	78	100	73	82	74

*cyclo-(D-Ala-L-Leu-D-Leu-L-Val-L-Val); ** cyclo-(L-Ala-D-Leu-L-Leu-D-Val-L-Val)

Table B.4 *In vitro* antibacterial activity of the synthetic linear peptide **3.6** and the synthetic cyclic peptide **3.3**.

Compounds	MIC and MBC against the following bacterial strains:											
	6538P		43300		29212		19606		27853		25922	
	MIC*	MBC	MIC	MBC	MIC	MBC	MIC	MBC	MIC	MBC	MIC	MBC
3.6	>128	>128	>128	>128	>128	>128	>128	>128	>128	>128	>128	>128
3.3	>128	>128	>128	>128	>128	>128	>128	>128	>128	>128	>128	>128

*Minimum Inhibitory Concentration (MIC) and Minimum Bactericidal Concentration (MBC), expressed in µg/mL. *Staphylococcus aureus* (ATCC **6538P**), *Staphylococcus aureus* MRSA (ATCC **43300**), *Enterococcus faecalis* (ATCC **29212**), *Enterococcus* spp. (ATCC **25922**), *Acinetobacter baumannii* (ATCC **19606**), and *Pseudomonas aeruginosa* (ATCC **27853**).
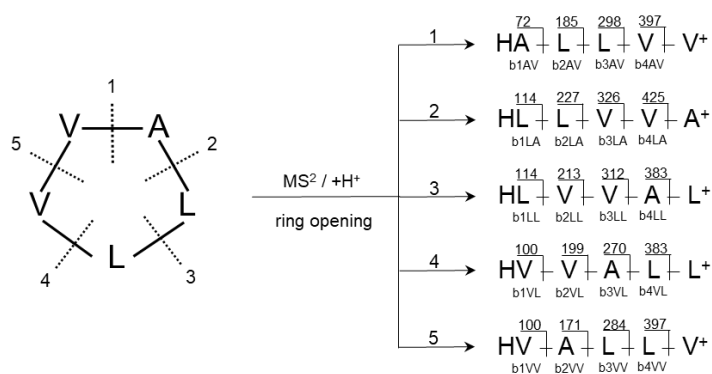
Scheme B.1. Theoretical **b**-ions (m/z) formed during fragmentation of microsporide A (**3.1a**).
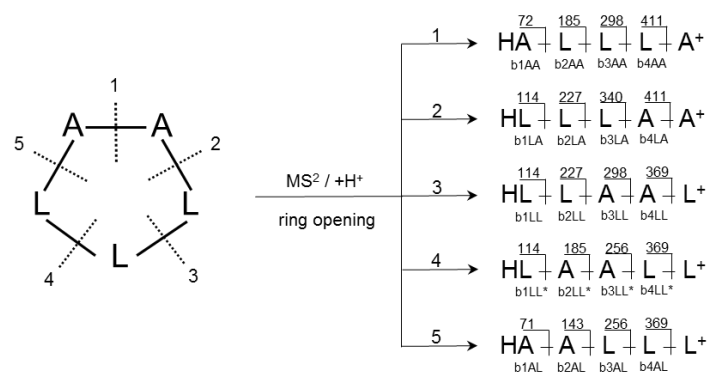
Scheme B.2. Theoretical **b**-ions (m/z) formed during fragmentation of microsporide B (**3.2**).
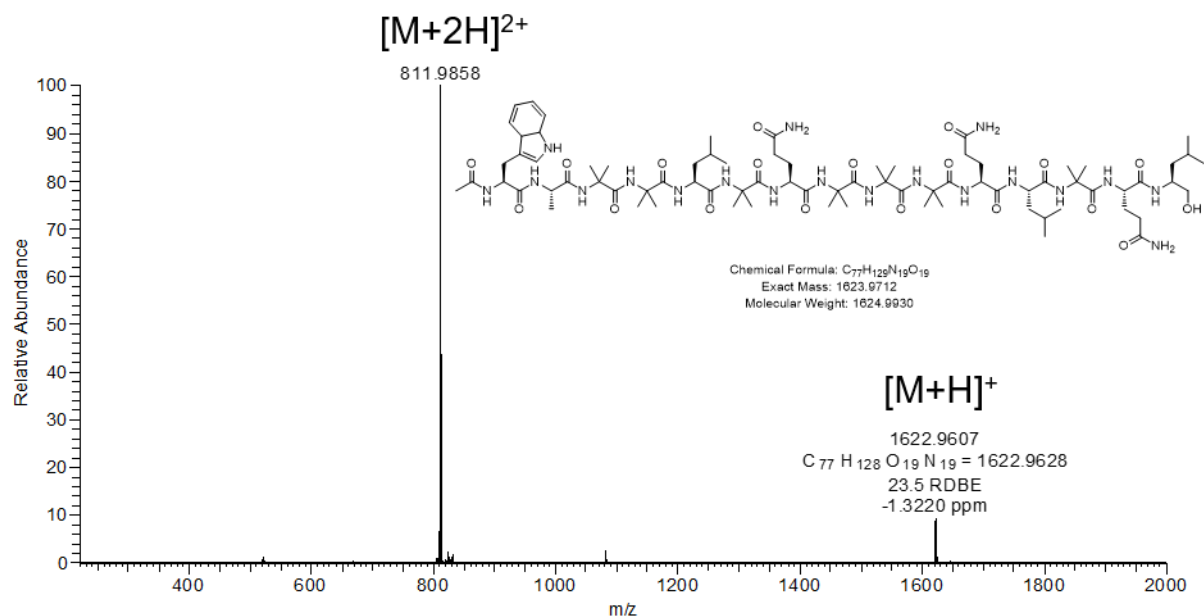
Figure B.1. (+)-ESI-HRMS spectrum of isolated ampullosporin A.
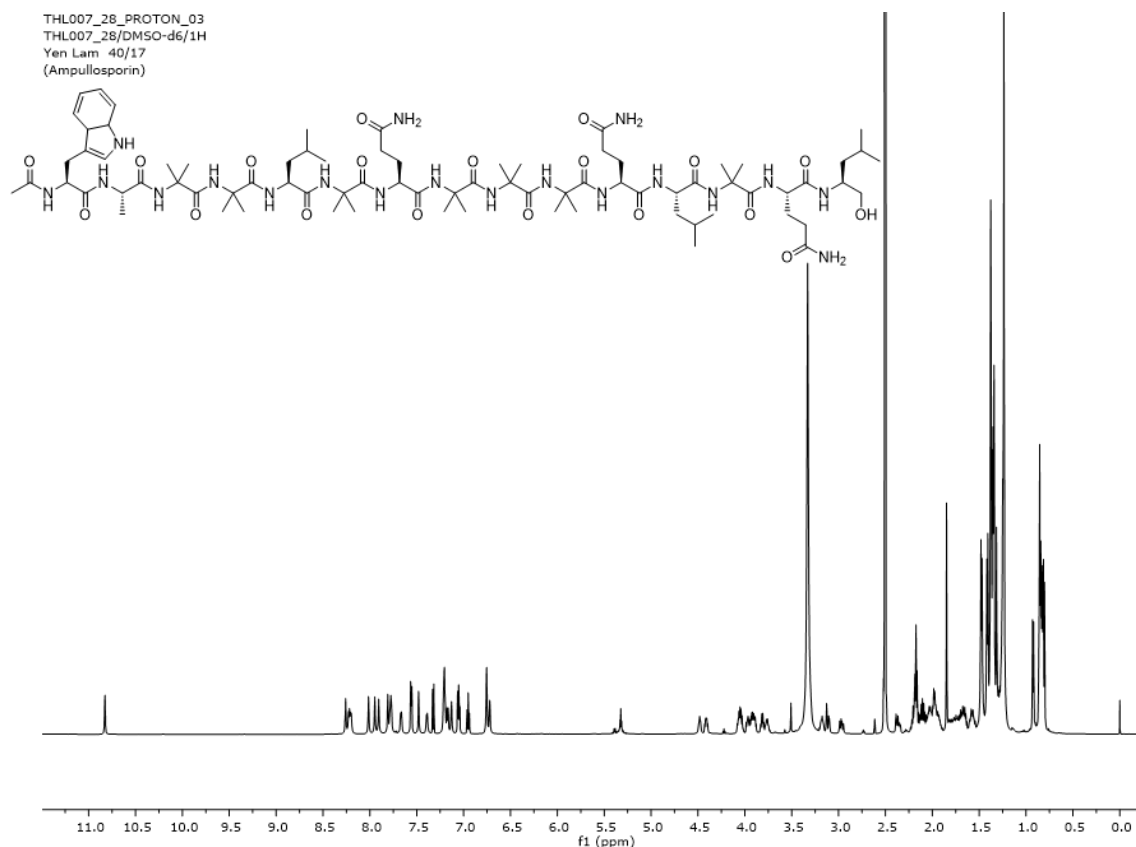
Figure B.2. ^{1}H NMR spectrum of isolated ampullosporin A (600 MHz, DMSO-*d6,* 25°C).
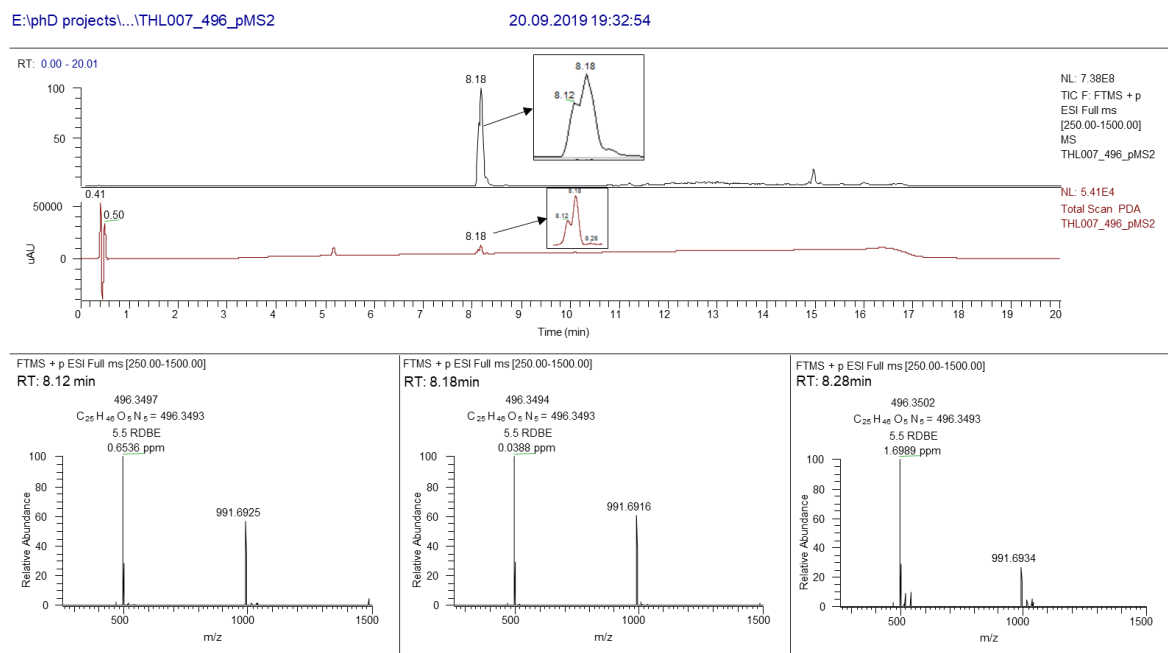
Figure B.3. UHPLC-(+)-ESI-HRMS and PDA spectrum of isomeric mixture of natural cyclic microsporide A (**3.1**).
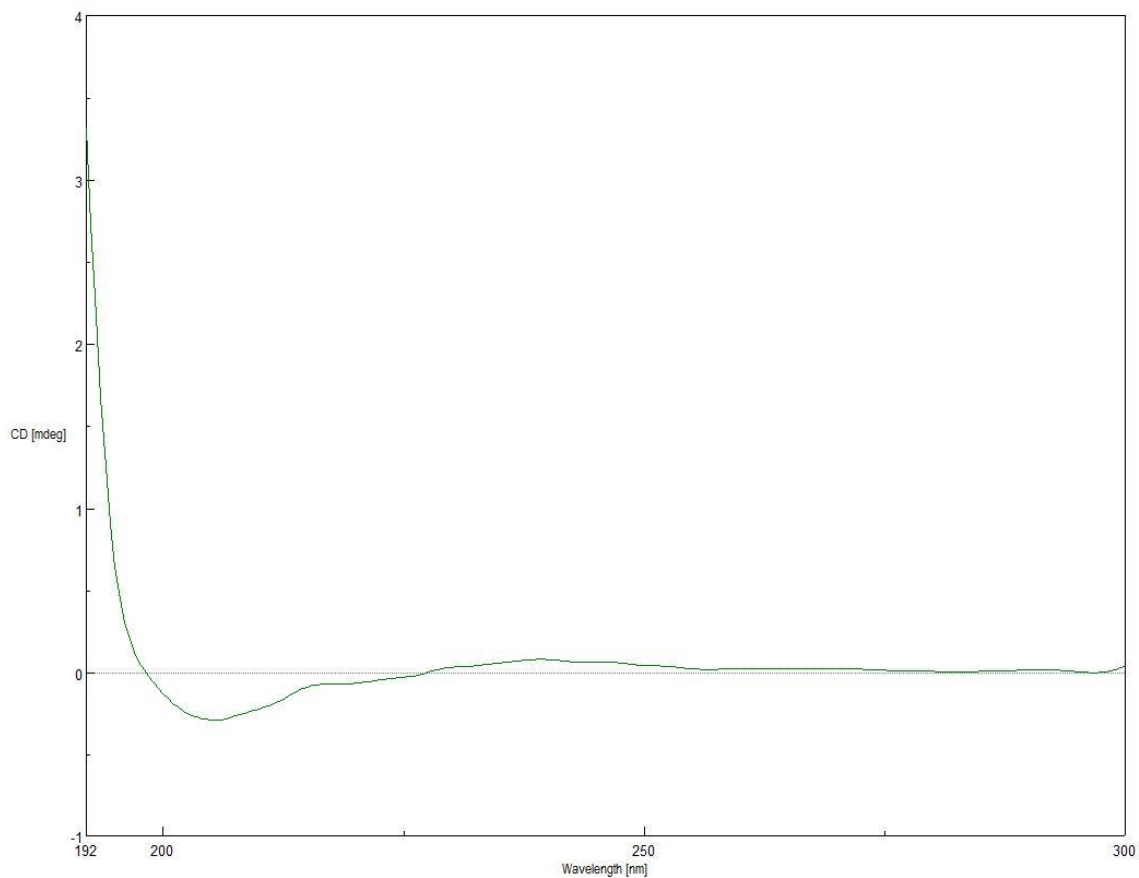
Figure B.4. CD spectrum of isomeric mixture of natural microsporide A (**3.1**) in MeOH.
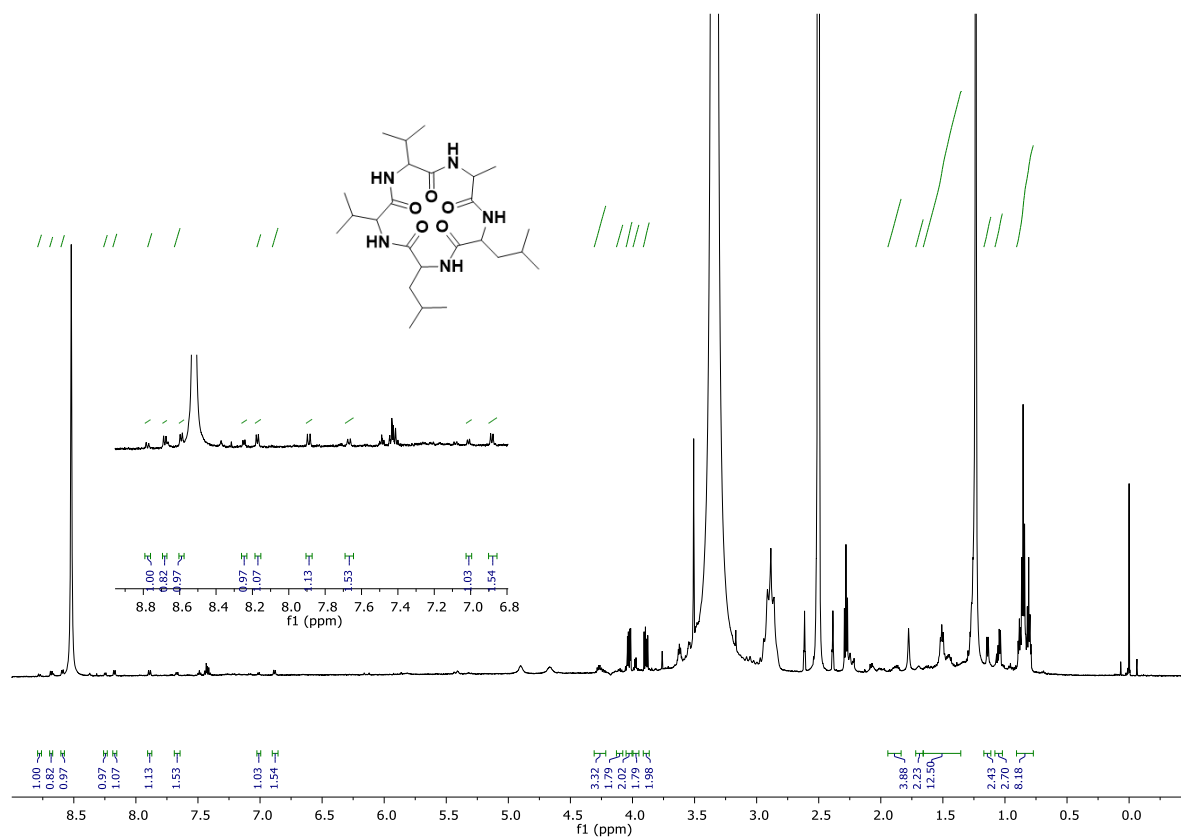
Figure B.5. ¹H NMR spectrum of isomeric mixture of natural microsporide A (**3.1**) (600 MHz, DMSO-*d6*, 40°C).
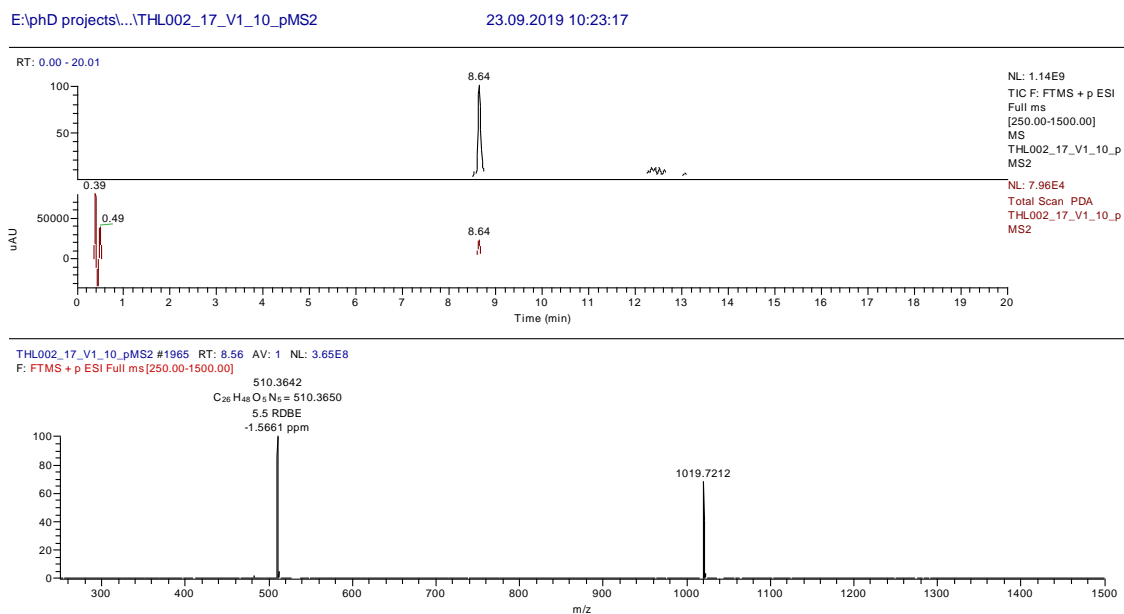
Figure B.6. UHPLC-(+)-ESI-HRMS and PDA spectrum of natural chrysosporide (**3.3**).
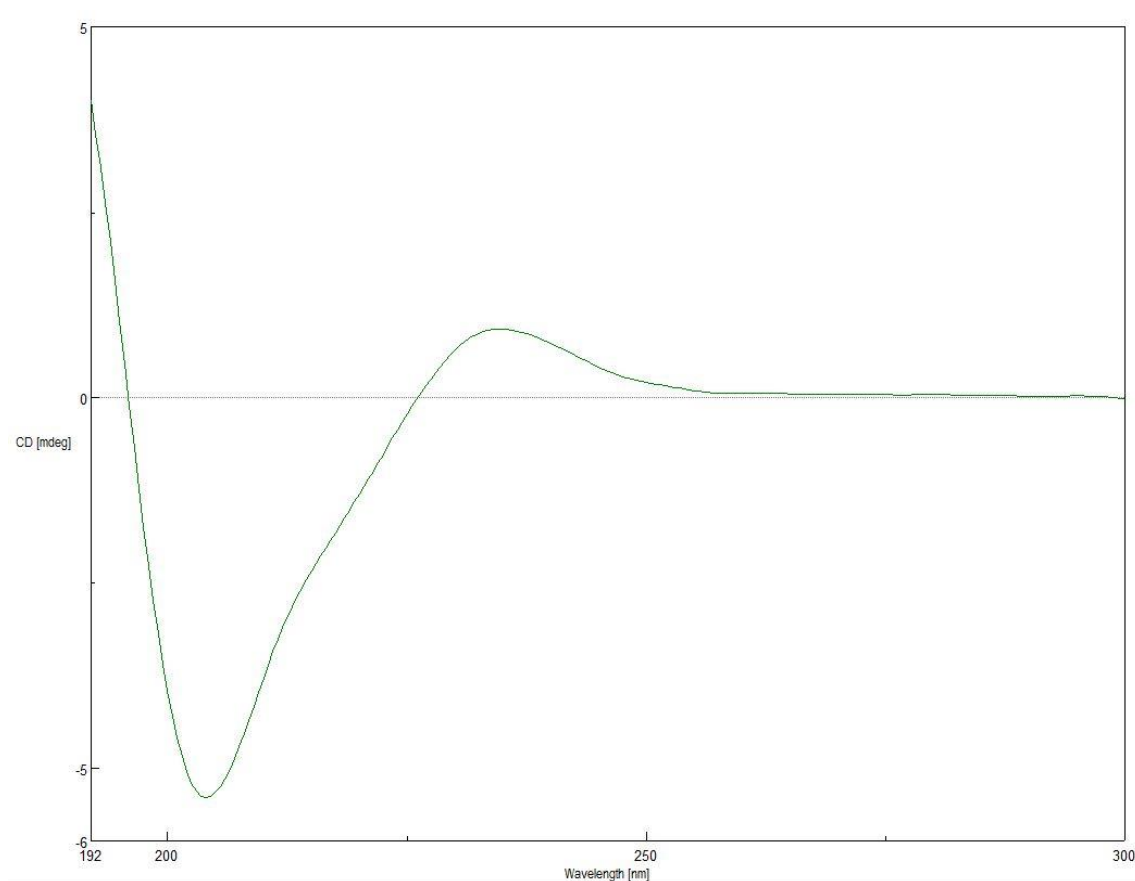
Figure B.7. CD spectrum of natural chrysosporide (**3.3**) in MeOH.
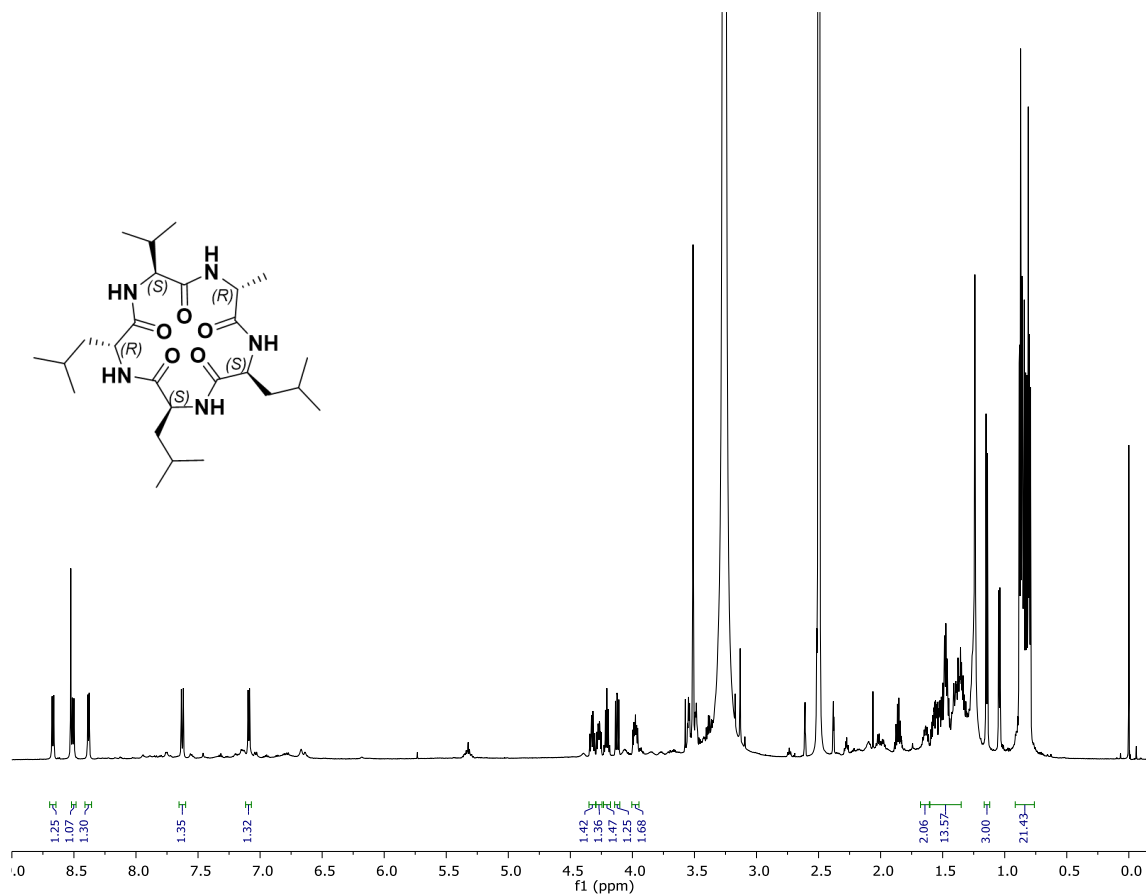
Figure B.8. ¹H NMR spectrum of natural chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C)
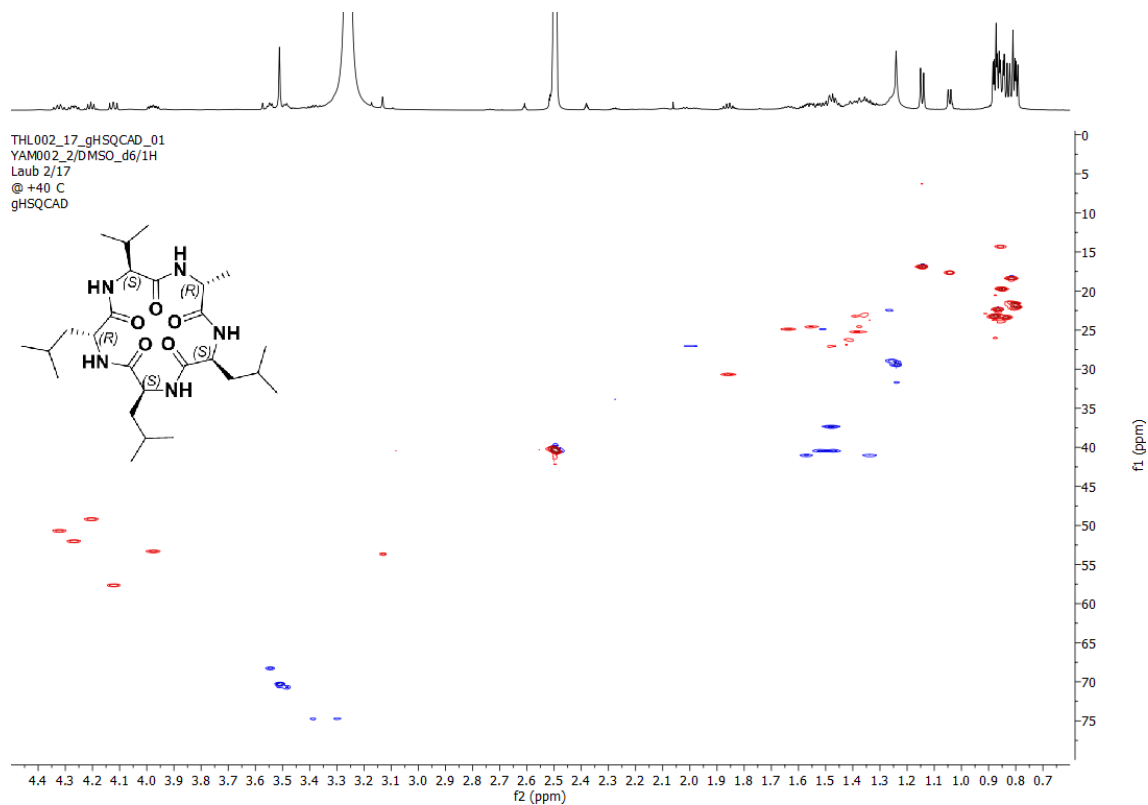
Figure B.9. ¹H, ¹³C HSQC spectrum of natural chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C).
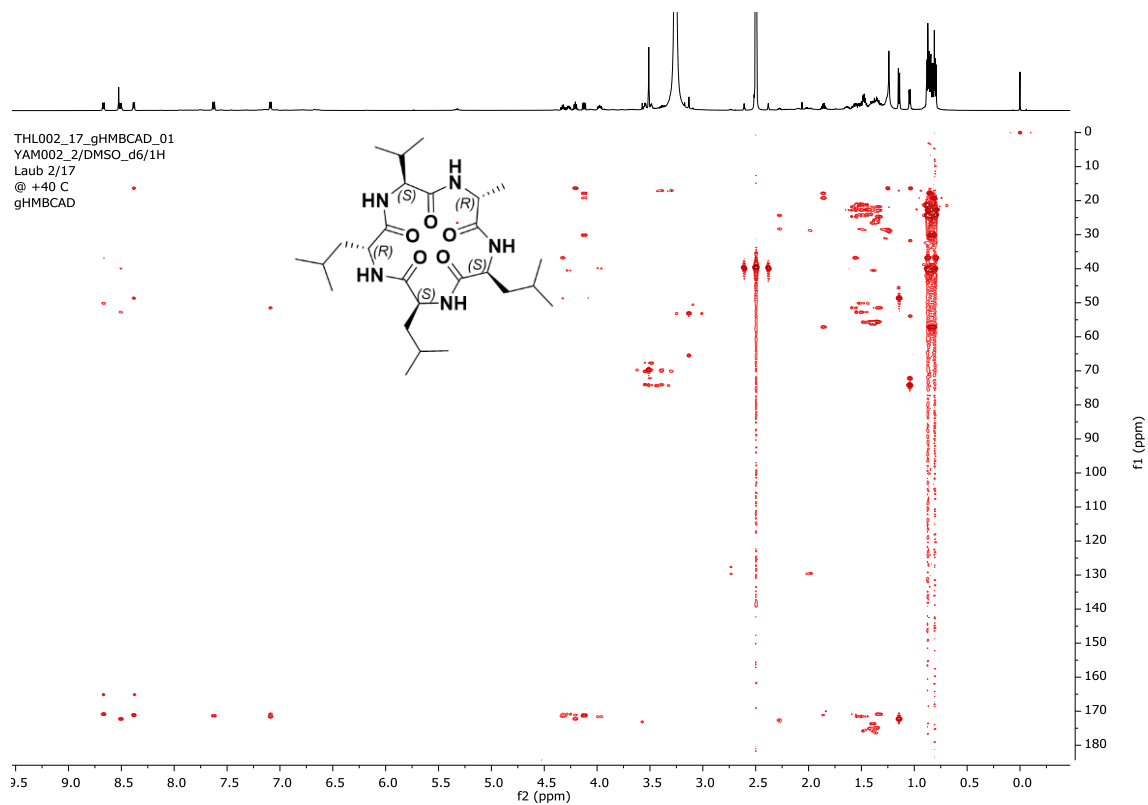
Figure B.10. ^{1}H, ^{13}C HMBC spectrum of natural chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C).
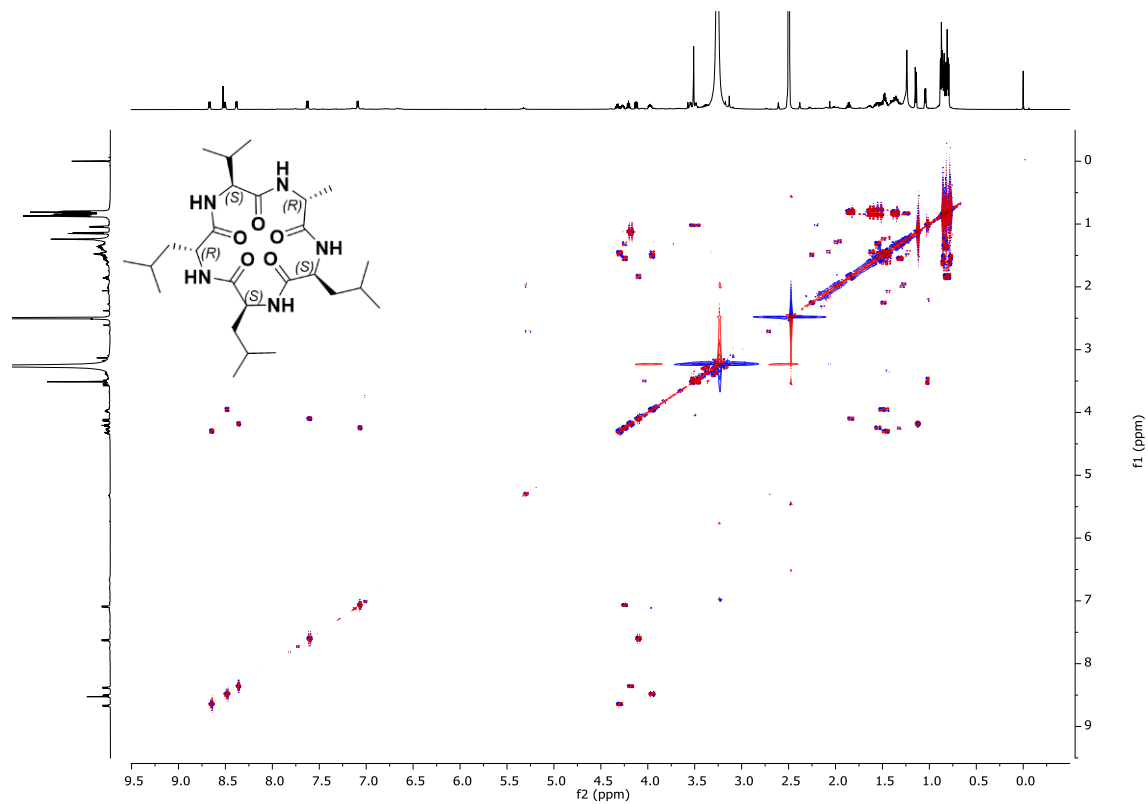
Figure B.11. ^{1}H, ^{1}H COSY spectrum of natural chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C).
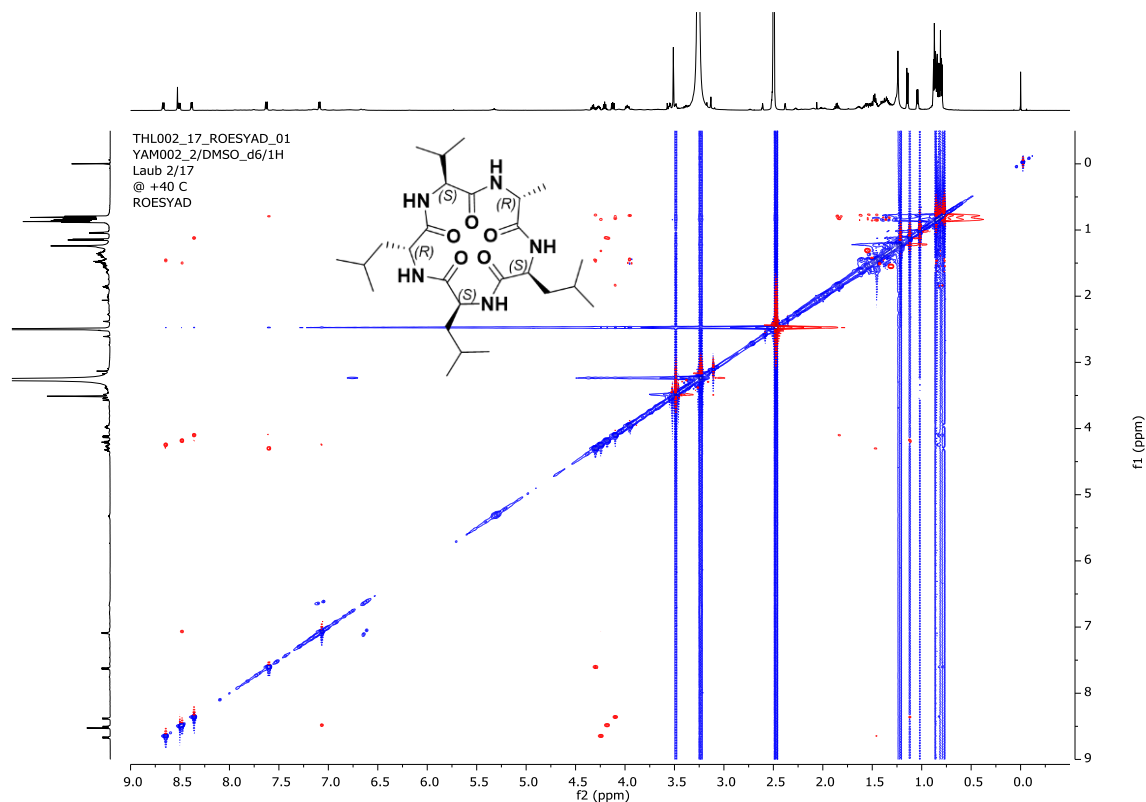
Figure B.12. [1]H, ^{1}H ROESY spectrum of natural chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C).
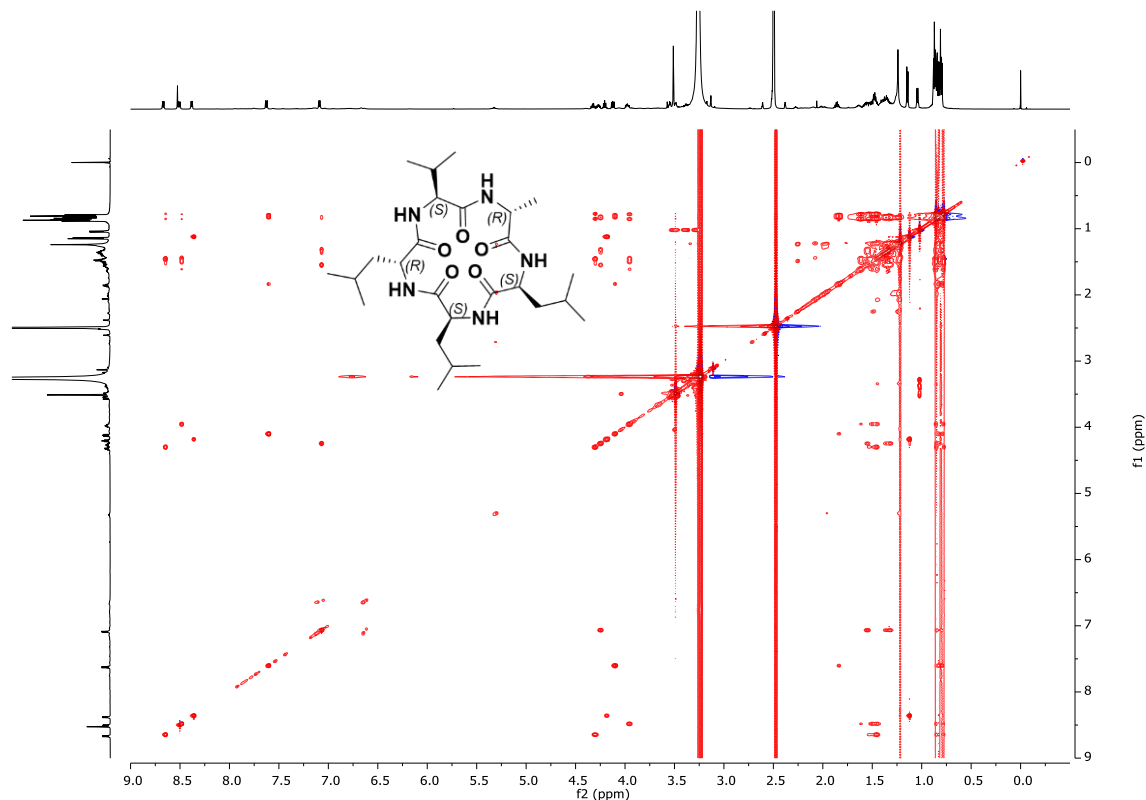
Figure B.13. [1]H, ^{1}H TOCSY spectrum of natural chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C).
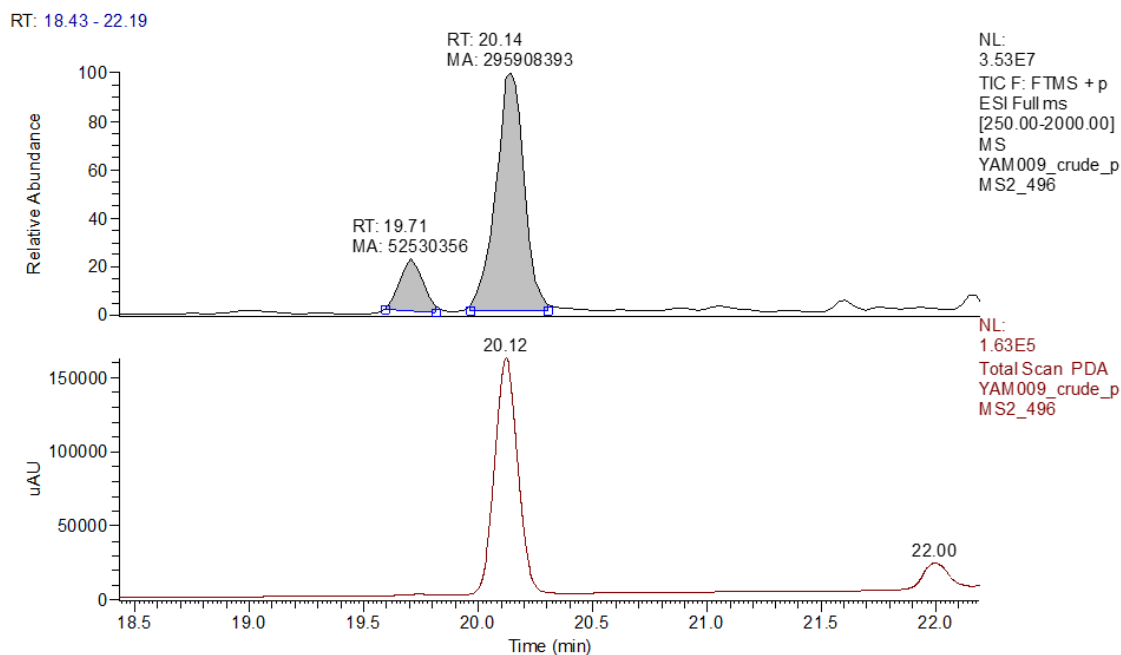
Figure B.14. Total ion chromatogram (TIC) obtained from the UHPLC-(+)-ESI-HRMS measurements in positive ion mode of the crude synthetic product of cyclic pentapeptide **3.1a**.
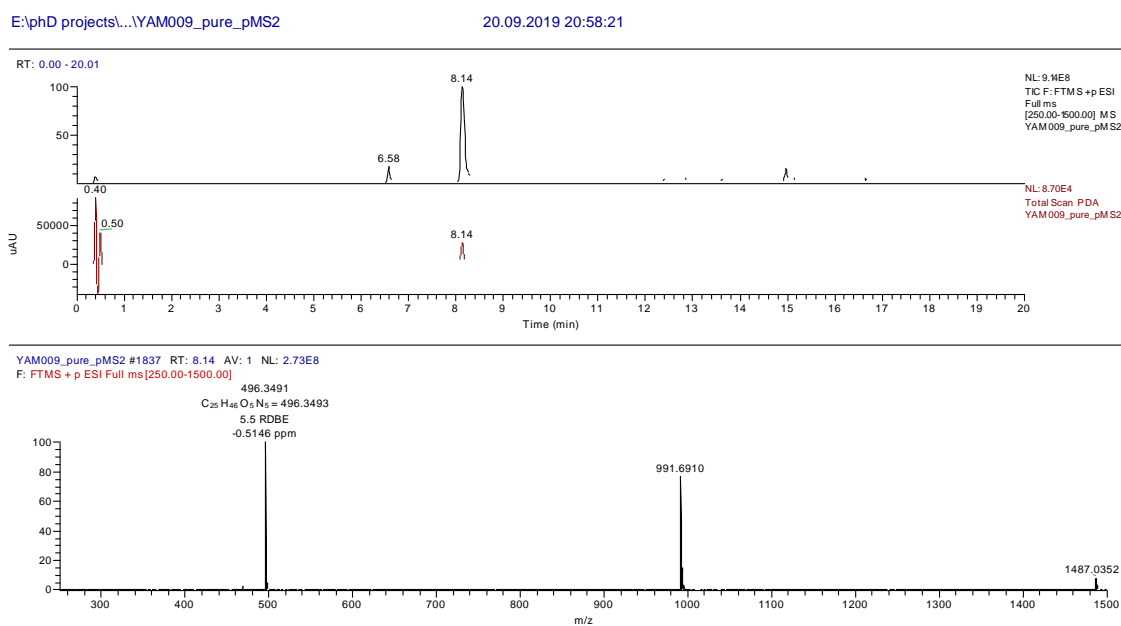
Figure B.15 UHPLC-(+)-ESI-HRMS and PDA spectrum of synthetic cyclic pentapeptide **3.1a**.

