Supplementary Information

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[†‡]*

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

* To whom correspondence should be addressed: ajlink@princeton.edu, Phone: 609-258-7191, 207 Hoyt Laboratory, Princeton, NJ 08544

Methods and Materials

Plasmid Expression Vector Construction

All PCR reactions in this study were done using the Picomaxx High Fidelity PCR System from Agilent with 200 μM of each dNTP and 2.5 U of Picomaxx enzyme mix in a total volume of 50 μL in 1x Picomaxx buffer. All primers were purchased from Integrated DNA Technologies (IDT) and listed in Table S1.

Gene Assembly of larA and larB1

The sequence of the lariatin gene cluster has been previously reported.[1] *larA* and *larB1* were synthesized using assembly PCR. First the primers required for assembly were designed using DNAWorks(v3.2.2).[2] Assembly PCR consisted of an initial assembly PCR step followed by another PCR step to amplify the assembled gene. The initial assembly PCR step contained 0.1 μM of each primer and carried out with an annealing temperature of 52 °C and thirty cycles. A 1 μL sample of the initial assembly reaction was then used as template in a second amplification PCR with 1 μM each of the first forward and last reverse primer used for the assembly. *larB1* was cloned via *Nco*I and *Bgl*II into pQE-60, which contains a His_6 C-terminal tag (pMC01). *larA* was cloned via *EcoR*I and *Hind*III into pQE-80, removing the His_6 N-terminal tag originally in the vector (pWC44).

Plasmid Expression Vector Cloning

LarB1-His_6 was expressed from the pQE-60 construct pMC01. All other constructs to express LarB1-His_6 variants were also cloned into pQE-60. Constructs to express LarA-IEGR-MBP (pWC45), IEGR-MBP (pWC46), and other LarA-MBP fusion variants were cloned into pQE-80. A complete table of all the constructs created can be found in Table S2.

To construct pWC45, *larA* was amplified from pWC44 and *malE* from pMAL-c2x with each PCR product containing an overlapping region with codons that encode for the Factor Xa Protease recognition site IEGR. The genes were then overlapped together in a second PCR step and cloned via *Eco*RI and *Hind*III into pQE-80. To construct pWC46, the *malE* gene was amplified from pMAL-c2x, with primers that append codons for MIEGR to the N-terminal end. The PCR product was then cloned via *Eco*RI and *Hind*III into pQE-80. To construct LarA-MBP azidophenylalanine variants, forward and reverse primers were designed to introduce a TAG amber stop codon to replace the codon at the desired position in *larA*. In the first PCR step, the fragment of *larA* upstream of the mutation site and the fragment of *larA-IEGR-malE* downstream of the mutation site were amplified from pWC45. The two fragments were then overlapped in a second PCR step, then cloned via *Eco*RI and *Hind*III into pQE-80. The construction of LarB1-His_6 azidophenylalanine variants was similar, but cloned into pQE-60 via *Nco*I and *Bgl*II. To construct LarA-MBP variants with truncations of LarA from the N-terminus of the leader peptide, forward primers were designed to exclude the designed number of codons but retaining the ATG start codon and *Eco*RI site upstream. To construct LarALP-GSSG-MBP and LarACP-GSSG-MBP, *larALP/larACP* and *malE* were amplified from pWC45 with primers to include a GSSG linker, overlapped in a second PCR step, then cloned via *Eco*RI and *Hind*III into pQE-80.

The astexin-3 cluster in pMM65 has been previously described.[3] The astexin-3 cluster with the AtxB Y41A substitution (pWC89) was cloned by generating the *atxB* Y41A mutation via overlap PCR (see Table S1) and then restriction digested and cloned into pMM65 between the *Eco*RI and *Aat*II sites.

Protein Expression

All protein expression was done in *E. coli* strain BL21. All cultures contained 100 mg/L ampicillin to maintain the pQE-based plasmids. For the expression of azidophenylalanine variants, the pQE-based plasmids were co-transformed with pEVOL-pAzF, which contains a tRNA synthetase/tRNA pair for *in vivo* incorporation of 4-azido-L-phenylalanine in response to the TAG codon[4]. To maintain the pEVOL-pAzF plasmid, 25 mg/L chloramphenicol was added to the cultures in addition to ampicillin.

Expression of LarB1-His_6

LB culture was inoculated with 1:100 volume of an overnight culture and grown at 37 °C, 250 rpm until the OD_{600} reached 1. The culture was then induced with 1 mM IPTG and expressed for 3 hours at 37°C. The cells were pelleted at 4000 x g for 20 min at 4°C and stored at -20°C until purification.

Expression of LarA-IEGR-MBP, IEGR-MBP, LarALP-GSSG-IEGR-MBP and LarACP-GSSG-IEGR-MBP

LB culture with 0.2 wt% glucose was inoculated with 1:100 volume of an overnight culture and grown at 37 °C, 250 rpm until the OD_{600} reached 0.5-0.7. The culture was then induced with 1 mM IPTG and expressed at 25 °C for 4 hours. The cells were pelleted at 4000 x g for 20 min at 4 °C, then resuspended in 25 mL of Column Buffer per 1L of culture (Column Buffer was made as described by New England Biolabs: 20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, pH 7.4) and frozen at -20 °C until purification.

Expression of LarA-MBP Azidophenylalanine Mutants

LB culture with 0.2 wt% glucose was inoculated with 1:100 volume of an overnight culture and grown at 37 °C, 250 rpm until the OD_{600} reached 0.6. Unnatural amino acid 4-azido-

L-phenylalanine (Chem-Impex International) was resuspended in water to 100 mM with a minimal amount of NaOH added dropwise to dissolve the powder. It was then added to the culture to a final concentration of 2 mM. The culture was allowed to shake at 25 °C for 10 minutes, after which it was simultaneously induced with 0.2 wt% arabinose and 1 mM IPTG. Expression was carried out for 5 hours at 25 °C and 250 rpm, after which the cells were pelleted, resuspended in Column Buffer, and frozen as described above for the expression of the other MBP fusions.

Expression of LarB1-His_6 Azidophenylalanine Variants

The expression conditions for LarB1-His_6 azidophenylalanine variants were the same as for LarA-MBP azidophenylalanine mutants except no glucose was added and expression was carried out for 5.5 hours. The cells were pelleted at 4000 x g for 20 min at 4 °C and stored at -20 °C until purification.

Expression of Astexin-3 Constructs

500 mL of M9 media was inoculated with an overnight culture of pMM65 or pWC89 to a starting OD_{600} of 0.02. The cultures were grown at 37 °C, 250 rpm until OD_{600} reached 0.25-0.3. The cultures were then induced with anhydrotetracycline and grown at 20 °C for 20 hours. The cells were pelleted at 4000 x g for 20 min at 4 °C, then washed with 10 mL PBS media. The cells were re-pelleted at 4000 x g for 20 min at 4 °C.

Protein Purification

LarB1-His_6

LarB1-His_6 and its azidophenylalanine-incorporated variants were purified based on Qiagen Ni-NTA batch purification under denaturing conditions and subsequently buffer-

exchanged and refolded. The frozen cell pellet was thawed and resuspended in 40 mL Buffer B (100 mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH 8.0) per 1 L of culture. The resuspended cells were then frozen at -80°C for 45 minutes then allowed to thaw at room temperature. The lysate was then centrifuged at 10,000 x g for 15 minutes to pellet the cell debris. The cleared lysate was incubated with 4 mL Ni-NTA slurry (Qiagen 50% Ni-NTA agarose slurry) at 4 °C on a rotary shaker for 1 hour then loaded into a chromatography column at room temperature. The resin was washed two times with 20 mL of Buffer C (Buffer B adjusted to pH 6.3), followed by two washes with 10 mL of Buffer D (Buffer B adjusted to pH 5.9), and eluted with Buffer E (Buffer B adjusted to pH 4.5).

MBP fusions

MBP and MBP-tagged LarA variants were purified based on the affinity chromatography protocol from NEB. Purification steps were all done at 4 °C. The frozen cells suspended in Column Buffer were thawed in an ice-water bath then sonicated in an ice-water bath with 10 x 15 second pulses. The lysate was then centrifuged at 20,000 x g for 20 minutes. The cleared lysate was diluted 1:6 with Column Buffer. Next 2 mL amylose resin bed (NEB) per 1L culture was loaded into a chromatography column and washed with five column volumes of Column Buffer. The diluted lysate was then loaded onto the column. After loading the lysate, the resin was washed twice with 10 column volumes per wash and then eluted with 10 mM maltose in Column Buffer.

Protein Analysis

For all protein purification above, the purity and quantity were first qualitatively assessed via a PAGE gel. The elutions with the highest amounts of protein were pooled together for a total volume of 3 mL, which were then buffer exchanged into 1x PBS buffer (pH 7.4) using Qiagen

Econo-Pac 10DG desalting columns. The purified proteins were quantified using a Nanodrop spectrophotometer with extinction coefficients calculated by the peptide property calculator provided by Northwestern University with the exception of LarB1-His_6 azidophenylalanine variants; these constructs were measured using a BCA assay (Thermo Scientific Pierce BCA Protein Assay Kit) because of an absorbance irregularity with the Nanodrop due to the unnatural amino acid. BCA assay was also used to measure some LarA-MBP variants and LarB1-His_6 to adjust for the concentration difference obtained using the two quantification methods. In general, the BCA assay measured approximately 60% higher than Nanodrop measurements.

LarA leader peptide

A crude synthesis of LarA leader peptide (LarALP) was purchased from Genscript and then purified via reverse-phase high-performance liquid chromatography (HPLC) using an Agilent LC system (1200 Series) on a semi-preparative column (Zorbax 300SB-C18, 9.4 X 250 mm from Agilent). The crude synthesis powder was resuspended in 10% acetonitrile/water with 0.1% trifluroacetic acid (TFA) to a concentration of approximately 1.7 mg/mL. For each run, 200 µL of the crude synthesis was injected. LarALP was purified using an acetonitrile/water gradient (with 0.1% TFA) from 10% to 50% over 19 minutes starting at the first minute with a 4 mL/min flow rate. The peak at 10 minutes was fraction collected. The collected fractions were then pooled together and frozen at -80 °C for a few hours and then lyophilized (Labconco FreeZone Freeze Dry System). The lyophilized peptide was then resuspended in 2 mL H_2O and the concentration was measured with a BCA assay (Pierce BCA Protein Assay Kit) in triplicate. The peptide was then aliquoted, frozen, and lyophilized again as described above.

Astexin-3

The cell pellets were extracted with 10 mL methanol (MeOH). The crude extracts were dried via rotavap, resuspended in 2 mL H_2O, then extracted through a 1 mL C8 column. The C8 MeOH extracts were dried via speedvac, then resuspended in 250 μL 50% acetonitrile/50% water. Production level was measured by HPLC, injecting 10 μL of the samples onto an analytical scale C18 column (Zorbax 300SB-C18, 4.6 × 150 mm). A linear gradient from 10% acetonitrile/water (0.1% TFA) to 50% acetonitrile/water (0.1% TFA) over 20 minutes followed by a linear gradient to 90% acetonitrile/water (0.1% TFA) in 5 minutes was used.

Bio-layer Interferometry Binding Assay

Binding kinetics were measured using bio-layer interferometry with the BLItz System and Ni-NTA biosensors from ForteBio. All samples were prepared so that they were in 1x Kinetics Buffer (1x PBS pH 7.4 with 0.13% BSA and 0.013% Tween-20). First all the biosensors (1 biosensor per sample) were rehydrated with 200 μL 1x Kinetics Buffer in a 96 well plate for at least 10 minutes, and kept hydrated until use. One biosensor was attached to the BLItz system at a time and the following binding kinetics program was used: 30 second baseline measurement where the biosensor was dipped into a fresh tube with 250 μL 1x Kinetics Buffer, 120 second loading of LarB1-His_6 or a variant of it where the biosensor was dipped into 4 μL of 1 μM LarB1-His_6 on a magnetic drop holder, 30 second baseline measurement where the biosensor was dipped into the tube with Kinetics Buffer, 120 second binding step where the biosensor was dipped into 4 μL of Kinetics Buffer (initial blank reference) or the analyte on a second magnetic drop holder, 120 second dissociation step where the biosensor was dipped into the tube with Kinetics Buffer. A new biosensor and buffer tube was used for each sample and the drop holders were washed with 10 μL Kinetics Buffer three times and dried with a kimwipe in

between runs. For quantitative kinetic measurements between wild-type LarB1 and LarA, four to nine concentrations of LarA-MBP spanning 50 nM to 2000 nM were tested each time, for a total of three sets of data over ten concentration values. The 18 curves were each step corrected for the start of association and dissociation, then fitted locally to obtain values for the association rate constant k_a ($M^{-1}s^{-1}$) , dissociation rate constant k_d (s^{-1}), and binding affinity $K_d = k_d/k_a$ (M). The reported K_d value is the average of the 18 K_d values and the corresponding standard deviation.

LarB1-His_6 and LarA-MBP Photocrosslinking

Photocrosslinking reaction

Photocrosslinking reactions were generally done using 2 μM each of LarB1-His_6 and LarA-MBP, unless otherwise specified, with one of the two constructs containing a site specific azidophenylanine substitution. Reactions were carried out in 75 μL 1x PBS pH 7.4 in a clear 96 well polystyrene plate with a lid. The plate was placed on top of a UVP transilluminator (model LMW-20) set to 365 nm UV light in a 4 °C cold room. Initially small-scale experiments were done with different times of UV exposure, using one well per time point. The reaction products were analyzed using PAGE. Constructs that showed good crosslinking at the standard concentrations (those that were easily visible after about 10 min of UV exposure) were then scaled-up to 40-50 x 75 μL reactions and exposed to UV for 5-7 minutes at a time with a 3-5 minute cooling period in between for a total exposure time of one hour.

