Supporting information for:

Structural and Functional Insight into an Unexpectedly Selective *N*-Methyltransferase Involved in Plantazolicin Biosynthesis

Jaeheon Lee[1*], Yue Hao[1,2*], Patricia M. Blair[3], Joel O. Melby[1,3], Vinayak Agarwal[1,4], Brandon J. Burkhart[1,3], Satish K. Nair[1,2,4,†], and Douglas A. Mitchell[1,3,5,†]

[1]Institute for Genomic Biology, [2]Department of Biochemistry, [3]Department of Chemistry, [4]Center for Biophysics and Computational Biology, [5]Department of Microbiology; all of the University of Illinois at Urbana-Champaign, Urbana, IL 61801

TABLE OF CONTENTS:

Materials anb Methods

General methods. Unless otherwise specified, all chemicals were purchased from either Fisher Scientific or Sigma-Aldrich. Oligonucleotides were synthesized by Integrated DNA Technologies and are listed in Table S3. All restriction endonucleases were purchased from New England BioLabs (NEB) and dNTPs were from Promega. PCR amplifications were performed using *Pfu*Turbo DNA polymerase (Agilent). Chemically competent DH5α and BL21(DE3)RIPL *E. coli* cells were used for routine subcloning and protein expression according to Sambrook *et al* (1). DNA sequencing was performed by ACGT, Inc. The "RPG" (RPGPGPIIPGPGPF), "RAT" (RATATTI), RAAA, and RGGG peptides were synthesized by GenScript.

Production and purification of desmethylPZN. An overnight seed culture of *B. amyloliquefaciens* FZB42 strain RS33 (2) was plated onto M9 minimal media agar plates (5 × 15 cm diameter dishes) containing spectinomycin (90 μg/mL), chloramphenicol (7 μg/mL), and erythromycin (1 μg/mL). These plate cultures were grown for 48 h at 37 °C prior to resuspending the cells in Tris-buffered saline (10 mM Tris, 150 mM NaCl, pH 8.0). The resuspended cells were transferred to a centrifuge tube and harvested (20,000 × *g*, 30 min, 4 °C). The cell pellet was resuspended in MeOH (7.5% culture volume) with vortexing. After equilibrating for 20 min at 22 °C, the agitated cells were again collected by centrifugation. The desmethylPZN-containing supernatant was filtered using Whatman filter paper and concentrated to dryness by rotary evaporation. The crude material was redissolved in 80% aqueous MeCN (2 mL) and passed through a syringe-driven sterile-filtration membrane (0.2 μm PES). This sample was subjected to purification on an Agilent 1200 series HPLC using a semi-preparative C_{18} column (Thermo BETASIL C_{18} column, 250 mm × 10 mm; pore size: 100 Å; particle size: 5 μm) at a flow rate of 4 mL/min. Solvent gradient: initial conditions 65% MeCN (aq) with 0.1% (v/v) formic acid were maintained for 6 min, followed by a linear gradient up to 85% MeCN (aq) with 0.1% (v/v) formic acid over 32 min. The fractions containing desmethylPZN were monitored by UV absorbance (272 nm), collected into 20 mL borosilicate vials, and dried by rotary evaporation. The fractions were analyzed on a Bruker Daltonics UltrafleXtreme MALDI TOF/TOF MS (see method below), lyophilized, and stored at -80 °C until needed.

Preparation of expression vectors. Protein expression vectors used in this study were based upon a modified pET28b vector (pET28b-MBP) containing an *N*-terminal maltose-binding protein (MBP) tag followed by thrombin and TEV protease cleavage sites. The *bamL* gene (PZN methyltransferase, locus tag RBAM_007500) was amplified from the genomic DNA of *B. amyloliquefaciens* FZB42 using primers BamLf and BamLr (Table S3), digested with the *BamH*I and *Not*I restriction enzymes, and cloned into similarly digested pET28b-MBP vector to generate pMBP-BamL (Table S3). The *bpumL* gene (locus tag BAT_2465) was amplified from the genomic DNA of *Bacillus pumilus* ATCC 7061 using primers BpumLf

and BpumLr (Table S3) and cloned into the pET28b vector to generate pET28b-BpumL (Table S3). Note that BpumL was not MBP-tagged.

QuikChange (Agilent) was employed to introduce site-directed mutations into *bamL* following the manufacturer's protocol. The sequences of all primers used are provided in Table S3. Transformants were selected from chemically competent DH5α cells using kanamycin (50 μg/mL); the plasmids were purified by miniprep (Epoch Biosciences) and then sequenced using the T7 terminator primer and a custom MBP forward primer (5' - ATGAAGCCCTGAAAGACG - 3').

Overexpression and purification of MBP-BamL/BpumL and mutant proteins. Each desired construct verified by DNA sequencing was transformed into chemically competent BL21(DE3)RIPL cells. Transformants were selected with kanamycin (50 μg/mL) and chloramphenicol (34 μg/mL) on LB agar plates. Starter cultures (10 mL) were grown overnight and used at a 1:2000 dilution to inoculate into LB containing the appropriate antibiotics. Cultures were grown at 37 °C until mid-log phase (OD_{600} ~0.5), at which point protein expression was induced by the addition of IPTG (isopropyl-β-D-thiogalactopyranoside, 0.4 mM final concentration). Cultures were shaken for an additional 17 h at 22 °C and harvested by centrifugation; the cell pellets were stored at -20 °C until purification.

For MBP-BamL and point mutant-bearing proteins, the cell pellets were resuspended in lysis buffer [50 mM Tris pH 7.5, 500 mM NaCl, and 2.5% glycerol (v/v)] with the following protease inhibitors: phenylmethanesulfonyl fluoride, benzamidine, leupeptin, and E64. The cells were lysed by sonication (3 × 30 s, continuous mode, < 20 W, 4 °C), followed by centrifugation at 38,000 × *g* for 45 min at 4 °C. The cleared lysates were applied to an amylose resin (NEB) pre-equilibrated with 5 column volumes of lysis buffer (lacking the protease inhibitors), washed with 20 column volumes of lysis buffer, and eluted by gravity flow with elution buffer [50 mM Tris pH 7.5, 150 mM NaCl, 2.5 % glycerol (v/v), and 10 mM maltose]. The eluates were analyzed by Coomassie-stained SDS-PAGE and the appropriate fractions were concentrated using Amicon centrifugal filters (50 kDa molecular weight cut-off, Millipore) with storage buffer [20 mM Tris pH 7.5, 50 mM NaCl, and 10% glycerol (v/v)]. Protein concentration was quantified by absorbance at 280 nm and Bradford analysis (Thermo Scientific). Typical protein yield was 3-5 mg/L, except for the destabilized mutants, which were purified at levels approximately 20- to 40-fold lower. Proteins were stored at -80 °C until use (Figure S6). Because the BpumL protein was His-tagged but not MBP-tagged, the cleared lysate was applied to a Ni-NTA column, washed, and protein fractions were eluted by gravity flow under a standard imidazole gradient. The fractions containing BpumL were pooled, concentrated, and stored at -80 °C until needed.

TEV cleavage of MBP-BamL. Purified MBP-BamL was digested with TEV protease (1 µg TEV to 12 µg MBP-BamL) for 17 h at 4 °C. The TEV-cleaved BamL protein was applied to Ni-NTA superflow resin (Qiagen) pre-equilibrated with 10 column volumes of His-lysis buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, and 15 mM imidazole) and washed with 6 column volumes of B buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, and 30 mM imidazole). The fractions containing BamL only were pooled and concentrated using Amicon centrifugal filters (10 kDa molecular weight cut-off, Millipore) with storage buffer. The purified protein was quantified by the methods described above. This procedure produced highly pure (~99%) protein, as assessed by Coomassie-stained SDS-PAGE (Figure S6).

