# Leptochelins A-C, Cytotoxic Metallophores Produced by Geographically Dispersed *Leptothoe* Strains of Marine Cyanobacteria

Nicole E. Avalon,[1+] Mariana A. Reis,[2+] Christopher C. Thornburg,[3+] R. Thomas Williamson,[4] Daniel Petras,[5,6,7] Allegra T. Aron,[5] George F. Neuhaus,[3] Momen Al-Hindy,[1] Jana Mitrevska,[1] Leonor Ferreira,[2] João Morais,[2] Yasin El Abiead,[5] Evgenia Glukhov,[1] Kelsey L. Alexander,[1,8] F. Alexandra Vulpanovici,[3] Matthew J. Bertin,[1] Syrena Whitner,[1] Hyukjae Choi,[9] Gabriella Spengler,[10] Kirill Blinov,[11] Ameen M. Almohammadi,[12] Lamiaa A. Shaala,[13] William R. Kew,[14] Ljiljana Paša-Tolić,[14] Diaa T. A. Youssef,[15,16] Pieter C. Dorrestein,[5] Vitor Vasconcelos,[2,*] Lena Gerwick,[1] Kerry L. McPhail,[3,*] William H. Gerwick[1,5,*]

*[1]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, UCSD, La Jolla, CA, USA; [2]CIIMAR/CIMAR, Interdisciplinary Centre of Marine and Environmental Research, University of Porto, Matosinhos, Portugal; [3]College of Pharmacy, Oregon State University, Corvallis, OR, USA; [4]Department of Chemistry and Biochemistry, University of North Carolina Wilmington, Wilmington, NC, USA; [5]Skaggs School of Pharmacy and Pharmaceutical Sciences, UCSD, La Jolla, CA, USA; [6]University of California Riverside, Department of Biochemistry, Riverside, CA, USA; [7]University of Tuebingen, CMFI Cluster of Excellence, Tuebingen, Germany; [8]Department of Chemistry and Biochemistry, UCSD, La Jolla, CA, USA; [9]College of Pharmacy, Yeungnam University, Gyeongsan, South Korea; [10]Department of Medical Microbiology, Albert Szent-Györgyi Health Center and Albert Szent-Györgyi Medical School, University of Szeged, Szeged, Hungary; [11]Molecule Apps, LLC, Corvallis, Oregon, USA; [12]Department of Pharmacy Practice, Faculty of Pharmacy, King Abdulaziz University, Jeddah, KSA; [13]Suez Canal University Hospital, Suez Canal University, Ismailia, Egypt; [14]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99354, USA; [15]Department of Natural Products, Faculty of Pharmacy, King Abdulaziz University, Jeddah, KSA; [16]Department of Pharmacognosy, Faculty of Pharmacy, Suez Canal University, Ismailia, Egypt*

*+ These authors contributed equally to the project.*
** Co-corresponding authors*

**Abstract:** Metals are important co-factors in the metabolic processes of cyanobacteria including photosynthesis, cellular respiration, DNA replication, and the biosynthesis of primary and secondary metabolites. In adaptation to the marine environment, cyanobacteria use metallophores to acquire trace metals when necessary as well as reduce potential toxicity from excessive metal concentrations. Leptochelins A-C were identified as structurally novel metallophores from three geographically dispersed cyanobacteria of the genus *Leptothoe*. The leptochelins are comprised of halogenated linear NRPS-PKS hybrid products with multiple heterocycles that have potential for hexadentate and tetradentate coordination with metal ions. The genomes of the three leptochelin producers were sequenced, and retrobiosynthetic analysis revealed one candidate biosynthetic gene cluster (BGC) consistent with the structure of leptochelin. The putative BGC is highly homologous in all three *Leptothoe* strains, and all possess genetic signatures associated with metallophores. Post-column infusion of metals using an LC-MS metabolomics workflow performed with leptochelin A and B revealed promiscuous binding of iron, copper, cobalt, and zinc, with greatest preference for copper. Iron depletion and copper toxicity experiments support the hypothesis that leptochelin metallophores may play a key ecological role in iron acquisition and in copper detoxification. In addition, the leptochelins possess significant cytotoxicity against several cancer cell lines.

## Introduction

Microorganisms have a remarkable ability to sense and adapt to their environment. As a result, they are ubiquitous and inhabit some of the most extreme habitats on Earth, from deep sea hydrothermal vents, desert crusts, high latitude ice floes, to acid lakes containing toxic levels of dissolved metals.[1–4] This capacity for adaptation to diverse environmental conditions is also a characteristic of one of the most ancient groups of microorganisms, the Cyanobacteria. Arising some 2.5 bya, these organisms have evolved specialized capacities to thrive in such diverse situations as sun-exposed tropical reef systems, concrete building walls, terrestrial hydrothermal vents, and oligotrophic ocean waters.[5–7] Cyanobacteria have requirements for enzymatic metal cofactors that are in very limited supply in oligotrophic environments. To flourish in these habitats, they and other microorganisms produce a variety of small molecule natural products, known generally as metallophores, that assist in their acquisition of diverse metal ions required for enzymatic catalysis.[8] Most metallophores exhibit tetradentate or hexadentate coordination when binding metal ions with high affinity. The best studied class of acquisition metallophores are those with a high-affinity for iron (III), known as siderophores.[9] There are several classifications of siderophores based on the structural groups involved in metal ion coordination, including hydroxymates such as putrebactin and bisucabactin, catecholates such as bacillibactin, enterobactin, and vibriobactin, phenolates such as yersiniabactin and pyochelin, carboxylates such as staphyloferrins, and mixed type such as pyoverdine and aerobactin.[10] Examples of previously reported acquisition metallophores from cyanobacteria include schizokinen and the synechobactins, both of which are hydroxamate metallophores, and anachelin, a mixed type metallophore.[11–13]

In environments where metal ion concentrations reach levels that would normally be toxic, microorganisms produce small molecules that are protective against this toxicity; examples are frankobactin, which was obtained from the actinobacterium *Frankia* sp. CH37, and a variety of metallothioneins isolated from diverse bacteria, including the cyanobacterium *Synechococcus* PCC 7942.[14,15] The capacity of cyanobacteria to resist metal toxicity has been observed previously, especially for copper and iron.[16–18]

In the present work, we characterize the intriguingly complex structures of leptochelins A-C as novel metallophores produced by three strains of *Leptothoe* cyanobacteria. The leptochelins collectively possess remarkable selectivity for copper chelation and are proposed to allow *Leptothoe* species to persist in conditions of elevated copper ion concentrations. At the same time, promiscuous metal binding by leptochelins could enable *Leptothoe* cyanobacteria to thrive in conditions where essential metal cofactors are limiting. These unique leptochelin molecules are also potently toxic to mammalian cancer cells, and thus may have therapeutic potential.

## Results and Discussion

### Geographically Dispersed Collections of *Leptothoe*

Living cultures of three strains of cyanobacteria of the genus *Leptothoe* (**SI Table 1**) were collected over a 25-year period from geographically dispersed regions [Sulawesi, Indonesia (in 1994), El Aruk, Egypt (in 2007), and Baía das Gatas, Republic of Cape Verde (in 2018)], and were found to produce the same series of structurally unique natural products, named here as the leptochelins. The first two of these collections were subtidal and made using SCUBA whereas the latter was collected by hand from shallow tide pools. Despite the geographical dispersion and littoral versus benthic nature of the collection sites, the cyanobacterial cultures all produced the same or a very similar suite of leptochelin natural products, suggesting that their production is highly conserved, and the compounds possess important adaptive roles.

All three strains of *Leptothoe* examined in this study possessed thin filaments, were unbranched and non-heterocystous, and formed a pinkish-red to brownish-red biofilm in laboratory cultures, consistent with the description by Konstantinou et al. 2019.[19] Phylogenetic analysis of the 16S rRNA gene (**Figure 1**) revealed that these three cyanobacterial strains all belong to the recently described genus *Leptothoe*[19] and share 99.9%-100% sequence identity of this key phylogenetic gene (**SI Table 2**). This genus comprises strains obtained from marine environments (supported by a bootstrap value of 100%) around the world. More distantly, the phylogenetic clade includes the genera *Rhodoploca*, *Salileptolyngbya*, *Cymatolege*, *Halomicronema* and *Nodosilinea,* which have less than 94% 16S rRNA sequence similarity to *Leptothoe*.

