UC San Diego
UC San Diego Previously Published Works

Title
A Maldiisotopic Approach to Discover Natural Products: Cryptomaldamide, a Hybrid Tripeptide from the Marine Cyanobacterium Moorea producens

Permalink
https://escholarship.org/uc/item/82m6d4db

Journal
Journal of Natural Products, 80(5)

ISSN
0163-3864

Authors
Kinnel, Robin B
Esquenazi, Eduardo
Leao, Tiago
et al.

Publication Date
2017-05-26

DOI
10.1021/acs.jnatprod.7b00019

Peer reviewed

A Maldiisotopic Approach to Discover Natural Products: Cryptomaldamide, a Hybrid Tripeptide from the Marine Cyanobacterium *Moorea producens*

Robin B. Kinnel[†,‡,*], **Eduardo Esquenazi**[§], **Tiago Leao**[†], **Nathan Moss**[†], **Emily Mevers**[⊥,‡], **Alban R. Pereira**[†], **Emily A. Monroe**[||], **Anton Korobeynikov**[∇], **Thomas F. Murray**[○], **David Sherman**[˟], **Lena Gerwick**[†], **Pieter C. Dorrestein**[□,*], and **William H. Gerwick**[†,□,*]

[†]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, USA

[‡]Department of Chemistry, Hamilton College, Clinton, NY, USA

[§]Department of Biology, University of California San Diego, USA

[⊥]Department of Chemistry and Biochemistry, University of California San Diego, USA

[||]Department of Biology, William Paterson University of New Jersey, USA

[∇]Faculty of Mathematics and Mechanics and Center for Algorithmic Biotechnology, Saint Petersburg State University, Russia

[○]Creighton University, School of Medicine, Department of Pharmacology, Omaha, NE, 68178, USA

[˟]Life Sciences Institute, University of Michigan, Ann Arbor, Michigan

[□]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, USA

Abstract

Genome sequencing of microorganisms has revealed a greatly increased capacity for natural products biosynthesis than was previously recognized from compound isolation efforts alone. Hence, new methods are needed for the discovery and description of this hidden secondary metabolite potential. Here we show that provision of heavy nitrogen ^{15}N-nitrate to marine cyanobacterial cultures followed by single filament MALDI analysis over a period of days was highly effective in identifying a new natural product with an exceptionally high nitrogen content.

[*]Corresponding Authors: R.B. Kinnel Tel: (315)-859-4725. rkinnel@hamilton.edu, P.C. Dorrestein: Tel: (858) 534-6607. pdorrestein@ucsd.edu, W.H. Gerwick Tel: (858)-534-0578. wgerwick@ucsd.edu.

[‡]Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, United States.

Notes. The authors declare no competing financial interest.

ASSOCIATED CONTENT

Supporting Information. The Supporting Information is available free of charge on the ACS Publications website at DOI: ^{1}H NMR, ^{13}C NMR, ^{1}H-^{1}H COSY, ^{1}H-^{1}H TOCSY, ^{1}H-^{1}H ROESY, ^{1}H-^{13}C HSQC spectra in D$_2$O for compound **1**, ^{1}H NMR and ^{13}C NMR in DMSO-d_6 for compound **1**, various phylogenetic analysis of putative amidinotransferase CpmA, Marfey's analysis of compound **1**, and MALDI analysis of *M. producens* filaments with and without ^{15}N-nitrate supplementation.

The compound, named cryptomaldamide, was subsequently isolated using MS to guide the purification process, and its structure determined by 2D NMR and other spectroscopic and chromatographic methods. Bioinformatic analysis of the draft genome sequence identified a 28.7 kB gene cluster which putatively encodes for cryptomaldamide biosynthesis. Notably, an amidinotransferase is proposed to initiate the biosynthetic process by transferring an amidino group from arginine to serine to produce the first residue to be incorporated by the hybrid NRPS-PKS pathway. The maldiisotopic approach presented here is thus demonstrated to provide an orthogonal method by which to discover novel chemical diversity from Nature.

Graphical Abstract
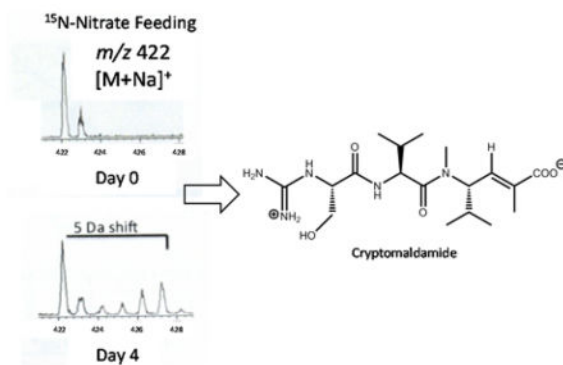
Discovery of secondary metabolites with unusual structural features and promising biological activities has been a productive avenue for research, either for pharmaceutical lead compounds or pharmacological tool compounds. More than 60% of all compounds used in antibiotic, anticancer or other therapeutic applications are either natural products or in some sense derived or inspired by natural products.[1] While traditional approaches of natural products drug discovery by collection, extraction, and biological assay have been highly productive and led to a significant number of clinically useful drugs, such as paclitaxel and yondelis,[2] this methodology is dependent on the nature of the biological assays as well as the presence of an adequate titer of compound to be detected by the screen. Compounds with lower activity but promising structures may not be detected in these screens, and extraction and fractionation schemes may bias the discovery process. Compounds in certain solubility classes, such as water soluble materials, may be missed using the traditional lipid extraction protocols that are most commonly employed.

As productive as traditional "top down" approaches have been and continue to be,[3] genome sequencing projects of microorganisms have revealed that many strains possess a great many more natural product biosynthetic gene clusters than compounds isolated to date. Thus, an opportunity exists to explore this expanded genetic capacity for natural products biosynthesis in diverse classes of microorganisms. A variety of innovative approaches have been developed to investigate this natural products biosynthetic capacity under the general rubric of "genome mining".[4] This includes activation of the expression of pathways through OSMAC approaches,[5] genomisotopic methods,[6] integration with mass spectrometric procedures,[7] co-cultivation with interacting and ecologically-relevant microbial strains,[8]

modulation of expression in the original strain through introduction of promoters or knocking out of repressors,[9] and functional expression in heterologous host strains.[10]

Cultured organisms, primarily microorganisms, have been the richest source of novel bioactive metabolites, a trend true both for terrestrial and marine derived samples.[11] In the latter case, marine cyanobacteria have been an especially rich source of chemically intriguing compounds, resulting in such valuable lead molecules as the dolastatins,[12] curacins,[13] jamaicamides,[14] largazole,[15] and the apratoxins.[16,17] Consequently, we have continued to explore the unique secondary metabolites of filamentous marine cyanobacteria for their bioactive constituents, especially those with anticancer,[18] anti-inflammatory[19] and neuromodulatory properties.[20]

