# Supplementary Information

# High-Yield Lasso Peptide Production in a *Burkholderia* Bacterial Host by Plasmid Copy Number Engineering

Hannah N. Fernandez,[1] Ashley Kretsch,[2] Sylvia Kunakom,[1] Adjo E. Kadjo,[1] Douglas A. Mitchell,[2] Alessandra S. Eustáquio[1*]

[1]Department of Pharmaceutical Sciences and Center for Biomolecular Sciences, College of Pharmacy, University of Illinois at Chicago, Chicago, IL 60607, USA

[2]Department of Chemistry and Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

# Table of Contents

**Table S1. Previously reported heterologous production of lasso peptides from Pseudomonadota.**

| Source of BGC (Genus) | Class and Family of BGC Source | G+C (%) | Host Used | Class and Family of Host | G+C (%) | Inducer | Codon Optimized | Yield (mg/L) | Compound | Ref |
|---|---|---|---|---|---|---|---|---|---|---|
| *Asticcacaulis* | Alphaproteobacteria Caulobacteraceae | 60 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | anhydrotetracycline | no | 0.26 | astexin-1 | 1 |
| *Asticcaucalis* | Alphaproteobacteria Caulobacteraceae | 60 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | ? | ? | 0.28 | benenodin I and II | 2,3 |
| *Brevundimonas* | Alphaproteobacteria Caulobacteraceae | 63 | *S. subterranea* | Alphaproteobacteria Sphingomonadaceae | 50 | none, constitutive *Sphingomonas* promoter | no | 10.2 | brevunsin | 4 |
| *Burkholderia / Mycetohabitans* | Betaproteobacteria Burkholderiaceae | 61 | *B. gladioli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | trace | burhizin-23 | 5 |
| *Burkholderia* | Betaproteobacteria Burkholderiaceae | 68 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | anhydrotetracycline note: stem loop between *capA* and *capB* removed | no | 1.6 | capistruin | 6 |
| *Burkholderia* | Betaproteobacteria Burkholderiaceae | 68 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 0.2 | capistruin | 7 |
| *Burkholderia* | Betaproteobacteria Burkholderiaceae | 68 | *Burkholderia* Sp. FERM BP-3421 | Betaproteobacteria Burkholderiaceae | 68 | L-arabinose | no | Non-outliers: 3 Outlier over-producer: 116 | capistruin | 8 |
| *Caulobacter* | Alphaproteobacteria Caulobacteraceae | 68 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 3.4 | caulonodin I-III | 5 |
| *Caulobacter* | Alphaproteobacteria Caulobacteraceae | 68 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 0.3 | caulosegnin I-III | 9 |
| *Citrobacter* | Gammaproteobacteria Enterobacteriaceae | 52 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG for *citA*, and constitutive promoter of *mcjBCD* for *citBCD* | yes | 2.7 | citrocin | 10 |
| *Klebsiella* | Gammaproteobacteria Enterobacteriaceae | 57 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | L-arabinose | no | 5 | klebsidin | 11 |
| *Sphingomonas* | Alphaproteobacteria Sphingomonadaceae | 65 | *S. subterranea* | Alphaproteobacteria Sphingomonadaceae | 63 | none, constitutive *Sphingomonas* promoter | no | 2.8 | koreensin | 12 |
| *Escherichia* | Gammaproteobacteria Enterobacteriaceae | 50 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | ? | yes | 12 | microcin J25 | 13 |
| *Salmonella* | Gamaproteobacteria Enterobacteriaceae | 52 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 9.61 | microcin Y | 14 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Mycetohabitans* | Betaproteobacteria Burkholderiaceae | 61 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | L-arabinose | no | 1 | mycetohabin-15 | 15 |
| *Panderoaea* | Betaproteobacteria Burkholderiaceae | 63 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 2 | pandonodin | 16 |
| *Rhodanobacter* | Gammaproteobacteria Rhodanobacteraceae | 65 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | trace | rhodanodin | 5 |
| *Rubrivivax* | Betaproteobacteria Burkholderiaceae | 70 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 2.5 | rubrinodin | 17 |
| *Rubrivivax* | Betaproteobacteria Burkholderiales order | 71 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 0.5 | rubrivinodin | 5 |
| *Sphingobium* | Alphaproteobacteria Sphingomonadaceae | 64 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 0.9 | sphingonodin I and II | 5 |
| *Sphingopyxis* | Alphaproteobacteria Sphingomonadaceae | 65 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 3.4 | sphingopyxin I and II | 5 |
| *Sphingobium* | Alphaproteobacteria Sphingomonadaceae | 64 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 5.2 | syanodin I | 5 |
| *Burkholderia* | Betaproteobacteria Burkholderiaceae | 68 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | yes | 1.8 | ubonodin | 18 |
| *Xanthomonas* | Gammaproteobacteria Xanthomonadaceae | 65 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | 12.7 | xanthomonin I-III | 19 |
| *Phenylobacterium* | Alphaproteobacteria Caulobacteraceae | 70 | *E. coli* | Gammaproteobacteria Enterobacteriaceae | 50 | IPTG | no | trace | zucinodin | 5 |

Table updated from.[8]
Yields isolated from liquid cultures.
IPTG, isopropyl β-D-1-thiogalactopyranoside
Blue rows, isolated yields greater than 10 mg/L.
Trace, only trace amounts of compound detected.

**Table S2**. **Identified differences between the reference genome and either pSK020 outlier overproducer or pSK020+48 low producer.** PATRIC was used for analysis.

| Reference Genome | Variation found | Type of mutation | Amino Acid shift | Frame-shift? | Function |
|---|---|---|---|---|---|
| Outlier Overproducer | | | | | |
| CGC | CGCGCGGC | insertion | Arg2_Gly3fs | yes | Protein ImpG/VasA |
| GCCGA | GCCCGA | insertion | Ser500_Ala501fs | yes | T6SS component TssC (ImpC/VipB) |
| CAAAAG | CCACTGCCCGA AGGCCACGCGG GACAAAAG | insertion | – | – | not in coding region |
| CGGT | GGGC | substitution (synonymous) | SerGly152SerGly | no | ABC transporter, substrate-binding protein |
| CCAC | CCACAC | insertion | Asp597_Val598fs | yes | hypothetical protein |
| ACCCTG | ACCCCTG | insertion | Arg835_Val836fs | yes | hypothetical protein |
| GCCCCCACG | GCCCCCCACG | insertion | Trp832_Gly833fs | yes | hypothetical protein |
| TCCCCGCCG CA | TCCCCCGCCGC A | insertion | Ala827_Gly828fs | yes | hypothetical protein |
| CCTC | CCTCTC | insertion | Val808_Glu809fs | yes | hypothetical protein |
| GT | GCTC | insertion | Thr801fs | yes | hypothetical protein |
| CCG | CCAGCT | insertion | Ala691delinsGluLeu | no | hypothetical protein |
| CGGGGCTTC G | CGGGGGGGGGC TTCG | insertion | – | – | not in coding region |
| CGGGGGC | CGGGGGGC | insertion | – | – | not in coding region |
| ACA | ACCA | insertion | – | – | not in coding region |
| G | A | substitution (synonymous) | Leu150Leu | no | Quinolinate phosphoribosyltrans-ferase [decarboxylating] (EC 2.4.2.19) |
| ATGCGA | ATCAAATGTGT AGGTATAAGGC GATGCGA | insertion | Met1_Arg2fs | yes | Transcriptional regulator, HxlR family |
| GGTCG | CGTT | deletion | – | – | not in coding region |
| CGACCGCGC GCGCGA | CGAGCTCGGTC GG | deletion | – | – | not in coding region |
| CCCGACGAA GCCTAGGCC GCCCGCGGC CC | CCCGACGAAGC CTAGGCCGCCC GCGCGGTATCG AGCGGGCGGCC C | insertion | – | – | not in coding region |
| CCCCG | CCCGCTCCCG | insertion | Gly69_Gly70fs | yes | hypothetical protein |
| TCCCCG | TCCCCCGC | insertion | – | – | not in coding region |
| ACCCG | ACCCCG | insertion | – | – | not in coding region |
| CGCCCGA | CCCGCCCCGA | insertion | Arg404delinsAlaGly | | CaiB/BaiF family protein |

| | | | | | |
|---|---|---|---|---|---|
| GGC | GGAGC | insertion | – | – | LSU rRNA 23S rRNA, large subunit ribosomal RNA |
| CAACC | CTTAACC | insertion | – | – | LSU rRNA 23S rRNA, large subunit ribosomal RNA |
| CCT | CCC | substitution (non-synonymous) | – | – | not in coding region |
| ACCGA | ACCCGA | insertion | Ser156_Val157fs | yes | hypothetical protein |
| CCGACCGA | CCCGA | deletion | Arg148_Ser149del | | hypothetical protein |
| A | T | substitution | Gln270Leu | no | Replicative DNA helicase (DnaB) (EC 3.6.4.12) |
| Nonoutlier Low Producer | | | | | |
| CCGGGTGA | CAATGTGATGCGCCGGGTGA | insertion | – | – | not in coding region |
| GGC | GGCAGC | insertion | Gln102_Pro103insGlnLeu | no | hypothetical protein |
| GGCGA | GAGCTGA | insertion | Ala14fs | yes | T6SS component TssK (ImpJ/VasE) |
| A | G | substitution (non-synonymous) | Ser16Gly | no | T6SS component TssK (ImpJ/VasE) |
| ACGCCTG | ACGCCCCCTG | insertion | Arg304_Arg305insArgGly | no | T6SS component TssM (IcmF/VasK) |
| TACG | TACACG | insertion | not indicated | not indicated | SSU rRNA 16S rRNA, small subunit ribosomal RNA |
| TCG | GTA | substitution (non-synonymous) | AlaArg60ValArg | no | T6SS component TssM (IcmF/VasK) |
| CCGCCG | CCTGCGCGCGCG | insertion | Gly292delinsAlaArgAla | no | T6SS outer membrane component TssL (ImpK/VasF) / OmpA/MotB domain |
| G | C | substitution (synonymous) | Ala2228Ala | no | hypothetical protein |
| T | C | substitution (synonymous) | Thr237Thr | no | VgrG protein |
| T | C | substitution (non-synonymous) | Leu252Pro | no | VgrG protein |
| A | G | substitution (synonymous) | Lys256Lys | no | VgrG protein |
| T | C | substitution (non-synonymous) | Val311Ala | no | VgrG protein |
| T | C | substitution (synonymous) | Tyr14Tyr | no | VgrG protein |
| AAGCCTAGGCCGCCCGCGGCCC | AAGCCTAGGCCGCCCGCGCGGTATCGAGCGGGCGGCCC | insertion | – | – | hypothetical protein |

| | | | | | |
|---|---|---|---|---|---|
| G | A | substitution (synonymous) | Val320Val | no | NADH-ubiquinone oxidoreductase chain N (EC 1.6.5.3) |
| CCCCGGA | TCCCGGA | substitution (non-synonymous) | Gly70Ser | no | hypothetical protein |
| CGCCCCAG | CGCCCCCAG | insertion | – | – | not in coding region |
| GCACCCGCA | GCACCCCGCA | insertion | – | – | not in coding region |
| GCCCCGCCC | GCCCCCGCCC | insertion | – | – | not in coding region |
| AGGGA | AGGGGA | insertion | not indicated | not indicated | LSU rRNA 23S rRNA, large subunit ribosomal RNA |
| A | C | substitution (synonymous) | Arg94Arg | no | hypothetical protein |
| CC | AT | substitution (non-synonymous) | GlnAla15582GlnSer | no | Polyketide synthase modules and related proteins |

