



Heterologous Expression of a Cryptic Gene Cluster from *Streptomyces leeuwenhoekii* C34^T Yields a Novel Lasso Peptide, LEEPTEIN

Juan Pablo Gomez-Escribano,^a Jean Franco Castro,^{a,b,c*} Valeria Razmilic,^{a,b} Scott A. Jarmusch,^c Gerhard Saalbach,^a Rainer Ebel,^c Marcel Jaspars,^c Barbara Andrews,^b Juan A. Asenjo,^b Mervyn J. Bibb^a

^aDepartment of Molecular Microbiology, John Innes Centre, Norwich, United Kingdom

^bCentre for Biotechnology and Bioengineering (CeBiB), Department of Chemical Engineering and Biotechnology, Universidad de Chile, Santiago, Chile

^cMarine Biodiscovery Centre, Department of Chemistry, University of Aberdeen, Old Aberdeen, United Kingdom

ABSTRACT Analysis of the genome sequence of *Streptomyces leeuwenhoekii* C34^T identified biosynthetic gene clusters (BGCs) for three different lasso peptides (Lp1, Lp2, and Lp3) which were not known to be made by the strain. Lasso peptides represent relatively new members of the RiPP (ribosomally synthesized and posttranslationally modified peptides) family of natural products and have not been extensively studied. Lp3, whose production could be detected in culture supernatants from *S. leeuwenhoekii* C34^T and after heterologous expression of its BGC in *Streptomyces coelicolor*, is identical to the previously characterized chaxapeptin. Lp1, whose production could not be detected or achieved heterologously, appears to be identical to a recently identified member of the citrulassin family of lasso peptides. Since production of Lp2 by *S. leeuwenhoekii* C34^T was not observed, its BGC was also expressed in *S. coelicolor*. The lasso peptide was isolated and its structure confirmed by mass spectrometry and nuclear magnetic resonance analyses, revealing a novel structure that appears to represent a new family of lasso peptides.

IMPORTANCE Recent developments in genome sequencing combined with bioinformatic analysis have revealed that actinomycetes contain a plethora of unexpected BGCs and thus have the potential to produce many more natural products than previously thought. This reflects the inability to detect the production of these compounds under laboratory conditions, perhaps through the use of inappropriate growth media or the absence of the environmental cues required to elicit expression of the corresponding BGCs. One approach to overcoming this problem is to circumvent the regulatory mechanisms that control expression of the BGC in its natural host by deploying heterologous expression. The generally compact nature of lasso peptide BGCs makes them particularly amenable to this approach, and, in the example given here, analysis revealed a new member of the lasso peptide family of RiPPs. This approach should be readily applicable to other cryptic lasso peptide gene clusters and would also facilitate the design and production of nonnatural variants by changing the sequence encoding the core peptide, as has been achieved with other classes of RiPPs.

KEYWORDS *Streptomyces*, natural products, RiPP, lasso peptide, heterologous expression

A group of strains belonging to the bacterial genus *Streptomyces*, subsequently named *Streptomyces leeuwenhoekii*, were isolated from a sample collected from the Chaxa Lagoon in the Salar de Atacama, an extreme environment located in northern Chile (1, 2). These strains have been shown to produce a range of previously unknown

Citation Gomez-Escribano JP, Castro JF, Razmilic V, Jarmusch SA, Saalbach G, Ebel R, Jaspars M, Andrews B, Asenjo JA, Bibb MJ. 2019. Heterologous expression of a cryptic gene cluster from *Streptomyces leeuwenhoekii* C34^T yields a novel lasso peptide, leepeptin. *Appl Environ Microbiol* 85:e01752-19. <https://doi.org/10.1128/AEM.01752-19>.

Editor Andrew J. McBain, University of Manchester

Copyright © 2019 Gomez-Escribano et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Mervyn J. Bibb, mervyn.bibb@jic.ac.uk.

* Present address: Jean Franco Castro, Instituto de Investigaciones Agropecuarias (INIA) Quilamapu, Chillán, Chile.

J.F.C., V.R., and S.A.J. contributed equally to this article.

Received 5 August 2019

Accepted 17 September 2019

Accepted manuscript posted online 27 September 2019

Published 14 November 2019

natural products. For example, *S. leeuwenhoekii* C38 produces the polyketide atacamycin (3); *S. leeuwenhoekii* C58 produces the lasso peptide chaxapeptin (4); *S. leeuwenhoekii* C79 produces four new specialized metabolites (H.-P. Fiedler, unpublished results); and the type species, *S. leeuwenhoekii* C34^T (2), produces two groups of polyketides: the chaxamycins (5) and the chaxalactins (6). The chaxamycins and the chaxalactins both possess antibacterial activity, while chaxamycin also displayed potential anticancer activity by inhibiting the chaperone protein Hsp90; the atacamycins show inhibitory activity toward phosphodiesterase PDE-4B2 (and thus may have anti-inflammatory potential) and, in the case of atacamycin A, antiproliferative activity against adeno carcinoma and breast carcinoma cells; chaxapeptin showed inhibitory activity toward human lung cancer cell line A549. Thus, despite few studies, it appears that streptomycetes (and probably other *Actinobacteria*) isolated from this extreme and little-scrutinized ecological niche are likely to prove to be an abundant source of new natural products with a range of potentially useful biological activities.

Our previous analysis of the genome sequence of *S. leeuwenhoekii* C34^T revealed the presence of 35 biosynthetic gene clusters (BGCs), including 3 encoding the putative lasso peptides Lp1, Lp2, and Lp3 (7, 8). Lasso peptides have been isolated from a wide range of bacteria, including *Burkholderia*, *Caulobacter*, *Enterobacteriaceae*, *Rhodococcus*, and *Streptomyces* species derived from terrestrial, freshwater, and marine environments (see reference 9 and references therein), although recent bioinformatic analysis suggests a much broader phylogenetic distribution (10). They represent a growing class of bioactive bacterial peptides, forming a subfamily within the much larger RiPP (ribosomally synthesized and posttranslationally modified peptides) family of natural products (11), and are characterized by a specific knotted structure, the lasso fold (12). Lasso peptides are produced initially as precursor peptides with an N-terminal leader sequence and a C-terminal retained core region of 14 to 24 amino acid residues (13) that contains a macrocyclic ring formed between the carboxylate side chain of an Asp or Glu residue located at position 7, 8, or 9 of the core peptide and the amino group of the N-terminal amino acid residue. While it was initially thought that there was a requirement for Gly to be the first core peptide residue, examples of lasso peptides with N-terminal Ala, Ser, and Cys have also been reported, and a recent bioinformatic analysis revealed even greater flexibility at this position (10). The C-terminal tail is threaded through the macrocycle, presumably prior to ring closure, and is generally retained in place sterically by the presence of C-terminal residues with bulky side chains (although disulfide bonds play that role in some lasso peptides). The resulting highly compact structure confers thermal stability and resistance to proteases and results in a broad range of biological activities that include enzyme inhibition and receptor antagonism; some lasso peptides possess antibiotic activity (13). The unusual lasso structure, which has thus far proved inaccessible to peptide synthesis, has also attracted attention as a scaffold for epitope grafting (14), providing the opportunity to develop novel lasso peptide-derived biological activities with therapeutic potential.

Lasso peptide biosynthesis requires a lasso cyclase (C-protein), homologous to asparagine synthase, and a leader peptidase (B-protein), homologous to transglutaminase. Biosynthesis also requires a RiPP recognition element (RRE [an E-protein]); note that we have adopted the nomenclature used in reference 10 for this protein, rather than the previously used B1) which binds the leader peptide and directs enzymatic modification. In most cases (an exception is provided in this work), lasso peptide BGCs also encode transporters for product export (15).

Here, we describe the analysis of three lasso peptide gene clusters of *S. leeuwenhoekii* C34^T and the heterologous expression of two of them, leading to the chemical characterization of one of the molecules, leepeptin, as a novel lasso peptide.