Herein we describe a new method for the detection of expressed natural products in cultured microorganisms that specifically targets molecules with interesting atomic compositions. This strategy involves providing cultures with the heavy isotope form of nitrogen, ^{15}N, in an assimilable form that is subsequently detected in the expressed metabolome of single filaments of the cultured cyanobacteria by MALDI mass spectrometry. In part, this builds on previous efforts in the literature using MALDI analysis of single cyanobacterial filaments to profile peptide-type natural produts.[21] As a variant of the genomisotopic method that we term here a "maldiisotopic" technique, the concept is predicated on the observation and hypothesis that biologically active natural products often contain significant nitrogen content, and that nitrogen atoms are present as components of many privileged structural fragments.[22]

Application of this method to cultured *Moorea producens JHB* (formerly *Lyngbya majuscula JHB*)[23] revealed the presence of a hitherto unknown nonribosomal peptide/polyketide hybrid natural product, given the name cryptomaldamide (**1**). Intriguingly, its structure, characterized by a combination of spectroscopic, degradation and chromatographic methods, is conceptually similar to the sponge derived natural product hemiasterlin,[24,25] a potent antimitotic agent. However, cryptomaldamide was not cytotoxic to cancer cells, possibly due to its predicted high total polar surface area.

Additionally, this discovery of cryptomaldamide suggests that the true source of hemiasterlin may derive from a microorganism, possibly a cyanobacterium, living in association with the sponge. There are several closely related hemiasterlins found in at least four different sponges, namely *Auletta, Siphonochlaina* sp., *Cymbastela* and *Hemiasterella*, which further supports the notion that their ultimate source is a microorganism. A genome sequence of the cryptomaldamide-producing cyanobacterial strain was subsequently analyzed, and a biosynthetic gene cluster located that is generally consistent with the predicted biosynthesis of cryptomaldamide (**1**).
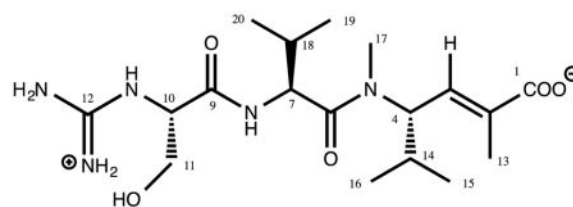
Cryptomaldamide (**1**)

RESULTS AND DISCUSSION

Cultured filaments of *M. producens* JHB, originally collected in Hector's Bay, Jamaica in 1996, were analyzed by MALDI MS and gave peaks for the sodium and potassium salts of jamaicamides A and B[14] as well as of hectochlorin[26] and pheophytin A, all compounds previously reported from this cyanobacterium (for methods see reference 27). In addition, a new peak for an unknown compound was observed at [M+H]$^+$ *m/z* 400 (Figure S13). When this strain was grown in the presence of ^{15}N-sodium nitrate followed by MALDI analysis of isolated filaments over a period of several days, the number of *N*-atoms in each of these metabolites was revealed by a careful inspection of the isotope cluster for their protonated molecules or sodium adduct ions.[27] Specifically, jamaicamides A and B each possessed two *N*-atoms as did hectochlorin whereas pheophytin A contained four *N*-atoms (Figure S13). In addition, the new compound at [M+H]$^+$ *m/z* 400, which had a sodium adduct at *m/z* 422, was found to possess five *N*-atoms (Figure 1). Based on this surprisingly high nitrogen content, we were inspired to isolate this new metabolite and characterize its molecular structure, biological properties and biosynthesis.

M. producens JHB filaments were harvested from a normal laboratory culture (e.g. cultured with ^{14}N-nitrate as the nitrogen source) and extracted first with CH$_2$Cl$_2$/MeOH (2:1) and then with EtOH/H$_2$O (1:1). The latter extract was separated into four fractions by SPE on RP-18; the second fraction, eluting with 1:3 MeOH/H$_2$O, contained the unknown compound with [M+H]$^+$ *m/z* 400 by MALDI MS analysis. This compound exhibited poor chromatographic behavior by LCMS analysis, giving a broad peak when separated by RP using unbuffered H$_2$O/MeCN, suggesting a possible zwitterionic structure. Further purification of this material by RP-HPLC gave a pure, optically active white solid showing an [M+H]$^+$ of *m/z* 400.2559 for a protonated molecular formula of C$_{18}$H$_{34}$N$_5$O$_5$. We named this new metabolite cryptomaldamide in recognition of its hidden presence in the organism that was revealed by MALDI mass spectrometry.

The 800 MHz ^{1}H NMR spectrum of isolated cryptomaldamide in D$_2$O showed four doublet methyl groups at δ$_H$ 0.65–0.85 that were coupled to two multiplets near δ$_H$ 2.0, indicating the presence of two isopropyl groups. By COSY, one of these was adjacent to a methine proton at δ$_H$ 4.45 appearing as an 8.5 Hz doublet and with an associated carbon at δ$_C$ 57.3. By HMBC, this unit was completed by an amide type carbonyl at δ$_C$ 173.8, thus defining a valine residue. The second isopropyl group methine proton was adjacent to another methine proton and appeared as a double doublet of 9.8 Hz and 8.4 Hz. The second coupling to this latter proton was from an olefinic proton resonating at δ$_H$ 6.21. In turn, this latter proton

showed a 1.2 Hz allylic coupling to a vinyl methyl group at δ_H 1.68. By HMBC, this methyl was proximate to a non-protonated vinyl carbon at δ_C 138.4 and near a carboxyl functionality at δ_C 177.7, thereby defining an unsaturated and methylated ketide-extended valine residue. By HMBC, the *N*-atom of this residue was modified with a pendant methyl group. A third amino acid residue possessed an alpha proton at δ_H 4.23 (δ_C 56.6) appearing as a triplet that was coupled to a deshielded methylene group with proton resonances at δ_H 3.78 and δ_H 3.82. These proton and carbon (δ_C 62.4) shifts indicated that this methylene group bore a hydroxy group, and thus with HMBC connections to an amide carbonyl at δ_C 170.5, defined a serine residue. Left unassigned at this point was a fragment comprised of CH_3N_2; however, none of these protons were visible in the D_2O spectrum, and the only NMR resonance associated with this fragment was a carbon appearing at the distinctive chemical shift of δ_C 157.6. Comparison with literature values[28] suggested that this belonged to a guanidine carbon, and completed the fourth partial structure that accounted for all atoms in cryptomaldamide (**1**).

