Supporting Information for:

Enzymatic reconstitution and biosynthetic investigation of the lasso peptide Fusilassin

Adam J. DiCaprio[†,‡], Arash Firouzbakht[†,‡], Graham A. Hudson[†,‡], Douglas A. Mitchell[†,‡]*

†Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, USA. ‡Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 1206 West Gregory Drive, Urbana, Illinois 61801, USA.

* Corresponding author:

Douglas A. Mitchell (douglasm@illinois.edu), phone: 1-217-333-1345, fax: 1-217-333-0508

Table of Contents:

Experimental Methods.

General materials and methods. Reagents used for molecular biology experiments were purchased from New England BioLabs (Ipswich, MA), Thermo Fisher Scientific (Waltham, MA), or Gold Biotechnology Inc. (St. Louis, MO). Other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). *Escherichia coli* DH5α and BL21 (DE3) strains were used for plasmid maintenance and protein overexpression, respectively. *Thermobifida fusca* was identified as a potential producer of a lasso peptide using previously published RODEO-based genome mining[1] and was obtained from the USDA Agriculture Research Service strain collection (Peoria, IL). Matrix-assisted laser desorption time of flight mass spectrometry MALDI-TOF-MS analysis was performed using a Bruker UltrafleXtreme mass spectrometer in reflector positive mode at the University of Illinois School of Chemical Sciences Mass Spectrometry Laboratory. Electrospray ionization (ESI)-MS/MS analyses were performed using a ThermoFisher Scientific Orbitrap Fusion ESI-MS using an Advion TriVersa Nanomate 100.

Molecular biology techniques. Oligonucleotides were purchased from Integrated DNA Technologies Inc. (Coralville, IA). *Thermobifida fusca* genomic DNA was isolated using an UltraClean Microbial DNA Isolation Kit (MO BIO, now available from Qiagen). Purified gDNA isolates were used to clone biosynthetic proteins using the primers listed in **Table S2**. Amplicons were purified using a QIAprep PCR Cleanup Spin Kit (Qiagen). The isolated DNA was then treated with appropriate restriction enzymes (**Table S2**, NEB). Digested inserts were re-purified using a QIAprep PCR Cleanup Spin Kit (Qiagen). The digested genes were reconstituted with a similarly digested and gel-purified pET28-MBP vector and subsequently ligated using T4 DNA ligase (NEB). Ligation reactions were transformed into chemically competent DH5α cells, which were plated on lysogeny broth (LB) agar plates containing 50 μg/mL kanamycin and grown at 37 °C overnight. Colonies were picked at random and grown in LB broth containing 50 μg/mL kanamycin for 16–20 h prior to plasmid isolation using a QIAprep Spin Miniprep Kit. All recombinant constructs contained an MBP-tag followed by a TEV-protease site and were sequenced using a custom MBP forward primer. All constructs were Sanger sequenced using a T7 terminator primer if necessary (**Table S2**).

For *E. coli* expression plasmids, FusB and FusE were cloned individually with the primers listed in **Table 2**. Individual amplicons shared an overlapping sequence (3' to FusE and 5' to FusB) containing an inserted ribosome binding site. These amplicons were purified using a QIAprep PCR Cleanup Spin Kit (Qiagen) and used as templates for a second round of PCR to combine the fragments using primers listed in **Table S2**. This amplicon was purified using a QIAprep PCR Cleanup Spin Kit (Qiagen). This fragment was inserted into the second multiple cloning site (MCS) of amplified pACYCDuet vector (primers in **Table S2**) using Gibson Assembly HiFi master mix (NEB). Ligation reactions were used to transform chemically competent DH5α cells, which were plated on LB agar plates containing 34 μg/mL chloramphenicol and grown at 37 °C overnight. Colonies were picked at random and grown in LB broth containing 34 μg/mL chloramphenicol for 16–20 h prior to plasmid isolation using a QIAprep Spin Miniprep Kit. This construct was verified by Sanger sequencing using a T7 terminator primer.

FusC was cloned from *T. fusca* gDNA using the primers listed in **Table S2.** Amplicons were purified using a QIAprep PCR Cleanup Spin Kit (Qiagen). The isolated DNA was then treated with appropriate restriction enzymes (**Table S2**, NEB). Digested inserts were re-purified using a QIAprep PCR Cleanup Spin Kit (Qiagen). The digested genes were reconstituted with an appropriate endonuclease-digested and gel-purified pACYCDuet-MCSI-FusEB vector and ligated using T4 DNA ligase (NEB). Ligation reactions were used to transform chemically competent DH5α cells, which were plated on LB agar plates containing 34 μg/mL chloramphenicol and grown at 37 °C overnight. Colonies were picked at random and grown in LB broth containing 34 μg/mL chloramphenicol for 16–20 h prior to plasmid isolation using a QIAprep Spin Miniprep Kit. Recombinant constructs contained an N-terminal His_6-tag. This construct was sequenced using custom forward and reverse primers (**Table S2**). The ABC transporter FusD wass not included in our cloning pipeline.

Site-directed mutagenesis on FusA, FusB, FusC and FusE was performed using the QuikChange method (Agilent) as per the manufacturer's instructions using Q5 DNA polymerase (NEB). The primers for each mutant are listed in **Table S2**. The mutant FusA-W1E was cloned using indicated primers and previously described site-directed ligase-independent mutagenesis (SLIM).[2] The mutations were verified by Sanger sequencing using the MBP forward primer or the T7 terminator primer.

***E. coli*-based expression of fusilassin.** *E. coli* (BL21(DE3) cells were co-transformed with pACYC-FusCEB and pET28-MBP-FusA. Cells were grown overnight on LB agar plates containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol at 37 °C. Single colonies were used to inoculate 10 mL of LB containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol and grown at 30 °C for 12 h. This culture was used to inoculate 250 mL of LB containing 25 μg/mL kanamycin and 17 μg/mL chloramphenicol which was grown at 37 °C to an optical density at 600 nm (OD_{600}) of 0.7-0.8. Expression was then induced by the addition of 0.5 mM (final concentration) isopropyl β-D-1-thiogalactopyranoside (IPTG). Expression was allowed to proceed for 3 h at 37 °C. Cells were harvested by centrifugation at 4,500 × *g* for 10 min. The harvested cells were extracted with 10 mL methanol with shaking for 3 h at room temperature. Extracts were clarified by centrifugation at 4,500 × *g* for 20 min and stored at -20 °C until analysis. Samples were analyzed by MALDI-TOF MS by spotting on a MALDI target with a 1:1 volume of saturated sinapinic acid in 7:3:0.1 acetonitrile (MeCN):H_2O:formic acid (FA).

MBP-tagged peptide and enzyme overexpression and purification. *E. coli* BL21(DE3) cells were transformed with a pET28 plasmid encoding the MBP-tagged enzyme or peptide of interest. Cells were grown overnight on LB agar plates containing 50 μg/mL kanamycin at 37 °C. Single colonies were used to inoculate 10 mL of LB containing 50 μg/mL kanamycin and grown at 30 °C for 12 h. This culture was used to inoculate 1 L of LB containing 37.5 μg/mL kanamycin and 25.5 μg/mL chloramphenicol which was grown at 37 °C to an OD_{600} of 0.7-0.8. Protein expression was then induced by the addition of 0.5 mM (final concentration) IPTG. Expression was allowed to proceed for 3 h at 37 °C. Cells were harvested by centrifugation at 4,500 × *g* for 10 min, flash-frozen, and stored at -80 °C for a maximum of 7 d before use.

Cell pellets were resuspended in lysis buffer (50 mM Tris-HCl pH 7.4, 500 mM NaCl, 2.5% glycerol (*v/v*), and 0.1% Triton X-100) containing 100-120 mg lysozyme and 500 μL protease inhibitor cocktail (PIC, 6 mM PMSF, 0.1 mM leupeptin, 0.1 mM E64). Expressions of MBP-FusB were not treated with PIC owing to a concern over inhibition of catalytic activity. After resuspension, cells were lysed by sonication (3 × 45 s with 10 min rocking periods at 4 °C). Insoluble debris was removed by centrifugation at 17,000 × *g* for 60 min. The supernatant was then applied to a pre-equilibrated amylose resin (NEB; 5 mL of resin per L of culture). The column was washed with 10 column volumes (CV) of lysis buffer followed by 10 CV of wash buffer (lysis buffer without Triton X-100). The MBP-tagged proteins were eluted with 15 mL elution buffer (lysis buffer with 300 mM NaCl, 10 mM maltose, and lacking Triton X-100) and collected into an appropriate molecular weight cutoff (MWCO) Amicon Ultra centrifugal filter (EMD Millipore). Protein eluent was concentrated to ~1.5 mL and exchanged with 10× volume of protein storage buffer [50 mM HEPES pH 7.5, 300 mM NaCl, 0.5 mM tris-(2-carboxyethyl)-phosphine (TCEP), 2.5% glycerol (*v/v*)]. Protein concentrations were assayed using 280 nm absorbance (theoretical extinction coefficients were calculated using the ExPASy ProtParam tool; http://web.expasy.org/protparam/protpar-ref.html). Final protein purity was assessed visually using a Coomassie-stained SDS-PAGE gel (**Figure S1**).

Expression and purification of MBP-FusC. Chemically competent *E. coli* BL21 (DE3) cells (50 μL) were co-transformed with pET28-MBP-FusC and pGro7 chaperone plasmid (Takara Bio USA, Inc.) and plated on LB agar plates supplemented with 50 μg/mL kanamycin and 37μg/mL chloramphenicol and grown at 37 °C overnight. A single colony was used to inoculate 10 mL of LB supplemented with kanamycin and chloramphenicol (as above), grown for 12 h at 30 °C. Cultures were used to inoculate 1L of LB containing 25 μg/mL kanamycin, 17 μg/mL chloramphenicol, and 0.5-4 mg/mL L-arabinose, which were grown at 37 °C to an OD_{600} of 0.7-0.8. Protein expression was induced by the addition of IPTG to a final concentration of 0.5 mM and cultures were grown at 37 °C for 3 h. Protein purification by amylose resin affinity chromatography was performed as above.

To express variants of MBP-FusC, chemically competent *E. coli* BL21 (DE3) cells were transformed with mutant pET28-MBP-FusC plated on LB agar plates supplemented with 50 μg/mL kanamycin and grown at 37 °C overnight. A single colony was used to inoculate 10 mL of LB supplemented with 50 μg/mL kanamycin and grown for 16 h at 37 °C. Cultures were used to inoculate 1 L of LB supplemented with 30 μg/mL kanamycin and 3% (*v/v*) molecular biology grade 100% ethanol. Given the influence of the Gro chaperone on proper folding of FusC, we simplified the dual plasmid expression system using ethanol-induced Gro expression.[3] Cells were grown at 37 °C to an OD_{600} of 0.7–0.8. Protein expression was induced by the addition of IPTG to a final concentration of 0.5 mM and cultures were grown at 37 °C for 3 h. Protein purification by amylose resin affinity chromatography was performed as above.

Synthetic peptides. Synthetic peptides were purchased from Genscript and were used without further purification [purity: 73% $FusA_{core}$; 91% $FusA\text{-}W1P_{core}$]. No interfering masses were observed by MS analysis.

***In vitro* enzymatic reconstitution of fusilassin biosynthesis.** Purified enzymes were reconstituted in 100 µL synthetase buffer (50 mM Tris-HCl pH 7.5, 125 mM NaCl, 20 mM $MgCl_2$, 10 mM DTT, 5 mM ATP) at the molar ratios of 1:5 MBP-FusBCE:MBP-FusA. Reactions were allowed to progress for 3 h at 37 °C. Samples were spotted on a MALDI target with an equal volume of saturated sinapinic acid in 7:3:0.1 MeCN:H_2O:FA. Reaction extent was estimated as the proportion of the masses corresponding to the linear precursor and the dehydrated, macrolactam-containing product (-18 Da).

Preparation and purification of citrulassin variants. *In vitro* biosynthetic reconstitution reactions (200 µL total volume) were iniated as above, except molar ratios of MBP-FusCEB:Precursor peptide was increased to 10:1. Reactions were diluted with 7:3 MeCN:H_2O (1% formic acid, *v/v*) to the HPLC starting conditions of the HPLC run. After clarification by centrifugation (17,000 × g for 15 mins) the full reaction volume was injected onto an Agilent 1200 series HPLC fitted with a pre-equilibrated 4.6 × 250 mm Zorbax C8 column (Agilent) operating at 1 mL/min. A gradient elution was used with solvent A (H_2O, 0.1% formic acid) and solvent B (MeCN, 0.1% formic acid) according to the following linear gradient combinations: 0-5 min held at 20% B, 5-25 min increased to 55% B, 25-28 min increased to 100% B, 28-33 min held at 100% B. Fractions were screened for the presence of the respective macrolactam product (CitA-D8E *m/z* = 1579 Da, CitA-L1W-ΔA5-D8E *m/z* = 1581 Da) by MALDI-TOF-MS. Positive fractions were concentrated to dryness under vacuum. Products were resuspended in 25 µL of 7:3 MeCN:H_2O and 1% formic acid and analyzed by MS.

High-resolution MS analysis of lasso peptide variants. Wild-type fusilassin and the ring expanded 10-mer variant produced through *in vitro* enzymatic reactions were desalted by ZipTip and eluted with 75% aq. MeCN. Citrulassin A analogs were first purified by HPLC (see above). Samples were directly infused into a ThermoFisher Scientific Orbitrap Fusion ESI-MS using an Advion TriVersa Nanomate 100. The MS was calibrated and tuned with Pierce LTQ Velos ESI Positive Ion Calibration Solution (ThermoFisher). The MS was operated using the following parameters: resolution, 100,000; isolation width (MS/MS), 2 *m/z*; normalized collision energy (MS/MS), 70; activation q value (MS/MS), 0.4; activation time (MS/MS), 30 ms. Fragmentation was performing using collision-induced dissociation (CID) at 70%. Data analysis was conducted using the Qualbrowser application of Xcalibur software (Thermo-Fisher Scientific).

Multiple sequence alignments. The sequence alignments shown in Figs S4 & S6 were generated using Clustal Omega[4] (https://www.ebi.ac.uk/Tools/msa/clustalo/). The MUSCLE module in MEGA-X[5] with the default parameters was used to prepare MSAs for larger datasets.

Enzymatic activity of Ala-substituted FusB and FusC proteins.
FusB variants. Enzymatic reconstitution of FusB variants was performed as above, but omitted FusC. Cleavage of the leader peptide was assessed by withdrawing 4 µL of the reaction, diluting with 40 µL 1% (*v/v*) aqueous formic acid and desalting by ZipTip. Samples were eluted with 3 µL of a saturated solution of α-cyano-4-hydroxycinnamic acid (CHCA) in 70:30:0.1 MeCN:H_2O:formic acid and spotted onto the MALDI target. Reaction efficiencies were estimated as a function of the ion intensity of *m/z* = 2287 Da ($FusA_{core}$) relative to *m/z* = 4955 Da ($FusA_{precursor}$).

FusC variants. Enzymatic reconstitution of FusC variants was performed as above, including FusC. Reaction efficiencies were estimated as a function of the ion intensity of *m/z* = 2269 Da (Fus) relative to that of *m/z* = 2287 Da ($FusA_{core}$).

Compendium of all lasso peptides found in GenBank. A non-redundant list of 3,661 leader peptidase protein accession identifiers from protein family PF13471 (transglut_core3) was gathered using four iterations of PSI-BLAST with known lasso leader peptidase proteins from Firmicutes, Proteobacteria, and Actinobacteria as the query. After removal of redundant sequences, each NCBI protein accession identifier was submitted to RODEO2 (https://github.com/thedamlab/rodeo2) using the scoring function for lasso peptide precursor prediction. Initial gene cluster analysis yielded ~125,000 putative peptides that co-occurred with the queried leader peptidase. Entries corresponding to redundant core peptides were removed by standard excel functions such that each unique core peptide sequence occurred only once per leader peptidase query. Hypothetical peptides receiving a score of <5 were deleted. Peptides with a score between 5-9 (~3,400 non-redundant sequences) are found in **Supplemental Dataset 1** on the "Lower Scoring" tab. The false-positive rate here is expected to be high although there are many valid-looking biosynthetic gene clusters present that harbor a non-conforming precursor peptide. Peptides receiving a

score of 10 or greater are considered to be most probable candidates for being a true lasso precursor peptide. This list was manually curated to remove entries that look incapable of forming a lasso peptide. These are included in **Supplemental Dataset 1** on the "Higher Scoring" tab, which contains ~3,000 non-redundant sequences.

Protein isoelectric point calculation. Isoelectric point calculations were performed using the Isoelectric Point Calculator[6], and reported values were the averages of the 14 separate formulae used in the algorithm.

Coevolutionary analysis of leader peptide:RRE and RRE:leader peptidase. Coevolutionary analysis of the leader peptide, RRE, and leader peptidase were performed using GREMLIN (http://gremlin.bakerlab.org/). First, the sequences of leader peptide, RRE, and the leader peptidase of the lasso peptide BGCs identified above that encode a discrete protease (i.e. non-fused) were retrieved. In cases were a BGC encoded >1 putative precursor peptide, only the top-scoring peptide was analyzed. Next, two paired alignments were generated using MUSCLE in MEGA-X. One included the leader peptide and RRE sequences while the second contained the RRE and leader peptidase sequences. Finally, the two alignments were separately submitted to GREMLIN by setting iterations to 0, filtering sequences that have less than 75% coverage by the template sequence, and removing sites on the alignment with <75% coverage. In both GREMLIN submissions, the Fus sequences were used as template (the first entry in the alignment).

Preparation of fluorescently labeled $FusA_{leader}$. HPLC-purified $FusA_{leader}$ that is modified to contain an N-terminal Cys was prepared by GenScript. One equivalent (5.0 mg) of the 85% pure peptide was dissolved in 1 mL reaction buffer [0.1 M phosphate pH 8.0, 10% (*v/v*) DMF, and 5.0 mg (6.3 eq) 5-iodoacetamide fluorescein (5-IAF). The mixture was allowed to react at room temperature in the dark for 2 h. Conjugation was verified via MALDI-TOF-MS by mixing 1 µL of the reaction with 1 µL of a saturated solution of CHCA in H_2O:MeCN:FA 49.5:49.5:0.5 and analyzed the conjugated product (*m/z* 3534.8 Da) The reaction was diluted to 5 mL with aqueous solution of 1% (*v/v*) FA and clarified by centrifugation. The supernatant was applied to a 200 mg HyperSep C18 cartridge equilibrated in 97:2:1 H_2O:MeCN:FA. The resin was washed with 10 mL of the equilibration solution and the labeled peptide was eluted with 5 mL of 89:20:1 H_2O:MeCN:FA. The eluent was dried under vacuum at room temperature. The solid was redissolved in 500 µL of 97:2:1 H_2O:MeCN:FA and clarified by centrifugation. The supernatant was injected into an Agilent 1200 HPLC system equipped with a 250 × 10mm (5 µm particle size, 100 Å pore size) C18 BetaSil column. Separation was performed at a flow rate of 4 mL/min and solvent phases composed of 99.9:0.1 H_2O:FA (A) and 99.9:0.1 MeCN:FA (B). Separation was achieved by the following gradient run on an equilibrated stationary phase: 2–40% B over 20 min then 40–100% B over 5 min. Fractions containing $Fl\text{-}FusA_{leader}$ were evaporated to dryness under vacuum. Solid was resuspended in MilliQ water and stored at -20°C. Labeled peptide was quantified by reading the absorbance at 392 nm of a solution of the labeled peptide in PBS and using $\varepsilon_{392} = 85{,}000\ M^{-1}\ cm^{-1}$.[7]

Fluorescence Polarization (FP) Assays.

Direct assays. Interactions between fluorescently labeled FusA leader peptide ($Fl\text{-}FusA_{leader}$) and FusE variants were quantified using fluorescence polarization. Polarization signals were maximized by leaving FusE as an MBP-fusion. Two-fold serial dilutions of MBP-FusE were made in binding buffer (50 mM Tris pH 7.5, 500 mM NaCl, 2.5% glycerol, 50 µM $ZnSO_4$) and mixed with 25 nM $Fl\text{-}FusA_{leader}$ in non-binding black polystyrene 384-well plates (Corning). Polarization measurements were taken with a FilterMax F5 multimode microplate reader (Molecular Devices) using $\lambda_{ex} = 485$ nm and $\lambda_{em} = 535$ nm. Data fitting was done by OriginPro using the dose-response curve fitting. The concentration of the titrant at the inflection point was recorded as K_D. All measurements were done in triplicate.

Competition assays. Interactions between FusE and variants of $FusA_{leader}$ were quantified by fluorescence polarization in a competition setting. In order to produce $MBP\text{-}FusA_{leader}$ variants, the first codon of the core region of FusA was changed to a stop codon (W1*) with site-directed mutagenesis. Using this construct as template, mutations of interest were introduced in the leader region of FusA and the leader peptides were expressed and purified as an MBP fusion. All the conditions remain the same as (a), except all the wells were supplemented with 200 nM MBP-FusE and the titrant was changed to the $MBP\text{-}FusA_{leader}$ variant of interest. The concentration of the titrant at the inflection point was recorded as IC_{50}.

Evaluation of threadedness

Biosynthetically reconstituted lasso peptides were concentrated and desalted by Zip Tip and eluted into 20 µL MeCN. This suspension was evaporated to dryness under vacuum and resuspended in 50 ng/µL carboxypeptidase Y in 25 mM NH_4HCO_3 pH 6.5. This mixture was briefly vortexed and allowed to react at room temperature for 18 h. Resulting peptidase mixtures were evaporated to dryness under vacuum and resuspended in 3 µL saturated sinapinic acid in 3:7:0.1 H_2O:MeCN:FA. The entire volume was spotted onto the MALDI target and analyzed by MALDI-TOF MS. For thermal unthreading assays, eluted lasso peptides were incubated at 95 °C for 2 h. This product was then treated as above.

Hydrogen-deuterium exchange MS.
Hydrogen-deuterium exchange was performed using previously established methods.[8] Briefly, biosynthetically reconstituted fusilassin was concentrated and desalted by Zip Tip and eluted into 20 µL acetonitrile. This suspension was evaporated to dryness under vacuum and resuspended in 10 µL 3:7 D_2O:MeCN (25 mM NH_4HCO_3, pH 5.5). A standard solution of 25 µg/mL synthetic $FusA_{core}$ was prepared identically to fusilassin. At indicated time points, 1.5 µL aliquots were removed and quenched by spotting onto the MALDI target with 1.5 µL saturated sinapinic acid in 3:7 D_2O:MeCN + 0.1% formic acid (pH = 2). Samples were then analyzed by MALDI-TOF MS.

