Supporting Information

Catalytic promiscuity of a bacterial α-N-methyltransferase

Qi Zhang and Wilfred A. van der Donk+

Chemicals

HCTU (O-(1H-6-chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate), Fmoc amino acids, and resins for solid phase peptide synthesis (SPPS) were purchased from Advanced ChemTech, Novabiochem, or Chem-Impex. Solvents commonly used in peptide synthesis and purification, including dimethylformamide (DMF), dichloromethane (DCM), trifluoroacetic acid (TFA), and acetonitrile (MeCN) were obtained in HPLC grade or better and used directly without further purification. Triisopropylsilane (TIPS), piperidine, and N-methyl morpholine were purchased from Acros. Other chemicals were purchased from Fisher Sciencific or Sigma Aldrich.

Biochemicals and Strains

Cosmid pIJ12404, which contains the entire cypemycin biosynthetic gene cluster [1], was a gift from Mervyn Bibb (John Innes Centre, Norwich, UK). Oligonucleotides were purchased from Integrated DNA Technologies. Restriction endonucleases, DNA polymerases and T4 DNA ligase were purchased from New England Biolabs or Invitrogen. S-adenosylhomocysteine hydrolase was purchased from Sigma Aldrich. *E. coli* DH5α was used as host for cloning and plasmid propagation, and *E. coli* BL21 (DE3) was used as a host for expression of C-terminal His-tagged CypM. Media components for bacterial cultures were purchased from Difco laboratories.

General methods

Polymerase chain reactions (PCR) were carried out on a C1000™ thermal cycler (Bio-Rad). DNA sequencing was performed by the Biotechnology Center at the University of Illinois at Urbana-Champaign, using appropriate primers. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) and tandem mass spectrometry (MALDI-TOF MS/MS) was carried out on a Bruker UltrafleXtreme™ MALDI-TOF/TOF mass spectrometer. Liquid chromatography electrospray ionization tandem mass spectrometry (LC/ESI-MS/MS) was carried out and processed using a Synapt ESI quadrupole TOF Mass Spectrometry System (Waters) equipped with an Acquity Ultra Performance Liquid Chromatography (UPLC) system (Waters).

Construct for expressing C-terminally 6 x His-tagged CypM (CypM-His$_6$). A 750bp PCR product containing *cypM* was amplified by PCR using the primers 5'-AAA CAT ATG AGT GAC CCG AGC GTG TAC GAC-3' and 5'-AAA CTC GAG CTG CCG CTC CAG AAC GAC CAG-3' from pIJ12404, a genomic library cosmid that contains the entire cypemycin (CYP) biosynthetic gene cluster [1]. After digestion with NdeI and XhoI, the resultant fragment was purified by agarose gel electrophoresis and cloned into pET24a to yield CypM-6xHis-pET24a for expressing the C-terminally His-tagged CypM (CypM-His$_6$).

Overexpression and purification of CypM-His$_6$. *E. coli* BL21 (DE3) cells were transformed via electroporation with CypM-6 x His-pET24a. A single colony transformant was used to inoculate a 30 mL culture of LB supplemented with 100 µg/mL kanamycin. The culture was grown at 37 °C for 12 h and was used to inoculate 3 L of LB medium containing 100 µg/mL kanamycin. Cells were grown at 37 °C and 220 rpm to an OD600 ≈ 0.6-0.8, and then IPTG was added to a final concentration of 0.1 mM. After additional 8 h of incubation, the cells were harvested by centrifugation at 12,000 × *g* for 15 min at 4 °C. The pellet was resuspended in 30 mL of start buffer (20 mM Tris (pH 8.0), 500 mM NaCl, 1 mM TCEP, 10% glycerol) and were lysed using a high pressure homogenizer (Avestin, Inc.). Cell debris was removed via centrifugation at 23,700 × *g* for 20 min at 4 °C. The supernatant was injected via a superloop onto a fast protein liquid chromatography (FPLC) system (ÄKTA, GE Heathcare Life Sciences) equipped with a 5 mL HisTrap HP IMAC column previously charged with Ni^{2+} and equilibrated in start buffer. The column was washed with 50 mL of buffer A (50 mM imidazole, 20 mM Tris, pH 7.5, 300 mM NaCl and the protein was eluted using a linear gradient of 0-100% B (buffer B = 500 mM imidazole, 20 mM Tris, pH 7.5, 300 mM NaCl) over 40 min at a 2 mL/min flow rate. UV absorbance (280 nm) was monitored and fractions were collected and analyzed by SDS-PAGE (4-20% Tris-glycine READY gel, BioRAD). The fractions containing CypM were combined and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit (10 kDa MWCO, Millipore). Gel filtration purification was used to further purify CypM. The concentrated protein sample was injected onto an FPLC system (ÄKTA) equipped with an XK16 16/60 (GE Healthcare Life Sciences) column packed with SuperDex 75 resin previously equilibrated in 20 mM HEPES (pH 7.5), 200 mM KCl. The protein was eluted with a flow rate of 0.9 mL/min. Both UV (280 nm) and conductance were monitored and fractions were collected. The desired fractions were combined and concentrated using an Amicon Ultra-15 Centrifugal Filter Unit. The resulting protein sample was stored at −80 °C. Protein concentration was determined using a Bradford Assay Kit (Pierce).