The sequence of residues in cryptomaldamide was established by analyses of both the HMBC and MS/MS datasets (Figure 2, 3). The lone *N*-methyl group showed correlations to both the valine carbonyl and α-proton of the ketide extended valine residue, linking these two components. Similarly, the valine α-proton showed an HMBC correlation to the serine carbonyl, providing this connection. Finally, the position of the guanidine group on the serine residue was clearly established by an HMBC correlation between its carbon atom and the α-proton at δ_H 4.23. MALDI MS fragment masses were observed for the Val-ketide (*m/z* 172), Val-Val-ketide (*m/z* 271) and amidino-Ser (*m/z* 131) sections, confirming these residue connections (Figure 3).

The chemical shifts of the alkene carbons and carboxyl functionality in cryptomaldamide (**1**) are somewhat unusual. The latter carbon is deshielded (δ_C 177.7) compared to the normal shifts of an unsaturated carboxylic acid. The shielded carbon of the alkene corresponds to the β position whereas the deshielded carbon corresponds to the α carbon; this is the opposite electron density of what is usually observed for α,β unsaturated carboxylic acids. This may result from the fact that the carboxyl group must be fully ionized in aqueous solution owing to the presence of the strongly basic terminal guanidine group in cryptomaldamide. It has been shown that the carboxyl carbon in aliphatic carboxylic acids is typically deshielded by about 5 ppm in aqueous solution when present as the anion.[29] The chemical shift for the carboxyl group in the structurally related hemiasterlins, which possess the same vinylogous valine amino acid end group, was reported at δ_C 172.1 in $CDCl_3$,[24] δ_C 170.7 ppm in CD_3OD[30] and at δ_C 168.5 in deuterated DMSO.[25] The observed chemical shift of this carbon in cryptomaldamide in D_2O thus corresponds to a deshielding of 6–9 ppm, consistent in both direction and magnitude predicted for an ionized carboxylate. Furthermore, a deshielding of 3–4 ppm is observed for the α-carbon and 1–2 ppm for the β-carbon in simple aliphatic carboxylates.[29] This would predict chemical shifts for the alkene in cryptomaldamide to resonate at about 137 and 140 ppm for the α and β carbons respectively, using hemiasterlin C as the starting reference. Conjugative electron release by the carboxylate anion should move the carbon shift of the β carbon to a more shielded position, leading to the observed inversion in chemical shifts compared to most enones.

Interestingly, the ^{13}C NMR spectrum of a similar alkene in criamide A, which also contains a guanidine group, has inverted chemical shifts, although the magnitude is not as great because the conjugation is with an amide rather than a carboxylic acid.[25]

Absolute configurations of the residues in cryptomaldamide (**1**) were established by modified Marfey's analysis.[31] Ozonolysis of cryptomaldamide, followed by oxidative workup, hydrolysis and derivatization with Marfey's reagent gave L-*N*-methylvaline, L-valine and L-serine derivatives when analyzed by LCMS and compared to authentic standards. In a separate reaction series, acid cleavage of cryptomaldamide without ozonolysis, followed by derivatization with Marfey's reagent, afforded the derivatives of L-valine and L-serine, providing additional proof for the configurations of these two residues.

The geometry of the alkene in cryptomaldamide (**1**) was assigned on the basis of the ^{13}C NMR chemical shift of the C-13 vinyl methyl group (δ_C 15.0) which was definitive for the *E*-configured double bond. The ^{13}C NMR chemical shift of the comparable vinyl methyl group in various hemiasterlin derivatives, all of which possess an *E*-configured olefin, ranges from 13.3 to 14.6 ppm in various solvents.[24,25,30] This matches well with a ChemDraw prediction for the vinyl methyl group on this *E* double bond (δ_C 14.5) whereas the *Z*-double bond is predicted to be deshielded by 5 to 6 ppm (ChemDraw prediction δ_C 20.5). Additional ROESY correlations between the *N*-CH$_3$ and the C-3 olefinic proton, as well as between the C-3 vinyl proton and the proximal central isopropyl proton, buttress the assignment of the olefin through comparisons to models as well as to the X-ray structure of the methyl ester of hemiasterlin,[32] confirming that the C-2-C-3 double bond in cryptomaldamide is *E*.

A number of intriguing structural elements exist in the structure of cryptomaldamide (**1**), especially the amidino-serine residue and the ketide extended valine residue. To explore the biosynthetic origin of these features, a genome sequencing and bioinformatics approach was undertaken. DNA was extracted from the cultured biomass of *M. producens* JHB using phenol/CHCl$_3$/isoamyl alcohol with cetyl trimethyl ammonium bromide (CTAB) to remove polysaccharides. Genome sequencing was accomplished using paired-end Illumina HiSeq 2x100, and reads were assembled using SPAdes 3.0.[33] Error corrections were accomplished using BayesHammer[34] followed by scaffold assembly with Opera and binning.[35] The quality of the assembly was further enhanced by development of a 'reference genome' from another *Moorea producens* strain from Palmyra Atoll (PAL-15AUG08-1, reported separately,[36] GenBank: NZ_CP017599.1). The genome was annotated using RAST and antiSMASH for secondary metabolite pathways.[37,38] The refined assembly was comprised of 210 contigs organized into a single chromosomal scaffold with an overall genome size of 9.38 Mb. The GC content was 43.67%, consistent with other *Moorea* strains,[39] and two plasmid scaffolds were identified to also be present in this strain. The N50 value was 552,468, and 101 of 102 housekeeping genes were present.[40] Remarkably, antiSMASH revealed the presence of 44 separate biosynthetic gene clusters (BGCs) in this genomic dataset, and these were manually evaluated for consistency with the structure of cryptomaldamide (**1**). One contig containing a partial NRPS gene cluster was identified as the best fit for encoding the amino acids present in cryptomaldamide, however, the lack of a logical upstream initiation region represented an obstacle in its analysis. Upon further

genome mining, another contig, containing a gene predicted to feature amidinotransferase activity, was found to be a strong candidate to complete the gene cluster. PCR amplification of overlapping fragments between the two contigs and Sanger sequencing confirmed their proximate connection. The putative cryptomaldamide BGC of 28.7 kb is available at the NCBI database (accession Genbank: KY315923).

