HHS Public Access

Author manuscript

J Am Chem Soc. Author manuscript; available in PMC 2022 April 21.

Published in final edited form as:

J Am Chem Soc. 2021 April 21; 143(15): 5917–5927. doi:10.1021/jacs.1c01452.

Cell-Free Biosynthesis to Evaluate Lasso Peptide Formation and Enzyme-Substrate Tolerance

Yuanyuan Si[1,2], **Ashley M. Kretsch**[1,2], **Laura M. Daigh**[1,2], **Mark J. Burk**[4,*], **Douglas A. Mitchell**[1,2,3,*]

[1]Carl R. Woese Institute for Genomic Biology, University of Illinois, Urbana, Illinois, 61801, United States of America

[2]Department of Chemistry, University of Illinois, Urbana, Illinois, 61801, United States of America

[3]Department of Microbiology, University of Illinois, Urbana, Illinois, 61801, United States of America

[4]Lassogen, Inc., San Diego, 92121, California, United States of America

Abstract

Lasso peptides are ribosomally synthesized and post-translationally modified peptide (RiPP) natural products that display a unique lariat-like, threaded conformation. Owing a locked three-dimensional structure, lasso peptides can be unusually stable towards heat and proteolytic degradation. Some lasso peptides have been shown to bind human cell-surface receptors and exhibit anticancer properties, while others display antibacterial or antiviral activities. All known lasso peptides are produced by bacteria and genome-mining studies indicate that lasso peptides are a relatively prevalent class of RiPPs; however, the discovery, isolation, and characterization of lasso peptides are constrained by the lack of an efficient production system. In this study, we employ a cell-free biosynthesis (CFB) strategy to address longstanding challenges associated with lasso peptide production. We report the successful use of CFB for the formation of an array of sequence-diverse lasso peptides that include known examples as well as a new predicted lasso peptide from *Thermobifida halotolerans*. We further demonstrate the utility of CFB to rapidly generate and characterize multisite precursor peptide variants to evaluate the substrate tolerance of the biosynthetic pathway. By evaluating more than 1,000 randomly chosen variants, we show that the lasso-forming cyclase from the fusilassin pathway is capable of producing millions of sequence-diverse lasso peptides via CFB. These data lay a firm foundation for the creation of large lasso peptide libraries using CFB to identify new variants with unique properties.

Graphical Abstract

[*]corresponding authors: Phone: 1-217-333-1345; douglasm@illinois.edu, Address: 600 South Mathews Avenue, 361 Roger Adams Laboratory, Urbana, Illinois, 61801; Phone: 1-858-362-8567; mburk@lassogen.com, Address: 4757 Nexus Center Drive, San Diego, California, 92121.

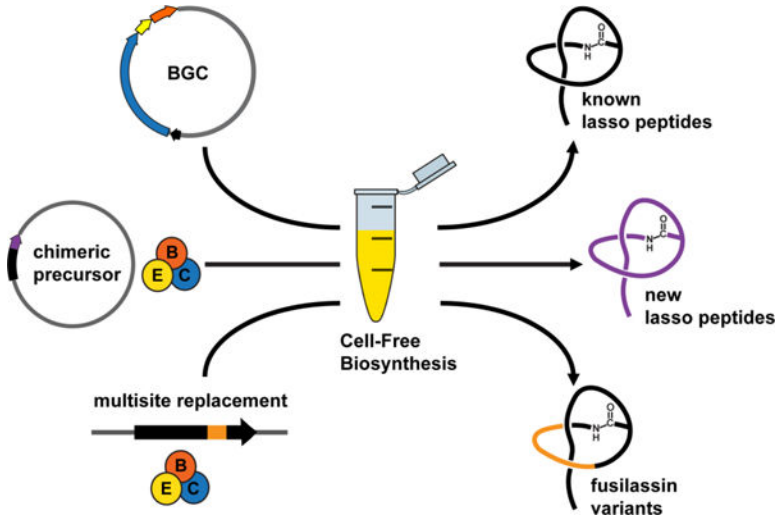
Supporting Information.

Experimental methods and supporting tables and figures (PDF)

Supplemental Dataset 1 (XLSX)

Conflict of Interest Statement

M.J.B and D.A.M are co-founders of Lassogen, Inc.



Introduction

Lasso peptides were first identified in the early 1990s and today comprise a large and growing class of ribosomally synthesized and post-translationally modified peptides (RiPPs).[1,2] Similar to other RiPP classes, lasso peptides are formed by conversion of precursor peptides that consist of a N-terminal leader region and a C-terminal core region.[3] After translation of the linear precursor peptide, lasso peptide biosynthesis begins with recognition of the leader region by a RiPP precursor recognition element (RRE) which guides the biosynthetic enzymes to the substrate.[4] The leader region is then proteolytically removed by a peptidase homologous to transglutaminase (Protein Family PF13471). Subsequently, an ATP-dependent lasso cyclase, homologous to asparagine synthetase (PF00733), catalyzes macrolactam formation between the N-terminus and the carboxyl side chain of a Glu or Asp residue within the core region.[1,2,5] To date, all characterized, naturally occurring lasso peptides contain 7–9 residues within the macrolactam ring region, and an unnatural 10-mer variant of a natural 9-mer lasso peptide has recently been reported.[6] During lasso peptide biosynthesis, the substrate peptide must be “pre-folded” with the C-terminal tail passing through the plane of the incipient ring, resulting in a lariat-like tertiary structure after macrolactam formation.[7] The threaded conformation, as opposed to an unthreaded “branched-cyclic” conformation, is maintained by a large, steric locking residue located below the ring (tail), which is occasionally further stabilized by disulfide bonds. An additional conformational constraint is sometimes provided by a large residue above the plane of the ring (loop).[8–10] Additional post-translational modifications have been observed in lasso peptides, such as glycosylation, phosphorylation, acetylation, deimination, methylation, epimerization, and hydroxylation.[1,11]

The distinctive and rigid globular shape of lasso peptides typically imparts substantial resistance to thermal and proteolytic degradation,[2,12] which are unusual traits for a peptide. Although nearly 80 lasso peptides have been described, less than 30 have a reported biological activity.[13] Many of the known activities involve antibacterial properties that entail binding a variety of distinct targets, including RNA polymerase,[14] prolyl endopeptidase,[15]

lipid II,[16] mycobacterial ClpC1 ATPase,[17] and blocking *fsr* quorum-sensing.[8] Other lasso peptides, discovered by targeted screening, antagonize various human GPCRs, such as the glucagon[18] and endothelin B receptors.[19] Additionally, anti-HIV[20] and anticancer[21] activities have been described.

Previous studies have explored the substrate tolerance of several lasso peptide biosynthetic enzymes in an effort to gain access to greater structural diversity and evaluate the potential of lasso peptides for medical applications. For example, multiple amino acid substitutions and systematic structure-activity analysis of microcin J25 found that the biosynthetic enzymes tolerated amino acid substitutions at over 80% of the lasso peptide residues.[22–24] Epitope grafting within the loop region endowed microcin J25 with integrin receptor-binding activity, whereas the natural scaffold was inactive.[25] Furthermore, an earlier study found that the core region of the lasso peptide citrulassin A (D8E) could be fused to the FusA leader region to create a chimeric precursor peptide. The fusilassin leader peptidase (FusB) and lasso cyclase (FusC) successfully produced the citrulassin A variant (D8E), which suggested the fusilassin pathway may be broadly tolerant of alternative substrates, although wildtype citrulassin A (containing Asp8 for macrolactam formation) was not tolerated.[6] Finally, unnatural amino acids have been incorporated into several lasso peptides, including capistruin[26] and microcin J25.[27]

The increase in publicly available genomes and availability of the open-access bioinformatic tool RODEO (Rapid ORF Description & Evaluation Online)[28] enabled the prediction of over 3,000 high-confidence lasso biosynthetic gene clusters (BGCs) in the NCBI database (as of 2018).[6] Despite the thousands of lasso peptides residing in databases, the isolation and characterization of new lasso peptides have been hindered by the lack of an efficient production method. Historically, lasso peptides have been isolated from the native producer,[19,29] which can suffer from slow cellular growth and low yields. More recently, lasso peptides have been produced through heterologous expression of the BGC.[30,31] However, heterologous expression requires time-consuming molecular cloning and growth steps and can be hampered by pathway-host incompatibilities,[32–35] giving rise to poor gene expression, pathway enzyme inactivity, and product toxicity.

Cell-free biosynthesis (CFB) offers an orthogonal, rapid, and high-throughput strategy to overcome several limitations of existing approaches for lasso peptide production. CFB reactions involve three components: cell extract, energy mix, and a DNA template. The crude cell extract provides the cellular machinery for *in vitro* transcription and translation of the DNA templates.[36] CFB eliminates the need for DNA transformation and protein purification, a historical challenge for lasso peptide enzymes.[6,37,38] In principle, CFB shortens the time and effort required to advance from gene sequence to chemical product. Also, the use of modular DNA templates in CFB allows straightforward access to precursor peptide variants,[39] streamlining the generation of lasso peptide variant libraries. Armed with an efficient production system, understanding the principles that govern lasso peptide formation and substrate tolerance could be more readily discerned, allowing for the future customization of lasso peptides with specific biological properties.

Cell-free protein synthesis is well-established for single protein production and numerous examples of high yielding proteins are available.[40] In contrast, there are few examples of CFB applied to small molecules and natural products, which require the production of multiple active enzymes to deliver a single product of interest. Some examples include 2,3-butanediol,[41] diverse glycan motifs,[42] and a cyclic dipeptide produced from two non-ribosomal peptide synthetases.[43] Several mature RiPPs have been produced using cell-free methods, such as goadsporin,[44] nisin,[45] thiocillin,[46] lactazole[47]. In this study, we evaluated CFB methodology to produce a diverse array of lasso peptides, including known examples and novel lasso peptides predicted from genomic data. Furthermore, we expanded the utility of CFB to rapidly interrogate the substrate tolerance of the fusilassin biosynthetic enzymes. To enable this work, we used CFB to create a randomized, multisite replacement library of the fusilassin precursor peptide. Our screening and sequencing results indicated that 1.85–2.05 million variants of FusA (at a 95% confidence interval) were tolerated by FusC in the ring region library of fusilassin. These data considerably expand the sequence-structure space accessible to lasso peptides and raises their potential as designer ligands for biotechnological applications.

