Plasticity and Evolution of Aeruginosin Biosynthesis in Cyanobacteria[▽][†]

Keishi Ishida,[1,2] Martin Welker,[3,4,5]* Guntram Christiansen,[1,6,7] Sabrina Cadel-Six,[3] Christiane Bouchier,[8]
Elke Dittmann,[1] Christian Hertweck,[2] and Nicole Tandeau de Marsac[3]

*Institute of Biology, Humboldt Universität Berlin, 10115 Berlin, Germany[1]; Leibniz Institute for Natural Product Research and
Infection Biology (HKI), 07745 Jena, Germany[2]; Institut Pasteur, Unité des Cyanobactéries, CNRS URA 2172, F75015 Paris,
France[3]; Institute of Chemistry, Technische Universität Berlin, 10587 Berlin, Germany[4]; Anagnostec GmbH, 14476 Potsdam,
Germany[5]; Department of Chemistry, University of Hawaii at Manoa, Honolulu, Hawaii 96822[6]; Institute of
Limnology, Österreichische Akademie der Wissenschaften, 5310 Mondsee, Austria[7]; and Institut Pasteur,
Pasteur Genopole Ile-de-France, F75015 Paris, France[8]*

Aeruginosins are bioactive oligopeptides that are produced in high structural diversity by strains of the bloom-forming cyanobacterial genera *Microcystis* and *Planktothrix*. A hallmark of aeruginosins is the unusual Choi moiety central to the tetrapeptides, while other positions are occupied by variable moieties in individual congeners. Here we report on three aeruginosin synthetase gene clusters (*aer*) of *Microcystis aeruginosa* (strains PCC 7806, NIES-98, and NIES-843). The analysis and comparison the *aer* gene clusters provide the first insight into the molecular basis of biosynthetic and structural plasticity in aeruginosin pathways. Major parts of the *aer* gene clusters are highly similar in all strains, particularly the genes coding for the first three nonribosomal peptide synthetase (NRPS) modules except for the region coding for the second adenylation domain. However, the gene clusters differ largely in genes coding for tailoring enzymes such as halogenases and sulfotransferases, reflecting structural peculiarities in aeruginosin congeners produced by the individual strains. Significant deviations were further observed in the C-terminal NRPS modules, suggesting two distinct release mechanisms. The architecture of the gene clusters is in agreement with the particular aeruginosin variants that are produced by individual strains, the structures of two of which (aeruginosins 686 A and 686 B) were elucidated. The *aer* gene clusters of *Microcystis* and *Planktothrix* are proposed to originate from a common ancestor and to have evolved to their present-day diversity largely through horizontal gene transfer and recombination events.

Bioactive oligopeptides exhibiting an exceptional structural diversity are produced by various cyanobacterial taxa (50). The majority of these peptides are considered to be synthesized by a nonribosomal peptide synthetase (NRPS) pathway (8), like the hepatotoxic microcystins (6) or the trypsin-inhibitory anabaenopeptilides/cyanopeptolins (33). Nevertheless, for several types of peptides, a pathway via ribosomal peptide synthesis followed by posttranslational modifications is anticipated (38, 51, 52). A characteristic of cyanobacterial nonribosomal peptides is the striking structural variety within particular peptide classes resulting from variations in the amino acid sequences and from modifications like N-methylation, chlorination, sulfatation, or glycosylation (50).

Aeruginosins represent a highly intriguing class of linear tetrapeptides with unusual (4-hydroxy)phenyllactic acid (Hpla) and 2-carboxy-6-hydroxyoctahydroindole (Choi) moieties and an arginine derivative at the C terminus (18) (Fig. 1). Most aeruginosins are strong inhibitors of serine proteases (7), and the inhibiting mechanisms have been elucidated by X-ray crystallography of aeruginosin-protease complexes (14, 37). The Choi moiety is present in all aeruginosins as well as the Hpla and phenyllactic acid (Pla) units that are found predominantly

in *Microcystis* and *Planktothrix*, respectively. However, the second position is occupied by a variable proteinogenic amino acid, generally in the D configuration. As C-terminal arginine mimetic, agmatine, argininol, and argininal units have been reported for *Microcystis* (18), the rare 1-amino-2-(*N*-amidino-Δ^3-pyrrolinyl)-ethyl (Aeap) has been found exclusively in *Planktothrix* (16). Because further modifications can occur at Hpla (mono- or dichlorination and sulfatation) and Choi (sulfatation, glycosylation, and chlorination), a potential number of more than 500 congeners is conceivable (49). Although only a fraction of these is known, this scenario is reflected by the high level of diversity of aeruginosins found in natural populations of *Microcystis* and *Planktothrix* (47, 48). Individual strains may produce different sets of peptide metabolites despite a very close relatedness, as established by comparisons of highly variable genomic sequences like the 16S-23S rRNA gene internal transcribed spacer (3).

Recently, the aeruginosin biosynthesis gene cluster of a *Planktothrix* strain (*Planktothrix agardhii* NIVA CYA 126-8) has been sequenced (GenBank accession number AM071396) (CAM59601.1 to CAM59614.1), and the structures of the corresponding peptides (aeruginosides 126 A and B) were elucidated (16). In the present study, the aeruginosin synthetase clusters of two *Microcystis* strains have been analyzed together with a third gene cluster identified in the publicly available genome sequence of *Microcystis aeruginosa* NIES-843 (19). Chemical analyses revealed that, in fact, all strains produce distinct aeruginosin congeners. The objective of the present study was to explore the molecular basis for the variability in
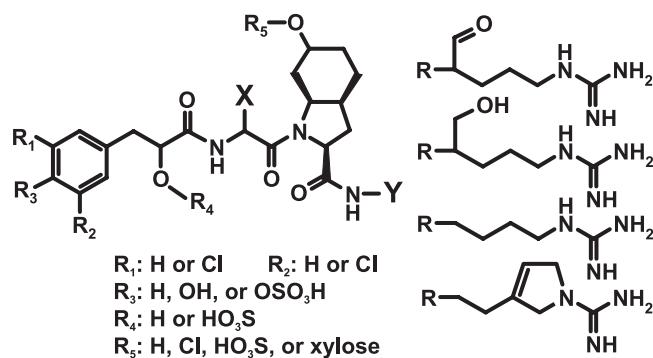
FIG. 1. Schematic structure of aeruginosin congeners. Variable sites at the N-terminal aryl acid (Hpla or Pla) and at the Choi are indicated. X denotes different amino acids that can occupy the second position: leucine, isoleucine, phenylalanine, tyrosine, and homotyrosine. In most aeruginosins, these amino acids are in the D configuration, but the L configuration has also been observed. Y denotes the C-terminal variable moieties as shown to the right (from top): argininal (linear tautomer), argininol, agmatine, and Aeap.

peptide production by closely related strains of *Microcystis*. The comparative analysis provides the first insight into the genetic basis for the structural variability of aeruginosin-type peptides in cyanobacteria. Insight into the evolution of the structural diversity of (cyanobacterial) nonribosomal peptides could offer valuable clues to the exploitation of the corresponding genes for the design of new therapeutics by pathway engineering approaches (36).

MATERIALS AND METHODS

Bacterial strains and culture conditions. *Microcystis aeruginosa* NIES-98, isolated in 1982 from Lake Kasumigaura (Japan), was obtained from the Microbial Culture Collection at the National Institute for Environmental Studies (Japan). Wild-type and mutant strains were grown at 23°C in MA or BG11 medium under continuous light provided by Osram Universal White fluorescent tubes (30 μE m^{-2} s^{-1}) with shaking at 110 rpm.

Microcystis aeruginosa PCC 7806, isolated in 1972 from the Braakman Reservoir (The Netherlands), is held as an axenic strain at the Pasteur Culture Collection of Cyanobacteria, Institut Pasteur (France). Strains were grown at 25°C in 50 ml of BG11$_0$ medium supplemented with 2 mM NaNO$_3$ and 10 mM NaHCO$_3$ (http://www.pasteur.fr/recherche/banques/PCC) under continuous light provided by Osram Universal White fluorescent tubes (30 μE m^{-2} s^{-1}). Cells in the exponential growth phase (optical density at 750 nm of 0.4 to 0.6) were harvested by centrifugation (10,000 $\times$ *g* for 10 min at 25°C) and lyophilized.

Microcystis aeruginosa NIES-843 was isolated from Lake Kasumigaura, Japan, in 1997. The complete genome of this strain has been sequenced (19).

All three *Microcystis* strains have to be considered to be very closely related as indicated by 16S-23S rRNA gene internal transcribed spacer sequence identity of 94 to 95% (3, 19).

Analysis of peptides. *M. aeruginosa* NIES-98 was previously described to be a producer of aeruginosin 98 (28). For the present study, dried cells were analyzed by matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) for oligopeptides in a mass range of 500 to 2,000 Da (48), as was done for *M. aeruginosa* NIES-843.