The putative cryptomaldamide BGC (Table 2) appears to be composed of only two biosynthetic Open Reading Frames (ORFs) (Figure 4). The first, *cpmA* (1095 nucleotides), was predicted to encode for a protein with a single enzymatic activity. This was annotated as having high sequence identity (58%) to that of the described amidinotransferase (StxG) involved in saxitoxin biosynthesis.[41, 42] CpmA forms a subgroup with StxG proteins with strong support (77% in Figure S10; 92% in PhyML tree, Figure S11)**,** but it is distinct from the other StxG sequences. Based on sequence analysis, CpmA contains the conserved residues of the catalytic triad (Asp 254, His 303, and Cys 407) but possesses substitutions at Met 302 and Asn 300 as was also noted for CyrA in cylindrospermopsin biosynthesis.[43] As suggested by Muenchhoff, these substitutions may affect substrate specificity, and the Met 302 substitution to cysteine is unique to CpmA. Collectively, these analyses provide strong support for CpmA's role as an amidinotransferase in cryptomaldamide biosynthesis.

The *cpmA* sequence was followed by a large gene, *cpmB*, comprised of 19,320 nucleotides, that is predicted to encode for a megasynthetase with three NRPS modules and one PKS module. Analysis of the amino acid pocket residues of the first adenylation domain predicts activation of cysteine whereas that of the second and third are predicted to activate valine. The second of the two valine encoding NRPSs also possesses a sequence annotated as an *N*-methylase, consistent with the structure of cryptomaldamide. The PKS module that follows this NRPS has all of the requisite domains for *C*-methylation, ketoreduction and dehydration to form the observed terminus of cryptomaldamide. The pathway terminates with a thioesterase that is predicted to catalyze simple water hydrolysis.

Although the first adenylation domain predicts activation of cysteine instead of serine, the structural similarity between cysteine and serine make it plausible that this domain is indeed involved in cryptomaldamide biosynthesis. At this point, it is uncertain if the amidinotransferase reaction occurs with the free amino acid serine, or with the ACP-bound substrate; however, based on precedence in the cylindrospermopsin[43] and saxitoxin cases,[41,42] we believe this likely occurs on the free amino acid, and hence it would be amidino-serine that is activated by the CpmB A1 adenylation domain. Such a modification may also explain why the bioinformatics analysis suggests that cysteine rather than serine is activated by this adenylation domain. Nevertheless, all required enzymatic reactions for cryptomaldamide biosynthesis are present in the putative biosynthetic gene cluster, and it is entirely co-linear with its structure and predicted order of assembly (Figure 4).

Cryptomaldamide (**1**) was evaluated for cytotoxicity to H-460 human lung cancer cells but was found to be inactive (IC_{50} = 40 μM) compared to the hemiasterlins and milnamide.[44] Predictions of its high Total Polar Surface Area using ChemDraw software (tPSA = 173.4; CLogP = −3.53) may explain the lack of activity at an intracellular target (e.g. it does not pass through the cell membrane). Considering that its guanidine group might resemble the

pharmacophore in saxitoxin,[45] cryptomaldamide was also evaluated for blocking effects of the mammalian voltage gated sodium channel Nav1.4; however, it was also inactive in this assay when evaluated at a maximum concentration of 10 μM.

CONCLUSION

The maldiisotopic approach described herein, involving the combination of ^{15}N-feeding to cyanobacterial cultures with MALDI analysis of the expressed nitrogen-containing metabolome, provides an orthogonal way by which to discover and isolate natural products of novel structure. The methodologies are very sensitive and require relatively little material, as it is possible to directly analyze individual filaments. Used in conjunction with genomic methods that can make tentative predictions about functionalities that might be present, this variant of the genomisotopic method[6] promises to be a sensitive and rapid technique for detecting and identifying new, potentially bioactive substances. Application of this method to a Jamaican collection of the cyanobacterium *Moorea producens* JHB resulted in the discovery of cryptomaldamide (**1**), a unique hybrid NRPS-PKS metabolite with a rare amidino-serine residue. Subsequent sequencing of the genome of this strain facilitated discovery of the putative biosynthetic gene cluster for cryptomaldamide (**1**). The pathway is co-linear with the structure of cryptomaldamide, and possesses a unique amidinotransferase at the start of the pathway which is proposed to produce the amidino-serine residue, in analogy to the biochemically described formation of guanidino-acetate as the initiating event in cylindrospermopsin biosynthesis.[43] To date, potent biological properties have not been discovered for this metabolite, possibly due to its highly polar nature which would likely preclude its cell penetration.

EXPERIMENTAL SECTION

General Experimental Procedures

Optical rotation was measured on a Jasco P-2000 polarimeter (JASCO Analytical Instruments, Easton, MD, USA). ^{1}H and ^{13}C NMR spectra as well as 2-D NMR data were collected on a Varian vs800 instrument (^{1}H NMR referenced to the HOD peak at δ_H 4.71; ^{13}C NMR was referenced to residual MeOH in D_2O at δ_C 49.5).[46] LC-MS data was obtained on a Thermo-Electron LTQ LC/MS instrument with a Phenomenex Luna 5 μm C18(2) 100 Å column (4.6 × 250 mm) and a gradient starting at 60% MeCN/40% H_2O and immediately ramping to 100% MeCN over 20 min, then holding at 100% MeCN for 5 min. MS fragmentation experiments were obtained using a Biversa Nanomate (Advion Biosystems) electrospray source for a Finnigan LTQ-FTICR-MS instrument (Thermo-Electron Corporation) running Tune Plus software version 1.0. HPLC purification was carried out with a Waters 515 HPLC pump with a Waters 996 photodiode array detector using Empower Pro software and a Phenomenex Synergi 4 μm C18 column (10 × 250 mm) and a gradient starting at 50:50 MeCN/H_2O for 10 min, ramping to 100% MeCN over 12 min and holding for 3 min. All solvents were HPLC grade.

Cultivation of the Cyanobacterium *Moorea producens* JHB

The original isolate was obtained from field collections made in 2 meters water depth at Hector's Bay, Southeast Jamaica in 1996. Cultures were established from single filaments grown in seawater (SW) BG11 medium as previously described.[14] The culture mass for chemical extraction was produced by growth in the same medium at 28 °C, 16 h light/8 h dark in either Nalgene pans (5 L of medium) covered with gas permeable plastic wrap or 2.8 L Fernbach flasks containing 0.5 to 0.75 L of SWBG11 medium. Harvesting of filaments was achieved by filtration over a Buchner funnel using #1 Whatman filter paper.

