ACS SyntheticBiology

pubs.acs.org/synthbio

Research Article

# Discovery and Heterologous Expression of Microginins from *Microcystis aeruginosa* LEGE 91341

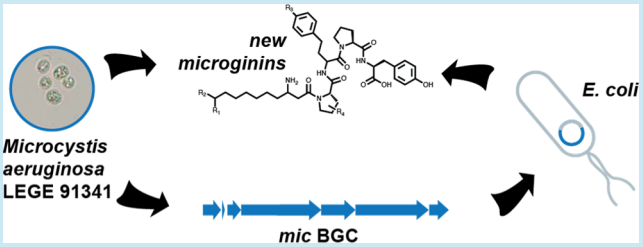
Nádia Eusébio,[§] Raquel Castelo-Branco,[§] Diana Sousa, Marco Preto, Paul D'Agostino, Tobias A. M. Gulder, and Pedro N. Leão*

Cite This: *ACS Synth. Biol.* 2022, 11, 3493–3503

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Microginins are a large family of cyanobacterial lipopeptide protease inhibitors. A hybrid polyketide synthase/non-ribosomal peptide synthetase biosynthetic gene cluster (BGC) found in several microginin-producing strains—*mic*—was proposed to encode the production of microginins, based on bioinformatic analysis. Here, we explored a cyanobacterium, *Microcystis aeruginosa* LEGE 91341, which contains a *mic* BGC, to discover 12 new microginin variants. The new compounds contain uncommon amino acids, namely, homophenylalanine (Hphe), homotyrosine (Htyr), or methylproline, as well as a 3-aminodecanoic acid (Ada) residue, which in some variants was chlorinated at its terminal methyl group. We have used direct pathway cloning (DiPaC) to heterologously express the *mic* BGC from *M. aeruginosa* LEGE 91341 in *Escherichia coli*, which led to the production of several microginins. This proved that the *mic* BGC is, in fact, responsible for the biosynthesis of microginins and paves the way to accessing new variants from (meta)genome data or through pathway engineering.




**KEYWORDS:** *cyanobacteria, natural products, direct pathway cloning (DiPaC), heterologous expression, microginins, biosynthesis*

## INTRODUCTION

Cyanobacteria are rich secondary metabolite producers. This is evidenced by the large number of natural products reported from these organisms[1] and by the presence of multiple biosynthetic gene clusters (BGCs) in their genomes.[2] In line with what is observed for most bacteria, cyanobacterial BGCs greatly outnumber the reported natural products for this phylum.[3] These observations highlight the potential for the discovery of new natural products from cyanobacteria, but they also underscore the lack of effective tools for revealing the natural products associated with orphan BGCs (i.e., those for which no compound can be ascribed). Deorphanizing BGCs from bacteria has been one of the main challenges in natural products discovery ever since next-generation sequencing became available at large.[4] For many natural products-rich organisms, genetic tools and appropriate heterologous hosts are readily available to reach this goal.[4] However, modifying cyanobacterial genomes, in particular of non-model strains, is a slow and complicated process, even if possible.[5] The heterologous expression of whole cyanobacterial BGCs in a non-cyanobacterial host has long been considered as a valid strategy but has progressed slowly, in particular due to challenges in finding suitable hosts.[5] Nevertheless, it has been possible to express over a dozen cyanobacterial BGCs, in particular in small-sized clusters, in *Escherichia coli* or cyanobacterial hosts.[5] Direct pathway cloning (DiPaC) is a recently developed strategy for BGC capture and heterologous expression that has shown promise in the expression of cyanobacterial BGCs, including complex polyketide synthase/non-ribosomal peptide synthetase pathways.[5−9]

Microginins are a large group of over 80 cyanobacterial secondary metabolites.[10] Their structures can be summarized as linear lipopeptides featuring a 3-amino fatty acyl residue connected to a peptide (typically three or four amino acid residues). Modifications of the fatty acyl residue such as terminal halogenation, hydroxylation, or *N*-methylation are common.[10,11] Several microginins were shown to inhibit a variety of proteases with considerable potency.[12] Rounge et al.[13] proposed that microginin production is carried out by enzymes encoded in the *mic* BGC (Figure 1a), found in the genome of the oscillaginin (a type of microginin)-producing strain *Planktothrix prolifica* NIVA-CYA 98. This proposal was grounded on the predicted functions of the *mic* genes, which were consistent with the generation of a metabolite with a structure similar to that of oscillaginins A and B. Specifically, the presence of NRPS adenylation domains, whose predicted specificity matched the amino acids found in the oscillaginins,

**Received:** July 21, 2022
**Published:** September 27, 2022





https://doi.org/10.1021/acssynbio.2c00389
*ACS Synth. Biol.* 2022, 11, 3493−3503

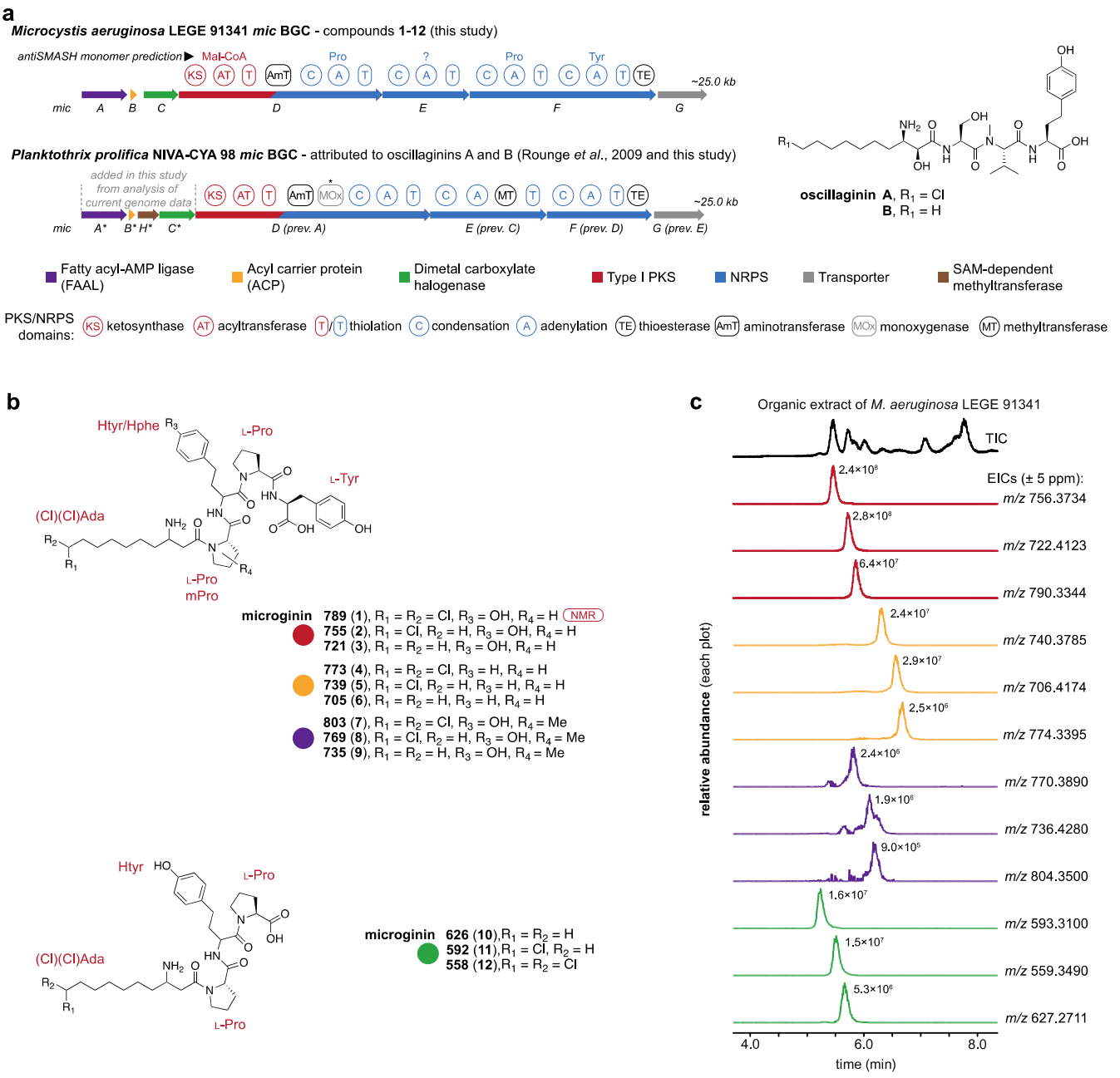



**Figure 1.** Microginin BGCs (*mic*) and discovery of microginins **1**−**12**. Architecture of the *mic* BGC from *M. aeruginosa* LEGE 91341 and of the initially proposed and now reannotated *mic* BGC from *P. prolifica* NIVA CYA 98 (asterisks indicate genes/domain annotated in this study), putatively encoding the production of oscillaginins A and B (a). Structures of the herein reported microginins **1**−**12** (b), colors of the circle for each group correspond to those in panel c. The compound isolated and structurally characterized using NMR is labeled with "NMR". LC−HRESIMS analysis of an organic extract of *M. aeruginosa* LEGE 91341 cells indicates the presence of four groups of microginins, each consisting of non-chlorinated, mono- and di-chlorinated variants, as illustrated by the represented extracted ion chromatograms for the $[M + H]^+$ ion of each of these compounds (c).

