UC San Diego
UC San Diego Previously Published Works

Title
Pagoamide A, a Cyclic Depsipeptide Isolated from a Cultured Marine Chlorophyte, Derbesia sp., Using MS/MS-Based Molecular Networking

Permalink
https://escholarship.org/uc/item/3fv3s77d

Journal
Journal of Natural Products, 83(3)

ISSN
0163-3864

Authors
Li, Yueying
Yu, Hao-Bing
Zhang, Yi
et al.

Publication Date
2020-03-27

DOI
10.1021/acs.jnatprod.9b01019

Peer reviewed

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

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access
Author manuscript
J Nat Prod. Author manuscript; available in PMC 2021 March 27.

Published in final edited form as:
J Nat Prod. 2020 March 27; 83(3): 617–625. doi:10.1021/acs.jnatprod.9b01019.

Pagoamide A, A Cyclic Depsipeptide Isolated from a Cultured Marine Chlorophyte, *Derbesia* sp., using MS/MS Based Molecular Networking

Yueying Li[†,‡], Hao-Bing Yu[‡,§], Yi Zhang[‡,⊥], Tiago Leao[‡], Evgenia Glukhov[‡], Marsha L. Pierce[‖], Chen Zhang[‡,∇], Hyunwoo Kim[‡], Huanru Henry Mao[∇], Fang Fang[‡], Garrison W. Cottrell[∇], Thomas F. Murray[‖], Lena Gerwick[‡], Huashi Guan[†], William H. Gerwick[‡,O,*]

[†]Key Laboratory of Marine Drugs, Chinese Ministry of Education, School of Medicine and Pharmacy, Ocean University of China, Qingdao 266003, People's Republic of China

[‡]University of California, San Diego, La Jolla, California 92093, United States

[§]Department of Biochemistry and Molecular Biology, College of Basic Medical Sciences, Second Military Medical University, Shanghai 200433, People's Republic of China

[⊥]College of Food Science and Technology, Guangdong Ocean University, Zhanjiang 524088, People's Republic of China

[‖]Department of Pharmacology, School of Medicine, Creighton University, Omaha, Nebraska 68178, United States

[∇]Department of Computer Sciences and Engineering, University of California, San Diego, La Jolla, California 92093, United States

[O]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California 92093, United States

Abstract

A thiazole-containing cyclic depsipeptide with 11 amino acid residues, named pagoamide A (**1**), was isolated from laboratory cultures of a marine Chlorophyte, *Derbesia sp*. This green algal sample was collected from America Samoa and pagoamide A was isolated using guidance by MS/MS based molecular networking. Cultures were grown in a light and temperature-controlled environment, and harvested after several months of growth. The planar structure of pagoamide A (**1**) was characterized by detailed 1D- and 2D-NMR experiments along with MS and UV analysis. The absolute configurations of its amino acid residues were determined by advanced Marfey's analysis following chemical hydrolysis and hydrazinolysis reactions. Two of the residues in

[*]To whom correspondence should be addressed. Tel: (858)-534-0578. wgerwick@ucsd.edu.

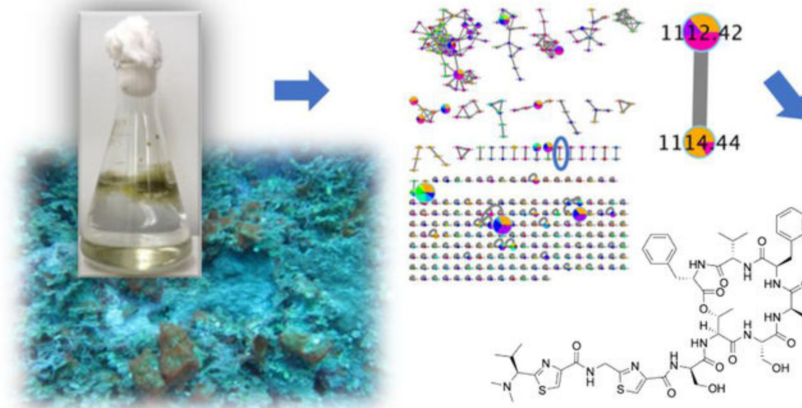
Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI:
MS/MS, ESIMS, HRESIMS, 1D/2D NMR spectra of compound **1**. LCMS chromatograms of Marfey's analysis hydrolysis, and derivatization products of **1** with and without ozonolysis, along with derivatized amino acid standards. Marfey's analysis hydrolysis, hydrazinolysis and derivatization products of **1**, along with derivatized amino acid standards. The MS^2-based molecular networking figure for the crude and VLC fractions, bioassay screening result.

The authors declare no competing financial interest.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

pagoamide A (**1**), phenylalanine and serine, each occurred twice in the molecule, once in the D- and once in the L-configuration. The biosynthetic origin of pagoamide A (**1**) was considered in light of other natural products investigations with coenocytic green algae.

Graphical Abstract



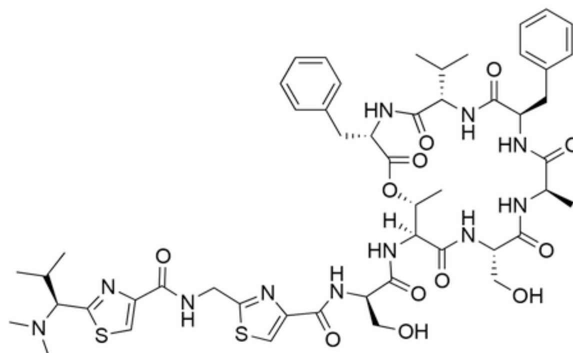
The green algal phylum Chlorophyta is broadly represented in the marine environment and has been a rich source of novel natural products with 400 published compounds and with a wide variety of biological activities. Interestingly, a number of these are of a peptide constitution, such as the the depsipeptide kahalalide F, first reported in 1993 from a Hawaiian *Bryopsis* species.[1] Subsequently a series of articles clarified stereochemical features of this molecule as well as reported on a number of related compounds, kahalalides A-Z2.[2–12] Most cyclic kahalalides possess unique structural features comprised of a macrocyclic region that forms from an ester bond between the C-terminus and the β-hydroxyl group of a threonine residue. Further, they possess a terminal polyketide moiety or a branched acyclic peptide. A recent study has shown that the kahalalides are produced by an intracellular bacterium of the Flavobacteriia family through a Non-Ribosomal Peptide Synthetase (NRPS) pathway.[13] These and related peptide molecules show interesting and highly diverse biological activities including cytotoxic and antitumor activity, antimicrobial activity, antileishmanial, immunosuppressive properties and protease inhibitory activity.[14–20]

Derbesia, a member of the Chlorophyta, is a siphonous filamentous green alga forming filaments approximately 100 μm in diameter.[21–23] The organism was collected from shallow coastal waters surrounding America Samoa and successfully adapted to laboratory culture for the purpose of exploring its natural products profile.[24,25] Previously, only two natural products have been isolated and reported from this genus, mebamamides A and B.[26] These latter two compounds possessed a similar macrocyclic arrangement to the kahalalides, however, neither were cytotoxic but one (B) did show weak cellular differentiation-inducing activity.

In the current study, we discovered a new cyclic non-toxic depsipeptide from cultures of a marine *Derbesia sp*. Isolation of this novel compound was facilitated and guided by using

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

MS/MS-based molecular networking as enabled through the Global Natural Product Social (GNPS) platform (https://gnps.ucsd.edu)[27,28] as well as a calcium influx/oscillation assay. This analysis revealed a small cluster comprised of two related compounds at *m/z* 1112 and 1114 that had no matches in the GNPS database. Subsequently, purification of the *m/z* 1112 metabolite afforded a novel cyclic depsipeptide named here as pagoamide A (**1**). The planar structure of compound **1** was determined by MS, UV and detailed 1D- and 2D-NMR analyses. Absolute configurations of the amino acid moieties were determined by advanced Marfey's analysis following chemical hydrolysis or hydrazinolysis reactions as well as DFT calculations in combination with ROESY results. While the extract and pagoamide-containing fractions showed inhibition activity in the calcium oscillation assay, pure pagoamide A was inactive, indicating that the bioactive compounds were other components in these fractions. The extract, pagoamide A-containing fractions and pure pagoamide A were also evaluated for H-460 human lung cancer cell cytotoxicity; however, all of these materials were inactive at the maximum concentrations tested (10 μg/mL for impure materials, 30 μM for pure pagoamide A).



