Published in final edited form as:

ACS Synth Biol. 2020 February 21; 9(2): 241–248. doi:10.1021/acssynbio.9b00438.

Heterologous Production of Lasso Peptide Capistrin in a *Burkholderia* host

Sylvia Kunakom, Alessandra S. Eustáquio*

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

Abstract

Burkholderia bacteria are an emerging source of natural products with applications in agriculture and medicine. Heterologous expression of biosynthetic gene clusters can streamline natural product discovery. However, production yields with the commonly used *Escherichia coli* host are usually low. Following the current paradigm that one host does not fit all, we aim to develop a *Burkholderia* host in order to ultimately tap into the biosynthetic potential of *Burkholderia* genomes, which can contain up to 27 biosynthetic gene clusters per genome. Because close phylogenetic relationship is expected to improve odds of success due to compatible gene expression and precursor supply, we tested *Burkholderia* sp. FERM BP-3421, a non-pathogenic isolate previously used to produce natural products at industrial scale. We show here that FERM BP-3421 can produce the model lasso peptide capistrin in yields that are at least 65-fold, and up to 580-fold higher than the previously used *E. coli* host.

Graphical Abstract

*Corresponding author: Tel.: (312) 413-7082. ase@uic.edu.

Supporting Information.

The Supporting Information is available free of charge online.

Biosynthetic gene clusters identified from *Burkholderia* genomes (Figure S1), confirmation of pSK020 vector (Figure S2), capistrin production and growth over time (Figure S3), effect of inducer concentration on capistrin titers (Figure S4), comparison of yields obtained after heterologous expression of lasso peptide gene clusters from Proteobacteria (Figure S5), previously reported heterologous production of lasso peptides (Table S1) (PDF).

Keywords

lasso peptide; capistrucin; natural product; bacteria; heterologous expression

Bacterial natural products are important sources of therapeutically and biotechnologically relevant compounds.¹ However, the large majority requires structural modification to tailor the compound to the desired application and/or alternative means of production that circumvent the native producer and address supply problems.² Although structural modifications and supply issues are often addressed through semi or total synthesis, synthetic biology promises to make a major contribution to the field of natural product discovery and development.³

Bacteria from the *Burkholderiaceae* family of β -Proteobacteria have emerged as a promising source of natural products. At least 66 compound classes that display a range of biological activities have been isolated from *Burkholderiaceae*.⁴ Yet, natural products previously isolated from these Gram-negative bacteria represent only a very small fraction of the biosynthetic potential predicted from the more than one thousand *Burkholderia* genomes available in public databases.⁵ For example, a query of the antiSMASH database⁶ identifies 11,581 biosynthetic gene clusters (BGCs) from 624 *Burkholderia* genomes (Figure S1). Although some level of redundancy is expected in the detected BGCs, it is clear *Burkholderia* remains largely unexplored as a source of natural products.

In order to tap into the biosynthetic potential of *Burkholderia*, a robust and high-yielding heterologous host would be advantageous; not only because many *Burkholderia* are pathogenic to humans, animals and plants, and some are slow-growing endosymbionts, but also to streamline discovery by going from DNA sequence to natural products without having to activate expression in native producers. *Escherichia coli* has been extensively used as a bacterial host. However, production yields obtained with *E. coli* are usually low. As such, the synthetic biology community is moving away from the idea that “one host fits all”. Factors that may influence success of heterologous expression include phylogenetic

relationship between the BGC source and the host, BGC regulatory parts (e.g. promoters and ribosome binding sites), G+C content and codon compatibility, and the intrinsic capacity of the host to produce natural products in high yields.⁷

With the goal of developing a heterologous host tailored for *Burkholderiaceae*, we chose to test *Burkholderia* sp. FERM BP-3421, a non-pathogenic soil isolate.^{8,9} Strain FERM BP-3421 was previously developed to produce native polyketide-nonribosomal peptides of the spliceostatin class of natural products in industrial scale and in up to 6 g/L yields.¹⁰

As a model BGC to test the host with, we chose the BGC encoding biosynthesis of the lasso peptide capistruiin. Our main rationale for choosing the capistruiin BGC was because it had been previously expressed in *E. coli*, allowing us to determine whether our chosen host offers any advantage over the broadly used *E. coli* host. Moreover, lasso peptides, which belong to the ribosomally synthesized and posttranslationally modified peptide (RiPP) family of natural products, are encoded in relatively small BGCs, facilitating DNA synthesis. Most importantly, lasso peptides have attracted significant attention due to their thermal¹¹ and proteolytic¹² stability, diverse biological activities,¹³ and susceptibility to structural modification such as epitope grafting^{14,15} and DNA-encoded compound libraries.¹⁶

Lasso peptides are characterized by a distinctive lariat knot-like topology or 1-rotaxane structure (Figure 1A). The N-terminal α -amino group of amino acid 1 (usually Gly) forms an isopeptide bond with the carboxyl side chain of an Asp or Glu located at positions 7, 8, or 9, while the C-terminal tail threads through the macrolactam ring, resulting in the lasso fold.¹³ Biosynthesis of lasso peptides involves three main steps (Figure 1A). First, the precursor peptide mRNA (from gene A) is translated by the ribosome. The leader sequence is then recognized and cleaved off by a protease (gene B). Finally, a macrocyclase (gene C) activates the carboxyl side chain of an internal Asp/Glu residue in an ATP-dependent manner before catalyzing isopeptide bond formation leading to the macrolactam threaded by the C-terminal tail. Moreover, lasso peptide BGCs may encode a transporter for lasso peptide secretion and immunity.¹³ The distinctive structure of lasso peptides not only confers thermal and proteolytic resistance (attractive characteristics for peptide therapeutics), but also poses synthetic challenges. The first total synthesis of a lasso peptide has only recently (in 2019) been reported with a molar yield of 2.5%.¹⁷