l
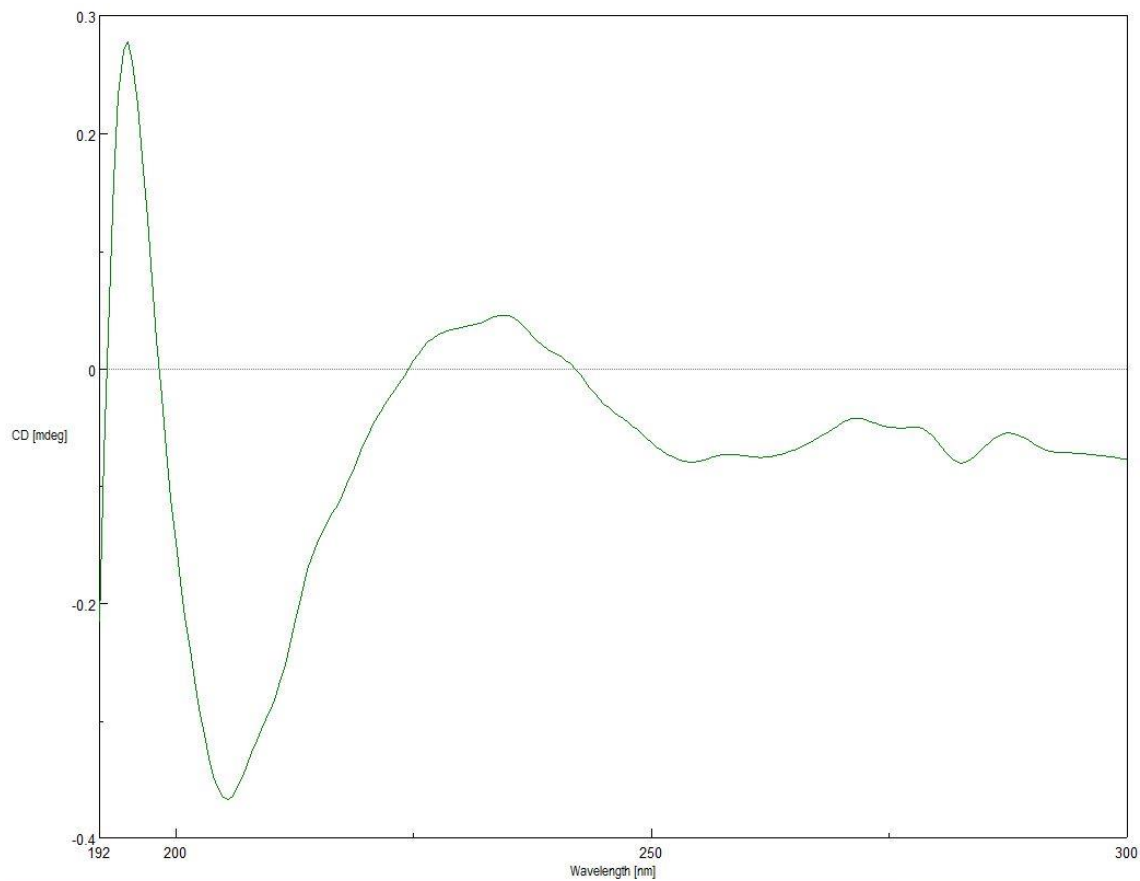
Figure B.16. CD spectrum of synthetic cyclic pentapeptide **3.1a** in MeOH.
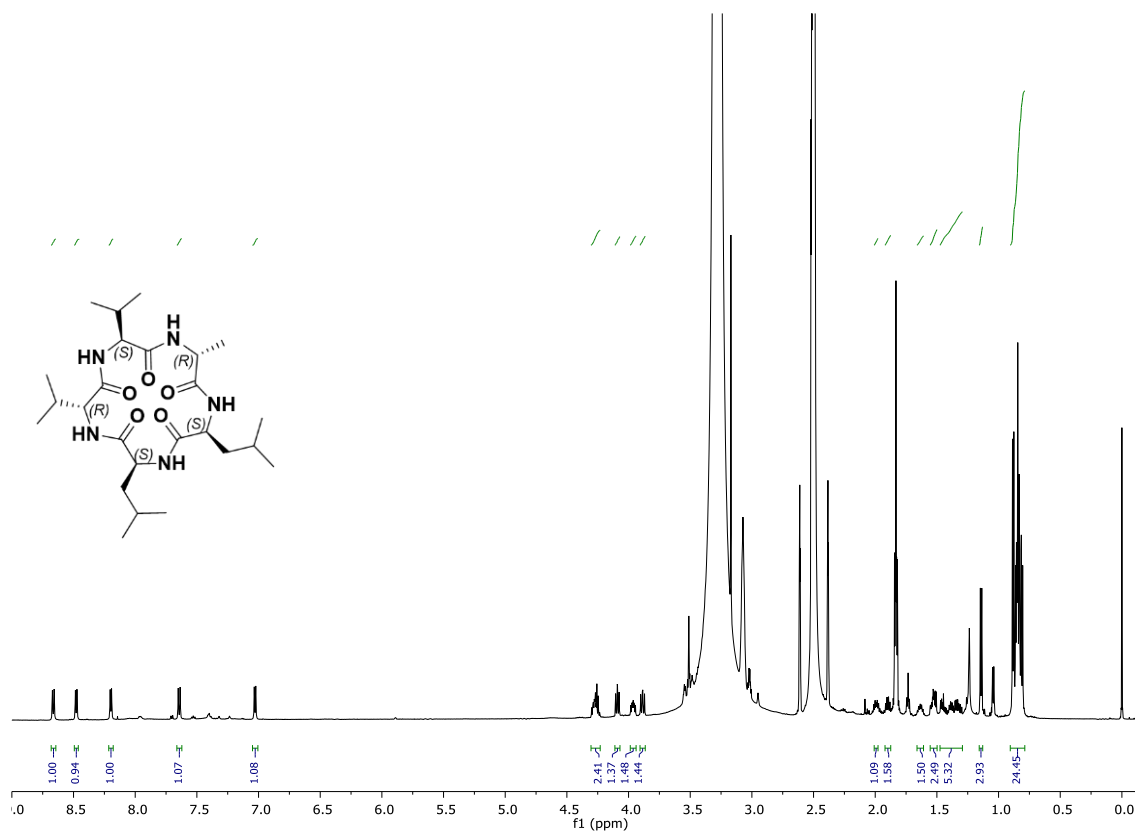
Figure B.17. ¹H NMR spectrum of cyclic pentapeptide **3.1a** (600 MHz, DMSO-*d6*, 40°C).
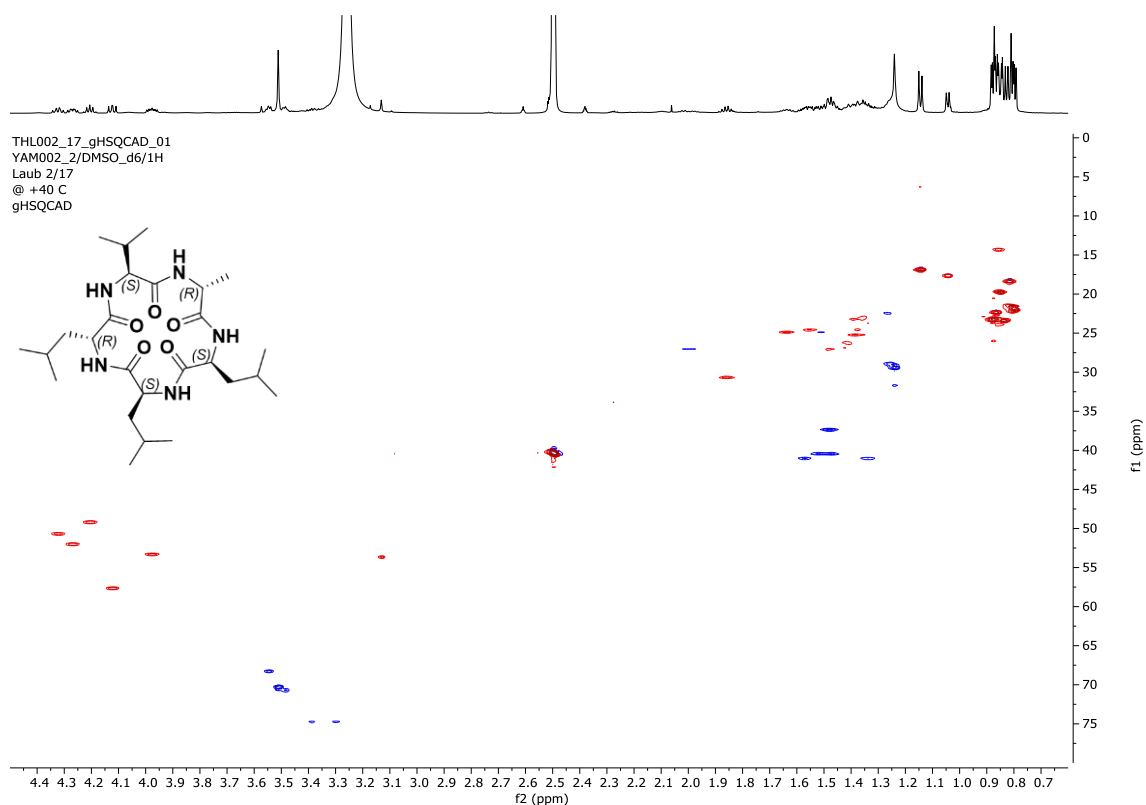
Figure B.18. ¹H, ¹³C HSQC spectrum of cyclic pentapeptide **3.1a** (600 MHz, DMSO-*d6*, 40°C).
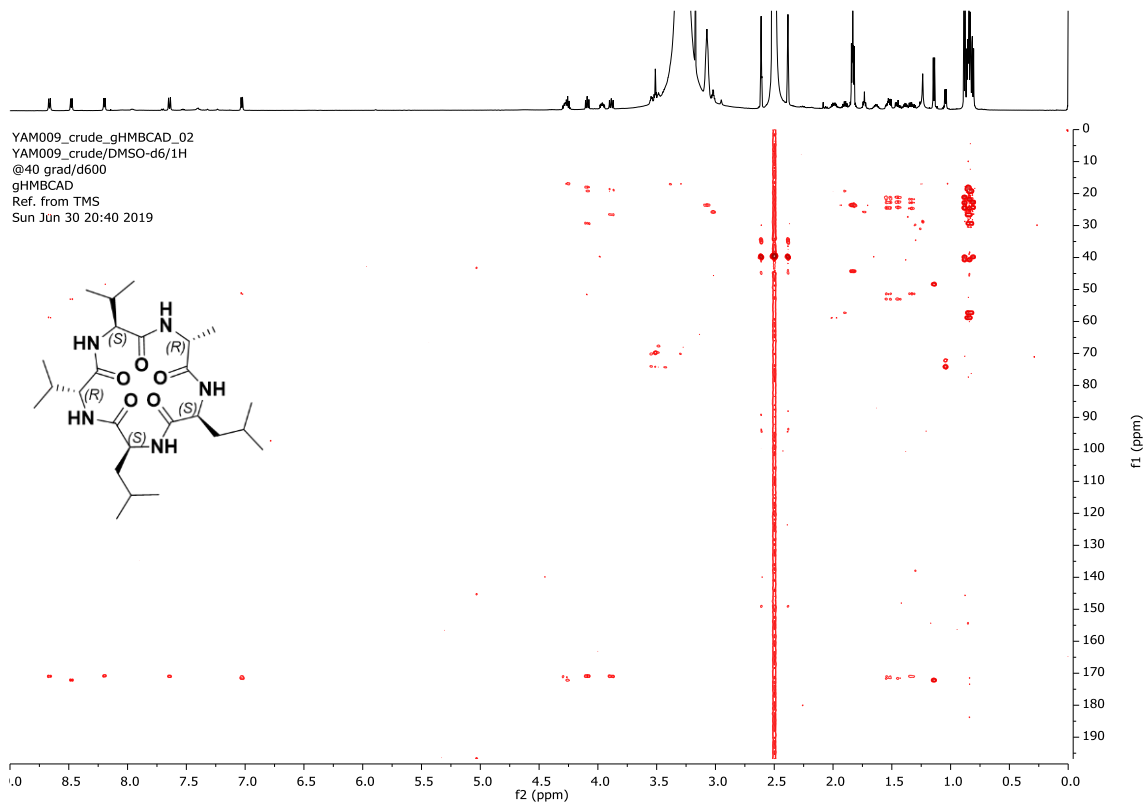
Figure B.19. ¹H, ¹³C HMBC spectrum of cyclic pentapeptide **3.1a** (600 MHz, DMSO-*d6*, 40°C).
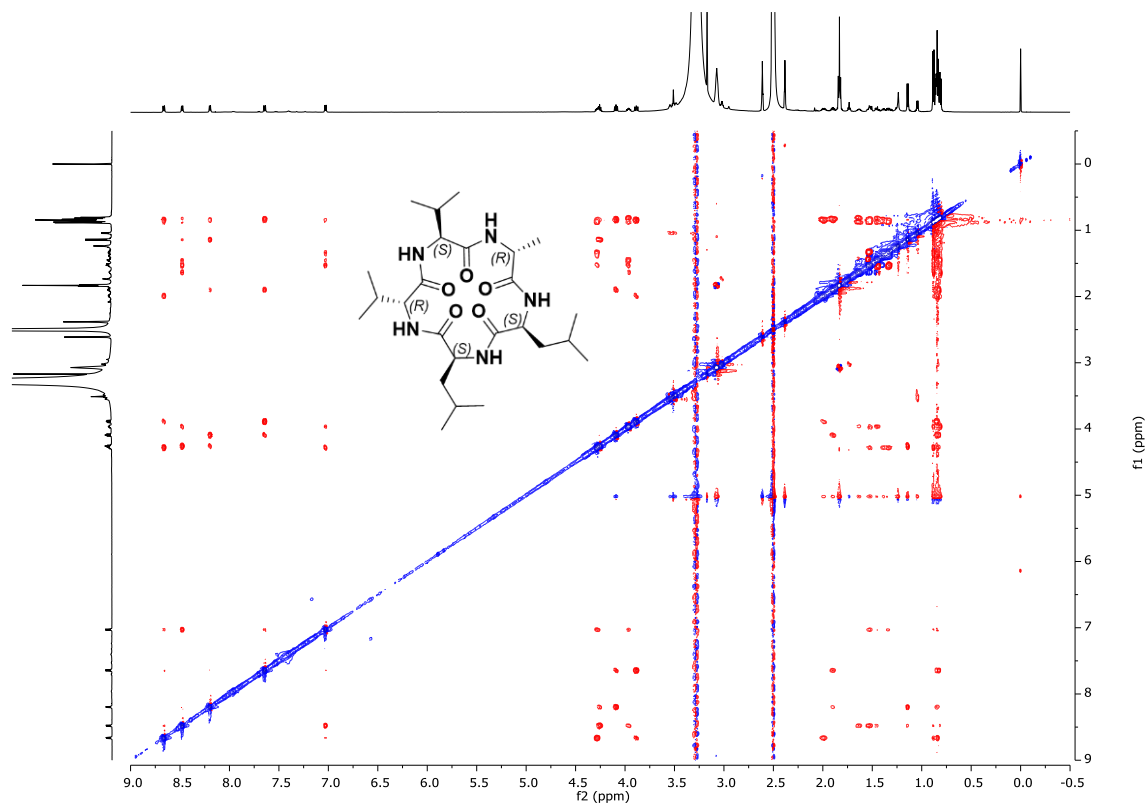
Figure B.20. [1]H, ^{1}H ROESY spectrum of cyclic pentapeptide **3.1a** (600 MHz, DMSO-*d6*, 40°C).
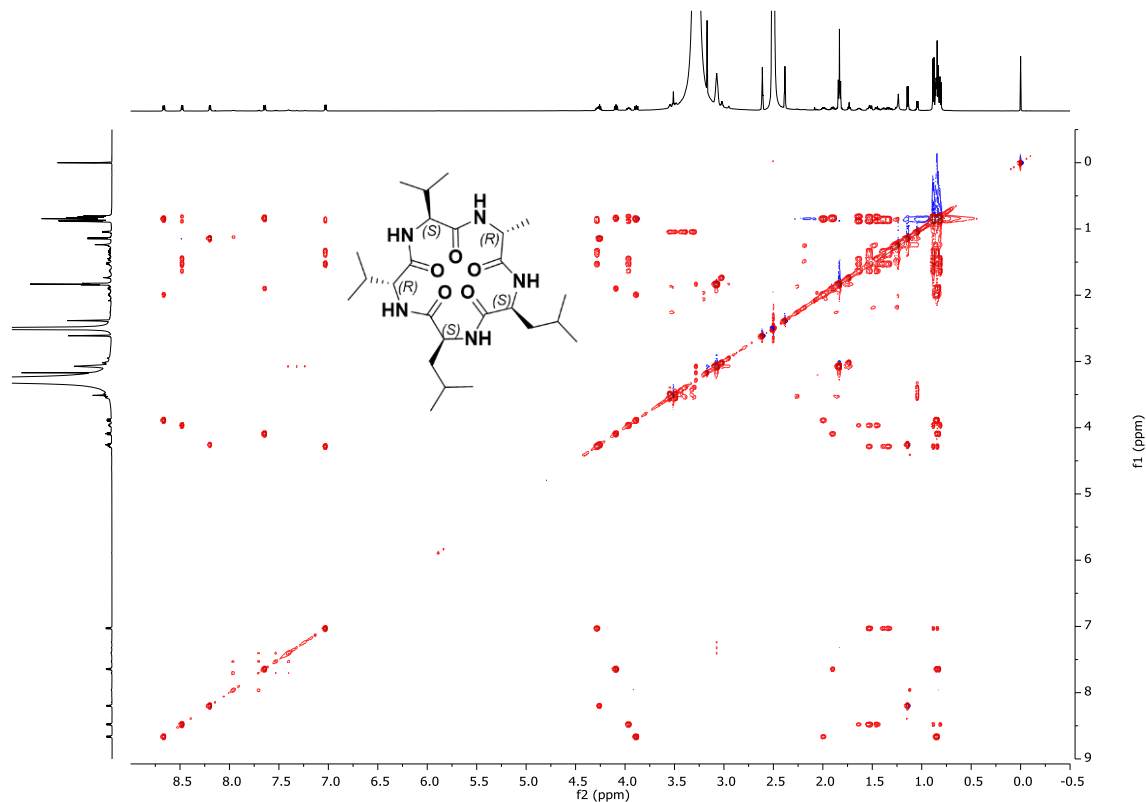
Figure B.21. [1]H, ^{1}H TOCSY spectrum of cyclic pentapeptide **3.1a** (600 MHz, DMSO-*d6*, 40°C).
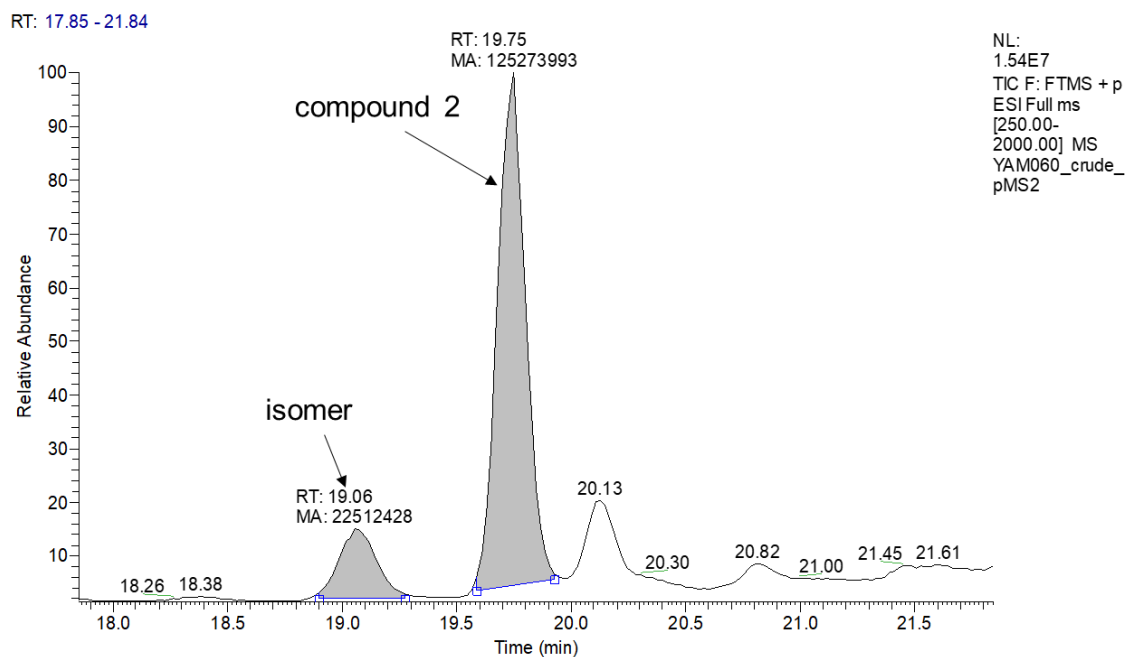
Figure B.22. Total ion chromatogram (TIC) obtained from the UHPLC-(+)-ESI-HRMS measurements of the crude synthetic product of cyclic pentapeptide **3.2**.
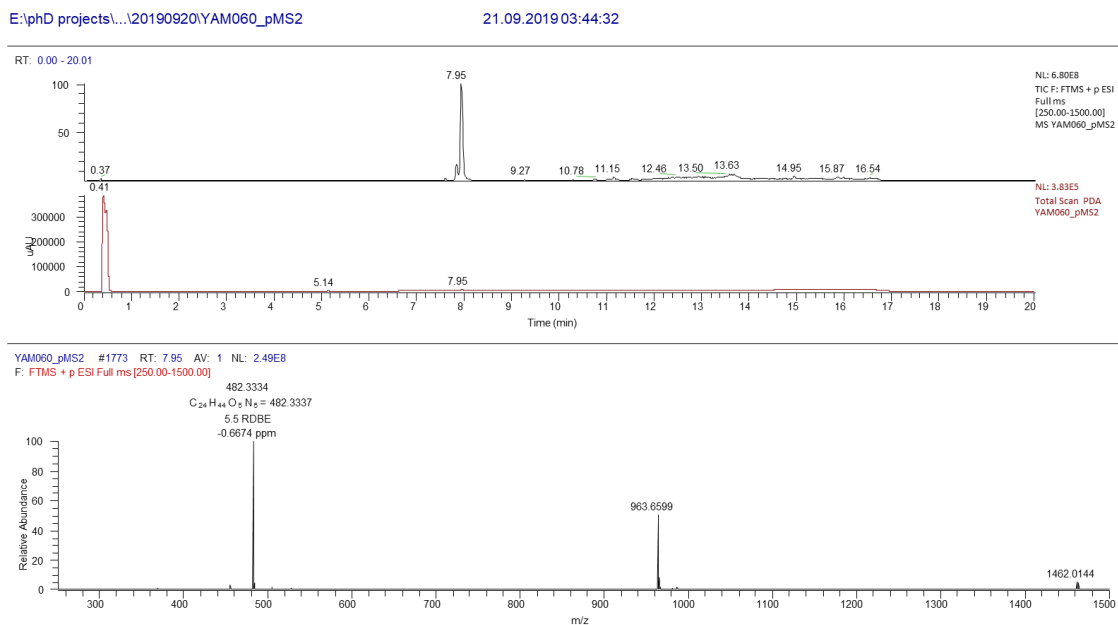
Figure B.23 UHPLC-(+)-ESI-HRMS and PDA spectrum of pure synthetic microsporide B (**3.2**).
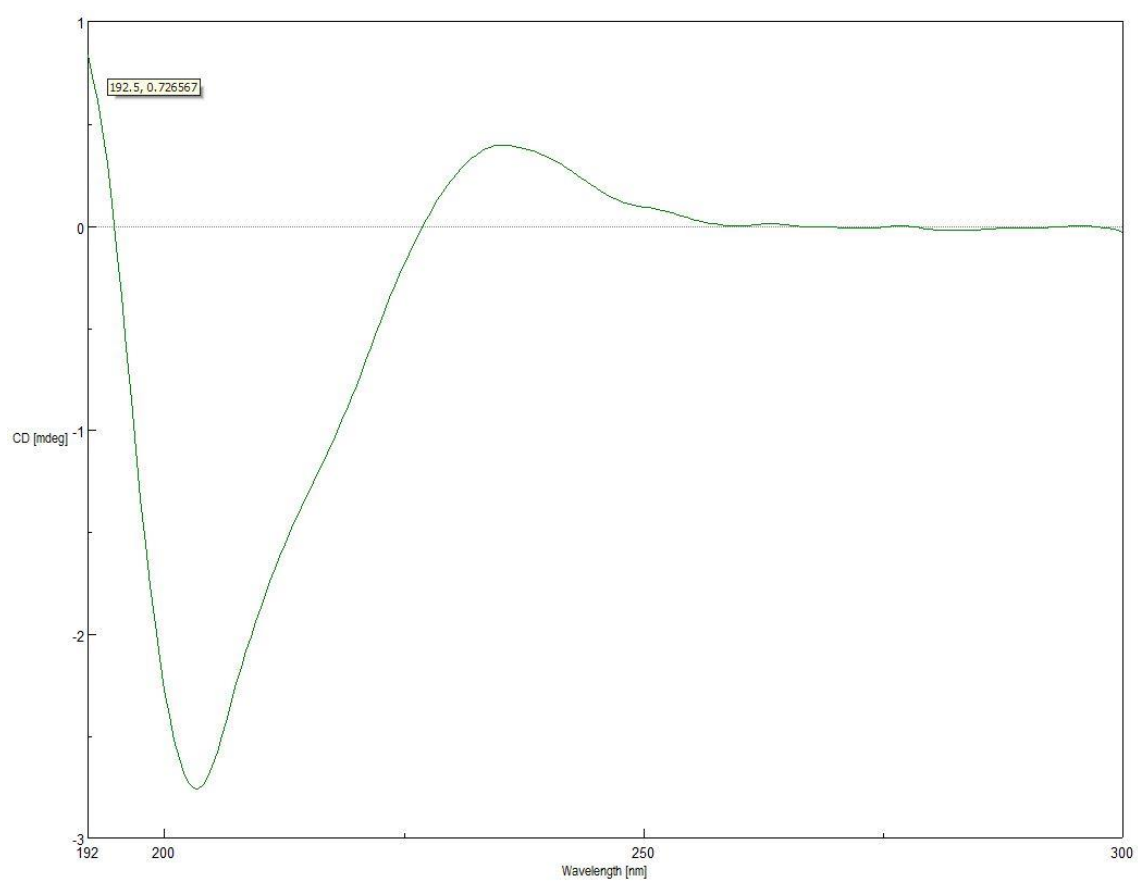
Figure B.24. CD spectrum of synthetic microsporide B (**3.2**) in MeOH.
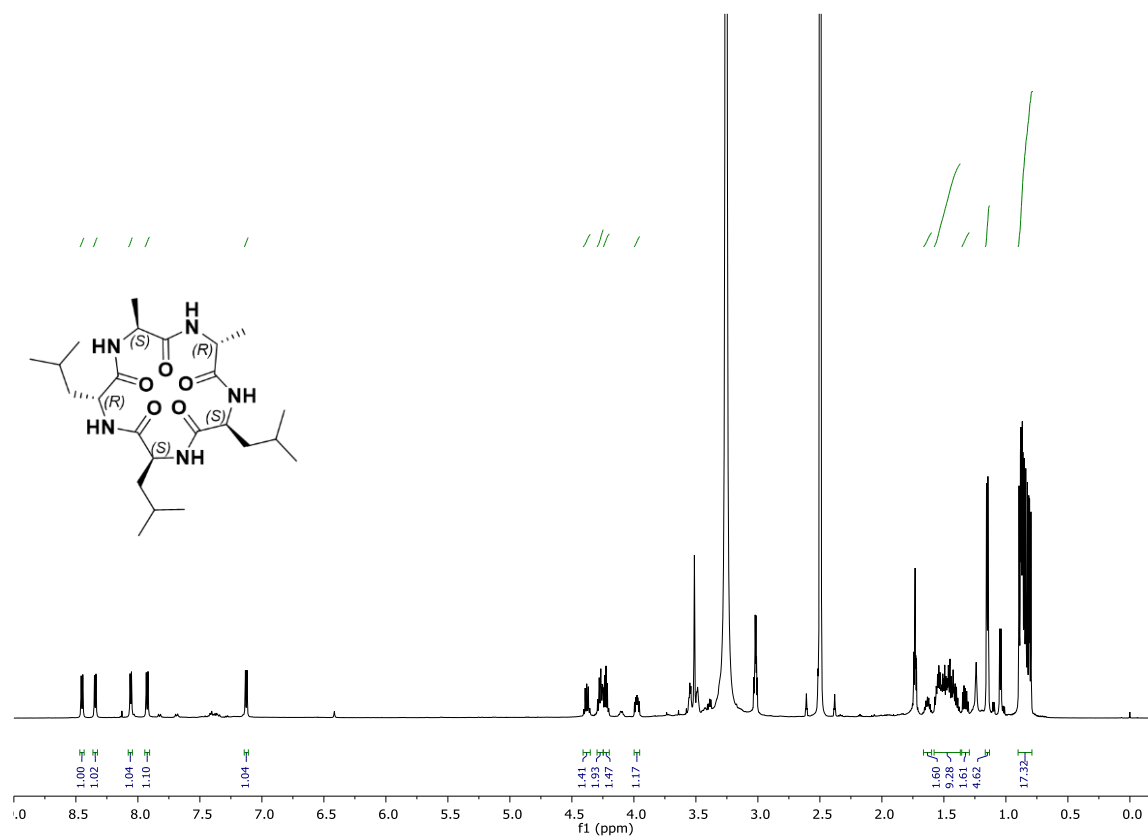
Figure B.25. ^{1}H NMR spectrum of synthetic cyclic microsporide B (**3.2**) (600 MHz, DMSO-*d6*, 40°C).
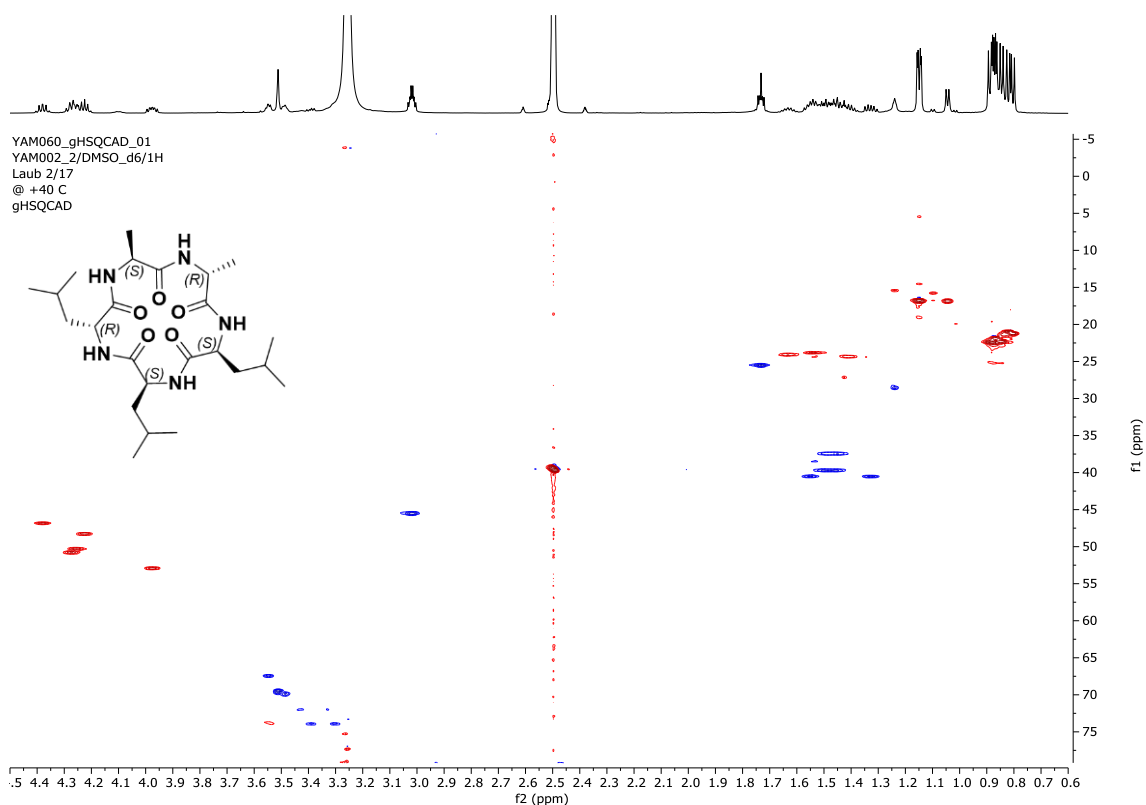
Figure B.26. ^{1}H, ^{13}C HSQC spectrum of synthetic cyclic microsporide B (**3.2**) (600 MHz, DMSO-*d6*, 40°C).
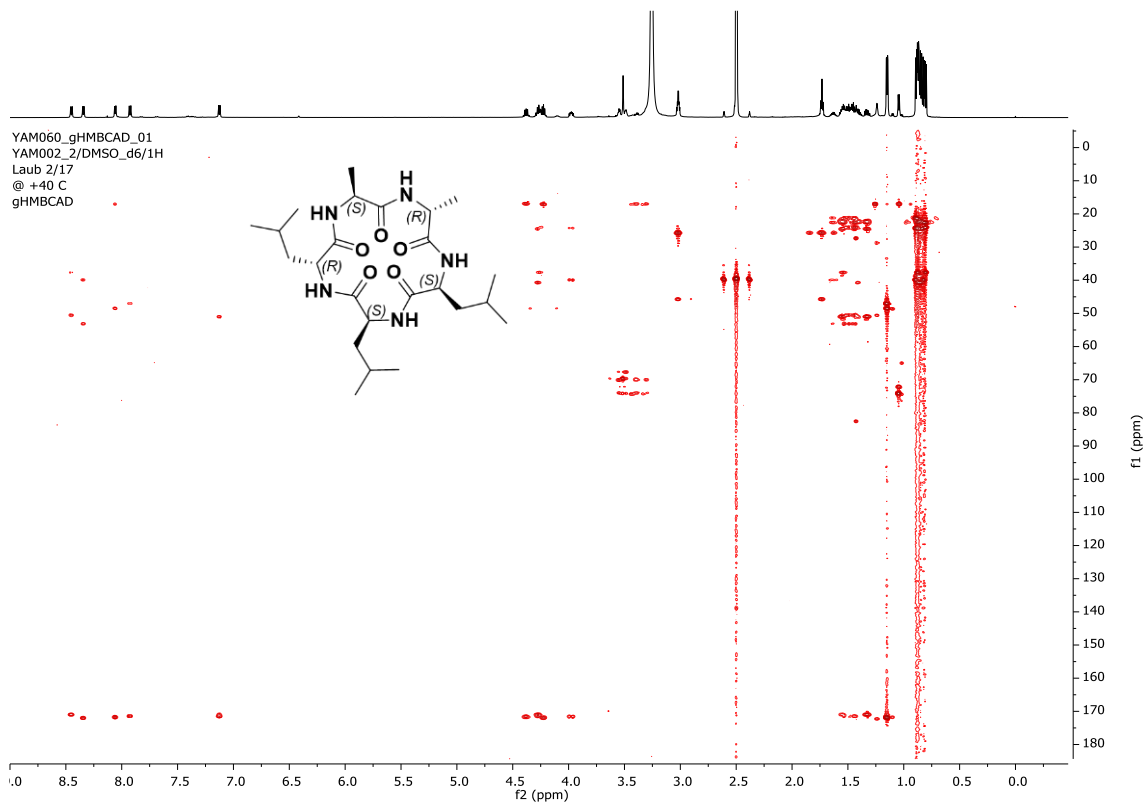
Figure B.27. ^{1}H, ^{13}C HMBC spectrum of synthetic cyclic microsporide B (**3.2**) (600 MHz, DMSO-*d6*, 40°C).
