ACS central science

Research Article

http://pubs.acs.org/journal/acscii

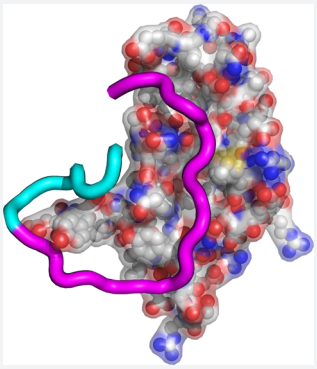
Lasso Peptide Biosynthetic Protein LarB1 Binds Both Leader and Core Peptide Regions of the Precursor Protein LarA

Wai Ling Cheung,[†] Maria Y. Chen,[†,‡] Mikhail O. Maksimov,[†] and A. James Link*[,†,§]

[†]Department of Chemical and Biological Engineering and [§]Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, United States

Supporting Information

ABSTRACT: Lasso peptides are a member of the superclass of ribosomally synthesized and posttranslationally modified peptides (RiPPs). Like all RiPPs, lasso peptides are derived from a gene-encoded precursor protein. The biosynthesis of lasso peptides requires two enzymatic activities: proteolytic cleavage between the leader peptide and the core peptide in the precursor protein, accomplished by the B enzymes, and ATP-dependent isopeptide bond formation, accomplished by the C enzymes. In a subset of lasso peptide biosynthetic gene clusters from Gram-positive organisms, the B enzyme is split between two proteins. One such gene cluster is found in the organism *Rhodococcus jostii*, which produces the antimicrobial lasso peptide lariatin. The B enzyme in *R. jostii* is split between two open reading frames, *larB1* and *larB2*, both of which are required for lariatin biosynthesis. While the cysteine catalytic triad is found within the LarB2 protein, LarB1 is a PqqD homologue expected to bind to the lariatin precursor LarA based on its structural homology to other RiPP leader peptide binding domains. We show that LarB1 binds to the leader peptide of the lariatin precursor protein LarA with a sub-micromolar affinity. We used photocrosslinking with the noncanonical amino acid *p*-azidophenylalanine and mass spectrometry to map the interaction of LarA and LarB1. This analysis shows that the LarA leader peptide interacts with a conserved motif within LarB1 and, unexpectedly, the core peptide of LarA also binds to LarB1 in several positions. A Rosetta model built from distance restraints from the photocrosslinking experiments shows that the scissile bond between the leader peptide and core peptide in LarA is in a solvent-exposed loop.



INTRODUCTION

The biosynthetic logic of many ribosomally synthesized and posttranslationally modified peptides (RiPPs)[1] involves the conversion of a precursor peptide into the final natural product by the action of one or more enzymes. The precursor peptide is composed of the core peptide, which becomes the natural product, and an N-terminal leader peptide, which is cleaved off during RiPP biosynthesis.[2] In the case of lasso peptides,[3] two enzymatic activities are required for formation of the final structure.[4] First, the leader peptide must be cleaved from the precursor peptide in order to generate a free core peptide N-terminus. That N-terminus is subsequently forged into an isopeptide bond that defines the lasso peptide fold. For most lasso peptides studied thus far, these two enzymatic activities can be mapped onto two distinct proteins.[5,6] The B enzymes, typified by McjB in microcin J25 biosynthesis, are cysteine proteases that cleave the leader peptide from the core peptide. The C enzymes, such as McjC, catalyze ATP-dependent isopeptide bond formation to generate the N-terminal "ring" portion of the lasso peptide.

In our previous studies to identify lasso peptides by genome mining,[7] we identified four conserved motifs within B enzymes. Motifs 2, 3, and 4 were found in the C-terminal portion of the B enzyme and are associated with the cysteine catalytic triad. Motif 1, on the other hand, was found in the N-terminal portion of the B enzyme (Figure S1). In the gene cluster for the antimycobacterial lasso peptide lariatin[8] (Figure 1), produced by *Rhodococcus jostii*, these four motifs are split between two ORFs, labeled *larC* and *larD* in the original report.[9] The *larC* ORF was deemed essential for lariatin production via complementation experiments.[9] In addition, the lariatin gene cluster includes a gene for the lasso peptide precursor, *larA*, an *mcjC* homologue, *larB*, and an ABC transporter, *larE*. Marahiel and co-workers have proposed renaming of these genes[10] to be in accordance with the nomenclature standards[1] for lasso peptides (Figure 1B, Figure S1). In the case of the "split" cysteine protease B enzyme, we will refer to the *larC* ORF as *larB1* and the *larD* ORF as *larB2*. Burkhart et al. have demonstrated previously that the LarB1 homologue in the lasso peptide cluster for streptomonomicin (Stm),[11] referred to as StmE, is a PqqD homologue that binds to the leader peptide of the Stm precursor protein, StmA.[12] These PqqD homologues appear in several different RiPP biosynthetic enzymes,[13−15] and thus have been termed RiPP precursor protein recognition elements (RRE).[12] We show here that the *larB1* gene product binds the leader peptide with sub-micromolar affinity and provide insights into the binding kinetics of this interaction. Furthermore, we map the interaction between the precursor protein LarA and LarB1 using photocrosslinking, allowing the