Methyltransferase Reactions. Unless otherwise indicated, methyltransferase reactions consisted of MBP-BamL (or BpumL, 10 µM), Pfs (SAH nucleosidase, 10 µM), desmethylPZN (50 µM), and SAM (3 mM). Reactions were carried out in Tris buffer (50 mM, pH 7.8) in 20 µL volumes. Reactions were initiated by the addition of SAM and allowed to proceed for 16 h at 37 °C. Reaction mixtures lacking SAM and/or enzyme were carried out as controls. Reactions were quenched by the addition of 40 µL of MeCN, which caused MBP-BamL and other buffer components to precipitate. Insoluble material was removed by centrifugation at 4 °C (16,000 × *g*, 20 min). The PZN-containing supernatants were decanted, transferred to microfuge tubes, and concentrated with an Eppendorf Vacufuge Concentrator (low heat). For samples containing non-PZN derived peptides, the reaction conditions were identical except that the peptide concentration was increased (up to 500 µM). Reaction mixtures were desalted via C_{18} ZipTip (Millipore) according to the manufacturer's instructions and eluted directly onto the MALDI target using 4 µL of a saturated solution of sinapic acid in 50% MeCN containing 0.1% TFA.

MALDI-TOF-MS Analysis. All samples were analyzed using a Bruker Daltonics UltrafleXtreme MALDI-TOF/TOF MS. For desmethylPZN and its analogs, 10 µL of MeCN was added to a microfuge tube to reconstitute lyophilized PZN products. The sample was then transferred onto the MALDI target and mixed with a saturated solution of α-cyano-4-hydroxycinnamic acid in 75% MeCN with 0.1% TFA. The samples were then allowed to dry under ambient conditions. All analyses were conducted using positive reflector mode. The instrument was calibrated daily before any data acquisition using a peptide calibration kit (AB SCIEX). To minimize erroneous readings, laser intensity was kept to a minimum and constant between all samples. Relative activity was determined by comparing the ion intensities of the starting material to product using the FlexAnalysis program. Although this analysis was performed in at least triplicate, we have opted to report relative activities in a semi-quantitative nature, mainly due to heterogeneity within the crystallized MALDI sample spots and subtle spot-to-spot variability. The relative activity scales used are reported in table legends.

Hydrolysis of desmethylPZN. To selectively hydrolyze the single methyloxazoline heterocycle of desmethylPZN back to threonine, thus producing hydrolyzed desmethylPZN (*m/z* 1326), 0.2% (v/v) aq. HCl was added to a solution containing HPLC-purified desmethylPZN. Conversion was completed after 16 h at 22 °C, as assessed by MALDI-TOF-MS. Samples were neutralized, lyophilized, and stored at -80 °C until use.

Determination of optimal of methyltransferase reaction buffer. To increase the yield of the *in vitro* methyltransferase reaction, various conditions and additives were supplied to reaction mixtures as described above. In this effort, temperature (30, 37, and 45 °C), pH (6.5, 7.0, 7.2, 7.5, 7.8, 8.0, and 8.6), and buffer (Tris, HEPES, Na_2HPO_4, and MOPS) were screened. Due to the relative hydrophobicity of desmethylPZN, the effects of organic co-solvents (n-propanol and MeCN) were tested at variable concentrations (1, 5, and 10% v/v). The use of 10 mM of $MgCl_2$ or 1 mM of DTT was also analyzed. Finally, Pfs (SAH nucleosidase) was omitted from reactions to establish if SAH (*S*-adenosyl-L-homocysteine) product inhibition was occurring (Table S1).

MS/MS analysis of reconstituted PZN by FTMS. Samples for high-resolution mass spectrometry were resuspended in 20 µL of 80% (v/v) aq. MeCN. An Advion Nanomate 100 directly infused the sample to a ThermoFisher Scientific LTQ-FT hybrid linear ion trap operating at 11 T (calibrated weekly). The FTMS was operated using the following parameters: minimum target signal counts, 5000; resolution, 100,000; *m/z* range detected, dependent on target *m/z*; isolation width (MS/MS), 5 *m/z*; normalized collision energy (MS/MS), 35; activation *q* value (MS/MS), 0.4; activation time (MS/MS), 30 ms.

Methyltransferase pseudo-single turnover reactions. Reactions were prepared as described above except that the desmethylPZN (25, 50, 100, and 200 µM) and SAM concentrations (50, 100, and 200 µM) were modified. See Figure S3.

Crystallization and Structure Determination. Co-crystals of all complexes were grown by the hanging drop vapor diffusion method under similar conditions. Typically, 1 µL of the BamL-ligand complex (10 mg/mL) was mixed with 1 µL precipitant solution containing 0.1 M NaOAc and 2.0 M $(NH_4)_2SO_4$ (pH 4.6), and the mixture drop was equilibrated over a well containing the same precipitant solution at 8 ºC. The same procedure was followed for SeMet-incorporated BpumL, except that the precipitant solution was 0.1 M HEPES, 0.2 M NaCl, 25% PEG3350 (pH 7.5) and the protein concentration was decreased to 6 mg/mL. Co-crystals were stepwise equilibrated with incremental concentrations of glycerol up to a final concentration of 30% prior to vitrification in liquid nitrogen.

Initial crystallographic phases were determined from a five-fold redundant data set collected from crystals of SeMet-substituted BpumL at the selenium absorption edge utilizing a Mar 300 CCD detector (LS-CAT, Sector 21 ID-D, Advanced Photon Source, Argonne, IL). Subsequent high-resolution data sets for BamL-SAH and BpumL-SAH were collected at Sectors ID-F and ID-G. All data were indexed and scaled using the HKL2000 package (3). Selenium sites were identified using HySS and the heavy atom substructure was imported to SHARP (4) for maximum likelihood refinement and phase calculation, yielding an initial figure of merit of 0.293 to 3.2 Å resolution. Solvent flattening using DM further improved the quality of the initial map (figure of merit = 0.665). Although the map was of marginal quality, a few α-helices could be identified and were manually docked. Further building of this initial model, carried out using Parrot (5) and Buccaneer (6), resulted in the addition of a few β-strands but with minimal registry of the primary sequence. However, this partial model could be successfully used to find a molecular replacement solution for the 1.75 Å resolution BamL data set. Automated and manual rebuilding yielded a nearly complete model. The structure of BpumL was likewise determined by molecular replacement using this partial model, followed by model building and refinement. Cross-validation used 5% of the data in the calculation of the free R factor (7). For each of the structures, stereochemistry of the model was monitored throughout the course of refinement using PROCHECK (8).

Formation of monomethylPZN by BamL-Y182F. MBP-BamL-Y182F (10 μM), Pfs (10 μM), desmethylPZN (50 μM), and SAM (3 mM) were added to Tris buffer (50 mM, pH 7.8) in a 20 μL reaction volume. The reaction was initiated by the addition of SAM and allowed to proceed for 1 h at 37 °C. The reaction was quenched by the addition of 80 μL MeCN, which caused protein precipitation. The sample was centrifuged at 22 °C (17,000 × *g* for 15 min). The supernatant was transferred to a clean microfuge tube and dried by high vacuum. The sample was redissolved in 5 μL MeCN and 1 μL was analyzed by MALDI-TOF-MS with 2 μL saturated solution of α-cyano-4-hydroxycinnamic acid in 75% MeCN. The remaining sample was dried by high vacuum before the addition of MBP-BamL wild-type (10 μM), Pfs (SAH nucleosidase, 10 μM), and SAM (3 mM) in Tris buffer (50 mM, pH 7.8) in a 20 μL reaction volume. The reaction was initiated by the addition of SAM and allowed to proceed for 16 h at 37 °C. The reaction was quenched and prepared for MALDI-TOF-MS analysis as previously described. See Figure S7.

Methyltransferase reactions using an unnatural, ethyl derivative of SAM. "Ethyl SAM", where the electrophilic methyl group of SAM is replaced with an ethyl group, was prepared using an established protocol (9). The crude material was purified by HPLC (Perkin Elmer Flexar HPLC) using a semi-preparative C_{18} column (Thermo BETASIL C_{18} column, 250 mm × 10 mm; pore size: 100 Å; particle size: 5 μm). At a flow rate of 4 mL/min, initial conditions (5% MeOH) were maintained for 6 min, followed by a linear gradient up to 95% MeOH over 25 min. The aqueous phase was 20 mM ammonium formate, pH 3.5.