Capistruiin, an antibiotic that targets RNA polymerase,^{18,19} was the first lasso peptide to be discovered by genome mining,²⁰ and the only lasso peptide (and the only RiPP) that has been successfully isolated from *Burkholderia* thus far.²¹ Moreover, the burhizin BGC from the endosymbiont *Mycetohabitans rhizoxinica* (basonym *Burkholderia rhizoxinica* HKI 454) was expressed in *E. coli* and the expected mass was detected by LC/MS; however, low yields precluded isolation and structure elucidation.²² Low yield is a frequent hurdle that stalls RiPP and natural product discovery in general.²³

Herein we demonstrate that *Burkholderia* sp. FERM BP-3421 can produce capistruiin in yields that are at least 65-fold higher, and up to 580-fold higher, under comparable BGC design, than the previously used *E. coli* host (Table S1).²¹

RESULTS AND DISCUSSION

Design of DNA construct.

The vector for expression in *Burkholderia* sp. FERM BP-3421 was designed to contain the four *cap* genes (*capABCD*) and their native ribosome binding sites (RBS) and terminators (Figure 1). No codon modifications were made except for one synonymous modification (GAA to GAG) that was introduced to remove an *Eco*RI restriction site that was subsequently used for cloning. According to the reported codon preferences for *Burkholderia mallei*,²⁴ the synonymous modification introduced is expected to be well tolerated. Transcription was driven from the L-arabinose inducible *araC*/P_{BAD} promoter that had been previously shown to work well in FERM BP-3421,¹⁰ which does not utilize L-arabinose as a carbon source.⁸ The placement of the transcriptional start site of P_{BAD} was 45 bp from the RBS, which was comparable to the location of the native promoter. The *araC*/P_{BAD} fragment was first ligated into the *E. coli*-*Burkholderia* replicative vector pUCP_neo¹⁰ to yield pSK019, after which synthetic *capABCD* was inserted into pSK019 to yield pSK020 (Figure 1B and Figure S2).

Establishment of an electroporation protocol.

The only reported DNA transfer method for FERM BP-3421 is conjugation from *E. coli*.^{10,26} In order to bypass *E. coli* and simplify the DNA transfer process, we aimed to establish an electroporation protocol. We adapted the protocol by Dennis & Sokol previously described for *Pseudomonas*.²⁷ The successful transfer of plasmid DNA pUCP_neo¹⁰ was confirmed by either colony PCR or restriction enzyme digest. Even though we were able to obtain correct clones, our transformation efficiency was initially low (~0.35 transformants/μg of DNA). Modification of the protocol by eliminating the use of magnesium chloride increased electroporation efficiency by 170-fold (60 transformants/μg of pUCP_neo plasmid DNA and 37 transformants/μg of pSK020). Although the electroporation efficiency is not conducive to library generation, it is sufficient for our purposes of BGC expression and natural product discovery/production.

Heterologous production of capistruin.

Burkholderia sp. FERM BP-3421/pSK020 was cultivated in the M20 capistruin optimized medium reported by Knappe *et al.*, using the native producer *B. thailandensis* E264 as positive control.²¹ *B. thailandensis* E264 was cultured at 37-42 °C as reported, whereas FERM BP-3421 was cultivated at 30 °C (the permissive growth temperature of FERM BP-3421 is <35 °C).⁸ To compensate for the decreased temperature, we extended the cultivation time from two to three days, when maximum titers were observed (Figure S3A). Cultures were extracted as reported by Knappe *et al.*²¹ and capistruin production was detected by HPLC by comparison with an authentic standard (Figure 2A). Analysis of the extracts by matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry identified *m/z* 2049 (expected for singly charged species [M+H]⁺) only in FERM BP-3421 containing plasmid pSK020 but not in the empty vector control (Figure 2B). Ion-trap time-of-flight (IT-TOF) mass spectrometry analysis detected *m/z* 1025.23 (expected for doubly charged species [M+2H]²⁺) (Figure 2C). The identity of capistruin was further confirmed by MS/MS analysis (Figure 2D).

Identification of optimal cultivation conditions.

We next cultured FERM BP-3421/pSK020 in the medium previously optimized for high-yield production of polyketide-nonribosomal peptides of the spliceostatin class of natural products. The 2S4G medium contains soy peptone (2%) and glycerol (4%) as cost-effective nitrogen and carbon sources respectively.¹⁰ On the other hand, the main source of carbon and nitrogen in the M20 medium is L-glutamic acid (2%). Growth and relative capistruin titers over time were compared for the two media (Figure S3). The 2S4G medium led to about 4-fold higher titers than the M20 medium, which correlated to the higher cell density obtained with the former (Figure S3). That higher titers in 2S4G medium correlate with higher cell density makes sense because use of the P_{BAD} inducible promoter is expected to decouple transcription from media components.