Given the large mass dispersions observed for deuterated masses of interest, centroid masses were calculated according to the following:

$$\sum_{i=1}^{n} x_i y_i + x_{i+1} y_{i+1} \ldots + x_n y_n$$

Where x = *m/z* value of peak, y = area value of mass peak, i = first *m/z* value of peak cluster with a S/N ratio >2 and n = last *m/z* value of peak cluster with S/N ratio >2. The centroid *m/z* of fusilassin (2270.2) or $FusA_{core}$ (2287.2) were subtracted from respective calculated centroid masses to determine the number of incorporated deuterons. This value was reported as a raw value or as a fraction of the maximal deuterium incorporation. Maximum deuterium number was calculated as the sum of all backbone amides, an exchangeable side chain protons, excluding side chains of Thr/Ser, Asp/Glu, α-NH_3^+ and the C-terminal carboxylate. Exclusions were based on established methods describing observable exchangeable protons.[9]

Table S1: Sequences of *Fus* biosynthetic genes for *E. coli* expression, empty vectors for heterologous expressions and the pACYCDuet vector. All sequences are provided 5′ to 3′. Restriction sites for cloning are underlined and restriction enzymes are indicated in parentheses.

***FusA* (BamHI, NotI)**

GGATCCATGGAAAAGAAGAAGTACACCGCTCCGCAGCTCGCTAAGGTCGGCGAATTCAAGGAGGCCACCGGCTGGTACACCGC
GGAATGGGGCCTCGAGCTGATCTTCGTCTTCCCGCGCTTCATCTGAGCGGCCGC

***FusB* (HindIII, NotI)**

AAGCTTGATGAGCGAGAACGTAGTGCTGCAGCGCAGCAACGTGCGCCTGTCGTGGCGGACGAAGTGGGCGGCGCGCTGCGCGG
TGGGAGCGGCCCGGCTCtTGGCCCGCAAACCCCCGGAGCGGATCCGCGCTACCCTGCTGCGTCTCCGGGGCGAGGTGCGGCCC
GCCACCTACGAAGAGGCCAAAGCCGCCCGGGACGCGGTGCTCGCGGTCAGCCTGCGCTGCGCCGGACTGCGGGCCTGCCTGCA
GCGTTCGCTGGCTATCGCACTGCTGTGCCGGATGCGCGGCACCTGGGCCACCTGGTGTGTGGGGGTTCCCCGCCGCCCGCCGT
TCATCGGACACGCCTGGGTGGAGGCCGAAGGACGCCTCGTGGAAGAAGGCGTCGGCTACGACTACTTCAGCCGTCTGATCACC
GTGGACTGAGCGGCCGC

***FusC* (HindIII, NotI)**

AAGCTTGATGGTCGGTTGCATCAGTCCTTACTTCGCTGTTTTCCCTGATAAGGACGTTCTCGGGCAGGCAACCGACCGCCTTC
CGGCAGCCCAGACGCTGGCGTCCCACCCTTCGGGACGCCCTTGGCTTGTAGGCGCGCTCCCCGCCGACCAGCTCCTGCTGGTC
GAGGCCGGGGAGCGCCGCCTCGCCGTGATCGGACACTGTTCGGCTGAGCCTGAGCGGCTGCGTGCGGAACTCGCCCAGATCGA
CGACGTGGCCCAGTTCGACCGCATCGCGCGCACCCTGGACGGCAGCTTTCACCTGGTCGTGGTCGTCGGCGACCAGATGCGGA
TCCAAGGAAGCGTGTCCGGGCTCCGCCGAGTCTTCCACGCCCACGTCGGTACAGCCCGAATCGCCGCGGACCGCTCCGACGTG
CTCGCCGCGGTACTCGGCGTCTCCCCCGACCCGGACGTGCTCGCGCTGCGGATGTTCAACGGCCTGCCCTACCCGCTCAGCGA
GCTGCCCCCATGGCCCGGGGTGGAGCACGTCCCCGCCTGGCACTACCTCAGCCTCGGCTTGCATGACGGCCGCCACCGGGTCG
TCCAGTGGTGGCATCCCCCGGAAGCGGAACTCGCCGTCACCGCCGCAGCTCCCCTGCTGCGCACTGCCCTGGCCGGAGCCGTG
GACACCCGCACCCGCGGTGGCGGAGTGGTTAGCGCCGACCTGTCCGGCGGGCTCGACTCCACTCCCCTCTGCGCCCTGGCCGC
TCGCGGCCCGGCCAAAGTCGTCGCCCTGACCTTCAGCAGCGGCCTTGACACGGACGATGATCTCCGCTGGGCGAAGATCGCCC
ACCAGTCCTTTCCCTCGGTGGAGCACGTGGTGCTCTCCCCCGAGGACATCCCCGGCTTCTACGCTGGCCTCGACGGGGAGTTC
CCCCTGTTGGACGAACCGTCCGTGGCCATGTTGAGCACGCCGCGCATTCTGAGCCGGCTGCACACGGCCCGCGCCCACGGCTC
CCGCCTGCACATGGACGGCCTGGGCGGCGACCAGCTCCTCACCGGATCGCTCAGCCTCTACCACGATCTGCTCTGGCAGCGGC
CGTGGACTGCCTTGCCGCTCATCCGCGGCCACCGCCTCCTGGCGGGACTGTCGCTGTCAGAGACCTTCGCTTCGCTCGCTGAC
CGCCGTGACCTCCGCGCATGGCTTGCCGATATCCGGCACTCCATCGCCACGGGAGAACCGCCGCGCCGCTCCCTGTTCGGATG
GGACGTCCTGCCCAAGTGCGGGCCGTGGCTGACCGCCGAGGCGCGCGAGCGCGTGCTGGCCCGCTTCGACGCGGTCCTGGAGT
CGCTGGAGCCTCTCGCCCCGACCCGAGGCAGGCACGCTGACCTGGCCGCGATCCGCGCGGCCGGCCGCGACCTGCGACTGCTG
CACCAGCTCGGCAGCAGCGACCTGCCCCGCATGGAATCGCCGTTCCTCGACGACCGCGTGGTGGAGGCCTGCCTTCAGGTGCG
CCACGAAGGCCGGATGAACCCGTTCGAATTCAAGTCCCTCATGAAAACGGCCATGGCTTCCCTCCTCCCCGCGGAGTTCTTGA
CCCGCCAGTCGAAGACCGACGGCACTCCGCTCGCTGCGGAGGGGTTCACCGAACAACGCGATCGCATCATCCAGATATGGCGG
GAGTCCCGACTGGCGGAACTGGGGCTCATCCACCCCGACGTGCTCGTGGAACGGGTCAAACAGCCCTACTCGTTCCGGGGCCC
GGACTGGGGCATGGAGCTCACCCTCACCGTCGAGCTGTGGCTGCGGTCCCGCGAACGTGTCCTCCAAGGAGCAAACGGTGGAG
ACAACAGGAGCTGAGCGGCCGC

***FusE* (BamHI, NotI)**

GGATCCATGGAGACAACAGGAGCTGAATTCCGGCTACGTCCCGAGATCTCAGTCGCGCAAACCGACTACGGCATGGTGCTGCT
GGACGGCCGCAGCGGTGAGTACTGGCAGCTCAACGACACCGCGGCGCTGATCGTGCAGCGGCTGCTGGACGGGCACTCCCCCG
CCGACGTGGCCCAGTTCCTCACCAGCGAATACGAGGTGGAGCGGACCGACGCGGAACGCGACATCGCCGCGCTCGTCACCAGC
CTGAAAGAGAACGGGATGGCGCTGCCATGCTGAGCGGCCGC

pET28-MBP-Empty (MCS underlined)

TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCA
GCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGG
GGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGG
GCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAA
CAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTG
ATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGG
AACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAAACTCATCGAGCATC
AAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTC
ACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATT
TCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGC
ATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCG

TGATTGCGCCTGAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGA
ACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCGGGGATCGCA
GTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAG
TCTGACCATCTCATCTGTAACATCATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCAT
ACAATCGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAA
TTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAG
TTTTATTGTTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCT
TCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGA
TCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGT
AGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGT
GGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTC
GTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTC
CCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAAC
GCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAG
CCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGT
TATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGC
GAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATA
TGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATG
GCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGT
GACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAG
CGTGGTCGTGAAGCGATTCACAGATGTCTGCCTGTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGG
CTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCAT
GGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTT
GTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACA
GATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGT
TTCCAGACTTTACGAAACACGGAAACCGAAGACCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTC
ACGTTCGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGC
ACGATCATGCGCACCCGTGGGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGAC
GAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGCGCTCCAGCGAAAGCGGTCCT
CGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATA
GTCATGCCCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGT
GAGCTAACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCG
GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATT
GCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGG
TGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCAGCCCG
GACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAG
CATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGA
GATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCC
AATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGAC
ATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATGATCA
GCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACC
ACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGC
AACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTT
CCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATAC
TCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCG
AAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGT
AGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCC
TGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGATAT
AGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGAGATCTCGATCCCGCGA
AATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAG
ATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGAAAATCGAAGAAG
GTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGA
ATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTAT
CTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACA
AGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTG
ATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGG
TAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGT
ATGAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTG
ATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGACCAT
CAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAAC

CATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCGAAAGAGTTCCTCGAA
AACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGA
AGAGTTGGCGAAAGATCCACGTATTGCCGCCACCATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGT
CCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCG
CAGACTAATTCGAGCTCCCACCATCACCATCACCACGCGAATTCGGTACCGCTGGTTCCGCGTGGATCTGAGAACCTGTACTT
CCAATCCGGATCCCTCGAGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAA
AGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGA
GGGGTTTTTTGCTGAA

pACYCDuet-FusC-FusEB

GGGGAATTGTGAGCGGATAACAATTCCCCTGTAGAAATAATTTTGTTTAACTTTAATAAGGAGATATACCATGGGCAGCAGCC
ATCACCATCATCACCACAGCCAGGATCCGAATTCGAGCTCGGCGCGCCTGCAGGTCGACATGGTCGGTTGCATCAGTCCTTAC
TTCGCTGTTTTCCCTGATAAGGACGTTCTCGGGCAGGCAACCGACCGCCTTCCGGCAGCCCAGACGCTGGCGTCCCACCCTTC
GGGACGCCCTTGGCTTGTAGGCGCGCTCCCCGCCGACCAGCTCCTGCTGGTCGAGGCCGGGGAGCGCCGCCTCGCCGTGATCG
GACACTGTTCGGCTGAGCCTGAGCGGCTGCGTGCGGAACTCGCCCAGATCGACGACGTGGCCCAGTTCGACCGCATCGCGCGC
ACCCTGGACGGCAGCTTTCACCTGGTCGTGGTCGTCGGCGACCAGATGCGGATCCAAGGAAGCGTGTCCGGGCTCCGCCGAGT
CTTCCACGCCCACGTCGGTACAGCCCGAATCGCCGCGGACCGCTCCGACGTGCTCGCCGCGGTACTCGGCGTCTCCCCCGACC
CGGACGTGCTCGCGCTGCGGATGTTCAACGGCCTGCCCTACCCGCTCAGCGAGCTGCCCCCATGGCCCGGGGTGGAGCACGTC
CCCGCCTGGCACTACCTCAGCCTCGGCTTGCATGACGGCCGCCACCGGGTCGTCCAGTGGTGGCATCCCCCGGAAGCGGAACT
CGCCGTCACCGCCGCAGCTCCCCTGCTGCGCACTGCCCTGGCCGGAGCCGTGGACACCCGCACCCGCGGTGGCGGAGTGGTTA
GCGCCGACCTGTCCGGCGGGCTCGACTCCACTCCCCTCTGCGCCCTGGCCGCTCGCGGCCCGGCCAAAGTCGTCGCCCTGACC
TTCAGCAGCGGCCTTGACACGGACGATGATCTCCGCTGGGCGAAGATCGCCCACCAGTCCTTTCCCTCGGTGGAGCACGTGGT
GCTCTCCCCCGAGGACATCCCCGGCTTCTACGCTGGCCTCGACGGGGAGTTCCCCCTGTTGGACGAACCGTCCGTGGCCATGT
TGAGCACGCCGCGCATTCTGAGCCGGCTGCACACGGCCCGCGCCCACGGCTCCCGCCTGCACATGGACGGCCTGGGCGGCGAC
CAGCTCCTCACCGGATCGCTCAGCCTCTACCACGATCTGCTCTGGCAGCGGCCGTGGACTGCCTTGCCGCTCATCCGCGGCCA
CCGCCTCCTGGCGGGACTGTCGCTGTCAGAGACCTTCGCTTCGCTCGCTGACCGCCGTGACCTCCGCGCATGGCTTGCCGATA
TCCGGCACTCCATCGCCACGGGAGAACCGCCGCGCCGCTCCCTGTTCGGATGGGACGTCCTGCCCAAGTGCGGGCCGTGGCTG
ACCGCCGAGGCGCGCGAGCGCGTGCTGGCCCGCTTCGACGCGGTCCTGGAGTCGCTGGAGCCTCTCGCCCCGACCCGAGGCAG
GCACGCTGACCTGGCCGCGATCCGCGCGGCCGGCCGCGACCTGCGACTGCTGCACCAGCTCGGCAGCAGCGACCTGCCCCGCA
TGGAATCGCCGTTCCTCGACGACCGCGTGGTGGAGGCCTGCCTTCAGGTGCGCCACGAAGGCCGGATGAACCCGTTCGAATTC
AAGTCCCTCATGAAAACGGCCATGGCTTCCCTCCTCCCCGCGGAGTTCTTGACCCGCCAGTCGAAGACCGACGGCACTCCGCT
CGCTGCGGAGGGGTTCACCGAACAACGCGATCGCATCATCCAGATATGGCGGGAGTCCCGACTGGCGGAACTGGGGCTCATCC
ACCCCGACGTGCTCGTGGAACGGGTCAAACAGCCCTACTCGTTCCGGGGCCCGGACTGGGGCATGGAGCTCACCCTCACCGTC
GAGCTGTGGCTGCGGTCCCGCGAACGTGTCCTCCAAGGAGCAAACGGTGGAGACAACAGGAGCTGAGCGGCCGCATAATGCTT
AAGTCGAACAGAAAGTAATCGTATTGTACACGGCCGCATAATCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGAT
ACAATTCCCCATCTTAGTATATTAGTTAAGTATAAGAAGGAGATATACATATGGAGACAACAGGAGCTGAATTCCGGCTACGT
CCCGAGATCTCAGTCGCGCAAACCGACTACGGCATGGTGCTGCTGGACGGCCGCAGCGGTGAGTACTGGCAGCTCAACGACAC
CGCGGCGCTGATCGTGCAGCGGCTGCTGGACGGGCACTCCCCCGCCGACGTGGCCCAGTTCCTCACCAGCGAATACGAGGTGG
AGCGGACCGACGCGGAACGCGACATCGCCGCGCTCGTCACCAGCCTGAAAGAGAACGGGATGGCGCTGCCATGCTGATTAGTT
AAGTATAAGAAGGAGATATACATATGAGCGAGAACGTAGTGCTGCAGCGCAGCAACGTGCGCCTGTCGTGGCGGACGAAGTGG
GCGGCGCGCTGCGCGGTGGGAGCGGCCCGGCTCtTGGCCCGCAAACCCCCGGAGCGGATCCGCGCTACCCTGCTGCGTCTCCG
GGGCGAGGTGCGGCCCGCCACCTACGAAGAGGCCAAAGCCGCCCGGGACGCGGTGCTCGCGGTCAGCCTGCGCTGCGCCGGAC
TGCGGGCCTGCCTGCAGCGTTCGCTGGCTATCGCACTGCTGTGCCGGATGCGCGGCACCTGGGCCACCTGGTGTGTGGGGGTT
CCCCGCCGCCCGCCGTTCATCGGACACGCCTGGGTGGAGGCCGAAGGACGCCTCGTGGAAGAAGGCGTCGGCTACGACTACTT
CAGCCGTCTGATCACCGTGGACTGAATGGCAGATCTCAATTGGATATCGGCCGGCCACGCGATCGCTGACGTCGGTACCCTCG
AGTCTGGTAAAGAAACCGCTGCTGCGAAATTTGAACGCCAGCACATGGACTCGTCTACTAGCGCAGCTTAATTAACCTAGGCT
GCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAACCTCAGGC
ATTTGAGAAGCACACGGTCACACTGCTTCCGGTAGTCAATAAACCGGTAAACCAGCAATAGACATAAGCGGCTATTTAACGAC
CCTGCCCTGAACCGACGACCGGGTCGAATTTGCTTTCGAATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGGCGTAGC
ACCAGGCGTTTAAGGGCACCAATAACTGCCTTAAAAAAATTACGCCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTCAT
TAAGCATTCTGCCGACATGGAAGCCATCACAGACGGCATGATGAACCTGAATCGCCAGCGGCATCAGCACCTTGTCGCCTTGC
GTATAATATTTGCCCATAGTGAAAACGGGGGCGAAGAAGTTGTCCATATTGGCCACGTTTAAATCAAAACTGGTGAAACTCAC
CCAGGGATTGGCTGAGACGAAAAACATATTCTCAATAAACCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAACACGCCACAT
CTTGCGAATATATGTGTAGAAACTGCCGGAAATCGTCGTGGTATTCACTCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGG
AAAACGGTGTAACAAGGGTGAACACTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAACTCCGGATGAGCATT
CATCAGGCGGGCAAGAATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATAT
CCAGCTGAACGGTCTGGTTATAGGTACATTGAGCAACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATATA
TCAACGGTGGTATATCCAGTGATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAATCTCGATAACTCAAAAAATACGCC
CGGTAGTGATCTTATTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTCATTTTCGCCAAAAGTTGGCCC
AGGGCTTCCCGGTATCAACAGGGACACCAGGATTTATTTATTCTGCGAAGTGATCTTCCGTCACAGGTATTTATTCGGCGCAA

AGTGCGTCGGGTGATGCTGCCAACTTACTGATTTAGTGTATGATGGTGTTTTTGAGGTGCTCCAGTGGCTTCTGTTTCTATCA
GCTGTCCCTCCTGTTCAGCTACTGACGGGGTGGTGCGTAACGGCAAAAGCACCGCCGGACATCAGCGCTAGCGGAGTGTATAC
TGGCTTACTATGTTGGCACTGATGAGGGTGTCAGTGAAGTGCTTCATGTGGCAGGAGAAAAAAGGCTGCACCGGTGCGTCAGC
AGAATATGTGATACAGGATATATTCCGCTTCCTCGCTCACTGACTCGCTACGCTCGGTCGTTCGACTGCGGCGAGCGGAAATG
GCTTACGAACGGGGCGGAGATTTCCTGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGAGGGCCGCGGCAAAGCCGTTT
TTCCATAGGCTCCGCCCCCCTGACAAGCATCACGAAATCTGACGCTCAAATCAGTGGTGGCGAAACCCGACAGGACTATAAAG
ATACCAGGCGTTTCCCCTGGCGGCTCCCTCGTGCGCTCTCCTGTTCCTGCCTTTCGGTTTACCGGTGTCATTCCGCTGTTATG
GCCGCGTTTGTCTCATTCCACGCCTGACACTCAGTTCCGGGTAGGCAGTTCGCTCCAAGCTGGACTGTATGCACGAACCCCCC
GTTCAGTCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGAAAGACATGCAAAAGCACCACTGGCAGC
AGCCACTGGTAATTGATTTAGAGGAGTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAAAGGACAAGTTTTGGTGAC
TGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGCTCAGAGAACCTTCGAAAAACCGCCCTGCAAGGCGGTTTT
TTCGTTTTCAGAGCAAGAGATTACGCGCAGACCAAAACGATCTCAAGAAGATCATCTTATTAATCAGATAAAATATTTCTAGA
TTTCAGTGCAATTTATCTCTTCAAATGTAGCACCTGAAGTCAGCCCCATACGATATAAGTTGTAATTCTCATGTTAGTCATGC
CCCGCGCCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTAATGAGTGAGCTAA
CTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACG
CGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTC
ACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAA
CGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATGTCCGCACCAACGCGCAGCCCGGACTCGG
TAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGC
ATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTT
ATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGA
CCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGA
AATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATGATCAGCCCACT
GACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGG
CACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCA
ATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTT
TTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGA
CATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTT
TTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATTAGGAAATTAATACGACTCACTATA

Table S2: Oligonucleotide primers used in this study. All sequences are provided 5′ to 3′. The suffices "f" and "r" indicate a forward or reverse primer, respectively. LP, leader peptide. CP, core peptide. SLIM, site-directed ligation-independent mutagenesis.

