Biochemistry

pubs.acs.org/biochemistry Communication

A Leader-Guided Substrate Tolerant RiPP Brominase Allows Suzuki−Miyaura Cross-Coupling Reactions for Peptides and Proteins

Nguyet A. Nguyen and Vinayak Agarwal*

Cite This: https://doi.org/10.1021/acs.biochem.3c00222 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Bioorthogonal derivatization of peptides and proteins enables investigations into their biological function and allows for exploitation of their therapeutic potential, among other varied deliverables. Herein, we describe a marine halogenating enzyme-assisted bioconjugation strategy in which an N-terminal leader peptide guides bromination of a C-terminal Trp residue in genetically encoded peptides and proteins, setting up further Trp arylation by Suzuki−Miyaura reactions. The bromination and subsequent cross-coupling reactions are residue-specific and regiospecific for the indole-6 position, occur under mild aqueous conditions, and do not require any modification of other Trp residues in the substrate peptide and/or protein. Workflows described herein demonstrate the applicability of halogenating enzymes in bioorthogonal conjugation chemistry.

Suzuki−Miyaura cross-coupling is a universal palladium-assisted carbon−carbon bond-forming reaction typically involving aryl halides and organoboron substrates that enables the bioorthogonal derivatization of peptides and proteins.[1−3] With an inventory of organoboron reaction partners already available,[4] the key consideration is the preparation of the peptidic aryl halides. Reported strategies for introducing halogens into peptides and proteins include expanding the genetic code to incorporate amino acids with halogenated side chains,[5,6] post-translational chemical modifications such as the alkylation of cysteine side chains to generate aryl halide thioethers,[7,8] and chemical synthesis of the peptidic substrates with preinstalled halogen handles.[9,10] For short synthetic peptides,[11] enzymatic halogenation of indolic and phenolic rings sets up subsequent derivatization via Suzuki−Miyaura coupling.[12,13] However, enzymatic halogenation of genetically encoded peptides and proteins has been out of reach as the repertoire of enzymes that halogenate peptidic substrates is limited.

Biocatalytic halogenation is rooted in natural product biosynthetic enzymology.[14] The flavin-dependent halogenase MibH that is involved in the biosynthesis of the ribosomally synthesized and post-translationally modified lanthipeptide antibiotic NAI-107 is a regiospecific tryptophan side chain chlorinase.[15] MibH is substrate selective; MibH chlorinated the Trp side chain indole only when all other post-translational modifications had been installed upon the NAI-107 precursor peptide, MibA, including the proteolytic removal of the modified core region from the MibA leader (Figure 1A).

We recently described a peptide brominase, SrpI, encoded in the microbiomes of marine sponges.[16] The likely physiological substrate for the SrpI was the SrpE peptide in which the three Cys residues in the SrpE core, -LCCCW, were modified into thiazoline rings by the YcaO cyclodehydratase SrpC (Figure 1B). Preparation of the post-translationally modified SrpI substrate, and its derivatives, was hampered by the poor activity and strict substrate selectivity of the SrpC. Recombinant SrpC was not amenable to purification, and co-expression of *srpC* and *srpE* genes in *Escherichia coli* yielded a mixture of partially modified SrpE products. This observation in turn precluded the *in vitro* activity reconstitution of SrpI, evaluation of its the substrate scope, querying whether the activity of SrpI was dependent on the SrpE leader, and realizing the potential of SrpI as a general-purpose biocatalyst for peptide and protein halogenation.

To address the challenge of the preparation of the physiological substrate for SrpI, we turned to the substrate promiscuous YcaO cyclodehydratase/azoline oxidase pair MprC/MprD that we had described for installing azol(in)e heterocycles into 10 different MprE substrate peptides (Figure 1B and Figure S1).[17] MprE and SrpE are proteusin peptides, characterized by long leader sequences that are similar to those of nitrile hydratases.[18] While the MprC/MprD demonstrated robust activities *in vivo* and *in vitro*, they were selective for the MprE leader sequences; in contrast, SrpI was tolerant to other proteusin leaders.[16,17] In light of these observations, we appended the SrpE -LCCCW core to the $MprE_X$ leader (a consensus leader sequence built from the 10 different MprE leader sequences[17]), thus creating a chimeric $MprE_X$−LCCCW substrate (Supplementary Note). Upon co-expression of the gene encoding this chimeric substrate with *mprC*/*mprD* in *E. coli*, we obtained the purified $MprE_X$−LCCCW peptide in which all three Cys residues in the core were neatly converted into thiazol(in)e heterocycles (Figure 2A and Figure S2). Using the thusly prepared substrate, the brominating activity of purified flavin-dependent brominase SrpI was successfully

Received: April 27, 2023

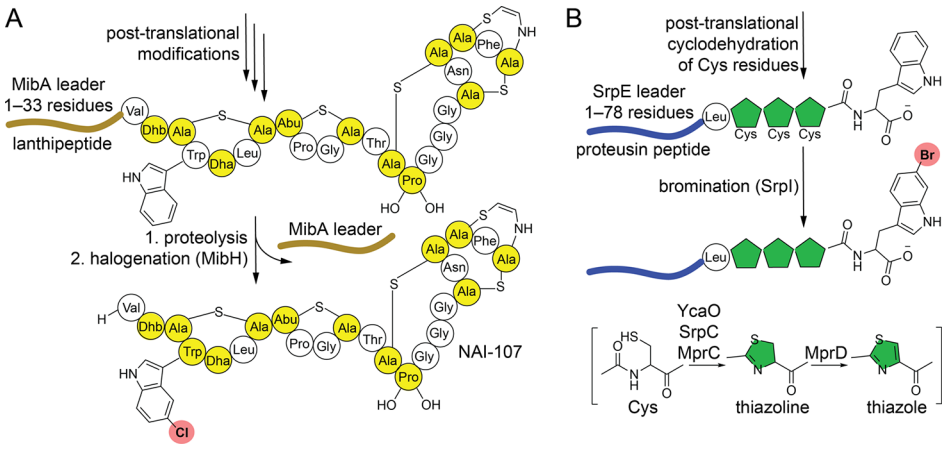


Figure 1. (A) Biosynthesis of lanthipeptide NAI-107. The 33-residue N-terminal leader sequence of the MibA substrate peptide guides post-translational modification of the C-terminal core. Proteolytic removal of the MibA leader is followed by halogenation by MibH. Post-translationally modified residues are colored yellow. Abbreviations: Dbh, dehydrobutyrine; Dha, dehydroalanine; Abu, aminobutyric acid. (B) The YcaO cyclodehydratase SprC catalyzes the cyclodehydration of three Cys residues to thiazoline in the SrpE substrate peptide. Bromination of the C-terminal Trp side chain is affected by SrpI.

reconstituted *in vitro* when paired with the flavin reductase RebF and the phosphite dehydrogenase PTDH (Figure 2B and Figure S3).[19,20] To mitigate potential cross reactivity with hydrogen peroxide that is produced when flavin cofactor redox cycling is uncoupled from halide oxidation,[21] catalase was included in all *in vitro* reactions.

Establishing an *in vitro* assay allowed us to evaluate several aspects of the SrpI activity. First, we queried the leader peptide dependence for SrpI. When the modified LCCCW core was excised from the $MprE_X$ core using the LahT150 peptidase[22] and provided by itself as a substrate to SrpI without the $MprE_X$ leader, no bromination was observed (Figure 2C). When the modified LCCCW core was provided *in trans* with the dissociated $MprE_X$ leader, as before, no bromination of the core was observed (Figure 2D). SrpI did not brominate free tryptophan either (Figure S4). While bromination of indole was observed, bromination regiospecificity changed from SrpI being an indole-6 brominase for RiPP substrates to halogenating position 3 of free indole (Figure S5).[16] Collectively, these data allow us to posit that the presence of the proteusin leader is an obligate requirement for the SrpI brominase. Both leaders, SrpE and $MprE_X$, support SrpI activity. This inference is in contrast to the RiPP chlorinase MibH that catalyzes tryptophan chlorination only after the MibA core has been removed from the MibA leader.[15]

We next evaluated the selectivity of SrpI for different core sequences. Conservative modifications in which an alanine residue was added after and before the terminal Trp residue in the LCCCW core were tolerated by SrpI, yielding brominated products in each case (Figure 2E,F and Figures S7 and S8). A tripeptide extension before the terminal Trp residue was also tolerated (Figure 2G and Figure S9). However, a tripeptide extension after the Trp residue ($MprE_X$–LCCCWAAA) was not processed by SrpI (Figure S10). Genes encoding all of the substrates mentioned above were co-expressed with *mprC*/*mprD* converting the Cys residues to thiazol(in)es. The consecutive thiazol(in)e sequence could be disrupted, and the $MprE_X$–GLCACCW substrate was brominated (Figure 2H and Figure S11). Crucially, moving the Trp residue away from the C-terminus, substrates $MprE_X$–GLCWCCC and $MprE_X$–GLCAWCC, did not result in bromination by SrpI (Figures S12 and S13).

Data presented above identify two requirements for SrpI activity: the presence of a proteusin leader and the Trp residue being present at the C-terminus of the core. To test whether meeting these requirements allows for extension of the substrate scope of SrpI, we turned our attention to tumor-homing (TH) hexa- and heptapeptides. The TH peptides can deliver payloads specifically to tumor cells, making them attractive vehicles for the delivery of therapeutic payloads.[23] We employed two TH peptides, here termed TH1 and TH2, LTVPLW and VLTVQPW, respectively, that possess terminal Trp residues.[24] In contrast to the physiological pentapeptide substrate SrpE, the TH1 and TH2 peptides are hexa- and heptapeptides, respectively. Note that while SrpI can modify octapeptides, as well [substrate core LCCCAAAW (Figure 2G)], the observed substrate turnover was lower. In contrast to the physiological substrate, the TH peptides bear no azol(in)e heterocycles, though the Pro residues in the TH peptides could serve as surrogates for azol(in)e heterocycles in the substrate peptide core, as has been observed for other RiPP-modifying enzymes.[25,26] The TH1 and TH2 sequences were appended to the SrpE proteusin leader. Bromination of these chimeric substrates was observed *in vitro* [for SrpE-TH2 (Figure 3A,B and Figure S14)] and *in vivo* upon co-expression of peptide-encoding genes with *srpI* [for SrpE-TH1 (Figure 3D,E and Figure S15)]. We also verified that the bromination of SrpE-TH1 proceeded *in vitro* in a time-dependent manner (Figure 3G). Despite the TH core sequences being divergent from the SrpE and bereft of azoline heterocycles, SrpI maintained regiospecificity for the terminal Trp bromination at the indole-6 position (Figure S16). SrpI also maintained its rigid specificity for bromination, and chlorination of either substrate was not observed *in vivo*, or *in vitro* (Figures S17 and S18). Though the Gln residue in TH2 was well tolerated, the current inventory of SrpI substrates generally consists of nonpolar peptides. An expanded investigation of the substrate scope of SrpI will involve investigating whether charged residues can also be accommodated in the substrate core.

In line with the extensive application of aryl halogenation as a reactive handle for late-stage chemical diversifica-

https://doi.org/10.1021/acs.biochem.3c00222
Biochemistry XXXX, XXX, XXX−XXX

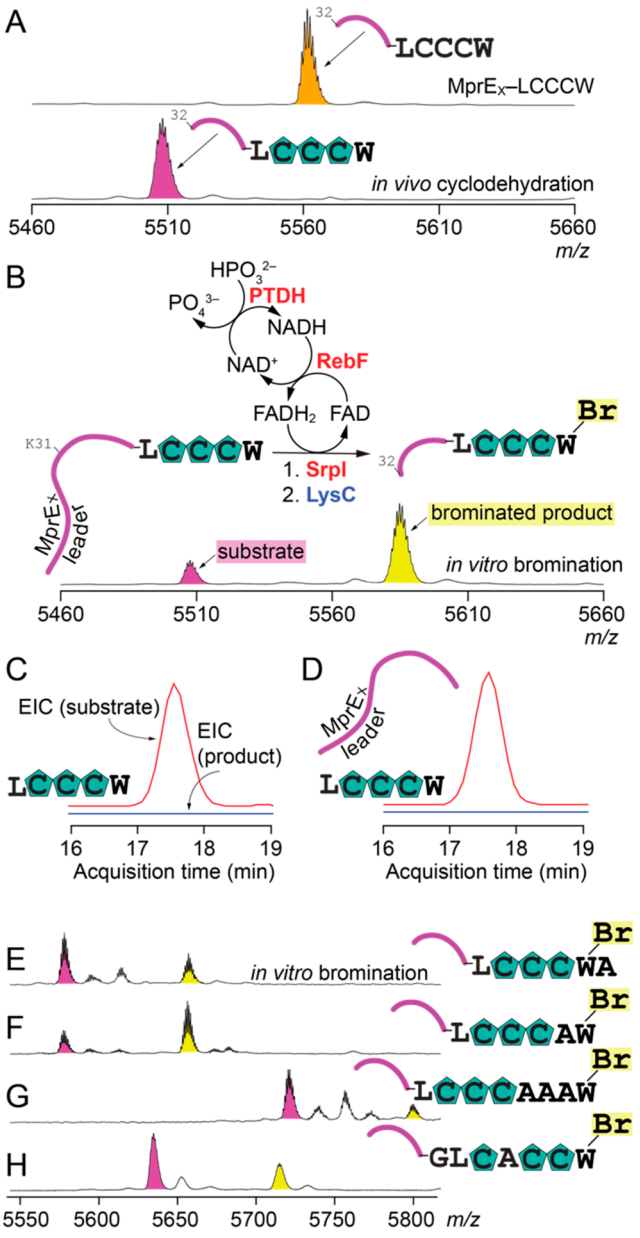


Figure 2. (A) MALDI-ToF MS spectra for the unmodified MprE$_X$–LCCCW chimeric peptide (top) and the MprE$_X$–LCCCW peptide in which the three Cys residues were converted to thiazol(in)e heterocycles (bottom). Digestion with the LysC protease removes the N-terminal 31 amino acids from the MprE$_X$ leader. (B) Reaction scheme for *in vitro* bromination by SrpI. LC/MS extracted ion chromatograms (EICs) demonstrating that bromination does not occur when (C) the modified LCCCW core is provided to SrpI by itself or (D) *in trans* with the dissociated MprE$_X$ leader peptide. The EIC for the modified LCCCW core is colored red, and that for the conceivable brominated product blue. The retention time for the brominated product lies within the acquisition time window illustrated here (Figure S6). (E–H) Bromination of modified cores appended to the MprE$_X$ leader.

tion,[10,12,13,28,29] we explored the Suzuki–Miyaura coupling of a panel of boronic acids to brominated peptides furnished by SrpI. For both *in vitro*-brominated SrpE-TH2 and *in vivo*-brominated SrpE-TH1, peptides that are >100 amino acids in length, coupling to boronic acids was observed (Figure 3C,F and Figures S19–S26). Obligate bromination by SrpI, without contaminating chlorination, allowed for mild reaction conditions in aqueous buffer. Qualitatively, in this proof-of-concept demonstration, benzylic boronic acids with electron-donating substituents delivered a higher yield of cross-coupling products. This observation was corroborated by the stoichiometric yield for coupling toluene and *p*-methoxyphenyl boronic acids to the *in vitro*-brominated SrpE-TH1 peptide (Figure 3H). SrpI also enabled the bromination and Suzuki–Miyaura coupling on large globular proteins. The SrpE-TH1 sequence was appended at the C-terminus of the 400-residue maltose binding protein (Figure 3I). The chimeric protein was a competent substrate for *in vitro* bromination by SrpI, followed by Suzuki–Miyaura coupling under conditions that did not require protein denaturation or the use of organic co-solvents (Figure 3J and Figures S27–S31). It is noteworthy that the SrpI-mediated strategy for peptide and protein labeling did not require the mutation of other Trp residues; SrpI itself maintains specificity for labeling only the C-terminal Trp. The MprE$_X$ leader possesses other Trp residues, as does the maltose binding protein, and they were not brominated and thus not conjugated.

While monitoring the halogenation assays mentioned above, we routinely observed the appearance of two brominated peptidic products, even in reactions in which the peptide/protein substrates were omitted. Using high-resolution mass spectrometry, we traced bromination to be occurring at two SrpI Tyr residues (SrpI Y102 and Y110); the brominated products mentioned above were generated by LysC digestion of SrpI in the reaction mixture (Figure 4A and Figure S32). Homology models indicate that these Tyr residues are proximal to the catalytic Lys residue (SrpI Lys84) that is implicated in forming a haloamine intermediate after halide oxidation at the flavin isoalloxazine ring or, as a proton donor, to facilitate resolution of the hypohalous acid intermediate (Figure S33).[21] Mutating either or both these Tyr residues did not compromise SrpI activity (Figure 4B). Our fortuitous discovery of SrpI self-halogenation was enabled by monitoring the progress of SrpI reactions using MALDI-ToF MS; it is conceivable that the self-halogenation could occur for other halogenating enzymes that bear electron rich amino acid side chains near the catalytic Lys residue.

Flavin-dependent halogenases have previously been used for the halogenation of short synthetic peptides containing tryptophan residues.[11] However, to the best of our knowledge, SrpI represents the first of its class enzyme for a leader peptide-guided bromination of genetically encoded peptides. The ability to deliver a bromide adduct selectively upon a single Trp residue in a ribosomally translated peptide and protein presents the opportunity to further develop SrpI as a biotechnology tool to facilitate bioorthogonal Suzuki–Miyaura cross-coupling reactions, and other conjugation reactions requiring halogenated peptide/protein precursors. These efforts will require an expanded investigation of the substrate scope for SrpI, as well as confirmation that brominating a fusion peptide at the C-terminus of a protein substrate does not alter the function or activity of the protein itself. As such, because the activity of SrpI is restricted to C-terminal residues and does not extend to internal Trp residues, SrpI likely serves to provide a route for peptide/protein labeling and bioconjugation, rather than modulation of the structure and activity of the biomolecular substrate.