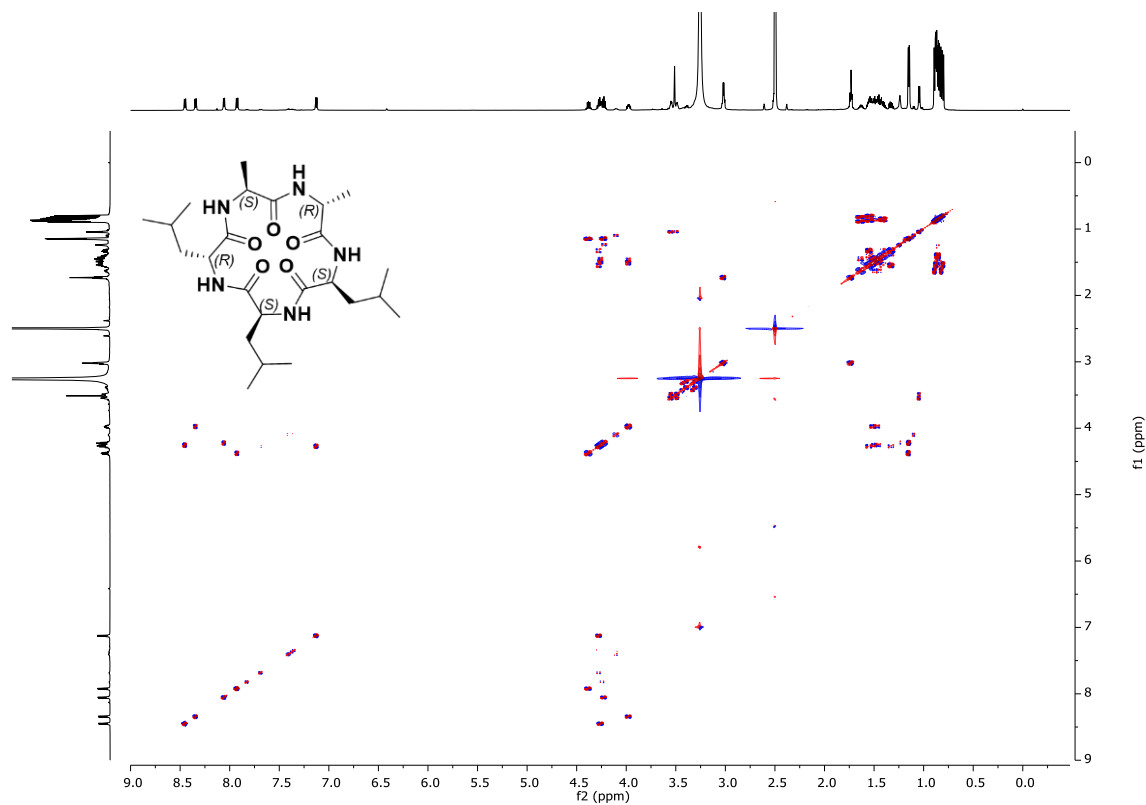
Figure B.28. ¹H, ¹H COSY spectrum of synthetic cyclic microsporide B (**3.2**) (600 MHz, DMSO-*d6*, 40°C).
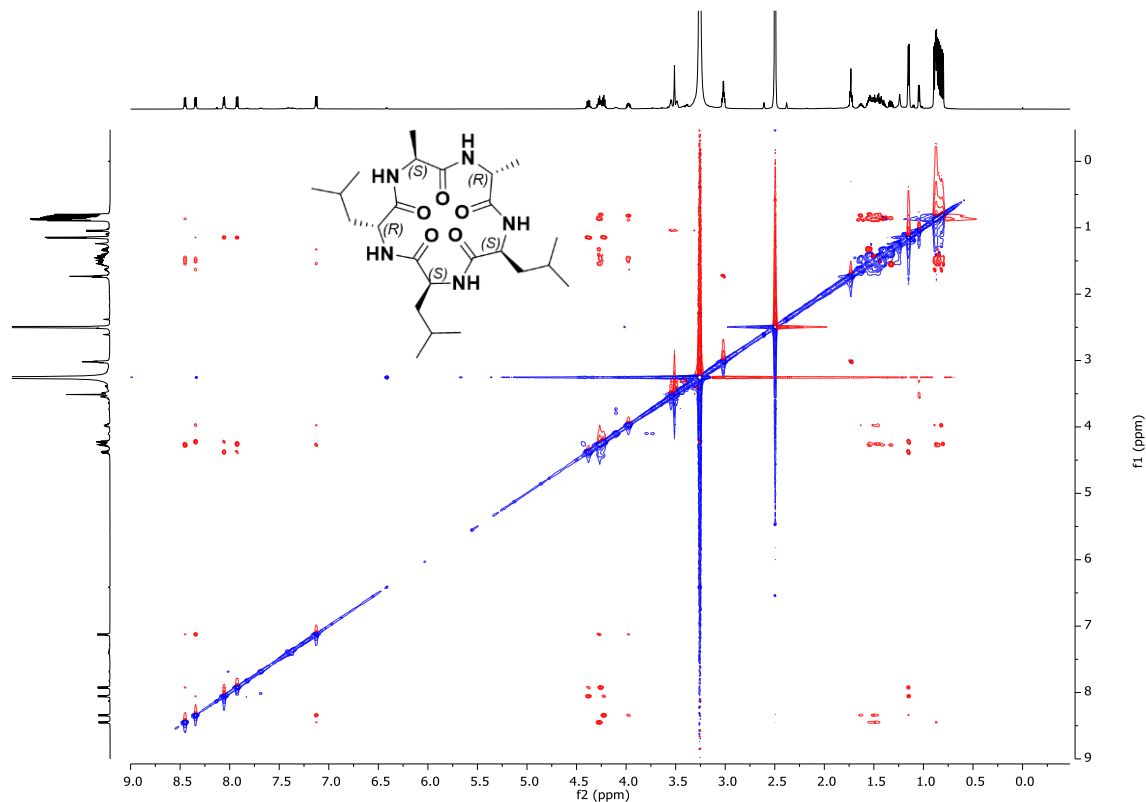
Figure B.29. ¹H, ¹H ROESY spectrum of synthetic cyclic microsporide B (**3.2**) (600 MHz, DMSO-*d6*, 40°C).
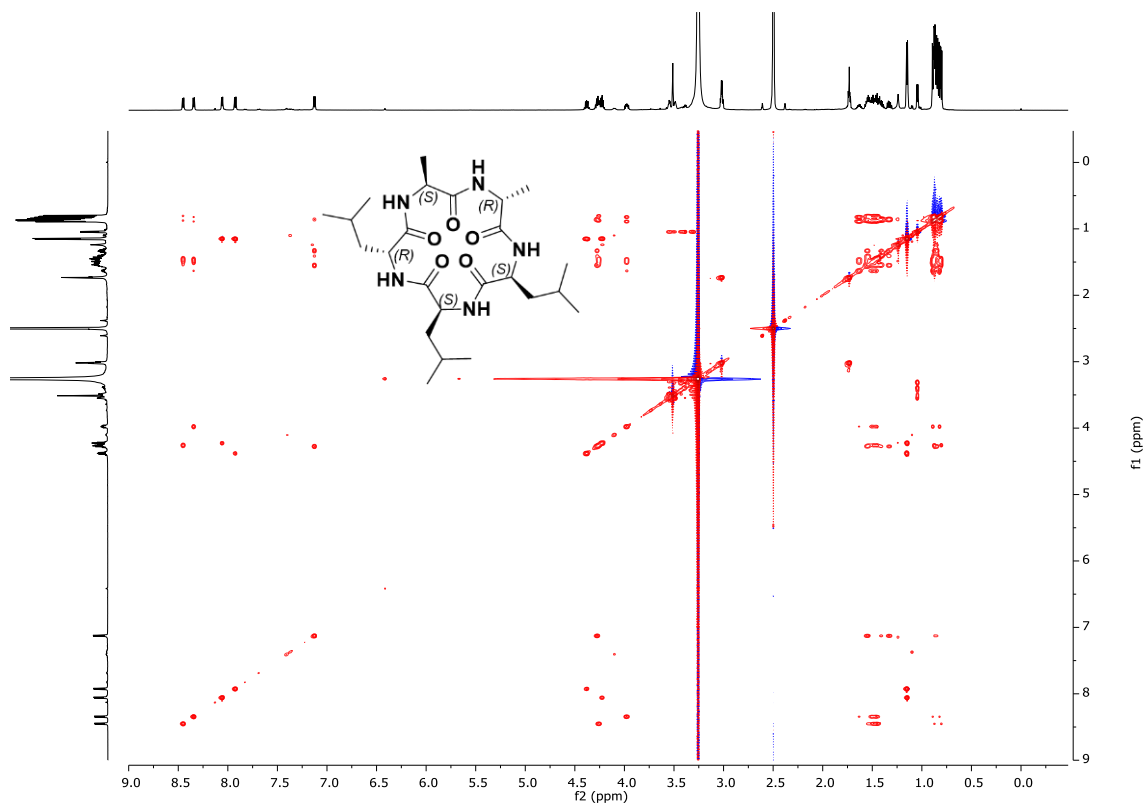
Figure B.30. ^{1}H, ^{1}H TOCSY spectrum of synthetic cyclic microsporide B (**3.2**) (600 MHz, DMSO-*d6*, 40°C).
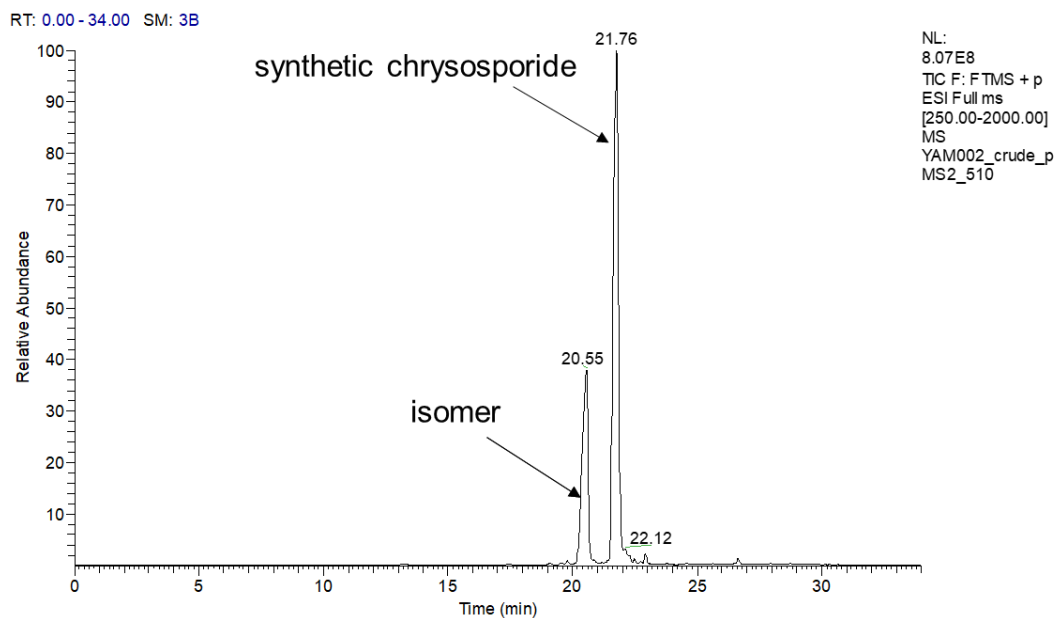
Figure B.31. Total ion chromatogram (TIC) obtained from the UHPLC-(+)-ESI-HRMS measurements of the crude synthetic chrysosporide (**3.3**).
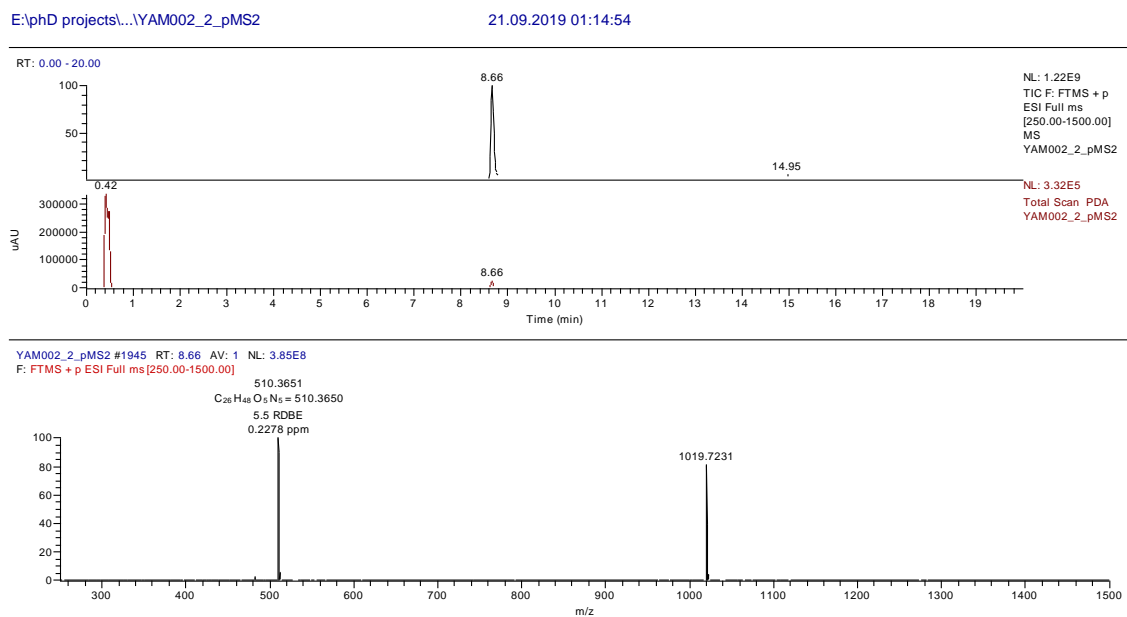
Figure B.32. UHPLC-(+)-ESI-HRMS and PDA spectrum of synthetic chrysosporide (**3.3**).
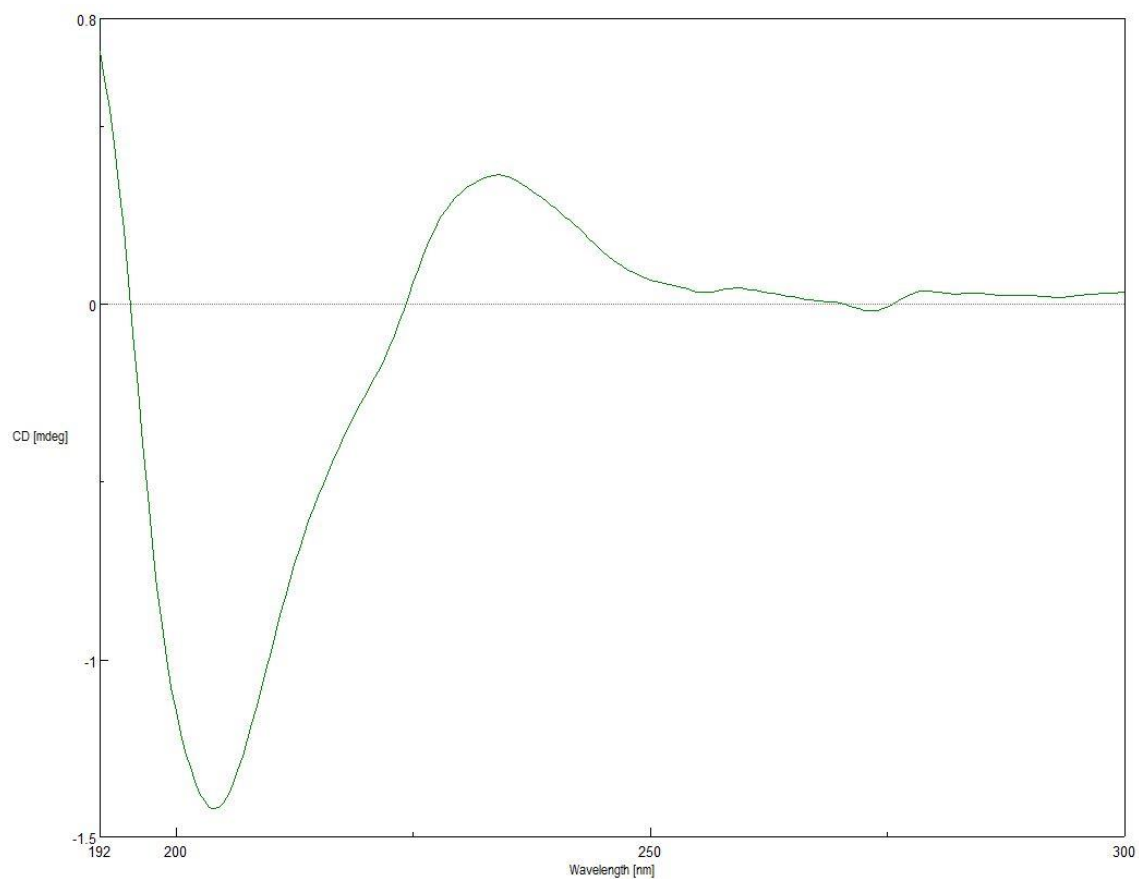
Figure B.33. CD spectrum of synthetic chrysosporide (**3.3**) in MeOH.
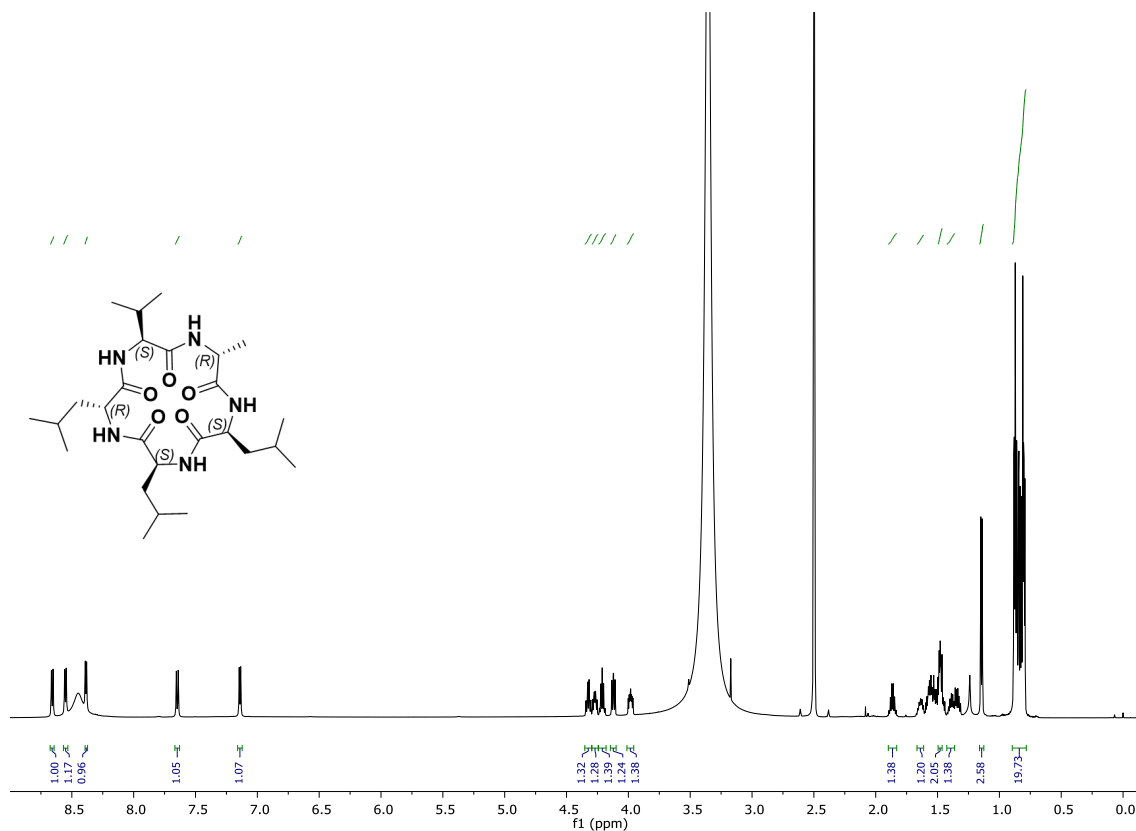
Figure B.34. ^{1}H NMR spectrum of synthetic chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C).
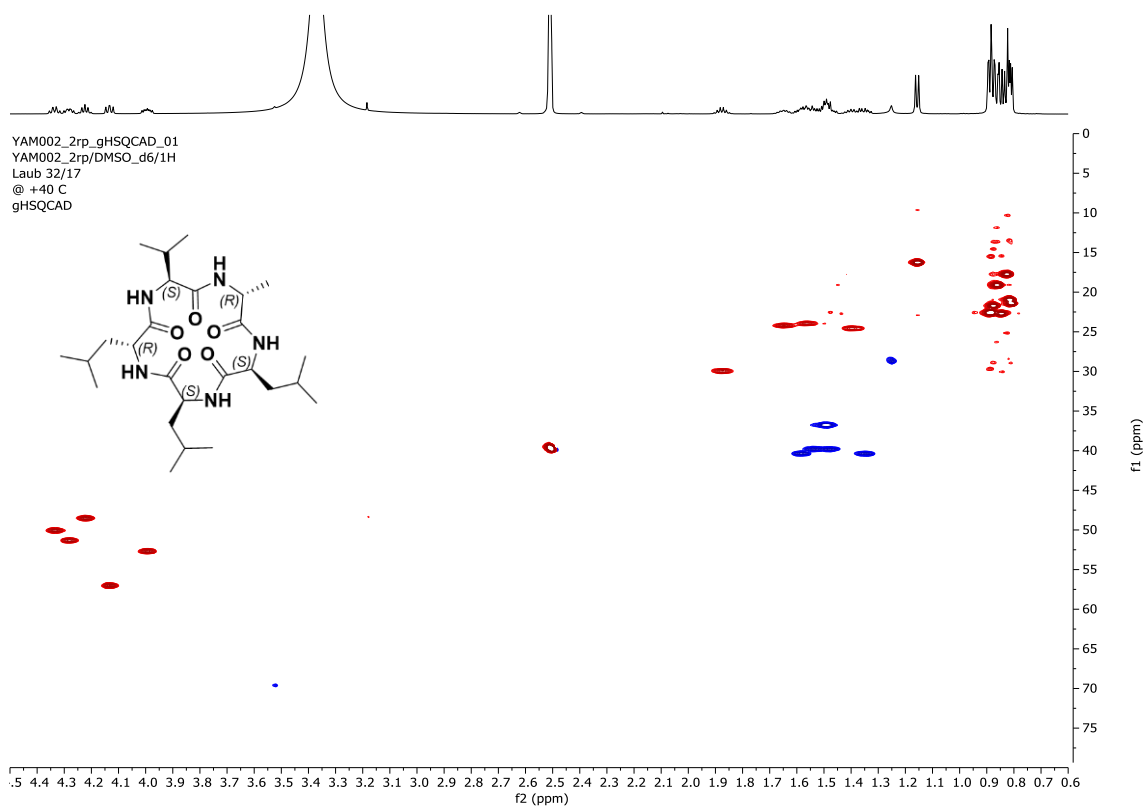
Figure B.35. ^{1}H, ^{13}C HSQC spectrum of synthetic chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C).
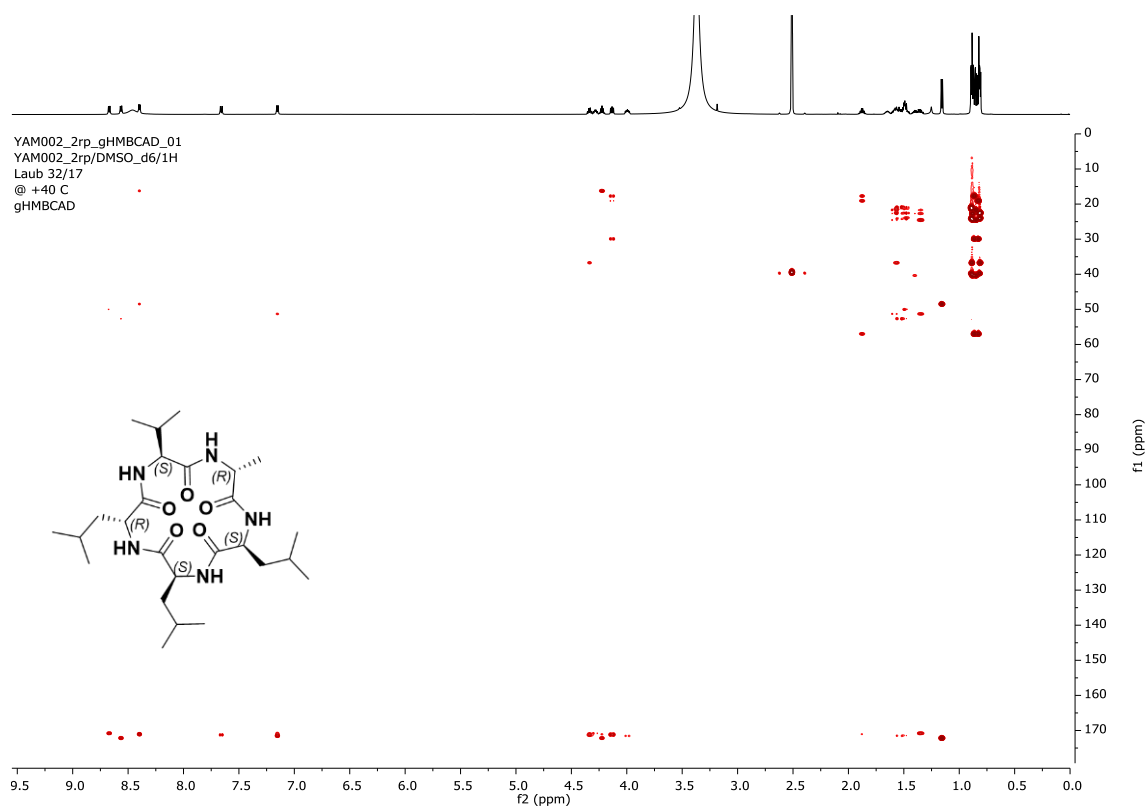
Figure B.36. ^{1}H, ^{13}C HMBC spectrum of synthetic chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C).
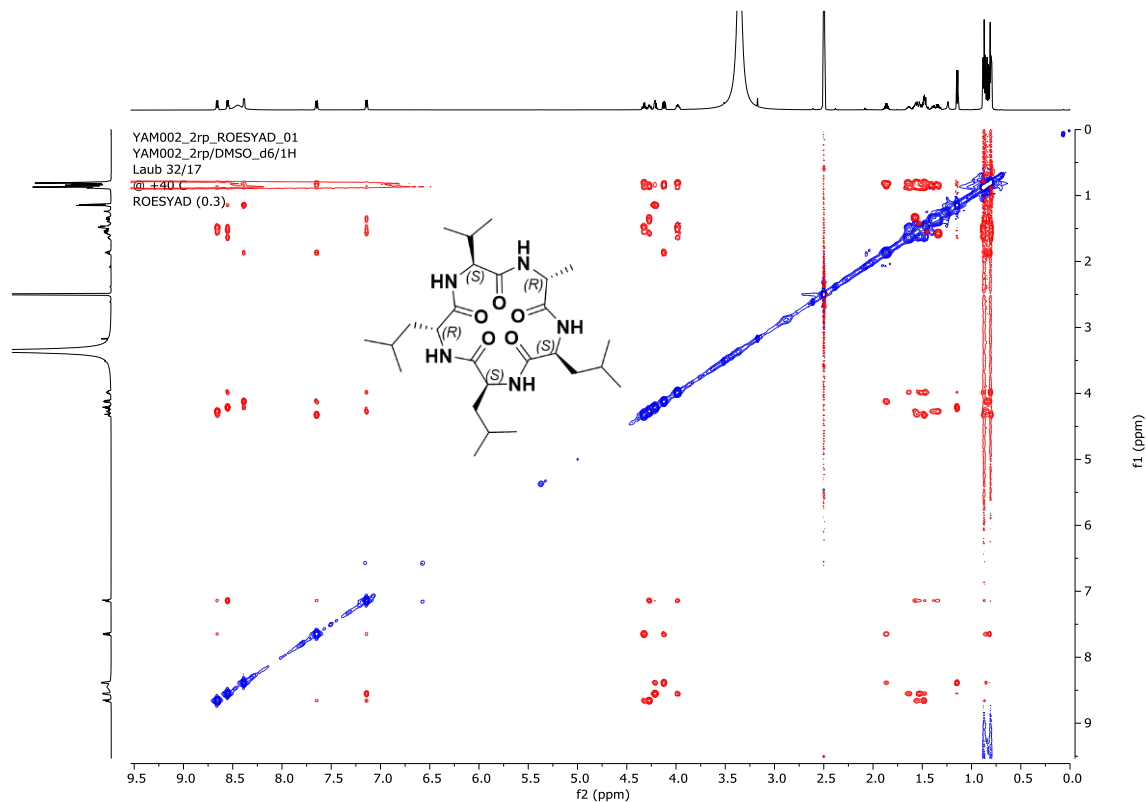
Figure B.37. [1]H, ^{1}H ROESY spectrum of synthetic chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C).
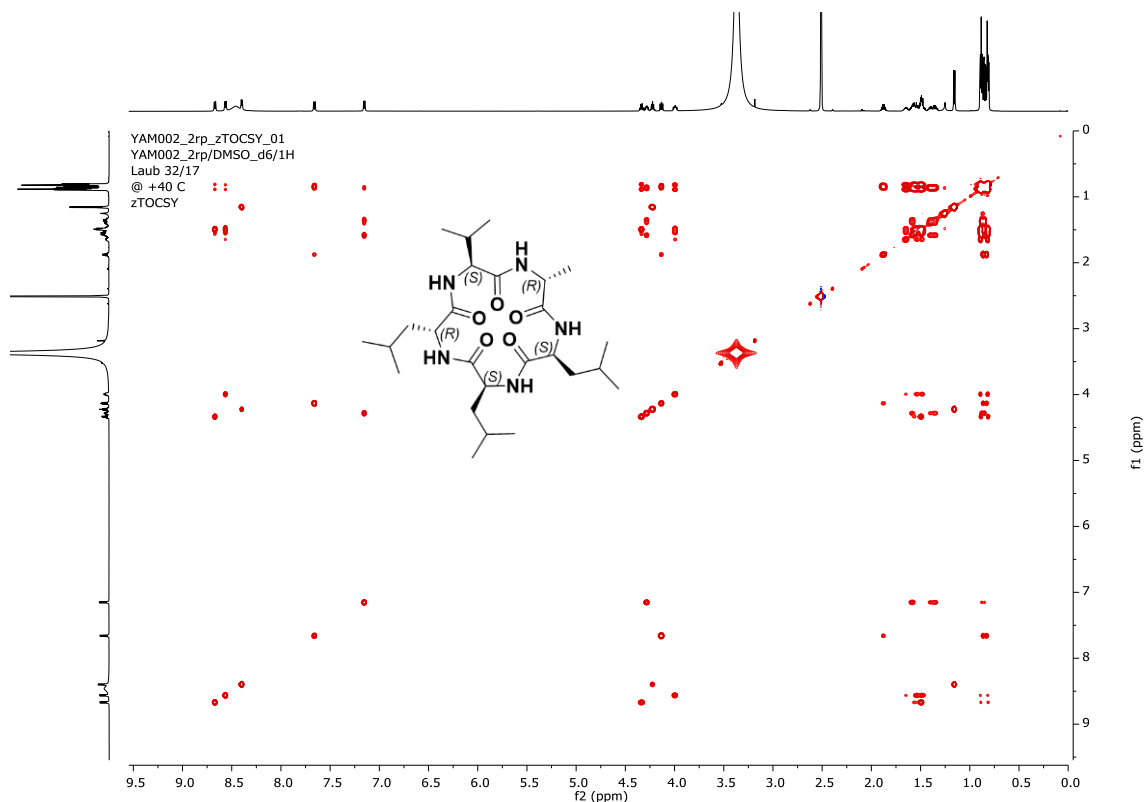
Figure B.38. ^{1}H, ^{1}H TOCSY spectrum of synthetic chrysosporide (**3.3**) (600 MHz, DMSO-*d6*, 40°C).
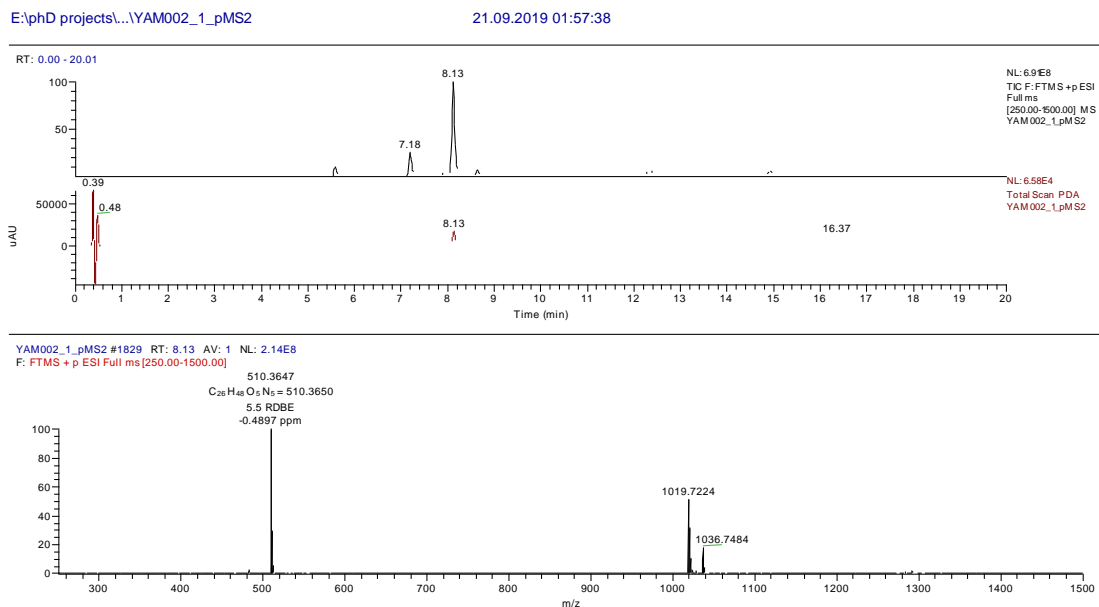
