



Cyanochelins, an Overlooked Class of Widely Distributed Cyanobacterial Siderophores, Discovered by Silent Gene Cluster Awakening

Tomáš Galica,^{a,b} Nicola Borbone,^c Jan Mareš,^{a,b,d} Andreja Kust,^{a,d,e} Alessia Caso,^c Germana Esposito,^c Kumar Saurav,^a Jan Hájek,^{a,b} Klára Řeháková,^{d,f} Petra Urajová,^a Valeria Costantino,^c Pavel Hrouzek^a

^aCentre Algatech, Institute of Microbiology, The Czech Academy of Sciences (CAS), Třeboň, Czech Republic

^bFaculty of Science, University of South Bohemia, České Budějovice, Czech Republic

^cDipartimento di Farmacia, Università degli Studi di Napoli Federico II, Naples, Italy

^dBiology Centre of the CAS, Institute of Hydrobiology, České Budějovice, Czech Republic

^eInnovative Genomics Institute, University of California, Berkeley, California, USA

^fInstitute of Botany of the Czech Academy of Sciences, Třeboň, Czech Republic

ABSTRACT Cyanobacteria require iron for growth and often inhabit iron-limited habitats, yet only a few siderophores are known to be produced by them. We report that cyanobacterial genomes frequently encode polyketide synthase (PKS)/nonribosomal peptide synthetase (NRPS) biosynthetic pathways for synthesis of lipopeptides featuring β -hydroxyaspartate (β -OH-Asp), a residue known to be involved in iron chelation. Iron starvation triggered the synthesis of β -OH-Asp lipopeptides in the cyanobacteria *Rivularia* sp. strain PCC 7116, *Leptolyngbya* sp. strain NIES-3755, and *Rubidibacter lacunae* strain KORDI 51-2. The induced compounds were confirmed to bind iron by mass spectrometry (MS) and were capable of Fe^{3+} to Fe^{2+} photoreduction, accompanied by their cleavage, when exposed to sunlight. The siderophore from *Rivularia*, named cyanochelin A, was structurally characterized by MS and nuclear magnetic resonance (NMR) and found to contain a hydrophobic tail bound to phenolate and oxazole moieties followed by five amino acids, including two modified aspartate residues for iron chelation. Phylogenomic analysis revealed 26 additional cyanochelin-like gene clusters across a broad range of cyanobacterial lineages. Our data suggest that cyanochelins and related compounds are widespread β -OH-Asp-featuring cyanobacterial siderophores produced by phylogenetically distant species upon iron starvation. Production of photolabile siderophores by phototrophic cyanobacteria raises questions about whether the compounds facilitate iron monopolization by the producer or, rather, provide Fe^{2+} for the whole microbial community via photoreduction.

IMPORTANCE All living organisms depend on iron as an essential cofactor for indispensable enzymes. However, the sources of bioavailable iron are often limited. To face this problem, microorganisms synthesize low-molecular-weight metabolites capable of iron scavenging, i.e., the siderophores. Although cyanobacteria inhabit the majority of the Earth's ecosystems, their repertoire of known siderophores is remarkably poor. Their genomes are known to harbor a rich variety of gene clusters with unknown function. Here, we report the awakening of a widely distributed class of silent gene clusters by iron starvation to yield cyanochelins, β -hydroxy aspartate lipopeptides involved in iron acquisition. Our results expand the limited arsenal of known cyanobacterial siderophores and propose products with ecological function for a number of previously orphan gene clusters.

KEYWORDS cyanobacteria, iron acquisition, lipopeptides, secondary metabolism, siderophores

Citation Galica T, Borbone N, Mareš J, Kust A, Caso A, Esposito G, Saurav K, Hájek J, Řeháková K, Urajová P, Costantino V, Hrouzek P. 2021. Cyanochelins, an overlooked class of widely distributed cyanobacterial siderophores, discovered by silent gene cluster awakening. *Appl Environ Microbiol* 87:e03128-20. <https://doi.org/10.1128/AEM.03128-20>.

Editor Hideaki Nojiri, University of Tokyo

Copyright © 2021 American Society for Microbiology. All Rights Reserved.

Address correspondence to Pavel Hrouzek, hrouzek@alga.cz.

Received 26 December 2020

Accepted 4 June 2021

Accepted manuscript posted online 16 June 2021

Published 11 August 2021

Iron, the fourth most abundant element in the Earth's crust, is an important metal cofactor to many metabolic enzymes and an essential nutrient to all known living organisms. Yet, due to the prevalent oxidative conditions on the surface of our planet, it is frequently present in its oxidized Fe^{3+} form that is generally insoluble and not easily acquired by living organisms (1, 2). Microorganisms have evolved various strategies to cope with this problem, one of which is the production and excretion of siderophores, which are low-molecular-weight metabolites able to reversibly bind an iron atom (1, 3). Excreted siderophores are able to pick up even trace amounts of iron from the environment and facilitate cellular iron uptake or prevent precipitation and loss of its bioavailable forms (1).

Cyanobacteria depend on iron as do all other microorganisms and live in various habitats, including those with a low availability of iron as a nutrient. Several mechanisms of cyanobacterial iron stress response have been described, such as remodeling of photosynthetic complexes, upregulation of iron transport systems, and siderophore production (2, 3). Yet, intriguingly, few cyanobacterial siderophores are known; only two hydroxamate siderophores (schizokinen and synechobactin) and a single catecholate siderophore (anachelin) have been structurally characterized (2). Cyanobacterial genomes are known to harbor a large number of gene clusters (4) that are silent in luxuriant cultivation medium under laboratory conditions. These clusters may be activated if an appropriate trigger is present and, possibly, some may encode production of unknown siderophores, as shown in other bacteria (5).

Knowledge of siderophores from other bacteria shows there are a number of structural motifs capable of iron chelation, some of which can be expected to also occur in cyanobacteria (1, 6). The carboxylate moiety of β -hydroxy aspartate (β -OH-Asp) is one such case (1, 6). Several bacterial siderophores, e.g., serobactin, cupriachelin, and pacifibactin, were found to employ β -OH-Asp for iron chelation (7–9). Their synthesis is driven by nonribosomal peptide synthetases (NRPS) with specialized modules for selective incorporation of β -OH-Asp into the nascent chain (7–9). The modules consist of an adenylation domain (A-domain) that activates aspartic acid (Asp), a dioxygenase that hydroxylates the activated Asp on its β carbon, and a specific condensation domain (C-domain) that does not effectively proceed to form a peptide bond prior to successful hydroxylation of Asp to β -OH-Asp (7, 10).

A common feature of β -OH-Asp siderophores is the presence of a prolonged hydrocarbon chain on the N terminus (1). The hydrocarbon chain is thought to play a role in adjusting diffusion rates via an enhanced association with biological membranes and so limit siderophore loss, especially in the case of microbes living in aquatic environments (1, 11). A prominent mechanism of N-terminal acylation of cyanobacterial nonribosomal peptides is the employment of fatty acyl-AMP ligase (FAAL) at the beginning of synthesis (12).

The genetic signature for both the NRPS module for β -OH-Asp and the FAAL domain can be identified in nucleotide sequences, and the growing number of publicly available cyanobacterial genomes provides suitable data for targeted genome mining. Our recent mining of cyanobacterial gene clusters that possibly encode biosynthesis of lipopeptides yielded, among others, a group of 10 similar clusters with features that suggest the presence of β -OH-Asp in the product and offered a direction for further research (13). In the present study, we performed a bioinformatic analysis of selected silent gene clusters, induced the production of a new type of cyanobacterial siderophores by iron deprivation, described the structure of a selected representative and, finally, took a closer look at the distribution across the cyanobacterial phylum.