Results and Discussion

Cell-free biosynthesis of known lasso peptides

In this study, we evaluated cell-free biosynthesis (CFB) as a potential solution to overcome longstanding challenges in lasso peptide production. To establish proof-of-concept for the utility of CFB, we first tested four previously reported lasso peptides whose BGCs were already cloned and available in our laboratory: burhizin (*Burkholderia rhizoxinica*),[31,34] capistruin (*Burkholderia thailandensis*),[48] fusilassin (synonymous with fuscanodin, *Thermobifida fusca*),[6,37] and cellulassin (*Thermobifida cellulosilytica*).[37] For burhizin and capistruin, a plasmid bearing the complete BGC was used as the DNA template. This included the genes encoding the precursor peptides (*burA/capA*), RRE-leader peptidase fusion proteins (*burB/capB*), lasso cyclases (*burC/capC*), and ABC transporters (*burD/capD*, Figures 1, S1). For fusilassin, a previously reported plasmid that fuses maltose-binding protein (MBP) to the N-terminus of the precursor peptide (*fusA*) was used in conjunction with a second plasmid (pACYC) encoding the lasso cyclase (*fusC*), RRE (*fusE*), and leader peptidase (*fusB*).[6] For cellulassin, the BGC with the native gene organization: MBP-tagged precursor peptide (*celA*), followed by the lasso cyclase (*celC*), RRE (*celE*), and leader peptidase (*celB*) (Figures 1, S1) was used as the DNA template. Each plasmid was purified by maxi-prep to achieve high DNA purity and concentration. Extracts for CFB reactions were prepared from *E. coli* BL21(DE3) cells (Supplemental Methods). After addition of energy mix and appropriate plasmids, the CFB reactions were allowed to proceed for 16 h prior to analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). In each case, we readily observed the expected product mass (Figure 1). Peptide identity was corroborated by high-resolution and tandem MS (HRMS/MS, Figures S2–4), and the positions of the macrolactam linkages of burhizin, capistruin, and fusilassin were consistent with previous reports (Gly1-Glu8, Gly1-Asp9, and Trp1-Glu9, respectively).[6,31,49] Tandem MS analysis of cellulassin had not been previously

reported; however, our data confirmed the expected macrolactam linkage between Trp1-Glu9 (Figure S5).

Encouraged by the production of four distinct lasso peptides by CFB, we sought to determine if multiple lasso peptides could be produced in a single CFB reaction, as this would render the method more attractive for library generation. We therefore carried out CFB to simultaneously produce capistruin, burhizin, and fusilassin (from a total of four plasmids). All three lasso peptides were detected by MALDI-TOF-MS with the relative production level of burhizin correlating with the amount of plasmid supplied (Figure S6).

To evaluate if lasso peptides produced by CFB were in the native threaded conformation, as opposed to the unthreaded, "branched-cyclic" conformation, we treated burhizin, fusilassin, and cellulassin with carboxypeptidase Y. All three were highly resistant to digestion, showing either no C-terminal trimming (burhizin) or low-efficiency removal of the last one or two residues, consistent with a lasso-like conformation (Figure S7). After prolonged heating (2 h at 95 °C), burhizin (eight-residue ring) retained protease resistance[31] while fusilassin and cellulassin (nine-residue rings) appeared to unthread, as evidenced by efficient carboxypeptidase Y digestion (Figure S7). This result was consistent with a previous report showing fusilassin (fuscanodin) adopted a branched-cyclic form after extended exposure to organic solvent.[37]

Scalability of CFB-produced lasso peptides

To determine if CFB can readily produce larger quantities of lasso peptides, we increased the capistruin-producing CFB reaction volume by 36-fold (0.75 mL). The yield of capistruin was determined using high-performance liquid chromatography (HPLC) against a standard curve of authentic capistruin.[50] Compared to the reported post-HPLC yield of 0.2 μg/mL culture from *E. coli* heterologous expression,[48] CFB resulted in ~200-fold higher yield (40 μg/mL from 0.75 mL of CFB reaction, Figure S8). HPLC retention time analysis and HRMS/MS spectra yielded indistinguishable data for authentic and CFB-produced capistruin (Figure S8). The biological activity of CFB-produced capistruin was then assessed using a modified transcription/translation inhibition assay that monitors the production of the fluorescent protein mCherry.[51] CFB-produced capistruin supplied at 10 μM yielded a 66% reduction in mCherry production compared to a control that omitted capistruin (Figure S8). Presumably, this was due to the known inhibitory activity of capistruin against RNA polymerase.[52] This assay could conceivably be used to evaluate any compound as a potential inhibitor of transcription and/or translation. Compared to microbroth antimicrobial assays, CFB can readily be conducted at 10-fold lower volumes and the use of the fluorescence reporter allows for greater sensitivity and less sample consumption. Compared to macromolecular biosynthesis inhibition assays, this assay also circumvents the need for using radiolabeled metabolic precursors of RNA and protein.[53,54]

CFB to produce sequence-diverse lasso peptides

The above examples demonstrated the ability of CFB to produce a diverse panel of known lasso peptides using an operationally simple procedure with native BGCs expressed from highly pure, concentrated plasmids. Our previous report demonstrated that the fusilassin

biosynthetic pathway is somewhat unusual in that it tolerates substitution at the first position of the core region (naturally Trp) and ring size variants (±1 residue).[6] This result, as previously reported, was obtained from enzymatic reconstitution studies that required molecular cloning, expression, and purification steps for each FusA variant. Herein we show that by simply adding the biosynthetic genes to the CFB mixture, these results were rapidly replicated without the need for any protein purification (Figure S9). In addition to single-site variants, the fusilassin enzymes are known to tolerate more divergent sequences. As described in previous reports, new RiPPs can be obtained by using a modular chimeric substrate strategy where the leader region from one precursor peptide is retained while the core region derives from a second precursor peptide.[55] Using this approach, citrulassin A (D8E) was successfully produced by the reconstituted fusilassin enzymes using a fusilassin leader-citrulassin core chimeric substrate.[6] Formation of the citrulassin variant by the non-cognate fusilassin enzymes suggested that one robust set of substrate-tolerant lasso peptide biosynthetic enzymes could allow the production of novel lasso peptides by simply replacing the core region of the precursor peptide.

We next sought to determine if CFB and the chimeric approach could be combined to produce lasso peptides predicted from genome-mining. We first conducted a BLASTP search of the non-redundant NCBI database using FusC as the query. All proteins with >40% identity to FusC (n = 90) were re-analyzed by RODEO, confirming all were within a high-confidence lasso peptide BGCs (i.e., homologs of FusB and FusE were present as well as a high-scoring precursor peptide). From the resulting list of BGCs containing FusC-like cyclases, we chose ten distinct core sequences to assess the tolerability of FusB/E and FusC for accepting chimeric substrates (Table S1). The genes for the chimeric peptides (FusA leader peptide, $FusA_{LP}$, fused to a non-cognate core peptide, $XxxA_{CP}$) were commercially synthesized with the native Glu acceptor residue retained (Table S2). The chimeric precursor peptides were then individually cloned into the pET28-MBP vector and purified by maxi-prep. As before, the precursor peptides, as well as FusB, FusC, and FusE (from pACYC) were produced by CFB using highly pure plasmid DNA. A mass corresponding to the cyclic product derived from *Thermobifida halotolerans* was detected by MALDI-TOF-MS ($FusA_{LP}$-$HalA_{CP}$, halolassin m/z 2047, Figure S10).

Since fusilassin was previously reconstituted *in vitro* using heterologously expressed and purified proteins,[6,37] we explored the effect of adding purified MBP-FusB, MBP-FusC, and MBP-FusE, rather than the pACYC plasmid, to CFB reactions (Figure S11). The addition of purified enzymes was expected to have positive benefits on the CFB process due to: *(i)* CFB energy resources could be preserved to elevate the production level of the precursor peptide, and *(ii)* the CFB extract contains endogenous proteases, which will compete for the precursor peptide. By direct addition of the biosynthetic proteins at a higher, researcher-defined concentration, the flux towards lasso peptide formation could be maximized relative to substrate degradation. As a positive control, the modified CFB reaction successfully produced fusilassin (Figure 2). Moreover, masses corresponding to the cyclized lasso peptides of two of the ten chimeric substrates, specifically halolassin and cellulassin from *Thermobifida cellulosilytica* ($FusA_{LP}$-$CelA_{CP}$, m/z 2277), also were detected by MALDI-TOF-MS. Like fusilassin, both cyclized lasso peptides produced from the chimeric precursors derive from *Thermobifida*, and their corresponding lasso cyclases show >70%

similarity to FusC (Table S1). The lasso peptide molecular formulas and sequences were confirmed by HRMS/MS and were consistent with a Tyr1-Glu8 and Trp1-Glu9 macrolactam for halolassin and cellulassin, respectively (Figures S12–13). A threaded conformation for halolassin and cellulassin was supported by resistance to carboxypeptidase Y digestion. Consistent with cellulassin produced from the native BGC (Figure S1), resistance to carboxypeptidase Y digestion was abolished after prolonged heating at 95 °C. Conversely, halolassin maintained a folded state after prolonged heating, as indicated by resistance to carboxypeptidase Y digestion before and after heat treatment. The enhanced thermal stability of halolassin relative to cellulassin is likely attributable to the smaller ring region for the former (Figure S14).[12]

