Supplementary Information

Discovery and biosynthesis of bosamycin from *Streptomyces* sp. 120454

Zi Fei Xu,† Sheng Tao Bo,† Mei Jing Wang, Jing Shi, Yang Sun, Rui Hua Jiao, Qiang Xu, Ren Xiang Tan, * Hui Ming Ge**
Table of Contents

Experimental procedures:
Heterologous expression of cosmid pHG06015 in *Streptomyces albus* J1074.
Co-expression of *bsmF*, *bsmG* and *bsmH* in *Streptomyces albus* J1074 /pHG06015 strain.
Isolation and purification of bosamycins from *S*. sp. 120454 wild-type and recombinant strains.
Physical data for bosamycins.
Preparation and analysis of Marfey’s derivatives.
Chemical synthesis of 5-OMe Tyr.
Chemical synthesis of 5-OH Tyr.
Chemical complementation of 5-OMe Tyr or 5-OH Tyr into ΔbsmF, ΔbsmG and ΔbsmH mutant.
Protein expression and purification.
In vitro assay of BsmH.
Biological activity assay of SHP2.

Supplementary tables:
Table S1. Bacterial plasmids and strains.
Table S2. Oligonucleotide primers used in this study.
Table S3. Substrate specificity predictions for the adenylation domains of the NRPSs encoded in the *bsm* gene cluster.
Table S4. 1H (600 MHz) and 13C (150 MHz) NMR data of compound 1 in DMSO-d6.
Table S5. 1H (600 MHz) and 13C (150 MHz) NMR data of compound 2 in DMSO-d6.
Table S6. 1H (600 MHz) and 13C (150 MHz) NMR data of compound 3 in DMSO-d6.
Table S7. 1H (600 MHz) and 13C (150 MHz) NMR data of compound 4 in DMSO-d6.
Table S8. 1H (600 MHz) and 13C (150 MHz) NMR data of compound 5 in DMSO-d6.
Table S9. 1H (600 MHz) and 13C (150 MHz) NMR data of compound 6 in DMSO-d6.
Table S10. 1H (600 MHz) and 13C (150 MHz) NMR data of compound 13 in DMSO-d6.
Table S11. 1H (600 MHz) and 13C (150 MHz) NMR data of compound 14 in DMSO-d6.
Table S12. 1H NMR (400 MHz) data of compound A-E in DMSO-d6.
Table S13. 1H NMR (400 MHz) data of G, I and J in DMSO-d6.
Table S14. Conserved sequence regions in the alignment comparisons of BsmF and other known P450s.

Supplementary figures:
Figure S1. Structures of bosamycins.
Figure S2. LC-MS analysis of L-FDAA and D-FDAA derivatives of the amino acid residues in 1 and 3.
Figure S3. Identification of compound 5 as a novel inhibitor of SHP2.
Figure S4. Generation of the S. sp. 120454 mutant strains.
Figure S5. MS/MS analysis of metabolite extracts from wild type and mutant strains.
Figure S6. Chemical complementation of 5-OMe Tyr into mutants.
Figure S7. Chemical complementation of 5-OH Tyr into mutants.
Figure S8. HPLC analysis of metabolite extracts from mutant strains and gene complementation strains.
Figure S9. Chemical complementation fed L-erythro-β-OH-Asp into ΔbsmC mutant.
Figure S10. Verification of cosmid phG06015.
Figure S11. SDS-PAGE analysis of proteins.
Figure S12. Sequence alignment of BsmF-P450 with other P450 proteins.
Figure S13. Domain analyze of BsmF.
Figure S14. Sequence alignment of BsmC with other β-hydroxylases.
Figure S15. Sequence alignment of BsmB-T1 with other T domains.
Figure S16. Sequence alignment of BsmA-C1 with other C domains.
Figure S17-22. MS/MS analysis of 1-6.
Figure S23-25. MS/MS analysis of 7-9 in ΔbsmF mutant strain.
Figure S26-28. MS/MS analysis of 7-9 in ΔbsmG mutant strain.
Figure S29-31. MS/MS analysis of 7-9 in ΔbsmH mutant strain.
Figure S32-34. MS/MS analysis of 10-12 in ΔbsmC mutant strain.
Figure S35-36. MS/MS analysis of 13.
Figure S37. HRESIMS spectrum of 1.
Figure S38-44. 1D and 2D NMR spectrum of compound 1 in DMSO-d6.
Figure S45. HRESIMS spectrum of 2.
Figure S46-52. 1D and 2D NMR spectrum of compound 2 in DMSO-d6.
Figure S53. HRESIMS spectrum of 3.
Figure S54-62. 1D and 2D NMR spectrum of compound 3 in DMSO-d6.
Figure S63. HRESIMS spectrum of 4.
Figure S64-70. 1D and 2D NMR spectrum of compound 4 in DMSO-d6.
Figure S71. HRESIMS spectrum of 5.
Figure S72-78. 1D and 2D NMR spectrum of compound 5 in DMSO-d6.
Figure S79. HRESIMS spectrum of 6.
Figure S80-86. 1D and 2D NMR spectrum of compound 6 in DMSO-d6.
Figure S87. HRESIMS spectrum of 13.
Figure S88-94. 1D and 2D NMR spectrum of compound 13 in DMSO-d6.
Figure S95. HRESIMS spectrum of 14.
Figure S96-102. 1D and 2D NMR spectrum of compound 14 in DMSO-d6.
Figure S103. HRESIMS spectrum of A.
Figure S104. 1H NMR (400 MHz) spectrum of compound A in DMSO-d6.
Figure S105. HRESIMS spectrum of B.
Figure S106. 1H NMR (400 MHz) spectrum of compound B in DMSO-d6.
Figure S107. HRESIMS spectrum of C.
Figure S108. 1H NMR (400 MHz) spectrum of compound C in DMSO-d6.
Figure S109. HRESIMS spectrum of D.
Figure S110. 1H NMR (400 MHz) spectrum of compound D in DMSO-d6.
Figure S111. HRESIMS spectrum of E.
Figure S112. 1H NMR (400 MHz) spectrum of compound E in DMSO-d6.
Figure S113. HRESIMS spectrum of G.
Figure S114-115. 1H NMR (400 MHz) and 13C (100 MHz) spectrum of compound G in DMSO-d6.
Figure S116. HRESIMS spectrum of I.
Figure S117-118. 1H NMR (400 MHz) and 13C (100 MHz) spectrum of compound I in DMSO-d6.
Figure S119. HRESIMS spectrum of J.
Figure S120-121. $^1$H NMR (400 MHz) and $^{13}$C (100 MHz) spectrum of compound J in DMSO-$d_6$.

References
Experimental Procedures

Heterologous expression of cosmId pHG06015 in Streptomyces albus J1074.

Cosmid library for strain 120454 was constructed according to the protocol. The cosmId pHG06015 that covers partial bsm gene cluster was obtained through PCR screening from 2000 clones using primers listed in Table 3. The cosmId pHG06015 was transferred into the E. coli ET12567 (pUZ8002) strain, and then introduced into Streptomyces albus J1074 by intergeneric conjugation. The recombinant strain was grown on MS agar medium supplemented with 50 μg/mL of apramycin for sporulation. The seed culture was prepared by inoculating incoilete fresh spores into 250-mL baffled flasks containing 50 mL of TSB medium (17.0 g tryptone, 3.0 g soyteine, 2.5 g glucose, 5.0 g sodium chloride, 2.5 g Na₂HPO₄ in 1 L water, pH 7.0) for 1 days at 30 °C and 250 rpm. Subsequently, 15 mL seed cultures were inoculated into 2 L baffled flasks containing 300 mL of the fermentation medium (dextrin 40 g, tomato paste 7.5 g, NZ Amine 2.5 g, primary yeast 5 g in 1 L distilled water, pH 7.0), and incubated for 7 days at 160 rpm and 30 °C. Finally, the fermentation broth was filtered and absorbed with XAD-16 resin. The resin was washed with water and eluted with acetone.

Co-expression of bsmF, bsmG and bsmH in Streptomyces albus J1074 /pHG06015 strain.

A DNA fragment containing three genes of bsmF, bsmG and bsmH was amplified from genomic DNA and subcloned into E. coli–Streptomyces expression shuttle vector pUWL201PWT plasmid to afford pHG06016 plasmid. The plasmid pHG06016 was transformed into E. coli ET12567/pUZ8002 and then introduced into S. albus/pHG06015 strain by conjugation. Clones harboring pHG06016 plasmid were selected by thiostrepton resistance and verified by diagnostic PCR. The resulting recombinant strain was then fermented for 7 days at 30 °C. The crude extract was analyzed by LC-MS and HPLC.

Isolation and purification of bosamycins from S. sp. 120454 wild-type and recombinant strains.

For isolation of compounds, a large scale fermentation (20 L) for wild-type strain was carried out using the same medium as mention above. The resin was harvested after seven days’ cultivation, and extracted with methanol for three times. The combined methanol phases were evaporated to dryness, and the resulting extract was subjected to silica-gel column, and eluted with the mixture of methylene dichloride and methanol (100:1 to 1:1). Fractions were combined according to HPLC analysis, and further separated by Sephadex LH-20 chromatography. Fractions containing the target compounds were finally purified by semi-preparative HPLC. Compounds 1-6 were purified from S. sp. 120454 wild-type strain. Compound 13 were purified from HG06012 strain. Compound 14 were purified from HG06013 strain.

Physical data for bosamycins.

Compound 1: white amorphous solid; [α]₂⁰⁺⁻26.0 (c 0.10, MeOH); UV (MeOH) λₘₐₓ (log ε) 218 (3.84), 228 (3.83), 278 (3.31); NMR data see Table S4; HRESIMS m/z 876.4166 [M+ H]⁺ (calcd for C₅ₐH₅ₙNₐOₐ₁₈, 876.3913).

Compound 2: white amorphous solid; [α]₂⁰⁺⁺30.8 (c 0.01, MeOH); UV (MeOH) λₘₐₓ (log ε) 212 (3.61), 226 (3.57), 279 (3.08); NMR data see Table S5; HRESIMS m/z 1082.4697 [M+ H]⁺ (calcd for C₅ₐH₅ₙNₐOₐ₁₈, 1082.4677).

Compound 3: white amorphous solid; [α]₂⁰⁺⁻20.0 (c 0.05, MeOH); UV (MeOH) λₘₐₓ (log ε) 218 (4.23), 228 (4.27), 281 (3.58); NMR data see Table S6; HRESIMS m/z 1126.4572 [M+ H]⁺ (calcd for C₅ₐH₅ₙNₐOₐ₂ₐ₂, 1126.4575).

Compound 4: white amorphous solid; [α]₂⁰⁺⁺8.7 (c 0.01, MeOH); UV (MeOH) λₘₐₓ (log ε) 218 (4.01), 229 (3.12), 280 (3.49); NMR data see Table S7; HRESIMS m/z 1081.4725 [M+ H]⁺ (calcd for C₅ₐH₅ₙNₐOₐ₂ₐ₂, 1081.4724).

Compound 5: white amorphous solid; [α]₂⁰⁺⁻6.7 (c 0.14, MeOH); UV (MeOH) λₘₐₓ (log ε) 218 (3.88), 229 (3.98), 278 (3.48); NMR data see Table S8; HRESIMS m/z 1140.4743 [M+ H]⁺ (calcd for C₅ₐH₅ₙNₐOₐ₂ₐ₂, 1140.4732).
Compound 6: white amorphous solid; [α]_D^25 +8.7 (c 0.02, MeOH); UV (MeOH) λ_{max} (log e) 218 (4.07), 227 (4.08), 280 (3.64); NMR data see Table S9; HRESIMS m/z 1095.4879 [M+H]^+ (calcd for C_{34}H_{32}N_{10}O_{16}, 1095.4881).