RESULTS

Identification of β -OH-Asp lipopeptide gene clusters. Three hybrid gene clusters that combine elements of polyketide synthesis (PKS) and NRPS (PKS/NRPS) putatively coding for β -OH-Asp-containing peptides were found (Fig. 1A, Table S1 to S3 in the supplemental material). One cluster was located on the chromosome of *Rivularia* sp.

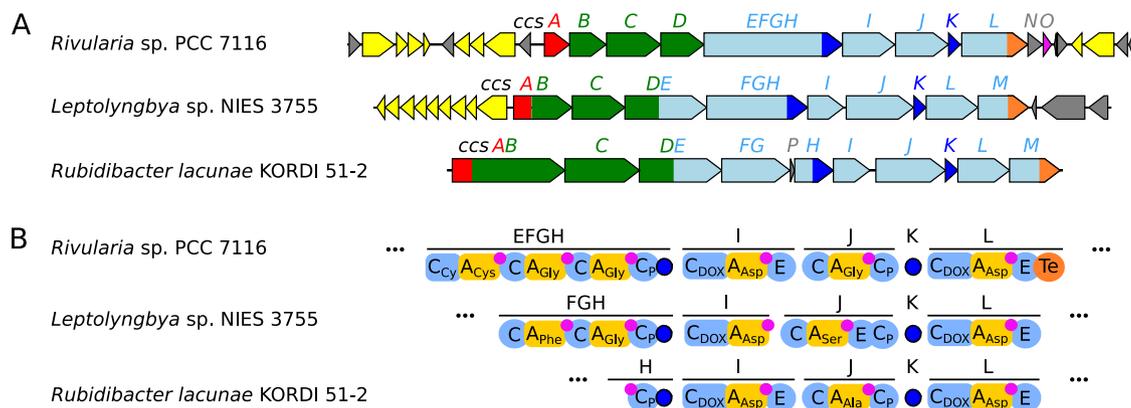


FIG 1 (A) Gene clusters encoding hybrid PKS/NRPS biosynthesis of cyanochelin with proportions of genes colored according to the type of biosynthetic step (FAAL, red; PKS, green; NRPS, blue; dioxygenase involved in β -OH-Asp formation, purple; thioesterase/terminal reductase, orange; aldo-keto reductase, pink) and genes related to siderophore transport (depicted in yellow). (B) Selected NRPS modules and related dioxygenases responsible for incorporation of two β -OH-Asp into cyanochelins. C, condensation domain; A, adenylation domain; E, epimerase domain; Te, thioesterase domain; pink, thiolation domain; dark blue, dioxygenase. Subscript specifies the type of C domain or specificity of A-domain.

PCC 7116 ([CP003549](#), 8.7 Mbp), the second was on a plasmid from *Leptolyngbya* sp. NIES-3755 ([AP017310.1](#), 97 kbp), and the third was on a whole-genome scaffold from *Rubidibacter lacunae* KORDI 51-2 ([ASSJ01000034.1](#), 54 kbp). Organization of the biosynthetic genes in one direction, with genes for tailoring enzymes scattered among the PKS/NRPS genes, suggests that the chronology of biosynthetic steps will reflect the cluster topology. All three gene clusters encode a fatty acyl-AMP ligase (FAAL) domain (in *Rivularia* as a standalone, in the other two fused to subsequent PKS modules) followed by several PKS genes (*ccsB* to *-D*), suggesting the presence of a fatty acyl residue elongated by several (modified) malonyl units at the N terminus of the compounds. Analysis of the specificity of the individual NRPS A-domains contained in *ccsE* to *-M* (Table S4) led to prediction of a peptide sequence composed of 6 or 7 amino acids.

The first NRPS module (CcsE) contains an A-domain with predicted specificity for cysteine preceded by a heterocyclization domain, suggesting the formation of a thiazole heterocycle attached to a nascent N-terminal chain. The following two amino acids activated by the downstream NRPS modules CcsFG vary among the gene clusters: Gly-Gly (*Rivularia*), Phe-Gly (*Leptolyngbya*), and Gly-Ala (*Rubidibacter*). Most importantly, in all three candidate clusters, the presence of two Asp-specific NRPS modules (*ccsI* and *ccsL*) with associated dioxygenases (*ccsH* and *ccsK*) suggests the incorporation of two β -OH-Asp in the putative product (Fig. 1B). The β -OH-Asp-specific modules are separated by NRPS modules (*ccsJ*) with predicted specificities to Gly (*Rivularia*) or Ser (*Leptolyngbya* and *Rubidibacter*). In *Rubidibacter* and *Leptolyngbya*, an additional NRPS module activating Glu and Gly, respectively, is encoded at the 3' end. While in *Rivularia* the gene cluster is terminated by a canonical thioesterase domain, the last module in *Leptolyngbya* and *Rubidibacter* instead contains a terminal reductase domain. An analogous organization of the terminal NRPS module was previously reported to result in the formation of a primary alcohol at the C terminus (14). Intriguingly, an enzyme of analogous function (aldo-keto reductase) is encoded within the *Rivularia* gene cluster by the *ccsN* gene, immediately following the terminal NRPS module.

Several genes encoding siderophore transporters and accessory processing proteins (colored yellow in Fig. 1, Tables S5 and S6) are localized in the immediate vicinity of the candidate clusters in *Rivularia* and *Leptolyngbya* and, thus, strengthen the evidence that the putative products of the investigated biosynthetic gene clusters (BGCs) are siderophores. In *Rubidibacter*, the biosynthetic genes span over the whole 54-kb genomic fragment and thus any further analysis of the surrounding genomic context is hindered.

In bacteria, iron-stress response is frequently controlled by a transcriptional repressor called ferric uptake regulator (FurA), which is bound to a DNA regulatory sequence called a