Received: June 30, 2016
Published: September 29, 2016

DOI: 10.1021/acscentsci.6b00184
ACS Cent. Sci. 2016, 2, 702−709

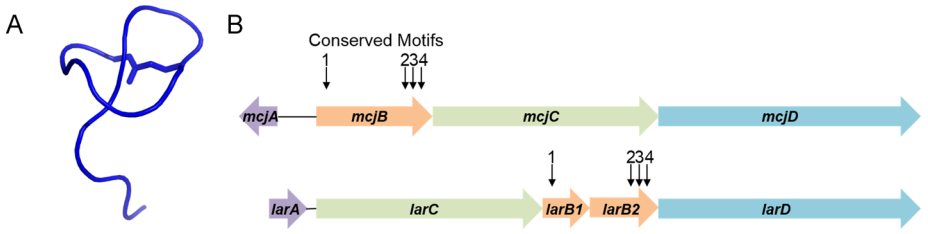
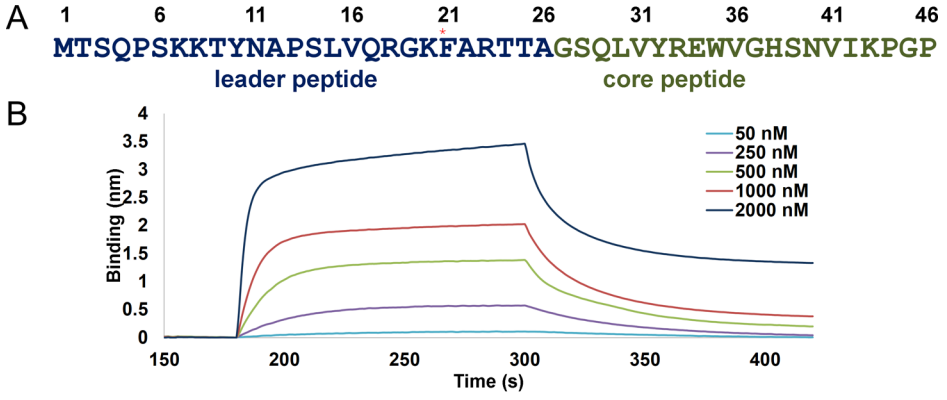


Figure 1. Structure and biosynthesis of the lasso peptide lariatin. (A) Lariatin A structure, drawn from coordinates in ref 8. (B) Biosynthetic gene clusters for microcin J25 (top) and lariatin (bottom). The conserved B enzyme motifs are split between the *larB1* and *larB2* genes in the lariatin gene cluster. Motifs 2, 3, and 4 correspond to the cysteine catalytic triad.



LarA-MBP conc. (nM)	K_d(M)	k_a($M^{-1}s^{-1}$)	k_d(s^{-1})
50	2.21×10^{-7}	$8.48 \pm 0.99 \times 10^{4}$	$1.87 \pm 0.02 \times 10^{-2}$
250	3.49×10^{-7}	$6.88 \pm 0.03 \times 10^{4}$	$2.40 \pm 0.003 \times 10^{-2}$
500	3.68×10^{-7}	$7.85 \pm 0.08 \times 10^{4}$	$2.89 \pm 0.01 \times 10^{-2}$
1000	4.43×10^{-7}	$7.28 \pm 0.06 \times 10^{4}$	$3.23 \pm 0.02 \times 10^{-2}$
2000	3.76×10^{-7}	$7.83 \pm 0.13 \times 10^{4}$	$2.95 \pm 0.04 \times 10^{-2}$

Figure 2. Binding of the lariatin precursor to LarB1. (A) Lariatin precursor (LarA) sequence. The leader peptide is shown in blue, and the core peptide is shown in green. The F21 residue where azidophenylalanine (AzF) was initially incorporated for the photocrosslinking experiments is marked with a red asterisk. (B) Biolayer interferometry (BLI) traces of LarA-MBP binding to immobilized LarB1-His at five different concentrations of LarA-MBP. Fits to the association data (k_a) and dissociation data (k_d) as well as the calculated equilibrium dissociation constant K_d are given in the table.

construction of a model of the LarA−LarB1 complex using Rosetta.

RESULTS AND DISCUSSION

Expression of LarB1. The protein product of the *larB1* ORF is 84 aa long, and the 15 aa motif 1 we have previously identified[7] corresponds to residues 16−30 of LarB1. We appended a C-terminal 6×His tag to the protein and were able to solubly express it in *Escherichia coli* and purify it under denaturing conditions (Figure S2). The protein remained soluble after removal of urea by buffer exchange with a yield of >10 mg/L of culture. Circular dichroism spectroscopy on this sample revealed a structure with α-helical content (Figure S3) in accordance with its function as a PqqD homologue.

Binding Kinetics. With purified LarB1 protein in hand, we next tested its ability to bind to the lariatin precursor peptide LarA using biolayer interferometry (BLI). In order to increase the signal in the BLI measurements, we fused LarA to the N-terminus of maltose binding protein (MBP). This fusion also includes a Factor Xa cleavage site (IEGR) between LarA and MBP. The LarA-MBP fusion was purified using amylose resin (Figure S4).

We first coated an anti-His biosensor with purified LarB1-His and used LarA-MBP as an analyte at 10 different concentrations ranging from 50 nM to 2 μM. LarA-MBP exhibited binding to LarB1 with an affinity of 440 ± 180 nM (standard deviation of 18 measurements) (Figure 2B). Using MBP alone or a fusion protein of the core peptide of LarA with MBP (LarACP-MBP) as analyte, we observed no binding (Figure S5). We also constructed a fusion protein between the leader peptide of LarA (LarALP, Figure 2A) with MBP. A four amino acid flexible linker, Gly-Ser-Ser-Gly, was added between LarALP and MBP. This construct was also tested using BLI with concentrations ranging from 100 nM to 2 μM of LarALP-MBP. Similar binding affinity was observed between LarALP and LarB1, with the caveat that BLI signal during the association phase did not reach a clean plateau at high LarALP-MBP concentrations (Figure S6). As a further confirmation that the interaction between LarA and LarB1 is mediated by the leader peptide, we used a 26 aa synthetic peptide corresponding to the leader peptide segment of LarA as the analyte in BLI experiments. Because of the small size of this analyte, we were able to detect binding only at a single, high concentration of the LarA leader peptide (4 μM). However, the association and dissociation rate constants were consistent with

DOI: 10.1021/acscentsci.6b00184
ACS Cent. Sci. 2016, 2, 702−709

the values determined for the MBP fusion constructs (Figure S5) and gave a K_d value of 280 nM.

Previously, we have found that the N-terminal portion of the leader peptide of the microcin J25 (MccJ25) precursor was dispensable with regard to the biosynthesis of MccJ25 in *E. coli*.[16] To test whether N-terminal truncations of LarA affected its binding to LarB1, we generated LarA-MBP constructs with 5−15 aa truncated from the N-terminus of LarA, not counting the N-terminal methionine. BLI experiments with these constructs showed that the first 5 aa (TSQPS, see Figure 2A) were dispensable with no effect on binding but the 9 aa, 10 aa, and 15 aa truncations abolished detectable binding at 2 μM analyte concentration (Figure S7). From these data we conclude that LarA and LarB1 interact with a sub-micromolar K_d, and that the interaction is mediated by the leader peptide portion of LarA.