For the eight chimeric substrates that failed to be processed by the fusilassin biosynthetic proteins, we sought to determine whether a lack of cyclization by FusC or leader peptidolysis by FusB/FusE was responsible. Due to the presence of endogenous proteases in the *E. coli* cell extract, unprocessed precursor peptides and uncyclized, linear core peptides were expected to undergo proteolytic degradation. To monitor linear core peptide formation, we used the PURExpress *in vitro* Protein Synthesis Kit (NEB), which provides the minimal components necessary for transcription and translation (*i.e.*, endogenous proteases are absent).[56] As a positive control, PURExpress reactions were first carried out to produce the FusA and $FusA_{LP}$-$HalA_{CP}$ precursor peptides. Purified MBP-FusB, MBP-FusE, MBP-FusC, and ATP were then added prior to analysis by MALDI-TOF-MS (Supplemental Methods). Fusilassin and halolassin were readily detected, although a significant amount of linear core (uncyclized) was observed in the halolassin sample (Figure S15). These data suggest that the fusilassin biosynthetic proteins tolerate $HalA_{CP}$, although the processing efficiency is lower relative to the native substrate. These results demonstrate that the endogenous proteases present in the *E. coli* extract will degrade short, uncyclized peptides, as the linear core of HalA was not detected during standard CFB (Figure 2). With the PURExpress system validated, we next tested the eight chimeric precursor peptides that were undetected using the earlier CFB procedure. After precursor peptide production by PURExpress, exogenous addition of biosynthetic proteins, and MALDI-TOF-MS analysis, linear core peptide was detected for 8/8 chimeric substrates and no evidence of cyclization was observed (Figure S15). Therefore, FusB and FusE properly removed $FusA_{LP}$ from the chimeric substrates, but FusC did not tolerate the alternative core sequences.

The core region of the two successful chimeric substrates ($HalA_{CP}$ and $CelA_{CP}$) encode FusA-like hydrophobic residues in the loop region (Table S1). In contrast, the ring regions of these substrates are more variable, suggesting that FusC might be more tolerant of sequence variation in the ring relative to the loop. To test this hypothesis, the loop regions of four failed chimeric substrates derived from *Marinactinospora thermotolerans*, *Nonomuraea candida*, *Streptomyces mobaraensis*, and *Streptomyces* sp. NBS 14/10 were replaced with the FusA loop sequence ($_{10}$LIFVFP$_{15}$, Figure S16). The latter two hybrid lasso peptides (i.e., $FusA_{LP}$-$MobA_{CP}$:LIFVFP and $FusA_{LP}$-$NbsA_{CP}$:LIFVFP) were successfully produced by CFB using exogenously supplied fusilassin biosynthetic proteins, suggesting that a native-like loop sequence was more critical than a native-like ring for FusC substrate tolerance. To evaluate the role of the ring in the successful maturation of these chimeric substrates, an additional set of four hybrid precursor peptides were derived from the same strains. While

these sequences retained their native (non-FusA) loop sequences, the ring regions were replaced with the FusA sequence ($_2$YTAEWG$_7$). No cyclized product was observed from any of these hybrid precursor peptides, in agreement that a native-like loop region was critical for FusC substrate tolerance (Figure S16).

Evaluating FusC substrate tolerance through Ala-substitution

The core sequences of cellulassin and halolassin share the highest identity to fusilassin (8/18 and 6/18 residues, respectively). Taken together with citrulassin A (3/18 allowing for a gap and non-wildtype replacement of Asp8 with Glu, Table S1),[6] the available data demonstrate that FusC can accept certain non-cognate chimeric substrates that differ considerably from FusA. To determine whether any single residue was necessary for cyclization by FusC, we prepared 16 Ala variants of FusA that individually replaced each core residue, except for Ala4 and the Glu9 macrolactam acceptor residue. Given that maxi-prep plasmid preparation is laborious, we developed a time- and cost-efficient method to obtain DNA templates for CFB. Instead of using plasmids, we PCR-amplified the linear DNA products comprising the T7 promoter, lac operon, *fusA* mutant, and T7 terminator for use as the precursor template (Figure S17). To prevent the degradation of linear DNA, CFB reactions were supplemented with the nuclease inhibitor GamS (NCBI accession CAA23978.1, Figure S11).[39] Similar to the chimeric strategy illustrated above, MBP-FusB, MBP-FusC, and MBP-FusE were supplied as purified proteins. The convenience of this optimized CFB strategy was demonstrated by the production of wildtype fusilassin and every Ala-substituted variant except for FusA-R16A (Figure 3). The lack of detection of FusA-R16A was consistent with a previous observation that conversion of FusA-R16G into a mature fusilassin variant by in vitro reconstitution was significantly impaired.[6] We attribute the lack of detection by CFB to inefficient cyclization of FusA-R16A with concomitant degradation by endogenous proteases. Furthermore, this variant removes the only basic residue in the peptide, likely reducing water solubility and rendering detection in positive mode MALDI-TOF-MS more difficult.[57] To evaluate if FusC could tolerate the R16A variant, we created a double variant that introduced Arg into the ring. The W6R/R16A variant of fusilassin was readily detected after CFB-based production, showing that Glu9 is the only position of fusilassin intolerant of Ala-substitution (Figure 3). Therefore, when assessing substrate tolerance by CFB, a number of factors must be considered, including solubility, the limit of detection, and the relative rates of lasso cyclization and proteolytic degradation. To probe more deeply into the substrate tolerance of FusC, a more comprehensive analysis with multi-site replacement was thus warranted.

CFB to probe lasso peptide substrate tolerance through multisite replacement

Based on the results of the chimeric substrate study, the data suggested that FusC is more tolerant to variation in the ring region compared to the loop region. To investigate the tolerance of FusC in more depth, we prepared two libraries that replaced five contiguous residues in these regions (ring positions 2–6 and loop positions 10–14, Table S3). These positions avoid the macrolactam-forming and steric plug residues. Each library was prepared using the degenerate NNK codon (comprises 32 codons that provide all 20 common amino acids and one stop codon) for a theoretical size of 3.2 million variants. In this design, ~16% of the sequences will contain a premature stop codon in the varied region.[58]

The ring and loop libraries were cloned into pET28 lacking the MBP fusion partner, and multiple transformations were performed with chemically competent *E. coli*. The transformed cells were then diluted and plated to obtain ~200 colonies per 10 cm plate with each colony presumed to encode a unique FusA variant. Colony PCR was performed on individual colonies in 96-well plate format to amplify linear DNA (Figure S17). After ethanol precipitation, the purified linear DNA was used in CFB reactions as before (Figures 4–5). The expected masses of the FusA variants range from 1923–2569 Da (ring library) and 1954–2599 Da (loop library) with the lowest and highest masses possessing Gly_5 and Trp_5 variable regions, respectively. A control CFB reaction lacking template DNA was conducted and confirmed that this mass range was clear of any potentially confounding entities (Figure S18). Owing to endogenous proteases present in the CFB extract, we anticipated that the majority of detected masses would correspond to FusC-tolerated cyclic products, although we also expected to occasionally detect linear, uncyclized core peptides (i.e., processed by the leader peptidase but not the lasso cyclase).

Individual colonies from the ring and loop NNK libraries were subjected to the above-described workflow with subsequent MALDI-TOF-MS analysis. Ions within the expected mass range were detected for 513/984 (~52%) ring variants and 3/57 (~5%) loop variants. The peptide sequences of all three successful loop variants were obtained, confirming that the observed masses correspond to the cyclized lasso peptides, instead of the linear core peptides or other truncated precursor peptides. Among the three fusilassin loop variants, no polar or charged residues were present, despite a high probability of occurrence from five the NNK codons (Figure 5).

Given the higher success rate observed for the ring library, we carried out a more thorough analysis. Adjusting for the expected frequency of stop codons within the library, the probability of a random ring sequence (core positions 2–6) being tolerated by FusC corresponds to ~61%.[29] Based on the fusilassin ring library size (3.2 million) and the success rate of lasso formation from our sample size of 984 (61%), the fusilassin biosynthetic proteins are capable of generating 1.85–2.05 million lasso peptides (at a 95% confidence interval) from a penta-substituted ring library (Supplemental Methods).[59] Since the ring region of lasso peptides represents a large fraction of the exposed surface area (Table S4), our studies indicate that ring libraries could be employed to engineer lasso peptides with customized binding affinities and activities. Among the 513 sequences expected to be converted to mature lasso peptides, we sequenced the DNA from more than half ($n = 280$), confirming that the observed mass precisely corresponded to the expected lasso peptide sequence (Figures 4, S19). Unlike previous CFB reactions where MBP precursor peptide fusions were used for fusilassin production, all library members used a DNA template excluding MBP. This flexibility in experimental approach underscores the versatility of the CFB method for creating lasso peptide diversity. In a minority of cases (34%), masses corresponding to the linear core peptides (FusB/E product) were detected, which likely arose from slower rates of lasso peptide formation and proteolytic degradation. The available data support a sequence-dependent rate of lasso peptide formation, which will be the subject of future study.

