

Chaxapeptin, a Lasso Peptide from Extremotolerant *Streptomyces leeuwenhoekii* Strain C58 from the Hyperarid Atacama Desert

Somayah S. Elsayed,[†] Franziska Trusch,[‡] Hai Deng,[†] Andrea Raab,[†] Ivan Prokes,[§] Kanungnid Busarakam,^{||} Juan A. Asenjo,[⊥] Barbara A. Andrews,[⊥] Pieter van West,[‡] Alan T. Bull,[∇] Michael Goodfellow,^{||} Yu Yi,[@] Rainer Ebel,[†] Marcel Jaspars,[†] and Mostafa E. Rateb^{*,†,#}

[†]Marine Biodiscovery Centre, University of Aberdeen, Meston Walk, Aberdeen AB24 3UE, Scotland, U.K.

[‡]Aberdeen Oomycetes Laboratory, Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, Scotland, U.K.

[§]Department of Chemistry, University of Warwick, Coventry CV4 7AL, U.K.

^{||}School of Biology, Newcastle University, Newcastle upon Tyne NE1 7RU, U.K.

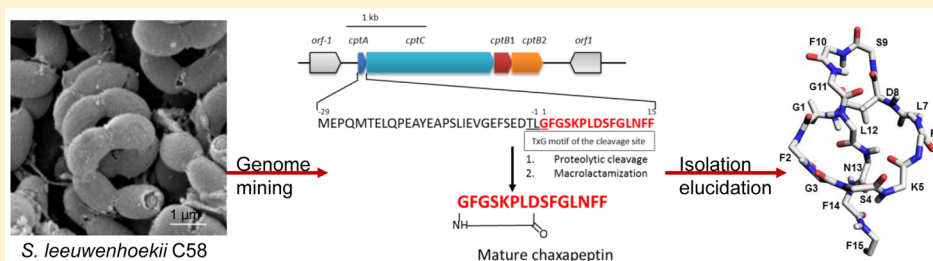
[⊥]Centre for Biotechnology and Bioengineering, CeBiB, University of Chile, Beauchef 850, Santiago, Chile

[∇]School of Biosciences, University of Kent, Canterbury, Kent CT2 7NJ, U.K.

[@]Key Laboratory of Combinatory Biosynthesis and Drug Discovery, School of Pharmaceutical Sciences, Wuhan University, 185 East Lake Road, Wuhan 430071, P. R. China

[#]Pharmacognosy Department, Faculty of Pharmacy, Beni-Suef University, Beni-Suef 32514, Egypt

S Supporting Information



ABSTRACT: Lasso peptides are ribosomally synthesized and post-translationally modified peptides (RiPPs) that possess a unique “lariat knot” structural motif. Genome mining-targeted discovery of new natural products from microbes obtained from extreme environments has led to the identification of a gene cluster directing the biosynthesis of a new lasso peptide, designated as chaxapeptin 1, in the genome of *Streptomyces leeuwenhoekii* strain C58 isolated from the Atacama Desert. Subsequently, 1 was isolated and characterized using high-resolution electrospray ionization mass spectrometry and nuclear magnetic resonance methods. The lasso nature of 1 was confirmed by calculating its nuclear Overhauser effect restraint-based solution structure. Chaxapeptin 1 displayed a significant inhibitory activity in a cell invasion assay with human lung cancer cell line A549.

INTRODUCTION

Traditional screening techniques such as bioassay-guided isolation have become less attractive for the discovery of compounds with novel structural frameworks. Consequently, a number of approaches have been developed, including genome mining,¹ PCR-based strain prioritization,² heterologous expression,³ reactivity-based screening,⁴ and mass spectrometry-based network analysis.⁵ One strategy that reduces the chance of the re-isolation of known molecules is to focus on non- and underexplored taxa or niches.^{6,7} The Atacama Desert is known for its extreme aridity that has persisted for at least ~15 million years.⁸ Some regions in the Atacama Desert have “Mars-like” soils that were once thought to be too extreme for life to exist.⁹ Despite this, novel actinomycetes were recently isolated even from the extreme hyperarid core of the Atacama Desert and were shown to produce new natural products belonging to

different structural classes and possessing various biological activities, including the antimicrobial chaxamycins and chaxalactins isolated from *Streptomyces leeuwenhoekii* C34^T, abenquines from *Streptomyces* sp. DB634, and the antitumor atacamycins from *Streptomyces* sp. C38.^{10–16}

Lasso peptides (LPs) are ribosomally synthesized and post-translationally modified peptides (RiPPs) that possess a unique “lariat knot” structural motif.^{17,18} The post-translational modification comprises an isopeptide bond forming an N-terminal macrolactam ring of seven to nine proteinogenic amino acids into which the remaining C-terminal chain is threaded and divided into a loop above and a tail below the ring.¹⁹ This isopeptide bond is formed between the N-terminal

Received: August 13, 2015

Published: September 24, 2015

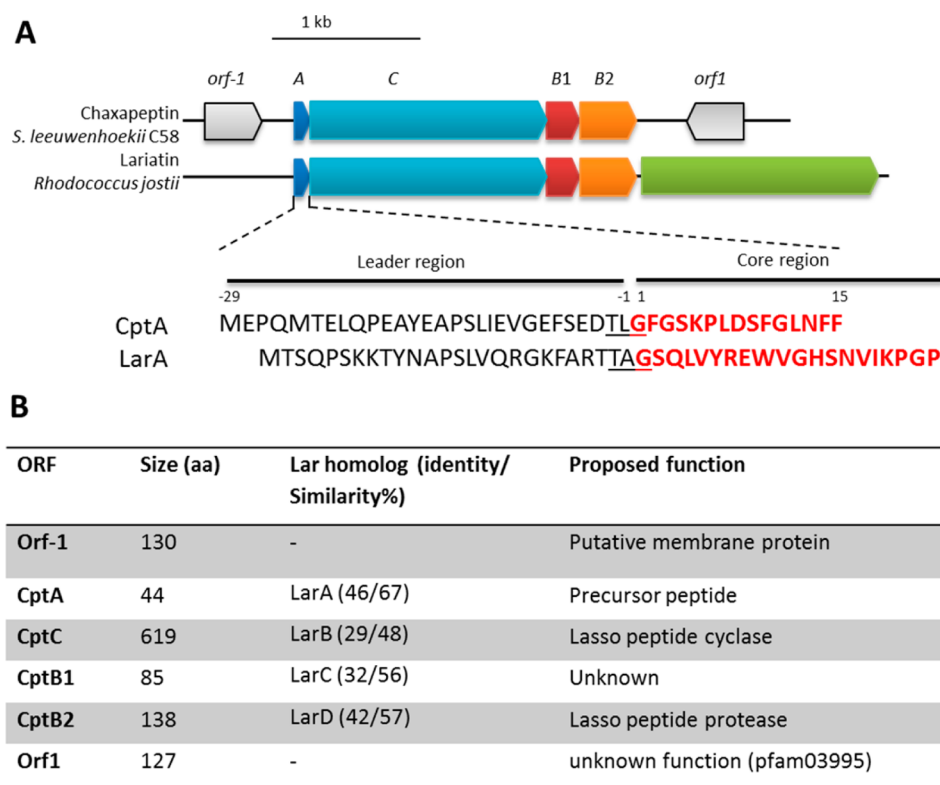


Figure 1. (A) Structure of the *cpt* biosynthetic gene cluster (accession number KT305743). The *cpt* cluster closely resembles the previously characterized *lar* gene cluster for the production of lariatrin in *R. jostii*.²⁷ The biosynthetic genes (*cptA*, *cptB1*, *cptB2*, and *cptC*) for chaxapeptin **1** are shown above together with the *lar*ABCDE genes of the *lar* gene cluster. Boundaries of the *cpt* cluster are defined by two genes, *orf-1* and *orf1*, which are predicted not to be required for the biosynthesis of **1**. (B) Proposed functions of the encoded proteins in the *cpt* cluster compared with those in the *lar* gene cluster.

