JOC
The Journal of Organic Chemistry

Downsizing Class II Lasso Peptides: Genome Mining-Guided Isolation of Huascopeptin Containing the First Gly1-Asp7 Macrocycle

Carlos Cortés-Albayay,[⊥] Scott A. Jarmusch,[⊥] Franziska Trusch, Rainer Ebel, Barbara A. Andrews, Marcel Jaspars,* and Juan A. Asenjo*

Cite This: *J. Org. Chem.* 2020, 85, 1661−1667

Read Online

ACCESS | 📊 Metrics & More | 📖 Article Recommendations | 🔵 Supporting Information

ABSTRACT: A new lasso peptide, huascopeptin, was isolated following genome-mined discovery of a new biosynthetic gene cluster in extremotolerant *Streptomyces huasconensis* HST28[T] from Salar de Huasco, Atacama Desert, Chile. Compound **1** is a 13-residue class II lasso peptide containing a novel Gly1-Asp7 macrolactam ring, a three-residue loop, and a three-residue tail, making it the smallest lasso peptide isolated to date. The lasso structure was confirmed using NOE restraint-based molecular dynamics simulations.

S. huasconensis HST28
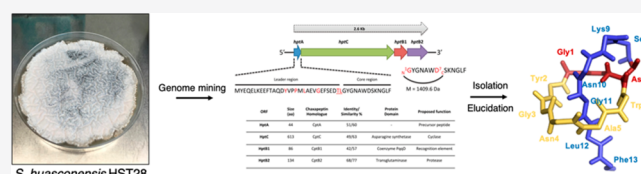
The development of high-throughput sequencing methods coupled with modern genome mining approaches has provided universal access to the biosynthetic potential of microbes and subsequently the potential to discover new chemistry.[1−3] The biosynthetic machineries producing ribosomally synthesized and post-translationally modified peptides (RiPPs) have conserved protein domains involved in the post-translational modification of the precursor peptide. Structural characteristics and putative structures can be determined on the basis of the predicted precursor peptides and the post-translational modification enzymes,[4] making RiPPs an ideal target for genome mining.

Lasso peptides, characterized by their unique lariat-knot conformation, have increasingly drawn attention due to their broad range of biological activities including antimicrobial,[5−7] antiviral,[8] and enzymatic inhibition.[9] Intriguingly, software tools such as RODEO and RiPPER predict that the overall chemical space of lasso peptides has barely been scratched.[10−12]

Streptomyces spp. are prolific sources of antimicrobial lasso peptides.[13−15] Lasso peptides are unique due to the formation of a macrolactam ring as a result of an isopeptide bond between the *N*-terminal α-amino group and the carboxylic group located in the side chain of an aspartic acid or a glutamic acid.[16] The lariat-knot is formed enzymatically when the macrolactam ring cyclizes around the tail, locking the carboxylic acid terminus.[17] This conformation is stabilized by disulfide bonds in class I and III lasso peptides, by steric hindrance in class II and by "handcuffing" in class IV.[11−13]

Extreme conditions in the Atacama Desert favor the development of unique actinobacterial diversity, forming a basis for new chemistry.[18,19] In continuation of our previous studies, we isolated a new species, *Streptomyces huasconensis*

HST28[T],[20] from Salar de Huasco, an athalossohaline high altitude wet-land (3800 m.a.s.l.), located in the Atacama Desert.[21] Genome mining was used to mine strain HST28[T] for RiPPs which confirmed the presence of one new lasso peptide gene cluster, *hpt*, encoding for the production of a novel class II chemotype. *In silico* predictions were confirmed *in vivo*, leading to the isolation and structure elucidation of huascopeptin **1** using MS[2] and NMR. Compound **1** contains the first reported Gly1-Asp7 macrolactam ring. Furthermore, the lasso nature of the molecule was evaluated by NOE and subsequent modeling, confirming the lasso motif laid out for the xanthomonins.[22]

In silico analysis of the *S. huasconensis* HST28[T] genome showed 39 putative secondary metabolite BGCs among which polyketides, nonribosomal peptides, and RiPPs were found.[20] The latter included four new lasso peptide BGCs,[20] among which the huascopeptin BGC (*hpt*) was assumed to be responsible for the production of **1** has a total size of 2.6 kb (Figure 1). The Hpt BGC houses the four main genes involved in class II lasso peptide biosynthesis and maturation:[23] (I) the HptA gene encodes for 44 amino acid residues including the precursor peptide of huascopeptin **1**. The precursor peptide of huascopeptin and most other lasso peptides contain highly conserved residues that allow for proper loading into further downfield biosynthetic machineries as shown by many studies.[24−28] HptA consists of 31 and 13 amino acid residues in the leader (*N*-terminal) and in the core peptide regions (*C*-
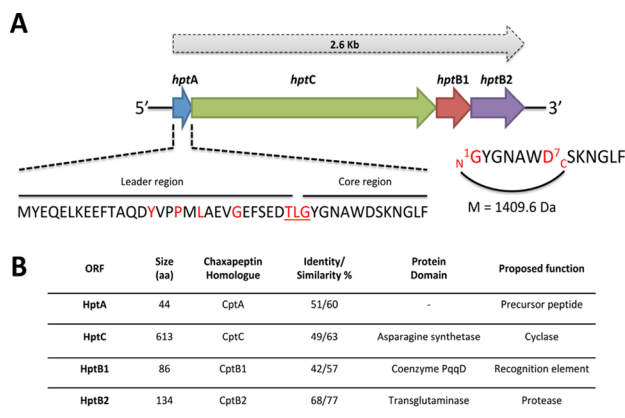
Figure 1. Huascopeptin biosynthetic gene cluster. (A) Genetic map of *hpt* gene cluster in the genome of *S. huasconensis* HST28[T]. The residues in red correspond to the leader conserved motif. (B) Functions and relatedness of the proteins encoded by the *hpt* genes with their homologues of the *cpt* gene cluster for chaxapeptin.[32]

Figure 2. MS[2] data showing observation of *b*-fragment ions of the peptide tail in accordance with those predicted in silico from the core peptide sequence.