Removal of MBP tag and purification of cross-linked peptide

The reactions were then pooled and excess LarB1-His_6 was removed via size exclusion using an Amicon Ultra- 4 mL 30 kDa centrifugal filter unit. The reaction was concentrated at 4000 rpm (Thermo Sorvall Legend RT+ centrifuge, swinging bucket rotor) for 10-15 minutes until an approximate volume of 300 μL. Next 4 mL of 20 mM Tris-HCl with 100 mM NaCl and

2 mM $CaCl_2$ (pH 8.0) was added and the filter unit was centrifuged again until 300 µL. The addition of Tris buffer and centrifugation was repeated two more times to obtain a buffer exchanged reaction with much of the excess LarB1-His_6 removed. The final 300 µL of buffer exchanged cross-linking reaction was digested with 2.4 µL of Factor Xa (1 mg/mL NEB) in a 1.5 mL tube for 12 hours at room temperature. The digested reaction was then incubated with 50 µL of Ni-NTA slurry at 4°C for two hours on a rotary shaker to bind the LarA-LarB1-His_6 cross-linked peptide. The resin was spun down briefly at 1000 x g and the supernatant was discarded by pipetting. The resin was then washed 3-4 times with 200 µL of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0), spinning down and decanting the supernatant each time. Finally the resin was eluted 3 times with 100 µL of elution buffer each (wash buffer with 250 mM imidazole), spinning down and saving the supernatant each time.

Trypsin Digestion

First 20 µL of purified LarA-LarB1-His_6 cross-linked peptide was denatured with 20 µL 8M urea, 50 mM Tris-HCl, pH 8.0 at room temperature for ten minutes. The peptide was then diluted five-folds with 50 mM ammonium bicarbonate and digested for 16 hours at 37 °C with 0.1 µg of sequencing grade trypsin (Promega). The digested product was then cleaned up for mass spectrometry using a C18 spin column (Pierce). The standard protocol from Pierce was used, resulting in digested peptide eluted in 40 µL 70% acetonitrile.

Mass Spectrometry

MALDI-TOF

Samples for MALDI-TOF were prepared by mixing 1 µL of trypsin digested peptide 50:50 with 2.5 mg/mL α-cyano-4-hydroxycinnamic acid (Sigma) matrix. The sample was prepped in duplicate and spotted onto an Applied Biosystems 384 Opti-TOF 123 mm x 81 mm

SS plate and allowed to dry. Molecular weights of digested peptide fragments were determined with an ABI-MDS SCIEX 4800 MALDI-TOF/TOF instrument in positive-ion mode.

LC-MS/MS

The rest of the trypsin digestion was sent to the Princeton University Department of Molecular Biology Proteomics & Mass Spectrometry Core Facility where LC-MS/MS was done using a Thermo Orbitrap Elite or QExactive instrument.

Circular Dichroism

Purified LarB1-His_6 was buffer exchanged into 10 mM potassium phosphate, 100 mM ammonium sulfate buffer (pH 7.3) using an Amicon Ultra- 4 mL 3 kDa centrifugal filter unit. The concentration was measured via a Nanodrop spectrophotometer as described in the protein purification section. The protein was then diluted to 10.5 μM and used for circular dichroism measurements. The blank buffer and protein samples were each measured five times at 20 °C in a 1 mm pathlength cuvette, scanning from 190 nm to 250 nm. The final LarB1 CD spectrum was obtained by averaging the LarB1 spectra and subtracting the average background spectrum.

Docking Model

The LarB1 I-TASSER homology model was first loaded into the standalone version of FoldIt and refined. PyMOL was then used to draw in residues 8-26 of the leader peptide (KTYNAPSLVQRGKFARTTA). The backbone of LarB1 was initially frozen in place to dock the leader peptide. The conserved Pro-13 residue in LarALP introduces a natural kink to the peptide, guiding the docking of the truncated LarALP into the binding groove of LarB1 between the β3 strand and α3 helix in a similar fashion to how NisA is bound to NisB. This placed LarA Phe-21 in close proximity to LarB1 Tyr-28, satisfying the first distance constraint from our

crosslinking data. The backbone of LarB1 was then unfrozen to allow refinement of the leader peptide-docked model. The calculated Rosetta energy improved with the peptide docked, supporting the model that LarB1 engages LarALP in a way analogous to how NisB and LynD engage the leader sequence of their substrates.

PyMOL was then used to extend the truncated LarALP up to Trp-35, the first residue of the tail in the core peptide. The backbone of LarB1 and of the docked leader peptide were initially frozen. The other five crosslinking interactions were used as distance restraints during manual docking of the core peptide. The core peptide was docked using the rubber-band feature of FoldIt to maintain the distance restraints. The model was then unfrozen and refined to both improve the docking energy and agreement with crosslinking data. The process to dock the core peptide was repeated a few times to explore alternative docking positions to satisfy the distance restraints but each trial either resulted in a similar final model state as shown in Figure 6 or had poorer Rosetta energy.

References

(1) 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.

(2) Hoover, D. M.; Lubkowski, J. Dnaworks: An automated method for designing oligonucleotides for PCR-based gene synthesis *Nucleic Acids Res.* **2002**, *30*, e43.

(3) Maksimov, M. O.; Koos, J. D.; Zong, C.; Lisko, B.; Link, A. J. Elucidating the specificity determinants of the AtxE2 lasso peptide isopeptidase *J. Biol. Chem.* **2015**, *290*, 30806-12.

(4) 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.

Supplementary Tables and Figures

Table S1.**Primers used in this study.**

Primer Name	Primer Sequence	Purpose
LarB1-1	GCAATGCCATGTCCATGGTTCTGCGCC	LarB1 gene assembly
LarB1-2	ATTCGGTCGGGGTTATAATCACATTTTTACGCAGGCGCAGAACCATGGAC	LarB1 gene assembly
LarB1-3	TGATTATAACCCCGACCGAATATGGCGCGGTGGCTTTAGATGAGCGTAGC	LarB1 gene assembly
LarB1-4	CGCGGTGCTATTCAGCTGGTAATAATCGCCGCTACGCTCATCTAAAGCCA	LarB1 gene assembly
LarB1-5	CAGCTGAATAGCACCGCGGCTCTGATTCTGGATCAACTGACCAAGAAAAT	LarB1 gene assembly
LarB1-6	ATACGGGCAGCAATGCTTTCCACCGGTATTTTCTTGGTCAGTTGATCCAG	LarB1 gene assembly
LarB1-7	AAAGCATTGCTGCCCGTATTGCGCTGGATTTTGAGGTGAGCAAAGCGCAG	LarB1 gene assembly
LarB1-8	CATACGCAGATACTCATCCAGATCCGCGCTAGCCTGCGCTTTGCTCACCT	LarB1 gene assembly
LarB1-9	TCTGGATGAGTATCTGCGTATGCTGCGTGAACAGGGCCTGTTGCGTAGAT	LarB1 gene assembly
LarB1-10	GTGTCAGAGCGAAGATCTACGCAACAGGCCCT	LarB1 gene assembly
larA oligo 1 F	CACACAGAATTCATTAAAGAGGAGAA	LarA gene assembly, pWC45, pWC47, pWC49
larA oligo 2 R	GGCTGGCTGGTCATAGTTAATTTCTCCTCTTTAATGAATTCTGTGT	LarA gene assembly
larA oligo 3 F	ATTAACTATGACCAGCCAGCCGAGCAAAAAAACCTATAATGCGCCG	LarA gene assembly
larA oligo 4 R	AAATTTACCACGCTGCACCAGGCTCGGCGCATTATAGGTTTTTTTG	LarA gene assembly
larA oligo 5 F	TGGTGCAGCGTGGTAAATTTGCGCGTACCACGGCGGGAAGTCAGCT	LarA gene assembly
larA oligo 6 R	TTGCTATGGCCCACCCATTCACGATACACCAGCTGACTTCCCGCCG	LarA gene assembly
larA oligo 7 F	TGGGTGGGCCATAGCAATGTGATTAAACCGGGTCCAATTGAAGGCC	LarA gene assembly
larA oligo 8 R	CTGTACAAGCTTACGGCCTTCAATTGGACCCG	LarA gene assembly
larA-Xa Reverse	ACGGCCTTCAATTGGACCCGG	pWC45
larA-Xa-malE Forward	GGTCCAATTGAAGGCCGTAAAATCGAAGAAGGTAAACTGGTAATCTG	pWC45
malE-HindIII Reverse	CTGCTCAAGCTTCCCGAGGTTGTTGTTATTGTTATTGTTG	All LarA-MBP constructs without azidophenylalanine
EcoRI-malE 1 F	GGAGAAATTAACTATGATTGAAGGCCGTAAAATCGAAGAAGGTAAACTGGTAATCTGG	pWC46
EcoRI-malE 2 F	CACACAGAATTCATTAAAGAGGAGAAATTAACTATGATTGAAGGCCG	pWC46
larA F21Azf For	GCGTGGTAAATAGGCGCGTACCAC	pWC47
larA F21Azf Rev	GTGGTACGCGCCTATTTACCACGC	pWC47
malE TAA HindIII Rev	CGACTAAGCTTACCCGAGGTTGTTGTTATTGTTATTGTTG	All LarA azidophenylalanine-MBP constructs
GSSG-malE For	CGGGCAGCAGCGGTATTGAAGGCCGTAAAATCGAAGAAGG	pWC49
GSSG-larALP Rev	CAATACCGCTGCTGCCCGCCGTGGTACGCGCAAATTTAC	pWC49
LarB1 Y28A	GTAGCGGCGATGCGTACCAGCTGAATAG	pWC57

For		
LarB1 Y28A Rev	CTATTCAGCTGGTACGCATCGCCGCTAC	pWC57
EcoRI LarA CP For	GAATTCATTAAAGAGGAGAAATTAACTATGGGAAGTCAGCTGG TGTATCGTGAATGGG	pWC58
pQE-60 NcoI For	GTGAGCGGATAACAATTTCACACAGAATTC	All LarB1 variants
pQE-60 BglII Rev	GCTTAGTGATGGTGATGGTGATGAGATC	All LarB1 variants
EcoRI LarA -10 For	GATATGAATTCATTAAAGAGGAGAAATTAACTATGGCGCCGAG CCTGGTGCAG	pWC60
LarA -15 For	GATATGAATTCATTAAAGAGGAGAAATTAACTATGCAGCGTGG TAAATTTGCGCGTACC	pWC61
LarA Y32AzF F	GTCAGCTGGTGTAGCGTGAATGGGTG	pWC66
LarA Y32AzF R	CACCCATTCACGCTACACCAGCTGAC	pWC66
LarA W35AzF F	GGTGTATCGTGAATAGGTGGGCCATAG	pWC67
LarA W35AzF R	CTATGGCCCACCTATTCACGATACACC	pWC67
EcoRI larA -5aa For	TATGAATTCATTAAAGAGGAGAAATTAACTATGAAAAAAACCT ATAATGCGCCGAGCCTG	pWC68
EcoRI LarA -9 For	GATATGAATTCATTAAAGAGGAGAAATTAACTATGAATGCGCC GAGCCTGGTGC	pWC71
larA Y10AzF For	CAAAAAAACCTAGAATGCGCCGAGCC	pWC72
larA Y10AzF Rev	GGCTCGGCGCATTCTAGGTTTTTTTG	pWC72
larB1 Y28AzF For	GTAGCGGCGATTAGTACCAGCTGAATAG	pWC73
larB1 Y28AzF Rev	CTATTCAGCTGGTACTAATCGCCGCTAC	pWC73
larB1 Y29AzF For	GCGGCGATTATTAGCAGCTGAATAGC	pWC74
larB1 Y29AzF Rev	GCTATTCAGCTGCTAATAATCGCCGC	pWC74
larB1 L31AzF For	CGATTATTACCAGTAGAATAGCACCGCG	pWC75
larB1 L31AzF Rev	CGCGGTGCTATTCTACTGGTAATAATCG	pWC75
larB1 N32AzF For	TTATTACCAGCTGTAGAGCACCGCGGC	pWC76
larB1 N32AzF Rev	GCCGCGGTGCTCTACAGCTGGTAATAA	pWC76
larB1 D27AzF For	CGGTGGCTTTATAGGAGCGTAGCGG	pWC77
larB1 D27AzF Rev	CCGCTACGCTCCTATAAAGCCACCG	pWC77
larB1 Q30AzF For	GGCGATTATTACTAGCTGAATAGCACCGCG	pWC78
larB1 Q30AzF Rev	CGCGGTGCTATTCAGCTAGTAATAATCGCC	pWC78
LarB1 Y16Z For	CCGACCGAATAGGGCGCGG	pWC79
LarB1 Y16Z Rev	CCGCGCCCTATTCGGTCGG	pWC79
LarB1 V19Z For	GGCGCGTAGGCTTTAGATGAG	pWC80
LarB1 V19Z Rev	CTCATCTAAAGCCTACGCGCC	pWC80

EcoRI Atx3B For	CTTCGCCAAAACGAATTCTACTCTC	pWC89
Atx3B Y41A Rev	GTTGAGGCAAAAGCACGGTTCTGCGATC	pWC89
Atx3B Y41A For	GATCGCAGAACCGTGCTTTTGCCTCAAC	pWC89
AatII Atx3B Rev	GCAGAGCGCCAAAGACGTC	pWC89

Table S2. Constructs used for this study.