α -amino group of cysteine, glycine, alanine, or serine and the carboxylic acid side chain of aspartic or glutamic acid. LPs have been classified on the basis of their N-terminal residues and the number of disulfide bridges into three main classes. Class I LPs have an N-terminal Cys and contain two disulfide bridges; class II LPs have an N-terminal Gly and are devoid of disulfide bridges, and class III LPs have an N-terminal Gly and contain one disulfide bridge.²⁰ The solution conformation is stabilized by either the disulfide bond(s) or “plug” amino acids formed in classes I and III, or the presence of “plug” amino acids only as in class II lasso peptides.^{21,22} This three-dimensional (3D) structure led to the assumption that loop folding precedes ring formation in their biosynthesis.²² The function of LPs in bacteria is still unclear, but they are thought to act as scavenging or signaling molecules.²³ They exhibit exceptional stability to proteolytic degradation and to thermal or chemical denaturation, which makes LPs highly attractive as potential scaffolds for epitope grafting.^{24,25}

In the course of targeting the isolation of new natural products possessing novel structural motifs from extreme environments using a genome mining approach, we interrogated the genome of *Streptomyces leeuwenhoekii* strain C58, isolated from Laguna de Chaxa soil in the Atacama Desert, which led to the discovery of a gene cluster directing the biosynthesis of a new lasso peptide, designated as chaxapeptin **1**. Subsequently, **1** was isolated and fully characterized by HRESIMS and NMR methods. Furthermore, the lasso nature of **1** was confirmed by calculating its NOE restraint-based solution structure. In addition to the chemical analyses, the taxonomic status of *Streptomyces* strain C58 and its closely

related isolates was determined in a polyphasic taxonomic study. Finally, an inhibitory effect of **1** on the invasion of human A549 cells was observed.

RESULTS AND DISCUSSION

Bacterial strain C58 was isolated from a soil sample of a hyperarid region in the Atacama Desert collected adjacent to the Laguna de Chaxa, Salar de Atacama, near Tocaño, Chile. Phenotypic and phylogenetic data indicated that isolate C58 represents a new member of the *S. leeuwenhoekii* clade.¹⁰ For full taxonomic identification, see the [supplementary data](#).

In the biosynthesis of lasso peptides, the asparagine synthetase B-like proteins are essential for the maturation of LPs, and their coding genes are commonly found in the biosynthetic gene clusters of LPs. To assess the biosynthetic potential, illumina-based genome scanning of *S. leeuwenhoekii* C58 was conducted and the draft genome of *S. leeuwenhoekii* C58 was annotated on the RAST server.²⁶ A homologue search for LarB, an asparagine synthetase B-like protein responsible for the lasso peptide cyclization in the lariatrin biosynthetic pathway,²⁷ in the annotated genome on the RAST server allowed the identification of an ORF for CptC (Figure 1A) that shares moderate sequence identity (28%) with LarB (Figure 1B). Analysis of the surrounding genetic environment of *cptC* allowed us to retrieve a candidate gene cluster (*cpt*) spanning approximately 2.6 kb (Figure 1A, accession number KT305743). The *cpt* cluster possesses only four biosynthetic genes (*cptAB1B2C*) that are homologous to those known to be involved in the biosynthesis of LPs. The genes *cptAB1B2C* most likely resemble *lar*ABCDE, which is responsible for the

Table 1. NMR Data for Chaxapeptin 1 (4 mM, DMSO- d_6 , 700 MHz, 298 K)^a

residue	position	δ_C/δ_N , type	δ_H , mult. [J (Hz)]	residue	position	δ_C/δ_N , type	δ_H , mult. [J (Hz)]
Gly ₁	α	44.3, CH ₂	a: 3.95, dd (18.0, 4.4) b: 3.26, d (18.0)	Ser ₉	α	59.9, CH	3.78, m
	CO	169.9, C			β	60.5, CH ₂	a: 3.36, m 3.10, m
Phe ₂	NH	113.9, NH	9.23, br s	Phe ₁₀	CO	168.9, C	
	α	56.7, CH	3.77, m		NH	111.6, NH	8.35, br s
	β	33.8, CH ₂	3.33, m		β -OH		4.91, dt (9.1, 4.9)
	γ	139.4, C			α	53.5, CH	4.51, dt (10.3, 4.3)
	δ/δ'	129.0, CH	7.04, d (7.5)		β	38.2, CH ₂	a: 3.14, dd (13.9, 4.3) b: 2.75, m
	ϵ/ϵ'	128.3, CH	7.27, t (7.5)		γ	137.9, C	
	ξ	126.1, CH	7.16, m		δ/δ'	127.8, CH	7.14, t (7.8)
Gly ₃	CO	170.6, C		ϵ/ϵ'	128.0, CH	7.21, d (7.8)	
	NH	108.9, NH	8.55, d (7.2)	ξ	126.1, CH	7.15, t (7.8)	
	α	43.6, CH ₂	a: 4.29, dd (16.8, 8.1) b: 3.07, m	CO	170.5, C		
	CO	167.6, C		Gly ₁₁	NH	112.5, NH	8.88, d (10.3)
NH	104.0, NH	7.74, d (8.1)	α		42.4, CH ₂	a: 4.18, m b: 3.47, m	
Ser ₄	α	55.7, CH	4.42, m		CO	168.1, C	
	β	62.8, CH ₂	a: 3.70, m b: 3.52, m	NH	104.0, NH	7.24, m	
	CO	169.3, C		Leu ₁₂	α	52.4, CH	3.87, m
	NH	112.7, NH	7.41, d (7.8)		β	36.9, CH ₂	a: 1.89, m b: 0.71, m
Lys ₅	β -OH		4.86, t (6.1)		γ	23.7, CH	1.53, m
	α	49.5, CH	4.91, dt (9.1, 4.9)		δ	19.1, CH ₃	0.52, d (6.5)
	β	32.6, CH ₂	a: 1.63, m b: 1.49, m	δ'	19.1, CH ₃	0.82, d (6.5)	
	γ	19.9, CH ₂	a: 1.31, m b: 1.22, m	CO	172.0, C		
	δ	26.4, CH ₂	a: 1.49, m b: 1.24, m	Asn ₁₃	NH	124.1, NH	8.66, d (6.9)
	ϵ	37.9, CH ₂	a: 2.75, m b: 2.60, m		α	49.7, CH	5.07, dt (8.8, 5.4)
	CO	171.2, C			β	40.8, CH ₂	a: 2.64, m b: 2.02, m
Pro ₆	NH	121.2, NH	8.73, d (9.1)	γ	171.9, C		
	ϵ -NH ₂	115.3, NH ₂	a: 8.22, s b: 4.81, br s	CO	172.8, C		
	α	57.8, CH	4.17, m	Phe ₁₄	NH	118.6, NH	9.64, d (8.8)
	β	29.0, CH ₂	a: 2.14, m b: 1.88, m		γ -NH ₂	108.5, NH ₂	a: 7.20, m b: 6.66, d (7.8)
	γ	25.7, CH ₂	a: 2.15, m b: 1.81, m	α	54.9, CH	4.74, q (5.6)	
δ	47.3, CH ₂	a: 3.79, m b: 3.61, m	β	35.8, CH ₂	a: 3.31, m b: 2.87, dd (13.7, 5.6)		
Leu ₇	CO	171.7, C		γ	136.3, C		
	α	53.7, CH	4.56, m	δ/δ'	130.2, CH	7.41, d (7.6)	
	β	43.8, CH ₂	a: 1.81, m b: 1.59, m	ϵ/ϵ'	128.3, CH	7.34, t (7.6)	
	γ	23.2, CH	1.63, m	ξ	126.7, CH	7.29, m	
	δ	22.9, CH ₃	0.79, d (6.5)	CO	167.2, C		
	δ'	23.4, CH ₃	0.73, d (6.5)	Phe ₁₅	NH	117.9, NH	6.87, d (5.6)
	CO	170.7, C			α	55.5, CH	4.21, m
	NH	118.2, NH	6.70, d (5.7)		β	37.9, CH ₂	a: 2.82, m b: 2.70, dd (13.1, 5.1)
Asp ₈	α	49.1, CH	4.62, m	γ	138.4, C		
	β	36.1, CH ₂	a: 3.40, m b: 2.63, m	δ/δ'	129.9, 2CH	6.66, d (7.8)	
	γ	169.9, C		ϵ/ϵ'	129.1, 2CH	7.14, t (7.8)	