Compared to amino acids such as Cys and Lys, chemical strategies for Trp arylation are sparse and almost exclusively restricted to position 2 of indole.[8] The regiospecificity of SrpI,

https://doi.org/10.1021/acs.biochem.3c00222
Biochemistry XXXX, XXX, XXX–XXX

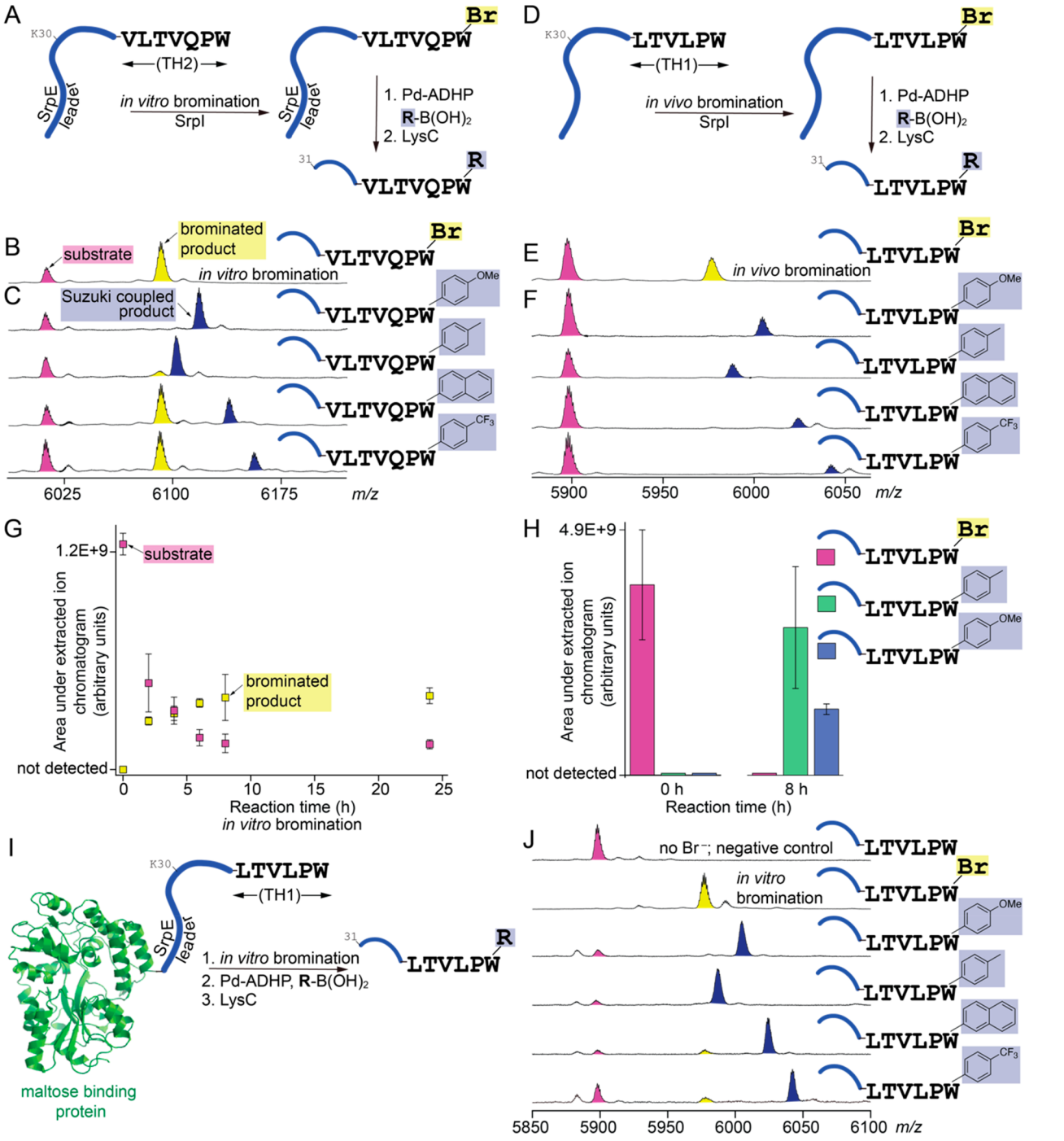


Figure 3. (A) Reaction scheme for *in vitro* bromination and Suzuki−Miyaura cross-coupling for SprE-TH2. Abbreviation: ADHP, 2-amino-4,6-dihydroxy-pyrimidine. (B and C) MALDI-ToF MS spectra demonstrating the unmodified peptide (pink peaks), brominated peptide (yellow peaks), and Suzuki−Miyaura coupling products (blue peaks) for the SprE-TH2 substrate. (D) Reaction scheme for *in vivo* bromination and Suzuki−Miyaura cross-coupling for SprE-TH1. (E and F) MALDI-ToF MS spectra demonstrating the unmodified peptide, brominated peptide, and coupling product for the SprE-TH1 substrate. (G) Time-dependent *in vitro* bromination of SrpE-TH1 by SrpI. Assays after proteolytic digestion were analyzed by liquid chromatography/mass spectrometry (LC/MS), and areas under the extracted ion chromatograms corresponding to the substrate peptide and the brominated product were determined. Means and standard deviations for three independent reactions are plotted. Note that the substrate and product peptides demonstrate disparate abundances using LC/MS. (H) Abundance of the brominated SrpE-TH1 peptide and after coupling to toluene and *p*-methoxyphenyl boronic acids monitored by LC/MS. (I) Reaction scheme for *in vitro* bromination and Suzuki−Miyaura cross-coupling for the SprE-TH1 peptide appended to the C-terminus of the maltose binding protein (illustrated as a green ribbon, Protein Data Bank entry 1FQD[27]). (J) MALDI-ToF MS spectra demonstrating a negative control reaction in which no bromide was added, and hence no brominated product was observed, followed by detection of the brominated and conjugated products.

bromination at indole-6, opens other sites on the indole side chain for modification. Several bottlenecks need to be overcome to improve the applicability of SrpI as a biocatalyst, among which is the limited activity of SrpI observed *in vivo*, contracting the leader peptide required for SrpI activity, and ameliorating the deleterious consumption of oxidized bromine

https://doi.org/10.1021/acs.biochem.3c00222
Biochemistry XXXX, XXX, XXX−XXX

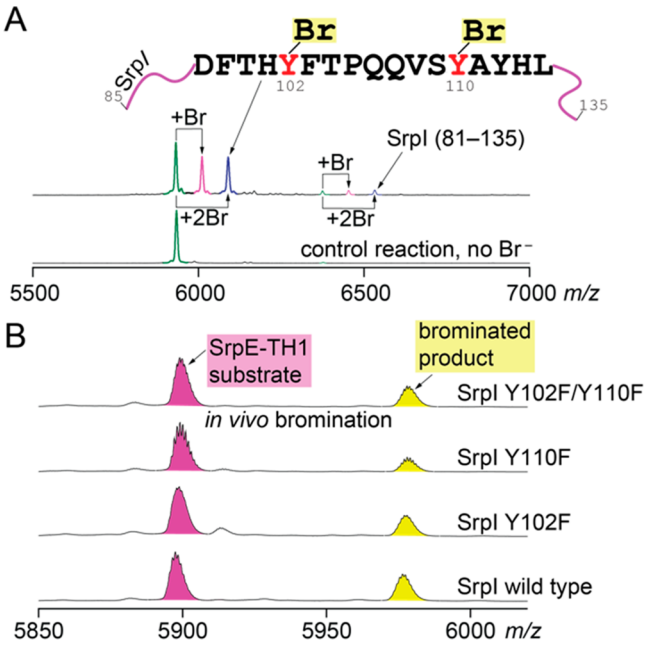


Figure 4. (A) MALDI-ToF MS spectra demonstrating the detection of mono- and dibrominated derivatives of SrpI 85−135 and SrpI 81−135 peptides generated by LysC digestion of SrpI. Self-bromination occurs at Tyr102 and Tyr110 residues. Bromide was omitted from the control reaction. (B) Bromination activity of wild type and mutant SrpI enzymes evaluated by *in vivo* bromination of the SrpE-TH1 chimeric substrate.

for self-halogenation of Tyr residues. Though in line with previous reports for cross-coupling reactions with peptidic substrates,[30−33] the organometallic catalyst loading in our reactions is currently high. A screening of organoboron reaction partners (Figure S34), core peptides of different lengths bearing the terminal Trp residue, and reaction conditions is currently underway.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.3c00222.

Experimental details for recombinant protein production, synthetic procedures, analytical procedures for compound characterization, Figures S1−S34, a supplementary note describing chimeric peptide sequences used in this study, and supplementary references (PDF)

Accession Codes

The SrpI and SrpE sequences are available from GenBank using the BioProject number PRJNA694437 and have been added to the supplementary note.

■ AUTHOR INFORMATION

Corresponding Author

Vinayak Agarwal − *School of Chemistry and Biochemistry and School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, United States;* orcid.org/0000-0002-2517-589X; Phone: (+1)404-385-378; Email: vagarwal@gatech.edu

Author

Nguyet A. Nguyen − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.biochem.3c00222

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank N. Arias, L. E. Roh, and N. Saha for technical assistance, N. Garg for acquiring mass spectrometry data, A. G. Roberts for insightful discussions, and the National Institutes of Health for financial support (R35GM142882 to V.A.).

■ REFERENCES

(1) Miyaura, N.; Suzuki, A. Palladium-catalyzed cross-coupling reactions of organoboron compounds. *Chem. Rev.* **1995**, *95* (7), 2457−2483.

(2) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. Palladium-catalyzed cross-coupling reactions in total synthesis. *Angew. Chem., Int. Ed.* **2005**, *44* (29), 4442−4489.

(3) Spicer, C. D.; Davis, B. G. Selective chemical protein modification. *Nat. Commun.* **2014**, *5* (1), 4740.

(4) Lennox, A. J. J.; Lloyd-Jones, G. C. Selection of boron reagents for Suzuki−Miyaura coupling. *Chem. Soc. Rev.* **2014**, *43* (1), 412−443.

(5) Spicer, C. D.; Triemer, T.; Davis, B. G. Palladium-mediated cell-surface labeling. *J. Am. Chem. Soc.* **2012**, *134* (2), 800−803.

(6) Zhao, Q.; Guo, G.; Zhu, W.; Zhu, L.; Da, Y.; Han, Y.; Xu, H.; Wu, S.; Cheng, Y.; Zhou, Y.; Cai, X.; Jiang, X. Suzuki cross-coupling reaction with genetically encoded fluorosulfates for fluorogenic protein labeling. *Chem. - Eur. J.* **2020**, *26* (68), 15938−15943.

(7) Chalker, J. M.; Wood, C. S. C.; Davis, B. G. A convenient catalyst for aqueous and protein Suzuki−Miyaura cross-coupling. *J. Am. Chem. Soc.* **2009**, *131* (45), 16346−16347.

(8) Zhang, C.; Vinogradova, E. V.; Spokoyny, A. M.; Buchwald, S. L.; Pentelute, B. L. Arylation chemistry for bioconjugation. *Angew. Chem., Int. Ed.* **2019**, *58* (15), 4810−4839.

(9) Ojida, A.; Tsutsumi, H.; Kasagi, N.; Hamachi, I. Suzuki coupling for protein modification. *Tetrahedron Lett.* **2005**, *46* (19), 3301−3305.

(10) Lee, T.-K.; Manandhar, B.; Kassees, K. J.; Ahn, J.-M. Peptide ligation via the Suzuki−Miyaura cross-coupling reaction. *Journal of Organic Chemistry* **2020**, *85* (3), 1376−1384.

(11) Schnepel, C.; Moritzer, A.-C.; Gäfe, S.; Montua, N.; Minges, H.; Nieß, A.; Niemann, H. H.; Sewald, N. Enzymatic late-stage halogenation of peptides. *ChemBioChem* **2023**, *24* (1), e202200569.

(12) Crowe, C.; Molyneux, S.; Sharma, S. V.; Zhang, Y.; Gkotsi, D. S.; Connaris, H.; Goss, R. J. M. Halogenases: a palette of emerging opportunities for synthetic biology−synthetic chemistry and C−H functionalisation. *Chem. Soc. Rev.* **2021**, *50* (17), 9443−9481.

(13) Gruß, H.; Sewald, N. Late-stage diversification of tryptophan-derived biomolecules. *Chem. - Eur. J.* **2020**, *26* (24), 5328−5340.

(14) Agarwal, V.; Miles, Z. D.; Winter, J. M.; Eustaquio, A. S.; El Gamal, A. A.; Moore, B. S. Enzymatic halogenation and dehalogenation reactions: pervasive and mechanistically diverse. *Chem. Rev.* **2017**, *117* (8), 5619−5674.

(15) Ortega, M. A.; Cogan, D. P.; Mukherjee, S.; Garg, N.; Li, B.; Thibodeaux, G. N.; Maffioli, S. I.; Donadio, S.; Sosio, M.; Escano, J.; Smith, L.; Nair, S. K.; van der Donk, W. A. Two flavoenzymes catalyze the post-translational generation of 5-chlorotryptophan and 2-aminovinyl-cysteine during NAI-107 biosynthesis. *ACS Chem. Biol.* **2017**, *12* (2), 548−557.

(16) Nguyen, N. A.; Lin, Z.; Mohanty, I.; Garg, N.; Schmidt, E. W.; Agarwal, V. An obligate peptidyl brominase underlies the discovery of

highly distributed biosynthetic gene clusters in marine sponge microbiomes. *J. Am. Chem. Soc.* **2021**, *143* (27), 10221−10231.
(17) Nguyen, N. A.; Cong, Y.; Hurrell, R. C.; Arias, N.; Garg, N.; Puri, A. W.; Schmidt, E. W.; Agarwal, V. A silent biosynthetic gene cluster from a methanotrophic bacterium potentiates discovery of a substrate promiscuous proteusin cyclodehydratase. *ACS Chem. Biol.* **2022**, *17* (6), 1577−1585.
(18) Haft, D. H.; Basu, M. K.; Mitchell, D. A. Expansion of ribosomally produced natural products: a nitrile hydratase- and Nif11-related precursor family. *BMC Biol.* **2010**, *8* (1), 70.
(19) Yeh, E.; Garneau, S.; Walsh, C. T. Robust *in vitro* activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebeccamycin biosynthesis. *Proceedings of the National Academy of Sciences U S A* **2005**, *102* (11), 3960−5.
(20) Johannes, T. W.; Woodyer, R. D.; Zhao, H. Efficient regeneration of NADPH using an engineered phosphite dehydrogenase. *Biotechnol. Bioeng.* **2007**, *96* (1), 18−26.
(21) Phintha, A.; Prakinee, K.; Jaruwat, A.; Lawan, N.; Visitsatthawong, S.; Kantiwiriyawanitch, C.; Songsungthong, W.; Trisrivirat, D.; Chenprakhon, P.; Mulholland, A.; van Pée, K.-H.; Chitnumsub, P.; Chaiyen, P. Dissecting the low catalytic capability of flavin-dependent halogenases. *J. Biol. Chem.* **2021**, *296*, 100068.
(22) Bobeica, S. C.; Dong, S.-H.; Huo, L.; Mazo, N.; McLaughlin, M. I.; Jiménez-Osés, G.; Nair, S. K.; van der Donk, W. A. Insights into AMS/PCAT transporters from biochemical and structural characterization of a double Glycine motif protease. *eLife* **2019**, *8*, e42305.
(23) Kondo, E.; Iioka, H.; Saito, K. Tumor-homing peptide and its utility for advanced cancer medicine. *Cancer Science* **2021**, *112* (6), 2118−2125.
(24) Shadidi, M.; Sioud, M. Identification of novel carrier peptides for the specific delivery of therapeutics into cancer cells. *FASEB J.* **2003**, *17* (2), 256−258.
(25) Sarkar, S.; Gu, W.; Schmidt, E. W. Expanding the chemical space of synthetic cyclic peptides using a promiscuous macrocyclase from prenylagaramide biosynthesis. *ACS Catal.* **2020**, *10* (13), 7146−7153.
(26) Mohanty, I.; Nguyen, N. A.; Moore, S. G.; Biggs, J. S.; Gaul, D. A.; Garg, N.; Agarwal, V. Enzymatic synthesis assisted discovery of proline-rich macrocyclic peptides in marine sponges. *ChemBioChem.* **2021**, *22* (16), 2614−2618.
(27) Duan, X.; Hall, J. A.; Nikaido, H.; Quiocho, F. A. Crystal structures of the maltodextrin/maltose-binding protein complexed with reduced oligosaccharides: flexibility of tertiary structure and ligand binding. *J. Mol. Biol.* **2001**, *306* (5), 1115−1126.
(28) Runguphan, W.; Qu, X.; O'Connor, S. E. Integrating carbon-halogen bond formation into medicinal plant metabolism. *Nature* **2010**, *468* (7322), 461−4.
(29) Pathak, T. P.; Miller, S. J. Chemical tailoring of teicoplanin with site-selective reactions. *J. Am. Chem. Soc.* **2013**, *135* (22), 8415−8422.
(30) Dumas, A.; Spicer, C. D.; Gao, Z.; Takehana, T.; Lin, Y. A.; Yasukohchi, T.; Davis, B. G. Self-liganded Suzuki−Miyaura coupling for site-selective protein PEGylation. *Angew. Chem., Int. Ed.* **2013**, *52* (14), 3916−3921.
(31) Spicer, C. D.; Davis, B. G. Palladium-mediated site-selective Suzuki−Miyaura protein modification at genetically encoded aryl halides. *Chem. Commun.* **2011**, *47* (6), 1698−1700.
(32) Li, N.; Lim, R. K. V.; Edwardraja, S.; Lin, Q. Copper-free Sonogashira cross-coupling for functionalization of alkyne-encoded proteins in aqueous medium and in bacterial cells. *J. Am. Chem. Soc.* **2011**, *133* (39), 15316−15319.
(33) Gao, Z.; Gouverneur, V.; Davis, B. G. Enhanced aqueous Suzuki−Miyaura coupling allows site-specific polypeptide 18F-labeling. *J. Am. Chem. Soc.* **2013**, *135* (37), 13612−13615.

https://doi.org/10.1021/acs.biochem.3c00222
Biochemistry XXXX, XXX, XXX−XXX

SUPPLEMENTARY INFORMATION FOR:

A leader-guided substrate tolerant RiPP brominase allows Suzuki–Miyaura cross-coupling reactions for peptides and proteins

Nguyet A. Nguyen,[1] and Vinayak Agarwal[1,2,*]

[1]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332, USA

[2]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA

*Correspondence: vagarwal@gatech.edu; Ph: (+1)404-385-378

SUPPLEMENTARY MATERIALS AND METHODS

Preparation of constructs for expression of SrpI, PtdH, RebF, and chimera peptides

Genes optimized for expression in *Escherichia coli* for SrpI, PtdH, RebF, and the chimeric substrate peptides were used as templates for the amplification and subcloning of PCR amplicons in plasmid vectors. The chimeric peptides were designed by incorporating the leader peptide SrpE or the consensus leader peptide sequence $MprE_X$ as described in the Supplementary Note (below).