We next investigated the effect of L-arabinose concentration (Figure 3A) and time of addition (Figure 3B) on capistruin titers. Capistruin production is highest after induction with 100 mM L-arabinose at the time of inoculation (Figure 3A). Higher amounts of inducer did not lead to improvement in titers (Figure S4). Regarding time of addition, the later L-arabinose was added, the lower the yields (Figure 3B).

Deletion of the autologous spliceostatin gene cluster increases production of heterologous capistruin.

Strain FERM BP-3421 produces up to 6 g/L spliceostatins in 2S4G medium, indicating that this organism utilizes a considerable amount of metabolic and gene expression resources toward spliceostatin production.¹⁰ Thus, we next tested whether blocking spliceostatin production would lead to an increase in capistruin titers. For this purpose, plasmid pSK020 was transferred into a spliceostatin-defective mutant strain (*fi9DEF*)²⁶ by electroporation. Extracts were analyzed by HPLC in comparison to the wild-type FERM BP-3421 containing pSK020. At least three independent transformants each were cultivated in both M20 and 2S4G media (Figure 4). Absence of spliceostatin production indeed led to a 4.3-fold increase in capistruin yields. This increase was observed only in the medium that enables spliceostatin production, 2S4G. In M20 medium, in which spliceostatins are not produced, no difference was observed, indicating that the yield improvement observed in 2S4G is indeed tied to the loss in spliceostatin production.

Burkholderia FERM BP-3421 produces more capistruin than previously reported with an *E. coli* host.

To-date, *E. coli* has been the only previously reported heterologous host used for expression of RiPP BGCs from *Burkholderia* sources.²² The merits of *E. coli* include rapid growth rate (20 min doubling time), and excellent genetic tractability. However, *E. coli* does not contain a large number of complex BGCs and is therefore not primed for natural product production. In contrast, *Burkholderia* genomes contain up to 27 BGCs, encoding compounds from diverse biosynthetic classes.⁴ In terms of growth rate, *Burkholderia* (~1 h doubling time) does not lag too far behind *E. coli*. Moreover, *E. coli* has a G+C DNA content (~50%) that is significantly lower than *Burkholderia* (~67%), leading to a potential requirement for codon optimization. Thus, functional and high-yield expression of BGCs from *Burkholderia* in *E. coli* may require rounds of refactoring and metabolic engineering.

Capistrin production after heterologous expression of the unaltered *cap* BGC in *E. coli* (as done here) was reported at yields of 0.2 mg/L by Knappe *et al.*²¹ Pan *et al.*²⁸ were later able to improve yields to 1.6 mg/L by adding an optimized RBS while removing the stem loop region found between *capA* and *capB* (Figure 1B) – the presence of such a terminator between the precursor peptide gene and the genes encoding post-translational modifying enzymes is a common and intuitive arrangement of lasso peptide BGCs^{29,30} that function to attenuate expression of the B and C genes.

Expression of the *cap* BGC containing its native RBSs and intergenic regions in the *Burkholderia* host led to more capistrin than initially reported for *E. coli* (0.2 mg/L)²¹ in all cases analyzed (Figure 4), and up to 65-fold more (13 mg/L for the *frgDEF*⁻ mutant in 2S4G medium). The results obtained in the first *E. coli* paper²¹ are more comparable to ours since we also used an unaltered BGC.

Moreover, during the process of obtaining and analyzing independent transformants to check for reproducibility, we serendipitously isolated one outlier transformant of the wild-type FERM BP-3421 (out of five) that consistently produced 116 mg/L capistrin (average of three independent experiments, each run with triplicate cultures), a 580-fold improvement over *E. coli* (Fig. 4). The heterologous production of lasso peptides at such high yields is unprecedented (Table S1 and Figure S5). The molecular basis behind such high production by the outlier is currently being investigated.

Even though unprecedented capistrin yield (up to 116 mg/L) was obtained with *Burkholderia* sp. FERM BP-3421, the yields we observed are well below the 6 g/L capability of this strain to produce autologous polyketide-nonribosomal peptide spliceostatsins.¹⁰ The two factors most likely to influence the disparate yields in these two cases are toxicity of the final product and biosynthetic class of the natural product.⁷ Capistrin is an antibiotic targeting bacterial RNA polymerase,¹⁹ whereas spliceostatsins have been shown to target the spliceosome,³¹ which is absent in bacteria. Even though the *cap* BGC does encode a transporter expected to confer immunity, it is conceivable that as capistrin production reaches a certain level, export may not be linearly correlated to the amount of transporter gene expression or that transporter expression may not be sufficient. Ultimately though, different biosynthetic classes will have different demands on the host. RiPPs are unique in that the substrate for enzymatic reactions is a protein itself. While for other biosynthetic classes gene expression leads only to protein catalysts, for RiPPs the product of the precursor peptide gene is itself used as substrate, so that the amount of precursor peptide – which is linked to the limit of protein production in cells – will quantitatively dictate yields of the final product.

Conclusions.

In this study, we have shown that *Burkholderia* sp. FERM BP-3421 is a promising host for heterologous production of *Burkholderia* natural products, which we demonstrated by expression of the BGC encoding lasso peptide capistrin. High-yield heterologous production of lasso peptides and other RiPPs remains a challenge. There are only three reported examples of yields in the 10 mg/L range and in all three cases the incoming DNA and the host are from the same bacterial class (Table S1). We envision that the *Burkholderia*

host tested here will serve as a good alternative to *E. coli* towards discovery and production of lasso peptides, other RiPPs, and other biosynthetic classes of compounds from *Burkholderia* and potentially from other β -Proteobacteria sources of natural products.

