**Acs Chemical Biology**
January 2019, Volume 14 Issue 1 Pages 67-75
https://doi.org/10.1021/acschembio.8b00918
https://archimer.ifremer.fr/doc/00481/59277/



# Unique Biosynthetic Pathway in Bloom-Forming Cyanobacterial Genus *Microcystis* Jointly Assembles Cytotoxic Aeruginoguanidines and Microguanidines

Pancrace Claire [1, 2], Ishida Keishi [3], Briand Enora [4], Pichi Douglas Gatte [5], Weiz Annika R. [5], Guljarnow Arthur [5], Scalvenzi Thibault [1], Sassoon Nathalie [1], Hertweck Christian [3, 6], Dittmann Elke [5, *], Gugger Muriel [1, *]

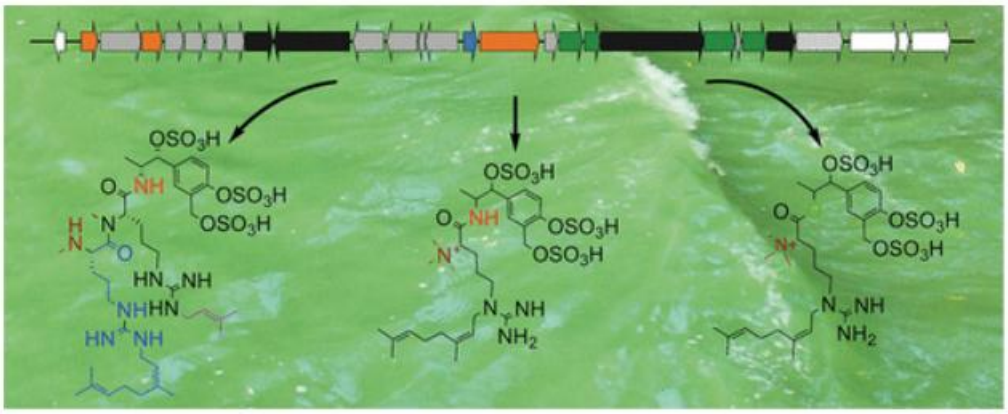
[1] Inst Pasteur, Collect Cyanobacteries, 28 Rue Dr Roux, F-75724 Paris 15, France.
[2] UPMC, UPEC, CNRS 7618, IEES Paris,UMR 113,IRD 242,INRA 1392,PARIS 7 113, 4 Pl Jussieu, F-75005 Paris, France.
[3] Hans Knoell Inst, Leibniz Inst Nat Prod Res & Infect Biol, Beutenbergstr 11a, D-07745 Jena, Germany.
[4] IFREMER, Lab Phycotoxines, Rue Ile Yeu, F-44311 Nantes, France.
[5] Univ Potsdam, Inst Biochem & Biol, Dept Microbiol, D-14476 Golm, Germany.
[6] Friedrich Schiller Univ Jena, Fac Biol Sci, D-07743 Jena, Germany.

* Corresponding authors : Elke Dittmann, email address : editt@uni-potsdam.de ; Muriel Gugger, email address : muriel.gugger@pasteur.fr

**Abstract :**

The cyanobacterial genus *Microcystis* is known to produce an elaborate array of structurally unique and biologically active natural products, including hazardous cyanotoxins. Cytotoxic aeruginoguanidines represent a yet unexplored family of peptides featuring a trisubstituted benzene unit and farnesylated arginine derivatives. In this study, we aimed at assigning these compounds to a biosynthetic gene cluster by utilizing biosynthetic attributes deduced from public genomes of *Microcystis* and the sporadic distribution of the metabolite in axenic strains of the Pasteur Culture Collection of Cyanobacteria. By integrating genome mining with untargeted metabolomics using liquid chromatography with mass spectrometry, we linked aeruginoguanidine (AGD) to a nonribosomal peptide synthetase gene cluster and coassigned a significantly smaller product to this pathway, microguanidine (MGD), previously only reported from two *Microcystis* blooms. Further, a new intermediate class of compounds named microguanidine amides was uncovered, thereby further enlarging this compound family. The comparison of structurally divergent AGDs and MGDs reveals an outstanding versatility of this biosynthetic pathway and provides insights into the assembly of the two compound subfamilies. Strikingly, aeruginoguanidines and microguanidines were found to be as widespread as the hepatotoxic microcystins, but the occurrence of both toxin families appeared to be mutually exclusive.

**Graphical abstract :**



## INTRODUCTION

*Microcystis* is a dominant bloom-forming cyanobacterium occurring in temperate freshwater ecosystems.[1] The genus is infamous for the production of the well-known hepatotoxin microcystin.[2] Both blooms and toxins cause ecosystem disturbance and public health threats, and constitute a growing concern in the frame of freshwater eutrophication and global warming. *Microcystis* has also been described as a producer of a multitude of bioactive natural products, some of interest for biotechnological and pharmaceutical application.[3-5]

Cytotoxic aeruginoguanidines (AGDs) represent one of the most remarkable families of compounds described for *Microcystis.*[6] The three AGD congeners reported for strain *M. aeruginosa* NIES-98 feature highly unprecedented characteristics such as a 1-(4-hydroxy-3-hydroxymethyl)-phenyl-1-hydroxy-2-propylamine-4',3',1-tri-*O*-sulfate (Hphpa trisulfate) moiety, along with geranylation and prenylation of arginines (Fig. 1A). While bloom-forming *Microcystis* belong to the most intensively studied cyanobacteria, AGDs were reported only twice from a bloom in Czech Republic and an isolate in Brazil,[7, 8] and never from any other cyanobacteria. Their intricate features confine AGDs into a unique compound family.[3]

Our recent genomic analysis of ten *Microcystis* strains revealed that the different genotypes share a highly similar core genome while their biosynthetic gene clusters (BGCs) involved in natural product (NP) formation show a sporadic distribution. Moreover, we uncovered three cryptic BGCs not associated with any cyanobacterial compound.[9] The continuously increasing number of publically available genomes of *Microcystis* further corroborates the high genetic diversity and patchy distribution of the NPs produced by this cyanobacterium.

Analysis of mass spectrometry (MS) data has been widely used for years in NP characterization efforts. Molecular networking computational approach uses tandem MS/MS data to group spectra based on their fragmentation patterns similarities, which gain strength in the frame of multi-strain comparison. Approaches combining molecular networking with genome mining highlight putative links between parent ions and pathways responsible for their biosynthesis. This combinatorial approach has been shown effective at linking NPs to their biosynthetic gene clusters in cyanobacteria and other prokaryotes such as *Salinospora.*[10, 11].

Here, we have utilized the sporadic distribution of BGCs in *Microcystis* to assign one of

the orphan BGCs to AGD. By integrating the genome sequence of the known AGD-producing strain *Microcystis aeruginosa* NIES-98,[12] we screened *Microcystis* public genomes and axenic PCC strains for the AGD and its candidate BGC using genome mining, PCR and untargeted metabolomics. These data were further combined with molecular networking and genome comparison to link AGD to its biosynthetic gene cluster and study its diversity at the genetic and the metabolite level. The integrative approach allowed to enlarge the AGD compound family with microguanidine amide congeners (MGAs) and new variants of microguanidines (MGDs), and provides comprehensive insights into the extraordinary versatility of this biosynthetic pathway.

## RESULTS AND DISCUSSION

**Candidate synthesis BGC for sulfated, geranylated and prenylated compounds.** Considering the chemical structure of aeruginoguanidine (Figure 1A), the BGC involved in its synthesis was expected to encode nonribosomal peptide synthetase (NRPS) modules with specificity for L-arginine and tailoring enzymes such as a prenyltransferase and a sulfatase/sulfotransferase. The genome of the AGD-producing strain *Microcystis aeruginosa* NIES-98 contained only one cluster with these features, which was homologous to the MIC2 cluster previously described in the genomes of *Microcystis aeruginosa* PCC 9806 and PCC 9717 and *Microcystis* sp. T1-4.[9] The candidate BGC encoded two mono-modular NRPS, one of which comprising an integrated *N*-methylation domain as anticipated for the *N*-methylation of the Arg moieties. Substrate prediction of the second NRPS was more ambiguous without excluding Arg (Table 1). The putative AGD BGC, which spans ~34kb in the genome of *Microcystis aeruginosa* NIES-98, includes 25 genes (Table 1) organized in three operons (Figure 1B). The two NRPS AgdE and AgdK are accompanied by a predicted hydroxybenzoate synthase (AgdH), an AMP-dependent-ligase (AgdA), a peptidyl carrier protein (AgdB), a radical SAM protein with decarboxylase function (AgdC) and two thioester reductases (AgdN and AgdU). Several proteins consistent with tailoring enzymes involved in AGD biosynthetic pathways are present such as two methyltransferases (AgdI, AgdM), an aminotransferase (AgdL), an isoprenyltransferase (AgdJ), several sulfatase/sulfotransferases (AgdD, AgdG, AgdP and AgdR), plus putative permease/transporters (AgdF, AgdO), and thiamine pyrophosphatase (AgdQ) genes.