Isolation of Cryptomaldamide

A 21.8 g (wet wt) sample of *Moorea producens* JHB, grown in laboratory culture as described above, was extracted five times with 2:1 CH_2Cl_2/MeOH by soaking for 15–20 min. The total volume of the extracts was 300 mL, and upon evaporation of the solvent, gave 0.129 g of a crude oily tar. The extracted cyanobacterial filaments were air-dried overnight giving 2.18 g of dry extracted biomass. The biomass was then extracted with 1:1 EtOH/H_2O by sonicating for 10 min and allowing the mixture to stand at −20 °C for three days, followed by filtration and evaporation to yield 0.52 g of a polar extract. After dissolution/ suspension in 6 mL of H_2O, the mixture was placed on an activated SPE C18 cartridge (Phenomenex polymeric) and washed successively with 20 mL each of H_2O, 3:1 H_2O/ MeOH, 1:3 H_2O/MeOH and 100% MeOH. MALDI analysis of the samples showed that the *m/z* 400 compound (cryptomaldamide) was present in the 3:1 H_2O/MeOH fraction (10.6 mg). Preparative HPLC (4 μ Phenomenex Synergi column, 10 × 250 mm, 70:30 MeCN/ H_2O) afforded 1.5 mg of pure cryptomaldamide (**1**) that had a t_r = 23–27 min under these conditions.

Cryptomaldamide (1)—White solid. $[\alpha]_D^{24°}$ −37.8 (*c* 0.52, MeOH); ^{1}H and ^{13}C NMR data, see Table 1; LRESI MS/MS *m/z* 271, 222, 194, 172, 141, 131; HRESIMS *m/z* 400.2559 [M+H]$^+$ (calcd for $C_{18}H_{34}N_5O_5$, *m/z* 400.2560).

Ozonolysis, Acid Hydrolysis and Marfey's Analysis of Cryptomaldamide (1)

Ozone was bubbled through a sample of cryptomaldamide (0.2 mg) dissolved in MeOH (300 μL) at −78 °C for 5 min, followed by the addition of 5 drops of 30% H_2O_2. The reaction mixture was warmed to room temperature and the solvent was removed under N_2 (g). The product was resuspended in 6 N HCl (300 μL), then microwaved at 160 °C for 5 min and dried under N_2 (g). The hydrolysate was treated with 1.1 mL of 1 mg/mL solution of D- FDAA [1-fluoro-2,4-dinitrophenyl-5-D-alanine amide] in acetone followed by the addition of 300 μL of 1 M NaHCO$_3$ (aq). After the mixture was maintained at 40 °C for 1 h, it was quenched by the addition of 300 μL of 1 M HCl (aq) and then dried under N_2 (g). The residue was resuspended in 100 μL of 50:50 MeCN/H_2O and 15 μL of this solution was analyzed by LC-ESIMS.

The Marfey's derivatives of the hydrolysate and standards (Val, *N*MeVal, and Ser) were prepared similarly with D-FDAA and analyzed by LC-ESIMS using a Phenomenex Luna 5 μm C_{18} column (4.6 × 250 mm). The HPLC conditions began with a 1% MeCN/H_2O + 0.1% formic acid (FA) isocratic run for 5 min followed by a gradient to 40% MeCN/H_2O

+ 0.1% FA over 70 min, followed by another gradient to 65% MeCN/H$_2$O + 0.1% FA over 15 min, and finally held for 5 min at a flow of 0.4 mL/min, monitoring from 200 to 600 nm. The retention times of the D-FDAA derivatives of the authentic amino acids were: L-Ser (67.7 min), D-Ser (66.7 min), L-Val (92.8 min), D-Val (88.0 min), L-*N*MeVal (94.2 min) and D-*N*MeVal (92.0 min); the derivatives from the hydrolysate product gave peaks with retention times of 67.7, 92.9, and 94.3 min, corresponding to L-Ser, L-Val, and L-*N*MeVal.

PCR Confirmation of Connection Between *cpmA* and *cpmB*

PCR confirmation of the connection between contigs containing *cpmA* and *cpmB* was carried out using Phusion polymerase with the following overlapping primer pairs: pair 1 (F1-CGACGGAGGTTGCTACGG, R1-ATAGTAACCGTGATTCGCG), and pair 2 (F2-TGTTCACACAATAATCTGC, R2-CCTCAATTAGGTTAGCATCTAC), which produced gel bands of predicted size: 3.7Kb and 2.8Kb, respectively. The amplification products of F1/R1 (3.6Kb) and F2/R1 (1.2Kb) were purified using the Zymo-5 Clean and Concentrate Kit as per manufacturer's instructions and sequenced by Sanger sequencing; this confirmed the connection between the *cpmA* and *cpmB*-containing contigs.

H-460 Human Lung Cancer Cell Toxicity Assay

H-460 cell viability was determined by mitochondrial-dependent reduction of MTT to formazan in 96-well plates, as previously described.[47] Cells were added at 3.33×10^4 cells/mL in Roswell Park Memorial Institute (RPMI) 1640 medium with fetal bovine serum (FBS) and 1% penicillin/streptomycin, compounds added in DMSO and then incubated for 48 h before staining with MTT (1 mg/mL, 37 °C, 25 min). In brief, the medium was aspirated and the cells resuspended in 100 μL DMSO to solubilize the formazan dye. Quantification was performed by measurements at 570 and 630 nm. The percent survival was determined by comparison with the control group (a volume of medium equal to that of the added DMSO was added to the test wells).

Voltage-Gated Sodium Channel (rNav1.4) Inhibition Assay

Assays followed previously published protocols.[48] In brief, rNav1.4 HEK 293 cells were plated in a 96 well plate at 50,000/cells per well, cultured overnight, and washed with Locke's buffer. Cells were incubated for 1 h at 37 °C with the Na$^+$ selective fluorescent dye, sodium-binding benzofuran isophthalate (SBFI), washed twice, and then supplemented with 2.5 mM probenecid. Cells were excited at 340 and 380 nm on a Flexstation II and Na$^+$-bound SBFI emission detected at 505 nm. Baseline was established and then 50 μL of cryptomaldamide solution was added in quadruplicate. Fluorescence was recorded for 5.5 minutes after the final addition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Newman DJ, Cragg GM. J Nat Prod. 2016; 79:629–661. [PubMed: 26852623]

2. Cragg, GM.Kingston, DGI., Newman, DM., editors. Anticancer Agents from Natural Products. 2. CRC Press/Taylor & Francis Group; Boca Raton, FL: 2012. xv + 751 pp. 16 × 26 cm. $179.96

3. Luo Y, Cobb RE, Zhao H. Current Opin Biotech. 2014; 30:230–237.

4. Ziemert N, Alanjary M, Weber T. Nat Prod Rep. 2016; 33:988–1005. [PubMed: 27272205]

5. Adpressa DA, Loesgen S. Chem Biodiversity. 2016; 13:253–259.

6. Gross H, Stockwell VO, Henkels MD, Nowak-Thompson B, Loper JE, Gerwick WH. Chem Biol. 2007; 14:53–63. [PubMed: 17254952]