**Table S3. Primer sequence list for plasmid construction requiring PCR.**

| Plasmid constructed | Forward primer | Reverse primer |
|---|---|---|
| **pKAE001** | P_CapBGC_*Bgl*II_f<br>GCC TAA AGA TCT CCT GCA TCG ATT TAT TAT GAC<br>*Bgl*II site underlined | P_CapBGC_*Nde*I_r<br>GAC TAA CAT ATG CAC ACA GGA AAC AGC TAT G<br>*Nde*I site underlined |
| **pHNF008** | PBAD_*Nhe*I_f<br>CGT AAC GCT AGC CCT GCA TCG ATT TAT TAT GAC<br>*NheI* site underlined | PBAD_*Bgl*II_r<br>GAT AGA TCT ACT CGA GGC TAT GAC ATG ATT ACG AAT TC<br>*BglII* site underlined |
| **pHNF009** | P_pBBR1_G159S_f<br>ACT ACC GAC CAG CCC CGG CGA | P_pBBR1_G159S_r<br>TGC TGC TCG CCC GGA TAC |
| **pAK343** | P_mlsA1A2B_OLEwithpUC_f<br>TAG AGG ATC CCC GGG TAC CGA GCT *atcaaggagattttc* ATG AAT AAG CAG CAA GAC GTG<br>Overlap region underlined, RBS site in italics | P_mlsA1A2B_OLEwithmlsC_r<br>ATA GGC GAT GCC GAG CAG GGT CAT TCA CAG GGG CAG TGC C<br>Overlap region underlined |
| | P_mlsC_OLEwithmlsAB__f<br>CCG CTG CTG GCA CTG CCC CTG TGA ATG ACC CTG CTC GGC ATC<br>Overlap region underlined | P_mlsC_OLEwithpUC_r<br>GCT ATG ACC ATG ATT ACG AAT TCG TCA TTG CCA TCC TTG AAC AAA CAG<br>Overlap region underlined |
| **pAK343 Sanger sequencing** | CGT CAC ACT TTG CTA TGC CA | N/A |
| **pHNF007** | P_B13BGC_*Bgl*II_ f<br>GAC TAA AGA TCT CCT GCA TCG ATT TAT TAT GAC<br>*Bgl*II site underlined | P_B13BGC_*Hind*III_ r<br>GAC TAG AAG CTT GCT ATG ACA TGA TTA CGA ATT C<br>*Hind*III site underlined |
| **pHNF012** | P_pBBR1_G159S_f<br>ACT ACC GAC CAG CCC CGG CGA | P_pBBR1_G159S_r<br>TGC TGC TCG CCC GGA TAC |
| ***araC* qPCR** | P_araC_qPCR_f<br>TAA CCT TTC ATT CCC AGC GGT C | P_araC_qPCR_r<br>CTG CCG GGA TAC TCG TTT AAT G |

| | | |
|---|---|---|
| ***kan*$^{R}$ qPCR** | P2_kanqPCR_f<br>CTC TGG TAA GGT TGG GAA GC | P2_kanqPCR_r<br>GTG CAA TCC ATC TTG TTC AAT CAT G |
| ***recA* qPCR** | P_recA_qPCR_f<br>ACA CCA CCT GGA TGT CCT CG | P_recA_qPCR_r<br>ATG GAA GAT AGC AAG AAG GGT TC |
| ***gyrA* qPCR** | P_gyrA_qPCR_f<br>ACA CCG AAG ACC TGA TCA CG | P_gyrA_qPCR_r<br>TCC AGT CGT CTT CCT TCA TCT G |

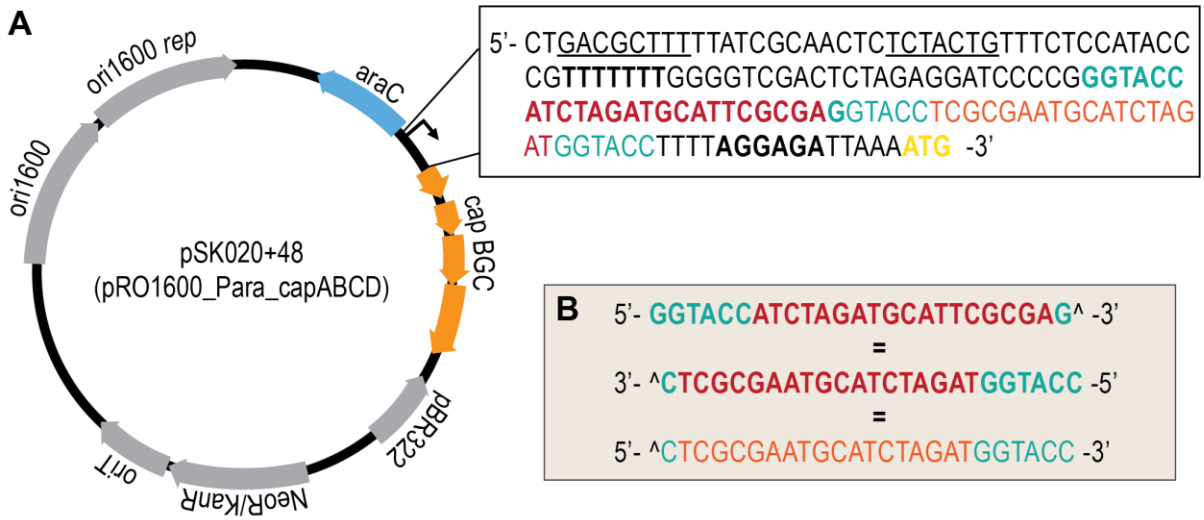



**Figure S1**. **Schematic of pSK020+48 plasmid and sequence of the promoter region. (A)** The capistruin BGC (*capABCD*, orange) under the control of the *araC* regulator (blue) and $P_{BAD}$ promoter in the pR01600-based vector. The sequence of the $P_{BAD}$ promoter is denoted by underlined -35 and -10 regions, respectively. The 48-bp insertion (dark and light red) is located upstream of the start codon of *capA* (yellow) and contains three *Kpn*I sites (teal). The thymine-rich region that is part of the translation enhancer and the RBS site upstream of *capA* are bolded. **(B)** We speculate the 48-bp insertion originated from the pUC57-Kan vector used to clone the synthetic *cap* BGC, which was used to generate pSK020. The 48-bp insertion (dark and light red) is composed of inverted repeats flanked by *Kpn*I sites (teal).

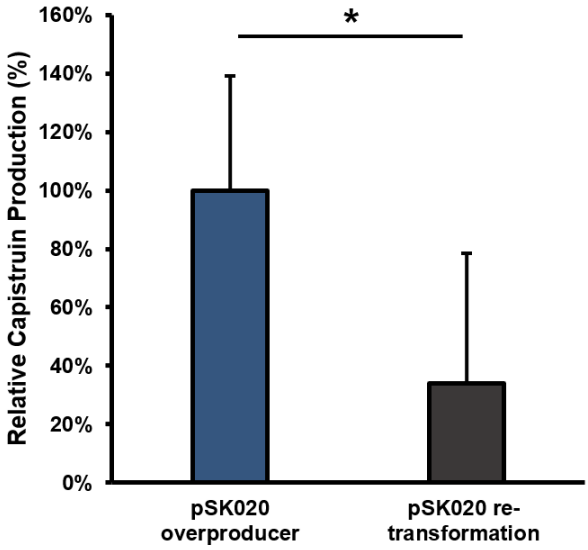



**Figure S2. Plasmid is not the only factor contributing to the overproducer phenotype.** Comparison of capistruin titers between pSK020 overproducer and clones that were obtained after retransformation of the plasmid isolated from the overproducer host into the *Burkholderia* sp. WT strain. Average of $N = 4 \pm$ standard deviation. For the overproducer clone, $N = 3$ indicates results from three production culture flasks inoculated from the same seed culture. For the cleared host, $N = 4$ indicates results from three independent clones. The plasmid from the pSK020 overproducer was reintroduced via electroporation and four independent clones were selected for analysis. T-test was implemented for statistical analysis. *, P = 0.09.

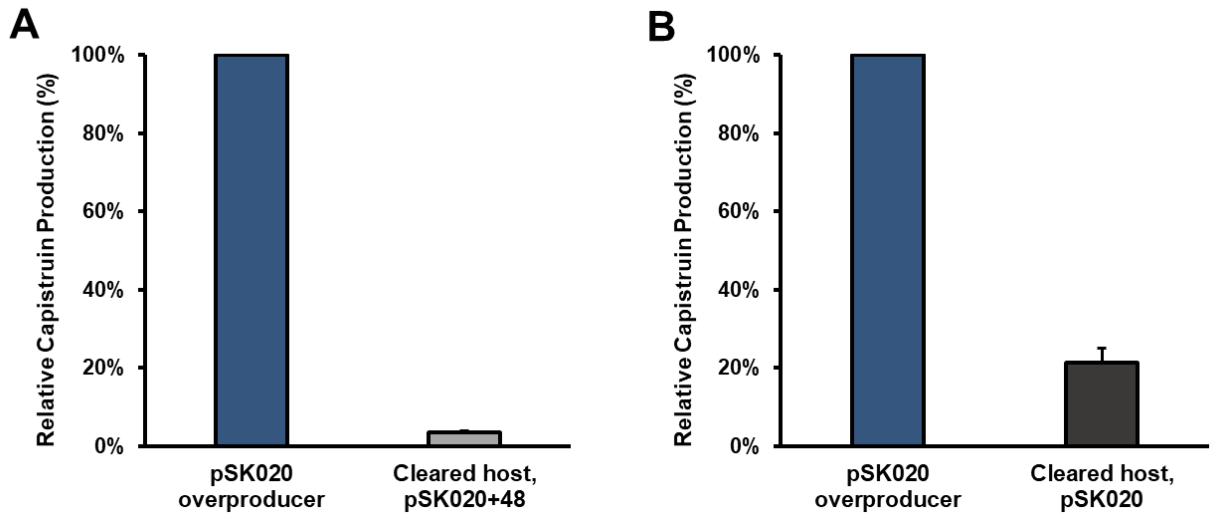



**Figure S3. Assessment of the overproducer clone's ability to recapitulate high yield capistruin production when used as a heterologous host.** Comparison of capistruin titers between pSK020 overproducer and clones that were obtained after clearing the pSK020 overproducer clone of plasmid and retransforming it. **(A)** After host clearing, pSK020+48 plasmid from *E. coli* was reintroduced into the host via electroporation. $N = 1$ for pSK020 overproducer and average of $N = 7 \pm$ standard deviation of 0.5% is shown for cleared host, pSK020+48. Biological replicates were utilized. **(B)** After plasmid clearing, the pSK020 plasmid from the overproducer was retransformed into the cleared host. $N = 1$ for pSK020 overproducer and average of $N = 2 \pm$ standard deviation of 3.5% is shown for cleared host, pSK020. Biological replicates were utilized.

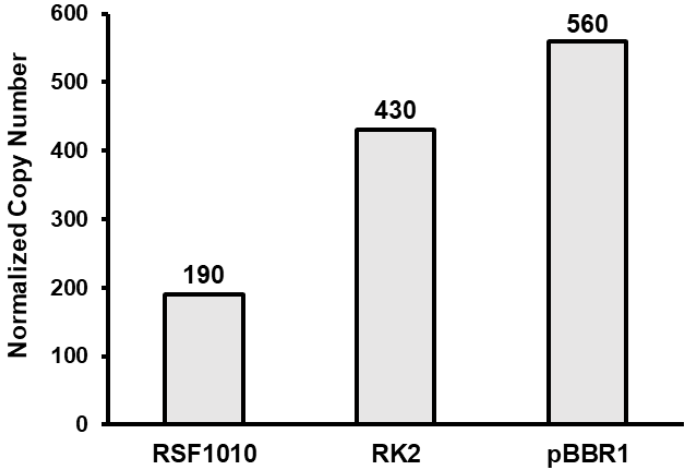



**Figure S4. Normalized copy number of different replicons in *Burkholderia* sp. FERM BP-3421.** Empty vectors containing the RSF1010 (pAM4891), RK2 (pSEVA227M), and pBBR1 (pMo168) replicons were transferred into *Burkholderia* sp. and genomic DNA was then isolated and sequenced using Illumina technology. Normalized plasmid copy number per genome equivalent is shown. All vectors contain a kanamycin resistance marker.