terminal), respectively; (II) the HptC gene encodes for an asparagine synthetase-like protein domain (HptC) which is responsible for isopeptide bond formation and subsequent macrocyclisation.[29] (III) the HptB1 gene product contains a pyrroloquinoline quinone protein domain D (PqqD), also known as the RiPP recognition element (RRE), which binds to the leader region and delivers the precursor peptide to the protease HptB2 for the further processing (Figure 1);[30,31] (IV) the HptB2 gene encodes a transglutaminase-like protease which is involved in the cleavage of the leader region in the precursor peptide.[31]

The core region amino acid sequence, GYGNAWDSK-NGLF, was predicted to have an accurate mass of 1409.6360 Da. The leader and core regions are divided by the characteristic TxG motif (Figure 1) needed to remove the leader peptide during lasso peptide maturation.[33]

The four genes of the *hpt* cluster (*hpt*ACB1B2) were found to be closely related to the *cpt*ACB1B2 biosynthetic gene cluster of chaxapeptin.[32] The HptA precursor peptide and HptC cyclase showed 51% and 49% identity with CptA and CptC while HptB1 protease and HptB2 had 42% and 68% of identity with CptB1 and CptB2, respectively (Figure 1). The presence of two nearby ABC transporters was observed, but whether they were involved in the BGC could not be confirmed.

The molecular formula of huascopeptin was established as $C_{65}H_{87}N_{17}O_{19}$ based on HRESIMS of the single and double charged ions, which matched the predicted core sequence (Figure S1). MS[2] of purified **1** (Figure 2) verified the sequence of tail residues. The sequence of the remaining amino acid residues within the macrocycle could not be established using MS[2] due to high stability of the macrocycle. ^{1}H, ^{13}C, and ^{15}N NMR data were all in agreement with in silico and MS predictions. Along with one-dimensional experiments, two-dimensional [1]H−^{13}C HSQC, [1]H−^{15}N HSQC, [1]H−^{13}C HMBC, [1]H−[1]H TOCSY, and [1]H−[1]H NOESY allowed for almost all signals (Table S1) in the 13-residue peptide to be assigned, including three glycines (Gly), two asparagines (Asn), one tyrosine (Tyr), one tryptophan (Trp), one alanine (Ala), one aspartic acid (Asp), one serine (Ser), one lysine (Lys), one leucine (Leu), and one phenylalanine (Phe) (Figures S3−S8a).
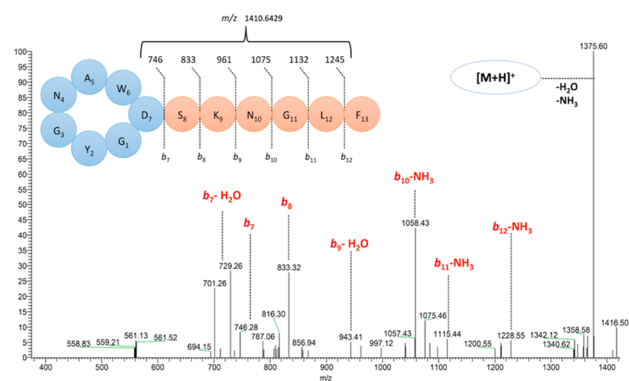
Using a combination of HRESI-MS/MS and NMR data, the planar structure of **1** was defined as a seven-membered amino acid cyclic isopeptide with a six-membered amino acid side chain (Figure 3). The peptide tail fragmented sequentially
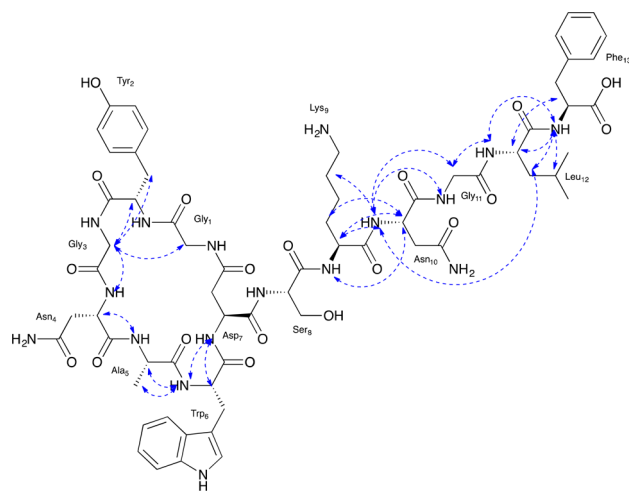
Figure 3. Key NOESY correlations (blue) confirming the majority of the connectivities for huascopeptin 1 peptide sequence.

under CID, typical of most peptides (Figure 2). Sequence connectivity of the lasso peptide tail was confirmed using NOESY correlations (Figure 3). The connection between Leu$_{12}$-Phe$_{13}$ was established using correlations between Leu$_{12}$-Hα to Phe$_{13}$-Hα and Leu$_{12}$-NH, Leu$_{12}$-Hα, Leu$_{12}$-Hβ, and Leu$_{12}$-Hγ all to Phe$_{13}$-NH. The connection between Gly$_{11}$-Leu$_{12}$ was determined with correlations between Gly$_{11}$-Hα (δ3.99 and δ2.99) and Leu$_{12}$-NH. Furthermore, connection between Asn$_{10}$-Gly$_{11}$ was established with correlations between Asn$_{10}$-NH to Gly$_{11}$-NH and Gly$_{11}$-Hα (δ3.99). The connection between Lys$_9$-Asn$_{10}$ was established using correlations between Lys$_9$-NH to Asn$_{10}$-Hα, Lys$_9$-Hα to Asn$_{10}$-NH and to Asn$_{10}$-Hα, Lys$_9$-Hβ (δ2.06 and δ1.68) to Asn$_{10}$-Hα, and last, Lys$_9$-Hδ (δ1.69) to Asn$_{10}$-NH. Connectivity between Ser$_8$-Lys$_9$ and Asp$_7$-Ser$_8$ was established with MS[2] data due to a lack of NMR evidence.

