Supporting Information for:

Reconstitution and substrate specificity of the radical SAM thiazole *C*-methyltransferase in thiomuracin biosynthesis

Nilkamal Mahanta,[†,‡] Zhengan Zhang,[†] Graham A. Hudson,[†] Wilfred A. van der Donk[†,‡,§*] and Douglas A. Mitchell[†,‡*]

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

* Corresponding authors:
Wilfred A. van der Donk (vddonk@illinois.edu), phone: 1-217-244-5360, fax: 1-217-244-8533
Douglas A. Mitchell (douglasm@illinois.edu), phone: 1-217-333-1345, fax: 1-217-333-0508

Table of Contents:

Experimental Methods.

General materials and methods. Reagents used for molecular biology experiments were purchased from New England BioLabs (Ipswich, MA), Thermo Fisher Scientific (Waltham, MA), or Gold Biotechnology Inc. (St. Louis, MO). Other chemicals were purchased from Sigma-Aldrich (St. Louis, MO). *Escherichia coli* DH5α and BL21(DE3) strains were used for plasmid maintenance and protein overexpression, respectively. Plasmid inserts were sequenced at the University of Illinois core sequencing facility and ACGT Inc. (Wheeling, IL). MALDI-TOF-MS analysis was performed using a Bruker UltrafleXtreme matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer (Bruker Daltonics) in reflector positive mode at the University of Illinois School of Chemical Sciences Mass Spectrometry Laboratory. ESI-MS/MS analyses were performed using a SYNAPT electrospray ionization (ESI) quadrupole TOF Mass Spectrometry System (Waters) equipped with an ACQUITY Ultra Performance Liquid Chromatography (UPLC) system (Waters). HiTrap columns for Ni-NTA affinity chromatography were purchased from GE Healthcare.

Molecular biology techniques. Oligonucleotides were purchased from Integrated DNA Technologies Inc. (Coralville, IA). Genes optimized for recombinant expression in *E. coli* were synthesized by GenScript (Piscataway, NJ) in pUC57 (kanamycin, Kan) vectors with BamHI and XhoI sites flanking each gene at the 5' and 3' ends, respectively. The GenBank locus tag and *E. coli* optimized sequence for *tbtI* is provided in **Sequence S1**. *E. coli* DH5α were transformed with pUC57-Kan vectors containing each gene for replication and subsequent isolation using a QIAprep Spin Miniprep Kit (Qiagen). The isolated DNA was then treated with BamHI-HF and XhoI-HF (New England Biolabs; NEB). The digested genes were separated on a 1% (*w/v*) agarose gel, purified using a QIAQuick gel extraction kit (Qiagen), and ligated into an appropriately endonuclease-digested and gel-purified pET28 vector with maltose binding protein (MBP) 5' to the multiple cloning site using T4 DNA ligase (NEB). Ligation reactions were used to transform chemically competent DH5α cells, which were plated on Luria-Bertani (LB) agar plates containing 50 μg/mL kanamycin and grown at 37 °C. Colonies were picked at random and grown in LB broth for 16–20 h prior to plasmid isolation using a QIAprep Spin Miniprep Kit. The construct encoding for *tbtA* with either one Cys at position 4 or two Cys at positions 2 and 4 of the core as well as all Cys replaced with Ala was previously described and the latter was used for downstream site-directed mutagenesis (SDM).[1] SDM was performed using the QuikChange method (Agilent) as per the manufacturer's instructions using PfuTurbo DNA polymerase. The primers for each mutant are listed in **Table S1**. The mutations were verified by sequencing using a custom MBP forward primer or the T7 reverse primer (**Table S1**).

MBP-tagged TbtI overexpression. *E. coli* BL21(DE3) cells were co-transformed with a pET28 plasmid encoding the MBP-tagged TbtI and the pSUF plasmid encoding the Fe-S biosynthetic operon.[2] Cells were grown for 24 h on LB agar plates containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol at 37 °C. Single colonies were used to inoculate 10 mL of LB containing the same concentration of antibiotics and grown at 37 °C for 16–18 h. This culture was used to inoculate 1 L of LB (5 g/L yeast extract, 10 g/L tryptone and 10 g/L NaCl) supplemented with the same concentration of antibiotics and grown to an optical density at 600 nm (OD_{600}) of 0.6 before being placed on ice for 15 min. Protein expression was then induced with the addition of 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and supplemented with 100 mg/L of ferrous ammonium sulfate ($Fe(NH_4)_2(SO_4)_2$) and 100 mg/L of cysteine. Expression was allowed to proceed for 12–16 h at 15 °C. Cells were harvested by centrifugation at 3,000 × *g* for 20 min, washed with phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4), and harvested by centrifugation. The cells were flash-frozen and stored at -80 °C for a maximum of one week before use.

MBP-tagged TbtI purification. Protein purification was performed in a Coy anaerobic chamber. All buffers were degassed and stored for 24–48 h in the anaerobic chamber before use. Cells were resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 2.5% glycerol (*v/v*), and 0.1% Triton X-100 (*v/v*)) containing 4 mg/mL lysozyme, 2 μM leupeptin, 2 μM benzamidine, and 2 μM E64. Cells were further lysed by sonication (3 × 45 s with 10 min agitation periods at 4 °C). Insoluble debris was removed by centrifugation at 20,000 × *g* for 40 min at 4 °C. The supernatant was then applied to a lysis-buffer pre-equilibrated amylose resin (NEB; 15 mL of resin per L of initial cell culture). The column was washed with 10 column volumes (CV) of lysis buffer followed by 10 CV of wash buffer (lysis buffer with 400 mM NaCl and lacking Triton X-100). The MBP-tagged proteins were eluted using elution buffer (lysis buffer with 300 mM NaCl, 10 mM maltose, but omitting Triton X-100) until the eluent

no longer contained protein detectable with the Bradford reagent. Eluent was concentrated using a 30 kDa molecular weight cut-off (MWCO) Amicon Ultra centrifugal filter (EMD Millipore). A buffer exchange with 10× volume of protein storage buffer [50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 300 mM NaCl, 2.5% glycerol (*v/v*)] was performed prior to final concentration, freezing in liquid nitrogen, and storage at -80 °C. Protein concentrations were determined using both absorbance at 280 nm (theoretical extinction coefficients were calculated using the ExPASy ProtParam tool; *http://web.expasy.org/protparam/protpar-ref.html*) and a Bradford colorimetric assay. Purity was assessed visually by analysis of Coomassie-stained SDS-PAGE gels (**Figure S1**). All wash, elution, and storage buffers were supplemented with 0.5 mM tris-(2-carboxyethyl)-phosphine (TCEP).

His_6-tagged cysteine desulfurase (IscS) overexpression. *E. coli* BL21(DE3) cells were co-transformed with a pET15 plasmid encoding His_6-tagged IscS. Cells were grown for 24 h on LB agar plates containing 100 µg/mL ampicillin at 37 °C. Single colonies were used to inoculate 10 mL of LB containing 100 µg/mL ampicillin and grown at 37 °C for 16–18 h. This culture was used to inoculate 1 L of LB (5 g/L yeast extract, 10 g/L tryptone and 10 g/L NaCl) and grown to an optical density at 600 nm (OD_{600}) of 0.6 before being placed on ice for 15 min. Protein expression was then induced with the addition of 0.5 mM IPTG. Expression was allowed to proceed for 12–16 h at 15 °C. Cells were harvested by centrifugation at 3,000 × *g* for 20 min, washed with PBS and harvested by centrifugation. The cells were flash-frozen and stored at -80 °C for a maximum of one week before use.

His_6-tagged IscS purification. Cells were resuspended in lysis buffer (20 mM Na_2HPO_4, 500 mM NaCl, 10 mM imidazole, pH 7.4, 2.5% glycerol (v/v), and 0.1% Triton X-100) containing 4 mg/mL lysozyme, 2 µM leupeptin, 2 µM benzamidine, and 2 µM E64. Cells were further lysed by sonication (3 × 45 s with 10 min agitation periods at 4 °C. Insoluble debris was removed by centrifugation at 20,000 × *g* for 40 min. The supernatant was then applied to a pre-equilibrated Ni-NTA resin (Thermo Scientific; 15 mL of resin per L of cells). The column was washed with 10 column volumes (CV) of lysis buffer followed by 10 CV of wash buffer (lysis buffer with 30 mM imidazole and lacking Triton X-100). The His_6-tagged proteins were eluted using elution buffer (lysis buffer with 300 mM NaCl, 300 mM imidazole and lacking Triton X-100) until the eluent no longer contained protein detectable with the Bradford reagent. Eluent was concentrated and stored and the protein concentration was assayed as described for TbtI (*vide supra*). Purity and possible truncation were assessed by Coomassie-stained SDS-PAGE gel (**Figure S1**). All wash, elution, and storage buffers were supplemented with 0.5 mM tris-(2-carboxyethyl)-phosphine (TCEP).

In vitro reconstitution of the Fe-S cluster in TbtI. TbtI (50 µM) was equilibrated with 1 mM ferrous ammonium sulfate [$Fe(NH_4)_2(SO_4)_2$], cysteine (2 mM), dithiothreitol (DTT, 2 mM) and IscS (2 µM) in 50 mM Tris-HCl pH 7.5 buffer in the anaerobic chamber. The reaction mixture was buffer exchanged using a 10 kDa MWCO gel filtration column (EMD Millipore) before concentration and use for assays. The reactions were allowed to proceed for 4–6 h.

Ferrozine assay for TbtI iron quantification. A previously published procedure was followed.[3] Briefly, the following aqueous reagents were prepared: 8 M guanidine hydrochloride, 2 M HCl, 10 mM ferrozine (3-(2-pridyl)-5,6-diphenyl-1,2,4-triazine-*p,p'*-disulfonic acid monosodium salt hydrate), and 100 mM L-ascorbic acid. To 100 µL of TbtI (0-100 µM), 100 µL of 8 M guanidine hydrochloride and 100 µL of 2 M HCl were added and the solution was diluted to 550 µL. Insoluble debris was removed by centrifugation and to 500 µL of the supernatant, 30 µL of the 10 mM ferrozine solution and 30 µL of 100 mM freshly prepared L-ascorbic acid were added. The resulting solution was mixed thoroughly and incubated at RT for 30 min. The absorbance of the iron-ferrozine complex was recorded at 562 nm using a Cary 4000 UV-Vis-NIR spectrophotometer (Agilent). The Fe content was determined by comparing this reading to a standard curve that was generated under identical conditions using ferrous ammonium sulfate ($Fe(NH_4)_2(SO_4)_2$) with a concentration range from 0 to 100 µM (**Figure S2**).

Methylene blue assay for TbtI sulfide quantitation. A previously published procedure was followed.[4] Briefly, to 300 µL of assay solution containing TbtI (0–100 µM), 1 mL of 1% (*w/v*) zinc acetate was added followed by 50 µL of 3 M NaOH. This mixture was agitated gently and 250 µL of 0.1 % *N, N*-dimethyl-*p*-phenylenediamine (DMPD) monohydrochloride in 5 M HCl and 50 µL of 23 mM $FeCl_3$ in 1.2 M HCl were added. The resulting solution was mixed vigorously for 5 min intervals for a total of 30 min. The samples were then centrifuged at 16,000 × *g* for 5

min at 25 °C. The supernatant was collected and the absorbance at 670 nm was recorded using a Cary 4000 UV-Vis-NIR spectrophotometer (Agilent). The sulfide content was determined by comparing the reading to a standard curve that was generated under identical conditions using a fresh solution of sodium sulfide (Na_2S) in 0.1 M NaOH with a concentration range of 0–100 μM (**Figure S3**).