Primer Name	Oligonucleotide Sequence
MBP-f (sequencing)	GAGGAAGAGTTGGCGAAAGATCCACGTA
T7-r (sequencing)	GCTAGTTATTGCTCAGCGG
FusC_1 (sequencing)	TGGAACGGGTCAAACAG
FusA_BamHI-f	AAAGGATCCATGGAAAAGAAGAAGTACACCGCTCCG
FusA_NotI-r	AAAGCGGCCGCTCAGATGAAGCGCGGGAAGAC
FusE_BamHI-f	AAAGGATCCATGGAGACAACAGGAGCTGAATTCCGG
FusE_NotI-r	AAAGCGGCCGCTCATGGCAGCGCCATCCCG
FusB_HindIII-f	AAAAAGCTTGATGAGCGAGAACGTAGTGCTGC
FusB_NotI-r	AAAGCGGCCGCTCAGTCCACGGTGATCAGACGG
FusC_HindIII-f	AAAAAGCTTGATGGTCGGTTGCATCAGTCCTTACTTC
FusC_NotI-r	AAAGCGGCCGCTCAGCTCCTGTTGTCTCCACCG
FusCEB_BamHI-f (Gibson Assembly)	CCATCACCATCATCACCACAGCCAGGATCCAATGGTCGGTTGCATCAG
FusCEB_NotI-r (Gibson Assembly)	CTTTCTGTTCGACTTAAGCATTATGCGGCCGCTCAGTCCACGGTGATCAG
FusB-C79A-f	GCCCTGCAGCGTTCGCTGGCTATCGCACTG
FusB-C79A-r	GGCCCGCAGTCCGGCGCAGCG
FusB-H112A-f	GCCGCCTGGGTGGAGGCCGAAGG
FusB-H112A-r	TCCGATGAACGGCGGGCGGC
FusC-D238A-f	GCCTCCACTCCCCTCTGCGCCCTGG
FusC-D238A-r	GAGCCCGCCGGACAGGTCGG
FusC-D340A-f	GCCCAGCTCCTCACCGGATCGCTCAG
FusC-D340A-r	GCCGCCCAGGCCGTCCATGTG
FusA_Core_1_Mutant_Screen-r	GCCGGTGGCCTCCTTGAATTCGCCGACCT
FusA-W1F-f	AGGAGGCCACCGGCTTCTACACCGCGGAATGGGGCC
FusA-W1L-f	AGGAGGCCACCGGCCTGTACACCGCGGAATGGGGCC
FusA-W1A-f	AGGAGGCCACCGGCGCGTACACCGCGGAATGGGGCC
FusA-W1P-f	AGGAGGCCACCGGCCCGTACACCGCGGAATGGGGCC
FusA-W1K-f	AGGAGGCCACCGGCAAATACACCGCGGAATGGGGCC
FusA-W1H-f	AGGAGGCCACCGGCCACTACACCGCGGAATGGGGCC
FusA-W1Y-f	AGGAGGCCACCGGCTATTACACCGCGGAATGGGGCC
FusA-W1E_SLIM-r	TTCGCCGGTGGCCTCCTTGAATTCGCCG
FusA_SLIM-r	GCCGGTGGCCTCCTTGAATTCGCCG
FusA-W1E_SLIM-f	GAATACACCGCGGAATGGGGCCTCGAGC
FusA_SLIM-f	TACACCGCGGAATGGGGCCTCGAGC
pET28-MBP-FusA(LP)-CitA(CP)_ins-r	ACGAGACGGTCGTTGCCCGCCAGGCCCAGCAGGCCGGTGGCCTCCTT
pET28-MBP-FusA(LP)-CitA(CP)_ins-f	AACGACCGTCTCGTCCTCAGCAAGAACTAAGCGGCCGCACTCGAGCACCA
pET28-MBP-FusA(LP)-CitA(CP)_D8E-r	GACGAGACGCTCGTTGCCCGCCAGGCCCAGCAG
pET28-MBP-FusA(LP)-CitA(CP)_D8E-f	CGAGCGTCTCGTCCTCAGCAAGAACTAAGCGGCCGC
FusA(LP)-CitA(CP)-r	GCCGGTGGCCTCCTTGAATTCGCCGACCT
pET28-MBP-FusA(LP)-CitA(CP)_L1W	AAGGAGGCCACCGGCTGGCTGGGCCTGGCGGGCAA

pET28-MBP-FusA(LP)-CitA(CP)_Ring+1-f	AGGCCACCGGCTGGCTGGGCCTGGCGGCGGGCAACGAGCGTCTCGTCCTC
pET28-MBP-FusA(LP)-CitA(CP)_Ring-1-f	GGAGGCCACCGGCTGGCTGGGCCTGGGCAACGAGCGTCTCGT
pET28-MBP-FusA(Ring+0)-r	GAGGCCCCATTCCGCGGTGTACCAGCCGGTGGCCT
pET28-MBP-FusA(Ring+1)-r	GAGGCCCCATTCCGCCGCGGTGTACCAGCCGG
pET28-MBP-FusA(Ring-1)-r	GAGGCCCCATTCGGTGTACCAGCCGGTGGCCTCC
pET28-MBP-FusA(Ring-2)-r	GAGGCCCCATTCGTACCAGCCGGTGGCCTCCTTG
pET28-MBP-FusA_Core_L10*-f	CCGCGGAATGGGGCCTCGAGTGAATCTTCGTCTTCCCGCGCTTC
pET28-MBP-FusA_Core_R16*-f	GGAATGGGGCCTCGAGCTGATCTTCGTCTTCCCGTGATTCATCTGAGCG
pET28-MBP-FusA_RingEx_E-f	GAATGGGGCCTCGAGCTGATCTTCGTCTTCCCGCGCTTCATC
pET28-MBP-FusA_RingEx_D-f	GAATGGGGCCTCGATCTGATCTTCGTCTTCCCGCGCTTCATC
pET28-MBP-FusA-I40G-f	CCCGCGCTTCGGTTGAGCGGCCGCACT
pET28-MBP-FusA-I40G-r	ACCGAAGCGCGGGAAGACGAAGATCAGCT
pET28-MBP-FusA-F39G-f	CCCGCGCGGTATCTGAGCGGCCGCACTCG
pET28-MBP-FusA-F39G-r	ATACCGCGCGGGAAGACGAAGATCAGCTCGAGGC
FusA-T-2A-r	GGCCTCCTTGAATTCGCCGACCTTAGCG
FusA-T(-2)A W1*-f	GCCGGCTGATACACCGCGGAATGGGG
FusA-K(-10)-r	AGCGAGCTGCGGAGCGGTG
FusA-K(-10)A-f	GCGGTCGGCGAATTCAAGGAGGCCAC
FusA-K(-10)E-f	GAGGTCGGCGAATTCAAGGAGGC
FusA-L(-12)-r	CTGCGGAGCGGTGTACTTCTTCTTTTCC
FusA-L(-12)M-f	ATGGCTAAGGTCGGCGAATTCAAGGAGGC
FusA-L(-12A)A-f	GCCGCTAAGGTCGGCGAATTCAAGGAGG
FusA-W1*-f	GGCCACCGGCTGATACACCGCGG
FusA-W1-r	TCCTTGAATTCGCCGACCTTAGCGAGCTG
FusA-P(-14)-r	AGCGGTGTACTTCTTCTTTTCCATGGATCCG
FusA-P(-14)A-f	GCGCAGCTCGCTAAGGTCGGC
FusA-P(-14)L-f	CTGCAGCTCGCTAAGGTCGGCG
FusA-Y(-17)-r	CTTCTTCTTTTCCATGGATCCGGATTGGAAG
FusA-Y(-17)A-f	GCCACCGCTCCGCAGCTCGC
FusA-Y(-17)D-f	GACACCGCTCCGCAGCTCGC
FusB-C79A-f	GCCCTGCAGCGTTCGCTGGCTATCGCACTG
FusB-C79A-r	GGCCCGCAGTCCGGCGCAGCG
FusB-E116A-f	GCGGCCGAAGGACGCCTCGTG
FusB-E116A-r	CACCCAGGCGTGTCCGATGAAC
FusB-E123A-f	GCAGAAGGCGTCGGCTACGACTACTTC
FusB-E123A-r	CACGAGGCGTCCTTCGGCC
FusB-E124A-f	GCAGGCGTCGGCTACGACTACTTCAG
FusB-E124A-r	TTCCACGAGGCGTCCTTCGGC
FusB-H112A-f	GCCGCCTGGGTGGAGGCCGAAGG
FusB-H112A-r	TCCGATGAACGGCGGGCGGC
FusB-W114A-f	GCGGTGGAGGCCGAAGGACGCC

FusB-W114A-r	GGCGTGTCCGATGAACGGCG
FusC-D238A-f	GCCTCCACTCCCCTCTGCGCCCTGG
FusC-D238A-r	GAGCCCGCCGGACAGGTCGG
FusC-D340A-f	GCCCAGCTCCTCACCGGATCGCTCAG
FusC-D340A-r	GCCGCCCAGGCCGTCCATGTG
FusC-G336A-f	GCCCTGGGCGGCGACCAGCTCCTC
FusC-G336A-r	GTCCATGTGCAGGCGGGAGCCG
FusC-S234A-f	GCCGGCGGGCTCGACTCCACTC
FusC-S234A-r	CAGGTCGGCGCTAACCACTCCG
FusC-S239A-f	GCCACTCCCCTCTGCGCCCTGG
FusC-S239A-r	GTCGAGCCCGCCGGACAGGTC
FusE-D74-r	GCGTTCCGCGTCGGTCCG
FusE-D74A-f	GCCATCGCCGCGCTCGTCACC
FusE-E32-r	ACCGCTGCGGCCGTCC
FusE-E32A-f	GCGTACTGGCAGCTCAACGACACCGC
FusE-E32K-f	AAGTACTGGCAGCTCAACGACACCGC
FusE-E63-r	AACTGGGCCACGTCGGCG
FusE-E63A-f	CCTCACCAGCGACTACGAGGTGGAG
FusE-E63K-f	CCTCACCAGCAAGTACGAGGTGGAGCG
FusE-G31-r	GCTGCGGCCGTCCAG
FusE-G31A-f	GCTGAGTACTGGCAGCTCAACG
FusE-G31V-f	GTTGAGTACTGGCAGCTCAACG
FusE-L26-r	CAGCACCATGCCGTAGTCG
FusE-L26A-f	GCGGACGGCCGCAGC
FusE-L78A-f	GCCGTCACCAGCCTGAAAGAGAACGG
FusE-L78P-f	CCCGTCACCAGCCTGAAAGAGAACGG
FusE-L78-r	CGCGGCGATGTCGCGTTCC
FusE-M87-r	CCCGTTCTCTTTCAGGCTGGTGACG
FusE-M87A-f	GCGGCGCTGCCATGAGCGGC
FusE-M87L-f	CTGGCGCTGCCATGAGCGGC
FusE-V24-r	CATGCCGTAGTCGGTTTGCGC
FusE-V24A-f	GCCCTGCTGGACGGCCGCAGC
FusE-Y33-r	CTCACCGCTGCGGCCGTC
FusE-Y33A-f	GCCTGGCAGCTCAACGACACCGCG

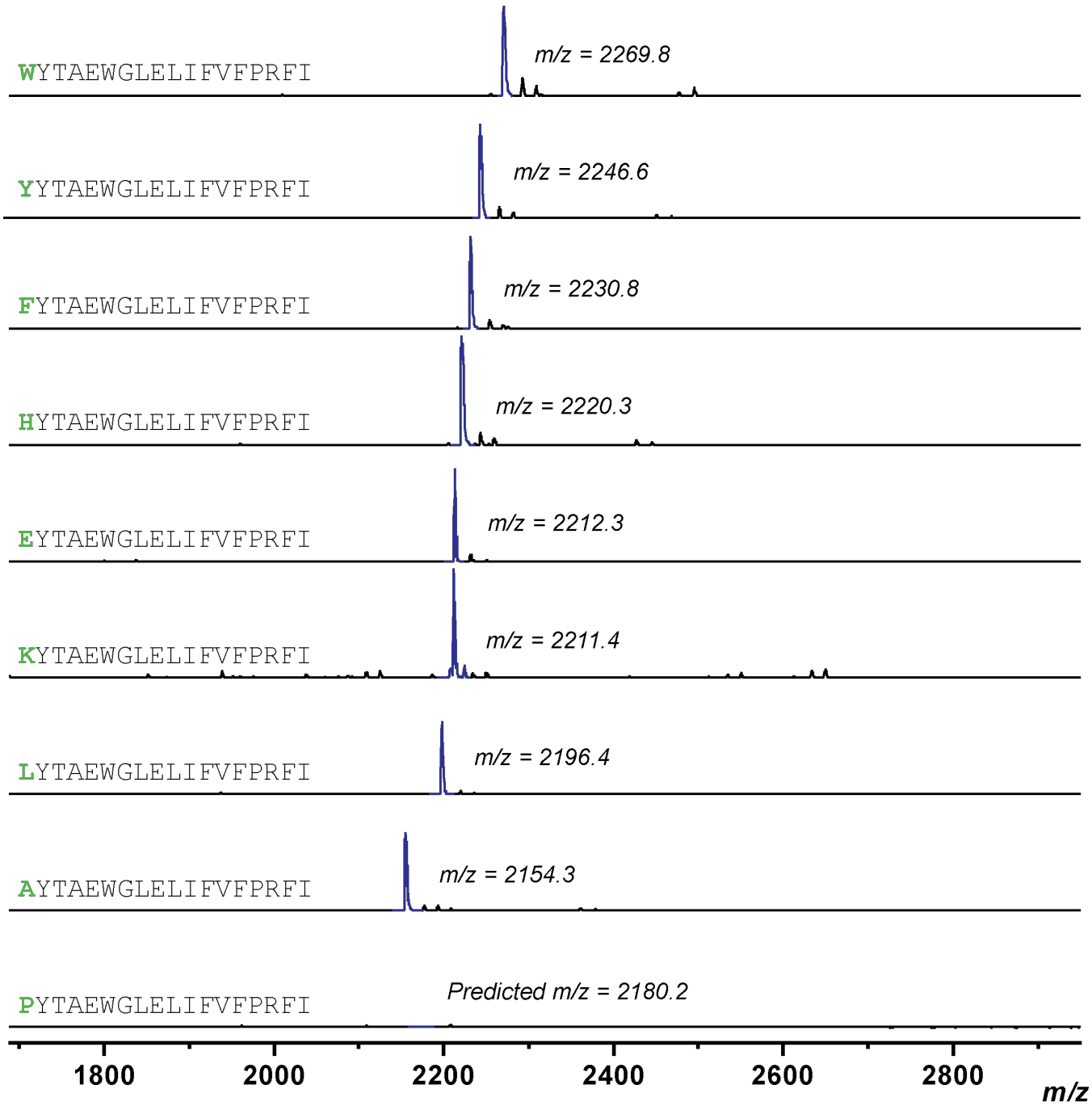


Figure S1. MALDI-TOF-MS of methanolic extracts of heterologously produced wild-type fusilassin and core position 1 variants. Precursor peptides were co-expressed with a pACYC-Duet vector FusB, FusC, and FusE. Crude methanolic extracts were analyzed by MALDI-TOF MS and masses corresponding to mature, cyclized products are indicated in blue. Only the W1P variant failed to generate a lasso peptide product.

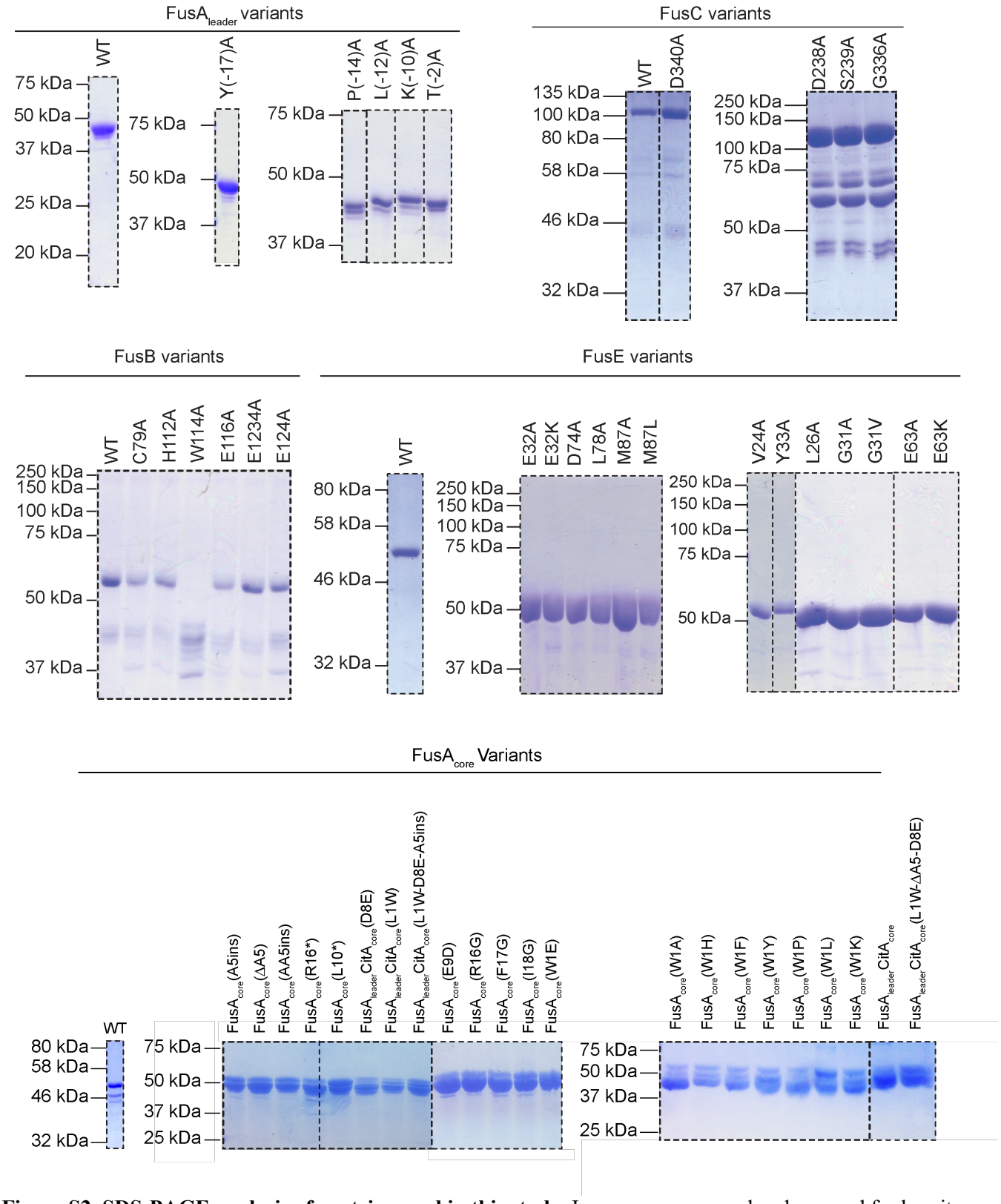
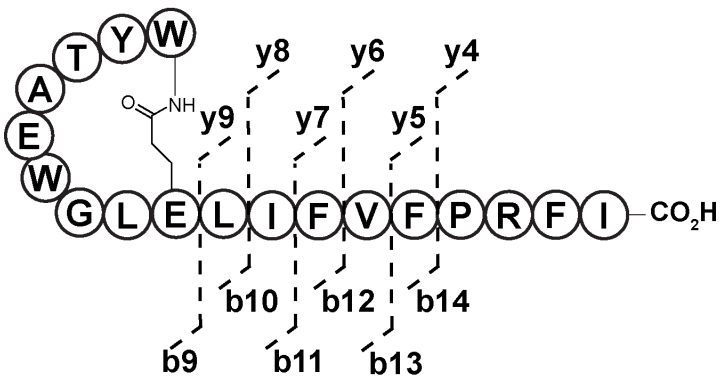


Figure S2. SDS-PAGE analysis of proteins used in this study. Images were merged and cropped for brevity – cropped and trimmed edges are indicated by dashed lines.



Species	Predicted mass	Observed mass	Error (ppm)
$[M+2H]^{2+}$	1135.5976	1135.5960	1.4
b9	1118.4942	1118.4923	1.7
b10	1231.5782	1231.5763	1.6
b11	1344.6623	1344.6598	1.9
b12	1491.7307	1491.7286	1.4
b13	1590.7991	1590.7966	1.6
b14	1737.8675	1737.8646	1.7
y4	532.3242	532.3234	1.5
y5	679.3926	679.3917	1.4
y6	778.4610	778.4598	1.5
y7	925.5294	925.5279	1.7
y8	1038.6135	1038.6114	2.0
y9	1151.6976	1151.6958	1.5

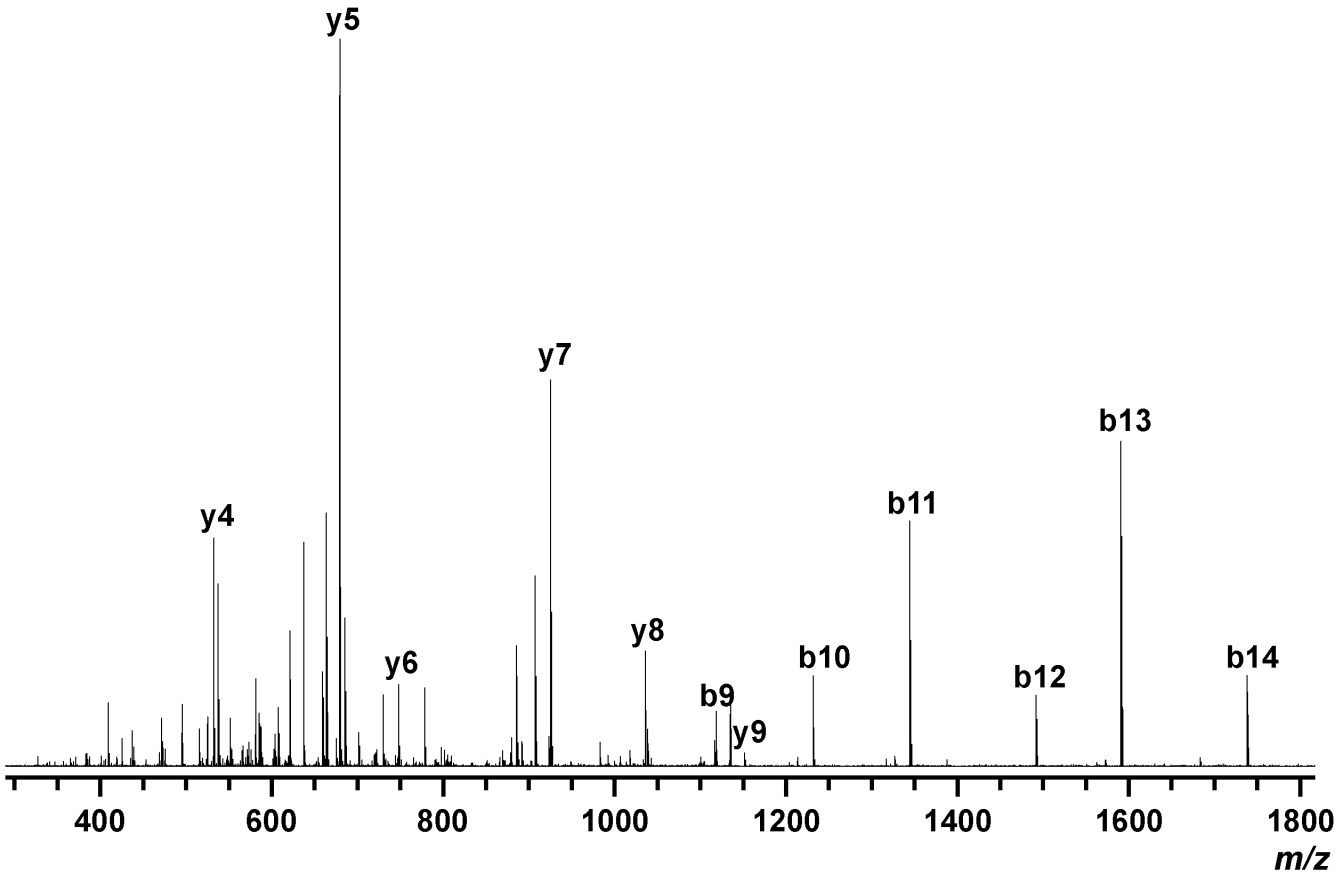


Figure S3. HR-ESI-MS/MS of biosynthetically reconstituted fusilassin. Fusilassin was subjected to collision-induced dissociation, resulting in a series of b^+ and y^+ ions consistent with a macrolactam linkage between Trp1 and Glu9 of the fusilassin core peptide.