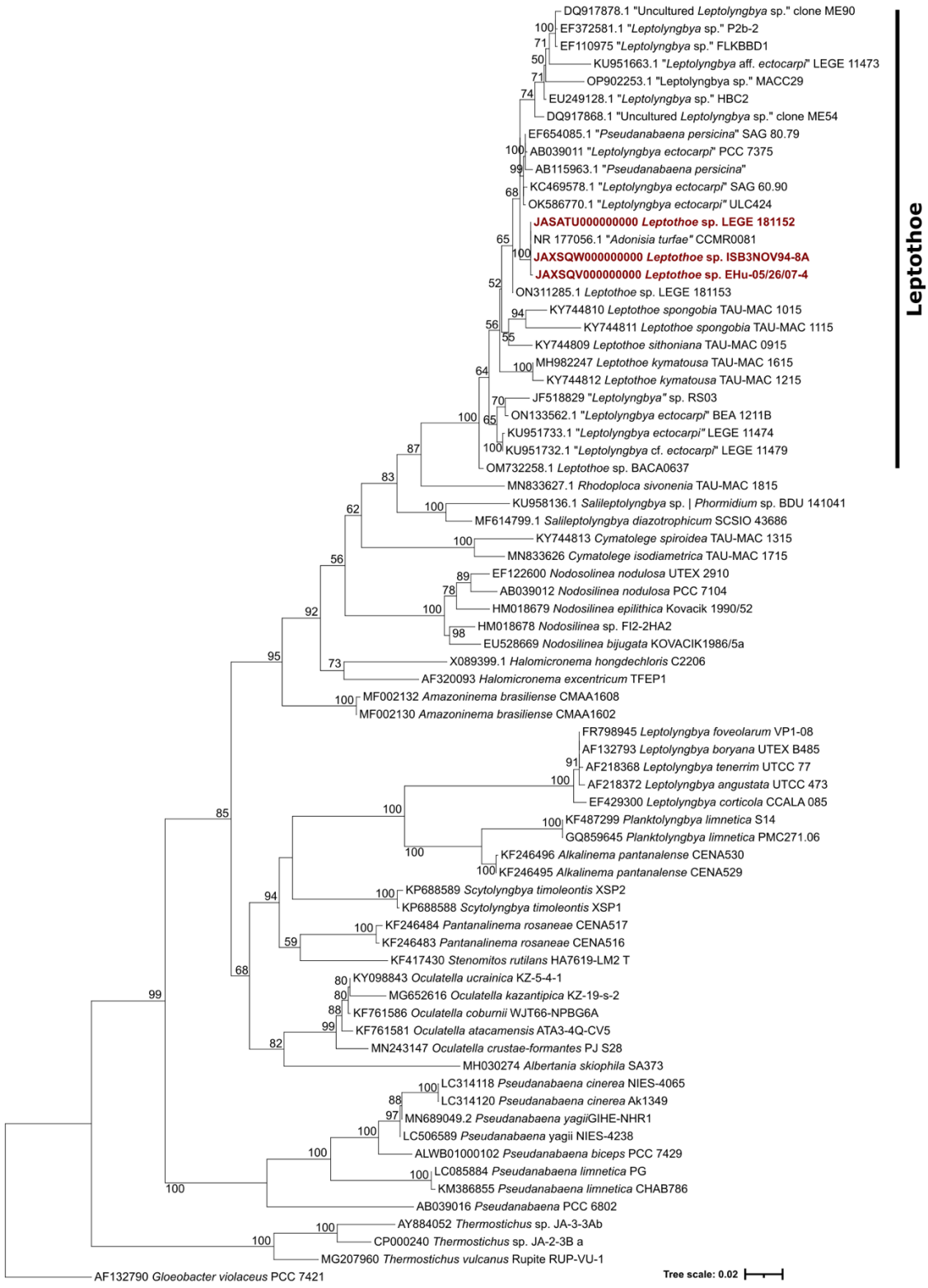



**Figure 1.** Maximum likelihood (ML) phylogenetic tree based on 73 16S rRNA gene sequences of cyanobacterial strains belonging to the orders Gloeobacterales, Thermostichales, Pseudanabaenales, Oculatellales, Nodosilineales, Leptolyngbyales and Synechococcales. *Gloeobacter violaceus* PCC 7421 was used as the outgroup. Phylogenetic positions of *Leptothoe* sp. LEGE 181152, *Leptothoe* sp. ISB3NOV94-8A and *Leptothoe* sp. EHU-05/26/07-4 are indicated in bolded red font. Bootstrap values over 50% are indicated at the nodes. Names of strains in quotation marks correspond to GenBank labels.

**Isolation and planar structures of leptochelins A-C (1-3).**

Interestingly, laboratory cultures of all three *Leptothoe* strains produced leptochelins A-C (**1**-**3**). Although they were cultured in three different laboratories, the common general strategy for isolation of the leptochelins included extraction using $CH_2Cl_2$:MeOH, followed by an initial separation using normal phase chromatography on silica gel as the stationary phase, followed by repeated iterations of reversed phase High Performance Liquid Chromatography (HPLC) using $C_{18}$ as the stationary phase. Zinc-bound leptochelins A–C were also isolated in parallel with leptochelins, especially when MS-grade water was not used in the HPLC purifications.

Pure leptochelin A (**1**), isolated as an optically active off-white amorphous solid $[\alpha]_D^{24}$ -89.1 (*c* 0.1, $CH_2Cl_2$), possessed a monoisotopic mass of *m/z* 895.0790 $[M+H]^+$ by high resolution ESIMS (**Figure 4A**). While the isotope pattern suggested two halogen atoms (either 2 Br, or 1 Br and 1 Cl), the mass accuracy of the Thermo QExactive (~1 ppm) did not allow for unambiguous calculation of the molecular formula. However, a 21 Tesla Fourier transform ion cyclotron resonance (21T FTICR) mass spectrometer permitted resolution of the isotopic fine structure and resultant high-confidence elemental composition for **1** as $C_{35}H_{41}Br_2N_6O_8S_2^+$, matching the experimentally observed mass with 0.1 ppm mass accuracy. This molecular formula indicated that leptochelin A (**1**) possessed 18 degrees of unsaturation, and by NMR analysis these were present as two substituted phenyl rings and six ester/amide/carboxyl-type carbons with chemical shifts between $\delta_C$ 165-185, leaving the presence of four additional rings to be determined. HMBC correlations established one of the rings as a trisubstituted epoxide ($\delta_C$ 61.5 qC and 64.3 CH). A combination of NMR chemical shift and 2D NMR data were subsequently used to construct 6 partial structures, A-F (**Figure 3**), which accounted for all atoms in the molecular formula as well as the three remaining rings.

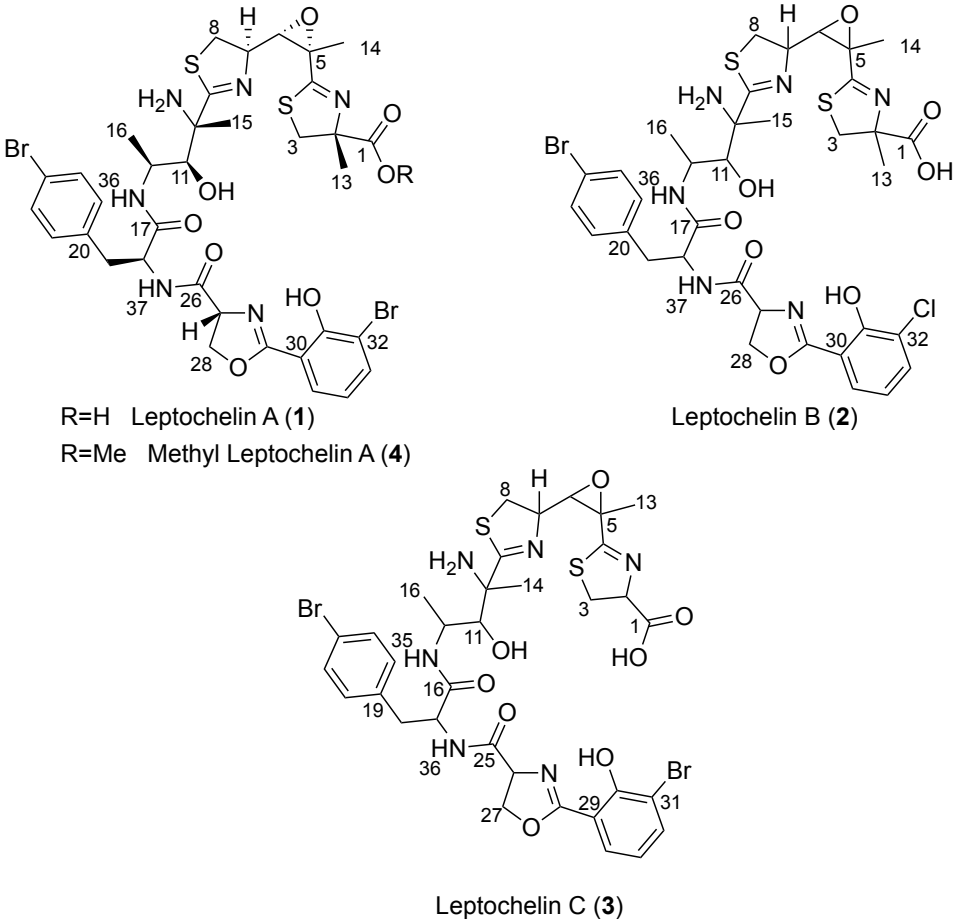



**Figure 2.** Structures of leptochelins A-C (**1**-**3**). Based on comparisons of $^{13}C$ NMR chemical shifts (**SI Tables 5-6**, **SI Figure 42**), the absolute configuration for leptochelin B (**2**) and leptochelin C (**3**) is proposed to be the same as leptochelin A (**1**) at comparable centers. The configuration at C2 of compound **3**, predicted to result from incorporation of L-cysteine, likely results in *R* configuration at this position.

Partial structure A (**Figure 3**) possessed a distinctive singlet methyl group with a shift of $\delta_H$ 1.46 that exhibited $^1$H-$^{13}$C HMBC correlations to a non-protonated carbon signal at $\delta_C$ 84.3, an isolated methylene at $\delta_C$ 42.1 and a deshielded carboxylic acid carbon signal at $\delta_C$ 178.9. The terminal nature of this carboxyl group was demonstrated by treatment of compound **1** with diazomethane ($CH_2N_2$), resulting in the observation of a new dibromo isotopic cluster $[M]^+/[M+2]^+/[M+4]^+$ at *m/z* 909.19/911.17/913.08 displaying a series of fragments diagnostic for methylation at the carboxyl terminus of compound **1** to yield compound **4** (**SI Table 8** and **SI Figure 4**). The methylene chemical shift ($\delta_C$ 42.1) was consistent with its attachment to a sulfur atom. This methylene also showed HMBC correlations to the non-protonated carbon signal $\delta_C$ 84.3 and a deshielded carbon signal ($\delta_C$ 176.2) that was assigned as a non-protonated carbon attached by a single bond to sulfur, a double bond to nitrogen, and a final single bond to a distal carbon atom. As a result, this constellation of atoms was formulated as a thiazoline ring with $\alpha$-methyl and $\alpha$-carboxylic acid functionalities (**Figure 3A**).

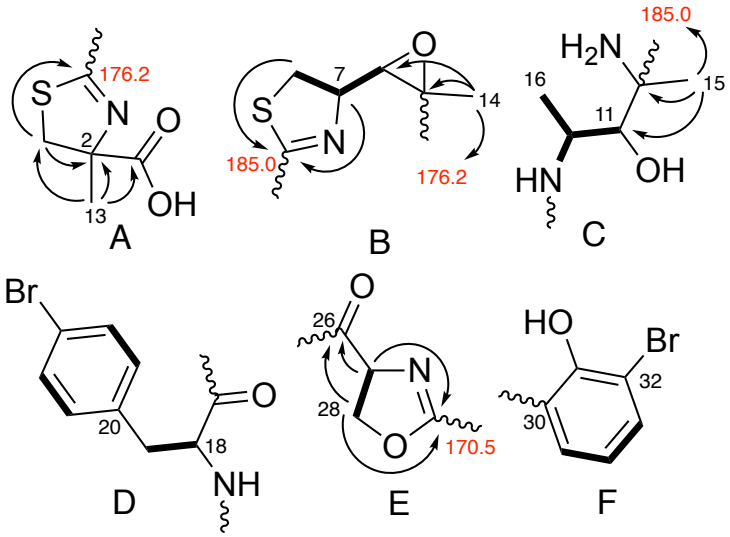



**Figure 3.** Partial structures A-F determined by $^1$H, $^{13}$C and 2D NMR experiments for leptochelin A (**1**). Bolded bonds show connectivity deduced by $^1$H-$^1$H COSY and arrows depict selected $^1$H-$^{13}$C HMBC correlations.

In partial structure B (**Figure 3B**), a -$CH_2$-CH-CH- motif was defined by $^1$H-$^1$H COSY correlations. The methylene group possessed $^1$H and $^{13}$C NMR shifts consistent with an attached sulfur atom while the central methine had shifts indicative of an attached nitrogen atom. HMBC correlations from H-7 and H-8b to a downfield carbonyl-like shifted resonance (C-9) identified this as a disubstituted thiazoline ring. The terminal methine of this motif was the protonated component of the trisubstituted epoxide described above. $^1$H–$^{13}$C HMBC correlations from $H_3$-14 to C-5 and C-6 indicated that the non-protonated carbon of the epoxide bore a methyl group ($\delta_H$ 1.53).