LarB1 Homology Model. Recently a prevalent precursor recognition element (RRE) in RiPPs biosynthesis has been identified.[12] These RREs are homologues of the PqqD protein involved in pyrroloquinoline quinone biosynthesis.[17] They may be found as standalone proteins as in the case of PqqD, StmE, and LarB1 or as domains in larger enzymes. PqqD[18] bound its substrate PqqA with sub-micromolar affinity, in agreement with what we report above for LarA binding to LarB1. PqqD has also been recently shown to be necessary for the function of radical *S*-adenosylmethionine protein PqqE.[19] However, no structure of the PqqA/PqqD complex is currently available.

Three crystal structures of RiPP biosynthetic enzymes with RRE domains with bound precursors or leader peptides have been reported. MccB is an enzyme responsible for the posttranslational adenylation of a heptapeptide to generate the antibiotic microcin C7.[20] The structure of this enzyme in complex with its heptapeptide substrate MccA[13] shows the peptide bound in an N-terminal domain of MccB which the authors term a "peptide clamp". When the peptide clamp domain was removed from MccB and then added back in *trans*, ∼20% of the MccB adenylation activity was recovered.[13] This is reminiscent of the two different solutions for leader peptide cleavage in lasso peptide biosynthesis: leader peptide binding and protease activity can be either localized to a single protein or split between two separate proteins. More recently, the structure of the lantibiotic biosynthetic enzyme NisB in complex with a portion of its precursor peptide NisA[14] and the structure of the cyanobactin biosynthetic enzyme LynD with the PatE precursor peptide[15] have been determined. NisB and LynD engage the leader peptide sequence in a similar pose while MccB engages MccA in an entirely different orientation.[21]

We constructed a homology model of LarB1 using I-TASSER,[22] and the highest ranking threading template was PDB file 3G2B, corresponding to PqqD. When we aligned our LarB1 homology model to either MccB or NisB, there was high structural similarity to the leader peptide binding domains with an rmsd of 3.33 Å over 64 amino acids for MccB and 3.45 Å over 72 amino acids for NisB (Figure 3). This homology model was used to guide our subsequent photocrosslinking experiments.

Photocrosslinking between LarA and LarB1. To examine the interaction between LarA and LarB1 in more detail, we next turned to a photocrosslinking approach.[23] Our BLI experiments above indicated that the N-terminal portion of a LarA was dispensable with regard to LarB1 binding. Thus, we first replaced the phenylalanine in position 21 (F21) near the C-terminus of the LarA leader peptide (Figure 2A) with *p*-azidophenylalanine (AzF) (Figure 4A) within the LarA-MBP fusion. This was accomplished by introducing an amber codon at the F21 position (Figure 2A) and utilizing an engineered aminoacyl-tRNA synthetase/tRNA pair[24] for the site-specific

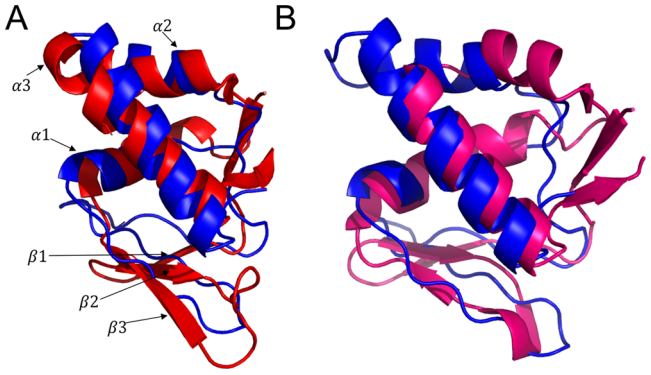


Figure 3. Comparison of LarB1 homology model to NisB and MccB leader peptide binding domain. (A) Homology model of LarB1 (blue) overlaid with NisB residues 139−225 (red). (B) Homology model of LarB1 (blue) overlaid with MccB chain B residues 1−81 (pink). NisB structures drawn from PDB file 4WD9. MccB drawn from PDB file 3H9G.

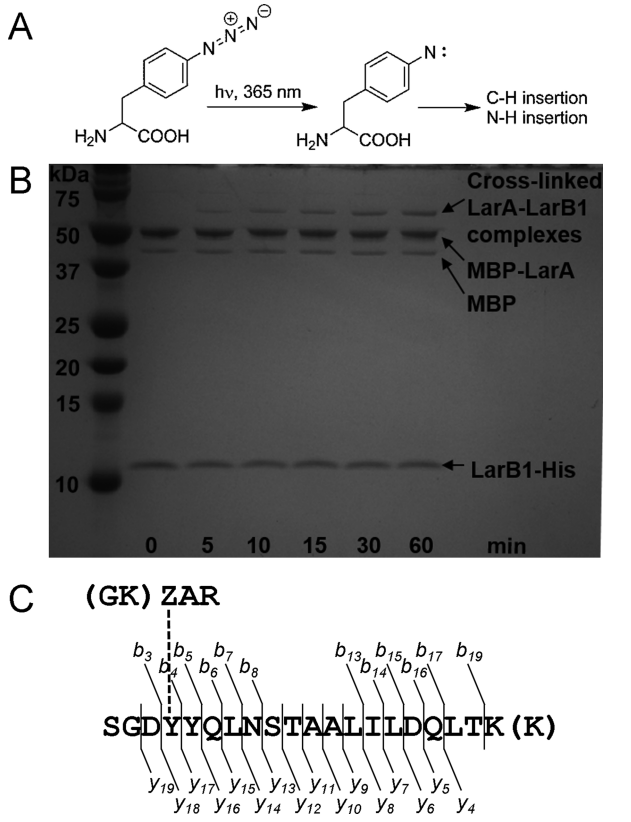


Figure 4. Photocrosslinking of LarA to LarB1. (A) Structure of azidophenylalanine (AzF) and the nitrene that forms upon photolysis. (B) SDS−PAGE analysis showing that the band for LarA−LarB1 adducts grows as expected with the length of UV exposure. (C) Location of LarA cross-linked to LarB1 as determined by LC−MS/MS. The AzF incorporated into LarA is labeled Z. Amino acids in parentheses correspond to missed cleavages in the tryptic digestion. The observed *b*- and *y*-ions are shown, indicating that the cross-linking occurred to Tyr-28 of LarB1. Spectra corresponding to this figure are in Figure S9.