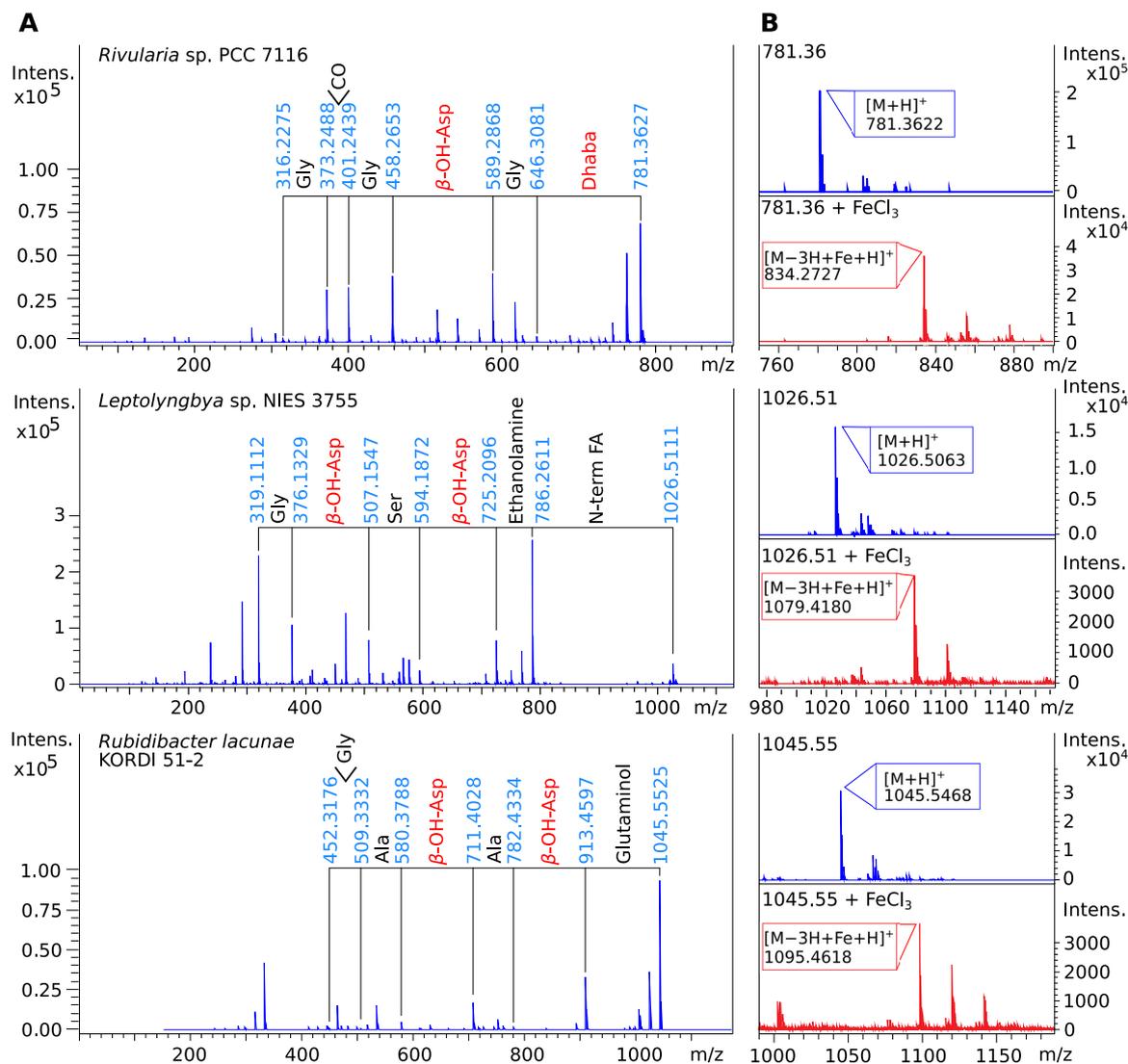


FIG 2 HRMS experiments demonstrating key features of cyanochelins. (A) MS/MS fragmentation spectra of β -OH-Asp siderophores detected at m/z 781.36, 1,026.51, and 1,045.55, demonstrating neutral loss characteristic for β -OH-Asp residues (see also Table S9 to S11 in the supplemental material). (B) Mass spectra obtained after injection of cyanochelin A1 and B1 with (red) or without (blue) supplemented Fe³⁺.

Fur box (3, 15). An *in silico* search yielded highly scoring putative Fur boxes located upstream of the first biosynthesis-related gene in both *Rivularia* and *Leptolyngbya* (Tables S7 and S8). In both aforementioned cases, the promoter region of the biosynthetic genes is shared with a cassette of genes coding for siderophore transporters in the opposite direction and an additional Fur box, likely regulating this cassette. In *Rivularia*, the Fur box for transporter cassette partly overlaps with the Fur box for biosynthetic genes. Such findings further support the hypothesis that the DNA region upstream of *ccsA* is likely to be a regulatory region responsive to iron stress via FurA that governs the expression of the clusters described above.

Detection of β -OH-Asp lipopeptides. Upon iron deprivation, cultures of all the three strains accumulated a compound(s) that showed tandem mass spectrometry (MS/MS) fragmentation patterns highly congruent with the bioinformatic prediction. Fragmentation analysis of compound with m/z at 781.36 found in *Rivularia* confirmed the partial amino acid sequence [Gly]-[Gly]-[β -OH-Asp]-[Gly] (Fig. 2, Fig. S1, Table S9). For candidate compounds from *Leptolyngbya* (m/z 1,026.51) and *Rubidibacter* (m/z 1,045.51), the analysis provided the AA sequences [Phe]-[Gly]-[β -OH-Asp]-[Ser]-[β -OH-Asp] and

[Gly]-[Ala]-[β -OH-Asp]-[Ala]-[β -OH-Asp], respectively (Fig. 2A, Fig. S2 and S3, Table S10 and S11). Cleavage of a β -OH-Asp moiety likely involved in the Fe chelation was represented by a distinctive neutral loss of 131.02 within the peptide sequence. Furthermore, all the compounds provided neutral loss corresponding to the presence of an amino acid with a carboxylic functional group reduced to its primary alcohol: *Rivularia* (135.05, Asp reduced to dihydroxy-amino-butyric acid [Dhaba]); *Leptolyngbya* (61.05, Gly reduced to ethanolamine); and *Rubidibacter* (132.09, Gln reduced to glutaminol). Additionally, all the detected compounds provided various a and b types of N-terminal fragments, suggesting the linkage of the thiazole/oxazole heterocycle to an aliphatic chain (Fig. S1 to S3 and Table S9 to S11). *Rivularia* and *Leptolyngbya* strains produced additional compounds at m/z 753.33 and 998.48 with the same partial AA sequence as 781.36 and 1,026.51, respectively, but with different N-terminal fragments (Fig. S4A and B). Based on their m/z , it was possible to conclude that these variants possess hydrocarbon chains shorter by two methylene groups ($\Delta \sim 28.0307$ Da, C_2H_4) (Fig. S4). Finally, the compound at m/z 1,046.54 (a variant of 1,045.55 from *Rubidibacter*) was found to differ in the last AA containing hydroxy-amino-pentanoic acid instead of glutaminol (Fig. S4C) as a result of the relaxed A-domain specificity, which activates either Glu or Gln, which is subsequently reduced to a corresponding alcohol.

Interestingly, all candidate compounds were found in their putative apo- and iron-complexed (m/z greater by 52.9120 Da) form, with the latter eluting earlier during analysis (Fig. S5). To verify siderophore activity, minute amounts of investigated compounds were isolated and analyzed by direct high-resolution mass spectrometry (HRMS), either with or without roughly equimolar amounts of Fe^{3+} ions (Fig. 2B). While samples without Fe^{3+} showed a compound in $[M+H]^+$ at the expected m/z along with common adducts (Na^+ , K^+), the samples with supplemented iron showed m/z greater by 52.91 Da from the original m/z , i.e., 834.27 for 781.36, 1,079.42 for 1,026.51, and 1,098.46 for 1,045.55, corresponding to $[M - 3H + Fe + H]^+$.

The candidate compounds share the key structural features, in particular the two β -OH-Asp residues (alternatively β -OH-Asp/Dhaba) separated by a small amino acid residue close to the C terminus, the hydrocarbon chain on the N terminus, and having the C-terminal amino acid reduced to a primary alcohol. Furthermore, all the candidate compounds are capable of binding iron; hence, we collectively refer to them as cyanochelins and encourage the future use of the term for similar compounds.

Photoreduction of iron by identified compounds. A prominent feature of β -OH-Asp siderophores is their ability to reduce ferric iron upon exposure to light, accompanied by cleavage of the siderophore at its peptide backbone close to the β -OH-Asp residue. Minute amounts of cyanochelins (m/z 781.36, 1,026.51, and 1,045.51) were isolated by high-performance liquid chromatography (HPLC) and subsequently exposed to direct sunlight in the presence of $FeCl_3$ (0.1 mM) and bathophenanthroline disulfonic acid (0.1 mM), which develops colorful complexes with divalent iron. Light-exposed aliquots clearly developed BPDS- Fe^{2+} complexes, manifested as pink coloration of the solution visible by eye and, therefore, confirmed the potential of the isolated compounds to reduce ferric iron. Moreover, in contrast to light-protected samples the HPLC high-resolution tandem mass spectrometry (HPLC-HRMS/MS) analysis of sunlight-exposed aliquots revealed the expected products of photolysis accompanied by a decreased amount of the candidate compounds. Molecular ions at m/z 475.29 and m/z 476.28 matched the expected molecular mass of photolytic products from m/z 781.36 cleaved between glycine (AA3) and β -OH-Asp (AA4). Furthermore, their MS/MS product ions correspond to that of cyanochelin m/z 781.36 (Fig. S6). Similarly, compounds with m/z 633.40 and m/z 634.39 originate from cyanochelin m/z 1,026.51 and m/z 598.39 from m/z 1,045.55 (Fig. S6).

