# Discovery of Ribosomally Synthesized Lipopeptides from Cyanobacteria

João Pedro Teixeira Pereira
Master's Degree in Applications in Biotechnology and Synthetic Biology
Department of Chemistry and Biochemistry and Department of Biology
2023

**Supervisor**
Pedro Nuno da Costa Leão, Principal Investigator, CIIMAR – Interdisciplinary Centre of Marine and Environmental Research


U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

*When you protect your dharma, your dharma protects you.*

- Manusmriti 8:15

# Acknowledgements

First and foremost, I would like to thank Dr. Pedro Leão for the opportunity to develop my master thesis in his lab. His guidance, trust, and the freedom he provided me to explore and experiment with various ideas have been instrumental in shaping the direction of this work.

I am deeply indebted to the Cyanobacterial Natural Products research team, with a special mention to Raquel. Raquel's support and mentorship over the course of an entire year were invaluable. She not only shared her knowledge and expertise but also nurtured me as a researcher, helping me grow and develop my skills. I am thankful for her unwavering mentorship and friendship.

To my girlfriend, Alexandra, I extend my heartfelt thanks for standing by my side during the challenging moments. Your support, patience, and understanding were my pillars of strength, and I couldn't have done it without you.

I also want to acknowledge my family, especially my parents and grandparents. Their unwavering love, belief in my abilities, and constant encouragement have been my driving force. Your support has been a constant source of motivation, and I am profoundly grateful for it.

Finally, I want to express my appreciation to all the friends, colleagues, and anyone who played a part, no matter how small, in this research journey. Your contributions, encouragement, and camaraderie have enriched this experience.

This thesis is a culmination of the collective support and encouragement of these wonderful individuals and groups. I am truly fortunate to have had them in my corner. Thank you all for being a part of this journey with me.

This work was funded by Fundação para a Ciência e a Tecnologia through strategic Funding UIDP/04423/2020 and UIDB/04565/2020 and through the project ATLANTIDA (ref. NORTE-01–0145-FEDER-000040), supported by the Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement and through the European Regional Development Fund (ERDF). This project has also received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 952374 and European Research Council (ERC) grant agreement No 759840.

# Resumo

As cianobactérias são um grupo de procariotas fotossintéticos com uma ampla distribuição geográfica e que se encontram numa grande diversidade de habitats, incluindo ambientes extremos. Para além disso, estes organismos têm sido reconhecidos como uma fonte promissora de metabolitos secundários ou também designados como produtos naturais (PN), com potencialidades aplicáveis em biotecnologia. À medida que a sequenciação de genomas se tornou mais acessível e com a evolução de técnicas bioinformáticas, a abordagem da *genome mining* ganhou um papel significativo na descoberta de novos produtos naturais. Isto é facilitado pelo facto de os genes biosintéticos estarem adjacentemente localizados em *clusters* de genes biossintéticos (BGCs, do inglês Biosynthetic Gene Clusters). O aumento de genomas de cianobactérias disponíveis em base de dados públicas enfatizou o potencial biossintético destes organismos em produzir produtos naturais; contudo grande parte desse potencial ainda está por desvendar.

Entre as diferentes categorias de compostos naturais produzidos por cianobactérias, os péptidos sintetizados diretamente pelo ribossoma e modificados pós-traducionalmente (RiPPs, do inglês *Ribosomally synthesized and Post-translationally modified Peptides*) têm atraído grande atenção devido às suas estruturas químicas e propriedades bioativas distintas. As cianobactinas, um grupo particular de RiPPs, são péptidos que podem ser cíclicos ou lineares, produzidos por uma gama diversificada de cianobactérias por meio de modificações pós-tradução de péptidos precursores. A prenilação, uma das modificações pós-traducionais mais comummente observadas nas cianobactinas, levam ao aumento da lipofilicidade dessas moléculas, permitindo uma interação mais efetiva com as membranas lipídicas. Além disso, as cianobactinas exibem notável variabilidade estrutural e possuem um amplo espectro de atividades biológicas, incluindo propriedades antibacterianas, antifúngicas, antivirais e anticancerígenas.

Neste sentido, o objetivo principal deste trabalho focou-se na deteção, isolamento, elucidação estrutural e atividade biológica de um novo grupo de cianobactinas preniladas da estirpe Oscillatoriales LEGE 16532. O ponto de partida deste trabalho foi a informação genómica desta estirpe sendo que o BGC responsável pela produção destes compostos foi identificado através de análise bioinformática. Neste trabalho, dez novos metabolitos foram identificados, incluindo cianobactinas cíclicas, mono e bis-preniladas, onde a prenilação ocorre no resíduo de tirosina, no sentido *forward*. De facto, até ao momento, poucas cianobactinas apresentaram bisprenilação, sendo que esta modificação no aminoácido da tirosina é pela primeira vez reportado neste trabalho. Assim, esta descoberta tem o potencial de servir como um instrumento biológico sintético, fornecendo propriedades semelhantes de fármacos a diversas pequenas moléculas. Globalmente, o estudo de RiPPs e cianobactinas representa uma área de investigação emocionante com aplicações promissoras em biotecnologia.

Palavras-chave: Cianobactérias, Produtos naturais, RiPPs, Biotecnologia, Biologia sintética, Bioatividade

# Abstract

Cyanobacteria are a group of photosynthetic prokaryotes with a wide geographical distribution and found in a great diversity of habitats, including extreme environments. These organisms have been recognised as a promising source of secondary metabolites, also known as natural products (NP), with potential applications in biotechnology. As genome sequencing has become more accessible and with the evolution of bioinformatics techniques, the genome mining approach has gained a significant role in the discovery of new NPs. This is facilitated by the fact that biosynthetic genes are adjacently located in Biosynthetic Gene Clusters (BGCs). The increase in cyanobacterial genomes available in public databases has emphasised the biosynthetic potential of these organisms to produce NPs, but much of this potential remains to be unravelled.

Among the different categories of natural compounds produced by cyanobacteria, peptides synthesised directly by the ribosome and post-translationally modified (RiPPs) have attracted great attention due to their distinct chemical structures and bioactive properties. Cyanobactins, a particular group of RiPPs, are peptides that can be cyclic or linear that are produced by a diverse range of cyanobacteria through post-translational modifications of precursor peptides. Prenylation, one of the most commonly observed post-translational modifications in cyanobactins, leads to an increase in the lipophilicity of these molecules, allowing them to interact more effectively with lipid membranes. In addition, cyanobactins exhibit remarkable structural variability and possess a broad spectrum of biological activities, including antibacterial, antifungal, antiviral and anticancer properties.

In this sense, the main objective of this work focused on the detection, isolation, structural elucidation and biological activity of a new group of prenylated cyanobactins from the strain Oscillatoriales cyanobacterium LEGE 16532. The starting point for this work was the genomic information of this strain and the BGC responsible for the production of these compounds was identified through bioinformatic analysis. In this work, ten new metabolites were identified, including cyclic, mono- and bis-prenylated cyanobactins, where forward prenylation occurs on the tyrosine (Tyr) residue. In fact, to date, few cyanobactins have shown bis-prenylation, and this modification on the Tyr is reported for the first time in this work. Thus, this discovery has the potential to serve as a synthetic biological tool, providing drug-like properties to various small molecules. Overall, the study of RiPPs and cyanobactins represents an exciting area of research with promising applications in biotechnology.

Keywords: Cyanobacteria, Natural Products, RiPPs, Biotechnology, Synthetic Biology, Bioactivity

# Table of Contents

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| ACN | Acetonitrile |
| Arg | Arginine |
| BGC | Biosynthetic Gene Cluster |
| BLAST | Basic Local Alignment Search Tool |
| Cys | Cysteine |
| EFI-EST | Enzyme Function Initiative - Enzyme Similarity Tool |
| FAS | Fatty Acid Synthase |
| GNAT | Gcn5-related N-acetyltransferase |
| HepG2 | Hepatocellular Carcinoma Cell Line |
| IPA | Isopropanol |
| LC-HRMESIMS | Liquid Chromatography - High-Resolution Mass Spectrometry |
| LC-HRMESIMS/MS | Liquid Chromatography - High-Resolution Mass Spectrometry/Mass Spectrometry |
| LEGE-CC | Blue Biotechnology and Ecotoxicology Culture Collection |
| MeOH | Methanol |
| MTT | 3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide |
| NPs | Natural products |
| NRP | Nonribosomal Peptides |
| NRP-PK | Nonribosomal Peptide Synthetase-Polyketide Synthase |
| PKS | Polyketide Synthase |
| PN | Produtos naturais |
| Ptase | Prenyltransferase |
| PTM | Post-Translational Modification |
| RiPPs | Ribosomally Synthesized and Post-translationally Modified Peptides |
| RS | Recognition sequence |
| Ser | Serine |
| SSN | Sequence Similarity Network |
| Thr | Threonine |
| Tyr | Tyrosine |

# 1. Introduction

## 1.1. Cyanobacteria

Cyanobacteria are a diverse group of photosynthetic prokaryotes that play a crucial role in various ecosystems. With their ability to convert solar energy into chemical energy through oxygenic photosynthesis, cyanobacteria have significantly shaped the Earth's atmosphere and influenced the evolution of life forms throughout history [1, 2].

These organisms exhibit a remarkable adaptability to different environmental conditions and can be found in a wide range of locations, including freshwater bodies, marine ecosystems, terrestrial habitats, and even extreme environments such as hot springs, deserts, glaciers, and hypersaline environments [3, 4]. Furthermore, cyanobacteria are known to form symbiotic relationships with various other organisms, including algae, plants, fungi, and marine invertebrates, influencing nutrient dynamics and community structures [5-8]. For many decades, cyanobacteria were considered algae due to their algae-type photosynthetic apparatus. However, by using electron microscopy, the researchers confirmed that these photosynthetic organisms were part of Prokaryotes, because of the similarities between the cell wall and ribosomal structure of others Gram-negative bacteria [9]. As consequence, their classification has been divided and attributed based on both the Bacteriological and Botanical Codes, leading to ongoing debates regarding the correct nomenclature system [10, 11].

In recent years, progresses in molecular techniques and genomic analysis have revolutionized the approach to cyanobacteria taxonomy, providing new insights into their phylogeny [9]. The latest taxonomic approach, using morphological, ecological, and genetic data, has classified cyanobacteria into various distinct orders. The first order, Gloeobacterales, comprises cyanobacteria lacking thylakoids, while Synechococcales, Spirulinales, Nodosilineales, Prochlorotrichales, Oculatellales, Leptolyngbyales, Geitlerinematales, Desertifilales, Coleofasciculales consist of unicellular and/or filamentous genera with parietal thylakoids. Aegeococcales, Gloeomargaritales, Acaryochloridales includes rod shaped unicellular species with parietal thylakoids. Chroococcales, and Pleurocapsales represent coccoid forms, and Oscillatoriales include filamentous species without cell differentiation. Chroococcidiopsidales and Thermostichales are mostly found in extreme habitats. Finally, the Pseudanabaenales, Nostocales order encompasses filamentous cyanobacteria featuring heterocysts and akinetes (Figure 1) [12, 13]

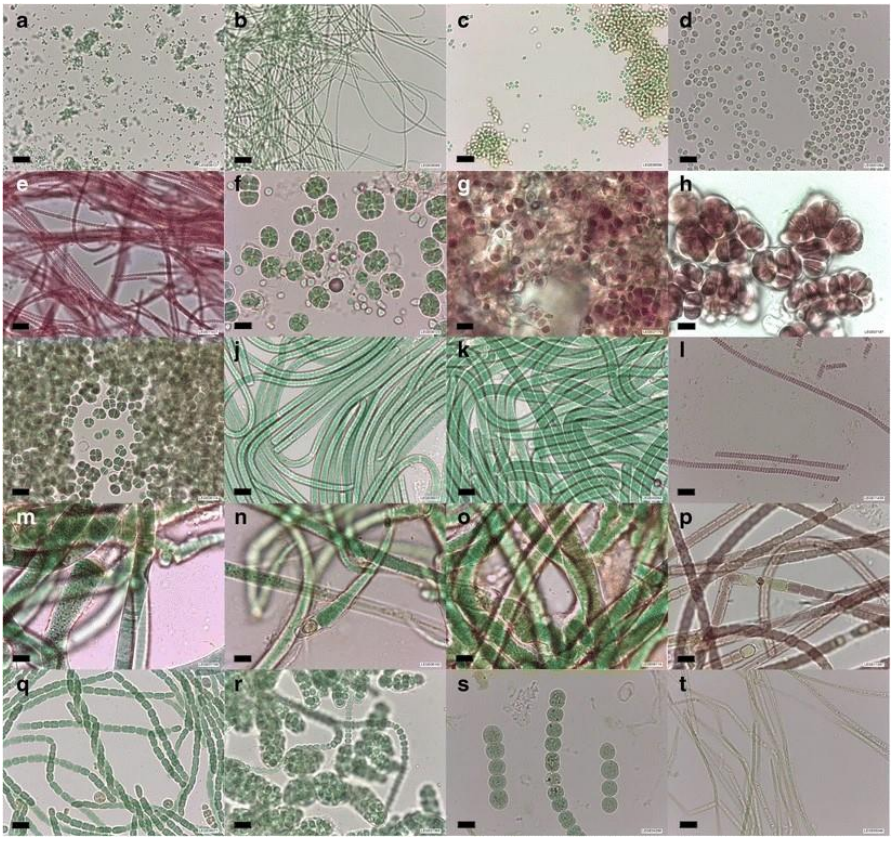



Figure 1 Example of morphological diversity among cyanobacterial strains. Strains belong to the orders: a–b Synechococcales, c–e Chroococcales, f Chroococcidiopsidales, g–i Pleurocapsales, j–l Oscillatoriales, and m–t Nostocales [14].

Heterocysts and akinetes are specialized cell variants in cyanobacteria that serve important functions. Heterocysts are involved in nitrogen fixation and are responsible for spatially separating the oxygen-sensitive process of nitrogen fixation from oxygen-producing photosynthesis in vegetative cells [15]. Akinetes, on the other hand, are spore-like dormant cells that are resistant to unfavourable environmental conditions. They play a crucial role in stress survival and recolonization after long periods of unfavourable conditions [16].

## 1.2. Natural Products Discovery Approaches

Throughout history, nature has drawn inspiration for new drugs and humans have extracted a variety of secondary metabolites from plants, fungi, and bacteria [17]. Antibiotics are one of the most successful of these products, these substances inhibit the growth of bacteria or fungi, by killing them or preventing their reproduction. The bulk of the earliest antibiotics, such as penicillin and streptomycin, were obtained from fungi and bacteria [18]. Unlike primary metabolites, secondary metabolites do not contribute directly to the basal metabolism of the organism, but rather operate as crucial elements in attracting, repelling, or killing other species, increasing their chances of survival in the

environment. Therefore, secondary metabolites are also known as specialized metabolites, since they are small molecules that perform a very specific function that is associated with the ecosystem [19].

Currently, the majority of efforts to discover natural products (NPs) can be categorized into two main strategies: top-down and bottom-up approaches (Figure 2) [20]. The top-down strategy involves starting at the organism level, without any prior knowledge of the genome, while the bottom-up strategy begins at the genomic level by identifying specific biosynthetic gene clusters (BGCs) of interest. These clusters consist of a group of genes physically located together on the genome, encoding the biosynthetic pathway for the production of a NP.

Traditionally, natural product discovery has relied on bioactivity-guided screening, which encompasses the cultivation of organisms, chemical extraction, bioassays, isolation, and elucidation of the chemical structure. This top-down approach has proven successful in the discovery of numerous compounds, without the need for genomic information. However, finding novel NPs had become more difficult, leading to the re-isolation of known NPs. Recent technological and scientific advancements, such as genome mining, genetic engineering strategies, and improvements in microbial culturing techniques, have revitalized the field of NP-based drug discovery [21]. In the post-genomics era, next-generation sequencing has become increasingly accessible due to a significant decrease in the cost per sequenced base, allowing for the exploration of a large number of genomes [22]. Genome mining has thus emerged as a valuable method for screening and analyzing sequenced genomes, specifically targeting NP BGCs. This approach has unveiled numerous cryptic BGCs, revealing the vast extent of chemical diversity associated with NPs [23, 24]

In addition, the accumulated understanding concerning conserved biosynthetic machineries of NPs, together with the significant advancements in bioinformatic platforms such as antiSMASH, RiPPminer, and BAGEL4, has expedited the detection of BGCs and the forecasting of the encoded NPs [25-27].

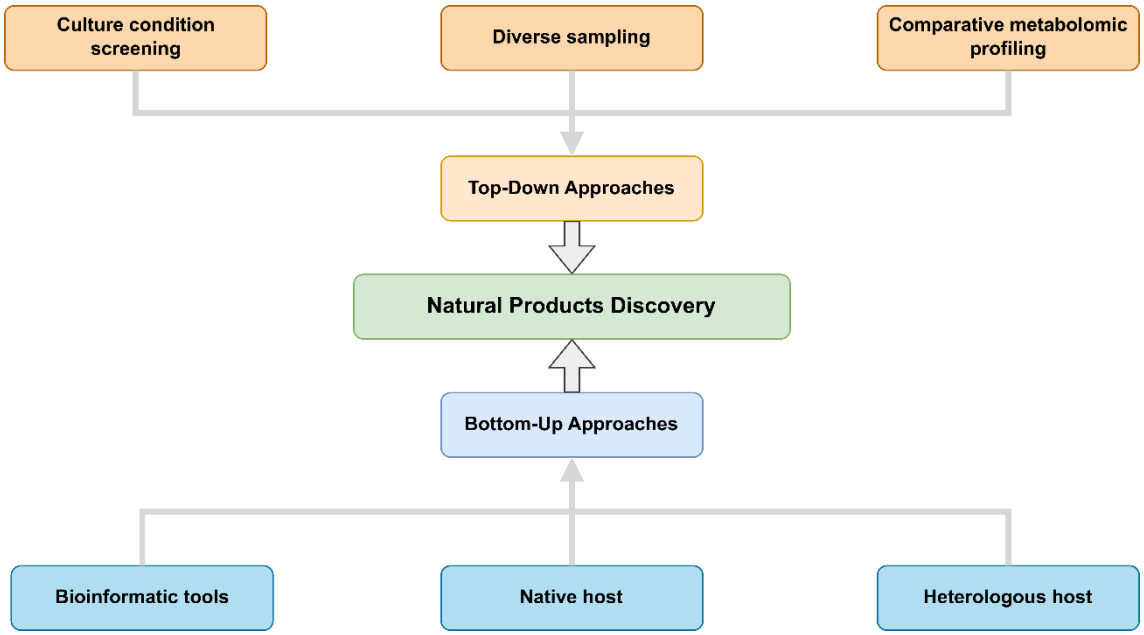



Figure 2 Overview of the recent strategies applied in NPs discovery (adapted from Lou *et al*. [20]).

In general, strategies involving the mining of genomes provide flexible approaches for the discovery of NPs. Additionally, these strategies facilitate the deliberate activation of dormant BGCs and establish a direct connection between the compound and its corresponding BGC. This linkage simplifies subsequent combinatorial biosynthesis endeavors [28-30].

## 1.3. Cyanobacterial Natural Products

Cyanobacteria are known for their production of cyanotoxins, which are toxic secondary metabolites. These cyanotoxins have been reported in harmful cyanobacterial blooms in aquatic environments worldwide [31]. These blooms can have ecological impacts and pose risks to human and animal health [31, 32]. In addition to cyanotoxins, cyanobacteria also produce a wide range of bioactive compounds with therapeutic potential. To this day, the majority of compounds have been discovered and isolated from the orders Oscillatoriales and Nostocales, while only a few compounds from the order Pleurocapsales have been described [33]. Out of the nearly 800 compounds that have been reported from marine cyanobacteria, approximately half of them come from the order Oscillatoriales (Figure 3) [34]. Within this order, the genus *Moorena* (formerly known as *Lyngbya* [35]) has proven to be the most abundant in terms of compound production, with the genera *Oscillatoria* and *Symploca* following closely behind [36].

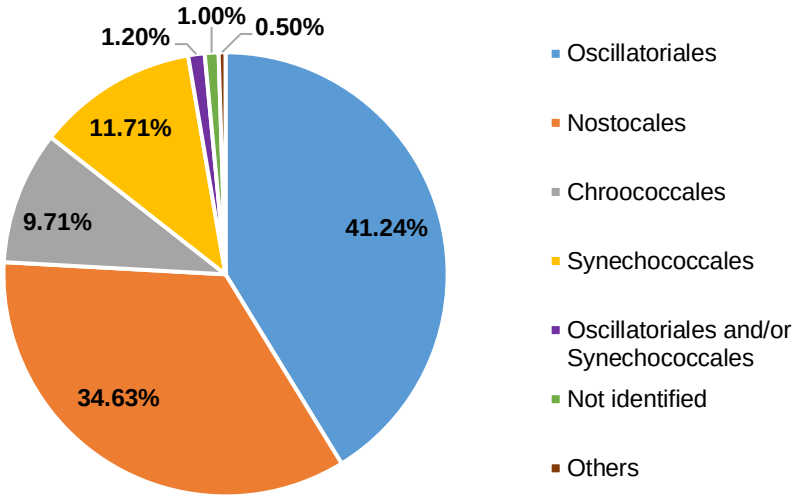



Figure 3 Distribution of compound families across taxonomical families. Adapted from Demay *et al.* 2019 [34].

The extensive studies conducted on this subject have primarily focused on the more complex species, particularly the filamentous forms that thrive in high densities within tropical and subtropical regions. These species are easily collected and cultured in large quantities, providing ample biomass for chemical investigations [37]. An additional factor that contributes to the preference for studying filamentous and colonial forms is the observation that they tend to have larger genomes compared to unicellular free-living species. These larger genomes are better equipped to accommodate the polyketide (PK) and nonribosomal peptide (NRP) pathways, which are known to produce a wide variety of bioactive compounds [38].

In a recent report, Demay *et al*. classified 260 families into 10 distinct chemical classes, comprising alkaloids, depsipeptides, lipopeptides, macrolides/lactones, peptides, terpenes, polysaccharides, lipids, polyketides, and several other compounds. Among them, peptides are the most representative family of secondary metabolites, constituting approximately one-third of the produced compounds (Figure 4) [34]. Notably, the vast diversity of peptides can be attributed to the presence of distinct biosynthetic pathways involved in their formation, such as Non-Ribosomal Peptide Synthase (NRPS) and Ribosomally synthesized and Post-translationally modified Peptides (RiPPs) [24].

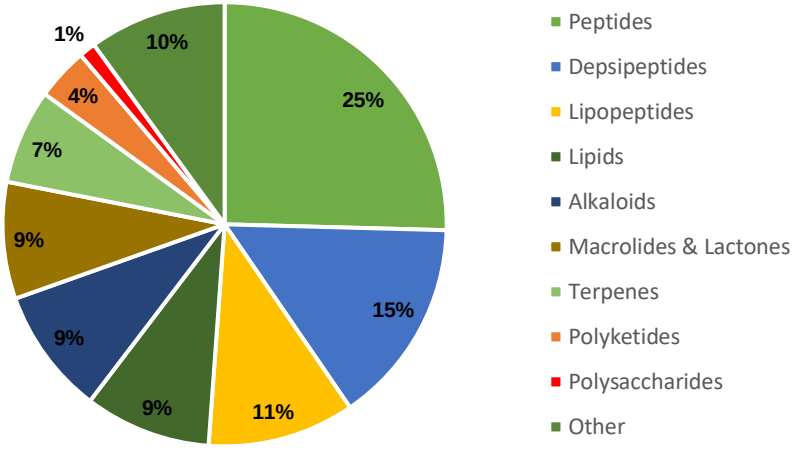



Figure 4 Cyanobacterial Metabolite Families Classified by Chemical Classes: Adapted from Demay *et al*. 2019 [36].

