Electronic Supplementary Material (ESI) for Chemical Communications

Supporting Information

The Role of a Conserved Threonine Residue in the Leader Peptide of Lasso Peptide Precursors

Si Jia Pan, [a, ‡] **Jakub Rajniak,** [a, ‡] **Mikhail O. Maksimov** [a] **and A. James Link*** [a,b]

Departments of [a]Chemical and Biological Engineering and [b]Molecular Biology, Princeton University, Princeton, NJ 08544

*to whom correspondence should be addressed, A207 Engineering Quadrangle, Princeton University, Princeton, NJ 08544, (609) 258-7191, ajlink@princeton.edu

[‡]These authors contributed equally.

Supplementary Figures and Table Mentioned in Main Text

***Figure S1, S2 and Table S1 appear on the following three pages.**

Figure S1: (a) Confirmed lasso peptides: microcin J25 (*E. coli* AY25), capistruin (*B. thailandensis* E264), SRO15-2005 (*S. roseosporus* NRRL 15998), SSV-2083 (*S. sviceus* ATCC 20983)[1]. (b) Putative lasso peptides found through genome mining by Severinov et al[2]. The leader portion of each peptide is shown in blue. The threonine found in the pentultimate position of the leader peptides is highlighted in yellow.

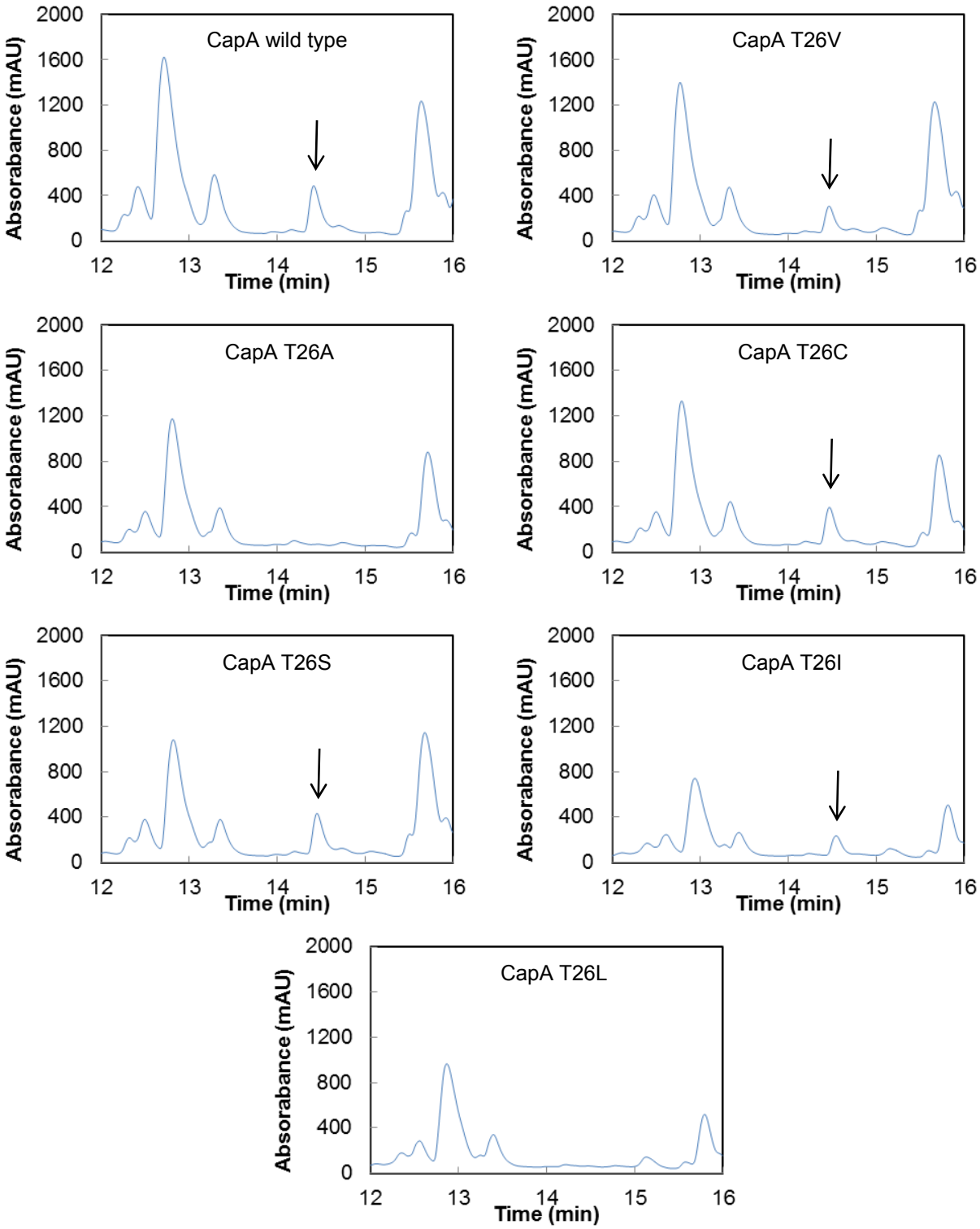

Figure S2: HPLC traces of culture supernatants from *E. coli* transformed with CapA leader peptide variants. Arrows indicate the position of the capistruin peak (~14.4 min). These peaks were integrated to generate the data presented in Table 1 of the main text.

Table S1. Production level of MccJ25 Variants (raw data for Figure 2B)

McjA Construct	HPLC Area (SD*)	Last Active Dilution
T35G	ND*	1
T35A	ND	8
T35S	515 (42)	64
T35C	4510 (132)	512
T35V	4659 (591)	512
Wild type	6583 (606)	1024
T35I	4788 (413)	512
T35L	ND	16
T35D	ND	NA*
T35M	ND	1
T35F	ND	NA

*ND: not determined, NA: no activity, SD: standard deviation

Supplementary Results

A Novel High-Yielding E. coli *Heterologous Expression System for Capistruin*