Figure B.39. UHPLC-(+)-ESI-HRMS and PDA spectrum of Valine-Epimer of synthetic chrysosporide (epi-**3.3**).
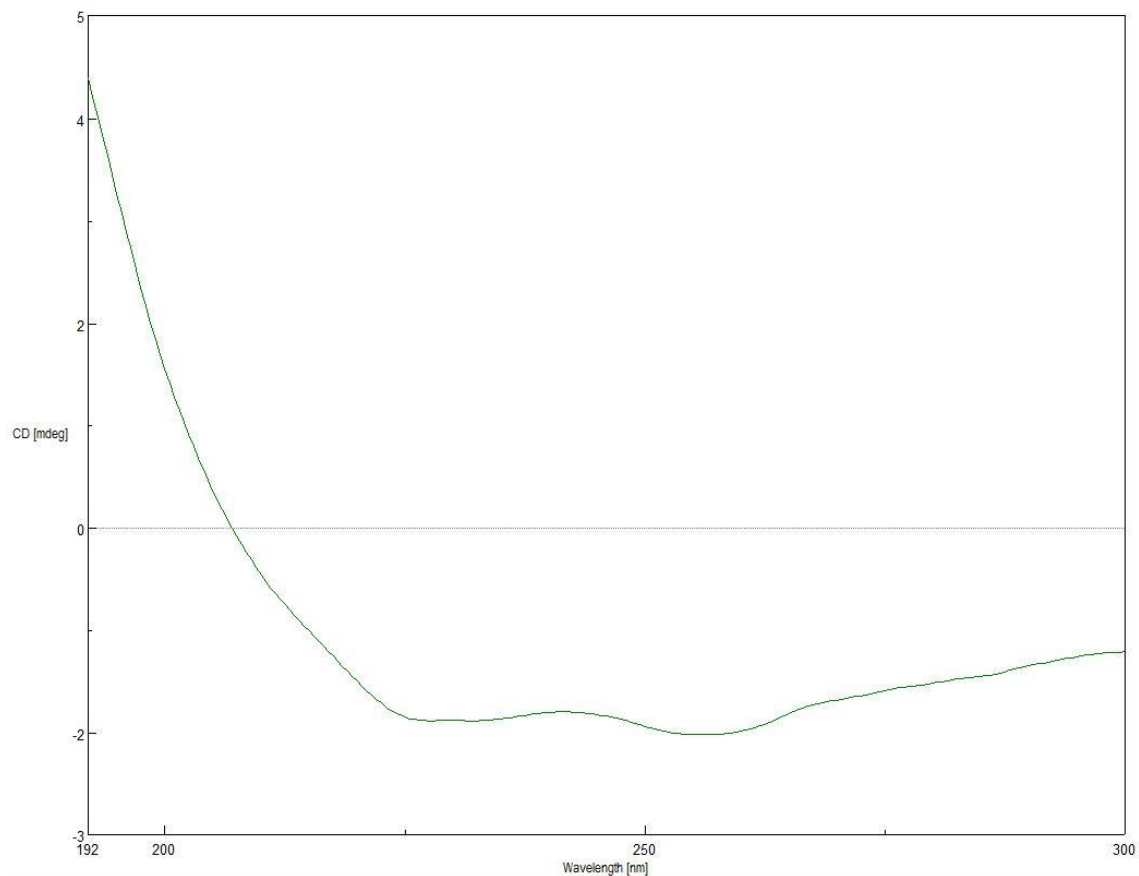
Figure B.40. CD spectrum of Valine-Epimer of synthetic chrysosporide (epi-**3.3**) in MeOH.
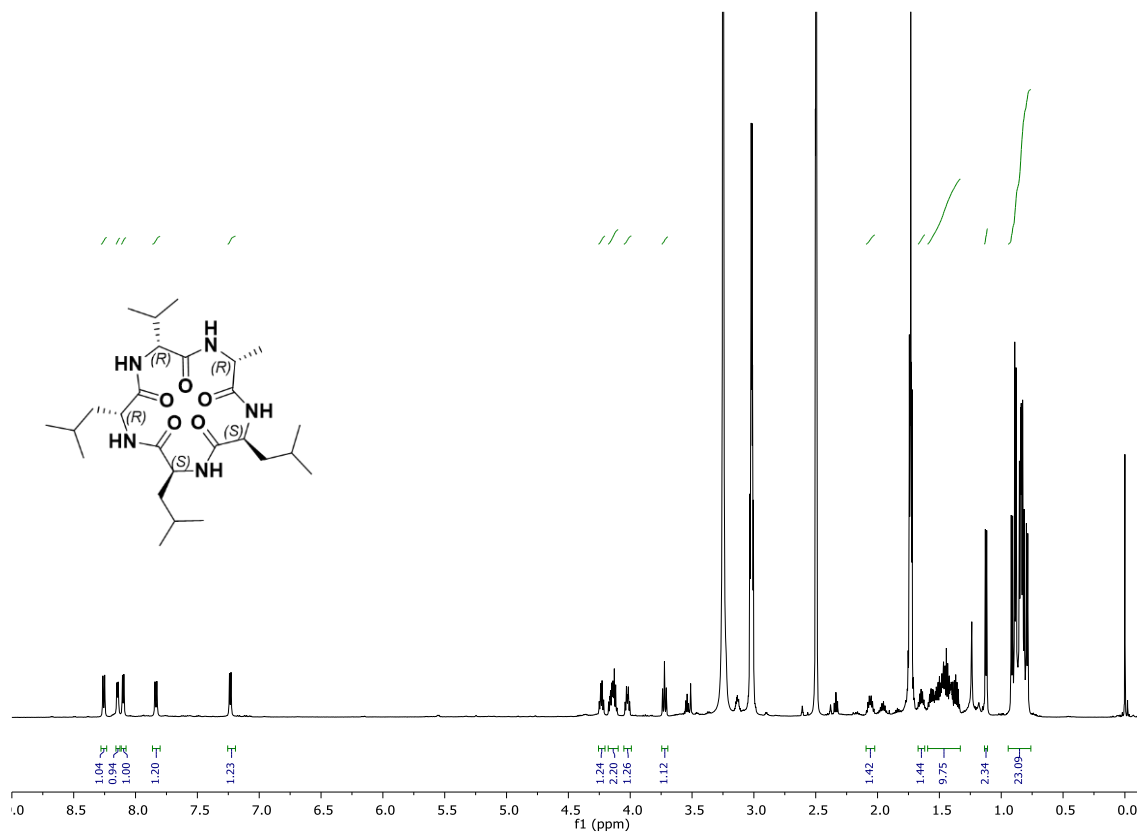
Figure B.41. ¹H NMR spectrum of Valine-Epimer of synthetic chrysosporide (epi-**3.3,** 600 MHz, DMSO-*d6*, 40°C).
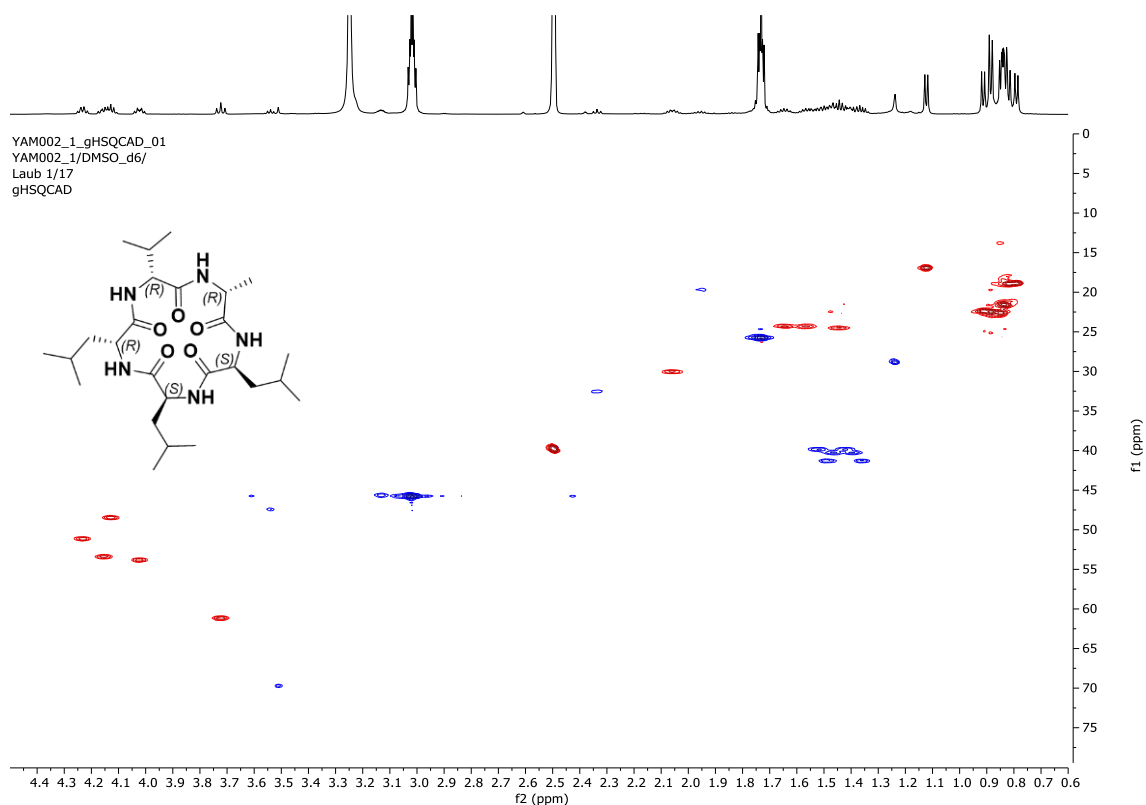
Figure B.42. ¹H, ¹³C HSQC spectrum of Valine-Epimer of synthetic chrysosporide (epi-**3.3,** 600 MHz, DMSO-*d6*, 40°C).
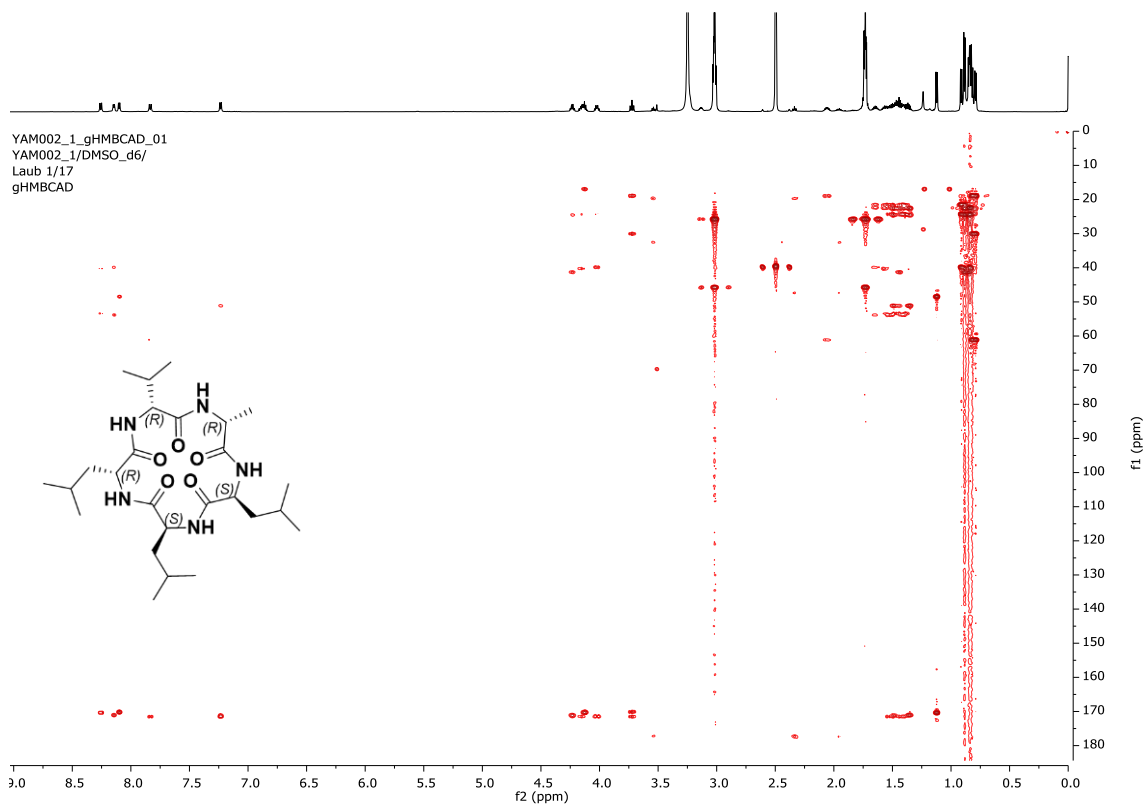
Figure B.43. ¹H, ¹³C HMBC spectrum of Valine-Epimer of synthetic chrysosporide (epi-**3.3,** 600 MHz, DMSO-*d6*, 40°C).
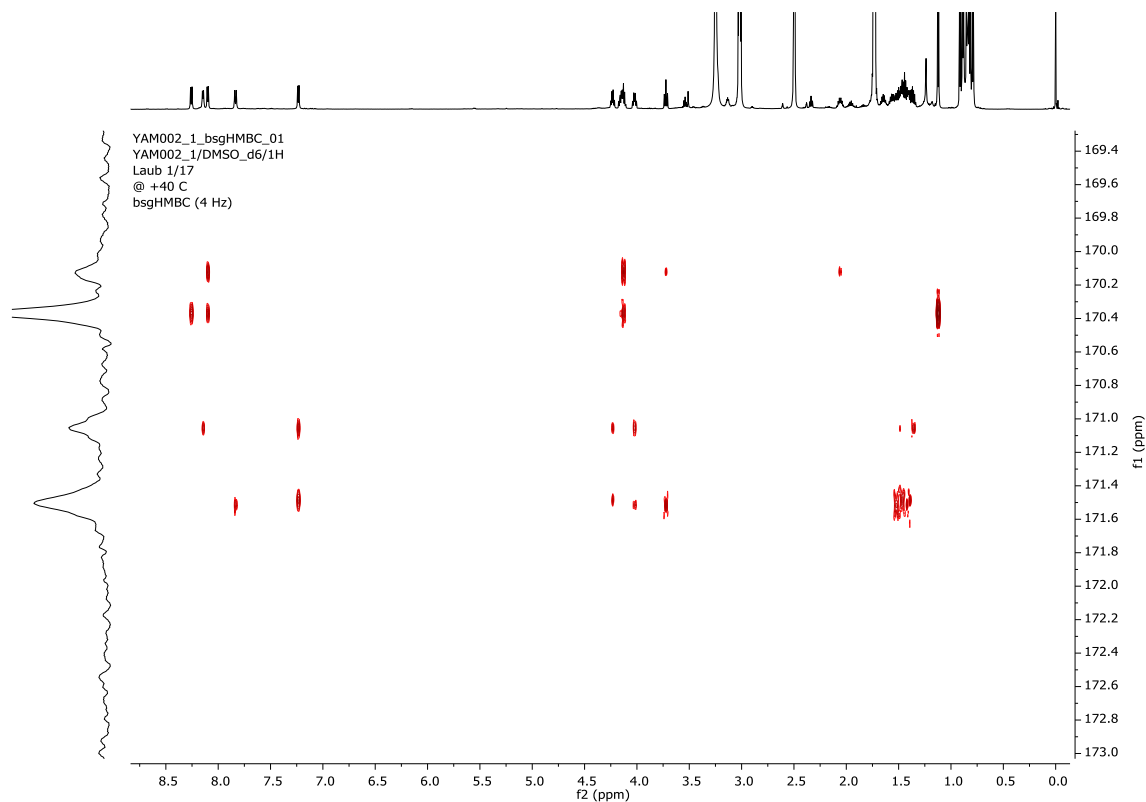
Figure B.44. ^{13}C band selective ^{1}H, ^{13}C HMBC spectrum of Valine-Epimer of synthetic chrysosporide (epi-**3.3,** 600 MHz, DMSO-*d6*, 40°C).
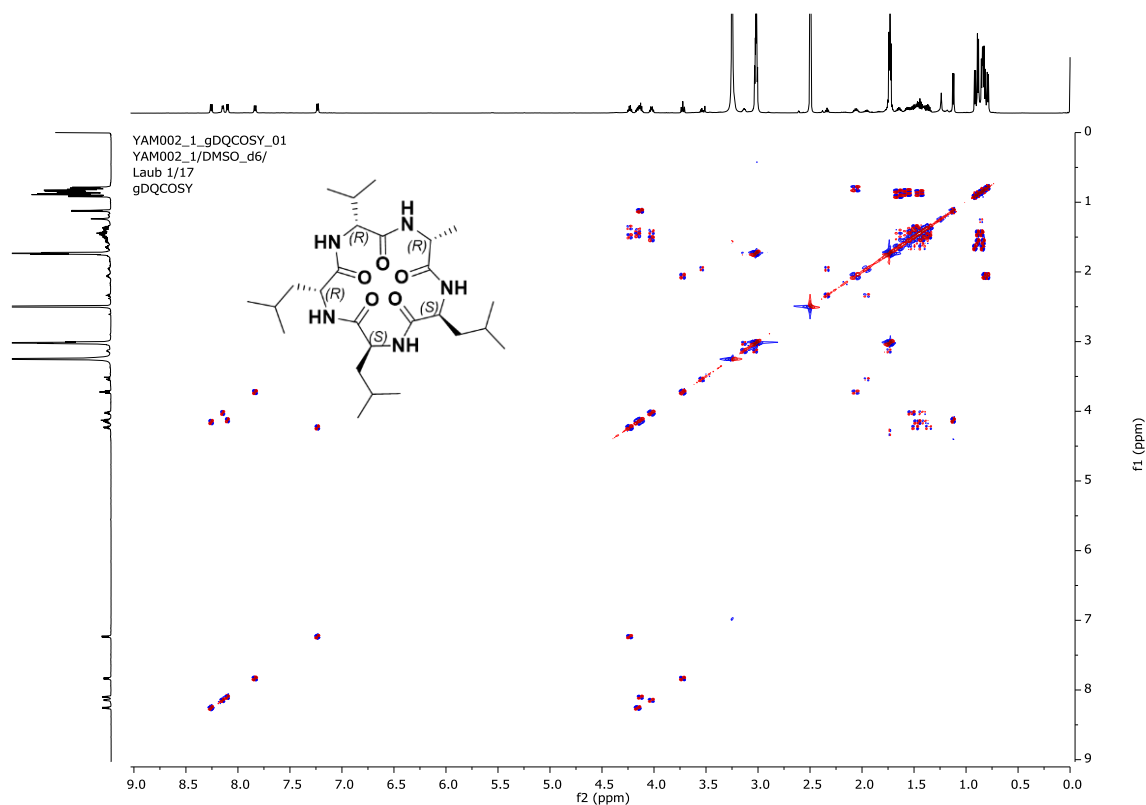
Figure B.45. ^{1}H, ^{1}H COSY spectrum of Valine-Epimer of synthetic chrysosporide (epi-**3.3,** 600 MHz, DMSO-*d6*, 40 °C).
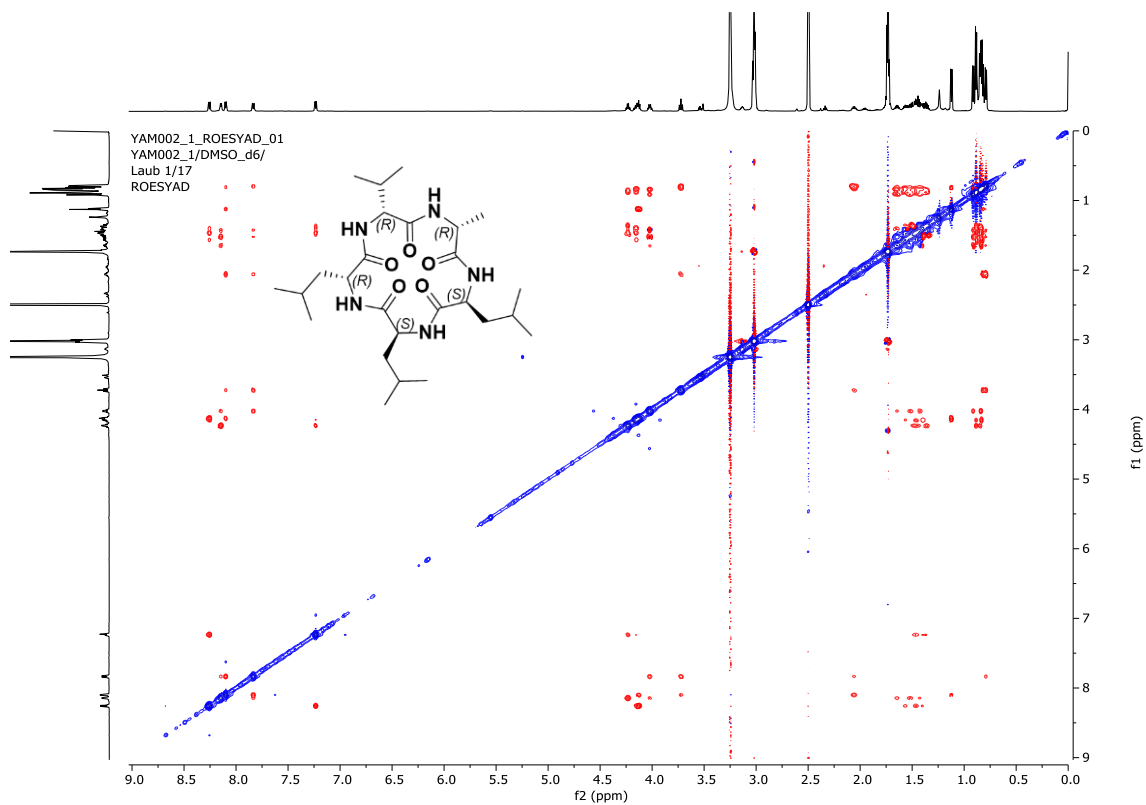
Figure B.46. [1]H, ^{1}H ROESY spectrum of Valine-Epimer of synthetic chrysosporide (epi-**3.3**, 600 MHz, DMSO-*d6*, 40°C).
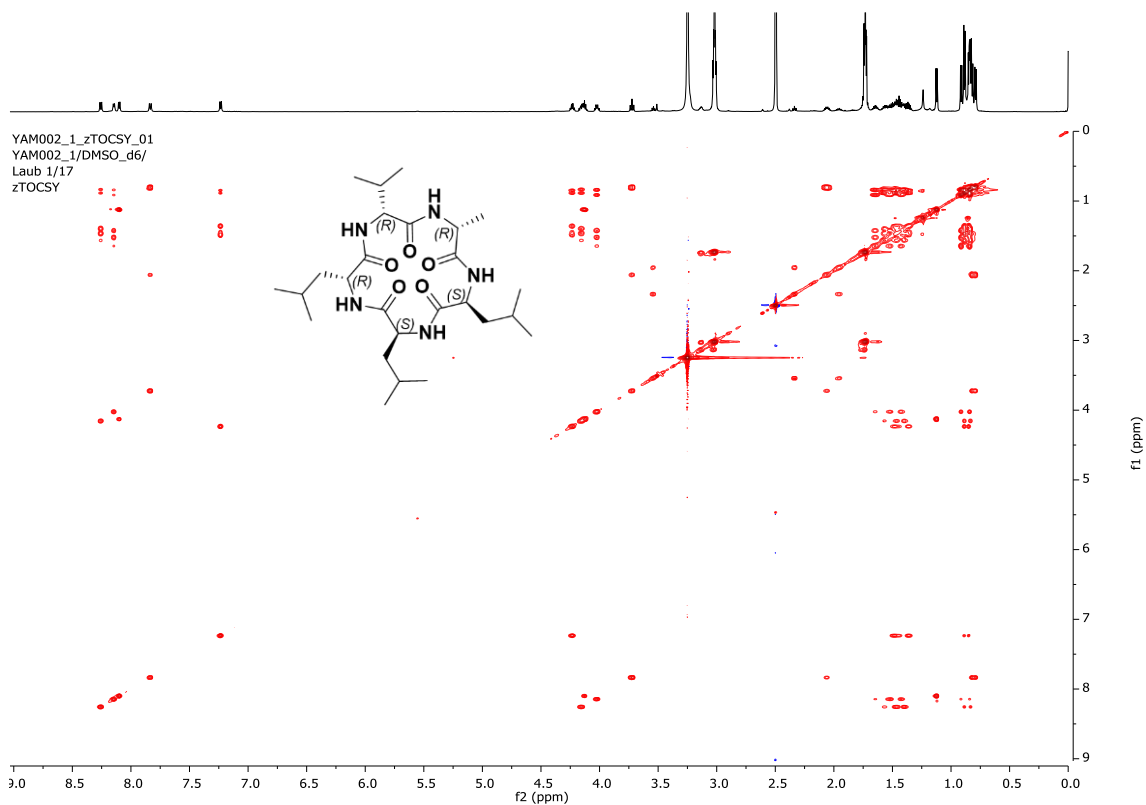
Figure B.47. [1]H, ^{1}H TOCSY spectrum of Valine-Epimer of synthetic chrysosporide (epi-**3.3**, 600 MHz, DMSO-*d6*, 40°C).
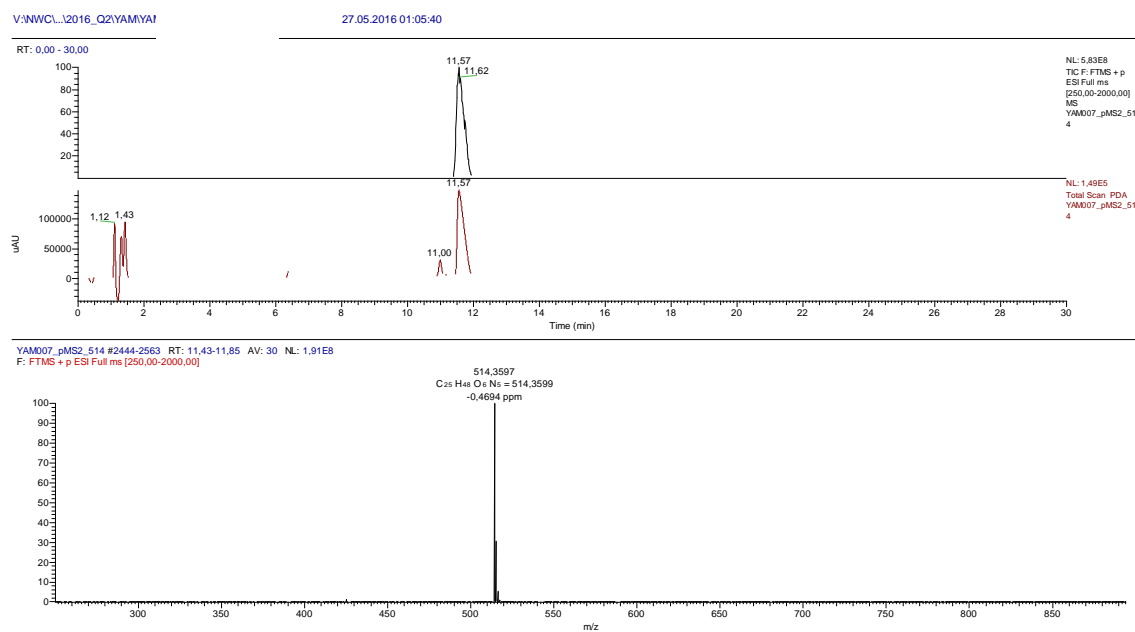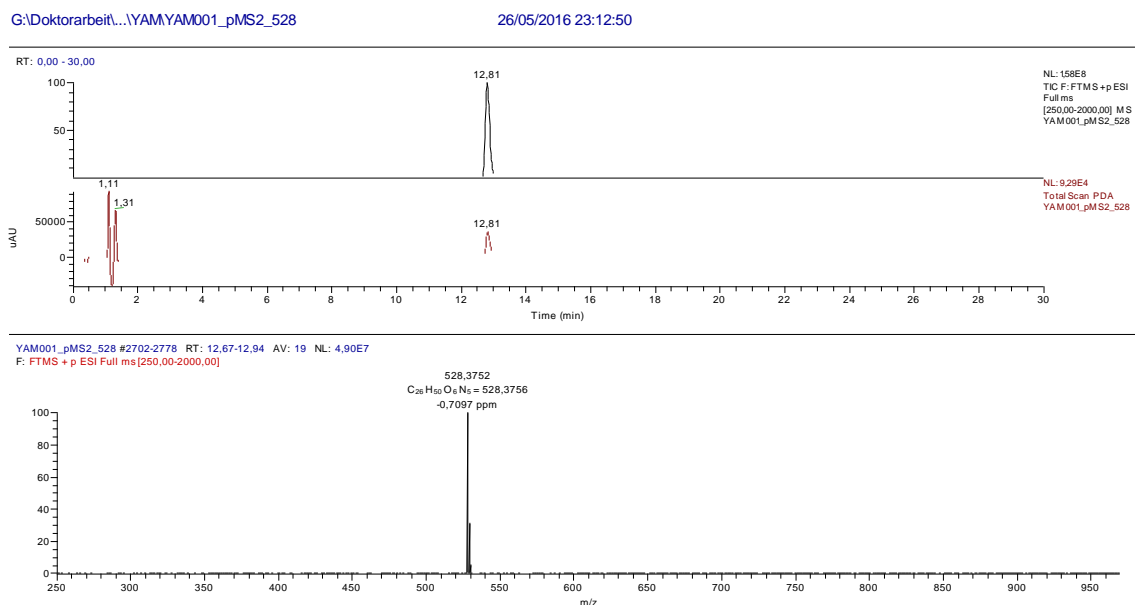
Figure B.48. UHPLC-(+)-ESI-HRMS and PDA spectrum of linear peptide precursor **I** of synthetic cyclic peptide **3.1a** (D-Leu-L-Leu-D-Val-L-Val-D-Ala).
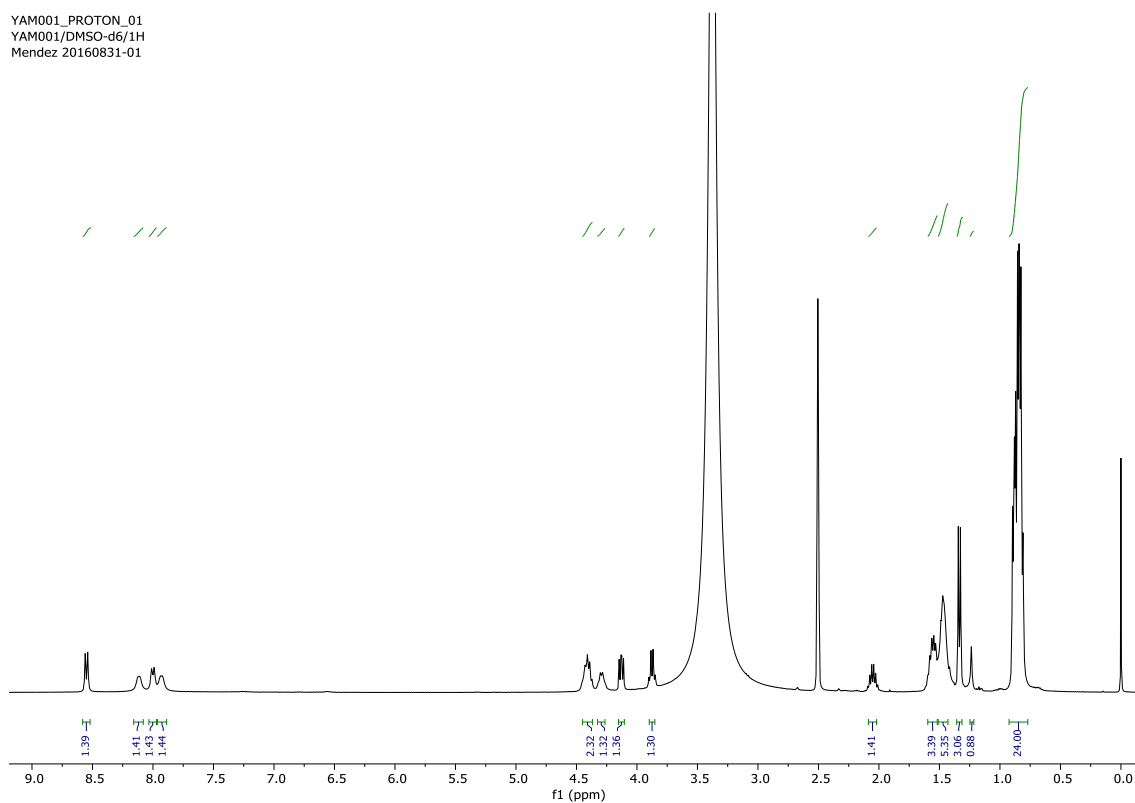