These metabolites exhibit a diverse range of bioactivities such as antibacterial, antifungal, anticancer, immunosuppressive, anti-inflammatory and antimycobacterial properties making them a valuable source of new drugs and templates for drug development [34]. Among these compounds, curacin and dolastatin stand out for their remarkable cytotoxic activity, making them promising candidates for cancer treatment. Curacin-A, a hybrid Non-Ribosomal Peptide-Polyketide (NRP-PK) isolated from *Lyngbya majuscula*, suppresses cell growth and mitotic activity by binding to tubulin [39]. Similarly, Dolastatin 10 is synthesised by *Symploca* sp. VP642 and induces cell apoptosis by binding the tubulin, arresting the cell cycle in the G2/M phase [40, 41]. The unique mechanisms of action of these compounds have led to their investigation in clinical trials as potential chemotherapeutics. Additionally, synthetic analogs of Dolastatin 10 have paved the way for the development of antibody-drug conjugates (ADCs) in cancer treatment [41]. Several ADCs, such as brentuximab vedotin, enfortumab vedotin, polatuzumab vedotin, and belantamab mafodotin, have been approved for specific cancer types, showcasing the successful translation of cyanobacteria-derived compounds into clinically effective drugs [42].

## 1.4. Ribosomally Synthesized Lipopeptides

The RiPPs are an important and diverse group of NPs, produced by all three branches of life [43]. In fact, over the past two decades, the rapid advancement of genome mining tools, along with improvements in genomic and metabolomic research areas, has uncovered an unexpected abundance of novel RiPP BGCs [44]. Due to their simple biosynthesis, RiPPs are capable of generating a vast structural diversity through extensive post-translational modifications of the precursor peptide [45].

RiPP biosynthesis (Figure 5) begins with a ribosomally synthesized precursor peptide, consisting of an N-terminal leader region and a C-terminal core region. The leader region is responsible to recruit different post-translational modifications (PTMs) enzymes, while the C-terminal core region encodes a sequence of amino acids that builds the structural backbone of the final peptide [46]. This precursor peptide is nearby processing enzymes that recognize the leader and act on the core for producing modified peptides. Eventually, the modified core peptide is liberated from the altered core by peptidases, resulting in the separation of the leader region. Subsequently, these modified core peptides can either exit the cell as mature peptides or undergo further tailoring modification facilitated by PTM enzymes independent of the leader peptide, before being prepared for export [45].

Moreover, the biosynthesis of certain classes of RiPPs, such as cyanobactins, dikaritins, orbitides, amatoxins, cyclotides, and lyciumins, involves the presence of multiple core peptides and recognition sequences (RSs) within the precursor peptide [47]. This unique arrangement allows for the engagement of distinct biosynthetic enzymes that recognize different segments of the leader/follower peptide. These RSs serve as specific cues for individual enzymes to catalyze precise chemical modifications on the core peptide. Notably, these modifications can occur at various locations along the peptide chain, with some cases even presenting recognition sequences at the atypical N-terminal or C-terminal positions. Furthermore, the RiPPs BGCs may harbor multiple core peptides, potentially separated by RSs, resulting in the production of multiple RiPP products from a single peptide precursor [45, 47]. This complex interplay of recognition sequences, core peptides, and enzymatic transformations contributes to the diverse array of RiPPs classes.

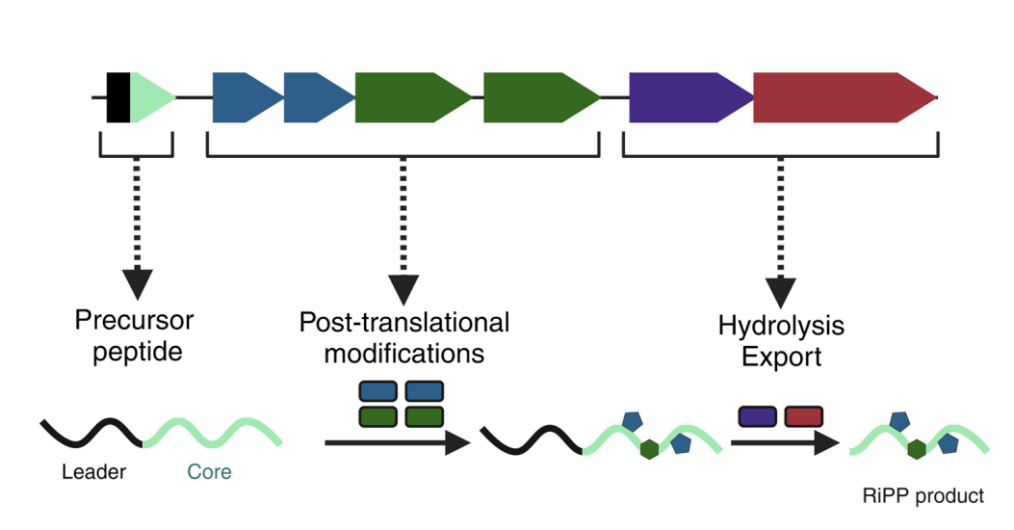



Figure 5 General RiPP biosynthesis pathway, adapted from Vogt et al. 2019 [48].

Lipidation of peptides and proteins is an established post-translational modification [49]. Lipopeptides are a class of NPs that have gathered significant attention due to their diverse ecological role, biological activities, and potential applications in biotechnology and synthetic biology. These molecules are composed by a hydrophobic lipid moiety of various lengths linked to a hydrophilic peptide that is often cyclic and/or linear [50]. The combination of these two structural features contributes to their unique properties and functions. Their amphiphilic properties facilitate his anchoring to the biological membranes, and the peptide head, which presents specific functional group is responsible for their bioactivity, allowing lipopeptides to exhibit remarkable bioactivities, such as surfactant, antifungal, antimicrobial, and anticancer, among other functionalities [51].

Cyanobacterial lipopeptides are synthesized through complex secondary metabolic pathways and the lipid residue can be incorporated in a number of ways [50]. The lipopeptides are typically extended through the successive additions of both proteinogenic and nonproteinogenic amino acids by nonribosomal peptide synthetases (NRPSs) [52]. NRPSs are multimodular enzymes that catalyze the biosynthesis of diverse peptides and the lipidic region is synthesized independently by either the fatty acid synthase (FAS) or dedicated polyketide synthetases (PKS) [53, 54]. Examples of lipopeptides produced by Cyanobacteria includes microginins [55], hassallidin [56, 57], puwainaphycin [58] and desmamides [59].

Recently, biochemical and genome-guided efforts have also identified several lipidated peptides from Cyanobacteria that fall into the category of RiPPs. Hubrich and co-workers have recently reported a new family of RiPP-derived lipopeptides, the

selidamides [60]. Kamptornamide and nostolysamides are the first members of the selidamide family of RiPP-derived lipopeptides that is characterized by fatty acyl moieties attached to the side chain of (hydroxy)ornithine or lysine, respectively. Heterologous expression experiments suggest that every peptide of the selidamide family is selectively modified by a fatty acid with fixed chain length by members of GCN5-related N-acetyltransferase (GNAT)-like family of peptide maturases [60].

Moreover, the most widely studied of ribosomally derived lipopeptides are the cyanobactin class of RiPP NPs which will be described in the next section.

## 1.5. Cyanobactins

Cyanobactins are diverse RiPP natural products produced by symbiotic and free-living cyanobacteria [61]. In fact, cyanobactin BGCs are found throughout the Cyanobacteria phylum, and it is estimated that 30% of all cyanobacteria encode a cyanobactin pathway [62], being one of the most common biosynthetic routes in cyanobacteria.

Although cyanobactins were described earlier as macrocyclic peptides which contain heterocyclized serine (Ser), threonine (Thr) and/or cysteine (Cys) (or their oxidized derivatives) and/or prenylated amino acids [61], recent discoveries have broadened this definition to include short linear cyanobactins and cyanobactins lacking heterocyclic residues, expanding the structural diversity of this family [63, 64]. Examples of cyanobactins (Figure 6) include the patellamides, trunkamides, aeruginosamides and anacyclamides [63, 65-67], among others, and exhibit bioactivities such as cytotoxic, antibacterial and antimalarial activity [68].

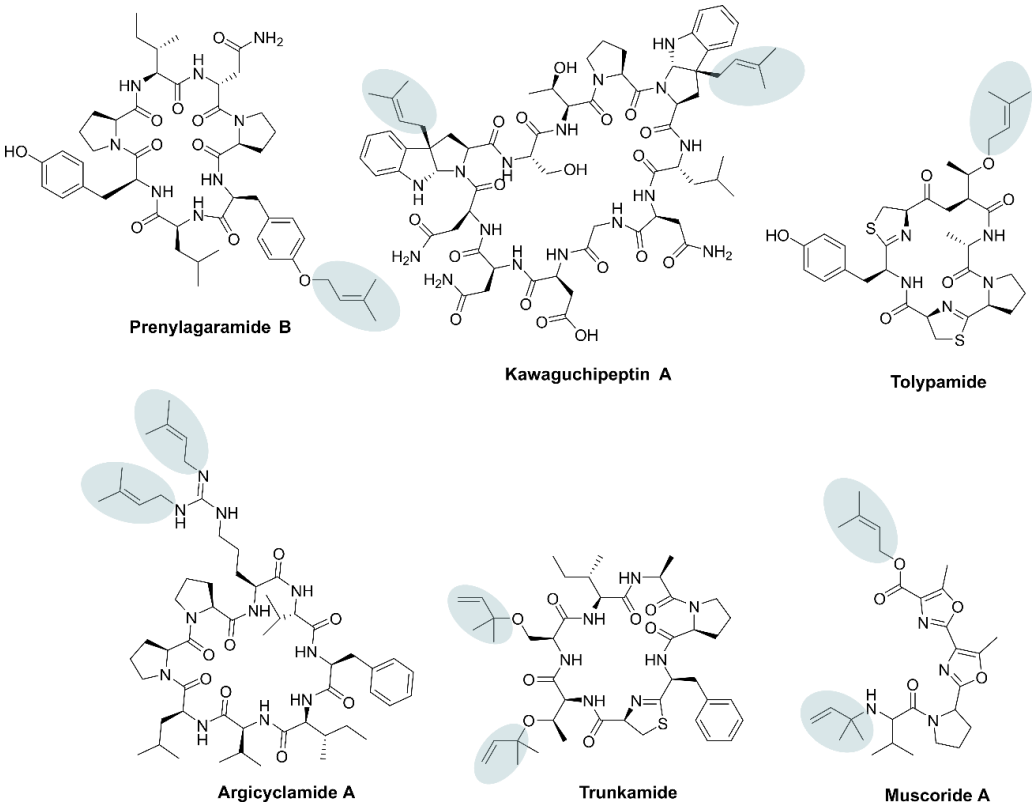



Figure 6 Chemical diversity of cyanobactins. Prenyl groups are highlighted on grey. Adapted from Zhang *et al* 2023 [69].

Cyanobactin BGCs (Figure 7) exhibit a length range of 8 to 19 kb and encode two proteases (“A” and “G” genes). The precursor peptide, encoded by the “E” gene contains the cyanobactin core amino acid sequence and serves as a substrate for post-translational modifications. Unlike other precursor peptides belonging to different RiPP classes, the cyanobactin precursor peptides can be composed of numerous cassettes, each with a core peptide that generates mature cyanobactin flanked by two highly conserved recognition sequences (RS), RSII and RSIII [70]. These sequences are accountable for directing the biosynthetic enzymes. The major division between cyanobactin subtypes is those with heterocyclase encoded by the “D” gene, and those that lack this enzyme. When the heterocyclase is present, this enzyme recognizes RSI (only present in cyanobactin BGCs harboring the heterocyclase) in the precursor peptide and converts Cys/Ser/Thr to azoline heterocycles [62]. Then, the C-terminal protease, encoded by the “G” gene, which present a macrocyclase domain exclusive to the G-protease cleaves off the RSIII and catalyzes the N-C macrocyclization (except on the linear cyanobactins). Cyanobactins BGC also encodes two genes “B” and “C” which encodes short proteins of unknown function in cyanobactin biosynthesis. Finally, tailoring reactions such as oxidation of the heterocycles, prenylation (carried out by the

prenyltransferase encoded in the "F" gene), methylation or phosphorylation (unpublished data) take place. [64, 68]

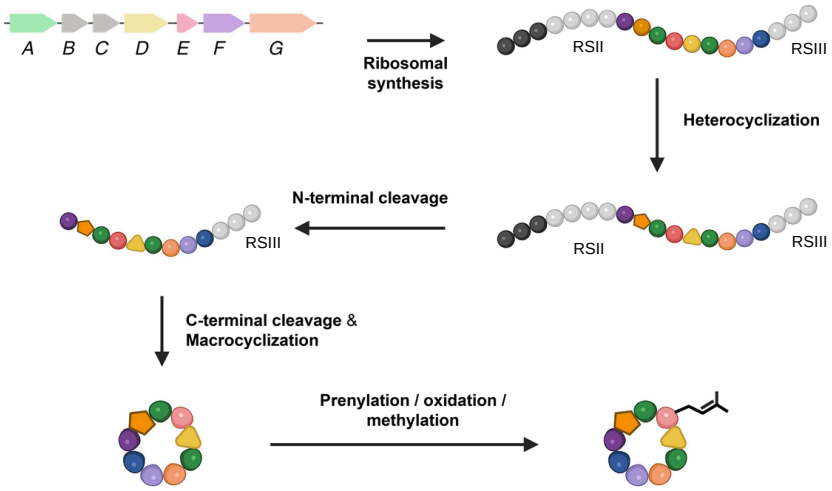



Figure 7 General scheme of cyanobactin biosynthesis, Patellamide like. The PTM is catalyzed by a prenyltransferase, led by the F gene. Adapted from [71].

In the biosynthesis of cyanobactins, prenyltransferases (PTases) play a crucial role in replacing the pyrophosphate group within the prenyl donor molecule with a carbon, nitrogen, or oxygen atom found in the peptide substrate. While there have not been any mechanistic investigations into cyanobactin prenyltransferases, this type of reaction could occur through one of two mechanisms [69, 72]. One possibility is an associative mechanism, where the acceptor nucleophile directly attacks the prenyl donor. Alternatively, a dissociative mechanism could be at play, involving an allylic carbocation intermediate [72]. This dissociative mode of action has been suggested based on extensive kinetic isotope experiments conducted on dimethylallyltryptophan synthase [73]. Consequently, it's reasonable to hypothesize that prenylation by cyanobactin PTases may also proceed through a dissociative mechanism [69, 74]. In this dissociative mechanism, PTases initially eliminate the pyrophosphate group from the prenyl donor, generating an allylic carbocation intermediate within the active site. Subsequently, this allylic carbocation is vulnerable to attack by an acceptor nucleophile, such as the phenolic hydroxy group found on Tyr, leading to the formation of prenylated products. The cyanobactin PTases exhibit selectivity, catalyzing prenylation in either the 'forward' or 'reverse' direction (Figure 8), where the primary carbon (C1) or the tertiary carbon (C3) of the prenyl donor attaches to the acceptor [69].

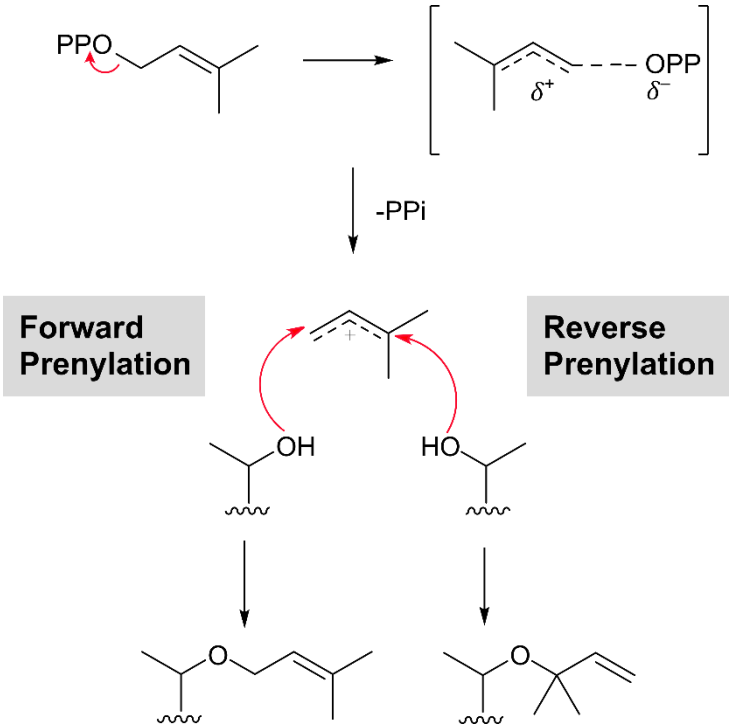



Figure 8 the prenylation mechanism executed by cyanobactin PTases, involving the attachment of a prenyl group to a tyrosine residue in both the forward and reverse direction (adapted from [69]).

The first characterized member of this family was LynF from *Lyngbya aestuarii*, which prenylates Tyr to generate reverse Tyr-O-prenylated peptides. This product undergoes a spontaneous Claisen rearrangement to yield forward Tyr-C2-prenylated peptides [74, 75]. In contrast, PagF from *Planktothrix agardhii* NIES-596 catalyzes forward Tyr-O-prenylation, and SphF from sphaerocyclamide is postulated to be another forward Tyr-O-prenyltransferase [76-78].

Furthermore, certain PTases also modify Tryptophan (Trp), Serine (Ser) and Threonine (Thr) residues in cyclic cyanobactin precursors. The KgpF from kawaguchipeptin A and AcyF from anacyclamide BGC which are responsible for Trp-C3-prenylation and reverse N-prenylation, respectively [63, 79-81]. Representative examples of Ser/Thr-O-prenylation include patellins and trunkamide A, which are produced by the trunkamide BGC using TruF1 to perform reverse on Ser/Thr [82, 83]. Most recently, TolF was discovered through genome mining and characterized as a member of the subfamily of Ser/Thr PTases, this enzyme is involved in tolypamide biosynthesis and is capable of performing a forward Thr-O-prenylation [84].

Similarly, two PTases, AgcF and AutF, were reported as responsible for the prenylation of arginine (Arg). AgcF, from the argicyclamide biosynthetic pathway was reported to catalyze the mono- and bis-N-prenylation of the guanidine moiety of L-Arg

and both compounds were isolated from *Microcystis aeruginosa* NIES-88 [85]. On the other hand, AutF targets the guanidinium moiety in arginine and homoarginine leading to the prenyl-D-Arg-containing autumnalamide A produced by *Phormidium autumnale* CCAP1446/10 [86]. It is clear that genome-guided approach is a valuable and effective strategy to discovery new enzymatic reactions in cyanobactin biosynthesis. Another example is the identification and characterization of LimF from *Limnothrix* sp. CACIAM 69d which was shown to catalyze His-C geranylation and Tyr-O geranylation [85, 87]. Additionally, there are also PTases involved in the prenylation of the N and C termini of some linear cyanobactins, with AgeMTPT being the first characterized member of this subgroup [88]. AgeMTPT is a didomain bifunctional enzyme consisting of a cyanobactin PTase fused to a methyltransferase (MT), and it catalyzes both N-terminal N-prenylation and C-terminal O-methylation [88]. MusF1 and MusF2 from *Nostoc* sp. PCC-7906 were identified as the enzymes responsible for the prenylation of muscoride A, with MusF1 catalyzing the forward O-prenylation of the C-terminal carboxylate and MusF2 being responsible for the reverse N-prenylation of the N-terminal amino group [89]. Lastly, some PTases use GPP instead of DMAPP as the prenyl donor and install C10-prenyl units, geranyl group, such as PirF from the piricyclamide biosynthesis that catalyzes Tyr-O-geranylation [90]

This post-translational modification is thus a crucial mechanism that enhances the lipophilicity of molecules, enabling them to interact more effectively with lipid membranes, which could be useful to explore since this could improve the biological activities of the secondary metabolites.

## 1.6. Combinatorial Biosynthesis and Chemical Diversity of RiPPs

The major goal of synthetic biology is to design and construct biological devices and systems for useful purposes [91]. This field applies engineering principles such as standardization, modularity, and abstraction to dismantle and reassemble biological cells and processes, enabling the creation of novel systems that can perform specific functions [92]. Additionally, the development of new tools, such as Gibson assembly and Direct Pathway Cloning combined (DiPaC) with Sequence - and Ligation-Independent Cloning (SLIC) have contributed to the field of combinatorial biosynthesis by providing efficient methods for cloning and expressing natural product BGCs [93]. These methods streamline the process of capturing and assembling gene clusters, allowing their manipulation and heterologous expression in different hosts.

In recent years, the RiPPs research field has revealed a significant diversity of biosynthetic enzymes and a vast range of new machineries [47]. These studies have provided a huge variety of tools for engineering innovative biosynthetic pathways and create novel and “new-to-nature” peptides [94]. RiPP NPs are particularly suitable for combinatorial biosynthesis and engineering, due to their simple biosynthesis, genetically encoded substrates, and modular post-translational modifications [95].

The leader peptide and core peptide can be separated as distinct modules, allowing RiPP modification enzymes to exhibit both substrate specificity and promiscuity [96]. Therefore, Schmidt *et al*. have demonstrated that the combination of enzymes from varied pathways, *in vitro*, can produce novel natural and nonnatural cyanobactins with customized post-translational modifications [96]. The modular logic of RiPP pathways also enables the creation of numerous core peptide libraries through saturation mutagenesis and site-directed mutagenesis of the core peptide [94]. These libraries serve as a first step in directed evolution experiments, enhancing the drug-like properties of the new chemical scaffolds and their bioactivities [97, 98]. Furthermore, apart from the precursor peptide, the engineering of enzymes involved in the maturation of the RiPP is also being explored. Zhao and colleagues elucidated a technique of high-throughput screening in order to acquire variations of the Nisin dehydratase (NisB) that possess customized substrate specificity. This research showcased an influential approach for identifying lanthipeptide modification enzymes with adjusted substrate specificity, which has the potential to augment the synthesis of lanthionine-stabilized therapeutic peptides with enhanced attributes, including heightened stability and improved target specificity [99].

The manipulation of BGCs by the combining and exchanging the biosynthetic enzymes from diverse pathways, could lead to the development of entirely novel structures. Recently, a demonstration of the artificial combination of RiPP enzymes from different classes has been carried out [100]. Also, Burkhart *et al*. have designed leader peptides that could facilitate recognition and processing by numerous enzymes from unrelated RiPP pathways [101]. They combined a thiazoline-forming cyclodehydratase with enzymes from the sactipeptide and lanthipeptide families, resulting in the creation of novel hybrid RiPPs that are new to nature. In particular, they merged an azoline-forming cyclodehydratase (HcaD/F) with sactipeptides (AlbA) and lanthipeptide synthetases (NisB/C and ProcM) to produce a thiazoline−sactipeptide, thiazolino-lanthipeptide and thiazolino- lanthionine hybrids, respectively. Moreover, they also

synthetized hybrids with leader-independent tailoring enzymes such as MibD and NpnJA [101].

The combinatorial biosynthesis of RiPPs presents a significant opportunity for the development of novel peptides or drug frameworks that do not currently exist or have widespread antibiotic resistance genes. Furthermore, these enzymes are readily adaptable for synthetic efforts aimed at increasing diversity, making them appealing catalysts for the creation of valuable peptidic therapeutics in the field of synthetic biology.

## 1.7. Aim Of This Work

CIIMAR hosts an important Culture Collection of Cyanobacteria, the Blue Biotechnology and Ecotoxicology Culture Collection (LEGE-CC), which comprises more than 600 cyanobacterial strains from different ecosystems and locations; dozens of these strains are unique within the phylum [14]. It is now well-established that LEGE-CC strains have the potential or proven capacity to produce a variety of biotechnologically interesting NPs [14]. However, only a small fraction of the diversity present at LEGE-CC was studied and chemodiversity and bioactivity potential of a large number of unique strains remains unknown. Considering this, the main goal of this work was to characterize a cyanobactin pathway detected in Oscillatoriales cyanobacterium LEGE 16532 genome and the characterization of the final products of this pathway.