The delineation of partial structure C (**Figure 3C**) commenced with a methyl doublet resonance coupled to a mid-field methine multiplet at $\delta_H$ 4.51 that was HSQC-correlated with a $^{13}$C resonance at $\delta_C$ 50.2. These mid-field chemical shifts implied a N-bound methine that was confirmed by observing homonuclear coupling between the methine multiplet and an NH doublet at $\delta_H$ 10.15 ($^3J_{H,H}$ = 7.7 Hz). The $\delta_H$ 4.51 multiplet was also coupled to an oxymethine $^1$H doublet ($\delta_H$ 3.85, $\delta_C$ 77.9). A methyl singlet at $\delta_H$ 1.53 displayed HMBC correlations to the oxymethine $^{13}$C resonance, as well as to two carbon atoms without attached hydrogen atoms, one resonating at $\delta_C$ 185.0 (C-9, part of the thiazoline ring system in partial structure B discussed above) and a second mid-field $^{13}$C signal at $\delta_C$ 61.7, consistent with its attachment to another nitrogen atom.

One of the two phenyl rings was 1,4-disubstituted as determined by a pair of 2H doublets at $\delta_H$ 7.23 (H-21/25) and $\delta_H$ 7.46 (H-22/24). A benzylic $CH_2$ moiety was indicated by reciprocal HMBC correlations between the diastereotopic $H_2$-19 signals and C-21/25 and H-21/25 and C-19. The *para* substituent was assigned as Br by virtue of a characteristically shielded $^{13}$C chemical shift of $\delta_C$ 121.0. Homonuclear coupling was also observed between $H_2$-19 and the adjacent H-18 methine ($\delta_H$ 4.40, $\delta_C$ 58.0), for which the mid-field $^{13}$C NMR shift, and an HMBC correlation from H-18 to an amide type carbonyl at $\delta_C$ 175.2 (C-17), indicated an alpha amino acid motif. Taken together, these moieties constituted a *para*-bromo-phenylalanine residue (partial structure D, **Figure 3D**).

Partial structure E possessed a motif of -O-$CH_2$-CH-N-, in which the oxymethylene was supported by a notably deshielded methylene $^{13}$C NMR shift ($CH_2$-28, $\delta_C$ 69.0, $\delta_H$ 4.33 and 4.45), while the comparable methine chemical shifts of CH-27 were consistent with an attached nitrogen atom ($\delta_C$ 67.1, $\delta_H$ 4.56). HMBC correlations from both H-27 and $H_2$-28 led to assignment

of an amide-type carbonyl C-26 ($\delta_C$ 172.8) as a substituent on C-27. The H-27 and H-28 protons also showed HMBC correlations to a second deshielded signal at $\delta_C$ 170.5, which allowed formulation of this oxazoline partial structure by comparison with literature values. [12,20]

For the second phenyl ring (partial structure F, **Figure 3F**), a 1,2,3-trisubstitution pattern was interpreted from the mutually coupled $^1$H NMR spin system with relatively large *ortho*-couplings between protons (7.8 Hz). Based on $^{13}$C NMR chemical shifts, hydroxy ($\delta_C$ 165.5), bromo ($\delta_C$ 109.7) and carbon ($\delta_C$ 118.4) substituents were inferred. HMBC correlations from H-34 to C-30/C-32 and H-33 to C-31/C-35 placed the oxygenated carbon between the brominated and carbon-substituted positions (**Table 1**).

Partial structures A-F (**Figure 3**) were assembled into an overall linear structure using a combination of NMR (**Table 1**) and MS$^2$ data (**Figure 4**). Partial structures A and B were connected by observation of HMBC correlations between $H_3$-14 and the deshielded imino carbon C-4. In turn, $H_3$-15 in partial structure C was correlated to the downfield thiazoline carbon of partial structure B, connecting these two through a C-9/C-10 bond. ROESY correlations between highly deshielded NH-36 to several resonances (H-18, H-19a, H-19b, H-21/25) associated with the bromophenylalanine residue of partial structure D connected these moieties. Partial structures D and E could be connected by HMBC correlations between $\alpha$-proton H-18 and carbonyl C-26, as well as ROESY correlations between H-27 and NH-37. Finally, connection of remaining partial structure F was enabled by an HMBC correlation observed between H-35 and C-29.

A range of MS$^2$ and MS$^3$ experiments were obtained from higher-energy C-trap dissociation (HCD) using a QExactive orbitrap MS; collision-induced dissociation (CID) with a 21T FT-ICR; and CID and HCD with an Orbitrap Elite MS (**Figure 4D**). With the precursor molecular formulas in hand, the mass differences in the MS$^2$ spectra were calculated and MS$^3$ fragmentation series were developed (**SI Figure 12**) with calculated product ion molecular formulas. Using SIRIUS and MSFinder, we constructed fragmentation trees for the different activation methods and derivatives (**SI Figures 2 and 3**). Characteristic b-ion fragments were observed for the brominated salicylate unit. An internal $C_5H_{10}NO^+$ fragment corresponded to neutral losses between major b- and y-ion fragments, indicating that this residue was located between the bromophenylalanine and thiazole moieties. With most ion molecular formulas identified, a fragmentation series of b- and y-ions was proposed (**Figure 4E**). Notably, a *diaminopropionyl* moiety in the fragmentation series was incongruent with the NMR derived structure. Guided by the NMR data, the diaminopropionyl moiety is proposed to derive from the oxazoline residue and the C-10 amino group. We hypothesize that a gas-phase transamination rearrangement between the C-10 amino group and the oxazoline occurs prior to MS$^2$ and MS$^3$ fragmentation (**Figure 4E**, products 3-8), which corresponds to the downstream b- and y-ion fragmentation series. Overall, this analysis provides additional support for the identity and sequence of residues that comprise leptochelin A (**1**).

Leptochelin B (**2**) analyzed for *m/z* 851.1294 for the $[M + H]^+$, and by 21T FT-ICR MS analysis, the nature of the halogen atoms was unambiguously resolved (**Figure 4B**), yielding a molecular formula of $C_{35}H_{41}BrClN_6O_8S_2$. NMR analysis (**SI Tables 3-6**) of **2** yielded nearly identical spectra to **1**, with small differences observed for the halogenated salicylate group (e.g., H-33 to H-35). From these data along with the comparable molecular formula to leptochelin A, we reasoned that the bromosalicylate group in **1** was replaced by a chlorosalicylate residue in **2**. Again, using a range of MS$^2$ and MS$^3$ experiments as described above (**Figure 4C**), a fragmentation pathway highly comparable to that for **1** was observed for **2** (**SI Figures 39** and **40**). Notably, the sequence of fragments for **2** differed from the sequence for **1** by -44.451 amu, the mass difference between $^{35}$Cl and $^{79}$Br. Thus, leptochelin B (**2**) was assigned as the chlorosalicylate equivalent of leptochelin A (**1**), and by its co-occurrence, we propose it to be of the same absolute configuration, as described below.

Leptochelin C (**3**) displayed a prominent $[M + H]^+$ ion peak by (+)-HRMS at *m/z* 881.0643, indicative of a molecular formula of $C_{34}H_{39}Br_2N_6O_8S_2$. Since the NMR data for compound **3** was very similar to those of **1** and **2**, an analogous elucidation strategy was used to assign its structure (**SI Table 3-6**). The mass difference of 14.01 Da between compounds **1** and **3** suggested the absence of a methyl group in leptochelin C (**3**). This variation was confirmed in the terminal thiazoline of **3**, for which the $\alpha$-carbon C-2 ($\delta_C$ 78.4), was protonated ($\delta_H$ 5.16, dd, $J$ = 11.7, 9.3 Hz) instead of being fully substituted, and resonating at a more shielded chemical shift than C-2 in **1** ($\Delta\delta_C$ = 6 ppm). The $^1$H-$^1$H COSY correlations between H-2 and the diastereotopic protons H-3a/H-3b ($\delta_H$ 3.67; 3.83), together with HMBC correlations to the C-1 carboxylic acid carbon ($\delta_C$ 176.7) and the methylene carbon C-3 ($\delta_C$ 35.0) confirmed the proposed planar structure of leptochelin C (**3**).

**Table 1.** NMR Data for leptochelin A (**1**) ($CDCl_3$, 400 MHz). ROESY spectrum was obtained in $CDCl_3$, on a 500 MHz instrument; Note: * tentative due to signal overlap and wk indicates a weak, but notable correlation.

| *Position* | $\delta_C$, *type* | $\delta_H$, *mult* (*J* in Hz) | *HMBC* | *ROESY* |
|---|---|---|---|---|
| *1* | 178.9, C | | | |
| *2* | 84.3, C | | | |
| *3a* | 42.1, $CH_2$ | 3.43, d (11.7) | 2, 4, 13 | 3b, 13 |
| *3b* | | 3.58, d (11.7) | 1, 2, 13 | 3a |
| *4* | 176.2, C | | | |
| *5* | 61.5, C | | | |
| *6* | 64.3, CH | 3.71, d (8.4) | 7, 8 | 14* |
| *7* | 75.0, CH | 4.00, dt (9.9, 8.4) | 6, 9 | 8a, 8b, 16 |
| *8a* | 38.0, $CH_2$ | 3.78, dd (11.3, 8.4) | 6, 7 | 7 |
| *8b* | | 3.92, dd (11.3, 9.9) | 6, 7, 9 | 7 |
| *9* | 185.0, C | | | |
| *10* | 61.7, C | | | |
| *11* | 77.9, CH | 3.85, bs | 16 | 12, 15*, 16 |
| *12* | 50.2, CH | 4.51, dd (10.7, 7.5) | 10, 11, 16 | 11, 16, 36 |
| *13* | 22.6, $CH_3$ | 1.46, s | 1, 2, 3 | 3a, 27, 28b |
| *14* | 20.9, $CH_3$ | 1.53, s | 4, 5, 6 | 6 |
| *15* | 30.8, $CH_3$ | 1.53, s | 9, 10, 11 | 11 |
| *16* | 15.5, $CH_3$ | 1.15, d (7.5) | 11, 12 | 7, 11, 12, 36 |
| *17* | 175.2, C | | | |
| *18* | 58.0, CH | 4.40, ddd (11.0, 7.0, 3.6) | 17, 19, 26 | 19a, 19b, 21/25, 36, 37 wk |
| *19a* | 37.0, $CH_2$ | 2.92, dd (14.2, 11.5) | 17, 18, 20, 21/25 | 21/25, 36, 37 |
| *19b* | | 3.24, dd (14.2, 3.6) | 18, 20, 21/25 | 18, 21/25, 37 wk |
| *20* | 135.6, C | | | |
| *21/25* | 130.9, CH | 7.23, d (8.3) | 19, 22/24, 23 | 18, 19a, 19b, 22/24, 37 |
| *22/24* | 131.9, CH | 7.46, d (8.3) | 20, 21/25,23 | 21/25 |
| *23* | 121.0, C | | | |
| *26* | 172.8, C | | | |
| *27* | 67.1, CH | 4.56, dd (9.5, 8.3) | 26, 28, 29 | 13, 28a, 28b, 37 |
| *28a* | 69.0, $CH_2$ | 4.33, t (9.0) | 26, 27, 29 | 27 |
| *28b* | | 4.45, t (9.0) | 26, 27, 29 | 13 wk, 37 wk |
| *29* | 170.5, C | | | |
| *30* | 118.4, C | | | |
| *31* | 165.5, C | | | |
| *32* | 109.7, C | | | |
| *33* | 137.6, CH | 7.58, dd (7.8, 1.9) | 30, 31, 35 | 34 |
| *34* | 112.8, CH | 6.29, t (7.8) | 30, 32 | 33, 35 |
| *35* | 130.4, CH | 7.61, dd (7.8, 1.9) | 29, 33 | 34 |
| *##* | 39.7, $NH_2$ | 1.52 | | |
| | | 3.29 | | |
| *36* | NH | 10.15, d (7.7) | 15, 12 | 12, 16, 19a, 37, 18 wk (3.51 exchangeable) |
| *37* | NH | 8.26, d (7.0) | | 18, 19a, 21/25, 27, 36, 28b wk, 19b wk |