and co-linearity between such adenylation and the sequence of amino acids in these peptides formed the basis for the proposal.[13] However, the *mic* BGC reported by Rounge et al.[13] with a single PKS module and three NRPS modules could not explain the origin of the Ahda moiety in oscillaginins. Later, Kramer[14] reported a fatty acyl AMP-ligase (FAAL) enzyme associated with a *Microcystis aeruginosa mic* BGC and proposed that it would activate and load octanoic acid onto an acyl carrier protein (ACP), which would then interact with the PKS module to generate the Ahda residue. The biosynthetic origin of the halogenations that are often observed in microginins[10] was proposed only after the discovery of dimetal carboxylate halogenases—Nakamura et al.[15] identified one putative halogenase of this type within a *mic* BGC from *M. aeruginosa* PCC 9432. In line with these observations, the updated genome data for the oscillaginin-producing *P. prolifica* NIVA-CYA 98 (NCBI: AVFZ00000000) include a longer contig harboring the *mic* BGC, extending its 5′ region, which features additional genes encoding FAAL, ACP, a dimetal-carboxylate halogenase, and a SAM-dependent methyltransferase (Figure 1a). In addition, a luciferase-like monooxygenase domain is annotated in its *micD* gene (Figure 1a) by the NCBI


https://doi.org/10.1021/acssynbio.2c00389
ACS Synth. Biol. 2022, 11, 3493−3503

Table 1. NMR Spectroscopic Data ($^1$H 400 MHz and $^{13}$C 100 MHz, DMSO-$d_6$) for Microginin 789 (1)

| | position | $\delta_C$ | type | $\delta_H$[a] | mult., $J$ (Hz) | HMBC[b] | COSY |
|---|---|---|---|---|---|---|---|
| Ada | 1 | 169.0 | C | | | | |
| | 2a | 36.0 | $CH_2$ | 2.67 | m | 1, 3, 4 | 2b, 3 |
| | 2b | | | 2.40 | m | 1, 3, 4 | 2a, 3 |
| | 3 | 47.1 | CH | 3.29 | m | 2 | 2a, 2b, 4a, 4b |
| | 4a | 32.8 | $CH_2$ | 1.58 | m | | 3, 5 |
| | 4b | | | 1.50 | m | | 3, 5 |
| | 5 | 24.4 | $CH_2$ | 1.31 | m | | 4a, 4b |
| | 6 | 28.6 | $CH_2$ | 1.26−1.23 | m | 7 | |
| | 7 | 27.7 | $CH_2$ | 1.26−1.23 | m | 6 | 8 |
| | 8 | 25.3 | $CH_2$ | 1.41 | m | 7 | 7, 9 |
| | 9 | 42.8 | $CH_2$ | 2.08 | m | 7, 8, 10 | 8, 10 |
| | 10 | 74.8 | $CHCl_2$ | 6.25 | t, 5.9 | 8 | 9 |
| Pro-1 | 11 | 171.6 | C | | | | |
| | 12 | 59.0 | CH | 4.36 | m | 13 | 13a, 13b |
| | 13a | 31.3 | $CH_2$ | 2.24 | m | 11 | 12, 13b, 14b |
| | 13b | | | 1.97 | m | 11 | 12, 13a |
| | 14a | 22.4 | $CH_2$ | 1.87 | m | 13 | 14b, 15 |
| | 14b | | | 1.80 | m | 13 | 13a, 14a, 15 |
| | 15 | 46.5 | $CH_2$ | 3.43 | m | | 14a, 14b |
| Htyr | 16 | 169.6 | C | | | | |
| | 17 | 52.2 | CH | 4.48 | m | | 17-NH, 18 |
| | 17-NH | | NH | 8.28 | d, 3.2 | 11 | 17 |
| | 18 | 32.4 | $CH_2$ | 1.79 | m | | 17, 19a, 19b |
| | 19a | 30.6 | $CH_2$ | 2.62 | m | 20, 21/25 | 18, 19b |
| | 19b | | | 2.42 | m | 20, 21/25 | 18, 19a |
| | 20 | 131.2 | C | | | | |
| | 21/25 | 129.0 | CH | 7.02 | d, 8.5 | 19, 22/24, 21/25, 23 | 22/24 |
| | 22/24 | 115.0 | CH | 6.65 | d, 8.5 | 20, 22/24, 23 | 21/25 |
| | 23 | 155.4 | C | | | | |
| | 23-OH | | Ar-OH | 9.15/9.04 | br s | | |
| Pro-2 | 26 | 169.2 | C | | | | |
| | 27 | 60.4 | CH | 4.14 | d, 8.1 | 26, 28, 29, 30 | 28b |
| | 28a | 30.5 | $CH_2$ | 2.05 | m | 26 | |
| | 28b | | | 1.88 | m | 26 | 27 |
| | 29 | 21.2 | $CH_2$ | 1.59 | m | | 30 |
| | 30 | 46.3 | $CH_2$ | 3.25 | m | | 29 |
| Tyr | 31 | 174.5 | C | | | | |
| | 32 | 55.9 | CH | 3.96 | m | 31 | 32-NH, 33a, 33b |
| | 32-NH | | NH | 8.12 | d, 8.4 | 26 | 31 |
| | 33a | 36.2 | $CH_2$ | 3.04 | dd, 13.8, 4.1 | 31, 32, 35/39 | 32, 33b |
| | 33b | | | 2.68 | m | 32, 35/39 | 32, 33a |
| | 34 | 130.3 | C | | | | |
| | 35/39 | 129.7 | CH | 6.89 | d, 8.5 | 33, 35/39, 36/38, 37 | 36/38 |
| | 36/38 | 114.5 | CH | 6.58 | d, 8.5 | 34, 35/39, 37 | 35/39 |
| | 37 | 155.2 | C | | | | |
| | 37-OH | | Ar-OH | 9.15/9.04 | br s | | |

[a]From HSQC. [b]From proton to indicated carbon.

Conserved Domain Search tool. Despite these bioinformatics-based data, direct biochemical evidence connecting microginins with their putative BGC (*mic*) is lacking.

Our group recently detected a putative *mic* BGC in the cyanobacterium *M. aeruginosa* LEGE 91341 (NCBI: JADEXY000000000) encoding a dimetal carboxylate halogenase homolog[16] (Figure 1a). Here, we report the discovery of 12 new microginins (compounds **1**−**12**, Figure 1b) from this strain, including mono- and di-chlorinated as well as non-halogenated congeners. The *mic* BGC from *M. aeruginosa* LEGE 91341 was heterologously expressed in *E. coli* using DiPaC, leading to the production of several of the microginins found in the cyanobacterium but also of new variants. Our work confirms that *mic* is the microginin BGC and opens the door for accessing novel microginin diversity, namely, through the expression of additional variants of *mic* BGCs captured from (meta)genomic data or by engineering the *mic* pathway. Furthermore, this study reinforces the usefulness of the DiPaC strategy for cyanobacterial NRPS/PKS BGC deorphanization.