DOI: 10.1021/acscentsci.6b00184
ACS Cent. Sci. 2016, 2, 702−709

incorporation of AzF. Of the photocrosslinking amino acids available for amber suppression in *E. coli*,[24−26] AzF is the most structurally similar to Phe and thus is expected to disrupt the native interaction between LarA and LarB1 minimally. Upon photolysis at 365 nm, the aryl azide side chain of AzF eliminates N_2 and generates a nitrene that can insert into nearby C−H or N−H bonds.[27]

We incubated 2 μM LarB1-His with 1 μM LarA-MBP(AzF) under UV illumination at 4 °C. SDS−PAGE analysis of the reaction showed clear UV-dependent cross-linking that increases over a period of 1 h (Figure 4B). To determine the location of the cross-link, we first removed free LarB1 using a 30 kDa cutoff centrifugal filter device. The remaining protein was digested with Factor Xa to cleave the MBP off of LarA. Finally, the protein mixture was purified using IMAC chromatography to isolate photocrosslinked LarA−LarB1 complexes (Figure S8). This protein was digested with trypsin, and the resulting peptides were analyzed using MALDI-TOF mass spectrometry (Figure S9A). Assuming total digestion, the tryptic peptide of LarA that contains AzF is only three amino acids, ZAR, where Z is AzF. We searched for conjugates of this peptide with all possible tryptic peptides of LarB1 and found a peptide with MH^+ = 2619.2 Da, close to the expected mass of 2619.3 Da (Figure S9B). In addition, we observed a related peptide with a missed tryptic cleavage prior to the site of AzF incorporation, GKZAR, also conjugated to the same LarB1 peptide (Figure S9B). The tryptic peptide that is photocrosslinked to the LarA leader peptide corresponds to amino acids 25−44 of LarB1.

To pinpoint the site of the cross-link between the leader peptide of LarA and LarB1, we next turned to LC−MS/MS analysis. Fragmentation of the masses we observed in MALDI indicated that AzF was covalently attached to Tyr-28 of LarB1 (Figure 4C, Figure S10). To confirm the specificity of the cross-linking reaction, LarB1 Tyr-28 was then mutated to Ala and photocrosslinking between LarA F21Z and LarB1 Y28A was repeated. No cross-linking was observed (Figure S11). Additionally, BLI measurements between LarB1 Y28A and LarA were made, showing an approximate 2-fold increase in K_d (Figure S12) in accordance with alanine scanning experiments carried out on StmE.[12] With the cross-linking reaction carried out above K_d, the modest change in binding affinity does not fully account for the abolishment of cross-linking. Instead, we propose that the aryl ring of Tyr-28 is the likely site of cross-linking with AzF, and substitution of Tyr-28 with alanine eliminates the site of cross-linking. Collectively, this data provides a strong confirmation that the C-terminal portion of the LarA leader peptide interacts with LarB1. It is also noteworthy that Tyr-28 is one of the more well-conserved residues within motif 1 (Figure S1).[7]

Successful spatial mapping of LarA Phe-21 to LarB1 Tyr-28 using photocrosslinking prompted the analysis of additional site-specifically incorporated AzF constructs, either in LarA or LarB1, with the homology model as a guide. The Tyr-28 residue of LarB1 is homologous to Val-169 of NisB. The segment of NisB including Val-169 is a β-strand and interacts directly with the NisA leader peptide, indicating that LarB1 may engage the LarA leader peptide in a similar fashion (Figure S13A). In the case of MccB, the Tyr-28 of LarB1 is homologous to Gln-27 in MccB. While Gln-27 is in the vicinity of the MccA precursor heptapeptide, it does not make direct contact with MccA (Figure S13B). Thus, we hypothesized that the LarA/LarB1 interaction is more similar to the NisA/NisB interaction than the MccA/MccB interaction.

A series of LarB1 constructs with AzF substitution were created (Figure 5), scanning for interactions between the β3 strand of LarB1 (Figure 5, Figure S1) and the leader peptide of LarA. The β3 strand is in closest proximity to the leader peptide in the NisA/NisB complex. The LarB1 Y28Z and LarB1 Q30Z constructs cross-linked well to LarA-MBP, while other substitutions in the β3 strand cross-linked weakly or not at all (Figure 5, Figure S14). LC−MS/MS analysis on the cross-linked LarB1 Y28Z and LarB1 Q30Z species showed that the cross-linking was occurring not in the leader peptide region of LarA but in the core peptide (Figure S15, Figure 5). This result was surprising given the BLI data presented above showing that LarACP-MBP does not bind to LarB1 (Figure S5). To verify this result, cross-linking between LarB1 Q30Z was repeated with full length LarA-MBP, LarALP-MBP, and LarACP-MBP. No cross-linking was observed with LarACP-MBP, very weak cross-linking with LarALP-MBP, and strong cross-linking with full length LarA-MBP (Figure S16). This is in good agreement with BLI data that showed that LarB1 binds tightly to the leader peptide but not to the core peptide, indicating that leader peptide binding would be necessary to position any interactions with the core peptide.

The finding that residues in β3 of LarB1 bind the core peptide suggested a model in which the core peptide must loop around in order to contact the surface of LarB1 (Figure 5). To test this model further, we introduced AzF to the Y16 and V19 positions of LarB1. These residues are located in the loop between β1 and β2, and in the β2 strand of LarB1, respectively. Both of these constructs cross-linked well to LarA-MBP (Figure S14), and LC−MS/MS analysis localized the cross-linking to the C-terminal portion of the LarA leader peptide (Figure 5, Figure S15). This experiment provided further evidence for a model in which the leader peptide docks next to β3 of LarB1, generating a four-stranded antiparallel β-sheet. The core peptide then curls back on the surface of LarB1, resulting in the specific interactions observed with the Y28Z and Q30Z constructs.