Whereas MccJ25 can be produced with yields of ca. 8 mg/L of culture,[3] capistruin is produced at significantly lower levels with reported yields of 700 µg/L of culture from its natural producer *Burkholderia thailandensis* and 100-200 µg/L from a previously described heterologous *E. coli* expression system using specialized M20 minimal media.[4, 5] In this heterologous expression system, Knappe *et al.* placed the entire *capABCD* cluster under the control of an inducible T7 promoter.[4] The low yield of capistruin makes quantitative analyses of production level, such as those brought about by amino acid substitutions in the leader peptide, more challenging than similar measurements with MccJ25.

The intergenic region between *capA* and *capB* is 78 bases long and includes an inverted repeat sequence that may function to regulate expression of genes in the cluster (Fig. S2a). We constructed a new heterologous expression system, harbored on plasmid pJR30, in which the intergenic region between *capA* and *capB* was removed and replaced with an *E. coli* optimized ribosome binding site (Fig. S2a). We compared capistruin production using this new heterologous system to capistruin production in the native host, *B. thailandensis* and the difference was striking. *B. thailandensis* M20 culture supernatants did not exhibit a discernible capistruin peak in HPLC analyses using a 215 nm UV detector (Fig. S2b), though capistruin was detected in MALDI-MS analyses of the supernatant. In contrast, culture supernatants from *E. coli* transformed with pJR30 and cultivated in M9 medium produced so much capistruin that it was one of the most abundant species in the extracellular fluid (Fig. S2b). To quantify the production of capistruin from this cluster, we grew a one liter culture of *E. coli* BL21 harboring pJR30, isolated pure capistruin by preparative HPLC, and confirmed its identity using tandem mass spectrometry (Fig. S3). The yield of capistruin from this experiment was 1.6 mg/L of culture, an increase of at least 8-fold over the previously reported heterologous expression system and a greater than 2-fold increase over the native producer strain yield.[4, 5]