Figure B.49. ^{1}H NMR spectrum of linear peptide precursor **3.4** of synthetic cyclic peptide **3.1a** (L-Leu-L-Leu-D-Val-L-Val-D-Ala, 400 MHz, DMSO-*d6*, 25 °C).
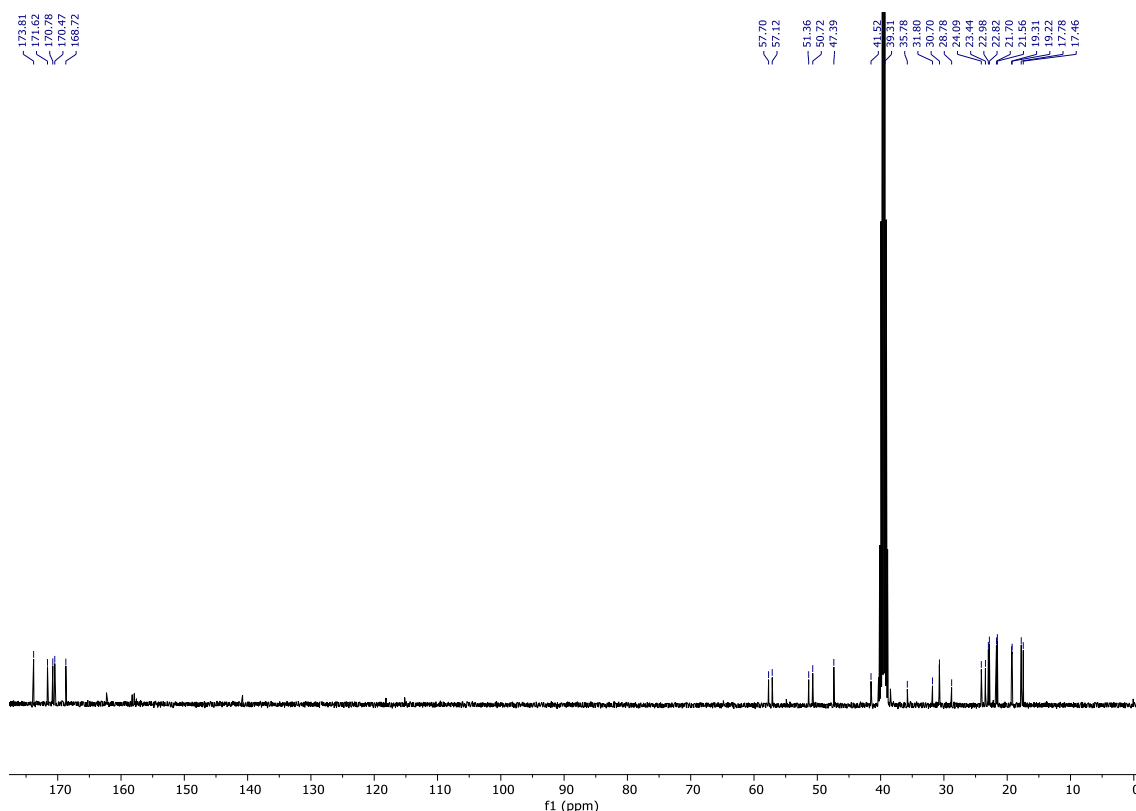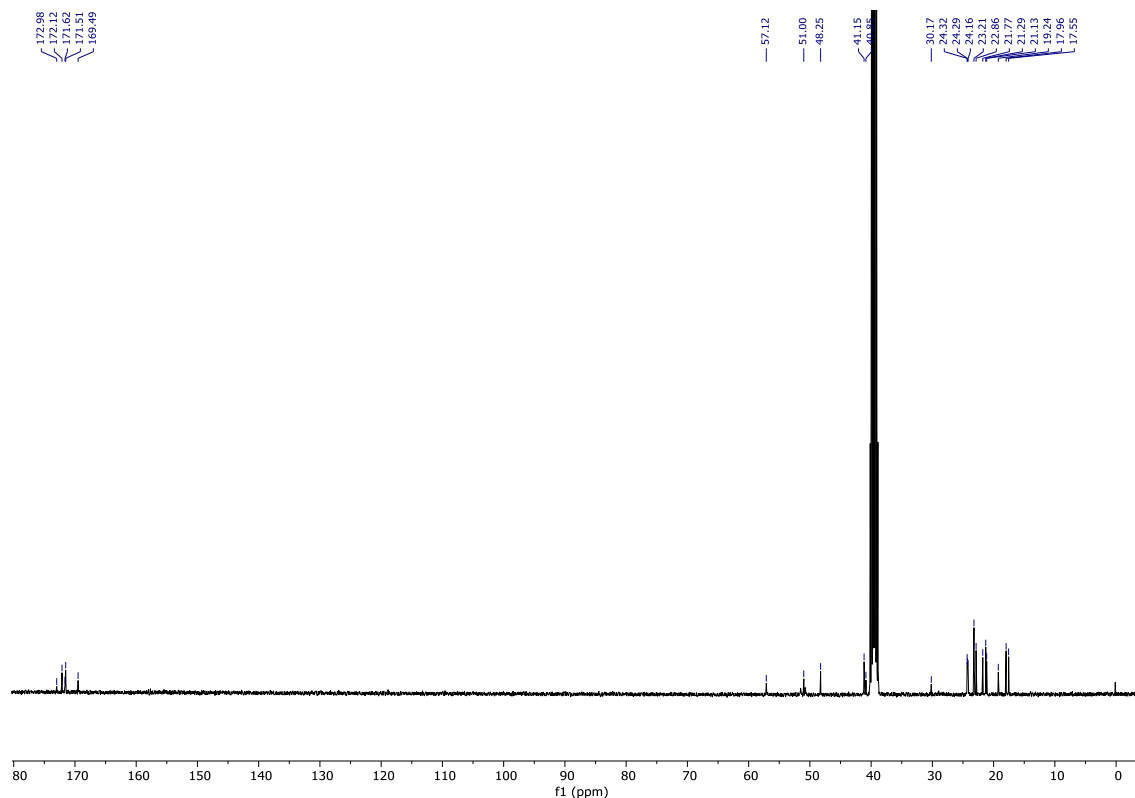
Figure B.50. ^{13}C NMR spectrum of linear peptide precursor 3.4 of synthetic cyclic peptide **3.1a** (L-Leu-L-Leu-D-Val-L-Val-D-Ala, 100 MHz, DMSO-*d6*, 25°C).
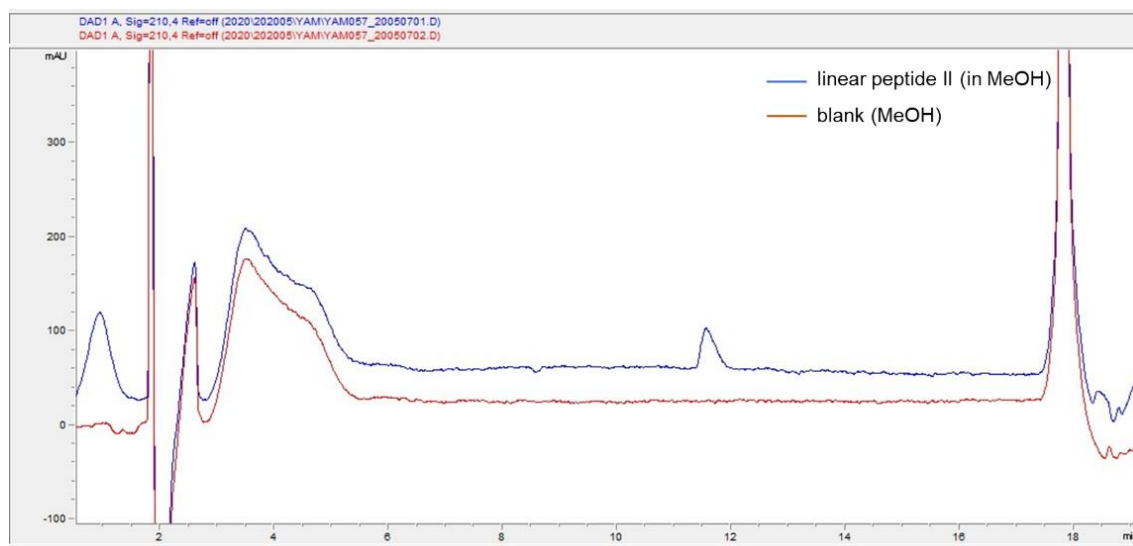
Figure B.51. HPLC profile of synthetic linear peptide **3.5**. Conditions: Agilent 1100/1260/1290 system, Macherey-Nagel ChromCart C18ec column (5 μm, 150 × 4.6 mm), flow rate 1.0 ml/min, 25°C, gradient 0-2 min, 20% B; 2-15 min, 20-25% B, 15-16 min 25-90%, isocratic at 90% B for further 7 min (A=H$_2$O + 0.1% FA; B=CH$_3$CN + 0.1% FA), absorbance detection at 210 nm.

