Environmental metabolomics characterization of modern stromatolites and annotation of ibhayipeptolides

George F. Neuhaus[1¶], Allegra T. Aron[2,3,7¶], Eric W. Isemonger[4], Daniel Petras[2,3,6], Samantha C. Waterworth[5], Luthando S. Madonsela[4], Emily C. Gentry[2,3], Xavier Siwe Noundou[4,8], Jarmo-Charles J. Kalinski[4], Alexandros Polyzois[4], Julius C. Habiyaremye[1], Margaret A. Redick[1], Jason C. Kwan[5], Rosemary A. Dorrington[4], Pieter C. Dorrestein[2,3], Kerry L. McPhail[1,*]

[1] Department of Pharmaceutical Sciences, College of Pharmacy, Oregon State University, Corvallis, Oregon 97331, United States

[2] Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, San Diego, CA, USA

[3] Collaborative Mass Spectrometry Innovation Center, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, USA

[4] Department of Biochemistry and Microbiology, Rhodes University, Makhanda, South Africa

[5] Division of Pharmaceutical Sciences, University of Wisconsin, Madison, WI, 53705

[6] Department of Biochemistry, University of California Riverside, Riverside, CA, USA

[7] Department of Chemistry and Biochemistry, University of Denver, Denver, CO, 80210, USA

[8] Department of Pharmaceutical Sciences, Sefako Makgatho Health Science University, Pretoria 0204, South Africa

* Corresponding author (kerry.mcphail@oregonstate.edu)

¶ These authors contributed equally to this work

Abstract

Lithified layers of complex microbial mats known as microbialites are ubiquitous in the fossil record, and modern forms are increasingly identified globally. A key challenge to developing an understanding of microbialite formation and environmental role is how to investigate complex and diverse communities *in situ*. We selected living, layered microbialites (stromatolites) in a peritidal environment near Schoenmakerskop, Eastern Cape, South Africa to conduct a spatial survey mapping the composition and small molecule production of the microbial communities from environmental samples. Substrate core samples were collected from nine sampling stations ranging from the upper point of the freshwater inflow to the lower marine interface where tidal overtopping takes place. Substrate cores provided material for parallel analyses of microbial community diversity by 16S rRNA gene amplicon sequencing and metabolomics using LC–MS^2. Species and metabolite diversities were correlated, and prominent specialized metabolites were targeted for preliminary characterization. A new series of cyclic hexadepsipeptides, named ibhayipeptolides, was most abundant in substrate cores of submerged microbialites. These results demonstrate the detection and identification of metabolites from mass-limited environmental samples and contribute knowledge about microbialite chemistry and biology, which facilitates future targeted studies of specialized metabolite function and biosynthesis.

Introduction

The role of metabolism on microbial community structure at local spatial scales is a central question in microbial chemical ecology [1] that is most often approached by (meta)genomic analyses of microbial metabolic capacity [2,3]. The potential for chemical communication networks to structure microbial communities is well recognized [4], although *in situ* detection of metabolites at low, highly variable abundances in complex environmental samples remains difficult. In addition to limitations in instrument technology, a lack of characterized metabolite structures in centralized databases makes reliable annotation of the hundreds to thousands of mass spectrometric (MS) features detected by untargeted liquid chromatography mass spectrometry (LC–MS) methods challenging. The coupling of untargeted metabolomics to microbial community data and metagenomics for analysis of complex microbial communities was demonstrated recently by Tuttle et al. who deployed adsorption resins for direct sampling of metabolites in marine sediments [5]. These integrated analyses have been enabled by technological advances in high resolution mass spectrometry and associated data analysis tools [6]. In particular, computational mass spectrometry tools enable predictions of metabolite structure or structure class based on tandem mass spectrometry (MS^2) fragmentation patterns, which require only nanograms of chemical extract from environmental samples. GNPS [7] is an open-access platform for comparative networking and searching of publicly shared MS^2 spectra and available spectral databases. The SIRIUS5 suite [8] incorporates a collection of software tools to perform molecular formula, structure, and compound class predictions from MS^2 spectra. Its core module SIRIUS analyzes isotope patterns and computes fragmentation trees to achieve molecular formula prediction, while the incorporated CSI:FingerID [9] uses these fragmentation trees to generate

molecular fingerprints that facilitate structure predictions. These molecular fingerprints are also used by the module CANOPUS [10–12] to predict structural classes of metabolites through a deep neural network. Such integrated approaches enable the structural annotation of microbial metabolites directly from chemically complex and scarce environmental samples [4] even when no spectral match to characterized metabolite structures in centralized databases is found.

Microbialites are organo-sedimentary deposits formed by the metabolic activity of microbial communities that induce the precipitation of minerals and trapped sediments to form structures that can be amorphous, clotted (thrombolites), or layered (stromatolites) [13]. They are ubiquitous in the fossil record, dating back to the Archaean and Precambrian eras [14–17]. Extant microbialites occur in temperate (marine and freshwater) as well as extreme environments that include hypersaline lagoons [18]. They have been documented primarily from sites in Western Australia [19], the Bahamas [20], South/Central America [21], North America [22,23], and South Africa [24,25], with highly diverse microbial community structures and metabolisms. Supratidal microbialites on rocky coasts have been described since 2003 [26], and these distinct siliciclastic environments are characterized as representative of ancient microbialite beds along Precambrian coastlines [27]. The presence of modern microbialites provides an opportunity to investigate biomineralized microbial communities as a source of specialized metabolites, which are small organic molecules produced by organisms that increase survivability and fecundity but are not essential for survival.

In studies of microbialite communities, molecular approaches have provided phylogenetic profiles, based on 16S/18S SSU rRNA genes that identified taxonomically diverse consortia, representing different functional guilds, for example, photosynthesis, sulfate reduction, sulfide oxidation, heterotrophy, carbohydrate metabolism, and carbonate accretion [28–33]. Metagenomic

studies of actively-growing microbialites have corroborated the presence of genes for the biochemical pathways expected for these functional guilds, and comparative metagenomic and metatranscriptomic analyses have been used to study the expression of genes associated with calcium carbonate accretion [34,35]. While metagenomic studies of microbialites may identify patterns of biosynthetic potential across different habitats, the context and levels of expression of metabolites and their potential ecological roles remain speculative. There are very few studies on the detection, characterization or biogenesis of small molecule metabolites produced by microbialites. Lipid profiles have been compared between microbialites and non-lithified microbial mats from the hypersaline Hamelin Pool in Shark Bay, Australia [36]. Putative assignment of cyanobacterial cyanopeptolin S and 21-bromo-oscillatoxin A by mass spectrometric analysis has been reported [37]. Production of natural microbial sunscreens, scytonemin and mycosporine-like amino acids, has also been studied in Shark Bay microbialites [38].

The first detailed report of coastal microbialites in South Africa documented the formation of tufa microbialites in upper intertidal rock pools fed by non-saline groundwater near the Kei River in the north of the Eastern Cape province [39]. Subsequently, similar occurrences of microbialites were recorded along 200 km (124 miles) of coastline in the Eastern Cape [24]. These formations are now known to be ubiquitous along the coastline of southern Africa, from Brandsebaai, South Africa, in the west to Tofo, Mozambique in the east [25]. The Schoenmakerskop tufa microbialite formation (**Fig S1**) has been the subject of a number of hydrochemical [40], microbial [24,41], and associated metazoan studies [42–44]. These studies provided evidence of new and known cyanobacterial and other taxa, some of which are related to known producers of specialized metabolites with varied biological functions. Thus, we selected the Schoenmakerskop microbialite system to demonstrate the utility of untargeted metabolomics

for analyzing environmental microbialite samples and to conduct an initial spatial survey of the study site that would inform future metagenomic deep sequence analyses.

At Schoenmakerskop, microbialites form a shallow barrage pool that is fed by diffuse freshwater input on the landward side and experiences marine tidal over-topping (flooding with seawater at high tide) at the seaward edge, leading to significant cycling in temperature and salinity [40]. Our primary goal was to assess the bacterial and small molecule composition of actively accreting microbialites by simultaneously comparing the diversity of prokaryotic species and chemistry in samples collected across the pool at established flag stations, spaced 3-5 m apart (**Fig S2**). We sought to structurally classify and identify prominent and or new specialized metabolites that may be the same or different from those reported from cyanobacteria, heterotrophic bacteria, or other microbes in non-lithified microbial mats. Chemical extraction of substrate cores (**Fig S3**) for LC–MS^2 profiling and computational analyses to compare and structurally classify MS features, were paired with DNA extraction of substrate cores for phylogenetic analyses of the microbial community using 16S rDNA sequencing. Metabolomic analyses of the variation between samples collected across the pool led to the identification of a new class of depsipeptides, named ibhayipeptolides for 'iBhayi', which is the regional name by which Algoa Bay is known. Spearman rank correlations of MS features with bacterial 16S rRNA amplicon sequences based on abundances indicated an association of the ibhayipeptolides with cyanobacterial and bacteroidete taxa.

Results and Discussion

Substrate cores obtained from collections made at two low tides on consecutive days were

categorized according to their source location, as either submerged or surface exposed. Samples assigned to the submerged category were either completely submerged in standing water or at the water interface, while samples assigned to the surface exposed category were exposed to air or flowing water. In general, cores collected from the same flag station all fit into the same category of either surface exposed or submerged (flag stations **1**, **2**, **4**, **5**, **7**, **8**, and **10**; **Figs 1A**, **1B**, and **S4-S12**). Two exceptions were flag stations **3** (**Fig S6**) and **6** (**Fig S9**), where substrate cores collected ranged from surface exposed to submerged in standing water. Based on bacterial community compositions discussed in detail below (**Fig 2**), and for clarity in presentation of this work, flag station **3** is considered to be surface exposed and flag station **6** as submerged (**Figs 1A** and **1B**).

A total of 1,567,727 reads, clustered into 12,498 OTUs, were recovered from all core samples collected across the stromatolite barrage pool (**Fig 1A**). The OTUs were classified within 613 different families, 361 orders, 150 classes and 51 bacterial phyla. At least 1,111 (8.9%) OTUs could not be classified beyond the bacterial kingdom. Approximately 30% of all OTUs were classified within Phylum Proteobacteria, and 16% classified within Phylum Bacteroidetes. The number of OTUs per phylum did not, however, correlate with the abundance of OTUs per phylum. The most abundant OTUs were classified within the Cyanobacteria phylum (33% relative abundance across all sites), with OTUs classified within the Proteobacteria and Bacteroidetes phyla accounting for 23% and 18% of total relative OTU abundance across all sites. Unclassified families in the Alphaproteobacteria, Gammaproteobacteria, and Bacteroidia, as well as the Omnitrophicaeota and Saprospiraceae families, accounted for 15% of all OTUs, with each consisting of approximately 3% of all OTUs. All other families accounted for 2% or less of all OTUs.

LC–MS^2 raw data for all substrate cores was processed with MZMine2. Subsequent blank subtraction and removal of MS features found solely in the washout phase of the LC gradient resulted in a total of 15,752 MS features and 5,997 pairs of different ion species (e.g., $[M+H]^+$, $[M+Na]^+$) of the same molecule being identified, based on the correlation of chromatographic peak shapes for MS features [45]. Feature based molecular networking (FBMN) using GNPS resulted in 15,752 nodes with 25,894 edges (21,165 MS^2 cosine scores greater than 0.7 and 4,729 MS^1 annotations). Of the total number of nodes, 6,005 shared an edge with another node, leaving 9,747 singletons. Only 487 nodes (3.1%) were annotated with spectral similarities and precursor *m/z* matches to metabolites in GNPS spectral libraries. MS features were assigned to 1,287 different structural classes using CANOPUS.