Compound 13: white amorphous solid; [α]_D^25 +8.0 (c 0.03, MeOH); UV (MeOH) λ_{max} (log e) 218 (3.86), 229 (3.97), 281 (3.40); NMR data see Table S10; HRESIMS m/z 1009.4564 [M+H]^+ (calcd for CuH_{40}N_{10}O_{16}, 1009.4513).

Compound 14: white amorphous solid; [α]_D^25 +5.3 (c 0.04, MeOH); UV (MeOH) λ_{max} (log e) 218 (3.65), 227 (3.67), 278 (3.26); NMR data see Table S11; HRESIMS m/z 1039.4631 [M+H]^+ (calcd for CuH_{42}N_{10}O_{17}, 1039.4619).

Preparation and analysis of Marfey’s derivatives.

The absolute configuration of the bosamycins was determined by advanced Marfey’s method. Briefly, Compound 3 (0.5 mg) was dissolved in 6 N HCl (1 mL) and hydrolyzed at 110 °C for 12 h. After cooling, the solution was evaporated to dryness and dissolved in H₂O (100 μL). To this mixture was added a 1% (w/v) solution (200 μL) of 1-fluoro-2,4-dinitrophenyl-5-L-alanine amide (L-FDAA) and 1-fluoro-2,4-dinitrophenyl-5-D-alanine amide (D-FDAA) in acetone respectively. After adding NaHCO₃ solution (1 M, 50 μL), the reaction mixture was heated at 45 °C for 1 h and then acidified with 2 N HCl (25 μL). The standards L-Leu, L-OMe Tyr, L-Ser, erythro L-OHAsp, threo L-OHAsp and L-Tyr were derivatized in a similar manner. Derivatized hydrolysate (20 μL) and standard amino acids (20 μL) were subjected to LC-MS analysis.

Because compound 3 has two tyrosines that have different configurations, whereas compound 1 only have one tyrosine that can be assigned unambiguously by Marfey’s method. Thus, 1 was hydrolyzed and derivatized in a similar manner. Results can be found in Figure S2.

Chemical synthesis of 5-OMe Tyr.

![Chemical Synthesis Diagram]

Compound A: 4-Hydroxy-2-methoxybenzaldehyde (1 g, 6.6 mmol), hippuric acid (1.2 g, 6.7 mmol) and acetic anhydride (1.5 mL, 1.6 g, 16 mmol) were added to a 100-mL round-bottom flask, heated in an oil bath at 100 °C for 2 h. The resulting solid mixture was cooled to room temperature before H₂O (10 mL) was added. The mixture was filtrated, washed with aqueous Na₂CO₃ and dried under vacuum. The desired compound A was then obtained after recrystallization from acetone/H₂O (2:1) as a yellow solid; yield: 1.2 g (53%).

¹H NMR data see Table S13. HRESIMS m/z: 338.0992 [M+H]^+ (calcd. for [C₉H₁₂NO₂H]^+) 338.0878.

Compound C: To a 100-mL round-bottom flask was added compound A (1 g, 3.38 mmol), 1:1 mixture of CH₃Cl₂ and MeOH (12 mL) and Na₂CO₃ (283 mg, 1.41 mmol). The mixture was stirred at room temperature for 12 hours, then filtered and concentrated to afford compound B (850mg, 85%). ¹H NMR data see Table S13. HRESIMS m/z: 350.1016 [M+Na]^+ (calcd. for [C₉H₁₂NO₂Na]^+) 350.1107). Compound B was dissolved in a solution of Pd/C (70 mg) in MeOH (15 mL) and acetic acid (2.5 mL). The reaction suspension was then hydrogenated (50 bar H₂) for 24 h. The mixture was filtered and concentrated to give compound C as a white crystalline solid (723 mg, 72%). ¹H NMR data see Table S13. HRESIMS m/z: 330.0549 [M+H]^+ (calcd. for [C₈H₁₂NO₂H]^+) 330.1297.

Compound D: Compound C (500 mg, 1.5 mmol) was dissolved in DMSO (1.5 mL) and diluted with acetone (12 mL) then diluted with Tris buffer (80 mM) pH 7.8. The mixture was warmed to 37 °C, and Alcalase 1ml (3 mL, > 2.4 U / mL, Sigma) from Bacillus licheniformis was added to the reaction system. The reaction was periodically adjusted to pH 7.8 by the addition of 1 M NaOH until conversion was stopped by HPLC (2 days). The volatiles were removed...
in vacuo. The organic phase was acidified with HCl to pH = 2, and then extracted with EtOAc (~1:1) to give compound D as a white needle (323 mg, 67%). $^1$H NMR data see Table S13. HRESIMS m/z: 338.1012 [M+Na]$^+$ (calcd. for [C$_{17}$H$_{17}$NO$_5$Na]$^+$ 338.1140).

Compound E (5-OMe Tyr): Compound D (200 mg, 0.63 mmol) was hydrolyzed with 4N HCl. The mixture was cooled, washed with EtOAc, and concentrated to give a crude brown compound E (75 mg, 56%). $^1$H NMR data see Table S13.

Chemical synthesis of 5-OH Tyr. $^4$-$^7$

5-OH Tyr (30 mg) was prepared using the method as mentioned in 5-OMe Tyr. NMR data see Table S13, Figures S120 and S121. HRESIMS m/z: 198.0873 [M+H]$^+$ (calcd. for [C$_{19}$H$_{15}$NO$_5$H]$^+$ 198.0688).

Chemical complementation of 5-OMe Tyr or 5-OH Tyr to ΔbsmF, ΔbsmG and ΔbsmH mutant.

ΔbsmF, ΔbsmG and ΔbsmH mutant was individually cultured in a 50 ml flasks containing 20 mL B medium. After 24 hours cultivation, 5-OMe Tyr (1.3 mg) or 5-OH Tyr (1.3 mg) dissolved in DMSO was supplemented into fermentation broth and cultured for another 6 days. After extraction by XAD-16, the eluted organic solution was then analyzed by LC-MS. The LC-MS analysis was performed using a 18 min solvent gradient from 10 % to 90 % methanol in water supplied with 0.1 TFA at a flow rate of 0.4 mL/min.

Chemical complementation of L-erythro-β-OH-Asp into ΔbsmC mutant.

ΔbsmC mutant was individually cultured in a 50 ml flasks containing 20 mL B medium. After 24 hours cultivation, L-erythro-β-OH-Asp (1.2 mg) dissolved in DMSO was supplemented into fermentation broth and cultured for another 6 days. After extraction by XAD-16, the eluted organic solution was then analyzed by LC-MS. The LC-MS analysis was performed using a 18 min solvent gradient from 10 % to 90 % methanol in water supplied with 0.1 TFA at a flow rate of 0.4 mL/min.

Protein expression and purification.

DNA fragments containing target genes including BsmA (A$_1$), BsmB (A$_4$-T$_4$), BsmD (C$_6$-A$_6$-T$_6$), BsmF (A$_0$-T$_0$) and BsmH were individually amplified from genomic DNA of S. sp. 120454 with primers listed in Table S3. The purified PCR product of BsmA (A$_1$), BsmB (A$_4$-T$_4$) and BsmF (A$_0$-T$_0$) were ligated with linearized pET28a (linearized by NdeI and HindIII) to afford pHG06017, pHG06018 and pHG06020, DNA fragments containing BsmD (C$_6$-A$_6$-T$_6$) and BsmH were ligated with linearized pET22b (treated with NdeI and HindIII) to afford pHG06019 and pHG06021. The obtained plasmids were transformed into E. coli BL21(DE3), respectively. A single colony was picked to inoculate a 4 mL LB starter culture grown overnight at 37 °C, 200 rpm. The following day, 0.4 L LB media supplemented with kanamycin or ampicillin was inoculated with the starter culture and incubated at 37 °C, 200 rpm until the OD600 reached 0.6. The culture was cooled to 4 °C and induced with 0.125 mM IPTG. Cultures were incubated at 16 °C, 200 rpm for 18 h. Cell pellets were resuspended in lysis buffer (100 mM Tris, pH 8.0, 15 mM imidazole, 300 mM
NaCl, 10 % glycerol) and sonicated on ice. After centrifugation at 15000 rpm for 30 min, the supernatant was filtered and loaded onto a 5 mL Histrap HP column (GE lifesciences). Fractions containing the proteins were pooled and desalted by a PD10 column (GE Healthcare) with 100 mM Tris-HCl buffer (pH 7.5) and 10% glycerol and stored at -80°C.

**Adenylation activities of A domain.**

The A domain specificity assays were conducted in a 50 µL reaction volume containing 100 mM Tris-HCl, pH 7.5, 20 µM NRPS protein, 12.5 mM MgCl₂, 2.0 mM TCEP, 2 mM amino acid, 4 mM ATP. After reaction at room temperature for 1 hour, an equal volume of the Master Reaction Mix (Sigma-Aldrich Pyrophosphate Assay Kit MAK 168) were added to each of the sample, and incubated for another 30 minutes. Then, the fluorescence intensity (λex=316 / λem=456 nm) was measured by a microplate reader (TECAN infinite M200PRO).

**In vitro assay of BsmH.**

Enzymatic reaction was performed in 100 mM phosphate buffer (pH 7.2) containing 10 µM BsmH, 0.5 mM 5-OH Tyr, 1 mM SAM. After incubation at 30 °C for 2 h, 50 µL acetonitrile were added to quench the reaction. Then the mixture was centrifuged at 14,000 g for 10 min and the supernatant was analyzed by analytic HPLC using a 20 min solvent gradient from 5% to 20% acetonitrile in water supplied with 0.1 TFA at a flow rate of 0.5 mL/min.

**Biological activity assay of SHP2.**

The catalytic activity of SHP2 was monitored using the surrogate substrate DiFMUP in a prompt fluorescence assay format. The phosphatase reactions were performed at room temperature in 96-well black polystyrene plate, flat bottom, low flange, nonbinding surface (Corning, cat. no. 3575) using a final reaction volume of 100 µl and the following assay buffer conditions: 60 mM HEPES, pH 7.2, 75 mM NaCl, 75 mM KCl, 1 mM EDTA, 0.05% P-20, 5 mM DTT. 1 nM of SHP2WT (residues 1-525) was co-incubated with of 1 µM of bisphosphorylated IRS1 peptide (sequence:H2N-LN(pY)IDL DLV(dPEG8)LST(pY)ASINFQK-amide) and 30 µM of tested compounds. Under the same buffer conditions, the phosphatase assays of SHP2E76K (0.5 nM) or SHP2PTP (1 nM) was incubated with Compound 5 at various concentrations. After 30-60 min incubation at 25 °C, the surrogate substrate DiFMUP (Invitrogen, cat. no. D6567, 100 µM) was added to the reaction and incubated at 25 °C for 30 min. The reaction was then quenched by the addition of 20 µl of a 160 µM solution of bpV (Phen) (Enzo Life Sciences cat. no. ALX-270-204). The fluorescence signal was monitored using a microplate reader (TECAN, M200PRO) using excitation and emission wavelengths of 340 and 450 nm, respectively.