Minimum Inhibitory Concentrations (MICs) of nisin and tri-methylated nisin. CypM was incubated with nisin under the conditions described above at 37 ^{o}C for 24-30 h to ensure complete conversion, which was confirmed by MS analysis. Methylated nisin was purified by HPLC using a Beckman Coulter System Gold HPLC equipped with a Grace-Vydac C18 column (5 µm, 300 Å, 250 mm x 4.6 mm). The column was equilibrated with 98% solvent A and 2% solvent B (solvent A = 0.1% TFA in water, solvent B = 0.0866% TFA in 80% ACN/20% water) and eluted at a flow rate of 1.0 mL/min and UV detection at 220 nm using a linear gradient from 98%A/2%B to 60%A/40%B over the first 15 min followed by 10 min of isocratic elution with 60% A/40% B, and a linear gradient to 10% A/90% B over 5 min. Tri-methylated nisin eluted at about 22 min. The product was collected, lyophilized and dissolved in deionized water to provide a stock solution (0.1 mM). The stock solution was serially diluted into 50 µL of LB broth in a 96-well microtiter plate to a final concentration ranging from 20 µM to 0 µM. The test strains *Lactococcus lactis* HP (ATCC11602) and *Bacillus subtilis* 168 (ATCC6633) (100-110 cfu/mL) were then added into each well of the microtiter plate. After incubation at 37 ^{o}C for 24 h, the MICs

were determined to be the lowest concentration of compound that inhibited visible bacterial growth. The MICs of nisin against the same strains were measured in parallel for comparison.

Figure S1 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 1 (AAPAAPA) as substrate.

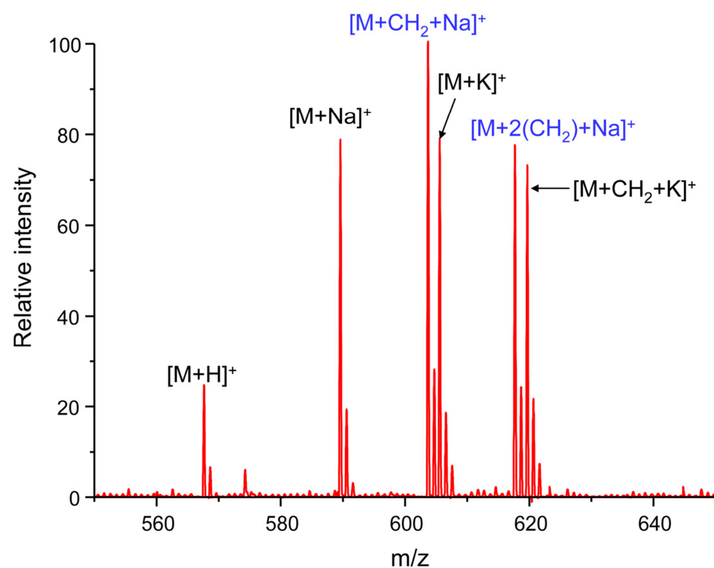

Figure S2 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 2 (AAPAAPS) as substrate.

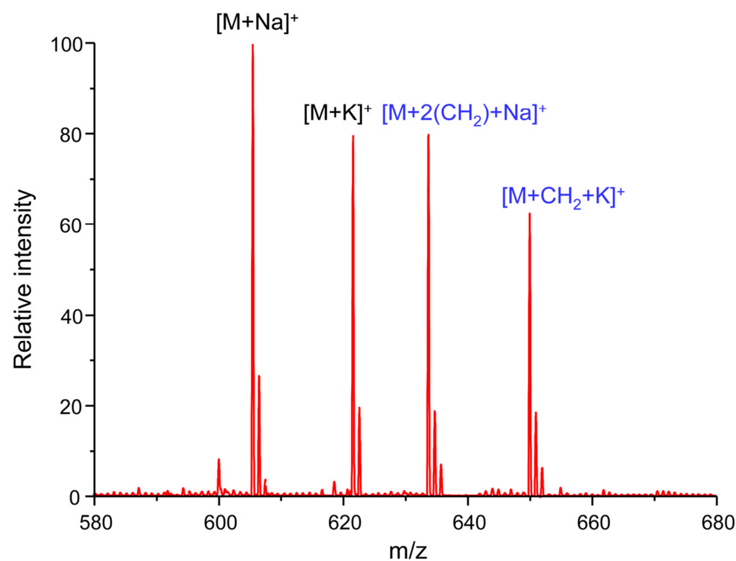

Figure S3 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 3 (ASPAAPA) as substrate.

