UC San Diego
UC San Diego Previously Published Works

Title
Heterologous Expression in Anabaena of the Columbamide Pathway from the Cyanobacterium Moorena bouillonii and Production of New Analogs

Permalink
https://escholarship.org/uc/item/0fb7k7cm

Journal
ACS Chemical Biology, 17(7)

ISSN
1554-8929

Authors
Taton, Arnaud
Rohrer, Sebastian
Diaz, Brienna
et al.

Publication Date
2022-07-15

DOI
10.1021/acschembio.2c00347

Supplemental Material
https://escholarship.org/uc/item/0fb7k7cm#supplemental

Copyright Information
This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at
https://creativecommons.org/licenses/by-nc-nd/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Heterologous expression in *Anabaena* of the columbamide pathway from the cyanobacterium *Moorena bouillonii* and production of new analogs

Arnaud Taton,[†,⊥] Sebastian Rohrer,[‡,⊥] Brienna Diaz,[†] Raphael Reher,[‡,#,¶] Andres Mauricio Caraballo Rodriguez,[§] Marsha L. Pierce,[¢] Pieter C. Dorrestein,[§] Lena Gerwick,[‡] William H. Gerwick,[‡,§,*] James W. Golden[†,*]

[†]School of Biological Sciences, [‡]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, and [§]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California 92093, United States, [#]Institute of Pharmacy, Martin-Luther-University Halle-Wittenberg, Halle (Saale) 06114, Germany, and [¶]Institute of Pharmaceutical Biology and Biotechnology, Philipps University of Marburg, Marburg 35037, Germany. [¢]Department of Pharmacology, Midwestern University, Downers Grove, IL 60515

[⊥]AT and SR contributed equally.

*Corresponding authors: wgerwick@ucsd.edu and jwgolden@ucsd.edu

ABSTRACT

Columbamides are chlorinated acyl amide natural products, several of which exhibit cannabinomimetic activity. These compounds were originally discovered from a culture of the filamentous marine cyanobacterium *Moorena bouillonii* PNG5-198 collected from the coastal waters of Papua New Guinea. The columbamide biosynthetic gene cluster had been identified using bioinformatics, but not confirmed by experimental evidence. Here we report the heterologous expression in *Anabaena* (*Nostoc*) PCC 7120 of the 28.5-kb biosynthetic gene cluster that encodes for columbamide biosynthesis. The production of columbamides in *Anabaena* was investigated under several different culture conditions and several new columbamide analogs were identified by liquid chromatography – tandem mass spectrometry (LC-MS/MS) and nuclear magnetic resonance (NMR). In addition to previously characterized columbamides A, B, and C, new columbamides I-M were produced in these experiments, and the structure of the most abundant monochlorinated analog, columbamide K (**11**), was fully characterized. The other new columbamide analogs were produced in only small quantities, and structures are proposed based on high resolution (HR)-MS, MS/MS, and ^{1}H NMR data. Overexpression of the pathway's predicted halogenases resulted in increased productions of di- and tri-chlorinated compounds. The most significant change in production of columbamides in *Anabaena* was correlated to the concentration of NaCl in the medium.

KEYWORDS: Cyanobacteria, natural products, heterologous expression, columbamides

INTRODUCTION

The columbamides comprise a family of chlorinated acyl amide natural products (NPs), and include analogs that exhibit strong binding to the cannabinoid receptors CB_1 and CB_2.[1-3] Columbamide analogs A, B, and C (**1**-**3**) and the biosynthetic pathway were originally discovered by combining genome mining and molecular networking approaches from a culture of the filamentous marine cyanobacterium *Moorena* (formerly *Moorea*)[4] *bouillonii* PNG5-198 (hereafter *M. bouillonii* PNG) isolated from coral reefs in New Ireland, Papua New Guinea.[1] Subsequently, columbamides D-H (**4**-**8**) were described from samples of *M. bouillonii* collected in Malaysia.[2, 3]

Upon structure elucidation of columbamides A-C (**1-3**) and a comprehensive analysis of *M. bouillonii* PNG's biosynthetic gene clusters (BGCs), the pathway encoding for the production of the columbamides was identified but lacked experimental evidence.[1] Columbamide biosynthesis (Figure 1) is proposed to start with an acyl-CoA synthetase, ColA, loading dodecanoic (and/or tetradecanoic) acid onto the acyl carrier protein ColC, while no functional role was associated to ColB. A unique feature of the columbamides is the presence of chlorine atoms at the ω1 and ω7 positions of the acyl chain. The incorporation of these chlorine atoms is attributed to two halogenase homologs, ColD and ColE, and involves the highly unusual chlorination of unactivated C-H bonds.[1] Next, ColF carries out two PKS elongation steps. This is followed by the ColG non-ribosomal peptide synthetase (NRPS) extension of the acyl-acyl carrier protein (ACP) with a serine residue and *N*- and *O*-methylations by the ColG methyltransferase domains. Finally, the peptidyl carrier protein (PCP)-bound product is released by a four-electron reductase-releasing mechanism, converting the carboxylate of serine to a primary alcohol. The biosyntheses of the *O*-acetylated compounds **1**, **2**, **6**, **7**, **9**, **11**, and **13** (Figure 1) are believed to occur as a post-modular assembly step in which the predicted acyltransferase ColI acetylates the primary alcohol. The function of ColH is unknown.[1]

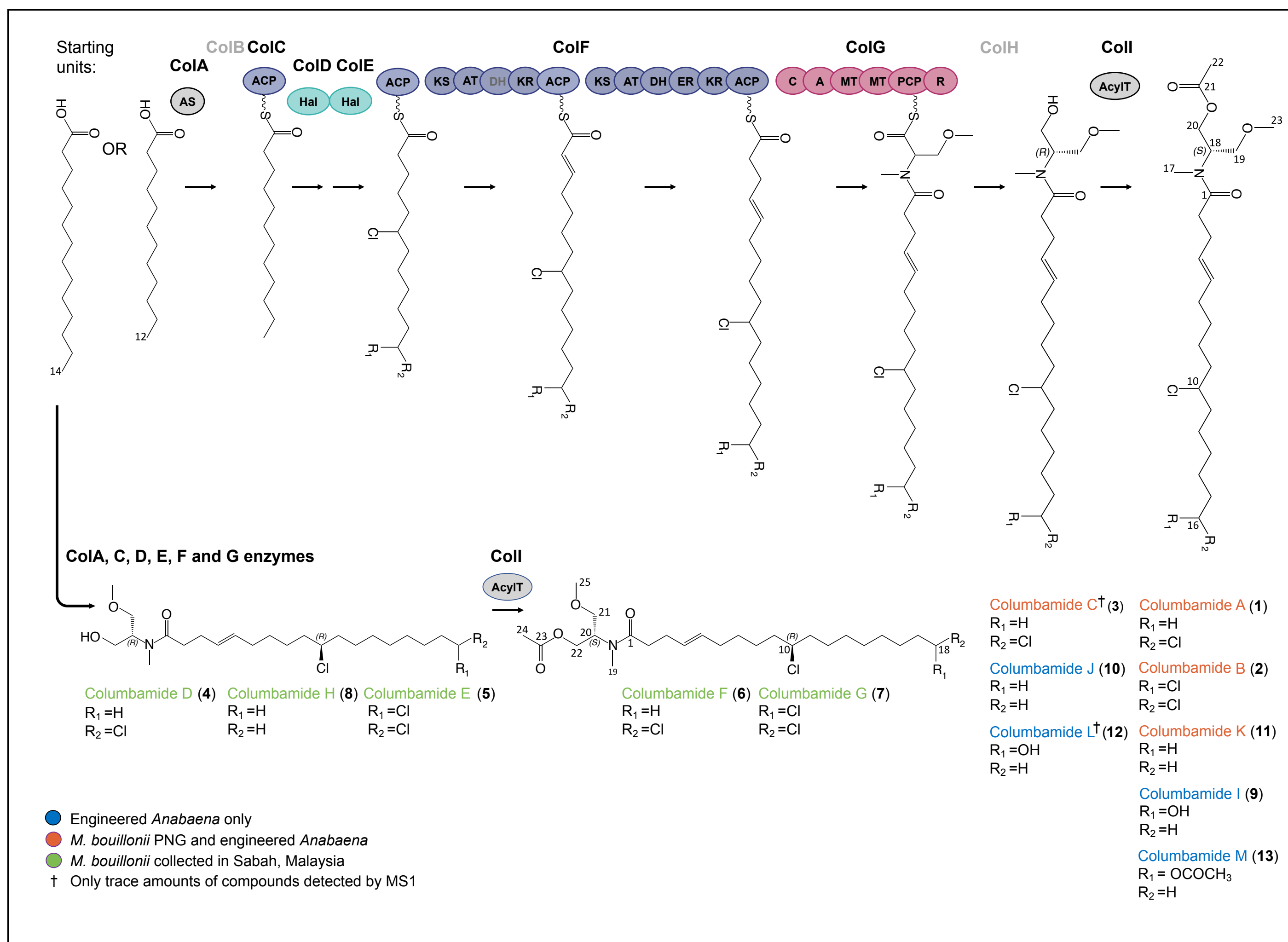


Figure 1. Biosynthetic pathway and structures of columbamide analogs produced natively in *Moorena* strains and by heterologous expression of the pathway in *Anabaena*. AS: acyl synthetase, ACP: acyl carrier protein, Hal: halogenase, KS: β-ketoacyl-ACP synthase; AT: acyl transferase, DH: β-hydroxy-acyl-ACP dehydratase, KR: β-ketoacyl-ACP reductase, ER: enoyl reductase, C: condensation domain, A: adenylation domain, MT: methyltransferase, PCP: peptidyl carrier protein, R: reductase, AcylT: acyltransferase. Structural variation in the columbamides is derived from two different starter units and activity of the halogenases ColD and ColE as well as acetyltransferase ColI. The function of ColB and ColH are unknown and depicted in gray. The first DH domain of ColF could not be identified using bioinformatics and is depicted in gray. It was proposed that the DH domain of the ColF second PKS module may provide this function.[1] All previously described analogs from *M. bouillonii* PNG[1] were detected in cultures of *Anabaena* carrying the columbamide BGC. In addition, five new columbamides (I-M) were detected in cultures of *Anabaena*. Analogs from *M. bouillonii* collected in Malaysia[2, 3] were not found in cultures of *Anabaena* transformed with the columbamide biosynthetic gene cluster. The configuration of the serinol portion was experimentally determined for columbamides A-H and for the secondary chlorine group for columbamides D-H.[2]

In addition to the pathway encoding for columbamide A–C (**1-3**), the pathways encoding for apratoxins A–C, lyngbyabellin A and 28 additional BGCs were identified in the genome of *M. bouillonii* PNG.[5] The genus *Moorena* is found in marine habitats of tropical regions around the globe and has been a very rich source of NPs.[6] Some strains carry over 40 different BGCs and close to 200 NPs have been identified from this genus alone.[6] Moreover, the NPs of *Moorena* possess highly diverse structures and many have potent biological

activities.[7] The presence of carbon-halogen bonds is another interesting characteristic of cyanobacterial NPs, leading to their further structural diversification. Of the 2025 reported cyanobacterial NPs (NPAtlas DB 2021_08), 21% are halogenated, and for those that are not peptidic, this number increases to 35%.[7] Besides having pleiotropic bioactivities, halogenated cyanobacterial NPs can also have a significant environmental impact. For example, the pentabrominated biindole alkaloid aetokthonotoxin (AETX) was identified as a cyanobacterial neurotoxin that acts as the causal agent of fatal vacuolar myelinopathy, and is responsible for the mass mortality of bald eagles and other animals in the southeastern US.[8] In general, cyanobacterial toxins have primarily been a concern in fresh water bodies, with examples such as the microcystins and saxitoxin.[9] However, marine cyanobacteria also produce environmentally important toxins, such as polybrominated diphenyl ethers that accumulate in the marine food chain and are potentially harmful to humans.[10]

The majority of cyanobacterial NPs are peptides or have peptidic components, and these are either ribosomally synthesized and post-translationally modified (RiPPs) or encoded by non-ribosomal peptide synthetases (NRPSs). Other cyanobacterial NPs are encoded by polyketide synthases (PKSs), and often by PKS/NRPS hybrid pathways with diverse tailoring steps.[11] Some of these NPs may be useful as drug leads in the treatment of cancer, neurological disorders, and infectious diseases,[12] while others have anti-inflammatory properties[13] or confer UV protection.[14, 15] However, the pharmaceutical development of marine cyanobacterial NPs has been restricted in part because natural resources are limited and most filamentous marine cyanobacteria grow very slowly in culture.[16, 17] Additionally, most cyanobacteria have so far not been amenable to genetic manipulations.