B protein accession	Organism	Lasso peptide
WP_011291590.1 (FusB)	*Thermobifida fusca*	fusilassin
WP_040732796.1	*Paenibacillus dendritiformis* C454	paeninodin
BAL72549.1	*Rhodococcus jostii*	lariatin
EWS93467.1	*Streptomyces filamentosus* NRRL 11379	SRO15-2005
KIH99639.1	*Streptomonospora alba*	streptomonomicin
WP_031079704.1	*Streptomyces sp.* NRRL S-118	lagmysin
OAL15485.1	*Streptomyces albulus* NRRL B-3066	citrulassin A
WP_052021341.1	*Actinokineospora spheciospongiae*	actinokineosin
WP_017534153.1	*Nocardiopsis alba* DSM 43377	LP2006
WP_078648027.1	*Streptomyces leeuwenhoekii*	chaxapeptin
EDY58509.1	*Streptomyces sviceus* ATCC 29083	sviceucin
WP_014140854.1	*Streptomyces cattleya* NRRL 8057	moomysin

```
                                                    *
WP_011291590.1      47     -RGEVRPATYEEAKAARDAVLAVSLRCAGLRACLQRSLAIALLCRMRGTWATWC--VGVP   104
WP_040732796.1      52     TPEVHDPSDTKTMKHIRSAIHIASRHTFWESKCLVRAIAGMKMLEKRNLHSTLYLGTGKN   112
BAL72549.1          48     -AAKARPASYREAEAAVVSITQYSKASAGPGSCLQRSISVCILMRLDGRWPTWC--VGVP   105
EWS93467.1          26     -ARGARPASYAETLEVLEAVTATSRRCASRYGCLPRSVAVALACRMSGVWPDWC--AGVW    83
KIH99639.1          46     -GRGGRAATPETVSSCRRAVVSVSARCAGE-GCVPRSVATFLLARSYGSVSSWR--AGIR   102
WP_031079704.1      47     -RKGARPATLAEAARARETVLAVSLAAGGQQGCLARSLATVLLCRLRGQWPTWC--VGVR   104
OAL15485.1          46     -RRGARPATAGRAQAARDAVCAVSLRCAGPRGCLPRSLGAALLCRLTGTWPTWC--TGVS   103
WP_052021341.1      31     -HRGARPASAELTQRARDTVCATSARCTGQ-GCLPRSVAVVVYCRLHGRAPDWC--TGFA    87
WP_017534153.1      16     -AGSAGPAGFAEALAARNAVVHVSRRCAGP-GCLQRSIATFLLCRIRGTVPEWC--TGVR    72
WP_078648027.1      46     -RRGAAPATVEQAMAARQAVVAVSQRCAGQ-ACLQRSIATALLCRARGSWPTWC--TGVR   102
EDY58509.1          46     -SRGSRPADAERALAARQAVVTVSLRCAGIAGCLQRSVATALLCRLAGRWPDWC--SGFR   103
WP_014140854.1      46     -CRGARPAGAESAARAREAVVAVSLRCAGP-YCLQRSVATALLCRLSGQWPDWC--TGVR   102
                                   :          ::   *        *: *::.     .  .         *

                                  *
WP_011291590.1             RRPPFIGHAWVEAEGRLVEEGVGY-DYFSRLITVD-------------------------   138
WP_040732796.1             KNGELIAHAWLRSGSIYITGAEMM-KQFVVVEKFAKSAGTDS------------------   153
BAL72549.1                 SKPPFRAHAWIEAGGQIVAELGDM-NSYSRLMTISTHAERTES-----------------   147
EWS93467.1                 ATPPFSPHAWVLAEGRTVGEQAEA-ADLRPLMVVTVREGVQGERGGGDDG-----GGGTG   137
KIH99639.1                 P-NPFLAHAWVEVDGVPVDEPEAV-RDFIPTLKVDPGSARPSA-----------------   143
WP_031079704.1             TQPPFAAHAWVEAEGVLVGEGAPA-DYFQRFFTVD-------------------------   138
OAL15485.1                 VVPPFAAHAWIEAEGRPVGEPMPD-DYFTRLLAVGPASPPGGAPHH--------------   148
WP_052021341.1             L-DPFTAHAWVEVDGAPVGEPPKV-AEYIISQAVRLRSPAAA------------------   127
WP_017534153.1             V-SPFRAHAWVEVDGSPVGELGEV-RHFHKVMTVRPVIVVRDDD----------------   114
WP_078648027.1             T-SPFEAHAWIEAEGRLIGEPYPD-GHYRPLLTVPPVS----------------------   138
EDY58509.1                 T-RPFGAHAWVEVDGTAIGEPGDM-TLFHTVLSVRHQDRDQHLHQGRRQARRQARAGRHE   161
WP_014140854.1             T-EPFRAHAWVEAEGVPVGEDAAGLRGFRVLLRVSASGRRA------------------   142
                               :  ***:   .  :              .
```

Figure S4. Sequence alignment of various discrete leader peptidases. A multiple sequence alignment of 12 discrete leader peptidases was generated using the Clustal Omega[4] web tool (https://www.ebi.ac.uk/Tools/msa/clustalo/) with the default parameters. Residues marked with a red asterisk were probed by mutagenesis owing to their potential role in catalysis (see Fig. S5 for MALDI-TOF-MS characterization). FusB-W114A (blue position) caused severe degradation in the protein after purification from *E. coli* (See Fig. S2).

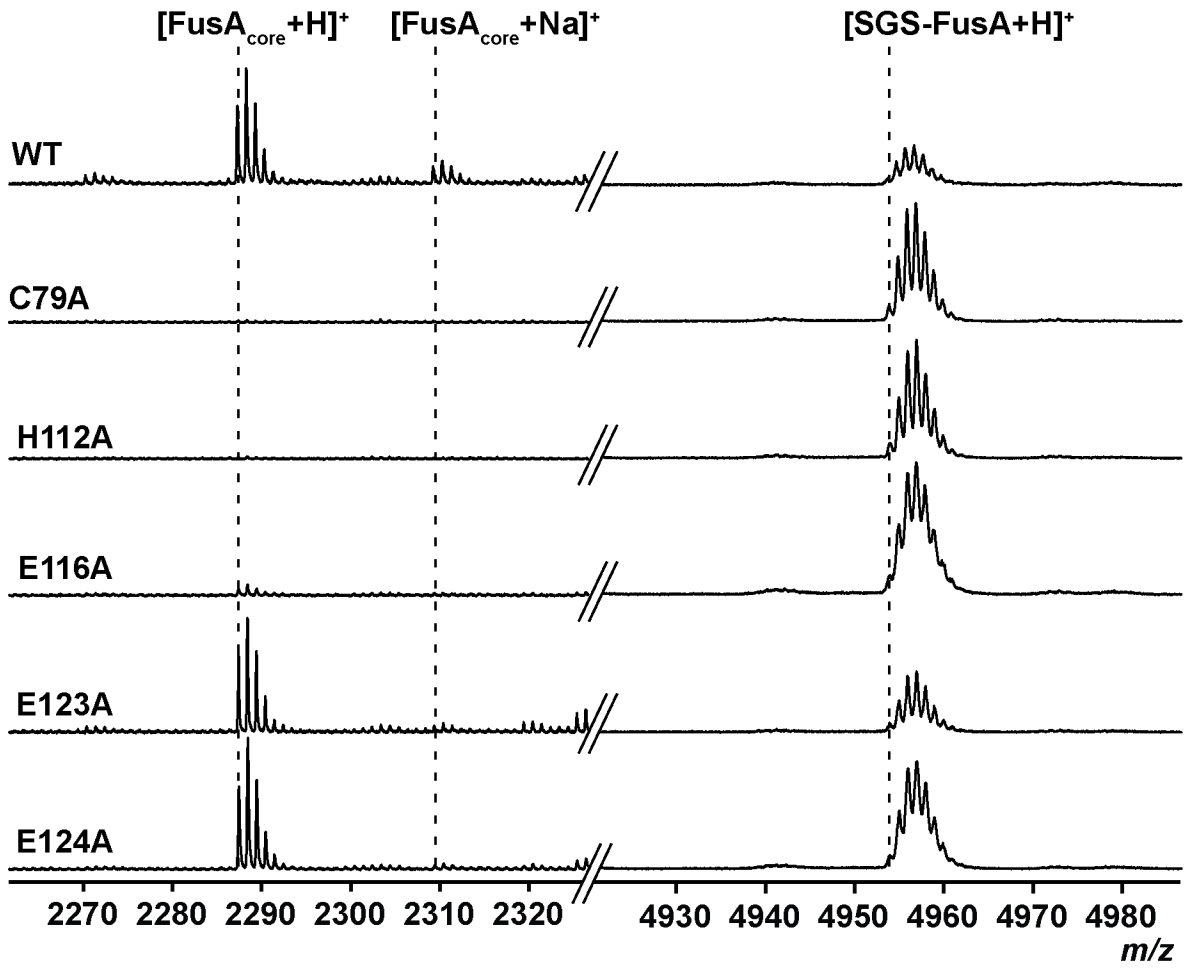


Figure S5. MALDI-TOF-MS characterization of FusB variants. Ala point substitutions of FusB were assayed for *in vitro* leader peptidase activity. The two predicted catalytic variants, C79A and H112A, are devoid of leader peptidase activity. The FusB-E116A variant displays considerable catalytic impairment while the other two variants, E123A and E124A, retain wild-type-like activity as observed by end-point assay. The "SGS" sequence remains on the N-terminus of the FusA precursor peptide as a result of cloning and TEV protease processing.

C protein accession	Organism	Lasso peptide
WP_104612995.1 (FusC)	*Thermobifida fusca*	fusilassin
WP_006678396.1	*Paenibacillus dendritiformis* C454	paeninodin
BAL72547.1	*Rhodococcus jostii*	lariatin
EWS93465.1	*Streptomyces filamentosus* NRRL 11379	SRO15-2005
KIH99641.1	*Streptomonospora alba*	streptomonomicin
WP_052411817.1	*Streptomyces sp.* NRRL S-118	lagmysin
WP_003988505.1	*Streptomyces albulus* NRRL B-3066	citrulassin A
WP_035284309.1	*Actinokineospora spheciospongiae*	actinokineosin
WP_017534152.1	*Nocardiopsis alba* DSM 43377	LP2006
WP_047121368.1	*Streptomyces leeuwenhoekii*	chaxapeptin
EDY58507.2	*Streptomyces sviceus* ATCC 29083	sviceucin
WP_014140852.1	*Streptomyces cattleya* NRRL 8057	moomysin

```
                                        *
FusC                222    RTRGGGVVSADLSGGLDSTPLCALAARGPAK------VVAL---------TFSSGLDTDD   266
BAL72547.1          221    QLSDHSAASCDISGGLDSSSIAAIAANAAKSGETHTVLHGT---------TSVSRDEFNS   271
KIH99641.1          178    RIQRRPNFSSDLSGGLDSTTLGFILAEDSVTRT--AGHHTL---------FLASVNRANR   226
EWS93465.1          184    HVGGLDRISCELSGGLDSTALAFAARDTGPAS-----LSLL---------TVAARDRYSE   229
WP_052411817.1      230    RVAGGGPVTADLG-GVDSTAVVCTAARQGAK------MVAY---------TAAVHDVLGD   273
WP_006678396.1      256    RLRTFRQVGAHLSGGLDSGSVVSFAAPMLKANN--KPLYTFSYYPVEGFSDFTSCNRFAD   313
WP_003988505.1      231    RTRQGGVVSCDLSGGLDSTSVCFLADRSAAH------VVAS---------TWPGRDPADT   275
WP_035284309.1       94    QSRSAGLVTADLSGGLDSTSICYALDDLGRR------FRAF---------RTSSTSPWND   138
WP_047121368.1      220    RNPVHGRLSADLSGGLDSTSLAFLAARTTPH------LLTH---------RWAEAEAGND   264
WP_017534152.1      219    RGRGRTTVSADLSGGLDSTGLCFVAAATGVP------LVTT---------RWQATDVADE   263
EDY58507.2           51    RTAGPGTVSADLSGGMDSTTLCFVAAAEGAD------LLTY---------HVEPLDRANE    95
WP_014140852.1      220    RTARGGVLSADLSGGLDSAGLCFLAADGPAR------LVTL---------HALAVDAGND   264
                           :        ..:. *:**  :

                                                                              *
FusC                288    PEDIPGFYAGLD---GEFPLLDEPSVAMLSTPRILSRLHTARAHGSRLHMDGLGGDQLLT   344
BAL72547.1          293    WNDMPKEYDDLDA--LASYDLDEPSIASISHSRFTHLINVARSKGSQVHLTGFGGDELFI   350
KIH99641.1          253    AL-----GTDA--APSVDLLPEGPGVSTVAIAATPRIRSVLKPTGSRIHLNGHGGDALFG   305
EWS93465.1          290    AGDAPYFYADLAS--TAAELNDEPLPVAPGRARAELLLGRAHATGSRYHLTGYGGDEIFL   347
WP_052411817.1      295    ATAVPLTFDGVH---ELDDALDTPSLYSVNRHRRMYIIERAAERGSSLHLSGMGGDELFS   351
WP_006678396.1      341    YSVIDEWL----------DTLEMPYKFFENSFWLKGIYEQAAARDIGVMLTGQRGNWTIS   390
WP_003988505.1      297    PDASPLVYEDLL---GIDDLLDEPTIGVMDRSRVLHHLPGLAARGSRLHLTGIGGDHVAW   353
WP_035284309.1      160    LSETSTAFTLD-ARADPVELAEGPVVWSASRGYLDVLAPAVAATGSRVHFTGLGGDELFD   218
WP_047121368.1      286    QSDIPPVFSDPH----LNGDPEQPYLFARTAARAVYTARMLVAHSSTRHIAGHGGDELFY   341
WP_017534152.1      285    RRSAPGWFTELLPRRPGVHDHEGPFPWIRTRGRLAHAAEAVADEGSRTHLTGHGADELFF   344
EDY58507.2          117    SQGASSWFDLPV---SDGHLGEGPLPWHGGRTHLEGLAHTVAEHGSRTHLTGLGGDELFG   173
WP_014140852.1      286    ADDGPRWFATPP---CADGDLEGPVVLARTRAKAAHAVAHAAREGSRIHLTGIGGDELFT   342
                                          : *                   .    : *  .:
```

Figure S6. Sequence alignment of various lasso cyclases. A multiple sequence alignment of 12 lasso cyclases was generated using the Clustal Omega web tool (https://www.ebi.ac.uk/Tools/msa/clustalo/) with the default parameters. Predicted ATP-binding residues probed by mutagenesis. The Asp residues also denoted with red asterisk were devoid of catalytic activity (see Fig. S7 for MALDI-TOF-MS characterization).

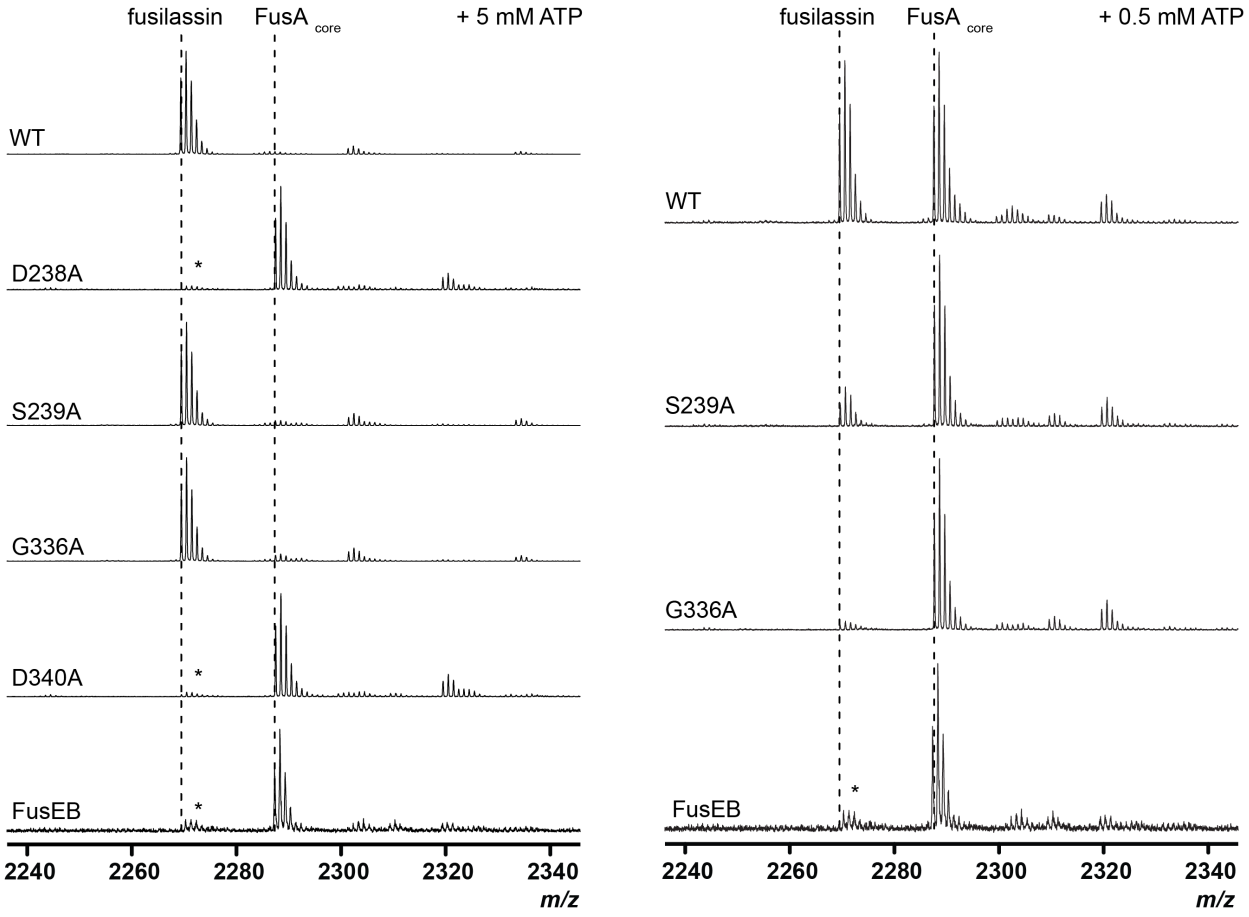


Figure S7. MALDI-TOF-MS characterization of FusC variants. Ala-substituted variants of FusC were assayed for *in vitro* lasso cyclase activity. FusC-D238A and -D340A were devoid of activity at 5 mM ATP concentration (*left*), while FusC-S239A and -G336A exhibited diminished turnover only at a reduced (500 μM) ATP concentration (*right*). The very low intensity mass marked with an asterisk in the FusB + FusE only controls (FusEB) and D238A and D340A samples is an unrelated contaminant (the observed mass is 0.7 Da distinct from fusilassin).

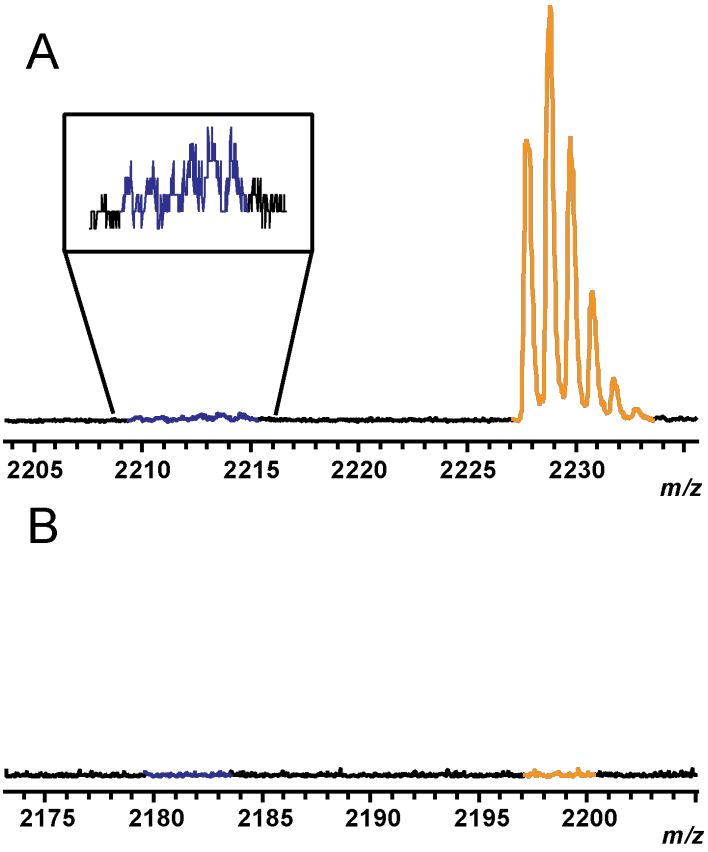


Figure S8. MALDI-TOF MS analysis of biosynthetic reconstitution of FusA variants. Expansion of the region corresponding to the linear (orange) and cyclized (blue) peptides. A peak consistent with the dehydrated form of FusA-W1E was observed at the limit of detection, as shown in the inset (A). FusA-W1P failed to generate linear core or the mature lasso peptide within the limit of detection (B).

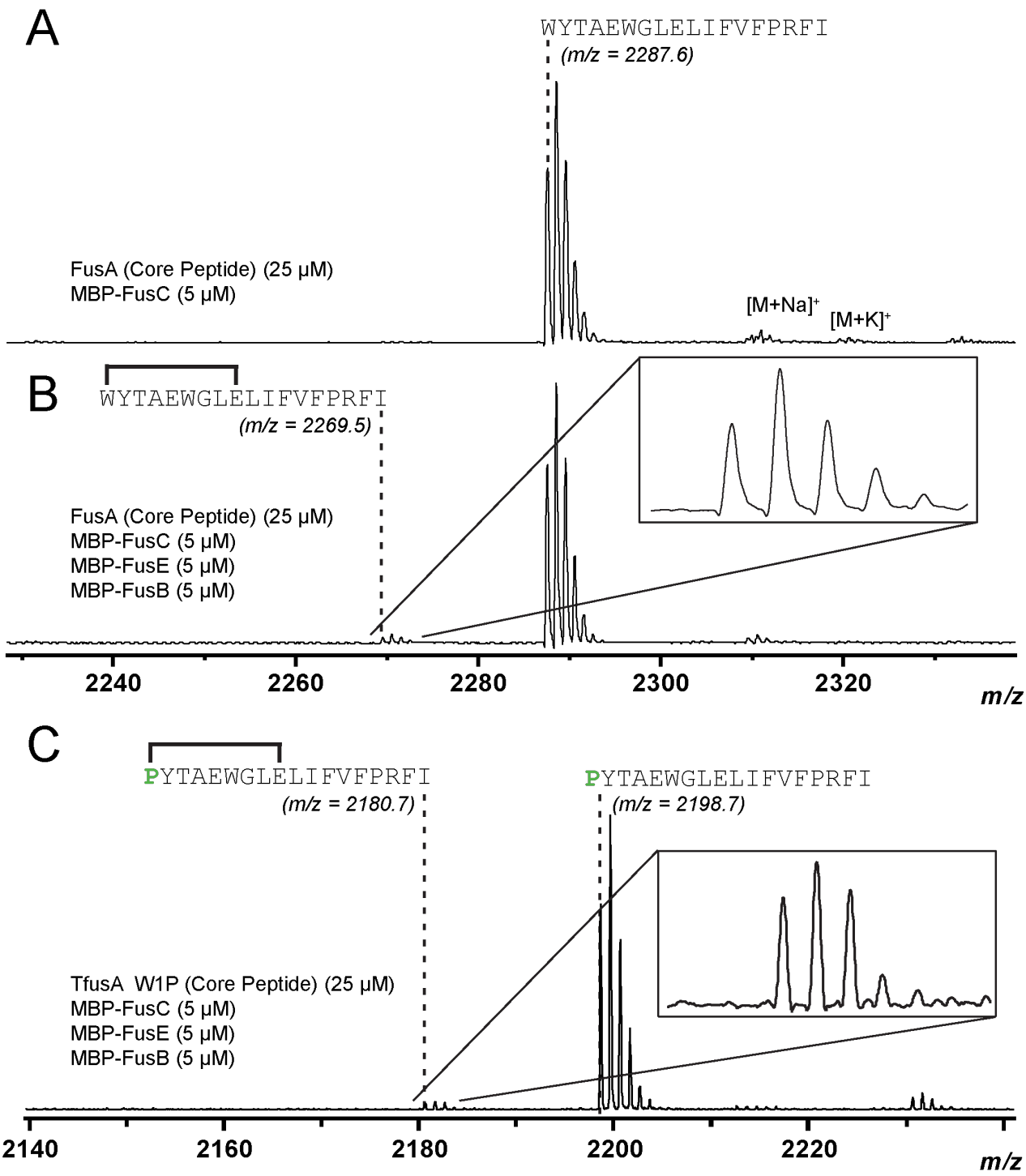


Figure S9. Biosynthetic reconstitution of synthetic core peptides. A synthetic core peptide that lacked the leader region of the native precursor peptide were not accepted as a substrate by FusC (A). Addition of FusB and FusE to the reaction led to a low but detectable amount of product formation. These data demonstrate that the leader region of the precursor peptide is critical for efficient processing by FusC (B). Same as panel B but synthetic FusA-W1P core peptide was employed as the substrate (C). Thus, FusC can tolerate a Pro1 variant of FusA. Insets are provided for low intensity peaks.