Plasmid	Protein
pMC01	LarB1-His_6
pWC45	LarA-IEGR-MBP
pWC46	IEGR-MBP
pWC47	LarA F21AzF
pWC49	LarALP-GSSG-IEGR-MBP
pWC57	LarA Y28A-MBP
pWC58	LarACP-IEGR-MBP
pWC60	LarA -10aa-IEGR-MBP (N-terminal truncation)
pWC61	LarA -15aa-IEGR-MBP (N-terminal truncation)
pWC66	LarA Y32AzF-IEGR-MBP
pWC67	LarA W35AzF-IEGR-MBP
pWC68	LarA -5aa-IEGR-MBP (N-terminal truncation)
pWC71	LarA -9aa-IEGR-MBP (N-terminal truncation)
pWC72	LarA Y9AzF-IEGR-MBP
pWC73	LarB1 Y28AzF-His_6
pWC74	LarB1 Y29AzF-His_6
pWC75	LarB1 L31AzF-His_6
pWC76	LarB1 N32AzF-His_6
pWC77	LarB1 D27AzF-His_6
pWC78	LarB1 Q30AzF-His_6
pWC79	LarB1 Y16AzF-His_6
pWC80	LarB1 V19AzF-His_6
pMM65	Astexin-3 cluster
pWC89	Astexin-3 cluster with AtxB Y41A mutation

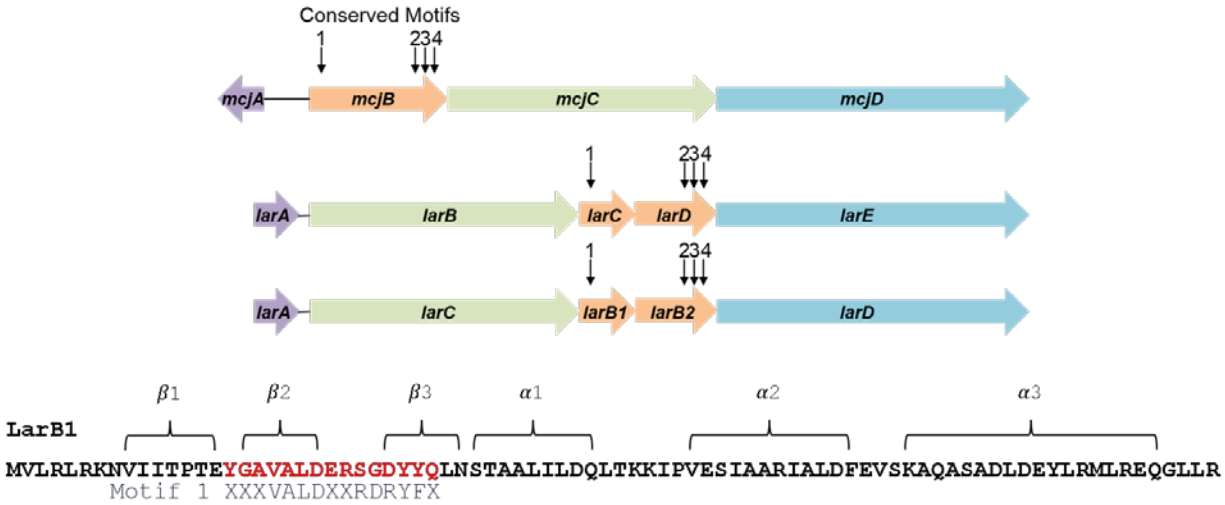


Figure S1. Lariatin biosynthetic gene cluster and LarB1 sequence. Top: Reproduction of Figure 1(b) with the original nomenclature of the lariatin gene cluster (middle) and the new nomenclature in accordance with lasso peptide standards (bottom). Bottom: Sequence of LarB1 with the secondary structure annotated above, see also Figure 5. The region corresponding to conserved motif 1 is highlighted in red and the motif sequence below it in blue.

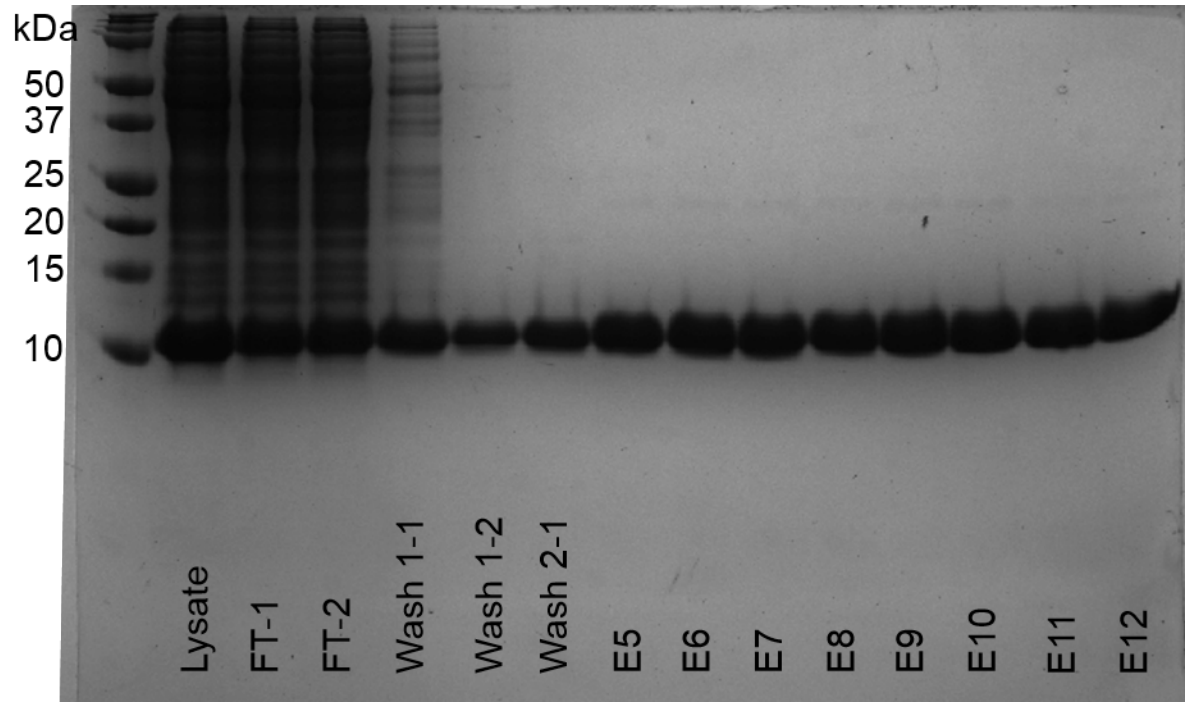


Figure S2. SDS-PAGE gel analysis of LarB1-His_6 purification. Elutions (lanes marked "E") show a pure band at the expected molecular weight of 10.5 kDa. FT: flowthrough

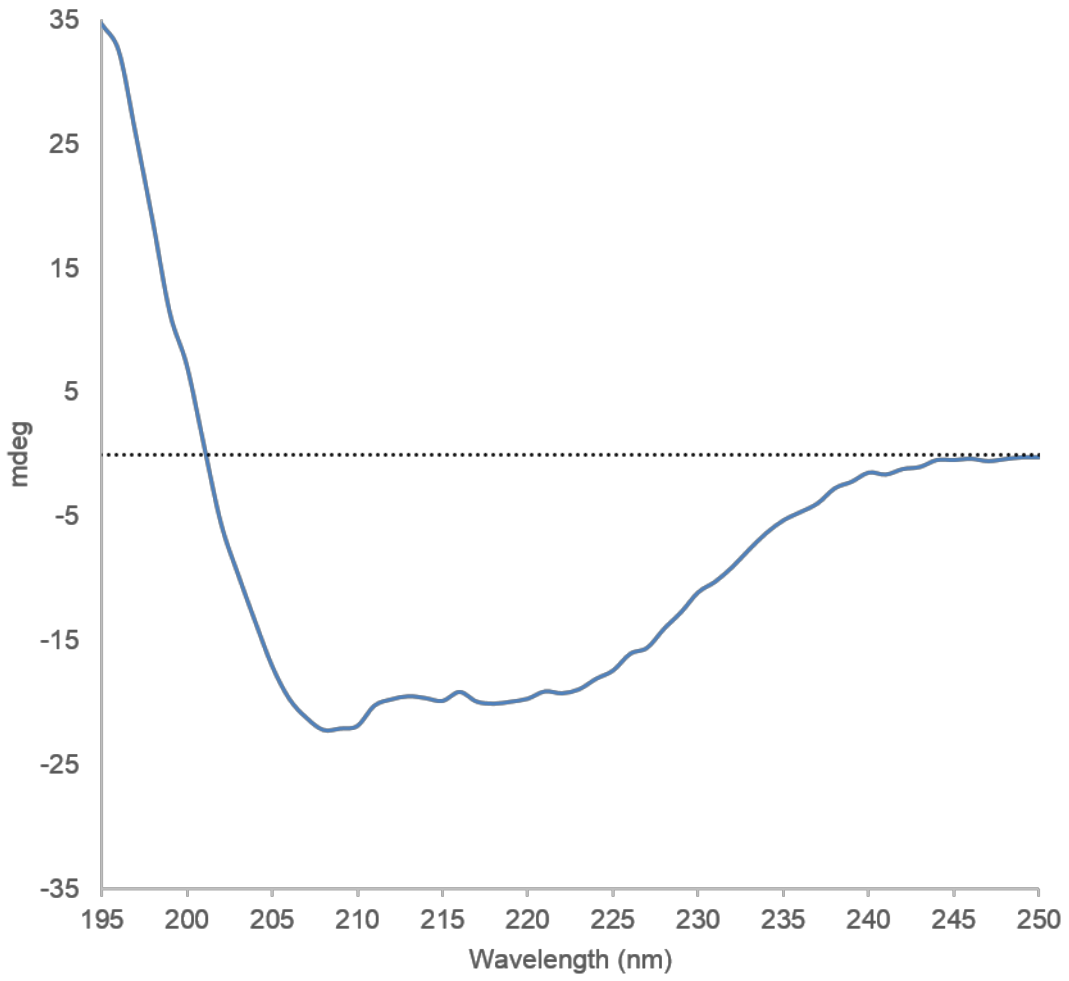


Figure S3. Circular dichroism spectrum of LarB1-His_6 showing the signature negative bands of an α-helix at 222 nm and 208 nm.

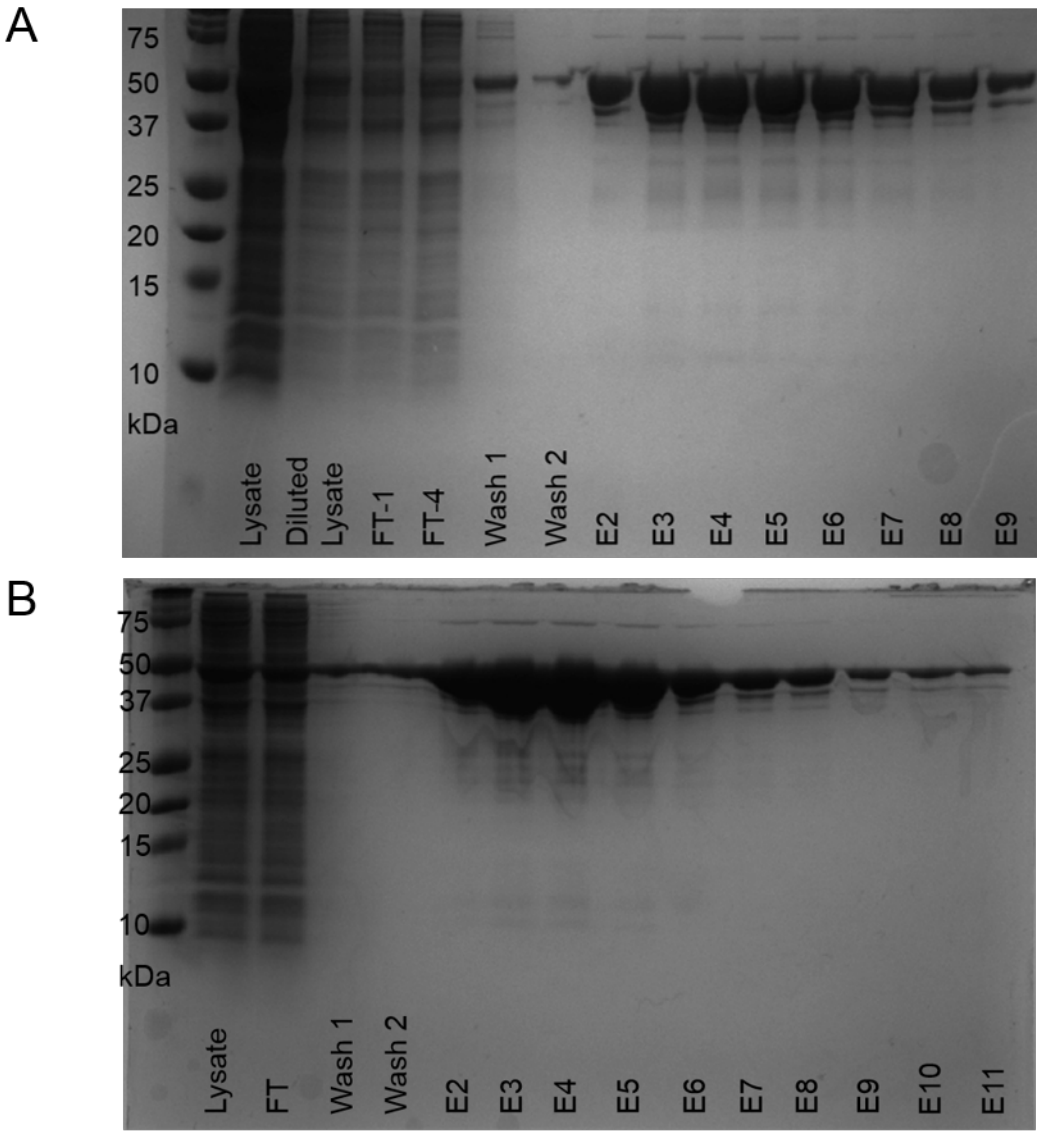
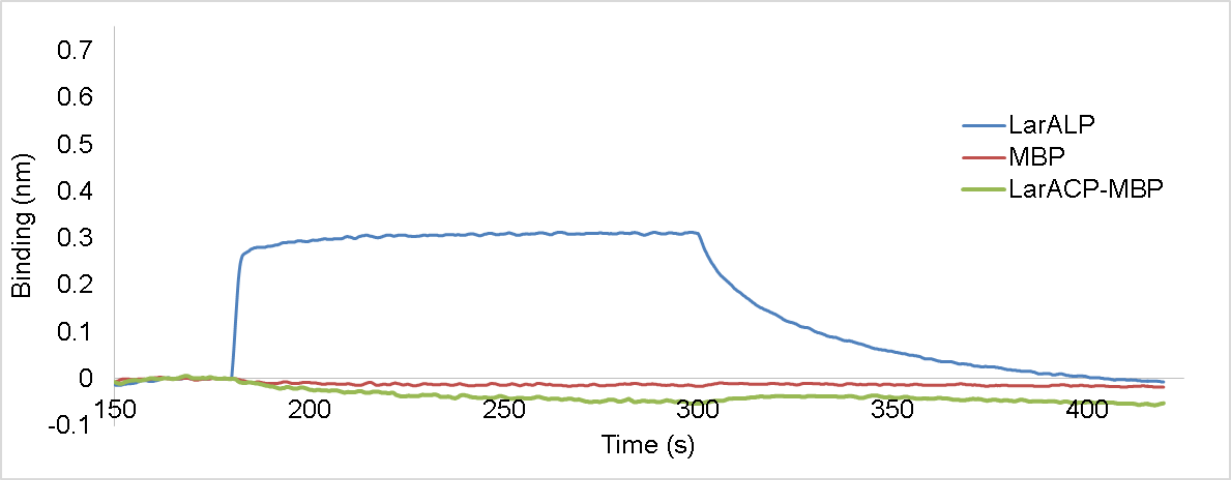
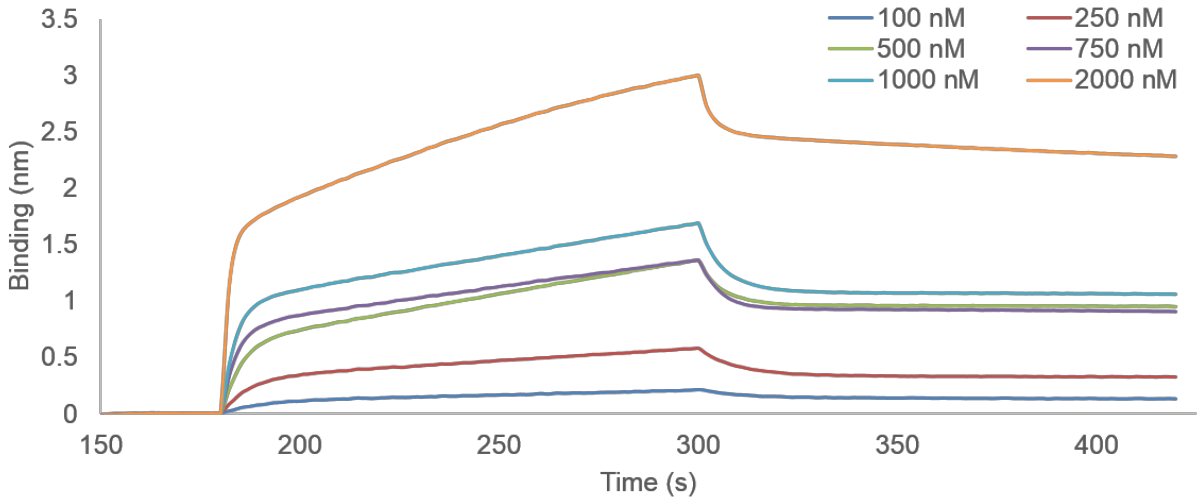