Ten successfully produced members of the ring library were selected for structure and stability assessment (Figure 4). HRMS/MS analysis supported a Trp1-Glu9 ring composition for each variant (Figure S20–29). Nine out of ten variants were resistant to carboxypeptidase Y digestion, suggestive of the threaded conformation (Figure S30). The LSSLM ring variant was susceptible to carboxypeptidase Y even prior to heating, indicating that this lasso peptide was especially prone to unfolding. Among the nine carboxypeptidase Y-resistant variants, a mass consistent with removal of Ile18 was observed at a ratio that varied dramatically with the mass intensity of the full-length lasso peptide. Also, a range of heat stabilities was noted, which illustrates that the stability of the lasso fold is sequence-dependent (Figure S30).[12]

For each of the 280 sequence-confirmed fusilassin ring variants, we tabulated how many substitutions each sequence contained relative to wildtype (Table S5). As expected from a diverse, randomized library, there was a strong bias towards having all five positions varied: one substitution ($n = 0$); two substitutions ($n = 1$); three substitutions ($n = 4$); four substitutions ($n = 62$); five substitutions ($n = 217$, 76% of the total). A sequence logo, generated from the 280 sequence-confirmed FusC substrates, showed little to no conservation at any of the five varied ring positions. Both analyses indicated a broad tolerance of FusC at each of the five varied ring positions (Figure S31).[60] We further analyzed the preference for specific amino acids by comparing the observed frequency at each position with the expected frequency based on the NNK codon (Figure 4). Overall, FusC displays a modest preference at positions 2, 4, and 6 compared to positions 3 and 5. There is a subtle bias towards hydrophobic residues at the even-numbered positions, such as Met, Phe, and Tyr, while charged residues like Glu, Asp, and Lys were disfavored. These results indicate that the FusC binding pocket for FusA is likely to be hydrophobic, and with the side chains of residues 2, 4, and 6 likely orienting on one face of the ring (odd-numbered residues oriented in the opposite direction), even-numbered ring positions may be more likely engaged by the enzyme. Several other residues show site-specific preferences. For example, Glu is favored at position 5 over all other positions while Lys is favored at position 6. To determine whether we could predict FusC tolerated sequences, we designed five predicted substrates according to the heatmap (core positions 2–6 varied with: MNMYM, TWTEM, NMQIY, MYTFQ, and FYNWK). Our predictions were correct for 5/5 substrates by MALDI-TOF-MS (Figure 4). Further, these FusA variants were readily produced from a single CFB reaction (Figure S32).

We next conducted a parallel analysis for FusC non-substrate sequences. We randomly sequenced 154 non-substrates and confirmed the formation of linear core peptide using PURExpress (Figure S33), indicating that FusB/E universally processed the leader region of all FusA variants. We then analyzed the preference at each position and generated a second heatmap (Figure S34). Analogous to the observations on true substrates, the presence of charged residues (e.g., Glu, Lys, Arg) were enriched in the non-substrate set. To test whether our analysis could predict FusC non-substrates, we designed an additional five FusA ring variants (IKEVT, QDWFM, FLRCL, IDRSY, and LKNFT). Our predictions were correct for 5/5 non-substrates. Similarly, PURExpress reactions confirmed FusB/E produced linear core in every case, and FusC failed to cyclize these peptides (Figure S34).

Bioinformatics predicts extensive diversity of naturally occurring lasso peptides

Given the broad substrate tolerance we observed in the ring region of fusilassin, and that previously reported for microcin J25,[23] we explored whether sequence variability in the ring region is characteristic of known and predicted lasso peptides. We first conducted a survey to update the list of all lasso peptides predicted from the NCBI database. This latest dataset used 11 known lasso leader peptidases as queries for Position-Specific Iterative (PSI)-BLAST (Table S6),[61] with class I-IV lasso peptide and all major phyla known to produce lasso peptides represented. Proteins homologous to the leader peptidase (n = 9,571, as of late 2020) were subjected to RODEO analysis.[28] This yielded 7,701 non-redundant, high-scoring lasso precursor peptides, and more than doubles the number previously reported.[6] Each entry was verified to locally encode the requisite RRE, leader peptidase, and lasso cyclase (Supplemental Dataset 1).

Among the 7,701 lasso precursor peptides, 4,485 unique core sequences remained after removal of identical entries (if two distinct organisms encode the same core sequence, only one was retained). To describe the residue variability at each core position, the Shannon entropy,[62] relative entropy,[63] and ConSurf values[64] were calculated (Table S7). To ascertain the amino acid specificity among naturally encoded lasso peptides, we compared the observed, position-dependent frequency of each amino acid to the expected frequency. The latter was calculated from the weighted average of codon usage at the genus-level (Table S8, Figure S35). For the set of predicted, 9-residue ring (same size as fusilassin) lasso peptides core sequences with all replicate entries removed, the Shannon entropy value at each position, was >3 (values >2 are considered variable). The only exception was the macrolactam acceptor site, which gave a Shannon entropy value <1. Within the ring region, positions 3–7 show the highest variability and the propensity to observe a particular amino acid in this region is near the expected frequency. Core position 7 was the most variable by all three metrics (Shannon and relative entropies and ConSurf score). This analysis further shows that, excluding the macrolactam-forming positions 1 and 9, the ring variability of natural lasso peptides is slightly higher than the loop-tail regions (Table S7). Notably, there are enrichments of specific amino acids, such as Gly/Ser within the ring, but not in the loop-tail region. Phe/Val are significantly enriched at position 11 while Gln/Phe are overrepresented at position 12 (Figure S35). Extensive sequence diversity is evident across all regions of natural lasso peptides; however, certain locations appear to be under evolutionary selection. This result could have been anticipated, as not every possible peptide sequence should be compatible with lasso peptide formation, nor will every possible sequence perform a beneficial function. The naturally encoded sequence space observed today is thus logically confined by substrate-enzyme co-evolution and conditional selection pressure that defines the breadth of the biological activities exhibited by lasso peptides in nature.

Conclusion

The data presented herein demonstrate the first application of cell-free biosynthesis (CFB) for the production of lasso peptides and shows the versatility of this method for enhancing our understanding of the rules for biosynthetic enzyme tolerance. We first produced four

known lasso peptides, capistruin, burhizin, fusilassin, and cellulassin using CFB. Capistruin was used as an example to demonstrate that the CFB reactions are scalable, as we produced 200-fold more capistruin by CFB relative to heterologous expression in live *E. coli*. Capistruin was then subjected to a CFB-based transcription/translation inhibition assay that monitored the production of the fluorescent protein mCherry. Considering the speed, ease of manipulation, and serviceable yields achieved for lasso peptides, CFB provides a new route to overcome production issues often encountered when attempting to isolate mature products from the native producer or through heterologous expression.

This work further advances the chimeric strategy of RiPP production by fusing variable core regions to a native leader peptide to permit substrate recognition by the employed biosynthetic proteins. We produced a new, genomically predicted lasso peptide by this method in addition to confirming wide substrate tolerance for the fusilassin biosynthetic proteins. The efficiency of CFB was leveraged to investigate the substrate scope of the fusilassin pathway by screening a diverse, penta-substituted NNK library of FusA variants. Over 1,000 FusA variants were rapidly assayed, showing FusC is widely permissive of ring substitution but less tolerant in the loop region. In contrast, a previous library screen of microcin J25 variants found a greater tolerance in the loop and tail regions, indicating that different lasso cyclases will exhibit varying degrees of substrate tolerance.[23] Our analysis of natural lasso peptide core sequences also suggests that other lasso peptide pathways may display a wider substrate tolerability in the loop region, with microcin J25 providing precedent.[23] Thus, when used together with the fusilassin biosynthetic proteins, library coverage of a large majority of the exposed surface area of a lasso peptide could be achievable. Our fusilassin variant screening provides a proof-of-concept for applying CFB to evaluate substrate compatibility with biosynthetic enzymes and its impact on the formation of lasso peptides. A comprehensive substrate preference analysis can likely be achieved on any isolated lasso peptide, if optimized CFB conditions are applied (e.g., cell lysate from a compatible organism). Although the library analysis previously reported on microcin J25 provided another approach for screening libraries of lasso peptides by evaluating their antibacterial activity,[23] the method reported herein does not require a known bioactivity.

With less than 2% of predicted lasso peptide sequences being characterized to date, this molecular class of natural product remains critically underexplored. CFB is a promising tool to access these molecules, particularly from difficult to cultivate or otherwise inaccessible organisms. Our data also provide a solid foundation to eventually permit the customization of lasso peptides for specific biotechnological applications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

We thank Julian Hegemann and Adam J. DiCaprio for providing expression constructs and Xiao Rui Guo for assistance with acquiring high-resolution mass spectral data (Univ. of Illinois at Urbana-Champaign). We thank Jozsef Kukolya (Agro-Environmental Research Institute) for providing *Thermobifida cellulosilytica* and Alessandra Eustáquio (Univ. of Illinois at Chicago) for providing capistruin-containing culture extracts. This work was

supported in part by the Ving and May Lee Chemistry Discovery Fund (Department of Chemistry, Univ. of Illinois at Urbana-Champaign) and the National Institutes of Health (R41 AT010852 to M.J.B. and D.A.M.).

References

(1). Montalbán-López M; Scott TA; Ramesh S; Rahman IR; van Heel AJ; Viel JH; Bandarian V; Dittmann E; Genilloud O; Goto Y; Grande Burgos MJ; Hill C; Kim S; Koehnke J; Latham JA; Link AJ; Martínez B; Nair SK; Nicolet Y; Rebuffat S; Sahl H-G; Sareen D; Schmidt EW; Schmitt L; Severinov K; Süssmuth RD; Truman AW; Wang H; Weng J-K; van Wezel GP; Zhang Q; Zhong J; Piel J; Mitchell DA; Kuipers OP; van der Donk WA New Developments in RiPP Discovery, Enzymology and Engineering. Nat. Prod. Rep 2020, 38 (1), 130–239 [PubMed: 32935693]

(2). Hegemann JD; Zimmermann M; Xie X; Marahiel MA Lasso Peptides: An Intriguing Class of Bacterial Natural Products. Acc. Chem. Res 2015, 48 (7), 1909–1919. [PubMed: 26079760]

(3). Arnison PG; Bibb MJ; Bierbaum G; Bowers AA; Bugni TS; Bulaj G; Camarero JA; Campopiano DJ; Challis GL; Clardy J; Cotter PD; Craik DJ; Dawson M; Dittmann E; Donadio S; Dorrestein PC; Entian K-D; Fischbach MA; Garavelli JS; Göransson U; Gruber CW; Haft DH; Hemscheidt TK; Hertweck C; Hill C; Horswill AR; Jaspars M; Kelly WL; Klinman JP; Kuipers OP; Link AJ; Liu W; Marahiel MA; Mitchell DA; Moll GN; Moore BS; Müller R; Nair SK; Nes IF; Norris GE; Olivera BM; Onaka H; Patchett ML; Piel J; Reaney MJT; Rebuffat S; Ross RP; Sahl H-G; Schmidt EW; Selsted ME; Severinov K; Shen B; Sivonen K; Smith L; Stein T; Süssmuth RD; Tagg JR; Tang G-L; Truman AW; Vederas JC; Walsh CT; Walton JD; Wenzel SC; Willey JM; van der Donk WA Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. Nat Prod Rep 2013, 30 (1), 108–160. [PubMed: 23165928]