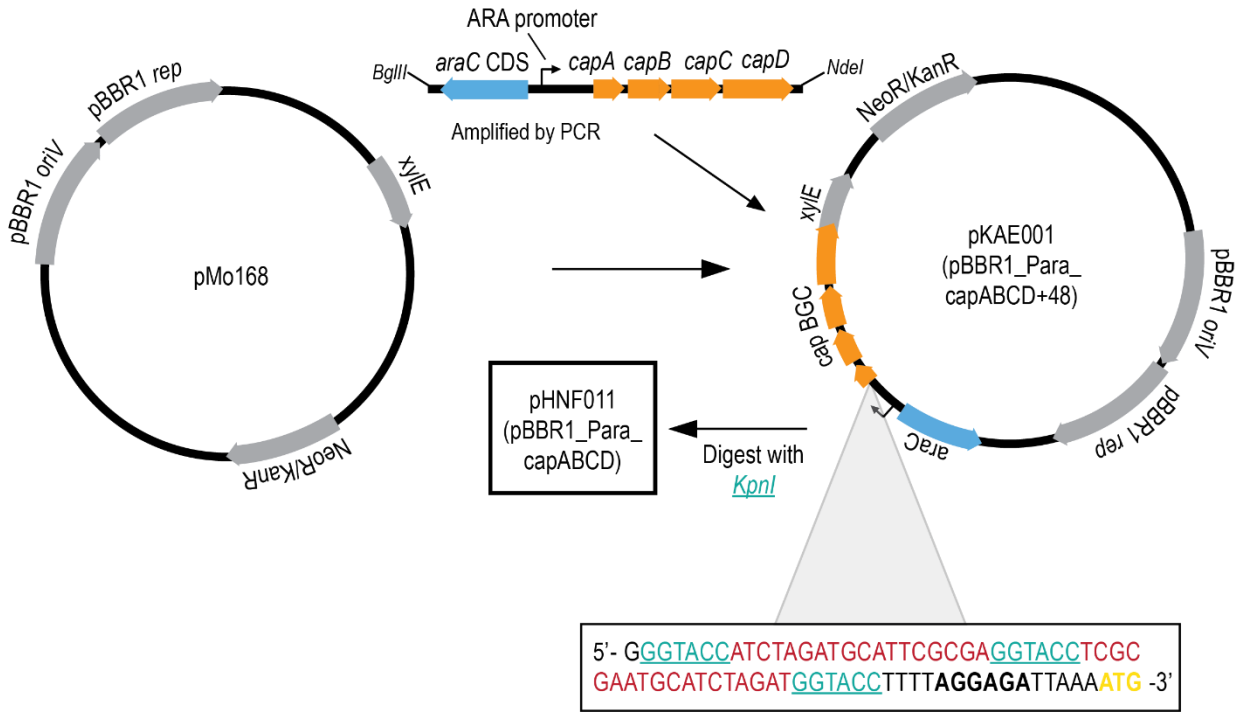



**Figure S5. Generation of pKAE001 and pHNF011.** The *cap* BGC (orange) under the control of the *araC*/$P_{BAD}$ promoter was amplified by PCR and subcloned into pMo168 (pBBR1-based vector) using the restriction sites *Bgl*II and *Nde*I to generate pKAE001. Cloning of the *cap* BGC disrupted the *xylE* reporter gene. The 48 bp insert (red) between the $P_{BAD}$ promoter and the RBS is present in pKAE001 because pSK020+48 was used as PCR template before we discovered the 48-bp insertion. The underlined sites in teal bordering and within the 48-bp insert denote *Kpn*I restriction sites. Isolated pKAE001 plasmid was digested with *Kpn*I and ligated to generate pHNF011. In the boxed sequence, the RBS site is bolded, and the start of transcription is shown in yellow.

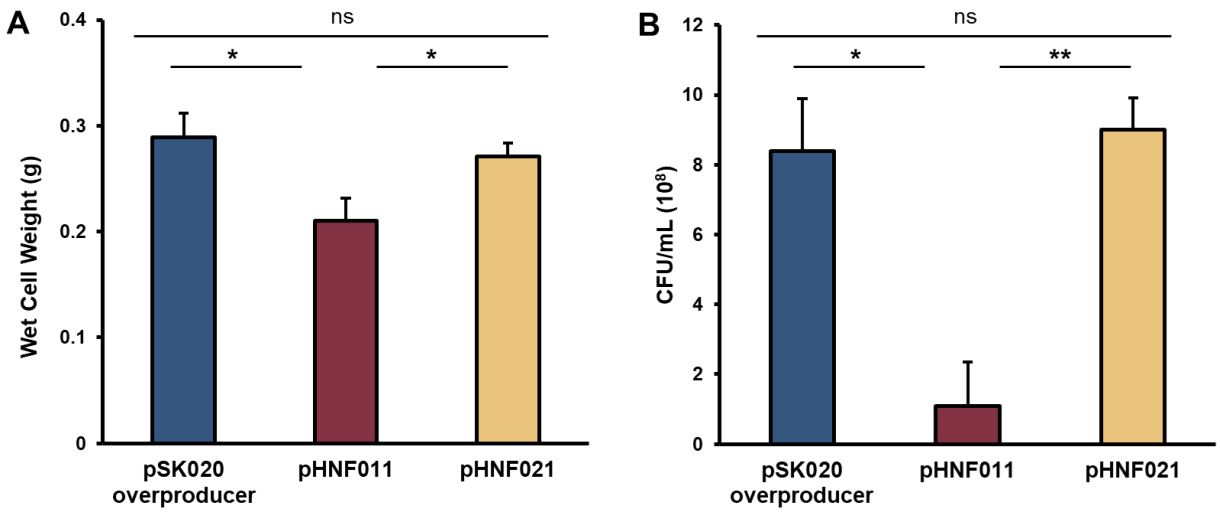



**Figure S6. Comparison of cell density between production cultures of *Burkholderia* sp. FERM BP-3421 harboring alternate capistruin expression vectors. (A)** The wet cell weight of production cultures of FERM BP-3421 harboring the pSK020, pHNF011, or pHNF021 vectors. Wet cell weights were acquired from aliquots (10 mL) of 48-h production cultures, which were centrifuged at maximum speed to separate the supernatant and cell mass. The average of $N = 3 \pm$ standard deviation is shown. **(B)** The number of colony forming units per milliliter (CFU/mL) from production cultures of FERM BP-3421 harboring the pSK020, pHNF011, or pHNF021 vectors. CFU/mL counts were acquired from serially diluting resuspended cell pellet that was harvested from 10 mL aliquots of 48-h production cultures. The average of $N = 3 \pm$ standard deviation is shown. T-test was implemented for statistical analysis. *, $P \leqslant 0.02$; **, $P \leq 1.0E^{-4}$; ns, no significant difference.

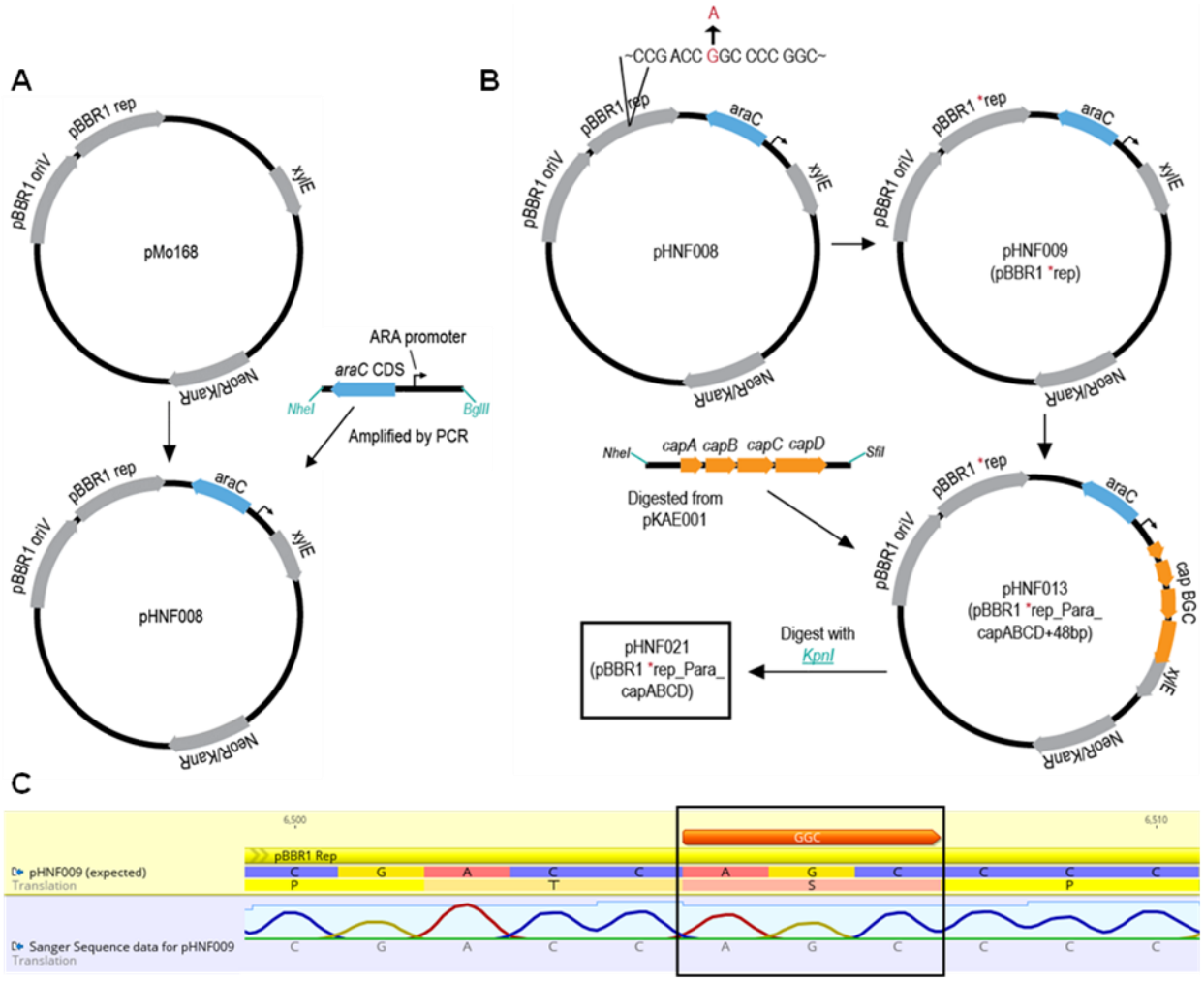



**Figure S7. Generation of pHNF021. (A)** The *araC*/$P_{BAD}$ L-arabinose inducible promoter was cloned into the pMo168 vector backbone to generate pHNF008. **(B)** Site-directed mutagenesis was used to introduce the G159S point mutation in the pBBR1 *rep* gene and generate pHNF009 (indicated as pBBR1 **rep*). pHNF013 was constructed by subcloning the capistruin BGC into pHNF009 using the restrictions sites *Nhe*I and *Sfi*I. pHNF013 contained the 48 bp insert carried over from the template plasmid before we found out about the 48-bp insertion. pHNF021 was then created by further digesting pHNF013 with *Kpn*I (teal) and ligating the digested plasmid. Cloning of the *cap* BGC disrupted the *xylE* reporter gene. All constructs were verified by restriction digest and either Sanger or whole plasmid sequencing. **(C)** Sanger sequence data indicating the introduction of the point mutation in the *rep* gene. The orange arrow annotation indicates the original sequence of the *rep* codon (GGC) before introduction of the point mutation (AGC).