Figure S4. Purification of LarA-MBP fusion proteins. (a) SDS-PAGE analysis of LarA-IEGR-MBP purification. Elutions (lanes marked E) show a main band at the expected molecular weight of 47.8 kDa, with a minor contaminant that corresponds to endogenous MBP. (b) SDS-PAGE analysis of LarALP-IEGR-MBP purification. Elutions show a main band at the expected molecular weight of 45.6 kDa, with a minor contaminant that corresponds to endogenous MBP. FT: flowthrough



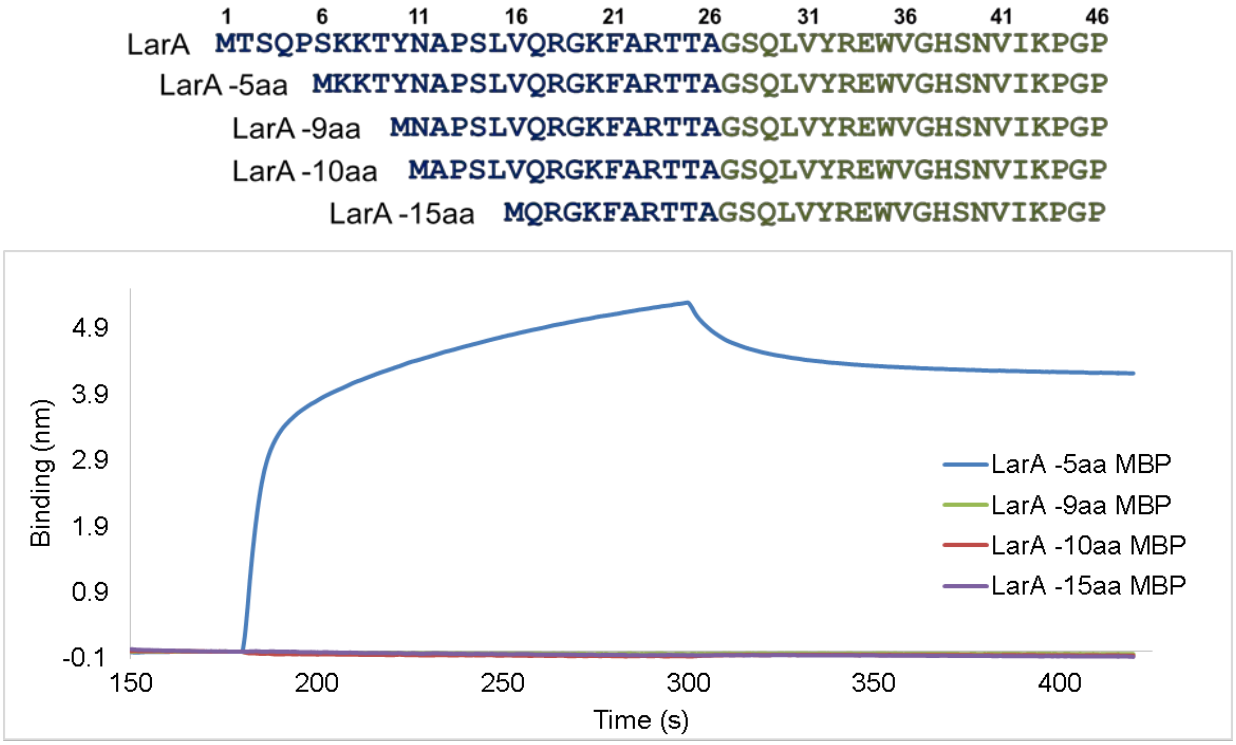
Sample	Conc. (nM)	K_d(M)	k_a($M^{-1}s^{-1}$)	k_d(s^{-1})
LarALP	4000	2.79×10^{-7}	$1.34 \pm 0.03 \times 10^{5}$	$3.72 \pm 0.02 \times 10^{-2}$
MBP	2000	N/A	N/A	N/A
LarACP-MBP	2000	N/A	N/A	N/A

Figure S5. BLI measurements between LarB1 and synthetic LarA leader peptide, LarB1 and MBP, and LarB1 and LarA core peptide-MBP (LarACP-MBP) with the corresponding kinetic fits. Note that the kinetic data obtained with the synthetic LarALP is consistent with that obtained with LarA-MBP (shown in Figure 2 in the main text). Also note that there is neither interaction of LarB1 with the negative control, MBP, nor with LarACP-MBP.



LarALP-MBP conc. (nM)	K_d(M)	k_a($M^{-1}s^{-1}$)	k_d(s^{-1})
100	6.35×10^{-8}	$1.51 \pm 0.22 \times 10^{5}$	$9.61 \pm 0.32 \times 10^{-3}$
250	2.44×10^{-7}	$6.73 \pm 1.42 \times 10^{4}$	$1.64 \pm 0.06 \times 10^{-2}$
500	9.69×10^{-8}	$6.33 \pm 0.77 \times 10^{4}$	$6.13 \pm 0.37 \times 10^{-3}$
750	1.27×10^{-7}	$8.12 \pm 0.82 \times 10^{4}$	$1.03 \pm 0.07 \times 10^{-2}$
1000	2.40×10^{-7}	$6.02 \pm 0.67 \times 10^{4}$	$1.45 \pm 0.09 \times 10^{-2}$
2000	6.77×10^{-8}	$4.88 \pm 0.30 \times 10^{4}$	$3.31 \pm 0.40 \times 10^{-3}$

Figure S6. BLI measurements between LarB1 and LarALP-MBP and the corresponding kinetic fits. Note that binding is observed but the curves do not reach a clean plateau during association especially at the higher concentrations.



Sample	Conc. (nM)	K_d(M)	k_a($M^{-1}s^{-1}$)	k_d(s^{-1})
LarA -5aa MBP	2000	7.81×10^{-8}	$5.2 \pm 0.17 \times 10^{4}$	$4.06 \pm 0.26 \times 10^{-3}$
LarA -9aa MBP	2000	N/A	N/A	N/A
LarA -10aa MBP	2000	N/A	N/A	N/A
LarA -15aa MBP	2000	N/A	N/A	N/A

Figure S7. BLI measurements between LarB1 and N-terminal truncations of LarA-MBP, and the corresponding kinetic fits. Truncation numbering does not include the N-terminal Met residue, which remains present in all constructs. LarA -5aa MBP binding does not reach a clean plateau but shows approximately wild-type binding affinity to LarB1. LarA -9aa MBP and further N-terminal truncations do not show binding.

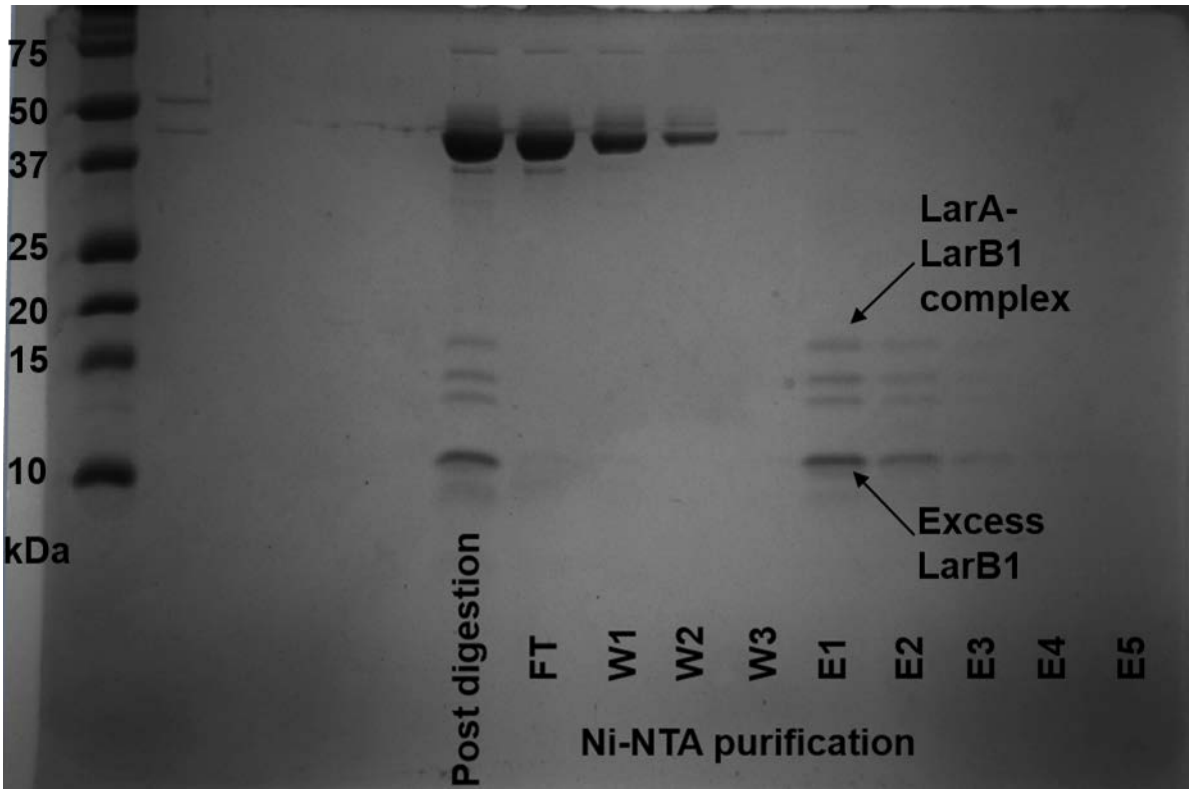


Figure S8. SDS-PAGE analysis of the purification of photocrosslinked LarA-LarB1 complexes. Post digestion: protein mixture after Factor Xa digestion. FT: flowthrough, W: washes, E: elutions