## RESULTS AND DISCUSSION

**Discovery of New Microginins from *M. aeruginosa* LEGE 91341.** Our previous observation of a putative microginin BGC (*mic*) in *M. aeruginosa* LEGE 91341



https://doi.org/10.1021/acssynbio.2c00389
ACS Synth. Biol. 2022, 11, 3493−3503

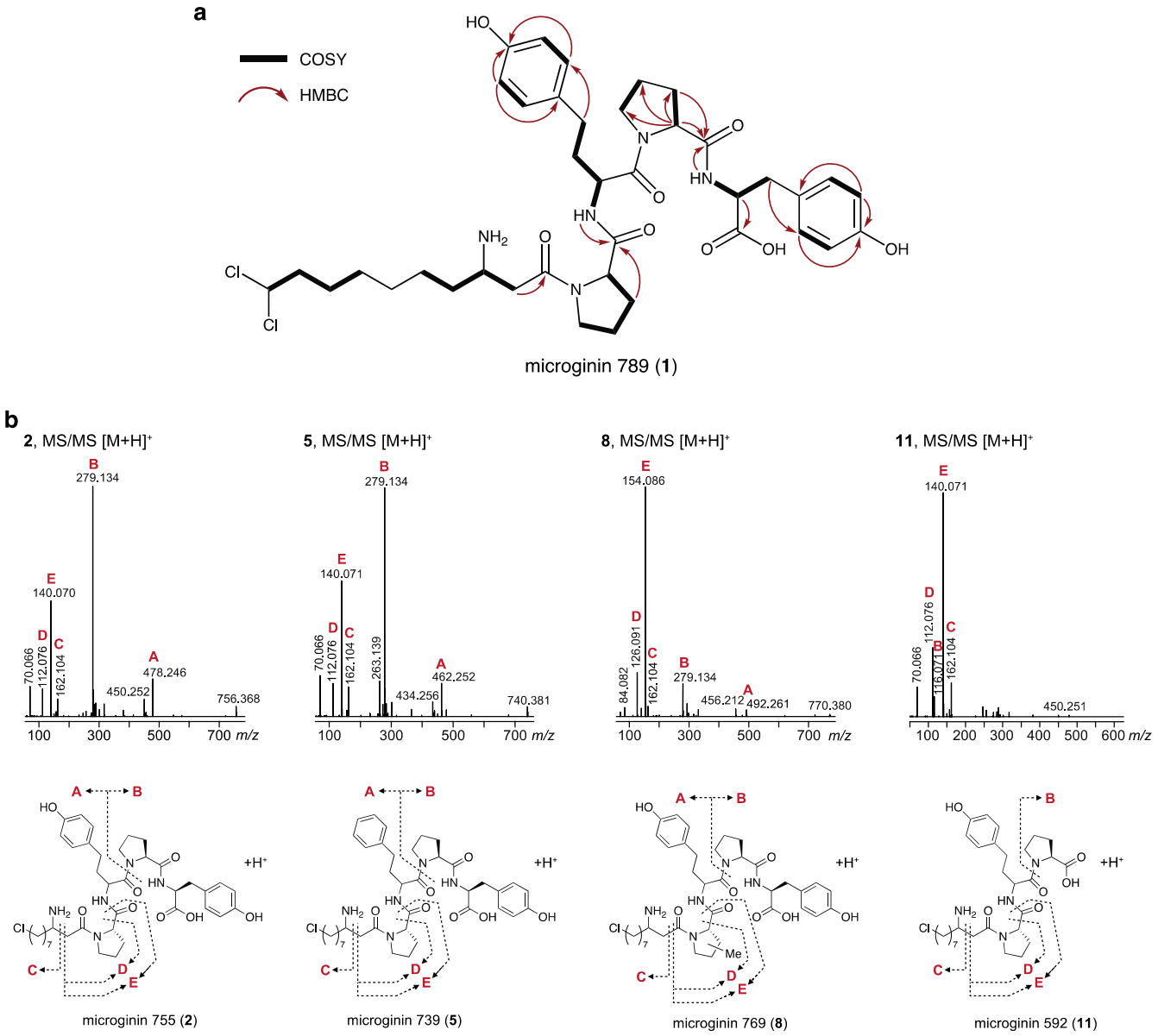



**Figure 2.** Structural elucidation of microginins. Key NMR correlations supporting the proposed structure of microginin 789 (**1**) (a); HRESIMS/MS-based deduction of the structures of the remaining microginins, as illustrated for the monochlorinated **2**, **5**, **8**, and **11** (b).

motivated a search for known or new microginins in this strain. AntiSMASH[17]-based annotation of this BGC allowed for predicting the specificity of three out of the four NRPS modules: two would incorporate Pro and one would incorporate Tyr. Manual inspection of liquid chromatography high-resolution electrospray ionization mass spectrometry (LC−HRESIMS) data for a methanolic extract of *M. aeruginosa* LEGE 91341 revealed 12 features corresponding to compounds with isotope patterns characteristic of one or two chlorinations or to their non-halogenated counterparts (Figure 1c). The most abundant features were in the *m*/*z* 600−800 range ([M + H]$^+$), which would be consistent with microginin-like metabolites. Overall, these mass features could be considered as making up four groups of analogues, each group containing variants with different degrees of halogenation (absent, one and two chlorine atoms). Searches for the accurate masses (±5 ppm) of the compounds in the NP Atlas[18] and Dictionary of Natural Products[19] returned no hits for cyanobacteria, suggesting that these were novel compounds.

**Isolation and Structure Elucidation of Microginin 789 (1).** Given the apparent abundance of some microginins in the small-scale organic extract of *M. aeruginosa* LEGE 91341 (Figure 1c), we focused next on the isolation of sufficient amounts of these metabolites to enable NMR-based structure elucidation. Fermentation of *M. aeruginosa* LEGE 91341 in 120 L of the culture medium yielded 42 g of biomass (d.w.), which was extracted by repeated percolation with methanol. The resulting crude extract (7.8 g) was subjected to reversed-phase ($C_{18}$) vacuum liquid chromatography (VLC) with a stepwise gradient of decreasing polarity from water to MeOH, to give seven fractions (A−G). LC−HRESIMS analysis indicated that fraction E, eluting with 100% MeOH, contained dihalogenated **1** in high abundance. Global Natural Products Social Molecular Networking[20] (GNPS) analysis of this fraction revealed what could be additional minor microginins (Figure S1), but their identification was not pursued in this study.