YAM057_PROTON_01
YAM057/DMSO-d6/1H
Mendez 20160831_05
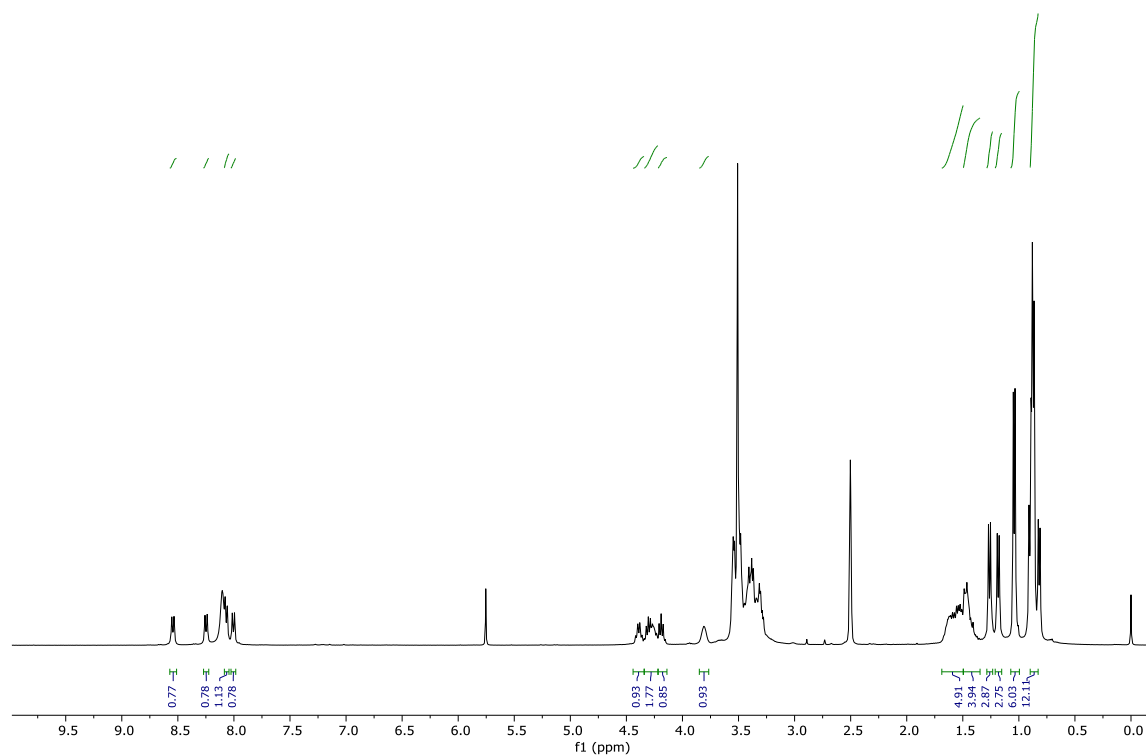
Figure B.52. ^{1}H NMR spectrum of linear peptide precursor **3.5** of synthetic cyclic peptide **3.2** (L-Leu-L-Leu-D-Leu-L-Ala-D-Ala, 400 MHz, DMSO-*d6*, 25 °C).
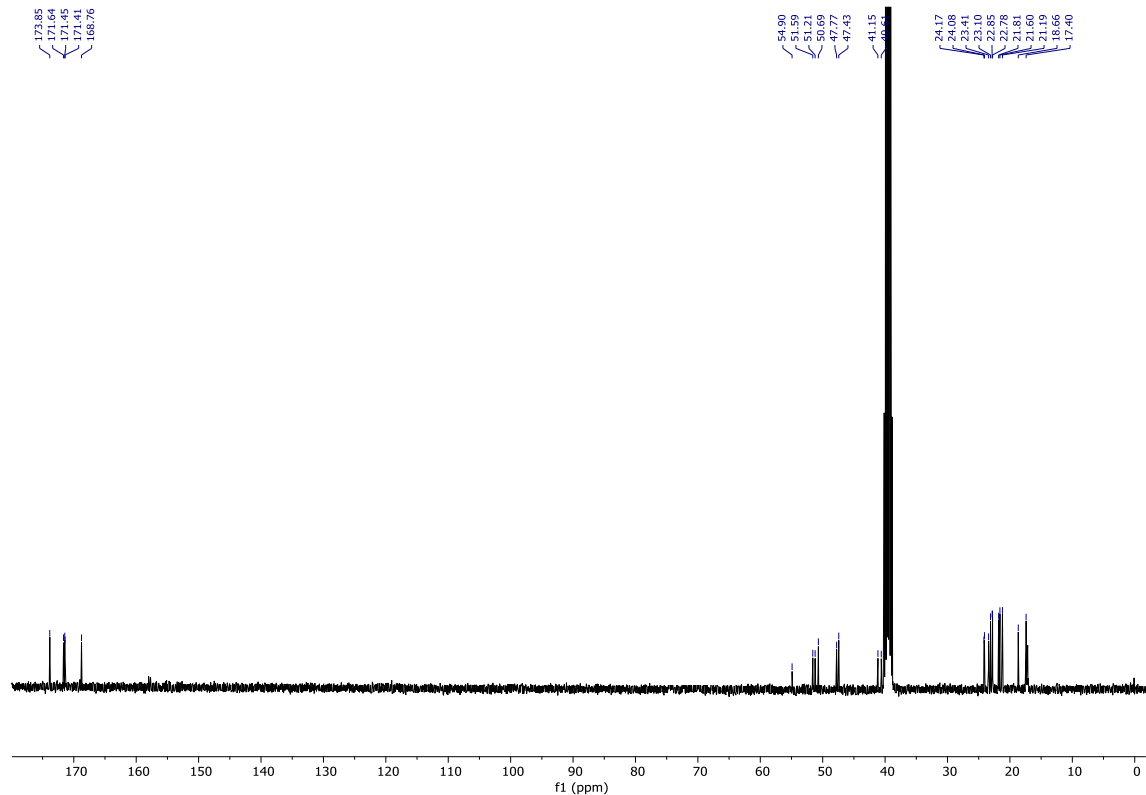
Figure B.53. ^{13}C NMR spectrum of linear peptide precursor **3.5** of synthetic cyclic peptide **3.2** (L-Leu-L-Leu-D-Leu-L-Ala-D-Ala, 100 MHz, DMSO-*d6*, 25°C).