This candidate BGC for AGD present in seven genomes, including the public ones of

*Microcystis aeruginosa* TAIHU98, *Microcystis* sp. SPC777 and CACIAM03, was used to optimize specific primers and PCR conditions to detect its presence in *Microcystis* strains. The two primer pairs designed were targeting two genes of the candidate BGC presumably involved in an early and a late stage of AGD biosynthesis. Both genes do not share homologies with other NRPS BGCs in *Microcystis* (*agdH* and *agdJ*, Table S1). The screening of these two selected genes revealed seven additional PCC *Microcystis* strains, whose on-going genome sequences helped to better define the limits of this BGC (Table S2). A close inspection of the 14 genomes revealed the candidate AGD BGC with 28 genes in perfect synteny, without rearrangement, and expanded the initial MIC2 cluster with conserved neighboring genes (Figure 1B). Noteworthy, the largest NRPS gene *agdK* of *Microcystis* sp. PCC 10613 was reduced to a remnant fragment, as confirmed by PCR. In addition, the gene *agdK* was split in two in the genomes of *Microcystis* sp. CACIAM03 and TAIHU98. Similarly, the gene *agdQ* was split in the genome of PCC 9624, while a contig border separated *agdP* and *agdQ* in the genomes of PCC 9624 and PCC 10613. The predicted aminotransferase gene *agdL* was lacking in the genomes of PCC 9717 and PCC 9810, also confirmed by PCR. Finally, the genes *agdS* and *agdT*, without known function, appeared duplicated in ten strains (Figure 1B).

**AGD and co-assignment of microguanidine by Molecular Networking.** Detection of AGD was performed by LC-MS/MS to assess its presence in the AGD producer strain NIES-98 and in ten strains of the PCC containing the candidate BGC, as well as in eight PCC strains that did not contain it in their genomes. Two molecular networks (MNs) were constructed from LC-MS/MS data, one in positive mode (MN(+)) and another in negative mode (MN(–)). In order to dereplicate the complex dataset, signatures of NPs previously found in some of these *Microcystis* strains were identified using high-performance liquid chromatography electrospray ionization mass spectrometry (HPLC-ESI-MS/MS). Specifically, MS/MS fragments were identified for the cyanopeptolins A, B and C in PCC 7806, aeruginosamides B and C and ferintoic acid (anabaenopeptins) in PCC 9432, and ferintoic acid in PCC 9701 as predicted from their genomes (Figure S1A).[9, 13] The MN(+), consisting of 1998 nodes, was thus reliable in finding the expected compounds. However, AGD was spread in several nodes of the MN(+) apart from each other. Indeed, AGD had a better fragmentation pattern in negative mode as it was collapsed into a single large node among the 1876 nodes of the MN(–) (Figure S1B). An

extraction of the AGD node in MN(–) encompassed all strains carrying the full candidate BGC for AGD synthesis, but neither the strain PCC 10613 nor the strains lacking this AGD candidate cluster (Figure 2A). Up to 20 different putative variants of AGD were found in these *Microcystis* strains, with strains NIES-98, PCC 9804, PCC 9805 and T1-4 able to produce the three known AGD standards, whereas the other strains produced one or two of those variants (Figure 2).

Strikingly, the MN(–) revealed that all the strains containing the AGD candidate BGC produced also a significantly smaller product of 772 Da (Figure 2B). Literature research revealed that a compound with this mass, microguanidine AL772, was previously reported for a *Microcystis* bloom.[14] Microguanidines (MGDs) share striking similarities with AGDs but display also considerable differences. Instead of the highly unusual Hphpa trisulfate moiety, MGDs contain 3-(4-hydroxy-3-hydroxymethylphenyl)-2-hydroxy-1-propanol (Hphpol). Further, MGDs feature a permethylation at the α-amino group of Arg that has not been observed in AGDs. Along with MGD AL772 (**4**, Figure 3**,** related Figures S3 and S4A), a new MGD variant, MGD-704 (**5**, Figure 3**,** related Figures S3 and S4B, and Table S3) was detected in the majority of strains differing from the two other characterized MGDs, KT636 and DA368.[14-16]

In addition, the structural elucidation of the MGD size range compounds by MS fragmentation and high-resolution MS analyses uncovered a novel intermediate class of metabolites mixing features of AGD and MGD. While both compounds contain the Hphpa trisulfate moiety linked with an amide bond to the arginine derivative as in AGDs they were lacking the second arginine moiety and carried the same permethylation at the α-amino group of Arg as in MGDs (**6** and **7**, Figure 3**,** related Figures S3 and S4C and D, and Table S3). To confirm the structure of **6**, several **4** and **6** producing strains were extracted and small amounts of **4** and **6** were purified by reversed-phase HPLC. The $^{1}$H NMR spectra of AGD 98-A (**1**), AGD 98-B (**2**), AGD 98-C (**3**), MGD AL772 (**4**), and MGA (**6**) showed highly similar signals (Figure S5-S14). Detailed comparison of $^{1}$H NMR signals between **4** and **6** revealed three notable differences, namely the appearance of new amide proton δ 8.48 (H11 in **6**), 1.02 and 0.18 ppm and high field shifted methine protons H8 δ 5.21 (**4**) to δ 4.23 (**6**) and H13 δ 4.11 (**4**) to δ 3.93 (**6**), respectively (Figure S5). The $^{1}$H-$^{1}$H COSY correlation from H8 to H11 and HSQC analysis of **6** indicated that C8 (δ 49.8 in **6**, δ 75.3 in **4**) is adjacent to nitrogen (Figure S15-S18, Table S4). These results strongly supported that the predicted structure of **6** indeed possesses an amide

bond instead of the ester bond in **4**. As the low amount of **6** did not enable a sufficient quality of $^{13}C$ NMR and other 2D NMR spectra, chemical shift assignment of **6** was performed by the comparison with NMR data of **4**. The stereochemistry of the geranyl group of **6** was determined as *Z*-form, judging from the close similarity of chemical shifts with **1**-**3** and $^{13}C$ NMR data of geraniol (*E*-form) and nerol (*Z*-form) (www.chemicalbook.com/). This result further revealed that the stereochemistry of geranyl group of MGD AL772 (4) also has *Z*-form. The new intermediate class of compounds was designated microguanidine amide, with MGA-771 and MGA-787.

Indeed, the MGA peptides and the two MGD depsipeptides were observed simultaneously with AGDs in four strains (PCC 9804, PCC 9805, PCC 9811 and T1-4). Thus, *Microcystis* harboring the Agd BGC may build two different condensations between the modified Arg residue and the phenetylalcohol (ester bond) in MGD congeners or the phenetylamine (amide bond) in all AGD congeners (Figure 3).

The co-existence of AGD and MGD in the majority of Agd BGC positive strains, the existence of a new intermediate class and the large overlap in anticipated biosynthetic features lead us to conclude that AGD and MGD represent alternative products of the same biosynthetic pathway. Remarkably, strain PCC 10613 lacking the NRPS gene agdK was found to produce the MGDs in the MN(–) (Figure 2). Noteworthy, strain PCC 9624 in which the Agd BGC differed at the level of the *agdQ* produced only the AGD-98A and the MGD-AL772. Similarly, PCC 9810, PCC 9811 and PCC 9717 that lack the predicted aminotransferase *agdL* and several Agd genes of unknown function (*agdS'*, *agdT'*) produced a lower diversity of AGD variants under the same growing conditions than other AGD producing *Microcystis* strains. None of the other *Microcystis* strains analyzed, notably the ones containing the Mcy gene cluster, produced AGD, MGA or MGD.

**Characterization of the BGC potentially involved in the AGD/MGD synthesis.** One of the most striking findings of our study is the extraordinary diversity of products concurrently generated by the AGD/MGD pathway in single strains. Considering the variations detected even in the backbone of AGDs and MGDs and in the linkage of their individual moieties, the biosynthesis pathway cannot be considered as a classic assembly line of NRPS. This pathway is rather a toolkit of enzymes optionally producing a cocktail of metabolites that share the same precursors and similar tailoring modifications but combine the different building blocks to alternative products. At the

same time, the unprecedented diversity of products and intermediates and the existence of natural mutants lacking individual biosynthetic genes allows for conclusions regarding a number of biosynthetic steps of the complex pathway.