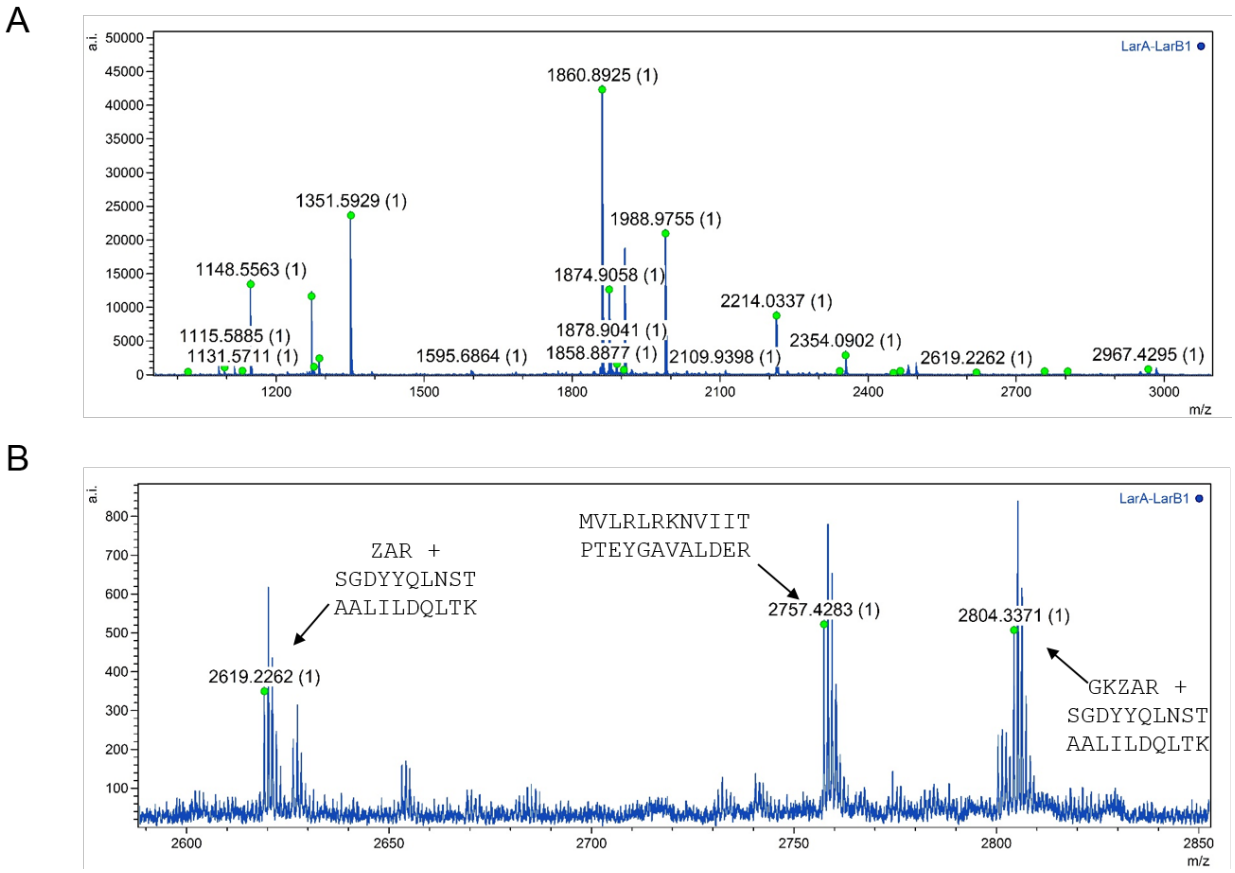
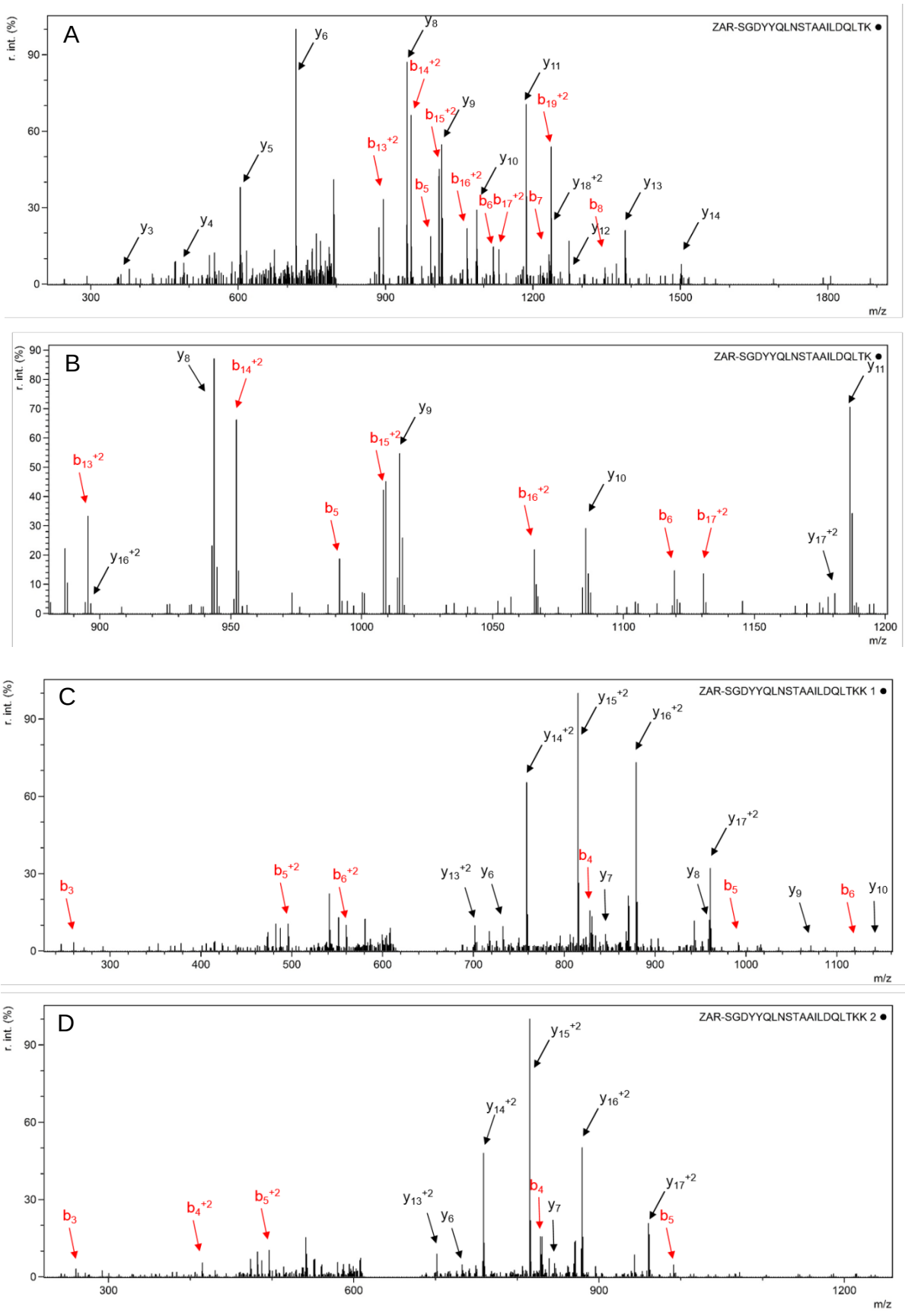


Figure S9. MALDI-TOF spectra of photocrosslinked LarA-LarB1 tryptic digestion. (a) Full spectrum; green dots indicate peaks that were identified as tryptic fragments of LarA-LarB1 adduct. (b) Zoom-in of the spectrum; the middle peak is a tryptic fragment from LarB1 while the other two peaks correspond to conjugates of LarA (ZAR/GKZAR) with amino acids 25-44 of LarB1. Z represents *p*-azidoPhe.


A
ZAR-SGDYYQLNSTAAILDQLTK
r. int. (%)
m/z
B
ZAR-SGDYYQLNSTAAILDQLTK
C
ZAR-SGDYYQLNSTAAILDQLTKK 1
D
ZAR-SGDYYQLNSTAAILDQLTKK 2

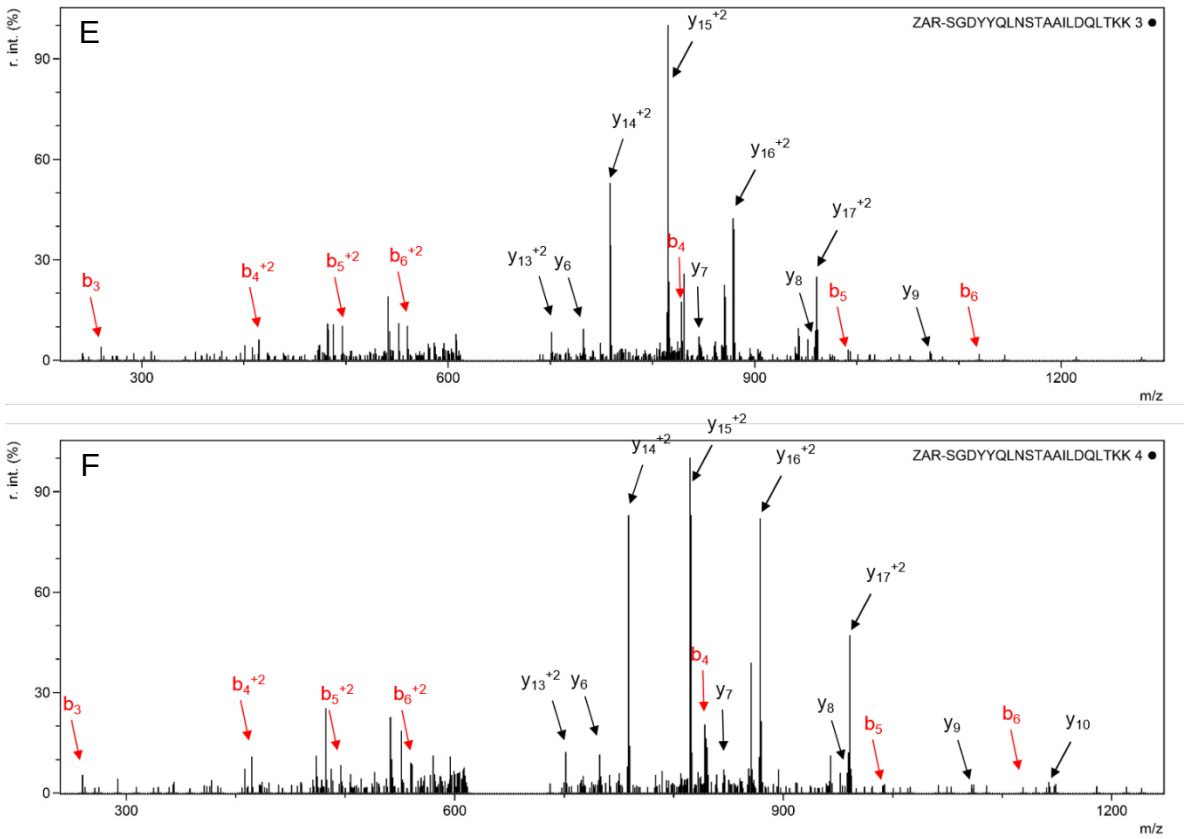


Figure S10. Sample LC-MS/MS spectra of LarA-LarB1 adduct tryptic peptide. Errors of the parent ion masses were less than 0.004 Da. (a) Spectrum of LarA (ZAR) conjugated to LarB1 amino acids 25-44. (b) Zoomed-in spectrum of (a) showing some lower intensity peaks. (c)-(f) Spectra of LarA (ZAR) conjugated to LarB1 amino acids 25-45; the retention time for the parent ion were 2187, 2201, 2206, and 2214 seconds respectively.

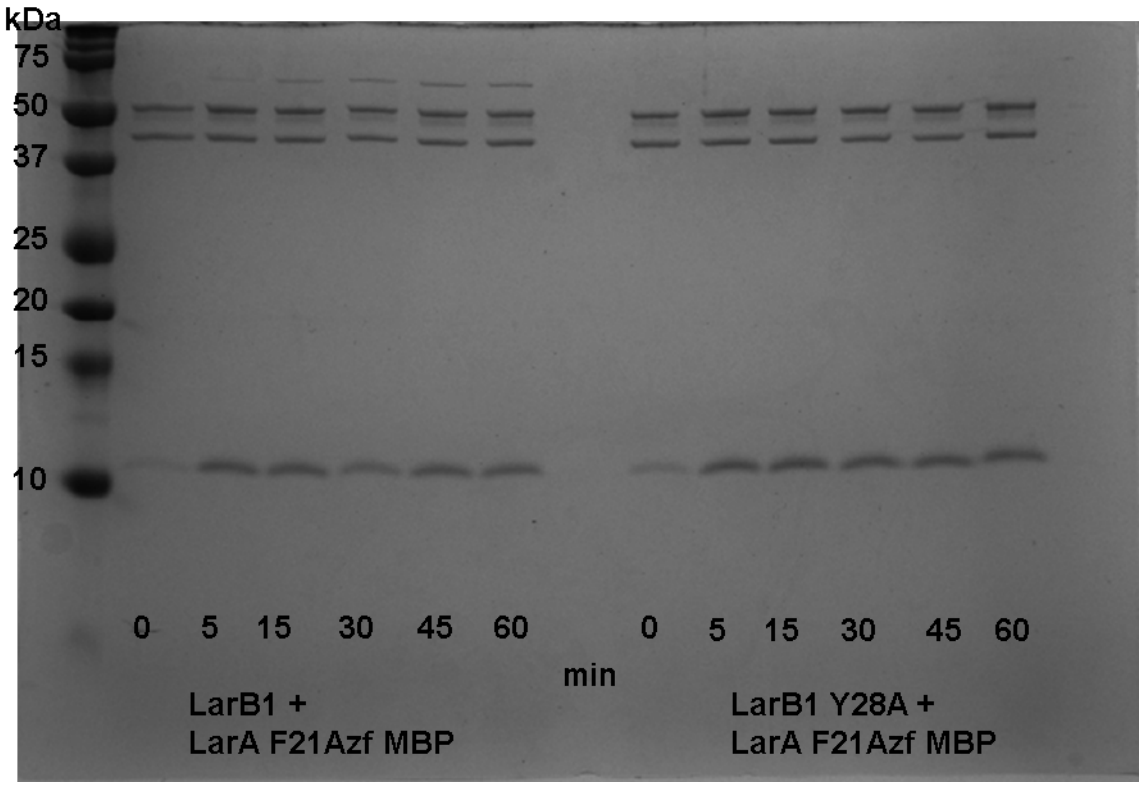
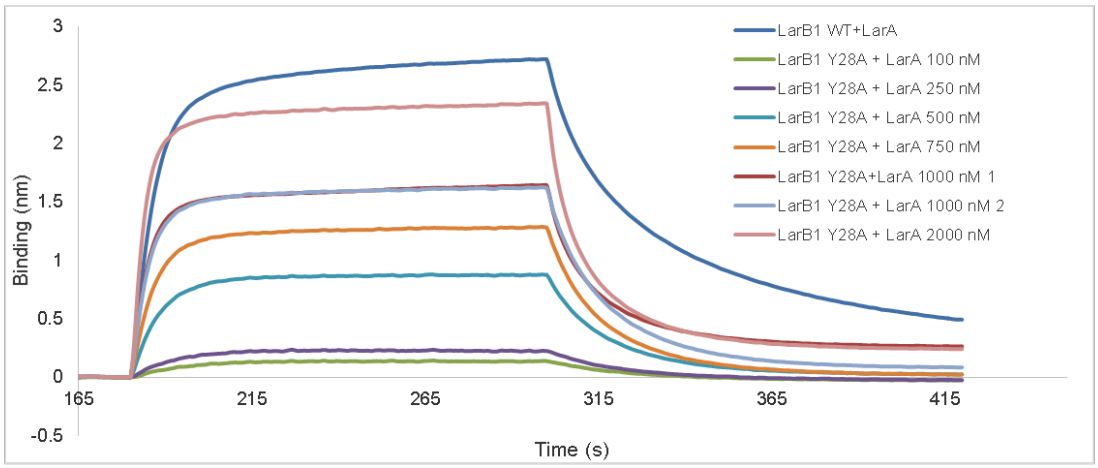


Figure S11. SDS-PAGE analysis of the photocrosslinking reaction between LarA F21AzF-MBP and wild-type LarB1 or LarB1 Y28A. LarA F21AzF-MBP crosslinked to LarB1 but did not crosslink to LarB1 Y28A.