Methods

Strains, Cultivation conditions, chemicals, and general procedures

Burkholderia sp. FERM BP-3421 (International Patent Organism Depository at the National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan) was routinely cultured in Luria-Bertani (LB) broth or LB agar at 30 °C, unless otherwise stated.^{10,26}

Burkholderia thailandensis E264 (ATCC No. 700388) was routinely cultured in LB broth/agar or nutrient broth at 37 °C, unless otherwise stated. Chemically competent *E. coli* DH5 α (Invitrogen, Cat. No 18265017) was used for plasmid construction and was cultured at 37 °C in LB medium containing appropriate antibiotics. All bacterial strains were preserved in 20% sterile glycerol and stored at –80 °C.

Oligonucleotide primers were synthesized by Sigma-Aldrich. Restriction enzymes and T4 ligase were from New England Biolabs. Phusion High Fidelity DNA polymerase (Thermo Fisher) and Dream Taq polymerase (Sigma) were used in PCR reactions unless stated otherwise. The accuracy of PCR inserts was confirmed by Sanger sequencing, performed by UIC DNA Facility services.

Plasmid DNA isolation was performed by using the Zymo Research plasmid miniprep kit (D4016). Standard molecular biology procedures were carried out according to the manufacturer's instructions.

The capistruin standard (0.2 mg) was a kind gift from Dr. Séverine Zirah (Muséum National d'Histoire Naturelle, Paris, France).

Capistruin gene cluster plasmid construction

The *cap* gene cluster was synthesized by GenScript. The restriction site *KpnI* was inserted upstream of the cognate ribosome binding site and the restriction site *EcoRI* was inserted downstream of the cognate terminator. The *EcoRI* site in *capC* was removed by codon modification, changing GAA to GAG. The *cap* gene cluster was inserted into the *EcoRV* cloning site of pUC57 to yield pUC57_*capBGC*.

For amplification of *araC*P_{BAD}, the PCR reaction (50 μ L) consisted of 10 ng of pIJ790 as template, 0.2 mM of dNTP (Thermo Fisher Scientific), 3% DMSO (Sigma Aldrich), 0.5 μ M of primer pair P3_araC_*HindIII*_f (TAG TCG AAG CTT ACC TGC ATC GAT TTA TTA TGA CAA C, *HindIII* site underlined) and P4_araC_*SaI*_r (AGG TAC GTC GAC CCC AAA AAA ACG GGT ATG GA, *SaI* site underlined), and 0.02 U/ μ L Phusion high-fidelity DNA polymerase (Thermo Fisher Scientific) in Phusion HF buffer supplied with the enzyme. PCR amplifications were performed using a Biorad T100 thermal cycler. The thermal cycling conditions were as follows: 30 s at 98 °C; 30 cycles of 98 °C for 10 s, 62 °C for 30 s, and 72 °C for 45 s; and a final hold for 5 min at 72 °C. The 1220-bp *araC*P_{BAD} PCR product was digested with *HindIII* and *SaI*, and ligated into the same sites of

pUCP_neo¹⁰ to yield pSK019. Selection of *E. coli* DH5 α clones was performed on LB agar containing kanamycin (50 μ g/mL), and clones were verified by restriction digest and Sanger sequencing.

To construct pSK020, pUC57_capBGC was digested with *KpnI* and *EcoRI* to release the *cap* gene cluster. The *cap* gene cluster was subsequently ligated into the same sites of pSK019 to yield pSK020, which was confirmed by restriction digest and gel electrophoresis (Fig. S1).

Electroporation of *cap* BGC in FERM BP-3421.

5 mL of LB were inoculated with 20 μ L of a FERM BP-3421 cryo stock and incubated at 30 $^{\circ}$ C, 180 rpm, for 24 hours. 1 mL of this overnight culture was added to 50 mL of LB broth in a 250-ml Erlenmeyer flask, and incubated at 30 $^{\circ}$ C, 180 rpm, until the OD₆₀₀ reached ~0.6 (approx. 4 hours). The cells were then harvested by centrifugation at 4 $^{\circ}$ C, 4000 rpm, for 5 min. The cell pellet was immediately washed twice with ice-cold 1 mM HEPES buffer pH 7.0 (Sigma Aldrich), the first wash with 50 mL and the second with 25 mL buffer. After the second wash, the cell pellet was resuspended in the remaining drops of buffer (yielding ~150 μ L of electrocompetent cell suspension). Electrocompetent FERM BP-3421 cells (50 μ L) were electroporated with up to 500 ng of plasmid DNA (pUCP_neo or pSK020) in a 0.2-cm ice-cold electroporation cuvette using a Bio-Rad Micropulser set to 2.5 kV. Obtained time constants ranged from 4.8 ms to 5.5 ms. Immediately after the pulse, 950 μ L of ice-cold LB medium was added and the cell suspension was transferred to a 1.5-ml tube and incubated horizontally at 30 $^{\circ}$ C, 180 rpm for 1 hr. The cells were then plated on LB agar containing kanamycin 500 μ g/mL to select for the incoming plasmid. Plates were incubated at 30 $^{\circ}$ C for 3-5 days.