General polymerase chain reaction protocol

PCR reactions (25 µL) contained 20 ng template DNA, 0.4 µM each of reverse and forward primers, 0.2 mM dNTPs, Phusion reaction buffer, and 0.25 U Phusion-high fidelity DNA polymerase (Thermo). PCR amplicons were subcloned into plasmid vectors using Gibson Assembly HiFi master mix (NEB). For co-expression with chimeric substrate peptide genes, the modification enzyme encoding genes (*mprC*, *mprD*, and *srpI*) were subcloned without any affinity chromatography tags. All constructs were verified by Sanger sequencing.

Protein expression and purification

General procedure for heterologous protein expression: Plasmid DNA (20 ng) was transformed in *E. coli* BL21(DE3). Colonies were grown under appropriate antibiotic selection on LB agar media for 16 h. A single colony was picked and inoculated in 10 mL of terrific broth (TB) supplemented with appropriate antibiotics for 16 h at 37 °C. This inoculum was used to initiate 1 L TB cultures supplemented with corresponding antibiotics. Cultures were incubated with shaking at 30 °C until the OD600 reached 0.6. Cultures were cooled at 18 °C for 1 h before induction of protein expression by adding 0.3 mM isopropyl-β-d-thiogalactopyranoside (IPTG). Cultures were incubated at 18 °C, 180 rpm for 24 h.

General purification protocol for N-His_6-substrate peptides: Cultures were harvested by centrifugation (5,000 rpm, 25 min, 4 °C) and resuspended in 50 mL lysis buffer A (20 mM Na-phosphate (pH 7.5), 50 mM NaCl, 4M Gdn-HCl). Cells were lysed by sonication, and the lysate was clarified by centrifugation at 18,000 rpm for 45 min at 4 °C. The supernatant was loaded onto a 5 mL His-Trap Ni-NTA column equilibrated with the lysis buffer. The column was washed with 10 mL wash buffer A1 (20 mM Na-phosphate (pH 7.5), 50 mM NaCl, 30 mM imidazole, 4M Gdn-HCl), 10 mL of 1:1 wash buffer A1:wash buffer A2 (20 mM Na-phosphate (pH 7.5), 50 mM NaCl, 30 mM imidazole), and 10 mL wash buffer A2. Bound proteins were then eluted using elution buffer A (20 mM Na-phosphate (pH 7.5), 50

mM NaCl, 1M imidazole) in three fractions of 5 mL volume each. The second fraction was found to contain the highest concentration of the peptide, and thus, collected. This fraction was desalted using Sephadex G-25 PD10 column in storage buffer B (20 mM Na-phosphate (pH 7.5), 100 mM NaCl). The peptide concentration was measured by Bradford assay. Aliquots were frozen and stored at –80 °C for future use.

Purification of PtdH and RebF was performed according to published protocols.[1-2]

Purification of N-His_6-SrpI: Chaperones (from chaperone plasmid pGro7, Takara) were co-expressed to assist the folding of SrpI. After addition of IPTG and 48 h of incubation at 18 °C, cultures were harvested by centrifugation as above, and cell pellets were resuspended in lysis buffer B (20 mM Na-phosphate (pH 7.5), 100 mM NaCl). Cells were lysed by sonication. The lysate was clarified by centrifugation at 18,000 rpm for 60 min. The supernatant was loaded onto a 5 mL His-Trap Ni-NTA column. The column was washed extensively with wash buffer B (20 mM Na-phosphate (pH 7.5), 100 mM NaCl, 30 mM imidazole), and protein was eluted using a linear gradient from 0% to 100% elution buffer B (20 mM Na-phosphate (pH 7.5), 100 mM NaCl, 250 mM imidazole). The purity of eluent fractions was checked by SDS-PAGE, and fractions containing protein of interest were pooled and concentrated by Amicon® Ultra Centrifugal Filter Unit. The protein was further purified by size exclusion chromatography on a Superdex 75 16/200 column with 20 mM Na-phosphate (pH 7.5), 100 mM NaCl buffer. The purity of eluent fractions was checked by SDS-PAGE and pure fractions pooled. The concentrations were measured by Bradford assay. Aliquots were frozen and stored at –80 °C for future use.

Enzymatic and chemo-enzymatic assays

Halogenation assays: Bromination reactions were performed in 200 µL volume containing 50 mM HEPES-Na (pH 7.5), 20 mM KBr, 25 µM FAD, 0.625 mM NAD^+, 6.25 mM Na_2HPO_3, 5 µM flavin reductase (RebF), 5 µM phosphite dehydrogenase (PTDH), 100 µM substrate peptide, 20 µM SrpI, 0.05 µg/µL catalase. After 24 h incubation at 30 °C, reactions were quenched by protease addition (LysC, LahT150, or carboxypeptidase A) according to published protocols.[3] The reactions treated with LysC were then desalted and analyzed by MALDI-ToF, the reactions treated with LahT150 and carboxypeptidase A were analyzed by HPLC-MS/MS.

Indole was used as a substrate at a final concentration of 2 mM in the above-mentioned assay. After 24 h of incubation, an equal volume of EtOAc was used to extract the assay. The mixture was then centrifuged at 16,000 rpm for 20 min to remove debris. The EtOAc extract was analyzed by GC-MS (1260G with 7890a MS; Agilent Technologies) in electron ionization (70 eV) mode using a DF-5ms ultra inert GC column (30 m length, 0.25 mm width and 0.5 µM film thickness). The column temperature conditions were as follows: 40 °C for 3 min, increased to 200 °C at 10 °C/min, and held for 1 min with a total run time of 20 min. The injection port, interface, and ion source were kept at 250 °C, 300 °C, and 230 °C, respectively. Helium was used as carrier gas at a 0.9 mL/min flow rate.

Time-course experiments to monitor the bromination of SrpE-TH1: The experimental procedure was conducted with a total volume of 600 µL, comprising a mixture of 50 mM HEPES-Na (pH 7.5), 20 mM KBr, 25 µM FAD, 0.625 mM NAD^+, 6.25 mM Na_2HPO_3, 5 µM flavin reductase (RebF), 5 µM phosphite dehydrogenase (PTDH), 20 µM substrate peptide SrpE-TH1, 10 µM SrpI, and 0.05 µg/µL catalase. To ensure optimal O_2 transfer, 100 µL assay was transferred and incubated in individual tubes. At 0, 2, 4, 6, 8, and 24 h, the reaction was quenched by adding 1 µL of 6N HCl, followed by brief centrifugation to remove precipitated protein. The reaction mixture was neutralized with 1 µL of 6N NaOH, and then 1 µL of 2 mg/mL GluC was added and incubated at 30 °C for an additional 2 h. Any precipitation was removed by centrifugation after adding 100 µL of MeOH. Finally, 100 µL of the quenched reaction mixture was analyzed by HPLC-MS, and the area under the curve for the starting material and product was calculated using the extracted chromatogram for $[M+2H]^{2+}$ ion at *m/z* 564.33 and 604.28, respectively.

d-labeling of indole: 2.4 mg indole was dissolved in 500 µL D_2O and stirred at 150 ^{0}C overnight. The product was analyzed by NMR using a 400 MHz Bruker NMR showing that conversion of indole to indole-3-*d* progressed with 85% yield. ^{1}H-NMR (400 MHz, D_2O): δ 7.61 (dt, 1H), δ 7.45 (dt, 1H), δ 7.31 (s, 1H), δ 7.15 (ddd, 1H), δ 7.06 (ddd, 1H). The sample was used for enzymatic halogenation without further purification.

Suzuki-Miyaura cross-coupling reactions

Several water soluble ligands have been developed to facilitate Suzuki-Miyaura cross coupling reactions, including phosphine-based ligands[4-5] and nitrogen-based ligands[6-7] (Fig. S34). Nitrogen-based Pd-ligands containing the guanidine moiety have been shown to be less sensitive to air than the

phosphine-based counterparts. Intensive studies in developing mild protocol for cross-coupling reaction utilizing guanidine moiety-containing ligands for proteins were conducted by Davis,[7-9] Lin,[10] and Zhang.[11] In our proof-of-concept demonstration, we adapted procedures for Suzuki-Miyaura cross-coupling reactions from Davis.[9]

Preparation of Pd-ADHP: To 5 mL 0.1 M NaOH, 13 mg 2-amino-4,6-dihydroxy pyrimidine (0.1 mmol) was added, followed by 11 mg palladium acetate (0.05 mmol). The solution was stirred at 65 ºC for 30 min to yield 0.01 M catalyst solution.

Brominated peptide SrpE-TH1 was obtained by co-expression of the substrate peptide (SrpE-TH1) encoding gene and *srpI*. The brominated peptide was purified and stored in 20 mM Na-phosphate (pH 7.5), 100 mM NaCl buffer.

The halogenation of the chimeric substrates SrpE-TH2 and MBP-SrpE-TH1 was performed *in vitro* on a 10 mL-scale reaction as described above. The reaction solution was desalted by Sephadex G-25 PD10 column using 20 mM Na-phosphate (pH 7.5), 100 mM NaCl buffer to remove all cofactors that have been shown to have a deleterious effect on Pd-catalyzed cross-coupling reaction.[12]

The cross-coupling reactions were performed in 200 µL volume containing 250 µM halogenated substrate peptide or proteins, 2.5 mM Pd-ADHP, 75 mM Na_2HPO_4, and 2.5 mM boronic acid. The reaction was incubated at 45 °C for 8 h. 1 µL 0.2 mg/mL LysC or GluC proteases were added and then incubated at 30 °C for further 90 min. 100 µL of 5 µL/mL 3-mercaptopropionic acid was added to chelate the excess Pd. The reactions treated with LysC were analyzed by MALDI-ToF, and the reactions treated with GluC were analyzed by LC/MS, as described below.

Mass spectrometry

For matrix-assisted laser desorption-ionization-time of flight mass spectrometry (MALDI-ToF MS), reactions samples were desalted using C_{18} ZipTips (Sigma) and spotted on a MALDI target using 2 µL saturated sinapinic acid (Sigma) in 7:3:0.1 MeCN:H_2O:TFA solvent for analysis by a rapifleX MALDI-ToF mass spectrometer (Bruker Daltonics) in reflectron positive ionization mode. The data were analyzed using flexAnalysis software.

The Suzuki-Miyaura coupling reaction products, the bromotryptophan standards, and the excision product generated by carboxypeptidase A were analyzed using Vanquish Flex UHPLC (Thermo) coupled to a Q Exactive HF hybrid quadrupole-orbitrap mass spectrometer (Thermo). Mass spectrometry data were collected in the positive ionization mode in the mass range *m/z* 100–2000 Da. Samples were analyzed using Acquity UPLC BEH C_{18} 1.7 µm column (100×2.1 mm) at a flow rate of 0.4 mL/min, and the chromatographic separation was achieved using solvent A: H_2O + 0.1% (v/v) formic acid, and solvent

B: MeCN + 0.1% (v/v) formic acid. The chromatography profile was as follows: 5% solvent B from 0–0.5 min, linear gradient to 100% solvent B from 0.5–7 min, 100% solvent B from 7–10 min, linear gradient to 5% solvent B from 10–11 min, 5% solvent B from 11–12 min.

The halogenated peptides after the digestion with LahT150 were analyzed by Bruker amaZon SL ion trap mass spectrometer coupled to an Agilent 1260 HPLC. Mass spectrometry data were collected in the positive ionization mode in the mass range *m/z* 100–1000 Da. Chromatography was performed using Luna 5 µm C_{18} reversed-phase HPLC column (100×4.6 mm) at a flow rate of 0.5 mL/min; solvent A: H_2O + 0.1% v/v formic acid, solvent B: MeCN + 0.1% v/v formic acid. The chromatography elution profile was as follows: 5% solvent B from 0–5 min, linear gradient to 100% solvent B from 5–35 min, 100% solvent B from 35–40 min, linear gradient to 5% solvent B from 40–41 min, and 5% solvent B from 41–42 min.

SUPPLEMENTARY NOTE

The sequence of the $MprE_X$ leader peptide:

MNEEQMQQYSQIVAKCWADAEFKAKLIADPKATLAAEGIAVPDGIELRVLENTATTVNLVLPPPP[§]AEGEL
SDEDL[‡]GAVTGG[†]

The sequence of the SrpE leader peptide:

MRSGDDMLQHLVEKSALDADFRQQLLADPKSTISQELGISIPESMTIRVHESDMETVHLALPPDP[§]NLTEE
QL[‡]EAISAG[†]

The double glycine motif[†] which marks the leader/core boundary, the proline-rich motif,[§] and the YcaO binding motif[‡] are underlined.

The sequence of the SrpI halogenase:

MIQPGSESLRKIAVIGRGTAGSLAAASVTRLHPDADHELHHIYDSRIPVIGVGEGSWPSLVQEVQQLTGL
PHETVQQRLKGTRKYGVAFEGWGRRGRDFTHYFTPQQVSYAYHLSADLLADMLHESSRARHIDAKVLDIA
RVDGGARVEFEGRAPERYDLVFDARGFPRELDTDEHIDISFIPTNTAVIRRCPAIVEEAAGPVLQHTYTR
AVARPHGWIFVIPLAVHTSYGYIFNRDVTGLDEVESDFDAFLETDGVPEFEQRAVLRFPNFVHRRIYDGA
VARIGNAAAFMEPLEATAIVSAQIQIGMVLKTRLGRSVEHLDRDAPAVNRFLVKNVLRYGLFVGWHYSCG
SRYDSPFWRFARDRTWPRYRSAADPAAVDCNALGEFDEMIRLLHQPVIDQGDWHRMCAVPLTSYAQMSQG
LGC

SUPPLEMENTARY TABLE

Table S1. Calculated and observed masses of substrate and product peptides in this study

Compound	Calculated mass	Observed mass
LCCCW	569.15	569.19
Br-LCCCW	647.06	647.06
3-Br-indole	194.97	194.86
Indole-3-*d*	118.06	117.96
LCCCWA	640.18	640.18
Br-LCCCWA	718.09	718.09
LCCCAW	640.18	640.18
Br-LCCCAW	718.09	718.09
LCCCAAAW	782.26	782.26
Br-LCCCAAAW	862.17	862.24
LCCCWAAA	5717.81	5717.52
GLCACCW	697.20	697.20
Br-GLCACCW	775.11	775.12
GLCWCCC	5720.78	5718.28
GLCAWCC	5688.81	5687.22
AISAGLTVLPW (TH1)	1127.65	1127.65
AISAGVLTVQPW (TH2)	1241.69	1241.70
Br-TH1	1205.55	1205.55
Br-TH2	1319.60	1319.59
4-Methoxyphenyl-TH1	1233.69	1233.69
4-Methylbenzene-TH1	1217.69	1217.69
2-Naphthalene-TH1	1253.69	1253.69
4-Trifluoromethylphenyl-TH1	1271.66	1271.66
4-Methoxyphenyl-TH2	1347.73	1347.73
4-Methylbenzene-TH2	1331.74	1331.74
2-Naphthalene-TH2	1367.74	1367.74
4-Trifluoromethylphenyl-TH2	1385.70	1385.71

SUPPLEMENTARY FIGURES

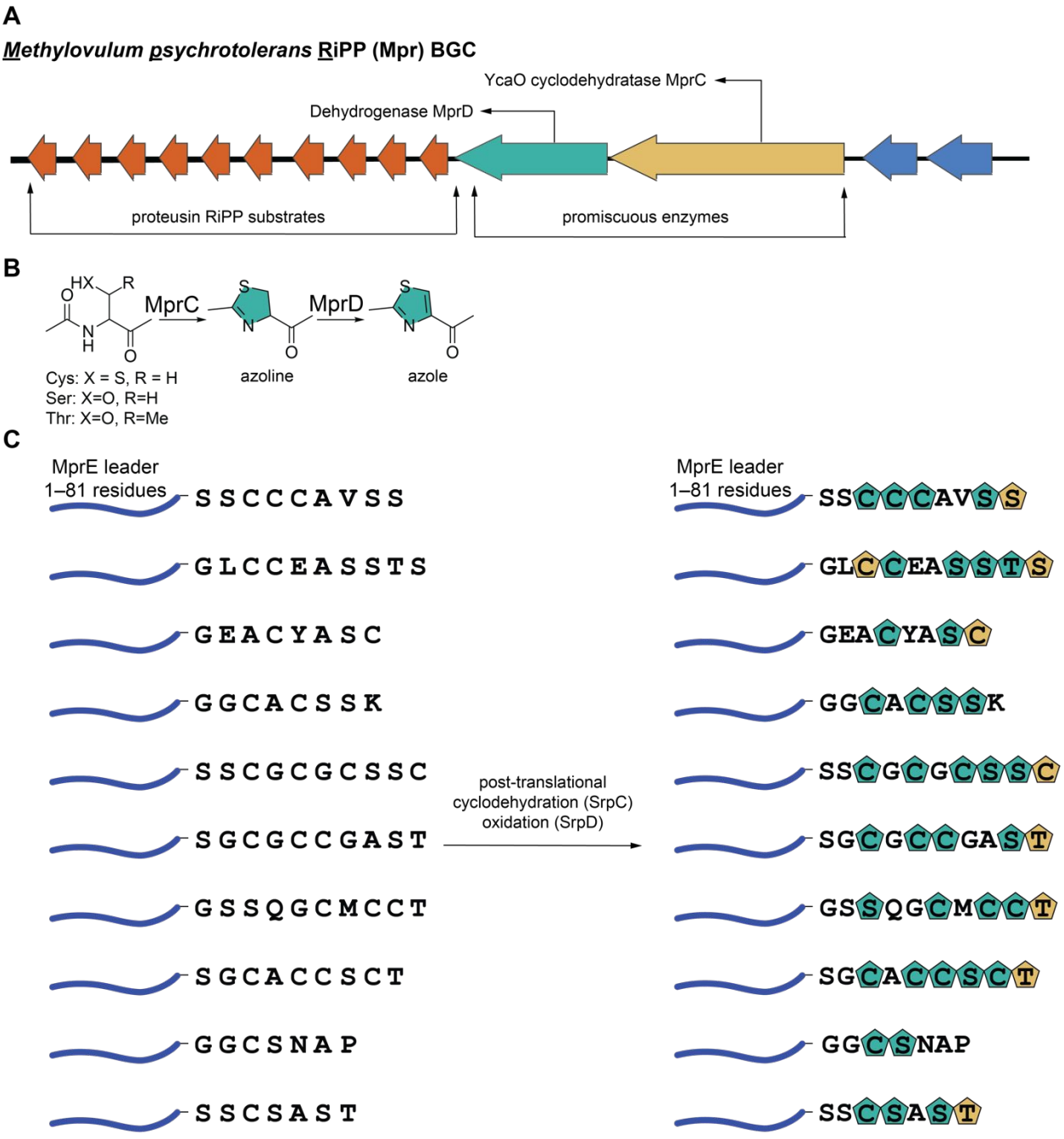


Fig. S1. (**A**) The *mpr* biosynthetic gene cluster showing the genes *mprC* and *mprD* gene clustered with genes encoding ten different MprE substrate peptides. (**B**) MprC catalyses the cyclodehydration of Cys, Ser, and Thr residues to form azoline rings. MprD catalyses the oxidation of azolines to azoles. (**C**) Promicuity activity of MprC/MprD allows for the installation of azol(in)e heterocycles in ten different MprE substrate peptides.