Initial Fe-S quantitation of the as-isolated protein using these assays yielded 1.6 ± 0.2 Fe and 1.9 ± 0.3 sulfides per monomer of TbtI. In vitro reconstitution using IscS, Cys, Fe(II) and DTT increased the Fe and S content to 3.7 ± 0.3 Fe and 4.1 ± 0.3 sulfides, near the expected values of a single [4Fe-4S] cluster. Indeed, the primary sequence of TbtI exhibits one CX_3CX_2C motif at the N-terminus (Sequence S1B), consistent with the incorporation of one [4Fe-4S] cluster. UV-Vis analysis also supports the proper incorporation of the Fe-S cluster (Figure S3).

Overexpression and purification of MBP-tagged TbtA precursor peptide variants. MBP-tagged variants of the TbtA precursor peptide were overexpressed and purified using affinity chromatography as previously described.[5]

Cyclodehydration of MBP-tagged TbtA precursor peptide variants. Thiazole installation on TbtA precursor peptide variants was carried out as previously described.[5] Reactions were monitored using MALDI-TOF-MS.

In vivo cyclodehydration and purification of TbtA variants. Production of cyclodehydrated TbtA variants was performed in *E. coli* as described previously.[1] Products were analyzed by MALDI-TOF-MS.

HPLC purification of TbtA-hexazole core peptide and variants. HPLC purification of TbtA hexazole core peptide prepared by GluC endoproteinase treatment of the full length hexazole intermediate as well as its variants were performed as described previously.[1]

In vitro methylation of TbtA by TbtI. Reaction mixtures generally included the following components: substrate (50 μM), purified TbtI from *E. coli* heterologous expression after in vitro reconstitution of the FeS cluster (2 μM), S-adenosylmethionine (SAM, 1 mM) and sodium dithionite (5 mM) in reaction buffer (50 mM Tris-HCl pH 7.5). The reaction was allowed to proceed for 6–12 h at 25 °C in an anaerobic chamber. Reaction mixtures were desalted via C18 ZipTip (EMD Millipore) per the manufacturer's instructions, and the product was eluted using a saturated solution of sinapinic acid in 60% aq. MeCN. Reactions were monitored using MALDI-TOF-MS.

UPLC MS and ESI-MS/MS. The sample was prepared as described previously.[1] LC-MS was performed using a Waters SYNAPT mass spectrometer outfitted with an ACQUITY UPLC, an ACQUITY Bridged Ethyl Hybrid C8 column (2.1 × 50 mm, 1.7 μm particle size, 200 Å; Waters), an ESI ion source, and a quadruple TOF detector. A gradient of 2–100% aq. MeCN with 0.1% formic acid (*v/v*) over 20 min was used. Fragmentation of the sample was performed using a collision-induced dissociation (CID) method. A ramping of cone voltage of 25–30 kV during the scan was performed to generate peptide fragments for MS/MS analysis.

Sequence S1. (A) Nucleotide sequence of codon-optimized *tbtI* for *E. coli* expression. The sequence is provided 5′ to 3′. Restriction sites for cloning are underlined (5′ BamHI, 3′ XhoI). This gene was synthesized by GenScript (Piscataway, NJ, USA). (B) Translated protein sequence of TbtI with Fe-S cluster residues highlighted in yellow. Length: 404 amino acids. NCBI Accession number: WP_050760412.1.

A

GGATCCATGACGCGCCCGCTGCTGCTGTATGTGAATATCCCGTTCTGTAACTCTAAATGCCACTTCTGCGACTGGG
TTGTCCAAATCCCGGTGCGTGATCTGCGCCTGGACCAGCAAGCTCCGGGTCGTATTGCGTATCTGGATGCAGTTCG
TGCTCAGATCCGCGGTCAAGCACCGGTCCTGCGCGAACATTATCAGCCGGCAATTATCTACTGGGGCGGTGGCACC
GCTAGCATTCTGGGTGAAAGCGAAATCGAATCTCTGTATACGTGCCTGCGTTCAGAATTCGATCTGTCGCACGTCC
GCGAAACCACGATTGAAGGTAGTCCGGAATCCCTGACCCCGCAGAAACTGCGTCGCCTGCGTGAACTGGGCTTTGA
TCGCATTAGCATCGGTGTTCAGTCTTTCGATGACCAACGTCTGCGTCGCCTGGGTCGTGCACATTCAGCAGAACAG
GCAGTGGAAGTGGTTAAAAACGCCCACGCGGCCGGTTTTCGTAACATTAATATCGATCTGATCGTTGGCTTCCCGG
GTCAAACCGACGCTGAAGTCGCGGAATCGGTGCGTACCGCGCTGACGCTGCCGATTAACCATTTTAGTATCTATCC
GTACCGTGCATCCCCGGGTACCATTCTGCGTAAACAGGTGGAACGCGGTGGCCACCTGGATCTGAATCGTCAACTG
GCAGCTTATTACATCACGCGCGACCTGCTGGAAGAAGCTGGCTTTCCGGAATATGCGATGAGCTACTTCGGTGCAC
CGCGTTGCGAAGCAGATCAGGCATATTACCGCCTGACCATGGACTGGATTGGTTTTGGCAGTGGTGCGAACTCCCT
GCTGGGCCATCGTTATCTGGCATTCCGTAAAGGTCGCCTGCATGCTTACAACCAGAATCCGCTGCGCTTTGATGTG
AATGCACCGGCCAGCTCTCCGCAGCTGACCCTGCATTGGCTGAGCCAAGCCCTGACCACGGTGGAAGGCATGGACG
CACGTGTTTATCAGGAACGTACCGGTACGCCGCTGCGTGTCGCATGTGAAGAACCGGAAGTGCAAGCCTACCTGCG
TCGCATGTCTGAACATGGCCGTCTGATTATCGATCGCAACGGTATTCGTATCCACCGCGAAGACATTGCCCGTGTT
ATTATCGCACTGAATTGGATCGATACCCCGGGTGGCGACCAGAAAGTTACCCGTCTGACGCCGGTCTCAGCAACGT
CGTAACTCGAG

B

MTRPLLLYVNIPF**CNSKCHFC**DWVVQIPVRDLRLDQQAPGRIAYLDAVRAQIRGQAPVLREHYQPAIIYWGGGTAS
ILGESEIESLYTCLRSEFDLSHVRETTIEGSPESLTPQKLRRLRELGFDRISIGVQSFDDQRLRRLGRAHSAEQAV
EVVKNAHAAGFRNINIDLIVGFPGQTDAEVAESVRTALTLPINHFSIYPYRASPGTILRKQVERGGHLDLNRQLAA
YYITRDLLEEAGFPEYAMSYFGAPRCEADQAYYRLTMDWIGFGSGANSLLGHRYLAFRKGRLHAYNQNPLRFDVNA
PASSPQLTLHWLSQALTTVEGMDARVYQERTGTPLRVACEEPEVQAYLRRMSEHGRLIIDRNGIRIHREDIARVII
ALNWIDTPGGDQKVTRLTPVSATS

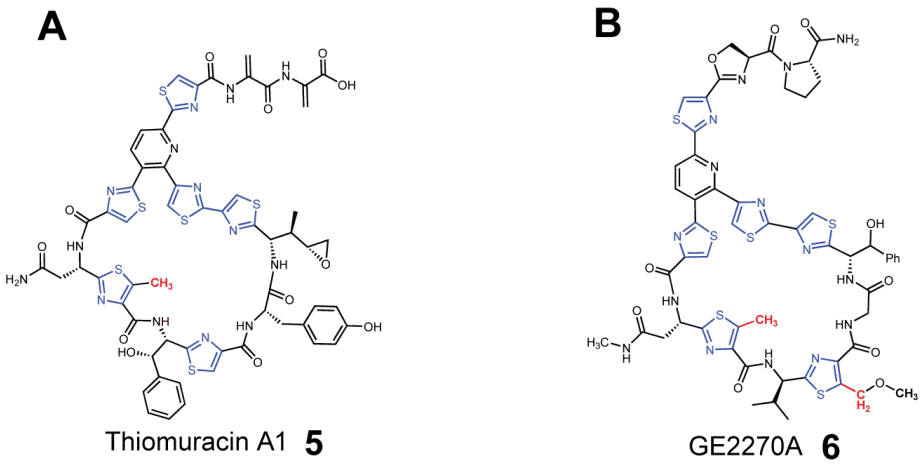
Table S1: **Oligonucleotide primers used in this study**. All sequences are provided 5′ to 3′. F indicates a forward primer while R indicates the reverse.

Primer Name	Oligonucleotide Sequence
tbtA N3A F	GTGCAAGCTGTGCTTGCTTTTGTTATATTTGTTGTAGCTGC
tbtA N3A R	TAACAAAAGCAAGCACAGCTTGCACCAACTTCGGTCATACC
tbtA N3D F	GTGCAAGCTGTGATTGCTTTTGTTATATTTGTTGTAGCTGC
tbtA N3D R	TAACAAAAGCAATCACAGCTTGCACCAACTTCGGTCATACC
tbtA N3Q F	GTGCAAGCTGTCAATGCTTTTGTTATATTTGTTGTAGCTGC
tbtA N3Q R	TAACAAAAGCATTGACAGCTTGCACCAACTTCGGTCATACC
tbtA C4A F	CAAGCTGTAATGCCTTTTGTTATATTTGTTGTAGCTGCAGC
tbtA C4A R	ATATAACAAAAGGCATTACAGCTTGCACCAACTTCGGTCAT
tbtA C4S F	CAAGCTGTAATAGCTTTTGTTATATTTGTTGTAGCTGCAGC
tbtA C4S R	ATATAACAAAAGCTATTACAGCTTGCACCAACTTCGGTCAT
tbtA S1N/C4A F	AAGTTGGTGCAAACTGTAATGCATTTTGTTATATTTGTTGT
tbtA S1N/C4A R	AATGCATTACAGTTTGCACCAACTTCGGTCATACCATGACC
tbtA F5N/C4A F	GCTGTAATGCAAATTGTTATATTTGTTGTAGCTGCAGCAGC
tbtA F5N/C4A R	CAAATATAACAATTTGCATTACAGCTTGCACCAACTTCGGT
tbtA I8N/C4A F	CATTTTGTTATAATTGTTGTAGCTGCAGCAGCGCCTAATGC
tbtA I8N/C4A R	CAGCTACAACAATTATAACAAAATGCATTACAGCTTGCACC
tbtA S11N/C4A F	ATATTTGTTGTAACTGCAGCAGCGCCTAATGCGGCCG
tbtA S11N/C4A R	GCGCTGCTGCAGTTACAACAAATATAACAAAATGCATTACA
tbtA C4,6 F	CGAATTGTTTTTGTTATATTGCGGCGAGCGCCAGCAGCGCC
tbtA C4,6 R	GCCGCAATATAACAAAAACAATTCGCGCTTGCACCAACTTC
tbtA C4,9 F	TTGCGTATATTTGTGCGAGCGCCAGCAGCGCCTAACTCGAG
tbtA C4,9 R	GCTGGCGCTCGCACAAATATACGCAAAACAATTCGCGCTTG
tbtA C4,10 F	CGTATATTGCGTGTAGCGCCAGCAGCGCCTAACTCGAGCAC
tbtA C4,10 R	CTGCTGGCGCTACACGCAATATACGCAAAACAATTCGCGC
tbtA C4,12 F	TTGCGGCGAGCTGTAGCAGCGCCTAACTCGAGCACCACCAC
tbtA C4,12 R	TAGGCGCTGCTACAGCTCGCCGCAATATACGCAAAACAATTC
tbtA S1N/N3F F	AAGTTGGTGCAAACTGTTTTTGCTTTTGTTATATTTGTTGTAGCTGCAGC
tbtA S1N/N3F R	AACAAAAGCAAAAACAGTTTGCACCAACTTCGGTCATACCATGACCTGCGG

Table S2: **Theoretical and observed masses for diagnostic fragment ions in compound 8.** These ions were used to locate the site of methylation in the TbtA hexazole after treatment with TbtI (**Figure 3**). During the MS/MS data acquisition, a high fragmentation voltage was required owing to the presence of the thiazoles. See the methods section for further details.