RESULTS

Identification of three lasso peptide gene clusters in the *S. leeuwenhoekii* C34^T genome sequence. Three putative lasso peptide BGCs were identified in the genome sequence of *S. leeuwenhoekii* C34^T (7). These BGCs were not identified at the time by

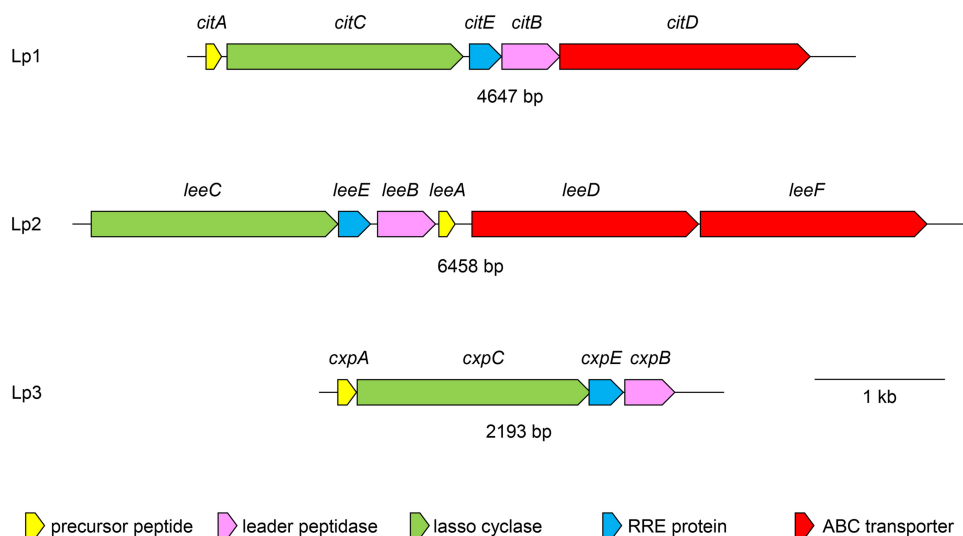


FIG 1 Schematic representation of the three lasso peptide gene clusters of *S. leeuwenhoekii* C34^T. Since the product of the Lp1 gene cluster appears to be identical to one of the members of the citrulassin family of lasso peptides, the genes have been named *cit*; similarly, since the product of the Lp3 gene cluster was shown to be chaxapeptin, the genes have been named *cxp*.

antiSMASH (16) but were revealed by searching initially for homologues of the lasso cyclase, LarB, involved in lariatin biosynthesis in *Rhodococcus jostii* (GenBank accession number [BAL72547.1](#)) (17) using blastP and the deduced proteome of *S. leeuwenhoekii* C34^T. Manual inspection of the genome sequences adjacent to identified homologues revealed, in three cases, all of the enzymatic functions required for the biosynthesis of three different lasso peptides. Moreover, the lack of substantial intergenic regions suggested that all of the biosynthetic genes for each peptide might be contained in a single transcription unit. A schematic representation of the BGCs is shown in Fig. 1. While the BGCs for Lp1 and Lp2 are chromosomally located (GenBank accession number [LN831790](#)), that for Lp3 occurs on pSLE2 (GenBank accession number [LN831789](#)) and thus is potentially capable of lateral transfer by conjugation.

Identification of Lp3 as chaxapeptin. Supernatants of *S. leeuwenhoekii* C34^T cultures grown in five different media were analyzed by matrix-assisted laser desorption ionization–time of flight (MALDI-ToF) mass spectrometry (MS) for the production of the expected peptides. When this work was initiated in early 2014, with the exception of one example of the use of Cys (10), Gly was the only residue known to be found at the N terminus of lasso core peptides; this apparent constraint influenced our predictions of the size of the mature peptides made by each of the three BGCs. The Lp1 BGC encodes the precursor peptide MKKAYEAPTLVRLGSRFRRTGLLQSRSGNDRLLILSKN (the core residues that we predicted, indicated in italics, would yield a molecule with an accurate mass of 1,764.9959 Da after cyclization). The Lp2 BGC encodes the precursor peptide MEHDEKTPYETPAVYGLGAFAEETGLYGVNRNDEEINWHFDYWT (the initially predicted core sequence, in italics, would yield a molecule with an accurate mass of 2,395.0658 Da after cyclization). The Lp3 BGC encodes the precursor peptide MTELQP EAYEAPSLIEVGEFSEDTLGFSGKPLDSFGLNFF (the predicted core residues, in italics, would yield a molecule with an accurate mass of 1,613.7879 Da after cyclization). Ions with masses matching those expected for Lp3 were readily detected and in high abundance (see Fig. S1 in the supplemental material); we could not detect any ions corresponding to the expected peptides for Lp1 and Lp2 or ions that could correspond to slightly different peptides (e.g., one residue more or less if we had not predicted the core peptide correctly).

Since Lp3 was readily detected in most of the culture media tested, we focused our initial attention on this peptide. Sequencing of the peptide using liquid chromatography–tandem high-resolution mass spectrometry (LC/tandem-HRMS) on a Synapt G2Si

platform (Waters) revealed the expected sequence of the C-terminal tail and identified the only aspartate in the core peptide as the residue likely to be involved in the macrocyclization (Fig. S2), as we had predicted (we could not detect fragmentation of the macrocycle, and so the last a, b, and y ions detected indicated the residue likely involved in ring formation with the N-terminal NH₂). These results, shown in more detail in the supplemental material, led us to conclude that Lp3 is identical to chaxapeptin produced by the closely related strain *S. leeuwenhoekii* C58 (4). Lp3 is also similar to sungsanpin (11/15 identical core residues), which is produced by a marine actinomycete, *Streptomyces* sp. strain SNJ013, that had been isolated from deep sea sediment collected off the Korean coast (9). A blastP search of the GenBank database also revealed homologues of Lp3 in the genome sequences of both *Streptomyces kanamyceticus* (GenBank accession number [WP_107099006](#)) and *Streptomyces cinnamoneus* (accession number [WP_104531718](#)) (Fig. S3).

Heterologous expression of the BGCs in *S. coelicolor*. The BGC for Lp3 is unusual in that it does not contain any gene encoding a possible transporter; it was unclear to us whether Lp3 could be secreted by ABC transporters encoded by the other two lasso peptide BGCs or by a nonspecific transporter encoded somewhere else in the genome. In contrast, the Lp2 BGC has two genes for possible ABC transporters, while that for Lp1 encodes a single ABC transporter. The three BGCs were cloned in the expression vector pIJ10257 (18) downstream from the constitutive *ermE** promoter (19) and were transferred to *Streptomyces coelicolor* expression hosts M1152 and M1154 (20) by conjugation (see the supplemental material for details).

We could readily detect Lp3 in culture supernatants of the heterologous hosts containing the Lp3 expression construct pIJ12815 by both MALDI-ToF and liquid chromatography–ion trap–time of flight mass spectrometry (LC-IT-ToF MS) (Fig. S4); the molecule had the same accurate mass and peptide sequence (determined by LC/tandem-HRMS on a Waters Synapt G2Si platform) as Lp3 produced by *S. leeuwenhoekii* C34^T. Since *S. coelicolor* does not contain any lasso peptide BGCs, this result suggests that export of Lp3 is not dependent on a dedicated lasso peptide transporter. We also transferred pIJ12815 to *Streptomyces albus* J1074 and found that we could also readily detect production of mature Lp3 in the culture supernatant (Fig. S4). These results suggest that the secretion of mature Lp3 can be mediated by a nonspecific transporter generally present in *Streptomyces* species.

The Lp2 BGC is also unusual in that the precursor peptide gene (*leeA*) lies downstream of the genes encoding biosynthetic activities (*leeCEB*) (Fig. 1). Our first expression construct contained just *leeCEBA* transcribed from the constitutive *ermE** promoter; transfer to heterologous hosts *S. coelicolor* M1152 and M1154 failed to yield the expected molecule or one of similar mass (data not shown). We then added the two putative ABC transporter genes *leeDF*, reconstructing the native gene arrangement (*leeCEBADF*; see Materials and Methods and Fig. S5 for details). Transfer of the resulting pIJ12819 plasmid to *S. coelicolor* M1152 and M1154 resulted in the production of a molecule (2,338.0440 Da) that was 57 Da lower in mass than that we had predicted for Lp2 (2,395.0658 Da) (Fig. 2a). Structure elucidation of this molecule (see below) resolved this anomaly and confirmed that the observed molecule was the product of the Lp2 BGC (Fig. S11). These results indicated that, in contrast to Lp3 (chaxapeptin), the two ABC transporters encoded within the BGC are required for peptide secretion (we did not test whether one of the genes would suffice).