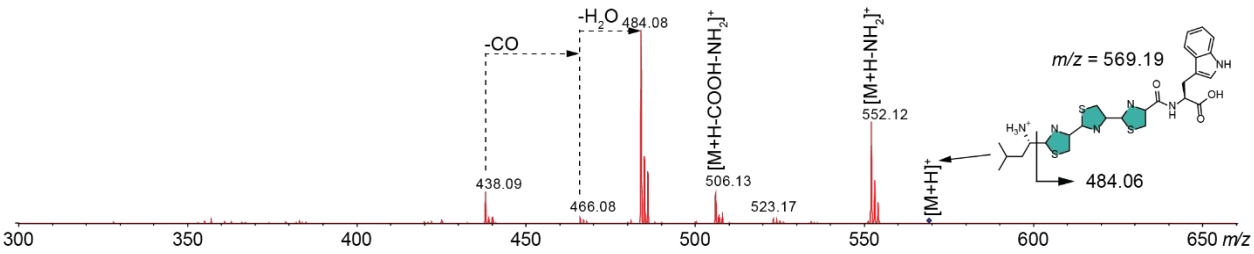


Fig. S2. Structural annotation of the MS^2 fragmentation spectra for the modified SrpE core (LCCCW) obtained by co-expression of the gene encoding the $MprE_X$-LCCCW chimeric substrate with *mprC and mprD*, followed by the treatment of the purified product with the protease LahT150 to remove the leader peptide and furnish the modified core as the product.

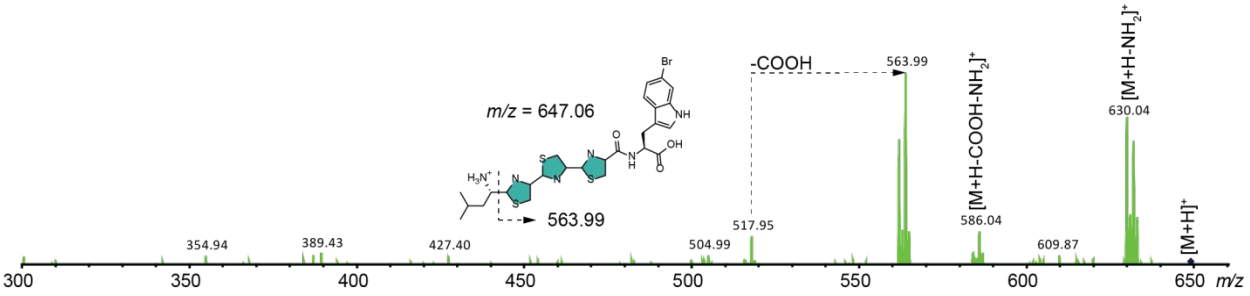
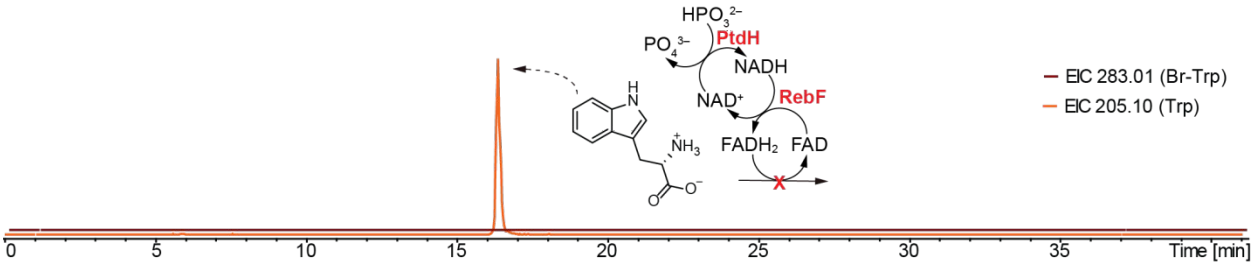


Fig. S3. Structural annotation of the MS^2 fragmentation spectra for the brominated SrpE core obtained by *in vitro* halogenation of modified $MprE_X$-LCCCW substrate by SrpI followed by the treatment of the product with LahT150.

184 

185 **Fig. S4.** LC-MS extracted ion chromatograms (EICs) for tryptophan and bromotryptophan showing that
186 the bromination of free tryptophan is not catalyzed by SrpI.

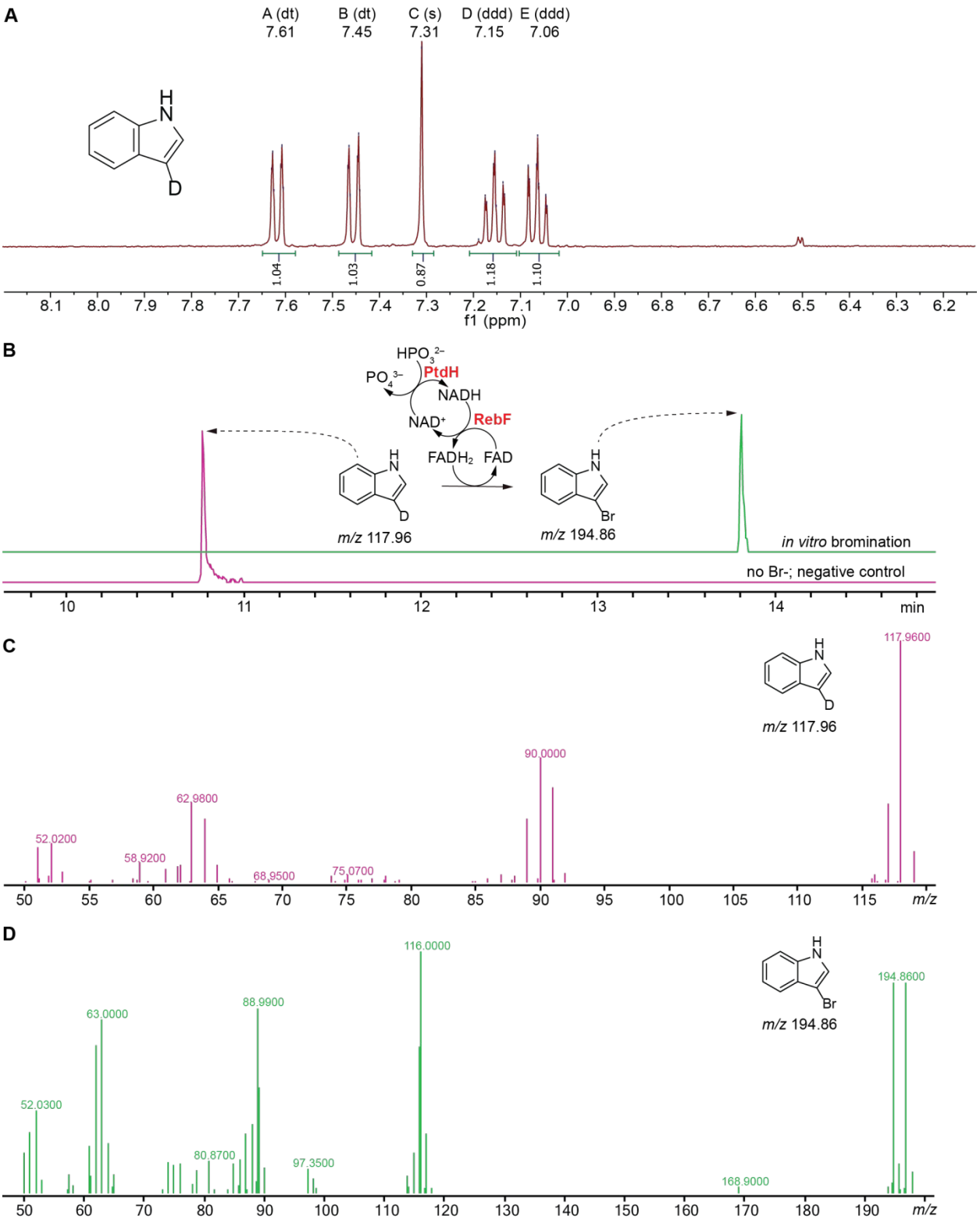
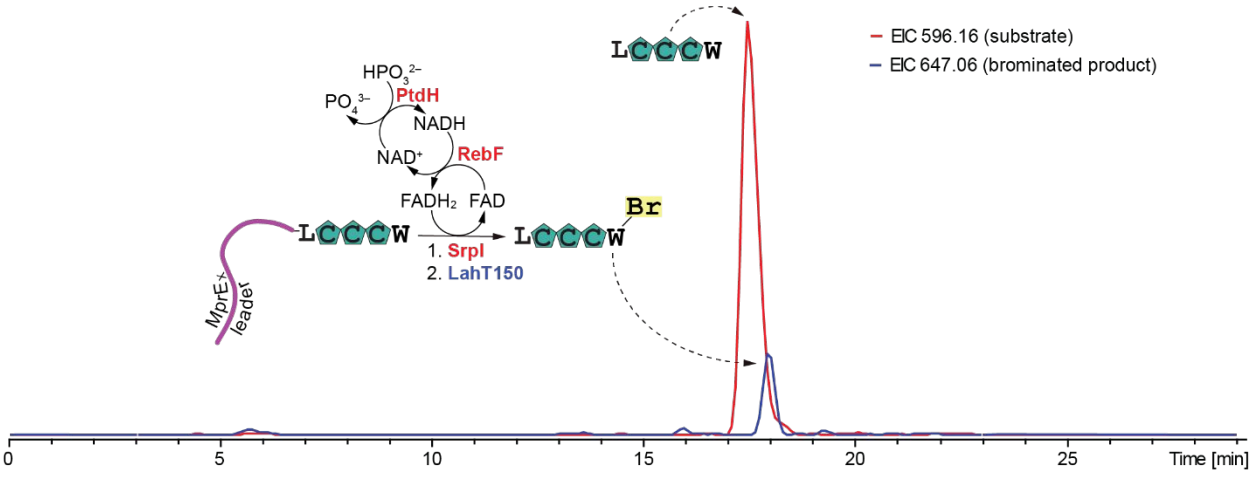


Fig. S5. (**A**) ^{1}H-NMR (400 MHz, D_2O) spectrum for indole-3-*d*. (**B**) GC-MS analysis for SrpI assays for indole-3-*d* showing the brominated product as 3-Br-indole. (**C**) MS spectra of indole-3-*d*. (**D**) MS spectra of 3-Br indole.



191

192 **Fig. S6.** Reaction scheme for *in vitro* bromination of the modified $MprE_X$-LCCCW chimeric substrate by
193 SrpI followed by the treatment of the product with LahT150. LC-MS EICs demonstrate that the substrate
194 and brominated products, after scission from the leader, are eluted within the time window from 16–19
195 min as is illustrated in Fig. 2C.

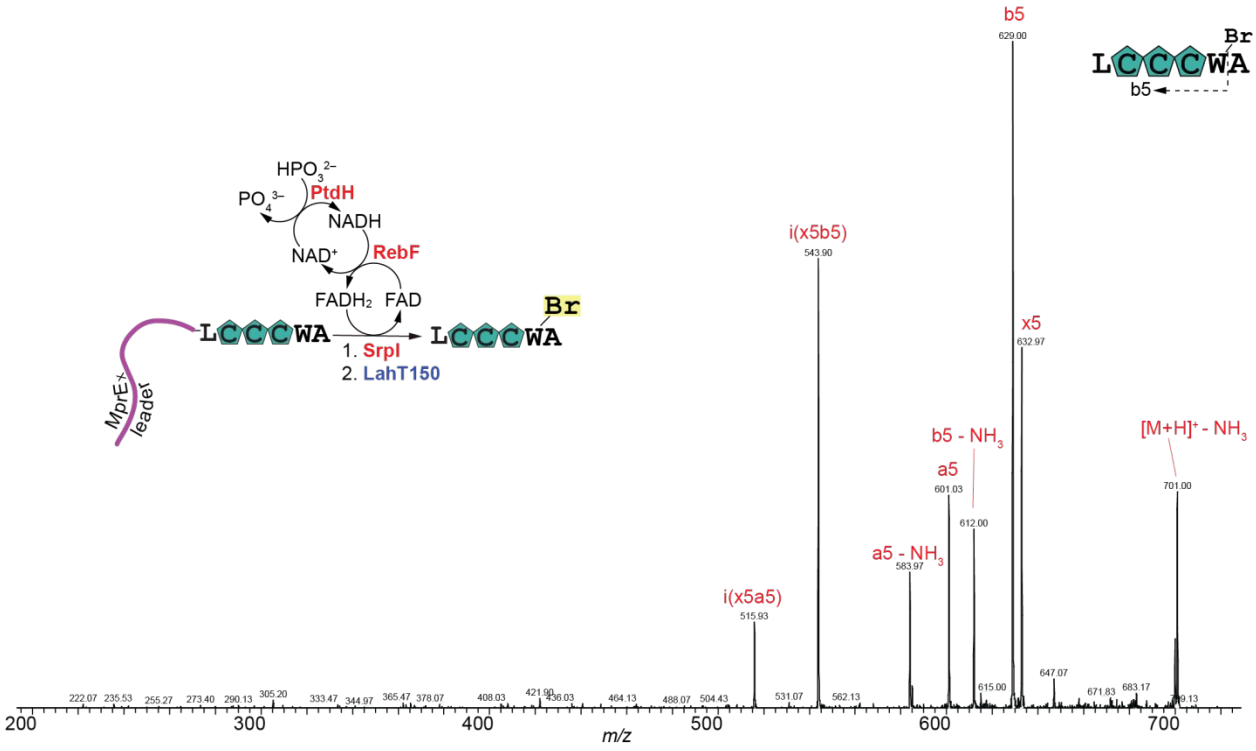


Fig. S7. Structural annotation of the MS^2 fragmentation spectra for the brominated L<u>CCC</u>WA core.

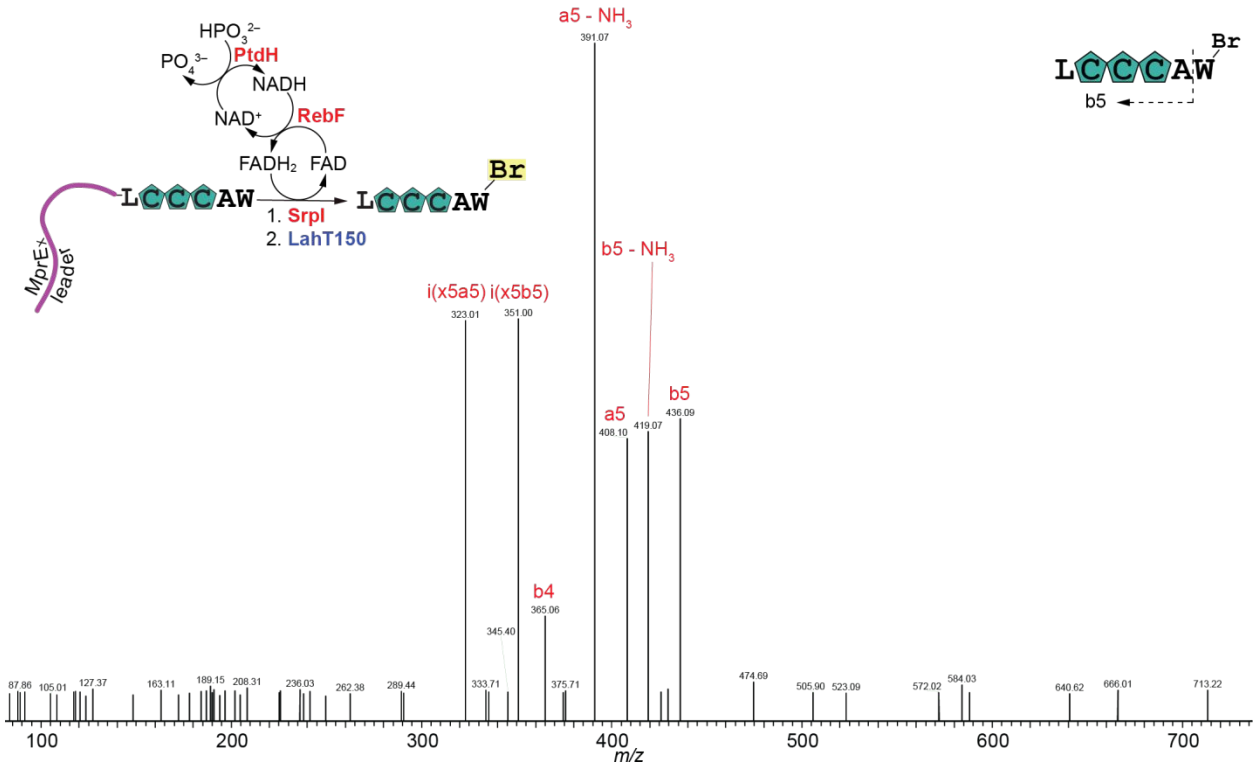
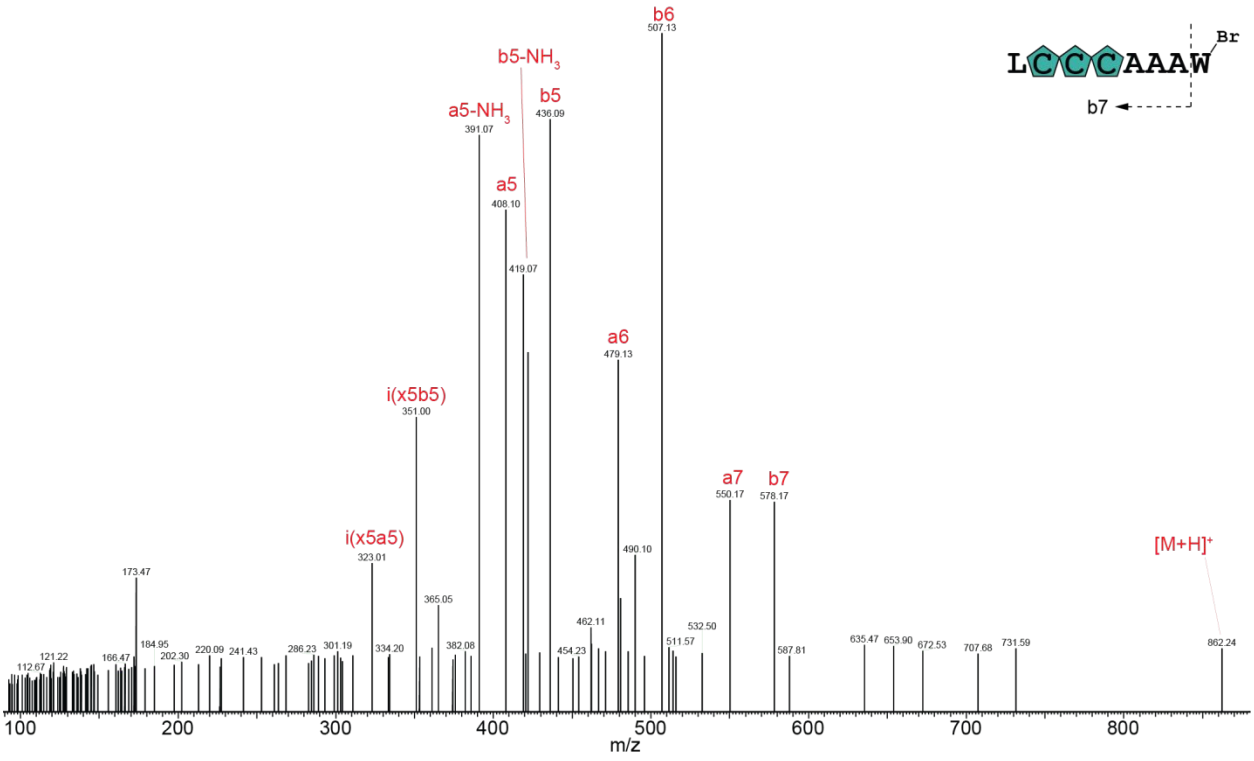


Fig. S8. Structural annotation of the MS^2 fragmentation spectra for the brominated LCCCAW core.