M. aeruginosa PCC 7806 was analyzed for aeruginosin-type peptides by high-performance liquid chromatography (HPLC) fractionation followed by MALDI-TOF MS. An extract was prepared by sequential extraction (50% methanol [MeOH], 75% MeOH, and 100% MeOH) of ~100 mg of dry cell material, which was fractionated by HPLC on an analytical rp-C$_{18}$ column (LiChroSpher; Merck, Darmstadt, Germany) by applying an acetonitrile-water (with 0.05% [vol/vol] trifluoroacetic acid [TFA]) gradient from 20 to 80% acetonitrile in 40 min. The absorbance was monitored at 220 nm, and peak fractions were collected, treated, and analyzed by MALDI-TOF MS as described previously (5).

Based on HPLC/MALDI-TOF MS data, aeruginosins from cells grown in 20-liter batch cultures were isolated by semipreparative HPLC. Freeze-dried

cells (~ 40 g [dry weight]) were extracted twice with 80% MeOH (10 ml/100 mg [dry weight]) and once with 100% MeOH (10 ml/100 mg [dry weight]). From combined methanolic extracts, methanol was evaporated, and the remaining aqueous phase was washed with diethylether. The aqueous layer was then extracted with *n*-butanol. The *n*-butanol fraction was evaporated under reduced pressure and dissolved in 20% aqueous acetonitrile, applied onto an ODS column (YMC GEL ODS A 60-S50, 5 by 28 cm), and eluted consecutively with aqueous acetonitrile (20%, 40%, and 80%), MeOH, and CH$_2$Cl$_2$. The 80% acetonitrile fraction was subjected to reversed-phase (RP) HPLC (Nucleosil 100-5C$_{18}$, 21 by 250 mm) with 10 min of loading (39% eluent B), followed by a linear gradient to 100% eluent B in 35 min at a flow rate of 10 ml/min and detection at 220 nm (eluent A, H$_2$O plus 0.1% [vol/vol] TFA; eluent B, 83% acetonitrile). The fraction containing aeruginosins 686 A and 686 B was resubjected to RP HPLC (Supelcosil LC-18, 21.2 by 250 mm) with the same gradient system to yield a fraction containing a mixture of aeruginosins 686 A and 686 B.

To elucidate the stereochemistry of tyrosine, a fraction containing aeruginosins (4.0 mg) was hydrolyzed with 6 M HCl containing 0.1% phenol (1.0 ml) at 105°C overnight. One hundred microliters of 1 M NaHCO$_3$ and Marfey's reagent (1-fluoro-2-4-dinitrophenyl-5-L-alanine amide [FDAA]; 50 μl of 10 mg L-FDAA ml^{-1} in acetone) was added to the reaction residue mixture, from which the solvent was removed by reduced pressure (23). The reaction mixture was incubated at 50°C for 1 h before being quenched by 50 μl 2 N HCl and diluted with 200 μl 50% aqueous acetonitrile. L-FDAA derivatives were analyzed by RP HPLC on a Cosmosil 5C$_{18}$-MS column (4.6 by 250 mm) at a flow rate of 1 ml min^{-1} and UV detection at 340 nm, applying a gradient program for eluent A (water plus 0.1% [vol/vol] TFA) and eluent B (acetonitrile plus 0.1% [vol/vol] TFA): A:B, 80:20 at 0 min, to A:B, 65:35 at 30 min, to A:B, 45:55 at 50 min, to A:B, 0.5:99.5 at 51 min. Retention times of standard amino acids were 28.88 min for L-FDAA-L-Tyr, 31.93 min for L-FDAA-D-Tyr, 48.42 min for *N*,*O*-di-L-FDAA-L-Tyr, and 51.72 min for *N*,*O*-di-L-FDAA-D-Tyr.

NMR analysis. ^{1}H and ^{13}C nuclear magnetic resonance (NMR) spectra were obtained with a Bruker Avance DRX 500 instrument in hexadeuterated dimethyl sulfoxide at 30°C.

Sequencing of *aer* biosynthesis gene clusters. The sequence of the genome of *M. aeruginosa* PCC 7806 has been obtained as described previously and was deposited in the EMBL database (11). In the genomic sequence, several NRPS genes were identified. Besides the expected microcystin and cyanopeptolin synthetase gene clusters (*mcy* and *mcn*, respectively) (44, 45), further NRPS modules that did not belong to either of these gene clusters were identified and were recognized as putative aeruginosin synthetase genes (GenBank accession number AM778955.1) (mic0223 to mic0237) (16).

Chromosomal DNA was isolated from *M. aeruginosa* NIES-98 (10). DNA manipulations were performed using *Escherichia coli* XL1 Blue cells grown in LB medium. *E. coli* XL1 Blue MRF′ Kan was used for the excision of lambda phagemids according to the suggestions of the manufacturer (Stratagene, La Jolla, CA). A lambda ZAP library was constructed using the supplied protocol (Stratagene, La Jolla, CA) and screened by Southern hybridization using a probe amplified from *aerB* of *Planktothrix agardhii* NIVA CYA 126-8 (Ae13fw [5′-CA GGAGACGGAGGTTATACCC-3′] and Ae13rv [5′-GGAGAAATCAATAAA TGGGTTTCTAGTTC-3′]; 571 bp). A single positive phagemid covered a region from *aerB* to *aerL*. The further sequencing was performed by primer walking with primers based on the *aer* gene cluster of *M. aeruginosa* PCC 7806 and an inverse PCR approach. Individual sequences of some 500 to 1,000 bp were assembled to yield the complete gene cluster.

A further aeruginosin synthetase gene cluster was identified in the genome of *M. aeruginosa* NIES-843 (19) by sequence comparison (GenBank accession number AP009552) (MAE56590 to MAE56480).

DNA sequencing and sequence analysis. Sequencing of phagemids and PCR products was performed by gene walking using primers with a Big Dye terminator cycle sequencing kit. Analysis of DNA sequences and deduced amino acid sequences were performed using the NCBI (National Institute for Biotechnology Information, Bethesda, MD) BLAST server (http://www.ncbi.nlm.nih.gov) and the ExPASy proteomics server (http://www.expasy.ch) of the Swiss Institute of Bioinformatics (Geneva, Switzerland). Sequence alignments were performed with ClustalX (43), and DNA polymorphism was assessed with DnaSP (35). Phylogenetic analyses of NRPS domains were done using different methods in MEGA4 (41). Binding pocket signature sequences (40) were extracted from adenylation domain sequences and compared according to data described previously by Rausch et al. (32) and May et al. (26) for aryl acid adenylation domains.

Construction of the gene disruption construct. A 3,611-bp gene fragment containing *aerKDEF* was amplified by PCR using primers Ae4rv4 (5′-TAG GTC AAT TTA CCA GCC ATC-3′) and Reduct_rv (5′-ATG CTG ACG AGG ATA
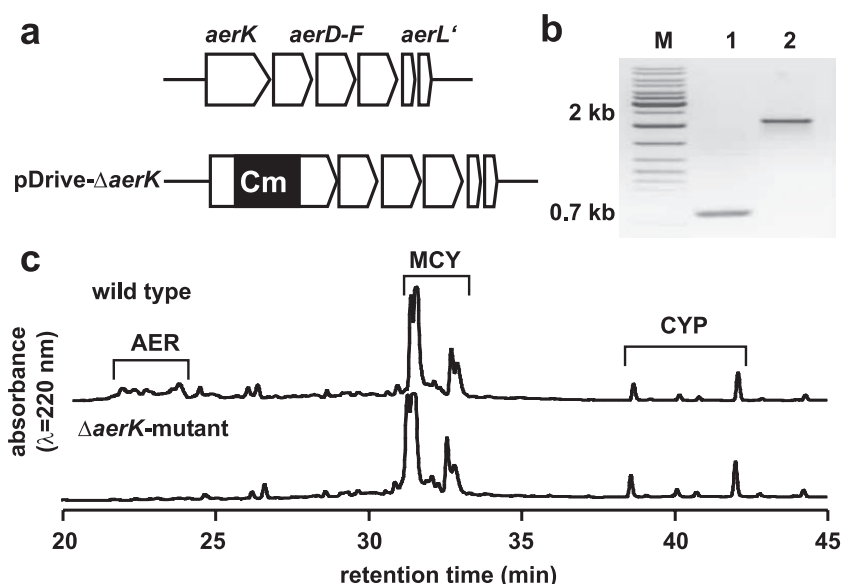
FIG. 2. ΔaerK mutant of *Microcystis aeruginosa* PCC 7806. (a) Schematic illustration of the knockout construct. (b) Agarose gel of PCR amplification products of *aerK* of the wild type (lane 1) and the ΔaerK mutant (lane 2). (c) HPLC chromatograms of extracts of the wild type and the ΔaerK mutant. Brackets indicate peptide types. AER, aeruginosins; MCY, microcystins; CYP, cyanopeptolins.