Given this model, we wanted to revisit the substitutions in the β3 strand of LarB1 that gave poor cross-linking results. BLI measurements were made with the LarB1 D27Z, Y29Z, L31Z, and N32Z constructs to determine if binding affinities were affected by the amino acid substitution. The measured K_d's were generally within a few fold of the wild-type interaction, with LarB1 L31Z exhibiting wild-type binding (Figure S17). For this construct, the AzF residue likely does not point toward LarA. For LarB1 D27Z, Y29Z, and N32Z, however, the drop in BLI signal was substantially greater than could just be accounted for from the drop in K_d. An explanation for this behavior is that these amino acid substitutions destabilize the protein to the point that much of the protein bound to the biosensor is misfolded. The small fraction of folded protein would still give near wild-type K_d values, while also giving a decreased signal in the BLI measurements. Thus, we propose that these residues on β3 of LarB1 play a role in promoting the stability of the protein.

Lastly, we generated several additional AzF-substituted LarA constructs to test their cross-linking to LarB1. The Y10Z LarA construct, with AzF located in the leader peptide well upstream of the Y21 position discussed earlier, did not cross-link to LarB1 (Figure 5). However, two variants in which AzF was inserted in the core peptide, Y32Z and W35Z, did exhibit cross-

DOI: 10.1021/acscentsci.6b00184
ACS Cent. Sci. 2016, 2, 702−709

A

Construct	Crosslinking	LC-MS/MS	Crosslinked Region on Partner	LarA Precursor Location
LarB1 Y16Z	Good	Y	T24-A26	End of leader
LarB1 V19Z	Good	Y	T24-A26	End of leader
LarB1 D27Z	Very weak	N	N/A	
LarB1 Y28Z	Good	Y	Y32-R33	Ring region of core
LarB1 Y29Z	Weak	N	N/A	
LarB1 Q30Z	Good	Y	Q29-R33	Ring region of core
LarB1 L31Z	Weak	N	N/A	
LarB1 N32Z	None	N	N/A	

B

Construct	Crosslinking	LC-MS/MS	Crosslinked Position on LarB1	Tertiary Location on LarB1
LarA Y10Z	None	N	N/A	N/A
LarA Y21Z	Good	Y	Y28	β3
LarA Y32Z	Good	N	N/A	N/A
LarA W35Z	Good	Y	Y16; A35-A36	β1-β2 turn; α1

C

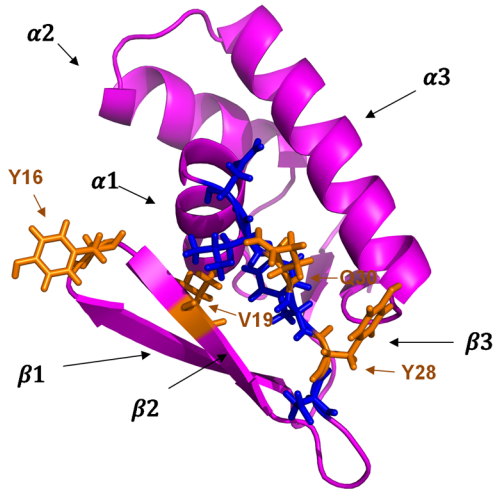


Figure 5. Photocrosslinking experiments tested. (A) Photocrosslinking results of LarB1 AzF constructs with wild-type LarA. (B) Photocrosslinking results of LarA AzF constructs with wild-type LarB1. (C) LarB1 homology model with AzF insertion locations highlighted in blue and orange. Blue positions gave little to no cross-linking while orange positions were strong cross-linkers.

linking (Figure 5, Figure S14). LC−MS/MS analysis on the W35Z cross-linked adduct (Figure S15) provided further support for the model in which the LarA core peptide curls back to interact with LarB1.

LarA−LarB1 Docking Model. Using the photocrosslinking and BLI data, we sought to build a structural model of the LarA and LarB1 interaction. The ability of LarA −5 aa to bind as well as full length LarA while LarA −9 aa lost binding (Figure S7) indicates that there are important residues between positions 7 and 10 that interact with LarB1. A sequence alignment of precursors from split-B gene clusters (Figure S18A) reveals a conserved Tyr/Trp, which is 15 residues upstream of the conserved Thr residue in lasso peptide leader peptides.[16] This corresponds to Tyr-10 in LarA (Figure 2A). In addition, there is a highly conserved proline 12 residue upstream of the conserved Thr residue, position 13 in LarA. Photocrosslinking data from LarA F21Z and the recent data on alanine scans of PqqD homologues acting as RREs[12] suggest that LarB1 likely engages the leader peptide of LarA in a fashion similar to how NisB and LynD engage their leader peptides. Hence we also aligned B1 protein sequences looking for conserved residues in β3 and α3 that are expected to interact with the leader peptide. This revealed a few well conserved residues (Figure S18B), including an Asp-69 residue in the upper region of the α3 that likely interacts with the conserved Tyr-10 and Pro-13 region of the leader peptide (Figure 6B). This interaction would allow the downstream region of LarA that includes Phe-21 to interact with the Tyr-28 region of LarB1 β3.

We turned to the interactive Rosetta user interface FoldIt[28] to generate the docking model (Figure 6A). FoldIt has been used to manually generate a structure based on NMR NOEs, which provide distance restraints.[29] We used our AzF cross-linking data as distance restraints during manual docking of a segment comprising residues 8−35 of LarA (Figure 2A). The leader peptide was first docked into the LarB1 model, satisfying the requirement for LarA Phe-21 to be close to LarB1 Tyr-28 (Figure 6B). The conserved Pro-13 introduced a kink in the leader peptide sequence that naturally guided the location of the docking, with the leader peptide docking in a similar fashion to NisB and LynD. This docking predicts an H-bond interaction between the conserved Asp-69 in LarB1 with relatively well-conserved Tyr-10 in LarA. If this interaction is crucial for leader peptide positioning, it may explain why we saw no cross-linking with the Y10Z LarA construct (Figure 5).