https://doi.org/10.1021/acssynbio.2c00389
ACS Synth. Biol. 2022, 11, 3493−3503

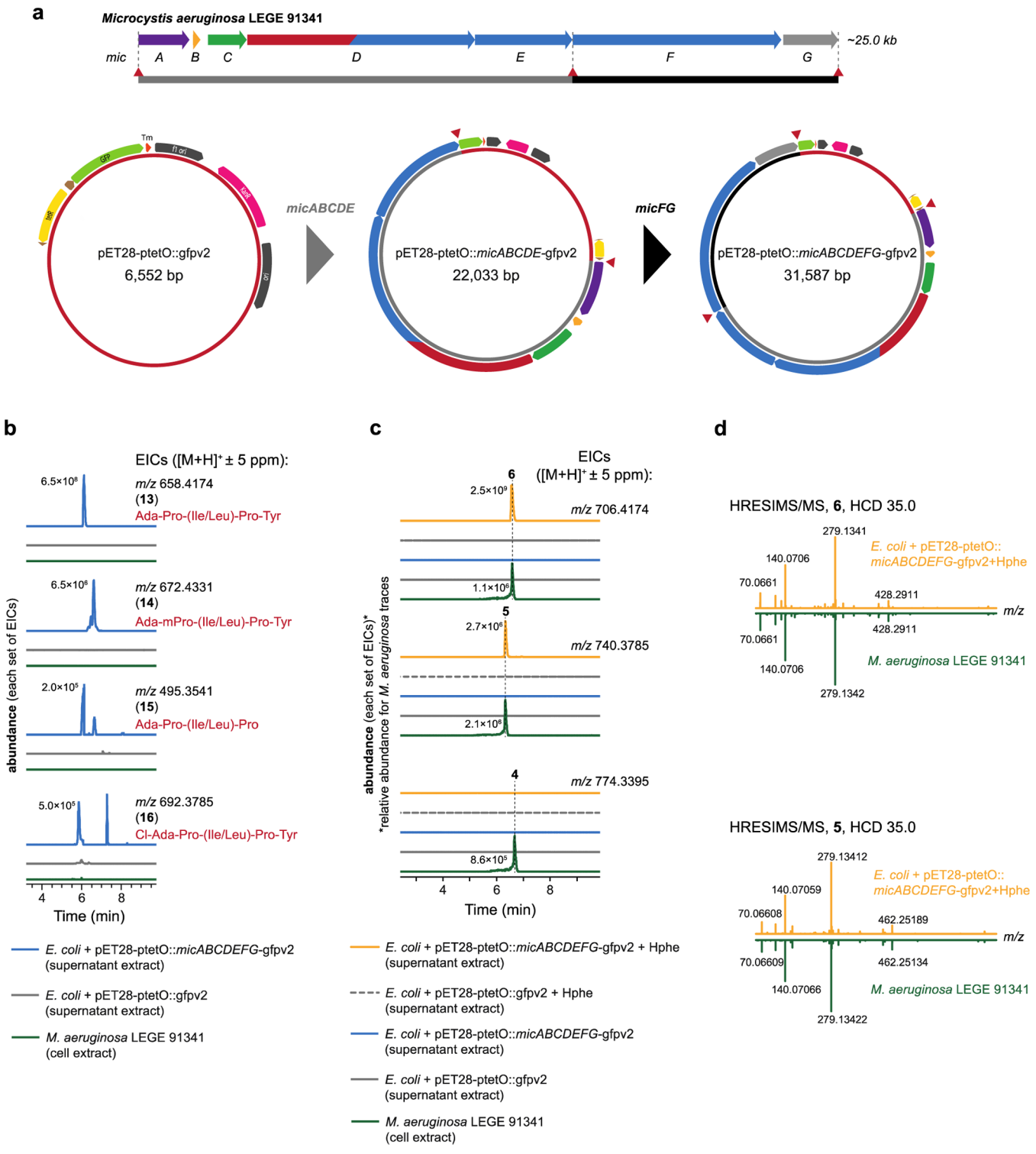



**Figure 3.** Heterologous expression of the *mic* BGC in *E. coli* leads to the production of microginins. Two-step SLIC-based DiPaC strategy used to clone and express the *mic* BGC from *M. aeruginosa* LEGE 91341 in *E. coli* (a). LC−HRESIMS analysis of *E. coli* culture extracts reveals the production of microginins in strains carrying the *mic* BGC that are not found in the producing cyanobacterium. Shown are EICs and proposed peptide sequences for novel compounds **13**−**16** (b). Supplementation with Hphe leads to the production of microginins **5** and **6** by *E. coli* harboring the *mic* BGC genes (c), which are identical to the ones produced by *M. aeruginosa* LEGE 91341, as confirmed by HRESIMS/MS analysis (d).

Subsequent reversed-phase flash chromatography and high-performance liquid chromatography (HPLC) steps afforded pure microginin 789 (**1**, 2.8 mg). This compound was subjected to 1D and 2D NMR-based structure elucidation. The $^{1}$H NMR data for **1** (Table 1) showed the presence of two broad singlets resonating at approx. 9 ppm, two exchangeable amide protons doublets ($\delta_H$ 8.28 and 8.12), several protons in the aromatic/olefin region, and resonances around $\delta_H \sim 5.0-3.5$ that could correspond to alpha protons. In addition, a methylene envelope was observable ($\delta_H \sim 1.4-1.2$), and a


https://doi.org/10.1021/acssynbio.2c00389
ACS Synth. Biol. 2022, 11, 3493−3503

terminal methyl triplet group resonance for a saturated alkyl chain was not present. The $^{13}$C NMR data (Table 1) notably indicated the presence of five carbonyls ($\delta_C \sim 175-169$) and two aromatic, likely phenol, moieties (eight resonances in the $\delta_C \sim 160-110$ range). These $^{1}$H and $^{13}$C NMR data would thus be consistent with a lipopeptide containing two aromatic residues. Additional structural information was obtained by analysis of the 2D NMR data. A combination of heteronuclear single quantum coherence (HSQC), heteronuclear multiple bond correlation (HMBC) and correlation spectroscopy (COSY) data (Table 1) was used to establish several substructures of **1**, as follows. Two very similar aromatic spin systems could be readily constructed. The first corresponded to a para-substituted phenol that further connected to an alpha proton (H-32, $\delta_H$ 3.96) through a diastereotopic methylene ($CH_2$-33, $\delta_{Ha}$ 3.04, $\delta_{Hb}$ 2.68) and was attributed to a Tyr residue. The second also corresponded to a para-substituted phenol moiety, which was HMBC-correlated to a COSY-derived spin system composed of two methylene groups ($CH_2$-19 and $CH_2$-18) and an alpha position (CH-17), thereby establishing a homotyrosine (Htyr) moiety. The two amide protons were COSY-correlated to each of the aromatic residue alpha protons (Table 1 and Figure 2a). Two proline residues could also be established (Table 1 and Figure 2a) and these explained the absence of additional amide protons. A conspicuous methine (CH-10, $\delta_C$ 74.8, $\delta_H$ 6.25) was assigned as a dichloromethyl group, taking into account the particular resonances and the presence of two unassigned chlorine atoms in the molecule. This was correlated (HMBC and COSY) to an alkyl moiety that degenerated into the methylene envelope at $CH_2$-7. A single carbonyl ($\delta_C$ 169.0) remained to be assigned and was correlated to an alkyl spin system containing a substituted methine (CH-3, $\delta_C$ 47.1, $\delta_H$ 3.29), which also degenerated into the methylene envelope at $CH_2$-9. A single methylene group, part of the methylene envelope, remained to be assigned in order to satisfy the molecular formula of **1** and was considered to be the bridging position between the two alkyl chain spin systems ($CH_2$-8). This established that **1** contains a dichlorinated 3-aminodecanoic acid (Ada) residue (Table 1 and Figure 2a). Inter-residue HMBC correlations supported that **1** corresponds to a linear peptide of sequence [$Cl_2$-Ada]-Pro-Htyr-Pro-Tyr (Table 1 and Figure 2a). The stereochemistry of all the proteinogenic amino acid residues in **1** was deduced, from the absence of epimerase domains, as being of the L configuration, while the stereochemistry of the Htyr residue and of the $\beta$-amino group were not elucidated. While angiotensin-converting enzyme (ACE) inhibitory activity has been reported for several microginins,[21] compound **1** did not show ACE inhibitory activity at a concentration of up to 1 $\mu$M (Figure S2).

**MS/MS-Based Structure Elucidation of Additional Microginins (3−12).** To obtain further structural information, we acquired MS/MS data (for the 12 new microginins, Figures S3−S14). To establish a comparative framework, we looked in detail at the MS/MS data for the monochlorinated representative of each of the four different groups (compounds **2**, **5**, **8**, and **11**, Figure 2b). Extensive analysis of microginin fragmentation under ESI conditions is available in the literature[11,22] and was used as the reference. Notably, the MS/MS spectra of **2**, **5**, **8**, and **11** contained an $m/z$ 162 ion corresponding to fragmentation of a chlorinated 3-amino decanoic acid-derived moiety. The three larger metabolites (**2**, **5**, and **8**) were composed of a C-terminal Pro−Tyr moiety, giving rise to a characteristic $m/z$ 279 fragment ion.[22] The smaller compound (microginin 592) did not show this ion and instead showed a $m/z$ 116 fragment ion corresponding to a C-terminal Pro. From this observation and because the mass difference between **2** and **11** is consistent with a Tyr residue, we propose that **10**, **11**, and **12** are truncated versions of **1**, **2**, and **3**, respectively, lacking the C-terminal Tyr.

Prominent MS/MS ions of $m/z$ 112 and 140, observed for **2**, **5**, and **11** and corresponding to Ada-Pro fragments were shifted by 14.015 a.m.u. ($CH_2$) in compound **8**. Given the absence of methyltransferase genes/domains in the *mic* BGC of *M. aeruginosa* LEGE 91341, which could lead to *N*-methylation, we propose that methyl proline (mPro) is present in **7**, **8**, and **9**, as observed in other cyanobacterial natural products.[23] Still, it is unclear how this modification takes place in the biosynthesis of **7**−**9**, and thus, we refrain from ascribing the methylation to a specific Pro carbon.

The MS/MS spectrum for compound **5** shows a prominent $m/z$ 462 peak that is not observed in **2**, while the $m/z$ 478 peak corresponding to the neutral loss of Pro−Tyr in **2** is not observed in **5**. The difference between these two fragments (15.994 a.m.u.) corresponds to an oxygen atom. Considering that, as described above, the [Cl-Ada] and Ada-Pro fragmentations in **2** and **5** are identical, we ascribe the oxygen atom difference to an Hphe residue being present in **4**, **5**, and **6**, instead of the Htyr residue found in **1**, **2**, and **3**. Analogues of cyanobacterial secondary metabolites with this type of modification have been described previously and are related to the relaxed specificity of the corresponding NRPS adenylation domain.[24]

Overall, our findings confirmed that *M. aeruginosa* LEGE 91341 produces 12 new microginins. These congeners contain the less common Ada fatty acyl moiety and a new combination of amino acids when compared to previously reported microginins. In addition, neither Hphe nor mPro had previously been encountered in this large group of cyanobacterial peptides.