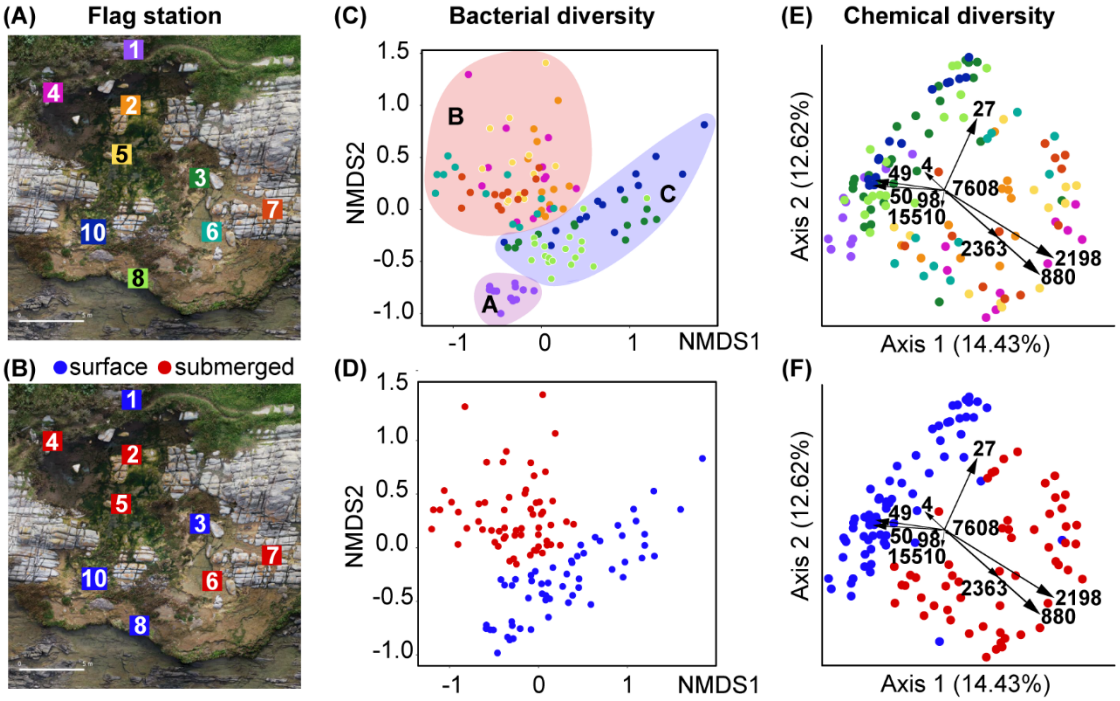


Fig 1. Schoenmakerskop microbialite collection site and component analyses. **(A)** Geographic information systems (GIS) image of the Schoenmakerskop barrage pool with superimposed numbered squares indicating the position of flagged collection stations (**1-8** and **10**) across the

pool. At each of the nine flag stations, 12 to 18 substrate sample cores were collected (**Figs S4-S12**). **(B)** GIS image of the Schoenmakerskop barrage pool displaying flag stations colored according to whether submerged or exposed at the surface. **(C)** NMDS plot of microbial 16S rDNA OTUs for all substrate cores colored by flag station, highlighting three visually apparent groupings, A: flag station **1** (purple, freshwater inflow); B: flag stations **2** and **4-7** (red, submerged); C: flag stations **3**, **8**, and **10** (blue, surface). (**D)** NMDS plot (ANOSIM, R 0.47, p-value 0.0001, Bray-Curtis) indicating a significant difference between surface (blue) and submerged (red) flag stations; **(E)** Bray–Curtis PCoA biplot of the LC–MS data for all substrate cores, represented by filled circles colored by source flag station. All flag sites were significantly different from one another (PERMANOVA, pseudo-F 5.68, p 0.001, Bray-Curtis distance) **(F)** Bray–Curtis PCoA biplot of the LC–MS data for all substrate cores colored by exposure to surface water or submerged. A greater degree of separation is observed between submerged and flowing flag sites (PERMANOVA, pseudo-F 14.64, p-value 0.001, Bray-Curtis distance). Vectors indicate driver MS features associated with Bray–Curtis distances between samples and are described in Table 1.

Bacterial diversity in microbialite cores across flag stations

Initially, we mapped the taxonomic diversity of the microbial communities to assess variability across the sampling stations, focusing on bacteria, which have been shown in a previous shotgun metagenomic study to comprise > 98% of the total microbial community [41]. We analyzed 16S rRNA gene amplicon libraries, acquired in parallel with LC–MS^2 data, for samples at each flag station (**Figs 1C and 1D**). An assessment of beta diversity in the bacterial community by NMDS and ANOSIM (both Bray–Curtis distance) revealed that all flag stations (**Fig 1C**) were significantly different from one another ($p < 0.05$, **Table S1**). However, the level of dissimilarity,

as indicated by the R statistic, suggested separate grouping of A) flag station **1** samples from the freshwater inflow as unique and unlike all other sites, B) samples from flag stations **3**, **8** and **10**, which were all described as exposed to flowing water and shared low R values, and C) samples from generally submerged flag stations **4**, **5**, **6**, and **7**, with similarly low R values, and flag station **2** samples, which shared some similarity with those from **4** and **5** (see **Table S1**). A canonical-correlation analysis (CCA, **Fig S13**) of the bacterial OTUs showed that OTUs classified within the Nostocales class of Cyanobacteria were prominent members of the surface group of taxa, responsible for the differences observed between surface (blue) and submerged (red) communities (PERMANOVA, pseudo-F = 15.98, p = 0.001, Bray-Curtis distance). Phylogenetic classification of the 16S rRNA gene OTUs showed that almost all the samples were dominated by Cyanobacteria, followed by taxa classified within the Bacteroidetes and Proteobacteria. Vericobacteria and Chloroflexi OTUs were also abundant (**Fig S14**). The bacterial communities were dominated by 20 OTUs, which accounted for between 50 to 80 % of all reads across the sampling stations (**Fig S15**); nine of these OTUs were conserved in all samples (**Fig 2**). Closer inspection revealed that the relative abundances of two conserved cyanobacterial OTUs classified as Nostocales (OTU1 and OTU5) were significantly more abundant in samples from the flag stations exposed to flowing water (**Fig 1B**: flags **1**, **3**, **8** and **10**; **Fig 2**). OTU1 shared the greatest nucleotide sequence identity with *Calothrix* species, and OTU5 was most closely related to mat-forming *Microcoleus* and *Phormidium* species (**Fig 3**). OTU5 is the only OTU that corresponds to a conserved OTU identified in a previous study of Schoenmakerskop and Cape Recife [41], suggesting that this species may be important for peritidal microbialite growth. Furthermore, *Microcoleus* species are reported to withstand high salinity [46], an attribute that may help these bacteria survive the periodic increases in salinity due to tidal over-topping of the stromatolite

barrage pool. OTU2 and OTU3 are conserved across all stations and are dominant in samples from the submerged sites (flags **2**, **4** - **7**; **Figs 1B** and **2**). OTU2 and OTU3 are closely related, and share greatest sequence identity with *Plectonema*, *Rivularia* and *Calothrix* species (**Fig 3**), which are all genera found in abundance in other microbialite formations that occur primarily in standing water bodies [47–51]. Additional data, such as metagenomic-assembled genomes, and associated transcriptomic data would be required to determine the effect that water flow rate has on the composition and behaviors of the mosaic of stromatolite-forming bacterial communities.

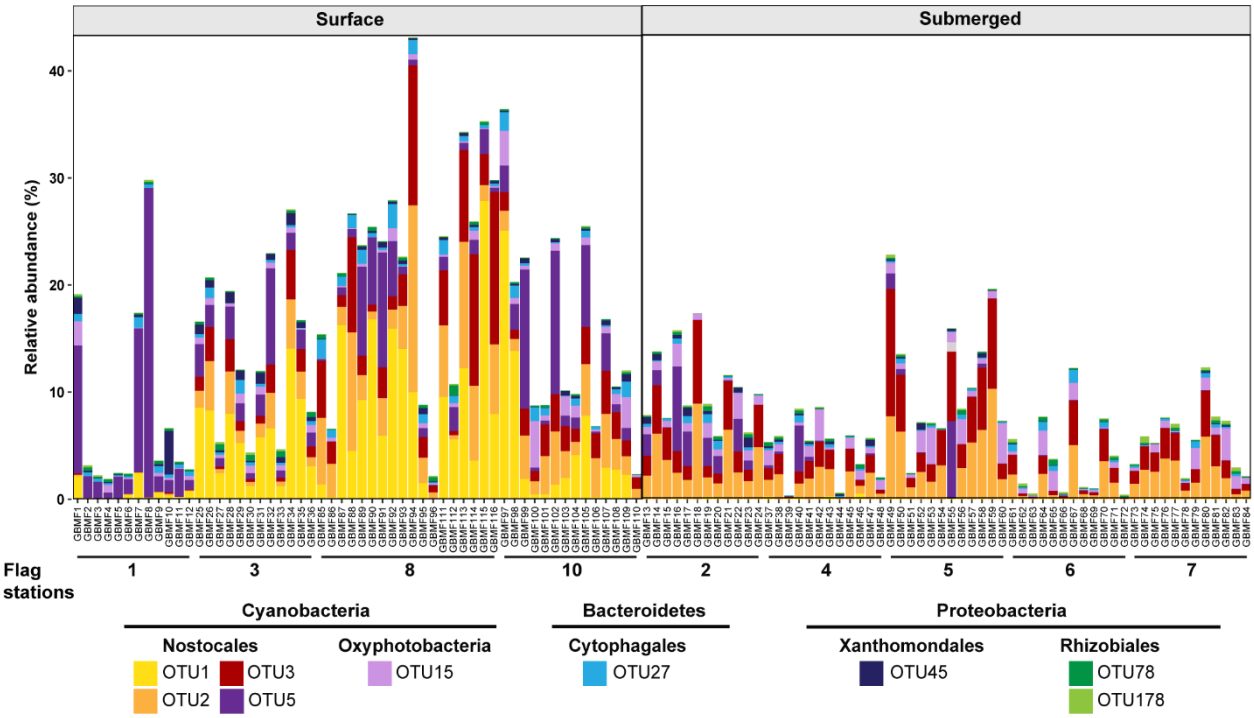


Fig 2. Analysis of 16S rRNA gene sequence amplicon data revealed the presence of nine conserved bacterial species across all microbialite core samples. Cores were grouped as either surface exposed or submerged, based on their source flag station, and these two groupings displayed distinct bacterial communities. In particular, cyanobacterial OTU1 characterized flag stations exposed to flowing water, and was largely absent from submerged flag stations.

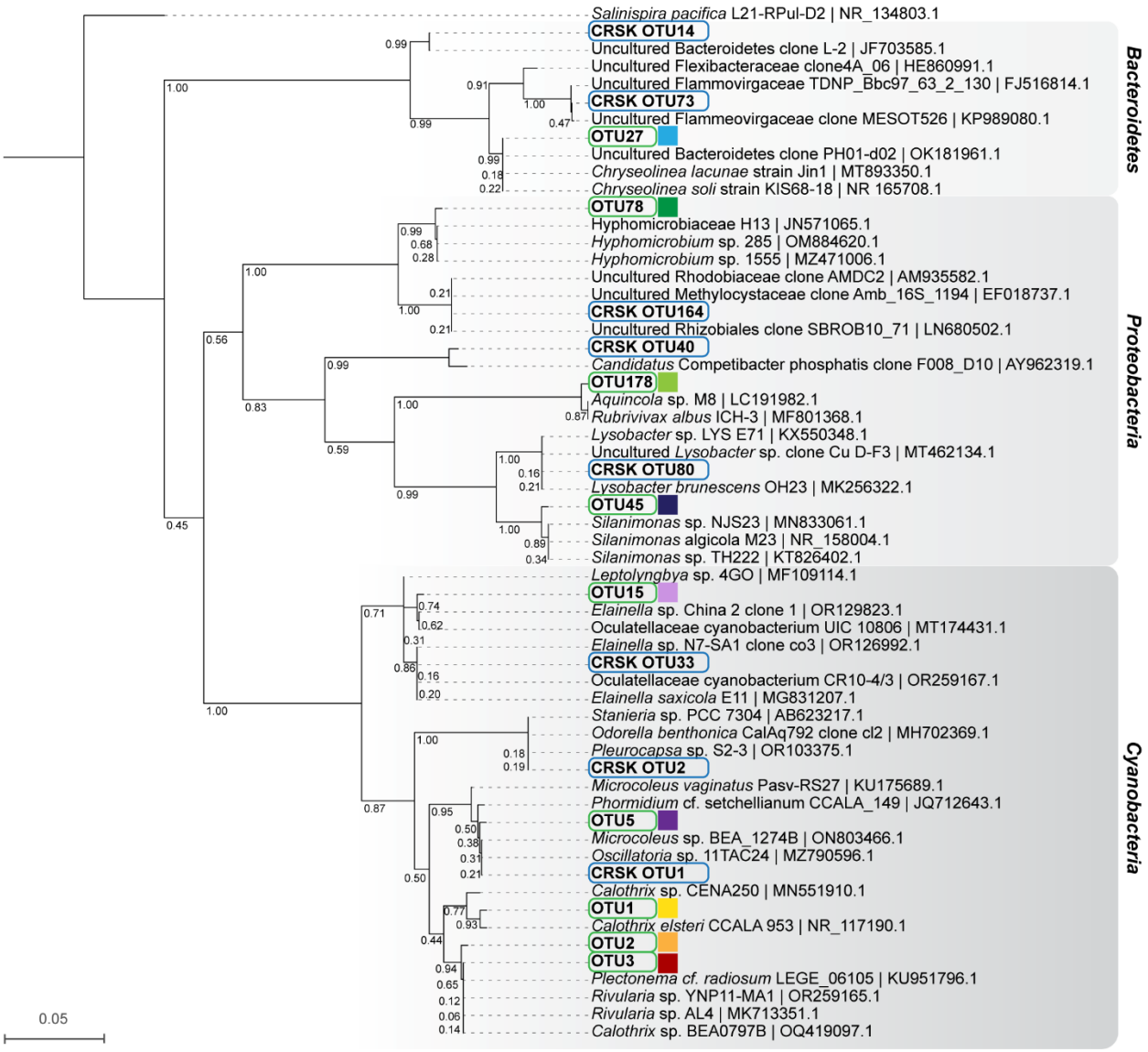


Fig 3. Phylogeny of operational taxonomic units (OTUs) conserved across all flag sites (highlighted with a green outline) relative to conserved OTUs from a previous study of bacterial communities in stromatolite formations from Schoenmakerskop and Cape Recife (highlighted with a blue outline). Colored blocks next to the conserved OTUs correlate with colors used in **Fig 2** for easy reference. Additional reference sequences from the NCBI nr database have been included for approximate taxonomic classification. Phylogeny was inferred using the Maximum-likelihood method with 1000 bootstrap replicates. A spirochete was chosen as an

outgroup. Bootstrap support is indicated on nodes and a scale has been provided for branch length, where the unit is the number of nucleotide substitutions per site.