The specific aims were:

1. Annotation the cyanobactin BGC identified in Oscillatoriales cyanobacterium LEGE 16532;
2. Investigation of the sequence-function and chemical space of cyanobactin prenylation;
3. Isolation and structural elucidation of the cyanobactins;
4. Characterization of the bioactivity profile of the isolated cyanobactins, using several bioassays to reveal potential pharmaceutical or biotechnological applications.

# 2. Material and Methods

## 2.1. Cyanobacterial culture conditions

The Oscillatoriales cyanobacterium LEGE 16532 strain (Figure 9) used in the current work, was isolated from a water tank in Serra do Monchique and was kindly provided by Blue Biotechnology and Ecotoxicology Culture Collection (acronym LEGE-CC), a curated collection of cyanobacteria and microalgae located at CIIMAR [14]. This strain was cultured in 50 mL of Z8 media [102], at 25 °C, with shaking, under a 14h/10h light (10−30 µmol photons s−1 m−2)/dark cycle. Biomass from the stationary phase was harvested by centrifugation (5000g, 10 min, 4 °C, Gyrozen 2236R) and frozen for genomic DNA extraction or small-scale methanolic extraction.

Large-scale culturing of this strain was carried out in 20 L of polycarbonate carboys (Nalgene) for a total of 60 L and cells were harvested during late exponential/early stationary phase through filtration, immediately frozen and freeze-dried until further usage.

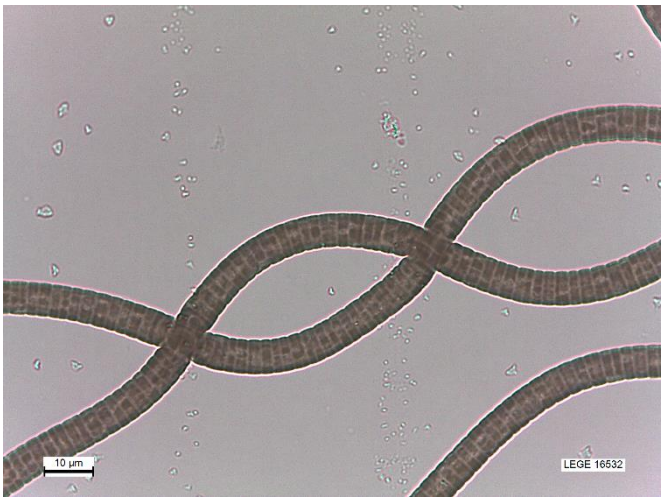



Figure 9 Microphotography of Oscillatoriales cyanobacterium LEGE 16532.

## 2.2. Genomic analysis of Oscillatoriales cyanobacterium LEGE 16532

### 2.2.1. Genome sequencing and annotation

The cyanobacterium Oscillatoriales cyanobacterium LEGE 16532 was sequenced through a specialized sequencing service (enhanced genome service) provided by the company MicrobesNG, using a combination of short-read Illumina and

long-read Nanopore sequencing technologies. Fresh biomass, harvested after a 15-day cultivation period, were sent to the service provider. Following gDNA extraction and sequencing, raw data was submitted to a bioinformatics pipeline. Briefly, identification of the closest reference genomes for reading mapping was done using Kraken 2 [103] BWA-MEM was used to check the quality of the reads and de novo assembly was performed using SPAdes [104]. Since the culture was not axenic, the genomic data was treated as a metagenome: the retrieved contigs were analyzed using the binning tool MaxBin 2.0 [105]. This procedure yielded a nearly complete genome with an estimated size of 9.64 Mb, distributed among 12 contigs. The genome was then submitted to antiSMASH v6.0 [106] using relaxed detection settings and all features selected to detect the secondary metabolites BGCs within this strain. The *osc* BGC from Oscillatoriales cyanobacterium LEGE 16532 was manually annotated using BlastP analysis to identify the homology and function of the open reading frames (ORFs). In order to predict the core peptides, the precursor peptides of the *osc* BGC were analysed using RiPPMiner Genome [26] and BAGEL4 [27].

#### 2.2.2. Generation of Sequence Similarity Network of prenyltransferase OscF

Sequence similarity networks (SSNs) were generated using the using the online Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST) [107], following a “Sequence BLAST (Basic Local Alignment Search Tool)” of OscF as input [108] and including the InterPro family of prenyltransferases (IPR031037) in the analysis. The alignment score threshold used was $10^{-86}$ as performed in previous works [84]. The generated network was visualized and analyzed using Cytoscape 3.7.1 for prenyltransferases occupying unique sequence-function space.

## 2.3. Chemical analysis of Oscillatoriales cyanobacterium LEGE 16532

### 2.3.1. General Experimental Procedures

LC−HRESIMS and LC−HRESIMS/MS analyses were conducted using an UltiMate 3000 UHPLC system (Thermo Fisher), which consisted of an LPG-3400SD pump, a WPS-3000SL autosampler, and a VWD-3100 UV/VIS detector. These instruments were connected to a Q Exactive Focus Hybrid Quadrupole Orbitrap mass spectrometer controlled by Q Exactive Focus Tune 2.9 and Xcalibur 4.1 software, also from Thermo Fisher Scientific. LC−HRESIMS data were obtained in full scan positive mode at a resolution of HRMS with a scan range of m/z 200-3000, with a capillary voltage of -3.5 kV, a capillary temperature of 262.5°C, and a sheath gas flow rate of 50 units.

LC−HRESIMS/MS analysis was performed using the same system, and fragmentation was achieved at 17,500 fwhm using higher-energy collision dissociation (HCD) with a 3.0 a.m.u. isolation window and a normalized collision energy of 35. Cyanobacterial extracts were separated using an ACE(r) UltraCoreTM SuperC18 column with a particle size of 2.5 µm, a pore size of 95 Å, and dimensions of 75 × 2.1 mm. The column oven was maintained at a temperature of 40 °C. Samples were eluted at 0.35 mL $min^{-1}$ using a linear gradient from 99.5% solution A (95% water, 5% methanol (Me, and 0.1% Formic acid v/v) and 0.5% solution B (95% isopropanol (IPA), 5% MeOH, and 0.1% Formic acid, v/v) to reach 10% solution B over 0.5 min, followed by an increase to 60% solution B in 8 min and then to 90% in 1 min; these conditions were maintained for 6 min before returning to the initial conditions. These chromatographic conditions were used in all LC−HRESIMS and LC−HRESIMS/MS subsequent analyses.

Flash chromatography was conducted using a PureC-850 Flash Prep chromatography system (Buchi). The separation of F5 and the purification of Cyanobactin 850 were performed using a Thermo Scientific UltiMate 3000 Basic system. For the final stage of the purification of Cyanobactin 1027, HPLC was carried out using a Thermo Scientific Vanquish High throughput system. All these instruments were controlled by Chromeleon software (Thermo Fisher).

1D and 2D NMR data were acquired at the Materials Center of the University of Porto (CEMUP) using a Bruker Ascend 600 14.1 T 600 MHz instrument controlled with TopSpin 4.11 software. Compounds were analyzed in DMSO-$d_6$ (Sigma), and the chemical shifts ($^{1}H$ and $^{13}C$) were expressed in δ (ppm), referenced to the residual non-deuterated solvent (δH 2.5000, δC 39.520). The NMR data were visualized using MestreNova 14.0.0 software. All solvents used were of MS grade for MS-based experiments, HPLC gradient grade for HPLC analysis and purification, and ACS grade for extraction, VLC, and flash chromatography.

### 2.3.2. Small-Scale Extraction of Oscillatoriales cyanobacterium LEGE 16532 and LC−HRESIMS Analysis

The strain Oscillatoriales cyanobacterium LEGE 16532 was grown under the conditions and biomass harvested as mentioned above. For small scale extraction, 30 mg of the freeze-dried biomass of the cyanobacterial strain was homogenized with liquid nitrogen and extracted using methanol (2 × 20 mL) at room temperature with shaking. The supernatant was obtained by centrifugation (5000 g, 15 min, 4ºC, Gyrozen 2236R) and the solvent was evaporated in a rotary evaporator. The resulting extract was

weighted and resuspended in MeOH LC-MS grade (5 mg $mL^{-1}$) and filtered through a 0.2 µm regenerated cellulose syringe filter. Five microliters (5 µL) of the extract were injected into the LC-HRESIMS system through the conditions mentioned. Identification of cyanobactins in the methanolic extract was based on the theoretical molecular weights calculated from the predicted core peptide from the cyanobacterial strain using ChemBioDraw.

## 2.4. Isolation of cyanobactins from Oscillatoriales cyanobacterium LEGE 16532

### 2.4.1. Organic extraction

The freeze-dried biomass (44.96 g, d.w.) obtained from the large cultivation of Oscillatoriales cyanobacterium LEGE 16532 (60 L) was extracted through 8 rounds of percolation with MeOH (8x500 mL of MeOH) until the biomass began to lose pigment, changing from ochre to a purplish tone (Figure 11A). The crude extract was then analyzed by LC-HRESIMS to confirm the presence of the masses detected in the small-scale extraction (Figure 10).

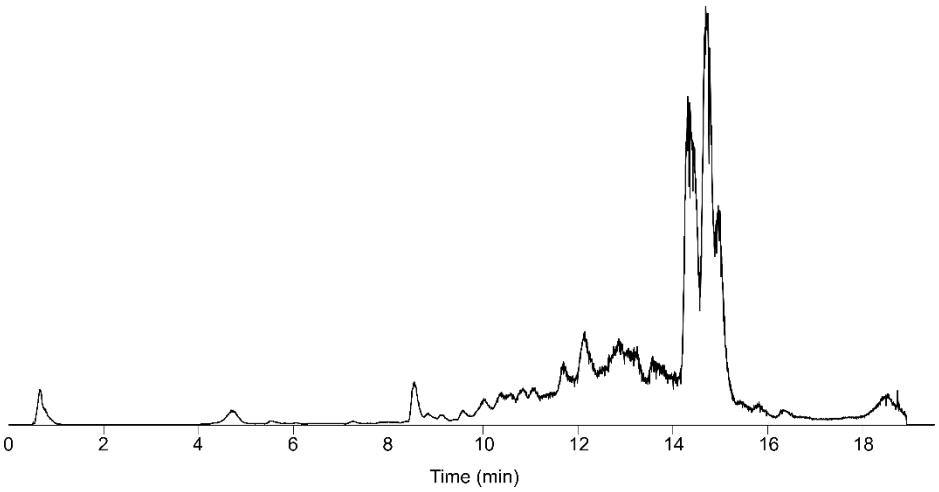



Figure 10 Total Ion Chromatograms obtained from LC-HRESIMS analysis of the crude extract.

### 2.4.2. Vacuum Liquid Chromatography (VLC)

In order to effectively separate the constituents of the crude extract, a process known as VLC (short for Vacuum Liquid Chromatography) was conducted utilizing a Reversed phase silica gel C18 (C18-RP; 17%C, approx. 0.7 mmol $g^{-1}$, part. Size: 0.045-0.065 mm, Acros Organics), as indicated in Figure 11B. This process was carried out through a gradual gradient of decreasing polarity, from water to methanol. The extract was solubilized with methanol and adsorbed onto silica particles via solvent evaporation. The extract-adsorbed silica was then loaded onto the top of the dry-packed column and

solvated with the initial elution mixture (1% MeOH/99% Water). A solvent gradient was implemented, beginning with a 1% MeOH/99% Water polarity, and gradually increasing to 100% MeOH, following a stepwise gradient of 1%>15%>30%>50%>70%>100%. This was then followed by a cleaning step with 100% Dichloromethane, which served to elute the remaining components and clean the column.

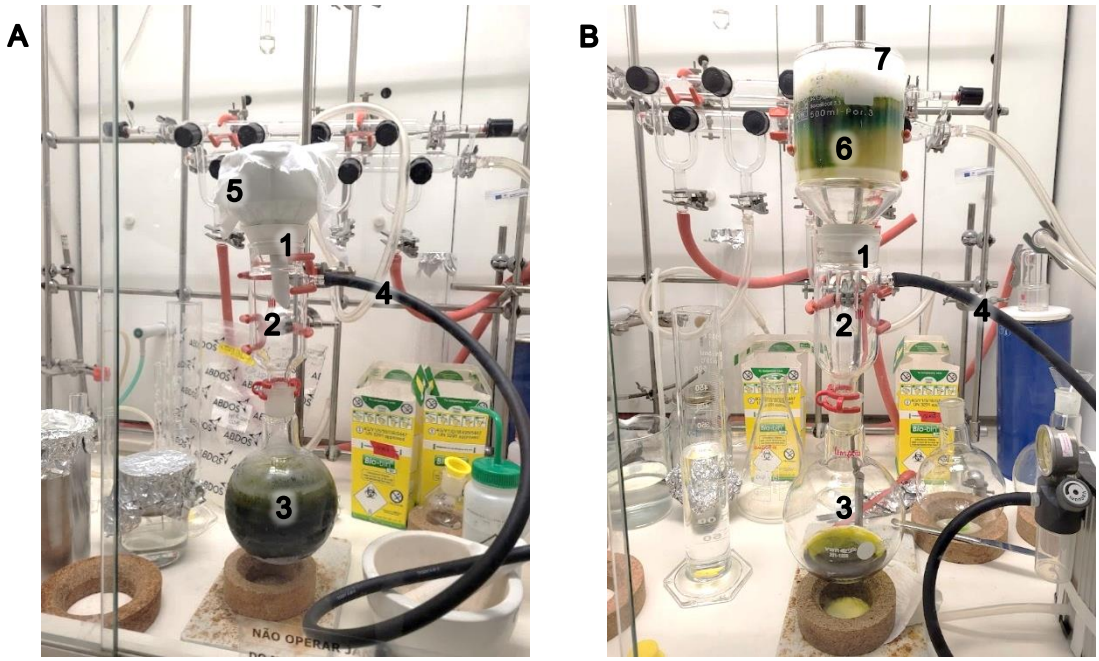



Figure 11 (A) Extraction system. (B) VLC system. 1- Büchner Funnel, 2 – vacuum adapter, 3 – round bottom flask, 4 – vacuum line, 5- cheese cloth, 6 – Reversed phase silica gel C18 (C18-RP; 17%C, approx. 0.7 mmol/g, part. Size: 0.045-0.065 mm, Acros Organics), 7 - Whatman No1 filter paper (also inside Bückner Funnel in extraction system)

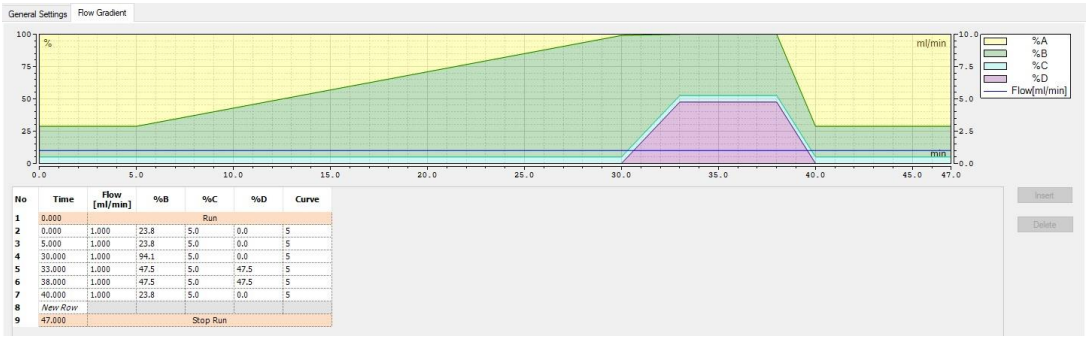
The VLC separation resulted in 6 distinguishable fractions, labelled as A-F and a Cleaning fraction that served as our backup, The fractions obtained were subjected to drying using a rotary evaporator and then transferred into pre-weighed vials. After each separation during the isolation process, a LC−HRESIMS analysis was performed in order to check in which fraction the target compounds were present.

### 2.4.3. Isolation of cyanobactin 850 and cyanobactin 1027

The fraction F was shown to contain the majority of the cyanobactins with higher abundance. Thus, this fraction (approx. 100 mg) was fractionated by automated flash chromatography using a C18 (17%) 40 g cartridge (Silicycle). The elution was performed, using a gradient of decreasing polarity from 30% acetonitrile (ACN) /70% water to 100% acetonitrile, followed by a stationary period with 100% acetonitrile and an increasing of

IPA, finishing with a stationary period with 50% ACN/50% IPA (Figure 11). The flow rate was 30 mL $min^{-1}$, and UV detection was performed at 254, 265, 280, and 320 nm.

This separation resulted in 70 tubes that were further pooled in 12 distinguishable fractions, based on the UV profile, labelled as F1-F12. The fractions were subsequently dried by rotary evaporation transferred into pre-weighed vials and the composition of each fraction was characterized via LC-HRESIMS. The *m/z* values for major cyanobactins were detected in subfraction 5. This fraction F5 (13.80 mg) was then fractionated by reverse-phase HPLC using a Thermo Scientific UltiMate 3000 with an Avantor Hichrom C18 analytical column (150 Å, 5 µm, 4.6 x 250 mm) with a constant flow rate of 1 mL min-1, at room temperature. The fraction F5 was solvated in a mixture 50% 1-propanol (IPA) /50% water and injected 50µL of the solution per run. Mobile phase A was H2O, mobile phase B was 1-propanol, mobile phase C was methanol and mobile phase D was acetonitrile. The method began with an isocratic step 71.2% A, 23.8% B and 5% C over 5 min followed by a gradient to 94.1% B over 25 min and held for 3 minutes, until the entrance of the 47.5% D. The composition of the mobile phase remains isocratic by a period of 7 minutes, with 47.5% B, 5% C, and 47.5% D, and returns to the initial conditions (Figure 12).



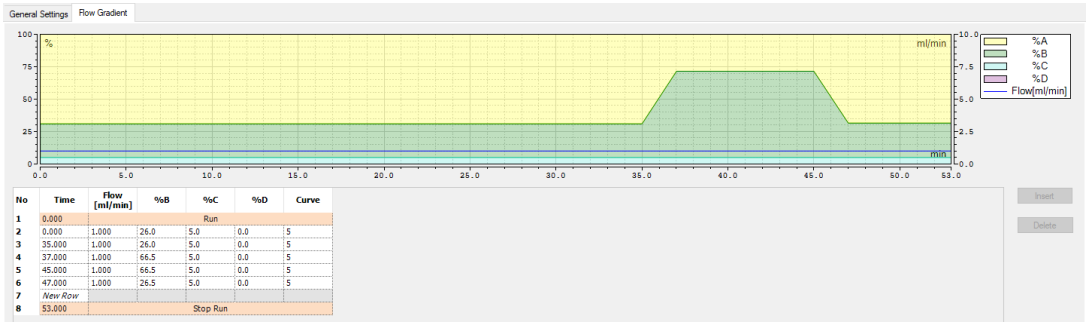

| No | Time | Flow [ml/min] | %B | %C | %D | Curve |
|---|---|---|---|---|---|---|
| 1 | 0.000 | Run | | | | |
| 2 | 0.000 | 1.000 | 23.8 | 5.0 | 0.0 | 5 |
| 3 | 5.000 | 1.000 | 23.8 | 5.0 | 0.0 | 5 |
| 4 | 30.000 | 1.000 | 94.1 | 5.0 | 0.0 | 5 |
| 5 | 33.000 | 1.000 | 47.5 | 5.0 | 47.5 | 5 |
| 6 | 38.000 | 1.000 | 47.5 | 5.0 | 47.5 | 5 |
| 7 | 40.000 | 1.000 | 23.8 | 5.0 | 0.0 | 5 |
| 8 | New Row | | | | | |
| 9 | 47.000 | Stop Run | | | | |

Figure 12 Elution program of fraction F5. A: Water, B: 1-propanol, C: methanol and D: acetonitrile.

This separation afforded 4 subfractions that were then analyzed by LC−HRESIMS: Fractions F5_A (2.83 mg), F5_B (which the most abundant mass corresponds to cyanobactin 850 (**1**), 3.52 mg) F5_C (which the most abundant mass corresponds to cyanobactin 1027 (**2**), 2.65 mg,) and F5_D (4.91 mg,).

For the last step of purification of cyanobactin 850, the fraction F5_B (3.52 mg) was then fractionated using the same UHPLC system and column. The mobile phase used was the same of the separation of fraction F5. The method (Figure 13) began with an isocratic step 69% A, 26% B and 5% C for 35 minutes, followed by a gradient step with an increasing of B till reach 66.5%. after that, we have an isocratic period to clean

the column and further this we return to initial conditions. 50 µL of the fraction was injected per run. This separation led to the isolation of cyanobactin 850 (**1**) yielding 1.1 mg and submitted to NMR analysis in order to determine in which residue the prenylation occurs.




| No | Time | Flow [ml/min] | %B | %C | %D | Curve |
|---|---|---|---|---|---|---|
| 1 | 0.000 | Run | | | | |
| 2 | 0.000 | 1.000 | 26.0 | 5.0 | 0.0 | 5 |
| 3 | 35.000 | 1.000 | 26.0 | 5.0 | 0.0 | 5 |
| 4 | 37.000 | 1.000 | 66.5 | 5.0 | 0.0 | 5 |
| 5 | 45.000 | 1.000 | 66.5 | 5.0 | 0.0 | 5 |
| 6 | 47.000 | 1.000 | 26.5 | 5.0 | 0.0 | 5 |
| 7 | New Row | | | | | |
| 8 | 53.000 | Stop Run | | | | |

Figure 13 Elution program of fraction F5_B. A: Water, B: 1-propanol, C: methanol and D: acetonitrile.

For the isolation of cyanobactin 1027 (**2**), which was the main component of fraction F5_C (2.65 mg), the purification was performed by reverse phase UHPLC using a Thermo Scientific Vanquish High throughput system with an Purospher® STAR RP-18 endcapped (5µm) LiChroCART analytical column (120 Å, 5 µm, 4.6 x 250 mm) with a constant flow rate of 1 mL $min^{-1}$, at room temperature and 35 µL of solubilized fraction was injected in each run. Mobile phase A was acetonitrile, mobile phase B was methanol, mobile phase C was water and mobile phase D was 1-propanol. The elution method (Figure 14) begins with an isocratic step with 5% B, 61% C and 34% D for 10 minutes, followed by a gradient with the increasing of polar phase, A till 61% and a decrease of C until 0%. The final part of the method is characterized by an isocratic period, mainly with non-polar solvent to clean the column. This purification step led to the isolation of 0.2 mg of cyanobactin 1027 (**2**), and the structure elucidation was performed based on LC−HRESIMS/MS data.


| No | Time | Flow [ml/min] | %B | %C | %D | Curve |
|---|---|---|---|---|---|---|
| 1 | -10.000 | Equilibration | | | | |
| 2 | -10.000 | 1.000 | 5.0 | 61.0 | 34.0 | 5 |
| 3 | New Row | | | | | |
| 4 | 0.000 | Run | | | | |
| 5 | 0.000 | 1.000 | 5.0 | 61.0 | 34.0 | 5 |
| 6 | 10.000 | 1.000 | 5.0 | 61.0 | 34.0 | 5 |
| 7 | 15.000 | 1.000 | 5.0 | 0.0 | 34.0 | 5 |
| 8 | New Row | | | | | |
| 9 | 22.000 | Stop Run | | | | |

Figure 14 Elution program of fraction F5_C. A: acetonitrile, B: methanol, C: water and D: 1-propanol.

### 2.4.4. Structural elucidation of cyanobactins

Since the unprenylated versions were produced in lower abundance compared with the prenylated ones, structure elucidation was performed only for the latter versions. Additionally, NMR analysis was performed only for cyanobactin 850 in order to determine the position and direction of prenylation, assuming that will be the same for the other prenylated cyanobactins. Cyanobactin 850 was dissolved in deuterated DMSO and analyzed by 1D and 2D NMR experiments ($^{1}H$, $^{13}C$ APT, HSQC, HMBC, COSY). The experiments were carried out at CEMUP, in 600MHz spectrometer equipped with a 5 mm cryoprobe. NMR spectra were processed in the Mnova software. All the other cyanobactins identified in this study had their structure elucidated based on LC−HRESIMS/MS data as described in General procedures section. The structural elucidation for these cyanobactins was performed through LC−HRESIMS/MS data since their isolation was not possible during the development of this work.