**Heterologous Expression of the mic BGC.** To obtain biochemical evidence for the involvement of *mic* genes in the biosynthesis of microginins, we sought to heterologously express the *mic* BGC from *M. aeruginosa* LEGE 91341 in *E. coli*. A few NRPS/PKS pathways from cyanobacteria have been successfully expressed heterologously (using cyanobacterial or *E. coli* hosts).[5] Using a Sequence- and Ligation-Independent Cloning (SLIC) DiPaC strategy,[7] we assembled the roughly 25 kb *mic* gene cluster from *M. aeruginosa* LEGE 91341 into a modified pET-28 vector backbone, which includes a PtetO promoter and generates a C-terminal *gfp* fusion for the last gene in the BGC (pET28-ptetO:gfpv2, Figure 3a). The cloning strategy was achieved through SLIC ligation of two PCR-derived amplicons (Figure 3a). The resulting construct was cloned into the *E. coli* BAP1 strain, which has a chromosomal copy of the *sfp* gene, encoding a promiscuous Ppant transferase.[25] This enables the post-translational phosphopantetheinylation of the several thiolation domains encoded in the *mic* BGC. Comparison of the LC−HRESIMS profiles of *E. coli* pET28-ptetO:*micABCDEFG*-gfpv2 transformants with those of empty vector transformants, both grown in the minimal M9 medium supplemented with a vitamin mix (see the Methods section and Table S1), revealed MS features that were only present in cells containing the *mic* BGC (Figure 3b). Full MS spectra for these peaks showed $[M + H]^+$ ions for non-halogenated compounds of $m/z$ 658.417 (**13**) and 672.433


https://doi.org/10.1021/acssynbio.2c00389
ACS Synth. Biol. 2022, 11, 3493−3503

Table 2. Strains and Plasmids Used in the Present Study[a]

| strain | description | reference or source |
|---|---|---|
| *E. coli* TOP10 | cloning host strain | ThermoFisher Scientific |
| *E. coli* BAP1 | strain for heterologous expression | Kerafast |
| *M. aeruginosa* LEGE 91341 | microginins native producer | LEGEcc culture collection |

| plasmids | description | reference or source |
|---|---|---|
| pET28-ptetO::gfpv2 (6552 bp) | tetracycline inducible expression plasmid, ColE1, KanR, gfp reporter gene downstream of the promoter | Duell et al.[28] |
| pET28-ptetO::micABCDE-gfpv2 (22,033 bp) | pET28b-ptetO-gfpv2 with micABCDE cloned as the first fragment between PtetO and gfp | this study |
| pET28-ptetO:: micABCDEFG-gfpv2 (31,587 bp) | construct built using pET28-ptetO::micABCDE-gfp2 as a vector and micFG as a single piece nucleotide insert between the ligation site and gfp | this study |

[a]Isolation of 1 from large-scale culturing of *M. aeruginosa* LEGE 91341.

(**14**) as well as minor amounts of 495.354 (**15**) (Figure 3b). A monochlorinated version of **13** (*m*/*z* 692.378, $[M + H]^+$) could also be detected (**16**). LC−HRESIMS/MS analysis of the most abundant compounds **13**, **14**, and **16** (Figure S15) indicated that these new metabolites are microginins related to **1**−**12**, in which Leu or Ile replaces Htyr/Hphe. The formation of **13**−**16** can be explained by the fact that the *mic* BGC does not include the *hphABCD* operon necessary for the production of Htyr/Hphe.[24] In fact, this set of genes is found elsewhere in the genome of LEGE 91341, as part of an anabaenopeptin BGC (Table S2). Supplementation of the *mic* BGC transformant cultures with exogenous Hphe led to the detectable production of **5** and **6** (Figure 3c)—as confirmed by MS/MS analysis (Figure 3d)—but did not lead to the production of dichlorinated **4** (Figure 3c). Dichlorinated microginin variants were also not observed in the absence of exogenously supplied Hphe. Increasing NaCl levels boosted monochlorinated but not dichlorinated microginin production (Figure S16). Interestingly, the addition of NaBr to the culture medium led to the production of mono- and di-brominated microginin variants (Figure S16), which prompts future investigations of the MicC halogenase. Using an LC−HRESIMS calibration curve for compound **1** (Figure S17) and assuming similar ionization behavior, we estimated that pellet and supernatant levels of the major microginin in *E. coli* (**13**) amount to 1.3 mg $L^{-1}$ in yield, compared to 2.8 mg $L^{-1}$ of compound **1** in the cyanobacterial cultures, as estimated from crude extracts. We tested the influence of induction and growth period on microginin production in *E. coli*. We found that tetracycline is required for expression (Figure S16) and that higher yields are obtained with a 2 day growth period (Figure S16, roughly corresponding to a 2-fold increase in supernatant titers of **13**). Overall, the heterologous expression data confirm experimentally that the *mic* BGC is responsible for microginin production and that all the structural features (including halogenation) of the natural microginins detected in *M. aeruginosa* LEGE 91341 can be generated in *E. coli*.

## CONCLUSIONS

In this study, we report a series of new microginins from a *M. aeruginosa* strain that can be summarized as non-halogenated mono- and di-chlorinated versions of Ada-Pro/mPro-Hphe/Htyr-Pro-(Tyr) scaffolds. The discovery of these new microginin congeners was motivated by genome mining of the producing strain, which had revealed the presence of a putative *mic* BGC predicted to encode microginins with a novel amino acid composition. We show that all core biosynthetic *mic* genes are functional in *E. coli* and confirm experimentally that these are, as previously postulated, responsible for the production of microginins. The heterologous expression of these compounds in *E. coli* opens the door for engineering variants for bioactivity screening, for example, by domain substitution.[26] In addition, it enables the straightforward discovery of additional natural microginin congeners from *mic* BGCs found in genome or metagenome data. Such strategies, which take advantage of the vast amount of BGC sequence data available, can prove useful for the discovery of novel, naturally evolved variants of bioactive compounds, with improved biological activity or resistance properties, as recently illustrated by Brady and coworkers.[27]

## METHODS

**General Experimental Procedures.** Flash chromatography was performed in a Pure C-850 Flash Prep chromatography system (Buchi). HPLC separations were performed in a Waters Alliance e2695 Separations Module instrument with a 2998 PDA module and an automatic Waters Fraction Collector III or on a Waters 1525 binary pump, coupled to a Waters 2487 UV/vis detector (for the last step of the purification of compound **1**). LC−HRESIMS and LC−HRESIMS/MS analyses were performed on an UltiMate 3000 UHPLC (Thermo Fisher Scientific) system composed of a LPG-3400SD pump, a WPS-3000SL autosampler, and a VWD-3100 UV/VIS detector coupled to a Q Exactive Focus Hybrid Quadrupole Orbitrap mass spectrometer controlled by Q Exactive Focus Tune 2.9 and Xcalibur 4.1 (Thermo Fisher Scientific). For LC−HRESIMS data, the full scan mode was used with a resolution of 70,000 full width at half-maximum (fwhm), a capillary voltage of −3.8 kV, a capillary temperature of 300 °C, and a sheath gas flow rate of 35 units. The LC−HRESIMS/MS analysis was performed on the same system, and fragmentation was carried out at 17,500 fwhm, using higher energy collision dissociation (HCD), a 3.0 a.m.u. isolation window, and a normalized collision energy of 35. Optical rotations were measured using a JASCO P-2000 polarimeter (JASCO) using SpectraManager 2.14.02 software. The UV spectra were acquired on a UV-1600PC spectrometer (VWR) controlled using MWAVE 1.0.20 software. 1D and 2D NMR data were obtained in the Materials Center of the University of Porto (CEMUP) on a Bruker Avance III, 400 MHz controlled using TopSpin 3.2. Compounds were analyzed in DMSO-$d_6$ (Sigma). Chemical shifts ($^1$H and $^{13}$C) are expressed in $\delta$ (ppm), referenced to the residual non-deuterated solvent ($\delta_H$ 2.5000, $\delta_C$ 39.520). NMR data were visualized in MestreNova 12.0.4. All solvents used were MS grade for MS-based experiments, HPLC gradient grade for


https://doi.org/10.1021/acssynbio.2c00389
ACS Synth. Biol. 2022, 11, 3493−3503

HPLC analysis/purification, and ACS grade for extraction, VLC, and flash chromatography. L-homophenylalanine was purchased from TCI Europe.

**Strains and Plasmids.** The strains and plasmids used in this work and their characteristics are summarized in Table 2.

**Small-Scale Extraction of *M. aeruginosa* LEGE 91341 and LC−HRESIMS Analysis.** The strain *M. aeruginosa* LEGE 91341 was grown for 20 days in 50 mL of Z8 medium, at 25 °C, under a 14 h/10 h light (10−30 $\mu$mol photons $s^{-1}$ $m^{-2}$)/dark cycle. Biomass from the stationary phase was harvested by centrifugation (5000*g*, 10 min, 4 °C, Gyrozen 2236R) and frozen. For small-scale extraction, 30 mg of the frozen biomass was homogenized using liquid nitrogen and extracted using methanol (2 × 20 mL) at room temperature with shaking. The supernatant was obtained by centrifugation, and the solvent was evaporated in a rotary evaporator. The resulting extract was weighted, resuspended in MeOH (1 mg $mL^{-1}$), and filtered through a 0.2 $\mu$m regenerated cellulose syringe filter. Ten microliters (10 $\mu$L) of the extract were injected into the LC−HRESIMS system fitted with an ACE UltraCore 2.5 Super C18 (75 mm × 2.1 mm) column. Samples were eluted at 0.35 mL $min^{-1}$ using a linear gradient from 99.5% solution A (95% $H_2O$, 5% MeOH, and 0.1% HCOOH v/v) and 0.5% solution B (95% isopropanol, 5% MeOH, and 0.1% HCOOH, v/v) to reach 10% solution B over 0.5 min, followed by an increase to 60% solution B in 8 min and then to 90% in 1 min; these conditions were maintained for 6 min before returning to the initial conditions. The column oven was set to 40 °C, and UV monitoring was carried out at 215 and 254 nm. These chromatographic conditions were used in subsequent analyses.

