How to cite: *Angew. Chem. Int. Ed.* **2021**, *60*, 10064–10072
International Edition: doi.org/10.1002/anie.202015105
German Edition: doi.org/10.1002/ange.202015105

Natural Products Discovery

Discovery of Cyanobacterial Natural Products Containing Fatty Acid Residues**

*Sandra A. C. Figueiredo, Marco Preto, Gabriela Moreira, Teresa P. Martins, Kathleen Abt, André Melo, Vitor M. Vasconcelos, and Pedro N. Leão**

Abstract: *In recent years, extensive sequencing and annotation of bacterial genomes has revealed an unexpectedly large number of secondary metabolite biosynthetic gene clusters whose products are yet to be discovered. For example, cyanobacterial genomes contain a variety of gene clusters that likely incorporate fatty acid derived moieties, but for most cases we lack the knowledge and tools to effectively predict or detect the encoded natural products. Here, we exploit the apparent absence of a functional β-oxidation pathway in cyanobacteria to achieve efficient stable-isotope-labeling of their fatty acid derived lipidome. We show that supplementation of cyanobacterial cultures with deuterated fatty acids can be used to easily detect natural product signatures in individual strains. The utility of this strategy is demonstrated in two cultured cyanobacteria by uncovering analogues of the multidrug-resistance reverting hapalosin, and novel, cytotoxic, lactylate-nocuolin A hybrids—the nocuolactylates.*

Introduction

It is now evident that we currently grasp but a small fraction of the small-molecule diversity produced by bacteria.[1,2] Revealing the compounds encoded in orphan biosynthetic gene clusters (BGCs) is a central theme in current natural products (NPs) research.[1–3] Genome-guided NP discovery relies on the ability to recognize BGCs in genome data and predict, at least partially, the function of their genes. This implies that NPs encoded by yet-unrecognized biosynthetic pathways are likely to remain hidden to genome-guided discovery.[1] In addition, BGCs with unusual organization or with unknown functionalities can be difficult to deorphanize through structural predictions.[4] In contrast, genome-data independent NP discovery strategies, in particular those making use of metabolomics, have the potential to reveal any kind of metabolite (including those with entirely unprecedented structure or biosynthetic logic), present in a given biological sample.[5] As an example, molecular networking strategies such as Global Natural Product Social Molecular Networking (GNPS)[6] have enabled the dereplication of known compounds and the identification and prioritization of potentially new NPs for isolation and structure elucidation.[7] However, annotation of the large amount of data generated in a single metabolomics experiment, in particular from less-studied organisms, is still challenging.[4]

Cyanobacteria are among the chemically rich bacterial phyla.[8] They occupy a differentiated ecological niche, which is reflected on their unique secondary metabolism.[9,10] Many cyanobacterial NPs have potent bioactivities and some are in clinical use, clinical trials or preclinical development.[9–11] Most known cyanobacterial NPs were discovered using traditional bioassay-guided isolation.[12] Still, from genomic data, it is clear that the vast majority of cyanobacterial NPs are yet to be found.[13] Unlike other NP-rich bacterial groups, genetic manipulation of cyanobacteria is challenging and lacks tools, despite promising recent efforts.[14] Hence, there is a need for strategies that can facilitate the discovery of the hidden majority of cyanobacterial NPs. Several stable-isotope-labeled precursor incorporation methods have been applied successfully to this end.[15–17]

A striking feature of cyanobacterial NPs is the common presence of fatty acid (FA) derived moieties.[12,18] These building blocks can be incorporated into secondary metabolites through the action of pathway-specific fatty acyl-AMP ligases (FAALs) that load the fatty acyl units into their cognate acyl carrier proteins (ACPs).[19,20] FAs can also be recruited from primary metabolism and incorporated into

[*] Dr. S. A. C. Figueiredo, Dr. M. Preto, G. Moreira, T. P. Martins, K. Abt, Prof. Dr. V. M. Vasconcelos, Dr. P. N. Leão
Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto
Avenida General Norton de Matos, s/n
4450-208 Matosinhos (Portugal)
E-mail: pleao@ciimar.up.pt
T. P. Martins, K. Abt
Institute of Biomedical Sciences Abel Salazar (ICBAS)
University of Porto
Rua de Jorge Viterbo Ferreira, 228, 4050-313 Porto (Portugal)
Prof. Dr. A. Melo
LAQV@REQUIMTE/ Department of Chemistry and Biochemistry
Faculty of Sciences, University of Porto
Rua do Campo Alegre, 4169-007 Porto (Portugal)
Prof. Dr. V. M. Vasconcelos
Department of Biology, Faculty of Sciences, University of Porto
Rua do Campo Alegre, 4169-007 Porto (Portugal)

[**] A previous version of this manuscript has been deposited on a preprint server (https://doi.org/10.26434/chemrxiv.13100564.v2).
Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under: https://doi.org/10.1002/anie.202015105.

cyanobacterial NPs by dialkylresorcinol-condensing enzymes or type III PKSs in the absence of a pathway-specific ACP/ACP-domain,[18] or by lipoyltransferases.[21] Free FAs can also be enzymatically esterified with alkyl halide bearing NPs.[22] Such FA-incorporating enzymes are abundant in cyanobacterial genomes (Figure S1), often within orphan BGCs[18] and represent opportunities for NP discovery. Nevertheless, it is difficult to predict the structure of NPs produced by most cyanobacterial FA-incorporating BGCs. For one, the exact type and size of fatty acyl moiety loaded by these enzymes cannot be accurately predicted.[23] Adding to this, different biosynthetic logics, tailoring enzymes and hypothetical proteins are found associated with these FA-incorporating enzymes.[18,22]

Cyanobacterial FA metabolism has another particular aspect: apparently, a functional β-oxidation pathway is not found in these organisms, as briefly noted by von Berlepsch and co-workers.[24] In our studies on bartoloside biosynthesis, we could not detect most of the β-oxidation enzyme-encoding genes in the draft genome data of *Synechocystis* sp. LEGE 06155.[25] Mills et al.[26] did not find a β-oxidation gene repertoire in *Synechocytsis* sp. PCC 6803. The same observation was made by Burkart et al.,[27] while suggesting that an alternative FA catabolism pathway might be present in this model cyanobacterium.

Considering both the apparent inability of cyanobacteria to carry out canonical β-oxidation of FAs and their capacity to elongate exogenously supplied FAs,[27,28] we envisioned that extensive labeling of their FA-derived lipidome (Figure S2) could be achieved using stable-isotope-labeled precursors (as previously observed for *Synechocystis* sp. PCC 6803[27]). In contrast, β-oxidation has been described for other bacteria[29] and several bacterial groups are unable to elongate exogenous FAs.[30] We recognized that lipidome labeling could be leveraged to uncover new cyanobacterial NPs in a lipid-version of the genomisotopic approach,[31] by targeting FA-incorporating BGCs (Figure S2). Metabolites incorporating FAs would be detected by LC-MS on the basis of their larger *m*/*z* isotope clusters. This strategy could, in principle, also be applied to the untargeted discovery of new metabolites without prior biosynthetic knowledge (Figure S2).

Here, we confirm that most cyanobacterial genomes indeed lack the genes encoding one or more of the key β-oxidation enzymes. We performed lipid analysis on several cyanobacterial strains and show that abundant labeling of lipids can be obtained through supplementation of their cultures with deuterated FAs. We used this approach to reveal new NPs, namely analogues of the multiple drug reversal resistance agent—hapalosin—from *Fischerella* sp. PCC 9431 and several 1,2,3-oxadiazine-lactylate hybrids—nocuolactylates—from *Nodularia* sp. LEGE 06071, with mild cytotoxic activity.