Sequence connectivity of the macrocycle was also confirmed using NOESY correlations (Figure 3). The connection between Tyr$_2$-Gly$_3$ was established using NOESY correlations between Tyr$_2$-Hα and Gly$_3$-H$_a$α and Tyr$_2$-Hβ and Gly$_3$-H$_a$α. Additionally, the connection between Gly$_3$-Asn$_4$ was estab-
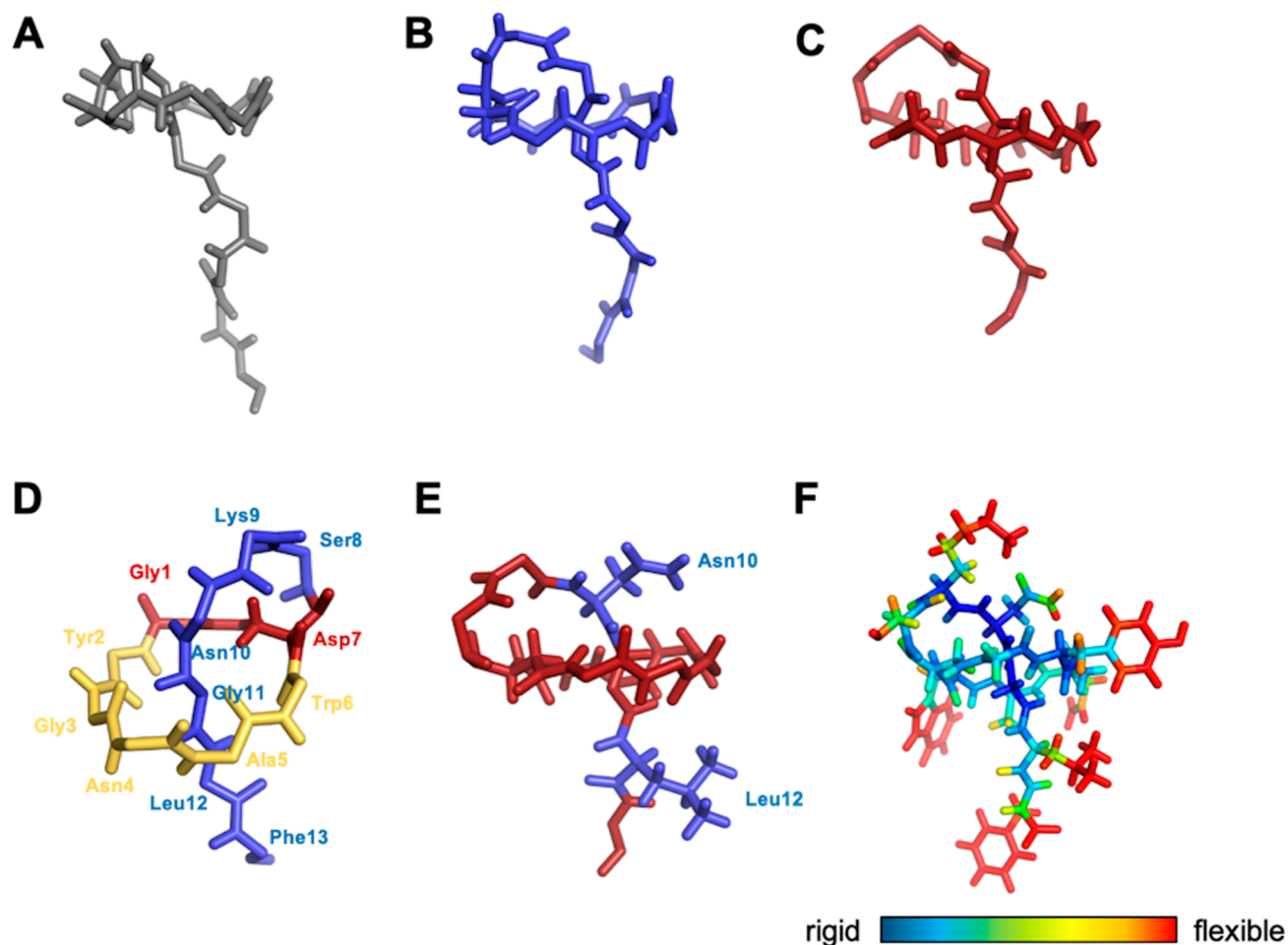
Figure 4. Model of huascopeptin. (A) Huascopeptin with a typical lasso peptide motif with a short loop (Ser8 only) and a long tail. (B) Huascopeptin with a lasso fold with Ser8 and Lys9 building the loop. (C) Huascopeptin with a lasso structure with a long loop comprising Ser8, Lys9, as well as Asn10 and a short tail. (D) Model C is energetically most favorable and likely to reflect the solution conformation of huascopeptin. It consists of a seven-residue ring, a three-residue loop, and a two-residue tail below the ring. (E) The plug amino acids in huascopeptin are Asn10 above and Leu12. (F) Flexibility map (blue: rigid, red: flexible) of huascopeptin over a period of 100 ns during MD simulation shows a rigid core structure typical for lasso peptides.

lished using the correlations between Gly_3-$H_a\alpha$ and Asn_4-NH. The connection between Gly_1-Tyr_2 was established through the link between Gly_1-$H\alpha$ and Gly_3-$H\alpha$. The connection between Asn_4-Ala_5 was determined using the correlation between Asn_4-$H\alpha$ and Ala_5-NH. Furthermore, connection between Ala_5-Trp_6 was determined using correlations between Ala_5-$H\alpha$ and Ala_5-$H\beta$ to Trp_6-NH. The connection between Trp_6-Asp_7 was established based on the correlations between Trp_6-NH and Trp_6-$H\alpha$ to Asp_7-NH. Finally, while there is no NMR evidence that Asp_7 and Gly_1 form the macrolactam ring, MS^2 data confirms that the unfragmented ring is missing the predicted aspartic acid residue.