Ligand	LarA-MBP conc. (nM)	K_d(M)	k_a($M^{-1}s^{-1}$)	k_d(s^{-1})
LarB1 wildtype	1000	2.85×10^{-7}	$9.68 \pm 0.09 \times 10^{4}$	$2.76 \pm 0.02 \times 10^{-2}$
LarB1 Y28A	100	9.85×10^{-7}	$6.34 \pm 2.43 \times 10^{4}$	$6.25 \pm 0.14 \times 10^{-2}$
LarB1 Y28A	250	5.81×10^{-7}	$1.00 \pm 0.06 \times 10^{5}$	$5.81 \pm 0.06 \times 10^{-2}$
LarB1 Y28A	500	4.48×10^{-7}	$1.21 \pm 0.01 \times 10^{5}$	$5.42 \pm 0.01 \times 10^{-2}$
LarB1 Y28A	750	5.8×10^{-7}	$1.02 \pm 0.01 \times 10^{5}$	$5.89 \pm 0.02 \times 10^{-2}$
LarB1 Y28A	1000 (run 1)	4.87×10^{-7}	$1.16 \pm 0.01 \times 10^{5}$	$5.67 \pm 0.03 \times 10^{-2}$
LarB1 Y28A	1000 (run 2)	4.94×10^{-7}	$1.10 \pm 0.01 \times 10^{5}$	$5.45 \pm 0.02 \times 10^{-2}$
LarB1 Y28A	2000	1.11×10^{-6}	$6.72 \pm 0.09 \times 10^{4}$	$7.48 \pm 0.05 \times 10^{-2}$

Figure S12. BLI measurements between LarB1 Y28A and LarA-MBP and the corresponding kinetic fits. A measurement between LarB1 wild-type and LarA-MBP was done as a control.

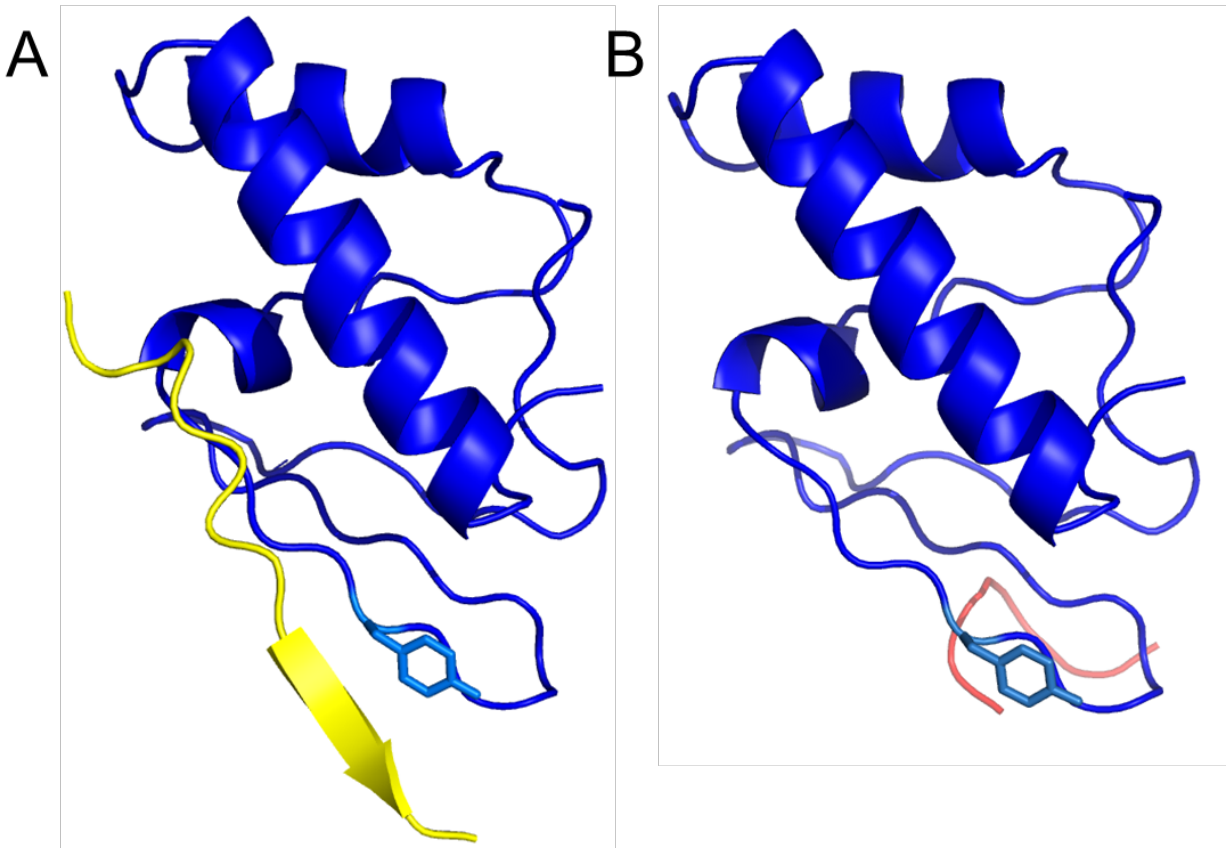


Figure S13. (a) LarB1 (blue) aligned to NisB (hidden) and shown with residues -21 to -9 of the NisA leader peptide (yellow). The Tyr-28 residue identified in crosslinking experiments is highlighted in light blue. NisB structures drawn from PDB file 4WD9. (b) LarB1 homology model aligned to MccB (hidden) showing MccA precursor peptide (salmon). The sidechain of Tyr-28 identified in the crosslinking experiments is shown in light blue. MccB drawn from PDB file 3H9G

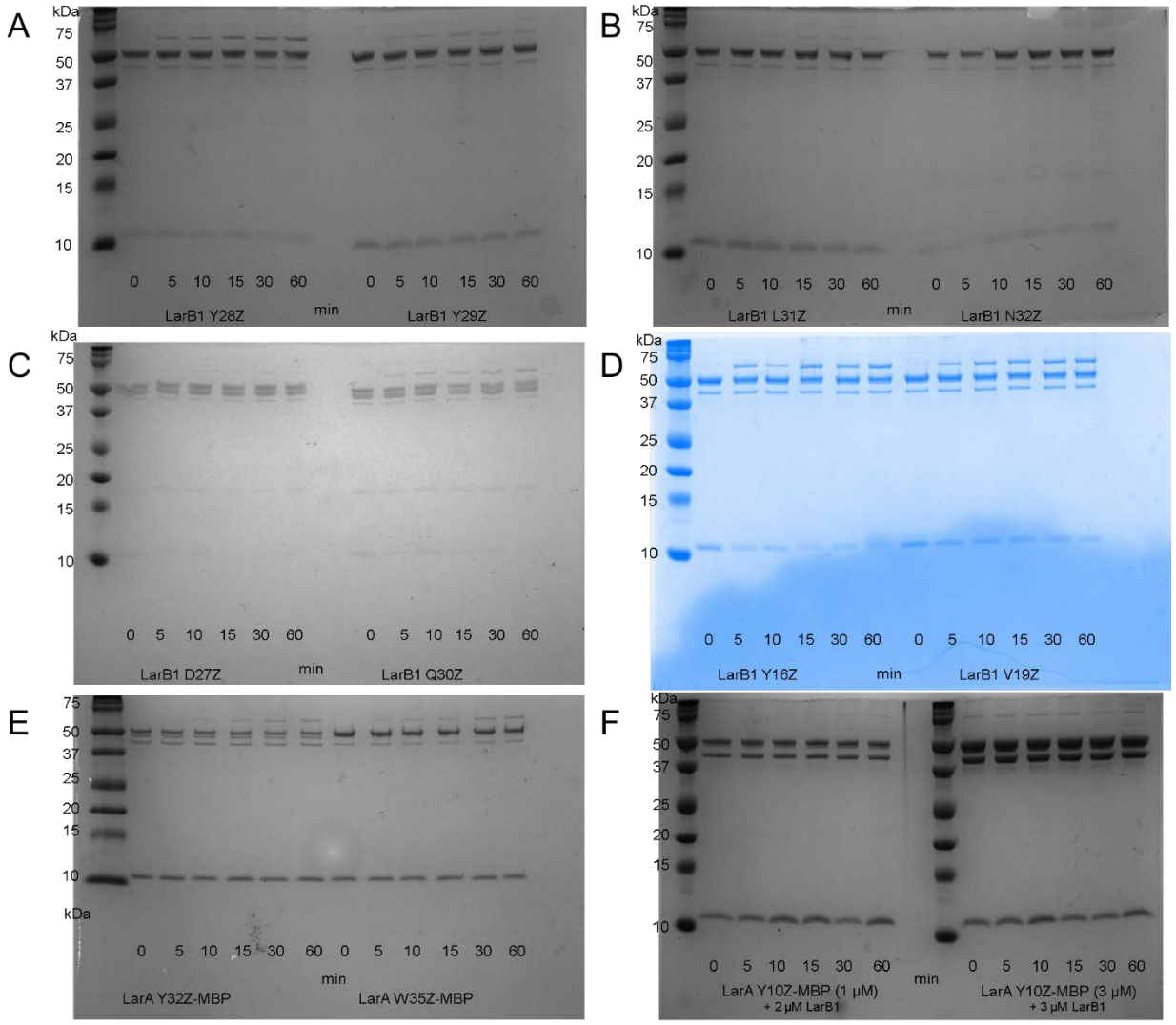
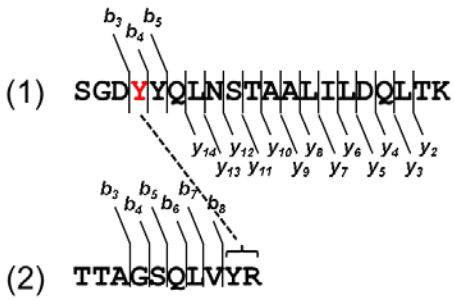
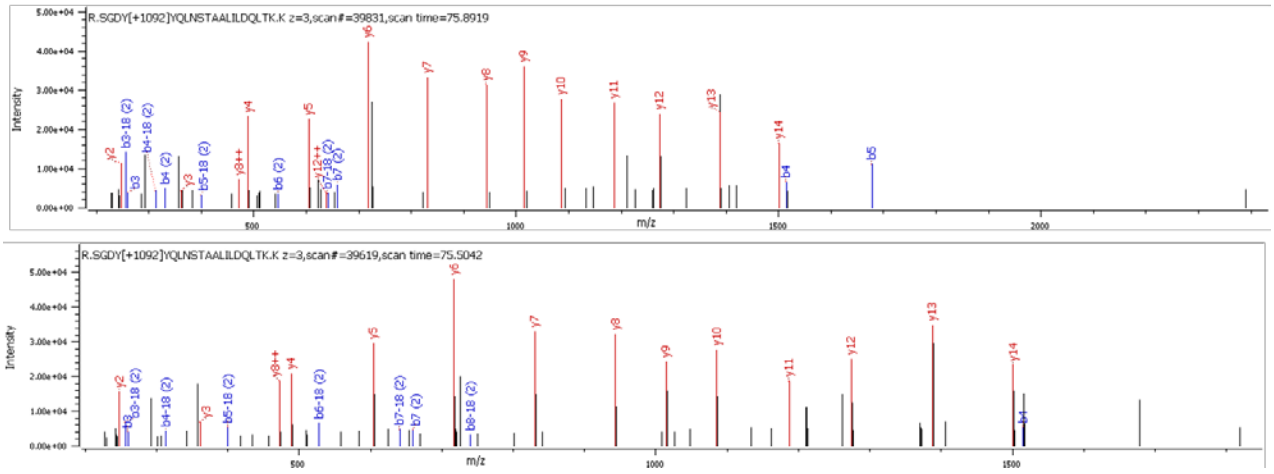


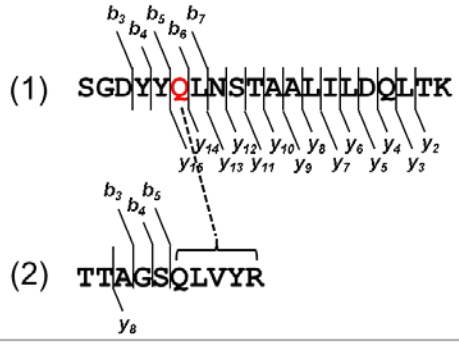
Figure S14. Photocrosslinking of various AzF constructs of LarB1 and LarA. 2 μM of LarB1 was incubated with 2 μM of LarA unless otherwise indicated. SDS-PAGE gels showing that the band for LarB1 and LarA adducts (~58 kDa) varies in strength according to the AzF construct but grows in strength with UV exposure time for those that crosslinked. (a) LarB1 Y28Z with LarA-MBP (left); LarB1 Y29Z with LarA-MBP (right); (b) LarB1 L31Z with LarA-MBP (left); LarB1 N32Z with LarA-MBP (right); (c) LarB1 D27Z with LarA-MBP (left); LarB1 Q30Z with LarA-MBP (right); (d) LarB1 Y16Z with LarA-MBP (left); LarB1 V19Z with LarA-MBP (right); (e) LarB1 with LarA Y32Z-MBP (left); LarB1 with LarA W35Z-MBP (right); (f) LarB1 with LarA Y10Z-MBP at two different concentrations.