Transformants were confirmed by colony PCR to screen for the *cap* genes, and DNA isolation and subsequent restriction enzyme digest analyses. The primer pairs used for screening of pSK020 clones were as follows: (1) neomycin/kanamycin resistance gene: P_neo_f (GCA AGG GCT GCT AAA GGA AG) and P_neo_r (ACG GAA ATG TTG AAT ACT CAT ACT C); (2) *capA* gene: P1_cap_f (AGT CCA CAT TGA TTA TTT G) and P1_cap_r (GAT ATC AAC GGC TTC AAT); and (3) *capB* gene: P2_cap_f (TCG ATT GAG TAA CGC TTC ATG) and P2_cap_r (TCC AGC ACG AGG AGG TAC). The PCR reactions (20 μ L) consisted of either 1 μ L of a colony suspension or ~10 ng of pSK020 as template, 0.2 mM of dNTP, 0.5 μ M of each primer, 3% DMSO, and 1.25 U DreamTaq DNA polymerase (Thermo Fisher Scientific) in reaction buffer supplied with the enzyme. For primer pair 1, the thermal cycling conditions were as follows: 2 min at 95 $^{\circ}$ C; 25 cycles of 95 $^{\circ}$ C for 30 s, 55 $^{\circ}$ C for 20 s, and 72 $^{\circ}$ C for 1 min; and a final hold for 5 min at 72 $^{\circ}$ C. For primer pair 2, the thermal cycling conditions were as follows: 2 min at 95 $^{\circ}$ C; 30 cycles of 95 $^{\circ}$ C for 30 s, 41 $^{\circ}$ C for 30 s, and 72 $^{\circ}$ C for 1 min; and a final hold for 5 min at 72 $^{\circ}$ C. For primer pair 3, the thermal cycling conditions were as follows: 2 min at 95 $^{\circ}$ C; 30 cycles of 95 $^{\circ}$ C for 30 s, 55 $^{\circ}$ C for 30 s, and 72 $^{\circ}$ C for 1 min; and a final hold for 5 min at 72 $^{\circ}$ C.

Burkholderia cultivation for capistruin production

***B. thailandensis* E264.**—Seed cultures were prepared by inoculating 5 mL LB with 20 μ L of a cryo stock of strain E264 and incubated at 37 $^{\circ}$ C, 180 rpm, 24 hours. Note that we

start the seed culture directly from frozen stocks rather than from single colonies for reproducibility purposes. Also note that this cultivation procedure follows inoculation method 3 recently described by Kragh *et al.*³² Aliquots of the seed culture were used to inoculate 50 mL of M20 medium²¹ [L-glutamic acid (Sigma Aldrich) 20 g/L, L-alanine (Sigma Aldrich) 0.2 g/L, sodium citrate tribasic dehydrate (Sigma Aldrich) 1.0 g/L, disodium hydrogen phosphate (Sigma Aldrich) 20 g/L, potassium chloride (Sigma Aldrich) 0.5 g/L, sodium sulfate anhydrous (Fisher Scientific) 0.5 g/L, magnesium chloride (Sigma Aldrich) 0.2 g/L, calcium chloride (Sigma Aldrich) 0.0076 g/L, iron (II) sulfate heptahydrate (MP Biomedicals) 0.01 g/L, manganese sulfate heptahydrate (Macron Chemicals) 0.0076 g/L, pH 7.0] until an OD₆₀₀ of ~0.01 was reached (~200–400 µL seed). Cultivations were performed in 250-mL Erlenmeyer flasks. Production cultures were incubated at 42 °C, 180 rpm, for one to three days.

B. sp. FERM BP-3421 cultivation in M20 production medium.—Cultivation in M20 medium was performed as described above for strain E264 except that the temperature used was 30 °C. Moreover, kanamycin at 500 µg/mL was used to maintain the selection of pUCP_neo or pSK020 in seed and production cultures. 100 mM of L-arabinose (Sigma Aldrich) was used for induction (Figure 4).

B. sp. FERM BP-3421 cultivation in 2S4G production medium.—5 mL of seed medium [peptone (Bacto Difco) 1% w/v, yeast extract (Bacto Difco) 0.5% w/v, and sodium chloride (Sigma Aldrich) 0.5% w/v]¹⁰ was inoculated with 20 µL of a cryo stock of strain FERM BP-3421 containing either pUCP_neo (negative control) or pSK020 and incubated at 30 °C, 180 rpm, for 1–2 days. The cells were subcultured to OD₆₀₀ of 0.01 into a 250 mL Erlenmeyer flask containing 50 mL of 2S4G production medium [glycerol (Fisher Scientific) 4% w/v, Hy-soy peptone (Kerry Bioscience) 2% w/v, ammonium sulfate (Sigma Aldrich) 0.2% w/v, magnesium sulfate (Sigma Aldrich) 0.006% w/v, and calcium carbonate (Acros Organics) 0.2% w/v]. Cultures were supplemented with kanamycin 500 µg/mL and 0–100 mM of L-arabinose (Sigma Aldrich) when appropriate, and incubated at 30 °C, 180 rpm, 1–5 days.

Capistruin extraction and analysis

Cultures were harvested by centrifugation (10 min, 4000 rpm). The cell pellet was discarded and the supernatant was mixed with 0.5% (5 g/L) XAD-16 resin (Alfa Aesar), followed by agitation for 1 hr at 500 rpm. Alternatively, aliquots of production culture (5 mL) were processed with 5% of XAD-16 (0.25 g). The resin was separated from the medium via centrifugation (5 min, 4000 rpm) and washed with one volume of water. Water was removed via centrifugation (5 min, 4000 rpm) and capistruin was extracted from the resin in one volume of methanol (e.g. 5 mL methanol for 5-mL cultures) by shaking for 30 min at 500 rpm. After centrifugation (5 min, 4000 rpm) or filtration, the methanol extract was dried under reduced pressure. The crude extracts were dissolved in methanol for HPLC analyses.