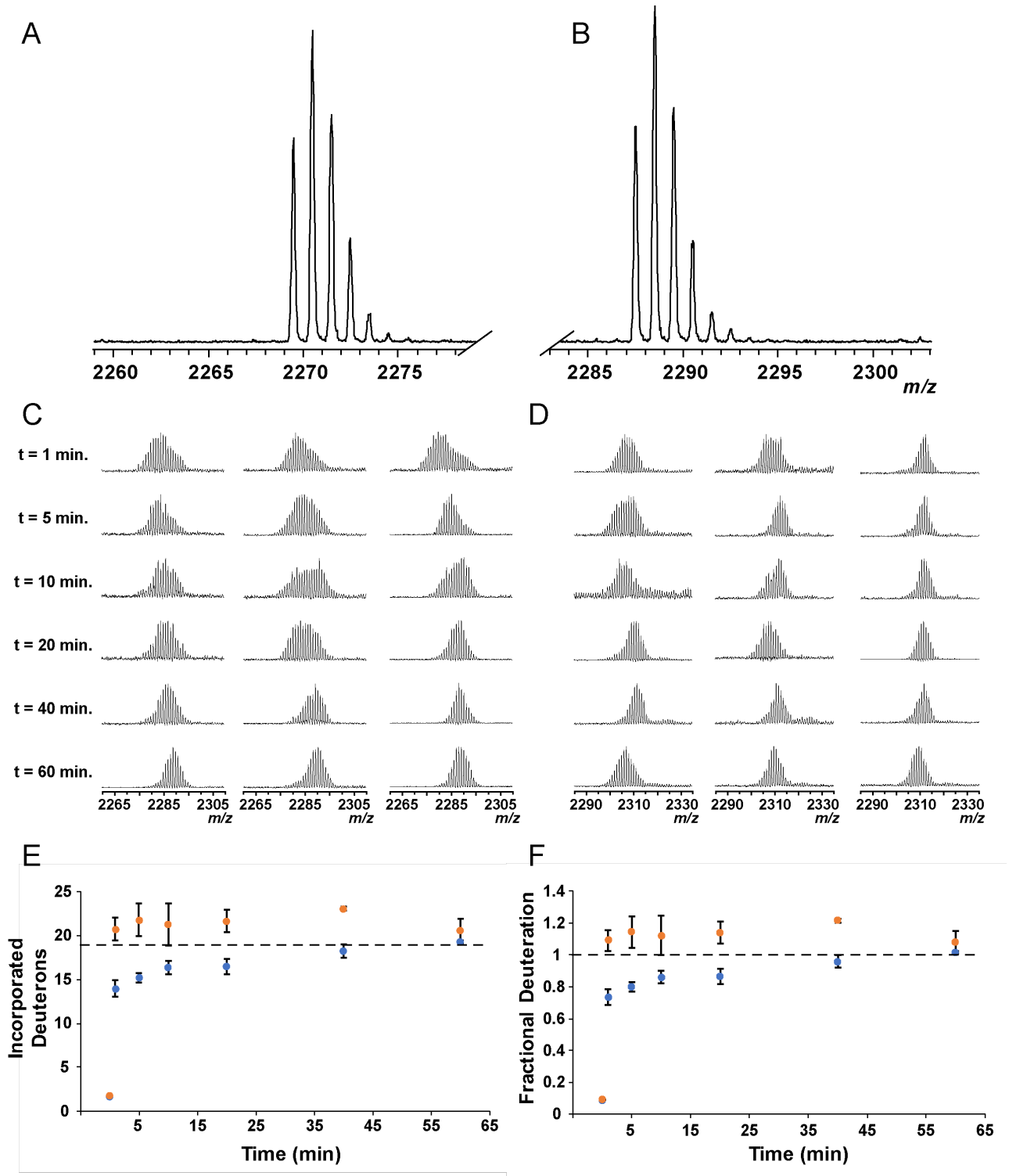


Figure S10. Hydrogen-deuterium exchange of biosynthetically reconstituted fusilassin. Centroid masses for natural abundance fusilassin (A) and linear core peptide (B) were calculated as described in methods and were used as single points for t=0. Raw data for fusilassin and linear core peptide deuteration vs. time was used to calculate centroid masses as described in methods (C and D, respectively). Full deuteration of the linear core peptide (orange circles) occurred within 1 min. Partial deuteration of fusilassin (blue circles) occurred rapidly, but the remaining exchangeable positions required an 1 h to reach the same level the control (E). Fractional deuteration rates are shown in (F) and support a threaded topology for fusilassin. Maximum fusilassin deuteration is indicated by the dotted line (panels E, F).

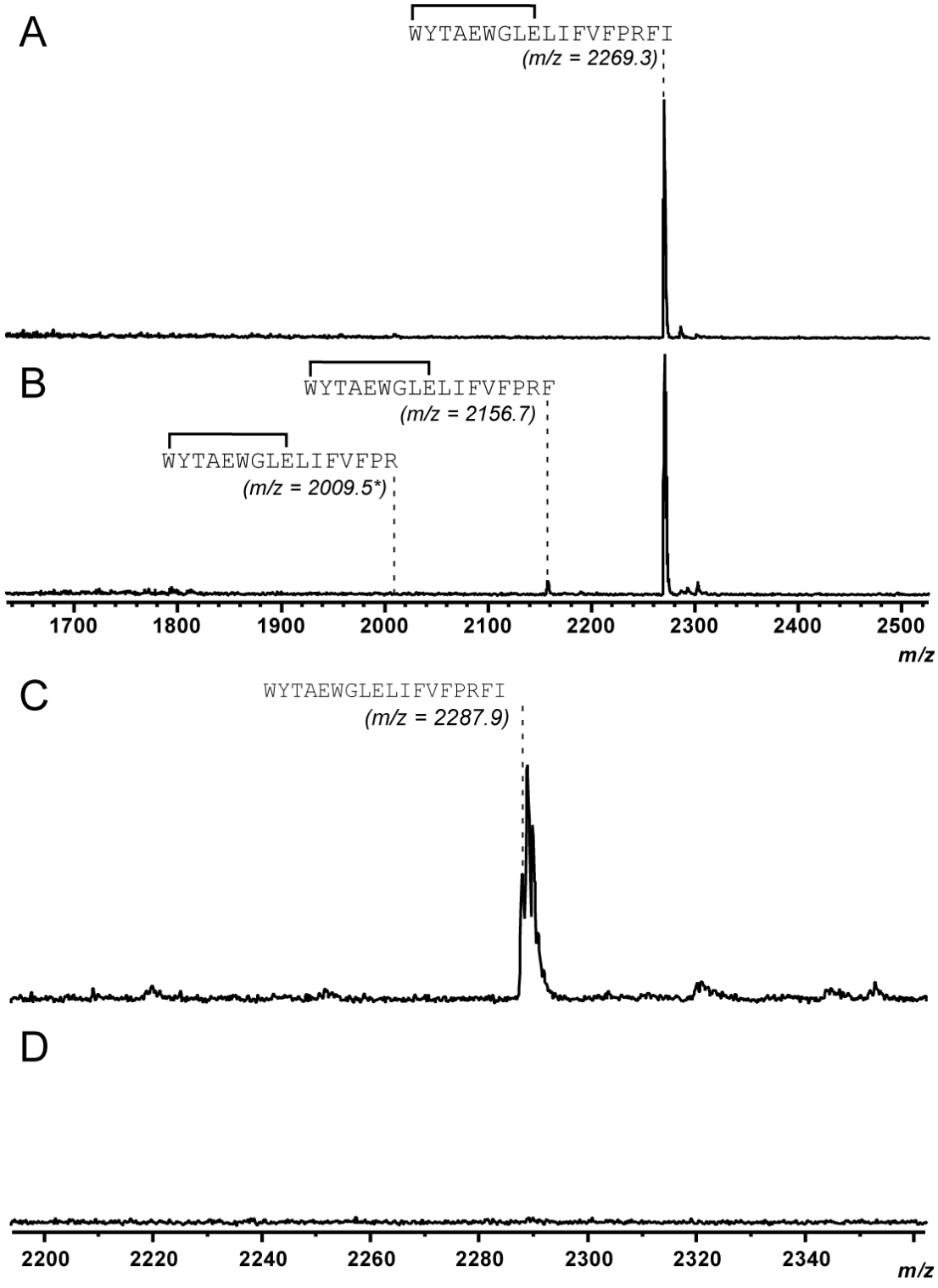


Figure S11. Proteolytic resistance of biosynthetically reconstituted fusilassin. Reconstituted fusilassin (A) was treated with carboxypeptidase Y for 18 h, yielding one additional mass corresponding to removal of the C-terminal residue Ile (*m/z* 2156.7 Da) while removal of both $Ile_{18}Phe_{17}$ (*m/z* 2009.5 Da, indicated by asterisk) was not observed (B). Identical treatment of linear core (C) led to complete proteolytic degradation, with no intermediate masses observed (D).

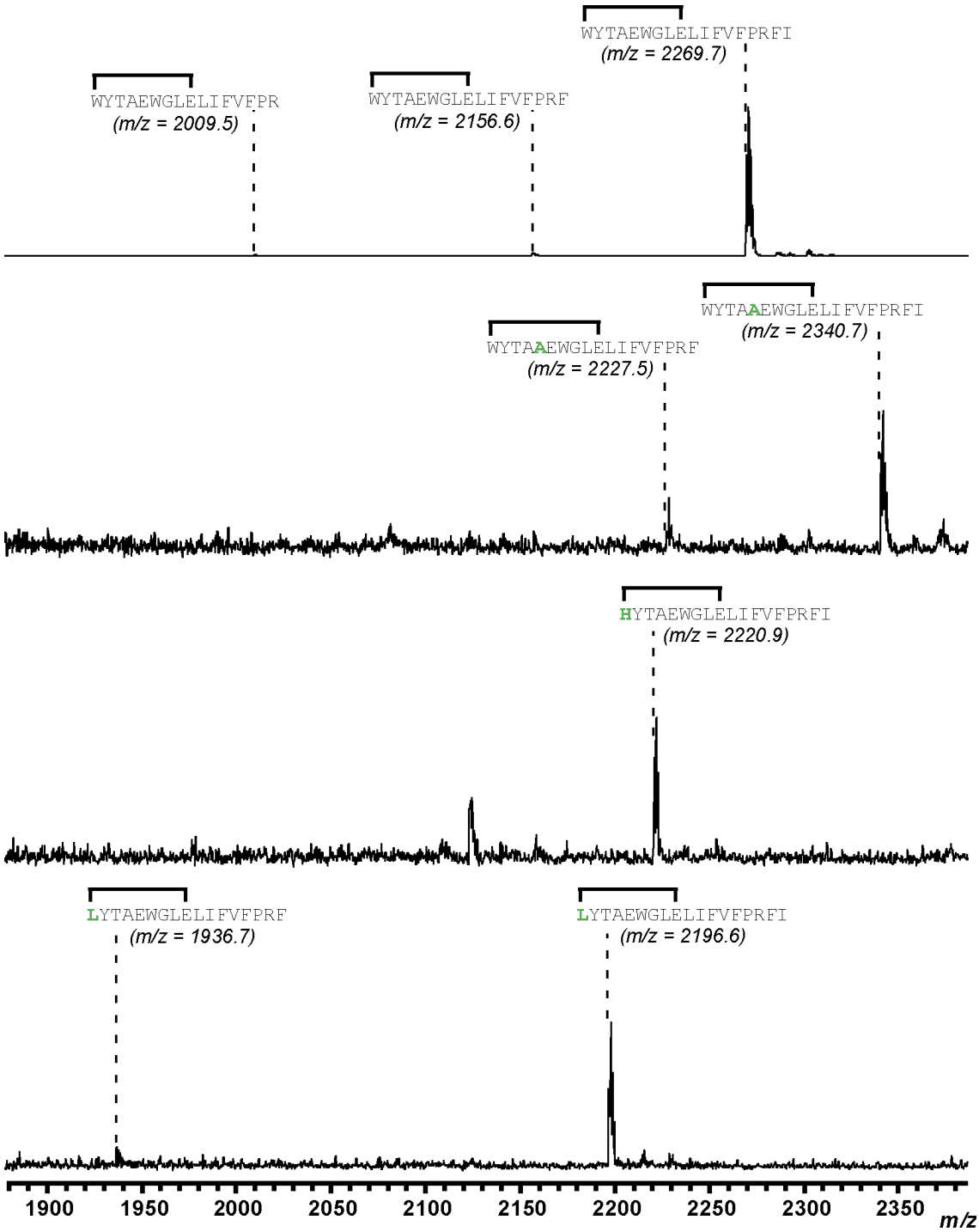


Figure S12. MALDI-TOF MS analysis of carboxypeptidase Y-treated biosynthetically reconstituted fusilassin variants. After heating at 95 °C for 2 h and a subsequent 18 h digestion with carboxypeptidase Y, all lasso peptides exhibited complete digestion of residual linear precursors (where appropriate) and incomplete proteolysis of dehydrated products, supporting the presence of a threaded conformation in all cases. No additional C-terminal proteolytic products were observed for any variant (non-wild-type residues are indicated in green).

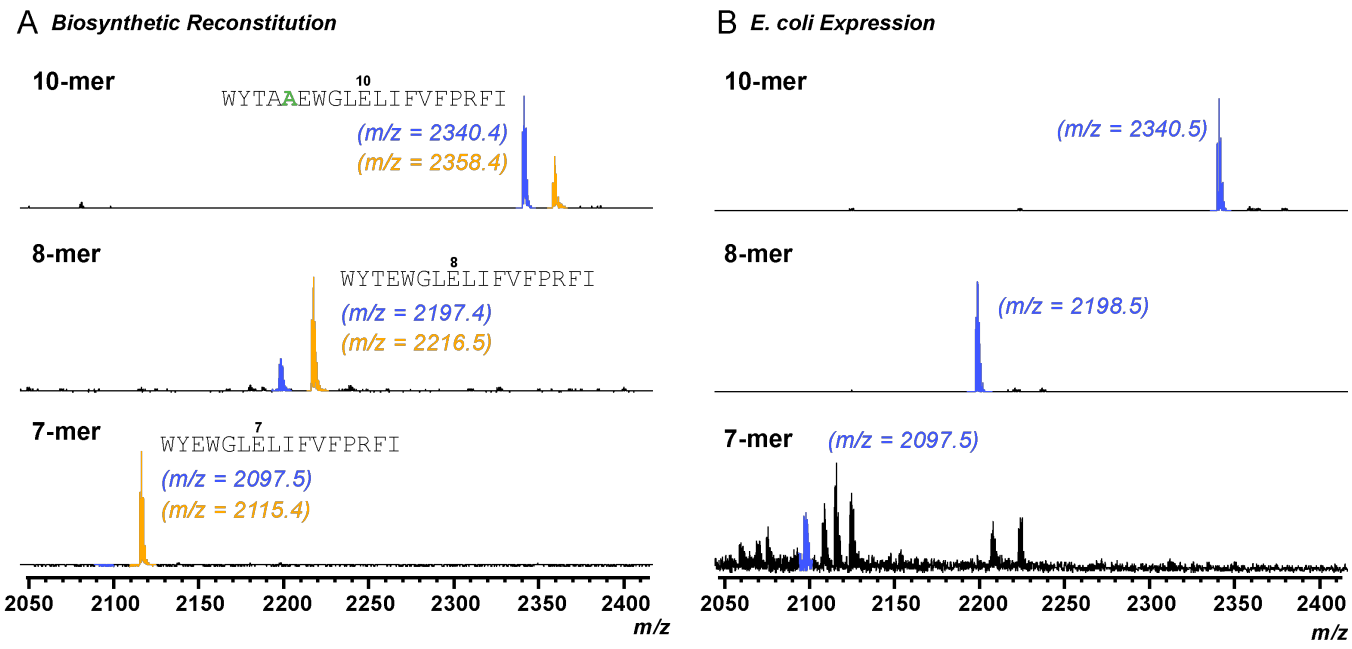
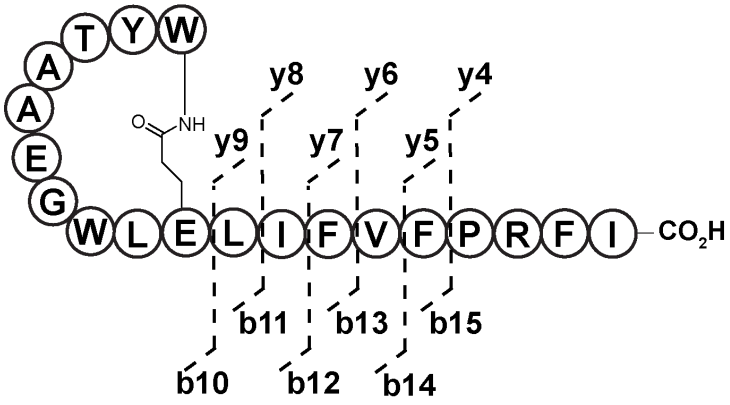


Figure S13. Enzymatic processing of FusA ring distortions. The fusilassin biosynthetic proteins (FusB, C, and E) were able to generate 10- and 8-mer macrolactam rings both by biosynthetic reconstitution (A) and *E. coli* expression (B). Contraction of the ring size to seven residues largely abrogated processing by the enzymatic machinery. Mature lasso peptides are indicated in blue while linear core peptide is indicated in orange.

A



Species	Predicted mass	Observed mass	Error (ppm)
$[M+2H]^{2+}$	1170.6144	1170.6115	2.5
b10	1189.5313	1189.5287	2.2
b11	1302.6154	1302.6123	2.3
b12	1415.6994	1415.6974	1.4
b13	1562.7678	1562.7649	1.9
b14	1661.8362	1661.8328	2.1
b15	1808.9047	1808.9004	2.4
y4	532.3242	532.3233	1.7
y5	679.3926	679.3914	1.8
y6	778.4610	778.4595	1.9
y7	925.5294	925.5276	2.0
y8	1038.6135	1038.6104	3.0
y9	1151.6976	1151.6967	0.8