(4). Burkhart BJ; Hudson GA; Dunbar KL; Mitchell DA A Prevalent Peptide-Binding Domain Guides Ribosomal Natural Product Biosynthesis. Nat. Chem. Biol 2015, 11 (8), 564–570. [PubMed: 26167873]

(5). Yan K-P; Li Y; Zirah S; Goulard C; Knappe TA; Marahiel MA; Rebuffat S Dissecting the Maturation Steps of the Lasso Peptide Microcin J25 in Vitro. Chembiochem 2012, 13 (7), 1046–1052. [PubMed: 22488892]

(6). DiCaprio AJ; Firouzbakht A; Hudson GA; Mitchell DA Enzymatic Reconstitution and Biosynthetic Investigation of the Lasso Peptide Fusilassin. J. Am. Chem. Soc 2019, 141 (1), 290–297. [PubMed: 30589265]

(7). Zong C; Maksimov MO; Link AJ Construction of Lasso Peptide Fusion Proteins. ACS Chem. Biol 2016, 11 (1), 61–68. [PubMed: 26492187]

(8). Li Y; Ducasse R; Zirah S; Blond A; Goulard C; Lescop E; Giraud C; Hartke A; Guittet E; Pernodet J-L; Rebuffat S Characterization of Sviceucin from Streptomyces Provides Insight into Enzyme Exchangeability and Disulfide Bond Formation in Lasso Peptides. ACS Chem. Biol 2015, 10 (11), 2641–2649. [PubMed: 26343290]

(9). Jeanne Dit Fouque K; Bisram V; Hegemann JD; Zirah S; Rebuffat S; Fernandez-Lima F Structural Signatures of the Class III Lasso Peptide BI-32169 and the Branched-Cyclic Topoisomers Using Trapped Ion Mobility Spectrometry–Mass Spectrometry and Tandem Mass Spectrometry. Anal. Bioanal. Chem 2019, 411 (24), 6287–6296. [PubMed: 30707269]

(10). Guerrero-Garzón JF; Madland E; Zehl M; Singh M; Rezaei S; Aachmann FL; Courtade G; Urban E; Rückert C; Busche T; Kalinowski J; Cao Y-R; Jiang Y; Jiang C; Selivanova G; Zotchev SB Class IV Lasso Peptides Synergistically Induce Proliferation of Cancer Cells and Sensitize Them to Doxorubicin. iScience 2020, 23 (12), 101785. [PubMed: 33294793]

(11). Oves-Costales D; Sánchez-Hidalgo M; Martín J; Genilloud O Identification, Cloning and Heterologous Expression of the Gene Cluster Directing RES-701-3, −4 Lasso Peptides Biosynthesis from a Marine Streptomyces Strain. Mar. Drugs 2020, 18 (5), 238.

(12). Hegemann JD Factors Governing the Thermal Stability of Lasso Peptides. ChemBioChem 2020, 21 (1–2), 7–18. [PubMed: 31243865]

(13). Kodani S; Unno K How to Harness Biosynthetic Gene Clusters of Lasso Peptides. J. Ind. Microbiol. Biotechnol 2020, 47(9–10), 703–714. [PubMed: 32705462]

(14). Mukhopadhyay J; Sineva E; Knight J; Levy RM; Ebright RH Antibacterial Peptide Microcin J25 Inhibits Transcription by Binding within and Obstructing the RNA Polymerase Secondary Channel. Mol. Cell 2004, 14 (6), 739–751. [PubMed: 15200952]

(15). Kimura K-I; Kanou F; Takahashi H; Esumi Y; Uramoto M; Yoshihama M Propeptin, a New Inhibitor of Prolyl Endopeptidase Produced by Microbispora. I. Fermentation, Isolation and Biological Properties. J. Antibiot. (Tokyo) 1997, 50 (5), 373–378. [PubMed: 9207905]

(16). Tan S; Ludwig KC; Müller A; Schneider T; Nodwell JR The Lasso Peptide Siamycin-I Targets Lipid II at the Gram-Positive Cell Surface. ACS Chem. Biol 2019, 14 (5), 966–974. [PubMed: 31026131]

(17). Gavrish E; Sit CS; Cao S; Kandror O; Spoering A; Peoples A; Ling L; Fetterman A; Hughes D; Bissell A; Torrey H; Akopian T; Mueller A; Epstein S; Goldberg A; Clardy J; Lewis K Lassomycin, a Ribosomally Synthesized Cyclic Peptide, Kills *Mycobacterium tuberculosis* by Targeting the ATP-Dependent Protease ClpC1P1P2. Chem. Biol 2014, 21 (4), 509–518. [PubMed: 24684906]

(18). Potterat O; Wagner K; Gemmecker G; Mack J; Puder C; Vettermann R; Streicher R BI-32169, a Bicyclic 19-Peptide with Strong Glucagon Receptor Antagonist Activity from *Streptomyces sp*. J. Nat. Prod 2004, 67 (9), 1528–1531. [PubMed: 15387654]

(19). Tanaka T; Tsukuda E; Nozawa M; Nonaka H; Ohno T; Kase H; Yamada K; Matsuda Y RES-701–1, a Novel, Potent, Endothelin Type B Receptor-Selective Antagonist of Microbial Origin. Mol. Pharmacol 1994, 45 (4), 724–730. [PubMed: 8183252]

(20). Helynck G; Dubertret C; Mayaux JF; Leboul J Isolation of RP 71955, a New Anti-HIV-1 Peptide Secondary Metabolite. J. Antibiot. (Tokyo) 1993, 46 (11), 1756–1757. [PubMed: 8270499]

(21). Um S; Kim Y-J; Kwon H; Wen H; Kim S-H; Kwon HC; Park S; Shin J; Oh D-C Sungsanpin, a Lasso Peptide from a Deep-Sea Streptomycete. J. Nat. Prod 2013, 76 (5), 873–879. [PubMed: 23662937]

(22). Ducasse R; Yan K-P; Goulard C; Blond A; Li Y; Lescop E; Guittet E; Rebuffat S; Zirah S Sequence Determinants Governing the Topology and Biological Activity of a Lasso Peptide, Microcin J25. ChemBioChem 2012, 13 (3), 371–380. [PubMed: 22287061]

(23). Pan SJ; Link AJ Sequence Diversity in the Lasso Peptide Framework: Discovery of Functional Microcin J25 Variants with Multiple Amino Acid Substitutions. J. Am. Chem. Soc 2011, 133 (13), 5016–5023. [PubMed: 21391585]

(24). Pavlova O; Mukhopadhyay J; Sineva E; Ebright RH; Severinov K Systematic Structure-Activity Analysis of Microcin J25. J. Biol. Chem 2008, 283 (37), 25589–25595. [PubMed: 18632663]

(25). Hegemann JD; De Simone M; Zimmermann M; Knappe TA; Xie X; Di Leva FS; Marinelli L; Novellino E; Zahler S; Kessler H; Marahiel MA Rational Improvement of the Affinity and Selectivity of Integrin Binding of Grafted Lasso Peptides. J. Med. Chem 2014, 57 (13), 5829–5834. [PubMed: 24949551]

(26). Al Toma RS; Kuthning A; Exner MP; Denisiuk A; Ziegler J; Budisa N; Süssmuth RD Site-Directed and Global Incorporation of Orthogonal and Isostructural Noncanonical Amino Acids into the Ribosomal Lasso Peptide Capistruin. ChemBioChem 2015, 16 (3), 503–509. [PubMed: 25504932]

(27). Piscotta FJ; Tharp JM; Liu WR; Link AJ Expanding the Chemical Diversity of Lasso Peptide MccJ25 with Genetically Encoded Noncanonical Amino Acids. Chem. Commun 2015, 51 (2), 409–412.

(28). Tietz JI; Schwalen CJ; Patel PS; Maxson T; Blair PM; Tai H-C; Zakai UI; Mitchell DA A New Genome-Mining Tool Redefines the Lasso Peptide Biosynthetic Landscape. Nat. Chem. Biol 2017, 13 (5), 470–478. [PubMed: 28244986]

(29). Wyss DF; Lahm H-W; Manneberg M; Labhardt AM Anantin-A Peptide Antagonist of the Atrial Natriuretic Factor(ANF). II. Determination of the Primary Sequence by NMR on the Basis of Proton Assignments. J. Antibiot. (Tokyo) 1991, 44 (2), 172–180. [PubMed: 1826288]

(30). Maksimov MO; Link AJ Discovery and Characterization of an Isopeptidase That Linearizes Lasso Peptides. J. Am. Chem. Soc 2013, 135 (32), 12038–12047. [PubMed: 23862624]

(31). Hegemann JD; Zimmermann M; Zhu S; Klug D; Marahiel MA Lasso Peptides from Proteobacteria: Genome Mining Employing Heterologous Expression and Mass Spectrometry. Biopolymers 2013, 100 (5), 527–542. [PubMed: 23897438]

(32). Mevaere J; Goulard C; Schneider O; Sekurova ON; Ma H; Zirah S; Afonso C; Rebuffat S; Zotchev SB; Li Y An Orthogonal System for Heterologous Expression of Actinobacterial Lasso Peptides in Streptomyces Hosts. Sci. Rep 2018, 8 (1), 8232. [PubMed: 29844351]

(33). Zong C; Cheung-Lee WL; Elashal HE; Raj M; Link AJ Albusnodin: An Acetylated Lasso Peptide from *Streptomyces albus*. Chem. Commun 2018, 54 (11), 1339–1342.