For further experiments, we isolated 2.2 mg of cyanochelin m/z 781.36. An absorbance spectrum of the isolated compound was measured in the presence of equimolar amounts of Fe^{3+} before and after exposure to direct sunlight. An absorbance peak at 318 nm, corresponding to the Fe^{3+} - α -hydroxycarboxylate charge transfer band

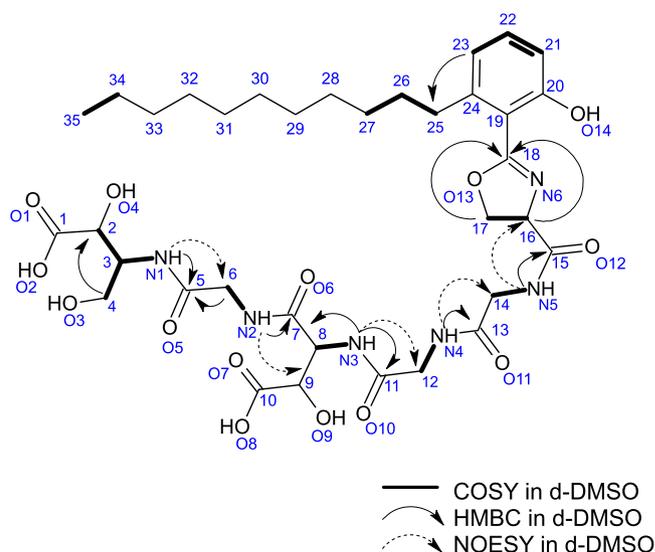


FIG 3 Structure of cyanochelin A. The most relevant 2D NMR correlations are depicted.

reported previously (9), was present before light exposure but was clearly diminished afterward (Fig. S6B).

Structural elucidation of cyanochelin A. Cyanochelin A, the major β -OH-Asp siderophore product from the *Rivularia* strain, was isolated and its chemical structure was elucidated using a combination of HRMS and NMR techniques. We recorded full sets of 1D and 2D NMR spectra (^1H , ^{13}C , COSY, HMBC, HSQC [Fig. S7 to S11, respectively], TOCSY, NOESY, and ROESY) in CD_3OD (*D*-methanol) and $(\text{CD}_3)_2\text{SO}$ (*D*-DMSO) deuterated solvents to take advantage of the spread of overlapping signals in different solvents. The analysis in *D*-DMSO also allowed us to trace the COSY connectivities involving NH protons of the amino acid chain. The most relevant COSY, HMBC, and NOESY correlations and all ^1H and ^{13}C chemical shift assignments are summarized in Fig. 3, Table S12, and Fig. S7 to S11. The ^1H -NMR spectrum of cyanochelin A was clearly suggestive of a lipopeptide molecule with distinctive signals of a saturated hydrocarbon chain, i.e., the presence of a methyl triplet typical for a methyl terminus at δ 0.85 ppm (C-35) (Fig. 3, Table S12), coupled in the COSY spectrum with a methylene at δ 1.26 ppm (C-34) that is part of a 10-methylene-long alkyl chain (δ_{H} 1.23 to 2.75 ppm, δ_{C} 22.1 to 34.5 ppm). The 11-carbon-long chain is directly linked to a phenol ring, as confirmed by the H-23 \rightarrow C-25 HMBC correlation shown in Fig. 3, which is in turn connected to C-18 of the 2-oxazoline moiety, likely resulting from the cyclization of a serine residue. The oxazoline substructure was confirmed by the COSY and HMBC correlations depicted in Fig. 3. The connection of the saturated hydrocarbon chain with a phenolate moiety and the 2-oxazoline was further confirmed by formation of the HRMS/MS fragment at m/z 316.2275 (Fig. 2, Fig. S1, Table S9). The peptide moiety was shown by resonances of five distinct amide-NH signals ($\delta_{\text{H-DMSO}}$ 8.42, 8.20, 7.97, 7.95, and 7.39), three α -proton signals ($\delta_{\text{H-DMSO}}$ 4.14, 4.64, and 4.94), three methylene resonances attributable to three glycine residues ($\delta_{\text{H-DMSO}}$ 3.68, 3.77, and 3.74), and two oxygenated methines, C-2 and C-9, resonating, respectively, at $\delta_{\text{H-DMSO}}$ 4.23 and 4.43. Finally, the presence of the 2,4-dihydroxy, 3-amino butyric acid (Dhaba) in the structure was confirmed by the full set of COSY and HMBC correlations (C-2–C-4) (Fig. 3, Table S12). The amino acid sequence was elucidated in combination with observed MS/MS fragmentation. Cyanochelin A monoisotopic $[\text{M}+\text{H}]^+$ peak at m/z 781.3619 (calc. for $\text{C}_{35}\text{H}_{53}\text{N}_6\text{O}_{14}^+$) provided fragment ions at m/z 646.31 (b5; loss of Dhaba) followed by a sequential loss of Gly (b4; m/z 589.28), β -OH-Asp (b3; m/z 458.26), and Gly (b2; m/z 401.24). Finally, the formation of the a2 and a1

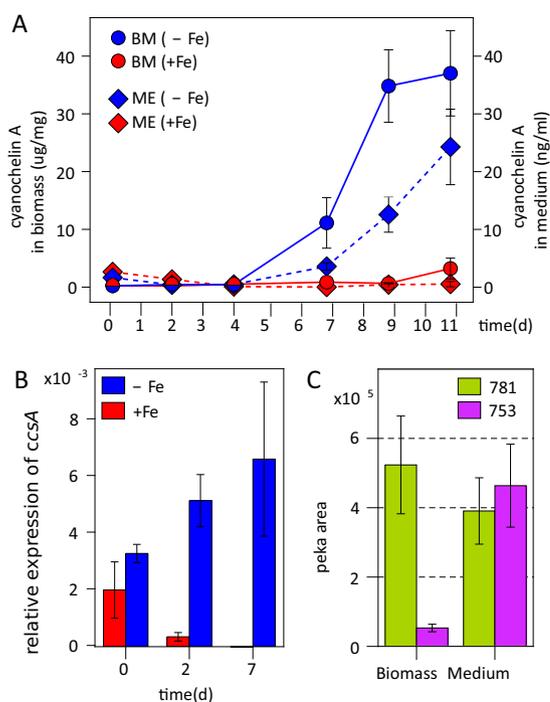


FIG 4 Cyanochelin A production in response to iron starvation. (A) Cultivation of *Rivularia* in iron-depleted (–Fe) and control (+Fe) media showed accumulation of cyanochelin A in the biomass (BM) and in the spent medium (ME). (B) Expression of *ccsA*, the first gene of the cyanochelin gene cluster, normalized to *rnpB* during batch cultivation with or without provided iron. (C) Cyanochelin A (*m/z* 781.36) and its variant with a two-carbon-shorter acyl chain (*m/z* 753.33) were found in the biomass and the spent medium in different ratios.

ions at *m/z* 373.25 and 316.23 can be attributed to fragmentations involving the C-13–C-14 and C-15–C-16 bonds, respectively (Fig. S1, Table S9).