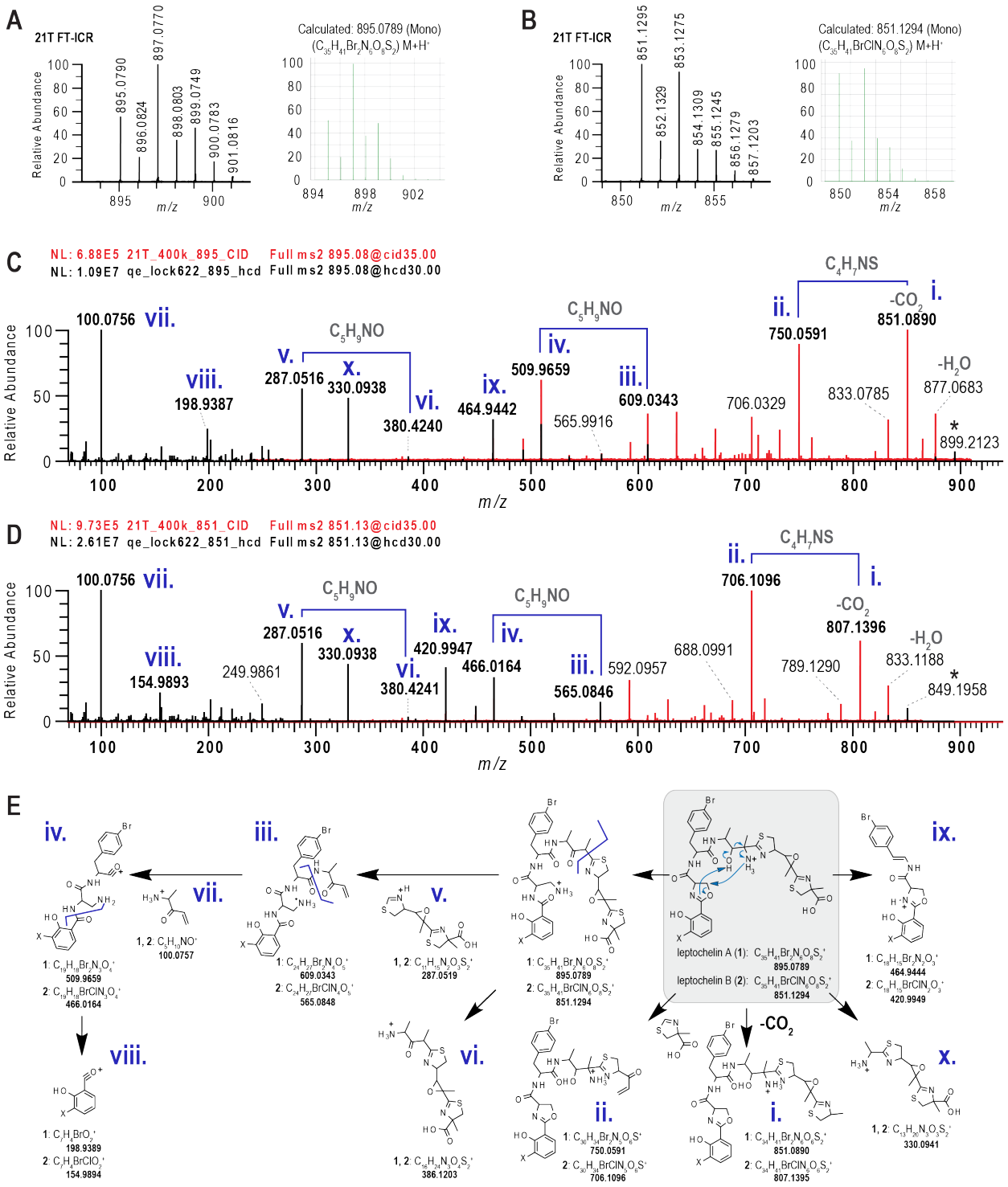
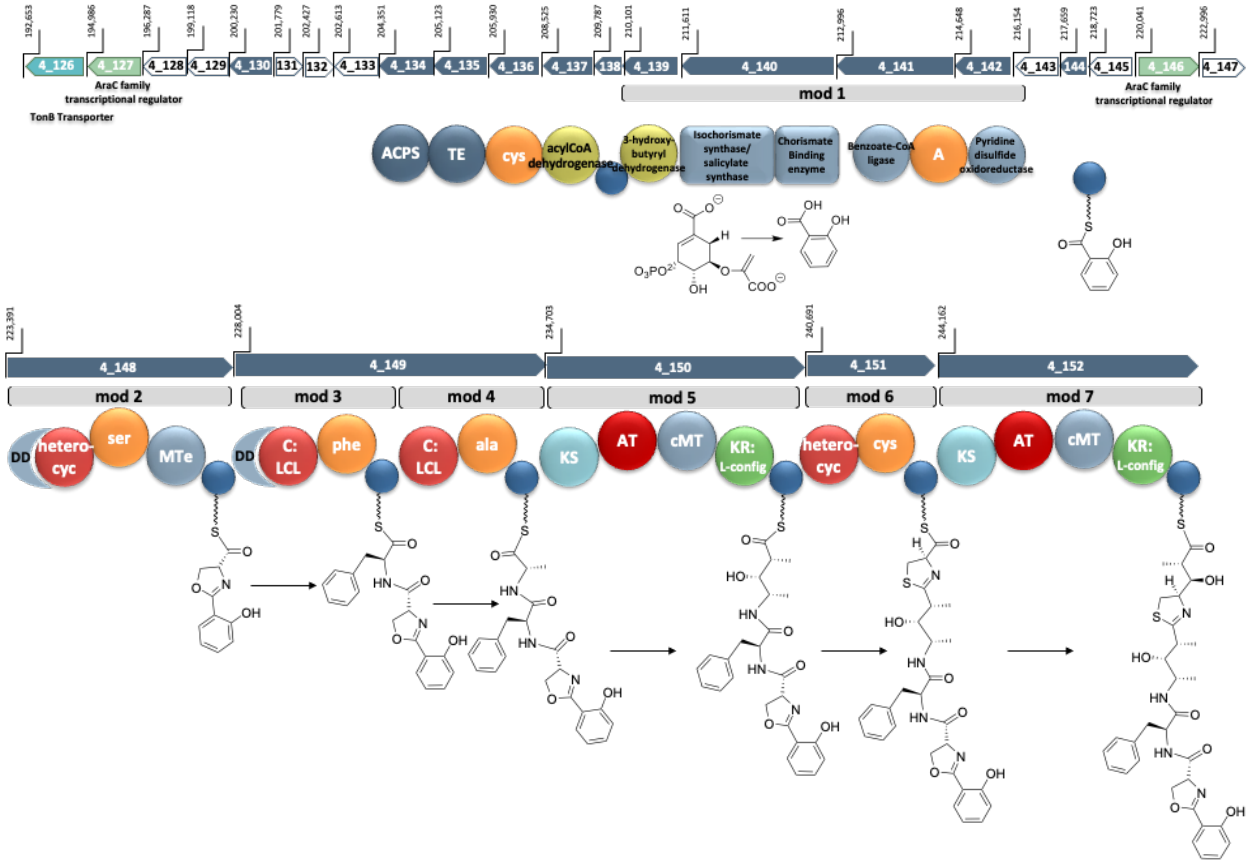



**Figure 4.** Exact mass, isotope pattern and $MS^2$ fragmentation patterns for leptochelins A (**1**) and B (**2**). (A, B) 21T FT-ICR spectra showing the isotopic fine structure for the monoisotopic molecular ions of leptochelin A (**1**) and leptochelin B (**2**) compared with the calculated values. (C) $MS^2$ fragments deriving from fragment mass *m/z* 851 that results from neutral loss of $CO_2$ from leptochelin A (**1**) with fragments from HCD in black and CID in red. Roman numerals refer to fragmentation tree depicted in 4E. (D) $MS^2$ fragments deriving from fragment mass *m/z* 807 that results from neutral loss of $CO_2$ from leptochelin B (**2**) with fragments from HCD in black and CID in red. Roman numerals refer to fragmentation tree depicted in 4E. (E) Depiction of proposed fragmentations for leptochelins A (**1**) and B (**2**). Proposed mechanisms for fragmentation of leptochelin A (**1**) and leptochelin B (**2**) are indicated by the blue arrows within the grey box in Panel E.

**The Leptochelin (*lec*) Biosynthetic Gene Cluster, Biosynthesis, and Genomic Insights on Configuration.**

To further confirm the structure of leptochelin A and explore its biosynthetic assembly process, we set out to identify the candidate genes responsible for its production. The biosynthetic gene clusters (BGCs) responsible for siderophore and other metallophore production often arise from non-ribosomal peptide synthetases (NRPSs) and NRPS-independent siderophore synthetases (NISs), and these are frequently associated with distinct transcriptional and other regulatory genes, such as TonB transporters and AraC transcriptional regulators.[10,21–23] The isolated DNA from *Leptothoe* sp. ISB3NOV94-8A was sequenced using Nanopore and Illumina technologies, and a hybrid assembly resulted in eight contigs with a total length of 8.51 Mbp; the longest contig was 7,062,414 bp and the GC content was 47.4%. The isolated DNA from *Leptothoe* sp. EHU-05/26/07-4 was sequenced using Nanopore technology, and the resulting 8.85 Mbp assembly was comprised of 11 contigs and the GC content was 47.4%. The final genome, *Leptothoe* sp. LEGE181152, also had a GC content of 47.4%, and was sequenced using Illumina technology, resulting in an 8.36 Mbp assembly comprised of 68 contigs (**SI Table 7**). Genomic signatures of siderophore production as well as a retrobiosynthetic analysis were used to screen the three cyanobacterial genomes for a candidate leptochelin BGC (*lec* BGC).