Results and Discussion

A canonical β-oxidation pathway is absent in most cyanobacteria. We initially aimed to clarify whether previous isolated observations or suggestions in the literature, namely that certain cyanobacteria are unable to degrade FAs through a canonical β-oxidation pathway,[24,26,27] were generalizable to the whole Cyanobacteria phylum. Previous work (Taylor,[32] doctoral thesis), provided substantial genomic and metabolomic evidence that this was the case. To confirm these findings, we searched, among the KEGG orthology-annotated genomes of cyanobacteria ($n = 129$), for the key genes in the β-oxidation pathway. Across the cyanobacterial representatives in the KEGG database, all strains appear to contain the *fadD* gene, which putatively encodes an acyl-CoA synthetase catalyzing the first step in β-oxidation.[29] However, all analyzed strains also lacked *fadB* (*fadJ*) and most did not encode *fadE, fadN* and *fadI* (*fadA*) or *atoB* homologs (Figure 1a). Among the expanded set of complete genomes deposited in the NCBI database, a similar picture emerges (Supporting Data 1). Overall, our findings from bioinformatics searches are consistent with an absent β-oxidation pathway in cyanobacteria.

To obtain experimental support for the lack of this catabolic pathway, we supplemented the cyanobacteria *Synechocystis* sp. PCC 6803, *Anabaena cylindrica* PCC 7122 and *Kamptonema formosum* PCC 6407 (representing different lineages of the phylum), with deuterated FAs of varying lengths (d_{11}-hexanoic, d_{23}-dodecanoic and d_{31}-hexadecanoic acids, hereafter d_{11}-C_6, d_{23}-C_{12} and d_{31}-C_{16}, respectively). For comparison, we performed supplementation with the same FAs in *E. coli* pET24d-*Synechococcus elongatus* PCC 7942 Aas (pET24d-*aas*7942), carrying a cyanobacterial acyl ACP synthetase (which expands the ability of *E. coli* to incorporate free FAs from C_{12} to C_{18})[28,33] and *E. coli* pET24d (empty vector) cultures. We then analyzed the labeling patterns of phosphatidylglycerols (PGs)—lipids shared by cyanobacteria and *E. coli.* As expected, supplementation with d_{11}-C_6, d_{23}-C_{12} or d_{31}-C_{16} led to incorporation of deuterium in cyanobacterial PG species, which exhibited isotope patterns with clear single or double incorporation of FA-derived moieties. In *E. coli* extracts, the incorporation was in lower abundance or in some cases not observed (Figure 1b, Figure S3). Surprisingly, an isotope pattern in PGs consistent with individual abstraction of deuterium atoms was observed for cyanobacterial strains in d_{23}-C_{12} and d_{31}-C_{16} supplementation experiments (Figure S3). This may suggest an alternative catabolic pathway for FAs or non-degradative exchange of deuterons and protons by a yet-unknown mechanism. We also observed persistent labeling of lipids when we cultured *Anabaena cylindrica* PCC 7122 under d_{11}-C_6 supplementation, after a 60-day culture period (Table S1). Taken together, the data from supplementation experiments show a much higher deuterium-label incorporation in the PGs of tested cyanobacteria than in the β-oxidation-capable *E. coli.* Exogenous FA-derived moieties were abundantly incorporated into membrane lipids, where they persist for a relatively long period, which is consistent with an absence of β-oxidation scenario.

Supplementation of cyanobacterial cultures with deuterated FAs leads to an abundantly labeled lipidome. We sought to explore the feasibility of the stable-isotope-labeled FA supplementation approach for NP discovery in cyanobacteria (Figure S2). We started by interrogating whether supplementation would result in labeling of the wider FA-derived

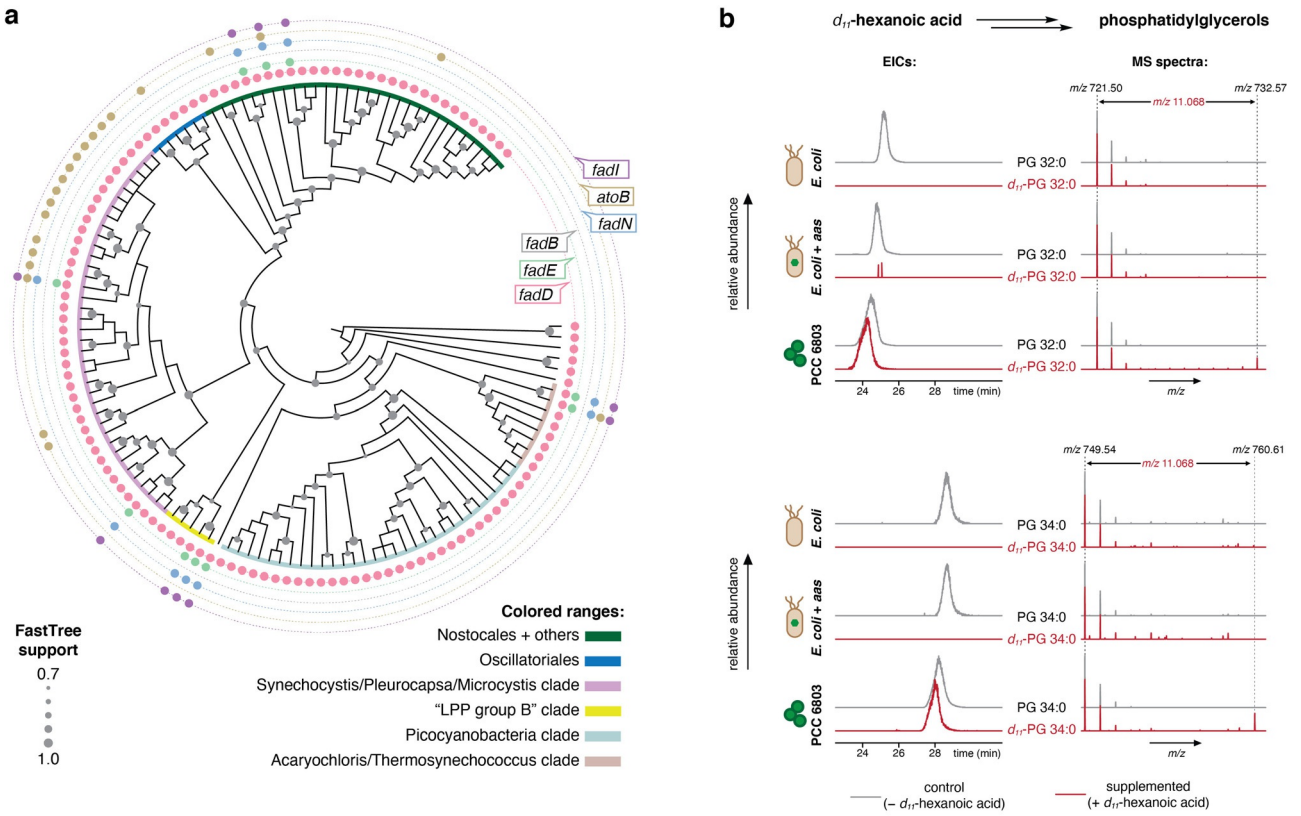

Figure 1. Cyanobacteria appear to lack a canonical β-oxidation pathway. a) Cladogram depicting the approximate maximum likelihood tree (computed with FastTree) of 16S rRNA genes from KEGG-annotated genomes of cyanobacteria, highlighting the presence or absence of each of the β-oxidation genes. b) LC-HRESIMS analysis of $[M-H]^-$ ions of phospholipids (phosphatidylglycerol groups PG 32:0 and PG 34:0) in extracts from cultures of *E. coli* BL21 pET24d and *E. coli* pET24d-aas7942, and in *Synechocystis* sp. PCC 6803, each cultured with or without d_{11}-C_6 supplementation, showing that no appreciable deuterium label is detected in *E. coli* strains.