Species	Theoretical mass (Da)	Observed Mass (Da)	Error (ppm)
$[M+H]^+$	1714.4828	1714.4898	4.1
$[M-2H_2O+H]^+$	1678.4616	1678.4490	7.5
b10	1002.3061	1002.3102	4.1
b16-H_2O	1520.3926	1520.3875	3.4
c5	415.1763	415.1808	10.8
c7	626.2179	626.2219	6.4
y2	177.0875	177.0882	4.0
y8	713.1845	713.1866	2.9
y8-H_2O	677.1634	677.1664	4.4

Figure S1: **Structures and gene clusters of thiopeptides bearing *C*-methylated thiazoles**. Thiomuracin and GE2270A are the only two thiopeptides bearing methylated thiazoles with known biosynthetic gene clusters. Shown are the structures of (**A**) thiomuracin and (**B**) GE2270A along with (**C**) the biosynthetic gene clusters from *Thermobispora bispora* and *Planobispora rosea*, respectively. Shown in pink are the class C rSAM methyltransferases (e.g. TbtI, the subject of the current study). The BGC of 6 additionally encodes two additional "classic" SAM-methyltransferases, PbtM1 and PbtM4. These methyltransferase act upon nucleophilic centers and proceed via by a polar mechanism unlike the rSAM enzymes.



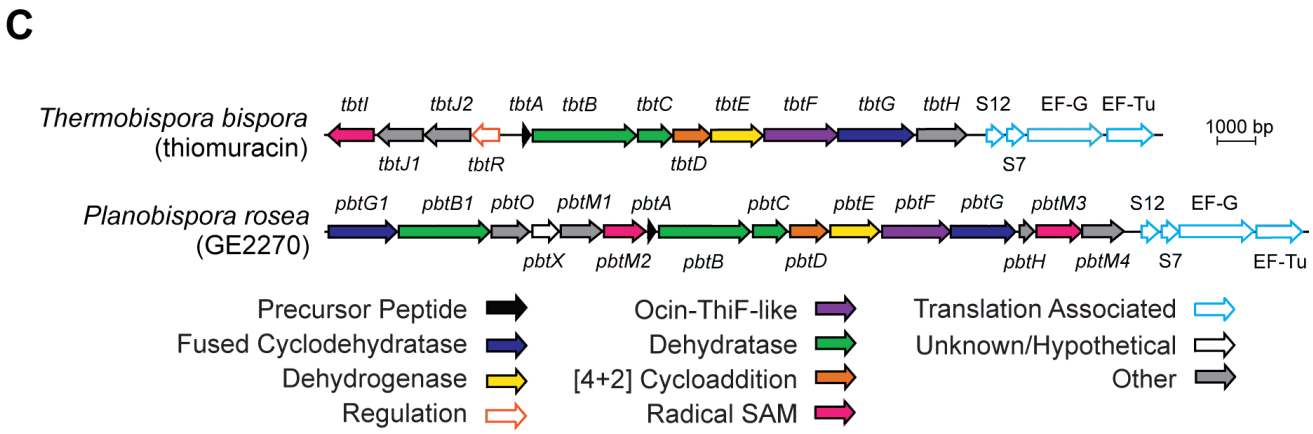


Figure S2: **SDS-PAGE analysis of proteins used in this study**.

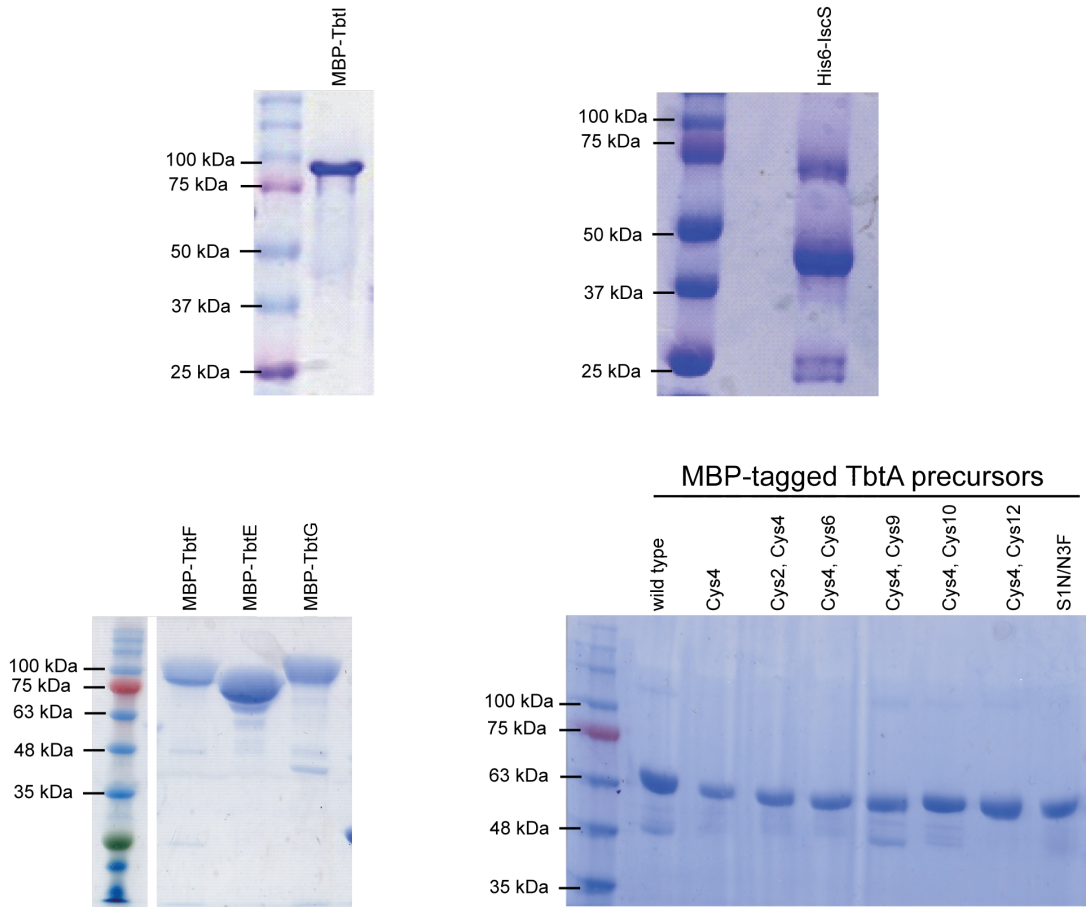


Figure S3: **UV-Vis spectrum of TbtI**. The UV-Vis spectrum of MBP-tagged TbtI (30 μM) collected on a Cary 4000 spectrophotometer. Shown are spectra of MBP-TbtI (green) as purified from *E. coli* and MBP-TbtI after in vitro Fe-S cluster reconstitution with IscS (purple).

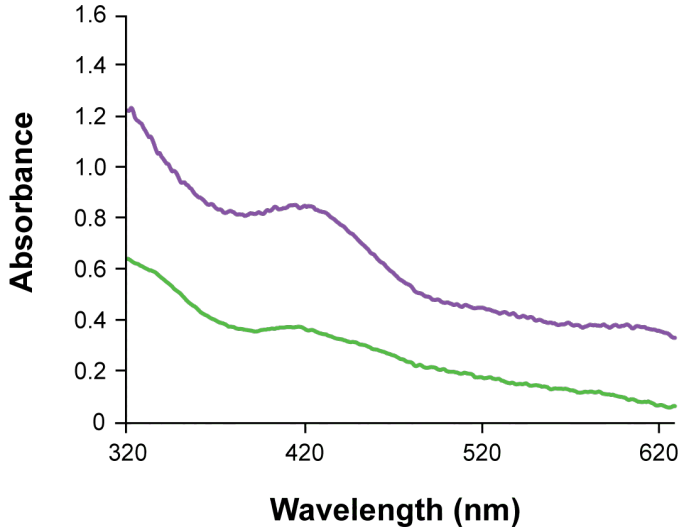


Figure S4: **TbtI reaction with Thiomuracin GZ (4)**. Shown are MALDI-TOF-MS spectra of thiomuracin GZ (top) and after reaction with TbtI supplemented with SAM and sodium dithionite (bottom). No methylation was observed even after extended reaction times. *m/z* assignments: 1366 Da, thiomuracin GZ (**4**).

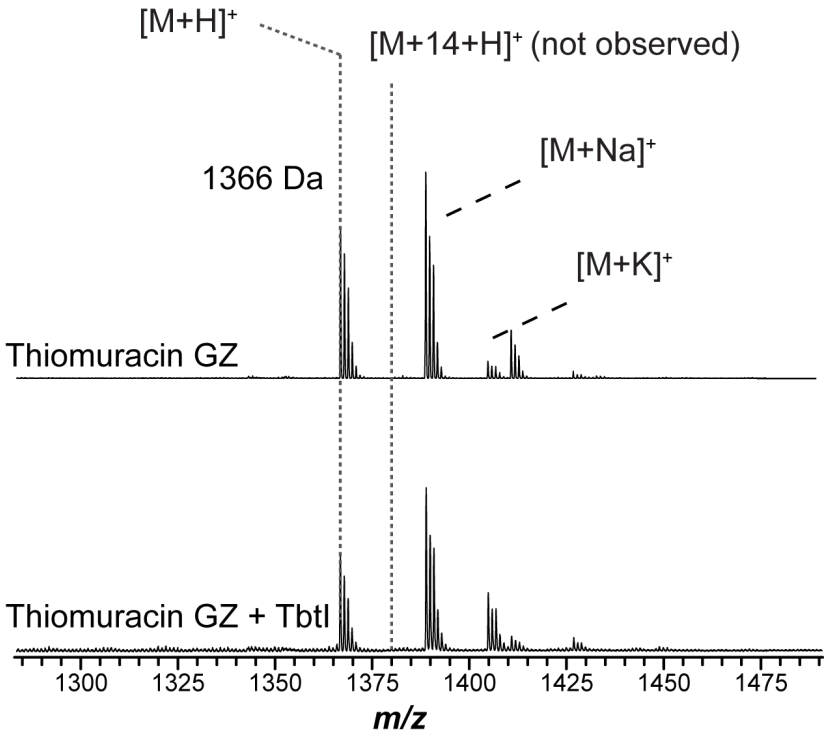


Figure S5: **TbtI reaction with TbtA hexazole, (2)**. TbtA hexazole (**2**) only showed apparent methylation (+14 Da) in the presence of TbtI, SAM, and dithionite as monitored by MALDI-TOF-MS. Omission of TbtI, SAM, or dithionite resulted in no reaction. *m/z* assignments: 6052 Da, TbtA hexazole (**2**) which at this resolution, forms a single peak with a minor product at *m/z* 6054, corresponding to TbtA pentazole/monoazoline, due to incomplete oxidation of one of the thiazolines; 6072 Da, TbtA pentazole with a hydrolyzed thiazoline (+18 Da); 6068 Da, methylated TbtA hexazole; 6086 Da, TbtA methylated pentazole with a hydrolyzed thiazoline.