7. Kersten RD, Yang YL, Xu Y, Cimermancic P, Nam SJ, Fenical W, Fischbach MA, Moore BS, Dorrestein PC. Nat Chem Bio. 2011; 7:794–802. [PubMed: 21983601]

8. Dashti Y, Grkovic T, Abdelmohsen UR, Hentschel U, Quinn RJ. Mar Drugs. 2014; 12:3046–3059. [PubMed: 24857962]

9. Yamanaka K, Reynolds KA, Kersten RD, Ryan KS, Gonzalez DJ, Nizet V, Dorrestein PC, Moore BS. Proc Natl Acad Sci U S A. 2014; 111:1957–1962. [PubMed: 24449899]

10. Li Y, Li Z, Yamanaka K, Xu Y, Zhang W, Vlamakis H, Kolter R, Moore BS, Qian PY. Sci Rep. 2015; 5:9383. [PubMed: 25807046]

11. Gerwick WH, Moore BS. Chem Biol. 2012; 19:85–98. [PubMed: 22284357]

12. Flahive, E., Srirangam, J. Anticancer Agents from Natural Products. 2. Cragg, GM.Kingston, DGI., Newman, DJ., editors. 2012. p. 263-289.

13. Yoo HD, Gerwick WH. J Nat Prod. 1995; 58:1961–19655.

14. Edwards DJ, Marquez BL, Nogle LM, McPhail K, Goeger DE, Roberts MA, Gerwick WH. Chem Biol. 2004; 11:817–33. [PubMed: 15217615]

15. Lee SU, Kwak HB, Pi SH, You HK, Byeon SR, Ying Y, Luesch H, Hong J, Kim SH. ACS Med Chem Lett. 2011; 2:248–251. [PubMed: 21666868]

16. Wu P, Cai W, Chen QY, Xu S, Yin R, Li Y, Zhang W, Luesch H. Org Lett. 2016; 18:5400–5403.

17. Paatero AO, Kellosalo J, Dunyak BM, Almaliti J, Gestwicki JE, Gerwick WH, Taunton J, Paavilainen VO. Cell Chem Biol. 2016; 23:561–566. [PubMed: 27203376]

18. Pereira AR, Kale AJ, Fenley AT, Byrum T, Debonsi HM, Gilson MK, Valeriote FA, Moore BS, Gerwick WH. ChemBioChem. 2012; 13:810–817. [PubMed: 22383253]

19. Choi H, Mascuch SJ, Villa FA, Byrum T, Teasdale ME, Smith JE, Preskitt LB, Rowley DC, Gerwick L, Gerwick WH. Chem Biol. 2012; 19:589–598. [PubMed: 22633410]

20. Choi H, Pereira AR, Cao Z, Shuman CF, Engene N, Byrum T, Matainaho T, Murray TF, Mangoni A, Gerwick WH. J Nat Prod. 2010; 73:1411–1421. [PubMed: 20687534]

21. Welker M, Erhard M. J Mass Spectrom. 2007; 42:1062–1068. [PubMed: 17605146]

22. Zhao Y, Gu Q, Morris-Natschke SL, Chen CH, Lee KH. J Med Chem. 2016; 59:9262–9268. [PubMed: 27676157]

23. Engene N, Rottacker EC, Kaštovský J, Byrum T, Choi H, Ellisman MH, Komarek J, Gerwick WH. Int J Syst Bacteriol. 2012; 62:1171–1178.

24. Talpir R, Benayahu Y, Kashman Y, Pannell L, Schleyer M. Tetrahedron Lett. 1994; 35:4453–4456.

25. Coleman JE, de Silva ED, Kong F, Andersen RJ, Allen T. Tetrahedron. 1995; 51:10653–10662.

26. Marquez BL, Watts KS, Yokochi A, Roberts M, Verdier-Pinard P, Jimenez JI, Hamel E, Scheuer PJ, Gerwick WH. J Nat Prod. 2002; 65:866–871. [PubMed: 12088429]

27. Esquenazi E, Jones AC, Byrum T, Dorrestein PC, Gerwick WH. Proc Natl Acad Sci U S A. 2011; 108:5226–5231. [PubMed: 21393570]

28. Munro MHG, Stroshane RM, Rinehart KL Jr. J Antibiot. 1982; 35:1331–1337. [PubMed: 6184357]

29. Hagen R, Roberts JD. J Am Chem Soc. 1969; 91:4504–4506.

30. Gamble WR, Durso NA, Fuller RW, Westergaard CK, Johnson TR, Sackett DL, Hamel E, Cardellina JH II. Bioorg Med Chem Lett. 1999; 7:1611–1615.

31. Marfey P. Carlsberg Res Commun. 1984; 49:591–596.

32. Coleman JE, Patrick BO, Andersen RJ, Rettig SJ. Acta Crystallographica, Section C. 1996; 52:1525–1527.

33. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. J Comput Biol. 2012; 19:455–477. [PubMed: 22506599]

34. Nikolenko S, Korobeynikov A, Alekseyev M. BMC Genomics. 2013; 14S1:S7.

35. Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH. Nat Biotechnol. 2013; 31:533–538. [PubMed: 23707974]

36. Leao T, Castelao G, Korobeynikov A, Monroe EA, Podell S, Glukhov E, Allen EE, Gerwick WH, Gerwick L. Proc Natl Acad Sci U S A. 2017; 114:3198–3203. [PubMed: 28265051]

37. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. Nucleic Acids Res. 2014; 42:D206–214. [PubMed: 24293654]

38. Weber T, Blin K, Duddela S, Krug D, Muller R, Kim HU, Lee SY, Bruccoleri R, Fischbach MA, Wohlleben W, Breitling R, Takano E, Medema MH. Nucleic Acids Res. 2015; 43:W237–243. [PubMed: 25948579]

39. Jones AC, Monroe EA, Podell S, Hess WR, Klages S, Esquenazi E, Niessen S, Hoover H, Rothmann M, Lasken RS, Yates JR III, Reinhardt R, Kube M, Burkart MD, Allen EE, Dorrestein PC, Gerwick WH, Gerwick L. Proc Natl Acad Sci U S A. 2011; 108:8815–8820. [PubMed: 21555588]

40. Puigbò P, Wolf YI, Koonin EV. J Biol. 2009; 8:59. [PubMed: 19594957]

41. Kellmann R, Mihali TK, Jeon YJ, Pickford R, Pomati F, Neilan BA. Appl Environ Microbiol. 2008; 74:4044–4053. [PubMed: 18487408]

42. Shigeki T, Yuko C, Keiichi K, Kazuo N, Yasukatsu O, Mari YY. Sci Rep. 2016; 6:20340. [PubMed: 26842222]

43. Muenchhoff J, Siddiqui KS, Poljak A, Raftery MJ, Barrow KD, Neilan BA. FEBS J. 2010; 277:3844–3860. [PubMed: 20718864]

44. Crews P, Farias JJ, Emrich R, Keifer PA. J Org Chem. 1994; 59:2932–2934.

45. Penzotti JL, Fozzard HA, Lipkind GM, Dudley SC Jr. Biophysical J. 1998; 75:2647–2657.

46. Fulmer GR, Miller AJM, Sherden NH, Gottlieb HE, Nudelman A, Stoltz BM, Bercaw JE, Goldberg KI. Organometallics. 2010; 29:2176–2179.

