UiT The Arctic University of Norway

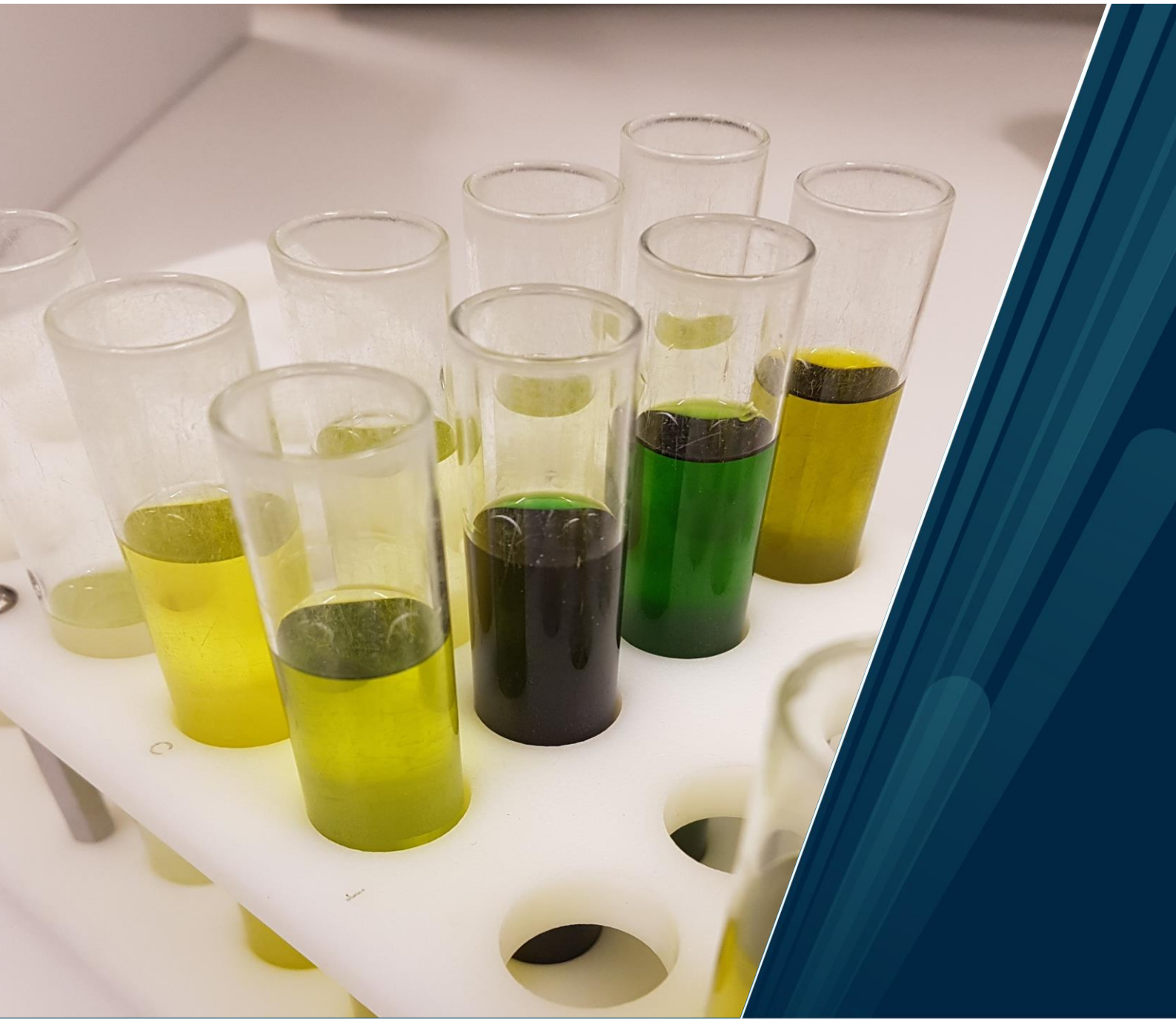
Faculty of Biosciences, Fisheries and Economics

Department of Arctic and Marine Biology

Metabolomic and genomic investigation of two North-Norwegian cyanobacterial isolates for bioprospecting of new compounds

Ane Norum Kirkesæther

Master's thesis in Biology BIO-3950 February 2022



Acknowledgement

I would like to thank my supervisors Anton Liaimer and Yannik K. Schneider for the provided guidance during the laboratory work and writing process. They have provided me with encouraging words and invaluable learning of which I am greatly thankful for and will never forget. I would like to especially acknowledge Yannik for the much-appreciated support for keeping me on track and immense patience for all my questions.

Also, a big thanks to the others at Marbio for a friendly environment and for always being happy to assist me and my fellow students. Lastly, a heartfelt appreciation for my fellow students at Marbio for creating a warm fellowship which has been greatly treasured.

Abstract

Cyanobacteria are an excellent source of bioactive natural products that can be used in the development of new medicinal drugs. The cyanobacterial genus *Nostoc* have proven to be prolific producers of molecules with exciting bioactivities, including anti-bacterial, anti-fungal, and anti-cancerous. This feature combined with the *Nostoc*'s complex life cycle and sizeable genomes make them interesting targets for bioprospecting.

During this thesis, a typical natural product discovery pipeline was carried out to investigate some of the genomic and metabolomic characteristics of two *Nostoc* sp. strains. The pipeline included prediction of the biosynthetic potential, liquid extraction of metabolites using a variation of solvents, reverse-phase fractionation, bioassay to test for bioactivity, and the use of reverse-phase UHPLC-MS^2 and an informatic tool for dereplication and molecular networking.

Some clear differences between the strains were observed in the prediction of metabolite production and in the extracted metabolites. Anti-bacterial activity was observed in extract fractions of both strains, and the dereplication process resulted in the discovery of a congener of the earlier described compound hapalosin which has been shown to reverse multi-drug resistance in cancer cells.

Table of Contents:

List of Tables

List of Figures

Abbreviations:

pH_2O	Purified water
°C	Degrees Celsius
µ	Micro
B.C	Before Christ
BGC	Biosynthetic gene cluster
BLAST	Basic Local Alignment Search Tool
C	Carbon
CID	Collision-induced dissociation
Da	Daltons
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
EI	Electron impact
ESI	Electrospray ionization
et al.	*et alii*
etc.	*et cetera*
EtOAc	Ethyl acetate
eV	Electron volt
FDA	Food and Drug Administration
g	Gram
GNPS	G
h	Hour
H	Hydrogen
i.e.	*id est*
kV	Kilovolt
KVJ10	*Nostoc* sp. KVJ10
KVJ20	*Nostoc* sp. KVJ20
L	Liter
LC	Liquid chromatography
LC	Liquid chromatography
m/*z*	Mass-to-charge ratio
MeOH	Methanol

mg	Milligram
mL	Milliliter
MS	Mass spectrometry
MS^2	Tandem mass spectrometry
N	Nitrogen
NCBI	National Center for Biotechnology Information
Ncp	Nostocyclopeptide
NMR	Nuclear magnetic resonance
NP	Natural product
NRPS	Non-ribosomal peptide synthase
OD	Optical density
PKS	Polyketide synthase
PRPS	Post-ribosomal peptide synthesis
RiPP	Ribosomally synthesized and post-translationally modified peptide
RNA	Ribonucleic acid
Rpm	Rounds per minute
RT	Retention time
s	Second(s)
sp.	Species
UHPLC	Ultra-high performance liquid chromatography
UV	Ultraviolet
Vis	Visible
w/	with

1 Introduction

1.1 Natural Products

The use of natural products (NPs) for medicinal purposes has followed the course of human history with records dating back as early as 2600 B.C. to ancient Mesopotamia, largely using plant-based remedies to treat about any form of ailment and disease. The path towards natural product discovery and natural-product-derived medicine development in the 21st century involved several pinnacles with the isolation of salicin (aspirin), morphine, and quinine, to name a few (Chin et al., 2006; Dias et al., 2012). Bioactive NPs continue to be a crucial source for pharmaceutical drug discovery up to this day (Dixit et al., 2010; Singh et al., 2005). To search for new sources of NPs with therapeutical properties is more relevant than ever with the ongoing rise of antibiotic resistance (Hudson & Mitchell, 2018; Ventola, 2015), an issue which has been deemed one of the major health threats against present humanity by the World Health Organization (WHO, 2020). Not to mention the rise of novel and highly contagious pathogens of which there is, so far, no known effective treatment (Carpine & Sieber, 2021; Wang & Yang, 2020).

Bioactive NPs can serve directly as drugs or as leads for the synthesis of new therapeutic compounds (Dixit et al., 2010; Singh et al., 2005). Bioactive NPs have contributed significantly to the development of novel antibiotics and other important drugs, for instance anticancer, antiparasitic and immunosuppressant agents. Of the 90 antibiotic drugs that became commercially available in the United States or were approved worldwide (between 1982 to 2002), 70 of these originated from natural products (Chin et al., 2006).

Microscopic fungi and bacteria became an important source for uncovering new therapeutic NPs after the work of Alexander Fleming, Howard W. Florey and Ernst B. Chain and the introduction of the first antibiotic drug, penicillin, in the 1940s, a secondary metabolite from the mould *Penicillium rubens* (*e.g.*, Abdel-Razek et al., 2020; Chin et al., 2006; Singh et al., 2011).

1.1.1 Secondary metabolites

NPs (natural products) can be defined as any naturally produced product of the metabolism from any organism and can be divided into primary and secondary metabolites. Primary metabolites are essential to organisms among all domains of life for survival and reproduction and include molecules such as carbohydrates, proteins, fatty acids, nucleic acids (DNA, RNA), *etc* (Dewick, 2001).

Secondary metabolites, on the contrary, are not directly required for growth, development, or reproduction and consist of exceptionally diverse chemical classes that differs in size, structure, and complexity, including terpenes, polyketides, peptides, alkaloids, and glycosides, to name a few. The biosynthetic pathways of primary metabolites are often strongly conserved and broadly shared across phylogenies while secondary metabolites are often distinct to a subset of organisms or a particular environmental niche (Dewick, 2009; Dias et al., 2012; Dittman et al., 2015; Kultschar & Llewelyn, 2018; Medema et al., 2021).

It is undoubtedly the secondary metabolites which have yielded the most pharmacologically active natural products (Chin et al., 2006; Dewick, 2009), so much in fact that natural products are in many cases synonymous to secondary metabolites (*e.g.*, Covington et al., 2021, Dias et al., 2012; Katz & Baltz, 2016; Maplestone et al., 1992).

Microorganisms express secondary metabolites to advance in their environment by adapting to environmental changes and increasing nutrient acquisition. Secondary metabolites also mediate the interaction with other microbes or organisms, for example in endophytic relationships with plants (Strobel et al., 2004), when communicating with fellow microbes or to wage war against competing microbial species (Covington et al. 2021; Maplestone et al., 1992).

Bioactive metabolites from both fungi and bacteria have for half a century been a field of research with great enthusiasm. And with good reason as many have been shown to possess not only antibacterial but also antiviral, anticancer, and anti-inflammatory (and other) properties (*e.g.*, Abdel-Razek et al., 2020; Bredholdt et al., 2008; Chin et al., 2006). Several such compounds and their derivatives have been approved by the Food and Drug Administration (FDA) in the US as drugs, *e.g.*, Paclitaxel derived from taxol (antitumor and antifungal), Fucidin from fusidic acid (bacterial infections), pimecrolimus (anti-inflammatory), and amrubicin (antitumor) (Abdel-Razek et al., 2020). Microbes have been an important source of secondary metabolites that have innovated many of the clinically used

antibiotics, including beta-lactams, tetracyclines, macrolides, and glycopeptides (Hudson & Mitchell, 2018)

Earlier screening efforts were intensified on a few well-known fungal and bacterial groups, for instance, actinobacteria which are responsible for about 45% of described microbe-derived bioactive compounds (Abdel-Razek et al., 2020; Purves et al., 2016). However, with genomic sequencing and bioinformatic tools it opened the possibility to predict and investigate the biosynthesis capacity of many less investigated taxa and revealed that several microbial phyla, including cyanobacteria, inherent a much greater biosynthetic potential than earlier though (Duncan et al., 2015; Purves et al., 2016; Ueoka et al., 2018).

1.1.2 Biosynthesis of natural products

In bacteria, natural product synthesis genes are often arranged in biosynthetic gene clusters (BGCs) on the chromosome. The physical nature of especially bacterial BGCs makes them suitable for genome mining for finding new biosynthetic pathways in the search for new secondary metabolites (Medema et al., 2021).

A big share of bioactive NPs from microbes are produced by non-ribosomal peptide synthases (NRPSs), polyketide synthases (PKSs), hybrid NRPS/PKS pathways (Adrover-Castellano et al., 2021; Kehr et al., 2011; Schwarzer & Marahiel, 2001), or by post-ribosomal peptide synthesis (PRPS) (Demay et al., 2019). The resulting products are individually classified as non-ribosomal peptides (NRPs), polyketides (PKs), NRP/PK hybrid molecules, or ribosomally synthesized and post-translationally modified peptides (RiPPs).
Generally during PRPS, a precursor peptide is ribosomally produced which will undergo several modifications before being transported and released from the biosynthetic machinery (Arnison et al., 2013; Hudson & Mitchell, 2018; Ortega & van der Donk, 2016).
NRPs and PKSs are multi-domain megaenzymes that consist of modules which contain distinct catalytic domains that control the biosynthesis initiation, chain extension, processing, and termination during the synthesis process. This process is like an assembly line where after initiation, the following modules adds on building blocks one at a time on the substrate compound and chemically modifies it along the way until the product reaches the releasing module. In the releasing module, the product is unhooked from the assembly line before it undergoes some final processing ending in a complete bioactive compound (Adrover-Castellano et al., 2021; Chang et al., 2002; Schwarzer & Marahiel, 2001; Weissman, 2015).

The resulting peptide or polyketide natural products from all these pathways are highly structurally diverse and can be linear, cyclic and/or branched in structure (Wang et al., 2014, Zhang et al., 2014). Example compounds are shown in table 1.1.

Table 1.1. Example of NRPS, PKS, PRPS, and hybrid NRPS/PKS products, their bioactivity and the producing organism.

Product	Bioactivity	Biosynthetic pathway	Producing organism	Ref.
Nisin	Anti-bacterial	PRPS	*Lactococcus lactis*	Cao et al., 2021
Penicillin	Anti-bacterial	NRPS	*Penicillium chrysogenum*	Du et al., 2001
Curacin	Cytotoxic	PKS	*Gloeobacter violaceus*	Gademann et al., 2007
Microcystin	Hepatotoxic	Hybrid NRPS/PKS	*Microcystis aeruginosa*	Kaebernick et al., 2000

1.1.2.1 Non-ribosomal peptide synthases

The non-ribosomal peptide synthases (NRPSs) use amino acids to build peptides through amide bonds (Miyanaga et al., 2018). A microbial NRPS module generally consist of adenylation (A), carrier protein (CP/T or peptidyl carrier protein (CPC), condensation (C), and thioesterase (TE) domain (Du et al., 2000; Weissman, 2015).

Typically, initiation begins with ATP-dependent activation of a specific amino acid substrate by the A domains into an aminoacyl adenylate and is transferred to PCP domains attached as thioesters. Further, the growing peptide chain might be modified by for example N-methyltransferase (N-MT) domains before or after reaching the C domains which couples the two aminoacyl-PCPs together by amide bond formation. After chain extension, optional structural modifications can be made to the amino acids such as configuration of a chiral center (epimerization) by E domains, formation of heterocyclic rings (heterocyclization) by HC domains, and oxidation by Ox domains. This reaction sequence is repeated until the polypeptide reaches the TE domain which catalyzes the release of the product (Adrover-Castellano et al., Dittman et al., 2015; Du et al., 2000; 2021; Weissman, 2015).

1.1.2.2 Polyketide synthases

The polyketide synthases (PKSs) construct polyketide chains from C-C bonding of simple acyl-CoA building blocks (Miyanaga et al., 2018; Weissman, 2015). PKSs are classified into three different groups based on their catalytic domain organization. Each module in type I PKSs consist of domains that recognize, activate and condensate acyl-CoA, while the catalytic domains of type II and type III PKSs are organized into separate proteins (Wang et al., 2014).

Generally, acyl transferase (AT) domains initiate the synthesis by recruitment of an appropriate acyl-CoA substrate and transfer it to an acyl carrier protein (ACP) of the same module. In the extension module, the starter unit is acylated by the module's ketosynthase, condensing the starter unit with the extender unit attached to the resident ACP of the extension module. After extension, the β-keto group on the chain can be modified by ketoreductase, dehydratase, and enoyl reductase domains. The cycle of chain growth and structural modifications are repeated until the complete product is released from the megaenzyme by a TE domain, similarly as in NRPSs (Adrover-Castellano et al., 2021; Miyanaga et al., 2018; Weissman, 2015).

1.1.2.3 Hybrid pathways

A high proportion of microbial natural products are mixed polyketide and non-ribosomal peptide compounds (Chang et al., 2002). These complex polyketide/peptide hybrid molecules are constructed through PKS-NRPS/NRPS-PKS hybrid pathways that incorporate both acyl and aminoacyl substrates into their products. This expands their biological activities through combined chemical properties (Miyanaga et al., 2018).

1.1.2.4 Post-ribosomal peptide synthesis

RiPPs are synthesized by unique biosynthetic pathways where a precursor peptide is ribosomally produced which consist of an N-terminal leader region that contain a recognition sequence and a C-terminal core region that have several sites for post-translational modifications. The N-terminal is recognized by a particular biosynthetic protein which inaugurates the many modifications in the C-terminal. The modifications can include processes such as methylation, cyclization, isomerization, and dehydration, which are

mediated by enzymes (Hudson & Mitchell, 2018). The N-terminal is then removed by peptidases and the mature RiPP is complete. Because of the diversity of the post-translational modifications, the resulting peptides have extensive structural diversity and have a broad variety of biological activities. Some examples of such RiPPs are lasso peptides, cyanobactins, and lanthipeptides (Arnison et al., 2013; Ortega & van der Donk, 2016; Zhang et al., 2014).

1.2 Strategies for Microbial Natural Product Discovery

The strategies for natural product discovery have been thoroughly reviewed and revised (*e.g.*, Carrano & Marinelli, 2015; Ito & Masubuchi, 2014) throughout the last 80 years. Especially with the standardization of genome sequencing as a readily available tool which gave rise to a set of new strategies and techniques (Katz & Baltz, 2016). And so, the workflow of natural product discovery can generally be divided into two principles: top-down and bottom-up approaches (Schneider, 2020).

1.2.1 Top-down approaches

The traditional route for microbial natural product discovery starts with the cultivation of an obtained and isolated microbial strain or biomass that undergo product isolation by various extraction methods. After extraction, the metabolites in the crude extract can be isolated further by fractionation. The fractions or the crude extracts can be investigated through chemical analyses, like mass spectrometry, and/or bioactivity tests (*e.g.*, target-based or phenotypic screening) (Katz & Baltz, 2016). Such analyses aim to discover new bioactive molecules or new bioactivities for a known compound that have not been described earlier. The identified compound is subsequently purified, and its structure elucidated, usually through mass spectrometry and nuclear magnetic resonance (Steckel & Schlosser, 2019). An example of a top-down pipeline in natural product discovery is illustrated in figure 1.1. Before cultivation the microbial strain is collected from an environment and isolated as a pure culture, or a complex sample consisting of several microbes may also be cultivated and extracted followed by analysis (Carrano & Marinelli, 2015). The microbial strain(s) can also be cultivated under several environmental conditions and medium compositions to provoke different phenotypic expression in metabolites produced. This can include adding or removing certain nutrients or co-cultivation with another microbe (Bladt et al., 2013).

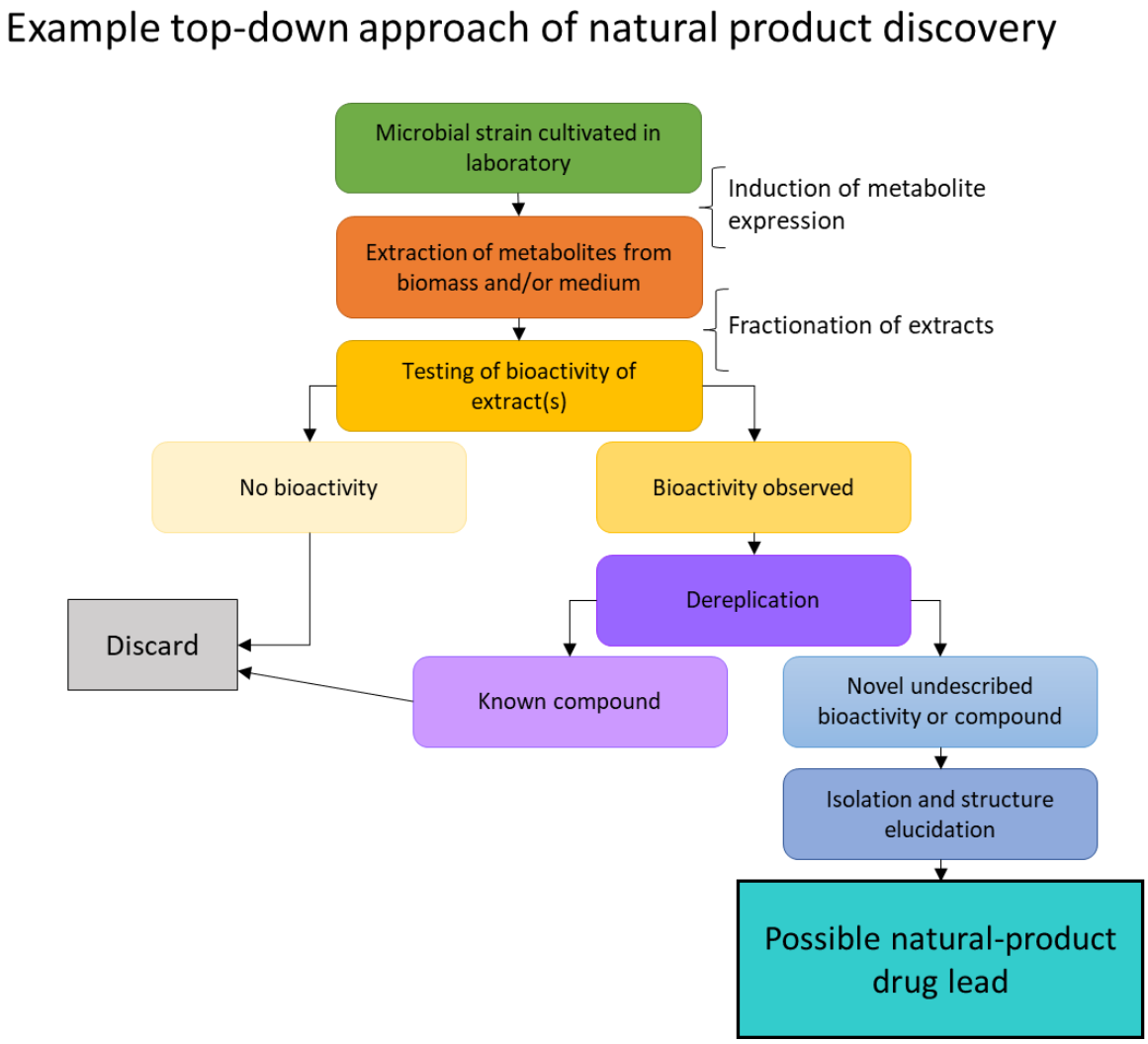


Figure 1.1. Illustration of an example pipeline of a top-down approach of natural product discovery. With inspiration of Schneider, 2020.

The screening for new natural products from complex microbial samples can be challenging and time-consuming work. There is a high probability for the re-discovery of molecules that have already been explored and characterized which could make even the most superlative researcher feel like valuable time and resources have been wasted on backing the wrong horse. This is one of the reasons why a number of big companies in the pharmaceutical industry have previously ceased natural-product discovery efforts (Carrano & Marinelli, 2015; Koehn 2008; Yang et al., 2013).

1.2.2 Bottom-up approaches

Bottom-up approaches starts with gene sequencing and subsequent analysis the sequence data by using various bioinformatic tools and/or homologous or heterologous expression screens (Schneider, 2020). The main purpose is to uncover the biosynthetic potential of a microorganism as a producer of novel natural products through genome mining which can uncover silent biosynthetic gene clusters that under standard cultivation conditions would not be expressed, or to explore the biosynthetic potential of novel or unexplored microbial strains (Medema et al., 2020; Katz & Baltz, 2016; Ueoka et al., 2018). In addition, by using gene and sequence information early in the bioprospecting process, the rate of re-discovery can be reduced by uncovering the biosynthetic potential of an organism before going through with cultivation and extraction. Thus, if a given microbial strain shows a lack of novel biosynthetic pathways they be disregarded early in the process as being likely no producers of novel secondary metabolites. It should be noted that bottom-up approaches are still dependent on chemical analyses to predict structural details of novel molecules and that these two approaches are only in theory based on different principles, therefore during the actual natural product discovery process, methods from both strategies are often complimented with each other to acquire a holistic understanding of the product, including mode of action and biosynthetic pathway (Katz & Baltz, 2016).

1.2.3 Dereplication

Dereplication is a vital part in the natural product discovery process. The goal is to identify known compounds or unspecific activity in natural product extracts to distinguish them from novel metabolites that are of interest and have not yet been analyzed. Dereplication can take place at various stages of the discovery process and there are several technologies in the dereplication toolbox (Carrano & Marinelli, 2015; Ito & Masubuchi, 2014). The most used tool in dereplication during natural product discovery is mass spectrometry (MS), usually coupled with liquid-chromatography (LC), and nuclear magnetic resonance (NMR). Although, LC-MS can be used to identify some compounds based on MS data and retention time, new compounds are often further elucidated through NMR to obtain specific molecular structure (Carrano & Marinelli, 2015; Ito & Masubuchi, 2014). NMR provides information on the number and chemical surrounding of hydrogens and carbons that are present in the analyzed molecule. This allows the analyst to conclude structural properties and ideally to elucidate a molecules chemical structure.

1.2.4 Chromatography

Chromatography refers to a collection of separation techniques which are used to separate different compounds in a sample. In liquid chromatography (LC), fluids are used to transport the sample (mobile phase) through solid particles which stays stationary (stationary phase) within a column. The liquid is transported through the column at a constant speed, but the stationary phase can slow down the speed of which the molecules of the sample transported by the mobile phase travels. Because of this and because the molecules in the sample have different characteristics, they will be transported through the stationary phase at different speeds, meaning that the different molecules will exit the column at different times and be separated. The time it takes from a sample is injected into the chromatograph system to a molecule in sample is detected is referred to as *retention time*. Retention time is a measurement of how long it takes a molecule to travel through the stationary phase. The retention time can be affected by various factors, including the speed of the mobile phase and the length of the column. Various chromatography techniques use different columns and solvents for the mobile and stationary phases (Pedersen-Bjergaard & Rasmussen, 2004). The type of metabolites that are detected through chromatography is strongly affected by, firstly, the preliminary method used for extraction of metabolites and secondly, the solvent gradient and separation technique used (Patti, 2011). The method of choice for metabolite extraction depends on the type of organism (*e.*g., plant, microbe, or animal) or material (medium, tissue, dry biomass, or leaves, *etc*.) the metabolites are extracted from, in addition to the nature (*e.g.*, polarity, bioactivity, or structure), if known, of the target compounds (Sarker et al., 2005).

1.2.4.1 High-performance liquid chromatography (HPLC)

LC (liquid chromatography) is commonly used in dereplication and characterization of crude mixtures or fractions from various natural sources during natural product discovery (Carrano & Marinelli, 2014; Patti, 2011, Sarker et al., 2005). The principle of separation used in HPLC is either normal-phase, reversed-phase, or size-exclusion chromatography. As the name suggests, HPLC uses liquids as mobile phase. The sample is injected into the mobile phase stream and is transported with the mobile phase through a column packed with a material that retard the compounds in a sample based on polarity or molecular size depending on the chromatographic separation method that is used. The molecules that are transported through the column, containing the stationary phase are detected with a detector that transfers the data

as electronic signals to a computer that processes and prints the signals as a chromatogram (Pedersen-Bjergaard & Rasmussen, 2004). As detector, ultraviolet/visible light (UV/Vis) detectors, such as photodiode array are most common, however a wide range of optical detectors such as refractive-index but also ESI-mass spectrometers or Geiger counters can be coupled to an HPLC. An advanced variation of HPLC that is called ultra-high performance liquid chromatography (UHPLC) is capable of handling much higher operating pressures, it is much quicker, separates analytes better, and is therefore more widely used than traditional HPLC (Rathod et al., 2019).

1.2.4.2 Fractionation by HPLC

HPLC can also be used as a means for fractioning a crude extract or mixture of compounds. Crude extracts contain a myriad of different compounds and it therefore beneficial to separate them into different fractions based on their characteristics, for example their mass, size, or polarity, to make it easier for subsequent isolation of pure compound or further analyses, for example bioassays (Sarker et al., 2005).

1.2.5 Mass spectrometry

MS (mass spectrometry) is done with an instrument called a *mass spectrometer*. In the mass spectrometer the introduced sample components are ionized into molecular ions that either lose an electron (positive ions) or gain an electron (negative ions). The mass of the molecular ions [M] corresponds to the mass of the original molecules in the sample because the weight of a gained or lost electron is negligible. Molecular ions are stable or unstable. Unstable molecular ions rapidly split into smaller fragments by the breakage of chemical bonds. This process is called fragmentation. Some of the fragments will keep being ionized and are referred to as fragment ions, while other will lose their ion charge. The ionization and the subsequent fragmentation take place either in an ionization source inside the spectrometer or directly outside the instrument. The mass of the molecular ion and fragments that are still charged are decided by a mass analyzer within the instrument that the ions are accelerated into. Inside the mass analyzer, the ions are separated according to their mass (m) to charge (z) ratio (*m*/*z*), either in a magnetic or electrostatic field. The ratio between mass and charge in abbreviated as *m*/*z*. Ions with low mass relative to charge will in the field of the mass analyzer have paths that are greatly curved while ions that have high mass relative to charge will have less curvation. In many cases the ion charge will be +/- 1 and so the curvation will only

depend on the mass of the ions, if not, isotopes can give cues about the charge since isotopes (such as C13) need to have mass different of full units. After the mass analyzer the ions are detected by a detector that register the number of ions that are formed and their *m*/*z*. The data from a mass spectrometry analysis is registered in a mass spectrum, where the intensity of the ions is plotted as a function of their mass-to-charge ratio (Pedersen-Bjergaard & Rasmussen, 2004).

MS can be used to identify organic compounds because different compounds have different mass spectra. The mass spectrum of a known molecule can thus be used to identify an unknown molecule if their mass spectra match. If this is not the case, mass spectrum of unknown compounds can still give valuable information about the structure of the compound. MS provides, with high accuracy and sensitivity, information about molecular characteristics *i.e.*, mass, formula, and fragmentation pattern (Bouslimani et al., 2014; Pedersen-Bjergaard & Rasmussen, 2004; Sarker et al., 2005; Steckel & Schlosser, 2019), which can be compared to compound libraries or literature for identification and characterization of an analyte.

1.2.5.1 Electrospray ionization mass spectrometry (ESI-MS)

Electrospray ionization (ESI)-MS have shown to be a suitable method for the analysis of secondary metabolites and is commonly used for this purpose (Ito & Masubuchi, 2014; Sarker et al., 2005). ESI is a so called “soft” ionization technique that causes minimal fragmentation and can accurately determine accurate mass and give detailed structures of fragile, polar, and higher molecular weight compounds than the so called “hard” ionization MS methods, such as electron impact (EI) ionization which are out of the scope of this work (Forcisi et al., 2013). However, fragmentation patterns are helpful to conclude structural features or the identify of an analyte and thus, collision-induced dissociation (CID) MS^2 is often used together with ESI-MS to generate fragmentation in ESI-mass spectrometry (Sarker et al., 2005; Steckel & Schlosser, 2019). ESI can be used in positive (ESI+) and negative ionization mode (ESI-) where typically, the former will easily ionize molecules containing basic functional groups, while the latter provides better spectra of molecules that have acidic functional groups and lacking basic ones (Steckel & Schlosser, 2019). MS coupled with separation techniques like liquid chromatography (LC) allows for additional separation of compounds from complex samples (Bouslimani et al., 2014).

1.2.6 HPLC-MS

MS (mass spectrometry) is often combined with LC (liquid chromatography) and provides a powerful tool for NP discovery from natural sources (Forcisi et al., 2013;). HPLC-MS combines the abilities of chemical separation of LC with the selective detection and molecular characterization of MS (Sarker et al., 2005). The compounds in the sample are first separated through a chromatography column before going separately into the spectrometer where they are ionized. In this way, the spectrometer functions as an advanced detector throughout the separation process (Pedersen-Bjergaard & Rasmussen, 2004).

1.3 The Cyanobacteria

Cyanobacteria is a phylum within the bacterial domain consisting of photosynthetic, Gram-negative prokaryotes. Cyanobacteria are distinguished from other photoautotroph bacteria by having phycobilins as photopigments and having both type I and type II photosystems. Algal and plant chloroplast are also present in the species of this phylum. These oxygenic photoautotrophs are estimated to appear on earth 2.7-3.5 billion years ago. They are crowned for playing a major role in the evolution of life as we know it by producing oxygen that is released in the atmosphere. They are also recognized for their contribution in the evolvement of plants and eukaryotic algae (Chlipala et al., 2011; Mandal & Rathm 2015) and they continue to function as major contributors to the global nitrogen and carbon cycles (Calteau et al., 2014). The cyanobacteria phylum consists of about 1579 species and 426 genera (Hauert & Komárek, 2022) among nine orders; Nostocales, Chroococcidiopsidales, Rubidibacter/Halothece, Spirulinales, Pleurocapsales, Chroococcales, Oscillatoriales, Synechhococcales, and Gleobacterales (Komárek et al., 2014).