His_6-TbtA: PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCSSA

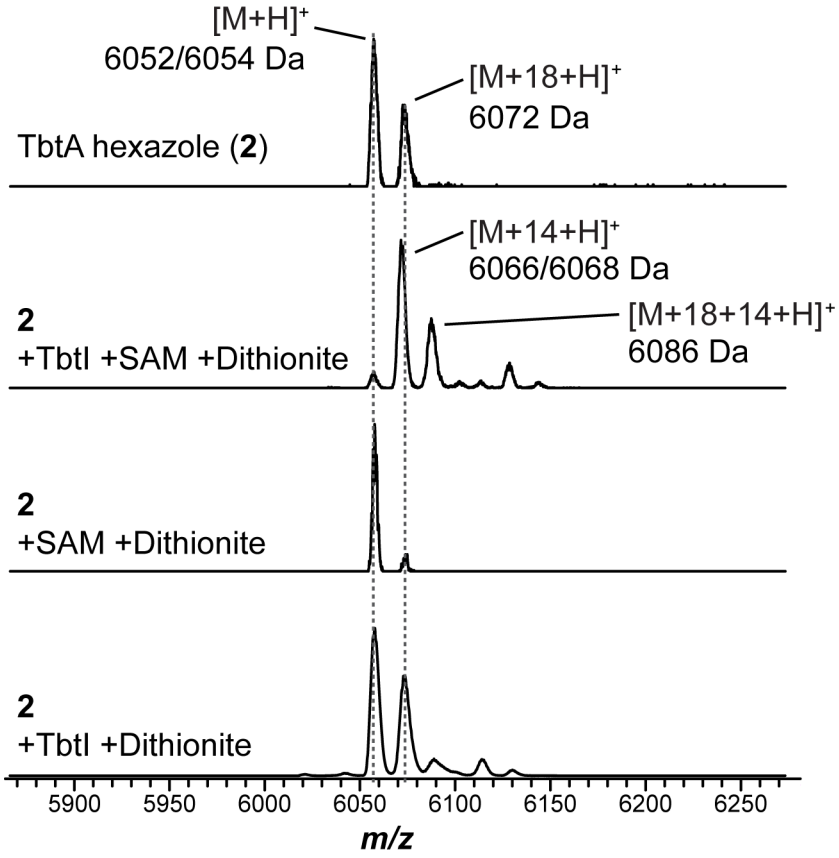
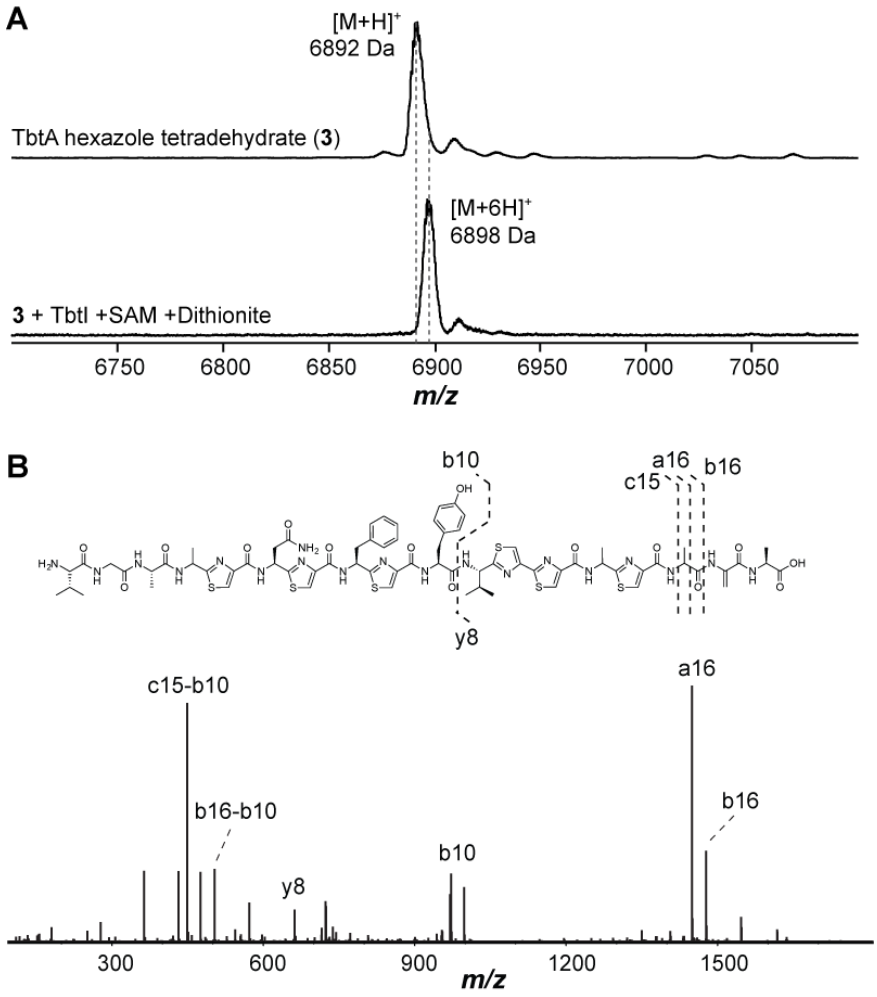


Figure S6: **TbtI reaction with TbtA hexazole tetradehydrate, (3)**. TbtA tetradehydrate was tested as a substrate for TbtI (**A**). The amino acid sequence shows the site of thiazoles (blue) and dehydrated Ser (green). No methylation (+14 Da) was observed in the presence of TbtI, SAM, and dithionite as monitored by MALDI-TOF-MS. Instead, a +6 Da species was observed. *m/z* assignments: 6892 Da, TbtA hexazole tetradehydrate (**3**); 6898 Da, off-pathway product of TbtI reaction with **3**. High-resolution tandem MS analysis of the +6 Da species indicates that **3** underwent reduction at the three most N-terminal alkene groups. This is a known reaction for dithionite,[6-8] which is required for the *in vitro* TbtI reactions.

His_6-TEV-TbtA: PHHHHHHSQENLYFQSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFCYICCSCSSA



C

Species	Theoretical mass (Da)	Observed Mass (Da)	Error (ppm)
a16	1448.4078	1448.4204	8.7
b10	972.2955	972.2918	3.8
b16	1476.4027	1476.4126	6.7
b16-b10	505.115	505.1219	13.7
c16-b10	451.1045	451.1105	13.3
y8	663.1842	663.1859	2.6

Figure S7: **TbtI reaction with TbtA-C4A and TbtA-C4S pentazoles**. TbtA pentazoles carrying the C4A or C4S substitutions were not substrates for TbtI, as visualized by MALDI-TOF-MS. Sites of thiazole installation are blue and the Ala/Ser substitutions are shown in red. *m/z* assignments: 6044 Da, TbtA-C4A pentazole; 6060 Da, TbtA-C4S pentazole.

His_6-TbtA (C4A): PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNAFCYICCSCSSA

His_6-TbtA (C4S): PHHHHHHSQVDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNSFCYICCSCSSA

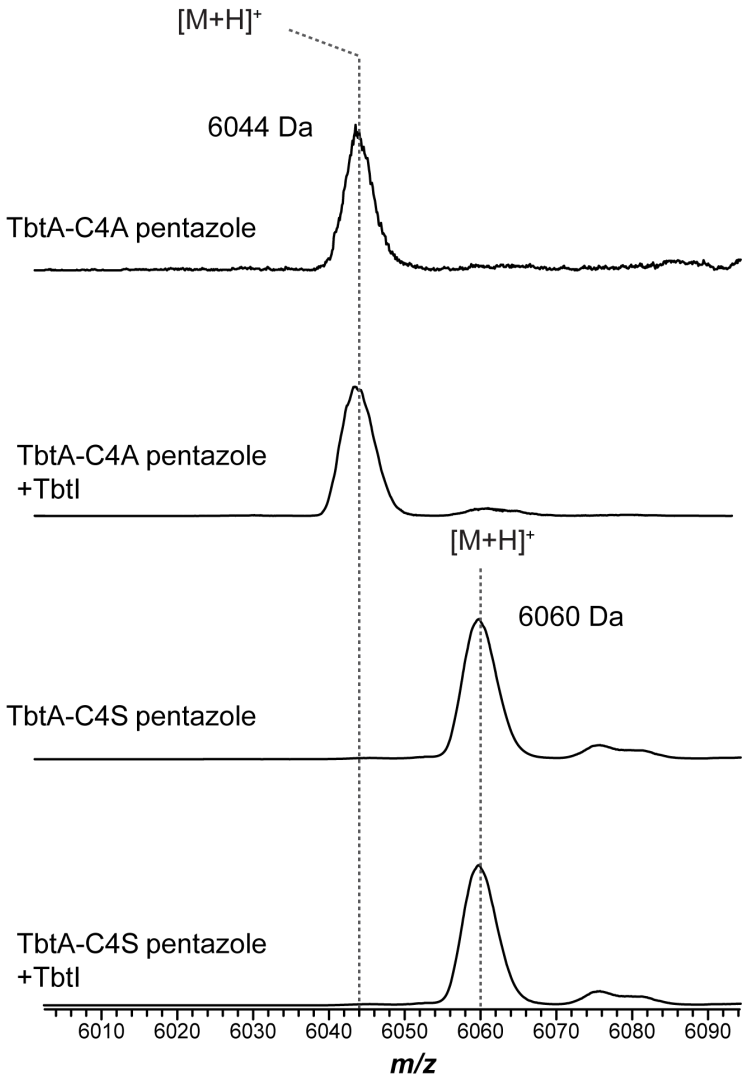


Figure S8: **TbtI reaction with TbtA-N3 variants**. GluC-treated TbtA hexazole variants were treated with TbtI. Black traces, starting material; Red traces, after treatment with TbtI. On the peptide sequences, thiazole sites are shown in blue and the amino acid substitutions are shown in red. While the wild-type (wt) substrate was properly processed, methylation was not observed for any Asn3 variant by MALDI-TOF-MS. *m/z* assignments: 1657 Da, TbtA-N3A hexazole core; 1701 Da, TbtA-N3D; 1700 Da, wt hexazole core; 1714 Da, TbtA-N3Q hexazole core, which is isobaric with methylated TbtA wt hexazole core.