(34). Bratovanov EV; Ishida K; Heinze B; Pidot SJ; Stinear TP; Hegemann JD; Marahiel MA; Hertweck C Genome Mining and Heterologous Expression Reveal Two Distinct Families of Lasso Peptides Highly Conserved in Endofungal Bacteria. ACS Chem. Biol 2020, 15 (5), 1169–1176. [PubMed: 31800204]

(35). Martin-Gómez H; Linne U; Albericio F; Tulla-Puche J; Hegemann JD Investigation of the Biosynthesis of the Lasso Peptide Chaxapeptin Using an *E. coli*- Based Production System. J. Nat. Prod 2018, 81 (9), 2050–2056. [PubMed: 30178995]

(36). Silverman AD; Karim AS; Jewett MC Cell-Free Gene Expression: An Expanded Repertoire of Applications. Nat. Rev. Genet 2020, 21 (3), 151–170. [PubMed: 31780816]

(37). Koos JD; Link AJ Heterologous and *in vitro* Reconstitution of Fuscanodin, a Lasso Peptide from *Thermobifida fusca*. J. Am. Chem. Soc 2019, 141 (2), 928–935. [PubMed: 30532970]

(38). Duquesne S; Destoumieux-Garzón D; Zirah S; Goulard C; Peduzzi J; Rebuffat S Two Enzymes Catalyze the Maturation of a Lasso Peptide in *Escherichia coli*. Chem. Biol 2007, 14 (7), 793–803. [PubMed: 17656316]

(39). Sun ZZ; Yeung E; Hayes CA; Noireaux V; Murray RM Linear DNA for Rapid Prototyping of Synthetic Biological Circuits in an Escherichia coli Based TX-TL Cell-Free System. ACS Synth. Biol 2014, 3 (6), 387–397. [PubMed: 24303785]

(40). Perez JG; Stark JC; Jewett MC Cell-Free Synthetic Biology: Engineering Beyond the Cell. Cold Spring Harb. Perspect. Biol 2016, 8 (12), a023853. [PubMed: 27742731]

(41). Kay JE; Jewett MC Lysate of Engineered *Escherichia coli* Supports High-Level Conversion of Glucose to 2,3-Butanediol. Metab. Eng 2015, 32, 133–142. [PubMed: 26428449]

(42). Kightlinger W; Duncker KE; Ramesh A; Thames AH; Natarajan A; Stark JC; Yang A; Lin L; Mrksich M; DeLisa MP; Jewett MC A Cell-Free Biosynthesis Platform for Modular Construction of Protein Glycosylation Pathways. Nat. Commun 2019, 10 (1), 5404. [PubMed: 31776339]

(43). Goering AW; Li J; McClure RA; Thomson RJ; Jewett MC; Kelleher NL *In vitro* Reconstruction of Nonribosomal Peptide Biosynthesis Directly from DNA Using Cell-Free Protein Synthesis. ACS Synth. Biol 2017, 6 (1), 39–44. [PubMed: 27478992]

(44). Ozaki T; Yamashita K; Goto Y; Shimomura M; Hayashi S; Asamizu S; Sugai Y; Ikeda H; Suga H; Onaka H Dissection of Goadsporin Biosynthesis by *In vitro* Reconstitution Leading to Designer Analogues Expressed *In vivo*. Nat. Commun 2017, 8 (1), 14207. [PubMed: 28165449]

(45). Liu R; Zhang Y; Zhai G; Fu S; Xia Y; Hu B; Cai X; Zhang Y; Li Y; Deng Z; Liu T A Cell-Free Platform Based on Nisin Biosynthesis for Discovering Novel Lanthipeptides and Guiding Their Overproduction *In vivo*. Adv. Sci 2020, 7 (17), 2001616.

(46). Fleming SR; Bartges TE; Vinogradov AA; Kirkpatrick CL; Goto Y; Suga H; Hicks LM; Bowers AA Flexizyme-Enabled Benchtop Biosynthesis of Thiopeptides. J. Am. Chem. Soc 2019, 141 (2), 758–762. [PubMed: 30602112]

(47). Vinogradov AA; Shimomura M; Goto Y; Ozaki T; Asamizu S; Sugai Y; Suga H; Onaka H Minimal Lactazole Scaffold for *In vitro* Thiopeptide Bioengineering. Nat. Commun 2020, 11 (1), 2272. [PubMed: 32385237]

(48). Knappe TA; Linne U; Zirah S; Rebuffat S; Xie X; Marahiel MA Isolation and Structural Characterization of Capistruin, a Lasso Peptide Predicted from the Genome Sequence of *Burkholderia thailandensis* E264. J. Am. Chem. Soc 2008, 130 (34), 11446–11454. [PubMed: 18671394]

(49). Knappe TA; Linne U; Robbel L; Marahiel MA Insights into the Biosynthesis and Stability of the Lasso Peptide Capistruin. Chem. Biol 2009, 16 (12), 1290–1298. [PubMed: 20064439]

(50). Kunakom S; Eustáquio AS Heterologous Production of Lasso Peptide Capistruin in a *Burkholderia* Host. ACS Synth. Biol 2020, 9 (2), 241–248. [PubMed: 31913601]

(51). Shaner NC; Campbell RE; Steinbach PA; Giepmans BNG; Palmer AE; Tsien RY Improved Monomeric Red, Orange and Yellow Fluorescent Proteins Derived from *Discosoma sp*. Red Fluorescent Protein. Nat. Biotechnol 2004, 22 (12), 1567–1572. [PubMed: 15558047]

(52). Braffman NR; Piscotta FJ; Hauver J; Campbell EA; Link AJ; Darst SA Structural Mechanism of Transcription Inhibition by Lasso Peptides Microcin J25 and Capistruin. Proc. Natl. Acad. Sci 2019, 116 (4), 1273–1278. [PubMed: 30626643]

(53). Polikanov YS; Osterman IA; Szal T; Tashlitsky VN; Serebryakova MV; Kusochek P; Bulkley D; Malanicheva IA; Efimenko TA; Efremenkova OV; Konevega AL; Shaw KJ; Bogdanov AA; Rodnina MV; Dontsova OA; Mankin AS; Steitz TA; Sergiev PV Amicoumacin A Inhibits Translation by Stabilizing mRNA Interaction with the Ribosome. Mol. Cell 2014, 56 (4), 531–540. [PubMed: 25306919]

(54). Travin DY; Watson ZL; Metelev M; Ward FR; Osterman IA; Khven IM; Khabibullina NF; Serebryakova M; Mergaert P; Polikanov YS; Cate JHD; Severinov K Structure of Ribosome-Bound Azole-Modified Peptide Phazolicin Rationalizes Its Species-Specific Mode of Bacterial Translation Inhibition. Nat. Commun 2019, 10 (1), 4563. [PubMed: 31594941]

(55). Burkhart BJ; Kakkar N; Hudson GA; van der Donk WA; Mitchell DA Chimeric Leader Peptides for the Generation of Non-Natural Hybrid RiPP Products. ACS Cent. Sci 2017, 3 (6), 629–638. [PubMed: 28691075]

(56). Shimizu Y; Inoue A; Tomari Y; Suzuki T; Yokogawa T; Nishikawa K; Ueda T Cell-Free Translation Reconstituted with Purified Components. Nat. Biotechnol 2001, 19 (8), 751–755. [PubMed: 11479568]

(57). Krause E; Wenschuh H; Jungblut PR The Dominance of Arginine-Containing Peptides in MALDI-Derived Tryptic Mass Fingerprints of Proteins. Anal. Chem 1999, 71 (19), 4160–4165. [PubMed: 10517141]

(58). Sieber T; Hare E; Hofmann H; Trepel M Biomathematical Description of Synthetic Peptide Libraries. PLOS ONE 2015, 10 (6), e0129200 [PubMed: 26042419]

(59). Fleiss JL; Levin B; Paik MC Statistical Methods for Rates and Proportions; Shewart WA, Wilks SS, Series Eds.; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2003.

(60). Crooks GE WebLogo: A Sequence Logo Generator. Genome Res 2004, 14 (6), 1188–1190. [PubMed: 15173120]

(61). Johnson M; Zaretskaya I; Raytselis Y; Merezhuk Y; McGinnis S; Madden TL NCBI BLAST: A Better Web Interface. Nucleic Acids Res 2008, 36 (Web Server), W5–W9. [PubMed: 18440982]

(62). Shannon CE A Mathematical Theory of Communication. Bell Syst. Tech. J 1948, 27 (3), 379–423.