Attempts to detect production of Lp1 by both *S. leeuwenhoekii* C34^T (MALDI-ToF analyses only; see Fig. S1) and *S. coelicolor* after introduction of the BGC (MALDI-ToF and LC-IT-ToF MS analyses; data not shown) failed. However, multiple blastP searches of the GenBank database performed using the predicted amino sequences of the Lp1 biosynthetic enzymes and ABC transporter as query sequences identified several *Streptomyces* strains that could potentially produce a similar lasso peptide. Four of these strains (*Streptomyces viridochromogenes* DSM 40736, *Streptomyces davaonensis* JCM 4913, *Streptomyces pristinaspiralis* ATCC 25486, and *Streptomyces roseochromoge-*

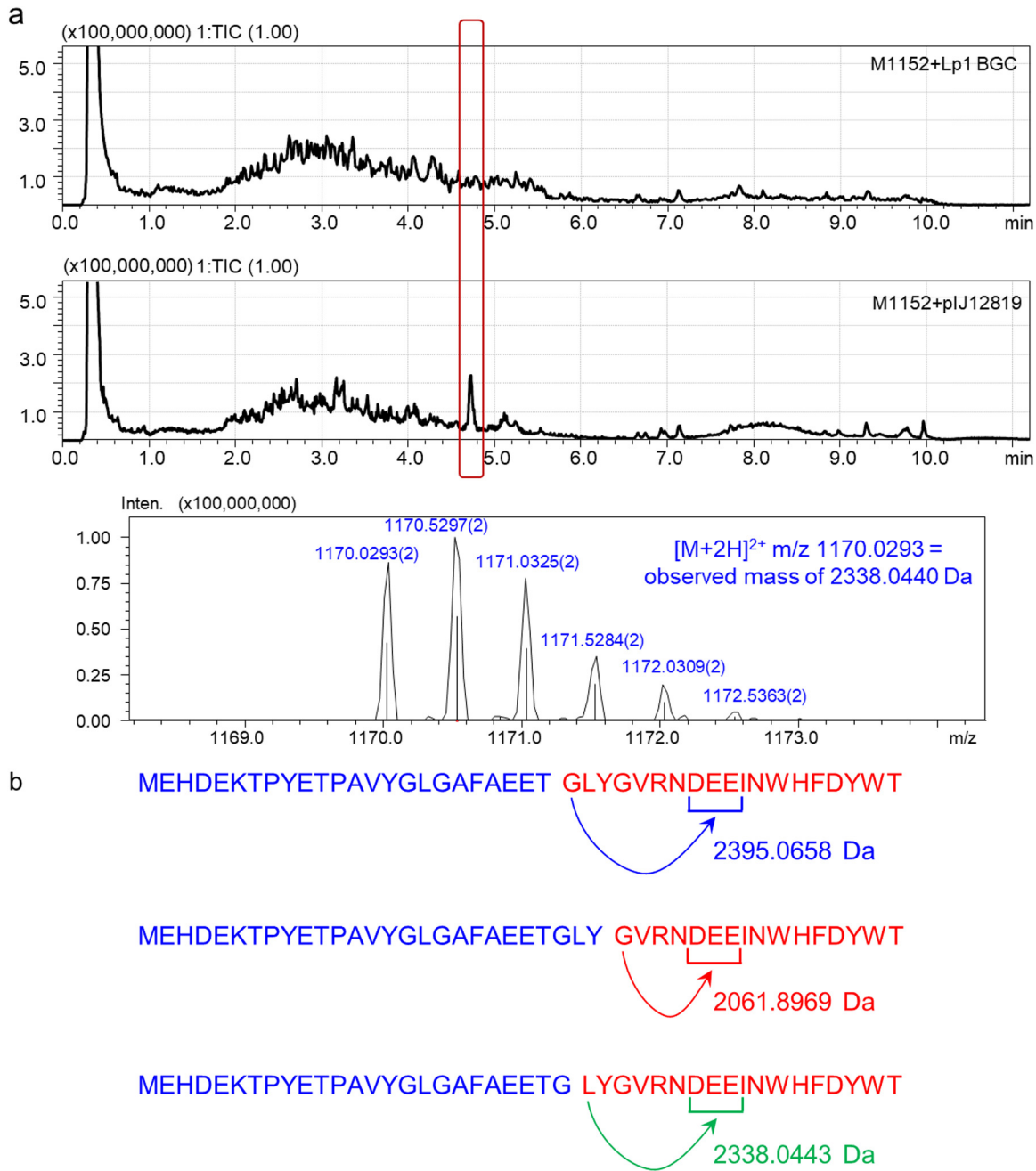


FIG 2 (a) Heterologous expression of the Lp2 BGC (contained in pJ12819) in *S. coelicolor* M1152. Total ion chromatogram (TIC) and zoomed-in full spectrum of the highlighted peak obtained by LC-IT-ToF MS (Shimadzu). The detected ion, $[M + 2H]^{2+}$, with 1,170.0293 m/z suggests an observed neutral mass of 2,338.0440 Da for the monoisotopic molecule. (b) Possible core peptides deduced from the nucleotide sequence of the *leeA* precursor peptide gene. Core peptides are shown in red and leader sequences in blue; the arrows indicate possible cyclization events. (c) Ions expected and observed after LC/tandem-HRMS analysis of leepeptin. The table of expected ions was calculated for the linear peptide after subtraction of a water molecule from the ions represented in panels b and a (using <http://db.systemsbio.net:8080/proteomicsToolkit/FraglonServlet.html>). The ions that were clearly detected are shaded in orange. The predicted structure of leepeptin is shown in red at the bottom, with the green arrow indicating cyclization between Leu₁ and Glu₈ (the leader sequence is shown in blue). Seq., sequence.

nes subsp. *oscitans* DS 12.976) were grown in a variety of different media, but we could not detect any possible candidate peptides in culture supernatants. Since these attempts were made, it has become apparent that the predicted core peptide of Lp1 is identical to that of one of the members of the citrulassin family of lasso peptides identified in reference 10. Interestingly, the study reported in reference 10 demon-