TbtA-N3A: **V**GA*S**C**A**C**F**C**YI**CC**S**C**SSA

TbtA-N3D: **V**GA*S**C**D**C**F**C**YI**CC**S**C**SSA

TbtA-N3Q: **V**GA*S**C**Q**C**F**C**YI**CC**S**C**SSA

TbtA-WT: **V**GA*S**C**N**C**F**C**YI**CC**S**C**SSA

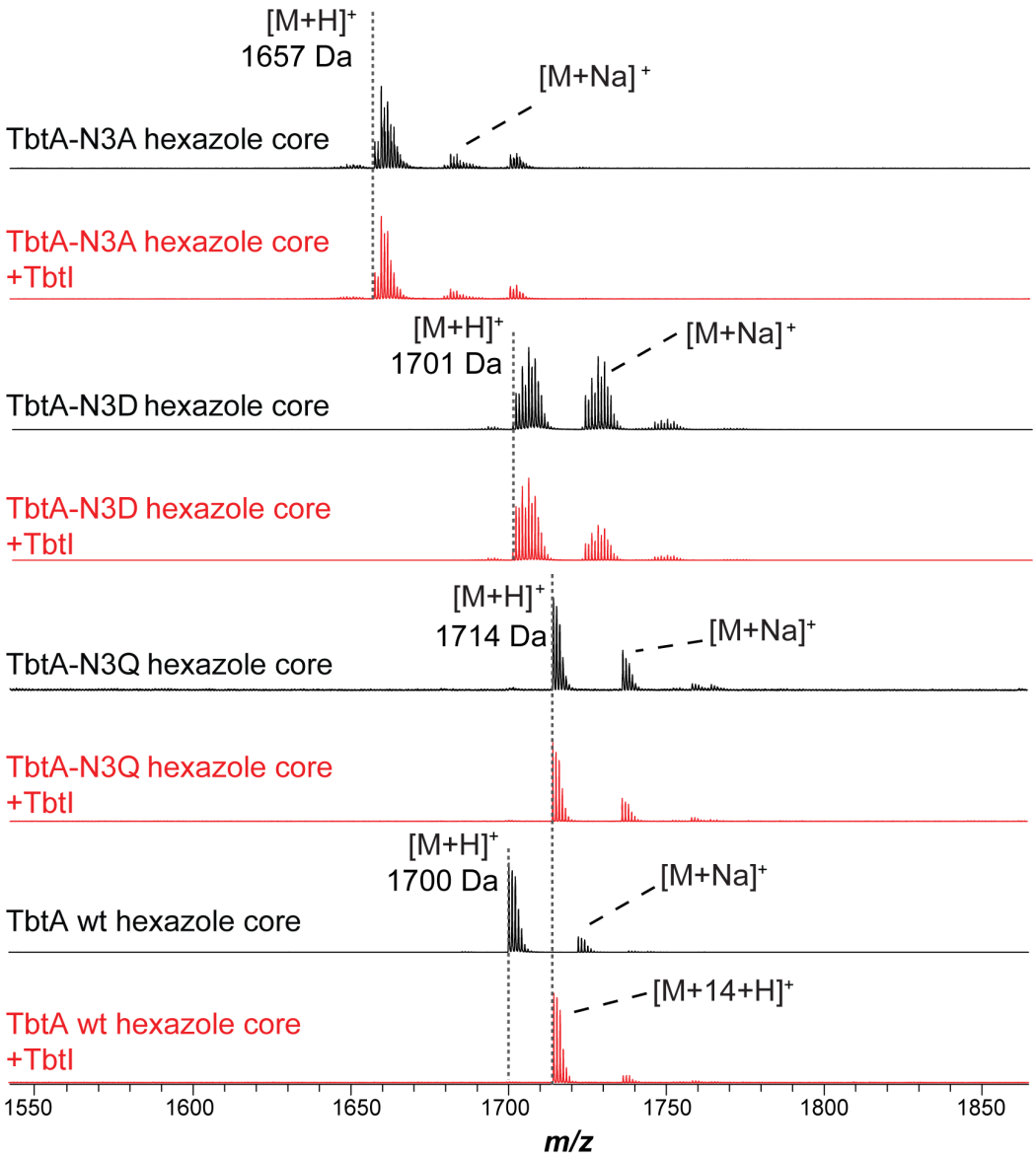


Figure S9: **TbtI reaction with TbtA-S1N/C4A and -C4A/F5N double variants**. GluC-treated TbtA pentazoles carrying the C4A substitution as well as S1N or F5N were not substrates for TbtI, as visualized by MALDI-TOF-MS. Sites of thiazole installation are blue and the substitutions are shown in red. *m/z* assignments: 1715 Da, TbtA-S1N/C4A hexazole core; 1655 Da, TbtA-C4A/F5N hexazole core.

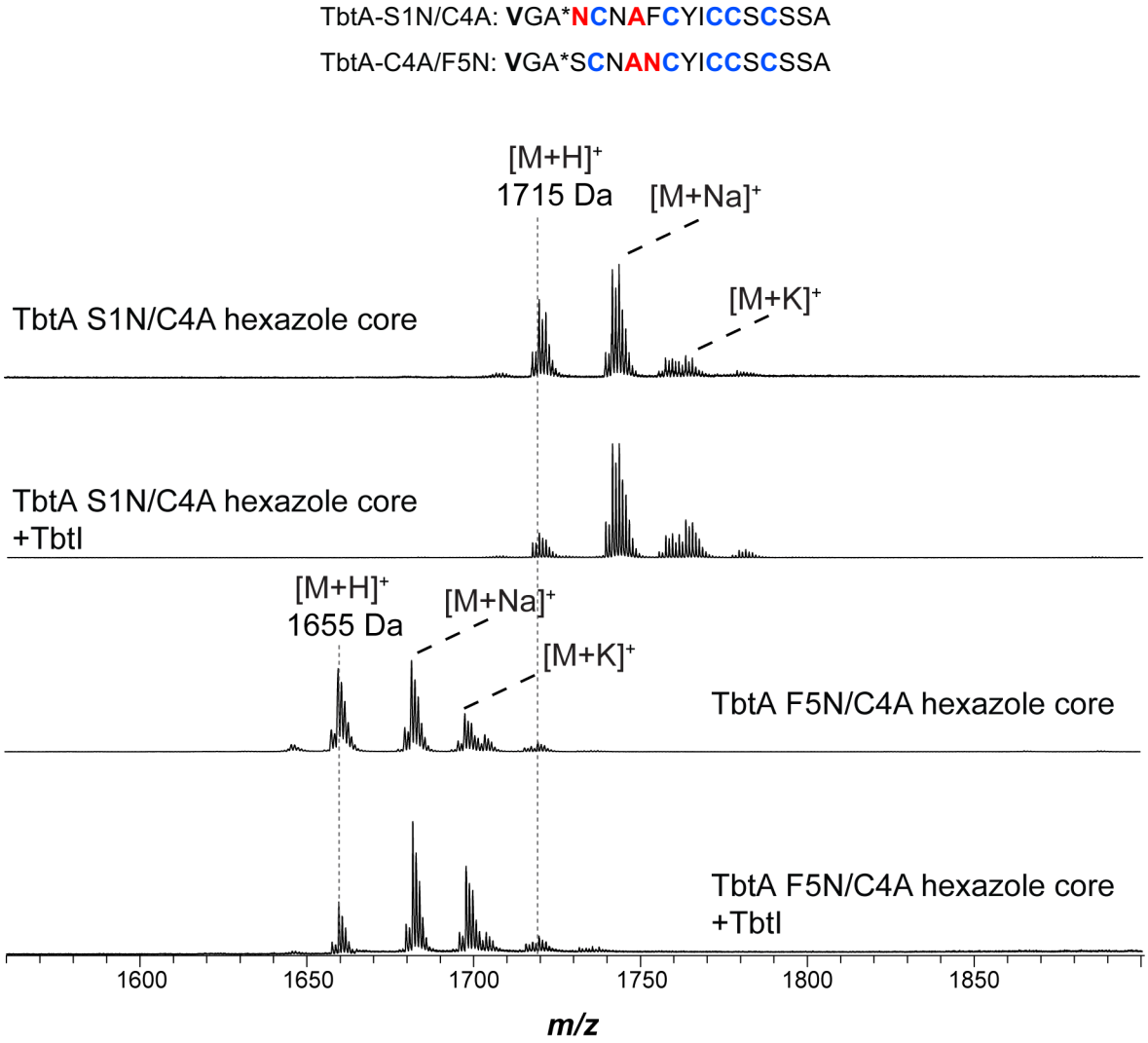


Figure S10: **TbtI reaction with TbtA-C4A/I8N and C4A/S11N double variants**. GluC-treated TbtA pentazoles carrying the C4A substitution as well as I8N or S11N were not substrates for TbtI, as visualized by MALDI-TOF-MS. Sites of thiazole installation are blue and the substitutions are shown in red. *m/z* assignments: 1689 Da, TbtA-C4A/I8N hexazole core; 1715 Da, TbtA-C4A/S11N hexazole core.

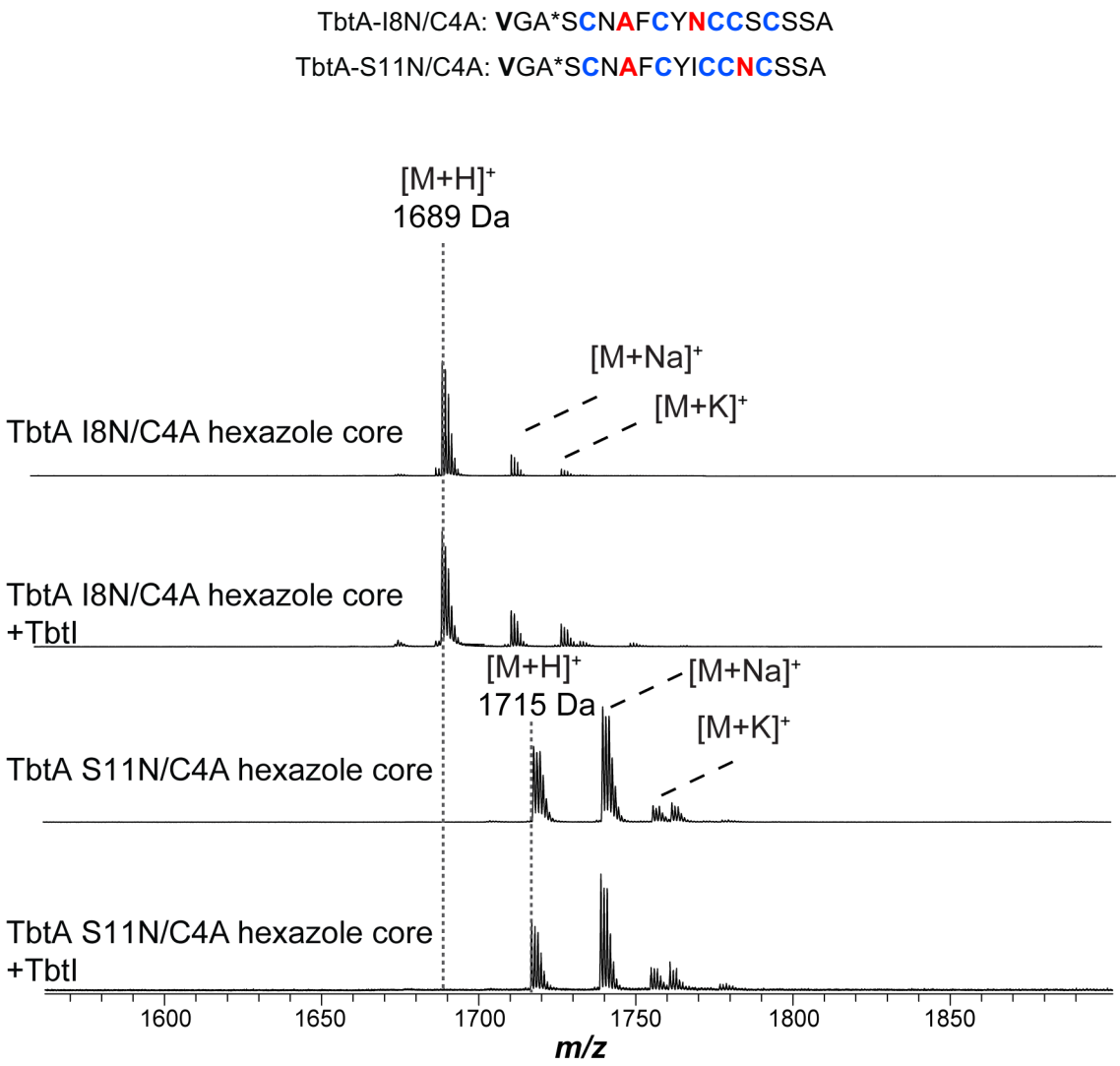


Figure S11: **TbtI reaction with TbtA-Thz4 monoazole**. TbtA featuring a single thiazole at position 4 was not a substrate for TbtI as visualized by MALDI-TOF-MS. All other Cys residues were substituted with Ala. Sites of thiazole installation are blue and substitutions are shown in red. Laser induced deamination artifacts are denoted with a red asterisk. *m/z* assignments: 5609 Da, unreacted TbtA (Cys4) precursor peptide; 5589 Da, TbtA (Thz4) monoazole.