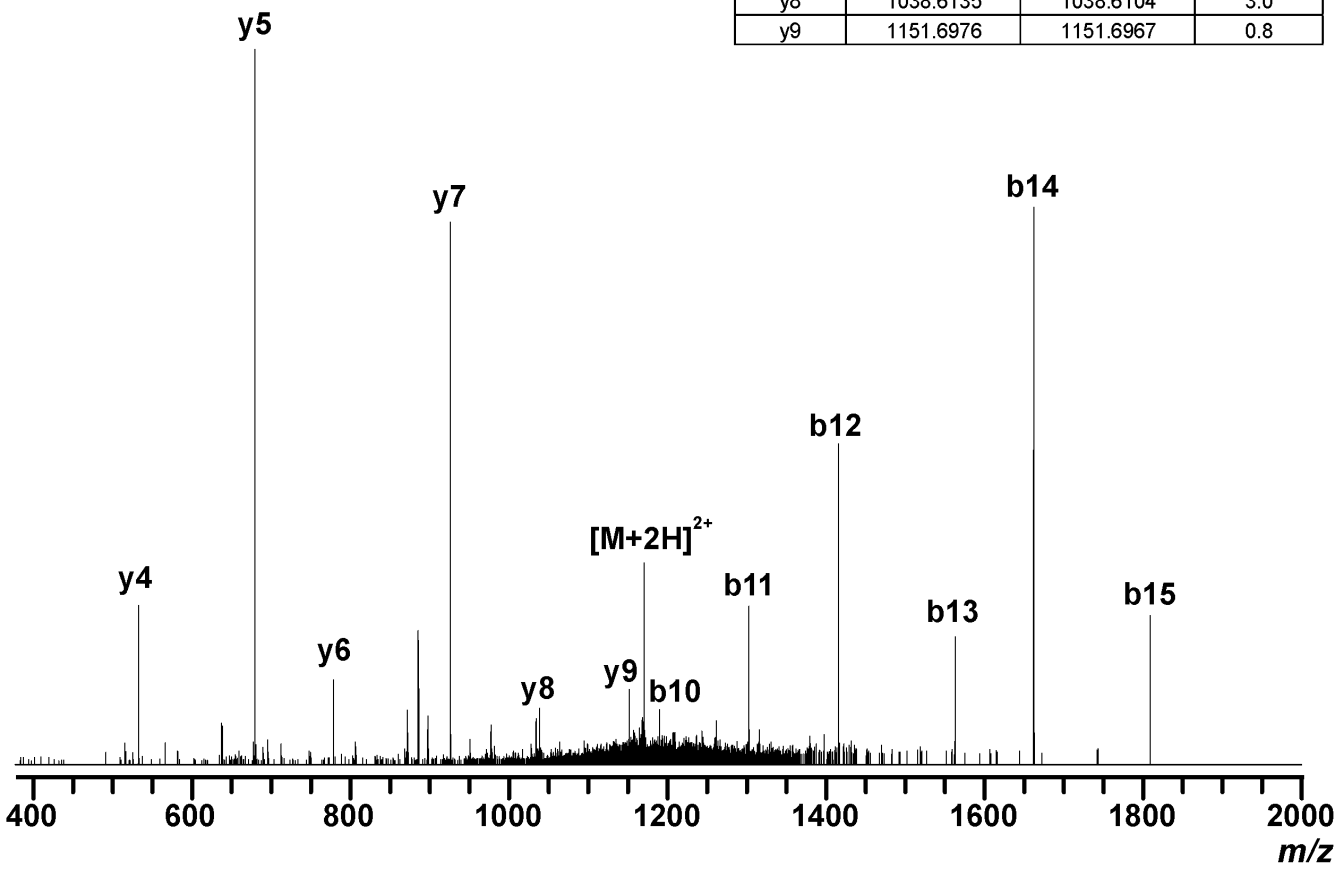
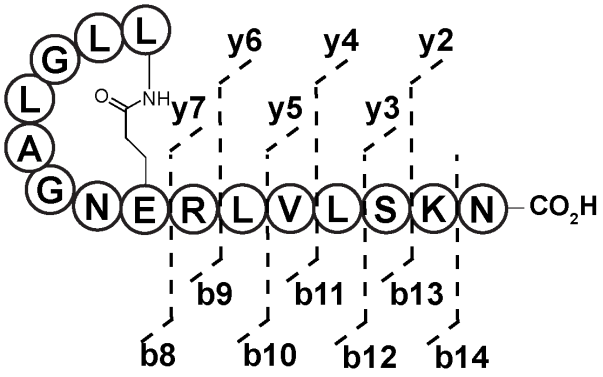
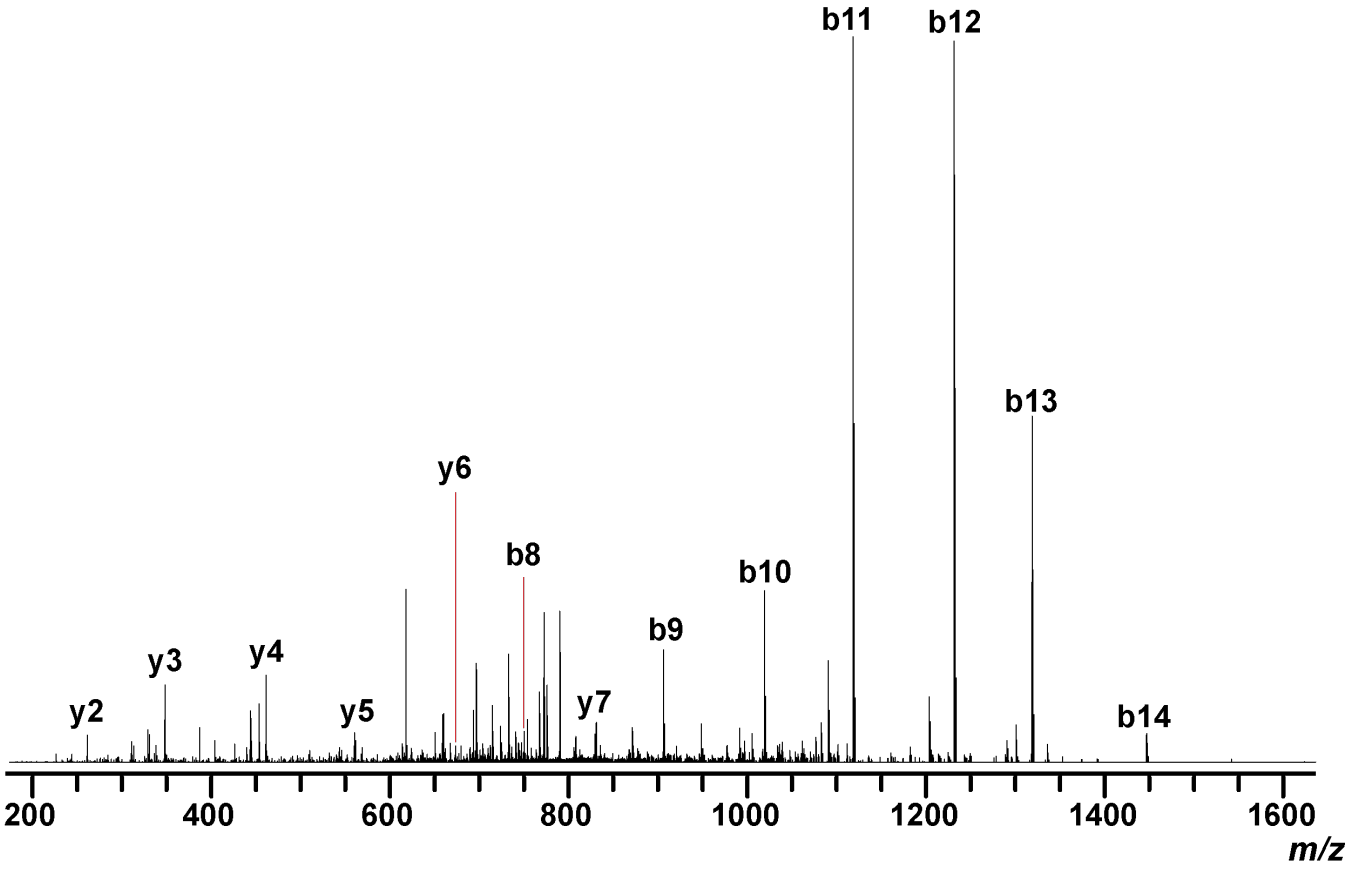


Figure S14: HR-ESI-MS/MS of biosynthetically reconstituted lasso peptide variants. Lasso peptides were subjected to collision-induced dissociation, resulting in a series of b^+ and y^+ ions consistent with the predicted site of macrolactam formation for each tested peptide. The ring expansion of fusilassin demonstrated a macrolatam formed between Trp1 and Glu10 (A). Substitution of the citrulassin A acceptor residue to Glu yielded a product with an 8-mer macrolactam between Leu1 and Gly8 (B, next page). Removal of Ala5 from CitA generated a product with the predicted 7-mer macrolactam (C, page after next).

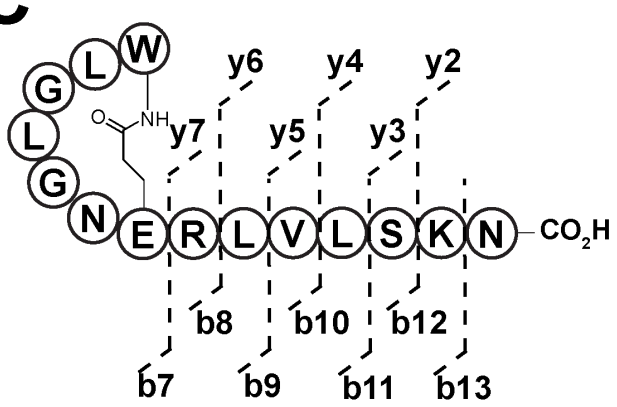
B



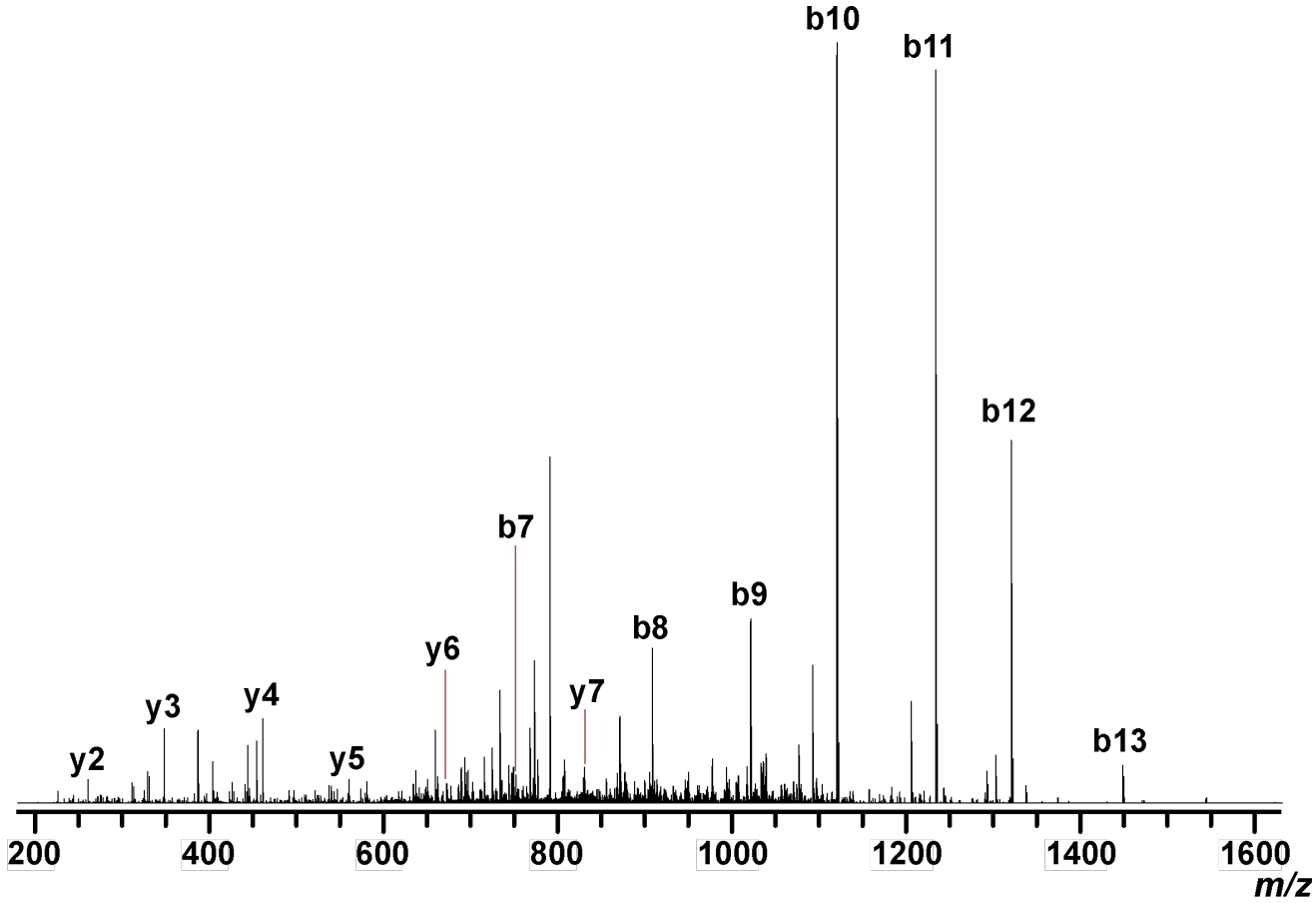
Species	Predicted mass	Observed mass	Error (ppm)
$[M+2H]^{2+}$	789.9699	789.9692	0.9
b8	750.4145	750.4140	0.6
b9	906.5156	906.5149	0.7
b10	1019.5996	1019.5987	0.9
b11	1118.6681	1118.6675	0.5
b12	1231.7521	1231.7514	0.6
b13	1318.7842	1318.7830	0.9
b14	1446.8791	1446.8780	0.7
y2	261.1557	261.1556	0.4
y3	348.1878	348.1875	0.7
y4	461.2718	461.2715	0.8
y5	560.3402	560.3398	0.9
y6	673.4243	673.4235	1.1
y7	829.5254	829.5248	0.7



C



Species	Predicted mass	Observed mass	Error (ppm)
$[M+2H]^{2+}$	790.9490	790.9479	1.4
b7	752.3726	752.3728	0.3
b8	908.4737	908.4730	0.8
b9	1021.5578	1021.5571	0.7
b10	1120.6262	1120.6254	0.7
b11	1233.7103	1233.7092	0.8
b12	1320.7423	1320.7407	1.2
b13	1448.8373	1448.8363	0.6
y2	261.1557	261.1556	0.5
y3	348.1878	348.1875	0.9
y4	461.2718	461.2714	1.0
y5	560.3402	560.3397	1.0
y6	673.4243	673.4238	0.8
y7	829.5254	829.5249	0.6



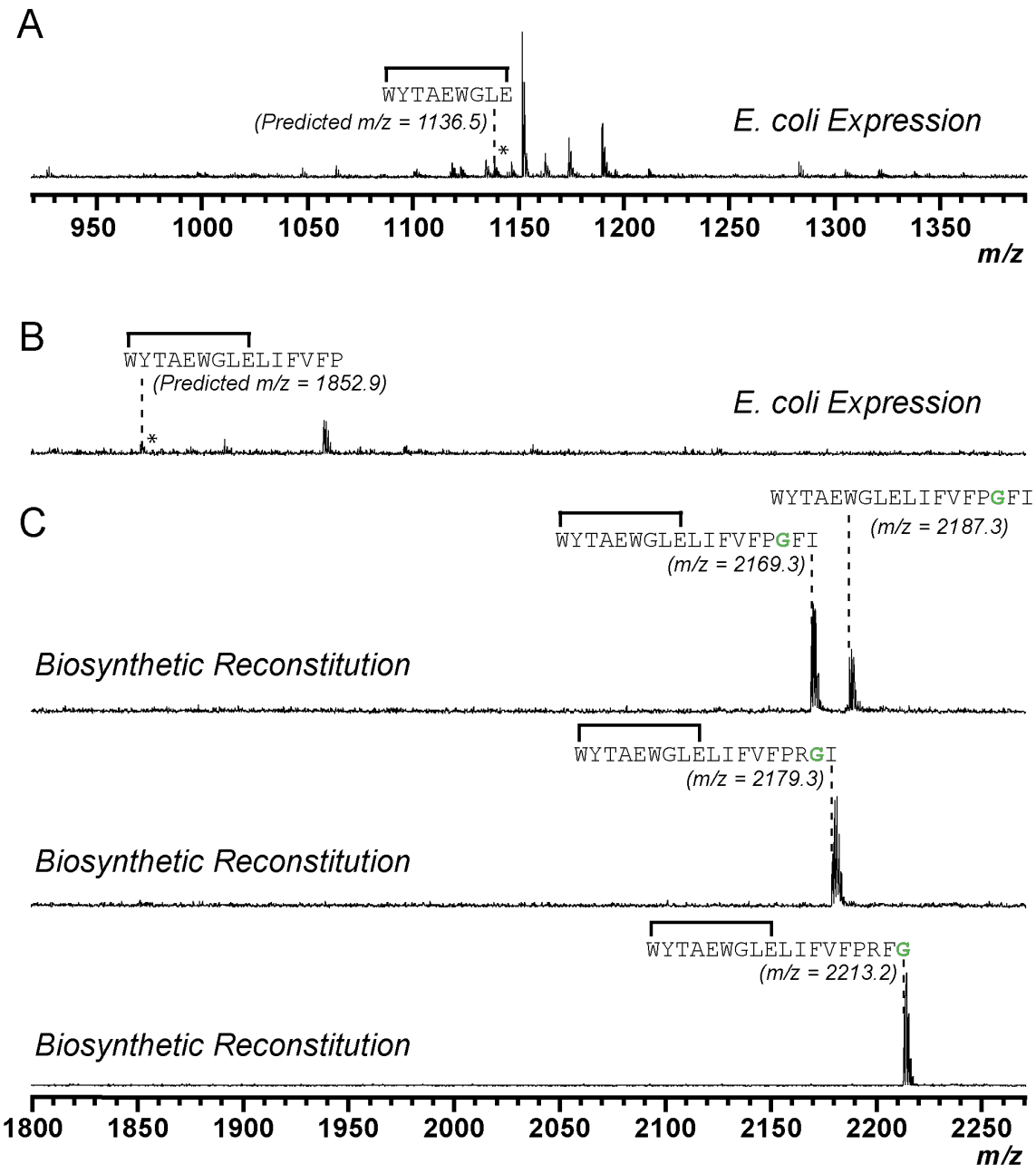


Figure S15. MALDI-TOF-MS analysis of fusilassin C-terminal truncations. Truncation of the C-terminal tail region of FusA abolished processing during *E. coli* expression (A). Removal of as few as 3 residues from the C-terminus ($_{16}$Arg-Phe-Ile$_{18}$) also abolished processing in *E. coli* (B). In panels A/B, the asterisks indicate unrelated contaminant ions of *m/z* = 1138.8 Da and 1851.3 Da, respectively. The importance of the three most C-terminal residues of FusA was probed by biosynthetic reconstitution after replacement with Gly. Substitution of Phe17 and Ile18 with Gly had no observed effect while Gly-substitution of Arg16 displayed impaired processing (C).

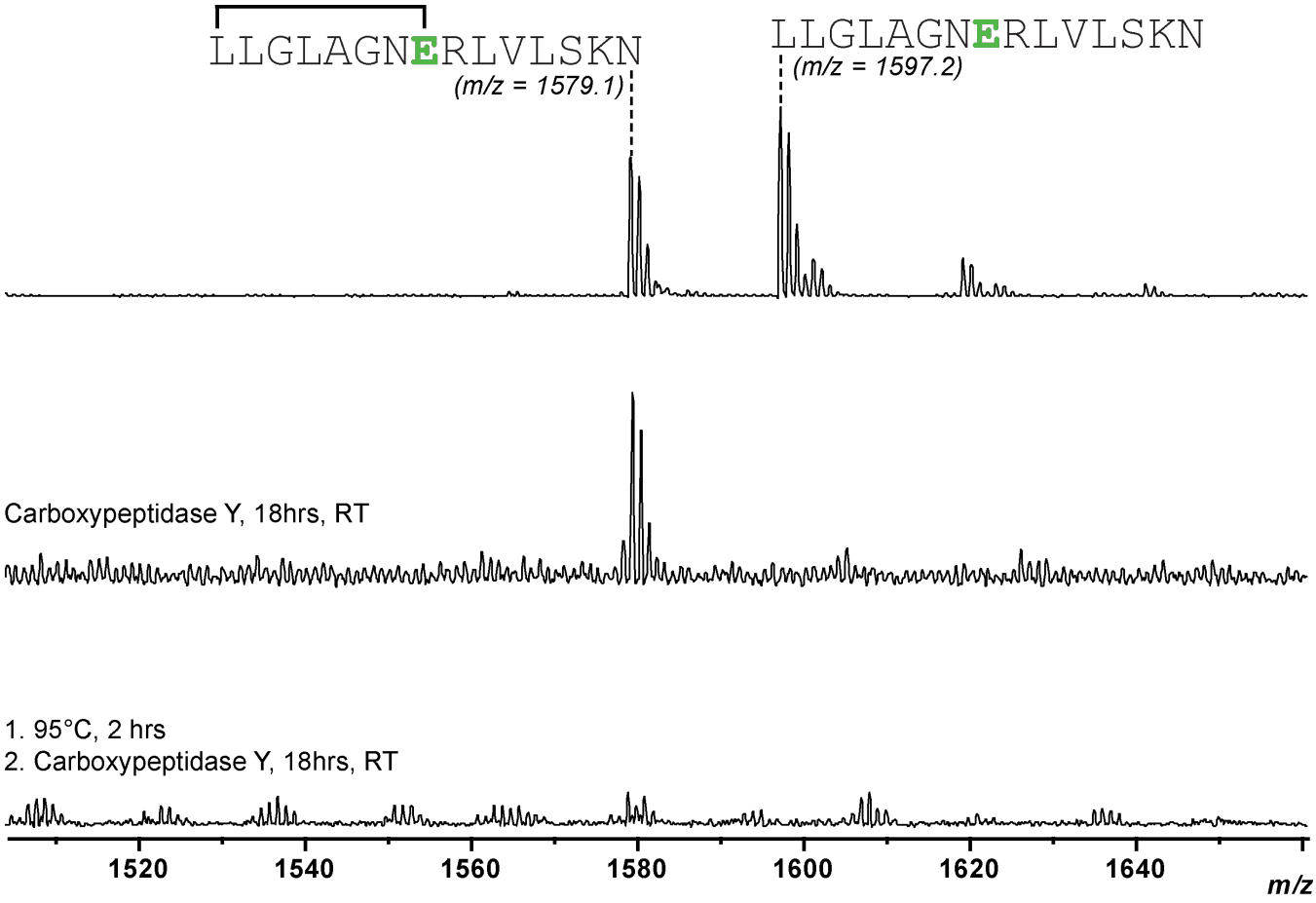


Figure S16. MALDI-TOF-MS of carboxypeptidase Y-treated CitA-D8E variant. Biosynthetically reconstituted lasso peptide (top) was incompletely digested after proteolysis (middle). Thermal unthreading of the lasso peptide and proteolysis completely degraded the mature product (bottom). Intermediate proteolysis products were not observed. The position of non-native residues (relative to citrulassin) are indicated in green. RT, room temperature.

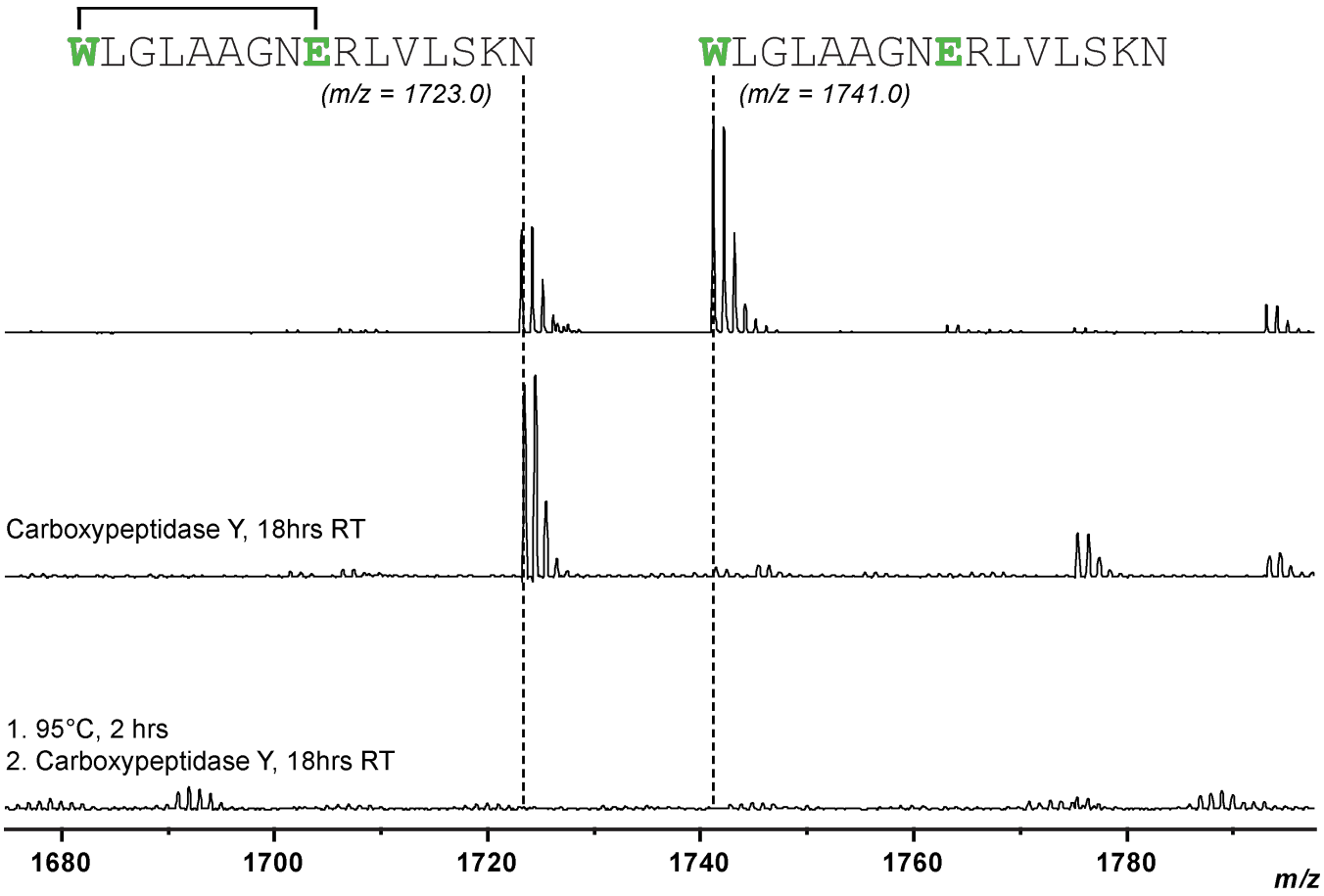


Figure S17. MALDI-TOF-MS of carboxypeptidase Y-treated CitA variant (L1W, D8E, A5ins). Proteolysis completely degraded the linear core peptide, though not the dehydrated product (middle). Thermal treatment potentiated complete proteolysis of the lasso peptide (bottom). Intermediate proteolysis products were not observed. The position of non-native residues (relative to citrulassin) are indicated in green. RT, room temperature.