lipidome. Because cyanobacteria will elongate exogenously supplied FAs, the smaller the supplemented deuterated FA, the higher the number of potential labeled metabolites. We opted for using d_{11}-C_6 supplementation in our experiments. While d_7-butyric acid could lead to detection of additional NPs, we anticipated that deuterons could be abstracted from the FA-derived portion of the metabolites during biosynthesis (e.g. unsaturation, tailoring), leading to substantial overlapping of isotope clusters from non-labeled and labeled metabolite species and complicating analysis. This would also be the case for d_3-acetate, together with extensive scrambling from its conversion to d_3-acetyl-CoA. Burkart et al.[27] also note that *Synechocystis* sp. PCC 6803 was not able to elongate short chain FA precursors (C_3-C_5). We thus added d_{11}-C_6 to cultures of six additional cyanobacteria strains from different orders. Because we had already observed extensive labeling for PG species, we looked instead for abundant cyanobacterial lipids such as glycolipids and sulfolipids. We detected the expected deuterium incorporation in these lipid families (Figure S4), indicating that supplementation of exogenous d_{11}-C_6 can be used to efficiently label FA-derived lipids in cyanobacteria.

Secondary metabolite labeling. Our next step was to understand whether this strategy could be generalized to the secondary metabolome, which is typically less abundant than membrane lipids. Nakamura et al.[34] showed that exogenously supplied deuterated FAs were incorporated into the cyanobacterial NP cylindrocyclophane F. We selected a set of cyanobacterial strains known to synthesize NPs with FA-derived moieties, namely the nocuolin A(**1**)-producing *Nodularia* sp. LEGE 06071,[35] the bartoloside A palmitates (**2a**/**2b**)-producing *Synechocystis salina* LEGE 06099,[22] and the hapalosin (or hapalosin A, **3**)-producing *Fischerella* sp. PCC 9431.[36] Upon supplementing *Nodularia* sp. LEGE 06071 cultures with d_{11}-C_6 (single supplementation of 0.1 mM), extraction and LC-HRESIMS analysis, we could detect d_{11}-**1**, but only at ca. 0.1 % abundance relative to **1**. We screened different supplementation regimes, initial cell densities and culturing times, and were able to increase the levels of d_{11}-**1** to ≈ 9 % of **1** (Figure 2a), when using a culture with a chlorophyll *a* content of 2.5 μg mL^{-1}, supplemented with d_{11}-C_6 by three pulses to a final concentration of 0.5 mM during 7 days. Under similar conditions, d_{11}-labeled versions of **2a**/**2b** and **3** were observed at comparable levels (≈ 12 % and ≈ 9 %, respectively) (Figure 2b,c). Overall, these results show that our strategy leads to labeling of different secondary metabolites in easily detectable amounts. As can be expected for metabolites under the control of distinct regulatory networks and with different turnovers, no single supplementation/cultivation strategy performed best for all strains/lipids. In our qualitative assessment of the different supplementation conditions, pulsed supplementation and culturing times of 7–15

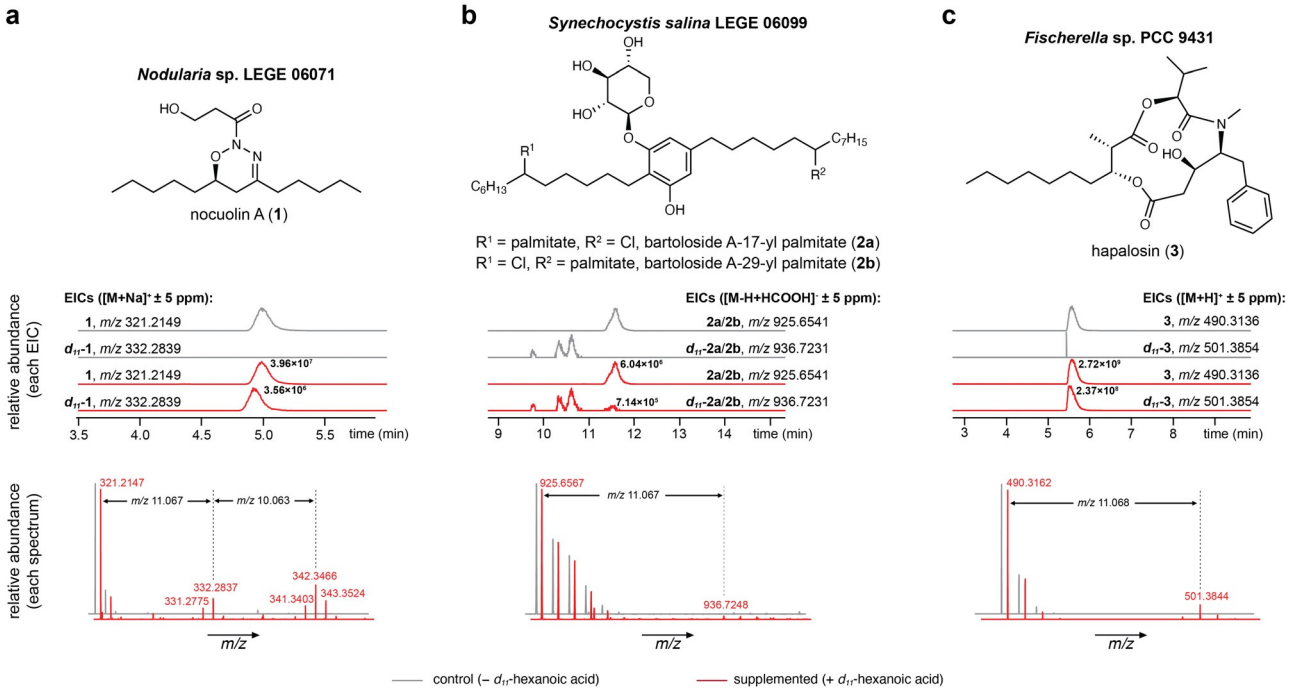

Figure 2. Supplementation of cyanobacteria with d_{11}-C_6 leads to labeling of different secondary metabolites. LC-HRESIMS-based detection of labeled **1** (a), **2a**/**2b** (b) and **3** (c) in d_{11}-C_6 supplemented cultures of *Nodularia* sp. LEGE 06071, *Synechocystis salina* LEGE 06099 and *Fischerella* sp. PCC 9431, respectively. Values next to chromatographic peaks of compounds **1**, **2a**/**2b** and **3** in EICs from d_{11}-C_6 supplemented cultures correspond to peak heights (ion counts).

days with shaking seemed to be important factors influencing labeling extent among the secondary metabolites analyzed (Table S2). The optimized conditions detailed above for *Nodularia* sp. LEGE 06071 were used for all subsequent supplementations.