The presence of a putative *p*-hydroxybenzoate synthase (AgdH) in the AGD cluster indicates that the trisubstituted benzene unit of Hphpa and Hphpol might be derived from chorismate [17]. Given that Hphpa and Hphpol possess a rare *m*-hydromethyl residue in the benzene ring, AgdH might act in a similar way as isochorismate mutase, which has been reported to catalyze the transformation of isochorismate to *m*-carboxyphenylpyruvate.[18, 19] We cannot dissect all individual steps towards the Hphpa and Hphpol moieties, but we propose that the AMP-dependent ligase AgdA might activate the *o*-carboxylic acid group of a *p*-hydroxyphenylpyruvate intermediate followed by the transfer to the free-standing PCP AgdB (Figure 4). The resulting thioester is presumably reduced to the corresponding alcohol either by thioester reductase AgdN or U through reductive chain termination as shown for myxochelin biosynthesis in *Stigmatella aurantiaca*.[20] A yet unassigned hydroxylation step at the β-position of the *m*-hydroxymethyl-*p*-hydroxyphenylpyruvate yields 3-hydroxy-*m*-hydroxymethyl-*p*-hydroxyphenylpyruvate as the precursor of both Hphpa and Hphpol. We hypothesize that this precursor represents a branching point where further transformation of the α-keto group by aminotransferase AgdL yields Hphpa, while transformation by a reductase (e.g. AgdN or U) yields Hphpol (Figure 4). This hypothesis is supported by the fact that the lack of agdL in strains PCC 9717 and PCC 9810 still permits production of MGD variants containing the Hphpol moiety (**4** and **5**) but not the alternative Hphpa moiety as in MGAs (**6** and **7**). It is of note, that some of the predicted biosynthetic steps for Hphpa and Hphpol biosynthesis (Figure 4) share similarities to enzyme reactions involved in biosynthesis of the characteristic Choi moiety in the aeruginosin pathway [21]. In this context, it is worth mentioning that the majority of AGD/MGD producers also harbor aeruginosin biosynthesis genes in their genome (Figure 5), thus not excluding the possibility of a joint use of precursors and enzymes.

Furthermore, the strain *M. aeruginosa* PCC 10613 can be considered as a natural *agdK* mutant, thus allowing deducing the roles of the two NRPSs in the pathway. The fact that the lack of AgdK in PCC 10613 still enables MGD production strongly suggests that AgdE is the responsible NRPS activating Arg in the MGD and MGA pathways (Figure 6). On the other hand, the NRPS AgdK harbouring an N-methyltransferase domain is likely

incorporating *N*-Me-Arg in the AGD pathway. Whether or not AgdK acts iteratively or cooperates with AgdE to yield the MeArg-MeArg-Hphpa moiety of AGDs cannot be dissected based on the current dataset. The biosynthetic intermediate(s) might be methylated and decarboxylated by the radical SAM enzyme AgdC. Since AgdC shows close homology to the oxygen-independent coproporphyrinogen III oxidase of *E.coli* (HemN) we propose that it utilizes a 5`-deoxyadenosyl radical to trigger a decarboxylation reaction as demonstrated for the HemN enzyme family. [22] The intermediate may further be modified by several tailoring enzymatic reactions such as *N*-methylation (methyltransferase; AgdI or M) of Arg residue, to the tri-sulfation (sulfotransferases; AgdD, P and R, sulfatase; AgdG) of the Hphpa residue, and the *N*-alkylation (isoprenyltransferase; AgdJ) of *N*-MeArg residues. Some of the proposed biosynthetic steps may occur while substrates are tethered on PCP-domains of NRPSs or the standalone peptidyl carrier protein AgdB. The fact that no desulfated intermediates were observed in the MS/MS networking may suggest that sulfation of the aromatic moiety occurs in the PCP-bound state.

The distinct alkylation pattern at the guanidinyl group of *N*-trimethyl Arg (ω for AGDs and ε for MGDs) may derive from alternative substrate specificities of the isoprenyltransferase AgdJ (Figure 6). Comparison of the distinct AGD/MGD product profiles of individual *Microcystis* strains thus suggests an outstanding versatility of the pathway. A complete assignment of biosynthetic steps will require biochemical characterization of participating enzymes and targeted feeding studies, yet the analysis of natural agdK and agdL mutants led to definite conclusions regarding the role of these two enzymes.

The example of the joint AGD/MGD pathway further strengthens the paradigm that cyanobacteria have evolved unique mechanisms to produce diverse NPs of high complexity in single strains using limited genetic resources. Other cyanobacterial mechanisms include the utilization of alternative starter modules for NRPS as shown for the anabaenopeptin synthetase of strain *Anabaena* 90,[23] the integration of multispecific adenylation domains of NRPS as shown for the anabaenopeptin synthetase of *Planktothrix* NIVA-CYA 126,[24] and the microcystin synthetase in *Microcystis aeruginosa* NIES 843.[25] Recently, a simultaneous production of anabaenopeptins and namalides allowed to reveal a single pathway for their synthesis.[26] We can only speculate whether

AGDs and MGDs act synergistically or fulfill parallel independent functions in the producing strains.

An interesting phenomenon observed during this study is that AGD/MGD production and MC production are almost mutually exclusive among *Microcystis* strains. The only exception was found in the genomes of two non-monoclonal Brazilian strains,[27, 28] that carry both clusters and for which the production of these compounds is not yet documented. There is increasing evidence that MCs are closely interfering with the primary metabolism of *Microcystis* in addition to their toxicity.[29] Whether or not AGD and MGD can complement for the loss of MC or reflect a different niche adaptation of their respective producers remains elusive.

Our study further suggests that the rare detection of AGD and MGD in only two *Microcystis aeruginosa* isolated in Japan and in Brazil (NIES 98[6] and NPCD-1[8]) and bloom materials of *Microcystis* in Israel[14-16] respectively is not due to the scarce occurrence of these metabolites among *Microcystis*, but rather to the lack of attention towards these peculiar NPs in previous studies. Thus, the AGD/MGD producers seem to be as dispersed worldwide as the MC producing strains, and therefore should be considered in future screening of *Microcystis* blooms and isolates.

## CONCLUSIONS

Cyanobacteria are infamous for worldwide bloom formation in freshwater bodies. Risk assessment of *Microcystis* blooms primarily considers the hepatotoxin microcystin (MC). The present study suggests that the neglected family of compounds, cytotoxic aeruginoguanidines and microguanidines, is more frequently produced than previously anticipated, mainly in non-MC producing *Microcystis* strains. Remarkably, the two structurally divergent groups of compounds are products of a branched and versatile biosynthetic pathway. The genetically constraint gene cluster generates a library of diverse products in single strains and further strengthens the paradigm that cyanobacteria have developed unique mechanisms to generate metabolic diversity. These findings open new perspectives for future studies on orphan natural products and evolution of their biosynthetic pathways.

## MATERIALS AND METHODS

**Strain cultures and detection of the cluster.** Axenic *Microcystis* strains from the PCC and from the NIES collections were grown at 25 °C in 40 mL $BG11_0$ medium[30] supplemented with 2 mM $NaNO_3$ and 10 mM $NaHCO_3$ under continuous light (Table S2). For nucleic acid extraction, chemical and PCR analysis, the details are described in Supporting information.

**Sequencing & genomics analysis.** For the strains suspected to carry the *agd* gene cluster, whole genome sequencing was performed by the Mutualized Platform for Microbiology at Institut Pasteur. Genomes were integrated in the MicroScope platform[31] for further analysis. The genome sequencing is described in Supplemental information. The species tree was generated by a concatenation of 586 conserved proteins selected from the phylogenetic markers previously validated for Cyanobacteria.[32] Phylogenetic analysis is detailed in Supplemental information. AntiSMASH 3.0[33] was used to identify the targeted BGC in each genome sequence. In cases where the *agd* gene cluster spanned several contigs/scaffolds PCRs were performed to confirm the colocalization of the gene cluster parts in the same genomic locus (Table S1).

**Cyanobacterial cell extraction**. Lyophilized cyanobacterial cells from 200 mL cultures of 19 *Microcystis aeruginosa* strains were extracted with 80% aqueous methanol (v/v, 25 mL) using a sonicator (Sonoplus MS73, Bandelin, 30% power, 5 cycles for 2 min at room temperature). Each extract was centrifuged at 8,000 × g for 15 min at 15 °C. The residues were extracted with 80% aqueous methanol (v/v, 25 mL) and methanol (25 mL), respectively, as the above-mentioned procedure. The extracts were combined and dried under a reduced pressure. The crude residues were dissolved in 50% aqueous methanol (v/v, 1 mL) and kept in a fridge until analysis.

**HPLC-MS measurement**. LC-MS/MS measurements were carried out by Bruker HCT Ultra ion trap mass spectrometry (BrukerDaltonics, Bremen, Germany) coupled with an Agilent Technologies 1100 series liquid chromatogram system (Agilent, Waldbronn, Germany). The HR-LCMS measurements were performed by HPLC-HRMS series of Thermo Accela (LC) and Thermo Exactive (HRMS), an ESI source operating in both polarity mode and an orbitrap analyzer (Thermo Fisher Scientific, Bremen). The details of both measurements are described in Supporting Information.