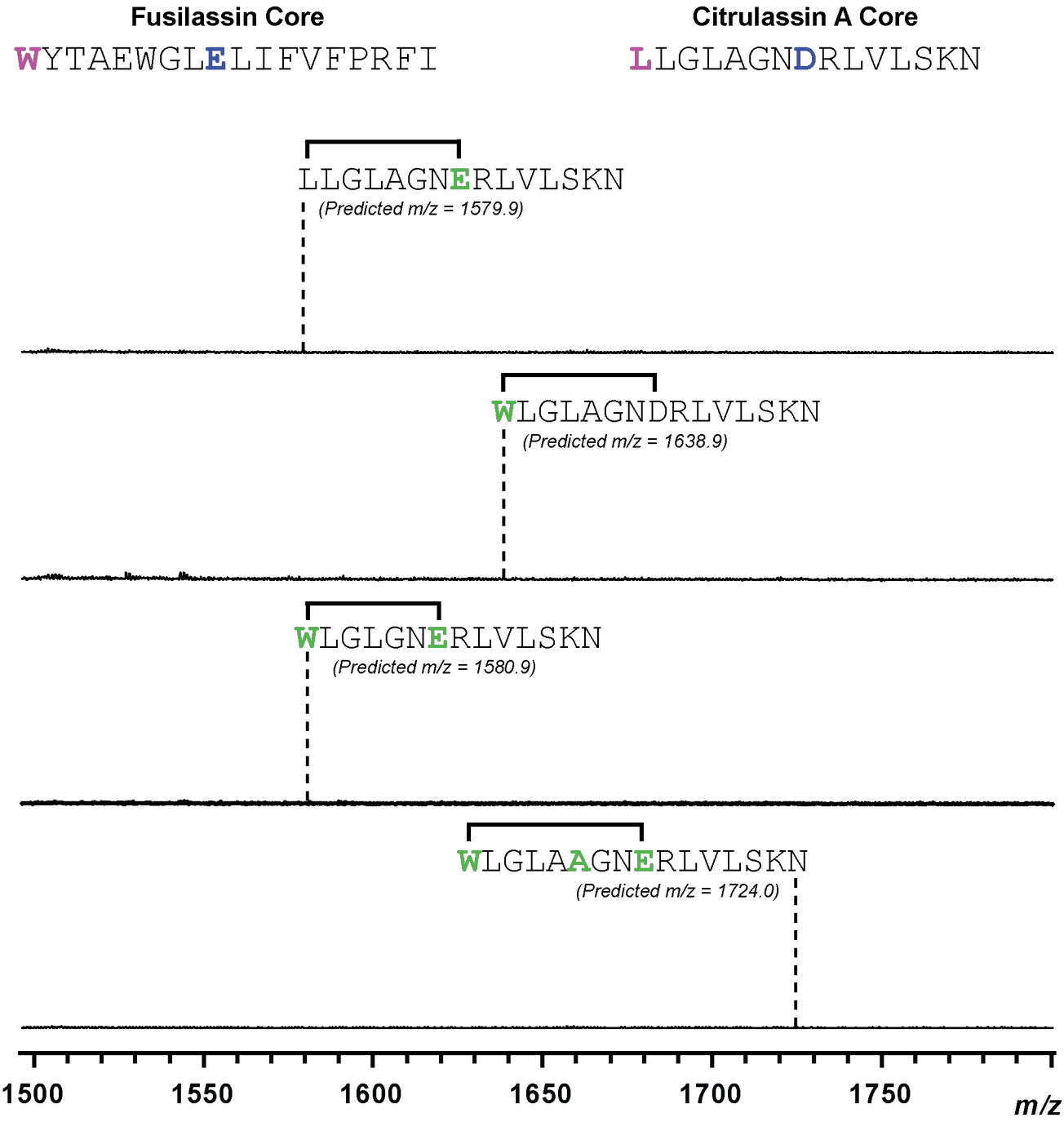


Figure S18. MALDI-TOF-MS of heterologously expressed citrulassin A variants. None of the tested constructs were processed by *E. coli* expression.

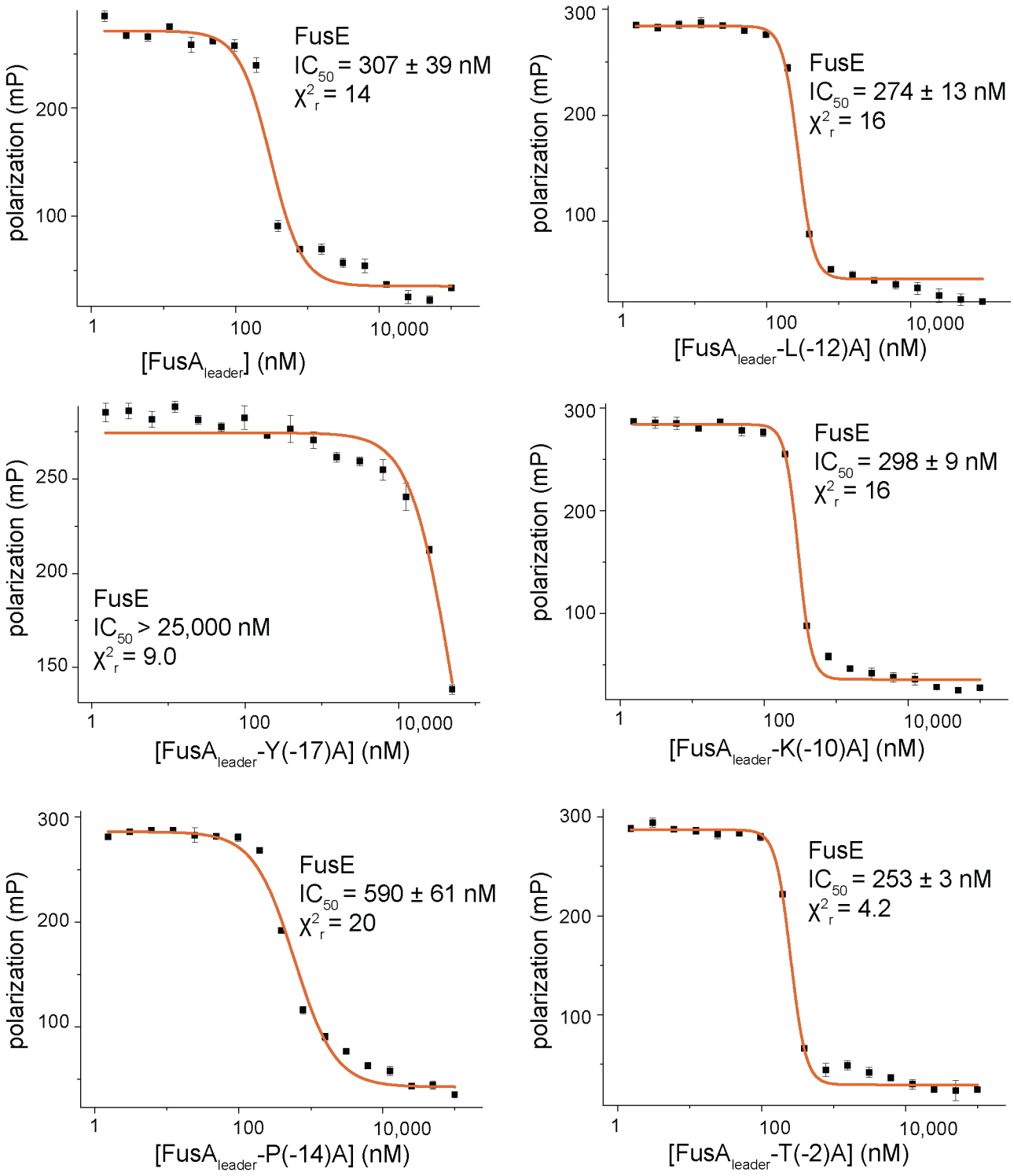


Figure S19. Competitive fluorescence polarization of $FusA_{leader}$ variants. Fluorescence polarization was performed using wild-type-FusE, fluorescein-labeled $FusA_{leader}$ and indicated $FusA_{leader}$ variants as competitive ligands. GREMLIN predictions of sites of interaction between the YxxP leader peptide motif of FusA and FusE (RRE) were confirmed, as they demonstrated decreased binding affinities. Remaining variants did not demonstrate significant perturbations to FusA-FusE interaction.

Table S3: Competition fluorescence polarization of $FusA_{leader}$ variants. Experiments were performed as described in methods. These data summarize the findings depicted in Figure S19. WT, wild-type.

$FusA_{leader}$ Variant	FusE	IC_{50} (nM)
WT	WT	307 ± 39
Y(-17)A	WT	>25,000
P(-14)A	WT	590 ± 61
L(-12)A	WT	247 ± 13
K(-10)A	WT	298 ± 9
T(-2)A	WT	253 ± 3

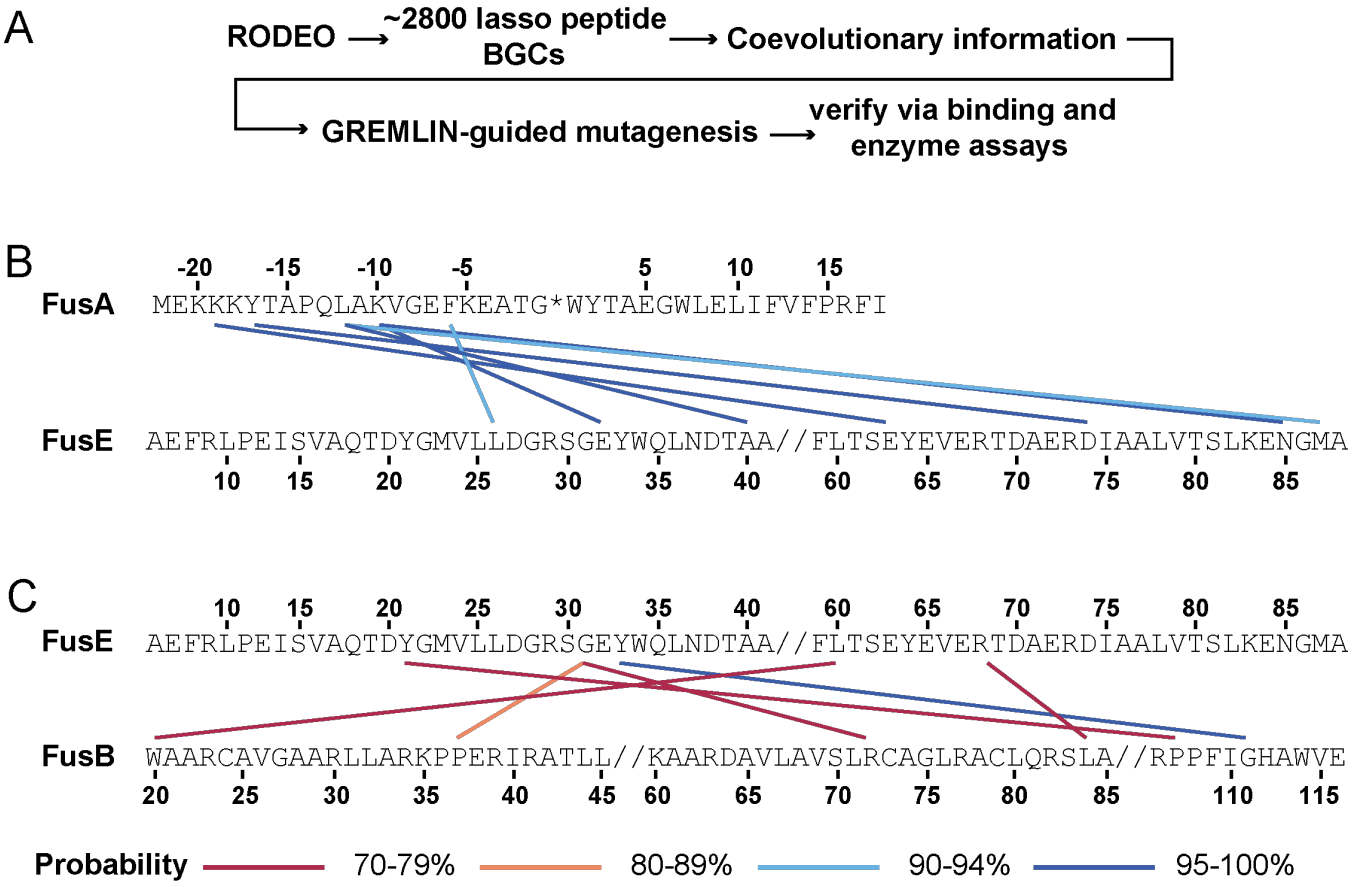
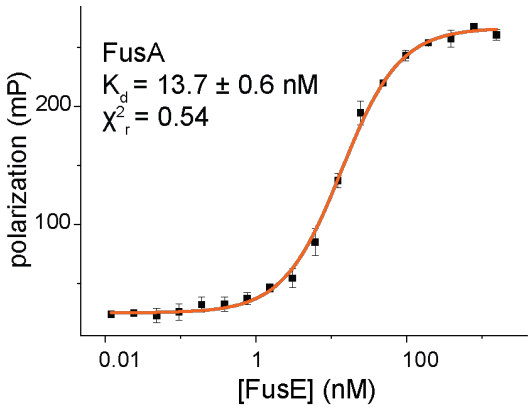
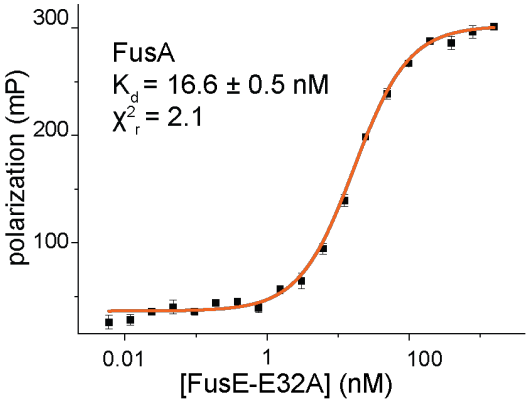
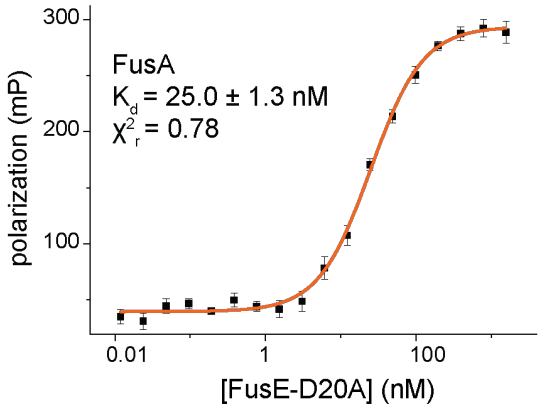
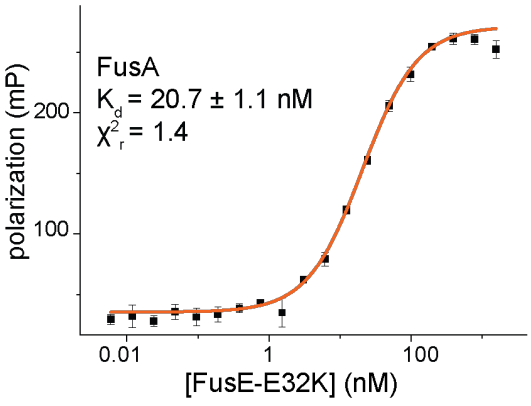


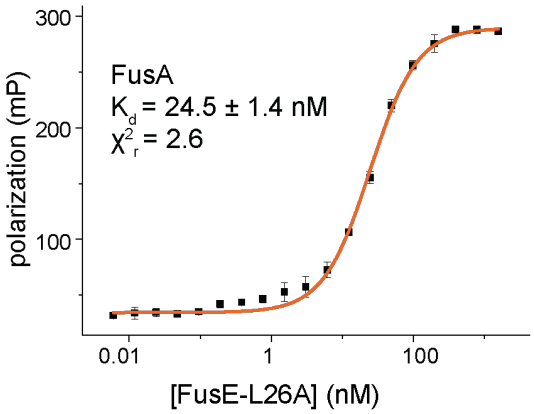
Figure S20. Coevolutionary analysis of $FusA_{leader}$:FusE and FusE:FusB. Workflow for generation of the GREMLIN data and experimental follow-up (A). GREMLIN analysis highlighting predicted sites of interprotein contact from coevolutionary (compensatory) mutations occurring between fusilassin biosynthetic proteins. Interacting residues between $FusA_{leader}$ and FusE (B) and interacting residues between FusE and FusB (C) were predicted by a bioinformatic data set derived from RODEO.

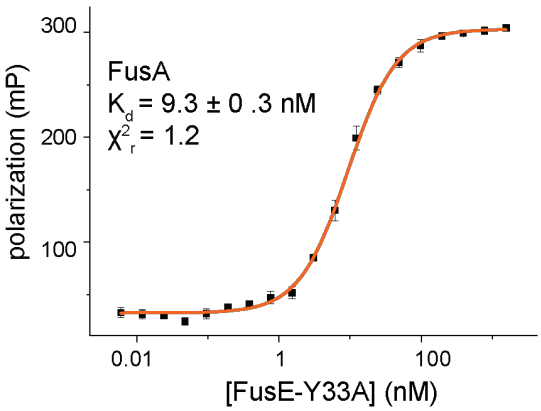

FusA
K_d = 13.7 ± 0.6 nM
χ^2_r = 0.54
polarization (mP)
100
200
0.01
1
100
[FusE] (nM)

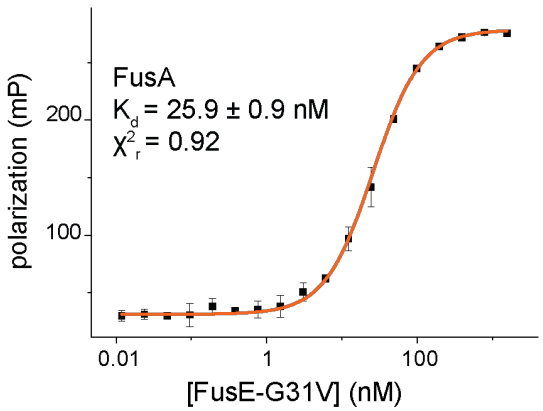

FusA
K_d = 16.6 ± 0.5 nM
χ^2_r = 2.1
polarization (mP)
100
200
300
0.01
1
100
[FusE-E32A] (nM)

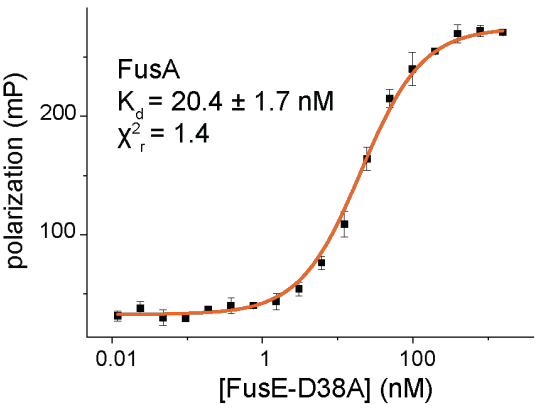

FusA
K_d = 25.0 ± 1.3 nM
χ^2_r = 0.78
polarization (mP)
100
200
300
0.01
1
100
[FusE-D20A] (nM)


FusA
K_d = 20.7 ± 1.1 nM
χ^2_r = 1.4
polarization (mP)
100
200
0.01
1
100
[FusE-E32K] (nM)


FusA
K_d = 24.5 ± 1.4 nM
χ^2_r = 2.6
polarization (mP)
100
200
300
0.01
1
100
[FusE-L26A] (nM)


FusA
K_d = 9.3 ± 0 .3 nM
χ^2_r = 1.2
polarization (mP)
100
200
300
0.01
1
100
[FusE-Y33A] (nM)


FusA
K_d = 25.9 ± 0.9 nM
χ^2_r = 0.92
polarization (mP)
100
200
0.01
1
100
[FusE-G31V] (nM)


FusA
K_d = 20.4 ± 1.7 nM
χ^2_r = 1.4
polarization (mP)
100
200
0.01
1
100
[FusE-D38A] (nM)

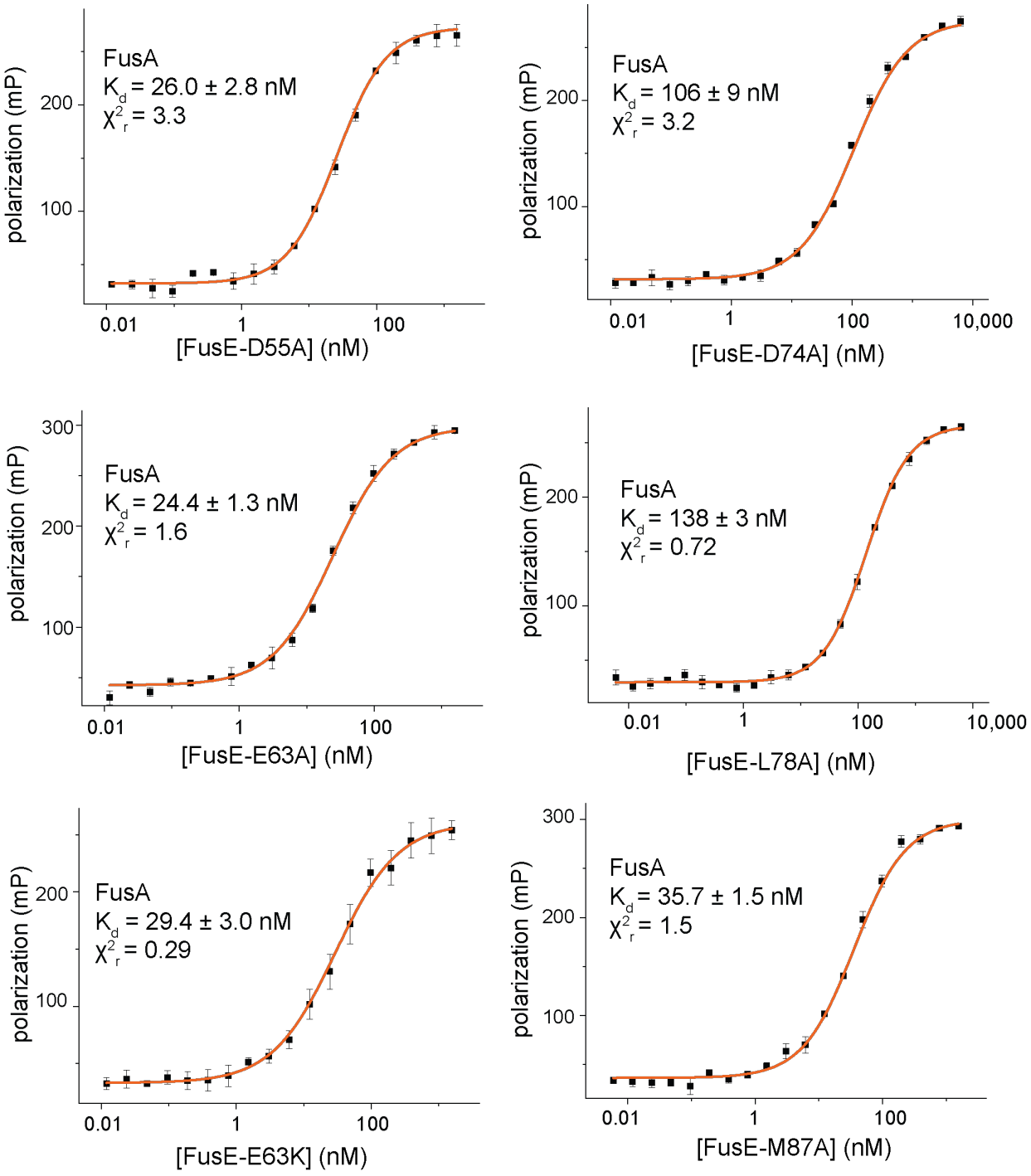


Figure S21. Direct fluorescence polarization of $FusA_{leader}$ and RRE variants. Fluorescence polarization was performed using fluorescein-labeled $FusA_{leader}$ and indicated FusE variants. Residues predicted to be sites of interaction between $FusA_{leader}$ and FusE (RRE) demonstrated lower binding affinities, whereas residues predicted to be sites of interaction between the FusE and FusB did not decrease binding affinity to $FusA_{leader}$.