TbtA (Cys4): SRRRGSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SANCFAYIAASASSA

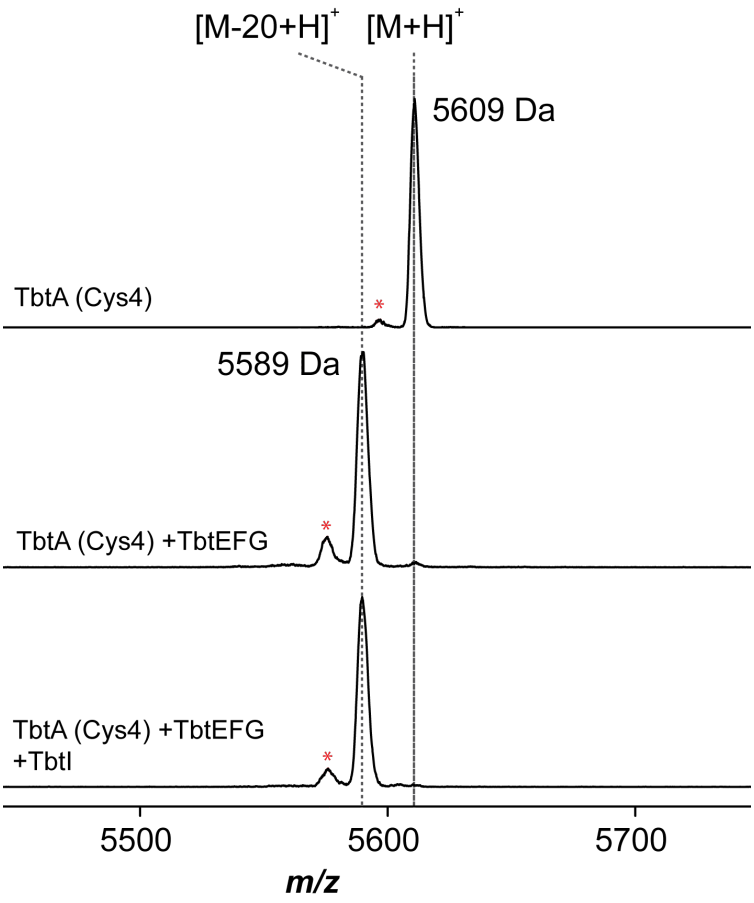


Figure S12: **TbtI reaction with TbtA-Thz2/Thz4 diazole**. TbtA featuring two thiazoles at positions 2 and 4 was not a substrate for TbtI as visualized by MALDI-TOF-MS. All other Cys residues were substituted with Ala. Sites of thiazoles installation are blue and substitutions are shown in red. Laser induced deamination artifacts are denoted with a red asterisk. *m/z* assignments: 5641 Da, unreacted TbtA (Cys2, Cys4) precursor peptide; 5601 Da, TbtA (Thz2/Thz4) diazole.

TbtA (Cys2, Cys4): SRRRGSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SCNCFAYIAASASSA

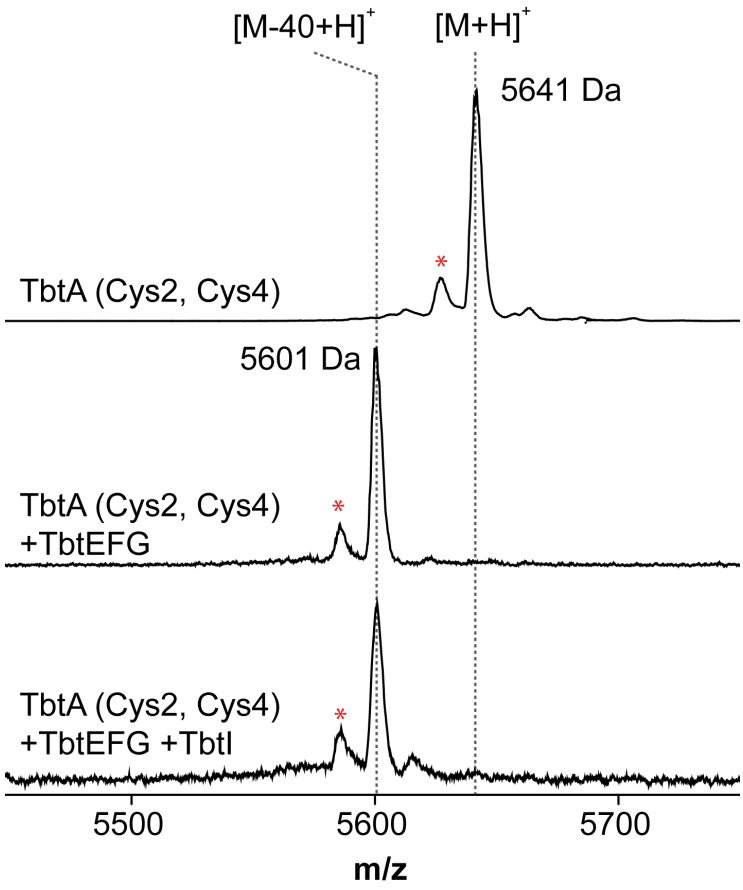


Figure S13: **TbtI reaction with TbtA-Thz4/Thz9 diazole**. TbtA featuring two thiazoles at positions 4 and 9 was partially processed by TbtI at the 30-minute time point as visualized by MALDI-TOF-MS. All other Cys residues were substituted with Ala. Sites of thiazole installation are blue and substitutions are shown in red. Laser induced deamination artifacts are denoted with a red asterisk and a putative off-pathway degradation product is denoted with a red §. *m/z* assignments: 5641 Da, unreacted TbtA (Cys4, Cys9) precursor peptide; 5601 Da, TbtA (Thz4/Thz9) diazole; 5615 Da, methylated TbtA (Thz4/Thz9) diazole.

TbtA (Cys4, Cys9): SRRRGSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*S**A**N**C**F**A**Y**I****C****A**S**A**SSA

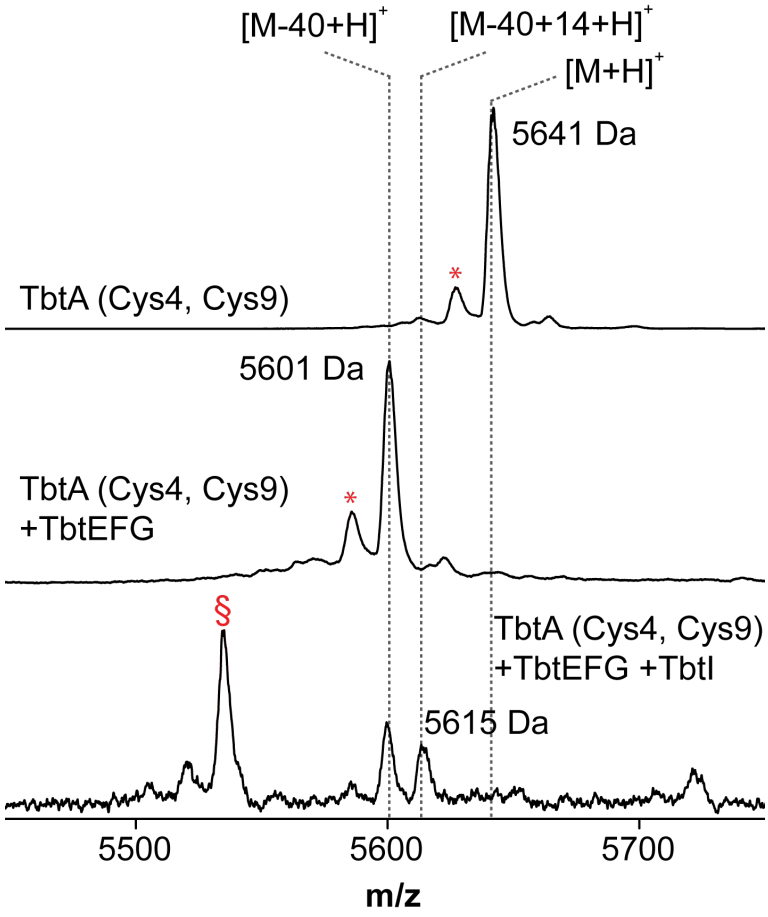


Figure S14: **TbtI reaction with TbtA-Thz4/Thz10 diazole**. TbtA featuring two thiazoles at positions 4 and 10 was not a substrate for TbtI as visualized by MALDI-TOF-MS. All other Cys residues were substituted with Ala. Sites of thiazole installation are blue and substitutions are shown in red. Laser induced deamination artifacts are denoted with a red asterisk. *m/z* assignments: 5641 Da, unreacted TbtA (Cys4, Cys10) precursor peptide; 5601 Da, TbtA (Thz4/Thz10) diazole.

TbtA (Cys4, Cys10): SRRRGSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SANCFAYIACSASSA

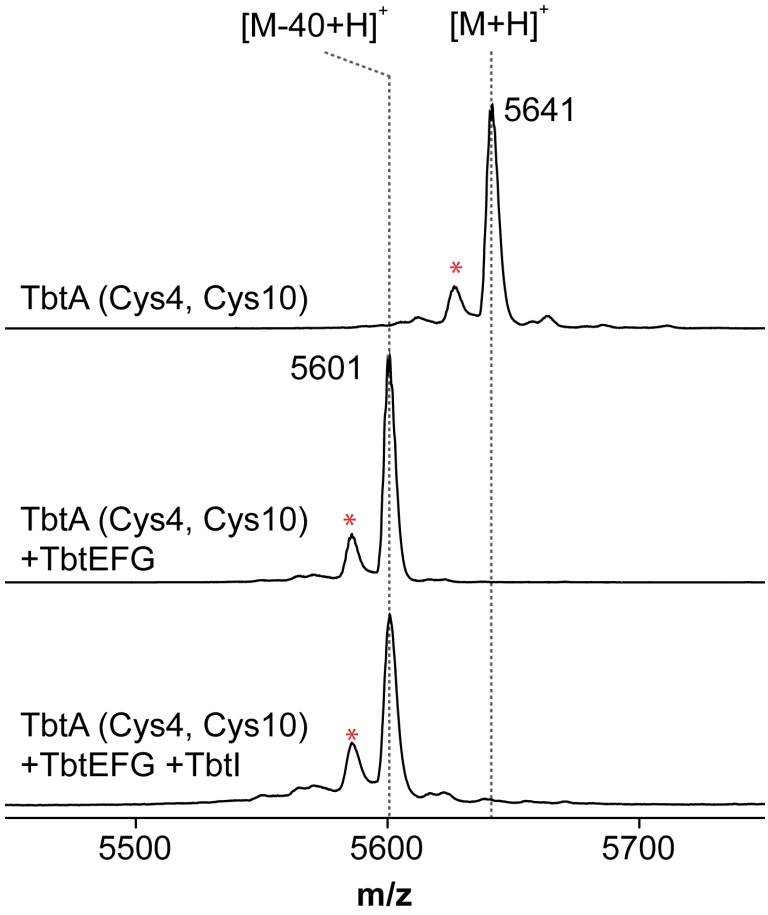


Figure S15: **TbtI reaction with TbtA-Thz4/Thz12 diazole**. TbtA featuring two thiazoles at positions 4 and 12 was not a substrate for TbtI as visualized by MALDI-TOF-MS. All other Cys residues were substituted with Ala. Sites of thiazole installation are blue and substitutions are shown in red. Laser induced deamination artifacts are denoted with a red asterisk and a putative off-pathway degradation product is denoted with a red §. *m/z* assignments: 5641 Da, unreacted TbtA (Cys4, Cys12) precursor peptide; 5601 Da, TbtA (Thz4/Thz12) diazole.