Compounds 1-6, 13-14 were screened at 30 µM against Src homology 2-containing protein tyrosine phosphatase 2 (SHP2), which is a major phosphatase involved in growth factor and cytokine-mediated signaling. Studies have shown that SHP2 allosteric inhibitors have shown remarkable anti-tumor benefits. In our initial single-concentration assays, only compound 5 had the inhibitory effect on the SHP2 (Figure S3, A). To further evaluate the acting mechanisms of compound 5 we used three SHP2 proteins to test sensitivities of 5: (1) wild-type (WT) SHP2 (residues 1–525); (2) SHP2E76K mutant with a partially open conformation in SHP2; (3) SHP2 PTP domain with a completely open conformation. Compound 5 was shown to inhibit SHP2 enzyme activity in a dose-dependent manner, and the IC₅₀ value of SHP2WT, SHP2E76K, SHP2PTP were 24.25, 45.56, and 89.98 µM, respectively (Figure S3, B and C), suggesting 5 could be a novel allosteric inhibitor of SHP2.
Table S1. Bacterial plasmids and strains.

| Plasmid/Strain | Relevant characteristics | Source |
|----------------|--------------------------|--------|
| **Plasmid**    |                          |        |
| pKC1139        | *E. coli*-Streptomyces shuttle plasmid used for gene disruption, temperature sensitive | 12     |
| PJTU2554       | Cosmid vector for genomic library construction | 13     |
| pSET152-kasOp* | pSET152 derived plasmid containing the promoter kasOp* | 14     |
| pUWL201PWT     | *E. coli*-Streptomyces expression shuttle vector harboring oriT (cloned into the PstI site) | 15     |
| pET28a         | Protein expression vector used in *E. coli*, encoding N-terminal His-tag, kanamycin resistance | Novagen |
| pET22b         | Protein expression vector used in *E. coli*, encoding C-terminal His-tag, ampicillin resistance | Novagen |
| pHG06001       | pKC1139 derived plasmid for disruption of bsmA-C<sub>1</sub> | This study |
| pHG06002       | pKC1139 derived plasmid for disruption of bsmC | This study |
| pHG06003       | pKC1139 derived plasmid for disruption of bsmD | This study |
| pHG06004       | pKC1139 derived plasmid for disruption of bsmF | This study |
| pHG06005       | pKC1139 derived plasmid for disruption of bsmG | This study |
| pHG06006       | pKC1139 derived plasmid for disruption of bsmH | This study |
| pHG06007       | pKC1139 derived plasmid for disruption of orf(-1) | This study |
| pHG06008       | pKC1139 derived plasmid for disruption of bsmI | This study |
| pHG06009       | pSET152-kasOp<sup>*</sup> derived plasmid for complementation of bsmC | This study |
| pHG06010       | pSET152-kasOp<sup>*</sup> derived plasmid for complementation of bsmF | This study |
| pHG06011       | pSET152-kasOp<sup>*</sup> derived plasmid for complementation of bsmG | This study |
| pHG06012       | pSET152-kasOp<sup>*</sup> derived plasmid for complementation of bsmF-T281A | This study |
| pHG06013       | pSET152-kasOp<sup>*</sup> derived plasmid for complementation of bsmF-F380A | This study |
| pHG06014       | pSET152-kasOp<sup>*</sup> derived plasmid for complementation of bsmF-C387A | This study |
| pHG06015       | Cosmid which contains *bsm* biosynthetic gene cluster | This study |
| pHG06016       | pUWL201PWT derived plasmid harboring *bsmF*, *bsmG*, *bsmH* | This study |
| pHG06017       | pET28a derived plasmid for expressing N-terminal His-tag BsmA (A<sub>1</sub>) | This study |
| pHG06018       | pET28a derived plasmid for expressing N-terminal His-tag BsmB (A<sub>2</sub>-T<sub>4</sub>) | This study |
| pHG06019       | pET22b derived plasmid for expressing C-terminal His-tag BsmD (C<sub>2</sub>-A<sub>6</sub>-T<sub>3</sub>) | This study |
| pHG06020       | pET28a derived plasmid for expressing N-terminal His-tag BsmF (A<sub>0</sub>-T<sub>0</sub>) | This study |
| pHG06021       | pET22b derived plasmid for expressing C-terminal His-tag BsmH | This study |
| **E. coli strains** |                           |        |
| DH5α           | General cloning host      | 16     |
| BL21 (DE3)     | Heterologous host for protein expression | NEB    |
| Strains | Description |
|---------|-------------|
| **ET12567 (pUZ8002)** | Methylation-deficient host used for *E. coli-Streptomyces* intergeneric conjugation | 3 |
| **S. albus J1074** | Model actinomycete used for gene heterologous expression | 17 |
| 120454 | Wild type strain for bosamycins production | This study |
| HG06001 | ΔbsmA-Ci, in-frame deletion mutant strain in WT, bosamycin D producing | This study |
| HG06002 | ΔbsmC, in-frame deletion mutant strain in WT, bosamycins non-producing | This study |
| HG06003 | ΔbsmD, in-frame deletion mutant strain in WT, bosamycins non-producing | This study |
| HG06004 | ΔbsmF, in-frame deletion mutant strain in WT, bosamycins non-producing | This study |
| HG06005 | ΔbsmG, in-frame deletion mutant strain in WT, bosamycins non-producing | This study |
| HG06006 | ΔbsmH, in-frame deletion mutant strain in WT, bosamycins non-producing | This study |
| HG06007 | Δorf-1), in-frame deletion mutant strain in WT, bosamycins producing | This study |
| HG06008 | Δbsml, in-frame deletion mutant strain in WT, bosamycin D producing | This study |
| HG06009 | complementation of ΔbsmC mutant by bsmC, bosamycins producing | This study |
| HG06010 | complementation of ΔbsmF mutant by bsmF, bosamycins producing | This study |
| HG06011 | complementation of ΔbsmG mutant by bsmG, bosamycins producing | This study |
| HG06012 | *Streptomyces albus* J1074 integrated with plasmid pJTU2554 | This study |
| HG06013 | *Streptomyces albus* J1074 integrated with plasmid pHG06015 which contains bsm biosynthetic gene cluster | This study |
| HG06014 | HG06013 containing plasmid pHG06016 | This study |
Table S2. Oligonucleotide primers used in this study.

| Oligonucleotide | Sequence<sup>a</sup> | Enzyme sites |
|-----------------|---------------------|--------------|
| bsmI-up-F       | CGGAGAACCGATTGCGCACAGCTTT | HindIII |
| bsmI-up-R       | ATCAGGGCACTACCAACCGACAGCTGAAG |  |
| bsmI-down-F     | CCAATCGATGACAGCAGCTGAAG |  |
| bsmI-down-R     | CCAATCGATGACAGCAGCTGAAG |  |
| bsmI-C<sub>1</sub>-up-F | GACGAGCCAGCGACAGCTGAAG | EcoRI |
| bsmI-C<sub>1</sub>-up-R | GACGAGCCAGCGACAGCTGAAG |  |
| bsmI-C<sub>1</sub>-down-F | GACGAGCCAGCGACAGCTGAAG |  |
| bsmI-C<sub>1</sub>-down-R | GACGAGCCAGCGACAGCTGAAG |  |
| bsmC-up-F       | CGGAGAACCGATTGCGACAGCTGAAG | HindIII |
| bsmC-up-R       | ATCAGGGCACTACCAACCGACAGCTGAAG |  |
| bsmC-down-F     | CCAATCGATGACAGCAGCTGAAG |  |
| bsmC-down-R     | CCAATCGATGACAGCAGCTGAAG |  |
| bsmC-C<sub>1</sub>-up-F | GACGAGCCAGCGACAGCTGAAG | EcoRI |
| bsmC-C<sub>1</sub>-up-R | GACGAGCCAGCGACAGCTGAAG |  |
| bsmC-C<sub>1</sub>-down-F | GACGAGCCAGCGACAGCTGAAG |  |
| bsmC-C<sub>1</sub>-down-R | GACGAGCCAGCGACAGCTGAAG |  |
| bsmD-up-F       | CGGAGAACCGATTGCGACAGCTGAAG | HindIII |
| bsmD-up-R       | ATCAGGGCACTACCAACCGACAGCTGAAG |  |
| bsmD-down-F     | CCAATCGATGACAGCAGCTGAAG |  |
| bsmD-down-R     | CCAATCGATGACAGCAGCTGAAG |  |
| bsmG-up-F       | CGGAGAACCGATTGCGACAGCTGAAG | EcoRI |
| bsmG-up-R       | ATCAGGGCACTACCAACCGACAGCTGAAG |  |
| bsmG-down-F     | CCAATCGATGACAGCAGCTGAAG |  |
| bsmG-down-R     | CCAATCGATGACAGCAGCTGAAG |  |
| bsmH-up-F       | CGGAGAACCGATTGCGACAGCTGAAG | EcoRI |
| bsmH-up-R       | ATCAGGGCACTACCAACCGACAGCTGAAG |  |
| bsmH-down-F     | CCAATCGATGACAGCAGCTGAAG |  |
| bsmH-down-R     | CCAATCGATGACAGCAGCTGAAG |  |

a. for amplification of homologous arms from genomic DNA for gene disruption (5′-3′)

| Oligonucleotide | Sequence<sup>a</sup> | Enzyme sites |
|-----------------|---------------------|--------------|
| orf(1)-up-F     | AAGCTT |  |
| orf(1)-up-R     | AAGCTT |  |
| orf(1)-down-F   | AAGCTT |  |
| orf(1)-down-R   | AAGCTT |  |
| bsmI-up-F       | AAGCTT |  |
| bsmI-up-R       | AAGCTT |  |
| bsmI-down-F     | AAGCTT |  |
| bsmI-down-R     | AAGCTT |  |
| bsmI-C<sub>1</sub>-up-F | AAGCTT |  |
| bsmI-C<sub>1</sub>-up-R | AAGCTT |  |
| bsmI-C<sub>1</sub>-down-F | AAGCTT |  |
| bsmI-C<sub>1</sub>-down-R | AAGCTT |  |
| bsmC-up-F       | AAGCTT |  |
| bsmC-up-R       | AAGCTT |  |
| bsmC-down-F     | AAGCTT |  |
| bsmC-down-R     | AAGCTT |  |
| bsmC-C<sub>1</sub>-up-F | AAGCTT |  |
| bsmC-C<sub>1</sub>-up-R | AAGCTT |  |
| bsmC-C<sub>1</sub>-down-F | AAGCTT |  |
| bsmC-C<sub>1</sub>-down-R | AAGCTT |  |
| bsmD-up-F       | AAGCTT |  |
| bsmD-up-R       | AAGCTT |  |
| bsmD-down-F     | AAGCTT |  |
| bsmD-down-R     | AAGCTT |  |
| bsmG-up-F       | AAGCTT |  |
| bsmG-up-R       | AAGCTT |  |
| bsmG-down-F     | AAGCTT |  |
| bsmG-down-R     | AAGCTT |  |
| bsmH-up-F       | AAGCTT |  |
| bsmH-up-R       | AAGCTT |  |
| bsmH-down-F     | AAGCTT |  |
| bsmH-down-R     | AAGCTT |  |

b. for screening of the correct mutants (5′-3′)
Letters highlighted in bold are sequences used for ligation independent cloning and the enzyme sites are indicated by underline.