47. Mevers E, Liu WT, Engene N, Mohimani H, Byrum T, Pevzner PA, Dorrestein PC, Spadafora C, Gerwick WH. J Nat Prod. 2011; 74:928–936. [PubMed: 21488639]

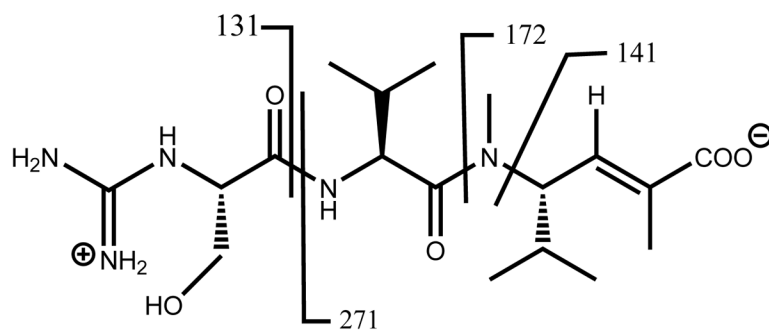48. Cao Z, Gerwick WH, Murray TF. BMC Neuroscience. 2010; 11:154. [PubMed: 21156065]
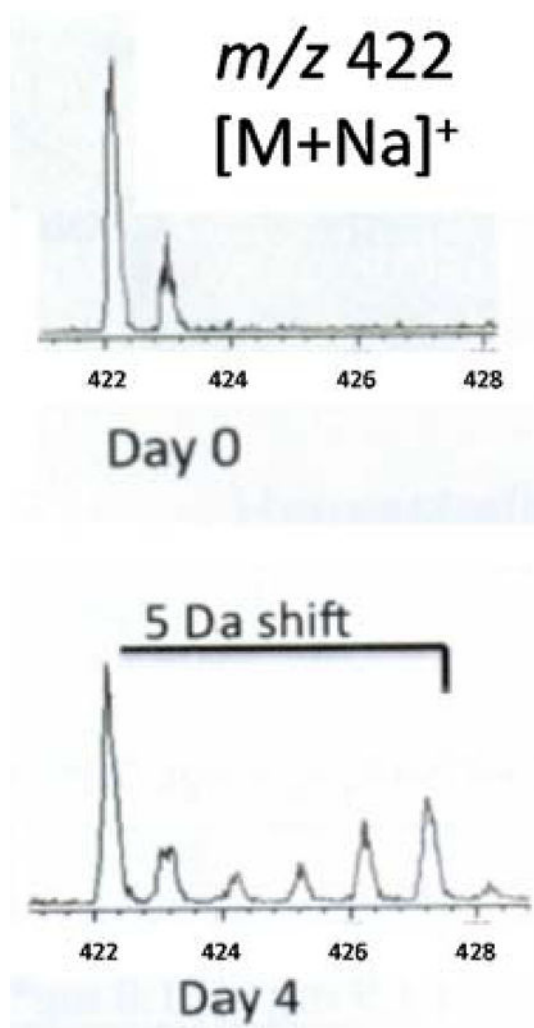
Figure 1.
Projection of mass range *m/z* 421–429 for the sodium adduct ion [M+Na]$^+$ for day 0 and day 4 showing a 5 Da increase from the incorporation of five ^{15}N-atoms into cryptomaldamide (**1**).
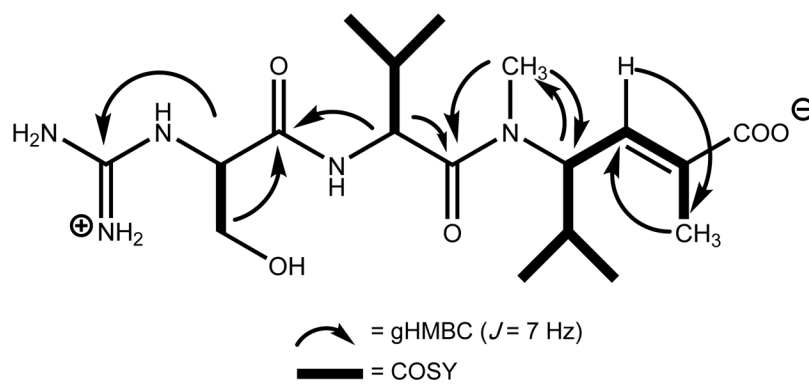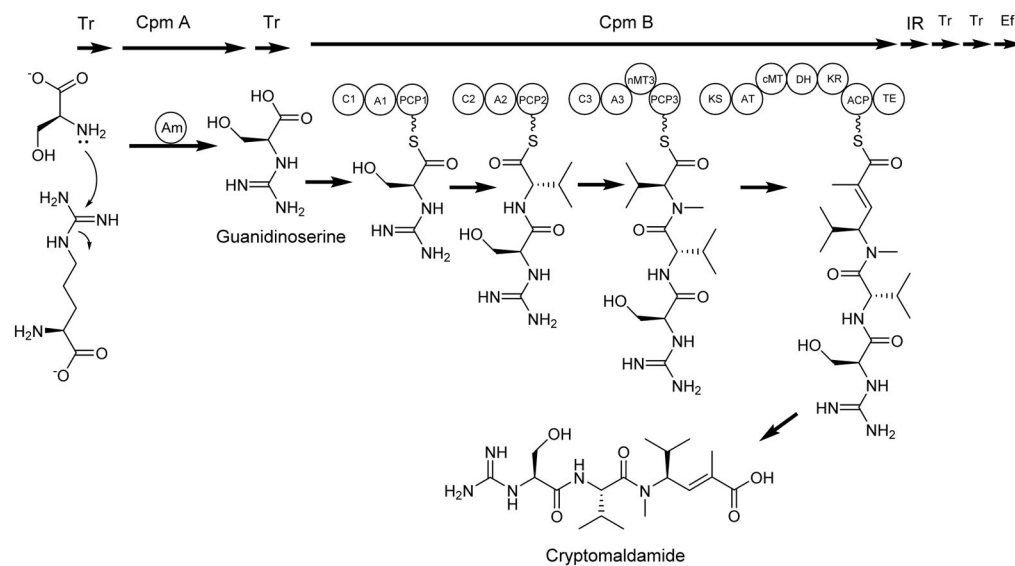
Figure 2.
Key COSY and HMBC correlations for cryptomaldamide (**1**).