The development of approaches for the heterologous expression of NP pathways, including in vitro DNA assembly methods, yeast TAR cloning, and the low cost of DNA sequencing have led to significant advances for the characterization of biosynthetic pathways and the production of NPs from heterotrophic bacteria,[18, 19] and to a lesser extent photoautotrophic cyanobacteria.[20] Small cyanobacterial RiPPs and NRPS pathways have been successfully expressed in *E. coli*.[14, 21-24] However, the heterologous expression of large cyanobacterial NRPS, PKS, and PKS/NRPS hybrid biosynthetic pathways has proven more difficult in heterotrophic bacteria such as *E. coli* or *Streptomyces*.[25-27] Nonetheless, a handful of compounds were produced in *E. coli* upon replacement of the pathway's native promoters.[28-31] The expression of heterologous proteins and pathways in a phylogenetically distant organism can require a significant amount of optimization related to codon usage, precursor concentrations, cofactors, chaperones, and phosphopantetheinyl transferases (PPTases).[32-34] To circumvent these limitations, the lyngbyatoxin A and cryptomaldamide BGCs were expressed in the well-studied cyanobacterium *Anabaena* (*Nostoc*) PCC 7120 (hereafter, *Anabaena*).[35, 36] The lyngbyatoxin A pathway was further engineered in *Anabaena* to produce pendolmycin and teleocidin B-4.[37] Finally, the pathway encoded by a 42-kb BGC responsible for production of hapalindole alkaloids by the cyanobacterium

Fischerella ambigua UTEX 1903 was engineered into *Synechococcus elongatus* UTEX 2973.[50] Advances in the heterologous expression of NP pathways from cyanobacteria were recently reviewed.[20]

Here, we engineered *Anabaena* to produce columbamides through heterologous expression of the *M. bouillonii* PNG columbamide pathway and identified several natural as well as novel analogs (Figure S1). This result is particularly notable as the columbamide pathway is a representative of a cyanobacterial PKS/NRPS hybrid pathway and the gene cluster is over 28 kb in length. Upon identification and characterization of the columbamides produced by the heterologous host, we further investigated different growth conditions that may influence the production of columbamides in *Anabaena*.

RESULTS AND DISCUSSION

Heterologous expression of the columbamide BGC in *Anabaena*: pathway instability and detection of several columbamide analogs

The *M. bouillonii* PNG columbamide BGC was amplified by PCR from genomic DNA as 6 overlapping fragments of 3 to 6 kb and cloned in one step into pAM5558 (Table S1), which carries sequences flanking the columbamide BGC, by transformation-associated recombination (TAR) in the yeast *Saccharomyces cerevisiae* VL6-48N (Figure 2).[36, 38] Positive yeast clones that contained the columbamide BGC were identified by PCR, and these plasmid constructs were transferred into *E. coli*. Plasmid preparation and analysis by NcoI restriction digests of a dozen clones identified 10 positive clones (Figure S2). Six of these plasmid clones were sequenced by next generation sequencing on a MiSeq platform. One clone, CO45 (hereafter pAM5562), did not carry any missense mutations and was selected for subsequent experiments. The other 5 clones carried 2 to 5 mutations at various positions in the columbamide BGC (Table S2).

Plasmid pAM5562 carrying the columbamide BGC was transferred by natural transformation into the AMC2566 strain of *Synechococcus elongatus* PCC 7942, which carries the *Bacillus subtilis* promiscuous Sfp-PPTase (Table S3). The resulting recombinant clones of *S. elongatus* carrying the columbamide BGC at chromosomal neutral-site 2 (S7942NS2) grew poorly compared to WT when exposed to a light intensity above ~ 50 μmol photon m^{-2} s^{-1} (Figure S3), but these clones did not produce any columbamide analogs at detectable levels. *S. elongatus* is a freshwater clone typically cultured in BG11 medium with only trace amounts of chlorine-containing salts, which we were concerned could affect production of columbamides. To provide additional chlorine for the halogenation of the columbamide acyl chain, the cultures were supplemented with 100 mM NaCl. However, the addition of NaCl to the medium did not result in the production of columbamides (data not shown).

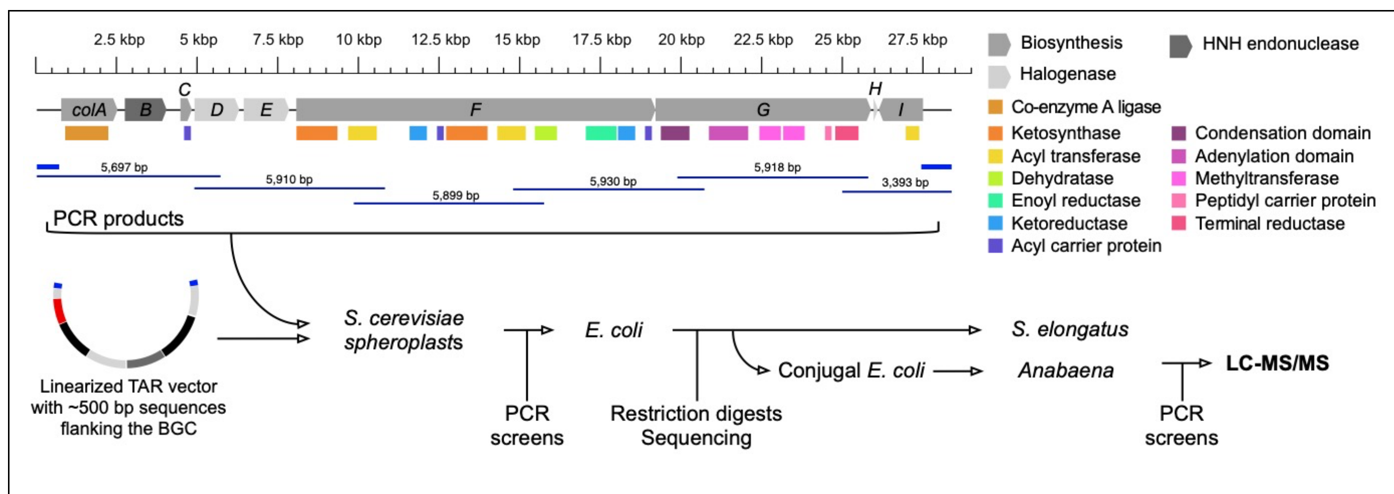


Figure 2. Columbamide BGC and cloning strategy for expression in the heterologous hosts *S. elongatus* and *Anabaena*. Annotations for genes and enzymatic domains are shown in the key. The experimental workflow is shown at the bottom. The columbamide BGC was amplified by PCR from genomic DNA as 6 overlapping fragments covering 28,372 bp starting 766 nucleotides upstream of the *colA* start codon to 893 nucleotides downstream of *colI*. The first and last PCR products carried 692 bp and 941 bp regions (blue bars) that overlap with the ends (blue segments) of the linearized TAR cloning vector pAM5558. The 6 PCR products and pAM5558 were assembled in *S. cerevisiae* by recombination. Yeast clones containing plasmids with the entire BGC were identified by PCR. Positive plasmids were transformed into *E. coli* and further verified by restriction digests with NcoI and next generation sequencing of the entire plasmid. Finally, a clone without any missense mutations was transformed into *S. elongatus* or conjugated in *Anabaena* from an *E. coli* donor strain. TAR cloning vector pAM5558 comprises components for plasmid replication and selection in yeast (light gray) and in *E. coli* (dark gray and red) and carries homologous sequences for recombination into *S. elongatus* and *Anabaena* chromosomes at a neutral site (black).[36]

As we previously reported for the heterologous expression of the cryptomaldamide pathway from *M. producens* JHB,[36] *S. elongatus* failed to produce cryptomaldamide but production in *Anabaena* was successful. Therefore, the columbamide BGC was next transferred into *Anabaena*. The plasmid clone pAM5562 carrying the columbamide BGC was conjugated into *Anabaena* AMC2556 by biparental mating from *E. coli*. The strain AMC2556 was engineered previously to harbor *S. elongatus* S7942NS2 neutral site homology regions flanking an antibiotic resistance gene for nourseothricin (Nt) at a previously identified neutral site in the all1697 gene.[36] Initially, numerous transconjugant colonies grew on selective BG11 plates but none could be recovered as single clones using standard plating methods.

To determine if the transconjugants produced columbamide, all colonies were pooled into a liquid culture and grown with antibiotic selection under relatively low light (35 µmol photons m^{-2} s^{-1}) conditions for *Anabaena* to avoid stressing the cells. Samples of this culture of pooled colonies along with a control strain, AMC2556, were grown in liquid cultures and then analyzed by liquid chromatography – tandem mass spectrometry (LC-MS/MS). The cultures were grown in 100 mL of liquid BG11 medium supplemented with 100 mM NaCl and bubbled with air under constant medium-low light intensity (40-50 μmol photon m^{-2} s^{-1}), which was empirically determined to allow good growth. The experiment was carried out in duplicate, and

each culture was grown under antibiotic selection; the pooled cultures were grown with neomycin (Nm) and the control strain was grown with Nt. Similar to the *S. elongatus* transformants, the pooled *Anabaena* transconjugant cultures initially grew very poorly, indicating that the columbamide BGC has a negative impact on the health of the heterologous hosts. However, the *Anabaena* transconjugant cultures recovered over time to produce dense cultures after 12 days. The cell biomass and the growth medium were extracted separately with butanol (BuOH) and the dried extracts were dissolved in MeOH at 1.0 mg/mL and then analyzed by low resolution LC-MS/MS. Peaks showing spectral similarity based on tandem mass spectrometry (MS/MS) data to columbamide A (**1**) and B (**2**) were detected in the cell biomass extracts of *Anabaena*. LC-MS/MS analyses also revealed a new major peak in these *Anabaena* cell biomass extracts with an *m/z* value of 432.3. This new peak could also be detected as a low abundance peak in the growth medium from the *Anabaena* pooled cultures and in the biomass of *M. bouillonii* PNG, but not in extracts of the control strain or the solvent blank. The new *m/z* feature was consistent with a new monochlorinated analog of columbamide A (**1**), termed here as columbamide K (**11**), and the MS/MS fragmentation pattern was similar to the fragmentation pattern of columbamide A (**1**). The relationships of the expressed metabolites were visualized in a molecular network constructed with the Global Natural Product Social Molecular Networking (GNPS) platform (Figure S4).[39]

Through serial dilutions of the recovered pooled cultures, we were able to isolate colonies and obtain clonal strains of *Anabaena* carrying the columbamide BGC, specifically strains AMC2557-2559. Unfortunately, this process led to the selection of clones that produced only very low quantities of the columbamides. Further selection using a CRISPR-based strategy (Figure S5) described previously[36] to obtain fully segregated double recombinant clones resulted in strains unable to produce detectable levels of columbamides (data not shown). Although our results demonstrated heterologous production of columbamides in *Anabaena*, they also showed that the introduction of the columbamide pathway into *Anabaena* was detrimental for the heterologous host and led to genetic instability.