1.3.1 Morphology

Cyanobacteria are unicellular microbes that come in a tremendous variety of morphologies. Based on species and environmental conditions they can occur as rods, spheres, or spirals, as single cells but more often as colonial aggregates or as filamentous colonies (Devillers et al., 2007; Singh et al., 2008). Different species have different ways of division including binary fission, multiple fission, and budding, and can form unbranched, pseudobranched or true branched filaments (Rippka et al., 1979). Many cyanobacteria show features of true multicellularity with cell-cell adhesion, intercellular communication, and cell differentiation

within filaments, as well as complex life cycles (Flores & Herrero, 2010).

1.3.2 Habitats

They can be found in a wide array of various marine, terrestrial, and freshwater ecosystems, ranging from the tropics to the arctic and even extreme environments (Chlipala et al., 2011; Fleischman, 2012; Singh et al., 2011). Heterocystous cyanobacteria can take residence in plants, either intracellularly in symbiotic organs or extracellularly in specialized cavities (Liaimer et al., 2016).

1.4 Natural Products from Cyanobacteria

Cyanobacteria are perhaps best known and previously studied for their cyanotoxins which can cause harm to both animals and humans during blooms in water reservoirs and drinking water supplies (Calteau et al., 2014; Carmichael et al., 2001). Despite this, cyanobacteria have as a source for new pharmaceutical drug leads been capturing the interest of scientists worldwide since the work of Richard E. Moore (1933-2007) in the 1970s to early 2000s. Moore provided findings that cyanobacteria contain a large number of bioactive compounds including fatty acid derivatives, saccharides and glycosides, peptides, polyketides, and lipopeptides (Burja et al., 2001; Cardellina & Moore, 2010; Dixit & Suseela, 2013; Tidgewell et al., 2010).

Cyanobacteria have throughout the years certainly proven to be an excellent source for the discovery of new bioactive compounds (*e.g.*, Burja et al., 2001; Jones et al., 2009; Schwartz et al., 1990; Singh et al., 2008, 2011) that can be used in pharmaceutic therapies (Dixit & Suseela, 2013; Mandal & Rath, 2015, Sainis et al., 2010). Their metabolites have shown to possess activities against microbes (Cardellina et al., 1979; Carpine & Sieber, 2021; Dixit et al., 2010; Pedersén & DaSilva, 1973), parasites (Xue et al., 2018), and cancer (Burja et al., 2001; Hamada & Shioiri 2005), and have enzyme inhibiting properties (Chlipala et al., 2011; Demay et al., 2019; Issa, 1999; Singh et al., 2011; Tidgewell et al., 2010). These cyanobacterial natural products belong to a wide range of chemical classes but are often produced through the various beforementioned biosynthetic pathways and have peptide and polyketide structural elements (Chlipala et al., 2011; Demay et al., 2019; Niedermeyer, 2015; Tidgewell et al., 2010).

In a review study, Demay and colleagues (2019) grouped 1630 unique molecules in 260 compound families which were further classified into 10 different chemical classes. They found that 51% of the compound families belonged to or derived from peptides (depsipeptides and lipopeptides). In addition, the authors reported that the most common bioactivity of these compound families were cytotoxicity, followed by lethality (toxicity) and antibacterial activity (see figure 1.2). However, as noted by Demay, the bioactivity profiles of the several of the compounds could be underestimated because of a skewedness in types of tests used among the reviewed studies and that some classes had few studies compared to others. Nonetheless, there did not seem to be any indication that one chemical class having specific bioactivity compared to other classes.

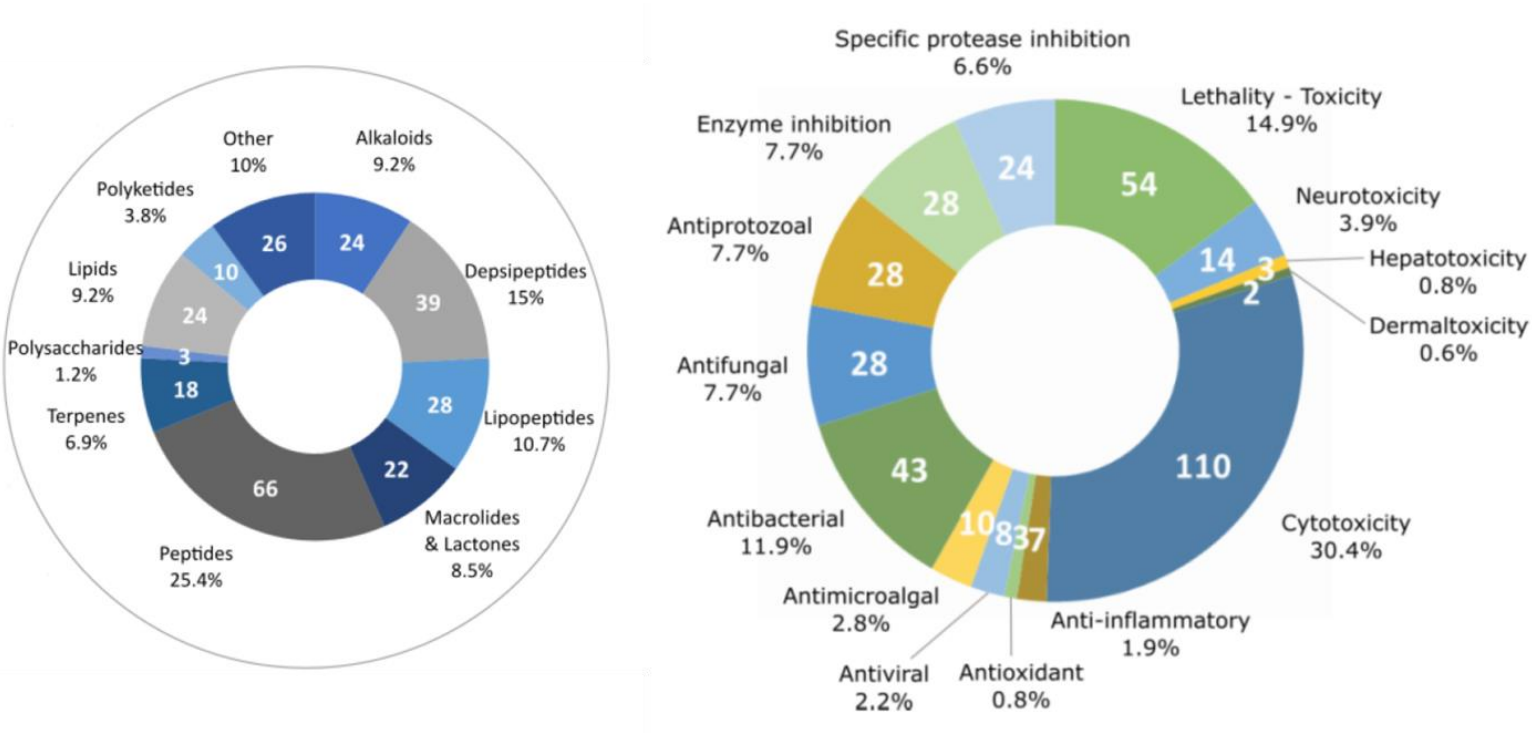


Figure 1.2. Classification of the 20 cyanobacterial metabolite families based on their respective chemical classes and bioactivities. Reprinted and adapted with permission from Demay et al., 2019.

The large variation of secondary metabolite production with great structural heterogeneity and powerful bioactivities seems to be partly due to their ability to adapt to tremendously varied and extreme environmental conditions (Chlipala et al., 2011; Kanekiyo et al., 2005; Kultschar & Llewelyn, 2018; Mandal & Rath, 2015; Rojas et al., 2020), including living as endophytes (Mazard et al., 2016; Meeks, 1998). But also because of intra- and inter-species interactions and the ability to deterring grazers (Leão et al., 2012; Mazard et al., 2016; Niedermeyer, 2015).

1.4.1 Cyanobacteria as producers of anti-microbial compounds

Anti-microbial compounds are molecules that have an inhibitory or deadly effect on microbes and are often naturally produced by many microbes. These natural product anti-microbials are thus of great interest in medicine as novel drugs or drug leads for the treatment of pathogens that have detrimental effects on human and animal health (Mullis et al., 2019; Rojas et al., 2020).

The compounds from cyanobacteria that have antimicrobial properties are often alkaloids, peptides or polyketides which are produced by NRPS, hybrid PKS-NRPS, or PRPS (Mullis et al., 2019). Other cyanobacterial compound classes that have shown antibacterial activities are terpenes, lipids, and polyphenols (Rojas et al., 2020; Swain et al., 2017; Xue et al., 2018) and more. The antimicrobial effect may not be the only or primary function of the metabolite and may serve as means of communication and inter- and intraspecific cooperation for the producing microbe (Mullis et al., 2019). Moreover, complex lifecycles and cellular organization, a trait of cyanobacteria, is often mediated by large genomes which correlates with the biosynthetic potential of a bacteria (Donadio et al., 2007; Nagle & Paul, 1999).

Table 1.2. Example of cyano-compounds that are antimicrobial.

Compound	Compound family	Producer	Target	Reference
Anabaenopeptin	Cyclic peptides	*Tychonema* sp.	*Mycobacterium tuberculosis*	Chlipala et al., 2011
Cryptophycin 1	Cyclic depsipeptides	*Nostoc* sp. ATCC 53789	Filamentous fungi	Hirsch et al., 1990
Hapalindole	Alkaloids	*Nostoc* CCC537 & *Fischerella* sp.	*S. aureus, S. typhi, P. aeruginosa, E. coli*	Swain et al., 2017
Hassallidin A and B	Lipopeptides	*Hassallia* sp.	*Candida sp. and Cryptococcus neoformans*	Chlipala et al., 2011; Swain et al., 2017
Noscomin	Terpenes	*Nostoc commune*	*Staphylococcys epidermis, Bacillus cereus*	Chlipala et al., 2011; Swain et al 2017
Sulfolipid 1	Sulfoglycolipid	*Lyngbya* sp., *Phormidium* sp., *Scytonema* sp.	HIV-virus	Niedermeyer, 2015

1.4.2 Cyclic peptides and depsipeptides from cyanobacteria

Some of the most important classes of bioactive metabolites from cyanobacteria are the cyclic peptides and depsipeptides. The majority of secondary metabolites isolated from cyanobacteria belong to these groups (Moore, 1996; Welker & Döhren, 2006) and are frequently synthesized by hybrid NRPS/PKS pathways (Jones et al., 2009; Kehr et al., 2011; Welker & Döhren, 2006). Many of these compounds show cytotoxic and protease inhibit activity which makes them interesting as anti-cancer and anti-viral agents (Dembitsky & Řezanka, 2005; Pandey, 2015; Tripathi et al., 2010; Sainis et al., 2010). Several also show anti-fungal properties (Frankmölle, et al., 1992; Hamada & Shioiri, 2005; Martin et al., 1993), but only a few shows anti-bacterial activity (Moore, 1996; Xue et al., 2018) like the cyclic lipopeptides Lyngbyazothrins C and D (Zainuddin et al., 2009) or the cyclic peptide aeruginazole A (Raveh & Carmeli, 2010).

Some highly prolific producers of bioactive cyclic peptides are cyanobacteria that belong to the order Nostocales (Demay et al., 2019; Swain et a., 2017, Tidgewell et al., 2010) and the genus *Nostoc*. Numerous cyclic peptides with pharmacological interesting properties have been isolated from the *Nostoc* genus (*e.g.*, Banker & Carmeli, 1998; Dembitsky & Řezanka, 2005; Kaebernick et al., 2000; Okino et al., 1997; Trimurtulu et al., 1994).

1.4.2.1 Hapalosin

Hapalosin is chemically a cyclic depsipeptide first isolated from *Hapalosiphon welwitschii* W. & G. S. West, 1897 and characterized in the 1990s by Stratman and colleagues (1994). They found that hapalosin has multi-drug resistance (MDR)-reversing properties superior to that of verapamil, an agent used to re-establish sensitivity of cancer cells that have acquired MDR. This re-established sensitivity is a result of the inhibition of p-glycoprotein transporters in the cell membrane that pump the chemotherapeutic out of the resistant cancer cells. Thus, inhibition of these p-glycoproteins makes the cancer cells unable to get rid of the chemotherapeutic compounds (Simpson, 1985, Stratman et al., 1994).

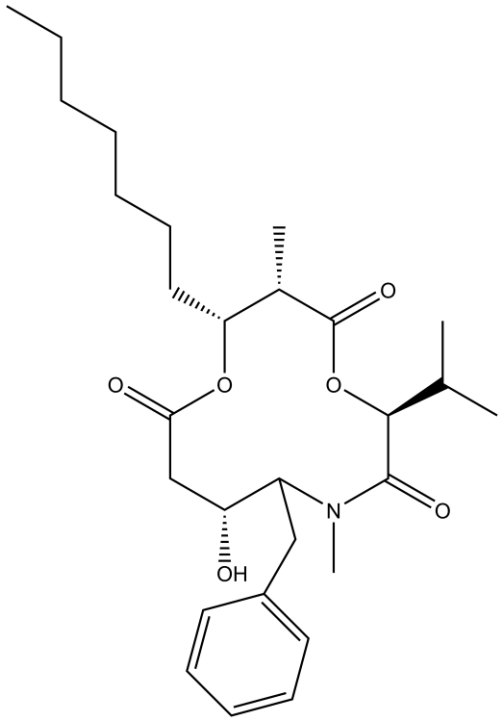


Figure 1.3. Structure of hapalosin

Since then, a handful of studies have been done with the MDR-reversing properties in mind, mostly synthesis of structural analogues of hapalosin (Dinh et al., 1996; Palomo et al., 2004) and biotesting of these on cancer cell lines (Dinh et al., 1997; Kashihara et al., 2000, O'Conell et al., 1999; Wagner et al., 1999).

In addition to *H. welwitchii*, homologous gene clusters for hapalosin biosynthesis have been identified in *Westiella intracata* and *Fischerella* sp. (Micallef et al., 2015), and more recently in a *Nostoc* sp. strain during a previous student's Master's project (Wilhelmsen, 2019).

1.5 Nostoc

The genus *Nostoc* belongs to the *Nostocaceae* family in the order Nostocales, Cyanobacteria. The *Nostoc* genus comprise of filamentous nitrogen-fixing species which are capable of cell differentiation (Kultschar & Llewellyn, 2018) depending on environmental cues and nutrient availability (Cohen, et al., 1994; Meeks & Elhai, 2002). *Nostoc* are arranged in unbranched chains of cells that forms a filament. When there is access to nutrients and combined nitrogen the chain may consist entirely of photosynthesizing vegetative cells that can release shorter filaments, hormogonia, that are motile and function as dispersal units for reproduction and the infection of plant symbionts. If there is a lack of available combined nitrogen, some of the vegetative cells can differentiate into distinguishable cells that are spaced out along the filament called heterocysts. The heterocyst cells are capable of nitrogen fixation by generating an almost completely anoxic intracellular environment by dismantling their own photosystem II, limiting the influx of gases, and having high oxidase activity. A third developmental option is spore-like cells called akinetes. Akinetes are formed under nutrient deprivation which enables cyanobacteria to survive under extreme conditions (Campbell & Meeks, 1989; Cohen et al., 1994; Liaimer et al., 2016; Meeks & Elhai, 2002).

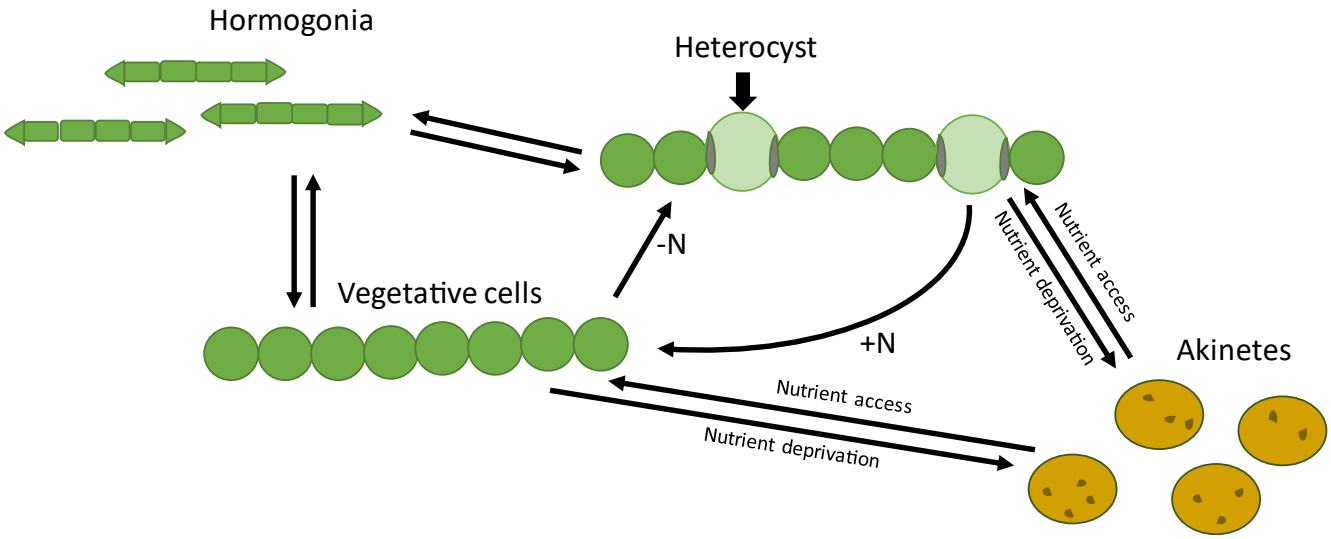


Figure 1.4. Simplified illustration of the developmental alternatives of Nostoc cyanobacterium. "N" refers to nitrogen.

They are widespread in terrestrial and in freshwater environments and a few were found in marine habitats (Dembitsky & Řezanka, 2005), around the globe in various climates (Dodds

et al., 1995). They are phenotypically highly diverse both macroscopic and microscopic levels (Rippka et al., 1979). Terrestrial *Nostoc* species are readily found in stable endophytic symbioses with bryophytes. The symbiotic interaction is restricted to specialized cavities, called domatia, on the gametophyte of the plant which is infected by *Nostoc* hormogonia that enters through pores in these symbiotic structures. *Nostoc* can form symbioses with a variety of plants and fungi and provides the symbiont with nitrogen fixed from the environment (Costa et al., 2001; Dodds et al., 1995).

Members of the *Nostoc* cyanobacteria have as early as 1500 BC been used to treat arthritis and several forms of cancer (Burja et al., 2001), but have only since the 1980s been a genus of interest in the targeted search for pharmaceuticals (Hirata et al., 2003). Species of *Nostoc* have been found to produce several types of bioactive secondary metabolites of therapeutic interest including but not limited to hydrocarbons, lipids, carotenoids, peptides, and saccharides (Dembitsky & Řezanka, 2005), that have various biological activities as antifungal (Kajiyama et al., 1998), anticancer (Chaganty et al., 2004; Magarvey et al., 2006), antiviral (Kanekiyo et al., 2005; Knübel et al., 1990), and antibacterial (Cano et al., 1990; Hirata et al, 2003) compounds (Burja et al., 2001).

The amount of various and bioactive peptides produced by species of *Nostoc* is immense. Of these, the cyclic peptides of *Nostoc* spp. have several times been shown to have drug-like anticancer activities (*e.g.*, Fewer et al., 2011; Golakoti et al., 2001, Okino et al., 1997; Trimurtulu et al., 1994).

In this project, secondary metabolites produced by two local strains belonging to the genus *Nostoc* were extracted and analyzed.

2 Objectives

The overall aim of this Master's thesis was to carry out the workflow typical for natural product discovery, from bioinformatically prediction of an organism's biosynthetic capacity to extraction and identifying compounds with bioactive properties. Based on earlier reported secondary metabolites diversity produced by *Nostoc* sp. KVJ10 and KVJ20, pilot studies on their biological activities, and predictions on their secondary metabolite production potential based on genome data, we set the following objectives:

I. To predict the biosynthetic gene clusters involved in production of secondary metabolites in *Nostoc* sp. KVJ10 and *Nostoc* sp. KVJ20, using available BGC prediction tools.
II. To test different extraction protocols and solvents in order to identify the most suitable ones for the isolation of compounds of interest.
III. To screen the extract fractions for antibacterial activity.
IV. Identify active and novel compounds in the extract(s), using available mass spectrometry based on-line tools. The results were to be matched to the genomic context of the strains.

We have set the goals of this study rather broad and open for adjustments and changes largely because the preliminary and pilot studies showed a high amount of BGCs not assigned to known compounds.

3 Materials and Methods

3.1 Organisms used in this study

The *Nostoc* sp. KVJ10 (NCBI: txid457941) and *Nostoc* sp. KVJ20 (NCBI: txid457944) used for this project were originally collected from symbiotic cavities of *Blasia pusilla* on Kvaløya island (69,64° N 18,73°E) near Tromsø, Northern Norway. Their phylogenetic affiliation has been established by Liaimer et al (2016). The total genome was sequences and characterized within previous research projects by Warshan et al (2018). Both strains have shown allelopathic properties against other strains (Liaimer et al., 2016) and were able to infect a wide range of host plants (Warshan et al., 2018).

3.2 Workflow

The first step in the workflow was to predict the biosynthetic potential for production of secondary metabolites in KVJ10 and KVJ20. This was done by running their 16S rDNA sequences was through the bioinformatic tool antiSMASH (Blin et al., 2021).
New KVJ10 cultures were started, cultivated, and harvested. The harvested and processed KVJ10 and KVJ20 cultures were extracted four times by cell extraction using methanol, a mixture of methanol and water, water, and ethyl acetate. The extracts were then processed and analyzed using ultra-high performance liquid chromatography-tandem mass spectrometry (UHPLC-MS^2). Based on the results from the chromatography and spectrometry, a few selected compounds in the ethyl acetate extract from KVJ10 were isolated and sent to the department of chemistry at The Arctic University of Norway (UiT) for nuclear magnetic resonance (NMR) analysis. The methanol and water extracts were fractioned and tested for antimicrobial activity. Any extracts that showed bioactivity were further analyzed in UHPLC-MS^2. The MS^2 spectra were also run through GNPS for dereplication and molecular networking.
Most of the workflow was executed by me. The steps which I did not participate in was the isolation of the compounds from the ethyl acetate extract, the NMR analysis, and the loading of samples into the HPLC-MS^2 system. During the bioassay most of the steps were done by the in-house chief engineer while I observed.

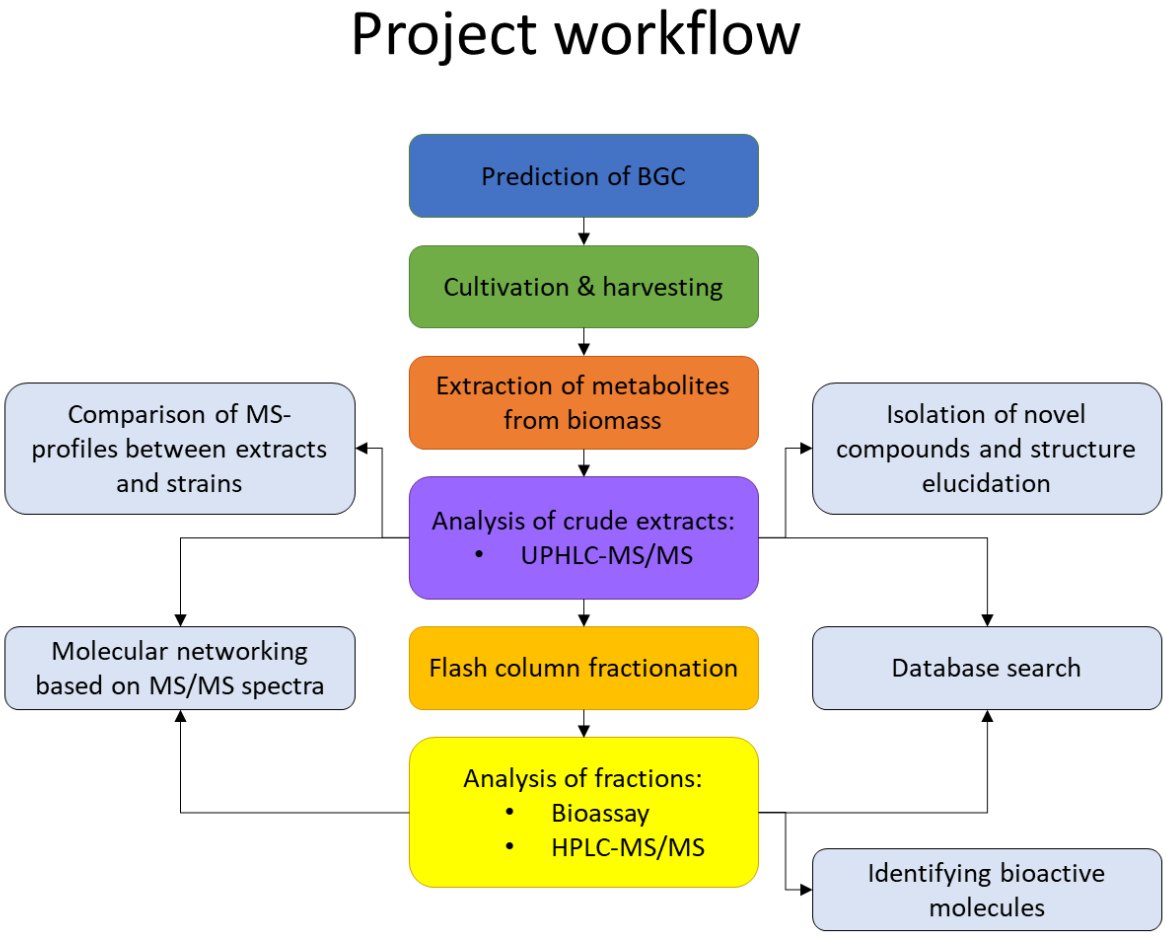


Figure 3.1. Simplified overview of the general workflow during the study project.

3.3 Prediction of biosynthetic gene clusters - AntiSMASH

The 16S rDNA sequences of both KVJ10 and KVJ20 were given in FASTA format from my supervisor. The FASTA files were run through antiSMASH 6.0 bacterial version web server (Blin et al., 2021) on "strict" detection strictness with ClusterBlast, KnownClusterBlast, SubClusterBlast, ActiveSiteFinder, and RREFinder as extra features. The clusters were revised and manually BLASTed in order to validate the identity and function of the individual genes of clusters that were of interest.

3.4 Cultivation and harvesting of *Nostoc* cultures

The *Nostoc* sp. cultures were cultivated in $BG11_0$ liquid medium before being harvested and prepared for cell extraction. $BG11_0$ is without sodium nitrate as opposed to BG11 medium, both of which are well-known stock solutions for cyanobacterial cultivation. The lack of sodium nitrate is suitable for cyanobacteria capable of aerobic nitrogen fixation (*i.e.*, heterocystous cyanobacteria like *Nostoc* sp.) because the absence of sodium nitrate prohibits the selection of mutants with abnormal heterocyst and that thus making them unable to fix nitrogen aerobically (Rippka et al., 1979)

All flasks with medium, pH_2O, heads with filters and silicone tubes used for the cultivation set-up were autoclaved to prevent the growth of other microbes in the system. The pH_2O used during this project was produced with the in-house Milli-Q® system.

3.4.1 Starting new cultures and cultivation of KVJ10

Liquid BG 11_0 medium was made by adding 1 mL of each stock solution (I.-VI.) to 1000 mL of pH_2O in gas washing bottles. The stock solutions used in the medium can be seen in Appendix 1: *Recipe for BG* 11_0 *medium*.

New KVJ10 cultures were started by inoculating 50 mL of BG 11_0 medium with 500 µL of liquid KVJ10 culture in 100 mL Erlenmeyer flasks and put on soft shaking at 24 ° C under constant artificial UV-light 30µmol $m^{-2}s^{-1}$ (36W/77 Osram Fluora).

After several weeks, the starter cultures were further cultivated in a larger cultivation system to increase growth. The system consisted of two gas wash bottles with heads, one containing approximately 1000 mL fresh medium and two KVJ10 liquid starter cultures and the other containing pH_2O which would serve as a humidity source for the system. In addition, the system was fitted with a Pall Acro™ Vent Device w/PTFE Membrane 0.2 µm and a coal filter made using a pipette tip cut in the narrowest end (to increase air outlet), cotton wool, and active coal. The coal filter was used to filtrate any potential gaseous toxins produced by the cyanobacteria. Starter cultures of KVJ10 were added to the 1000 mL liquid BG 11_0 medium and the separate parts of the system were fitted together using silicone hoses and secured with parafilm. A single assembly is illustrated in figure 3.2 and 3.3.

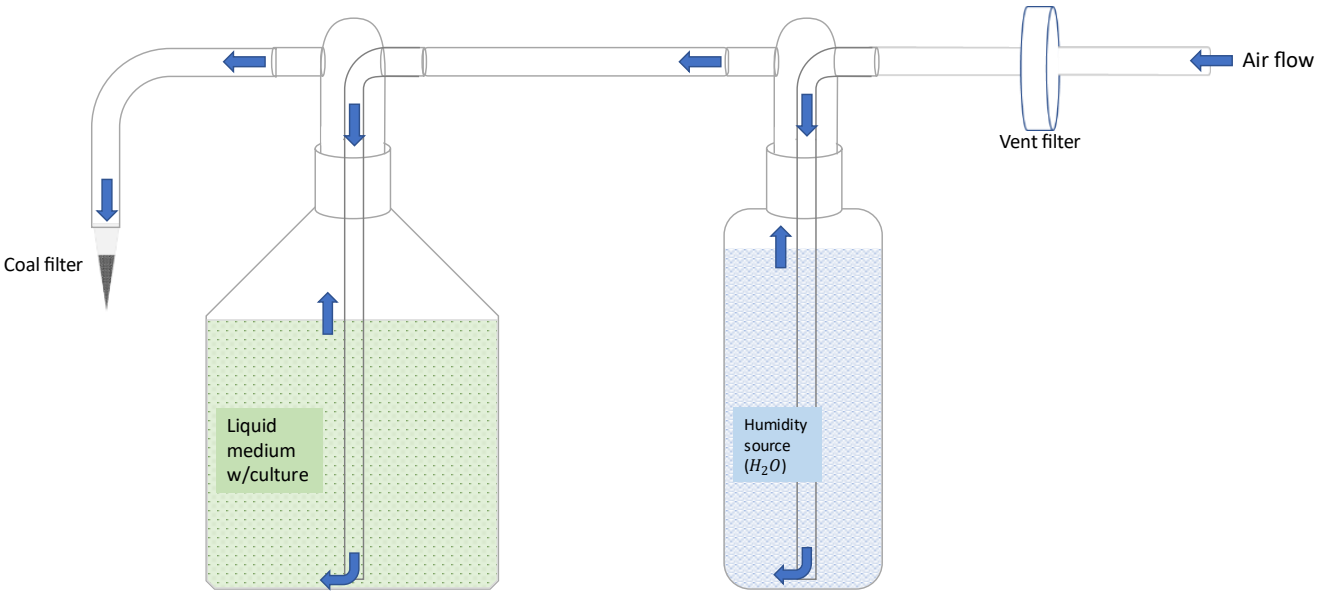


Figure 3.2. Schematic of the cultivation assembly. The flow of air is illustrated by blue arrows. Air is continuously forced into the system which goes through a vent filter and further into the humidity source, generating bubbling which transport the humid air into the next gas wash bottle containing medium and starter culture. The air flow then generates bubbling in the liquid medium containing cyanobacteria culture.

The cultivation assemblies were then connected to a continuous air flow generated by a Stellar air pump. The air flow generated bubbling in the humidity source which further transports the air, and thus bubbling, in the liquid medium containing KVJ10 culture. This is important for a continuous input of carbon dioxide to the bacterial culture (Sánchez-Bayo et al., 2020). One cultivation set-up consisting of seven assemblies were set in 24°C for approximately six weeks under constant artificial UV-light 30µmol $m^{-2}s^{-1}$ (36W/77 Osram Fluora) before being harvested.