Figure 3.
Fragmentations observed in the ESI MS/MS spectrum of cryptomaldamide (**1**).

Figure 4.
Proposed biosynthetic assembly of cryptomaldamide (**1**) from the Cpm gene cluster present in *Moorea producens* JHB (Tr = transposase, Am = amidinotransferase, C = condensation domain, A = adenylation domain, PCP = peptidyl carrier protein, nMT = *N*-methyl transferase, KS = ketosynthase, AT = acyl transferase, cMT = *C*-methyl transferase, DH = dehydratase, KR = ketoreductase, ACP = acyl carrier protein, TE = thioesterase, IR = integrase-resolvase, Ef = efflux transporter).

Table 1

NMR Spectroscopic Data for Cryptomaldamide (1) in D_2O (^1H 800 MHz, ^{13}C 200 MHz).

Position	δ_C, type	δ_H (*J* in Hz)	COSY	HMBC
1	177.7, C			3, 13
2	138.4, C			4, 13, 14
3	133.0, CH	6.21, dq (8.4, 1.2)	4, 13	1, 4, 13
4	58.7, CH	4.72, dd (9.8, 8.4)	3,14	2, 3, 6, 14, 15, 17
6	173.8, C			4, 7, 17
7	57.3, CH	4.45, d (8.5)	18	6, 9,18, 19, 20
9	170.5, C			7, 10, 11
10	56.6, CH	4.23, dd (5.9, 4.3)	11a, 11b	9, 11, 12
11a	62.4, CH_2	3.78, dd (12.0, 5.2)	11b, 10	9, 10
11b		3.82, dd (12.0,4.1)	11a, 10	9, 10
12	157.6, C			10
13	15.0, CH_3	1.68, d (1.2)	3	2, 3
14	30.1, CH	1.85, m	15, 16, 4	3, 15, 17
15	19.2, CH_3	0.65, d (6.7)	14	4, 14, 16
16	19.7, CH_3	0.76, d (6.6)	14	4, 14, 15
17	31.3, CH_3	2.95, s		4, 6
18	31.5, CH	1.95, m	19, 20, 7	7, 19, 20
19	19.3, CH_3	0.78, d (6.7)	18	7, 18, 20
20	19.6, CH_3	0.85, d (6.7)	18	7, 18, 19

Table 2

Predicted Enzymatic Activities in the Cryptomaldamide (*cpm*) Biosynthetic Gene Cluster and Their Closest Annotated Relatives.

Protein	Amino Acids	Proposed Function	Sequence Similarity (Protein, Origin)	Identity/Similarity	Accession Number
ORF18	91	transposase	transposase, *Anabaena cylindrica*	42%, 62%	WP_042466367.1
CpmA (ORF19)	**364**	**Amidinotransferase**	**amidinotransferase, *Dolichospermum circinale***	**58%, 73%**	**WP_051347453.1**
ORF24	120	transposase	transposase, *Anabaena cylindrica*	42%, 54%	WP_042466367.1
CpmB	**6439**	**NRPS/PKS**	**NRPS, *Rivularia sp.* PCC 7116**	**43%, 59%**	**WP_015122075.1**
NRPS_1	**21-1076**		**NRPS, *Anabaena cylindrica***	63%, 77%	WP_015217394.1
C_1	297	Condensation	NRPS, *Anabaena cylindrica*	63%, 78%	WP_015217394.1
A_1	403	Adenylation	NRPS, *Anabaena cylindrica*	73%, 85%	WP_015217394.1
PCP_1	67	PCP	NRPS, *Streptomyces sp. NBRC 109706*	56%, 78%	WP_062218182.1
NRPS_2	**1122-2154**		**mcyB, *Microcystis aeruginosa***	39%, 57%	BAA83993.1
C_2	297	Condensation	NRPS, *Nostoc punctiforme*	41%, 61%	WP_012409016.1
A_2	402	Adenylation	NRPS, amino acid adenylation enzyme, *Brevibacillus sp. BC25*	48%, 63%	EJL20544.1
PCP_2	69	PCP	NRPS, *Pedosphaera parvula*	54%, 80%	WP_007415652.1
NRPS_3	**2173-3620**		**non-ribosomal peptide synthase, *Nostoc sp.* 'Peltigera membranacea cyanobiont'**	41%, 58%	ADL59761.1
C_3	297	Condensation	NRPS, *Nostoc sp. PCC 7524*	48%, 61%	KKD37616.1
A_3	404	Adenylation	thioester reductase, *Nostoc punctiforme*	50%, 65%	WP_012412785.1
MT_3	219	N-methyltransferase	PuwA, *Cylindrospermum alatosporum CCALA 988*	51%, 66%	AIW82277.1
PCP_3	68	PCP	hybrid NRPS-PKS, *Paenibacillus tyrfis*	53%, 80%	WP_036675273.1
PKS	**3647-6431**		**PKS, Cystobacter sp. SBCb004**	**35%, 50%**	**ADH04682.1**
KS	425	KS	PKS, partial, *Oscillatoria sp. PCC 6506*	58%, 72%	ACJ46057.1
AT	306	AT	type I PKS, *Crocosphaera watsonii*	47%, 65%	WP_021836604.1
MT	192	MT	PKS, *Cystobacter sp. SBCb004*	46%, 63%	ADH04682.1
KR	152	KR	TubF protein, *Angiococcus disciformis*	48%, 70%	CAF05651.1
DH	211	DH	PKS, *Cystobacter sp. SBCb004*	39%, 51%	ADH04682.1
ACP	69	ACP	NRPS, *Kibdelosporangium aridum*	56%, 72%	WP_051896182.1
TE	269	TE	amino acid adenylation enzyme/thioester reductase family protein, *Cylindrospermum stagnale PCC 7417*	37%, 55%	AFZ26071.1
ORF26	162	integrase-resolvase	putative site-specific integrase-resolvase, *Moorea producens 3L*	89%, 93%	EGJ33692.1
ORF27	277	transposase	transposase, *Limnoraphis robusta*	82%, 88%	WP_049560375.1
ORF28	79	transposase	transposase, IS605 OrfB family, central region, *Moorea producens 3L*	81%, 84%	EGJ34035.1

Protein	Amino Acids	Proposed Function	Sequence Similarity (Protein, Origin)	Identity/Similarity	Accession Number
ORF29	497	transporter	MATE family efflux transporter, *Kamptonema*	47%, 67%	WP_007353095.1