DOI: 10.1021/acscentsci.6b00184
ACS Cent. Sci. 2016, 2, 702−709

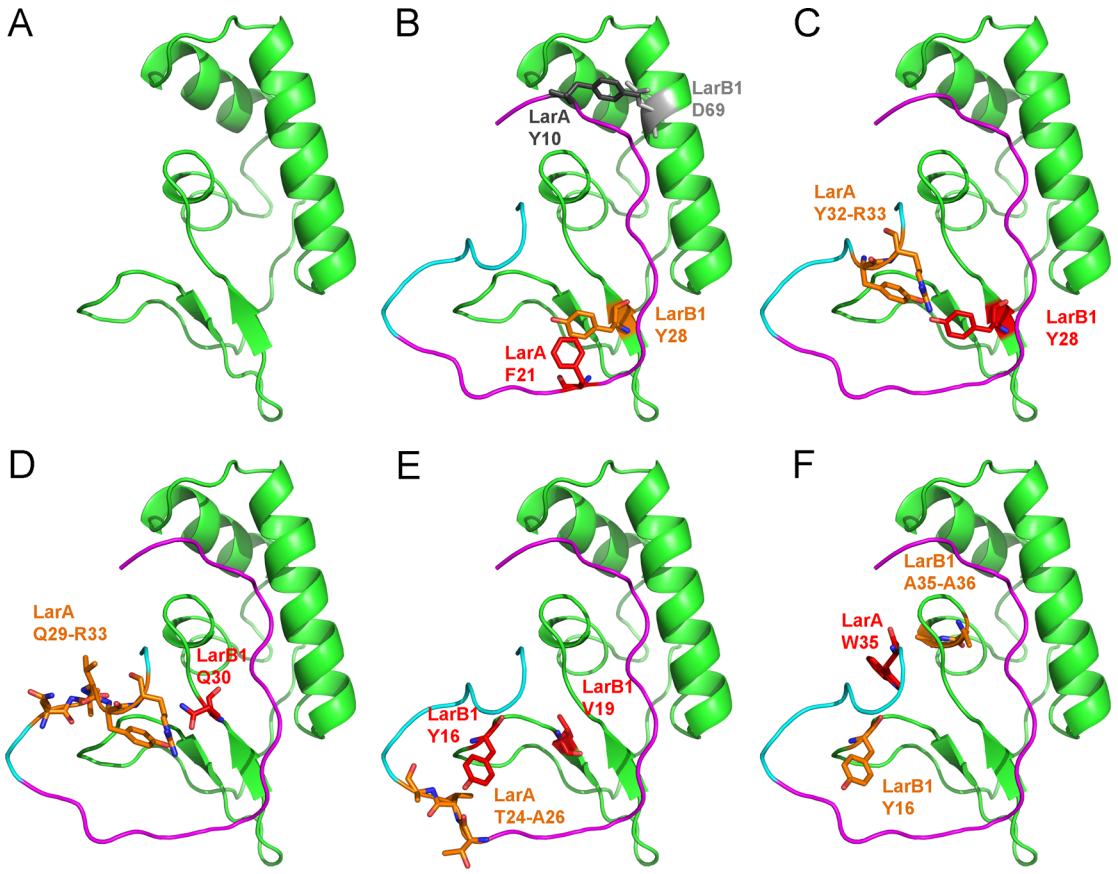


Figure 6. Docking model of LarA and LarB1. Model manually generated using FoldIt with homology models, BLI data, and cross-linking data as distance restraints. AzF constructs analyzed are shown in red, and corresponding regions of cross-linking are shown in orange. LarB1 is shown in green, LarA leader peptide residues 8−26 are shown in magenta, and LarA core peptide residues 27−35 are shown in cyan. (A) LarB1 homology model after LarA docking (LarA hidden). Note the shift of the β1 and β2 strands as compared to the homology model before docking (Figure 5C). (B) Predicted interaction from bioinformatics analysis between LarA Y10 (dark gray) and LarB1 D69 (light gray); LarA F21Z cross-linked to LarB1 Y28. (C) LarB1 Y28Z cross-linked to the Y32−R33 region of LarA. (D) LarB1 Q30Z cross-linked to the Q29−R33 region of LarA. (E) LarB1 Y16Z and LarB1 V19Z both cross-linked to the T24−A26 region of LarA. The LarB1 V19Z restraint could not be satisfied, likely because substitution of V19, a buried hydrophobic residue, with the larger AzF caused a structural change to LarB1. (F) LarA W35Z cross-linked to two regions, LarB1 Y16 and LarB1 A35−A36, that are far apart in amino acid sequence but close together in space.

The replacement of a hydroxyl group in Tyr with the azide group in AzF would abrogate the H-bond.

Core peptide residues were then added to the precursor peptide up to Trp-35. The AzF cross-linking data were used as distance restraints to dock this region of the core peptide (Figures 6C−6F). Except for the restraint from the LarB1 V19Z construct, all other distance restraints were satisfied in the model with the lowest estimated energy. Val-19 is in the general vicinity of the Thr-24−Ala-26 region of LarA but is facing inward into the hydrophobic core of LarB1. We suspect that substitution of the larger AzF residue at Val-19 can cause a significant structural change, perhaps flipping the residue out of the core due to steric clashes.

The photocrosslinking data showed significant interactions between the LarB1 protein and the core peptide of LarA. This biochemical data allowed us to generate a model of 28 aa of the 46 aa precursor in the bound form. The model shows that, in addition to the expected leader peptide binding function of LarB1, the LarB1 protein is also positioning the core peptide, causing it to loop back and interact with the β1 strand and α1 helix of the protein. The scissile bond between the leader and core peptide regions of LarA is solvent-exposed in this model, presumably positioning it for cleavage by the cysteine protease LarB2. It would be of interest to study the cross-linking behavior of LarA in the presence of both LarB1 and LarB2. We attempted to produce LarB2 on its own and as a fusion protein, but were unsuccessful in generating soluble protein. The aggregation propensity of B2 enzymes has been previously noted,[12] and intact B enzymes such as McjB copurify with chaperones.[6] We cannot rule out the possibility that some of the LarA−LarB1 contacts we observe would be disturbed in the presence of LarB2. Given that LarB2 is a relatively small protein at 147 aa long, we argue that it is unlikely that LarB2 would disrupt all of the LarA-LarB1 contacts observed.