AAC C-3′) with *Pfu* polymerase (Stratagene, La Jolla, CA). The PCR product, which was treated with *Taq* polymerase to add A overhangs (72°C for 30 min), was purified by use of a gel purification kit (QIAquick; Qiagen, Hilden, Germany) and cloned into the pDrive vector (Qiagen, Hilden, Germany). The obtained plasmid was digested with MscI (cutting once in *aerK*) and end blunted by Klenow treatment. The 1.4-kb BsaAI fragment from pACYC184 containing the chloramphenicol resistance cassette was ligated into the blunted plasmid, thus yielding the gene disruption plasmid pDrive-ΔaerK (Fig. 2).

DNA manipulation of strain PCC 7806. The knockout construct (15 to 20 μg of plasmid DNA) was linearized with BamHI at the multiple-cloning site of the vector. The column-purified linear construct was heated at 95°C for 10 min and immediately put on ice. A 50-ml culture of PCC 7806 in the exponential phase was centrifuged at $4,600 \times g$, followed by three washing steps with 1 mM HEPES. The cell pellet was resuspended in 500 μl 1 mM HEPES, 200 μl of which was subjected to electroporation with 10 μg of the knockout plasmids (1.0 kv, 25 μF, and 200 Ω).

The thus-treated cells were spread onto BG11 agarose plates and incubated under the above-mentioned conditions. After 2 days, 200 μl of chloramphenicol (1 mg ml^{-1}) was applied onto one side of the plate below the agarose. After a few weeks, antibiotic-resistant colonies were selected and transferred onto a fresh BG11 agarose plate containing 0.5 μg ml^{-1} chloramphenicol. Transformants were purified by stepwise increasing of the chloramphenicol concentration up to 5 μg ml^{-1} (4 weeks to 8 weeks). Selected resistant colonies were inoculated in BG11 medium with initially 0.5 μg ml^{-1} chloramphenicol, increasing to 5 μg ml^{-1} chloramphenicol.

Potential knockout mutants were first screened by MALDI-TOF MS directly from colonies on agarose plates (47). Colonies that were negative for aeruginosins were grown in liquid medium until the cell biomass was sufficient for further analyses. The peptide content of ΔaerK mutants was analyzed in detail by analytical RP HPLC (YMC-Pack R&D ODS, 4.6 by 250 mm, 5 μm, 120 Å; YMC Co., Ltd., Kyoto, Japan) with photodiode array detection and a gradient of 10 to 100% acetonitrile in 40 min.

Nucleotide sequence accession number. The sequence of the *aer* gene cluster of *M. aeruginosa* NIES-98 is available under GenBank accession number FJ609416.

RESULTS AND DISCUSSION

Aeruginosins produced by *Microcystis aeruginosa* PCC 7806 and NIES-98. To survey the peptide production of *Microcystis* strains, especially with respect to the aeruginosin congeners,

cell extracts were investigated by MALDI-TOF MS and HPLC-MS, respectively.

In cells of *M. aeruginosa* NIES-98, aeruginosins 98 A and B (18) were detected by MALDI-TOF MS. In positive ion extraction mode, both were detected as M − SO$_3$ + H$^+$ (*m/z* 609.32 and *m/z* 575.40, respectively), while aeruginosin 98 C (with a brominated Hpla) was not detected. A further congener was detected with *m/z* 643.32 and identified by postsource-decay fragmentation as aeruginosin 101-SO$_3$, a congener with a dichlorination at Hpla but otherwise identical to aeruginosin 98 A. All three aeruginosins were detected as molecular ions (M − H$^−$) in negative ion extraction mode with an Δ*m/z* of 77.94 Da compared to the positive ion extraction mode, corresponding to the sulfate group that is lost in the positive ion extraction mode. Besides aeruginosins, aeruginoguanidins were also detected (17).

For *M. aeruginosa* PCC 7806, the mass spectrometric analysis of whole cells revealed mass signals of microcystin-LR and cyanopeptolins A, B, and 963A (2, 24) with high intensities, while a peak at *m/z* 687.29, corresponding to aeruginosin 686, which was previously detected in field samples (49), was only of low intensity. Only in HPLC fractions (Fig. 2) could sufficient material be collected to obtain a fragmentation pattern that corresponded to that of a sulfate-deficient but chlorinated aeruginosin 102 (18). For further analyses, peak fractions were collected to yield a pure compound. The rapid tautomerization, however, hampered its analysis by NMR, a problem that was encountered previously with aeruginosins bearing an arginal unit at the C terminus (18, 25).

Attempts to modify the aldehyde in the arginal unit to avoid tautomerization (18) and to isolate a single tautomer were unfruitful (25). Therefore, the mixture of tautomers of aeruginosin 686 was investigated by NMR spectroscopy. The ^{1}H and ^{13}C NMR spectra indicated the presence of two aerugi-
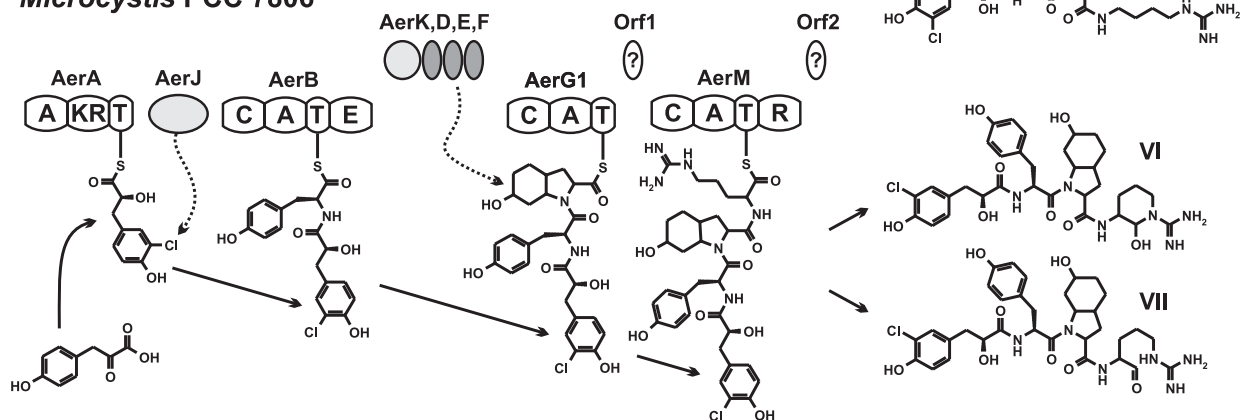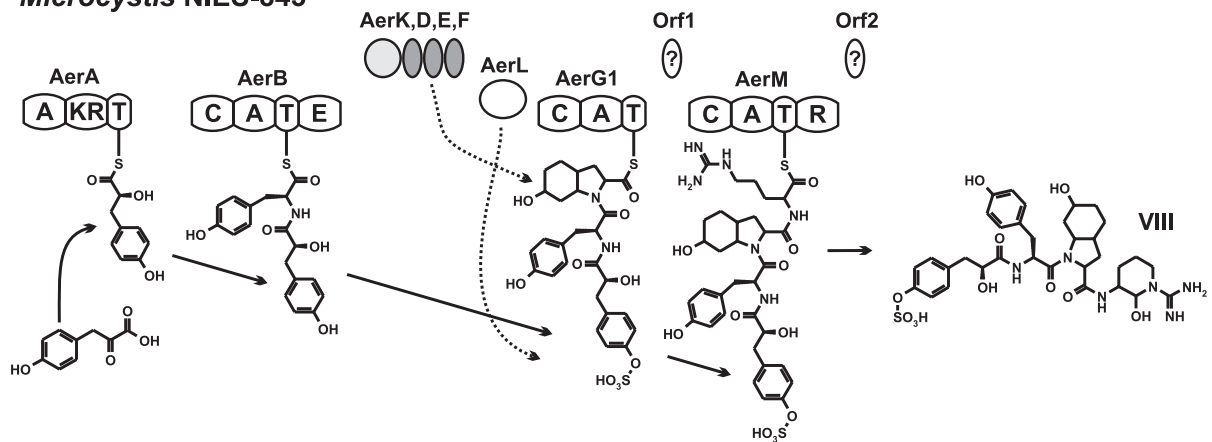
Planktothrix NIVA CYA 126-8
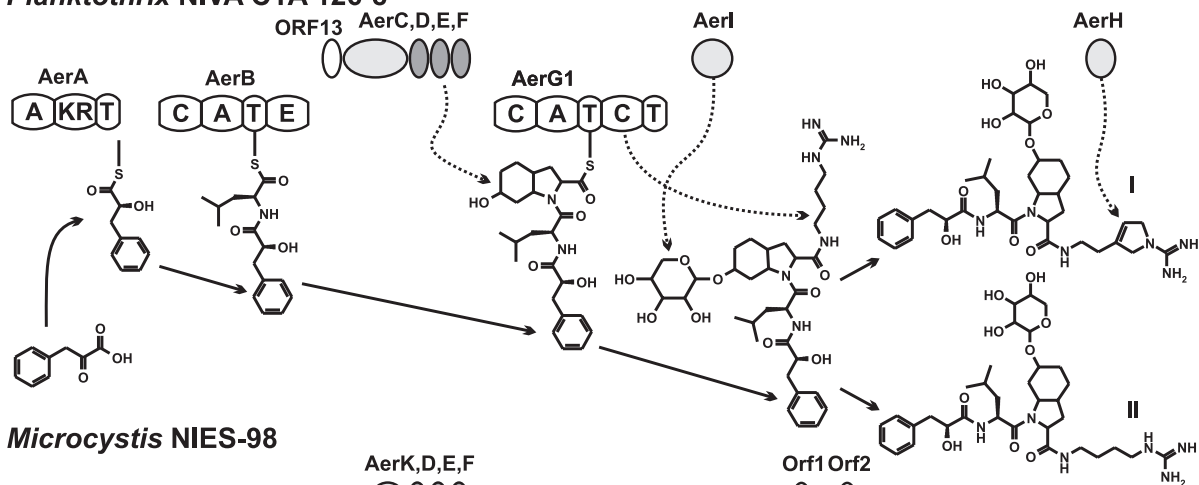
Microcystis NIES-98
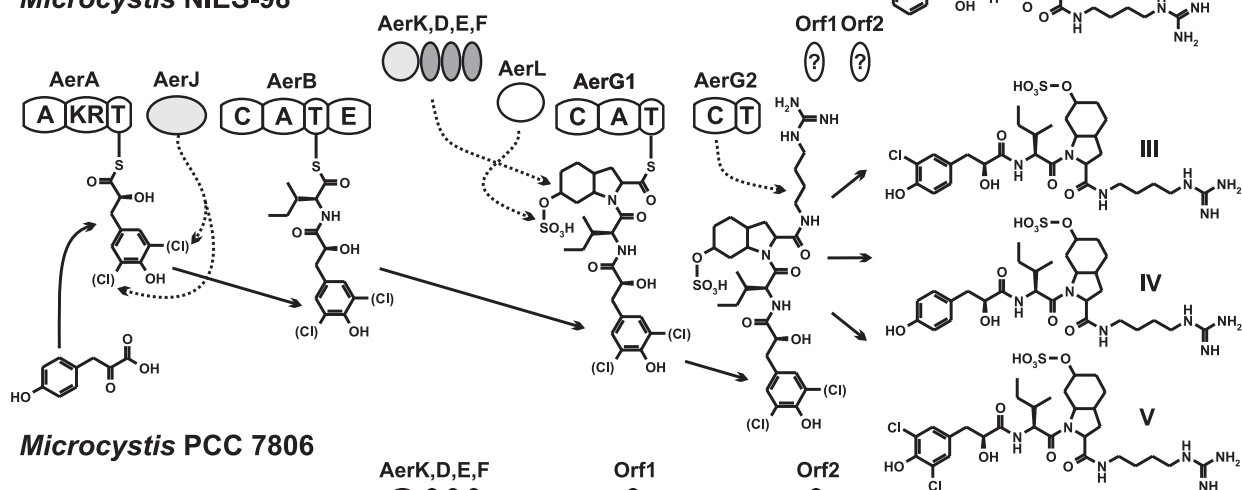
Microcystis PCC 7806