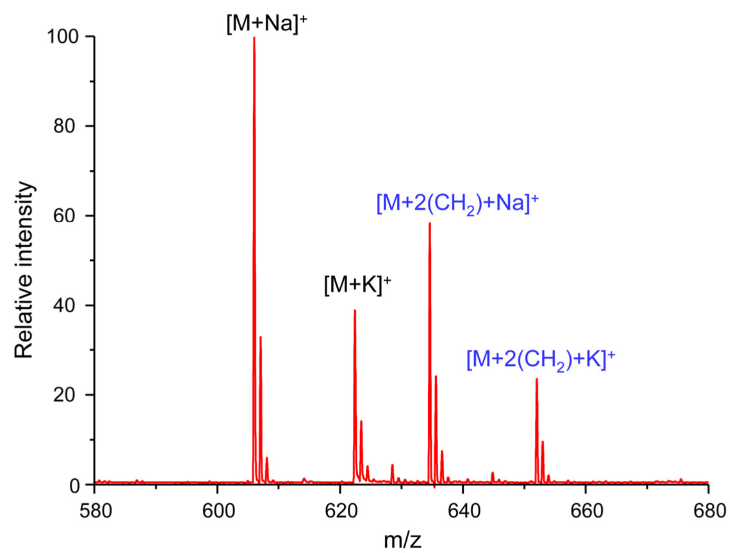

Figure S4 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 4 (ATPAAPA) as substrate.

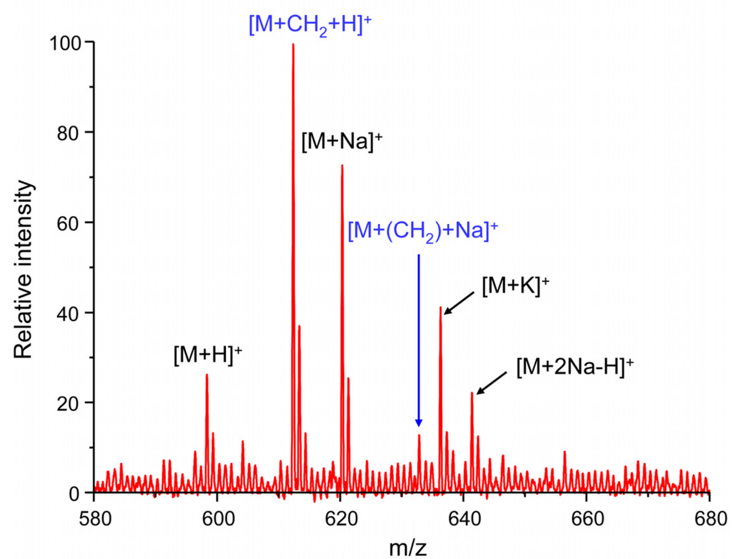

Figure S5 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 5 (ATPATPA) as substrate.

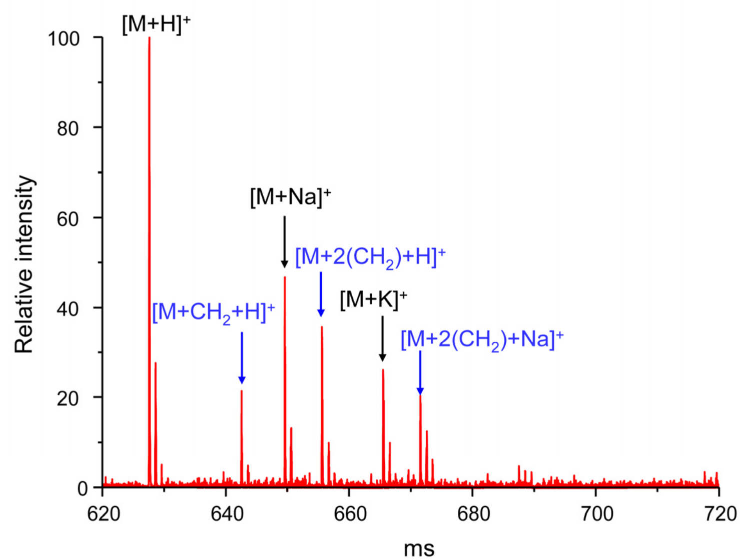

Figure S6 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 6 (GAPAAPA) as substrate.

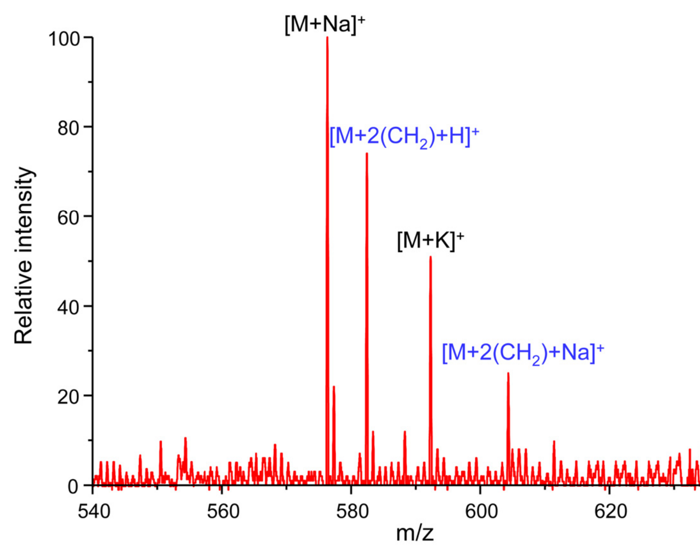

Figure S7 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 7 (SAPAAPA) as substrate.