To obtain clonal strains of *Anabaena* with higher production levels of columbamides, we repeated the conjugal transfer of pAM5562 into *Anabaena* strain AMC2556 but maintained the conjugation plates and downstream cultures under low light to achieve slower growth and reduce metabolic stress on the cells. A dozen colonies isolated on selective BG11 plates were grown as small cultures and then scaled up to 100-mL and then 300-mL liquid cultures grown under low light (Figure 3B). PCR analyses showed that the 12 new clones carried the columbamide BGC (Figure 3A). However, conjugations of our CRIPSR-based segregation plasmid pAM5565 into those strains repeatedly failed, and thus we were unable to obtain fully segregated double recombinant clones (Figure S6), which again suggested that the columbamide BGC is not well tolerated by *Anabaena*.

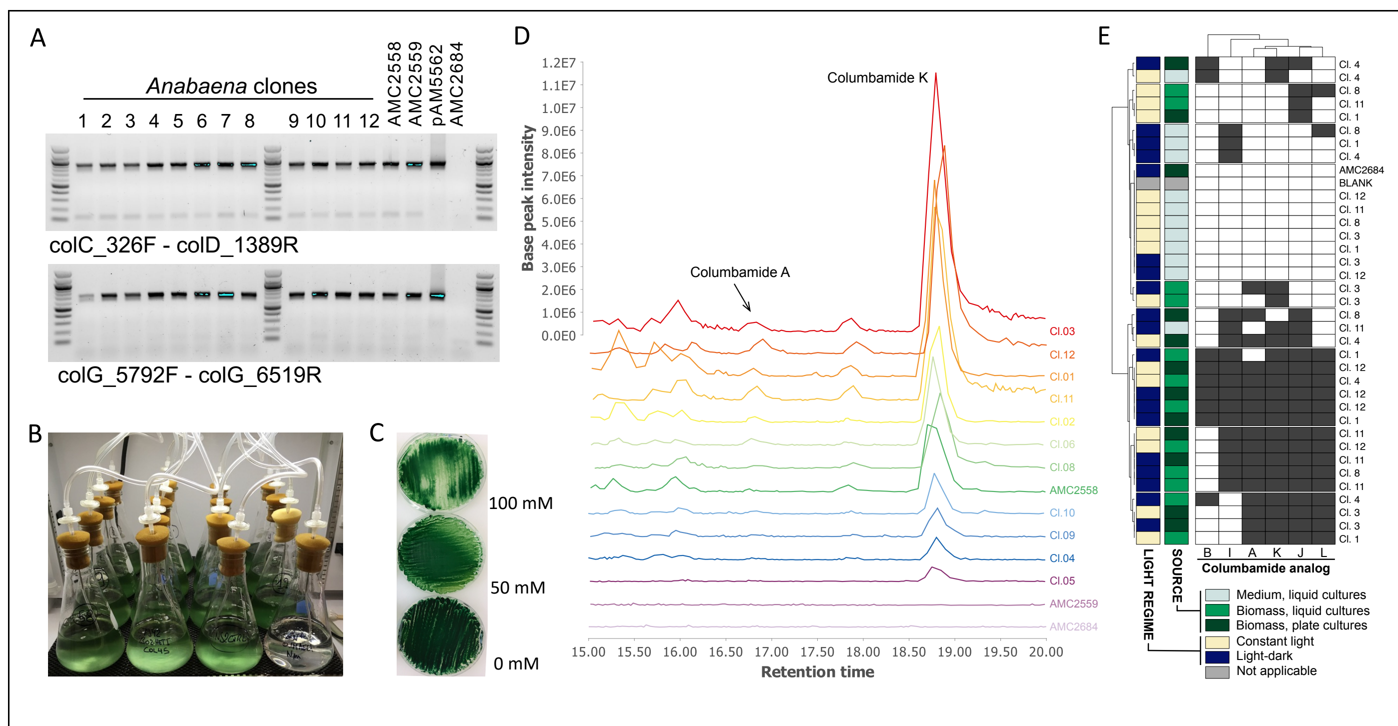


Figure 3. Detection of columbamides in *Anabaena* harboring the columbamide BGC. (A) PCR identification of the columbamide BGC in 12 *Anabaena* clones after conjugation with plasmid pAM5562 that carries the columbamide BGC. Strains AMC2558 and AMC2559 are positive controls and carry the columbamide BGC in the chromosome. Strain AMC2684 is a negative control. (B) Cultures of *Anabaena* in BG11 liquid medium supplemented with 100 mM NaCl; (C) Cultures of *Anabaena* on BG11 agar medium supplemented with different concentration of NaCl; (D) LC-MS extracted ion chromatograms (XICs, 389.3 – 488.9) of extracts obtained from different *Anabaena* cultures grown on BG11 agar plates supplemented with 100 mM NaCl (the *Anabaena* clone or strain numbers are indicated on the right of the chromatogram). (E) Clustering analysis based on the detection (displayed as dark-gray cells) of putative columbamide analogs (identified by one letter codes across bottom) in different samples collected from cultures of *Anabaena* (including extracts from 6 clones carrying the columbamide BGC, a control strain, and a blank extraction) grown in liquid medium or on agar plates. Columbamide M could not be reliably found in this dataset. The cultures were maintained under constant illumination or under 12-hour light/dark cycles. The clustering analysis was done with ClustVis.[40] Rows and column were clustered based on the average distance of all possible pairs using Manhattan and correlation distance methods, respectively.

All 12 new clones, the strains AMC2558 and AMC2559 that were constructed and archived earlier, as well as a negative control strain AMC2684 that carried only sequences flanking the columbamide BGC, were grown for 3 weeks under antibiotic selection (Nm) with low light (~ 20 μmol photon m^{-2} s^{-1}) on BG11 agar plates supplemented with 100 mM NaCl. LC-MS/MS analyses showed the production of various columbamides in all 12 clones. While the strain AMC2558 also produced columbamides, strain AMC2559 had lost this ability. No columbamides were detected in the control strain. As noted above, the major peak detected from extracts of *Anabaena* matched the molecular ion *m*/*z* value of a new monochlorinated analog, columbamide K (**11**) (Figure 3D). In addition, a smaller peak for columbamide A (**1**) was also detected. LC-MS/MS analyses of these 12 clones showed large differences in the columbamide production levels. This likely

resulted from our inability to obtain fully segregated double recombinant clones of *Anabaena* carrying the columbamide BGC. Although we could not obtain segregated clones, we proceeded to further characterize the columbamides produced in *Anabaena* while taking precautions to avoid problems that could be caused by genome instability including analyzing multiple independent clones and maintaining cultures under low-light slow-growth conditions.

Isolation and chemical characterization of new columbamide analogs

To obtain enough material for isolation of the new columbamides, as well as confirm identity of the known analogs, the remaining extracts from cultures of engineered *Anabaena* strains (clones 1-12 and derivative *colDE* overexpression clones described below) that showed the presence of columbamides were pooled into a combined extract (Figure 4). The combined extract was subjected to normal phase vacuum liquid chromatography (NP-VLC) to yield 9 fractions. By LC-MS, three fractions (5, 6, and 7) contained the columbamide analogs I, J, K, L, and M as determined from the distinctive chlorination patterns in their molecular ions. Fraction 5, possessing the largest quantity of columbamides, was fractionated by high performance liquid chromatography (HPLC) into 25 subfractions, and yielded two subfractions containing pure columbamides K (1.3 mg) and A (~ 0.1 mg). Reversed phase (RP) HPLC was performed with VLC fraction 7 and yielded pure columbamide M (**13**) (~ 0.1 mg). The small amounts of columbamides I, J, and L described below were isolated from the initial cultures of pooled colonies.

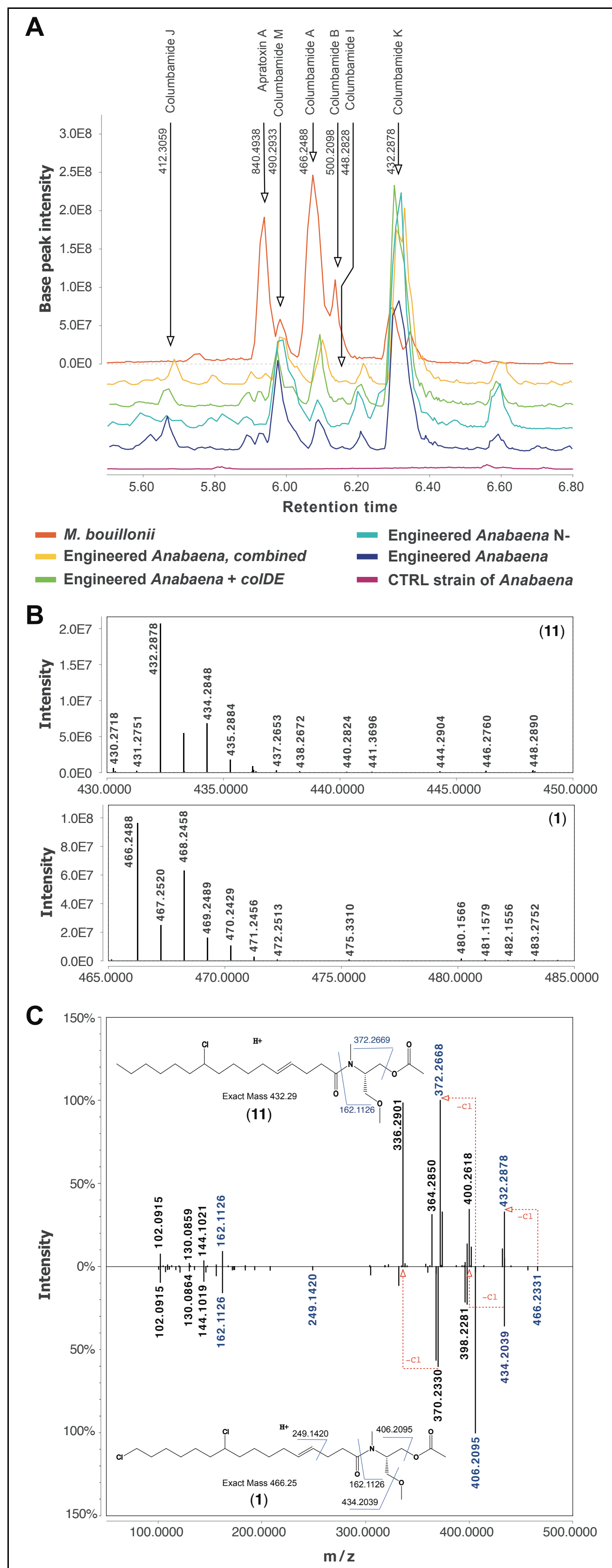


Figure 4. Identification by HR-LC-MS/MS of naturally occurring columbamides as well as novel analogs from heterologous expression of the columbamide BGC in *Anabaena*. (A) Extracted ion chromatogram (*m/z*: 152.0706 - 1028.3202). Columbamide L co-eluted with columbamide A. Cultures of *Anabaena* were grown under different conditions summarized in the key. (B) Representative MS1 spectra for columbamide analogs K (**11**) and A (**1**) from *Anabaena*. (C) MS2 spectra of columbamide analogs K (**11**) and A (**1**) from *Anabaena* in a mirror plot.