Metabolite diversity of microbialite cores across flag stations

Another goal was to broadly characterize and compare the small molecule composition of substrate cores across the nine flag stations of the barrage pool. A molecular formula and prediction of compound structural class were obtained, using SIRIUS4 [8] and CANOPUS [11], respectively, for each MS feature detected in the pre-processed LC–MS^2 data for all core samples. The *in silico* annotation tool CANOPUS utilizes deep neural network prediction based on fragmentation spectra to assign each detected MS feature to a superclass, class, and subclass. A preliminary assessment of differences between flagged stations across the barrage pool based on the CANOPUS output indicated that flag station **1** (landward freshwater inflow) and flag station **7** (mid-pool) displayed marked differences at the superclass and class level in metabolites present at each sampling station (**Fig S16**).

Using Principal Coordinate Analysis (PCoA, Bray–Curtis distances) to assess the separation of MS features by individual flag site revealed that all flag stations (**Fig 1E**) were significantly different from one another (PERMANOVA, pseudo-F = 5.68, p = 0.001, Bray–Curtis distance). However, a greater degree of separation was observed between stations that were submerged in water versus those that were surface exposed, over which water was flowing (**Fig 1F**). Statistical significance of the separation between submerged and exposed stations was confirmed by PERMANOVA analysis, which revealed a pseudo-F score of 14.64 (p-value = 0.001, Bray–Curtis distance), in agreement with the bacterial community data.

To understand the factors distinguishing submerged and surface exposed flag stations, we inspected the loading vectors in **Fig 1F** (**Table 1**). Discrete MS features can be visualized in the molecular network as nodes colored by submerged versus surface exposed metadata (**Figs 4**, **S17**). Molecular families are observed as subnetworks of nodes connected by edges that represent MS^2 spectra with cosine similarity scores greater than 0.7 (solid light gray lines) or MS^1 annotations of different ion species of the same molecule (dashed dark gray lines). Ion species relationships are based on retention time matches and Pearson correlations of greater than 85% for chromatographic peak shape. Only subnetworks containing MS features responsible for separation of samples between surface and submerged stations are shown (**Fig 4**). Analysis of the MS^2 spectra for ten top metabolite vectors with SIRIUS5 and CANOPUS (**Table 1**) predicted they belong to seven different chemical classes. Diradylglycerols, thiazole-containing macrolides, and depsipeptides were found primarily in the submerged cores, while fatty alcohols were present primarily in the surface exposed cores (**Figs 4** and **S17**). Notably, the CANOPUS output for MS feature 7608 predicted epothilone (**Table 1**), although the large *m/z* (734.4649) and SIRIUS molecular formula prediction suggest that MS features in this subnetwork are larger thiazole-containing macrolides (**Figs 4B and S17B**). Further structure elucidation is precluded by low abundance and thus poor quality MS^2 spectra (**Fig S18**).

We found that key MS features responsible for this separation, including MS feature IDs 2198, 880, and 2363, belong to the same family of depsipeptides (**Fig 4C**). Notably, a DEREPLICATOR [52] search matched MS feature ID 2198 with the depsipeptide serrawetin, for which the low score of 7 indicated a match in structural class only, further supporting the prediction by SIRIUS and CANOPUS (**Table 1**). Manual inspection of MS^2 spectra for these MS features confirmed a peptidic molecular family and suggested they represent a new family of cyclic

depsipeptides. Therefore, we targeted these prominent MS features for further structural characterization, assigning them the trivial name of ibhayipeptolides.

Table 1. Prominent MS features prioritized as PCoA vectors responsible for Bray-Curtis dissimilarity between submerged and surface (exposed) samples.

Ion cluster ID	MS Feature *m/z*, adduct	Subclass proposed by CANOPUS	Proposed molecular formula (SIRIUS5)
4	565.4028, $[M+H]^+$	Tetraterpenoid	$C_{40}H_{52}O_2$
27	535.2704, $[M+H]^+$	Carboxylic acid Derivative	$C_{32}H_{38}O_7$
49	539.4642, $[M+Na]^+$	Fatty alcohol	$C_{31}H_{64}O_5$
50	517.4825, $[M+H]^+$		
98	183.1495, $[M+H]^+$	Fatty amide	$C_{10}H_{18}N_2O$
880	799.4619, $[M+Na]^+$	Depsipeptide	$C_{44}H_{64}N_4O_8$
2198	777.4806, $[M+H]^+$		
2363	785.4488, $[M+Na]^+$	Depsipeptide	$C_{43}H_{62}N_4O_8$
7608	734.4649, $[M+H]^+$	Epothilone	$C_{41}H_{67}NO_8S$
15510	549.4484, $[M+Na]^+$	Diradylglycerol	$C_{32}H_{62}O_5$

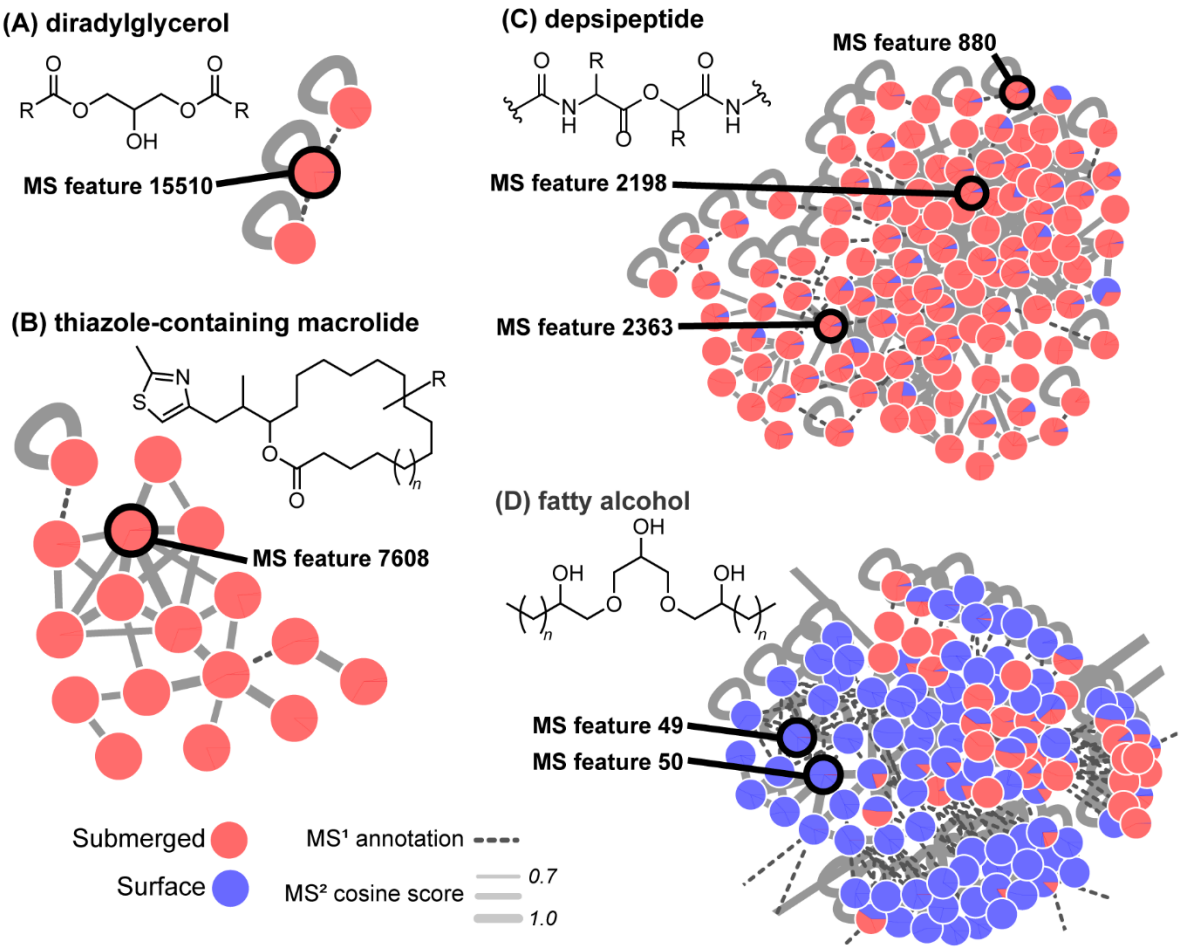


Fig 4. GNPS subnetworks containing MS features from Table 1 (PCoA biplot vectors). Feature-based and ion identity molecular networking in GNPS were used to create a global molecular network illustrating the chemical diversity in Schoenmakerskop microbialite substrate core samples. Nodes represent individual MS features and are connected by edges representing MS^2 spectra with cosine similarity scores greater than 0.7 (solid light gray lines) or MS^1 annotations of different ion species of the same molecule based on retention time and Pearson correlations of greater than 85% for chromatographic peak shape (dashed dark gray lines). Node color indicates contribution to an MS feature from either a surface exposed (blue) or a submerged (red) substrate core. Molecular family subnetworks that contain MS features responsible for

separation of surface and submerged samples in the PCoA (**Fig 1F**) are displayed, and the potential drivers listed in **Table 1** are circled and annotated in black with their MS feature ID. These MS features are predicted by CANOPUS to be, **(A)** diradylglycerols, **(B)** thiazole-containing macrolides, **(C)** depsipeptides, and **(D)** fatty alcohols, with representative chemical motifs as shown.

Spatial distribution of depsipeptides across the barrage pool

Preliminary assessment of the chemical class differences between flag stations using CANOPUS indicated that stations **4**, **5**, and **7** were predicted to contain the highest abundance of depsipeptides, although this structural class was still present at flag stations **1** and **10**. CANOPUS plots of flag stations **1** and **7** are provided as representative examples in **Fig S16**. GNPS feature-based molecular networking was used to confirm this CANOPUS prediction, and to visualize the number and extent of depsipeptide MS features across the flag stations (**Fig 5**). The molecular network revealed that MS features 2198 and 2363 are most abundant at flag stations **4** (pink), **5** (yellow), and **7** (dark orange). This trend in the presence of depsipeptides across the barrage pool may correspond with variation in the bacterial community.

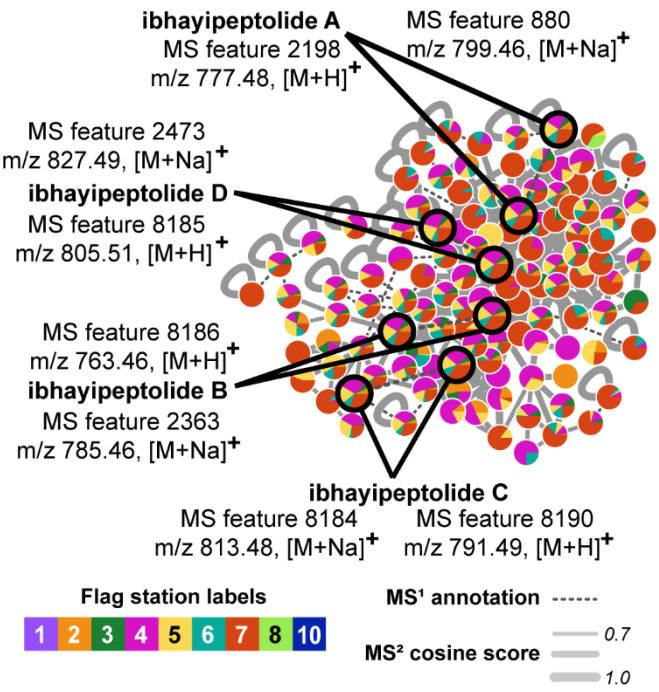


Fig 5. Distribution of depsipeptide MS features across barrage pool flag stations. GNPS feature-based molecular subnetwork for the ibhayipeptolide molecular family, incorporating ion identity networking of related mass adducts. MS features are represented as nodes labeled with MS^1. Nodes are connected by edges that represent MS^2 spectra with cosine similarity scores greater than 0.7 (solid light gray lines) or MS^1 annotations of different ion species of the same molecule based on retention time and Pearson correlations of greater than 85% for chromatographic peak shape (dashed dark gray lines). Node contributions to the ibhayipeptolide family of depsipeptides are colored according to the source flag station of the parent core sample. MS features representing the major ibhayipeptolides (A-D) are circled in black, enlarged for clarity, and annotated with MS feature ID, *m/z* value, and assigned ion species.