The fractions containing ethyl SAM were monitored by A_{260}, collected into 20 mL borosilicate vials, concentrated by rotary evaporation, and lyophilzed overnight to yield a white solid (1.67 mg, 16%); ^{1}H NMR (500 MHz, D_2O): δ 8.34 (s, 1H), 8.24 (s, 1H), 6.11 (d, ^3J=4.6 Hz, 1H), 4.91 (m, 1H), 4.35-4.45 (m, 2H), 3.91 (m, 2H), 3.69 (t, ^3J=7.6 Hz, 1H), 3.55 (m, 2H), 2.98 (m, 2H), 2.63 (t, ^3J=7.7 Hz, 3H), 2.02 (q, ^3J=7.4 Hz, 2H); HRMS (ESI) for $C_{16}H_{25}N_6O_5S$, $[M+H]^+$ 413.1601 (experimental); $[M+H]^+$ 413.1607 (theoretical, Δ1.5 ppm). Ethyltransferase reactions were completed as described above except that ethyl SAM (3 mM) was substituted for SAM. To test the substrate acceptance of monoethylPZN, additional ethyl SAM was supplemented into the reaction after 16 h and removal of a sample for MALDI-MS analysis. After an additional 4 h, reactions were quenched and analyzed as previously described. When no additional ethylation was seen, SAM was dosed into parallel ethyltransferase reactions and analyzed following the same procedure.

Sequence alignment. ClustalW alignment of five PZN methyltransferases was performed (10). Local alignment with ClustalW to the respective sequences allowed identification of proposed active site residues and conserved motifs. Percent identity and similarity were calculated based on ClustalW2 alignments. See Figure S4.

Isothermal titration calorimetry. Calorimetry experiments were conducted at 22 ˚C on a VP-ITC titration microcalorimeter (Microcal, Inc., Northampton, MA). The buffer solution was degassed for 10 min prior to use. MBP-BamL was diluted to 40 µM in ITC buffer [50 mM HEPES pH 7.5, 150 mM NaCl, 2.5% (v/v) glycerol]. The reference cell was filled with Millipore water. The sample cell (effective volume = 1.45 mL) was filled with protein and stirred continuously at 270 rpm during the titration. The protein in the sample cell was titrated with 36 aliquots (8 µL each) of 250 µM SAM (or SAH) in the same buffer with a 300 s equilibration period between titrations. The heat of dilution of SAM (or SAH) into buffer was subtracted from the titration data. Integration of the area under each peak in the graph of heat change over time was used to determine the heat produced per injection. The MicroCal version of Origin was used to integrate, baseline correct, and normalize raw data as described elsewhere (11).

Table S1. Effect of buffer components on the efficiency of the BamL methyltransferase reaction.

Protein	Pfs	Buffer (50 mM)	pH	Temp (°C)	Additive	Activity
MBP-BamL	+	Tris	7.8	30	-	+++
BamL (no MBP)	+	Tris	7.8	30	-	++
MBP-BamL	+	Tris	7.8	37	-	+++
MBP-BamL	-	Tris	7.8	37	-	++
MBP-BamL	+	Tris	7.8	37	$MgCl_2$ (10 mM)	+++
MBP-BamL	+	Tris	7.8	37	DTT (1 mM)	+++
MBP-BamL	+	Tris	7.8	37	1% *n*-propanol	++
MBP-BamL	+	Tris	7.8	37	5% *n*-propanol	++
MBP-BamL	+	Tris	7.8	37	10% *n*-propanol	+
MBP-BamL	+	Tris	7.8	37	1% MeCN	+++
MBP-BamL	+	Tris	7.8	37	5% MeCN	++
MBP-BamL	+	Tris	7.8	37	10% MeCN	+
MBP-BamL	+	Tris	7.8	45	-	++
MBP-BamL	+	HEPES	8.0	37	-	++
MBP-BamL	+	Na_2HPO_4	7.2	37	-	+++
MBP-BamL	+	MOPS	6.5	37	-	++
MBP-BamL	+	MOPS	7.0	37	-	+++
MBP-BamL	+	MOPS	7.5	37	-	+++
MBP-BamL	+	MOPS	8.0	37	-	+++
MBP-BamL	+	MOPS	8.6	37	-	++

Activity was determined by MALDI-TOF-MS analysis through a semi-quantitative assay and compared to results obtained with wild-type MBP-BamL in Tris pH 7.8 buffer. Activity scale (as compared to the highest level of product formation observed): +++ denotes > 85% activity; ++ denotes 50-85% activity; + denotes 10-49% activity; - denotes < 10% activity. Percent activity was determined from the relative intensities of peaks corresponding to starting material (desmethylPZN, *m/z* 1308) and product (PZN, *m/z* 1336) by MALDI-TOF-MS after a 16 h reaction. Pfs, SAH nucleosidase.

Table S2. Data collection, phasing, and refinement statistics

	BamL	**BpumL**
PDB Codes	4KVZ	4KWC
Data collection		
Space Group	$C222_1$	$P2_12_12$
a, b, c (Å),	78.1, 80.1, 88.2	67.2, 77.9, 44.9
Resolution (Å)[1]	20-1.75 (1.75-1.7)	50-2.0 (2.07-2.0)
R_{sym} (%)[2]	8.1 (62.7)	9.3 (44.4)
I/σ(I)	15.2 (2.9)	21.9 (2.2)
Completeness (%)	99.8 (99.2)	94.8 (68.9)
Redundancy	7.4 (6.0)	9.2 (5.4)
Refinement		
Resolution (Å)	19.5-1.75	38.9-2.0
No. reflections	26,763	15,777
R_{work} / R_{free}[3]	21.1/24.6	18.9/25.6
Number of atoms		
Protein	2158	2042
SAH	26	26
Water	189	234
B-factors		
Protein	26.8	24.2
SAH	20.1	18.9
Water	35.6	28.9
R.M.S. deviations		
Bond lengths (Å)	0.006	0.007
Bond angles (°)	1.12	1.133

1. Highest resolution shell is shown in parenthesis.
2. $R_{sym} = \Sigma |(I_i - \langle I_i \rangle | \Sigma I_i$ where I_i = intensity of the *i*th reflection and $\langle I_i \rangle$ = mean intensity.
3. R-factor = $\Sigma(|F_{obs}|-k|F_{calc}|)/\Sigma |F_{obs}|$ and R-free is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.

Table S3. Oligonucleotides (primers) used for cloning and site-directed mutagenesis.