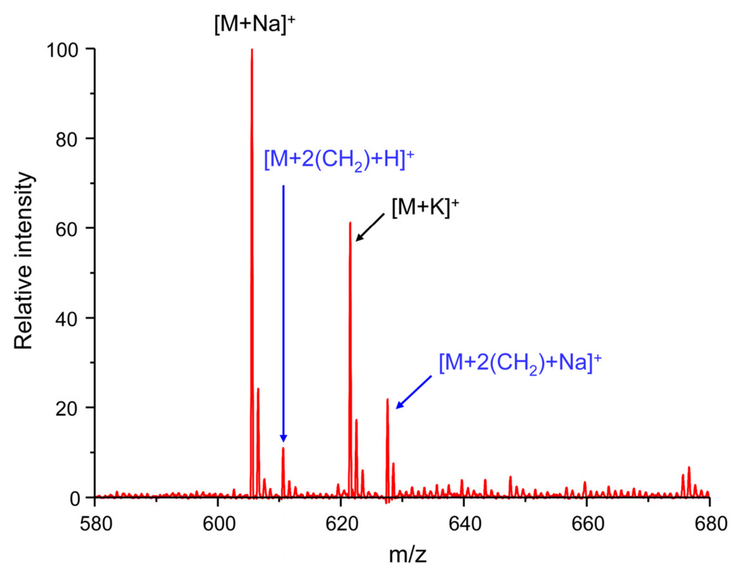

Figure S8 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 8 (MAPAAPA) as substrate.

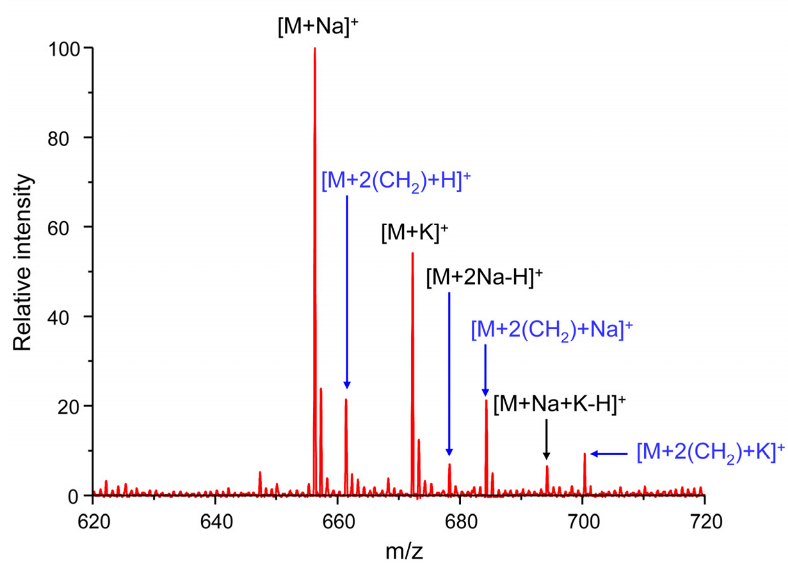

Figure S9 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 9 (KAPAAPA), which was not methylated by CypM.

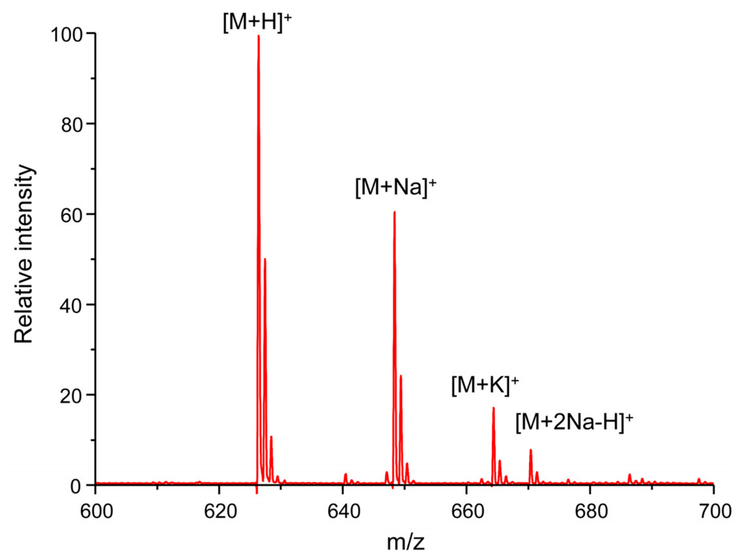

Figure S10 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 10 (AKPAAPA), which was not methylated by CypM.