Table S4: Direct fluorescence polarization of FusE variants. Experiments were performed as described in methods. These data summarize the findings depicted in Figure S21. WT, wild-type.

$FusA_{leader}$	FusE Variant	K_d (nM)
WT	WT	13.7 ± 0.6
WT	D20A	24.5 ± 1.3
WT	L26A	24.5 ± 1.4
WT	G31V	25.9 ± 0.9
WT	E32A	16.6 ± 0.5
WT	E32K	20.7 ± 1.1
WT	Y33A	9.3 ± 0.3
WT	D38A	20.4 ± 1.7
WT	D55A	20.6 ± 2.8
WT	E63A	24.4 ± 1.3
WT	E63K	29.4 ± 3.0
WT	D74A	106 ± 9
WT	L78A	138 ± 3
WT	M87A	35.7 ± 1.5

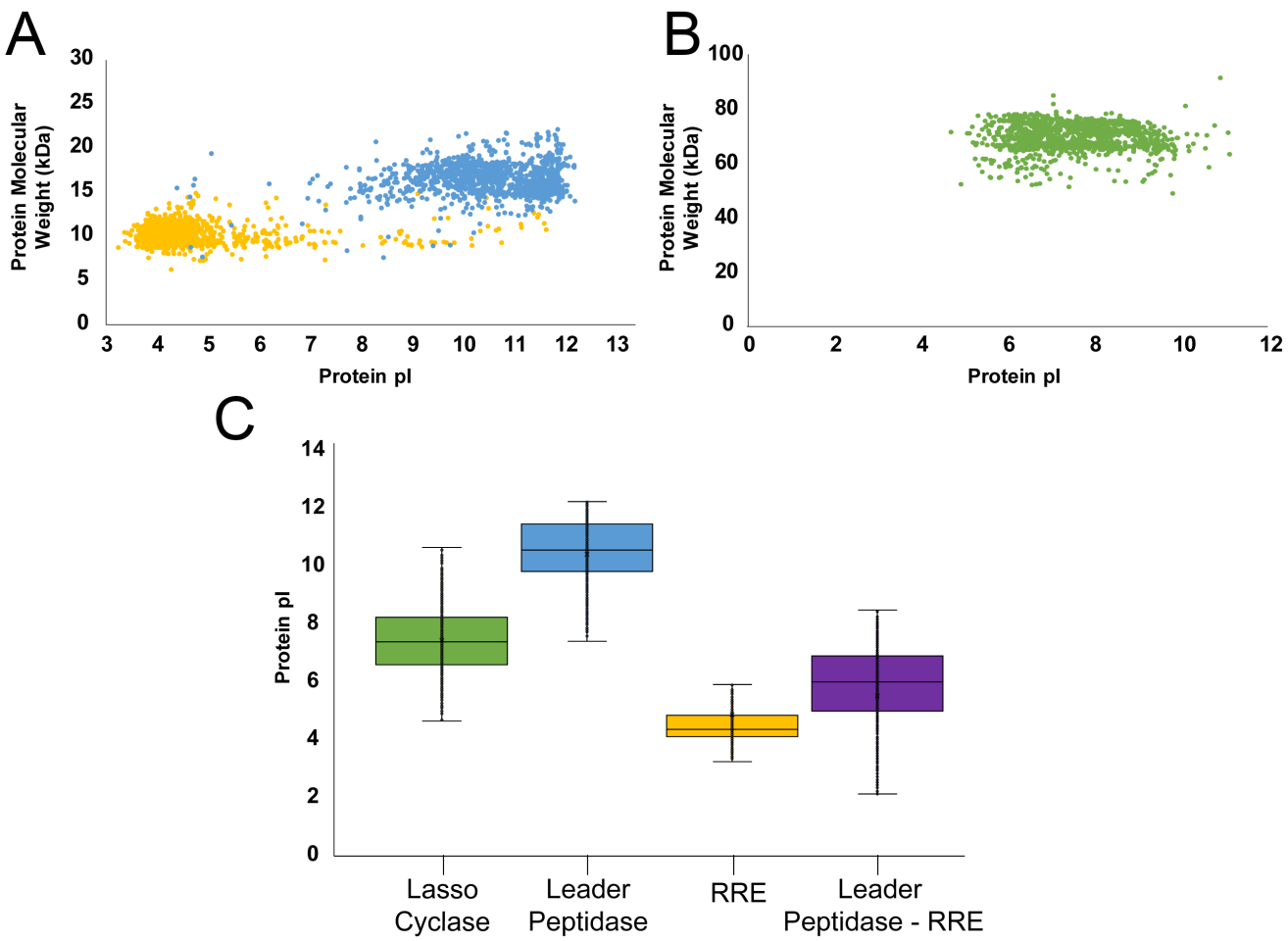


Figure S22. Calculated isoelectric points (pIs) of bioinformatically predicted lasso peptide biosynthetic enzymes. RRE proteins of molecular weight < 14.0 kDa exhibit acidic pIs (yellow), whereas leader peptidases of molecular weight < 22.5 kDa exhibit basic pIs (blue) (A). Lasso cyclases of molecular weight > 50 kDa exhibit neutral pIs (B). When subtracting the pI of the RRE from the leader peptidase within the same BGC (n > 1,000) the ΔpI's were approximately neutral (C, purple). This suggests the RRE-leader peptidase interaction is highly electrostatically driven. Statistical outliers were omitted.

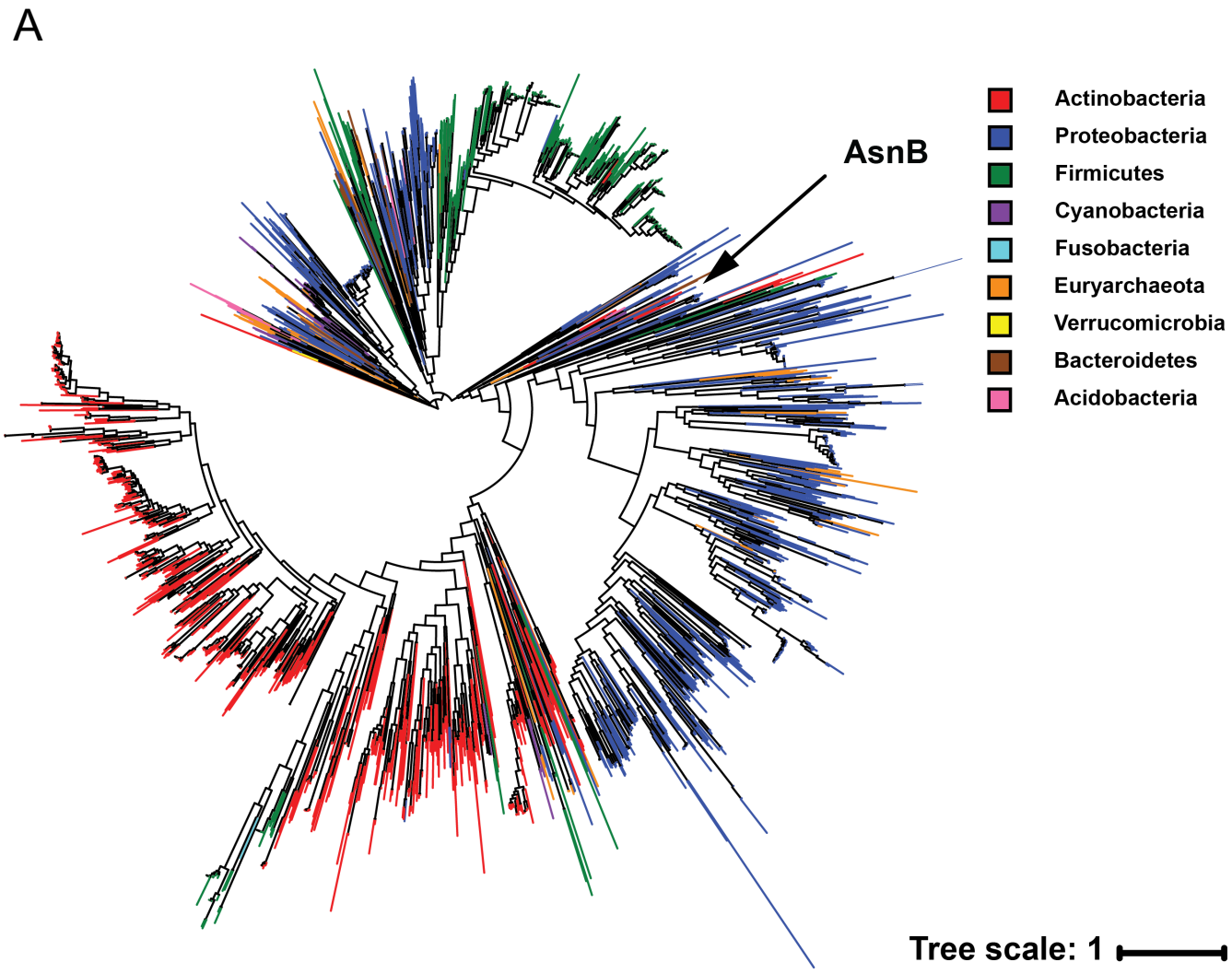
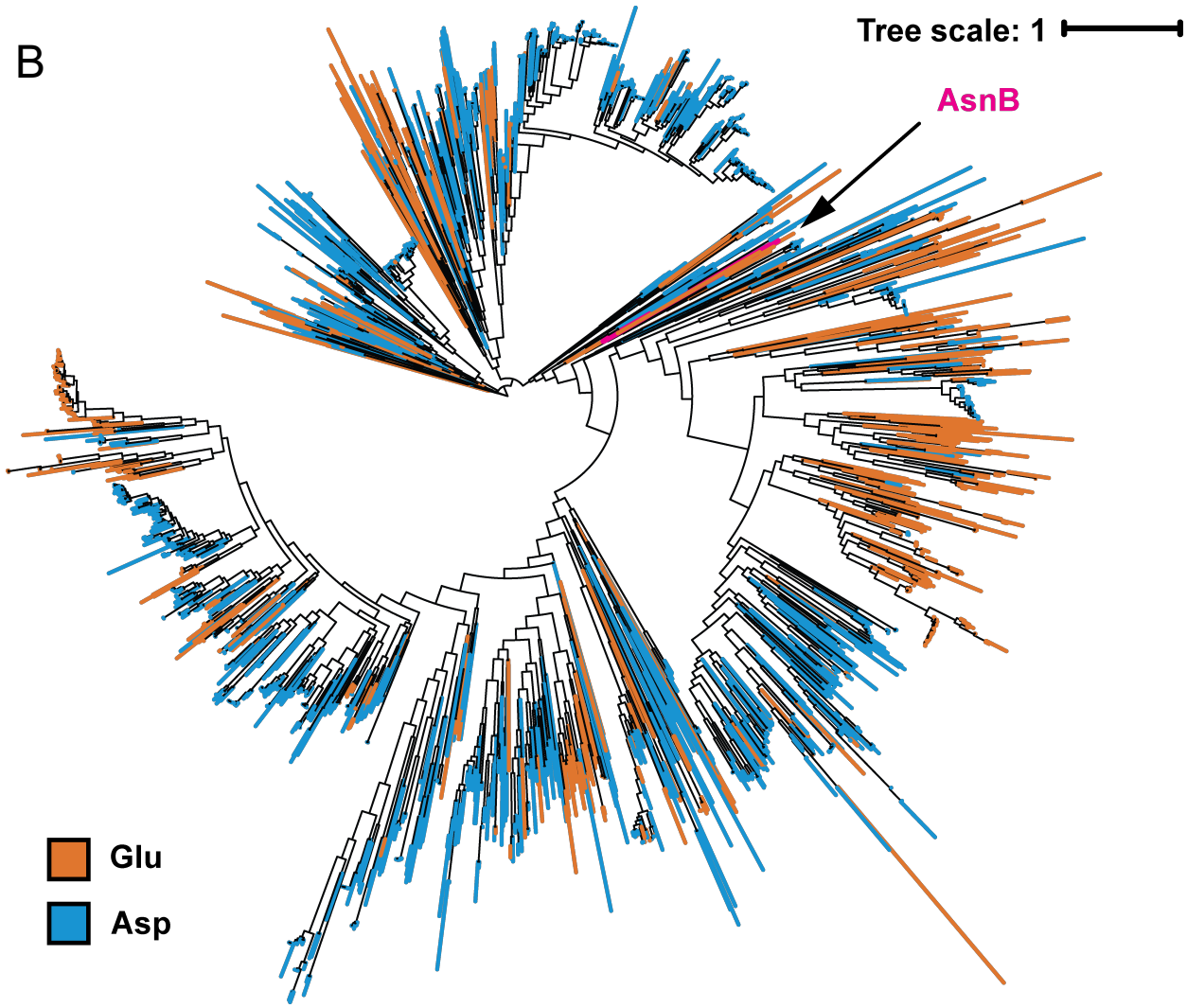


Figure S23: Phylogenetic trees of lasso cyclases. Shown is a maximum likelihood tree of RODEO-identified lasso cyclases colored by phylum (A). The Actinobacteria, Proteobacteria and Firmicutes are the most populated phyla. Given the monochromatic nature of the branches, these data suggest that lasso peptides have been largely transmitted vertically. The same tree has been color-coded based on the identity of the macrolactam acceptor residue (Asp vs Glu) (B). The branches are less monochromatic, suggesting that lasso cyclases that prefer Glu over Asp (or vice versa) evolved many times during evolution. For each tree, the location of *E. coli* aspargine synthetase B (AsnB) is given for reference.


B
Tree scale: 1
AsnB
Glu
Asp

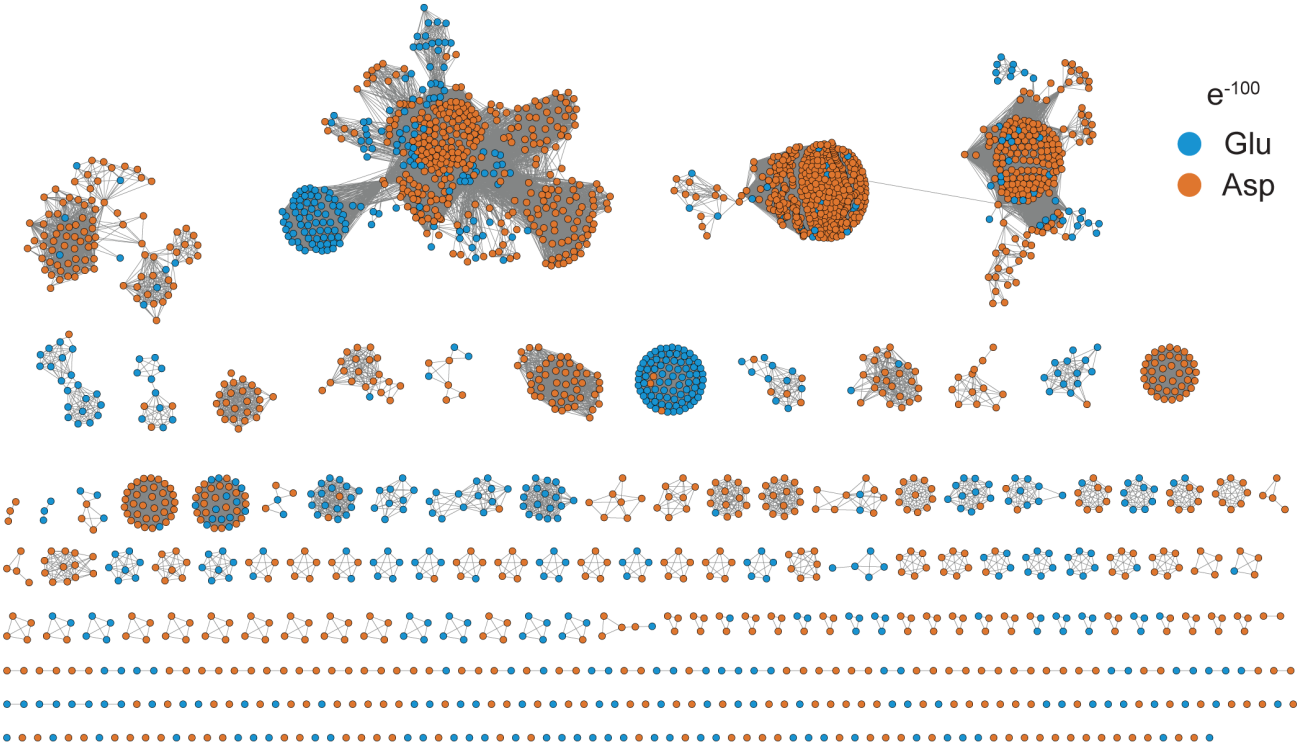


Figure S24: Sequence similarity network of lasso cyclases. At an alignment score of 100 (BLAST expectation value of e^{-100}), some lasso cyclases separate into clusters that correlate well with the identity of the macrolactam acceptor residue while other clusters contain a mixture of Asp (orange) and Glu (blue) acceptor types.

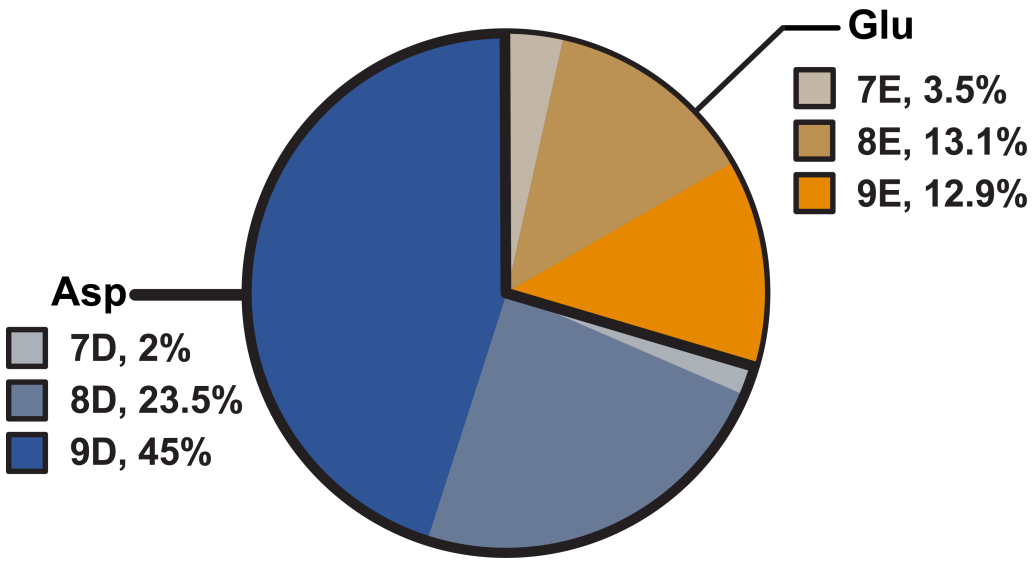


Figure S25: Ring architecture of bioinformatically predicted lasso peptides. The majority of lasso peptides possess an Asp acceptor residue. Among both acceptor types, the largest macrolactams (9-mers) are most frequently found while 7-mers are less frequently found and rare overall.

References

(1) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.* **2017**, *13* (5), 470–478.

(2) Chiu, J.; March, P. E.; Lee, R.; Tillett, D. Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100% efficiency in 4 h. *Nucleic Acids Res.* **2004**, *32* (21), e174.

(3) Chhetri, G.; Kalita, P.; Tripathi, T. An efficient protocol to enhance recombinant protein expression using ethanol in *Escherichia coli*. *MethodsX* **2015**, *2*, 385–391.

(4) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D., Higgins, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.

(5) Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K.; Battistuzzi, F. U. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35* (6), 1547–1549.

(6) Kozlowski, L. P. IPC - Isoelectric Point Calculator. *Biol. Direct.* **2016**, *11*, 55.

(7) Bogdanova, L. N.; Mchedlov-Petrossyan, N. O.; Vodolazkaya, N. A.; Lebed, A. V. The influence of beta-cyclodextrin on acid-base and tautomeric equilibrium of fluorescein dyes in aqueous solution. *Carbohydr. Res.* **2010**, *345* (13), 1882–1890.

(8) Pirrone, G. F.; Wang, H.; Canfield, N.; Chin, A. S.; Rhodes, T. A.; Makarov, A. A. Use of MALDI-MS Combined with Differential Hydrogen-Deuterium Exchange for Semiautomated Protein Global Conformational Screening. *Anal. Chem.* **2017**, *89* (16), 8351–8357.

(9) Morgan, C. R.; Engen, J. R. Investigating Solution-Phase Protein Structure and Dynamics by Hydrogen Exchange Mass Spectrometry. *Curr. Protoc. Protein Sci.* **2009**, *58*, 17.6.1-17.6.17.