In order to obtain sufficient amounts of microginins for NMR-based structure elucidation from the biomass of *M. aeruginosa* LEGE 91341, a large-scale cultivation was carried out. A total of 120 L of *M. aeruginosa* cultures in Z8 medium was carried out in 20 L of polycarbonate carboys (Nalgene). Biomass from stationary-phase batch cultures was harvested by centrifugation at 8500*g*, lyophilized, and stored at −20 °C until extraction. The freeze-dried biomass (42 g) was extracted with MeOH (6 × 500 mL). The resulting crude extract (7.8 g) was fractionated by $C_{18}$ reversed-phase VLC using a gradient of decreasing polarity—10, 30, 50, and 70% MeOH (aq), 100% MeOH, 100% *i*PrOH, and 100% $CH_2Cl_2$, yielding a total of seven fractions (A−G). After each separation during the isolation process, an LC−HRESIMS analysis was performed in order to check in which fraction the target compounds were present. The fractions D−F were shown to contain the microginins with a higher abundance. Thus, these fractions were pooled (1.7 g) and fractionated by automated flash chromatography, using a $C_{18}$ 120 g cartridge (Silicycle). Elution was carried out by a gradient of increasing MeOH proportion in $H_2O$/MeOH mixtures: 0−7 min: 10%, 8−14 min: 30%, 15−25 min: 40%, 26−36 min: 50%, 37−44 min: 60%, 45−51 min: 80%, and lastly 100% MeOH. The flow rate was 25 mL $min^{-1}$, and UV detection was performed at 210, 254, 280, and 320 nm. Pooling of subfractions was carried out according to the UV profile. Sixteen subfractions have resulted from this flash chromatography step; these were then analyzed by LC−HRESIMS. The *m*/*z* values for major microginins were detected in subfractions 2−12. These were again pooled and fractionated by reverse-phase HPLC using a semi-preparative ACE 10 C18-AR (250 × 10 mm) column. The chromatography was carried out at a 3 mL $min^{-1}$ flow at 30 °C with *i*PrOH (solution A), water (solution B), and acetonitrile (solution C). Samples were eluted by a linear gradient of B and C, from 0 to 66.6% of solution C, over 20 min. This was followed by a linear gradient of A and B solutions, from 0 to 70% of solution A over 5 min, then held at 70% solution A for 1 min, before returning to initial conditions. This chromatography yielded 16 fractions, and after LC−HRESIMS analysis, the target compounds indicated that the interaction with the stationary phase was poor and that the compounds eluted mostly in the baseline, which had also been collected (130 mg). This fraction solubilized poorly in common HPLC solvents or their mixtures, so a liquid−liquid extraction (2× $H_2O$/EtOAc, 1:1, v/v) was performed. The aqueous layer was collected and showed precipitation. After filtration, the precipitate could be dissolved in a 1:1 mixture (v/v) of iPrOH/$H_2O$. LC−HRESIMS analysis showed that several microginins were present and abundant in both the filtered aqueous layer and in the solubilized precipitate. We used the aqueous layer for semi-preparative HPLC (Luna 10u C18 100A, 250 × 10 mm, Phenomenex) using eluents A (95% $H_2O$ and 5% MeOH) and B (95% *i*PrOH and 5% MeOH). The elution program was carried out at a flow of 3 mL $min^{-1}$ and involved a gradient from 15% B to 25% B over 10 min, 25% B was held for 26 min, and a gradient to 90% B was applied over 4 min and held at 90% B for 5 min before returning to the initial conditions. UV monitoring was carried out at 254 and 280 nm. This separation afforded pure compound **1** (2.8 mg, $t_R$ = 30.5 min).

**Microginin 789 (1) (White Amorphous Solid).** $[\alpha_D^{24}]$ −58 (*c* 0.1, MeOH); UV (MeOH) $\lambda_{max}$ (log $\varepsilon$) 228.5 (3.26), 275.5 (3.18); $^1H$ and $^{13}C$ NMR data, see Table 1 and Figures S18−S22; HRESIMS *m*/*z*: 790.3333 $[M + H]^+$ (calcd for $C_{39}H_{54}N_5O_8Cl_2^+$, 790.3344).

**Molecular Networking.** To investigate if other microginins are being produced by *M. aeruginosa* LEGE 91341, each VLC fraction was analyzed by LC−HRESIMS/MS. Raw data files of each fraction but also the blanks of the run were converted to the mzML format using MSConvert, using the parameters recommended for GNPS molecular networking. The appropriate files were uploaded to the GNPS web platform using the default parameters recommended for high-resolution MS data. The molecular network was created using feature-based molecular networking analysis on GNPS and visualized in Cytoscape 3.9.0.

**Cloning of the *mic* BGC from *M. aeruginosa* LEGE 91341.** DiPaC-SLIC is based on long amplification PCR combined with homologous nucleotide overhangs, which allows for in vitro DNA assembly via SLIC of the BGC directly into the selected expression vector, in this case pET28-ptetO:gfpv2. DNA extraction from *M. aeruginosa* LEGE 91341 and genome analysis were described previously.[16] The first step of the cloning strategy was to identify putative terminators in the *mic* BGC using the ARNold tool,[29,30] but none were detected. Because after several attempts, we were unable to amplify the entire 25 kb *mic* BGC, it was split in two fragments (Figure 3a) to facilitate their insertion into pET28-ptetO:gfpv2. The primer pairs used in the cloning strategy are detailed in Table S3. The vector backbone (pET28-ptetO:gfpv2) was linearized by PCR amplification. The *micABCDE* fragment was PCR-amplified with homology tails for pET28-ptetO:gfpv2, but a restriction site for PmlI (NEB) was added to the 3′ end of the amplicon to allow linearization of the large pET28-ptetO:*micABCDE*-gfpv2 construct through endonuclease digestion. The *micFG* fragment was obtained


https://doi.org/10.1021/acssynbio.2c00389
ACS Synth. Biol. 2022, 11, 3493−3503

through PCR with homology tails to ptetO:*micABCDE*-gfpv2, which enabled assembly through SLIC. The ptetO:*micABCDEFG*-gfpv2 construct was then used for heterologous expression (Figure 3a). Amplification of all fragments was carried out as standard 50 μL PCR reactions with a final concentration of 1× Q5 High-Fidelity 2× Master Mix (NEB), 0.5 μM of forward and reverse primers, and 50−100 ng of DNA template using a Veriti Thermal Cycler. Thermal cycling programs varied according to the template and were as follows: initial denaturation at 98 °C for 30 s; denaturation at 98 °C for 10 s; annealing at 62 °C for linearization of vector backbone pET28-ptetO:gfpv2, 63 °C for amplification of the *micABCD* amplicon, or 57 °C for amplification of the *micEF* amplicon; and extension at 72 °C, 6 min for pET28-ptetO:gfpv2 or 8 min for both *micABCD* and *micFG*, followed by a final extension of 10 min at 72 °C. The PCR product of vector pET28-ptetO:gfpv2 was first treated with DpnI (NEB) (37 °C 1 h and 65 °C for 20 min) in order to remove the template plasmid that could act as the transformation background. Prior to cloning, all linear fragments were purified by gel excision using the NZYGelpure kit (NZYTech) and eluted in 30 μL. For the construction of ptetO:*micABCDE*-gfpv2, the SLIC method was performed using 1× buffer 2.1 (NEB), 0.5 μL of T4 polymerase (NEB), a variable volume of linear DNA fragments for assembly (concentrations calculated as described by Greunke et al.[6] ranging from 0.02 to 0.5 pmol), and molecular biology-grade water in a total reaction volume of 10 μL. The SLIC reaction mixture was incubated for 30 s at 50 °C and then for 10 min on ice. A total of 5 μL of the reaction mixture was transformed by heat shock into chemically competent *E. coli* TOP10 cells. The presence of positive clones was evaluated by colony PCR (Table S3), using a primer pair amplifying a portion of the ptetO:*micA* region. Clones leading to amplicons of the expected size were grown overnight at 37 °C in liquid LB medium supplemented with 50 μg $mL^{-1}$ of kanamycin, with shaking at 180 rpm. Plasmids were isolated from overnight cultures using the NZYMiniprep kit (NZYTech) and sent for sequencing to confirm the integration site. Next, the vector backbone ptetO::*micABCDE*-gfpv2 (22 Kb) was treated with PmlI (NEB) (ca. 1 μg of DNA of ptetO::*micABCDE*-gfpv2 at 37 °C for 1 h and 65 °C for 20 min). The *micFG* amplicon was purified by gel excision, as indicated above, and the construction of ptetO::*micABCDEFG*-gfpv2 (31 Kb) was performed by the SLIC reaction using the same conditions described above. To confirm the positive transformants, colony PCR was performed using two primer pairs amplifying *micEF* and *micG*-gfpv2 regions. Purified plasmids were obtained and sequenced as detailed above. The positive transformants confirmed by sequencing were then cryopreserved and used for heterologous expression. The entire BGC in the construct was sequenced by primer walking, and no mutations were detected (Figure S23).