Microcystis NIES-843

nosins, aeruginosins 686 A and 686 B. However, more-detailed analyses of two-dimensional NMR data suggested two major and two minor aeruginosins; the latter ones have minor chemical shifts and are a mixture of rotamers (25, 46).

The two-dimensional NMR data including ^{1}H-^{1}H correlation spectroscopy (COSY), heteronuclear multiple-quantum correlation, and heteronuclear multiple-bond correlation (HMBC) spectra suggested that aeruginosin 686 A consists of four units: 3-Hpla, tyrosine, Choi, and argininal (Argal). Chloro-Hpla and Tyr were well assigned by HMBC correlations (see Fig. S1 and Table S1 in the supplemental material), while the overlapping of methylene proton signals from the Choi and Argal units hampered further signal assignments. Typical chemical shifts, however, such as three methine δ_H 4.22/δ_C 59.4, δ_H 4.19/δ_C 56.8 (Choi H2/C2 and H7a/C7a), and δ_H 3.92/δ_C 45.3 (Argal H2/C2), adjacent to nitrogen; oxymethine δ_H 3.31/δ_C 66.5 (Choi H6/C6); and oxymethine δ_H 5.27/δ_C 74.9 (Argal H1/C1), adjacent to nitrogen, were observed in the heteronuclear multiple-quantum correlation spectrum. One spin system from Choi H2 to H7a was observed in the ^{1}H-^{1}H COSY spectrum, and some HMBC correlations supported this spin system (see Fig. S1 in the supplemental material). Another spin system from Argal OH to H4/NH was revealed by ^{1}H-^{1}H COSY data. Although no ^{1}H-^{1}H COSY and HMBC correlations from methylene (Argal H5/C5) adjacent to nitrogen were observed, the correlation was solved (δ_H 3.17, 3.47/δ_C 39.4) by a comparison of the NMR data with those for other aeruginosins (18, 21, 28). The sequences of these units were determined by HMBC correlations, but the correlations between Tyr and Choi were not observed. The proposed structure of aeruginosin 686 A is in full agreement with PSD fragmentation data and amino acid analysis.

Chiral chromatography of a hydrolysate revealed two peaks at 31.92 and 51.71 min, corresponding to FDAA derivatives of D-Tyr, while no peaks indicating L-Tyr were encountered.

The overall structure of aeruginosin 686 B was also elucidated, differing from aeruginosin 686 A by the tautomerization of the argininal moiety (Fig. 3). Two further minor (rotamer) components, however, could not be identified due to weak signals.

In *M. aeruginosa* NIES-843, an aeruginosin-type peptide was detected with m/z 653.33 in the positive ion extraction mode and m/z 731.30 in the negative ion extraction mode by MALDI-TOF MS. This corresponds to aeruginosin 102, and the PSD fragment pattern was found to be identical to that of aeruginosin 102 detected in cell material of *M. aeruginosa* NIES-102.

In three *Microcystis* strains, six aeruginosin congeners were identified (congeners I to VI) (Fig. 3), varying in the degree of chlorination and sulfatation of Hpla, the amino acid in the second position, the sulfatation of Choi, and the C-terminal moiety.

Physical organization of the aeruginosin synthetase gene cluster in four cyanobacterial strains. The total size of the *aer* gene clusters is 25 kb in *M. aeruginosa* NIES-98 and NIES-843 and 27 kb in *M. aeruginosa* PCC 7806, respectively (Fig. 4). In all three clusters, NRPS genes were found together with genes for tailoring enzymes (see Table S2 in the supplemental material). The first gene, *aerA*, coding for a hybrid NRPS/polyketide synthase enzyme, and two further genes coding for NRPS modules (*aerB* and *aerG1*) are present in all three *Microcystis* strains. These genes are also found in *Planktothrix agardhii* NIVA CYA 126-8 with the constraint that *aerG1* corresponds to the 5′ part of a larger gene in *Planktothrix*, *aerG*. The second gene in *M. aeruginosa* strains NIES-98 and PCC 7806 is a putative halogenase gene, *aerJ* (3), with high similarity to the putative halogenase gene in the anabaenopeptilide and cyanopeptolin synthetases of *Anabaena* sp. strain 90 (33) and *Microcystis* sp. strain NIVA CYA 172-5 (45), respectively, that is absent in *M. aeruginosa* NIES-843 and *Planktothrix agardhii* NIVA CYA 126-8. A sporadic distribution of halogenase genes in the *aer* and *mcn* clusters in *Microcystis* strains was discussed in detail elsewhere previously (3).

Downstream of *aerB*, a gene is present in *Microcystis* strains (*aerK*) that has no homologue in the *aer* gene cluster of *Planktothrix* (see Table S2 in the supplemental material). A chloramphenicol-resistant Δ*aerK* mutant of *M. aeruginosa* PCC 7806 was tested for the expected insertion of the resistance cassette into *aerK* and PCR amplicons were found to have the correct size (Fig. 2). In HPLC profiles (and liquid chromatography/mass spectrometry traces) of cell extracts, several early-eluting peaks were missing in the Δ*aerK* mutant compared to the wild-type strain that could be attributed to tautomers of aeruginosin 686 (Fig. 2). Thus, *aerK* is apparently essential for the synthesis of aeruginosins in *M. aeruginosa* PCC 7806 but not in *Planktothrix agardhii* NIVA CYA 126-8, where in the corresponding region a gene of unknown function is found together with the oxygenase gene *aerC*. An inactivation of *aer* genes of *M. aeruginosa* NIES-98 with different constructs could not be achieved despite repeated attempts.

The genes *aerDEF* were found in the *Microcystis* strains as well as in *Planktothrix agardhii* NIVA-CYA 126-8 and have been shown to be involved in the formation of the Choi moiety (16).