Table 1. continued

residue	position	δ_C/δ_N , type	δ_H , mult. [J (Hz)]	residue	position	δ_C/δ_N , type	δ_H , mult. [J (Hz)]
	CO	170.8, C			ξ	125.9, CH	7.20, d (7.8)
	NH	117.9, NH	8.96, d (8.9)		CO	174.5, C	
					NH	118.7, NH	6.64, d (7.8)

^aAssignments based on COSY, TOCSY, HSQC, and HMBC experiments.

production of lariat in *Rhodococcus jostii* K01-B0171.²⁷ The *cptA* gene encodes the 44-residue precursor peptide (CptA), comprised of a 29-amino acid N-terminal leader sequence and the 15-amino acid C-terminal core peptide, with the characteristic TxG motif for lasso peptide precursors, necessary for removal of the leader peptide by the B proteins (lasso peptide proteases).^{17,28,29} CptB1 is homologous to a protein of unknown function and shares moderate sequence identity (32%) with LarC, while CptB2 is a homologue (42% sequence identity) of LarD. LarC and D resemble the N-terminal and C-terminal domains of full length B proteins in other lasso peptide systems, respectively.³⁰ To use the recent community consensus, we, therefore, annotated the homologues of LarC and D as CptB1 and CptB2, respectively.^{23,30,31} The upstream and downstream genes of *cptAB1B2C*, namely, *orf-1* and *orf1*, respectively, are not found in other lasso peptide gene clusters. Therefore, we propose that the four ORFs between them define the boundaries of the *cpt* cluster. No export/self-immunity gene encoding an ABC transporter could be identified, which is in line with recent studies that also demonstrated a lack of immunity genes in the gene clusters responsible for production of lasso peptides.^{23,24,31} The bulky side chains of two Phe residues (Phe₁₄ and Phe₁₅) are predicted to sterically entrap the C-terminal tail.

To determine whether the identified gene cluster indeed directs the biosynthesis of chaxapeptin **1**, *S. leeuwenhoekii* C58 cultivation was optimized in different media (see the Supporting Information for details) and expression of **1** was monitored by LC–MS using its predicted mass. **1** was traced in two of five media tested, ISP2 with a titer of 1.0 mg L⁻¹ and modified ISP2 with a titer of 0.3 mg L⁻¹ (Figure S5). Large scale fermentation in ISP2 medium, extraction, and purification using a series of chromatographic techniques, including liquid–liquid partitioning, Sephadex LH-20, and subsequent HPLC, led to the isolation of a pure sample of **1**.

The molecular formula of chaxapeptin **1** was established as C₇₉H₁₀₇N₁₇O₂₀ based on the combined analyses of HRESIMS of the molecular ion and doubly charged ion peaks as well as ¹H, ¹³C, and ¹H–¹⁵N HSQC NMR spectra, which also were in agreement with the predicted core peptide in the *cpt* cluster. The cyclic nature of chaxapeptin was supported by the 35 degrees of unsaturation required by its molecular formula. NMR spectra were recorded on a 4.0 mM sample of **1** that led to a well-resolved set of signals. The combined analysis of one-dimensional (Table 1) and detailed two-dimensional NMR spectral data obtained from ¹H–¹³C HSQC, COSY, TOCSY, and HSQC–TOCSY experiments allowed the full assignment of individual amino acids in **1** and revealed a 15-amino acid peptide, including four phenylalanines (Phe), three glycines (Gly), two serines (Ser), two leucines (Leu), one proline (Pro), one aspartic acid (Asp), one asparagine (Asn), and one lysine (Lys) (Figure S13).

Sequence connectivity between different amino acids was established on the basis of the analysis of the observed ¹H–¹³C HMBC and NOESY correlations (Figure 2). The connectivity

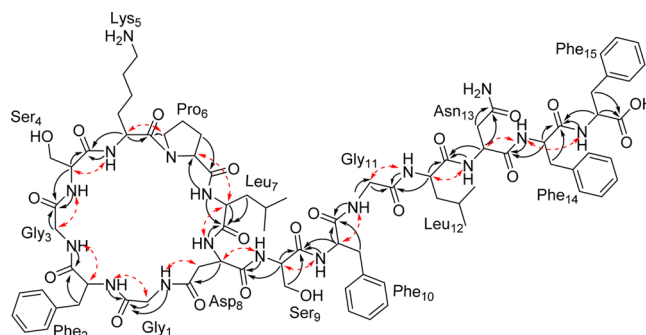


Figure 2. Key HMBC (fx1) and NOESY (fx2) correlations confirming the chaxapeptin **1** peptide sequence and indicating the point of cyclization.

of Gly₁–Phe₂ was evident from the HMBC correlations of Gly₁–H₂α, Gly₁–NH, and Phe₂–NH to Gly₁–CO and confirmed by the NOE between Gly₁–H₂α and Phe₂–NH. Leu₇–Asp₈ and Gly₁₁–Leu₁₂ were assigned in a similar fashion. The connection between Phe₂–Gly₃ was established through the HMBC correlations of Phe₂–H₂β and Gly₃–NH to Phe₂–CO and the NOE between Phe₂–H_α and Gly₃–NH. Ser₉–Phe₁₀ and Phe₁₀–Gly₁₁ were obtained in a similar fashion. Additionally, the connectivity of Gly₃–Ser₄ was established through the HMBC correlations of Gly₃–H₂α, Ser₄–H_α, and Ser₄–NH to Gly₃–CO and the NOE between Gly₃–H₂α and Ser₄–NH. Ser₄–Lys₅ and Leu₁₂–Asn₁₃ connectivities were assigned in a similar fashion. Furthermore, the connectivity of Lys₅–Pro₆ was established through the HMBC correlations of Lys₅–H_α and Pro₆–H_δ to Lys₅–CO and the NOE between Lys₅–H_α and Pro₆–H_δ. Pro₆–Leu₇ was established through the HMBC correlation of Leu₇–NH/Pro₆–C_α, and the NOE between Pro₆–H_α and Leu₇–H_α. The connectivities of Asp₈–Ser₉ and Asn₁₃–Phe₁₄ could not be established through HMBC alone as the HMBC correlations to the carbonyl carbons of Asp₈ and Asn₁₃ were only from the NH and α-protons of Ser₉ and Phe₁₄, respectively, with no correlations from any of the Asp₈ and Asn₁₃ protons, respectively. Thus, these connectivities were confirmed through NOE correlations Ser₉–NH/Asp₈–H_α and Phe₁₄–NH/Asn₁₃–H_α. The remaining Phe₁₄–Phe₁₅ connectivity was established through the HMBC correlations of Phe₁₄–H₂β, Phe₁₄–H_α, Phe₁₅–NH, and Phe₁₅–H_α to Phe₁₄–CO carbon and NOESY correlation Phe₁₄–NH/Phe₁₅–NH. Finally, the NOE between Gly₁–NH and Asp₈–H_β established the isopeptide bond between Gly₁ and Asp₈, which forms the characteristic macrolactam ring of lasso peptides.