Columbamide K (**11**), the most abundant analog produced by heterologous expression of the columbamide BGC in *Anabaena*, was subjected to full planar structure characterization including comparison to previously

described columbamides. Interestingly, the yield of columbamide K from the pooled extract was 1.3 mg, which if present equally in all contributing extracts (212.3 mg) represents a percentage yield of 0.6%; this is highly comparable to the yield of the major compound present in the native producer, columbamide A (2 mg from 389 mg of extract, 0.5% yield).[1] The ^{1}H and ^{13}C nuclear magnetic resonance (NMR) spectra for columbamide K (**11**) showed a high degree of homology to the published ^{1}H and ^{13}C NMR for columbamide A (**1**) (Table 1, Figures S7 and S8). The greatest discrepancies arose from shifts for atoms at the terminus of the acyl chain (C13 through C16), indicating that this was where columbamides A (**1**) and K (**11**) differed. Columbamide K (**11**) lacked the δ_H 3.53 resonance present in columbamide A (**1**) for the ω1-chlorination site, and this was replaced by a new terminal methyl resonance (δ_H 0.88 ppm). These structural differences were further confirmed with a full complement of 2D NMR and high resolution (HR) LC-MS/MS data analyses (Figures 4A-C, 5 and S9-S14), establishing the structure of columbamide K (**11**) as 16-dechloro-columbamide A. Due to the co-occurrence with authentic columbamide A (**1**), we assume that the head group of columbamide K (**11**) also derives from L-serine, and thus is of *S* configuration at C-18. Additionally, we speculate that the secondary chlorine group in columbamide K (**11**) at C-10 is of *R* configuration as was shown for columbamides D-H (**4**-**8**).[2,3]

Table 1. ^{1}H and ^{13}C NMR spectroscopic data for columbamides K (11) and A (1)[1] in $CDCl_3$ at 600 MHz for ^{1}H and 150 MHz for ^{13}C

	Columbamide K					**Columbamide A**					**Difference (δA-δK)**	
No.	**δC**	**C type**	**δH**	**M***	**(J [Hz])**	**δC**	**C type**	**δH**	**M***	**(J [Hz])**	**Δ δC**	**Δ δH**
1	173.58	C				173.5	C				-0.1	0
2	34.10	CH_2	2.40	m		34.2	CH2	2.40	m		0.1	0
3	28.30	CH_2	2.33	m		28.6	CH2	2.32	m		0.3	-0.01
4	129.30	CH	5.46	m		129.3	CH	5.45	m		0.0	-0.01
5	131.18	CH	5.46	m		131.1	CH	5.45	m		-0.1	-0.01
6	32.53	CH_2	1.99	m		32.5	CH2	1.99	m		0.0	0
7	29.08	CH_2	1.38	m		29.1	CH2	1.35	m		0.0	-0.03
8	26.11	CH_2	1.39	m		26.2	CH2	1.35	m		0.1	-0.04
9	38.64	CH_2	1.70	dt	(8.4, 5.3)	38.5	CH2	1.70	h	(8.9, 7.7)	-0.1	0
10	64.32	CH	3.88	tt	(7.9, 4.9)	64.3	CH	3.88	ddd	(12.8, 8.1, 4.9)	0.0	0
11	38.64	CH_2	1.70	dt	(8.4, 5.3)	38.5	CH2	1.70	h	(8.9, 7.7)	-0.1	0
12	26.58	CH_2	1.51	m		26.5	CH2	1.53	td	(9.1, 5.0)	-0.1	0.02
13	29.08	CH_2	1.27	m		28.2	CH2	1.35	m		-0.9	0.08
14	31.90	CH_2	1.29	m		26.9	CH2	1.45	dq	(13.5, 6.7)	-5.0	0.16
15	22.66	CH_2	1.29	m		32.6	CH2	1.78	p	6.9	9.9	0.49
16	14.21	CH3	0.88	t	(6.7, 4.7)	45.2	CH	3.53	t	6.7	31.0	2.65
17	31.90	CH3	2.96	s		31.9	CH3	2.95	s		0.0	-0.01
	27.52		2.82	s		27.6		2.82	s		0.1	0
18	52.10	CH	4.90	h	(7.1, 5.3)	52.1	CH	4.90	tt	(6.8, 6.8)	0.0	0
	55.23		4.26	m		55.3		4.27	m		0.1	0.01
19	71.05	CH_2	3.58	dd	(10.3, 6.8)	71.1	CH2	3.58	dd	(10.3, 6.9)	0.0	0
			3.48	dd	(10.3, 5.0)			3.48	dd	(10.4, 5.3)	0.0	0
20	62.12	CH_2	4.27	dd	(11.6, 7.9)	62.2	CH2	4.26	dd	(11.5, 7.9)	0.1	-0.01
			4.20	dd	(11.7, 5.0)			4.20	dd	(11.9, 4.8)	0.0	0
21	170.89	C				170.9	C				0.0	0
											0.0	0
22	20.26	CH3	2.04	s		21.0	CH3	2.04	s		0.7	0
			2.05	s				2.05	s		0.0	0
23	58.46	CH3	3.32	s		59.1	CH3	3.32	s		0.6	0
			3.34	s				3.34	s		0.0	0

M* - Multiplicity

Columbamide M (**13**), the second most abundant analog in *Anabaena* cultures, was also isolated by HPLC and structurally characterized by MS and ^{1}H NMR. HR-LC-MS/MS of columbamide M (**13**) suggested a molecular formula of $C_{25}H_{44}ClNO_6$, a difference of $C_2H_2O_2$ from columbamide K (**11**), which together with a resonance at δ_H 4.38 for the C16 position suggested the presence of an ω-acetoxy group (Figure S15 and Table 2). HR-LC-MS/MS provided fragments that reinforced this molecular connectivity (Figure S16). Fragmentation between C9 and C10 and loss of HCl generated the $C_6H_{16}O_2^{\bullet+}$ fragment with an *m/z* 156.1019 (calcd as 156.1145). Cleavage of the amide bond yielded a fragment with *m/z* 162.1126, a feature present in all other columbamide analogs that possess an acetylated serinol residue. The isotopic pattern indicating the presence of a single chlorine atom and the increase in mass of 58.0060 amu of the precursor ion when compared to the precursor mass of columbamide K (**11**) suggested that columbamide M (**13**) was the ω-acetoxy analog of columbamide K (**11**).

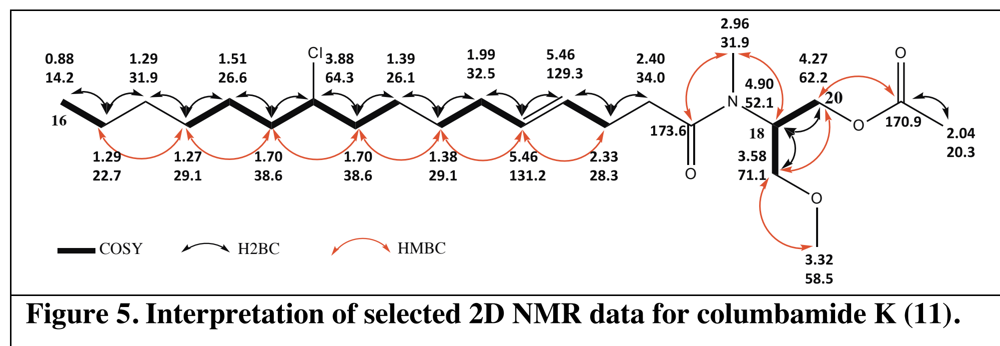


Figure 5. Interpretation of selected 2D NMR data for columbamide K (11).

While detectable amounts of columbamide I (**9**), J (**10**), and L (**12**) were found in cultures of *Anabaena* that express the columbamide BGC, these analogs were present in insufficient quantity for complete chemical characterization.

Table 2. ^{1}H and ^{13}C NMR spectroscopic data for columbamide K (11) and ^{1}H NMR spectroscopic data for proposed columbamides I (9), J (10), and M (13) in $CDCl_3$ at 600 MHz for ^{1}H and 150 MHz for ^{13}C

	Columbamide K			Columbamide I	Columbamide J	Columbamide M
No.	δC [ppm]	C type	δH [ppm]	δH [ppm]	δH [ppm]	δH [ppm]
1	173.58	C				
2	34.10	CH_2	2.40	2.41	2.42	2.38
3	28.30	CH_2	2.33	2.33	2.33	2.35
4	129.30	CH	5.46	5.46	5.46	5.46
5	131.18	CH	5.46	5.46	5.46	5.46
6	32.53	CH_2	1.99	1.99	1.99	1.99
7	29.08	CH_2	1.38	1.38	1.38	1.35-1.41
8	26.58	CH_2	1.39	1.39	1.39	1.35-1.41
9	38.64	CH_2	1.7	1.7	1.7	n.d.
10	64.32	CH	3.88	3.88	3.89	3.98
11	38.64	CH_2	1.7	1.7	1.7	n.d.
12	26.11	CH_2	1.51	1.51	n.d.	n.d.
13	29.08	CH_2	1.29	1.28	1.29	1.27-1.33
14	22.66	CH_2	1.29	1.28	1.29	1.27-1.33
15	22.66	CH_2	1.29	1.28	1.29	1.35-1.41
16	22.66	CH3	0.88	4.12 (t), CH_2	0.89	4.38
17	31.9	CH3	2.96	3.02	3.02	2.96
18	52.1	CH	4.9	4.37	4.37	4.98
19	71.05	CH_2	3.58	3.57	3.57	3.58
			3.46	3.48	n.d.	n.d.

20	62.12	CH_2	4.27	3.79	3.79	4.26
			4.2	N/A	N/A	N/A
21	170.89	C				
22	20.26	CH3	2.04	2.01	N/A	2.05
23	58.46	CH3	3.32	3.34	3.34	3.32
24						
25						2.06

n.d. - Not determined, N/A - Not available

Columbamide I (**9**) analyzed by HR-LC-MS/MS for $C_{23}H_{42}ClNO_5$ showed an $[M+H]^+$ *m/z* 448.2827 (calcd as 448.2824), which indicated that it is a hydroxylated analog of columbamide K (**11**). The acetylated serinol residue was again found intact from the diagnostic amide cleavage that results in a *m/z* 162.1127 peak. Moreover, the combination of its molecular ion being 42 Daltons less than columbamide M (**13**) and possessing only a single chlorine atom suggested that it was the ω-hydroxy analog of columbamide K (**11**). This was confirmed in part from a new resonance for the ω-terminus methylene protons which resonated at δ_H 4.12. In all other respects, it closely matched the chemical shifts and coupling patterns of columbamide K (**11**), and thus is circumspectly identified as the ω-hydroxy analog of columbamide K (**11**).

Columbamide J (**10**) analyzed by HR-LC-MS/MS for $C_{21}H_{40}ClNO_3$ showing an $[M+H]^+$ *m/z* 390.2770 (calcd as 390.2769), which indicated it could be the ω-dechlorinated analog of columbamide C (**3**), which is analogous to the relationship between columbamides K (**11**) and A (**1**). This observation was supported by the absence of any methyl acetate peak in the ^1H NMR spectrum of columbamide J (**10**), the occurrence of the diagnostic *m/z* 120.1019 peak for the des-acetyl serinol fragment, and the presence of a high field methyl triplet at δ_H 0.89 for the ω-methyl group.