Downstream of these genes, open reading frames (ORFs) that are absent in *Planktothrix agardhii* NIVA-CYA 126-8 were found in *Microcystis* strains. In *M. aeruginosa* NIES-98 and NIES-843, *aerL* codes for a protein of 314 amino acids for which a conserved domain (CD) search revealed a sulfotransferase domain (heparan sulfate 2-*O*-sulfotransferase) (pfam05040) with a rather high E

FIG. 3. Proposed biosynthesis pathways of aeruginosins in *Microcystis aeruginosa* strains PCC 7806, NIES-98, and NIES-843 and *Planktothrix agardhii* NIVA CYA 126-8 (the latter was constructed based on data from reference 16). NRPS domains are indicated as described in the legend of Fig. 1 for *aer* gene clusters. Full arrows indicate NRPS-related steps in biosynthesis, while dotted arrows indicate a proposed reaction catalyzed by non-NRPS genes. The particular aeruginosin variants are denoted by roman numerals. I, aeruginoside 126 A; II, aeruginoside 126 B; III, aeruginosin 98-A; IV, aeruginosin 98-B; V, aeruginosin 101; VI, aeruginosin 686 A; VII, aeruginosin 686 B; VIII, aeruginosin 102. The unknown role of two hypothetical proteins in aeruginosin synthesis (Orf1 and Orf2) is indicated by a question mark.
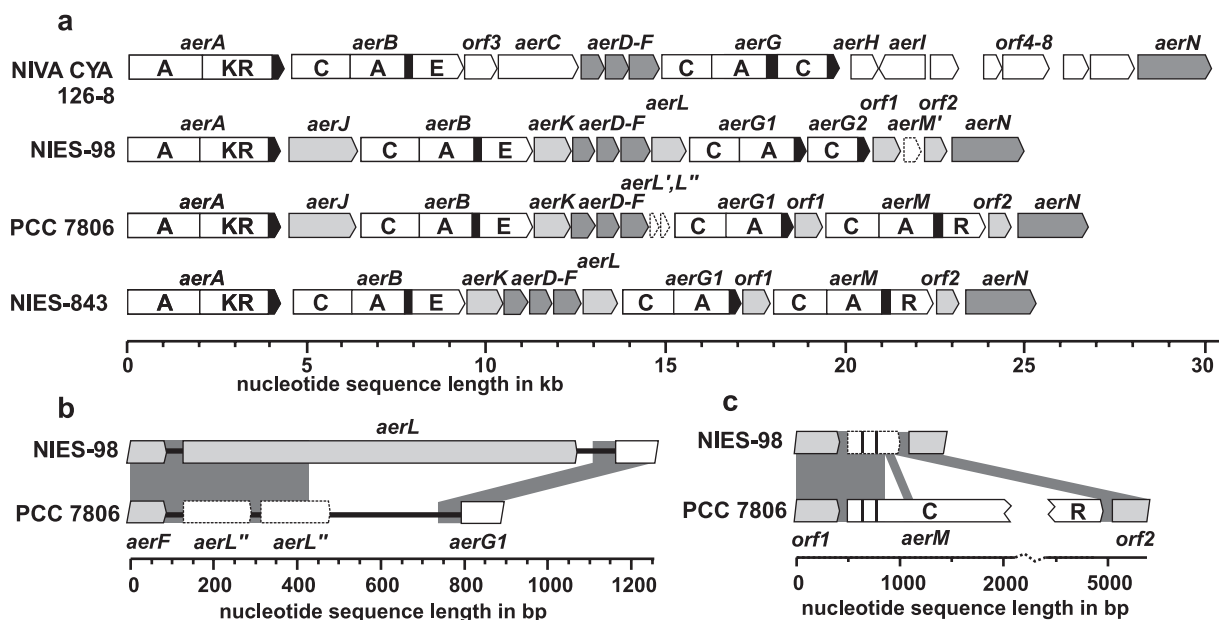
FIG. 4. Aeruginosin biosynthesis gene clusters from *Planktothrix agardhii* NIVA CYA 126-8 and three *Microcystis aeruginosa* strains, NIES-98, PCC 7806, and NIES-843. (a) NRPS-coding gene regions are labeled according to their function as revealed by the recognition of domain-specific core motifs (22): adenylation (A), condensation (C), epimerization (E), reductase (R), and ketoreductase (KR) domains. Thiolation domains (PCPs) are indicated in black. The genes are named *aerABCDEFGHI* according to data reported previously (16) and extended for new genes in *Microcystis* (*aerJKLMN*). Non-NRPS genes present in all four clusters are marked in dark gray, while those present only in *Microcystis* are marked in light gray. (b and c) Two regions of the *aer* gene clusters of NIES-98 and PCC 7806 are enlarged, with regions showing high nucleotide sequence similarity (>90% identity) indicated by a gray background. In *aerM*, the first two core motifs (C1 and C2) are indicted by black bands.

value of 0.95. Sequence similarity, however, was observed only for the 3′ part of the gene between the two strains, while the 5′ is considerably divergent, indicating a differing function in aeruginosin biosynthesis in both strains. In *M. aeruginosa* PCC 7806, two small pseudogenes, *aerL′* and *aerL″*, that are translated to 52 and 65 amino acid peptides, respectively, each with high similarity to the N-terminal part of AerL, were found in this region (Fig. 4).

While *aerG1* is present in all three *Microcystis* aeruginosin synthetase clusters, *aerG2*, which displays high similarity to the 3′ end of *aerG* of *Planktothrix agardhii* NIVA-CYA 126-8, is unique in *M. aeruginosa* NIES-98 and completely absent in the *M. aeruginosa* PCC 7806 and NIES-843 *aer* loci. Downstream of *aerG2* or *aerG1*, an ORF (*orf1*) is found in *Microcystis* strains coding for a 258-residue protein with similarity to oxidoreductases (CD, short chain dehydrogenase, PRK06172; E value of 8e−48). In *M. aeruginosa* PCC 7806 and NIES-843, *orf1* is followed downstream by a gene coding for a complete NRPS module, *aerM*, consisting of condensation, adenylation, thiolation, and reduction domains. In *M. aeruginosa* NIES-98, a terminal thioesterase or reduction domain is absent as in *Planktothrix agardhii* NIVA CYA 126-8, but fragments of *aerM* that form a pseudogene are present (*aerM′*) (Fig. 4).

In the three *Microcystis* strains, an ORF (*orf2*) coding for a 202-residue protein with similarity to (putative) acetyltransferases of various organisms is found (CD, LbetaH, cd00208; E value of 0.001) downstream of *aerM* or *aerM′*, respectively. A corresponding gene is not present in *Planktothrix agardhii* NIVA-CYA 126-8, where downstream of *aerG*, two genes, *aerH* and *aerI*, that code for a dioxygenase and a glycosyltransferase, respectively, are present.

In all three *Microcystis* strains, a predicted protein of 664

amino acids is encoded at the 3′ end of the aeruginosin clusters, *aerN*. The presence of the Walker A and B and Q-loop motifs clearly suggests that these proteins are ABC transporters. Orthologous genes are present in other cyanobacterial NRPS/polyketide synthase clusters and were found to be essential for the synthesis of microcystins, for example (30).

Plasticity in the *aer* gene clusters and the putative aeruginosin biosynthesis pathways. The aeruginosin synthetase gene clusters in all four strains have large parts in common, in particular the NRPS coding genes, suggesting a largely similar biosynthetic pathway. Nonetheless, the peptide products differ at several moieties, as do particular genes in the clusters. Below, we describe biosynthesis pathways leading to particular aeruginosin congeners that are reconstructed from the *aer* gene cluster sequences of the individual strains and summarized in Fig. 3.

Aeruginosin biosynthesis starts presumably with the activation of an aryl acid, phenylpyruvate in *Planktothrix agardhii* NIVA CYA 126-8, the keto group of which is reduced by the ketoreductase domain of the first module (16). However, in *Microcystis* strains, a similar pathway can be assumed with initiation of the biosynthesis with hydroxy-phenylpyruvate that is reduced to Hpla. Slight differences of binding pocket signatures (26) in the corresponding domains in the translated proteins (AerA) of the *Microcystis* and *Planktothrix* aeruginosin synthases may account for this observation (Table 1). In a phylogenetic tree of cyanobacterial adenylation domain sequences (see Fig. S2 in the supplemental material), a sister clade of the AerA adenylation domains comprises the initiating adenylation domains of microcystin synthetases (McyG) that have been shown to preferentially activate and load as-

TABLE 1. Binding pocket signatures ("nonribosomal code") conferring specificity to adenylation domains encoded in the *aer* gene clusters of
Microcystis aeruginosa PCC 7806, NIES-98, and NIES-843 and *Planktothrix agardhii* NIVA CYA 126-8[a]

Domain	Strain	Binding pocket signature	Best match	Amino acid
AerA, A1[b]	NIES-98	I G D W L A A S G K	NM	Aryl acid
	PCC 7806	I G D W L A A S G K	NM	Aryl acid
	NIES-843	I G D W L A A S G K	NM	Aryl acid
	CYA 126-8	I G E W L A A S G K	NM	Aryl acid
AerB, A2	NIES-98	D A F F L G V T F K	ApdD, AD1 Ile	Ile
	PCC 7806	D A S T I A A V C K	NosD, AD1 Tyr	Tyr
	NIES-843	D A S T I A A V C K	NosD, AD1 Tyr	Tyr
	CYA126-8	D A W F L G N V V K	SrfAB, AD3 Leu	Leu
AerG, A3	NIES-98	D V H I I A Y I V K	NM	Choi
	PCC 7806	D V H I I A Y I V K	NM	Choi
	NIES-843	D V H I I A F I V K	NM	Choi
	CYA 126-8	D V H I I A F L V K	NM	Choi
AerM, A4	PCC 7806	D V E N A G V V T K	NM	Arg?
	NIES-843	D V E N I G A I T K	NM	Arg?