HRESIMS/MS analysis of chaxapeptin **1** did not show the full sequence identified by NMR but only confirmed the sequence of the peptide tail (Figure 3). Thus, the planar structure of **1** was defined as a cyclic isopeptide containing an eight-amino acid ring and a remaining seven-amino acid side chain. The macrocyclic ring of **1** proved to be highly stable because no fragmentation could be observed under standard or vigorous MS conditions. Moreover, **1** was found to be

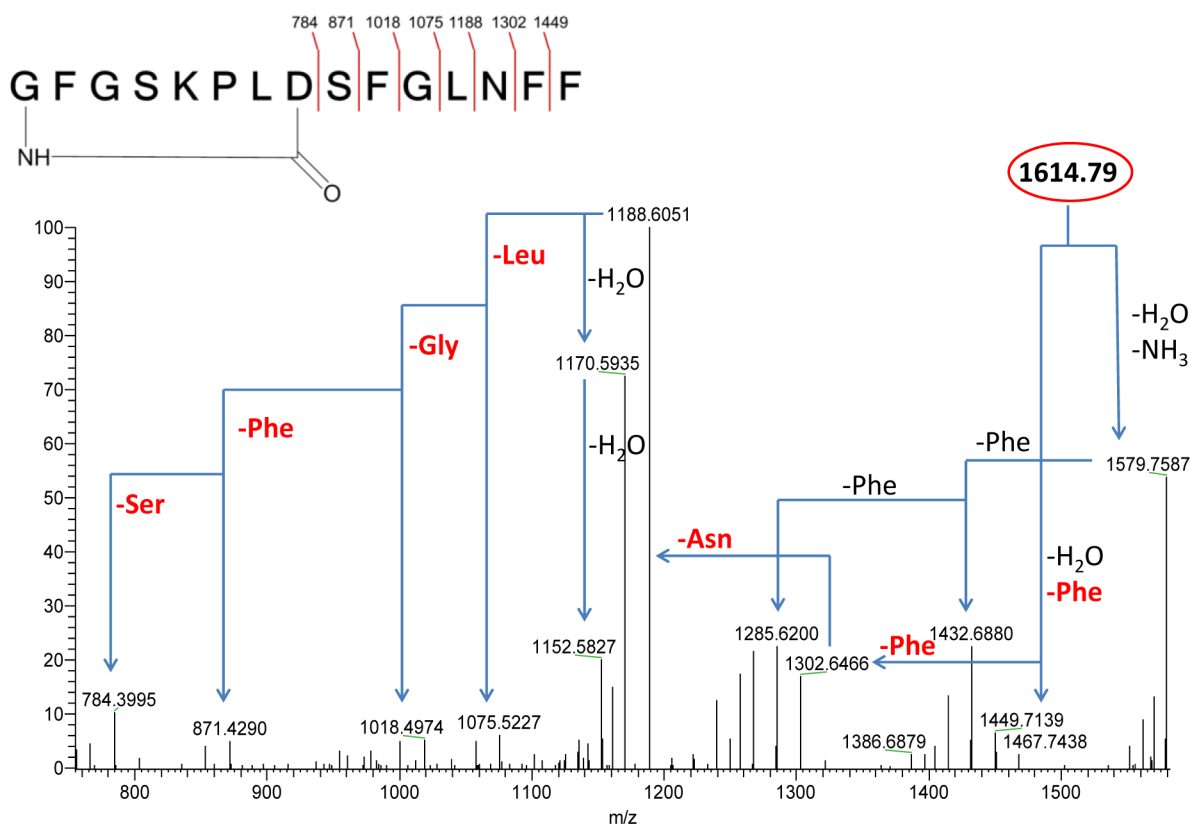


Figure 3. MS² data for m/z 1614.8 ($[M + H]^+$) in chaxapeptin 1. Only peptide tail fragments can be detected while the macrocycle is stable against MS fragmentation.

completely resistant to proteolytic degradation when assessed against different proteases (trypsin, chymotrypsin, pepsin, and papain) under standard conditions.

Pro₆ in chaxapeptin 1 adopted a *trans* peptide bond, which was established through the NOESY correlation between Lys₅-H α (δ_H 4.91) and Pro₆-H δ (δ_H 3.79 and 3.61), together with the absence of a correlation between Lys₅-H α and Pro₆-H α (δ_H 4.17). To establish the absolute configuration of all amino acid residues, the advanced Marfey's method was used.³² The retention times obtained for FDLA-derivatized hydrolyzed amino acid residues of 1 as compared to the retention times of FDLA-derivatized standard amino acids allowed the assignment of an L configuration to all the amino acids as expected from the gene cluster. Further investigation of the NOESY spectrum of 1 showed more correlations than the sequence-confirming ones expected for a cyclic branched peptide, with the additional correlations suggesting that the chain of 1 was threaded through the macrocyclic ring (Figure 4).

The complete assignment of all proton resonances in chaxapeptin 1 and available distance information from the NOESY spectra were used to determine its solution structure at 25 °C. We converted NOE cross-peak volumes into distance constraints with CYANA2.1, and 156 unambiguous distance constraints (average of <10 distance constraints per residue) were obtained, including 34 NOEs for the backbone, 46 sequential NOEs, 14 medium-range NOEs, and 62 NOEs for long-range interactions (Figure S3a,b). In addition, we predicted dihedral angles from C α resonances with CYANA2.1. To avoid the artificial enforcing of the lasso structure, we refrained from inserting a linkage between Gly₁ and Asp₈ but used the observed NOE distance constraints between both residues. The NOE and dihedral angle-based

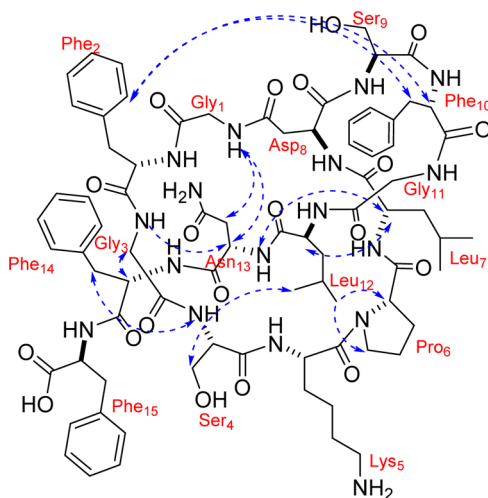


Figure 4. Key NOESY correlations for chaxapeptin 1 showing the chain part folding through the ring.

constraints were used for structure calculations with CYANA2.1 by simulated annealing starting with 200 random conformers. The averaged backbone atom root-mean-square deviation (rmsd) of the 15 conformers with the lowest energy values was 0.2 Å (Figure 5a; for the Ramachandran plot, see Figure S3c). All structures adopt the typical lasso peptide fold. The corresponding structures had no violations for NOE constraints, van der Waals bonds, or angles and were used for further processing. To realize the ring structure, we inserted a single *in trans* bond between the nitrogen of Gly₁ and the side chain C γ (CG) of Asp₈ with YASARA and subsequently