**Molecular networking.** LC-MS/MS data acquired from Bruker instrument were used for molecular networking. Two molecular networks (MNs) were performed with LC-

MS/MS data, one in positive mode (MN(+)) and another with negative mode data (MN(–)) with LC-MS/MS data from *Microcystis* strains and AGD A, B and C standards. The steps followed for both MNs are described in Supporting Information.

**REFERENCES**

1. Harke, M. J., Steffen, M. M., Gobler, C. J., Otten, T. G., Wilhelm, S. W., Wood, S. A., and Paerl, H. W. (2016) A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp, *Harmful Algae 54*, 4—20.

2. Merel, S., Walker, D., Chicana, R., Snyder, S., Baures, E., and Thomas, O. (2013) State of knowledge and concerns on cyanobacterial blooms and cyanotoxins, *Environ. Int. 59*, 303—327.

3. Welker, M., and von Dohren, H. (2006) Cyanobacterial peptides - nature's own combinatorial biosynthesis, *FEMS Microbiol. Rev. 30*, 530—563.

4. Kehr, J. C., Gatte Picchi, D., and Dittmann, E. (2011) Natural product biosyntheses in cyanobacteria: A treasure trove of unique enzymes, *Beilstein J. Org. Chem. 7*, 1622—1635.

5. Dittmann, E., Gugger, M., Sivonen, K., and Fewer, D. P. (2015) Natural product biosynthetic diversity and comparative genomics of the Cyanobacteria, *Trends Microbiol. 23*, 642—652.

6. Ishida, K., Matsuda, H., Okita, Y., and Murakami, M. (2002) Aeruginoguanidines 98-A-98-C: cytotoxic unusual peptides from the cyanobacterium *Microcystis aeruginosa*, *Tetrahedron 58*, 7645—7652.

7. Welker, M., Marsalek, B., Sejnohova, L., and von Dohren, H. (2006) Detection and identification of oligopeptides in *Microcystis* (cyanobacteria) colonies: Toward an understanding of metabolic diversity, *Peptides 27*, 2090—2103.

8. Silva-Stenico, M., da Silva, C., Lorenzi, A., Shishido, T., Etchegaray, A., Lira, S., Moraes, L., and Fiore, M. (2011) Non-ribosomal peptides produced by Brazilian cyanobacterial isolates with antimicrobial activity, *Microbiol. Res. 166*, 161—175.

9. Humbert, J. F., Barbe, V., Latifi, A., Gugger, M., Calteau, A., Coursin, T., Lajus, A., Castelli, V., Oztas, S., Samson, G., Longin, C., Medigue, C., and de Marsac, N. T. (2013) A tribute to disorder in the genome of the bloom-forming freshwater cyanobacterium *Microcystis aeruginosa*, *PLoS One 8*, e70747.

10. Duncan, K. R., Crusemann, M., Lechner, A., Sarkar, A., Li, J., Ziemert, N., Wang, M., Bandeira, N., Moore, B. S., Dorrestein, P. C., and Jensen, P. R. (2015) Molecular networking and pattern-based genome mining improves discovery of biosynthetic gene clusters and their products from *Salinispora* species, *Chem. Biol. 22*, 460—471.
11. Moss, N. A., Bertin, M. J., Kleigrewe, K., Leao, T. F., Gerwick, L., and Gerwick, W. H. (2016) Integrating mass spectrometry and genomics for cyanobacterial metabolite discovery, *J. Ind. Microbiol. Biotechnol. 43*, 313—324.
12. Yamaguchi, H., Suzuki, S., Sano, T., Tanabe, Y., Nakajima, N., and Kawachi, M. (2016) Draft genome sequence of *Microcystis aeruginosa* NIES-98, a non-microcystin-producing cyanobacterium from Lake Kasumigaura, Japan, *Genome Announc. 4*, e01187—01116.
13. Briand, E., Bormans, M., Gugger, M., Dorrestein, P. C., and Gerwick, W. H. (2016) Changes in secondary metabolic profiles of *Microcystis aeruginosa* strains in response to intraspecific interactions, *Environ. Microbiol. 18*, 384—400.
14. Gesner-Apter, S., and Carmeli, S. (2008) Three novel metabolites from a bloom of cyanobacterium *Microcystis* sp., *Tetrahedron 64*, 6628—6634.
15. Adiv, S., and Carmeli, S. (2013) Protease inhibitors from *Microcystis aeruginosa* bloom material collected from the Dalton Reservoir, Israel, *J. Nat. Prod. 76*, 2307—2315.
16. Lifshits, M., and Carmeli, S. (2012) Metabolites of *Microcystis aeruginosa* bloom material from Lake Kinneret, Israel, *J. Nat. Prod. 75*, 209—219.
17. Siebert, M., Severin, K., and Heide, L. (1994) Formation of 4-hydroxybenzoate in *Escherichia coli*: Characterization of the ubiC gene and its encoded enzyme chorismate pyruvate-lyase., *Microbiology 140*, 897-904.
18. Zamir, L. O., Nikolokakis, A., Bonner, C. A., and Jensen, R. A. (1993) Evidence for enzymatic formation of isoprephenate from isochorismate, *Bioorganic Med. Chem. Lett. 3*, 1441—1446.
19. Blasiak, L. C., and Clardy, J. J. (2010) Discovery of 3-formyl-tyrosine metabolites from *Pseudoalteromonas tunicata* through heterologous expression, *Am. Chem. Soc. 132*, 926—927.
20. Li, Y., Weissman, K., and Müller, R. (2008) Myxochelin biosynthesis: Direct evidence for two- and four-electron reduction of a carrier protein-bound thioester, *J. Am. Chem. Soc. 130*, 7554—7555.

21. Ishida, K., Christiansen, G., Yoshida, W., Kurmayer, R., Welker, M., Valls, N., Bonjoch, J., Hertweck, C., Börner, T., Hemscheidt, T., and Dittmann, E. (2007) Biosynthesis and structure of aeruginoside 126A and 126B, cyanobacterial peptide glycosides bearing a 2-carboxy-6-hydroxyoctahydroindole moiety, *Chem. Biol. 14*, 565—576.

22. Layer, G., Pierik, A. J., Trost, M., Rigby, S. E., Leech, H. K., Grage, K., Breckau, D., Astner, I., Jansch, L., Heathcote, P., Warren, M. J., Heinz, D. W., and Jahn, D. (2006) The substrate radical of *Escherichia coli* oxygen-independent coproporphyrinogen III oxidase HemN, *J. Biol. Chem. 281*, 15727—15734.

23. Rouhiainen, L., Jokela, J., Fewer, D. P., Urmann, M., and Sivonen, K. (2010) Two alternative starter modules for the non-ribosomal biosynthesis of specific anabaenopeptin variants in *Anabaena* (Cyanobacteria), *Chem. Biol. 17*, 265—273.

24. Kaljunen, H., Schiefelbein, S., Stummer, D., Kozak, S., Meijers, R., Christiansen, G., and Rentmeister, A. (2015) Structural elucidation of the bispecificity of A domains as a basis for activating non-natural amino acids, *Angew. Chem. Int. Ed. Engl. 54*, 8833—8836.

25. Meyer, S., Kehr, J., Mainz, A., Dehm, D., Petras, D., Süssmuth, R., and Dittmann, E. (2016) Biochemical dissection of the natural diversification of microcystin provides lessons for synthetic biology of NRPS, *Cell Chem. Biol. 23*, 462—471.

26. Shishido, T. K., Jokela, J., Fewer, D. P., Wahlsten, M., Fiore, M. F., and Sivonen, K. (2017) Simultaneous production of anabaenopeptins and namalides by the cyanobacterium *Nostoc* sp. CEN543, *ACS Chem. Biol. 12*, 2746—2755.

27. Fiore, M. F., Alvarenga, D. O., Varani, A. M., Hoff-Risseti, C., Crespim, E., Ramos, R. T., Silva, A., Schaker, P. D., Heck, K., Rigonato, J., and Schneider, M. P. (2013) Draft genome sequence of the Brazilian toxic bloom-forming cyanobacterium *Microcystis aeruginosa* strain SPC777, *Genome Announc. 1*, e00547—00513.

28. Castro, W. O., Lima, A. R., Moraes, P. H., Siqueira, A., Aguiar, D., Baraúna, A., Martins, L., Fuzii, H., de Lima, C., Vianez-Júnior, J., Nunes, M., Dall'Agnol, L., and Gonçalves, E. (2016) Draft genome sequence of *Microcystis aeruginosa* CACIAM 03, a cyanobacterium isolated from an Amazonian freshwater environment, *Genome Announc. 4*, e01299.

29. Neilan, B., Pearson, L., Muenchhoff, J., Moffitt, M., and Dittmann, E. (2013) Environmental conditions that influence toxin biosynthesis in cyanobacteria, *Environ. Microbiol. 15*, 1239—1253.

30. Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M., and Stanier, R. Y. (1979) Generic assignments, strain histories and properties of pure cultures of Cyanobacteria, *J. Gen. Microbiol. 111*, 1—61.
31. Vallenet, D., Calteau, A., Cruveiller, S., Gachet, M., Lajus, A., Josso, A., Mercier, J., Renaux, A., Rollin, J., Rouy, Z., Roche, D., Scarpelli, C., and Medigue, C. (2017) MicroScope in 2017: An expanding and evolving integrated resource for community expertise of microbial genomes, *Nucleic Acids Res. 45*, D517—D528.
32. Pancrace, C., Barny, M., Ueoka, R., Calteau, A., Scalvenzi, T., Pedron, J., Barbe, V., Piel, J., and Humbert, J. (2017) Insights into the *Planktothrix* genus: Genomic and metabolic comparison of benthic and planktic strains, *Sci. Rep. 7*, 41181.
33. Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., Lee, S. Y., Fischbach, M. A., Muller, R., Wohlleben, W., Breitling, R., Takano, E., and Medema, M. H. (2015) antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters, *Nucleic Acids Res. 43*, W237—243.

**FIGURE LEGENDS**

Figure 1. Aeruginoguanidines and the corresponding biosynthetic gene cluster. *(A)* The structure of aeruginoguanidines (AGDs), **1**; AGD-98A, **2**; AGD-98B, **3**; AGD-98C; *(B)* AGD biosynthetic gene cluster of *Microcystis aeruginosa* NIES-98 and its variation in 13 other *Microcystis* genomes sharing 94 to 98% of similarity. The genes are color-coded with orange for carbohydrate sulfotransferase, sulfotransferase and sulfatase; black for NRPS and thioesterase; green for methyl-, isoprenyl- and aminotransferase; blue for permease; grey for proteins with putative and unknown function. The blue line indicates the span of the MIC2 cluster previously described.[9]

Figure 2. Molecular network of AGD (A) and of MGD with MGA (B). Characterized structural variants are indicated as red-colored nodes and new congeners characterized in this study are highlighted in green. Diversity and distribution of AGD and MGD variants for each strain. Details of MN(+), MN(–), and the AGD and MGD networks are presented in Figures S1-S3.

Figure 3. Microguanidine and microguanidine amide variants detected in strains containing the AGD biosynthetic gene cluster. The MGD depsipeptides contain ester-bonds, while the peptidic MGAs contain amide-bonds in their structures. Details of the high-resolution MS data of MGA and MGD are presented in Table S3.

Figure 4. Proposed Hphpa and Hphpol biosynthesis. 3-hydroxy-*m*-hydroxymethyl-*p*-hydroxyphenylpyruvate is synthesized as a precursor of both Hphpa and Hphpol via several steps from isochorismate. The resulting intermediate is further transformed by either an aminotransferase (AgdL) or a reductase (AgdN or U) to yield Hphpa or Hphpol, respectively and further transferred to the free-standing PCP (AgdB) after adenylation by AgdA. The AgdL enzyme for which natural mutants were identified in the course of this study is highlighted in red.

Figure 5. Distribution of the known and unknown BGCs in the frame of the phylogeny of the 23 *Microcystis* genomes based on maximum likelihood tree built upon 586 marker genes. The known BGCs are involved in the synthesis of aeruginoguanidine (AGD) and microguanidine (MGD) and/or MGA only, of microcystin, of cyanobactins including aeruginosamide, of aeruginosin, of microviridin, of cyanopeptolin, of anabaenopeptin including ferintoic acid, and of microginin. One BGC only predicted in one strain is involved in synthesis of puwainaphycin.[12] The numbers indicate the unknown BGCs detected in the genome; the origin of each strain is indicated in parenthesis.

Figure 6. Proposed AGD, MGA and MGD biosynthetic pathways. **Top line**; AGD biosynthesis route: Hphpa which is linked to the free-standing PCP AgdB is transferred to AgdE and condensed with the dipeptide, which is derived from AgdK and E. The thioester-tethered intermediate is methylated by a radical SAM enzyme (AgdC) followed by decarboxylation and released from the enzyme. The resulting molecule is further modified by sulfation and farnesylation. **Middle line**: MGA (**6** and **7**) biosynthesis route, almost the same pathway as AGD biosynthesis, but only AgdE is used and the α-amino group of Arg is permethylated by AgdI or M. **Bottom line**; MGD (**4** and **5**) biosynthesis route, almost the same pathway as MGAs, but using Hphpol as the intermediate instead of Hphpa. The enzyme AgdK for which a natural mutant was identified in the course of this study is highlighted in red.

**TABLE**

Table 1. Proposed function of proteins encoded in the AGD gene cluster and flanking ORFs in *Microcystis aeruginosa* NIES-98. The strand position and the size of gene in amino acids are indicated with the corresponding Best BLASTp hit and identity, all found in *Microcystis* genomes. NRPS domains: C for condensation, A for adenylation with substrate prediction, PCP for peptidyl carrier protein, and nMT for N-methyl transferase.

## ASSOCIATED CONTENT

### Supporting Information

The supporting Information is available free of charge *via* the ACS Publications website at DOI

Methods of preparation of the extracts and of recovering complete AGD cluster, HLPC-MS measurement and molecular networking; four supporting tables and 18 supporting figures on the detailed molecular network and the spectra of the new structures, as indicated in the text (PDF).

### Accession Codes

New sequence data are archived in GenBank under accession numbers MH049490 to MH049500.

## AUTHOR INFORMATION


### Corresponding Author

* email: muriel.gugger@pasteur.fr

* email: editt@uni-potsdam.de

### ORCID

Muriel Gugger: 0000-0001-6728-1976

Enora Briand: 0000-0001-8996-0072

Elke Dittmann: 0000-0002-7549-7918

Christian Hertweck: 0000-0002-0367-337X

Douglas Gatte Pichi: 0000-0001-9164-8969

Thibault Scalvenzi: 0000-0002-5760-1574


### Notes

The author declare no competing financial interest

## ACKNOWLEDGMENTS

CP was supported by the Ile-de-France ARDoC Grant for PhD. Funding was provided by the Institut Pasteur. ED was supported by a grant of the German Research Foundation (DFG, Di910/10-1). Financial support by the DFG-funded Collaborative Research Centre ChemBioSys (SFB 1127) to ED and CH is gratefully acknowledged. We thank A. Perner and H. Heinecke for Thermo Exactive LC-MS measurements. All PCC cyanobacteria of this study are available from the Institut Pasteur. All data are contained in the main text and supplementary materials.

Table 1. Proposed function of proteins encoded in the AGD gene cluster and flanking ORFs in *Microcystis aeruginosa* NIES-98. The strand position and the size of gene in amino acids are indicated with the corresponding Best BLASTp hit and identity, all found in *Microcystis* genomes. NRPS domains: C for condensation, A for adenylation with substrate prediction, PCP for peptidyl carrier protein, and nMT for N-methyl transferase.

| **Gene (Strand)** | Size (aa) | Proposed function (NRPS with substrat prediction) | Best BLASTp hit (Accession number) | Identity (%) |
|---|---|---|---|---|
| **Orf (-)** | 160 | Conserved protein of unknown function | Hypothetical protein O53_4696 (ELP52967.1) | 100 |
| **agdP (+)** | 238 | Carbohydrate sulfotransferase II | Hypothetical protein O53_4419 (ELP52967.1) | 100 |
| **agdQ (+)** | 589 | Thiamine pyrophosphate enzyme | Acetolactate synthase large subunit (EPF22845.1) | 100 |
| **agdR (+)** | 296 | Sulfotransferase I | Sulfotransferase domain protein (ELP52945.1) | 100 |
| **agdS (+)** | 271 | Conserved protein of unknown function | Hypothetical protein O53_4433 (ELP52708.1) | 95 |
| **agdT (+)** | 274 | Conserved protein of unknown function | Conserved hypothetical protein (CCH98454.1) | 99 |
| **agdS' (+)** | 268 | Conserved protein of unknown function | Hypothetical protein MAESPC_01420 (EPF22841.1) | 99 |
| **agdT' (+)** | 270 | Conserved protein of unknown function | Putative uncharacterized ORF3 domain protein (ELP52673.1) | 99 |
| **agdU (+)** | 405 | Thioester reductase | Polyketide synthase hetM (CCI12982.1) | 98 |
| **agdE (-)** | 1093 | NRPS ($A_{Arg,/Lys/Orn}$-PCP-C) | Linear gramicidin synthase subunit D (EPF22838.1) | 98 |
| **agdD (-)** | 441 | Sulfotransferase III | Zinc chelation protein SecC (WP_069474152.1) | 100 |
| **agdC (-)** | 438 | Radical SAM | Radical SAM superfamily protein (ELP52520.1) putative oxygen-independent coproporphyrinogen III synthase | 100 |
| **agdB (-)** | 94 | Peptidyl carrier protein | Phosphopantetheine attachment site family protein (ELP52599.1) | 100 |
| **agdA (-)** | 473 | AMP-dependent synthetase and ligase | AMP-dependent synthetase (WP_069474153.1) | 100 |
| **agdF (+)** | 196 | Permease | Conserved hypothetical protein (CCI31673.1) | 97 |
| **agdG (+)** | 852 | Sulfatase | Sulfatase family protein (ELP52537.1) | 99 |
| **agdH (+)** | 191 | 4-Hydroxybenzoate synthetase | Hypothetical protein O53_4514 (ELP52787.1) | 100 |
| **agdI (+)** | 342 | O-Methyltransferase | Methyltransferase (WP_069474155.1) | 100 |
| **agdJ (+)** | 231 | Isoprenyl-transferase | Di-trans,poly-cis-decaprenylcistransferase (ELP52925.1) | 99 |
| **agdK (+)** | 1588 | NRPS ($A_{Arg}$-nMT-PCP-C) | Chondramide synthase cmdD (EPF22828.1) | 99 |
| **agdL (+)** | 455 | Aminotransferase | Uncharacterized aminotransferase yodT (CCI31679.1) | 99 |
| **Orf (+)** | 71 | Hypothetical protein | Hypothetical protein (WP_069474158.1) | 100 |
| **agdM (+)** | 346 | O-Methyltransferase | O-Methyltransferase family protein (ELP53140.1) | 99 |
| **agdN (+)** | 401 | Thioester reductase | Thioester reductase domain protein (ELP52682.1) | 99 |
| **agdO (+)** | 671 | ABC transporter | ABC Transporter transmembrane region 2 family protein (ELP52531.1) | 99 |
| **Orf (+)** | 671 | Conserved protein of unknown function | Hypothetical protein O53_4447 (ELP52722.1) | 99 |
| **Orf (+)** | 156 | Conserved protein of unknown function | Hypothetical protein O53_4299 (ELP52574.1) | 100 |
| **Orf (+)** | 554 | GUN4-like family protein | Hypothetical protein (WP_069474163.1) | 100 |