Figure B.54. UHPLC-(+)-ESI-HRMS and PDA spectrum of linear peptide precursor **3.6** of synthetic chrysosporide (**3.3**) (D-Ala-L-Leu-L-Leu-D-Leu-L-Val).

Figure B.55. ^{1}H NMR spectrum of linear precursor **3.6** of synthetic chrysosporide (**3.3**) (D-Ala-L-Leu-L-Leu-D-Leu-L-Val, 400 MHz, DMSO-*d6*, 25 °C).

Figure B.56. ^{13}C NMR spectrum of linear peptide precursor **3.6** of synthetic chrysosporide (**3.3**) (D-Ala-L-Leu-L-Leu-D-Leu-L-Val, 100 MHz, DMSO-*d6*, 25 °C).
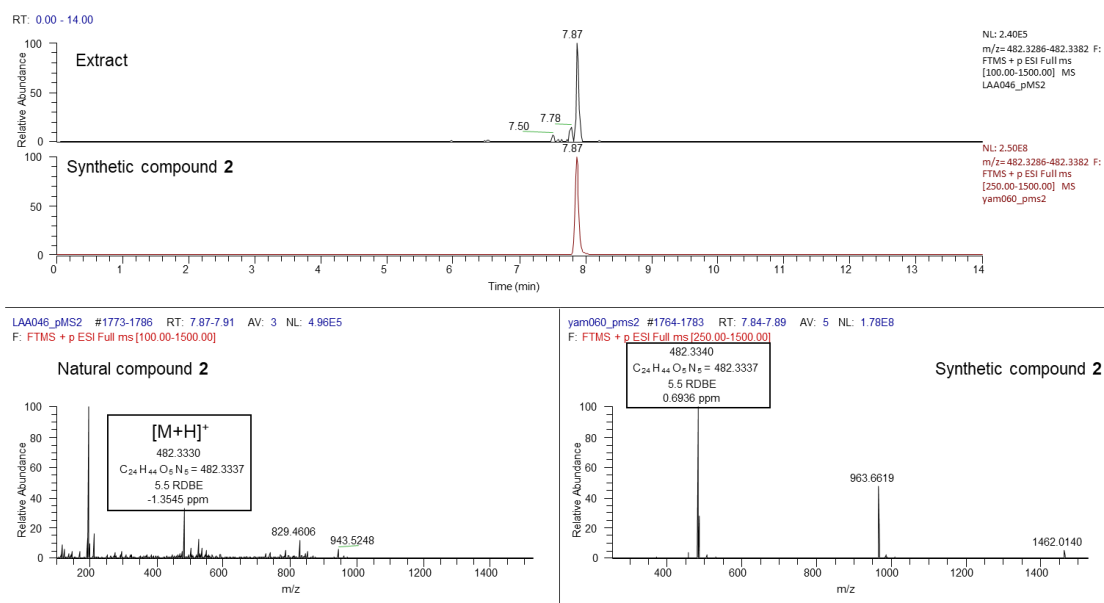
Figure B.57. UHPLC-(+)-ESI-HRMS comparison of extracted ion chromatograms *m/z* 482 [M+H]$^+$ of the natural microsporide B (**3.2**) within in the crude extract of *S. microspermum* and the corresponding synthetic cyclic pentapeptide (**3.2**).
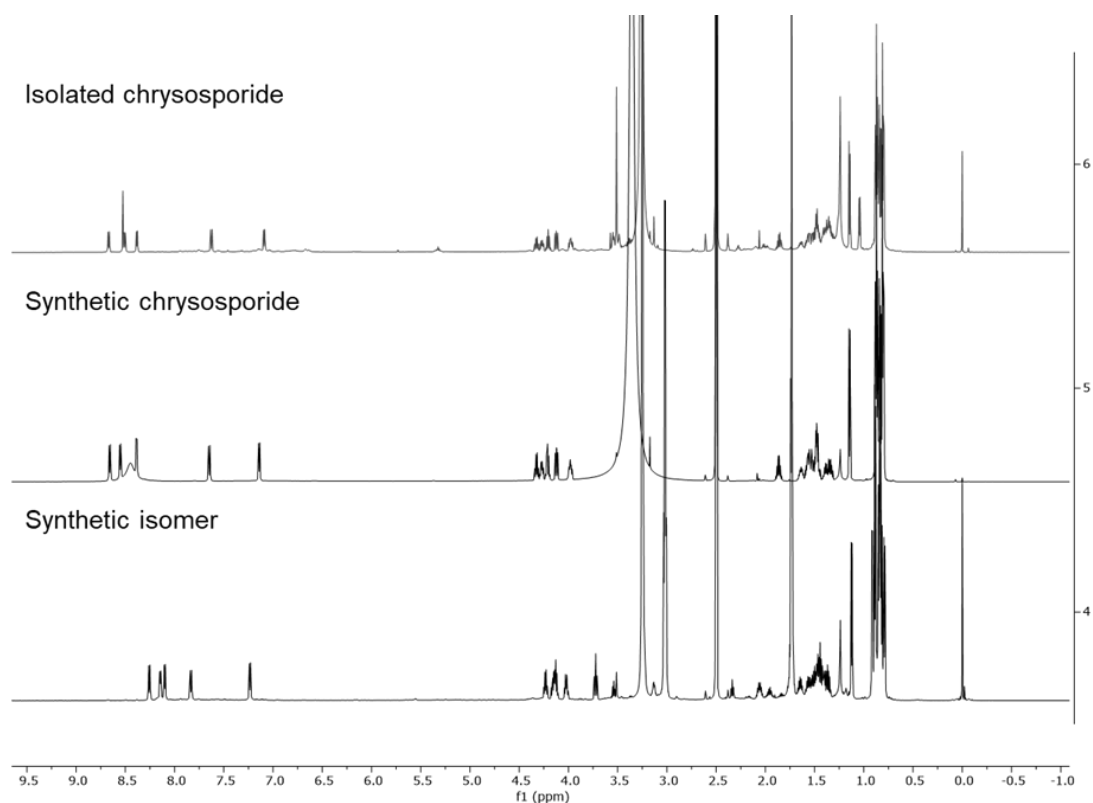
Figure B.58. Comparison of ^{1}H NMR spectra of the isolated and synthetic chrysosporide (**3.3**) and Valine-Epimer of synthetic chrysosporide (epi-**3.3**) (600 MHz, DMSO-*d6*, 40 °C).
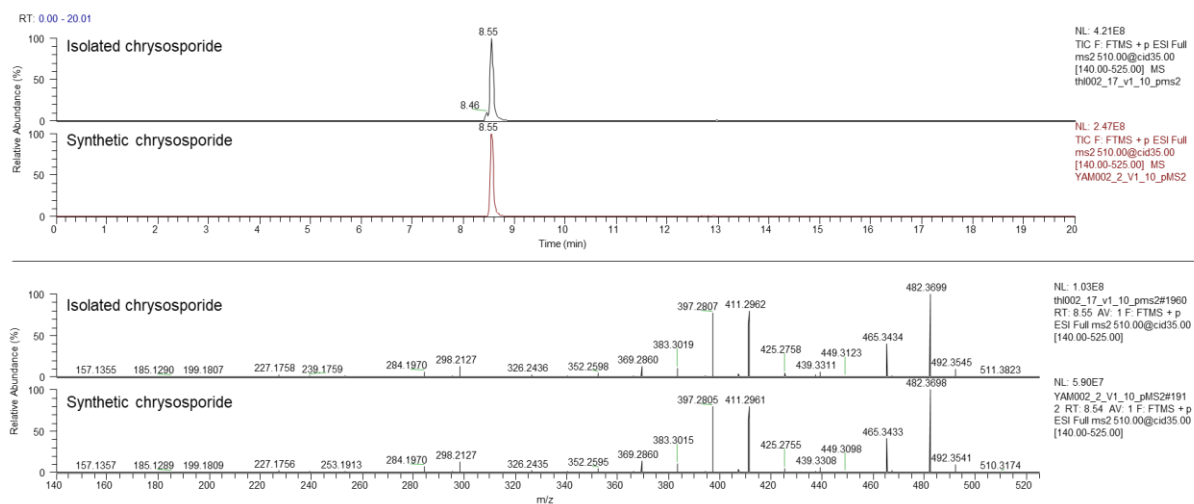
Figure B.59. Comparison of the TICs and the fragmentation pattern obtained from the UHPLC-(+)-ESI-HRMS^2 measurements of the isolated and synthetic chrysosporide (**3.3**).
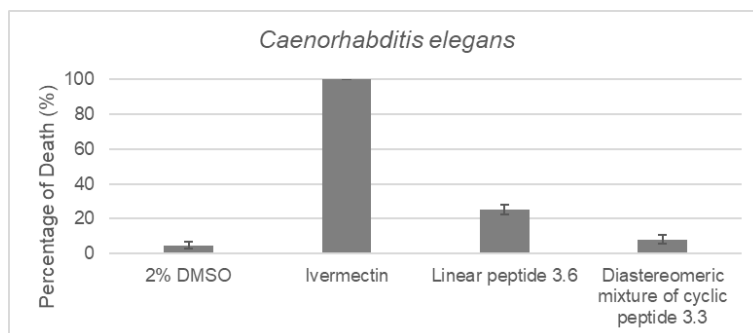
Figure B.60. Anthelmintic activity of synthetic linear peptide **3.6** and diastereomeric mixture of cyclic peptide **3.3**.
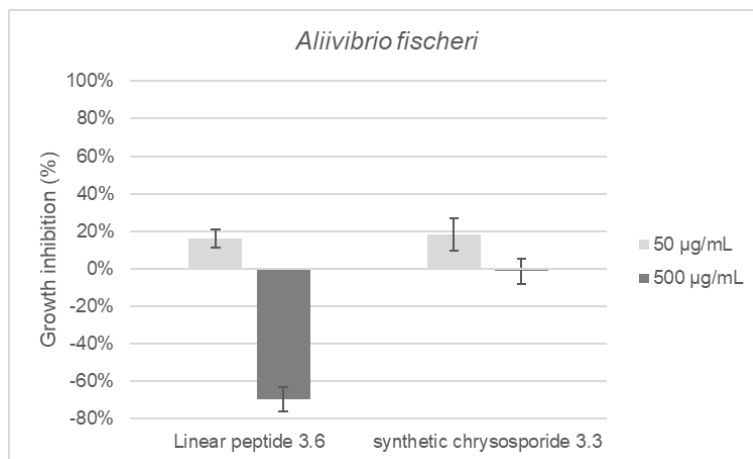

Figure B.61. Antibacterial activity of synthetic linear peptide **3.6** and synthetic chrysosporide **3.3**.

C.   Appendix Chapter 4

Table C.1. Origin of fungal material.

Compound	Elemental composition	[M-H]⁻ Theoretical mass	[M-H]⁻ Experimental mass (error)					
			C. (D.) *austronanceiensis*	*C.* (D.) *icterina*	*C.* (D.) *icterinula*	*C. (D) obscuro-olivea*	*C.* (D.) spec.	*C.* (D.) *viridulifolius*
emodin (**4.1**)	$C_{15}H_9O_5^-$	269.0455	269.0450 (-2.0 ppm)	269.0451 (-0.8 ppm)	269.0448 (-2.6 ppm)	269.0450 (-1.9 ppm)	269.0449 (-2.2 ppm)	269.0453 (-0.8 ppm)
physicion (**4.2**)	$C_{16}H_{11}O_5^-$	283.0612	283.0611 (-0.5 ppm)	283.0610 (-0.8 ppm)	n.d.	283.0615 (-0.7 ppm)	283.0605 (-2.6 ppm)	283.0610 (-0.6 ppm)
endocrocin (**4.3**)	$C_{16}H_9O_7^-$	313.0354	313.0349 (-1.5 ppm)	313.0353 (-0.3 ppm)	313.0351 (-1.0 ppm)	313.0348 (-1.7 ppm)	n.d.	313.0351 (-1.0 ppm)
dermolutein (**4.4**)	$C_{17}H_{11}O_7^-$	327.0510	327.0505 (-1.7 ppm)	327.0505 (-1.6 ppm)	327.0503 (-2.1 ppm)	327.0506 (-1.4 ppm)	327.0501 (-2.9 ppm)	327.0507 (-1.1 ppm)
hypericin (**4.5**)	$C_{30}H_{15}O_8^-$	503.0772	503.0763 (-1.8 ppm)	n.d.	n.d.	503.0766 (-1.3 ppm)	n.d.	503.0767 (-1.1 ppm)
skyrin (**4.6**)	$C_{30}H_{17}O_{10}^-$	537.0827	537.0817 (-1.8 ppm)	n.d.	n.d.	537.0819 (-1.6 ppm)	n.d.	537.0822 (-1.0 ppm)

[1] Leg. = (lat.) legit, det. = (lat.) determinavit.

Table C.2. Detected anthraquinones (**4.1**-**4.6**), their elemental composition and exact masses.

Species	Koll.	Date	Place	Leg./det.[1]
C. (D.) *austronanceiensis*	Chile 34/12	05.05.2012	Conguillío National Park	Arnold, Palfner
C. (D.) *icterina*	Chile 24/12	05.05.2012	Conguillío National Park	Arnold, Palfner
C. (D.) *icterinula*	Chile 3/12	27.04.2012	Curacautin, Laguna Blanca	Arnold, Palfner
C. (D.) *obscuro-olivea*	Chile 37/14	June 2014	Curacautin, Laguna Blanca	Arnold, Palfner
C. (D.) spec.	Chile 32/12	05.05.2012	Conguillío National Park	Arnold
C. (D.) *viridulifolius*	Chile 44/11	June 2011	Quillon, Cayumanque	Arnold

Table C.3. Key ions in the negative ion ESI-MSⁿ spectra of skyrin (**4.6**).

| Compound	Method	Scan Mode [*m/z*]	*m/z* [relative Intensity (%)]
skyrin (**4.6**) (fungal extract)	DESI	MS² (50%) *m/z* 537	493.0923 ([M-H-CO₂]⁻, **100**), 469.0926 ([M-H-C₃O₂]⁻, 80)
skyrin (**4.6**) (authentic reference compound)		MS² (35%) *m/z* 537	520.0783 ([M-H-OH]·⁻,), 519.0716 ([M-H-H₂O]⁻), 509.0861 ([M-H-CO]⁻), 493.0920 ([M-H-CO₂]⁻), 475.0809 ([M-H-H₂O-CO₂]⁻), 469.0922 ([M-H-C₃O₂]⁻), 465.0956 [M-H-CO-CO₂]⁻, 449.1021 ([M-H-₂CO₂]⁻)
		MS³ (35%) *m/z* 537-493	493.0908 ([M-H- M-H-CO]⁻, 64), 475.0805 ([M-H-H₂O-CO₂]⁻, 46), 465.0963 ([M-H-CO-CO₂]⁻, **100**), 449.1014 ([M-H-2CO₂]⁻, 97)
skyrin (**4.6**) (authentic reference compound)	direct infusion	MS² (30%) *m/z* 537	537.0829 ([M-H]⁻, 96), 520.0801 ([M-H-OH]·⁻, 16), 519.0723 ([M-H-H₂O]⁻, 19), 509.0879 ([M-H-CO]⁻, 11), 493.0930 ([M-H-CO₂]⁻, **100**), 475.0824 ([M-H-H₂O-CO₂]⁻, 7), 469.0931 ([M-H-C₃O₂]⁻, 72), 465.0981 ([M-H-CO-CO₂]⁻, 8), 449.1033 ([M-H-2CO₂]⁻, 14)
		MS³ (40%) *m/z* 537-493	493.0930 ([M-H-CO₂]⁻, 34), 475.0824 ([M-H-H₂O-CO₂]⁻, 61), 465.0980 ([M-H-CO-CO₂]⁻, 86), 449.1032 ([M-H-2CO₂]⁻, **100**), 421.1083 ([M-H-2CO₂-CO]⁻, 8)
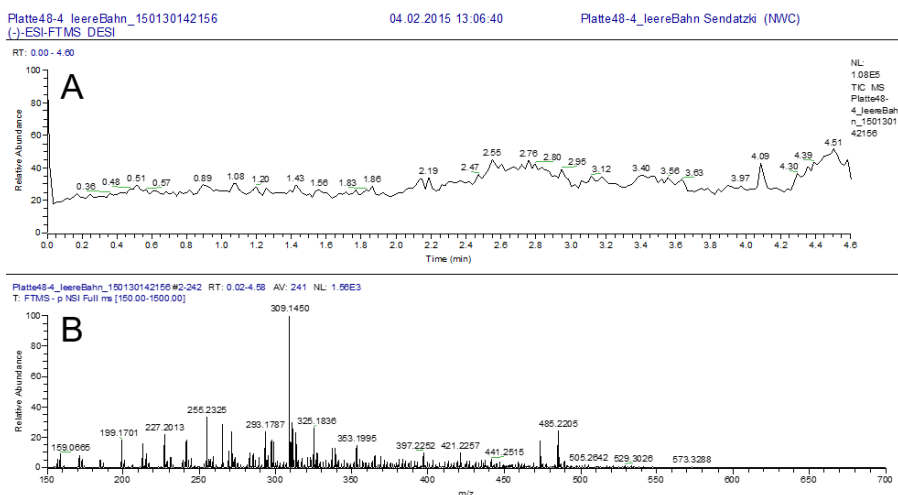
Figure C.1. A) Total ion chromatogram of an empty HPTLC band after development with eluent system (toluene: ethyl formate: formic acid (10:5:3; *v/v/v*)) , B) Corresponding Full MS spectrum to A (averaged over t_R 0 - 4.6 min) showing background related peaks.
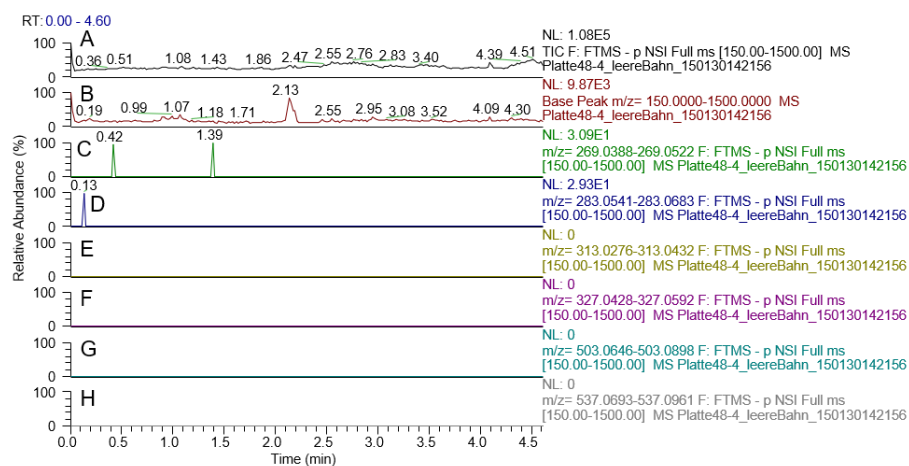

Figure C.2. Total ion (A), base peak (B) and extracted ion chromatograms of an unspotted HPTLC band showing no anthraquinone related peaks (C - H).
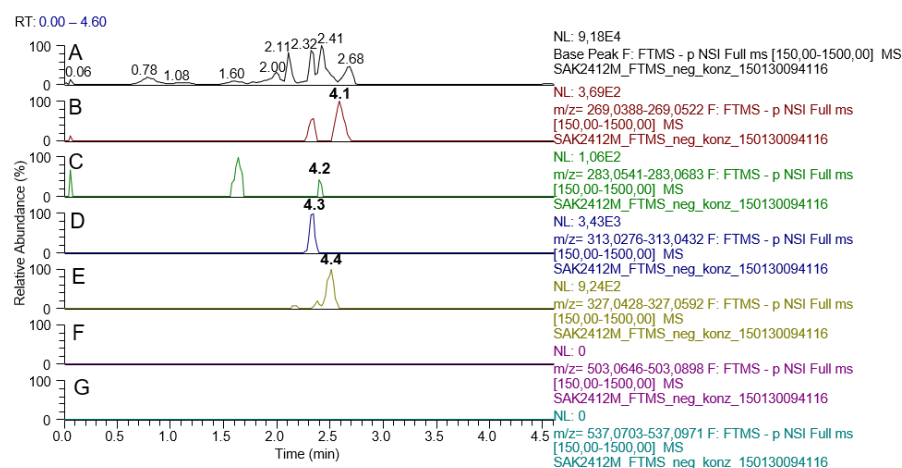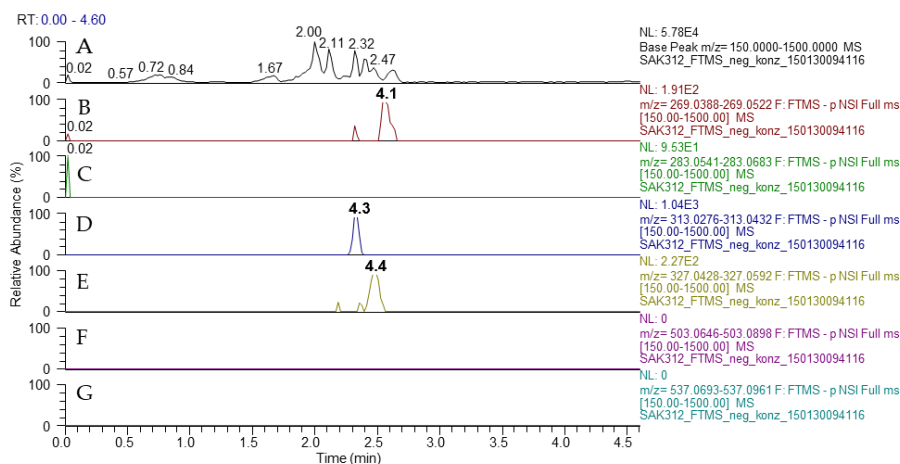

Figure C.3. Base peak chromatogram (A) and extracted ion chromatograms (EICs, B - G) based on the theoretical masses of the investigated anthraquinones (**4.1** – **4.6**) obtained from the methanolic crude extract of *Cortinarius* (D.) *icterina*.