Due to the limited amount of sample available, we were not able to determine of absolute stereochemistry using Marfey's reagent. With the exception of rare epimerization occurring at the C-terminus,[34,35] all remaining lasso peptides characterized thus far only contain L amino acids,[32,36] and additionally, there is no explicit epimerase in the *hpt* cluster; therefore, L amino acids were placed in the planar structure.

The overall 3D structure of **1** was established by analyzing NOESY spectra which displayed a total of 33 sequential and 9 long-range interactions ($\geq i + 3$). Obtained distance restraints could confirm the ring structure of the peptide but were not

sufficient to locate the position of tail relative to the ring. However, due to the presence of tail to ring NOE correlations as well as a wide range of amide NH chemical shifts (9.22–7.53) characteristic of the lasso topology,[37] we used a lasso motif instead of a branched cycle for models (Figure S3). Hence, YASARA was used to model the structure of huascopeptin with chaxapeptin as a template, a typical class II lasso peptide for which the structure was calculated by NMR.[32]

The ring structure was realized by inserting a single in trans bond between the nitrogen of Gly_1 and the side chain $C\gamma$ of Asp_7. Subsequently, an energy minimization step was performed with a YASARA2 force field to reduce any remaining tension after modeling. Following this, we designed three models of huascopeptin with various lengths of the loop region (Figure 4A–C). To explore their dynamics, MD simulations were performed with an AMBER03 force field over 100 ns. During this time, model A (loop: Ser8 only) was unstable and lost the typical lasso peptide structure which is not in agreement with the high stability of lasso peptides under experimental conditions.[36] Therefore, this model was discarded. After comparing energies between model B (loop: Ser8 and Lys9) and model C (loop: Ser8, Lys9 and Asn10), the

latter showed lower energies of bonds, angles, dihedrals, planarity and van der Waals energies (Table S2). Further evidence for model C is the lasso motif established with the xanthomonins,[21] which consists of a threaded glycine through the ring, as well as NOEs from the plug residues Asn10 and Leu12 (Figure S2).

Thus, the 3D structure of compound **1** comprises a seven-residue ring with a short two-residue tail below the macrolactam ring (Figure 4D). Asn10 and Leu12 are assumed to operate as the plug amino acids located above and below the ring (Figure 4E). Additionally, huascopeptin possesses hydrogen bonds in the loop region (Asp7-Ser8 and Asp7-Lys9) and between residues in the tail and the ring (Gly1-Asn10, Gly3-Gly11, Gly3-Asn10), which results in a characteristic short β-sheet known to promote the observed high stabilities of lasso peptides.[6] Furthermore, the structure of huascopeptin does not show any flexible regions, which is probably due to the short tail in comparison to other lasso peptides with longer tails (Figure 4F).[17,38] Lastly, 400 μg of compound **1** was evaluated for antimicrobial activity against *MRSA* and *P. aeruginosa* and cytotoxicity against MRC5 cells but was found inactive.

A seven-membered macrocycle has only been described one other time for xanthomonins I–III.[22] In the previous study, 25 xanthomonin II site-directed mutants were generated to evaluate the biosynthetic machinery of this group of lasso peptides. Multiple mutants were generated to determine if Asp could replace Glu in position 7 of the macrocycle, but the mutants were unable to produce a mature peptide. Nevertheless, the authors did foresee the possibility of a seven-membered macrocycle with Asp_7 which is first described here. Blastp homology searches showed biosynthetic similarity to chaxapeptin as well as the anantins[11] (Figure S10). DiCaprio et al. showed 2% of predicted lasso peptide chemical space contains seven-membered aspartate macrocycles,[12] making this discovery an important step in further understanding the structural diversity of this RiPP class. Huascopeptin not only represents a new macrocycle linkage but represents the smallest macrocycle opening as 22 atoms, one less than the xanthomonins, which was previously seen as physio-chemically improbable due to the limitation in ring size to thread the tail through the macrocycle.

In conclusion, genome mining continues to be a viable approach for the discovery of new chemistry. In addition to this, the previously unexplored source of Salar de Huasco and in a wider context, the Atacama Desert, continues to be a rich source for specialized metabolites. In this study, genome mining has led to the discovery of huascopeptin **1**, a new class II lasso peptide containing the first evidence for a Gly1-Asp7 macrolactam ring, adding to the already structurally diverse group of RiPPs. This study also represents the first biosynthetic example of this lasso peptide structural chemotype, an important step in better understanding the biosynthetic machinery of lasso peptides and how that translates to new chemistry.

■ EXPERIMENTAL SECTION

Isolation, Culture, And Maintenance. *S. huasconensis* HST28[T] was isolated from an arid soil sample collected in sectors surrounding a salt lake (−20.328611, −68.838611) in complete absence of vegetation. This site is marked by a negative water balance,[21] a salinity gradient from freshwater to saturation, and elevated levels of solar radiation (<1100 W/m²). The soil samples were treated according to the procedure described by Okoro et al.,[39] and strain HST28 was

recovered from Starch-Casein-KNO₃ agar plates supplemented with 50 μg/mL of nystatin and 50 μg/mL of cycloheximide as selective agents.[40] A pure culture was obtained and maintained on Starch-Casein-KNO₃ agar plates at 28 °C for 5 days.