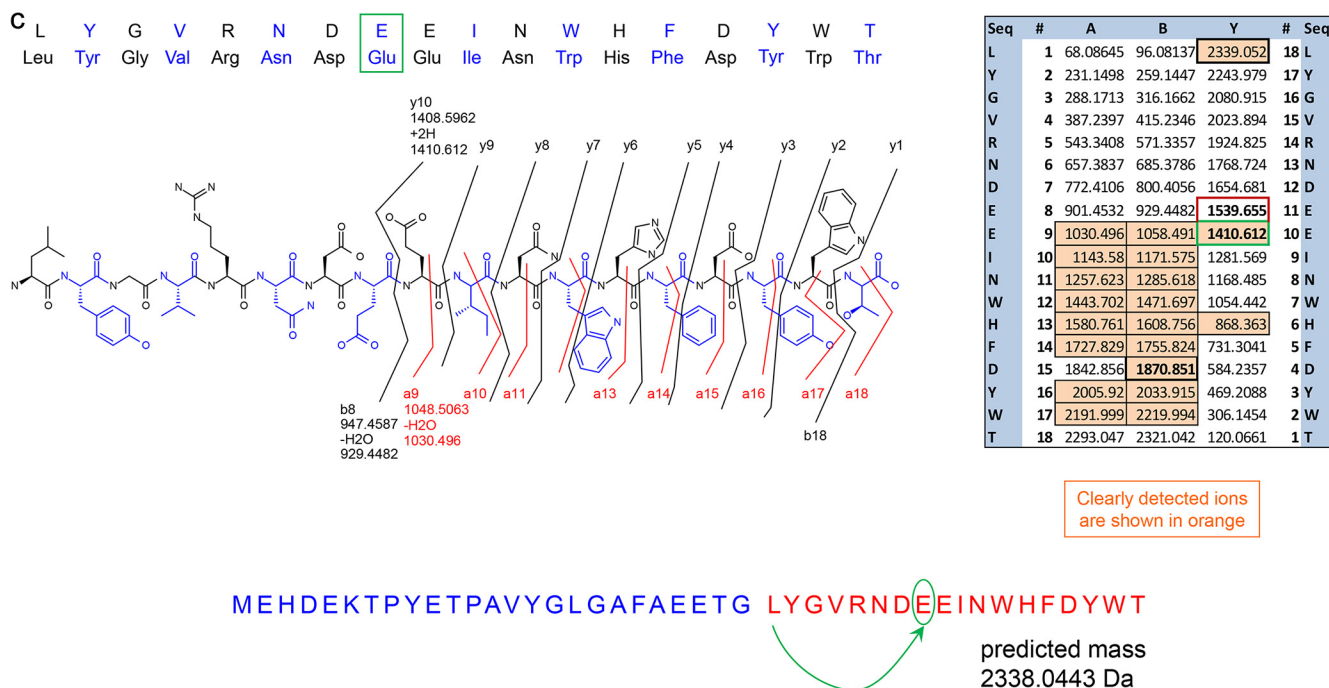


FIG 2 (Continued)

strated that for family member citrulassin A, Arg9 of the core sequence (LLGLAGNDR LVLSKN) is modified to citrulline. Arg9 is conserved across the large citrulassin family of lasso peptides (10), suggesting that this unusual modification may occur in many if not all family members. As those published data established Lp1 as a member of a previously identified class of lasso peptide, we abandoned further attempts to express its BGC heterologously.

Characterization of Lp2 by liquid chromatography—high-resolution mass spectrometry (LC-HRMS) and LC/tandem-HRMS. Heterologous expression of the Lp2 BGC in *S. coelicolor* produced a molecule with a mass of 2,338.0440 Da (the monoisotopic ion detected by LC-IT-ToF MS was 1,170.0293 m/z for $[M + 2H]^{2+}$) (Fig. 2a). This molecule was present in culture supernatants from all *S. coelicolor* M1152 and M1154 derivatives containing the Lp2 BGC but absent in samples from the host without the BGC (in Fig. 2a, the negative control is M1152 carrying the Lp1 BGC). The Lp2 precursor peptide (MEHDEKTPYETPAVYGLGAF A EETGLYGV R N D E E I N W H F D Y W T) could yield several possible core peptides (Fig. 2b). At the time when we started this project in early 2014, with the single exception of Cys (10), Gly was the only amino acid known to be used to form the macrocyclic ring; consequently, we predicted two possible core peptides, namely, GLYGV R N D E E I N W H F D Y W T, which would yield a molecule with an accurate mass of 2,395.0658 Da after cyclization, and GVRNDEEINWHFDYWT, which would yield a molecule with an accurate mass of 2,061.8970 Da after cyclization. However, analysis of the molecule detected by LC-IT-ToF MS indicated that the core peptide was likely to be LYGV R N D E E I N W H F D Y W T. This matched, within the accepted range of instrument error, the predicted mass (2,338.0443 Da after cyclization) to the observed mass (2,338.0440 Da). The partial sequence of the peptide was observed by LC/tandem-HRMS on a Synapt G2Si platform (Waters); all the expected a, b, and y ions for the 10 C-terminal residues were clearly identified in the deconvoluted fragmentation spectrum (Fig. 2c; see also Fig. S6 and S7) confirming that the heterologously produced peptide was the product of the Lp2 BGC. We named this new natural product, with (at the time) an unprecedented N-terminal Leu, leepeptin.

The LC/tandem-HRMS fragmentation data also allowed us to identify the internal acidic residue likely to be involved in macrocyclization. Three acidic residues could

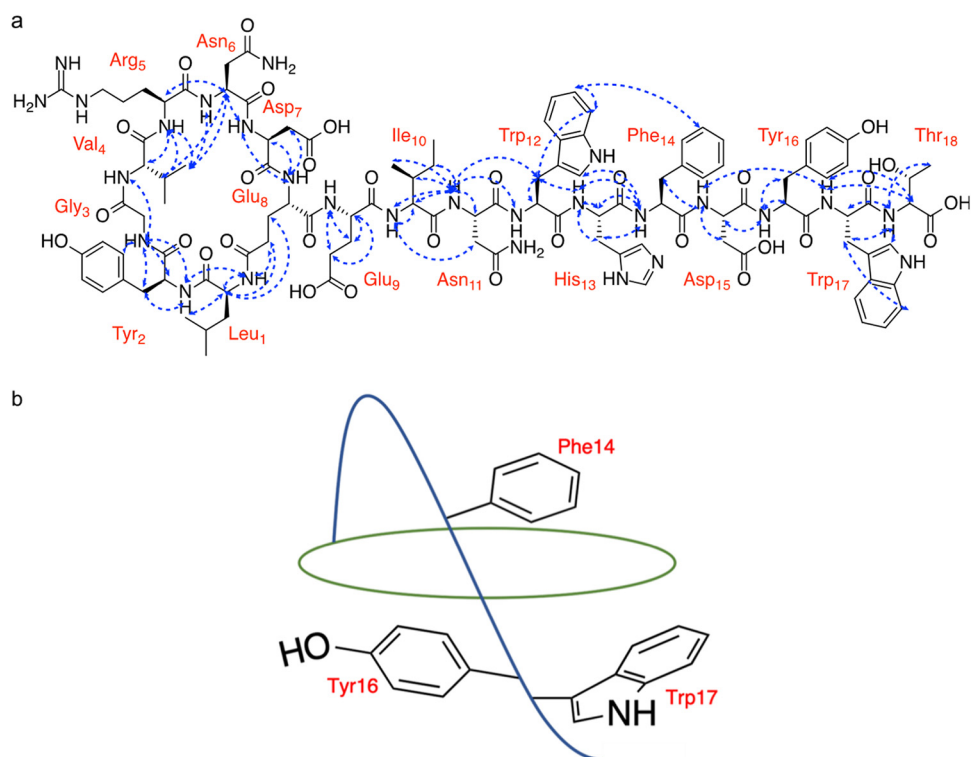


FIG 3 (a) NMR structure determination of leepeptin. Key NOESY correlations (blue) establishing sequence connectivity between amino acid residues. (b) Stylized representation of leepeptin. Tail residues deduced to be plugs, based on tail-to-ring NOE correlations, are displayed.

potentially be involved based on the rules of lasso peptide macrocycle formation: Asp₇, Glu₈, and Glu₉ (positions refer to the core peptide). The last clearly detected fragment ions were a⁹ (predicted, 1,030.4960; observed, 1,030.4958) and b⁹ (predicted, 1,058.4910; observed, 1,058.4921). We did not observe any ions that matched those expected for the a⁸, b⁸, and y¹¹ ions. Altogether, these results suggested that it is the side chain of Glu₈ that closes the macrocycle with the NH₂ of the N-terminal Leu (Fig. 2c; see also Fig. S7).

NMR structure elucidation of leepeptin. All of the ¹H and ¹³C nuclear magnetic resonance (NMR) data were in agreement with *in silico* and MS predictions. Along with one-dimensional (1D) experiments, 2D ¹H-¹³C heteronuclear single quantum correlation (HSQC), ¹H-¹³C heteronuclear multiple-bond correlation (HMBC), ¹H-¹H total correlated spectroscopy (TOCSY), ¹H-¹³C HSQC-TOCSY, and ¹H-¹H nuclear Overhauser effect spectroscopy (NOESY) allowed assignment of nearly all hydrogens (see Table S2 in the supplemental material) in the 18-residue peptide, including two Trp, two Tyr, two Asn, two Asp, two Glu, one Leu, one Gly, one Val, one Arg, one Ile, one His, one Phe and one Thr.