Planar structure elucidation of the ibhayipeptolides

Ibhayipeptolides A (networked MS features 2198 $[M+H]^+$ and 880 $[M+Na]^+$) and B (MS features 8186 $[M+H]^+$ and 2363 $[M+Na]^+$) were targeted for additional structure elucidation and potential chromatographic isolation since they were prominent PCoA vectors representing metabolic differences between flag stations and could not be assigned as known compounds from their LC–MS^2 data. Elucidation of their planar structures was assisted by comparison with the MS^2 data for two additional congeners designated as ibhayipeptolides C (MS features 8190 $[M+H]^+$ and 8184 $[M+Na]^+$) and D (MS features 8185 $[M+H]^+$ and 2473 $[M+Na]^+$). We anticipated that knowledge of the ibhayipeptolide structures could be used to guide future investigations of the ecological role of these prominent microbialite metabolites. For ibhayipeptolide A (MS feature 2198), SIRIUS5 analysis of the molecular ion isotope (**Fig S19**) and MS^2 fragmentation patterns (**Figs S20** and **S21**) of both positive ($[M+H]^+$ and $[M+Na]^+$) and negative ionization ($[M-H]^-$) spectra provided a molecular formula of $C_{44}H_{64}N_4O_8$ with 15 degrees of unsaturation. The identity and sequence of the six residues in this predicted depsipeptide (**Fig 6A**) were determined by manual analysis of the MS^2 spectrum for the $[M+Na]^+$ ion (MS feature 880). Two major MS fragmentation pathways were described for ibhayipeptolide A, one initiated by b-type cleavage of an ester and the other initiated by a-type ring cleavage of a α-C/C=O bond (**Fig 6A**). In pathway 1, ring-opening lactone cleavage of the sodiated ion is followed by sequential loss of amino and hydroxy acid residues beginning with Phe, then Lxx (representing Leu, Ile or MeVal), an α-hydroxybutyric acid (Hba) or equivalent moiety ($C_4H_6O_2$), and a second Phe, yielding a sodiated b_2 ion (*m/z* 306.2035, $C_{16}H_{29}NO_3Na^+$). In the congruent fragmentation pathway 2, an a-type sodiated ion (*m/z* 799.4562) results from ring cleavage at the α-C/C=O bond of a putative long chain α-hydroxy acid ($C_{10}H_{18}O_2$). A neutral loss of *m/z* 142.1345 is explained by subsequent b-type cleavage and loss of hydroxydecanoic acid (Hda) from the linearized ion. Subsequent

sequential loss of amino and hydroxy acid units culminates in a residual sodiated x_1 ion (*m/z* 164.0677, $C_7H_{11}NO_2Na^+$) assigned as Lxx–CO. Further support for these fragmentation pathways was provided by the computed fragmentation tree in SIRIUS, for which each ion in a given pathway is found in a unique branch (**Fig S23**). Together these analyses support the sequence of Phe–Lxx–Hba–Phe–Lxx–Hda.

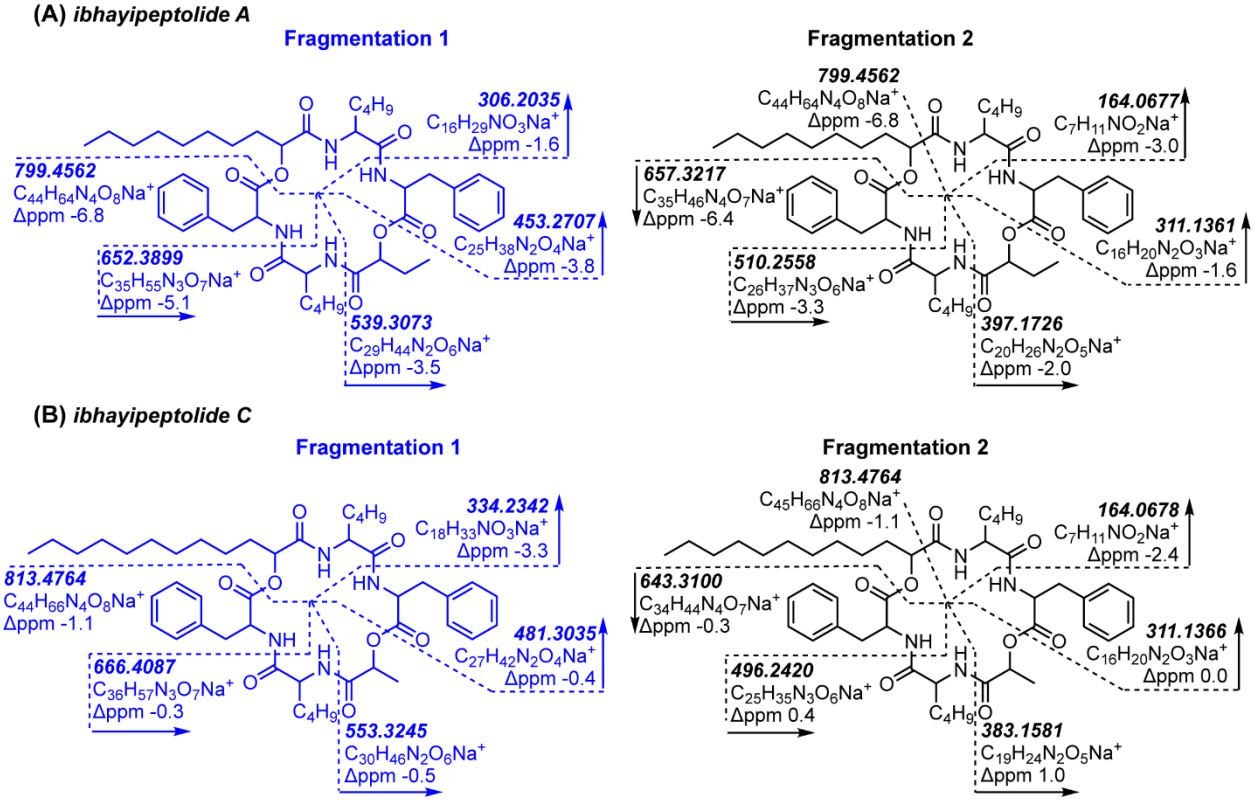
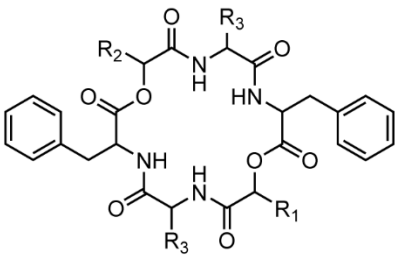


Fig 6. Assignment of ibhayipeptolide planar structures from high resolution MS^2 data. (A) Proposed MS^2 fragmentation pathways for ibhayipeptolide A (MS feature 2198) labeled with calculated mass and ppm error of fragment ions. **(B)** Proposed MS^2 fragmentation pathways for ibhayipeptolide C (MS feature 8190) labeled with calculated mass and ppm error of fragment ions.

Analysis of the isotope pattern and MS^2 fragmentation of $[M+H]^+$ and $[M+Na]^+$ with SIRIUS5 provided molecular formulas of $C_{43}H_{62}N_4O_8$ for ibhayipeptolide B (MS feature 8186),

$C_{45}H_{66}N_4O_8$ for ibhayipeptolide C (MS feature 8190), and $C_{46}H_{68}N_4O_8$ for ibhayipeptolide D (MS feature 8185), respectively. Sufficient quantities of ibhayipeptolides could not be isolated for NMR analysis, yet nearly identical fragmentation patterns in the MS^2 spectra for sodiated ions provided consistent planar structure assignments for all four congeners (**Figs S22**, **S31**, **S38**, and **S43**). Ibhayipeptolide B (MS feature 8186) with *m/z* 785.4438 ($[M+Na]^+$) differed from ibhayipeptolide A only in the loss of 72.0208 Da between the b_4 and b_3 ions, instead of 86.0365 Da, in pathway 1, indicating the presence of a lactic acid residue in ibhayipeptolide B in place of the Hba in ibhayipeptolide A (**Fig S31**). Ibhayipeptolide C (MS feature 8184, *m/z* 813.4764 ($[M+Na]^+$), also showed two major fragmentation pathways (**Fig 6B**). Fragmentation pathway 1 again indicates the presence of a lactic acid residue (b_4 - b_3 = 72.0215 Da), as for ibhayipeptolide B. In fragmentation pathway 2 for ibhayipeptolide C, a loss of 170.1664 Da follows the initial a-type cleavage, instead of 142.1345 Da for ibhayipeptolide A, indicating substitution of a hydroxydodecanoic acid (Hdda) for the Hda moiety in ibhayipeptolides A and B. An Hba residue was evident (pathway 1, b_4 - b_3 = 86.0365 Da) in the spectra for ibhayipeptolide D (MS feature 8185) with *m/z* 827.4910 ($[M+Na]^+$), while an initial loss of 170.1658 Da in fragmentation pathway 2 was consistent with the presence of Hdda (**Fig S43**). Together these data indicate that all four peptides share the same sequence and have α-hydroxy acids with varying sidechains (**Fig 7**).



ibhayipeptolide A: $R_1 = CH_2CH_3$, $R_2 = C_8H_{17}$, $R_3 = C_4H_9$ (one Leu, one Ile)
ibhayipeptolide B: $R_1 = CH_3$, $R_2 = C_8H_{17}$, $R_3 = C_4H_9$ (Leu or Ile)
ibhayipeptolide C: $R_1 = CH_3$, $R_2 = C_{10}H_{21}$, $R_3 = C_4H_9$ (Leu or Ile)
ibhayipeptolide D: $R_1 = CH_2CH_3$, $R_2 = C_{10}H_{21}$, $R_3 = C_4H_9$ (Leu or Ile)

Fig 7. Molecular structure summary for ibhayipeptolides A-D. The ibhayipeptolide planar structures contain either α-hydroxybutyric acid (Hba, A and D) or lactic acid (B and C), and α-hydroxydecanoic acid (Hda, A and B) or α-hydroxydodecanoic acid (Hdda, C and D) residues as determined by analysis of MS^2 fragmentations. Limited NMR data indicates the presence of both Leu and Ile residues in ibhayipeptolide A (**Fig S27**), whereas these are unassigned in B-D because they are not distinguishable from MS^2 data.

The ^1H NMR spectrum (**Fig S24**) for ibhayipeptolide A (MS feature 2198) displayed two NH (δ_H 7.45, 7.48), ten aromatic (δ_H 7.1-7.3), two oxymethine (δ_H 5.15, 5.34), four amino acid α-H resonances (δ_H 3.6-4.6), and a congested aliphatic region (δ_H 0.6-2.4). A high quality, comprehensive NMR data set could not be obtained for ibhayipeptolide A due to a paucity of material isolated (~50 μg). Nevertheless, correlations observed in COSY (**Fig S25**) and HSQC (**Fig S26**) NMR experiments supported the identity of the six residues (**Fig S27**, **Table S2**). Spin systems for two Phe residues and an Hba, incorporating a deshielded α-H signal (δ_H 5.15), were present in the COSY spectrum. A Leu residue was delineated by COSY correlations between two CH_3 doublets (δ_H 0.89, 0.93) and a γ-H multiplet (δ_H 1.58), along with a spin system including an NH (δ_H 7.45), an α-H (δ_H 4.19), and diastereotopic β-CH_2 (δ_H 1.73, 1.82) signals. The presence of an additional oxymethine moiety (δ_H 5.34, δ_C 73.7) was consistent with a second hydroxy acid, as

proposed from MS^2 data ($C_{10}H_{18}O_2$). This was further supported by two incomplete COSY spin systems suggesting a terminal methyl and α-hydroxy substituent. Finally, despite an overlapped upfield region in the ^1H NMR spectrum, correlations supporting an Ile residue could be delineated. COSY spin systems incorporating a δ-CH_3 triplet (δ_H 0.84, δ_C 10.9) correlated to diastereotopic γ-CH_2 resonances (δ_H 1.29, 1.01, δ_C 26.1), and COSY-correlated NH (δ_H 7.48) and α-H (δ_H 3.67) signals were both observed. While both Leu and Ile are present, their relative location within the ring could not be determined without a high quality HMBC spectrum.