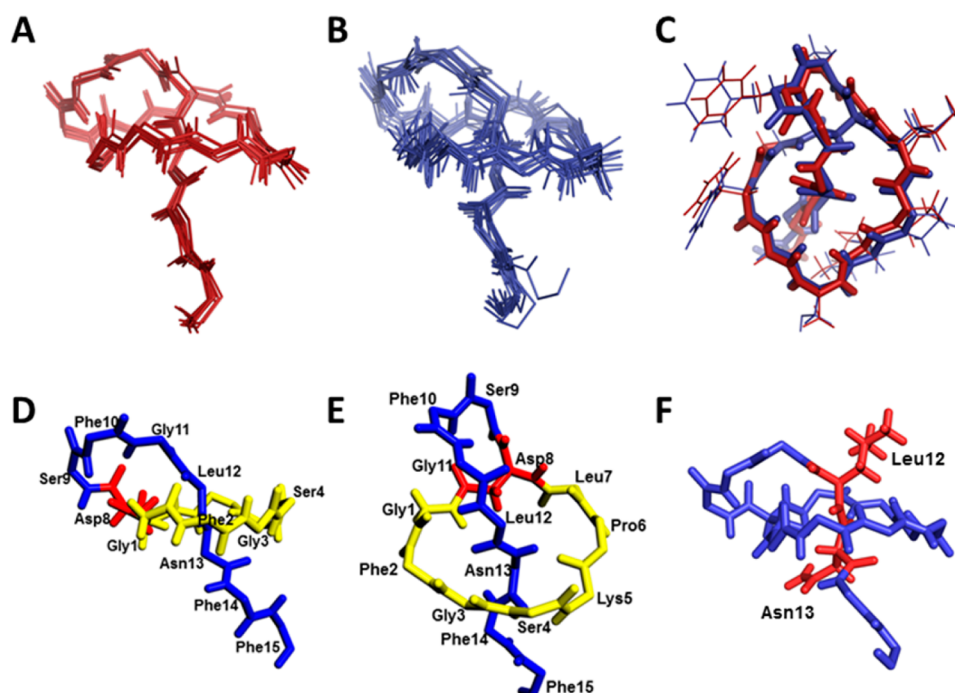


Figure 5. NMR-based structure of chaxapeptin 1. (A) Superimposition of the 15 lowest-energy NMR structures. (B) Superimposition of the 15 structures after insertion of the isopeptide bond and energy minimization. (C) Top view superimposition of the lowest-energy NMR structure without the peptide bond (red) and the structure after insertion of the isopeptide bond and subsequent energy minimization (blue). (D and E) Side and top views, respectively, of **1** belonging to the class II lasso peptides. Isopeptide bond-forming Asp₈ is colored red; the ring-forming residues are colored yellow, and the tail is colored blue. (F) Plug amino acids (Leu₁₂ and Asn₁₃, red) of **1** above and below the ring in opposite directions.

performed an energy minimization step with a YASARA2 force field to reduce any tension due to the insertion of the isopeptide bond (Figure 5b). By comparison of the chaxapeptin structure without the isopeptide bond [NMR structure of the lowest-energy model (Figure 5c, red)] with the structure after insertion of the isopeptide bond and energy minimization (Figure 5c, blue), the high level of identity of backbone atoms and the small differences in the side chain atoms validated the accuracy of our approach. The 15 energy-minimized conformers showed an average rmsd of 0.3 Å and are kept to represent the solution structure of **1** at 25 °C (Protein Data Bank entry 2n5c and BMRB entry 25704). The 3D structure of **1** confirmed a class II lasso peptide motif (Figure 5d,e) with the usual right-handed topology.^{19–21} The structure comprises an eight-residue ring with a four-residue loop above and a three-residue tail below the macrolactam ring. Furthermore, the high stability of the peptide may be due to Leu₁₂ and Asn₁₃ operating as the plug amino acids located directly above and below the ring, respectively, in opposite orientations (Figure 5f). Additionally, **1** possesses hydrogen bonds in the loop region (Asp₈-Ser₉ and Asp₈-Gly₁₁) and between residues in the tail and the ring (Lys₅-Asn₁₃ and Ser₄-Asn₁₃), which results in a characteristic short β -sheet known to result in the observed high stabilities of lasso peptides.³³ Furthermore, the structure of **1** does not show any flexible regions, which is probably due to the short tail below the ring in comparison to other lasso peptides with longer tails.^{18,28}

With the structure of chaxapeptin **1** in hand, screens for its antimicrobial effects were performed. Chaxapeptin **1** was found to show a weak antibacterial effect against the Gram-positive strains, *Staphylococcus aureus* and *Bacillus subtilis* (30–35 $\mu\text{g mL}^{-1}$), and no activity against the Gram-negative *Escherichia coli* (maximal concentration up to 50 $\mu\text{g mL}^{-1}$). This lack of

activity may be explained by the absence of the ABC-transporter protein in the gene cluster of *S. leuwenhoekii* C58,³⁰ which confers immunity by the export of the mature lasso peptide. In addition, no significant growth inhibition on the oomycete *Saprolegnia parasitica* by **1** could be observed (data not shown).

The sequence as well as the structure of chaxapeptin **1** is similar to that of another class II lasso peptide, sungsanpin (GFGSKPIDSFGLSWL),³⁴ although both are produced by actinomycetes recovered from completely different habitats (the Atacama Desert and the deep-sea sediment collected off the Korean coast, respectively). From a taxonomic point of view, the sungsanpin-producing actinomycete, *Streptomyces* strain SNJ013, is only distantly related to *S. leuwenhoekii* C58, and the two strains share a 16S rRNA gene sequence similarity of 94%. Both LPs share an eight-residue ring, a four-residue loop, and a three-residue tail. Interestingly, three of four nonidentical residues are located in the tail (residues 13–15, SWL vs NFF). Moreover, both peptides share the same glycine and hydrophobic contents, while the aromatic contents of **1** and sungsanpin were different (27 and 20%, respectively). The high degree of sequence similarity observed for **1** and sungsanpin (alignment score of 73.3) can be regarded as unique among class II LPs, which contain the highest number of identified LPs to date. Previously, sungsanpin had been demonstrated to exhibit an inhibitory activity in a cell invasion assay with human lung cancer cell line A549.³⁴ To determine whether the changes in the sequence had an effect on biological activity, **1** was tested with the same invasion assay using A549 cells (Figure 6).

The concentration-dependent inhibition of A549 cell invasion through the extracellular matrix (ECM) by chaxapeptin **1** was comparable to that of sungsanpin (28 and 25% at

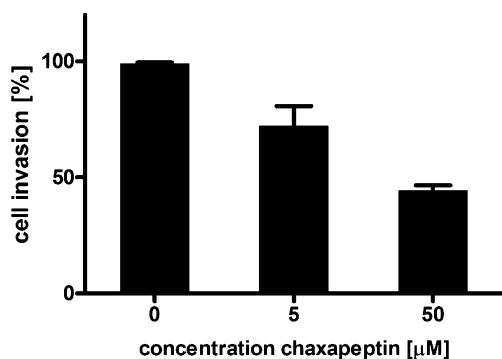


Figure 6. Effect of chaxapeptin **1** on invasion of A549 cells. Cells were incubated for 36 h with the given concentrations. Control cells were treated with 0.7% (v/v) DMSO. Cell invasion was quantified by staining invaded cells after lysis with CyQuant dye (excitation at 428 nm, emission at 520 nm). Each bar shows the standard deviation about the mean of three independent experiments.