200

201 **Fig. S9.** Structural annotation of the MS^2 fragmentation spectra for the brominated LCCCAAAW core.

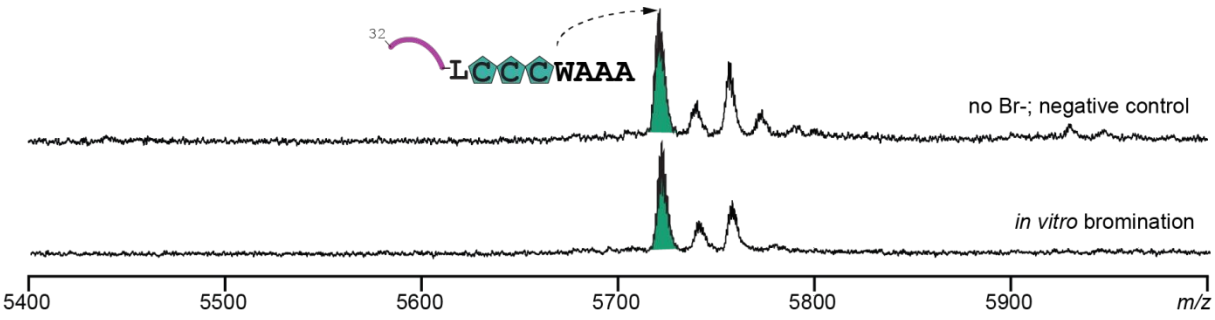


Fig. S10. MALDI-ToF MS spectra for the *in vitro* bromination reaction of $MprE_X$-LCCCWAAA followed by the treatment of the reaction with LysC showing that $MprE_X$-LCCCWAAA is not processed by SrpI.

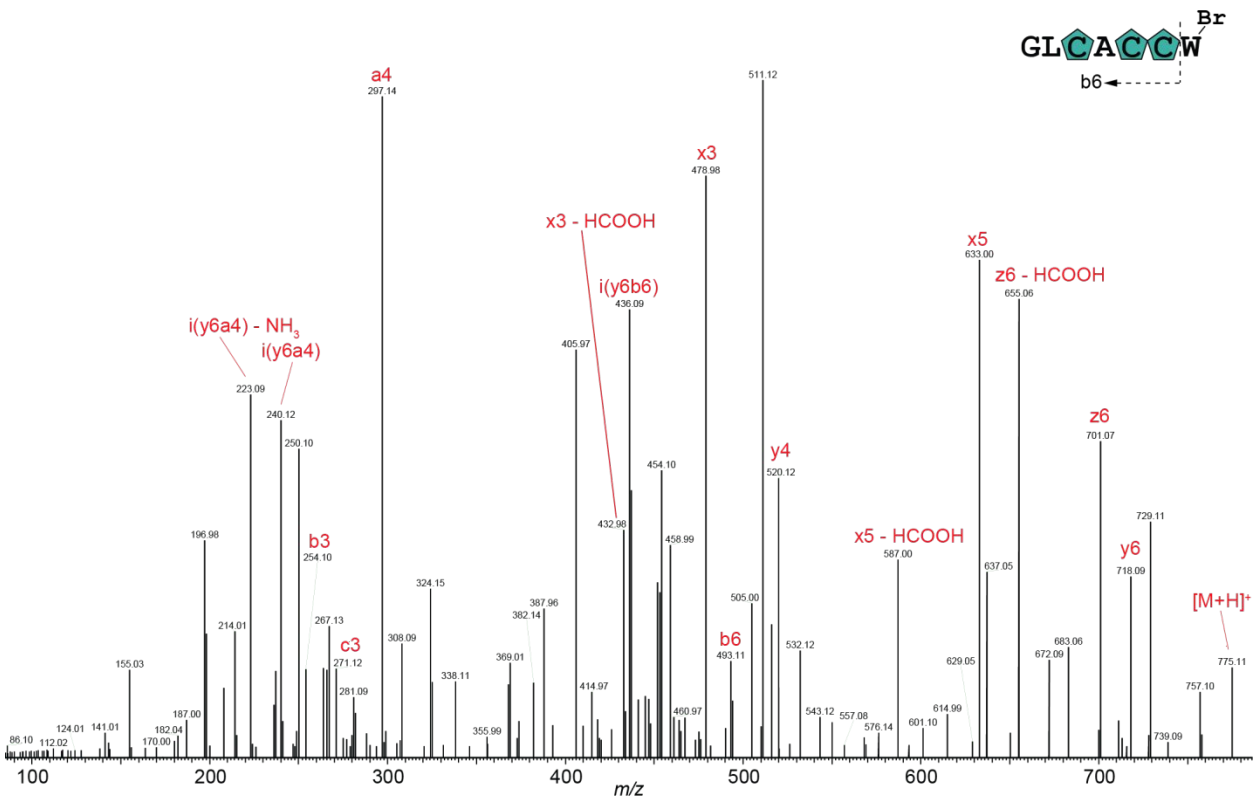
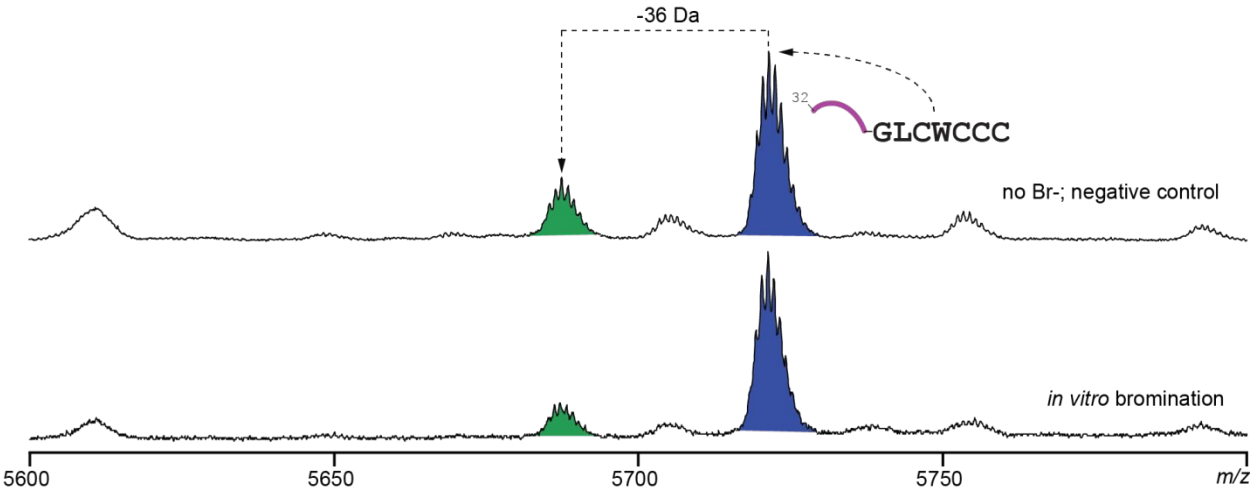


Fig. S11. Structural annotation of the MS^2 fragmentation spectra for the brominated GLCACCW core.



208

209 **Fig. S12.** MALDI-ToF MS spectra for the *in vitro* bromination of MprE$_X$-GLCWCCC followed by the

210 treatment of the reaction with LysC shows that MprE$_X$-GLCWCCC is not processed by SrpI.

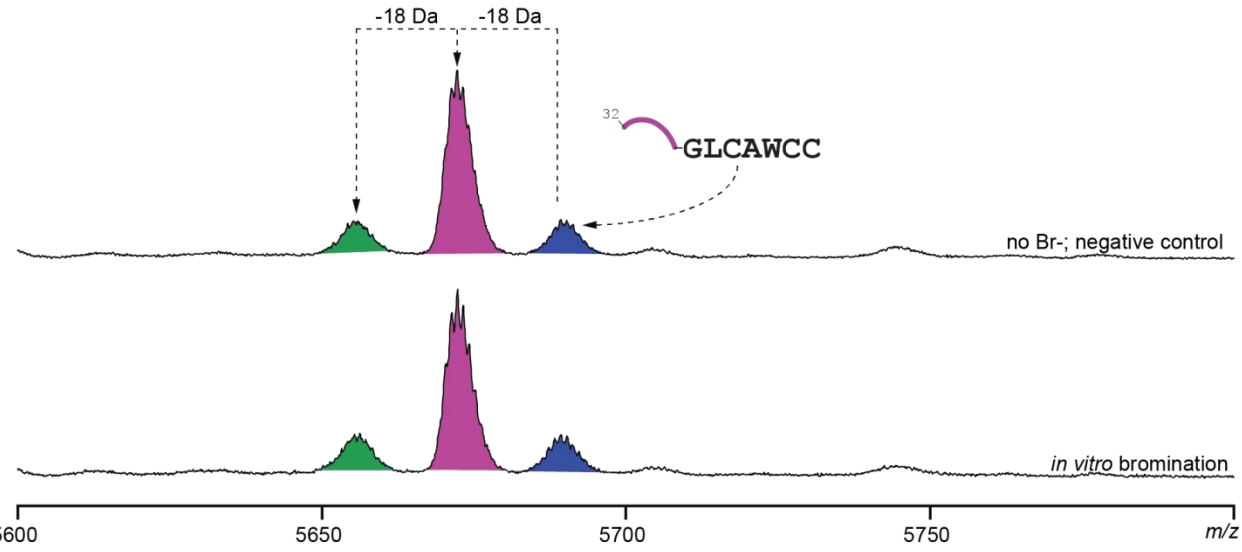
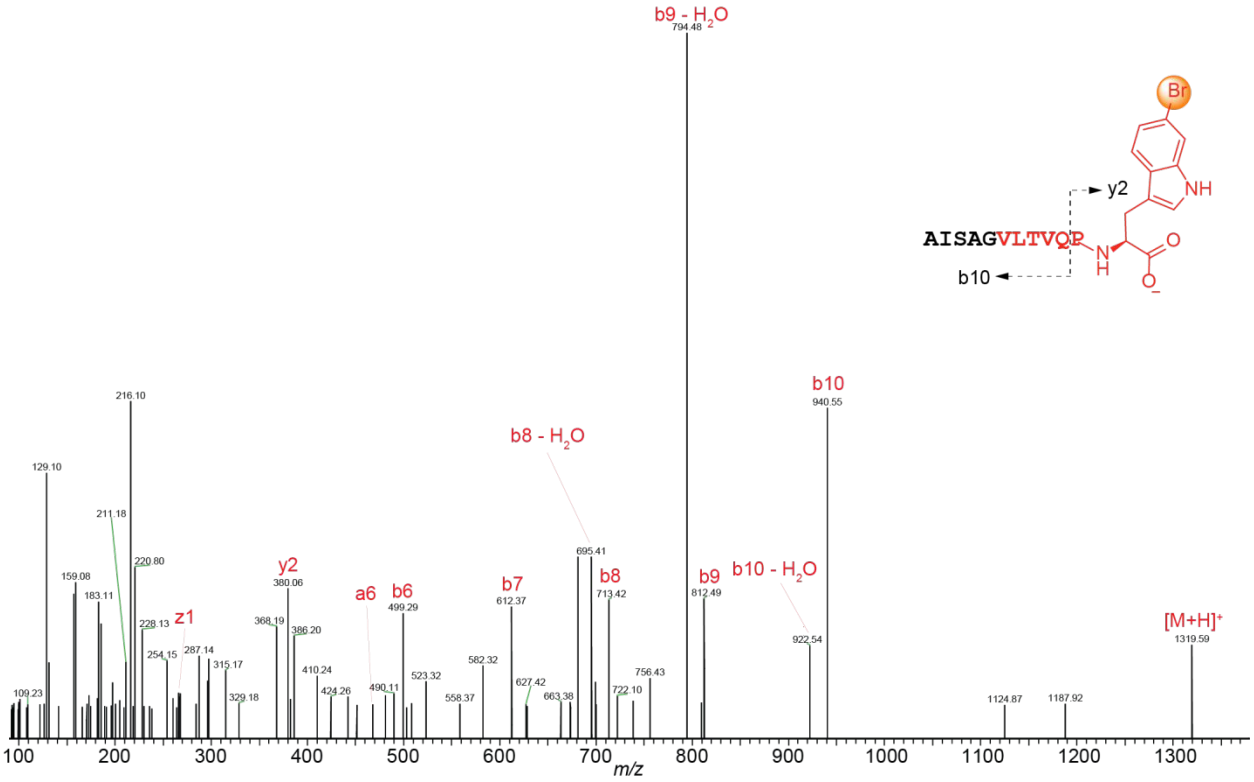


Fig. S13. MALDI-ToF MS spectra for the *in vitro* bromination of $MprE_X$-GLCAWCC followed by the treatment of the reaction with LysC shows that $MprE_X$-GLCAWCC is not processed by SrpI.



214

215 **Fig. S14.** Structural annotation of the MS^2 fragmentation spectra for the C-terminal fragment containing
216 tumor homing peptide sequence TH2 (VLTVQPW) after *in vitro* halogenation by SrpI followed by the
217 treatment of the product with protease GluC. *m/z* cal.: 1319.60, *m/z* obs.: 1319.59.

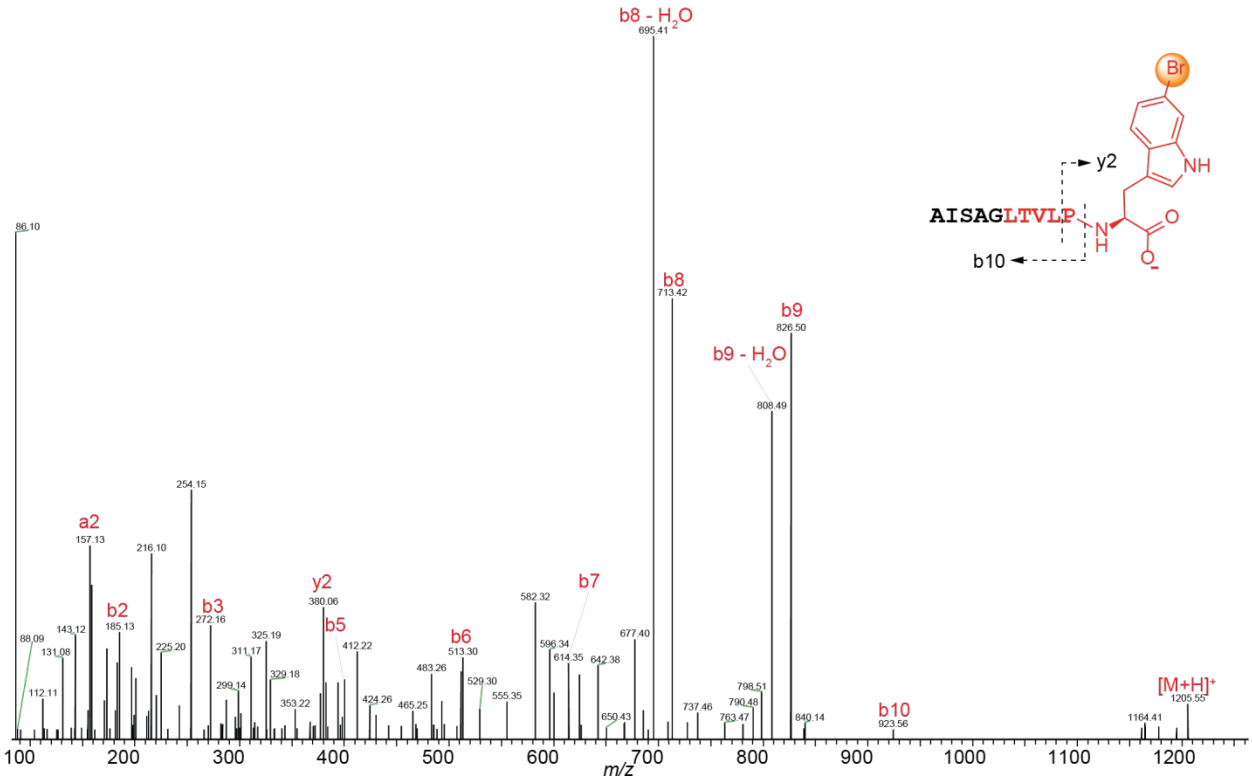


Fig. S15. Structural annotation of the MS^2 fragmentation spectra for the C-terminal fragment containing brominated tumor homing peptide sequence TH1 (LTVLPW) by co-expression of gene encoding SrpE-TH1 with *sprI* followed by the treatment of the purified product with protease GluC. *m/z* cal.: 1205.55, *m/z* obs.: 1205.55

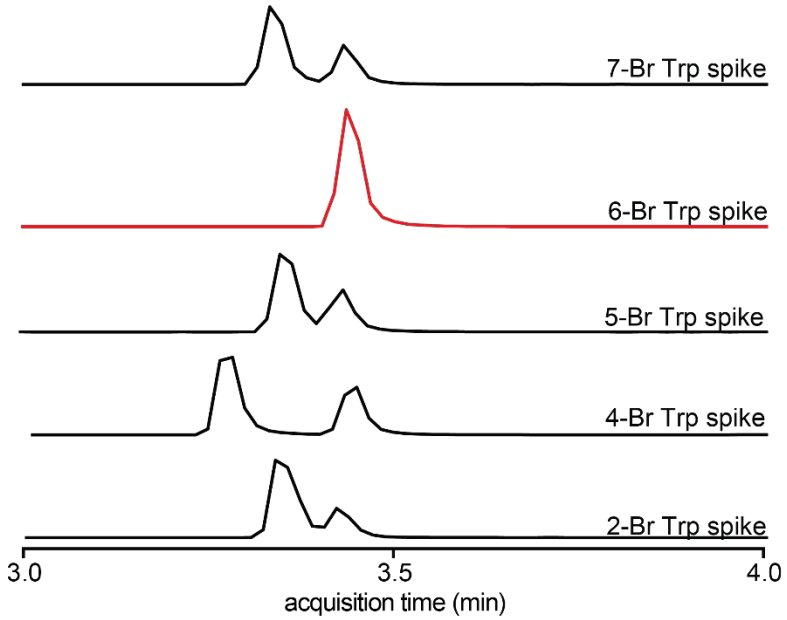


Fig. S16. LC-MS extracted ion chromatograms for *m/z* 283.01 ± 0.1 Da corresponding to bromotryptophan. From bottom to top: co-injection of the carboxypeptidase digestion reaction of brominated SrpE-TH2 with 2-, 4-, 5-, 6-, and 7-bromotryptophan standards. Observation of a single peak upon co-injection with 6-bromotryptophan demonstrates that bromination of the SrpE-TH2 substrate occurred at the indole-6 position.