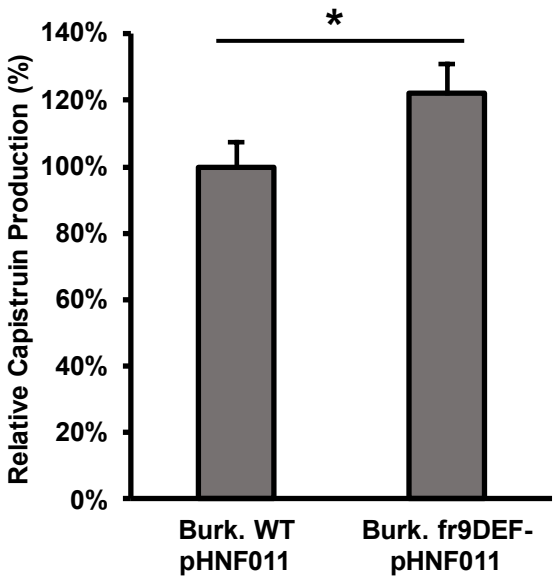



**Figure S8. Production of capistruin in wild-type and spliceostatin-defective mutant of *Burkholderia* sp. FERM BP-3421.** Relative capistruin titers when using the WT background versus the *fr9DEF*- background. The average of triplicates ± standard deviation is shown for both WT and mutant. T-test was implemented for statistical analysis. *, P = 0.03.

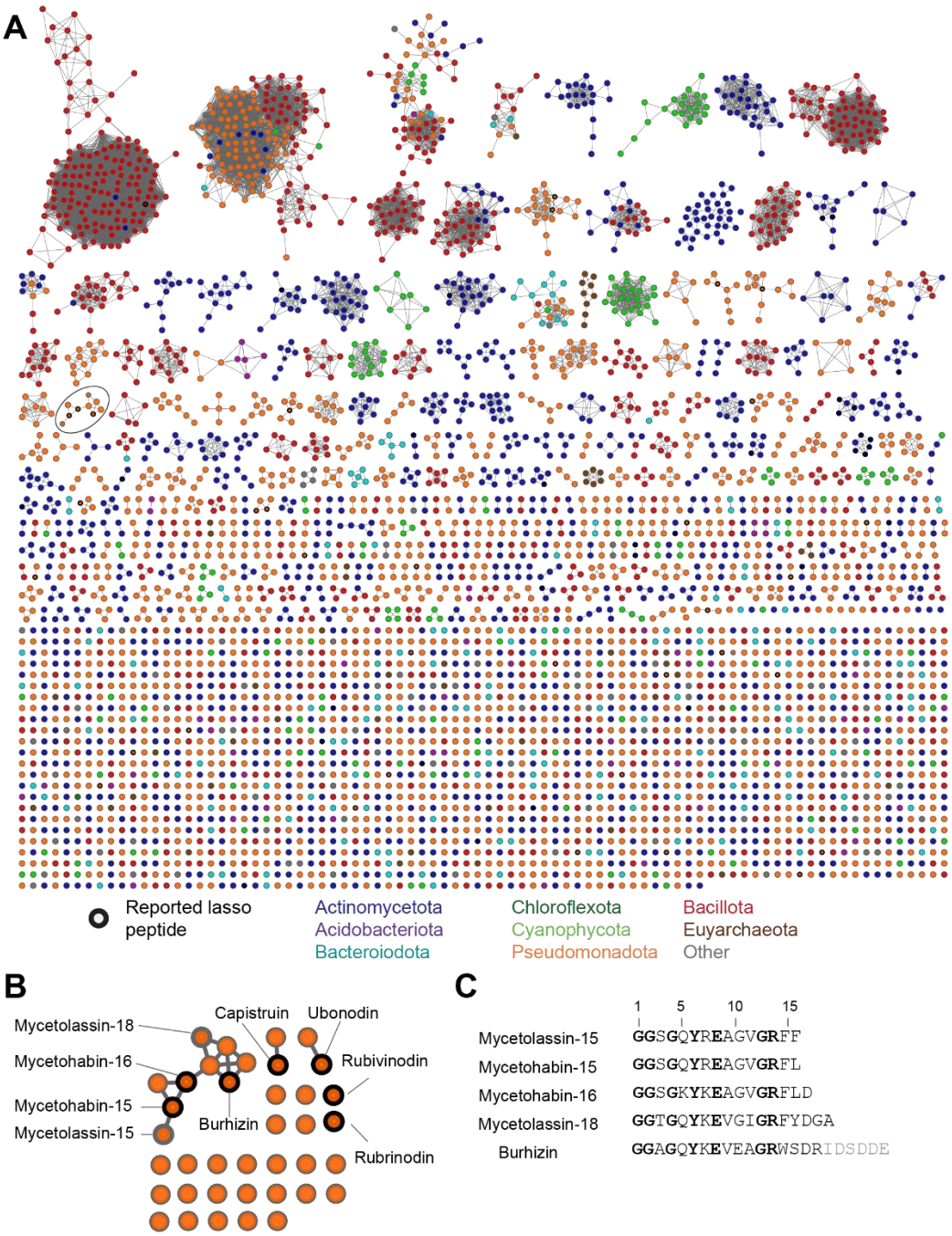
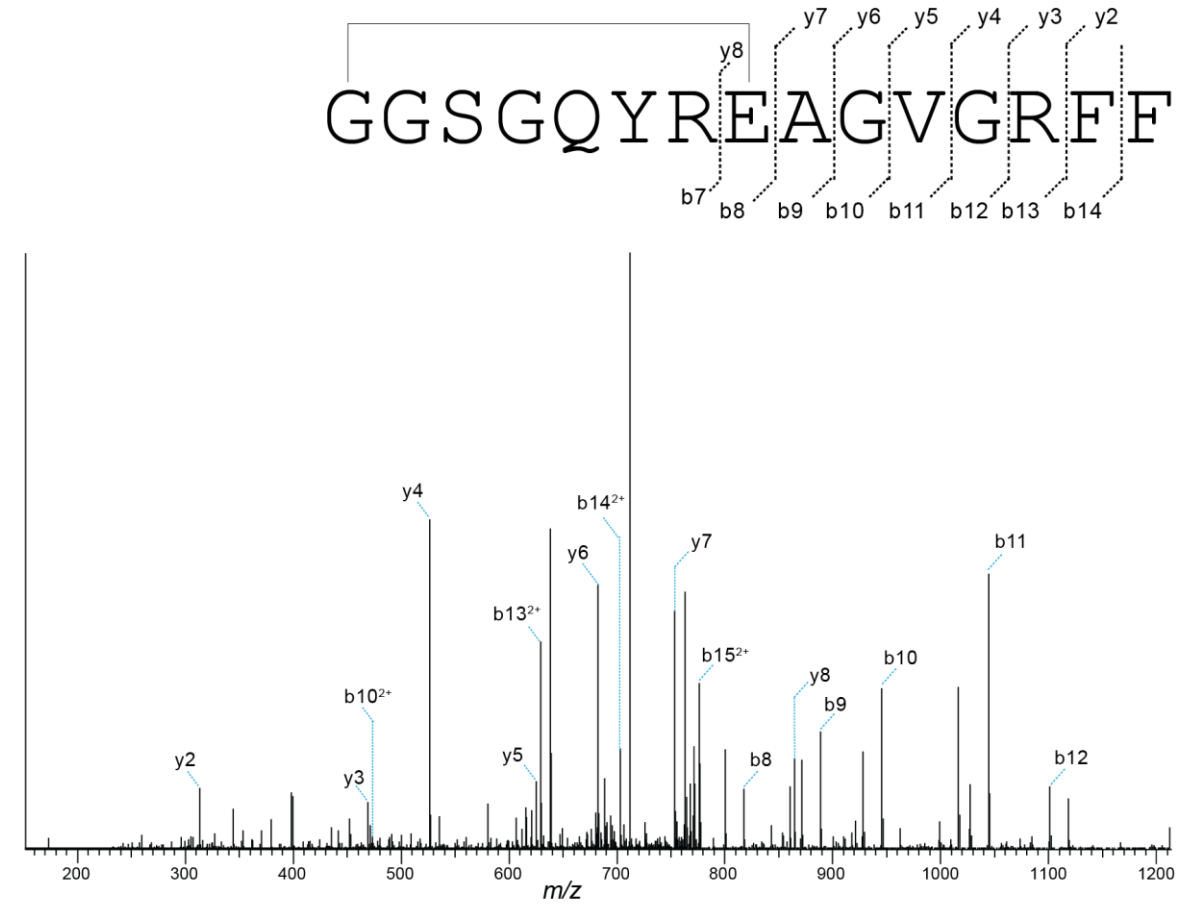
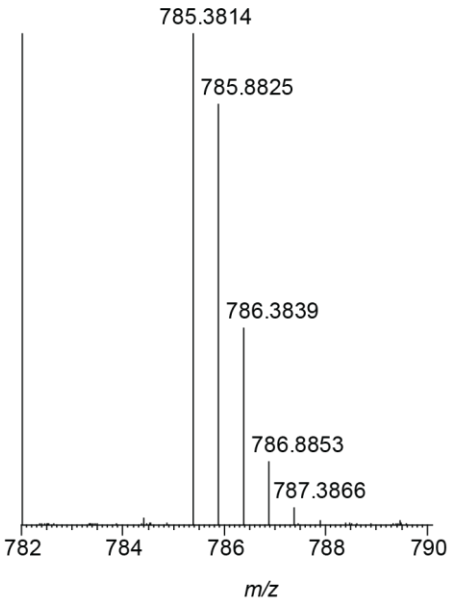


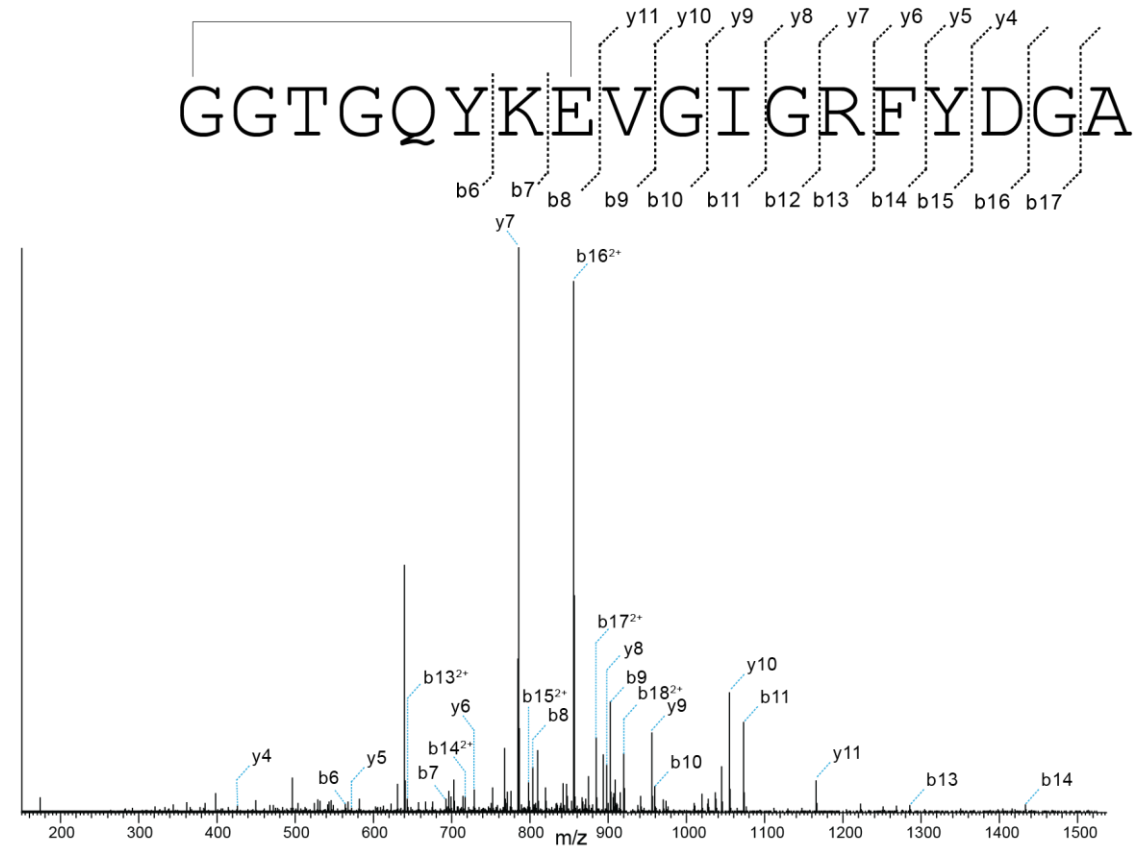

**Figure S9. Sequence similarity networks of predicted lasso core peptides.** Core peptides from a previous curation of predicted lasso peptides.[20] The non-identical set of curated sequences ($N$ = 4,485) was used to generate a sequence similarity network (SSN) using EFI-EST.[21] **(A)** SSN at alignment score 5 is displayed and colored by phylum of the encoding organism. Bolded nodes are core sequences from experimentally validated lasso peptides. The clade containing mycetolassin-15 and -18 is circled. **(B)** SSN at alignment score 5 of core peptides from *Burkholderiaceae* only. Literature reported lasso peptides are indicated. **(C)** A comparison of the sequences of several reported lasso peptides and mycetolassin-15 and -18. Conserved residues are in bold. Residues not observed in the final lasso product are grey.