Results and Discussion

A culture sample of a *Derbesia* sp. green macroalga (Figure 1) was obtained in 12–18 m water depth by scuba diving in Fagatele Bay, America Samoa and codified as ASF-14JUL14–1A after separation and purification from co-collect cyanobacteria under experimental environment. Purification of the living collection was achieved using standard methods in microbiology, and cultures were maintained under a light and temperature-controlled environment in Erlenmeyer flasks with salt water BG11 media.[29–31] Morphologically, this cultured material was consistent with a *Derbesia* species (Figure 1); unfortunately, DNA sequencing of the metagenomic DNA failed to capture a phylogenetically informative sequence. After a number of scale-up experiments, the culture was harvested and extracted using CH_2Cl_2/MeOH (2:1). The crude extract was washed with distilled H_2O to remove polysaccharides and salts, and afforded both aqueous and organic extracts. The organic extract was separated using silica gel vacuum liquid chromatography (VLC) to obtain nine fractions. These were subjected to untargeted LC-MS/MS analysis as well as evaluation in a calcium influx/oscillation assay.[32–34] The biological assay indicated that VLC fractions H and I potently suppressed the amplitude of spontaneous calcium oscillations in neocortical neurons, suggesting the presence of neuromodulatory secondary metabolites. By detailed comparison of the LC-MS/MS and molecular networking outcomes

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

with the biological assay results, a small cluster of compounds was targeted for isolation with parent masses of *m/z* 1112 and 1114. The MS/MS fragmentations suggested these compounds were peptides, and ensuing reversed-phase silica gel column chromatography followed by HPLC was successful in the isolation of 3.2 mg of pure pagoamide A (**1**). Unfortunately, upon isolation of pure pagoamide A, it was found to be inactive in the calcium oscillation assay, and thus other trace components of these fractions must constitute the active materials.

Pagoamide A (**1**) was obtained as a white powder and the molecular formula was determined by HRESIMS as $C_{51}H_{67}N_{11}O_{12}S_2$ from observation of the sodium adduct ion peak at *m/z* 1112.4292. Of the 24 degrees of unsaturation inherent in this molecular formula, nine were revealed as amide-type carbonyls from ^{13}C NMR resonances at δ_C 168–174. The ^{1}H NMR, ^{13}C NMR and MS/MS fragmentations exhibited characteristic profiles for a peptide-type structure (Table 1, Figures 2 and 3). The ^{1}H NMR spectrum contained two signals at δ_H 8.10 (H-39) and 8.31 (H-44) with associated ^{13}C NMR resonances at δ_C 124.2 and δ_C 124.4, and these could be assigned to the methine groups of two substituted thiazole rings; this was confirmed by diagnostic HMBC correlations and accounted for six additional degrees of unsaturation (Table 1).[35–37] One Gly-Cys thiazole ring was confirmed by Marfey's analyses either with or without ozonolysis. Two phenyl rings, accounting for 8 more degrees of unsaturation, were revealed by 10 protons appearing between δ_H 7.00–7.30. The remaining unaccounted degree of unsaturation was thus deduced to be due to an overall macrocyclic ring structure.

An intense singlet at δ_H 2.15 (H_3-50, H_3-51) integrated for six protons, and suggested the presence of two *N*-methyl groups in pagoamide A. These methyl protons and an α-methine carbon at δ_c 71.9 (C-46) were correlated in the HMBC spectrum. This α-methine proton was coupled to a β-methine proton at δ_H 2.26 (H-47), which in turn was coupled to two γ-methyl groups at δ_H 0.75 and 0.99, respectively (H_3-48, H_3-49), thus constructing a terminating residue as *N*,*N*-dimethyl valine. The correlations between eight amide NH protons (Table 1) and their corresponding α-methine protons were determined by the COSY and HMBC experiments, and these spin systems could be extended to the amino acid side chains in each case. This analysis revealed that in addition to the two cysteine-derived thiazoles and one *N*,*N*-dimethyl valine, pagoamide A possessed two serines, two phenylalanines, one valine, one alanine, one glycine and one threonine. The sequence of residues in the peptide was determined by HMBC correlations from the amide NH protons to the carbonyl equivalents in the adjacent residue, or in some cases, from β-protons to carbonyls (Table 1, Figure 2), resulting in the series *N*,*N*-dimethyl-valine-cysteine thiazole, glycine-cysteine thiazole, serine, threonine, serine, alanine, phenylalanine, valine and phenylalanine. Finally, the phenylalanine carboxyl terminus could be connected as an ester to the β-hydroxy group of threonine through observation of a key HMBC correlation. Thus, the planar structure of pagoamide A was revealed as a new cyclic despipeptide formed from 11 amino acid moieties (Figure 2).

The configurations of the amino acid moieties in pagoamide A (**1**) were investigated using advanced Marfey's analysis, and involved ozonolysis, hydrolysis, chemical derivatization with D-FDAA, and LC-MS comparison against available standards.[37–40] The configuration

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of the *N,N*-dimethyl valine residue was determined as L by comparing the hydrolyzed product with synthetic D/L *N,N*-dimethyl valine standards by chiral-phase HPLC [(D)-penicillamine stationary phase]. The remainder of the hydrolyzed products was subjected to chemical derivatization with D-FDAA and comparison with the D-DAA derivatives of the authentic amino acid standards. The valine and threonine residues were of L-configuration whereas the alanine was D-configured. However, the two serine and two phenylalanine residues were each comprised of one D and one L configured residue, a result that provided four possibilities for the absolute configuration of compound **1**. Subsequently, mild base hydrolysis of pagoamide A provided the acyclic peptide with phenylalanine-1 as the carboxy terminus.[6] A hydrazinolysis reaction was employed to selectively release this terminal phenylalanine residue,[11] which was then subjected to Marfey's analysis to reveal its L configuration.[38,39] This result necessarily defined the phenylalanine-2 residue to be D -configured (C15-C23) which leaves two possible absolute configurations for compound **1**.

Conformational analysis using Discovery Studio 2.5 gave 8 conformers in total [5 for structure A and 3 for structure B (Figure S24)]. Reoptimization was carried out using DFT calculations on the Gaussian 09 program for the 8 conformers. In the calculated lowest energy conformers, conformer A1and conformer B1 showed Boltzmann distributions with Gibbs free energies of more than 99.99 percent. (Figure 4A) The other 6 conformers were A2 (8.505e-15%), A3 (3.621e-17%), A4 (2.737e-20%), A5 (5.391e-24%), B2 (0.004%), B3 (1.300e-13%), each with very minor Boltzmann distributions (Figure S24). It was observed that the side chain in A1 extends out from the macrocycle versus partially folding back over the macrocycle as in B1 (Figure 4A). Discrimination between these two possibilities (A1 and B1, Figure 4B, 4C) was achieved using ROESY correlations. The ROESY correlations observed between the α-proton of Ser-2 (H-35), β-protons of Ser-2 (H-36) and NH of the Phe-1 residue as well as the correlations between the α-proton of Val (H-11), NH of the Val and the NH of the Thr were consistent with the stable conformer wherein serine-1 is L and serine-2 is D (B1, Figure 4B); conversely, these long range correlations were inconsistent with the major conformer where serine-1 is D and serine-2 in the side chain is L (A1, Figure 4B). Consistent with this result, ROESY correlations were observed between H-28 of Ser-1 and the NH of both the Ala and the Phe-2 residues, this was only consistent with structure B1 wherein Ser-1 is of L-configuration (B1, Figure 4C).