TbtA (Cys4, Cys12): SRRRGSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*SANCFAYIAASCSSA

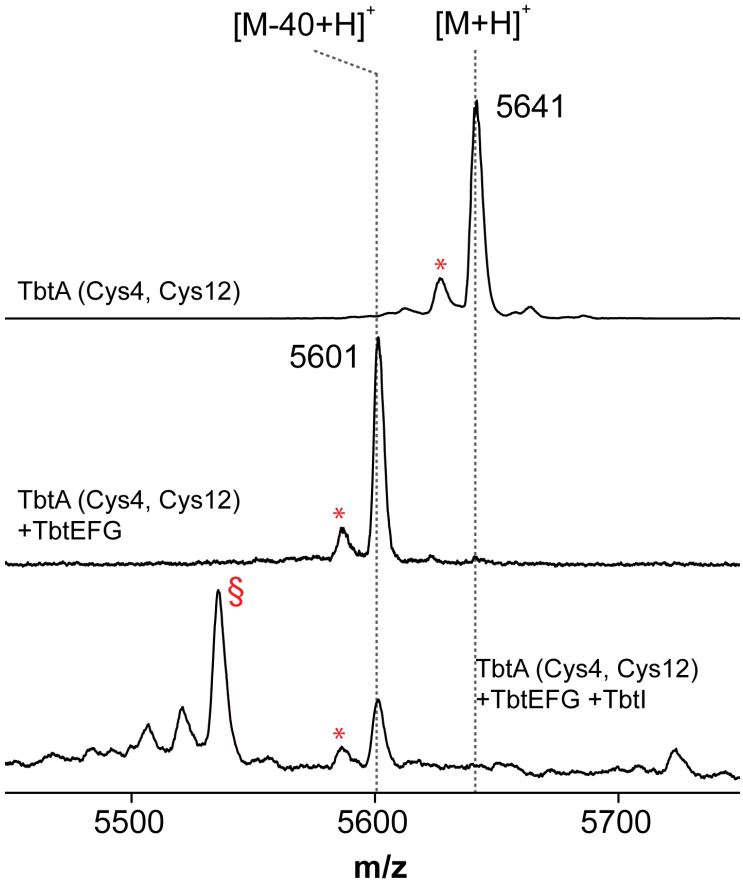


Figure S16: **TbtI reaction with TbtA S1N/N3F double variant**. TbtA hexazole featuring the substitutions Ser1Asn and Asn3Phe was not a substrate for TbtI as visualized by MALDI-TOF-MS. Sites of thiazole installation are blue and substitutions are shown in red. Laser induced deamination artifacts are denoted with a red asterisk. *m/z* assignments: 5830 Da, unreacted TbtA (S1N, N3F) precursor peptide; 5710 Da, TbtA (S1N, N3F) hexazole.

TbtA S1N/N3F: SRRRGSMDLNDLPMDVFELADSGVAVESLTAGHGMTEVGA*NCFCFCYICCSCSSA

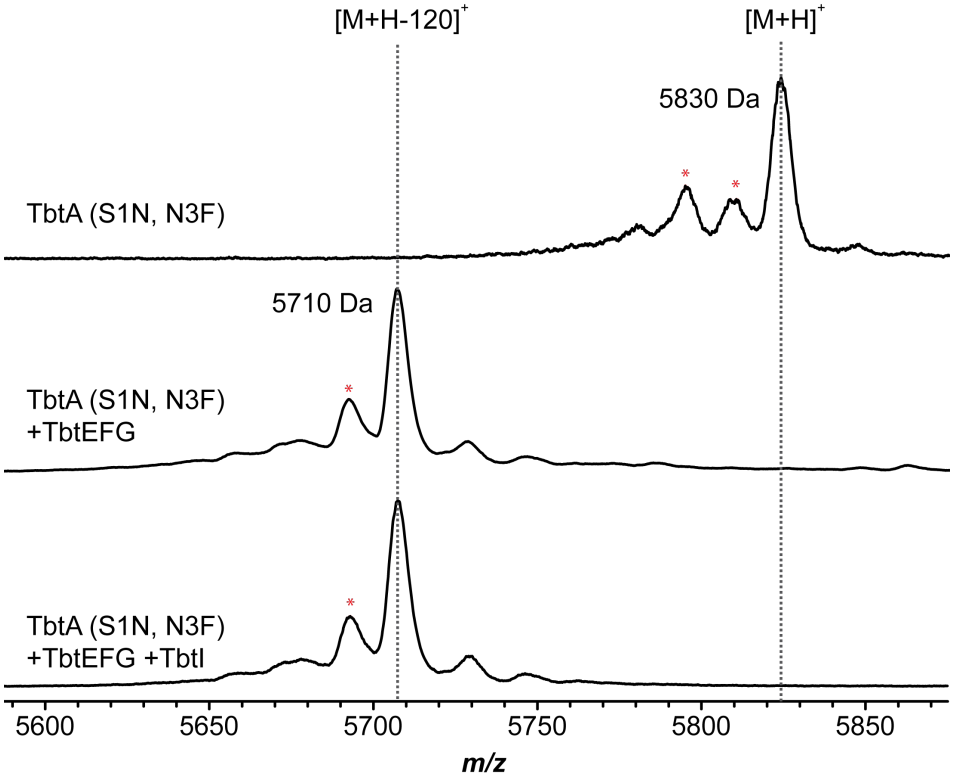


Figure S17: Sequence similarity network (SSN) of the top 20,000 proteins identified by a BLAST-P search against the non-redundant GenBank database using TbtI as the query sequence. The SSN was generated by the Enzyme Function Initiative Enzyme Similarity Tool and connects nodes with an alignment score of at least 55. Highly similar proteins (i.e. ≥ 50% identity) are conflated into a single node. The numeric values next to each group shows their non-conflated membership.

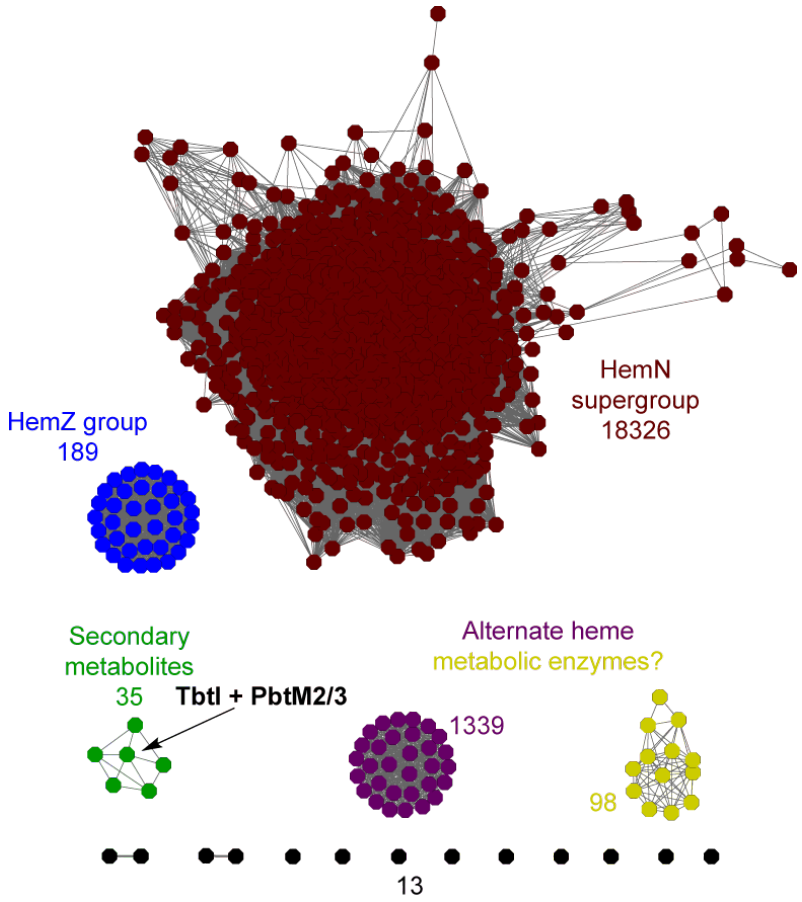


Figure S18: **Biosynthetic gene clusters identified containing TbtI homologs**. TbtI homologs potentially involved in natural product biosynthesis by sequence similarity analysis (**Figure 5**) were further analyzed for local genomic context. Several gene clusters were identified that encode for a putative (**A**) goadsporin-like[9] cluster found in *Janthinobacterium* sp. CG23_2 and several (**B**) linear azole-containing peptides.[10] The leader peptide-cleavage site for the latter peptides is predicted to be: (…PGA*SSC…).

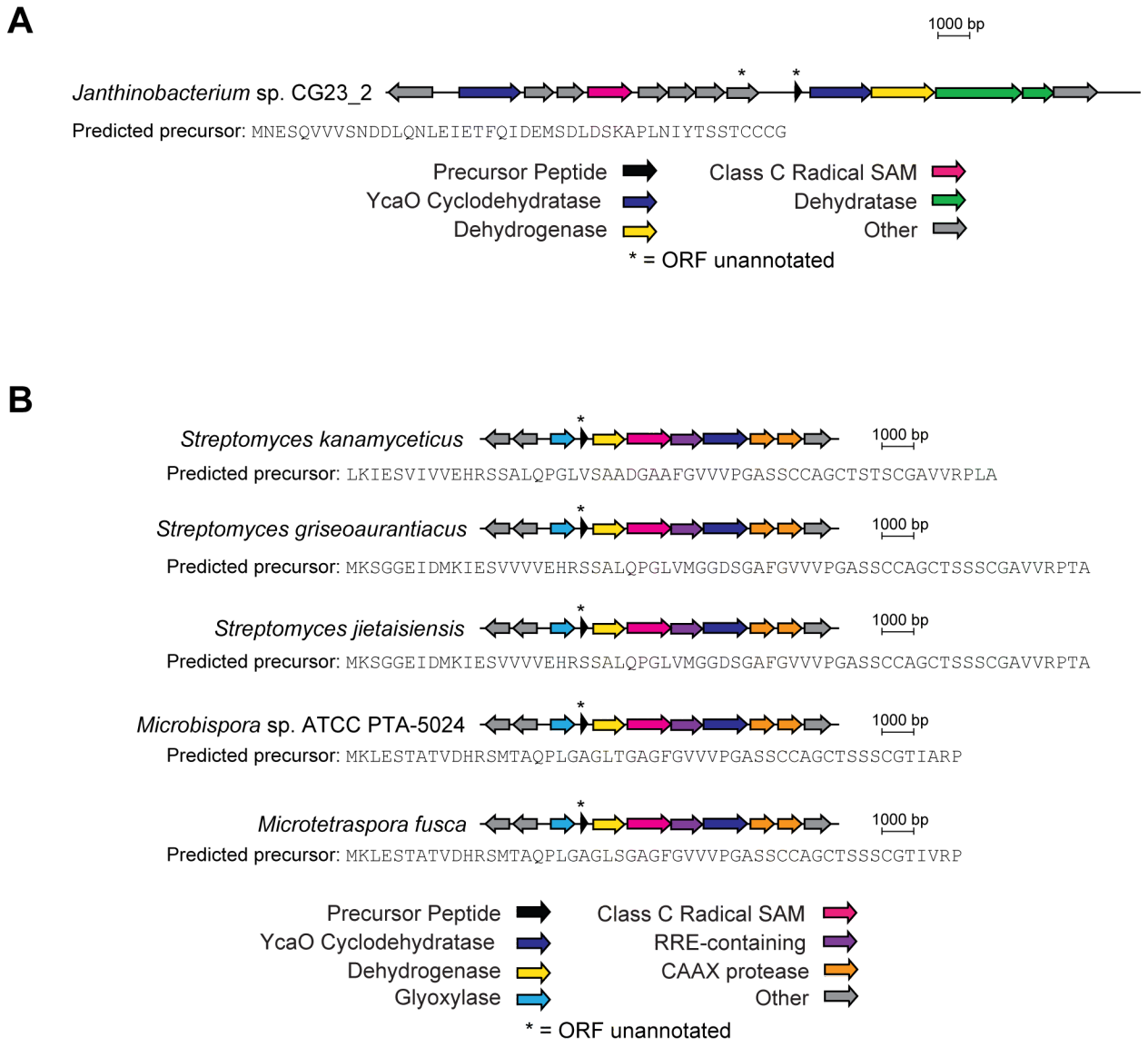


Figure S19: **Maximum likelihood tree of non-HemN/Z TbtI homologs**. Maximum likelihood tree of TbtI homologs with HemN/HemZ sequences omitted. Coloring is based on anticipated natural product class. YtkT and Jaw5 (lime green) were manually added, given that they were not retrieved by BLAST-P search of the top 20,000 TbtI homologs (see main text). For reference, TbtI appears at approximately 8:30 on this radial tree while NosN appears at approximately 7:00. The tree was generated in MEGA7 using the James-Taylor-Thornton model and visualized using EMBL's Interactive Tree of Life software.[11-13] Although they were not within the top 20,000 matched sequences, YtkT and Jaw5 clustered within clades known to be involved in natural product biosynthesis.[14] *Abbreviations*: LAP, linear azol(in)e-containing peptide; PKS, polyketide synthase; NRPS, non-ribosomal peptide synthetase; BGC, biosynthetic gene cluster.