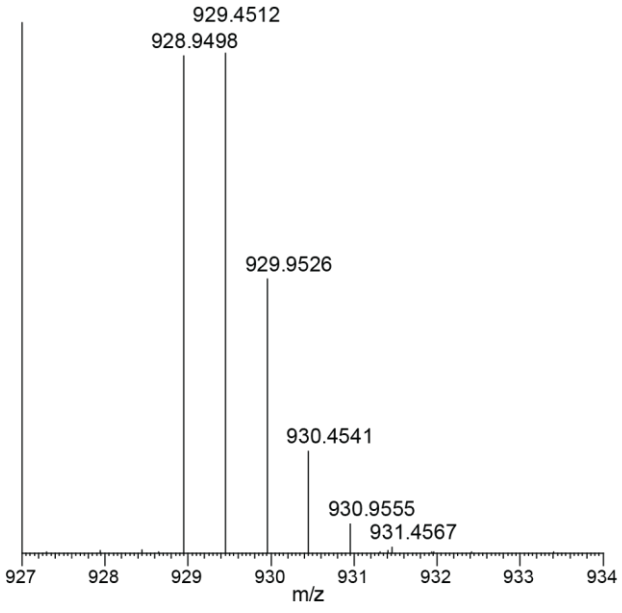



| Ion | Calc. mass | Obs. Mass | error (ppm) |
|---|---|---|---|
| $[M+2H]^{2+}$ | 785.3792 | 785.3814 | 0.52 |
| $b15^{2+}$ | 776.3765 | 776.3759 | 0.80 |
| $b14^{2+}$ | 702.8423 | 702.8415 | 1.17 |
| $b13^{2+}$ | 629.3081 | 629.3074 | 1.13 |
| $b12^{2+}$ | 551.2576 | 551.2570 | 1.00 |
| b12 | 1101.5078 | 1101.5061 | 1.53 |
| $b11^{2+}$ | 522.7468 | 522.7462 | 1.19 |
| b11 | 1044.4863 | 1044.4847 | 1.55 |
| $b10^{2+}$ | 473.2126 | 473.2120 | 1.31 |
| b10 | 945.4179 | 945.4165 | 1.49 |
| $b9^{2+}$ | 444.7019 | 444.7012 | 1.53 |
| b9 | 888.3965 | 888.3952 | 1.41 |
| $b8^{2+}$ | 409.1833 | 409.1827 | 1.54 |
| b8 | 817.3593 | 817.3584 | 1.14 |
| b7 | 706.3268 | 706.3262 | 0.79 |
| y2 | 313.1547 | 313.1545 | 0.67 |
| y3 | 469.2558 | 469.2554 | 0.90 |
| y4 | 526.2773 | 526.2769 | 0.74 |
| y5 | 625.3457 | 625.3452 | 0.80 |
| y6 | 682.3672 | 682.3665 | 0.97 |
| y7 | 753.4043 | 753.4038 | 0.64 |
| y8 | 864.4369 | 864.4357 | 1.34 |

**Figure S10. HR-ESI-MS/MS of purified mycetolassin-15**. Mycetolassin-15 was subjected to collision induced dissociation (30 keV), resulting in a series of $b^{+}$, $b^{2+}$, and $y^{+}$ ions consistent with a macrolactam linkage between Gly1 and Glu8 of the mycetolassin-15 core peptide. HR-MS of purified mycetolassin-15 is consistent with the predicted mass of the core peptide with a single cyclization.




| Ion | Calc. mass | Obs. Mass | error (ppm) |
|---|---|---|---|
| $[M+2H]^{2+}$ | 928.9503 | 928.9498 | 0.48 |
| $b18^{2+}$ | 919.9450 | 919.9436 | 1.49 |
| $b17^{2+}$ | 884.4264 | 884.4252 | 1.37 |
| $b16^{2+}$ | 855.9157 | 855.9142 | 1.73 |
| b16 | 1710.8240 | 1710.8222 | 1.07 |
| $b15^{2+}$ | 798.4022 | 798.4012 | 1.27 |
| b15 | 1595.7971 | 1595.7952 | 1.18 |
| $b14^{2+}$ | 716.8705 | 716.8694 | 1.59 |
| b14 | 1432.7338 | 1432.7322 | 1.09 |
| $b13^{2+}$ | 643.3363 | 643.3353 | 1.62 |
| b13 | 1285.6654 | 1285.6636 | 1.36 |
| $b12^{2+}$ | 565.2858 | 565.2848 | 1.73 |
| $b11^{2+}$ | 536.7751 | 536.7743 | 1.40 |
| b11 | 1072.5428 | 1072.5407 | 1.94 |
| $b10^{2+}$ | 480.2330 | 480.2322 | 1.71 |
| b10 | 959.4587 | 959.4569 | 1.89 |
| b9 | 902.4373 | 902.4357 | 1.72 |
| $b8^{2+}$ | 402.1881 | 402.1874 | 1.69 |
| b8 | 803.3688 | 803.3676 | 1.54 |
| b7 | 692.3363 | 692.3354 | 1.24 |
| b6 | 564.2413 | 564.2406 | 1.24 |
| y4 | 425.1667 | 425.1663 | 1.01 |
| y5 | 572.2352 | 572.2345 | 1.14 |
| y6 | 728.3363 | 728.3354 | 1.18 |
| y7 | 785.3577 | 785.3568 | 1.17 |
| y8 | 898.4418 | 898.4408 | 1.10 |
| y9 | 955.4633 | 955.4621 | 1.20 |
| y10 | 1054.5317 | 1054.5302 | 1.38 |
| y11 | 1165.5642 | 1165.5621 | 1.84 |

**Figure S11. HR-ESI-MS/MS of purified mycetolassin-18**. Mycetolassin-18 was subjected to collision induced dissociation (30 keV), resulting in a series of $b^{+}$, $b^{2+}$, and $y^{+}$ ions. The significantly lower intensity of the b7, b6, b5, and y12 ions is consistent with a macrolactam linkage between Gly1 and Glu8 of the mycetolassin-18 core peptide. HR-MS of purified mycetolassin-18 is consistent with the predicted mass of the core peptide with a single cyclization.

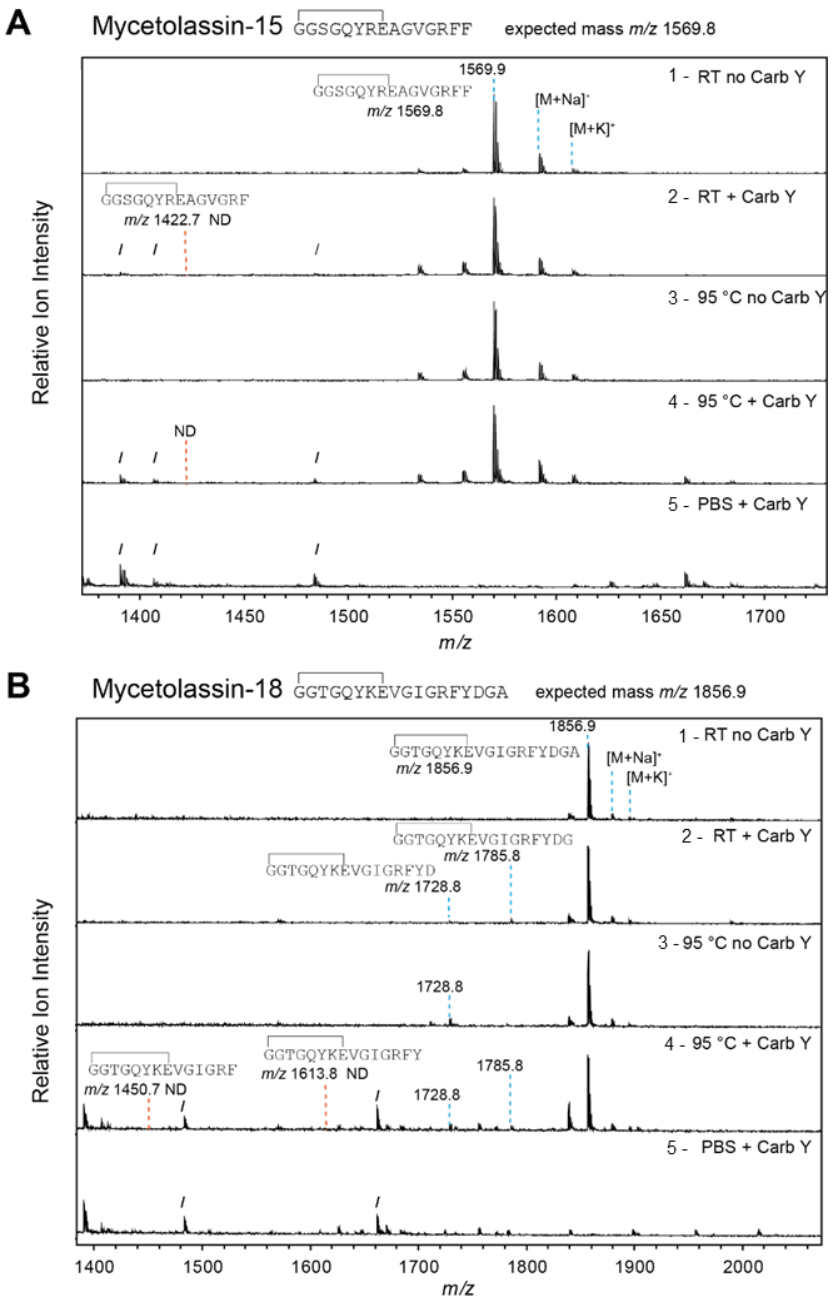



**Figure S12. Proteolytic resistance of purified mycetolassins.** Panel 1 and 2, purified **(A)** mycetolassin-15 and **(B)** mycetolassin-18 treated with either PBS control or carboxypeptidase Y and incubated for 18 h at room temperature. Panel 3 and 4, mycetolassins subjected to initial heat treatment for 3h at 95 ºC followed by carboxypeptidase treatment identical to above. **(A)** After either treatment of mycetolassin-15, removal of the C-terminal residue F15 (*m/z* 1422.7) was not detected (ND, orange). **(B)** Carboxypeptidase treatment of mycetolassin-18 with/without heat exposure yielded masses corresponding to the removal of the C-terminal Ala (*m/z* 1785.8) and Gly-Ala (*m/z* 1728.8), shown in panel 2 and 4 in blue. Heat treatment of mycetolassin-18 alone yielded a mass corresponding to the removal of the two C-terminal residues G17A18 (*m/z* 1728.8), indicated in panel 3 in blue. Removal of three or four residues was not observed (ND, orange). Masses indicated with / were identified in the PBS, carboxypeptidase Y sample and were not attributed to the mycetolassins, panel 5.