## 2.5. Bioactivity Assays

Since CIIMAR has several biological assays available, the pure compounds were delivered and tested in collaboration with different colleagues. The cytotoxicity assays were carried out by research technician Sara Freitas, while the antiamoebic assays were conducted by Dr Ana Vieira, in collaboration with the CYAmoeba project and the protocols used are described in the next sections.

### 2.5.1. Cytotoxicity assays

Human liver hepatocellular carcinoma (HepG2) and human colon carcinoma cell line HCT 116, were obtained from Sigma-Aldrich, while the neuroblastoma cell line SH-SY5Y was obtained from American Type Culture Collection (ATCC). The three human cancer cell lines were cultured at 37 °C in a humidified atmosphere of 5% $CO_2$, in Dulbecco's modified Eagle medium (DMEM) (Gibco, Thermo Fisher Scientific), McCoy's 5A medium (Sigma-Aldrich) and 1:1 mixture of DMEM and Ham's F12 medium (brand), respectively. All media were complemented with 10% fetal bovine serum, 1% penicillin/streptomycin and 0.1% amphotericin.

Cell viability was assessed by the MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay. HepG2 and SH-SY5Y were seeded in 96-well plates at a density of $3.0\times10^{4}$ cells $mL^{-1}$, while HCT116 was seeded at a density of $3.3x10^{4}$ cells $mL^{-1}$ and incubated overnight before adding the the pure compounds in triplicates at a concentration of 10 µg $mL^{-1}$. DMSO served as solvent (0.1%) and staurosporine (0.5%) was used as the positive control. After 48h, 20 µL of MTT (1 mg $mL^{-1}$) was added to each well and incubated for 4h. Media was removed, and formazan

crystals were dissolved in 100 µL of DMSO and absorption was read at 570 nm on a microplate reader (Synergy HT, Biotek, Germany).

### 2.5.2. Antimicrobial assays

The isolated compounds were tested against five references strains: the Gram-positive *Staphylococcus aureus* ATCC 29213 and *Bacillus subtilis* ATCC 6633, the Gram-negative *Escherichia coli* ATCC 25922 and *Salmonela typhimurium ATCC* 25241 and the yeast *Candida albicans* ATCC 10231. The antimicrobial activity was screened against all strains through the agar disc-diffusion method [109]. For the bacterial inoculation we used the Mueller-Hinton agar medium (Oxoid Ldt., Wade Road, UK) and for fungi inoculation the Sabouraud agar medium (Oxoid Ldt., Wade Road, UK) was used. Petri dishes containing the respective media and inocula were incubated at 37°C for 24 hours, with the turbidity adjusted to 0.5 McFarland standard ($OD_{600nm}$ = 0.095-0.110). The plates were then inoculated with the adjusted culture, and 6mm blank discs (Oxoid Ldt., Wade Road, UK) were placed on the inoculated plates and impregnated with 15 µl of the 1 mg $mL^{-1}$ stock solution of pure compound dissolved in DMSO. Blank discs impregnated with 15 µL DMSO were utilized as negative controls. Following a 30 minute incubation at room temperature, the plates were further incubated at 37°C for 24 hours, and the antimicrobial activity was determined by the development of inhibition halos.

### 2.5.3. Antiamoebic assays

*Acanthamoeba castellanii*, *Acanthamoeba polyphaga* (both freshwater amoebae) and *Dictyostelium discoideum* (soil amoeba) were grown in 5 mL of PYG medium in 25$cm^2$ flasks and incubated at 25ºC for one week. Then, the suspension was transferred to a 75 $cm^2$ flask with fresh PYG medium and incubated for more 2-3 days at 25ºC. For the experiments, each strain of amoebae was quantified by adding 10 µL of 0.4% Trypan Blue solution to 10 µL of culture and the viable cells were determined by counting using a disposable slide hemocytometer. To study the anti-amoebic effect of the pure cyanobactins against *A. castellanii*; *A. polyphaga* and *D. discoideum*, a colorimetric protocol was used by evaluating the cell viability using the Alamar Blue assay. Briefly, $10^6$ cells $mL^{-1}$ of each strain of amoeba was seeded on a 96-well microtiter plate and incubated for 1 hour at 25ºC. A final concentration of 10 µM (0.5 µL of a stock solution of 2 mM) of the pure compound was tested for each strain, DMSO (0.5 µL of DMSO 0.5%) and the detergent chlorhexidine digluconate (0.5 µL of a stock solution at 20 mM) were used as negative and positive control, respectively and the plate was then incubated for 24 hours at 25ºC. Finally, 10 µL of the alamarBlue™ reagent (ThermoFisher) was added to each well to measure the cells viability. Plates were

incubated for 4h at 25ºC. Subsequently, the plates were analyzed using a microplate reader (BioTek Cytation 5 Cell Imaging Multimode Reader, Agilent) and the fluorescence wavelengths of excitation and emission were set at 560 and 590 nm, respectively. All assays were performed in triplicate.

# 3. Results and Discussion

## 3.1. Discovery of new cyanobactins from Oscillatoriales cyanobacterium LEGE 16532

Genome mining represents an alternative strategy in comparison to traditional methods for the identification of new secondary metabolites. These compounds act as scaffold for further enhancement through medicinal chemistry and combinatorial biosynthesis. This catalyzes the development of valuable solutions for human health, animal health, crop protection, and numerous biotechnological applications [110].

In this work, we were focused on cyanobactins, the most widely studied of ribosomally derived lipopeptides. After a previous genome mining performed in cyanobacterial strains from LEGE Culture Collection, a candidate cyanobactin BGC was detected in Oscillatoriales cyanobacterium LEGE 16532; however, this cluster was split in two contigs. To overcome this, a new attempt of genome sequencing from the cyanobacterial strain was performed using a combination of short-read Illumina and long-read Nanopore sequencing technologies. The analysis of the obtained sequencing data using antiSMASH [106], currently the most widely used tool for detecting and characterizing BGCs in bacteria and fungi, has led to discovery of a complete 11 Kb cyanobactin BGC (*osc*) (Figure 15).

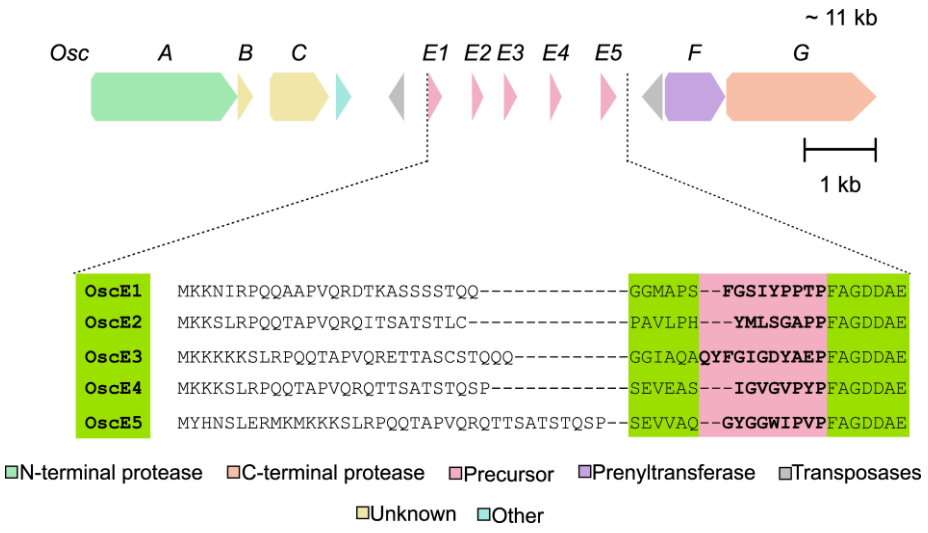



Figure 15 Architecture of the *osc* BGC from Oscillatoriales cyanobacterium LEGE 16532. Highlighted in light green are the predicted RSII and RSIII and in pink are the predicted core peptides.

A BlastP analysis was conducted in order to annotate the function of each protein based on their homology with entries in the NCBI database (Table 1). The *osc* BGC encodes a total of 13 proteins which include the OscA protease, OscG macrocyclase, OscE1-E5 precursor peptides, OscF prenyltransferase, as well as the OscB and OscC proteins and other three open reading frames (ORFs) encoding transposases or other proteins with unclear role in the biosynthesis of cyanobactin.

Table 1 Annotation of the *osc* BGC from Oscillatoriales cyanobacterium LEGE 16532

| Protein | Predicted function | Description | Identity/similarity (%) | Organism | Accession number |
|---|---|---|---|---|---|
| **OscA** | N-terminal protease | PatA/PatG family cyanobactin maturation protease | 77/84 | *Lyngbya* sp. CCAP 1446/10 | WP_265234509.1 |
| **OscB** | Unknown | Cyanobactin biosynthesis system PatB/AcyB/McaB family protein | 90/93 | *Microcoleus sp.* PH2017_19_SFW_U_A | MCC3506896.1 |
| **OscC** | Unknown | cyanobactin biosynthesis PatC/TenC/TruC family protein | 88/94 | *Oscillatoriales cyanobacterium* | TAG93318.1 |
| **ORF1** | Protein antitoxin (HicB) | Type II toxin-antitoxin system HicB family antitoxin | 82/91 | *Scytonema* sp. UIC 10036 | WP_155750940.1 |
| **ORF2** | Putative transposase | Transposase | 56/69 | *Dolichospermum* sp. UHCC 0352 | WP_168650423.1 |
| **OscE1** | Percursor peptide | Anacyclamide/piricyclamide family prenylated cyclic peptide | 56/75 | *Calothrix* sp. NIES-2100 | WP_096598589.1 |
| **OscE2** | Percursor peptide | Anacyclamide precursor | 53/62 | *Anabaena sp.* TR232 | ADA63241.1 |
| **OscE3** | Percursor peptide | Anacyclamide precursor | 53/71 | *Dolichospermum lemmermannii* PH256 | ADA63228.1 |
| **OscE4** | Percursor peptide | Anacyclamide/piricyclamide family prenylated cyclic peptide | 54/68 | *Calothrix sp.* NIES-2100 | WP_096598589.1 |
| **OscE5** | Percursor peptide | Anacyclamide precursor | 57/68 | *Anabaena sp.* BIR260 | ADA63229.1 |
| **ORF3** | Putative transposase | Transposase | 76/85 | *Microcystis aeruginosa* BS13-02 | NCS26882.1 |
| **OscF** | Prenyltransferase | LynF/TruF/PatF family peptide O-prenyltransferase | 50/71 | *Symploca sp.* SIO1B1 | NER99685.1 |
| **OscG** | C-Terminal protease | PatA/PatG family cyanobactin maturation protease | 71/81 | *Microcoleus asticus* | WP_172192417.1 |

The biosynthesis of cyanobactins involves the participation of various proteins such as the protease A and the macrocyclase G. Among these, OscA and OscG have been found to exhibit homology to the proteins involved in the biosynthesis of other cyanobactins, including PatA/G from patellamides which was the first cyanobactin biosynthetic pathway to be characterized [65]. PatA and PatG are subtilisin-like serine proteases that play a critical role in the maturation process of cyclic cyanobactins. They are regulated by recognition sequences (RS) such as RSII and RSIII, respectively [111]. The RSII and RSIII are recognized by PatA and PatG, respectively, and are composed of amino acid sequences such as AVLAS, GVDAS, GLTPH, GLEAS, GVEPS, AYD, AYDGE, FAGDDAE, SYD, SYDD, and SYEGDEAE [112]. PatA can cleave the peptide by releasing the N-terminal end of the peptide, while PatG frees the C-terminal end of the core peptide and further catalyses the C-N macrocyclization of the final peptide sequence [111].

The RSII and RSIII of PatA and PatG have proven to be useful tools in the discovery of new precursor peptides and in the prediction of various core peptide sequences [64, 112]. We used RSIII (FAGDDAE) to identify five distinct precursor peptides. However, we were unable to identify homologous sequences to RSII as described in the literature, prompting us to utilize bioinformatics tools such as RiPPMiner-Genome and BAGEL4 to analyze the precursor amino acid sequences [26, 27]. The resulting sequence of the core peptides were named as OscE1 (FGSIYPPTP), OscE2 (YMLSGAPP), OscE3 (QYFGIGDYAEP), OscE4 (IGVGVPYP), OscE5 (GYGGWIPVP) and are indicated in pink in in Figure 15. Furthermore, the OscF shows a low homology when compared to other prenyltransferases present in cyanobactin pathways reported until now and the function of OscB and OscC remains unknown, even though they display homology to proteins encoded in nearly all the cyanobactin gene clusters [111].

Since RiPPs are directly encoded in the genome as are proteins, a proteomic approach could be used to connect a RiPP genotype with its chemotype [113]. Considering all the information retrieved from the annotation of the *osc* BGC, we predicted the structures of each core with possible modifications (macrocyclization and prenylation). Then, after manually examining the data obtained from LC−HRESIMS on a methanolic extract of freeze-dried biomass from Oscillatoriales cyanobacterium LEGE 16532, it was observed that there were five groups of cyanobactins, (Figure 16, Supplementary table 1), each consisting of unprenylated and mono-prenylated variants of each core peptide, supported by the prenyl neutral loss (*m/z* 68.03). However, the detection of the unprenylated version of the OscE3 core peptide (retention time of 7.83

min) as presented in supplementary table 1, was not achievable when analyzing the methanolic extract. Furthermore, the EIC of the ion with a *m/z* 1241.5473 (Figure 16), corresponding to the unprenylated version of OscE3, suggests in source fragmentation of cyanobactin 1308 (*m/z* 1309.6099), that correspond to a prenylated version of the OscE3 core peptide, since they present a similar retention time. .

Finally, at this point, we were not able to identify in which residue the prenylation occurs as well as the direction of the prenylation. The exact mass of each compound was searched in the Natural Product Atlas (NAP) [25], but no results were found, indicating that these compounds are new-to-science.

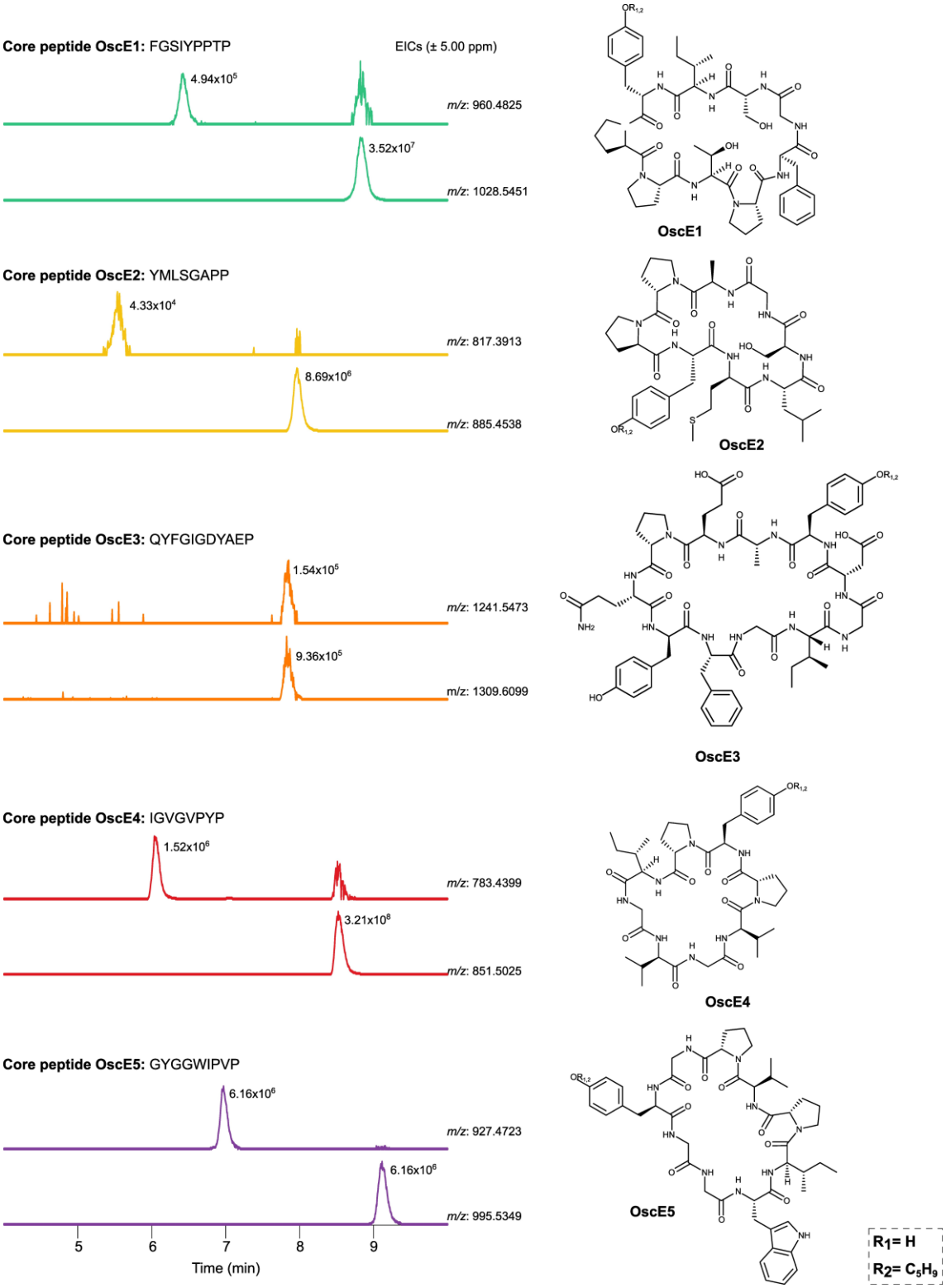



Figure 16 Detection of the cyanobactins are illustrated by the extracted ion chromatograms (EICs) for the $[M+H]^+$ ion of each compound. Each group of cyanobactins are composed by the unprenylated (Top) and mono-prenylated (bottom) version. The proposed structures of each cyanobactin produced by each core are illustrated in Figure 15.

## 3.2. Sequence-function and chemical space of cyanobactin prenylation

F-type prenyltransferases play a crucial role in modifying a notable portion of cyanobactins. These enzymes add an isoprene unit to specific amino acid residues such as Ser, Thr, Tyr, and Trp, either at the N- or C-terminus of linear peptides [74, 81, 114]. In recent years, new modifications have been reported as the case of LimF, which catalyzes a histidine-C-geranylation as well as Tyrosine-O-geranylation [115], AgcF which catalyzes the mono- and bis-N-prenylation of the guanidine moiety of arginine [85], while AutF targets the guanidinium moiety in Arg and homoarginine resulting in the prenyl-D-Arg-containing autumnalamide [86]. Other example is TolF present in *tol* BGC responsible for the production of tolypamide which is forward prenylated on a threonine residue, representing an unprecedented modification [84]. The *osc* BGC of Oscillatoriales cyanobacterium LEGE 16532 also encodes a prenyltransferase (OscF) which exhibits only 49.83% identity to *Symploca* sp. SIO1B1. This led to investigate the sequence-function and chemical space of cyanobactins prenylation including the OscF. For that, a sequence similarity network (SSN) using the online Enzyme Function Initiative-Enzyme Similarity Tool (EFI-ESI) [116] was generated using the OscF as query and add to the analysis the InterPro family of prenyltransferases data (IPR031037), with an alignment score threshold of $10^{-86}$, as used in previous works [84]. This analysis resulted in the network represented in Figure 17.

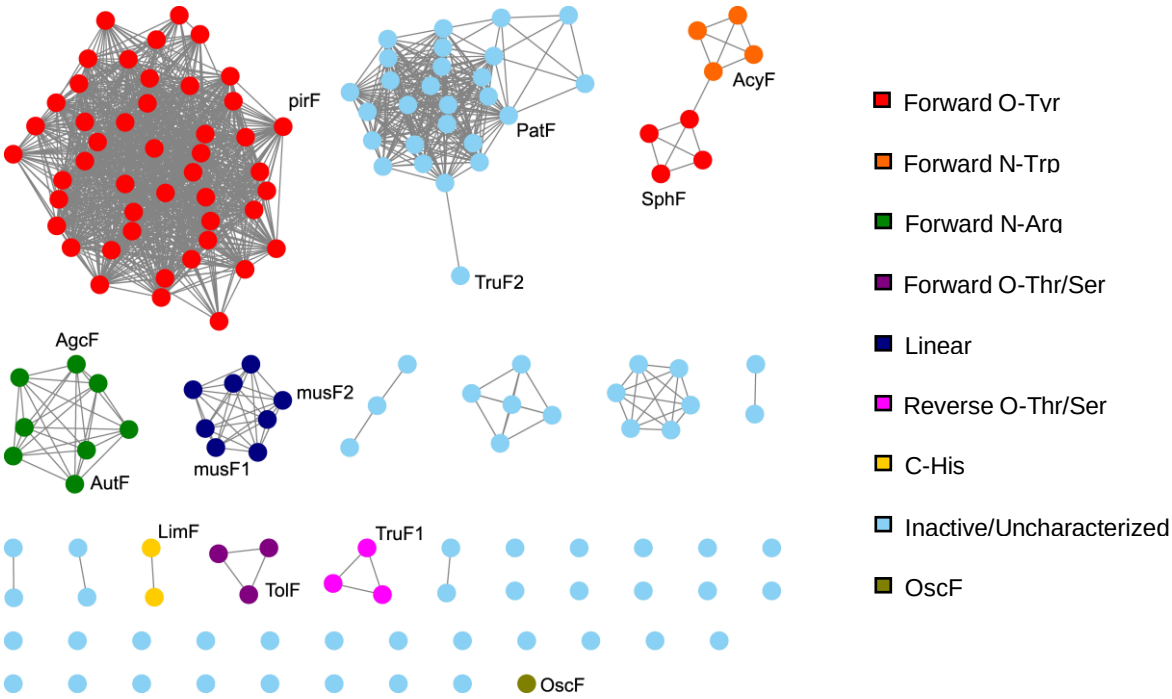



Figure 17 Sequence similarity network demonstrating the clustering of cyanobactin prenyltransferases according to their regiospecificity using an alignment score threshold of $10^{-86}$.

On the SSN we observed two large clusters, one containing inactive enzymes (PatF and TruF2) and the other with enzymes catalyzing forward O-prenylation on Tyr (PirF and SphF) [74, 77, 82, 117]. The SSN also contained smaller clusters representing other regiospecific transformations such as N-prenylation on tryptophan (AcyF) [63], N-prenylation on arginine (AgcF and AutF) [85, 118], prenylation on the terminal amino acid (musF1, musF2) [89], C-geranylation on histidine (LimF) [115], O-prenylation on threonine and serine residues (TolF) [84] and O-prenylation on threonine and serine residues (TruF1) [74]. Often, clustering is diagnostic for the function of uncharacterized enzymes, and indeed all characterized enzymes fell into clusters predictive of their chemistry. The OscF from Oscillatoriales cyanobacterium LEGE 16532 did not group in any cluster. In fact, OscF from Oscillatoriales cyanobacterium LEGE 16532 is a singleton, demonstrating that it might have a distinct activity compared with the other proteins belonging to the InterPro family of prenyltransferases (IPR031037). Thus, at this point, it was still not possible to assign which is the amino acid acceptor and what is the orientation of the prenylation either forward or reverse. In this way, we proceeded with MS-guided isolation and structural elucidation of the cyanobactins detected in the methanolic extract from Oscillatoriales cyanobacterium LEGE 16532.