Discovery of hapalosins C–G. Next, we used a comparative metabolomics approach to detect new NPs among the pool of labeled metabolites for a particular strain. The previously acquired LC-HRESIMS data for *Fischerella* sp. PCC 9431 cultures supplemented with d_{11}-C_6, was analyzed using MZmine2[37] and compared to non-supplemented controls. To expedite analysis, positive- and negative-mode exclusion lists of labeled metabolite features found in at least one strain among *Synechocystis* sp. PCC 6803, *Anabaena cylindrica* PCC 7122 or *Kamptonema formosum* PCC 6407 (Supporting Data 2 and 3) were used. These features were assumed to correspond mainly to primary lipids. From over 124 labeled MS features in positive mode (Supporting Data 4), we focused on a group of four unlabeled m/z values, with $[M+H]^+$ ions ranging from m/z 462 to 504, corresponding to unknown metabolite species incorporating the d_{11} label. Their retention times and mass defects suggested relatedness to **3**. The corresponding Extracted Ion Chromatograms (EICs) (Figure 3a), indicated the presence of more than four compounds; optimized LC gradient conditions revealed eight hapalosin congeners (Figure 3b, Figure S5). One of the compounds likely corresponds to the previously reported hapalosin B[38] (**4**); two were minor components and overlapped with major isobaric congeners, precluding their structure elucidation by MS/MS. The remaining five molecules are disclosed herein for the first time as hapalosins C-G (**5**–**9**). Their structures were deduced from comparative LC-HRESIMS/MS analysis (Text S1, Figure 3c) and precursor supplementation experiments (Text S1, Figure 3d, Figure S6).

By expanding the hapalosin class of cyanobacterial metabolites we show that this supplementation method can be used to swiftly detect new FA-incorporating NPs. Curiously, although compounds **6** and **9** had not been reported as NPs, they were synthesized as part of a medicinal chemistry study[39] and, like **3**, were found to have multidrug-resistant reversal activity and general cytotoxicity.

Discovery of the nocuolactylates. To further demonstrate the usefulness of the FA-supplementation method for NP discovery, we looked into the d_{11}-C_6 supplementation data for *Nodularia* sp. LEGE 06071. Through comparative metabolomics of supplemented and non-supplemented extracts (Figure S7), we detected a series of m/z values that were present only in supplemented cultures (Supporting Data 5). Their EICs showed close retention times and corresponded to labeled ($+9$, $+10$ or $+11$ Da) versions of compounds with m/z values ranging from 551 to 619 ($[M+H]^+$). From the respective isotopic patterns, one of the compounds was dichlorinated, one was monochlorinated and the other was not chlorinated/brominated. These metabolites incorporated up to three of the supplemented d_{11}-C_6 moieties (Figure 4a). Dereplication using several databases (SI Materials and Methods) did not retrieve NP matches. We pursued the isolation of the potentially new compounds to enable their Nuclear Magnetic Resonance (NMR)-based structure elucidation. Following large-scale (140 L) culturing of *Nodularia*

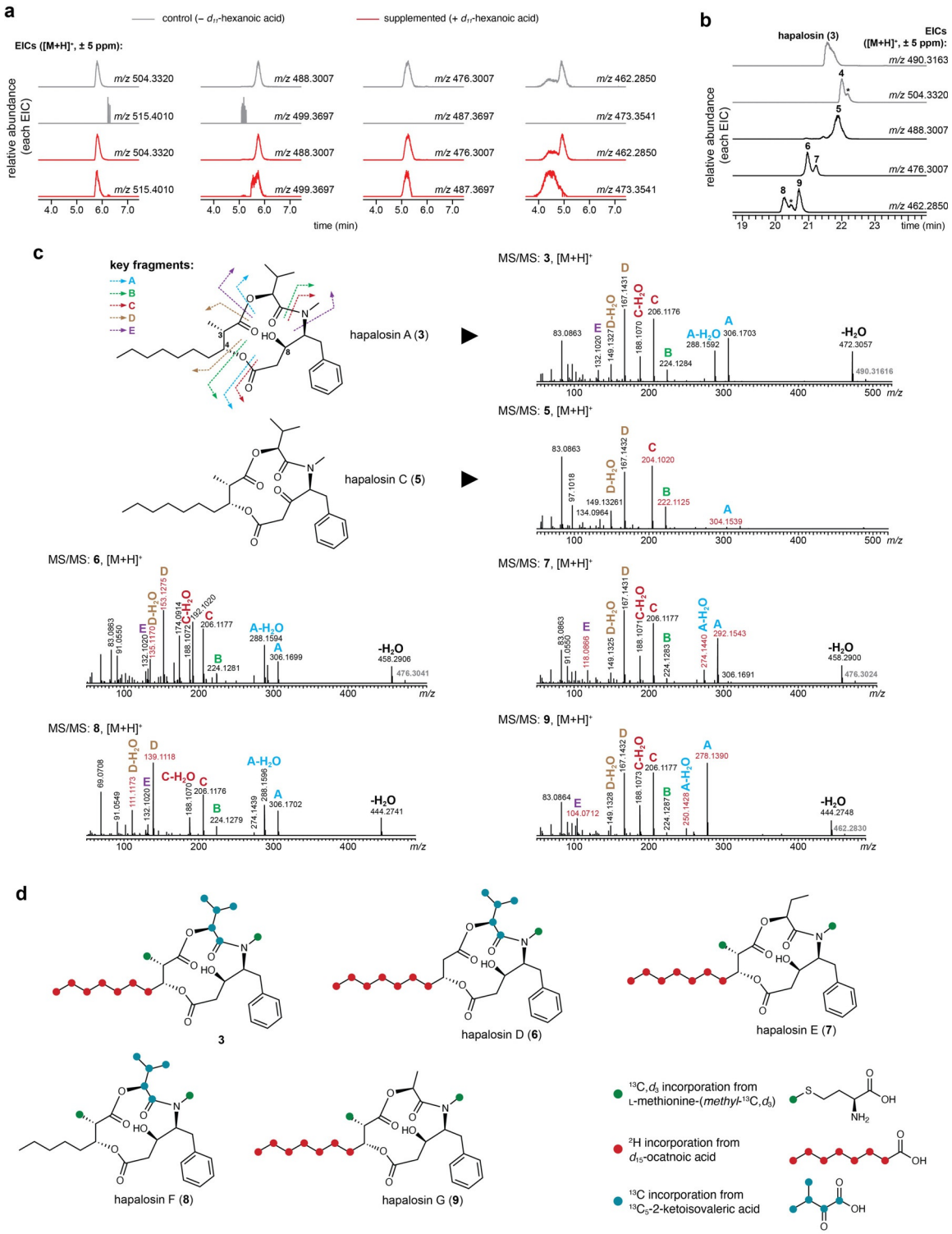

Figure 3. Discovery of hapalosin analogues following supplementation of *Fischerella* sp. PCC 9431 with d_{11}-C_6. a) LC-HRESIMS features corresponding to incorporation of a d_{11} label were detected in extracts of *Fischerella* sp. PCC 9431 supplemented with d_{11}-C_6. b) Optimized separation conditions in LC-HRESIMS analysis reveals eight metabolites associated with the detected features, including compounds **4**–**9**; EIC peaks marked with an asterisk (*) correspond to minor compounds with substantial overlap with isobaric metabolites and for which structure elucidation was not pursued. c) Structure elucidation of hapalosins **5**–**9** (see Text S1): LC-HRESIMS/MS spectra ($[M+H]^+$ ions) are shown for the new hapalosins and **3**, as a reference. Peak letters correspond to the double fragmentations indicated as "key fragments"; *m/z* values in red indicate diagnostic fragmentations supporting the structural proposals. d) Proposed labeling pattern for compounds **3**, **6**, **7**, **8** and **9**, deduced from supplementation studies with the indicated stable-isotope-labeled substrates.