Pagoamide A (**1**) possesses two thiazole ring moieties that are predicted to derive from the cyclization between cysteine and *N,N*-dimethyl valine in one case, and with glycine in a second instance. It was interesting that only the terminal valine moiety was *N*-methylated, as *N*-methylation is frequently observed in marine-derived cyclic and acyclic peptides.[41,42] Additionally, three out of 11 amino acids were of D-configuration. These types of structural modifications are somewhat more consistent with a biogenesis through the ribosomally synthesized and post-translationally modified peptide (RiPPs) pathway rather than a nonribosomal peptide synthetase (NRPS) route.[43,44] Partial support for this comes from preliminary genome sequence data for this *Derbesia sp.* (Illumina HiSeq 400 assembled with metaSPAdes 3.12). AntiSmash analysis of the biosynthetic gene clusters in this eukaryotic alga failed to identify any NPRS pathways that could potentially encode for pagoamide A (**1**) biosynthesis. However, as noted above, the somewhat related kahalalides were recently

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

shown to be produced via NRPS metabolism by an endosymbiotic bacterium in another coenocytic green alga, a *Bryopsis* sp.[13]

To explore this aspect in somewhat greater depth, we applied our recently described Small Molecule Automatic Recognition Technology (SMART) program which is a machine learning tool that recognizes structurally related compounds on the basis of their HSQC spectra.[45] The pagoamide A HSQC image was evaluated in SMART that had been trained with 30,000 marine and terrestrial natural products, and provided close hits for echinomycin A, bistratamide F, didmolamide A, bistratamide E, ulongamide A, patellamide E, banyascyclamide A, xentrivalpeptide A and guineamide B (Figure 5). Ulongamide A, banyascyclamide A and guineamide B are metabolites of free living marine cyanobacteria whereas bistratamides E, F, didmolamide A and patellamide E are produced by cyanobacteria living within an ascidian.[37,46–50] Echinomycin A and xentrivalpeptide A are metabolites reported from actinomycetes and proteobacteria, respectively.[51,52] The dominant feature being recognized by the SMART 2.0 algorithm appears to be that the molecule is a cyclic peptide with thiazole/oxazole and aromatic ring structures. This analysis provides some additional support for the hypothesis that pagoamide A (**1**) is the result of prokaryotic, possibly symbiotic, metabolism within the context of a eukaryotic host. Nevertheless, additional investigations are needed to confirm the origin of pagoamide A **(1)** in this cultured *Derbesia* sp.

Conclusions

In this work, we report the isolation and structure determination of the cyclic depsipeptide pagoamide A (**1**) from a cultured *Derbesia sp*. marine Chlorophyte. This finding expands the scope of secondary metabolites to be described from this under-explored genus, previously only investigated for its lipids, carotenoids and in one case, two structurally unique cyclic peptides, mebamamides A and B.[26] Detailed cultivation experiments were performed to obtain monocultures suitable for long-term scale-up experiments, and this enabled the accumulation of sufficient biomass for chemical investigations. MS/MS-based molecular networking facilitated the efficient isolation of pagoamide A through its detection in the various repeat cultures and sub-fractions produced during the chromatographic process. This was complemented by calcium influx/oscillation and cytotoxicity assays that helped to guide the isolation and purification of this new metabolite. By integrating multiple analytical methods and synthetic reaction procedures, the complete stereostructure of pagoamide A (**1**) was fully elucidated.

Relatively large and complex natural products such as pagoamide A impose significant challenges in structure elucidation due to NMR signal degeneracy and a multitude of stereogenic centers that require assignment. These difficulties are compounded when dealing with relatively small amounts of the natural product material. And while laboratory cultivation can in principle overcome the small quantity issue, it poses its own unique problems as oftentimes the cultured organism grows slowly requiring months for scale-up.

However, in some cases, associated or intracellular bacterial symbionts have been found to be responsible for the production of bioactive compounds associated with macroorganisms

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(e.g. algae, sponges, tunicates), and thus either through microbial culture or genetic methods, other production strategies are possible. For example, the cyclic peptide kahalalide F, originally isolated from a coenocytic *Bryopsis* sp. green alga. as well as from a nudibranch that feeds on the alga, *Elysia rufescens*, was shown to derive from an NRPS pathway found within a bacterium living within the cells of the alga.[13] This severely genome-reduced bacterium was deduced to be of the class Flavobacteriia, and given the name *Candidatus Endobryopsis kahalalidefaciens*. Given the related structures of pagoamide A and kahalalide F in the sense that both are cyclic peptides comprised of D and L amino acids (kahalalides H, J, P and Q possess both D- and L-phenylalanine residues), it may point to a general trend of intracellular bacterial symbionts with significant natural product biosynthetic capacity associated with coenocytic algae. These as well as questions concerning other bioactive constituents of this cultured *Derbesia* sp. (e.g. those causing modulation of calcium oscillations) are the focus of ongoing investigations.

Experimental Section

General Experimental Procedures.

Optical rotations were measured in MeOH with a JASCO P-2000 polarimeter operated at 25°C. Ultraviolet and visible (UV–Vis) spectra were recorded using a Beckman Coulter DU 800 spectrophotometer. All NMR data except ^{13}C NMR spectra were obtained from a JEOL ECZ 500 MHz spectrometer with a 3 mm detection probe. The ^{13}C NMR spectrum was recorded with a Varian VX 500 MHz spectrometer with a Varian XSens 2 channel (^{1}H and ^{13}C) NMR cold probe optimized for direct observe ^{13}C NMR. The NMR solvent was DMSO-d_6 purchased from Cambridge Isotope Laboratories, Inc. with 99.8% deuterium and using the solvent peaks for the internal standard (δ_H 2.50 and δ_C 39.1). LCMS data were obtained with a Thermo Finnigan LCQ Advantage Max mass spectrometer coupled to a Thermo Finnigan Surveyor HPLC system with a Kinetex C18 100 × 4.6 mm × 5 μm analytical column (Phenomenex) (monitoring 200–600 nm and *m/z* 200–1600 in positive ion mode). HR-ESI-MS data were recorded using an Agilent 6530 Accurate-Mass QTOF mass spectrometer in positive ion mode. Semi-preparative HPLC purification was performed on a Thermo Dionex UltiMate 3000 HPLC system with a YMC-Triart C18 (250 mm × 10 mm × 5 μm) column. Analytical HPLC was performed on a Thermo Dionex UltiMate 3000 HPLC system with a Phenomenex Chirex 3126 D-Penicillamine column (250 mm × 4.6 mm, 5 μm). HPLC-grade and LC-MS grade CH_3CN were purchased from Thermo Fisher Scientific. HPLC-grade H_2O was obtained by filtration using a Milli-Q Direct water purification system (Millipore). Optical Density (OD) values were measured on ThermoElectron Multiskan Ascent plate reader.

Organism Collection, Identification and Cultivation.

The sample was collected from 12–18 m water depth in Fagatele Bay, American Samoa (S 14° 21’ 47.714”, W 170° 45’ 43.902” on July 14, 2014). Cultures of a *Symploca sp.* marine cyanobacterium were initially collected, however, the cyanobacterium died leaving a fine siphonous green unialga alive in the culture flask, codified as ASF14JUL14–1A. Based on its microscopic morphology, this sample was identified as a *Derbesia sp.*[21,22,53,54] No attempt was made to produce axenic cultures of this cultured Chlorophyte, however, there

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

was no visible evidence of microbial growth. The cultures were initially grown in 12 × 250 mL autoclaved flasks with SWBG11 media under 27.0–27.6°C and 16h/8h light/dark cycle at 400–650 lx condition with semi-continuous dilution.[24,25,29] After 2-weeks of growth, the cultures were transferred into autoclaved 2 L flasks for another one month until they fully occupied the culture volume. The culture was harvested by filtration from the SWBG11 culture media and lyophilized overnight before grinding into a fine powder.

DNA Extraction and Sequencing.

A DNA sample was obtained from live cultures of *Derbesia sp*. by carefully grinding freshly harvested filaments using a mortar and pestle under liquid nitrogen. DNA was extracted using the QIAGEN Bacterial Genomic DNA Extraction Kit using the standard kit protocol. The library was prepared with the miniaturized version of the KAPA HyperPlus Kit, followed by quantification using the PicoGreen ds DNA assay. The metagenomic DNA was sequenced with Illumina HiSeq 4000 and subsequently assembled with metaSPAdes 3.12.

Extraction and Separation.