## 3.3. MS-guided isolation of Cyanobactins

After the detection of the unprenylated and mono-prenylated version of cyanobactins in the methanolic extract from a small-scale culture of Oscillatoriales cyanobacterium LEGE 16532, it was not possible to determine in which residue the prenylation occurs. Also, after the SSN analysis, the OscF appears as a singleton, suggesting a potentially new enzymatic activity of this prenyltransferase. Thus, to confirm the position of the prenylation, a large-scale cultivation of the cyanobacterial strain was carried out in order to obtain enough biomass to isolate the cyanobactins and thus, confirm the position of the modification. The large-scale culturing of the Oscillatoriales Cyanobacterium LEGE 16532 in Z8 culture medium yielded 44.96 g of biomass (d.w.) that was extracted with methanol using repeated percolation, resulting in a crude extract of 4.00 g.

To isolate the cyanobactins, we employed reversed-phase (C18) vacuum liquid chromatography (VLC) with a stepwise gradient of decreasing polarity from water to MeOH, resulting in six fractions (A−F). In each step of fractionation, a LC−HRESIMS analysis was performed and after the VLC, fraction F showed the presence of the five groups of cyanobactins previously identified in the methanolic extraction.

Subsequent reversed-phase flash chromatography of fraction F resulted in 12 fractions with a distinct distribution of different groups of cyanobactins. Interestingly, when searching for the masses of core peptide OscE3 with the aminoacidic sequence QYFGIGDYAEP, it was found that alongside the mono-prenylated peptide, there was also a less intense peak corresponding to a bis-prenylated cyanobactin (Figure 18A), with a difference of +68.03 a.m.u. to the mono-prenylated molecule observed in MS spectrum (Figure 18B). The two versions of cyanobactins encoded by the OscE3, ion $[M+H]^+$ *m/z* 1309.60993 for the mono and *m/z* 1377.67681 (calculated mass $[M+H]^+$ *m/z* 1377.67253, error -0.00428 Δppm) for the bis-prenylated version, were detected in fraction F2.

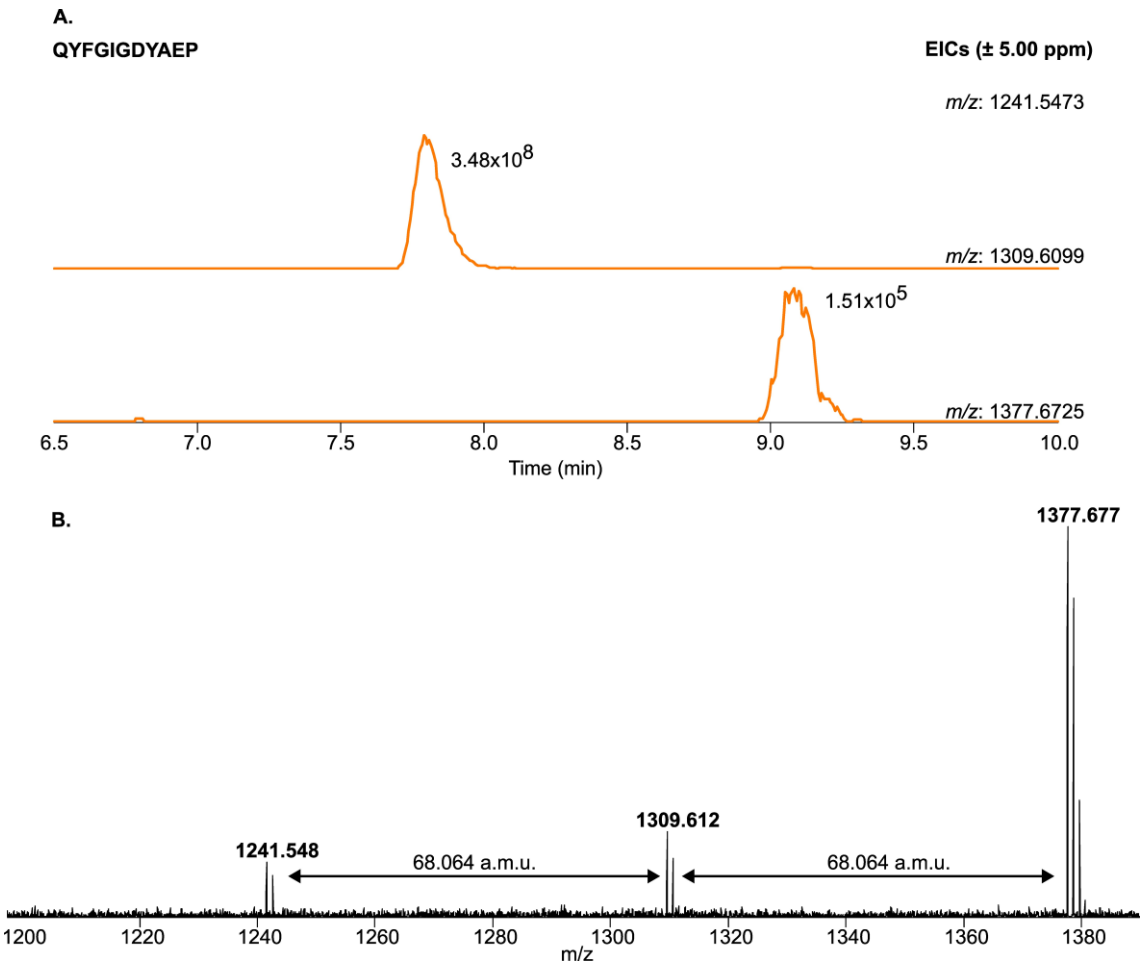



Figure 18 Detection of the mono- (top) and bis-prenylated (bottom) cyanobactins in the F2 fraction are illustrated by EICs for the $[M+H]^+$ on A., and the mass spectrum is represented in B.

Despite the high carry-over of the different groups of cyanobactins throughout the fractions, fraction F5 exhibited a higher abundance of cyanobactin 850 (*m/z* 851.5025) and cyanobactin 1027 (*m/z* 1028.5451), as represented in Figure 19.

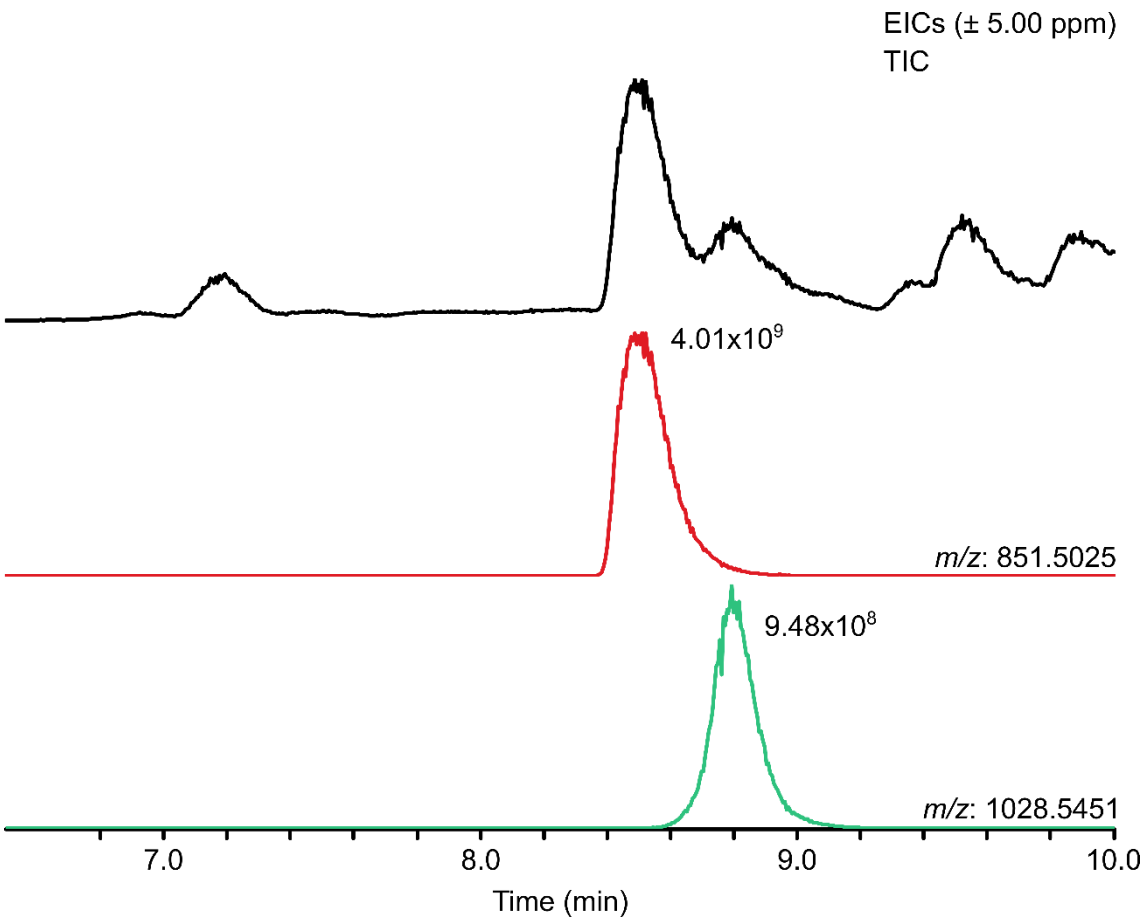



Figure 19 EICs for the [M+H]+ ion of cyanobactin 850 (**1**) and cyanobactin 1027 (**2**) in the F5 fraction.

Due the higher abundance of cyanobactin 850 (**1**) and cyanobactin 1027 (**2**), additional chromatographic steps were performed in order to isolate these compounds for structural elucidation through NMR spectroscopy since the direction of prenylation would only be confirmed using this analysis. Therefore, using High-Performance Liquid Chromatography (HPLC) coupled to a reversed-phase analytical column (Avantor Hichrom, C18, 150 Å, 5 µm, 4.6 x 250 mm) was carried out. This separation has provided high selectivity of the peaks and facilitated the subsequent isolation of the compounds, as shown in Figure 20. Subsequently, a second purification process was conducted using an isocratic program for the F5_B fraction, resulting in a mass yield of 1.1 mg of pure cyanobactin 850 (**1**). Additionally, a similar purification process for fraction F5_ C using an isocratic program, leading to a mass yield of 0.2 mg of pure cyanobactin 1027 (**2**). Following these purification steps and based on the mass obtained from the pure compounds, the pure compound **1** was analysed via NMR spectroscopy using a combined analysis of 1D and 2D (HSQC, HMBC and COSY).

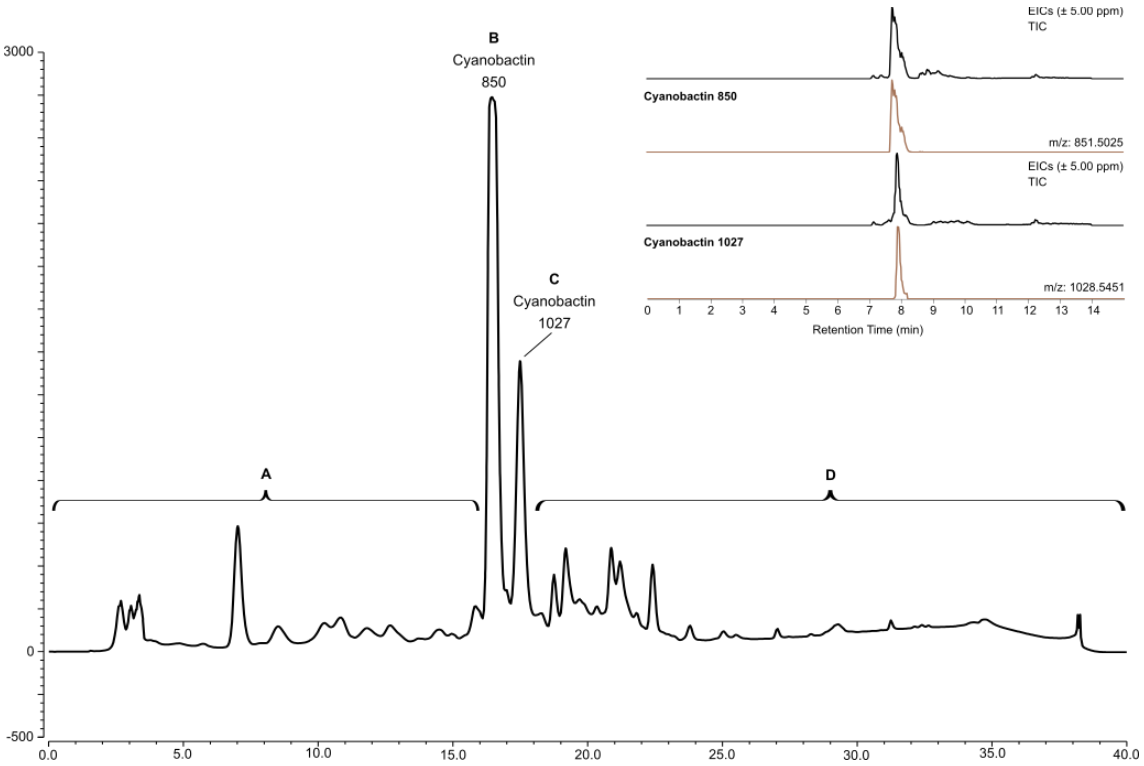



Figure 20 Fractionation scheme divided into 4 fractions (F5_A-D). Fractions B and C correspond to cyanobactin 850 (**1**) and cyanobactin 1027 (**2**), respectively.

A summary of the cyanobactins isolation process from Oscillatoriales cyanobacterium LEGE 16532 is represented in Figure 21.

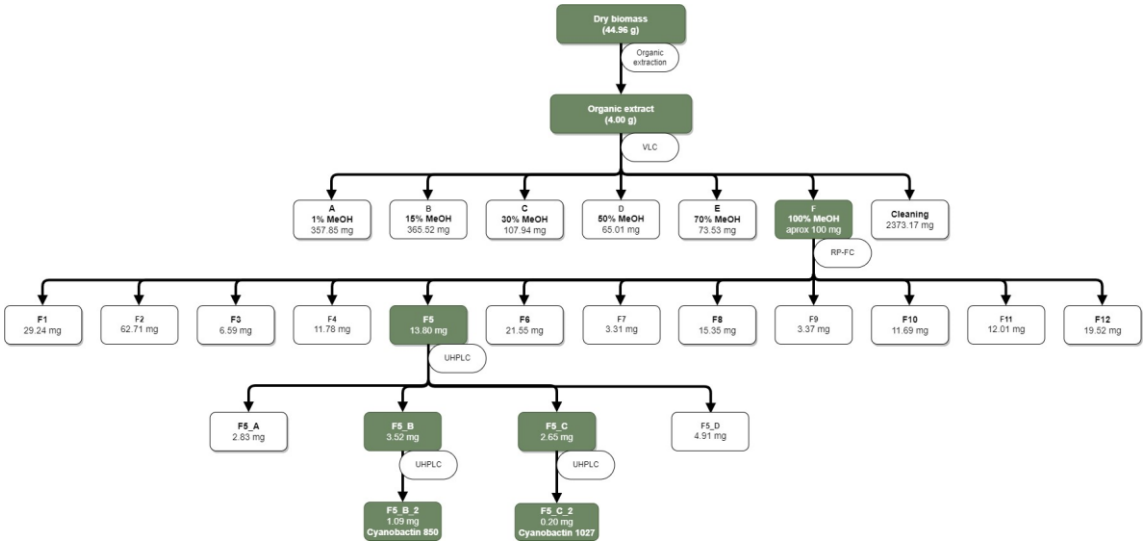



Figure 21 Fractionation scheme leading to the purification of cyanobactin 850 (**1**) and cyanobactin 1027 (**2**)

## 3.4. Structural Elucidation of Cyanobactin 850 (**1**) by NMR

Through NMR analysis, it would be now possible to assign the prenylated amino acid but also the orientation of the prenylation. Since RiPPs are genetically encoded, the structural elucidation by NMR was only focused on the amino acid where the prenylation occurs. The confirmation of the amino acid sequence was performed through HRESIMS/MS as performed for the other prenylated cyanobactins.

The 1H NMR spectrum of compound 1, as described in Table 2 and Supplementary Information S1 and S2, exhibited characteristic features of a forward prenylation. This included the distinctive gem-dimethyl moiety signals from the prenyl group, each integrating to three protons, and resonated at δH 1.69 and 1.73. Moreover, the combination of heteronuclear single quantum coherence (HSQC), heteronuclear multiple bond correlation (HMBC) and correlation spectroscopy (COSY) data (Table2, Supplementary Figure S3-S5), was carried out and used to assign spin systems for the tyrosine and for the prenyl group that constitute compound 1 (Figure 22).

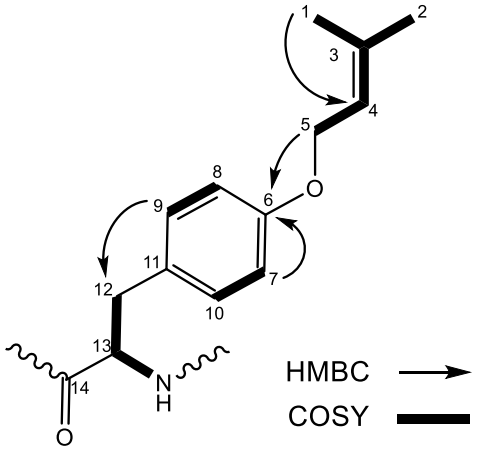



Figure 22 Figure 22 NMR-based structure analysis of Cyanobactin 850 (**1**).

The NMR analysis provided strong evidence for the prenylation of Tyr, chiefly by observing the correlation on HMBC between the $H_2$-5 protons and the aromatic C-6 carbon, present on the tyrosine side chain. in fact, these values were in agreement with the literature for other forward prenylation on tyrosine residue [77]. In this way, we can confirm that the prenyltransferase acts on tyrosine residue in forward direction.

Table 2 NMR spectroscopic data (600 MHz for $^{1}H$ and 150 MHz for $^{13}C$, DMSO-$d_6$, 25 ºC) for Cyanobactin 850.

| No | $\delta_C$ | Group | $\delta_H$ [mult., *J* (Hz)] | | HMBC | COSY |
|---|---|---|---|---|---|---|
| 1 | 17.87 | $CH_3$ | 1.69 | d (3.4) | 2, 3, 4 | 4, 5 |
| 2 | 25.41 | $CH_3$ | 1.73 | s | 1, 3, 4 | 4, 5 |
| 3 | 136.81 | C | | | | |
| 4 | 120.05 | CH | 5.39 | m | 1, 2 | 1, 2, 5 |
| 5 | 64.21 | $CH_2$ | 4.47 | d (7.02) | 3, 4, 6 | 1, 2, 4 |
| 6 | 157.2 | C | | | | |
| 7 | 114.5 | $C_H$ | 6.86 | d (8.49) | 6, 8, 9/10 | 9/10 |
| 8 | 114.5 | CH | 6.86 | d (8.49) | 6, 7, 9/10 | 9/10 |
| 9 | 129.63 | CH | 7.14 | dd (2.89; 8.69) | 12a, 11, 6 | 7/8 |
| 10 | 129.63 | CH | 7.14 | dd (2.89; 8.69) | 12a, 11, 6 | 7/8 |
| 11 | 129.78 | C | | | | |
| 12a | 37.86 | $CH_2$ | 2.58 | d (5.54) | 10 | 12b, 13 |
| 12b | | | 3.06 | m | 10 | 12a |
| 13 | 52.1 | CH | 4.74 | m | | 12a, 12b |
| 14 | N.A. | C=O | | | | |
| | | NH | 7.53 | d (7.8) | | 13 |

## 3.5. Structural Elucidation of Cyanobactins by MS/MS

In numerous instances, RiPPs exhibit molecular weights that surpass 2,500 Da, thereby posing a significant challenge for the expeditious analysis of these molecules through the utilization of NMR spectroscopy. Consequently, mass spectrometry (MS) emerges as the most suitable and convenient technique for the structural characterization of RiPPs due to its inherent capability to effectively handle and investigate compounds of such substantial size [119]. In addition, the effectiveness of tandem mass spectrometry in structural elucidation of cyanobacterial peptides is well

established [62, 120, 121] and is especially useful in the analysis of compounds presenting low abundance in complex samples. For the structural elucidation, only the mono-prenylated and bis-prenylated version of the cyanobactins identified were considered since the unprenylated version present much less intensity and was not possible to pursue with the HRESIMS/MS analysis. In fact, the cyclic cyanobactin, is the substrate of the enzymatic activity of the prenyltransferase, justifying the higher abundance of prenylated version of the cyanobactins than the cyclic ones.

The structural elucidation of the prenylated cyanobactins by MS/MS analysis has expanded the knowledge of these molecules. The observed MS/MS spectra of each prenylated cyanobactin revealed a series of fragment ions, each corresponding to specific cleavage sites along the peptide's backbone. By comparing these observed fragments to the expected fragment ions based on the predicted structure drawn on ChemDraw, it was possible to establish a high degree of sequence confidence.

Regarding the cyanobactin 850 (**1**) (*m/z* 851.5026), after the confirmation of the position and direction of the prenylation based on NMR data, it was possible to confirm the aminoacidic sequence of this compound as observed in Figure 23. For instance, the fragment ion with *m/z* 426.23 corresponds to the sequence "IGVGV". Moreover, the fragment ion with *m/z* 329.18 corresponds to the sequence “PY(preny)l”.

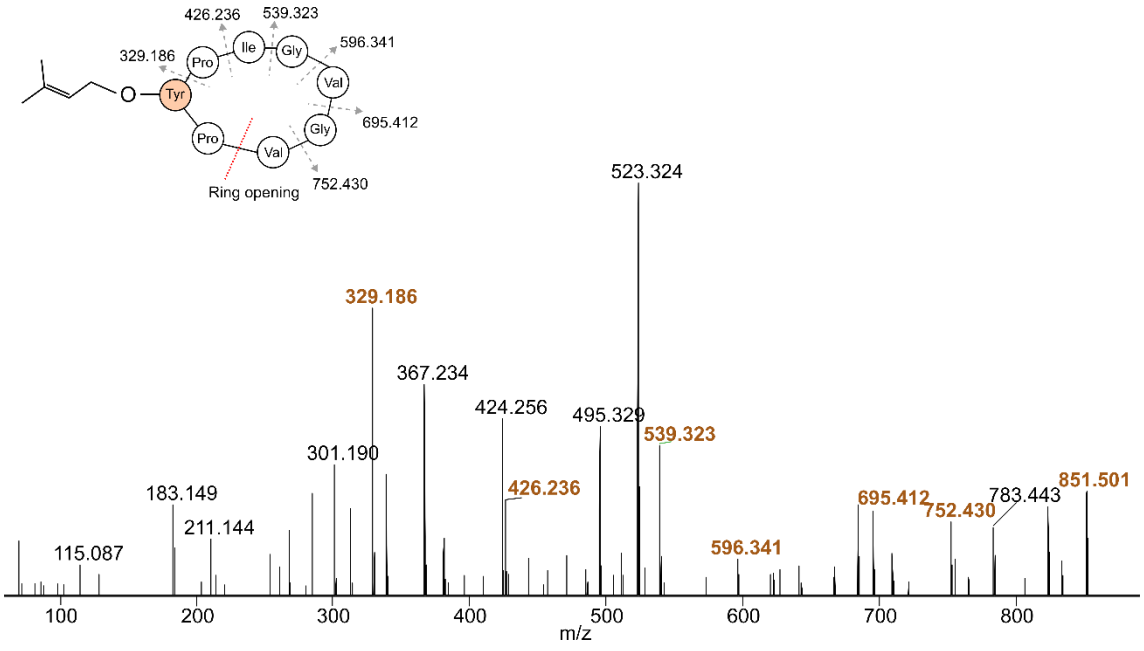



Figure 23 MS/MS data from cyanobactin 850 (**1**) (m/z 851.5026).