Sequence connectivity was established using NOESY correlations (Fig. 3a). The connection between Leu₁ and Tyr₂ was established using the correlations between Leu₁-NH and Tyr₂-NH. The connection between Tyr₂ and Gly₃ was established using the correlations among Tyr₂-NH, αH, βH (δ 2.35), and δH (δ 6.71) and Gly₃-NH. Furthermore, the Gly₃-Val₄ connection was determined using Gly₃-NH to Val₄-NH and Val₄-βH correlations as well as the Gly₃-αH (all) to Val₄-NH correlations. The Val₄-Arg₅ connection was determined using Val₄-NH, αH, βH, and γH (δ 0.95) correlations (all) to Arg₅-NH correlations. The Arg₅-Asn₆ connection was established using Arg₅-αH and Asn₆-αH correlations as well as Val₄-βH and γH (δ 0.95) connections to both Asn₆-NH and Asn₆-αH. Furthermore, the Asn₆-Asp₇ connection was determined using the Asn₆-αH and Asp₇-NH correlations. The Asp₇-Glu₈ connection was established using the Asp₇-NH, αH, and βH (δ 2.27) correlations (all) to Glu₈-NH and Asp₇-NH to Glu₈-βH (δ 1.75).

<i>Micromonospora carbonacea</i>	MKETYQAPTLPGLVFAEETG	AYGVRNQDEIVWFFDTWS
Leepeptin	MEHDEKTPYETPAVYGLGFAEETG	LYGVRNDEEINWHFDYWT
<i>Streptomyces</i> sp. L-9-10	MEKESYEAPVVMELGSFGEETG	VFGIRNGDEITWFFDTWQ
<i>Amycolatopsis xylanica</i>	MVKNGYEAPAVIELGAFDEETG	VFYVRNGEVLWFFDTWI
<i>Actinomadura fibrosa</i>	MDAFS YEAP EVIDLGDFTEATG	NTGIRNND E I Y M P F D E W H

FIG 4 Alignment of the amino acid sequences of the precursor peptide of leepeptin (GenBank accession number [CQR60041.1](#)) with homologues found in the GenBank database, including *Micromonospora carbonacea* ([SCF49478.1](#)), *Streptomyces* sp. L-9-10 ([RYJ28011.1](#)), *Amycolatopsis xylanica* ([SDZ36359.1](#)), and *Actinomadura fibrosa* ([WP_131760281.1](#)). Leader sequences are shown in blue and core peptide sequences in red.

The Glu₈-Leu₁ macrocyclization was determined through analysis of NOEs from Glu₈-βH (δ 1.54) and γH (δ 1.67) to Leu₁-αH as well as Glu₈-γH (δ 1.67) to Leu₁-NH.

There is no NMR evidence for the sequence connectivities of Glu₈-Glu₉ and Glu₉-Ile₁₀, but there is tandem MS (MS²) (Fig. 2c) evidence for these linkages. The Ile₁₀-Asn₁₁ connection was determined using the correlations between Ile₁₀-NH to Asn₁₁-βH (δ 2.66) and Ile₁₀-NH, αH, βH, γH's, and ε (δ 0.82) (all to Asn₁₁-NH). Furthermore, the Asn₁₁-Trp₁₂ connection was established using the correlations between Asn₁₁-NH to Trp₁₂-NH and βH (δ 2.95). The Trp₁₂-His₁₃ connection was established using the correlations from Trp₁₂-βH (δ 2.95) and His₁₃-NH. The His₁₃-Phe₁₄ connection was determined using the correlations between His₁₃-NH, αH, and βH (δ 1.05) to Phe₁₄-NH. Furthermore, the Phe₁₄-Asp₁₅ connection was determined using the correlations between Phe₁₄-βH's to Asp₁₅-NH. The Asp₁₅-Tyr₁₆ connection was determined using the correlation between Asp₁₅-αH and Tyr₁₆-NH. The Tyr₁₆-Trp₁₇ connection was difficult to assign using NOEs due to overlapping chemical shifts of multiple residues with those signals assigned in Trp₁₇. The Tyr₁₆-ε (δ 6.38) correlation to Trp₁₇-NH and the longer-range correlation between Asp₁₅-NH to Trp₁₇-NH help to place this residue. Lastly, the Trp₁₇-Thr₁₈ connection was established using the correlations between Trp-αH, βH's, and δH (δ 7.37) (all to Thr₁₈-NH) as well as from Trp₁₇-NH to Thr₁₈-αH.

NOESY data are generally used to confirm the residues in class II lasso peptides that prevent unthreading of the structure (see reference 15 for a discussion of the different classes of lasso peptides). Purely on the basis of the number of NOE correlations from tail residues back to macrocyclic residues, we were able to predict which amino acid residues in the tail of the peptide play a role in locking the tail in place. Working from the C-terminal Thr, Trp₁₇ and Tyr₁₆ each contain seven NOE correlations (Fig. S8). Both of these two residues seem to act as lower plugs, in a "harpoon"-like fashion, which is unusual among class II lasso peptides (Fig. 3b). Further up the tail, Phe₁₄ contains seven NOE correlations, suggesting that it acts as the upper plug (Fig. S8).

In an attempt to establish the absolute configuration of the amino acid residues of leepeptin, the methodology used in determining the absolute stereochemistry of chaxapeptin was followed (4). Marfey's analysis was used to determine the L conformation for all amino acids, with the exception of Leu₁, Trp₁₂ and Trp₁₇ (overlapping peaks).

Leepeptin represents a new subfamily of lasso peptides. While we could not detect any close homologues of leepeptin in the data set reported in reference 10, blastP searches (21) of the GenBank database revealed four homologues encoded by the genome sequences of *Streptomyces* sp. L-9-10, *Micromonospora carbonacea*, *Amycolatopsis xylanica*, and *Actinomadura fibrosa* (Fig. 4). Generation of a sequence similarity network (SSN) (22) using the precursor peptide sequences identified in Data Set 2 in reference 10 but also including these four newly identified sequences confirmed the emergence of this new subfamily of lasso peptides (Fig. S9). blastP searches of precursor or core peptide data sets derived from reference 10 but also including these four newly identified sequences and using the precursor and core sequences of each of the five peptides as query sequences confirmed the results of the SSN analyses; the only significant hits were other members of this new subfamily (data not shown).

DISCUSSION

Analysis of the genome sequence of *S. leeuwenhoekii* C34^T had revealed three BGCs with the potential to produce the lasso peptides Lp1, Lp2, and Lp3 (7). Each of these putative lasso peptides was also later identified at the sequence level in the bioinformatic analysis reported in reference 10 but had not been characterized further. Lp3 was shown in this work to be identical to the previously published chaxapeptin from the closely related strain *S. leeuwenhoekii* C58, also isolated from the Chaxa Lagoon in the Salar de Atacama. Production by *S. leeuwenhoekii* C34^T and, after heterologous expression, by *S. coelicolor* was detected. Unusually, no transporters were encoded by the Lp3 BGC, and the results of our heterologous expression experiments suggest that export of chaxapeptin is mediated by a nonspecific secretion system commonly present in *Streptomyces* species; we are not aware of any other experimental reports of a lack of a specific transporter for lasso peptide export. Genome sequence information is available for three of the four strains that possess genes encoding Lp3 homologues (see Fig. S3 in the supplemental material); while neither of the BGCs identified in *S. leeuwenhoekii* C58 and in *S. kanamyceticus* encodes identifiable transporters, that present in *S. cinnamoneus* encodes a two-component ABC transporter (data not shown). Why some members of this newly identified family of lasso peptides should apparently utilize a specific transporter while others do not is unknown.

Although we could not detect production of Lp1 by *S. leeuwenhoekii* C34^T or elicit its synthesis in *S. coelicolor*, the predicted product appears to be identical to a member of the recently identified citrulassin family of lasso peptides (10).

Lp2, which appeared to represent a novel lasso peptide which we subsequently named leepeptin, could not be detected in the culture supernatant of *S. leeuwenhoekii* C34^T but could be readily isolated after heterologous expression in *S. coelicolor*, enabling its structural elucidation. Lack of production in the native strain but successful expression of the Lp2 BGC in the heterologous host under the control of the constitutive *ermE** promoter presumably reflects lack of activation of the regulatory mechanism required for leepeptin biosynthesis in *S. leeuwenhoekii* C34^T under the growth conditions used. This may indicate that a specific environmental signal that is required to elicit leepeptin production in the native host was not provided under the experimental conditions used. Leepeptin appears to represent a new subfamily of lasso peptides. All three of the *S. leeuwenhoekii* C34^T lasso peptides contain the conserved leader sequence motif YxxPxLxxxGxxxxTx, with the sole exception of the replacement of Leu₆ with Val in the leader peptide of leepeptin, and all three leader peptides possess the “invariant” (with one known exception; see reference 10) Thr as the penultimate amino acid residue (Fig. 5).

Lasso peptides are currently an understudied family of natural products. Although the number of isolated compounds is relatively small (44 were identified in the supplementary information of reference 10), genome mining has revealed gene clusters for over 1,300 unique lasso peptides occurring across a wide range of bacterial genera (10). Over 35% of these BGCs are found in *Actinobacteria*, and many may not be expressed under laboratory conditions at the levels required for their detection. Consequently, we believe that the approach of heterologous expression described here could provide a useful means to unlock this untapped source of novel chemical diversity. Moreover, our discovery of a structurally novel lasso peptide, leepeptin (whose biological activity remains to be explored), in samples from the extreme environment of the Chaxa Lagoon in the high-altitude Atacama Desert provides further support for the hypothesis that the isolation of microorganisms from relatively little-explored environments such as this may result in a plethora of new natural products.

MATERIALS AND METHODS

Strains, culture conditions, and general methods. Bacterial strains used and generated during this study are listed in Table 1. *Escherichia coli* and *Streptomyces* strains were cultured, maintained, and manipulated genetically following the methods described in reference 23 and reference 24, respectively. Spore stocks of *S. leeuwenhoekii* were prepared following standard methods from cultures grown on mannitol soya flour (SFM) agar medium at 37°C for 5 to 7 days. The culture media used for lasso peptide

TABLE 2 Plasmids used and constructed during this study

Plasmid	Description	Reference or source
pBluescript II KS(+)	General cloning vector	34
pJJ10257	Expression vector for <i>Streptomyces</i> , with <i>ermE</i> *p, <i>hyg</i> , conjugative (<i>oriT</i> from RK2), integrative (<i>phiC31 attP</i>)	18
pJJ12815	pJJ10257 with the Lp3 BGC	This work
pJJ12819	pJJ10257 with the complete Lp2 BGC	This work

below are underlined, and nucleotides not present in the genome sequence of *S. leeuwenhoekii* are lowercase) and JP128_Pacl (ttaatCGTGTGACCGGTGTCAGG) and genomic DNA from *S. leeuwenhoekii* C34^T as the template. The PCR product was first cloned into SmaI-digested pBluescript II KS(+) and the sequence verified by Sanger sequencing with universal primers M13F-24mer (CGCCAGGGTTTCCCAGT CACGAC) and M13R-22mer (TCACACAGGAAACAGCTATGAC). The verified insertion from one clone was excised with NdeI and PacI and cloned in pJJ10257 digested with the same enzymes, generating pJJ12815. pJJ12815 was introduced into the heterologous *Streptomyces* expression hosts by conjugation selecting for hygromycin-resistant integrants (24).

Construction of pJJ12819. A 2.8-kb DNA fragment containing the biosynthetic genes for Lp2 (*leeCEBA*) was PCR amplified with primers JP138_NdeI (aacatATGGAACCTGCGTCCCG) and JP139_HindIII (aaaagcttAGACCTCATCCGCGCAATG) and genomic DNA from *S. leeuwenhoekii* C34^T as the template. Independently, a 3.8-kb DNA fragment containing the precursor peptide and transport genes for Lp2 (*leeADF*) was PCR amplified with primers JP143_AvrII (aactagGGCGGGGTGACCTGAGGTGGATG) and JP144_NdeI (aacatATGGAGCAGCAGGAGAGAGC) and genomic DNA from *S. leeuwenhoekii* C34^T as the template. In both cases, the PCR product was first cloned into SmaI-digested pBluescript II KS(+) and the sequence verified by Sanger sequencing with universal primers M13F-24mer (CGCCAGGGTTTCCC AGTCACGAC) and M13R-22mer (TCACACAGGAAACAGCTATGAC). The *leeADF* fragment was excised by digestion with BsiWI (internal to *leeA*) and NotI (provided by the vector polylinker) and ligated to the pBluescript II KS(+) derivative containing *leeCEBA* cut with the same enzymes, reconstructing the complete Lp2 *leeCEBADF* BGC (see Fig. S5 in the supplemental material). The insertion from one clone was excised with NdeI and AvrII and cloned in pJJ10257 digested with the same enzymes, generating pJJ12819. pJJ12819 was introduced into the heterologous *Streptomyces* expression hosts by conjugation selecting for hygromycin-resistant integrants (24).

Large-scale production and purification of leepeptin. *S. coelicolor* M1623 (*S. coelicolor* M1152 containing pJJ12819) was cultured in 25 ml of production medium (12.5 ml tryptone soya broth [TSB; Difco] plus 12.5 ml yeast extract-malt extract (YEME) broth (YEME composition per liter, 3 g yeast extract, 3 g malt extract, 5 g bacteriological peptone, 10 g glucose, and 170 g sucrose) in a 250-ml Erlenmeyer flask for 8 days. An aliquot of the culture was used to inoculate 150 ml TSB/YEME (50:50) in a 500-ml Erlenmeyer flask and grown for 7 days at 30°C and with shaking at 200 rpm. This culture was used to inoculate five 2-liter baffled Erlenmeyer flasks containing 300 ml TSB/YEME (50:50) and incubated for 7 days at 30°C and with shaking at 200 rpm. The purity of the culture was verified at each step by streaking out an aliquot of the culture on tryptic soy agar (Becton, Dickinson). The cultures were harvested at 4,000 rpm for 10 min and the supernatant fractions pooled and mixed with 6% (wt/vol) Diaion HP-20 resin beads (Mitsubishi) and left overnight. The resin beads were recovered by filtering the solution through glass wool and then soaked in methanol. The methanol fraction was filtered through filter paper in a vacuum system. The eluent was concentrated at 40°C at reduced pressure, and the resulting solution was fractionated through an ion exchange column (Strata SAX) (55- μ m pore size, 70 Å, 500 mg/3 ml) using 10% of the packing weight as the loading fraction. The column was equilibrated with 20 column volumes of water before the sample was loaded and washed with 10 column volumes (30 ml) of different mixtures of methanol-water solutions i to v as follows: (i) water; (ii) 75% water and 25% methanol; (iii) 50% water and 50% methanol; (iv) 25% water and 75% methanol; (v) 100% methanol; (vi) 100% methanol and 0.1% trifluoroacetic acid (TFA). An aliquot of each fraction was concentrated and analyzed by liquid chromatography-mass spectrometry (LC-MS), and the peptide was detected with highest intensity in the fraction containing 75% methanol. This fraction was further concentrated and subjected to semipreparative high-performance liquid chromatography (HPLC; Agilent 1200) through a C₁₈ column (Waters SunFire) (5- μ m particle size, 100-Å pore size, 250 by 10 mm) connected to a binary pump and a photodiode array detector set at 280 nm. Solvents used were Milli-Q water–0.1% (vol/vol) formic acid (solvent A) and methanol–0.1% (vol/vol) formic acid (solvent B). A gradient was set from 0% to 100% solvent B in 30 min and then 100% solvent B for 5 min, giving a total run time of 35 min, with a flow rate of 2 ml/min. The fraction containing the peptide eluted at around 22 to 23 min. The purified peptide was obtained in a volume of 3 mg, and the purity of the peptide was assessed by LC-MS (Fig. S10).

Mass spectrometric analyses. Values corresponding to the ions expected from peptide fragmentation by tandem mass spectrometry (MS/MS) were calculated with Fragment Ion Calculator software (<http://db.systemsbio.org:8080/proteomicsToolkit/FragIonServlet.html>). For the calculation of A and B ions, 18.0106 Da was deducted to account for the loss of a water molecule during cyclization. For MALDI-ToF analyses, the peptide samples were spotted onto a prespotted AnchorChip MALDI target plate (α -cyano-4-hydroxycinnamic acid matrix; Bruker Daltonics, Coventry, United Kingdom), and the spots were washed briefly with 10 mM ammonium phosphate–0.1% TFA according to the manufacturer's instructions. After drying, the samples were analyzed using MALDI-ToF and an Ultraflex TOF/TOF mass spectrometer (Bruker). The instrument was calibrated using the prespotted standards (ca. 200 laser

shots). Samples were analyzed using a method optimized for peptide analysis, and spectra were summed from ca. 30×15 laser shots. The data were processed in FlexAnalysis (Bruker). Several liquid chromatography-mass spectrometry instruments and methods were employed. LC-IT-ToF MS methods were performed as previously published (28); briefly, samples were analyzed on a Shimadzu NexeraX2 LC instrument fitted with a Prominence photo diode array detector and with an LC-IT-ToF mass spectrometer set for positive-ion-mode detection; samples were injected in a Kinetex XB-C₁₈ column (part no. 00B-4496-AN; Phenomenex, USA) (2.6- μ m particle size, 100-Å pore size, 50 by 2.10 mm) and eluted with a gradient of 0.1% formic acid–water–methanol (2% to 100% methanol) over 9.5 min, at a flow rate of 0.6 ml per min; data acquisition and analysis were performed with LCMSsolution version 3 (Shimadzu). High-resolution mass spectra were acquired by LC-HRMS on a Synapt G2 mass spectrometer equipped with an Acquity ultraperformance liquid chromatograph (UPLC; Waters, Wilmslow, United Kingdom). Aliquots of the samples were injected onto an Acquity UPLC ethylene bridged hybrid (BEH) C₁₈ column (Waters) (1.7- μ m particle size, 100-Å pore size, 1 by 100 mm) and eluted with a gradient of (solvent B) acetonitrile–0.1% formic acid mixed with (solvent A) water–0.1% formic acid with a flow rate of 0.4 ml/min at 45°C. The concentration of solvent B was kept at 3% for 1 min followed by a gradient up to 50% solvent B in 12 min. MS data were collected with the following parameters: resolution mode; positive-ion mode; scan time, 1 s; mass range, m/z 50 to 2,000 (calibrated with sodium formate); capillary voltage, 3.0 kV; cone voltage, 40 V; source temperature, 150°C; desolvation temperature, 500°C. Leu-enkephalin peptide was used to generate a lock-mass calibration with 556.2766 m/z measured every 10 s during the run. For MS/MS fragmentation, a data directed analysis (DDA) method was used with the following parameters: precursor selection from inclusion list only (807.5; 1,170.5), MS/MS threshold, 7,000; scan time, 5 s; no dynamic exclusion; collision energy (CE) ramped between 20 and 30 at low mass (m/z 50) and 30 and 60 at high mass (m/z 2,000). HRMS was also performed on an Agilent Technology 1290 Infinity II system fitted with a Kinetex C₁₈ column (part no. 00D-4462-AN; Phenomenex Inc.) (2.6- μ m particle size, 100-Å pore size, 100 by 2.1 mm), which was kept at 40°C. The system was coupled to a quadrupole time of flight (QTOF) maXis II mass spectrometer (Bruker Inc.). For liquid chromatography, a gradient of 95% Milli-Q water to 100% solvent B (acetonitrile–0.1% formic acid) was used over 10 min with a 5- μ l injection volume. Mass spectrometry was deployed in positive-ionization mode with a mass/charge range of 50 to 2,000. MS/MS fragmentation of purified leupeptin was obtained using direct injection into a LTQ XL Orbitrap mass spectrometer using 35% relative collision energy. Marfey's analysis was carried out using Marfey's FDLA [N α -(2,4-dinitro-5-fluorophenyl)-L-leucylamide] standard (Sigma-Aldrich) and the protocol described in reference 4.

Nuclear magnetic resonance. All 1D and 2D NMR experiments were performed using a Bruker Avance III HD 600-MHz spectrometer (Ascend 14.1 Tesla) with a Prodigy TCI cryoprobe at 25°C.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/AEM.01752-19>.

SUPPLEMENTAL FILE 1, PDF file, 1.8 MB.

ACKNOWLEDGMENTS

We are grateful to Michael Goodfellow and Alan Bull for providing *S. leeuwenhoekii* C34^T and to Michael Fischbach and Jan Claesen for *S. viridochromogenes* and *S. pristinaspiralis*, Matthias Mach for *S. davawensis*, and Kristian Apel for *S. roseochromogenes*. We thank Govind Chandra for advice on blastP analyses of the lasso peptide data sets, Solène Rollet for technical support in the isolation of leupeptin, and Andrew Truman for his comments on the manuscript.

J.F.C. and V.R. received national Ph.D. scholarships (21110356 and 21110384, respectively) and visiting student scholarships (Becas Chile, 2013 to 2014) from the National Commission for Scientific and Technological Research (CONICYT). S.A.J. thanks the University of Aberdeen for an Elphinstone scholarship. This work was supported financially by the Biotechnological and Biological Sciences Research Council (BBSRC; United Kingdom) Institute strategic program grant "Understanding and Exploiting Plant and Microbial Secondary Metabolism" (BB/J004561/1), by the Basal Program of CONICYT (Chile) funding of the Centre for Biotechnology and Bioengineering, by CeBiB (project FB0001), and by the United Kingdom Newton Project for United Kingdom-Chile collaboration (grant JIC CA586).

REFERENCES

- Okoro CK, Brown R, Jones AL, Andrews BA, Asenjo JA, Goodfellow M, Bull AT. 2009. Diversity of culturable actinomycetes in hyper-arid soils of the Atacama Desert, Chile. *Antonie Van Leeuwenhoek* 95:121–133. <https://doi.org/10.1007/s10482-008-9295-2>.
- Busarakam K, Bull AT, Girard G, Labeda DP, van Wezel GP, Goodfellow M. 2014. *Streptomyces leeuwenhoekii* sp. nov., the producer of chaxalactins and chaxamycins, forms a distinct branch in *Streptomyces* gene trees. *Antonie Van Leeuwenhoek* 105:849–861. <https://doi.org/10.1007/s10482-014-0139-y>.
- Nachtigall J, Kulik A, Helaly S, Bull AT, Goodfellow M, Asenjo JA, Maier A,