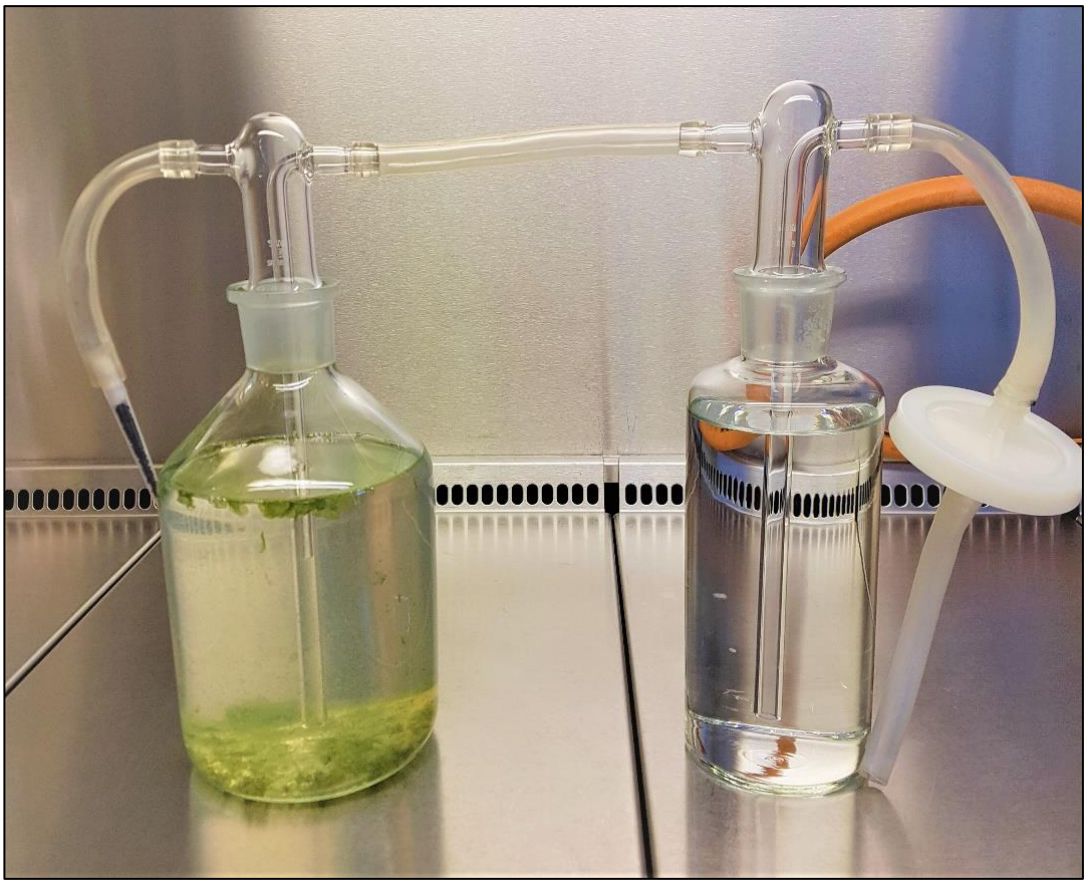


Figure 3.3. Photograph taken of the cultivation assembly. Starting from the right; the inlet hose (pale opaque) that was connected to the air flow, following was the disc-shaped vent filter connected to the humidity source bottle-head which was further connected to the gas wash bottle-head of the liquid medium, which was finally connected to the coal filter seen to the far left by silicone tubes (the orange hose is not part of the assembly). The green masses seen in the medium bottle is the added KVJ10 starter culture. Source: private.

3.4.2 Harvesting biomass from *Nostoc* cultures

As we wanted to analyze the intracellular metabolites, only the cyanobacterial cells were prepared for and extracted and not the supernatant (Pinu et al., 2017).

The liquid *Nostoc* cultures were divided into falcon tubes and centrifugated at 5000 rounds per minute (rpm) for 5 minutes. During centrifugation, the *Nostoc* cells had accumulated as a solid biomass at the bottom of the tube, with the liquid medium on top. The accumulated biomass and liquid medium were separated. All the harvested *Nostoc* biomass in the assemblies from one cultivation set-up were combined as a single strain sample. The biomass was divided into two separate containers, before being frozen at -70°C and then freeze-dried, to make the freezing and subsequent freeze-drying process more efficient. The two divisions were then combined and gravimetrically determined.

3.5 Cell extraction of secondary metabolites

KVJ20 cultures had been cultivated, harvested, and freeze dried, prior to the project.

3.5.1 Preparation of *Nostoc* samples for extraction

The harvested and freeze-dried biomass was rehydrated in ca. 300-350 mL pH_2O before being sonicated for 1 minute, 150 mL at a time, (at pulse 50, 10 Ampl 85%) using Sonics Vibra Cell™ VCX130 (Sonics®). Sonication lyses the cell by using pulses of high frequency sound waves that disrupts the cell's membrane releasing the cell contents out to its environment. The lysate was frozen at -85°C and freeze dried again. The freeze-dried and lysed biomass was gravimetrically determined before extraction.

The prepared KVJ10 sample resulted in 0.84 g freeze-dried biomass, which was only sufficient for MeOH/H_2O extraction. Therefore, an additional KVJ10 culture was needed for the ethyl acetate extraction. This KVJ10 culture was cultivated from the same colony as before and under the same conditions. The processing of this culture resulted in 4.52 g biomass of which 2.0 g was used in ethyl acetate extraction.

Each of the strain samples were extracted using dilutions of 100% Methanol, 1:2 Methanol (and pH_2O), 100% pH_2O, and 100% Ethyl acetate, respectively. This resulted in four different extracts per strain (eight extracts in total).

3.5.2 Extraction using methanol and H_2O

From the lysed and freeze-dried biomass 2 g (or 0.84g) was mixed with 150 mL 100% Methanol (MeOH) in a beaker under a fume hood. The mixture was sonicated for 30 seconds with the same settings as beforementioned. The sample was covered and left at room temperature on magnetic stirring for at least one hour. After stirring, the sample was divided into falcon tubes and centrifuged for 20 minutes at 4500 rpm. A distinct separation occurred in the centrifuged tubes between a solid "pellet" consisting of accumulated biomass at the bottom and a liquid supernatant at the top. The supernatant was collected from the tubes and combined in a round bottomed flask, attempting to get no solids from the pellet with the liquid. The remaining pellets of solid biomass were collected in a single beaker to be extracted again, this time with a 150 mL 1:2 dilution consisting of 75 mL MeOH + 75 mL pH_2O (50% MeOH and 50% pH_2O). The mixture was left at room temperature for at least one

hour but due to my own forgetfulness it was not put on magnetic stirring. The mixture was again centrifuged and processed the same way as the 100% MeOH extract, but with the last step consisting of adding 100% pH_2O to the remaining pellet which was then processed as the earlier 100% MeOH extracts with magnetic stirring. When the evaporation of the solvent was not done the same day (see *Evaporation of solvents*), the collected supernatant was stored at -20°C for later processing and investigation. The described procedure was used for both KVJ10 and KVJ20 samples, resulting in a 100% MeOH extract, a 50% MeOH (50% pH_2O) extract, and a 100% pH_2O (aqueous) extract from each strain.

3.5.3 Extraction using Ethyl Acetate

150 mL Ethyl Acetate (EtOAc) was added to 2.0 g freeze-dried biomass and left on magnetic stirring at room temperature for at least one hour. The mixture was then filtrated by vacuum using a 500 mL filtration flask equipped with a ceramic Büchner funnel and Whatman® qualitative filter paper, Grade 1 (pore size 11 µm). The collected filtrate was poured into a round bottomed flask, while the powdery biomass atop the filtration paper was removed and mixed with an additional 150 mL of EtOAc. The mixture was left for at least one hour with magnetic stirring. This mixture was filtrated in the same fashion as before and the second filtrate was combined with the first. The described procedure was used for both KVJ10 and KVJ20, resulting in one EtOAc extract for each strain.

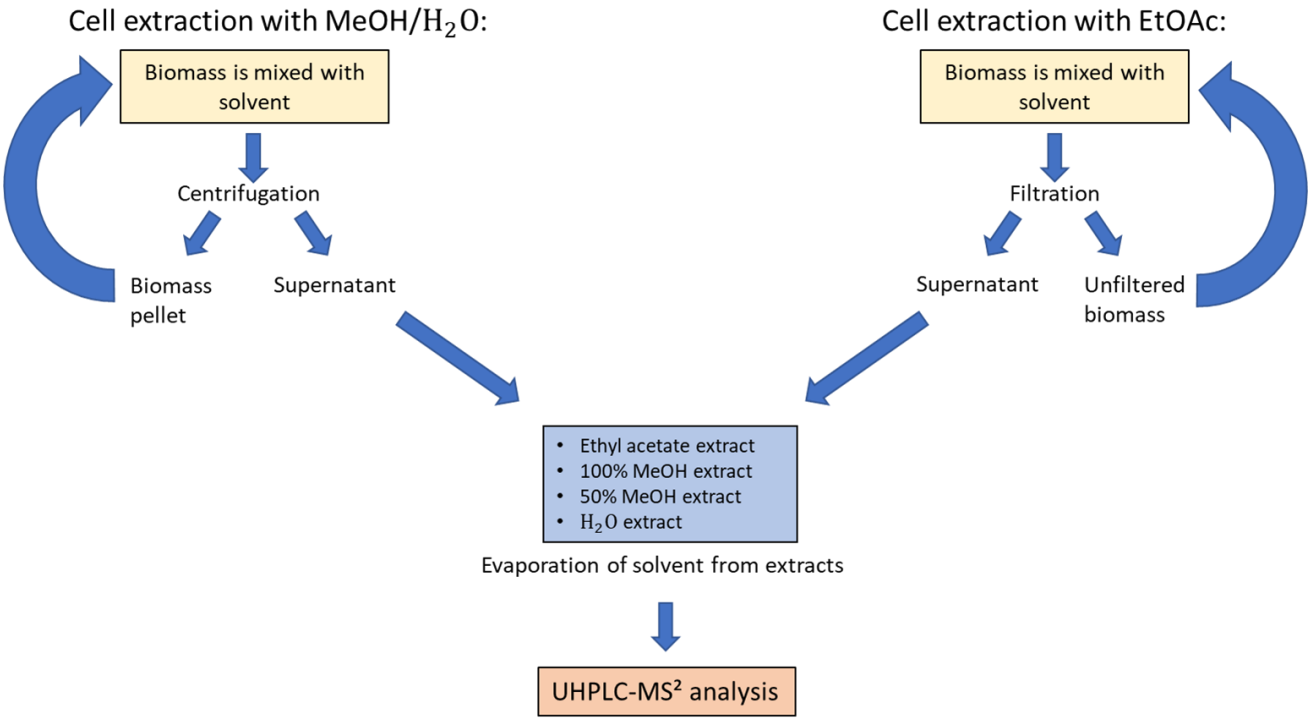


Figure 3.4. Simplified overview of the extraction process. MeOH: Methanol. EtOAc: Ethyl acetate. H_2O: water (aqueous).

The extraction processes produced four different extracts for each strain: a 100% MeOH, a 50% MeOH, an aqueous, and a EtOAc extract.

3.5.4 Evaporation of solvents

The eight extracts were evaporated to dryness under reduced pressure in a rotary evaporator (Laborota 4002 Control, Heidolph™ rotavapor system) set at 40°C with slow rotation. Starting at 180 vacuum and gradually decreasing the vacuum without causing boiling of the extracts. When the extracts had been evaporated sufficiently to fill a 50 mL round bottomed flask no more than halfway, 100 µL were pipetted from each extract and put in Eppendorf tubes and would be used for later (See *Preparation of extract samples for UHPLC-MS analysis*). The extracts were then reduced to dryness and the yield was determined gravimetrically.

3.6 Ultra-High Performance Liquid Chromatography – Mass Spectrometry

The crude extracts from KVJ10 and KVJ20 were analyzed for dereplication using reversed phase UHPLC (ultra-high performance liquid chromatography)-ESI (electrospray ionization)-MS^2 (tandem mass spectrometry). Reversed phase UHPLC separates analytes based on their polarity through a non-polar stationary phase and a polar mobile phase which takes place in the HPLC column. The column consisted of silica with nonpolar functional C18 groups for the stationary phase and the solvents used for the mobile phase were LC-MS grade acetonitrile (non-polar) and pH_2O (polar) in gradients, in addition to 0.1% formic acid to suppress polarity alternation of the analytes' functional group charge and thus improving their separation. The mobile phase started at 90% pH_2O.

3.6.1 Preparation of extract samples for UHPLC-MS

20 µL from the previously pipetted 100 µL during *Evaporation of solvents* were mixed with 180 µL 80% MeOH in appropriate vials. Their stability was checked by putting the vials in freezing temperatures for a few minutes and seeing if any particles separated from the supernatant. All samples but one dissolved properly. The KVJ20 EtOAc sample was dense with particles and after several attempts to dissolve the particles using more 80% MeOH and acetonitrile, new samples with KVJ20 and KVJ10 EtOAc extract were made using 10 µL of sample (instead of 20 µL) and 180 µL EtOAc (instead of 80% MeOH), which proved to be successful after a stability check.

The samples were then loaded and run through VION® IMS QToF on electrospray ionization (ESI) on both positive and negative polarization mode. Detailed parameters are listed in table 3.1.

Table 3.1. Parameters used for VION® IMS QToF during analysis of the extracts.

Parameters	Specification
Capillary voltage	3.00 kV
Cone gas flow	50 L/h
Desolvation gas flow	600 L/h
Desolvation temperature	250°C
High collision energy	20-60 eV
Mass range	150-1500 m/z
Scan time	0.20 s
Source temperature	100°C

The output from the UHPLC-MS² was transferred to UNIFI™ software platform. The most prominent observed masses were elucidated by calculating their elemental composition and comparing their spectra and elemental composition to compound libraries.

3.6.2 Nuclear Magnetic Resonance analysis

From the UHPLC-MS, a couple of masses of particular interest were isolated from the KVJ10 EtOAc extract and sent to the Nuclear Magnetic Resonance (NMR) facility of the department of Chemistry at the UiT – The arctic University of Norway. The NMR analysis and structure elucidation was executed by Dr. Johan Isaksson.

3.7 Fractionation

The 100% MeOH, 50% MeOH, and aqueous extracts from KVJ10 and KVJ20 were fractionized prior to testing for bioactivity. The extracts were fractionized using FLASH-liquid chromatography by Biotage® SP4 Flash Purification System which fractionize after polarity. FLASH fractionation was used in this manner to roughly separate components from the extracts into several sample fractions which was then screened for activity. The EtOAc crude extracts from KVJ10 and KVJ20 were not fractionized and not tested for bioactivity.

3.7.1 Preparation of SNAP cartridge

6.5 g Diaion® HP-20SS poly-benzyl-resin was mixed with 75 mL 100% MeOH in a 100 mL Erlenmeyer flask and left for ca. 30 minutes. The methanol was then discarded and replaced with pH_2O. An empty Biotage® SNAP Flash cartridge were mounted on top of a vacuum manifold. pH_2O was added to the cartridge before the treated resin was put into the cartridge and the vacuum turned on. When the resin material had accumulated at the bottom with the water at least 1 cm above the material, the cartridge was detached from the manifold and lidded. The prepared SNAP cartridge was stored at 4°C until later use.

3.7.2 Preparation of samples for FLASH fractionation

The 100% MeOH, 50% MeOH, and aqueous extracts from KVJ10 and KVJ20 were dissolved in 100% MeOH, 50% MeOH (50% pH_2O), and pH_2O, respectively, with shaking by hand. The dissolved extracts from each strain were combined in two round bottomed flasks, resulting in one sample of KVJ10 extract and one of KVJ20 extract. 2.0 g Diaion® HP-20SS resin was added to each sample and dried under vacuum at 40°C using a rotary evaporator until completely dry.

3.7.3 Fractionation by FLASH-liquid chromatography

The dried KVJ10 extract sample were loaded into the prepared SNAP cartridge (see *preparation of SNAP cartridge*) and put in the Biotage® SP4 Flash Purification System with settings at 20 mL/min flow rate and 80 mL sample volume. The reagents used for the mobile phase were pH_2O , MeOH, and acetone. The sample was eluted with the mobile phase gradient according to table 3.2, resulting in 27 Flash tubes of sample. The Flash tubes were pooled into six fractions in Polyvap tubes. Fraction number 6 was pooled into two Polyvap tubes and later combined during drying because the volume of the combined flash tubes 16-27 were too large for a single tube. The same procedure was done for KVJ20.

Table 3.2. Overview of the flash tubes, fraction number, and their respective mobile phase gradients during fractionation by FLASH-liquid chromatography.

Flash tubes	Fraction number	pH_2O	MeOH	Acetone
1-3	1	95 %	5 %	0 %
4-6	2	75 %	25 %	0 %
7-9	3	50 %	50 %	0 %
10-12	4	25 %	75 %	0 %
13-15	5	0 %	100 %	0 %
16-18	6	0 %	50 %	50 %
19-27	6	0 %	0 %	100 %

The fractions were reduced to dryness using a Büchi Syncore Polyvap heated to 45°C. When the fractions had completely dried, the yield was determined gravimetrically, and the yields were used to calculate the amount of Dimethyl Sulfoxide (DMSO) to add to each fraction to get a dilution with a concentration of 40.0 mg/mL. If the amount of DMSO to add was below 500 µL, the concentration was adjusted to 20.0 mg/mL, 10.0 mg/mL, *etc*. The tubes containing the fractions with DMSO were put on a universal shaker at 160 rpm and rotated about 90° every hour until adequately dissolved. After about five hours the fractions were transferred into individual 1.8 mL Cryotubes and stored at -20°C until further use.

3.8 Anti-bacterial Growth Inhibition Assay

3.8.1 Preparation of fractions for bioassays

For the bioassays, solutions of the fractions were made using 2.5% DMSO (solvent) to make 1 mL samples with a solute concentration of 2.0 mg/mL. The amount of each fraction to be used in the dilutions were calculated. For example, for a fraction with a concentration of 40.0 mg/mL, 50 µL of the fraction sample (and 950 µL of solvent) would be needed to make a 1 mL solution with a concentration of 2.0 mg/mL, see Equation 1 below.

$$\frac{2.0\ mg/mL \times 1\ mL}{40.0\ mg/mL} = 0.05\ mL = 50\ \mu L \qquad \textit{Equation 1.}$$

The solutions were made in Eppendorf tubes and stored at -20°C until biotesting.

3.8.2 Preparation of stock solution of samples for biotesting

The fractions from KVJ20 and KVJ10 were tested on five different bacteria: *Enterococcus faecalis* (LGC Standards, ATCC 29212), *Escherichia coli* (LGC Standards, ATCC 25922), *Pseudomonas aeruginosa* (LGC Standards, ATCC 27853), *Staphylococcus aureus* (LGC Standards, ATCC 25923), and *Streptococcus agalactia* (LGC Standards, ATCC 12386). The growth media used during the assay was Mueller-Hinton (MH) broth (Difco, 275730) for *E. coli*, *P. aeruginosa*, and *S. aureus* and Brain Heart Infusion (BHI) (Sigma Aldrich, 53286) for *E. faecalis* and *S. agalactiae*.

A scoop of bacteria from blood agar plates were transferred to 8.0 mL of respective growth media and incubated at 37 °C for approximately 20 hours. 2.0 mL of the liquid cultures were transferred to separate Erlenmeyer flasks containing 25.0 mL growth media and incubated for 1.5 hours (*E. faecalis*, *E. coli*, and *S. agalactiae*) or 2.5 hours (*P. aeruginosa* and *S. aureus*) in a shaking incubator at 37 °C and 100 rpm.

The bacterial suspensions were then diluted by transferring 100 µL bacteria culture to 9.9 mL new growth medium giving a dilution factor of 1:100 and from this dilution taking 4.0 mL and transfer it to 36.0 mL of new growth medium making a dilution of 1:10. 50 µL of these dilutions were added to the designated 96-well microtiter plates in column 2-12. One plate for each pathogen. In column 1 and 12, 50 µL pH_2O was added and 50 µL of KVJ10 and KVJ20

DMSO-fraction samples were added in duplicates to the plates in column 2-3 and 4-5, respectively. 50 µL growth media were added to column 1, either BHI or MH broth for the respective bacteria. The plate layout is illustrated in figure 3.5. The plates were incubated at 37 °C overnight.

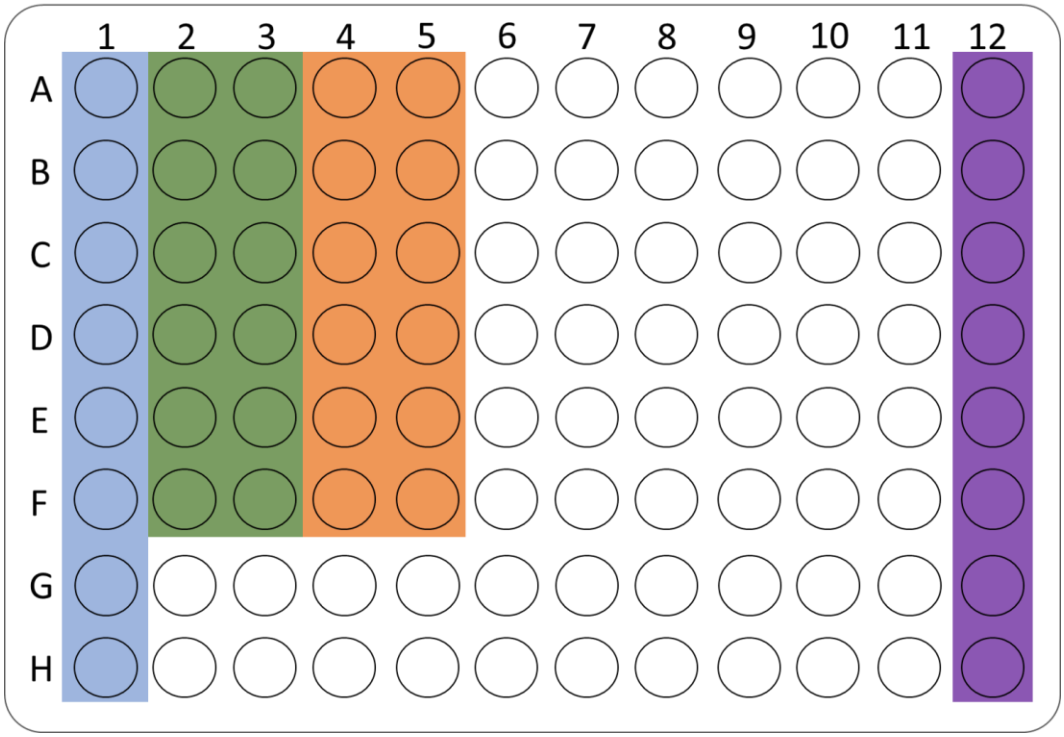


Figure 3.5. Microtiter plate layout. Highlighted columns (col.) represent the following: Blue (col. 1): media blank, containing pH_2O + growth media. Green (col. 2 & 3): KVJ10 fractions 1-6 (A-F) + bacterial suspension. Orange (col. 4 & 5): KVJ20 fractions 1-6 (A-F) + bacterial suspension. Purple (col. 12): growth control, containing pH_2O + bacterial suspension. Only highlighted columns were used for deciding activity or inactivity.

3.8.3 Gentamicin control

In addition to the five KVJ-assay plates, a gentamycin control plate was made as a control for the assay and normal growth of the bacterial strains tested. The control was plated with serial dilutions of gentamycin and autoclaved pH_2O on a 96-well microtiter plate. To the dilution series, 50 µL bacterial suspension for each bacterium was added row-wise. The gentamicin control then had a column-wise final test concentrations of 0.02, 0.03, 0.07, 0.13, 0.25, 0.50, 1.00, 2.00, 4.00, 8.00, 10.00 and 16.00 µg/mL. The control plate was incubated under the same conditions as the anti-bacterial growth inhibition assay plates.

3.8.4 Plate reading

The next day the microtiter plates were visually inspected and was measured for optical density using Victor Multilabel Counter (PerkinElmer) at 600 nm (OD_{600}) and classified as active (< 0.05), inactive (> 0.09, or questionable (0.05 - 0.09) according to the measured absorbance (OD_{600}). The level of absorbance thus represents the density of bacterial cells in each well of the microtiter plate. The higher OD_{600} the higher density of cells. Fractions that were classified as active were analyzed using UHPLC-MS and major peaks observed in the produced BP chromatograms were dereplicated to attempt identifying the compound(s) with bioactivity.

The gentamicin control plate was read in the same way and the minimum inhibition concentration (MIC) was determined and compared to reference MIC values for the pathogens.

3.9 UHPLC-MS of bioactive fractions

Fractions that showed anti-bacterial activity and the two "neighboring" fractions were run through UHPLC-MS. The parameters used were the same as in UHPLC-MS during dereplication of the crude KVJ extracts.

The most prominent peaks were elucidated by calculating their elemental composition and comparing their mass spectra and elemental composition to compound libraries and available literature.

3.10 Molecular Networking using Global Natural Products Social Molecular Networking (GNPS)

GNPS (Global Natural Products Social Molecular Networking) is an open-access platform for analyzing, organization, and sharing of tandem mass (MS^2) data. GNPS provides third-party MS^2 spectral libraries relevant to NPs, as well as community shared raw, processed, or identified MS^2 spectrometry data, that can be used for online dereplication and automated molecular networking. Inputted data sets are matched to reference spectral libraries to annotate molecules and discover putative analogs (Wang et al., 2016).

The main purpose for using GNPS in this project was to gain experience in molecular networking and using GNPS as a tool for dereplication.

Spectra from extracts and fractions were exported from UNFI as processed openMS data (.mzML) files and converted into centroide data using MSConvert software (Chambers et al., 2012; Kessner et al., 2008). In MSConvert the files were filtrated using "peakPicking" with MS levels set to 1-2 (as according to GNPS documentation "Mass Spectrometry Files Conversion" under "Data Preparation/Upload") and outputted as mzML files with 32-bit binary encoding precision. The converted spectra files were then uploaded and run through GNPS METABOLOMICS-SNETS-V2 workflow version release_30 (Wang et al., 2016) in Mozilla Firefox web browser with default parameters and the extract and fraction spectra grouped together in the following way: Spectrum Files G1: KVJ10 crude extracts, G2: KVJ20 crude extracts, G3: KVJ10 fraction 4-6, and G4: KVJ20 fraction 3-5.

4 Results

For this project, the primary focus and work was dedicated to find new metabolites with bioactivity.

4.1 AntiSMASH

AntiSMASH predicted in total 24 BGC regions for KVJ20 and 87 for KVJ10, however, many of the predicted gene clusters from the KVJ10 antiSMASH runs were shown to be fragmented and incomplete. Also, for both strains there were many clusters without any matches, *i.e.*, 0% similarity to similar known biosynthetic clusters (not shown).

Table 4.1. Biosynthetic gene clusters prediction by antiSMASH ver.6.0 for KVJ10 genome. nt=nucleotides.

Most similar known cluster	Similarity	Type	Region	Location (nt)	
				from	to
Meilingmycin	2 %	Phoshonate	2636.1	1	26 031
Polyoxypeptin	8%	NRPS, PKS	2442.1	1	2355
Putrebactin/Avaroferrin	20 %	Siderophore	2101.1	1	10 298
Crocacin	23 %	T1PKS	150.1	1	9 978
Carotenoid	25 %	Terpene	2094.1	93 944	114 147
Heterocyst Glycolipids	42 %	Hgle-KS	273.1	1	7 710
Nosperin	46 %	T3PKS	607.1	1	5 185
Nostopeptolide A2	50 %	T1PKS, NRPS	173.1	1	21 791
Hapalosin	60 %	NRPS, T1PKS	15.1	1	56 162
Ectoine (Other)	66 %	Ectoine	1972.1	39 140	49 553
1-Heptadecene	100 %	T1PKS	1621.1	1	1 864
Anabaenopeptin Nz857/Nostamide A	100 %	NRPS	3.1	1	17 370
Anabaenopeptin Nz857/Nostamide A	100 %	NRPS	546.1	1	6 068
Anabaenopeptin Nz857/Nostamide A	100 %	NRPS	1042.1	1	2 559
Anabaenopeptin Nz857/Nostamide A	100 %	NRPS	1352.1	1	2 313
Carotenoid	100 %	Terpene	1791.1	1 187	22 263

Geosmin	100 %	Terpene	1466.1	1	2 065
Geosmin	100 %	Terpene	1661.1	1	1 142
Geosmin	100 %	Terpene	6724.1	1	1 089
Hapalosin	100 %	T1PKS, NRPS	31.1	1	41 520
N-Acylalanine	100 %	Homoserine lactone	7460.1	1	1 114

Table 4.2. Biosynthetic gene clusters prediction by antiSMASH ver.6.0 for KVJ20 genome. nt=nucleotides.

Most similar known cluster	Similarity	Type	Region	Location (nt) from	to
Plpa1/Plpa2/Plpa3	33 %	Lanthipeptide (RiPP)	67.1	20 763	59 748
Aeruginoside 126B/ Aeruginoside 126A	41 %	NRPS	26.1	21 372	83 223
Nostopeptolide A2	50 %	NRPS, T1PKS	15.1	43 687	97 697
Nostopeptolide A2	50 %	NRPS, T1PKS	18.1	1 953	87 899
Micropeptin K139	50 %	NRPS, betalactone	103.1	16 027	74 841
Heterocyst Glycolipids	57 %	Hgle-KS, T1PKS	88.1	5 942	56 931
Nostocyclopeptide A2	57 %	NRPS, microviridin	150.1	1	13 934
Anabaenopeptin NZ857/Nostamide A	100 %	NRPS	146.1	6 811	35 062
Geosmin (Terpene)	100 %	Terpene	35.1	92 433	114 685
Heterocyst Glycolipids (Other)	100%	Hgle-KS, T1PKS	74.1	1	30 617

Several of the fragmented KVJ10 clusters were manually pieced together by Anton Liamier. Some of the gene clusters that were predicted by antiSMASH to be incomplete were in fact not. An example of an incomplete prediction is the BGC for a heterocyst glycolipid shown in figure 4.1. Gene clusters that have been correctly and manually glued together will be referred to as “revised” BGC(s).

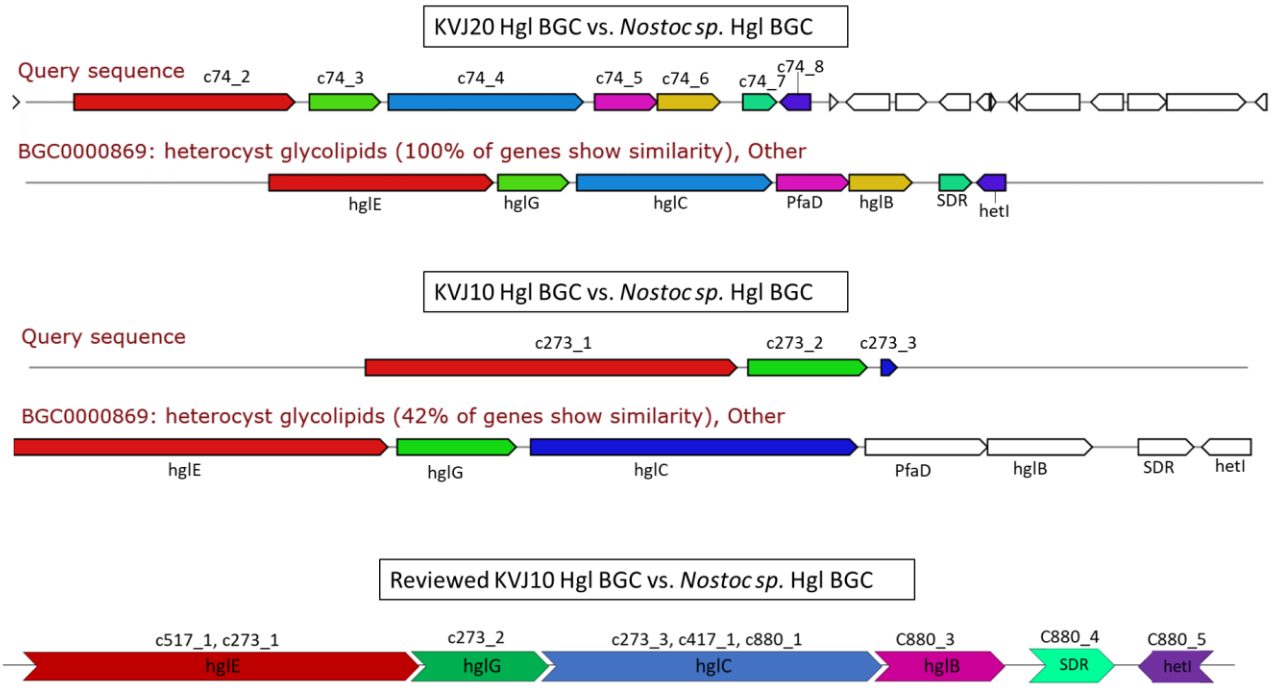


Figure 4.1. antiSMASH prediction of the biosynthetic gene cluster for a heterocyst glycolipid (BGC0000869) in KVJ20 and KVJ10, respectively, and the revised KVJ10 heterocyst glycolipid BGC.

Figure 4.1 shows an example of a BGC for a heterocyst glycolipid predicted with antiSMASH with 42% gene similarity in KVJ10 region 42 and the same BGC prediction with 100% gene similarity in KVJ20 region 74 (and additionally region 88 at 57% gene similarity) (see table 4.2). Also shown is the revised BGC which were considerably more intact in KVJ10 than predicted by antiSMASH.