| Gene | Sequence | Enzyme |
|------|----------|--------|
| bsmC | CACAGGAACACGAGGAAGATG | SpeI |
| bsmC | GCAAGGTGAACGAGTACCC | EcoRI |
| bsmD | TCCGAACACCTGATGA | |
| bsmD | CAACTCAGTGACGAGAACAC | |
| bsmF | TGCGTACATGGAAGAACAG | |
| bsmF | GTGTCGCGCTTGCCTTC | |
| bsmH | CTTGACCCGATAGCCTTTC | |
| bsmH | GATGCCAATTCGGGACAG | |
| bsmC - F | TGCTGCATGCATACGTACTAGT | |
| bsmC - R | CTATGACATGATTACGAATTC | |
| bsmF - F | TGCTGCATGCATACGTACTAGT | |
| bsmF - R | CTATGACATGATTACGAATTC | |
| bsmG - F | TGCTGCATGCATACGTACTAGT | |
| bsmG - R | CTATGACATGATTACGAATTC | |
| bsmF-T281A - F | TGCTGCATGCATACGTACTAGT | |
| bsmF-T281A - R | CTATGACATGATTACGAATTC | |
| bsmF-C387A - F | TGCTGCATGCATACGTACTAGT | |
| bsmF-C387A - R | CTATGACATGATTACGAATTC | |
| bsmF-T281A - F | TGCTGCATGCATACGTACTAGT | |
| bsmF-T281A - R | CTATGACATGATTACGAATTC | |
| bsmF-C387A - F | TGCTGCATGCATACGTACTAGT | |
| bsmF-C387A - R | CTATGACATGATTACGAATTC | |
| bsmF-C387A - F | TGCTGCATGCATACGTACTAGT | |
| bsmF-C387A - R | CTATGACATGATTACGAATTC | |
| bsmF-C387A - F | TGCTGCATGCATACGTACTAGT | |
| bsmF-C387A - R | CTATGACATGATTACGAATTC | |

**c. for genes complementation (5'-3')**

| Sequence | Enzyme |
|----------|--------|
| 152-bsmC-F | TGCTGCATGCATACGTACTAGT | |
| 152-bsmC-R | CTATGACATGATTACGAATTC | |
| 152-bsmF-F | TGCTGCATGCATACGTACTAGT | |
| 152-bsmF-R | CTATGACATGATTACGAATTC | |
| 152-bsmG-F | TGCTGCATGCATACGTACTAGT | |
| 152-bsmG-R | CTATGACATGATTACGAATTC | |
| 152-bsmF-T281A-F | TGCTGCATGCATACGTACTAGT | |
| 152-bsmF-T281A-R | CTATGACATGATTACGAATTC | |
| 152-bsmF-C387A-F | TGCTGCATGCATACGTACTAGT | |
| 152-bsmF-C387A-R | CTATGACATGATTACGAATTC | |
| 201-bsmFGH-F | AAGAGAGAGAAATACATATG | Ndel |
| 201-bsmFGH-R | CAGGATTCGATATCGTCCTTC | HindIII |

**d. for protein expression**

| Sequence | Enzyme |
|----------|--------|
| BsmA (A<sub>1</sub>-T<sub>28a</sub>) - F | GTGCCGCGCGCAGCATATGTGCTTGCGAGCCGCGGTT | Ndel |
| BsmB (A<sub>4</sub>-T<sub>4</sub>) - F | GTGCCGCGCGCAGCATATGTGCTTGCGAGCCGCGGTT | HindIII |
| BsmB (A<sub>4</sub>-T<sub>4</sub>) - R | GTGCCGCGCGCAGCATATGTGCTTGCGAGCCGCGGTT | HindIII |
| BsmB (A<sub>4</sub>-T<sub>4</sub>) - F | GTGCCGCGCGCAGCATATGTGCTTGCGAGCCGCGGTT | HindIII |
| BsmB (A<sub>4</sub>-T<sub>4</sub>) - R | GTGCCGCGCGCAGCATATGTGCTTGCGAGCCGCGGTT | HindIII |

<sup>a</sup> Letters highlighted in bold are sequences used for ligation independent cloning and the enzyme sites are indicated by underline.
| Module | Substrate recognition sequence | Corresponding amino acid | Predicted amino acid |
|--------|--------------------------------|--------------------------|---------------------|
| M1     | DASTIAAVCK                     | Tyr                      | Tyr                 |
| M2     | DASTIAAVCK                     | Tyr                      | Tyr                 |
| M3     | DAWMVGAVCK                     | Leu                      | Phe                 |
| M4     | DLTKLGVVK                     | Asp                      | Asn                 |
| M5     | DVWHFSLVDK                    | Ser                      | Ser                 |
| M6     | DASTIGAVCK                     | OMe-Tyr                  | Tyr                 |
| M7     | DAWMVGAVCK                     | Leu                      | Phe                 |
| M8     | DILQLGVIWK                     | Gly                      | Gly                 |
| M0     | DGSIAALVWK                     | Tyr                      | Tyr                 |

The prediction of the substrate specificity was based on NRPS Predictor2.18
Table S4. $^1$H (600 MHz) and $^{13}$C (150 MHz) NMR data of compound 1 in DMSO-$d_6$.

| No | δ$_C$ (type) | δ$_H$, multi. ($J$ in Hz) | No | δ$_C$ (type) | δ$_H$, multi. ($J$ in Hz) |
|----|--------------|--------------------------|----|--------------|--------------------------|
|    |              |                          |    |              |                          |
| d-Tyr | 169.44, C | 54.67, CH 3.92, d (7.4) | D-OCH$_2$-Tyr | 171.49, C | 54.28, CH 4.30, dd (14.6, 7.1) |
|     | 36.85, CH$_2$ 2.86, d (7.2) | 31.57, CH$_2$ 2.92, dd (13.5, 6.5) |
|     | 125.39, C | 130.69, CH 7.00, d (8.1) | 115.49, C | 130.69, CH 7.00, d (8.1) | 157.74, C |
|     | 54.67, CH 3.92, d (7.2) | 115.65, CH 6.68, d (8.0) | 55.55, CH$_3$ 3.70, s |
|     | 156.96, C | 154.28, CH 4.30, dd (14.6, 7.1) | 99.06, CH 6.31, s |
|     | 131.51, CH | 131.49, CH 6.83, d (8.1) | 130.69, CH 7.00, d (8.1) | 158.49, C |
| L-Leu | 171.47, C | 51.27, CH 4.28, m | 106.87, CH 6.18, d (8.0) |
|     | 39.41, CH$_2$ 1.39, 1.26, m | NH 8.12, d (6.5) |
|     | 23.99, CH 1.02, m | L-Leu 172.32, C | 23.63, CH$_3$ 0.75, d (6.5) | 51.55, CH 4.14, dd (15.8, 7.8) |
|     | 23.7, CH$_3$ 0.77, d (6.5) | 40.59, CH$_2$ 1.41, m |
|     | NH 8.67, d (7.4) | 24.15, CH 1.26, m |
| L-erythro-OHAsp | 169.72, C | 21.49, CH$_3$ 0.68, d (6.4) | Gly 171.86, C |
|     | 57.63, CH 4.29, m | 21.69, CH$_3$ 0.70, d (6.4) |
|     | 72.86, CH 3.94, d (4.1) | NH 7.88, d (8.2) |
|     | 174.47, C | 174.47, C |
|     | NH 8.04, d (5.7) | 41.68, CH$_2$ 3.70, d (5.4) |
| L-Ser | 170.33, C | NH 8.21, t (5.4) |
|     | 55.91, CH 4.12, m | 61.92, CH$_2$ 3.60, m |
|     | NH 7.78, d (7.2) |
Table S5. $^1$H (600 MHz) and $^{13}$C (150 MHz) NMR data of compound 2 in DMSO-d$_6$.

| No | $\delta_{\text{C}}$ (type) | $\delta_{\text{H}}$ multi. ($J$ in Hz) | No | $\delta_{\text{C}}$ (type) | $\delta_{\text{H}}$ multi. ($J$ in Hz) |
|----|----------------|----------------|----|----------------|----------------|
| L-Tyr | 172.41, C | 54.82, CH 4.25, m | L-Ser | 170.29, C | 55.76, CH 4.18, q (6.0) |
|     | 38.04, CH$_2$ 2.57, dd (13.8, 4.0), 2.35, dd (13.9, 8.6) | 62.06, CH$_2$ 3.56, dd (10.6, 5.3), 3.48, dd (10.1, 3.6) |
|     | 128.32, C | 171.62, C | 171.31, C |
|     | 130.62, CH 6.70, d (8.1) | 54.2, CH 4.33, dd (14.7, 7.4) |
| D-OCH$_3$-Tyr | 115.16, CH 6.54, d (8.4) | 31.87, CH$_2$ 2.82, dd (13.6, 4.3) |
|     | 115.16, CH 6.54, d (8.4) | 2.71, dd (13.3, 7.4) |
|     | 130.62, CH 6.70, d (8.1) | 115.53, C |
|     | 158.74, C | 157.84, C |
| D-Tyr | 171.62, C | 55.53, CH$_3$ 3.70, s |
|     | 54.92, CH 4.46, dd (14.2, 8.7) | 99.07, CH 6.30, d (1.9) |
|     | 37.67, CH$_2$ 2.86, 2.59, m | 158.52, C |
|     | 128.17, C | 131.44, CH 6.79, d (8.1) |
|     | 130.69, CH 7.06, d (8.3) | 106.88, CH 6.17, dd (8.1, 2.2) |
|     | 115.23, CH 6.62, d (8.4) | NH 7.93, br s |
|     | 156.31, C | L-Leu |
|     | 115.23, CH 6.62, d (8.4) | 51.31, CH 4.11, dd (15.4, 8.1) |
|     | 130.69, CH 7.06, d (8.3) | 40.73, CH$_2$ 1.35, m |
|     | NH 8.24, d (8.6) | 24.08, CH 1.17, m |
| L-Leu | 172.55, C | 21.69, CH$_3$ 0.68, d (6.5) |
|     | 51.27, CH 4.42, m | 21.74, CH$_3$ 0.77, d (6.5) |
|     | 41.35, CH$_2$ 1.46, 1.41, m | NH 7.88, d (8.2) |
|     | 24.45, CH 1.42, m | Gly 171.53, C |
|     | 23.62, CH$_3$ 0.76, d (5.7) | 41.13, CH$_2$ 3.71, s |
|     | 23.71, CH$_3$ 0.80, d (5.7) | NH 8.11, br s |
|     | NH 8.19, d (9.0) | |
| L-erythro-OHAsp | 169.09, C | 55.53, CH 4.70, t (5.4) |
|     | 71.46, CH 4.08, d (5.3) | 173.27, C |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
|     | 171.53, C | |
|     | 172.88, C | |
Table S6. $^1$H (600 MHz) and $^{13}$C (150 MHz) NMR data of compound 3 in DMSO-$d_6$.