HPLC analyses were performed on an Agilent 1260 Infinity system equipped with a Kinetex C₁₈ column (150 x 4.6 mm, 5 µm particle size, 100 Å pore size, Phenomenex). Elution was performed in water/0.1% trifluoroacetic acid (TFA) [v/v] (solvent A) and acetonitrile

(solvent B) at a flow rate of 1 mL/min with a solvent gradient as follows: initial hold at 40% B for 2 min, linear gradient from 40% to 55% B for 23 min, linear gradient from 55% to 100% for 5 min, and terminal hold at 100% B for 5 min. The detection wavelength range was 200-600 nm and chromatograms were extracted at $\lambda = 210$ nm.

MALDI-TOF MS measurements were performed on an Autoflex Speed LRF mass spectrometer (Bruker Daltonics) equipped with a Smartbeam™ Laser. Spectra that recorded the detection of capistruin was performed in positive reflectron mode (1000 total shots; ion source voltage 19 kV; ion source 2 voltage: 16.61 kV; reflector 1 voltage 20.86 kV; reflector 2 voltage: 9.71 kV; lens voltage 8.3 kV, mass range 200-3500 Da; pulsed ion extraction 200 ns). Calibration was performed externally using the Bruker Daltonics peptide calibration standard: Angiotensin II (1046.5418 Da), Angiotensin I (1296.6848 Da), Substance P (1347.7354 Da), Bombesin (1619.8223 Da), ACTH clip 1-17 (2093.0862 Da), ACTH clip 18-39 (2465.1983 Da). Data acquisitions were performed using FlexControl software v. 3.4.135.0 (Bruker Daltonics) and FlexAnalysis software v. 3.4.

Low resolution MS experiments were analyzed on a Finnigan LCQ Advantage MAX Mass spectrometer (Thermo Electro Corporation) with a Hewlett Packard series 1050 HPLC, equipped with a Kinetex® 5 μ m C18 100 Å-column (150 $\times$ 4.6 mm, 5 μ m particle size, Phenomenex®). Solvent A was 0.02% [v/v] formic acid in water, and solvent B was 0.02% [v/v] formic acid in methanol. The solvent gradient was set as follows: initial hold at 10% B for 2 min, linear gradient from 10% to 100% B within 38 min, hold at 100% B for 4 min with a flow rate of 1.0 mL/min. The detection wavelength was 210 nm. The detection mass range was 200-2000 Da in positive mode. Mass spectrometry analyses was performed when the sample was eluting from 2–30 minutes. Spectra that recorded the detection of capistruin was obtained in positive reflectron mode.

Capistruin isolation and yield determination

Production titers were calculated based on a standard curve of authentic capistruin kindly provided by Dr. Séverine Zirah (Muséum National d'Histoire Naturelle, Paris, France).²¹ To confirm that the estimated yields based on the calibration curve were accurate, we performed capistruin isolation using the high-yielding outlier clone. For this purpose, a crude extract from a 100-ml culture (0.24 g) was resuspended in methanol and purified by HPLC as follows. A semi-preparative Kinetex® C18 -column (250 $\times$ 10.0 mm, 5 μ m particle size, 100 Å pore size, Phenomenex®) was used and elution was performed in 0.1% aqueous formic acid [v/v] (solvent A) and acetonitrile (solvent B) at a flow rate of 2.5 mL/min with a solvent gradient as follows: 10 to 20% B for 5 min, 20% to 55% B for 23 min, 55% to 10% B for 1 min, and terminal hold at 10% B for 7 min. The detection wavelength range was 200-600 nm and chromatograms were extracted at $\lambda = 210$ nm. The fraction containing capistruin was dried under reduced pressure to yield 16 mg, 80% pure by UV. After correcting for purity the isolated yield would be ~128 mg/L which is in agreement with the 116 mg/L average we initially estimated based on a calibration curve.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank the International Patent Organism Depository of the National Institute of Technology and Evaluation (Japan) for strain FERM BP-3421, Edmund Graziani and Mark Noe (Pfizer, Groton, CT) for the *fr9DEF⁻* mutant of FERM BP-3421, Dr. Séverine Zirah (Muséum National d'Histoire Naturelle, Paris, France) for the capistrin standard, Dr. Laura Sanchez (University of Illinois at Chicago) for mass spectrometry access and discussions, and Christian San Andres for help with *Burkholderia* cultivation and metabolite extraction. Financial support for this work was provided by the National Center for Advancing Translational Sciences, National Institutes of Health (NIH), under grant KL2TR002002 (to A.S.E.), by a UIC Provost Graduate Award and the Charles Wesley Petranek Memorial Scholarship (to S.K.), and by startup funds from the Department of Pharmaceutical Sciences and the Center for Biomolecular Sciences, University of Illinois at Chicago (to A.S.E.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