Relaxed substrate specificity of the NRPS adenylation domains [53–55] may be responsible for structural microheterogeneity (**Fig 7**) and co-elution of closely related metabolites. Microheterogeneity is supported by the large number of isobaric MS features in the molecular network: specifically, there are seven nodes for *m/z* 777.48 ± 0.03 with retention times between 7.7 and 8.0 min.

Correlation of select metabolites with microbialite bacteria.

Given the variation in microbial community and metabolites between flag stations in the Schoenmakerskop barrage pool, we calculated (nonparametric) Spearman rank correlations between abundance of bacterial 16S rRNA OTUs and MS features (**Fig 8**) to identify groups of organisms associated with specific metabolites, whether positively or negatively correlated. These correlations in abundance may indicate producer or utilizer bacteria of specific metabolite families. It is also noteworthy that a fungal producer or utilizer of the targeted metabolites cannot be discounted in the absence of fungal ITS sequencing.

The ibhayipeptolide molecular family (e.g., MS features 2198, 880, and 2363) were significantly positively correlated with *Geitlerinema*_PCC-9228 (Cyanobacteria, OTU290),

unknown Oxyphotobacteria_Incertae_Sedis (Cyanobacteria, OTU42), Cryomorphaceae (Bacteroidota, OTU302) and Chitinophagales (Bacteroidota, OTU25), while they were negatively correlated with *Phormidesmis*_ANT.L52.6 and Chamaesiphon_PCC-7430 (Cyanobacteria, OTUs 56 & 62 respectively), Anaerolineae A4b (Chloroflexota, OTU538), and Flavobacteriaceae (Bacteroidota, OTU92). It is noteworthy that none of these bacterial OTUs were dominant community members that could be expected to influence gross stromatolite morphology and there was no distinct stromatolite morphology associated with the highest abundance of depsipeptolides at flag stations 4-7. Interestingly, the fatty alcohol molecular family (e.g., MS features 49 and 50) also displayed significant but inverse correlations with some of the same organisms (**Fig 8**).

Cyanobacteria, heterotrophic bacteria, and fungi are all known to produce nonribosomal depsipeptides, comprising diverse amino and hydroxy acid residues, that exhibit a variety of biological properties. Many of these metabolites with alternating peptide and ester bonds comprise repeating depsipeptide oligomers. Cyclic hexadepsipeptides from cyanobacteria include antanapeptins [56], which like ibhayipeptolides consist of one short and one longer chain hydroxy acids, separated by a dipeptide. Unlike ibhayipeptolides, the antanapeptins and a variety of other cyanobacterial depsipeptides contain branched β-hydroxy acids [56–60]. There are many examples of branched longer chain hydroxy acids in the extensive set of known cyanobacterial cyclic depsipeptides and it is notable that the linear long chain α-hydroxy acid in the ibhayipeptolide structural motif appears unusual for a cyanobacterial metabolite [61,62]. Cyclic hexadepsipeptides containing long chain hydroxy acids from heterotrophic bacteria include the icosalides produced by *Burkholderia* species from diverse habitats [63]. A relevant group of pentadepsipeptides derived from both cyanobacteria and heterotrophic bacteria constitutes the unnarmicins [64,65] and solonamides [66], which comprise a general structural motif of one β-hydroxy acid cyclized

with a tetrapeptide sequence of alternately repeating amino acids. Derived from a marine *Photobacterium* (Gammaproteobacteria), unnarmicins A and C possess β-hydroxy-octanoyl and -hexanoyl units, respectively, preceding a Leu-Phe-Leu-Phe sequence. In contrast, unnarmicin D was isolated from a marine *Trichodesmium* cyanobacterium and comprises a β-hydroxy-dodecanoyl moiety cyclized with a Gly-Tyr-Gly-Phe tetrapeptide. Solonamides A and B were also isolated from a marine *Photobacterium* and possess β-hydroxy-hexanoyl and -octanoyl units, respectively, preceding a Phe-Leu-Ala-Leu sequence. The solonamides have been investigated as biofilm inhibitors of *Staphylococcus aureus* given their inhibition of the accessory gene regulator (agr) quorum sensing system [66–69]. Cyclic depsipeptides reported from fungi mostly contain one or multiple relatively short and branched chain α-hydroxy acids, and hexadepsipeptides are the largest subgroup of this class [70,71]. A notable octadepsipeptide exception is verticilide, which is a cyclotetramer of the dipeptidol 2*R*-hydroxyheptanoic acid–N-methyl-L-Ala [72].

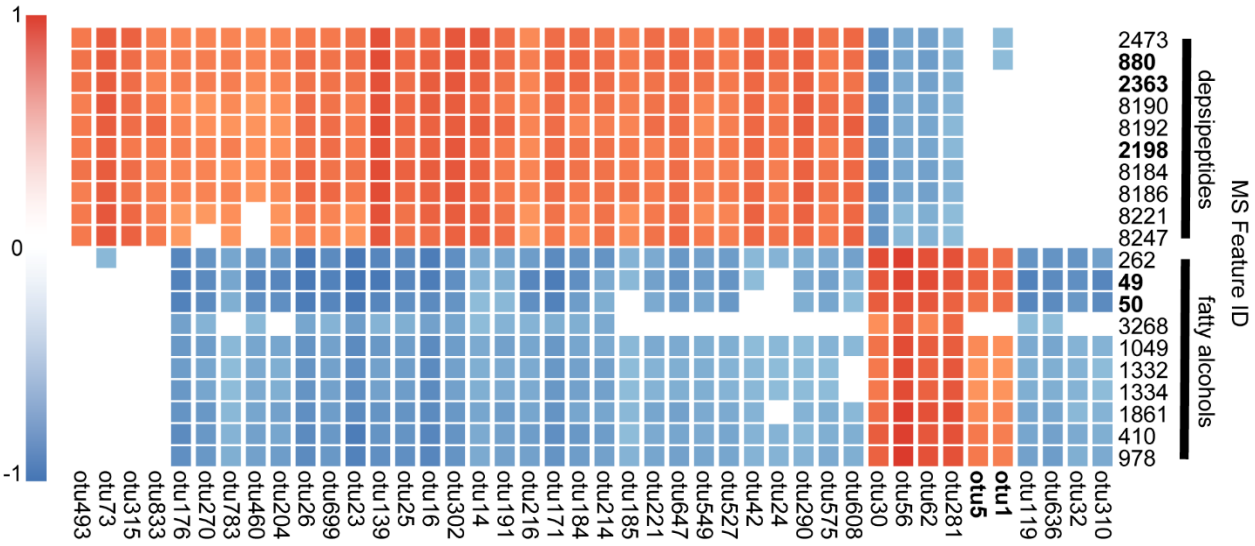


Fig 8. Correlation analysis heatmap for select metabolites with microbialite bacteria. Heatmap representation of top 20 most significant Spearman rank correlations and anticorrelations

for both the ibhayipeptolide depsipeptides and fatty alcohol metabolites with bacteria present in microbialite substrate samples. Due to significance in driving microbial diversity and high correlation values, OTUs 5 and 1 (rank 23rd and 36th highest, respectively) were also included. Spearman correlations > 0, $p < 0$ are colored red, and anticorrelations > 0, $p < 0$ colored blue. MS features and OTUs that drive spatial metabolite and microbial diversity are bolded.

Conclusion

The Schoenmakerskop microbialite system comprises mature microbialite formations with an abundance of superficial cyanobacteria as indicated by the diverse mosaic of bacterial taxa reported previously [24,41] and detected here (**Figs 2, S14, S15**). We found that metabolites and microorganisms differ between core samples of microbialites submerged in the pool throughout the tidal cycle compared to surface microbialites exposed to flowing water or air between tides. The different chemical profiles observed between submerged and surface exposed core samples could be due to variation in UV exposure [73,74], oxidation, predator grazing, and or age and microbial composition of the microbialite accretions. The new ibhayipeptolide family of cyclic depsipeptides was most abundant in substrate cores of submerged microbialites, and correlated positively with cyanobacterial (OTUs 290, 42) and bacteroidete (OTUs 302, 25) taxa. This spatial survey of a mature stromatolite bed guided prioritization of substrate cores for future deep metagenomic sequencing, which is anticipated to reveal the ibhayipeptolide-producing organism, as well as biosynthetic templates for these and other MS features detected in low abundance and assigned to putative structural classes. Importantly, deep metagenomic sequencing may provide knowledge of the metabolic functional guilds of known and new organisms in the complex mosaic

of microbialite communities and facilitate hypotheses for future studies on the role of chemical signaling in microbialite formation and maintenance. Given the microgram quantities of ibhayipeptolides and other more minor metabolites isolated by HPLC, investigation of their biological function awaits the ongoing chemical syntheses of representative structures.

Materials and Methods

Data availability

All LC–MS^2 data are available in the MassIVE repository with the identifier MSV000083729.

Results of the molecular networking analysis through GNPS can be viewed and downloaded under the following URLs:

https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=90288a3e0cb44704bb1df00b440acba1

All code and the minimal source datasets are available in a Zenodo repository with DOI: 10.5281/zenodo.10775759, which includes the MS feature table (quant table), input mgf files for FBMN and SIRIUS, the OTUn table, and correlation analysis input table.

All 16S rDNA sequences are available as an NCBI BioProject with the accession number PRJNA901469.

Inclusivity in global research

Additional information regarding the ethical, cultural, and scientific considerations specific to inclusivity in global research is included in the Supporting Information (**Checklist S1**. Inclusivity in Global Research).

Study site description

The Schoenmakerskop site (34°2'29" S 25°32'21" E; **Fig S1**) is located adjacent to the metropole of Gqerberha (previously Port Elizabeth), close to human settlements and is exposed to moderate anthropogenic influences on ambient water chemistry [75]. Microbialite formations at this site begin from the freshwater inflow in the supratidal zone (**Fig 1A**). Freshwater from seeps feed through a short region of fluvial microbialites (**Fig 1A**, stations **1** and **4**) and then into a barrage pool (**Fig 1A**, stations **2**, **3**, **5** and **6**) separated from the subtidal zone by a microbialite plateau (**Fig 1A**, stations **7**, **8** and **10**). The formations are relatively flat, forming a shallow pool (mean depth 0.35 +/- 0.06 m) and the difference in elevation between the inflow formations and the plateau formations is ~1 m vertically, stretching over a horizontal distance of ~15 m. The flag stations range from ~3-5 m apart.

Sample collection

Prior to field sampling, collection permits (RES2018/44; RES2021/81) were acquired from the South African Department Environmental Affairs (DEA) and the South African Department of Environment, Forestry and Fisheries (DFFE). The study site at Schoenmakerskop is public property and no protected species were sampled for this study. Microbialite substrate cores (**Fig S3**) were collected from nine flagged stations (**1-8** and **10**, **Figs 1A**, **1B**, and **S4-S12**) in

Schoenmakerskop barrage pool in the Eastern Cape of South Africa (**Figs S1** and **S2**) at low tide across two days (November 9 and 10, 2018; Permit No RES2018/44). Twelve substrate cores were collected from each of the flag stations **1**-**7**, while 18 cores were collected from flag station **8**, and 14 cores were collected from flag station **10**, for a total of 116 core samples. In each case a custom-made one-inch diameter, hollow metal corer was hammered into the substrate to a depth of about 5 inches with a mallet to provide cores at three different positions along the water flowpath (upper, middle, lower), to the left and right of each marker flag (**Fig S3**). On each day, at least six cores were obtained per flag station (**Figs S4-S12**). Additional cores were collected from stations **8** and **10** given their position at the edge of the seaward barrage wall. As cores were obtained, they were subsampled for separate chemical and DNA extractions. Subsamples for chemical extraction were placed in 20 mL glass vials on ice immediately in the field and frozen at -20 ℃ within two hours of collection. Subsamples for DNA extraction were placed in Falcon tubes containing RNALater and flash-frozen in the field, then stored at –20 °C within two hours of collection. Two sea water samples were also collected from an adjacent beach access (assigned as station **9**), as environmental controls for this study.

Two larger microbialite substrate samples for isolation and planar structure elucidation of metabolites were collected from flag station **7** in March 2021 (Permit No: RES2021/81), wrapped in aluminum foil and stored on ice immediately in the field until freezing at –20 °C for subsequent chemical extraction.