Of the 18 BGCs identified in the genome assembly for *Leptothoe* sp. ISB3NOV94-8A, three were NRPS-PKS hybrid systems and only one contained the requisite isochorismate synthase needed for biosynthesis of the salicylate group found in the leptochelins. This hybrid NRPS-PKS pathway (**Figure 5**) had excellent congruence with the retrobiosynthetic scheme for leptochelin A (**SI Figure 43**) as well as regulation and transport genetic signatures for siderophore production. A nearly identical hybrid PKS-NRPS BGC was present in the genomes of both *Leptothoe* sp. EHU-05/26/07-4 and *Leptothoe* sp. LEGE181152, and all three showed remarkable synteny and a high degree of nucleotide identity for each gene (98-100%). This is notable as the three collections were obtained from geographically dispersed regions (Celebes Sea, South Pacific Ocean; Red Sea, Indian Ocean; Republic of Cabo Verde, Eastern Mid-Atlantic Ocean) over a 30-year period (**Figure 6, SI Table 9**). Such a profound preservation of the BGC and secondary metabolite production despite the divergent temporal and geographic distribution of the three collections indicates that the leptochelins must confer a strong adaptive advantage. Additionally, modules 5 and 7 within the cluster appear to be the result of a gene duplication event with neofunctionalization (amino acid percent identity of 48% and similarity of 63%). Also present in the *lec* BGC are two AraC family transcriptional regulators and a TonB-dependent receptor transporter domain (PF07715.17), which provide insight into how *the lec* BGC is transcriptionally regulated as well as how the expressed molecule is transported through the cell membrane.[21–23]



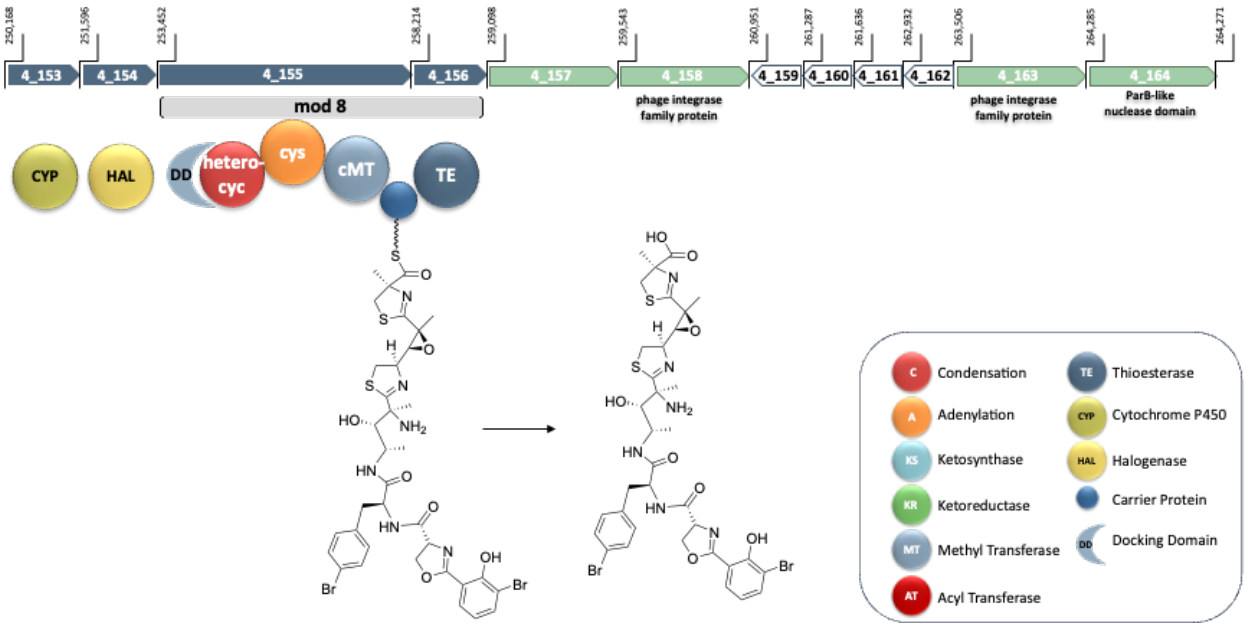



**Figure 5**. The putative *lec* BGC for the biosynthesis of the leptochelins. The 71.6 kbp gene cluster (54 kbp of which represents core biosynthetic genes) is an NRPS-PKS hybrid system that begins with a cassette containing an isochorismate synthase followed by a series of NRPS and PKS modules to form the elongated chain. A proposed *trans*-acting halogenase is likely responsible for at least one of the halogenations seen on the aromatic rings. Additionally, a proposed *trans*-acting cytochrome P450 is likely responsible for one or two epoxidations (C-5–C-6 and possibly C-10–C-11). A TonB-dependent transporter associated with metallophore-specific transport and several regulatory proteins are encoded. Also encoded are RTX calcium-binding nonapeptide repeats (PF00353.21) (ctg4_128), a sulfite exporter with homology to TauE and SafE (PF01925.21) (ctg4_129), an oxoacyl-ACP synthase III (ctg4_130), chlorophyllase (ctg4_133), along with several hypothetical proteins which fall within the designated gene neighborhood. In the schematic above, dark blue arrows represent core biosynthetic genes. Green arrows represent genes with ancillary functions (e.g., transcriptional regulation). White arrows represent genes that encode hypothetical proteins. The light blue arrow represents a gene associated with transport, in this case the TonB-dependent transporter.

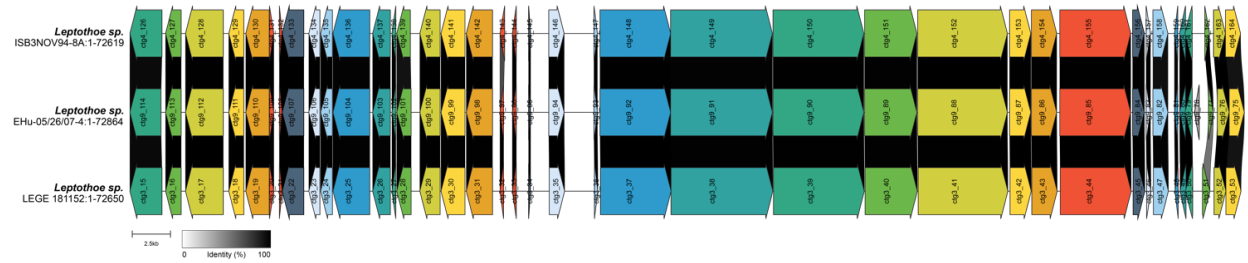



**Figure 6.** Clinker analysis of the putative *lec* BGC from the assembled genomes of the three leptochelin producers showing high synteny and near 100% identity between the BGCs. Similarities between genes are indicated by the shaded links between the genes, where the gradient ranges from 0% identity (white) to 100% identity (black). The figure shows only links for genes with >50% identity, which in the case of the three putative leptochelin BGCs is at least 94% for all of the core biosynthetic and tailoring genes (**SI Table 9**). Intergenic regions are shown by lines without arrows and without links.

The first complete biosynthetic module of the *lec* BGC (**Figure 5**, module 1) encodes a cassette consistent with the formation of salicylic acid, and includes an isochorismate synthase, chorismate binding enzyme, and salicylate synthase, along with a benzoate-CoA ligase that is proposed to tether the salicylate to a carrier protein via a phosphopantetheinyl arm. The second AraC family transcriptional regulator is then encoded before a series of modules responsible for production of the remaining sections of leptochelin A (**1**). In module 2, an adenylation domain specific for serine (cysteine) is followed by a C-methyl transferase (cMT) that likely serves as a methyltransferase-like epimerase (MTe)(see below).[24,25] This is followed by a heterocyclic condensation domain that forms the oxazoline ring from the tethered serine and salicylate residues. A hydrophobic amino acid, phenylalanine, is subsequently incorporated into module 3. It is uncertain if this is incorporated as bromophenylalanine, or if the bromination occurs later in the assembly process. Next, an alanine subunit is incorporated followed by elongation with acetate from an PKS module. This latter module contains a C-methyltransferase domain, placing a methyl group at the $\alpha$-position between the two transient carbonyl functionalities. The $\beta$-carbonyl is next reduced to a hydroxy group by a ketoreductase, which may be transformed to an intermediate 10,11-epoxy species that is subsequently opened by an ammonolysis reaction.[26] In module 6, an adenylation domain with specificity for cysteine (**SI Table 10**) is paired with a heterocyclic condensation domain, producing a thiazoline ring. This is followed by another PKS elongation in module 7 that also adds a methyl group to the $\alpha$-position from a C-methyl transferase. This module again contains a ketoreductase that

results in a C-6 hydroxy group, and then this is subsequently transformed into a 5,6-epoxide, possibly by action of the adjacent CyP450 gene. This CyP450 may be responsible for both the enzymatic epoxidation proposed between C-10 and C-11 and the observed epoxidation between C-5 and C-6. A second and possibly *trans* acting tryptophan halogenase is encoded adjacent to the CyP450. This may be responsible for both aromatic ring halogenations[27,28] (the salicylic acid monomer and the benzyl ring of phenylalanine) resulting in dibromination for leptochelins A (**1**) and C (**3**), and in monobromination and monochlorination for leptochelin B (**2**). Alternatively, a second currently unrecognized *trans* acting halogenase may be present in the genome. The final elongation step is encoded by NRPS module 8 in which cysteine is incorporated followed by $\alpha$-methylation from SAM and heterocyclization to form a methyl- and carboxyl-substituted thiazoline ring in leptochelins A (**1**) and B (**2**). These core biosynthetic genes terminate with a thioesterase (TE) domain that is proposed to hydrolytically liberate the completed leptochelin molecule. A phylogenetic analysis of hydrolyzing versus cyclizing TEs from cyanobacteria indicates that the leptochelin TE catalyzes hydrolysis, consistent with the results of $CH_2N_2$ treatment of leptochelin A (**1**) to produce methyl leptochelin A (**4**), and providing another level of confirmation that the leptochelins have an overall linear architecture (**SI Figure 44**).