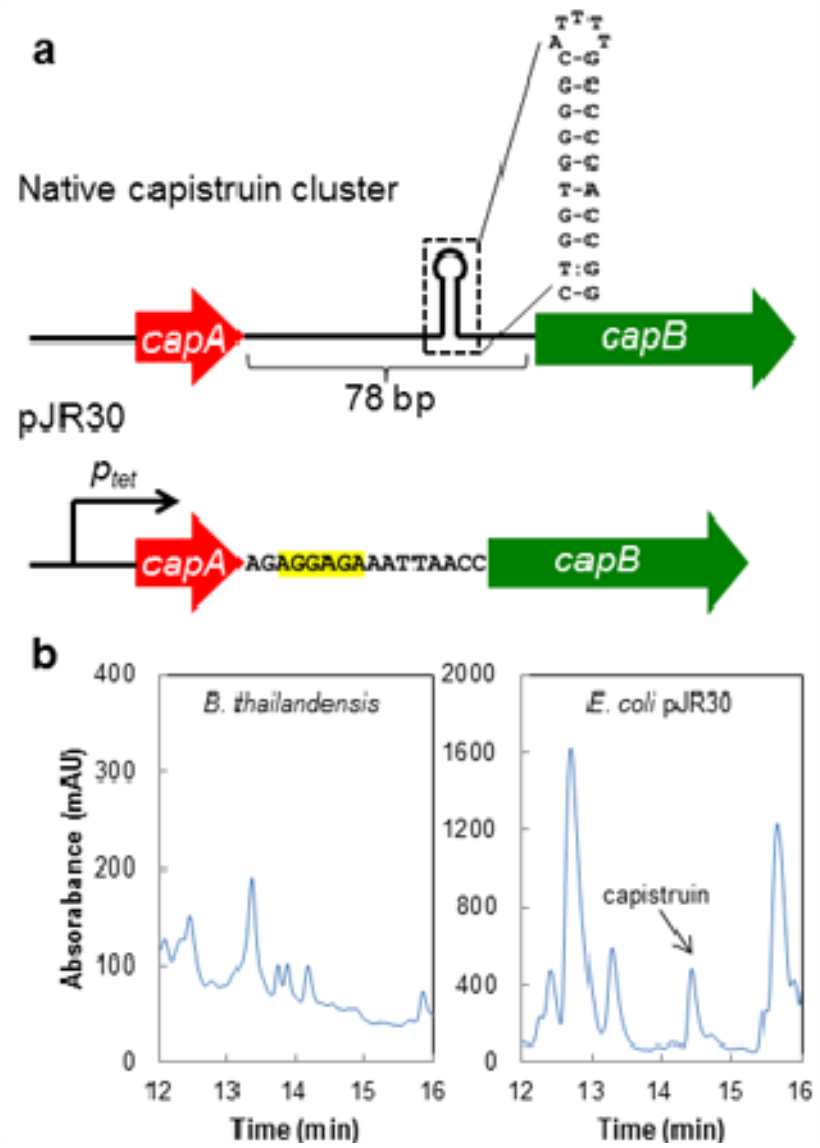

Figure S2 Heterologous production system for capistruin. (a) The native capistruin gene cluster contains a 78 bp intergenic region that includes an inverted repeat sequence. In the pJR30 plasmid, the cluster is under the control of an inducible tetracycline promoter and the intergenic region is replaced by a short sequence containing an *E. coli* Shine-Dalgarno sequence (yellow). The length of the CapB gene in pJR30 is 241 aa (see Supporting Information). (b) Comparison of HPLC traces (215 nm UV detector) from culture supernatants of native capistruin producer B. thailandensis (left) and E. coli harboring pJR30 (right). The peak at 14.4 minutes is capistruin. Note that the scales on the y-axes are not equivalent.