Molecular Analysis. Genomic DNA of strain HST28 was extracted from 10 mL of GYM liquid culture incubated for 5 days at 30 °C while shaking at 150 rpm (glucose, yeast extract, malt extract, and CaCO₃) media. DNA extraction was completed using an UltraClean Microbial DNA isolation kit (MoBio Laboratories, Carlsbad, CA) following the manufacturer procedure. The *16S rRNA-PCR amplification was examined using the primer* Eub9-27F and Eub1542R and according to the conditions by Stackebrandt and Liesack.[41] BLAST of an almost complete 16S rRNA gene sequence (1429pb) of strain HST28 was performed in the EzTaxon server.[42] The assignment of the strain HST28 to the genus *Streptomyces* was confirmed with 99.0% of 16S rRNA gene sequence similarity with *Streptomyces aureus* NBRC 100912[T].[43]

Genome Mining. The whole genome of strain HST28 was sequenced using an Ion Torrent PGM system at the Strathclyde Institute of Pharmacy and Biomedical Sciences (SIPBS, Glasgow, United Kingdom) and submitted in GenBank under accession MK636692.1. Digital DNA:DNA hybridization analysis was carried out with the type strain of close phylogenetic neighbor *S. aureus* species using the Genome-to-Genome distance calculator tool available at DSMZ server.[44]

The genome sequence of strain HST28 was annotated using the RAST web server version 2.0 (Rapid Annotation using Subsystem Technology).[45] The prediction of BGCs for specialized metabolites, including lasso peptides BGCs and their putative final chemical structures was completed using PRISM and complimented against combinatorial structure libraries.[46] The reconstruction and identification of all the enzymes in the BGC for huascopeptin **1** was developed using ARTEMIS.[47] This was used to identify the ORFs of each gene present in the BGC and determine all conserved protein domains of the class II lasso peptides using the chaxapeptin BGC sequence as reference. The chaxapeptin BGC sequence was obtained using the web tool "CDS search" of NCBI.[48] The corresponding amino acid sequence to the core region of the predicted precursor peptide was used to calculate the molecular weight and the web tool Protein Prospector[49] (http://prospector.ucsf.edu) was used to predict potential fragments from MS^2, considering the loss of one water molecule as result of the isopeptide bond formed during the cyclization.

General Experimental Procedures. Optical rotations were measured using a polarimeter. UV was measured using a UV/vis spectrometer. NMR spectra were recorded on a 2.5 mM sample of **1** and obtained with an 800 MHz Bruker spectrometer with 5 mm TCI He ¹H/¹³C/¹⁵N CryoProbe and ¹H−¹⁵N HSQC data was collected using 600 MHz Bruker spectrometer with TCI He ¹H/¹³C/¹⁵N CryoProbe.. High-resolution mass spectrometric data were acquired using a LTQ XL Orbitrap (Thermo) and Bruker MAXIS II QTOF (both in tandem with an Agilent 1290 Infinity UHPLC). LC-LTQ XL Orbitrap utilizes a Phenomenex Kinetex XB-C₁₈ (2.6 μM, 100 × 2.1 mm) with a mobile phase of 5% ACN+0.1% formic acid to 100% ACN+0.1% formic acid in 25 min. LC-QToF utilizes a Phenomenex Kinetex XB-C₁₈ (2.6 μM, 100 × 2.1 mm) column with a mobile phase of 5% ACN+0.1% formic acid to 100% ACN + 0.1% formic acid in 11 min. Semipreperative HPLC separations were conducted using a C₁₈ column (5 μM, 100 Å, 250 × 10 mm column), connected to a binary pump, and monitored using a photodiode array detector. NMR solvents were purchased from VWR international as well as all chromatography solvents.

Fermentation, Extraction, and Isolation. To evaluate the presence of in silico predicted lasso peptides, a single colony of strain HST28 was inoculated in 50 mL (×3) of GYM and SPM (Starch, Soy peptone and CaCl₂ x H₂O) liquid media supplemented with 2% of artificial sea salts and incubated for 7 days at 30 °C with shaking at 150 rpm. The resultant cultures were centrifuged at 4600 rpm for 20 min; the wet biomass was washed twice with sterile Milli-Q water and then homogenized and extracted with methanol while the resultant

supernatants were shaken at 130 rpm overnight with 20 g/L of HP-20 resin at room temperature. The resin was filtered and washed twice with Milli-Q water, extracted with methanol to exhaustion, and all resultant extracts analyzed by LC-MS. For the large-scale culture, 2 L baffled flasks (×8) containing 500 mL of GYM media, were inoculated with 100 mL of the seed culture and incubated for 10 days at 30 °C with shaking at 150 rpm. Large scale fermentation was extracted using HP-20 resin as explained above, except no centrifugation was completed as compound **1** retained on the resin. Crude extract was partitioned using modified Kupchan partitioning[50] yielding five fractions. Solid-phase extraction using 10g C18 (Phenomenex) was conducted on the *sec*-butanol fraction, yielding 5 fractions eluted with increasing methanol. The 50:50 methanol/water fraction was then separated using reversed-phased flash chromatography (Biotage, 12g C18) where huascopeptin was detected in fraction 6 (t_R 50 min). Lastly, fraction 6 was repurified using an Agilent 1200 HPLC with a Waters Sunfire semipreparative column (t_R 63.5 min) with a solvent system of A, 95/5 MeOH/H$_2$O, and B, MeOH. Agilent 1200 HPLC utilized a solvent gradient of 30% B to 100% B at 2.5 mL/min over 85 min. The overall yield of product was 0.6 mg/4 L.

3D Modeling of Huascopeptin. The structure of huascopeptin was modeled with the lasso peptide chaxapeptin as the structural template with YASARA WhatIf (version 11.12.31).[32] Final cyclization was done by deleting one HN of Gly1 and one Oδ (OD) atom of the side chain of Asp7 and inserting a single bond in trans between N of Gly1 and CG of Asp7 to mimic the isopeptide bond of the lasso peptide. The length of the newly created isopeptide bond and position of the replaced amino acids after modeling were corrected by energy minimization with YASARA2 force field. In total, three models differing in the position of the tail through the ring were generated (Figure 4A−C). With all models, molecular dynamics (MD) simulations were performed with AMBER03 over 100 ns in a cubic box (30 Å length) with periodic boundaries and drift correction at 298 K with pH 7.4, 0.9% NaCl, a van der Waals cutoff of 7.86 Å, and PME for long-range Coulomb forces. Hydrogen bonds were analyzed with YASARA WhatIf (version 11.12.31). The flexibility was visualized by color coding the RMSF (Root Mean Square Fluctuation) value over 100 ns and mapping on the structure of huascopeptin for each atom.