#	Primer Name	Primer Sequence (5' – 3')
1	BamLf (cloning)	aaaaggatccatggaaattgaaacaattgtcagagagt
2	BamLr (cloning)	aaaagcggccgctcacgtatacctttgtttttataatccaac
3	BpumLf (cloning)	aagcagccgcatatgatacaagaaaaaatcaaagagcttgaa
4	BpumLr (cloning)	ctagctcgagttagccagaatttttaataagccacgcatgta
5	W20Af	ccaacagaattcaagcacaaacagcgttttctcatcctgagaaaagta
6	W20Ar	tacttttctcaggatgagaaaacgctgtttgtgcttgaattctgttgg
7	F21Af	caacagaattcaagcacaaacatgggcttctcatcctgagaaaagtaaagtg
8	F21Ar	cactttacttttctcaggatgagaagcccatgtttgtgcttgaattctgttg
9	D34Af	agtgagctttcggtatgctgagagggaaacctcat
10	D34Ar	atgaggtttccctctcagcataccgaaagctcact
11	T38Af	cggtatgatgagagggaagcctcatccattagaag
12	T38Ar	cttctaatggatgaggcttccctctcatcataccg
13	T38Ff	gctttcggtatgatgagagggaattctcatccattagaagtatttca
14	T38Fr	tgaaatacttctaatggatgagaattccctctcatcataccgaaagc
15	R42Af	gatgagagggaaacctcatccattgcaagtatttcaattgaaacatttctg
16	R42Ar	cagaaatgtttcaattgaaatacttgcaatggatgaggtttccctctcatc
17	D91Af	cgaactgacaggtattgcttcctctgcccaagcaa
18	D91Ar	ttgcttgggcagaggaagcaatacctgtcagttcg
19	D126Af	tatgcaatatgtgtctaagaagcaagccattatattcatacatctatgttttg
20	D126Ar	caaaacatagatgtatgaatataatggcttgcttcttagacacatattgcata
21	L132Af	tatgtgtctaagaagcaagacattatattcatacatgcatgttttggactctttaag
22	L132Ar	cttaaagagtccaaaacatgcatgtatgaatataatgtcttgcttcttagacacata
23	L132Ff	gtgtctaagaagcaagacattatattcatacatttctgttttggactctttaagaatc
24	L132Fr	gattcttaaagagtccaaaacagaaatgtatgaatataatgtcttgcttcttagaca
25	D161Af	cagatcaatcatgtatctatattgtagccttggacaggaacagtttg
26	D161Ar	caaactgttcctgtccaaggctacaatatagatacatgattgatctg
27	L162Af	atcatgtatctatattgtagacgcggacaggaacagtttgggagag
28	L162Ar	ctctcccaaactgttcctgtccgcgtctacaatatagatacatgat
29	L162Ff	agatcaatcatgtatctatattgtagacttcgacaggaacagtttgg
30	L162Fr	ccaaactgttcctgtcgaagtctacaatatagatacatgattgatct
31	R164Af	ctatattgtagacttggacgcgaacagtttgggagagggc
32	R164Ar	gccctctcccaaactgttcgcgtccaagtctacaatatag
33	Y182Af	gcaatctcgtgaggaagaagccgctttaaaagatcaatatcgtgct
34	Y182Ar	agcacgatattgatcttttaaagcggcttcttcctcacgagattgc
35	Y182Ff	aatctcgtgaggaagaagccttttttaaaagatcaatatcgtgc
36	Y182Fr	gcacgatattgatcttttaaaaaggcttcttcctcacgagatt
37	D185Af	tgaggaagaagcctatttaaaagctcaatatcgtgcttctttaacaa
38	D185Ar	ttgttaaagaagcacgatattgagcttttaaataggcttcttcctca
39	Q186Af	cgtgaggaagaagcctatttaaaagatgcatatcgtgcttctttaacaatgg
40	Q186Ar	ccattgttaaagaagcacgatatgcatcttttaaataggcttcttcctcacg
41	S190Af	cctatttaaaagatcaatatcgtgctgctttaacaatggaagaatttaaacag
42	S190Ar	ctgtttaaattcttccattgttaaagcagcacgatattgatcttttaaatagg

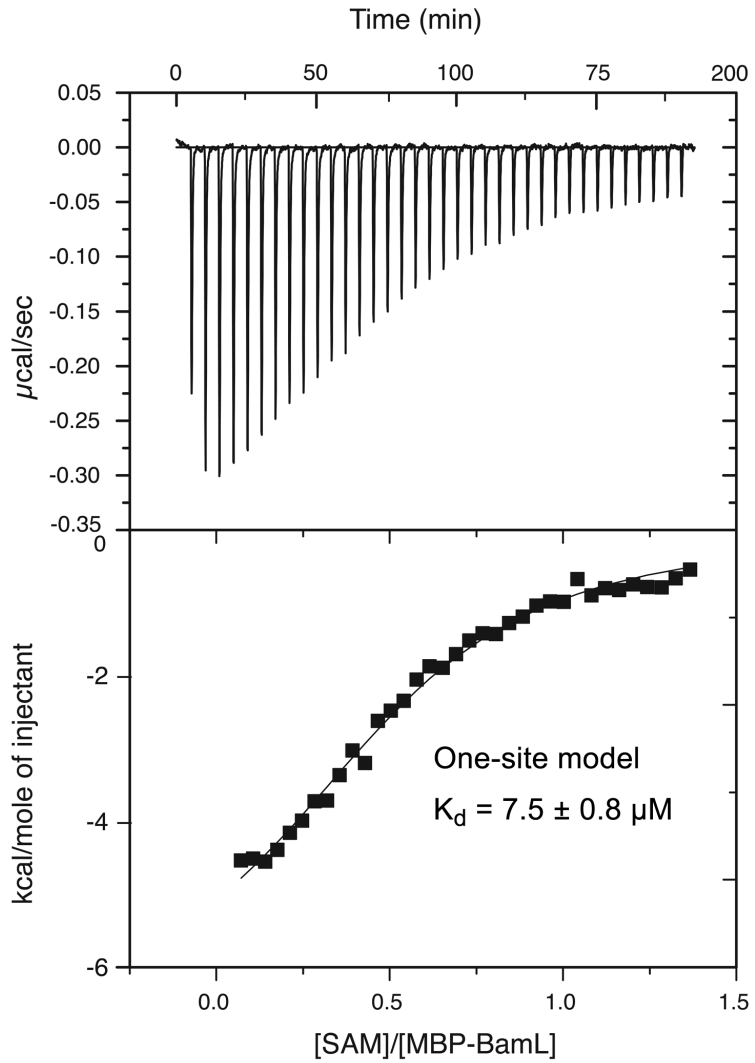


Fig. S1. Isothermal titration calorimetry of MBP-BamL with SAM. A buffered solution of SAM (250 µM) was titrated into a solution of MBP-BamL (40 µM) containing the same buffer. The resulting isotherm data were fit to a single-site model, which gave a dissociation constant of 7.5 µM. Numerous attempts of titrations with SAH gave were unsuccessful, as the shape of the isotherm curve was unusual and the data could not be fit to a reasonable binding model.