In the case of cyanobactin 1027 (**2**) (*m/z* 1028.5451), the fragmentation of the prenylated cyanobactin using an energy of 35 eV led to a few ion fragments (Figure 24). For example, it was possible to identify the parent mass *m/z* 1028.54, and the ion fragment of *m/z* 733.39, which corresponds to "PFGSIY(prenyl)" and the fragment *m/z* 502.26, with a difference of 231.12 that indicates a loss of the prenylated Tyr residue.

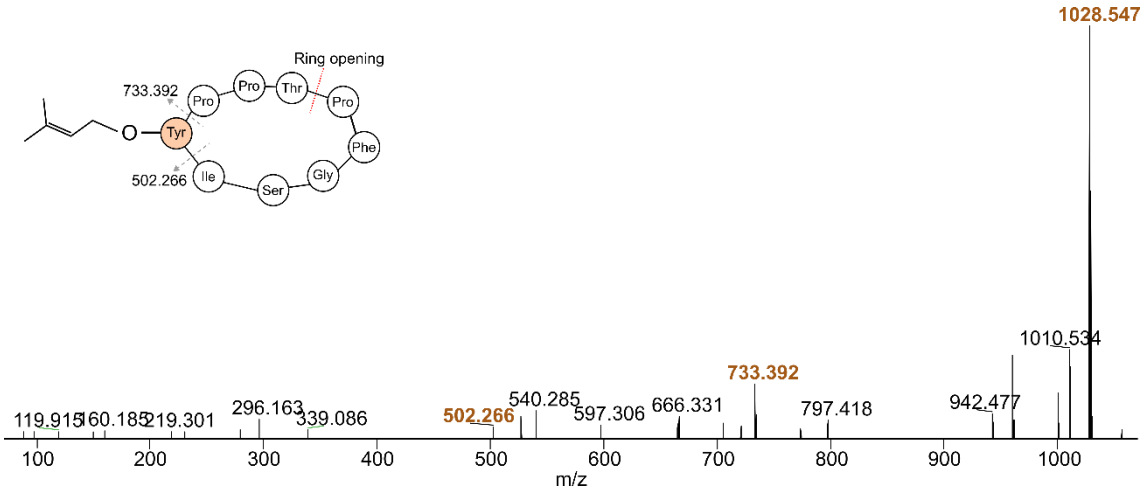



Figure 24 MS/MS data from cyanobactin 1027 (**2**) (m/z 1028.5451).

By the analysis of MS/MS data from cyanobactin 994 (*m/z* 995.5349), it was also possible to confirm the aminoacidic sequence of this peptide. Looking to the MS/MS spectra, additionally to the parent ion (*m/z* 995.533) another prominent ion of *m/z* 799.41 was detected corresponding to "PGY(prenyl)GGWI" sequence (Figure 25).

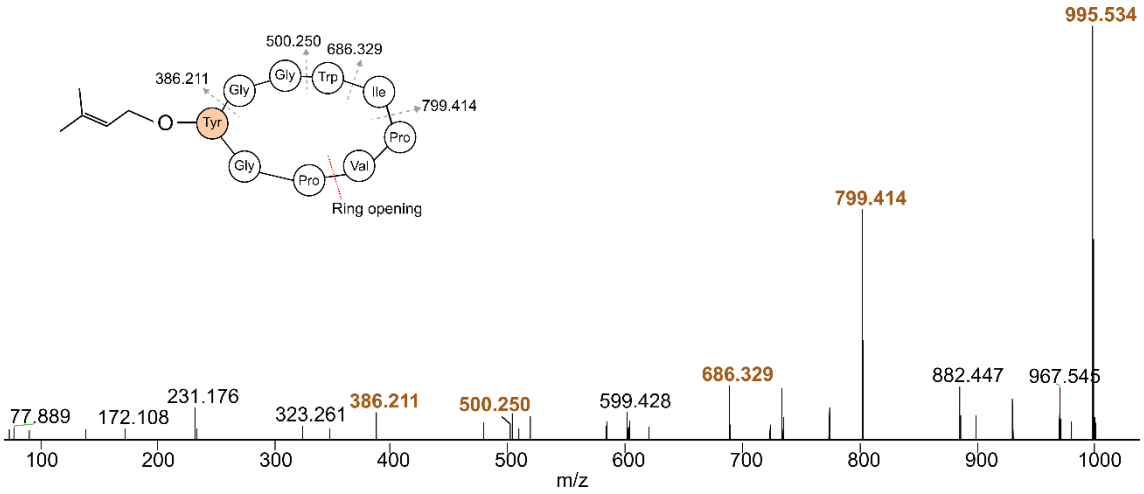



Figure 25 MS/MS data from cyanobactin 994 (m/z 995.5349).

The aminoacidic sequence of the peptide cyanobactin 884 (*m/z* 885.4539) was successfully validated through the examination of MS/MS data. For instance, it was possible to identify the parent mass *m/z* 885.453, and the ion fragment of *m/z* 557.280, which corresponds to “LSGA” and the fragment *m/z* 460.218, which represents the “PY(prenyl)M” sequence (Figure 26).

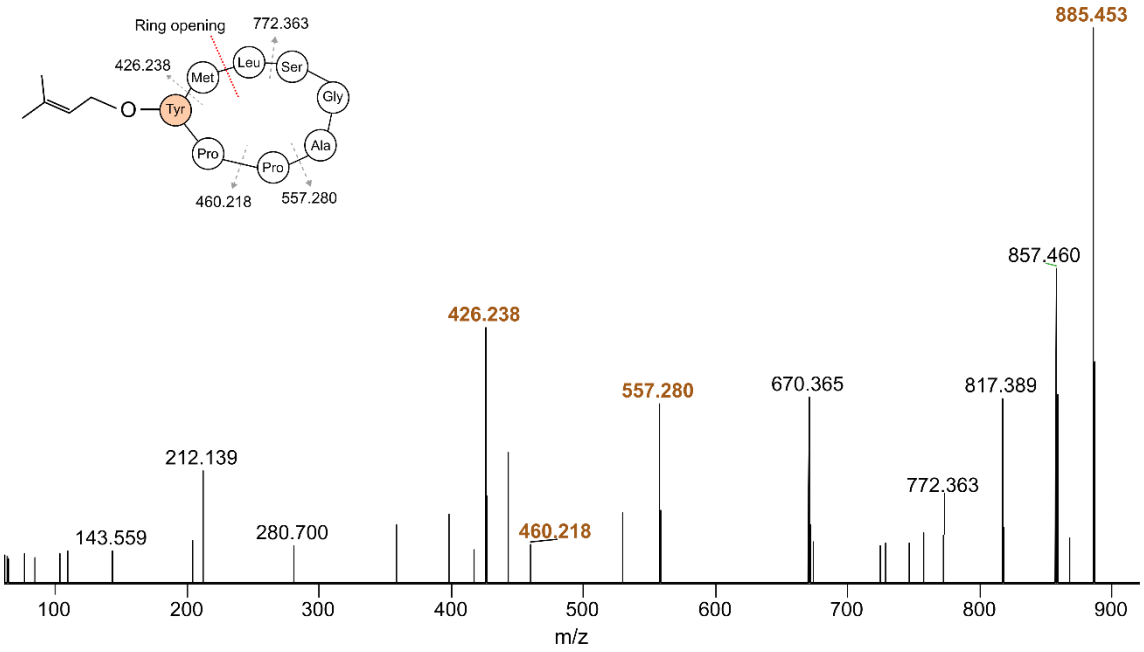



Figure 26 MS/MS data from cyanobactin 884 (m/z 885.4539)

In the case of the cyanobactins encoded by the OscE3, which presents a mono- and bis-prenylated versions named 1308 (*m/z* 1309.6099) and 1376 (*m/z* 1377.6725), respectively, the bis-prenylated version was firstly elucidated. As observed in the Figure 27, it is possible to identify an evident peak with *m/z* 774.418 which corresponds to the ion fragment of the sequence “GIGDY(prenyl)AE”. This fragment was not present on the MS/MS spectra of mono-prenylated cyanobactin 1308, which facilitated the identification in which tyrosine residue this modification occurs, being the tyrosine 43 (position based on the precursor sequence) the prenylated residue in cyanobactin 1308. In fact, an ion *m/z* 706.35 fragment was observed on MS/MS spectra of cyanobactin 1308, which corresponds to the sequence “GIGDYAE” without prenylation on the Tyr residue.

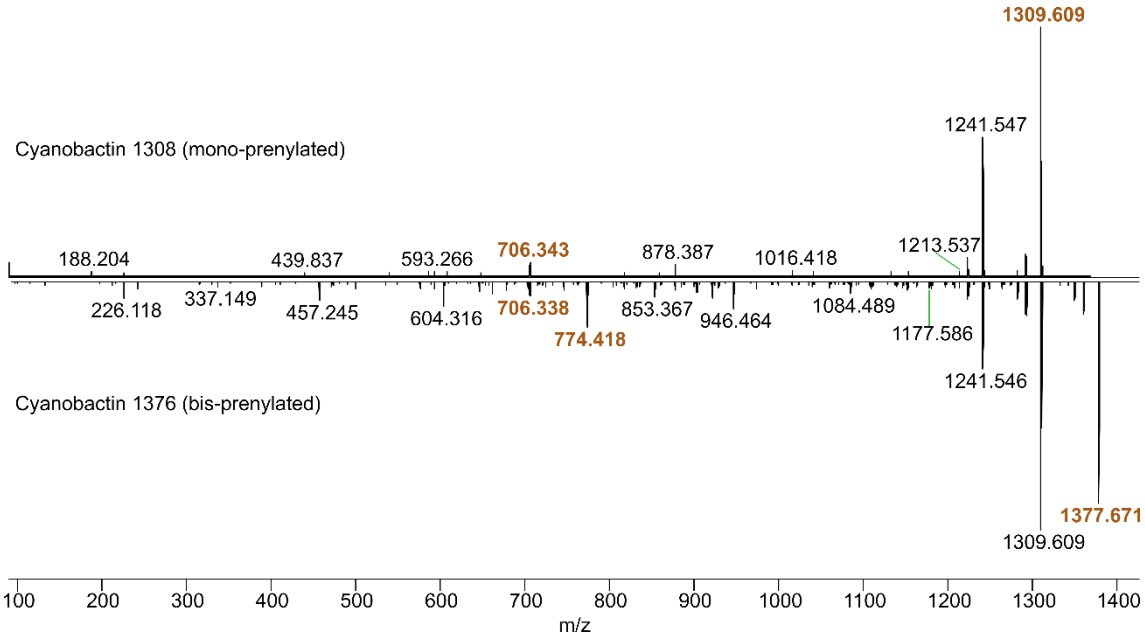



Figure 27 Comparison of MS/MS data from mono-prenylated cyanobactin 1308 (m/z 1309.6099, top) and bis-prenylated cyanobactin 1376 (m/z 1377.6725, bottom).

Despite the intensity of the majority of the fragments were very low, it was possible to identify all the aminoacidic sequence from the cyanobactins. In addition, it was indeed feasible to determine the existence of the prenyl group in each cyanobactin since the loss of 68.03 was detected in all prenylated versions and a difference of 132.12 in case of cyanobactin 1376. It is important to emphasize that despite the prenylation was confirmed on Tyr residue, the direction is assumed to be in forward orientation for the cyanobactins with structure elucidation based on MS/MS data since this was only confirmed by NMR analysis in cyanobactin 850. Forward prenylation on Tyr residue was already reported, also in a LEGE strain, *Sphaerospermopsis* sp. LEGE 00249, which produces sphaerocyclamide, a macrocyclic prenylated cyanobactin. However, the double prenylation on Tyr performed by only one prenyltransferase was never reported, demonstrating the novelty of this work. However, at this point, we cannot ensure that this double prenylation is occurring in the same orientation and to clarify this question additional studies are necessary. Furthermore, in the future, it will be also interesting to perform the biochemical characterization of OscF and study their substrate selectivity.

## 3.6. Bioactivity Assays

### 3.6.1. Cytotoxic Properties of Cyanobactin 850 (**1**) and Cyanobactin 1027 (**2**)

The most distinct activity associated with cyanobactins is their cytotoxicity [61]. Ma Yadanar Phyo *et al*. isolated three new cyanobactins from the marine cyanobacterium *Symploca hydnoides*, and found that two of them, trikoramides B and D, exhibited cytotoxicity against MOLT-4 acute lymphoblastic leukemia cell line [122].

In this work, the cytotoxic properties of two isolated mono-prenylated cyanobactins, namely cyanobactin 850 (**1**) and cyanobactin 1027 (**2**) were investigated against HCT116, HepG2, and SH-SY5Y cancer cells (Figure 28). Following a 24-hour exposure, the compound **1** resulted in a reduction of HCT116 cell viability to 84.1%±1.9%. However, no significant cytotoxicity was observed in the other cell lines. On the other hand, compound **2** consistently exhibited cytotoxicity, causing a decrease in cell viability to 87.1%±4.2% across all the cell lines studied. Overall, our findings indicate that the isolated cyanobactins displayed limited efficacy against the three cell lines, leading us to conclude that neither of them possesses cytotoxic properties in the cell lines tested.

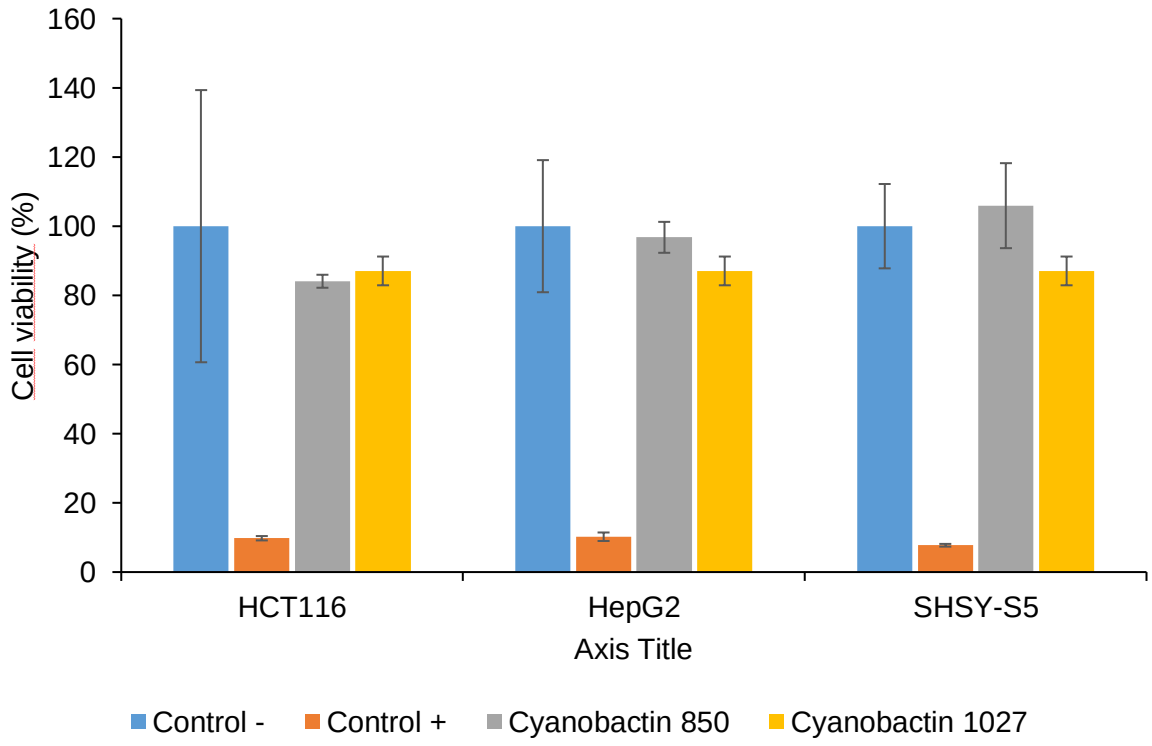



Figure 28 Cytotoxicity assays results in HCT116, HepG2 and SHSY-S5cell lines for cyanobactin 850 (**1**) and cyanobactin 1027 (**2**).

### 3.6.2. Antimicrobial properties of Cyanobactin 850 (**1**)

Beside cytotoxic compounds, cyanobacteria are also known to produce a wide range of peptides with antibacterial activities [34]. Recently, the discovery of new prenylated cyanobactins, such as sphaerocyclamide, has expanded the knowledge of their bioactivity [77]. These compounds have been tested in biological assays and have demonstrated moderate antimicrobial activity against bacteria such as *Halomonas aquamarina* CECT 5000 [77].

As a result of the limited supply of cyanobactin 1027, the remaining bioactivity assays were conducted testing only cyanobactin 850 (**1**). To assess the antimicrobial potential of cyanobactin 850 (**1**) several pathogenic strains described in the literature as a model for bioactivity assays were selected, namely *Staphylococcus aureus* ATCC 29213 and *Bacillus subtilis* ATCC 6633*, Escherichia coli* ATCC 25922 and *Salmonela typhimurium ATCC* 25241 and the yeast *Candida albicans* ATCC 10231. However, no relevant anti-bacterial and anti-fungal were found as it possible to observe in Figure 28.

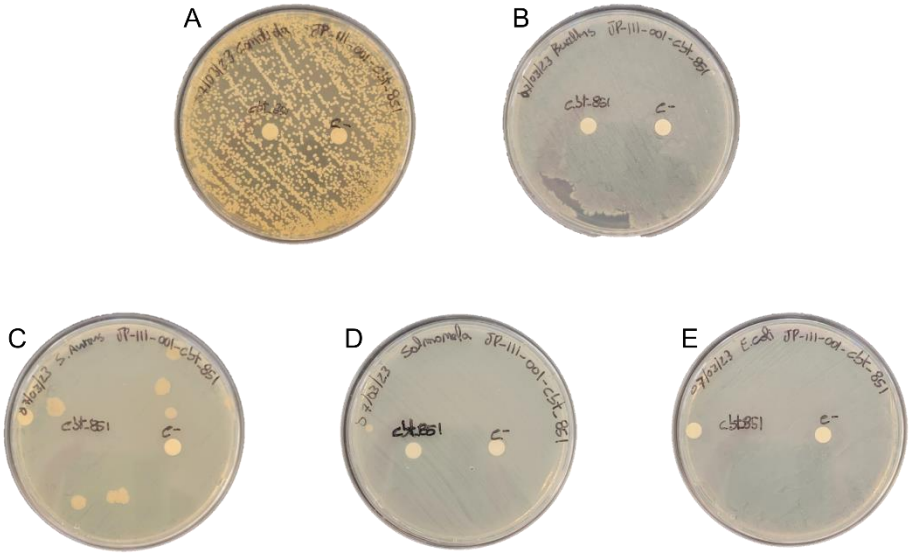



Figure 29 Antibiograms of cyanobactin 850 (**1**) against *Candida albicans ATCC 10231* (A), *Bacillus subtilis ATCC 6633* (B), *Staphylococcus aureus ATCC 29213* (C), Salmonela typhimurium ATCC 25241 (D) and *Escherichia coli* ATCC 25922 (E).

### 3.6.3. Antiamoebic properties of Cyanobactin 850 (**1**)

In the environment, amoeba is classified as a natural grazer of cyanobacteria, however, it has been observed that certain strains of cyanobacteria have exhibited a degree of resistance to such consumption [123, 124]. Since cyanobactin 850 (**1**) was one of the most abundant compounds produced by Oscillatoriales cyanobacterium LEGE 16532, this led to the investigation of the antiamoebic properties of compound **1** in order to understand the possible ecological role of this cyanobactin. The assessment

of antiamoebic activity of compound **1** did not reveal any direct inhibitory effect against the tested amoeba strains (Figure 29). However, the presence of cyanobactin 850 (**1**) appeared to enhance the viability of amoebae, leading to increased amoebic growth and survival. This was also observed by Urrutia-Cordero *et al*, where they exposed pure microcystin and six different strains of cyanobacteria to the freshwater amoeba *Acanthamoeba castellanii*. Among these strains, three were known to produce microcystin while the other three contained other oligopeptides such as anabaenopeptins and cyanopeptolins [125]. In that study, *Anabaena lemmermannii* NIVA-CYA 426 was exposed to amoebae, and population densities were slightly decreased during the first five days of the incubation period. However, growth was subsequently recovered to rates higher than controls, indicating that it was represented as a suitable food for amoebae by promoting its growth [125].

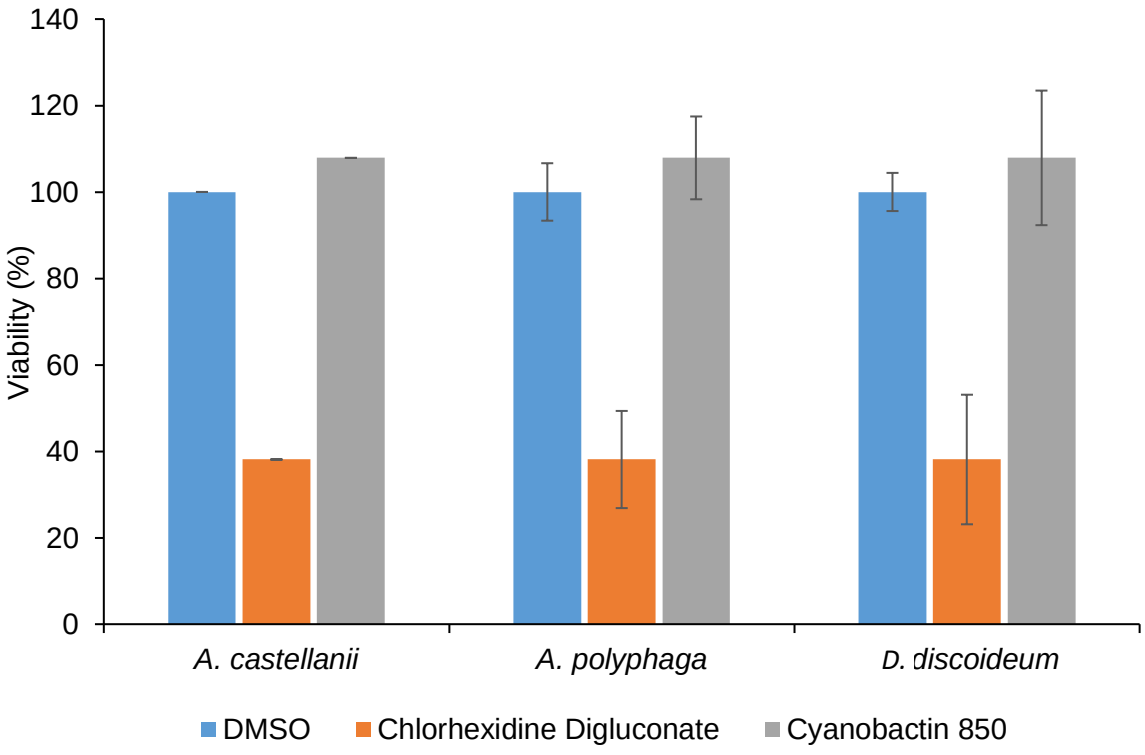



Figure 30 Dose response assay to detect anti-amoebic activity of cyanobactin 850 (**1**).

# 4. Conclusion and future perspetives

Cyanobacteria are a diverse group of photosynthetic microorganisms that have become a promising source of NPs. Among the various classes of cyanobacterial NPs, RiPPs are of particular interest due to their simple biosynthesis. RiPPs are capable of

generate a vast structural diversity through extensive post-translational modifications of the precursor peptide. Cyanobactins, a subclass of RiPPs, are small cyclic and linear peptides that are produced by a diverse selection of cyanobacteria and their functions are likewise diverse and include antibacterial, antifungal, cytotoxic, antibacterial and antimalarial activity that would make them important molecules for clinical and pharmacological applications [61, 68].

Over the past two decades, the rapid advancement of genome mining tools, along with improvements in genomic and metabolomic research areas, has uncovered an unexpected abundance of novel RiPP BGCs [44]. Moreover, many bioinformatic tools have been developed to identify BGCs, most of which have no corresponding NPs. Hence, opportunities abound for the discovery of new chemistry and enzymology associated with NPs.