5 µM and 56 and 53% at 50 µM for **1** and sungsanpin, respectively) without affecting the cell morphology (data not shown). Hence, the difference in the tail of both lasso peptides does not affect their inhibitory effect on cell invasion of A549 cells. Because **1** and sungsanpin have very similar effects, we abstained from further detailed investigation.

Recent approaches to the discovery of new drug candidates from microbial sources proved to be efficient, and their application using intelligently selected strains makes them more successful. In the search for new sources of natural products, a group of actinomycete strains was isolated from the poorly explored hyperarid Atacama Desert, whose environmental conditions were previously regarded as unfavorable for the presence of life. Some of the isolated extremotolerant actinomycetes have been shown to produce interesting new natural products,^{12–16} and the study presented here confirms the importance of the Atacama Desert environments for biodiscovery campaigns. Genome mining has led to the discovery of chaxapeptin **1**, a new member of the structurally unique group of lasso peptides. Chaxapeptin belongs to class II of the LPs and was intriguingly found to share a high degree of sequence similarity with the previously isolated class II lasso peptide sungsanpin despite being isolated from completely different habitats. Recently, several lasso peptides have been discovered. However, they were mainly produced by proteobacteria with only three examples from actinomycetes,¹⁷ and therefore, the study presented here helps in a broader understanding of lasso peptides.

EXPERIMENTAL SECTION

General Experimental Procedures. Optical rotations were recorded using a polarimeter. UV and IR spectra were measured on a UV/vis spectrometer and a FT/IR spectrometer, respectively. NMR data were acquired on 600 and 700 MHz spectrometers. High-resolution mass spectrometric data were obtained using a LTQ Orbitrap coupled to the HPLC system (PDA detector, PDA autosampler, and pump). The following conditions were used: capillary voltage of 45 V, capillary temperature of 260 °C, auxiliary gas flow rate of 10–20 arbitrary units, sheath gas flow rate of 40–50 arbitrary units, spray voltage of 4.5 kV, and mass range of 100–2000 amu (maximal resolution of 30000). For LC/MS, a C₁₈ analytical HPLC column (5 µm, 4.6 mm × 150 mm) was used with a mobile phase of 0 to 100% MeOH over 30 min at a flow rate of 1 mL min⁻¹. Preparative HPLC separations were conducted using a C₁₈ column (5

µm, 100 Å, 10 mm × 250 mm), connected to a binary pump, and monitored using a photodiode array detector.

Microbial Taxonomy. Isolate C58 and associated strains C38 and C79 were isolated from a hyperarid soil sample collected next to the Laguna de Chaxa, Salar de Atacama, near Tocaño, Chile (23°17'S/68°10'W) by one of us (A.T.B.) in November 2008. Strains C58 and C79 were isolated on raffinose-histidine agar, and strain C38 was isolated on glucose-yeast extract agar [both media were supplemented with the antifungal antibiotics cycloheximide and nystatin (each at 25 µg mL⁻¹)] using the procedure described by Okoro and colleagues.¹⁰ In addition, the strains were shown to have morphological properties consistent with their assignment to the genus *Streptomyces* and to be members of the novel *S. leeuwenhoekii* 16S rRNA gene clade that encompassed additional strains isolated from arid Atacama Desert soils. It was also shown that *Streptomyces* strain SNJ013 is only distantly related to strains assigned to the *S. leeuwenhoekii* 16S rRNA clade.

Fermentation, Extraction, and Isolation. Different media were used for small scale fermentation to screen for the production of chaxapeptin **1** (Table S1). The seed culture of *S. leeuwenhoekii* strain C58 [EU551701 (Figures S1 and S2)] was prepared by inoculating 50 mL of liquid ISP2 medium with a single colony of the organism and incubating for 3 days at 30 °C with shaking at 180 rpm. The primary culture was used to inoculate 250 mL of each of the five media for initial screening of the production of **1**. Incubation of the secondary culture was conducted for 7 days at 30 °C with shaking at 180 rpm. At the end of the incubation period, cultures were centrifuged (3000 rpm for 20 min) where the cell mass was washed with distilled water and extracted with MeOH while HP-20 resin was added to the culture media and shaken for 6 h and then filtered, washed with distilled water, and extracted with MeOH. Each of the cell mass and culture medium extracts was subjected to LC–HRESIMS analysis.

For large scale production, the seed culture of *S. leeuwenhoekii* strain C58 was prepared as described above. Then this primary culture was used to inoculate 4 L of ISP2 medium as a production culture. Incubation of the secondary culture was conducted for 7 days at 30 °C with shaking at 180 rpm as before.

Culture broths were centrifuged (3000 rpm for 20 min), and the cell mass was washed with distilled water and extracted with MeOH (4 × 500 mL). The successive MeOH extracts were combined and concentrated under reduced pressure to 250 mL of total extract. This extract was fractionated successively with *n*-hexane (3 × 250 mL), CH₂Cl₂ (3 × 300 mL), and then EtOAc (3 × 250 mL). Each solvent fraction was evaporated *in vacuo* and subjected to LC–MS and ¹H NMR analysis, which revealed that the CH₂Cl₂ fraction is the one of interest to follow. This CH₂Cl₂ fraction was loaded on a Sephadex LH-20 column (1:150 sample:Sephadex ratio) equilibrated with a MeOH/CH₂Cl₂ mixture (1:1) ~24 h before application of the fraction with a flow rate of 0.5 mL min⁻¹, and two 150 mL fractions were collected, evaporated *in vacuo*, and monitored by ¹H NMR analysis. The second fraction (30 mg) was purified by semipreparative HPLC using a gradient of MeOH in H₂O (50 to 100% over 30 min) followed by 100% MeOH for 10 min at a flow rate of 1.25 mL min⁻¹ to afford pure compound **1** (3.5 mg), which was acquired at a *t*_R of 35.5 min at 25 °C using 220 and 254 nm as monitoring collection wavelengths.

Mass Spectrometric Analysis. Mass spectrometric data were obtained using an LTQ Orbitrap MS system and the same parameters stated earlier in the general experimental procedures. Different fragmentation attempts were made with and without in-source fragmentation using three different voltages in the CID mode at activation Q of 0.25 and an activation time of 30 ms: (1) 10–30 V normalized collision energy (NCE), no fragmentation observed; (2) 35–45 V NCE, fragmentation observed for the tail but not the ring; (3) 50–100 V NCE, similar to previous, but no additional fragmentation observed for the ring. Attempts (MS³) to further fragment the MS² fragments at *m/z* 1075.5 and 871.4 were also unsuccessful at 40–50 V NCE and trap fill times between 50 and 500 ms. Other attempts using the PQD mode were not successful at different energies and trapping times.

Proteolytic Stability Assay. To assess the proteolytic stability of chaxapeptin **1**, assays were performed using 10 μg of **1** and different proteases (trypsin, chymotrypsin, pepsin, and papain) under standard conditions described in manufacturers' information sheets. For trypsin, the assay was performed with 0.5 μg of trypsin in a buffer containing 50 mM Tris-HCl and 1 mM CaCl_2 at pH 7.6 for 6 h at 37 °C. For chymotrypsin, the assay was performed with 0.5 μg of chymotrypsin in a buffer containing 100 mM Tris-HCl and 10 mM CaCl_2 at pH 8.0 for 6 h at 37 °C. For pepsin, the assay was performed with 1.0 mg of pepsin in 0.1 M HCl at pH 3.0 for 18 h at 25 °C. For papain, the assay was performed with 1.0 mg of papain in 0.1 M sodium acetate buffer at pH 5.0 for 18 h at 25 °C. At the end of the incubation times, the samples were subjected to LC–HRESIMS analysis.