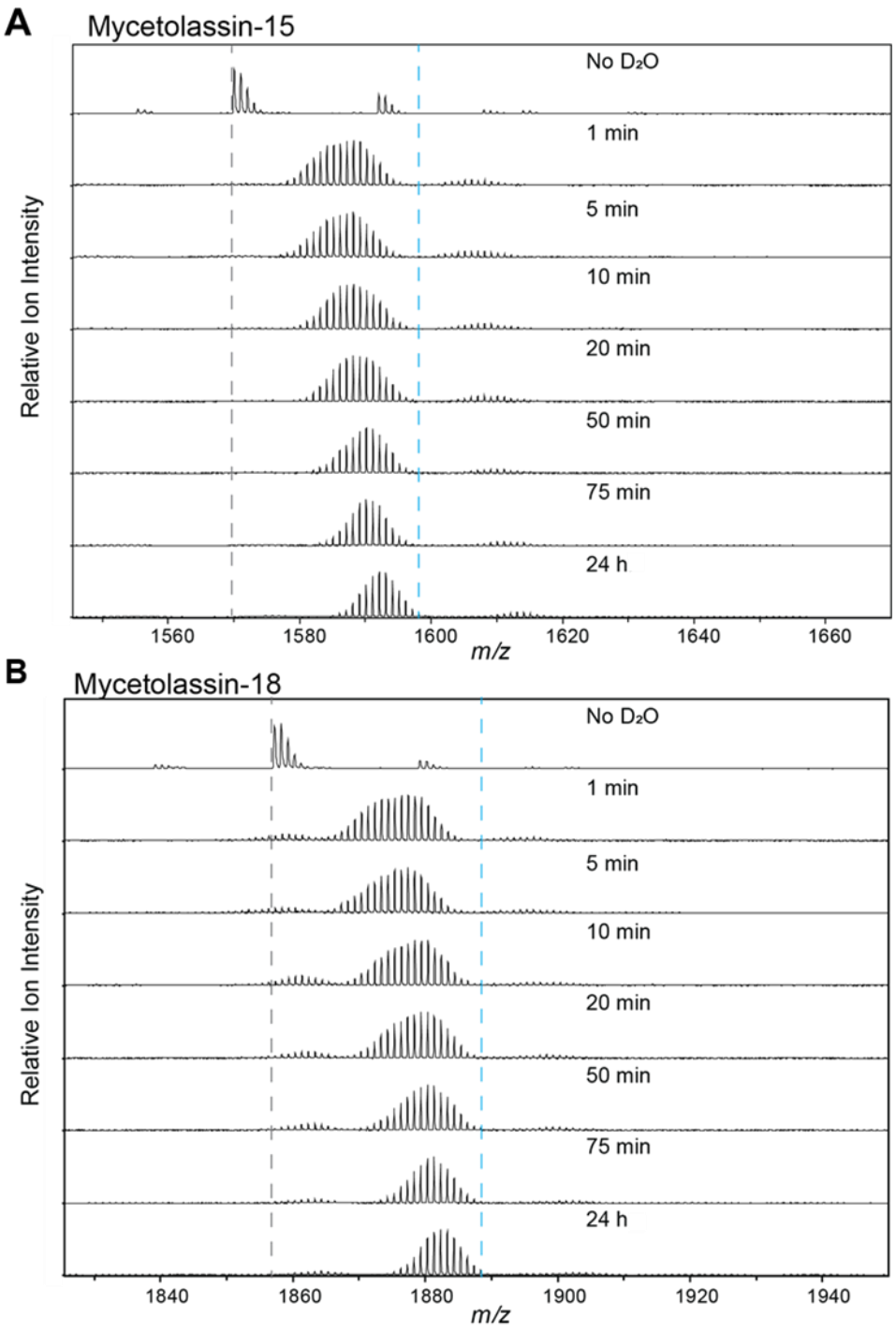



**Figure S13. Hydrogen-deuterium exchange of mycetolassin-15 and mycetolassin-18**. Raw data for mycetolassin-15 **(A)** and mycetolassin-18 **(B)** peptide deuteration over time. Partial deuteration occurred rapidly, but deuteration continued to progress for 24 h, which supports a threaded conformation. The expected mass of mycetolassin-15 and -18 is $[M+H]^+$ *m/z* = 1569.9 and 1856.9, respectively (shown in grey, dashed lines). With full deuteration the expected mass is $[M+H]^+$ *m/z* = 1598 and 1888 for mycetolassin-15 and mycetolassin-18, respectively (shown in blue, dashed line), approximated using MSTools.[22]

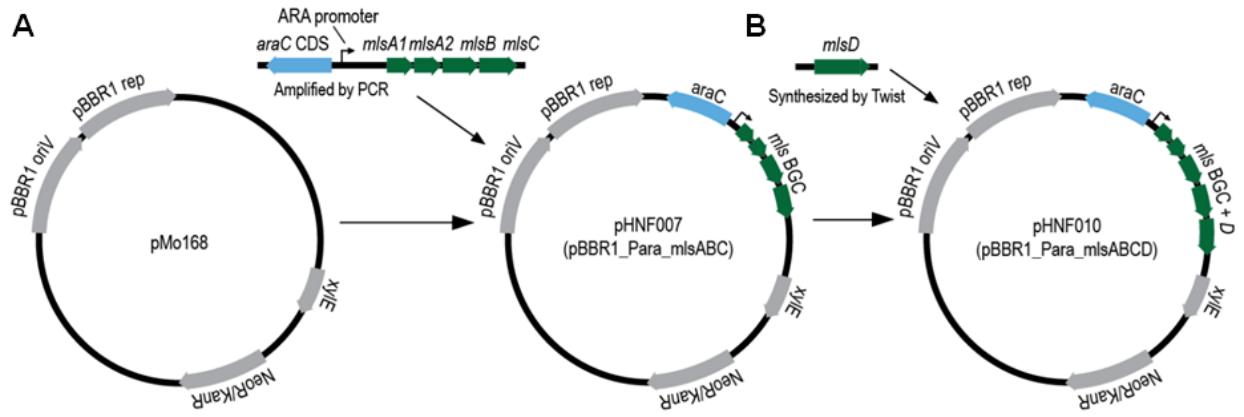



**Figure S14. Generation of pHNF007 and pHNF010. (A)** The mycetolassin BGC (*mlsA*$_1$*A*$_2$*BC* green) under the control of the *araC*/P$_{BAD}$ promoter was amplified by PCR and subcloned into pMo168 (pBBR1-based vector) to generate pHNF007. **(B)** The native *mlsD* transporter gene was synthetized by Twist and *Hind*III and *Sph*I restriction sites were used to subclone *mlsD* into pHNF007 to generate pHNF010.

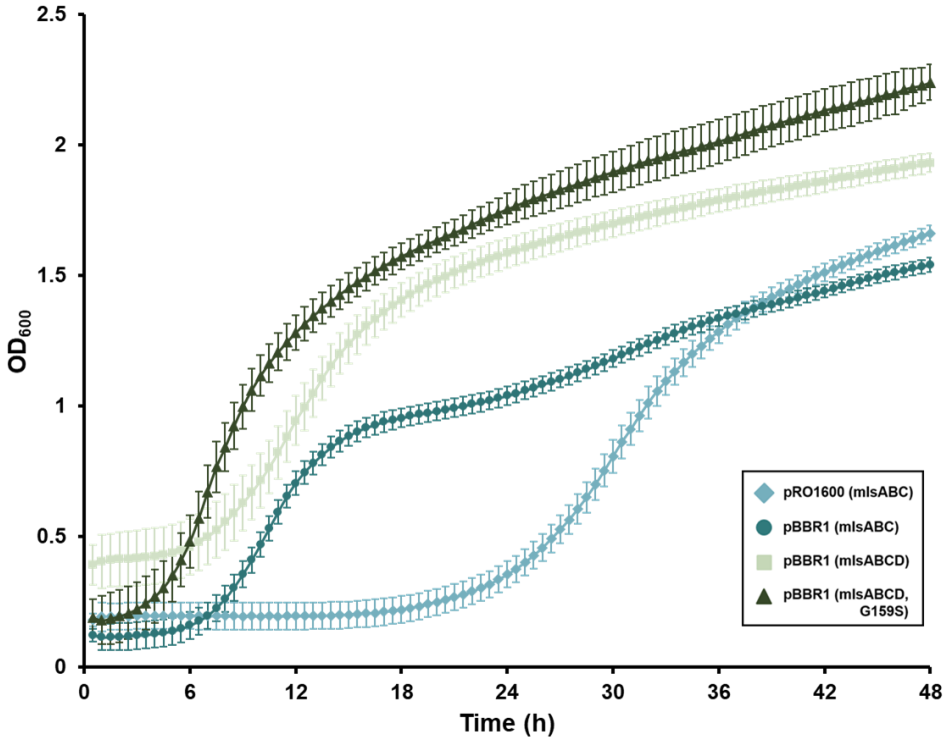



**Figure S15. Growth curve comparison between production with alternate *mls* expression vectors.** Forty-eight-hour growth comparison in production media between the *Burkholderia* sp. FERM BP-4321 spliceostatin-defective mutant strain (*fr9DEF*$^-$) containing alternate *mls* expression vectors. The average of $N = 10 \pm$ standard deviation is shown.

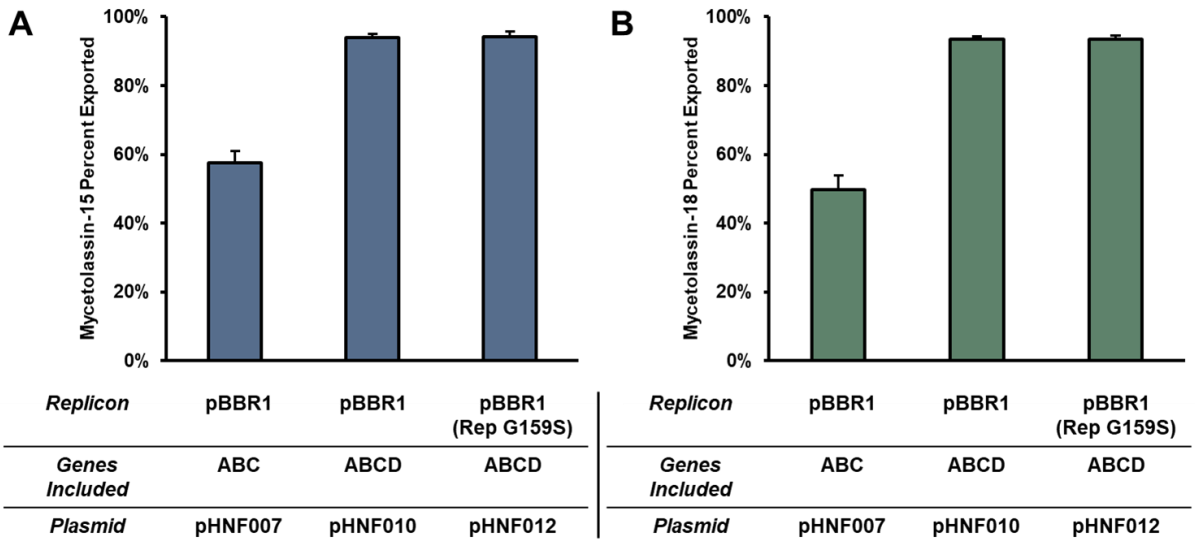



**Figure S16. Export of mycetolassins with and without inclusion of transporter gene. (A)** The percent of mycetolassin-15 detected in the supernatant of production cultures. **(B)** The percent of mycetolassin-18 detected in the supernatant of production cultures. Production was done in the *Burkholderia* sp. *fr9DEF*- strain. The average of triplicates ± standard deviation is shown.