In this work, ten new metabolites, including cyclic, mono, and bis-prenylated cyanobactins were discovered and based on NMR analysis it was confirmed that prenylation occurs on the tyrosine residue presenting forward orientation, which is assumed to be present in all the other cyanobactins that had their structure elucidated only by MS/MS data. This post-translational modification is a crucial mechanism that enhances the lipophilicity of molecules, enabling them to interact more effectively with lipid membranes, thus improving the biological activities of the secondary metabolites. Although the isolated cyanobactins exhibited no cytotoxicity or antimicrobial activity in various bioactivity assays, further investigations are warranted to expand the scope of testing for these compounds.

In conclusion, the prenyltransferase identified in this study holds the potential to be utilized as a valuable tool in the field of synthetic biology and drug development. Additionally, the utilization of combinatorial biosynthesis techniques is considered a promising approach for expanding the chemical diversity of cyanobactins by providing a platform for the rapid and efficient production of customized peptide-based therapeutics.

Looking ahead, it is imperative to fully harness the potential of this prenyltransferase. In this way, future work includes the use of combinatorial biosynthesis to broaden the chemical diversity of cyanobactins and produce new-to-nature peptides.

# 5. References

1. Schirrmeister, B.E., M. Gugger, and P.C. Donoghue, *Cyanobacteria and the Great Oxidation Event: evidence from genes and fossils.* Palaeontology, 2015. **58**(5): p. 769-785.
2. Nutman, A.P., et al., *Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures.* Nature, 2016. **537**(7621): p. 535-538.
3. Mehdizadeh Allaf, M. and H. Peerhossaini, *Cyanobacteria: Model Microorganisms and Beyond.* Microorganisms, 2022. **10**(4): p. 696.
4. Garcia-Pichel, F., *Cyanobacteria*, in *Encyclopedia of Microbiology (Third Edition)*, M. Schaechter, Editor. 2009, Academic Press: Oxford. p. 107-124.
5. Mutalipassi, M., et al., *Symbioses of Cyanobacteria in Marine Environments: Ecological Insights and Biotechnological Perspectives.* Mar Drugs, 2021. **19**(4): p. 227.
6. Bahareh, N., et al., *Plant-cyanobacteria interactions: Beneficial and harmful effects of cyanobacterial bioactive compounds on soil-plant systems and subsequent risk to animal and human health.* Phytochemistry, 2021. **192**: p. 112959.
7. Aschenbrenner, I.A., T. Cernava, G. Berg, and M. Grube, *Understanding Microbial Multi-Species Symbioses.* Front Microbiol, 2016. **7**: p. 180.
8. Foster, R.A., et al., *Nitrogen fixation and transfer in open ocean diatom-cyanobacterial symbioses.* ISME J, 2011. **5**(9): p. 1484-93.
9. Dextro, R.B., et al., *Trends in Free-access Genomic Data Accelerate Advances in Cyanobacteria Taxonomy.* J Phycol, 2021. **57**(5): p. 1392-1402.
10. Palinska, K.A. and W. Surosz, *Taxonomy of cyanobacteria: a contribution to consensus approach.* Hydrobiologia, 2014. **740**(1): p. 1-11.
11. Oren, A. and S. Ventura, *The current status of cyanobacterial nomenclature under the "prokaryotic" and the "botanical" code.* Antonie Van Leeuwenhoek, 2017. **110**(10): p. 1257-1269.
12. Strunecky, O., A.P. Ivanova, and J. Mares, *An updated classification of cyanobacterial orders and families based on phylogenomic and polyphasic analysis.* J Phycol, 2023. **59**(1): p. 12-51.
13. Komárek J., K.J., Mareš J. & Johansen J. R. , *Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach.* Preslia, 2014. **86**(4): p. 295-335.

14. Ramos, V., et al., *Cyanobacterial diversity held in microbial biological resource centers as a biotechnological asset: the case study of the newly established LEGE culture collection.* J Appl Phycol, 2018. **30**(3): p. 1437-1451.
15. Singh, P., A. Khan, and A. Srivastava, *Heterocyst and akinete differentiation in cyanobacteria: a view toward cyanobacterial symbiosis*, in *Advances in Cyanobacterial Biology*, P.K. Singh, A. Kumar, V.K. Singh, and A.K. Shrivastava, Editors. 2020, Academic Press. p. 235-248.
16. Garg, R. and I. Maldener, *The Formation of Spore-Like Akinetes: A Survival Strategy of Filamentous Cyanobacteria.* Microb Physiol, 2021. **31**(3): p. 296-305.
17. Dias, D.A., S. Urban, and U. Roessner, *A historical overview of natural products in drug discovery.* Metabolites, 2012. **2**(2): p. 303-36.
18. Nicolaou, K.C. and S. Rigol, *A brief history of antibiotics and select advances in their synthesis.* J Antibiot (Tokyo), 2018. **71**(2): p. 153-184.
19. Petersen, L.-E., M.Y. Kellermann, and P.J. Schupp, *Secondary Metabolites of Marine Microbes: From Natural Products Chemistry to Chemical Ecology*, in *YOUMARES 9 - The Oceans: Our Research, Our Future*. 2020, Springer International Publishing. p. 159-180.
20. Luo, Y., R.E. Cobb, and H. Zhao, *Recent advances in natural product discovery.* Curr Opin Biotechnol, 2014. **30**: p. 230-7.
21. Atanasov, A.G., et al., *Natural products in drug discovery: advances and opportunities.* Nat Rev Drug Discov, 2021. **20**(3): p. 200-216.
22. Didelot, X., et al., *Transforming clinical microbiology with bacterial genome sequencing.* Nat Rev Genet, 2012. **13**(9): p. 601-612.
23. Katz, L. and R.H. Baltz, *Natural product discovery: past, present, and future.* J Ind Microbiol Biotechnol, 2016. **43**(2-3): p. 155-76.
24. Dittmann, E., M. Gugger, K. Sivonen, and D.P. Fewer, *Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria.* Trends Microbiol, 2015. **23**(10): p. 642-652.
25. van Santen, J.A., et al., *The Natural Products Atlas: An Open Access Knowledge Base for Microbial Natural Products Discovery.* ACS Cent Sci, 2019. **5**(11): p. 1824-1833.
26. Agrawal, P., et al., *RiPPMiner-Genome: A Web Resource for Automated Prediction of Crosslinked Chemical Structures of RiPPs by Genome Mining.* J Mol Biol, 2021. **433**(11): p. 166887.
27. van Heel, A.J., et al., *BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins.* Nucleic Acids Res, 2018. **46**(W1): p. W278-W281.

28. Yan, F., et al., *Synthetic biology approaches and combinatorial biosynthesis towards heterologous lipopeptide production.* Chem Sci, 2018. **9**(38): p. 7510-7519.
29. Xie, Y., et al., *Activation and enhancement of caerulomycin A biosynthesis in marine-derived Actinoalloteichus sp. AHMU CJ021 by combinatorial genome mining strategies.* Microb Cell Fact, 2020. **19**(1): p. 159.
30. Soldatou, S., et al., *Linking biosynthetic and chemical space to accelerate microbial secondary metabolite discovery.* FEMS Microbiol Lett, 2019. **366**(13).
31. Gugger, M., A. Boullie, and T. Laurent, *Cyanotoxins and Other Bioactive Compounds from the Pasteur Cultures of Cyanobacteria (PCC).* Toxins (Basel), 2023. **15**(6): p. 388.
32. Zhang, W., et al., *The Impact of Cyanobacteria Blooms on the Aquatic Environment and Human Health.* Toxins (Basel), 2022. **14**(10): p. 658.
33. Tidgewell, K.J., B.R. Clark, and W.H. Gerwick. *The Natural Products Chemistry of Cyanobacteria*. 2010.
34. Demay, J., C. Bernard, A. Reinhardt, and B. Marie, *Natural Products from Cyanobacteria: Focus on Beneficial Activities.* Mar Drugs, 2019. **17**(6): p. 320.
35. Engene, N., et al., *Moorea producens gen. nov., sp. nov. and Moorea bouillonii comb. nov., tropical marine cyanobacteria rich in bioactive secondary metabolites.* Int J Syst Evol Microbiol, 2012. **62**(Pt 5): p. 1171-1178.
36. Gerwick, W.H., et al., *Giant Marine Cyanobacteria Produce Exciting Potential Pharmaceuticals.* Microbe Magazine, 2008. **3**: p. 277-284.
37. Schlegel, I., N.T. Doan, N. de Chazal, and G.D. Smith, *Antibiotic activity of new cyanobacterial isolates from Australia and Asia against green algae and cyanobacteria.* Journal of Applied Phycology, 1998. **10**(5): p. 471-479.
38. Shih, P.M., et al., *Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing.* Proc Natl Acad Sci U S A, 2013. **110**(3): p. 1053-8.
39. Chang, Z., et al., *Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium Lyngbya majuscula.* J Nat Prod, 2004. **67**(8): p. 1356-67.
40. Luesch, H., et al., *Isolation of dolastatin 10 from the marine cyanobacterium Symploca species VP642 and total stereochemistry and biological evaluation of its analogue symplostatin 1.* J Nat Prod, 2001. **64**(7): p. 907-10.
41. Gao, G., et al., *Marine Antitumor Peptide Dolastatin 10: Biological Activity, Structural Modification and Synthetic Chemistry.* Mar Drugs, 2021. **19**(7): p. 363.

42. Joubert, N., A. Beck, C. Dumontet, and C. Denevault-Sabourin, *Antibody-Drug Conjugates: The Last Decade.* Pharmaceuticals (Basel), 2020. **13**(9): p. 245.
43. Arnison, P.G., et al., *Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature.* Nat Prod Rep, 2013. **30**(1): p. 108-60.
44. Russell, A.H. and A.W. Truman, *Genome mining strategies for ribosomally synthesised and post-translationally modified peptides.* Comput Struct Biotechnol J, 2020. **18**: p. 1838-1851.
45. Zhong, G., et al., *Recent Advances in Discovery, Bioengineering, and Bioactivity-Evaluation of Ribosomally Synthesized and Post-translationally Modified Peptides.* ACS Bio Med Chem Au, 2023. **3**(1): p. 1-31.
46. Ortega, M.A. and W.A. van der Donk, *New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products.* Cell Chem Biol, 2016. **23**(1): p. 31-44.
47. Montalban-Lopez, M., et al., *New developments in RiPP discovery, enzymology and engineering.* Nat Prod Rep, 2021. **38**(1): p. 130-239.
48. Vogt, E. and M. Kunzler, *Discovery of novel fungal RiPP biosynthetic pathways and their application for the development of peptide therapeutics.* Appl Microbiol Biotechnol, 2019. **103**(14): p. 5567-5581.
49. Jiang, H., et al., *Protein Lipidation: Occurrence, Mechanisms, Biological Functions, and Enabling Technologies.* Chem Rev, 2018. **118**(3): p. 919-988.
50. Fewer, D.P., et al., *Chemical diversity and cellular effects of antifungal cyclic lipopeptides from cyanobacteria.* Physiol Plant, 2021. **173**(2): p. 639-650.
51. Hutchinson, J.A., S. Burholt, and I.W. Hamley, *Peptide hormones and lipopeptides: from self-assembly to therapeutic applications.* J Pept Sci, 2017. **23**(2): p. 82-94.
52. Roongsawang, N., K. Washio, and M. Morikawa, *Diversity of nonribosomal peptide synthetases involved in the biosynthesis of lipopeptide biosurfactants.* Int J Mol Sci, 2010. **12**(1): p. 141-72.
53. Miller, B.R. and A.M. Gulick, *Structural Biology of Nonribosomal Peptide Synthetases*, in *Methods in Molecular Biology*. 2016, Springer New York. p. 3-29.
54. McErlean, M., J. Overbay, and S. Van Lanen, *Refining and expanding nonribosomal peptide synthetase function and mechanism.* J Ind Microbiol Biotechnol, 2019. **46**(3-4): p. 493-513.

55. Eusebio, N., et al., *Discovery and Heterologous Expression of Microginins from Microcystis aeruginosa LEGE 91341.* ACS Synth Biol, 2022. **11**(10): p. 3493-3503.
56. Neuhof, T., et al., *Hassallidin A, a glycosylated lipopeptide with antifungal activity from the cyanobacterium Hassallia sp.* J Nat Prod, 2005. **68**(5): p. 695-700.
57. Vestola, J., et al., *Hassallidins, antifungal glycolipopeptides, are widespread among cyanobacteria and are the end-product of a nonribosomal pathway.* Proc Natl Acad Sci U S A, 2014. **111**(18): p. E1909-17.
58. Mares, J., et al., *A hybrid non-ribosomal peptide/polyketide synthetase containing fatty-acyl ligase (FAAL) synthesizes the beta-amino fatty acid lipopeptides puwainaphycins in the Cyanobacterium Cylindrospermum alatosporum.* PLoS One, 2014. **9**(11): p. e111904.
59. Freitas, S., et al., *Structure and Biosynthesis of Desmamides A-C, Lipoglycopeptides from the Endophytic Cyanobacterium Desmonostoc muscorum LEGE 12446.* J Nat Prod, 2022. **85**(7): p. 1704-1714.
60. Hubrich, F., et al., *Ribosomally derived lipopeptides containing distinct fatty acyl moieties.* Proc Natl Acad Sci U S A, 2022. **119**(3): p. e2113120119.
61. Martins, J. and V. Vasconcelos, *Cyanobactins from Cyanobacteria: Current Genetic and Chemical State of Knowledge.* Mar Drugs, 2015. **13**(11): p. 6910-46.
62. Donia, M.S. and E.W. Schmidt, *Linking chemistry and genetics in the growing cyanobactin natural products family.* Chem Biol, 2011. **18**(4): p. 508-19.
63. Leikoski, N., et al., *Highly diverse cyanobactins in strains of the genus Anabaena.* Appl Environ Microbiol, 2010. **76**(3): p. 701-9.
64. Leikoski, N., et al., *Genome mining expands the chemical diversity of the cyanobactin family to include highly modified linear peptides.* Chem Biol, 2013. **20**(8): p. 1033-43.
65. Schmidt, E.W., et al., *Patellamide A and C biosynthesis by a microcin-like pathway in Prochloron didemni, the cyanobacterial symbiont of Lissoclinum patella.* Proc Natl Acad Sci U S A, 2005. **102**(20): p. 7315-20.
66. Caba, J.M., et al., *Solid-phase total synthesis of trunkamide A(1).* J Org Chem, 2001. **66**(23): p. 7568-74.
67. Lawton, L.A., L.A. Morris, and M. Jaspars, *A Bioactive Modified Peptide, Aeruginosamide, Isolated from the Cyanobacterium Microcystis aeruginosa.* J Org Chem, 1999. **64**(14): p. 5329-5332.

68. Sivonen, K., N. Leikoski, D.P. Fewer, and J. Jokela, *Cyanobactins-ribosomal cyclic peptides produced by cyanobacteria.* Appl Microbiol Biotechnol, 2010. **86**(5): p. 1213-25.
69. Zhang, Y., Y. Goto, and H. Suga, *Discovery, biochemical characterization, and bioengineering of cyanobactin prenyltransferases.* Trends Biochem Sci, 2023. **48**(4): p. 360-374.
70. Gu, W., D. Sardar, E. Pierce, and E.W. Schmidt, *Roads to Rome: Role of Multiple Cassettes in Cyanobactin RiPP Biosynthesis.* J Am Chem Soc, 2018. **140**(47): p. 16213-16221.
71. Zheng, Y., Y. Cong, E.W. Schmidt, and S.K. Nair, *Catalysts for the Enzymatic Lipidation of Peptides.* Acc Chem Res, 2022. **55**(9): p. 1313-1323.
72. Tanner, M.E., *Mechanistic studies on the indole prenyltransferases.* Nat Prod Rep, 2015. **32**(1): p. 88-101.
73. Luk, L.Y. and M.E. Tanner, *Mechanism of dimethylallyltryptophan synthase: evidence for a dimethylallyl cation intermediate in an aromatic prenyltransferase reaction.* J Am Chem Soc, 2009. **131**(39): p. 13932-3.
74. McIntosh, J.A., M.S. Donia, S.K. Nair, and E.W. Schmidt, *Enzymatic basis of ribosomal peptide prenylation in cyanobacteria.* J Am Chem Soc, 2011. **133**(34): p. 13698-705.
75. McIntosh, J.A., Z. Lin, M.D.B. Tianero, and E.W. Schmidt, *Aestuaramides, a Natural Library of Cyanobactin Cyclic Peptides Resulting from Isoprene-Derived Claisen Rearrangements.* ACS Chemical Biology, 2013. **8**(5): p. 877-883.
76. Murakami, M., Y. Itou, K. Ishida, and H.J. Shin, *Prenylagaramides A and B, new cyclic peptides from two strains of oscillatoria agardhii.* J Nat Prod, 1999. **62**(5): p. 752-5.
77. Martins, J., et al., *Sphaerocyclamide, a prenylated cyanobactin from the cyanobacterium Sphaerospermopsis sp. LEGE 00249.* Sci Rep, 2018. **8**(1): p. 14537.
78. Hao, Y., et al., *Molecular basis for the broad substrate selectivity of a peptide prenyltransferase.* Proceedings of the National Academy of Sciences, 2016. **113**(49): p. 14037-14042.
79. Parajuli, A., et al., *A Unique Tryptophan C-Prenyltransferase from the Kawaguchipeptin Biosynthetic Pathway.* Angew Chem Int Ed Engl, 2016. **55**(11): p. 3596-9.

80. Ishida, K., H. Matsuda, M. Murakami, and K. Yamaguchi, *Kawaguchipeptin A, a novel cyclic undecapeptide from cyanobacterium Microcystis aeruginosa (NIES-88).* Tetrahedron, 1996. **52**(27): p. 9025-9030.
81. Dalponte, L., et al., *N-Prenylation of Tryptophan by an Aromatic Prenyltransferase from the Cyanobactin Biosynthetic Pathway.* Biochemistry, 2018. **57**(50): p. 6860-6867.
82. Tianero, M.D., et al., *Metabolic model for diversity-generating biosynthesis.* Proc Natl Acad Sci U S A, 2016. **113**(7): p. 1772-7.
83. Carroll, A., et al., *Patellins 1-6 and Trunkamide A: Novel Cyclic Hexa-, Hepta- and Octa-peptides From Colonial Ascidians, <I>Lissoclinum</I> sp.* Australian Journal of Chemistry, 1996. **49**(6): p. 659-667.
84. Purushothaman, M., et al., *Genome-Mining-Based Discovery of the Cyclic Peptide Tolypamide and TolF, a Ser/Thr Forward O-Prenyltransferase.* Angew Chem Int Ed Engl, 2021. **60**(15): p. 8460-8465.
85. Phan, C.S., et al., *Argicyclamides A-C Unveil Enzymatic Basis for Guanidine Bis-prenylation.* J Am Chem Soc, 2021. **143**(27): p. 10083-10087.
86. Clemente, C., et al., *Biochemical characterization of a cyanobactin arginine-N-prenylase from the autumnalamide biosynthetic pathway.* Chem Commun (Camb), 2022. **58**(86): p. 12054-12057.
87. Zhang, Y., et al., *LimF is a versatile prenyltransferase catalyzing histidine-C-geranylation on diverse nonnatural substrates*. 2021, American Chemical Society (ACS).
88. Cong, Y., P.D. Scesa, and E.W. Schmidt, *AgeMTPT, a Catalyst for Peptide N-Terminal Modification.* ACS Synth Biol, 2022. **11**(11): p. 3699-3705.
89. Mattila, A., et al., *Biosynthesis of the Bis-Prenylated Alkaloids Muscoride A and B.* ACS Chem Biol, 2019. **14**(12): p. 2683-2690.
90. Morita, M., et al., *Post-Translational Tyrosine Geranylation in Cyanobactin Biosynthesis.* J Am Chem Soc, 2018. **140**(19): p. 6044-6048.
91. Santos-Merino, M., A.K. Singh, and D.C. Ducat, *New Applications of Synthetic Biology Tools for Cyanobacterial Metabolic Engineering.* Front Bioeng Biotechnol, 2019. **7**: p. 33.
92. Garner, K.L., *Principles of synthetic biology.* Essays Biochem, 2021. **65**(5): p. 791-811.
93. D'Agostino, P.M. and T.A.M. Gulder, *Direct Pathway Cloning Combined with Sequence- and Ligation-Independent Cloning for Fast Biosynthetic Gene Cluster*

*Refactoring and Heterologous Expression.* ACS Synth Biol, 2018. **7**(7): p. 1702-1708.

94. Do, T. and A.J. Link, *Protein Engineering in Ribosomally Synthesized and Post-translationally Modified Peptides (RiPPs).* Biochemistry, 2023. **62**(2): p. 201-209.
95. Ruijne, F. and O.P. Kuipers, *Combinatorial biosynthesis for the generation of new-to-nature peptide antimicrobials.* Biochem Soc Trans, 2021. **49**(1): p. 203-215.
96. Sardar, D., Z. Lin, and E.W. Schmidt, *Modularity of RiPP Enzymes Enables Designed Synthesis of Decorated Peptides.* Chem Biol, 2015. **22**(7): p. 907-16.
97. Pan, S.J. and A.J. Link, *Sequence diversity in the lasso peptide framework: discovery of functional microcin J25 variants with multiple amino acid substitutions.* J Am Chem Soc, 2011. **133**(13): p. 5016-23.
98. Schmitt, S., et al., *Analysis of modular bioengineered antimicrobial lanthipeptides at nanoliter scale.* Nat Chem Biol, 2019. **15**(5): p. 437-443.
99. Zhao, X., et al., *High-Throughput Screening for Substrate Specificity-Adapted Mutants of the Nisin Dehydratase NisB.* ACS Synth Biol, 2020. **9**(6): p. 1468-1478.
100. Glassey, E., et al., *Functional expression of diverse post-translational peptide-modifying enzymes in Escherichia coli under uniform expression and purification conditions.* PLoS One, 2022. **17**(9): p. e0266488.
101. Burkhart, B.J., et al., *Chimeric Leader Peptides for the Generation of Non-Natural Hybrid RiPP Products.* ACS Cent Sci, 2017. **3**(6): p. 629-638.
102. Kotai, J., *Instructions for preparation of modified nutrient solution Z8 for algae.* 1972.
103. Wood, D.E. and S.L. Salzberg, *Kraken: ultrafast metagenomic sequence classification using exact alignments.* Genome Biol, 2014. **15**(3): p. R46.
104. Prjibelski, A., et al., *Using SPAdes De Novo Assembler.* Curr Protoc Bioinformatics, 2020. **70**(1): p. e102.
105. Wu, Y.W., B.A. Simmons, and S.W. Singer, *MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets.* Bioinformatics, 2016. **32**(4): p. 605-7.
106. Blin, K., et al., *antiSMASH 6.0: improving cluster detection and comparison capabilities.* Nucleic Acids Res, 2021. **49**(W1): p. W29-W35.
107. Gerlt, J.A., et al., *Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks.* Biochim Biophys Acta, 2015. **1854**(8): p. 1019-37.