| No | δC (type) | δH, multi. (J in Hz) | No | δC (type) | δH, multi. (J in Hz) |
|----|-----------|----------------------|----|-----------|----------------------|
| L-Tyr | 169.84, C | L-erythro-OHAsp | 168.65, C | 53.85, CH | 4.47, m |
|      | 36.59, CH$_2$ | 2.69, 2.58, m | 71.85, CH | 61.56, CH$_2$ | 3.56, dd (10.4, 5.0), |
|      | 127.11, C | 171.14, C | 130.11, CH | 6.69, d (8.3) | 3.49, dd (10.4, 4.0) |
|      | 114.82, CH | 6.53, d (8.4) | 115.34, NH | 7.76, d (6.6) | |
|      | 155.92, C | 55.24, CH | 114.93, NH | 8.20, d (8.6) | |
|      | 158.84, C | 54.20, CH | 115.96, C | 7.18, dd (10.4, 4.5) | |
| D-Tyr | 170.85, C | D-OCH$_3$-Tyr | 170.89, C | 54.92, CH | 4.54, m |
|      | 37.68, CH$_2$ | 2.88, dd (13.4, 4.5), | 115.23, C | 128.05, C | 55.09, CH$_3$ | 3.70, s |
|      | 2.63, m | 156.31, C | 130.23, CH | 7.07, d (8.4) | 98.62, CH | 6.30, d (2.0) |
|      | 114.79, CH | 6.63, d (8.4) | 157.35, C | 114.79, CH | 6.63, d (8.4) | 106.43, CH | 6.18, dd (8.1, 2.0) |
|      | 155.79, C | 131.04, CH | 114.79, CH | 6.63, d (8.4) | 106.43, CH | 6.18, dd (8.1, 2.0) |
|      | 130.23, CH | 7.07, d (8.4) | 120.75, NH | 7.96, br s | |
|      | 117.76, NH | 8.40, d (8.3) | L-Leu | 172.43, C | 50.97, CH | 4.12, dd (15.7, 7.9) |
| L-Leu | 172.07, C | 50.67, CH | 40.27, CH$_2$ | 1.36, m | |
|      | 41.16, CH$_2$ | 1.43, m | 23.68, CH | 1.23, m | |
|      | 24.09, CH | 1.43, m | 21.29, CH$_3$ | 0.69, d (6.4) | |
|      | 23.17, CH$_3$ | 0.80, d (6.6) | 21.22, CH$_3$ | 0.78, d (6.4) | |
|      | 23.27, CH$_3$ | 0.76, d (6.6) | 120.37, NH | 7.89, d (8.3) | |
|      | 119.15, NH | 8.25, d (8.4) | Gly | 172.84, C | 40.85, CH$_2$ | 3.71, s |
Table S7. $^1$H (600 MHz) and $^{13}$C (150 MHz) NMR data of compound 4 in DMSO-$d_6$.

| No | δ$_C$ (type) | δ$_H$, multi. ($J$ in Hz) | No | δ$_C$ (type) | δ$_H$, multi. ($J$ in Hz) |
|----|--------------|----------------------------|----|--------------|----------------------------|
| L-Tyr | 170.89, C | 54.73, CH 4.47, m | L-erythro `-OHAsp | 168.66, C |
| | 37.27, CH$_2$ 2.37, 2.58, m | 128.45, C | | |
| | 130.5, CH 6.88, d (7.8) | 155.8, C | 55.69, CH 4.19, m |
| | 115.34, CH 6.61, d (8.1) | NH 8.322, d (7.6) |
| | NH 8.27, d (8.2) | D-OCH$_3$-Tyr 171.11, C |
| | D-Tyr | 169.73, C | 115.57, C |
| | 54.68, CH 4.54, dd (13.4, 8.3) | 155.33, CH 6.59, d (7.8) |
| | 37.95, CH$_2$ 2.85, 2.60, m | 128.02, C | 99.13, CH 6.31, s |
| | 130.67, CH 7.01, d (7.9) | 115.33, CH 6.59, d (7.8) |
| | 156.28 | NH 7.93, br s |
| | 130.67, CH 7.01, d (7.9) | NH 8.40, d (7.9) |
| | NH 8.19, d (7.8) | L-Leu 172.4, C |
| | L-Leu | 172.15, C | 51.11, CH 4.41, m |
| | 41.63, CH$_2$ 1.43, m | 24.58, CH 1.42, m |
| | 23.65, CH$_3$ 0.77, d (6.4) | NH 7.89, d (8.0) |
| | 23.74, CH$_3$ 0.80, d (6.4) | Gly 172.73, C |
| | NH 8.19, d (7.8) | 41.15, CH$_2$ 3.71, s |
| | NH 8.12, br s | | |
Table S8. $^1$H (600 MHz) and $^{13}$C (150 MHz) NMR data of compound 5 in DMSO-$d_6$.

| No     | $\delta_C$ (type) | $\delta_H$, multi. ($J$ in Hz)   | No     | $\delta_C$ (type) | $\delta_H$, multi. ($J$ in Hz)   |
|--------|-------------------|---------------------------------|--------|-------------------|---------------------------------|
| L-Tyr  | 170.24, C         |                                 | L-erythro-ODAsp | 169.24, C         |                                 |
|        | 54.43, CH         | 4.48, m                         |        | 56.49, CH         | 4.63, m                         |
|        | 37.67, CH$_2$     | 2.72, 2.62, m                   |        | 72.98, CH         | 3.98, d (3.8)                   |
|        | 127.52, C         |                                 |        | 170.04, C         |                                 |
|        | 130.56, CH        | 6.71, d (8.3)                   | L-Ser  | 170.39, C         |                                 |
|        | 115.26, CH        | 6.53, d (8.4)                   |        | 55.66, CH         | 4.11, m                         |
|        | 156.37, C         |                                 |        | 62.01, CH$_2$     | 3.60, 3.52, m                   |
|        | 115.26, CH        | 6.53, d (8.4)                   | NH     | 7.67, d (6.9)     |                                 |
|        | 130.56, CH        | 6.71, d (8.3)                   |        | 54.94, CH         | 4.52, m                         |
|        | NH                | 8.25, d (9.1)                   | D-OCH$_2$-Tyr | 171.51, C         |                                 |
|        | 159.31, C         |                                 |        | 54.18, CH         | 4.34, dd (14.9, 7.8)            |
|        | 157.67, C         |                                 |        | 31.64, CH$_2$     | 2.70, m                         |
| D-Tyr  | 171.30, C         |                                 |        |                  | 2.98, dd (13.3, 4.7)            |
|        | 54.94, CH         | 4.52, m                         |        | 115.23, C         |                                 |
|        | 37.97, CH$_2$     | 2.89, 2.66, m                   |        | 156.59, C         |                                 |
|        | 128.18, C         |                                 |        | 55.54, CH$_3$     | 3.70, s                         |
|        | 130.56, CH        | 7.06, d (8.4)                   |        | 99.05, CH         | 6.30, d (1.9)                   |
|        | 115.25, CH        | 6.63, d (8.4)                   |        | 158.49, C         |                                 |
|        | 156.27, C         |                                 |        | 131.59, CH        | 6.82, d (8.2)                   |
|        | 115.25, CH        | 6.63, d (8.4)                   |        | 106.88, CH        | 6.18, dd (8.1, 1.9)             |
|        | 130.56, CH        | 7.06, d (8.4)                   | NH     | 8.15, br s        |                                 |
|        | NH                | 8.40, d (8.3)                   | L-Leu  | 173.25, C         |                                 |
| L-Leu  | 172.35, C         |                                 |        | 51.68, CH         | 4.16, dd (14.6, 9.1)            |
|        | 51.25, CH         | 4.43, m                         |        | 41.11, CH$_2$     | 1.44, m                         |
|        | 40.68, CH$_2$     | 1.44, m                         |        | 24.27, CH         | 1.314 m                         |
|        | 24.46, CH         | 1.44, m                         |        | 21.79, CH$_3$     | 0.73, d (6.4)                   |
|        | 23.75, CH$_3$     | 0.82, d (5.4)                   |        | 21.77, CH$_3$     | 0.79, d (6.4)                   |
|        | 23.52, CH$_3$     | 0.80, d (5.4)                   | NH     | 7.96, d (8.3)     |                                 |
|        | NH                | 8.22, d (9.1)                   | Gly    | 52.06, CH$_3$     | 3.58, s                         |
|        |                  |                                 |        | 170.58, C         |                                 |
|        |                  |                                 |        | 41.36, CH$_2$     | 3.81, dd (7.6, 6.6)             |
Table S9. \(^{1}H\) (600 MHz) and \(^{13}C\) (150 MHz) NMR data of compound 6 in DMSO-d\(_{6}\).

| No | \(\delta_c\) (type) | \(\delta_H\), multi. \((J\ \text{in Hz})\) | No | \(\delta_c\) (type) | \(\delta_H\), multi. \((J\ \text{in Hz})\) |
|----|------------------|------------------|----|------------------|------------------|
| L-Tyr | 171.82, C | L-erythro-OHAsp | 169.04, C |
| | 54.8, CH | 4.38, m | 55.45, CH | 4.73, dd (8.77, 5.9) |
| | 37.35, CH\(_2\) | 2.38, dd (13.7, 9.6) | 71.38, CH | 4.10, d (5.8) |
| | 2.55, dd (13.2, 9.4) | 173.08, C |
| | 128.45, C | NH | 8.32, d (6.5) |
| | 130.53, CH | 6.88, d (8.4) | 170.37, C |
| | 115.32, CH | 6.61, d (8.3) | 55.69, CH | 4.20, dd (12.4, 5.0) |
| | 156.15, C | 62.06, CH\(_2\) | 3.50, 3.55, dd (11.0, 5.3) |
| | 115.34, CH | 6.61, d (8.3) | NH | 7.77, d (7.5) |
| | 130.51, CH | 6.88, d (8.3) | D-OCH\(_3\)-Tyr | 171.4, C |
| | NH | 7.91, d (7.9) | 54.24, CH | 4.30, dd (14.7, 7.4) |
| | 169.79, C | 31.82, CH\(_2\) | 2.71, dd (13.3, 7.1), |
| | 22.86, CH\(_3\) | 1.73, s | 2.82, dd (13.6, 5.1) |
| D-Tyr | 171.65, C | | 115.57, C |
| | 54.91, CH | 4.46, d (14.3, 8.5) | 157.86, C |
| | 37.71, CH\(_2\) | 2.83, 2.61, m | 55.54, CH\(_3\) | 3.70, s |
| | 128.02, C | 99.03, CH | 6.31, d (6.3) |
| | 130.67, CH | 7.01, d (8.5) | 158.52, C |
| | 115.33 CH | 6.59, d (8.3) | 131.41, CH | 6.77, d (8.2) |
| | 156.28, C | 106.88, CH | 6.18, dd (8.1, 2.0) |
| | 115.33, CH | 6.59, d (8.5) | NH | 7.96, d (6.7) |
| | 130.67, CH | 7.01, d (8.5) | L-Leu | 173.11, C |
| | NH | 8.29, d (7.8) | 51.23, CH | 4.11, d (5.8) |
| L-Leu | 172.64, C | | 40.57, CH\(_2\) | 1.34, m |
| | 51.29, CH | 4.37, m | 24.04, CH | 1.15, m |
| | 41.17, CH\(_2\) | 1.46, m | 21.74, CH\(_3\) | 0.68, d (6.5) |
| | 24.41, CH | 1.42, m | 21.58, CH\(_3\) | 0.75, d (6.5) |
| | 23.57, CH\(_3\) | 0.77, d (5.6) | NH | 7.93, d (5.5) |
| | 23.68, CH\(_3\) | 0.79, d (5.6) | Gly | 52.26, CH\(_3\) | 3.60, s |
| | NH | 8.19, d (8.2) | 170.57, C |
| | | | 41.05, CH\(_2\) | 3.80, d (5.9) |
| | | | NH | 8.16, d (6.0) |
Table S10. $^1$H (600 MHz) and $^{13}$C (150 MHz) NMR data of compound 13 in DMSO-$d_6$.