Figure C.4. Base peak chromatogram (A) and extracted ion chromatograms (EICs, B - G) based on the theoretical masses of the investigated anthraquinones (**4.1** – **4.6**) obtained from the methanolic crude extract of *Cortinarius* (D.) *icterinula*.
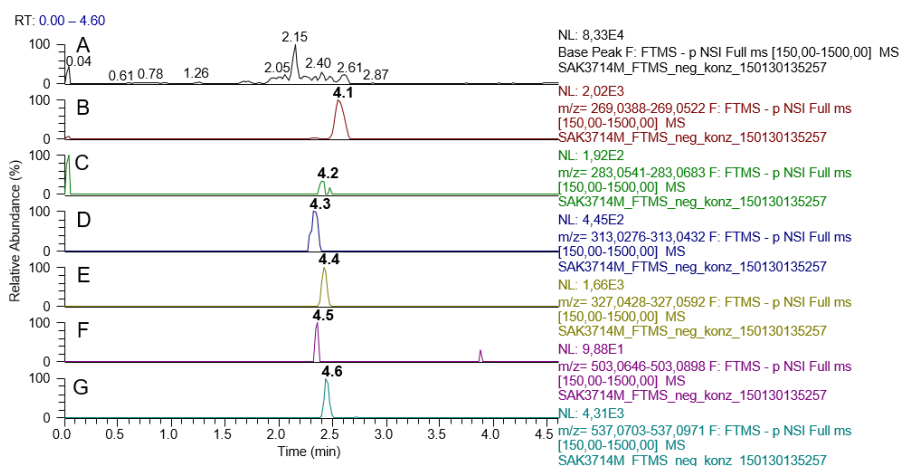

Figure C.5. Base peak chromatogram (A) and extracted ion chromatograms (EICs, B - G) based on the theoretical masses of the investigated anthraquinones (**4.1** – **4.6**) obtained from the methanolic crude extract of Cortinarius (D.) obscuro-olivea.
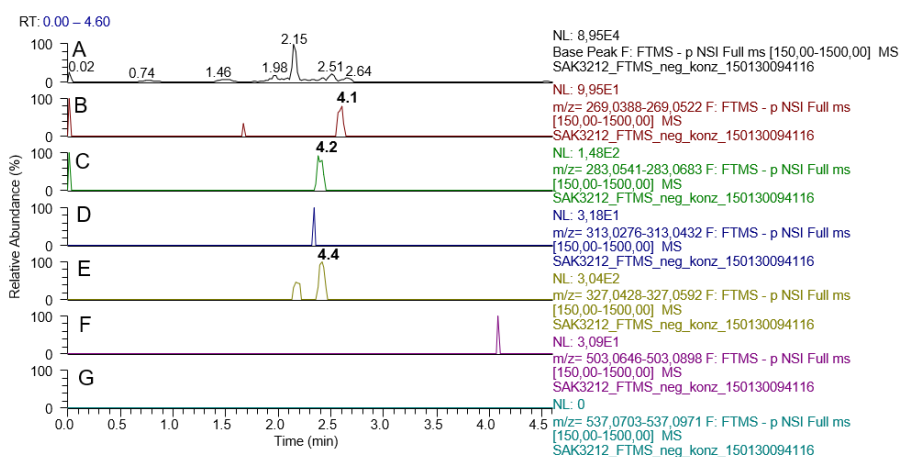

Figure C.6. Base peak chromatogram (A) and extracted ion chromatograms (EICs, B - G) based on the theoretical masses of the investigated anthraquinones (**4.1** – **4.6**) obtained from the methanolic crude extract of *Cortinarius* (D.) spec.
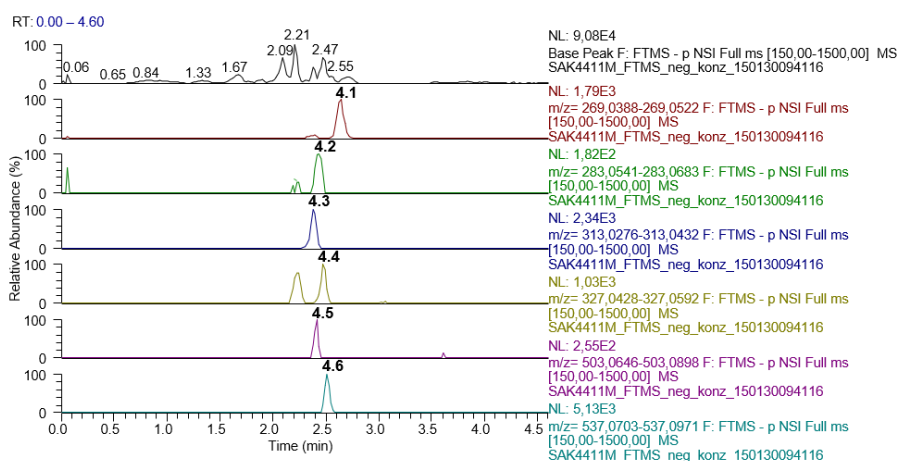
Figure C.7. Base peak chromatogram (A) and extracted ion chromatograms (EICs, B - G) based on the theoretical masses of the investigated anthraquinones (**4.1** – **4.6**) obtained from the methanolic crude extract of of *Cortinarius* (D.) *viridulifolius.*
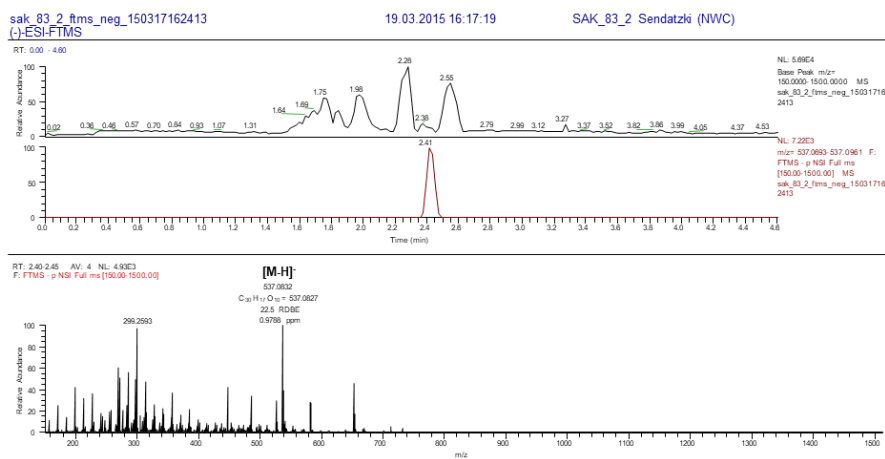

Figure C.8. Total ion and extracted ion chromatogram (EIC) of reference compound skyrin (**4.6**) and the corresponding HRMS spectrum.
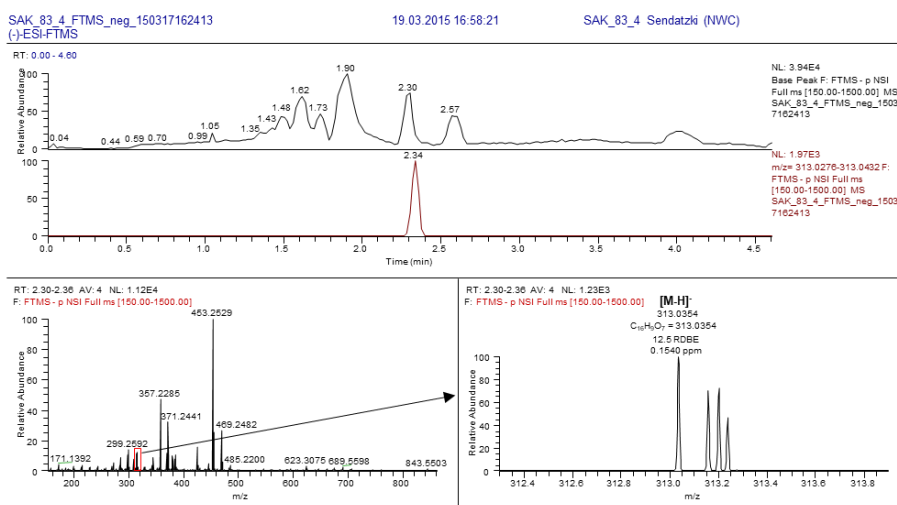

Figure C.9. Total ion and extracted ion chromatogram EIC) of reference compound endocrocin (**4.3**) and the corresponding HRMS spectrum.
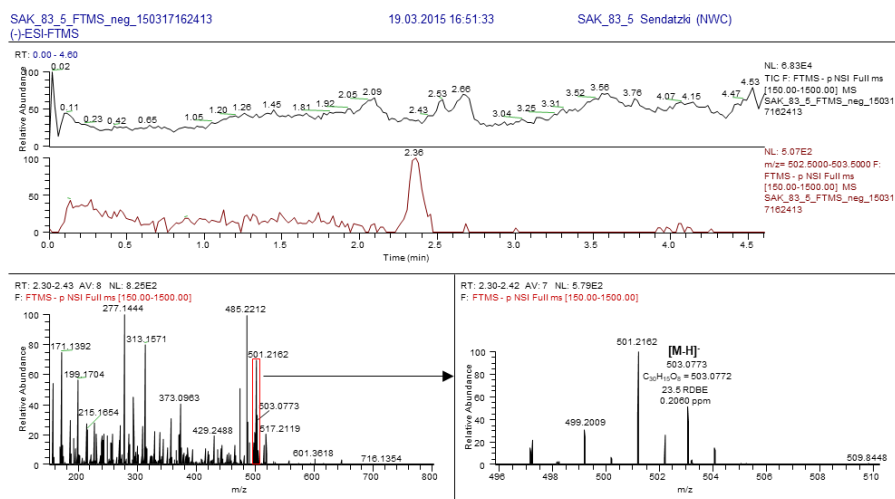
Figure C.10. Total ion and extracted ion chromatogram (EIC) of reference compound hypericin (**4.5**) and the corresponding HRMS spectrum.
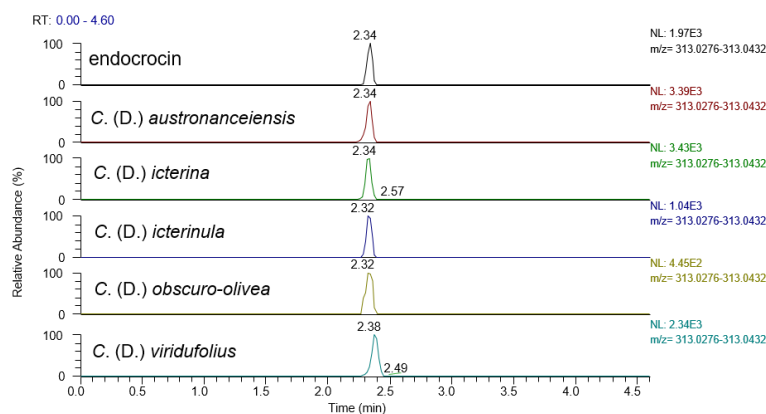

Figure C.11. Comparison of extracted ion chromatograms (EICs) of endocrocin (4.3, *m/z* 313) acquired during DESI-HRMS measurement of methanolic extract from *C.* (*D.*) species and the reference compound.

Declaration on the authors contribution

Chapter 2. Laub, A.; Franke, K.; Schmidt, J.; Porzel, A.; Abrahamczyk, S.; Weigend, M.; and Wessjohann, L.A. *Secondary metabolic profiling of 31 Impatiens species and Hydocera triflora (Balsaminaceae) leaf extracts by LC-HRMS and NMR in combination with chemometrics.* Manuscript in preparation.

In this project, Annegret Laub performed all practical work, including all analytical measurements, and data analysis. Andrea Porzel performed the NMR measurements of the isolated and reference compounds. Katrin Franke and Jürgen Schmidt provided help with data analysis, interpretation and mentoring. Stefan Abrahamczyk and Maximilian Weigend provided the plant material. Annegret Laub wrote the first draft, edited by Katrin Franke, Jürgen Schmidt, and Ludger A. Wessjohann. This project was planned, supervised, and edited by Ludger Wessjohann and Katrin Franke.

Chapter 3. Laub, A.; Lam, Y.T.H; Mendez, Y.; Vasco Vidal, A.; Porzel, A.; Schmidt, J.; Wessjohann, L.A.; Westermann, B.; and Arnold, N. *Identification and Total Synthesis of Two New Cyclic Pentapeptides from Sepedonium microspermum Besl.* Manuscript in preparation.

Annegret Laub carried out the cultivation, extraction, and screening of 16 *Sepedonium* strains and measurements of the isolated and synthesis products via UHPLC-ESI-HRMS^n as well as the ESI-HRMS^n and NMR data analysis. Yen Lam performed the cultivation of *S microspermum* in big scale and the isolation of the cyclic peptides, final purification of the natural products via LC-MS was performed by Annegret Laub. The total synthesis of the cyclic pentapeptides was performed by Yanira Gomez and the final purification was performed by Annegret Laub. Aldrin Vasco Vidal and Andrea Porzel performed the NMR measurements and supported the NMR data analysis. Aldrin Vasco Vidal performed the CD measurements and CD data evaluation. Annegret Laub wrote the major part of the manuscript, Aldrin Vasco Vidal contributed to the CD part and Yanira Mendez Gomez to the synthesis part. This project was designed and supervised by Ludger Wessjohann, Bernhard Westermann, Jürgen Schmidt, and Norbert Arnold.

Chapter 4. Laub, A.; Sendatzki, A.-K.; Palfner, G.; Wessjohann, L.A.; Schmidt, J.; Arnold, N.; HPTLC-DESI-HRMS-Based Profiling of Anthraquinones in Complex Mixtures — A Proof-of-Concept Study Using Crude Extracts of Chilean Mushrooms *Foods* **2020**, *9*, 156.

In this project Annegret Laub performed the method development and measurements via HPTLC-DESI-HRMS as well as the data evaluation. A.-K. Sendatzki supported the work with the preparation of the fungal extracts and the development of the HPTLC plates under supervision of Annegret Laub, Norbert Arnold, and Ludger A. Wessjohann. Götz Palfner provided a part of the fungal material. Annegret Laub wrote the manuscript. This project was designed, supervised, and edited by Ludger Wessjohann, Jürgen Schmidt, and Norbert Arnold.

Publications

Publications in peer-reviewed journals

Adjedje, V., Schell, E., Wolf, Y., **Laub, A.**, Weissenborn, M., Binder, W. Enzymatic degradation of synthetic polyisoprenes via surfactant-free polymer emulsification, *Green Chemistry*, **2021**, *23*, 9433-9438.

Corrêa dos Santos, C.H., Geraldo de Carvalho, M., **Laub, A.**, Franke, K., Wessjohann, L.A., UHPLC-ESI-Orbitrap-HRMS analysis of cyclopeptide alkaloids from *Ziziphus joazeiro*., *Natural Product Communications*, **2021**, *16* (11), DOI: 10.1177/1934578X211054955.

Nugraha, A.S.; Untari, L.F.; **Laub, A.**; Porzel, A.; Franke, K.; Wessjohann, L.A. Anthelmintic and antimicrobial activities of three new depsides and ten known depsides and phenols from Indonesian lichen: *Parmelia cetrata* Ach. *Nat. Prod. Res.* **2020**, 1-10.

Laub, A.; Sendatzki, A.-K.; Palfner, G.; Wessjohann, L.A.; Schmidt, J.; Arnold, N. HPTLC-DESI-HRMS-Based Profiling of Anthraquinones in Complex Mixtures—A Proof-of-Concept Study Using Crude Extracts of Chilean Mushrooms. *Foods* **2020**, *9*, 156.

Schmidt, J.; Kuck, D.; Franke, K.; Sultani, H.; Laub, A.; Wessjohann, L.A. The unusual fragmentation of long-chain feruloyl esters under negative ion electrospray conditions. *J. Mass Spectrom.* **2019**, *54*, 549-556.

Coors, A., Brosch, M., Kahl, E., Khalil, R., Michels, B., **Laub, A.**, Franke, K., Gerber, B., Fendt, M. *Rhodiola rosea* root extract has antipsychotic-like effects in rodent models of sensorimotor gating, *J. Ethnopharmacol.* **2019**, *235*, 320-328.

Farag, M. A., Ali, S. A., Hodaya, R.H., El-Seedi, H. R., Sultani, H. N., **Laub, A.**, Eissa, T.F., Abou-Zaid, F. O. F., Wessjohann, L. A. Phytochemical profiles and antimicrobial activities of *Allium cepa* Red cv. and *A. sativum* subjected to different drying methods: A comparative MS-based metabolomics. *Molecules* **2017**, *22*, 761.

Morejón, M. C.; **Laub, A.**; Kaluđerović, G.; Puentes, A. R.; Hmedat, A. N.; Otero-González, A. J.; Rivera, D. G.; Wessjohann, L. A. A Multicomponent Macrocyclization Strategy to Natural Product-Like Cyclic Lipopeptides: Synthesis and Anticancer Evaluation of Surfactin and Mycosubtilin Analogues. *Org. Biomol. Chem.* **2017**, *15*, 3628–3637.

Morejón, M. C.; **Laub, A.**; Westermann, B.; Rivera, D. G.; Wessjohann, L. A. Solution and Solid-Phase Macrocyclization of Peptides by the Ugi-Smiles Multicomponent Reaction: Synthesis of *N*-Aryl-Bridged Cyclic Lipopeptides. *Org. Lett.* **2016**, *18*, 4096–4099.

Otto, A., **Laub, A.**, Haid, M., Porzel, A., Schmidt, J., Wessjohann, L., Arnold, N., Tulasporins A–D, 19-Residue Peptaibols from the Mycoparasitic Fungus *Sepedonium tulasneanum*. *Nat. Prod. Commun.* **2016**, *11*, 1821–1824

Otto, A., **Laub, A.**, Wendt, L., Porzel, A., Schmidt, J., Palfner, G., Becerra, J., Krüger, D., Stadler, M., Wessjohann, L., Westermann, B., Arnold, N. Chilenopeptins A and B, Peptaibols from the Chilean *Sepedonium* aff. *chalcipori* KSH 883. *J. Nat. Prod.* **2016**, *79*, 928–938.

Otto, A., **Laub, A.**, Porzel, A., Schmidt, J., Wessjohann, L., Westermann, B., Arnold, N., 2015. Isolation and Total Synthesis of Albupeptins A–D: 11-Residue Peptaibols from the Fungus *Gliocladium album. Eur. J. Org. Chem.* **2015**, 7449–7459.

Oral presentations

Matousek, J., Kroftka, K., **Laub, A.**, Kocabek, T., Patzak, J., Killi, U.K., Mishra, A.K., Hagemann, M.H., Svoboda, P., Wünsche, J.N., Wessjohann, L., Koncz, C. Some metabolome changes in hop glandular trichomes due to transgenes *HI*WRKY*IHI*WDR1, MBW, gene tissue-specific expression and light. International Hop Growers´ Convention I.H.G.C., 2019, Bischoffsheim, Alsace, France.

Laub, A. Metabolic profiling of bioactive plants from Indonesia. Begrüßungstreffen der Stipendiaten des DAAD-BMBF-Programms "Biodiversität & Gesundheit", 2018, Bonn, Germany.

Laub, A., Otto, A., Porzel, A., Schmidt, J., Palfner, G., Becerra, J., Wessjohann, L., Westermann, B., Arnold, N. Antifungal peptaibols from *Sepedonium* species. 13th Plant Science Student Conference (PSSC), 2017, Halle (Saale), Germany.

Laub, A., Otto, A., Wendt, L., Porzel, A., Schmidt, J., Palfner, G., Becerra, J., Krüger, D., Stadler, M., Wessjohann, L., Westermann, B., Arnold, N. Peptide-like compounds from fungi – a polythetic approach. 53. Doktorandenworkshop "Naturstoffe: Chemie, Biologie und Ökologie", 2017, München, Germany.

Laub, A. Antifungal peptaibols from *Sepedonium* species. Postgraduate workshop of the section „Natural Products" of the German Society for Plant Sciences (DBG), 2016, Meisdorf, Germany.

Poster presentations

Laub, A., Sendatzki, A.-K., Palfner, G., Arnold, N., Wessjohann, L.A., Schmidt, J. Screening for anthraquinones in Chilean Mushrooms by HPTLC-negative ion high-resolution mass spectrometry. At 12th International Mass Spectrometry Conference (IMSC), 2018, Florence, Italy.

Laub, A., Otto, A., Arnold, N., Schmidt, J., Westermann, B. Screening for anthraquinones in Chilean mushrooms by HPTLC-negative ion DESI high-resolution mass spectrometry. 49th meeting of the German society of mass spectrometry (DGMS), 2016, Hamburg, Germany.

Laub, A., Otto, A., Roth, P., Arnold, N., Schmidt, J., Westermann, B. Total synthesis and biological evaluation of new peptaibols and their derivatives. 11th Plant Science Student Conference (PSSC), 2015, Halle (Saale), Germany.

Laub, A., Otto, A., Arnold, N., Schmidt, J., Westermann, B. Synthesis and mass spectrometric characterization of aminoisobutyric acid containing, antibiotic peptides. 48th meeting of the German society of mass spectrometry (DGMS), 2015, Wuppertal, Germany.

<u>Awards</u>

Poster award from Springer – Analytical and Bioanalytical Chemistry (ABC), 49th meeting of the German society of mass spectrometry (DGMS), 2016, Hamburg, Germany.

Curriculum vitae

Personal details

Name: Annegret Laub

Date/place of birth: 9th November 1990 in Querfurt

Nationality: German

School education

09/2001 – 06/2009 Staatliches Gymnasium „Klosterschule" Roßleben

Degree: General qualification for University entrance

Academic studies

10.2009 – 04.2014 Diploma Studies Food Chemistry – Faculty of Chemistry, Martin-Luther-University Halle-Wittenberg

11.2013 – 04.2014 Diploma Thesis

Title: Synthesis and biological evaluation of aminoisobutyric acid containing peptides

Supervisor: Prof. Dr. Bernhard Westermann

01. 2015 – Present PhD Candidate (Dr. rer. nat.) – Leibniz Institute of Plant Biochemistry, Department of Bioorganic Chemistry, IPB, Halle (Saale), Germany

Title: Chemical fingerprints of natural origin: Analyses of secondary metabolite profiles from plants, micro- and macrofungi

Supervisor: Prof. Dr. Ludger A. Wessjohann

Work experience

05.2014 – 07.2014 Research Assistant – Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry, IPB, Halle (Saale), Germany

08.2014 – 12.2014 Research Assistant – Department of Stress and Developmental Biology, Leibniz Institute of Plant Biochemistry, IPB, Halle (Saale), Germany

01.2015 – 12.2020 Research Assistant – Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry, IPB, Halle (Saale), Germany

01.2021 – present Responsible Senior Scientist for Mass Spectrometry, Leibniz Institute of Plant Biochemistry, IPB, Halle (Saale), Germany

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Hilfsmittel und Quellen angefertigt habe.

Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen sind, habe ich als solche kenntlich gemacht. Des Weiteren versichere ich, dass ich diese Arbeit an keiner anderen Institution eingereicht habe.

Mir ist bekannt, dass bei Angabe falscher Aussagen die Prüfung als nicht bestanden gilt. Mit meiner Unterschrift versichere ich die Richtigkeit dieser Angaben und erkenne die rechtlichen Grundlagen an.

Halle,

Datum	Annegret Laub