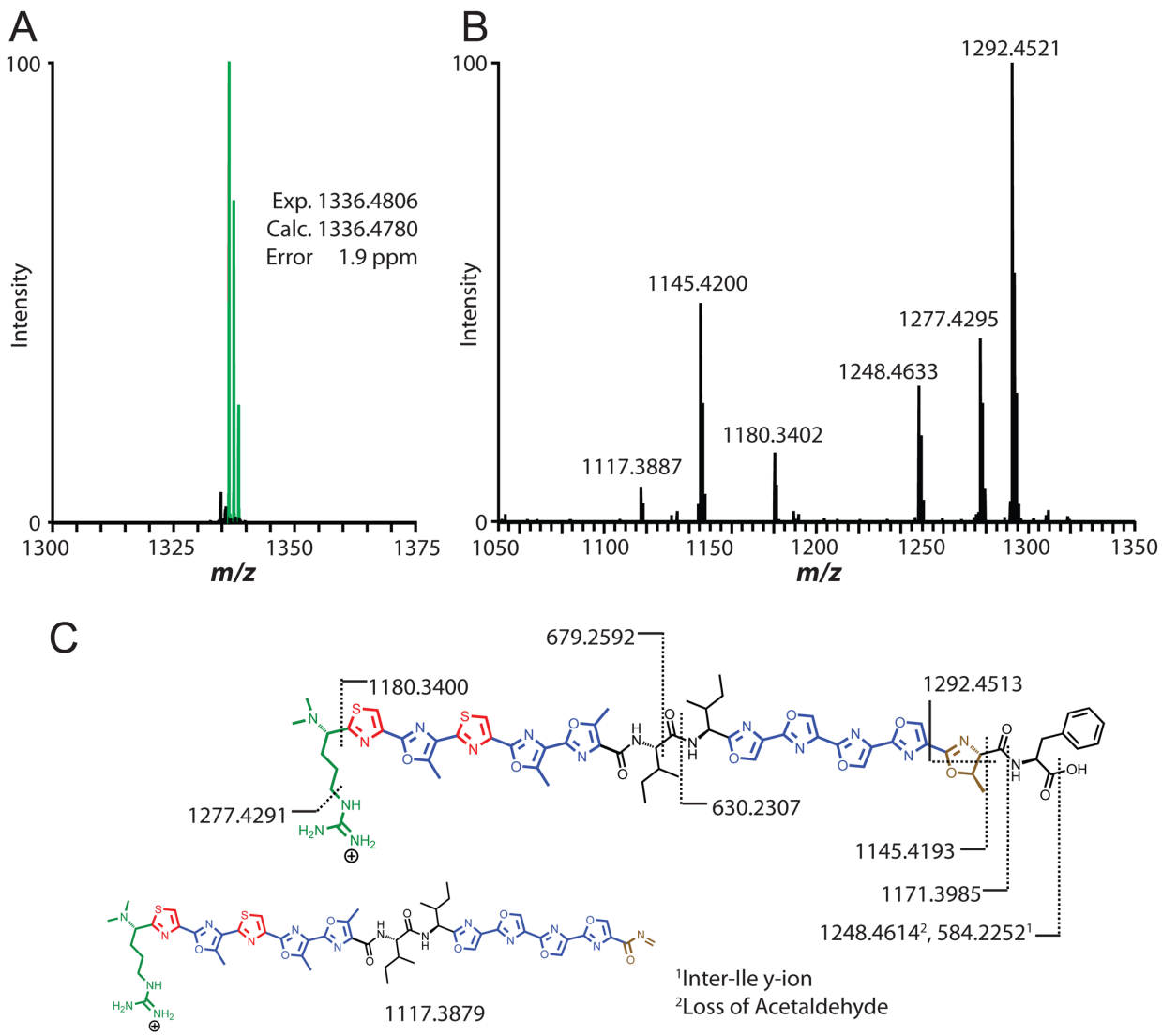


Fig. S2. Mass spectrometric analysis of BamL reaction. ***A*.** High-resolution, Fourier transform MS was used to analyze MBP-BamL reaction products. First, the proteinaceous material was precipitated with MeCN and removed via centrifugation. The supernatant was then taken to dryness, resuspended in 80% MeCN containing 1% formic acid, and directly infused into an 11 T LTQ-FT hybrid linear ion trap using an Advion Nanomate 100. A *m/z* scan showed an ion in the 1+ charge state with an observed monoisotopic *m/z* value with < 2 ppm error of the calculated PZN *m/z*. ***B***. After isolating the peak observed in panel A, collision-induced dissociation fragmentation produced the given spectrum. Again, data were collected in the FTMS. ***C***. Predicted fragments and their calculated monoisotopic *m/z* for observed ions from MS/MS of the BamL reaction product (shown in panel B) were consistent with native PZN. In particular, the peaks at 1277 and 1180 localized the two methylations to the *N*-terminus.

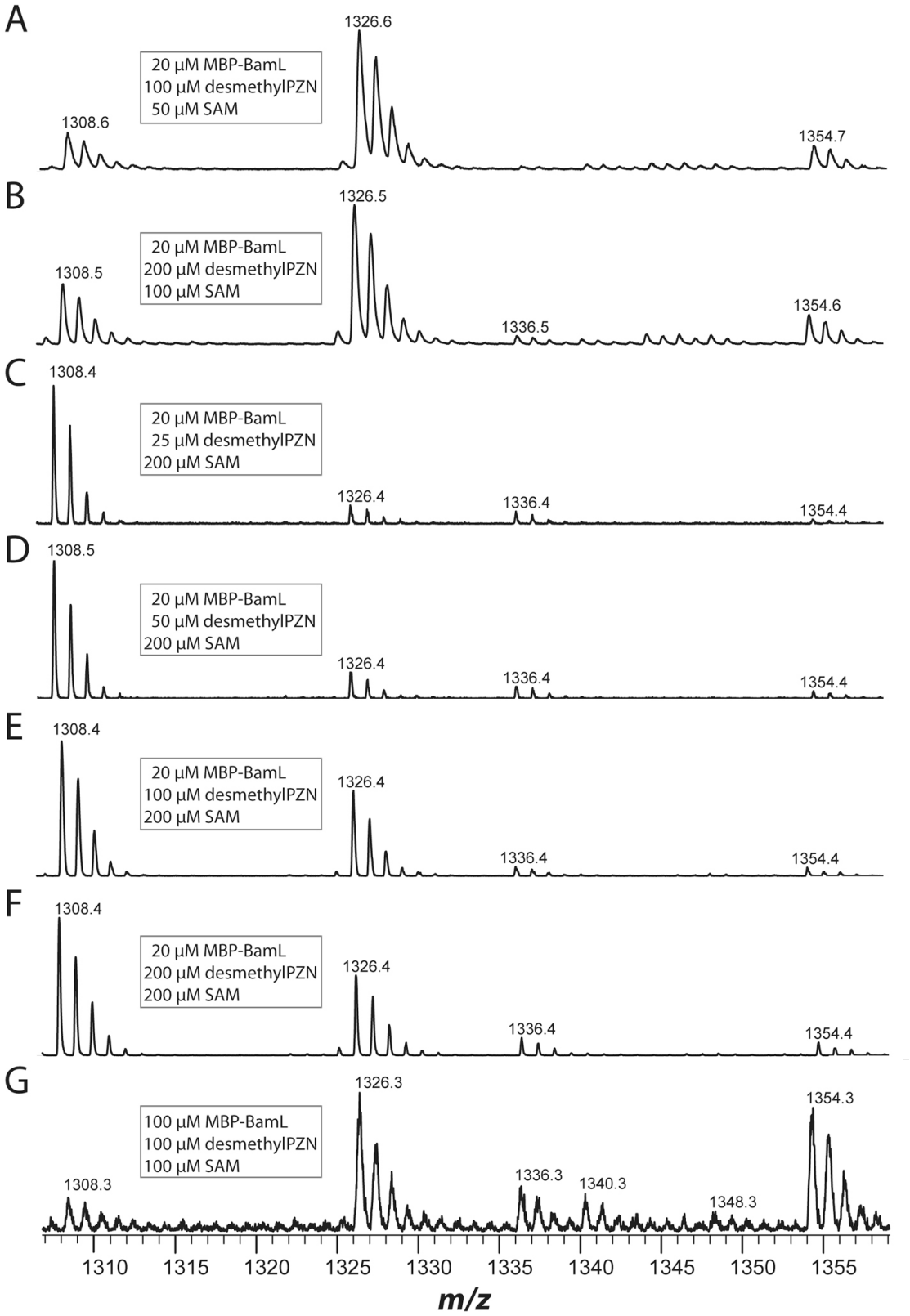
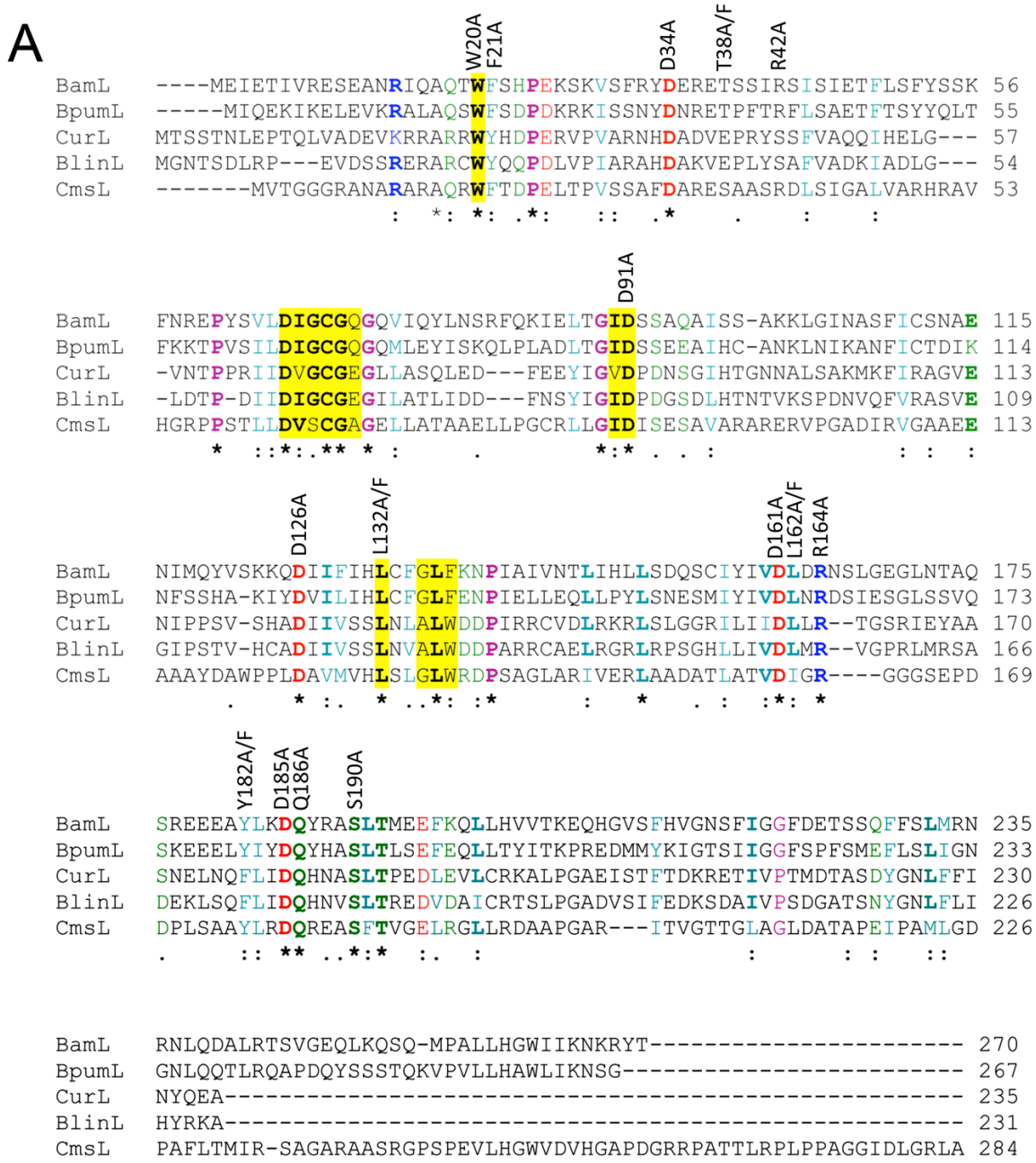