Cyanochelin A production dynamics. To further support the hypothesis that the detected β -OH-Asp lipopeptides are the final products of the candidate gene clusters expressed under iron deprivation, we monitored relative mRNA levels of the starter module gene *ccsA* in *Rivularia* (normalized to *rnpB*) along with accumulation of the cyanochelin A and its putative analogue compound *m/z* 753.33 (Fig. 4). We observed a 2-fold increase in the expression of *ccsA* at 7 days after transferring the *Rivularia* culture to iron-deprived medium (Fig. 4B). In contrast, culture inoculated into fresh Fe-full medium exhibited a gradual downregulation of *ccsA* (Fig. 4B) and finally accounted for a >1,000-fold difference in *ccsA* expression between treated and control cultures after 7 days of the experiment. Along with the changes in expression, we also detected the accumulation of cyanochelin A and its variant compound with *m/z* 753.33 in the starved culture. The maximal observed concentration of cyanochelin A at the end of the experiment reached 35 μ g per mg of biomass and 25 ng per ml of culture medium. Interestingly, the relative amount of 753.33 to cyanochelin A, evaluated as ion peak areas in HRMS analysis at the end of the experiment, differed between biomass and culture medium. In the culture medium, amounts of both compounds were approximately equal, whereas the amount of 753.33 was proportionally 10 times lower in the biomass.

Phylum-wide distribution of cyanochelin-like gene clusters in cyanobacteria.

The cyanochelins reported here were found in strains from three divergent lineages and, hence, indicated a possible widespread distribution in cyanobacteria. Therefore, we investigated publicly available cyanobacterial genomes for the presence of PKS/NRPS gene clusters that would include FAAL and two or more Asp/Asn-specific A-domains with associated β -OH-Asp-specific C-domain in such an arrangement that would be compatible with the putative biosynthesis of cyanochelins. An additional 27 clusters from 29 strains (Table S13) matched the criteria and further interpretation of the clusters split the batch into two groups (Table S14). The first contained the eight clusters with

the two β -OH-Asp residues positioned on the N terminus of the expected products, close to the FA, similar to as is found in cupriachelin (7). The clusters in the second group (19 cases) were expected to encode biosynthesis of lipopeptides with β -OH-Asp residues on the C terminus, as in the isolated cyanochelin A. In 17 cases of the latter group, the β -OH-Asp-specific module is the last PKS/NRPS module; moreover, in 14 cases it is fused to a thioesterase domain, further supporting its C-terminal position. In a majority of the expected products, the β -OH-Asp are separated by Gly (18 cases), analogous to cyanochelin A. Other AA predicted to be found between the two β -OH-Asp are Ser (two cases), Thr (two cases), Orn (two cases), and Asn (one case); in the remaining two cases, the prediction software could not resolve the specificity of the A-domain. The genetic surroundings of the clusters were investigated for the presence of genes encoding transport-related proteins in a similar manner as was done for *Rivularia* and *Leptolyngbya*. In particular, we looked for TonB-dependent siderophore receptors, as they are known to play a key role in siderophore retrieval in bacteria. TonB-dependent transporter genes were found in the vicinity of 25 out of 27 clusters, mostly (17/25) in close proximity (\sim 1,000 bp) upstream of the starting FAAL-encoding gene (Table S15).

Phylogenomic analysis of the available cyanobacterial genomes (Fig. 5) has shown that the three confirmed cyanochelin-producing strains fall into three divergent phylogenetic clades: a deep-branching lineage of simple filamentous cyanobacteria of the genus *Leptolyngbya*, the derived lineage of halotolerant coccoid cyanobacteria (*Rubidibacter*), and the most derived group of heterocytous cyanobacteria (*Rivularia*). The remaining 29 strains containing cyanochelin-like biosynthetic gene clusters were widely distributed throughout the phylum, encompassing both basal and derived clades, with slightly higher abundance in the heterocytous clade.

DISCUSSION

Cyanobacteria depend on considerable amounts of iron as an essential cofactor for basic metabolic processes such as photosynthesis, respiration, or nitrogen fixation (2, 3). Therefore, it is surprising that only three types of cyanobacterial siderophores (with limited structural variability) have been documented in the literature. In particular, schizokinen and synechobactin both employ hydroxyamate, while anachelin uses catecholate residues for iron chelation (2). The recent accumulation of sequenced cyanobacterial genomes and silent gene clusters detected therein offers a great waypoint for the identification of novel siderophores, if iron-depleted medium is employed for the cultivation. In the present study, we report novel cyanobacterial lipopeptides that employ β -OH-Asp for chelation of iron and, hence, broaden the repertoire of chelating residues and siderophores known from cyanobacteria. The most striking finding was the actual distribution of cyanochelins. The compounds investigated in this study come from vastly separated phylogenetic clades across three cyanobacterial orders. The growth mode among the producing strains varies from simple trichal soil-inhabiting cyanobacteria (*Leptolyngbya*) through unicellular in marine sediment (*Rubidibacter*) to heterocytous filaments that form macroscopic colonies on rocks in the oceanic splash zone (*Rivularia*). An additional 27 gene clusters putatively encoding biosynthesis of cyanochelins or related double β -OH-Asp-containing lipopeptides were found across the major cyanobacterial lineages, spanning from basal clades to the most derived ones (Fig. 5). The observed pattern suggests the wide phylogenetic and ecological distribution of β -OH-Asp-containing siderophores, including cyanochelins. Since the specific cyanochelin genes do not show a close homology between the three confirmed producer strains (Table S16), the production of cyanochelins does not seem to be a result of recent horizontal gene transfer. The question of whether the characteristic structure of cyanochelins is maintained as an ancestral metabolic capacity in cyanobacteria or, rather, serially introduced by functional convergence remains to be addressed when more records of cyanochelin producers are available. Previous studies suggested that strong selective pressure is important in siderophore evolution, as

illustrated by the divergent biosynthetic origin of hydroxamate siderophores that can be assembled both by NRPS and NRPS-independent biosynthetic machineries (16).

The two β -OH-Asp separated by a single amino acid residue are expected to participate in iron chelation by alterobactin and have been experimentally proven to provide four coordination bonds for Fe chelation in the bacterial siderophore pacifibactin (9, 17). The participation of β -OH-Asp in iron chelation is manifested in the absorption spectra as a charge transfer band around 300 nm and also as an ability to reduce ferric iron accompanied by lysis of the siderophore upon exposure to light (9). The cyanochelins share these key structural features and properties with the above-mentioned siderophores, which implies that they also employ a double β -OH-Asp motif to provide four out of the six electron pairs required for chelation. As for the remaining two electron pairs, it was not possible to unambiguously determine the donors from the collected data. In the case of cyanochelin A, it is most probable that the phenolate moiety and nitrogen atom form the oxazole ring as reported for mycobactin (1, 18).

The incorporation of β -OH-Asp into cyanochelins is likely implemented in the same manner as in the biosynthesis of cupriachelin and pacifibactin (7, 9) and requires the same enzymatic elements, i.e., dioxygenase participating in conversion of Asp to β -OH-Asp and a C domain with selective affinity for β -OH-Asp. Cupriachelin and pacifibactin have the dioxygenase encoded either as part of a long PKS or NRPS gene or as a standalone gene in a cassette of accessory genes apart from the main body of NRPS genes (7, 9). In the aforementioned cyanochelin gene clusters, the whole machinery that converts Asp to β -OH-Asp and incorporates it into the peptide is integrated among NRPS genes, either fused or standalone, but the order of the genes is fully colinear with the biosynthesis. Such topology is observed in all the other putative cyanochelin gene clusters. Therefore, the observed topology facilitates the identification of further β -OH-Asp-containing NRPS peptides, at least in cyanobacteria, and may improve bioinformatic NRPS prediction tools that still lack accuracy when it comes to modified amino acids.