Columbamide L (**12**) analyzed by HR-LC-MS/MS for $C_{21}H_{40}ClNO_4$ showing an $[M+H]^+$ *m/z* 406.2724 (calcd as 406.2719), which indicated it could be the des-acetyl analog of columbamide I (**9**). This was supported by observation of cleavage of the amide C-N bond at *m/z* 120.1019, indicating the presence of the des-acetyl serinol head group. Tentatively, columbamide L (**12**) is therefore assigned a structure analogous to previously reported columbamide C (**3**) but wherein the ω-chloro group is replaced with a ω-hydroxy group.

The columbamides were mostly found in the cell biomass of *Anabaena* cultures

To determine the best growth and sampling conditions for the recombinant *Anabaena* strains, we investigated the production of columbamides in the biomass and growth medium from liquid cultures and in the cell biomass from plates. In addition, it has been proposed that some halogenases found in natural product pathways, such as the halogenase responsible for the bromination of jamaicamide A in *Moorena producens* JHB (formerly *Lyngbya majuscula*), are principally active in the dark.[41] Therefore, the cultures were grown under constant light and 12 h:12 h light/dark cycles. For these experiments, the cells were grown under

antibiotic selection at 30°C for 3 weeks in BG11 medium supplemented with 100 mM NaCl. We included the control strain AMC2684 and columbamide BGC clones 1, 3, 4, 8, 11, and 12.

Several columbamide analogs were detected in those samples including columbamide A (**1**), B (**2**), and new analogs I-L (**9-12**) (Figure 1). In liquid cultures, columbamide analogs were mostly detected in the cell biomass (Figure 3E). A few small columbamide peaks were detected in five of the growth medium extracts, likely as the result of cell lyses because the other seven extracts from culture media did not contain detectable quantities of columbamides. These results indicate that the compounds remain in the cell biomass and are not released into the culture medium. We did not find major differences in the composition of columbamides between extracts obtained from cells grown on plates or from cells grown in liquid cultures; neither did we find major differences in the composition or the relative quantities of columbamides in extracts of cultures grown under constant light or light–dark cycles (Figure 3E). Columbamide K (**11**) remained the major analog in extracts from cultures grown under constant light or light–dark regimes (Figure S17).

Heterologous production of columbamides increased with chloride availability

Preliminary experiments indicated that the addition of NaCl to the freshwater BG11 medium was important for the production of columbamides in *Anabaena*. Seawater is approximately 600 mM NaCl. To test the NaCl requirement for *Anabaena* to produce columbamides, cells were grown on agar plates supplemented with 0 mM, 50 mM, or 100 mM NaCl, which is the highest concentration well tolerated by *Anabaena* (Figure 3C). *Anabaena* cultures were grown for 3 weeks under constant low light (20 μmol photon m^{-2} s^{-1}) at 30°C under antibiotic selection. For these experiments, we included the clones 1, 3, 4, 8, 11, and 12 carrying the columbamide BGC. Each clone was treated separately and analyzed in triplicate by LC-MS/MS. The relative quantities of columbamide analogs were determined using palmitoylethanolamide (PEA) as an internal standard.

The results indicated a strong positive correlation between the NaCl concentration in the medium and the columbamides produced. The number of columbamide analogs detected and their relative amounts were consistently higher in cultures supplemented with 100 mM NaCl (Figures 6A and 6B). NaCl concentration is clearly an important factor for the production of the columbamides and may explain the relative prevalence of monochlorinated analogs in cultures of engineered *Anabaena* grown under low chloride salt concentrations.

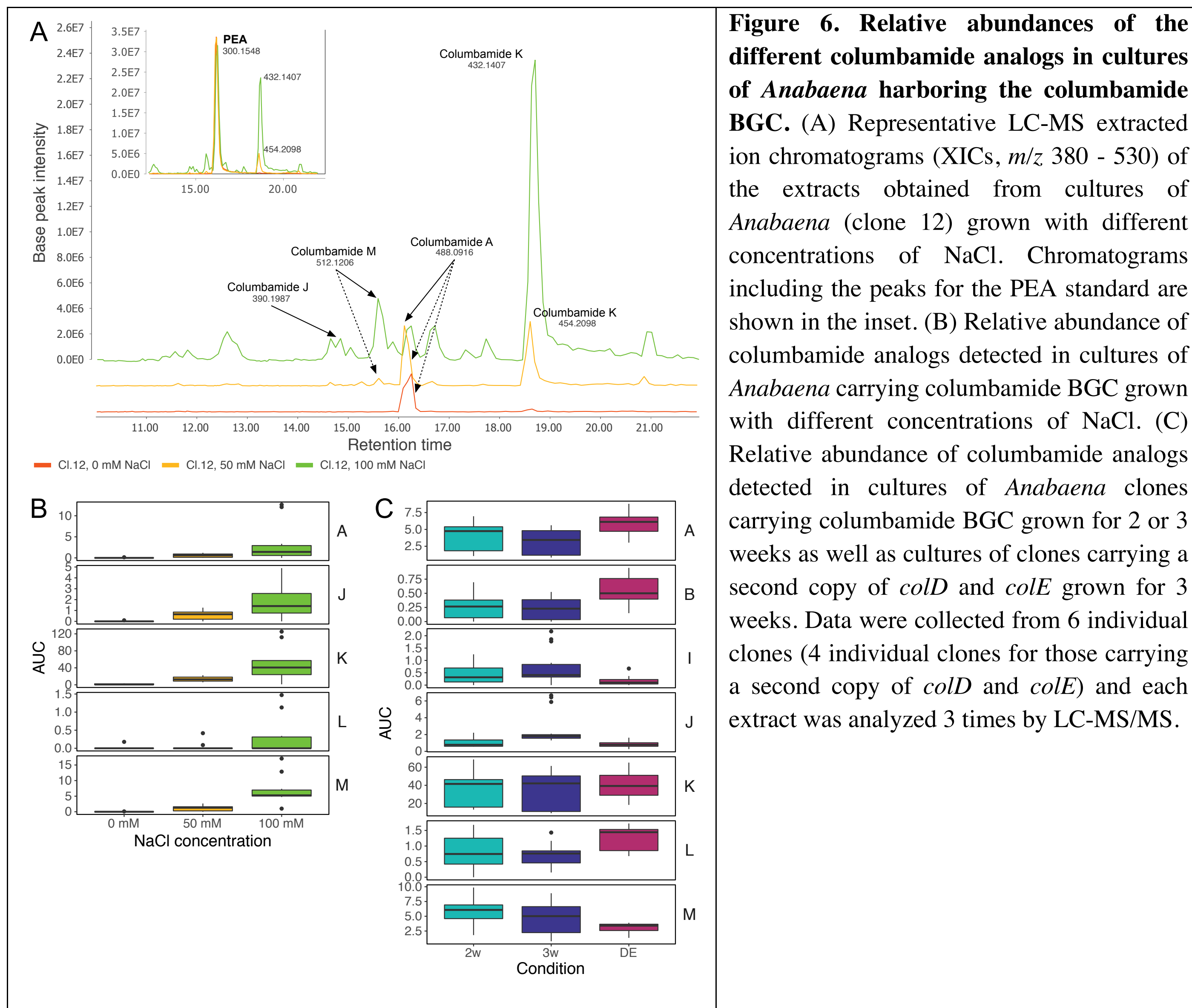


Figure 6. Relative abundances of the different columbamide analogs in cultures of *Anabaena* harboring the columbamide BGC. (A) Representative LC-MS extracted ion chromatograms (XICs, *m/z* 380 - 530) of the extracts obtained from cultures of *Anabaena* (clone 12) grown with different concentrations of NaCl. Chromatograms including the peaks for the PEA standard are shown in the inset. (B) Relative abundance of columbamide analogs detected in cultures of *Anabaena* carrying columbamide BGC grown with different concentrations of NaCl. (C) Relative abundance of columbamide analogs detected in cultures of *Anabaena* clones carrying columbamide BGC grown for 2 or 3 weeks as well as cultures of clones carrying a second copy of *colD* and *colE* grown for 3 weeks. Data were collected from 6 individual clones (4 individual clones for those carrying a second copy of *colD* and *colE*) and each extract was analyzed 3 times by LC-MS/MS.

Heterologous production of columbamides was not increased by growth in medium lacking nitrate

Cyanobacteria do not accumulate lipids under nitrogen starvation as do eukaryotic microalgae.[42] Nevertheless, it was reported that nitrogen starvation of the unicellular cyanobacterium *Synechocystis* PCC 6803 leads to a 2-fold increase in acetyl-CoA,[43] the principal building block of fatty acids. In addition, deprivation of combined nitrogen in filamentous nitrogen-fixing cyanobacteria results in a modest increase of lipid content.[44] *Anabaena* is a nitrogen fixing cyanobacterium that differentiates specialized cells, heterocysts, to fix atmospheric nitrogen when starved for combined nitrogen.[45] The presence of the C16 or C18 fatty acyl chain in columbamides suggested that nitrogen starvation might increase the fatty acid pool available for their biosynthesis, thereby increasing production. We investigated this possibility on $BG11_0$ plates, which lack nitrate, under constant low light growth conditions. Precultures were grown as previously described and the cell aliquots were washed twice with $BG11_0$ before being spread onto $BG11_0$ plates. We included the same set

of clones and the relative quantities of columbamide analogs were determined as described above. We obtained substantially smaller cell biomass from diazotrophic cultures collected after 3 weeks compared to cultures grown with nitrate and detected smaller relative amounts of columbamides (Figure S18). Moreover, two of these cultures from different clones lacked detectable levels of columbamides. A similar result was found for the expression of lyngbyatoxin in *Anabaena*.[35]

Overexpression of *colDE* increased production of di- and tri-chlorinated columbamide analogs

A special feature of the columbamides is the presence of chlorine atoms at the ω1 and ω7 positions of the 16 or 18-carbon acyl chain. This chlorination profile is attributed to the presence of two predicted halogenases, ColD and ColE, in the columbamide BGC.

Because previous analyses revealed that the monochlorinated columbamide K (**11**) was the most abundant analog in cultures of *Anabaena*, we reasoned that overexpressing the pathway's two halogenase homologs encoded by *colD* and *colE* might enhance the halogenation process. Therefore, to increase the amount of di- and tri-chlorinated columbamides produced in *Anabaena* and obtain experimental evidence for ColD and ColE being halogenases, we expressed a second copy of these enzymes on a broad host range plasmid, which is based on an RSF1010 replicon. RSF1010-based replicons are known to stably replicate in *Anabaena* with about 3 copies per chromosome.[46] Both genes were cloned downstream of a conII promoter that functions well in *Anabaena*, as well as other strains of cyanobacteria.[47, 48] The *colDE*-containing plasmid pAM5764 was introduced in *Anabaena* clones 1, 3, 4, 8, 11, and 12 by conjugation. To reduce the risk of selecting derivative clones that had lost the ability to produce columbamides, 3 colonies for each clone were selected, grown as individual cultures, first as liquid cultures, then on agar plates, and finally pooled before extraction and LC-MS/MS analyses. The relative quantities of columbamide analogs were determined as described above.