Functional Significance of Core Peptide−B1 Protein Interactions. We wished to determine whether residues in LarB1 that are implicated in core peptide binding had any effect on the biosynthesis of lasso peptides. Unfortunately, there is no heterologous expression for lariatin currently available. However, Y28 of LarB1, one of the residues that cross-linked the core peptide, is highly conserved among LarB1 homologues (Figure S18B) and lasso peptide B proteins more broadly (Figure S19A). As discussed above (Figure S12) and elsewhere,[12] the Y28A substitution has a minimal effect on leader peptide binding, so any effects this substitution has on lasso peptide production can instead be ascribed to tyrosine's

DOI: 10.1021/acscentsci.6b00184
ACS Cent. Sci. 2016, 2, 702−709

interaction with the core peptide. We introduced the corresponding Tyr to Ala substitution into the B protein in the astexin-3 heterologous expression system (Figure S19A).[30] This substitution reduced the titer of astexin-3 produced by more than 15-fold (Figure S19B), clearly showing a role for this Tyr residue in lasso peptide biosynthesis.

CONCLUSION

Here we have used photocrosslinking to map the interaction of the lariatin precursor, LarA, with its PqqD homologue LarB1. LarB1 binds the leader peptide region of the lariatin precursor protein LarA with an affinity of ∼400 nM, in agreement with other recently characterized PqqD homologues. Using the photocrosslinking amino acid AzF, we mapped the interaction of LarA with LarB1. This demonstrated that the leader sequence interaction likely resembles the interaction of the lantibiotic precursor NisA with its maturation enzyme NisB. More interestingly, this work demonstrated that LarB1 interacts with the core sequence as well, likely positioning it to be processed by the maturation enzymes LarB2 and perhaps LarC. In the docking model generated using AzF data as distance restraints, the N-terminus of the core sequence is in a solvent exposed loop (Figure 6), accessible for cleavage by the LarB2 enzyme. These new data further elucidate the mechanism of lasso peptide B enzymes: the conserved N-terminal motif of B enzymes, whether fused to the proteolytic domain or as a standalone protein, interacts with the leader peptide with high affinity. The docking of the leader peptide presumably can dramatically increase the low affinity binding of the core peptide, which was not detectable without the leader peptide, via a local concentration effect. The combination of leader peptide and core peptide binding may line up the precursor peptide substrate for subsequent cleavage at the cysteine catalytic triad of B2 enzymes. The photocrosslinking approach is particularly well-suited for observing weaker interactions such as the LarA core peptide−LarB1 interaction. Because these peptide regions are not expected to be well-ordered, they may not be observable using either NMR or crystallographic techniques. In addition, it is likely that this photocrosslinking approach can find utility in studying large protein complexes that are not amenable to crystallographic approaches.

The prevalence of so-called "split-B" enzymes, in which the lasso peptide protease activity is divided between two proteins, in lasso peptide biosynthetic gene clusters has been previously noted in our genome mining studies.[31] The split-B enzymes appear largely in lasso peptide gene clusters from Gram-positive bacteria where there is a major clade of lasso peptide producers among the actinobacteria.[7] Experimentally confirmed examples include *R. jostii* discussed herein, *Streptomyces roseosporus*,[32] and *Streptomonospora alba*, the producer of streptomonomicin.[11] In contrast, most of the Gram-negative proteobacterial lasso peptide producers[10,30] use a single protein for lasso peptide protease activity. With our demonstration that LarB1 binds both the leader peptide and core peptide of its cognate precursor protein, there is mounting evidence that there are two different catalytic solutions to the problem of precursor peptide binding and cleavage. However, it remains unclear why a subset of Gram-positive bacteria prefers to use a split-B enzyme whereas many other bacteria utilize a single protein.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscentsci.6b00184.

Detailed methods and figures (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: ajlink@princeton.edu. 207 Hoyt Laboratory, Princeton University, Princeton, NJ 08544. Phone: (609) 258-7191.

Present Address

‡M.Y.C.: Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030.

Author Contributions

W.L.C. and A.J.L. conceived of the study. W.L.C. and M.Y.C. performed all experiments. W.L.C. and M.O.M. performed bioinformatics analysis. W.L.C. and A.J.L. wrote the paper.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by NIH Grant GM107036. A.J.L. is a Sloan Research Fellow. We thank Tharan Srikumar and Saw Kyin for assistance with LC/MS−MS. We thank Galia Debelouchina and Tom Muir for assistance with circular dichroism. We thank Grant Murphy for assistance with FoldIt.

REFERENCES

(1) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Goransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Muller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Sussmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **2013**, *30*, 108−160.

(2) Oman, T. J.; van der Donk, W. A. Follow the leader: The use of leader peptides to guide natural product biosynthesis. *Nat. Chem. Biol.* **2010**, *6*, 9−18.

(3) Maksimov, M. O.; Pan, S. J.; Link, A. J. Lasso peptides: Structure, function, biosynthesis, and engineering. *Nat. Prod. Rep.* **2012**, *29*, 996−1006.

(4) Duquesne, S.; Destoumieux-Garzon, 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*, 793−803.

(5) Pan, S. J.; Rajniak, J.; Cheung, W. L.; Link, A. J. Construction of a single polypeptide that matures and exports the lasso peptide microcin J25. *ChemBioChem* **2012**, *13*, 367−370.

(6) Yan, K. P.; Li, Y.; Zirah, S.; Goulard, C.; Knappe, T. A.; Marahiel, M. A.; Rebuffat, S. Dissecting the maturation steps of the lasso peptide microcin J25 in vitro. *ChemBioChem* **2012**, *13*, 1046−1052.

(7) Maksimov, M. O.; Pelczer, I.; Link, A. J. Precursor-centric genome-mining approach for lasso peptide discovery. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 15223−15228.