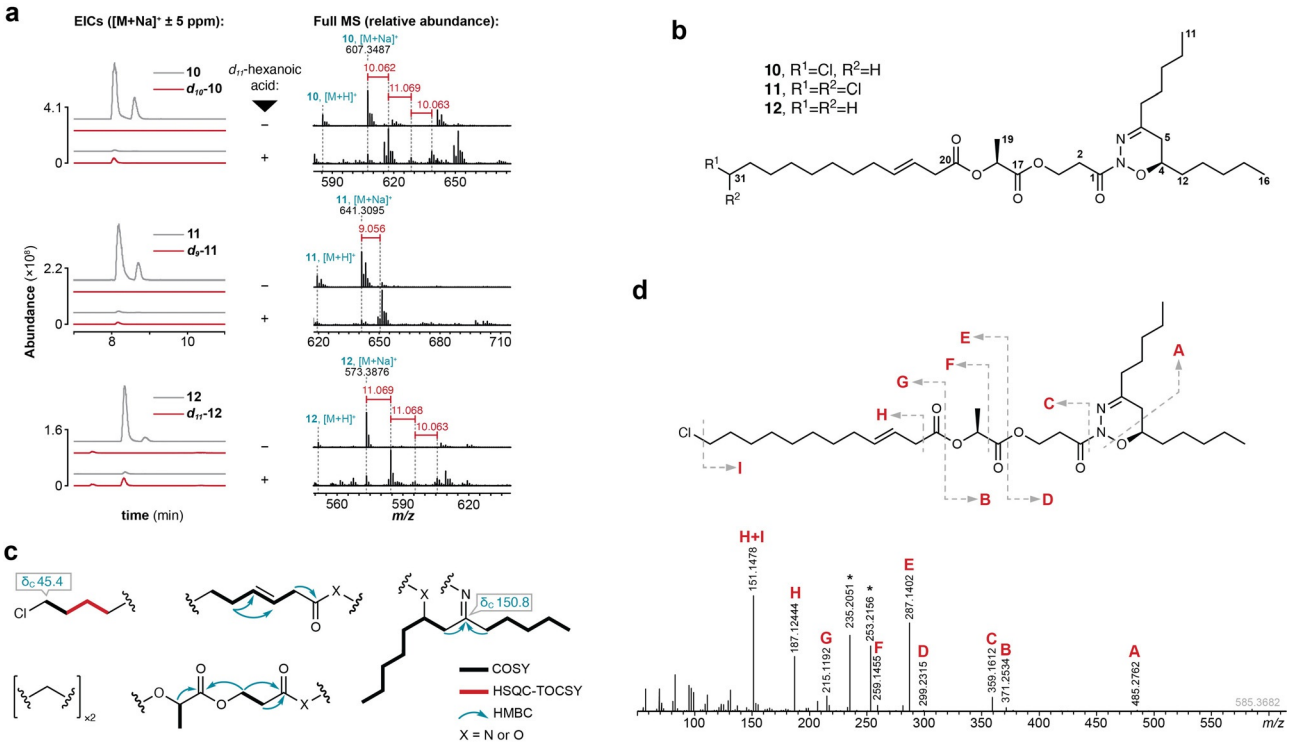

Figure 4. Discovery of nocuolactylates A–C (**10**–**12**). a) Detection of **10**–**12** following supplementation of *Nodularia* sp. LEGE 06071 with d_{11}-C_6, extraction and metabolome analysis, as described. Shown are EICs and the respective MS spectra, illustrating the difference between non-supplemented and supplemented extracts that enabled detection. b) Structure of **10**–**12**. c) NMR-derived substructures of **10** with key correlations and ^{13}C chemical shifts. d) HRESIMS/MS analysis of **10**, establishing the connectivity between the NMR-derived substructures (asterisks denote major fragments present in all nocuolactylates which result from fragmentation of the **1**-derived portion of the nocuolactylates and likely involve subsequent oxadiazine ring opening, rearrangement and concomitant N_2 loss, as previously proposed for **1**.[40]

sp. LEGE 06071, cell extraction using CH_2Cl_2/MeOH (2:1, v/v) and successive rounds of MS-guided fractionation, purified nocuolactylates A and B (**10**–**11**, 0.4 and 0.3 mg, respectively, Figure 4b) were obtained.

Compound **10** was isolated with highest purity (^1H NMR) and was selected for further characterization. A combination of 1D and 2D NMR experiments with HRESIMS/MS data was used to elucidate its planar structure (Figure 4c,d). Compound **10** showed an HRESIMS *m/z* of 585.3663 for its $[M+H]^+$ ion and an isotope pattern consistent with mono-chlorination, corresponding to a molecular formula $C_{31}H_{53}N_2O_6Cl$ (calcd 585.3665, $[M+H]^+$) and six degrees of unsaturation. Analysis of the ^1H, ^{13}C NMR and HSQC data in [D$_6$]DMSO for **10** (Table S3, Figures S8–S16) indicated the presence of an *E*-configured olefin ($\delta_H = 5.43/\delta_C = 134.2$ and $\delta_H = 5.57/\delta_C = 121.5$, $^3J_{HH} = \approx 14$ Hz), one halogenated methylene ($\delta_H = 3.62$, $\delta_C = 45.4$), one oxygenated methylene ($\delta_{Ha/b} = 4.36/4.28$, $\delta_C = 60.6$), two oxygenated methines ($\delta_H = 4.95$, $\delta_C = 68.4$ and $\delta_H = 4.02$, $\delta_C = 75.0$), three ester or amide functionalities ($\delta_C = 170.8$, 170.2 and 163.5), one imine ($\delta_C = 150.8$), along with three methyl groups and, as expected, a large number of methylenes, most within a prominent aliphatic envelope around $\delta_H = 1.35$–1.20 and $\delta_C = 30.0$–20.0. HMBC, COSY and HSQC-TOCSY correlations allowed the establishment of four substructures (Figure 4c), namely an aliphatic moiety with a terminal chlorination, a heteroatom-bound β,γ-unsaturated acyl, a lactyl group connected to a 3-hydroxy-propanoyl moiety and 6,8-disubstituted tridecane, one of the substituents being an imine. Connectivity between these substructures was obtained through HRESIMS/MS data (Figure 4d) and determined that the terminal alkyl chloride aliphatic moiety was connected to the 3,4-unsaturated acyl (in agreement with NMR data, which showed that both substructures were connected to the methylene envelope, Table S3). This, in turn, was found to be esterified to the 3-hydroxypropanoyl lactyl moiety, which itself connected to the disubstituted tridecane. To satisfy the molecular formula, it was necessary to assign the positioning of one nitrogen and one oxygen. We considered that the disubstituted tridecane moiety could be related to nocuolin A (**1**), which is produced by this cyanobacterium, because a 1,2,3-oxadiazine was consistent with the NMR and HRESIMS/MS data for **10**. Furthermore, in the proposed structure for **1**, the oxadiazine is connected to a 3-hydroxypropanoyl moiety. Gratifyingly, we detected an in-source fragment of **10** that matched the molecular formula of **1**; MS/MS data for this species was identical to the MS/MS data for purified **1** (Figure S17, see Text S2, Table S4 and Figures S18–S26 for details on the structure elucidation of **1** isolated from *Nodularia* sp. LEGE 06071 supporting the previously reported 1,2,3-oxadiazine moiety[40]). These findings clarified the nature of the substituted moiety and the proposed MS/MS-based connectivity for NMR-derived fragments, thereby establishing the planar structure of **10**.