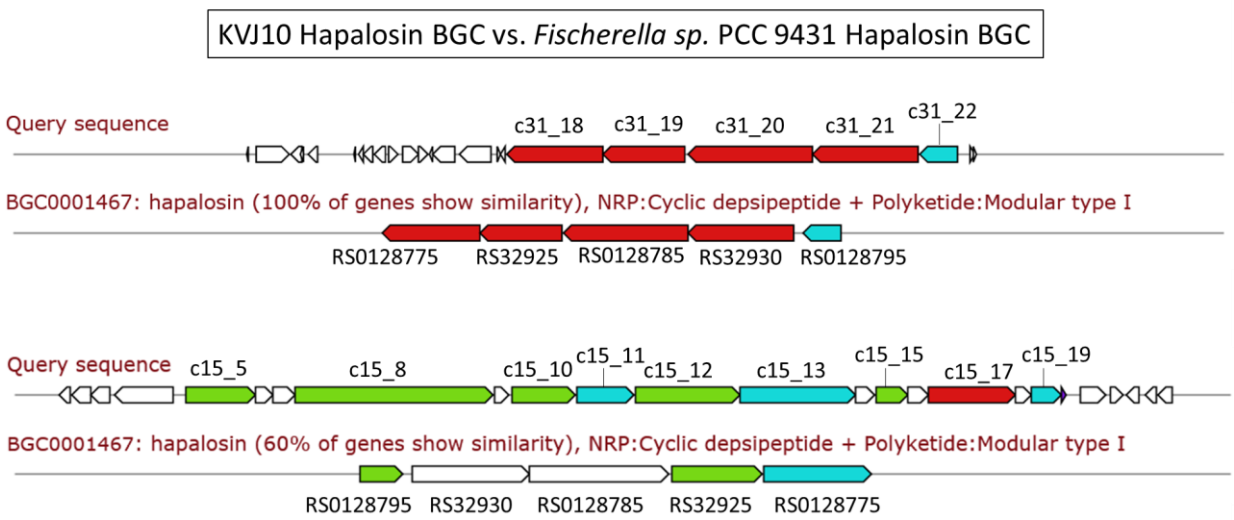


Figure 4.2. AntiSMASH prediction of two BGCs for hapalosin, with gene similarities at 100% and 60%, respectively.

AntiSMASH predicted two separate BGC for hapalosin in KVJ10 for 100% gene similarity at region 31 and 60% similarity at region 15. No hapalosin gene clusters were predicted in KVJ20.

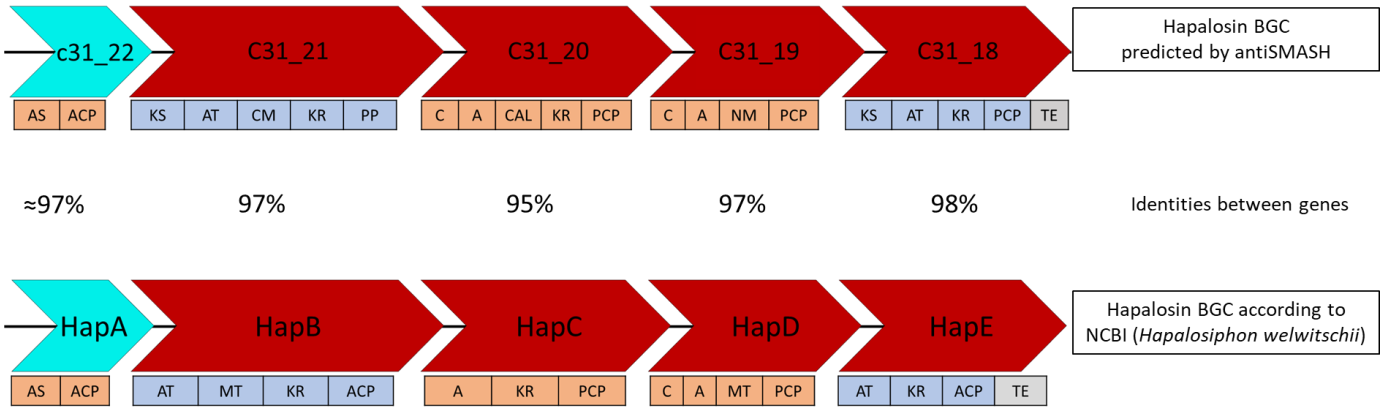


Figure 4.3. Illustration of hapalosin gene cluster predicted by antiSMASH and the hapalosin biosynthesis genes registered in the National Center for Biotechnology Information (NCBI) protein database as HapA, HapB, HapC, HapD, and HapE from Hapalosiphon welwitschia, and their similarities in percent (%).

The identities between the predicted gene cluster for hapalosin production in KVJ10 and the hapalosin biosynthetic genes of *Hapalosiphon welwitschii* ranged from 95% up to 98% percent.

4.2 Yield from cell extraction

Table 4.3. Overview of the KVJ10 and KVJ20 cell dry weight used in the extraction, the yield (dry weight) and the recovery rate for each extraction method, i.e., how many percent is "yield dry weight" of "dry weight before extraction".

Strain	Dry weight before extraction	Extract	Yield dry weight (recovery rate)
KVJ10	0.84 g	100% MeOH	0.10 g (11.9 %)
		50% MeOH	0.08 g (9.5 %)
		Aqueous	0.05 g (6.0 %)
	2.0 g	EtAc	0.04 g (2.0 %)
KVJ20	2.0 g	100% MeOH	0.18 g (9.0 %)
		50% MeOH	0.11 g (5.5 %)
		Aqueous	0.19 g (9.5 %)
		EtAc	0.05 g (2.5 %)

The recovery rate varied between two percent to around 12 percent. The EtOAc extracts of both KVJ10 and KVJ20 had the lowest yield with 0.04 and 0.05 g, respectively, while 100% MeOH had the highest yield for KVJ10 and the aqueous extracts the highest for KVJ20. For the respective extraction-method used, the yield from the cell extraction was as expected from earlier experiments with bacteria.

4.3 UHPLC-MS of crude extracts and dereplication

The base peak intensity (BPI) chromatogram of KVJ10/KVJ20 EtOAc and 50% MeOH will be presented and discussed in the results section as these had the most noteworthy differences that portray the strain profiles and variations between extract methods. Observations of the KVJ10 UHPLC-MS results lead to the isolation of two compounds from the EtOAc extract and subsequent NMR analysis and structure elucidation of one of the two compounds. The third compound turned out to be obtained at too low yields for structural analysis via NMR.

All BPI chromatograms of all crude extracts from both KVJ10 and KVJ20 are available in Appendix 2.

A BPI chromatogram is generated by the plotting of ion signals with the greatest detected response (*i.e.*, base peaks) in a string of recorded mass spectra over time. Because the spectrometry method used was reversed phase chromatography, the compounds in the samples are separated based on their polarity, with the most polar molecules appearing first and the most non-polar appearing last. Examples of polar molecules are simple sugars and small metabolites while fatty acids, lipids and alkanes are examples of non-polar molecules. Furthermore, compounds with similar observed masses, RT (retention time), and fragmentation patterns can, with a high likelihood, be assumed to be the same molecule.

Prominent peaks without annotation in the chromatograms indicates that the compound represented is of no interest in accordance with the aim(s) of this project. For example, if a molecule of great intensity were shown to most likely be a common cell constituent of cyanobacteria it will not be given much attention (expect from compound 4* which is used as an example) as the molecules of focus are those that are unique to the strains and those that show bioactivity (see *Signals of potential bioactivity*). Unidentified compounds that were found in both strains is also set aside. Compounds that are unknown but only found in one of the two strains is annotated but not given attention because the calculation of elemental composition did not give any satisfactory results or have an insufficient fragmentation spectrum, both of which makes it difficult to make suggestions of type of molecule.

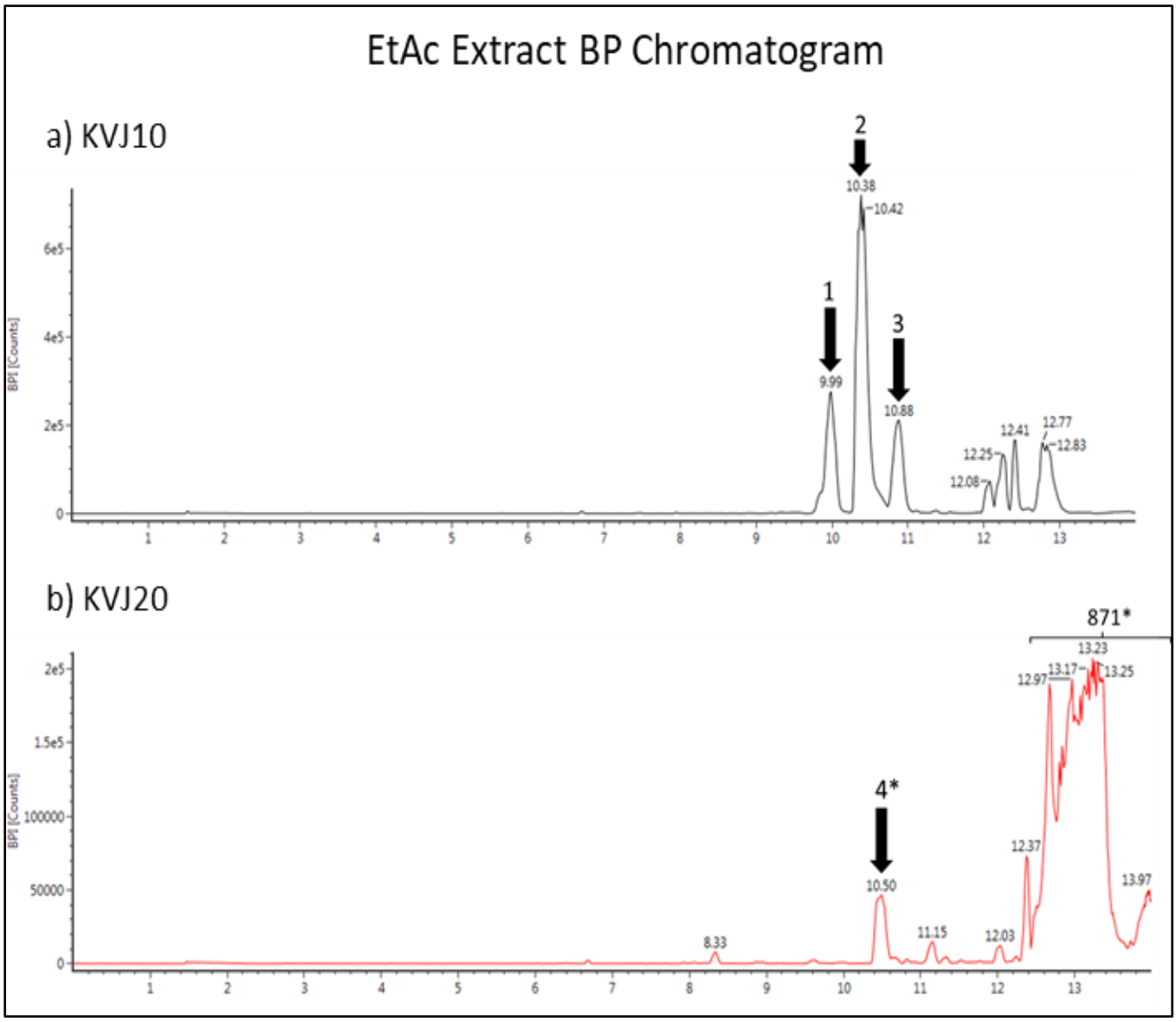


Figure 4.4. Base peak intensity chromatograms as a function of time of KVJ10 (a) and KVJ20 (b) Ethyl Acetate extract from UHPLC-MS analysis. The x-axis shows Base peak intensity (BPI), while the y-axis shows Retention time in minutes. Arrows and brackets with number point out base peaks that will be discussed further. Number with one star (*) indicates a peak which represents an identified compound that is found in both strains.

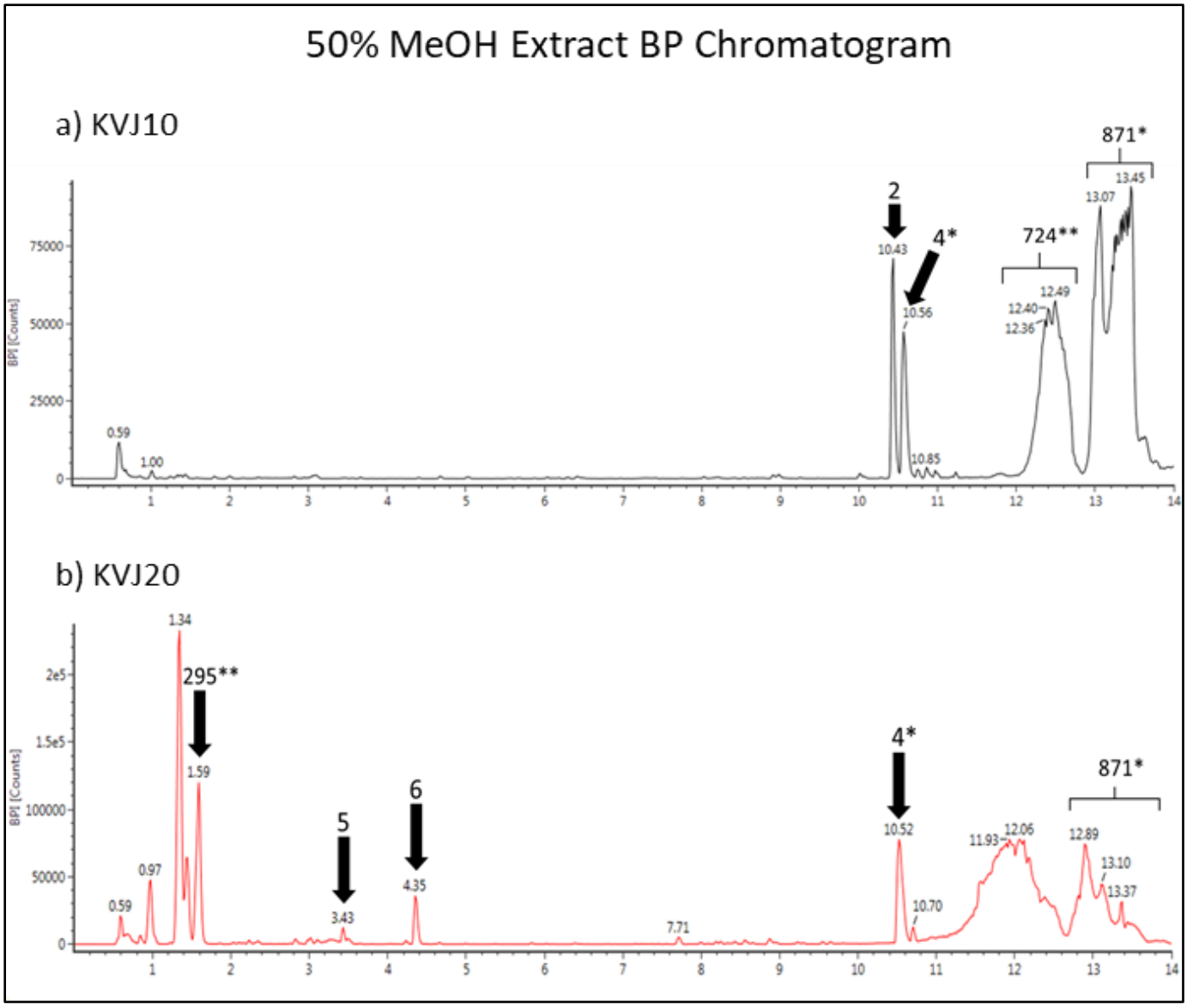


Figure 4.5. Base peak intensity chromatograms as a function of time of KVJ10 (a) and KVJ20 (b) 50% MeOH extract from UHPLC-MS analysis. The x-axis shows Base peak intensity (BPI), while the y-axis shows Retention time in minutes. Arrows and brackets with number point out base peaks that will be discussed further. Number with one star (*) indicates a peak which represents an identified compound that is found in both strains, while number with double stars (**) indicates an unknown compound that is unique to the strain of which the extract is from.

Table 4.4. Overview of dereplicated compounds that were unique to KVJ10 or KVJ20 and one common compound which is used as an example. RT: Retention Time; Peak number: refers to numbered peaks in chromatograms in figure 4.4 and 4.5.

Compound	M+H (*m/z*)	RT	Suggested composition	Suggested compound	Present in:	Suggestion based on:
1	490.32	9.5-10.5	$C_{28}H_{43}NO_6$	Hapalosin 489	KVJ10	Library match + MS² fragment pattern match (Stratman et al., 1994)
2	504.33	10.0-11.5	$C_{29}H_{45}NO_6$	Hapalosin 503	KVJ10	MS² fragments compared with hapalosin 489(1), NMR
3	502.32	10.5-11.5	$C_{29}H_{43}NO_6$	Hapalosin 501	KVJ10	MS² fragments compared with hapalosin 489(1)
4(*)	577.47	10.0-11.5	$C_{32}H_{64}O_8$	Heterocyst glycolipid	Both	MS² fragment pattern match (Bauersachs et al., 2009)
5	1077.46	3.5-4.5	$C_{45}H_{72}N_8O_{20}S$	Suomilide B	KVJ20	MS² fragment pattern match (Riba et al., 2020)
6	1007.66	3.0-3.5	$C_{41}H_{66}N_8O_{19}S$	Suomilide-1006	KVJ20	MS² fragment pattern match (Schneider, 2020)

The majority of ions, and predominantly those of highest count, were generally towards the apoloar side in both KVJ10 and KVJ20. There were less polar compounds detected in KVJ10 than in KVJ20. KVJ20 had some major peaks towards the polar side of the range, except from the EtOAc extract, while KVJ10 had no noteworthy peaks in the highly polar area in any of the extracts. One of the major polar peaks, with mass 295 *m/z*, was unique to KVJ20 and not detected in KVJ10 at all. The elemental composition calculated for this compound was $C_{14}H_{18}N_2O_5$, but no library hits were made for a specific molecule and literature search gave no results. However, two other unique KVJ20 peaks (compound **5** and **6**), though not major, were identified as suomilides (see figure 4.5 and table 4.4).

Two well-represented compounds detected as individual major peaks in both KVJ10 and KVJ20 were observed at 871 and 577 *m/z* $[M + H]^+$, respectively. The mass observed at 871 *m/z* corresponded to a molecular formula of $C_{56}H_{70}N_8O$ and was identified as pheophytin *a*. The mass at 577 *m/z* (compound **4**) corresponded to 1-(*O*-hexose)-2,25-hexacosanediol with a formula of $C_{32}H_{64}O_8$. Their respective identities were confirmed by MS² fragmentation pattern. Pheophytin *a* is an electron transporter in photosystem II (Allakhverdiev et al., 2010) and the latter a heterocyst glycolipid (Bauersachs et al., 2009), both of which are common constituents of cyanobacteria. An example of how a compound identity were proposed through a MS² spectrum is shown in figure 4.6. The suggested identity of most compounds in table 4.4, and figures 4.12 and 4.13 were made based on such spectra.

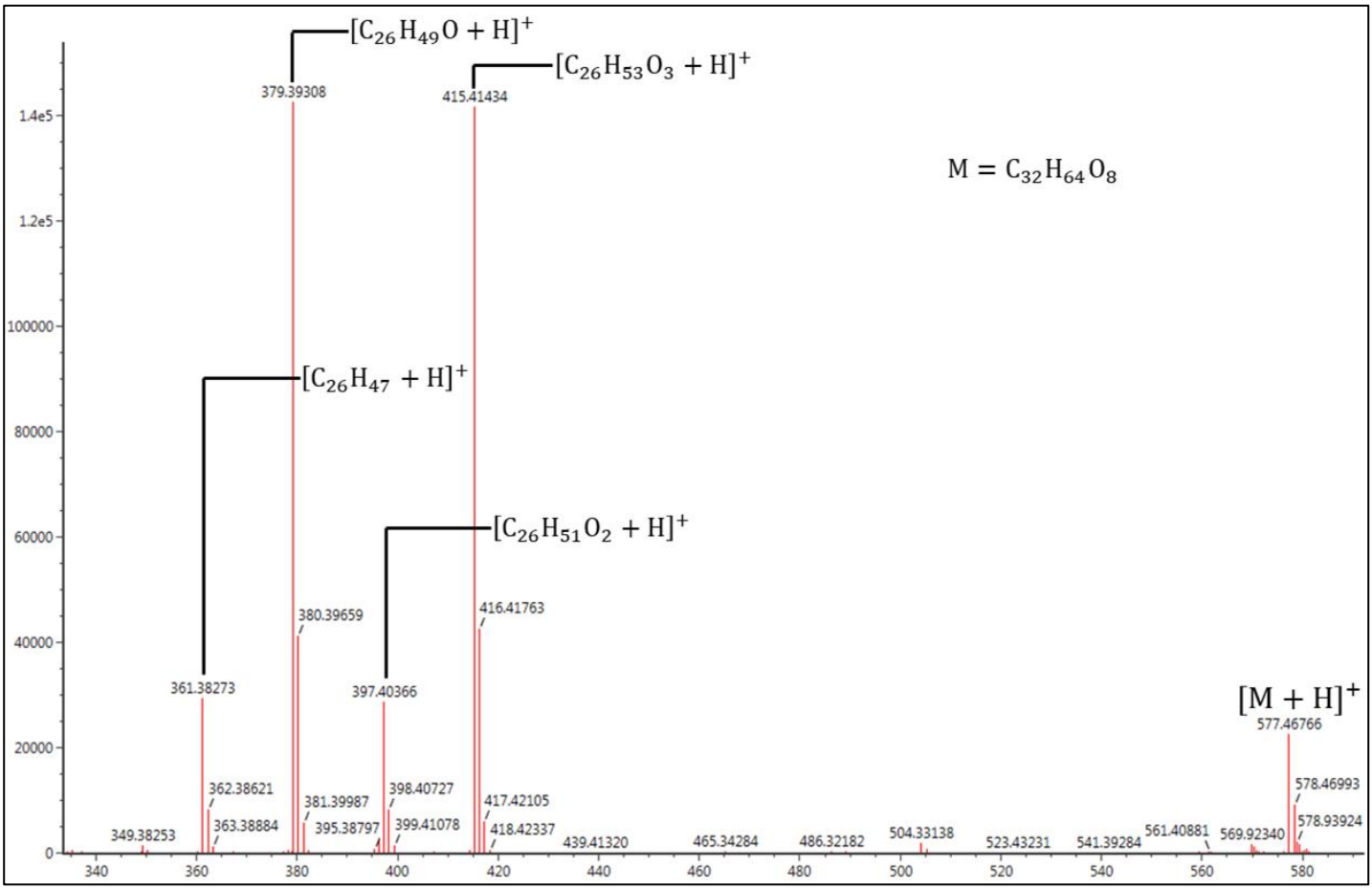


*Figure 4.6. MS² spectrum for compound **4** which was found in both strains at 577 m/z $[M + H]^+$. The x-axis shows Intensity (counts), while the y-axis shows Observed mass [m/z]. Elemental composition and structure of the compound is shown. Precursor ion and product ions used for identification is annotated.*

The protruding distinction that caught our interest were the three peaks observed in the BPI chromatogram of KVJ10 EtOAc extract which are annotated with number **1**, **2**, and 3 in figure 4.4. The observed masses of these three peaks were *m/z* 490, 504, and 502 $[M + H]^+$, respectively. Calculation of the elemental composition and the MS² spectrum of 490 gave a library hit matching that of hapalosin 489 (**1**). The similarity in retention time and observed mass of **2** and 3, confirmed by elemental composition calculations of their precursor ions and product ions, in addition to the spectra patterns, revealed that they were most likely variants of hapalosin 489 (**1**), one of which had to our knowledge only been described recently by Figueiredo et al. (2021) and the other not found described in any literature. Compound **2** and 3 thus became the main attraction in this project. The two discovered variants of hapalosin 489 (**1**) will, in addition to their compound number, be referred to as hapalosin 503 (**2**) and hapalosin 501 (3). The large peak at retention time 10.42 in figure 4.4 that was seen in several KVJ10 crude extract chromatograms had a mass of *m/z* 1007 $[M + H]^+$ and was identified as a dimer of compound **2** (see figure 4.8 b).

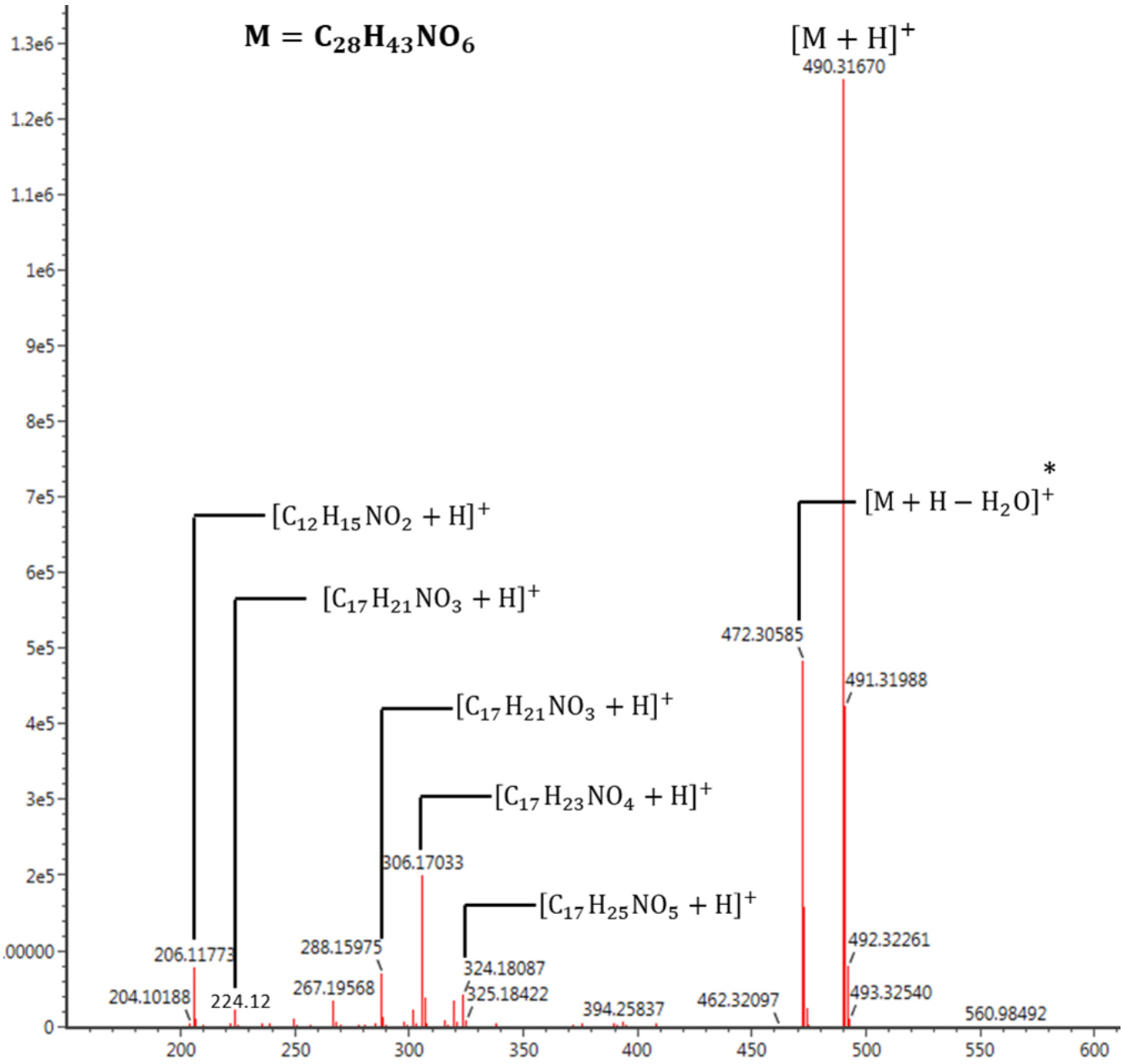


*Figure 4.7. Elemental composition and MS² fragment ion pattern of Hapalosin 489 (**1**). The x-axis shows Intensity (counts), while the y-axis shows Observed mass [m/z]. Precursor ion is annotated with* $[M + H]^+$*. Product ion elemental compositions annotated with "*" are uncertain.*

Hapalosin 489 (**1**) observed at *m/z* 490.32 had a corresponding elemental composition of $C_{28}H_{43}NO_6$. Major product ions of **1** were *m/z* 472.30 $[M + H - H_2O]^+$, *m/z* 324.18 $[C_{17}H_{25}NO_5 + H]^+$, *m/z* 306.17 $[C_{17}H_{23}NO_4 + H]^+$, *m/z* 288.16 $[C_{17}H_{21}NO_3 + H]^+$, *m/z* 206.12 $[C_{12}H_{15}NO_2 + H]^+$.

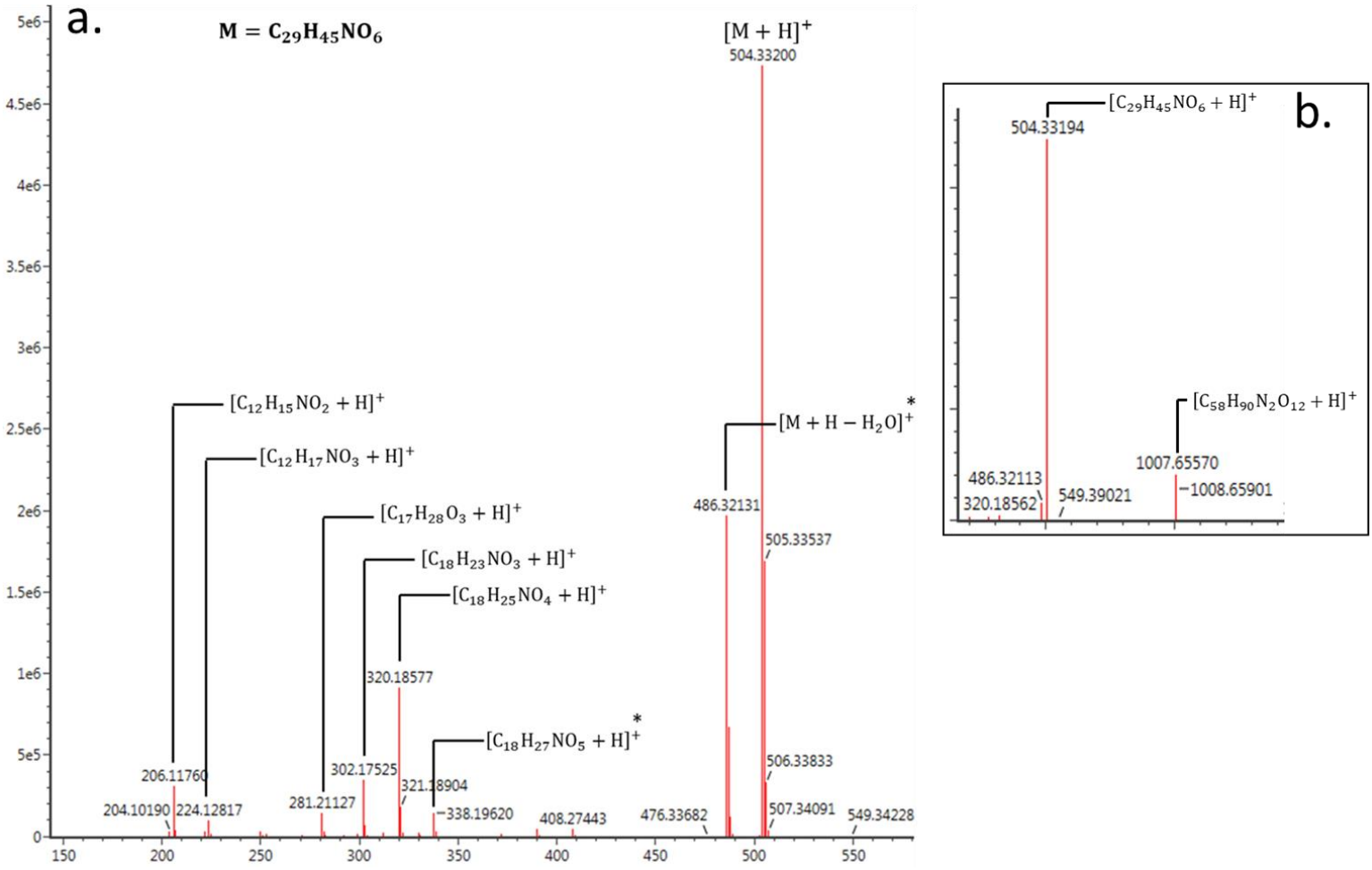


*Figure 4.8. a) Elemental composition and MS² fragment ion pattern of compound **2**. Elucidated structure can be found in figure 4.10. Detailed NMR results can be found in the Appendix 3. b) MS² spectrum annotated with elemental composition of compound **2** (m/z 504 $[M+H]^+$) and its dimer at m/z 1007 $[M+H]^+$. The x-axis shows Intensity (counts), while the y-axis shows Observed mass [m/z]. Precursor ion is annotated with $[M+H]^+$. Product ion elemental compositions annotated with "*" are uncertain.*

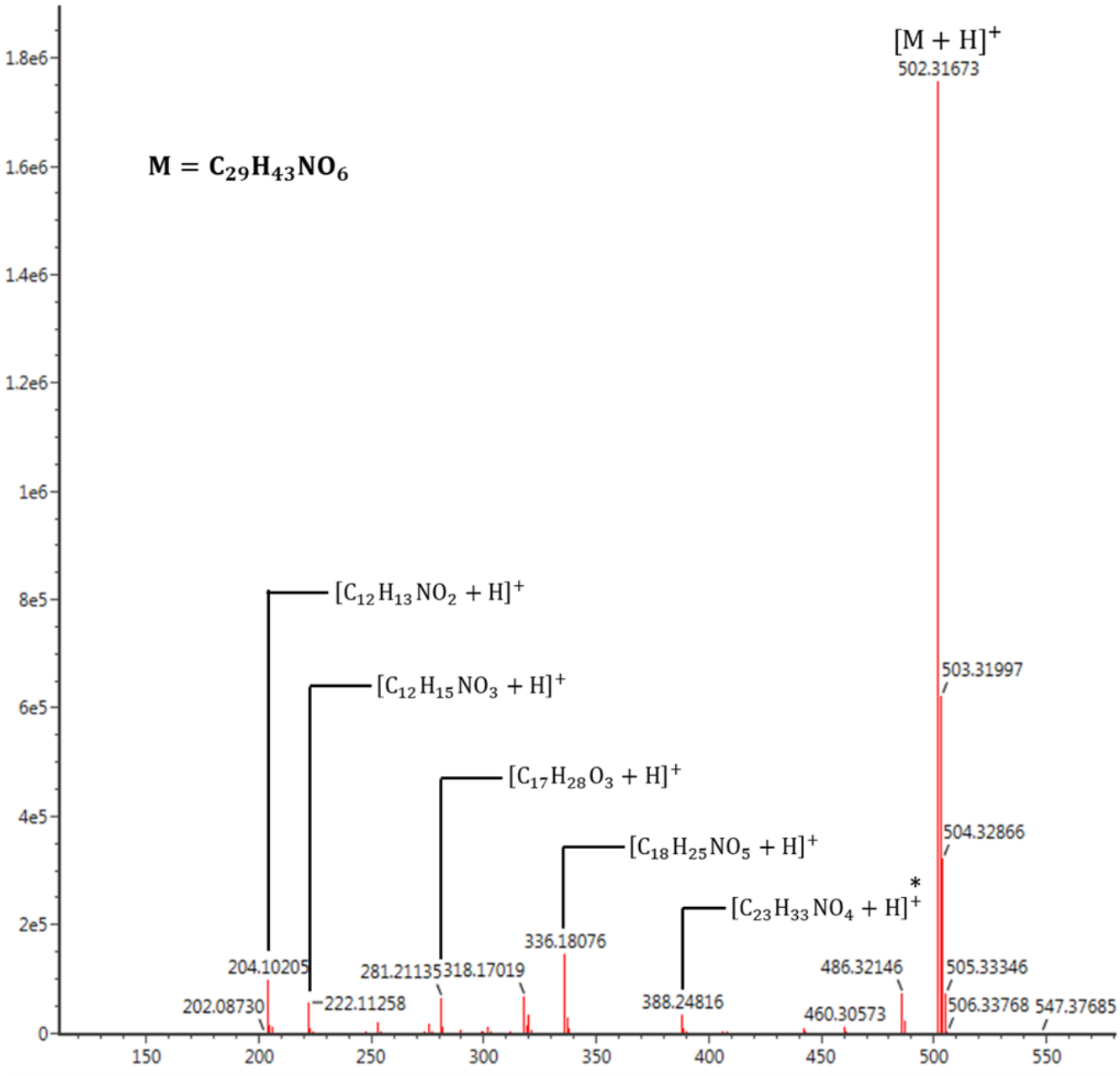


Figure 4.9. Elemental composition and MS² fragment ion pattern of compound 3 (m/z 502 $[M + H]^+$). The x-axis shows Intensity (counts), while the y-axis shows Observed mass [m/z]. Precursor ion is annotated with $[M + H]^+$. Fragment elemental composition annotated with "" is uncertain.*

Compound **2** had an observed mass of *m/z* 504.33 $[M + H]^+$ corresponding to an elemental composition of $C_{29}H_{45}NO_6$ ($+CH_2$ of hapalosin 489 (**1**)), while compound 3 observed at *m/z* 502.32 matched a composition of $C_{29}H_{43}NO_6$ ($+C$ of hapalosin 489(**1**)). Major product ions of compound **2** were *m/z* 486.32 $[C_{29}H_{43}NO_5]^{+}$*, *m/z* 320.19 $[C_{18}H_{25}NO_4 + H]^+$, *m/z* 302.18 $[C_{18}H_{23}NO_3 + H]^+$, *m/z* 281.21 $[C_{17}H_{28}O_3 + H]^+$, *m/z* 206.12 $[C_{12}H_{15}NO_2 + H]^+$. Major product ions of compound 3 were *m/z* 388.25 $[C_{23}H_{33}NO_4 + H]^{+}$*, *m/z* 336.18 $[C_{18}H_{25}NO_5 + H]^+$, *m/z* 281.21 $[C_{17}H_{28}O_3 + H]^+$, *m/z* 222.11 $[C_{12}H_{15}NO_3 + H]^+$, *m/z* 204.10 $[C_{12}H_{13}NO_2 + H]^+$.