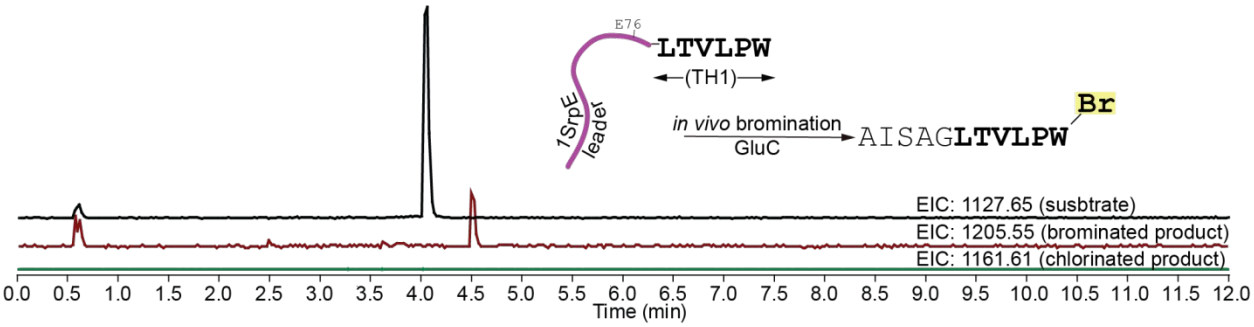


Fig. S17. Scheme for *in vivo* bromination of SrpE-TH1 by SrpI followed by GluC digestion. LC-MS extracted ion chromatograms of the substrate, brominated product, and the conceivable chlorinated product shows that SrpI is strictly specific for incorporation of bromine.

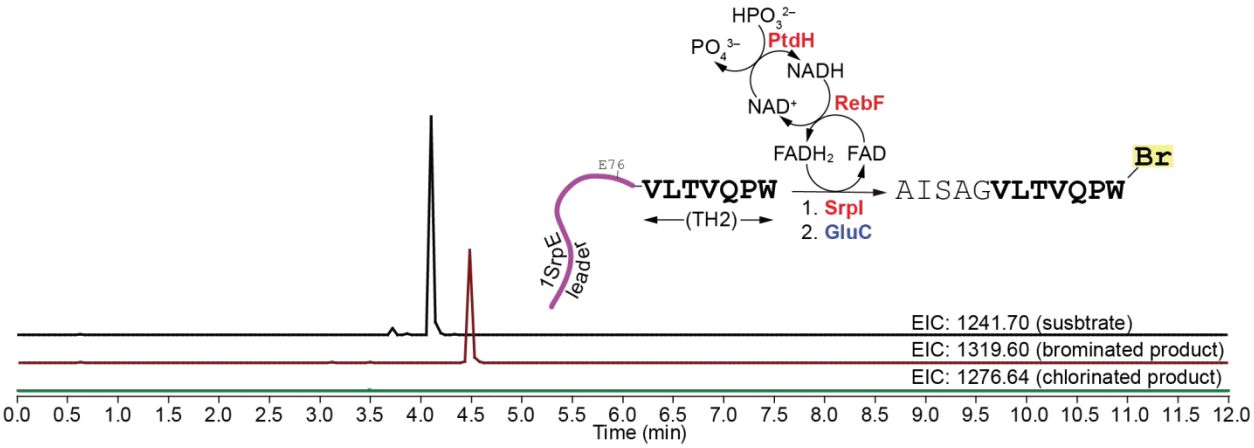
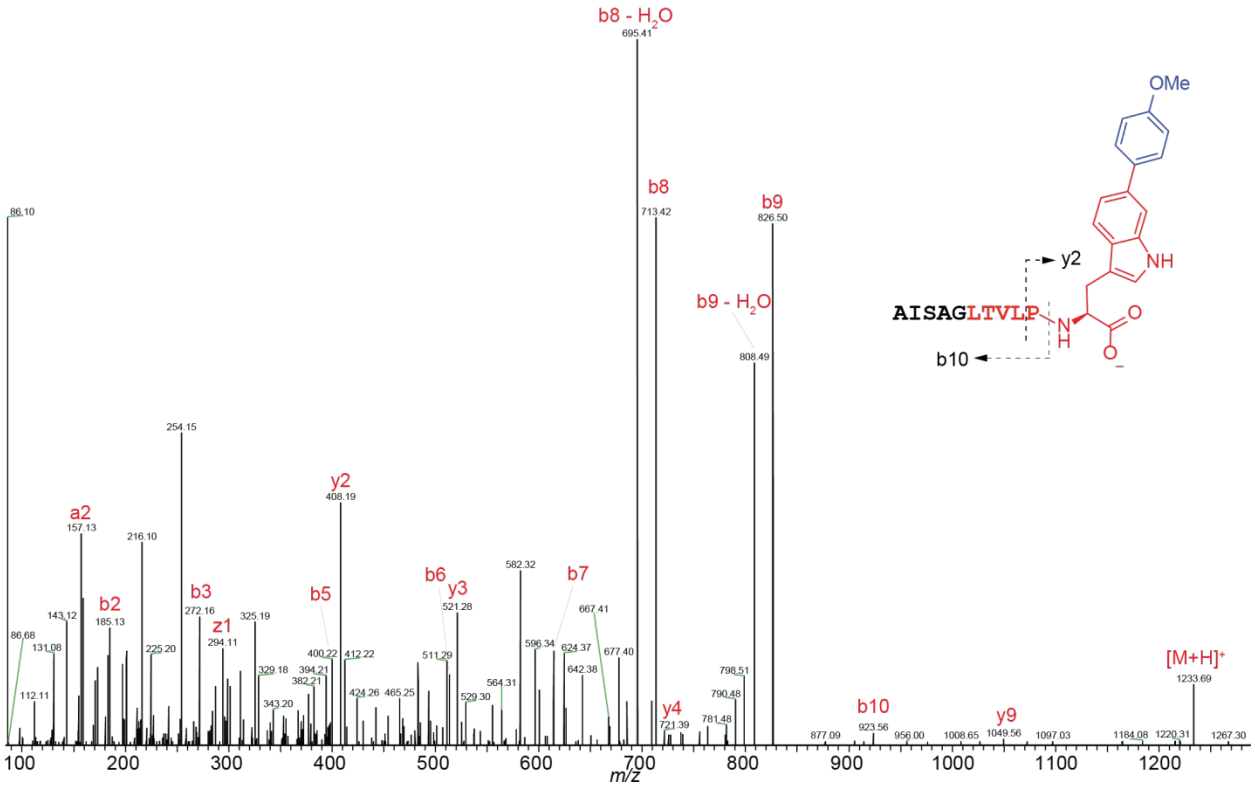


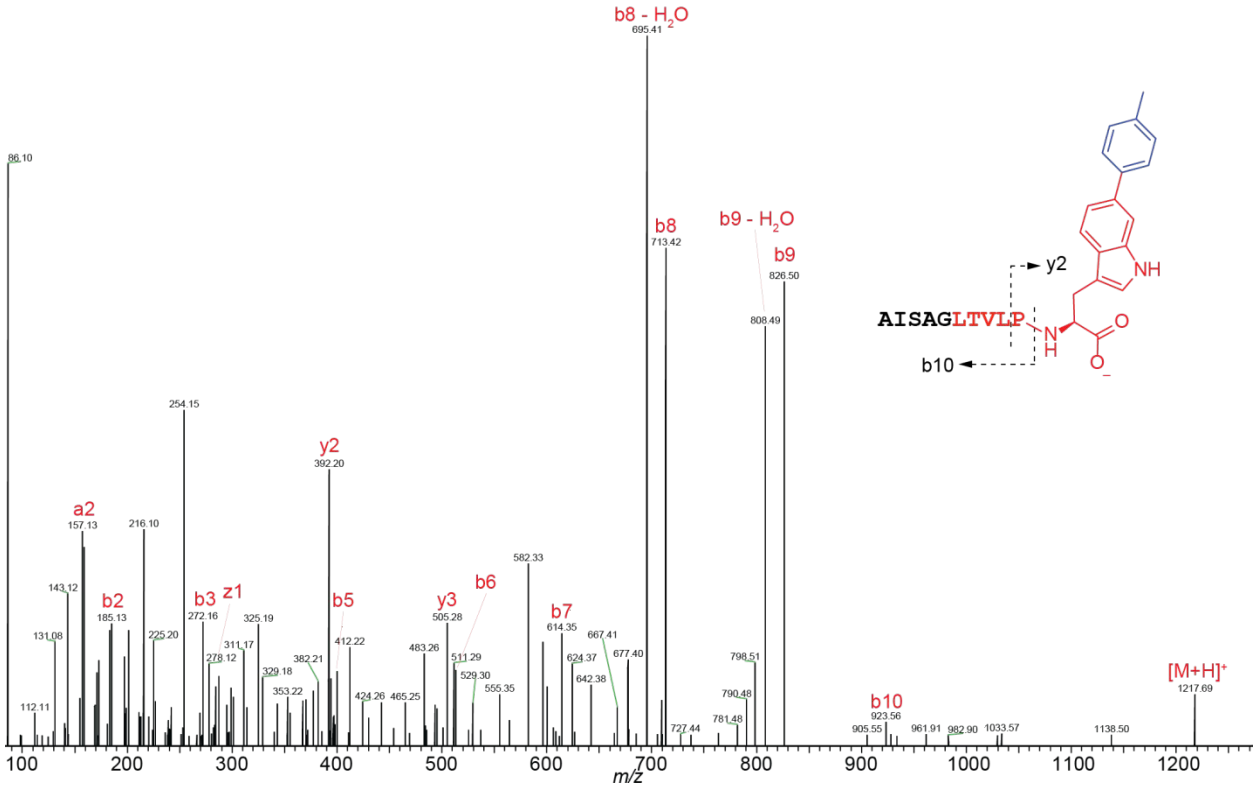
Fig. S18. Reaction scheme for *in vitro* bromination of SrpE-TH2 by SrpI followed by GluC digestion. LC-MS extracted ion chromatograms of the substrate, brominated product, and the conceivable chlorinated product shows that SrpI is strictly specific for incorporation of bromine.



238

239 **Fig. S19.** Structural annotation of the MS^2 fragmentation spectra for TH1-derived Suzuki-Miyaura

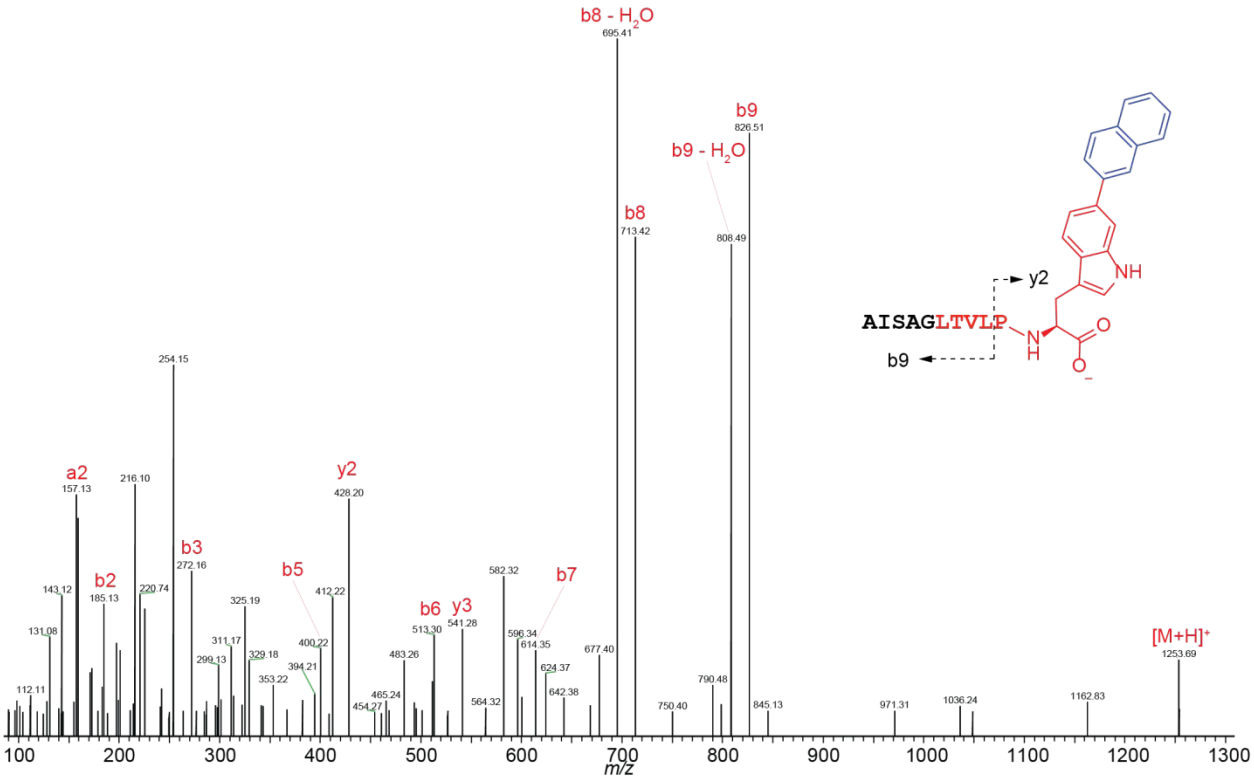
240 coupling product, as illustrated. *m/z* cal.: 1233.69, *m/z* obs.: 1233.69.



241

242 **Fig. S20.** Structural annotation of the MS^2 fragmentation spectra for TH1-derived Suzuki-Miyaura

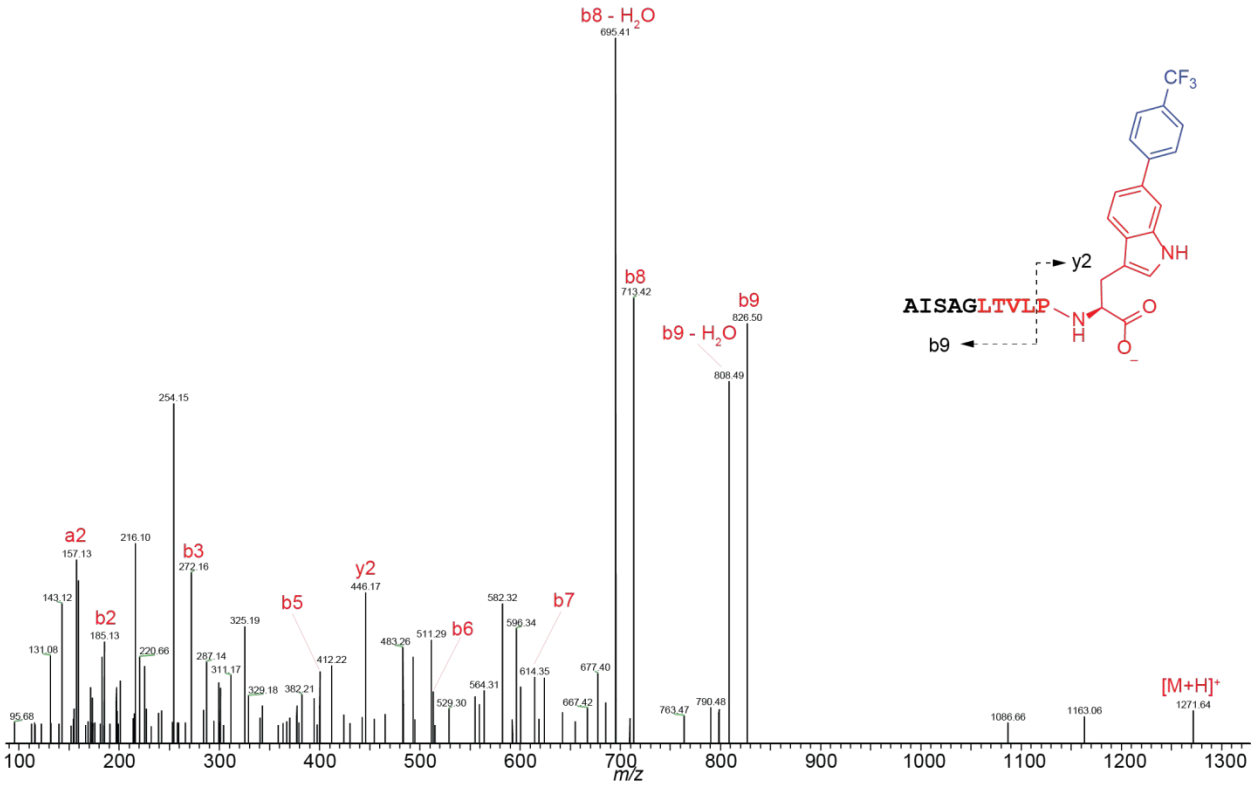
243 coupling product, as illustrated. *m/z* cal.: 1217.69, *m/z* obs.: 1217.69.