**Heterologous Expression.** For heterologous expression experiments, *E. coli* BAP1 cells were transformed by electroporation with pET28-ptetO:gfpv2 (empty vector) or with ptetO:*micABCDEFG*-gfpv2. Expression cultures were carried out using M9 minimal medium [17.1 g $L^{-1}$ $Na_2HPO_4{\cdot}12H_2O$, 3.0 g $L^{-1}$ $KH_2PO_4$, 1.0 g $L^{-1}$ $NH_4Cl$, 0.5 g $L^{-1}$ NaCl, 1.0 mL $L^{-1}$ $MgSO_4$ (2 M), and 0.2 mL $L^{-1}$ $CaCl_2$ (0.5 M), pH 7.0] to which 10.0 mL of glucose solution (40% (w/v) and 2 mL of Vitamin mix (500×) were added after autoclaving, and which was then filtered (0.2 μm) and supplemented with 50 μg $mL^{-1}$ kanamycin. For cultures additionally supplemented with L-homophenylalanine (Hphe), this amino acid was dissolved in the culture medium at a concentration of 1.0 mM prior to inoculation. For cultures additionally supplemented with NaCl (final concentration of 2 g $L^{-1}$) or NaBr at equivalent concentrations, the salts were dissolved in the culture medium prior to inoculation.

For heterologous expression, 500 μL of a 37 °C LB overnight starter culture (supplemented with 50 μg $mL^{-1}$ kanamycin) was inoculated in 50 mL of M9 minimal medium in 100 mL Erlenmeyer flasks. Cultures were grown at 37 °C with shaking (180 rpm) until reaching an OD600 of 0.25−0.3, at which point the cultures were placed at 4 °C for ∼1 h. The cultures were then induced with 50 μL of 0.5 mg $mL^{-1}$ tetracycline. Expression was carried out for 4 days at 20 °C with shaking (180 rpm). To test the microginin production over time, cultures were harvested on the 2nd, 3rd, 4th, and 5th day of growth. In experiments testing the effect of tetracycline induction in microginin production, four conditions were tested: no induction and final concentrations of 0.2, 0.5, and 1 μg $mL^{-1}$ of tetracycline in 50 mL of culture. The cultures were harvested by centrifugation (5000*g*, 10 min at 4 °C). Cell pellets were extracted with EtOAc (20 mL) for 1−2 h with shaking and then centrifuged. The supernatants were transferred to a previously weighted vial and dried in a rotary evaporator. A second extraction was performed using 100% MeOH for 1−2 h with shaking and then centrifuged, and the supernatants were transferred to the same vial. In the case of culture supernatants, a liquid−liquid extraction was performed using 3 × 3 volumes of EtOAc. The organic phases were transferred to a round-bottom flask, dried in a rotary evaporator, and then transferred to a previously weighted vial. The extracts were weighted, dissolved in MeOH to a concentration of 5 mg $mL^{-1}$, filtered (0.2 μm), and used for LC−HRESIMS (or LC−HRESIMS/MS) analysis. Differences between the control (*E. coli* carrying the empty vector pET28-ptetO:gfpv2) and *E. coli* carrying the pET28-ptetO:*micABCDEFG*-gfpv2 plasmid were detected by manual inspection of the resulting data.

**Quantification of Microginins in *M. aeruginosa* and *E. coli*.** To quantify the production of microginins in M. aeruginosa LEGE 91341 and *E. coli*, a calibration curve for microginin 789 (**1**) was obtained from LC−HRESIMS analysis. Briefly, standards of **1** were prepared at concentrations of 0.05, 0.01, 0.005, and 0.001 mg $mL^{-1}$. Five microliters (5 μL) of each standard solution were injected under the same conditions as those used for the analysis of cyanobacterial and *E. coli* extracts, as detailed above. The area under the curve (AUC) from the extracted ion chromatograms (EIC) for **1** was plotted (0.05 mg $mL^{-1}$ to 0.001 mg $mL^{-1}$), and a linear regression was fitted. This was then used to extrapolate (rough estimation assuming a similar ionization behavior) the concentration of each microginin in cyanobacteria and *E. coli* extracts.

**ACE Inhibition Assay.** ACE inhibitory activity was evaluated using the ACE activity assay test (Sigma-Aldrich CS0002). This assay is based on the hydrolysis of angiotensin I by the ACE to yield active angiotensin II. This test uses a synthetic fluorogenic peptide as the substrate, and the measured fluorescence is directly proportional to the ACE activity present. Briefly, all reagents were diluted in the assay buffer, according to the kit instructions. A standard curve was prepared in 100 μL of assay buffer per well ranging from 0 to 8 nmol for calculating the enzymatic activity. A total of 10 μL of


https://doi.org/10.1021/acssynbio.2c00389
ACS Synth. Biol. 2022, 11, 3493−3503

**1** (final concentration of 1 $\mu$M), an inhibitor of ACE activity (lisinopril, Cayman Chemical) (a final concentration of 12 nM), a positive control of the ACE (provided by the kit), and a negative control (assay buffer) were pipetted into a 96-well black opaque plate. Then, 40 $\mu$L of the ACE, provided in the kit, were added to the sample and control wells and incubated at 37 °C for 5 min. The reaction was initiated by adding 50 $\mu$L of the fluorogenic substrate to experimental, control, and blank sample wells. The reaction was carried out at 37 °C, and fluorescence was read every minute for 5 min using a microplate reader (BioTek Cytation 5 Cell Imaging Multimode Reader, Agilent). The absorbance wavelengths of excitation and emission were set at 320 and 405 nm, respectively. All analyses were performed in duplicate. To calculate the enzymatic activity, a linear regression was determined.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.2c00389.

Molecular network, MS/MS spectra supporting the structure elucidation, culture media, details of the primer and bioinformatics searches, microginin quantification data, NMR spectra, and the cloned *mic* BGC sequence (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

**Pedro N. Leão** − *Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto, 4450-208 Matosinhos, Portugal;* orcid.org/0000-0001-5064-9164; Email: pleao@ciimar.up.pt

### Authors

**Nádia Eusébio** − *Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto, 4450-208 Matosinhos, Portugal*

**Raquel Castelo-Branco** − *Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto, 4450-208 Matosinhos, Portugal*

**Diana Sousa** − *Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto, 4450-208 Matosinhos, Portugal*

**Marco Preto** − *Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto, 4450-208 Matosinhos, Portugal*

**Paul D'Agostino** − *Chair of Technical Biochemistry, Technical University of Dresden, 01069 Dresden, Germany;* orcid.org/0000-0002-8323-5416

**Tobias A. M. Gulder** − *Chair of Technical Biochemistry, Technical University of Dresden, 01069 Dresden, Germany;* orcid.org/0000-0001-6013-3161

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.2c00389


### Author Contributions

[§]N.E. and R.C.-B. contributed equally to this work.

### Notes

The authors declare no competing financial interest.

The GNPS data set is available in MassIVE (MSV000088978). HRESIMS/MS spectra for **1**−**12** have been added to the GNPS Library under accessions CCMSLIB00009918963-CCMSLIB00009918973 and CCMSLIB00010011901. Raw NMR data for **1** are available at DOI:10.6084/m9.figshare.19303271.

## ■ ACKNOWLEDGMENTS

We acknowledge the funding from Fundação para a Ciência e a Tecnologia through grants PTDC/BIA-BQM/29710/2017, UIDB/04423/2021, and UIDP/04423/2021 to P.N.L. and scholarship SFRH/BD/136367/2018 to R.C.B. This project has also received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952374. Generous funding by the German Research Foundation and the State Ministry of Science and Cultural Affairs of Saxony (SAB 100589439) to T.A.M.G. and from the Hans-Fischer-Gesellschaft to P.D.A. are gratefully acknowledged.