The scale-up cultivation of the *Derbesia* sp. was repeated eight times, and for each, a separate extraction and isolation procedure was performed. The freeze-dried and powdered biomass was exhaustively extracted in 2:1 CH_2Cl_2/MeOH (six times). The extracts (#2245EX) were dried under vacuum and then washed with DI H_2O (3x) to remove residual salt (#2245AQ), affording 1.6 g of organic portion (#2245CR). The organic portion was subjected to VLC separation over grade H silica gel. The column was eluted to provide nine sub-fractions: 2245CR_A (100% hexanes, 64.5 mg), 2245CR_B (10% EtOAc/ hexanes, 107.6 mg), 2245CR_C (20% EtOAc/hexanes, 316.4 mg), 2245CR_D (40% EtOAc/hexanes, 185.1 mg), 2245CR_E (60% EtOAc/hexanes, 32.4 mg), 2245CR_F (80% EtOAc/hexanes, 45.2 mg), 2245CR_G (100% EtOAc, 13.7 mg), 2245CR_H (25% MeOH/EtOAc, 507.5 mg), 2245CR_I (100% MeOH, 211.3 mg). Fractions 2245CR_H and 2245CR_I were combined and then applied to a 5000 mg/20 mL Gracepure C18 max SPE cartridge to produce seven sub-fractions: 2245CR_HI_1(100% H_2O, 263.2 mg), 2245CR_HI_2 (20% CH_3CN/H_2O, 20.6 mg), 2245CR_HI_3 (40% CH_3CN/H_2O, 4.9 mg), 2245CR_HI_4 (60% CH_3CN/H_2O, 82.6 mg), 2245CR_HI_5 (80% CH_3CN/H_2O, 76.7 mg), 2245CR_HI_6 (100% CH_3CN, 56.2 mg) and 2245CR_HI_7 (2:1 CH_2Cl_2/MeOH, 187.6 mg). Accordingly, 2245CR_HI_2 and 2245CR_HI_3 were combined and further purified by semi-preparative RP-HPLC using a YMC-Triart C18 (250 mm × 10 mm × 5 μm) column eluted with mixtures of H_2O with 0.1% formic acid and CH_3CN with 0.1 formic acid using the following gradient: 10% CH_3CN/H_2O for 5 min, 10% CH_3CN/H_2O to 99% CH_3CN /H_2O in 15 min, 99% CH_3CN /H_2O for 5 min, 99% CH_3CN /H_2O back to 10% CH_3CN /H_2O in 1 min and then maintained in 10% CH_3CN /H_2O for 4 min at the flow rate of 2 mL/min. The aqueous portion (2245AQ) was subjected to semi-preparative RP-HPLC using the same conditions to accumulate small additional amounts of pagoamide A (**1**), eluting at 15.2 min and yielding 3.2 mg in total from all fractions.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Molecular Networking.

The organic portion (2245CR) and fractions 2245CR_A–I were analyzed by LCMS/MS. Samples were dissolved in MeOH at 1 mg/mL and then filtered through a Gracepure C18 SPE cartridge before LCMS analysis. A 10 μL aliquot of each sample was injected and eluted with a gradient of H_2O containing 0.1% formic acid and CH_3CN containing 0.1% formic acid with a gradient method as follows: 30% CH_3CN/H_2O for 5 min, 30% CH_3CN/H_2O to 99% in 17 min, held at 99% CH_3CN/H_2O for 3 min, then 99% CH_3CN /H_2O to 30% CH_3CN/H_2O in 1 min, and finally held at 30% CH_3CN/H_2O for 4 min with the flow rate of 0.7 mL/min. Mass spectra were recorded in positive ESI mode (*m/z* 200–1600) and with an automated full dependent MS/MS scan enabled. All chromatograms were converted digitally to .mzXML files using MSConvert software (www.proteowizard.sourceforge.net). The organic extract (#2245CR) and the fractions (#2245CR_A-I) were submitted to GNPS molecular networking for dereplication.[27,28,55] The node and table file from GNPS were inputted into Cytoscape software to generate a molecular networking. The cluster containing the sodiated pagoamide A ion at *m/z* 1112 was used to direct further purification.

Pagoamide A (1): white powder; [α]25 D +5.5 (*c* 0.10, MeOH); UV (MeOH) *λmax* (log ε) 215 (4.37), 220 (4.39), 234 (4.59); ^{1}H and ^{13}C NMR data, Table 1; HRESIMS *m/z* 1112.4292 $[M + Na]^+$ (calcd for $C_{51}H_{67}N_{11}O_{12}S_2Na^+$, 1112.4304).

Synthesis of *N,N*-Dimethyl-valine Standards.

D-Valine (1.00 g, 8.541 mmol) was dissolved in DI H_2O (45 mL) and then added formaldehyde (37% w/w, 1.4 mL, 17.1 mmol) and 5% Pd-C (1.0 g). The reaction mixture was hydrogenated at rt for 5 days under H_2 gas, and then reflux for 30 min. The mixture was filtered through a bed of Celite and the filtrate dried *in vacuo*. The product was dissolved in a 7:1 mixture of EtOH/H_2O and evaporated to dryness *in vacuo*. A sample of L-Val was subjected to the same reaction conditions. The identities of both products were confirmed by comparing their ^{1}H NMR features with literature values.[56,57]

N,N-dimethyl-L-valine white powder; ^{1}H NMR (500 MHz, DMSO-d_6 TMS) δ2.65 (d, *J* = 9.4 Hz, 1H), 2.28 (s, 6H), 1.90 (m, 1H), 0.92 (d, *J* = 6.7 Hz, 2H), 0.83 (d, *J* = 6.5 Hz, 2H).

N,N-dimethyl-D-valine white powder; ^{1}H NMR (500 MHz, DMSO-d6 TMS) δ2.64 (d, *J* = 9.4 Hz, 1H), 2.28 (s, 6H), 1.90 (m, 1H), 0.92 (d, *J* = 6.8 Hz, 2H), 0.83 (d, *J* = 7.0 Hz, 2H).

Chemical Degradation of Pagoamide A (1) and Absolute Configuration Analysis.

A 0.1 mg sample of pagoamide A (**1**) was dissolved in 1 mL CH_2Cl_2 and then treated with ozone for 10 min at rt.[37,40] Next, the product was subjected to acid hydrolysis in a sealed vial at 110°C in with 200 μL of 6 M HCl for 16 h. The reaction mixture was cooled to rt and evaporated to dryness under a stream of N_2. A 20 μL aliquot of this hydrolysis product as well as the synthetic *N,N*-dimethyl-valine standards were injected onto an analytical HPLC system using a Phenomenex Chirex 3126 (D)-Penicillamine column (250 mm × 4.6 mm with 5 μm particles and 110 Å pore size) eluting with 2 M $CuSO_4$ for 15 min at a flow rate of 2 mL/min. The remainder of the hydrolysis product and both D- and L- amino acid standards

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(serine, threonine, alanine, phenylalanine and valine) were treated with 100 μL of 1 M $NaHCO_3$ and 200 μL of D-FDAA acetone solution (1.0%) at 40°C for 1 h and then neutralized with 100 μL of 1 M HCl to obtain the D-DAA derivatives. The D-DAA derivatives were then evaporated to dryness under N_2 and the products dissolved in 1 mL of MeOH and filtered over C18 SPE cartridges before LCMS analysis. A 10 μL aliquot was taken and injected into an analytical LCMS to compare against the derivatized authentic amino acid standards.[38,39] Retention times for the D-DAA derivatized amino acid standards were as follows: L-Ser (2.99 min), D-Ser (2.97 min), L-Thr (4.07 min), D-Thr (3.24 min), L-*allo*-Thr (3.67 min), D-*allo*-Thr (3.37 min), L-Ala (6.12 min), D-Ala (4.71 min), L-Phe (13.54 min), D-Phe (12.37 min), L-Val (12.13 min), and D-Val (9.52 min).

Another portion of pagoamide A (0.1 mg) was treated with the same acid hydrolysis and derivatization protocol described above, but without the prior ozonolysis step. The retention time of the D-DAA derivatized Gly-Cys moiety was 5.99 min and showed a *m/z* 410 (Gly-Cys Thiazole)(Supporting Information S23).

A third portion of pagoamide A (0.1 mg) was hydrolyzed in a sealed vial with 0.1 N NaOH in aqueous MeOH (50 μL) at 37°C for 90 min. The partially hydrolyzed product was added to 1 mg of Amberlite GC50 under argon, followed by 40 μL of distilled hydrazine. The reaction mixture was heated at 80°C for 5 days. After cooling to rt, the reaction mixture as well as both the D- and L- authentic phenylalanine amino acid standards were treated with 100 μL of 1 M $NaHCO_3$ and 200 μL of D-FDAA acetone solution (1.0%) at 40°C for 1 h and then neutralized with 100 μL of 1 M HCl to obtain the D-DAA derivatives. A 20 μL aliquot of this product and the phenylalanine standards were injected into an analytical HPLC coupled with a YMC-Triart C18 column (250mm × 10.0 mm, 5 μm) and their retention times compared as follows: L-Phe (20.04 min), D-Phe (19.56 min).[6,11,38,39]

Conformer and Stereochemical Calculations.

The conformational analyses were carried out in the Discovery Studio 2.5 software package. The results showed the 8 lowest energy conformers for 2 proposed structures (Figure S24) whose relative energies were within 10.0 kcal/mol. Subsequently, the conformers were re-optimized using DFT at the B3LYP/6–31+G(d,p) level in the gas phase by the GAUSSIAN 09 program. The B3LYP/6–31+G(d,p) harmonic vibrational frequencies were also calculated to confirm their stability.

Biological Evaluation.