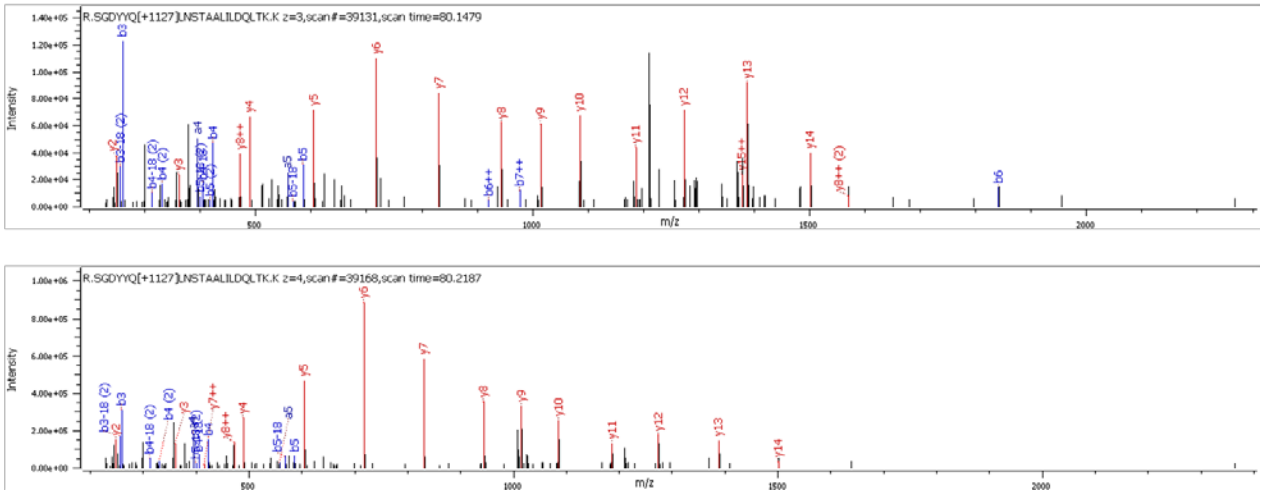
A


(1) SGDYYQLNSTAALILDQLTK
(2) TTAGSQLVYR

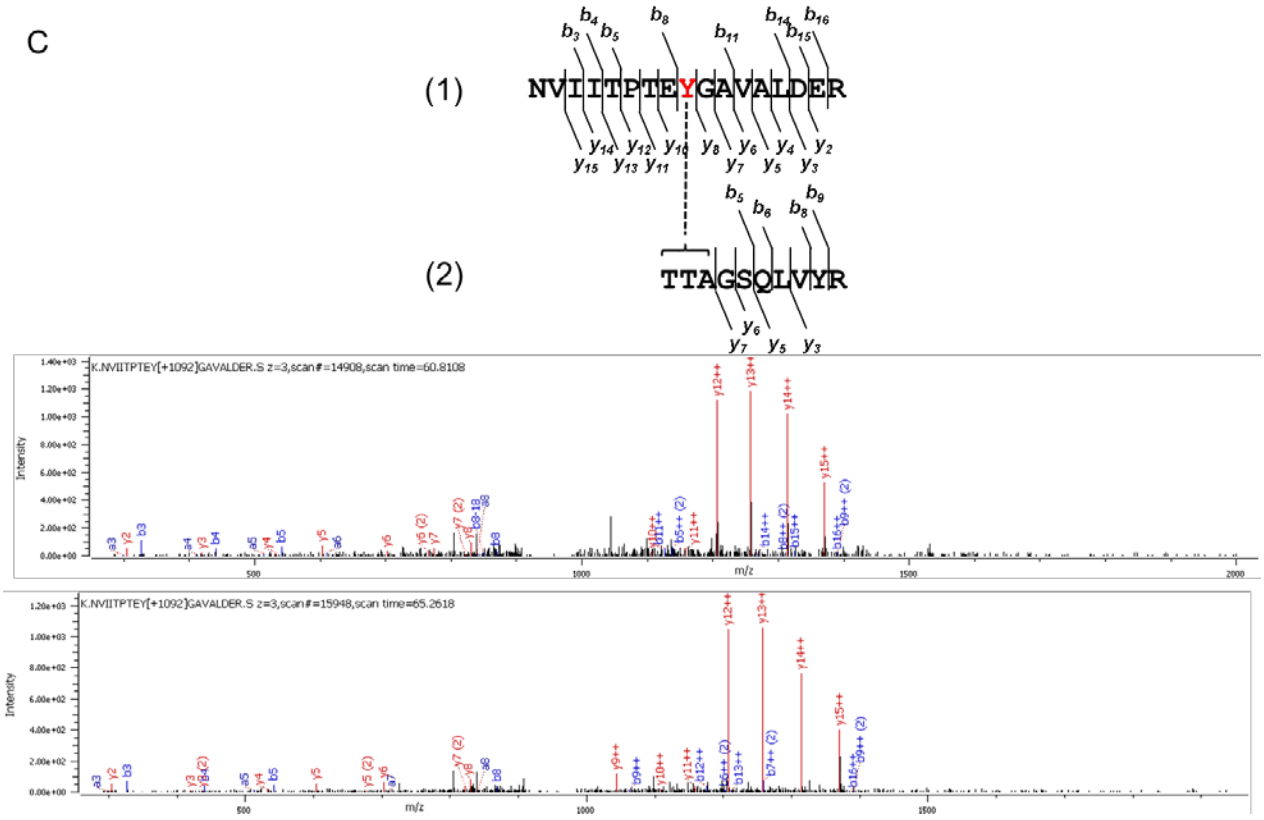

R.SGDY[+1092]YQLNSTAALILDQLTK.K z=3,scan#=39831,scan time=75.8919
R.SGDY[+1092]YQLNSTAALILDQLTK.K z=3,scan#=39619,scan time=75.5042
Intensity
m/z

B

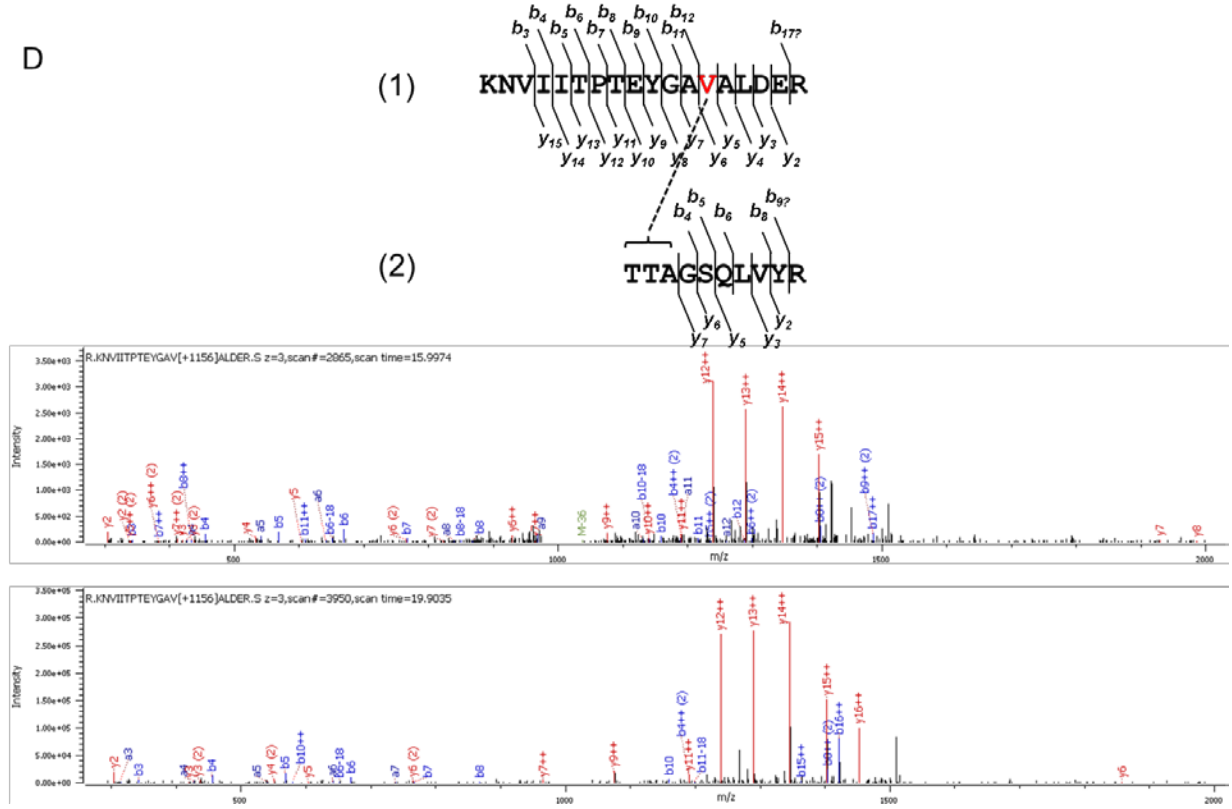

(1) SGDYYQLNSTAALILDQLTK
(2) TTAGSQLVYR


R.SGDYYQ[+1127]LNSTAALILDQLTK.K z=3,scan#=39131,scan time=80.1479
R.SGDYYQ[+1127]LNSTAALILDQLTK.K z=4,scan#=39168,scan time=80.2187
Intensity
m/z

C


(1)
NVIITPTEYGAVALDER
(2)
TTAGSQLVYR
K.NVIITPTEY[+1092]GAVALDER.S z=3,scan#=14908,scan time=60.8108
K.NVIITPTEY[+1092]GAVALDER.S z=3,scan#=15948,scan time=65.2618
Intensity
m/z

D


(1)
KNVIITPTEYGAVALDER
(2)
TTAGSQLVYR
R.KNVIITPTEYGAV[+1156]ALDER.S z=3,scan#=2865,scan time=15.9974
R.KNVIITPTEYGAV[+1156]ALDER.S z=3,scan#=3950,scan time=19.9035
Intensity
m/z

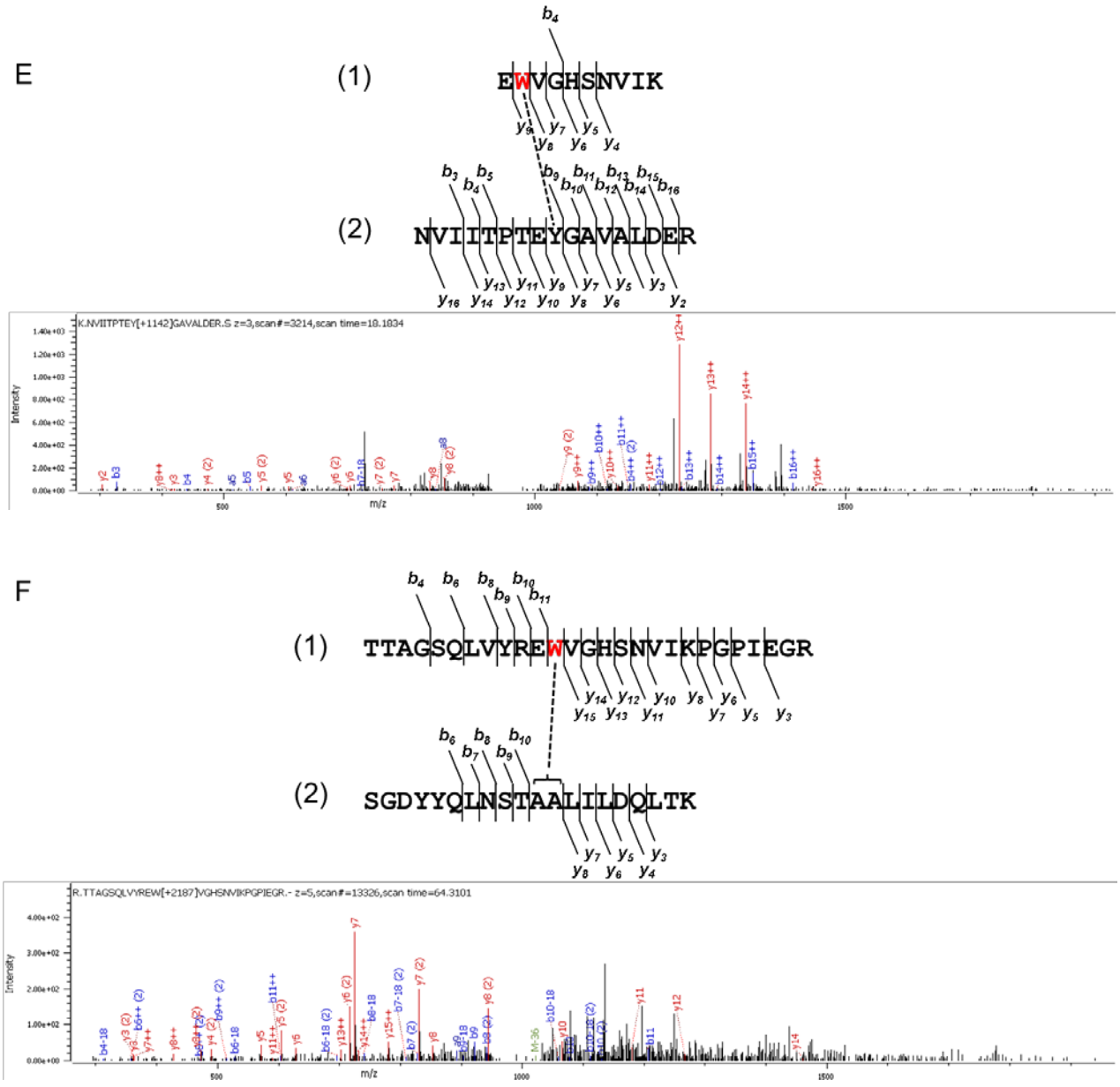


Figure S15. LC-MS/MS spectra of LarA-LarB1 adduct tryptic peptides. Spectra and corresponding diagram of (a) LarB1 Y28Z tryptic fragment (1) conjugated to the C-terminal end of LarA amino acids 24-33 (2), (b) LarB1 Q30Z tryptic fragment (1) conjugated to the C-terminal half of LarA amino acids 24-33 (2), (c) LarB1 Y16Z tryptic fragment (1) conjugated to the N-terminal end of LarA amino acids 24-33 (2), (d) LarB1 V19Z trypic fragment (1) conjugated to the N-terminal end of LarA amino acids 24-33 (2) (Note: b_{17} (1) and b_9 (2) have the same expected m/z) , (e) LarA W35Z tryptic fragment (1) conjugated to LarB1 amino acids 8-24 at Y16 (2) and (f) LarA W35Z tryptic fragment (1) also conjugated to LarB1 amino acids 25-44.