To limit diffusibility, siderophores from aquatic environments frequently feature a fatty acyl residue linked by a peptide bond (6). In cyanochelin A, the peptide bond is subsequently included into the oxazole ring during the heterocyclization of the first amino acid (serine), further strengthening the binding of the hydrophobic moiety to the peptide core. In an analogous manner, cysteine forms a thiazole ring in cyanochelins from *Rubidibacter* and *Leptolyngbya*. Interestingly, siderophores are often found in multiple variant forms differing only in the size of the hydrophobic moiety and resulting polarity. The N-terminal hydrophobic residue of cyanobacterial siderophores consists of 8 to 14 carbons (synechobactin) (2), whereas in the majority of other bacterial siderophores it consists of 12 to 18 carbons (e.g., marinobactin or amphibactin) (19). Cyanochelins produced by *Rivularia* possess a linear hydrophobic moiety containing 9 to 11 carbons, and in *Rubidibacter* and *Leptolyngbya* this is expected to account for C₂₂ and C₁₈ to C₂₀ alkyl chains, respectively. The variability of hydrophobic chains is likely achieved by the extended specificity of FAAL, which recognizes several FAs and introduces them into the biosynthetic assembly line, as reported previously from other cyanobacterial lipopeptides (12). Notably, in *Rivularia*, the ratio between cyanochelin variants possessing different hydrophobic moieties varies between biomass and the culture medium, suggesting they may also exhibit distinct or complementary functions during the iron acquisition. Keeping the siderophore close to the producer prevents losses of the metabolite and may be especially important considering its photolability. Ferrous iron released during photolytic cleavage of the siderophore is likely to be con-

FIG 5 Legend (Continued)

gene clusters (highlighted in blue) are distributed throughout the cyanobacterial tree of life. The phylogenetic tree was inferred using approximately maximum-likelihood in FastTree2 based on a concatenated alignment of 120 universally conserved bacterial proteins employing the Genome Taxonomy Database toolkit *de novo* workflow. The phylogeny includes 607 publicly available and quality-checked cyanobacterial genome assemblies. The tree is rooted by *Vampirovibrio chlorellavorus* from a lineage sister to photosynthetic cyanobacteria; bootstrap supports are shown at nodes (100% support is marked with asterisks).

sumed by other organisms unless the reaction happens in close proximity to the producer. On the other hand, siderophores diffusing further away may present a service to the whole community (20).

The discovery of cyanochelins substantially broadens the collection of cyanobacterial siderophores and provides valuable waypoints for understanding of the orphan gene clusters in this specific phylum. We show that genetic elements encoding synthesis of acylated peptides with two β -OH-Asp functioning as siderophores are widespread in cyanobacteria. In detail, however, the presence of photolabile siderophores in cyanobacteria as phototrophic organisms is a bit surprising and raises intriguing questions about the possible recipients of the benefits implied by the siderophore presence in the environment. The nonchelating residues can provide specific recognition patterns to ensure exclusive uptake of the Fe-siderophore by the producer. In contrast, the ferrous iron generated by photolytic cleavage of the compound is much harder to monopolize in complex microbial communities. Identifying the recipient of ferrous iron fluxes generated by the photolysis of cyanobacterial photoreactive siderophores will require innovative approaches to cyanobacterial cultivation methods, most likely along with other members of their natural microbial communities.

MATERIALS AND METHODS

Bioinformatic analysis and genome mining. Gene clusters from *Rivularia*, *Leptolyngbya*, and *Rubidibacter* (CP003549, AP017310, and ASSJ01000034.1, respectively) (for nucleotide positions see Table S1 to S3) were selected from the previous study (13) and interpreted by identification of the closest known functional homologue in the MIBiG repository 2.0 (21) in combination with antiSMASH 5.0.0 (22) and functional analysis of individual domains of multidomain proteins by CDD (23). Promoter regions responsive to FUR (ferric uptake regulator) were identified by sequence comparison to a weight matrix (15).

FAAL domains from proteins AFY58527, BAU16016, and ERN42091, found to be encoded at the beginning of the candidate gene clusters, were aligned by MAFFT (24) v. 7.308 (as provided via Geneious 10.0.9 MAFFT plugin 1.3.6) and used as a template for PSSM creation and subsequent PSI-BLAST (25), a search taxonomically limited to cyanobacteria. The search from 27 August 2019 with bitscore cutoff set to 250 yielded 943 hits. Regions spanning 40 kbp upstream and 80 kbp downstream from the start of the hit-encoding gene were obtained and inspected for the presence of two β -OH-Asp-specific NRPS modules. A module was evaluated as β -OH-Asp specific if it included an A-domain specific for Asp/Asn and a C-domain aligning to PSSM of β -OH-Asp-specific C-domains constructed from *Rivularia*, *Leptolyngbya*, and *Rubidibacter* with a bitscore of >300. Clusters encoding two such modules separated by a single NRPS module, along with a FAAL domain, from the first PSI-BLAST search were considered cyanochelin-like clusters. Cluster surroundings were inspected for the presence of siderophore transport-related genes by antiSMASH and by identification of the closest known functional homologue.

Phylogenomic analysis. To assess the phylogenetic distribution of cyanobacterial strains containing cyanochelin-like gene clusters, a phylogenomic tree was constructed utilizing the *de novo* workflow available in the Genome Taxonomy Database toolkit (26, 27). The GTDB-Tk v0.3.2 release from July 2019, offering 600 quality-checked cyanobacterial genomes, was used to produce an approximately maximum-likelihood tree (FastTree2-v2.1.9) (28) based on 120 concatenated conserved bacterial markers, and rooted by the closely related outgroup *Vampirovibrio chlorellavorus*.

Strains and cultivation. *Rivularia* sp. strain PCC 7116 and *Leptolyngbya* sp. strain NIES-3755 were obtained from respective culture collections. *Rubidibacter lacunae* strain KORDI 51-2 was kindly provided by Dong Han Choi (Korea Institute of Ocean Science and Technology, Republic of Korea). Strains were grown in BG-11 medium, with *Rivularia* and *Rubidibacter* requiring an addition of Turks Islands Salts (TIS) solution (NaCl 112 g/liter, KCl 2.68 g/liter, MgSO₄·7 H₂O 27.7 g, CaCl₂·2 H₂O 5.8 g/liter) at a ratio of 3:1. All strains were grown under ambient temperature (21 ± 2°C) and continuous dispersed light.