Animal care and handling complied with protocols approved by the Creighton University Institutional Animal Care and Use Committee and employed measures to minimize pain and discomfort (0801.3). Swiss-Webster gravid dams were euthanized by CO_2 asphyxiation and embryonic day 16 mice were removed under sterile conditions. Cerebrocortices were dissected, stripped of meninges, minced, and incubated with trypsin for 25 min at 37 °C. Cells were dissociated by trituration in isolation buffer containing soybean trypsin inhibitor and DNase, centrifuged, and resuspended in Eagle's minimal essential medium with Earle's salt (MEM) and supplemented with 1 mM L-glutamine, 10% fetal bovine serum, 10% horse serum, 100 IU/mL penicillin, and 0.10 mg/mL streptomycin (pH 7.4). Cells were plated onto

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

poly-L-lysine-coated 96-well (9 mm), clear-bottomed, black-well culture plates (MidSci, St. Louis, USA) at a density of 1.5×10^5 cells/well. Cells were then incubated at 37 °C in a 5% CO_2 and 95% humidity atmosphere. Cytosine arabinoside (ARA-C; 10 μM) was added to the culture medium 24 h after plating to prevent proliferation of non-neuronal cells. From day 4 *in-vitro*, culture media was changed every other day using a serum free growth medium containing Neurobasal Medium supplemented with B-27, 100 IU/mL penicillin, 0.10 mg/mL streptomycin, and 0.2 mM L-glutamine. Experiments were performed on cerebrocortical cultures at 12 days *in-vitro*.

Cerebrocortical neuron cultures were used for [Ca^{2+}] measurements at day 12 *in vitro* (DIV12). Briefly, the growth medium was removed and replaced with dye loading medium (100 μL per well) containing 4 μM fluo-3 AM and 0.04% Pluronic acid in Locke's buffer (8.6 mM HEPES, 5.6 mM KCl, 154 mM NaCl, 5.6 mM glucose, 1.0 mM $MgCl_2$, 2.3 mM $CaCl_2$, 0.1 μM glycine, pH 7.4). After 1 h of incubation in dye loading medium, the neurons were washed 4x with fresh Locke's buffer (180 μL per well) using an automated microplate washer (Bio-Tek Instruments Inc) and transferred into a FLIPR II (Molecular Devices, Sunnyvale, USA). Cells were excited at 488 nm and Ca^{2+}-bound fluo-3 emission was recorded at 515–575 nm at 1.2 second intervals for calcium influx and 0.5 second intervals for calcium oscillations. For calcium influx, baseline fluorescence was recorded for 1 min, 20 μL of 10X concentrations of compounds were added to wells at a rate of 20 μL/s, and the fluorescence was monitored for an additional 5 min. For calcium oscillations, baseline fluorescence was recorded for 2 min, 20 μL of 10X concentrations of compounds were added to wells at a rate of 20 μL/s, and the fluorescence was monitored for an additional 2 min. Graphpad Prism 7 software was used to analyze time and concentration response relationships. Calcium oscillation data is a plot of the raw Fluo3 fluorescence reads over time. With the calcium influx graph, Fluo3 fluorescence was expressed as (Fmax - $F0$), where Fmax is the maximum and Fo is the baseline line fluorescence measured in each well.[32–34]

The extract 2245CR as well as all nine VLC fractions 2245CR_A–I were evaluated for cytotoxicity using NCI-H460 human large cell lung carcinoma cells for 48 h at two concentrations (10 and 1 μg/mL) using the MTT staining method for cytotoxicity.[58] Pagoamide A was screened in this assay system at 10 concentrations prepared by half-logarithmic dilutions starting at 30 μM, run in duplicate. Doxorubicin and 1% DMSO were used as positive and negative controls, respectively. The optical density of the stained lysate was measured at 570 and 630 nm.

SMART Analysis of Pagoamide A (1).

The HSQC spectrum for pagoamide A (**1**) was pre-processed using MestreNova, phase corrected and referenced to the solvent peak. The vertical gain was adjusted to facilitate automatic peak picking and peak list generation. The peak list table was used as the input to SMART, a deep CNN-based 2D HSQC classification system.[45] The deep CNN was trained with SqueezNet[59] using the HSQC spectra of a mixture of over 30,000 real ^{1}H-^{13}C chemical shift tables and predicted chemical shifts. The predicted data was generated using ACD/labs predictor tool. Details regarding the latest SMART 2.0 system will be published separately.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The molecular structures and cosine similarity of the top nine hits as output of SMART are shown in the Figure 5. Predicted data are shown as (Calcd), with the solvent indicated; otherwise, data are from experimental 2D NMR spectra available from the literature. The cosine similarity scores are presented below each chemical structure.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This research was supported by NIH grant GM107550 (to W.H.G). We acknowledge the Donald C. and Elizabeth M. Dickinson Foundation for generously providing the JEOL NMR spectrometer used in this study. At UC San Diego, we appreciate A. Mrse and B. Duggan for facilitating acquisition of NMR data and Y. Su for help acquiring HRESIMS data. We thank G. Li for providing conformer and stereochemical calculations. Y.L. acknowledges and appreciates financial support provided by the fellowship from the Ocean University of China, PR China.

REFERENCES

(1). Hamann MT; Scheuer PJ J. Am. Chem. Soc 1993, 115, 5825–5826.

(2). Hamann MT; Otto CS; Scheuer PJ; Dunbar DC J. Org. Chem 1996, 63, 6594–6600.

(3). Albadry MA; Elokely KM; Wang B; Bowling JJ; Abdelwahab MF; Hossein MH; Doerksen RJ; Hamann MT J. Nat. Prod 2013, 76, 178–185. [PubMed: 23363083]

(4). Horgen FD; Delos Santos DB; Goetz G; Sakamoto B; Yukiko K; Nagai H; Scheuer PJ J. Nat. Prod 2000, 63, 152–154. [PubMed: 10650101]

(5). Rao KV; Na M; Cook JC; Peng J; Matsumoto R; Hamann MT J. Nat. Prod 2008, 71, 772–778. [PubMed: 18407693]

(6). Dmitrenok A; Iwashita T; Nakajima T; Sakamoto B; Namikoshi M; Nagai H Tetrahedron 2006, 62 (6), 1301–1308.

(7). Kan Y; Fujita T; Sakamoto B; Hokama Y; Nagai H J. Nat. Prod 1999, 62, 1169–1172. [PubMed: 10479330]

(8). Tilvi S; Naik CG J. Mass Spectrom 2007, 42, 70–80. [PubMed: 17149796]

(9). Carbone ML; Kiss R; Mathieu V; Casapullo A; Devi P; Ciavatta ML; Gavagnin M Tetrahedron 2016, 72 (5), 625–631.

(10). Ashour M; Edrada R; Ebel R; Wray V; Wa W; Padmakumar K; Mu WEG; Lin WH; Proksch P J. Nat. Prod 2006, 69, 1547–1553. [PubMed: 17125219]

(11). Goetz G; Nakao Y; Scheuer PJ J. Nat. Prod 1997, 60, 562–567.

(12). Goetz G; Yoshida WY; Scheuer PJ Tetrahedron 1999, 55 (25), 7739–7746.

(13). Zan J; Li Z; Tianero MD; Davis J; Hill RT; Donia MS Science 2019, 364, eaaw6732.

(14). Vervoort H; Fenical W; Epifanio RDA J. Org. Chem 2000, 65 (3), 782–792. [PubMed: 10814011]

(15). Matthew S; Ross C; Rocca JR; Paul VJ; Luesch H J. Nat. Prod 2007, 70, 124–127. [PubMed: 17253864]

(16). Pelay-Gimeno M; Tulla-Puche J; Albericio F Marine Drugs. 2013, 11, 1693–1717. [PubMed: 23697952]

(17). Gao J; Hamann MT Chem. Rev 2011, 111, 3208–3235. [PubMed: 21480581]

(18). Nogle LM; Williamson RT; Gerwick WH J. Nat. Prod 2001, 64, 716–719. [PubMed: 11421730]

(19). Medina RA; Goeger DE; Hills P; Mooberry SL; Huang N; Romero LI; Ortega-Barría E; Gerwick WH; McPhail KL J. Am. Chem. Soc 2008, 130, 6324–6325. [PubMed: 18444611]

(20). Linington RG; Edwards DJ; Shuman CF; McPhail KL; Matainaho T; Gerwick WH J. Nat. Prod 2008, 71, 22–27. [PubMed: 18163584]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(21). Leliaert F; Verbruggen H; D'Hondt S; López-Bautista JM; De Clerck O Cryptogam. Algol 2014, 35, 207–219.