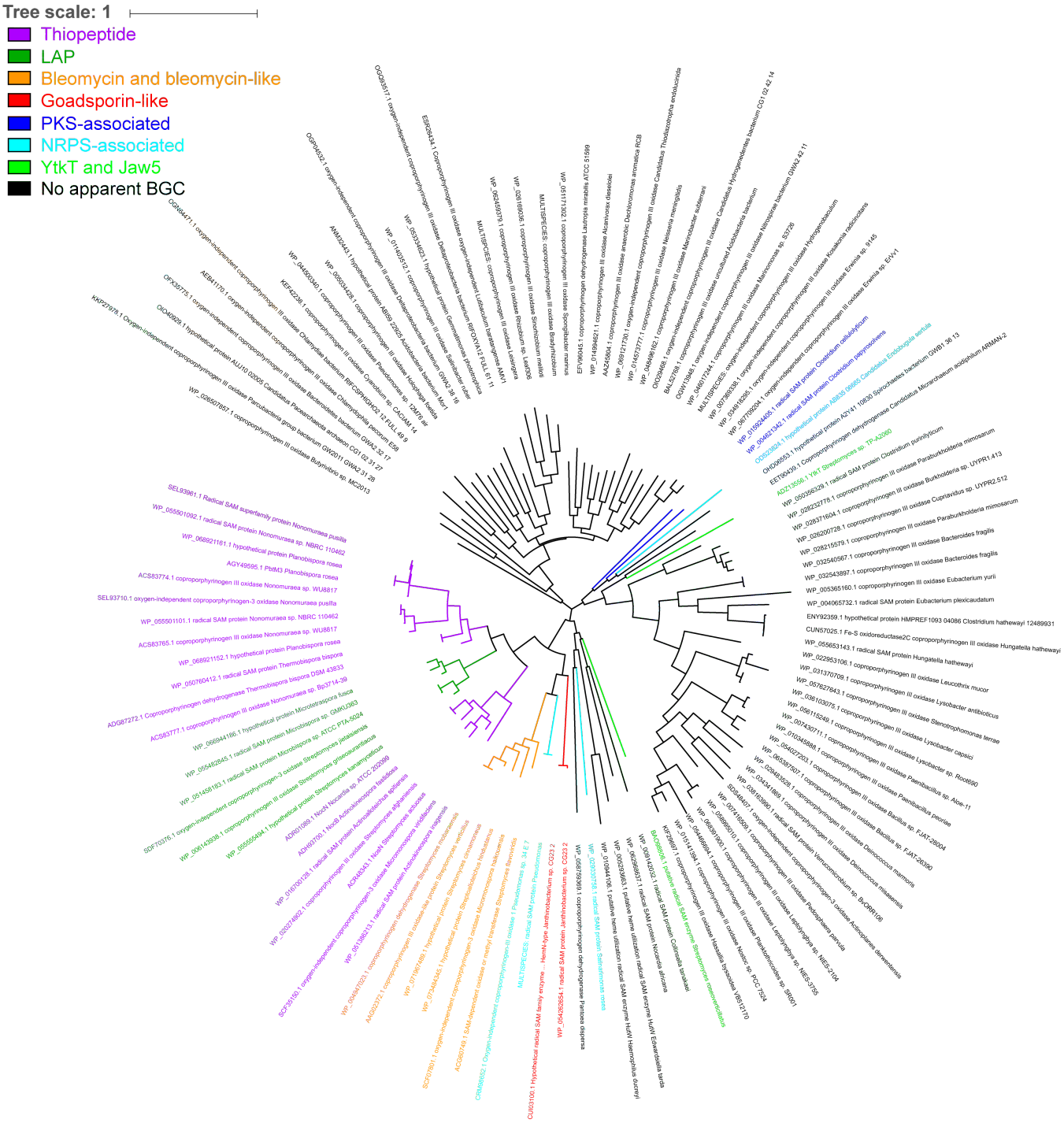
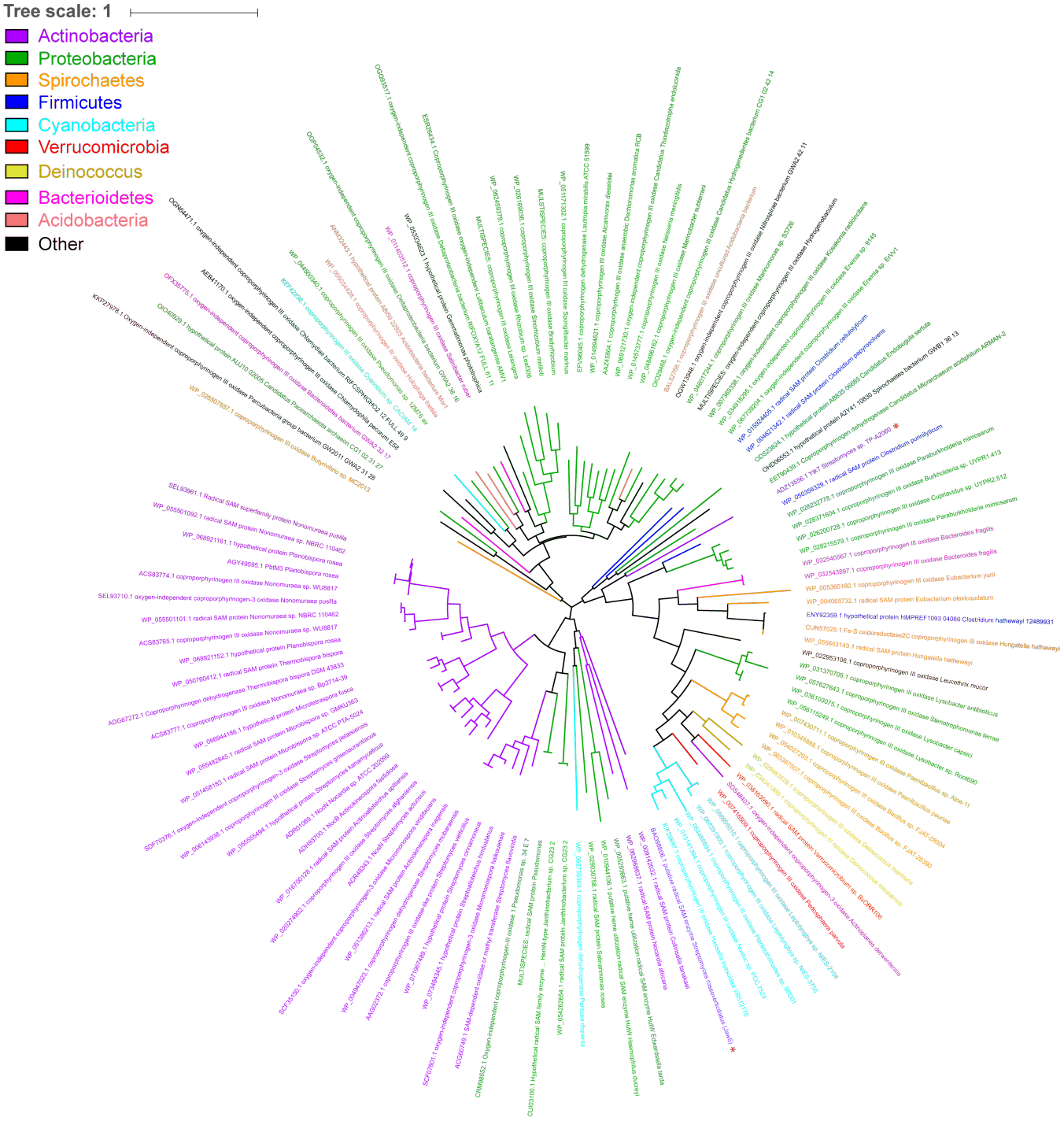


Figure S20: **Phylogenetic classification of non-HemN/Z TbtI homologs**. Maximum likelihood tree of TbtI homologs with HemN/HemZ sequences omitted. Coloring is based on bacterial phylum. As in Figure S18, YtkT and Jaw5 were manually added and here denoted with peripheral red asterisks at approximately 2:00 and 5:30. The tree was generated in MEGA7 using the James-Taylor-Thornton model and visualized using EMBL's Interactive Tree of Life software.[11-13] The majority of these proteins derive from the actinobacteria and proteobacteria phyla. The examples from actinobacteria are enriched in TbtI homologs with suspected involvement in natural product biosynthesis (compare to Fig. S18). The non-BGC-associated homologs also tend to be appear in multiphyletic clades while the BGC-associated homologs appear in more monophyletic clades.



Supporting References (full citation for reference 1 of the main text is given here as reference 10)

1. Zhang, Z.; Hudson, G. A.; Mahanta, N.; Tietz, J. I.; van der Donk, W. A.; Mitchell, D. A. *J. Am. Chem. Soc.* **2016,** *138* (48), 15511.

2. Jang, S.; Imlay, J. A. *Mol Microbiol.* **2010,** *78* (6), 1448.

3. Hidalgo, E.; Bollinger, J. M., Jr.; Bradley, T. M.; Walsh, C. T.; Demple, B. *J. Biol. Chem.* **1995,** *270* (36), 20908.

4. Brumby, P. E.; Miller, R. W.; Massey, V. *J. Biol. Chem.* **1965,** *240*, 2222.

5. Hudson, G. A.; Zhang, Z.; Tietz, J. I.; Mitchell, D. A.; van der Donk, W. A. *J. Am. Chem. Soc.* **2015,** *137* (51), 16012.

6. Camps, F.; Coll, J.; Guerrero, A.; Guitart, J.; Riba, M. *Chem. Lett.* **1982,** *11* (5), 715.

7. Camps, F.; Coll, J.; Guitart, J. *Tetrahedron* **1986,** *42* (16), 4603.

8. Louisandre, O.; Gelbard, G. *Tetrahedron Lett.* **1985,** *26* (7), 831.

9. Onaka, H.; Nakaho, M.; Hayashi, K.; Igarashi, Y.; Furumai, T. *Microbiology* **2005,** *151* (12), 3923.

10. Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D.; Fischbach, M. A.; Garavelli, J. S.; Goeransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Mueller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Suessmuth, R. D.; Tagg, J. R.; Tang, G.-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. *Nat. Prod. Rep.* **2013,** *30* (1), 108.

11. Jones, D. T.; Taylor, W. R.; Thornton, J. M. *Comput. Appl. Biosci.* **1992,** *8* (3), 275.

12. Kumar, S.; Stecher, G.; Tamura, K. *Mol. Biol. Evol.* **2016**, *33* (7), 1870.

13. Letunic, I.; Bork, P. *Nuc. Acids Res.* **2016,** *44* (W1), W242.

14. Huang, W.; Xu, H.; Li, Y.; Zhang, F.; Chen, X.-Y.; He, Q.-L.; Igarashi, Y.; Tang, G.-L. *J. Am. Chem. Soc.* **2012,** *134* (21), 8831.