## ■ REFERENCES


(1) Tidgewell, K.; Clark, B. R.; Gerwick, W. H. The Natural Products Chemistry of Cyanobacteria. In *Comprehensive Natural Products II Chemistry and Biology*; Mander, L., Lui, H.-W., Eds.; Elsevier: Oxford, 2010; Vol. 2, pp 141−188.

(2) Dittmann, E.; Gugger, M.; Sivonen, K.; Fewer, D. P. Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria. *Trends Microbiol.* **2015**, *23*, 642−652.

(3) Popin, R. V.; Alvarenga, D. O.; Castelo-Branco, R.; Fewer, D. P.; Sivonen, K. Mining of Cyanobacterial Genomes Indicates Natural Product Biosynthetic Gene Clusters Located in Conjugative Plasmids. *Front. Microbiol.* **2021**, *12*, 684565.

(4) Covington, B. C.; Xu, F.; Seyedsayamdost, M. R. A Natural Product Chemist's Guide to Unlocking Silent Biosynthetic Gene Clusters. *Annu. Rev. Biochem.* **2021**, *90*, 763−788.

(5) Dhakal, D.; Chen, M.; Luesch, H.; Ding, Y. Heterologous Production of Cyanobacterial Compounds. *J. Ind. Microbiol. Biotechnol.* **2021**, *48*, kuab003.

(6) Greunke, C.; Duell, E. R.; D'Agostino, P. M.; Glöckle, A.; Lamm, K.; Gulder, T. A. M. Direct Pathway Cloning (DiPaC) to Unlock Natural Product Biosynthetic Potential. *Metab. Eng.* **2018**, *47*, 334−345.

(7) D'Agostino, P. M.; Gulder, T. A. M. Direct Pathway Cloning Combined with Sequence- and Ligation-Independent Cloning for Fast Biosynthetic Gene Cluster Refactoring and Heterologous Expression. *ACS Synth. Biol.* **2018**, *7*, 1702−1708.

(8) Duell, E. R.; Milzarek, T. M.; El Omari, M.; Linares-Otoya, L. J.; Schäberle, T. F.; König, G. M.; Gulder, T. A. M. Identification, Cloning, Expression and Functional Interrogation of the Biosynthetic Pathway of the Polychlorinated Triphenyls Ambigol A−C from *Fischerella Ambigua* 108b. *Org. Chem. Front.* **2020**, *7*, 3193−3201.

(9) D'Agostino, P. M.; Seel, C. J.; Ji, X.; Gulder, T.; Gulder, T. A. M. Biosynthesis of Cyanobacterin, a Paradigm for Furanolide Core Structure Assembly. *Nat. Chem. Biol.* **2022**, *18*, 652−658.

(10) Jones, M. R.; Pinto, E.; Torres, M. A.; Dörr, F.; Mazur-Marzec, H.; Szubert, K.; Tartaglione, L.; Dell'Aversano, C.; Miles, C. O.; Beach, D. G.; et al. CyanoMetDB, a Comprehensive Public Database of Secondary Metabolites from Cyanobacteria. *Water Res.* **2021**, *196*, 117017.

(11) Zervou, S.-K.; Gkelis, S.; Kaloudis, T.; Hiskia, A.; Mazur-Marzec, H. New Microginins from Cyanobacteria of Greek Freshwaters. *Chemosphere* **2020**, *248*, 125961.

(12) Kraft, M.; Schleberger, C.; Weckesser, J.; Schulz, G. E. Binding Structure of the Leucine Aminopeptidase Inhibitor Microginin FR1. *FEBS Lett.* **2006**, *580*, 6943−6947.

(13) Rounge, T. B.; Rohrlack, T.; Nederbragt, A. J.; Kristensen, T.; Jakobsen, K. S. A Genome-Wide Analysis of Nonribosomal Peptide Synthetase Gene Clusters and Their Peptides in a Planktothrix Rubescens Strain. *BMC Genom.* **2009**, *10*, 396.

(14) Kramer, D. Microginin Producing Proteins and Nucleic Acids Encoding a Microginin Gene Cluster as Well as Methods for Creating Novel Microginins. U.S. Patent 0,034,680 A1, December 2, 2011.
(15) Nakamura, H.; Schultz, E. E.; Balskus, E. P. A New Strategy for Aromatic Ring Alkylation in Cylindrocyclophane Biosynthesis. *Nat. Chem. Biol.* **2017**, *13*, 916−921.
(16) Eusebio, N.; Rego, A.; Glasser, N. R.; Castelo-Branco, R.; Balskus, E. P.; Leão, P. N. Distribution and Diversity of Dimetal-Carboxylate Halogenases in Cyanobacteria. *BMC Genom.* **2021**, *22*, 633.
(17) Blin, K.; Shaw, S.; Kloosterman, A. M.; Charlop-Powers, Z.; van Wezel, G. P.; Medema, M. H.; Weber, T. AntiSMASH 6.0: Improving Cluster Detection and Comparison Capabilities. *Nucleic Acids Res.* **2021**, *49*, W29−W35.
(18) van Santen, J. A.; Jacob, G.; Singh, A. L.; Aniebok, V.; Balunas, M. J.; Bunsko, D.; Neto, F. C.; Castaño-Espriu, L.; Chang, C.; Clark, T. N.; et al. The Natural Products Atlas: An Open Access Knowledge Base for Microbial Natural Products Discovery. *ACS Cent. Sci.* **2019**, *5*, 1824−1833.
(19) Dictionary of Natural Products 32.0. https://dnp.chemnetbase.com (accessed 20 12, 2021).
(20) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; et al. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, 828−837.
(21) Paiva, F.; Ferreira, G.; Trossini, G.; Pinto, E. Identification, In Vitro Testing and Molecular Docking Studies of Microginins' Mechanism of Angiotensin-Converting Enzyme Inhibition. *Molecules* **2017**, *22*, 1884.
(22) Zervou, S.-K.; Kaloudis, T.; Hiskia, A.; Mazur-Marzec, H. Fragmentation Mass Spectra Dataset of Linear Cyanopeptides - Microginins. *Data Brief* **2020**, *31*, 105825.
(23) Liu, L.; Jokela, J.; Herfindal, L.; Wahlsten, M.; Sinkkonen, J.; Permi, P.; Fewer, D. P.; Døskeland, S. O.; Sivonen, K. 4-Methylproline Guided Natural Product Discovery: Co-Occurrence of 4-Hydroxy- and 4-Methylprolines in Nostoweipeptins and Nostopeptolides. *ACS Chem. Biol.* **2014**, *9*, 2646−2655.
(24) Lima, S. T.; Alvarenga, D. O.; Etchegaray, A.; Fewer, D. P.; Jokela, J.; Varani, A. M.; Sanz, M.; Dörr, F. A.; Pinto, E.; Sivonen, K.; Fiore, M. F. Genetic Organization of Anabaenopeptin and Spumigin Biosynthetic Gene Clusters in the Cyanobacterium *Sphaerospermopsis torques-reginae* ITEP-024. *ACS Chem. Biol.* **2017**, *12*, 769−778.
(25) Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of *E. coli*. *Science* **2001**, *291*, 1790−1792.
(26) Calcott, M. J.; Owen, J. G.; Ackerley, D. F. Efficient Rational Modification of Non-Ribosomal Peptides by Adenylation Domain Substitution. *Nat. Commun.* **2020**, *11*, 4554.
(27) Wang, Z.; Kasper, A.; Mehmood, R.; Ternei, M.; Li, S.; Freundlich, J. S.; Brady, S. F. Metagenome-Guided Analogue Synthesis Yields Improved Gram-Negative-Active Albicidin- and Cystobactamid-Type Antibiotics. *Angew. Chem., Int. Ed.* **2021**, *60*, 22172−22177.
(28) Duell, E. R.; D'Agostino, P. M.; Shapiro, N.; Woyke, T.; Fuchs, T. M.; Gulder, T. A. M. Direct Pathway Cloning of the Sodorifen Biosynthetic Gene Cluster and Recombinant Generation of Its Product in *E. coli*. *Microb. Cell Factories* **2019**, *18*, 32.
(29) Macke, T. J. RNAMotif, an RNA Secondary Structure Definition and Search Algorithm. *Nucleic Acids Res.* **2001**, *29*, 4724−4735.
(30) Gautheret, D.; Lambert, A. Direct RNA Motif Definition and Identification from Multiple Sequence Alignments Using Secondary Structure Profiles. *J. Mol. Biol.* **2001**, *313*, 1003−1011.


https://doi.org/10.1021/acssynbio.2c00389
*ACS Synth. Biol.* 2022, 11, 3493−3503