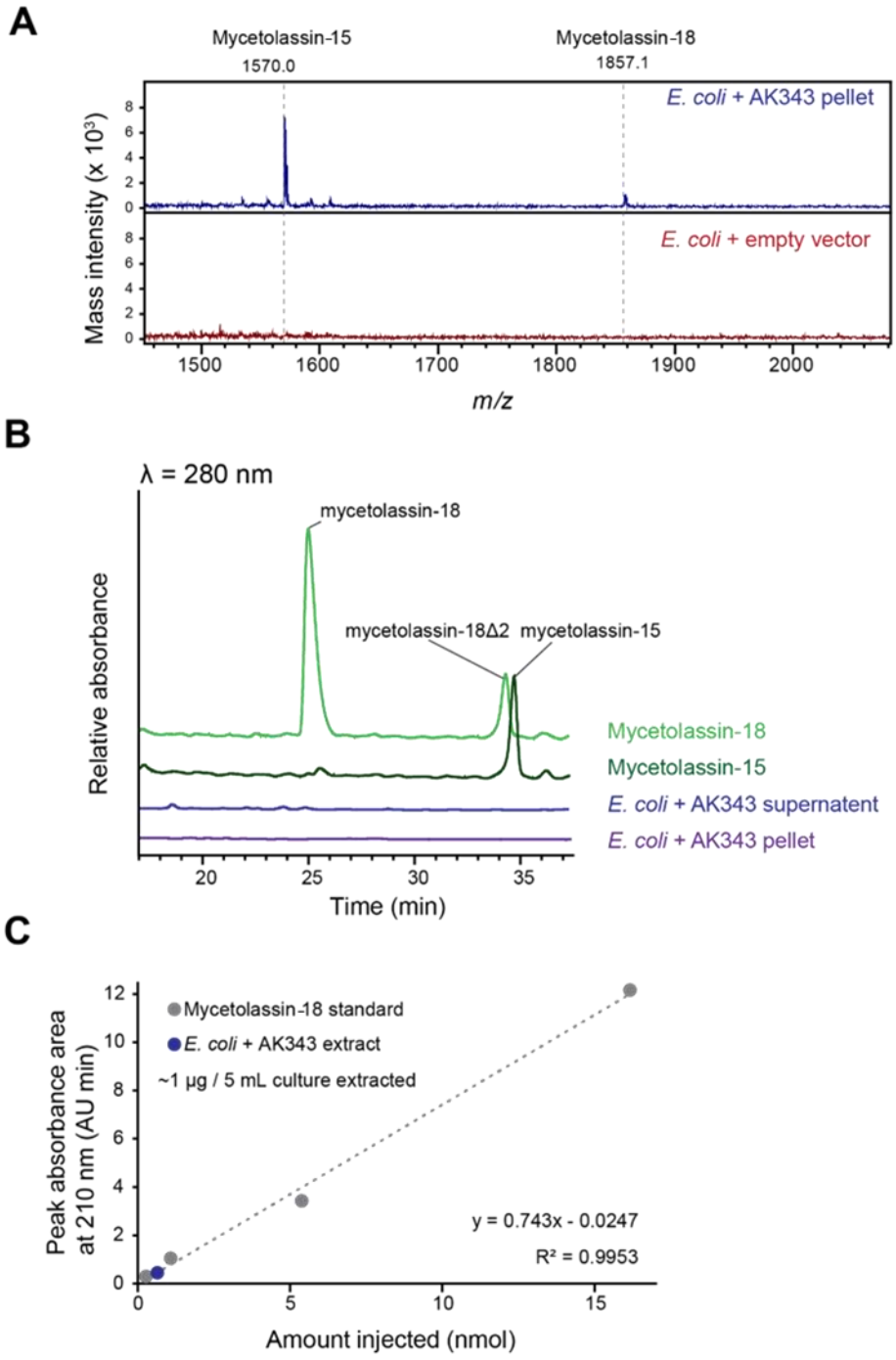



**Figure S17. Production of mycetolassins in *E. coli*. (A)** Masses corresponding to mycetolassin-15 and mycetolassin-18 (expected mass $[M+H]^+$ *m/z* = 1569.9 and 1856.9, respectively) were observed in trace amounts by MALDI-TOF-MS in *E. coli* expressing pAK343 (containing *mlsA1A2BC*) and not in the empty vector control. **(B)** HPLC traces of the pellet and supernatant extracts from *E. coli* do not show significant peaks at 280 nm at the retention times consistent with purified mycetolassin-15 and mycetolassin-18. **(C)** Quantification of mycetolassin-18 from a 500 mL *E. coli* culture expressing pAK343. 1% of the extract (eq. ~5 mL culture) was injected onto the HPLC and the area of the 220 nm absorbance peak corresponding to retention time of myceotolassin-18 was compared to a standard curve to determine yield ($N$ = 1). Blue circle, *E. coli* extract; grey circles, authentic standard.

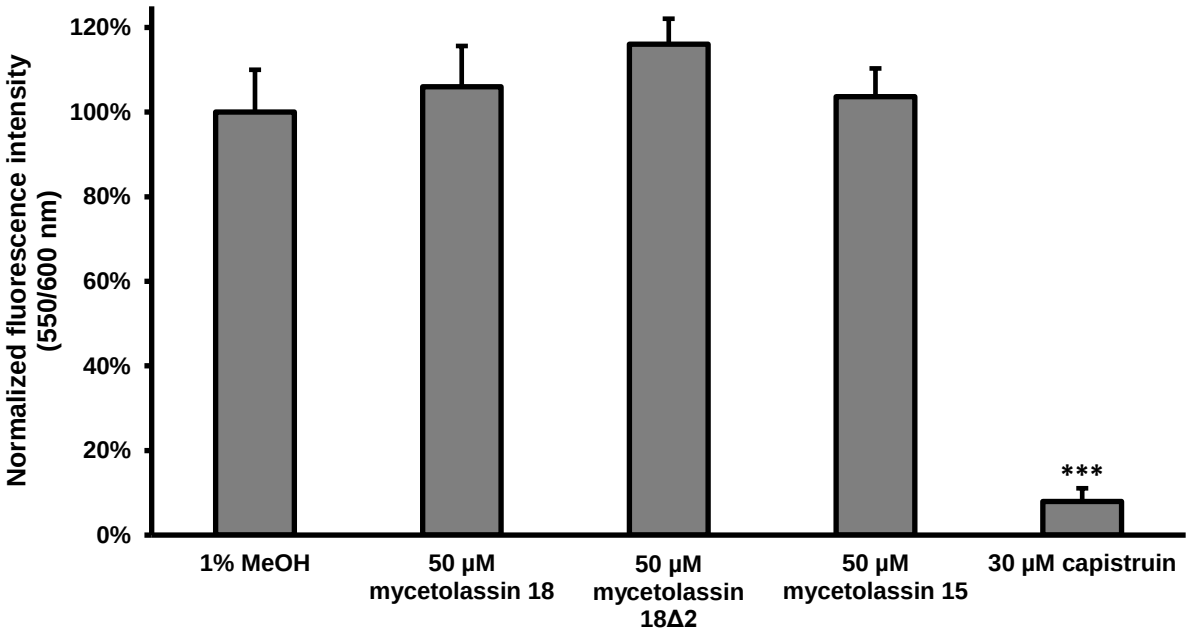



**Figure S18. Inhibition of transcription and/or translation assessed by cell free production of mCherry.** mCherry under the control of an endogenous *E. coli* σ70 promoter was produce in the presence of negative control 10% methanol, and 50 µM final concentration of mycetolassin-15, a C-terminal truncation of mycetolassin-18 (mycetolassin-18Δ2), and mycetolassin-18. Capistruin (30 µM), a known RNA polymerase inhibitor, is also shown. Error bars represent standard error of the mean ($N$ = 4), and only samples containing capistruin exhibited a statistically significant difference in fluorescent intensity. T-test was implemented for statistical analysis. ***, P = 0.0001.

Sequences for mycetolassin expression plasmids

***mlsA₁:***

5'- atgaataagcagcaagacgtgaagcatgaagtgtcggaattcctgttggatgacgaatcgctgatggaactgtccgcttccgaatcgacgctgggc ggttcgggacaataccgagaagctggcgttggccgctttttctga -3'

**Intergenic region:**

5'- catagggcgaccagccgcctgcctaccgggtagcggtttcctttccactacgggatcaaggagattctc -3'

***mlsA₂:***

5'- atgaataagcaacaagacgtgacgtatgaagtgtcgaacttcctgttggacgatgaatcgctgatggaactgtgcgcttccgagtcgacgctgggtg gtacggggcaatacaaagaagtcggcattggccgcttctatgatggagcgtaa -3'

**Intergenic region:**

5'- tcttagatgtattgtgacgcatactaaggttttgcttagcgttctcttgtagtcacgatacagaacaaaaaa -3'

***mlsB:***

5’- atgccgcatcatttgatgcccaatgtgttgggcgcctgctttgctaagcaaattatcttgctcgatgtggtgcgtgatcggtacgttgtcttgaccgagag
gcaatccgcagcattggcgtcgcagcttggtctggatattgagcgcttttcccaggctggcacgattcagcattcggccattctggaagaattgaagcgtg
ccaactggattagcgccacctcgtttccggaagctgacgctaaactgaacacgttgagcccctcgccgcaaacgggggggaatgtcggccaatggttgg
aagttgccggacgatgcttttgcgcatcgaccggcttggacgaacatcgccaaagccacgctcatcttgaggtccgtgcaccgatgcgcagcgcggcg
cggaatggccgggttgttgcagttgtgcaatcaagctcaccgtgaggccggagcaggcttgcgtgtcagccagccggaagattatgcacccctttgtgaa
cacgctcaattgggcctgcctgttttacccccggccgaccaagtgcttggagtggtcggtcgcgctgaccgtgctgtgcgcgcgtgccggcttggcattg
aaattggtgatcggcgtgcaatcctttccgttttatgcacacgcctggtcggaagccggagaggtggtgattggcgatagcccgctcaggcgcaaagag
ctgtccgttattctggaatgcccgcatccgctgctggcactgcccctgtga -3’