- (1). Cragg GM, and Newman DJ (2013) Natural products: A continuing source of novel drug leads. *Biochim. Biophys. Acta - Gen. Subj* 1830, 3670–3695.
- (2). Harvey AL, Edrada-Ebel R, and Quinn RJ (2015) The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov* 14, 111–129. [PubMed: 25614221]
- (3). Kunakom S, and Eustáquio AS (2019) Natural products and synthetic biology: Where we are and where we need to go. *mSystems* 4, e00113–19. [PubMed: 31120026]
- (4). Kunakom S, and Eustáquio AS (2019) *Burkholderia* as a source of natural products. *J. Nat. Prod* 82, 2018–2037. [PubMed: 31294966]
- (5). Winsor GL, Khaira B, Van Rossum T, Lo R, Whiteside MD, and Brinkman FSL (2008) The *Burkholderia* genome database: Facilitating flexible queries and comparative analyses. *Bioinformatics* 24, 2803–2804. [PubMed: 18842600]
- (6). Blin K, Pascal Andreu V, de los Santos ELC, Del Carratore F, Lee SY, Medema MH, and Weber T (2019) The antiSMASH database version 2: A comprehensive resource on secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res.* 47, D625–D630. [PubMed: 30395294]
- (7). Zhang MM, Wang Y, Ang EL, and Zhao H (2016) Engineering microbial hosts for production of bacterial natural products. *Nat. Prod. Rep* 33, 963–987. [PubMed: 27072804]
- (8). Nakajima H, Sato B, Fujita T, Takase S, Terano H, and Okuhara M (1996) New antitumor substances, FR901463, FR901464 and FR901465. I. Taxonomy, fermentation, isolation, physico-chemical properties and biological activities. *J. Antibiot* 49, 1196–1203. [PubMed: 9031664]
- (9). He H, Ratnayake AS, Janso JE, He M, Yang HY, Loganzo F, Shor B, O'Donnell CJ, and Koehn FE (2014) Cytotoxic spliceostatin from *Burkholderia* sp. and their semisynthetic analogues. *J. Nat. Prod* 77, 1864–1870. [PubMed: 25098528]
- (10). Eustáquio AS, Chang L-P, Steele GL, O'Donnell CJ, and Koehn FE (2016) Biosynthetic engineering and fermentation media development leads to gram-scale production of spliceostatin natural products in *Burkholderia* sp. *Metab. Eng* 33, 67–75. [PubMed: 26620532]
- (11). Hegemann JD (2019) Factors governing the thermal stability of lasso peptides. *ChemBioChem.* Epub 7 27, 2019. DOI: 10.1002/cbic.201900364.
- (12). Martin-Gómez H, and Tulla-Puche J (2018) Lasso peptides: Chemical approaches and structural elucidation. *Org. Biomol. Chem.* 16, 5065–5080. [PubMed: 29956707]
- (13). Hegemann JD, Zimmermann M, Xie X, and Marahiel MA (2015) Lasso peptides: An intriguing class of bacterial natural products. *Acc. Chem. Res* 48, 1909–1919. [PubMed: 26079760]
- (14). Knappe TA, Manzenrieder F, Mas-Moruno C, Linne U, Sasse F, Kessler H, Xie X, and Marahiel MA (2011) Introducing lasso peptides as molecular scaffolds for drug design: Engineering of an integrin antagonist. *Angew. Chem* 50, 8714–8717. [PubMed: 21812076]
- (15). Hegemann JD, De Simone M, Zimmermann M, Knappe TA, Xie X, Di Leva FS, Marinelli L, Novellino E, Zahler S, Kessler H, et al. (2014) Rational improvement of the affinity and selectivity of integrin binding of grafted lasso peptides. *J. Med. Chem* 57, 5829–5834. [PubMed: 24949551]