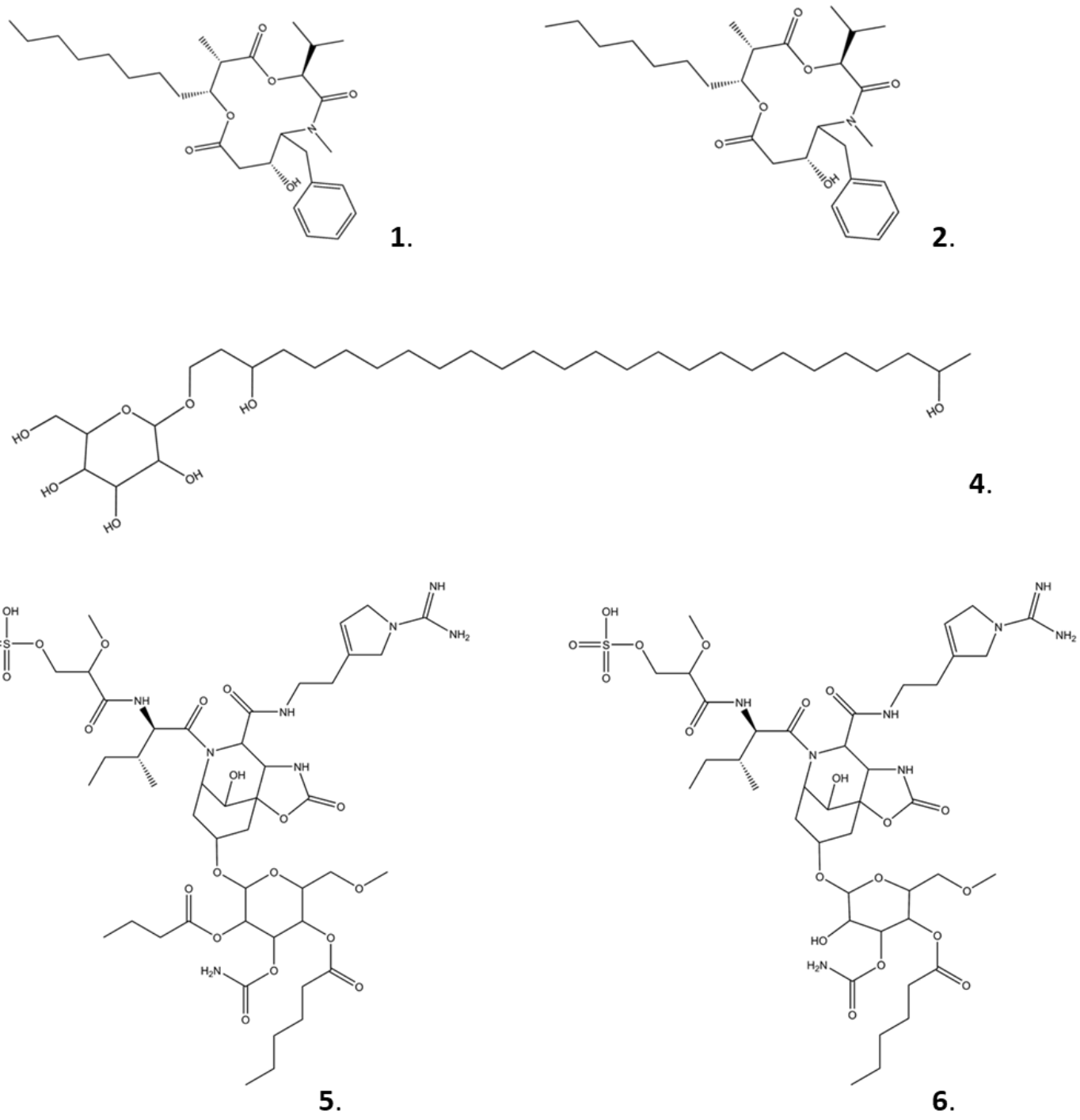


Figure 4.10. Structures of dereplicated compounds in the crude extracts of KVJ10 and KVJ20. ***1.*** *Hapalosin 489,* ***2****. Hapalosin 503,* **4**. 1-(*O-hexose-3,25-hexacosanediol)(heterocyst glycolipid at m/z 577),* ***5****. Suomilide B,* ***6****. Suomilide 1006. Details about each compound is found in table 4.4.*

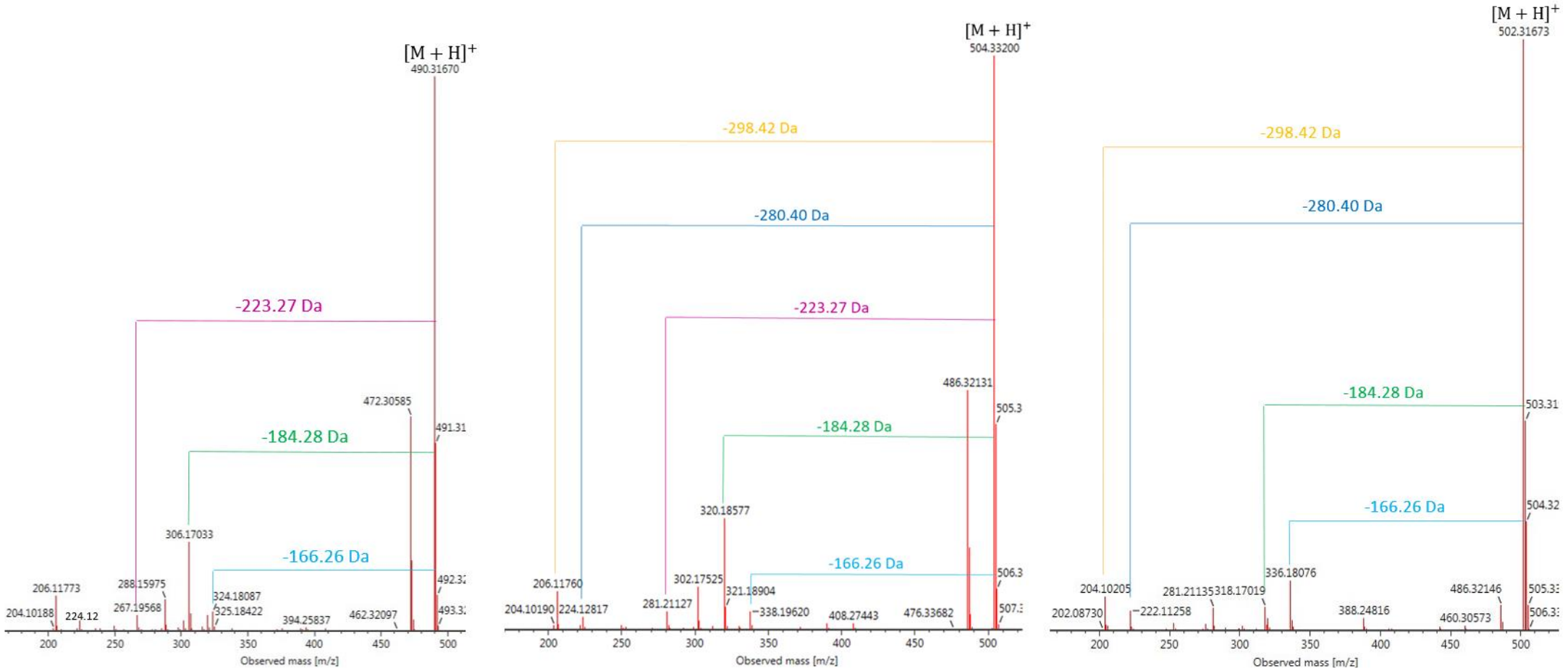


*Figure 4.11. MS² of compound **1**, **2** and 3. Colored brackets indicate neutral losses in Da that are common for two or more of the compounds.*

4.4 Yield from FLASH-chromatography fractionation

Table 4.5. Dry weight yield from extract fractionation by FLASH-chromatography.

Fraction	Solvent(s)	KVJ10	KVJ10
		Dry weight (mg)	Dry weight (mg)
1	95% pH_2O + 5% MeOH	22.3	22.6
2	75% pH_2O + 25% MeOH	24.1	77.8
3	50% pH_2O + 50% MeOH	3.0	11.6
4	25% pH_2O + 75% MeOH	1.3	11.7
5	100% MeOH	15.3	21.5
6	25% MeOH + 75% Acetone	16.9	58.8

The various gradients of the mobile phase reagents that were used during the fractionation chromatography resulted in fractions that contained compounds with different polarity, with fraction 1 having the most polar compounds with the subsequent fractions having increasing number of non-polar molecules with fraction 6 having the most non-polar ones, *i.e.*, if there were lipids present in the extracts, these will likely be fractioned into fraction 5 and 6.

The dry weight yield of the fractions varied from 1.3 to 24.1 mg in KVJ10 and from 11.6 to 58.8 mg in KVJ20. For both KVJ10 and KVJ20, fractions 3 and 4 were the ones with the lowest dry wight yield, and fraction 2 were the fraction with the highest. Fraction 5 was had the third lowest yield for both strains, while 1 and 6 were either the second or third highest depending on the strain.

4.5 Anti-bacterial assay

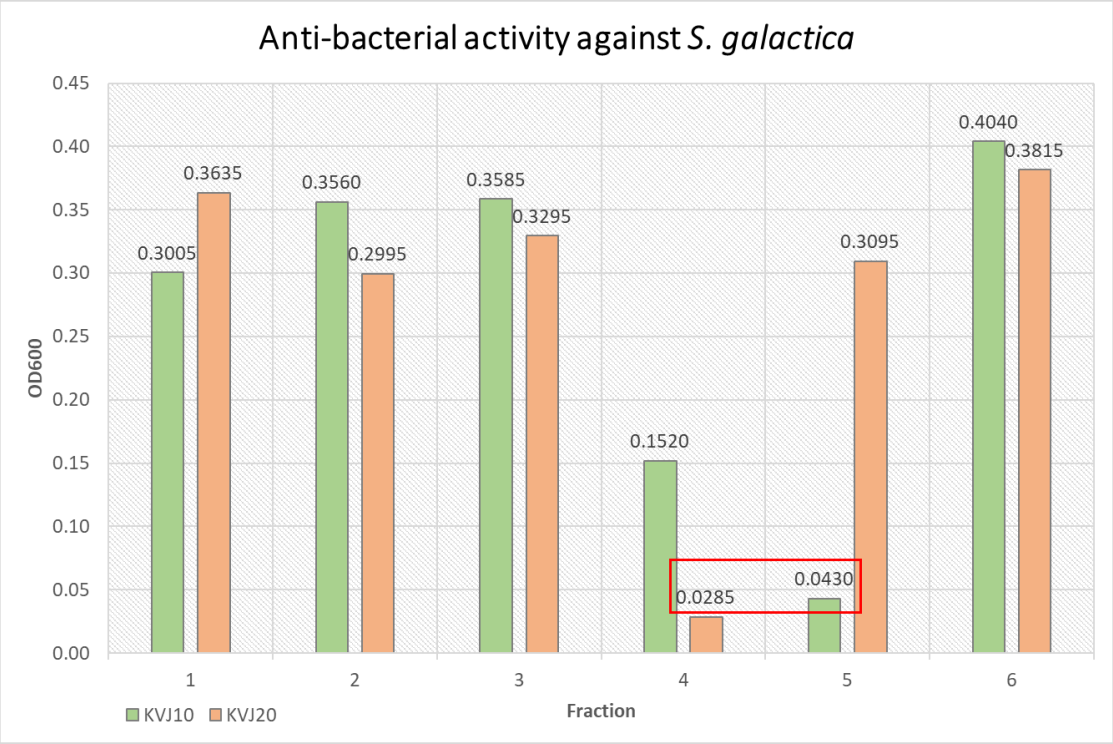


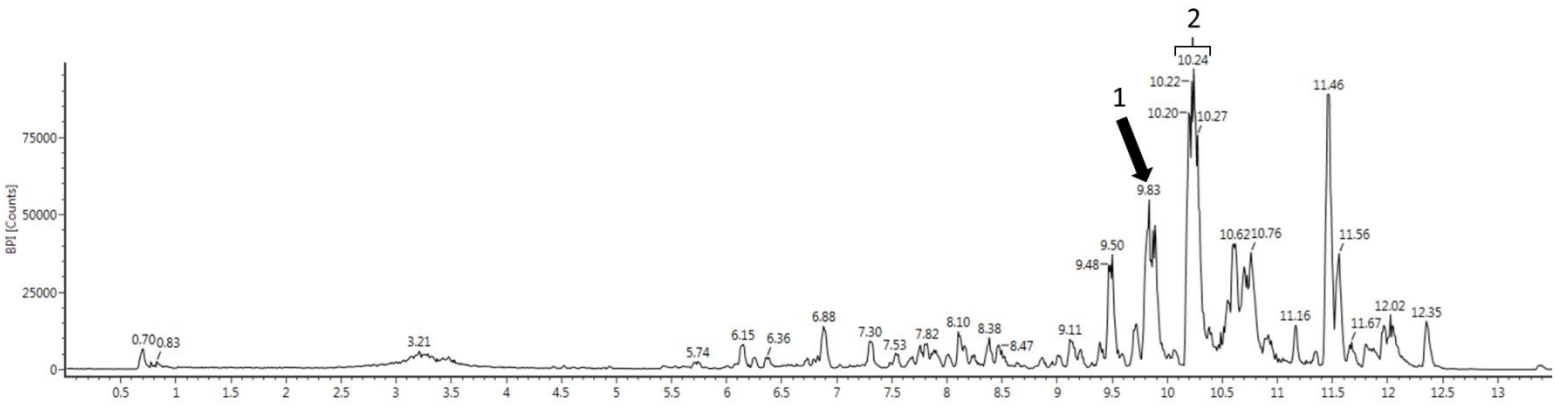
Figure 4.12. Growth of S. agalactiae (OD600) in media w/KVJ10 or KVJ20 fractions 1-6 with red box enclosing KVJ20 fraction 4 and KVJ10 fraction 5 that have OD600 values that classified them as "active".

Fraction 5 of KVJ10 and fraction 4 of KVJ20 showed anti-bacterial activity against *Streptococcus agalactia* with an average OD_{600} of 0.043 and 0.029, respectively (figure 4.11). In comparison, the media blank had an average of 0.035 and the growth control an average of 0.310 (not shown). None of the other fractions that were tested were recognized as "active" or "questionable" against *S. agalactia* or any of the other bacteria.

4.6 UHPLC-ESI-MS² of flash fractions

4.6.1 Signals of potential bioactive compounds

Prominent peaks that are not annotated are unknown compounds that have been classified as uninteresting and are most likely not responsible for the observed bioactivity. For example, if an unknown molecule had lipid-like characteristics it was not further investigated because lipids have in case of bioactivity mostly interspecific bioactivity and thus do not have bioactive properties that are of interest.

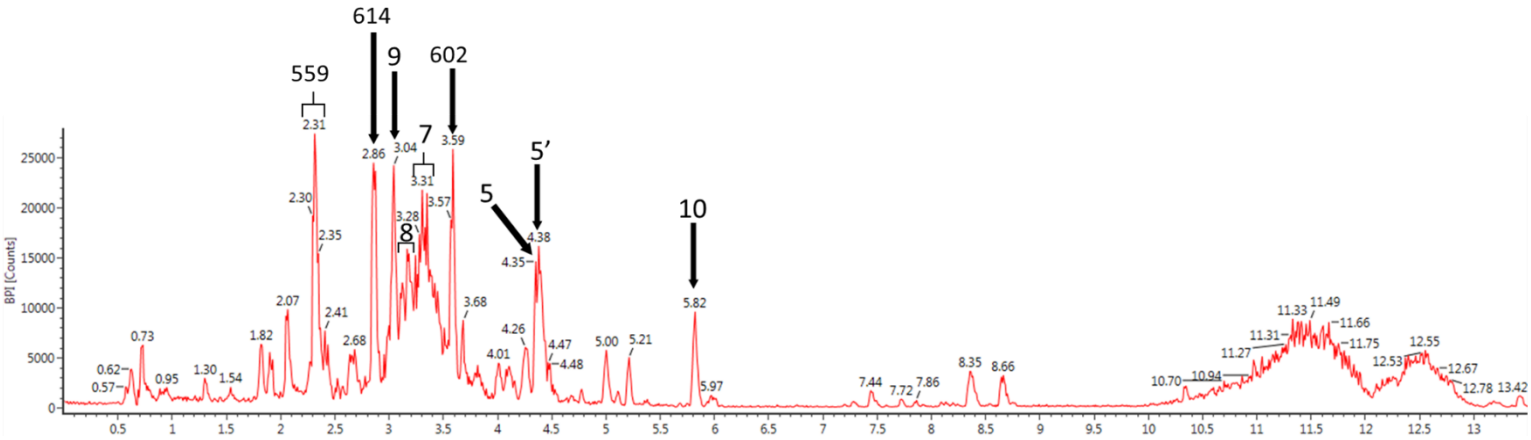


Intensity

Peak number	m/z (M+H)	RT	Suggested composition	Suggested compound	Suggestion based on:
2	504.33	10.0-11.5	$C_{29}H_{45}NO_6$	Hapalosin 503 (**2**)	MS² fragments compared with hapalosin 489 (1), NMR
1	490.22	9.5-10.5	$C_{28}H_{43}NO_6$	Hapalosin 489 (**1**)	Library match + MS² fragment pattern match (Stratman et al., 1994)

Figure 4.13. BPI chromatogram of fraction 5 of KVJ10 that showed bioactivity against S. agalactia. The x-axis represents retention time in minutes, while the y-axis shows base peak intensity. Peaks of observed masses that were deemed likely for being responsible for the observed bioactivity are annotated and detailed in the supplied table.

Two prominent peaks in the bioactive fraction of KVJ10 was identified as compound **1** and **2**, respectively. Other major peaks were deemed unlikely to be responsible for the observed bioactivity. Compounds **1** and **2** had both been detected in KVJ10 crude extracts and appeared as major peaks in the EtOAc chromatogram (figure 4.4). The major peak observed at RT 11.46 was a compound that was also observed in the "blank" (80% MeOH) injection.



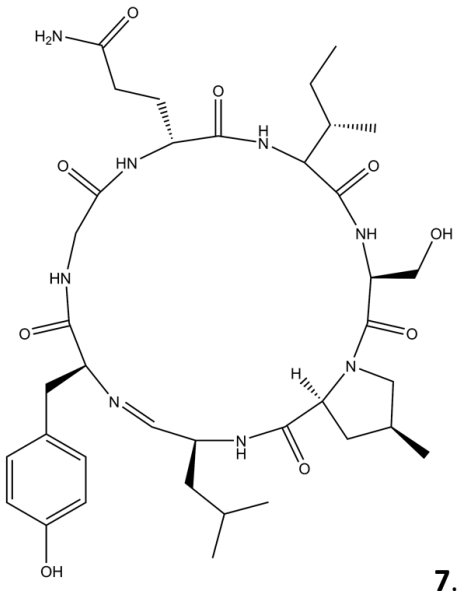
Peak number	m/z (M+H)	RT	Suggested composition	Suggested compound	Suggestion based on:
7	827.43 (809)	3.0-4.0	$C_{40}H_{56}N_8O_{10}$	Nostocyclopeptide A2 (**7**)	Fragment pattern match (Fidor et al., 2020)
8	775.44	3.0-3.5	$C_{37}H_{58}N_8O_{10}$	Nostocyclopeptide A1 (**8**)	Fragment pattern match (Fidor et al., 2020)
5 (5')	1077.47 (997)	4.0-4.5	$C_{45}H_{72}N_8O_{20}S$	Suomilide B (**5**)	Fragment pattern match (Riba et al., 2020)
9	777.42	3.0-3.5	$C_{39}H_{52}N_8O_9$	Nostocyclopeptide E1	Fragment pattern match (Fidor et al., 2020)
614	614.39	2.5-3.0	$C_{28}H_{51}N_7O_8$	Unknown peptide	Library match
602	602.56	3.5-4.0	n/a	Unknown compound	n/a
559	559.96	2.0-2.5	n/a	Unknown compound	n/a
10	792.47	5.5-6.0	$C_{42}H_{61}N_7O_8$	Schizopeptin 791	Fragment pattern match (Schneider, 2020)

Intensity

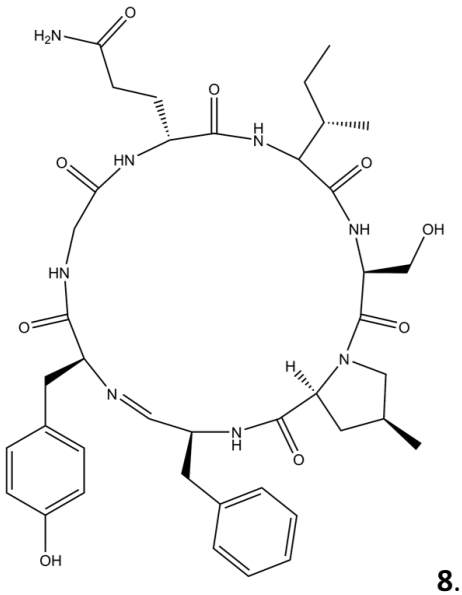
Figure 4.14. BPI chromatogram of fraction 4 of KVJ20 during that showed bioactivity. The x-axis represents retention time in minutes, while the y-axis shows base peak intensity. Major peaks that were deemed likely be responsible for the observed bioactivity are annotated and detailed in the supplied table.

The active fraction of KVJ20 had many major peaks towards the polar end of the chromatogram. Several of which remains unidentified as no satisfactory library matches were made and in addition were not detected in any crude extracts of KVJ20, this includes compounds with peak number 614, 602, and 559 in figure 4.13. Some were also detected in crude extracts of KVJ20, this includes compound with peak number **7**, **8**, **5** and 5'. Desulphated suomilide B (5') was detected as a relatively prominent peak together with sulphated suomilide B (**5**) in the fraction, while in the crude extracts of KVJ20, compound **5** was distinguishable as a peak while 5' was detected but not as a separable peak in the chromatogram.

Two identified compounds that was not detected in any crude extracts was nostocyclopeptide E1 (compound 9) and schizopeptin 791 (compound 10).



7.



8.

*Figure 4.15. Structure of compound **7** and **8** that was dereplicated from the active fraction of KVJ20. More details about the compounds can be found in figure 4.14.*

4.7 GNPS

GNPS predicted in total eight different compounds, including two β-carotenes, two cholane steroids, an acylglycerol, a galactolipid, one heterocyst glycolipid, and pheophytin *a*, from the crude extracts and fractions that were analyzed using UHPLC-MS. Masses that gave library hits in GNPS but were not observed as peaks in the chromatograms were not given attention or explored further.

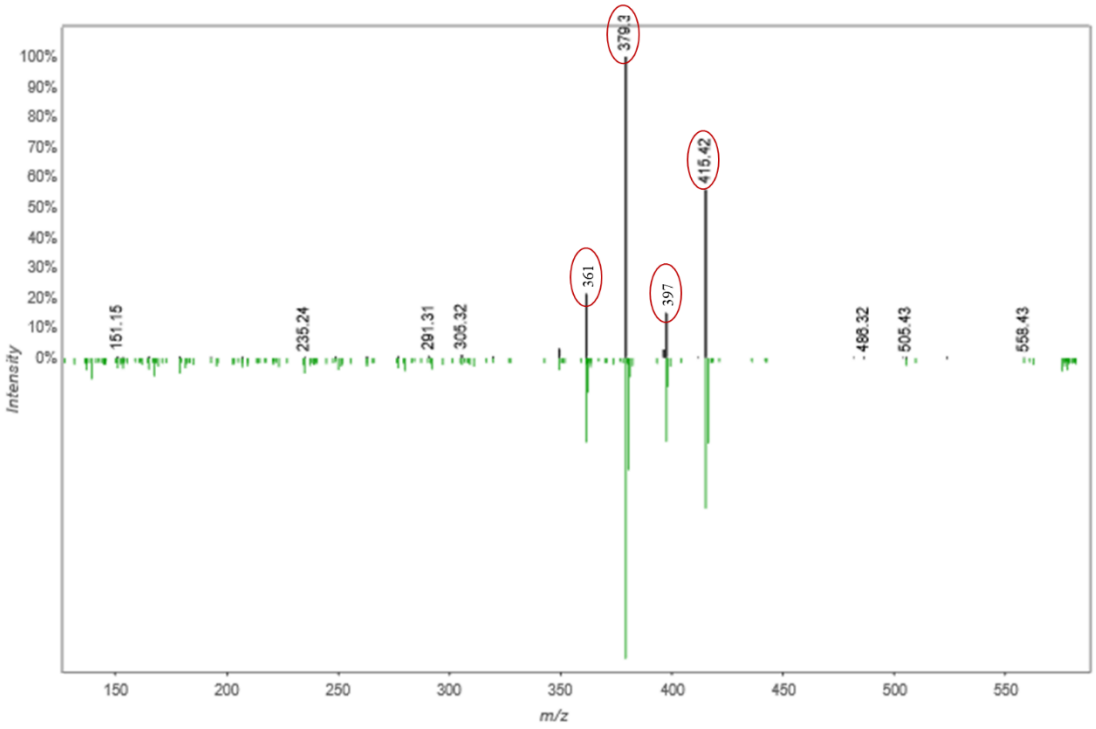


*Figure 4.16. Raw spectrum (upwards, black bars) of the heterocyst glycolipid predicted with GNPS (downwards, green bars). Masses of product ions that are similar to product ions of the MS² spectrum (figure 4.6) of compound **4** are encircled.*

From the inputted MS² spectral data, GNPS generated in total 31 molecular networks. Molecular networks that only contained spectra with molecular ions that had not been given any attention earlier were not considered for further exploration and discussion. Only networks containing beforementioned compounds will be shown, and only networks containing more than one previously mentioned compound will be discussed further.