Biological Assays. MIC assays were completed on *P. aeruginosa* ATCC 27853 and *MRSA* ATCC 33591 using the methodology established by Ingebrigtsen et al.,[51] with the exception that MRSA was cultivated in Brain heart infusion instead of Mueller-Hinton broth. Cell viability assay (MTS assay) was also completed using the methodology by Ingebrigtsen et al.[51] MRC5 lung fibroblast cells (ATCC CCL-171) were seeded in 96-well microtiter plates at 4000 cells/well in Earl's MEM (EMEM) supplemented with 10% FBS, 10 ug/mL gentamycin, 1% nonessential amino acids, 1% L-glutamine, 1% sodium pyruvate and 1% sodium bicarbonate.

Spectroscopic Data. Huascopeptin **1**: colorless amorphous solid; $[\alpha]^{25}_D$ +20 (*c* 0.05, MeOH); UV (MeOH) λ_{max} 228, 279 nm^{-1}; ^{1}H and ^{13}C NMR data in Table S1; HRMS (ESI-QTOF) *m/z* [M + H]$^+$ calcd for C$_{65}$H$_{88}$N$_{17}$O$_{19}$ 1410.6440, found 1410.6429, Δ = −0.7 ppm.[27]

ASSOCIATED CONTENT

Ⓢⓘ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.joc.9b02231.

NMR table, model energies table, HRESIMS, and spectra of **1**; BGC homology comparison (PDF)

YASARA structures (PDB)

AUTHOR INFORMATION

Corresponding Authors

Marcel Jaspars − *Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Old Aberdeen AB24 3UE, Scotland, U.K.;* Email: m.jaspars@abdn.ac.uk

Juan A. Asenjo − *Centre for Biotechnology and Bioengineering, CeBiB, Department of Chemical Engineering, Biotechnology and Materials, University of Chile, Santiago 8370450, Chile;* Email: juasenjo@ing.uchile.cl

Authors

Carlos Cortés-Albayay − *Centre for Biotechnology and Bioengineering, CeBiB, Department of Chemical Engineering, Biotechnology and Materials, University of Chile, Santiago 8370450, Chile;* ⓞ orcid.org/0000-0001-9312-3162

Scott A. Jarmusch − *Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Old Aberdeen AB24 3UE, Scotland, U.K.;* ⓞ orcid.org/0000-0002-1021-1608

Franziska Trusch − *Institute of Medical Sciences, University of Aberdeen, Foresterhill AB25 2ZD, Scotland, U.K.; Division of Plant Sciences, College of Life Science, University of Dundee (at JHI), Dundee DD2 5DA, U.K.*

Rainer Ebel − *Marine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Old Aberdeen AB24 3UE, Scotland, U.K.*

Barbara A. Andrews − *Centre for Biotechnology and Bioengineering, CeBiB, Department of Chemical Engineering, Biotechnology and Materials, University of Chile, Santiago 8370450, Chile*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.joc.9b02231

Author Contributions

[⊥]C.C.-A. and S.A.J. contributed equally to this work.

Author Contributions

C.C.A., B.A., and J.A. conceived the present idea and developed the theory of this project. C.C.A. and S.A.J. developed the experimental methodology with support of M.J. C.C.A. carried out the culture and maintenance of the strain, optimized and performed the fermentations and developed all the in silico genome analysis. S.A.J. and C.C.A. completed the isolation and purification of huascopeptin. S.A.J. and R.E. completed NMR and LCMS structural characterization. Molecular modeling and calculations were carried out by F.T. The manuscript was drafted by S.A.J., C.C.A., F.T., R.E., and M.J. with input from B.A. and J.A. All the authors read and approved the final manuscript.

Notes

The authors declare no competing financial interest.

The Arctic University of Norway) for conducting biological assays. Finally, we thank Dr. Imen Nouioui and Dr. Cristina Dorador for their feedback during manuscript preparation.

■ REFERENCES

(1) Challis, G. L. Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology* **2008**, *154*, 1555−1569.

(2) Ziemert, N.; Alanjary, M.; Weber, T. The evolution of genome mining in microbes - a review. *Nat. Prod. Rep.* **2016**, *33*, 988−1005.

(3) Gomez-Escribano, J. P.; Castro, J. F.; Razmilic, V.; Chandra, G.; Andrews, B.; Asenjo, J. A.; Bibb, M. J. The Streptomyces leeuwenhoekii genome: de novo sequencing and assembly in single contigs of the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2. *BMC Genomics* **2015**, *16*, 485−015.

(4) Hetrick, K. J; van der Donk, W. A Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era. *Curr. Opin. Chem. Biol.* **2017**, *38*, 36−44.

(5) Li, Y.; Ducasse, R.; Zirah, S.; Blond, A.; Goulard, C.; Lescop, E.; Giraud, C.; Hartke, A.; Guittet, E.; Pernodet, J. L.; Rebuffat, S. Characterization of Sviceucin from Streptomyces Provides Insight into Enzyme Exchangeability and Disulfide Bond Formation in Lasso Peptides. *ACS Chem. Biol.* **2015**, *10*, 2641−2649.

(6) Iwatsuki, M.; Tomoda, H.; Uchida, R.; Gouda, H.; Hirono, S.; Omura, S. Lariatins, antimycobacterial peptides produced by Rhodococcus sp. K01-B0171, have a lasso structure. *J. Am. Chem. Soc.* **2006**, *128*, 7486−7491.

(7) Gavrish, E.; Sit, C. S.; Cao, S.; Kandror, O.; Spoering, A.; Peoples, A.; Ling, L.; Fetterman, A.; Hughes, D.; Bissell, A.; Torrey, H.; Akopian, T.; Mueller, A.; Epstein, S.; Goldberg, A.; Clardy, J.; Lewis, K. Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. *Chem. Biol.* **2014**, *21*, 509−518.

(8) Chokekijchai, S.; Kojima, E.; Anderson, S.; Nomizu, M.; Tanaka, M.; Machida, M.; Date, T.; Toyota, K.; Ishida, S.; Watanabe, K. NP-06: a novel anti-human immunodeficiency virus polypeptide produced by a Streptomyces species. *Antimicrob. Agents Chemother.* **1995**, *39*, 2345−2347.