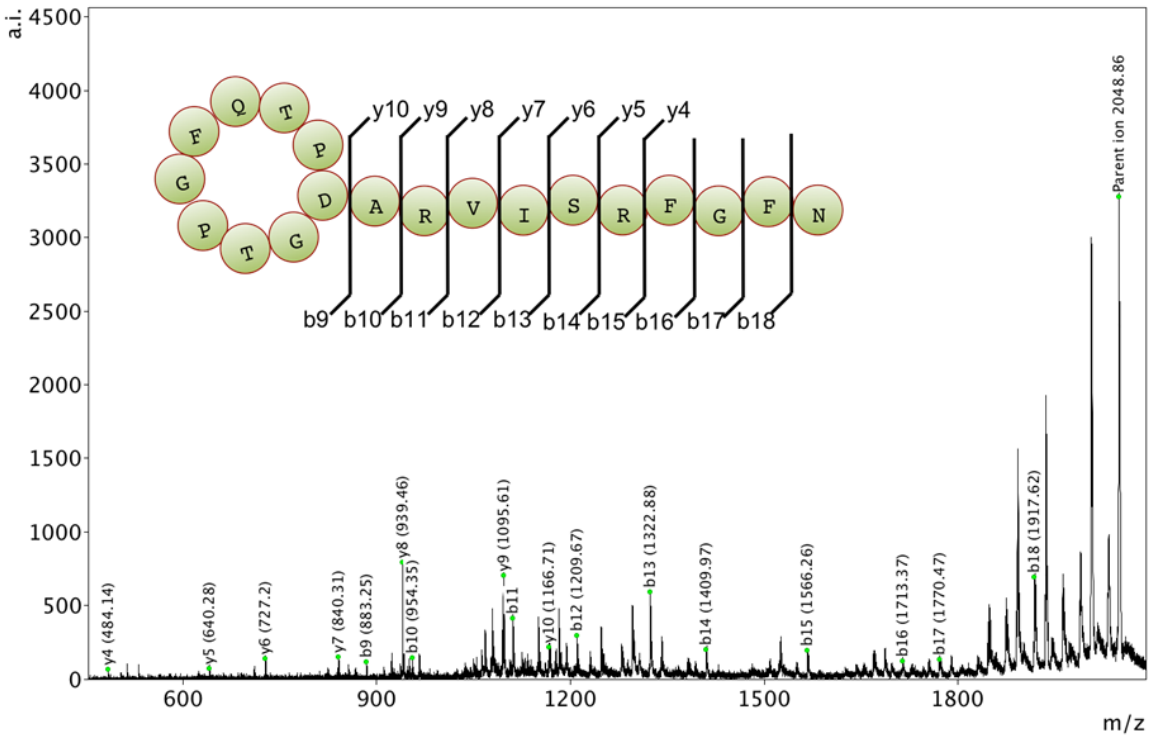

Figure S3: MS/MS analysis of purified capistruin. The parent ion and y and b series ions corresponding to capistruin are labeled on the spectrum, and the observed fragments are noted on the drawing.

Revision of the ORF Length of CapB

In the original description of the capistruin gene cluster, the length of the open reading frame (ORF) encoding the CapB protein was reported as 221 aa.[4] Analysis of the cluster sequence upstream of the start codon of the putative 221 aa ORF revealed two additional start codons in the same reading frame, one resulting in a 241 aa ORF and the second resulting in a 255 aa ORF. The inverted repeat hairpin sequence we describe above (Figure S4) is immediately upstream of the start codon of the 241 aa ORF, so we hypothesized that the true ORF length of CapB was 241 aa rather than 221 aa. Two "hairpin-less" constructs were made (see Materials and Methods below); the plasmid pJR29 harbors a hairpin-less capistruin cluster with the 221 aa CapB ORF and the plasmid pJR30 is identical to pJR29 except that it harbors the 241 aa CapB ORF. In our tests of capistruin production, *E. coli* transformed with pJR29 were unable to produce an amount of capistruin detectable by HPLC (Figure S3) or by MALDI-MS. In contrast, *E. coli* harboring pJR30 were able to produce large amounts of capistruin that were readily evident in HPLC analyses (Figure 3 in the main text, Figure S3). This result indicates that the 241 aa CapB ORF is both necessary and sufficient for the biosynthesis of capistruin.