[a] Signature sequences were extracted from protein sequences according to data described previously by Rausch et al. (32). NM, no match, i.e., less than 70% similarity to signature sequences of known specificity.
[b] Signature sequences for aryl acid-activating adenylation domains according to data described previously by May et al. (26). Binding pocket residues were extracted manually after alignment of AerA protein sequences with DhbE, PchD, EntE, and YbtE. AerA adenylation domains did not produce an alignment exploitable by the NRPS predictor.

sorted phenylpropanoids onto the peptidyl carrier protein (PCP) carrier domain (15).

The putative halogenase gene (*aerJ*) is nearly identical in two *Microcystis* strains (96% identity), although the peptide products differ. In *M. aeruginosa* NIES-98, the Hpla moiety is non-, mono-, and dichlorinated, while in *M. aeruginosa* PCC 7806, only monochlorinated Hpla was found. Aeruginosins with non-, mono-, and dichlorinated Hpla have frequently been found to be coproduced in other *Microcystis* strains (49). For the functioning of the downstream peptide assembly line, the chlorination of Hpla, however, seems to be not essential, at least not in particular clones, thus giving rise to three structural variants as in *Microcystis aeruginosa* NIES-98 (variants III, IV, and V) (Fig. 3).

The first 15 kbp (*aerABCDEF*) are nearly identical in the three *Microcystis* strains, including the intergenic regions. The only exception is the region coding for the adenylation domain in *aerB* (Fig. 5 and Table 1). The condensation, thiolation, and epimerization domains in *aerB* are nearly identical (average $\pi = 0.03$), while the region coding for the adenylation domain deviates markedly (average $\pi = 0.38$). The region coding for the binding pocket (downstream of core motif A4) is especially divergent ($\pi > 0.6$ in windows of 100 nucleotides [nt], comparing *M. aeruginosa* PCC 7806 and NIES-98). The transition from the nearly identical condensation domain to the dissimilar adenylation domain is very sharp and can be narrowed to a region of some few nucleotides. More-conserved parts can be found around the core motifs. In the 3′ direction, high values of identities are reached again only shortly upstream of the region coding for the thiolation domain core motif. The predicted specificities of the adenylation domain binding pockets in AerB of all strains match with the amino acids found at the corresponding positions in the peptide products (Table 1). All other NRPS genes are very similar in *Microcystis* strains ($\pi < 0.05$ for entire genes). The sequence divergence of adenylation domain coding regions of *aerB* is less pronounced when *Mi-crocystis* sequences are compared to corresponding sequences in *Planktothrix agardhii* NIVA CYA 126-8 but is nonetheless evident (Fig. 5).

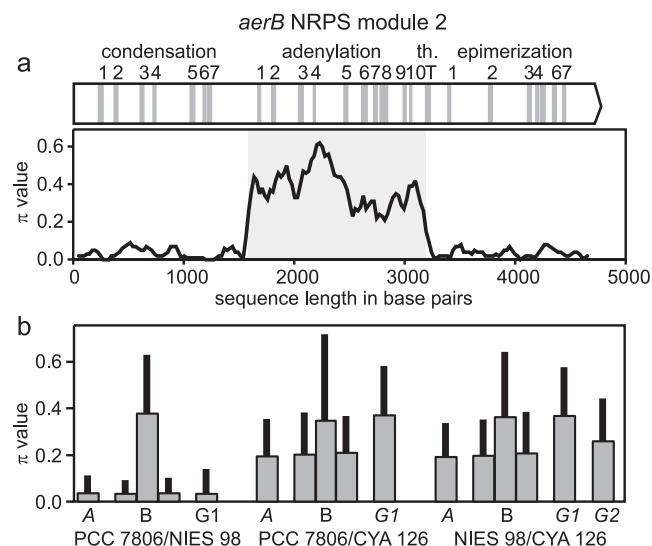The genes coding for AerK, AerD, AerE, and AerF are

FIG. 5. Nucleotide sequence polymorphism of the genes coding for the second NRPS module in the *aer* gene clusters of *Microcystis aeruginosa* NIES-98 and PCC 7806 and *Planktothrix agardhii* NIVA CYA 126-8. (a) Schematic illustration of *aerB* (module 2), with domain-specific core motifs indicated by bands and numbers (C1 to C7, A1 to A10, T, and E1 to E7) and π values (average number of nucleotide differences per site between two sequences) (29) for pairwise comparisons of aligned sequences of NIES-98 and PCC 7806 computed in the sliding window mode in DnaSP (width, 100 nt; step, 25 nt). (b) π values for pairwise comparisons of aligned sequences of NRPS-coding *aer* genes of all three strains. Columns, average π; lines, maximum π for 100-nt windows. Note that *aerB* has been split into three parts as indicated above (a).

likely involved in the formation of Choi, as has been shown by analyses of knockout mutants in *M. aeruginosa* PCC 7806 (Δ*aerK*) and *Planktothrix agardhii* NIVA CYA 126-8 (Δ*aerD* and Δ*aerEF*). A precursor of Choi has not yet been defined by experimental work, but for *Planktothrix agardhii* NIVA CYA 126-8, arogenate has been suggested (16). The pathway may, however, differ in both species as indicated by the exclusive presence of *aerK* in *Microcystis*.

The (weak) similarity of AerL to sulfotransferases together with a sulfatation of the aeruginosins produced by *M. aeruginosa* NIES-98 and NIES-843, and the absence of sulfated aeruginosins in *M. aeruginosa* PCC 7806 suggests that AerL is indeed a sulfotransferase. The location of the sulfatation, however, is presumably different with Choi sulfatation in NIES-98 and Hpla sulfatation in NIES-843.

The similarity of the gene coding for the third NRPS module (*aerG1*) and the identical binding pocket signature sequences of the third adenylation domain suggests that in all three aeruginosin synthetases, Choi (and not a precursor) is selected and activated (Table 1). Thus, in *M. aeruginosa* NIES-98, Choi sulfatation is likely accomplished after the formation of the Choi-phosphopantetheine thioester.

The incorporation of the last moiety into the aeruginosin molecule in *M. aeruginosa* NIES-98 is arguably similar to that in *Planktothrix agardhii* NIVA CYA 126-8. AerG2 (and the C-terminal part of AerG in *Planktothrix agardhii* NIVA CYA 126-8) would thus bind an agmatine to the PCP-bound carboxyl group of Choi and, in the same step, release the completed aeruginosin in a manner similar to that proposed for the release of aeruginosides in *Planktothrix* (16). Whether the last condensation domain is actually involved in amide bond formation in analogy to amide bond formation by a free-standing condensation domain (VibH) during vibriobactin biosynthesis (20) remains to be studied.

In contrast, the C-terminal moiety (argininal) is incorporated into the aeruginosin molecule by a full NRPS module in *M. aeruginosa* PCC 7806 and NIES-843. The binding pocket signature sequences revealed no significant match with the "NRPS predictor" (32), but a high level of similarity was found for signature sequences of arginine, lysine, and ornithine activating domains by applying the neighbor-joining algorithm to the 10 amino acid signature sequences (data not shown). The precursor of argininal should possess a carboxy group that is thioesterized after activation by the adenylation domain and that is modified by a reduction step to yield argininal. The simplest pathway would involve the incorporation of arginine that is then modified by the reductase domain, thus releasing the aeruginosin from the NRPS complex in analogy to the reductive release of myxochelin from the PCP domain of MxcC (13). A similar mechanism has been proposed for another cyanobacterial peptide, nostocyclopeptide, in *Nostoc* sp. strain ATCC 53789 (1). Interestingly, in an aeruginosin-like peptide from *Nodularia spumigena*, spumigin (12), an intact C-terminal arginine moiety has been identified.

In *Microcystis* strains, the *aer* gene clusters contain two ORFs of unknown function, *orf1* and *orf2*. Their location next to *aerM* suggests that they may be involved in the modification of the last moiety. If this was the case, they are probably without function in *M. aeruginosa* NIES-98, where the last moiety is likely added to the peptide molecule by a mechanism

similar to that in *Planktothrix agardhii* NIVA CYA 126-8, the *aer* gene cluster of which does not contain homologues of *orf1* and *orf2*.

The glycosyltransferase gene present in the *aer* gene cluster in *Planktothrix agardhii* NIVA-CYA 126-8 is absent from the *aer* gene clusters in *Microcystis* strains, and accordingly, the aeruginosins are not glycosylated. However, the glycosylation of aeruginosins has been observed exclusively and frequently in *Planktothrix* strains to date (16, 39, 48), indicating a possible genus-specific association of this gene with the *aer* gene cluster.