| No   | $\delta_C$ (type) | $\delta_H$, multi. (J in Hz) | No  | $\delta_C$ (type) | $\delta_H$, multi. (J in Hz) |
|------|------------------|-----------------------------|-----|------------------|-----------------------------|
|      |                  |                             | L-Tyr | 170.64, C       |                             |
|      |                  |                             |       | 54.12, CH        | 4.70, s                     |
|      |                  |                             |       | 36.19, CH$_2$    | 2.57, m                     |
|      |                  |                             |       | 127.15, C        |                             |
|      |                  |                             |       | 130.56, CH       | 6.79, d (8.7)               |
|      |                  |                             |       | 114.81, CH       | 6.61, d (8.7)               |
|      |                  |                             |       | 156.55, C        |                             |
|      |                  |                             |       | 114.81, CH       | 6.61, d (8.7)               |
|      |                  |                             |       | 130.56, CH       | 6.79, d (8.7)               |
| D-Tyr | 167.73, C        |                             |       | 114.81, CH       | 6.61, d (8.7)               |
|      |                  |                             |       | 54.80, CH        | 4.41, m                     |
|      |                  |                             |       | 37.05, CH$_2$    | 2.69, 2.78, m               |
|      |                  |                             |       | 124.60, C        |                             |
|      |                  |                             |       | 130.16, CH       | 6.96, d (8.4)               |
|      |                  |                             |       | 114.81, CH       | 6.61, d (8.4)               |
|      |                  |                             |       | 155.84, C-OH     | 9.16, s                     |
|      |                  |                             |       | 114.81, CH       | 6.61, d (8.4)               |
|      |                  |                             |       | 130.16, CH       | 6.96, d (8.4)               |
|      |                  |                             |       | 8.47, d (8.2)    |                             |
| L-Leu | 172.34, C        |                             |       | 21.11, CH$_3$    | 0.75, d (6.3)               |
|      |                  |                             |       | 50.79, CH        | 4.38, m                     |
|      |                  |                             |       | 4.38, m          |                             |
|      |                  |                             |       | 171.08, C        |                             |
|      |                  |                             |       | 24.09, CH        | 1.44, m                     |
|      |                  |                             |       | 23.17, CH$_3$    | 0.82, d (6.2)               |
|      |                  |                             |       | 23.28, CH$_3$    | 0.79, d (6.2)               |
|      |                  |                             |       | 7.94, d (7.5)    |                             |
| L-erythro-OHAsp | 168.60, C        |                             |       | 54.10, CH        | 4.68, m                     |
|      |                  |                             |       | 71.00, CH        | 4.07, d (5.7)               |
|      |                  |                             |       | 172.66, C        |                             |
|      | NH               |                             |       | 8.33, d (7.4)    |                             |
Table S11. $^1$H (600 MHz) and $^{13}$C (150 MHz) NMR data of compound 14 in DMSO-$d_6$.

| No  | $\delta_C$ (type) | $\delta_H$, multi. ($J$ in Hz) | No  | $\delta_C$ (type) | $\delta_H$, multi. ($J$ in Hz) |
|-----|-------------------|-----------------------------|-----|-------------------|-----------------------------|
| L-Tyr | 171.26, C | 55.00, CH 4.71, m | L-Ser | 170.08, C | 55.50, CH 4.21, m |
|      |                  | 38.46, CH$_2$ 2.58, m |                  |                  | 61.90, CH$_2$ 3.60, dd (10.9, 5.0), |
|      |                  | 2.81, m |                  |                  | 3.64, dd (11.0, 5.4) |
|      |                  | 127.60, C |                  | NH | 7.78, d (7.7) |
|      |                  | 130.74, CH 7.03, d (8.4) | D-OCH$_3$-Tyr | 171.04, C | |
|      |                  | 115.44, CH 6.62, d (8.4) |                  | 54.09, CH | 4.35, dd (14.7, 7.4) |
|      |                  | 156.50, C |                  | 32.05, CH$_2$ | 2.79, m |
|      |                  | 115.44, CH 6.62, d (8.4) |                  | 2.71, m |
|      |                  | 130.74, CH 7.03, d (8.4) |                  | 115.44, C | |
|      |                  | NH | 8.78, d (8.7) |                  | 158.53, C |
| D-Tyr | 168.16, C | 55.55, CH$_3$ 3.70, s |                  |                  | |
|      |                  | 54.03, CH | 4.40, m | 99.08, CH | 6.31, d (1.3) |
|      |                  | 36.64, CH$_2$ | 2.77, 2.80, m | 157.87, C-OH | 9.20, s |
|      |                  | 125.08, C |                  | 131.41, CH | 6.79, dd (8.1, 3.9) |
|      |                  | 131.03, CH 6.79, d (8.1) |                  | 106.88, CH | 6.18, d (8.0) |
|      |                  | 115.44, CH 6.62, d (8.3) |                  | NH | 7.88, br s |
|      | 157.87, C-OH | 9.17, s |                  | L-Leu | 172.77, C |
|      |                  | 115.44, CH 6.62, d (8.3) |                  | 51.26, CH | 4.13, m |
|      |                  | 131.03, CH 6.79, d (8.3) |                  | 40.81, CH$_3$ | 1.34, m |
|      |                  | NH | 8.47, d (8.2) |                  | 24.05, CH | 1.17, m |
| L-Leu | 172.64, C |                  |                  |                  | 21.77, CH$_3$ | 0.69, d (6.5) |
|      | 51.18, CH | 4.40, m |                  |                  | 21.54, CH$_3$ | 0.79, d (6.5) |
|      | 41.40, CH$_2$ | 1.44, m |                  |                  | NH | 7.90, d (7.5) |
|      | 24.55, CH | 1.44, m |                  | Gly | 173.07, C |
|      | 23.62, CH$_3$ | 0.76, d (5.7) |                  |                  | 41.08, CH$_2$ | 3.71, s |
|      | 23.72, CH$_3$ | 0.83, d (5.7) |                  |                  | NH | 8.10, d (5.8) |
|      | NH | 7.90, d (9.0) |                  |                  | |
| L-erythro-OHAsp | 169.06, C | 55.00, CH 4.71, m |                  |                  | |
|      | 71.44, CH | 4.07, d (5.8) |                  |                  | |
|      | 172.89, C | NH | 8.37, d (8.7) |                  | |
Table S12. $^1$H NMR (400 MHz) data of A-E in DMSO-$d_6$.

| No | A         | B         | C         | D         | E         |
|----|-----------|-----------|-----------|-----------|-----------|
|    | $\delta_H$ (multi, $J$, Hz) | $\delta_H$ (multi, $J$, Hz) | $\delta_H$ (multi, $J$, Hz) | $\delta_H$ (multi, $J$, Hz) | $\delta_H$ (multi, $J$, Hz) |
| 1' | 3.67 (s)  | 3.60 (s)  |           |           |           |
| 2  |           | 4.58 (dd, 13.6, 9.0) | 4.55 (m)  | 3.96 (t, 5.5) |           |
| 2'-NH |     | 8.66 (d, 7.6) | 8.49 (d, 8.0) | 8.12 (s) |           |
| 3  | 7.51 (s)  | 7.70 (s)  | 2.86 (dd, 13.5, 9.7) | 2.82 (dd, 13.5, 10.5) | 2.86 (dd, 14.0, 7.1) |
|    |           |           | 3.10 (dd, 13.5, 5.5) | 3.16 (dd, 13.5, 4.4) | 3.30 (dd, 14.0, 6.4) |
| 4  | 8.81 (d, 8.6) | 7.48 (d, 7.7) | 6.97 (d, 8.1) | 7.00 (d, 8.0) | 6.91 (d, 8.1) |
| 5  | 6.94 (dd, 8.6, 1.8) | 6.20 (d, 8.6) | 6.97 (d, 8.1) | 6.23 (dd, 8.1, 1.5) | 6.29 (dd, 8.1, 2.2) |
| 6-OH  | 9.75 (brs) | 9.35 (brs) | 9.24 (brs) | 9.47 (brs) |           |
| 6' | 2.31 (s)  |           |           |           |           |
| 7  | 7.00 (d, 1.8) | 6.32 (s)  | 6.37 (d, 1.5) | 6.36 (d, 1.5) | 6.40 (d, 2.1) |
| 8' | 3.91 (s)  | 3.75 (s)  | 3.74 (s)  | 3.73 (s)  | 3.71 (s)  |
| Ph | 8.12 (d, 7.4) | 7.95 (d, 7.5) | 7.79 (d, 7.4) | 7.77 (d, 7.5) |           |
|    | 8.12 (d, 7.4) | 7.95 (d, 7.5) | 7.79 (d, 7.4) | 7.77 (d, 7.5) |           |
|    | 7.73 (m)  | 7.56 (t, 7.1) | 7.53 (t, 7.2) | 7.53 (t, 7.2) |           |
|    | 7.64 (t, 7.6) | 7.48 (t, 7.6) | 7.46 (t, 7.1) | 7.45 (t, 7.5) |           |
|    | 7.64 (t, 7.6) | 7.48 (t, 7.6) | 7.46 (t, 7.1) | 7.45 (t, 7.5) |           |
Table S13. $^1$H NMR (400 MHz) data of G, I and J in DMSO-$d_6$.

| No | G | I | J |
|----|---|---|---|
| 1' | $\delta_H$ (multi, $J$, Hz) | $\delta_H$ (multi, $J$, Hz) | $\delta_H$ (multi, $J$, Hz) |
| 2 | 3.69 (s) | 4.55 (m) | 3.95 (t, 5.5) |
| 2'-NH | 8.49 (brd, 7.7) | 8.28 (brd, 7.7) |  |
| 3 | 7.70 (s) | 2.80 (dd, 13.6, 10.3) | 2.85 (dd, 13.8, 7.4) |
|  | | 3.08 (dd, 13.6, 4.3) | 2.98 (dd, 13.8, 6.6) |
| 4 | 7.52 (d, 8.6) | 6.92 (d, 8.2) | 6.83 (d, 8.2) |
|  | | 6.19 (dd, 8.6, 2.1) |  |
| 5 | 9.75 (s) | 9.03 (brs) |  |
|  | | 10.10 (brs) | 9.47 (brs) |
| 6-OH | 6.37 (d, 2.1) | 6.27 (d, 2.3) | 6.39 (d, 2.3) |
| 8-OH | 7.96 (d, 7.4) | 7.77 (d, 7.4) | 7.77 (d, 7.5) |
| Ph | 7.96 (d, 7.4) | 7.77 (d, 7.4) | 7.52 (t, 7.5) |
| | 7.58 (t, 7.4) | 7.48 (t, 7.4) | 7.45 (t, 7.5) |
| | 7.48 (t, 7.4) | 7.45 (t, 7.5) |  |
Table S14. Conserved sequence regions in the alignment comparisons of BsmF and other known P450s (numbering indicated for BsmF).19,20

| P450     | B-B2 loop N-term | B-B2 loop C-term | F-helix | G-helix | I-helix | β-1 sheet |
|----------|------------------|------------------|---------|---------|---------|-----------|
|          | 87  90           | 107 108 109 111 114 115 | 194 195 196 197 | 216 217 218 219 222 | 270 271 277 | 324 325 326 |
| OxyD     | G I              | S G G M V S      | H A F G | A H T E V | N C G   | A M H     |
| CloI     | G L              | A S G M V T      | H A W S | A K N E L | N C G   | S L H     |
| NovI     | G L              | A S G M V T      | H A W S | A K N E L | N C G   | S L H     |
| SimD1    | G L              | A S R M L T      | H A L S | A K N E L | N C G   | S L H     |
| Sky32    | G L              | A A G M V T      | S A L S | A R N E L | N C G   | A M H     |
| Consensus*| G L              | A (1) G M V T   | H A (2) S | A (3) N E (4) | N C G   | (5) LM H  |
| BsmF-P450| A G              | M G S Q F N      | S Y E R | L L D K A | N A G   | S Q Y     |