- (16). Pan SJ, and Link AJ (2011) Sequence diversity in the lasso peptide framework: Discovery of functional microcin J25 variants with multiple amino acid substitutions. *J. Am. Chem. Soc* 133, 5016–5023. [PubMed: 21391585]
- (17). Chen M, Wang S, and Yu X (2019) Cryptand-imidazolium supported total synthesis of the lasso peptide BI-32169 and its D-enantiomer. *Chem. Commun* 55, 3323–3326.
- (18). Braffman NR, Piscotta FJ, Hauver J, Campbell EA, Link AJ, and Darst SA (2019) Structural mechanism of transcription inhibition by lasso peptides microcin J25 and capistrui. *Proc. Natl. Acad. Sci. U. S. A* 116, 1273–1278. [PubMed: 30626643]
- (19). Kuznedelov K, Semenova E, Knappe TA, Mukhamedyarov D, Srivastava A, Chatterjee S, Ebright RH, Marahiel MA, and Severinov K (2011) The antibacterial threaded-lasso peptide capistrui inhibits bacterial RNA polymerase. *J. Mol. Biol* 412, 842–848. [PubMed: 21396375]
- (20). Cheung-Lee WL, and Link AJ (2019) Genome mining for lasso peptides: Past, present, and future. *J. Ind. Microbiol. Biotechnol* 46, 1371–1379. [PubMed: 31165971]
- (21). Knappe TA, Linne U, Zirah S, Rebuffat S, Xie X, and Marahiel MA (2008) Isolation and structural characterization of capistrui, a lasso peptide predicted from the genome sequence of *Burkholderia thailandensis* E264. *J. Am. Chem. Soc* 130, 11446–11454. [PubMed: 18671394]
- (22). Hegemann JD, Zimmermann M, Zhu S, Klug D, and Marahiel MA (2013) Lasso peptides from Proteobacteria: Genome mining employing heterologous expression and mass spectrometry. *Biopolymers* 100, 527–542. [PubMed: 23897438]
- (23). Zhang Y, Chen M, Bruner SD, and Ding Y (2018) Heterologous production of microbial ribosomally synthesized and post-translationally modified peptides. *Front. Microbiol* 9, 1801. [PubMed: 30135682]
- (24). Zhao S, Zhang Q, Chen Z, Zhao Y, Z. J. (2007) The factors shaping synonymous codon usage in the genome of *Burkholderia mallei*. *J. Genet Genomics* 34, 362–372. [PubMed: 17498635]
- (25). Greenfield L, Boone T, and Wilcox G (1978) DNA sequence of the *araBAD* promoter in *Escherichia coli* B/r. *Proc. Natl. Acad. Sci. U. S. A* 75, 4724–4728. [PubMed: 368797]
- (26). Eustáquio AS, Janso JE, Ratnayake AS, O'Donnell CJ, and Koehn FE (2014) Spliceostatin hemiketal biosynthesis in *Burkholderia* spp. is catalyzed by an iron/ α -ketoglutarate-dependent dioxygenase. *Proc. Natl. Acad. Sci. U. S. A* 111, E3376–E3385. [PubMed: 25097259]
- (27). Dennis JJ, and Sokol PA (1995) Electroporation of *Pseudomonas*. In *Electroporation Protocols for Microorganisms* Methods in Molecular Biology (Nickoloff JA, Ed.), pp 125–134, Humana Press, New Jersey.
- (28). Pan SJ, Rajniak J, Maksimov MO, and Link AJ (2012) The role of a conserved threonine residue in the leader peptide of lasso peptide precursors. *Chem. Commun.* 48, 1880–1882.
- (29). Severinov K, Semenova E, Kazakov A, Kazakov T, and Gelfand MS (2007) Low-molecular-weight post-translationally modified microcins. *Mol. Microbiol.* 65, 1380–1394. [PubMed: 17711420]
- (30). Genilloud O, Moreno F, and Kolter R (1989) DNA sequence, products, and transcriptional pattern of the genes involved in production of the DNA replication inhibitor microcin B17. *J. Bacteriol* 171, 1126–1135. [PubMed: 2644225]
- (31). Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, Ishigami K, Watanabe H, Kitahara T, Yoshida T, Nakajima H, et al. (2007) Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat. Chem. Biol* 3, 576–583. [PubMed: 17643111]
- (32). Kragh KN, Alhede M, Rybtke M, Stavnsberg C, Jensen PØ, Tolker-Nielsen T, Whiteley M, and Bjarnsholt T (2017) The inoculation method could impact the outcome of microbiological experiments. *Appl. Environ. Microbiol* 84, e02264–17.

Figure 1. Capistruin biosynthesis and *cap* BGC construct

(A) Generic lasso peptide biosynthesis using capistruin as example. The A gene encodes the precursor peptide that consists of the leader (light blue) and core (dark blue) sequences. The B gene encodes a protease that cleaves off the leader sequence. The C protein catalyzes the formation of the macrolactam ring.¹³ (B) DNA construct used for *cap* BGC expression. The native promoter was replaced with the *araC*/ P_{BAD} promoter.¹⁰ The sequence of the promoter is annotated in red as reported by Greenfield *et al.*²⁵ The distance between the transcriptional start site and the ribosome binding site is 45 bp. The synthesized construct included the native ribosome binding sites (RBS), terminators, and intergenic regions. The asterisk indicates the site at which a codon was modified to remove an *EcoRI* restriction site in *capC*. Kan^R, kanamycin resistance gene, *oriT*, origin of transfer.

Figure 2. Production of capistruin in *Burkholderia* sp. FERM BP-3421.

(A) High performance liquid chromatography analysis of extracts from cultures grown in the previously reported M20 capistrin production medium. The arrow indicates the capistrin peak identified using an authentic standard. (B) MALDI-TOF analysis of extracts from cultures grown in the M20 capistrin medium. The m/z signals of 2049.43 Da and 2071.52 Da in FERM BP-3421/pSK020 correspond to proton and sodium adducts, respectively. (C) Ion-trap MS analysis of FERM BP-3421/pSK020. Detection of doubly charged capistrin species at m/z 1025.23 Da $[M+2H]^{2+}$ and 1036.25 $[M+H+Na]^{2+}$. (D) Ion-trap MS² analysis of m/z 1025.23 from panel C. Fragmentation m/z values that coincide with the literature²¹ are highlighted in red.

Figure 3. Effect of inducer concentration and time of addition on capistrain titers. (A) Effect of inducer (L-arabinose) concentration on capistrain production in 2S4G medium. For this experiment, the inducer was added at the time of inoculation (day 0). (B) Effect of time of addition of inducer (L-arabinose) on capistrain production in 2S4G medium. The inducer concentration used in this experiment was 50 mM. Experiments were performed in triplicates. Error bars indicate standard deviation.

Figure 4. Effect of deletion of the autologous spliceostatin BGC on capistrain titers.

Comparison of capistrain titers obtained using either FERM BP-3421 wild-type or the *fr9DEF*⁻ spliceostatin-defective mutant as hosts, and after cultivation in two different media, M20 and 2S4G. Results from four independently obtained transformants of the wild-type strain are shown in dark blue, whereas results obtained with three independent transformants of the *fr9DEF*⁻ spliceostatin-defective mutant are shown in red. In addition, one high-producing, “wild-type” outlier was obtained and is shown in light blue (average of triplicate experiments each run with triplicate flasks).