Evolution of aeruginosin synthetase gene clusters. The presence of *aer* biosynthesis gene clusters in taxa as distant as *Microcystis* and *Planktothrix* indicates that they could originate from a common ancestor, thus representing ancient genes, as has been suggested for the microcystin (*mcy*) and the cyanopeptolin synthetase (*mcn*, *apd*, and *oci*) gene clusters (31, 34). Although the horizontal transfers of complete gene clusters cannot be completely ruled out, there is no evidence that this mechanism could explain the sporadic distribution of homologue NRPS operons among genera, species, and strains of cyanobacteria. The degree of DNA polymorphism in homologous *aer* genes in the four clusters also indicates an independent evolution of an ancestral gene in lineages leading to modern genera, with *Microcystis* sequences being more similar to each other than those of *Planktothrix*. Since aeruginosins are not produced by all isolates of *Microcystis* and *Planktothrix*, their distribution in these genera may be caused by repeated gene loss (31) or inactivation (4), as has been shown for the microcystin synthetase gene cluster.

On the other hand, the differences between the *aer* gene sequences cannot be explained solely by gradual mutations. Recombination events within cyanobacterial NRPS clusters have been demonstrated, for example, for the *mcy* cluster (27, 42). In trees inferred with maximum-parsimony and minimum-evolution models, the variable adenylation domains from AerB cluster with cyanobacterial adenylation domains of the same substrate specificities (see Fig. S2 in the supplemental material). The region coding for this adenylation domain is apparently subject to recombination, although it cannot be deduced which sequence should be considered as ancestral with the data at hand.

As for the region from *aerF* to *aerG1*, the presence of a functional sulfotransferase gene appears to be ancestral in *M. aeruginosa* NIES-98 and NIES-843 compared to *M. aeruginosa* PCC 7806, where only remnant pseudogenes (*aerL′* and *aerL″*) are found. In contrast to this, the presence of only fragments of *aerM*, coding for a full NRPS module, in *M. aeruginosa* NIES-98 (*aerM′*) portends that in *M. aeruginosa* PCC 7806 and NIES-843, the ancestral sequence is contained. The reconstruction of the evolution of the *aer* gene cluster is further complicated by the observed similarity of the region coding for the fourth (incomplete) NRPS modules, *aerG2* in *M. aeruginosa* NIES-98 and 3′ *aerG* in *Planktothrix agardhii* NIVA CYA 126-8, suggesting that *aerG2* in *M. aeruginosa* NIES-98 has been acquired by horizontal gene transfer, possibly from the *Planktothrix* lineage.

While "natural" combinatorial biosynthesis is apparently not unusual, it is still difficult to be achieved in vitro, where peptide production is generally strongly impaired, e.g., by domain exchange (9). Although the *aer* gene clusters are arguably ancient

and have likely evolved independently in major cyanobacterial lineages, there is evidence for horizontal gene transfer of partial sequences between these lineages. Relative sequence dissimilarities and the presence or absence of particular genes among the four *aer* gene clusters suggest that recombination (adenylation domain coding region of *aerB*), horizontal gene transfer (*aerG2*), and gene decay (*aerL* and *aerM*) have been important for the evolution of individual biosynthesis gene clusters.

The metabolic profiling of three closely related strains of *Microcystis aeruginosa* in combination with the analysis of the requisite *aer* biosynthesis gene clusters provides new insights into the biosynthetic assembly lines and their evolution. The diversity of aeruginosin structures can be deduced from the organization and gene content of the *aer* gene clusters in cyanobacterial strains. While all strains share a tetrapeptide backbone, there is a high degree of flexibility in the building blocks and the amino acid-tailoring reactions, including oxidoreductions, glycosylation, halogenation, and sulfatation. These structural peculiarities are well reflected in the nucleotide sequences of the *aer* biosynthesis gene clusters. Differences between the individual cyanobacterial *aer* biosynthesis gene clusters suggest that these are susceptible to sequence modifications by various mechanisms such as horizontal gene transfer, recombination, and gene decay. Apparently, the selective pressure does not effectuate a streamlining of the biosynthetic gene clusters but, on the contrary, promotes the diversity of genes and peptide products.

The analysis of related pathways implicated in the biosynthesis of structurally distinct aeruginosin variants could provide clues to the activity of cryptic genes that are present or absent in individual aeruginosin synthetase gene clusters. Sequence comparisons of homologue biosynthesis genes together with the thorough analysis of secondary metabolites enable the elucidation of gene and protein functions. These insights can then be used for genome-mining approaches to identify organisms with promising new natural products.

REFERENCES

1. **Becker, J. E., R. E. Moore, and B. S. Moore.** 2004. Cloning, sequencing, and biochemical characterization of the nostocyclopeptide biosynthetic gene cluster: molecular basis for imine macrocyclization. Gene **325**:35–42.
2. **Bister, B., S. Keller, H. I. Baumann, G. Nicholson, S. Weist, G. Jung, R. D. Süssmuth, and F. Jüttner.** 2004. Cyanopeptolin 963A, a chymotrypsin inhibitor of *Microcystis* PCC 7806. J. Nat. Prod. **67**:1755–1757.
3. **Cadel-Six, S., C. Dauga, A. M. Castets, R. Rippka, N. Tandeau de Marsac, and M. Welker.** 2008. Halogenase genes in two non-ribosomal peptide synthetase gene clusters of *Microcystis* (Cyanobacteria): sporadic distribution and evolution. Mol. Biol. Evol. **25**:2031–2041.
4. **Christiansen, G., R. Kurmayer, Q. Liu, and T. Börner.** 2006. Transposons inactivate biosynthesis of the nonribosomal peptide microcystin in naturally occurring *Planktothrix* spp. Appl. Environ. Microbiol. **72**:117–123.
5. **Czarnecki, O., I. Lippert, M. Henning, and M. Welker.** 2006. Identification of peptide metabolites of *Microcystis* (Cyanobacteria) that inhibit trypsin-like activity in planktonic herbivorous *Daphnia* (Cladocera). Environ. Microbiol. **8**:77–87.
6. **Dittmann, E., B. A. Neilan, M. Erhard, H. von Döhren, and T. Börner.** 1997. Insertional mutagenesis of a peptide synthetase gene that is responsible for hepatotoxin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. Mol. Microbiol. **26**:779–787.
7. **Ersmark, K., J. R. Del Valle, and S. Hanessian.** 2008. Chemistry and biology of the aeruginosin family of serine protease inhibitors. Angew. Chem. Int. Ed. Engl. **47**:1202–1223.
8. **Finking, R., and M. A. Marahiel.** 2004. Biosynthesis of nonribosomal peptides. Annu. Rev. Microbiol. **58**:453–488.
9. **Fischbach, M. A., J. R. Lai, E. D. Roche, C. T. Walsh, and D. R. Liu.** 2007. Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes. Proc. Natl. Acad. Sci. USA **104**:11951–11956.
10. **Franche, C., and T. Damerval.** 1988. Tests on *nif* probes and DNA hybridizations. Methods Enzymol. **167**:803–808.
11. **Frangeul, L., P. Quillardet, A. M. Castets, J. F. Humbert, H. C. P. Matthijs, D. Cortez, A. Tolonen, C. C. Zhang, S. Gribaldo, J. C. Kehr, Y. Zilliges, N. Ziemert, S. Becker, E. Talla, A. Latifi, A. Billault, A. Lepelletier, E. Dittmann, C. Bouchier, and N. Tandeau de Marsac.** 2008. Highly plastic genome of *Microcystis aeruginosa* PCC7806, an ubiquitous toxic freshwater cyanobacterium. BMC Genomics **9**:274.
12. **Fujii, K., K. Sivonen, K. Adachi, K. Noguchi, H. Sano, K. Hirayama, M. Suzuki, and K.-I. Harada.** 1997. Comparative study of toxic and non-toxic cyanobacterial products: novel peptides from toxic *Nodularia spumigena* AV1. Tetrahedron Lett. **38**:5525–5528.
13. **Gaitatzis, N., B. Kunze, and R. Müller.** 2001. In vitro reconstitution of the myxochelin biosynthetic machinery of *Stigmatella aurantiaca* Sg a15: biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases. Proc. Natl. Acad. Sci. USA **98**:11136–11141.
14. **Hanessian, S., J. F. Tremblay, and J. F. W. Petersen.** 2004. The N-acyloiminium ion Aza-Prins route to octahydroindoles: total synthesis and structural confirmation of the antithrombotic marine natural product oscillarin. J. Am. Chem. Soc. **126**:6064–6071.
15. **Hicks, L. M., M. C. Moffitt, L. L. Beer, B. S. Moore, and N. L. Kelleher.** 2006. Structural characterization of in vitro and in vivo intermediates on the loading module of microcystin synthetase. ACS Chem. Biol. **1**:93–102.
16. **Ishida, K., G. Christiansen, W. Y. Yoshida, M. Welker, J. Bonjoch, C. Hertweck, T. Börner, T. K. Hemscheidt, and E. Dittmann.** 2007. Biosynthetic pathway and structure analysis of aeruginoside 126A and B, cyanobacterial peptides bearing an unusual 2-carboxy-6-hydroxyoctaindole moiety. Chem. Biol. **14**:565–576.
17. **Ishida, K., H. Matsuda, Y. Okita, and M. Murakami.** 2002. Aeruginoguanidines 98-A-98C: cytotoxic unusual peptides from the cyanobacterium *Microcystis aeruginosa*. Tetrahedron **58**:7645–7652.
18. **Ishida, K., Y. Okita, H. Matsuda, T. Okino, and M. Murakami.** 1999. Aeruginosins, protease inhibitors from the cyanobacterium *Microcystis aeruginosa*. Tetrahedron **55**:10971–10988.
19. **Kaneko, T., N. Nakajima, S. Okamoto, I. Suzuki, Y. Tanabe, M. Tamaoki, Y. Nakamura, F. Kasai, A. Watanabe, K. Kawashima, Y. Kishida, A. Ono, Y. Shimizu, C. Takahashi, C. Minami, T. Fujishiro, M. Kohara, M. Katoh, N. Nakazaki, S. Nakayama, S. Yamada, S. Tabata, and M. M. Watanabe.** 2007. Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. DNA Res. **14**:247–256.
20. **Keating, T. A., C. G. Marshall, C. T. Walsh, and A. E. Keating.** 2002. The structure of VibH represents nonribosomal peptide synthetase condensation and epimerization domains. Nat. Struct. Biol. **9**:522–526.
21. **Kodani, S., K. Ishida, and M. Murakami.** 1998. Aeruginosin 103-A, a thrombin inhibitor from the cyanobacterium *Microcystis viridis*. J. Nat. Prod. **61**:1046–1048.
22. **Konz, D., and M. A. Marahiel.** 1999. How do peptide synthetases generate structural diversity? Chem. Biol. **6**:R39–R48.
23. **Marfey, P.** 1984. Determination of D-amino acids. II. Use of a bifunctional reagent, 1,5-difluoro-2,4-dinitrobenzene. Carlsberg Res. Commun. **49**:591–596.
24. **Martin, C., L. Oberer, T. Ino, W. A. König, M. Busch, and J. Weckesser.** 1993. Cyanopeptolins, new depsipeptides from the cyanobacterium *Microcystis* sp. PCC 7806. J. Antibiot. **46**:1550–1556.
25. **Matsuda, H., T. Okino, M. Murakami, and K. Yamaguchi.** 1996. Aeruginosins 102-A and B, new thrombin inhibitors from the cyanobacterium *Microcystis viridis* (NIES-102). Tetrahedron **52**:14501–14506.
26. **May, J. J., N. Kessler, M. A. Marahiel, and M. T. Stubbs.** 2002. Crystal structure of DhbE, an archetype for aryl acid activating domains of modular nonribosomal peptide synthetases. Proc. Natl. Acad. Sci. USA **99**:12120–12125.
27. **Mikalsen, B., G. Boison, O. M. Skulberg, J. Fastner, W. Davies, T. M. Gabrielsen, K. Rudi, and K. S. Jakobsen.** 2003. Natural variation in the microcystin synthetase operon *mcyABC* and impact on microcystin production in *Microcystis* strains. J. Bacteriol. **185**:2774–2785.
28. **Murakami, M., K. Ishida, T. Okino, Y. Okita, H. Matsuda, and K. Yamaguchi.** 1995. Aeruginosins 98-A and B, trypsin inhibitors from the blue-green alga *Microcystis aeruginosa* (NIES-98). Tetrahedron Lett. **36**:2785–2788.
29. **Nei, M.** 1987. Molecular evolution genetics. Columbia University Press, New York, NY.
30. **Pearson, L. A., M. Hisbergues, T. Börner, E. Dittmann, and B. A. Neilan.**