(22). Neveux N; Magnusson M; Maschmeyer T; de Nys R; Paul NA GCB Bioenergy 2015, 7, 673–689.

(23). Magnusson M; Mata L; de Nys R,; Paul NA Mar. Biotechnol 2014, 16, 456–464. [PubMed: 24510373]

(24). Page JZ; Sweeney BM J. Phycol 1968, 4, 253–260. [PubMed: 27068197]

(25). Gosch BJ; Magnusson M; Paul NA; De Nys R GCB Bioenergy 2012, 4, 919–930.

(26). Iwasaki A; Ohno O; Sumimoto S Matsubara T; Shimada S; Sato T; Suenaga K J. Nat. Prod 2015, 78, 901–908. [PubMed: 25768725]

(27). Wang M; Carver JJ; Phelan VV; Sanchez LM; Garg N; Peng Y; Nguyen DD; Watrous J; Kapono CA; Luzzatto-Knaan T; Porto C; Bouslimani A; Melnik AV; Meehan Mi. J.; Liu W-T; Crusemann M; Boudreau PD; Esquenazi E; Sandoval-Calderon M; Kersten RD; Pace LA; Quinn RA; Duncan KR; Hsu C-C; Floros DJ; Gavilan RG; Kleigrewe K; Northen T; Dutton RJ; Parrot D; Carlson EE; Aigle B; Michelsen CF; Jelsbak L; Sohlenkamp C; Pevzner P; Edlund A; McLean J; Piel J; Murphy BT; Gerwick L; Liaw C-C; Yang Y-L; Humpf H-U; Maansson M; Keyzers RA; Sims AC; Johnson AR; Sidebottom AM; Sedio BE; Klitgaard A; Larson CB; Boya CA; Torres-Mendoza D; Gonzalez DJ; Silva DB; Marques LM; Demarque DP; Pociute E; O'Neill EC; Briand E; Helfrich EJN; Granatosky EA; Glukhov E; Ryffel F; Houson H; Mohimani H; Kharbush JJ; Zeng Y; Vorholt JA; Kurita KL; Charusanti P; McPhail KL; Nielsen KF; Vuong L; Elfeki M; Traxler MF; Engene N; Koyama N; Vining OB; Baric R; Silva RR; Mascuch SJ; Tomasi S; Jenkins S; Macherla V; Hoffman T; Agarwal V; Williams PG; Dai J; Neupane R; Gurr J; Rodriguez AMC; Lamsa A; Zhang C; Dorrestein K; Duggan BM; Almaliti J; Allard P-M; Phapale P; Nothias L-F; Alexandrov T; Litaudon M; Wolfender J-L; Kyle JE; Metz TO; Peryea T; Nguyen D-T; Van Leer D; Shinn P; Jadhav A; Muller R; Waters KM; Shi W; Liu X; Zhang L; Knight R; Jensen PR; Palsson BO; Pogliano K; Linington RG; Gutierrez M; Lopes NP; Gerwick WH; Moore BS; Dorrestein PC; Bandeira N Nat. Biotechnol 2016, 34, 828–837. [PubMed: 27504778]

(28). Yang JY; Sanchez LM; Rath CM; Liu X; Boudreau PD; Bruns N; Glukhov E; Wodtke A; De Felicio R; Fenner A; Wong WR; Linington RG; Zhang L; Debonsi HM; Gerwick WH; Dorrestein PC J. Nat. Prod 2013, 76, 1686–1699. [PubMed: 24025162]

(29). Moss NA; Leao T; Glukhov E; Gerwick L; Gerwick WH Collection, Culturing, and Genome Analyses of Tropical Marine Filamentous Benthic Cyanobacteria, 1st ed.; Elsevier Inc., 2018; Vol. 604.

(30). Li T; Li CT; Butler K; Hays SG; Guarnieri MT; Oyler GA; Betenbaugh MJ Biotechnol. Biofuels 2017, 10, 1–11. [PubMed: 28053662]

(31). Zhang P; Zhai C; Wang X; Liu C; Jiang J; Xue Y J. Appl. Phycol 2013, 25, 555–565.

(32). Cao Z; George J; Gerwick WH; Baden DG; Rainier JD; Murray TF J. Pharmacol. Exp. Ther 2008, 326, 604–613. [PubMed: 18448863]

(33). Cao Z; LePage KT; Frederick MO; Nicolaou KC; Murray TF Toxicol. Sci 2010, 114, 323–334. [PubMed: 20047973]

(34). Cao Z; Shafer TJ; Murray TF J. Pharmacol. Exp. Ther 2011, 336, 197–205. [PubMed: 20881019]

(35). Han B; McPhail KL; Gross H; Goeger DE; Mooberry SL; Gerwick WH Tetrahedron 2005, 61, 11723–11729.

(36). Ishida K; Nakagawa H; Murakami M J. Nat. Prod 2000, 63, 1315–1317. [PubMed: 11000050]

(37). Perez LJ; Faulkner DJ J. Nat. Prod 2003, 66, 247–250. [PubMed: 12608858]

(38). Bhushan R; Brückner H J. Chromatogr. B 2011, 879, 3148–3161.

(39). Bhushan R; Brückner H Amino Acids 2004, 27, 231–247. [PubMed: 15503232]

(40). Luesch H; Yoshida WY; Moore RE; Paul VJ J. Nat. Prod 2000, 63, 1437–1439. [PubMed: 11076573]

(41). Medina RA; Goeger DE; Hills P; Mooberry SL; Huang N; Romero LI; Ortega-Barría E; Gerwick WH; McPhail KL J. Am. Chem. Soc 2008, 130, 6324–6325. [PubMed: 18444611]

(42). Pettit GR; Kamano Y; Herald CL; Tuinman AA; Boettner FE; Kizu H; Schmidt JM; Baczynskyj L; Tomer KB; Bontems RJ J. Am. Chem. Soc 1987, 109, 6883–6885

(43). Sardar D; Pierce E; McIntosh JA; Schmidt EW ACS Synth. Biol 2015, 4, 167–176 [PubMed: 24625112]

(44). Ortega MA; van der Donk WA Cell Chem. Biol 2016, 23, 31–44. [PubMed: 26933734]

(45). Zhang C; Idelbayev Y; Roberts N; Tao Y; Nannapaneni Y; Duggan BM; Min J; Lin EC Gerwick EC; Cottrell GW; Gerwick WH Sci. Rep 2017, 7, 14243. [PubMed: 29079836]

(46). Rudi A; Chill L; Aknin M; Kashman Y J. Nat. Prod 2003, 66, 575–577. [PubMed: 12713422]

(47). Luesch H; Williams PG; Yoshida WY; Moore RE; Paul VJ J. Nat. Prod 2002, 65, 996–1000. [PubMed: 12141859]

(48). McDonald LA; Ireland CM J. Nat. Prod 1992, 55, 376–379. [PubMed: 1593285]

(49). Ploutno A; Carmeli S Tetrahedron 2002, 58, 9949–9957.

(50). Tan LT; Sitachitta N; Gerwick WH J. Nat. Prod 2003, 66, 764–771. [PubMed: 12828459]

(51). Corbaz R; Ettlinger L; Gaumann E; Keller-Schierlein W; Kradolfer F; Neipp L; Prelog V; Reusser P; Zahner H Helv. Chim. Acta 1957, 40, 199–204.

(52). Zhou Q; Dowling A; Heido H; Wohnert J; Brandt U; Baum J; ffrench-Constant R; Bode HB J. Nat. Prod 2012, 75, 1717–1722. [PubMed: 23025386]

(53). Verbruggen H; Ashworth M; LoDuca ST; Vlaeminck C; Cocquyt E; Sauvage T; Zechman FW; Littler DS; Litter MM; Leliaert F; De Clerck O Mol. Phylogenet. Evol 2009, 50, 642–653. [PubMed: 19141323]

(54). Martin E; Pickett-Heaps J; Kim GH; West J Algae 2006, 21, 109–124.