Sample extraction, data acquisition and processing

16S amplicon sequencing

Genomic DNA (gDNA) was extracted from microbialite samples using the Zymo Quick-DNA™ Fecal/Soil Microbe Microprep Kit (Catalog No. D6012) as per the manufacturer's instructions. DNA was extracted from approximately 50 – 100 mg of microbialite material, yielding 40 µL volumes containing between 2 – 4 µg of gDNA. Polymerase chain reaction (PCR) amplification of template DNA was performed with Miseq primers E517F (5'-CAGCAGCCGCGGTAA-3') and E969-984R (5'-GTAAGGTTCYTCGCGT-3'), which target the V4-V5 region of the bacterial 16S rDNA gene (approximately 435 nt). Each 50 µL PCR reaction was prepared with 15-20 ng gDNA template, 10 µM of each primer, 5X PCR buffer (containing $MgCl_2$), 10 µM DNTP's, 0.5 units KAPA HIFI Hotstart DNA polymerase (cat no. 07958838001). MilliQ water was used to bring the total volume of the reaction to 50 µL. Thermal cycling parameters employed were as follows; initial denaturation at 95 °C for 5 min followed by amplification with 5 cycles at 94 °C for 30 s, 45 °C for 20 s, 72 °C for 1 min and additional amplification with 18 cycles at 94 °C for 30 s, 50 °C for 20 s and 72 °C for 1 min, followed by a final elongation step of 72 °C for 5 min. Samples were gel purified using the Bioline Isolate II gel and PCR kit (cat. no. BIO-52059). Samples were sequenced on an Illumina Miseq platform, generating ~250 nt amplicon libraries (only forward reads were determined to be of a high enough quality to use in analysis). NCBI BioProject accession number: PRJNA901469.

Chemical extraction and tandem mass spectrometry of microbialite core samples

Microbialite core samples were lyophilized at Rhodes University in South Africa in November/December 2018 and shipped to UC San Diego in January (received 01/2019). Of 116 vials containing core samples for the spatial survey, six vials were broken during shipping and were discarded. Approximately 1 cm^3 (0.2 g) of each sample was placed in a scintillation vial and

5 mL 100% MeOH were added, including an empty vial as a control. The vials were left to stand for 1-2 h at room temperature before being placed at 4 °C overnight. Aliquots of these methanol extracts (160 µL) were transferred to vials, and dried in vacuo for weighing. The range of extract weights was estimated by selecting forty representative samples, from all flag stations, judged by their color and apparent mass as low, medium and high mass. Aliquots (160 µL) of each MeOH extract were added to 96-well plates (Thermo 0.5 mL, U-bottom, PP). Controls included background seawater from a remote non-stromatolite site (#9), and methanol solvent and glass scintillation vial. The plates were dried *in vacuo* using a vacuum centrifuge (Centrivap, Labconco) and resuspended in 200 µL 9:1 MeOH/H_2O per well. LC–MS^2 analysis of the 2018 core sample collection was performed as described previously [76]. In short, 10 µL of each sample were injected into a Vanquish UHPLC system coupled to a Q-Exactive quadrupole orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). For the chromatographic separation, a reversed phase C18 porous core column (Kinetex C18, 150 x 2.1 mm, 1.8 µm particle size, 100 Å pore size, Phenomenex, Torrance, USA) was used. For gradient elution a high-pressure binary gradient system was used. The mobile phase consisted of solvent A H_2O + 0.1 % formic acid (FA) and solvent B acetonitrile (ACN) + 0.1 % FA, and the flow rate was 0.5 mL/min. After injection, the samples were eluted with a linear gradient from 0-1 min at 5 % B, 1-4 min 5-60 % B, 4-10 min 60-99 % B, followed by a 3 min washout phase at 99% B and a 2 min re-equilibration phase at 5 % B. Data dependent acquisition (DDA) of MS^2 spectra was performed in positive mode. Electrospray ionization (ESI) parameters were set to 52 AU sheath gas flow, 14 AU auxiliary gas flow, 0 AU sweep gas flow and 400 °C auxiliary gas temperature. The spray voltage was set to 3.5 kV and the inlet capillary to 320 °C. 50 V S-lens level was applied. MS scan range was set to *m/z* 150-1500 with a resolution at *m/z* 200 ($R_{m/z\ 200}$) of 17,500 with one micro-scan. The maximum ion

injection time was set to 100 ms with automated gain control (AGC) target of 5E5. Up to 5 MS^2 spectra per MS^1 survey scan were recorded in DDA mode with $R_{m/z\ 200}$ of 17,500 with one micro-scan. The maximum ion injection time for MS^2 scans was set to 100 ms with a AGC target of 5E5 ions. The MS^2 precursor isolation window was set to *m/z* 1. Normalized collision energy was set to a stepwise increase from 20 to 30 to 40 % with z = 1 as default charge state. MS^2 scans were triggered at the apex of chromatographic peaks within 2 to 15 s from their first occurrence. Dynamic precursor exclusion was set to 5 s. Ions with unassigned charge states were excluded from MS^2 acquisition as well as isotope peaks.

For MS^2 data analysis, raw spectra were converted to .mzXML files using MSconvert (ProteoWizard). MS^1 and MS^2 feature extraction was performed using MZmine2.37, an IIN enabled version. The parameters used in MZmine2 are listed in **Table S3**. The MS feature table .csv and .mgf files were exported and uploaded to GNPS (gnps.ucsd.edu) for feature-based molecular networking (FBMN). For spectrum library matching and spectral networking the minimum cosine score to define spectral similarity was set to 0.7. The precursor and fragment ion mass tolerances were set to 0.01 Da, minimum matched fragment ions to 6 and minimum cluster size to 1 (MS cluster off). Molecular networks were visualized with Cytoscape 3.7.2 [77] and node information was enriched with the MS^1 peak areas from the MS feature table. The link to the GNPS job is as follows:

https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=90288a3e0cb44704bb1df00b440acba1

Data Analysis

16S amplicon data analysis

Amplicon reads were curated using Mothur [78]. Reads with a quality window average below 20, length shorter than 250 bases or greater than 500 bases, any ambiguous bases, or homopolymeric runs greater than 7, were removed from the dataset. Chimeric sequences were identified using VSEARCH and were removed from the dataset. The remaining reads were classified against the SILVA database (v138.1) and all reads classified as Chloroplasts, Mitochondria, Unknown, Archaea or Eukaryota were removed. Sequences were aligned and all that did not fit into the region into which 95% of all sequences were aligned were removed. Aligned sequences were checked again for chimeric sequences which were removed upon identification. Singletons were removed to reduce computational expense. While some believe that this may bias the analysis in terms of removing spurious sequences [79], there is evidence to suggest that even when extremely rare species are the focus of research, removal of singletons aids in chimera removal, and increases accuracy of alpha and beta diversity while decreasing computational requirements [80]. Sequences were once again filtered to remove any reads classified as Chloroplasts, Mitochondria, Unknown, Archaea or Eukaryota as anecdotal experience has shown that some contaminants, primarily "unknowns", are still present at this stage and a second sweep is required to thoroughly clean up the dataset. The remaining sequences were clustered into operational taxonomic units (OTUs) at a distance of 0.03 and counted per sample. Representative sequences for each OTU were similarly extracted and classified against the SILVA database (V138.1). OTU counts per sample were converted to relative abundance and used to generate non-metric multidimensional scaling (NMDS) plots in R using the vegan, plotly, and ggplot packages. Associated pairwise Analysis of Similarity (ANOSIM) scores were calculated for the different groups of samples (e.g., samples that were submerged or exposed to flowing water) using the vegan package in R. Canonical-correlation analysis (CCA) of bacterial OTUs was performed on

normalized data using custom functions adapted from [81] and available at https://userweb.eng.gla.ac.uk/umer.ijaz/bioinformatics/ecological.html, which utilizes vegan [82] in R. Results were visualized in R using ggplot2 v3.4.4 (**Fig S13**) [83].

Relative abundance plots of bacterial OTUs in the stromatolite core samples (**Figs 2**, **S14**, **S15**) were generated using ggplot2 v3.4.4 in R [83]. Phylogeny of the conserved OTUs in Schoenmakerskop, relative to reference sequences and conserved OTUs from a previous study, was inferred using maximum likelihood (**Fig 3**). Briefly, conserved OTU sequences from this study and a previous study were aligned against the NR database using NCBI BLASTn [84,85] to find reference sequences. A spirochete sequence was chosen as an outgroup. All sequences were aligned using muscle (v 5.1) [86]. The alignment was then used to build a phylogenetic tree in MEGA 11 [87], using the Maximum Likelihood method with 1000 bootstrap replicates. The Tamura-Nei model was used, with uniform rates amongst sites where all sites were used. The initial tree was generated using the Neighbor-Joining method and the Nearest-Neighbor-Interchange was used for the heuristic method. The bootstrapped tree was visualized in iTol [88] and re-rooted at the outgroup.

LC–MS^2 data preparation for analysis

The LC–MS^2 data table was exported from MZmine2 and was further processed in the statistical software R (version 4.1.2) using modified code from a previously published analysis [89]. This table was blank-subtracted with a cutoff of 0.3 (i.e., MS features with a ratio of mean intensity in blanks vs mean intensity of MS features less than 30% were considered background noise and removed). After blank subtraction, the data were normalized by total ion count. The blank subtracted, normalized data were used for subsequent analysis in QIIME2.

LC–MS^2 data analysis in QIIME 2

Subsequent visualization and statistical analysis of LC–MS^2 data (**Fig 1**) was performed in QIIME2 (Bolyen et al. 2019). Unsupervised principal coordinate analysis (PCoA) was performed using the Bray-Curtis distance and the qiime beta and diversity pcoa plugin. PERMANOVA analysis was performed to quantify the significance of separation. All code is provided in the accompanying Zenodo repository (see Data Availability).

Spearman rank correlation analysis of MS features and 16S rRNA sequence data

The tables of MS feature abundances and bacterial OTU counts were used to generate a correlation analysis visualized as a heatmap (**Fig 8**). Prior to correlation, MS features observed in less than 10% of samples were removed. Amplicon sequencing data was transformed to proportions by dividing the counts for each sample by the total sequencing depth of the sample. Both datasets were mean centered and scaled, and combined in a single input file. Spearman rank correlations were calculated between each pair (microbe~metabolite) of MS features using R package "Hmisc". P values were adjusted according to the Benjamini-Hochberg method to control the false discovery rate (FDR). FDR < 0.05 was considered significant. This linear statistical correlation analysis method was chosen over newer neural networking covariance methods due to the benchmarking performed by Quinn and Erb [90]. All code is provided in the accompanying Zenodo repository (see Data Availability).

Compound class prediction using SIRIUS and CANOPUS

Molecular formulas were assigned using the SIRIUS GUI. MzMine2 was used to generate and export an .mgf file specific for SIRIUS4 analysis of the 2018 core sample set. For individual analysis of the four major ibhayipeptolide depsipeptides, HRMS and MS^2 (20, 40 and 60 eV)

spectra of both $[M+H]^+$ and $[M+Na]^+$ ($[M-H]^-$ was also included for MS feature 2198) were manually uploaded as csv files to SIRIUS5. SIRIUS, CSI:FingerID, and CANOPUS analyses were performed with the parameters listed in **Table S4** and **S5**.

Compound isolation and structure elucidation

General experimental procedures

NMR data were acquired on a Bruker Avance III 800 MHz spectrometer equipped with a 5 mm TCI cryoprobe, with the residual solvent used as an internal standard ($CDCl_3$, δ_H 7.26, δ_C 77.16 ppm). High resolution (HR)TOFMS (ESI^+) and tandem MS data were recorded on an Agilent 1260 infinity II LC coupled to a 6545 QToF MS. The mobile phase consisted of ultra-pure H_2O (A) and ACN (B) with 0.1 % formic acid. A gradient method from 15 % B to 90 % B in 9 min at a flow rate of 0.4 mL/min was used. The column (Phenomenex Kinetex C_{18}, 2.6 µm, 100 Å, 50 mm x 2.1 mm) was re-equilibrated before each injection and the column compartment was maintained at 40 ˚C throughout each run. Semi-preparative HPLC (Phenomenex Kinetex C_{18}, 5 µm, 100 Å, 250 mm x 10 mm) utilized isocratic elution conditions or a gradient system with a flow rate of 4 mL/min on a Shimadzu LC-20AD HPLC system operating at room temperature, equipped with an SPD-M20A photodiode array detector. All samples were filtered through a 0.2 µm nylon filter or centrifuged at 14,000 rpm for 5 min before LC–MS and HPLC analysis. General reagents were from VWR International.

Extraction and chromatographic purification of depsipeptides

Two larger samples collected in March 2021 from flag station 7 for characterization of pure depsipeptides were deep frozen before lyophilization for 48 h and each processed separately according to the following general protocol for separate LC–MS profiling in case they varied in chemistry. The dry biomass (1.91 kg total for both samples) was crushed by mortar and pestle into pieces smaller than 0.5 cm. The resulting powdered biomass was extracted twice in DCM-MeOH (2:1) by soaking for 48 h at room temperature in the dark and without agitation. Approximately 2.8 L of solvent was used per kg of dry biomass. After each extraction, the supernatant was filtered through filter paper and concentrated under reduced pressure to afford two organic extracts (4.13 g total).