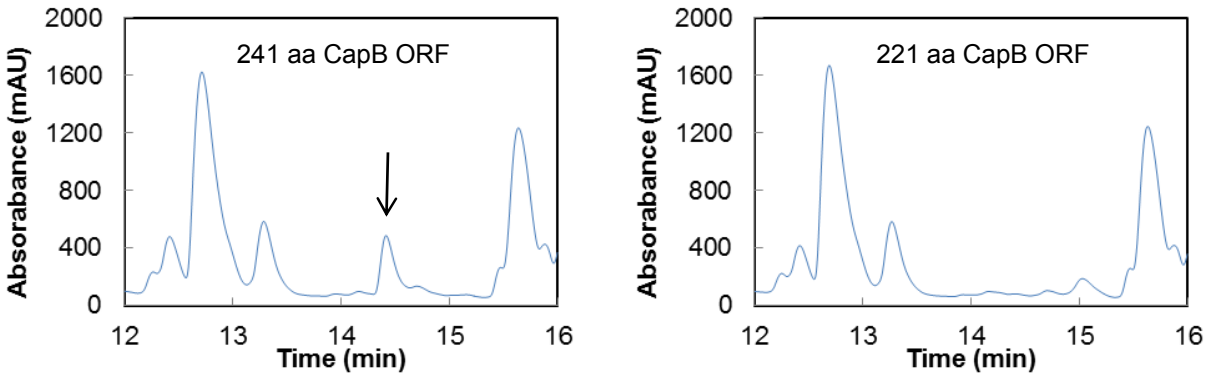

Figure S4: The correct length of CapB is 241 aa. Left: HPLC trace of culture supernatant from *E. coli* BL21 pJR30, which is a capistruin cluster harboring a 241 aa CapB ORF. Arrow shows the location of the capistruin peak. Right: As in the left panel, but with *E. coli* BL21 pJR29, which contains a 221 aa CapB ORF. No capistruin production is observed.

Materials and Methods

General

PicoMaxx DNA polymerase was used for all PCR manipulations. Restriction enzymes were from New England Biolabs. Oligonucleotides were from IDT, and the integrity of all constructs was confirmed by DNA sequencing at Genewiz. All recombinant DNA steps were carried out in the *E. coli* strain XL-1 Blue. Ampicillin was added to *E. coli* cultures when necessary to a concentration of 100 mg/L.

Plasmids

Mutations at Thr-35 in the leader peptide of *mcjA* were introduced in pJP3, a pQE60-derived ampicillin-resistant plasmid carrying the engineered MccJ25 biosynthetic gene cluster described previously [3]. To construct the T35S variant plasmid, the region in pJP3 flanked by restriction enzymes *Xho*I and *Hind*III containing *mcjA* and its upstream promoter region was amplified by PCR as two fragments with overlapping end regions that contain the T35S mutation. The two sets of primers used are listed in Table 1: *mcjA* forward and T35S reverse, T35S forward and *mcjA* reverse. In a second PCR step, the two amplified fragments were joined into one piece to produce the mutant *mcjA* gene. The resulting PCR fragment and the plasmid pJP3 were both digested with *Xho*I and *Hind*III and then ligated to form the T35S plasmid. The other *mcjA* mutants were constructed in a similar fashion using the primers listed in Table 1.

The pJR4 plasmid contains the entire capistruin gene cluster in the pASK-75 vector. To construct pJR4, primers *capABCD* forward and *capABCD* reverse, containing *XbaI* and *KpnI* restriction sites, respectively, (see Table 2) were used to amplify the entire 4.5 kb *capABCD* gene cluster from *B. thailandensis* along with a 70 bp region directly upstream of the *capA* gene. Whole genomic DNA was isolated from *B. thailandensis* using a standard protocol (Qiagen DNeasy Blood & Tissue Kit). The resulting PCR amplicon and ampicillin-resistant plasmid pASK75[6] were both digested with *XbaI* and *KpnI* restriction enzymes and subsequently ligated to form pJR4.

To construct pJR29, two pairs of primers were designed: *capA XbaI* forward and *capB* RBS reverse, and *capB* 221 RBS forward and *capB BsmBI* reverse (see Table 2). Primer *capA XbaI* forward anneals just upstream of the *capA* gene and contains an *XbaI* site, while primer *capB BsmBI* reverse anneals in the middle of the *capB* gene and contains an endogenous *BsmBI* site. Primers *capB* RBS forward and *capB* (221) RBS reverse are partially complementary, and anneal to the end of the *capA* gene and the beginning of the *capB* gene, separated by an RBS sequence transplanted from the pQE plasmid series (Qiagen). Initially, two fragments with overlapping end regions were amplified by PCR from pJR4 using the two primer pairs listed above; the two fragments were joined in a subsequent PCR step. The resulting final PCR amplicon and pJR4 were both digested with *XbaI* and *BsmBI* and subsequently ligated to form pJR29, which lacks both the RNA hairpin sequence found naturally between the *capA* and *capB* genes and most of the 70 bp region from *B. thailandensis* genomic DNA found upstream of *capA* in pJR4.