Figure 1. Aeruginoguanidines and the corresponding biosynthetic gene cluster. *(A)* The structure of aeruginoguanidines (AGDs), **1**; AGD-98A, **2**; AGD-98B, **3**; AGD-98C; *(B)* AGD biosynthetic gene cluster of *Microcystis aeruginosa* NIES-98 and its variation in 13 other *Microcystis* genomes sharing 94 to 98% of similarity. The genes are color-coded with orange for carbohydrate sulfotransferase, sulfotransferase and sulfatase; black for NRPS and thioesterase; green for methyl-, isoprenyl- and aminotransferase; blue for permease; grey for proteins with putative and unknown function. The dashed arrows under the cluster indicate the three operons. The blue line indicates the span of the MIC2 cluster previously described.[9]

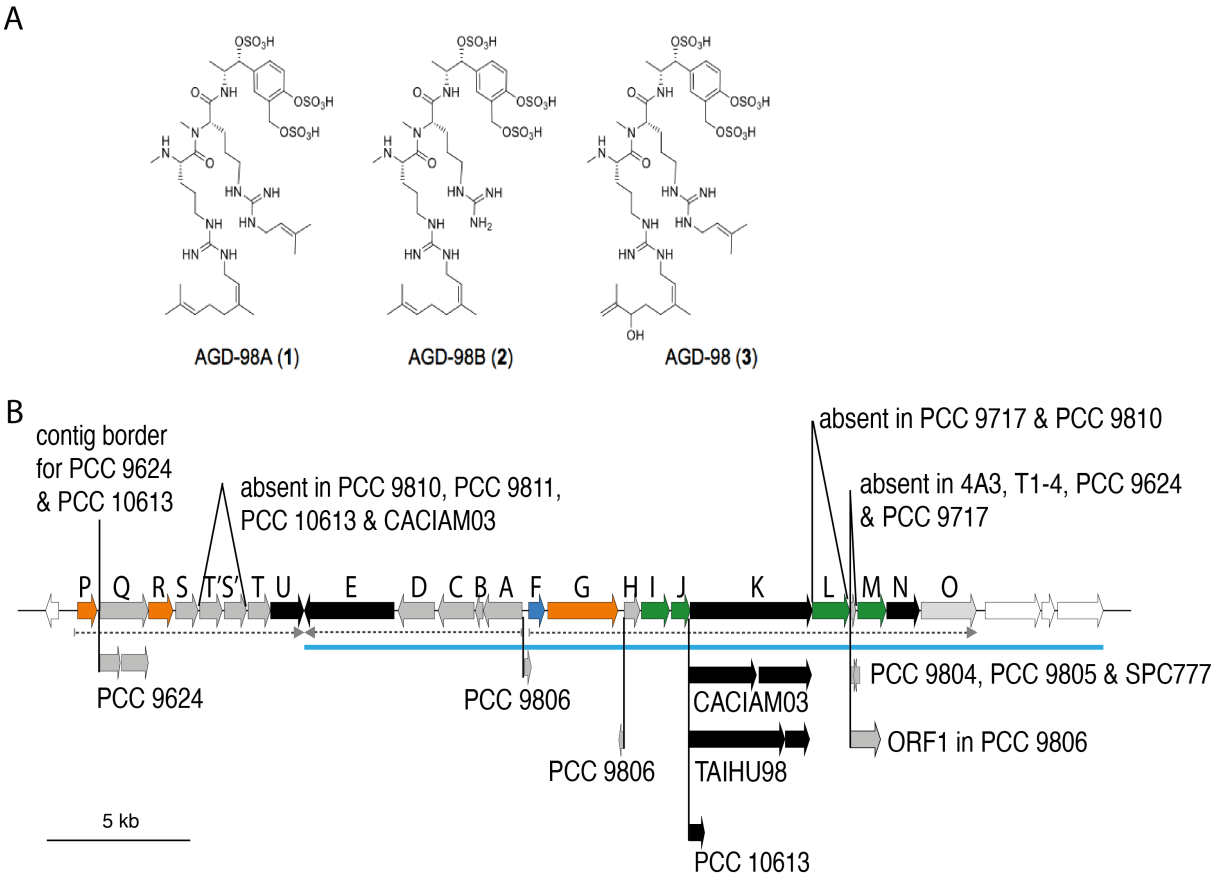
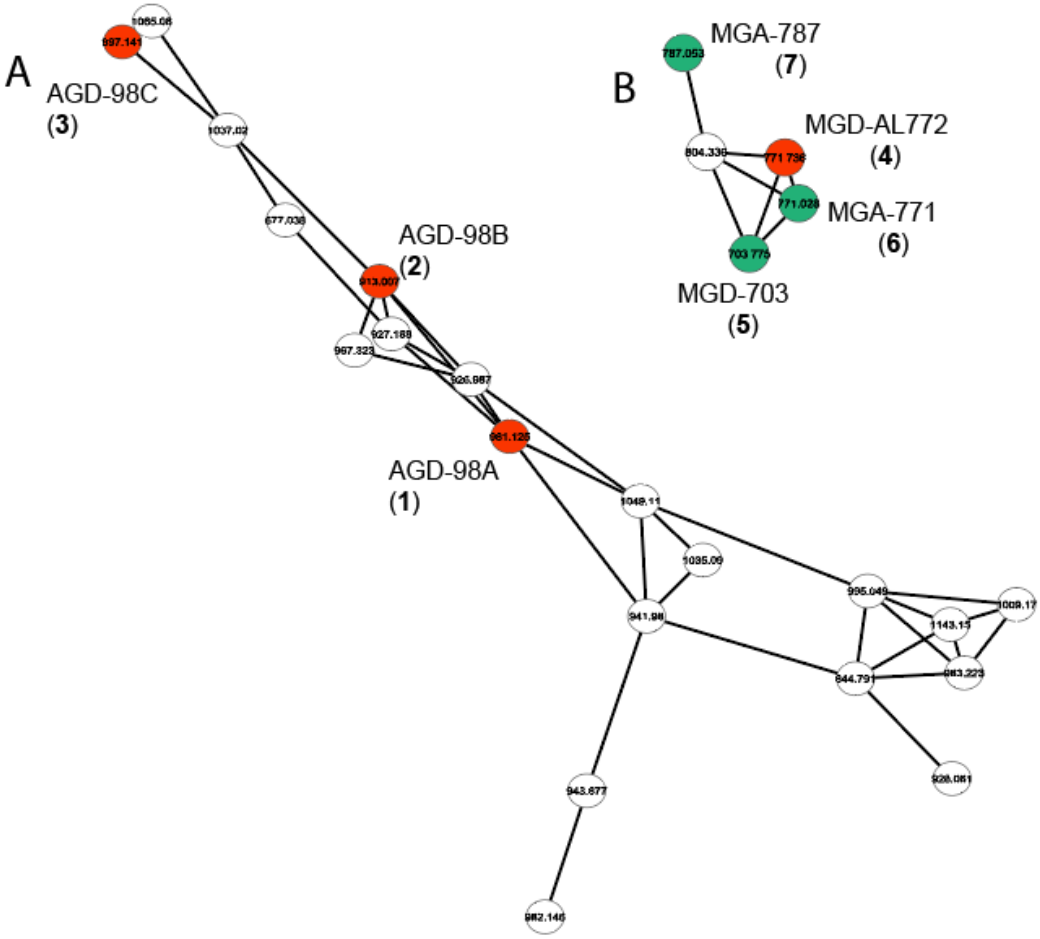
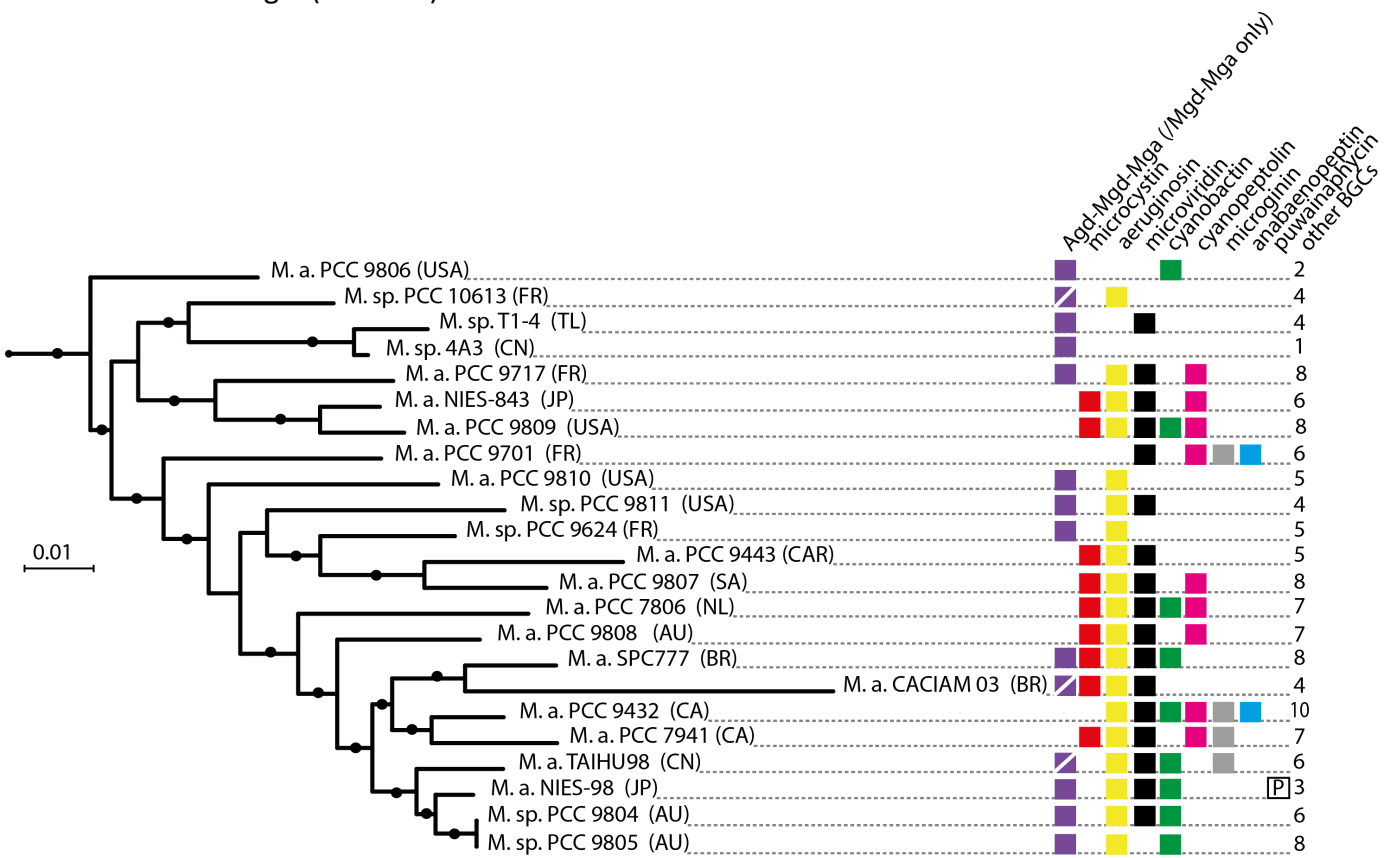