***mlsC:***

5’- atgaccctgctcggcatcgcctatccgaagacgcaggattatttgacgcaaaaacgggattggaggttcgtgcgcgacccgtccaaccagattgca
ggtgcactttggcatcaacaagacagggccaagacgcgtgccctgtgcgggcaatccggcacgcacttctactgcttttatggttcggcattttcgaagct
ggaccagagcgaaattaaggaagacgacttgcaggcaattgcggcagtgggtcccgcgcatctcatggctcgattctggggtaaatatttgtttatctgctt
tgacctgaccgccaaaaagttgatctgctgcagcgatccgagcaaccaatggccggtgtattgggcatggtccagccgtcacggtttgctttttcggatac
gatcacgtgcctgcatcaaaccttggtggagagcggcgagcgtccggcatggaataaaaagtttttacgacgtggttgcgcacgggtaccgtgcaaaat
ggtgagctgccgtttgatcgtatccaagaaattccggcaggctgcgctgtctcctatatcgatgatggtcaaccgcaaatcgtgacggtttgggatcccctc
aagcacgcagtggaagatgacacccaaaccccgttcggcattttgaagaattatcttaaccggtttgtgtcgttggccgatcggcccgtgctggaactctc
cggaggtctggaatccagctccgtgctgctggcaatgcgcgcagtctcctccgctgaccatccgctgtcctgcgcgcattattaccatgccggtgtcgctt
cgtccaacgagctggaacatgcgcgtcacgttgcacaaagcgtgcaggctcatctccatgaggtggatacgagggtgttgtcgtttaccccggttcagtc
cgttccgcgggtggctaaaccgcatatccgatattgcttgcttgcgtttaatcagtattttgcaagccaactcgagaaaggagagcatacccagttgatcag
cgggcatggaggggattcgctctttttggcgccgccgccgttcagcgcgcttgccgatgcggcactgacgtggcaatggcgacgattggtgcgcgtgg
cgatggacctggcgctcatttacagggcgccgctgattcaagtggcaacgggcgcagcccgatcgttgttttcctcggatccgtttgaaaataacggcag
cacgctcgatcttctgtccgaagaaacgcagcacgagtcgttcgatcaacgagcccatctgcatccgtttctccagcaaaataaggttcgtttgcccggtca
gctctatcagctggctttggctttttgacgctcgacgacattcgatcgcccgcctatccgtttacccaacccacgcattatccgtttctttgccagccgatggt
cgagtttgcgttgtccatccccagctacgatcatttcgaaggagcgcataatcgcattgtgctgaggaaatccgtttcggcggccacggggtatccgaacc
tgtggcgacgaaacaaaggcgaaacgtcgggtatcgatttgttgggcattcgtgaccacaaggatcatatgatggcgctctgcctcgagggatttttggcc
aaagagggctatgtcgacgtgaaacgcacgcgtgctgcgattcaacatagcagcaagggacatagcaaatattttatgggcctccttcacctctatgccg
ccgaactgtttgttcaaggatggcaatga -3’

**Intergenic region:**

5’- cgaattcgtaatcatgtcatagcaagcttctagcgagtaaagcgtg -3’

***mlsD:***

5’- atgcagcatggtctcttcaaaacaatcttgtttttgctgcattgcttgtggacgcgtgaatggaaaacgcgtcttcagatcatcgcagccattggctgtac
aggatgctcagtcgggctgaccattatcgcaccgatgttgctcaagtggctcatcgacggtttgtccaacgcggcgctctcggcagtgagcgtctacctgc
tctgtgtagcttacggcggcgcgtggtttctgtcacaagcgacgaaccgttgcgaacagtatttgttcgtgttggtcaatgaacgtattaaacgcacgattac

gttgaattacgtcagccgagtgctcggtcagcccgctgcggcattgagtgctcaccaaacgggcgcttctgtgagcgaactctcgcgcgtacaagacag
tgcgatgcgtgcaattatcgggctcttttggagcctcgcaccgattgttgtcgaaatcgcggtcgtcggcatcgtcgtgttctcattgtttggcctcgtttatgc
agctgtcttgctgggggttctgacccttttatatttactttagtgtatcaacggcgaagcgctgtgaagcgctgcaacgcaagaaaaaccgagcgtcggatc
gtgcctctacgcggctggtcgatacgctgacgcacgccgatacggtcaaagcgttttcgaatgaagccagcgaattgacccgactcgatcgtgtgttcaa
aatgcgcgcctgggcgagtatccatatcacgcgcaagatggaaatcgttggcgccactcaattgctgattgtcggtatcgggttgaccacgcttacggctt
tggcaggacaagatatcttggcacaccggctcaccattggtgatttcgtgctgttgaacagttacctgctgcgatttactttgccgctcagcacgttcggctat
gtggtgcgggaagttcagatcggtttgtttaatctgtcggaattgctgcacactatccgtgttagatggcacttcgaaccgttgggtcctcccctcttcaagca
agcaccgtcgattgaatttcgcagcgccagctttcgccatccgtgcggtagacgactgttgcaccacgcgaatttttcgattcctgctggcgccttctgtgc
gctggtggggccgagcggcgcggggaaatctacgctcatcaagctgttgttgcgtctgtataagctcgatgaagggcaaatccttatcgatgatcaaacg
atcgaatcgttttccgaatcgcagttgcgtcagtcgatcagttatgtgacacaagacgcgcatttgctggatcgctcgatccgcgaaaacatctgctttgcct
gcccgcaagcctctaattcggaacttgaacaagttatccattcctgtgcgctgacatccgttattgctgcggcgagccagggcctggacacgccgattggc
gagcgaggcaacaaattgtcgggtggtgaacgacaacgtgtgggactggcccgcgctttgttgcgcaatagccctttgctgctgttggacgaaccaact
gcggcgctcgatgcacacacggaggccgatatcttcaagcatctctatgagcgttatcgtcacaccacgcggctcgtgattgcgcaccgactttcggcca
ttcagcacgcggacttgattttggtcatgcaagagggtgcgatcgtggagacgggcacgcatagtagtttgctgaggcgacaaggtgtttatgccaatcttt
ggcataaacagacccagctgtcgacagaacaacctgacttgtcagcggcacgataa -3’

**Figure S19. Sequences of mycetolassin genetic constructs *mlsA$_1$A$_2$B, mlsC,* and *mlsD*.** The *mlsA$_1$A$_2$B* and *mlsC* constructs were purchased from GenScript and utilized to generate pAK343. To avoid rare codons from both *Burkholderia* sp. and *E. coli*, codons were swapped as follows: UUA → UUG, UCU → UCC, UCA → UCG, UAA → UGA, UGU → UGC, CUA → CUG, CCU → CCG, CCA → CCG, AUA → AUC, ACU → ACC, ACA → ACG, AGU → AGC, AGA → AGG, GUA → GUG. The native *mlsD* sequence was synthesized by Twist Bioscience and utilized in the pHNF010 and pHNF012 constructs. The native sequence originates from *Mycetohabitans* sp. (basonym *Burkholderia* sp.) B13 (NZ_FTPM01000001). Accession codes for each gene are as follows: *A$_1$,* WP_175972411*; A$_2$,* WP_175972412*; B,* WP_083705989; *C*, WP_076785565; *D*, WP_076785566.

## References


1. Maksimov, M. O.; Pelczer, I.; Link, A. J. Precursor-centric genome-mining approach for lasso peptide discovery. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 15223-15228.
2. Chekan, J. R.; Koos, J. D.; Zong, C.; Maksimov, M. O.; Link, A. J.; Nair, S. K. Structure of the lasso peptide isopeptidase identifies a topology for processing threaded substrates. *J. Am. Chem. Soc.* **2016**, *138*, 16452-16458.
3. Zong, C.; Wu, M. J.; Qin, J. Z.; Link, A. J. Lasso peptide benenodin-1 is a thermally actuated [1]rotaxane switch. *J. Am. Chem. Soc.* **2017**, *139*, 10403-10409.
4. Kodani, S.; Hemmi, H.; Miyake, Y.; Kaweewan, I.; Nakagawa, H. Heterologous production of a new lasso peptide brevunsin in *Sphingomonas subterranea*. *J. Ind. Microbiol. Biotechnol.* **2018**, *45*, 983–992.
5. Hegemann, J. D.; Zimmermann, M.; Zhu, S.; Klug, D.; Marahiel, M. A. Lasso peptides from proteobacteria: Genome mining employing heterologous expression and mass spectrometry. *Biopolymers* **2013**, *100*, 527-542.
6. Pan, S. J.; Rajniak, J.; Maksimov, M. O.; Link, A. J. The role of a conserved threonine residue in the leader peptide of lasso peptide precursors. *Chem. Commun.* **2012**, *48*, 1880-1882.
7. Knappe, T. A.; Linne, U.; Zirah, S.; Rebuffat, S.; Xie, X.; Marahiel, M. A. Isolation and structural characterization of capistruin, a lasso peptide predicted from the genome sequence of *Burkholderia thailandensis* E264. *J. Am. Chem. Soc.* **2008**, *130*, 11446-11454.
8. Kunakom, S.; Eustáquio, A. S. Heterologous production of lasso peptide capistruin in a *Burkholderia* host. *ACS Synth. Biol.* **2020**, *9*, 241-248.
9. Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. Caulosegnins i-iii: A highly diverse group of lasso peptides derived from a single biosynthetic gene cluster. *J. Am. Chem. Soc.* **2013**, *135*, 210-222.
10. Cheung-Lee, W. L.; Parry, M. E.; Cartagena, A. J.; Darst, S. A.; James Link, A. Discovery and structure of the antimicrobial lasso peptide citrocin. *J. Biol. Chem.* **2019**, *294*, 6822-6830.
11. Metelev, M.; Arseniev, A.; Bushin, L. B.; Kuznedelov, K.; Artamonova, T. O.; Kondratenko, R.; Khodorkovskii, M.; Seyedsayamdost, M. R.; Severinov, K. Acinetodin and klebsidin, RNA polymerase targeting lasso peptides produced by human isolates of *Acinetobacter gyllenbergii* and *Klebsiella pneumoniae*. *ACS Chem. Biol.* **2017**, *12*, 814-824.
12. Fuwa, H.; Hemmi, H.; Kaweewan, I.; Kozaki, I.; Honda, H.; Kodani, S. Heterologous production of new lasso peptide koreensin based on genome mining. *J. Antibiot.* **2021**, *74*, 42-50.
13. Yu, H.; Ding, X.; Shang, L.; Zeng, X.; Liu, H.; Li, N.; Huang, S.; Wang, Y.; Wang, G.; Cai, S.; et al. Protective ability of biogenic antimicrobial peptide microcin J25 against enterotoxigenic *Escherichia coli*-induced intestinal epithelial dysfunction and inflammatory responses IPEC-J2 Cells. *Front. Cell. Infect. Microbiol.* **2018**, *8*, 242-242.
14. Li, Y.; Han, Y.; Zeng, Z.; Li, W.; Feng, S.; Cao, W. Discovery and bioactivity of the novel lasso peptide microcin y. *J. Agric. Food. Chem.* **2021**, *69*, 8758-8767.

15. Bratovanov, E. V.; Ishida, K.; Heinze, B.; Pidot, S. J.; Stinear, T. P.; Hegemann, J. D.; Marahiel, M. A.; Hertweck, C. Genome mining and heterologous expression reveal two distinct families of lasso peptides highly conserved in endofungal bacteria. *ACS Chem. Biol.* **2020**, *15*, 1169-1176.

16. Cheung-Lee, W. L.; Cao, L.; Link, A. J. Pandonodin: A proteobacterial lasso peptide with an exceptionally long c-terminal tail. *ACS Chem. Biol.* **2019**, *14*, 2783-2792.

17. Xiu, H.; Wang, M.; Fage, C. D.; He, Y.; Niu, X.; Han, M.; Li, F.; An, X.; Fan, H.; Song, L.; et al. Discovery and characterization of rubrinodin provide clues into the evolution of lasso peptides. *Biochemistry* **2022**, *61*, 595-607.

18. Cheung-Lee, W. L.; Parry, M. E.; Zong, C.; Cartagena, A. J.; Darst, S. A.; Connell, N. D.; Russo, R.; Link, A. J. Discovery of ubonodin, an antimicrobial lasso peptide active against members of the *Burkholderia cepacia* complex. *ChemBioChem* **2020**, *21*, 1335-1340.

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

20. Si, Y.; Kretsch, A. M.; Daigh, L. M.; Burk, M. J.; Mitchell, D. A. Cell-free biosynthesis to evaluate lasso peptide formation and enzyme-substrate tolerance. *J. Am. Chem. Soc.* **2021**, *143*, 5917-5927.

21. Zallot, R.; Oberg, N.; Gerlt, J. A. The EFI web resource for genomic enzymology tools: Leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry* **2019**, *58*, 4169-4182.

22. Kavan, D.; Man, P. MStools—web based application for visualization and presentation of HXMS data. *Int. J. Mass spectrom.* **2011**, *302*, 53-58.