(55). Shannon paul; Markiel A; Ozier O; Baliga NS; Wang JT; Ramage D; Amin N; Schwikowski B; Ideker T Genome Res. 2003, 13, 2498–2504. [PubMed: 14597658]

(56). Yao G; Pan Z; Wu C; Wang W; Fang L; Su W J. Am. Chem. Soc 2015, 137, 13488–13491. [PubMed: 26469305]

(57). Naresh Chary V; Sudarshana Reddy B; Kumar CD; Srinivas R; Prabhakar S J. Mass Spectrom 2015, 50, 771–781. [PubMed: 26259661]

(58). Alley MC; Scudiero DA; Monks A; Hursey M; Czerwinski MJ; Fine DL; Abbott BJ; Mayo JG; Shoemaker R; Boyd MR Cancer Res. 1988, 48, 589–601. [PubMed: 3335022]

(59). Iandola FN; Han S; Moskewicz MW; Ashraf K; Dally WJ; Keutzer K arXiv Prepr. arXiv1602.07360 2016.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

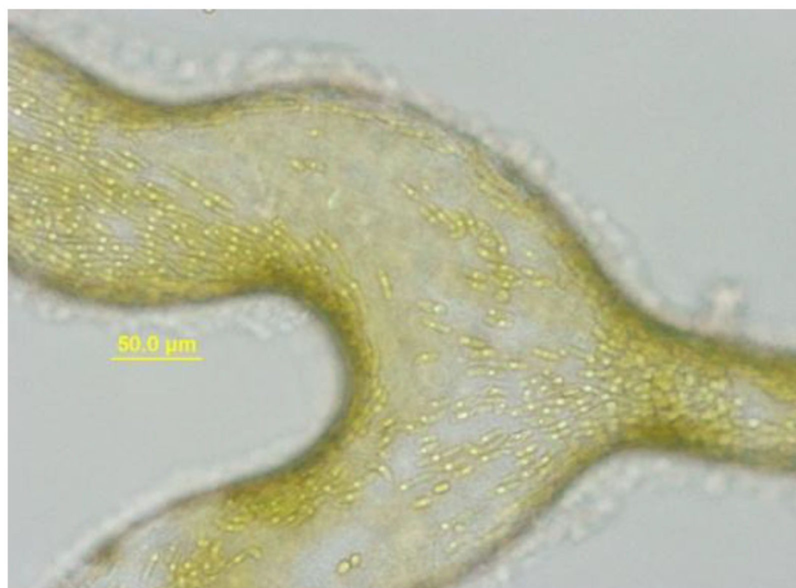


Figure 1.
Photomicrograph of ASF14JUL14–1A *Derbesia* sp. from culture (bar = 50 µm).

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

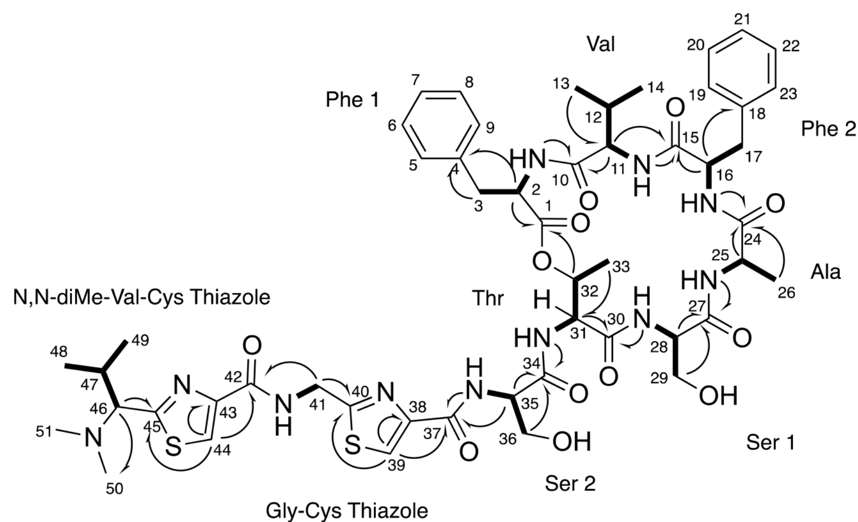


Figure 2.
Select correlations used to determine the planar structure of pagoamide A (**1**). Bolded bonds represent proton-based spin systems deduced by COSY correlations. Arrows represent correlations observed in the HMBC spectrum.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

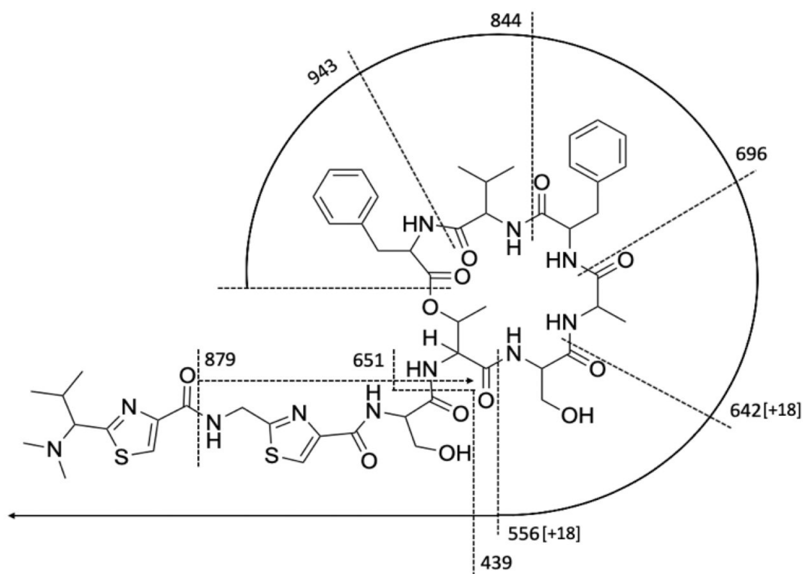


Figure 3.
Selected MS/MS fragments of pagoamide A (**1**).

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

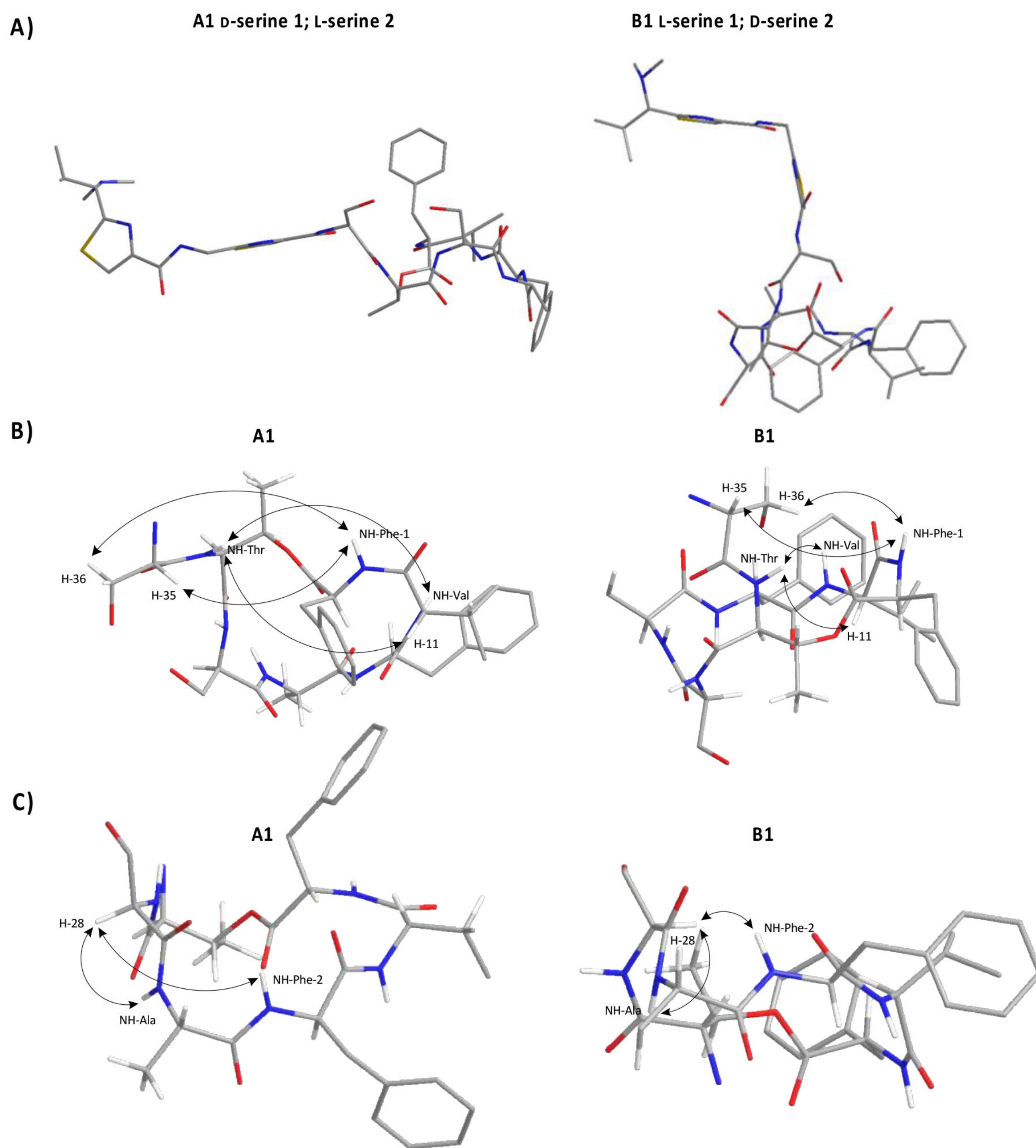


Figure 4.