Fig. S3. MALDI-MS of BamL reactions under pseudo-single turnover conditions. All reactions were carried out at 37 °C for 16 h. Specific concentrations of reactants employed are given on individual spectra. Observed in these spectra are desmethylPZN (*m/z* 1308), hydrolyzed desmethylPZN (*m/z* 1326), PZN (*m/z*

1336), and hydrolyzed PZN (*m/z* 1354). Under no condition was monomethylPZN (*m/z* 1322) observed. However, when using MBP-BamL, SAM, and desmethylPZN all at 100 μM, a minor peak consistent with hydrolyzed monomethylPZN (*m/z* 1340) was observed (panel G). Unfortunately, this species was not intense enough for tandem MS verification. Also observed in the panel G spectrum is the sodium adduct of hydrolyzed desmethylPZN (*m/z* 1348), which arises from the large amount of enzyme (100 μM MBP-BamL) used for the reaction (also provides an explanation for the poorer signal to noise ratio). Intriguingly, when the concentration of SAM was at or below 100 μM (panels A, B, and G), the extent of hydrolyzed desmethylPZN dramatically increased. This suggests that BamL may catalyze the hydrolysis of desmethylPZN when SAM is not enzyme-bound.

A



B

	Bam	Bpum	Cms	Cur	Blin
Bam	100	48	25	22	18
Bpum	82	100	25	21	16
Cms	60	59	100	20	22
Cur	56	54	54	100	56
Blin	55	52	49	83	100

% Similarity **% Identity**

Fig. S4. (previous page). Amino acid sequence alignment of the PZN methyltransferases. ***A.*** ClustalW alignment. The specific name for each protein is based on the producing organism: BamL, *Bacillus amyloliquefaciens*; BpumL, *Bacillus pumilus*; CurL, *Corynebacterium urealyticum*; BlinL, *Brevibacterium linens*; CmsL, *Clavibacter michiganensis* subsp. *sepedonicus*. As predicted from alignment with a closely related X-ray structure (Protein Data Bank entry 3E8S), the residues highlighted in yellow are predicted to comprise the SAM-binding pocket. Other color-coding represents conserved basic (blue), acidic (red), hydrophobic (teal), and polar (green) residues. Conserved Gly/Pro are magenta. When 4/5 sequences contain an identical residue, they are shown in boldface. Overall, nineteen mutations, highlighted above each residue, were selected for mutations to probe their relevance in BamL enzymatic activity. ***B***. Similarity (blue) identity (green) matrix from pairwise alignments of PZN methyltransferase amino acid sequences.