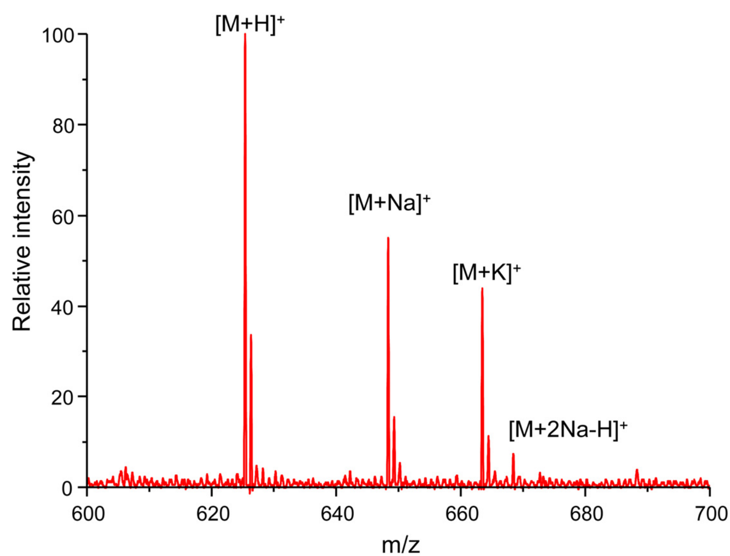

Figure S11 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 11 (AAPAAPK) as substrate.

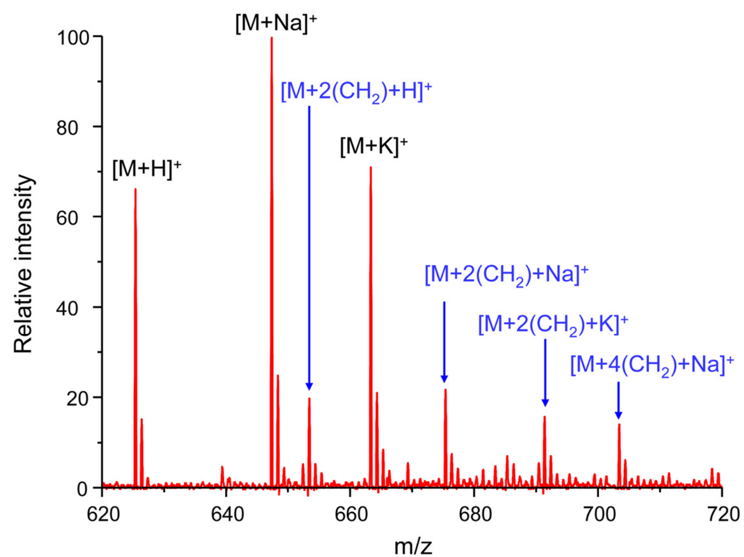

Figure S12 MS-MS analysis of di-methylated peptide 11 (AAPAAPK). The methylated amino groups are highlighted in red.

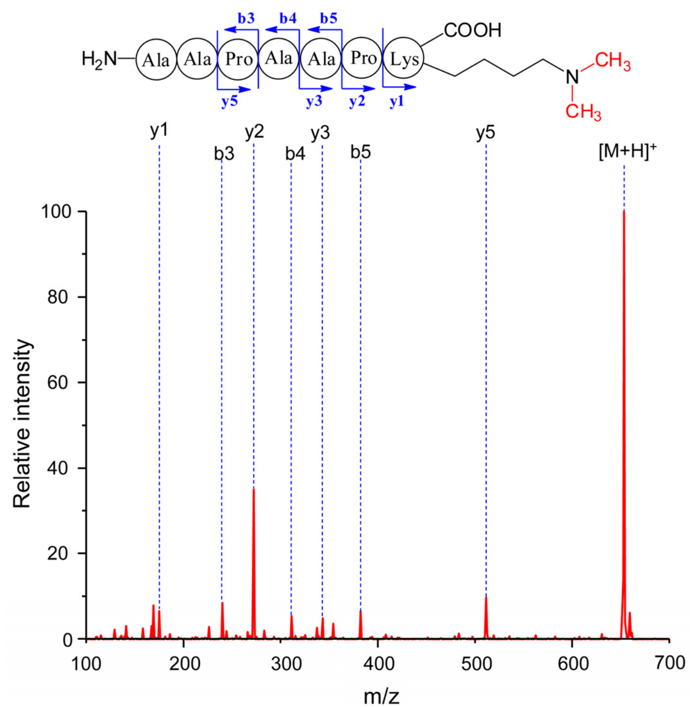

Figure S13 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 12 (AAAATPT), which was not methylated by CypM.