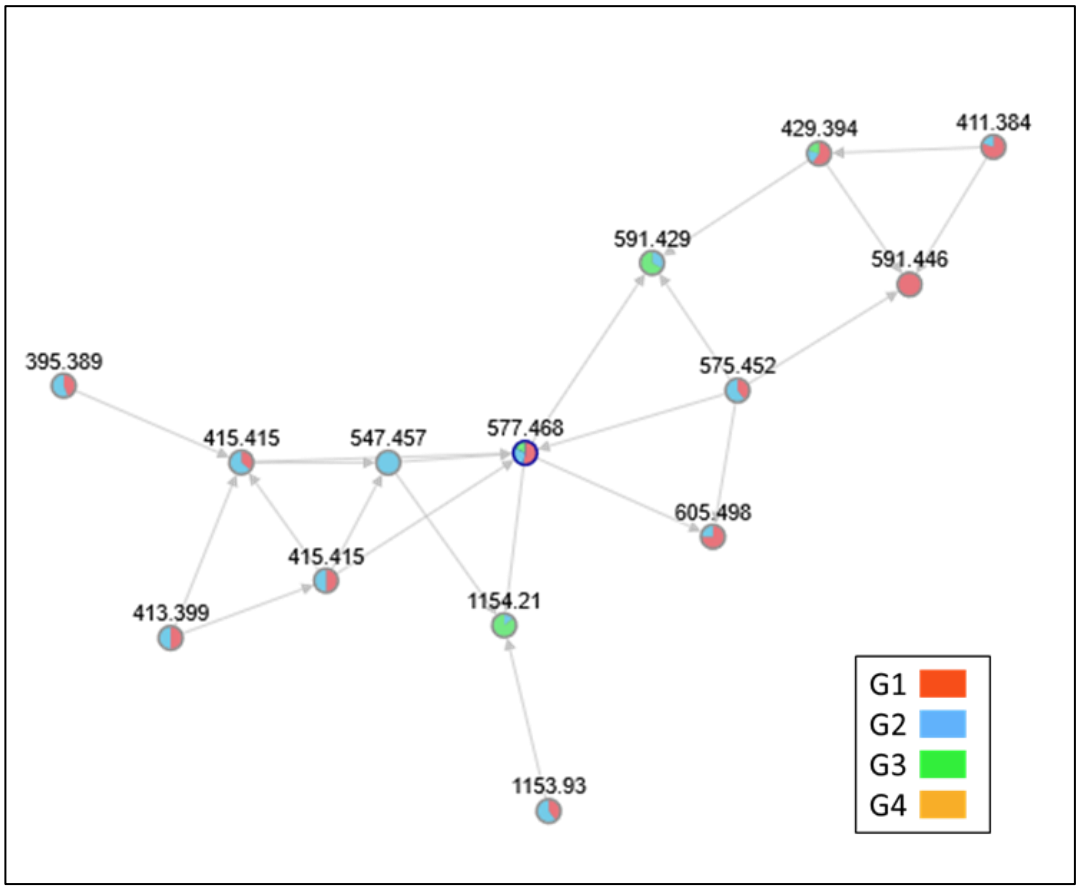


*Figure 4.17. Molecular network generated with GNPS of molecules related to compound **4** detected at 577 m/z $[M+H]^+$. Nodes represents each MS² spectrum of the network and are respectively annotated with the mass of the precursor ion. Edges represents spectrum-to-spectrum alignments. Node coloration shows the origin of the set of spectra; G1: KVJ10 crude extracts, G2: KVJ20 crude extracts, G3: KVJ10 fraction 4-6, KVJ20 fraction 3-5.*

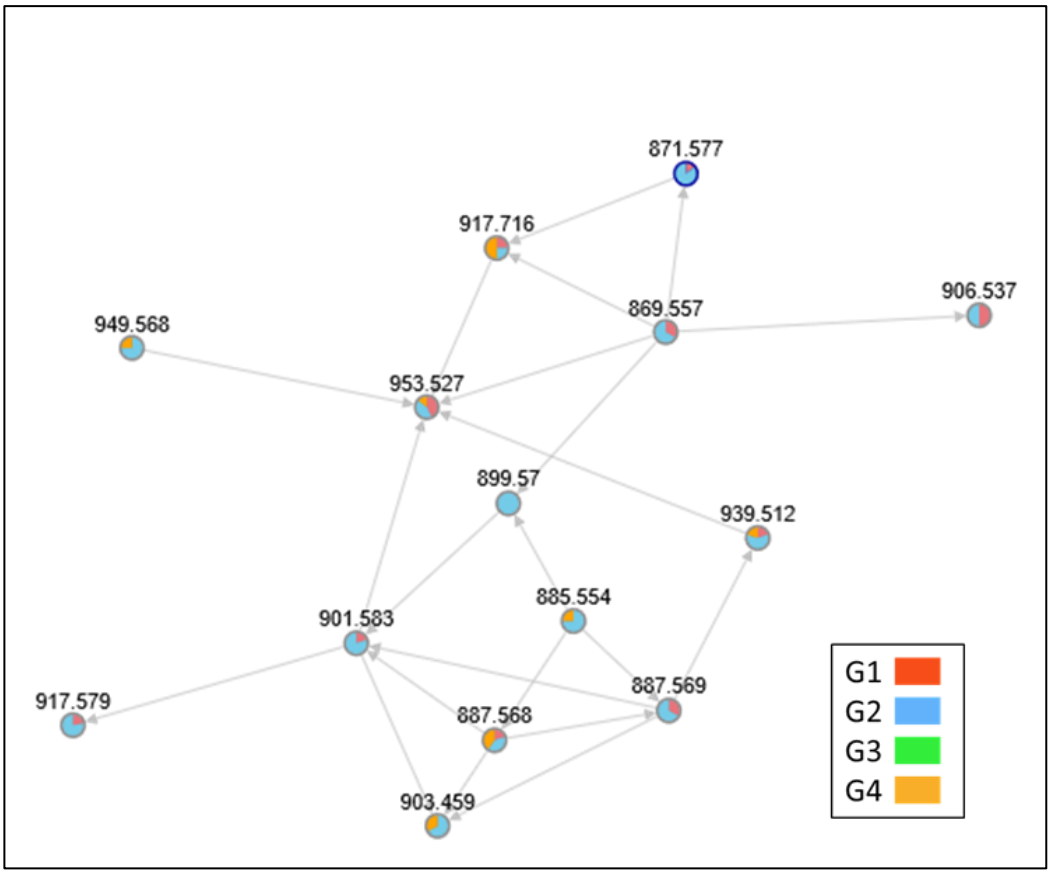


Figure 4.18. Molecular network generated with GNPS of molecules related to pheophytin a at 871 m/z $[M+H]^+$*. Nodes represents each MS² spectrum of the network and are annotated with the mass of the molecular ion. Edges represents spectrum-to-spectrum alignments. Node coloration shows the origin of the set of spectra; G1: KVJ10 crude extracts, G2: KVJ20 crude extracts, G3: KVJ10 fraction 4-6, KVJ20 fraction 3-5.*

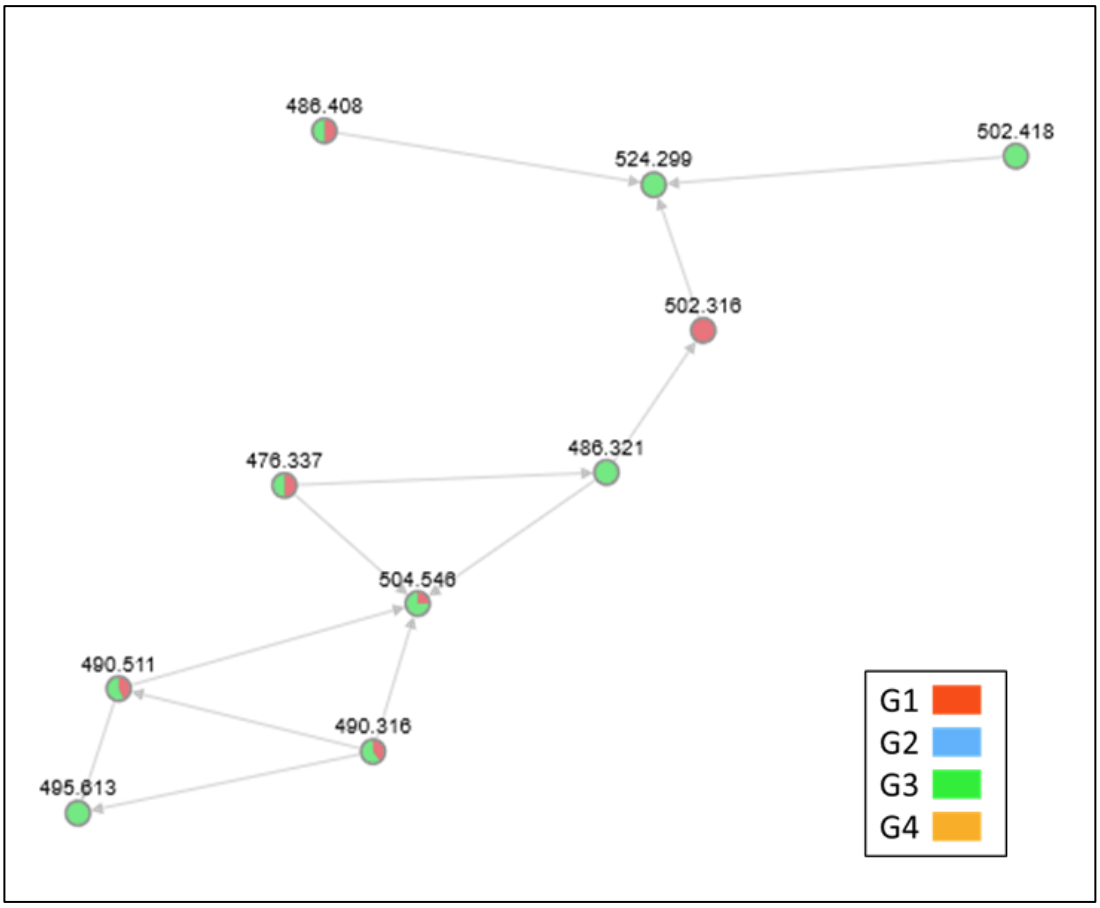


*Figure 4.19. Molecular network generated with GNPS of molecules related to compound **1**, **2**, and 3, detected at m/z 490, 504, and 502 $[M + H]^+$ respectively. Nodes represents each MS² spectrum of the network and are annotated with the mass of the molecular ion. Edges represents spectrum-to-spectrum alignments. Node coloration shows the origin of the set of spectra; G1: KVJ10 crude extracts, G2: KVJ20 crude extracts, G3: KVJ10 fraction 4-6, KVJ20 fraction 3-5.*

The GNPS predicted molecular network which consist of compound **1**-3, shows that there are in total ten related molecules in the network, all of which originates from KVJ10.

5 Discussion

5.1 Efficiency of solvents for cell extraction

When extracting secondary metabolites from bacteria for subsequent HPLC-MS analysis, there are a number of different methods and solvents that can be used. For bacterial natural products methanol, water and ethyl acetate are widely employed solvents that have can be used for liquid extraction, that are practical to use and that are relatively harmless for the environment (Haque et al., 2017; Rao et al., 2007).

Ethyl acetate is a non-polar solvent that extract lipophilic compounds. Ethyl acetate extracts have been shown to be especially suitable for further purification because they contain less media components and unwanted polar metabolites (Schneider et al., 2021). Water is the most polar solvent that is used in natural product extraction and extract hydrophilic compounds which compared to methanol, is less likely to contain contaminants (Haque et al., 2017). Methanol is a polar solvent that is highly effective in extraction and are broadly used in water-methanol mixes for metabolite extraction because of the wide range of molecules that can be extracted, however, methanol extracts often contain contaminants that appear as "noise" on chromatograms.

The attempted isolation of hapalosin 501 did not produce a sufficient yield of the compound for NMR analysis. However, experience with cell extraction using ethyl acetate do not recommend additional rounds of extraction, *i.e.*, there have been no increased metabolite yield from prolonging or re-extracting the cells with additional ethyl acetate, thus re-isolations of hapalosin 501 should be done in order to obtain more compound.

5.2 AntiSMASH predictions and dereplicated compounds

Many of the predicted biosynthetic gene clusters did not appear to be expressed during this project. One explanation is that the gene clusters were "silent" and thus not expressed under the cultivation conditions used in this project (Bloor & England, 1989); that the conditions did not stimulate sufficient production for the compounds to be detected during UHPLC-MS, or that the preparation or analysis method used degraded or changed the molecule(s). Another part of the explanation could be that antiSMASH have done wrongful predictions that were not included in the revise-process of the BGCs. The beforementioned work of piecing together gene clusters did give somewhat different results than what was predicted by

antiSMASH (see figure 4.1 in *Results*) and considering the amount of BGC regions predicted, it is likely that this is the case for several of the BGCs. Additionally, the genome sequence of KVJ10 was done on samples that were not axenic, and are indeed deposited as metagenome, thus some of the BGCs predicted could've originated from other organisms' genomes and not a representation of KVJ10 genes.

Suomilide B (**5)** and 1006 (**6**) production was not predicted by antiSMASH but was identified in KVJ20 extracts. No suomilides were detected in KVJ10.

One of the main differences predicted by antiSMASH between KVJ10 and KVJ20 biosynthesis was the production of hapalosin in KVJ10 and the production of nostocyclopeptide in KVJ20. This projection fitted well with the chemical analyses done during this project. There was no trace of any variation of hapalosin (**1**, **2**, or 501) in KVJ20 and KVJ10 did not show any detectable production of nostocyclopeptides A1 (**8**) and A2 (**7**).

5.2.1 Prediction of hapalosin production

Two biosynthetic gene clusters predicted the production of hapalosin, one with 100% similarity and the other with 60% similarity. After closer inspection, the latter was deemed unlikely to be a hapalosin operon for any hapalosin congeners as the percent identities between the predicted and the reference hapalosin BGC genes ranged from 48 to 63%, which is considered too low. Also, by visually comparing the query and reference sequence, there was clearly too many dissimilarities for the predicted BGC for it to produce hapalosin.

5.3 UHPLC-MS² analysis of KVJ10 and KVJ20 crude extracts

The hapalosins (**1**, **2**, and 501) seemed to be most efficiently extracted from the KVJ10 cells using ethyl acetate. Even though the EtOAc extract of KVJ10 (and KVJ20) had the lowest dry weight extraction yield, the intensity of the hapalosins were highest in the KVJ10 EtOAc chromatogram than in any other KVJ10 extracts that also contained hapalosin. However, because UHPLC-MS is not a quantitative method, the intensity of peaks should not be considered a measurement of the concentration of metabolite in the sample, but the high intensity could mean that the hapalosins are more stable and less subjected to ionization in ethyl acetate compared to methanol and water. Nevertheless, the presence of the *m*/*z* 1007

$[M + H]^+$ proposed dimer of hapalosin 503, could indicate that this hapalosin variant was highly concentrated in the extract (Steckel & Schlosser, 2019).

There were also fewer base peaks of low intensity in the EtOAc chromatograms. The peaks that appeared were of relatively high intensity and less "cluttered" than the other extraction methods, especially in KVJ10. Importantly, the response or ion-count of a respective analyte does not necessarily correlate quantitatively with the presence of the compound within a sample because different compounds have different ionization efficiencies. The discount of minor signals is a compromise in order to handle the high number of signals.

While pheophytin *a* was represented as a major peak in KVJ20 EtOAc extract, it did not produce an observable peak in KVJ10 EtOAc, instead, it was a major peak in the 50% MeOH extract of KVJ10 as well as a notable peak in 50% MeOH of KVJ20. As pheophytin *a* was found in both KVJ20 and KVJ20, it would be expected that the same extraction method would recover similar quantities of the compound, since this is not the case it could indicate that the KVJ10 cells in the EtOAc released their cell content to a lesser degree than the KVJ20 cells. This could be due to variations between the strains, for example in pH causing the KVJ10 EtOAc extract to be more acidic than the EtOAc extract of KVJ20, or because something was done slightly different between the EtOAc extraction process of the strains by accident. However, it needs to be taken into account that UHPLC-ESI-MS is not giving reliable quantitative analyses without a reference-sample and reference runs.

5.3.1 Relatedness of compound 1, 2, and 3

As shown in figure 4.10 there are several neutral losses in common between compound **1**, **2**, and 3. In addition, compound **2** shared two product ions with hapalosin 489 at *m*/*z* 224.13 $[C_{12}H_{15}NO_3]^+$ and *m*/*z* 206.11 $[C_{12}H_{15}NO_2]^+$, and one product ion with compound 3 at *m*/*z* 281.21 $[C_{17}H_{28}O_3]^+$. Compound 3 shared additionally one product ion with hapalosin 489 at *m*/*z* 222.11 $[C_{12}H_{15}NO_3]^+$. These common neutral losses and product ions, together with the similarity in mass and retention times, is strong evidence for the relatedness of the three compounds. Moreover, the NMR elucidation of compound **2**, shows a structure that is very similar to that of hapalosin 489.

5.4 Structure of hapalosin 501 and 503

The suggestion for the structure of compound **2** was first made based on the calculated elemental composition for the precursor ion and product ions from the MS analysis and MS^2 spectrum. The structure was then confirmed through the isolation of the pure compound followed by NMR analysis (results from NMR can be found in Appendix 3) which indicated that the additional methylene group ($+CH_2$) was located on the hydrocarbon chain (see figure 4.10 with structure of compound **2**).

Based on the elemental composition, fragment ions, and elucidated structure of hapalosin 503, the structure of hapalosin 501 was theorized to be similar, but with a C=C bond somewhere on the ring structure which would explain the two hydrogens in difference between the two variants.

Figueiredo et al., (2021) suggested, based on LC-HRESI (High Resolution Electrospray Ionization)-MS^2 analysis that the hapalosin variant they observed at *m*/*z* 504.33 $[M + H]^+$ had the methylene on the $CH_3-CH_2R-CH_3$ chain connected to the ring structure of the molecule. This was however in conflict with the structure provided by NMR analysis done during this project.

5.5 Bioactive fractions

Both KVJ10 and KVJ20 had fractions that showed anti-bacterial activity against *Streptococcus agalactia*. It is unlikely that it is the same compound that is responsible for bioactivity in KVJ20 and KVJ10. Both because there were two different fractions that showed activity meaning they were fractionized using different solvent ratios of different polarity and because there is no common major base peak with similar observed mass. The major peaks in the active fraction of KVJ10 is towards the non-polar area of the range, while in KVJ20 the major peaks lie in the polar range of the chromatogram. It can also be assumed that a major base peak that is shared with an inactive fraction is not responsible for the bioactivity. There is a possibility that there is a combination of compounds that is responsible for the observed bioactivity through synergistic effects. This can be investigated through biotesting of pure compounds to see if the same activity is observed then.

5.5.1 Bioactivity in KVJ10 fraction

Hapalosin 489 and especially hapalosin 503 were detected as major peaks in the active fraction of KVJ10 (fraction 5). This could mean that hapalosin is responsible for the anti-bacterial effect on *P. agalactica*. Cyanobacteria have for some time been known to produce a wide variety of antibacterial compounds, but as far as we known, this is the first time antibacterial activity of a hapalosin have been indicated, and anti-bacterial properties of cyclic depsipeptides have not been widely discussed. Earlier investigations on the molecular family (cyclic depsipeptides) of which hapalosin belongs, have predominantly shown protease inhibiting or cytotoxic properties (*e.g*., Hamada & Shioiri, 2005; Luesch et al., 2002; Mevers et al., 2011; Weckesser et al., 1996; Williams et al., 2003) and very few anti-bacterial activities, although a trait more often found in other types of cyclic peptides and lipopeptides (*e.g*., Carpine & Sieber, 2021; Pergament & Carmeli, 1994; Rojas et al., 2020; Zainuddin et al., 2009). However, because it was not isolates of pure hapalosin variants that was assayed, one cannot say for certain that the anti-bacterial activity comes from hapalosin, any of its variants, or a collaborative effect of them. Biotesting of the isolated compounds is needed. Because of the reported MDR-reversing potency of hapalosin 489 from earlier research, it would be of great interest to test pure samples of hapalosin 501 and 503, individually, for similar activity.

5.5.2 Bioactivity in KVJ20 fraction

The compounds that are likely for being responsible for the observed anti-bacterial activity in KVJ20 is shown in figure 4.13 under *Results*. However, because suomilide B and 1006 have in a previous study by Schneider (2020) been shown to have no growth inhibiting effect on *S. agalactia* it can be deemed highly unlikely that suomilide B is accountable for the observed activity in the KVJ20 fraction during this project. It should also be noted that nostocyclopeptide A1 and A2 have earlier been shown to have no antibacterial activity when tested against for example *Pseudomonas aeruginosa* and *Mycobacterium turberculosis* strains (Golakoti et al., 2001). Nostocyclopeptide A2 (Ncp-A2) was largely detected in the bioactive fraction of KVJ20 at 827 *m*/*z* $[M + H]^+$, a mass that does not match that of Ncp-A2 according to the findings of Fidor and colleagues (2020), however, the MS^2 pattern of observed mass 827 was identical to that of linear Ncp-A2 which was observed at 809 *m*/*z,* at high intensity in fraction 4 of KVJ20 as well, and matched in mass and fragmentation pattern to that of Fidor et al.

Nostocyclopeptide E1 have not yet been tested for anti-bacterial or any other inhibiting properties and could be responsible for the observed inhibition of bacterial growth. A closely related molecule cyclic nostocyclopeptide E2 was tested by Fidor et al., (2021) against human 20S proteasome but no effects were observed. Schizopeptin 791 have been shown to have protease inhibiting activity which could affect microorganisms (Reshef & Carmeli, 2002).

Several compounds, like Schizopeptin 791, the unknown compounds 602 and 559 were detected in flash fractions but not in any crude KVJ20 (or KVJ10) extracts, this could be because that in the crude extracts schizopeptin 791 was under the detection limit, the compound was poorly dissolved in the extracts, or there were resolution issues during HPLC-MS. The first one seemed like a probable issue for the crude extracts as there was several compounds with high intensity in the fractions of KVJ10 and KVJ20 that was not detected in any of the crude extracts. There is also the possibility that these molecules do not originate from the cyanobacteria extracts but are so-called “artifact” particles that are a result of the sample preparation or analysis process. However, this would not explain the positive identification of schizopeptin 791 in the active fraction of KVJ20 and the lack of it in all crude extracts.

The unknown peptide 614 was found in both crude extracts and in all the fractions of KVJ20 but at highest intensity in the bioactive fraction. Library search and HR-MS gave a suggestion for elemental composition (figure 4.14) and type of molecule, but further investigation and literature search did not provide any other characterizations.

5.6 GNPS

GNPS can be used for several useful purposes in metabolomic studies. The three functions most relevant for this project was dereplication, molecular networking, and comparison of MS^2 data between samples. GNPS can be used in dereplication by matching obtained MS^2 data with spectral libraries. GNPS also use the inputted MS^2 data to create molecular networks that display relatedness between compounds, that are not easily detected by visual inspection of individual MS^2 spectrum, even if the compounds have no library hits. Another useful function is the comparison of compounds across samples, in which GNPS recognizes if a compound's MS^2 spectrum is present in one or more samples. This makes it possible to effortlessly prospect samples for specific molecules as this feature can be visualized in the generated networks.

5.6.1 Dereplication using GNPS

The identity of the mass observed at *m*/*z* 871 $[M + H]^+$ was provided by GNPS as pheophytin *a*. Initial library search through the UNIFI software only produced an elemental composition with no compound specification given. The identity provided by GNPS was confirmed through comparing the product ions in the MS^2 spectrum of pheophytin *a* from literature (Zvezdanović et al, 2014) with the MS^2 spectrum obtained of *m*/*z* 871 during the UHPLC-ESI-MS^2. GNPS also correctly identified the previously dereplicated heterocyst glycolipid with observed mass *m*/*z* 577 $[M + H]^+$. However, GNPS suggested a cholane steroid for *m/z 1007* which had already been deemed to most likely be a dimer of compound **2** (see figure 4.8 b under *Results*). When looking at the raw MS^2 spectrum (figure 5.1) used to make this prediction, there are clearly product ions that match the MS^2 spectrum of compound **2**. The cholane steroid prediction from GNPS is likely wrong and this could perhaps be because hapalosin 489 and its variants are not registered in the GNPS database.

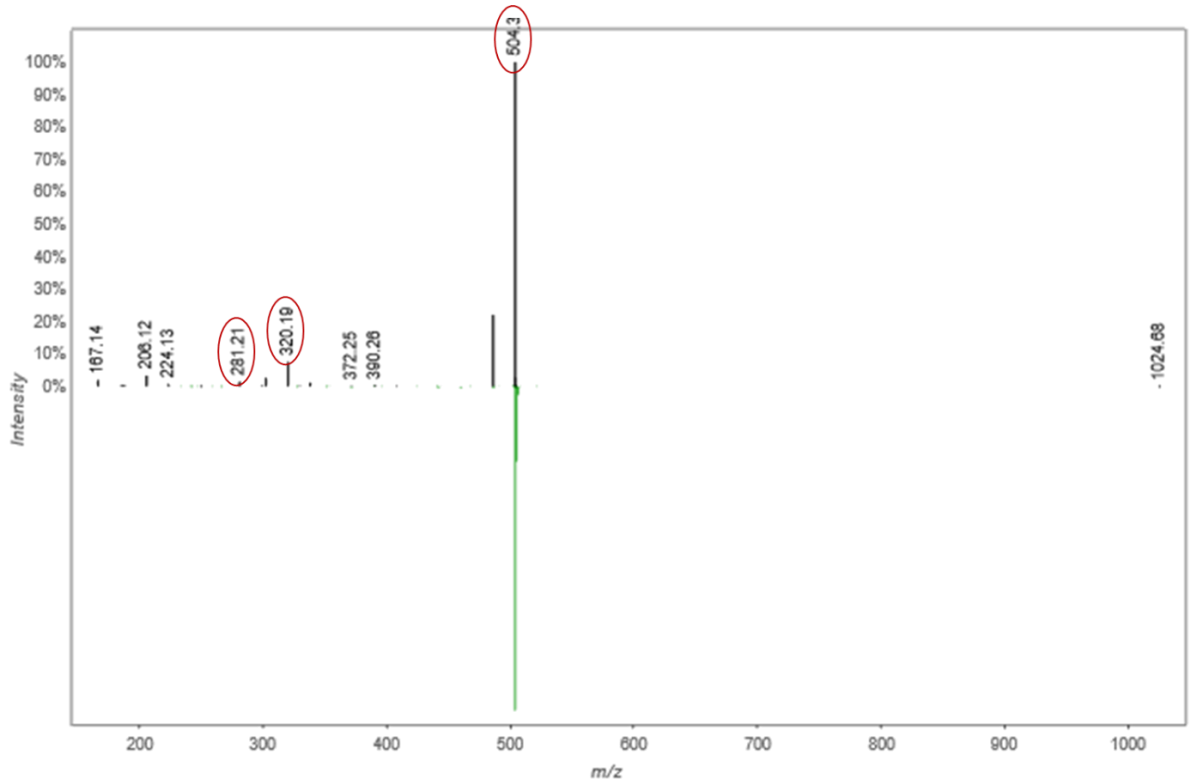


Figure 5.1. Raw spectrum (upwards, black bars) used for the library match of a cholane steroid (downwards, green bars) in GNPS. The mass of fragment ions that match that of hapalosin 503 is encircled.

Although GNPS can be used for automatic identification of compounds, the experience gained during this project of using GNPS indicates that there are few library hits for cyanobacterial or at least *Nostoc* compounds. Therefore, dereplication of *Nostoc* metabolites are still dependent on the visual comparison of obtained MS^2 spectra with reference spectra from literature. The accuracy of library matches predicted by GNPS depend on the correct identification and cataloguing of compounds registered in the libraries. Additions to the libraries can be made by the users who utilize various spectrometry methods and instrumentation. Therefore, the library MS^2 spectra do not only include spectra from for instance ESI+/- and if there is a lack of library entries the use of GNPS as a dereplication tool is limited.

5.6.2 Molecular networks generated with GNPS

The molecular networks showed relatedness of detected compounds from all the samples analyzed with UHPLC-MS². Several networks correctly displayed relatedness between known compounds and some known congeners, for example for suomilide 1006 (**6**) (not shown). But suomilide B was not found to be in any networks. It is difficult to say why this was the case, and it would've been interesting to see if GNPS would have put suomilide B in the same network as suomilide 1006. If there were any other compounds that were left out of the generated networks was not investigated closer, but most of the discussed compounds seemed to be present.

Even though GNPS did not provide a library match with hapalosin 489, the molecular networking of hapalosin 489 confirmed the relatedness of compound 1, 2, and 3. In addition, the network displayed several other related compounds to the hapalosins which could've been interesting to investigate further in a different study. A number of networks consisted of only unknown compounds and was therefore not further investigated in this study. In addition, time restriction prevented deeper exploration of the networks of interest as the use of GNPS was greatly delayed and not considered the main topic of this project.

The molecular networking of GNPS is a valuable tool for explore the relatedness of various compounds within and across samples. However, it was observed that for the nostocyclopeptides, there were several different networks generated (not shown) in which Ncp-A2 was put in a separate network from Ncp-A1 and Ncp-E1. Perhaps there is a logical explanation for this, but it was not examined further.

5.6.3 Comparing samples using GNPS

The molecular networking of hapalosin 489, its variants, and related molecules indicated that none of them originated from any KVJ20 extract or fraction, which is consistent to the chemical analyses. The same was observed for suomilide 1006. Unfortunately, this useful function was not as extensively used as it could have been due to the beforementioned time and topic limitations.

The spectrum files were set in groups depending on which strain they originated from and if the sample came from crude extracts or fractions, as described in section 3.10. By changing

the group set-up, it could've been possible to generate networks with nodes colored according to their presence in KVJ10, KVJ20 or both. For example, by organizing file group according to strain, *i.e.*, setting all KVJ10 extracts and fractions into a single group and all KVJ20 extracts and fractions into a second group, one could easily compare which compound MS^2 spectra was detected in which strain. As it was at times challenging and time-consuming to take note of which compounds were detected in either or both strains, this could have been a much quicker way of observing compounds which were unique to the respective strains. Other group set-ups to compare other aspects of the MS^2 data could also been done, for example grouping the extraction methods separately to see which compounds varied between the extraction method used.

6 Conclusion

6.1 Extraction methods

Using water, methanol, and ethyl acetate as solvents for metabolite extraction provided samples with a wide range of detectable metabolites that had various polarities and structure, and some of which were shown to be antagonistic against a bacterium (*S. agalactia*). Based on the findings during this project, I would suggest using ethyl acetate for the extraction of hapalosin and its variants from *Nostoc* sp. KVJ10 cells. Even if the dry weight yield was low, the intensities of hapalosin 489, 503, and 501 were highest in this extract. Additionally, this method gave the "cleanest" extract containing few other uninteresting metabolites and contaminants and was thus suitable for subsequent isolation of the pure compounds.

6.2 Prediction of secondary metabolite production

AntiSMASH was used to predict the biosynthetic potential of both strains with varying accuracy. Because of silent gene clusters it is hard to determine the accuracy of all predictions. The cluster prediction and similarity results have to be taken with caution when values are not suggesting an exact match.

6.3 Bioactive compounds

Both KVJ10 and KVJ20 contain one or more compounds that have anti-bacterial properties. Isolation of pure compounds and additional anti-bacterial testing of these must be done in order to find which one(s) is responsible for the observed activity. The chromatograms of the bioactive fractions of KVJ10 and KVJ20 gives strong indication that the compound(s) responsible for activity in KVJ10 is non-polar, while the bioactive compound(s) in KVJ20 is polar. It is likely that it is either hapalosin 489, hapalosin 503, or both in combination that is responsible for the activity observed in KVJ10. Of the possible bioactive compounds in KVJ20, several were unknown while the rest lacked documented anti-bacterial activity. It is therefore difficult to say more specifically about which compounds from KVJ20 that could be responsible for the inhibiting activity against *S. agalactia*.

6.4 Identification of compounds

Several compounds which were common and individually unique to KVJ10 and KVJ20 were identified through comparing UHPLC-ESI-MS^2 data with UNIFI library matches and literature. Some direct matches were made in the UNIFI software, including hapalosin 489, while others were matched through comparing their MS^2 spectra with literature, and one compound was identified through GNPS and subsequent comparison of MS^2 spectra. The comparison of acquired MS^2 data with reference spectra from previous studies is a time-consuming process and requires literature research for finding relevant cyanobacterial compound MS^2 spectra to be used as reference, but the identifications are trustworthy. Moreover, new variants of known compounds are often indicated through fragmentation patterns as was the case with hapalosin 503 and 501, but this requires knowledge and experience within such chemical analysis methods.

In this project, GNPS was used to predict the relatedness of molecules detected in KVJ10 and KVJ20 samples which was illustrated as molecular networks. Because there was limited time and it was not the main purpose of this project to confirm the accuracy of every network generated, it is difficult at this time to say how successful these predictions were. Nonetheless, there were certainly some correct predictions as seen in the network generated for hapalosin 489 and related compounds which correctly included hapalosin 501 and 503. This network could be used to explore and discover additional variants of hapalosin in addition to already described.

GNPS is a very useful tool, but not perfect. For example, GNPS was not able to recognize the dimer of hapalosin 503. The main reason GNPS was not used as a tool for identifying compounds was due to the lack of library matches with identified molecules to the inputted MS^2 data. Perhaps in the future when there will be more entries of spectra belonging to cyanobacterial products, there will be more to be gained from using GNPS as a dereplication tool on the KVJ10 and KVJ20 samples.

6.5 Identified compounds

The prediction that KVJ10 produced hapalosin was confirmed through HPLC-ESI-MS^2 analysis with library match of observed mass and interpretation of MS^2 spectra. It was also confirmed through MS^2 and NMR analysis that compound **2** is a variant of hapalosin 489, that have an additional carbon and two hydrogens on the hapalosin hydrocarbon tail. The MS^2 spectrum of compound 3 strongly indicates that this too is a variant of hapalosin, but the exact structure is yet to be elucidated.

There was no bioinformatic or chemical indication of hapalosin production in KVJ20, but instead produced other cyclic depsipeptides and peptides that was not found in KVJ10, including nostocyclopeptide A1 and A2, and suomilide B and 1006. KVJ10 and KVJ20 produced many of the same compounds, *e.g.*, compound **4** and pheophytin *a*. However, the compounds common for both strains were not prioritized in this study.

The fractions were only tested for anti-bacterial activity. As hapalosin 489 has earlier been shown to have MDR-reversing activity against multidrug resistant cancer cells, it would be interesting to test hapalosin 501 and hapalosin 503 for the same activity in further studies.

For further investigation of the potential anti-bacterial activity of hapalosin it should be isolated and tested for minimal inhibitory concentration (MIC) to see if they are efficient enough to be of therapeutic value, in addition to toxicity against mammalian cells. Even if the anti-bacterial or anti-cancer effects of hapalosin is good, it cannot be too detrimental for other cells in the body. The beneficial effects must overgo any harmful activity. Biotesting of the hitherto isolated hapalosin samples was not done in order to save the compound for NMR experiments.

It still needs to be proven that, indeed, hapalosin 489 or 503 have inhibiting effect on *S. agalactia*. The battle against antibiotic-resistant bacteria is of need of novel broad-spectrum antibiotics. Is this anti-bacterial activity broad or important enough for being further developed as a potential drug?

6.6 The use of antiSMASH and GNPS

It’s probably safe to say that tools like antiSMASH and GNPS are not flawlessly able to account for every variable. At the current state both tools are very good for initial predictions, sorting and organizing sequences and MS data respectively. The exact predictions of BGCs and compounds are possible only when those are already described and logged in the databases. In a way, these tools are perfect to point to the novelty of the submitted data.

References

(WHO), W. H. O. (2020, 13. october 2020). *Antimicrobial resistance*. Retrieved 18.10.2021 from https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance

Abdel-Razek, A. S., El-Naggar, M. E., Allam, A., Morsy, O. M., & Othman, S. I. (2020). Microbial Natural Products in Drug Discovery. *Processes*, *8*(4), 470. https://www.mdpi.com/2227-9717/8/4/470

Adrover-Castellano, M. L., Schmidt, J. J., & Sherman, D. H. (2021). Biosynthetic Cyclization Catalysts for the Assembly of Peptide and Polyketide Natural Products. *ChemCatChem*, *13*(9), 2095-2116. https://doi.org/10.1002/cctc.202001886

Allakhverdiev, S. I., Tomo, T., Shimada, Y., Kindo, H., Nagao, R., Klimov, V. V., & Mimuro, M. (2010). Redox potential of pheophytin a in photosystem II of two cyanobacteria having the different special pair chlorophylls. *Proceedings of the National Academy of Sciences*, *107*(8), 3924-3929. https://doi.org/10.1073/pnas.0913460107

Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P. C., Entian, K.-D., Fischbach, M. A., Garavelli, J. S., . . . van der Donk, W. A. (2013). Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural Product Reports*, *30*(1), 108-160. https://doi.org/10.1039/c2np20085f

Banker, R., & Carmeli, S. (1998). Tenuecyclamides A−D, Cyclic Hexapeptides from the Cyanobacterium Nostoc spongiaeforme var. tenue. *Journal of Natural Products*, *61*(10), 1248-1251. https://doi.org/10.1021/np980138j

Bauersachs, T., Hopmans, E. C., Compaoré, J., Stal, L. J., Schouten, S., & Damsté, J. S. S. (2009). Rapid analysis of long-chain glycolipids in heterocystous cyanobacteria using high-performance liquid chromatography coupled to electrospray ionization tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, *23*(9), 1387-1394. https://doi.org/https://doi.org/10.1002/rcm.4009

Bladt, T. T., Dürr, C., Knudsen, P. B., Kildgaard, S., Frisvad, J. C., Gotfredsen, C. H., Seiffert, M., & Larsen, T. O. (2013). Bio-Activity and Dereplication-Based Discovery of Ophiobolins and Other Fungal Secondary Metabolites Targeting Leukemia Cells. *Molecules*, *18*(12), 14629-14650. https://www.mdpi.com/1420-3049/18/12/14629

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., van Wezel, G. P., Medema, Marnix H., & Weber, T. (2021). antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Research*, *49*(W1), W29-W35. https://doi.org/10.1093/nar/gkab335

Bloor, S., & England, R. R. (1989). Antibiotic production by the cyanobacteriumNostoc muscorum. *Journal of Applied Phycology*, *1*(4), 367-372. https://doi.org/10.1007/BF00003474

Bouslimani, A., Sanchez, L. M., Garg, N., & Dorrestein, P. C. (2014). Mass spectrometry of natural products: current, emerging and future technologies. *Nat Prod Rep*, *31*(6), 718-729. https://doi.org/10.1039/c4np00044g

Bredholt, H., Fjaervik, E., Johnsen, G., & Zotchev, S. B. (2008). Actinomycetes from sediments in the Trondheim fjord, Norway: diversity and biological activity. *Mar Drugs*, *6*(1), 12-24.