In constructing pJR29, we assumed that the *capB* gene encodes a 221 aa protein, as initially reported.[4] However, inspection of the region upstream of the putative start codon revealed two additional start codons in the same reading frame, potentially resulting in proteins 241 or 255 aa in length. We constructed an analogue of pJR29, termed pJR30, containing the 241 aa *capB* gene. The construction method was identical to that described above, except that primer *capB* (241) RBS forward was used in place of primer *capB* (221) RBS forward (Table S1).

Mutations at Thr-26 in the CapA leader peptide were introduced into pJR30 in a similar manner as for the McjA Thr-35 mutants. Briefly, to construct the T26S mutant, the region in pJR30 flanked by *XbaI* and *BsmBI* restriction sites containing the entire *capA* gene was amplified by PCR as two fragments with overlapping end regions, with one fragment containing the desired mutation. The two primer pairs used were *capA XbaI* forward and T26S reverse, and T26X forward and *capB BsmBI* reverse. The two fragments were joined in a subsequent PCR step, and the resulting PCR amplicon and pJR30 were digested with *XbaI* and *BsmBI* and subsequently ligated to form the T26S plasmid. The other *capA* mutants were constructed in a similar fashion using the primers listed in Table 1; note primer T26X forward is generic, and was therefore used in cloning all *capA* mutants.

Table S2. Oligonucleotides used in this study

McjA T35 variants	
mcjA forward	GCCCTTTCGTCTTCACCTCGAG
mcjA reverse	CCGGCGGCAACCGAGCGTTCTGAAC
T35S forward	CAGCATCGCAACTCTCAAAAGGTGGTGCAGG
T35S reverse	CCTGCACCACCTTTTGAGAGTTGCGATGCTG
T35C forward	CAGCATCGCAACTCTGCAAAGGTGGTGCAGG
T35C reverse	CCTGCACCACCTTTGCAGAGTTGCGATGCTG
T35D forward	CAGCATCGCAACTCGATAAAGGTGGTGCAGG
T35D reverse	CCTGCACCACCTTTATCGAGTTGCGATGCTG
T35V forward	CAGCATCGCAACTCGTAAAAGGTGGTGCAGG
T35V reverse	CCTGCACCACCTTTTACGAGTTGCGATGCTG
T35G forward	CAGCATCGCAACTCGGTAAAGGTGGTGCAGG
T35G reverse	CCTGCACCACCTTTACCGAGTTGCGATGCTG
T35A forward	CAGCATCGCAACTCGCAAAAGGTGGTGCAGG
T35A reverse	CCTGCACCACCTTTTGCGAGTTGCGATGCTG
T35I forward	CAGCATCGCAACTCATCAAAGGTGGTGCAGG
T35I reverse	CCTGCACCACCTTTGATGAGTTGCGATGCTG
T35L forward	CAGCATCGCAACTCTTAAAAGGTGGTGCAGG
T35L reverse	CCTGCACCACCTTTTAAGAGTTGCGATGCTG
T35M forward	CAGCATCGCAACTCATGAAAGGTGGTGCAGG
T35M reverse	CCTGCACCACCTTTCATGAGTTGCGATGCTG
T35F forward	CAGCATCGCAACTCTTCAAAGGTGGTGCAGG
T35F reverse	CCTGCACCACCTTTGAAGAGTTGCGATGCTG
pJR4 cloning	
capABCD forward	GGCGATGCTCTAGACCTAACGCCGGATGC
capABCD reverse	GCGCAAGCGCGGGTACCCGCGCATCAC
pJR29 and pJR30 cloning	
capA XbaI forward	GGACACCGGCACACTCGATATCTAGAAGGAGATTAAAATGG
capB RBS reverse	CATGGTTAATTTCTCCTCTTTAATTGAACCCGAAGCGCGAAATGAC
capB (221) RBS forward	GTTCAATTAAAGAGGAGAAATTAACCATGCAACGGTCGCGCTATTTTC
capB (241) RBS forward	GTTCAATTAAAGAGGAGAAATTAACCATGACGCCAGCCAGCCATTGCC
capB BsmBI reverse	GAGCCGAGCGAGACGGTCGAGC