Transconjugant clones containing a second copy of *colDE* in the original clones 8 and 11 made no columbamide at detectable levels. However, the other clones carrying a second copy of *colDE* produced 1.5 to 2-fold more di- and tri-chlorinated columbamides A (**1**) and B (**2**) than the parent clones (Figure 6C). While these analyses showed only small increases in the halogenation of the acyl chain, they support the role of *colD* and *colE* in halogenation of columbamides. Additional genetic experiments are needed to establish whether ColD or ColE introduces one or two chlorine atoms at the terminal end of the acyl chain, and which one introduces a chlorine atom at the ω-7 position.[1]

Molecular relationships between columbamides produced by the heterologous host and *M. bouillonii* PNG

To obtain more accurate chemical data and confirm the molecular relationships between columbamides produced by the heterologous host and *M. bouillonii* PNG, extracts obtained from cultures of both organisms

were analyzed by HR-LC-MS/MS and molecular networking using the GNPS platform.[39] For these analyses, extracts obtained from the experiments described above of clones 1, 3, 4, 8, 11, and 12 and the clones that overexpress a second copy of *colD* and *colE* were pooled for each condition.

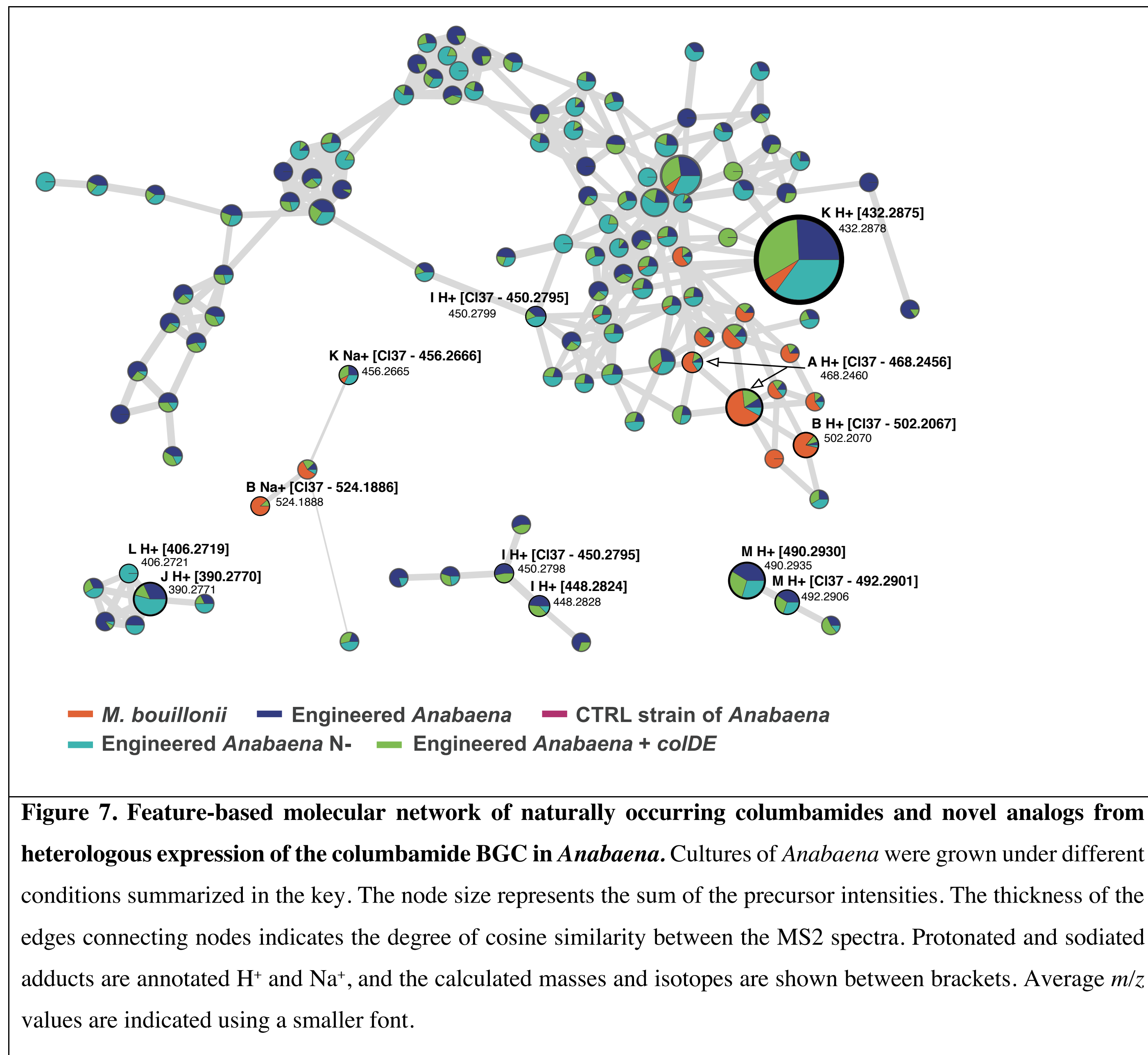


Figure 7. Feature-based molecular network of naturally occurring columbamides and novel analogs from heterologous expression of the columbamide BGC in *Anabaena*. Cultures of *Anabaena* were grown under different conditions summarized in the key. The node size represents the sum of the precursor intensities. The thickness of the edges connecting nodes indicates the degree of cosine similarity between the MS2 spectra. Protonated and sodiated adducts are annotated H^+ and Na^+, and the calculated masses and isotopes are shown between brackets. Average *m/z* values are indicated using a smaller font.

HR-LC-MS/MS and molecular networking analyses confirmed the production of a diversity of columbamide analogs (Figures 4A, 7, S16, and S19), as observed previously in the cultures of engineered *Anabaena*. Both columbamide A (**1**) and K (**11**) were found in cultures of *M. bouillonii* PNG and *Anabaena,* but their relative abundances differed (Figures 4A and 7). While columbamide A (**1**) is the dominant analog in cultures of *M. bouillonii*, columbamide K (**11**) was the major analog in extracts of *Anabaena*. Columbamide K (**11**) was identified above as a new monochlorinated columbamide analog based on its *m/z* value,

fragmentation pattern, and NMR data. In addition, columbamide B (**2**) and the new analogs I (**9**), J (**10**), L (**12**), and M (**13**) were identified (Figures 4A, 7, and S16). Columbamide M (**13**), *m/z* 490.2932, was more abundant than the other analogs in the extracts of engineered *Anabaena*.

The structural diversity among columbamide analogs can be rationalized through an analysis of the biosynthetic pathway (Figure 1). Major structural differences among analogs include the degree of terminal ω1 chlorination, ω1 oxidation, acyl chain length, and acetylation at either terminus. For example, columbamide J (**10**) and K (**11**) differ only in the presence or absence of acetylation of the hydroxy group present at C20 (Figure 1). The HR-LC-MS/MS data show fragments corresponding to acetylated columbamide K (**11**) and nonacetylated columbamide J (**10**) that are consistent with the biosynthetic pathway and similar to the previously reported acetylated/non-acetylated pairs columbamide A/C (**1**,**3**) or D/F (**4**,**6**) and E/G (**5**,**7**), respectively (Figure 1).[1-3] The different degrees of acetylated products could be explained by the variable functioning of the predicted *trans*-acting acyltransferase ColI in the columbamide biosynthetic gene cluster. More intriguing is the presence of an oxygen atom in place of the ω1 chlorine in columbamides I (**9**), L (**12**), and M (**13**), which are detected only in the engineered *Anabaena* extracts. The mechanism by which ColD and ColE chlorinate the columbamides was proposed to be similar to that of non-heme diiron monooxygenase AurF.[1] The occurrence of columbamides I (**9**), L (**12**), and M (**13**) in the cultures of *Anabaena* further suggests that ColD or ColE is responsible for oxygenation of the unactivated terminus of the acyl chain, similarly to AurF.[49] In columbamide M (**13**) the resulting C16 hydroxy group is proposed to also be acetylated by ColI.

Biological activity of columbamide K

From prior studies using a radioligand binding assay and a membrane preparation from transfected HEK-293 cells, it was shown that columbamides A and B were potent binders to CB1 and CB2 receptors.[1] In the current study, columbamide K and newly reisolated columbamide A were evaluated in a spontaneous calcium oscillation assay using mature murine primary neocortical cultures. These cells give distinct responses to modulators of G-protein coupled receptors (GPCRs), including the cannabinoid receptors CB1 and CB2. Surprisingly, no effect was observed for either columbamide A or K in this assay, either in naïve cells or those pretreated with CB1 antagonist AN-281, over a range of concentrations from 1 nM to 10 μM (Figures S20-S25). In this assay configuration, however, only agonist activity would be detected, and thus it remains a possibility that the columbamides may possess antagonist properties through binding to these receptors.

CONCLUSIONS

The slow growing and genetically intractable marine cyanobacterium *M. bouillonii* PNG produces columbamides, which are a family of chlorinated acyl amide natural products. Through heterologous expression in *Anabaena,* we confirmed experimentally that the 28.5-kb BGC proposed by Kleigrewe et al.[1] to encode for columbamide biosynthesis is responsible for its production.

Successful heterologous production of columbamides required overcoming several obstacles. For unknown reasons, expression of the columbamide BGC in *S. elongatus* was unsuccessful. Engineering *Anabaena* to express the columbamide BGC was difficult because of its apparent harmful effects on the strain. The isolation and cultivation of *Anabaena* strains carrying the columbamide BGC under low light intensity was necessary for host viability and genetic stability of the pathway. Production of columbamides in the heterologous host required modification of the culture growth conditions. The culture conditions that generally result in faster growth and higher biomass production, such as higher light intensities, increased CO_2 concentrations, and good mixing conditions resulted in the loss of production of columbamides. Additionally, the columbamides are chlorinated molecules natively produced by a marine cyanobacterium and we found that *Anabaena*'s freshwater medium needed to be supplemented with NaCl to obtain increased production of columbamides.

Unexpectedly, in addition to columbamides A (**1**) and B (**2**) that were previously identified from *M. bouillonii* PNG, the expression of the columbamide BGC in *Anabaena* resulted in the production of several new columbamide analogs including columbamides I-M (**9**-**13**). We later also found columbamide K (**11**) in cultures of *M. bouillonii* but the proportions of columbamide A (**1**) and K (**11**) were very different in the native and heterologous hosts. In *M. bouillonii* the major analog was columbamide A (**1**), whereas in cultures of engineered *Anabaena* the most abundant analog was columbamide K (**11**). Our study suggests that the low concentration of NaCl added to *Anabaena* cultures in comparison to seawater were limiting for the halogenases ColD and ColE to fully chlorinate the acyl chain of the columbamides and may explain the higher proportion of monochlorinated analogs produced by *Anabaena* in comparison to the analogs produced by *M. bouillonii*. Increases in di- and tri-chlorinated columbamides in cultures of *Anabaena* strains overexpressing ColD and ColE further suggest that the level of expression or activity of ColD and ColE in the heterologous host may be another factor that contributes to the production of different analogs by *Anabaena*.

While the activities and mechanisms of ColD and ColE need to be further investigated, our results support that one of these is involved in the halogenation of the columbamides at the terminal end of the acyl chain. Furthermore, the identification of new columbamides I (**9**), L (**12**) and M (**13**) harboring an oxygen atom in place of the ω1 chlorine supports that either ColD or ColE is responsible for oxygenation of the unactivated terminus of the acyl chain using a mechanism similar to the non-heme diiron monooxygenase AurF, as proposed previously.[1] Intriguingly, the halogenase enzyme introducing the secondary chlorine group appears to select C-10 regiochemistry on the basis of distance from the carboxy terminus, as this location is ω7 in columbamides A-C and I-M whereas it is ω9 for columbamides D-H.