* Identity residues shown in bold and underlined, mismatching residues or similar residues indicated in normal font. Exceptions are: (1) Small residue (S, G, A), (2) large hydrophobic (W, F, L), (3) positively charged residue (K, H, R), (4) majority hydrophobic (L, V; also S and G), (5) majority small (S, A; also V). Protein accession number: OxyD (3MGX_A); CloI (AAN65225); NovI (Q9L9F9); SimD1 (AAK06805); Sky32 (4L0F_A).
Figure S1. Structures of bosamycins.
Figure S2. LC-MS analysis of L-FDAA and D-FDAA derivatives of the amino acid residues in 1 and 3. Panel A indicates 3\textsuperscript{rd} and 7\textsuperscript{th} Leu in 3 is L-type; Panel B indicates 6\textsuperscript{th} OMe-Tyr in 3 is D-type, Panel C indicates 5\textsuperscript{th} Ser in 3 is L-type; Panel D indicates 4\textsuperscript{th} OH-Asp in 3 is erythro-L-OH-Asp; Panel E indicates 2\textsuperscript{nd} Tyr in 1 and 3 is D-type, 1\textsuperscript{st} Tyr in 3 is L-type. The deduced d-type configurations in 2\textsuperscript{nd}, and 6\textsuperscript{th} amino acid residues are consistent with the presence of E domains in their corresponding modules (Scheme 1).
Figure S3. Identification of compound 5 as a novel inhibitor of SHP2. A) Primary screen the biological activity of compounds on SHP2 enzyme activity were examined. SHP2 was screened in the presence of 1 μM 2P-IRS-1 and 30 μM of each compound. B) Phosphatase activities of SHP2<sup>WT</sup>, SHP2<sup>E76K</sup>, SHP2<sup>PTP</sup> were assessed in the presence of compound 5 at various concentrations. C) The IC<sub>50</sub> value of 5 against SHP2<sup>WT</sup>, SHP2<sup>E76K</sup>, SHP2<sup>PTP</sup>. 

| Enzyme     | IC<sub>50</sub> (μM) |
|------------|----------------------|
| SHP2<sup>WT</sup> | 24.25                |
| SHP2<sup>E76K</sup> | 46.56                |
| SHP2<sup>PTP</sup> | 89.98                |
Figure S4. Generation of the S. sp. 120454 mutant strains. A) Gene disruption with homologous recombination strategies. B) The S. sp. 120454 HG06001 mutant (∆bsmA-C). Lane 1, mutant strain; Lane 2, S. sp. 120454 WT; C) The S. sp. 120454 HG06002 mutant (∆bsmC). Lane 1, mutant strain; Lane 2, S. sp. 120454 WT; D) The S. sp. 120454 HG06003 mutant (∆bsmD). Lane 1, mutant strain; Lane 2, S. sp. 120454 WT; E) The S. sp. 120454 HG06004 mutant (∆bsmF). Lane 1, mutant strain; Lane 2, S. sp. 120454 WT; F) The S. sp. 120454 HG06005 mutant (∆bsmG). Lane 1, mutant strain; Lane 2, S. sp. 120454 WT; G) The S. sp. 120454 HG06006 mutant (∆bsmH). Lane 1, mutant strain; Lane 2, S. sp. 120454 WT; H) The S. sp. 120454 HG06007 mutant (∆orf-1). Lane 1, mutant strain; Lane 2, S. sp. 120454 WT; I) The S. sp. 120454 HG06008 mutant (∆bsml). Lane 1, mutant strain; Lane 2, S. sp. 120454 WT; Lane M, Trans2K® Plus II DNA marker.
Figure S5. MS/MS analysis of metabolite extracts from wild type and mutant strains. A) S. sp. 120454 WT. B) The S. sp. 120454 HG06004 mutant (ΔbsmF). C) The S. sp. 120454 HG06005 mutant (ΔbsmG). D) The S. sp. 120454 HG06005 mutant (ΔbsmH). E) The S. sp. 120454 HG06002 mutant (ΔbsmC).
**Figure S6.** Chemical complementation of 5-OMe Tyr into mutants. LC-MS analyses of A) compound 2, B) compound 3 and C) compound 4 for different mutant strains. i) wt; ii) ΔbsmF mutant fed with 5-OMe Tyr; iii) ΔbsmG mutant fed with 5-OMe Tyr; iv) ΔbsmH mutant fed with 5-OMe Tyr; v) ΔbsmH mutant; vi) ΔbsmG mutant; vii) ΔbsmF mutant.

**Figure S7.** Chemical complementation of 5-OH Tyr into mutants. LC-MS analyses of A) compound 2, B) compound 3 and C) compound 4 for different mutant strains. i) wt; ii) ΔbsmF mutant fed with 5-OH Tyr; iii) ΔbsmG mutant fed with 5-OH Tyr; iv) ΔbsmH mutant fed with 5-OH Tyr; v) ΔbsmH mutant; vi) ΔbsmG mutant; vii) ΔbsmF mutant.
Figure S8. HPLC analysis of metabolite extracts from mutant strains and gene complementation strains. i) wild type; ii) ΔbsmF mutant strain; iii) ΔbsmG mutant strain; iv) ΔbsmC mutant strain; v) complementation of ΔbsmF mutant by bsmF; vi) complementation of ΔbsmF-T281A mutant by bsmF; vii) complementation of ΔbsmF-F281A mutant by bsmF. viii) complementation of ΔbsmF−C281A mutant by bsmF; ix) complementation of ΔbsmC mutant by bsmC; x) complementation of ΔbsmG mutant by bsmG. 8 and 9 have identical retention time.

Figure S9. Chemical complementation fed L-erythro-β-OH-Asp into ΔbsmC mutant. LC-MS analyses of A) compound 2, B) compound 3 and C) compound 4 for different strains. i) wt; ii) ΔbsmC mutant fed L-erythro-β-OH-Asp; iii) ΔbsmC mutant.
Figure S10. Verification of cosmid pHG06015. A) Physical map of pJTU2554 harboring the bsm gene cluster. B) agarose gel electrophoresis of pHG06015 cosmid. Legend: M1, 15K marker; 1, pHG06015 digested by BamHI; 2, pHG06011 digested by KpnI; M2, 8K marker.

Figure S11. SDS-PAGE analysis of proteins. A) BsmA (A\textsubscript{1}) (calculated molecule weight: 44.6 KDa); B) BsmB (A\textsubscript{4}-T\textsubscript{4}) (calculated molecule weight: 71.6 KDa); C) BsmD (C\textsubscript{6}-A\textsubscript{6}-T\textsubscript{6}) (calculated molecule weight: 112.9 KDa); D) BsmF (A\textsubscript{0}-T\textsubscript{0}) (calculated molecule weight: 70.8 KDa). E) BsmH (calculated molecule weight: 37.0 KDa).
Figure S12. Sequence alignment of BsmF-P450 with other P450 proteins. The Glu/Thr residues are important in interactions with and proton transfer to iron-oxo intermediates in the P450 catalytic cycle. The heme-binding motif conserved Cys residues that act as the proximal ligand to the heme iron, the conserved phenylalanines as a regulator of heme iron potential.21,22 Protein accession number: OxyD (3MGX_A); Sky32 (4LOF_A); HmtT (CB24154); HmtN (5XW2_A); EryK (P48635.3); Clol (AAN65225); NovI (Q9L9F9); SimD1 (AAK06805); Biol (AGG62423.1). The multiple alignment was generated by ClustalW servers and rendered with ESPript 3.0.23
**Figure S13.** Domain analyze of BsmF. P450 domain (27-436 aa); A domain (537-930 aa); T domain (1062-1131 aa); linker regions (437-536 aa and 931-1061 aa).
Figure S14. Sequence alignment of BsmC with other β-hydroxylases. Fe(II) and α-ketoglutarate binding residues are marked with red arrows and blue arrows. Protein accession number: SyrP_Syringomycin (AKF46133.1); ThaF_Thanamycin (ALG65284.1); NupP_Nunamycin (KPN90375.1); PSEEN3233_Pyoverdine-L48 (WP_011534378.1); SyrP_Pyoverdine-21245 (AJW67533.1); HcsE_Histicorrugatin (WP_053122094.1). The multiple alignment was generated by ClustalW servers and rendered with ESPript 3.0.
Figure S15. Sequence alignment of BsmB-T2 with other T domains. Strictly conserved GGHSL motif in the thiolation domain are marked with black box. Protein accession number: NupE-T3 (KPN90369.1); SyrE-T4 (AY37647.1); ThaB-T3 (AED90003.1). The multiple alignment was generated by ClustalW servers and rendered with ESPript 3.0.

Figure S16. Sequence alignment of BsmA-C1 with other C domains. The conserved histidine and aspartate residues are marked by red arrows. Protein accession number: CDAPS1-C1 (CAB38517); SrfAA-C1 (CAE02630); DptA-C1 (AHX36919); Cdel-C1 (QBC75021); MlcK-C1 (ARU08073); MycA-C1 (Q9R9J1); ItuA-C1 (BAA69698). The multiple alignment was generated by ClustalW servers and rendered with ESPript 3.0.
Figure S17. MS/MS analysis of 1.

Figure S18. MS/MS analysis of 2.

Figure S19. MS/MS analysis of 3.

Figure S20. MS/MS analysis of 4.
Figure S21. MS/MS analysis of 5.

Figure S22. MS/MS analysis of 6.

Figure S23. MS/MS analysis of 7 in ΔbsmF mutant strain.

Figure S24. MS/MS analysis of 8 in ΔbsmF mutant strain.
Figure S25. MS/MS analysis of 9 in ΔbsmF mutant strain.

Figure S26. MS/MS analysis of 7 in ΔbsmG mutant strain.

Figure S27. MS/MS analysis of 8 in ΔbsmG mutant strain.

Figure S28. MS/MS analysis of 9 in ΔbsmG mutant strain.
Figure S29. MS/MS analysis of 7 in ΔbsmH mutant strain.

Figure S30. MS/MS analysis of 8 in ΔbsmH mutant strain.

Figure S31. MS/MS analysis of 9 in ΔbsmG mutant strain.

Figure S32. MS/MS analysis of 10 in ΔbsmC mutant strain.
Figure S33. MS/MS analysis of 11 in ΔbsmC mutant strain.

Figure S34. MS/MS analysis of 12 in ΔbsmC mutant strain.

Figure S35. MS/MS analysis of 13.

Figure S36. MS/MS analysis of 14.
Figure S37. HRESIMS spectrum of 1.

Figure S38. $^1$H NMR (600 MHz, DMSO-$d_6$) spectrum of compound 1.
Figure S39. $^{13}$C NMR (150 MHz, DMSO-$d_6$) spectrum of compound 1.

Figure S40. DEPT NMR (150 MHz, DMSO-$d_6$) spectrum of compound 1.
Figure S41. $^1$H−$^1$H COSY NMR (600 MHz, DMSO-d$_6$) spectrum of compound 1.

Figure S42. HSQC NMR (600 MHz, DMSO-d$_6$) spectrum of compound 1.
Figure S43. HMBC NMR (600 MHz, DMSO-$d_6$) spectrum of compound 1.

Figure S44. NOESY NMR (600 MHz, DMSO-$d_6$) spectrum of compound 1.
Figure S45. HRESIMS spectrum of 2.

Figure S46. $^1$H NMR (600 MHz, DMSO-$d_6$) spectrum of compound 2.
Figure S47. $^{13}$C NMR (150 MHz, DMSO-$d_6$) spectrum of compound 2.

Figure S48. DEPT NMR (150 MHz, DMSO-$d_6$) spectrum of compound 2.
Figure S49. $^1$H–$^1$H COSY NMR (600 MHz, DMSO-$d_6$) spectrum of compound 2.

Figure S50. HSQC NMR (600 MHz, DMSO-$d_6$) spectrum of compound 2.
Figure S51. HMBC NMR (600 MHz, DMSO-d$_6$) spectrum of compound 2.