244

245 **Fig. S21.** Structural annotation of the MS^2 fragmentation spectra for TH1-derived Suzuki-Miyaura

246 coupling product, as illustrated. *m/z* cal.: 1253.69, *m/z* obs.: 1253.69.



247

248 **Fig. S22.** Structural annotation of the MS^2 fragmentation spectra for TH1-derived Suzuki-Miyaura

249 coupling product, as illustrated. *m/z* cal.: 1271.66, *m/z* obs.: 1271.64.

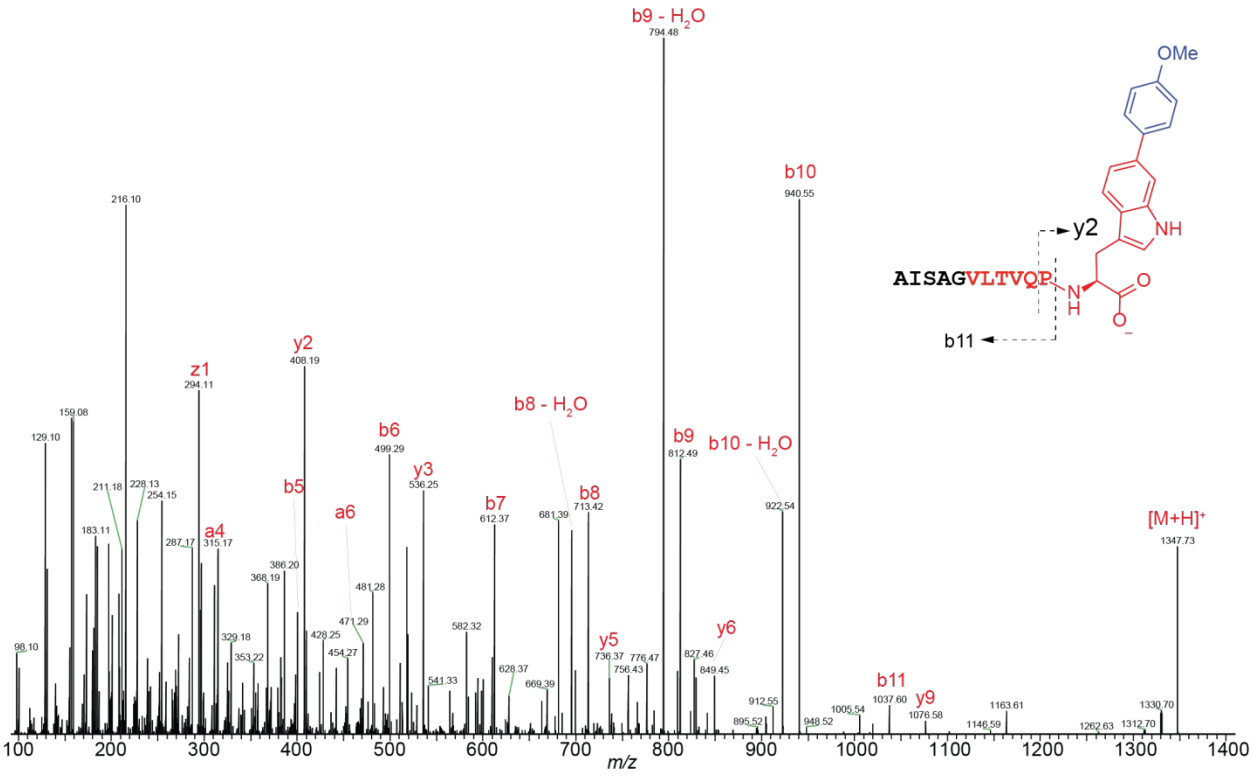
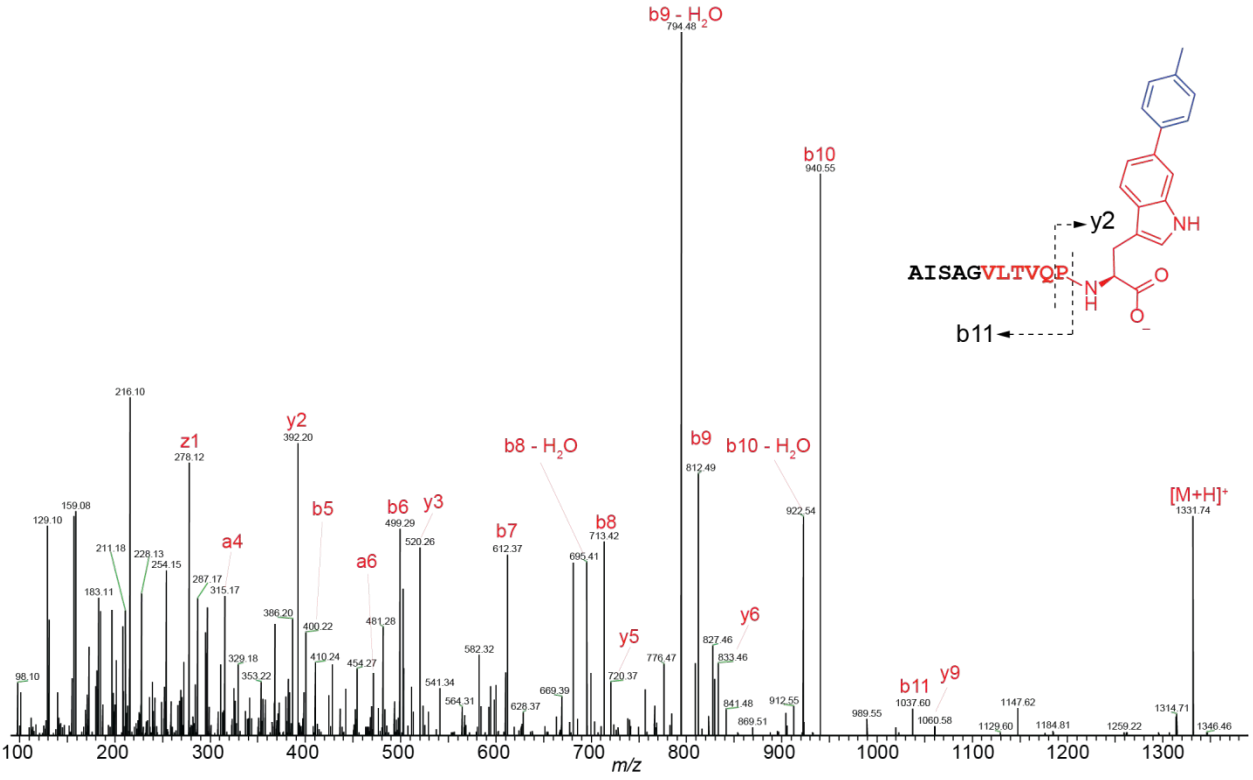


Fig. S23. Structural annotation of the MS^2 fragmentation spectra for TH2-derived Suzuki-Miyaura coupling product, as illustrated. *m/z* cal.: 1347.73, *m/z* obs.: 1347.73.



253

254 **Fig. S24.** Structural annotation of the MS^2 fragmentation spectra for TH2-derived Suzuki-Miyaura

255 coupling product, as illustrated. *m/z* cal.: 1331.74, *m/z* obs.: 1331.74.

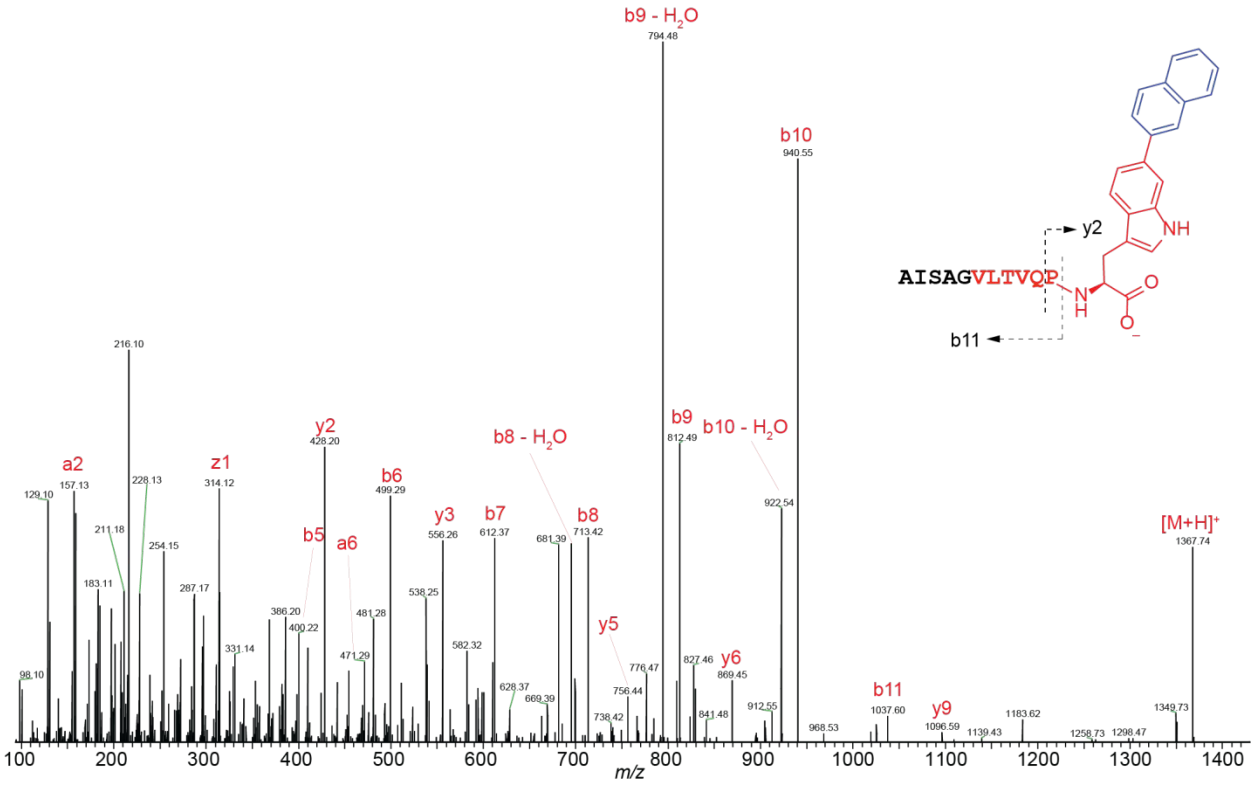


Fig. S25. Structural annotation of the MS^2 fragmentation spectra for TH2-derived Suzuki-Miyaura coupling product, as illustrated. *m/z* cal.: 1367.74, *m/z* obs.: 1367.74.

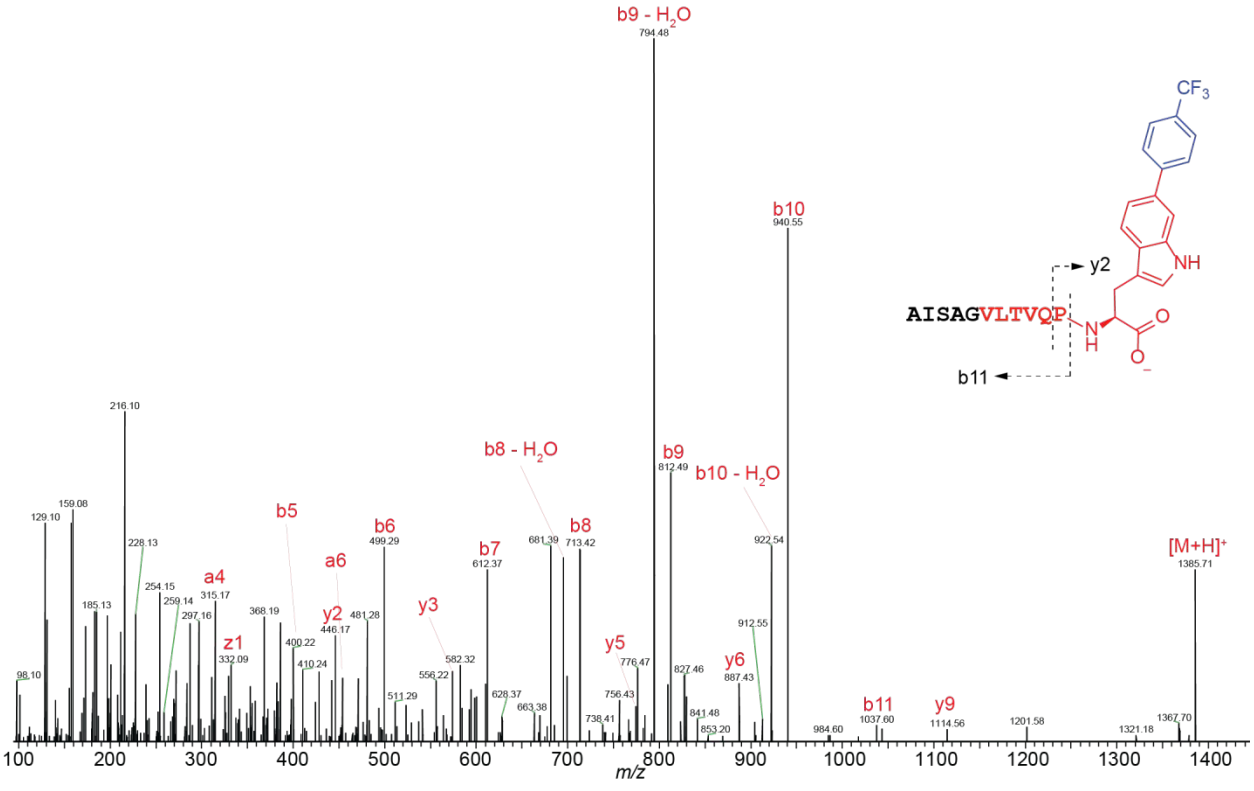


Fig. S26. Structural annotation of the MS^2 fragmentation spectra for TH2-derived Suzuki-Miyaura coupling product, as illustrated. *m/z* cal.: 1385.70, *m/z* obs.: 1385.71.

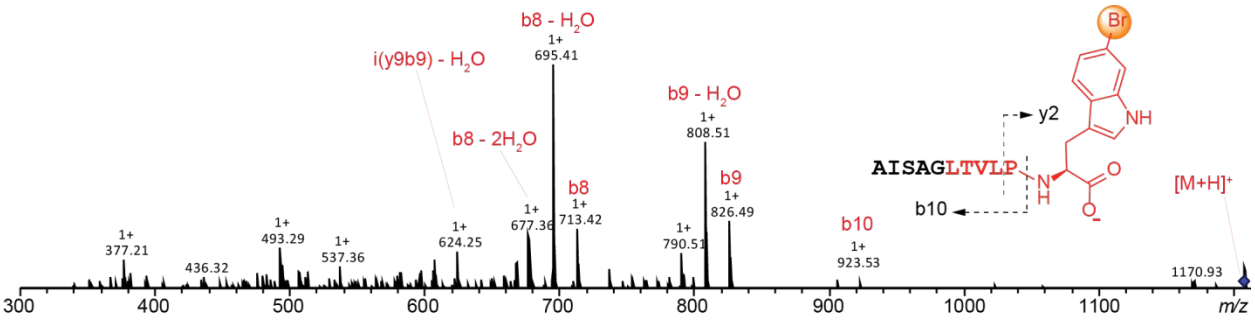


Fig. S27. Structural annotation of the MS^2 fragmentation spectra for the brominated C-terminal fragment of the MBP-SrpE-TH1 substrate delivered by digestion by GluC protease. *m/z* cal.: 1205.55, *m/z* obs.: 1205.55.

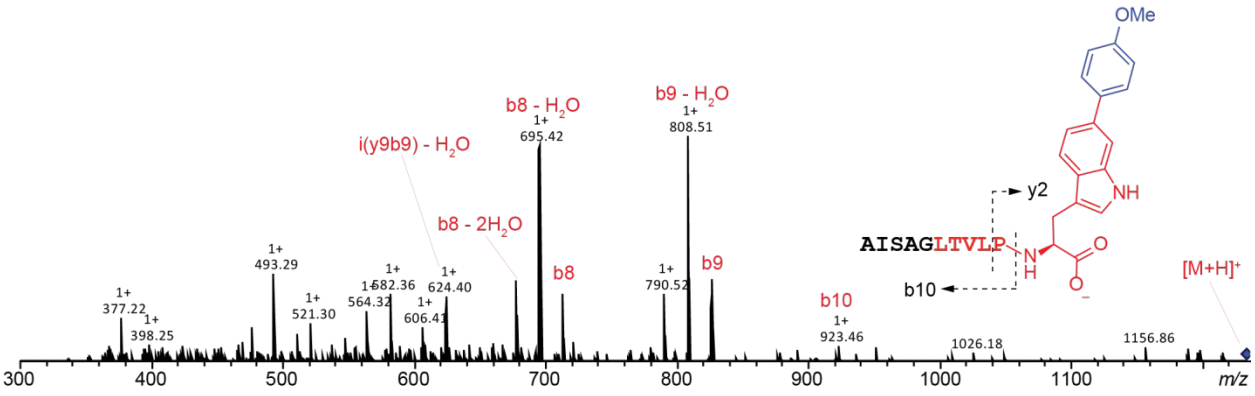


Fig. S28. Structural annotation of the MS^2 fragmentation spectra for the modified C-terminal fragment of the MBP-SrpE-TH1 substrate delivered by digestion by GluC protease. *m/z* cal.: 1233.69, *m/z* obs.: 1233.69.