Chaxapeptin **1**: yellowish white powder, $[\alpha]_{\text{D}}^{20} - 21$ (c 0.1, CH_3OH); UV (MeCN) λ_{max} (log ϵ) 219 (3.74), 258 (1.78) nm; IR (film) ν_{max} 3295, 2963, 1648, 1532, 1242 cm^{-1} ; ^1H and ^{13}C NMR data in Table 1; HRESIMS m/z 1614.7926 $[\text{M} + \text{H}]^+$ (calcd for $\text{C}_{79}\text{H}_{108}\text{O}_{20}\text{N}_{17}$, m/z 1614.7951, $\Delta = -1.5$ ppm).

Determination of the Absolute Configurations of the Amino Acid Residues of 1. Chaxapeptin **1** (0.2 mg) was hydrolyzed in 0.4 mL of 6 N HCl at 110 °C for 24 h. The reaction mixture was subsequently cooled, the solvent evaporated *in vacuo*, and residual HCl completely removed by the addition of 0.4 mL of water and removal of the solvent three times. The hydrolysate was completely dried under N_2 for 24 h. The hydrolysate was dissolved in 50 μL of H_2O . To 50 μL of a 50 mM standard amino acid solution (aqueous) or the hydrolysate were added 20 μL of 1 M NaHCO_3 and 100 μL of 1% FDFA in acetone. The reaction mixtures were incubated at 40 °C for 1 h, with frequent mixing. After the mixtures had been cooled at room temperature, the reactions were quenched by the addition of 10 μL of 2 N HCl. The samples were diluted by increasing the volume to 1 mL with MeOH. A 5 μL aliquot of each sample was analyzed by LC–MS using a SunFire C_{18} column (5 μm , 100 \AA , 4.6 mm \times 150 mm), eluted with a solvent gradient of 10 to 80% CH_3CN in H_2O containing 0.1% formic acid over 45 min, at a flow rate of 1 mL min^{-1} . Comparison of the retention times of derivatized hydrolyzed amino acids with the retention times of derivatized D- and L-amino acid standards revealed L configuration for all chaxapeptin **1** amino acid residues.

NMR Solution Structure of Chaxapeptin 1. NMR-based structure calculation was performed using CYANA 2.1.³⁵ Manually picked NOESY cross-peaks from a ^1H – ^1H NOESY spectrum of chaxapeptin **1** were automatically assigned by CYANA and used for calibration and successive calculation of distance constraints, while torsion angles were determined from $^{13}\text{C}\alpha$ chemical shifts of each residue. Distance and angle constraints were used for seven cycles of combined automated NOESY assignment and structure calculation, followed by a final structure calculation. For each structure, 200 conformers were calculated using the standard simulated annealing schedule with 10000 torsion angle dynamics steps per conformer. Structural statistics of the 20 lowest-target function value conformers are listed in Table S2. The 15 conformers with the lowest final target function values, holding already the characteristic lasso peptide structure, were used for final cyclization of the peptide with YASARA WhatIf (version 11.12.31). For this, one $\text{O}\delta$ (OD) atom of Asp₈ and the HN of Gly₁ were deleted and a single bond was inserted *in trans* between CG of Asp₈ and N of Gly₁ to mimic the isopeptide bond of the lasso peptide. The length of the bond was corrected by energy minimization with YASARA2. Finally, to test the stability of **1** and to detect possible hydrogen bonds, molecular dynamics (MD) simulations were performed with Amber03 over 10 ns in a cubic box (40 \AA 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 \AA , and PME for long-range Coulomb forces.

Antibacterial Screening. The antibacterial activity of **1** was evaluated against Gram-positive strains *St. aureus* ATCC 25923 and *B. subtilis* NCTC 2116 and Gram-negative strain *E. coli* ATCC 25922 using the agar diffusion method and regression line analysis.³⁶ Filter paper disks containing ampicillin (10 μg) were used as positive controls. MICs against species were calculated using the method described previously with minor modifications.³⁷ In brief, tested strains

were grown in Müller-Hinton (MH) broth³⁸ to early stationary phase and then diluted to an OD_{600} of 0.005. The assays were performed in a 96-well microtiter plate format in duplicate, with two independent cultures for each strain. All compounds were dissolved in DMSO (Sigma) and added to the cultures in wells so that the final concentration of DMSO was 10% that did not affect the growth of any of the tested strains. The effect of different dilutions of **1** on the growth was assessed after incubation at 37 °C for 18 h using a Labsystems iEMS reader MF plate reader at OD_{620} . The MIC was determined as the lowest concentration showing no growth compared to the MH broth control.

Cell Culture. Human lung cancer cell line A549 was purchased from the American Type Culture Collection (ATCC). Cells were grown in DMEM supplemented with 10% fetal bovine serum, 200 units mL^{-1} penicillin, and 200 $\mu\text{g mL}^{-1}$ streptomycin and incubated at 37 °C in a humidified atmosphere with 5% CO_2 .

Cell Invasion Assay. Cell invasion of A549 cells in the presence and absence of chaxapeptin **1** was investigated with a cell invasion assay kit according to the manufacturer's protocol. In brief, cells (0.5×10^6 cells mL^{-1}) were seeded into the upper ECM gel-coated chamber in 300 μL of serum-free medium, containing the indicated chaxapeptin **1** concentrations. For chemotraction, the lower chamber was filled with 500 μL of DMEM containing 10% FBS. After incubation for 36 h (extended to increase the signal-to-noise ratio), invasive cells on the bottom of the ECM membrane were detached and lysed successively and stained with CyQuant for quantitative analysis; 100 μL of each sample was used to measure the CyQuant dye emission at 520 nm with a Synergy HT at 27 °C. Absolute values were normalized to the control sample with 0.7% DMSO and corrected by the buffer background. Statistical analysis was conducted with GraphPad Prism 5. Each bar shows the standard deviation about the mean of three independent experiments.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.joc.5b01878.

Experimental procedures, strain taxonomy, HR-ESIMS, detailed NMR table, and spectra of **1** (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: m.rateb11@aberdeen.ac.uk. Telephone: +44 1224 274536.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the College of Physical Sciences, University of Aberdeen, for provision of infrastructure and facilities in the Marine Biodiscovery Centre. S.S.E. thanks the Egyptian Government for the Ph.D. scholarship. K.B. is grateful for a scholarship from the Thai Royal Government, A.T.B. for support from the Royal Society (International Joint Project Grant JP100654), and M.G. for an Emeritus Fellowship from the Leverhulme Trust. J.A.A., B.A.A., and A.T.B. thank Conicyt for the Basal Centre Grant for the CeBiB (Project FB0001). We are indebted to Dr. Byung-Yong Kim (ChunLab, Seoul, Korea) for conducting the fatty acid analyses. F.T. thanks Prof. Peter Bayer for technical advice.

■ REFERENCES

(1) Doroghazi, J. R.; Albright, J. C.; Goering, A. W.; Ju, K. S.; Haines, R. R.; Tchalukov, K. A.; Labeda, D. P.; Kelleher, N. L.; Metcalf, W. W. *Nat. Chem. Biol.* **2014**, *10*, 963–968.