Burja, A. M., Banaigs, B., Abou-Mansour, E., Grant Burgess, J., & Wright, P. C. (2001). Marine cyanobacteria—a prolific source of natural products. *Tetrahedron*, *57*(46), 9347-9377. https://doi.org/https://doi.org/10.1016/S0040-4020(01)00931-0

Calteau, A., Fewer, D. P., Latifi, A., Coursin, T., Laurent, T., Jokela, J., Kerfeld, C. A., Sivonen, K., Piel, J., & Gugger, M. (2014). Phylum-wide comparative genomics unravel the diversity of secondary metabolism in Cyanobacteria. *BMC Genomics*, *15*(1), 977. https://doi.org/10.1186/1471-2164-15-977

Campbell, E. L., & Meeks, J. C. (1989). Characteristics of Hormogonia Formation by Symbiotic Nostoc spp. in Response to the Presence of Anthoceros punctatus or Its Extracellular Products. *Applied and environmental microbiology*, *55*(1), 125-131. https://doi.org/10.1128/aem.55.1.125-131.1989

Cao, L., Do, T., & Link, A. J. (2021). Mechanisms of action of ribosomally synthesized and posttranslationally modified peptides (RiPPs). *Journal of Industrial Microbiology and Biotechnology*, *48*(3-4). https://doi.org/10.1093/jimb/kuab005

Cardellina , J. H., & Moore, B. S. (2010). Editorial: Richard E. Moore (1933−2007). *Journal of Natural Products*, *73*(3), 301-302. https://doi.org/10.1021/np100045f

Cardellina, J. H., Moore, R. E., Arnold, E. V., & Clardy, J. (1979). Structure and absolute configuration of malyngolide, an antibiotic from the marine blue-green alga Lyngbya majuscula Gomont. *The Journal of Organic Chemistry*, *44*(23), 4039-4042. https://doi.org/10.1021/jo01337a003

Carmichael, W. W., Azevedo, S. M., An, J. S., Molica, R. J., Jochimsen, E. M., Lau, S., Rinehart, K. L., Shaw, G. R., & Eaglesham, G. K. (2001). Human fatalities from cyanobacteria: chemical and biological evidence for cyanotoxins. *Environmental health perspectives*, *109*(7), 663-668. https://doi.org/10.1289/ehp.01109663

Carpine, R., & Sieber, S. (2021). Antibacterial and antiviral metabolites from cyanobacteria: Their application and their impact on human health. *Current Research in Biotechnology*, *3*, 65-81. https://doi.org/https://doi.org/10.1016/j.crbiot.2021.03.001

Carrano, L., & Marinelli, F. (2015). The relevance of chemical dereplication in microbial natural product screening. *Journal of Applied Bioanalysis*, *1*(2), 55-67. https://doi.org/https://doi.org/10.17145/jab.15.010

Chaganty, S., Golakoti, T., Heltzel, C., Moore, R. E., & Yoshida, W. Y. (2004). Isolation and Structure Determination of Cryptophycins 38, 326, and 327 from the Terrestrial Cyanobacterium Nostoc sp. GSV 224. *Journal of Natural Products*, *67*(8), 1403-1406. https://doi.org/10.1021/np0499665

Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M.-Y., Paulse, C., Creasy, D., . . . Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, *30*(10), 918-920. https://doi.org/10.1038/nbt.2377

Chang, Z., Flatt, P., Gerwick, W., Nguyen, V.-A., Willis, C., & Sherman, D. (2002). The barbamide biosynthetic gene cluster: A novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit. *Gene*, *296*, 235-247. https://doi.org/10.1016/S0378-1119(02)00860-0

Chin, Y.-W., Balunas, M. J., Chai, H. B., & Kinghorn, A. D. (2006). Drug discovery from natural sources. *The AAPS journal*, *8*(2), E239-E253. https://doi.org/10.1007/BF02854894

Chlipala, G. E., Mo, S., & Orjala, J. (2011). Chemodiversity in freshwater and terrestrial cyanobacteria - a source for drug discovery. *Current drug targets*, *12*(11), 1654-1673. https://doi.org/10.2174/138945011798109455

Cohen, M. F., Wallis, J. G., Campbell, E. L., & Meeks, J. C. (1994). Transposon mutagenesis of Nostoc sp. strain ATCC 29133, a filamentous cyanobacterium with multiple cellular differentiation alternatives. *Microbiology (Reading), 140 (Pt 12)*, 3233-3240. https://doi.org/10.1099/13500872-140-12-3233

Costa, J. L., Paulsrud, P., Rikkinen, J., & Lindblad, P. (2001). Genetic diversity of Nostoc symbionts endophytically associated with two bryophyte species. *Applied and environmental microbiology*, *67*(9), 4393-4396. https://doi.org/10.1128/AEM.67.9.4393-4396.2001

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

Demay, J., Bernard, C., Reinhardt, A., & Marie, B. (2019). Natural Products from Cyanobacteria: Focus on Beneficial Activities. *Marine drugs*, *17*(6), 320. https://doi.org/10.3390/md17060320

Dembitsky, V. M., & Rezanka, T. (2005). Metabolites produced by nitrogen-fixing Nostoc species. *Folia Microbiol (Praha)*, *50*(5), 363-391. https://doi.org/10.1007/bf02931419

Devillers, J., Doré, J. C., Guyot, M., Poroikov, V., Gloriozova, T., Lagunin, A., & Filimonov, D. (2007). Prediction of biological activity profiles of cyanobacterial secondary metabolites. *SAR and QSAR in Environmental Research*, *18*(7-8), 629-643. https://doi.org/10.1080/10629360701698704

Dewick, P. (2001). Secondary Metabolism: The Building Blocks and Construction Mechanisms.

Dewick, P. M. (2009). *Medicinal natural products : a biosynthetic approach* (3rd ed. ed.). Wiley.

Dias, D. A., Urban, S., & Roessner, U. (2012). A historical overview of natural products in drug discovery. *Metabolites*, *2*(2), 303-336. https://doi.org/10.3390/metabo2020303

Dinh, T. Q., Du, X., & Armstrong, R. W. (1996). Synthesis and Conformational Analysis of the Multidrug Resistance-Reversing Agent Hapalosin and Its Non-N-methyl Analog. *J Org Chem*, *61*(19), 6606-6616. https://doi.org/10.1021/jo9608329

Dinh, T. Q., Du, X., Smith, C. D., & Armstrong, R. W. (1997). Synthesis, Conformational Analysis, and Evaluation of the Multidrug Resistance-Reversing Activity of the Triamide and Proline Analogs of Hapalosin. *The Journal of Organic Chemistry*, *62*(20), 6773-6783. https://doi.org/10.1021/jo9708396

Dittmann, E., Gugger, M., Sivonen, K., & Fewer, D. P. (2015). Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria. *Trends Microbiol*, *23*(10), 642-652. https://doi.org/10.1016/j.tim.2015.07.008

Dixit, R., Sharma, N., & Rai, A. (2010). Anticancerous/Antiviral compounds from cyanobacteria. *Algas*, *43*, 4-6.

Dixit., R. B., & Suseela., M. R. (2013). Cyanobacteria: potential candidates for drug discovery. *Antonie van Leuuwenhoek*, *103*. https://doi.org/https://doi.org/10.1007/s10482-013-9898-0

Dodds, W. K., Gudder, D. A., & Mollenhauer, D. (1995). THE ECOLOGY OF NOSTOC. *Journal of Phycology*, *31*(1), 2-18. https://doi.org/https://doi.org/10.1111/j.0022-3646.1995.00002.x

Donadio, S., Monciardini, P., & Sosio, M. (2007). Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics. *Nat Prod Rep*, *24*(5), 1073-1109. https://doi.org/10.1039/b514050c

Du, L., Sánchez, C., & Shen, B. (2001). Hybrid peptide-polyketide natural products: biosynthesis and prospects toward engineering novel molecules. *Metab Eng*, *3*(1), 78-95. https://doi.org/10.1006/mben.2000.0171

Duncan, K. R., Crüsemann, M., Lechner, A., Sarkar, A., Li, J., Ziemert, N., Wang, M., Bandeira, N., Moore, B. S., Dorrestein, P. C., & Jensen, P. R. (2015). Molecular networking and pattern-based genome mining improves discovery of biosynthetic gene clusters and their products from Salinispora species. *Chem Biol*, *22*(4), 460-471. https://doi.org/10.1016/j.chembiol.2015.03.010

Fewer, D., Osterholm, J., Rouhiainen, L., Jokela, J., Wahlsten, M., & Sivonen, K. (2011). Nostophycin Biosynthesis Is Directed by a Hybrid Polyketide Synthase-Nonribosomal Peptide Synthetase in the Toxic Cyanobacterium Nostoc sp. Strain 152. *Applied and environmental microbiology*, *77*, 8034-8040. https://doi.org/10.1128/AEM.05993-11

Fidor, A., Cekała, K., Wieczerzak, E., Cegłowska, M., Kasprzykowski, F., Edwards, C., & Mazur-Marzec, H. (2021). Nostocyclopeptides as New Inhibitors of 20S Proteasome. *Biomolecules*, *11*(10). https://doi.org/10.3390/biom11101483

Figueiredo, S. A. C., Preto, M., Moreira, G., Martins, T. P., Abt, K., Melo, A., Vasconcelos, V. M., & Leão, P. N. (2021). Discovery of Cyanobacterial Natural Products Containing Fatty Acid Residues*. *Angewandte Chemie (International ed. in English)*, *60*(18), 10064-10072. https://doi.org/10.1002/anie.202015105

Fleischman, D. (2012). Chapter 51 - Photosynthesis. In N. Sperelakis (Ed.), *Cell Physiology Source Book (Fourth Edition)* (pp. 909-924). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-387738-3.00051-2

Flores, E., & Herrero, A. (2010). Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat Rev Microbiol*, *8*(1), 39-50. https://doi.org/10.1038/nrmicro2242

Forcisi, S., Moritz, F., Kanawati, B., Tziotis, D., Lehmann, R., & Schmitt-Kopplin, P. (2013). Liquid chromatography–mass spectrometry in metabolomics research: Mass analyzers in ultra high pressure liquid chromatography coupling. *Journal of Chromatography A*, *1292*, 51-65. https://doi.org/https://doi.org/10.1016/j.chroma.2013.04.017

Frankmölle, W. P., Knübel, G., Moore, R. E., & Patterson, G. M. (1992). Antifungal cyclic peptides from the terrestrial blue-green alga Anabaena laxa. II. Structures of laxaphycins A, B, D and E. *J Antibiot (Tokyo)*, *45*(9), 1458-1466. https://doi.org/10.7164/antibiotics.45.1458

Gademann, K., Portmann, C., Blom, J. F., Zeder, M., & Jüttner, F. (2010). Multiple Toxin Production in the Cyanobacterium Microcystis: Isolation of the Toxic Protease Inhibitor Cyanopeptolin 1020. *Journal of Natural Products*, *73*(5), 980-984. https://doi.org/10.1021/np900818c

Golakoti, T., Yoshida, W. Y., Chaganty, S., & Moore, R. E. (2001). Isolation and structure determination of nostocyclopeptides A1 and A2 from the terrestrial cyanobacterium Nostoc sp. ATCC53789. *J Nat Prod*, *64*(1), 54-59. https://doi.org/10.1021/np000316k

Hamada, Y., & Shioiri, T. (2005). Recent Progress of the Synthetic Studies of Biologically Active Marine Cyclic Peptides and Depsipeptides. *Chemical Reviews*, *105*(12), 4441-4482. https://doi.org/10.1021/cr0406312

Haque, F., Banayan, S., Yee, J., & Chiang, Y. W. (2017). Extraction and applications of cyanotoxins and other cyanobacterial secondary metabolites. *Chemosphere*, *183*, 164-175. https://doi.org/https://doi.org/10.1016/j.chemosphere.2017.05.106

Hauert, T., & Komárek, J. (2022). *CyanoDB 2.0 - On-line database of cyanobacterial genera*. Univ. of South Bohemia & Inst. of Botany AS CR, . http://www.cyanodb.cz

Hirata, K., Yoshitomi, S., Dwi, S., Iwabe, O., Mahakhant, A., Polchai, J., & Miyamoto, K. (2003). Bioactivities of nostocine a produced by a freshwater cyanobacterium Nostoc

spongiaeforme TISTR 8169. *Journal of Bioscience and Bioengineering*, *95*(5), 512-517. https://doi.org/https://doi.org/10.1016/S1389-1723(03)80053-1
Hirsch, C. F., Liesch, J. M., Salvatore, M. J., Schwartz, R. E., & Sesin, D. F. (1990). *Antifungal fermentation product and method*. U. Patent.
Hudson, G. A., & Mitchell, D. A. (2018). RiPP antibiotics: biosynthesis and engineering potential. *Current Opinion in Microbiology*, *45*, 61-69. https://doi.org/https://doi.org/10.1016/j.mib.2018.02.010
Hutchinson, C. R. (2003). Polyketide and non-ribosomal peptide synthases: Falling together by coming apart. *Proceedings of the National Academy of Sciences*, *100*(6), 3010. https://doi.org/10.1073/pnas.0730689100
Issa, A. A. (1999). Antibiotic production by the cyanobacteria Oscillatoria angustissima and Calothrix parietina. *Environmental Toxicology and Pharmacology*, *8*(1), 33-37. https://doi.org/https://doi.org/10.1016/S1382-6689(99)00027-7
Ito, T., & Masubuchi, M. (2014). Dereplication of microbial extracts and related analytical technologies. *The Journal of Antibiotics*, *67*(5), 353-360. https://doi.org/10.1038/ja.2014.12
Jones, A. C., Gu, L., Sorrels, C. M., Sherman, D. H., & Gerwick, W. H. (2009). New tricks from ancient algae: natural products biosynthesis in marine cyanobacteria. *Current opinion in chemical biology*, *13*(2), 216-223. https://doi.org/10.1016/j.cbpa.2009.02.019
Kaebernick, M., Neilan, B. A., Börner, T., & Dittmann, E. (2000). Light and the transcriptional response of the microcystin biosynthesis gene cluster. *Applied and environmental microbiology*, *66*(8), 3387-3392. https://doi.org/10.1128/AEM.66.8.3387-3392.2000
Kajiyama, S.-i., Kanzaki, H., Kawazu, K., & Kobayashi, A. (1998). Nostofungicidine, an antifungal lipopeptide from the field-grown terrestrial blue-green alga Nostoc commune. *Tetrahedron Letters*, *39*(22), 3737-3740. https://doi.org/https://doi.org/10.1016/S0040-4039(98)00573-5
Kanekiyo, K., Lee, J. B., Hayashi, K., Takenaka, H., Hayakawa, Y., Endo, S., & Hayashi, T. (2005). Isolation of an antiviral polysaccharide, nostoflan, from a terrestrial cyanobacterium, Nostoc flagelliforme. *J Nat Prod*, *68*(7), 1037-1041. https://doi.org/10.1021/np050056c
Kashihara, N., To-e, S., Nakamura, K., Umezawa, K., Yamamura, S., & Nishiyama, S. (2000). Synthesis and biological activities of hapalosin derivatives with modification at the C12 position. *Bioorg Med Chem Lett*, *10*(2), 101-103. https://doi.org/10.1016/s0960-894x(99)00647-2
Katz, L., & Baltz, R. H. (2016). Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol*, *43*(2-3), 155-176. https://doi.org/10.1007/s10295-015-1723-5
Kehr, J. C., Gatte Picchi, D., & Dittmann, E. (2011). Natural product biosyntheses in cyanobacteria: A treasure trove of unique enzymes. *Beilstein J Org Chem*, *7*, 1622-1635. https://doi.org/10.3762/bjoc.7.191
Kessner, D., Chambers, M., Burke, R., Agus, D., & Mallick, P. (2008). ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics*, *24*(21), 2534-2536. https://doi.org/10.1093/bioinformatics/btn323
Knübel, G., Larsen, L. K., Moore, R. E., Levine, I. A., & Patterson, G. M. L. (1990). Cytotoxic, Antiviral Indolocarbazoles from a Blue-Green Alga Belonging to the Nostocaceae. *The Journal of Antibiotics*, *43*(10), 1236-1239. https://doi.org/https://doi.org/10.7164/antibiotics.43.1236

Koehn, F. E. (2008). New Strategies and Methods in the Discovery of Natural Product Anti-Infective Agents: The Mannopeptimycins. *Journal of Medicinal Chemistry*, *51*(9), 2613-2617. https://doi.org/10.1021/jm070432l

Komárek, J. l., Kbatovský, J., Marea, J., & Johansen, J. R. (2014). Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach. *Preslia*, *86*.

Kultschar., B., & Llewellyn., C. (2018). Secondary Metabolites in Cyanobacteria. In R. Vijayakumar. & S. S. S. Raja (Eds.), *Secondary Metabolites - Sources and Applications*. https://doi.org/http://dx.doi.org/10.5772/intechopen.75648

Leão, P. N., Engene, N., Antunes, A., Gerwick, W. H., & Vasconcelos, V. (2012). The chemical ecology of cyanobacteria. *Nat Prod Rep*, *29*(3), 372-391. https://doi.org/10.1039/c2np00075j

Liaimer, A., Jensen, J. B., & Dittmann, E. (2016). A Genetic and Chemical Perspective on Symbiotic Recruitment of Cyanobacteria of the Genus Nostoc into the Host Plant Blasia pusilla L [Original Research]. *Frontiers in Microbiology*, *7*(1693). https://doi.org/10.3389/fmicb.2016.01693

Liu, X. J., & Chen, F. (2003). Cell differentiation and colony alteration of an edible terrestrial cyanobacterium Nostoc flagelliforme, in liquid suspension cultures. *Folia Microbiol (Praha)*, *48*(5), 619-626. https://doi.org/10.1007/bf02993468

Luesch, H., Yoshida, W. Y., Moore, R. E., & Paul, V. J. (2002). New apratoxins of marine cyanobacterial origin from Guam and Palau. *Bioorg Med Chem*, *10*(6), 1973-1978. https://doi.org/10.1016/s0968-0896(02)00014-7

Magarvey, N. A., Beck, Z. Q., Golakoti, T., Ding, Y., Huber, U., Hemscheidt, T. K., Abelson, D., Moore, R. E., & Sherman, D. H. (2006). Biosynthetic Characterization and Chemoenzymatic Assembly of the Cryptophycins. Potent Anticancer Agents from Nostoc Cyanobionts. *ACS Chemical Biology*, *1*(12), 766-779. https://doi.org/10.1021/cb6004307

Mandal., S., & Rath., J. (2015). *Extremophilic Cyanobacteria For Novel Drug Development* (1 ed.). Springer, Cham. https://doi.org/https://doi-org.mime.uit.no/10.1007/978-3-319-12009-6

Maplestone, R. A., Stone, M. J., & Williams, D. H. (1992). The evolutionary role of secondary metabolites — a review. *Gene*, *115*(1), 151-157. https://doi.org/https://doi.org/10.1016/0378-1119(92)90553-2

Martin, C., Oberer, L., Ino, T., König, W. A., Busch, M., & Weckesser, J. (1993). Cyanopeptolins, new depsipeptides from the cyanobacterium Microcystis sp. PCC 7806. *J Antibiot (Tokyo)*, *46*(10), 1550-1556. https://doi.org/10.7164/antibiotics.46.1550

Mazard, S., Penesyan, A., Ostrowski, M., Paulsen, I. T., & Egan, S. (2016). Tiny Microbes with a Big Impact: The Role of Cyanobacteria and Their Metabolites in Shaping Our Future. *Marine drugs*, *14*(5), 97. https://doi.org/10.3390/md14050097

Medema, M. H., de Rond, T., & Moore, B. S. (2021). Mining genomes to illuminate the specialized chemistry of life. *Nature Reviews Genetics*, *22*(9), 553-571. https://doi.org/10.1038/s41576-021-00363-7

Meeks, J. C. (1998). Symbiosis between Nitrogen-Fixing Cyanobacteria and Plants. *BioScience*, *48*(4), 266-276. https://doi.org/10.2307/1313353

Meeks, J. C., & Elhai, J. (2002). Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. *Microbiol Mol Biol Rev*, *66*(1), 94-121; table of contents. https://doi.org/10.1128/mmbr.66.1.94-121.2002

Mevers, E., Liu, W.-T., Engene, N., Mohimani, H., Byrum, T., Pevzner, P. A., Dorrestein, P. C., Spadafora, C., & Gerwick, W. H. (2011). Cytotoxic Veraguamides, Alkynyl Bromide-Containing Cyclic Depsipeptides from the Marine Cyanobacterium cf. Oscillatoria margaritifera. *Journal of Natural Products*, *74*(5), 928-936. https://doi.org/10.1021/np200077f

Miyanaga, A., Kudo, F., & Eguchi, T. (2018). Protein-protein interactions in polyketide synthase-nonribosomal peptide synthetase hybrid assembly lines. *Nat Prod Rep*, *35*(11), 1185-1209. https://doi.org/10.1039/c8np00022k

Moore, R. E. (1996). Cyclic peptides and depsipeptides from cyanobacteria: A review. *Journal of Industrial Microbiology*, *16*(2), 134-143. https://doi.org/10.1007/BF01570074

Mullis, M. M., Rambo, I. M., Baker, B. J., & Reese, B. K. (2019). Diversity, Ecology, and Prevalence of Antimicrobials in Nature [Review]. *Frontiers in Microbiology*, *10*(2518). https://doi.org/10.3389/fmicb.2019.02518

Niedermeyer, T. H. (2015). Anti-infective Natural Products from Cyanobacteria. *Planta Med*, *81*(15), 1309-1325. https://doi.org/10.1055/s-0035-1546055

O'Connell, C. E., Salvato, K. A., Meng, Z., Littlefield, B. A., & Schwartz, C. E. (1999). Synthesis and evaluation of hapalosin and analogs as MDR-reversing agents. *Bioorg Med Chem Lett*, *9*(11), 1541-1546. https://doi.org/10.1016/s0960-894x(99)00243-7

Okino, T., Qi, S., Matsuda, H., Murakami, M., & Yamaguchi, K. (1997). Nostopeptins A and B, Elastase Inhibitors from the Cyanobacterium Nostoc minutum. *Journal of Natural Products*, *60*(2), 158-161. https://doi.org/10.1021/np960649a

Ortega, M. A., & van der Donk, W. A. (2016). New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products. *Cell chemical biology*, *23*(1), 31-44. https://doi.org/10.1016/j.chembiol.2015.11.012

Palomo, C., Oiarbide, M., García, J. M., González, A., Pazos, R., Odriozola, J. M., Bañuelos, P., Tello, M., & Linden, A. (2004). A Practical Total Synthesis of Hapalosin, a 12-Membered Cyclic Depsipeptide with Multidrug Resistance-Reversing Activity, by Employing Improved Segment Coupling and Macrolactonization. *The Journal of Organic Chemistry*, *69*(12), 4126-4134. https://doi.org/10.1021/jo0497499

Pandey, V. D. (2015). Cyanobacterial natural products as antimicrobial agents. *International Journal of Current Microbiology and Applied Sciences*, *4*(1), 310-317.

Patti, G. J. (2011). Separation strategies for untargeted metabolomics. *J Sep Sci*, *34*(24), 3460-3469. https://doi.org/10.1002/jssc.201100532

Paul, D. G. N., & J., V. (1999). Production of Secondary Metabolites by Filamentous Tropical Marine Cyanobacteria: Ecological Functions of the Compounds. *Journal of Phycology*, *35*(6), 1412-1421. https://doi.org/https://doi.org/10.1046/j.1529-8817.1999.3561412.x

Pedersen-Bjergaard, S., & Rasmussen, K. E. (2004). *Legemiddelanalyse*. Fagbokforl.

Pedersén, M., & Dasilva, E. J. (1973). Simple brominated phenols in the bluegreen alga Calothrix brevissima West. *Planta*, *115*(1), 83-86. https://doi.org/10.1007/bf00388608

Pergament, I., & Carmeli, S. (1994). Schizotrin A; a novel antimicrobial cyclic peptide from a cyanobacterium. *Tetrahedron Letters*, *35*(45), 8473-8476. https://doi.org/https://doi.org/10.1016/S0040-4039(00)74436-4

Pinu, F. R., Villas-Boas, S. G., & Aggio, R. (2017). Analysis of Intracellular Metabolites from Microorganisms: Quenching and Extraction Protocols. *Metabolites*, *7*(4), 53. https://doi.org/10.3390/metabo7040053

Purves, K., Macintyre, L., Brennan, D., Hreggviðsson, G., Kuttner, E., Ásgeirsdóttir, M. E., Young, L. C., Green, D. H., Edrada-Ebel, R., & Duncan, K. R. (2016). Using

Molecular Networking for Microbial Secondary Metabolite Bioprospecting. *Metabolites*, *6*(1). https://doi.org/10.3390/metabo6010002
Rao, M., Malhotra, S., Fatma, T., & Rattan, A. (2007). Antimycobacterial Activity from Cyanobacterial Extracts and Phytochemical Screening of Methanol Extract of Hapalosiphon. *Pharmaceutical Biology*, *45*(2), 88-93. https://doi.org/10.1080/13880200601105319
Rathod, R. H., Chaudhari, S. R., Patil, A. S., & Shirkhedkar, A. A. (2019). Ultra-high performance liquid chromatography-MS/MS (UHPLC-MS/MS) in practice: analysis of drugs and pharmaceutical formulations. *Future Journal of Pharmaceutical Sciences*, *5*(1), 6. https://doi.org/10.1186/s43094-019-0007-8
Raveh, A., & Carmeli, S. (2010). Aeruginazole A, a Novel Thiazole-Containing Cyclopeptide from the Cyanobacterium Microcystis sp. *Organic Letters*, *12*(15), 3536-3539. https://doi.org/10.1021/ol1014015
Reshef, V., & Carmeli, S. (2002). Schizopeptin 791, a New Anabeanopeptin-like Cyclic Peptide from the Cyanobacterium Schizothrix sp. *Journal of Natural Products*, *65*(8), 1187-1189. https://doi.org/10.1021/np020039c
Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M., & Stanier, R. Y. (1979). Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Microbiology*, *111*(1), 1-61. https://doi.org/https://doi.org/10.1099/00221287-111-1-1
Rojas, V., Rivas, L., Cárdenas, C., & Guzmán, F. (2020). Cyanobacteria and Eukaryotic Microalgae as Emerging Sources of Antibacterial Peptides. *Molecules*, *25*(24). https://doi.org/10.3390/molecules25245804
Sainis, I., Fokas, D., Vareli, K., Tzakos, A. G., Kounnis, V., & Briasoulis, E. (2010). Cyanobacterial cyclopeptides as lead compounds to novel targeted cancer drugs. *Marine drugs*, *8*(3), 629-657. https://doi.org/10.3390/md8030629
Sánchez-Bayo, A., Morales, V., Rodríguez, R., Vicente, G., & Bautista, L. F. (2020). Cultivation of Microalgae and Cyanobacteria: Effect of Operating Conditions on Growth and Biomass Composition. *Molecules*, *25*(12), 2834. https://www.mdpi.com/1420-3049/25/12/2834
Sarker, S. D., Latif, Z., Gray, A.I. (2005). *Natural Product Isolation* (2 ed., Vol. 20). Humana Press. https://doi.org/https://doi.org/10.1385/1592599559
Schneider, Y. K. (2020). *Bioactive secondary metabolites from bacteria: Natural products from marine and terrestrial bacteria, dereplication, isolation and investigation of bacterial secondary metabolites* UiT The Arctic University of Norway]. Munin. https://hdl.handle.net/10037/19264
Schneider, Y. K., Jørgensen, S. M., Andersen, J. H., & Hansen, E. H. (2021). Qualitative and Quantitative Comparison of Liquid–Liquid Phase Extraction Using Ethyl Acetate and Liquid–Solid Phase Extraction Using Poly-Benzyl-Resin for Natural Products. *Applied Sciences*, *11*(21), 10241. https://www.mdpi.com/2076-3417/11/21/10241
Schwartz, R. E., Hirsch, C. F., Sesin, D. F., Flor, J. E., Chartrain, M., Fromtling, R. E., Harris, G. H., Salvatore, M. J., Liesch, J. M., & Yudin, K. (1990). Pharmaceuticals from cultured algae. *Journal of Industrial Microbiology*, *5*(2), 113-123. https://doi.org/10.1007/BF01573860
Schwarzer, D., & Marahiel, M. A. (2001). Multimodular biocatalysts for natural product assembly. *Naturwissenschaften*, *88*(3), 93-101. https://doi.org/10.1007/s001140100211
Simpson, W. G. (1985). The calcium channel blocker verapamil and cancer chemotherapy. *Cell Calcium*, *6*(6), 449-467. https://doi.org/https://doi.org/10.1016/0143-4160(85)90021-1

Singh, S., Kate, B. N., & Banerjee, U. C. (2005). Bioactive Compounds from Cyanobacteria and Microalgae: An Overview. *Critical Reviews in Biotechnology*, *25*(3), 73-95. https://doi.org/10.1080/07388550500248498

Singh., R. K., Tiwari., S. P., Rai., A. K., & Mohapatra, T. M. (2011). Cyanobacteria: an emergin source for drug discovery [Review]. *The Journal of Antibiotics*, *64*, 401-412. https://doi.org/https://doi.org/10.1038/ja.2011.21

Steckel, A., & Schlosser, G. (2019). An Organic Chemist's Guide to Electrospray Mass Spectrometric Structure Elucidation. *Molecules*, *24*(3), 611. https://www.mdpi.com/1420-3049/24/3/611

Stratmann, K., Burgoyne, D. L., Moore, R. E., Patterson, G. M. L., & Smith, C. D. (1994). Hapalosin, a Cyanobacterial Cyclic Depsipeptide with Multidrug-Resistance Reversing Activity. *The Journal of Organic Chemistry*, *59*(24), 7219-7226. https://doi.org/10.1021/jo00103a011

Strobel, G., Daisy, B., Castillo, U., & Harper, J. (2004). Natural products from endophytic microorganisms. *J Nat Prod*, *67*(2), 257-268. https://doi.org/10.1021/np030397v

Swain, S. S., Paidesetty, S. K., & Padhy, R. N. (2017). Antibacterial, antifungal and antimycobacterial compounds from cyanobacteria. *Biomed Pharmacother*, *90*, 760-776. https://doi.org/10.1016/j.biopha.2017.04.030

Tidgewell, K., Clark, B. R., & Gerwick, W. H. (2010). 2.06 - The Natural Products Chemistry of Cyanobacteria. In H.-W. Liu & L. Mander (Eds.), *Comprehensive Natural Products II* (pp. 141-188). Elsevier. https://doi.org/https://doi.org/10.1016/B978-008045382-8.00041-1

Trimurtulu, G., Ohtani, I., Patterson, G. M. L., Moore, R. E., Corbett, T. H., Valeriote, F. A., & Demchik, L. (1994). Total Structures of Cryptophycins, Potent Antitumor Depsipeptides from the Blue-Green Alga Nostoc sp. Strain GSV 224. *Journal of the American Chemical Society*, *116*(11), 4729-4737. https://doi.org/10.1021/ja00090a020

Tripathi, A., Puddick, J., Prinsep, M. R., Lee, P. P. F., & Tan, L. T. (2010). Hantupeptins B and C, cytotoxic cyclodepsipeptides from the marine cyanobacterium Lyngbya majuscula. *Phytochemistry*, *71*(2), 307-311. https://doi.org/https://doi.org/10.1016/j.phytochem.2009.10.006

Ueoka, R., Bhushan, A., Probst, S. I., Bray, W. M., Lokey, R. S., Linington, R. G., & Piel, J. (2018). Genome-Based Identification of a Plant-Associated Marine Bacterium as a Rich Natural Product Source. *Angewandte Chemie International Edition*, *57*(44), 14519-14523. https://doi.org/https://doi.org/10.1002/anie.201805673

Ventola, C. L. (2015). The antibiotic resistance crisis: part 1: causes and threats. *P t*, *40*(4), 277-283.