Columbamides A and B obtained from *M. bouillonii* PNG are potent ligands of the cannabinoid receptor CB1 and CB2, with columbamide B (**2**) being slightly more active than A.[1] The cannabinoid receptor CB2 is

involved in the regulation of inflammation and has been proposed as a potential therapeutic target.[50, 51] In this current study, an assay that can detect agonists to CB1 or CB2 cannabinoid receptors failed to detect any appreciable activity; however, it remains a possibility that these unusual lipid species possess antagonistic properties. Hundreds of NPs, including NPs with potent biological activities, have been identified from filamentous marine cyanobacteria.[7] However, the slow growth rate and lack of genetic methods to study filamentous marine cyanobacteria have limited the development of their NPs for the pharmaceutical industry. Therefore, heterologous expression of cyanobacterial NP BGCs in *Anabaena* is a promising approach for the identification and development of valuable new NPs. In addition, we found that heterologous expression can result in the production of new analogs, which could further expand the spectrum of compounds available for pharmaceutical evaluation.

Supporting information

Materials and methods; structures of all known columbamide analogs; cloning of the columbamide BGC; growth assay illustrating light sensitivity of *S. elongatus* clones carrying the columbamide BGC; molecular network of columbamides; segregation of double recombinant clones; spectra of columbamide K by ^{1}H NMR, ^{13}C NMR, COSY, HSQC, HMBC, H2BC, LR-HSQMBC, and UV; stacked ^{1}H NMR of new columbamides I, J, K, and M; MS1 and MS2 data of columbamide analogs; chromatograms from light-dark cycle experiments; production of columbamides in the absence of combined nitrogen; plasmids used in this study; mutations for 6 plasmids carrying the columbamide gene cluster; feature-based molecular network of columbamide analogs; structures for cannabinoid receptor agonists and antagonist; dose and time-response relationship for cannabinoid agonist-induced changes in spontaneous calcium oscillations for columbamide analogs; strains and oligonucleotides used in this study. (File: Columbamides_SI.docx)

GenBank format data for maps of the plasmids that were constructed in this study including pAM5558, pAM5562, and pAM5564. (Files: pAM5558.gb, pAM5562.gb and pAM5564.gb)

AUTHOR INFORMATION

Corresponding Authors

William H. Gerwick – *Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography and Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0003-1403-4458; Email: wgerwick@ucsd.edu

James W. Golden – *School of Biological Sciences, University of California San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0001-5463-3207; Email: jwgolden@ucsd.edu

Authors

Arnaud Taton – *School of Biological Sciences, University of California San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0001-5494-5964

Sebastian Rohrer – *Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0003-0107-088X

Brienna Diaz – *School of Biological Sciences, University of California San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0003-4358-9727

Raphael Reher – *Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, United States and Institute of Pharmacy, Martin-Luther-University Halle-Wittenberg, Halle (Saale) 06114, Germany and Institute of Pharmaceutical Biology and Biotechnology, Philipps University of Marburg, Marburg 35037, Germany;* orcid.org/0000-0002-5858-1173

Andres M. Caraballo-Rodriguez – *Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0001-5499-2728

Marsha L. Pierce – *Department of Pharmacology, Midwestern University, Downers Grove, IL 60515, United States;* orcid.org/0000-0002-7265-0104

Pieter C. Dorrestein – *Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0002-3003-1030

Lena Gerwick – *Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0001-6108-9000

Author Contribution

⊥AT and SR contributed equally

A.T., J.W.G., L.G., and W.H.G. conceived the project. A.T. and B.D. constructed the recombinant plasmids and strains. S.R., R.R., and A.M-C. performed LC-MS/MS and NMR experiments. M.L.P. performed bioassays. P.C.D. provided laboratory support for HR-LC-MS/MS analysis. A.T., S.R., R.R., J.W.G, L.G., and W.H.G. analyzed the data and wrote the paper, which was reviewed and edited by all authors.

Conflict of Interest

W.H.G. has an equity interest in Sirenas Marine Discovery, Inc., a company that may potentially benefit from the research results and serves on the company's Scientific Advisory Board. The terms of this

arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict-of-interest policies.

ACKNOWLEDGMENTS

We thank K. Jepsen at the UC San Diego IGM sequencing facility for technical support. We thank E. Glukhov for maintaining the laboratory culture of *M. bouillonii*. We thank Brendan M. Duggan and Nicole E. Avalon for helping with acquisition of NMR data and Jihad Aburas for help with bioassays. We thank S.S. Golden for general laboratory financial support. Funding was provided by the National Institute of General Medical Sciences of the U.S. National Institutes of Health under award number R01GM118815 (J.W.G., L.G., and W.H.G.) and R01GM107550 (W.H.G and L.G.), and by the Midwestern University One Health Initiative Funds (M.L.P. and W.H.G.). A.M.C.-R. and P.C.D. were supported by the Gordon and Betty Moore Foundation through grant GBMF7622, the U.S. National Institutes of Health for the Center P41GM103484 and R01GM107550, and U.S. Department of Energy award DE-SC0021340 subaward 1070261-436503. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

(1) Kleigrewe, K.; Almaliti, J.; Tian, I. Y.; Kinnel, R. B.; Korobeynikov, A.; Monroe, E. A.; Duggan, B. M.; Di Marzo, V.; Sherman, D. H.; Dorrestein, P. C.; et al. Combining Mass Spectrometric Metabolic Profiling with Genomic Analysis: A Powerful Approach for Discovering Natural Products from Cyanobacteria. *J Nat Prod* **2015**, *78* (7), 1671-1682. DOI: 10.1021/acs.jnatprod.5b00301

(2) Lopez, J. A. V.; Petitbois, J. G.; Vairappan, C. S.; Umezawa, T.; Matsuda, F.; Okino, T. Columbamides D and E: Chlorinated Fatty Acid Amides from the Marine Cyanobacterium *Moorea bouillonii* Collected in Malaysia. *Org Lett* **2017**, *19* (16), 4231-4234. DOI: 10.1021/acs.orglett.7b01869

(3) Mehjabin, J. J.; Wei, L.; Petitbois, J. G.; Umezawa, T.; Matsuda, F.; Vairappan, C. S.; Morikawa, M.; Okino, T. Biosurfactants from Marine Cyanobacteria Collected in Sabah, Malaysia. *J Nat Prod* **2020**, *83* (6), 1925-1930. DOI: 10.1021/acs.jnatprod.0c00164

(4) Tronholm, A.; Engene, N. *Moorena* gen. nov., a valid name for "*Moorea* Engene & al." nom. inval. (Oscillatoriaceae, Cyanobacteria). *Notulae algarum* **2019**, *122*, 1-2.

(5) Grindberg, R. V.; Ishoey, T.; Brinza, D.; Esquenazi, E.; Coates, R. C.; Liu, W. T.; Gerwick, L.; Dorrestein, P. C.; Pevzner, P.; Lasken, R.; et al. Single cell genome amplification accelerates identification of the apratoxin biosynthetic pathway from a complex microbial assemblage. *PLoS One* **2011**, *6* (4), e18565. DOI: 10.1371/journal.pone.0018565

(6) Leao, T.; Castelao, G.; Korobeynikov, A.; Monroe, E. A.; Podell, S.; Glukhov, E.; Allen, E. E.; Gerwick, W. H.; Gerwick, L. Comparative genomics uncovers the prolific and distinctive metabolic potential of the

cyanobacterial genus *Moorea*. *Proc Natl Acad Sci U S A* **2017**, *114* (12), 3198-3203. DOI: 10.1073/pnas.1618556114

(7) Jones, M. R.; Pinto, E.; Torres, M. A.; Dorr, F.; Mazur-Marzec, H.; Szubert, K.; Tartaglione, L.; Dell'Aversano, C.; Miles, C. O.; Beach, D. G.; et al. CyanoMetDB, a comprehensive public database of secondary metabolites from cyanobacteria. *Water Res* **2021**, *196*, 117017. DOI: 10.1016/j.watres.2021.117017

(8) Breinlinger, S.; Phillips, T. J.; Haram, B. N.; Mares, J.; Martinez Yerena, J. A.; Hrouzek, P.; Sobotka, R.; Henderson, W. M.; Schmieder, P.; Williams, S. M.; et al. Hunting the eagle killer: A cyanobacterial neurotoxin causes vacuolar myelinopathy. *Science* **2021**, *371* (6536). DOI: 10.1126/science.aax9050

(9) Dittmann, E.; Fewer, D. P.; Neilan, B. A. Cyanobacterial toxins: biosynthetic routes and evolutionary roots. *FEMS Microbiol Rev* **2013**, *37* (1), 23-43. DOI: 10.1111/j.1574-6976.2012.12000.x

(10) Agarwal, V.; Blanton, J. M.; Podell, S.; Taton, A.; Schorn, M. A.; Busch, J.; Lin, Z.; Schmidt, E. W.; Jensen, P. R.; Paul, V. J.; et al. Metagenomic discovery of polybrominated diphenyl ether biosynthesis by marine sponges. *Nat Chem Biol* **2017**, *13* (5), 537-543. DOI: 10.1038/nchembio.2330

(11) Choi, H.; Pereira, A. R.; Gerwick, W. H. The chemistry of marine algae and cyanobacteria. In *Handbook of Marine Natural Products*, Springer Netherlands, 2012; pp 55-152.

(12) Gerwick, W. H.; Coates, R. C.; Engene, N.; Gerwick, L.; Grindberg, R. V.; Jones, A. C.; Sorrels, C. M. Giant marine cyanobacteria produce exciting potential pharmaceuticals. *Microbe* **2008**, *3* (6), 277-288.

(13) Villa, F. A.; Lieske, K.; Gerwick, L. Selective MyD88-dependent pathway inhibition by the cyanobacterial natural product malyngamide F acetate. *Eur J Pharmacol* **2010**, *629* (1-3), 140-146. DOI: 10.1016/j.ejphar.2009.12.002

(14) Balskus, E. P.; Walsh, C. T. The genetic and molecular basis for sunscreen biosynthesis in cyanobacteria. *Science* **2010**, *329* (5999), 1653-1656. DOI: 10.1126/science.1193637

(15) Proteau, P. J.; Gerwick, W. H.; Garcia-Pichel, F.; Castenholz, R. The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experientia* **1993**, *49* (9), 825-829. DOI: 10.1007/BF01923559

(16) Rossi, J. V.; Roberts, M. A.; Yoo, H. D.; Gerwick, W. H. Pilot scale culture of the marine cyanobacterium *Lyngbya majuscula* for its pharmaceutically-useful natural metabolite curacin A. *J Appl Phycol* **1997**, *9* (3), 195-204. DOI: 10.1023/A:1007973009869

(17) Engene, N.; Rottacker, E. C.; Kastovsky, J.; Byrum, T.; Choi, H.; Ellisman, M. H.; Komarek, J.; Gerwick, W. H. *Moorea producens* gen. nov., sp. nov. and *Moorea bouillonii* comb. nov., tropical marine cyanobacteria rich in bioactive secondary metabolites. *Int J Syst Evol Microbiol* **2012**, *62* (Pt 5), 1171-1178. DOI: 10.1099/ijs.0.033761-0

(18) Gomez-Escribano, J. P.; Bibb, M. J. *Streptomyces coelicolor* as an expression host for heterologous gene clusters. *Methods Enzymol* **2012**, *517*, 279-300. DOI: 10.1016/B978-0-12-404634-4.00014-0

(19) Zhang, H.; Boghigian, B. A.; Armando, J.; Pfeifer, B. A. Methods and options for the heterologous production of complex natural products. *Nat Prod Rep* **2011**, *28* (1), 125-151. DOI: 10.1039/c0np00037j