(63). Capra JA; Singh M Predicting Functionally Important Residues from Sequence Conservation. Bioinformatics 2007, 23 (15), 1875–1882. [PubMed: 17519246]

(64). Ashkenazy H; Abadi S; Martz E; Chay O; Mayrose I; Pupko T; Ben-Tal N ConSurf 2016: An Improved Methodology to Estimate and Visualize Evolutionary Conservation in Macromolecules. Nucleic Acids Res 2016, 44 (W1), W344–350. [PubMed: 27166375]

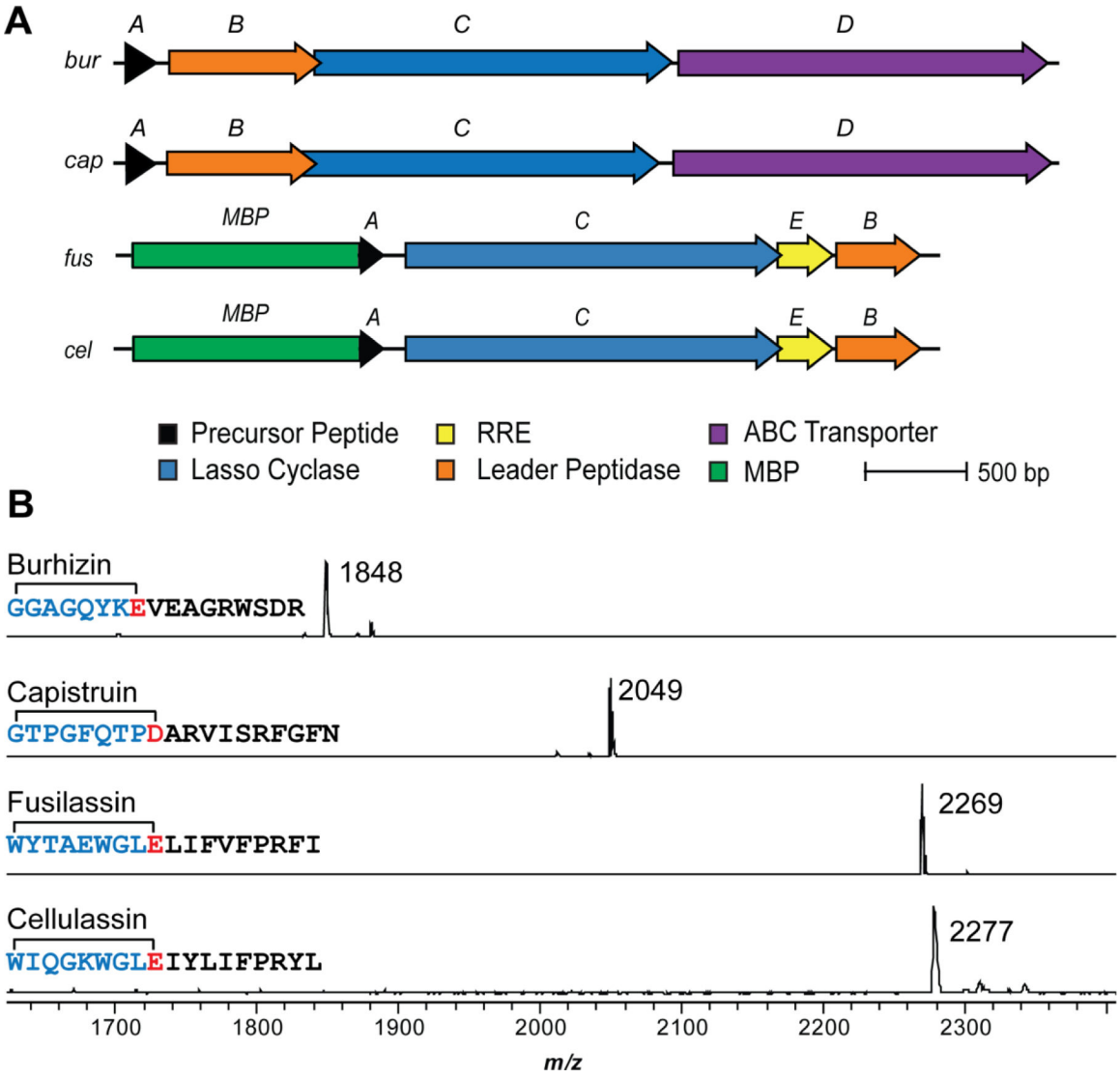


Figure 1. ***In vitro*** **lasso peptide production through CFB.**

(A) The biosynthetic gene clusters provided in the CFB reaction for the production of burhizin, capistruin, fusilassin, and cellulassin respectively. RRE: RiPP leader peptide Recognition Element. MBP: maltose binding protein. **(B)** Endpoint MALDI-TOF-MS assay of burhizin (*m/z* 1848), capistruin (*m/z* 2046), fusilassin (*m/z* 2269), and cellulassin (*m/z* 2277) produced from CFB. Blue, macrolactam ring residues; Red, acceptor site. The mass label corresponds to the $[M+H]^+$ ion of the lasso peptides.

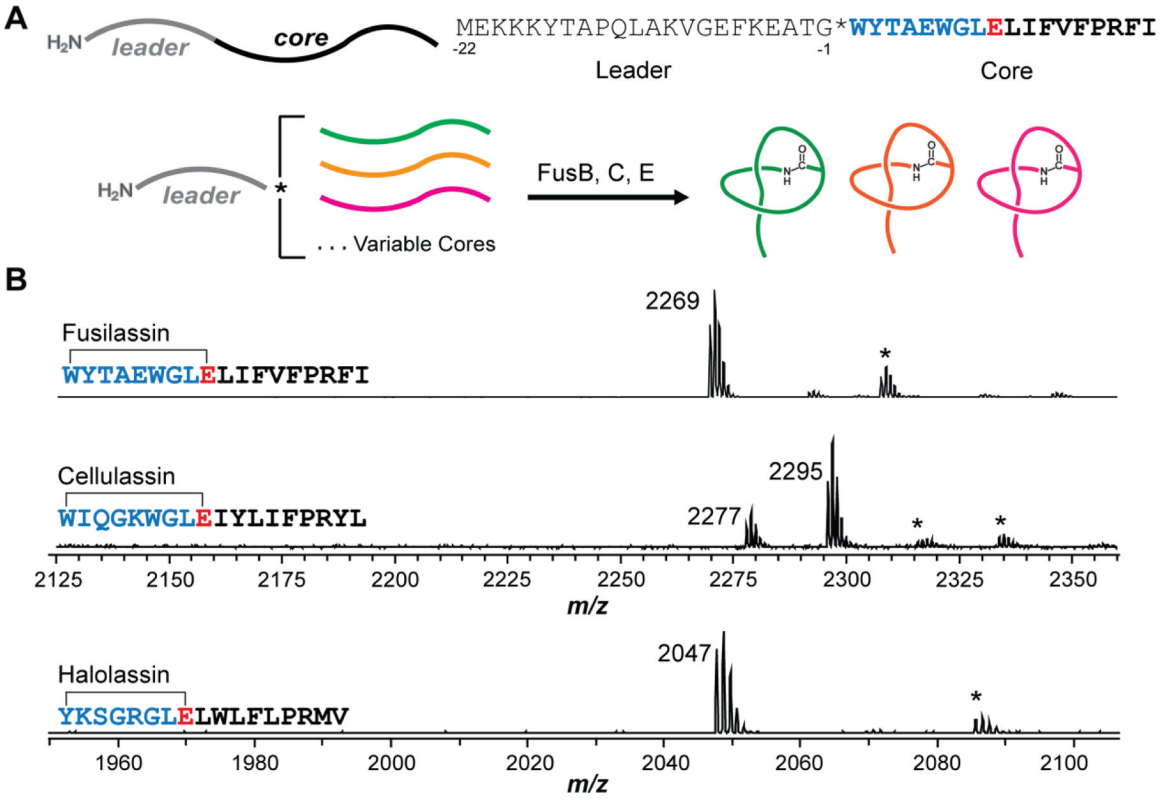


Figure 2. Chimeric substrate strategy to generate diverse lasso peptides.

(A) The chimeric substrate principle fuses diverse (non-cognate) core peptides to the leader peptide sequence of FusA (residues −1 to −22). This permits a single set of biosynthetic proteins (FusB, C, and E) to produce various lasso peptides. **(B)** Endpoint MALDI-TOF-MS assay of CFB reactions with fusilassin as the positive control; $FusA_{LP}$-$CelA_{CP}$ (cellulassin, 2277 Da; uncyclized core 2295 Da) and $FusA_{LP}$-$HalA_{CP}$ (halolassin, 2047 Da) as the chimeric substrates. Precursor peptides were produced from CFB and reacted with purified MBP-FusB, MBP-FusC, and MBP-FusE to form the mature lasso peptide. Blue, macrolactam ring residues; Red, acceptor site. The mass label corresponds to the $[M+H]^+$ ion of the lasso peptide. * indicates the $[M+K]^+$ ion for the lasso peptide.

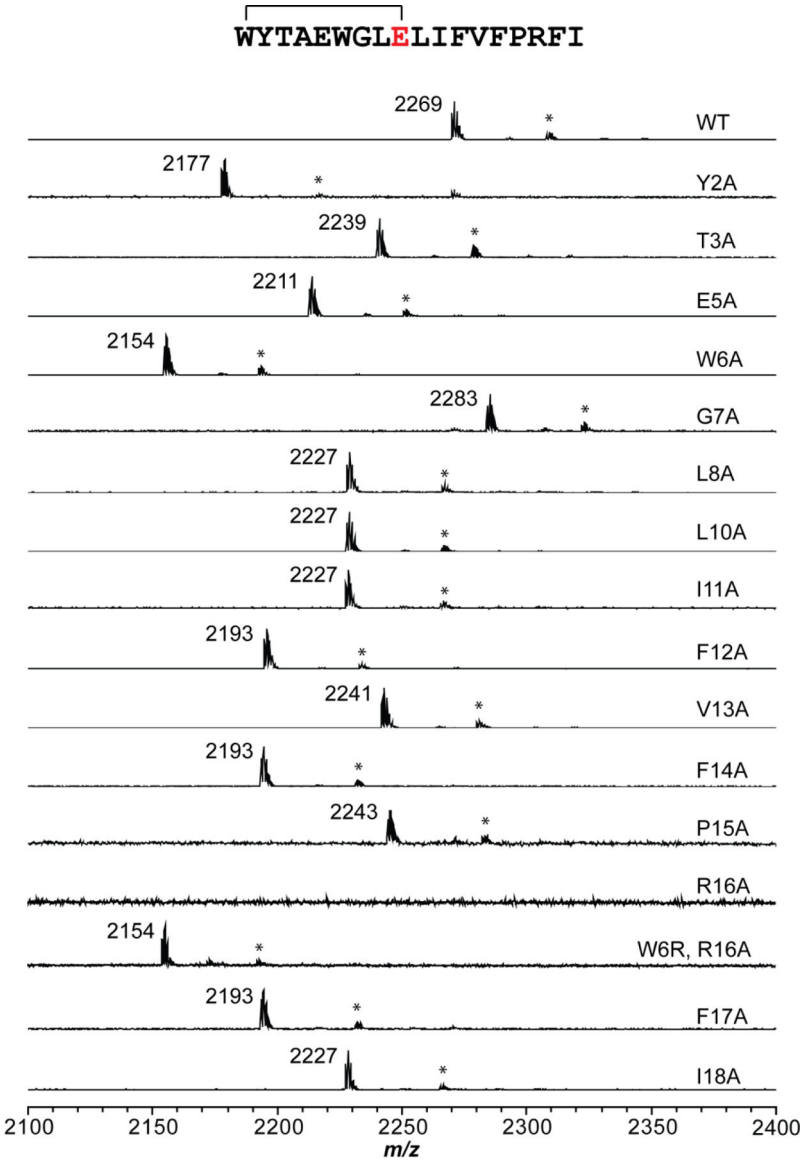


Figure 3. Evaluating FusC substrate tolerance through Ala-substitution.