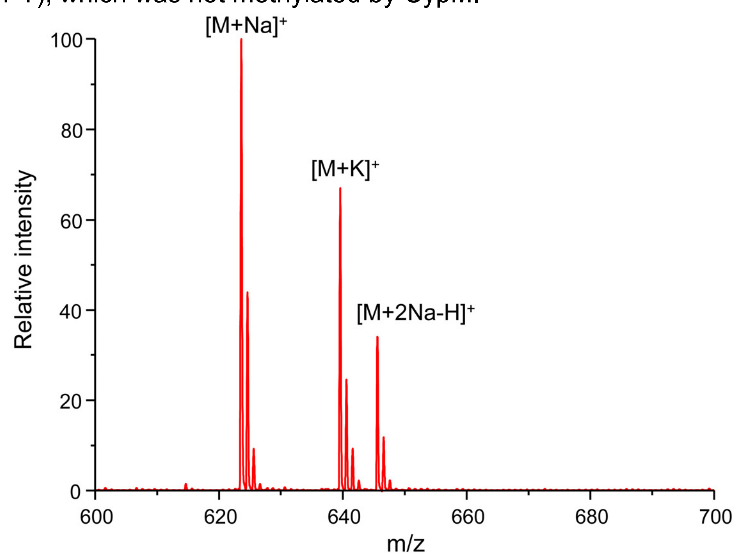

Figure S14 MALDI-TOF mass spectrum of CypM reaction mixture using peptide 13 (APKAAPA), which was not methylated by CypM.

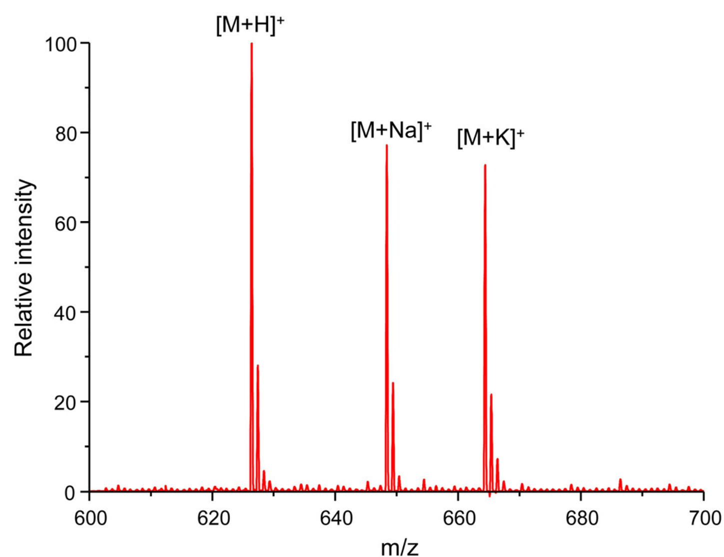

Figure S15 MS-MS analysis of tri-methylated nisin. The methylated amino groups are highlighted in red.

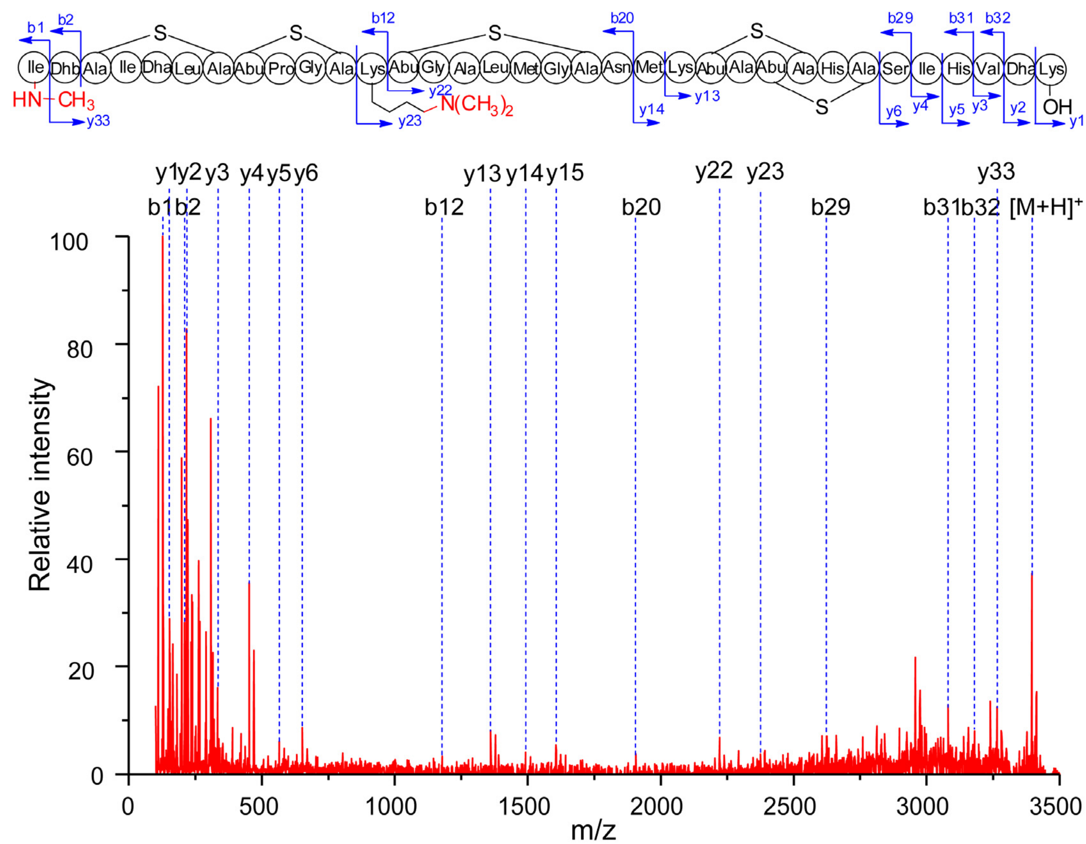

Figure S16 MS-MS analysis of methylated Halα. The disulfide bond of Halα was reduced by TCEP. The methylated amino groups are highlighted in red.