- (2) Hindra, T.; Huang, Y.; Yang, D.; Rudolf, J. D.; Xie, P.; Xie, G.; Teng, Q.; Lohman, J. R.; Zhu, X.; Huang, Y.; Zhao, L.-X.; Jiang, Y.; Duan, Y.; Shen, B. *Nat. Prod.* **2014**, *77*, 2296–2303.
- (3) Feng, Z. Y.; Kim, J. H.; Brady, S. F. *J. Am. Chem. Soc.* **2010**, *132*, 11902–11903.
- (4) Cox, C. L.; Tietz, J. I.; Sokolowski, K.; Melby, J. O.; Doroghazi, J. R.; Mitchell, D. A. *ACS Chem. Biol.* **2014**, *9*, 2014–2022.
- (5) Nguyen, D. D.; Wu, C. H.; Moree, W. J.; Lamsa, A.; Medema, M. H.; Zhao, X. L.; Gavilan, R. G.; Aparicio, M.; Atencio, L.; et al. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, E2611–E2620.
- (6) Pidot, S. J.; Coyne, S.; Kloss, F.; Hertweck, C. *Int. J. Med. Microbiol.* **2014**, *304*, 14–22.
- (7) Bull, A. T. In *Extremophiles Handbook*; Horikoshi, K., Antranikian, G., Bull, A. T., Robb, F., Stetter, K., Eds.; Springer-Verlag: Tokyo, 2011; Vol. 2, pp 1203–1240.
- (8) Gómez-Silva, B.; Rainey, F. A.; Warren-Rhodes, K. A.; McKay, C. P.; Navarro-González, R. Atacama Desert Soil Microbiology. In *Microbiology of Extreme Soils, Soil Biology*; Dion, P., Nautiyal, C. S., Eds.; Springer: Berlin, 2008; Vol. 13, pp 117–132.
- (9) Navarro-González, R.; Rainey, F. A.; Molina, P.; Bagaley, D. R.; Hollen, B. J.; De La Rosa, J.; Small, A. M.; Quinn, R. C.; Grunthaner, F. J.; Cáceres, L.; Gómez-Silva, B.; McKay, C. P. *Science* **2003**, *302*, 1018–1021.
- (10) Okoro, C. K.; Brown, R.; Jones, A. L.; Andrews, B. A.; Asenjo, J. A.; Goodfellow, M.; Bull, A. T. *Antonie van Leeuwenhoek* **2009**, *95*, 121–133.
- (11) Bull, A. T.; Asenjo, J. A. *Antonie van Leeuwenhoek* **2013**, *103*, 1173–1179.
- (12) Rateb, M. E.; Houssen, W. E.; Arnold, M.; Abdelrahman, M. H.; Deng, H.; Harrison, W. T. A.; Okoro, C. K.; Asenjo, J. A.; Andrews, B. A.; Ferguson, G.; Bull, A. T.; Goodfellow, M.; Ebel, R.; Jaspars, M. *J. Nat. Prod.* **2011**, *74*, 1491–1499.
- (13) Rateb, M. E.; Houssen, W. E.; Harrison, W. T. A.; Deng, H.; Okoro, C. K.; Asenjo, J. A.; Andrews, B. A.; Bull, A. T.; Goodfellow, M.; Ebel, R.; Jaspars, M. *J. Nat. Prod.* **2011**, *74*, 1965–1971.
- (14) Busarakam, K.; Bull, A. T.; Girard, G.; Labeda, D. P.; van Wezel, G. P.; Goodfellow, M. *Antonie van Leeuwenhoek* **2014**, *105*, 849–861.
- (15) Schulz, D.; Beese, P.; Ohlendorf, B.; Erhard, A.; Zinecker, H.; Dorador, C.; Imhoff, J. F. *J. Antibiot.* **2011**, *64*, 763–768.
- (16) Nachtigall, J.; Kulik, A.; Helaly, S.; Bull, A. T.; Goodfellow, M.; Asenjo, J. A.; Maier, A.; Wiese, J.; Imhoff, J. F.; Süßmuth, R. D.; Fiedler, H.-P. *J. Antibiot.* **2011**, *64*, 775–780.
- (17) Zimmermann, M.; Hegemann, J. D.; Xie, X.; Marahiel, M. A. *Chem. Sci.* **2014**, *5*, 4032–4043.
- (18) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; et al. *Nat. Prod. Rep.* **2013**, *30*, 108–160.
- (19) Hegemann, J. D.; Zimmermann, M.; Zhu, S.; Steuber, H.; Harms, K.; Xie, X.; Marahiel, M. A. *Angew. Chem., Int. Ed.* **2014**, *53*, 2230–2234.
- (20) Maksimov, M. O.; Link, A. J. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 333–344.
- (21) Xie, X.; Marahiel, M. A. *ChemBioChem* **2012**, *13*, 621–625.
- (22) Oman, T. J.; Van Der Donk, W. A. *Nat. Chem. Biol.* **2010**, *6*, 9–18.
- (23) Maksimov, M. O.; Link, A. J. *J. Am. Chem. Soc.* **2013**, *135*, 12038–12047.
- (24) Maksimov, M. O.; Pan, S. J.; Link, A. J. *Nat. Prod. Rep.* **2012**, *29*, 996–1006.
- (25) Knappe, T. A.; Manzenrieder, F.; Mas-Moruno, C.; Linne, U.; Sasse, F.; Kessler, H.; Xie, X.; Marahiel, M. A. *Angew. Chem., Int. Ed.* **2011**, *50*, 8714–8717.
- (26) Blin, K.; Medema, H. M.; Kazempour, D.; Fischbach, M. A.; Breitling, R.; Takano, E.; Weber, T. *Nucleic Acids Res.* **2013**, *41*, W204–W212.
- (27) Inokoshi, J.; Matsuhama, M.; Miyake, M.; Ikeda, H.; Tomoda, H. *Appl. Microbiol. Biotechnol.* **2012**, *95*, 451–60.
- (28) Knappe, T. A.; Linne, U.; Robbel, L.; Marahiel, M. A. *Chem. Biol.* **2009**, *16*, 1290–1298.
- (29) Pan, S. J.; Rajniak, J.; Maksimov, M. O.; Link, A. J. *Chem. Commun.* **2012**, *48*, 1880–1882.
- (30) Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. *Acc. Chem. Res.* **2015**, *48*, 1909–1919.
- (31) Hegemann, J. D.; Zimmermann, M.; Zhu, S.; Klug, D.; Marahiel, M. A. *Biopolymers* **2013**, *100*, 527–542.
- (32) Fujii, K.; Ikai, Y.; Oka, H.; Suzuki, M.; Harada, K. *Anal. Chem.* **1997**, *69*, 5146–5151.
- (33) Iwatsuki, M.; Tomoda, H.; Uchida, R.; Gouda, H.; Hirono, S.; Omura, S. *J. Am. Chem. Soc.* **2006**, *128*, 7486–7491.
- (34) Um, S.; Kim, Y.; Kwon, H.; Wen, H.; Kim, S.; Kwon, H. C.; Park, S.; Shin, J.; Oh, D. *J. Nat. Prod.* **2013**, *76*, 873–879.
- (35) Herrmann, T.; Guntert, P.; Wuthrich, K. *J. Mol. Biol.* **2002**, *319*, 209–227.
- (36) Kronvall, G. *J. Clin. Microbiol.* **1983**, *17*, 975–980.
- (37) Domenech, P.; Kobayashi, H.; LeVier, K.; Walker, G. C.; Barry, C. E. *J. Bacteriol.* **2009**, *191*, 477–485.
- (38) Müller, H. J.; Hinton, J. *Exp. Biol. Med.* **1941**, *48*, 330–333.