(9) Yano, K.; Toki, S.; Nakanishi, S.; Ochiai, K.; Ando, K.; Yoshida, M.; Matsuda, Y.; Yamasaki, M. MS-271, a novel inhibitor of calmodulin-activated myosin light chain kinase from Streptomyces sp.–I. Isolation, structural determination and biological properties of MS-271. *Bioorg. Med. Chem.* **1996**, *4*, 115−120.

(10) Hudson, G.; Burkhart, B.; DiCaprio, A.; Schwalen, C.; Kille, B.; Pogorelov, T.; Mitchell, D. Bioinformatic Mapping of Radical S-Adenosylmethionine-Dependent Ribosomally Synthesized and Post-Translationally Modified Peptides Identifies New Cα, Cβ, and Cγ-Linked Thioether-Containing Peptides. *J. Am. Chem. Soc.* **2019**, *141*, 8228−8238.

(11) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H. C.; Zakai, U. I.; Mitchell, D. A. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.* **2017**, *13*, 470−478.

(12) DiCaprio, A. J.; Firouzbakht, A.; Hudson, G. A.; Mitchell, D. A. Enzymatic Reconstitution and Biosynthetic Investigation of the Lasso Peptide Fusilassin. *J. Am. Chem. Soc.* **2019**, *141*, 290−297.

(13) Potterat, O.; Stephan, H.; Metzger, J. W.; Gnau, V.; Zähner, H.; Jung, G. Aborycin − A Tricyclic 21-Peptide Antibiotic Isolated from Streptomyces griseoflavus. *Liebigs Annalen der Chemie* **1994**, *1994*, 741−743.

(14) Weber, W.; Fischli, W.; Hochuli, E.; Kupfer, E.; Weibel, E. K. Anantin–a peptide antagonist of the atrial natriuretic factor (ANF). I. Producing organism, fermentation, isolation and biological activity. *J. Antibiot.* **1991**, *44*, 164−171.

(15) Detlefsen, D. J.; Hill, S. E.; Volk, K. J.; Klohr, S. E.; Tsunakawa, M.; Furumai, T.; Lin, P. F.; Nishio, M.; Kawano, K.; Oki, T. Siamycins I and II, new anti-HIV-1 peptides: II. Sequence analysis and structure determination of siamycin I. *J. Antibiot.* **1995**, *48*, 1515−1517.

(16) Maksimov, M. O.; Pan, S. J.; James Link, A. Lasso peptides: structure, function, biosynthesis, and engineering. *Nat. Prod. Rep.* **2012**, *29*, 996−1006.

(17) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Goransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Muller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Sussmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **2013**, *30*, 108−160.

(18) Rateb, M. E.; Ebel, R.; Jaspars, M. Natural product diversity of actinobacteria in the Atacama Desert. *Antonie van Leeuwenhoek* **2018**, *111*, 1467−1477.

(19) Goodfellow, M.; Nouioui, I.; Sanderson, R.; Xie, F.; Bull, A. T. Rare taxa and dark microbial matter: novel bioactive actinobacteria abound in Atacama Desert soils. *Antonie van Leeuwenhoek* **2018**, *111*, 1315−1332.

(20) Cortés-Albayay, C.; Dorador, C.; Schumann, P.; Andrews, B.; Asenjo, J.; Nouioui, I. *Streptomyces huasconensis* sp. nov., an haloalkalotolerant actinobacterium isolated from a high altitude saline wetland at the Chilean Altiplano. *Int. J. Syst. Evol. Microbiol.* **2019**, *69*, 2315−2322.

(21) Dorador, C.; Vila, I.; Remonsellez, F.; Imhoff, J. F.; Witzel, K. P. Unique clusters of Archaea in Salar de Huasco, an athalassohaline evaporitic basin of the Chilean Altiplano. *FEMS Microbiol. Ecol.* **2010**, *73*, 291−302.

(22) Hegemann, J. D.; Zimmermann, M.; Zhu, S.; Steuber, H.; Harms, K.; Xie, X.; Marahiel, M. A. Xanthomonins I-III: a new class of lasso peptides with a seven-residue macrolactam ring. *Angew. Chem., Int. Ed.* **2014**, *53*, 2230−2234.

(23) Maksimov, M. O.; Link, A. J. Prospecting genomes for lasso peptides. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 333−344.

(24) Cheung, W. L.; Chen, M. Y.; Maksimov, M. O.; Link, A. J. Lasso Peptide Biosynthetic Protein LarB1 Binds Both Leader and Core Peptide Regions of the Precursor Protein LarA. *ACS Cent. Sci.* **2016**, *2*, 702−709.

(25) Hegemann, J. D.; Schwalen, C. J.; Mitchell, D. A.; van der Donk, W. A. Elucidation of the roles of conserved residues in the biosynthesis of the lasso peptide paeninodin. *Chem. Commun.* **2018**, *54*, 9007−9010.

(26) Koos, J. D.; Link, A. J. Heterologous and in Vitro Reconstitution of Fuscanodin, a Lasso Peptide from Thermobifida fusca. *J. Am. Chem. Soc.* **2019**, *141*, 928−935.

(27) Sumida, T.; Dubiley, S.; Wilcox, B.; Severinov, K.; Tagami, S. Structural Basis of Leader Peptide Recognition in Lasso Peptide Biosynthesis Pathway. *ACS Chem. Biol.* **2019**, *14*, 1619−1627.

(28) Martin-Gomez, H.; Tulla-Puche, J. Lasso peptides: chemical approaches and structural elucidation. *Org. Biomol. Chem.* **2018**, *16*, 5065−5080.

(29) Maksimov, M. O.; Link, A. J. Discovery and characterization of an isopeptidase that linearizes lasso peptides. *J. Am. Chem. Soc.* **2013**, *135*, 12038−12047.

(30) Zhu, S.; Fage, C. D.; Hegemann, J. D.; Mielcarek, A.; Yan, D.; Linne, U.; Marahiel, M. A. The B1 Protein Guides the Biosynthesis of a Lasso Peptide. *Sci. Rep.* **2016**, *6*, 35604.