**Stereochemical Insights from the putative leptochelin BGC and Marfey's Analysis.**

Informatic analyses of the biosynthetic gene cluster provided stereochemical insights at 7 of the 9 asymmetric centers in leptochelin A (**1**). In module 2, the first amino acid monomer, L-serine, is selected and undergoes heterocyclization and epimerization from a MTe domain that would result in *R*-configuration at C-27.[24,25] This is consistent with the Marfey's analysis that established the presence of D-serine in the hydrolysis product of **1** (**SI Figure 45**). The condensation domains in modules 3 and 4 specify for the incorporation of L-phenylalanine and L-alanine, corresponding with *S* configurations at C-12 and C-18. Marfey's analysis of hydrolyzed **1**, in comparison with authentic standards, confirmed the *S* configuration of the bromophenylalanine residue (**SI Figure 46**). The ketoreductases in modules 5 and 7 both have active site predictions for L-hydroxy products at C-11 and C-6, and transiently, an L-methyl at C-10 with the KR in module 5 identified as a C2-Type.[29] The result is an *S* configuration at C-11, and predicting that there is no alteration of the C-6 configuration upon conversion to a 5,6-epoxide by a *trans*-acting CypP450 enzyme, an *S*-configuration at C-6. The relative configuration of C-6 and the fully substituted C-5 center was demonstrated by ROESY analysis to be *cis*, and thus the C-5 center is also predicted to be of *S* configuration. There are no epimerases present in the sequences for modules 6 or 8, and the adenylation domain specificities have consensus predictions for L-cysteine in both cases, resulting in an initial assignment of *R* configurations for C-7 and C-2; however, subsequent methylation at C-2 by a cMT make its absolute configuration uncertain from this biosynthetic analysis. Marfey's analysis after sequential ozonolysis (**SI Figure 47**) and hydrolysis of **1**, revealed the presence of 2-Me-L-cysteic acid in the reaction product, supporting an R configuration for C-2, and implying that the C-methylation event occurs with retention of configuration. The *S* configuration at C-10 was inferred from observation of NOE correlations between $H_3$-15 /H-11, $H_3$-16 /H-8, $H_3$-16 /H-11, H-12/H-11, combined with an intense HMBC correlation from H-12 to C-10 (**SI Figure 48**). These various chemical, spectroscopic and bioinformatic analyses support a 2*R*,5*S*,6*S*,7*R*,10*S*,11*S*,12*S*,18*S*,27*R* configuration for leptochelin A (**1**).

## Discovery of chelating properties of leptochelin A (1) and subsequent metal-based culture studies to increase leptochelin production and measure strain resilience.

The leptochelins were isolated over a span of 25 years from three geographically distinct regions. While leptochelin A (**1**) was first isolated and identified around 2001 from the Indonesian strain (*Leptothoe* sp. ISB3NOV94-8A), it initially appeared absent from the Red Sea strain (*Leptothoe* sp. EHU-05/26/07-4). Instead, the Red Sea strain contained only the brominated macrolide, phormidolide,[30–32] which was also isolated from *Leptothoe* sp. ISB3NOV94-8A. However, a careful inspection of the (+)-HRESIMS isotope pattern for a minor metabolite (*m/z* 956.9924) from the Red Sea isolate led to the discovery of a zinc complexed form of leptochelin A. A notable difference between the culture media for these two samples was the inclusion of an uncharacterized soil extract that was used in the Red Sea strain culture medium to enrich and potentially induce the production of secondary metabolites.[33] Based on these observations, in addition to the presence of iron, copper and cobalt-complexed leptochelins in subsequent LCMS analyses, several metal-based culture experiments were designed to explore the role of these metabolites as metallophores.

In an iron-depleted medium, the *Leptothoe* sp. ISB3NOV94-8A strain exhibited a substantial enhancement in production of the leptochelin A as detected by LCMS analysis. Interestingly, in medium possessing a range of copper concentrations (0 to 1,000 ppb) compared with standard SWBG11 medium (20 ppb)[34] or native coastal seawater (2 ppb and as high as 25 ppb for anthropogenically affected sites), no enhancement in leptochelin production was noted. Alternatively, as leptochelin has bromine incorporated into its structure, we also evaluated the effect of bromide concentration in the growth medium on leptochelin production. When the normal SWBG11 medium with ~5.4 ppm of bromide (with normal seawater concentration ~65 ppm) was enriched to 0.5 g/L (500 ppm) and 1.0 g/L (1,000 ppm), this also led to a roughly fourfold enhancement in the production of leptochelin A from 4 mg/L to 15 mg/L and 16 mg/L, respectively.

A native metabolomics analysis[35,36] and post-LC infusion of metals was used to explore the binding of metals to leptochelin A (**Figure 7A**) as well as leptochelin B and a mixture of leptochelins A and B (**SI Figure 53 and 54**). Iron, copper, cobalt, and zinc were separately infused as well as an equimolar mixture of metal salts. When each metal was infused separately, the corresponding metal adduct along with the protonated adduct were observed for both leptochelins A and B by LC-$MS^2$. In the mixed metal experiment, the copper-bound species clearly predominated, indicating a strong preference for binding of Cu by

leptochelins A and B. However, zinc and cobalt adduct peaks were also observed in this mixed metal infusion experiment, indicating that the leptochelins are promiscuous in their metal binding ability. This effect was least prominent with iron as an iron adduct was observed only when it was the only metal infused.

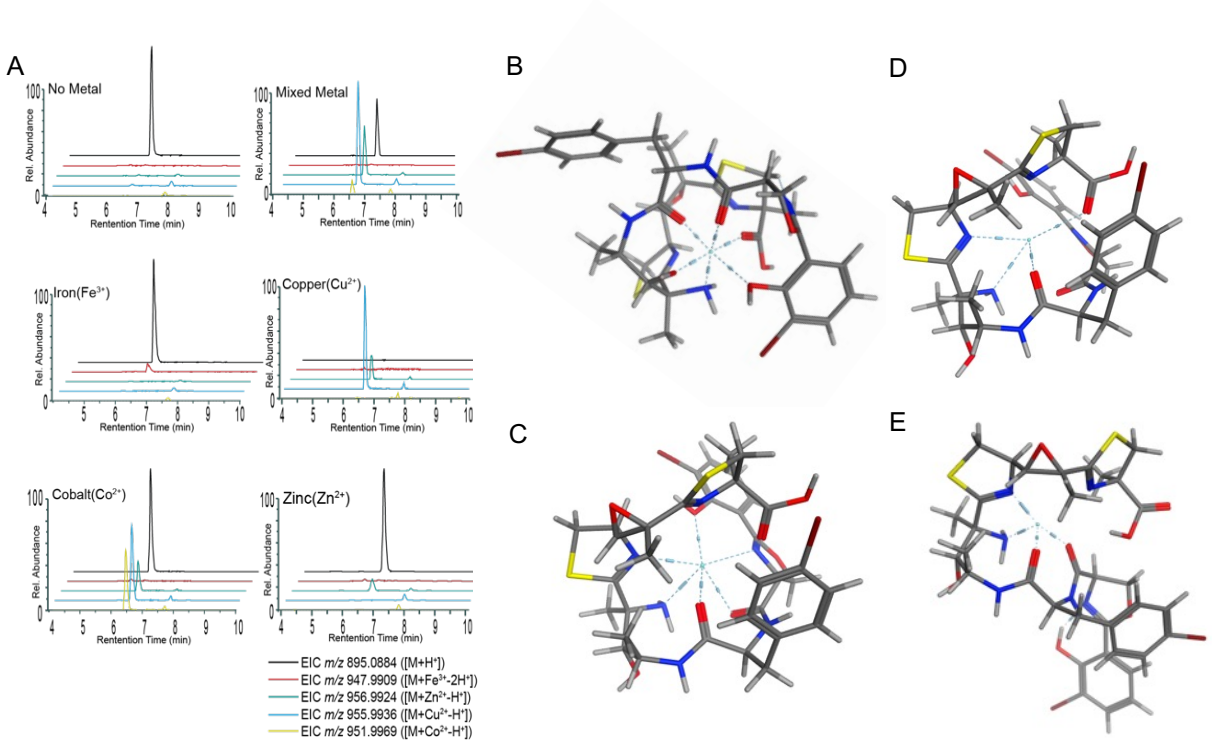



**Figure 7. A.** Extracted Ion Chromatograms from metal-binding studies using pure leptochelin A (**1**). Note that the cobalt injection shows small peaks for copper and zinc adducts. These latter peaks may result from in-instrument contamination. Molecular modeling in MOE (Amber10:EHT) as seen in **B - E** indicates that the carbonyl groups, amine group, phenolic oxygen, and nitrogen atoms of the thiazoline and oxazoline rings participate in the hexadentate and tetradentate coordination of metals. **B.** Modeled pose of iron(III) complexed with leptochelin A (**1**); coordinating residues: C-7 N, C-10 N, C-17 O, and C-26 O. **C.** Modeled pose of cobalt(II) complexed with leptochelin A (**1**); coordinating residues: C-7 N, C-10 N, C-17 O, C-26 O, C-27 N, and C-31 O. **D.** Modeled pose of copper(II) complexed with leptochelin A (**1**); coordinating residues: C-1 O, C-7 N, C-10 N, and C-17 O. **E.** Modeled pose of zinc(II) complexed with leptochelin A (**1**); coordinating residues: C-7 N, C-10 N, C-17 O, and C-26 O. Additionally, when coordinated with metals, the leptochelins produce a pseudo-cyclic conformation, an observation that is consistent with through-space correlations observed by ROESY NMR with zinc-bound leptochelin A (**1**) (e.g., rOe from $H_3$-13 to H-27, **Table 1**, **SI Figure 49**).

The observed selective copper binding but lack of enhanced leptochelin production when copper concentrations were reduced led us to consider other potential roles for the leptochelins. Cyanobacteria are quite sensitive to copper-induced toxicity,[16,18,37] and several commercial “algicides” that target “blue green algae” (= cyanobacteria) contain various formulations of copper. Therefore, we evaluated the ability of the Indonesian leptochelin-producing strain of *Leptothoe* sp. for its ability to tolerate elevated levels of copper, present as copper(II) sulfate. The visually apparent health of cultures of *Leptothoe* in culture media containing elevated levels of copper was charted (**Figure 8**, **SI Table 12**, and **SI Figures 55-62**). Additionally, the relative concentrations of intracellular and extracellular leptochelin production with and without trace metal complexation were measured by LC-MS analysis. A range of copper concentrations were evaluated in SWBG11 media, with native SWBG11 media forming the control condition. These studies were performed with *Leptothoe* sp. ISB3NOV94-8A compared to other taxonomically similar strains for which we possessed genome sequence information. A bioinformatic analysis of these other strains indicated that the strain demonstrating susceptibility to elevations in copper levels lacked genes for siderophore production (*Leptolyngbya* sp. PAP09SEP10-2A), while the other three organisms with siderophore-related genes (e.g., TonB-dependent receptor genes) demonstrated resistance to elevated copper levels. Remarkably, *Leptothoe* sp. ISB3NOV94-8A has the ability to withstand elevated copper concentrations that are 125 - 250 times the average coastal seawater concentrations of 2 µg/L (**Figure 8** and **SI Figures 55-62**), while the strain without siderophore-like biosynthetic genes was negatively impacted in its growth and visible health at elevations of only 1.5 – 2 times the average coastal seawater concentration. Elevation of the copper concentration in the culture medium did not result in an observable change in leptochelin levels, based on LC-MS analysis of the media and the biomass of the resultant cultures (see **SI Figure 63**).