CapA T26 variants	
capA XbaI forward	GGACACCGGCACACTCGATATCTAGAAGGAGATTAAAATGG
capB BsmBI reverse	GAGCCGAGCGAGACGGTCGAGC
T26X forward	CACGGCACCCCGGGTTTCCAAACCCC
T26S reverse	GGGTTTGGAAACCCGGGGTGCCGTGCGACGAGGAAAC
T26C reverse	GGGTTTGGAAACCCGGGGTGCCGTGGCACGAGGAAAC
T26V reverse	GGGTTTGGAAACCCGGGGTGCCGTGCACCGAGGAAAC
T26A reverse	GGGTTTGGAAACCCGGGGTGCCGTGCGCCGAGGAAAC
T26I reverse	GGGTTTGGAAACCCGGGGTGCCGTGGATCGAGGAAAC
T26L reverse	GGGTTTGGAAACCCGGGGTGCCGTGCAACGAGGAAAC

MccJ25 production and antimicrobial assay

MccJ25 from various leader peptide mutants was produced in *E. coli* strain DH5α. Cells were cultured in LB medium supplemented with ampicillin at 37°C with shaking. IPTG was added to induce MccJ25 production when cultures reached OD_{600} of ~0.4-0.5. Cells were grown overnight after induction and then culture supernatants were obtained by centrifugation.

The antimicrobial assay as described previously[7] was used to detect the level of MccJ25 in culture supernatants. Briefly, 5 µL drops of culture supernatants were spotted on M63 plates overlaid with soft agar inoculated with 10^7 colony-forming units per mL of MccJ25-susceptible *Salmonella Newport*. The presence of MccJ25 in supernatants resulted in zones of growth inhibition on the plates. Serial dilution of the culture supernatants was performed by dilution in LB medium and followed by the antimicrobial assay.

HPLC quantification of MccJ25

The quantification of MccJ25 in culture supernatants using analytical HPLC has been described previously.[3] Briefly, culture supernatants were extracted with *n*-butanol. 1 mL of each butanol extract was dried under reduced pressure, re-dissolved in 200 µL of water, and then applied to HPLC using previously described gradient and column. The HPLC chromatogram was analyzed by integrating the area under the peak corresponding to MccJ25 to determine the relative level in the supernatants.

Capistruin production in native producer *B. thailandensis*

Burkholderia thailandensis E264 cells (ATCC) were initially grown in 5 mL of LB medium at 37°C with shaking overnight. Cells were then subcultured into 50 mL of unsupplemented M20 medium (20 g/L glutamic acid, 0.2 g/L alanine, 1.0 g/L sodium citrate, 20 g/L anhydrous Na_2HPO_4, 0.5 g/L KCl, 0.5 g/L anhydrous Na_2SO_4, 0.2 g/L $MgCl_2{\cdot}6H_2O$, 0.0076 g/L anhydrous $CaCl_2$, 0.01 g/L $FeSO_4{\cdot}7H_2O$, 0.0076 g/L $MnSO_4{\cdot}H_2O$) to an initial OD_{600} of 0.01 and grown at 42°C with shaking for 48 hours, whereupon the culture supernatant was obtained by centrifugation.

Heterologous capistruin production

E. coli BL21 cells transformed with a capistruin-producing plasmid were initially grown in 5 mL of LB medium supplemented with ampicillin at 37°C with shaking overnight. Cells were then subcultured into 50 mL of M9 minimal medium (3 g/L Na_2HPO_4, 1.5 g/L KH_2PO_4, 0.5 g/L NH_4Cl,

0.25 g/L NaCl, 2 g/L glucose, 1 mM $MgSO_4$, and 500 µg/L thiamine) to an initial OD_{600} of 0.01. Cultures were induced with anhydrotetracycline (0.2 mg/L) upon reaching an OD_{600} of 0.3, induced for 48 hours at 37°C with shaking, and supernatants obtained by centrifugation.

Reversed-phase extraction and quantification of capistruin

Supernatants from all cultures grown in M9 or M20 medium were subjected to solid-phase extraction on Strata C8 columns (Phenomenex). After passing all supernatant (50 mL) through the resin and washing with water (12 mL), elution was performed with 6 mL of methanol. Eluates were evaporated to dryness under reduced pressure and dissolved in 600 µL of 10% acetonitrile. Reconcentrated eluates were analyzed by HPLC as described previously.[3] Under the conditions used, capistruin elutes at approximately 14.4 minutes.