Isolation of cyanochelins. Freeze-dried iron-starved biomass of *Rivularia* sp. PCC 7116 was extracted using 60% acetonitrile (ACN) three times in final biomass-to-solvent ratio of 10 g/liter. The crude extract was subjected to liquid-liquid partitioning using an equal volume of ethyl acetate in a separatory funnel. The retained water phase containing the compound was acidified using formic acid to pH 3.5 and subjected to a second liquid-liquid partitioning step using four volumes of ethyl acetate/ACN (in ratio 3:5). The drop in pH caused the compound to prefer the organic phase, which was subsequently dried, dissolved in dimethyl sulfoxide (DMSO), and subjected to preparative reverse phase HPLC on a phenyl column (μ Bondapak Phenyl, 10 μ m, 125 Å, 7.8 × 300 mm). Water (A) and ACN (B), both acidified using formic acid to pH 2, were used as solvents in the following gradient: 0 min 15% B, 1 min 15% B, 8 min 37% B, 31 min 54% B, 35 min 100% B, 40 min 100% B at flow rate 2.5 ml/min. The diode array detector (DAD) was set to 254 nm and used for detection of the target compounds. Cyanochelin A eluted at 30 min and its variant 753.33 eluted at 25 min. The isolation procedure yielded 2.2 mg of cyanochelin A from *Rivularia* sp. PCC 7116; the variant 753.33 was not collected in a sufficient amount that it could be weighed. Analogously, cyanochelins from *Leptolyngbya* and *Rubidibacter* were isolated for iron chelation experiments using the identical semipreparative column, flow rate, and detection settings as described above. For *Leptolyngbya*, the following gradient was applied. (A) water/(B) ACN (both

acidified using 0.1% formic acid): 0 min 15% B, 8 min 48% B, 21 min 56% B, 26 min 100% B, 31 min 100%, and 35 min 15%. The compound of interest (m/z 1,026.51) eluted at 24.2 min. For *Rubidibacter*, the following gradient was used: (A) 25% ACN/(B) 96% ACN (both acidified using trifluoroacetic acid): 0 min 0% B, 1 min 0% B, 6 min 30% B, 46 min 100% B, 51 min 100% B, 52 min 0%. The compound of interest (m/z 1,045.55) eluted at 24.5 min. The compounds from *Leptolyngbya* and *Rubidibacter* were collected in insufficient amounts to be weighed. The purity of the collected fractions was confirmed by HPLC-HRMS/MS analysis by the method described below.

Structural characterization of cyanochelin A. NMR experiments were performed at 25°C on a Bruker AvanceNeo 700 MHz spectrometer (Billerica, MA, US) equipped with a triple resonance CHN cryoprobe, using CD₃OD or *D*-DMSO (Sigma-Aldrich, Milan, Italy) as solvents and the 1D and 2D standard pulse sequences provided by the manufacturer. ¹H chemical shifts were referenced to the residual solvents' protons resonating at 3.31 (CHD₂OD) or 2.50 (CHD₂SOCD₃) ppm. ¹³C-NMR chemical shifts were referenced to the solvents' methyl carbons resonating at 49.01 (CD₃OD) or 39.51 ppm (*D*-DMSO). Abbreviations for signal couplings are as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, b = broad.

Iron depletion experiment. *Rivularia* biomass homogenized by pipetting was distributed equally into three flasks of each standard or iron-deprived medium supplemented with 10 μM desferrioxamine B. Expression of the candidate gene cluster along with content of cyanochelins in the medium and biomass was monitored at six time points of 0, 2, 4, 7, 9, and 11 days. Solid phase extractions (SPE) of supernatants were performed on a Supelco Discovery DSC-18 cartridge (500 mg, 3 ml). The spent media samples (5 ml) were loaded on preconditioned SPE cartridges (two volumes of methanol and one volume of water) and eluted to 1 ml of 100% methanol. Samples of biomass were spun at 1,600 × *g* for 5 min to remove residual supernatant and either supplied with 1 ml of PGTX and frozen at -80°C for RNA extraction, or frozen, freeze-dried, weighted, and extracted using 60% ACN in a biomass-to-solvent ratio of 4 mg/ml for HPLC-HRMS/MS analysis (see below). RNA was isolated using Direct-zol RNAMiniprep (Zymo Research) and genomic DNA was degraded with the TURBO DNA-free kit (Life Technologies) according to the manufacturer's protocol. cDNA was generated from 500 ng of total RNA using Transcriptor reverse transcriptase (Roche). A specific primer pair was designed to amplify *ccsA* (q_AFY58527-R [CCAAGCTA TAGTCGCTCTAA] and -F [ATACGGTTTGGATTCACTGAT]), while primer pair *rnpB*-F (CTGCTGGTGCCTC TTACC) and *rnpB*-R (GTGAGGATAGTGCCACAGAA) was used to amplify a housekeeping reference gene (*rnpB*). Reverse transcriptase quantitative PCR (RT-qPCR) was performed using LightCycler 96 (Roche) and FastStart Essential DNA Green Master (Roche). A 20-μl reaction contained 10 μl FastStart Essential DNA Green Master (Roche), 500 nM primers, and 1 μl of cDNA. qPCR cycling parameters included preincubation at 1 min for 95°C, 40 times 95°C for 10 s, 60°C for 10 s, and 72°C for 20 s, melting at 95°C for 10 s, 65°C for 60 s, 97°C for 1 s, with a final cooling step to 37°C. The relative gene expression of *ccsA* was normalized to *rnpB* and expressed as an average of three replicates.

HPLC-HRMS/MS. The biomass samples in batch cultures, as well as those of media and biomass obtained from iron-deprivation experiments, were analyzed using the HPLC system (Dionex Ultimate 3000) connected to HRMS with electrospray ionization (Bruker Impact HD II). The separation was performed on the reverse-phase phenyl column (μBondapak Phenyl, 10 μm, 125 Å, 7.8 × 300 mm) using an ACN/water gradient (A/B) of 0 min 15/85%, 1 min 15/85%, 25 min 100/0%, 30 min 100/0%, and 33 min 15/85% at a flow rate of 0.6 ml/min. To ensure that the vast majority of the compounds were present in the nonchelated form, the pH of the mobile phase was kept below 3.0 by the addition of trifluoroacetic acid into the mobile phases (0.1% of TFA in water and 0.002% of TFA in ACN). The following settings for the mass spectrometer were used: dry temperature 200°C; drying gas flow 12 liters/min; nebulizer 3 bar; capillary voltage 4,500 V; endplate offset 500 V. The spectra were collected in the range of 100 to 1,500 m/z with a spectral rate of 0.5 Hz. Collision-induced dissociation (CID) was set as a ramp from 20 to 80 eV on masses of 200 to 1,500, respectively, using the automated precursor ion selection. Calibration of the instrument was performed using CH₃COONa clusters at the beginning of each analysis. For analysis where elemental composition analysis was required, lock mass (m/z 622) was added to the ion source. The individual spectra were analyzed using Bruker Data Analysis software (version 5.2) to obtain elemental formulas. Manual direct injections were applied to obtain high-quality fragmentation spectra, as well as to test the cyanochelin capability to bind iron. Cyanochelin A (25 μg/ml) in 50% MeOH/water (vol/vol) with 0.5% formic acid and lockmass (0.2 μg/ml) was injected to the ionization source using an automated syringe at a flow rate of 1,200 μl/h. To test whether it would bind Fe³⁺, the solution was supplied with approximately 1.5 to 2 molar equivalents of FeCl₃. The following settings of the mass spectrometer were used: dry temperature 200°C; drying gas flow 3 liters/min; nebulizer 0.5 bar; capillary voltage 4,500 V; endplate offset 500 V. The compound of interest was isolated in quadrupole using an isolation window of 10 Da and fragmented at different fragmentation energies (20 to 80 eV). Analogously, cyanochelins from *Leptolyngbya* and *Rubidibacter* were analyzed with the compound concentration roughly estimated based on the comparison to the HRMS signal of the cyanochelin A.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 1.9 MB.

ACKNOWLEDGMENTS

The study was supported by the Czech Science Foundation projects "Bioactive cyanobacterial lipopeptides: genome mining, detection, and structure-activity relationships"

(number 16-09381S) and “Co-existence of quorum sensing and quorum sensing inhibition in cyanobacteria: Chemical warfare or cooperation?” (number 19-17868Y), the Ministry of Education, Youth and Sports of the Czech Republic, National Program of Sustainability I, number LO1416 and RVO 67985939, and by the Ministry of Regional Development of the Czech Republic—Cross-Border cooperation Czech-Bavaria project no. 41. N.B., A.C., G.E., and V.C. were supported by Regione Campania, PO FESR2014 to 2020, O.S. 1.2, Project “Campania Oncoterapie” no. B61G18000470007.