Compound **11** presented an $[M+H]^+$ peak at *m/z* 619.3278 (HRESIMS), corresponding to a molecular formula of $C_{31}H_{52}N_2O_6Cl_2$ (calcd 619.3275, $[M+H]^+$), with an additional chlorine atom when compared to **10**. The 1D and 2D NMR data for **11** were similar to those for **10** (Table S5, Figures S27–S31), but indicated that, in place of a chloromethyl group, a dichloromethyl moiety was present in **11**. This assignment was supported by HRESIMS/MS analysis (Figure S32). In addition, we obtained 0.3 mg of partially purified nocuolactylate C (**12**) which, from HRESIMS data, had the molecular formula $C_{31}H_{54}N_2O_6$, (*m/z* 551.4053, $[M+H]^+$; calcd 551.4055). HRESIMS/MS data were entirely consistent with a non-chlorinated version of **10** and **11** (Figure S33). Finally, upon reinspection of the LC-HRESIMS and MS/MS data for a crude extract of *Nodularia* sp. LEGE 06071, we detected the two minor nocuolactylates D (**13**) and E (**14**) (Figure S34). These were also among the labeled features list (Supporting Data 5) and likely correspond to saturated versions of **10** and **12**, respectively (Figure S34).

The structures of nocuolactylates are reminiscent of two recently reported cyanobacterial metabolites: nocuolin A (**1**), produced by several Nostocales strains[40] and chlorosphaerolactylates, described from *Sphaerospermopsis* sp. LEGE 00249[41] (Figure 5a). A portion of the ≈40 kb BGC putatively assigned to **1** (*noc*), on the basis of comparative genomics,[40] was recently proposed to produce the chlorosphaerolactylates and renamed as *cly*.[42] We sequenced and searched the genome of *Nodularia* sp. LEGE 06071 for a *noc/cly* locus, which we found within a 226 kb contig. The *Sphaerospermopsis* sp. LEGE 00249 *cly* BGC encodes two halogenases, which is consistent with the two halogenated positions in the alkyl chain of the chlorosphaerolactylates (Figure 5b).[42] Likewise, the *Nodularia* sp. LEGE 06071 *cly* cluster encodes one halogenase, and a single halogenated carbon is present in **10** and **11** (Figure 5b). The *noc/cly* locus exhibits an overall similar architecture in these and other strains that produce **1**[40] (Figure 5b). Hence, the nocuolactylates could be products of the *noc/cly* BGCs, which co-localize in cyanobacterial genomes. However, further evidence is required to unequivocally connect *noc* genes to these metabolites.

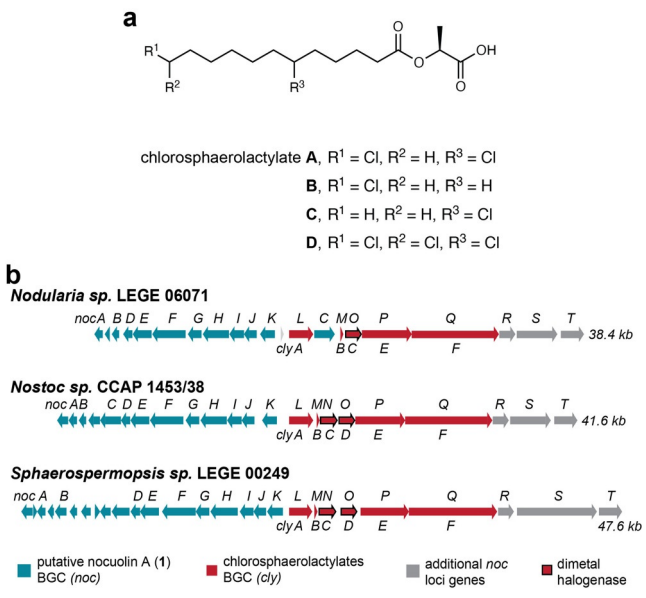

Figure 5. Putative biosynthesis of the nocuolactylates. a) Structure of the chlorosphaerolactylates, cyanobacterial secondary metabolites that are structurally related to the nocuolactylates. b) Architecture of the *noc* loci in the genomes of the cyanobacterial *Nodularia* sp. LEGE 06071, *Nostoc* sp. CCAP 1453/38 and *Sphaerospermopsis* sp. LEGE 00249, from where nocuolactylates, nocuolin A and chlorosphaerolacylates, respectively, were first reported.

To assign the configuration of the two stereogenic centers in **10–14**, we first focused on the nocuolin A (**1**)-related moiety. The configuration of the single chiral center in **1** had hitherto not been reported. Using pure **1**, we determined that the single stereogenic center in **1** exhibits an (*R*) configuration from comparison of theoretical and experimental ECD spectra (Figure S35). Given that **1** is identical to the non-lactylate portion of **10–14** and because these compounds are produced by the same organism, we propose a 4*R* configuration for **10–14**. Regarding the lactylate moiety of **10–14**, we used bioinformatics analysis to approach the configuration of position 18. In the structurally related chlorosphaerolactylates from *Sphaerospermopsis* sp. LEGE 00249, the lactyl stereogenic center was determined experimentally to be of *S* configuration.[41] This matched the bioinformatics prediction for the KR domain in the ClyF depsipeptide synthetase,[42] responsible for the generation of the stereogenic center. In *Nodularia* sp. LEGE 06071, ClyF-KR (83% identity, 94% similarity to ClyF-KR from *Sphaerospermopsis* sp. LEGE 00249) is also predicted to generate an *S*-configured product (Figure S36). With this, the absolute configuration of compounds **10–14** is proposed as 4*R*, 18*S*.

Compounds **10** and **11** were tested, in parallel with **1**, for cytotoxicity towards four cell lines, one non-tumor (hCMEC/D3) and three cancer cell lines (HCT-116, MG-63 and MCF7); **10** and **11** showed cytotoxicity with GR_{50} between 0.97–6.39 μM and IC_{50} between 0.77–22.20 μM but were less potent than **1** (Table S6). Furthermore, **10** produced a small inhibition halo in an agar disc diffusion assay against *Bacillus subtilis* (Figure S37), but the MIC could not be determined due to the limited amount of isolated metabolite. The function of these hybrids and of their individual units is unknown at this stage. One possibility is that **10–14** represent prodrugs of the more toxic **1**, as observed for example for zwittermicin A[43] and colibactin.[44] The discovery of additional molecules with the proposed 1,2,3-oxadiazine moiety, the potent activity, low yields and likely novel mode of action for **1**,[35] taken together with the elusiveness of this heterocyclic system[40,45] will hopefully motivate synthetic chemistry efforts towards this scaffold. Accessing 1,2,3-oxadiazines synthetically would also enable confirmation of the NMR- and MS/MS-based structure proposals for **1** and **10–14**. Alternatively, these compounds could be suitable for structure confirmation by the crystalline sponge method[46] or by Microcrystal Electron Diffraction.[47]