DOI: 10.1021/acscentsci.6b00184
ACS Cent. Sci. 2016, 2, 702−709

(8) Iwatsuki, M.; Tomoda, H.; Uchida, R.; Gouda, H.; Hirono, S.; Omura, S. Lariatins, antimycobacterial peptides produced by *Rhodococcus sp*. K01-B0171, have a lasso structure. *J. Am. Chem. Soc.* **2006**, *128*, 7486−7491.
(9) Inokoshi, J.; Matsuhama, M.; Miyake, M.; Ikeda, H.; Tomoda, H. Molecular cloning of the gene cluster for lariatin biosynthesis of *Rhodococcus jostii* K01-B0171. *Appl. Microbiol. Biotechnol.* **2012**, *95*, 451−460.
(10) Hegemann, J. D.; Zimmermann, M.; Zhu, S.; Klug, D.; Marahiel, M. A. Lasso peptides from proteobacteria: Genome mining employing heterologous expression and mass spectrometry. *Biopolymers* **2013**, *100*, 527−542.
(11) Metelev, M.; Tietz, J. I.; Melby, J. O.; Blair, P. M.; Zhu, L.; Livnat, I.; Severinov, K.; Mitchell, D. A. Structure, bioactivity, and resistance mechanism of streptomonomicin, an unusual lasso peptide from an understudied halophilic actinomycete. *Chem. Biol.* **2015**, *22*, 241−250.
(12) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol.* **2015**, *11*, 564−570.
(13) Regni, C. A.; Roush, R. F.; Miller, D. J.; Nourse, A.; Walsh, C. T.; Schulman, B. A. How the MccB bacterial ancestor of ubiquitin E1 initiates biosynthesis of the microcin C7 antibiotic. *EMBO J.* **2009**, *28*, 1953−1964.
(14) Ortega, M. A.; Hao, Y.; Zhang, Q.; Walker, M. C.; van der Donk, W. A.; Nair, S. K. Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB. *Nature* **2015**, *517*, 509−512.
(15) Koehnke, J.; Mann, G.; Bent, A. F.; Ludewig, H.; Shirran, S.; Botting, C.; Lebl, T.; Houssen, W. E.; Jaspars, M.; Naismith, J. H. Structural analysis of leader peptide binding enables leader-free cyanobactin processing. *Nat. Chem. Biol.* **2015**, *11*, 558−563.
(16) Cheung, W. L.; Pan, S. J.; Link, A. J. Much of the microcin J25 leader peptide is dispensable. *J. Am. Chem. Soc.* **2010**, *132*, 2514−2515.
(17) Velterop, J. S.; Sellink, E.; Meulenberg, J. J.; David, S.; Bulder, I.; Postma, P. W. Synthesis of pyrroloquinoline quinone in vivo and in vitro and detection of an intermediate in the biosynthetic pathway. *J. Bacteriol.* **1995**, *177*, 5088−5098.
(18) Latham, J. A.; Iavarone, A. T.; Barr, I.; Juthani, P. V.; Klinman, J. P. PqqD is a novel peptide chaperone that forms a ternary complex with the radical S-adenosylmethionine protein PqqE in the pyrroloquinoline quinone biosynthetic pathway. *J. Biol. Chem.* **2015**, *290*, 12908−12918.
(19) Barr, I.; Latham, J. A.; Iavarone, A. T.; Chantarojsiri, T.; Hwang, J. D.; Klinman, J. P. Demonstration that the radical S-adenosylmethionine(SAM) enzyme PqqE catalyzes de novo carbon-carbon cross-linking within a peptide substrate PqqA in the presence of the peptide chaperone PqqD. *J. Biol. Chem.* **2016**, *291*, 8877−8884.
(20) Roush, R. F.; Nolan, E. M.; Lohr, F.; Walsh, C. T. Maturation of an Escherichia coli ribosomal peptide antibiotic by ATP-consuming N-P bond formation in microcin C7. *J. Am. Chem. Soc.* **2008**, *130*, 3603−3609.
(21) Link, A. J. Biosynthesis: Leading the way to RiPPs. *Nat. Chem. Biol.* **2015**, *11*, 551−552.
(22) Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein structure and function prediction. *Nat. Methods* **2015**, *12*, 7−8.
(23) Futran, A. S.; Kyin, S.; Shvartsman, S. Y.; Link, A. J. Mapping the binding interface of ERK and transcriptional repressor Capicua using photocrosslinking. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 8590−8595.
(24) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of p-azido-L-phenylalanine to the genetic code of Escherichia coli. *J. Am. Chem. Soc.* **2002**, *124*, 9026−9027.
(25) Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11020−11024.
(26) Tippmann, E. M.; Liu, W.; Summerer, D.; Mack, A. V.; Schultz, P. G. A genetically encoded diazirine photocrosslinker in Escherichia coli. *ChemBioChem* **2007**, *8*, 2210−2214.
(27) Leyva, E.; Platz, M. S.; Persy, G.; Wirz, J. Photochemistry of phenyl azide: The role of singlet and triplet phenylnitrene as transient intermediates. *J. Am. Chem. Soc.* **1986**, *108*, 3783−3790.
(28) Cooper, S.; Khatib, F.; Treuille, A.; Barbero, J.; Lee, J.; Beenen, M.; Leaver-Fay, A.; Baker, D.; Popovic, Z.; Players, F. Predicting protein structures with a multiplayer online game. *Nature* **2010**, *466*, 756−760.
(29) Kier, B. L.; Anderson, J. M.; Andersen, N. H. Disulfide-mediated beta-strand dimers: Hyperstable beta-sheets lacking tertiary interactions and turns. *J. Am. Chem. Soc.* **2015**, *137*, 5363−5371.
(30) Maksimov, M. O.; Link, A. J. Discovery and characterization of an isopeptidase that linearizes lasso peptides. *J. Am. Chem. Soc.* **2013**, *135*, 12038−12047.
(31) Maksimov, M. O.; Link, A. J. Prospecting genomes for lasso peptides. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 333−344.
(32) Kersten, R. D.; Yang, Y. L.; Xu, Y.; Cimermancic, P.; Nam, S. J.; Fenical, W.; Fischbach, M. A.; Moore, B. S.; Dorrestein, P. C. A mass spectrometry-guided genome mining approach for natural product peptidogenomics. *Nat. Chem. Biol.* **2011**, *7*, 794−802.

DOI: 10.1021/acscentsci.6b00184
ACS Cent. Sci. 2016, 2, 702−709