Determination of heterologous capistruin yield

E. coli BL21 harboring pJR30 cells were grown overnight in 10 mL M9 minimal media as previously described. This starter culture was used to inoculate 1 L of M9 minimal media to an initial OD_{600} of 0.003. Capistruin expression was induced at OD_{600} of 0.3 by the addition of 200 µg/L of anhydrotetracycline, and was allowed to proceed for 48 hours at 37 ˚C with shaking at 250 rpm. The culture supernatant was obtained by centrifuging the culture for 12 min at 6000 x g. The cell-free supernatant was split into 5 aliquots of 200 mL each. Using a peristaltic pump, an aliquot was applied to a 6 mL Phenomenex Strata C8 SPE column containing 0.5 g of packing. The column was washed with two column volumes (12 mL) of water and eluted with 8 mL methanol. This was repeated with the other four aliquots of culture supernatant to afford a final eluent volume of 40 mL that was subsequently dried *in vacuo*. Water (8.5 mL) was used to reconstitute the extract which was then applied to a fresh C8 column. The column was eluted in a stepwise fashion by adding 1.8 mL of 13% acetonitrile (ACN) in water, 1.8 mL of 23% ACN/water, 1.8 mL of 30% ACN/water, and finally 1.8 mL of methanol. These fractions were evaporated to dryness using a speedvac and reconstituted in 0.5 mL of 50% ACN/water. Analysis by analytical and semi-preparative HPLC using the gradients previously described[3] of the fractions confirmed the presence of capistruin in the 30% ACN/water and methanol fractions (retention time 14.4 min on analytical column and 14.1 min on semi-prep). This retention time was used to purify capistruin from the 30% ACN/water fractions. The eluent from the HPLC (25 mL) was reduced to 1.5 mL under reduced pressure and then lyophilized to a fluffy white powder affording a final dry weight of 1.6 mg.

Mass Spectrometry

Mass spectra were obtained at the Rutgers/CABM core facility using an ABI MDS SCIEX 4800 MALDI-TOF/TOF mass spectrometer. The instrument was calibrated with external standards to an accuracy of 50 ppm or better. Samples were dried, reconstituted in 50/50 ACN/water and spotted (1 µL) onto an Applied Biosystems (ABI) 384 Opti-TOF 123mm x 81mm SS plate and allowed to dry. An equal volume of a 2.5 mg/mL solution of the matrix α-Cyano-4-hydroxycinnamic acid (Sigma), dissolved in 50% ACN/water 0.1% trifluoroacetic acid (TFA) was mixed with the spotted sample and allowed to dry. For tandem mass spectrometry of the purified capistruin sample (Figure S1), MS^1 spectra were obtained in positive reflector mode (300 cm ion path length). Singly protonated m/z 2048.99 capistruin was selected for fragmentation in 1 kV positive MS/MS mode (240 cm ion path length). Metastable ion suppression was turned on to suppress precursor ions and metastable fragment ions from

reaching the detector. 2000 shots were collected per MS^2 spectrum. The data were analyzed using mMass software.[8]

References

1. R. D. Kersten, Y. L. Yang, Y. Xu, P. Cimermancic, S. J. Nam, W. Fenical, M. A. Fischbach, B. S. Moore and P. C. Dorrestein, *Nat Chem Biol*, 2011, **7**, 794-802.
2. K. Severinov, E. Semenova, A. Kazakov, T. Kazakov and M. S. Gelfand, *Molecular Microbiology*, 2007, **65**, 1380-1394.
3. S. J. Pan, W. L. Cheung and A. J. Link, *Protein Expr. Purif.*, 2010, **71**, 200-206.
4. T. A. Knappe, U. Linne, S. Zirah, S. Rebuffat, X. L. Xie and M. A. Marahiel, *J. Am. Chem. Soc.*, 2008, **130**, 11446-11454.
5. T. A. Knappe, U. Linne, L. Robbel and M. A. Marahiel, *Chem. Biol.*, 2009, **16**, 1290-1298.
6. A. Skerra, *Gene*, 1994, **151**, 131-135.
7. W. L. Cheung, S. J. Pan and A. J. Link, *J. Am. Chem. Soc.*, 2010, **132**, 2514-2515.
8. M. Strohalm, D. Kavan, P. Novak, M. Volny and V. Havlicek, *Anal Chem*, 2010, **82**, 4648-4651.