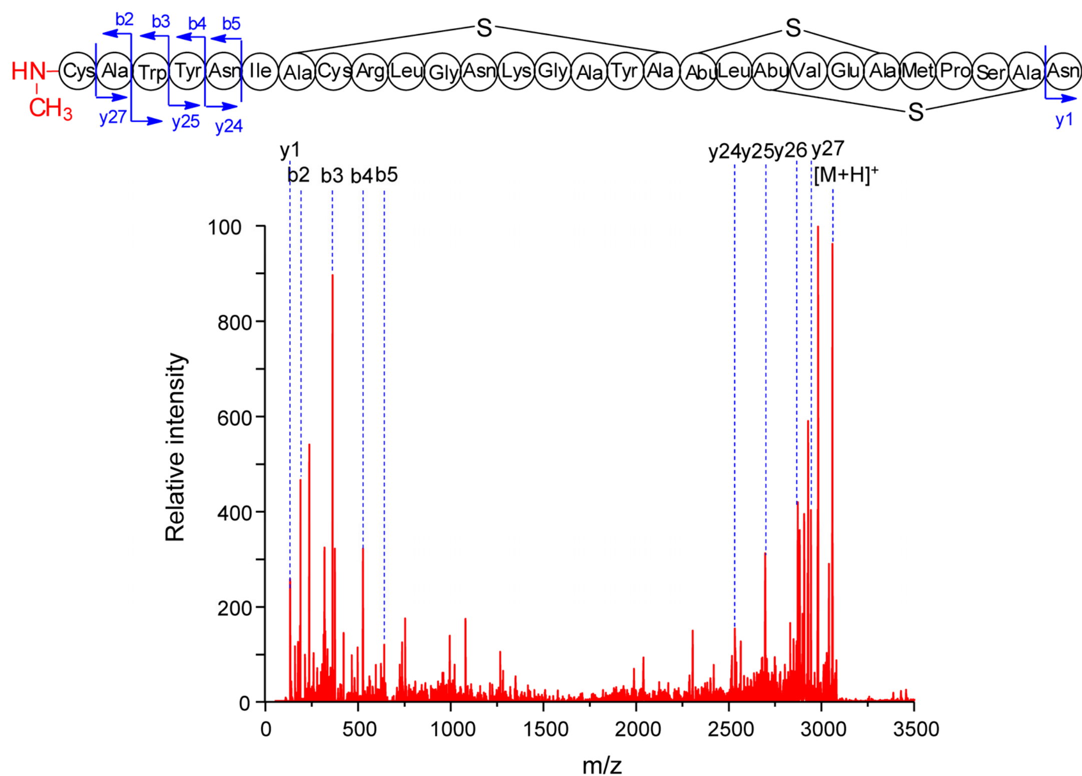

Figure S17 MS-MS analysis of methylated Halβ. The methylated amino groups are highlighted in red.