(20) Dhakal, D.; Chen, M.; Luesch, H.; Ding, Y. Heterologous production of cyanobacterial compounds. *J Ind Microbiol Biotechnol* **2021**, *48* (3-4). DOI: 10.1093/jimb/kuab003

(21) Donia, M. S.; Hathaway, B. J.; Sudek, S.; Haygood, M. G.; Rosovitz, M. J.; Ravel, J.; Schmidt, E. W. Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians. *Nat Chem Biol* **2006**, *2* (12), 729-735. DOI: 10.1038/nchembio829

(22) Long, P. F.; Dunlap, W. C.; Battershill, C. N.; Jaspars, M. Shotgun cloning and heterologous expression of the patellamide gene cluster as a strategy to achieving sustained metabolite production. *Chembiochem* **2005**, *6* (10), 1760-1765. DOI: 10.1002/cbic.200500210

(23) Schmidt, E. W.; Nelson, J. T.; Rasko, D. A.; Sudek, S.; Eisen, J. A.; Haygood, M. G.; Ravel, J. Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclinum patella*. *Proc Natl Acad Sci U S A* **2005**, *102* (20), 7315-7320. DOI: 10.1073/pnas.0501424102

(24) Ziemert, N.; Ishida, K.; Weiz, A.; Hertweck, C.; Dittmann, E. Exploiting the natural diversity of microviridin gene clusters for discovery of novel tricyclic depsipeptides. *Appl Environ Microbiol* **2010**, *76* (11), 3568-3574. DOI: 10.1128/AEM.02858-09

(25) Ongley, S. E.; Bian, X.; Neilan, B. A.; Muller, R. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Nat Prod Rep* **2013**, *30* (8), 1121-1138. DOI: 10.1039/c3np70034h

(26) Jones, A. C.; Ottilie, S.; Eustaquio, A. S.; Edwards, D. J.; Gerwick, L.; Moore, B. S.; Gerwick, W. H. Evaluation of *Streptomyces coelicolor* A3(2) as a heterologous expression host for the cyanobacterial protein kinase C activator lyngbyatoxin A. *The FEBS journal* **2012**, *279* (7), 1243-1251. DOI: 10.1111/j.1742-4658.2012.08517.x

(27) Kim, E. J.; Lee, J. H.; Choi, H.; Pereira, A. R.; Ban, Y. H.; Yoo, Y. J.; Kim, E.; Park, J. W.; Sherman, D. H.; Gerwick, W. H.; et al. Heterologous production of 4-O-demethylbarbamide, a marine cyanobacterial natural product. *Org Lett* **2012**, *14* (23), 5824-5827. DOI: 10.1021/ol302575h

(28) Ongley, S. E.; Bian, X.; Zhang, Y.; Chau, R.; Gerwick, W. H.; Muller, R.; Neilan, B. A. High-titer heterologous production in *E. coli* of lyngbyatoxin, a protein kinase C activator from an uncultured marine cyanobacterium. *ACS Chem Biol* **2013**, *8* (9), 1888-1893. DOI: 10.1021/cb400189j

(29) Liu, T.; Mazmouz, R.; Ongley, S. E.; Chau, R.; Pickford, R.; Woodhouse, J. N.; Neilan, B. A. Directing the heterologous production of specific cyanobacterial toxin variants. *ACS Chem Biol* **2017**, *12* (8), 2021-2029. DOI: 10.1021/acschembio.7b00181

(30) D'Agostino, P. M.; Gulder, T. A. M. Direct Pathway Cloning Combined with Sequence- and Ligation-Independent Cloning for Fast Biosynthetic Gene Cluster Refactoring and Heterologous Expression. *ACS Synth Biol* **2018**, *7* (7), 1702-1708. DOI: 10.1021/acssynbio.8b00151

(31) Greunke, C.; Duell, E. R.; D'Agostino, P. M.; Glockle, A.; Lamm, K.; Gulder, T. A. M. Direct Pathway Cloning (DiPaC) to unlock natural product biosynthetic potential. *Metab Eng* **2018**, *47*, 334-345. DOI: 10.1016/j.ymben.2018.03.010

(32) Mutka, S. C.; Carney, J. R.; Liu, Y.; Kennedy, J. Heterologous production of epothilone C and D in *Escherichia coli*. *Biochemistry* **2006**, *45* (4), 1321-1330. DOI: 10.1021/bi052075r

(33) Menzella, H. G.; Reisinger, S. J.; Welch, M.; Kealey, J. T.; Kennedy, J.; Reid, R.; Tran, C. Q.; Santi, D. V. Redesign, synthesis and functional expression of the 6-deoxyerythronolide B polyketide synthase gene cluster. *J Ind Microbiol Biotechnol* **2006**, *33* (1), 22-28. DOI: 10.1007/s10295-005-0038-3

(34) Gao, X.; Wang, P.; Tang, Y. Engineered polyketide biosynthesis and biocatalysis in *Escherichia coli*. *Appl Microbiol Biotechnol* **2010**, *88* (6), 1233-1242. DOI: 10.1007/s00253-010-2860-4

(35) Videau, P.; Wells, K. N.; Singh, A. J.; Gerwick, W. H.; Philmus, B. Assessment of *Anabaena* sp. strain PCC 7120 as a heterologous expression host for cyanobacterial natural products: Production of lyngbyatoxin A. *ACS Synth Biol* **2016**, *5* (9), 978-988. DOI: 10.1021/acssynbio.6b00038

(36) Taton, A.; Ecker, A.; Diaz, B.; Moss, N. A.; Anderson, B.; Reher, R.; Leao, T. F.; Simkovsky, R.; Dorrestein, P. C.; Gerwick, L.; et al. Heterologous Expression of Cryptomaldamide in a Cyanobacterial Host. *ACS Synth Biol* **2020**, *9* (12), 3364-3376. DOI: 10.1021/acssynbio.0c00431

(37) Videau, P.; Wells, K. N.; Singh, A. J.; Eiting, J.; Proteau, P. J.; Philmus, B. Expanding the natural products heterologous expression repertoire in the model cyanobacterium *Anabaena* sp. strain PCC 7120: Production of pendolmycin and teleocidin B-4. *ACS Synth Biol* **2020**, *9* (1), 63-75. DOI: 10.1021/acssynbio.9b00334

(38) Kouprina, N.; Larionov, V. Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nat Protoc* **2008**, *3* (3), 371-377. DOI: 10.1038/nprot.2008.5

(39) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* **2016**, *34* (8), 828-837. DOI: 10.1038/nbt.3597

(40) Metsalu, T.; Vilo, J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res* **2015**, *43* (W1), W566-570. DOI: 10.1093/nar/gkv468

(41) Edwards, D. J.; Marquez, B. L.; Nogle, L. M.; McPhail, K.; Goeger, D. E.; Roberts, M. A.; Gerwick, W. H. Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*. *Chem Biol* **2004**, *11* (6), 817-833. DOI: 10.1016/j.chembiol.2004.03.030

(42) Piorreck, M.; Pohl, P. Preparatory Experiments for the Axenic Mass-Culture of Microalgae .2. Formation of Biomass, Total Protein, Chlorophylls, Lipids and Fatty-Acids in Green and Blue Green-Algae during One Growth-Phase. *Phytochemistry* **1984**, *23* (2), 217-223. DOI: 10.1016/S0031-9422(00)80305-2

(43) Anfelt, J.; Kaczmarzyk, D.; Shabestary, K.; Renberg, B.; Rockberg, J.; Nielsen, J.; Uhlen, M.; Hudson, E. P. Genetic and nutrient modulation of acetyl-CoA levels in *Synechocystis* for n-butanol production. *Microb Cell Fact* **2015**, *14*, 167. DOI: 10.1186/s12934-015-0355-9

(44) Vargas, M. A.; Rodriguez, H.; Moreno, J.; Olivares, H.; Del Campo, J. A.; Rivas, J.; Guerrero, M. G. Biochemical composition and fatty acid content of filamentous nitrogen-fixing cyanobacteria. *Journal of Phycology* **1998**, *34* (5), 812-817. DOI: 10.1046/j.1529-8817.1998.340812.x

(45) Haselkorn, R. Heterocysts. *Annu Rev Plant Phys* **1978**, *29*, 319-344. DOI: 10.1146/annurev.pp.29.060178.001535

(46) Chen, Y.; Taton, A.; Go, M.; London, R. E.; Pieper, L. M.; Golden, S. S.; Golden, J. W. Self-replicating shuttle vectors based on pANS, a small endogenous plasmid of the unicellular cyanobacterium *Synechococcus elongatus* PCC 7942. *Microbiology* **2016**, *162* (12), 2029-2041. DOI: 10.1099/mic.0.000377

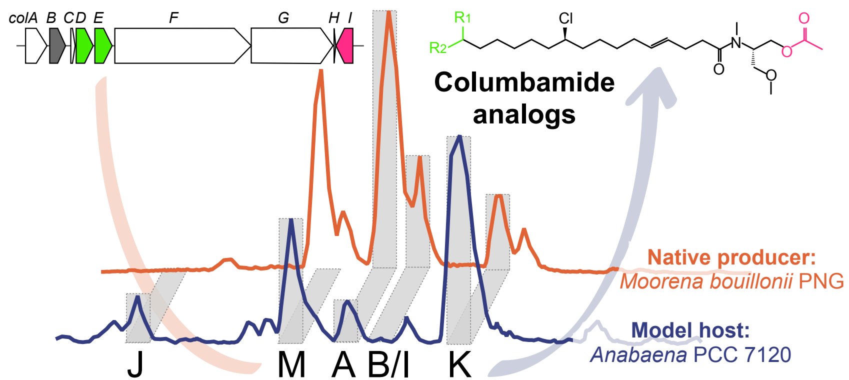
(47) Wei, T. F.; Ramasubramanian, T. S.; Pu, F.; Golden, J. W. *Anabaena* sp. strain PCC 7120 bifA gene encoding a sequence-specific DNA-binding protein cloned by in vivo transcriptional interference selection. *J Bacteriol* **1993**, *175* (13), 4025-4035. DOI: 10.1128/jb.175.13.4025-4035.1993

(48) Taton, A.; Unglaub, F.; Wright, N. E.; Zeng, W. Y.; Paz-Yepes, J.; Brahamsha, B.; Palenik, B.; Peterson, T. C.; Haerizadeh, F.; Golden, S. S.; et al. Broad-host-range vector system for synthetic biology and biotechnology in cyanobacteria. *Nucleic Acids Res* **2014**, *42* (17), e136. DOI: 10.1093/nar/gku673

(49) Choi, Y. S.; Zhang, H.; Brunzelle, J. S.; Nair, S. K.; Zhao, H. In vitro reconstitution and crystal structure of p-aminobenzoate N-oxygenase (AurF) involved in aureothin biosynthesis. *Proc Natl Acad Sci U S A* **2008**, *105* (19), 6858-6863. DOI: 10.1073/pnas.0712073105

(50) Cabral, G. A.; Griffin-Thomas, L. Emerging role of the cannabinoid receptor CB2 in immune regulation: therapeutic prospects for neuroinflammation. *Expert Rev Mol Med* **2009**, *11*, e3. DOI: 10.1017/S1462399409000957

(51) Turcotte, C.; Blanchet, M. R.; Laviolette, M.; Flamand, N. The CB2 receptor and its role as a regulator of inflammation. *Cell Mol Life Sci* **2016**, *73* (23), 4449-4470. DOI: 10.1007/s00018-016-2300-4


colA
B
C
D
E
F
G
H
I
Cl
R1
R2
Columbamide analogs
Native producer:
Moorena bouillonii PNG
Model host:
Anabaena PCC 7120
J
M
A
B/I
K