Figure S52. NOESY NMR (600 MHz, DMSO-d$_6$) spectrum of compound 2.
Figure S53. HRESIMS spectrum of 3.

Figure S54. $^1$H NMR (600 MHz, DMSO-d$_6$) spectrum of compound 3.
Figure S55. $^{13}$C NMR (150 MHz, DMSO-d$_6$) spectrum of compound 3.

Figure S56. DEPT NMR (150 MHz, DMSO-d$_6$) spectrum of compound 3.
Figure S57. $^1$H−$^1$H COSY NMR (600 MHz, DMSO-$d_6$) spectrum of compound 3.

Figure S58. HSQC NMR (600 MHz, DMSO-$d_6$) spectrum of compound 3.
Figure S59. HMBC NMR (600 MHz, DMSO-\textit{d}_6) spectrum of compound 3.

Figure S60. NOESY NMR (600 MHz, DMSO-\textit{d}_6) spectrum of compound 3.
Figure S61. N-H HSQC NMR (600 MHz, DMSO-$d_6$) spectrum of compound 3.

Figure S62. N-H HMBC NMR (600 MHz, DMSO-$d_6$) spectrum of compound 3.
Figure S63. HRESIMS spectrum of 4.

Figure S64. $^1$H NMR (600 MHz, DMSO-$d_6$) spectrum of compound 4.
Figure S65. $^{13}$C NMR (150 MHz, DMSO-$d_6$) spectrum of compound 4.

Figure S66. DEPT NMR (150 MHz, DMSO-$d_6$) spectrum of compound 4.
Figure S67. $^1$H-$^1$H COSY NMR (600 MHz, DMSO-$d_6$) spectrum of compound 4.

Figure S68. HSQC NMR (600 MHz, DMSO-$d_6$) spectrum of compound 4.
Figure S69. HMBC NMR (600 MHz, DMSO-\textit{d}_6) spectrum of compound 4.

Figure S70. NOESY NMR (600 MHz, DMSO-\textit{d}_6) spectrum of compound 4.
Figure S71. HRESIMS spectrum of 5

Figure S72. $^1$H NMR (600 MHz, DMSO-$d_6$) spectrum of compound 5.
Figure S73. $^{13}$C NMR (150 MHz, DMSO-$d_6$) spectrum of compound 5.

Figure S74. DEPT NMR (150 MHz, DMSO-$d_6$) spectrum of compound 5.
Figure S75. $^1$H−$^1$H COSY NMR (600 MHz, DMSO-$d_6$) spectrum of compound 5.

Figure S76. HSQC NMR (600 MHz, DMSO-$d_6$) spectrum of compound 5.
Figure S77. HMBC NMR (600 MHz, DMSO-d$_6$) spectrum of compound 5.

Figure S78. NOESY NMR (600 MHz, DMSO-d$_6$) spectrum of compound 5.
Figure S79. HRESIMS spectrum of 6.

Figure S80. $^1$H NMR (600 MHz, DMSO-$d_6$) spectrum of compound 6.
Figure S81. $^{13}$C NMR (150 MHz, DMSO-$d_6$) spectrum of compound 6.

Figure S82. DEPT NMR (150 MHz, DMSO-$d_6$) spectrum of compound 6.
Figure S83. $^1$H–$^1$H COSY NMR (600 MHz, DMSO-$d_6$) spectrum of compound 6.

Figure S84. HSQC NMR (600 MHz, DMSO-$d_6$) spectrum of compound 6.
Figure S85. HMBC NMR (600 MHz, DMSO-\textit{d}_6) spectrum of compound 6.

Figure S86. NOESY NMR (600 MHz, DMSO-\textit{d}_6) spectrum of compound 6.
Figure S87. HRESIMS spectrum of 13.

Figure S88. $^1$H NMR (600 MHz, DMSO-$d_6$) spectrum of compound 13.
Figure S89. $^{13}$C NMR (150 MHz, DMSO-$d_6$) spectrum of compound 13.

Figure S90. DEPT NMR (150 MHz, DMSO-$d_6$) spectrum of compound 13.
Figure S91. $^1$H–$^1$H COSY NMR (600 MHz, DMSO-$d_6$) spectrum of compound 13.

Figure S92. HSQC NMR (600 MHz, DMSO-$d_6$) spectrum of compound 13.
Figure S93. HMBC NMR (600 MHz, DMSO-\textit{d}_6) spectrum of compound 13.

Figure S94. NOESY NMR (600 MHz, DMSO-\textit{d}_6) spectrum of compound 13.
Figure S95. HRESIMS spectrum of 14.

Figure S96. $^1$H NMR (600 MHz, DMSO-$d_6$) spectrum of compound 14.
Figure S97. $\text{^{13}C}$ NMR (150 MHz, DMSO-$d_6$) spectrum of compound 14.

Figure S98. DEPT NMR (150 MHz, DMSO-$d_6$) spectrum of compound 14.
Figure S99. $^1$H−$^1$H COSY NMR (600 MHz, DMSO-$d_6$) spectrum of compound 14.

Figure S100. HSQC NMR (600 MHz, DMSO-$d_6$) spectrum of compound 14.
Figure S101. HMBC NMR (600 MHz, DMSO-\textit{d}_6) spectrum of compound 14.

Figure S102. NOESY NMR (600 MHz, DMSO-\textit{d}_6) spectrum of compound 14.
Figure S103. HRESIMS spectrum of A.

Figure S104. $^1$H NMR (400 MHz, DMSO-$d_6$) spectrum of compound A.
Figure S105. HRESIMS spectrum of B.

Figure S106. $^1$H NMR (400 MHz, DMSO-$_d_6$) spectrum of compound B.
**Figure S107.** HRESIMS spectrum of C.

**Figure S108.** $^1$H NMR (400 MHz, DMSO-$d_6$) spectrum of compound C.
Figure S109. HRESIMS spectrum of D.

Figure S110. $^1$H NMR (400 MHz, DMSO-$d_6$) spectrum of compound D.
Figure S111. HRESIMS spectrum of E.

Figure S112. $^1$H NMR (400 MHz, DMSO-$d_6$) spectrum of compound E.
Figure S113. HRESIMS spectrum of G.

Figure S114. $^1$H NMR (400 MHz, DMSO-$d_6$) spectrum of compound G.
Figure S115. $^{13}$C NMR (100 MHz, DMSO-$d_6$) spectrum of compound G.

Figure S116. HRESIMS spectrum of I.
Figure S117. $^1$H NMR (400 MHz, DMSO-$d_6$) spectrum of compound I.

Figure S118. $^{13}$C NMR (100 MHz, DMSO-$d_6$) spectrum of compound I.
Figure S119. HRESIMS spectrum of \( J \).

Figure S120. \(^1\)H NMR (400 MHz, DMSO-\( d_6 \)) spectrum of compound \( J \).
Figure S121. $^{13}\text{C}$ NMR (100 MHz, DMSO-$d_6$) spectrum of compound J.
References

1. Y. N. Song, R. H. Jiao, W. J. Zhang, G. Y. Zhao, H. Dou, R. Jiang, A. H. Zhang, Y. Y. Hou, S. F. Bi, H. M. Ge and R. X. Tan, Org. Lett., 2015, 17, 556-559.
2. J. Sambrook, E. F. Fritsch and T. Maniatis, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1989.
3. R. Bhushan and H. Bruckner, Amino Acids 2004, 27, 231-247.
4. K. Kim, K. Parang, O. D. Lau and P. A. Cole, Bioorg. Med. Chem., 2000, 8, 1263-1268.
5. F. M. Bautista, J. M. Campelo, A. Garcia, D. Luna, J. M. Marines and A. A. Romero, J. Chem. Soc. Perkin Trans., 2002, 2, 227-234.
6. U. Nagel, Angew. Chem., Int. Ed., 1984, 96, 425-426.
7. M. Tagaki and K. Yamamoto, Tetrahedron 1991, 47, 8869-8882.
8. Y. N. Chen, M. J. LaMarche, H. M. Chan, P. Fekkes, J. Garcia-Fortanet, M. G. Acker, B. Antonakos, C. H. Chen, Z. Chen, V. G. Cooke, J. R. Dobson, Z. Deng, F. Fei, B. Firestone, M. Fodor, C. Fridrich, H. Gao, D. Grunenfelder, H. X. Hao, J. Jacob, S. Ho, K. Hsiao, Z. B. Kang, R. Karki, M. Kato, J. Larrow, L. R. La Bonte, F. Lenoir, G. Liu, S. Liu, D. Majumdar, M. J. Meyer, M. Palermo, L. Perez, M. Pu, E. Price, C. Quinn, S. Shaha, M. D. Shultz, J. Slisz, K. Venkatesan, P. Wang, M. Warmuth, S. Williams, G. Yang, J. Yuan, J. H. Zhang, P. Zhu, T. Ramsey, N. J. Keen, W. R. Sellers, T. Stams and P. D. Fortin, Nature 2016, 535, 148-152.
9. Q. Liu, C. Qu, M. Zhao, Q. Xu and Y. Sun, Pharmacol. Res., 2020, 152, 104595-104609.
10. H. Ran, R. Tsutsumi, T. Araki and B. G. Neel, Cancer Cell 2016, 30, 194-196.
11. A. Flemming, Nat. Rev. Drug Discov., 2016, 15, 530-531.
12. M. Bierman, R. Logan, O’Brien, K. Seno, E. T. Rao and R. N. Schoner, B. E. Gene 1992, 116, 43-49.
13. L. Li, Z. Xu, J. Wu, Y. Zhang, X. He, T. M. Zabriskie and Z. Deng, ChemBioChem., 2008, 9, 1286-1294.
14. L. Li, G. Pan, X. Zhu, K. Fan, W. Gao, G. Ai, J. Ren, M. Shi, C. Olano, J. A. Salas and K. Yang, Appl. Microbiol. Biot., 2017, 101, 5291-5300.
15. J. D. Rudolf, L. B. Dong, T. Huang and B. Shen, Mol. BioSyst., 2015, 11, 2717-2726.
16. M. Green, J. Sambrook, Molecular Cloning: A Laboratory Manual 4th edn (Cold Spring Harbor Laboratory Press, 2012).
17. F. Xu, Nazari, B. Moon, K. L. B. Bushin and M. R. Seyedsayamdost, J. Am. Chem. Soc., 2017, 139, 9203-9212.
18. M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch and O. Kohlbacher, Nucleic Acids Res., 2011, 39, 362-367.
19. M. J. Cryle, A. Meinhart and I. Schlichting, J. Biol. Chem., 2010, 285, 24562-24574.
20. K. Haslinger, C. Brieke, n S. Uhlman, L. Sieverling, R. D. Süßmuth, M. J. Cryle. Angew. Chem., Int. Ed., 2014, 53, 8518-8522.
21. H. Yeom, S. G. Sligar, H. Li, T. L. Poulos and A. J. Fulco, Biochemistry 1995, 34, 14733-14740.
22. T. W. Ost, C. S. Miles, A. W. Munro, J. Murdoch, C. Reig and K. A. Chapman, Biochemistry 2001, 40, 13421-13429.
23. X. Robert and P. Gouet., Nucleic Acids Res., 2014, 42, W320-W324.
24. Z. L. Reitz, C. D. Hardy, J. Suk, J. Bouvet and A. Butler, Proc. Natl. Acad. Sci. U. S. A., 2019, 116, 19805-19814.
25. R. D. Süßmuth and A. Mainz, Angew. Chem., Int. Ed., 2017, 56, 3770-3821.