A. Conformationally optimized stereostructures of the two possible absolute configurations of pagoamide A (**1**) A1) Serine-1 as D and serine-2 as L, B1) serine-1 as L and serine-2 as D. Protons were omitted for clarity. **B.** Conformationally optimized stereostructures of two possible absolute configurations for pagoamide A (**1**) were discriminated using selected ROESY correlations (e.g. between Ser-2 and Phe-1 residues and between Thr and Val residues). The Gly-Cys thiazole, N,N-diMe-Val-Cys thiazole and unrelated protons were omitted for clarity. **C.** Conformationally optimized stereostructures of two possible absolute configurations of pagoamide A (**1**) were discriminated using additional selected ROESY

correlations (e.g. between Ser-1 and NH's of Ala and the Phe-2). The side chains and unrelated protons were omitted for clarity.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

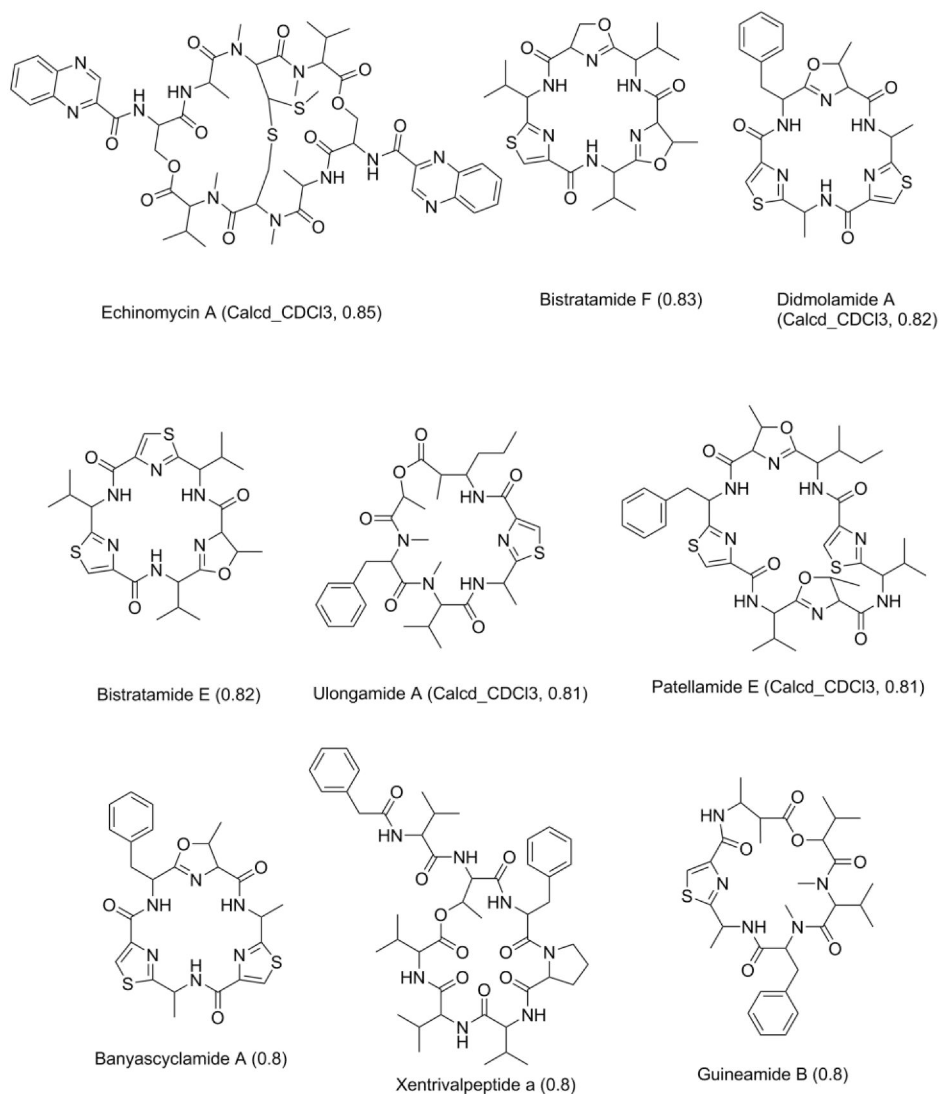


Figure 5.
The nine most similar natural products to pagoamide A (**1**) using the SMART tool based on ^{1}H-^{13}C HSQC spectra (similarity coefficient above 0.8 out of 1).[45]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

^{1}H and ^{13}C NMR Data for Pagoamide A (**1**) in DMSO-d_6[a]

Residue	Position	δ_C, type	δ_H, mult. (*J*, Hz)	HMBC
Phenylalanine-1	1	170.4, C		
	2	52.4, CH	4.85, m	1
	3	36.1, CH_2	3.18, m	2, 4, 5–9
	4	137.9, C		
	5, 9	128.1, CH	7.19, m	3, 4, 6–8
	6, 8	129.2, CH	7.19, m	3–5, 7, 9
	7	126.2, CH	7.13, m	5–9
	NH		7.74, d *(9.35)*	10
Valine	10	170.0, C		
	11	58.7, CH	3.87, dd *(7.9, 4.6)*	10, 12, 13, 14, 15
	12	29.3, CH	1.84, m	13, 14
	13	19.1, CH_3	0.59, d *(6.9)*	11, 12, 14
	14	16.4, CH_3	0.27, d *(6.9)*	11, 12, 13
	NH		7.38, *d (7.9)*	15
Phenylalanine-2	15	172.5, C		
	16	53.0, CH	4.58, m	15, 18
	17	34.1, CH_2	3.16, m	16
	18	138.1, C		
	19, 23	128.0, CH	7.19, m	17, 18, 20–22
	20, 22	129.1, CH	7.19, m	17–19, 21, 23
	21	126.1, CH	7.13, m	19–23
	NH		8.21, d *(8.9)*	24
Alanine	24	173.5, C		
	25	50.1, CH	4.03, *quint (6.8)*	24, 26
	26	16.7, CH_3	1.15 d *(7.3)*	24, 25
	NH		9.12, d *(5.7)*	25, 26, 27
Serine-1	27	171.4, C		
	28	57.4, CH	4.14, t *(7.0)*	27, 29
	29	60.7, CH_2	3.57, m	27, 28
	NH		8.53, s	30
Threonine	30	169.1, C		
	31	54.1, CH	4.85, m	
	32	71.3, CH	5.65, qd *(6.6, 3.2)*	1
	33	16.0, CH_3	1.11, d *(6.6)*	31, 32
	NH		6.86, d *(9.5)*	34
Serine-2	34	169.9, C		
	35	54.5, CH	4.87, m	34, 37
	36	62.4, CH_2	3.75, dd *(11.0, 4.3)*; 3.93, dd *(11.1, 4.1)*	34, 35

Residue	Position	δ_C, type	δ_H, mult. (*J*, Hz)	HMBC
	NH		8.07, d *(8.0)*	37
Gly-Cys Thiazole	37	159.7, C		
	38	148.8, C		
	39	124.2, CH	8.10, s	37, 38, 40
	40	170.3, C		
	41	40.8, CH_2	4.74, d *(6.4)*	40, 42
	NH		9,26, t *(6.2)*	42
N,N-diMe-Val-Cys Thiazole	42	161.1, C		
	43	148.5, C		
	44	124.4, CH	8.31, s	42, 43, 45
	45	168.4, C		
	46	71.9, CH	3.51, m	45, 47, 48, 49, 50, 51
	47	29.5, CH	2.26, m	45, 46
	48	19.4, CH_3	0.75, d *(6.5)*	46, 47, 49
	49	20.1, CH_3	0.99, d *(6.5)*	46, 47, 48
	50	41.3, CH_3	2.15, s	46, 51
	51	41.3, CH_3	2.15, s	45, 50

[a] Data recorded at 500 MHz (^{1}H NMR) and 125 MHz (^{13}C NMR).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript