doi.org/10.26434/chemrxiv.14544213.v1

Discovery of the Class I Antimicrobial Lasso Peptide Arcumycin

Lydia Stariha, Dewey G. McCafferty

Submitted date: 05/05/2021 • Posted date: 10/05/2021

Citation information: Stariha, Lydia; McCafferty, Dewey G. (2021): Discovery of the Class I Antimicrobial Lasso Peptide Arcumycin. ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.14544213.v1

Lasso peptides are a structurally diverse superfamily ofconformationally-constrained peptide natural products, of which asubset exhibits broad antimicrobial activity. Although advances inbioinformatics have increased our knowledge of strains harboringthe biosynthetic machinery for lasso peptide production, relatingpeptide sequence to bioactivity remains a continuous challenge.Towards this end, a structure-driven genome mining investigationof Actinobacteria-produced antimicrobial lasso peptides wasperformed to correlate predicted primary structure with antibioticactivity. Bioinformatic evaluation revealed eight putative novelclass I lasso peptide sequences. This subset is predicted topossess antibiotic activity as characterized members of this classhave both broad spectrum and potent activity against Gram positivestrains. Fermentation of one of these hits, StreptomycesNRRL F-5639, resulted in the production of a novel class I lassopeptide, arcumycin, named for the Latin word for bow or arch,arcum. Arcumycin exhibited antibiotic activity against Gram positivebacteria including Bacillus subtilis (4 μg/mL),Staphylococcus aureus (8 μg/mL), and Micrococcus luteus (8μg/mL). Arcumycin treatment of B. subtilis liaI-β-gal promoterfusion reporter strain resulted in upregulation of the system liaRSby the promoter liaI, indicating arcumycin interferes with lipid IIbiosynthesis. Cumulatively, the results illustrate the relationshipbetween phylogenetically related lasso peptides and theirbioactivity as validated through the isolation, structuraldetermination, and evaluation of bioactivity of the novel class Iantimicrobial lasso peptide arcumycin.

File list (2)

Discovery of the Class I Antimicrobial Lasso Peptide Arcumycin

Lydia M. Stariha and Dewey G. McCafferty*

Department of Chemistry
Duke University
Durham, NC, 27708-0346 (USA)
E-mail: dewey.mccafferty@duke.edu

Supporting information for this article is given via a link at the end of the document.

Abstract: Lasso peptides are a structurally diverse superfamily of conformationally-constrained peptide natural products, of which a subset exhibits broad antimicrobial activity. Although advances in bioinformatics have increased our knowledge of strains harboring the biosynthetic machinery for lasso peptide production, relating peptide sequence to bioactivity remains a continuous challenge. Towards this end, a structure-driven genome mining investigation of Actinobacteria-produced antimicrobial lasso peptides was performed to correlate predicted primary structure with antibiotic activity. Bioinformatic evaluation revealed eight putative novel class I lasso peptide sequences. This subset is predicted to possess antibiotic activity as characterized members of this class have both broad spectrum and potent activity against Gram-positive strains. Fermentation of one of these hits, *Streptomyces* NRRL F-5639, resulted in the production of a novel class I lasso peptide, arcumycin, named for the Latin word for bow or arch, arcum. Arcumycin exhibited antibiotic activity against Gram-positive bacteria including *Bacillus subtilis* (4 μg/mL), *Staphylococcus aureus* (8 μg/mL), and *Micrococcus luteus* (8 μg/mL). Arcumycin treatment of *B. subtilis liaI-β-gal* promoter fusion reporter strain resulted in upregulation of the system *liaRS* by the promoter *liaI*, indicating arcumycin interferes with lipid II biosynthesis. Cumulatively, the results illustrate the relationship between phylogenetically related lasso peptides and their bioactivity as validated through the isolation, structural determination, and evaluation of bioactivity of the novel class I antimicrobial lasso peptide arcumycin.

Introduction

The discovery of antibiotics revolutionized the treatment of infectious disease. However, over the past thirty years, the rate of introduction of new-in-class antibiotics has flattened, yet the simultaneous rate of clinical cases of infections due to bacteria that are resistant to front-line antibiotics has steadily increased. Given our dwindling arsenal of effective antibiotics, it is not difficult to see a time when our most serious infectious threats may be untreatable.[1] In the United States alone, there are 2.8 million reported cases and 35,000 deaths annually due to resistant bacterial infections.[2] Natural products have proven to be critical sources for antibiotics, as over 70% of clinically approved treatments are natural products or natural product derivatives.[3] Bioinformatic analyses have estimated there are thousands of predicted compounds with novel structures and biological activities that are yet to be discovered and functionally validated.[4]

Lasso peptides are a superfamily of ribosomally synthesized and post-translationally modified peptide natural products (RiPPs).[5] First discovered in the 1990s, lasso peptides are produced by numerous bacterial species, including members of the Proteobacteria, Actinobacteria, and Firmicutes phyla.[6],[7] These peptides have a wide range of bioactivities including antiviral,[8] anticancer,[9] cell-surface receptor agonist,[10] and antimicrobial.[11–15] To date, there are over 80 characterized lasso peptides. Each shares a general structure comprised of 16–25 amino acids that are characterized by an intramolecular macrocycle through which is threaded a C-terminal tail region to create a knot structure. Adoption of its unusual lariat-like topology is essential for bioactivity,[16][17] renders the lasso peptide thermostable,[18] and protects the peptide backbone against proteolytic degradation.[19]

Within the superfamily, a subset of peptides also contains intramolecular disulfide bonds that increase stability. Four structural classes of lasso peptides have been classified based in part on the number and location of the disulfide bonds (**Figure 1**).[20] Class I members are characterized by two disulfide bonds between residues Cys_1-Cys_{13} and Cys_7-Cys_{19}, (e.g. siamycin I) (Figure 1). Class II members lack disulfides (e.g. propeptin), while class III and IV peptides contain a single disulfide with class III between residues Cys_7-Cys_{20} (e.g. BI-32169), and class IV between Cys_{13}-Cys_{17} (e.g. LP-2006). Within each class, members exhibit some diversity of amino acid sequence,[21],[22] peptide length,[23],[24] macrocycle size,[25] post-translational modification,[26],[27] and residues involved in macrocyclization. [28]

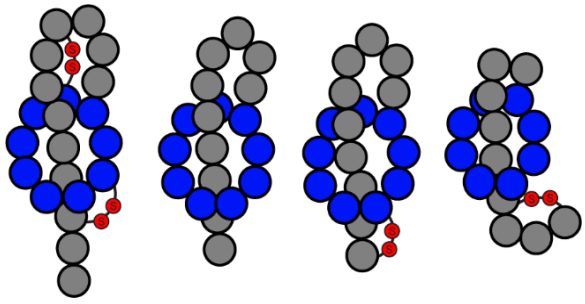

Figure 1. Representative structures of the four lasso peptide classes. Blue residues denote amino acids within macrocycle regions, grey residues belong to tail regions, and red indicates disulfide bonds.

Advances in genome mining have accelerated the study of lasso peptides by vastly expanding our knowledge of strains containing the biosynthetic machinery to produce this class of natural products.[29–31] Sequence comparison algorithms such as RODEO, developed by Mitchell and coworkers,[31] and decRiPPter, developed by Medema and colleagues,[32] have significantly advanced the discovery of RiPP biosynthetic gene clusters (BGCs) within bacterial genomes. Genome mining studies of lasso peptides have been used to explore novel producers like the fungal-associated bacterium *Rhizopua microspores*[33] and to discover unique **post-translational modification**s such as acylation.**[34]** While genome mining has expanded the known chemical space of lasso peptides, relating structure to bioactivity has remained a challenge with fewer than 30 of the 80 characterized lasso peptides having reported activity.[35]

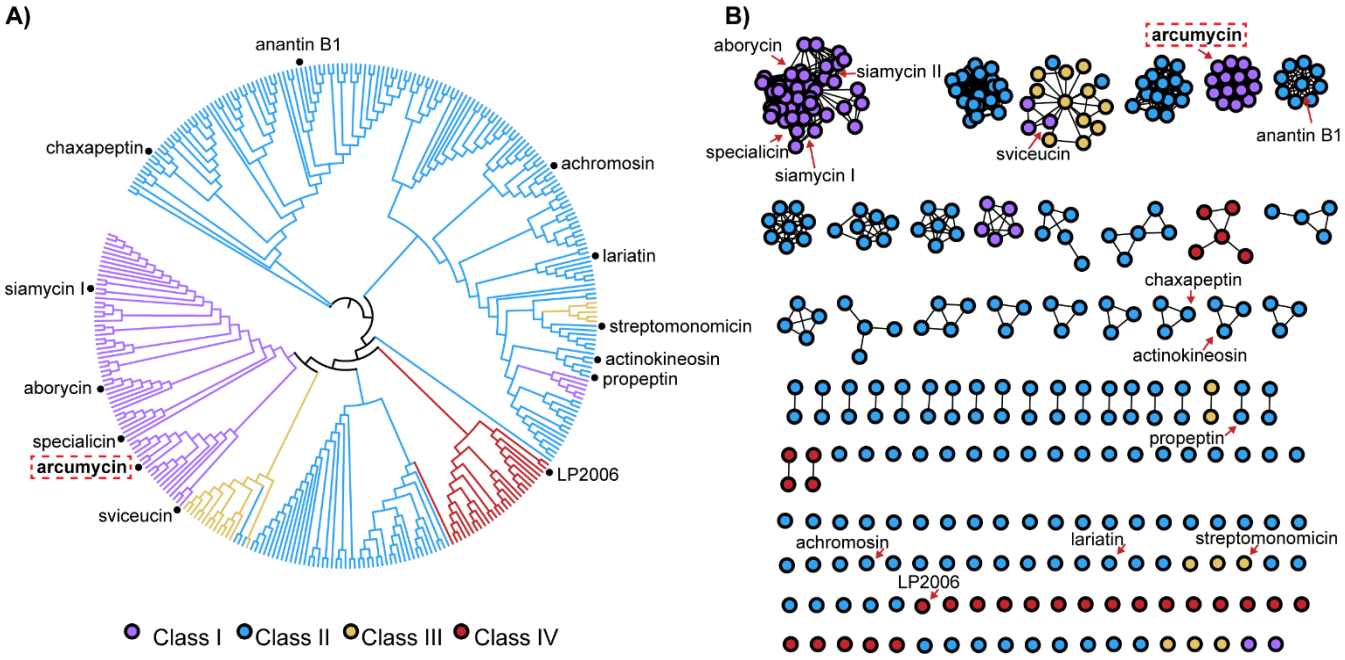

Figure 2. Phylogenetic analyses of predicted antimicrobial lasso peptides from Actinobacteria. (**A**) The multi-locus maximum-likelihood phylogenetic tree was generated from 331 BGCs from unique strains using CORASON. (**B**) Sequence similarity network of the 331 lasso peptide BGCs generated by BiG-SCAPE.

Results and Discussion

We hypothesized that lasso peptides identified in strains that are phylogenetically related to known antimicrobial lasso peptides producing strains may result in the discovery of novel lasso peptides with conserved structural features resulting in similiar bioactivities. As such, we designed a focused, structure-driven genome mining study to identify lasso peptide BGCs that are related to lasso peptides with well-established antimicrobial activity. We analyzed known antimicrobial lasso peptides produced by Actinobacteria to expand the chemical space surrounding these natural products. There are currently fifteen lasso peptides produced by Actinobacteria that have reported antimicrobial activity, and twelve have completely annotated BGCs available on the NCBI database (**Table S1**).[7],[26],[31],[36–44] By studying biosynthetic diversity within this one phylum, the genetic diversity of the input data is lowered, allowing for the identification of more discrete variations within BGC sequences and phylogenetic relationships.

Lasso peptide biosynthesis requires four proteins for production and maturation of the natural product: 1) the precursor peptide, comprised of the leader peptide that is recognized by tailoring proteins and the core peptide that contains the primary amino acid sequence of the mature natural product; 2) a lasso cyclase, which forms the isopeptide macrocyclic bond within the core peptide; 3) the RiPP recognition element, which binds to the leader peptide; and 4) the leader peptidase, which cleaves the leader peptide from the core peptide.[45],[46] To explore structural diversity directly related to lasso peptide sequence, we elected to use the precursor peptide from each of the twelve characterized antimicrobial lasso peptides as queries in alignment searches. We conducted BLASTp[47] searches against the NCBI database for each of the precursor peptides, and all resulting sequences query were collected as preliminary hits. Our hits were dereplicated across the different search queries and the surrounding biosynthetic machinery for lasso peptide biosynthesis was detected using antiSMASH 5.0.[48]

A total of 331 unique strains harboring the machinery for lasso peptide biosynthesis were identified, with several containing two different precursor peptides within the same cluster, providing a total of 345 predicted core peptide sequences (**Dataset 1**). Most of these BGCs were found among *Streptomyces,* but 39 genera and over 250 unique species were also predicted to harbor BGCs for lasso peptide biosynthesis. Interestingly, we detected four previously characterized lasso peptides that have either not been tested for antimicrobial activity (Res-701-3, ulleungdin, humidimycin) or were tested for antimicrobial activity against only a small number of strains (cattlecin), potentially indicating undiscovered antibiotic bioactivities.[9],[49–51] Among the 345 predicted lasso peptide sequences identified, we observed representation of each structural class as determined by the location and number of predicted disulfide bonds (70 class I, 217 class II, 27 class III, 31 class IV). The hits from our genome mining efforts were further analyzed through the use of bioinformatic tools allowing for the evaluation of sequence relationships based on evolutionary proximity.

To complement the traditional division of lasso peptides into classes based on disulfide bonds, we sought to use evolutionary proximity to evaluate the relationships of antimicrobial lasso peptides. To determine the phylogenetic relationships and sequence similarity for all lasso peptide BGCs within each unique strain, CORASON and BiG-SCAPE were employed, respectively.[52] While the data output for the phylogenetic tree and sequence similarity network (SSN) contains many similarities, using both of these tools facilitated the tracing of evolutionary proximity and relationships between BGC sequences. As CORASON and BiG-SCAPE perform multi-locus analyses for an entire BGC instead of a singular protein, these bioinformatic methods are particularly advantageous for lasso peptides that generally have smaller BGCs and lower molecular weight biosynthetic proteins. The lasso cyclase was used as the query for CORASON as it retains high homology across all input strains and is at the center of the BGC, allowing for the complete comparison of all strains and full clusters. Using this analysis, a multi-locus unrooted phylogenetic maximum-likelihood tree was generated tracing the evolutionary relationships for our 331 lasso peptide BGCs (**Figure 2A**). Lasso peptide classes form separate clades within the phylogenetic tree while previously characterized antimicrobial peptides are distributed throughout the clades and share lineages with other putative peptides within their class. Consistent with this result, the BGC SSN generated by BiG-

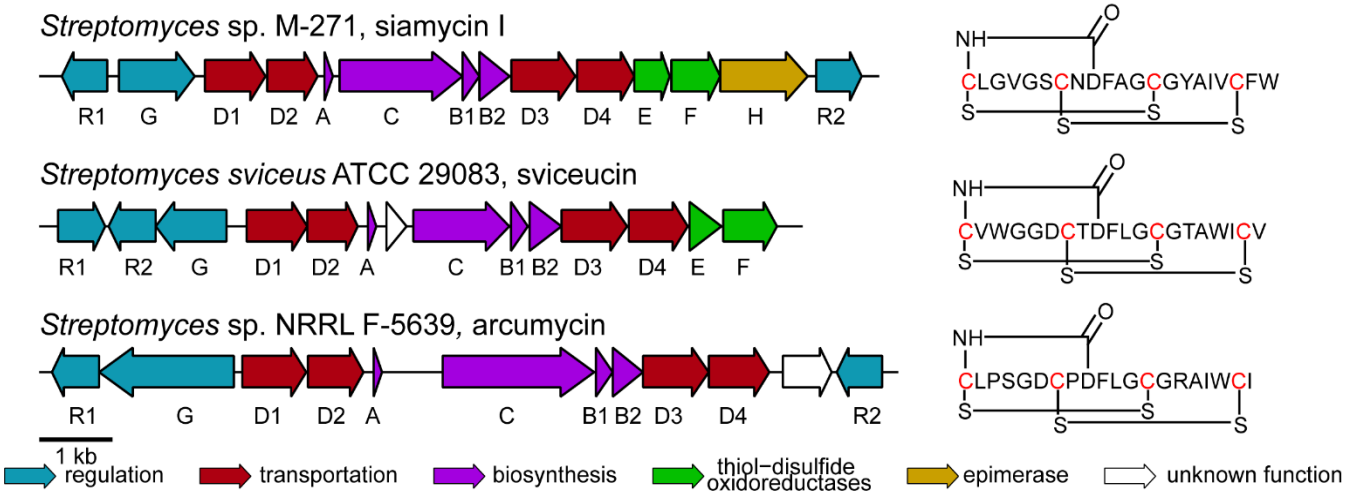

Figure 3. **Biosynthetic and structural comparison of class one lasso peptides.** Open reading frame arrow diagrams and amino acid sequences of class I lasso peptide natural products including previously characterized structures and putative novel member arcumycin.

SCAPE indicates predominant clustering of nodes based on structural class (**Figure 2B**).

While the structural class II lasso peptides, which are devoid of disulfide linkages, represent the majority of identified sequences at 61%, the disulfide-containing lasso peptides (classes I, III, and IV) are more highly represented in our results compared to previous genome mining studies.[31] This result illustrates the vast chemical space for disulfide bond-containing lasso peptides that is yet to be explored. This is significant as lasso peptides with disulfide bonds have been demonstrated to display interesting bioactivities. Generally, class I lasso peptides have shown broad-spectrum and potent antimicrobial activity. As an example, siamycin I, the most widely studied of the disulfide bond-containing lasso peptides, has been shown to bind cell wall biosynthesis intermediate lipid II,[53] inhibit the futalosine pathway,[54] and attenuate quorum sensing in *Enterococcus faecalis*.[55] Within the siamycin I clade, we identified several novel putative class I lasso peptides, providing naturally-produced congeners of this antimicrobial lasso peptide to explore for antimicrobial activity.

As several of the putative class I lasso peptide BGCs presented novel arrangements, we sought to interrogate which enzymatic features were conserved and which were not broadly required for lasso peptide production. To compare the BGCs encoding class I lasso peptides, we first removed BGCs containing redundant precursor peptides to specifically investigate variations in natural product primary sequence, resulting in 22 unique BGCs (**SI Table 2**). We then determined the composition of each BGC, which revealed four distinct archetypes based on putative enzymes present in the BGC: siamycin, sviceucin, novel BGCI, and novel BGCII (**SI Figure 1**). In comparing the sequences of the proteins encoded in these BGCs to that of siamycin I, it was illustrated that the core biosynthesis genes have a high level of homology, especially for the precursor peptide (**SI Table 3**). Additionally, the flanking four ABC transporter proteins retained high levels of sequence similarity, indicating a conserved mechanism for self-resistance.[56] All BGCs within the siamycin family encoded the hypothetical epimerase for the conversion of the terminal tryptophan residue to the D-Trp conformation, but this putative gene was not present in the other archetypes. BGCs belonging to both the siamycin and sviceucin-type BGCs contain two thiol-disulfide oxidoreductases, a protein found only in BGCs for class I lasso peptides. These proteins have been previously investigated for their role in the biosynthesis of class I lasso peptides, with in-frame deletions of these genes resulting in lower production of the final natural product, but not full elimination of production or loss of disulfide bond formation. [38],[57] Interestingly, the archetypes of novel BGCI and novel BGCII from our analysis do not contain any putative thiol-disulfide oxidoreductases, indicating these genes may not be essential for full maturation. To determine if alterations in BGC composition affect the production of class I lasso peptides or have any relationship to the bioactivity of the natural products, we selected to isolate a novel lasso peptide from BGCII.

To validate the predictions made through our analysis of class I lasso peptides produced by Actinobacteria, we elected to isolate, verify the structure, and determine the bioactivity of a new class I lasso peptide. The bacterial strain *Streptomyces* NRRL F-5639, identified as harboring a BGC encoding a unique lasso peptide, was obtained from the Agricultural Research Service Culture Collection. This strain was selected as the BGC belongs to the previously uncharacterized archetype BGCII, containing the four core genes for lasso peptide biosynthesis, four ABC transporter proteins, and three genes for transcriptional regulation. Putative epimerases and thiol-disulfide oxidoreductases are absent from this cluster (**Figure 3**). Using antiSMASH with the RODEO extension,[31] this novel class I lasso peptide—that we have named arcumycin after the Latin word for bow or arch—was predicted to be 20 amino acids in length with a macrocycle of nine amino acids cyclized between Cys_1-Asp_9 and two disulfide bridges between Cys_1-Cys_{13} and Cys_7-Cys_{19} (**Figure 3**). The peptide sequence of arcumycin was determined to have an overall amino acid identity to both the siamycin family and sviceucin of 55%, classifying this peptide as a novel subfamily of the class I lasso peptides while still having the potential to retain key structural elements for bioactivity.[38],[58]

To determine if the BGC for arcumycin could be activated, we first screened for peptide production from *Streptomyces* NRRL F-5639 using media conditions that have been used to produce other class I lasso peptides.[40],[38],[54],[55] None of the conditions resulted in the production of the novel lasso peptide as detected by MALDI-MS. We next screened numerous liquid media formulations optimized for *Actinomycete* metabolome profiling,[59] and arcumycin was found to be produced in seven of these conditions with the predicted mass detected by MALDI-MS and ESI-MS (**SI Figure 2**). We selected RAM media in our scale-up conditions as it showed the highest concentration of peptide (**SI Figure 3**). Guided by lasso peptide structural prediction tools, we experimentally validated the structure of arcumycin. We observed a $[M+2H]^{2+}$ of 1052.4519 *m/z* by high-resolution LC-MS, confirming the predicted chemical formula of $C_{92}H_{134}N_{24}O_{25}S_4$ (calculated $[M+2H]^{2+}$ 1052.4490). We next sought to determine connectivity by MS/MS. Typically, only the tail regions of lasso peptides have been elucidated with MS/MS analysis due to the stability of the macrolactam, which leads to poor levels of fragmentation within the macrocyclic region. However, recent advancements using higher-energy C-trap dissociation (HCD) over collisional-induced dissociation (CID) have resulted in complete connectivity validation for some peptides.[60] We performed an HCD-MS/MS analysis on a reduced

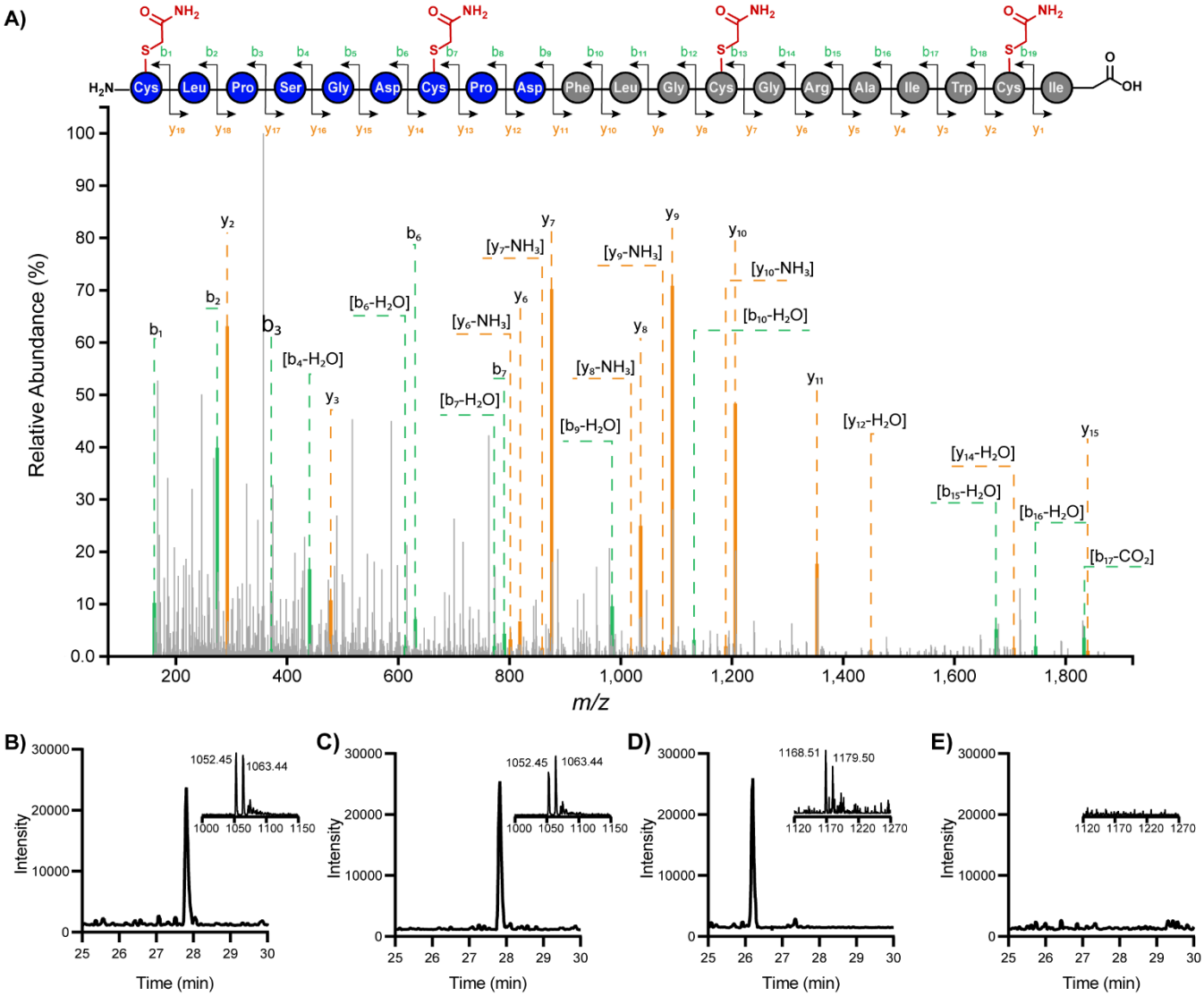

Figure 4. Mass spectral verification of arcumycin sequence and topology. (**A**) MS/MS spectrum of reduced and alkylated arcumycin (macrocycle blue, tail grey, alkylated sulfurs red) with diagnostic fragmentation pattern b- (green) and y-ions (yellow). Extraction ion chromatograms and mass spectra of **(B)** arcumycin ($[M+2H]^{2+}$ =1052.4490), **(C)** arcumycin exposed to carboxypeptidase Y ($[M+2H]^{2+}$ =1052.4490), **(D)** reduced and alkylated arcumycin ($[M+2H]^{2+}$ =1168.5077), **(E)** reduced and alkylated arcumycin after treatment with carboxypeptidase Y ($[M+2H]^{2+}$ =1168.5077).

and alkylated arcumycin and the data was analyzed through the Interactive Peptide Spectral Annotator.[61] The HCD-MS/MS analysis generated the corresponding b- and y-ions to confirm the lasso peptide amino acid sequence for the linear peptide that would arise from a ring opening at the site of macrocyclization between Cys_1-Asp_9 (**Figure 4A, SI Table 4**). We further examined other possible ring opening sites for the macrocycle and found prominent fragmentation for cycle opening between Cys_1-Asp_9 and Leu_2-Pro_3 for the corresponding nine amino acid linear peptide of the macrocycle (**SI Figures 4 and 5, SI Tables 5 and 6**). Collectively, the HCD-MS/MS data allowed for the validation of the amino acid sequence and connectivity of arcumycin. Additionally, we aimed to confirm the topology of the lasso peptide fold. Previous studies of lasso peptides have shown that fully maturated lasso peptides have resistance to proteolytic degradation, yet when unfolded to a branched linear form the natural product is vulnerable to proteases.[38],[62] Analysis by LC-MS showed that arcumycin was resistant to degradation by protease carboxypeptidase Y in its native form; however, after disrupting this secondary structure through chemical reduction and alkylation of the disulfide bonds the peptide was degraded, which suggests that it adopts a lariat fold (**Figure 4B, SI Figures 6-13**).

Using broth microdilution assays, we evaluated the ability of arcumycin to inhibit the growth of bacteria. As arcumycin shares high levels of sequence similarity to both the siamycins and sviceucin, establishing its killing profile is an interesting way to increase the knowledge of how the structures of class I lasso peptides contribute to biological activity. We tested arcumycin against the four Gram-positive strains *Bacillus subtilis* ATCC 6051, *Staphylococcus aureus* ATCC 25923, *Enterococcus faecalis* ATCC 29212, and *Micrococcus luteus* ATCC 4698, and one representative Gram-negative strain *Escherichia coli* ATCC 25922. Our results show that arcumycin was not active against the Gram-negative bacterium, but induced inhibition against multiple Gram-positive bacteria at a concentration of 4–8 μg/mL similar to other class I lasso peptides (**Figure 5A**). This indicates that although arcumycin is not identical in sequence to other class I lasso peptides, it does retain important structural features for bioactivity. Interestingly, the class I lasso peptides exhibit high variability in the strains each peptide is active against. Arcumycin showed no activity against *E. faecalis*, but both siamycin I and sviceucin are active against this strain. The evaluation of bioactivity shows that, through SAR-driven genome mining, we are able to use the phylogenetic relationships of antimicrobial lasso peptides to discover a new member of class I with antimicrobial bioactivity.

To further pursue the hypothesis that phylogenetically related lasso peptides will have similar bioactivities, we performed investigations into the mechanism of action of arcumycin. Previous studies have shown that seven Proteobacteria-produced lasso peptides exhibit a similar mechanism of action that induces bactericidal activity via inhibition of RNA polymerase.[14] While a similar conserved mechanism among Actinobacteria-produced lasso peptides has not been validated, initial evidence has shown that both siamycin I[53] and streptomonomicin[63] have similar resistance mutants, both within genes associated with cell wall biosynthesis. Further, siamycin I has been fully validated as a lipid II binder.[53] To evaluate the ability of arcumycin to act as an inhibitor of cell wall biosynthesis, we employed a reporter assay using a *B. subtilis* strain harboring a *liaI-β-gal* promoter fusion, which relies on the upregulation of the two component system *liaRS* by the promoter *liaI* when cell wall envelope is under stress.[64] When the fused *liaI-β-gal* promoter is induced β-galactosidase is expressed and able to cleave the X-gal dye resulting in the formation of a blue-colored product. Treatment of the *B. subtilis* reporter strain with arcumycin produced a blue ring around the zone of inhibition, similiar to the known lipid II binder vancomycin (**Figure 5B**). While further analysis of the association of lipid II with arcumycin will help elucidate the molecular details of this interaction, these initial results suggest a shared mechanism of action may exist among structurally diverse *Actinobacteria*-produced lasso peptides.

A)

Strain Name	MIC (μg/mL)
Bacillus subtilis ATCC 6051	4
Staphylococcus aureus ATCC 25923	8
Micrococcus luteus ATCC 4698	8
Enterococcus faecalis ATCC 29212	>32
Escherichia coli ATCC 25922	>32

B)

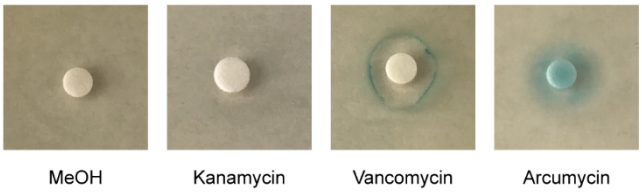

Figure 5. Antimicrobial activity of arcumycin. (**A**) MIC values of arcumycin against select microbial strains. (**B**) Activation of transcriptional *liaI* reporter in *B. subtilis* induced by treatment with arcumycin and positive control vancomycin.

Conclusion

In summary, through bioinformatic analyses of antimicrobial lasso peptides from Actinobacteria, we have demonstrated the potential of these techniques to relate lasso peptide sequence to antibiotic bioactivity. Analysis of the phylogenetic relationships of lasso peptides also illustrated a close relationship between genetic sequence and the previously described structure-based classification system. Using bioinformatics we identified several putative class I lasso peptides that were predicted to possess antimicrobial activity. Fermentation of the representative *Streptomyces* sp. NRRL F-5639 strain hit resulted in the production of a novel class I lasso peptide, arcumycin, which was structurally characterized and subsequently shown to possess antimicrobial activity against several Gram-positive bacterial test strains. Mode of action studies suggested that arcumycin interferes with lipid II biosynthesis, a mode of action shared with the class I lasso peptide siamycin I.[53] Together, these studies may help inform the bioinformatics-guided discovery of new lasso peptides with biological activities.

Experimental Section

General methods and materials

Bacterial media components were obtained from Affymetrix, Fisher Scientific, Millipore-Sigma, and BD Difco Laboratories. A sample of Pharmamedia was obtained from Archer Daniels Midland Company, and fish meal was purchased from Coyote Creek Organic Feed Mill and Farm. Tomato paste, oatmeal, cornmeal, V8 juice, whole wheat flour, millet flour, and nutritional yeast were purchased from Harris Teeter. All chemical solvents and reagents were purchased through Millipore-Sigma with no further purification. Preparative and analytical reverse-phase HPLC purifications were performed on a Waters Prep 150B system with a Phenomenex octadecyl silica (C18) column (250 mm x 21 mm, 10 μm, 300 Å) or Phenomenex C18 column (250 x 4.6 mm, 5 μm, 300 Å). Mass spectrometry analyses were collected on the following instruments: tandem MS/MS spectrometry was performed using Fusion Lumos Orbitrap mass spectrometer, matrix-assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF) was performed using a Bruker Autoflex Speed LRF MALDI-TOF System, LC-MS analysis was done on an Agilent 6460 Triple Quadrupole system, and HR-MS spectra were collected on an Agilent 6224 LC/MS-TOF instrument.

Bioinformatics

The full biosynthetic gene loci for twelve of the fifteen known antimicrobial lasso peptides are deposited in the NCBI database (**Table S1**). The precursor peptide within each cluster was identified and used as the query for twelve individual BLASTp[47] searches against the NCBI database. All resulting sequences query were selected as initial hits. To verify whether the BLASTp hits were indeed part of lasso peptide BGCs, the available genomic sequences (whole genome or BGC) from NCBI GenBank for each hit were analyzed by antiSMASH 5.0.[48] A total of 331 unique BGCs were identified (**Dataset 1**). The phylogenetic visualization tool CORASON[52] was used to perform multi-locus phylogenetic analyses to understand the evolutionary relationships of lasso peptide BGCs. For this analysis, the lasso cyclase (Protein Accession WP_103143102.1) from the chaxapeptin biosynthetic pathway was selected as the query sequence from the reference cluster of *Streptomyces leeuwenhoekii* strain C58 (NZ_LFEH01000114). The gbks files collected from antiSMASH for all 331 lasso peptide BGCs were included within the analysis using an E-value of 10^{-15}, e-cluster and e-core of 10^{-3}, rescale value of 85000, bitscore of 0, and a cluster ratio of 10. The phylogenetic tree generated from CORASON was exported using MUSCLE v.3.8.31[65] and visualized using iTOL[66] to generate an unrooted maximum-likelihood tree. To further visualize the lasso peptide BGC landscape, the program BiG-SCAPE[52] was used to generate the sequence similarity network for the 331 lasso peptide BGCs using the gbks files from antiSMASH. The cut-off distance was set at 0.3, global alignment mode was activated, the singleton parameter was used to visualize nodes that did not contain an edge to the network, and clan-cutoff was employed to deactivate the second layer of cluster to ensure all BGCs were included in the RiPP family analysis. The network generated through BiG-SCAPE was visualized using Cytoscape version 3.7.1.[67]

Media screen for lasso peptide production from *Streptomyces* sp. NRRL F-5639

A frozen glycerol stock of *Streptomyces* sp. NRRL F-5639, obtained from the Agricultural Research Service Culture Collection, was used to inoculate ISP II agar plates and incubated at 28 °C for 5 days. Spores were collected through scraping of the agar plate and used to inoculate 10 mL seed cultures. Either reported seed media conditions[38,40,54,55] or ISP1 were used for seed cultures. Seed cultures were grown for 5 days at 28 °C with orbital shaking at 200 rpm then used to inoculate 10 mL production cultures at 5% v/v. Production media included previously reported media conditions for the production of class I lasso peptides as well as media conditions disclosed for Actinomycete metabolite[59] profiling. Production cultures were incubated at 28 °C with orbital shaking at 200 rpm for 7 days. To screen each condition, after production, cells were collected through centrifugation for 10 min at 4225 x g at RT. The aqueous supernatant was separated from the cell pellet, and the pellet was resuspended in MeOH and extracted overnight by stirring the solution at RT. To monitor for lasso peptide production, both the aqueous supernatant and MeOH extract of cell mass were screened using MALDI-TOF-MS.

Scale-up production, isolation, and purification of arcumycin

For large-scale production of arcumycin, RAM media conditions were selected.[59] A frozen vegetative stock of *Streptomyces* sp. NRRL F-5639 was prepped by mixing equal volumes of ISP1 seed media with 40% glycerol followed by flash freezing with liquid nitrogen and storage at -80 °C. The vegetative stock was used to inoculate ISP II agar plates and incubated at 28 °C for 5 days. Spores were collected and used to inoculate 30 mL of ISPI seed medium, which was grown for 5 days at 28 °C with orbital shaking at 200 rpm. Seed culture was used to inoculate 500 mL of RAM production medium at 5% v/v. Production media was incubated at 28 °C with orbital shaking at 250 rpm for 7 days. Cells were collected through centrifugation, resuspended in MeOH for cell lysis, and extracted by stirring overnight at RT. After extraction, cell debris was collected through centrifugation and MeOH extracts were dried under reduced pressure to a residue. After resuspension in an H_2O/MeCN solution, the crude extract was subjected to RP-HPLC purification using a Jupiter C18, 250×21.2 mm column with a linear gradient of 10–90% B over 60 minutes, where mobile phase A was 0.1% TFA in H_2O and mobile phase B was 0.06% TFA in MeCN. Semi-pure arcumycin was further purified by HPLC using a Vydac C18 250×10 mm column using the same gradient and run-time giving a yield of 570 µg of pure arcumycin.

Carboxypeptidase Y digestion of arcumycin

Analyses were performed using crude MeOH extracts of arcumycin. For each reaction, 50 µL of crude lasso peptide was dried under reduced pressure. The residue was resuspended in 15 µL of buffer containing 50 mM MES (pH 6.7), 1 mM $CaCl_2$, and 50 mM DTT and reduced for 1 h at room temperature. To alkylate reduced peptide, iodoacetamide (15 mM, 5 µL) in 50 mM MES buffer (pH 6.7) with 1 mM $CaCl_2$ was added and allowed to react for 1 h at room temperature. Both samples containing intact or reduced/alkylated arcumycin were digested with carboxypeptidase Y (5 µg) overnight at RT. Samples were analyzed using RP-HPLC-MS with a 2.6 µm EVO–C18, 100×3 mm column with a gradient of 10‑90% B over 45 minutes, where mobile phase A was 100: 3:0.3 H_2O/MeOH/TFA and mobile phase B was 100: 3:0.3 MeCN/H_2O/TFA. ESI-MS was performed in positive ion mode and extracted ion chromatograms (EIC) were collected for the doubly charged peptide species.

ESI-MS/MS based structural determination.

For ESI-MS/MS, crude reduced/alkylated peptide was dissolved in 70% aq. MeCN with 0.1% TFA. The sample was analyzed with a Thermo Exploris 480 mass spectrometer at 2.5 µL/min. The sample was run over a 1 x 10 mm Water ACQUITY Premier CSH column with a gradient of 3‑60% B over 45 minutes where mobile phase A was H_2O and mobile phase B was 0.1% TFA in MeCN. Data was collected at 120 K for full MS scans and 120 K for MS/MS scans. Arcumycin was analyzed through MS/MS higher-energy C-trap dissociation (HCD) fragmentation for $[M+2H]^{2+}$ charge state using a normalized collision energy of 40. Spectra data was analyzed using Interactive Peptide Spectral Annotator.[61]

MIC assays

The antimicrobial activity of arcumycin was determined using a broth microdilution assay with ampicillin acting as a positive control. *Bacillus subtilis* ATCC 6051, *Staphylococcus aureus* ATCC 25923, *Enterococcus faecalis* ATCC 29212, *Micrococcus luteus* ATCC 4698, and *Escherichia coli* ATCC 25922 were grown at 37 °C, 220 rpm in cation adjusted Mueller-Hinton Broth (MHB), except *M. luteus,* which was grown at 28 °C. Antimicrobial peptides were dissolved in MeOH and tested from a range of 0.125–32 µg/mL. The bacteria culture was diluted to a concentration of $2x10^5$ CFU/mL using MHB. Plates were incubated overnight at 37 °C, except *M. luteus,* which was grown at 28 °C, and inspected visually to determine the MIC. Each reported MIC value is the average of two biological replicates.

***B. subtilis lial-β-gal* promoter fusion reporter assay**

The ability of arcumycin to induce cell envelope stress was evaluated using the strain 1A980 *Bacillus subtilis* BSF2470 obtained from the Bacillus Genetic Stock Center. *B. subtilis* BSF2470 was grown overnight at 37 °C, 220 rpm in Luria Broth (LB) containing erythromycin (1 µg/mL) for bacterial selection. For the reporter assay, the saturated *B. subtilis BSF2470* culture (40 uL) was plated on freshly prepared LB agar plates supplemented with erythromycin (1 µg/mL) and X-gal (200 µg/mL). Each disk was loaded with either arcumycin (6.4 µg), vancomycin (6.4 µg), kanamycin (6.4 µg), or MeOH (20 µL). Plates were visualized after 22 h for production of zones of inhibition.

Acknowledgements

We wish to thank Duke University the generous support of this research and members of the McCafferty lab for thoughtful discussion and critical reading of this manuscript. We also wish to thank Drs. Peter Silinski, Arthur Moseley, and Matt Foster for mass spectrometry assistance.

Keywords: lasso peptides • antimicrobials • bioinformatics • *Actinobacteria* • natural products

[1] R. I. Aminov, *Front. Microbiol.* **2010**, *1*, 134.

[2] Centers for Disease Control and Prevention., **2019**.

[3] O. Genilloud, *Nat. Prod. Rep.* **2017**, *34*, 1203–1232.

[4] M. H. Medema, M. A. Fischbach, *Nat. Chem. Biol.* **2015**, *11*, 639–648.

[5] P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K. D. Entian, M. A. Fischbach, J. S. Garavelli, U. Göransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Müller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. T. Reaney, S. Rebuffat, R. P. Ross, H. G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Süssmuth, J. R. Tagg, G. L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey, W. A. Van Der Donk, *Nat. Prod. Rep.* **2013**, *30*, 108–160.

[6] W. L. Cheung-Lee, A. J. Link, *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 1371–1379.

[7] M. Tsunakwa, S.-L. Hu, Y. Hoshino, D. J. Detlefson, S. E. Hill, T. Furumai, R. J. White, K. Kawano, S. Yamamoto, Y. Fukagawa, T. Oki, *J. Antibiot. (Tokyo).* **1995**, *48*, 433–434.

[8] H. Nakashima, K. Ichiyama, K. Inazawa, M. Ito, H. Hayashi, Y. Nishihara, E. Tshuhii, T. Kino, *Biol. Pharm. Bull.* **1996**, *19*, 405–412.

[9] S. Son, M. Jang, B. Lee, Y. S. Hong, S. K. Ko, J. H. Jang, J. S. Ahn, *J. Nat. Prod.* **2018**, *81*, 2205–2211.

[10] T. A. Knappe, F. Manzenrieder, C. Mas-Moruno, U. Linne, F. Sasse, H. Kessler, X. Xie, M. A. Marahiel, *Angew. Chemie Int. Ed.* **2011**, *50*, 8714–8717.

[11] T. A. Knappe, U. Linne, S. Zirah, S. Rebuffat, X. Xie, M. A. Marahiel, *J. Am. Chem. Soc.* **2008**, *130*, 11446–11454.

[12] M. O. Maksimov, I. Pelczer, A. J. Link, *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 15223–15228.

[13] M. Metelev, A. Arseniev, L. B. Bushin, K. Kuznedelov, T. O. Artamonova, R. Kondratenko, M. Khodorkovskii, M. R. Seyedsayamdost, K. Severinov, *ACS Chem. Biol.* **2017**, *12*, 814–824.

[14] S. Tan, G. Moore, J. Nodwell, *Antibiotics* **2019**, *8*, 117.

[15] P. Zhao, Y. Xue, W. Gao, J. Li, X. Zu, D. Fu, S. Feng, X. Bai, Y. Zuo, P. Li, *Peptides* **2018**, *103*, 48–59.

[16] M. O. Maksimov, S. J. Pan, A. J. Link, *Nat. Prod. Rep.* **2012**, *29*, 996–1006.

[17] R. Ducasse, K.-P. Yan, C. Goulard, A. Blond, Y. Li, E. Lescop, E. Guittet, S. Rebuffat, S. Zirah, *ChemBioChem* **2012**, *13*, 371–380.

[18] J. D. Hegemann, *ChemBioChem* **2020**, *21*, 7–18.

[19] M. Zimmermann, J. D. Hegemann, X. Xie, M. A. Marahiel, *Chem.*

Biol. **2013**, *20*, 558–569.

[20] H. Martin-Gómez, J. Tulla-Puche, *Org. Biomol. Chem.* **2018**, *16*, 5065.

[21] W. L. Cheung-Lee, M. E. Parry, A. J. Cartagena, S. A. Darst, A. J. Link, *J. Biol. Chem.* **2019**, *294*, 6822–6830.

[22] J. D. Hegemann, C. J. Schwalen, D. A. Mitchell, W. A. Van der Donk, *Chem. Commun.* **2018**, *54*, 9007–9010.

[23] W. L. Cheung-Lee, M. E. Parry, C. Zong, A. J. Cartagena, S. A. Darst, N. D. Connell, R. Russo, A. J. Link, *ChemBioChem* **2020**, *21*, 1335–1340.

[24] S. Um, Y. J. Kim, H. Kwon, H. Wen, S. H. Kim, H. C. Kwon, S. Park, J. Shin, D. C. Oh, *J. Nat. Prod.* **2013**, *76*, 873–879.

[25] C. Cortés-Albayay, S. A. Jarmusch, F. Trusch, R. Ebel, B. A. Andrews, M. Jaspars, J. A. Asenjo, *J. Org. Chem.* **2020**, *85*, 1661–1667.

[26] E. Gavrish, C. S. Sit, S. Cao, O. Kandror, A. Spoering, A. Peoples, L. Ling, A. Fetterman, D. Hughes, A. Bissell, H. Torrey, T. Akopian, A. Mueller, S. Epstein, A. Goldberg, J. Clardy, K. Lewis, *Chem. Biol.* **2014**, *21*, 509–518.

[27] F. Xu, Y. Wu, C. Zhang, K. M. Davis, K. Moon, L. B. Bushin, M. R. Seyedsayamdost, *Nat. Chem. Biol.* **2019**, *15*, 161–168.

[28] J. P. Gomez-Escribano, J. F. Castro, V. Razmilic, S. A. Jarmusch, G. Saalbach, R. Ebel, M. Jaspars, B. Andrews, J. A. Asenjo, M. J. Bibb, *Appl. Environ. Microbiol.* **2019**, *85*.

[29] M. A. Skinnider, C. W. Johnston, R. E. Edgar, C. A. Dejong, N. J. Merwin, P. N. Rees, N. A. Magarvey, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, E6343–E6351.

[30] M. O. Maksimov, A. J. Link, *J Ind Microbiol Biotechnol* **2014**, *41*, 333–344.

[31] J. I. Tietz, C. J. Schwalen, P. S. Patel, T. Maxson, P. M. Blair, H.-C. Tai, U. I. Zakai, D. A. Mitchell, *Nat. Chem. Biol.* **2017**, *13*, 470–478.

[32] A. M. Kloosterman, P. Cimermancic, S. S. Elsayed, C. Du, M. Hadjithomas, M. S. Donia, M. A. Fischbach, G. P. van Wezel, M. H. Medema, *PLOS Biol.* **2020**, *18*, e3001026.

[33] E. V. Bratovanov, K. Ishida, B. Heinze, S. J. Pidot, T. P. Stinear, J. D. Hegemann, M. A. Marahiel, C. Hertweck, *ACS Chem. Biol.* **2020**, *15*, 1169–1176.

[34] C. Zong, W. L. Cheung-Lee, H. E. Elashal, M. Raj, A. J. Link, *Chem. Commun.* **2018**, *54*, 1339–1342.

[35] S. Kodani, K. Unno, *J. Ind. Microbiol. Biotechnol.* **2020**, *47*, 703–714.

[36] S. S. Elsayed, F. Trusch, H. Deng, A. Raab, I. Prokes, K. Busarakam, J. A. Asenjo, B. A. Andrews, P. Van West, A. T. Bull, M. Goodfellow, Y. Yi, R. Ebel, M. Jaspars, M. E. Rateb, *J. Org. Chem.* **2015**, *80*, 10252–10260.

[37] N. Takasaka, I. Kaweewan, M. Ohnishi-Kameyama, S. Kodani, *Lett. Appl. Microbiol.* **2017**, *64*, 150–157.

[38] Y. Li, R. Ducasse, S. Zirah, A. Blond, C. Goulard, E. Lescop, C. Giraud, A. Hartke, E. Guittet, J. L. Pernodet, S. Rebuffat, *ACS Chem. Biol.* **2015**, *10*, 2641–2649.

[39] I. Kaweewan, M. Ohnishi-Kameyama, S. Kodani, *J. Antibiot. (Tokyo).* **2017**, *70*, 208–211.

[40] I. Kaweewan, H. Hemmi, H. Komaki, S. Harada, S. Kodani, *Bioorganic Med. Chem.* **2018**, *26*, 6050–6055.

[41] M. Iwatsuki, H. Tomoda, R. Uchida, H. Gouda, S. Hirono, S. Omura, *J. Am. Chem. Soc.* **2006**, *128*, 7486–7491.

[42] Y. Esumi, Y. Suzuki, Y. Itoh, M. Uramoto, K. I. Kimura, M. Goto, M. Yoshihama, T. Ichikawa, *J. Antibiot. (Tokyo).* **2002**, *55*, 296–300.

[43] T. Zyubko, M. Serebryakova, J. Andreeva, M. Metelev, G. Lippens, S. Dubiley, K. Severinov, *Chem. Sci.* **2019**, *10*, 9699–9707.

[44] S. Kodani, Y. Inoue, M. Suzuki, H. Dohra, T. Suzuki, H. Hemmi, M. Ohnishi-Kameyama, *European J. Org. Chem.* **2017**, *2017*, 1177–1183.

[45] S. Zhu, C. D. Fage, J. D. Hegemann, A. Mielcarek, D. Yan, U. Linne, M. A. Marahiel, *Sci. Rep.* **2016**, *6*, 1–12.

[46] C. Cheng, Z. C. Hua, *Front. Bioeng. Biotechnol.* **2020**, *8*, 1131.

[47] NCBI Resource Coordinators, *Nucleic Acids Res.* **2015**, *43*, D6-17.

[48] K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema, T. Weber, *Nucleic Acids Res.* **2019**, *47*, W81–W86.

[49] Y. Morishita, S. Chiba, E. Tsukuda, T. Tanaka, T. Ogawa, M. Yamasaki, M. Yoshida, Y. Matsuda, I. Kawamoto, *J. Antibiot. (Tokyo).* **1994**, *47*, 269–275.

[50] V. Valiante, M. C. Monteiro, J. Martín, R. Altwasser, N. El Aouad, I. González, O. Kniemeyer, E. Mellado, S. Palomo, N. De Pedro, I. Pérez-Victoria, J. R. Tormo, F. Vicente, F. Reyes, O. Genilloud, A. A. Brakhage, *Antimicrob. Agents Chemother.* **2015**, *59*, 5145–5153.

[51] S. Sugai, M. Ohnishi-Kameyama, S. Kodani, *Appl. Biol. Chem.* **2017**, *60*, 163–167.

[52] J. C. Navarro-Muñoz, N. Selem-Mojica, M. W. Mullowney, S. A. Kautsar, J. H. Tryon, E. I. Parkinson, E. L. C. De Los Santos, M. Yeong, P. Cruz-Morales, S. Abubucker, A. Roeters, W. Lokhorst, A. Fernandez-Guerra, L. T. D. Cappelini, A. W. Goering, R. J. Thomson, W. W. Metcalf, N. L. Kelleher, F. Barona-Gomez, M. H. Medema, *Nat. Chem. Biol.* **2020**, *16*, 60–68.

[53] S. Tan, K. C. Ludwig, A. Mü, T. Schneider, J. R. Nodwell, *ACS Chem. Biol* **2019**, *14*, 31.

[54] T. Yamamoto, H. Matsui, K. Yamaji, T. Takahashi, A. Øverby, M. Nakamura, A. Matsumoto, K. Nonaka, T. Sunazuka, S. Ōmura, H. Nakano, *J. Infect. Chemother.* **2016**, *22*, 587–592.

[55] J. Nakayama, E. Tanaka, R. Kariyama, K. Nagata, K. Nishiguchi, R. Mitsuhata, Y. Uemura, M. Tanokura, H. Kumon, K. Sonomoto, *J. Bacteriol.* **2007**, *189*, 1358–1365.

[56] S. Duquesne, D. Destoumieux-Garzón, S. Zirah, C. Goulard, J. Peduzzi, S. Rebuffat, *Chem. Biol.* **2007**, *14*, 793–803.

[57] Z. Feng, Y. Ogasawara, S. Nomura, T. Dairi, *ChemBioChem* **2018**, *19*, 2045–2048.

[58] M. Shao, J. Ma, Q. Li, J. Ju, *Mar. Drugs* **2019**, *17*, 127.

[59] J. R. Tormo, J. B. García, M. DeAntonio, J. Feliz, A. Mira, M. T. Díez, P. Hernández, F. Peláez, *J. Ind. Microbiol. Biotechnol.* **2003**, *30*, 582–588.

[60] S. A. Jarmusch, I. Feldmann, B. Blank-Landeshammer, C. Cortés-Albayay, J. F. Castro, B. Andrews, J. A. Asenjo, A. Sickmann, R. Ebel, M. Jaspars, *J. Antibiot. (Tokyo).* **2020**, *73*, 772–779.

[61] D. R. Brademan, N. M. Riley, N. W. Kwiecien, J. J. Coon, *Mol. Cell. Proteomics* **2019**, *18*, S193–S201.

[62] M. O. Maksimov, A. J. Link, *J. Am. Chem. Soc.* **2013**, *135*, 12038–12047.

[63] M. Metelev, J. I. Tietz, J. O. Melby, P. M. Blair, L. Zhu, I. Livnat, K. Severinov, D. A. Mitchell, *Chem. Biol.* **2015**, *22*, 241–250.

[64] T. Mascher, S. L. Zimmer, T. A. Smith, J. D. Helmann, *Antimicrob. Agents Chemother.* **2004**, *48*, 2888–2896.

[65] F. Madeira, Y. mi Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R. N. Tivey, S. C. Potter, R. D. Finn, R. Lopez, F. Madeira, Y. mi Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, R. Lopez, *Nucleic Acids Res.* **2019**, *47*, W636–W641.

[66] I. Letunic, P. Bork, *Nucleic Acids Res.* **2019**, *47*, W256–W259.

[67] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, *Genome Res.* **2003**, *13*, 2498–2504.

Discovery of the Class I Antimicrobial Lasso Peptide Arcumycin

Lydia M. Stariha and Dewey G. McCafferty*

Table of Contents

I. Supplemental tables

Table S1. *Actinobacteria* produced antimicrobial lasso peptides.

Lasso peptide	Producing strain	Accession number	Protein ID	Sequence	Bioactivity (μg/mL)
siamycin I[1]	*Streptomyces* sp. M271	LC381634	BBD82040.1	CLGVGSCNDFAGCGYAIVCFW	*B. subtilis* 1.6 *S. epidermis* 4.2 *S. saprophyticus* 65 *S. aureus* 8 *E. faecalis* 8 *M. luteus* 1.6
siamycin II[1]	*Streptomyces sp.* AA389 (ATCC 55286)	not available	not available	CLGIGSCNDFAGCGYAIVCFW	*B. subtilis* 1.6 *S. aureus* 3.1 *M. luteus* 1.6
aborycin[2]	*Streptomyces* sp. SCIO Z50098	MKCP01000039.1	allorf_022746_022874	CLGIGSCNDFAGCGYAVVCFW	*B. brevis* 25 *B. subtilis* 2 *S. viridochromeogenes 2* *E. faecalis* 8 *S. aureus* 8 *P. saccharophilia* 15
specialicin[3]	*Streptomyces specialis* GW41-1564	NZ_LN929757	WP_104530897.1	CLGVGSCVDFAGCGYAVVCFW	*M. luteus* 8
sviceucin[4]	*Streptomyces sviceus* ATCC 29083	NZ_CM000951	WP_106433542.1	CLGVGSCVDFAGCGYAVVCFW	*B. megaterium* 2.6 *E. faecalis* 20.82 *L. brevis* 20.82 *L. bulgaricus* 2.6 *S. aureus* 5.2 *L. sakei* 10.41 *M. luteus* 20.82 *Streptomyces* sp 523 5.2
anantin B1/B2[5]	*Streptomyces* sp. NRRL S-146	NZ_JOAW01000056	WP_158761618.1	CVWGGDCTDFLGCGTAWICV	*B. subtilis* 20.8 *E. coli* 166
lariatin A/B[6]	*Rhodococcus jostii*	AB593691	BAL72546.1	GFIGWGNDIFGHYSGGF	*M. smegmatis* 3.13
lassomycin[7]	*Lentzea kentuckyensis* sp IS009804	not available	not available	GSQLVYREWVGHSNVIKPGP	*M. tuberculosis* 0.4 *M. avium* 0.78 *M. smegmatis* 0.7
propeptin 1/2[8]	*Microbispora rosea*	NZ_JNZQ01000012	WP_142556472.1	GLRRLFADNLVGRRNLMe	*M. phlei* - 40 μg/disk *X. orzyae* - 40 μg/disk *P. aeruginosa* - 40 μg/disk
streptomonomicin[9]	*Streptomonospora alba* strain YIM 90003	NZ_JROO01000009	WP_103055249.1	GYPWWDYRDLFGGHTFISP	*B. anthracis* 4 *B. halodurans* 8 *B. cereus* 8 *B. subtilis* 64 *L. monocytogenes* 32 *E. faecalis* 64 *S. aureus* 128
achromosin[10]	*Streptomyces achromogenes* strain NRRL B-2120	NZ_JODT01000002	WP_157846926.1	SLGSSPYNDILGYPALIVIYP	*M. luteus* - 10 μg/disk
chaxapeptin[11]	*Streptomyces leeuwenhoekii* strain C58	NZ_LFEH01000114	WP_103143102.1	GIGSQTWDTIWLWD	*S. aureus* 30 *B. subtilis* 35
actinokineosin[12]	*Actinokineospora spheciospongiae* strain EG49	AYXG01000139	EWC60872.1	GFGSKPLDSFGLNFF	*M. luteus* - 50 μg/disk
sphaericin[13]	*Planomonospora sphaerica* strain JCM 9374	NZ_BDCX01000012 (not complete, not used as query)	allorf_158350_158469	GLPIGWWIERPSGWYFPI	*M. luteus* - 50 μg/disk
LP2006[5]	*Nocardiopsis alba* DSM 43377	NZ_ANAC01000010.1	WP_014911574.1	GRPNWGFENDWSCVRVC	*B. anthracis* 12.5 *B. subtilis* 100 *E. faecalis* 25 *M. smegmatis* 25

Table S2. Unique BGCs of class I lasso peptides.

BGCs	Protein ID	Strains
siamycin I	BBD82040.1	*Streptomyces* sp. M-271
siamycin II	WP_107308646.1	*Streptomyces* sp. F-7
specialicin	WP_104530897.1	*Streptomyces specialis* strain GW41-1564T
humidimycin	QHZ32192.1	*Streptomyces humidus* strain CA-100629
aborycin	WP_106963308.1	*Streptomyces* sp. ZS0098
		Streptomyces griseorubens strain JSD-1
		Streptomyces sp. SMS_SU21
		Streptomyces sp. Z38
		Streptomyces sp. HNS054
		Streptomyces sp. E2N171
siamycin BGC 1	WP_107070498.1	*Streptomyces nodosus* strain ATCC 14899
siamycin BGC 2	WP_106969722.1	*Streptomyces* sp. NRRL F-6674
		Streptomyces sp. NRRL F-6676
		Streptomyces sp. NRRL B-12105
		Streptomyces sp. NRRL F-6677
siamycin BGC 3	WP_103079659.1	*Streptomyces* sp. E1N211
		Streptomyces sp. CB02056
		Kitasatospora xanthocidica strain MMS17-GH009
siamycin BGC 4	WP_153484251.1	*Streptomyces katsurahamanus* strain T-272
siamycin BGC 5	WP_123992009.1	*Streptomyces* sp. SAI-083
siamycin BGC 6	WP_107121080.1	*Streptomyces* sp. BK205
		Streptomyces griseorubiginosus strain DSM 40469
		Streptomyces griseorubiginosus strain 3E-1
		Streptomyces griseorubiginosus strain DSM 40125
		Streptomyces olivochromogenes strain NBRC 3561
siamycin BGC 7	SFF02595.1	*Streptomyces mirabilis* strain OK461
siamycin BGC 8	KPI21041.1	*Actinobacteria bacterium* OK006
siamycin BGC 9	WP_103079660.1	*Streptomyces olivochromogenes* strain NBRC 3561
		Streptomyces sp. 93
		Streptomyces sp. 3212.2
		Streptomyces sp. 94
		Streptomyces sp. Ag82_O1-15
		Streptomyces sp. OV198
		Streptomyces sp. 42
		Streptomyces sp. 74
		Streptomyces sp. RLB3-6
		Streptomyces sp. S1D4-23
		Streptomyces sp. S1A1-7
		Streptomyces sp. OK228
		Streptomyces sp. RLB1-9
		Streptomyces sp. RLB3-5
siamycin 10	WP_141635469.1	*Kitasatospora* sp. MMS16-CNU292
sviceucin	WP_106433542.1	*Streptomyces sviceus* ATCC 29083
		Streptomyces sp. CC0208
		Streptomyces sp. SID5470
sviceucin BGC 1	WP_130347269.1	*Herbihabitans rhizosphaerae* strain DSM 101727
novel BGCI 1	WP_156966645.1	*Actinomyces israelii* DSM 43320
novel BGCI 2	WP_115737004.1	*Actinomyces denticolens* strain NCTC11490
		Actinomyces denticolens strain DSM 20671
		Actinomyces denticolens strain DSM 20671
		Actinomyces denticolens strain DSM 20671
arcumycin	WP_106967485.1	*Streptomyces* sp. NRRL F-5639
		Streptomyces albus strain NRRL B-2465
		Streptomyces albus strain NRRL B-12378
		Streptomyces albus strain NRRL B-1335
		Streptomyces albus strain NRRL B-1685
		Streptomyces graminis strain NRRL B-1570
		Streptomyces albus strain NRRL B-1811
		Streptomyces albus strain NBRC 13014
novel BGCII 1	WP_107488709.1	*Streptomyces* sp. CNZ279
novel BGCII 2	WP_106963780.1	*Streptomyces albus* strain CAS922
		Streptomyces sp. HPH0547
		Streptomyces sp. NRRL F-5917
		Streptomyces albus strain NRRL B-2362
		Streptomyces albus strain DSM 40763

Table S3. Open reading frame homology comparison of percent identity to siamycin I for predicted proteins in class I lasso peptide BGCs.

BGCs	R1	G	D1	D2	A	C	B1	B2	D3	D4	E	F	H	R2
siamycin I	-	-	-	-	-	-	-	-	-	-	-	-	-	-
siamycin II	58.80	71.09	72.85	72.29	88.10	77.05	69.77	89.78	76.64	82.64	72.55	65.67	72.71	84.75
specialicin	65.91	69.48	73.94	70.00	80.49	75.8	68.60	88.03	78.50	81.51	67.31	70.82	74.49	85.65
humidimycin	63.35	73.84	72.96	77.51	88.10	77.61	74.42	87.86	81.31	85.86	80.00	71.55	74.89	87.50
aborycin	60.89	71.467	72.51	73.49	85.71	76.72	67.44	86.33	75.70	83.39	71.24	65.40	72.94	86.10
siamycin BGC I	62.01	64.93	75.50	72.80	94.87	77.26	71.76	87.23	81.50	80.41	73.33	67.01	75.06	86.10
siamycin BGC 2	63.51	69.13	76.63	74.70	90.48	76.69	74.42	85.70	79.13	80.46	71.19	65.27	74.55	82.59
siamycin BGC 3	60.89	72.68	72.96	73.09	80.95	77.48	67.44	86.43	77.26	87.22	76.30	70.59	77.4	87.44
siamycin BGC 4	63.48	73.13	73.94	71.60	90.48	35.29	69.77	86.18	81.10	83.85	76.92	68.49	75.11	87.50
siamycin BGC 5	89.29	84.62	83.86	92.40	97.62	87.48	93.10	87.16	91.28	90.66	80.47	85.95	86.59	91.03
siamycin BGC 6	86.96	83.97	81.66	94.00	95.24	86.83	91.95	86.49	90.34	92.04	76.24	86.70	86.26	91.03
siamycin BGC 7	96.52	94.96	92.38	97.20	100.00	96.57	100.00	88.96	96.57	97.23	89.60	91.77	94.09	99.11
siamycin BGC 8	99.13	98.67	95.87	99.60	100.00	98.69	100.00	100.00	98.75	99.31	99.42	97.12	97.50	100.00
siamycin BGC 9	99.57	99.20	95.24	99.60	100.00	98.86	100.00	100.00	99.38	99.31	98.34	98.35	97.95	100.00
siamycin BGC 10	62.21	N.A.	N.A.	N.A.	67.74	68.56	70.11	83.80	79.13	75.09	N.A.	N.A.	73.80	N.A.
sviceucin	40.83	45.27	55.74	42.97	54.05	50.98	50.57	61.15	67.20	55.71	46.04	31.73	N.A.	65.18
sviceucin BGC 1	36.41	46.08	61.03	40.16	61.11	53.36	46.99	55.56	64.44	57.36	45.71	44.17	N.A.	60.55
novel BGCI 1	33.33	N.A	54.67	30.49	48.28	32.17	N.A.	50.83	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
novel BGCI 2	25.89	N.A	54.63	28.50	50.00	32.80	28.49	44.44	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
arcumycin	32.14	46.13	57.95	40.16	63.41	48.68	49.41	59.26	61.90	54.83	N.A.	N.A.	N.A.	41.70
novel BGCII 1	43.26	42.36	57.58	45.16	60.98	45.36	49.4	58.75	61.59	52.83	N.A.	N.A.	N.A.	62.4
novel BGCII 2	32.14	46.13	57.95	39.18	60.98	48.68	49.41	59.26	61.90	54.83	N.A.	N.A.	N.A.	41.70

Table S4. Calculated and observed ions of MS/MS analysis of reduced and alkylated arcumycin with opening between Cys_1 and Asp_9 in macrocycle.

Fragment	Predicted mass	Observed mass	Error (ppm)	Fragment	Predicted mass	Observed mass	Error (ppm)
b_1	161.0380	161.0379	0.5881	y_8-NH_3	1018.4600	1018.4600	-0.0604
b_2	274.1222	274.1220	0.7377	y_8	1035.4870	1035.4860	0.4128
y_2	292.1328	292.1326	0.9863	y_9	1092.5080	1092.5080	0.4104
b_3	371.1748	371.1748	0.2183	y_9-NH_3	1839.7920	1839.7980	-3.2581
b_4-H_2O	440.1966	440.1962	0.8535	b_{10}-H_2O	1131.4240	1131.4230	0.4948
y_3	478.2120	478.2119	0.3581	y_{10}-NH_3	1188.5660	1188.5650	0.3639
b_6-H_2O	612.2458	612.2446	2.0043	y_{10}	1205.5920	1205.5920	0.5113
b_6	630.2554	630.2552	0.2918	y_{11}	1352.6610	1352.6600	0.4187
b_7-H_2O	772.2767	772.2753	1.8079	y_{12}-H_2O	1449.6800	1449.6770	2.1467
b_7	790.2861	790.2858	0.2922	b_{15}-H_2O	1674.6830	1674.6820	0.5751
y_6-NH_3	801.4073	801.4076	-0.3895	y_{14}-H_2O	1706.7630	1706.7600	1.8693
y_6	818.4351	818.4341	1.1115	b_{16}-H_2O	1745.7210	1745.7190	1.0386
y_7-NH_3	858.4286	858.4291	-0.4821	b_{17}-CO_2	1832.8160	1832.8240	-4.5708
y_7	875.4562	875.4556	0.7138	y_{15}	1075.4810	1075.4810	0.0181
b_9-H_2O	984.3557	984.3550	0.7443				

Table S5. Calculated and observed ions of MS/MS analysis of reduced and alkylated arcumycin cyclic portion with opening between Cys_1 and Asp_9 in macrocycle.

Fragment	Predicted mass	Observed mass	Error (ppm)
b_1	161.0380	161.0379	0.5881
y_2-H_2O	213.0871	213.0870	0.5478
b_2	274.1222	274.1220	0.7377
b_3	371.1748	371.1748	0.2183
y_3-H_2O	373.1179	373.1176	0.7657
b_4-H_2O	440.1966	440.1962	0.8535
y_4-H_2O	488.1461	488.1446	3.2255
b_6-H_2O	612.2458	612.2446	2.0043
b_6	630.2554	630.2552	0.2918
b_7-H_2O	772.2767	772.2753	1.8079
b_7	790.2861	790.2858	0.2922

Table S6. Calculated and observed ions of MS/MS analysis of reduced and alkylated arcumycin cyclic portion with opening between Leu_2 and Pro_3 in macrocycle.

Fragment	Predicted mass	Observed mass	Error (ppm)
b_2-H_2O	167.0816	167.0815	0.7764
b_2	185.0922	185.0921	0.7533
b_3-CO_2	198.1237	198.1236	-0.4680
b_3-H_2O	224.1031	224.1030	0.3977
b_3	242.1137	242.1135	0.6597
y_2	292.1328	292.1326	0.9863
b_4-H_2O	339.1299	339.1299	-0.0766
b_4	357.1408	357.1405	0.8516
b_5-H_2O	499.1610	499.1606	0.8979
b_5	517.1716	517.1711	0.8558
b_6-H_2O	596.2142	596.2133	1.5211
b_6	614.2242	614.2239	0.4736
y_5-H_2O	646.2296	646.2323	-4.2301
b_7-H_2O	711.2404	711.2403	0.2123
y_6-H_2O	761.2578	761.2593	-2.0182
b_8-H_2O	871.2729	871.2709	2.2587

II. Supplemental figures

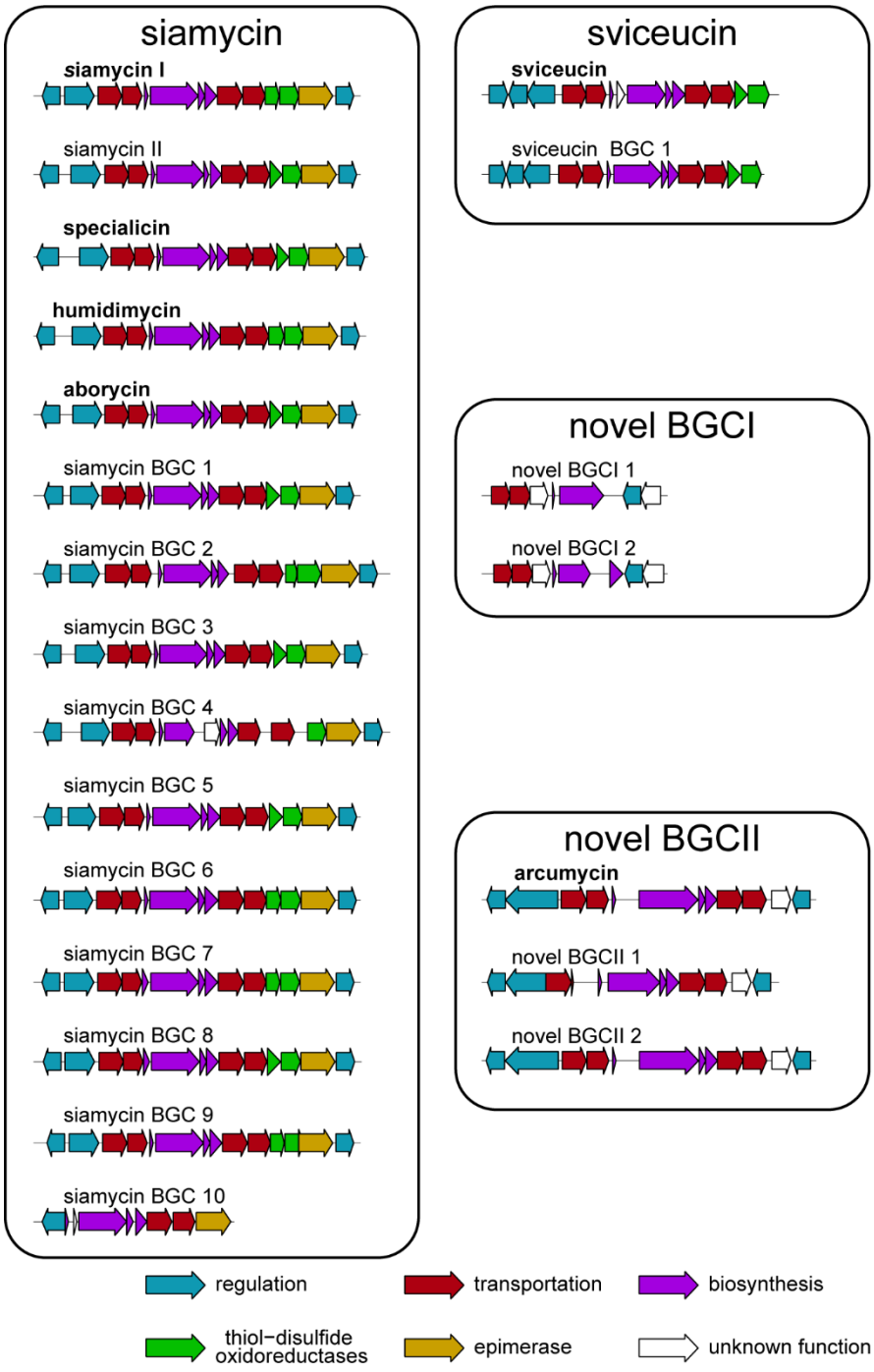

Figure S1. Open reading frame arrow diagram of 22 unique class I lasso peptide BGCs, which reveals four compositional archetypes: siamycin, sviceucin, novel BGCI, and novel BGCII. Lasso peptides that have been structurally characterized from the unique BGCs are bolded.

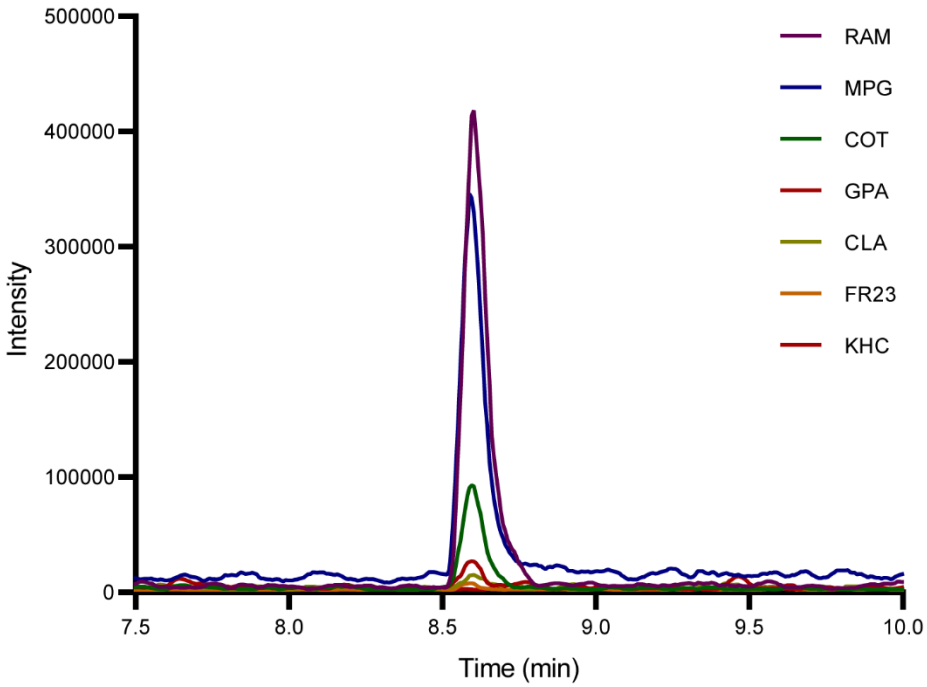

Figure S2. Extracted ion chromatogram of arcumycin production in media from Actinomycete metabolite profiling media screen. Extracts with arcumycin detected by MALDI-MS were dried to a residue, dissolved in 1 mL of MeOH, and analyzed to determine media with the highest production level. Only media conditions determined by MALDI to contain arcumycin were analyzed.

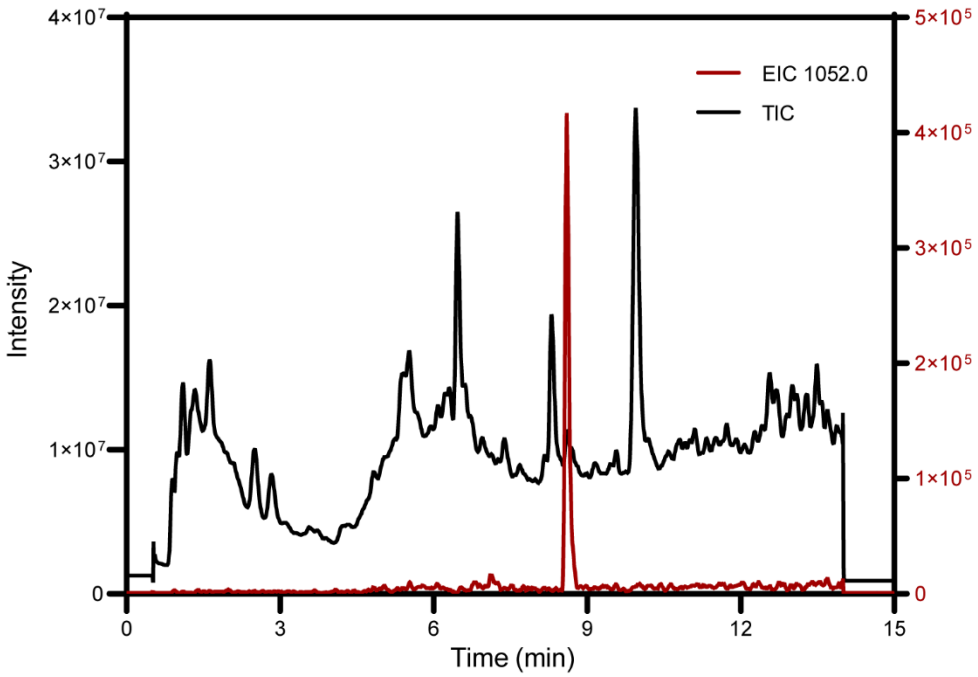

Figure S3. Identification of arcumycin in RAM production media. Total ion chromatogram of MeOH extract of cell pellet from *Streptomyces* NRRL F-5639 grown in RAM production media and extracted ion chromatogram for arcumycin (m/z = 1052 $[M+2H]^{2+}$).

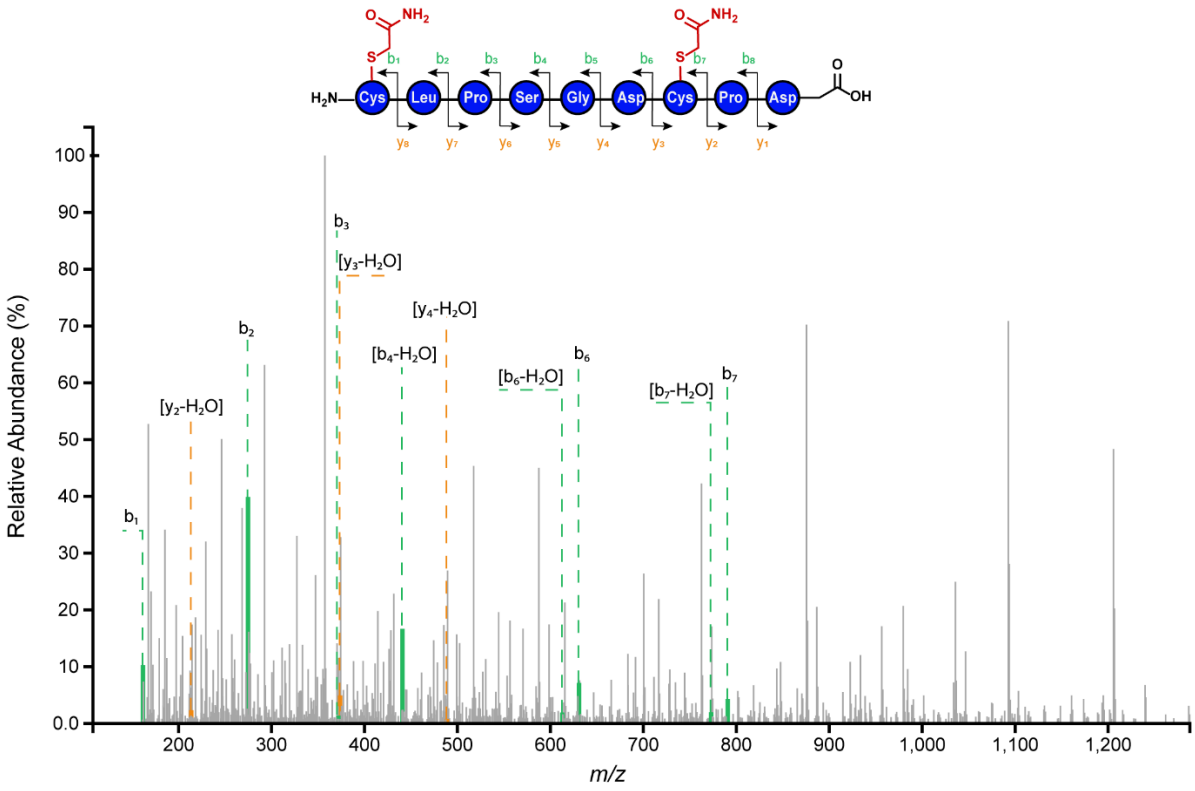

Figure S4. MS/MS fragmentation of reduced and alkylated arcumycin linear macrocycle opened at Cys_1 and Asp_9.

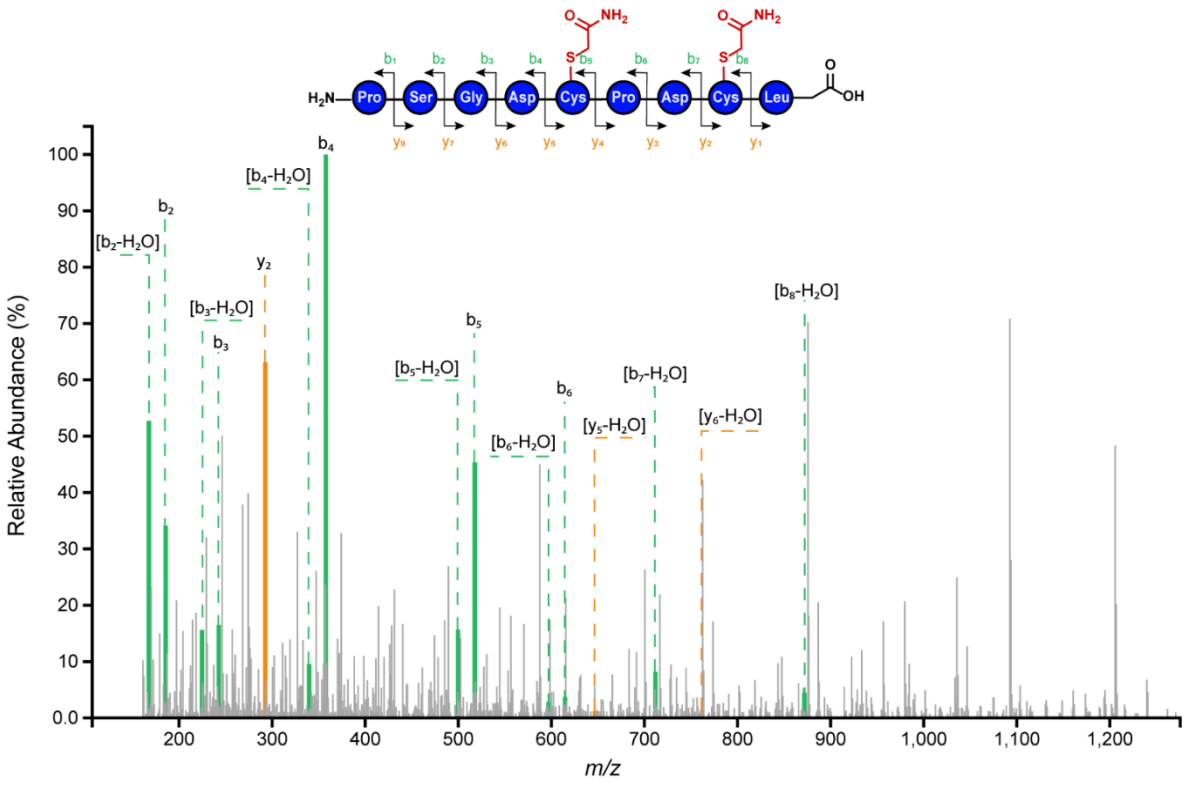

Figure S5. MS/MS fragmentation of reduced and alkylated arcumycin linear macrocycle opened at Leu_2 and Pro_3.

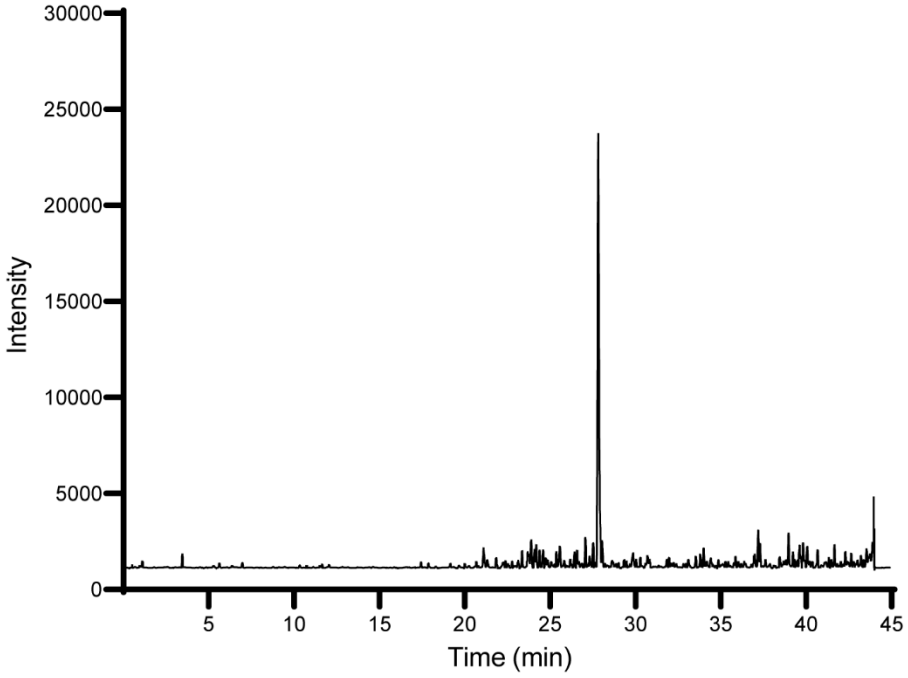

Figure S6. Extracted ion chromatogram of intact arcumycin (m/z = 1052 $[M+2H]^{2+}$).

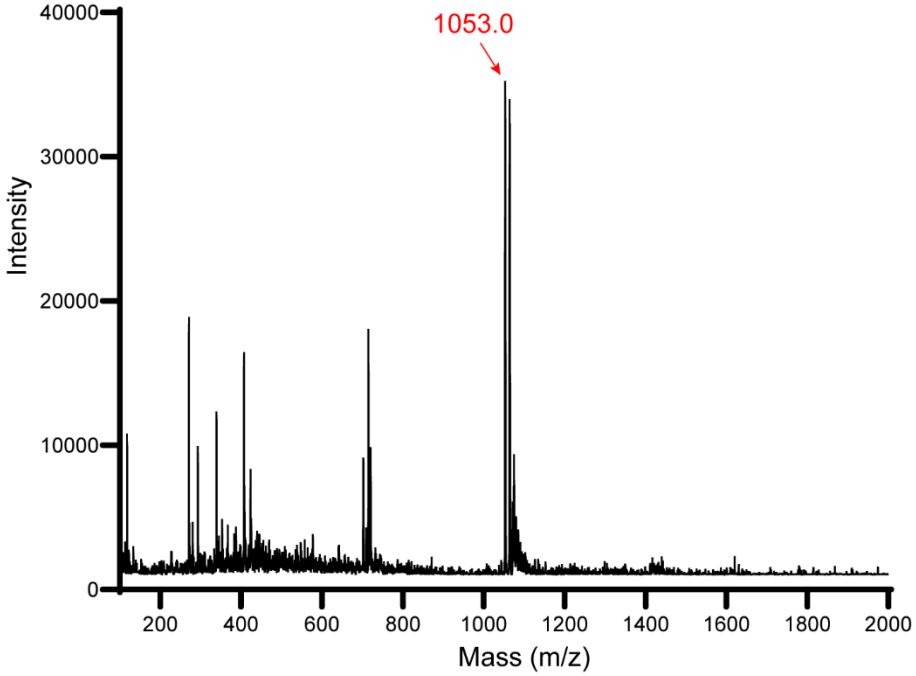

Figure S7. ESI-MS of intact arcumycin at rt = 27.8 min untreated (calc. $[M+2H]^{2+}$ =1052.4490).

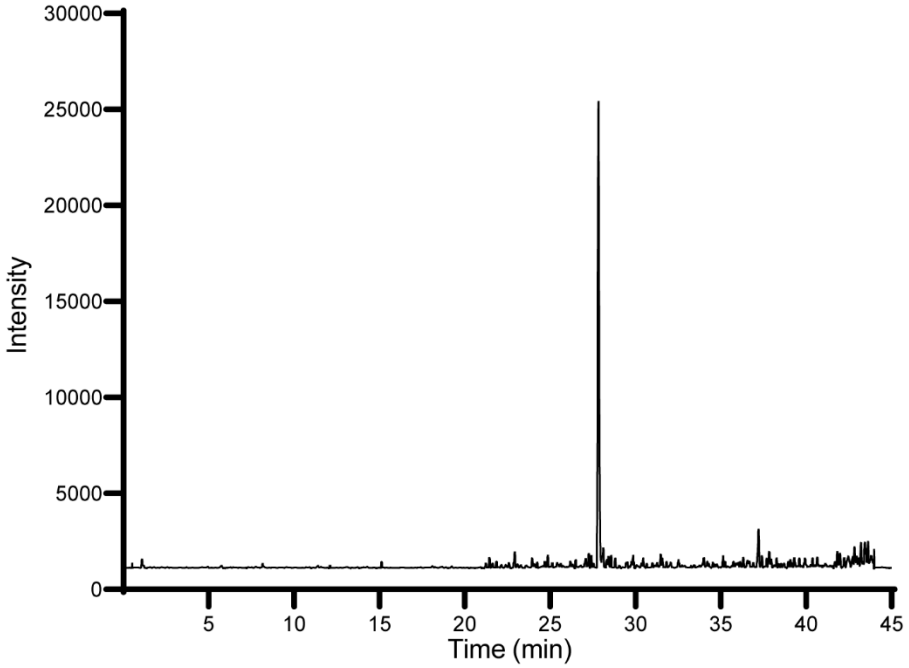

Figure S8. Extracted ion chromatogram of arcumycin treated with carboxypeptidase Y (*m*/*z* = 1052 $[M+2H]^{2+}$).

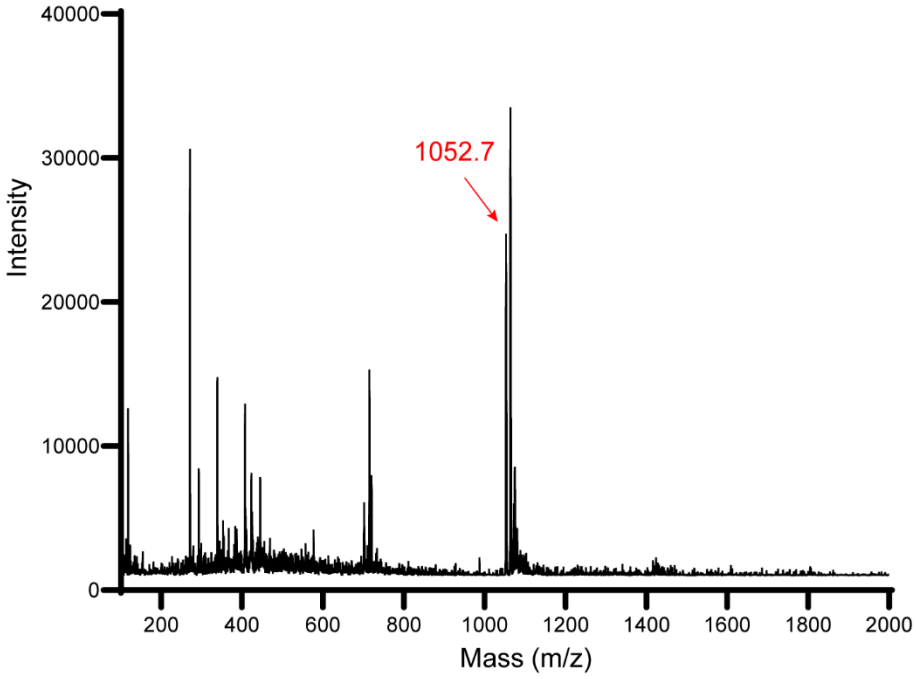

Figure S9. ESI-MS of arcumycin treated with carboxypeptidase Y at rt = 27.8 min (calc. $[M+2H]^{2+}$ =1052.4490).

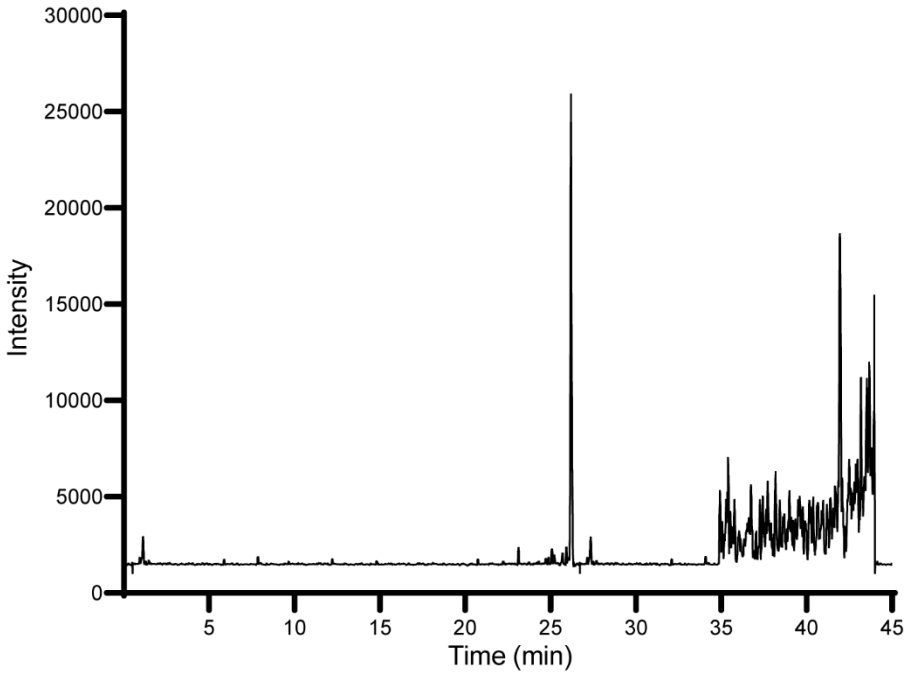

Figure S10. Extracted ion chromatogram of reduced and alkylated arcumycin (*m/z* = 1168 $[M+2H]^{2+}$).

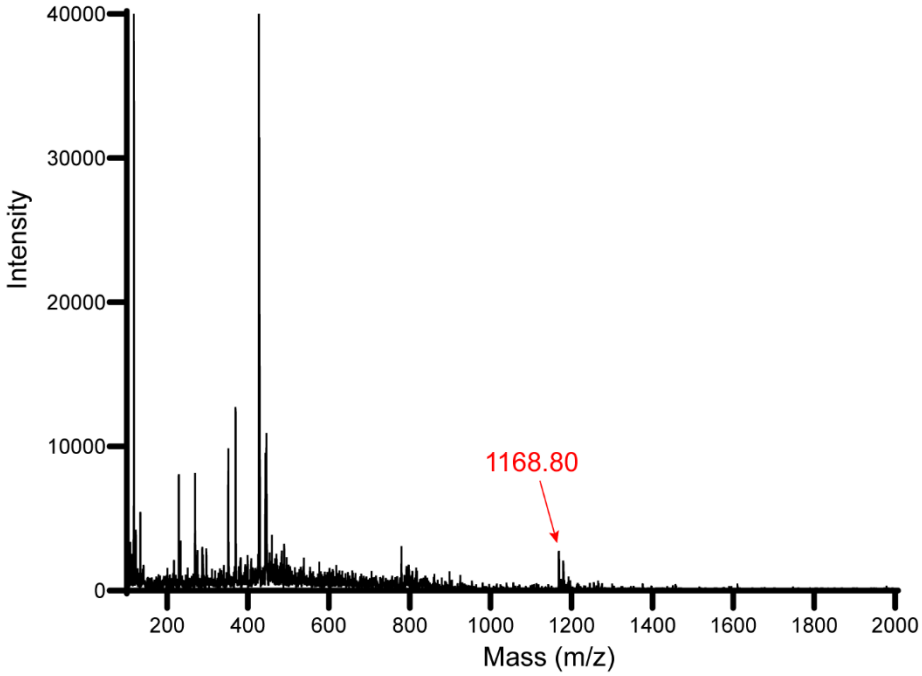

Figure S11. ESI-MS of reduced and alkylated arcumycin at rt = 26.2 min (calc. $[M+2H]^{2+}$ =1168.5077)

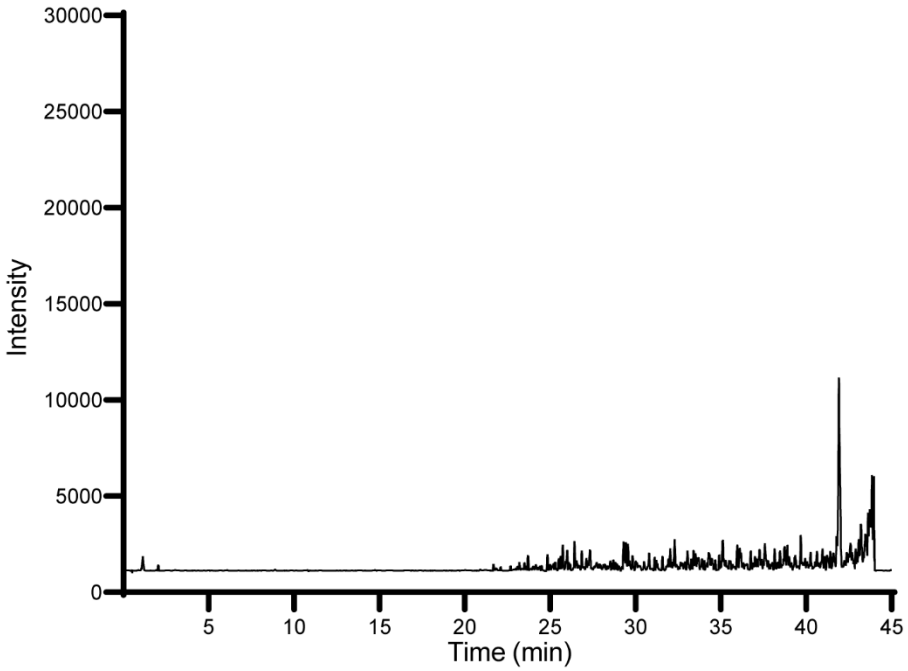

Figure S12. Extracted ion chromatogram of reduced and alkylated arcumycin treated with carboxypeptidase Y (*m*/*z* = 1168 $[M+2H]^{2+}$). No significant signal was detected for reduced and alkylated arcumycin indicating full degradation by protease treatment.

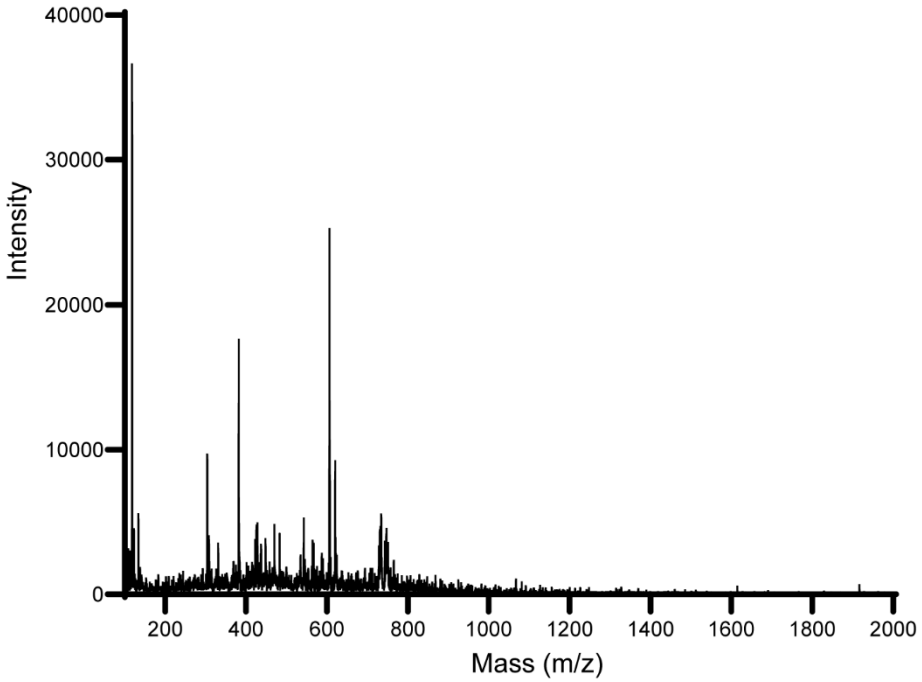

Figure S13. ESI-MS of reduced and alkylated arcumycin treated with carboxypeptidase Y (calc. $[M+2H]^{2+}$ = 1168.5077). No mass signal was detected for reduced and alkylated arcumycin for the doubly charged state.

III. References

[1] M. Tsunakwa, S.-L. Hu, Y. Hoshino, D. J. Detlefson, S. E. Hill, T. Furumai, R. J. White, K. Kawano, S. Yamamoto, Y. Fukagawa, T. Oki, *J. Antibiot. (Tokyo).* **1995**, *48*, 433–434.

[2] M. Shao, J. Ma, Q. Li, J. Ju, *Mar. Drugs* **2019**, *17*, 127.

[3] I. Kaweewan, H. Hemmi, H. Komaki, S. Harada, S. Kodani, *Bioorganic Med. Chem.* **2018**, *26*, 6050–6055.

[4] Y. Li, R. Ducasse, S. Zirah, A. Blond, C. Goulard, E. Lescop, C. Giraud, A. Hartke, E. Guittet, J. L. Pernodet, S. Rebuffat, *ACS Chem. Biol.* **2015**, *10*, 2641–2649.

[5] J. I. Tietz, C. J. Schwalen, P. S. Patel, T. Maxson, P. M. Blair, H.-C. Tai, U. I. Zakai, D. A. Mitchell, *Nat. Chem. Biol.* **2017**, *13*, 470–478.

[6] M. Iwatsuki, H. Tomoda, R. Uchida, H. Gouda, S. Hirono, S. Omura, *J. Am. Chem. Soc.* **2006**, *128*, 7486–7491.

[7] E. Gavrish, C. S. Sit, S. Cao, O. Kandror, A. Spoering, A. Peoples, L. Ling, A. Fetterman, D. Hughes, A. Bissell, H. Torrey, T. Akopian, A. Mueller, S. Epstein, A. Goldberg, J. Clardy, K. Lewis, *Chem. Biol.* **2014**, *21*, 509–518.

[8] Y. Esumi, Y. Suzuki, Y. Itoh, M. Uramoto, K. I. Kimura, M. Goto, M. Yoshihama, T. Ichikawa, *J. Antibiot. (Tokyo).* **2002**, *55*, 296–300.

[9] M. Metelev, J. I. Tietz, J. O. Melby, P. M. Blair, L. Zhu, I. Livnat, K. Severinov, D. A. Mitchell, *Chem. Biol.* **2015**, *22*, 241–250.

[10] I. Kaweewan, M. Ohnishi-Kameyama, S. Kodani, *J. Antibiot. (Tokyo).* **2017**, *70*, 208–211.

[11] S. S. Elsayed, F. Trusch, H. Deng, A. Raab, I. Prokes, K. Busarakam, J. A. Asenjo, B. A. Andrews, P. van West, A. T. Bull, M. Goodfellow, Y. Yi, R. Ebel, M. Jaspars, M. E. Rateb, *J. Org. Chem.* **2015**, *80*, 10252–10260.

[12] N. Takasaka, I. Kaweewan, M. Ohnishi-Kameyama, S. Kodani, *Lett. Appl. Microbiol.* **2017**, *64*, 150–157.

[13] S. Kodani, Y. Inoue, M. Suzuki, H. Dohra, T. Suzuki, H. Hemmi, M. Ohnishi-Kameyama, *European J. Org. Chem.* **2017**, *2017*, 1177–1183.