Wagner, B., Gonzalez, G. I., Tran Hun Dau, M. E., & Zhu, J. (1999). Total synthesis and conformational studies of hapalosin, N-desmethylhapalosin and 8-deoxyhapalosin. *Bioorg Med Chem*, *7*(5), 737-747. https://doi.org/10.1016/s0968-0896(98)00208-9

Wang, H., Fewer, D. P., Holm, L., Rouhiainen, L., & Sivonen, K. (2014). Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. *Proceedings of the National Academy of Sciences*, *111*(25), 9259. https://doi.org/10.1073/pnas.1401734111

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen, D. D., Watrous, J., Kapono, C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W.-T., Crüsemann, M., Boudreau, P. D., Esquenazi, E., Sandoval-Calderón, M., . . . Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology*, *34*(8), 828-837. https://doi.org/10.1038/nbt.3597

Wang, Z., & Yang, L. (2020). Turning the Tide: Natural Products and Natural-Product-Inspired Chemicals as Potential Counters to SARS-CoV-2 Infection [Mini Review]. *Frontiers in Pharmacology*, *11*(1013). https://doi.org/10.3389/fphar.2020.01013

Warshan, D., Liaimer, A., Pederson, E., Kim, S. Y., Shapiro, N., Woyke, T., Altermark, B., Pawlowski, K., Weyman, P. D., Dupont, C. L., & Rasmussen, U. (2018). Genomic Changes Associated with the Evolutionary Transitions of Nostoc to a Plant Symbiont. *Mol Biol Evol*, *35*(5), 1160-1175. https://doi.org/10.1093/molbev/msy029

Weckesser, J., Martin, C., & Jakobi, C. (1996). Cyanopeptolins, depsipeptides from cyanobacteria. *Systematic and Applied Microbiology*, *19*(2), 133-138. https://doi.org/https://doi.org/10.1016/S0723-2020(96)80038-5

Weissman, K. J. (2015). The structural biology of biosynthetic megaenzymes. *Nature Chemical Biology*, *11*(9), 660-670. https://doi.org/10.1038/nchembio.1883

Welker, M., Dittmann, E., & von Döhren, H. (2012). Chapter Two - Cyanobacteria as a Source of Natural Products. In D. A. Hopwood (Ed.), *Methods in Enzymology* (Vol. 517, pp. 23-46). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-404634-4.00002-4

Welker, M., & von Döhren, H. (2006). Cyanobacterial peptides - nature's own combinatorial biosynthesis. *FEMS Microbiol Rev*, *30*(4), 530-563. https://doi.org/10.1111/j.1574-6976.2006.00022.x

Wilhelmsen, O. S. B. (2019). *Expression of Secondary Metabolite Gene Clusters and Production of Secondary Metabolites in Three Nostoc strains Subjected to Deprivation on Nutrients and Competition* UiT The Arctic University of Norway]. Munin. https://hdl.handle.net/10037/16119

Williams, P. G., Yoshida, W. Y., Quon, M. K., Moore, R. E., & Paul, V. J. (2003). The Structure of Palau'amide, a Potent Cytotoxin from a Species of the Marine Cyanobacterium Lyngbya. *Journal of Natural Products*, *66*(12), 1545-1549. https://doi.org/10.1021/np034001r

Xue, Y., Zhao, P., Quan, C., Zhao, Z., Gao, W., Li, J., Zu, X., Fu, D., Feng, S., Bai, X., Zuo, Y., & Li, P. (2018). Cyanobacteria-derived peptide antibiotics discovered since 2000. *Peptides*, *107*, 17-24. https://doi.org/https://doi.org/10.1016/j.peptides.2018.08.002

Yang, J. Y., Sanchez, L. M., Rath, C. M., Liu, X., Boudreau, P. D., Bruns, N., Glukhov, E., Wodtke, A., de Felicio, R., Fenner, A., Wong, W. R., Linington, R. G., Zhang, L., Debonsi, H. M., Gerwick, W. H., & Dorrestein, P. C. (2013). Molecular Networking as a Dereplication Strategy. *Journal of Natural Products*, *76*(9), 1686-1699. https://doi.org/10.1021/np400413s

Zainuddin, E. N., Jansen, R., Nimtz, M., Wray, V., Preisitsch, M., Lalk, M., & Mundt, S. (2009). Lyngbyazothrins A−D, Antimicrobial Cyclic Undecapeptides from the Cultured Cyanobacterium Lyngbya sp. *Journal of Natural Products*, *72*(8), 1373-1378. https://doi.org/10.1021/np8007792

Zhang, Q., Ortega, M., Shi, Y., Wang, H., Melby, J. O., Tang, W., Mitchell, D. A., & van der Donk, W. A. (2014). Structural investigation of ribosomally synthesized natural products by hypothetical structure enumeration and evaluation using tandem MS. *Proceedings of the National Academy of Sciences*, *111*(33), 12031-12036. https://doi.org/10.1073/pnas.1406418111

Appendix 1: Stock solutions in Liquid BG 11_0 medium

Table A.1. Stock solutions I.-VI. used in BG 11_0 *medium.*

Stock	*Compound*	*g/L*
I.	$K_2HPO_4 \times 3H_2O$	0.04
II.	$MgSO_4 \times 7H_2O$	0.075
III.	$CaCl_2 \times 2H_2O$	0.036
	Citric acid	0.006
IV.	Ferric ammonium citrate	0.006
	Na_2Mg EDTA	0.001
V.	Na_2CO_3	0.004
VI.	Trace elements (1 mL/L):	
	H_3BO_3	2.86
	$MnCl_2 \times 4H_2O$	1.81
	$ZnSO_4 \times 7H_2O$	0.222
	$Na_2MoO_4 \times 2H_2O$	0.39
	$CuSO_4 \times 5H_2O$	0.079
	$Co(NO_3)_2 \times 6H_2O$	0.0404

Appendix 2: Chromatograms of Crude Extracts

All chromatograms of all KVJ20 and KVJ10 crude extracts from UHPLC analysis

Chromatograms KVJ20:

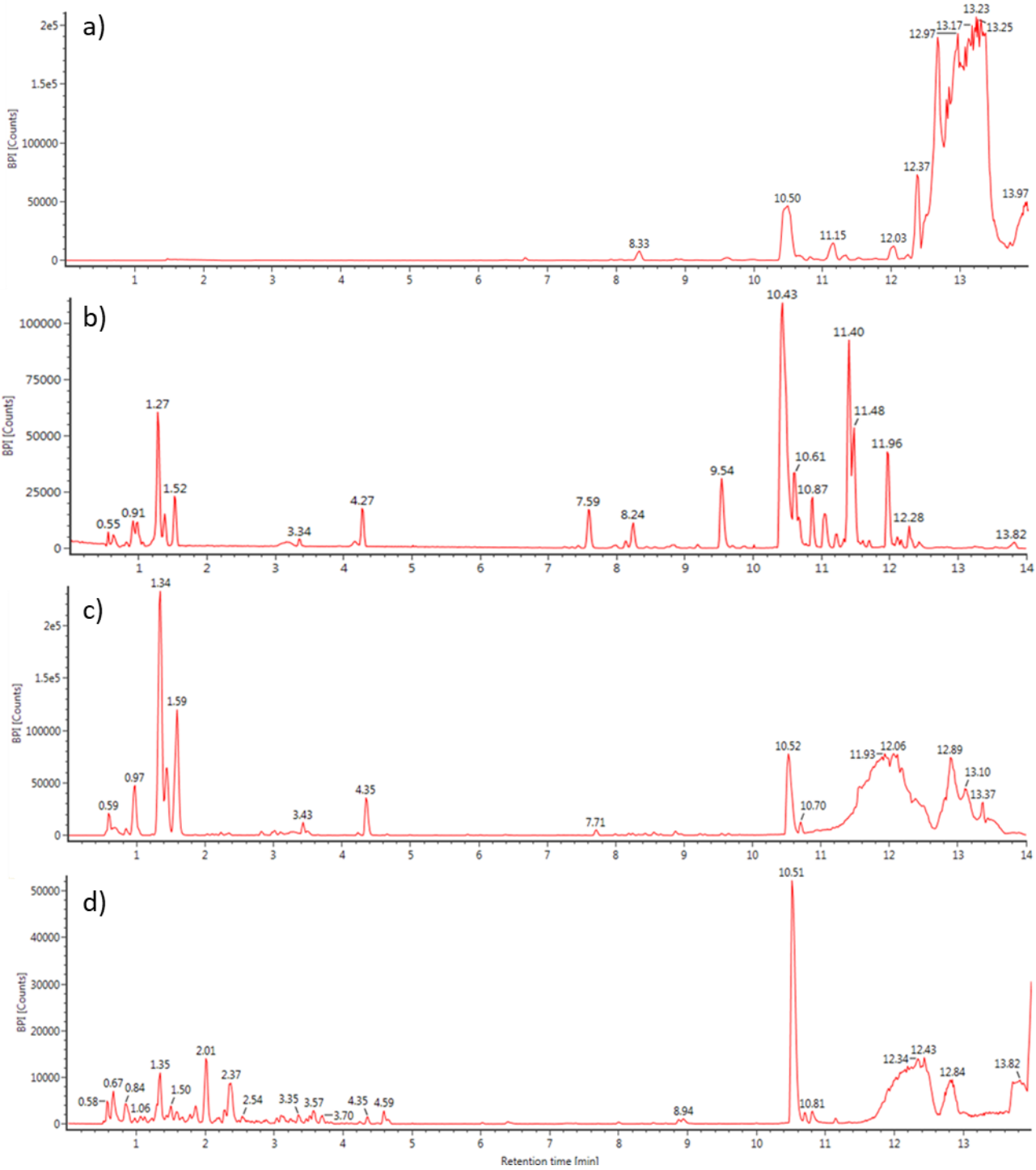


Figure A.1. Chromatograms of crude extracts of KVJ20. a) EtOAc, b) 100% MeOH, c) 50% MeOH, and d) Aqueous extract.

Chromatograms KVJ10:

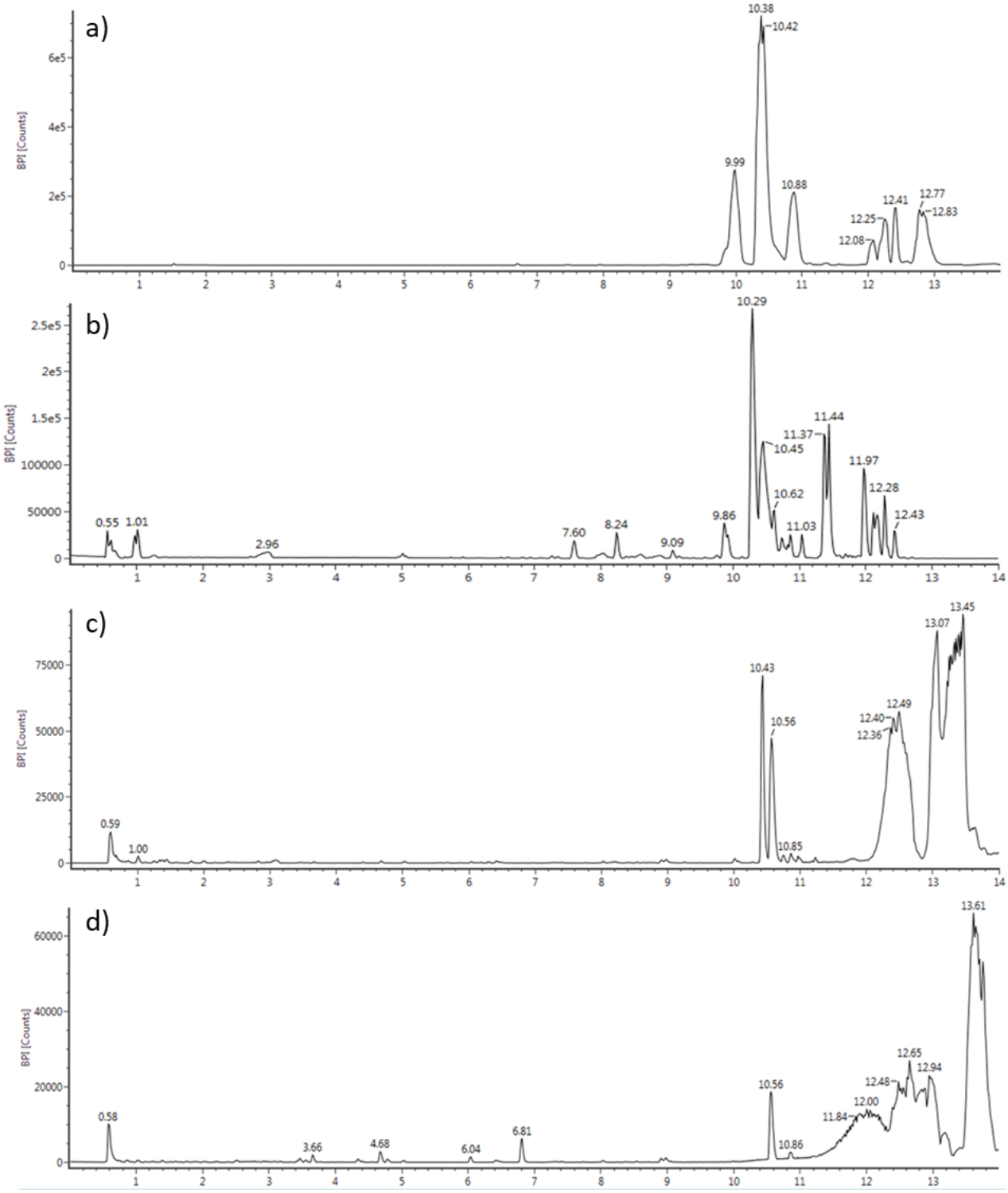


Figure A.2. Chromatograms of crude extracts of KVJ10. a) EtOAc, b) 100% MeOH, c) 50% MeOH, and d) Aqueous extract.

Appendix 3: Nuclear Magnetic Resonance (NMR) Spectrometry of compound 2 (Hapalosin 503)

Proton NMR (nuclear magnetic resonance):

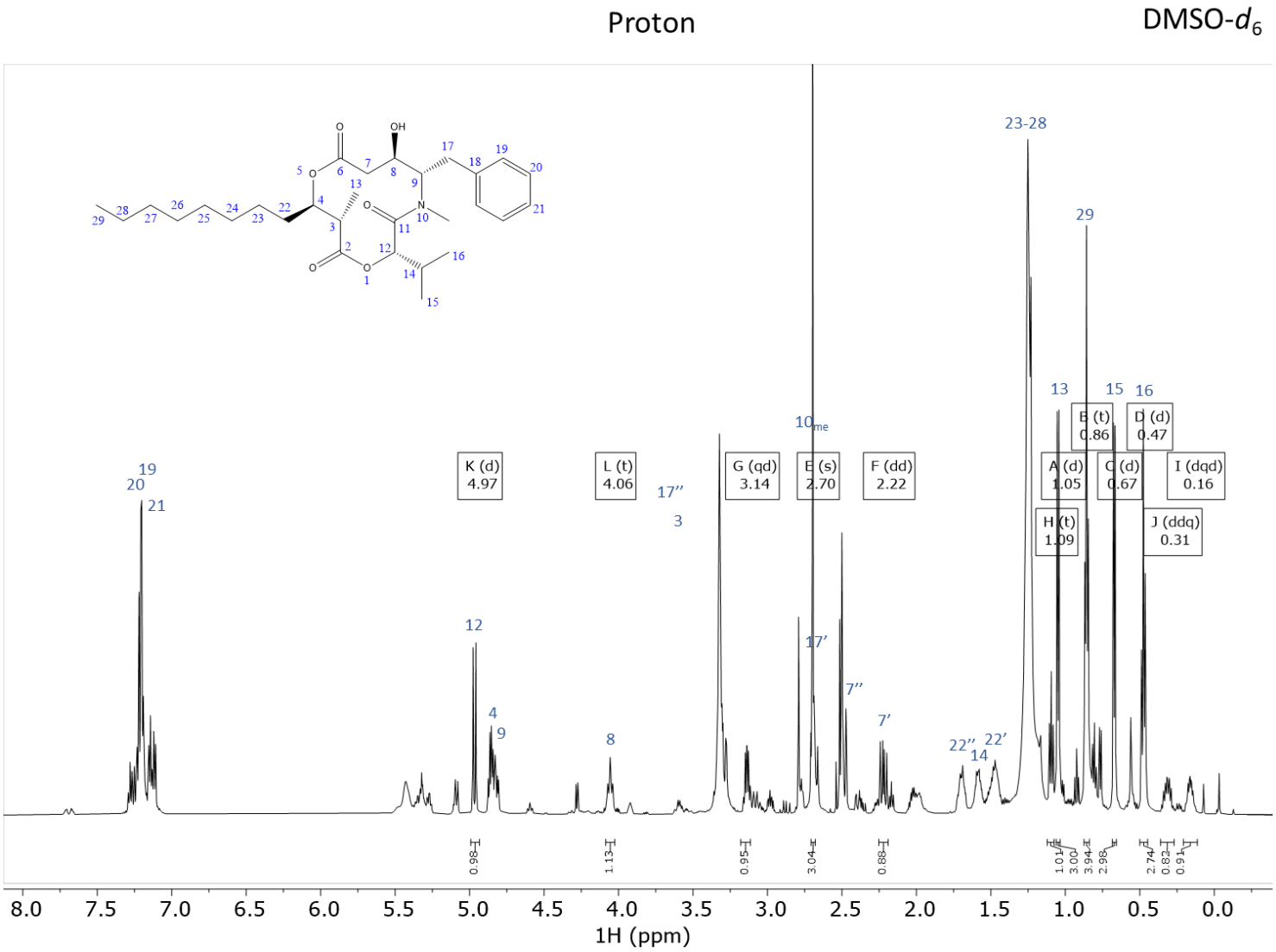


*Figure A.3. Proton NMR (nuclear magnetic resonance) spectrum of compound **2**.*

^{13}C NMR:

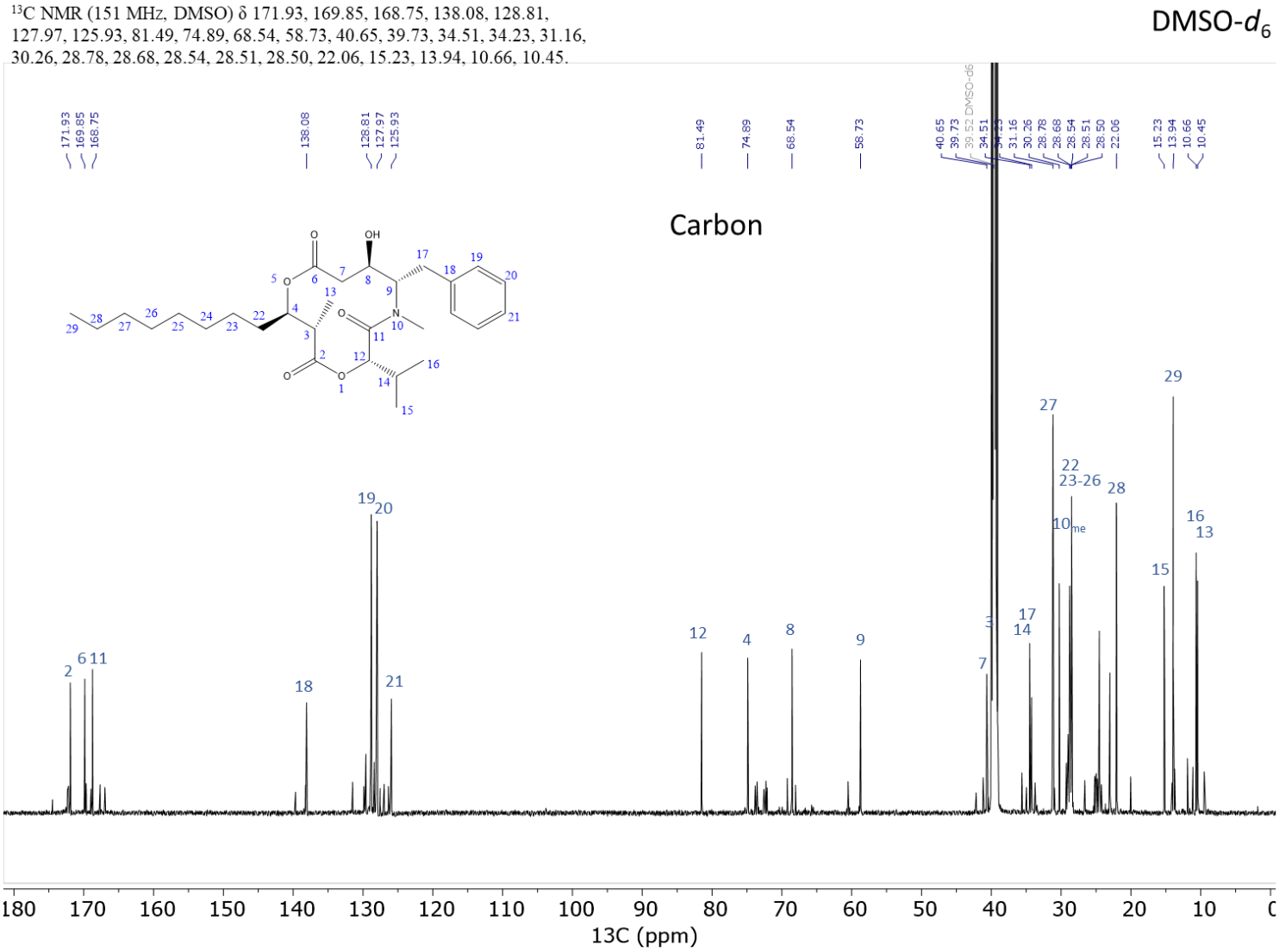


*Figure A.4. ^{13}C NMR of compound **2**.*

HSQC+HMBC NMR:

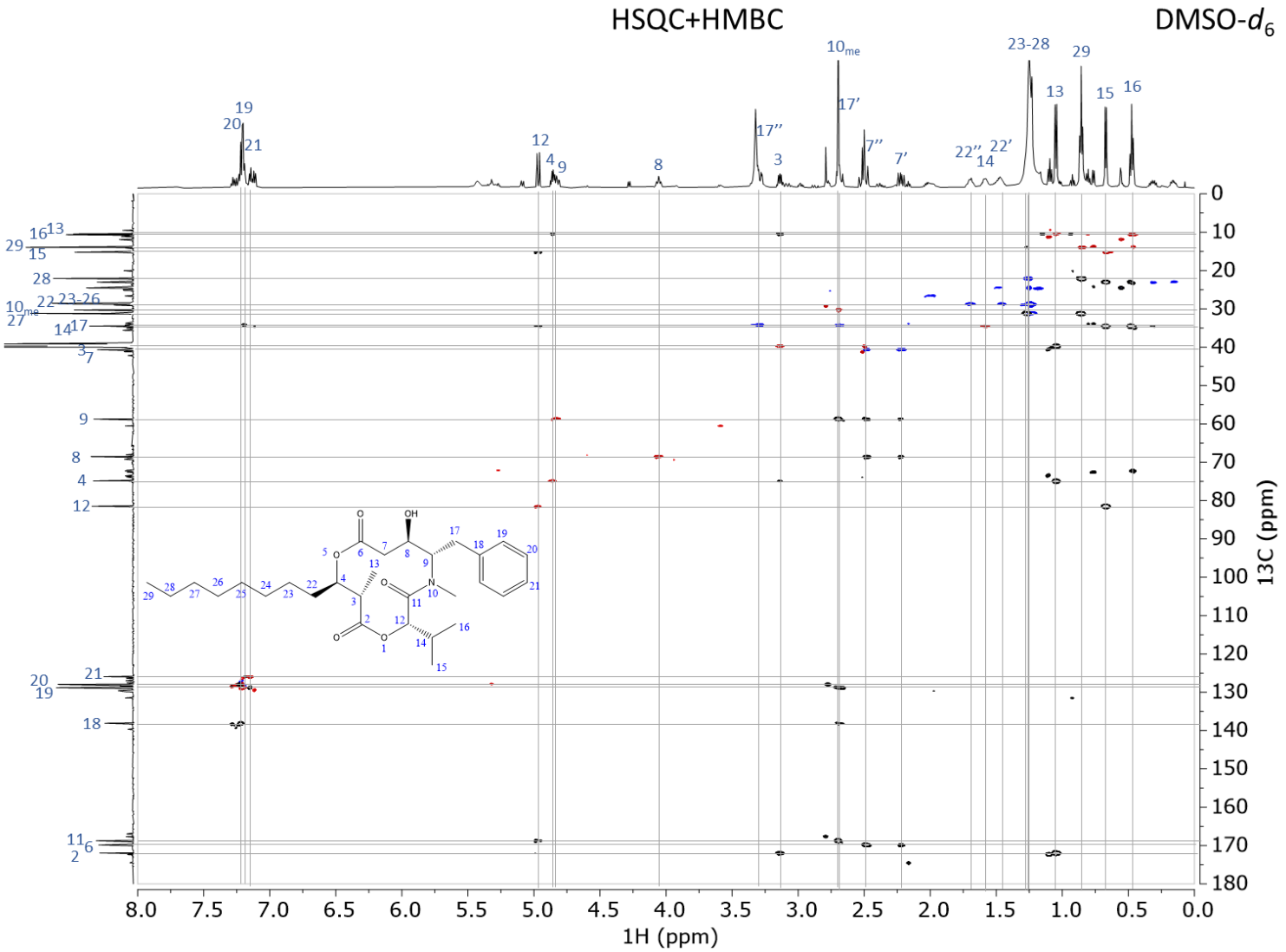


Figure A.5. HSQC (heteronuclear single quantum coherence) combined with HMBC (heteronuclear multiple bond correlation) spectrum of compound 2.

DQF-COSY NMR:

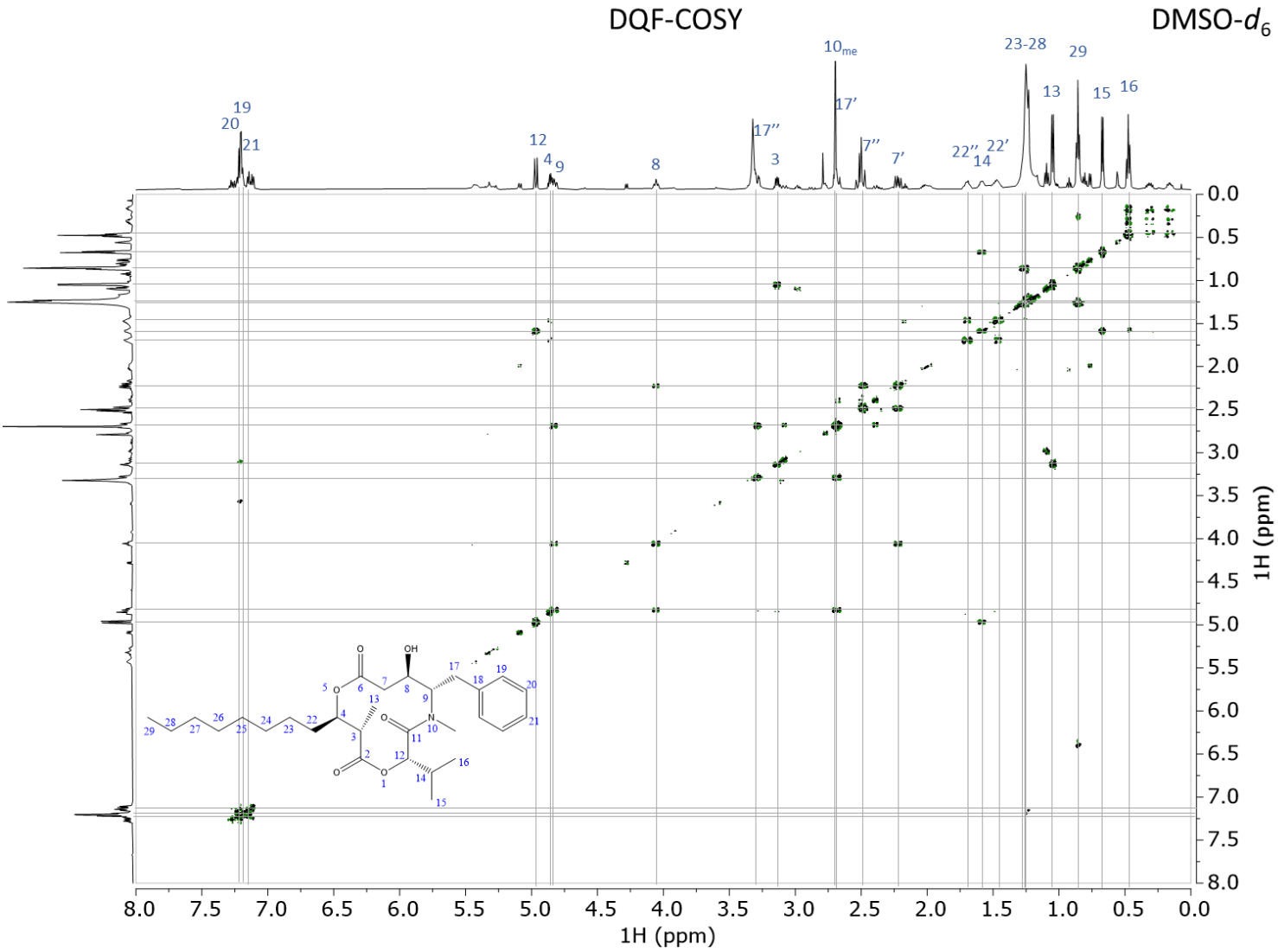


Figure A.6. DQF (double-quantum filtered)-COSY (correlated spectroscopy) spectrum of compound ***2***

ROESY NMR:

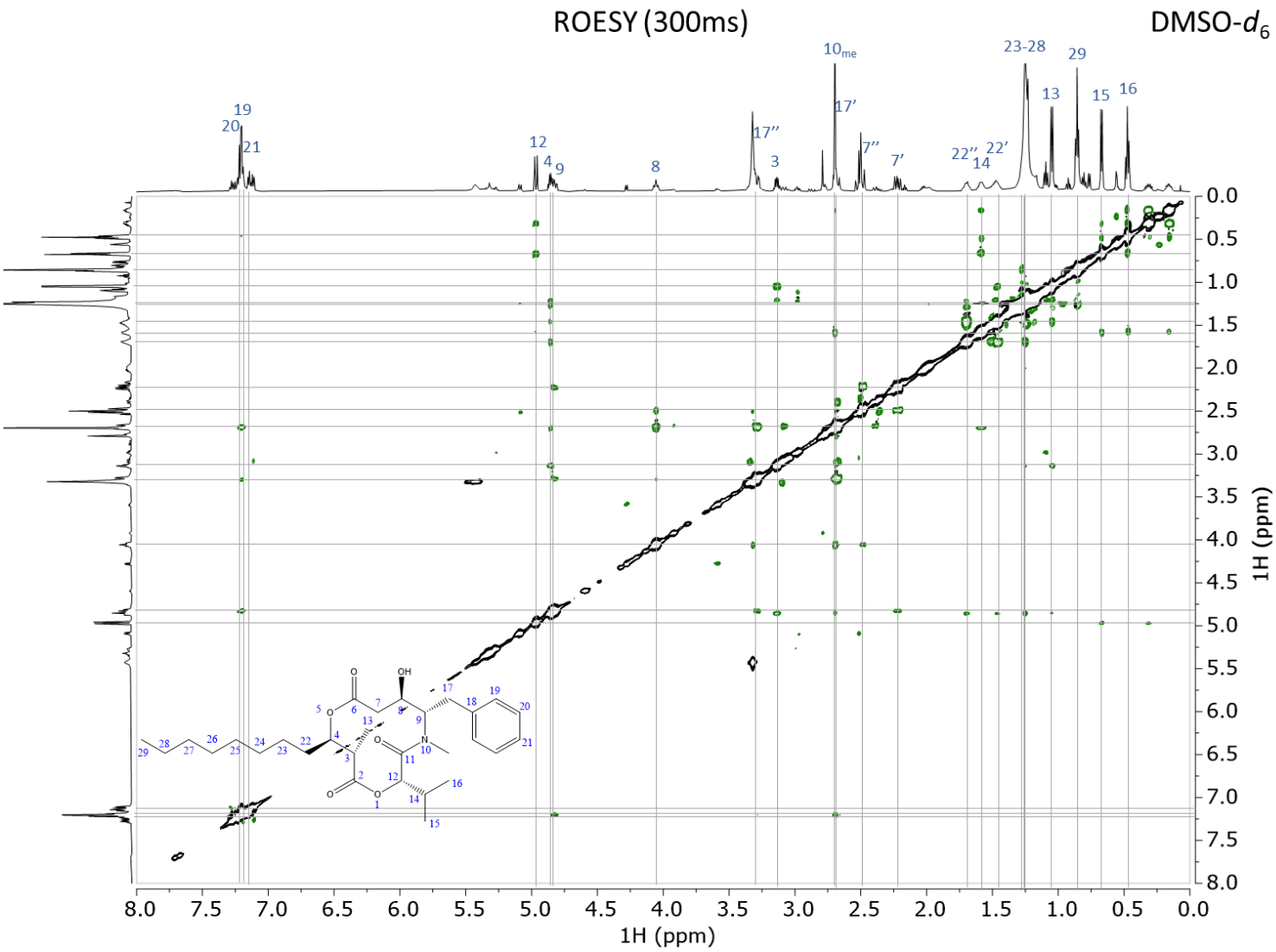


Figure A.7. ROESY (rotating frame overhauser enhancement spectroscopy) spectrum of compound ***2***