Figure 2. Molecular network of AGD (A) and of MGD with MGA (B). Characterized structural variants are indicated as red-colored nodes and new congeners charactized in this study are highlighted in green. Diversity and distribution of AGD and MGD variants for each strain. Details of MN(+), MN(–), and the AGD and MGD networks are presented in Figures S1-S3.




| Parent Mass | Strains | | | | | | | | | | Assignment |
|---|---|---|---|---|---|---|---|---|---|---|---|
| m/z (M-H)⁻ | NIES_98 | PCC 9624 | PCC 9717 | PCC 9804 | PCC 9805 | PCC 9806 | PCC 9810 | PCC 9811 | PCC 10613 | T1-4 | |
| 667.038 | | | | | | ■ | | | | | |
| 844.791 | | | | | | | | | | | |
| 913.007 | ■ | | | ■ | ■ | | | ■ | | ■ | AGD-98B (2) |
| 926.987 | | | | | | | | ■ | | | |
| 927.186 | | | | | | | | ■ | | ■ | |
| 928.061 | | | ■ | | | | ■ | | | | |
| 941.98 | | | | | | | ■ | | | | |
| 949.877 | | | | | | | ■ | | | | |
| 967.323 | ■ | | | ■ | ■ | | | | | | |
| 981.125 | ■ | ■ | ■ | ■ | ■ | ■ | | | | ■ | AGD-98A (1) |
| 982.146 | | | | | | ■ | | | | | |
| 983.223 | | | | | | ■ | | | | | |
| 995.049 | ■ | | | ■ | ■ | | | | | ■ | |
| 997.141 | ■ | | | ■ | ■ | ■ | | | | ■ | AGD-98C (3) |
| 1009.17 | | | | ■ | ■ | ■ | | | | | |
| 1035.09 | | | | ■ | ■ | ■ | | | | ■ | |
| 1037.02 | | | | ■ | ■ | | | | | | |
| 1049.11 | | | | ■ | ■ | | | | | ■ | |
| 1065.08 | | | | ■ | ■ | ■ | | | | | |
| 1143.15 | | | | | | | | | | ■ | |
| 703.775 | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | MGD-704 (5) |
| 771.028 | ■ | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | MGA-771 (6) |
| 771.736 | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | MGD-AL772 (4) |
| 787.053 | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | MGA-787 (7) |
| 804.336 | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | |

Figure 5. Distribution of the known and unknown BGCs in the frame of the phylogeny of the 23 *Microcystis* genomes based on maximum likelihood tree built upon 586 marker genes. The known BGCs are involved in the synthesis of aeruginoguanidine (AGD) and microguanidine (MGD) and/or microguanidine amide (MGA) only, of microcystin, of cyanobactins including aeruginosamide, of aeruginosin, of microviridin, of cyanopeptolin, of anabaenopeptin including ferintoic acid, and of microginin. One BGC only predicted in one strain is involved in synthesis of puwainaphycin.[12] The numbers indicate the unknown BGCs detected in the genome; the abbreviations in parenthesis after the name of the strain indicate its origin (Table S2).