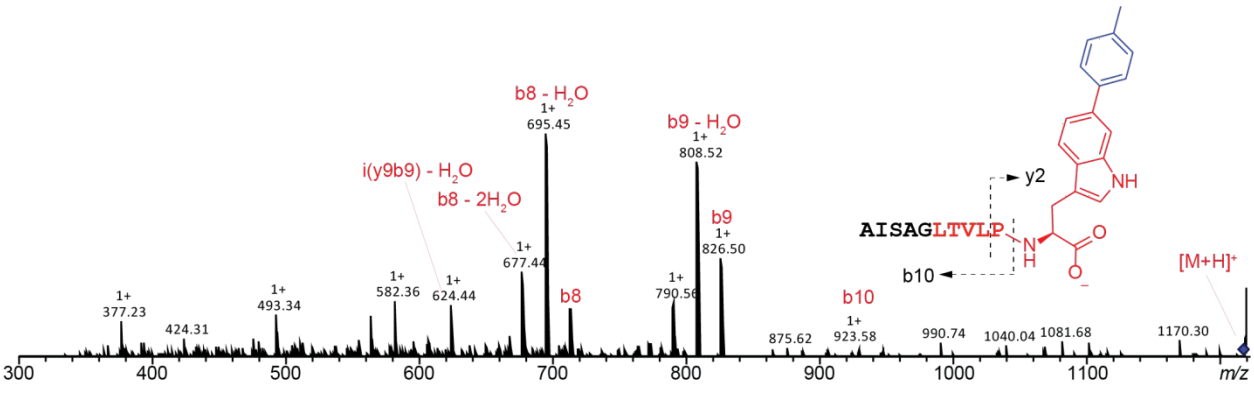
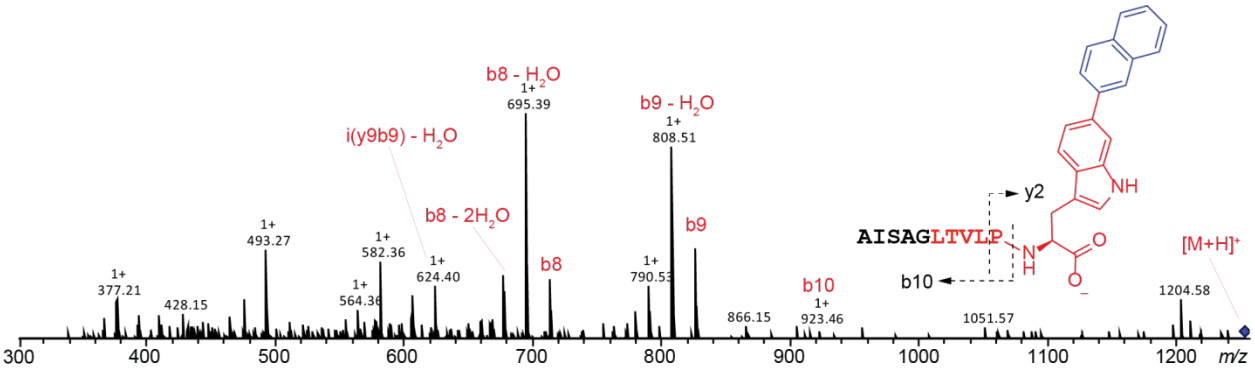
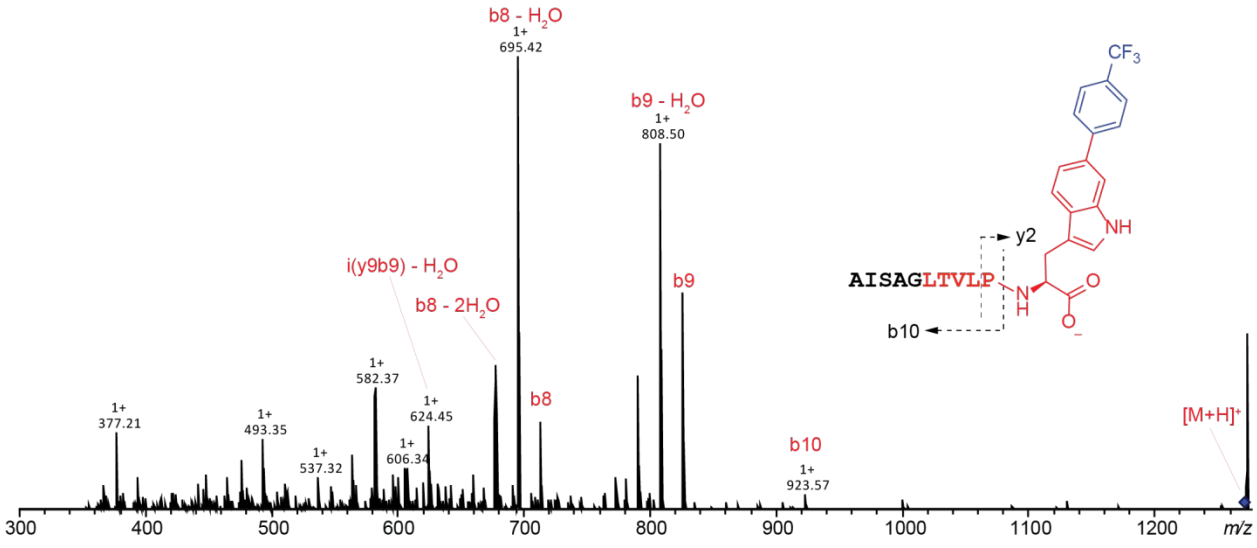


Fig. S29. Structural annotation of the MS^2 fragmentation spectra for the modified C-terminal fragment of the MBP-SrpE-TH1 substrate delivered by digestion by GluC protease. *m/z* cal.: 1217.69, *m/z* obs.: 1217.69.



274

275 **Fig. S30.** Structural annotation of the MS^2 fragmentation spectra for the modified C-terminal fragment of
276 the MBP-SrpE-TH1 substrate delivered by digestion by GluC protease. *m/z* cal.: 1253.69, *m/z* obs.:
277 1253.69.



278

279 **Fig. S31.** Structural annotation of the MS^2 fragmentation spectra for the modified C-terminal fragment of

280 the MBP-SrpE-TH1 substrate delivered by digestion by GluC protease. *m/z* cal.: 1271.66, *m/z* obs.:

281 1271.64.

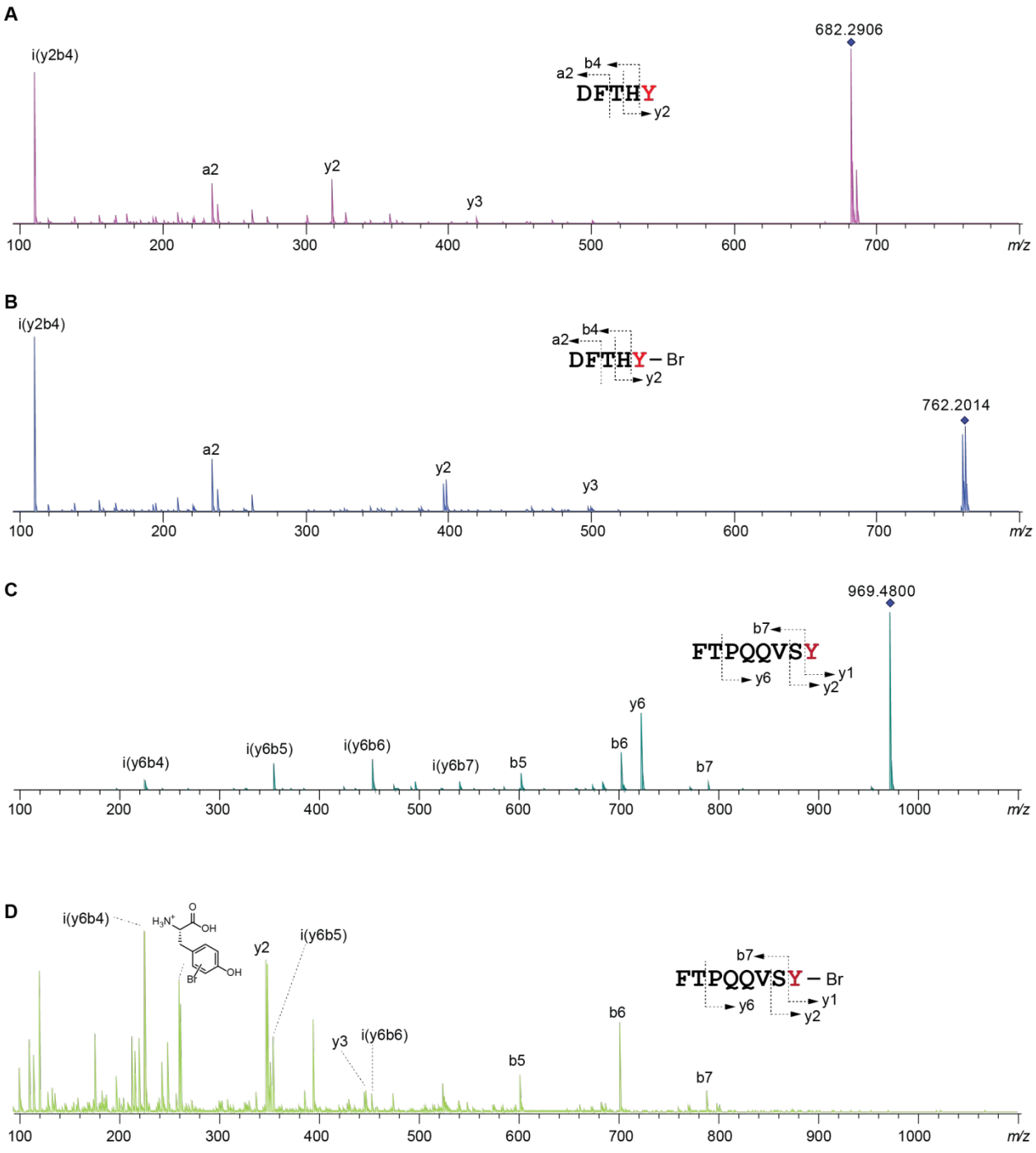


Fig. S32. LC-MS analysis of self-bromination assay products. MS^2 spectrum of the SrpI fragment (**A**) DFTHY and (**B**) the corresponding brominated fragment at *m/z* 682.283 and *m/z* 762.194, respectively. MS^2 spectrum of (**C**) FTPQQVSY and (**D**) the corresponding brominated fragment at *m/z* 969.467 and *m/z* 1047.378, respectively.

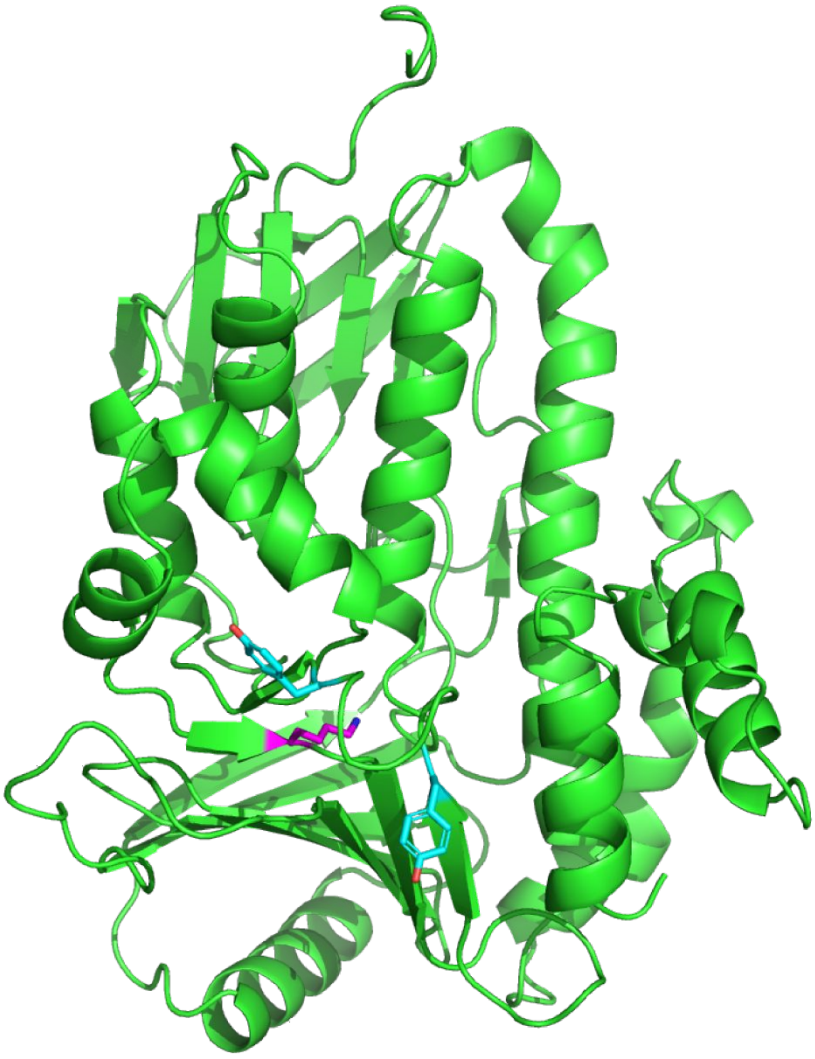
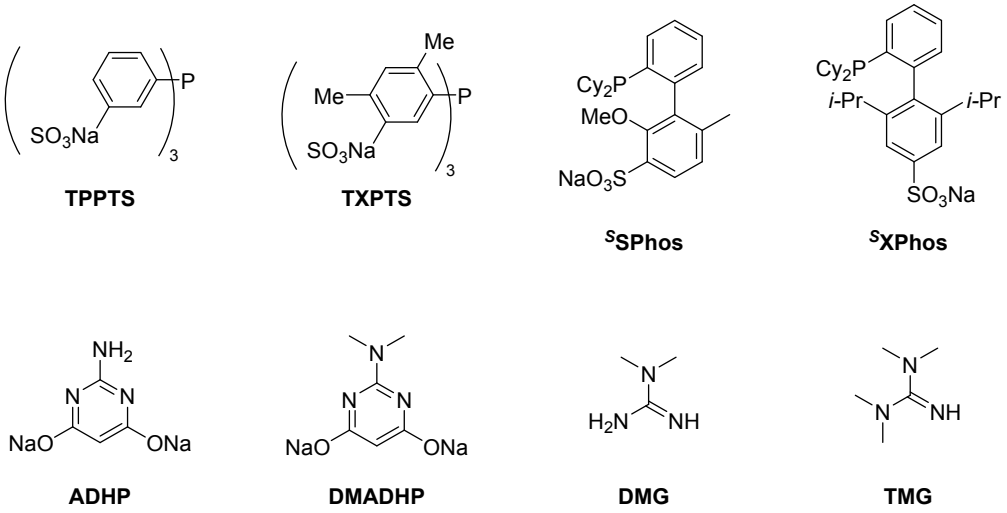


Figure S33. Structural model of SrpI generated by AlphaFold. Catalytic Lys residue is shown in magneta. Two Tyr residues which are self-brominated are shown in cyan.



291

292 **Fig. S34.** Selected water-soluble ligands used for aqueous palladium-catalyzed cross coupling reactions.

SUPPLEMENTARY REFERENCES

1. Johannes, T. W.; Woodyer, R. D.; Zhao, H., Efficient regeneration of NADPH using an engineered phosphite dehydrogenase. *Biotechnol. Bioeng.* **2007,** *96* (1), 18-26.

2. Yeh, E.; Garneau, S.; Walsh, C. T., Robust *in vitro* activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebeccamycin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102* (11), 3960-3965.

3. Nguyen, N. A.; Lin, Z.; Mohanty, I.; Garg, N.; Schmidt, E. W.; Agarwal, V., An Obligate Peptidyl Brominase Underlies the Discovery of Highly Distributed Biosynthetic Gene Clusters in Marine Sponge Microbiomes. *J. Am. Chem. Soc.* **2021,** *143* (27), 10221-10231.

4. Western, E. C.; Daft, J. R.; Johnson, E. M., 2nd; Gannett, P. M.; Shaughnessy, K. H., Efficient one-step Suzuki arylation of unprotected halonucleosides, using water-soluble palladium catalysts. *J. Org. Chem.* **2003,** *68* (17), 6767-6774.

5. Anderson, K. W.; Buchwald, S. L., General catalysts for the Suzuki-Miyaura and Sonogashira coupling reactions of aryl chlorides and for the coupling of challenging substrate combinations in water. *Angew. Chem., Int. Ed. Engl.* **2005,** *44* (38), 6173-6177.

6. Li, J.-H.; Zhang, X.-D.; Xie, Y.-X., Efficient and Copper-Free Sonogashira Cross-Coupling Reaction Catalyzed by Pd(OAc)2/Pyrimidines Catalytic System. *Eur. J. Org. Chem.* **2005,** *2005* (20), 4256-4259.

7. Gao, Z.; Gouverneur, V.; Davis, B. G., Enhanced Aqueous Suzuki–Miyaura Coupling Allows Site-Specific Polypeptide 18F-Labeling. *J. Am. Chem. Soc.* **2013,** *135* (37), 13612-13615.

8. Dumas, A.; Spicer, C. D.; Gao, Z.; Takehana, T.; Lin, Y. A.; Yasukohchi, T.; Davis, B. G., Self-liganded Suzuki-Miyaura coupling for site-selective protein PEGylation. *Angew. Chem., Int. Ed. Engl.* **2013,** *52* (14), 3916-21.

9. Spicer, C. D.; Davis, B. G., Palladium-mediated site-selective Suzuki–Miyaura protein modification at genetically encoded aryl halides. *Chem. Comm.* **2011,** *47* (6), 1698-1700.

10. Li, N.; Lim, R. K. V.; Edwardraja, S.; Lin, Q., Copper-Free Sonogashira Cross-Coupling for Functionalization of Alkyne-Encoded Proteins in Aqueous Medium and in Bacterial Cells. *J. Am. Chem. Soc.* **2011,** *133* (39), 15316-15319.

11. Li, S.; Lin, Y.; Cao, J.; Zhang, S., Guanidine/Pd(OAc)2-Catalyzed Room Temperature Suzuki Cross-Coupling Reaction in Aqueous Media under Aerobic Conditions. *J. Org. Chem.* **2007,** *72* (11), 4067-4072.

12. Latham, J.; Henry, J. M.; Sharif, H. H.; Menon, B. R.; Shepherd, S. A.; Greaney, M. F.; Micklefield, J., Integrated catalysis opens new arylation pathways via regiodivergent enzymatic C-H activation. *Nat. Commun.* **2016,** *7*, 11873.