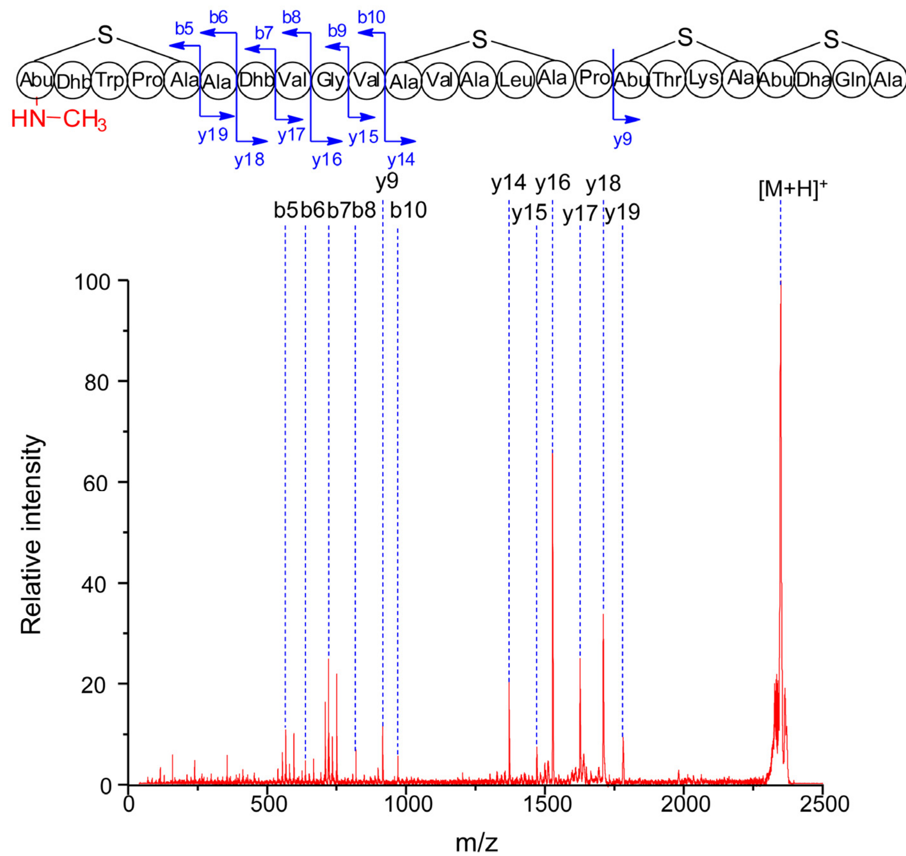

Table S1 Proteins for phylogenetic analysis in this study.

Genebank Accession no	Organism	Function
CCF04252.1	Bacillus amyloliquefaciens subsp. plantarum	Plantozolicin biosynthesis
ADR72966.1	Streptomyces sp. OH-4156	Cypemycin biosynthesis
YP_001827874.1	Streptomyces griseus subsp. griseus	Grisemycin biosynthesis
ZP_07305499.1	Streptomyces viridochromogenes DSM 40736	Methyltransferases for Biosynthesis of other cypemycin-like peptides
YP_001823029.1	Streptomyces griseus subsp. griseus	
CBG75362.1	Streptomyces scabiei 87.22	
ABX00606.1	Streptomyces lincolnensis	Lincomycin biosynthesis
ADB92558.1	Streptomyces caelestis	Celesticetin biosynthesis
YP_004641211.1	Paenibacillus mucilaginosus KNP414	UbiE Ubiquinone biosynthesis
CCF05633.1	Bacillus amyloliquefaciens subsp. plantarum	
YP_004877755.1	Bacillus subtilis subsp. spizizenii	
NP_371995.1	Staphylococcus aureus subsp. aureus Mu50	
CCG45629.1	Halobacillus halophilus DSM 2266	
YP_003518485.1	Pantoea ananatis LMG 20103	
ADR62353.1	Pseudomonas putida BIRD-1	
ABB68305.1	Shigella boydii Sb227	
AAX76384.1	Brucella abortus bv. 1 str. 9-941	
ZP_06655937.1	Escherichia coli B354	
ZP_08027196.1	Actinomyces sp. oral taxon 178	
YP_004225197.1	Microbacterium testaceum StLB037	
ZP_06806555.1	Brevibacterium mcbrellneri ATCC 49030	
YP_121345.1	Nocardia farcinica IFM 10152	
YP_003313526.1	Sanguibacter keddieii DSM 10542	
YP_004599821.1	Cellvibrio gilvus ATCC 13127	
NP_417725.1	Escherichia coli str. K-12	PrmA Ribosomal protein L11 methyltransferase
CAD84565.1	Nitrosomonas europaea	
ADR62163.1	Pseudomonas putida BIRD-1	
YP_003521870.1	Pantoea ananatis LMG 20103	
EHN10704.1	Patulibacter sp. I11	
ADB51263.1	Conexibacter woesei DSM 14684	
BAA17877.1	Synechocystis sp. PCC 6803	
CAO90565.1	Microcystis aeruginosa PCC 7806	
EDK32955.1	Clostridium kluyveri DSM 555	
YP_002531814.1	Bacillus cereus Q1	
YP_003599723.1	Bacillus megaterium DSM 319	
YP_001421966.1	Bacillus amyloliquefaciens FZB42	
YP_004509989.1	Porphyromonas gingivalis TDC60	RsmE 16S ribosomal RNA methyltransferase
EGT70312.1	Escherichia coli O104:H4	
ADX05192.1	Acinetobacter baumannii 1656-2	
YP_005248843.1	Salmonella enterica subsp. enterica	
YP_002236603.1	Klebsiella pneumoniae 342	

YP_002468150.1	Buchnera aphidicola str. 5A
YP_203819.1	Vibrio fischeri ES114
YP_001090945.1	Prochlorococcus marinus MIT 9301
ZP_06382657.1	Arthrospira platensis str. Paraca
YP_476943.1	Synechococcus sp. JA-2-3B
YP_005542264.1	Bacillus amyloliquefaciens TA208
YP_003599722.1	Bacillus megaterium DSM 319
YP_004204315.1	Bacillus subtilis BSn5
YP_004911121.1	Streptomyces cattleya NRRL 8057
NP_626790.1	Streptomyces coelicolor A3(2)
ZP_07313056.1	Streptomyces griseoflavus Tu4000

References

1. Claesen J, Bibb M (2010) Genome mining and genetic analysis of cypemycin biosynthesis reveal an unusual class of posttranslationally modified peptides. *Proc Natl Acad Sci U S A* **107:** 16297-16302