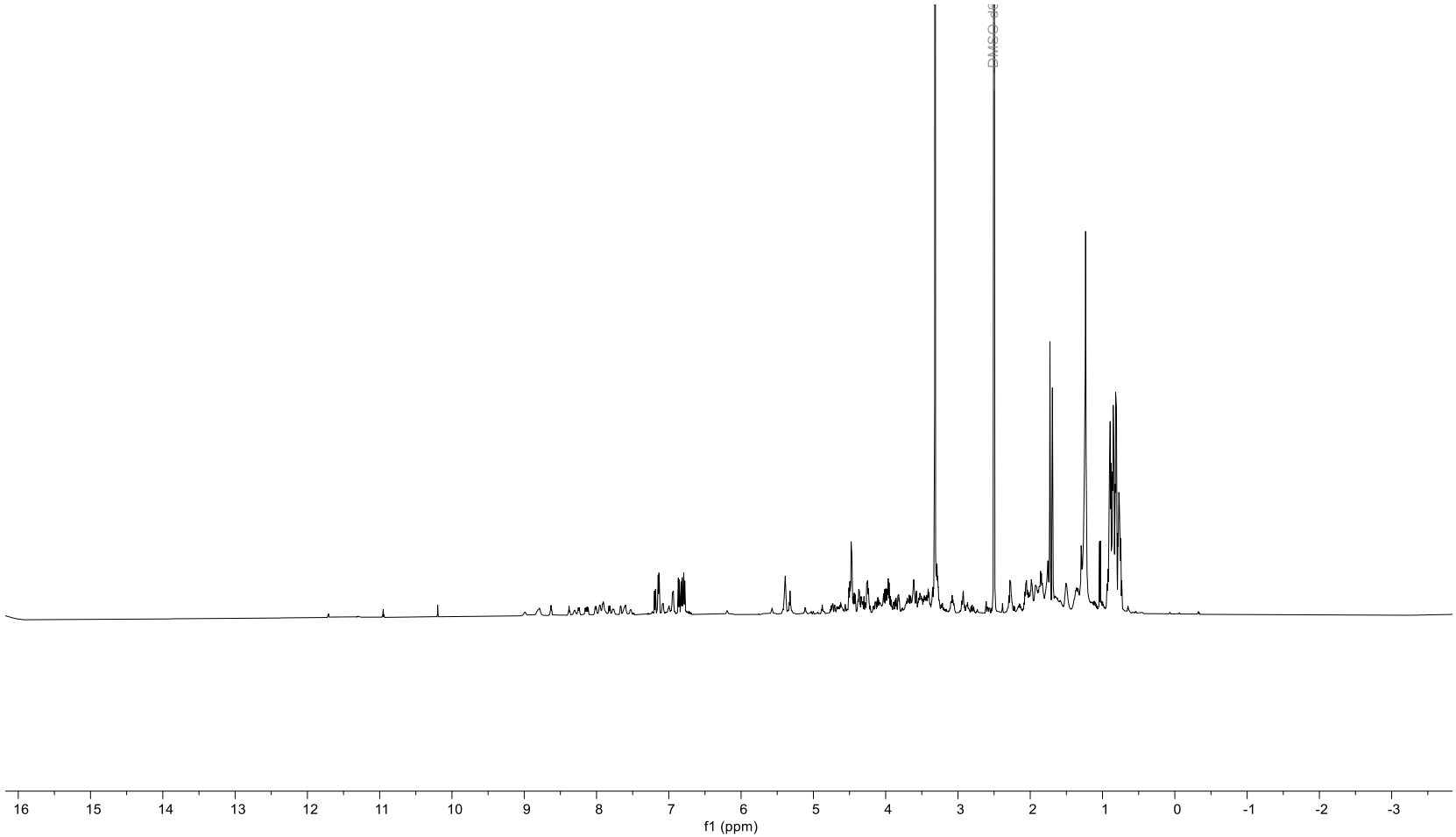
108. Camacho, C., et al., *ElasticBLAST: accelerating sequence search via cloud computing.* BMC Bioinformatics, 2023. **24**(1): p. 117.
109. Dickert, H., K. Machka, and I. Braveny, *The uses and limitations of disc diffusion in the antibiotic sensitivity testing of bacteria.* Infection, 1981. **9**(1): p. 18-24.
110. Bachmann, B.O., S.G. Van Lanen, and R.H. Baltz, *Microbial genome mining for accelerated natural products discovery: is a renaissance in the making?* J Ind Microbiol Biotechnol, 2014. **41**(2): p. 175-84.
111. Gu, W., et al., *The Biochemistry and Structural Biology of Cyanobactin Pathways: Enabling Combinatorial Biosynthesis*, in *Methods in Enzymology*. 2018, Elsevier. p. 113-163.
112. Sardar, D., E. Pierce, J.A. McIntosh, and E.W. Schmidt, *Recognition sequences and substrate evolution in cyanobactin biosynthesis.* ACS Synth Biol, 2015. **4**(2): p. 167-76.
113. Mohimani, H., et al., *Automated genome mining of ribosomal peptide natural products.* ACS Chem Biol, 2014. **9**(7): p. 1545-51.
114. Sarkar, S., W. Gu, and E.W. Schmidt, *Expanding the chemical space of synthetic cyclic peptides using a promiscuous macrocyclase from prenylagaramide biosynthesis.* ACS Catal, 2020. **10**(13): p. 7146-7153.
115. Zhang, Y.C., et al., *LimF is a versatile prenyltransferase for histidine-C-geranylation on diverse non-natural substrates.* Nature Catalysis, 2022. **5**(8): p. 682-+.
116. Oberg, N., R. Zallot, and J.A. Gerlt, *EFI-EST, EFI-GNT, and EFI-CGFP: Enzyme Function Initiative (EFI) Web Resource for Genomic Enzymology Tools.* J Mol Biol, 2023. **435**(14): p. 168018.
117. Bent, A.F., et al., *Structure of PatF from Prochloron didemni.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2013. **69**(Pt 6): p. 618-23.
118. Audoin, C., et al., *Autumnalamide, a prenylated cyclic peptide from the cyanobacterium Phormidium autumnale, acts on SH-SY5Y cells at the mitochondrial level.* J Nat Prod, 2014. **77**(10): p. 2196-205.
119. Zhang, Q., et al., *Structural investigation of ribosomally synthesized natural products by hypothetical structure enumeration and evaluation using tandem MS.* Proc Natl Acad Sci U S A, 2014. **111**(33): p. 12031-6.
120. Fujii, K., et al., *Mass Spectrometric Studies of Peptides from Cyanobacteria under FAB MS/MS Conditions.* Journal of the Mass Spectrometry Society of Japan, 2000. **48**(1): p. 56-64.

121. Rouhiainen, L., et al., *Two alternative starter modules for the non-ribosomal biosynthesis of specific anabaenopeptin variants in Anabaena (Cyanobacteria).* Chem Biol, 2010. **17**(3): p. 265-73.
122. Tan, L.T. and M.Y. Phyo, *Marine Cyanobacteria: A Source of Lead Compounds and their Clinically-Relevant Molecular Targets.* Molecules, 2020. **25**(9): p. 2197.
123. Van Wichelen, J., et al., *Strong effects of amoebae grazing on the biomass and genetic structure of a Microcystis bloom (Cyanobacteria).* Environ Microbiol, 2010. **12**(10): p. 2797-813.
124. Xinyao, L., et al., *Feeding characteristics of an amoeba (Lobosea: Naegleria) grazing upon cyanobacteria: food selection, ingestion and digestion progress.* Microb Ecol, 2006. **51**(3): p. 315-25.
125. Urrutia-Cordero, P., et al., *Effects of harmful cyanobacteria on the freshwater pathogenic free-living amoeba Acanthamoeba castellanii.* Aquat Toxicol, 2013. **130-131**: p. 9-17.

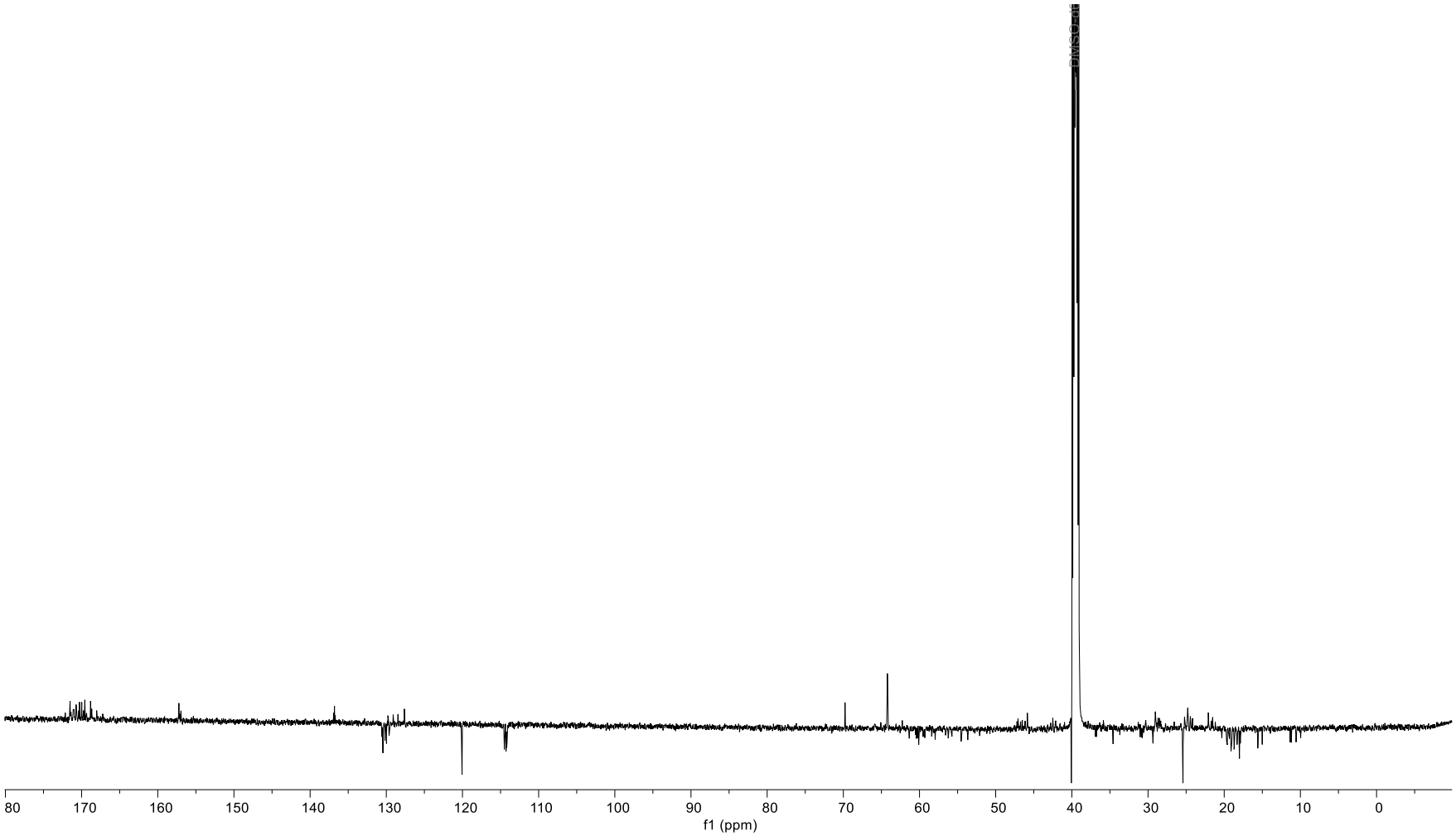
# Supplementary Information

Supplementary Table 1 Identification of the cyclo- and mono-prenylated cyanobactins on methanolic extract.

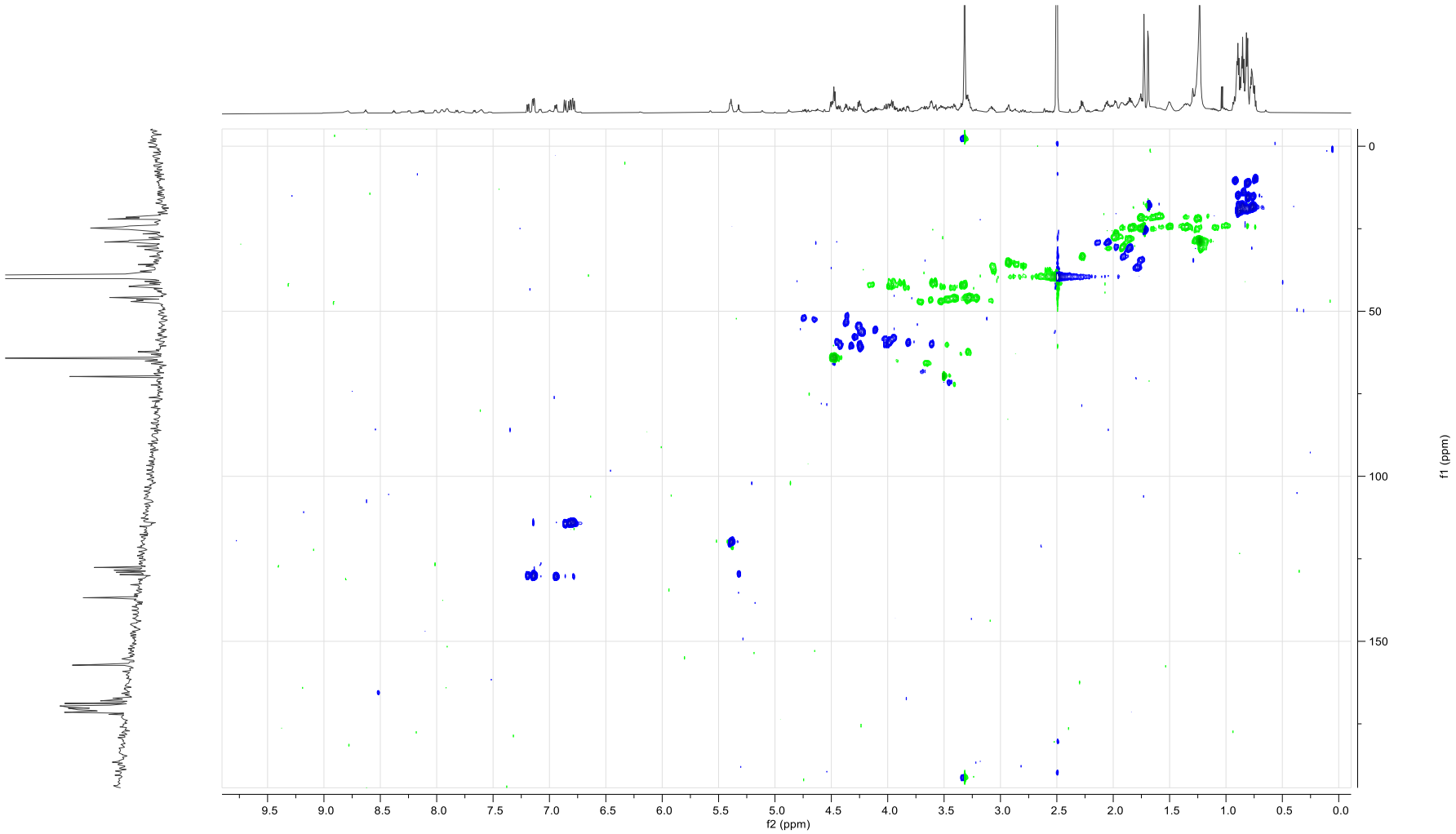
| Cyanobactins | Retention Time (min) | Calculated $[M+H]^+$ | Observed $[M+H]^+$ | Error (Δppm) |
|---|---|---|---|---|
| cyclo(OscE1) | 6.42 | 960.4825 | 960.4855 | 0.003 |
| prenyl(OscE1) | 8.82 | 1028.5451 | 1028.5459 | 0.0008 |
| cyclo(OscE2) | 5.55 | 817.3913 | 817.3904 | -0.0009 |
| prenyl(OscE2) | 7.96 | 885.4539 | 885.4553 | 0.0014 |
| cyclo(OscE3) | 7.83 | 1241.5473 | 1241.5477 | 0.0004 |
| prenyl(OscE3) | 7.84 | 1309.6099 | 1309.6130 | 0.0031 |
| cyclo(OscE4) | 6.05 | 783.4400 | 783.4413 | 0.0013 |
| prenyl(OscE4) | 8.53 | 851.5026 | 851.5030 | 0.0004 |
| cyclo(OscE5) | 6.97 | 927.4723 | 927.4745 | 0.0022 |
| prenyl(OscE5) | 9.12 | 995.5349 | 995.5360 | 0.0011 |



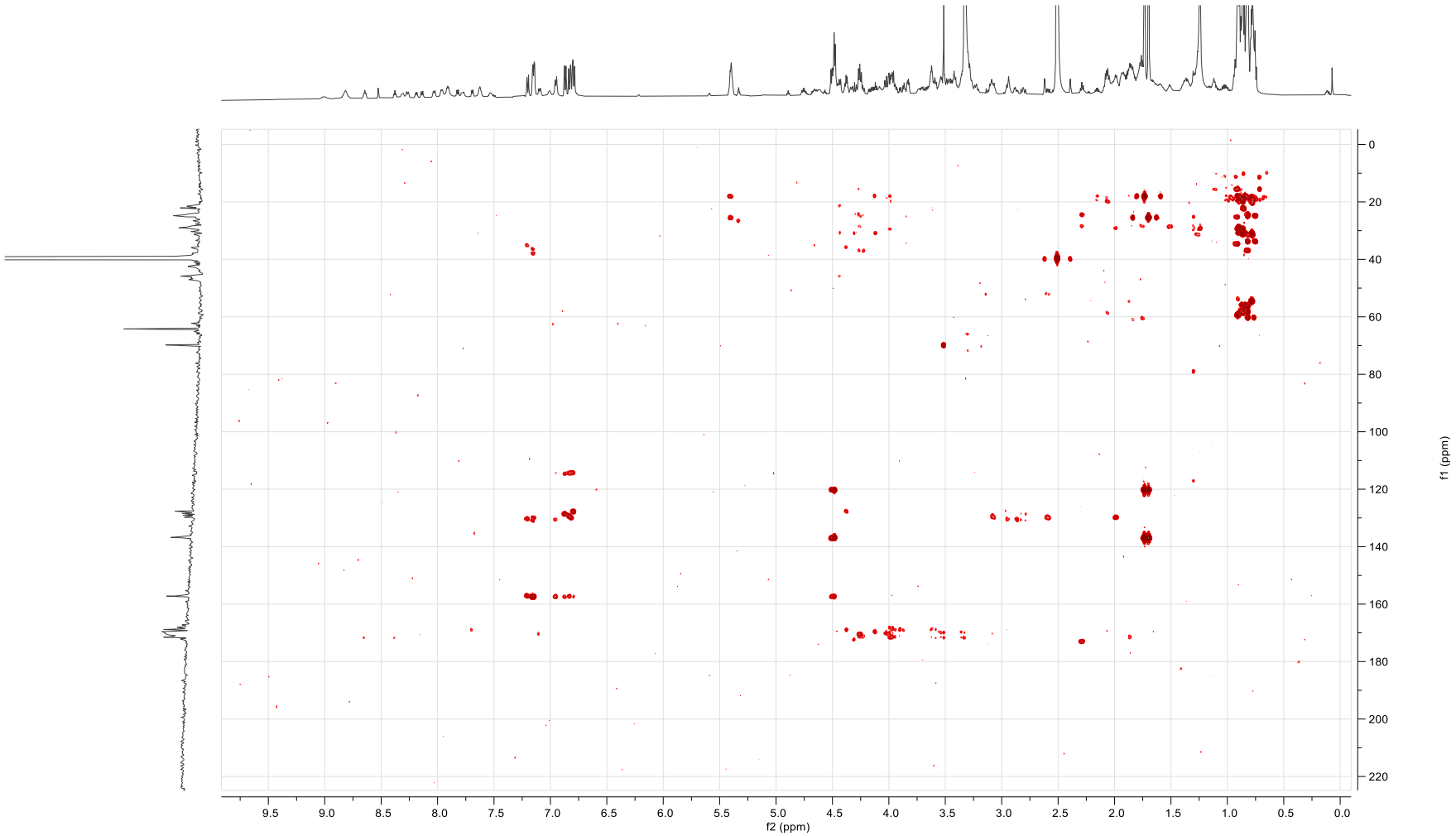

Supplementary Figure 1 $^{1}$H NMR spectrum (600 MHz, DMSO-$d_6$) of cyanobactin 850 (**1**).



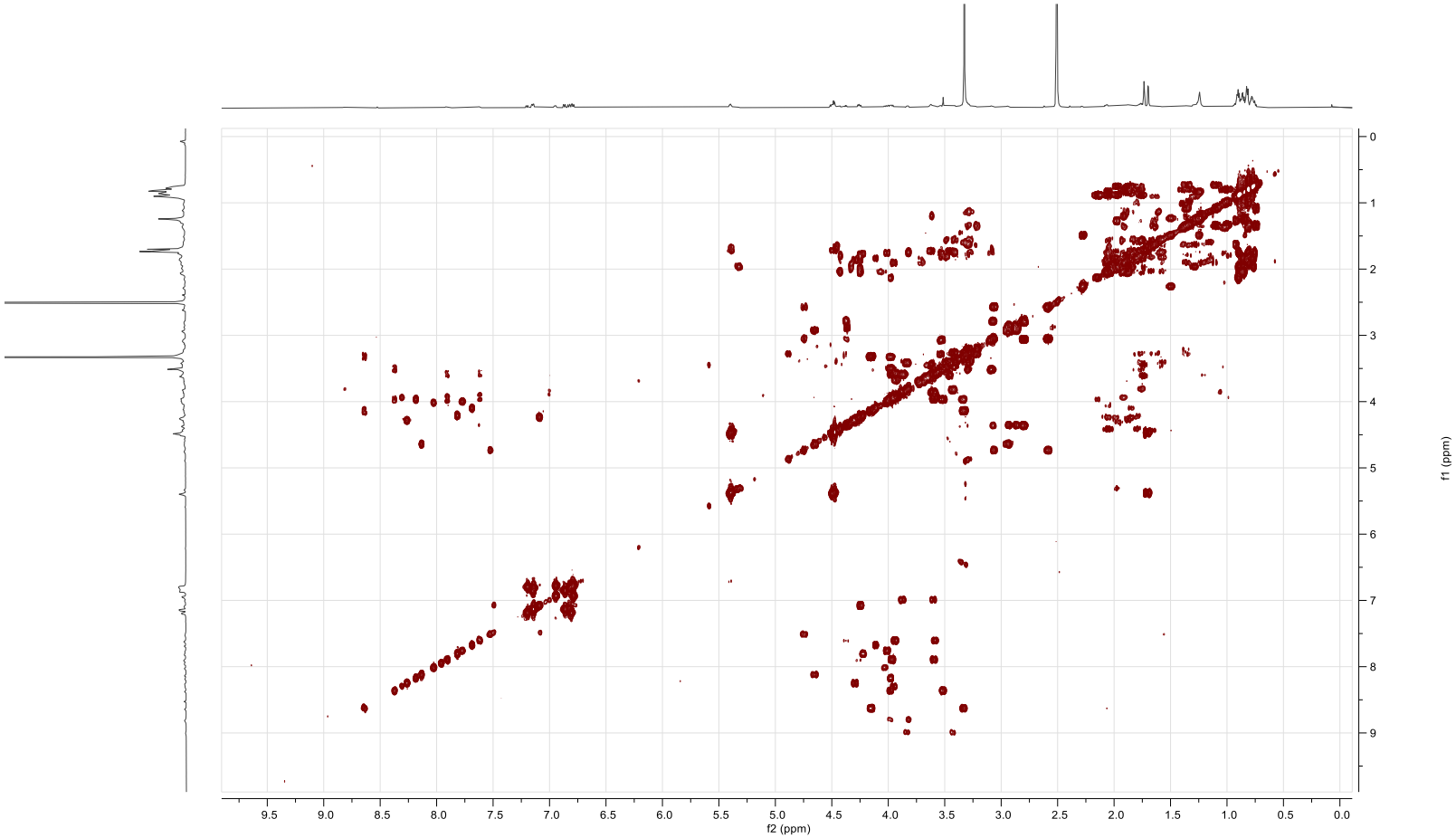

Supplementary Figure 2 $^{13}C$ NMR spectrum (150 MHz, DMSO-$d_6$) of cyanobactin 850 (**1**).




Supplementary Figure 3 HSQC spectrum (600 MHz, DMSO-$d_6$) of cyanobactin 850 (**1**).




Supplementary Figure 4 HMBC spectrum (600 MHz, DMSO-$d_6$) of cyanobactin 850 (**1**).




Supplementary Figure 5 COSY spectrum (600 MHz, DMSO-$d_6$) of cyanobactin 850 (**1**).