Each organic extract was initially fractionated into 10 fractions (A-G) by normal phase chromatography on a Teledyne ISCO CombiFlash R_f 200 system using a RediSep column (silica 40 g Gold). The mobile phase gradient used was 100% hexanes to 100% EtOAc over 23 min, holding at 100% EtOAc for 3.0 min, then ramping to 75% EtOAc-MeOH over 7 min and holding for 5 min. Due to the largest amount of ibhayipeptolide A present, fraction B (7-15 min) was split into 5 subfractions (B1-B5) using RP_{18} chromatography on the same Combiflash system. A RediSep column (30 g HP C_{18}) was used with a gradient of MeOH-H_2O (50% for 2 min before ramping to 100% MeOH over 5 min and holding for 9 min). Ibhayipeptolides A and B (MS features 2198 and 8186) were in highest abundance in subfraction B3 (4-8 min). Ibhayipeptolides A (t_R 9.1 min) and B (t_R 8.0 min) were isolated by semi-preparative HPLC (Kinetex C_{18} 250x10 mm) from subfraction B3 utilizing a shallow gradient (90% ACN to 100% ACN over 17 min). Ibhayipeptolides A (80% ACN, Kinetex C_{18} 100x4.6 mm, 8.6-9.5 min) and B (80% ACN, Kinetex C18 250x10 mm, 17.1-17.7 min) were purified further using isocratic HPLC conditions. Less than 100 μg of each depsipeptide were isolated.

Analytical data for ibhayipeptolides with assigned planar structures

*Ibhayipeptolide A (MS feature 2198)***:** white amorphous solid; HRESI$^+$MS: *m/z* 777.4806 $[M+H]^+$ (calcd for $C_{44}H_{65}N_4O_8^+$, 777.4797; Δ ppm = 1.2), *m/z* 799.4562 $[M+Na]^+$ (calcd for $C_{44}H_{64}N_4O_8Na^+$, 799.4616; Δ ppm = 6.8). *m/z* 775.4674 $[M-H]^-$ (calcd for $C_{44}H_{63}N_4O_8^-$, 775.4651; Δ ppm = 3.0). SIRIUS score: $[M+H]^+$ 54.2%, $[M+Na]^+$ 59.4%, $[M-H]^-$ 40.2%.

*Ibhayipeptolide B (MS feature 8186)***:** white amorphous solid; HRESI$^+$MS: *m/z* 763.4666 $[M+H]^+$ (calcd for $C_{43}H_{63}N_4O_8^+$, 763.4640; Δ ppm = 3.4), *m/z* 785.4488 $[M+Na]^+$ (calcd for $C_{43}H_{62}N_4O_8Na^+$, 785.4460; Δ ppm = 3.6). SIRIUS score: $[M+H]^+$ 87.4%, $[M+Na]^+$ 60.3%.

*Ibhayipeptolide C (MS feature 8190)***:** white amorphous solid; HRESI$^+$MS: *m/z* 791.5013 $[M+H]^+$ (calcd for $C_{45}H_{67}N_4O_8^+$, 791.4953; Δ ppm = 7.6), *m/z* 813.4831 $[M+Na]^+$ (calcd for $C_{45}H_{66}N_4O_8Na^+$, 813.4773; Δ ppm = 7.1). SIRIUS score: $[M+H]^+$ 18.5%, $[M+Na]^+$ 49.1%.

*Ibhayipeptolide D (MS feature 8185)***:** white amorphous solid; HRESI$^+$MS: *m/z* 805.5135 $[M+H]^+$ (calcd for $C_{46}H_{69}N_4O_8^+$, 805.5110; Δ ppm = 3.1, *m/z* 827.4955 $[M+Na]^+$ (calcd for $C_{46}H_{68}N_4O_8Na^+$, 827.4929; Δ ppm = 3.1). SIRIUS score: [M+H]+ 92.3%, [M+Na]+ 85.1%.

Acknowledgments

The authors acknowledge the South African Institute for Aquatic Biodiversity Aquatic Genomics Research Platform for support in NGS sequencing data analysis. The authors also

acknowledge the Centre for High Performance Computing, South Africa, for provision of computing resources to curate the Illumina-sequenced DNA datasets, and the OSU NMR facility.

Author Contributions

KLM, RAD, PCD, ATA, EWI, DP, ECG contributed to conceptualization of the study; KLM, RAD, PCD, JCK contributed to funding acquisition; all authors contributed to the investigation; ATA, SCW and LSM contributed to code curation; GFN, ATA, EWI, SCW, DP, and LSM contributed to formal analysis and data visualization; ATA, GFN, KLM contributed to writing – original draft; all authors contributed to writing – review and editing.

Conflicts of interest

There are no conflicts to declare.

References

1. Schmidt R, Ulanova D, Wick LY, Bode HB, Garbeva P. Microbe-driven chemical ecology: past, present and future. ISME J. 2019 Nov;13(11):2656–63.
2. Atkinson S, Williams P. Quorum sensing and social networking in the microbial world. J R Soc Interface. 2009 Nov 6;6(40):959–78.
3. Kost C, Patil KR, Friedman J, Garcia SL, Ralser M. Metabolic exchanges are ubiquitous in natural microbial communities. Nat Microbiol. 2023 Dec;8(12):2244–52.
4. Shaffer JP, Nothias LF, Thompson LR, Sanders JG, Salido RA, Couvillion SP, et al. Standardized multi-omics of Earth's microbiomes reveals microbial and metabolite

diversity. Nat Microbiol. 2022 Dec;7(12):2128–50.

5. Tuttle RN, Demko AM, Patin NV, Kapono CA, Donia MS, Dorrestein P, et al. Detection of Natural Products and Their Producers in Ocean Sediments. Appl Environ Microbiol [Internet]. 2019 Apr 15;85(8). Available from: http://dx.doi.org/10.1128/AEM.02830-18

6. Bauermeister A, Mannochio-Russo H, Costa-Lotufo LV, Jarmusch AK, Dorrestein PC. Mass spectrometry-based metabolomics in microbiome investigations. Nat Rev Microbiol. 2022 Mar;20(3):143–60.

7. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat Biotechnol. 2016 Aug 9;34(8):828–37.

8. Dührkop K, Fleischauer M, Ludwig M, Aksenov AA, Melnik AV, Meusel M, et al. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. Nat Methods. 2019 Apr;16(4):299–302.

9. Dührkop K, Shen H, Meusel M, Rousu J, Böcker S. Searching molecular structure databases with tandem mass spectra using CSI:FingerID. Proc Natl Acad Sci U S A. 2015 Oct 13;112(41):12580–5.

10. Djoumbou Feunang Y, Eisner R, Knox C, Chepelev L, Hastings J, Owen G, et al. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. J Cheminform. 2016 Nov 4;8:61.

11. Dührkop K, Nothias LF, Fleischauer M, Reher R, Ludwig M, Hoffmann MA, et al. Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. Nat Biotechnol. 2021 Apr;39(4):462–71.

12. Kim HW, Wang M, Leber CA, Nothias LF, Reher R, Kang KB, et al. NPClassifier: A Deep Neural Network-Based Structural Classification Tool for Natural Products. J Nat Prod. 2021 Nov 26;84(11):2795–807.

13. Riding R. Microbial carbonates: the geological record of calcified bacterial-algal mats and biofilms. Sedimentology. 2000 Feb;47:179–214.

14. Bosak T, Knoll AH, Petroff AP. The Meaning of Stromatolites. Annu Rev Earth Planet Sci. 2013 May 30;41(1):21–44.

15. Vasconcelos C, Dittrich M, McKenzie JA. Evidence of microbiocoenosis in the formation of laminae in modern stromatolites. Facies. 2014 Jan 1;60(1):3–13.

16. Hohl SV, Viehmann S. Stromatolites as geochemical archives to reconstruct microbial habitats through deep time: Potential and pitfalls of novel radiogenic and stable isotope systems. Earth-Sci Rev. 2021 Jul 1;218:103683.

17. Stolz JF. Stromatolites: Linking the Past to the Future. Environ Microbiol [Internet]. 2022 Nov 1; Available from: http://dx.doi.org/10.1111/1462-2920.16276

18. White RA. The Global Distribution of Modern Microbialites: Not So Uncommon After All. In: Souza V, Segura A, Foster J, editors. Astrobiology and Cuatro Ciénegas Basin as an Analog of Early Earth. Springer; 2020. p. 107–34.

19. Suosaari EP, Reid RP, Mercadier C, Vitek BE, Oehlert AM, Stolz JF, et al. The microbial carbonate factory of Hamelin Pool, Shark Bay, Western Australia. Sci Rep. 2022 Jul 28;12(1):12902.

20. Myshrall KL, Mobberley JM, Green SJ, Visscher PT, Havemann SA, Reid RP, et al. Biogeochemical cycling and microbial diversity in the thrombolitic microbialites of Highborne Cay, Bahamas. Geobiology. 2010 Sep;8(4):337–54.

21. Albarracín VH, Gärtner W, Farias ME. Forged Under the Sun: Life and Art of Extremophiles from Andean Lakes. Photochem Photobiol. 2016 Jan;92(1):14–28.

22. Belan MA, Brady AL, Kim ST, Lim DSS, Slater GF. Spatial Distribution and Preservation of Carbon Isotope Biosignatures in Freshwater Microbialite Carbonate. ACS Earth Space Chem. 2019 Mar 21;3(3):335–43.

23. Pepe-Ranney C, Berelson WM, Corsetti FA, Treants M, Spear JR. Cyanobacterial construction of hot spring siliceous stromatolites in Yellowstone National Park. Environ Microbiol. 2012 May;14(5):1182–97.

24. Perissinotto R, Paul-Pierre BTG, Miranda NAF, Dorrington RA, Matcher GF, Strydom N, et al. Tufa stromatolite ecosystems on the South African south coast. South African Journal of Science. 2014;110(9/10):1–8.

25. Rishworth GM, Cawthra HC, Dodd C, Perissinotto R. Peritidal stromatolites as indicators of stepping-stone freshwater resources on the Palaeo-Agulhas Plain landscape. Quat Sci Rev. 2020 May 1;235:105704.

26. Smith AM, Uken R. Living marine stromatolites at Kei River mouth. South African Journal of Science. 2003;99(5):200.

27. Cooper A, Smith A, Rishworth G, Dodd C, Forbes M, Cawthra H, et al. Microbialites of modern siliciclastic rock coasts. J Sediment Res. 2022 Jul 31;92(7):619–34.

28. Dillon JG. The Role of Sulfate Reduction in Stromatolites and Microbial Mats: Ancient and Modern Perspectives. In: Tewari VC, Seckbach J, editors. Stromatolites: interaction of microbes with sediments. Springer; 2011. p. 573–90.

29. Reid RP, Visscher PT, Decho AW, Stolz JF, Bebout BM, Dupraz C, et al. The role of microbes in accretion, lamination and early lithification of modern marine stromatolites. Nature. 2000 Aug 31;406(6799):989–92.

30. Visscher PT, Pamela Reid R, Bebout BM. Microscale observations of sulfate reduction: Correlation of microbial activity with lithified micritic laminae in modern marine stromatolites. Geology. 2000 Oct 1;28(10):919–22.

31. Strauss H. Sulfur Isotopes in Stromatolites. In: Tewari V, Seckbach J, editors. STROMATOLITES: Interaction of Microbes with Sediments. Dordrecht: Springer Netherlands; 2011. p. 687–701.

32. Louyakis AS, Mobberley JM, Vitek BE, Visscher PT, Hagan PD, Reid RP, et al. A Study of the Microbial Spatial Heterogeneity of Bahamian Thrombolites Using Molecular, Biochemical, and Stable Isotope Analyses. Astrobiology. 2017 May;17(5):413–30.

33. Lamérand C, Shirokova LS, Bénézeth P, Rols JL, Pokrovsky OS. Carbon sequestration potential of Mg carbonate and silicate biomineralization in the presence of cyanobacterium Synechococcus. Chem Geol. 2022 Jun 20;599:120854.

34. Khodadad CLM, Foster JS. Metagenomic and metabolic profiling of nonlithifying and lithifying stromatolitic mats of Highborne Cay, The Bahamas. PLoS One. 2012 May 25;7(5):e38229.

35. Casaburi G, Duscher AA, Reid RP, Foster JS. Characterization of the stromatolite microbiome from Little Darby Island, The Bahamas using predictive and whole shotgun metagenomic analysis. Environ Microbiol. 2016 May;18(5):1452–69.

36. Allen MA, Neilan BA, Burns BP, Jahnke LL, Summons RE. Lipid biomarkers in Hamelin Pool microbial mats and stromatolites. Org Geochem. 2010 Nov 1;41(11):1207–18.