- Wiese J, Imhoff JF, Süßmuth RD, Fiedler H-P. 2011. Atacamycins A-C, 22-membered antitumor macrolactones produced by *Streptomyces* sp. C38. *J Antibiot* 64:775–780. <https://doi.org/10.1038/ja.2011.96>.
4. Elsayed SS, Trusch F, Deng H, Raab A, Prokes I, Busarakam K, Asenjo JA, Andrews BA, van West P, Bull AT, Goodfellow M, Yi Y, Ebel R, Jaspars M, Ratab ME. 2015. Chaxapeptin, a lasso peptide from extremotolerant *Streptomyces leeuwenhoekii* strain C58 from the hyperarid Atacama Desert. *J Org Chem* 80:10252–10260. <https://doi.org/10.1021/acs.joc.5b01878>.
 5. Ratab ME, Houssen WE, Arnold M, Abdelrahman MH, Deng H, Harrison WTA, Okoro CK, Asenjo JA, Andrews BA, Ferguson G, Bull AT, Goodfellow M, Ebel R, Jaspars M. 2011. Chaxamycins A-D, bioactive ansamycins from a hyper-arid desert *Streptomyces* sp. *J Nat Prod* 74:1491–1499. <https://doi.org/10.1021/np200320u>.
 6. Ratab ME, Houssen WE, Harrison WTA, Deng H, Okoro CK, Asenjo JA, Andrews BA, Bull AT, Goodfellow M, Ebel R, Jaspars M. 2011. Diverse metabolic profiles of a *Streptomyces* strain isolated from a hyper-arid environment. *J Nat Prod* 74:1965–1971. <https://doi.org/10.1021/np200470u>.
 7. Gomez-Escribano JP, Castro JF, Razmilic V, Chandra G, Andrews B, Asenjo JA, Bibb MJ. 2015. The *Streptomyces leeuwenhoekii* genome: *de novo* sequencing and assembly in single contigs of the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2. *BMC Genomics* 16:485. <https://doi.org/10.1186/s12864-015-1652-8>.
 8. Castro JF, Razmilic V, Gomez-Escribano JP, Andrews B, Asenjo J, Bibb MJ. 2018. The 'gifted' actinomycete *Streptomyces leeuwenhoekii*. *Antonie Van Leeuwenhoek* 111:1433–1448. <https://doi.org/10.1007/s10482-018-1034-8>.
 9. Um S, Kim Y-J, Kwon H, Wen H, Kim S-H, Kwon HC, Park S, Shin J, Oh D-C. 2013. Sungsanpin, a lasso peptide from a deep-sea streptomycete. *J Nat Prod* 76:873–879. <https://doi.org/10.1021/np300902g>.
 10. Tietz JI, Schwalen CJ, Patel PS, Maxson T, Blair PM, Tai HC, Zakai UI, Mitchell DA. 2017. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat Chem Biol* 13:470–478. <https://doi.org/10.1038/nchembio.2319>.
 11. Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J, Cotter PD, Craik DJ, Dawson M, Dittmann E, Donadio S, Dorrestein PC, Entian KD, Fischbach MA, Garavelli JS, Göransson U, Gruber CW, Haft DH, Hemscheidt TK, Hertweck C, Hill C, Horswill AR, Jaspars M, Kelly WL, Klinman JP, Kuipers OP, Link AJ, Liu W, Marahiel MA, Mitchell DA, Moll GN, Moore BS, Müller R, Nair SK, Nes IF, Norris GE, Olivera BM, Onaka H, Patchett ML, Piel J, Reaney MJ, Rebuffat S, Ross RP, Sahl HG, Schmidt EW, Selsted ME, et al. 2013. Ribosomally synthesized and post-translationally modified peptide natural products. *Nat Prod Rep* 30:108–160. <https://doi.org/10.1039/C2NP20085F>.
 12. Maksimov MO, Pan SJ, Link AJ. 2012. Lasso peptides: structure, function, biosynthesis, and engineering. *Nat Prod Rep* 29:996–1006. <https://doi.org/10.1039/c2np20070h>.
 13. Tan S, Moore G, Nodwell J. 2019. Put a bow on it: knotted antibiotics take center stage. *Antibiotics* 8:117. <https://doi.org/10.3390/antibiotics8030117>.
 14. Hegemann JD, De Simone M, Zimmermann M, Knappe TA, Xie X, Di Leva FS, Marinelli L, Novellino E, Zahler S, Kessler H, Marahiel MA. 2014. Rational improvement of the affinity and selectivity of integrin binding of grafted lasso peptides. *J Med Chem* 57:5829–5834. <https://doi.org/10.1021/jm5004478>.
 15. Hegemann JD, Zimmermann M, Xie X, Marahiel MA. 2015. Lasso peptides: an intriguing class of bacterial natural products. *Acc Chem Res* 48:1909–1919. <https://doi.org/10.1021/acs.accounts.5b00156>.
 16. Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, Weber T. 2013. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res* 41:W204–W212. <https://doi.org/10.1093/nar/gkt449>.
 17. Inokoshi J, Matsuhama M, Miyake M, Ikeda H, Tomoda H. 2012. Molecular cloning of the gene cluster for lariat biosynthesis of *Rhodococcus jostii* K01-B0171. *Appl Microbiol Biotechnol* 95:451–460. <https://doi.org/10.1007/s00253-012-3973-8>.
 18. Hong H-J, Hutchings MI, Hill LM, Buttner MJ. 2005. The role of the novel Fem protein VanK in vancomycin resistance in *Streptomyces coelicolor*. *J Biol Chem* 280:13055–13061. <https://doi.org/10.1074/jbc.M413801200>.
 19. Bibb MJ, White J, Ward JM, Janssen GR. 1994. The mRNA for the 23S rRNA methylase encoded by the *ermE* gene of *Saccharopolyspora erythraea* is translated in the absence of a conventional ribosome-binding site. *Mol Microbiol* 14:533–545. <https://doi.org/10.1111/j.1365-2958.1994.tb02187.x>.
 20. Gomez-Escribano JP, Bibb MJ. 2011. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol* 4:207–215. <https://doi.org/10.1111/j.1751-7915.2010.00219.x>.
 21. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
 22. Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, Whalen KL. 2015. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): a Web tool for generating protein sequence similarity networks. *Biochim Biophys Acta* 1854:1019–1037. <https://doi.org/10.1016/j.bbapap.2015.04.015>.
 23. Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. [https://doi.org/10.1016/0167-7799\(91\)90068-5](https://doi.org/10.1016/0167-7799(91)90068-5).
 24. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA. 2000. *Practical Streptomyces genetics*. John Innes Foundation, Norwich, United Kingdom.
 25. Shima J, Hesketh A, Okamoto S, Kawamoto S, Ochi K. 1996. Induction of actinorhodin production by *rpsL* (encoding ribosomal protein S12) mutations that confer streptomycin resistance in *Streptomyces lividans* and *Streptomyces coelicolor* A3(2). *J Bacteriol* 178:7276–7284. <https://doi.org/10.1128/jb.178.24.7276-7284.1996>.
 26. Gavriš E, Sit CS, Cao S, Kandró 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. 2014. Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. *Chem Biol* 21:509–518. <https://doi.org/10.1016/j.chembiol.2014.01.014>.
 27. Staden R, Beal KF, Bonfield JK. 1999. The Staden package, 1998, p 115–130. *In* Misener S, Krawetz SA (ed), *Bioinformatics methods and protocols*. Humana Press, Totowa, New Jersey. <https://doi.org/10.1385/1592591922>.
 28. Castro JF, Razmilic V, Gomez-Escribano JP, Andrews B, Asenjo J, Bibb MJ. 2015. Identification and heterologous expression of the chaxamycin biosynthetic gene cluster from *Streptomyces leeuwenhoekii*. *Appl Environ Microbiol* 81:5820–5831. <https://doi.org/10.1128/AEM.01039-15>.
 29. Grant SG, Jessee J, Bloom FR, Hanahan D. 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc Natl Acad Sci U S A* 87:4645–4649. <https://doi.org/10.1073/pnas.87.12.4645>.
 30. MacNeil DJ, Gewain KM, Ruby CL, Dezeny G, Gibbons PH, MacNeil T. 1992. Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* 111:61–68. [https://doi.org/10.1016/0378-1119\(92\)90603-M](https://doi.org/10.1016/0378-1119(92)90603-M).
 31. Chater KF, Carter AT. 1978. Restriction of a bacteriophage in *Streptomyces albus* P (CMI72566) by endonuclease SalPI. *J Gen Microbiol* 109:181–185. <https://doi.org/10.1099/00221287-109-1-181>.
 32. Jankowitsch F, Schwarz J, Ruckert C, Gust B, Szczepanowski R, Blom J, Pelzer S, Kalinowski J, Mack M. 2012. Genome sequence of the bacterium *Streptomyces davawensis* JCM 4913 and heterologous production of the unique antibiotic roseoflavin. *J Bacteriol* 194:6818–6827. <https://doi.org/10.1128/JB.01592-12>.
 33. Rückert C, Kalinowski J, Heide L, Apel AK. 2014. Draft genome sequence of *Streptomyces roseochromogenes* subsp. *oscitans* DS 12.976, producer of the aminocoumarin antibiotic clorobiocin. *Genome Announc* 2:e01147-13. <https://doi.org/10.1128/genomeA.01147-13>.
 34. Altig-Mees MA, Sorge JA, Short JM. 1992. pBluescriptII: multifunctional cloning and mapping vectors. *Methods Enzymol* 216:483–495. <https://doi.org/10.1093/nar/17.22.9494>.