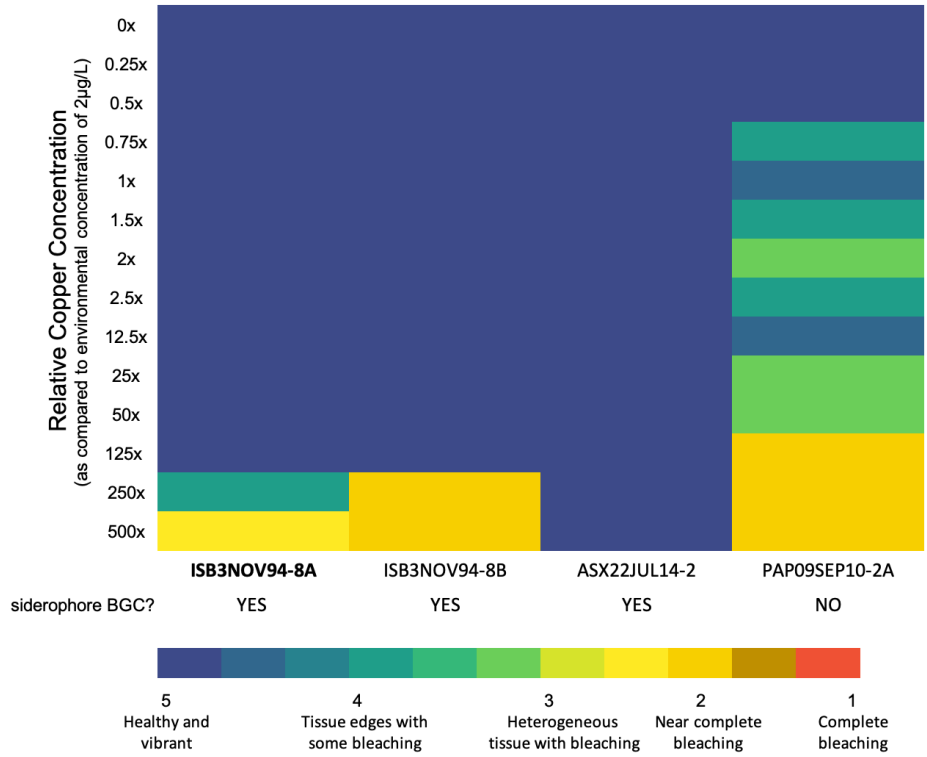



**Figure 8.** Summary of copper toxicity assay reflecting organism health at Day 14 post-inoculation for *Leptothoe* sp. ISB3NOV94-8A and four *Leptolyngbya* spp. Relative concentrations are calculated based on an average coastal seawater concentration of ~2 µg/L.[38,39] Color coding was based on visual inspection of culture health on a scale from Grade 1 (complete bleaching and death) to Grade 5 (healthy and vibrant) on Days 0, 2, 4, 7, 9, 10, and 14. The genome assemblies of each organism were evaluated for the presence of siderophore-producing genes. The three organisms showing resilience to increases in copper concentrations possessed genes associated with siderophore production.

### Biological Properties of Leptochelins A-C

The cytotoxicity of metal-free leptochelin A (**1**) was evaluated against a panel of human cancer cell lines, including NCI-H460 lung carcinoma, HeLa cervical carcinoma, SF188 glioblastoma, D283-med medulloblastoma, and HCT116 colon carcinoma (**SI Table 13**, **SI Figures 65-68**). Free leptochelin A (**1**) was cytotoxic to all the cell lines tested with dose-dependent cytotoxicity. D283-med medulloblastoma cells ($IC_{50}$ = 390 ± 20 nM) and monolayers of HCT116 colon carcinoma cells ($IC_{50}$ = 400 ± 69 nM) were the most sensitive to leptochelin A. The cytotoxic effects of free and zinc-bound forms of leptochelins A-C (**1-3**) were compared in HCT 116 cells; the zinc-bound form of each was significantly less cytotoxic (**SI Table 13**). In addition to evaluation of the leptochelins in 2D monolayer cultures, we also evaluated their activity in three-dimensional (3D) multicellular tumour spheroids (**SI Table 13**, **SI Figure 69**). Metal free leptochelins caused a decrease in proliferating and quiescent cells and an increase in the necrotic core in three-dimensional multicellular tumour spheroids of HCT 116 cells (see **SI Section 2**).

## Discussion

Despite investigations of the unique natural products of marine cyanobacteria for the past 50 years, they continue to be a rich source of structurally novel and biologically active metabolites.[40–42] In the last few decades, this has been assisted and enriched by biosynthetic investigations, and most recently, the genetic basis for their production has been pursued.[25] Indeed, genome sequencing projects have revealed some strains of marine cyanobacteria to use over 20% of their relatively large genomes (7-10 MB) to potentially encode for over 40 different natural products.[43] They are especially rich in hybrid NRPS-PKS pathways which display a strong collinearity between gene order and the assembly process, allowing for their relatively accurate deciphering into predicted chemical structures. Additionally, it is quite frequent that unique biochemical transformations are observed in these otherwise interpretable pathways, and these distinctive enzymatic reactions are enriching the tool kits for synthetic chemical biology.[44–46]

In the current work, we discovered a series of structurally novel natural products named leptochelins A-C (**1**-**3**) from three geographically dispersed cyanobacterial collections of the genus *Leptothoe*. Complete structure elucidation of these complex molecules was challenging, and required multiple orthogonal strategies including NMR, mass spectrometry, bioinformatic deduced stereochemical insights, Marfey's analysis, and chemical reactivities, in concert with molecular modeling. Moreover, genome sequencing of all three of these collections resulted in identification of the putative biosynthetic gene cluster that encodes production of the leptochelins. Interestingly, this hybrid NRPS-PKS biosynthetic gene cluster is extremely well conserved across these geographically disparate *Leptothoe* collections, signifying the importance of the encoded metabolites to the ecology and physiology of the producing strains. The genome sequence data also revealed signatures for metal-dependent regulation in close proximity to the BGC and suggested that the leptochelins could have metallophore properties.

This recognition of the potential metallophore properties of the leptochelins inspired a deeper investigation of their capacity to bind metals. This property was initially recognized by a careful inspection of LC-MS traces of the leptochelins, and recognition that iron, cobalt and zinc adducts were present. Subsequently, culture experiments were conducted with reduced available iron, leading to an increase in production of leptochelin A. This capacity to bind metal ions was probed in a native metabolomics investigation with introduction of individual metal species as well as mixed metal formulations into the flow system following HPLC separation. This was highly insightful as it revealed a strong preference of leptochelin A (**1**) for the binding of copper. However, a culture experiment with media deficient in copper failed to induce upregulation of leptochelin A production, suggesting that leptochelin A might be protective of copper toxicity by sequestration and detoxification of this metal species, thus serving multiple roles in the regulation of local metal ion concentrations. A range of copper concentrations were formulated in liquid SWBG11 media and cultures were grown and evaluated for growth. This revealed that *Leptothoe* possesses a remarkable tolerance to elevated copper concentrations up to 125 – 250-fold over those of normal coastal seawater, and that the leptochelins can therefore potentially serve diverse roles such as helping to acquire trace metals needed for growth, such as iron, and protecting against the toxic effects of elevated copper levels. As anthropomorphic impacts increase in coastal waters, metallophore production may be a critical mechanism for microbial resilience against elevated toxic metal concentrations, such as seen here with these three cyanobacterial strains of *Leptothoe*.

Distinctive features of the structure and biosynthesis of leptochelin A include the two aromatic positions of halogenation, the alpha-methylation of the terminal cysteine-derived thiazoline ring, and two six-carbon moieties that appear to result from a gene duplication of an NRPS-PKS bimodular duet and subsequent neofunctionalization to create distinct yet recognizably similar fragments (C4-8+14; C9-12+15,16). Also present in the leptochelins is a salicylate-oxazoline fragment, along with two additional thiazoline motifs, that represent signature moieties found in other several other metallophores such as amychelin and yersiniabactin.[47,48] In addition to their metal binding properties, the leptochelins have potential pharmaceutical value in that they show relatively potent cancer cell toxicity to multiple cancer cell lines. The potential utility of the leptochelins as cancer chemotherapeutics is under continuing investigation.

## Funding

Research reported in this publication was supported in part by the National Center for Complementary and Integrative Health of the NIH under award number F32AT011475 to N.E.A., and NIH 5R01GM107550-10 to L.G, W.H.G and P.D. D.P. was supported by the Deutsche Forschungsgemeinschaft through the CMFI Cluster of Excellence (EXC 2124) and the Collaborative Research Center CellMap (TRR 261). M.A.R., L.F., J.M., V.V. were supported by FCT (Fundação para a Ciência e a Tecnologia) through CIIMAR research unit strategic fund support (UIDB/04423/2020 and UIDP/04423/2020) and also by the Innovation Pact, Project No. C644915664-00000026 (WP9- Portuguese Blue Biobank under the Blue Economy Pact), known as the "Blue Bioeconomy Pact", resulting from the submission of the application to Notice No. 02/C05-i01/2022, within the scope of the Recovery and Resilience Plan (PRR) co-funded by the Portuguese Republic and the European Union. L.F. is also supported by the Fundação para a Ciência e Tecnologia grant 2022.11979.BD. G.S. was supported by the János Bolyai Research Scholarship (BO/00158/22/5) of the Hungarian Academy of Sciences. K.L.A. was funded by T32 CA009523 and T32 GM067550. Y.E. was funded by NIH 1R03OD034493-01. A portion of this research was performed at the Environmental Molecular Sciences Laboratory, a DOE Office of Science User Facility sponsored by the BER program under Contract No. DE-AC05-76RL01830. A portion of this research was performed under the U.S.-Egypt Science and Technology joint fund number BIO6-002-010 to D.T.A.Y. and the Institutional Fund Projects number IFPIP-601-166-1443 to A.M.A. We acknowledge NSF grant 2116395 for the NMR spectrometer utilized in this work.

## Acknowledgements

We thank N. Arakawa (UCSD – ECAL) for his assistance with $MS^3$ experiments, T. Leão (UCSD) for his assistance with the genome sequencing of *Leptothoe* sp. ISB3NOV94-8A, B. Duggan for NMR experiment acquisition (UCSD), and the University of California Davis Genomics Core for the sequencing and assembly of *Leptothoe* sp. EHU-05/26/07-4. We also appreciate the contributions for biological activity testing at OSU (T. Okino, D. Goeger, A.M. Hau, J.D. Serrill, J.M. Malmo, J. Sikorska, and J. Ishmael) and at the University of Szeged (N. Szemerédi and G. Tóth at the University of Szeged). We thank J. Diaz for use of the SpectraMax microplate reader for cytotoxicity assays. We thank T. Hall and T. Burger for their laboratory assistance. We thank B.J. Philmus for advice on Marfey's analysis and M. K. Gilson for advice on molecular modeling. We thank Raquel Silva for her assistance with DNA extraction and amplification at CIIMAR.