Conclusion

We report a new strategy for NP discovery in cyanobacteria. By applying this method to two cultured strains, we detected and elucidated the structure of five new hapalosin analogues and of a family of new cytotoxic lipids, the nocuolactylates. A considerable number of additional unknown metabolites were labeled and detected by our approach—we have yet to attempt their isolation. Analysis of further cyanobacteria strains will likely reveal multiple new targets for isolation. Albeit limited to a subset of NPs and to a particular group of organisms, this strategy enables efficient and fast discovery of metabolites featuring FA-derived moieties, which are abundant in cyanobacteria. It enables the targeted deorphanization of FA-incorporating BGCs of interest but also untargeted discovery efforts, can be used to swiftly screen multiple strains for isolation targets and can also be integrated with traditional bioassay-guided isolation to potentiate bioactive compound discovery. Powerful metabolomics methods such as GNPS-based molecular networking[6] detected (as could be expected) some of the metabolite features reported here (Figures S38 and S39). However, when applicable, our method circumvents challenges associated with annotation of MS/MS-derived molecular network data,[4] and, consequently, prioritization of isolation targets. Combining our strategy with molecular networking and other metabolomics tools is possible and could help streamline the analysis of labeled feature lists.[e.g.,48] Based on the relatively high incorporation efficiencies reported here, precursor directed biosynthesis strategies are likely to allow the straightforward generation of NP analogues by supplementation of cyanobacteria with pre-functionalized FAs that are either directly incorporated or that can still undergo elongation.

The apparent absence of β-oxidation in cyanobacteria deserves further investigation. The incomplete set of β-oxidation enzymes observed throughout the cyanobacterial tree of life suggests that the full pathway was absent early in the radiation of cyanobacteria. Still, our data do not entirely rule out the presence of an alternative route for FA-degradation in these organisms.[27] For example, another central metabolic pathway—the tricarboxylic acid cycle—was thought to be incomplete in cyanobacteria until recently.[49] In any case, the absent or unique FA catabolism of cyanobacteria most likely underpins the prevalence of FA-derived moieties in their secondary metabolism.

Acknowledgements

We acknowledge funding from the European Research Council (Starting Grant 759840) to PNL. Parts of the work were supported by Fundação para a Ciência e a Tecnologia through grants PTDC/BIA-BQM/29710/2017, PTDC/QUI-QIN/30649/2017, UIDB/04423/2020, UIDP/04423/2020 and UID/QUI/50006/2019, and scholarships SFRH/BD/138308/2018 to TPM and SFRH/BD/146003/2019 to KA. We thank Martin Fulda for the pET24d-aas7942 construct, Rui Rodrigues and António Vicente for ECD measurements and Lígia Sousa and Ralph Urbatzka for assistance with cytotoxicity assays.

Conflict of interest

The authors declare no conflict of interest.

Keywords: cyanobacteria · isotopic labeling · natural products discovery · oxadiazines · oxidation

[1] T. A. Scott, J. Piel, *Nat. Rev. Chem.* **2019**, *3*, 404 – 425.
[2] J. J. Hug, D. Krug, R. Müller, *Nat. Rev. Chem.* **2020**, *4*, 172 – 193.
[3] N. Ziemert, M. Alanjary, T. Weber, *Nat. Prod. Rep.* **2016**, *33*, 988 – 1005.
[4] J. J. J. van der Hooft, H. Mohimani, A. Bauermeister, P. C. Dorrestein, K. R. Duncan, M. H. Medema, *Chem. Soc. Rev.* **2020**, *49*, 3297 – 3314.
[5] A. E. F. Ramos, L. Evanno, E. Poupon, P. Champy, M. A. Beniddir, *Nat. Prod. Rep.* **2019**, *36*, 960 – 980.
[6] M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A. V. Melnik, M. J. Meehan, W.-T. Liu, M. Crüsemann, P. D. Boudreau, E. Esquenazi, M. Sandoval-Calderón, R. D. Kersten, L. A. Pace, R. A. Quinn, K. R. Duncan, C.-C. Hsu, D. J. Floros, R. G. Gavilan, K. Kleigrewe, T. Northen, R. J. Dutton, D. Parrot, E. E. Carlson, B. Aigle, C. F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B. T. Murphy, L. Gerwick, C.-C. Liaw, Y.-L. Yang, H.-U. Humpf, M. Maansson, R. A. Keyzers, A. C. Sims, A. R. Johnson, A. M. Sidebottom, B. E. Sedio, A. Klitgaard, C. B. Larson, C. A. Boya, P. D. Torres-Mendoza, D. J. Gonzalez, D. B. Silva, L. M. Marques, D. P. Demarque, E. Pociute, E. C. O'Neill, E. Briand, E. J. N. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J. J. Kharbush, Y. Zeng, J. A. Vorholt, K. L. Kurita, P. Charusanti, K. L. McPhail, K. F. Nielsen, L. Vuong, M. Elfeki, M. F. Traxler, N. Engene, N. Koyama, O. B. Vining, R. Baric, R. R. Silva, S. J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P. G. Williams, J. Dai, R. Neupane, J. Gurr, A. M. C. Rodríguez, A. Lamsa, C. Zhang, K. Dorrestein, B. M. Duggan, J. Almaliti, P.-M. Allard, P. Phapale, L.-F. Nothias, T. Alexandrov, M. Litaudon, J.-L. Wolfender, J. E. Kyle, T. O. Metz, T. Peryea, D.-T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Müller, K. M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P. R. Jensen, B. Ø. Palsson, K. Pogliano, R. G. Linington, M. Gutiérrez, N. P. Lopes, W. H. Gerwick, B. S. Moore, P. C. Dorrestein, N. Bandeira, *Nat. Biotechnol.* **2016**, *34*, 828 – 837.
[7] R. A. Quinn, L.-F. Nothias, O. Vining, M. Meehan, E. Esquenazi, P. C. Dorrestein, *Trends Pharmacol. Sci.* **2017**, *38*, 143 – 154.
[8] K. Blin, V. Pascal Andreu, E. L. C. de los Santos, F. Del Carratore, S. Y. Lee, M. H. Medema, T. Weber, *Nucleic Acids Res.* **2019**, *47*, D625 – D630.
[9] P. N. Leão, N. Engene, A. Antunes, W. H. Gerwick, V. Vasconcelos, *Nat. Prod. Rep.* **2012**, *29*, 372.
[10] K. Kleigrewe, L. Gerwick, D. H. Sherman, W. H. Gerwick, *Nat. Prod. Rep.* **2016**, *33*, 348 – 364.
[11] J. Demay, C. Bernard, A. Reinhardt, B. Marie, *Mar. Drugs* **2019**, *17*, 320.
[12] K. Tidgewell, B. R. Clark, W. H. Gerwick, in *Compr. Nat. Prod. II*, *Chem. Biol.* (Eds.: L. Mander, H.-W. Lui), Elsevier, Oxford, **2010**, pp. 141 – 188.
[13] C. Cassier-Chauvat, V. Dive, F. Chauvat, *Appl. Microbiol. Biotechnol.* **2017**, *101*, 1359 – 1364.