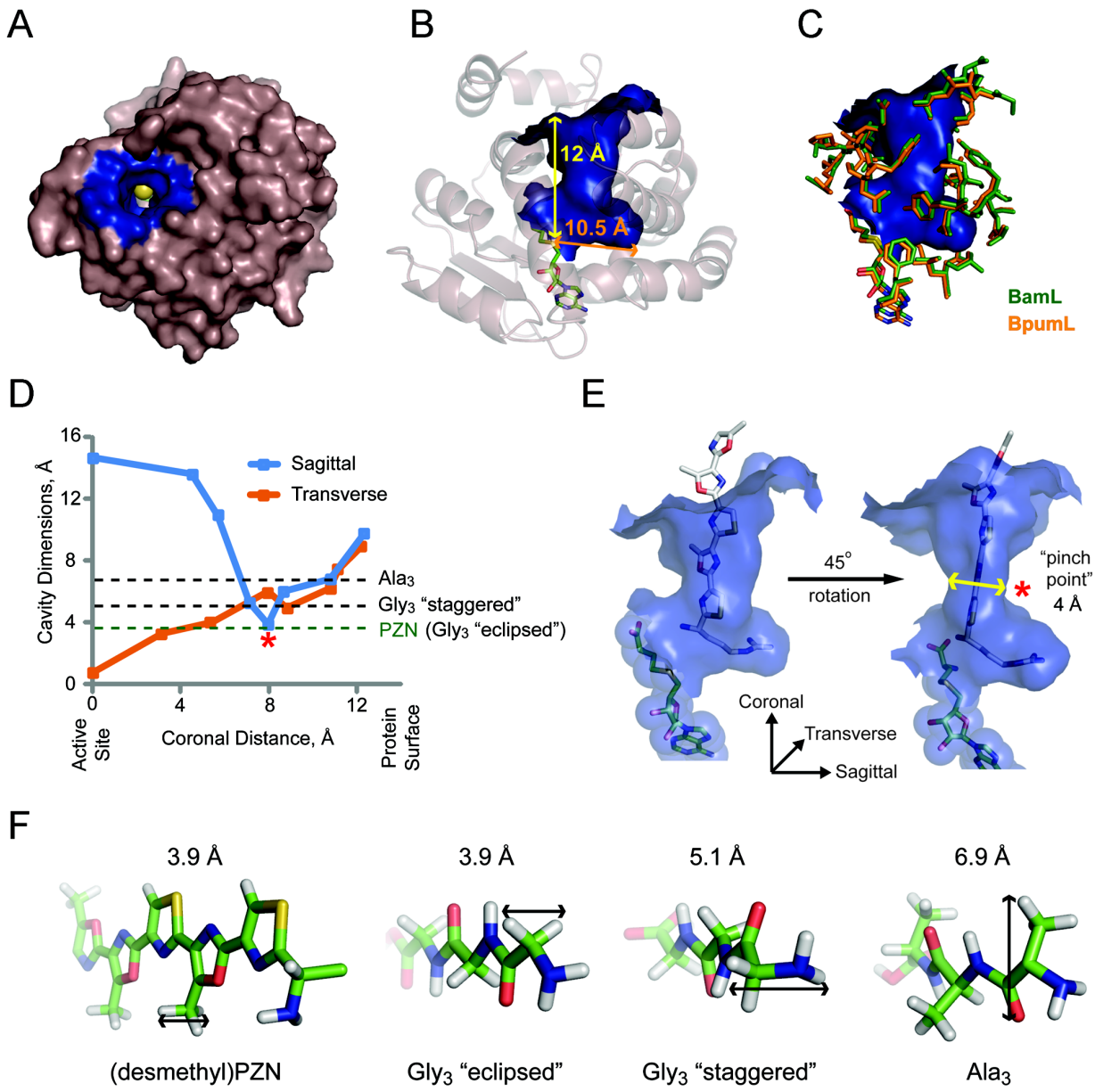


Fig. S5. Dimensions and modeling of the putative substrate-binding tunnel. ***A.*** Surface-rendering of BpumL viewed down the axis of the putative substrate-binding tunnel. The sulfur atom of SAH (yellow sphere) is visible at the floor of the tunnel. Residues that define the walls of the tunnel are blue. ***B.*** BpumL ribbon diagram superimposed on the surface display of residues defining the tunnel walls (blue). SAH is visible at the floor of the tunnel and is colored by element. The shortest distance from the sulfur atom of SAH to the protein surface is ~12 Å (15 Å if through the center of the tunnel) and 10.5 Å to the end of a perpendicularly oriented cavity (presumably occupied by the Arg sidechain). ***C.*** A structural alignment of the residues defining the tunnel walls (blue surface) of BamL (green sticks) and BpumL (orange sticks) demonstrates a striking similarity between the two proteins. ***D.*** Dimensions of the BpumL tunnel. The distance between atoms lining the tunnel walls were measured and adjusted for their respective van der Waals radii to obtain the cavity dimensions (see panel E for orientation). The red asterisk indicates the "pinch point" (*i.e.* the

narrowest part of the tunnel as measured in the BpumL crystal structure). We postulate that such a structure helps to define the extreme substrate selectivity of PZN methyltransferases. Dashed lines indicate the width of potential substrates, omitting Arg for clarity (desmethylPZN, RGGG, and RAAA, see also panel F). (Desmethyl)PZN, and a coplanar, higher energy conformation of Gly_3 (where the C=O substituent of an amide bond "eclipses" the N-H of an adjacent amide bond in a Newman projection), will fit past the pinch point as measured from the crystal structure; however, the tunnel dimensions could be different when the protein is in solution. Gly_3, where the abovementioned substituents are in a "staggered" (lower energy) conformation, is considerably wider due to pleating of the backbone and would presumably be less likely to pass through the pinch point. With additional steric interactions from β-methyl groups, Ala_3 would struggle to fit regardless of conformation. The *N*-terminal Arg may pass the pinch point by coplanar alignment of its side chain with the polyazole moiety, rendering the conformation (and width) similar to eclipsed Gly_3. ***E*.** Manual docking of desmethylPZN (white sticks) into the putative substrate-binding tunnel of BpumL (blue surface). Shown also are the axis definitions that were used to measure tunnel dimensions. Rotating the view 45^o clearly shows the pinch point. We predict that coplanarity is important for passing through the pinch point, but additional flexibility not depicted by a crystal structure may alter the dimensions of this structural filter. ***F*.** Dimensions of potential substrates (desmethyl)PZN, RGGG, and RAAA (Arg omitted). The width of each substrate was measured (arrows) and adjusted for van der Waals radii (see also panel D). This figure was generated using PyMol, Microsoft Excel, and Adobe Illustrator/Photoshop.

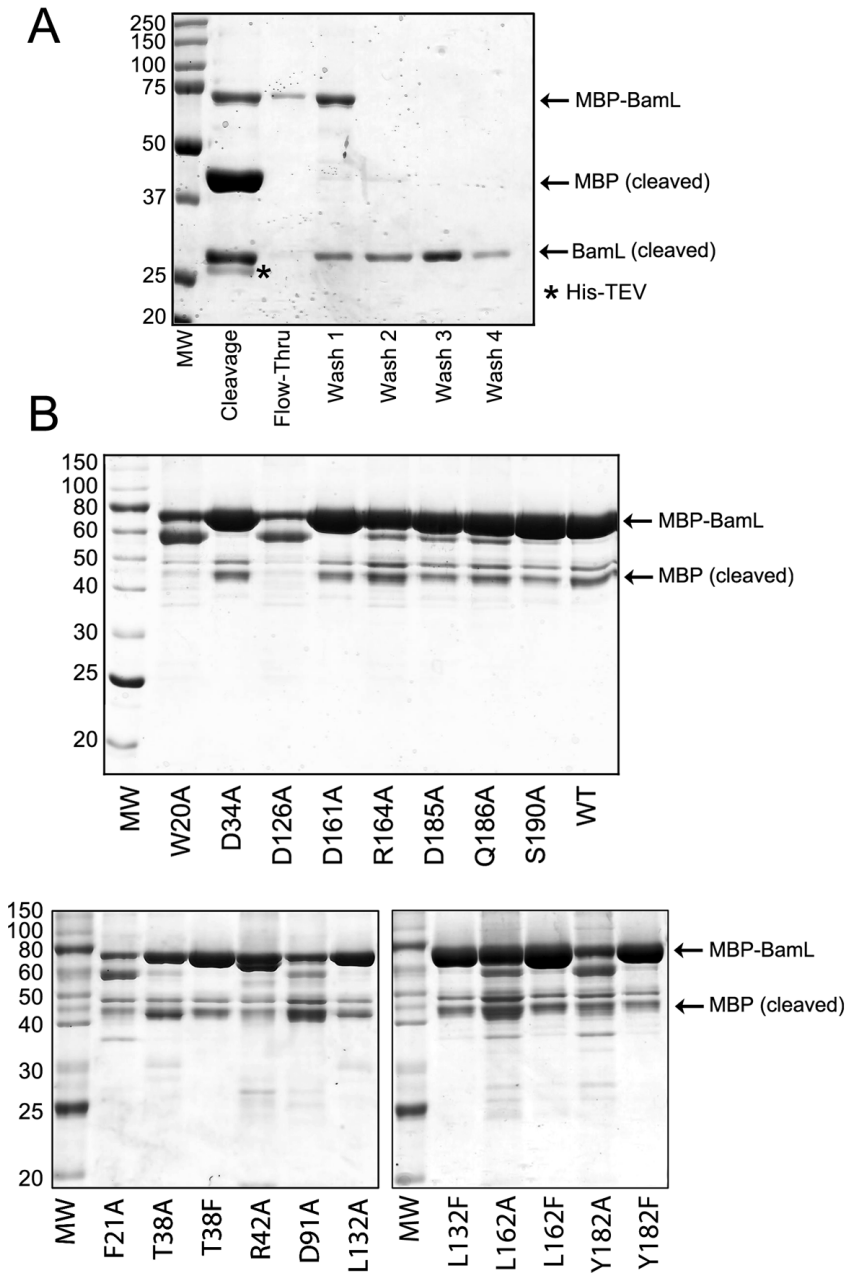


Fig. S6. Coomassie-stained SDS-PAGE gels of MBP-BamL, TEV cleavage reaction, and mutants. Shown is the relative purity of enzymes used for this study. Full length, maltose-binding protein (MBP)-tagged BamL proteins (76 kDa) were purified using amylose affinity chromatography. ***A***. Results of the TEV cleavage reaction. His-tagged TEV protease was added to MBP-BamL (the MBP portion also being His-tagged). A negative Ni-NTA purification was then performed, which bound TEV and MBP but allowed cleaved BamL to pass through the column. This protein was used for crystallization. ***B***. Relative purity of MBP-BamL site-directed mutants (10 µg per lane loaded). These proteins were not TEV treated. During overexpression in *E. coli*, a low but detectable amount of proteolysis of MBP (45 kDa) was observed for all proteins, including wild-type (WT) BamL. As shown in Table S1, the presence of the MBP tag actually enhanced the catalytic reaction, presumably by stabilizing the BamL fusion partner. Such effects of *N*-terminal MBP tags are well

known (12). Mutation of W20A, F21A, D91A, D126A, and Y182A led to structurally destabilized proteins, as indicated by their greater proteolytic processing (yield was also 20- to 40-fold lower than the other mutants). Because those residues are of structural importance, their contribution to catalysis cannot be assessed at the present time.