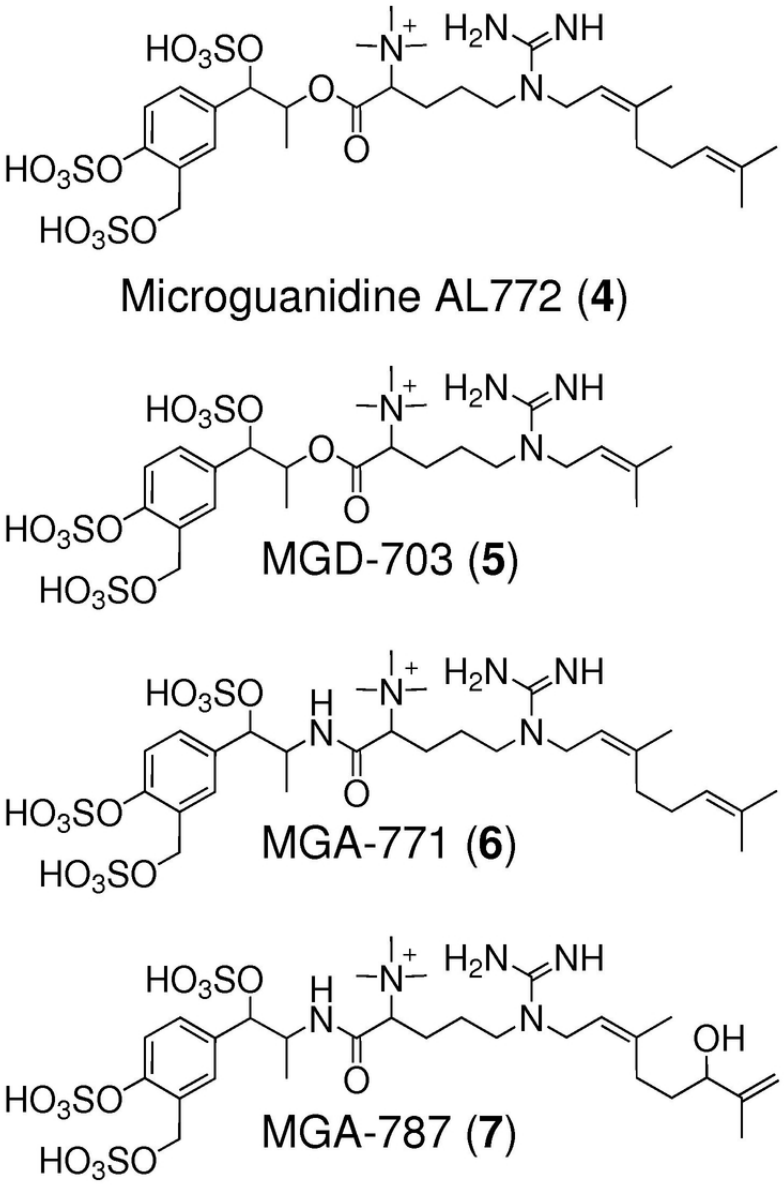



Figure 3. Microguanidine and microguanidine amide variants detected in strains containing the AGD biosynthetic gene cluster. The MGD depsipeptides contain ester-bonds, while the peptidic MGAs contain amide-bonds in their structures. Details of the high-resolution MS data of MGA and MGD are presented in Table S3.

63x94mm (300 x 300 DPI)

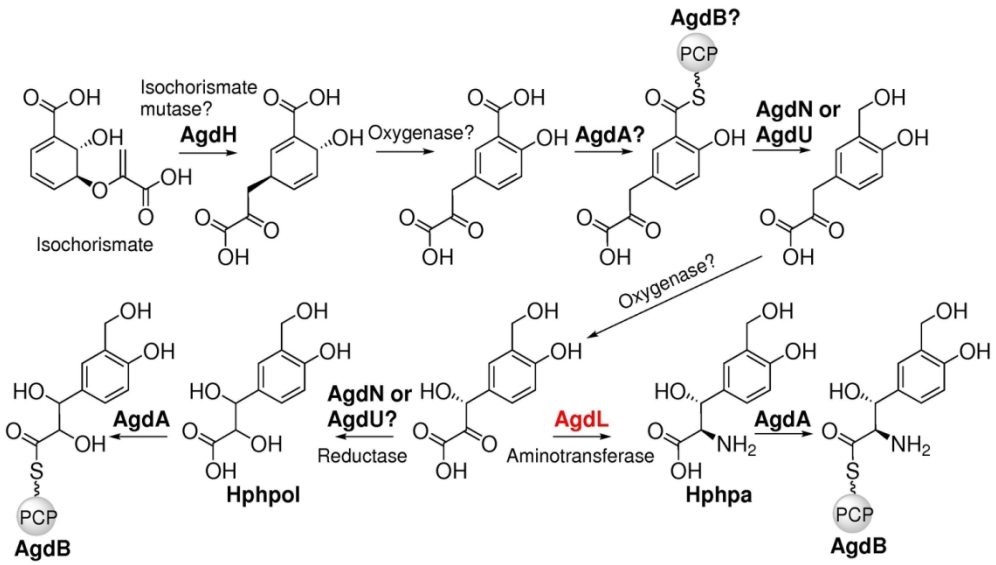



Figure 4. Proposed Hphpa and Hphpol biosynthesis. 3-hydroxy-m-hydroxymethyl-p-hydroxyphenylpyruvate is synthesized as a precursor of both Hphpa and Hphpol via several steps from isochorismate. The resulting intermediate is further transformed by either an aminotransferase (AgdL) or a reductase (AgdN or U) to yield Hphpa or Hphpol, respectively and further transferred to the free-standing PCP (AgdB) after adenylation by AgdA. The AgdL enzyme for which natural mutants were identified in the course of this study is highlighted in red.

130x74mm (300 x 300 DPI)

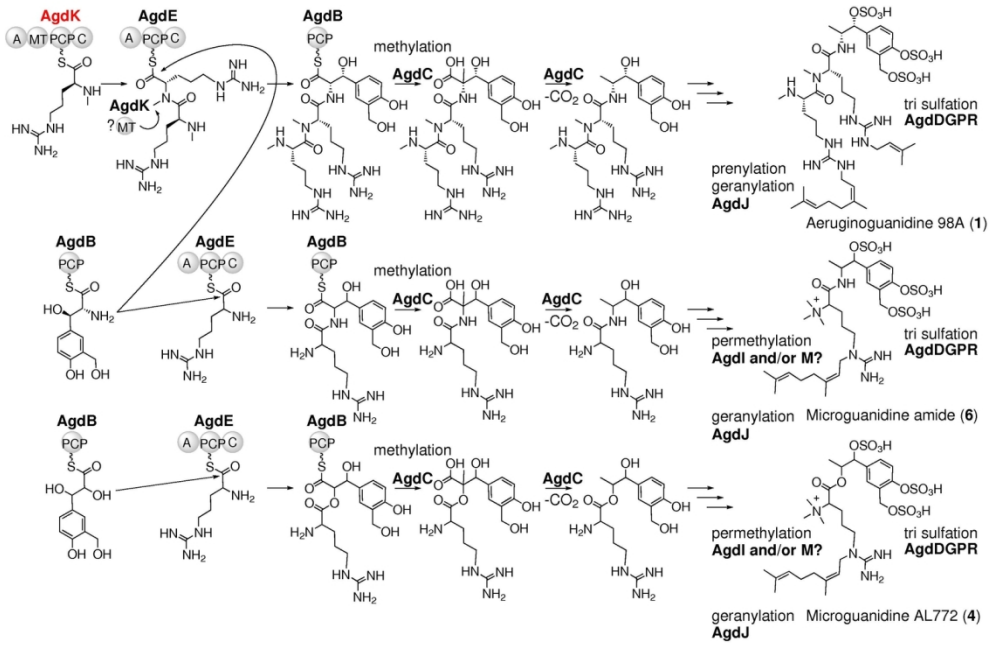
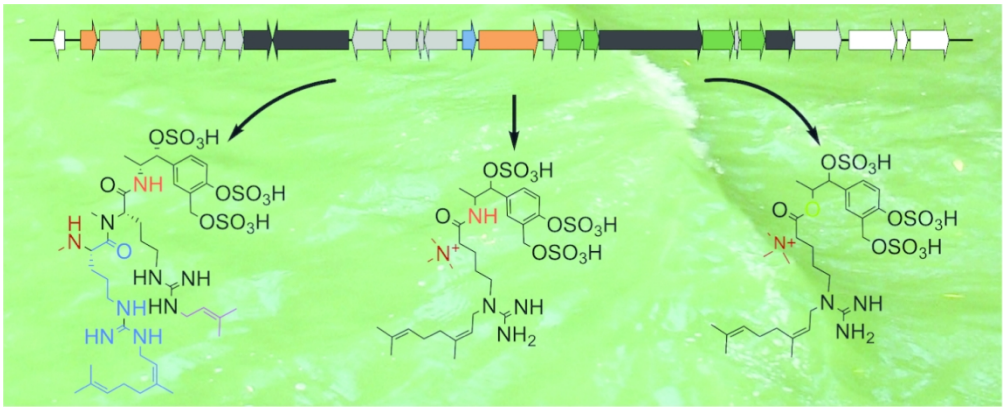



Figure 6. Proposed AGD, MGA and MGD biosynthetic pathways. Top line; AGD biosynthesis route: Hphpa which is linked to the free-standing PCP AgdB is transferred to AgdE and condensed with the dipeptide, which is derived from AgdK and E. The thioester-tethered intermediate is methylated by a radical SAM enzyme (AgdC) followed by decarboxylation and released from the enzyme. The resulting molecule is further modified by sulfation and farnesylation. Middle line: MGA (6 and 7) biosynthesis route, almost the same pathway as AGD biosynthesis, but only AgdE is used and the α-amino group of Arg is permethylated by AgdI or M. Bottom line; MGD (4 and 5) biosynthesis route, almost the same pathway as MGAs, but using Hphpol as the intermediate instead of Hphpa. The enzyme AgdK for which a natural mutant was identified in the course of this study is highlighted in red.

138x89mm (300 x 300 DPI)



A Unique Biosynthetic Pathway in Bloom-Forming *Microcystis* Jointly Assembles Cytotoxic Aeruginoguanidines and Microguanidines