## Disclosures

PCD is an advisor and holds equity in Sirenas and Cybele, consulted for MSD animal health in 2023, and he is a Co-founder, scientific advisor and holds equity in Ometa Labs, Arome, and Enveda with prior approval by UC San Diego.

## Data Availability

Raw and processed LC-$MS^2$ and $MS^3$ data are available through the MassIVE repository (massive.uscd.edu) under the following identifier: MSV000084765. The genome assemblies have been submitted to NCBI with accession numbers SAMN38524433 for *Leptothoe* sp. ISB3NOV94-8A, SAMN38764028 for *Leptothoe* sp. EHU-05/26/07-4, and SAMN34340027 for *Leptothoe* sp. LEGE 181152. Additionally, the putative BGCs from each genome assembly have been submitted to the MIBiG repository (https://mibig.secondarymetabolites.org/) under Accession Numbers BGC0002819, BGC0002820 and BGC0002821. NMR Data is available through NP-MRD.

## References:


[1] H. W. Jannasch, M. J. Mottl, *Science* **1985**, *229*, 717–725.
[2] I. Perera, S. R. Subashchandrabose, K. Venkateswarlu, R. Naidu, M. Megharaj, *Appl. Microbiol. Biotechnol.* **2018**, *102*, 7351–7363.
[3] A. L. Hauptmann, M. Stibal, J. Bælum, T. Sicheritz-Pontén, S. Brunak, J. S. Bowman, L. H. Hansen, C. S. Jacobsen, N. Blom, *Extremophiles* **2014**, *18*, 945–951.
[4] N.-U.-H. Ghori, M. J. Wise, A. S. Whiteley, *Front. Microbiol.* **2021**, *12*, 649594.
[5] J. W. Schopf, *Proc. Natl. Acad. Sci.* **1994**, *91*, 6735–6742.
[6] C. E. Blank, *Geobiology* **2004**, *2*, 1–20.
[7] M. M. Tice, D. R. Lowe, *Nature* **2004**, *431*, 549–552.
[8] R. C. Hider, X. Kong, *Nat. Prod. Rep.* **2010**, *27*, 637.
[9] K. N. Raymond, G. Müller, B. F. Matzanke, in *Struct. Chem.*, Springer Berlin Heidelberg, Berlin, Heidelberg, **1984**, pp. 49–102.
[10] M. Miethke, M. A. Marahiel, *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 413–451.
[11] Y. Ito, A. Butler, *Limnol. Oceanogr.* **2005**, *50*, 1918–1923.
[12] H. Beiderbeck, K. Taraz, H. Budzikiewicz, A. E. Walsby, *Z. Für Naturforschung C* **2000**, *55*, 681–687.
[13] K. B. Mullis, J. R. Pollack, J. B. Neilands, *Biochemistry* **1971**, *10*, 4894–4898.
[14] J. F. Mohr, F. Baldeweg, M. Deicke, C. F. Morales-Reyes, D. Hoffmeister, T. Wichard, *J. Nat. Prod.* **2021**, *84*, 1216–1225.
[15] C. A. Blindauer, *JBIC J. Biol. Inorg. Chem.* **2011**, *16*, 1011–1024.
[16] D. M. McKnight, F. M. M. Morel, *Limnol. Oceanogr.* **1980**, *25*, 62–71.
[17] S. W. Wilhelm, C. G. Trick, *Limnol. Oceanogr.* **1994**, *39*, 1979–1984.
[18] M. Huertas, L. López-Maury, J. Giner-Lamia, A. Sánchez-Riego, F. Florencio, *Life* **2014**, *4*, 865–886.
[19] D. Konstantinou, E. Voultsiadou, E. Panteris, S. Zervou, A. Hiskia, S. Gkelis, *J. Phycol.* **2019**, *55*, 882–897.
[20] N. Okujo, M. Saito, S. Yamamoto, T. Yoshida, S. Miyoshi, S. Shinoda, *Biometals* **1994**, *7*, DOI 10.1007/BF00140480.
[21] O. Mirus, S. Strauss, K. Nicolaisen, A. Von Haeseler, E. Schleiff, *BMC Biol.* **2009**, *7*, 68.
[22] F. C. Beaumont, H. Y. Kang, T. J. Brickman, S. K. Armstrong, *J. Bacteriol.* **1998**, *180*, 862–870.
[23] L. Michel, N. González, S. Jagdeep, T. Nguyen-Ngoc, C. Reimmann, *Mol. Microbiol.* **2005**, *58*, 495–509.
[24] H. M. Patel, J. Tao, C. T. Walsh, *Biochemistry* **2003**, *42*, 10514–10527.
[25] Q. Wu, B. A. Bell, J.-X. Yan, M. G. Chevrette, N. J. Brittin, Y. Zhu, S. Chanana, M. Maity, D. R. Braun, A. M. Wheaton, I. A. Guzei, Y. Ge, S. R. Rajski, M. G. Thomas, T. S. Bugni, *J. Am. Chem. Soc.* **2023**, *145*, 58–69.
[26] E. J. De Vries, D. B. Janssen, *Curr. Opin. Biotechnol.* **2003**, *14*, 414–420.
[27] S. Mori, A. H. Pang, N. Thamban Chandrika, S. Garneau-Tsodikova, O. V. Tsodikov, *Nat. Commun.* **2019**, *10*, 1255.
[28] S. Gäfe, H. H. Niemann, *Acta Crystallogr. Sect. Struct. Biol.* **2023**, *79*, 596–609.
[29] J. Zheng, A. T. Keatinge-Clay, *J. Mol. Biol.* **2011**, *410*, 105–117.
[30] R. T. Williamson, A. Boulanger, A. Vulpanovici, M. A. Roberts, W. H. Gerwick, *J. Org. Chem.* **2002**, *67*, 7927–7936.
[31] I. E. Ndukwe, X. Wang, N. Y. S. Lam, K. Ermanis, K. L. Alexander, M. J. Bertin, G. E. Martin, G. Muir, I. Paterson, R. Britton, J. M. Goodman, E. J. N. Helfrich, J. Piel, W. H. Gerwick, R. T. Williamson, *Chem. Commun.* **2020**, *56*, 7565–7568.
[32] M. J. Bertin, A. Vulpanovici, E. A. Monroe, A. Korobeynikov, D. H. Sherman, L. Gerwick, W. H. Gerwick, *ChemBioChem* **2016**, *17*, 164–173.
[33] C. C. Thornburg, M. Thimmaiah, L. A. Shaala, A. M. Hau, J. M. Malmo, J. E. Ishmael, D. T. A. Youssef, K. L. McPhail, *J. Nat. Prod.* **2011**, *74*, 1677–1685.

[34] N. A. Moss, T. Leao, E. Glukhov, L. Gerwick, W. H. Gerwick, in *Methods Enzymol.*, Elsevier, **2018**, pp. 3–43.
[35] A. T. Aron, D. Petras, R. Schmid, J. M. Gauglitz, I. Büttel, L. Antelo, H. Zhi, S.-P. Nuccio, C. C. Saak, K. P. Malarney, E. Thines, R. J. Dutton, L. I. Aluwihare, M. Raffatellu, P. C. Dorrestein, *Nat. Chem.* **2022**, *14*, 100–109.
[36] R. Reher, A. T. Aron, P. Fajtová, P. Stincone, B. Wagner, A. I. Pérez-Lorente, C. Liu, I. Y. B. Shalom, W. Bittremieux, M. Wang, K. Jeong, M. L. Matos-Hernandez, K. L. Alexander, E. J. Caro-Diaz, C. B. Naman, J. H. W. Scanlan, P. M. M. Hochban, W. E. Diederich, C. Molina-Santiago, D. Romero, K. A. Selim, P. Sass, H. Brötz-Oesterhelt, C. C. Hughes, P. C. Dorrestein, A. J. O'Donoghue, W. H. Gerwick, D. Petras, *Nat. Commun.* **2022**, *13*, 4619.
[37] S. Lima, J. Matinha-Cardoso, J. Giner-Lamia, N. Couto, C. C. Pacheco, F. J. Florencio, P. C. Wright, P. Tamagnini, P. Oliveira, *J. Hazard. Mater.* **2022**, *431*, 128594.
[38] A. G. Lewis, P. H. Whitfield, **1974**.
[39] R. N. McNeely, V. P. Neimanis, L. Dwyer, *Water Quality Sourcebook: A Guide to Water Quality Parameters*, Inland Waters Directorate, Water Quality Branch, **1979**.
[40] M. R. Jones, E. Pinto, M. A. Torres, F. Dörr, H. Mazur-Marzec, K. Szubert, L. Tartaglione, C. Dell'Aversano, C. O. Miles, D. G. Beach, P. McCarron, K. Sivonen, D. P. Fewer, J. Jokela, E. M.-L. Janssen, *Water Res.* **2021**, *196*, 117017.
[41] P. Nandagopal, A. N. Steven, L.-W. Chan, Z. Rahmat, H. Jamaluddin, N. I. Mohd Noh, *Biology* **2021**, *10*, 1061.
[42] R. Carpine, S. Sieber, *Curr. Res. Biotechnol.* **2021**, *3*, 65–81.
[43] T. Leão, M. Wang, N. Moss, R. Da Silva, J. Sanders, S. Nurk, A. Gurevich, G. Humphrey, R. Reher, Q. Zhu, P. Belda-Ferre, E. Glukhov, S. Whitner, K. L. Alexander, R. Rex, P. Pevzner, P. C. Dorrestein, R. Knight, N. Bandeira, W. H. Gerwick, L. Gerwick, *Mar. Drugs* **2021**, *19*, 20.
[44] K. Shimoda, N. Kubota, H. Hamada, M. Kaji, T. Hirata, *Tetrahedron Asymmetry* **2004**, *15*, 1677–1679.
[45] M. Górak, E. Żymańczyk-Duda, *Green Chem* **2015**, *17*, 4570–4578.
[46] A. L. Lukowski, F. M. Hubert, T.-E. Ngo, N. E. Avalon, W. H. Gerwick, B. S. Moore, *J. Am. Chem. Soc.* **2023**, jacs.3c05750.
[47] M. R. Seyedsayamdost, M. F. Traxler, S.-L. Zheng, R. Kolter, J. Clardy, *J. Am. Chem. Soc.* **2011**, *133*, 11434–11437.
[48] B. A. Pfeifer, C. C. C. Wang, C. T. Walsh, C. Khosla, *Appl. Environ. Microbiol.* **2003**, *69*, 6698–6702.

**TOC Graphic:**