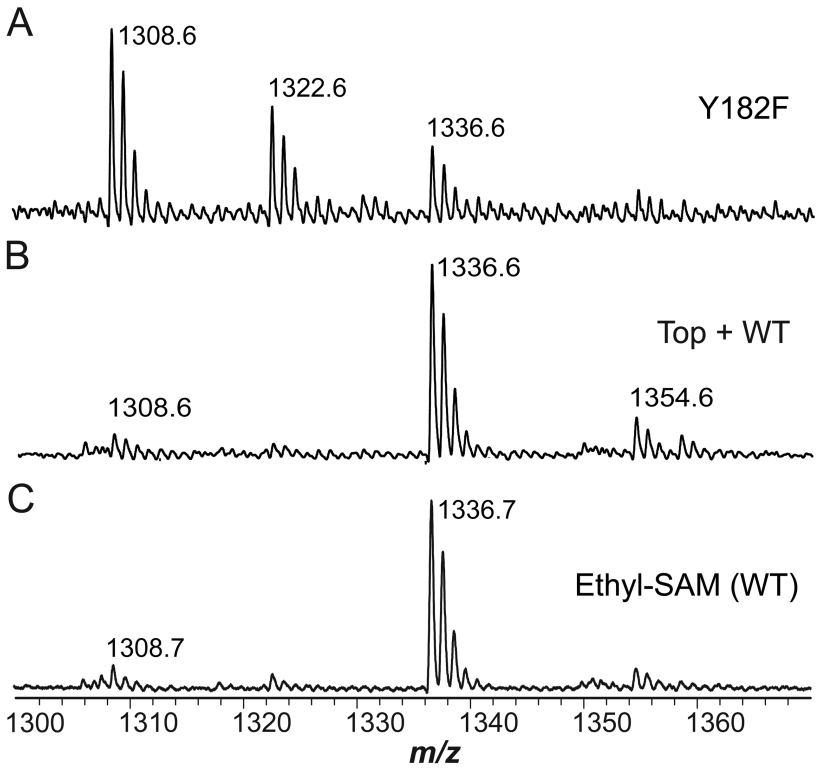


Fig. S7. Observation of monoalkylated PZN analogs. ***A.*** MALDI-TOF-MS was employed to monitor the conversion of desmethylPZN (*m/z* 1308) through monomethylPZN (*m/z* 1322) to PZN (*m/z* 1336) by the BamL mutant, Y182F. ***B.*** Addition of wild-type BamL to the reaction shown in panel A demonstrated product convergence to PZN, thus demonstrating monomethylPZN as a discrete, on-pathway substrate. ***C.*** Upon replacing SAM with an ethyl-bearing derivative, wild-type BamL produced monoethylPZN (*m/z* 1336) but not diethylPZN (*m/z* 1364).

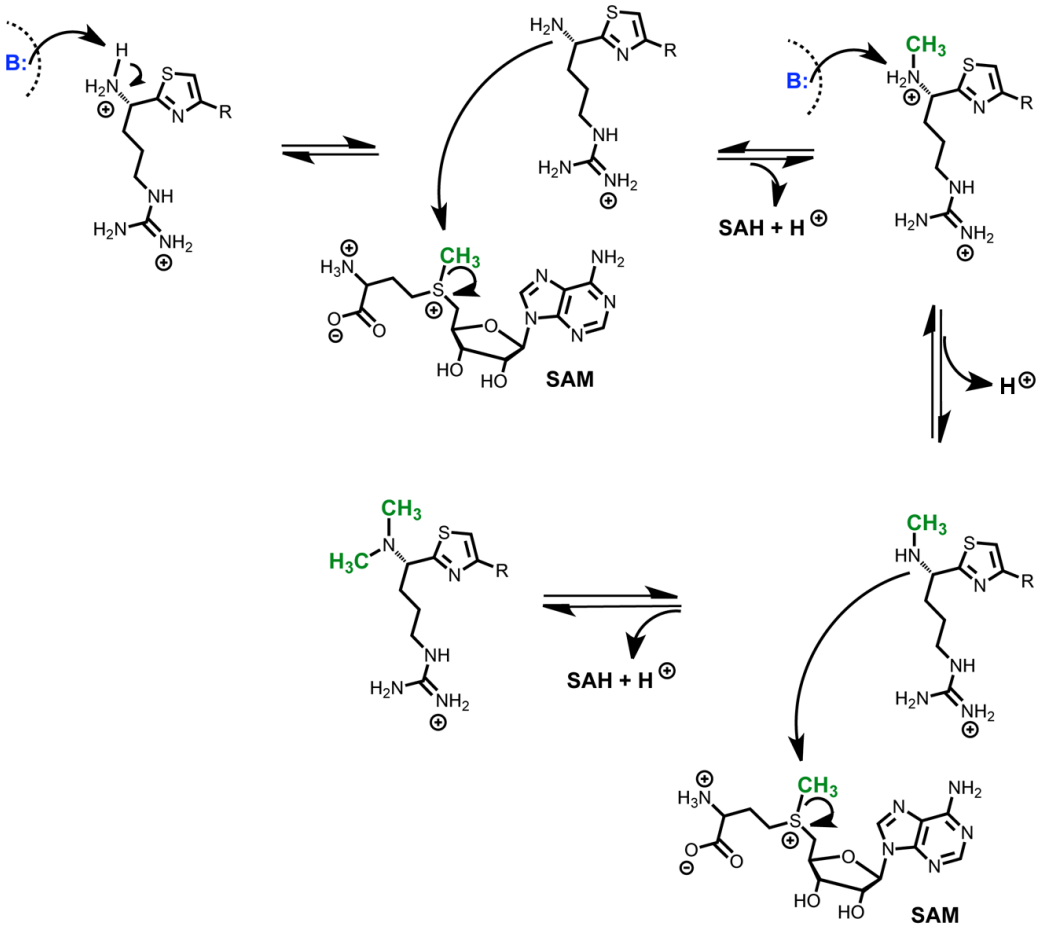


Fig. S8. Proposed BamL mechanism. Upon binding of desmethylPZN to BamL, a catalytic base (blue) deprotonates the positively charged amino terminus of desmethylPZN. The resultant primary amine then attacks the electrophilic methyl group of SAM to yield SAH and a secondary amine (monomethylPZN). The SAH byproduct must then dissociate from the BamL active site to allow binding of another molecule of SAM. Unlike most methyltransferases, BamL does not exhibit potent inhibition by SAH (as evidenced by significant catalytic activity in the absence of SAH nucleosidase, see Table S1). After another round of deprotonation, this time on the amino terminus of monomethylPZN, the second methylation event occurs. Based on our difficulties in observing monomethylPZN, even under limiting SAM, we speculate that the second methylation reaction occurs more rapidly than the first.

Supporting References

1. Sambrook J & Russell DW (2001) *Molecular Cloning: A Laboratory Manual.* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor).

2. Scholz R*, et al.* (2011) Plantazolicin, a novel microcin B17/streptolysin S-like natural product from Bacillus amyloliquefaciens FZB42. *J Bacteriol* 193(1):215-224.

3. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromolec Crystallogr A* 276:307-326.

4. Bricogne G, Vonrhein C, Flensburg C, Schiltz M, & Paciorek W (2003) Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr D Biol Crystallogr* 59:2023-2030.

5. Zhang KY, Cowtan K, & Main P (1997) Combining constraints for electron-density modification. *Methods Enzymol* 277:53-64.

6. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* 62:1002-1011.

7. Kleywegt GJ & Brunger AT (1996) Checking your imagination: Applications of the free R value. *Structure* 4(8):897-904.

8. Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, & Thornton JM (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8(4):477-486.

9. Islam K, Zheng W, Yu H, Deng H, & Luo M (2011) Expanding cofactor repertoire of protein lysine methyltransferase for substrate labeling. *ACS Chem Biol* 6(7):679-684.

10. Larkin MA*, et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947-2948.

11. Wiseman T, Williston S, Brandts JF, & Lin LN (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* 179(1):131-137.

12. Kapust RB & Waugh DS (1999) Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* 8(8):1668-1674.