2004. Inactivation of an ABC transporter gene, *mcyH*, results in loss of microcystin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. Appl. Environ. Microbiol. **70:**6370–6378.

31. **Rantala, A., D. Fewer, M. Hisbergues, L. Rouhiainen, J. Vaitomaa, T. Börner, and K. Sivonen.** 2004. Phylogenetic evidence for the early evolution of microcystin synthesis. Proc. Natl. Acad. Sci. USA **101:**568–573.

32. **Rausch, C., T. Weber, O. Kohlbacher, W. Wohlleben, and D. H. Huson.** 2005. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). Nucleic Acids Res. **33:**5799–5808.

33. **Rouhiainen, L., L. Paulin, S. Suomalainen, H. Hyytiainen, W. Buikema, R. Haselkorn, and K. Sivonen.** 2000. Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in *Anabaena* strain 90. Mol. Microbiol. **37:**156–167.

34. **Rounge, T. B., T. Rohrlack, A. Tooming-Klunderud, T. Kristensen, and K. S. Jakobsen.** 2007. Comparison of cyanopeptolin genes in *Planktothrix*, *Microcystis*, and *Anabaena* strains: evidence for independent evolution within each genus. Appl. Environ. Microbiol. **73:**7322–7330.

35. **Rozas, J., J. C. Sánchez-DelBarrio, X. Messeguer, and R. Rozas.** 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics **19:**2496–2497.

36. **Sánchez, C., L. Zhu, A. F. Brana, A. P. Salas, J. Rohr, C. Méndez, and J. A. Salas.** 2005. Combinatorial biosynthesis of antitumor indolocarbazole compounds. Proc. Natl. Acad. Sci. USA **102:**461–466.

37. **Sandler, B., M. Murakami, and J. Clardy.** 1998. Atomic structure of the trypsin-aeruginosin 98-B complex. J. Am. Chem. Soc. **120:**595–596.

38. **Schmidt, E. W., J. T. Nelson, D. A. Rasko, S. Sudek, J. A. Eisen, M. G. Haygood, and J. Ravel.** 2005. Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of Lissoclinum patella. Proc. Natl. Acad. Sci. USA **102:**7315–7320.

39. **Shin, H. J., H. Matsuda, M. Murakami, and K. Yamaguchi.** 1997. Aeruginosins 205A and -B, serine protease inhibitory glycopeptides from the cyanobacterium *Oscillatoria agadhii* (NIES-205). J. Org. Chem. **62:**1810–1813.

40. **Stachelhaus, T., H. D. Mootz, and M. A. Marahiel.** 1999. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem. Biol. **6:**493–505.

41. **Tamura, K., J. Dudley, M. Nei, and S. Kumar.** 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. **24:**1596–1599.

42. **Tanabe, Y., K. Kaya, and M. M. Watanabe.** 2004. Evidence for recombination in the microcystin synthetase (*mcy*) genes of toxic cyanobacteria *Microcystis* spp. J. Mol. Evol. **58:**633–641.

43. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The Clustal_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25:**4876–4882.

44. **Tillett, D., E. Dittmann, M. Erhard, H. von Döhren, T. Börner, and B. A. Neilan.** 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-poliketide synthetase system. Chem. Biol. **7:**753–764.

45. **Tooming-Klunderud, A., T. Rohrlack, K. Shalchian-Tabrizi, T. Kristensen, and K. S. Jakobsen.** 2007. Structural analysis of a non-ribosomal halogenated cyclic peptide and its putative operon from *Microcystis*: implications for evolution of cyanopeptolins. Microbiology **153:**1382–1393.

46. **Valls, N., M. Lopez-Canet, M. Vallribera, and J. Bonjoch.** 2001. First total syntheses of aeruginosin 298-A and aeruginosin 298-B, based on a stereocontrolled route to the new amino acid 6-hydroxyoctahydroindole-2-carboxylic. Chem. Eur. J. **7:**3446–3460.

47. **Welker, M., M. Brunke, K. Preussel, I. Lippert, and H. von Döhren.** 2004. Diversity and distribution of *Microcystis* (Cyanobacteria) oligopeptide chemotypes from natural communities studied by single colony mass spectrometry. Microbiology **150:**1785–1796.

48. **Welker, M., G. Christiansen, and H. von Döhren.** 2004. Diversity of coexisting *Planktothrix* (Cyanobacteria) chemotypes deduced by mass spectral analysis of microcystins and other oligopeptides. Arch. Microbiol. **182:**288–298.

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

50. **Welker, M., and H. von Döhren.** 2006. Cyanobacterial peptides—nature's own combinatorial biosynthesis. FEMS Microbiol. Rev. **30:**530–563.

51. **Ziemert, N., K. Ishida, A. Liaimer, C. Hertweck, and E. Dittmann.** 2008. Ribosomal synthesis of tricyclic depsipeptides in bloom-forming cyanobacteria. Angew. Chem. Int. Ed. Engl. **47:**7756–7759.

52. **Ziemert, N., K. Ishida, P. Quillardet, C. Bouchier, C. Hertweck, N. Tandeau de Marsac, and E. Dittmann.** 2008. Microcyclamide biosynthesis in two strains of *Microcystis aeruginosa*: from structure to genes and vice versa. Appl. Environ. Microbiol. **74:**1791–1797.