37. Burns BP, Pomati F, Goh F, Yasar SA, Neilan BA. Stromatolites as a resource for novel natural products. Orig Life Evol Biosph. 2006 Dec;36(5-6):623–4.

38. D'Agostino PM, Woodhouse JN, Liew HT, Sehnal L, Pickford R, Wong HL, et al. Bioinformatic, phylogenetic and chemical analysis of the UV-absorbing compounds scytonemin and mycosporine-like amino acids from the microbial mat communities of Shark Bay, Australia. Environ Microbiol. 2019 Feb;21(2):702–15.

39. Smith AM, Uken R, Thackeray Z. Cape Morgan peritidal stromatolites: the origin of lamination. South African Journal of Science. 2005;101(3):107–8.

40. Dodd C, Anderson CR, Perissinotto R, du Plooy SJ, Rishworth GM. Hydrochemistry of peritidal stromatolite pools and associated freshwater inlets along the Eastern Cape coast, South Africa. Sediment Geol. 2018 Oct 1;373:163–79.

41. Waterworth SC, Isemonger EW, Rees ER, Dorrington RA, Kwan JC. Conserved bacterial genomes from two geographically isolated peritidal stromatolite formations shed light on potential functional guilds. Environ Microbiol Rep. 2021 Apr;13(2):126–37.

42. Rishworth GM, Perissinotto R, Bird MS, Strydom NA, Peer N, Miranda NAF, et al. Non-reliance of metazoans on stromatolite-forming microbial mats as a food resource. Sci Rep. 2017 Feb 16;7:42614.

43. Rishworth GM, Perissinotto R, Miranda NAF, Bornman TG, Steyn PP. Phytoplankton community dynamics within peritidal pools associated with living stromatolites at the freshwater–marine interface. Aquat Sci. 2017 Apr 1;79(2):357–70.

44. Weston RLA, Perissinotto R, Rishworth GM, Steyn PP. Macroinvertebrate variability between microhabitats of peritidal stromatolites along the South African coast. Marine Ecology Progress Series. 2018;605:37–47.

45. Schmid R, Petras D, Nothias LF, Wang M, Aron AT, Jagels A, et al. Ion identity molecular networking for mass spectrometry-based metabolomics in the GNPS environment. Nat Commun. 2021 Jun 22;12(1):3832.

46. Chen LZ, Li DH, Song LR, Hu CX, Wang GH, Liu YD. Effects of salt stress on carbohydrate metabolism in desert soil alga *Microcoleus vaginatus* Gom. J Integr Plant Biol. 2006 Aug;48(8):914–9.

47. Mlewski EC, Pisapia C, Gomez F, Lecourt L, Soto Rueda E, Benzerara K, et al. Characterization of Pustular Mats and Related Rivularia-Rich Laminations in Oncoids From the Laguna Negra Lake (Argentina). Front Microbiol. 2018 May 22;9:996.

48. Sabater S, Guasch H, Romaní A, Muñoz I. Stromatolitic communities in Mediterranean streams: adaptations to a changing environment. Biodiversity & Conservation. 2000 Mar 1;9(3):379–92.

49. Cooper JAG, Smith AM, Arnscheidt J. Contemporary stromatolite formation in high intertidal rock pools, Giant's Causeway, Northern Ireland: preliminary observations. Journal of Coastal Research. 2013 Apr;65(sp2):1675–80.

50. Freytet P, Verrecchia EP. Freshwater organisms that build stromatolites: a synopsis of biocrystallization by prokaryotic and eukaryotic algae. Sedimentology. 1998 May;45(3):535–63.

51. Chacón E, Berrendero E, Garcia Pichel F. Biogeological signatures of microboring cyanobacterial communities in marine carbonates from Cabo Rojo, Puerto Rico. Sediment Geol. 2006 Mar 15;185(3):215–28.

52. Mohimani H, Gurevich A, Mikheenko A, Garg N, Nothias LF, Ninomiya A, et al. Dereplication of peptidic natural products through database search of mass spectra. Nat Chem Biol. 2017 Jan;13(1):30–7.

53. Crawford JM, Portmann C, Kontnik R, Walsh CT, Clardy J. NRPS substrate promiscuity diversifies the xenematides. Org Lett. 2011 Oct 7;13(19):5144–7.

54. Takeda K, Kemmoku K, Satoh Y, Ogasawara Y, Shin-Ya K, Dairi T. N-Phenylacetylation and Nonribosomal Peptide Synthetases with Substrate Promiscuity for Biosynthesis of Heptapeptide Variants, JBIR-78 and JBIR-95. ACS Chem Biol. 2017 Jul 21;12(7):1813–9.

55. Cai X, Nowak S, Wesche F, Bischoff I, Kaiser M, Fürst R, et al. Entomopathogenic bacteria use multiple mechanisms for bioactive peptide library design. Nat Chem. 2017 Apr;9(4):379–86.

56. Nogle LM, Gerwick WH. Isolation of four new cyclic depsipeptides, antanapeptins A-D, and dolastatin 16 from a Madagascan collection of Lyngbya majuscula. J Nat Prod. 2002 Jan;65(1):21–4.

57. Tripathi A, Puddick J, Prinsep MR, Lee PPF, Tan LT. Hantupeptin A, a cytotoxic cyclic depsipeptide from a Singapore collection of Lyngbya majuscula. J Nat Prod. 2009 Jan;72(1):29–32.

58. Tripathi A, Puddick J, Prinsep MR, Lee PPF, Tan LT. Hantupeptins B and C, cytotoxic cyclodepsipeptides from the marine cyanobacterium Lyngbya majuscula. Phytochemistry. 2010 Feb;71(2-3):307–11.

59. Bunyajetpong S, Yoshida WY, Sitachitta N, Kaya K. Trungapeptins A-C, cyclodepsipeptides from the marine cyanobacterium Lyngbya majuscula. J Nat Prod. 2006 Nov;69(11):1539–42.

60. Levert A, Alvariño R, Bornancin L, Abou Mansour E, Burja AM, Genevière AM, et al. Structures and Activities of Tiahuramides A-C, Cyclic Depsipeptides from a Tahitian Collection of the Marine Cyanobacterium Lyngbya majuscula. J Nat Prod. 2018 Jun 22;81(6):1301–10.

61. Zhao P, Xue Y, Gao W, Li J, Zu X, Fu D, et al. Actinobacteria-Derived peptide antibiotics since 2000. Peptides. 2018 May;103:48–59.

62. Demay J, Bernard C, Reinhardt A, Marie B. Natural Products from Cyanobacteria: Focus on Beneficial Activities. Mar Drugs. 2019 May 30;17(6):320.

63. Dose B, Niehs SP, Scherlach K, Flórez LV, Kaltenpoth M, Hertweck C. Unexpected Bacterial Origin of the Antibiotic Icosalide: Two-Tailed Depsipeptide Assembly in Multifarious Burkholderia Symbionts. ACS Chem Biol. 2018 Sep 21;13(9):2414–20.

64. Oku N, Kawabata K, Adachi K, Katsuta A, Shizuri Y. Unnarmicins A and C, new antibacterial depsipeptides produced by marine bacterium Photobacterium sp. MBIC06485. J Antibiot . 2008 Jan;61(1):11–7.

65. Bertin MJ, Roduit AF, Sun J, Alves GE, Via CW, Gonzalez MA, et al. Tricholides A and B and Unnarmicin D: New Hybrid PKS-NRPS Macrocycles Isolated from an Environmental Collection of Trichodesmium thiebautii. Mar Drugs [Internet]. 2017 Jun

30;15(7). Available from: http://dx.doi.org/10.3390/md15070206

66. Mansson M, Nielsen A, Kjærulff L, Gotfredsen CH, Wietz M, Ingmer H, et al. Inhibition of virulence gene expression in Staphylococcus aureus by novel depsipeptides from a marine photobacterium. Mar Drugs. 2011 Dec;9(12):2537–52.

67. Kjaerulff L, Nielsen A, Mansson M, Gram L, Larsen TO, Ingmer H, et al. Identification of four new agr quorum sensing-interfering cyclodepsipeptides from a marine Photobacterium. Mar Drugs. 2013 Dec 12;11(12):5051–62.

68. Baldry M, Kitir B, Frøkiær H, Christensen SB, Taverne N, Meijerink M, et al. The agr Inhibitors Solonamide B and Analogues Alter Immune Responses to Staphylococccus aureus but Do Not Exhibit Adverse Effects on Immune Cell Functions. PLoS One. 2016 Jan 5;11(1):e0145618.

69. Zhang SD, Lindqvist LL, Isbrandt T, Borre IL, Wibowo M, Nielsen MW, et al. Solonamides, a Group of Cyclodepsipeptides, Influence Motility in the Native Producer Photobacterium galatheae S2753. Appl Environ Microbiol. 2022 Sep 13;88(17):e0110522.

70. Wang X, Gong X, Li P, Lai D, Zhou L. Structural Diversity and Biological Activities of Cyclic Depsipeptides from Fungi. Molecules. 2018 Jan 15;23(1):169.

71. Süssmuth R, Müller J, von Döhren H, Molnár I. Fungal cyclooligomer depsipeptides: from classical biochemistry to combinatorial biosynthesis. Nat Prod Rep. 2011 Jan;28(1):99–124.

72. Shiomi K, Matsui R, Kakei A, Yamaguchi Y, Masuma R, Hatano H, et al. Verticilide, a new ryanodine-binding inhibitor, produced by Verticillium sp. FKI-1033. J Antibiot . 2010 Feb;63(2):77–82.

73. Popall RM, Bolhuis H, Muyzer G, Sánchez-Román M. Stromatolites as Biosignatures of Atmospheric Oxygenation: Carbonate Biomineralization and UV-C Resilience in a Geitlerinema sp. - Dominated Culture. Front Microbiol. 2020 May 19;11:948.

74. Zannier F, Portero LR, Douki T, Gärtner W, Farías ME, Albarracín VH. Proteomic Signatures of Microbial Adaptation to the Highest Ultraviolet-Irradiation on Earth: Lessons From a Soil Actinobacterium. Front Microbiol. 2022 Mar 15;13:791714.

75. Rishworth GM, Perissinotto R, Bornman TG, Lemley DA. Peritidal stromatolites at the convergence of groundwater seepage and marine incursion: Patterns of salinity, temperature and nutrient variability. J Mar Syst. 2017 Mar 1;167:68–77.

76. Petras D, Koester I, Da Silva R, Stephens BM, Haas AF, Nelson CE, et al. High-Resolution Liquid Chromatography Tandem Mass Spectrometry Enables Large Scale Molecular Characterization of Dissolved Organic Matter. Frontiers in Marine Science. 2017;4:405.

77. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003 Nov;13(11):2498–504.

78. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009 Dec;75(23):7537–41.

79. Reitmeier S, Hitch TCA, Treichel N, Fikas N, Hausmann B, Ramer-Tait AE, et al. Handling of spurious sequences affects the outcome of high-throughput 16S rRNA gene amplicon profiling. ISME Commun. 2021 Jun 29;1(1):31.

80. Auer L, Mariadassou M, O'Donohue M, Klopp C, Hernandez-Raquet G. Analysis of large 16S rRNA Illumina data sets: Impact of singleton read filtering on microbial community description. Mol Ecol Resour. 2017 Nov;17(6):e122–32.

81. Torondel B, Ensink JHJ, Gundogdu O, Ijaz UZ, Parkhill J, Abdelahi F, et al. Assessment of the influence of intrinsic environmental and geographical factors on the bacterial ecology of pit latrines. Microb Biotechnol. 2016 Mar;9(2):209–23.

82. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara R, et al. Package "vegan". Community ecology package, version 2. The comprehensive R network (CRAN) [Google Scholar]. 2013;

83. Wickham H. Data Transformation. In: ggplot2, Elegant Graphics for Data Analysis. Springer International Publishing; 2016. p. 203–20.

84. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 Oct 5;215(3):403–10.

85. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2013 Jan;41(Database issue):D8–20.

86. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004 Aug 19;5:113.

87. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol. 2021 Jun 25;38(7):3022–7.

88. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021 Jul 2;49(W1):W293–6.

89. Pakkir Shah AK, Walter A, Ottosson F, Russo F, Navarro-Díaz M, Boldt J, et al. The Hitchhiker's Guide to Statistical Analysis of Feature-based Molecular Networks from Non-Targeted Metabolomics Data. ChemRxiv [Internet]. 2023 Nov 1 [cited 2023 Dec 26]; Available from: https://chemrxiv.org/engage/chemrxiv/article-

details/6540eb2548dad23120c52242

90. Quinn TP, Erb I. Examining microbe-metabolite correlations by linear methods. Nat Methods. 2021 Jan;18(1):37–9.