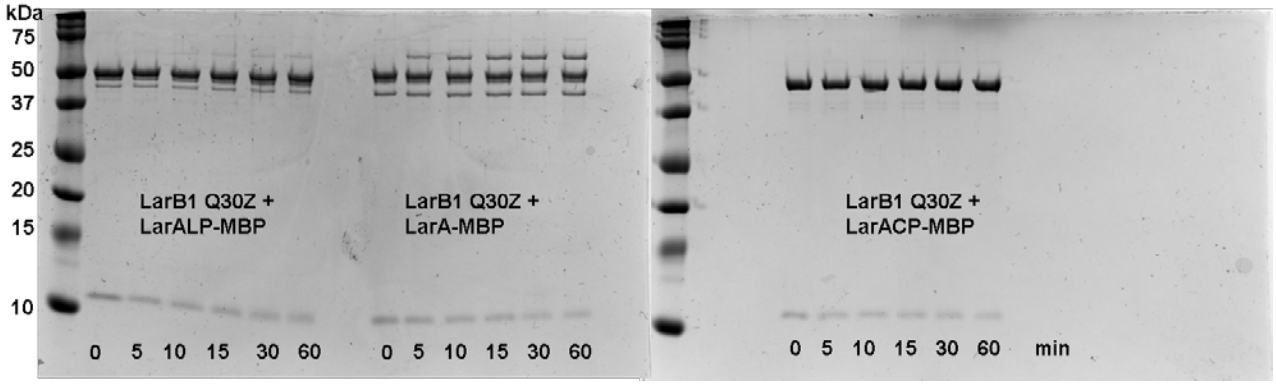
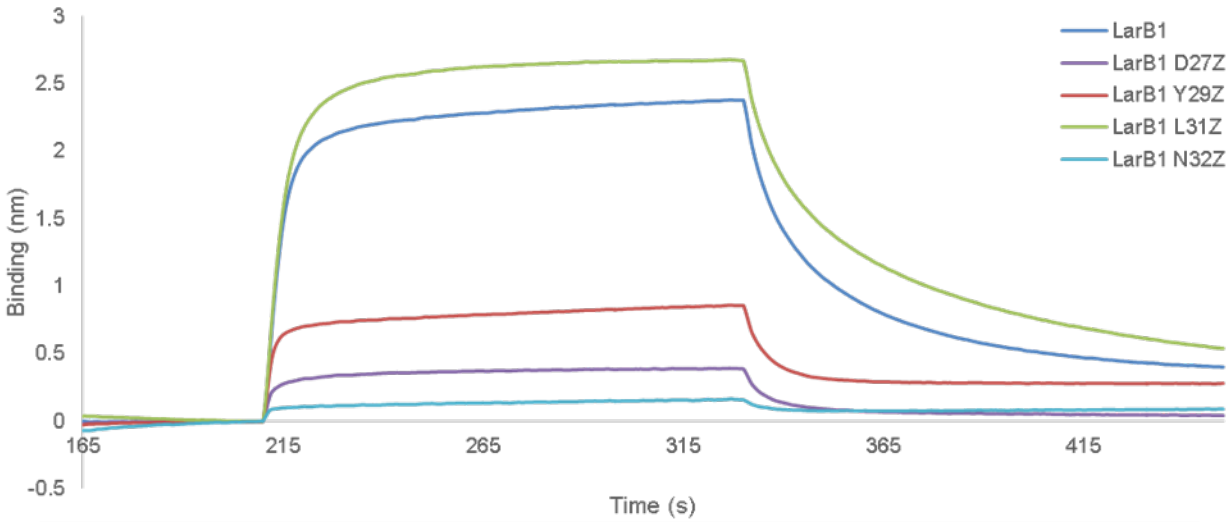


Figure S16. SDS-PAGE analysis of the photocrosslinking reaction between LarB1 Q30Z and full length LarA-MBP, LarA leader peptide-MBP, or LarA core peptide-MBP fusion proteins. LarB1 Q30Z crosslinked very weakly or not at all to LarALP-MBP and LarACP-MBP respectively, but crosslinked strongly to full length LarA-MBP.



Construct	K_d(M)	k_a($M^{-1}s^{-1}$)	k_d(s^{-1})
LarB1 WT	5.91×10^{-7}	$6.98 \pm 0.10 \times 10^{4}$	$4.13 \pm 0.03 \times 10^{-2}$
LarB1 D27Z	1.72×10^{-6}	$6.4 \pm 0.26 \times 10^{4}$	$1.1 \pm 0.02 \times 10^{-1}$
LarB1 Y29Z	9.24×10^{-7}	$9.74 \pm 0.34 \times 10^{4}$	$9.00 \pm 0.19 \times 10^{-2}$
LarB1 L31Z	4.26×10^{-7}	$6.92 \pm 0.08 \times 10^{4}$	$2.95 \pm 0.02 \times 10^{-2}$
LarB1 N32Z	1.41×10^{-6}	$4.35 \pm 1.12 \times 10^{4}$	$6.15 \pm 0.47 \times 10^{-2}$

Figure S17. BLI measurements between LarB1 AzF variants and 2 µM LarA-MBP and the corresponding kinetic fits. Note that for LarB1 D27Z, LarB1 Y29Z, and LarB1 N32Z, each calculated K_d from the measured on and off rates is close to the LarB1 WT value but the BLI signal plateau is significantly lower than expected based on the K_d value in a 1:1 binding model.

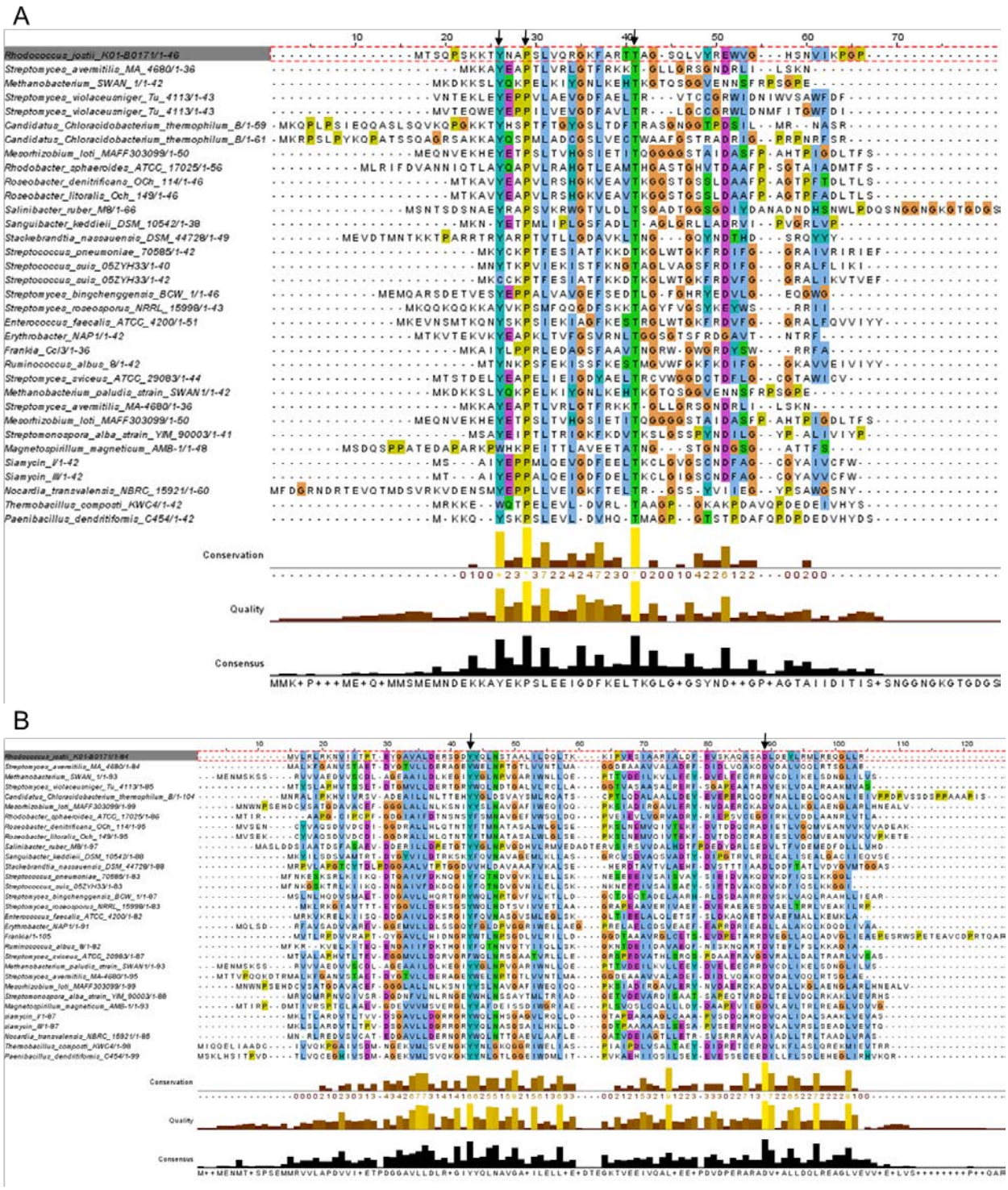
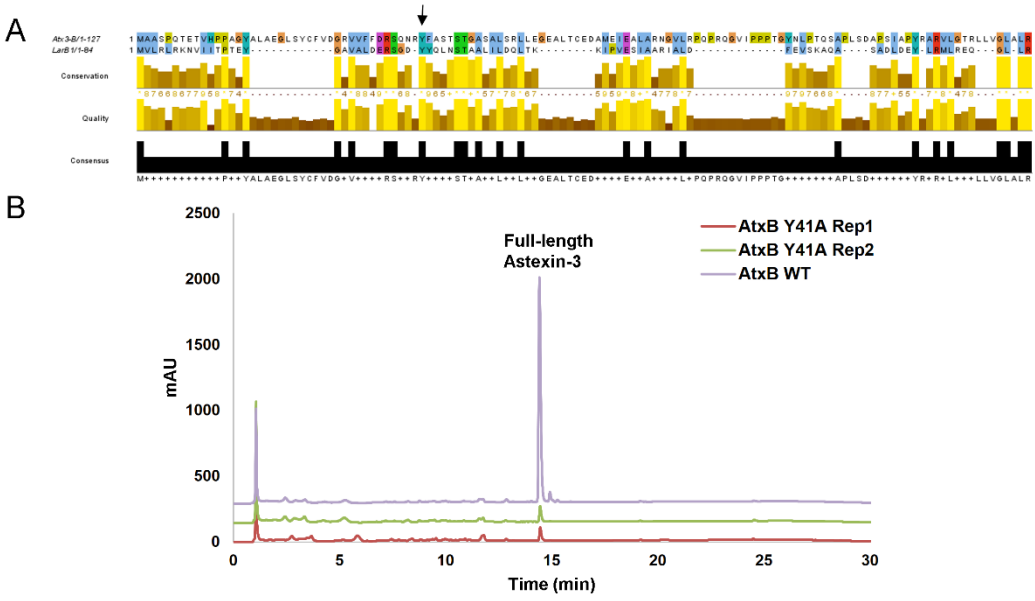


Figure S18. ClustalWS alignment of lasso peptide precursors and B1 proteins in lasso peptide clusters with split-B genes. (a) Alignment of lasso peptide precursors. Note that in addition to the highly conserved penultimate T, there are highly conserved P and Y/W residues in the leader sequence as indicated by the arrows. (b) Alignment of B1 proteins. Note the highly conserved Y and D residue.



Construct	Area (mAU*min)	% of WT Production
AtxB WT	11409	100
AtxB Y41A Rep1	604	5.3
AtxB Y41A Rep2	726	6.4

Figure S19. Heterologous astexin-3 production from the wild-type astexin-3 cluster and from the astexin-3 cluster with an AtxB Y41A substitution. (a) Alignment of the first 127 residues of astexin-3 B protein (AtxB) and LarB1. The conserved LarB1 Y28 and AtxB Y41 are indicated with an arrow. (b) Top: HPLC traces of the cell extracts after 20 hours of expression. Two biological replicates were done for the AtxB Y41A construct. Bottom: corresponding areas of the full-length astexin-3 peak. The AtxB Y41A substitution causes approximately a 15- to 19-fold drop in lasso peptide production.