The sequence of wildtype fusilassin is provided at the top with endpoint MALDI-TOF-MS assay data for CFB-produced fusilassin variants shown below. The Glu9 acceptor site is intolerant to substitution.[6] The mass label denotes the lasso peptide $[M+H]^+$ ion. * indicates the lasso peptide $[M+K]^+$ ion.

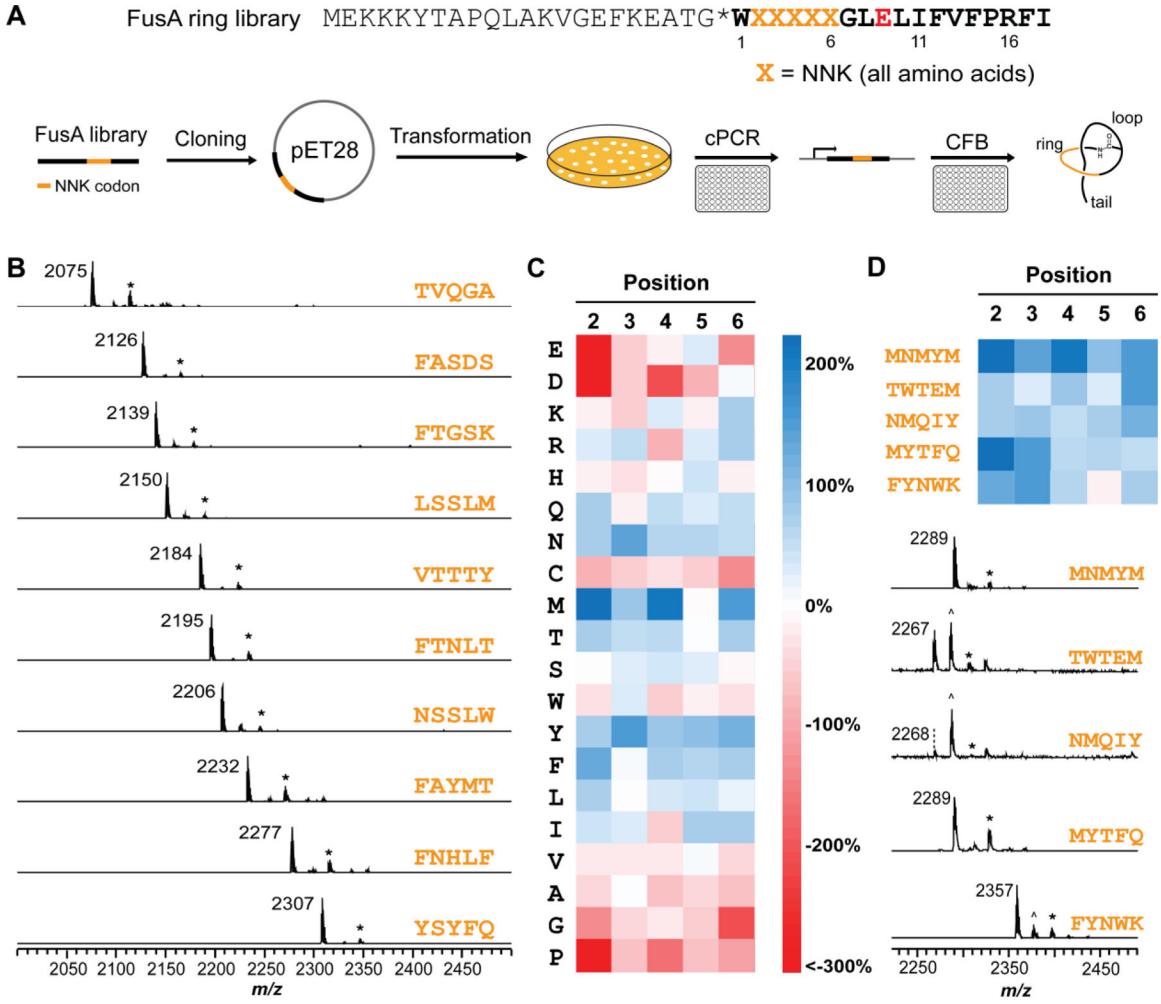


Figure 4. Substrate compatibility evaluation of FusA ring.

(A) Experimental workflow for fusilassin ring library screening. Positions 2–6 of the FusA core region were diversified using 5 NNK codons (orange, library size of 3.2 M unique peptide sequences). After library cloning and transformation into *E. coli*, colony PCR was performed in 96-well plates. The linear DNA products were then used in CFB reactions to produce lasso peptides. **(B)** MALDI-TOF mass spectra of 10 representative fusilassin ring variants. The sequences of the varied region are orange. **(C)** Heatmap analysis on confirmed FusC substrates (lasso-compatible sequences) showing the percent difference between the occurrence of a residue by core position and the expected frequency based on the NNK codon. The analysis was conducted from 280 substrates. **(D)** MALDI-TOF mass spectra showing the lasso peptide formation of five predicted FusC substrates from the heatmap. At least four residues in designed sequences are predicted to be favored in FusC substrates. The precursor peptides were synthesized through CFB and reacted with heterologously expressed and purified FusB, C, and E. The mass labels denote the $[M+H]^+$ ion of the lasso peptides. ^ indicates the uncyclized linear core peptide. * indicates the $[M+K]^+$ ion for the lasso peptide.

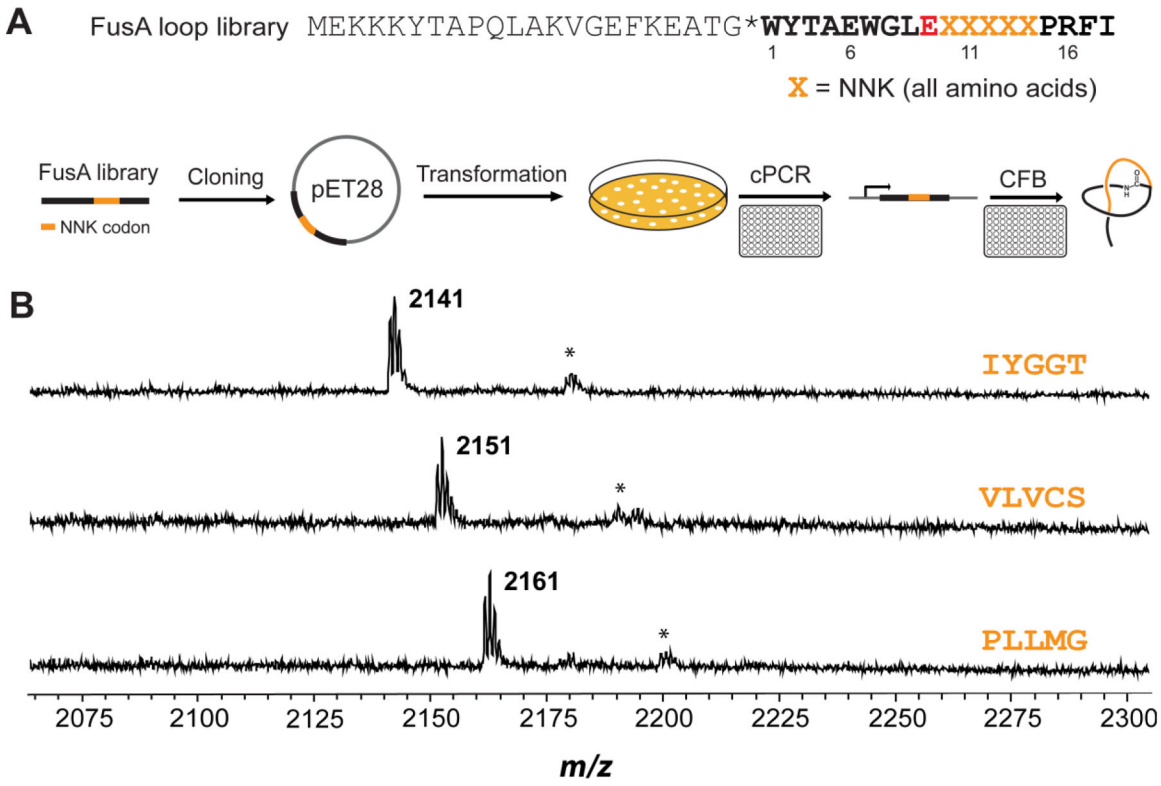


Figure 5. Substrate compatibility evaluation of FusA loop.

(A) Experimental workflow for fusilassin loop library evaluation. Positions 10–14 of the FusA core region were diversified using 5 contiguous NNK codons (orange, theoretical library size of 3.2 M unique peptide sequences). After library cloning and transformation into *E. coli*, colony PCR (cPCR) was performed in 96-well plates. The linear DNA products were then used in CFB reactions to produce lasso peptides. **(B)** From 57 randomly chosen clones, lasso peptide formation was only detected for three loop variants. Orange depicts the varied sequence at positions 10–14 of the FusA loop region. The mass labels denote to the $[M+H]^+$ ion of the lasso peptide. * indicates $[M+K]^+$ ion for the lasso peptide.