[14] J. J. Zhang, X. Tang, B. S. Moore, *Nat. Prod. Rep.* **2019**, *36*, 1313 – 1332.
[15] A. Kampa, A. N. Gagunashvili, T. A. M. Gulder, B. I. Morinaka, C. Daolio, M. Godejohann, V. P. W. Miao, J. Piel, Ó. S. Andrésson, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E3129 – E3137.
[16] R. B. Kinnel, E. Esquenazi, T. Leao, N. Moss, E. Mevers, A. R. Pereira, E. A. Monroe, A. Korobeynikov, T. F. Murray, D. Sherman, L. Gerwick, P. C. Dorrestein, W. H. Gerwick, *J. Nat. Prod.* **2017**, *80*, 1514 – 1521.
[17] D. S. May, C. M. Crnkovic, A. Krunic, T. A. Wilson, J. R. Fuchs, J. E. Orjala, *ACS Chem. Biol.* **2020**, *15*, 758 – 765.
[18] T. P. Martins, C. Rouger, N. R. Glasser, S. Freitas, N. B. de Fraissinette, E. P. Balskus, D. Tasdemir, P. N. Leão, *Nat. Prod. Rep.* **2019**, *36*, 1437 – 1461.
[19] O. A. Trivedi, P. Arora, V. Sridharan, R. Tickoo, D. Mohanty, R. S. Gokhale, *Nature* **2004**, *428*, 441 – 445.
[20] J. Mareš, J. Hájek, P. Urajová, A. Kust, J. Jokela, K. Saurav, T. Galica, K. Čapková, A. Mattila, E. Haapaniemi, P. Permi, I. Mysterud, O. M. Skulberg, J. Karlsen, D. P. Fewer, K. Sivonen, H. H. Tønnesen, P. Hrouzek, *Appl. Environ. Microbiol.* **2019**, *85*, https://doi.org/10.1128/AEM.02675-18.
[21] N. A. Moss, T. Leão, M. R. Rankin, T. M. McCullough, P. Qu, A. Korobeynikov, J. L. Smith, L. Gerwick, W. H. Gerwick, *ACS Chem. Biol.* **2018**, *13*, 3385 – 3395.
[22] J. P. A. Reis, S. A. C. Figueiredo, M. L. Sousa, P. N. Leão, *Nat. Commun.* **2020**, *11*, 1458.
[23] S. L. Robinson, B. R. Terlouw, M. D. Smith, S. J. Pidot, T. P. Stinear, M. H. Medema, L. P. Wackett, *J. Biol. Chem.* **2020**, *295*, 14826 – 14839.
[24] S. von Berlepsch, H.-H. Kunz, S. Brodesser, P. Fink, K. Marin, U.-I. Flügge, M. Gierth, *Plant Physiol.* **2012**, *159*, 606 – 617.
[25] P. N. Leão, H. Nakamura, M. Costa, A. R. Pereira, R. Martins, V. Vasconcelos, W. H. Gerwick, E. P. Balskus, *Angew. Chem. Int. Ed.* **2015**, *54*, 11063 – 11067; *Angew. Chem.* **2015**, *127*, 11215 – 11219.
[26] L. A. Mills, A. J. McCormick, D. J. Lea-Smith, *Biosci. Rep.* **2020**, *40*, BSR20193325.
[27] J. Beld, R. Abbriano, K. Finzel, M. Hildebrand, M. D. Burkart, *Mol. Biosyst.* **2016**, *12*, 1299 – 1312.
[28] D. Kaczmarzyk, M. Fulda, *Plant Physiol.* **2010**, *152*, 1598 – 1610.
[29] L. Jimenez-Diaz, A. Caballero, A. Segura, in *Aerob. Util. Hydrocarb. Oils Lipids* (Ed.: F. Rojo), Springer International Publishing, Cham, **2017**, pp. 1 – 23.
[30] J. Yao, C. O. Rock, *Biochimie* **2017**, *141*, 30 – 39.
[31] H. Gross, V. O. Stockwell, M. D. Henkels, B. Nowak-Thompson, J. E. Loper, W. H. Gerwick, *Chem. Biol.* **2007**, *14*, 53 – 63.
[32] G. Taylor, *Fatty Acid Metabolism in Cyanobacteria*, University of Exeter, **2012**.
[33] Y. Jiang, R. M. Morgan-Kiss, J. W. Campbell, C. H. Chan, J. E. Cronan, *Biochemistry* **2010**, *49*, 718 – 726.
[34] H. Nakamura, H. A. Hamer, G. Sirasani, E. P. Balskus, *J. Am. Chem. Soc.* **2012**, *134*, 18518 – 18521.
[35] M. L. Sousa, M. Preto, V. Vasconcelos, S. Linder, R. Urbatzka, *Front. Oncol.* **2019**, *9*, 224.
[36] P. M. D'Agostino, T. A. M. Gulder, *ACS Synth. Biol.* **2018**, *7*, 1702 – 1708.
[37] T. Pluskal, S. Castillo, A. Villar-Briones, M. Orešič, *BMC Bioinf.* **2010**, *11*, 395.
[38] J. Yang, *Natural Product Anticancer Drug Discovery and Mechanistic Studies on Hapolosin and Silvestrol*, thesis, University of Illinois at Chicago, **2013**.
[39] C. E. O'Connell, K. A. Salvato, Z. Meng, B. A. Littlefield, C. E. Schwartz, *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1541 – 1546.
[40] K. Voráčová, J. Hájek, J. Mareš, P. Urajová, M. Kuzma, J. Cheel, A. Villunger, A. Kapuscik, M. Bally, P. Novák, M. Kabeláč, G. Krumschnabel, M. Lukeš, L. Voloshko, J. Kopecký, P. Hrouzek, *PLoS One* **2017**, *12*, e0172850.
[41] I. Gutiérrez-del-Río, N. Brugerolle de Fraissinette, R. Castelo-Branco, F. Oliveira, J. Morais, S. Redondo-Blanco, C. J. Villar, M. J. Iglesias, R. Soengas, V. Cepas, Y. L. Cubillos, G. Sampietro, L. Rodolfi, F. Lombó, S. M. S. González, F. López Ortiz, V. Vasconcelos, M. A. Reis, *J. Nat. Prod.* **2020**, *83*, 1885 – 1890.
[42] K. Abt, R. Castelo-Branco, P. N. Leão, *J. Nat. Prod.* **2021**, *84*, 278 – 286.
[43] B. M. Kevany, D. A. Rasko, M. G. Thomas, *Appl. Environ. Microbiol.* **2009**, *75*, 1144 – 1155.
[44] C. A. Brotherton, E. P. Balskus, *J. Am. Chem. Soc.* **2013**, *135*, 3359 – 3362.
[45] M. O. Hunsen, in *Compr. Heterocycl. Chem. III*, (Eds.: A. R. Katritzky, C. A. Ramsden, E. F. V. Scriven, R. J. K. Taylor), Elsevier, Oxford, **2008**, pp. 291 – 299.
[46] Y. Inokuma, T. Ukegawa, M. Hoshino, M. Fujita, *Chem. Sci.* **2016**, *7*, 3910 – 3913.
[47] E. Danelius, S. Halaby, W. A. van der Donk, T. Gonen, *Nat. Prod. Rep.* **2021**, https://doi.org/10.1039/D0NP00035C.
[48] T. Hautbergue, E. L. Jamin, R. Costantino, S. Tadrist, L. Meneghetti, J.-C. Tabet, L. Debrauwer, I. P. Oswald, O. Puel, *Anal. Chem.* **2019**, *91*, 12191 – 12202.
[49] S. Zhang, D. A. Bryant, *Science* **2011**, *334*, 1551 – 1553.

Manuscript received: November 16, 2020
Accepted manuscript online: February 17, 2021
Version of record online: March 22, 2021