We declare no competing interests.

REFERENCES

- Sandy M, Butler A. 2009. Microbial iron acquisition: marine and terrestrial siderophores. *Chem Rev* 109:4580–4595. <https://doi.org/10.1021/cr9002787>.
- Årstøl E, Hohmann-Marriott MF. 2019. Cyanobacterial siderophores—physiology, structure, biosynthesis, and applications. *Mar Drugs* 17:281. <https://doi.org/10.3390/md17050281>.
- Morrissey J, Bowler C. 2012. Iron utilization in marine cyanobacteria and eukaryotic algae. *Front Microbiol* 3:43. <https://doi.org/10.3389/fmicb.2012.00043>.
- Shih P, Wu D, Latifi A, Axen S, Fewer D, Talla E, Calteau A, Cai F, de Marsac N, Rippka R, Herdman M, Sivonen K, Coursin T, Laurent T, Goodwin L, Nolan M, Davenport K, Han C, Rubin E, Eisen J, Woyke T, Gugger M, Kerfeld C. 2013. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc Natl Acad Sci U S A* 110:1053–1058. <https://doi.org/10.1073/pnas.1217107110>.
- Gross H. 2007. Strategies to unravel the function of orphan biosynthesis pathways: recent examples and future prospects. *Appl Microbiol Biotechnol* 75:267–277. <https://doi.org/10.1007/s00253-007-0900-5>.
- Butler A, Theisen RM. 2010. Iron(III)-siderophore coordination chemistry: reactivity of marine siderophores. *Coord Chem Rev* 254:288–296. <https://doi.org/10.1016/j.ccr.2009.09.010>.
- Kreutzer MF, Kage H, Nett M. 2012. Structure and biosynthetic assembly of cupriachelin, a photoreactive siderophore from the bioplastic producer *Cupriavidus necator* H16. *J Am Chem Soc* 134:5415–5422. <https://doi.org/10.1021/ja300620z>.
- Rosconi F, Davyt D, Martínez V, Martínez M, Abin-Carriquiry JA, Zane H, Butler A, de Souza EM, Fabiano E. 2013. Identification and structural characterization of serobactins, a suite of lipopeptide siderophores produced by the grass endophyte *Herbaspirillum seropedicae*. *Environ Microbiol* 15:916–927. <https://doi.org/10.1111/1462-2920.12075>.
- Hardy CD, Butler A. 2019. Ambiguity of NRPS structure predictions: four bidentate chelating groups in the siderophore pacifibactin. *J Nat Prod* 82:990–997. <https://doi.org/10.1021/acs.jnatprod.8b01073>.
- Singh GM, Fortin PD, Koglin A, Walsh CT. 2008. beta-Hydroxylation of the aspartyl residue in the cytotoxin syringomycin E: characterization of two candidate hydroxylases AspH and SyrP in *Pseudomonas syringae*. *Biochemistry* 47:11310–11320. <https://doi.org/10.1021/bi801322z>.
- Boiteau RM, Mende DR, Hawco NJ, McIlvin MR, Fitzsimmons JN, Saito MA, Sedwick PN, DeLong EF, Repeta DJ. 2016. Siderophore-based microbial adaptations to iron scarcity across the eastern Pacific Ocean. *Proc Natl Acad Sci U S A* 113:14237–14242. <https://doi.org/10.1073/pnas.1608594113>.
- Mareš J, Hájek J, Urajová P, Kust A, Jokela J, Saurav K, Galica T, Čapková K, Mattila A, Haapaniemi E, Permi P, Myrsterud I, Skulberg OM, Karlsen J, Fewer DP, Sivonen K, Tønnesen HH, Hrouzek P. 2019. Alternative biosynthetic starter units enhance the structural diversity of cyanobacterial lipopeptides. *Appl Environ Microbiol* 85:e02675-18. <https://doi.org/10.1128/AEM.02675-18>.
- Galica T, Hrouzek P, Mareš J. 2017. Genome mining reveals high incidence of putative lipopeptide biosynthesis NRPS/PKS clusters containing fatty acyl-AMP ligase genes in biofilm-forming cyanobacteria. *J Phycol* 53:985–998. <https://doi.org/10.1111/jpy.12555>.
- Manavalan B, Murugapiran SK, Lee G, Choi S. 2010. Molecular modeling of the reductase domain to elucidate the reaction mechanism of reduction of peptidyl thioester into its corresponding alcohol in non-ribosomal peptide synthetases. *BMC Struct Biol* 10:1. <https://doi.org/10.1186/1472-6807-10-1>.
- González A, Angarica VE, Sancho J, Fillat MF. 2014. The FurA regulon in *Anabaena* sp. PCC 7120: in silico prediction and experimental validation of novel target genes. *Nucleic Acids Res* 42:4833–4846. <https://doi.org/10.1093/nar/gku123>.
- Carroll CS, Moore MM. 2018. Ironing out siderophore biosynthesis: a review of non-ribosomal peptide synthetase (NRPS)-independent siderophore synthetases. *Crit Rev Biochem Mol Biol* 53:356–381. <https://doi.org/10.1080/10409238.2018.1476449>.
- Reid RT, Live DH, Faulkner DJ, Butler A. 1993. A siderophore from a marine bacterium with an exceptional ferric ion affinity constant. *Nature* 366:455–458. <https://doi.org/10.1038/366455a0>.
- Sritharan M. 2016. Iron homeostasis in *Mycobacterium tuberculosis*: mechanistic insights into siderophore-mediated iron uptake. *J Bacteriol* 198:2399–2409. <https://doi.org/10.1128/JB.00359-16>.
- Kem MP, Butler A. 2015. Acyl peptidic siderophores: structures, biosyntheses and post-assembly modifications. *Biomaterials* 28:445–459. <https://doi.org/10.1007/s10534-015-9827-y>.
- Barbeau K, Rue EL, Bruland KW, Butler A. 2001. Photochemical cycling of iron in the surface ocean mediated by microbial iron(III)-binding ligands. *Nature* 413:409–413. <https://doi.org/10.1038/35096545>.
- Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, de Bruijn I, Chooi YH, Claesen J, Coates RC, Cruz-Morales P, Duddela S, Düsterhus S, Edwards DJ, Fewer DP, Garg N, Geiger C, Gomez-Escribano JP, Greule A, Hadjiiothas M, Haines AS, Helfrich EJN, Hillwig ML, Ishida K, Jones AC, Jones CS, Jungmann K, Kegler C, Kim HU, Kötter P, Krug D, Masschelein J, Melnik AV, Mantovani SM, Monroe EA, Moore M, Moss N, Nützmann H-W, Pan G, Pati A, Petras D, Reen FJ, Rosconi F, Rui Z, Tian Z, Tobias NJ, Tsunematsu Y, Wiemann P, Wyckoff E, Yan X, et al. 2015. Minimum information about a biosynthetic gene cluster. *Nat Chem Biol* 11:625–631. <https://doi.org/10.1038/nchembio.1890>.
- Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 47:W81–W87. <https://doi.org/10.1093/nar/gkz310>.
- Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43:D222–D226. <https://doi.org/10.1093/nar/gku1221>.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
- Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 36:996–1004. <https://doi.org/10.1038/nbt.4229>.
- Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/btz848>.
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.