(31) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol.* **2015**, *11*, 564−570.

(32) Elsayed, S. S.; Trusch, F.; Deng, H.; Raab, A.; Prokes, I.; Busarakam, K.; Asenjo, J. A.; Andrews, B. A.; van West, P.; Bull, A. T.; Goodfellow, M.; Yi, Y.; Ebel, R.; Jaspars, M.; Rateb, M. E. Chaxapeptin, a Lasso Peptide from Extremotolerant Streptomyces leeuwenhoekii Strain C58 from the Hyperarid Atacama Desert. *J. Org. Chem.* **2015**, *80*, 10252−10260.

(33) Pan, S. J.; Rajniak, J.; Maksimov, M. O.; Link, A. J. The role of a conserved threonine residue in the leader peptide of lasso peptide precursors. *Chem. Commun. (Cambridge, U. K.)* **2012**, *48*, 1880−1882.

(34) Feng, Z.; Ogasawara, Y.; Nomura, S.; Dairi, T. Biosynthetic Gene Cluster of a d-Tryptophan-Containing Lasso Peptide, MS-271. *ChemBioChem* **2018**, *19*, 2045−2048.

(35) Kaweewan, I.; Hemmi, H.; Komaki, H.; Harada, S.; Kodani, S. Isolation and structure determination of a new lasso peptide specialicin based on genome mining. *Bioorg. Med. Chem.* **2018**, *26*, 6050−6055.

(36) Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. Lasso peptides: an intriguing class of bacterial natural products. *Acc. Chem. Res.* **2015**, *48*, 1909−1919.

(37) Xie, X.; Marahiel, M. A. NMR as an effective tool for the structure determination of lasso peptides. *ChemBioChem* **2012**, *13*, 621−625.

(38) Knappe, T. A.; Linne, U.; Robbel, L.; Marahiel, M. A. Insights into the Biosynthesis and Stability of the Lasso Peptide Capistruin. *Chem. Biol.* **2009**, *16*, 1290−1298.

(39) Okoro, C. K.; Brown, R.; Jones, A. L.; Andrews, B. A.; Asenjo, J. A.; Goodfellow, M.; Bull, A. T. Diversity of culturable actinomycetes in hyper-arid soils of the Atacama Desert, Chile. *Antonie van Leeuwenhoek* **2009**, *95*, 121−133.

(40) Küster, E.; Williams, S. T. Selection of Media for Isolation of Streptomycetes. *Nature* **1964** 202928−929.

(41) Stackebrandt, E. Nucleic Acids and Classification. In *Handbook of New Bacterial Systematics*; Goodfellow, M., O'Donnell, A. G., Eds.; Academic Press, 1993; pp 151−194.

(42) Kim, O. S.; Cho, Y. J.; Lee, K.; Yoon, S. H.; Kim, M.; Na, H.; Park, S. C.; Jeon, Y. S.; Lee, J. H.; Yi, H.; Won, S.; Chun, J. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int. J. Syst. Evol. Microbiol.* **2012**, *62*, 716−721.

(43) Manfio, G. P.; Atalan, E.; Zakrzewska-Czerwinska, J.; Mordarski, M.; Rodriguez, C.; Collins, M. D.; Goodfellow, M. Classification of novel soil streptomycetes as Streptomyces aureus sp. nov., Streptomyces laceyi sp. nov. and Streptomyces sanglieri sp. nov. *Antonie van Leeuwenhoek* **2003**, *83*, 245−255.

(44) Meier-Kolthoff, J. P.; Auch, A. F.; Klenk, H. P.; Goker, M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinf.* **2013**, *14*, 60−2105.

(45) Aziz, R. K.; Bartels, D.; Best, A. A.; DeJongh, M.; Disz, T.; Edwards, R. A.; Formsma, K.; Gerdes, S.; Glass, E. M.; Kubal, M.; Meyer, F.; Olsen, G. J.; Olson, R.; Osterman, A. L.; Overbeek, R. A.; McNeil, L. K.; Paarmann, D.; Paczian, T.; Parrello, B.; Pusch, G. D.; Reich, C.; Stevens, R.; Vassieva, O.; Vonstein, V.; Wilke, A.; Zagnitko, O. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **2008**, *9*, 75−2164.

(46) Skinnider, M. A.; Merwin, N. J.; Johnston, C. W.; Magarvey, N. A. PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res.* **2017**, *45*, W49−W54.

(47) Carver, T.; Harris, S. R.; Berriman, M.; Parkhill, J.; McQuillan, J. A. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* **2012**, *28*, 464−469.

(48) Marchler-Bauer, A.; Derbyshire, M. K.; Gonzales, N. R.; Lu, S.; Chitsaz, F.; Geer, L. Y.; Geer, R. C.; He, J.; Gwadz, M.; Hurwitz, D. I.; Lanczycki, C. J.; Lu, F.; Marchler, G. H.; Song, J. S.; Thanki, N.; Wang, Z.; Yamashita, R. A.; Zhang, D.; Zheng, C.; Bryant, S. H. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **2015**, *43*, D222−6.

(49) Chalkley, R. J.; Baker, P. R.; Huang, L.; Hansen, K. C.; Allen, N. P.; Rexach, M.; Burlingame, A. L. Comprehensive analysis of a multidimensional liquid chromatography mass spectrometry dataset acquired on a quadrupole selecting, quadrupole collision cell, time-of-flight mass spectrometer: II. New developments in Protein Prospector allow for reliable and comprehensive automatic analysis of large datasets. *Mol. Cell. Proteomics* **2005**, *4*, 1194−1204.

(50) Houssen, W. E.; Jaspars, M.. In *Natural Products Isolation*; Sarkar, S., Ed.; Humana Press, 2007; p 380.

(51) Ingebrigtsen, R. A.; Hansen, E.; Andersen, J. H.; Eilertsen, H. C. Light and temperature effects on bioactivity in diatoms. *J. Appl. Phycol.* **2016**, *28*, 939−950.